Jet fragmentation study in ALICE

FKPPN workshop

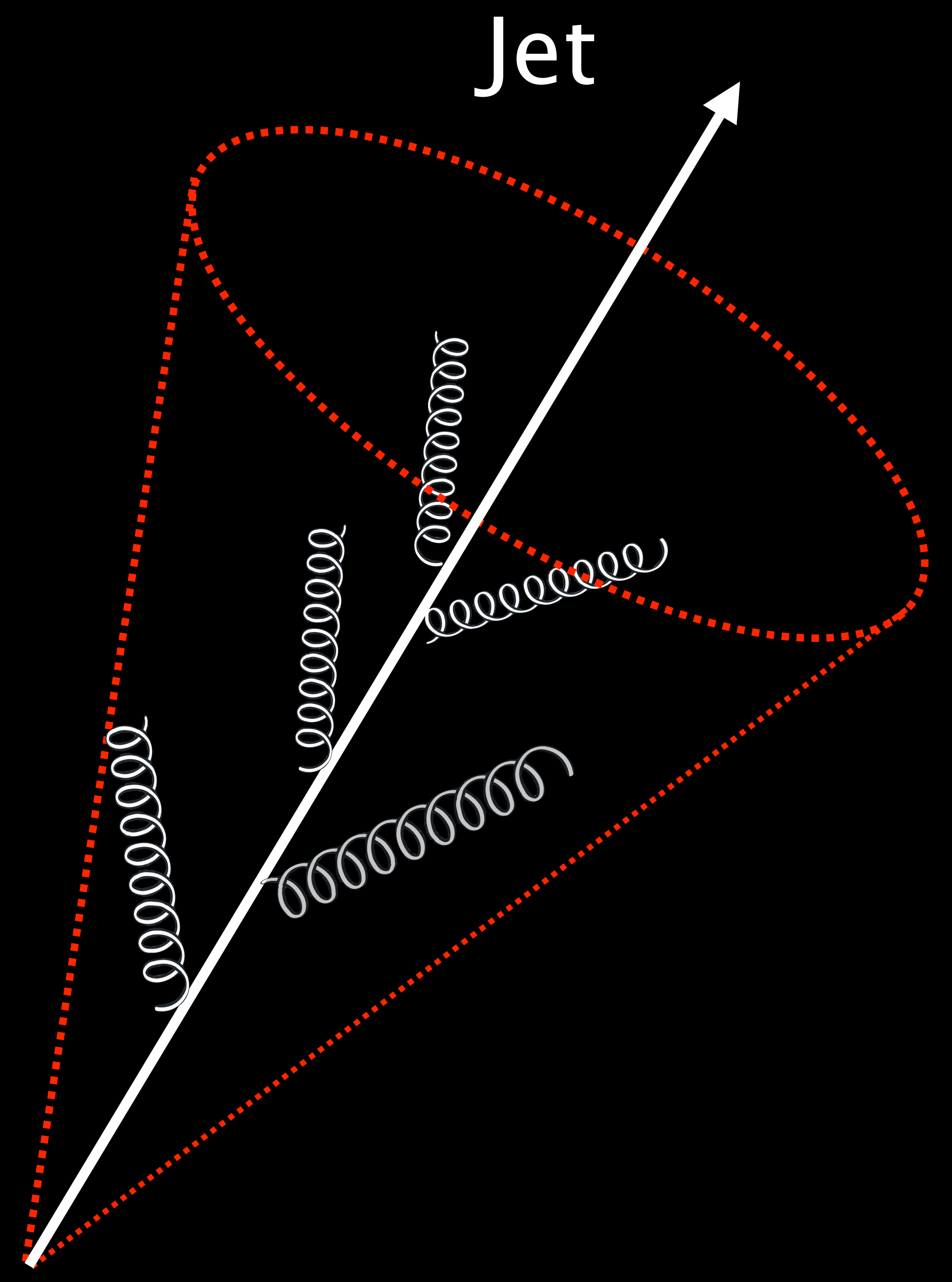
Jaehyeok Ryu

Pusan National University

15 Dec 2025

Introduction

What is jet substructure?

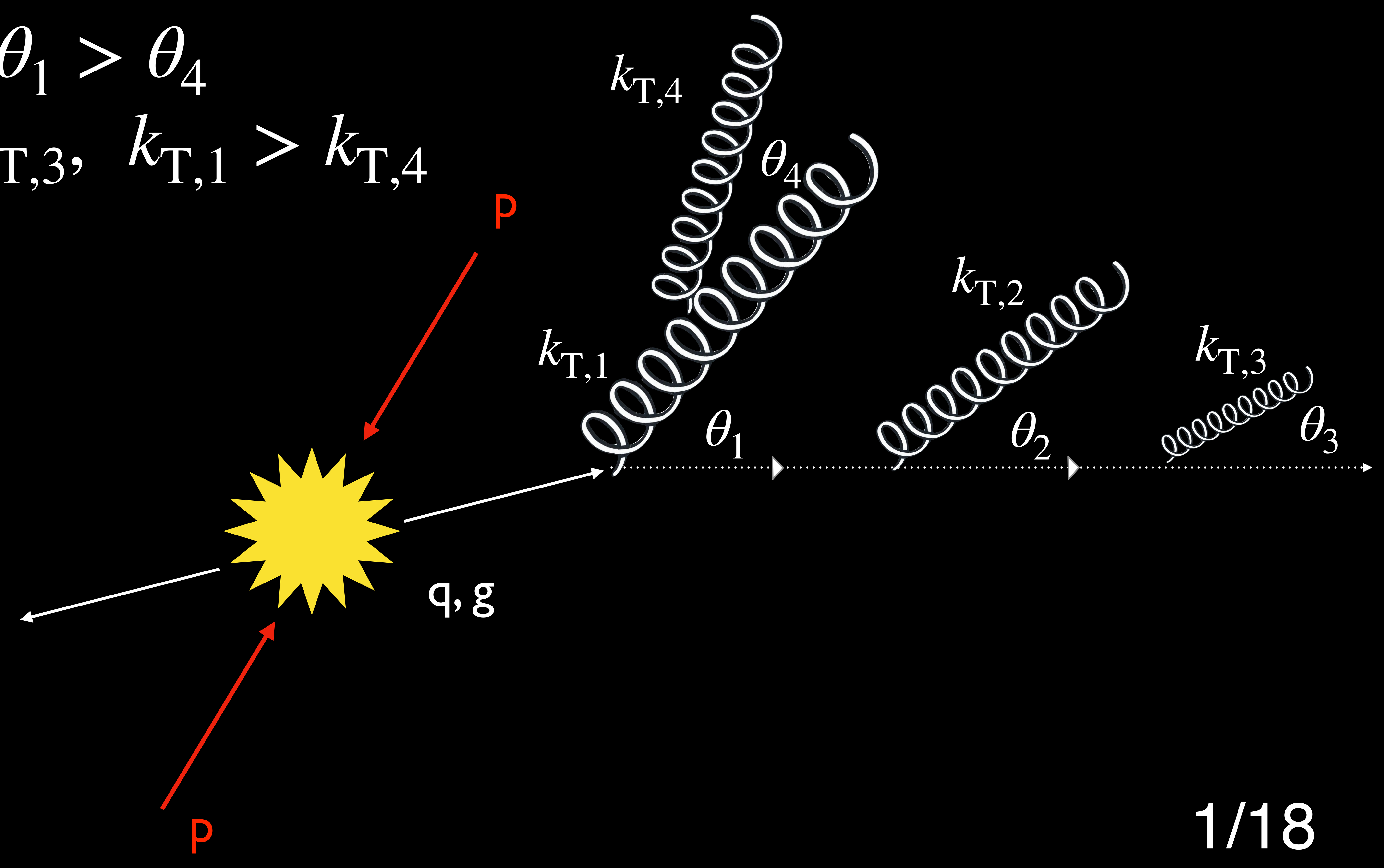


- Internal dynamics of particles constituting jets in angular + momentum space

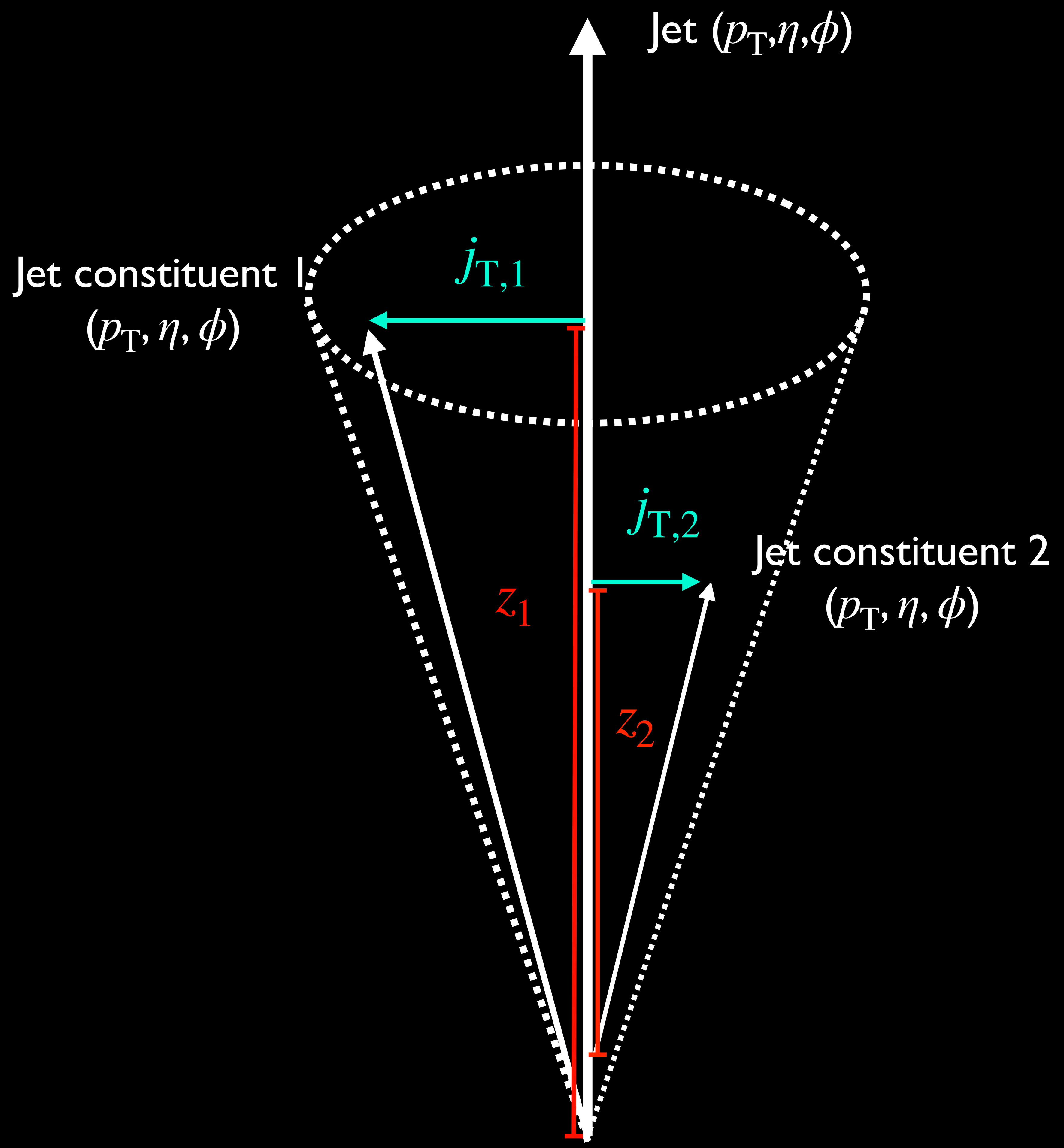
QCD angular ordering

- QCD prefers smaller emitting angles and lower virtuality with increasing number of emitting gluons
- Called angular ordering

$$\theta_1 > \theta_2 > \theta_3,\ \ \theta_1 > \theta_4$$
$$k_{T,1} > k_{T,2} > k_{T,3},\ \ k_{T,1} > k_{T,4}$$



Introduction

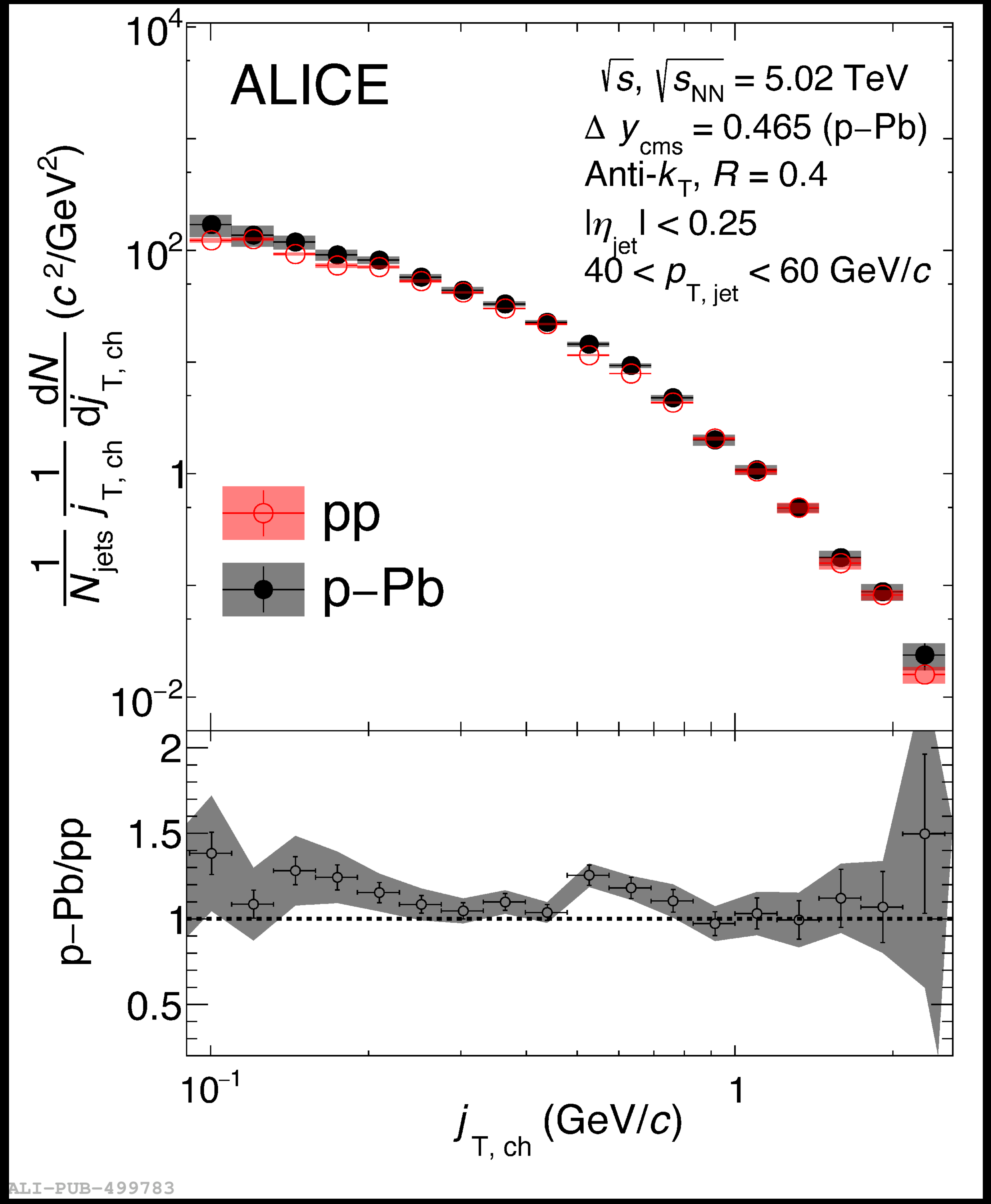


- $j_T = \dfrac{|\vec{p}_{jet} \times \vec{p}_{track}|}{|\vec{p}_{jet}|}$
- $z = \dfrac{\vec{p}_{jet} \cdot \vec{p}_{track}}{p^2_{\ jet}}$

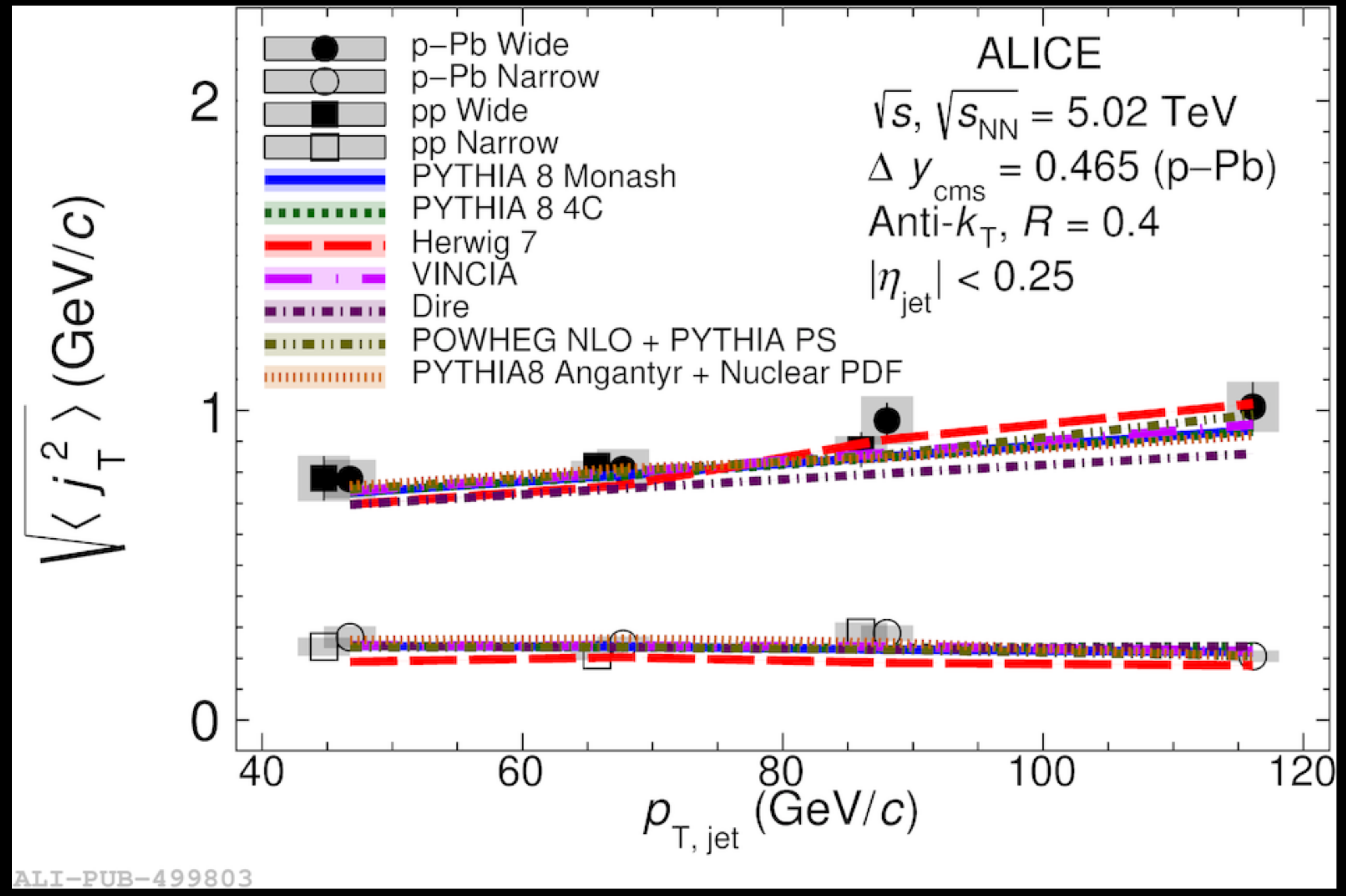
Understanding of QCD jet evolution

- Test our current understanding of QCD theory by measuring differential distributions of charged-particle jet fragments in pp collisions and comparing results to model predictions
- Naive expectation that dominance of high j_T, z components at the early stage (Larger angle) and low j_T, z components at the late stage (Smaller angle)
- Possibly disentangle jet fragmentation and hadronisation processes

JHEP09 (2021) 211

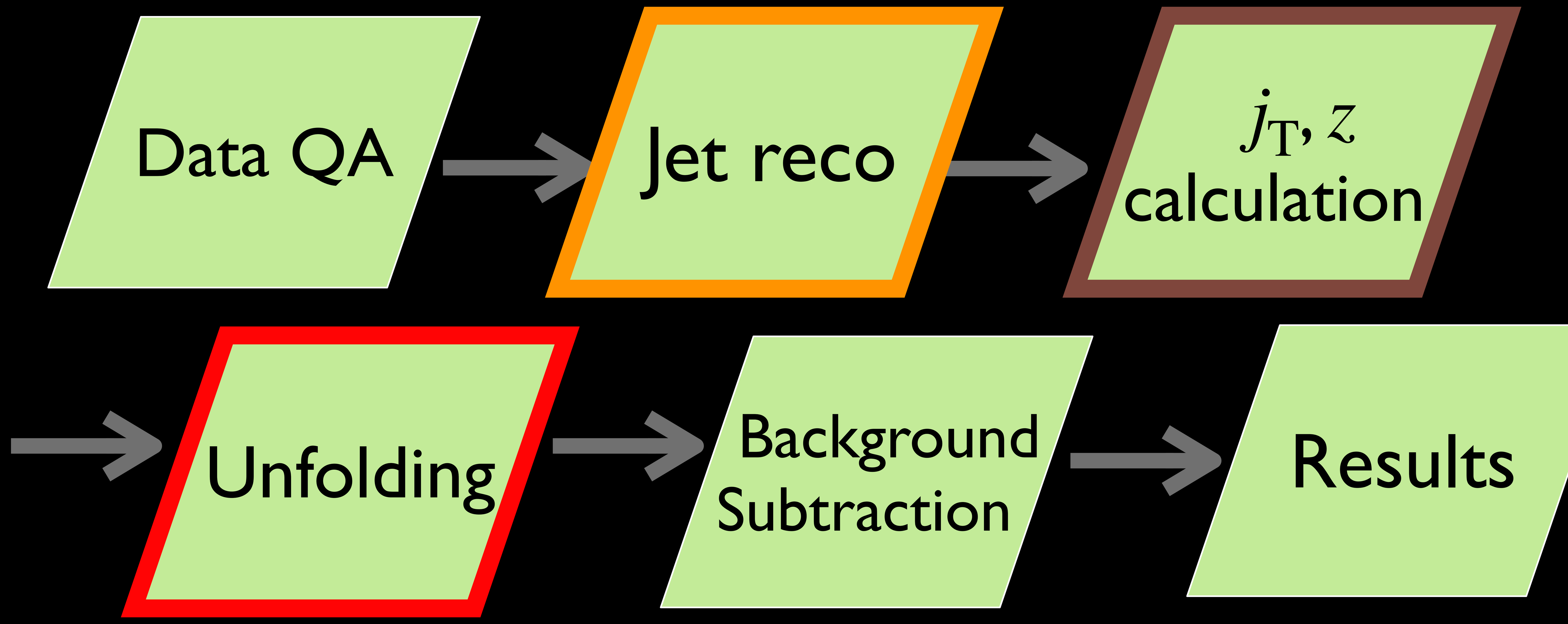


j_{T} in MB pp and p–Pb



Two-components fit

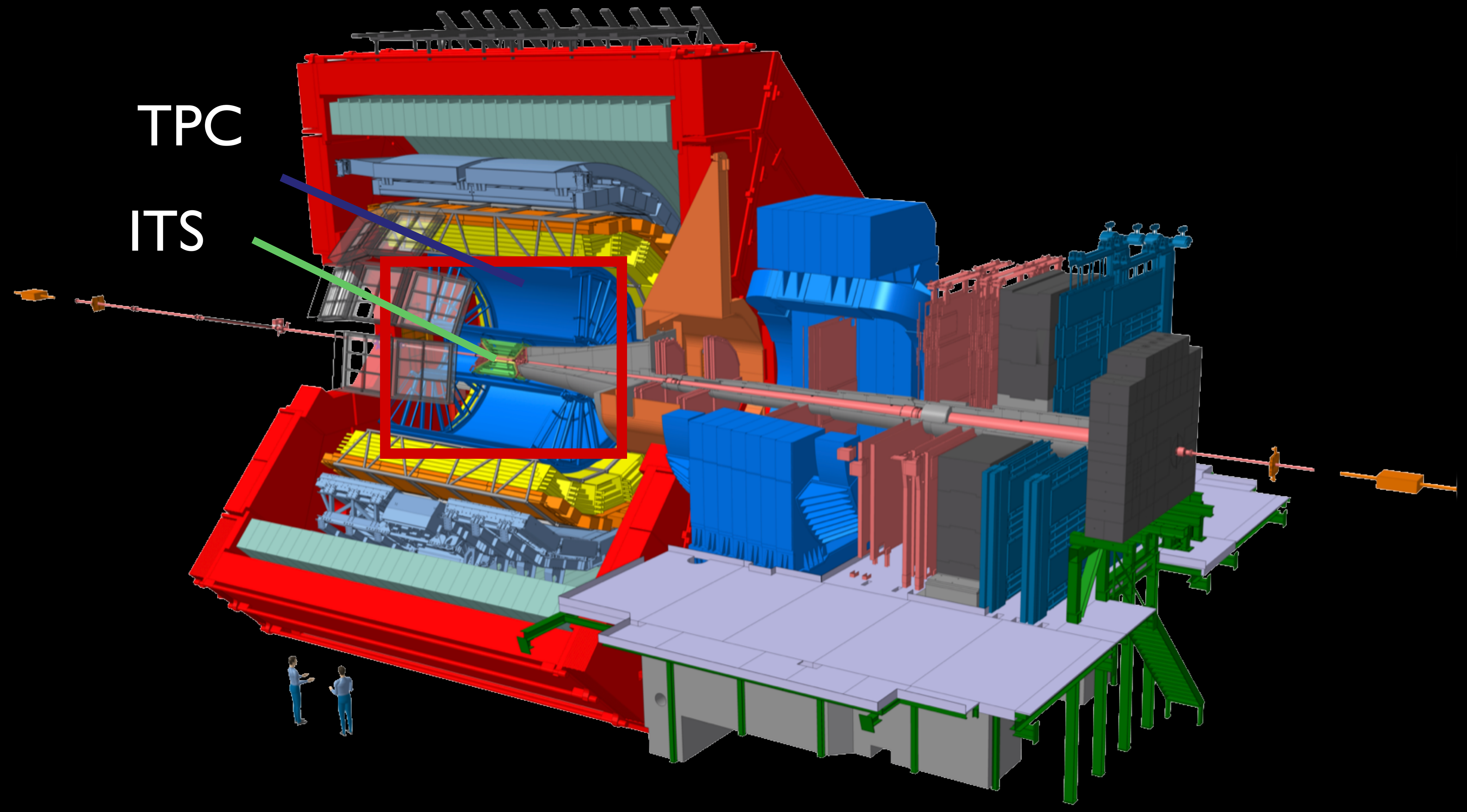
- j_{T} distribution in inclusive z are comparable in pp and p–Pb collision within the uncertainty in the previous ALICE measurement
- Narrow and Wide components of the $\sqrt{\langle j_{\mathrm{T}}^2 \rangle}$ distributions from the two-component fit to disentangle the jet fragmentation and hadronisation process are also similar in both collision systems (No clear medium effect)



- 2017 LHC pp collisions at $\sqrt{s}$ = 5.02 TeV

Dataset - LHC 17p-pass1 FAST + CENT woSDD

MC set - LHC 18b8



Jet reconstruction

- Charged jets in $|\eta| < 0.5$ are reconstructed with charged tracks in the ITS/TPC ($p_T > 0.15$ GeV/c and $|\eta| < 0.9$, $0 < \phi < 2\pi$)
- Anti-k_T algorithm with R = 0.4

j_T, z calculation

- j_T and z are calculated with charged jets and constituent charged tracks reconstructed in ALICE ITS/TPC ($|\eta| < 0.9$, $0 < \phi < 2\pi$)
- Minimum p_T = 0.15 GeV/c for charged particles

3-D Unfolding

- Correct detector effects that smear in jet p_T, z, and j_T by switching to a 3D unfolding procedure

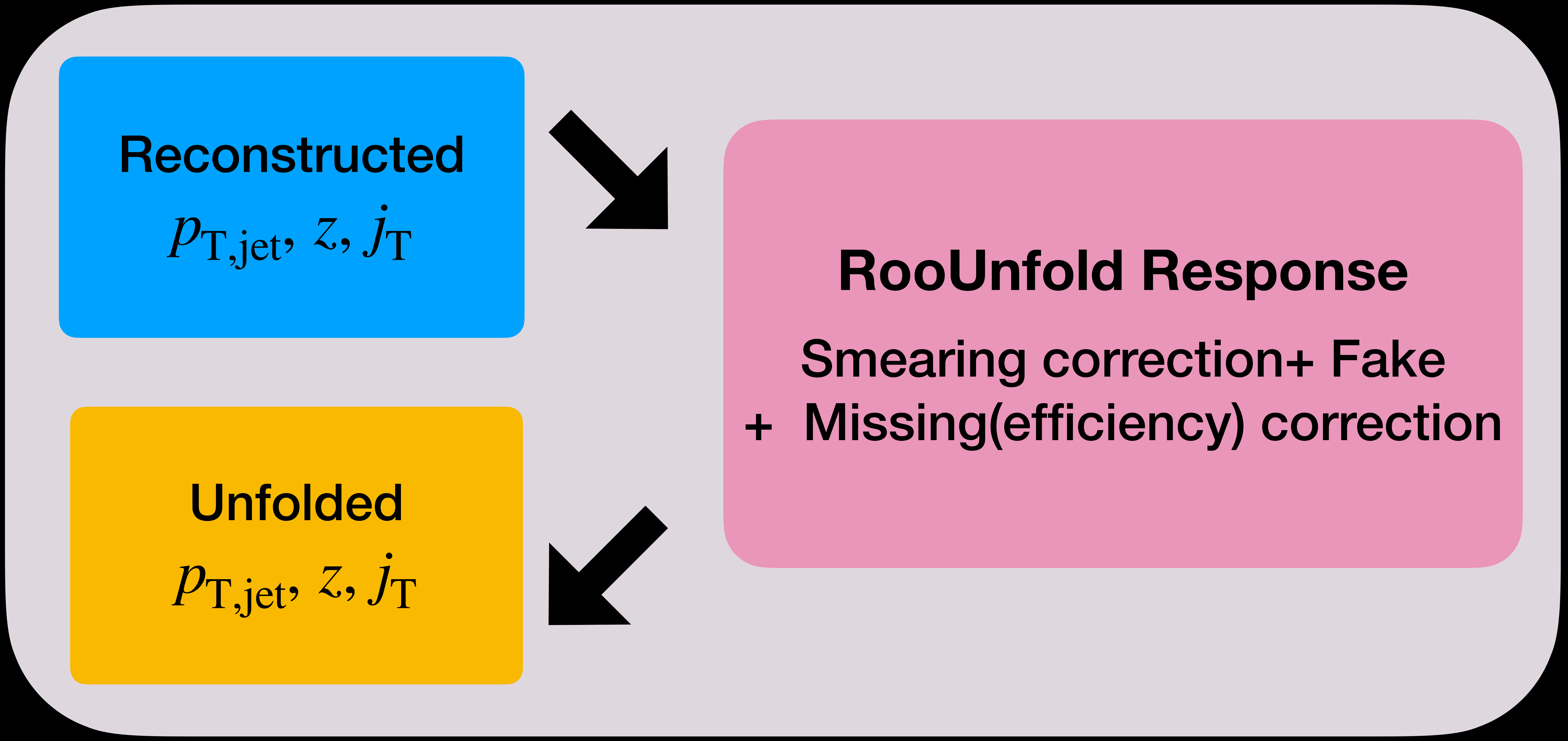
  4-D response matrix (Previous analysis)

  $$(p^{obs}_{T,\,jet},\ j^{obs}_{T},\ p^{true}_{T,\,jet},\ j^{true}_{T})$$

  ↓

  6-D response matrix (This analysis)

  $$(p^{obs}_{T,\,jet},\ z^{obs},\ j^{obs}_{T},\ p^{true}_{T,\,jet},\ z^{true},\ j^{true}_{T})$$

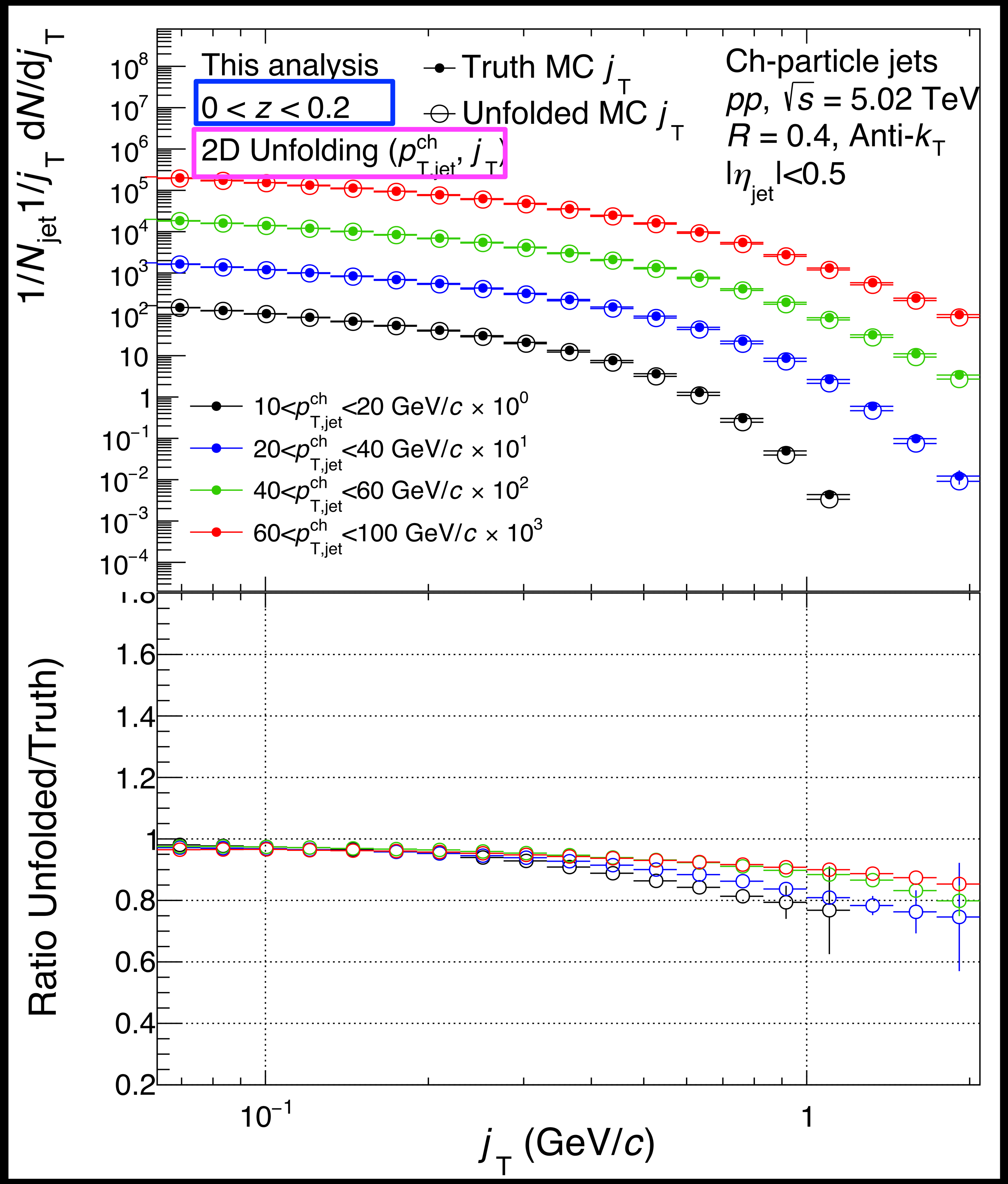


- Iterative Bayesian unfolding method
- Fake correction - Correct fake jets and fake tracks
- Missing corrections - Correct missing jets and missing tracks

Low z bin

$(0 < z < 0.2)$

2-D unfolding



- 2-D and 3-D unfolding closure test results are compared to validate the 3-D unfolding
- Closure test is also done in other z bins

