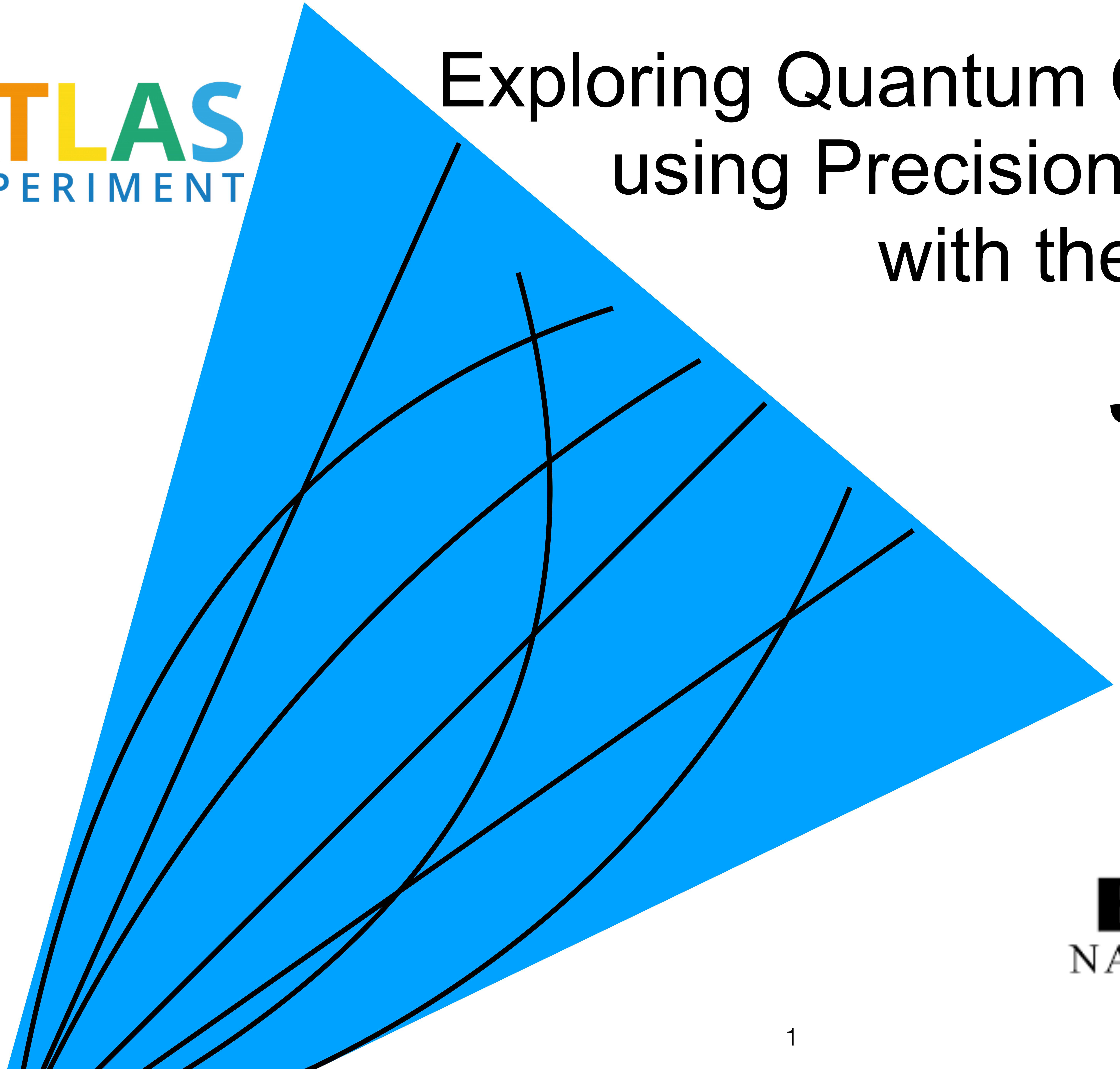


Exploring Quantum Chromodynamics using Precision Jet Substructure with the ATLAS Detector

Jennifer Roloff

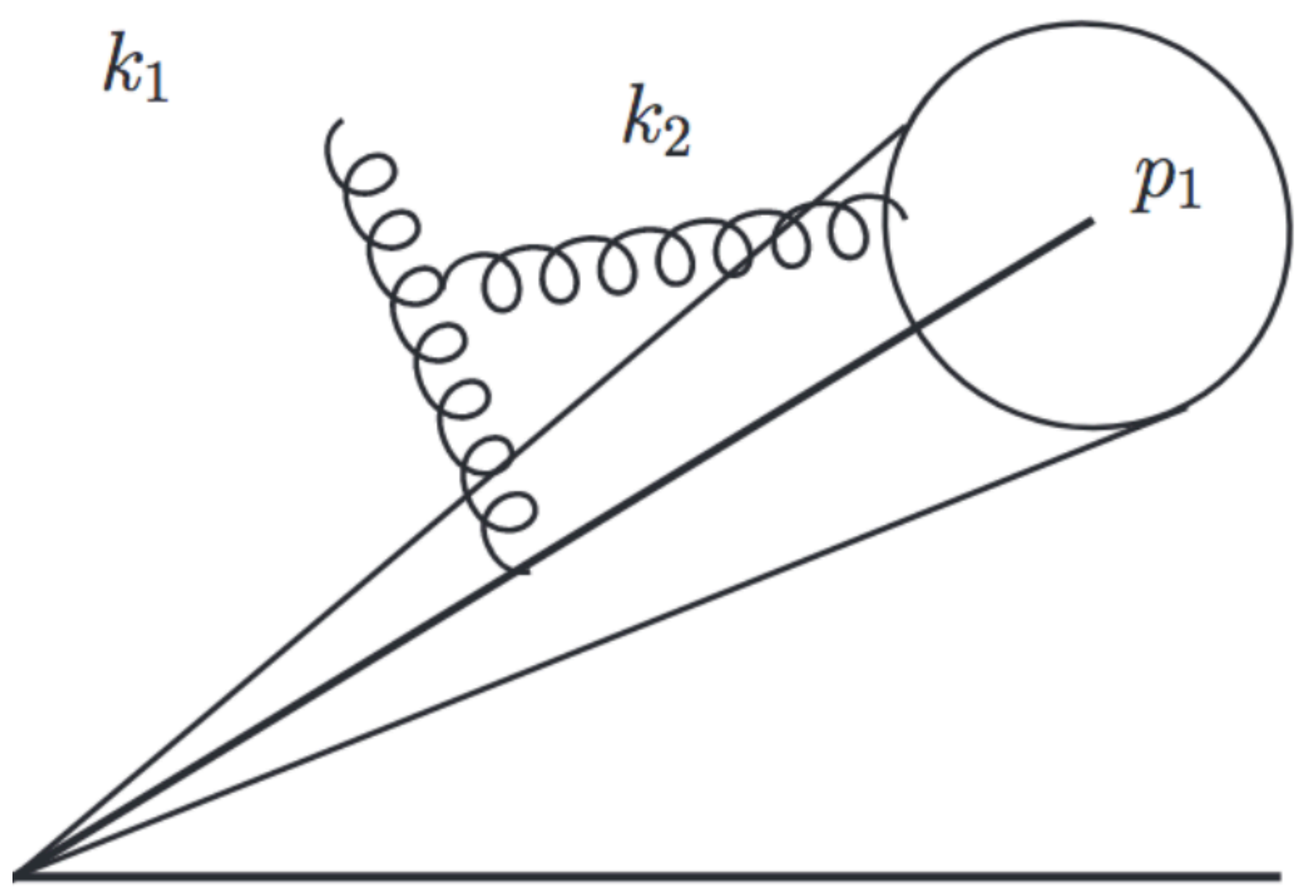
September 21, 2020

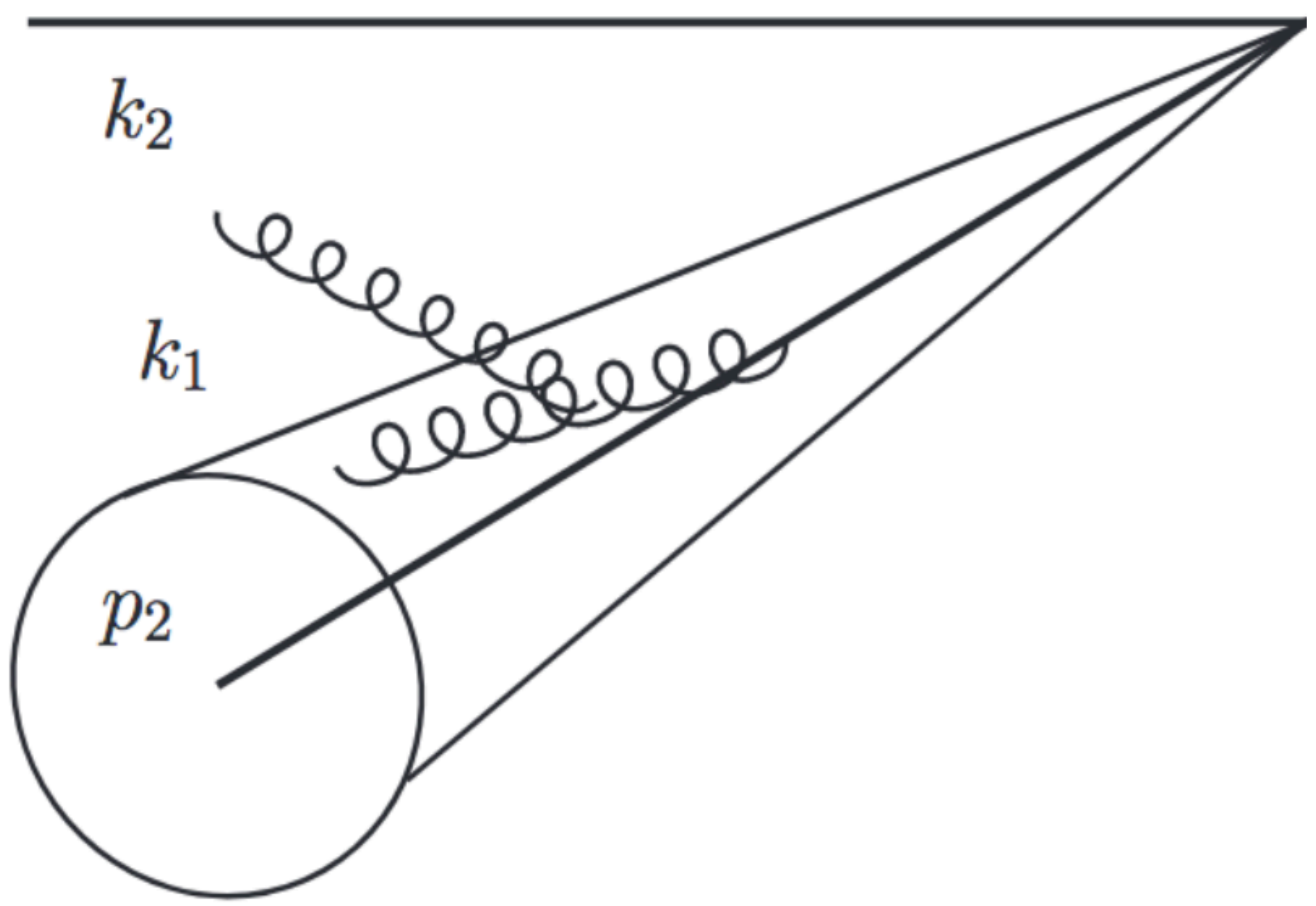
BROOKHAVEN NATIONAL LABORATORY

why jet substructure measurements?

- Jet substructure provides insight into several *different scales of QCD*
  - Can be used to understand everything from fixed order effects to parton showers to hadronization
- *Jet modeling* is one of the dominant sources of uncertainties for many analyses
  - Deeper understanding of jet formation can be used to develop better models of jets, and to provide better tuning of Monte Carlo predictions

why jet substructure measurements?

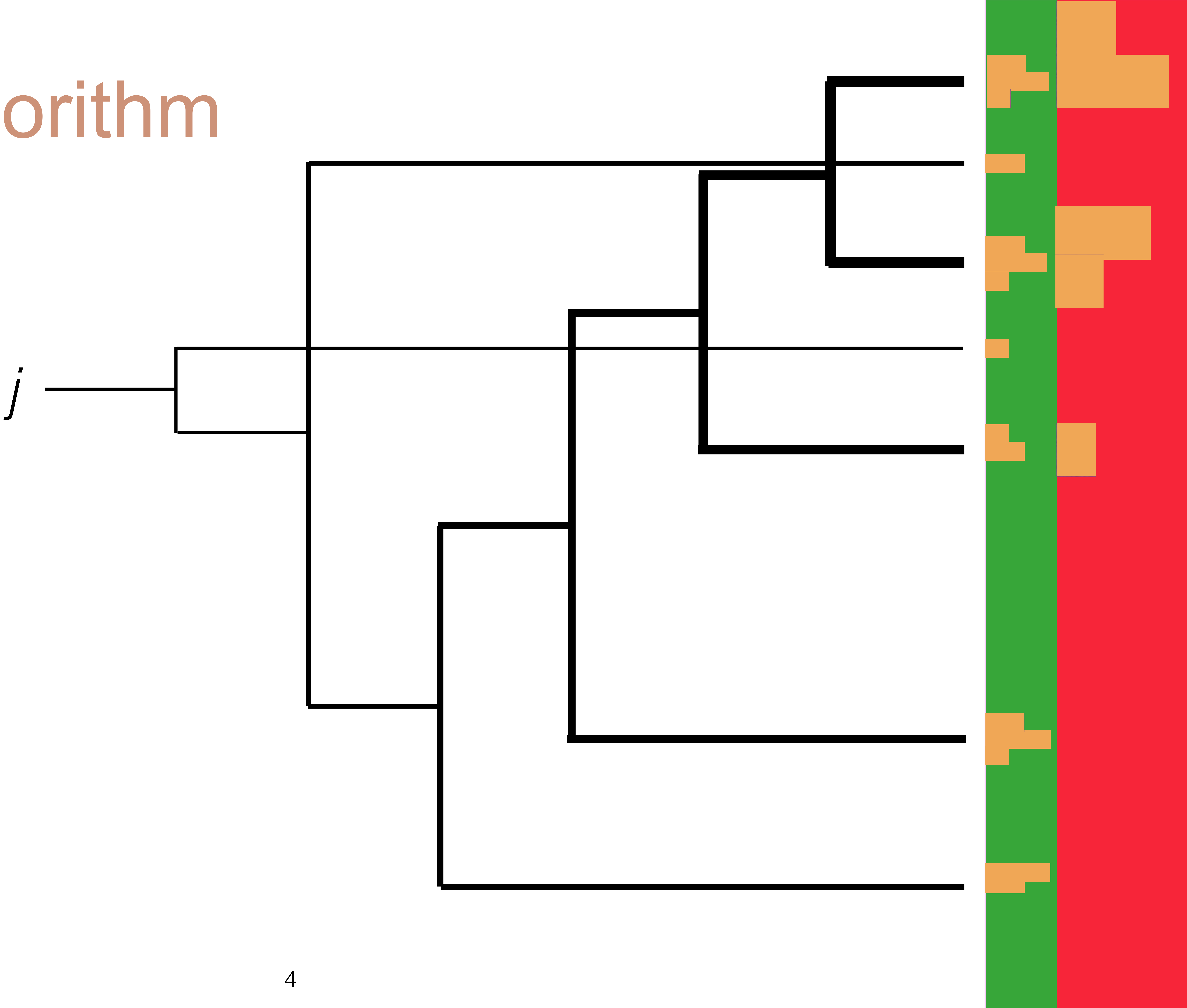




- Calculations of substructure observables are complicated by the presence of non-global logarithms
  - These tend to be soft and wide-angle radiation
- Grooming algorithms remove soft and wide-angle radiation from jets
  - ***Soft drop*** is a grooming algorithm which removes these non-global logarithms → able to perform precision calculations to beyond leading logarithmic accuracy
- Enables precision measurements of Standard Model parameters like the top mass or the strong coupling constant α_s
- Provides measurements sensitive to different parton shower models

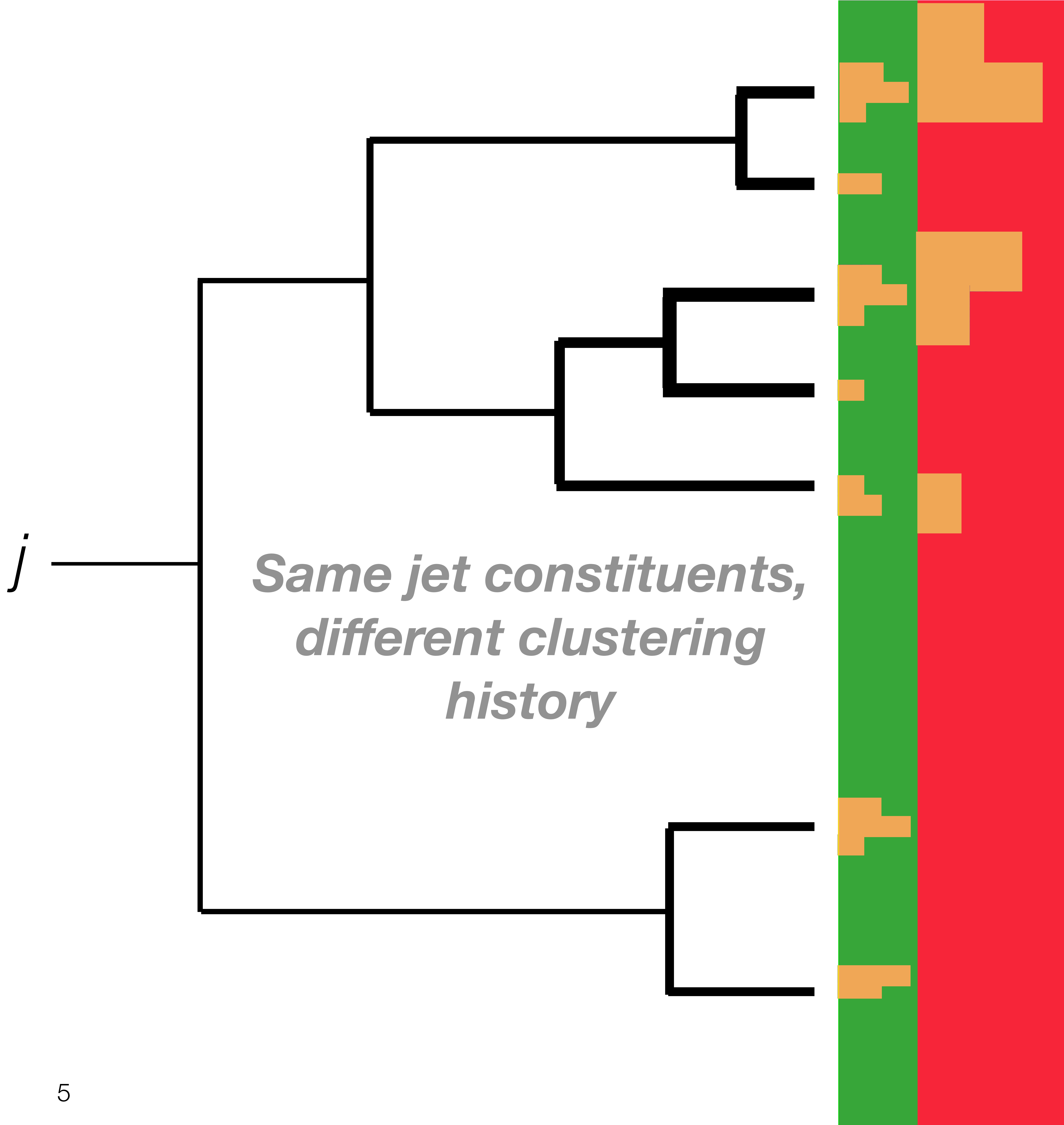
the soft drop algorithm

- Run jet finding using the **anti-k_t algorithm**

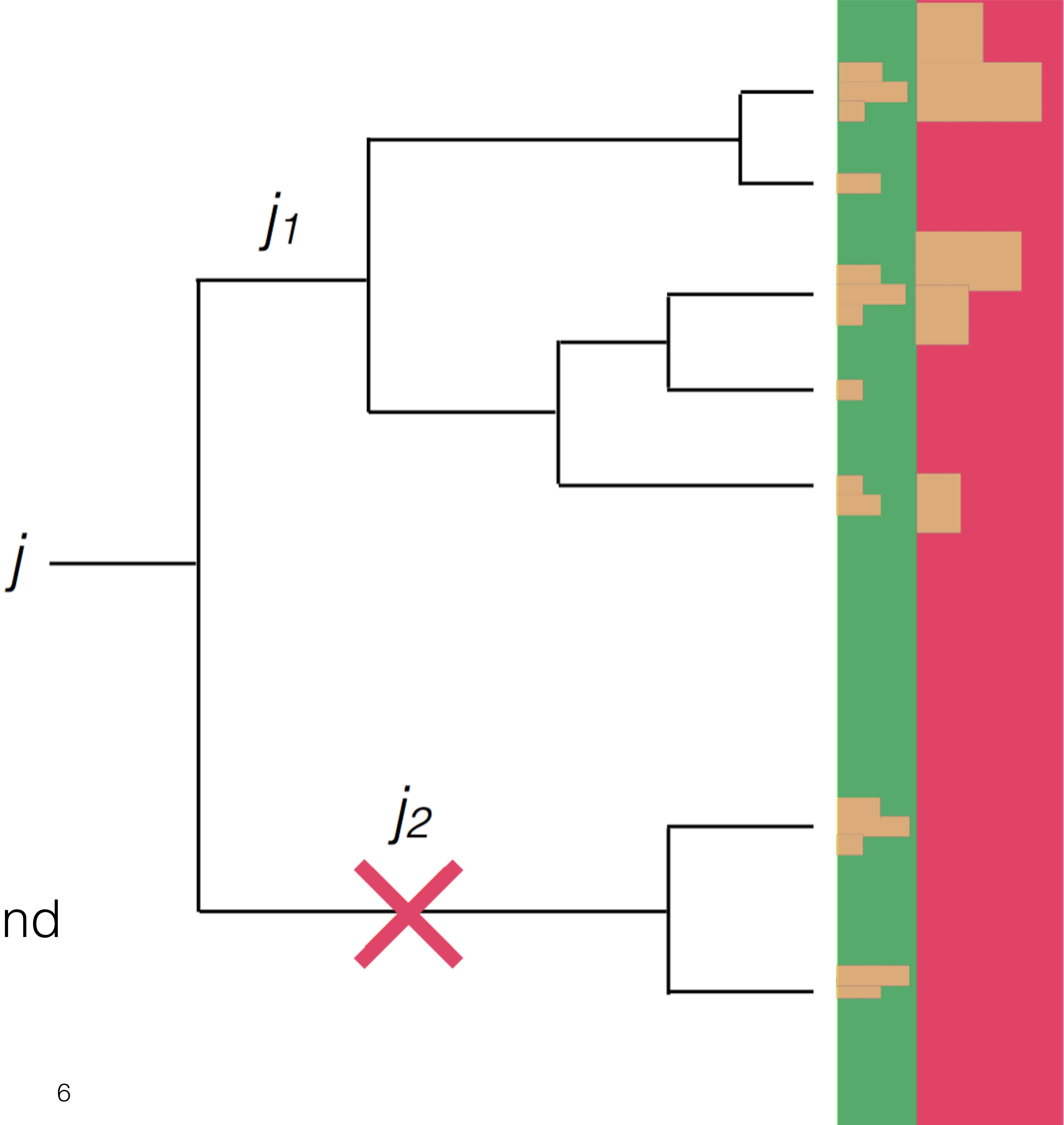


the soft drop algorithm

- Recluster its constituents with the Cambridge/Aachen algorithm to get an **angular-ordered shower history**

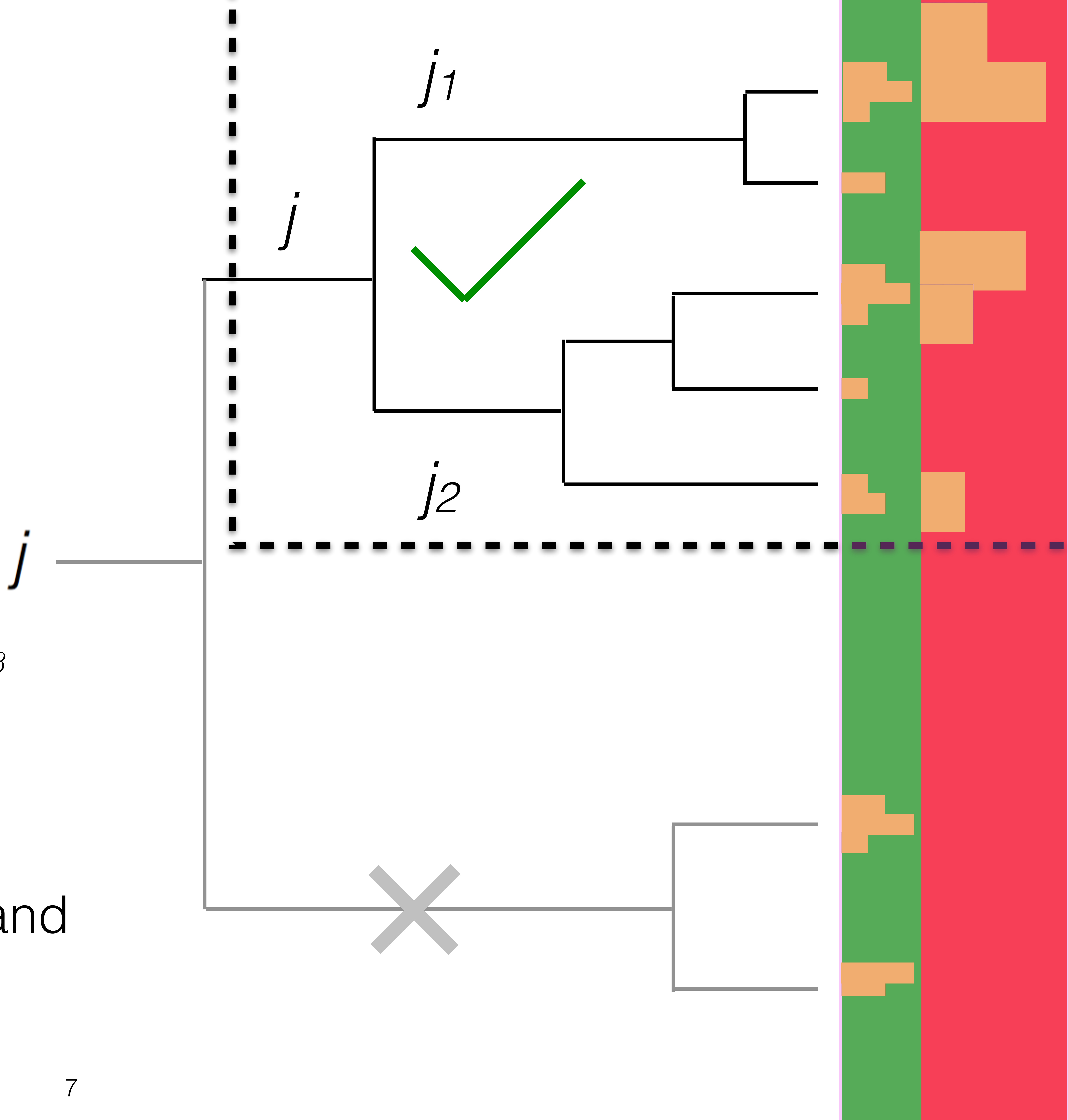


the soft drop algorithm



- Check if $$\frac{min(p_{T,j1}, p_{T,j2})}{(p_{T,j1} + p_{T,j2})} > z_{cut}(\frac{\Delta R_{j1,j2}}{R})^{\beta}$$
    - If not, drop the softer branch (*j2*), and repeat with the harder branch (*j1*)

the soft drop algorithm

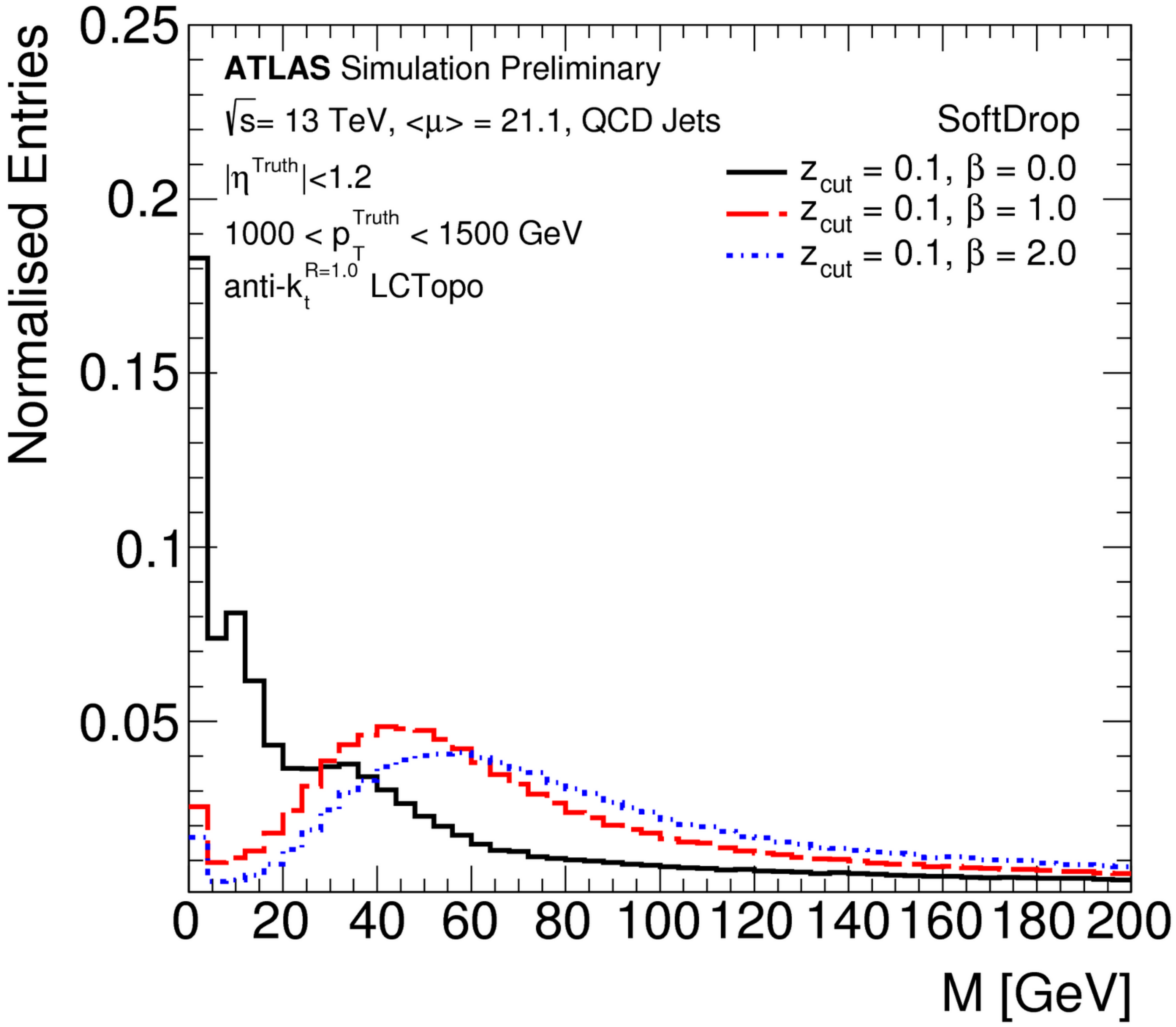


- Check if $\frac{min(p_{T,j1}, p_{T,j2})}{(p_{T,j1} + p_{T,j2})} > z_{cut}(\frac{\Delta R_{j1,j2}}{R})^{\beta}$
  - If not, drop the softer branch (*j2*), and repeat with the harder branch (*j1*)

the soft drop algorithm

$$\frac{min(p_{T,j1}, p_{T,j2})}{(p_{T,j1} + p_{T,j2})} > z_{cut}(\frac{\Delta R_{j1,j2}}{R})^{\beta}$$

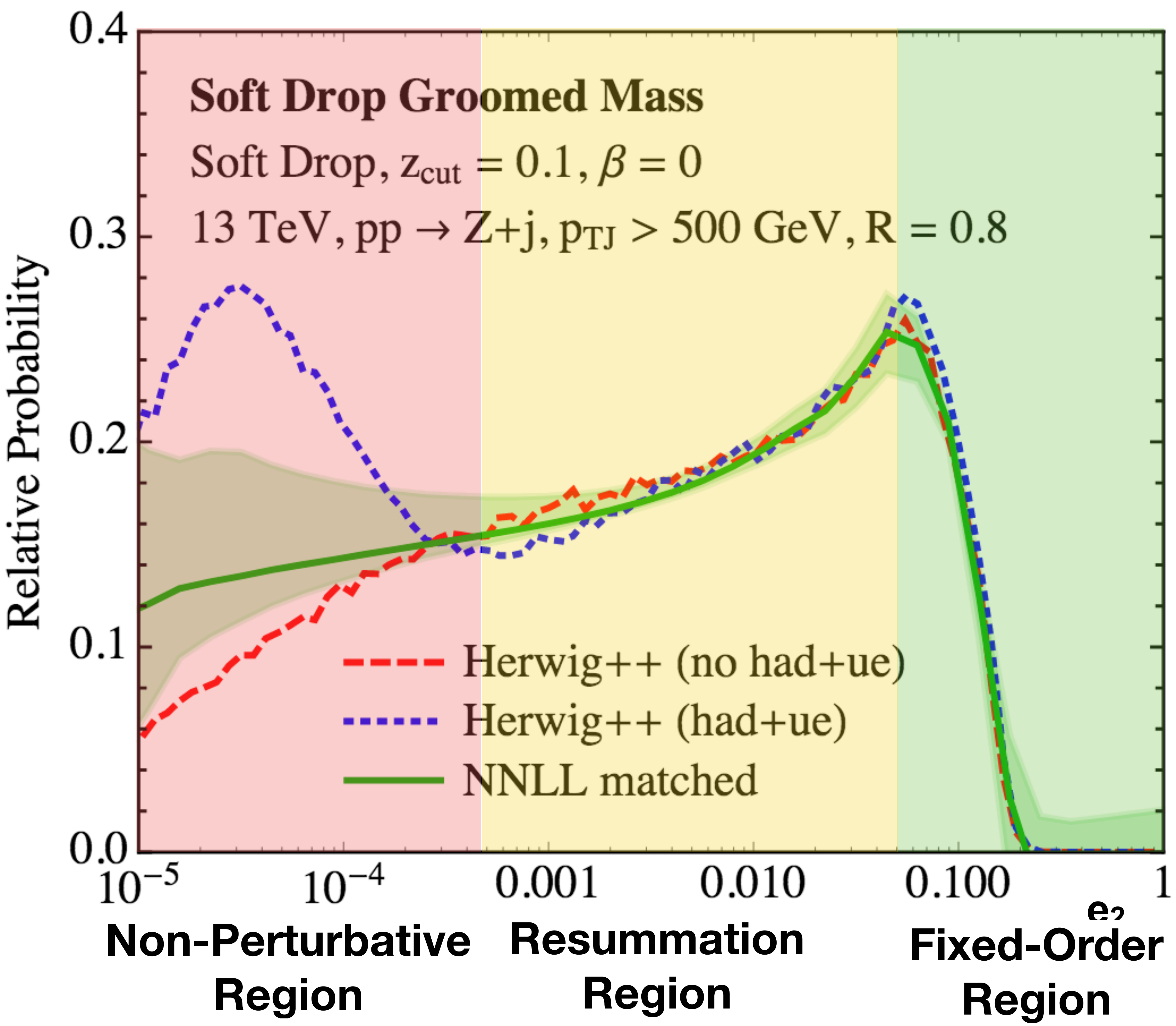
- Two free parameters: z_{cut} and β
- z_{cut} sets the **scale of energy removal**
  - Larger values of z_{cut} mean the more of the jet is groomed away
- β determines the **sensitivity to wide-angle radiation**
  - Smaller values of β mean that more aggressive grooming is applied




ATLAS
EXPERIMENT

the jet mass

- The jet mass is one of the most commonly used jet substructure observables
- Quarks and gluons are very light, but jets can be very massive because of fragmentation
- Measuring $\boldsymbol{\rho} = \log[(m^{\text{Soft Drop}} / p_T^{\text{Ungroomed}})^2]$
  - Using m / p_T results in less dependence on underlying p_T spectrum



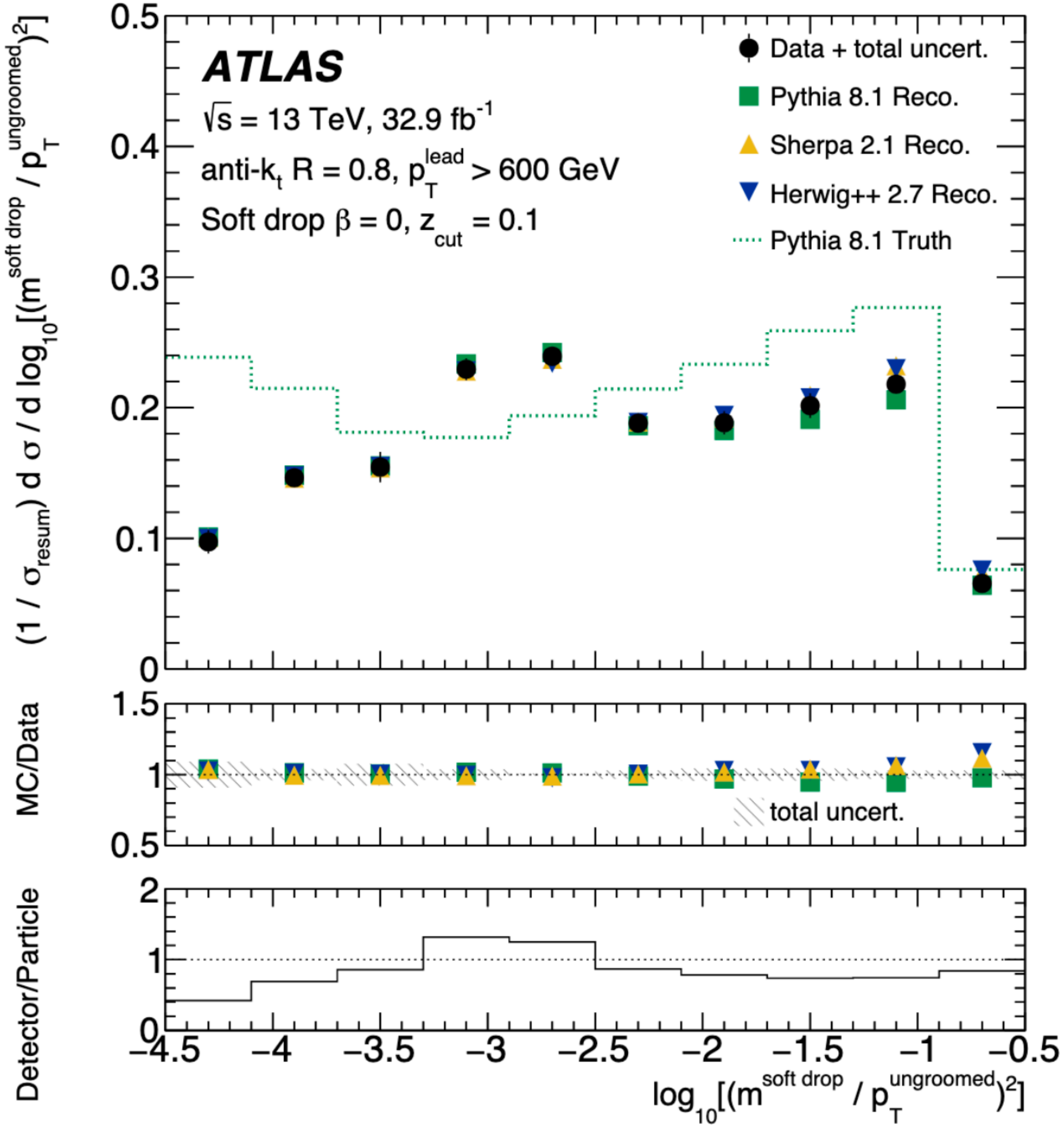
C. Frye, A. Larkoski, M. Schwartz, K. Yan, JHEP 07 (2016) 064

- High-mass region dominated by single hard splitting
  - Use *log-scale binning* to understand the resummation region
- The jet mass calculation is *factorizable* → different effects dominant in specific places

the jet mass

[1711.08341](1711.08341)

