Background Subtraction and Fluctuations Update (Long Version)

Anthony Hodges, Christal Martin, David Stewart, Muhammad Khan, Sookhyun Lee, Tanner Mengel

Reconstruction Details

- 10 + 30 GeV dijet samples.
  - Event selection based on leading R=0.4 anti-k_T truth jet.
- Kinematic cuts on reconstructed jets:
  - $p_T^{reco} > 5$ GeV
- Kinematic cuts on truth jets:
  - $10.0 < p_T^{truth} < 100$ GeV
  - $|\eta_{truth}| < 1.1. - R$
- Reconstructed jets using TOWERINFO containers.
  - CEMC_TOWERINFO_RETOWER, HCALIN_TOWERINFO, HCALOUT_TOWERINFO.
- Centrality determined using HIJANG centrality.
  - Needs to be modified for data embedding. Changes added
- Available code JetBkgdSub
  - Produces TTree with all subtracted $p_T^{Area}, p_T^{Mult}, p_T^{Iter.}$, and p_T^{truth} as well as η and ϕ.
  - Added multiplicity curves for R = 0.3 and R = 0.5.
  - Benjamin pointed out an error in the un-subtracted $p_T^{Iter.}$. Implemented Virginia's fix.
  - Reco cut was being performed on raw jet momentum, thanks to Benjamin for pointing this out. Fixed.
  - Adding switch for simulation vs. data for embedding.

Matching Details

- Using event weights from wiki.
  - p+p no pileup sample had half as many events as the embedded sample. Lead to lower jet yield.
    - Chris added 10 million events to this sample to make the comparison even
- Kinematic cuts on matched reco jets:
  - $p_T^{reco} > 5$ GeV
  - $\Delta R < 0.75R$
- Kinematic cuts on truth jets:
  - $10.0 < p_T^{truth} < 100$ GeV
- Available code will be added to JetBkgdSub/Offline
  - Macro for matching all jet samples for each subtraction type
  - More plotting code will be added (need to switch from python to root)

Analysis Status

- Jet Energy Scale
  - Area
  - Iterative
  - Multiplicity
  - p+p (no embedding)
- Jet Energy Resolution
  - Area
  - Iterative
  - Multiplicity
  - p+p (no embedding)
- Reconstructed jet spectra
  - Area
  - Iterative
  - Multiplicity

- Missed/Fake Ratios
  - Area
  - Multiplicity
  - Iterative
- Method Plots
  - Area:
    - Rho
    - Subtracted pT
  - Multiplicity:
    - Rho
    - N signal
    - Subtracted pT
  - Iterative:
    - Subtracted pT

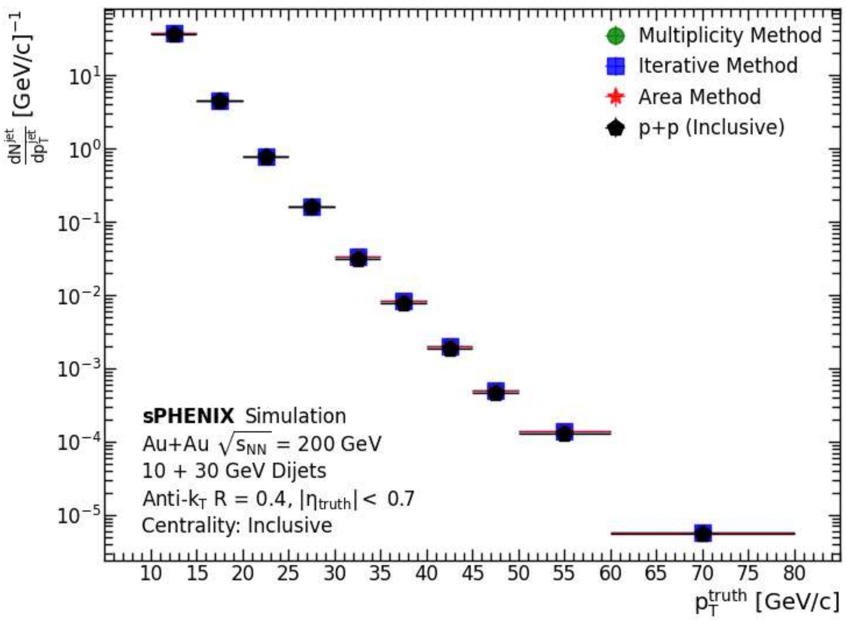
- Background fluctuations
  - Area
  - Iterative
  - Multiplicity

TODO

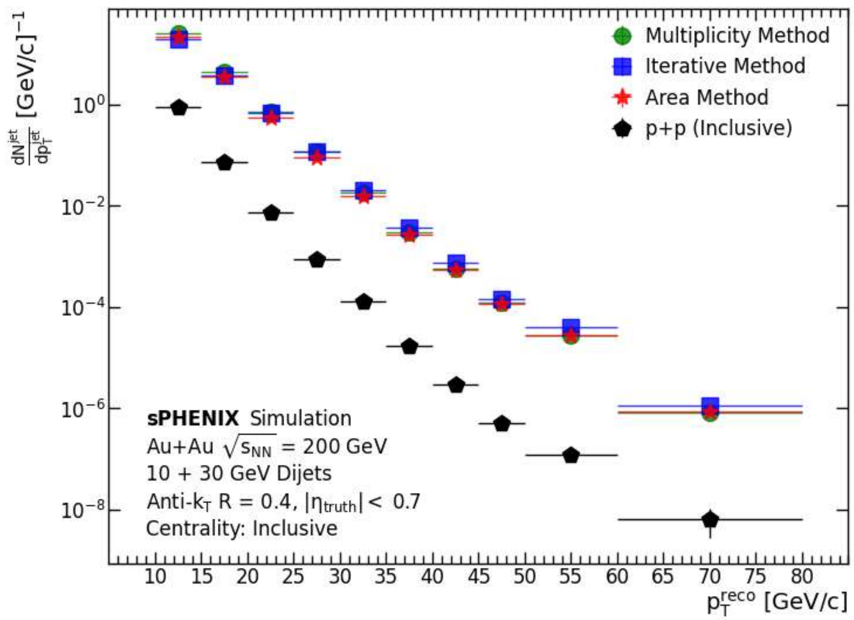
- Random cone analysis for each method to quantify background fluctuations
- Upload plotting code to GitHub.
- Work with Benjamin for results in data
- Run iterative method with EPD event plane.

Reconstructed Spectra

- Inclusive reconstructed spectra (0-100% central)
- Truth spectra for p+p without embedding agree with truth from embedded samples.
- Discrepancies between reco pp spectra and reco AuAu spectra are from background fluctuations.



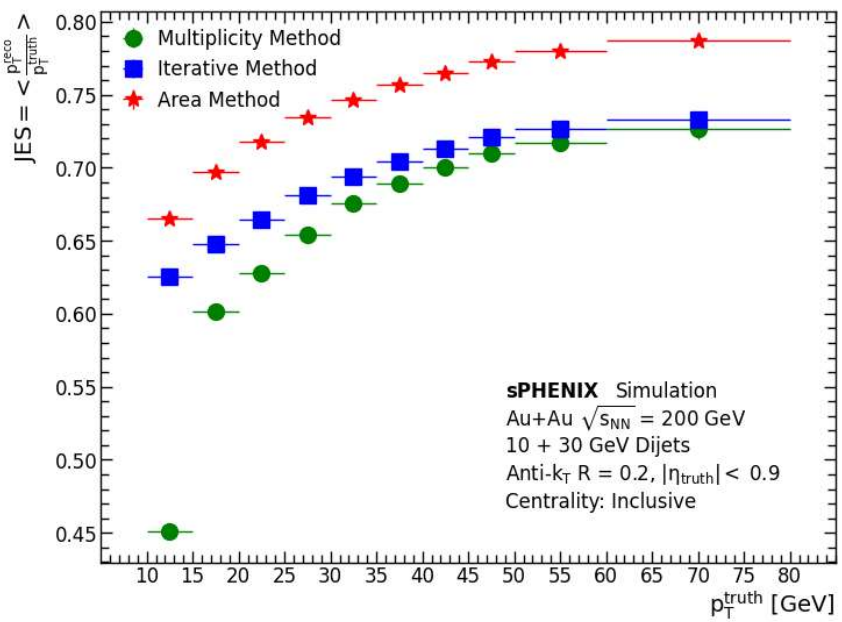
Comparisons of truth yield for R = 0.4 jets