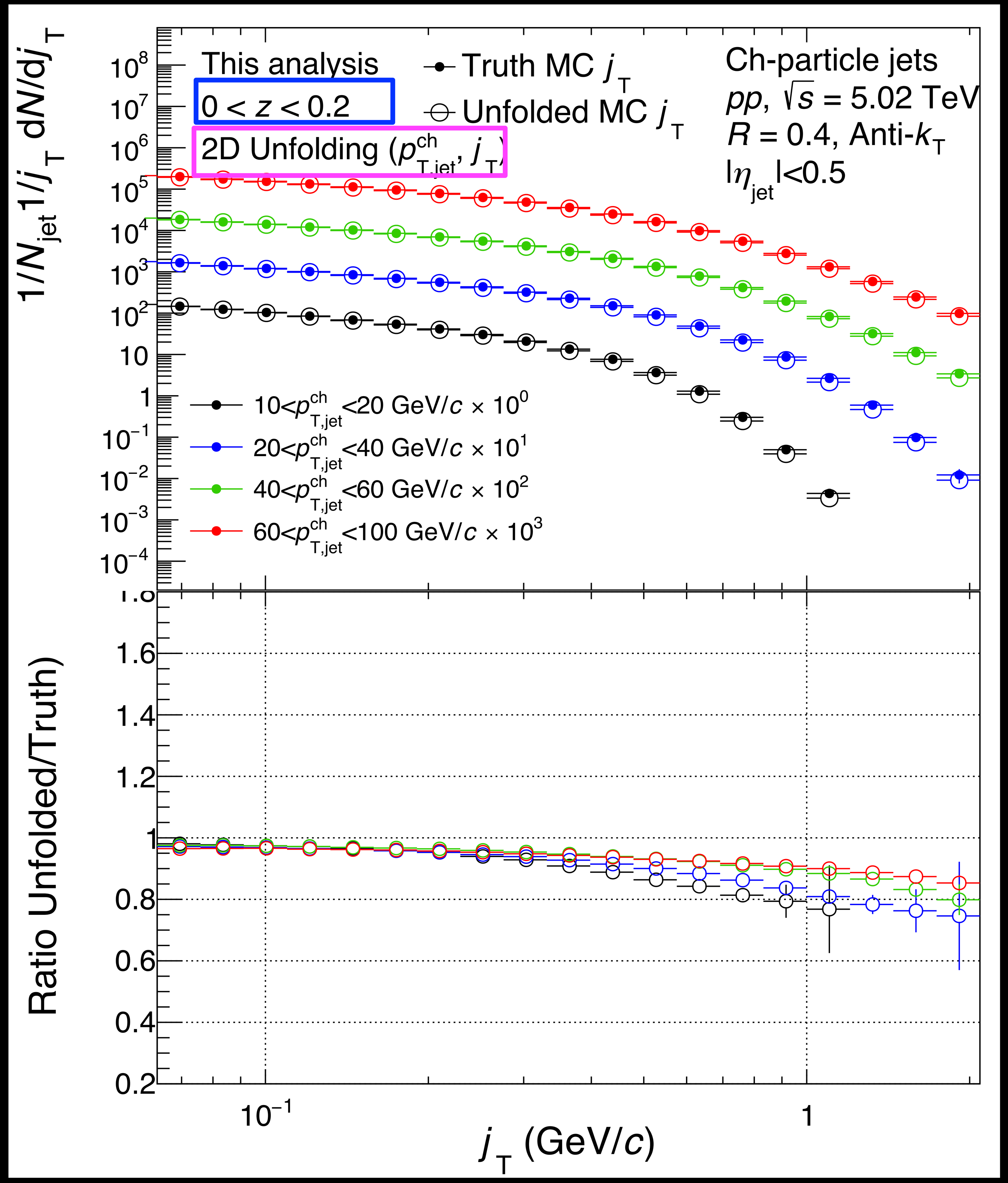
Low z bin

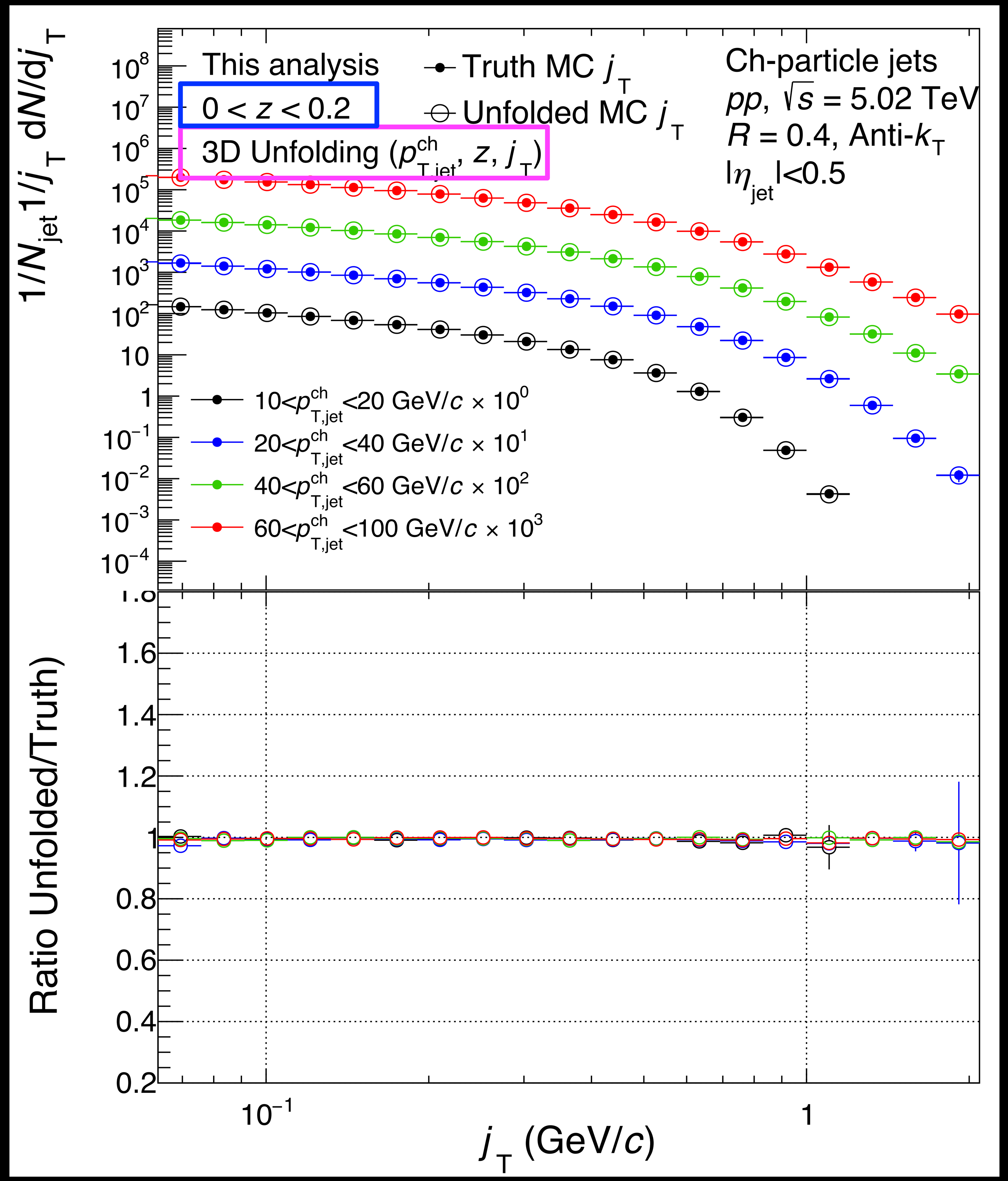
$(0 < z < 0.2)$

ALICE

2-D unfolding

3-D unfolding



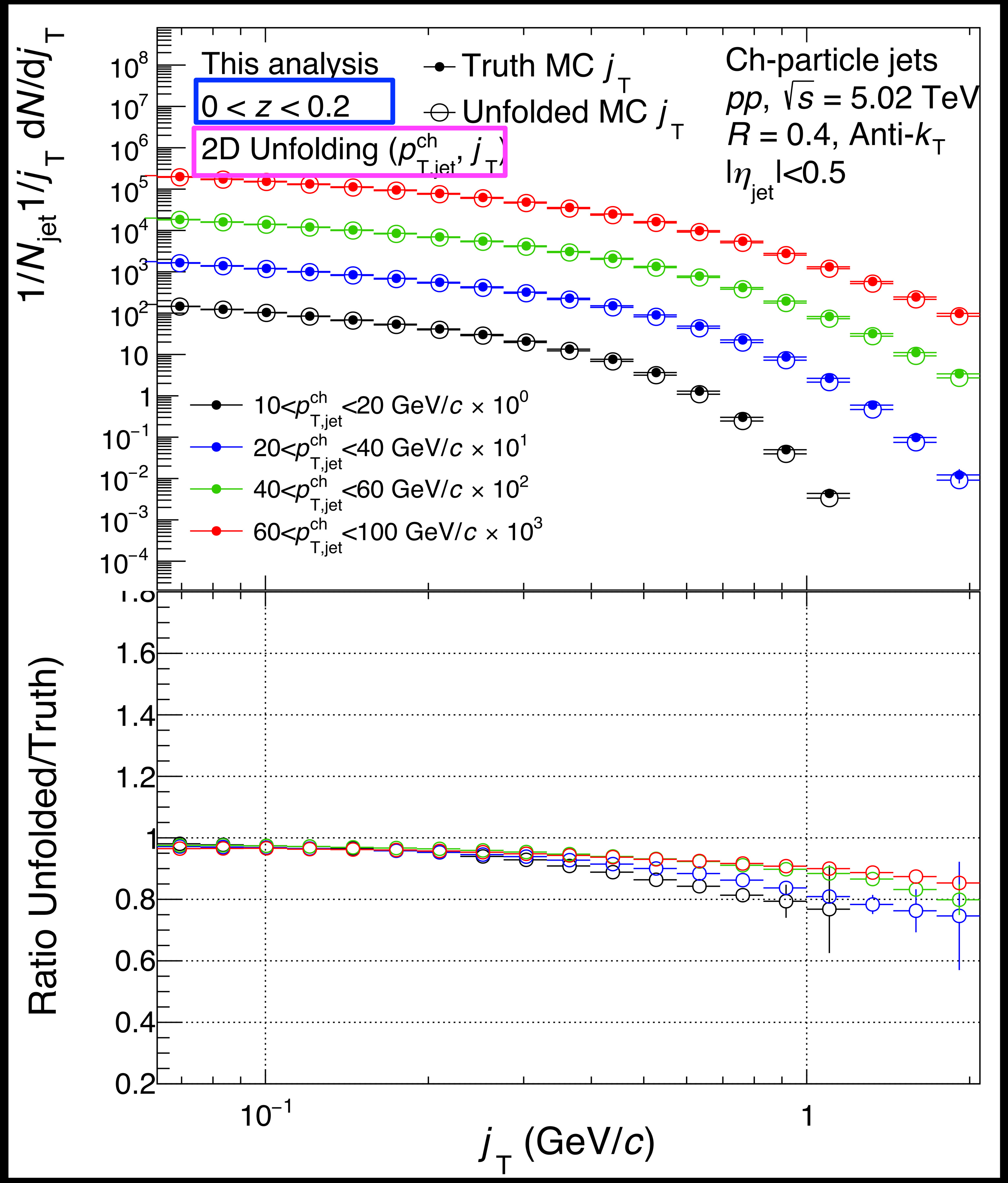


- 2-D and 3-D unfolding closure test results are compared to validate the 3-D unfolding
- Closure test is also done in other z bins

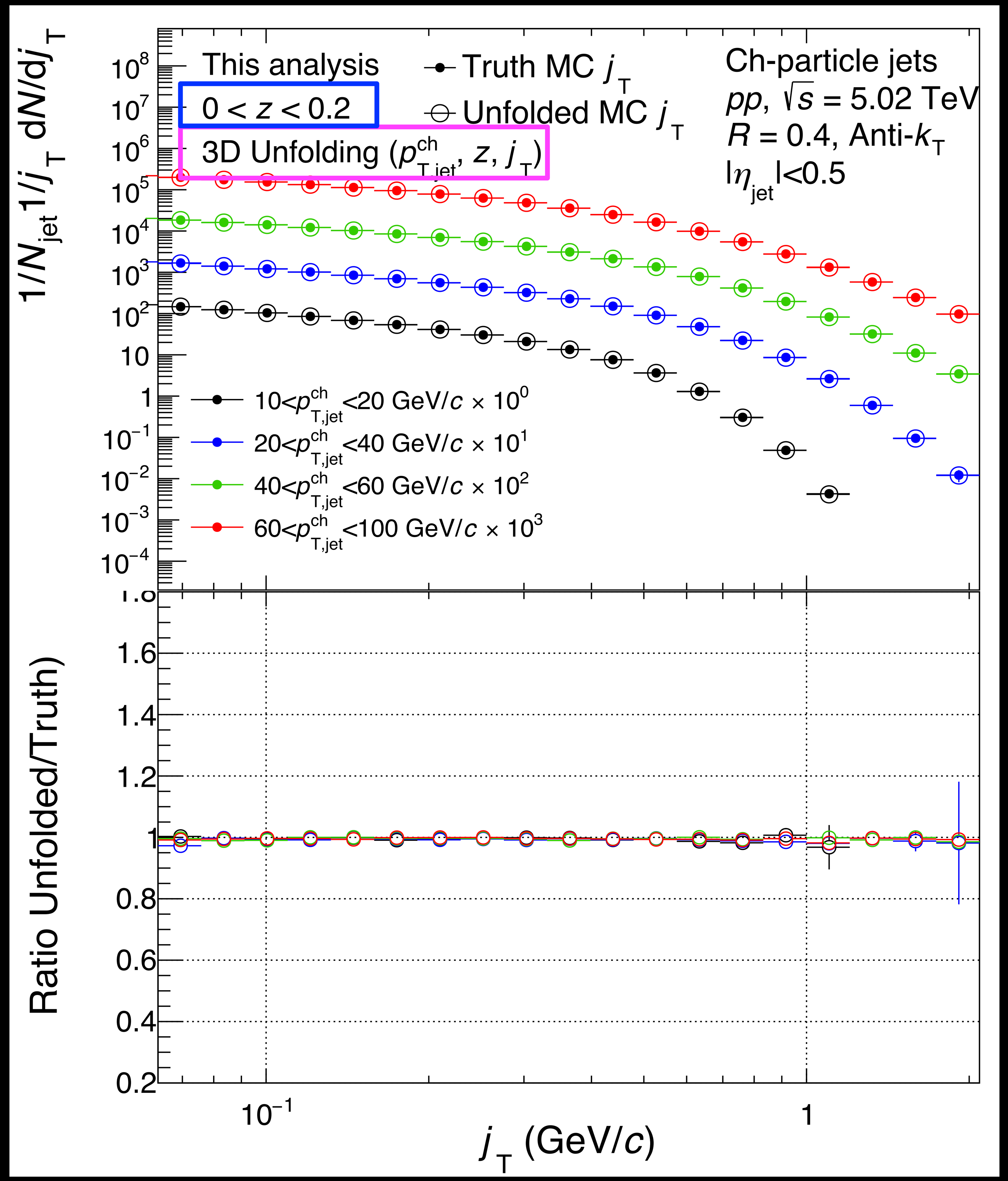
Low z bin

$(0 < z < 0.2)$

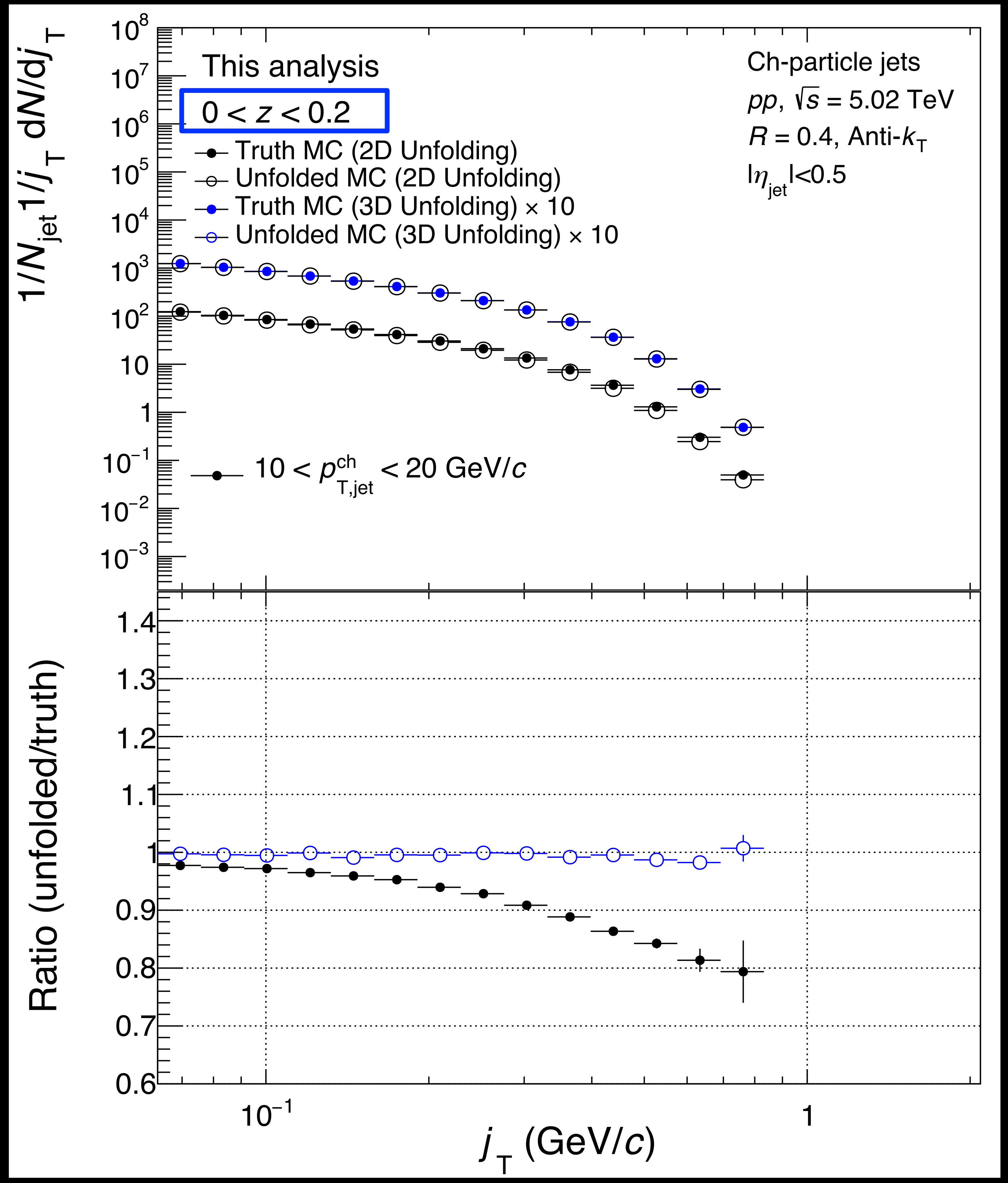
2-D unfolding



3-D unfolding



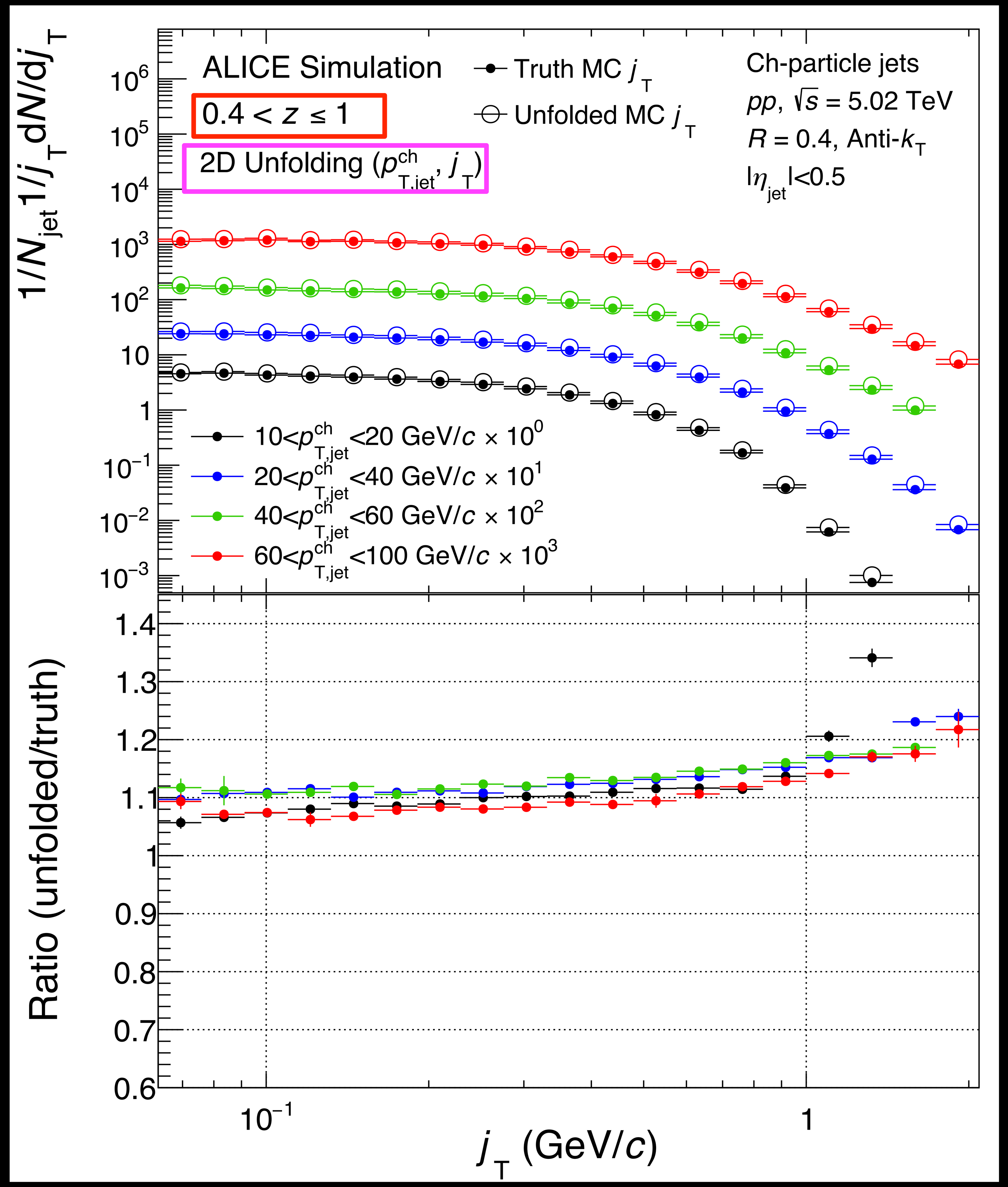
2-D vs 3-D



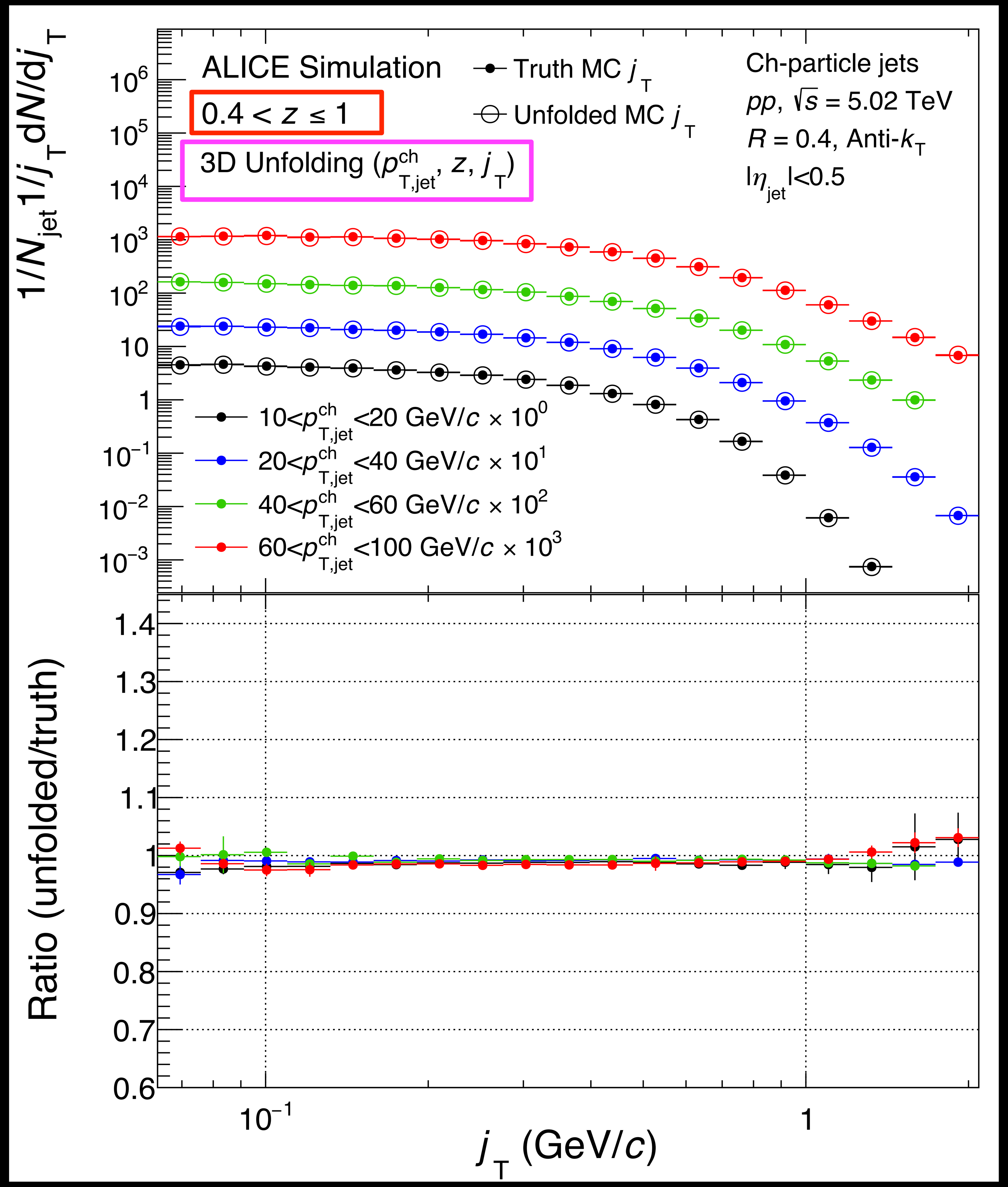
- 2-D and 3-D unfolding closure test results are compared to validate the 3-D unfolding
- Closure test is also done in other z bins

High z bin
$(0.4 < z \leq 1)$

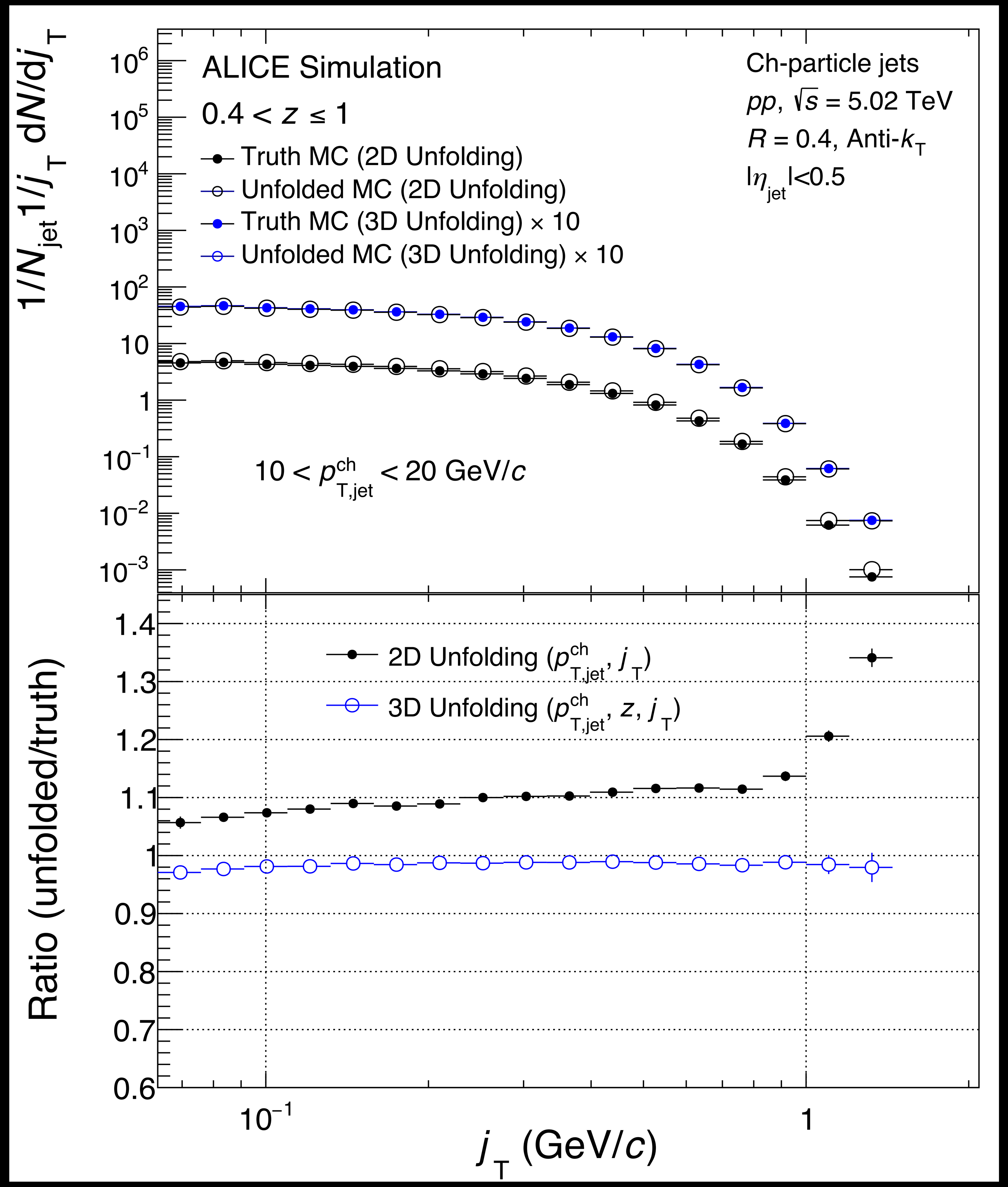
2-D unfolding



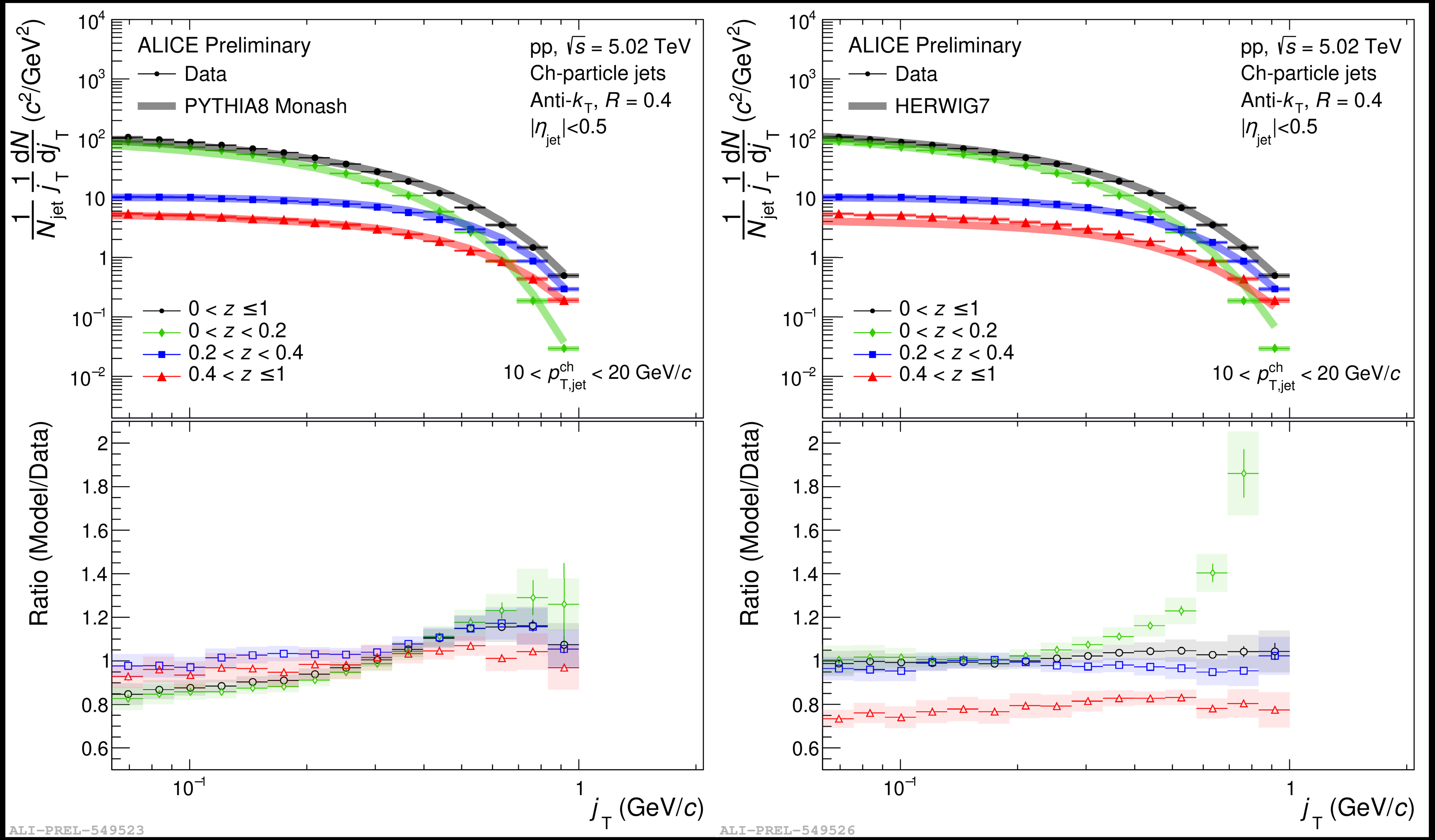
3-D unfolding



2-D vs 3-D

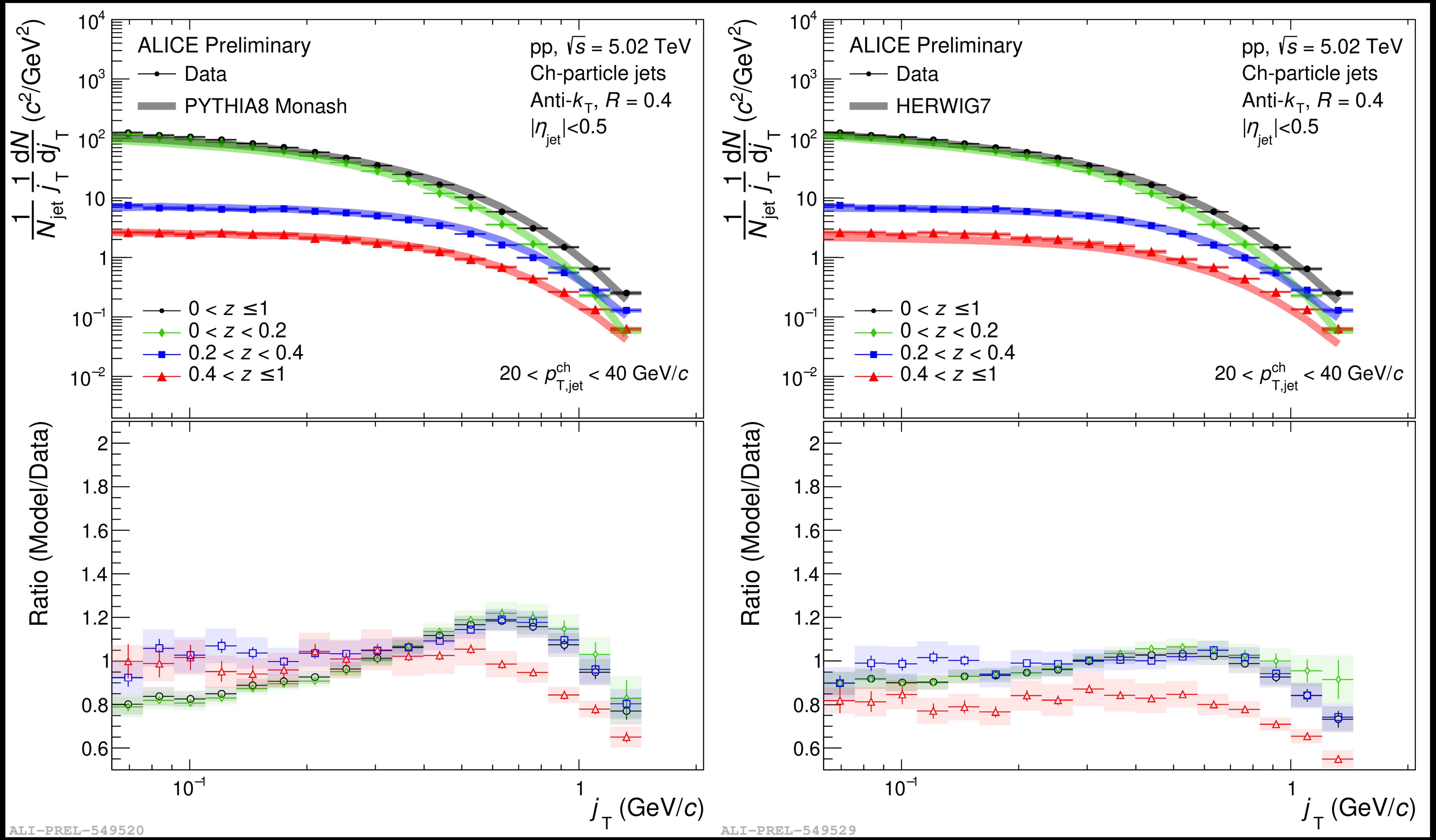


- 2-D and 3-D unfolding closure test results are compared to validate the 3-D unfolding
- Closure test is also done in other z bins



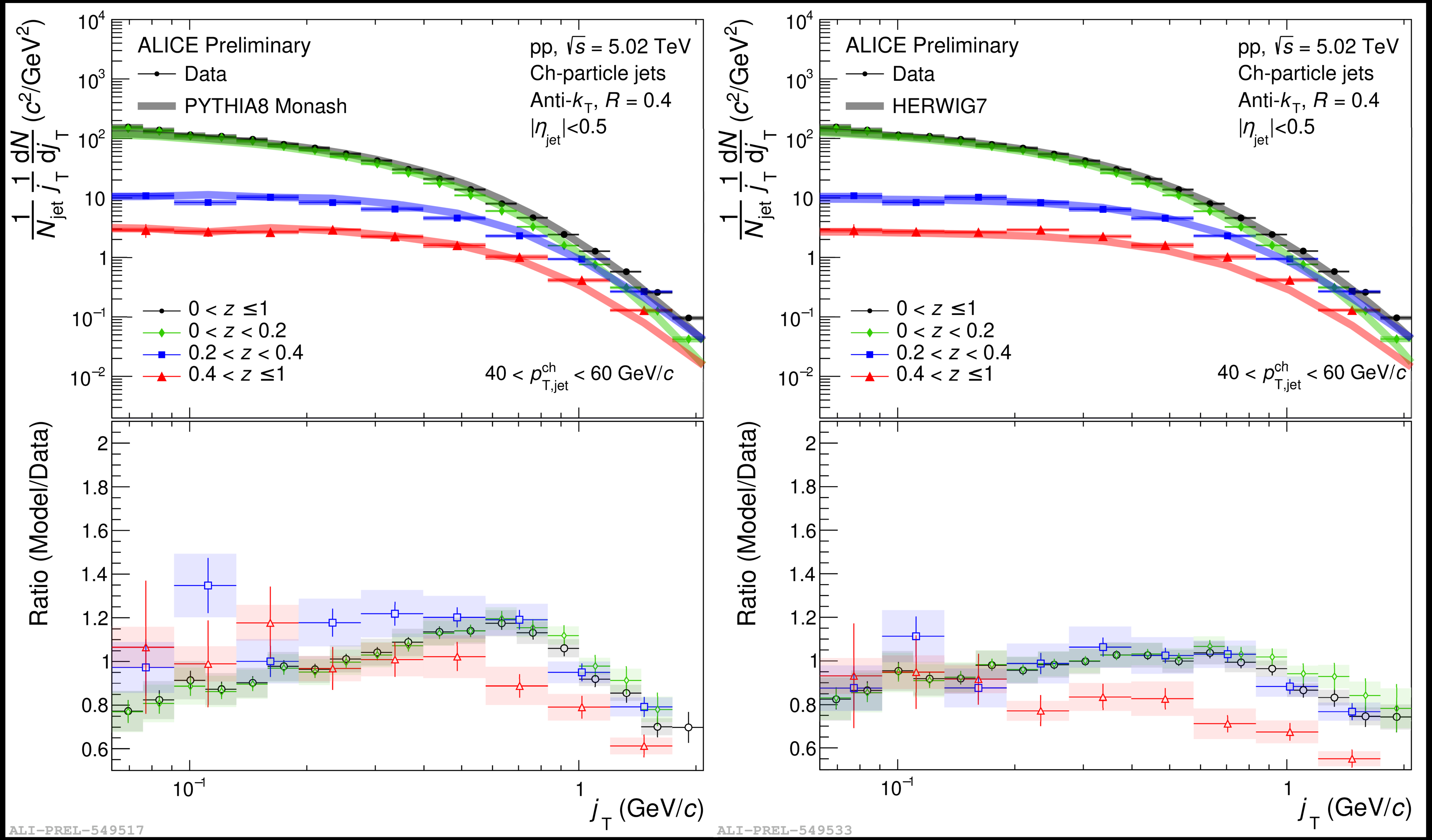
$10 < p_{\mathrm{T,jet}} < 20$ GeV/c

- The j_{T} distributions for different $p_{\mathrm{T,jet}}$ compared with PYTHIA8 Monash and HERWIG 7 for different z ranges
- PYTHIA8 shows the general trend of an increase below the ratio value of 1 at low j_{T} Crossing above the ratio 1 at mid j_{T} Decreasing at high j_{T} in inclusive and low z
- PYTHIA8 shows consistency with the data in mid and high z regions
- Herwig underestimate the high z region and overestimate the low z, high j_{T} region



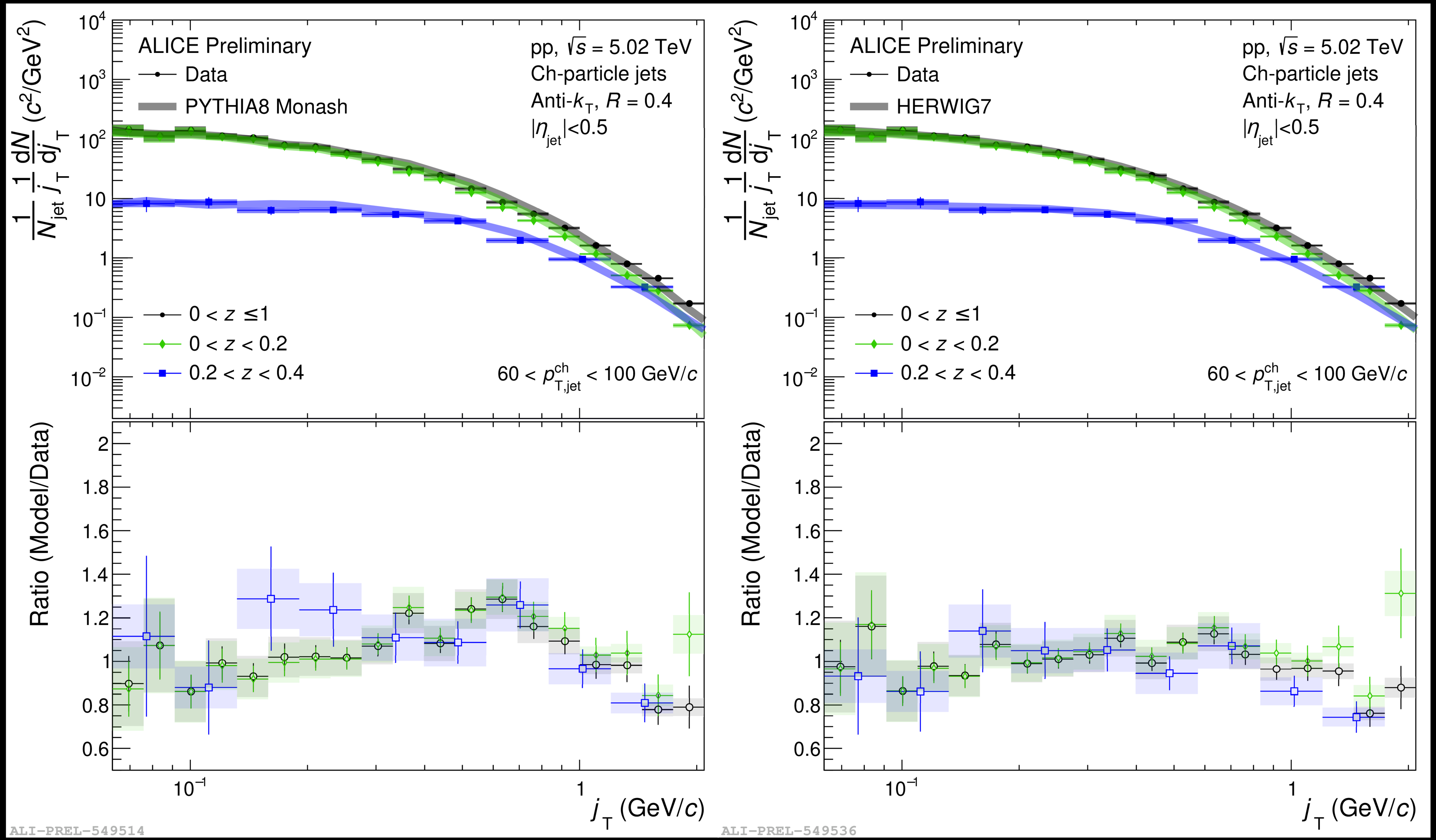
$20 < p_{\mathrm{T,jet}} < 40$ GeV/c

- The j_{T} distributions for different $p_{\mathrm{T,jet}}$ compared with PYTHIA8 Monash and HERWIG 7 for different z ranges
- PYTHIA8 shows the general trend of an increase below the ratio value of 1 at low j_{T} Crossing above the ratio 1 at mid j_{T} Decreasing at high j_{T} in inclusive and low z
- PYTHIA8 shows consistency with the data in mid and high z regions for the lower j_{T}
- Herwig underestimate the high z region and overestimation as lower $p_{\mathrm{T,jet}}$ was disappeared