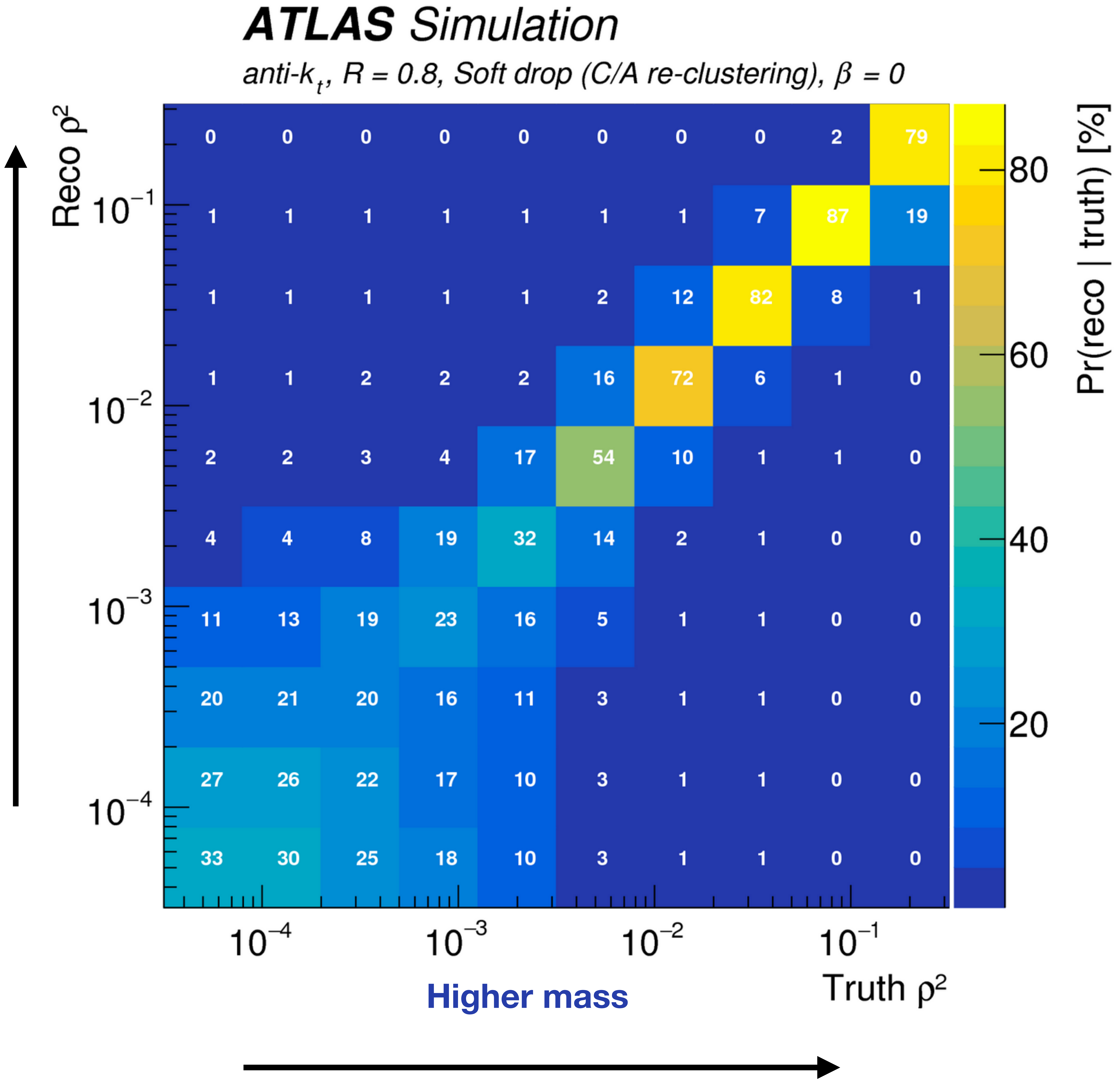
- The calorimeter-based jet mass is affected by **non-trivial detector corrections**
- Unfolding creates a mapping between a detector-level measurement and a truth-level distribution
  - Corrects for several detector effects, reconstruction efficiencies, and fake rates
- **Simultaneously unfold ρ and p_T** using Bayesian unfolding



the jet mass

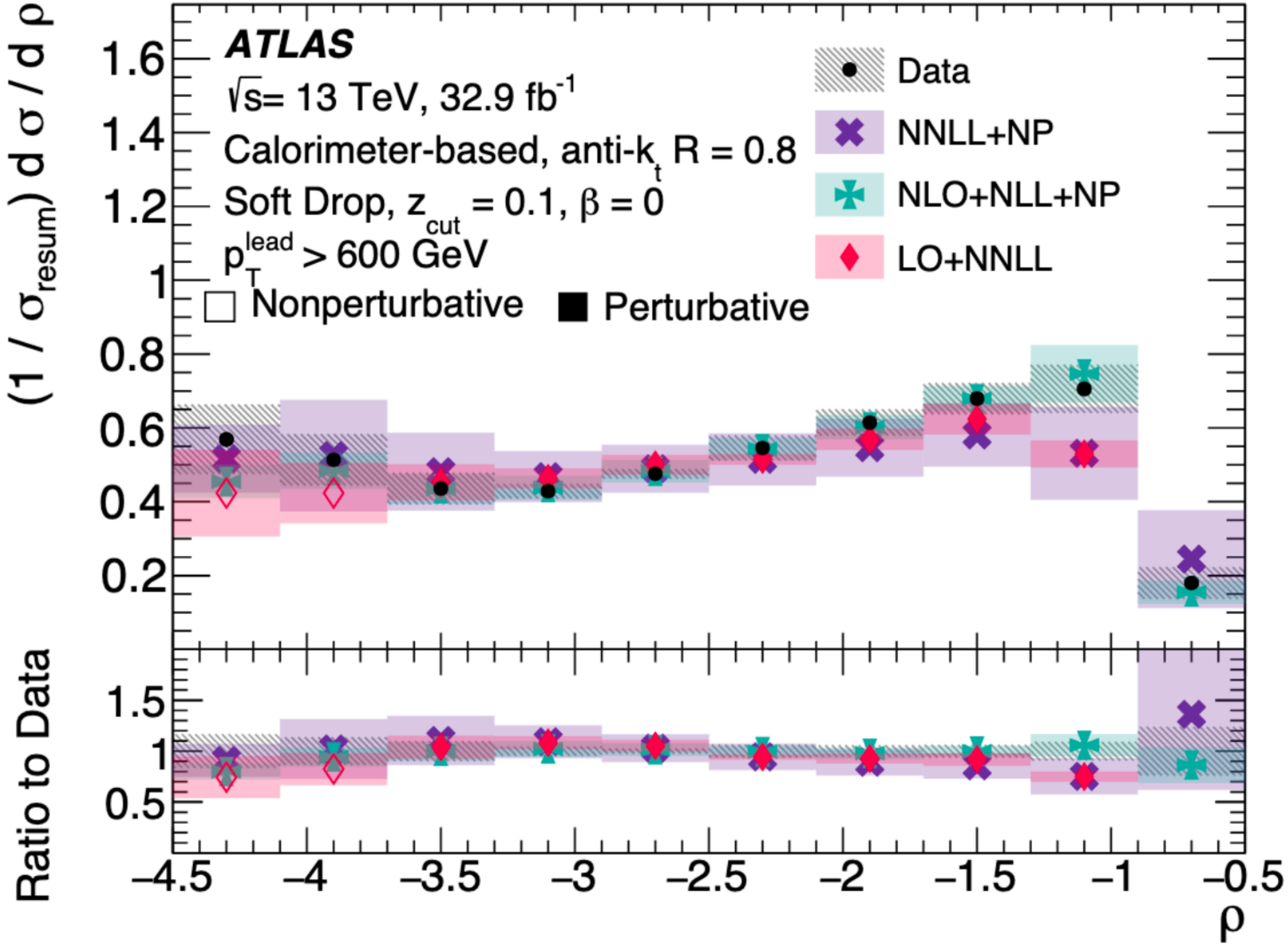
1711.08341

- The calorimeter-based jet mass is affected by **non-trivial detector corrections**
- Unfolding creates a mapping between a detector-level measurement and a truth-level distribution
  - Corrects for several detector effects, reconstruction efficiencies, and fake rates
- **Simultaneously unfold ρ and p_T** using Bayesian unfolding



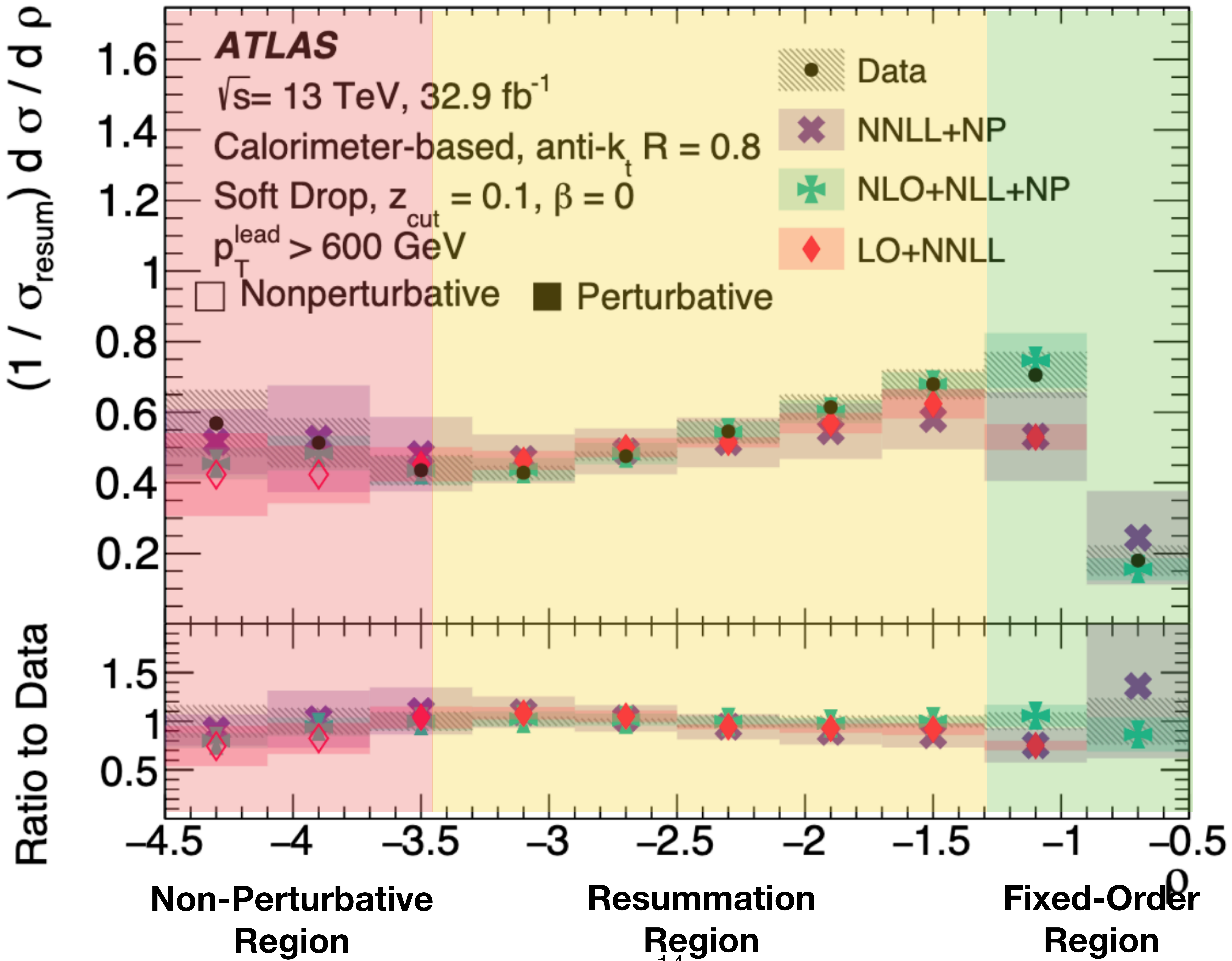
1912.09837

the unfolded jet mass



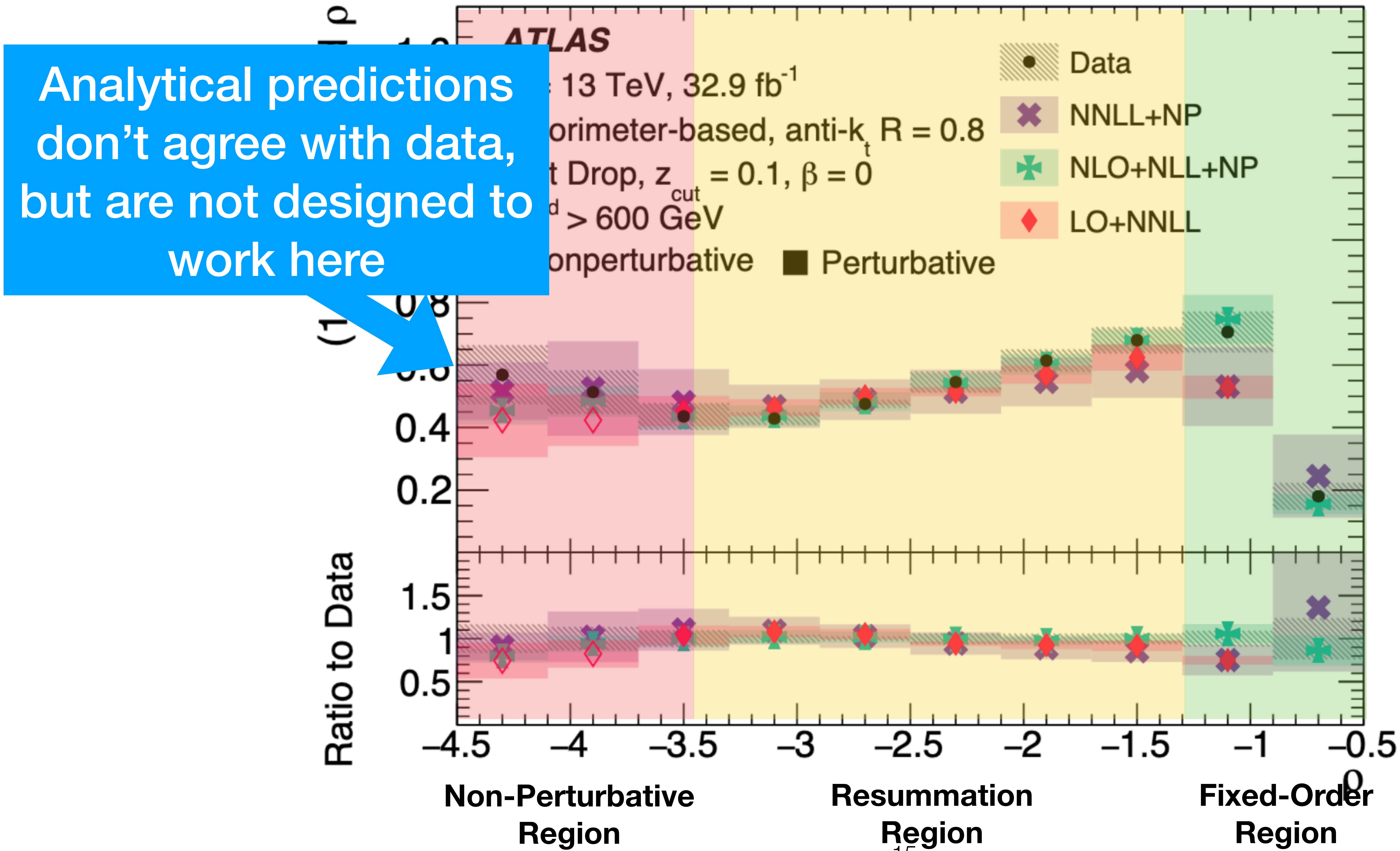
the unfolded jet mass

1912.09837



1912.09837

the unfolded jet mass

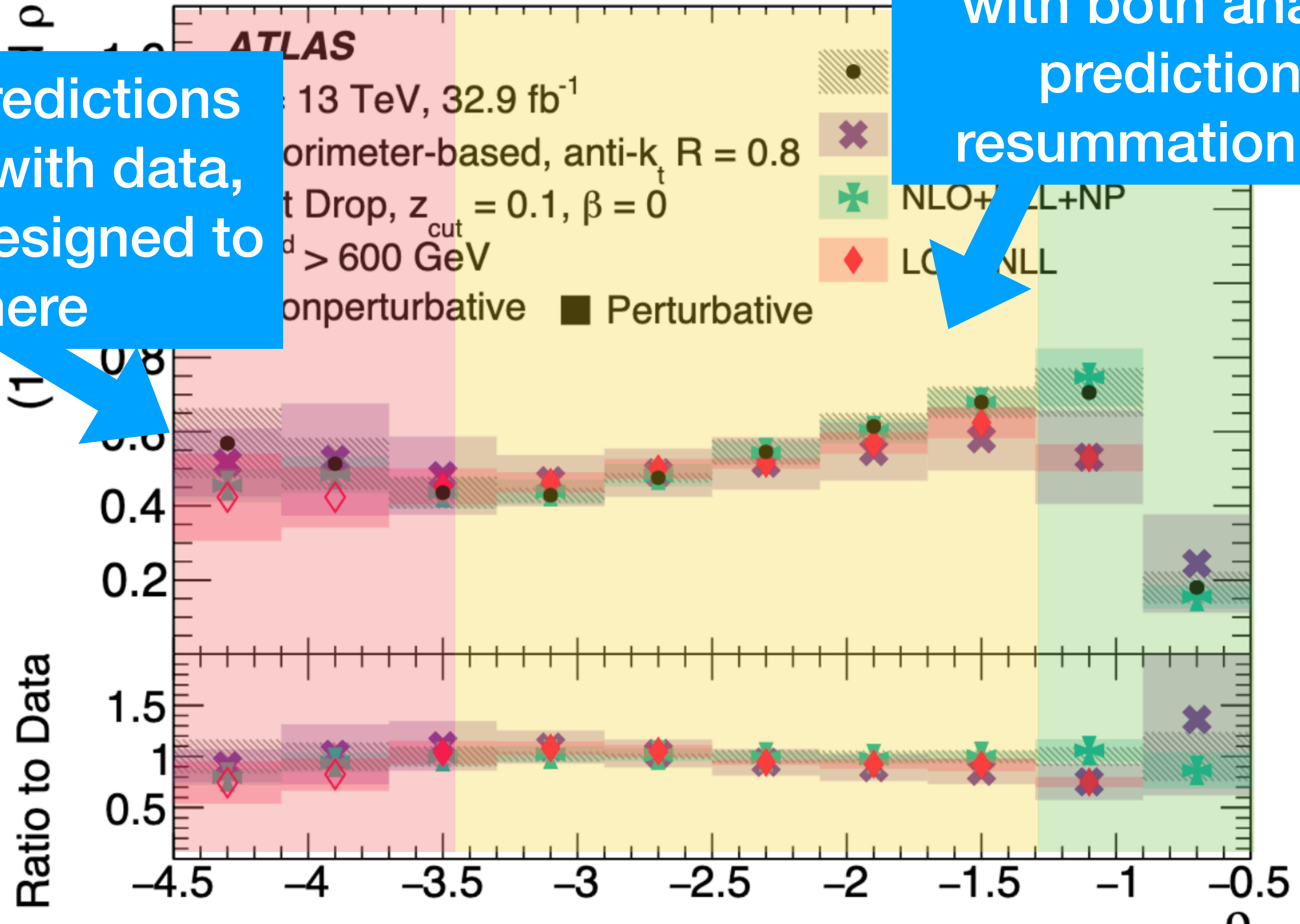


1912.09837

the unfolded jet mass

Analytical predictions don’t agree with data, but are not designed to work here

Very good agreement with both analytical predictions in resummation region



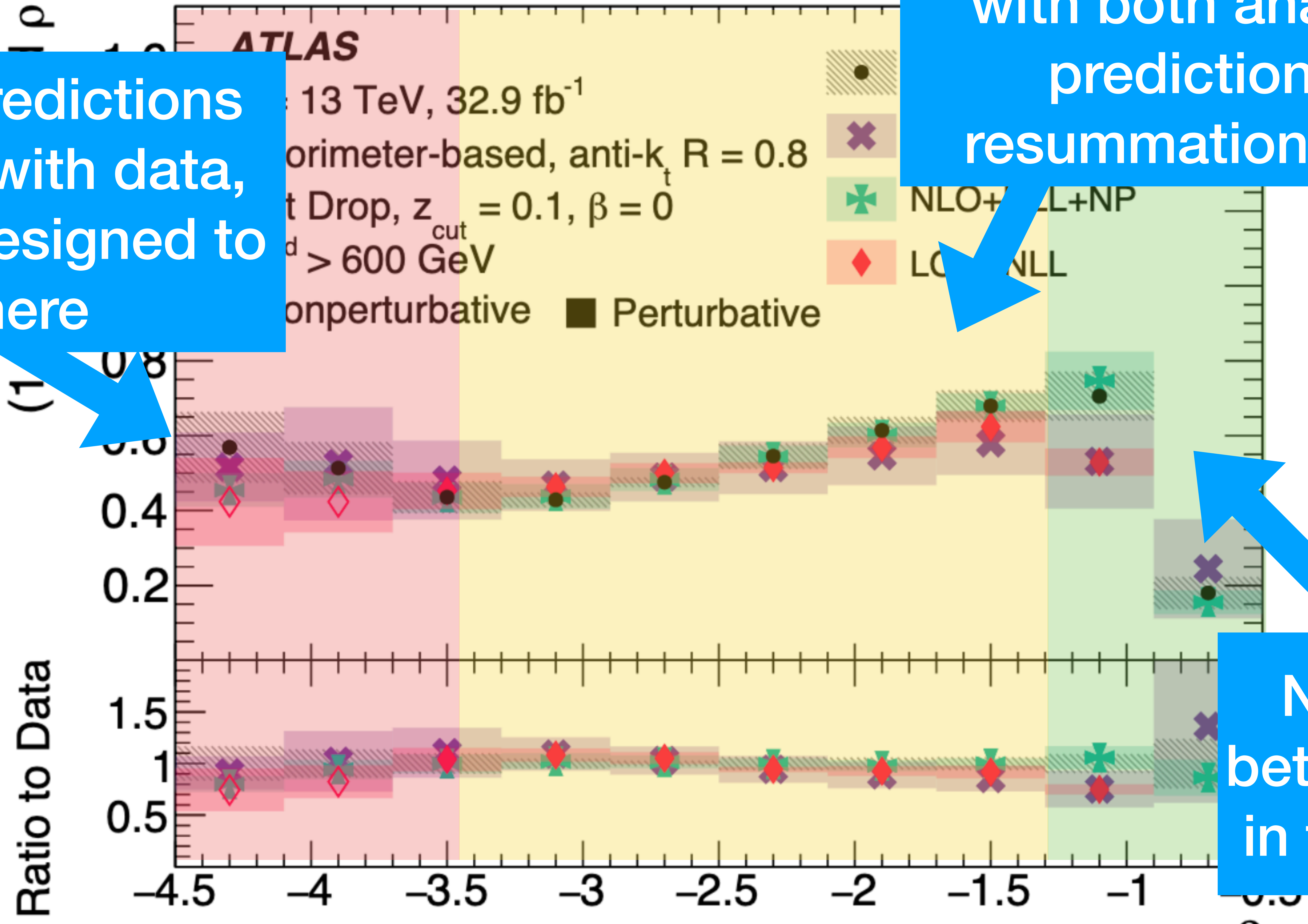
Non-Perturbative Region

Resummation Region

Fixed-Order Region

1912.09837

the unfolded jet mass



Analytical predictions don't agree with data, but are not designed to work here

Very good agreement with both analytical predictions in resummation region

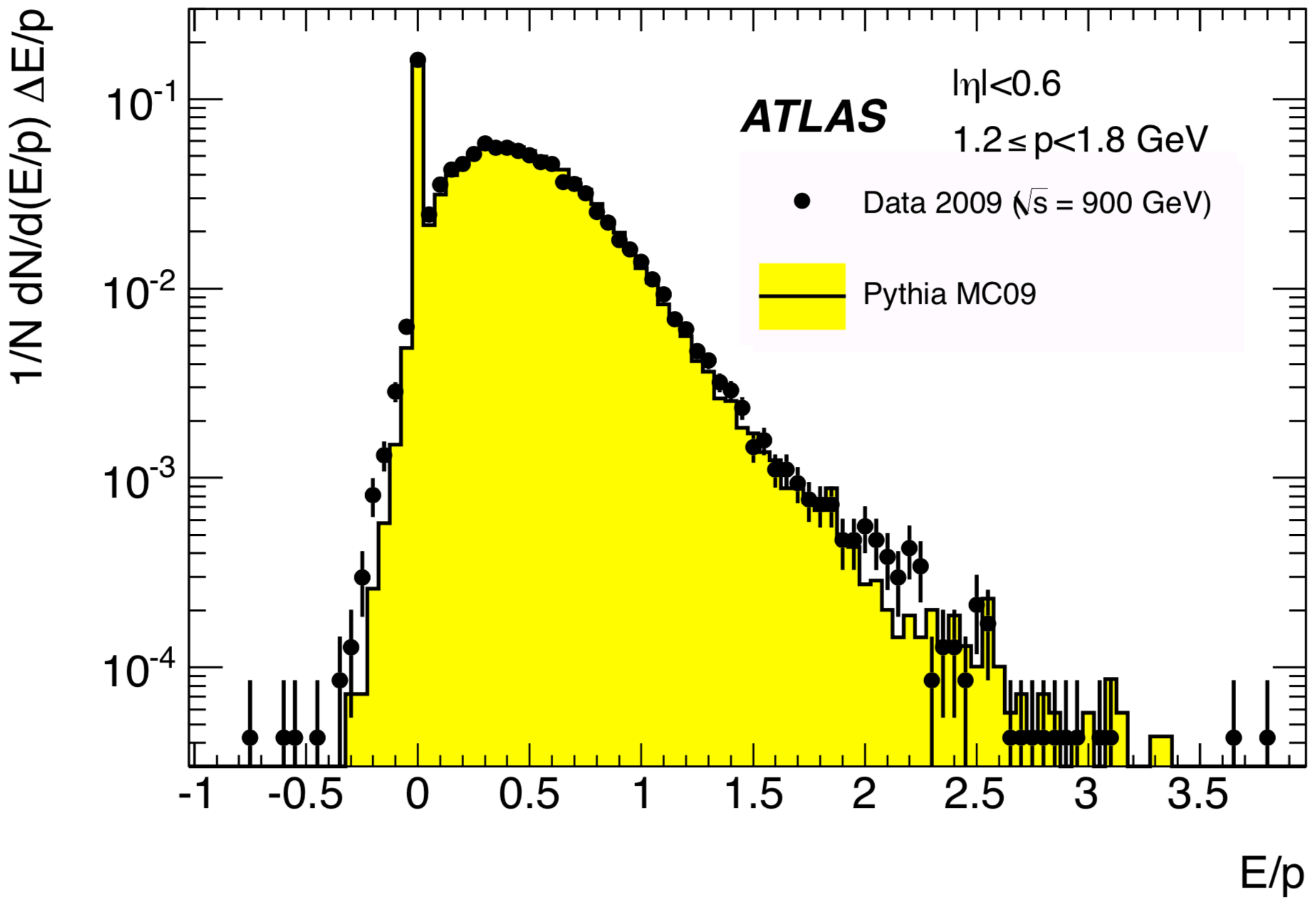
NLO+NLL agrees better than LO+NNLL in fixed order region

Non-Perturbative Region

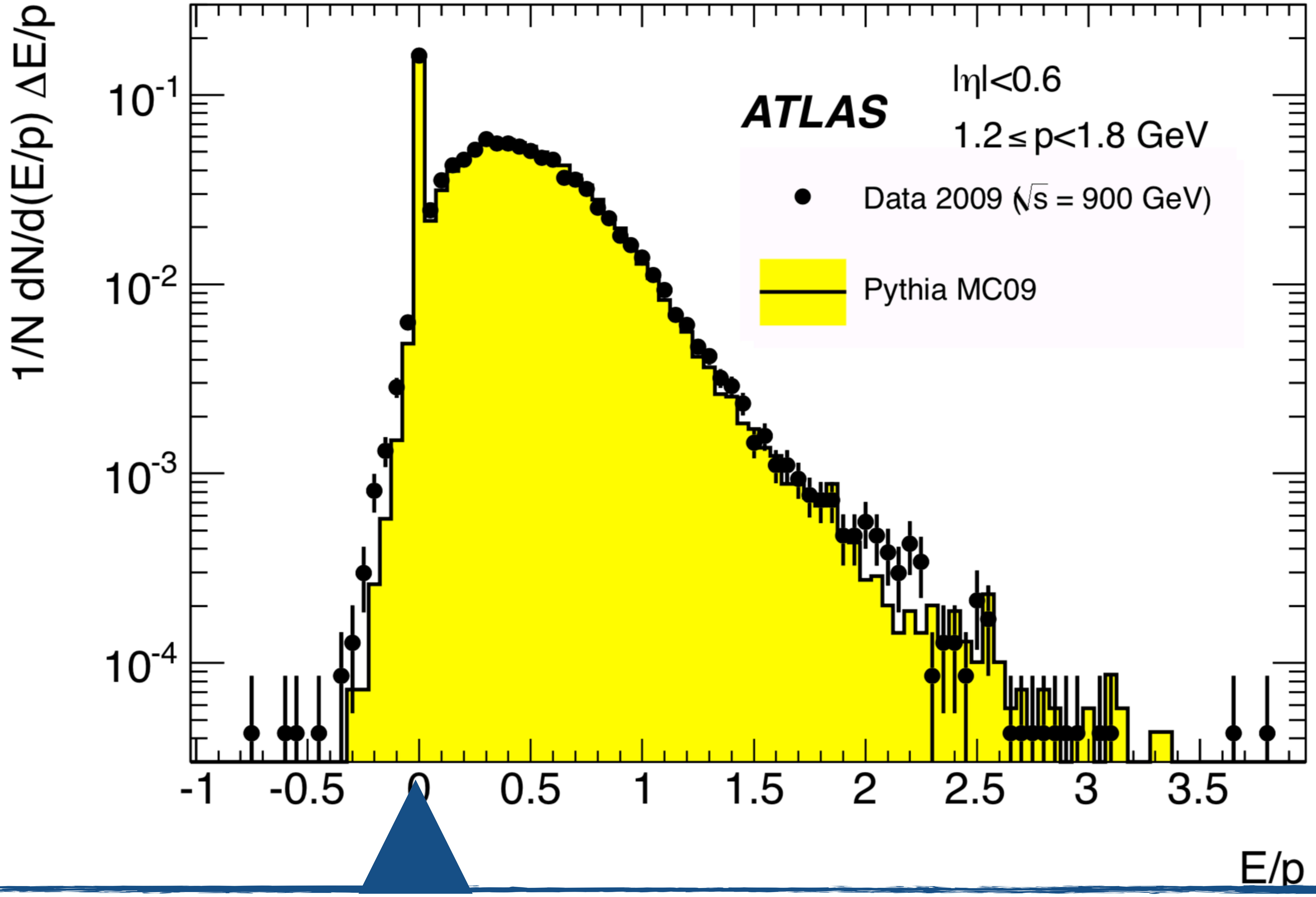
Resummation Region

Fixed-Order Region

cluster-based uncertainties



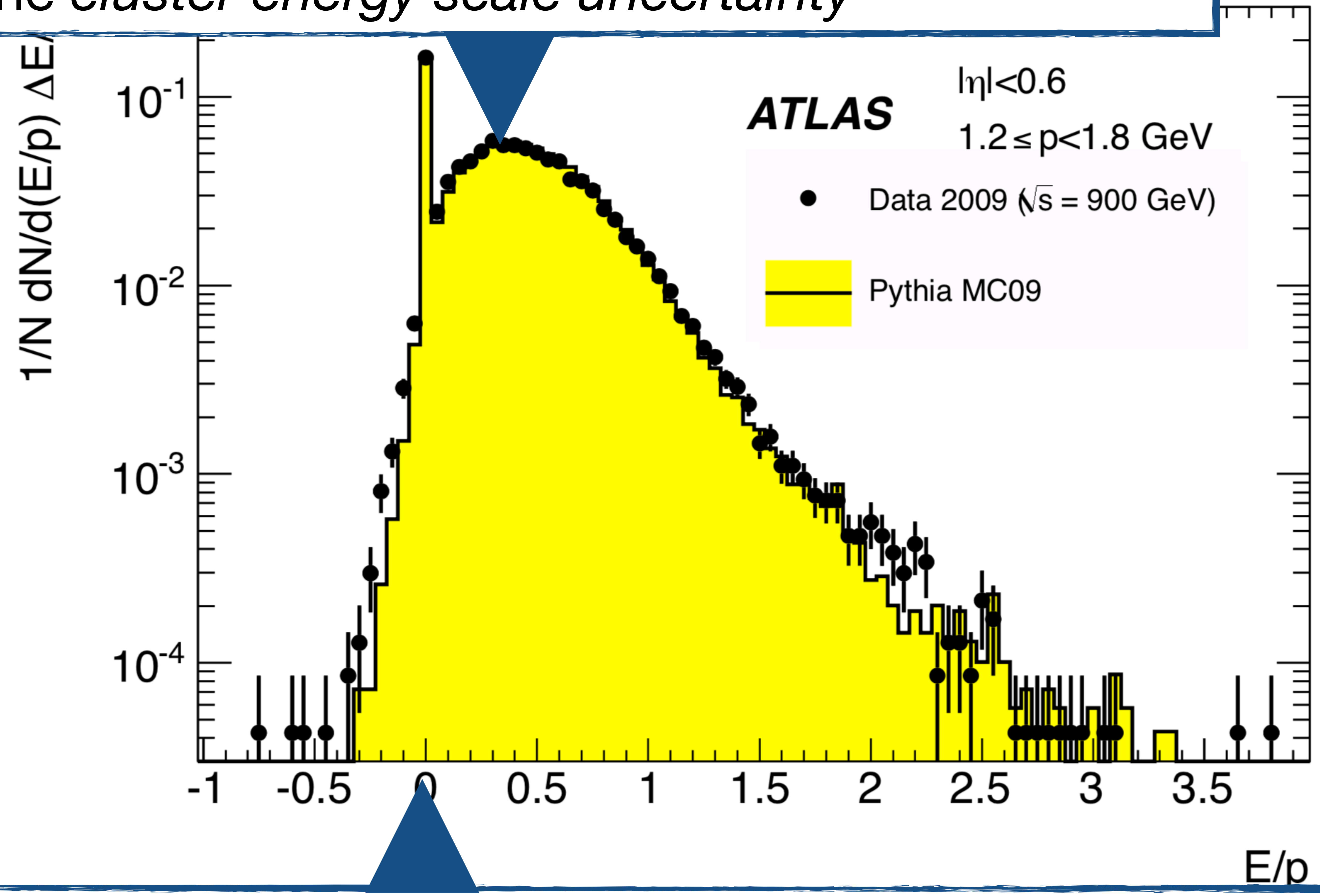
cluster-based uncertainties



- Percentage of events in the 0 bin describes the probability of not having a cluster → Data-MC difference gives *cluster efficiency uncertainty*

cluster-based uncertainties

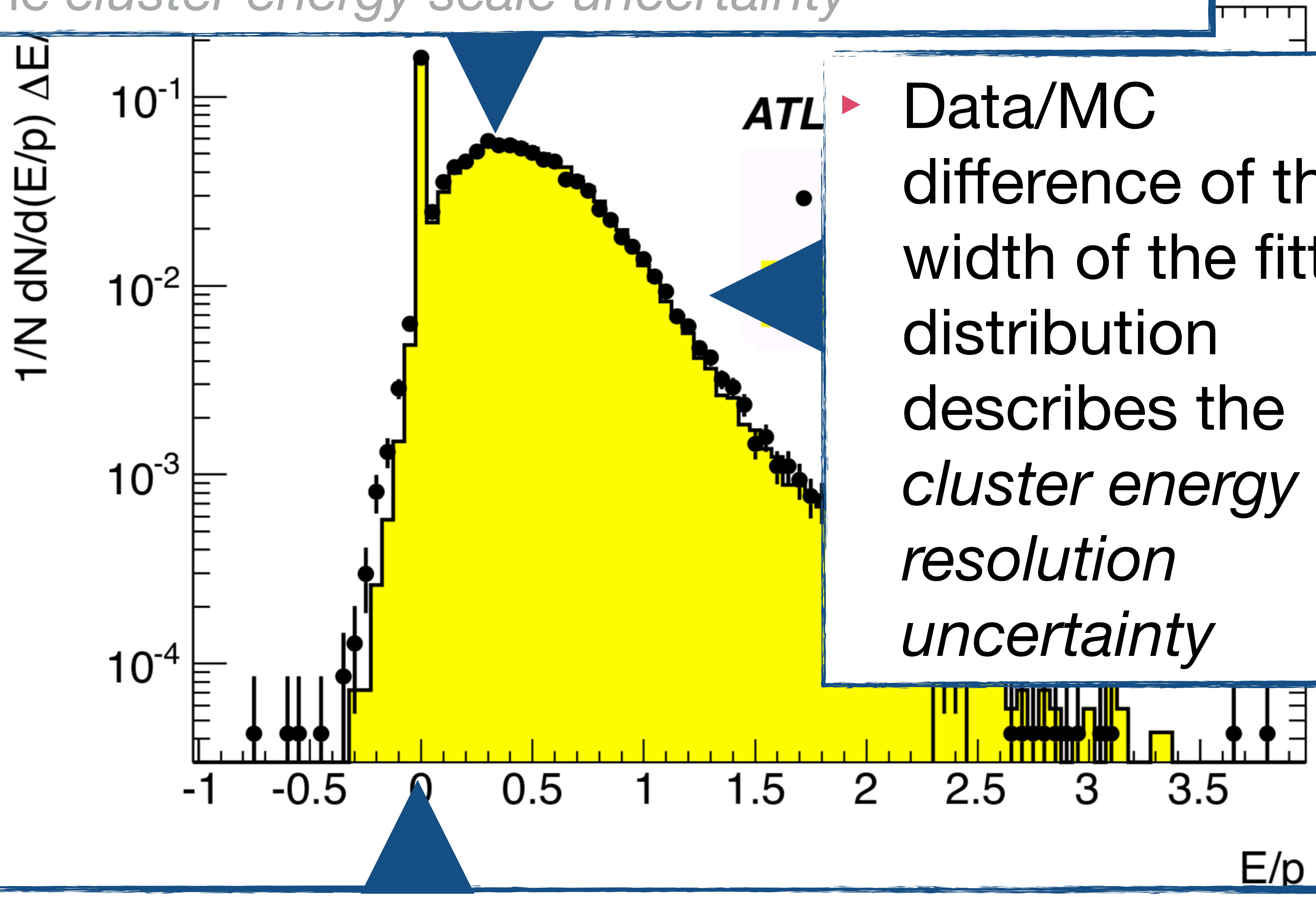
- Data/MC difference of the mean of the fitted distribution describes the *cluster energy scale uncertainty*



- Percentage of events in the 0 bin describes the probability of not having a cluster → Data-MC difference gives *cluster efficiency uncertainty*

cluster-based uncertainties

- Data/MC difference of the mean of the fitted distribution describes the *cluster energy scale uncertainty*

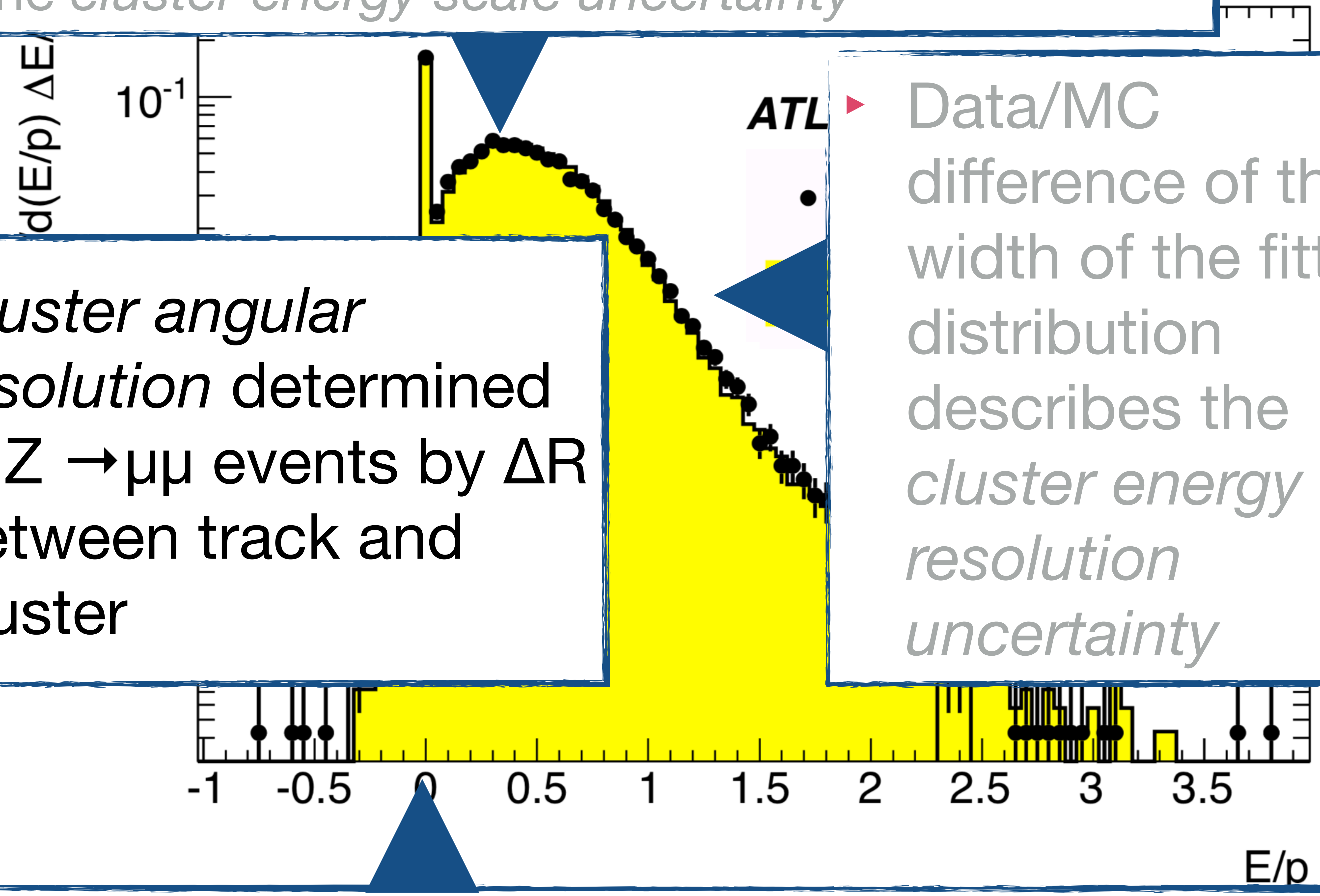


- Data/MC difference of the width of the fitted distribution describes the *cluster energy resolution uncertainty*

- Percentage of events in the 0 bin describes the probability of not having a cluster → Data-MC difference gives *cluster efficiency uncertainty*

cluster-based uncertainties

- Data/MC difference of the mean of the fitted distribution describes the *cluster energy scale uncertainty*



- *Cluster angular resolution* determined in Z →μμ events by ΔR between track and cluster

- Data/MC difference of the width of the fitted distribution describes the *cluster energy resolution uncertainty*

- Percentage of events in the 0 bin describes the probability of not having a cluster → Data-MC difference gives *cluster efficiency uncertainty*

cluster-based uncertainties

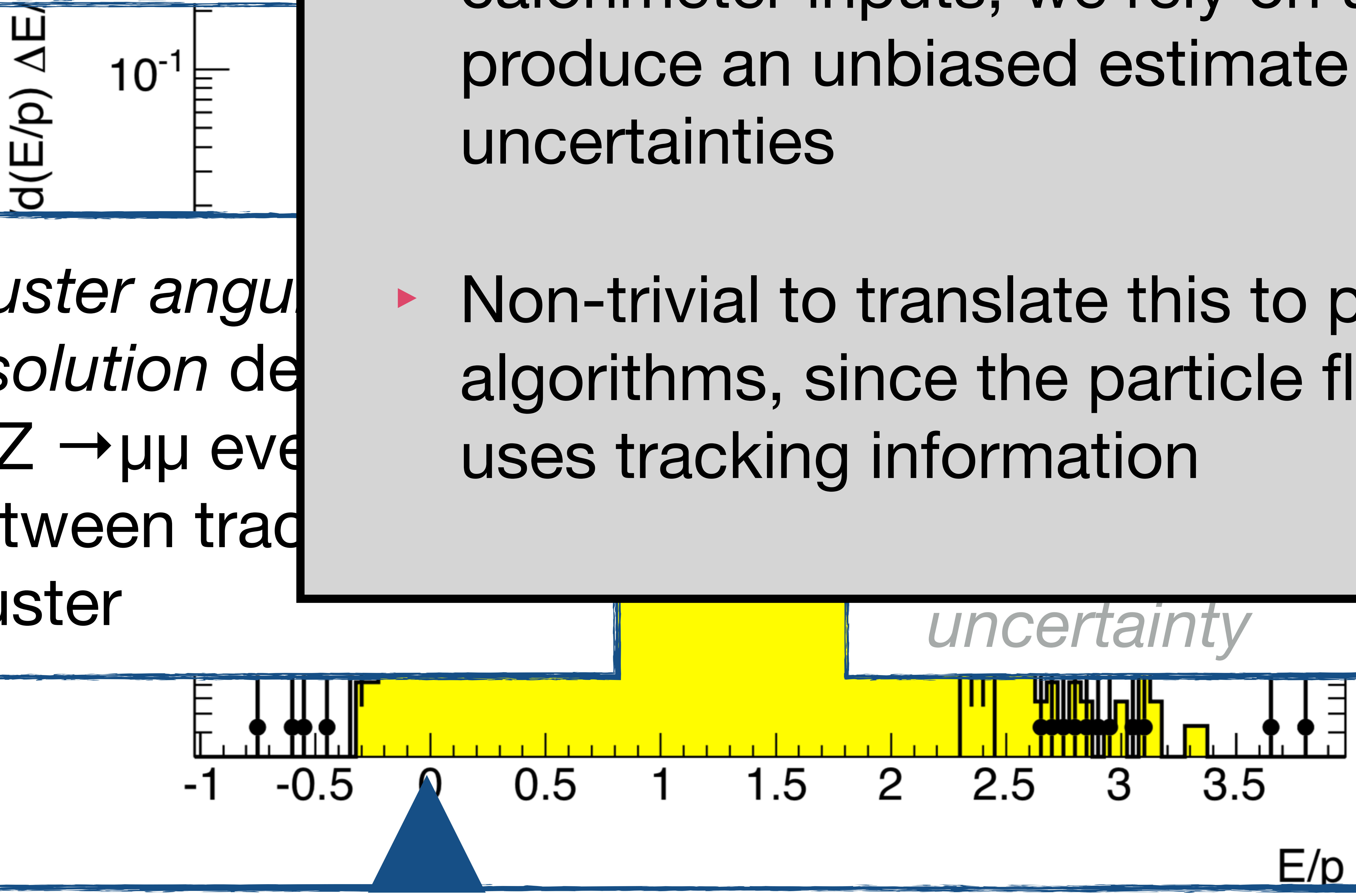
- Data/MC difference of t
describes the *cluster er*