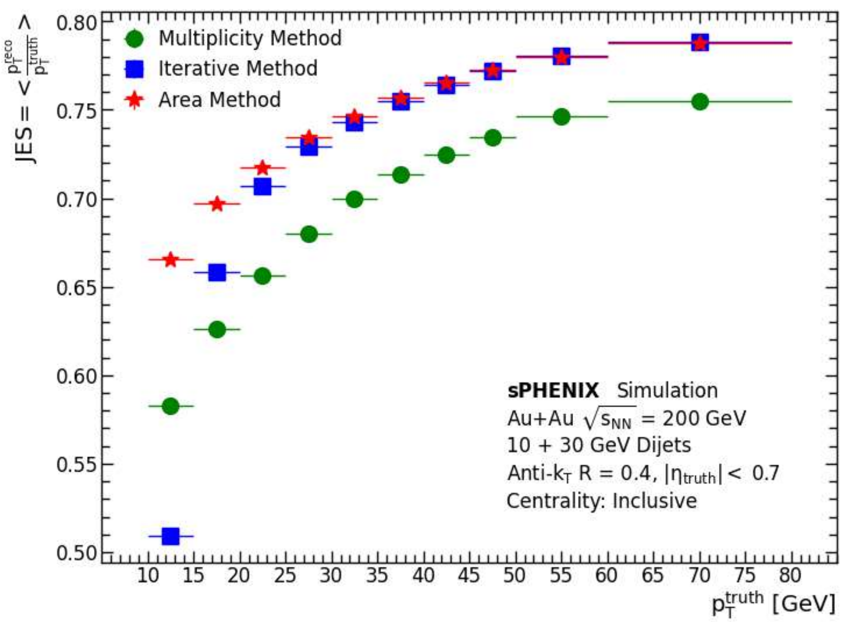
Comparisons of Reco yield for R = 0.4 jets

Jet energy scale

- $JES \equiv \left\langle \frac{p_T^{reco}}{p_T^{truth}} \right\rangle$
- Area method performs best over entire centrality range for both R = 0.2 and R = 0.4 jets.



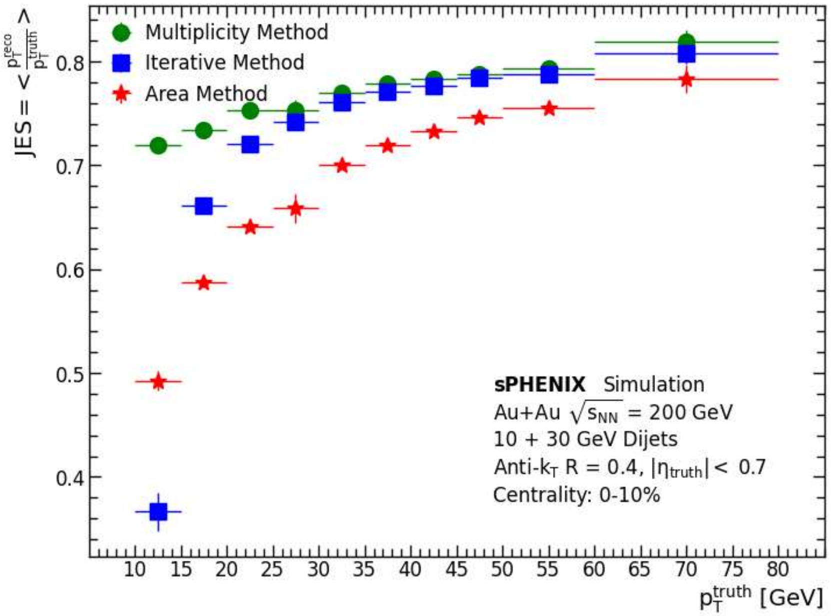
Comparisons of JES for all centralities for R = 0.2 jets



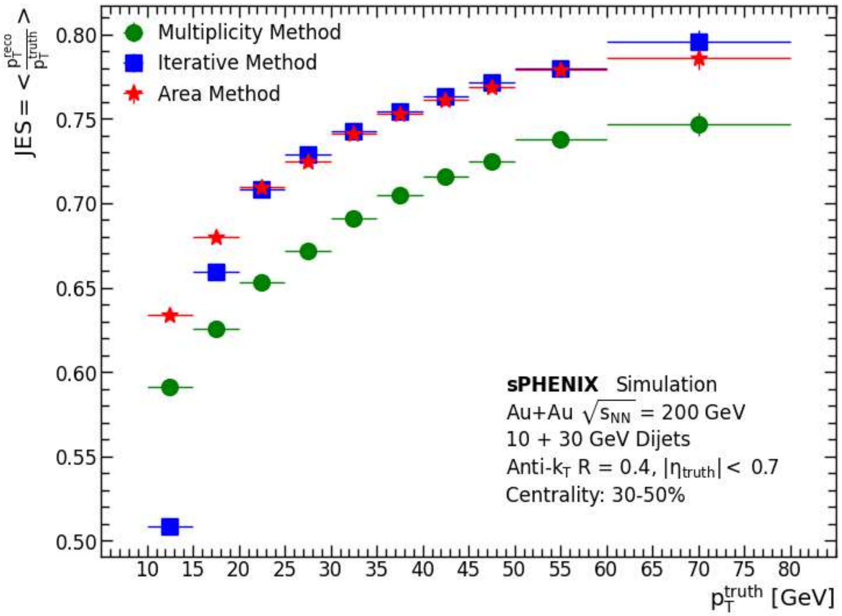
Comparisons of JES for all centralities for R = 0.4 jets

JES in different centralities

- Area and multiplicity methods vary in performance from central to peripheral events.
- Multiplicity performs best in central.
- Iterative method falls off steeply at low jet momentum.



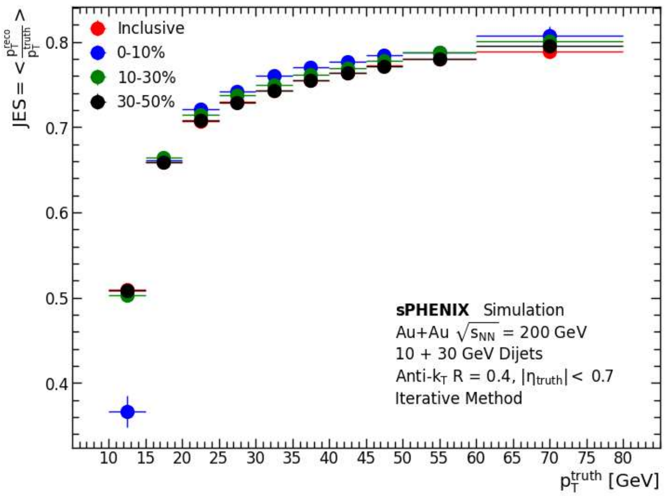
Comparisons of JES for central events for R = 0.4 jets