40 < $p_{\mathrm{T,jet}}$ < 60 GeV/c

- The j_{T} distributions for different $p_{\mathrm{T,jet}}$ compared with PYTHIA8 Monash and HERWIG 7 for different z ranges
- PYTHIA8 shows the general trend of an increase below the ratio value of 1 at low j_{T} Crossing above the ratio 1 at mid j_{T} Decreasing at high j_{T} in inclusive and low z
- j_{T} distributions widen with increasing z, $p_{\mathrm{T,jet}}$
- Herwig underestimate the high z region and overestimation as lower $p_{\mathrm{T,jet}}$ was disappeared

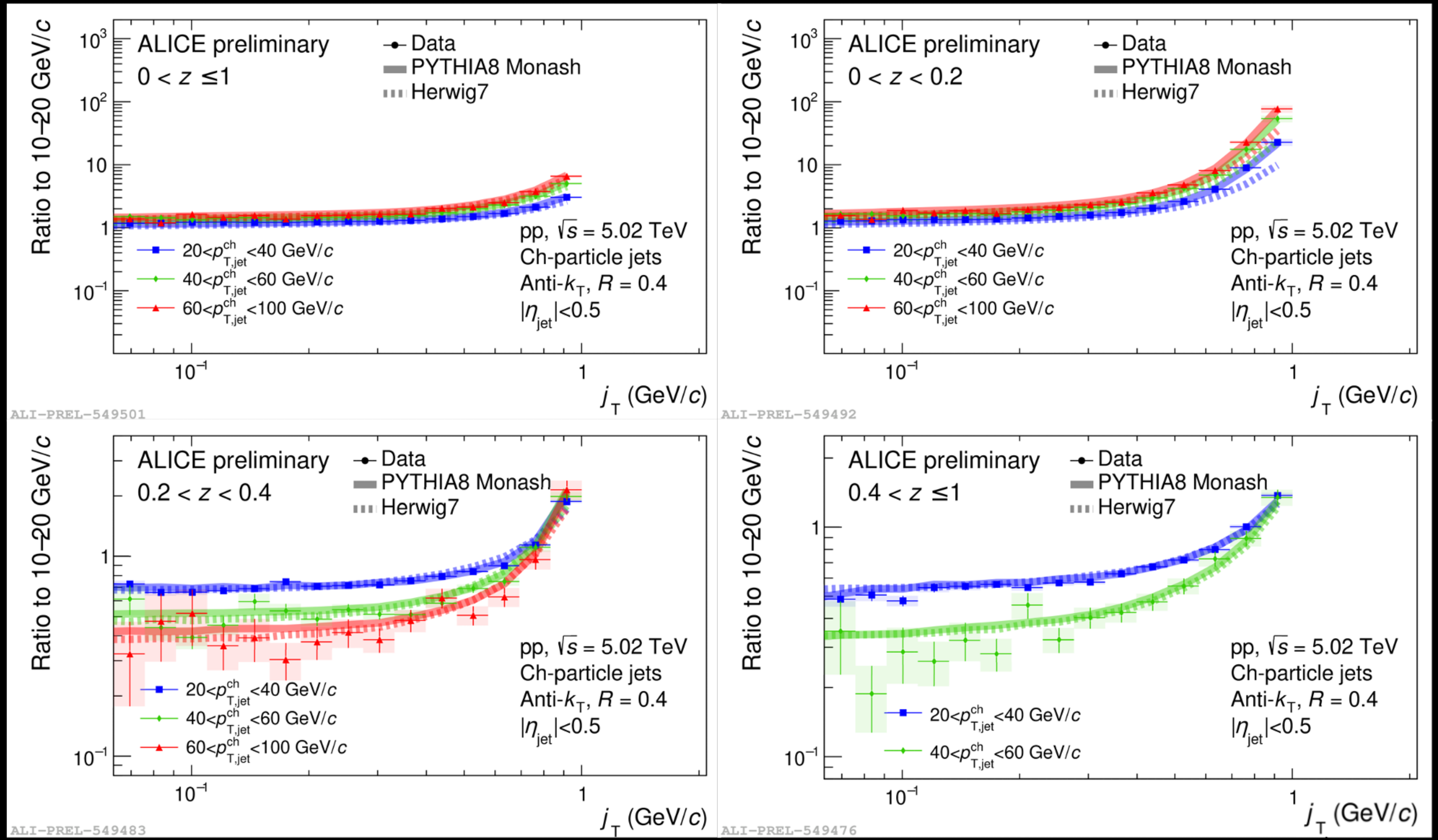


$60 < p_{T,jet} < 100$ GeV/c

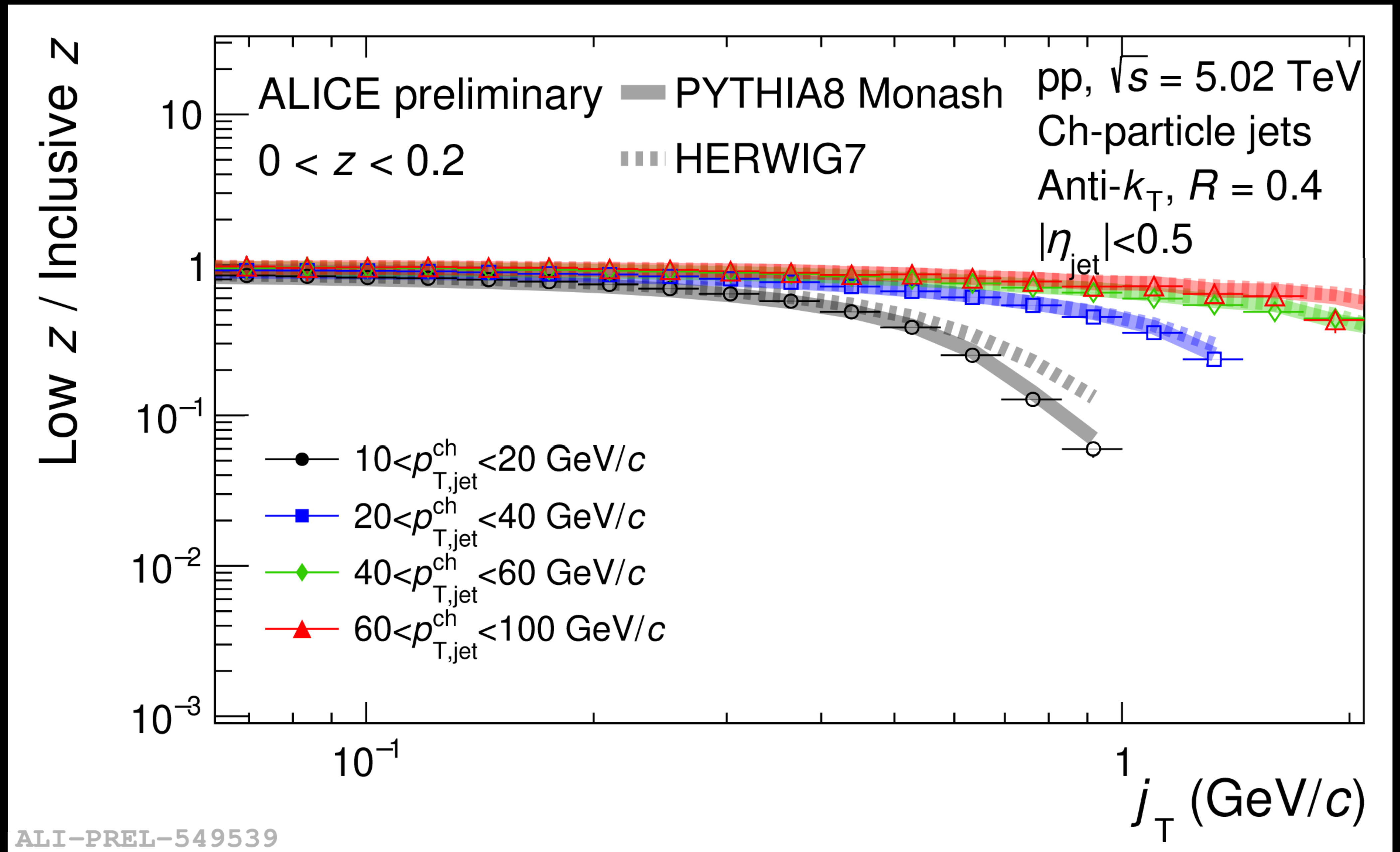
- The j_T distributions for different $p_{T,jet}$ compared with PYTHIA8 Monash and HERWIG 7
- j_T distributions widen with increasing z, $p_{T,jet}$
- Descriptions of models are different in the different kinematic ranges

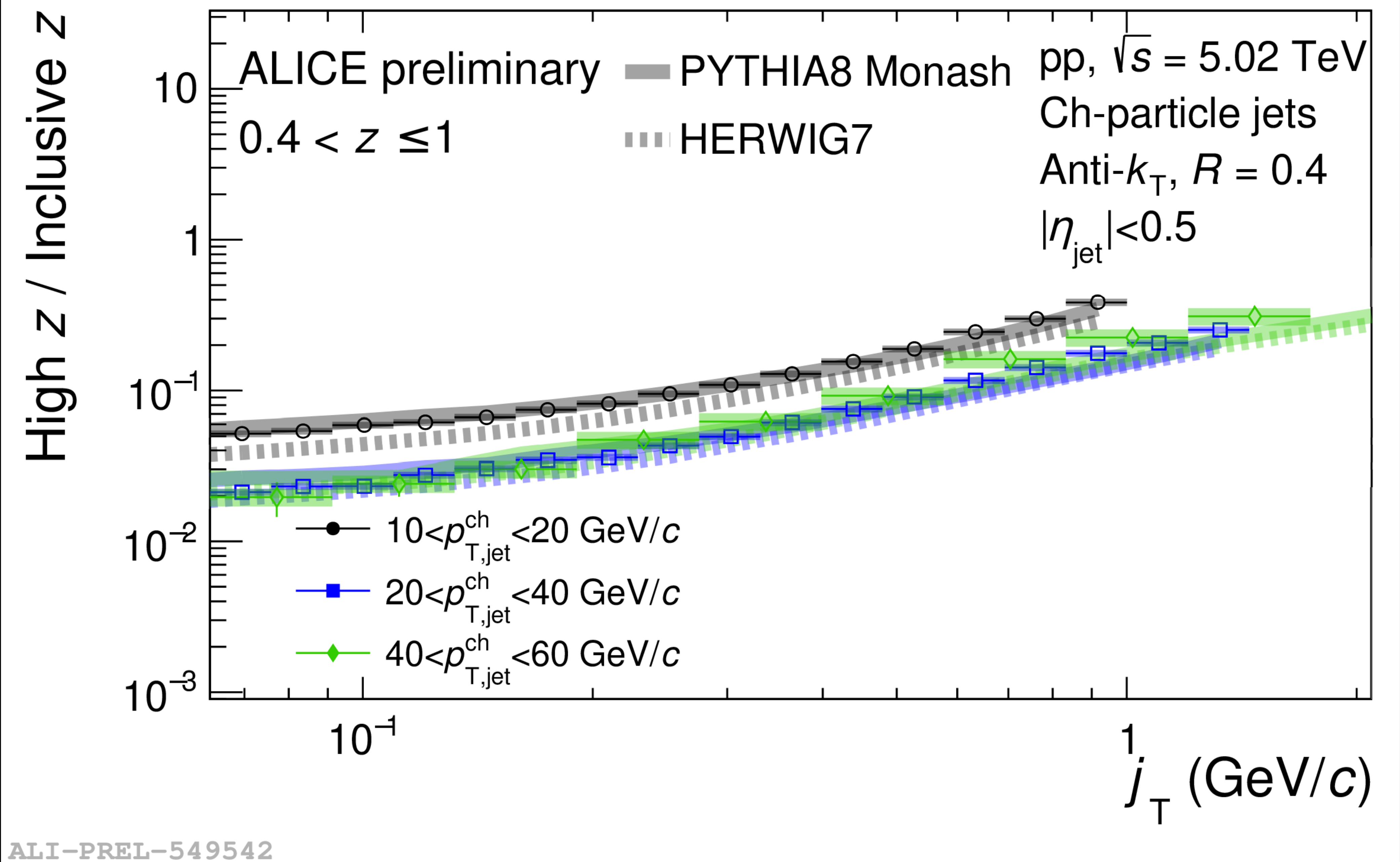
- Comparison with POWHEG + PYTHIA will also be in the paper
- POWHEG + PYTHIA shows similar trends that overestimate in low z, high j_T region and under-estimate the high z region tells the trend is from the difference between the LO and NLO calculation of initial hard scattering process

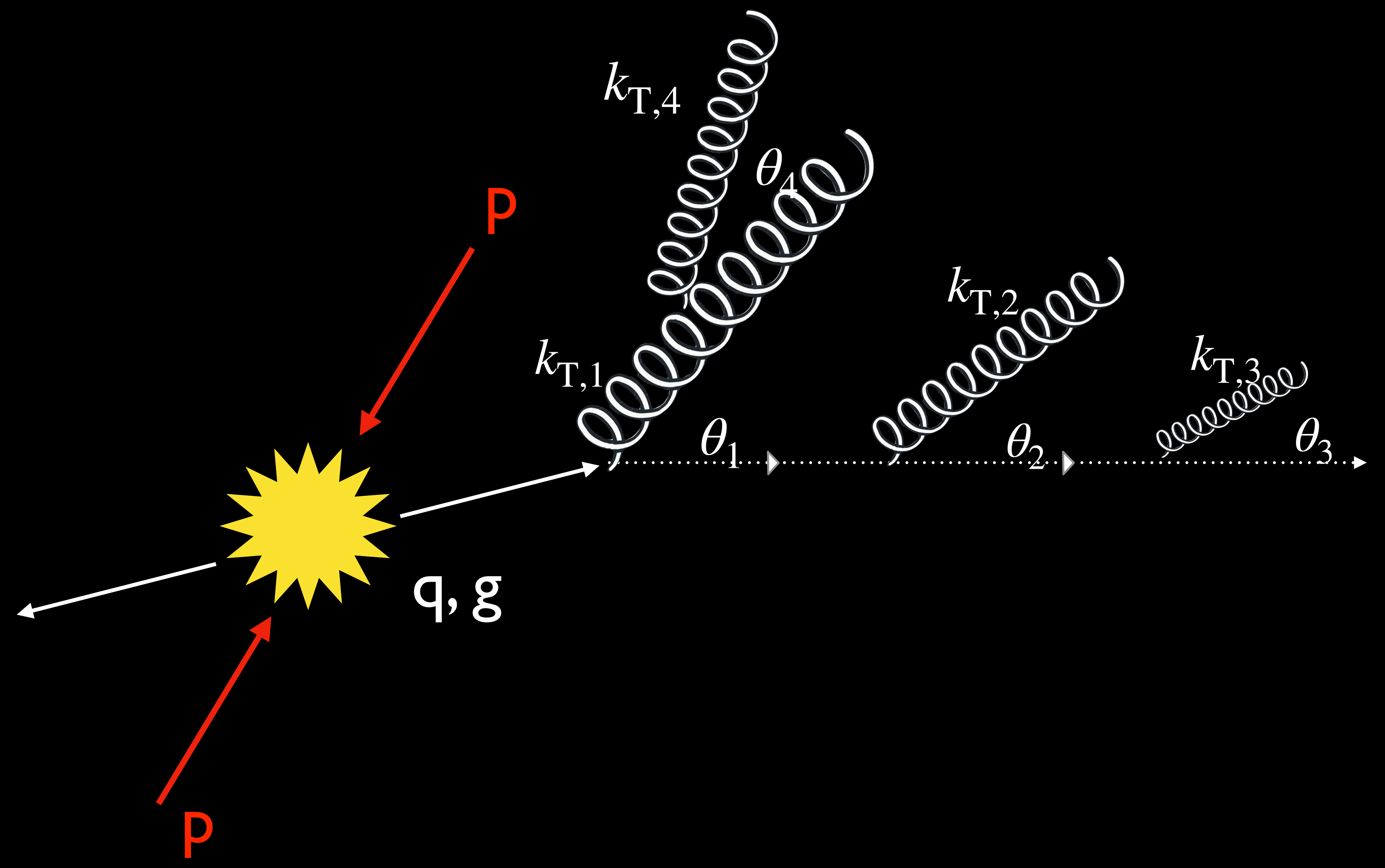
- Comparison with POWHEG + PYTHIA will also be in the paper
- POWHEG + PYTHIA shows similar trends that overestimate in low z, high j_T region and under-estimate the high z region tells the trend is from the difference between the LO and NLO calculation of initial hard scattering process



- j_T distributions in other $p_{\mathrm{T,jet}}$ regions compared to the 10-20 GeV/*c*
- While both models couldn't describe the j_T distribution, they have good descriptions on the trend of the ratio in all z ranges
- Indicate the difference between the model and data is not from the $p_{\mathrm{T,jet}}$ dependence
- Comparisons with MC generators set constraints on models



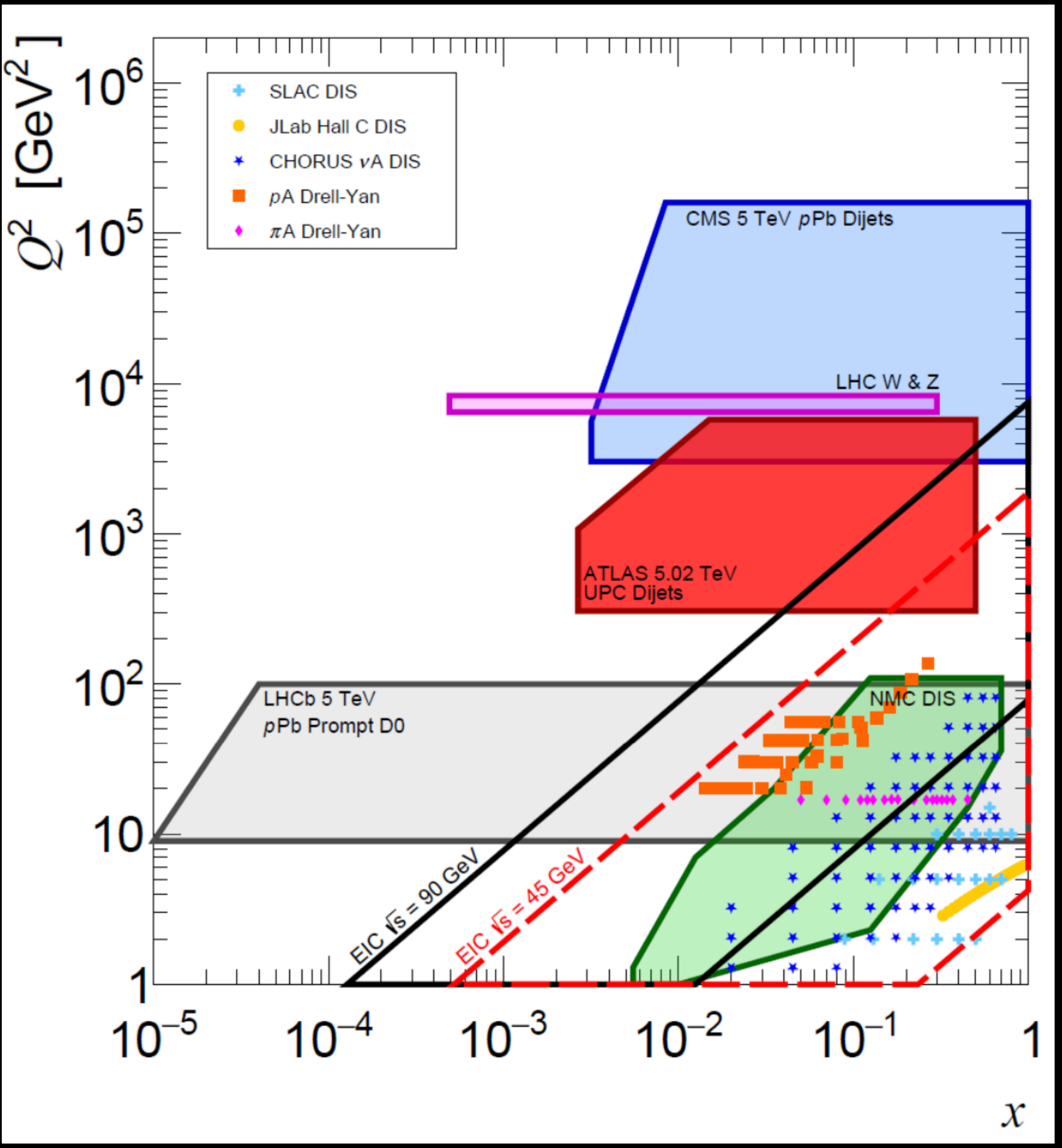




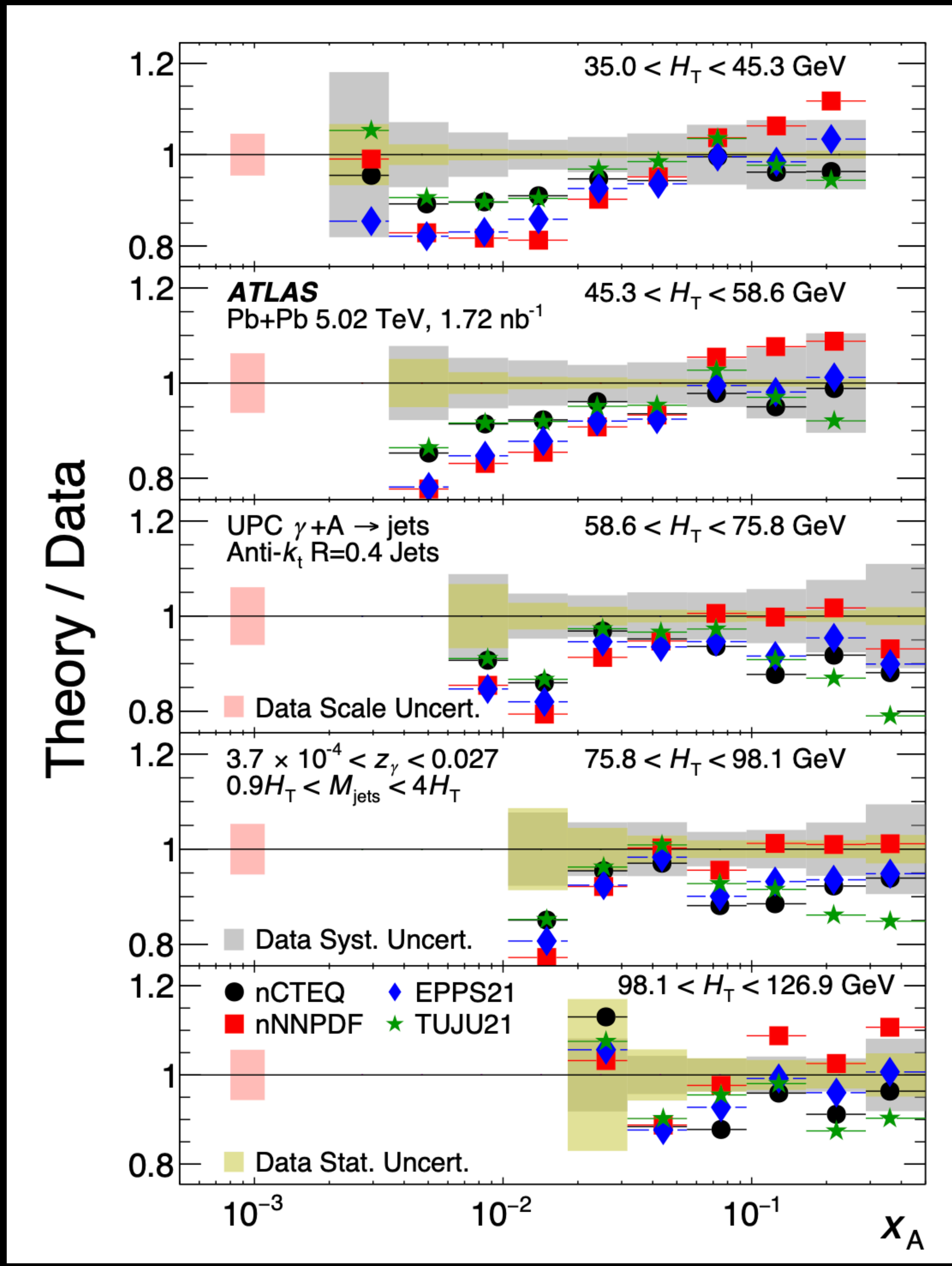
- Low j_T components are dominant in the low z, high j_T components are dominant in the high z which is consistent with QCD theory
- Models qualitatively explain the data
- Further research on the z dependent j_T in specific kinematic ranges might allow us to explore the small system jet quenching

Inclusive z | Low z | High z

- First reconstructed-level j_T distributions in ALICE Run 3 using MC data are shown for inclusive, low, and high-z region
- Based on jet-enhanced MC sample without correction (LHC24g4)
- Plan to investigate how the j_T evolves with event multiplicity or jet flavor in future steps of the analysis



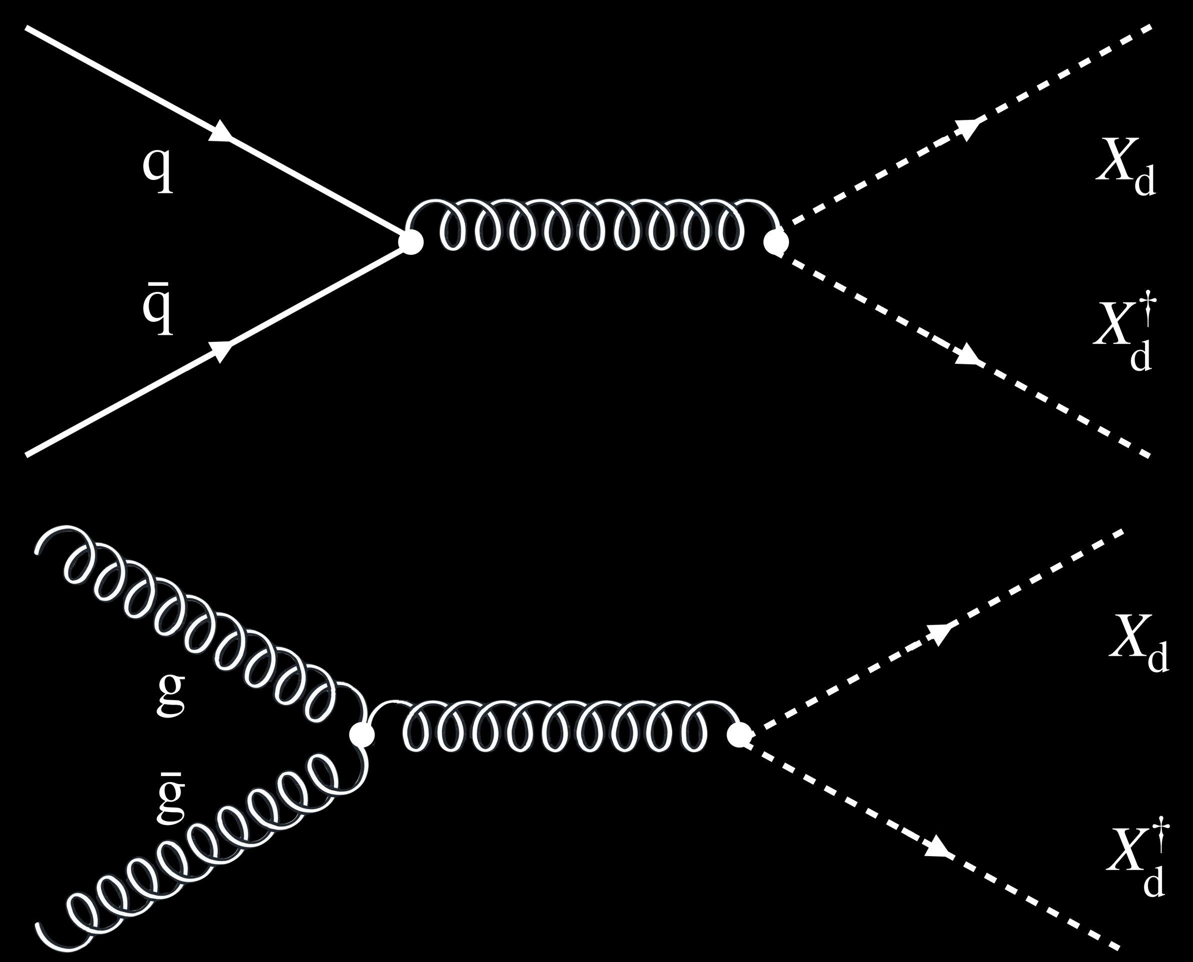
- ATLAS recently measured photonuclear dijet and multijet production in Pb+Pb UPC
- Covering an x_A-Q^2 region that nicely bridges existing LHC p+Pb jet data and the projected γA/eA coverage at the EIC
- In the shadowing region at low x_A, all current nPDF sets systematically underpredict the UPC γ Pb jet yields
- No single nPDF describes the full (x_A, H_T) range, so these UPC dijet data provide strong new constraints on the nuclear gluon density



Forked from the ATLAS slide

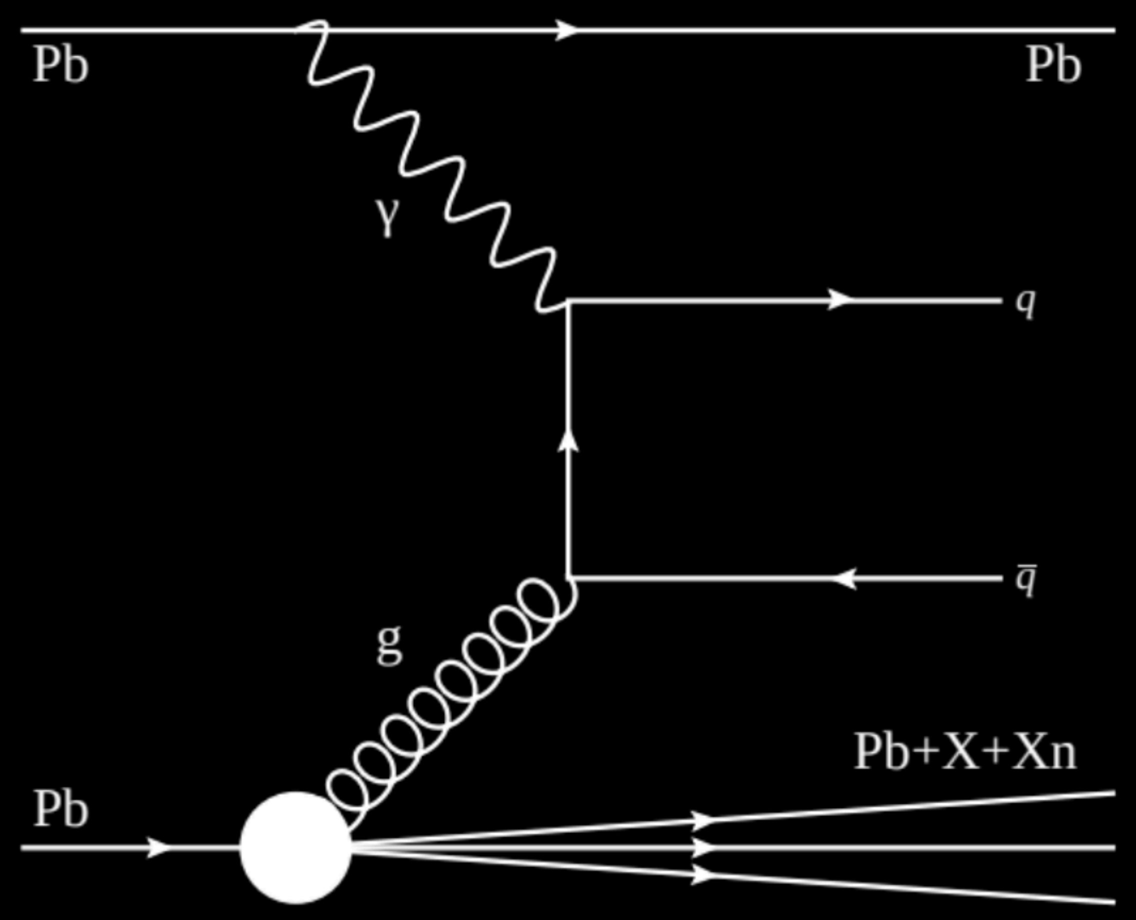
Derived from Arrive:2112.12462

pp



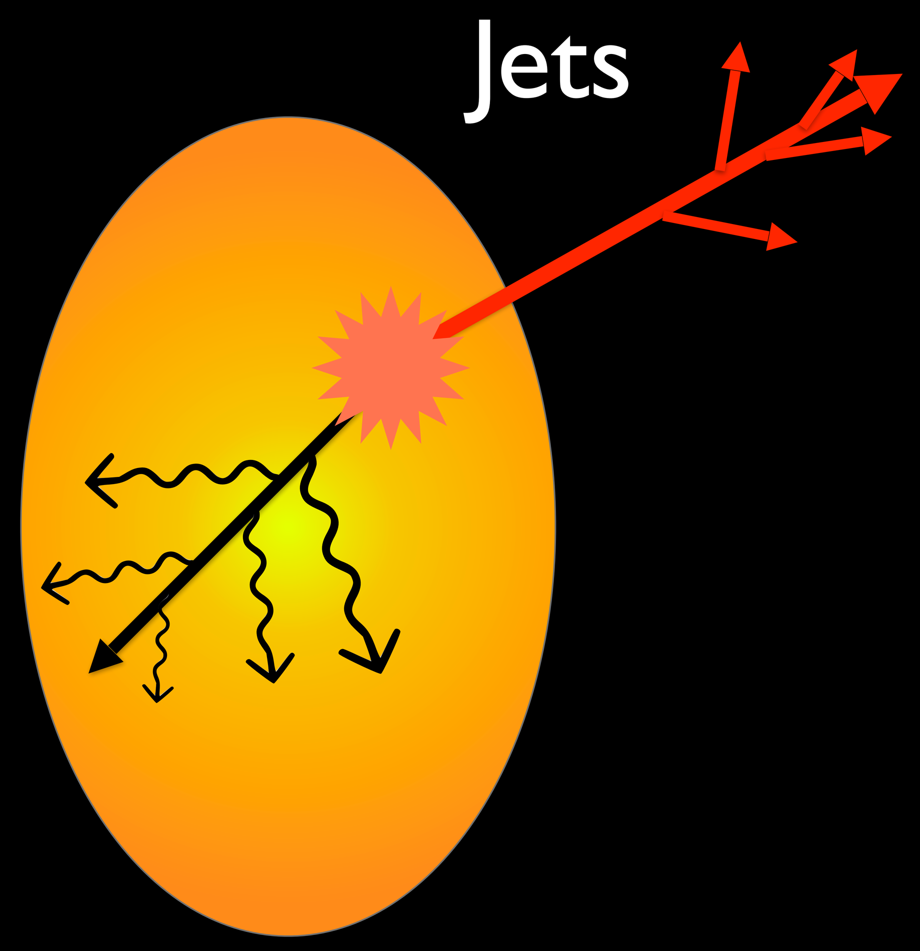
- Minimum-bias pp case, hard scattering and jet evolution in (almost) pure vacuum

γPb



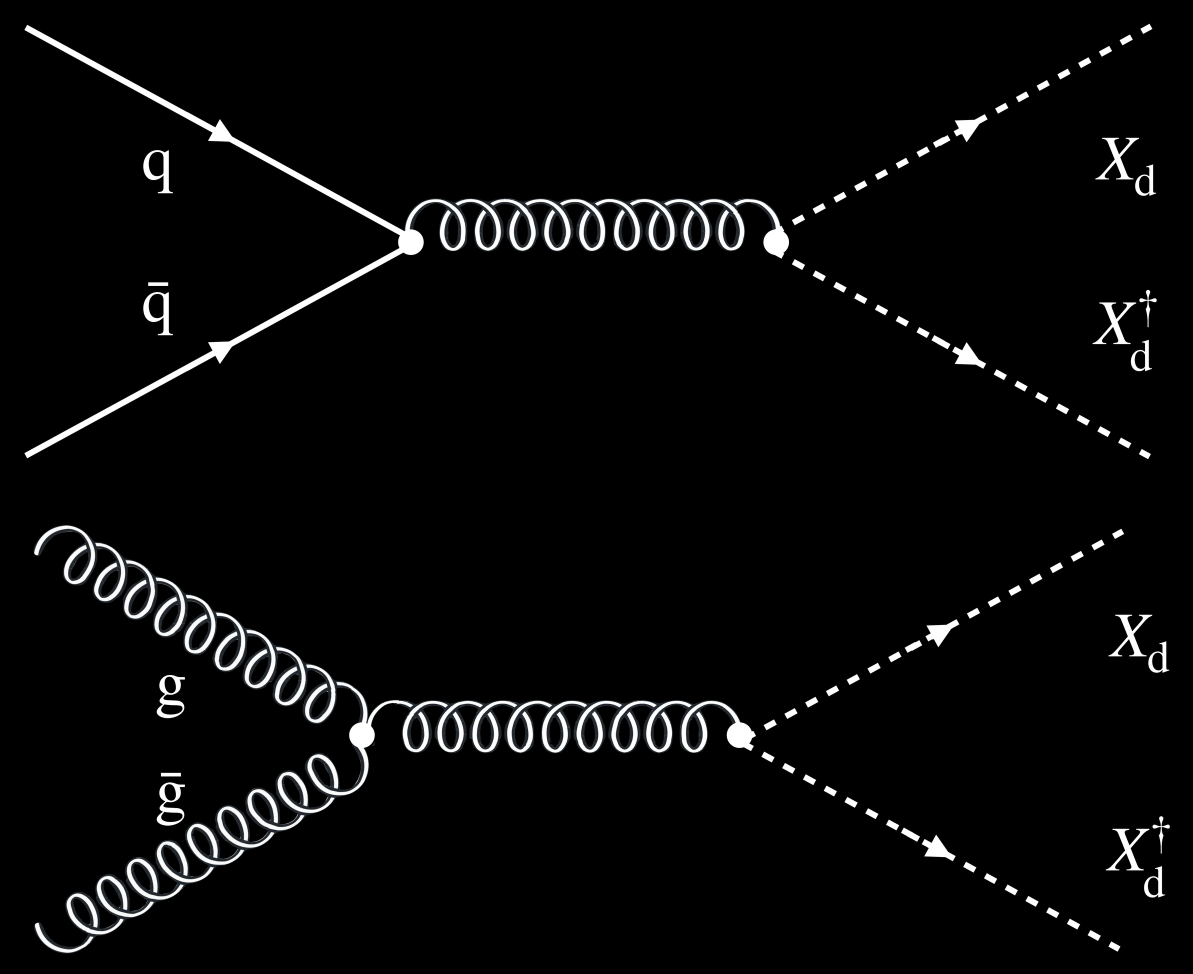
- Cold nuclear initial state (nPDF, small-*x*, …)
- Vacuum-like jet shower (no hot medium)