- *Cluster angu*
resolution de
in Z →μμ eve
between trac
cluster

- For substructure measurements with calorimeter inputs, we rely on using tracks to produce an unbiased estimate of our uncertainties

- Non-trivial to translate this to particle flow algorithms, since the particle flow subtraction uses tracking information

uncertainty

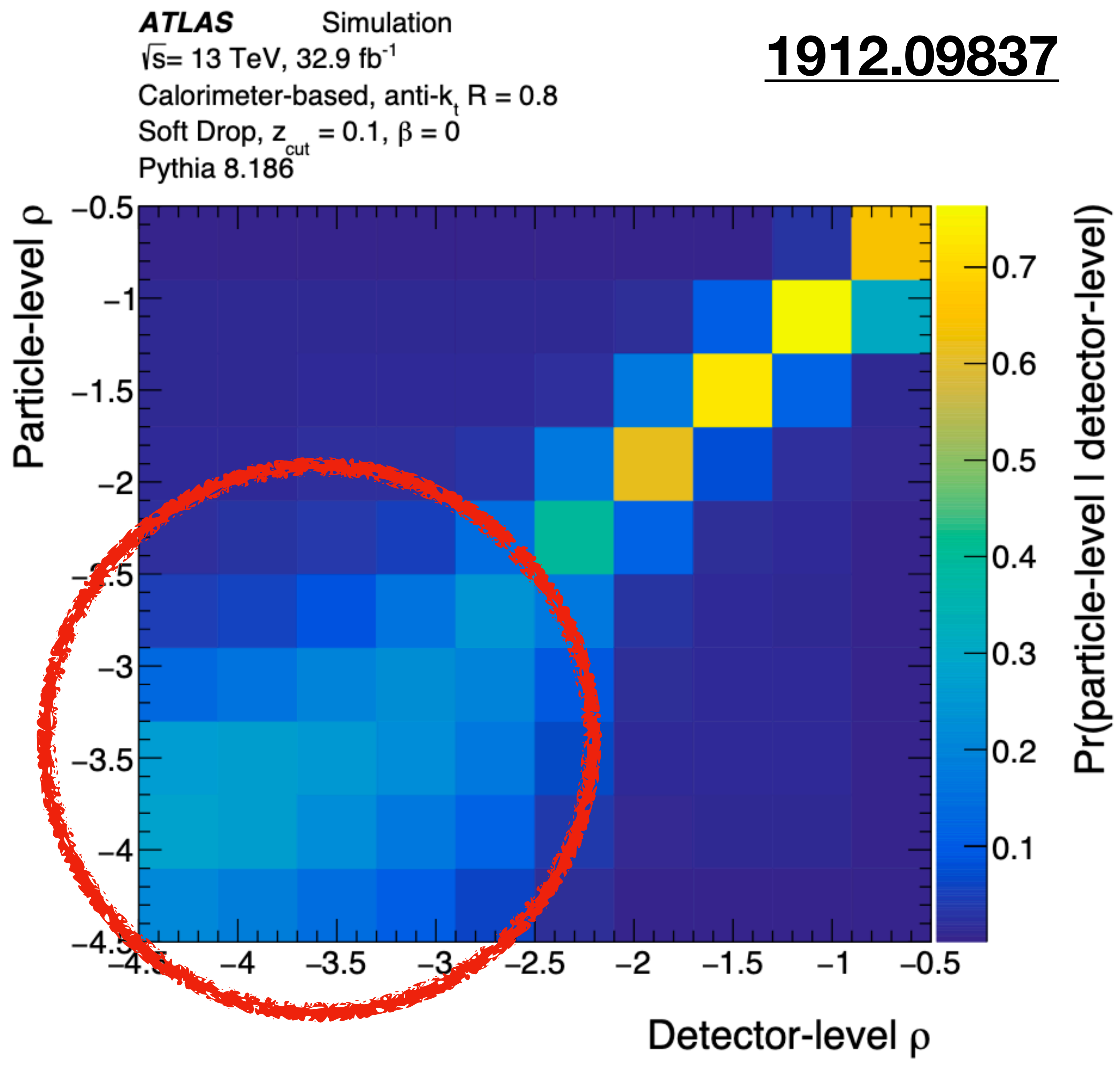


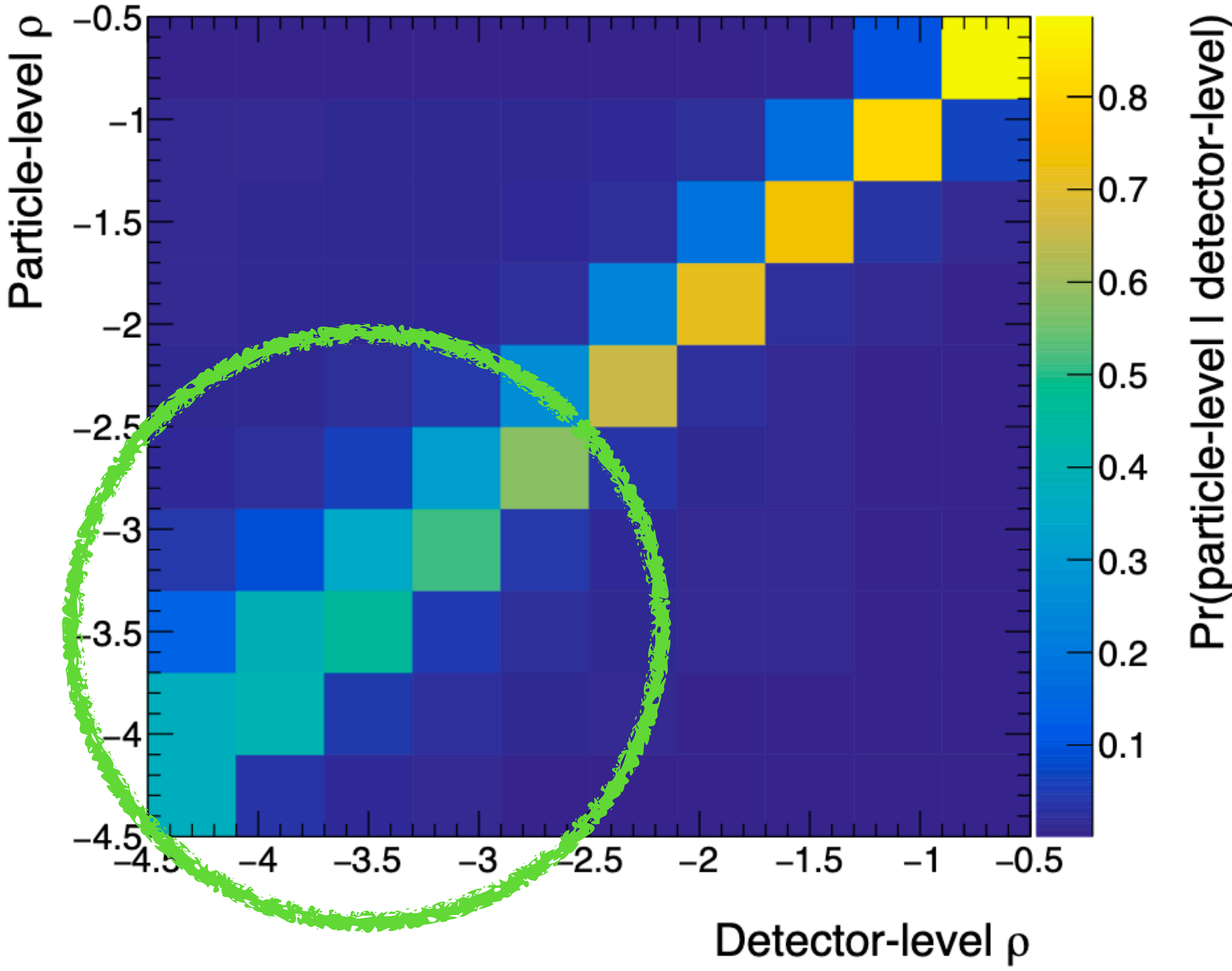
- Percentage of events in the 0 bin describes the probability of not having a cluster → Data-MC difference gives *cluster efficiency uncertainty*

the unfolded jet mass

1912.09837

- Can also measure the jet mass using only *charged* particles
  - The overall behavior should be very similar on average to the all-particle case, because of isospin symmetry
- Tracks have better angular resolution → smaller uncertainties in non-perturbative region
  - Smaller migrations in migration matrices, particularly in low-mass regions where angular resolution is important



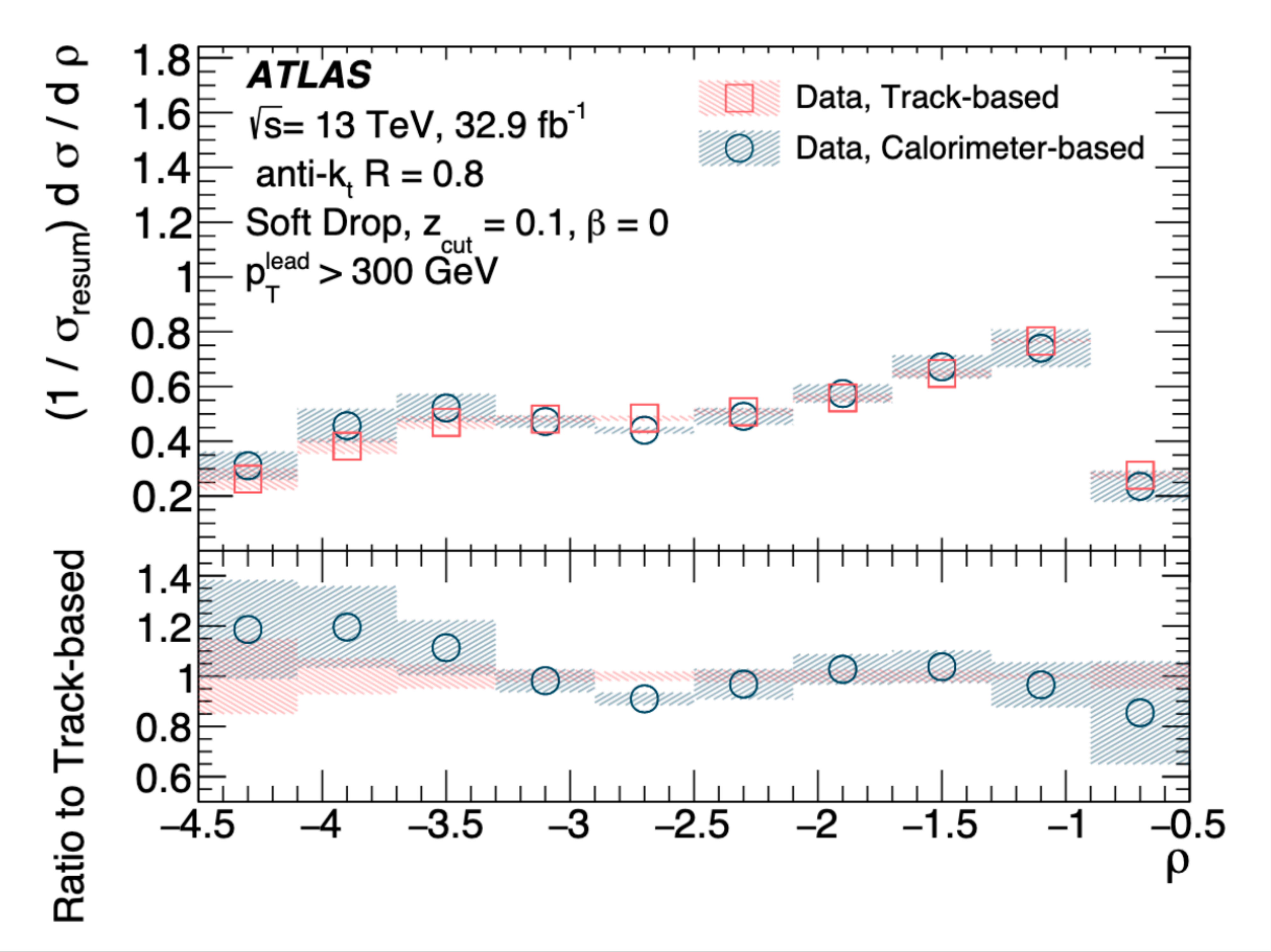


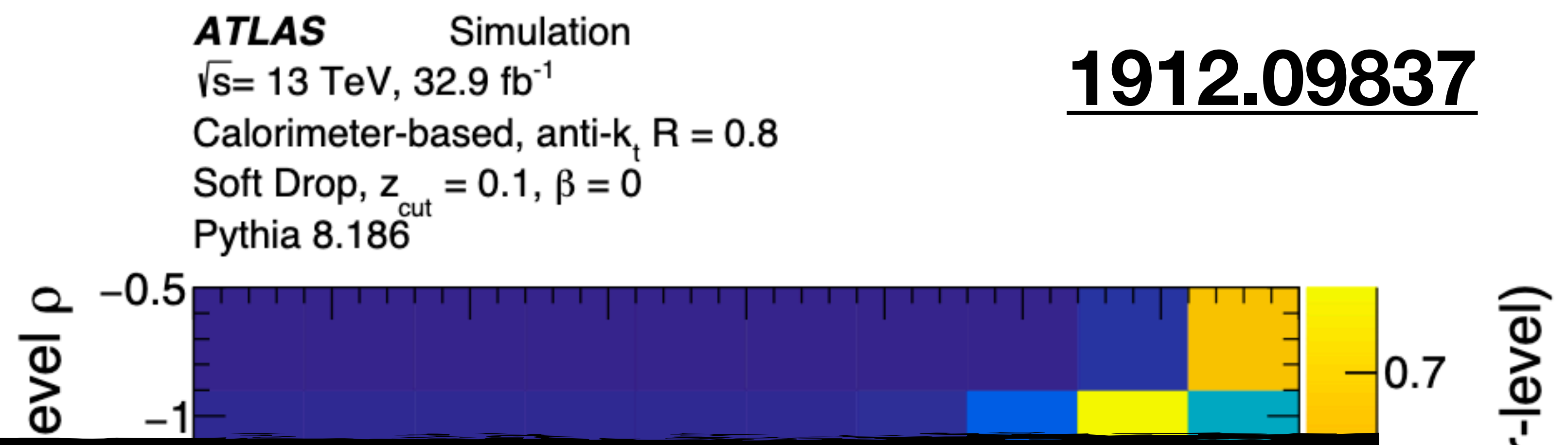
the unfolded jet mass

1912.09837

- Can al particl
  - The ave isos
- Track unce
  - Sma par resolution is important

- Some track and cluster observables look similar on average
  - Particularly true where non-perturbative effects are small
  - No calculations exist for track-based observables (yet), but would be powerful experimentally
- Track-based observables typically have much smaller uncertainties

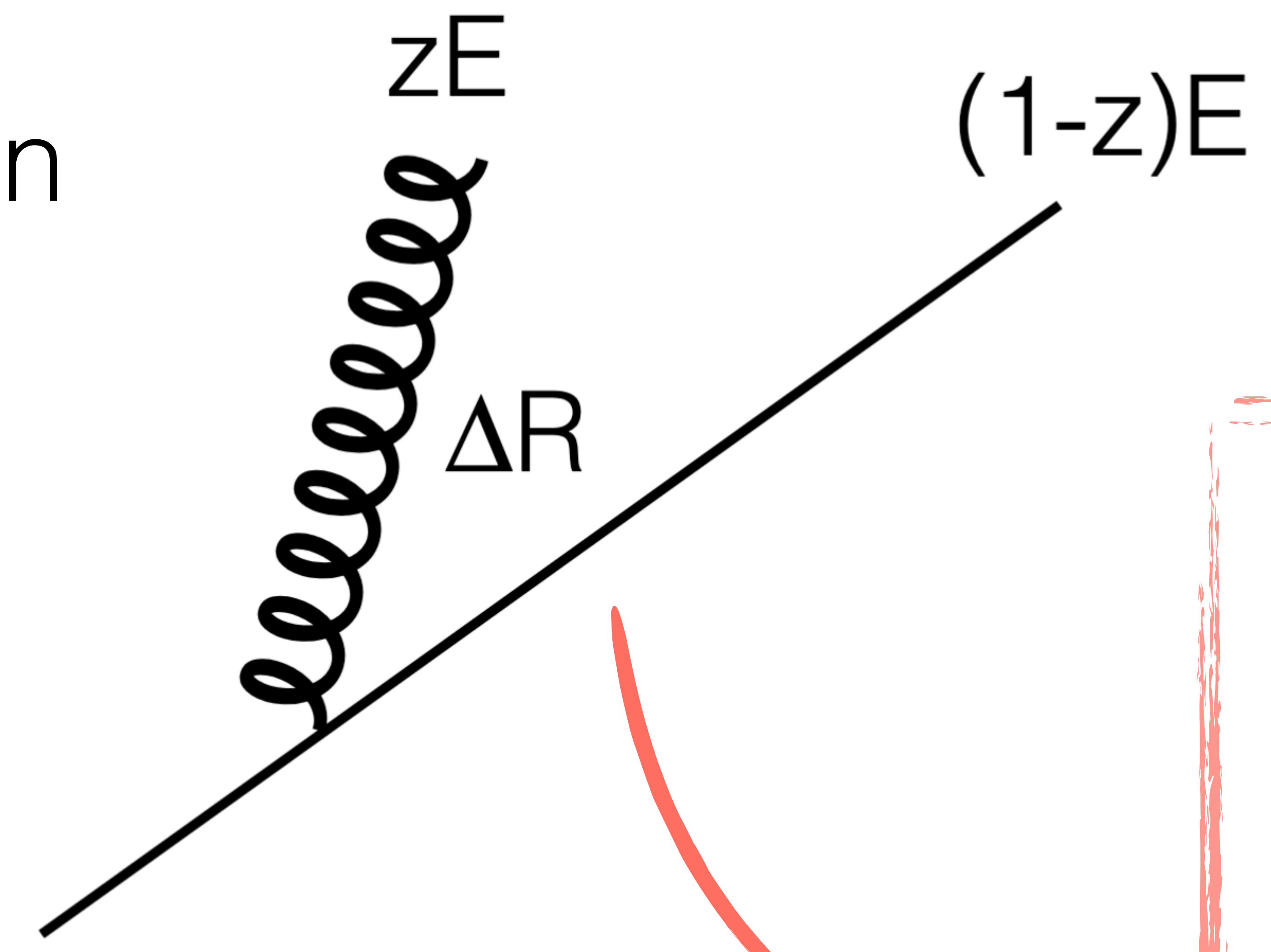




jet substructure

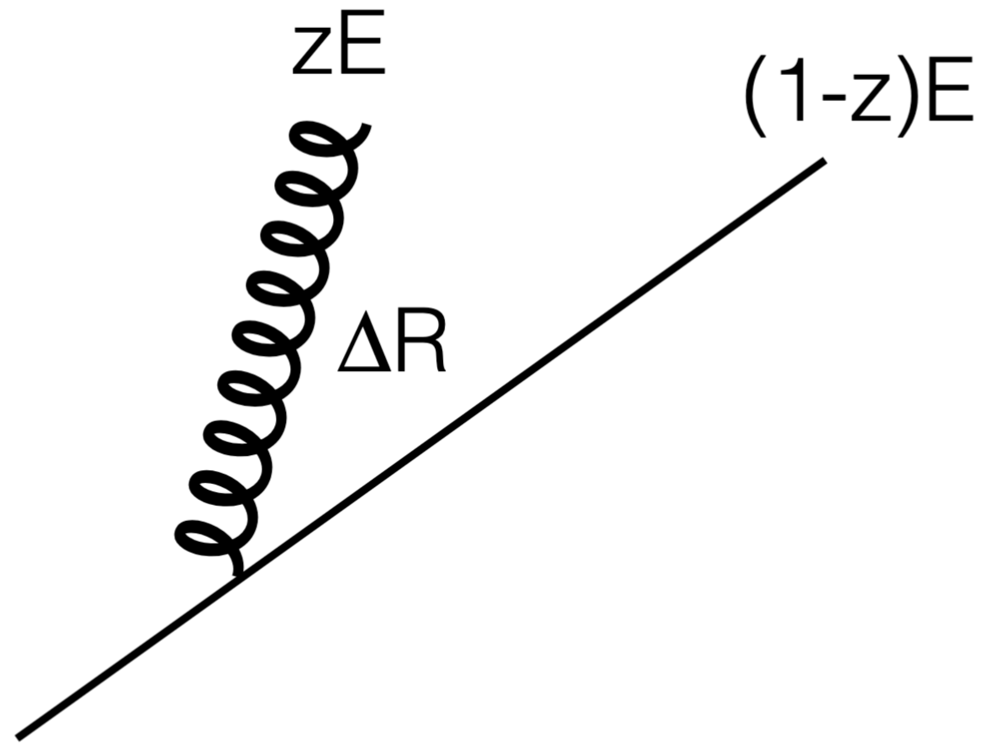
the lund plane

- A jet may be approximated as soft emissions around a hard core which represents the originating quark or gluon
- Emissions may be characterized by
  - z = relative momentum of emission relative to the jet core
  - ΔR = angle of emission relative to the jet core

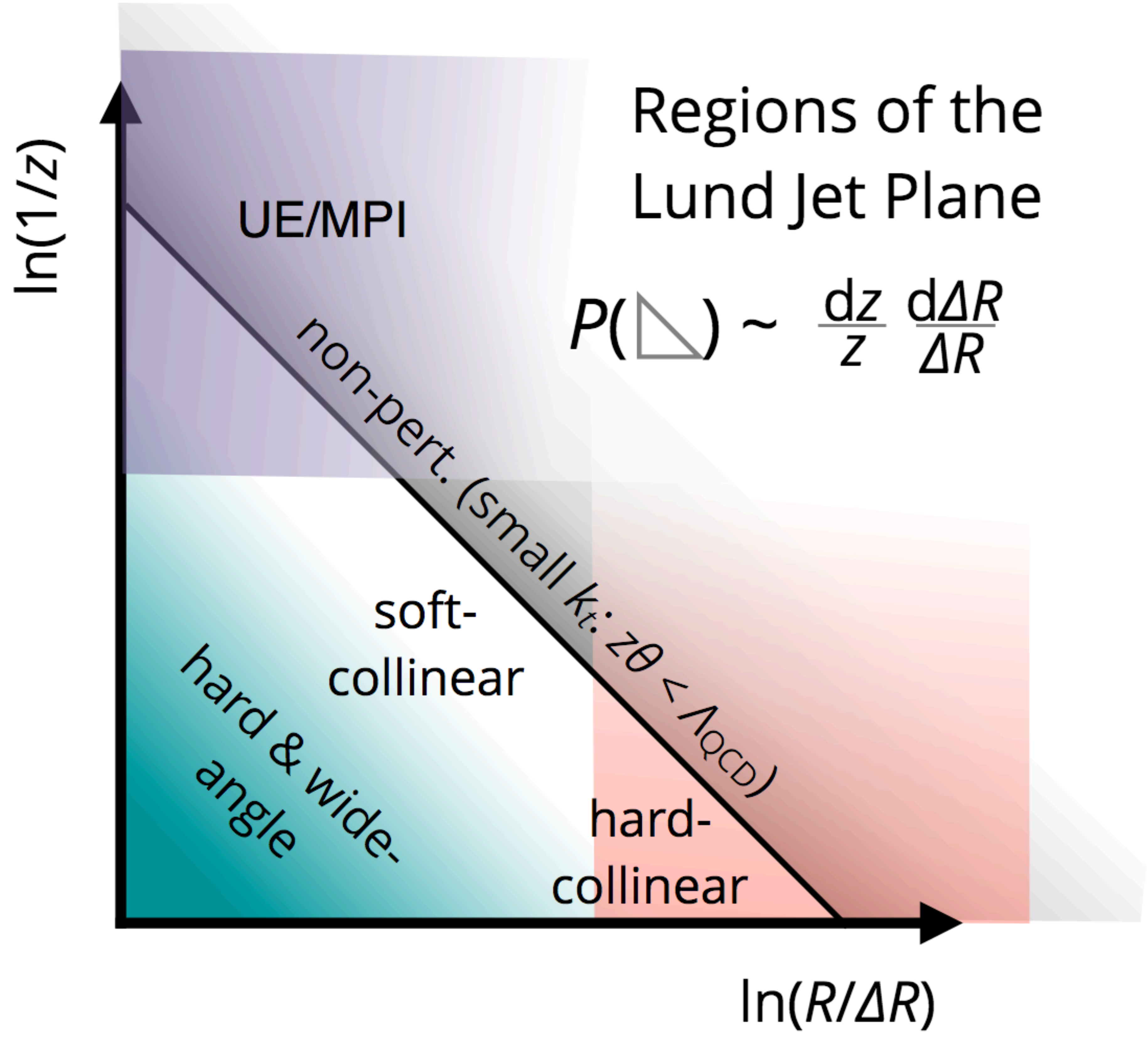


The Lund Plane is the phase space of these emissions: it naturally factorises perturbative and non-perturbative effects, UE/MPI, etc.

the lund plane

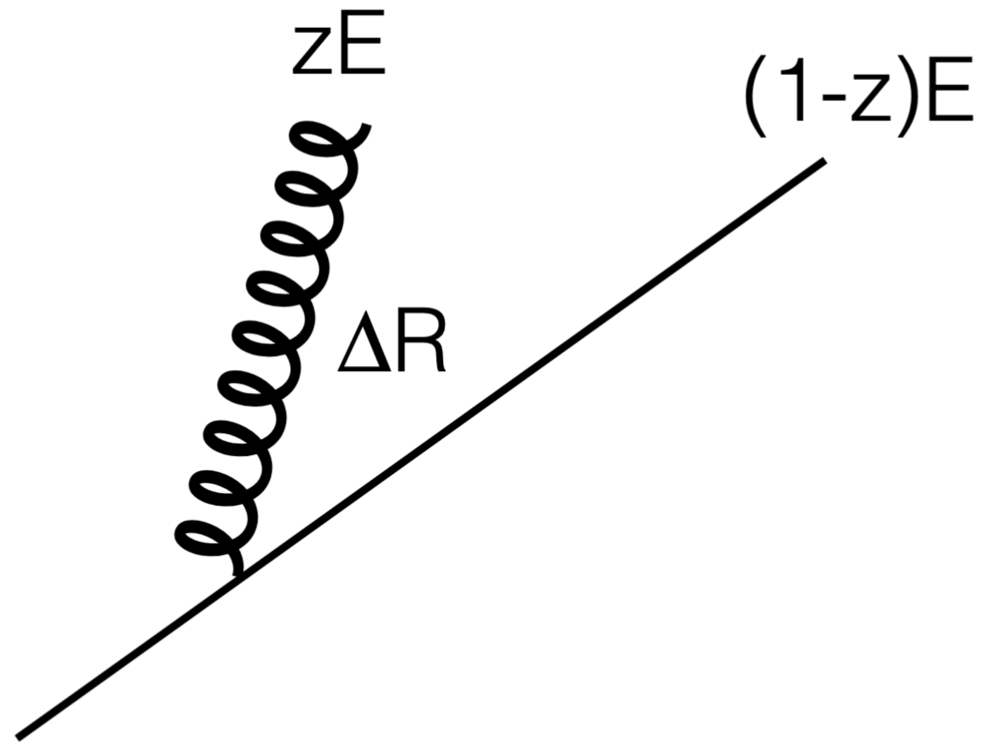


- A jet may be approximated as soft emissions around a hard core which represents the originating quark or gluon
- Emissions may be characterized by
  - z = relative momentum of emission relative to the jet core
  - ΔR = angle of emission relative to the jet core

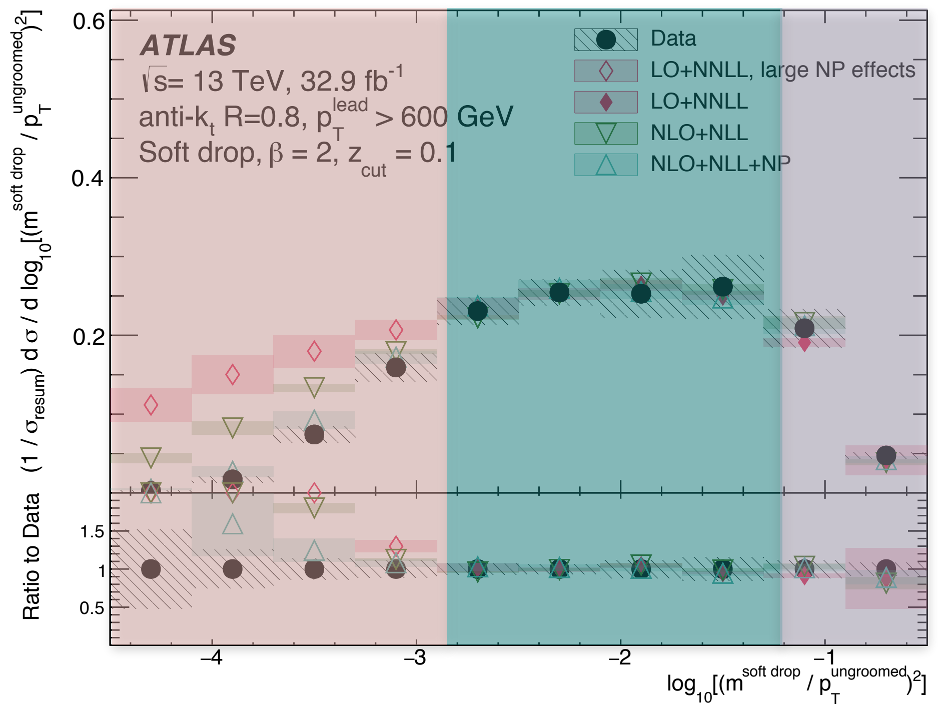


the lund plane

- A jet may be approximated as soft emissions around a hard core which represents the originating quark or gluon

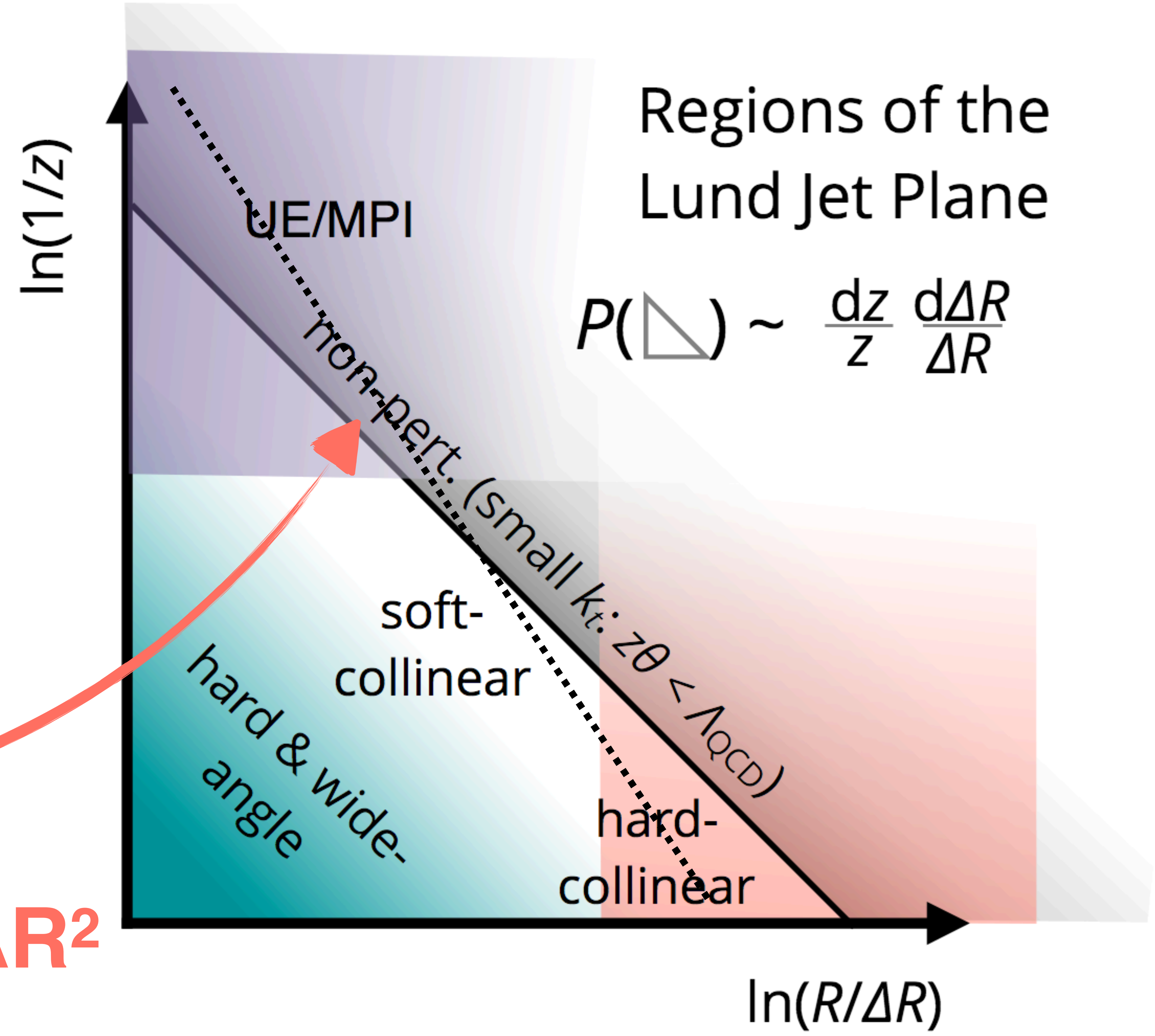


- Emissions may be characterized by
  - z = relative momentum of emission relative to the jet core
  - ΔR = angle of emission relative to the jet core



The jet mass is just one diagonal line in this space

$m \sim z^{*}\Delta R^2$



$P(\triangle) \sim \frac{dz}{z}\frac{d\Delta R}{\Delta R}$

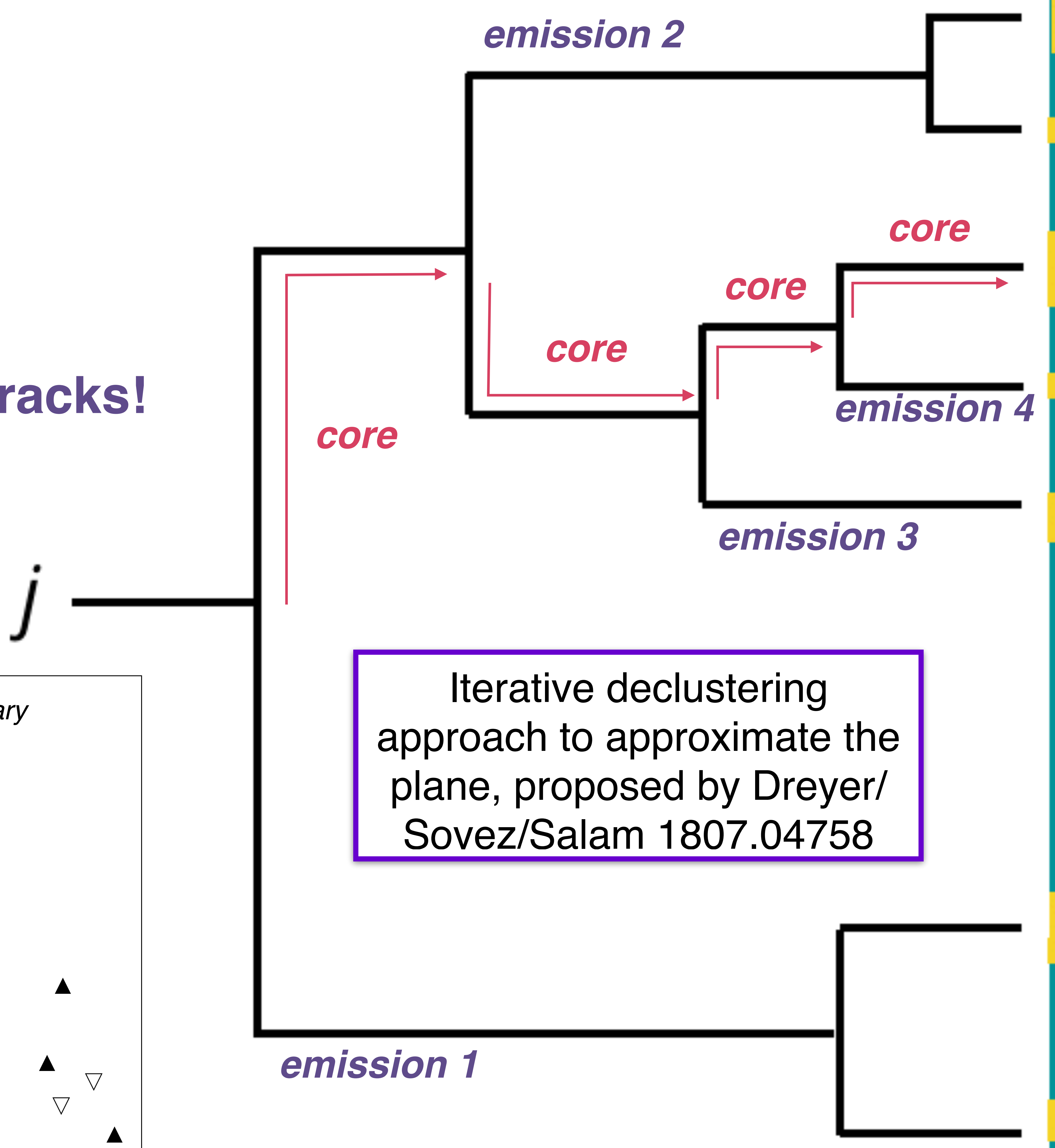
the lund jet plane

ATLAS-CONF-2019-035

1. Jet Finding:
Cluster jets using your favorite jet algorithm

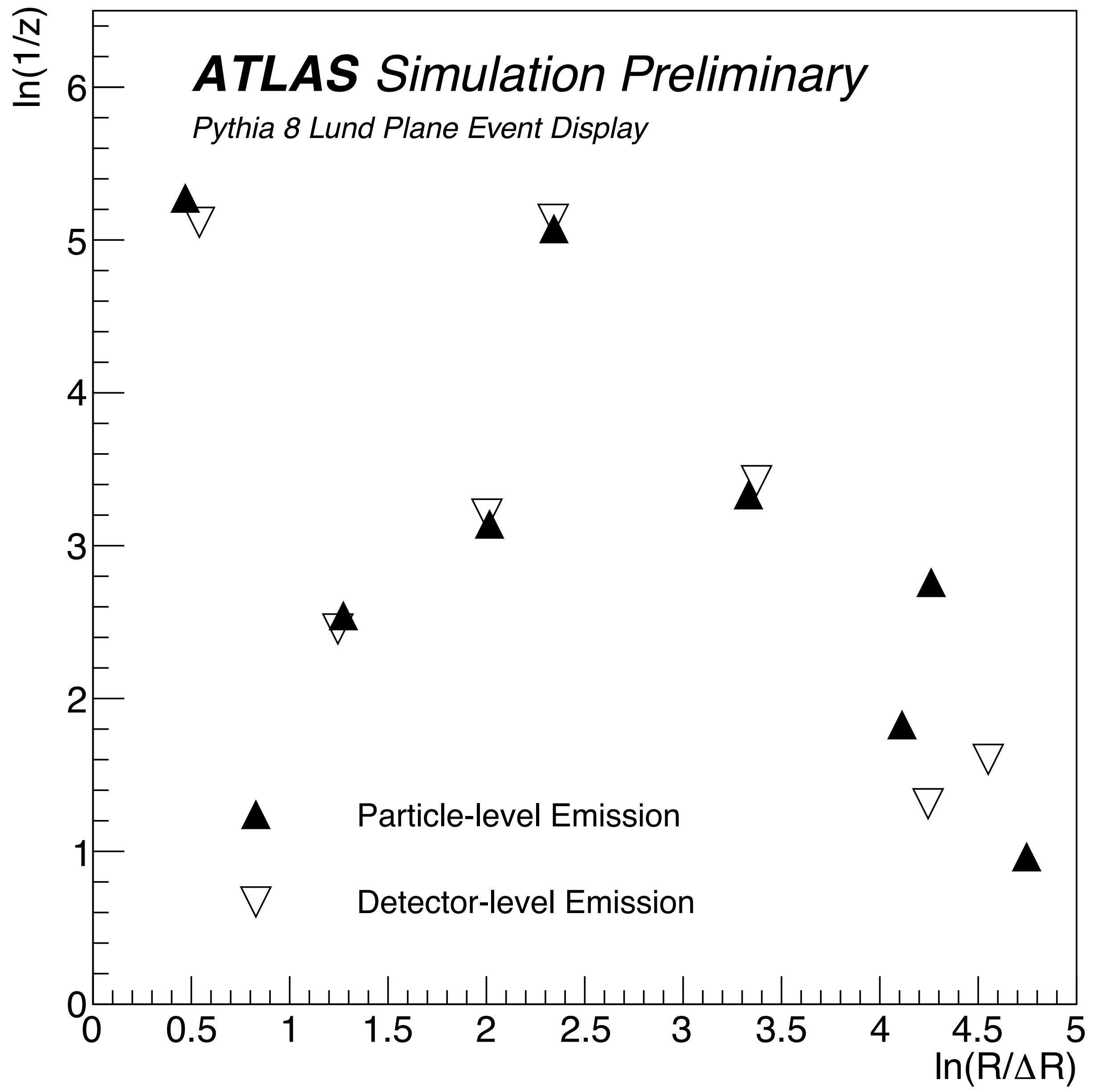
2. C/A Reclustering:
Combine closest pairs of **charged particles or tracks!**