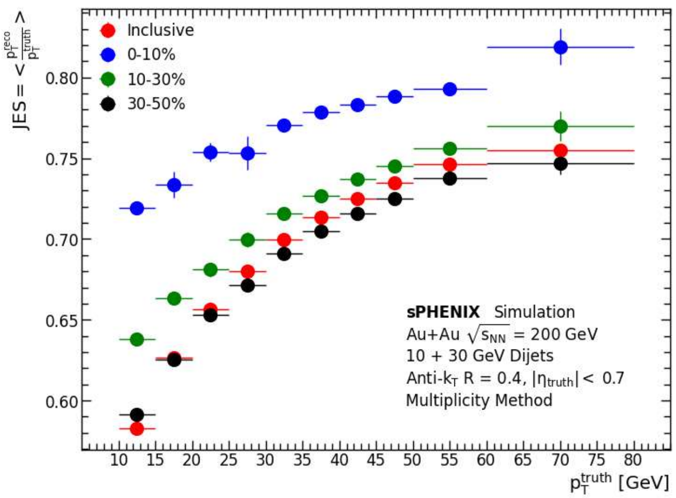
Comparisons of JES for semi-peripheral events for R = 0.4 jets

Centrality dependence on JES for each method

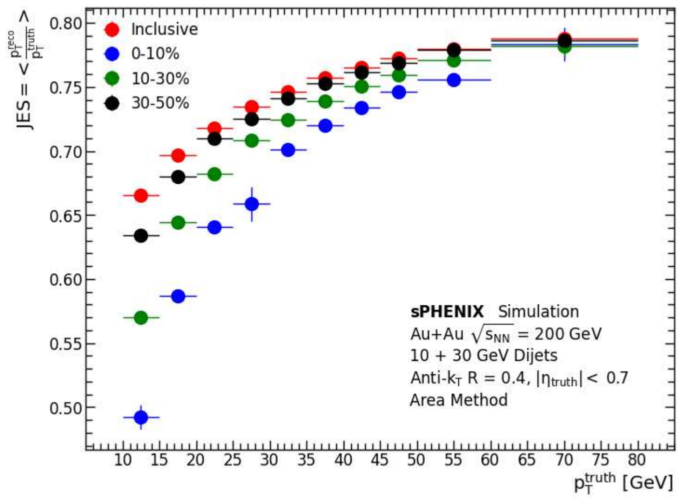
- Iterative method is constant across event centrality.
- Area and multiplicity have opposite behavior from central to peripheral events.



Comparisons of JES across centrality for iterative method



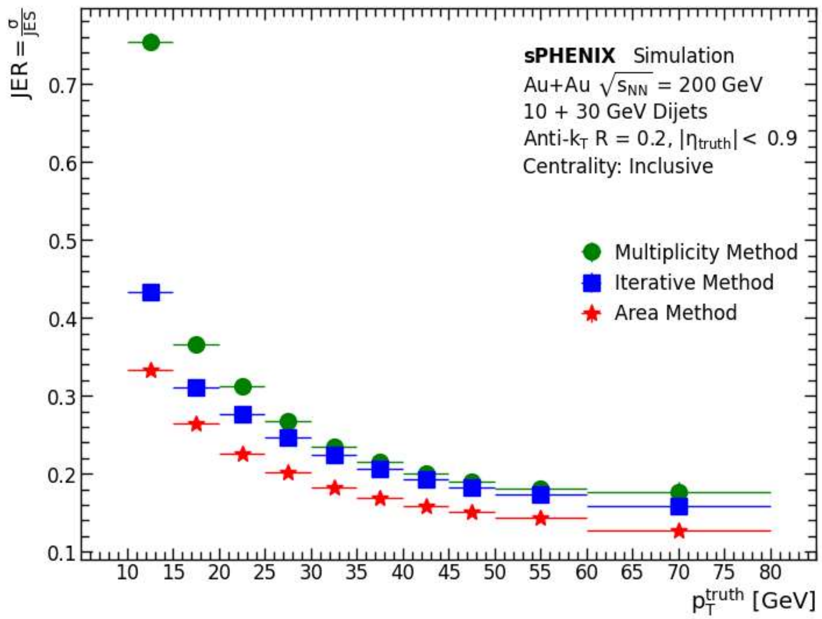
Comparisons of JES across centrality for multiplicity method



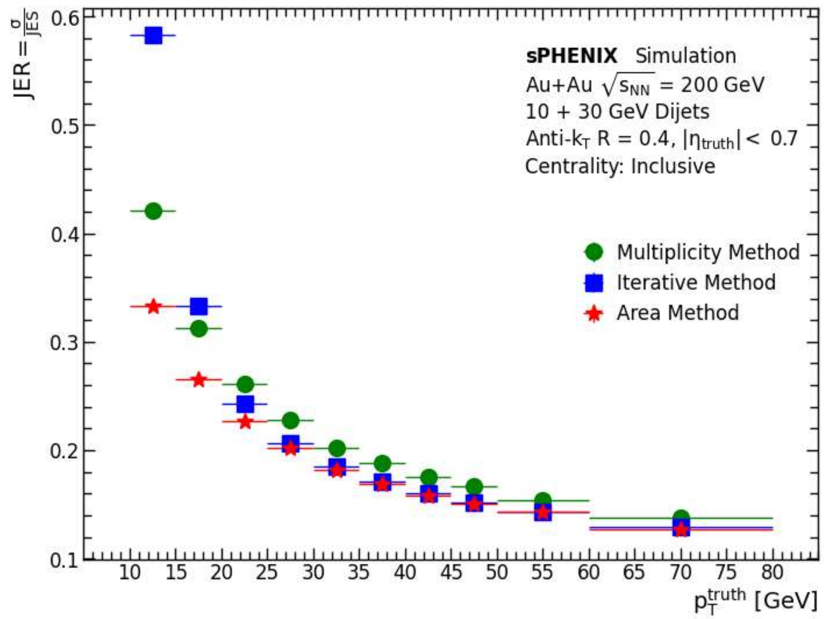
Comparisons of JES across centrality for area method

Jet energy resolution

- $JER \equiv \sigma\left(\frac{p_T^{reco}}{p_T^{truth}}\right) \Big/ \left\langle \frac{p_T^{reco}}{p_T^{truth}} \right\rangle$
- Area method has lowest inclusive JER at low momentum and is comparable to iterative method at high momentum in R =0.4 jets.



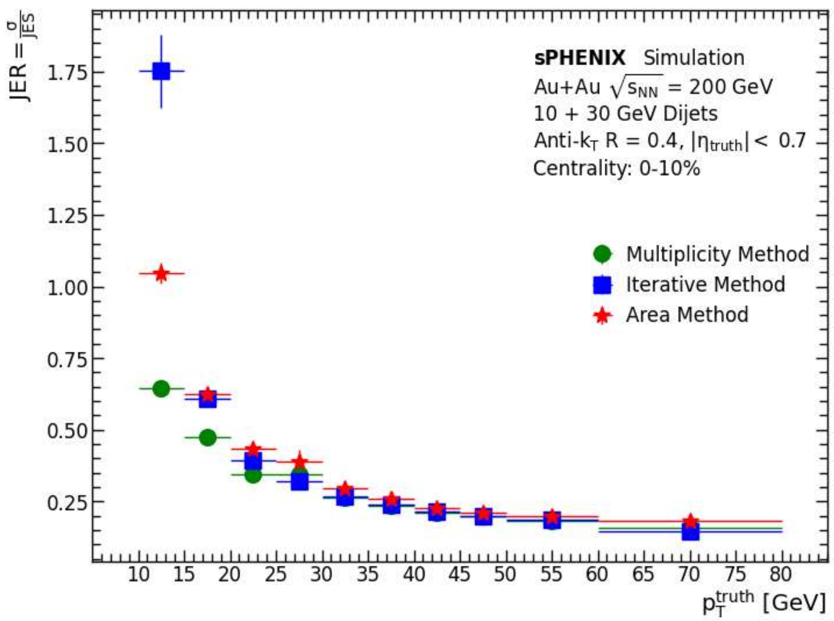
Comparisons of JER for all centralities for R = 0.2 jets