PbPb



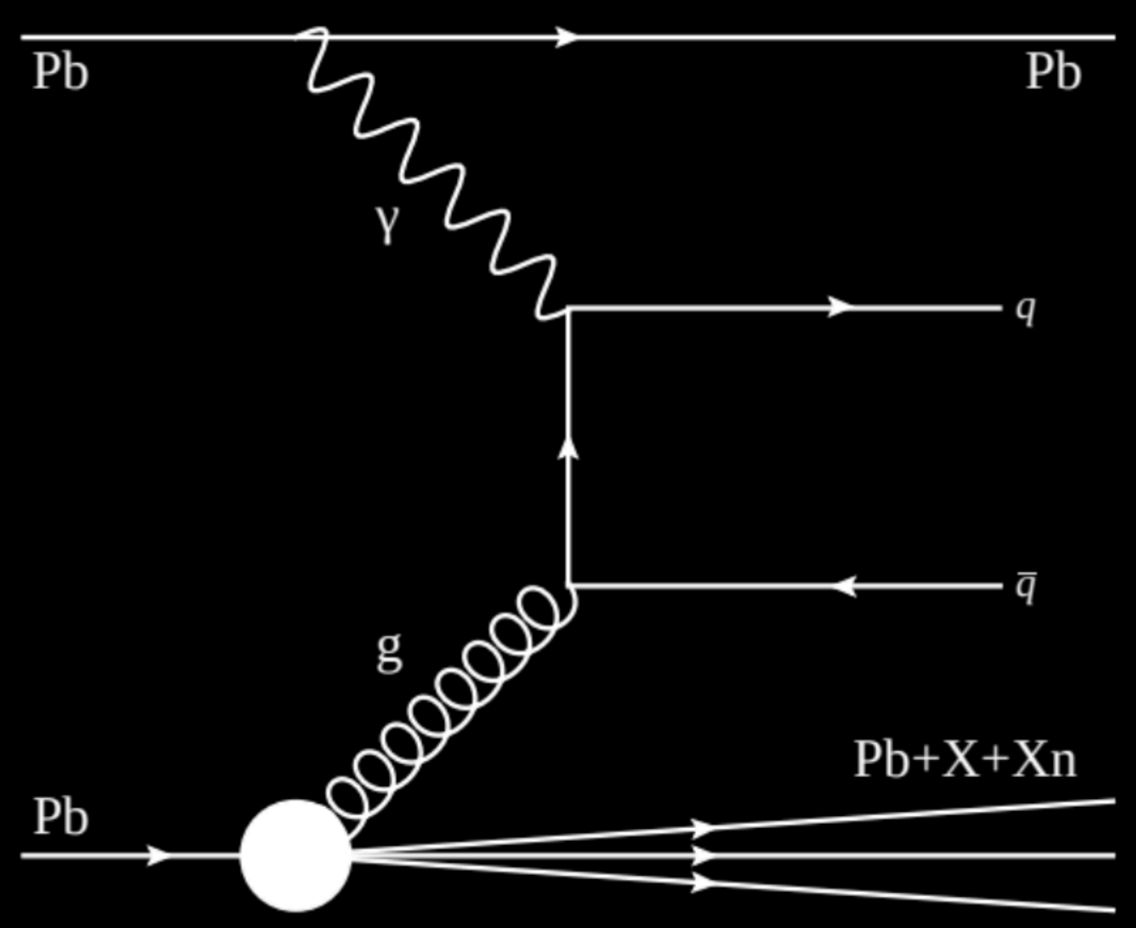
- Cold nuclear PDFs + hot QGP medium
- Strong jet quenching and substructure modification

pp



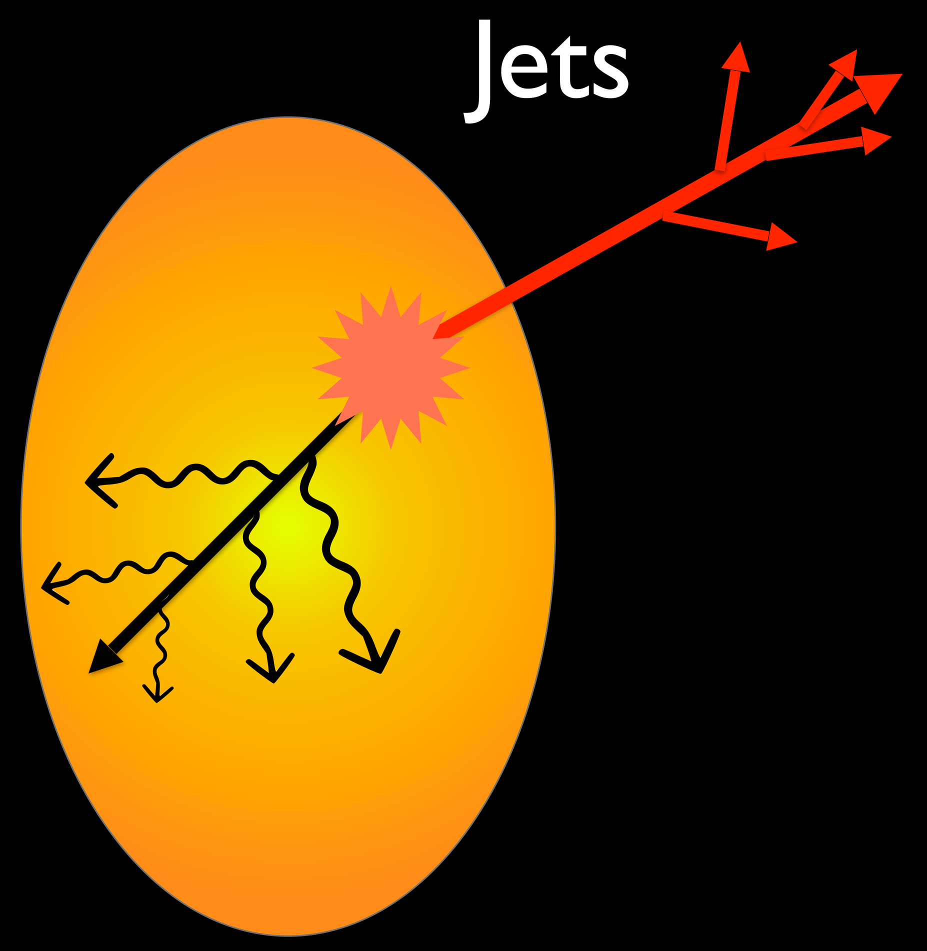
- Minimum-bias pp case, hard scattering and jet evolution in (almost) pure vacuum

γPb



- Cold nuclear initial state (nPDF, small-*x*, …)
- Vacuum-like jet shower (no hot medium)

PbPb



- Cold nuclear PDFs + hot QGP medium
- Strong jet quenching and substructure modification

- UPC γPb substructure provides a clean cold-nuclear baseline between pp (vacuum) and Pb+Pb (cold + hot, quenching) for jet fragmentation observables

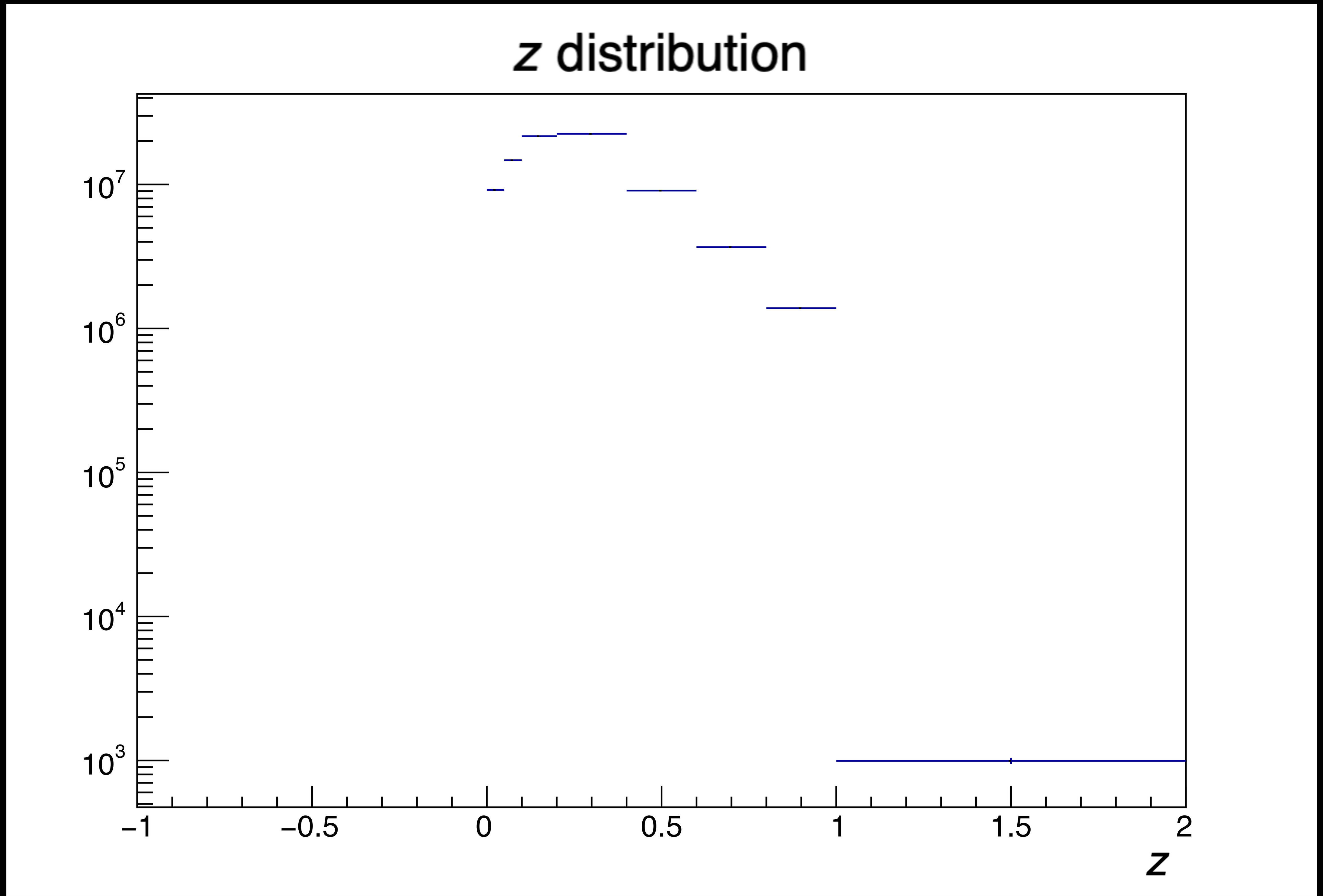
Summary & Outlook

ALICE

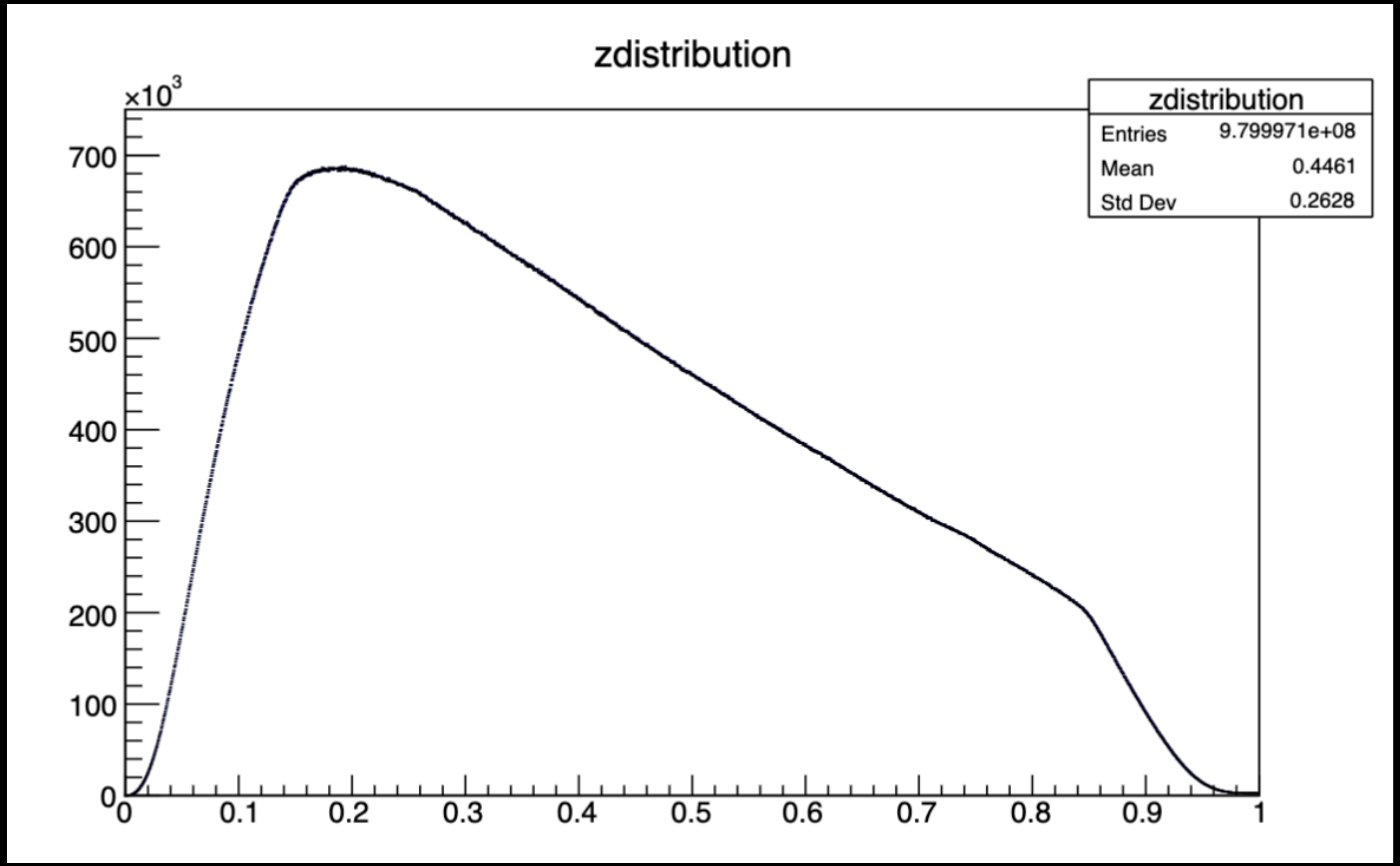
- ALICE has been investigating the medium-induced jet quenching effect in a small system with various observables
- Through the recent ALICE measurements, there was a difference between MB and HM pp but from the EA selection bias
  - This bias should be taken into account for the further small system jet-medium interaction studies
- There were no clear differences in jet substructure variables between pp and p−Pb by far
- Outlook
  - Results of these recent studies and more are coming with new ALICE papers
  - Stay tuned for further researches includes multiplicity-dependent j_{T} and UPC jet fragmentation with ALICE Run 3 data

Thank you!

Back-up slides

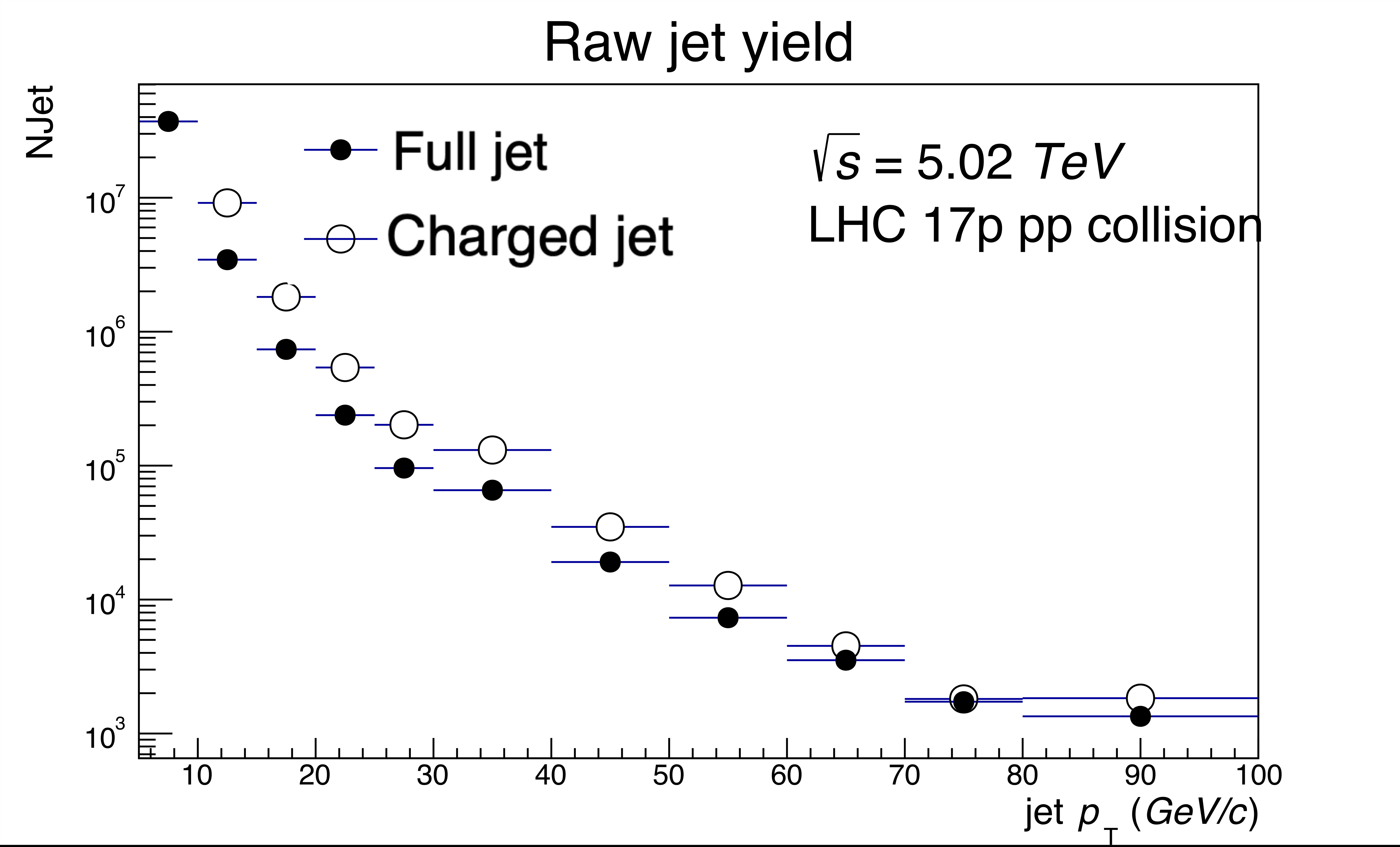


z distribution from LHC 17p FAST
(Binning used in this analysis)
(This analysis)



z distribution from LHC 17pAll
(Finely binned)
(Previous analysis)

Advantage of using charged jet



- Much more charged jets than full jets due to the limited acceptance of ALICE EMCal
- Acceptance of Tracking detector ($|\eta| < 0.9$, $0 < \phi < 2\pi$), EMCal ($|\eta| < 0.7$, $\Delta\phi = 110^\circ$)
- More differential study can be done in various z bin

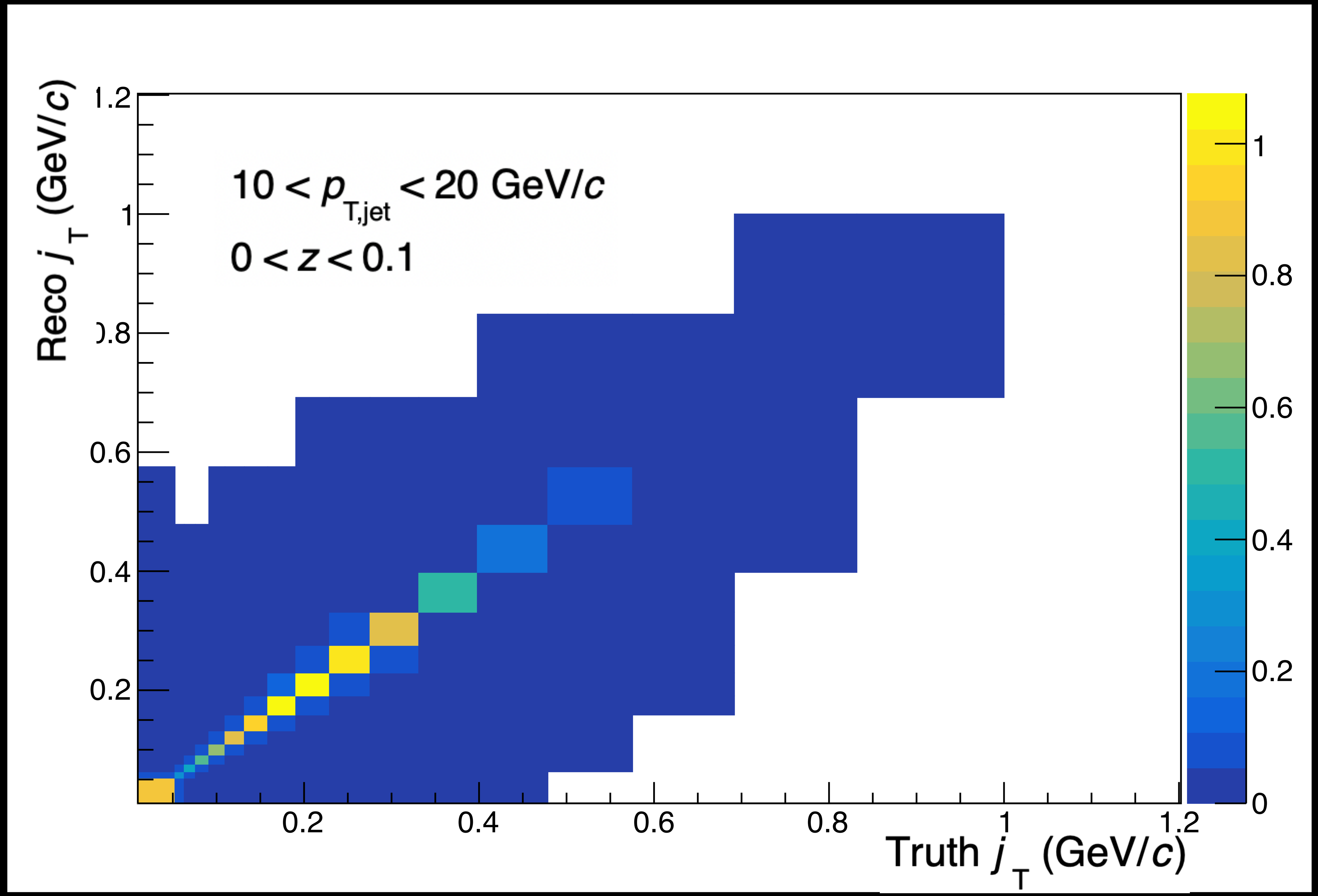
Event selection

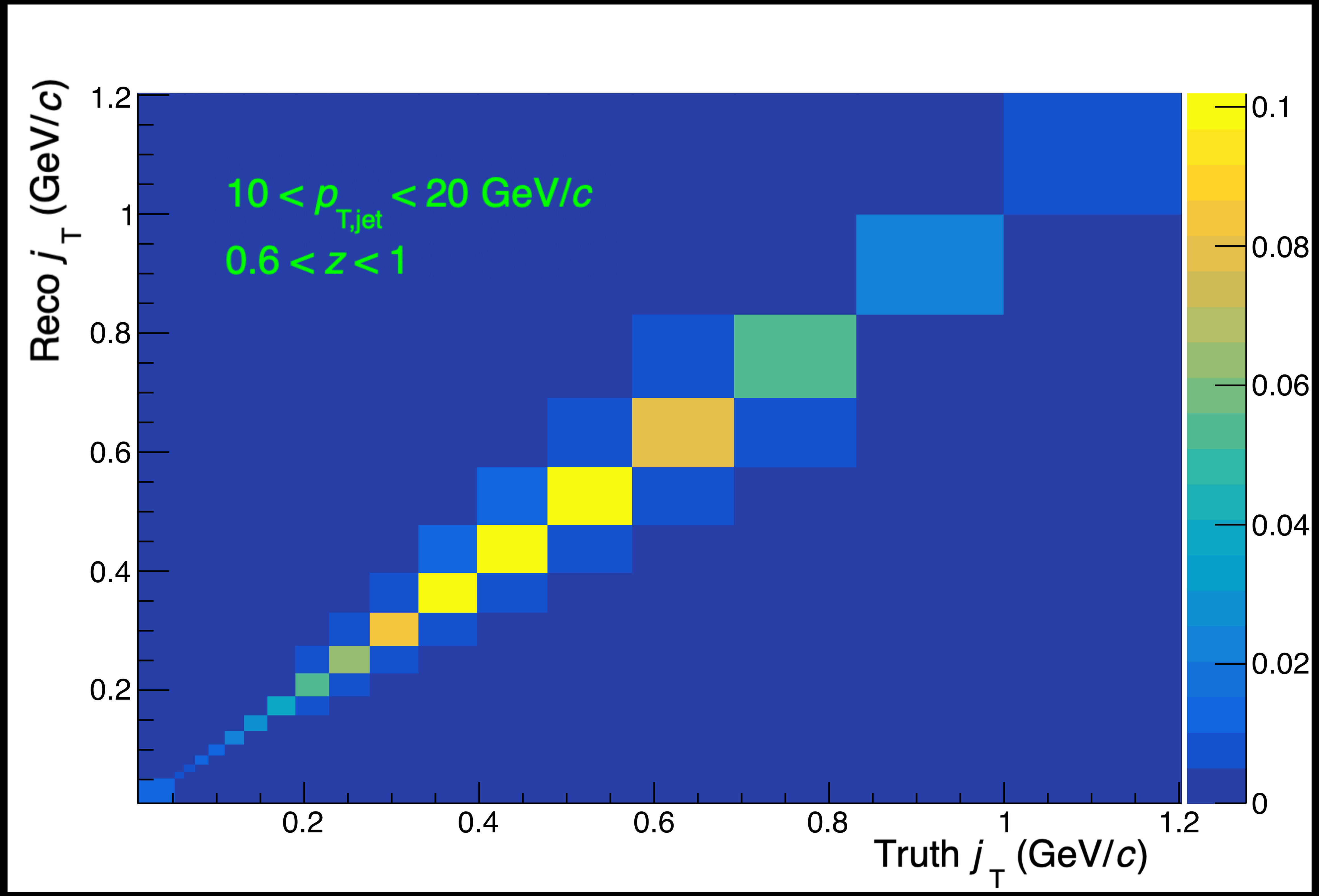
Cuts	Parameter range
Trigger	kINT7
# of vertex contributors	> 0
Z_{vtx}	<10 cm
$Z_{vtx} - Z_{vtx,\ SPD}$	<0.5 cm

Track selection - Hybrid tracks(filterbit 768)

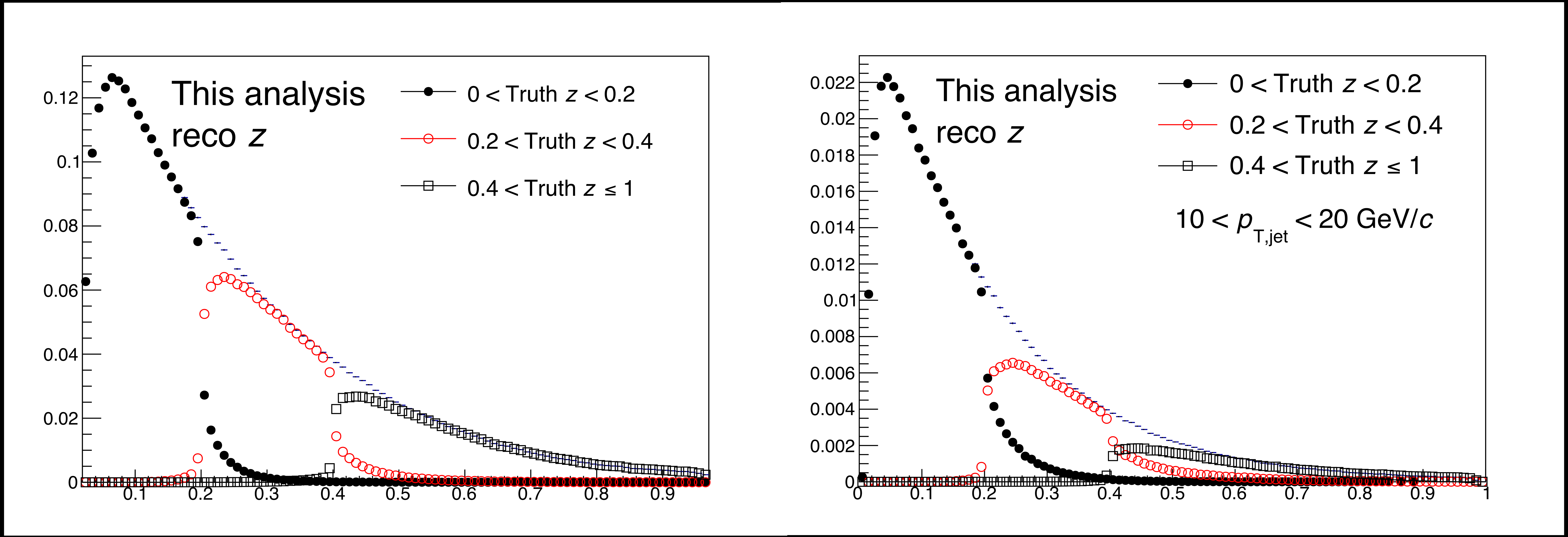
Cuts	**Parameter range**	**or Cuts**	**Parameter range**
DCA_z	< 3.2 cm	DCA_z	< 3.2 cm
DCA_r	< 2.4 cm	DCA_r	< 2.4 cm
# of minimum crossed rows in TPC	> 50	# of minimum crossed rows in TPC	> 50
Ratio of crossed rows over findable clusters in TPC Fraction of shared TPC clusters	> 0.8 < 0.4	Ratio of crossed rows over findable clusters in TPC Fraction of shared TPC clusters	> 0.8 < 0.4
χ^2 per TPC cluster	< 4	χ^2 per TPC cluster	< 4
No kinks, TPC refit, ITS refit		Requirements of ITS refit	
At least one hit in the SPD		No requirements on the SPD hits	
χ^2 per ITS cluster	< 36	χ^2 per ITS cluster	< 36

- Fake jets -> A reconstructed jet is not matched
  ($j_{\mathrm{T}}^{obs}, p_{\mathrm{T}}^{obs}, z^{obs}$) are set as fake
- Fake tracks -> For detector level tracks that have no corresponding particle level
  ($j_{\mathrm{T}}^{obs}, z^{obs}$) are set as fake
- Missing jets -> A generated jets that are not matched to any reconstructed jet
  ($j_{\mathrm{T}}^{true}, p_{\mathrm{T}}^{true}, z^{true}$) are set as miss
- Missing -> When do not find a corresponding track
  ($j_{\mathrm{T}}^{true}, z^{true}$) are set as miss





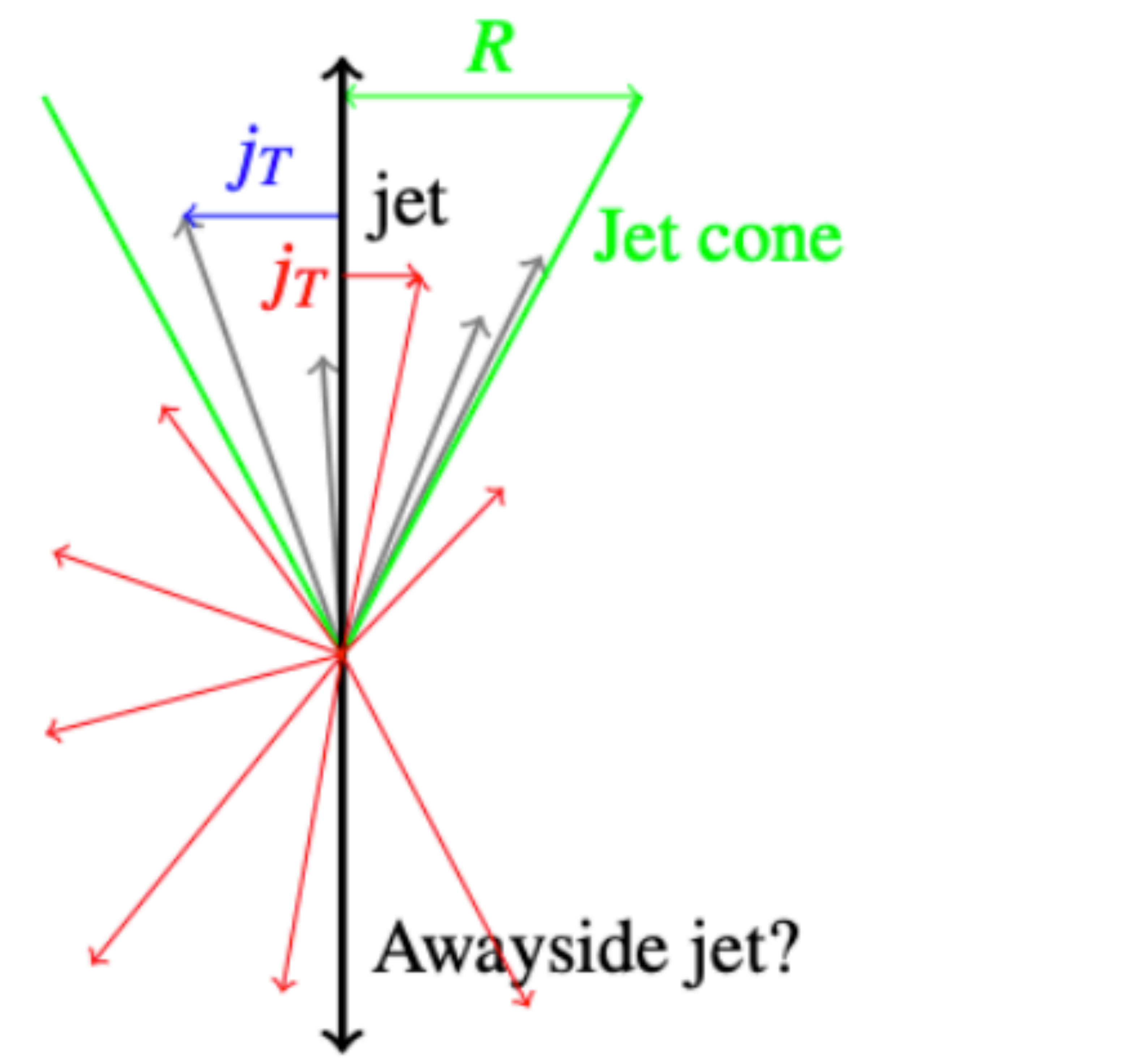
- j_T response with respect to z and jet p_T
- j_T distributions at each z bin are clearly different
- Thus, we need to consider z when unfolding j_T in various z bins

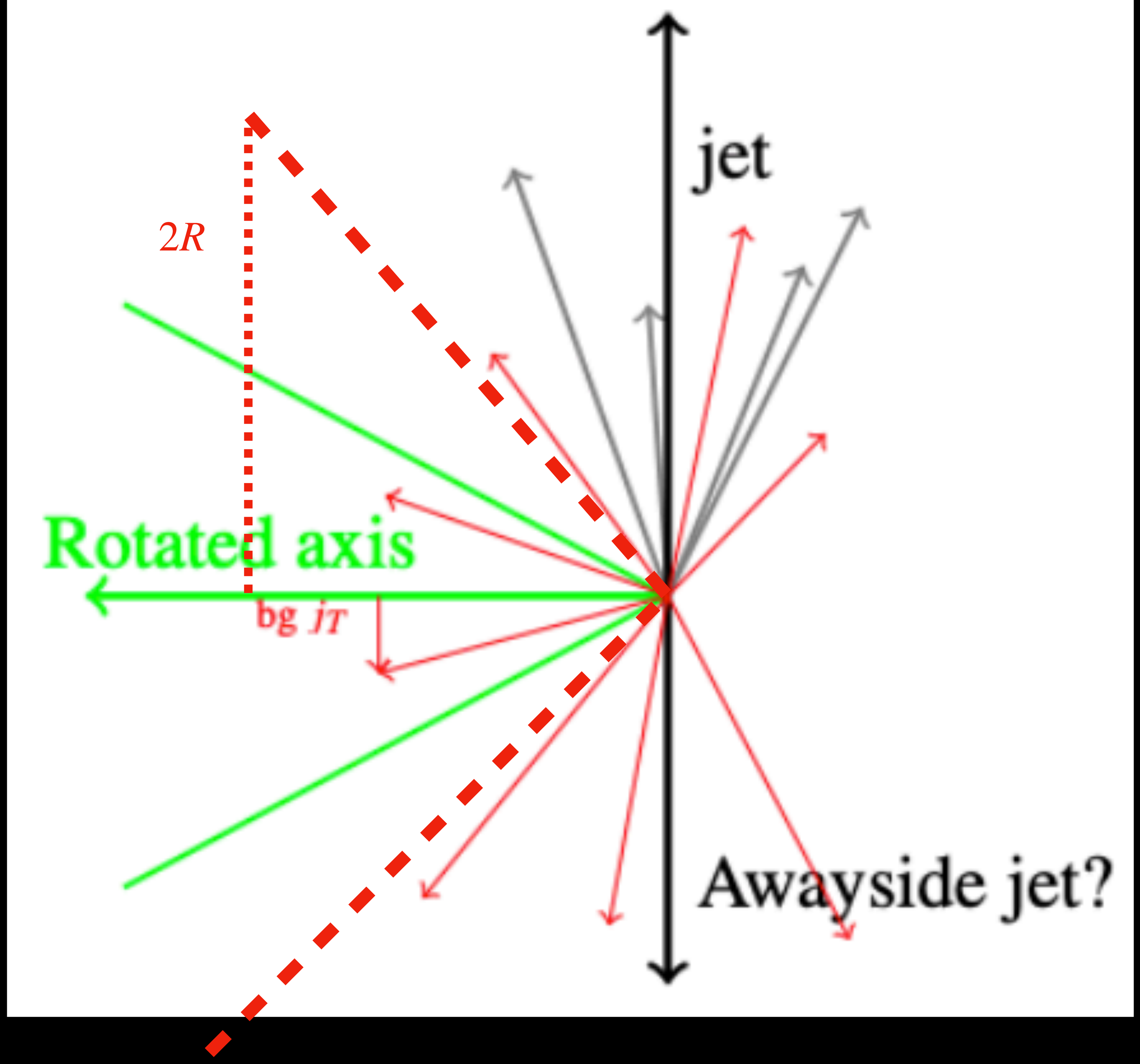


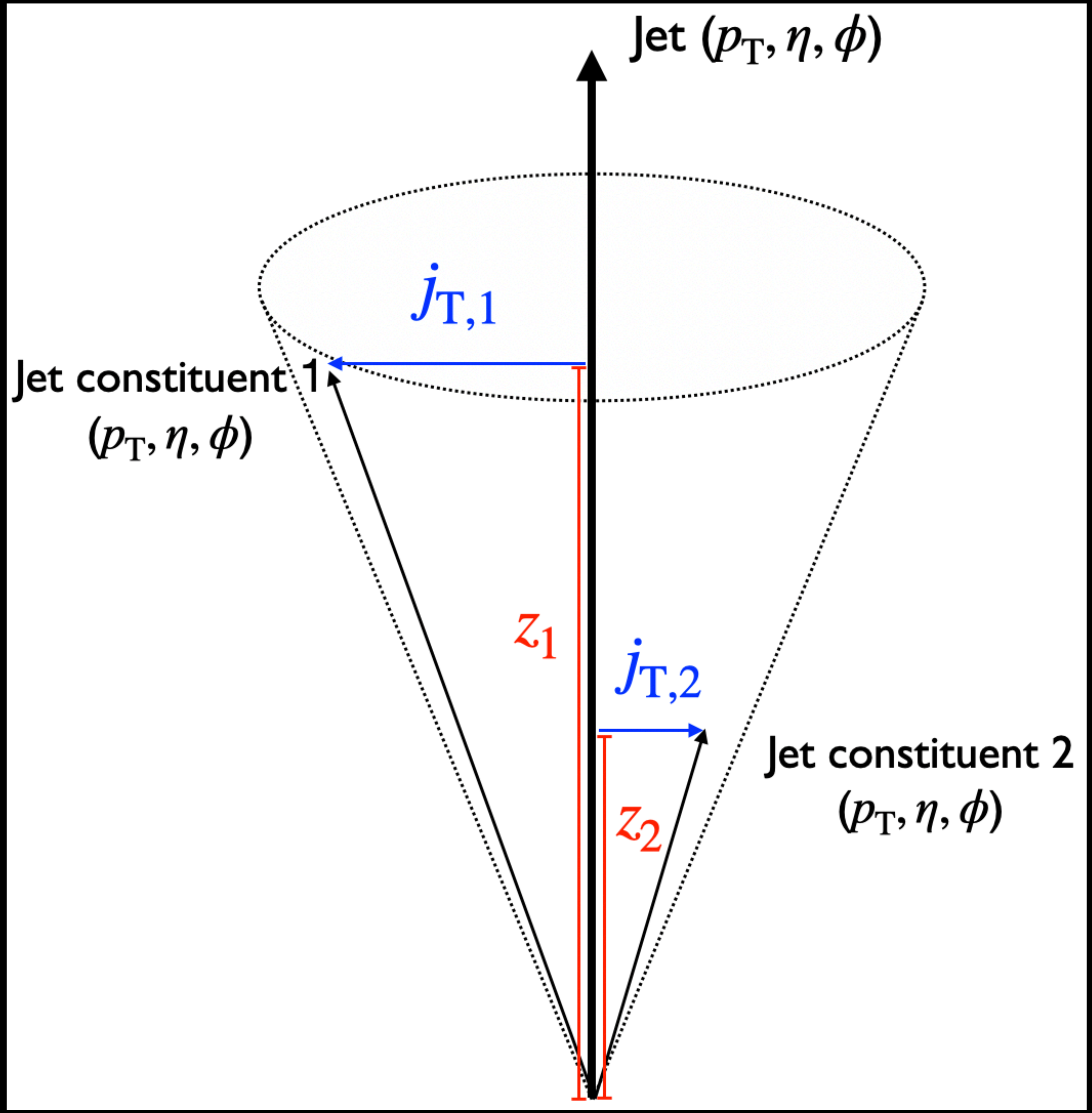
- Reconstructed MC z w.r.t. differential truth z
- Smearing effect can be seen in the distribution
- Thus, we need to consider z when unfolding j_T in various z bins