3. C/A Declustering:
Unwind, widest angles first. Each step is an **emission**, or, a point in the Lund Jet Plane!



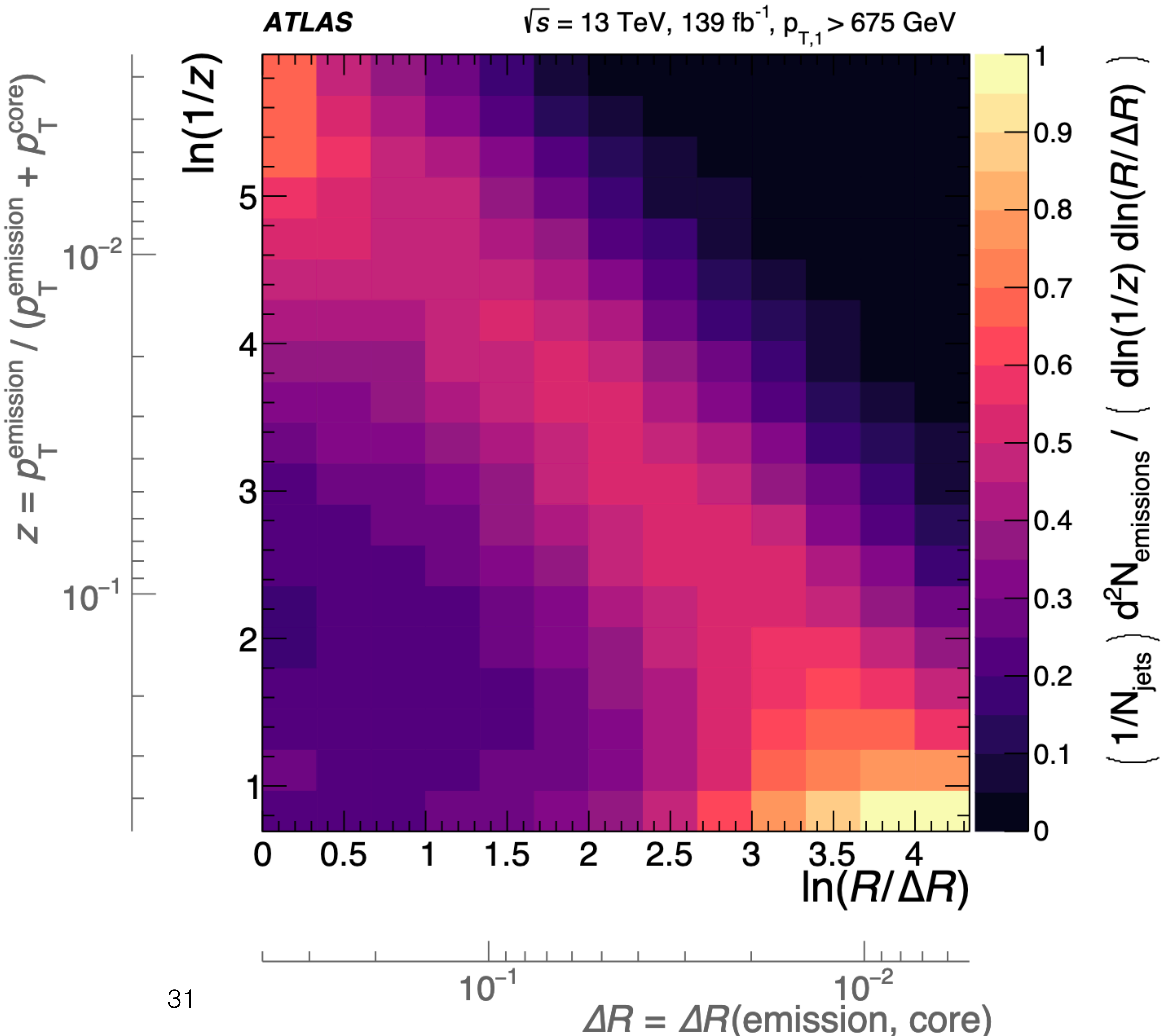
Iterative declustering approach to approximate the plane, proposed by Dreyer/Sovez/Salam 1807.04758

4. Plot Emissions:
Characterize emissions based on their angle (ΔR), and the hardness of the splitting and $z = p_T^{emission} / p_T$



the lund jet plane

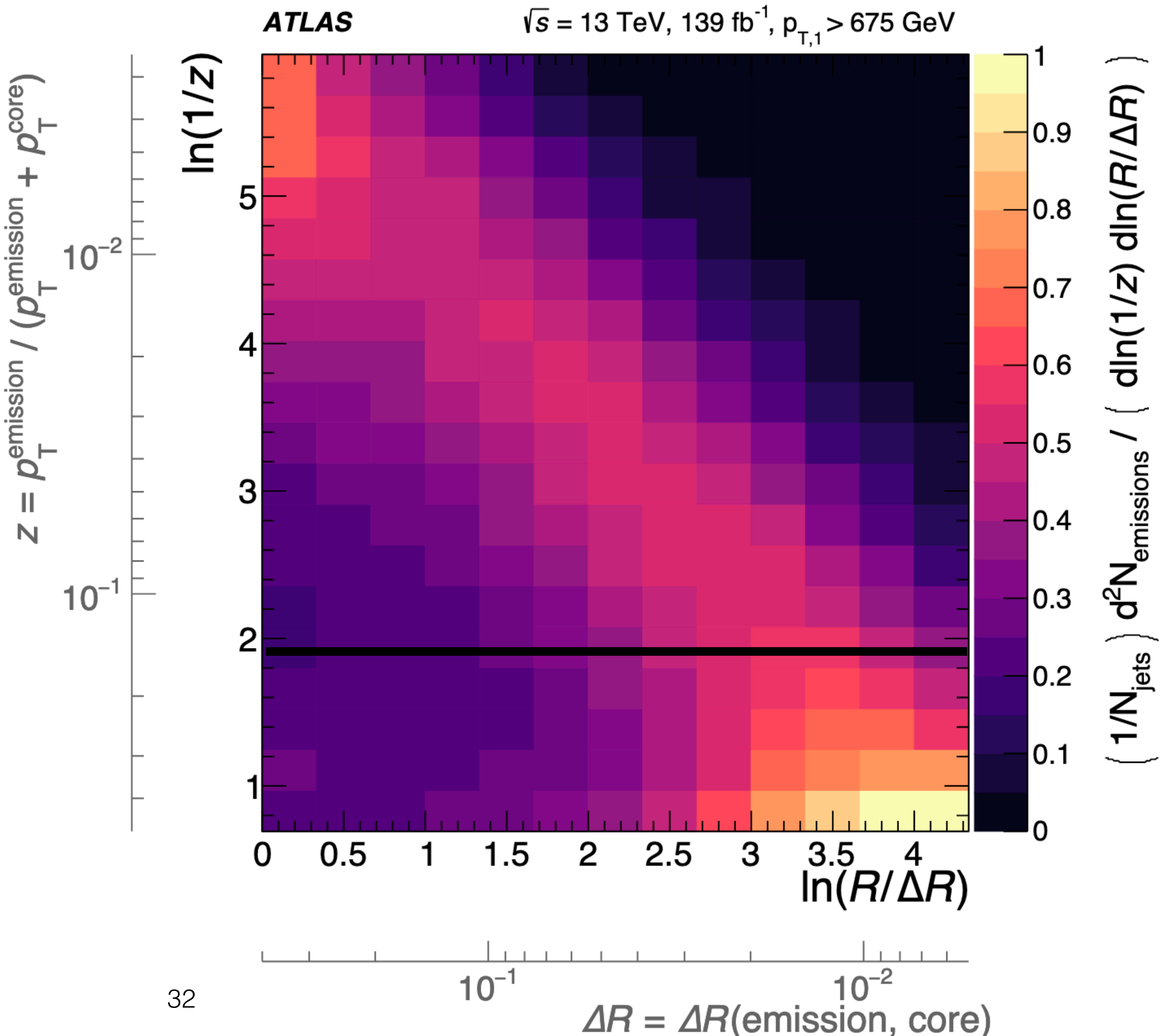
- Unfolded the primary Lund plane in dijet events
- Use tracks associated to the jets in order to have precise measurements for small splittings
    - Unfolded to charged particle level



2004.03540

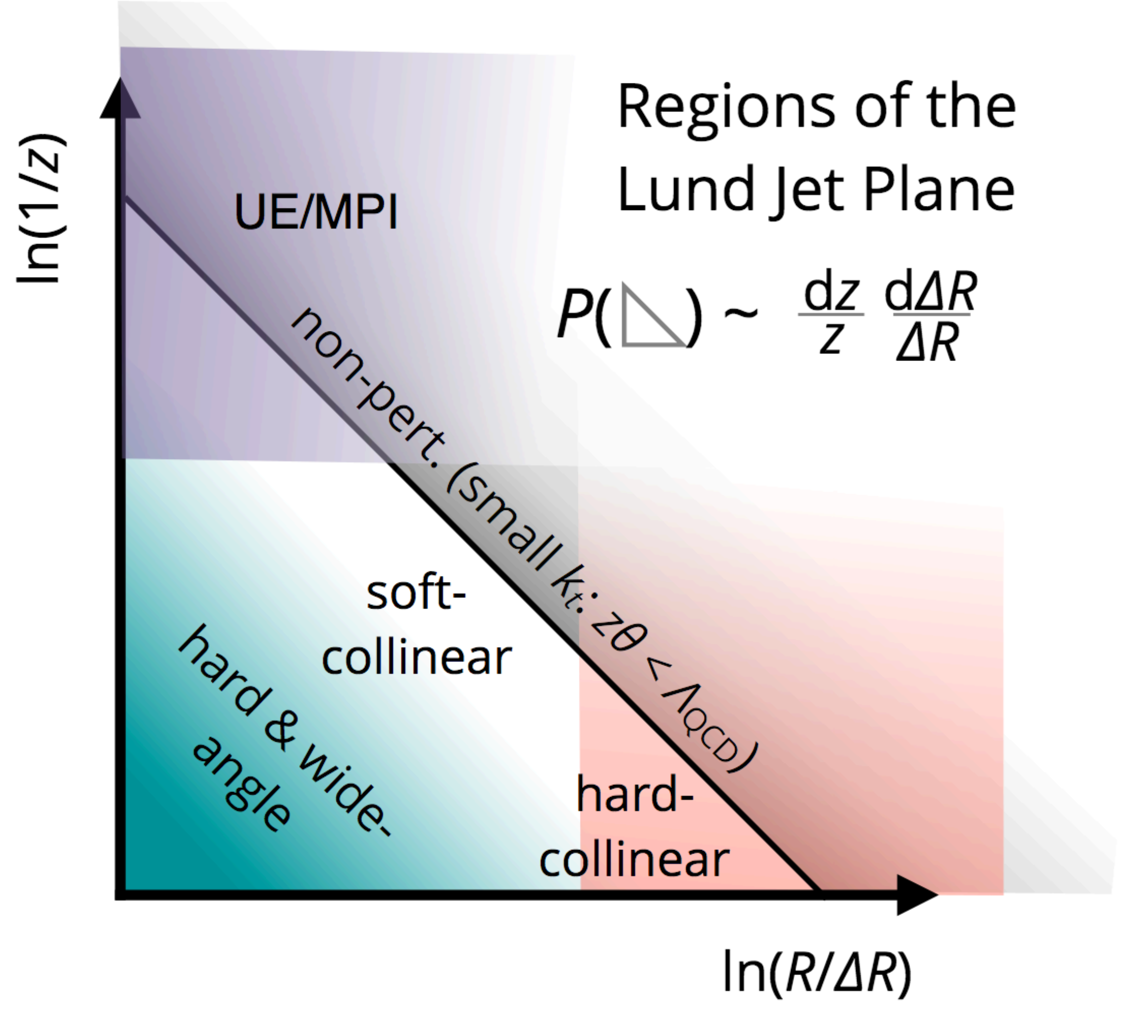
the lund jet plane

- Unfolded the primary Lund plane in dijet events
- Use tracks associated to the jets in order to have precise measurements for small splittings
  - Unfolded to charged particle level

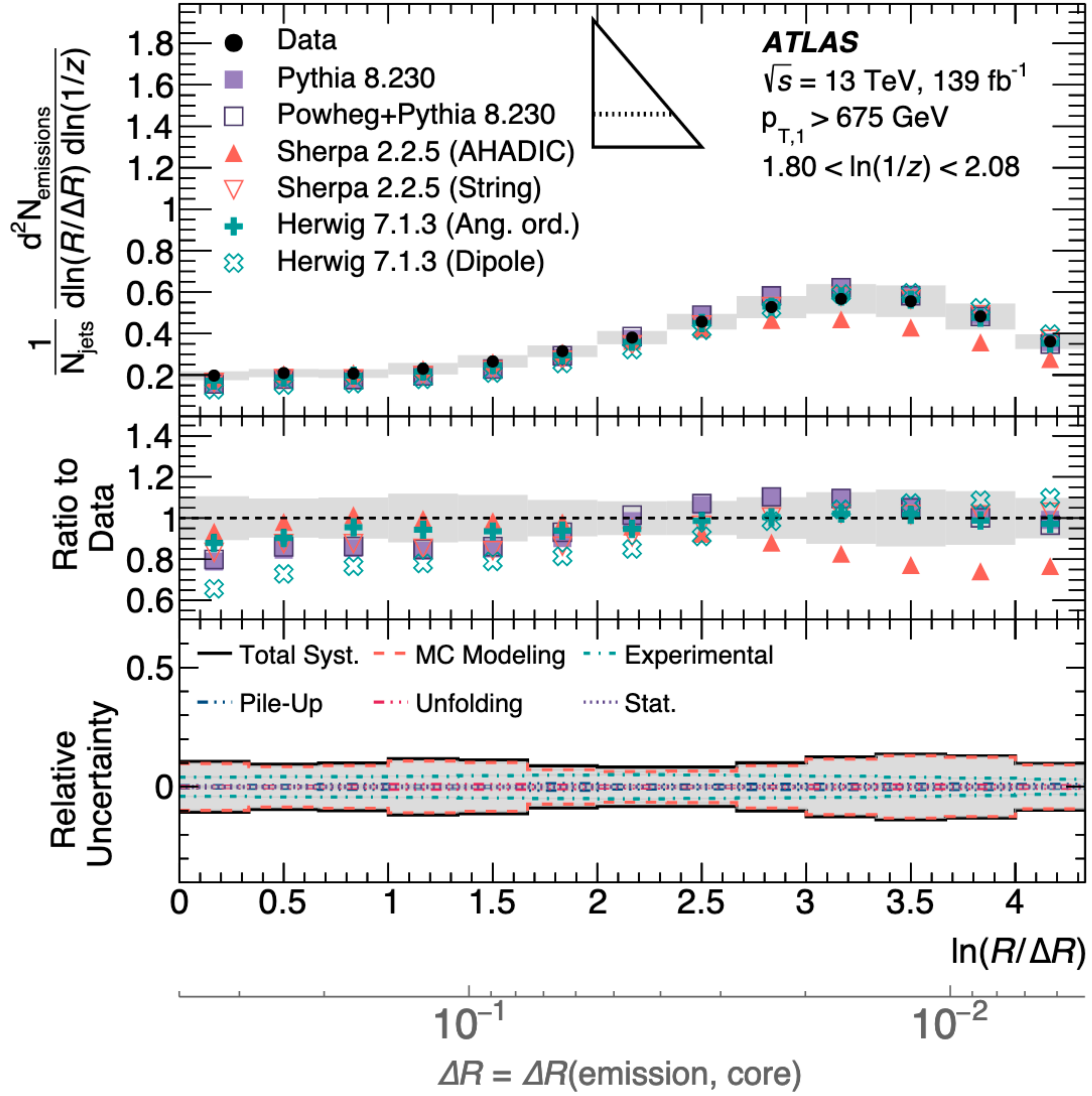


the lund jet plane

2004.03540

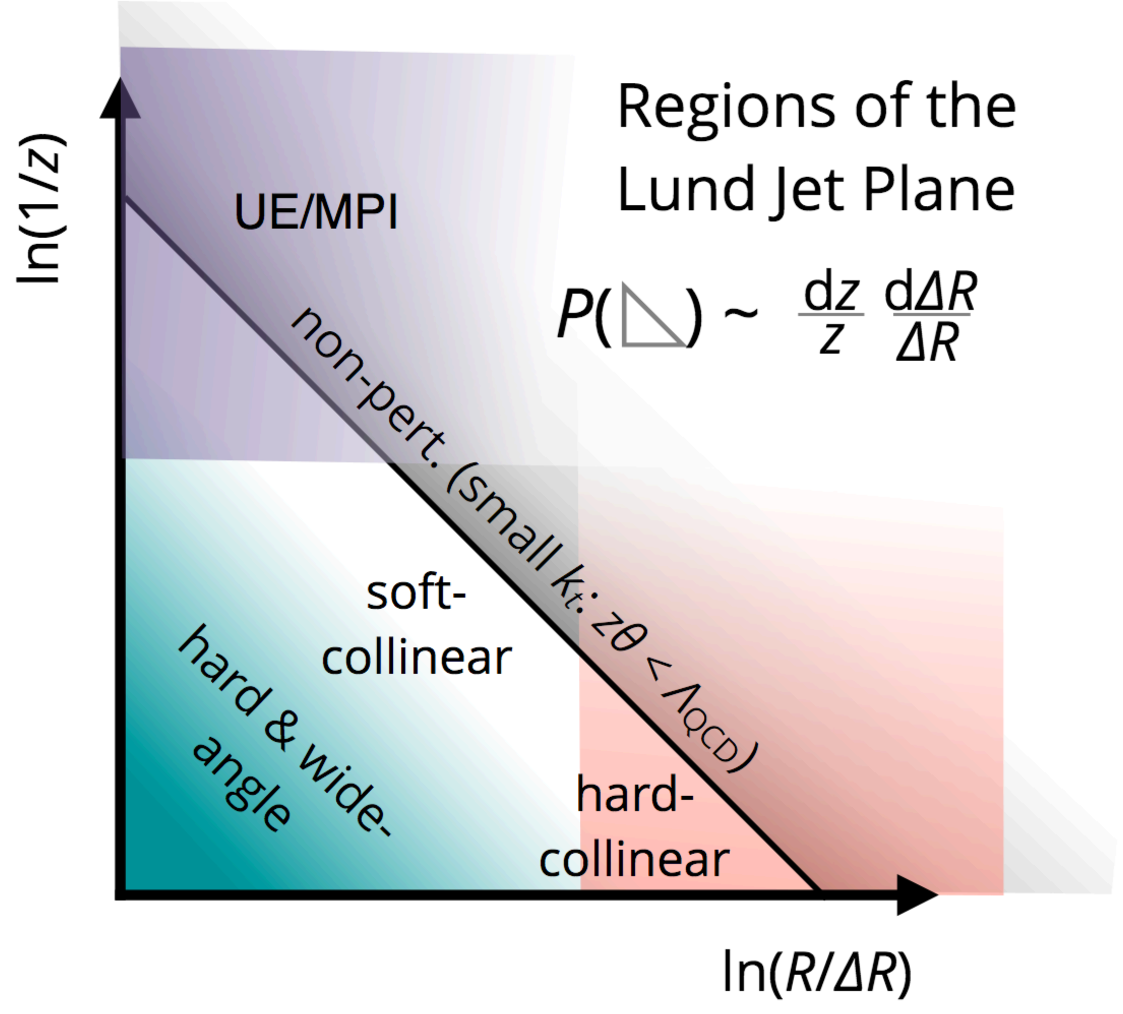


- Non-trivial differences between different generators and unfolded data
- Region dominated by hard and wide-angle splitting is affected by parton shower
- Hadronization effects in region with non-perturbative effects

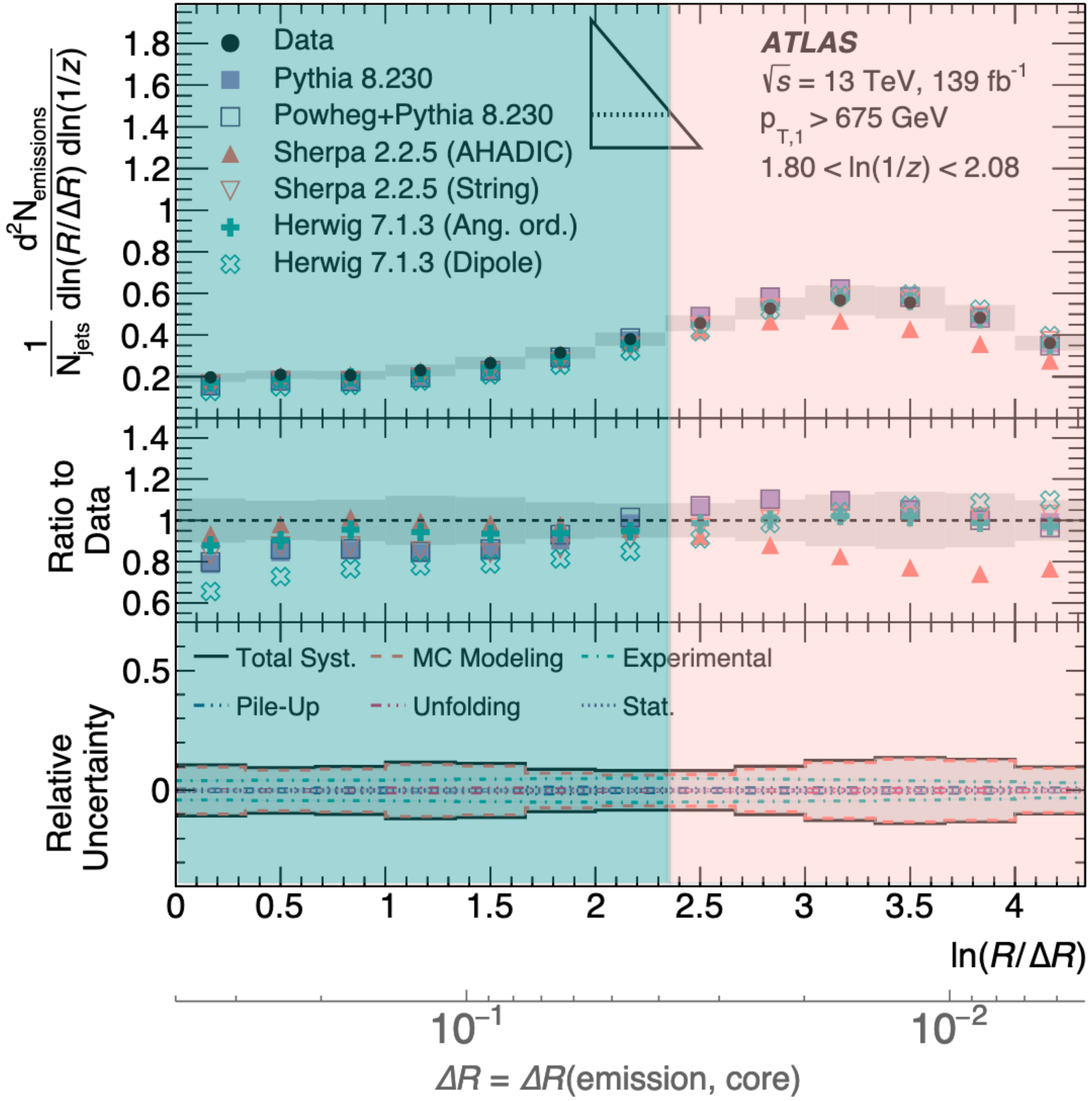


2004.03540

the lund jet plane



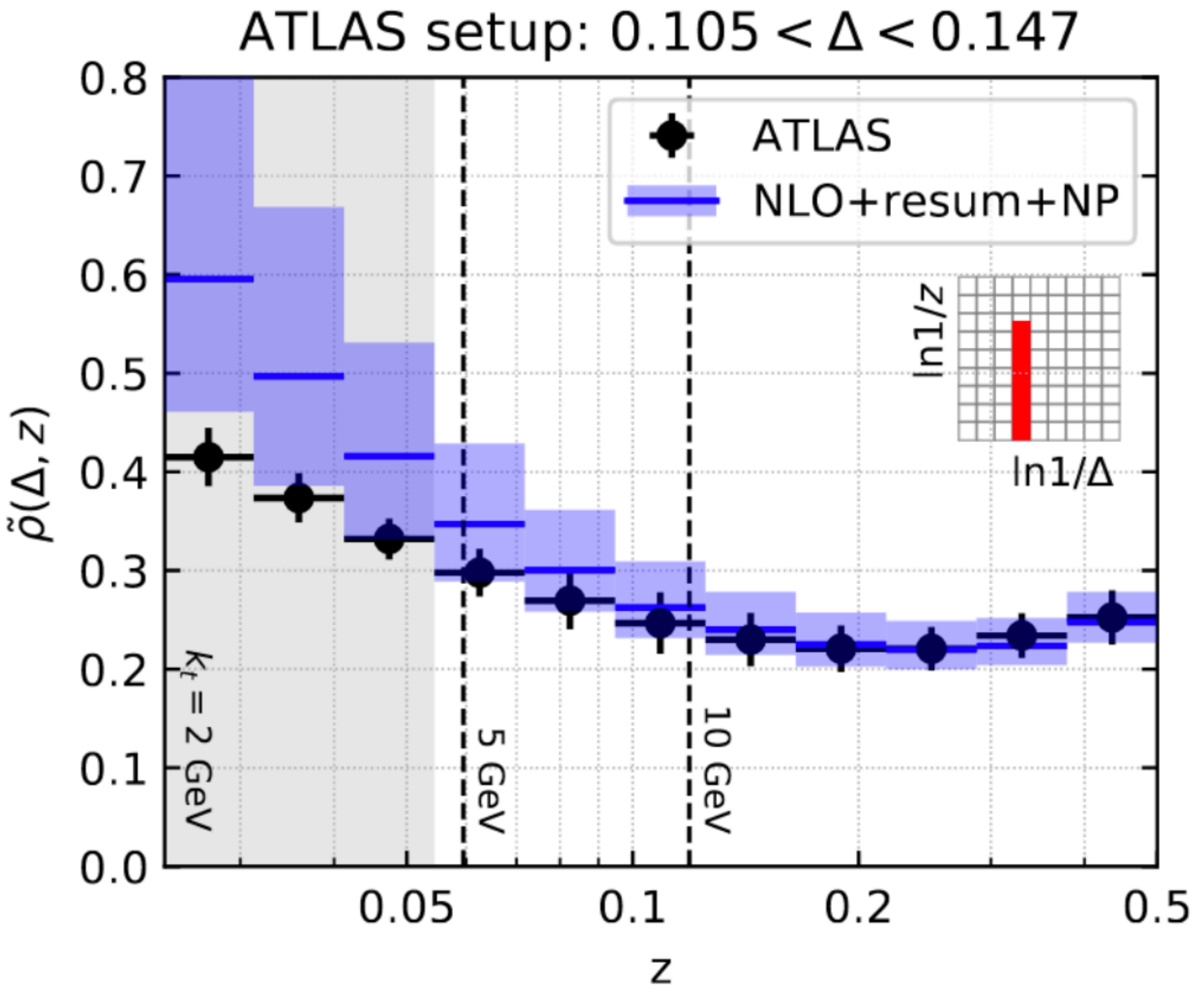
- Non-trivial differences between different generators and unfolded data
- Region dominated by hard and wide-angle splitting is affected by parton shower
- Hadronization effects in region with non-perturbative effects



the lund jet plane

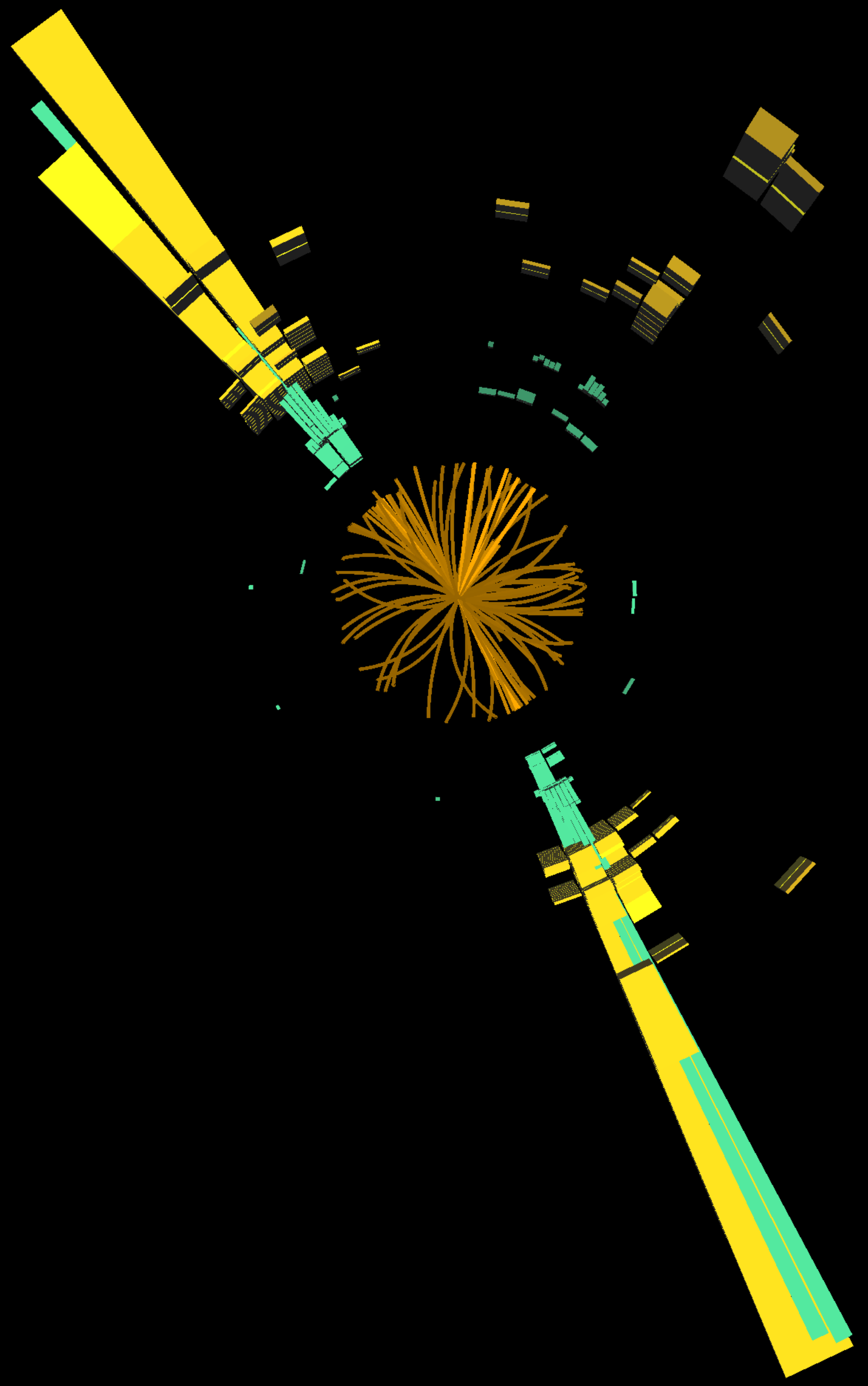
2007.06578

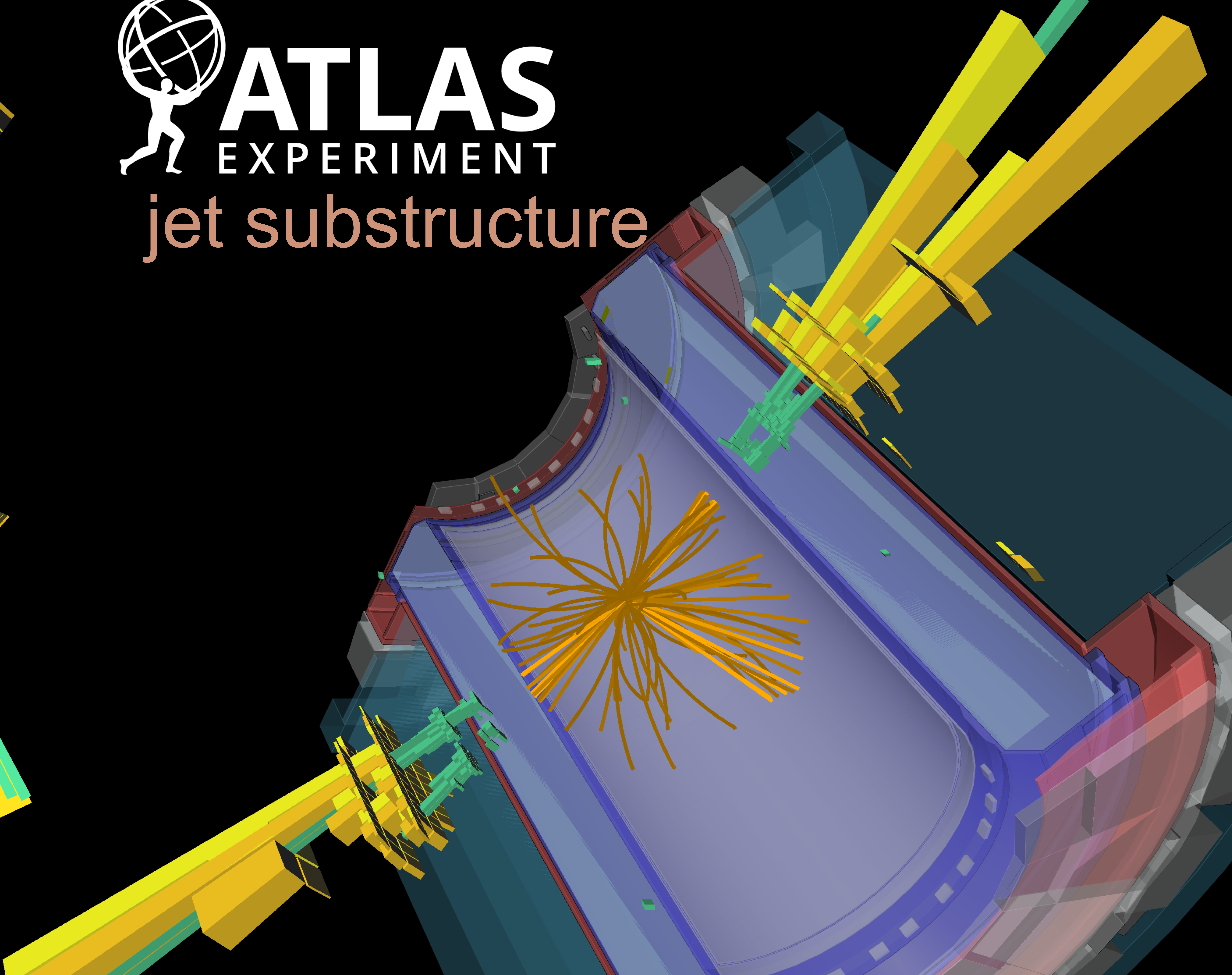
- Possible to produce predictions for parts of the LJP
  - All-order calculation is accurate down to k_t of ~5 GeV
- Example of substructure prediction without grooming algorithm!
  - Work ongoing to extend this to higher logarithmic accuracy



ATLAS EXPERIMENT

jet substructure

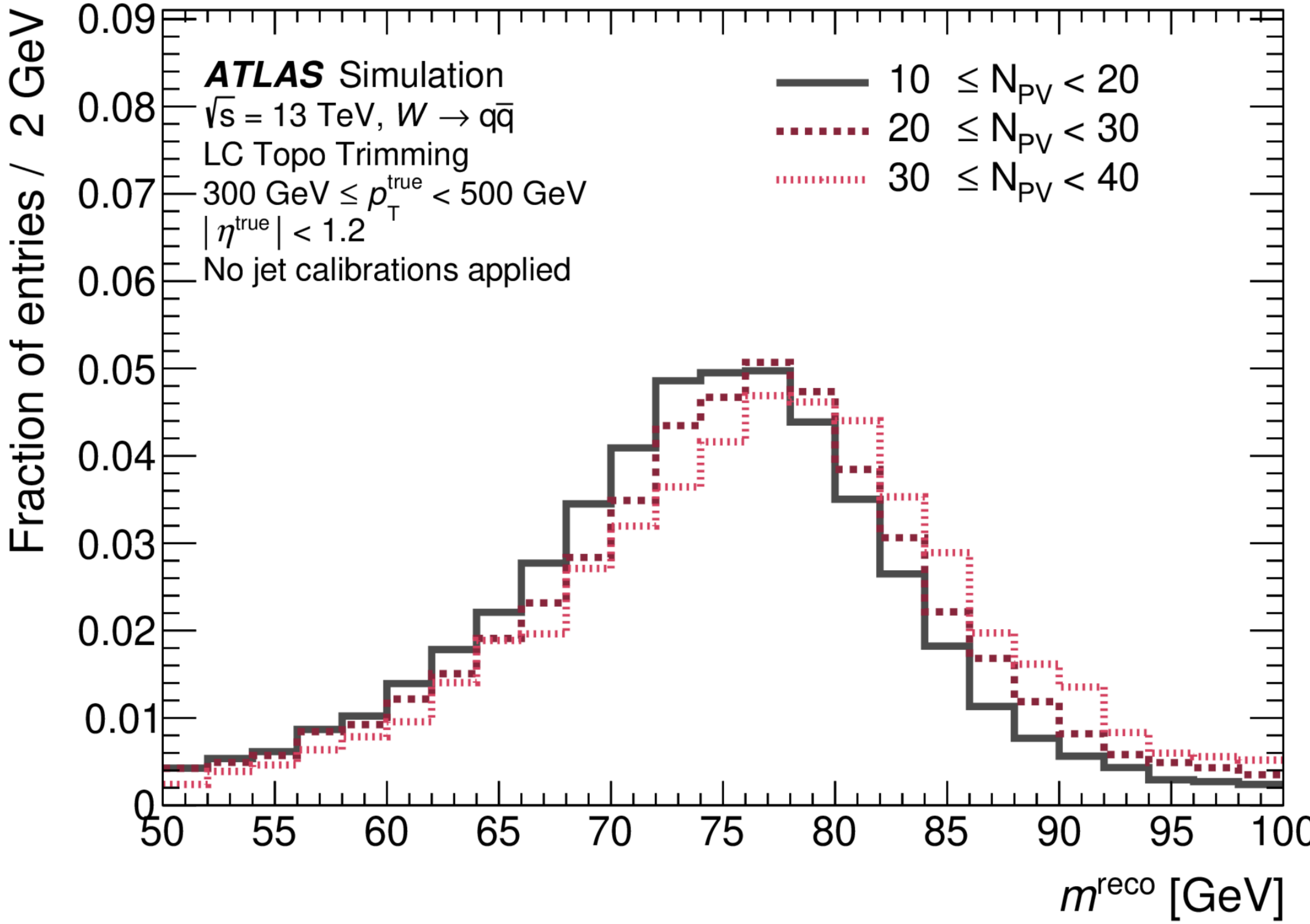




pileup mitigation

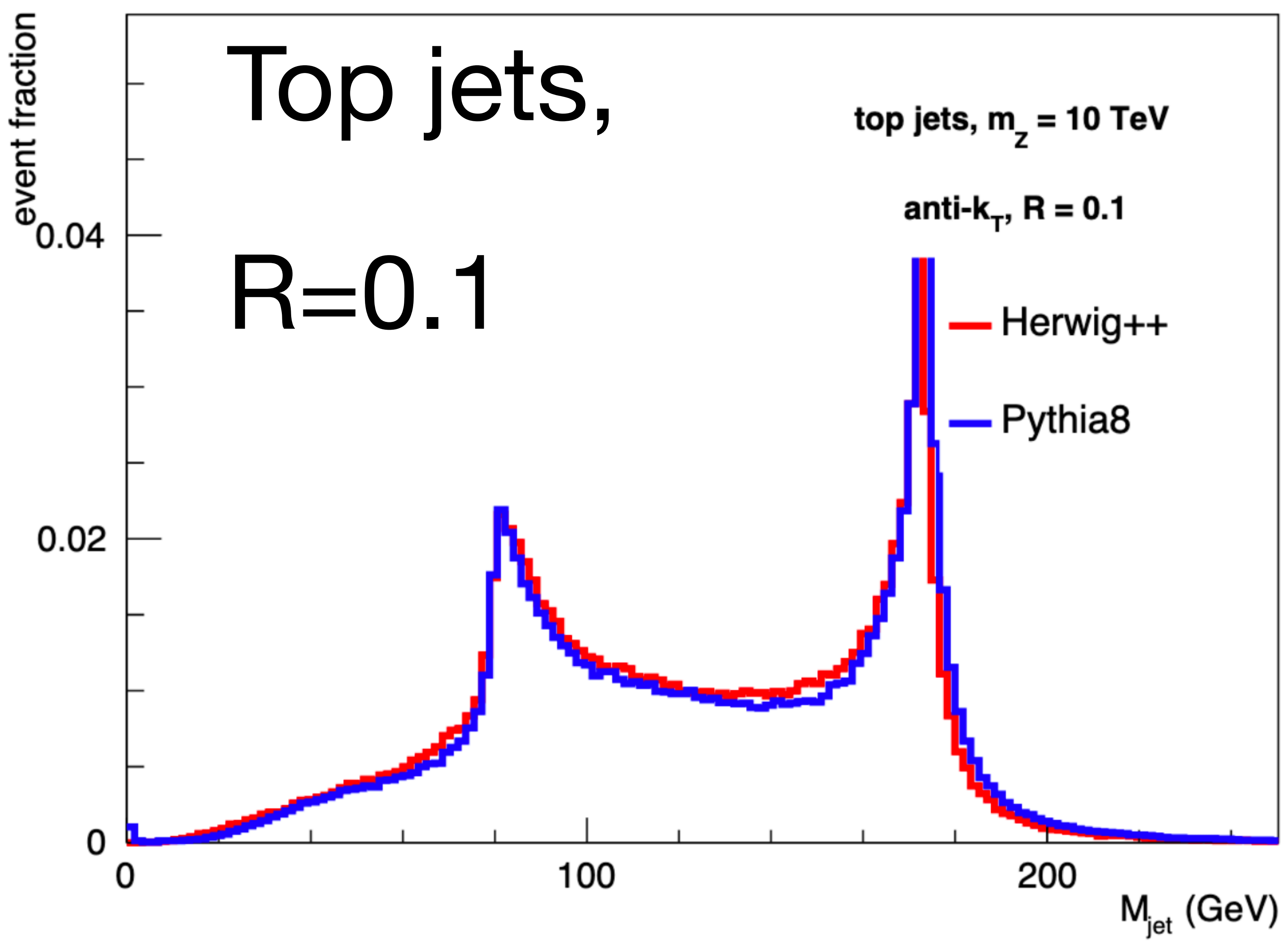
JETM-2018-06

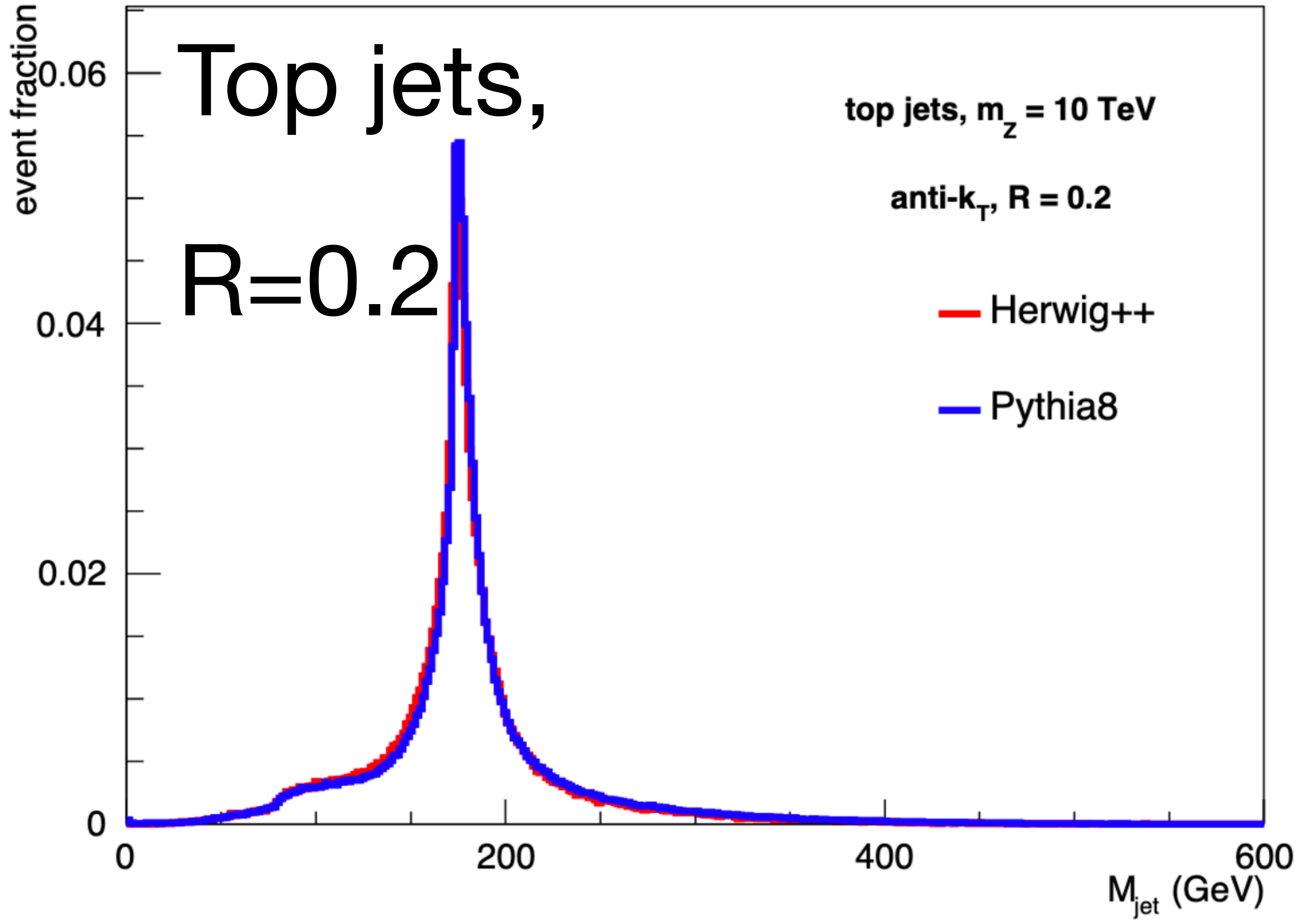
- Jet substructure observables are sensitive to the soft radiation in the jet
- Need pileup mitigation to reduce these effects
  - Good object and detector design helps to minimize these effects
  - Techniques like Constituent Subtraction or SoftKiller also lessen the effects of pileup
- Future colliders will require that we have a better understanding of how to mitigate pileup effects
  - FCC-hh could have up to 1000 interactions per bunch crossing!
  - Muon colliders have beam-induced backgrounds, which will result in similar issues