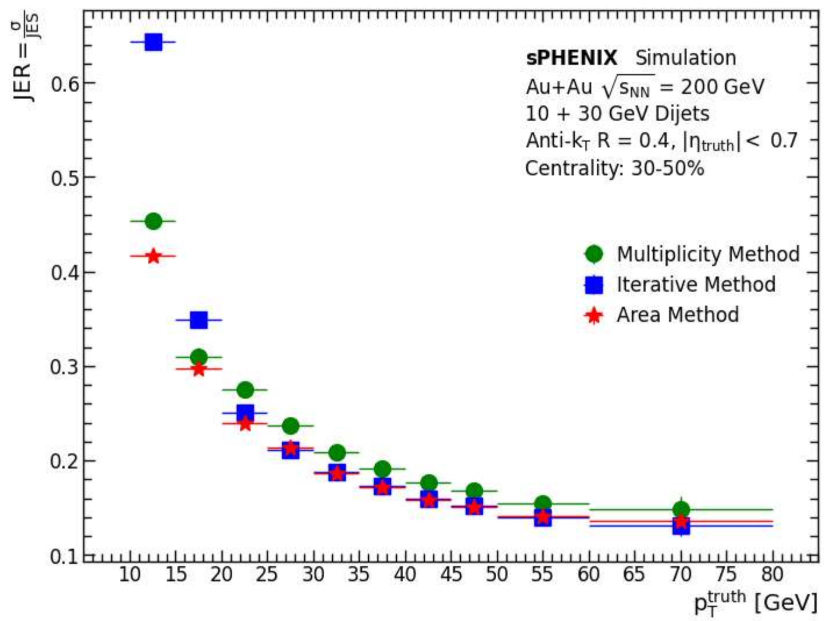
Comparisons of JER for all centralities for R = 0.4 jets

JER in different centralities

- Area and multiplicity methods vary in performance from central to peripheral events.
- Multiplicity performs best in central.
- Area method better/comparable to iterative method at all jet momentum.
- Same behavior where iterative method falls off steeply at low jet momentum.



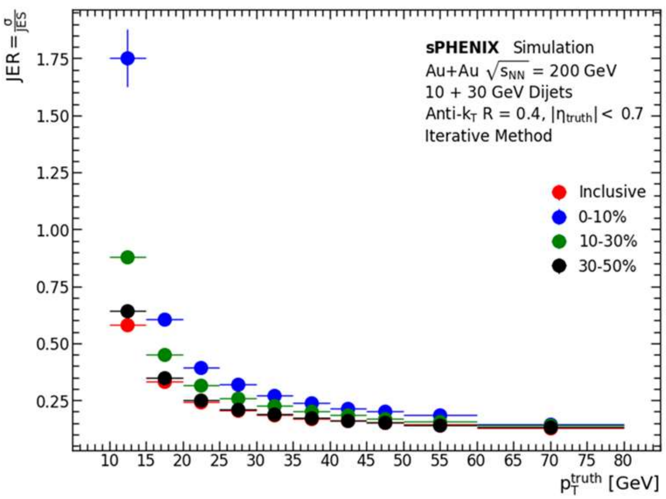
Comparisons of JER for central events for R = 0.4 jets



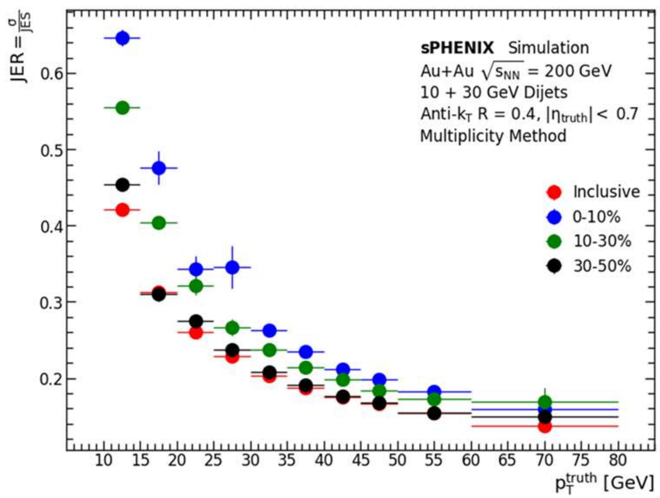
Comparisons of JER for semi-peripheral events for R = 0.4 jets

Centrality dependence on JER for each method

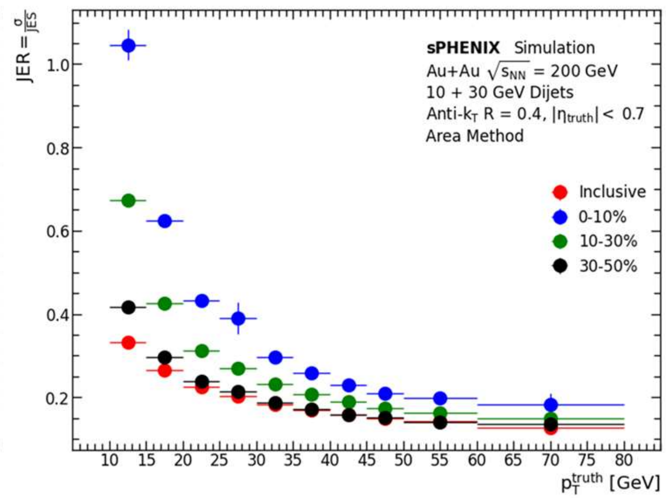
- Iterative method is constant across event centrality.
- Area and multiplicity have opposite behavior from central to peripheral events.
- Multiplicity method performs best for low momentum jets compared to other methods.



Comparisons of JER across centrality for iterative method



Comparisons of JER across centrality for multiplicity method



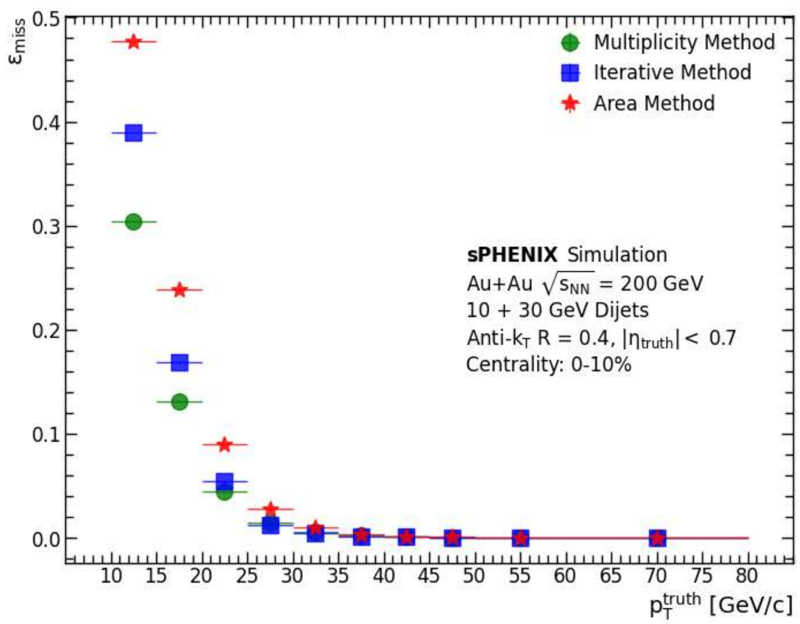
Comparisons of JER across centrality for area method