Background estimation

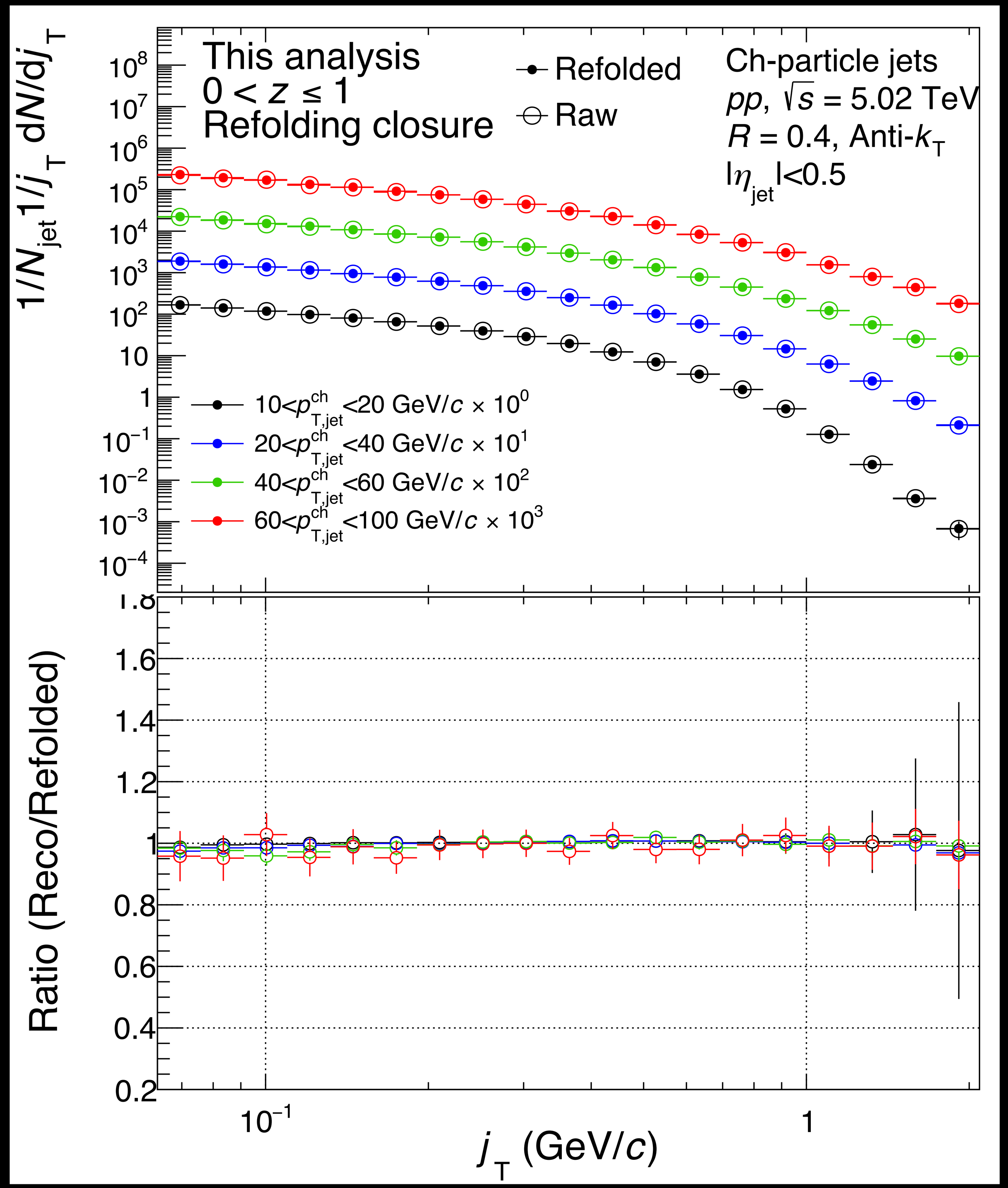
- Perpendicular cone (Default)
  -Rotate the jet axis by 90° in a positive ϕ direction
  - If there are no other reconstructed jets around the rotated axis(Delta R<0.8), calculate j_T, z w.r.t the rotated axis
- j_T, z calculated with a perpendicular cone method was unfolded separately
- Used random background method for systematic check



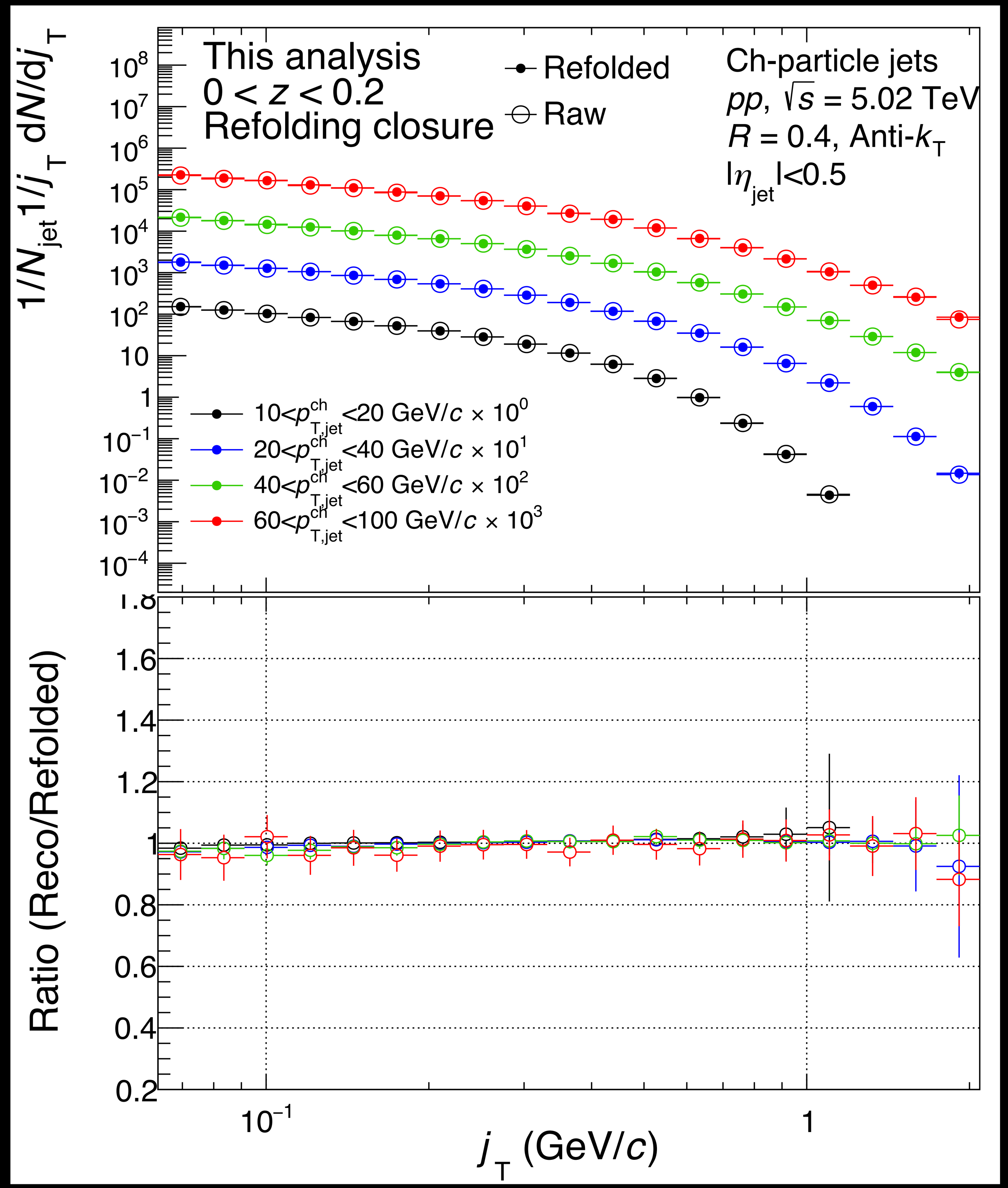




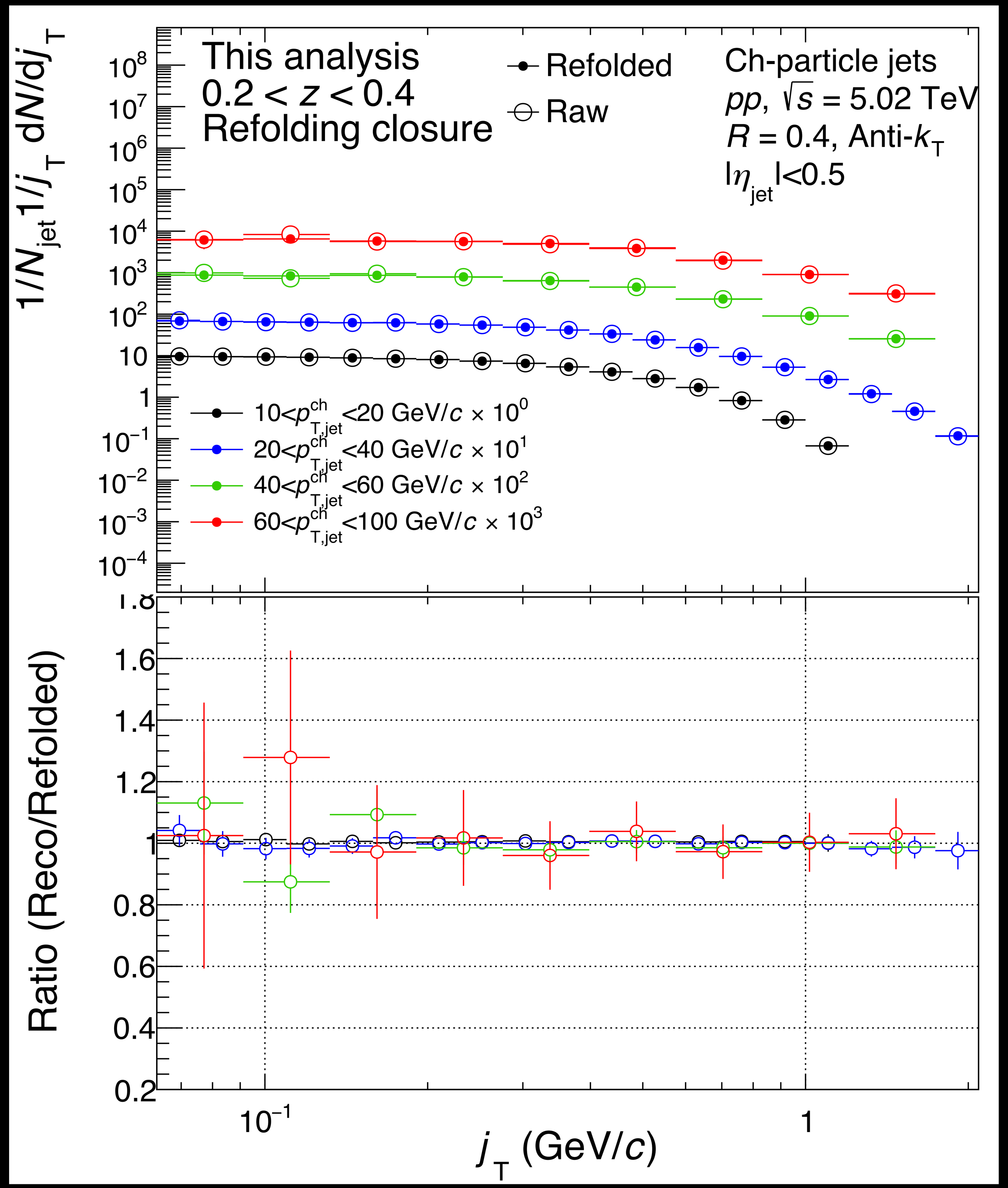
Jet $(p_{\mathrm{T}}, \eta, \phi)$
$j_{\mathrm{T},1}$
Jet constituent 1
$(p_{\mathrm{T}}, \eta, \phi)$
z_1
$j_{\mathrm{T},2}$
Jet constituent 2
$(p_{\mathrm{T}}, \eta, \phi)$
z_2



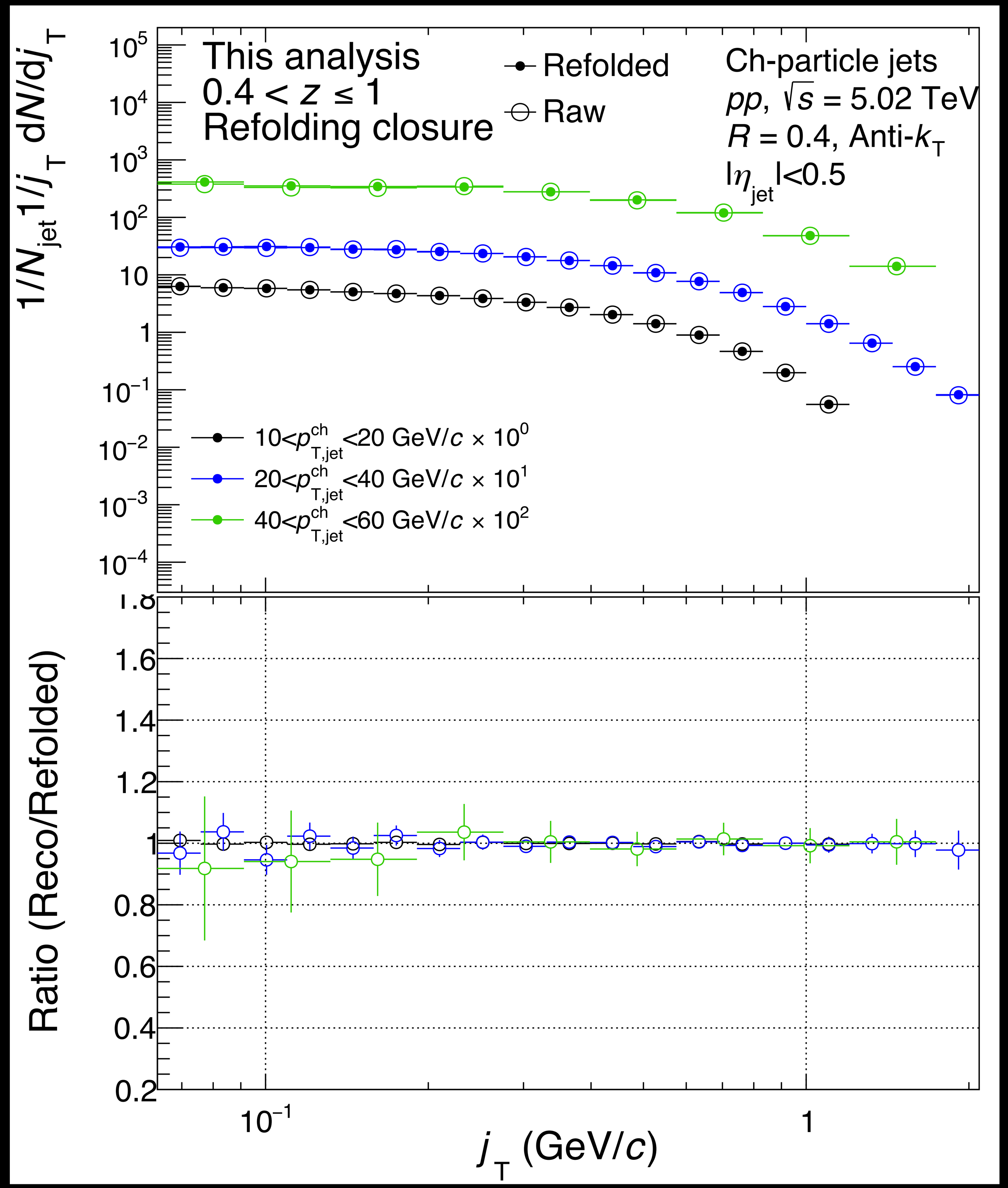
Inclusive z (0 < z ≤ 1)



Low z (0 < z < 0.2)



Mid z (0.2 < z < 0.4)



High z (0.4 < z ≤ 1)

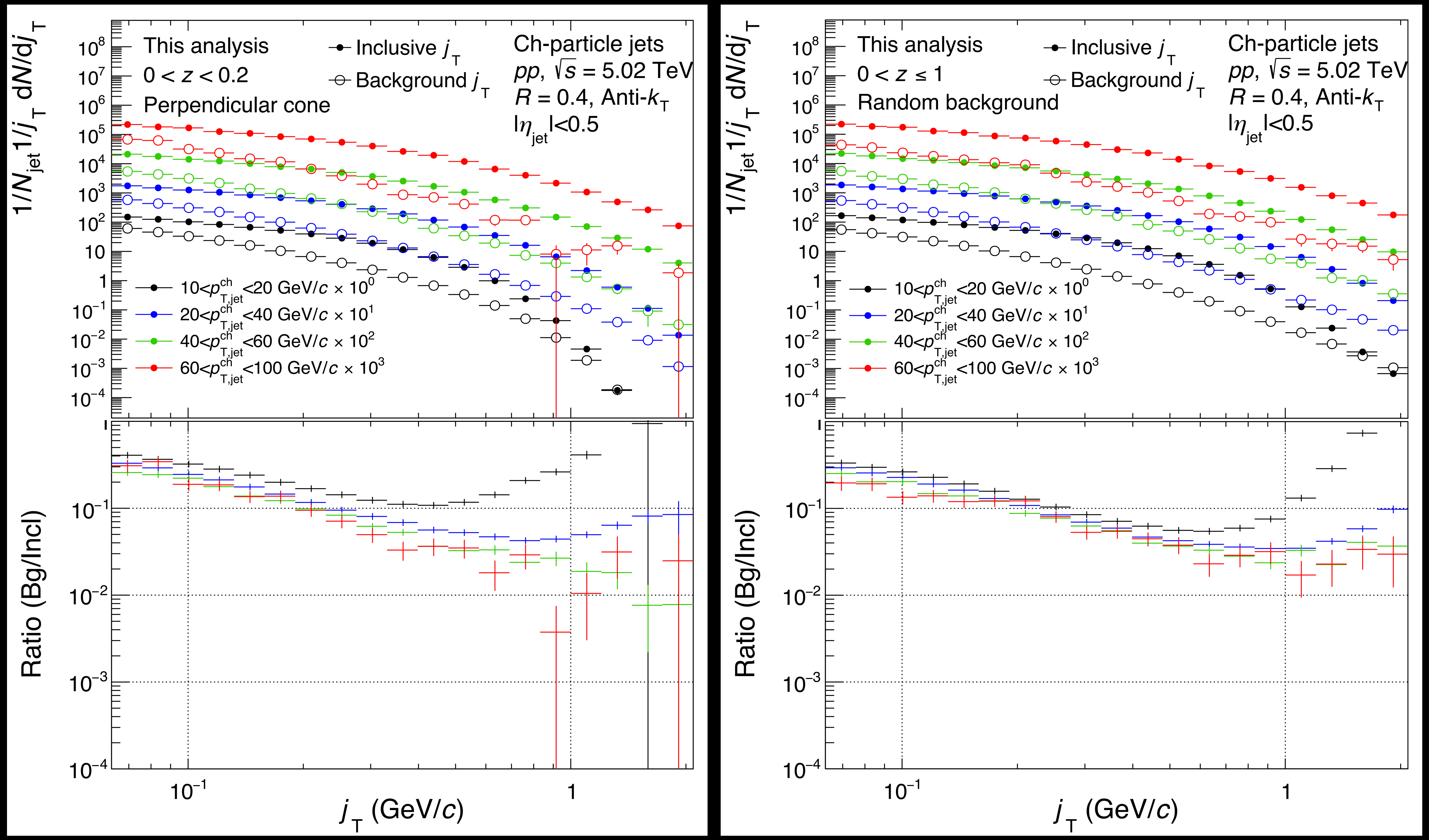
Background subtraction

- Auto-correlations in random background method could affect the small system j_T backgrounds
- Thus, we removed the auto-correlations
- Now the Bg fractions of perp and random gets more similar

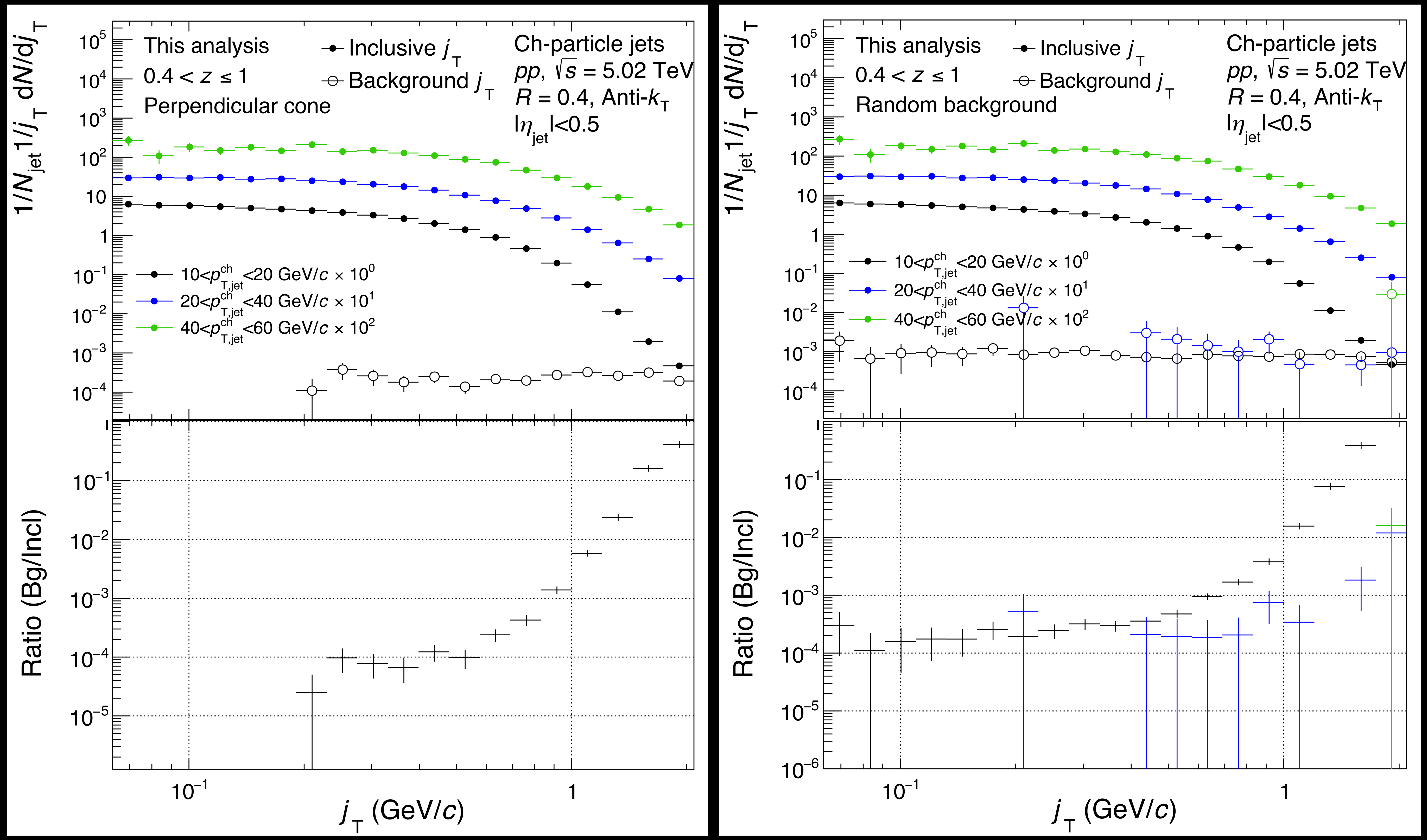
$0 < z < 0.2$ $\quad$ $0.4 < z \leq 1$

Background subtraction

- Auto-correlations in random background method could affect the small system j_{T} backgrounds
- Thus, we removed the auto-correlations
- Now the Bg fractions of perp and random gets more similar



$0 < z < 0.2$



$0.4 < z \leq 1$

Rebinning

- Re-binned distributions in the response matrix
- Bin numbers are increased two times
- Default bin and changed bin are as below

```
auto binsjetpt = AxisVar("binjetpt", {5,10,15,20,25,30,40,50,60,70,80,100,500})
auto binsz = AxisVar("zbin", {-1,0,0.05,0.1,0.2,0.4,0.6,0.8,1,2});
```

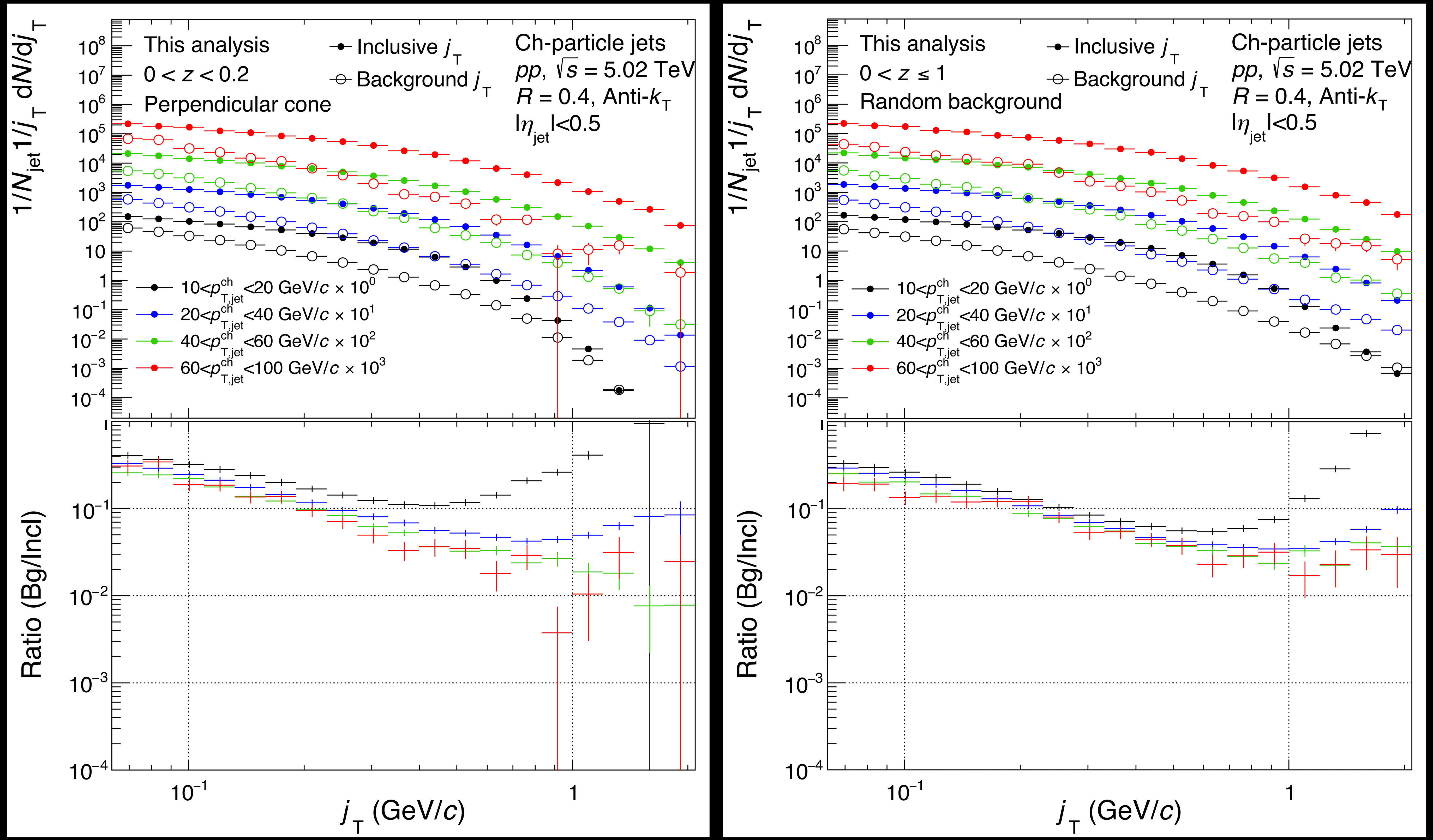
```
auto rebinsjetpt = AxisVar("binjetpt", {5,7, 10,12,15,17,20,23,25,27,30,35,40,45,50,55,60,65,70,75,80,90,100,200,500});
auto rebinsz = AxisVar("zbin", {-1,0,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0,1.1,1,2});
```

– j_{T} default : 0.01,0.0524807,0.0630957,0.0758577,0.0912006,0.109648,0.131825,0.15849,0.190546, 0.229087,0.275423,0.331131,0.398108,0.47863,0.575439,0.691831,0.831764,0.999998,1.20226 ,1.44544,1.7378,2.08929,2.51189,3.01996,3.63078,4.36516,10

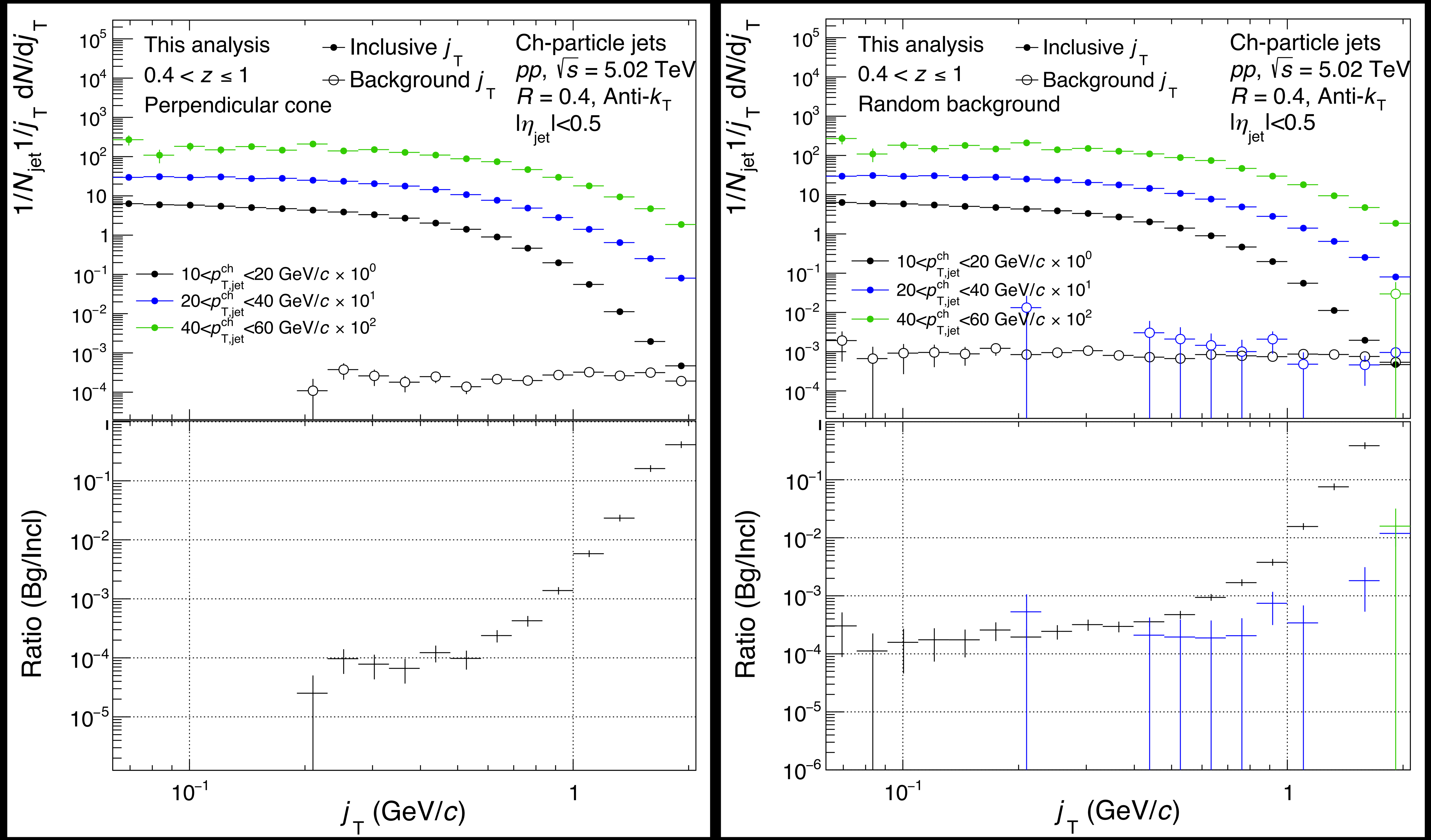
– j_{T} changed : 0.01,0.03,0.0524807,0.057,0.0630957,0.069,0.0758577,0.082,0.0912006,0.0959, 0.109648,0.12,0.131825,0.14,0.15849,0.17,0.190546,0.205,0.229087,0.25,0.275423,0.3,0.331131, 0.36,0.398108,0.436,0.47863,0.52,0.575439,0.63,0.691831,0.76,0.831764,0.91,0.999998,1.1,1.20226, .3,1.44544,1.6,1.7378,1.9,2.08929,2.34,2.51189,2.8,3.01996,3.33,3.63078,3.96,4.36516,5.7,10

Background subtraction

- Background fraction of each background subtraction method



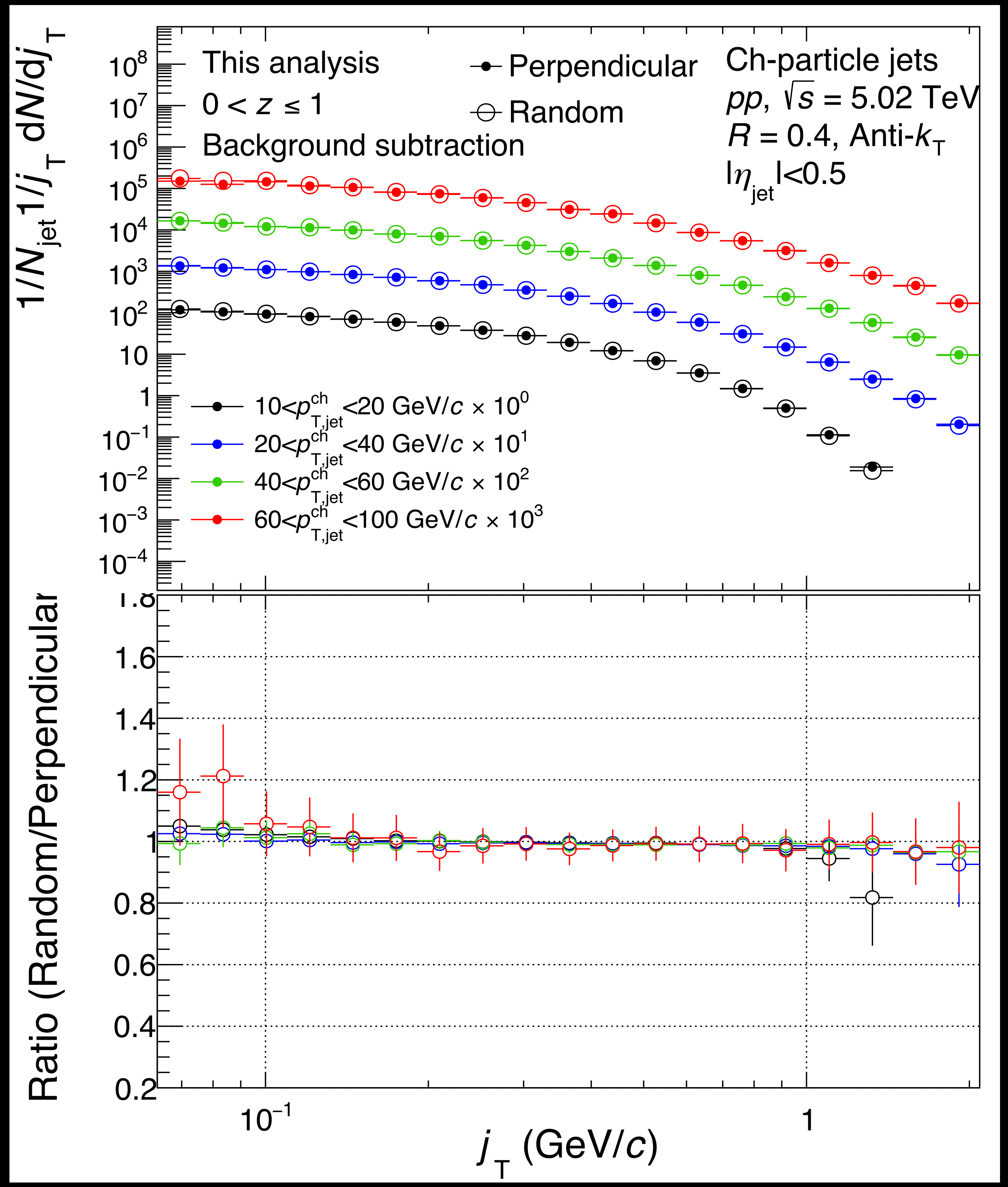
$0 < z < 0.2$



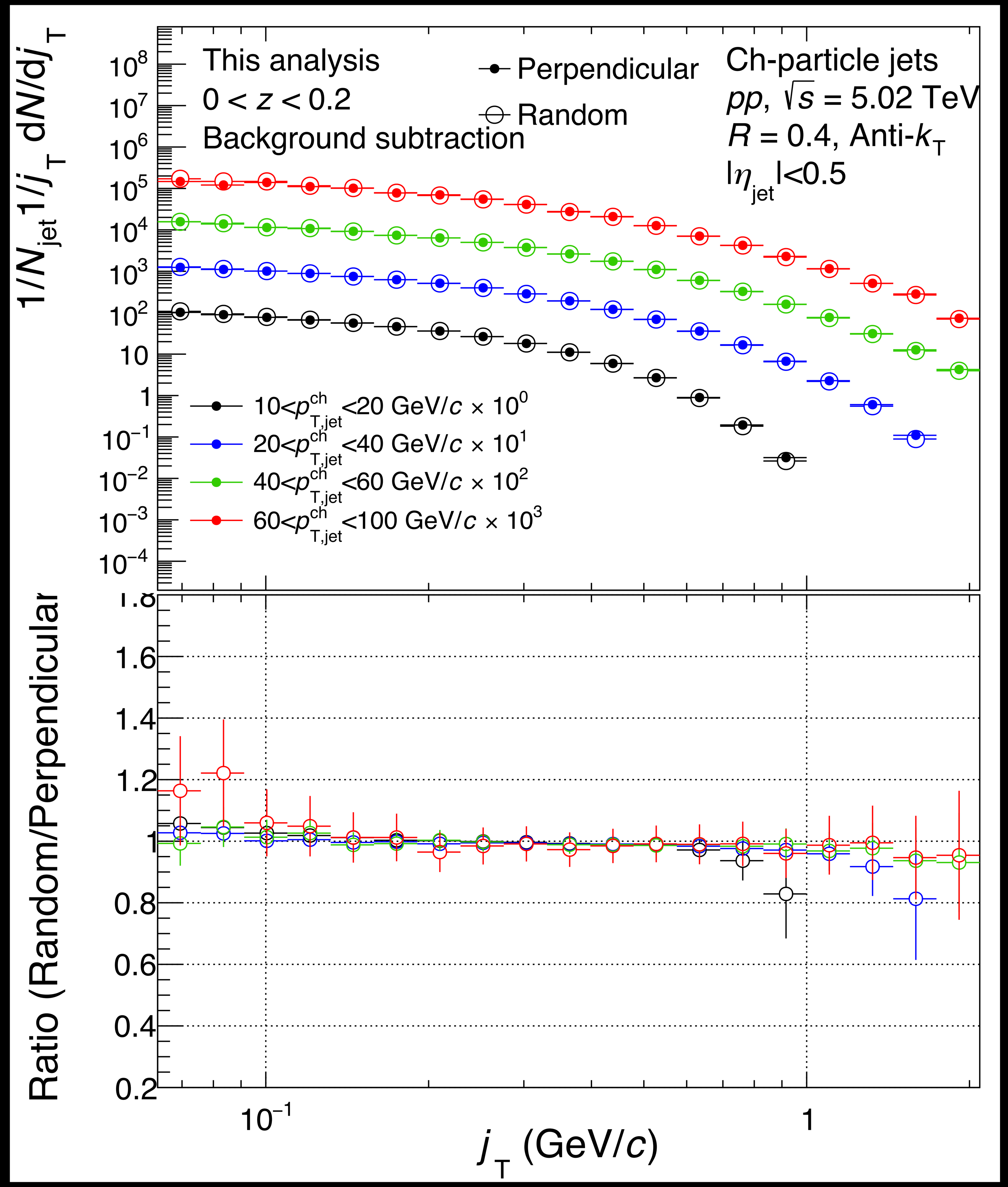
$0.4 < z \leq 1$

Background subtraction

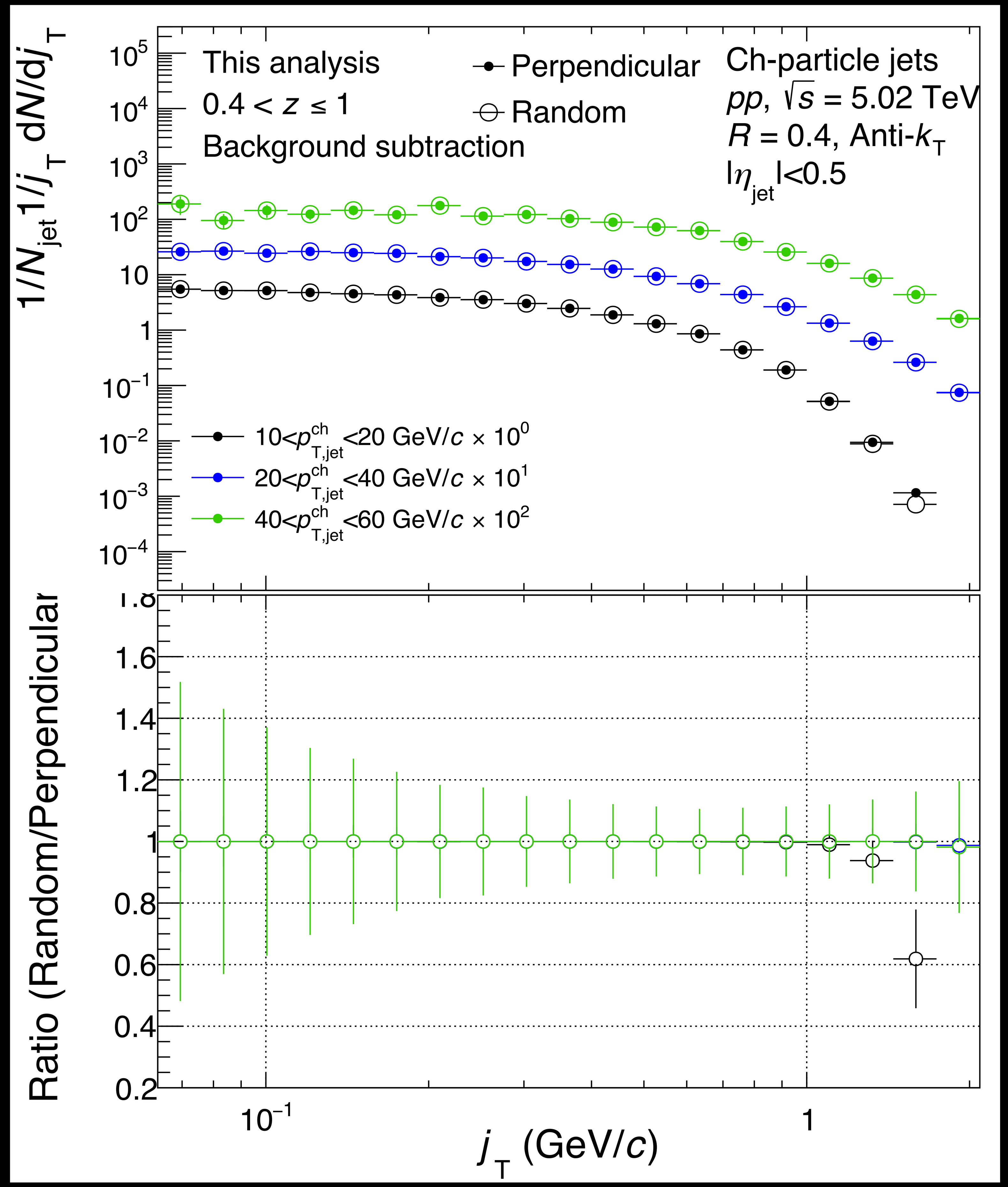
- Systematic errors from different background subtraction methods.
- Perpendicular cone method as the default and random background method used to extract systematic error values