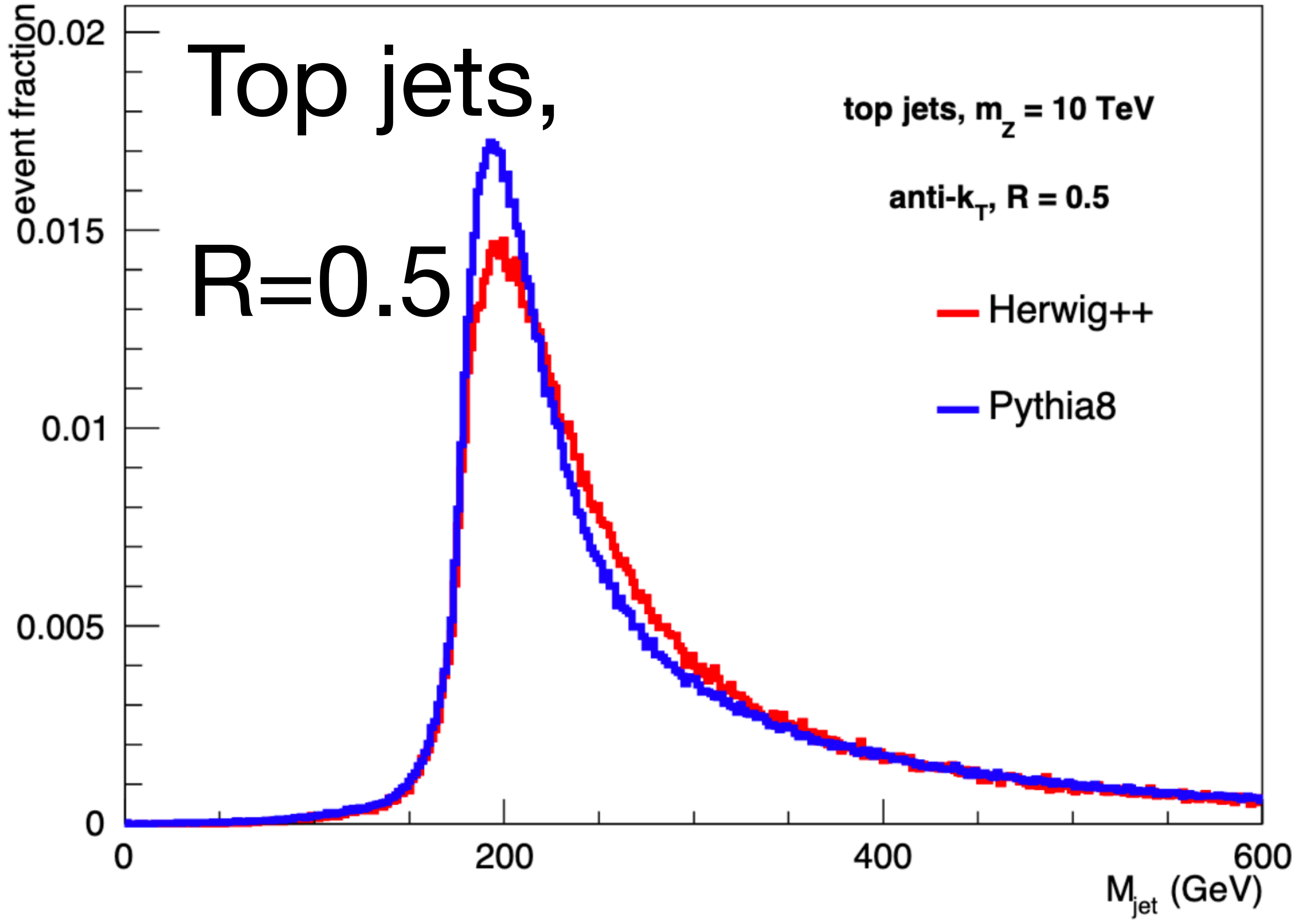


future colliders

- Need excellent boosted object reconstruction to be able to tag W/Z/top jets
- Boosted jets at FCC/SPPC energies will look significantly different than at the LHC
  - Containment will happen for much smaller jet radii than at the LHC
  - The decay products of W/Z decays could conceivably be collimated within a single calorimeter cell!
  - Need more studies to understand the full implications of this for future colliders



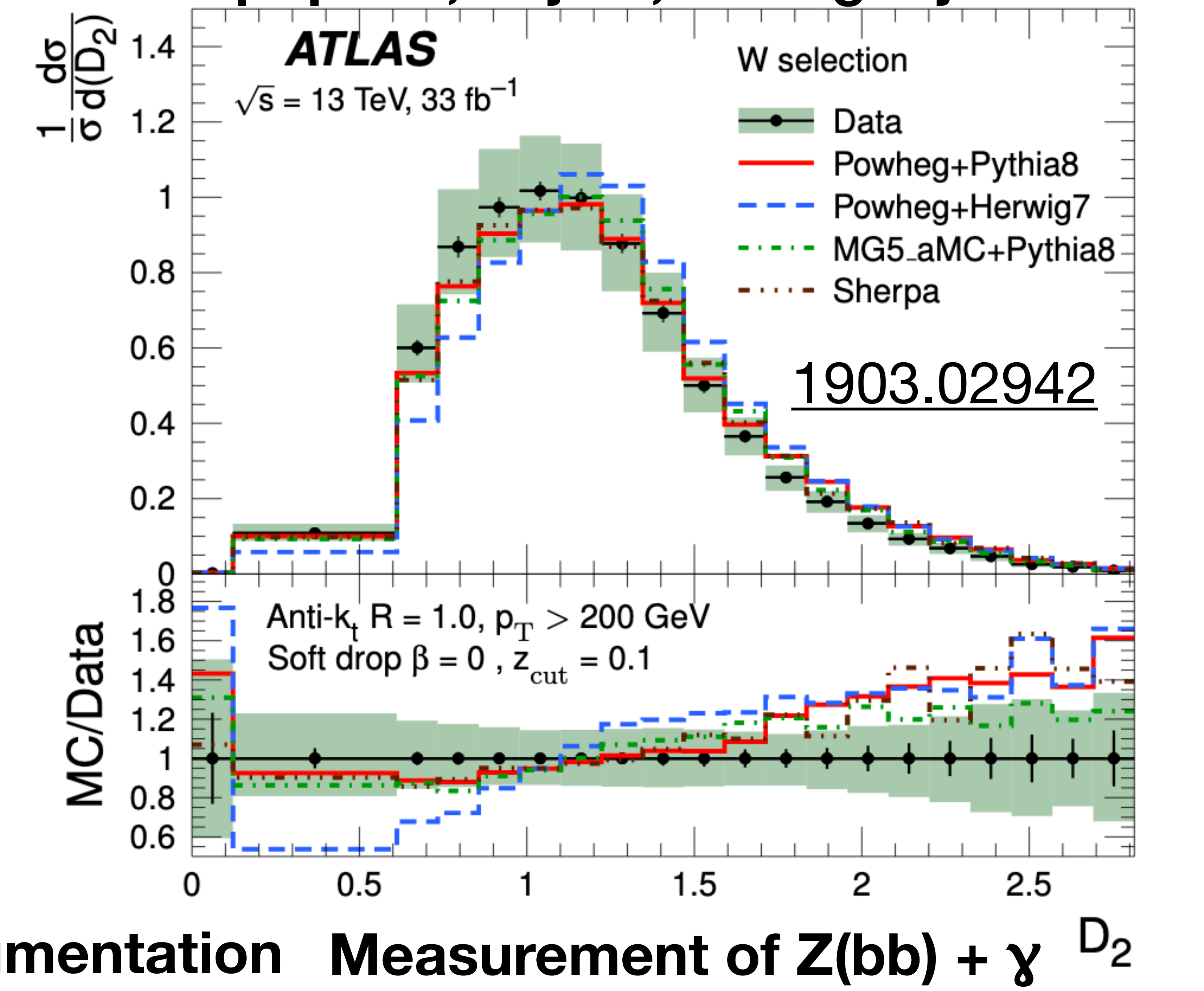




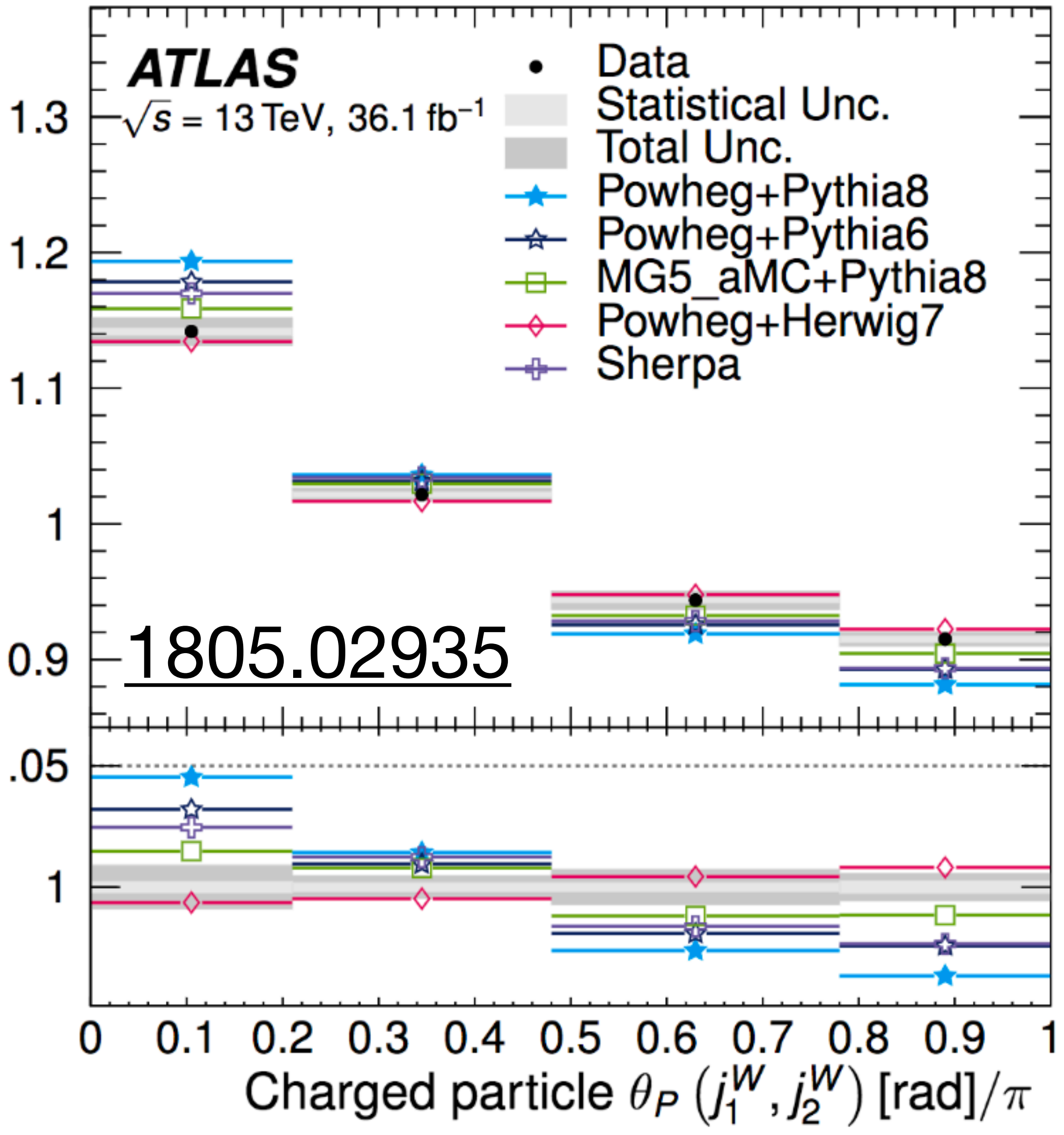
other substructure measurements

- Covering the measurements most relevant for pQCD, but many interesting measurements of jet substructure from ATLAS
- Including links here for anyone interested in learning more, and more information in the backup

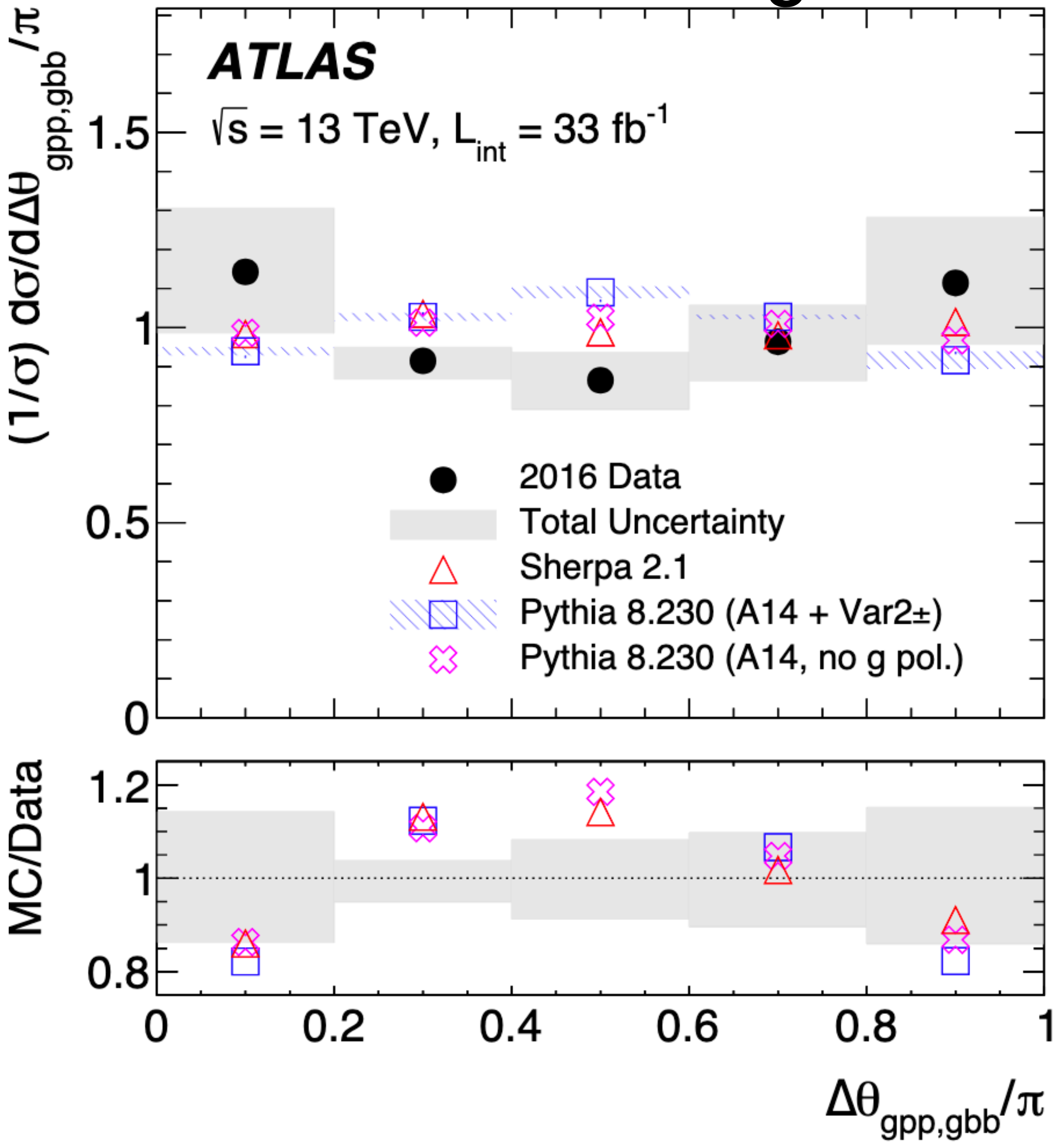
Measurement of jet substructure observables for top quark, W jets, and light jets



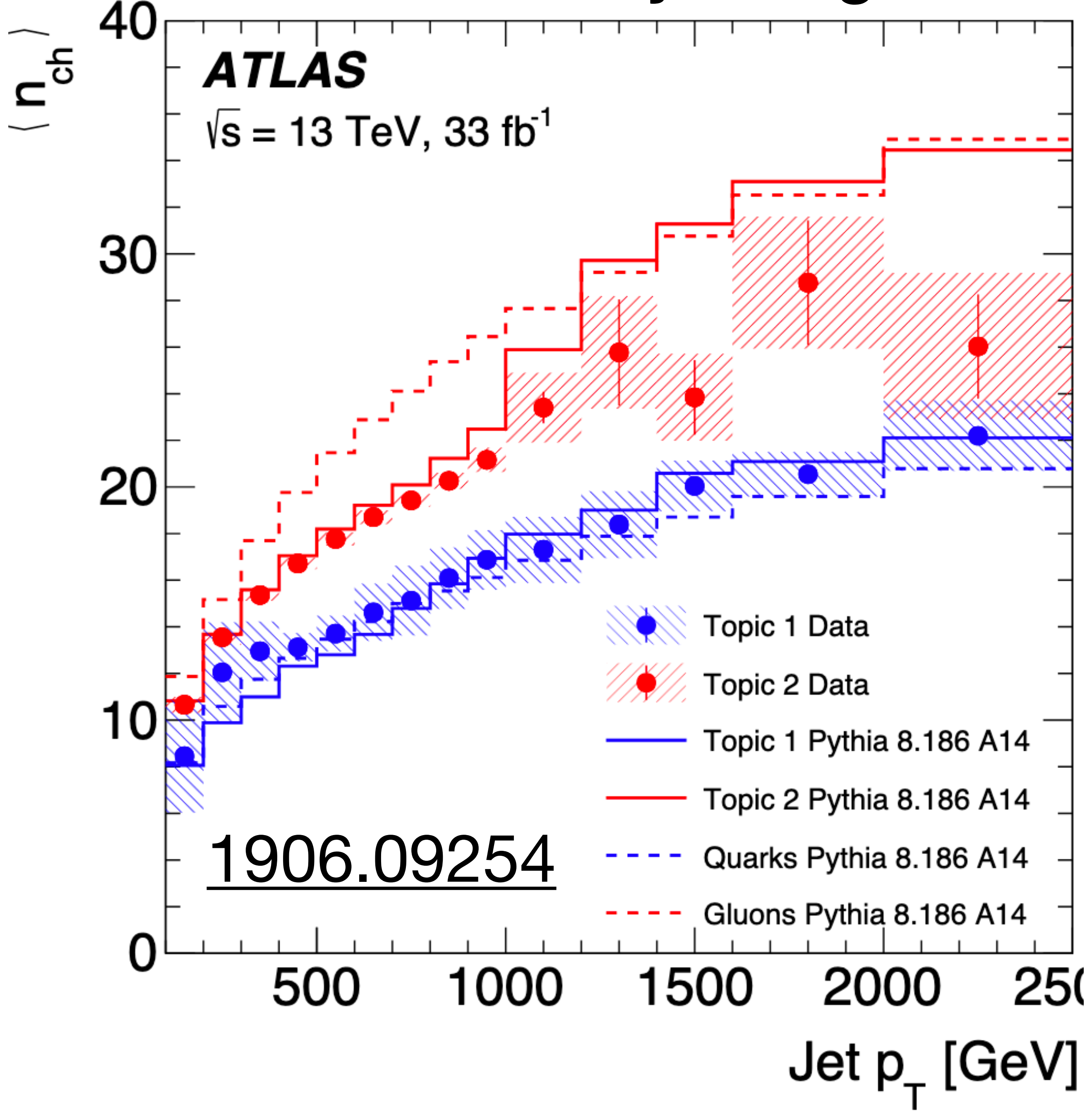
Measurement of the jet pull



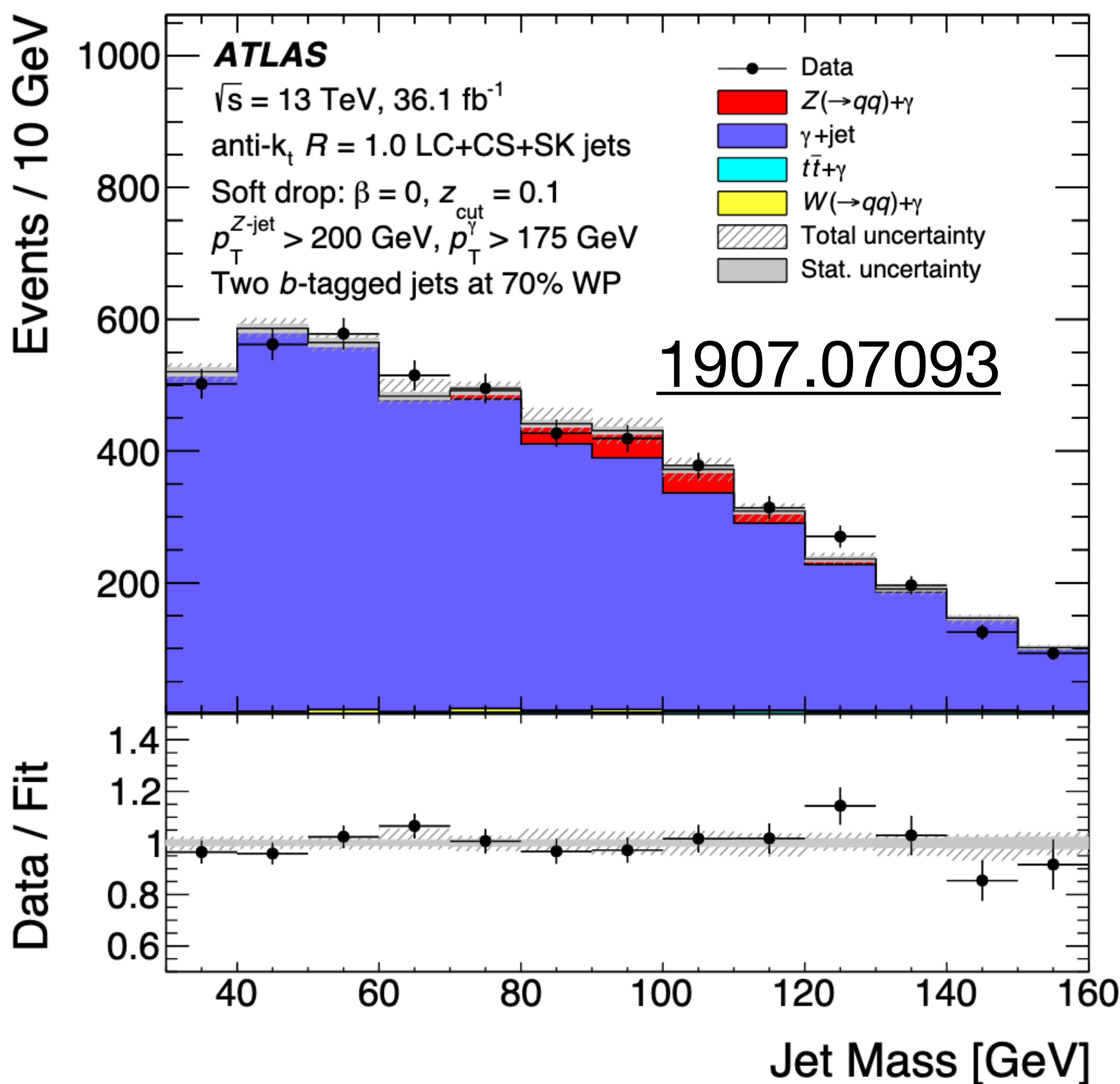
Measurement of g → bb



Measurement of jet fragmentation

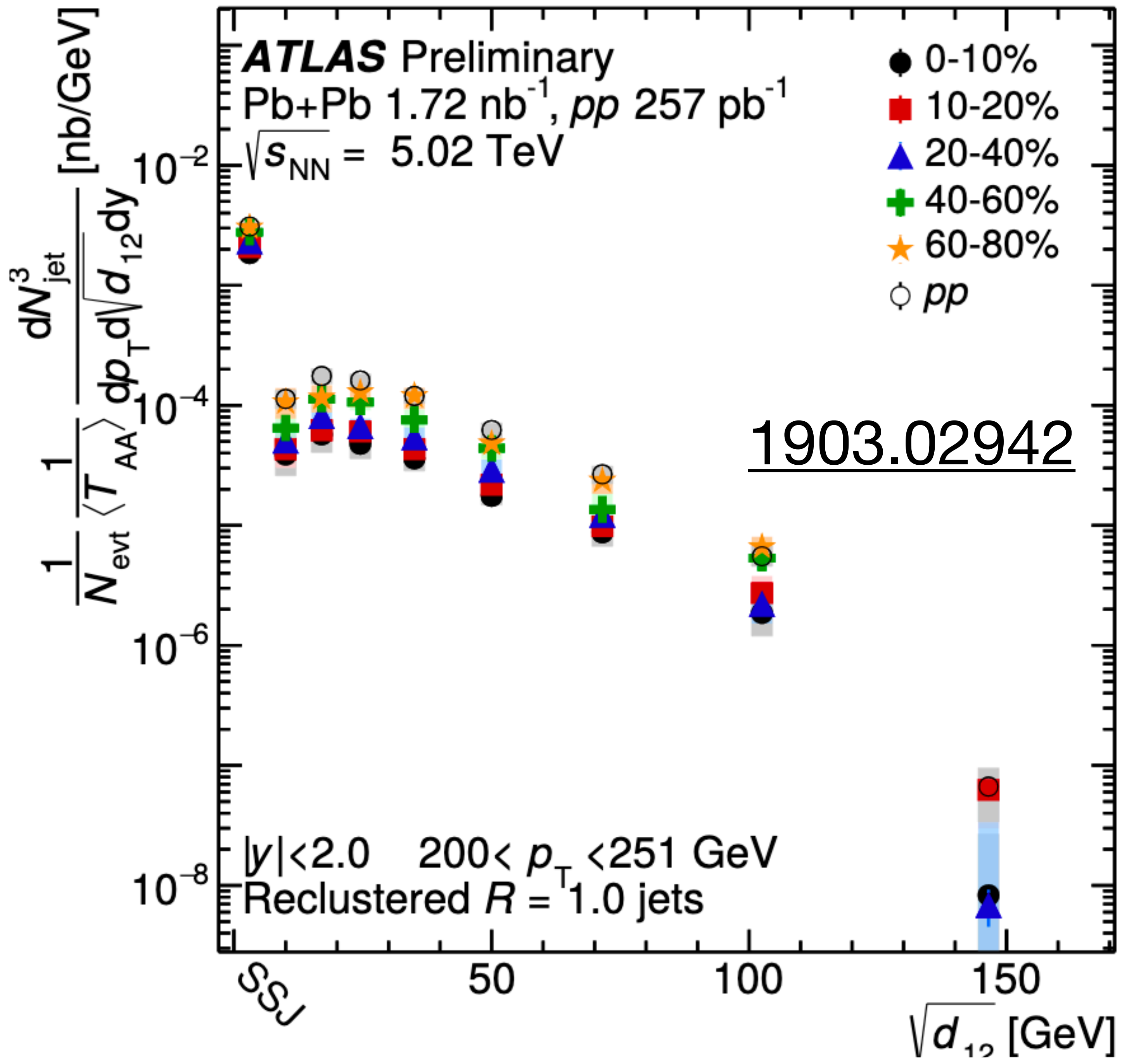


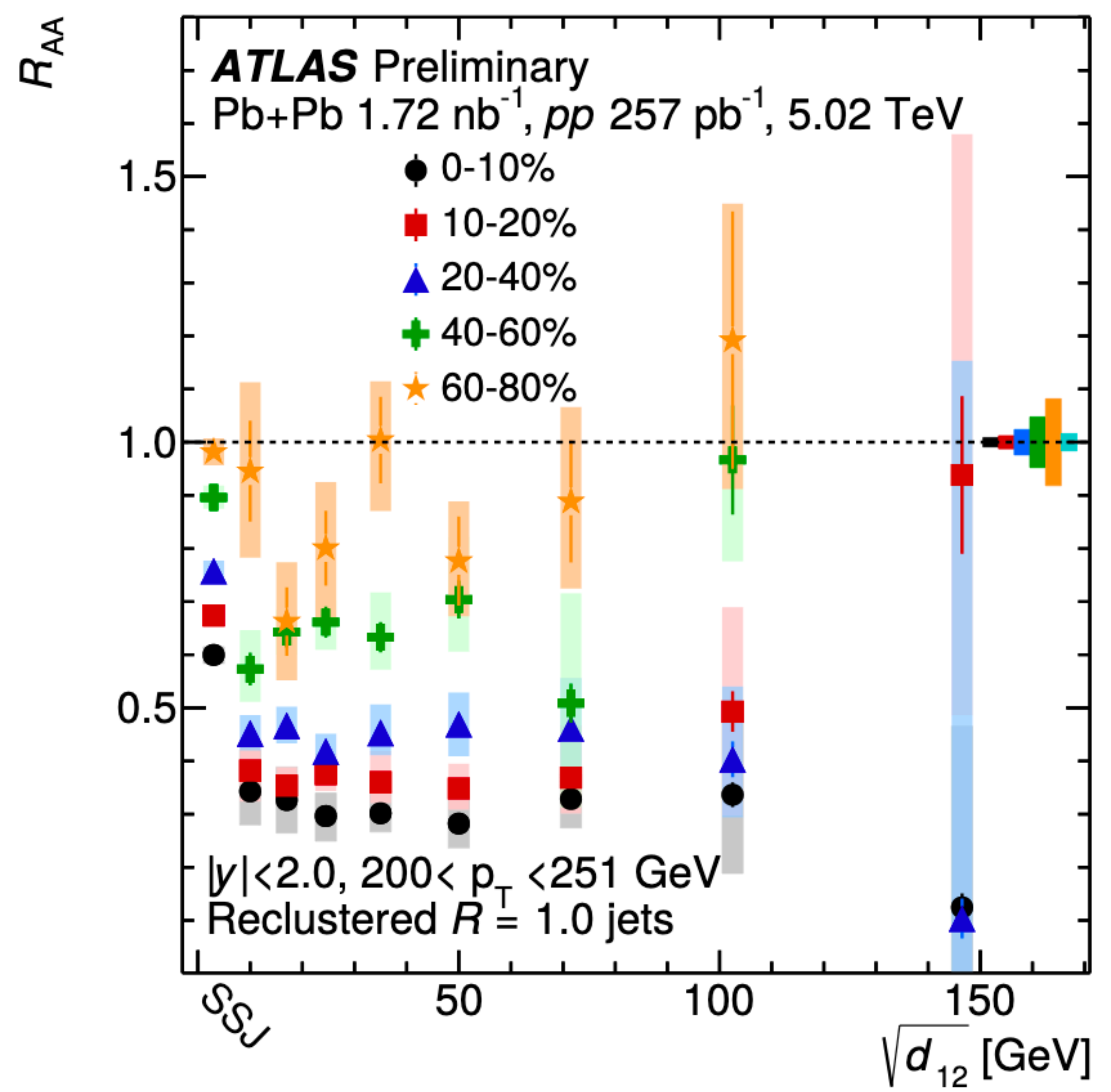
Measurement of Z(bb) + γ



other substructure measurements

- Lots of recent developments in heavy ions, which are beyond the scope of today's talk
- Studies of jet substructure help provide insight into jet quenching
  - Measurements are relatively recent, and lots of rapid development in this area





concluding thoughts

- ***Detector and experimental developments are crucial for improving precision of substructure measurements***
  - Advances in detector design (like timing detectors), and object reconstruction open new doors for more advanced substructure reconstruction
  - New colliders will make substructure reconstruction more challenging, with higher energies and increased pileup
    - Need to design detectors with substructure in mind in order to take full advantage of their capabilities!
- ***Tracking will continue being an important part of substructure measurements***
  - Provides simple and robust way of measuring substructure observables, at the cost of only measuring charged particles
  - Need more discussions between theorists and experimentalists on possibilities for using these for predictions
    - *See Ian Moult's talk for ideas of how this can be addressed in the future!*

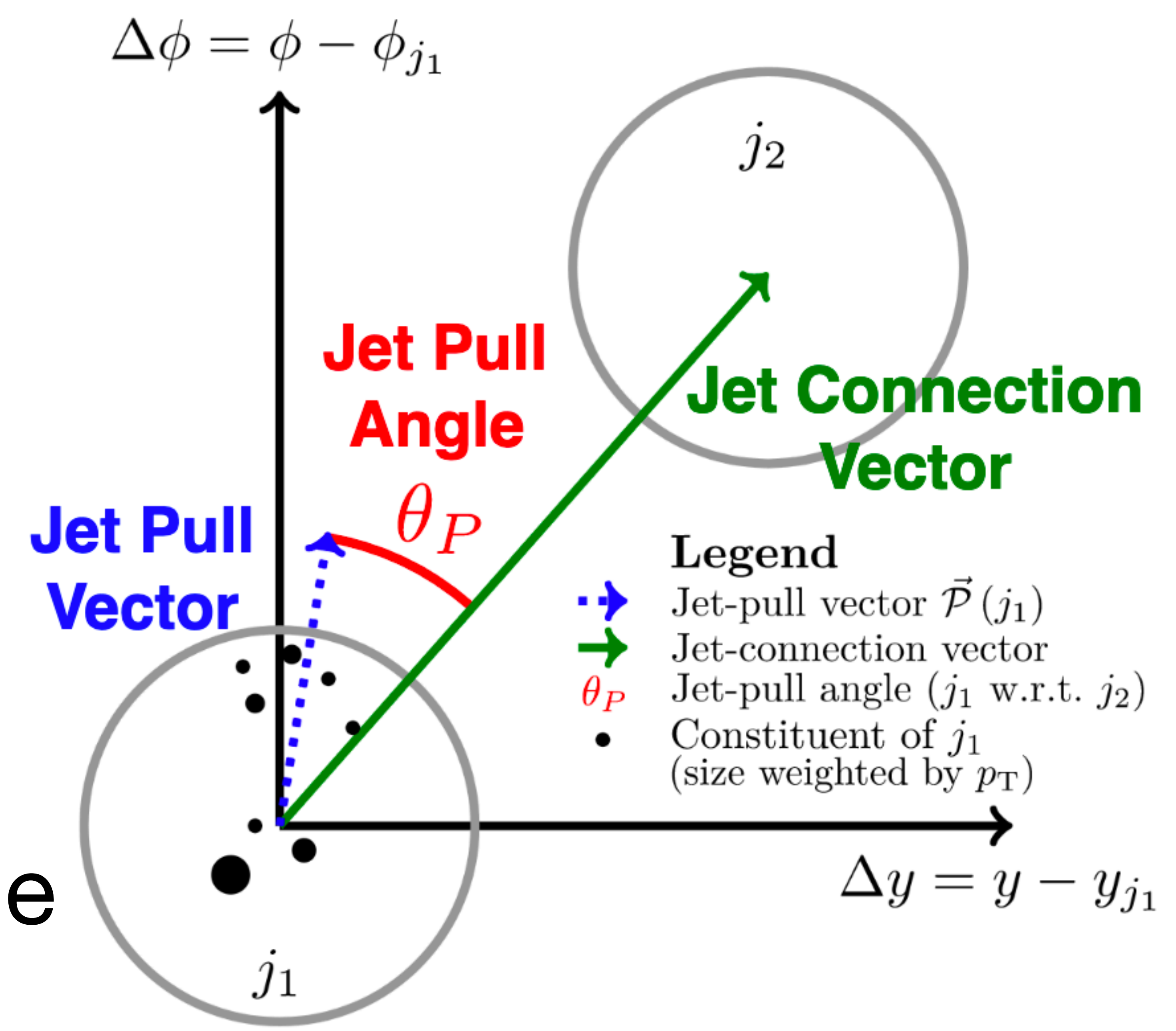
concluding thoughts

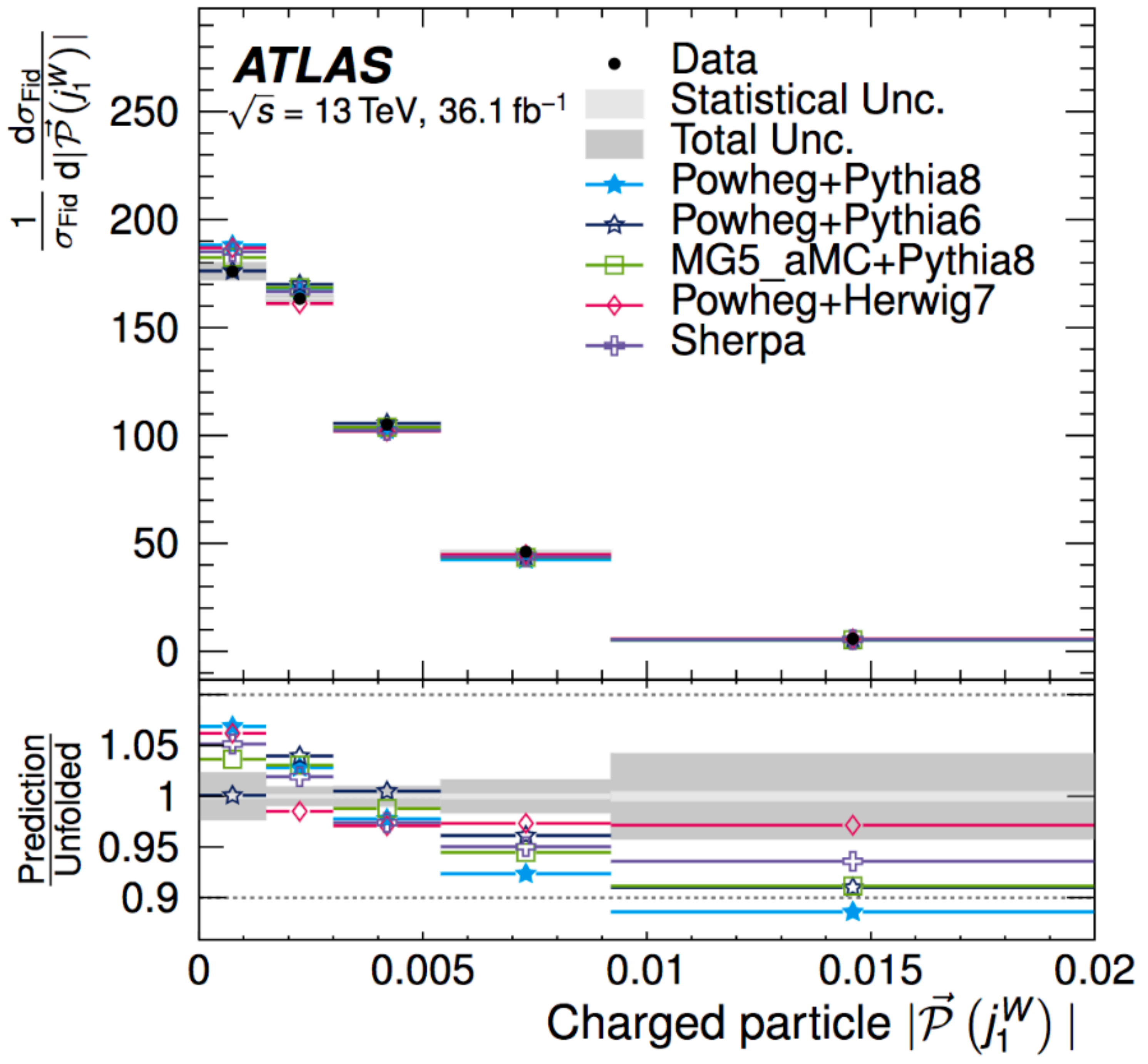
- ***Jet substructure is a quickly developing field***
  - Jet mass measurement demonstrates experimental and theoretical understanding of jet substructure beyond LL accuracy
  - New ideas and predictions are frequent — the first calculations beyond leading logarithmic accuracy were completed a few years ago, and the Lund jet plane was only proposed a two years ago
  - Current measurements of jet substructure lay the foundations for broader explorations of QCD, including measurements of α_S, better parton showers, and more

thanks!