Fakes and Misses

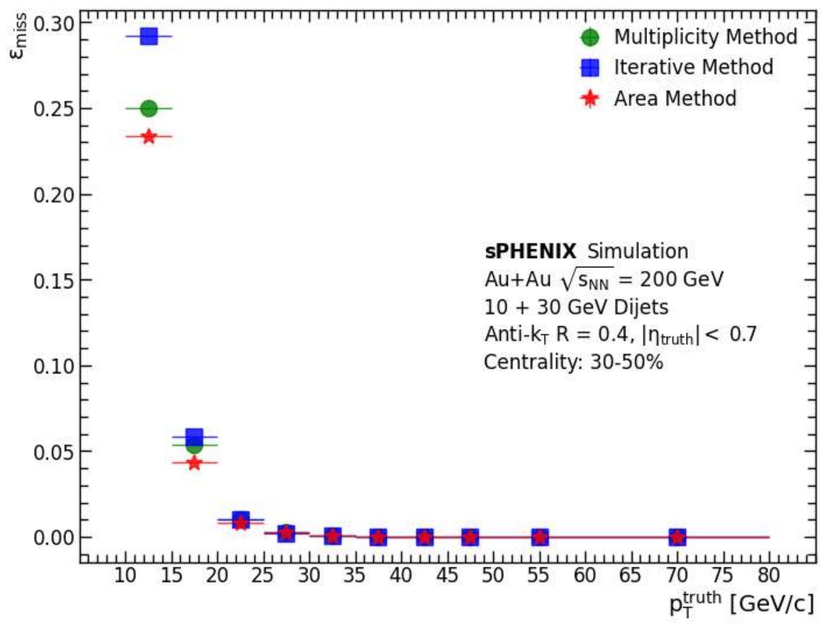
- Gauge each methods ability to suppress combinatorial jets while not losing many truth jets.
- Truth missing efficiency:
  - $\varepsilon_m = \left. \frac{dN^{miss}}{dp_T} \middle/ \frac{dN^{total\ truth}}{dp_T} \right.$
- Unmatched reco efficiency (proxy for combinatorial jets) :
  - $\varepsilon_f = \left. \frac{dN^{unmatched}}{dp_T} \middle/ \frac{dN^{total\ reco}}{dp_T} \right.$
- Lower ratios lead to easier unfolding and extend kinematic range of measurement

ε_m in different centralities

- Multiplicity has lowest ratio for central events
- Area has highest miss ratio at low pT in central events
- Iterative has highest ratio in semi-peripheral events.



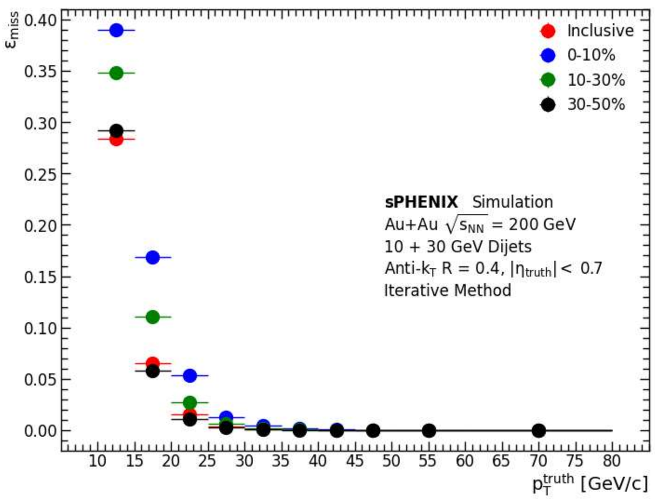
Comparisons of ε_m for central events for R = 0.4 jets



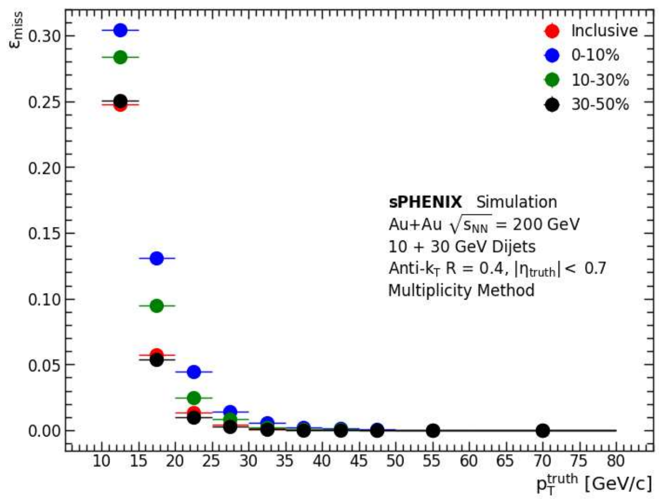
Comparisons of ε_m for semi-peripheral events for R = 0.4 jets

ε_m centrality dependence

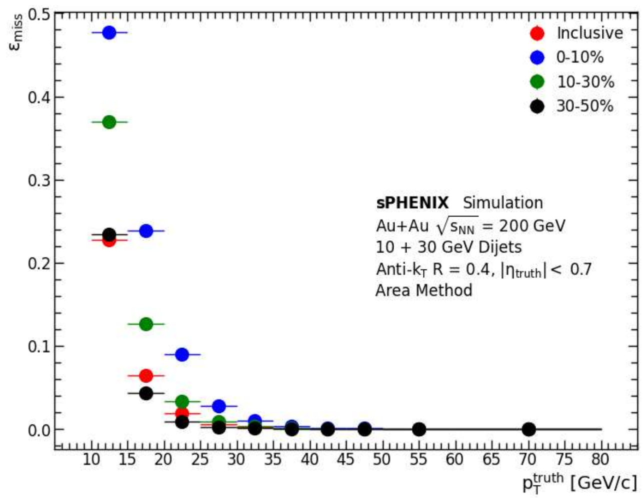
- Multiplicity performers best at low jet momentum in central events
- Area performers best at low jet momentum in semi-peripheral events
- Iterative method has least variation across centrality.



Comparisons of ε_m across centrality for Iterative method



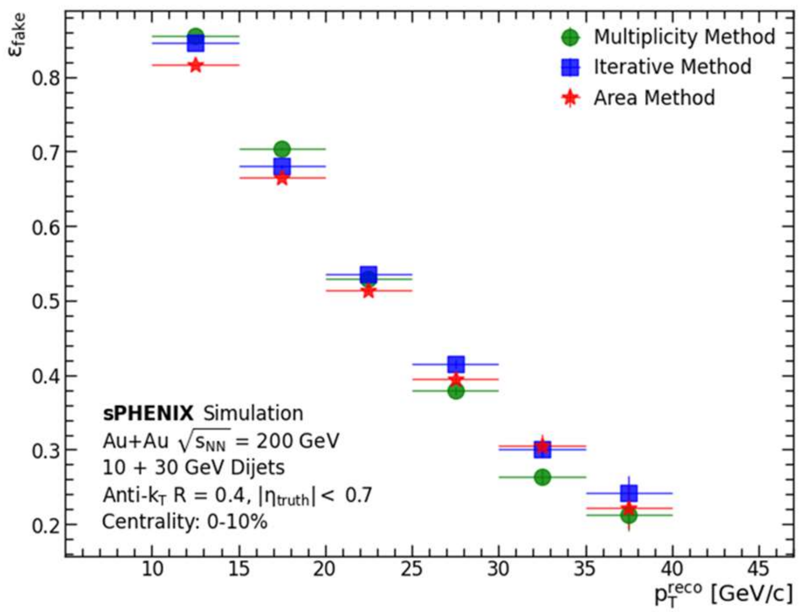
Comparisons of ε_m across centrality for multiplicity method