Inclusive z



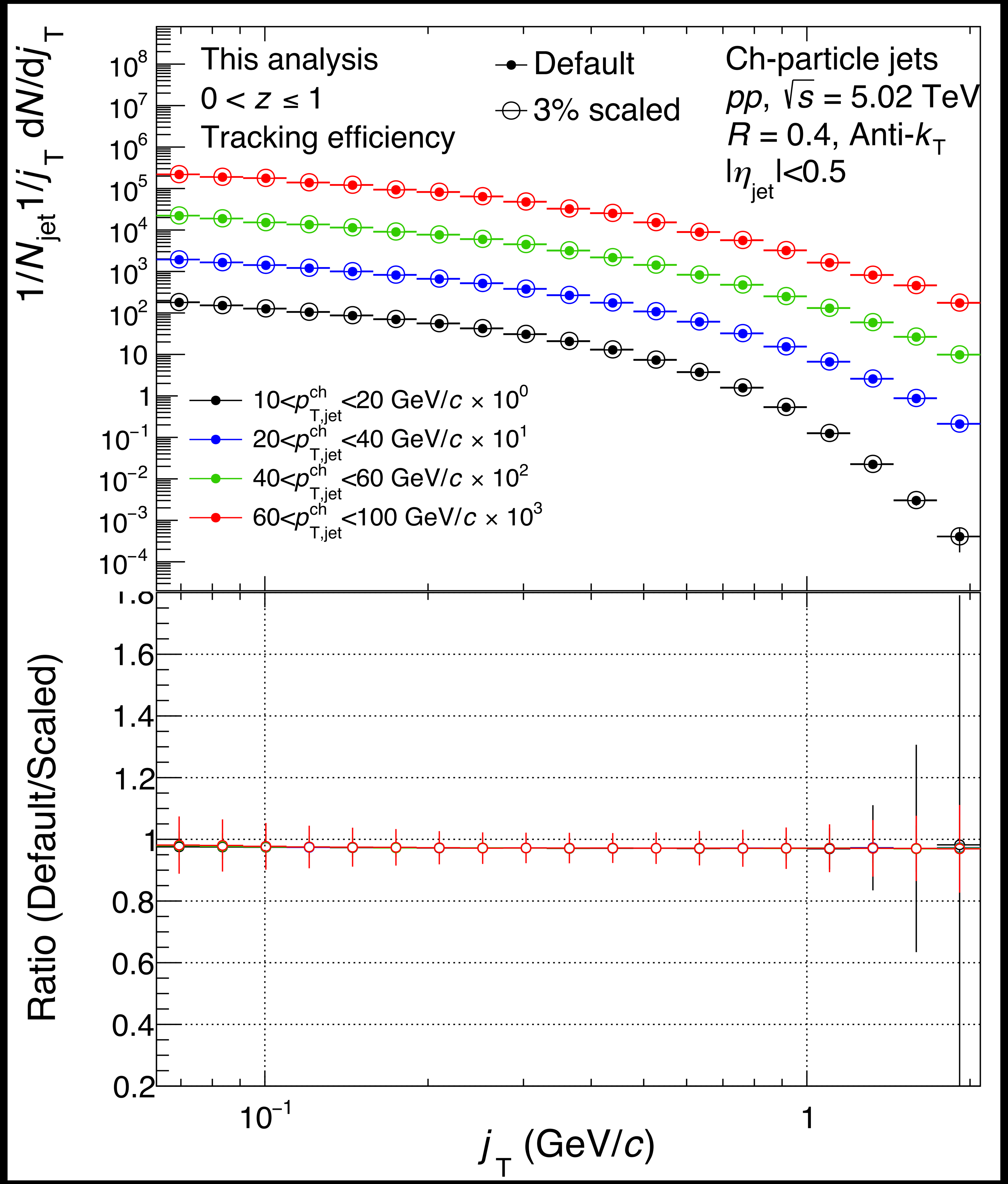
$0 < z < 0.2$



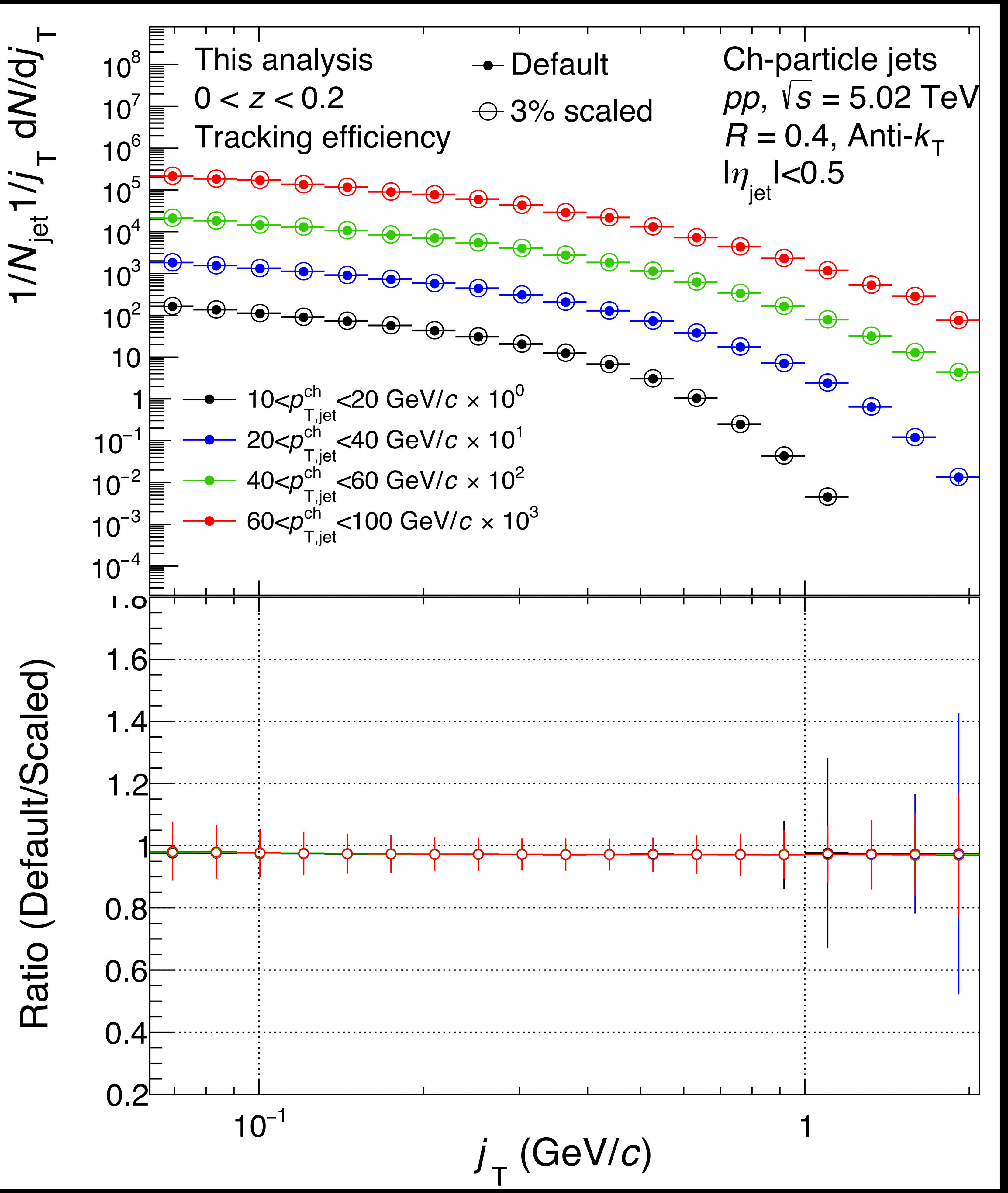
$0.4 < z \leq 1$

Tracking efficiency

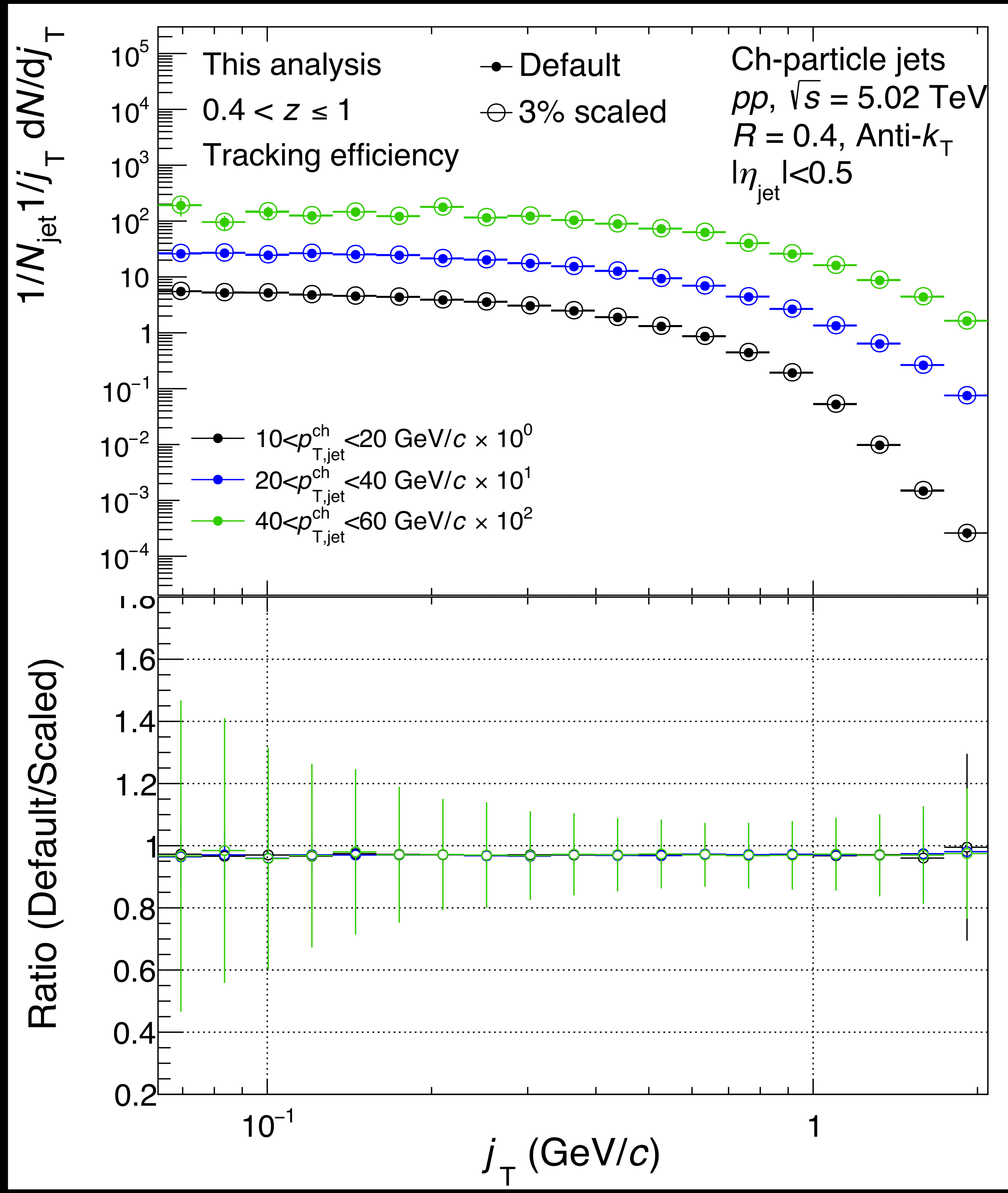
- We made MC set with a 3% reduced efficiency and compare it to default MC set



Inclusive z



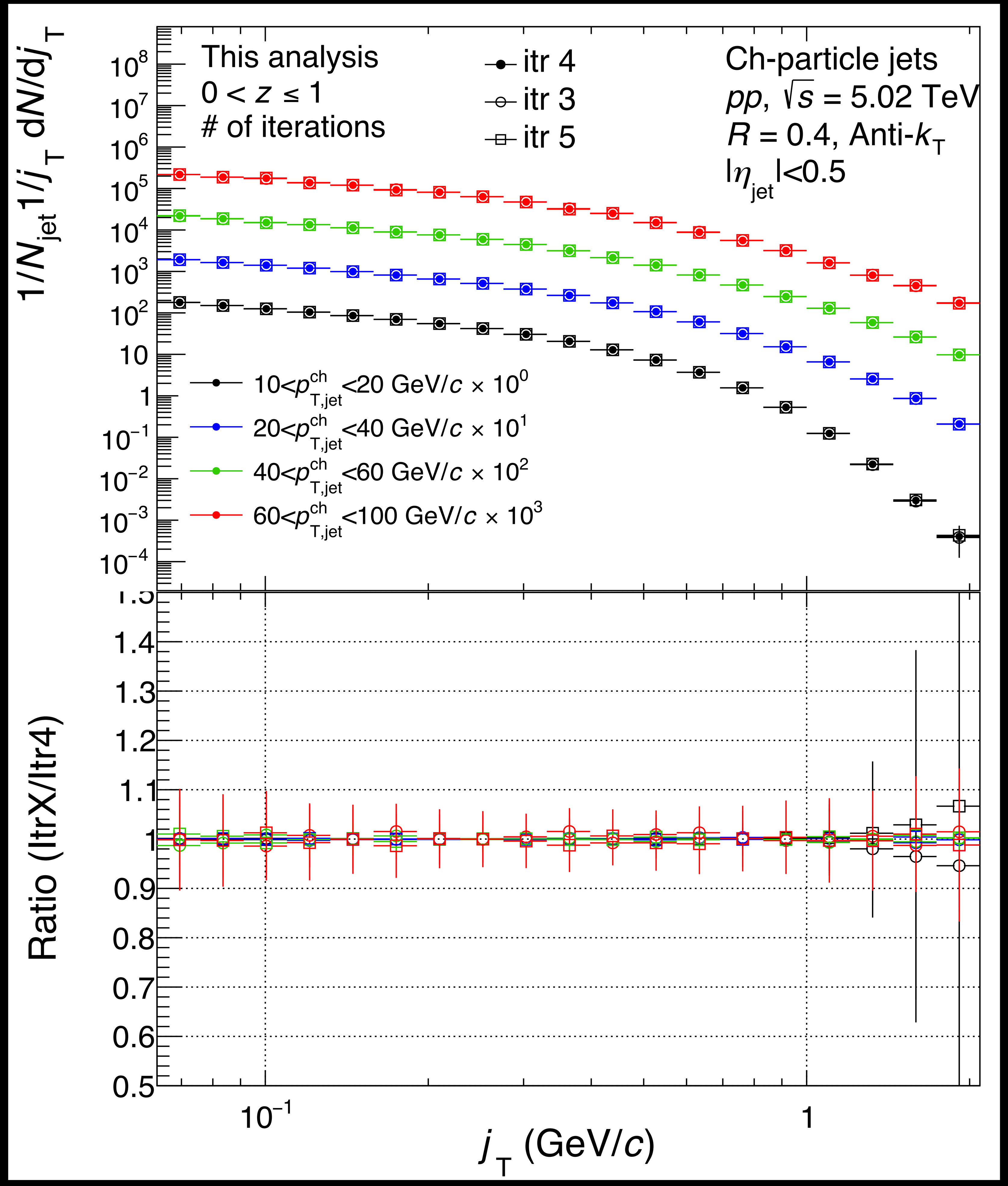
$0 < z < 0.2$



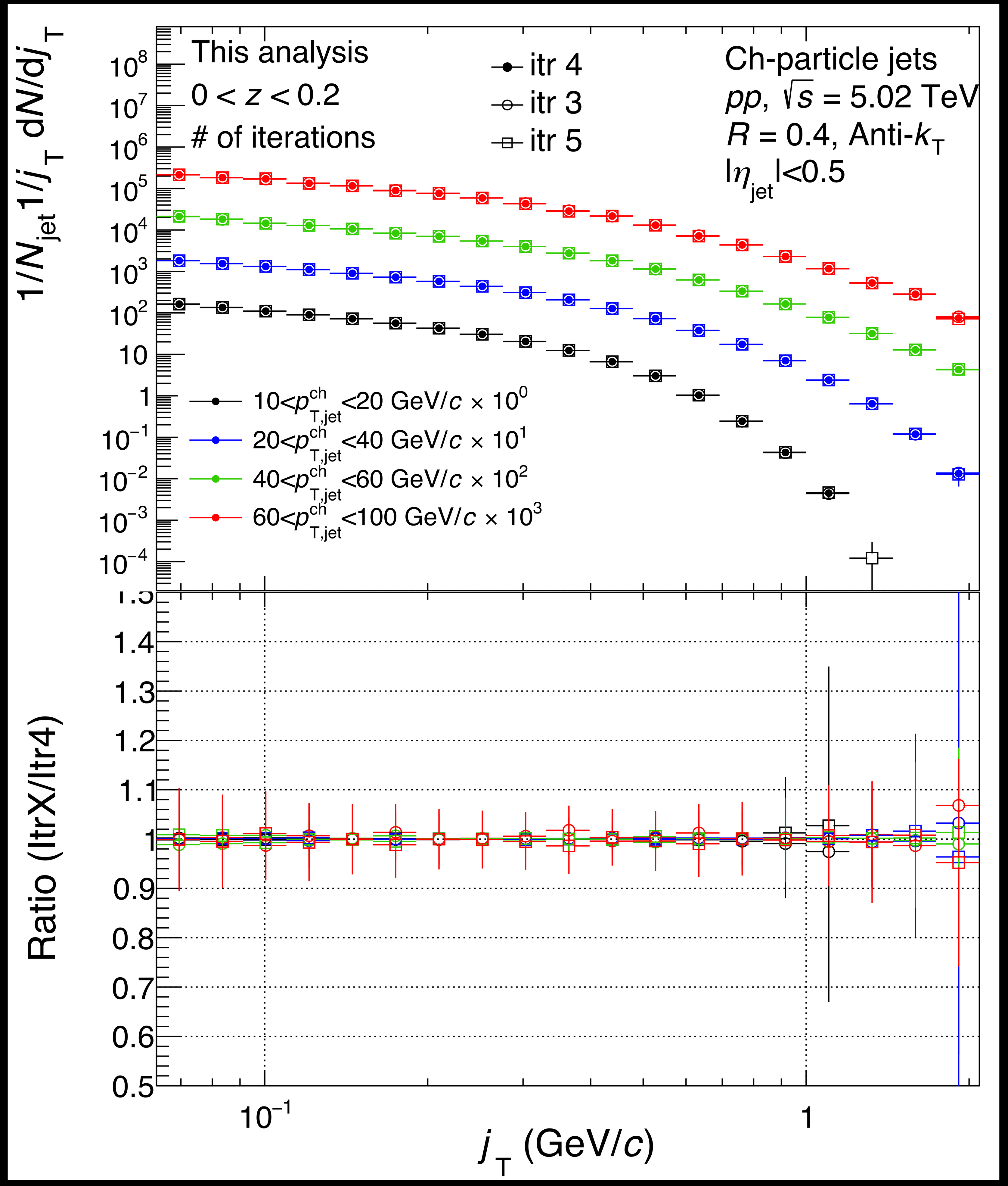
$0.4 < z \leq 1$

The number of iteration

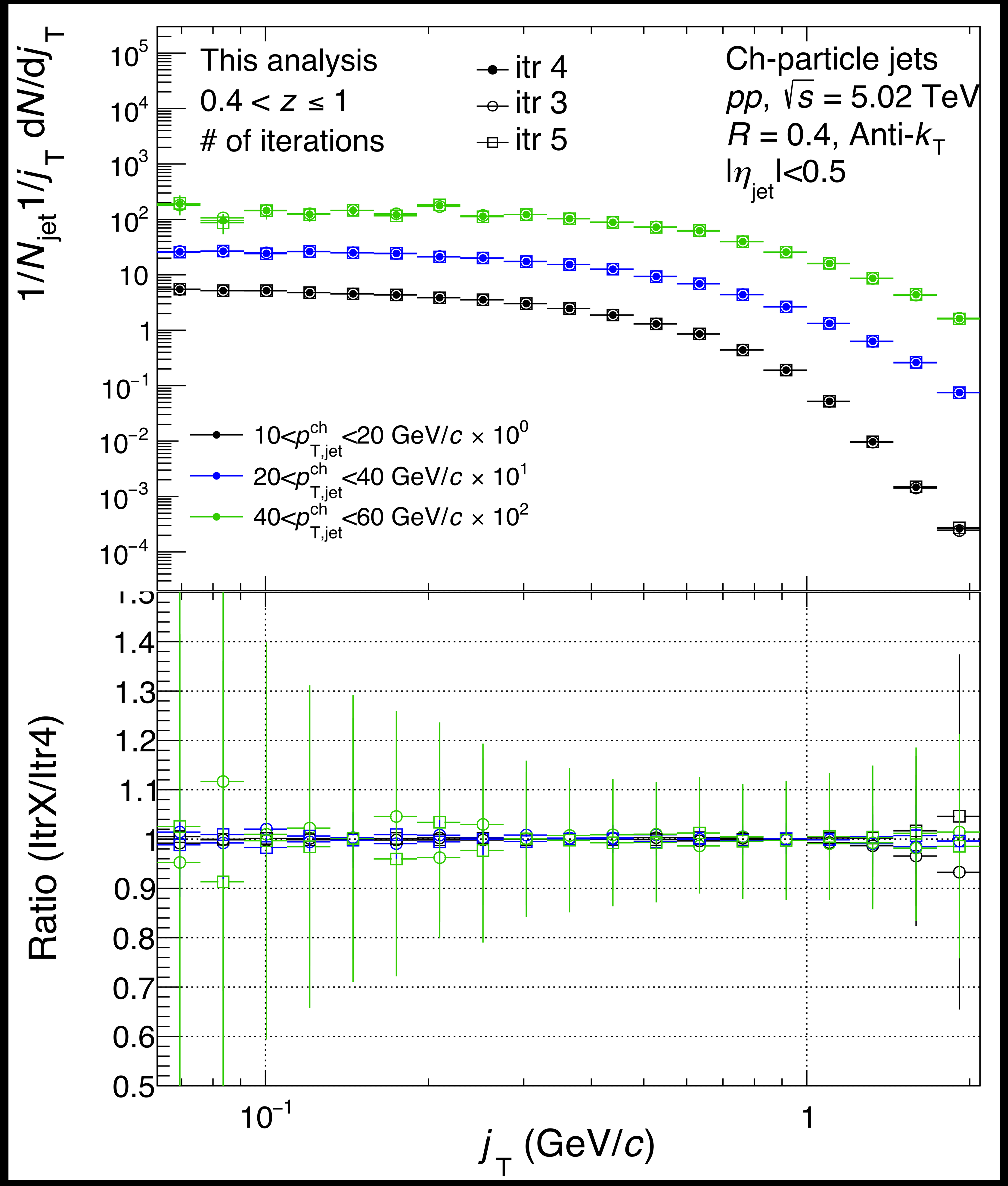
- Variation of the number of iterations
- Set default as 4 times of iterations and vary it from 3 to 5



Inclusive z



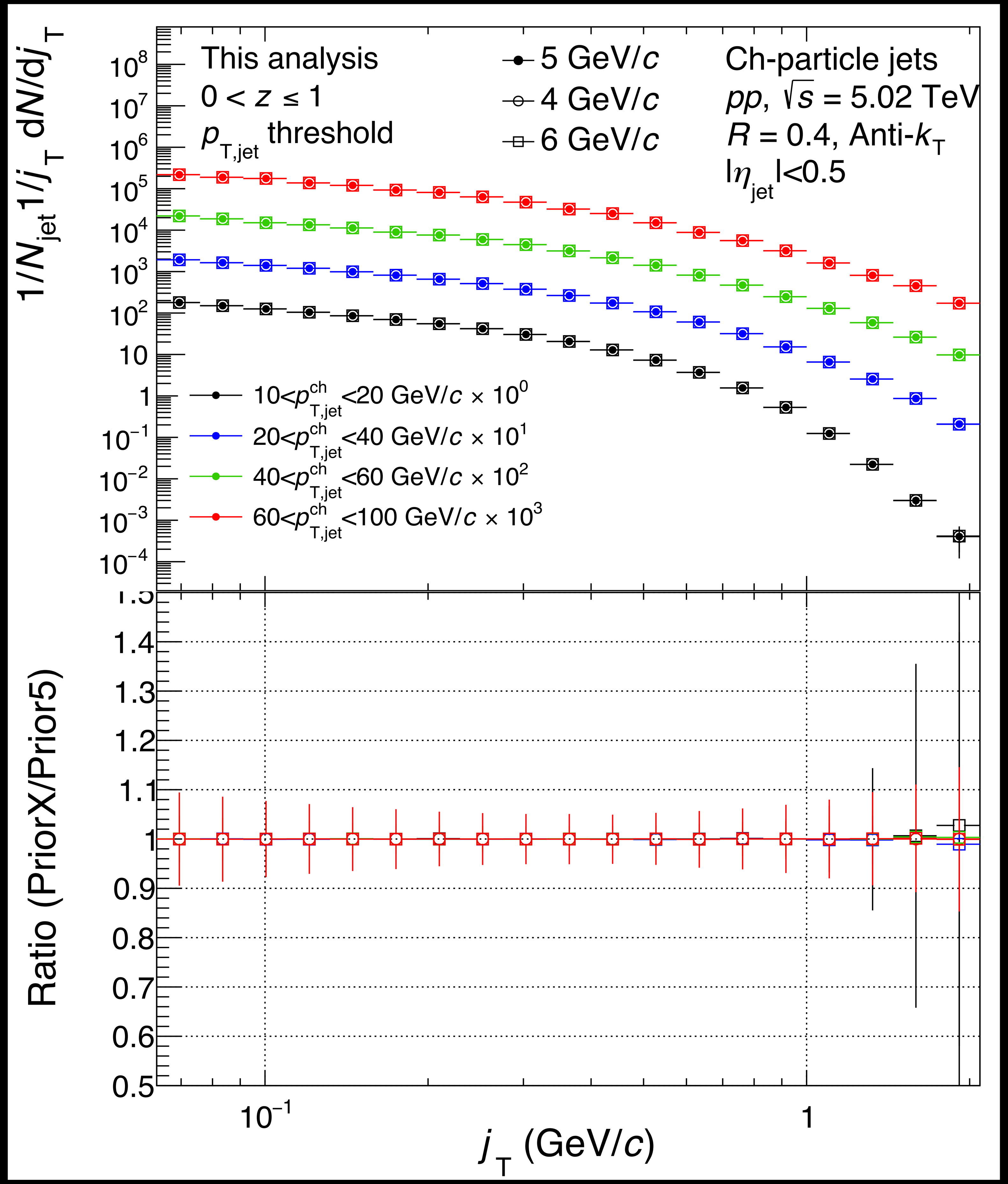
$0 < z < 0.2$



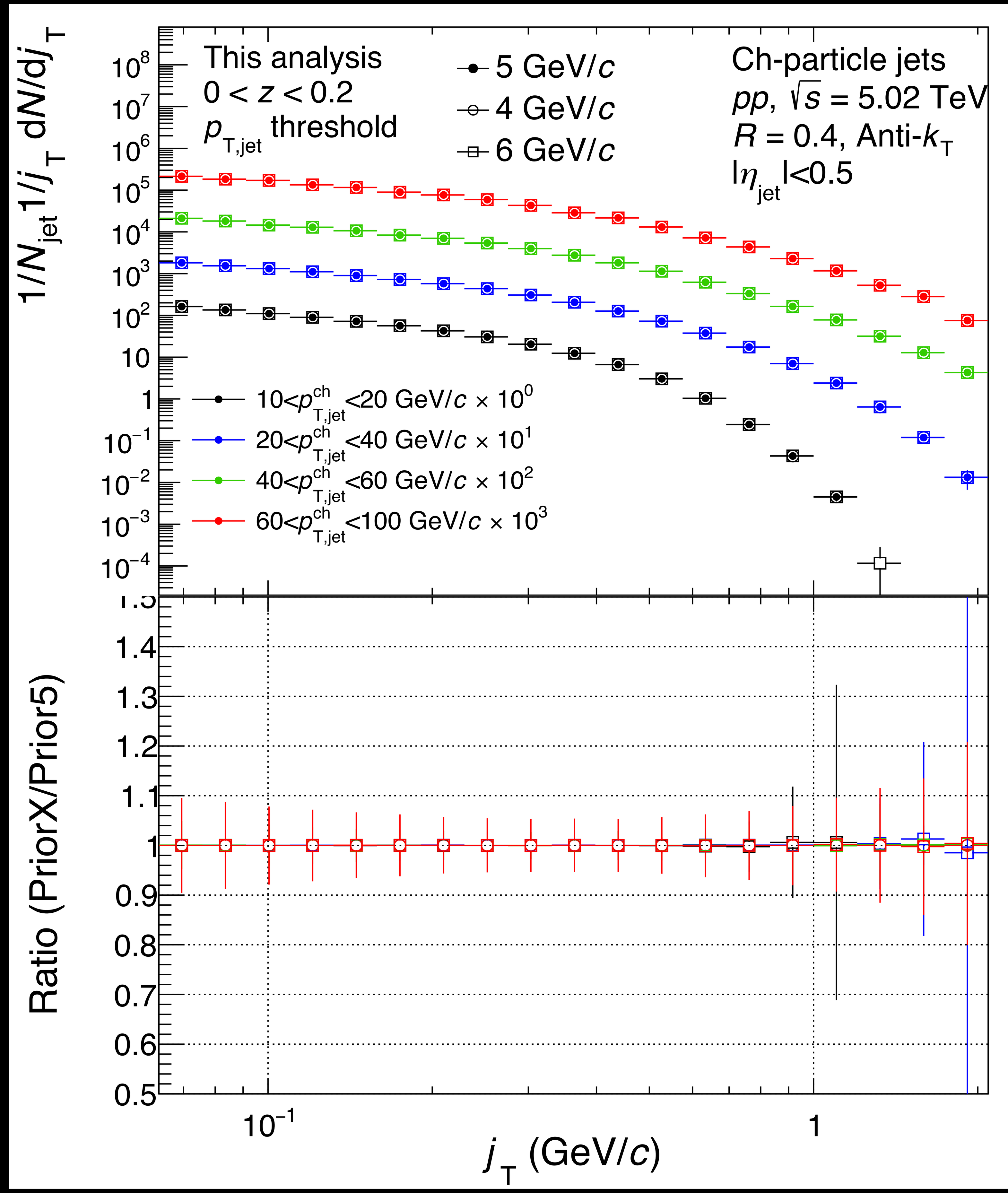
$0.4 < z \leq 1$

Jet p_T threshold

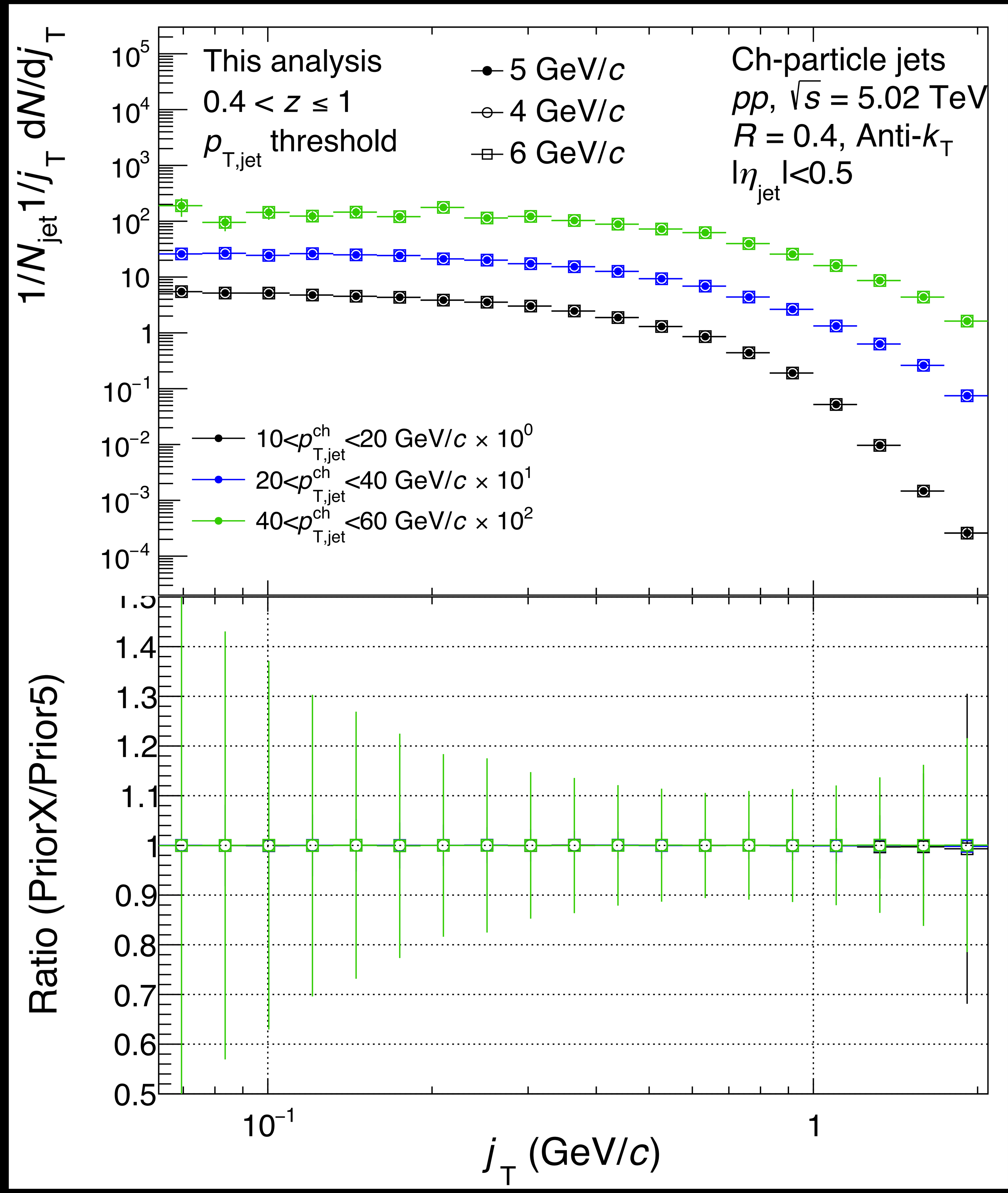
- Variations of the jet p_T threshold for the unfolding matrix
- Set 5 GeV/c as a default jet p_T threshold and varies it from 4 to 6 GeV/c



Inclusive z ($0 < z \leq 1$)



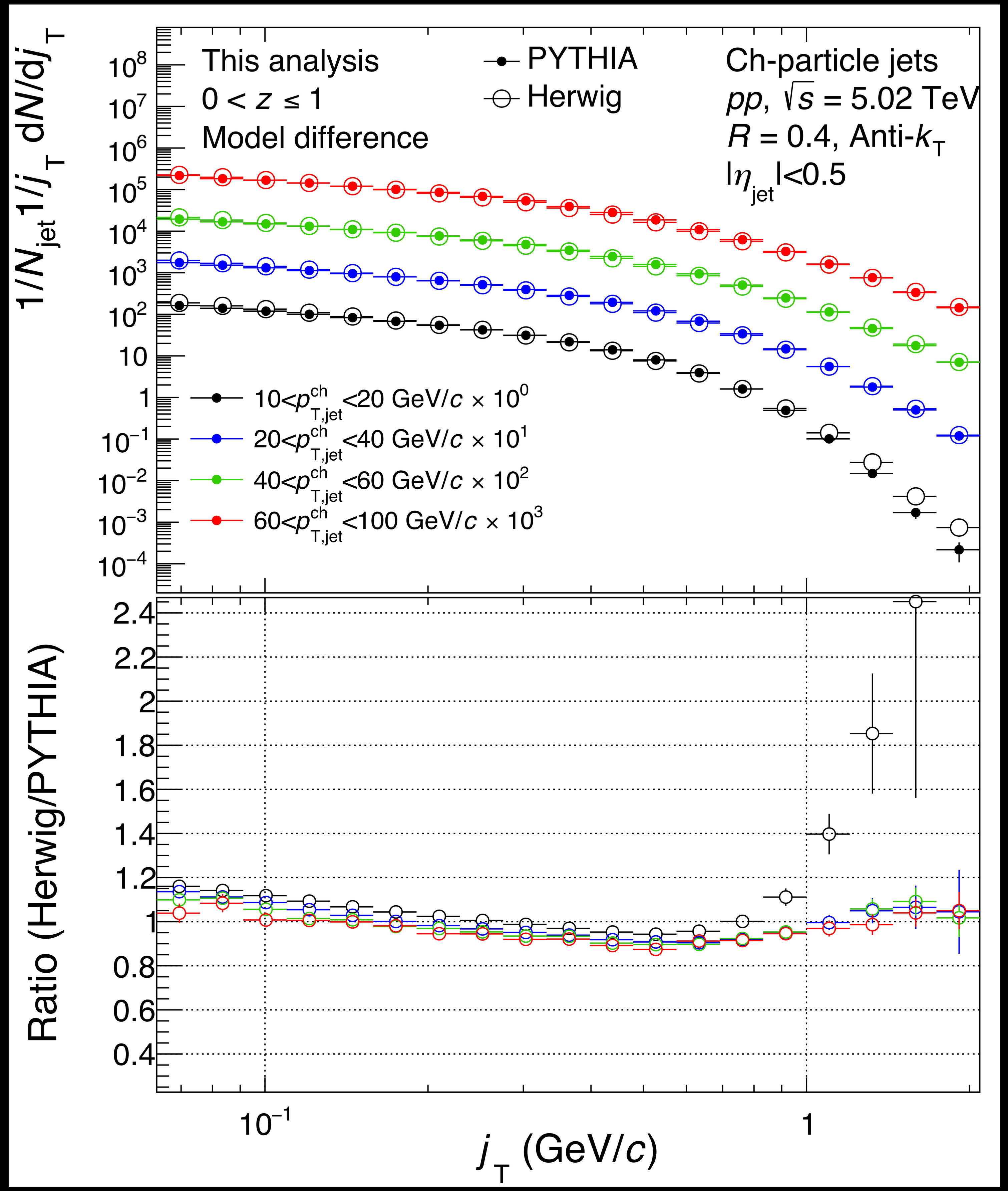
Low z ($0 < z < 0.2$)



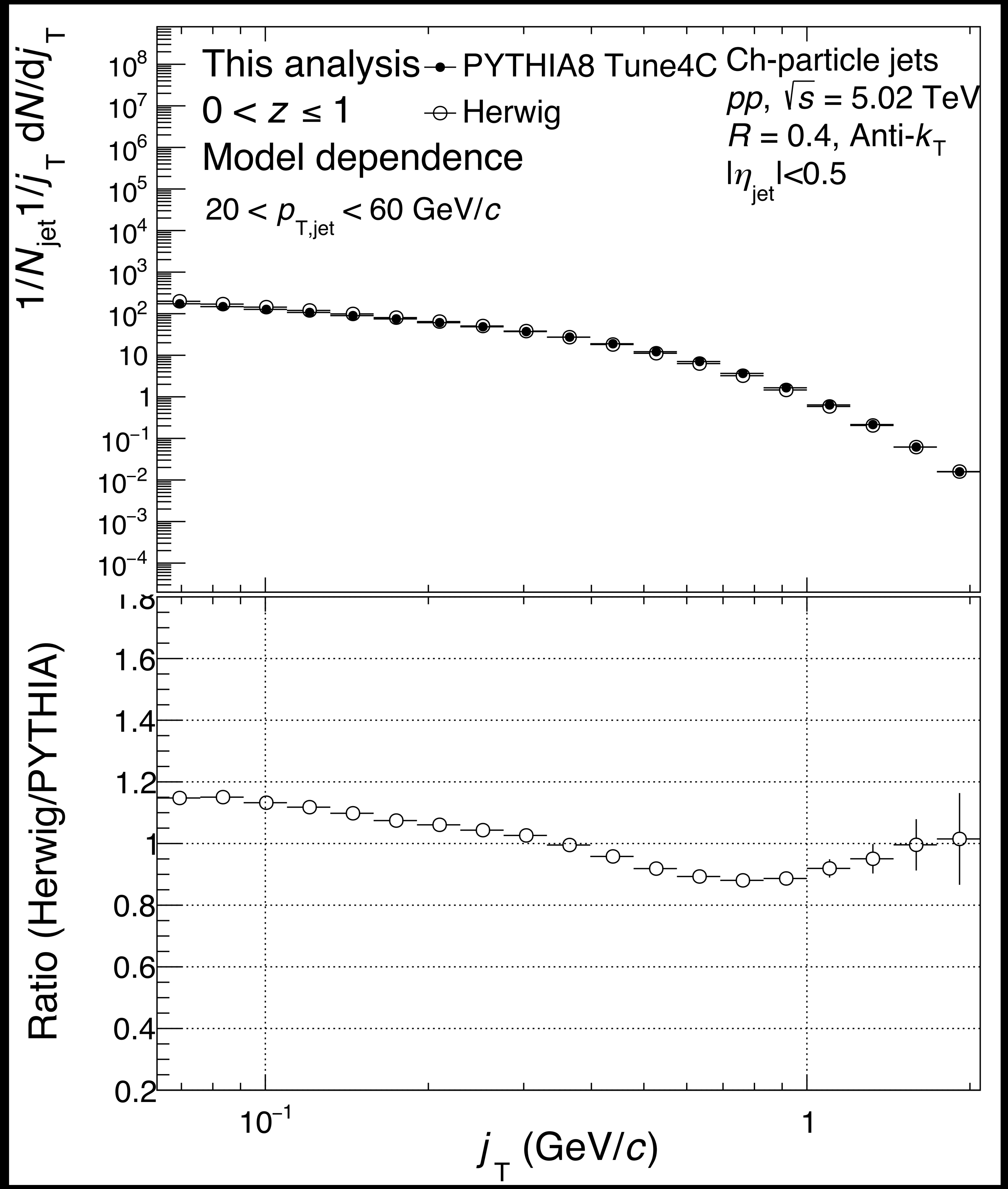
High z ($0.4 < z \leq 1$)