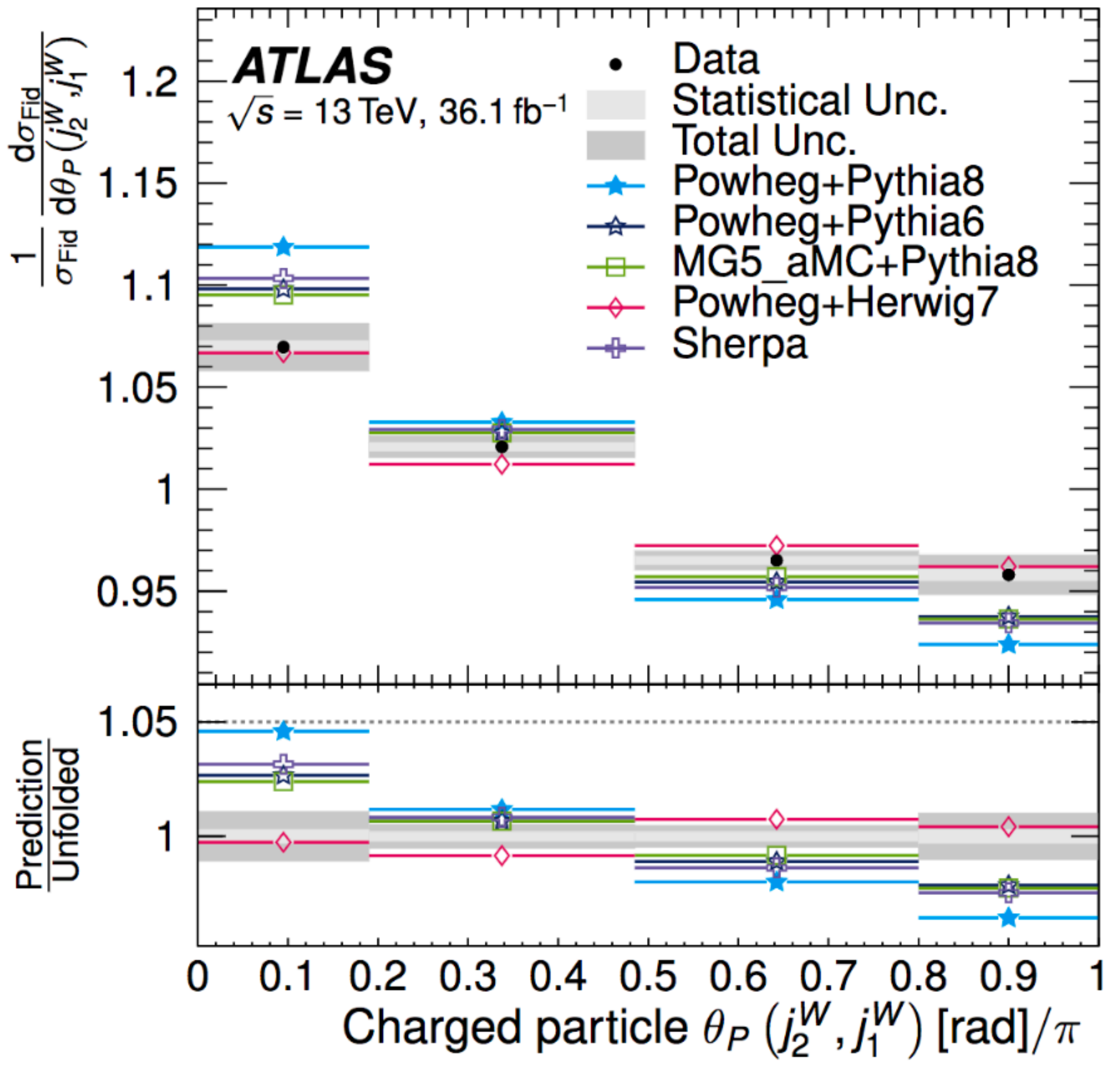
Measurement of the jet pull

- The jet pull is sensitive to color connection in QCD
- ttbar events can contain both color connected and non-connected jets
- No prediction is able to model both the pull angle and magnitude

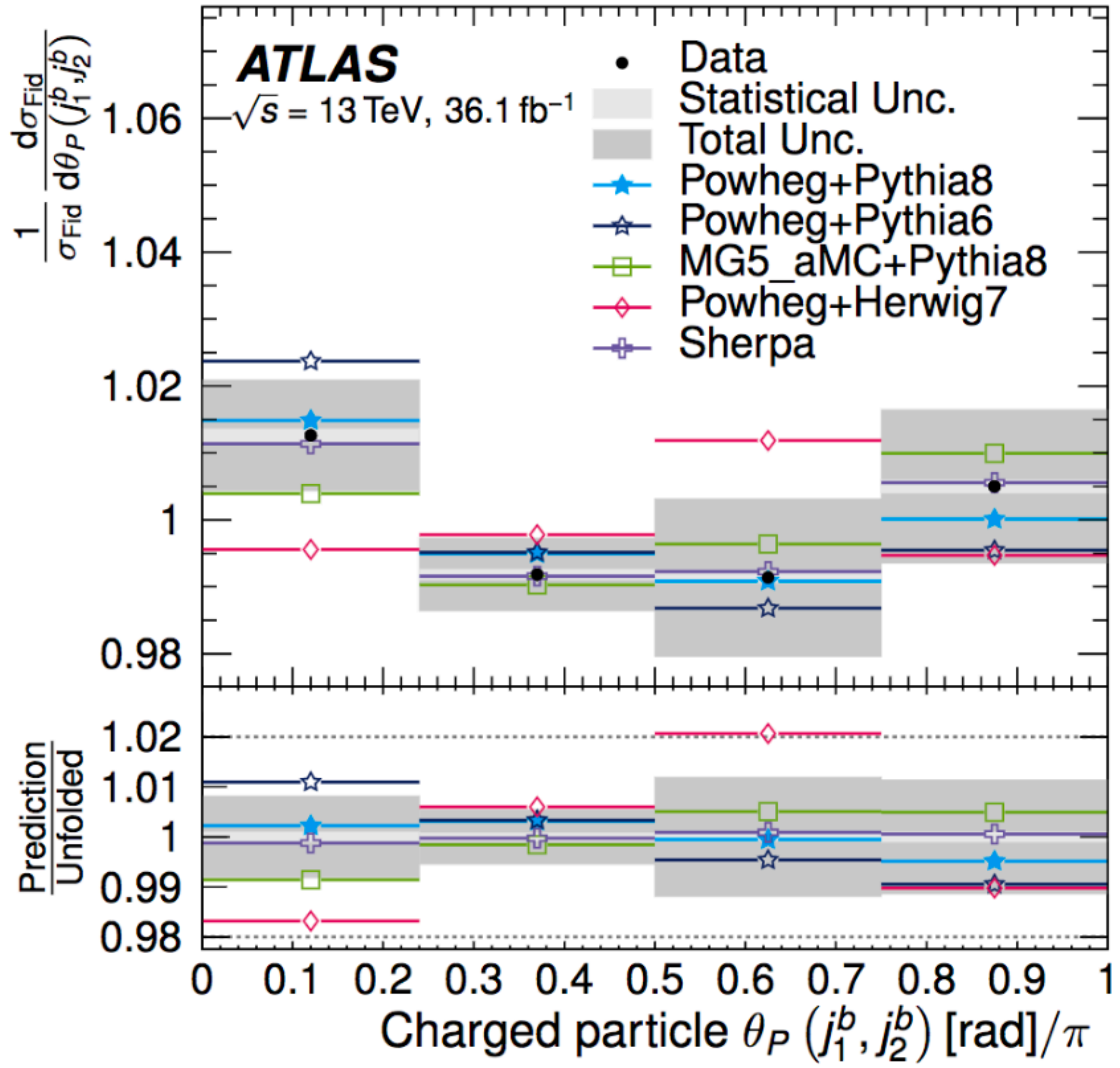




Magnitude of the pull vector



Backward pull angle

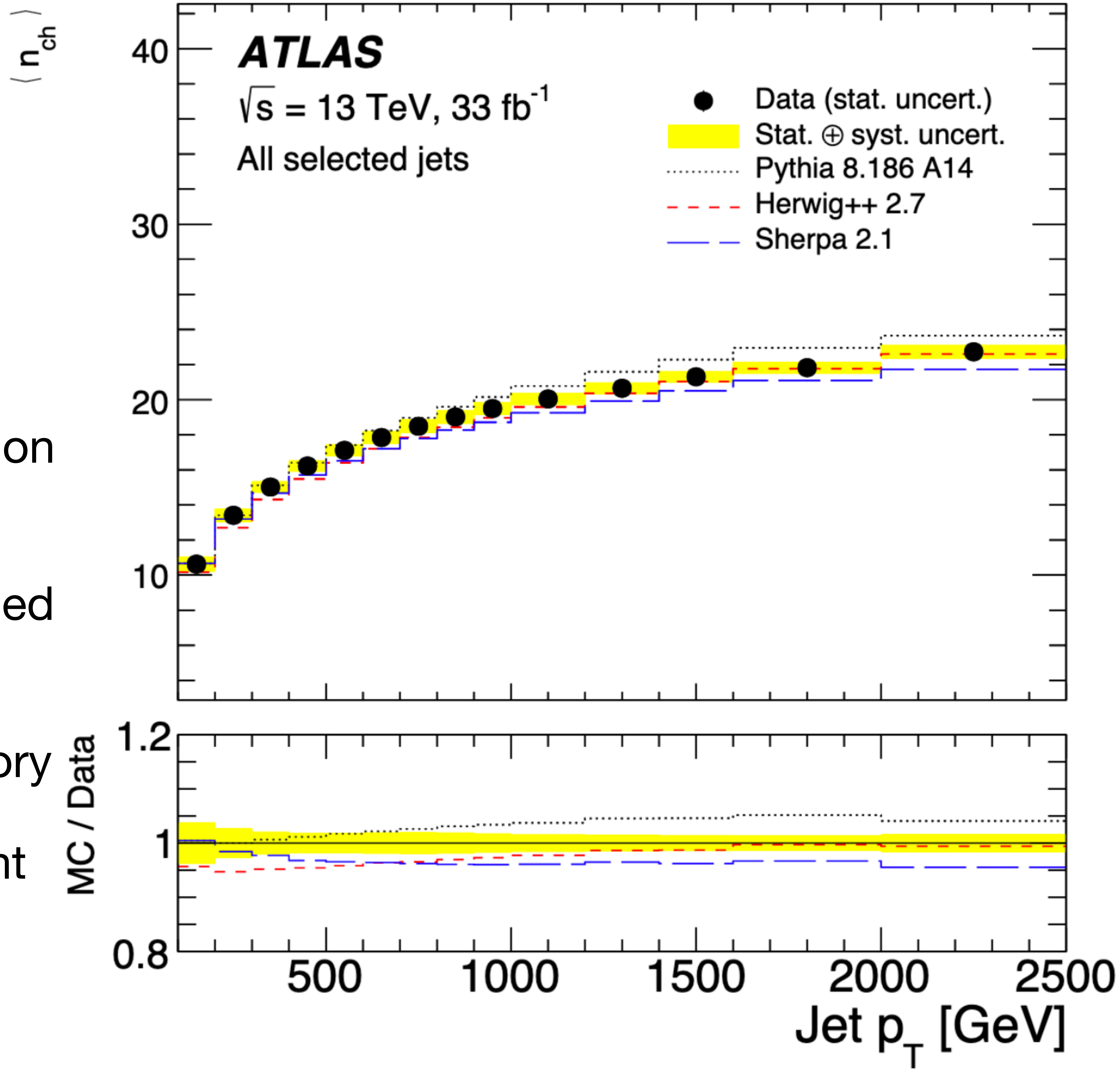


b-quarks (not color connected) Forward Pull Angle

jet fragmentation

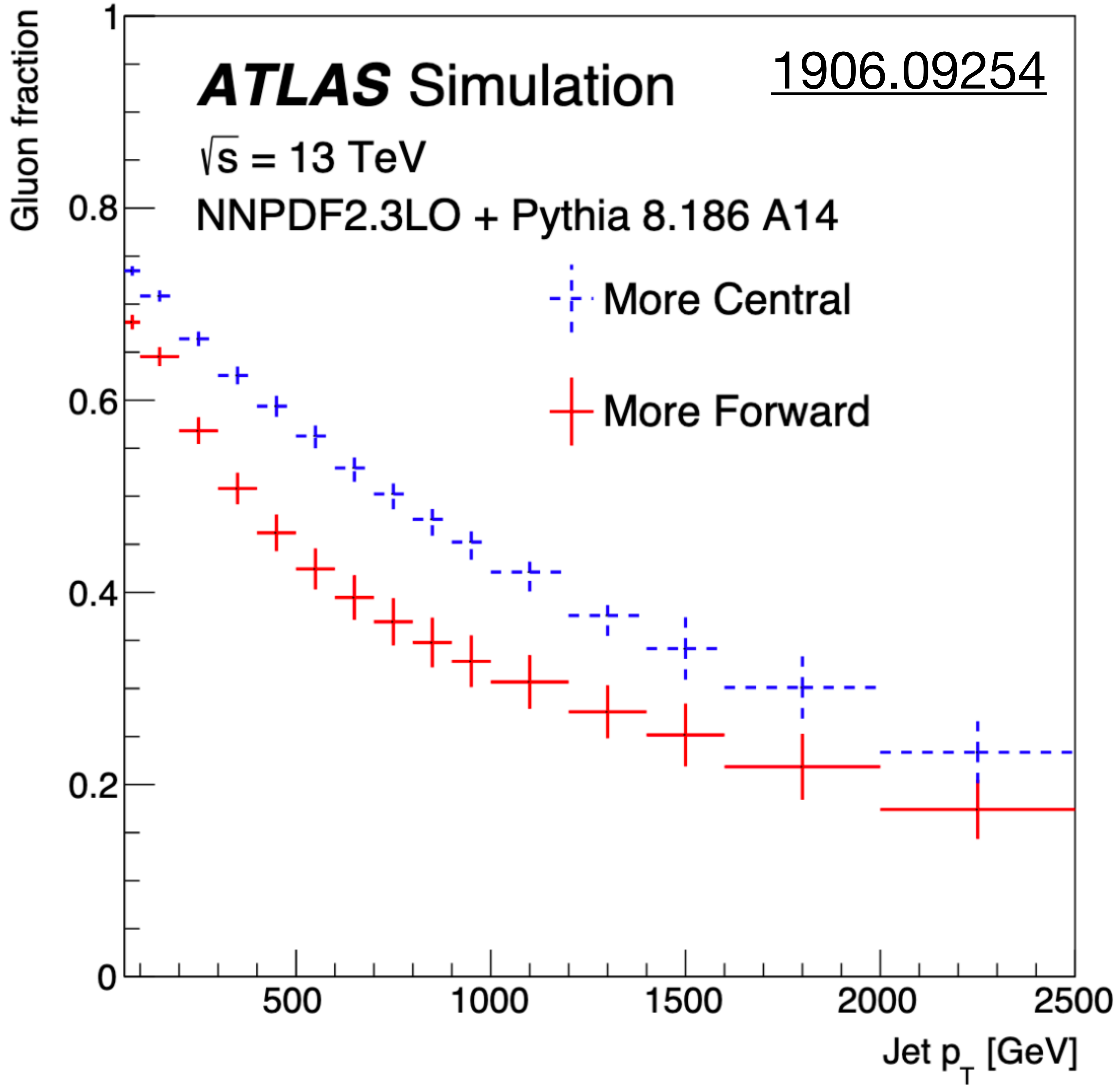
1906.09254

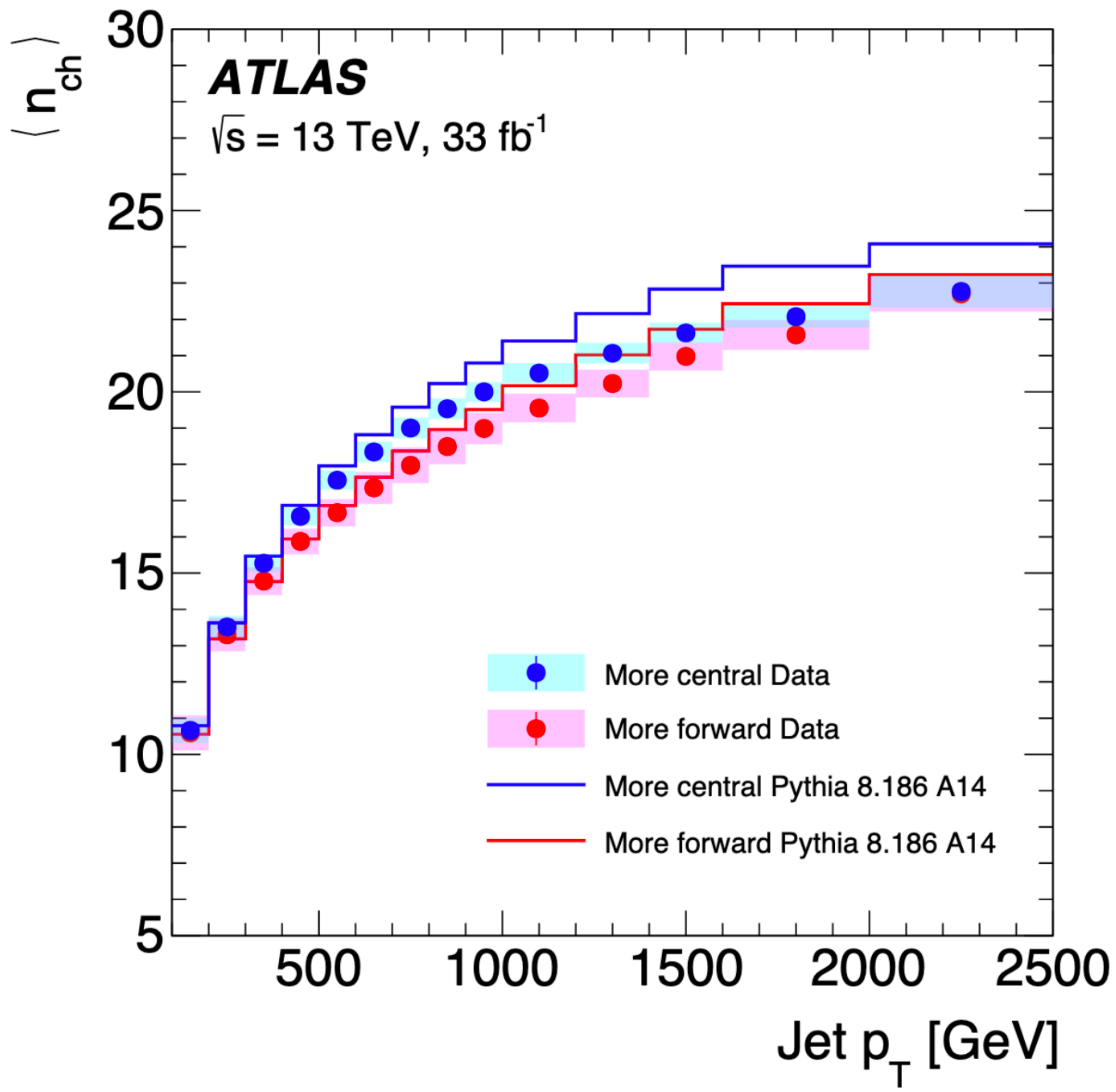
- Jet formation is complicated, and is not fully describable by perturbation theory
  - Rely on Monte Carlo models in order to produce predictions involving jets
- Jet fragmentation measurements study the distribution of particles within a jet
  - Includes observables such as the number of charged particles, the radial profile, and more
  - Energy dependence calculable in perturbation theory
  - Important input for tuning MC, and some significant disagreements between data and MC
- Using tracks to calculate fragmentation to improve precision



jet fragmentation

- Jet fragmentation does not depend strongly on η, just on the initiating parton
- Central jets tend to be gluon initiated more often than forward jets



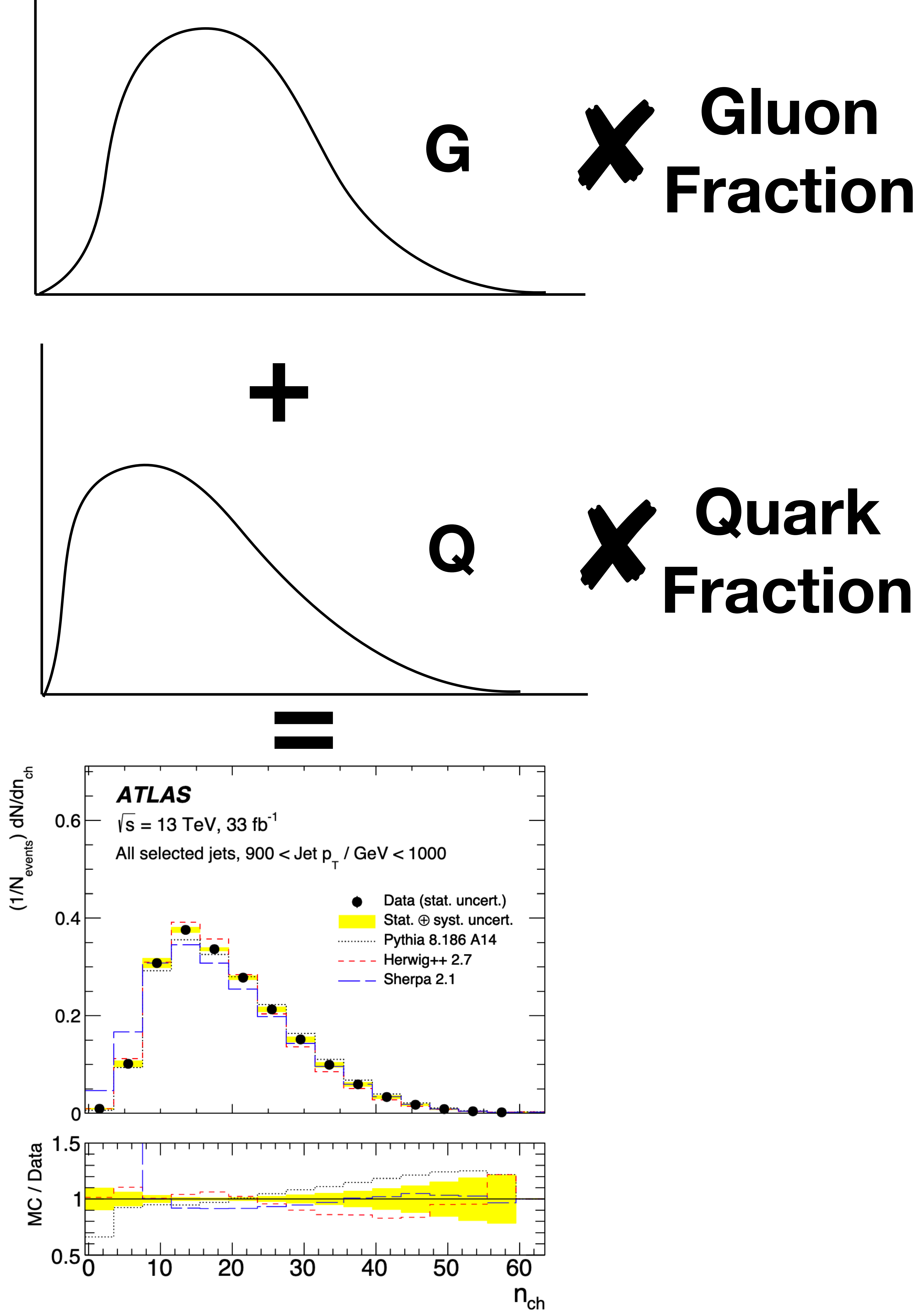


- Measuring forward and central jets separately gives us access to differences between quarks and gluons

jet fragmentation

1906.09254

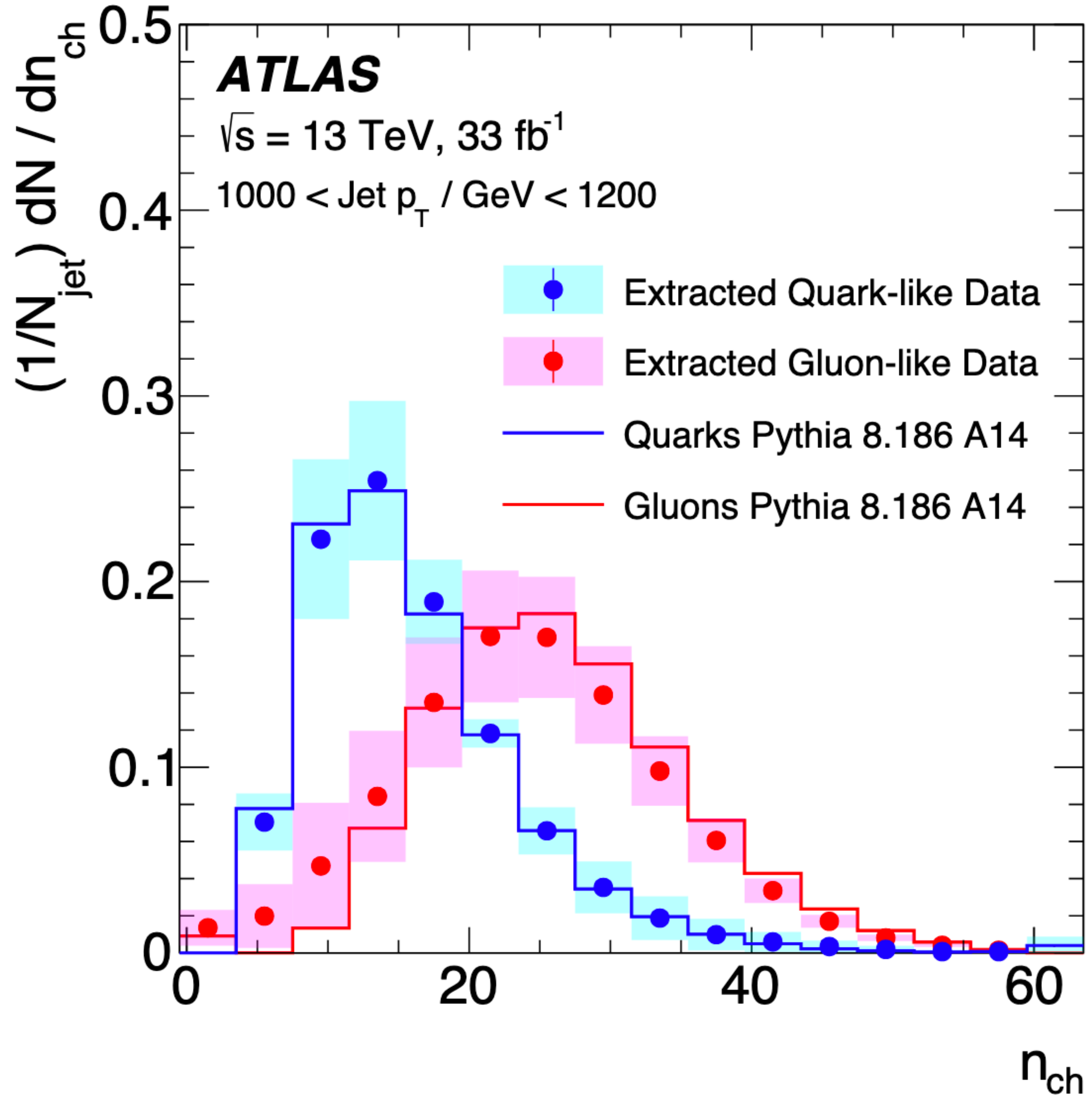
- The measured distributions are a linear combination of the quark and gluon distributions, multiplied by the fraction of quarks and gluons
  - Can invert this to extract the quark and gluon distributions in data
- Two methods:
  - Use the quark and gluon fractions determined in an MC generator (e.g. Pythia)
  - Use topic modeling to extract the distributions, which uses a minimization to separate mutually irreducible distributions

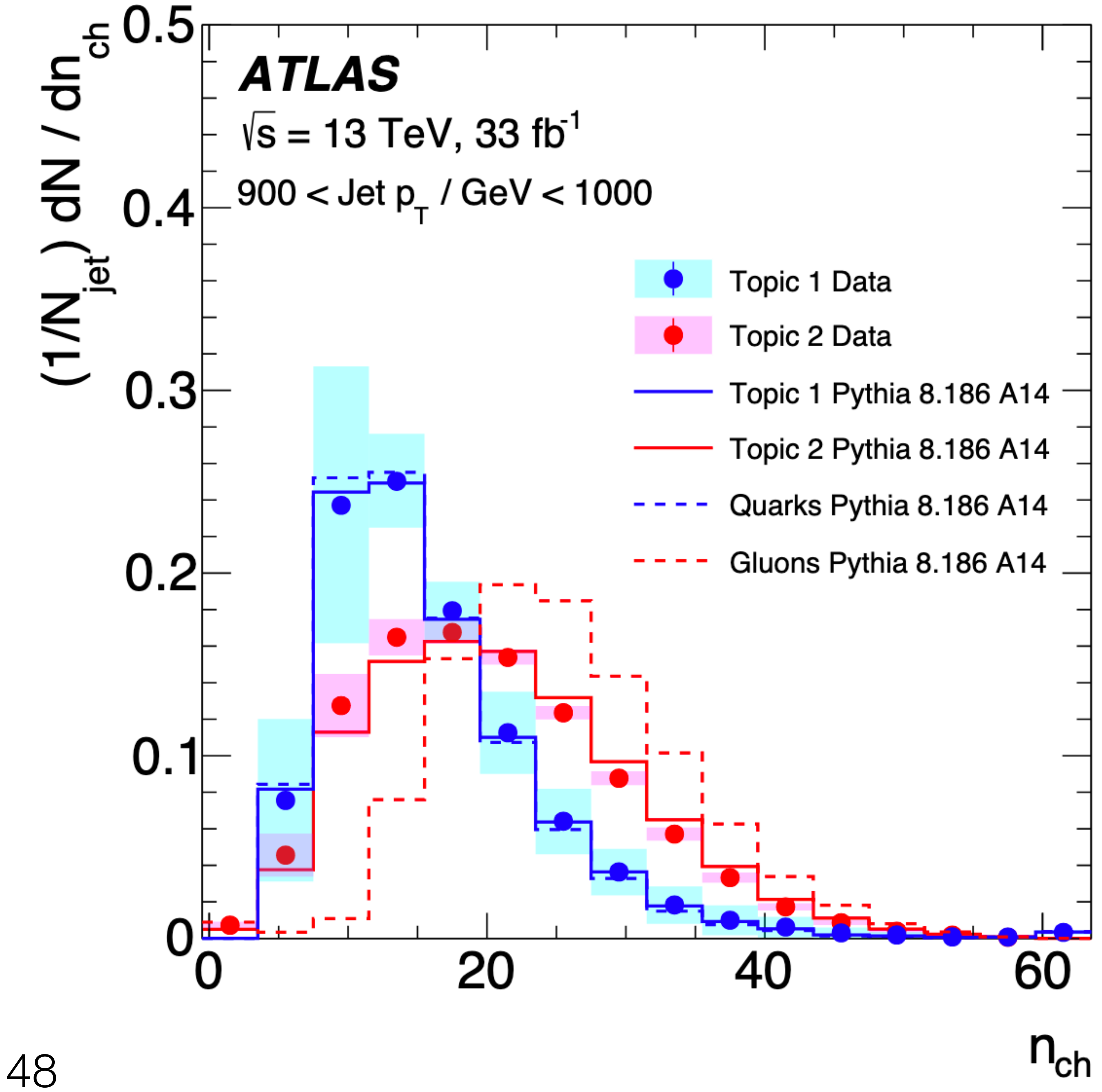


1906.09254

jet fragmentation

- Both methods provide similar results for the extracted quark and gluon distributions
- First time topic modeling has been used in a measurement!
  - Provides more model-independent way of extracting this information

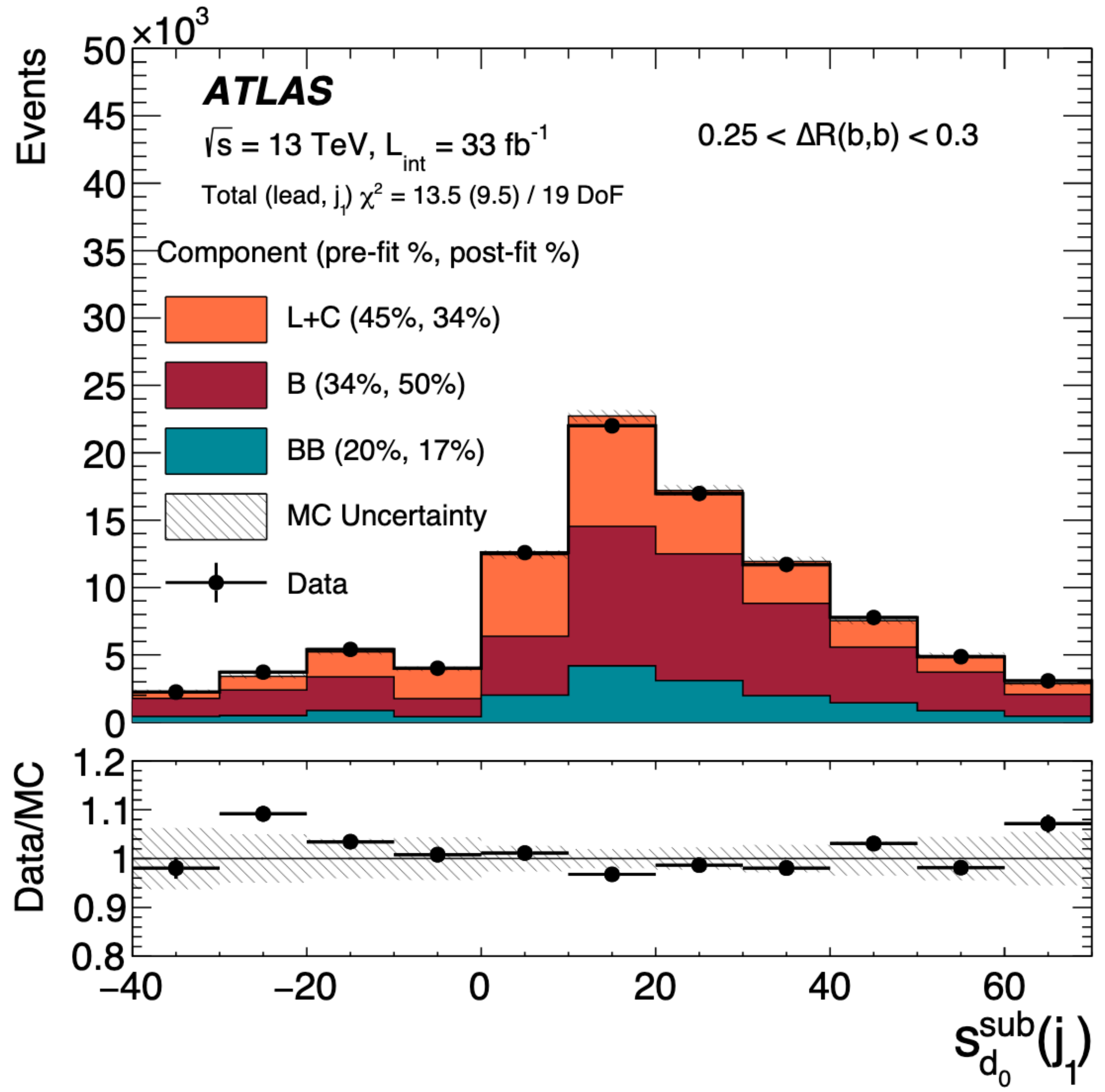


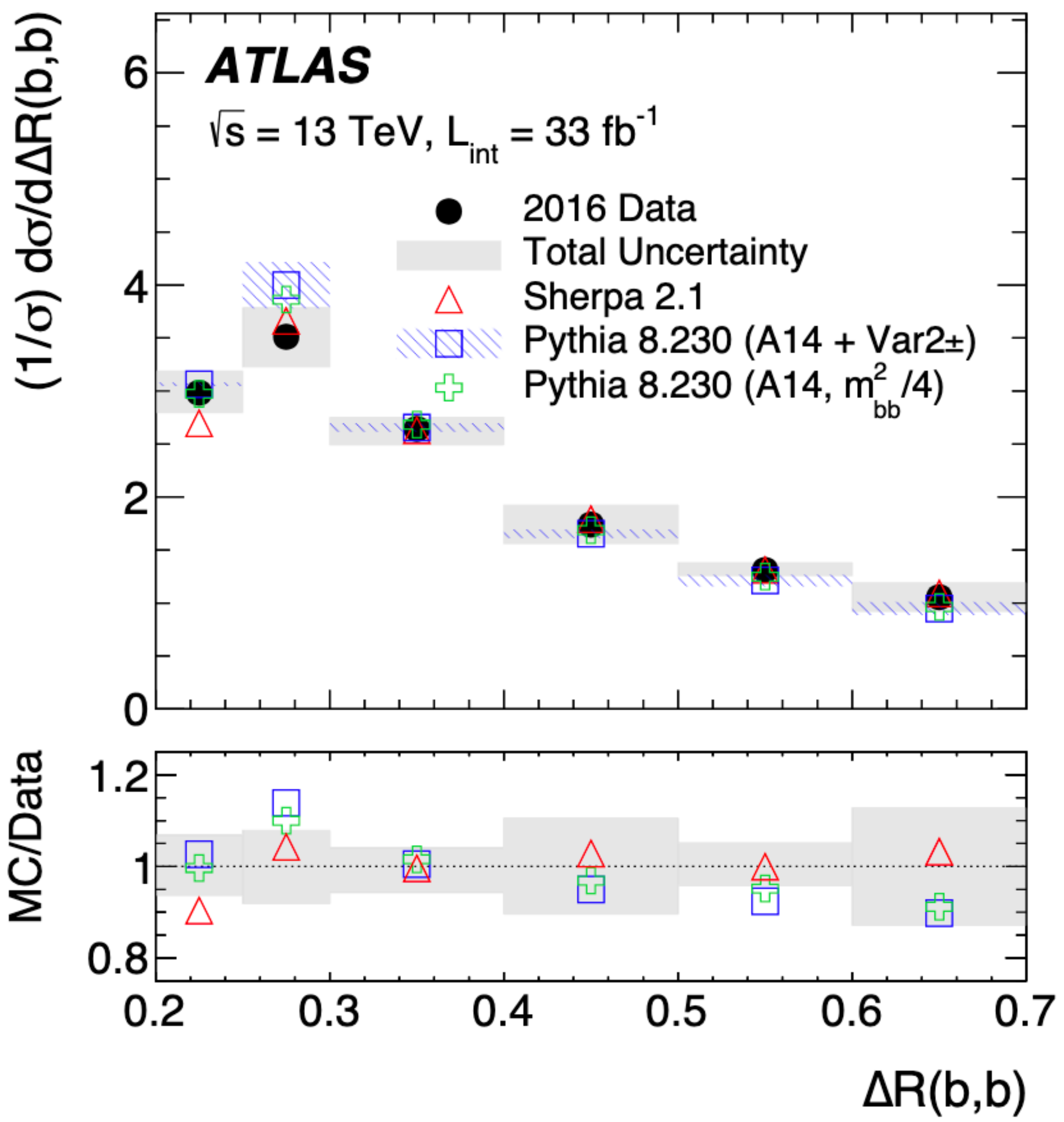


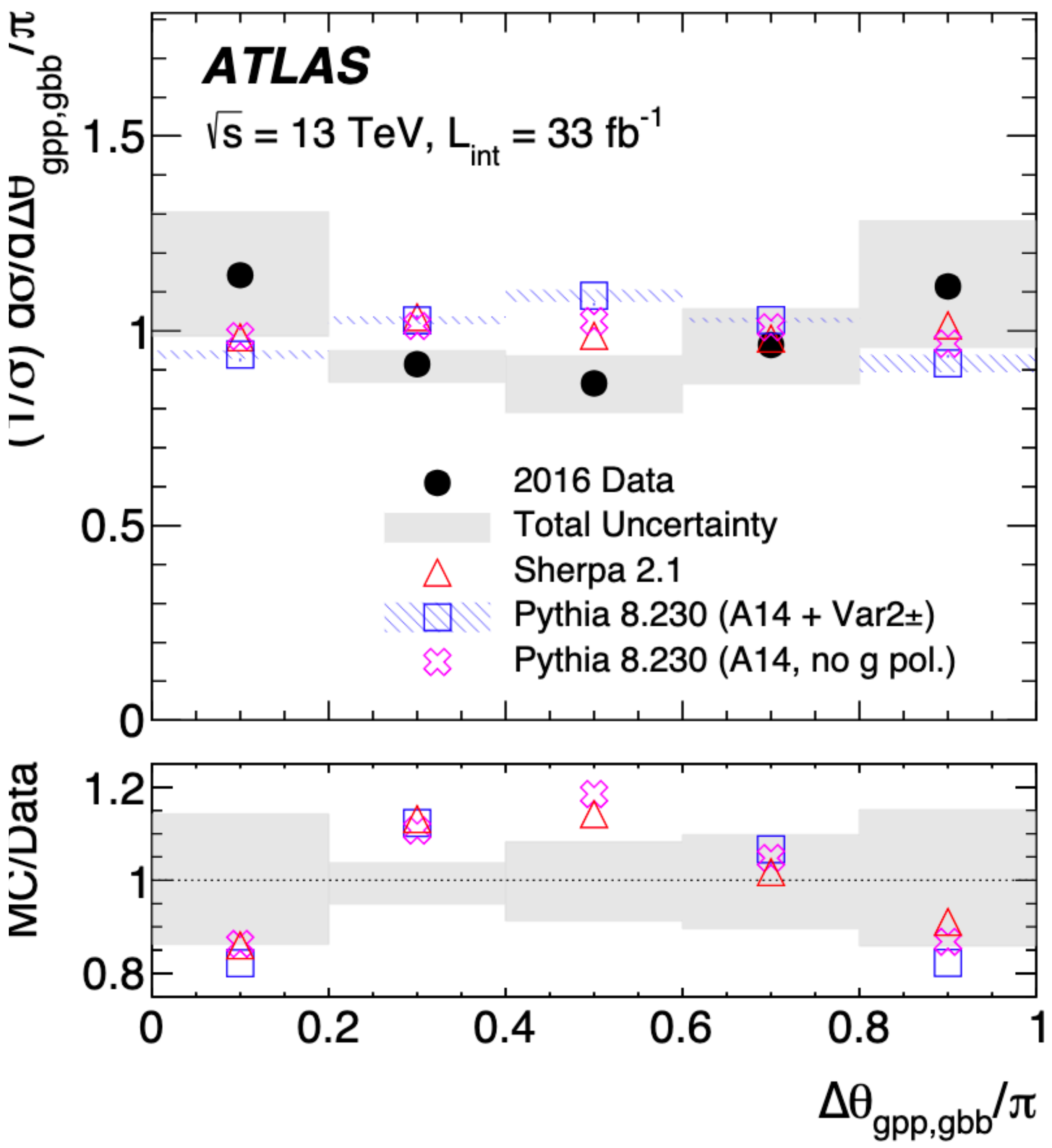
1812.09283

Measurement of g→bb properties

- Gluon fragmentation is challenging to measure, and also important for background modeling
- Using templates to estimate the background for g→bb events





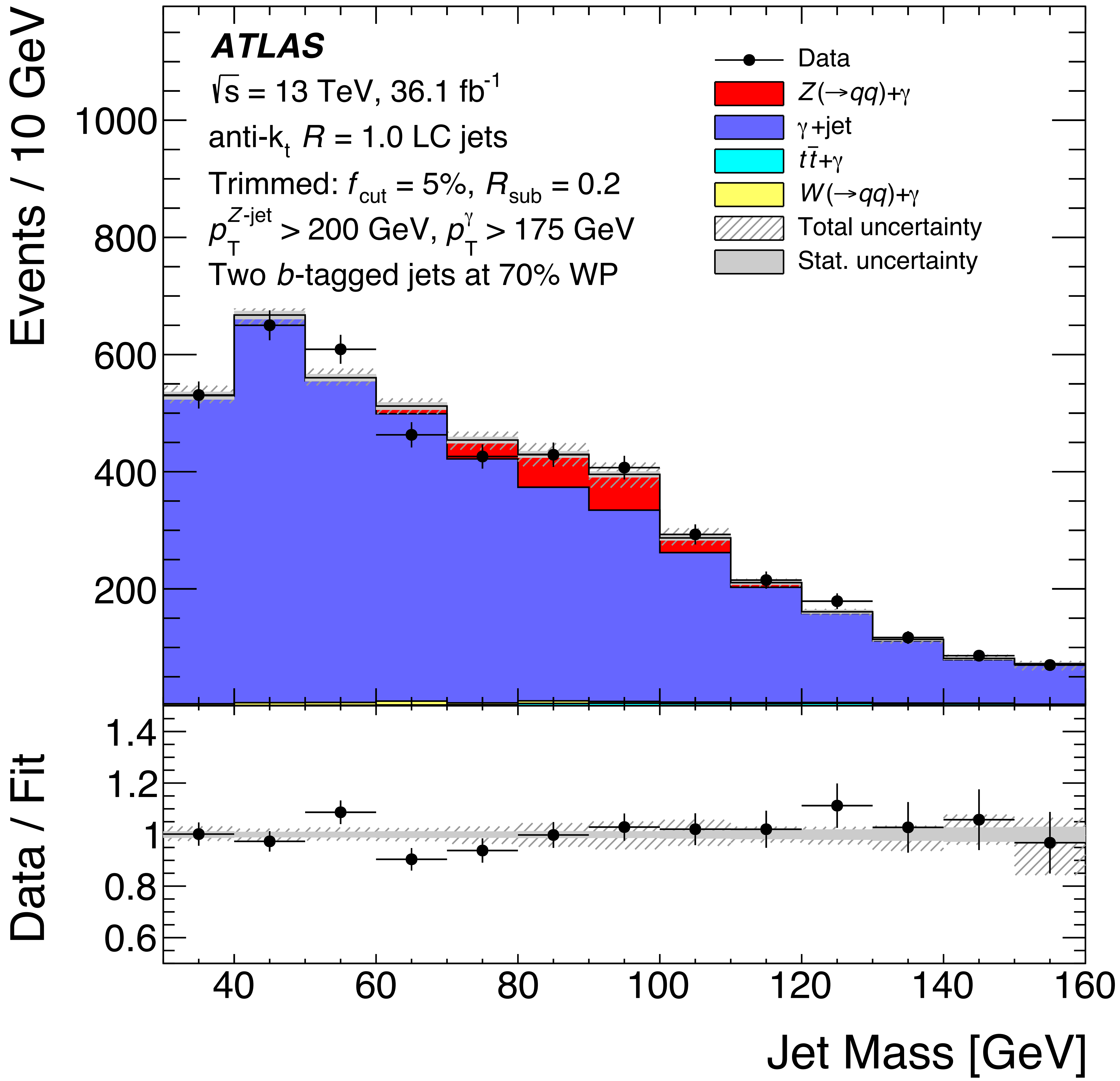


- Certain aspects of the gluon fragmentation are not well-modeled by any of the studied MC predictions