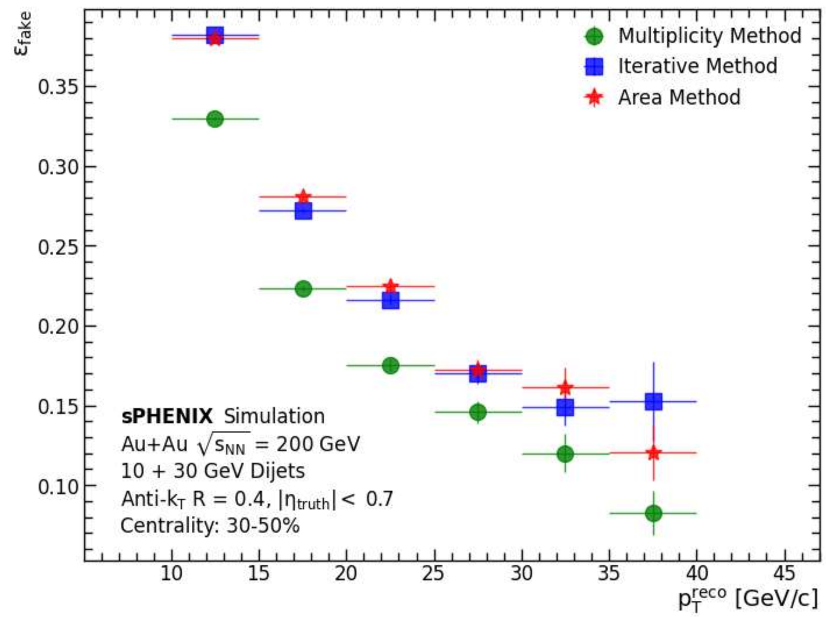
Comparisons of ε_m across centrality for area method

ε_f in different centralities

- Multiplicity method suppresses combinatorial best for most momentum bins
- Area performs better at low momentum in central events
- Iterative is never suppresses better than the area and/or multiplicity method.



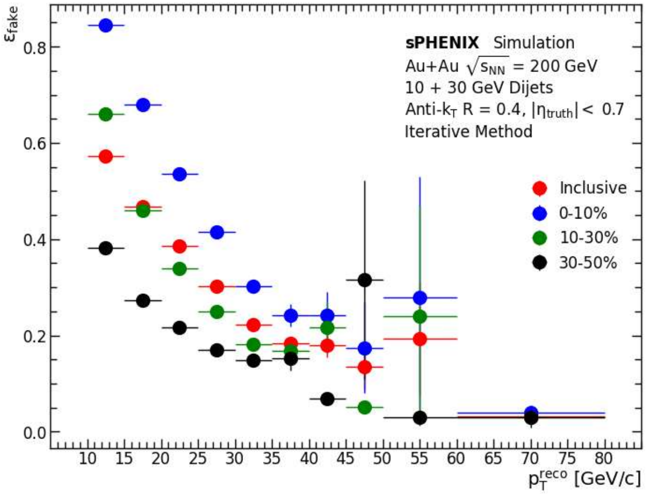
Comparisons of ε_f for central events for R = 0.4 jets



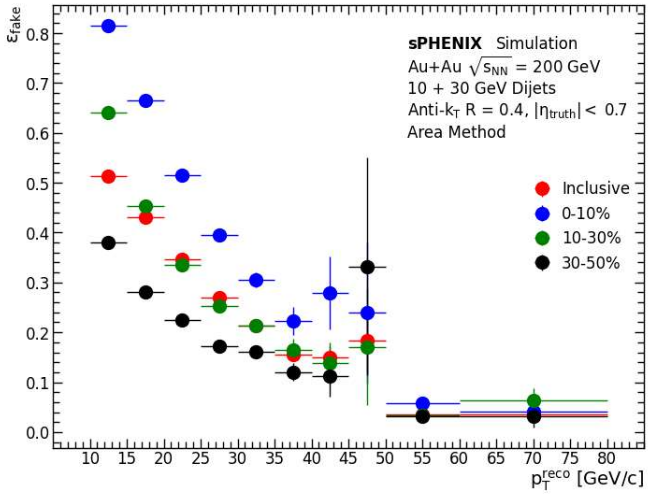
Comparisons of ε_f for semi-peripheral events for R = 0.4 jets

ε_f centrality dependence

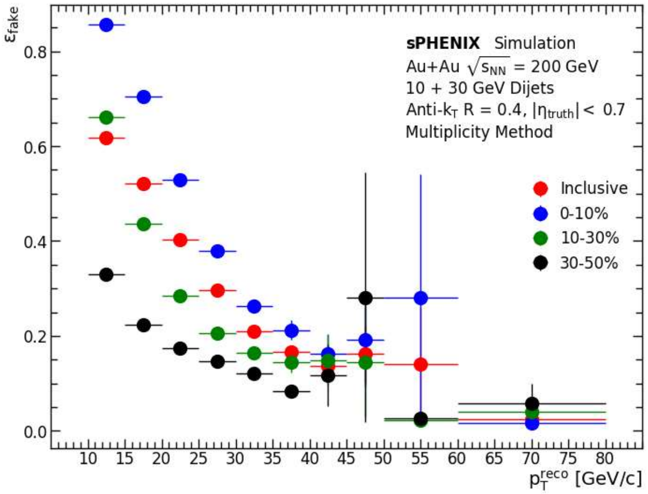
- Similar behavior across centrality for all methods
- Multiplicity has best overall performance (red) especially at low jet momentum



Comparisons of ε_f across centrality for Iterative method



Comparisons of ε_f across centrality for area method



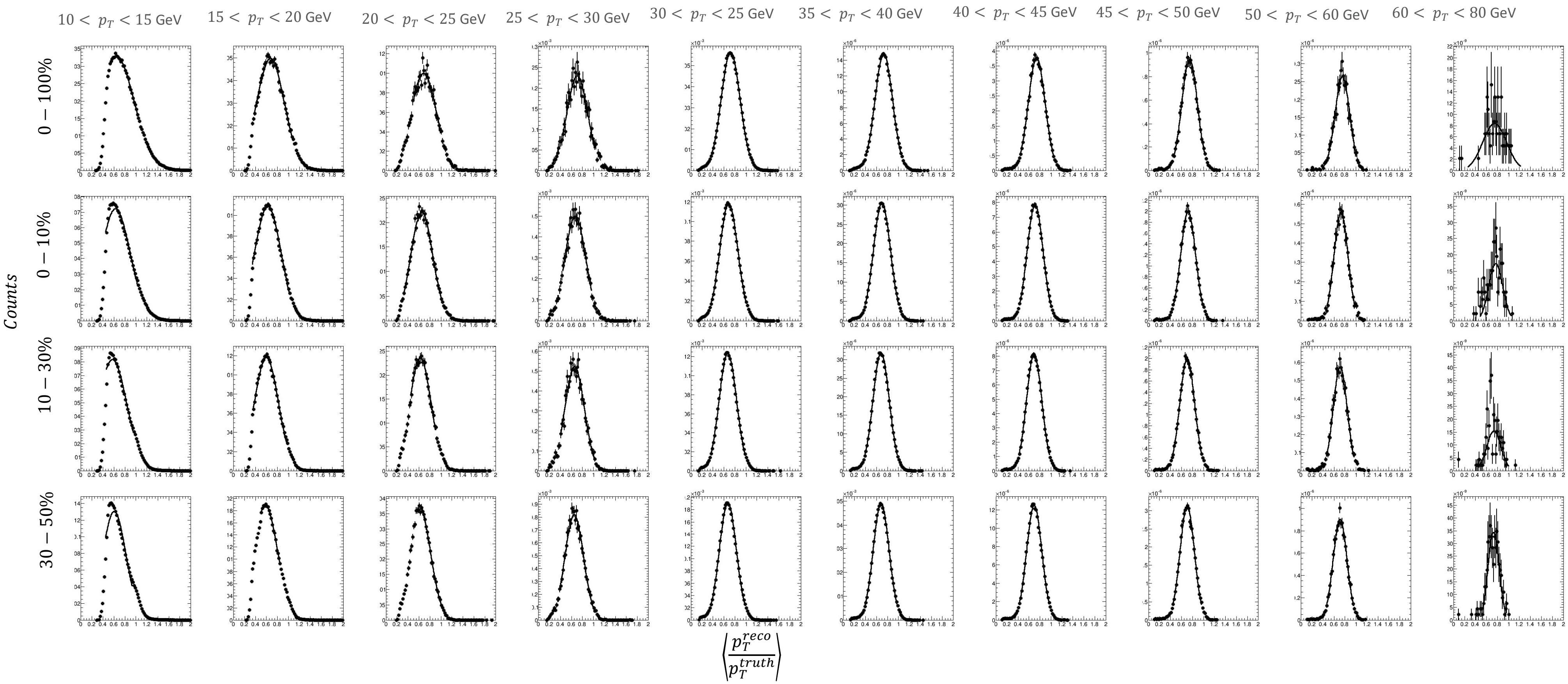
Comparisons of ε_f across centrality for multiplicity method