Modified response matrix Herwig

- Modified response matrix based on the ratio of truth j_T
- Since high j_T at low $p_{T,jet}$ increased in Herwig so we got scaling factors from $20 < p_{T,jet} < 60$ GeV/c



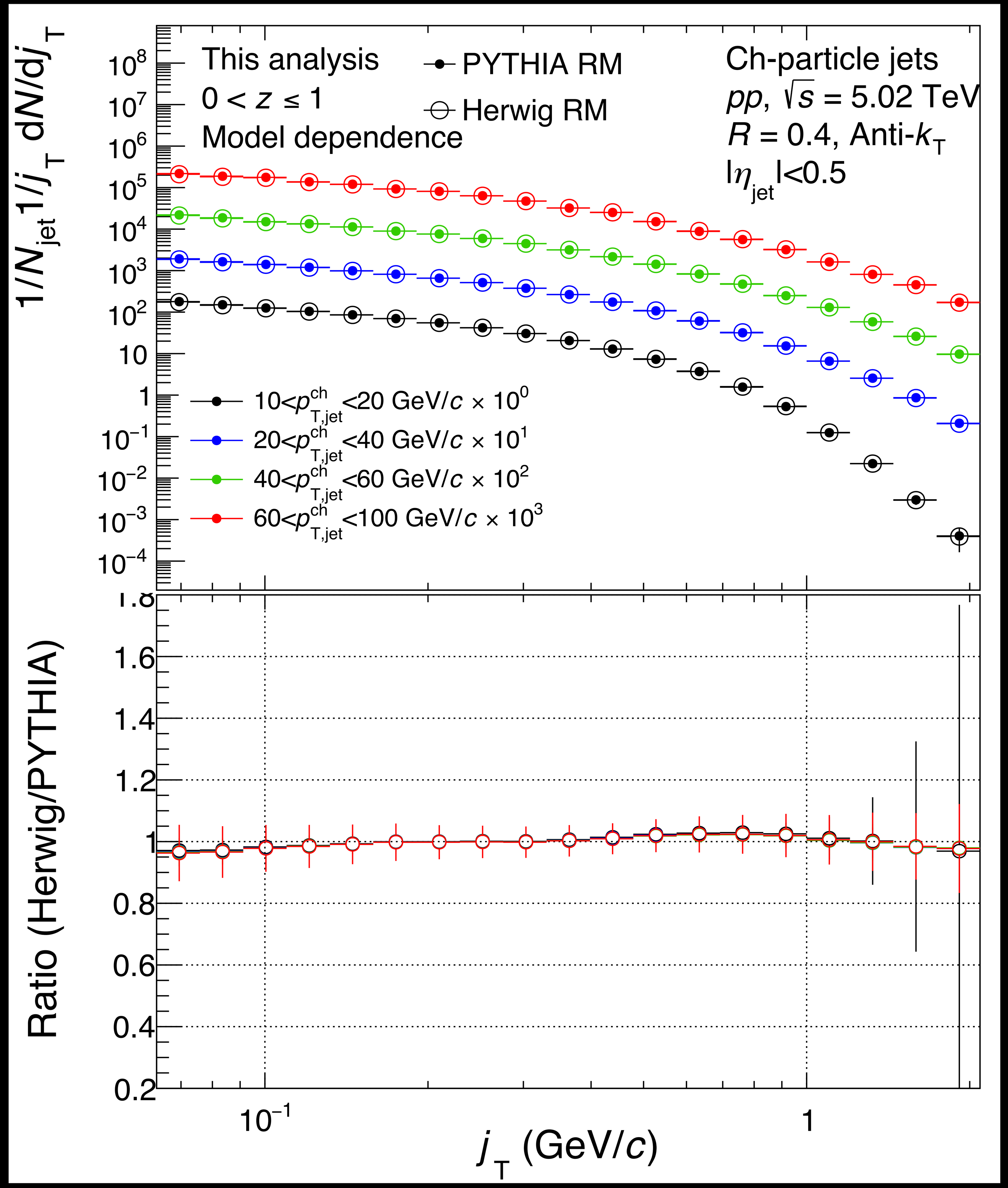
PYTHIA vs. Herwig



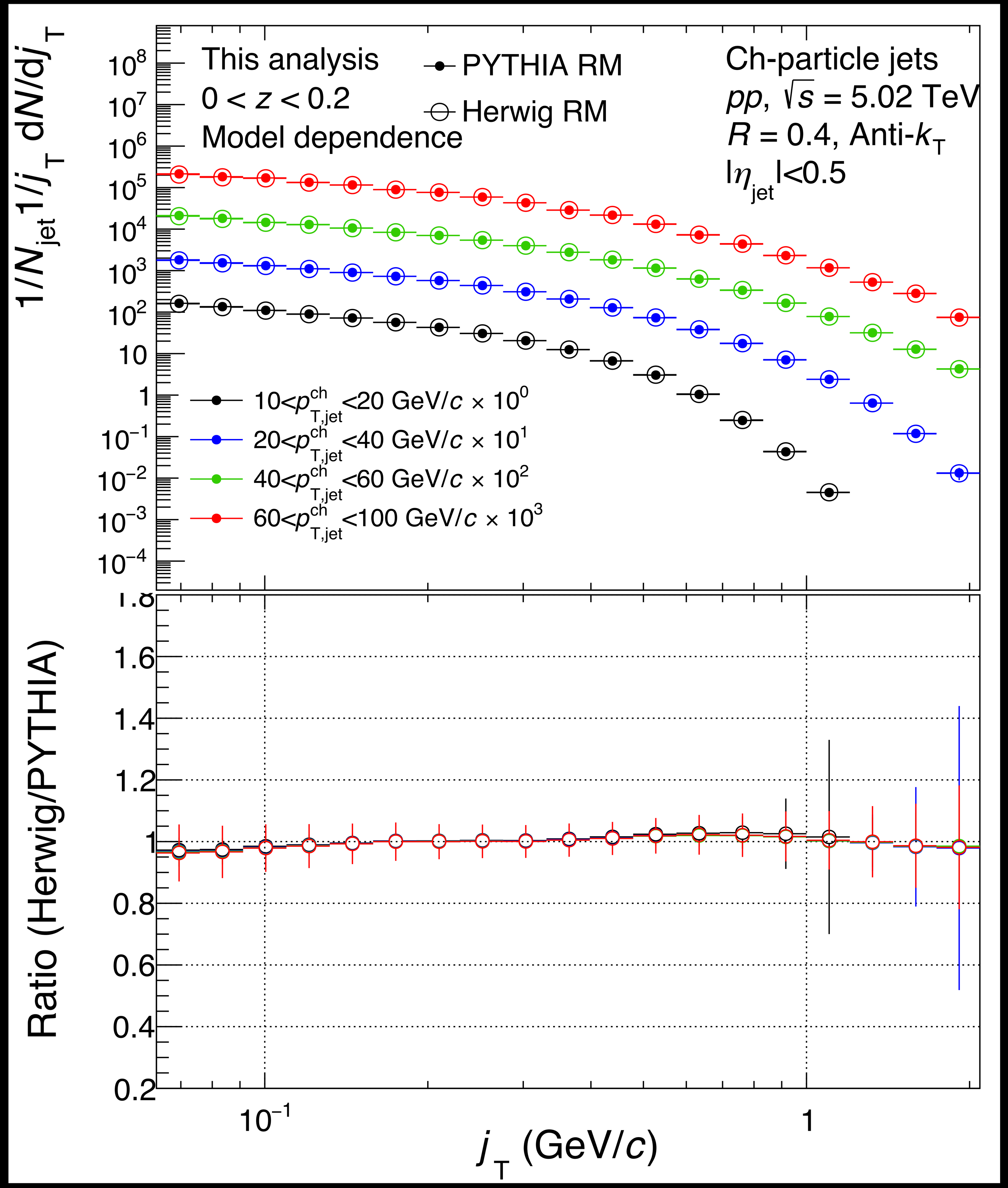
Scaling factor

Modified response matrix Herwig

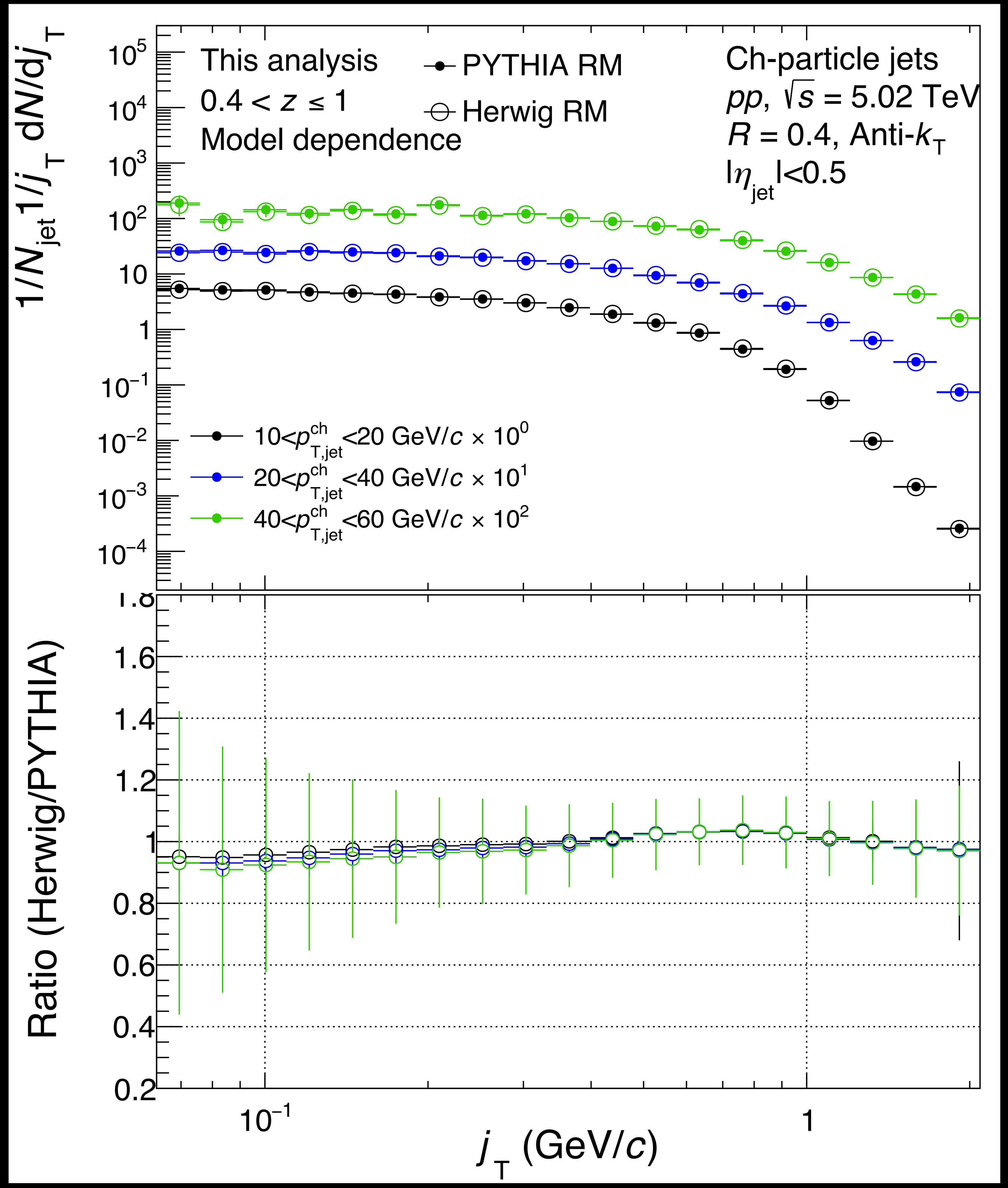
- Modified response matrix based on the ratio of truth j_T



Incl z ($0 < z \leq 1$)



Low z ($0 < z < 0.2$)

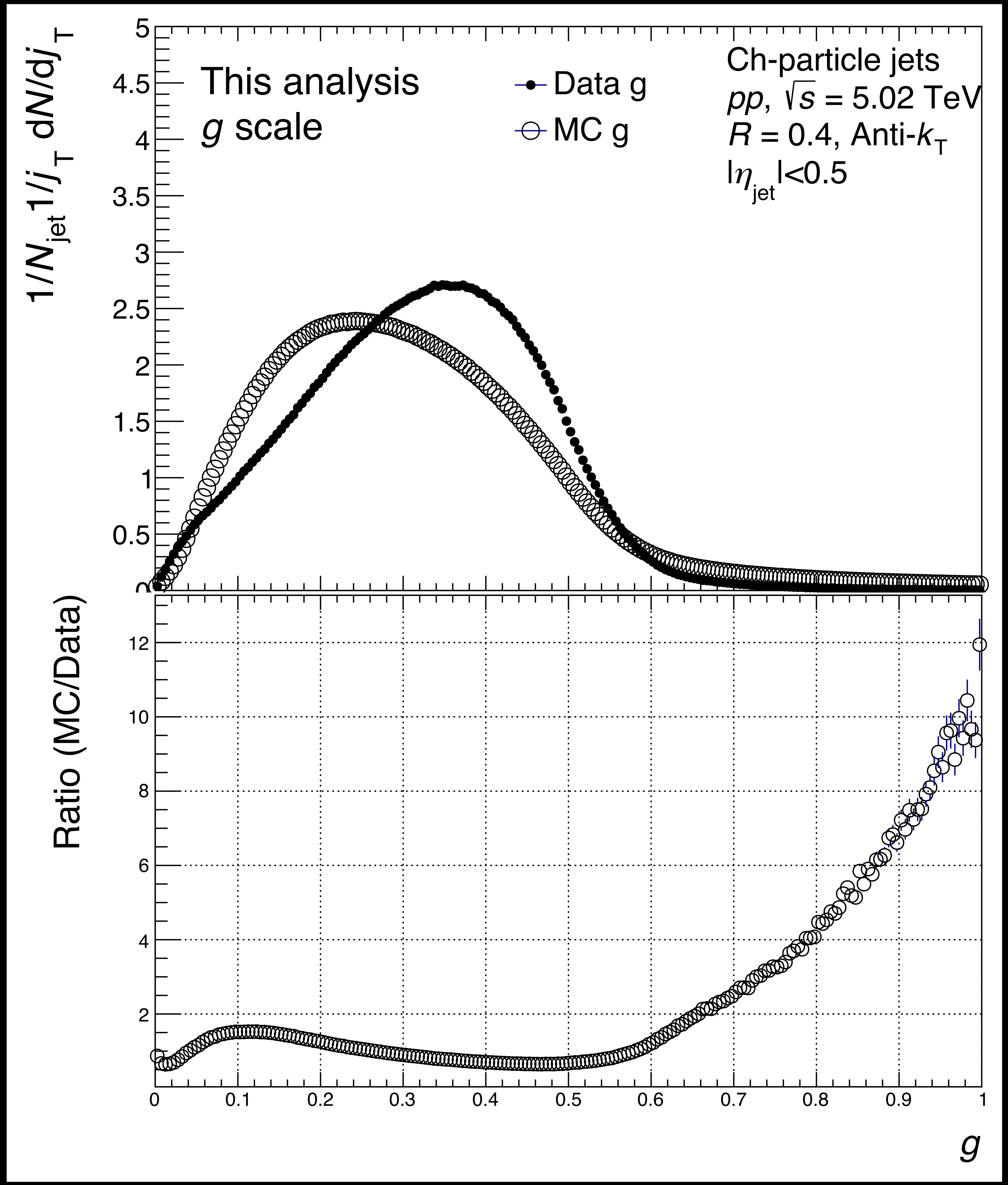


High z ($0.4 < z \leq 1$)

Modified response matrix with jet angularity (g)

- Scale the number of jets and constituents with the difference between data and MC jet angularity g
- $$g = \sum_{i \in jet} z_i \theta_i, \quad z_i = \frac{p_{\mathrm{T,i}}}{p_{\mathrm{T,jet}}} \quad \theta_i = \frac{\Delta R_{i,jet}}{R}$$

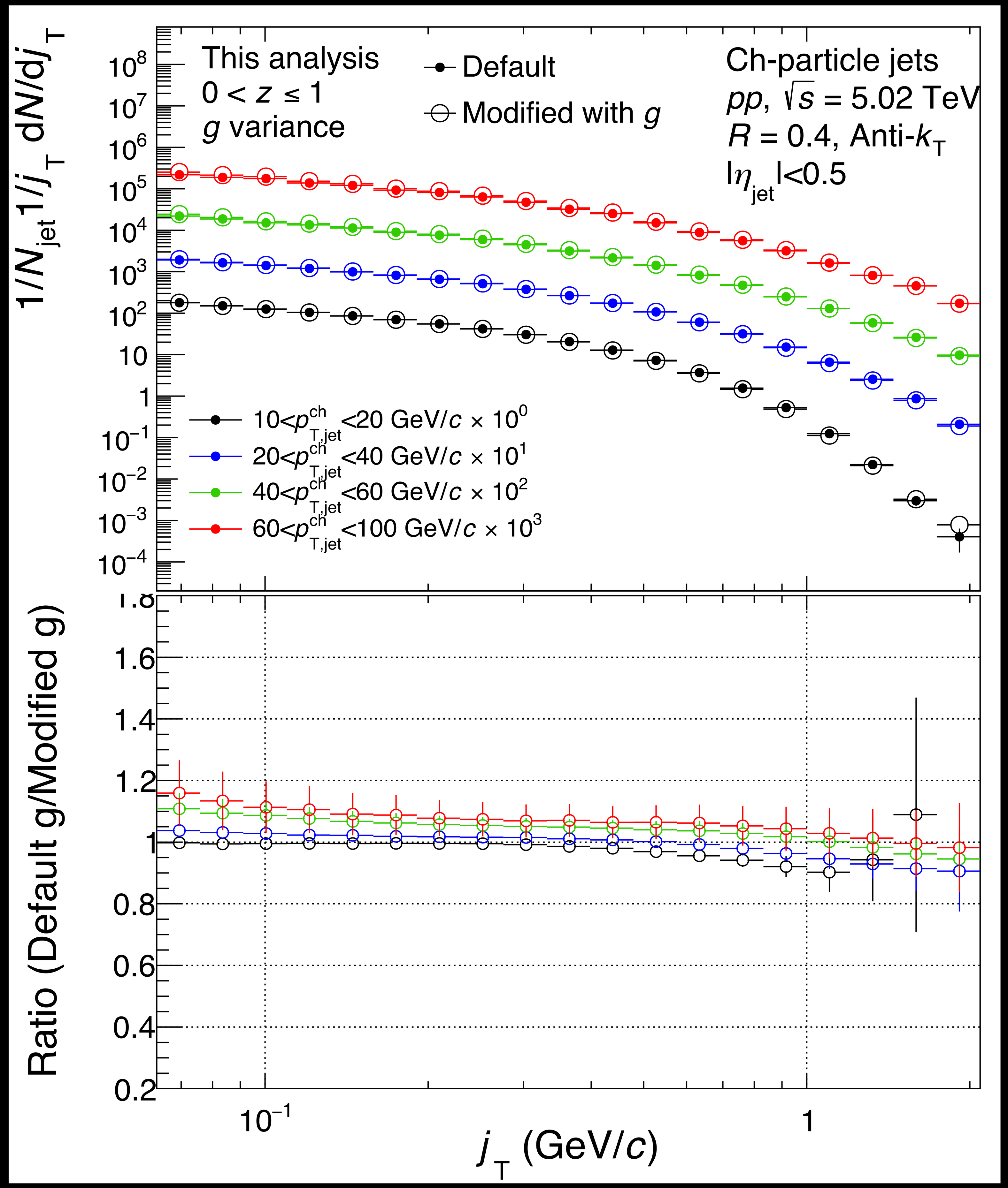
$$\Delta R_{i,jet} = \sqrt{(\eta_{jet} - \eta_i)^2 + (\phi_{jet} - \phi_i)^2}$$



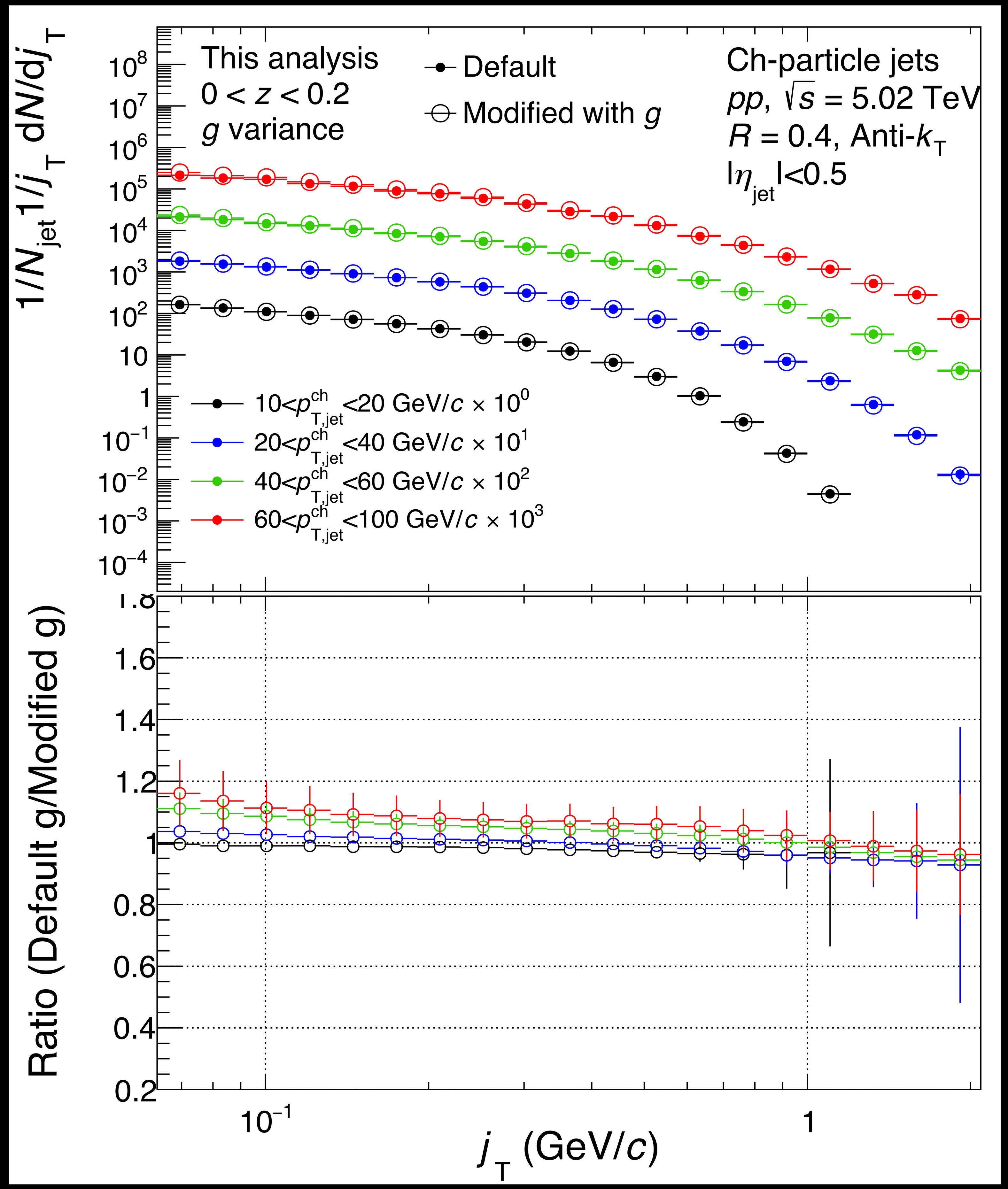
Getting scaling factor

Modified response matrix with jet angularity (g)

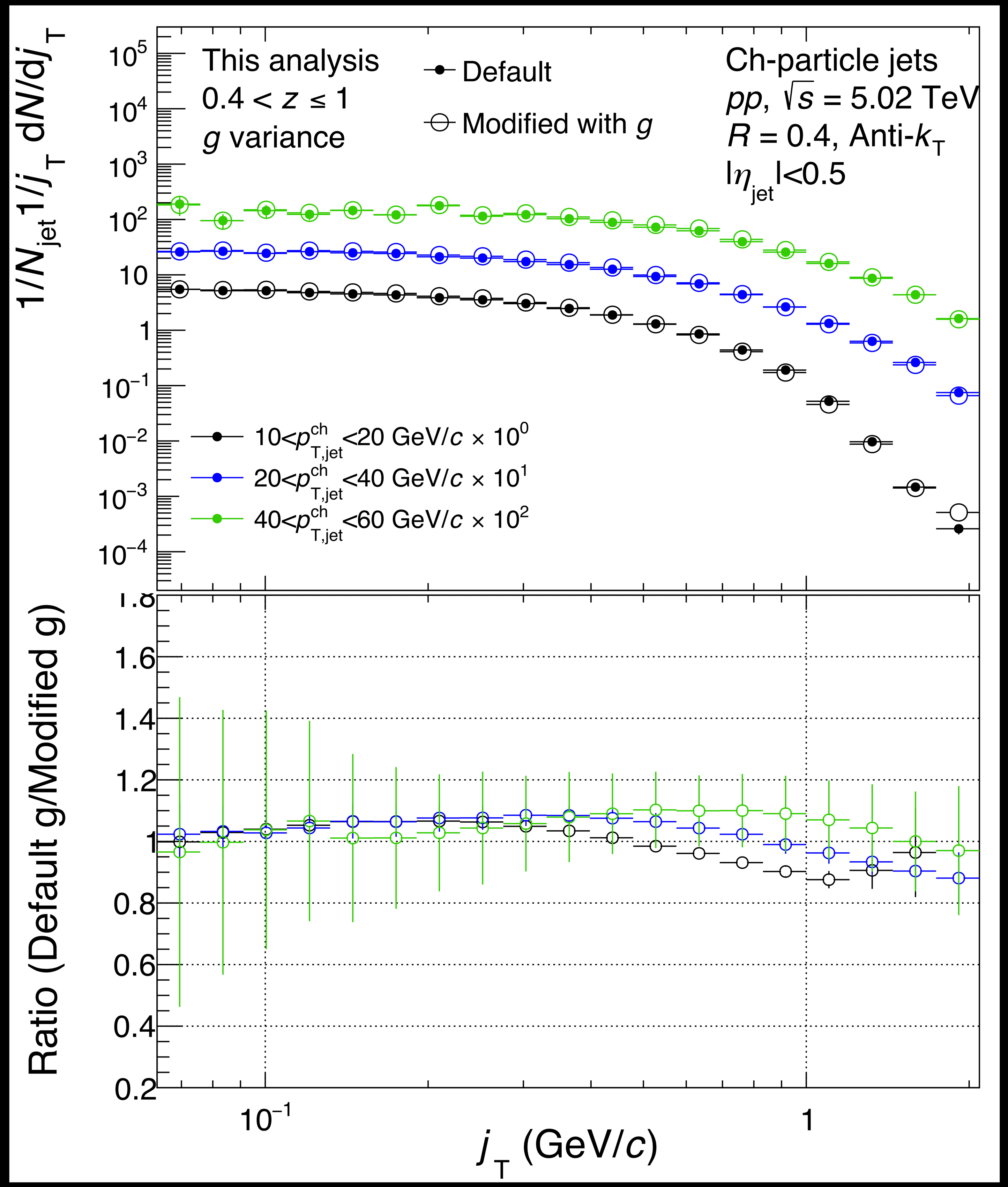
- Modified response matrix based on the ratio of truth j_{T} modified with g



Inclusive z (0 < z ≤1)



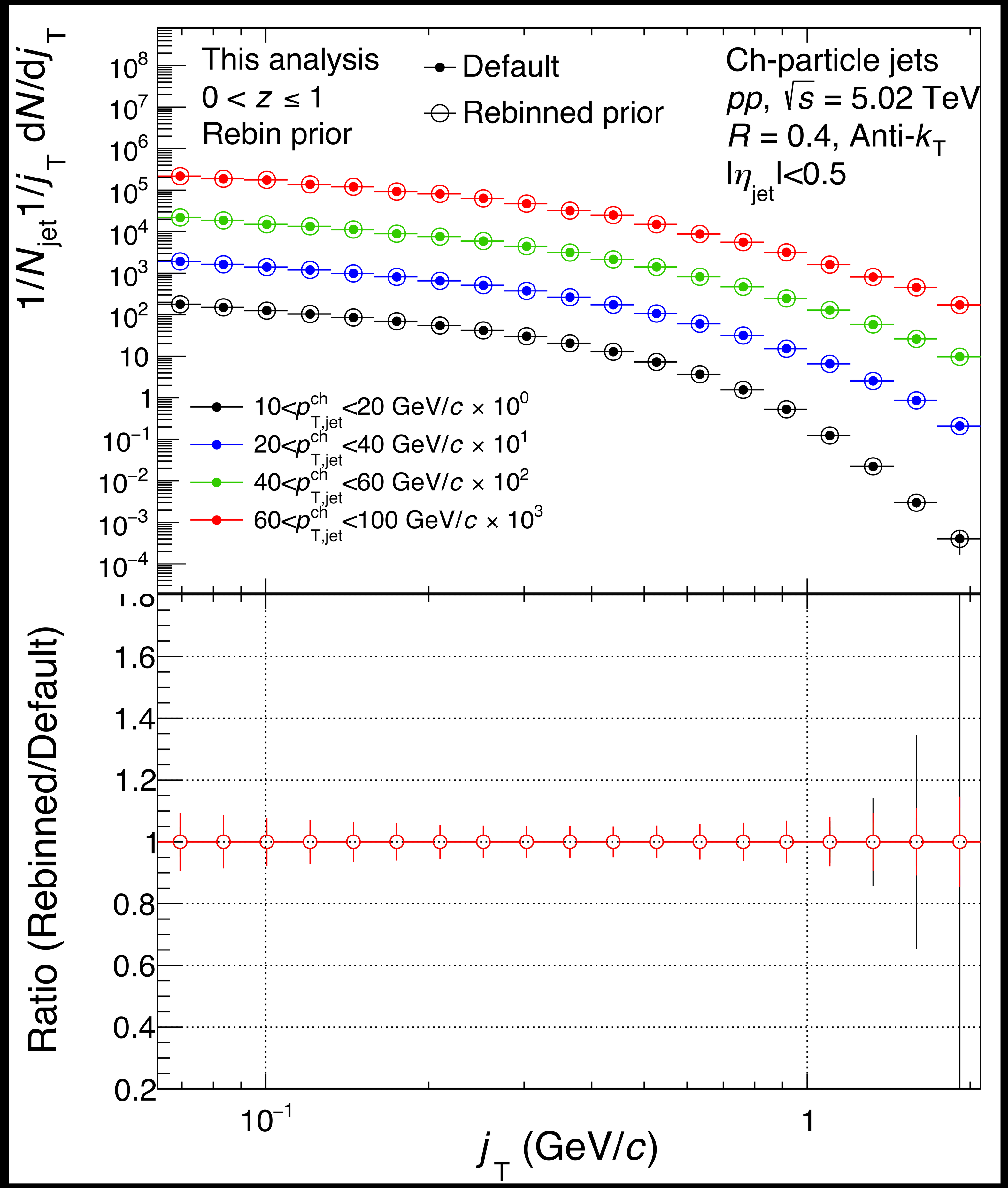
Low z (0 < z < 0.2)



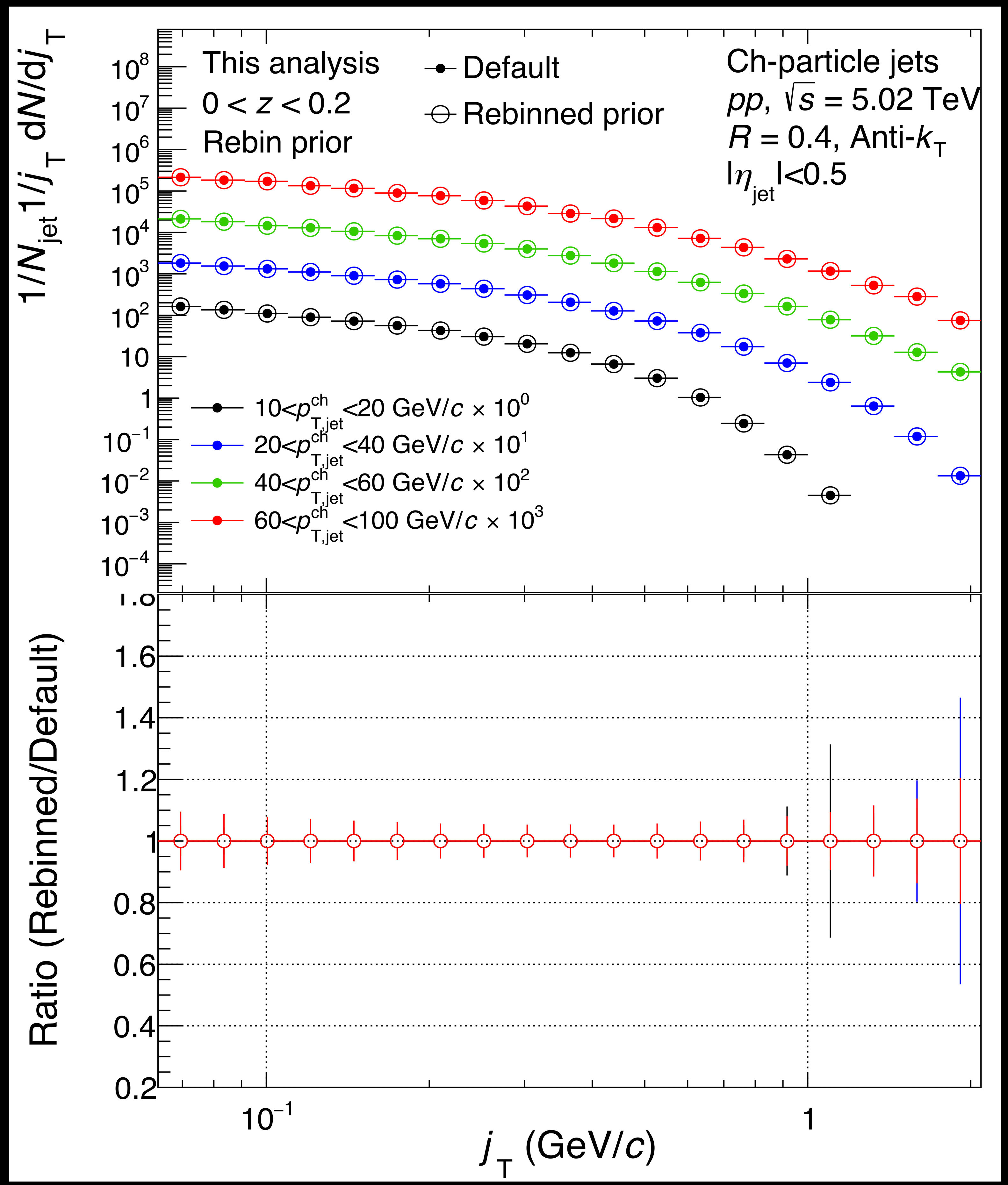
High z (0.4 < z ≤ 1)

Rebinning

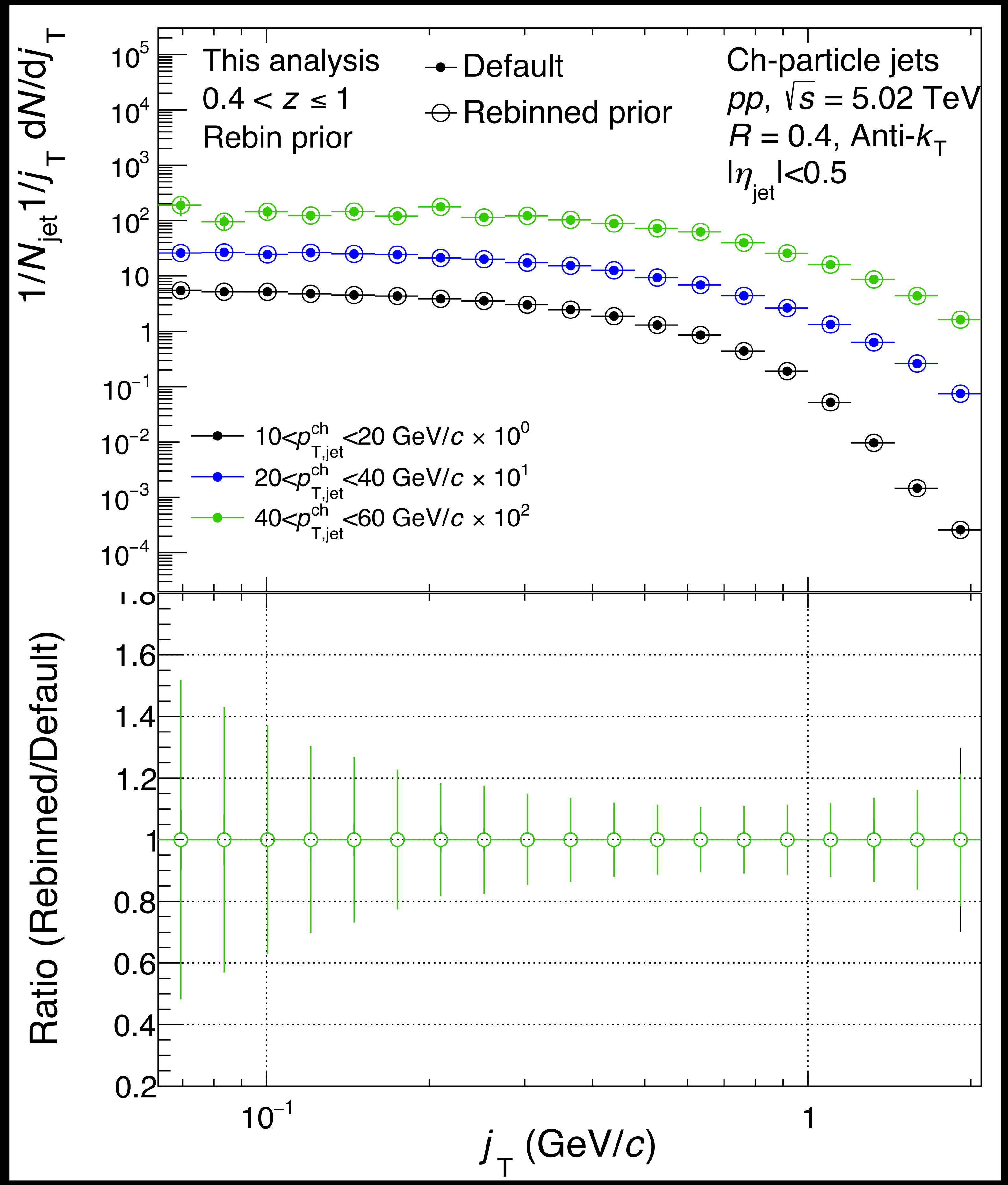
- Rebinned(twice) the prior distributions and investigated the systematic difference
- There were no systematic differences so we decided not to use these results in the systematic error



Inclusive z (0 < z ≤1)