Z(→bb) + γ

1907.07093

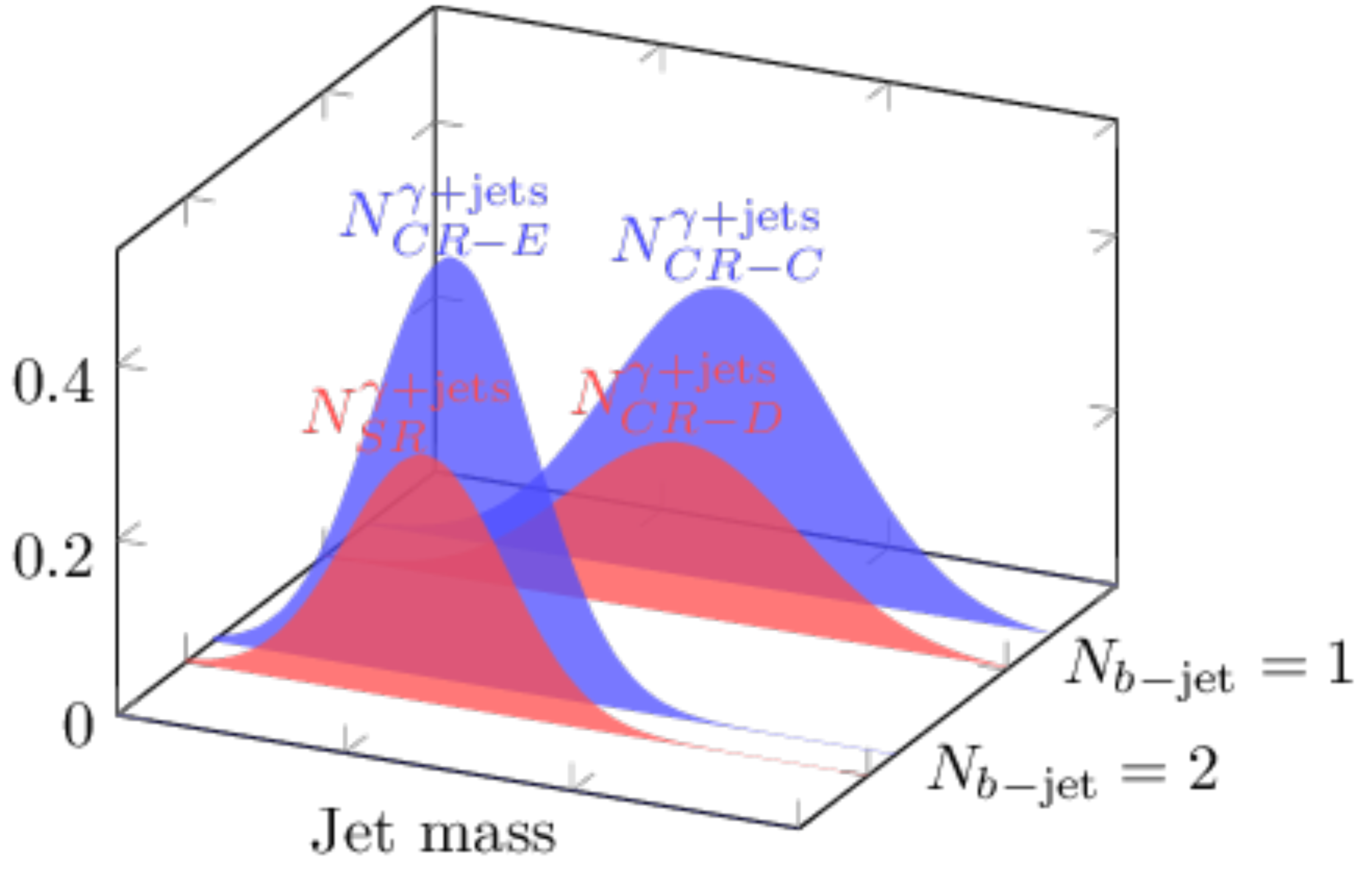
- First measurement of unfolded jet mass spectrum of hadronically decaying Z bosons at the LHC
- Important for understanding boosted boson hadronic decays (color singlets)
  - Have measurements of color octet states such as g→bb



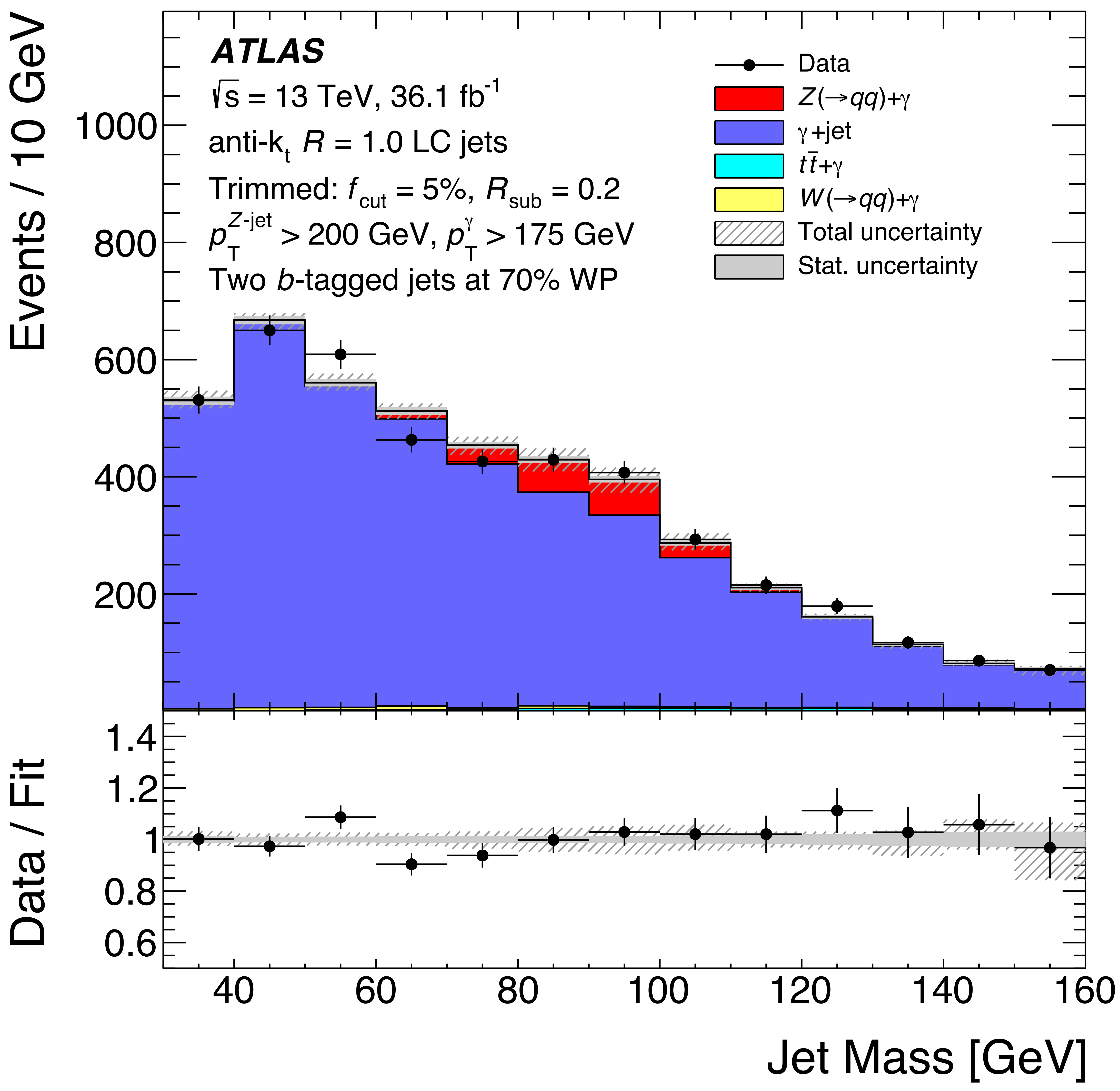
Z(→bb) + γ

1907.07093

- Use γ to trigger on events and for background estimation
- Reconstructing both Z boson decay products within a single large-R jet
- Require two R=0.2 b-tagged subjets to be associated to the large-R jet
- Simultaneously fit signal and background templates to Z(→bb) mass distribution



non-tight γ, tight γ



Jet substructure observables in top quark, W boson, and light jet production

1903.0294

- Slightly different event selections for each different final state
-

	Detector level	Particle level
Dijet selection:		
Two trimmed anti-k_t $R = 1.0$ jets	$p_T > 200$ GeV $\lvert\eta\rvert < 2.5$	$p_T > 200$ GeV $\lvert\eta\rvert < 2.5$
Leading-p_T trimmed anti-k_t $R = 1.0$ jet	$p_T > 450$ GeV	
Top and W selections:		
Exactly one muon	$p_T > 30$ GeV $\lvert\eta\rvert < 2.5$ $\lvert z_0 \sin(\theta)\rvert < 0.5$ mm and $\lvert d_0/\sigma(d_0)\rvert < 3$	$p_T > 30$ GeV $\lvert\eta\rvert < 2.5$
Anti-k_t $R = 0.4$ jets	$p_T > 25$ GeV $\lvert\eta\rvert < 4.4$ JVT output > 0.5 (if $p_T < 60$ GeV)	$p_T > 25$ GeV $\lvert\eta\rvert < 4.4$
Muon isolation criteria	If $\Delta R(\mu, \text{jet}) < 0.04 + 10\text{ GeV}/p_{T,\mu}$: muon is removed, so the event is discarded	None
E_T^{miss}, m_T^W	$E_T^{miss} > 20$ GeV, $E_T^{miss} + m_T^W > 60$ GeV	
Leptonic top	At least one small-radius jet with $0.4 < \Delta R(\mu, \text{jet}) < 1.5$	
Top selection:		
Leading-p_T trimmed anti-k_t $R = 1.0$ jet	$\lvert\eta\rvert < 1.5$, $p_T > 350$ GeV, mass > 140 GeV ΔR(large-radius jet, b-tagged jet) < 1 $\Delta\phi(\mu$, large-radius jet) > 2.3	
W selection:		
Leading-p_T trimmed anti-k_t $R = 1.0$ jet	$\lvert\eta\rvert < 1.5$, $p_T > 200$ GeV, mass > 60 GeV and mass < 100 GeV $1 < \Delta R$(large-radius jet, b-tagged jet) < 1.8 $\Delta\phi(\mu$, large-radius jet) > 2.3	

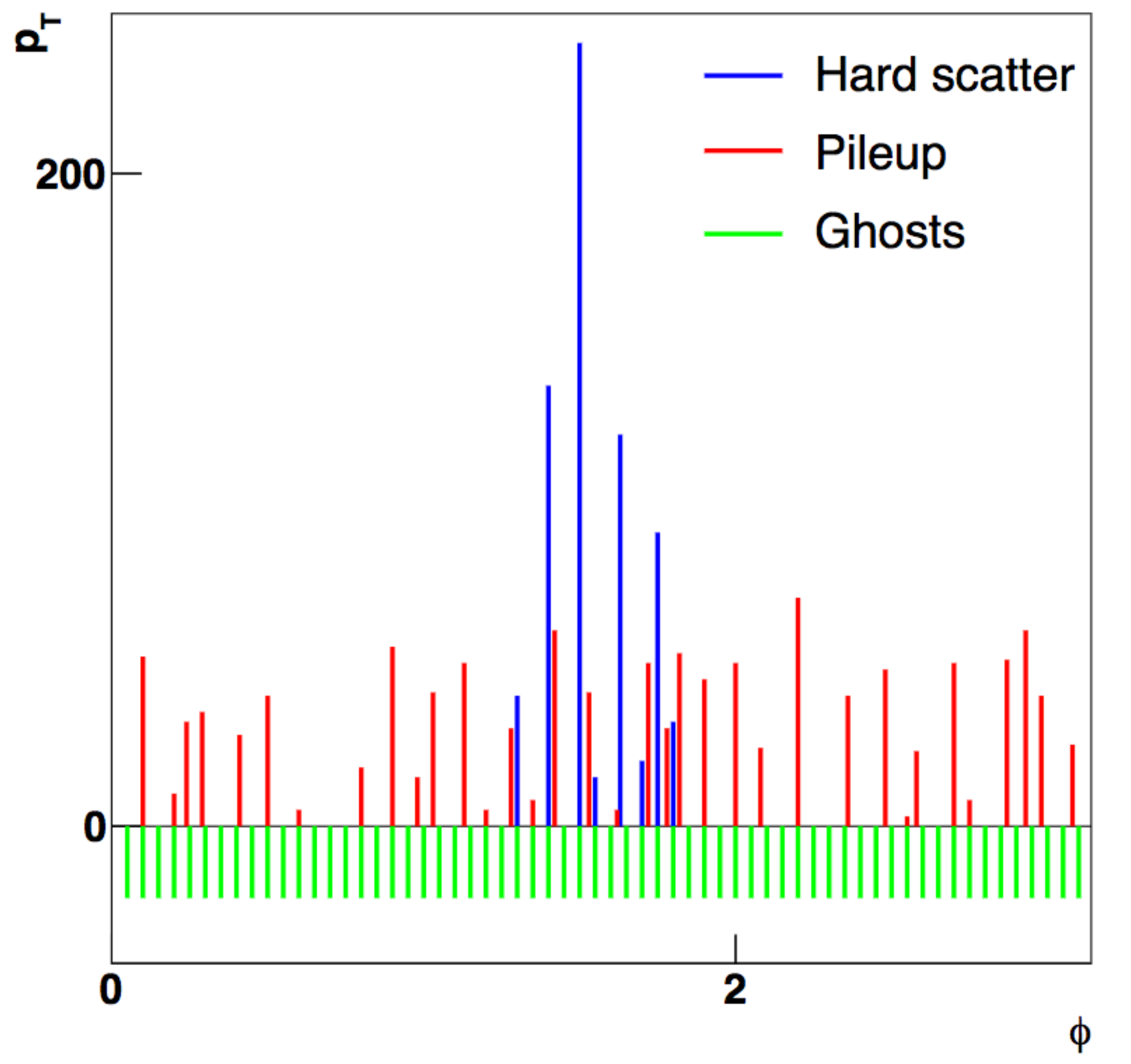
Jet substructure observables in top quark, W boson, and light jet production

1903.0294

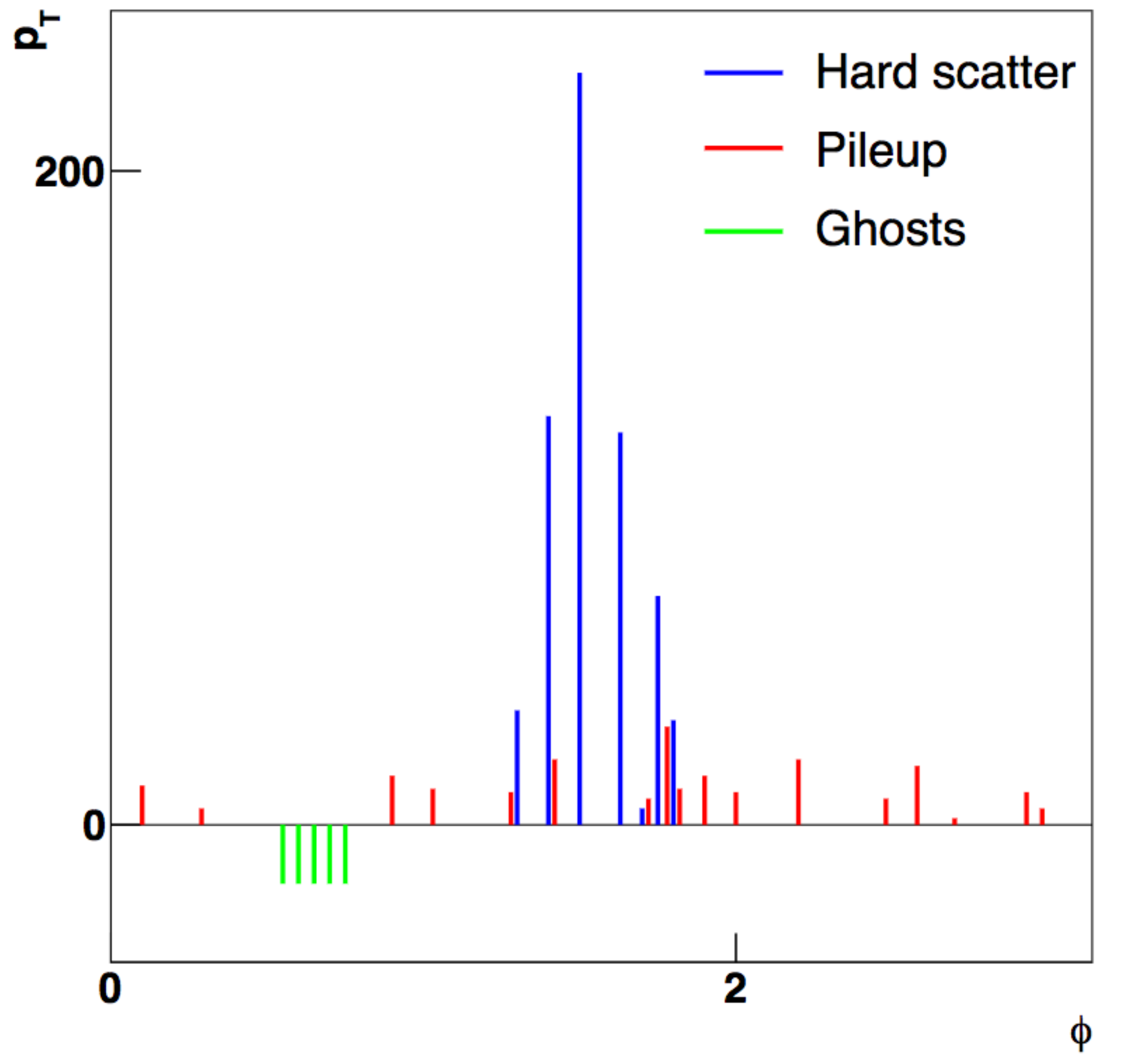
- Number of subjets follows expectation
- Most results consistent with the different generators

Pileup Mitigation: Constituent Subtraction

- Constituent-level pileup mitigation technique which *rescales* the constituent 4-momentum
- Adds ghosts evenly throughout an event with p_T density equal to the median energy density ρ
- Ghosts matched to constituents, and the ghost p_T is subtracted off
  - Only matched within some maximum ΔR of the constituent
- After subtraction, the median energy density should be approximately zero
- Note that for PFlow and TCCs, only the neutrals are used and corrected



Whole event before correction



Whole event after correction

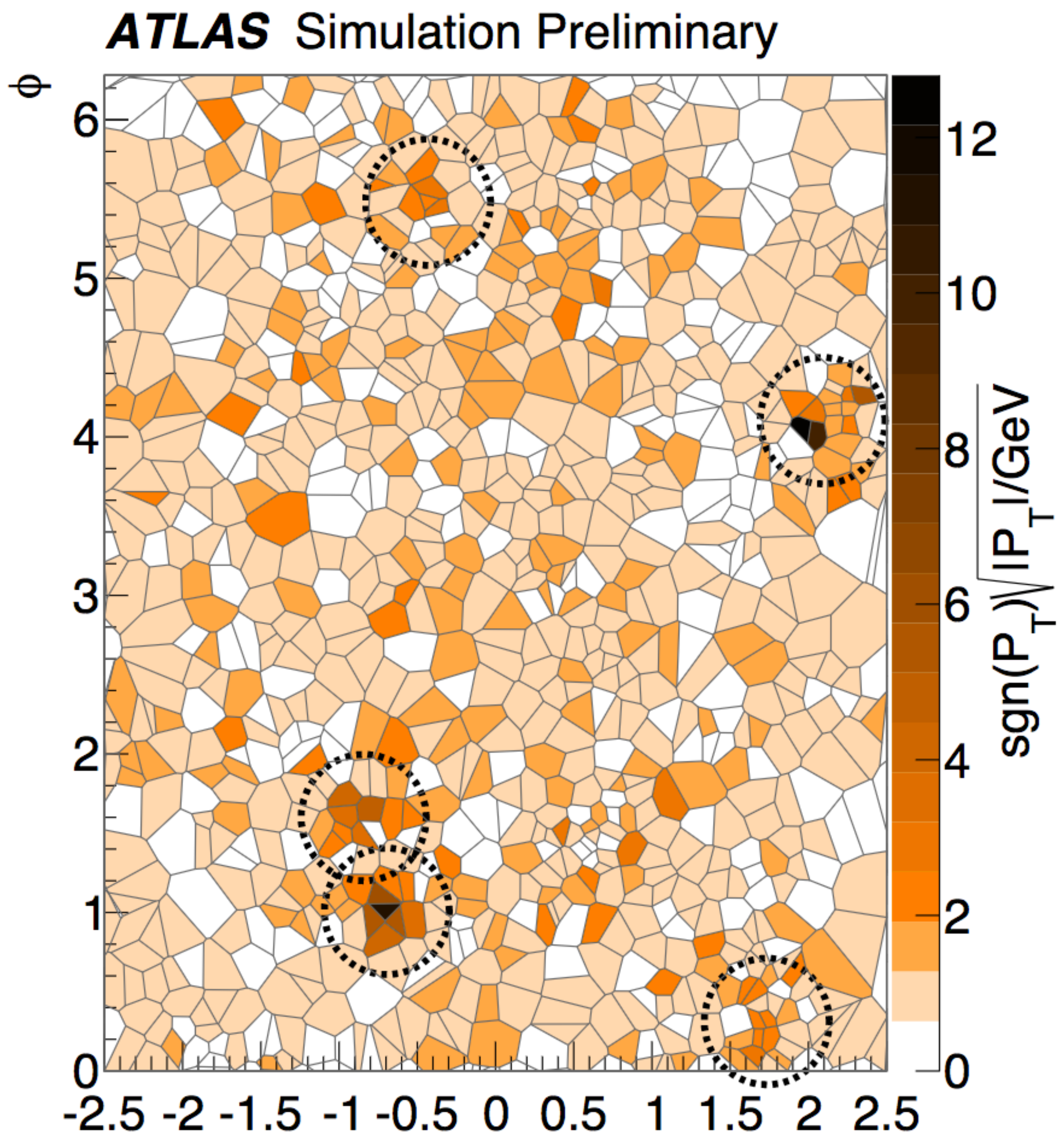
Constituent Subtraction Paper

CONF note on pileup mitigation

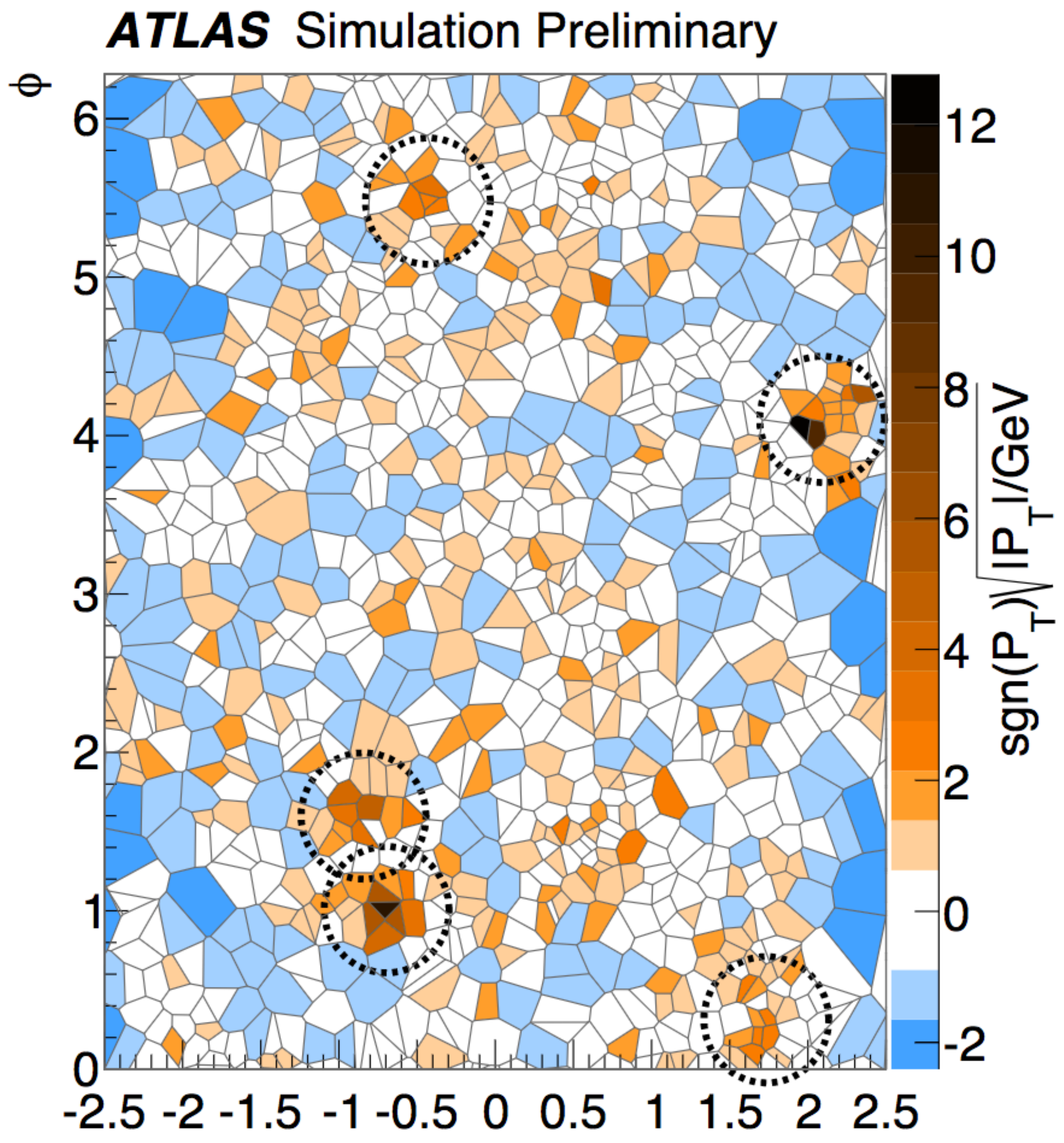
Pileup Mitigation: Voronoi Subtraction

CONF note on pileup subtraction

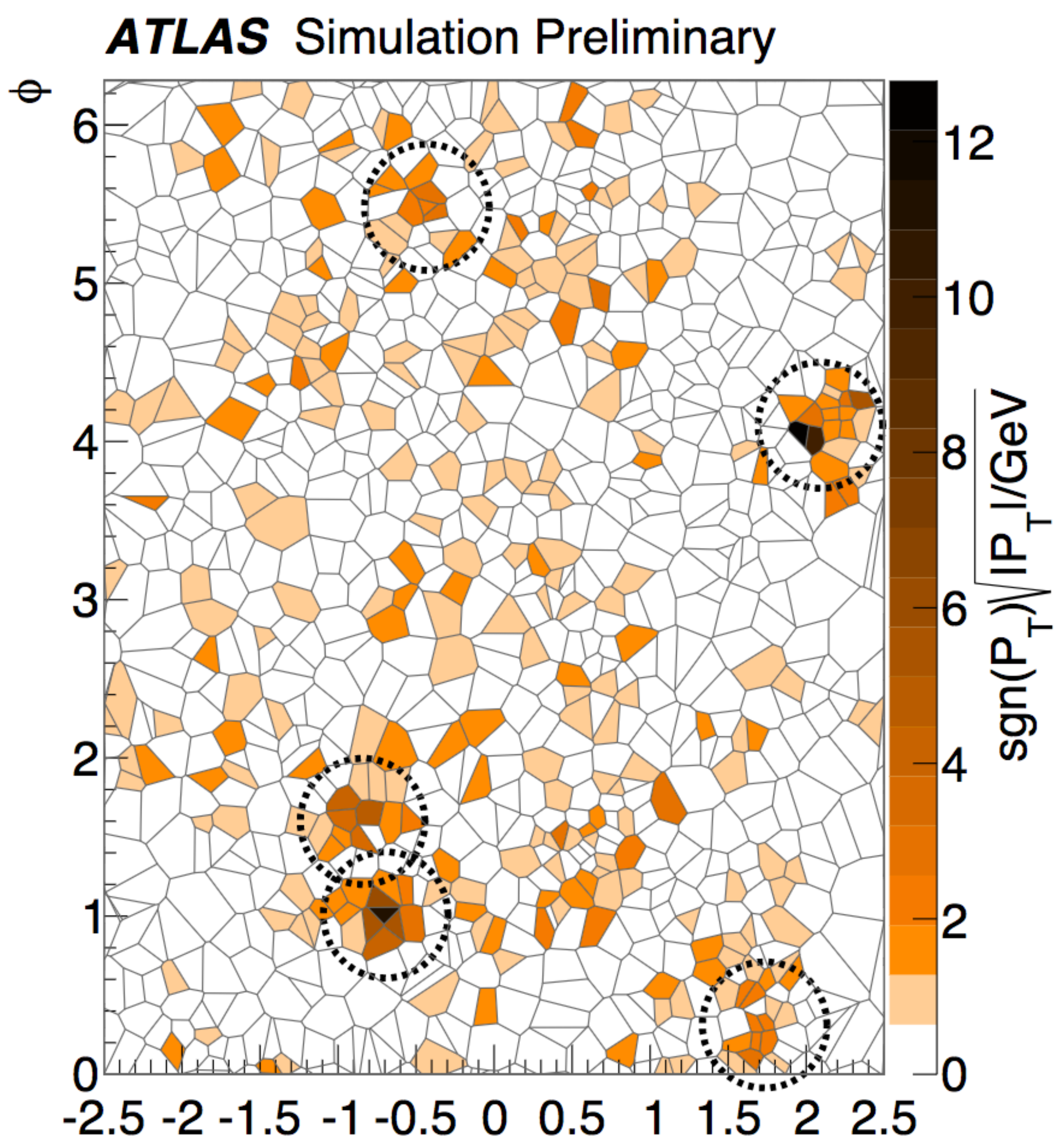
- Voronoi subtraction is a type of constituent-level pileup mitigation which uses the **median energy density** (ρ) and the **Voronoi area** to reweight constituents
  - Voronoi area is the area of points in η-ϕ space which are closer to a constituent than any other
- Leaves some constituents with negative p_T — **Voronoi suppression** discards any constituents with negative p_T



Uncorrected



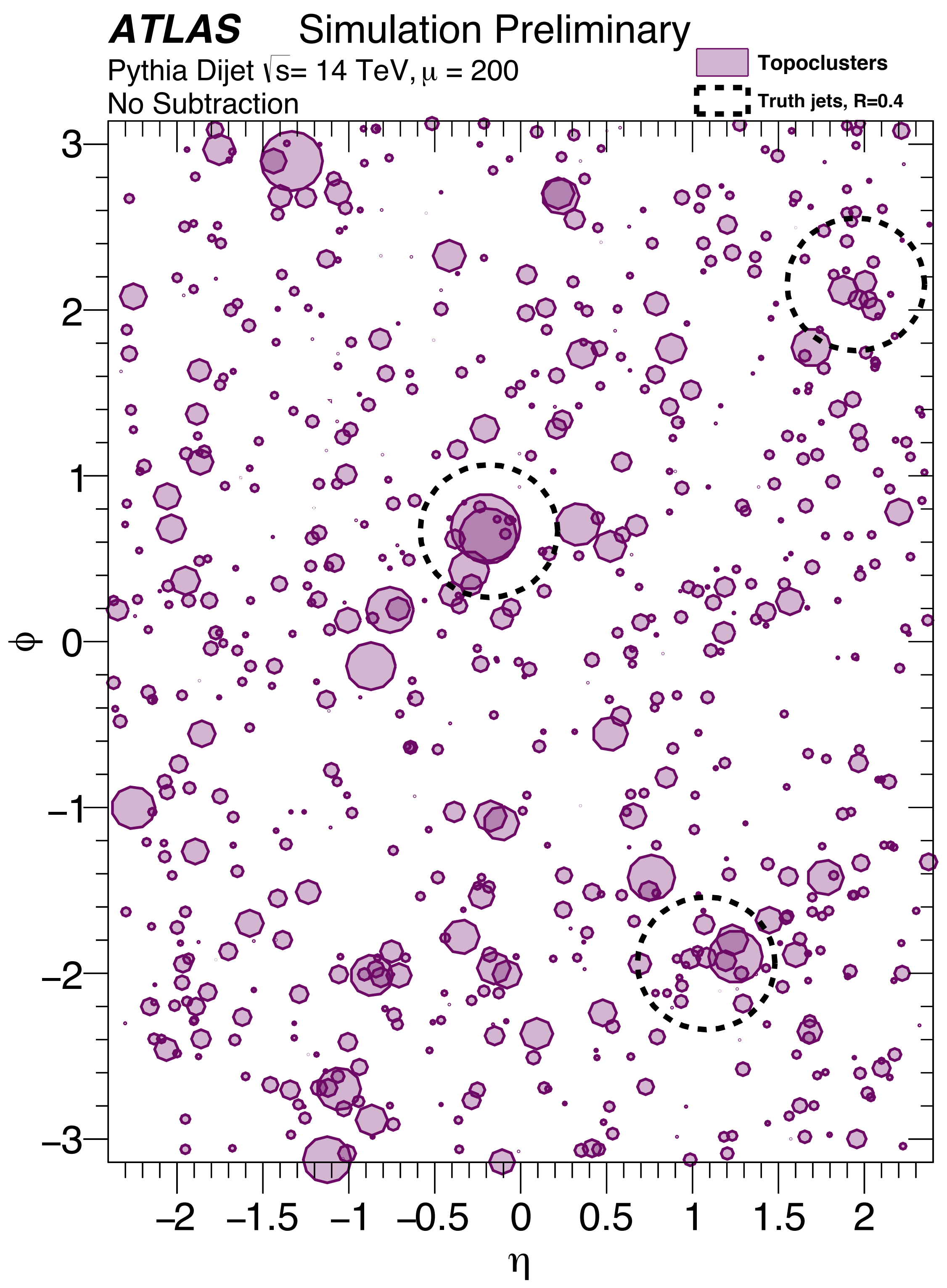
Subtraction

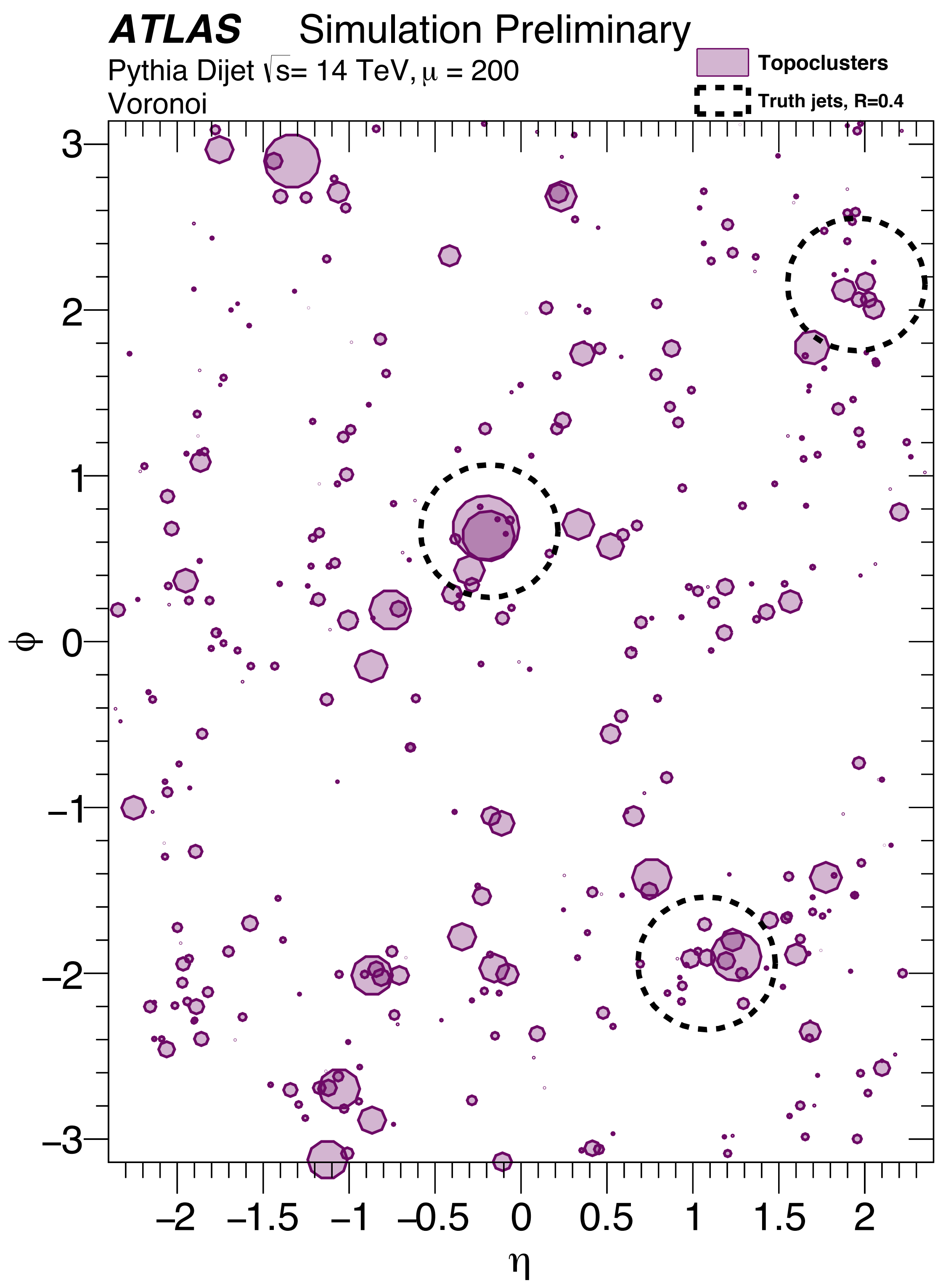


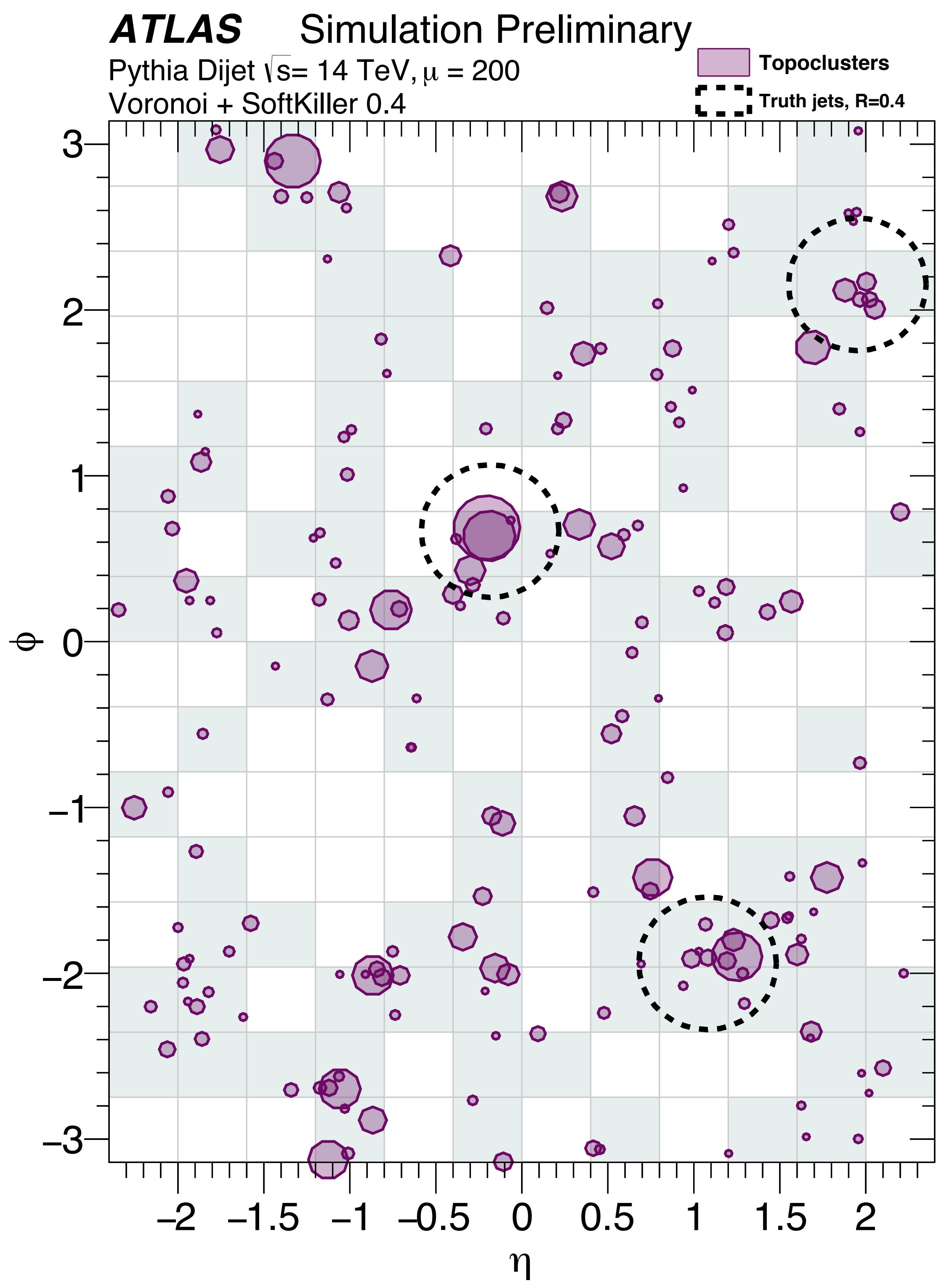
Suppression

Pileup Mitigation: SoftKiller

- Determines an event-by-event p_T cut for constituents
  - Should apply either Voronoi Subtraction or Constituent Subtraction first
- Makes a grid, finds p_T cut where half of grid cells are empty afterwards
- Note that for PFlow and TCCs, only the neutrals are used and corrected