Conclusions

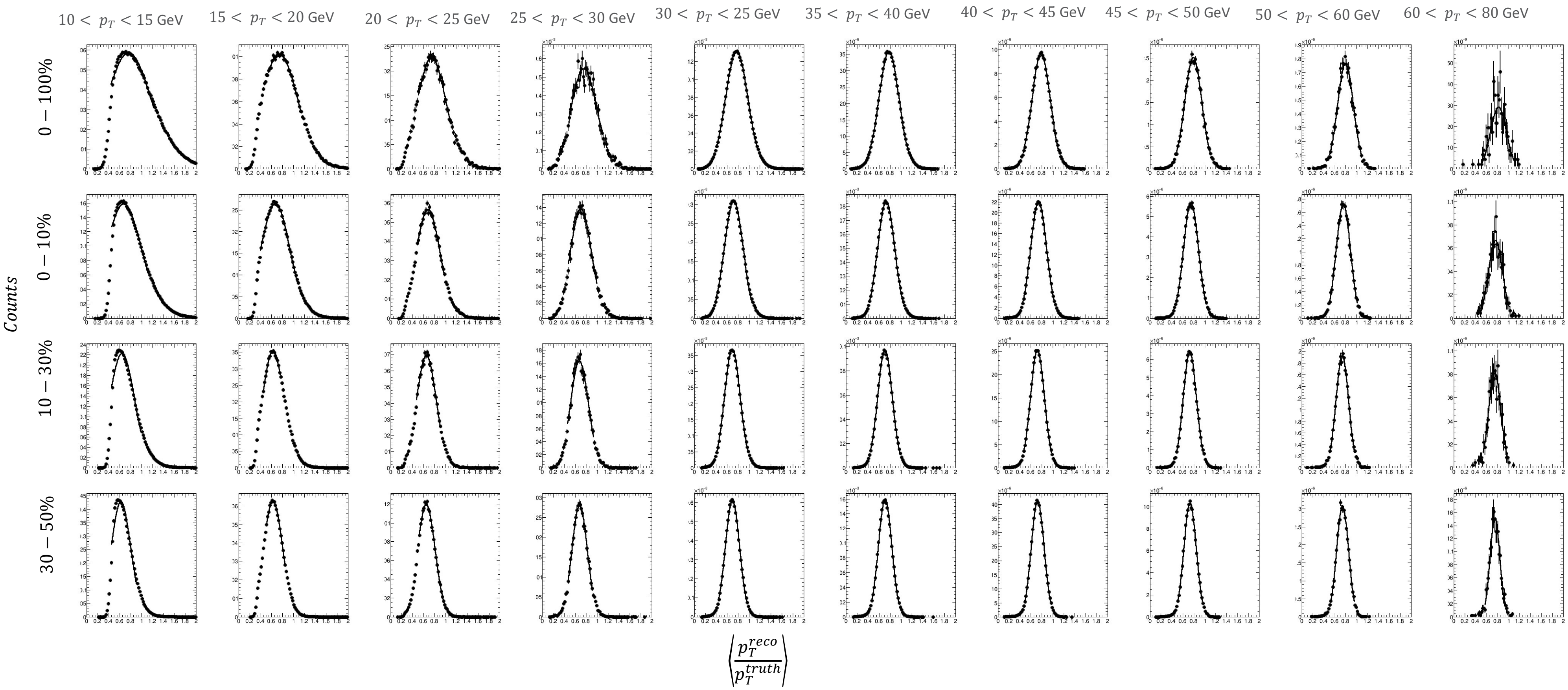
- In almost all pT bins, across all centrality regions the best performing background subtraction method is either the area or multiplicity method.
- The Iterative method has the advantage of being less dependent on centrality but takes a hit on performance.
- The area and multiplicity method have inverse behavior from central to peripheral events
  - Multiplicity method performs best in central events, Area in peripheral events.
- The multiplicity method performs much better at low momentum in both JER and combinatorial suppression in central events
- The area method performs as good/better than the iterative method in more peripheral events.
- Combinatorial jet suppression and JER in the low jet momentum regime should be prioritized given the target kinematic range of the sPHENIX jet program.
- These results do not suggest that there is one clear ‘standard’ method we should adopt for all jet measurements but rather use a background subtraction method that is tailored to a given observable.

JES fit plots

Multiplicity (R = 0.2)

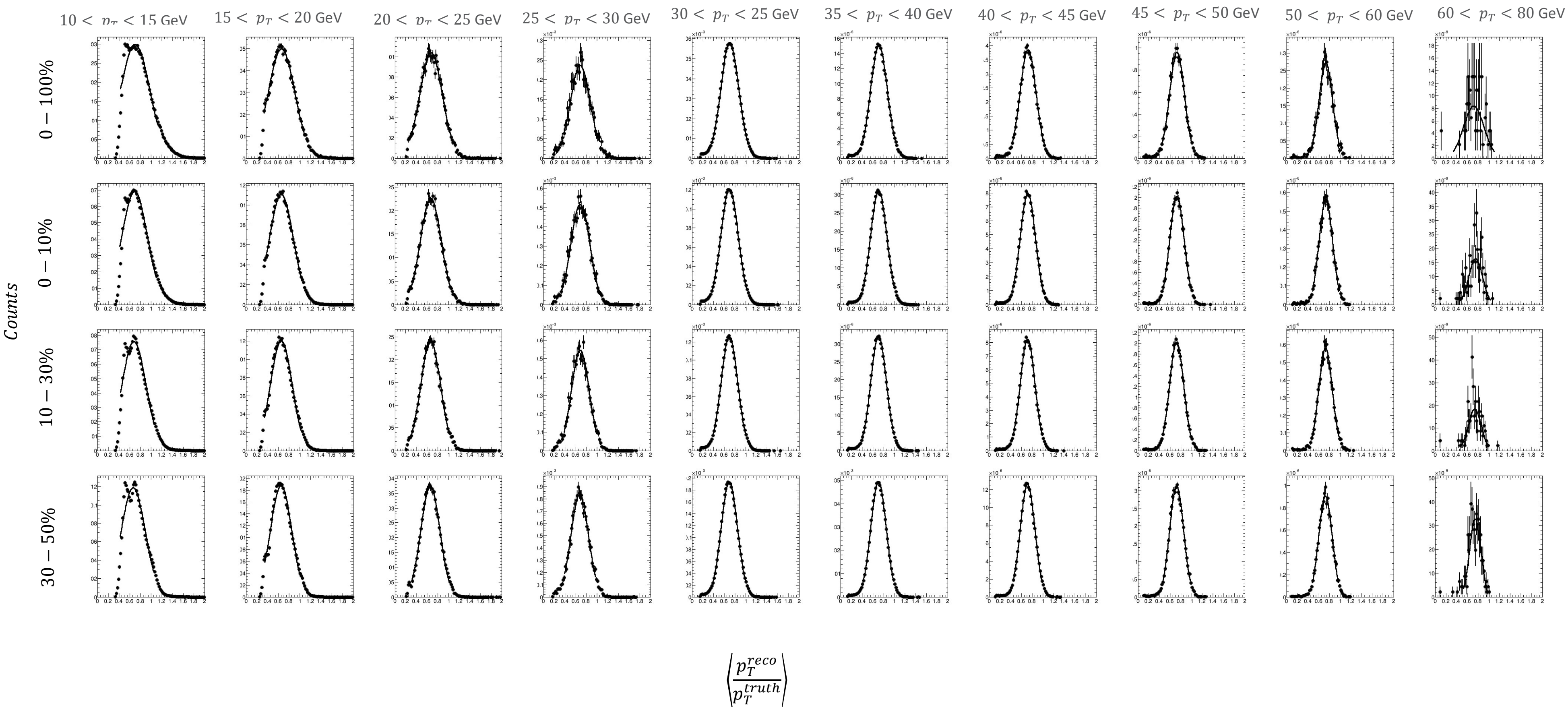


Multiplicity (R = 0.4)