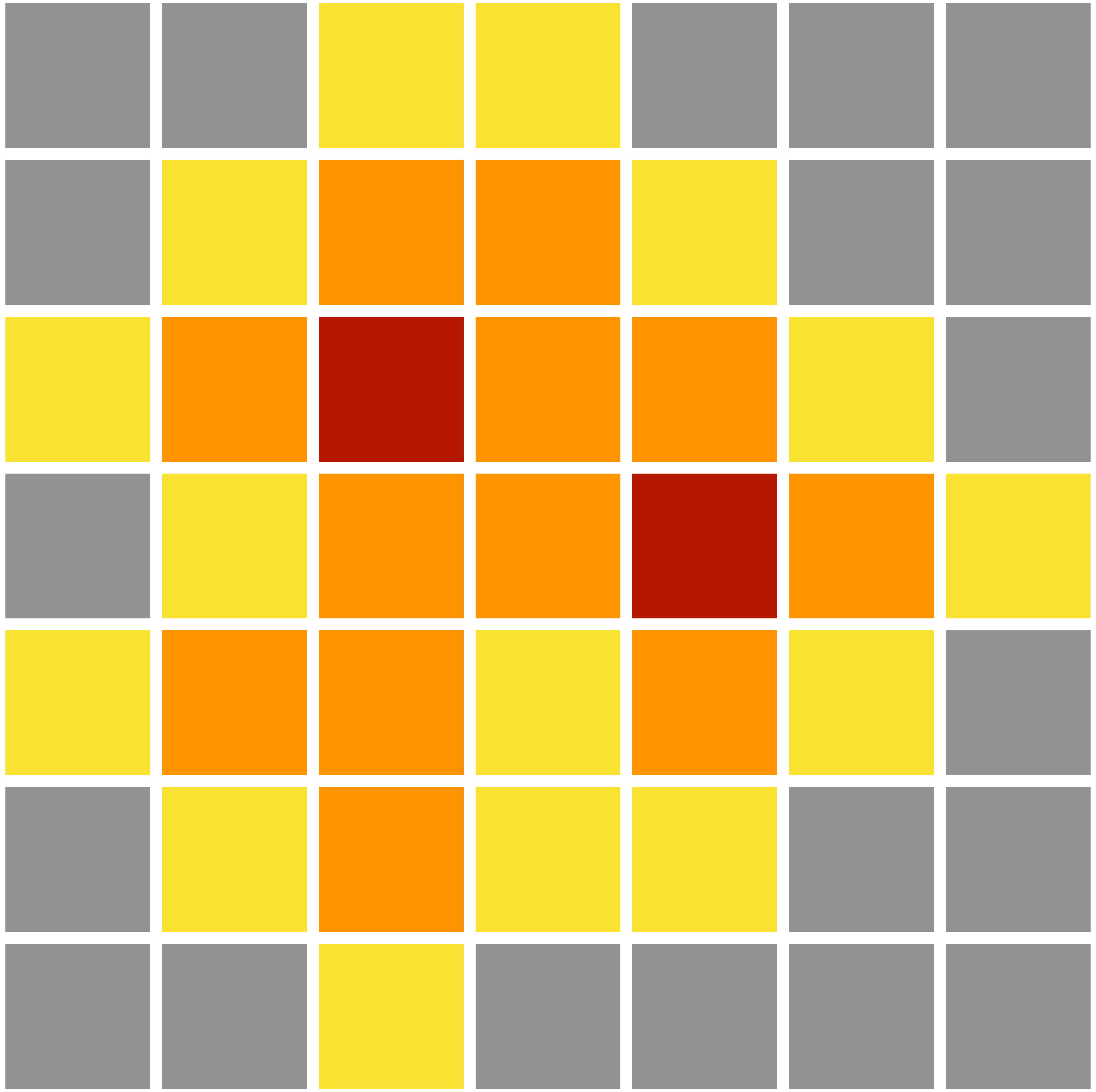


SoftKiller Paper

CONF note on pileup mitigation

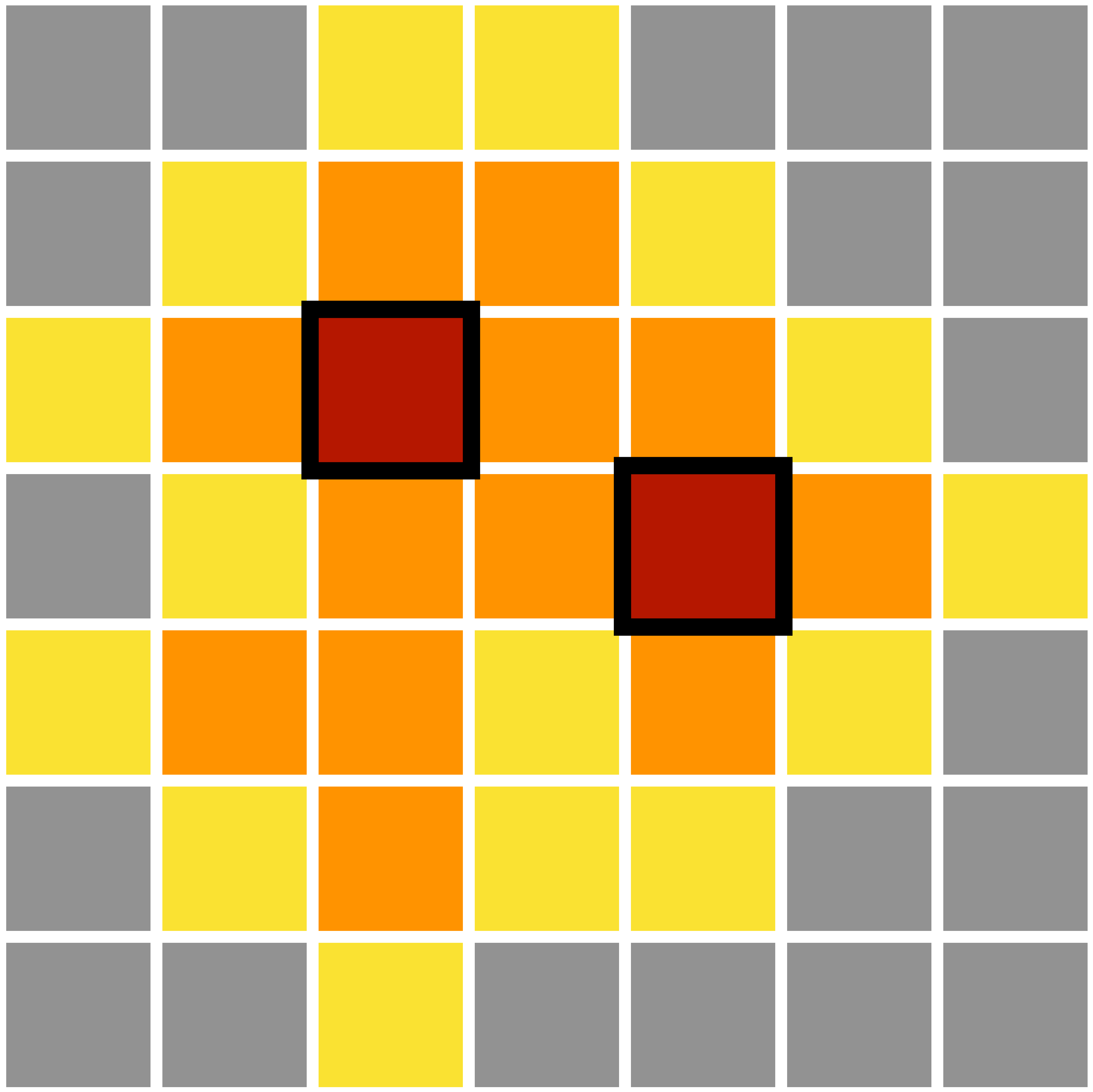
jet inputs: topoclusters

- Jets are composed of both charged and neutral hadrons
  - Isospin symmetry means that observables should be similar when constructed from all particles or only charged particles
- Analytical predictions only for all-particles
  - Need to use the information from the calorimeter
- Combine nearby groups of calorimeter cells into clusters in order to produce object which approximately corresponds to a single particle



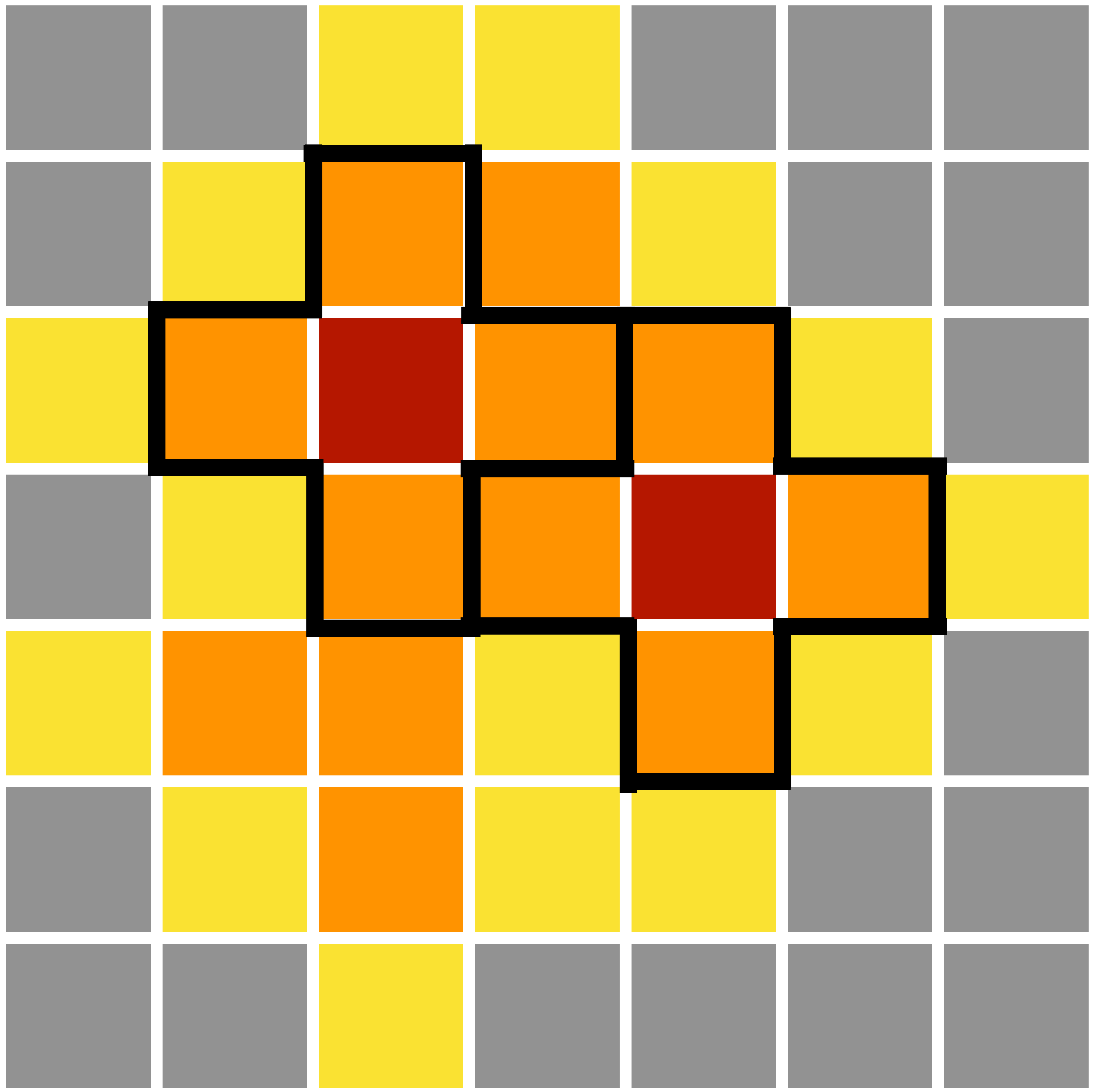
jet inputs: topoclusters

- Use the ‘4-2-0’ algorithm for reconstruction
  - Clusters are seeded by cells with energy of 4σ above the expected noise



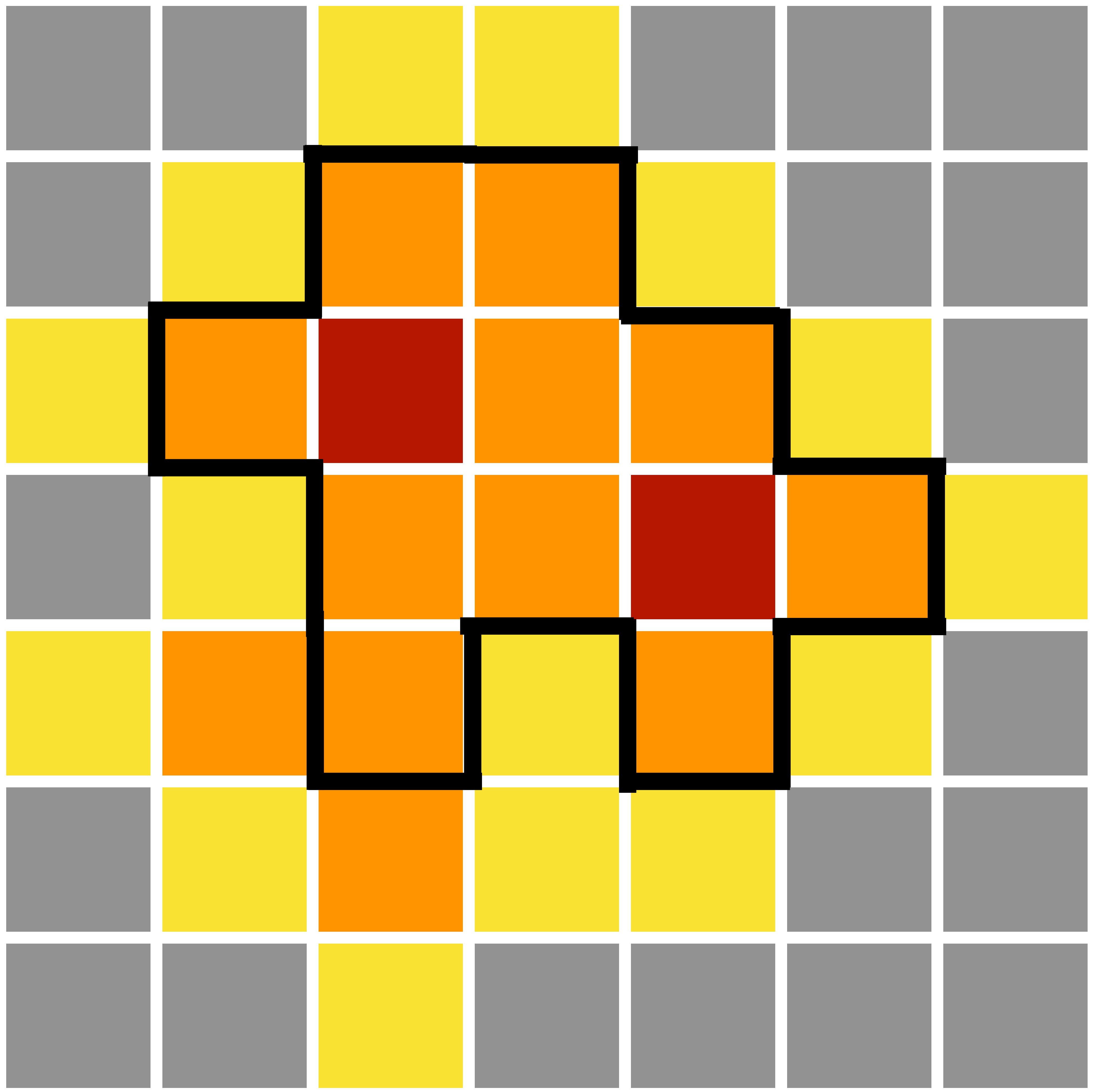
jet inputs: topoclusters

- Use the '4-2-0' algorithm for reconstruction
  - Clusters are seeded by cells with energy of 4σ above the expected noise
  - Any neighboring cells with E>2σ are added recursively until no high-energy neighboring cells remain



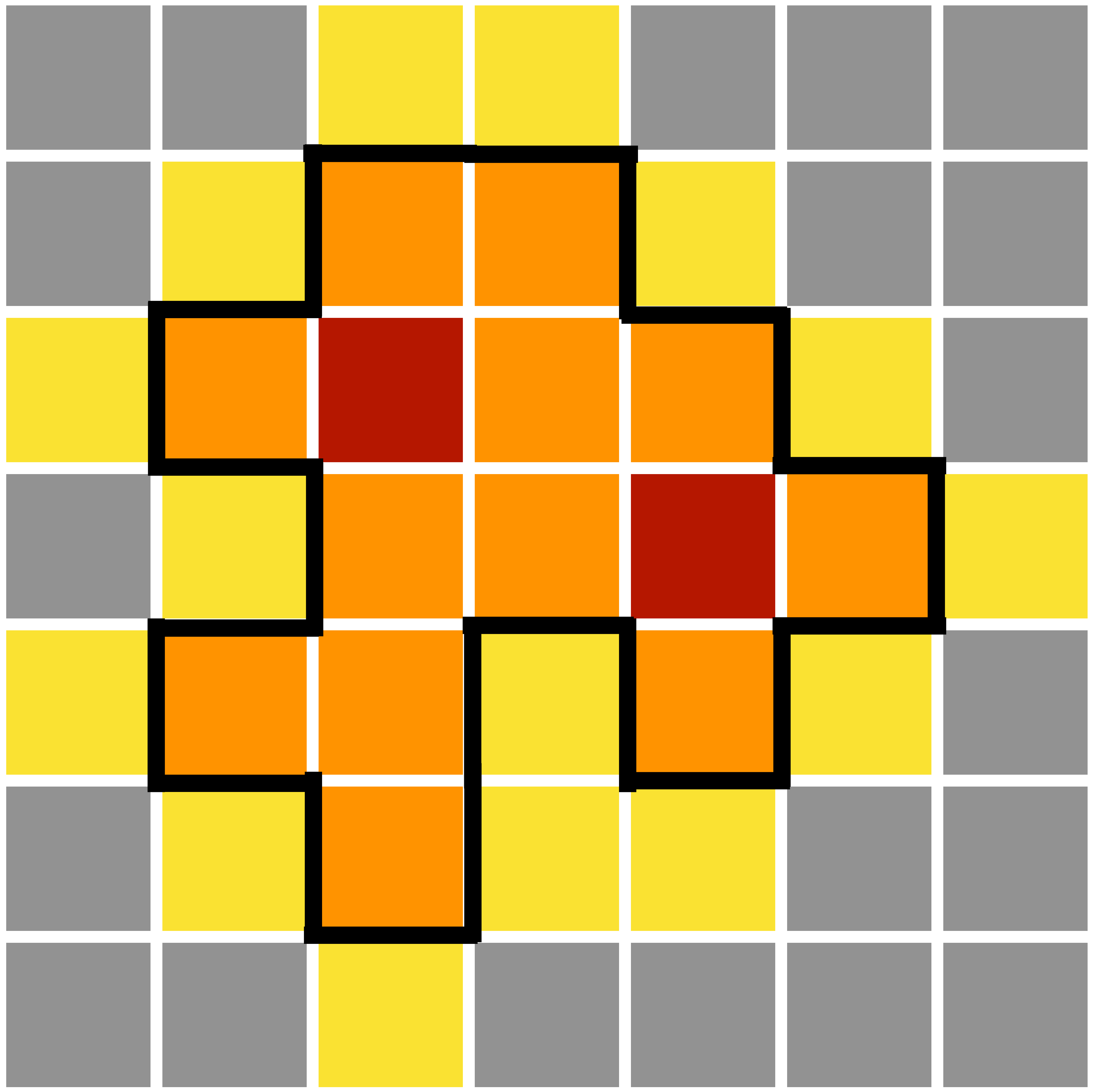
jet inputs: topoclusters

- Use the '4-2-0' algorithm for reconstruction
  - Clusters are seeded by cells with energy of 4σ above the expected noise
  - Any neighboring cells with E>2σ are added recursively until no high-energy neighboring cells remain



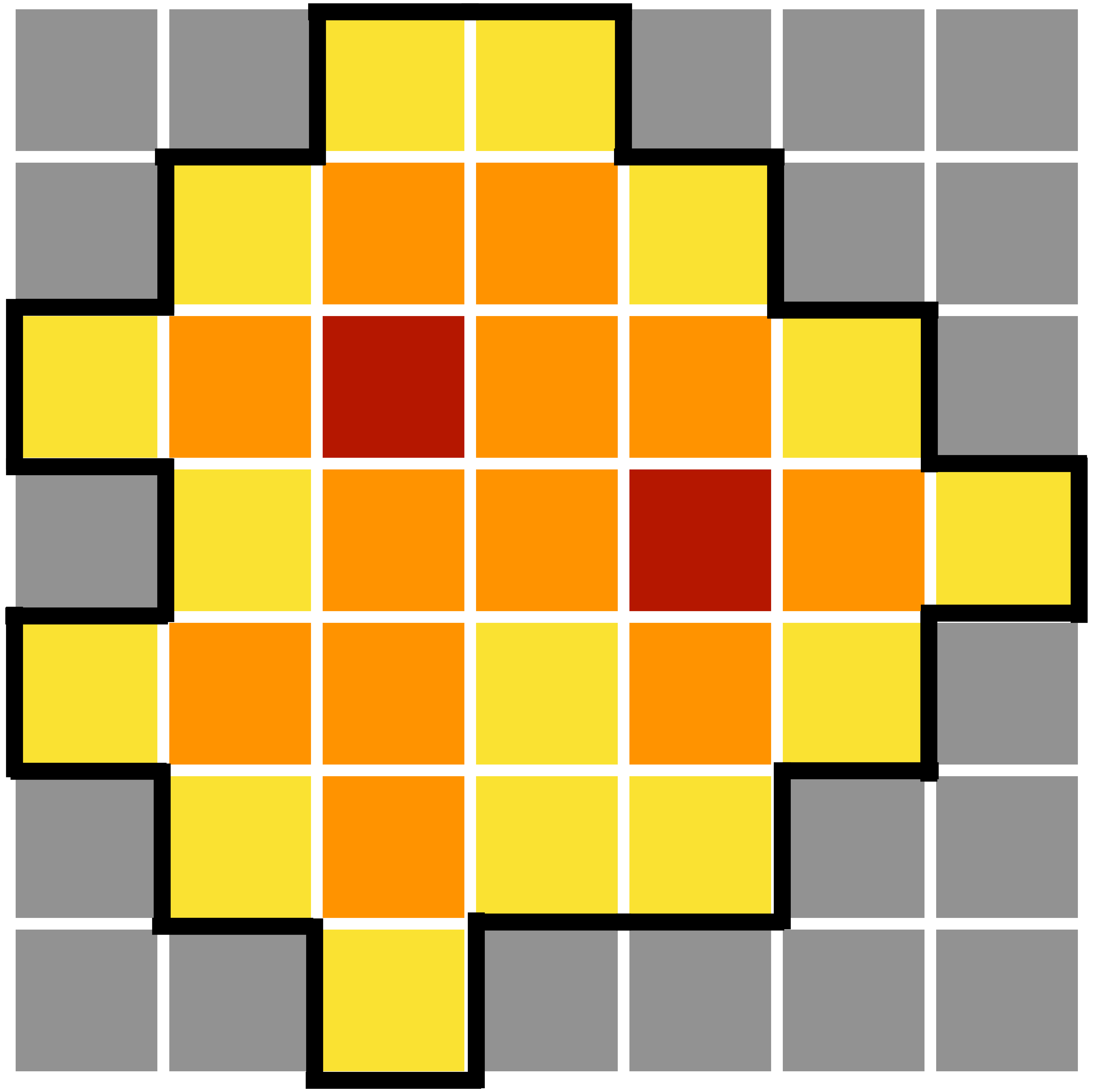
jet inputs: topoclusters

- Use the ‘4-2-0’ algorithm for reconstruction
    - Clusters are seeded by cells with energy of 4σ above the expected noise
    - Any neighboring cells with E>2σ are added recursively until no high-energy neighboring cells remain



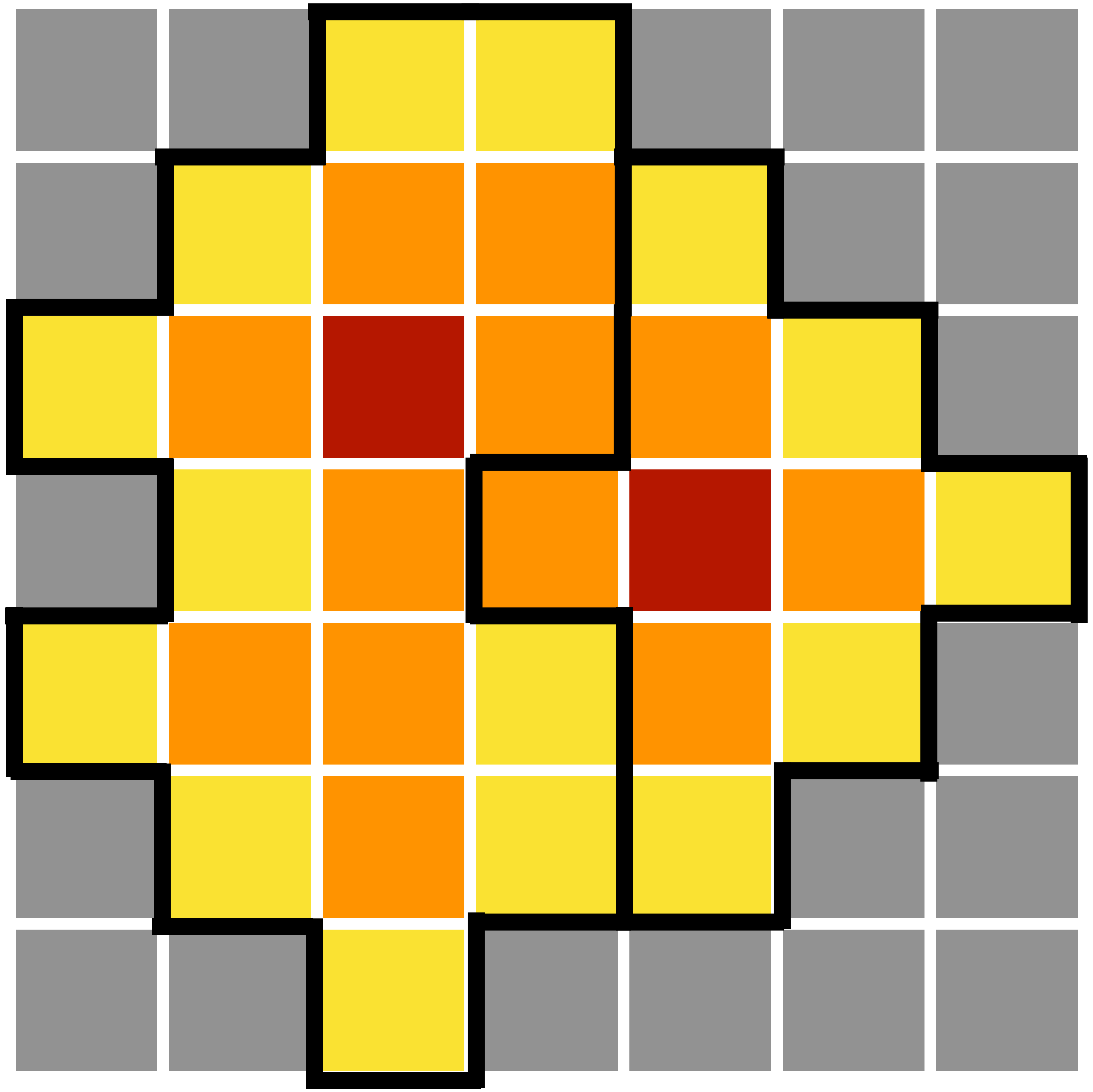
jet inputs: topoclusters

- Use the '4-2-0' algorithm for reconstruction
  - Clusters are seeded by cells with energy of 4σ above the expected noise
  - Any neighboring cells with E>2σ are added recursively until no high-energy neighboring cells remain
  - All neighboring cells are added, regardless of their energy



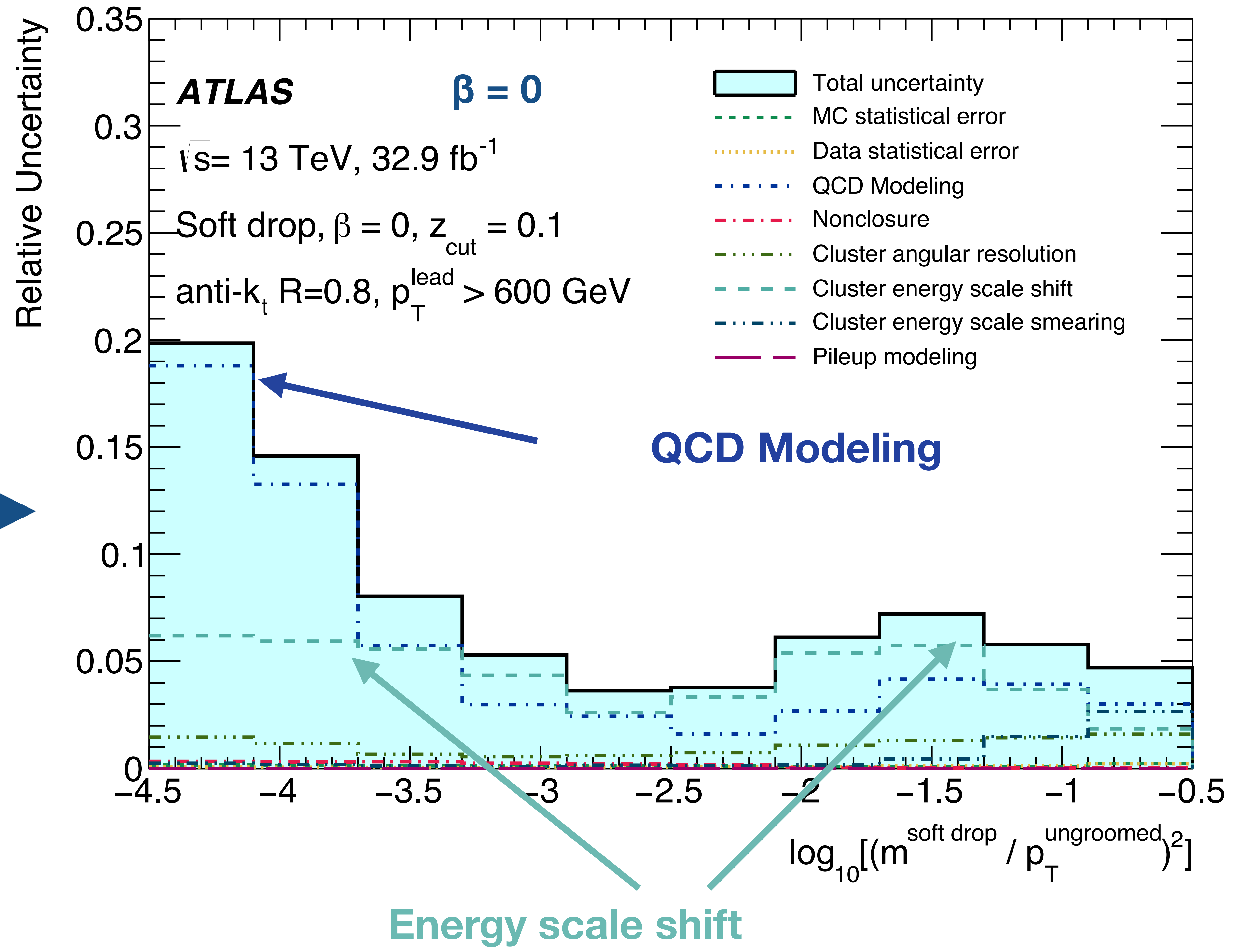
jet inputs: topoclusters

- Use the '4-2-0' algorithm for reconstruction
  - Clusters are seeded by cells with energy of 4σ above the expected noise
  - Any neighboring cells with E>2σ are added recursively until no high-energy neighboring cells remain
  - All neighboring cells are added, regardless of their energy
  - Clusters with multiple local maxima are split



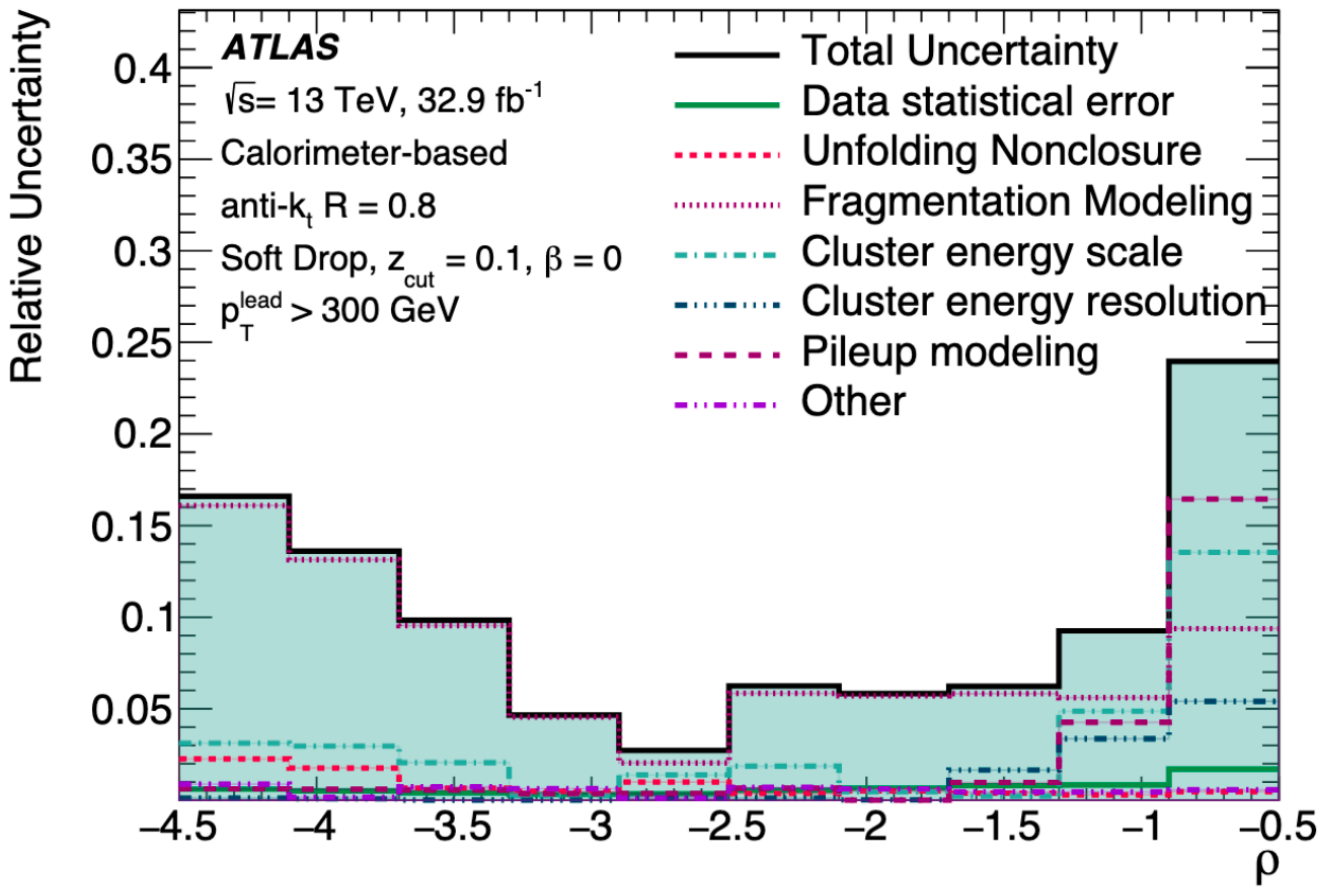
uncertainties

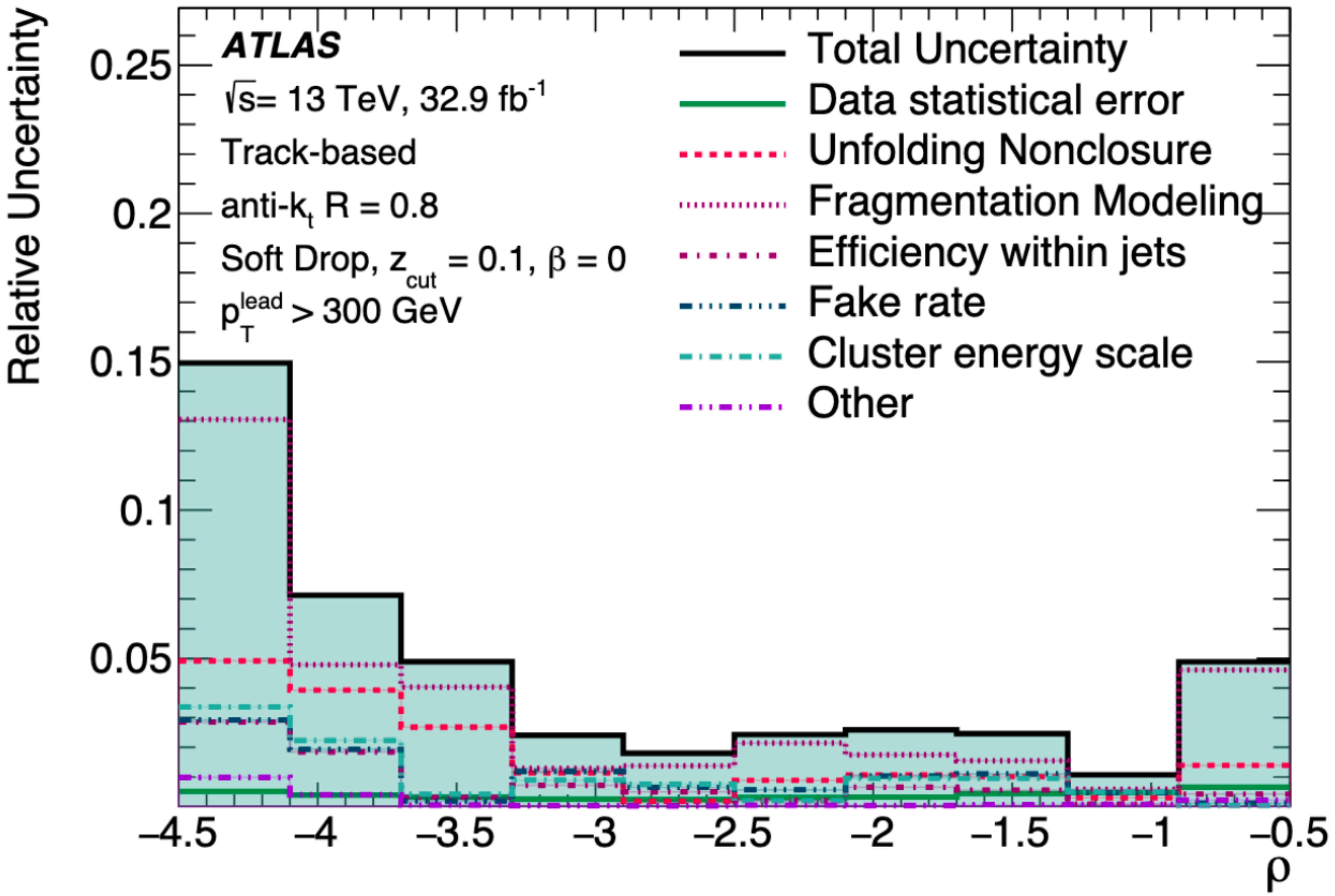
- **QCD modeling** uncertainties dominate, especially in the non-perturbative region
- **Cluster energy scale shift** uncertainty large at lower masses where there are few clusters per jet
- **Cluster energy scale smearing** and **cluster energy scale shift** become more important at higher masses where the energy of hard prongs dominates



uncertainties

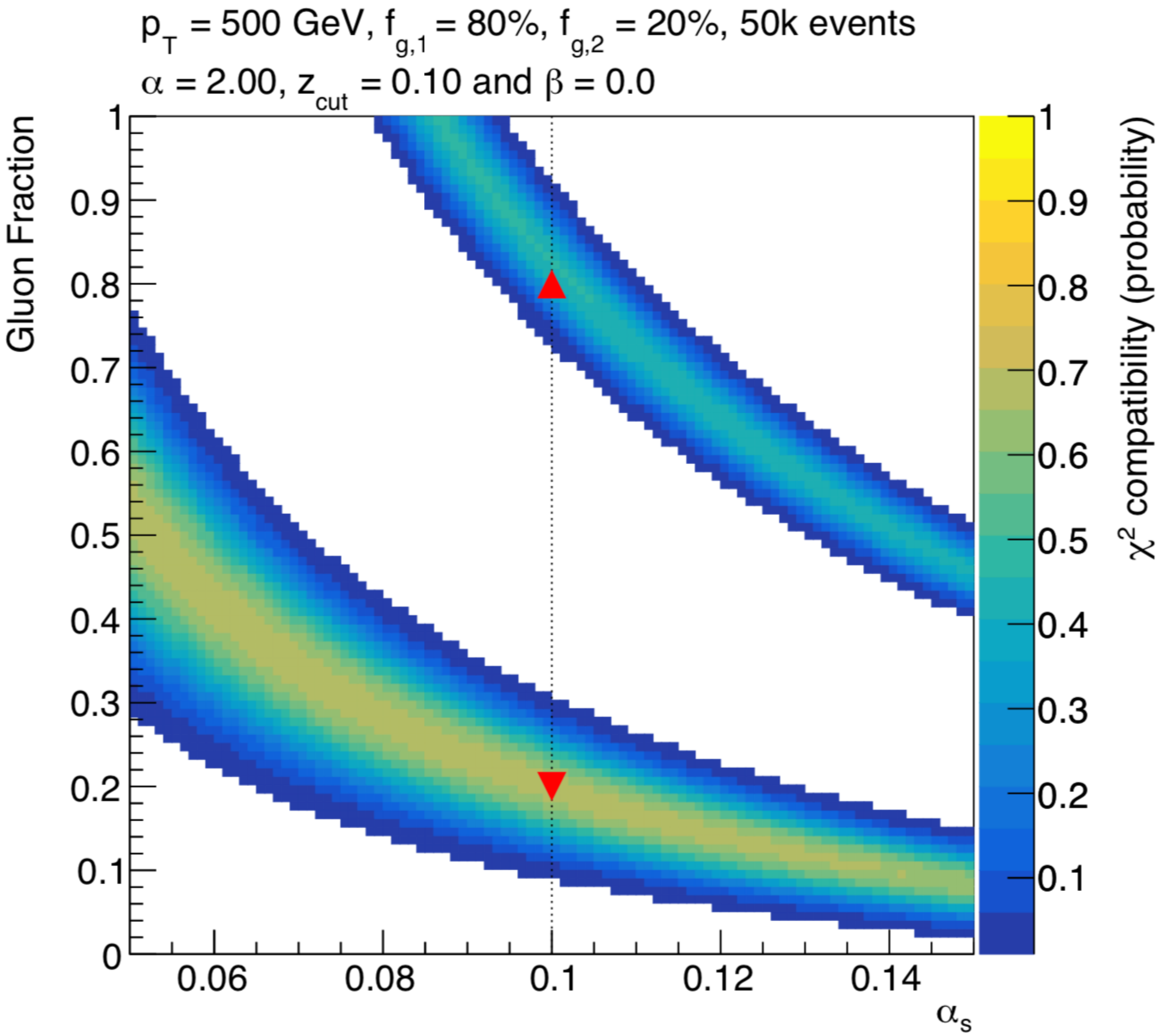
- Tracking uncertainties are typically smaller than cluster uncertainties
  - This is more apparent for larger values of beta, where more soft radiation is included in the jet
- Fragmentation modeling uncertainty is smaller for track-based observable, which is related to the smaller migrations in the response matrix





looking forward: α_s

- Differential cross section for jet mass is proportional to α_s x C_i in the resummation region
- Measurements of mass with multiple samples with different quark/gluon fractions could be used to extract α_s
- May be able to get somewhere around 5-10% precision



Les Houches 2017

- Not competitive with precise measurements, could be used to better understand discrepancies between existing measurements

- Also could provide measurement of running of α_s