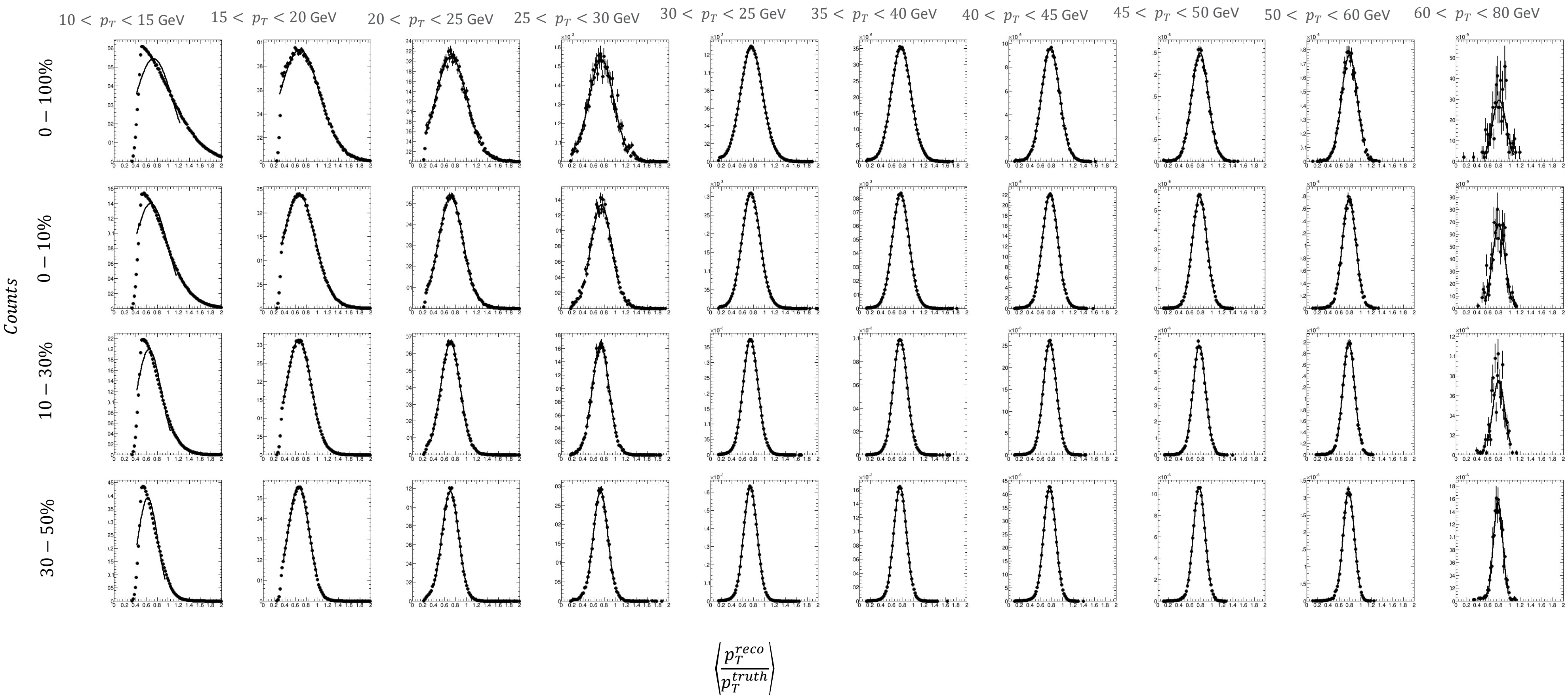
Iterative (R = 0.2)

- Iterative



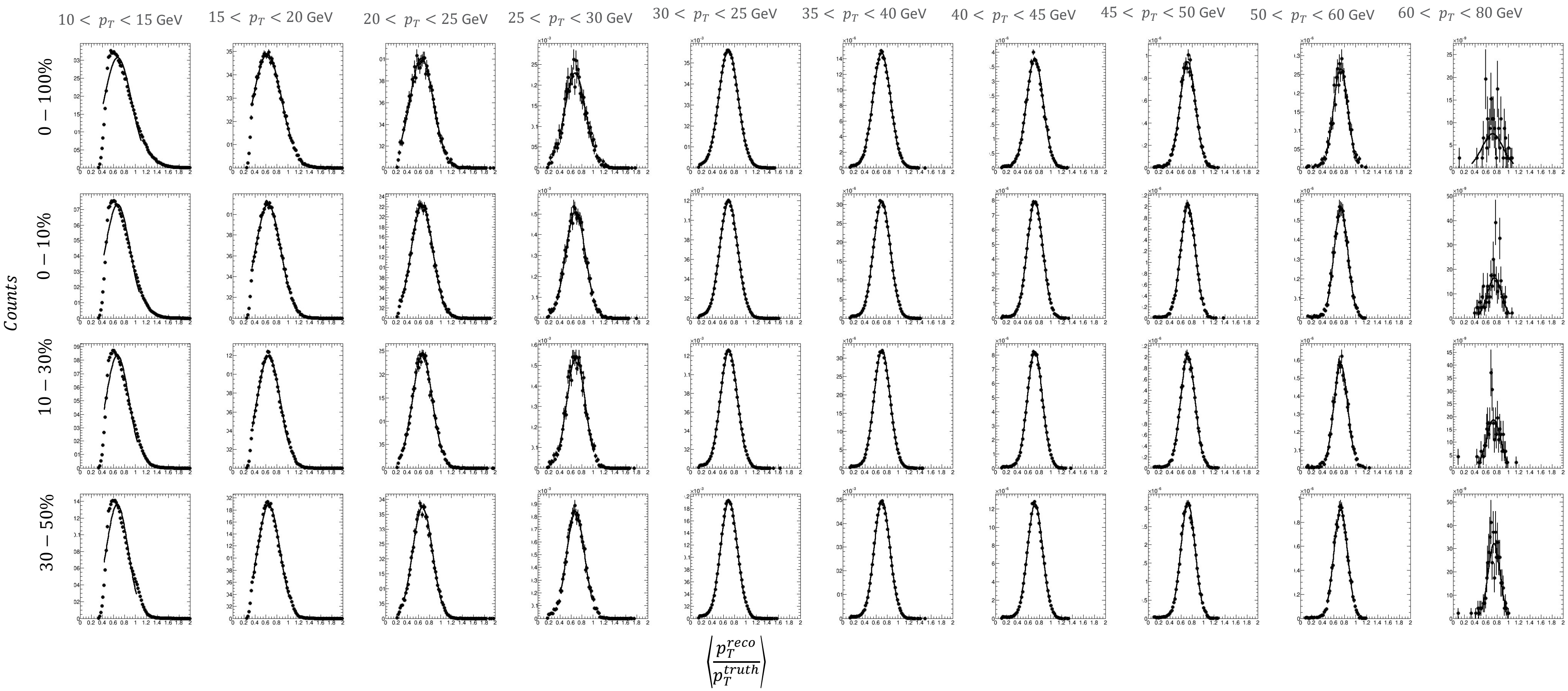
Iterative (R = 0.4)

- Iterative



Area (R = 0.2)

- Area



Area (R = 0.4)

- Area