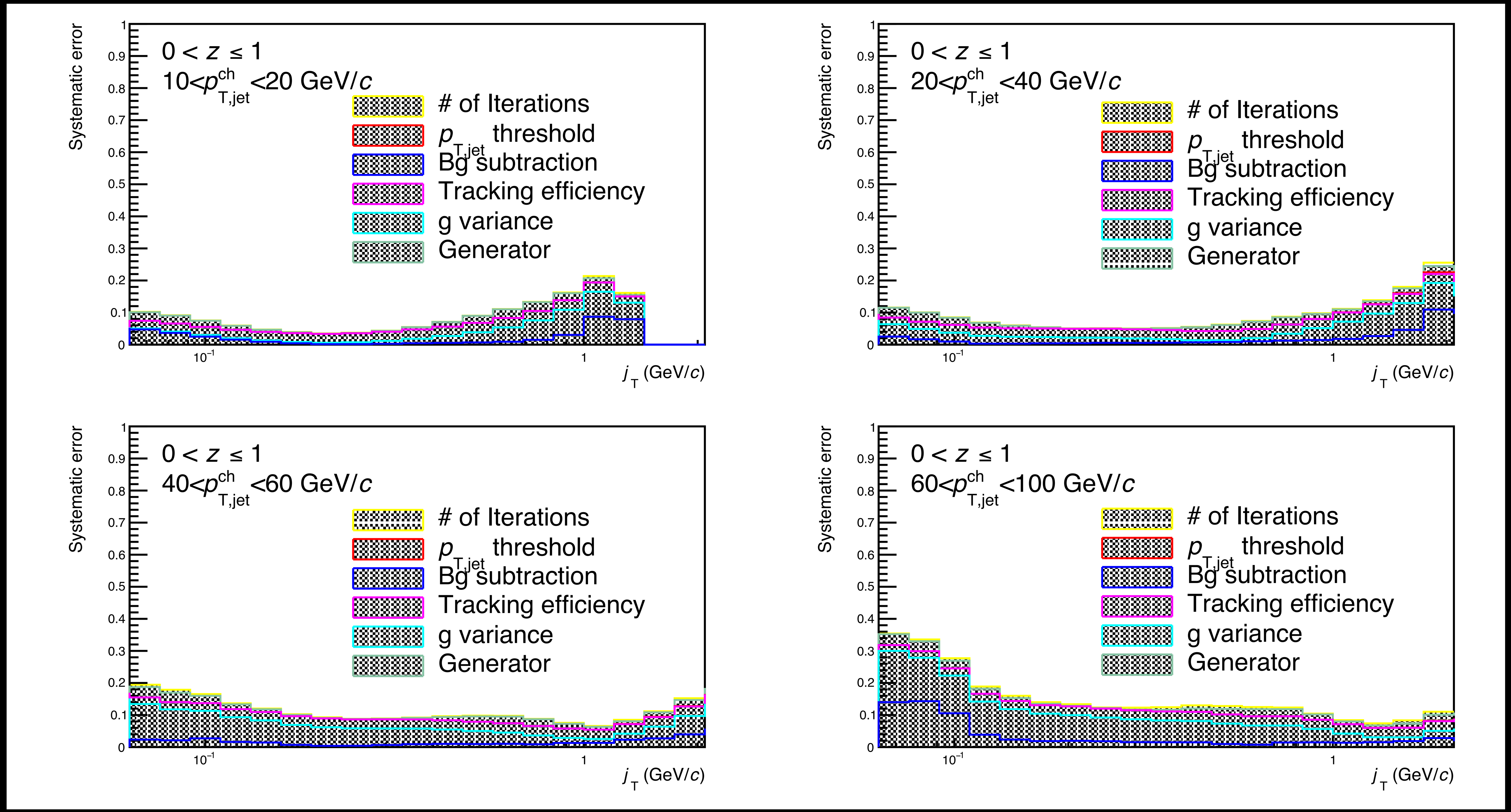
Low z (0 < z < 0.2)



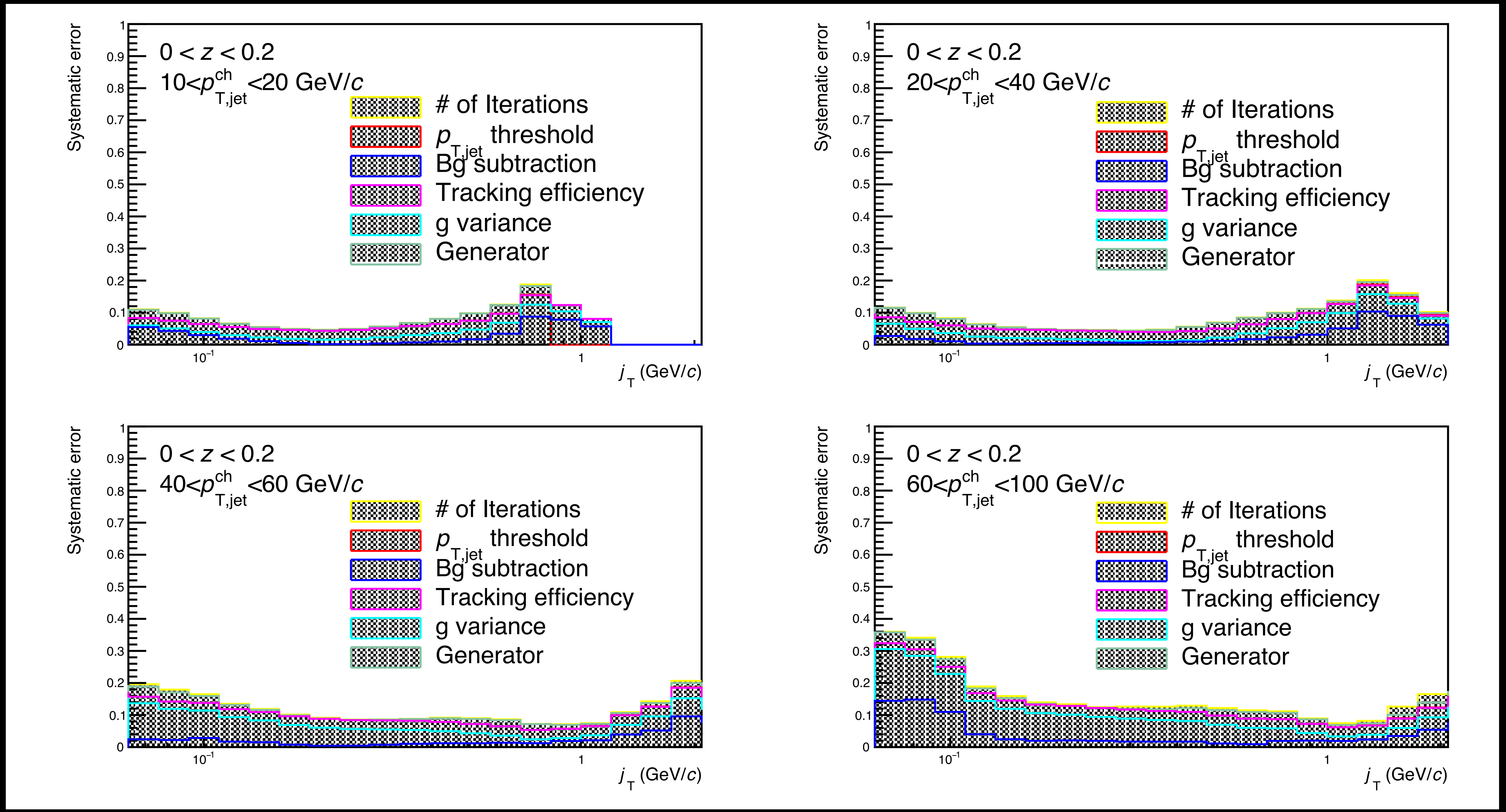
High z (0.4 < z ≤ 1)

Systematic errors from each source

- Each systematic error went through a smoothening process.
  (Add the values of the bin and both side bins and divide it by 3)
- Fraction of the systematic w.r.t. central value



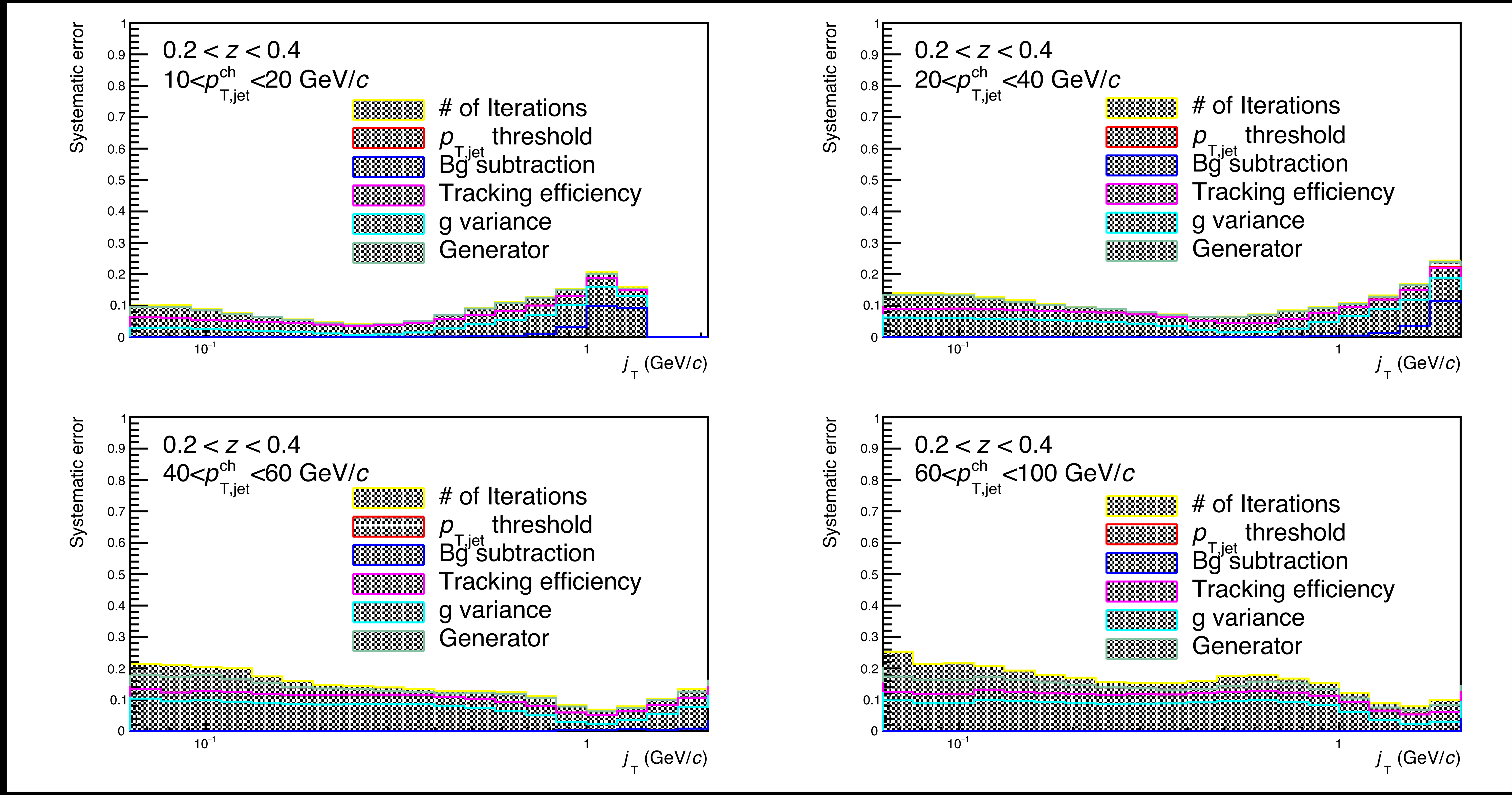
Inclusive z (0 < z ≤1)



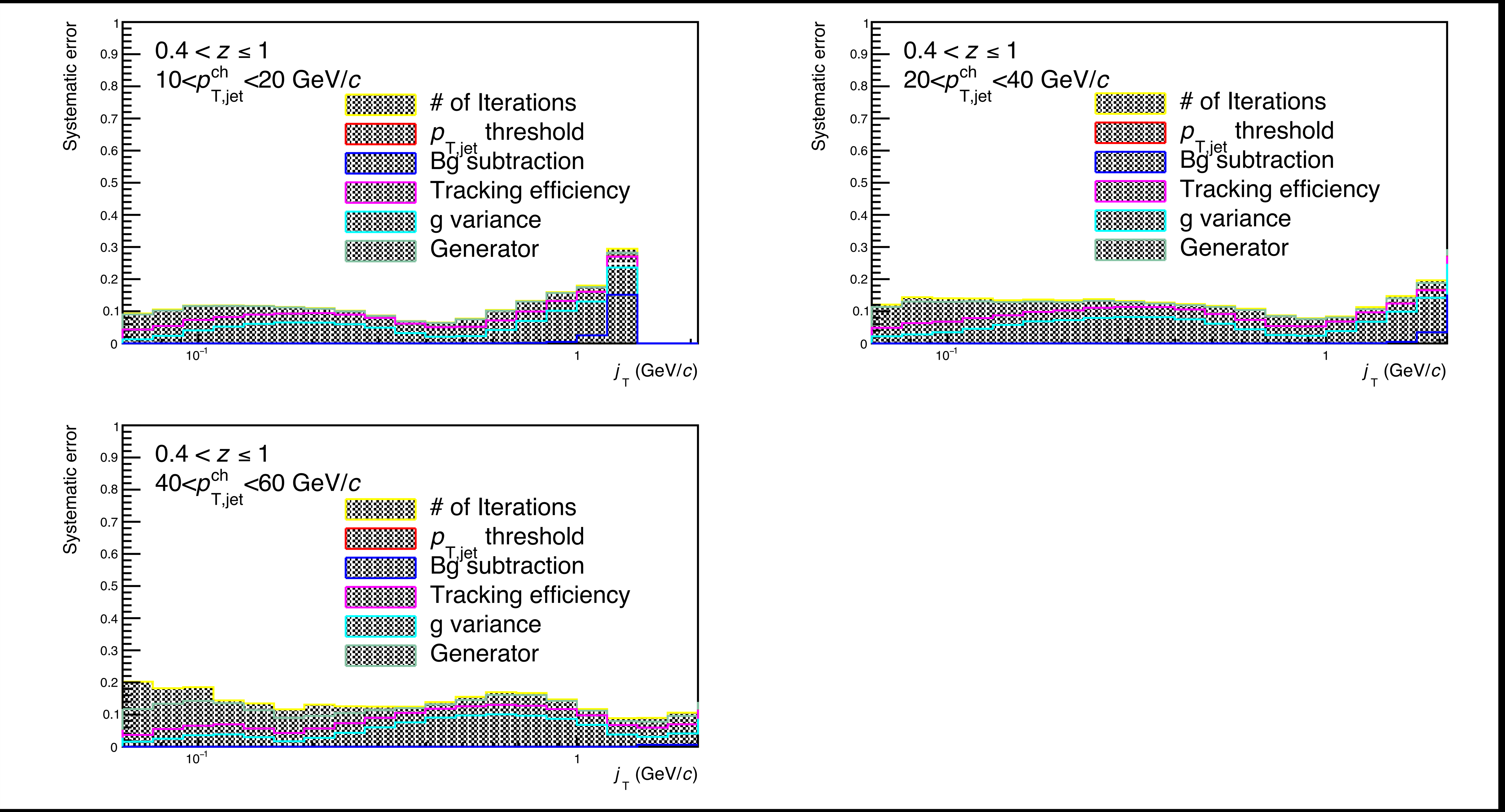
Low z (0 < z < 0.2)

Systematic errors from each source

- Each systematic error went through a smoothening process.
(Add the values of the bin and both side bins and divide it by 3)
- Fraction of the systematic w.r.t. central value



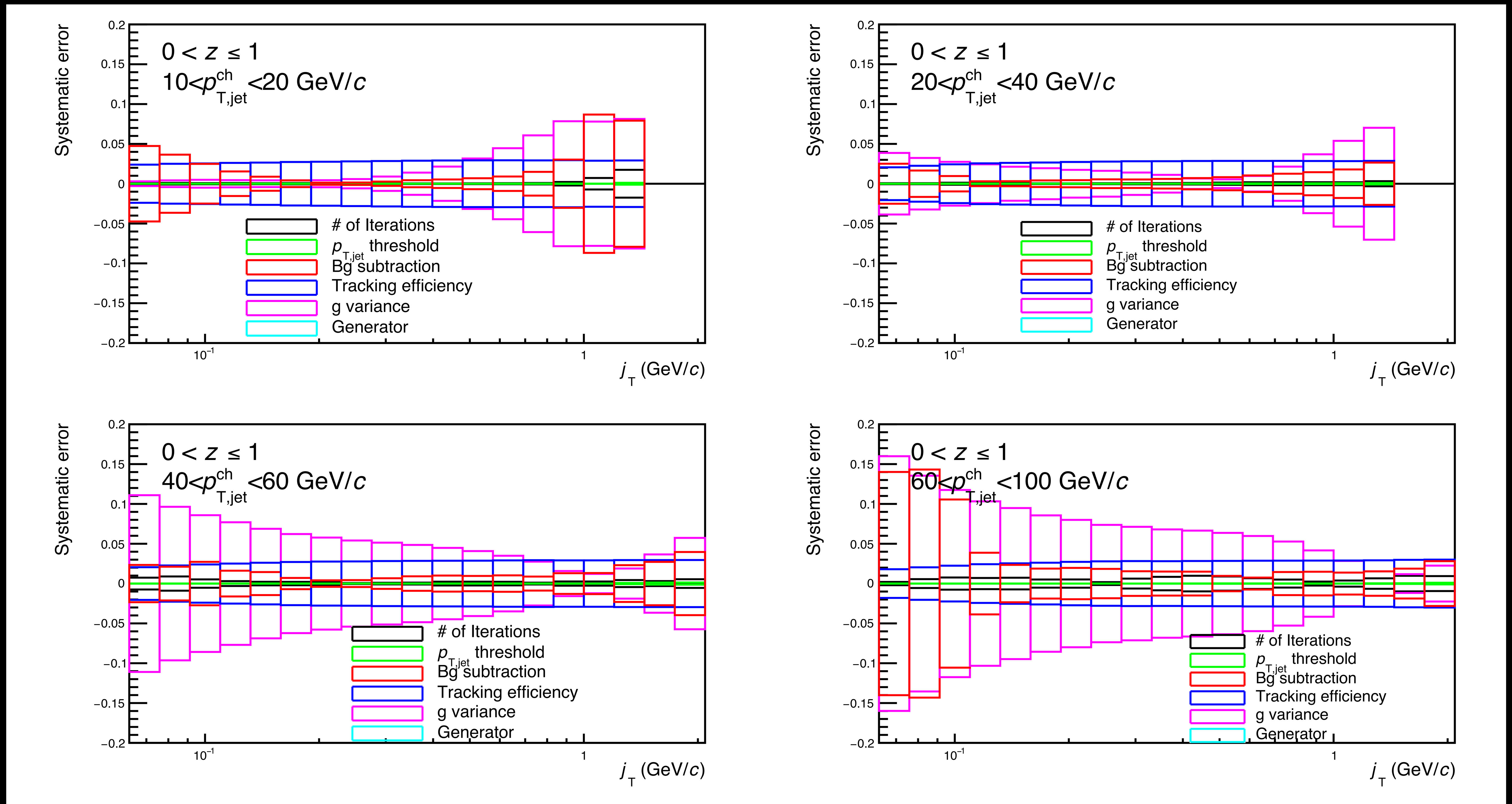
Mid z (0.2 < z < 0.4)



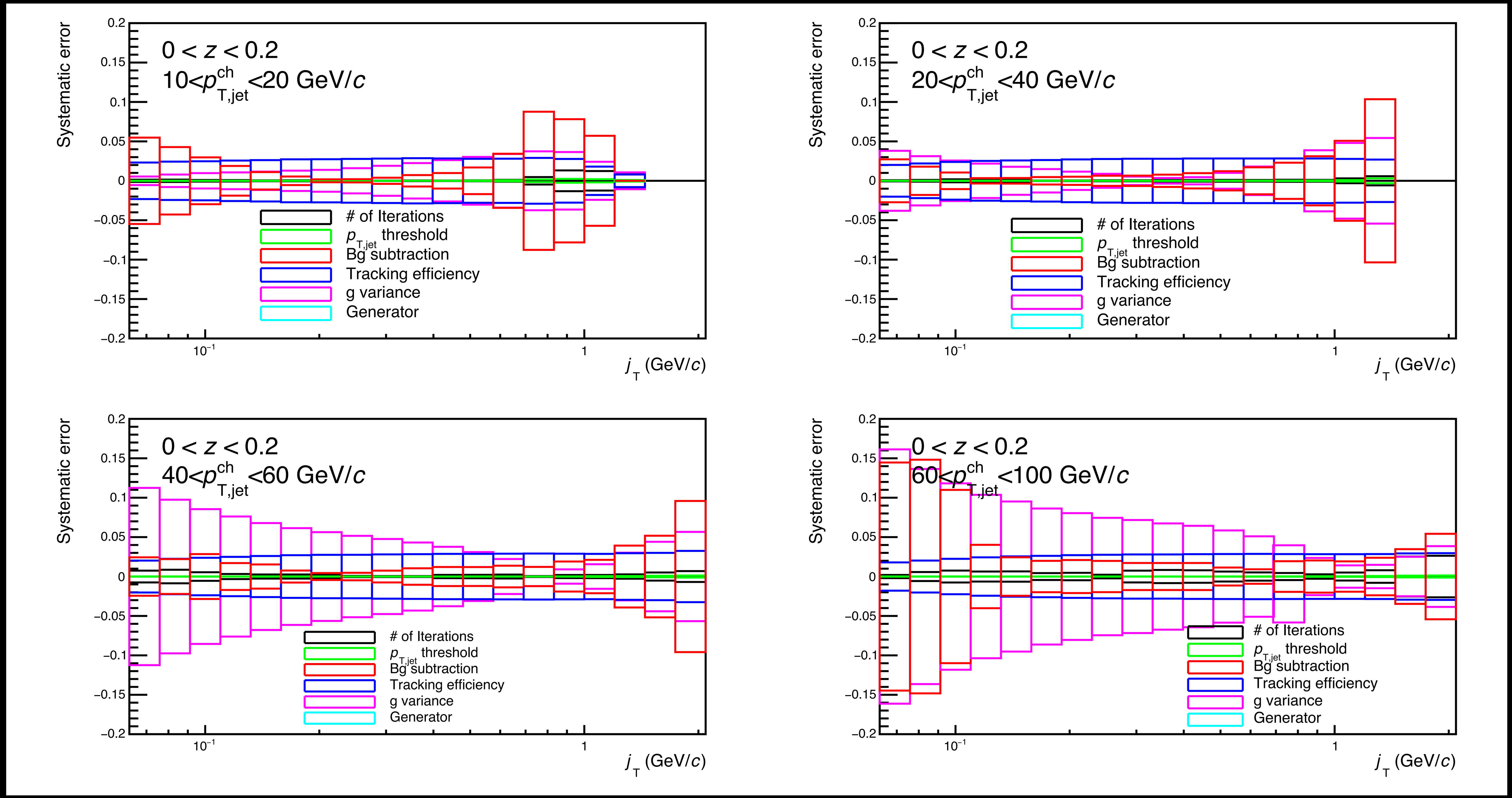
High z ($0.4 < z \leq 1$)

Systematic errors from each source

- Each systematic error went through a smoothening process.
- Fraction of the systematic w.r.t. central value



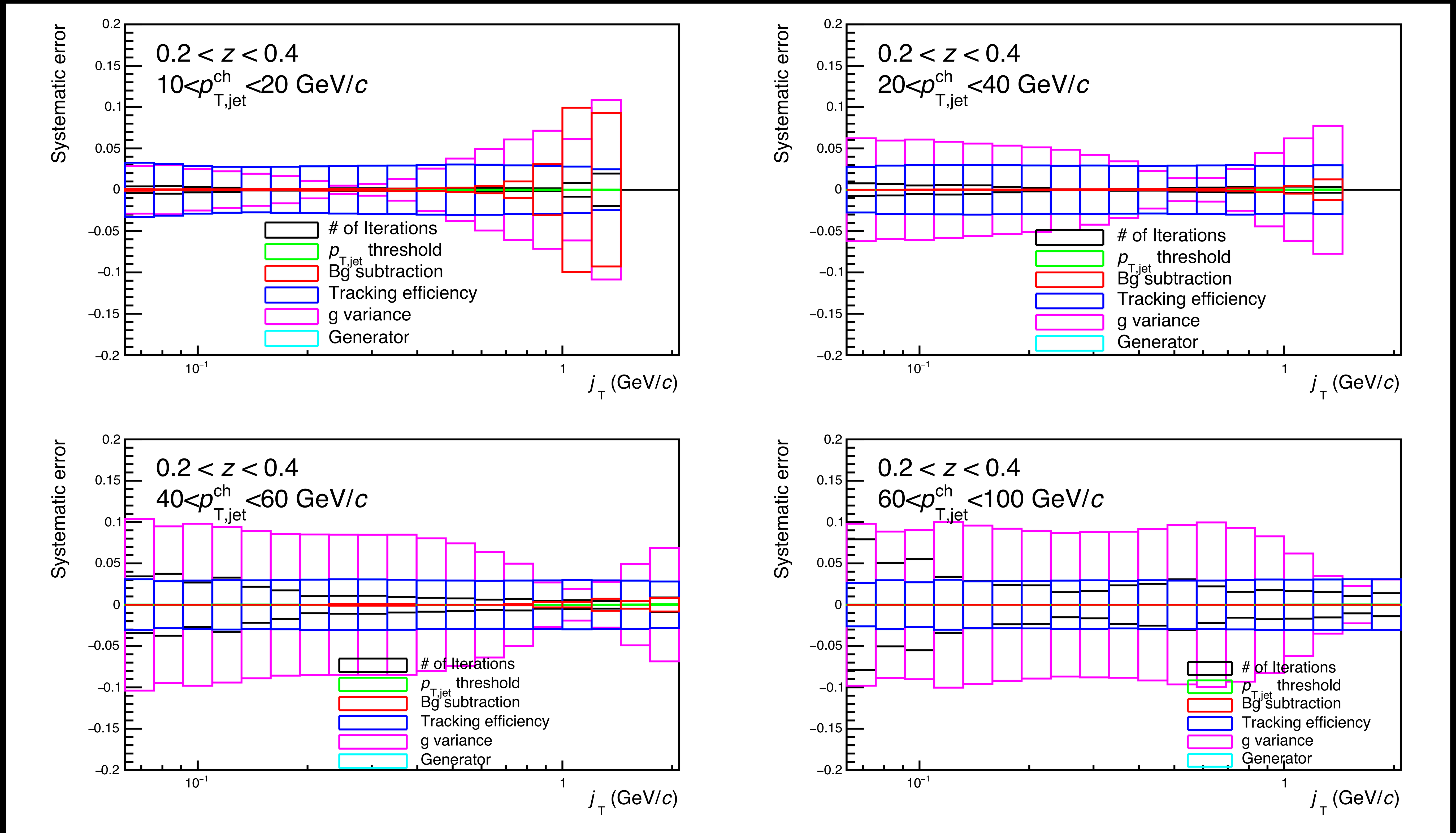
Inclusive z ($0 < z \leq 1$)



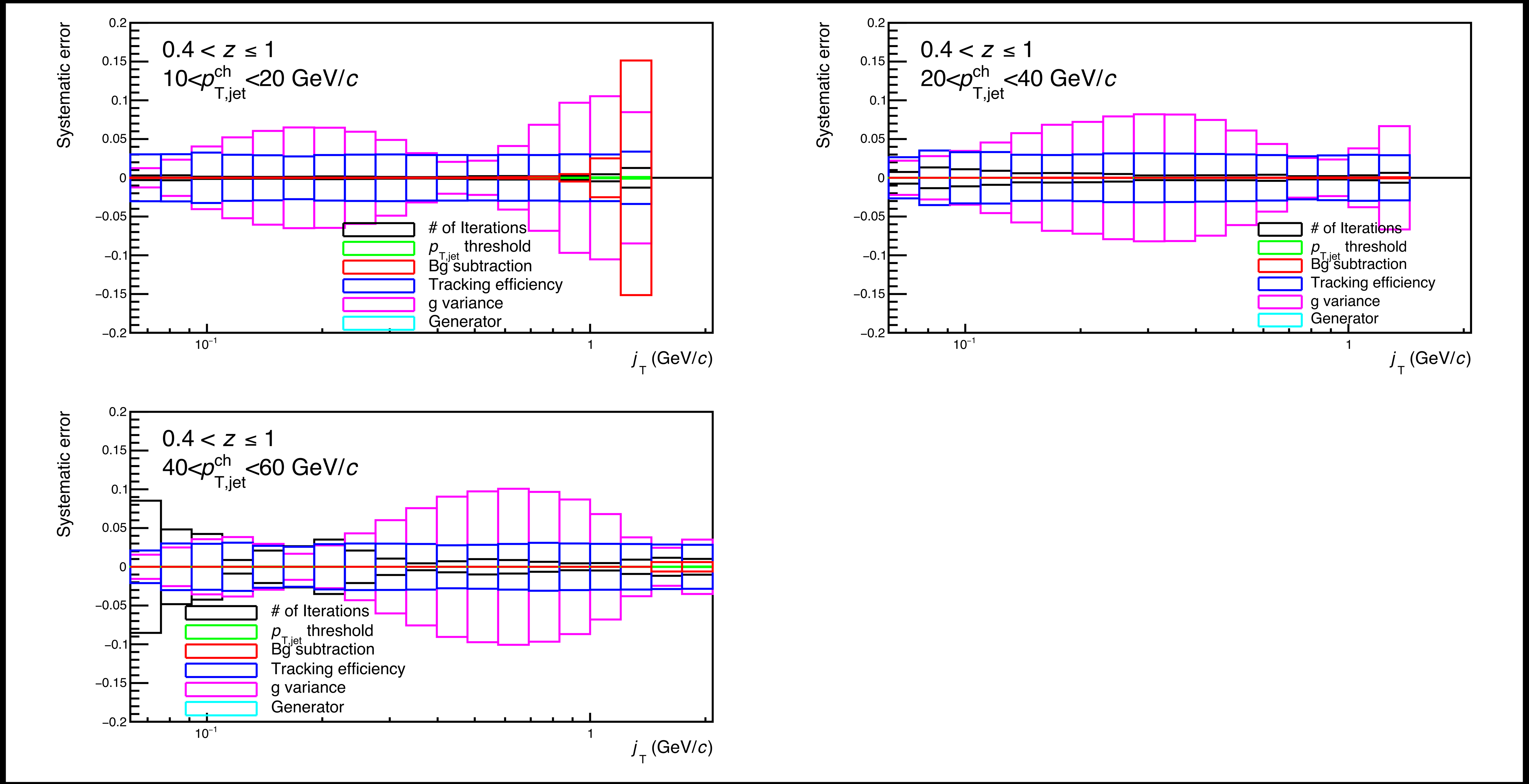
Low z ($0 < z < 0.2$)

Systematic errors from each source

- Each systematic error went through a smoothening process.
- Fraction of the systematic w.r.t. central value



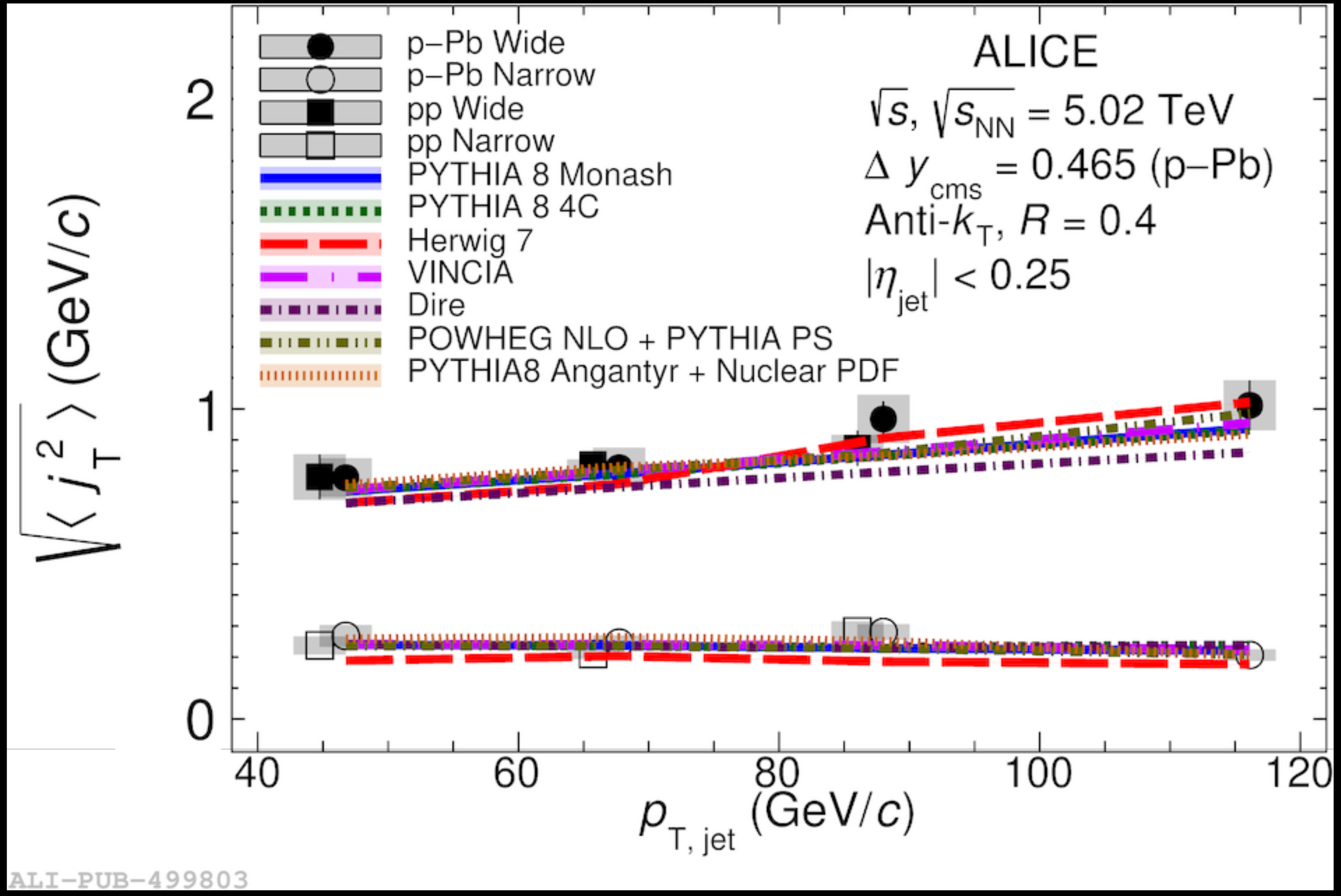
Mid z (0.2 < z < 0.4)

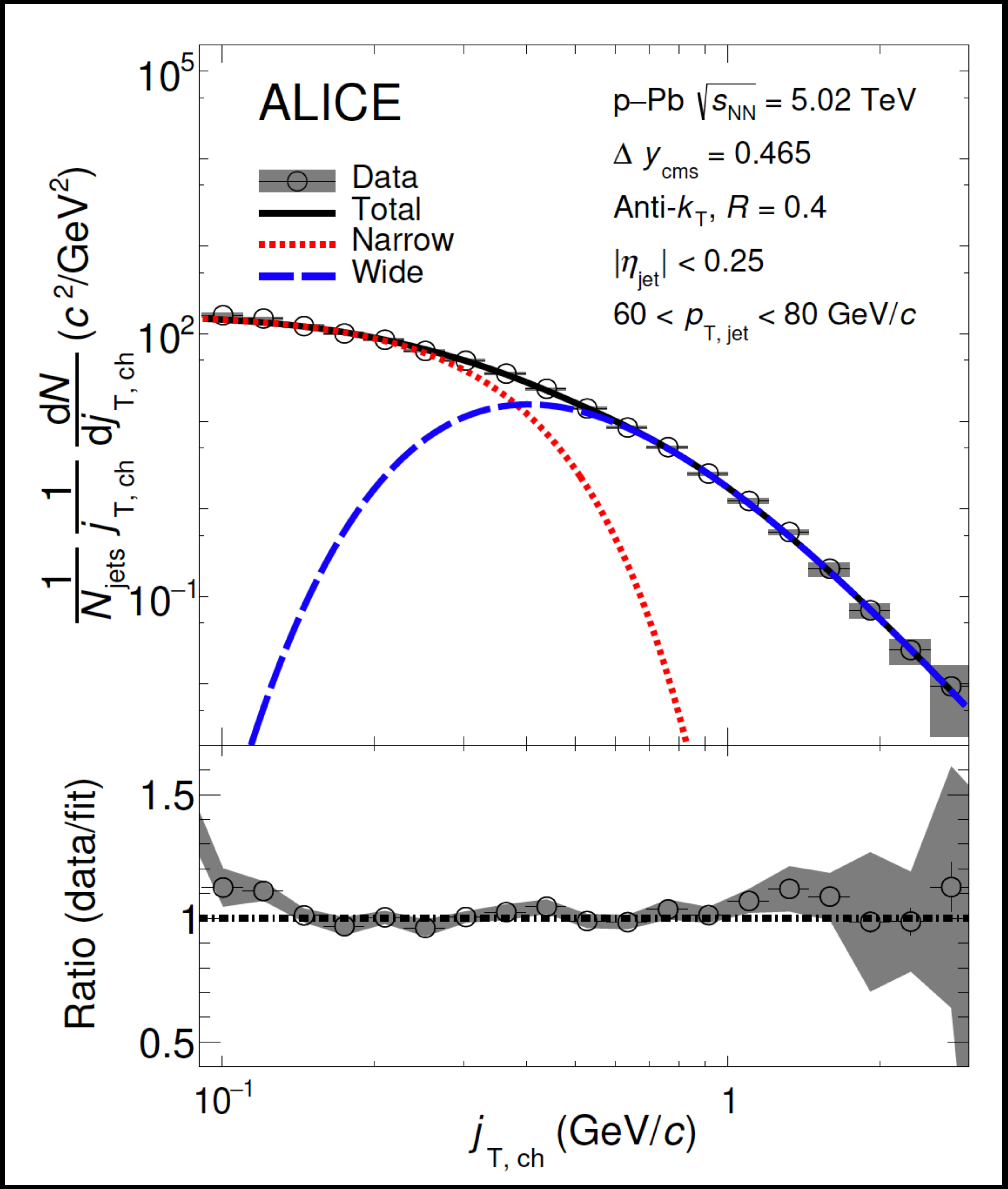


High z (0.4 < z ≤ 1)

Two-component fit

JHEP09(2021)211





Gaussian distribution centered at j_T = 0 GeV/c for the lower j_T

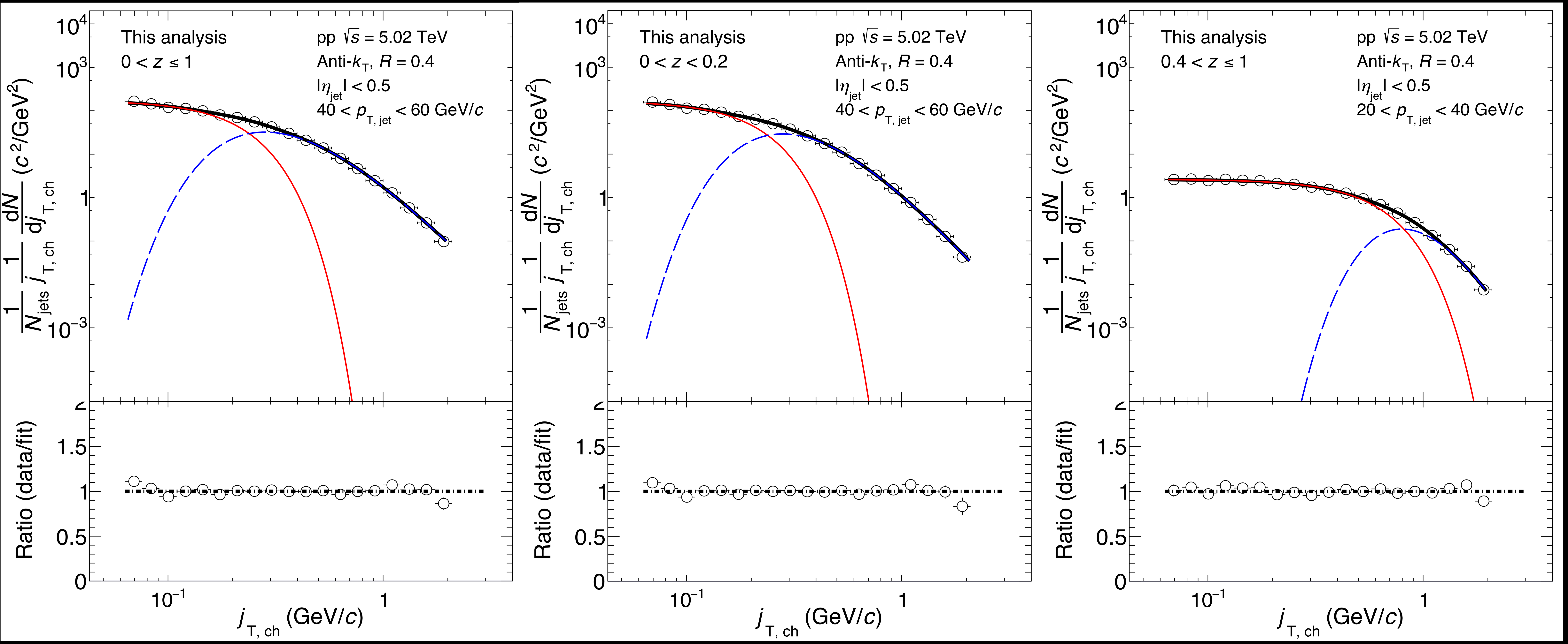
$$\frac{1}{N_{jets}} \frac{dN}{j_{T,ch} dj_{T,ch}} = \frac{B_2}{B_1\sqrt{2\pi}} e^{-\frac{j_T^2}{2B_1^2}} + \frac{B_3 B_5^{B_4}}{\Gamma(B_4)} \frac{e^{-\frac{B_5}{j_T}}}{j_T^{B_4+1}} \quad (3.2)$$

Inverse gamma function for the higher j_T

B1 : σ, B2 : normalization   B3 : normalization, B4 : shape, B5 : scale

Two-component fit

Data

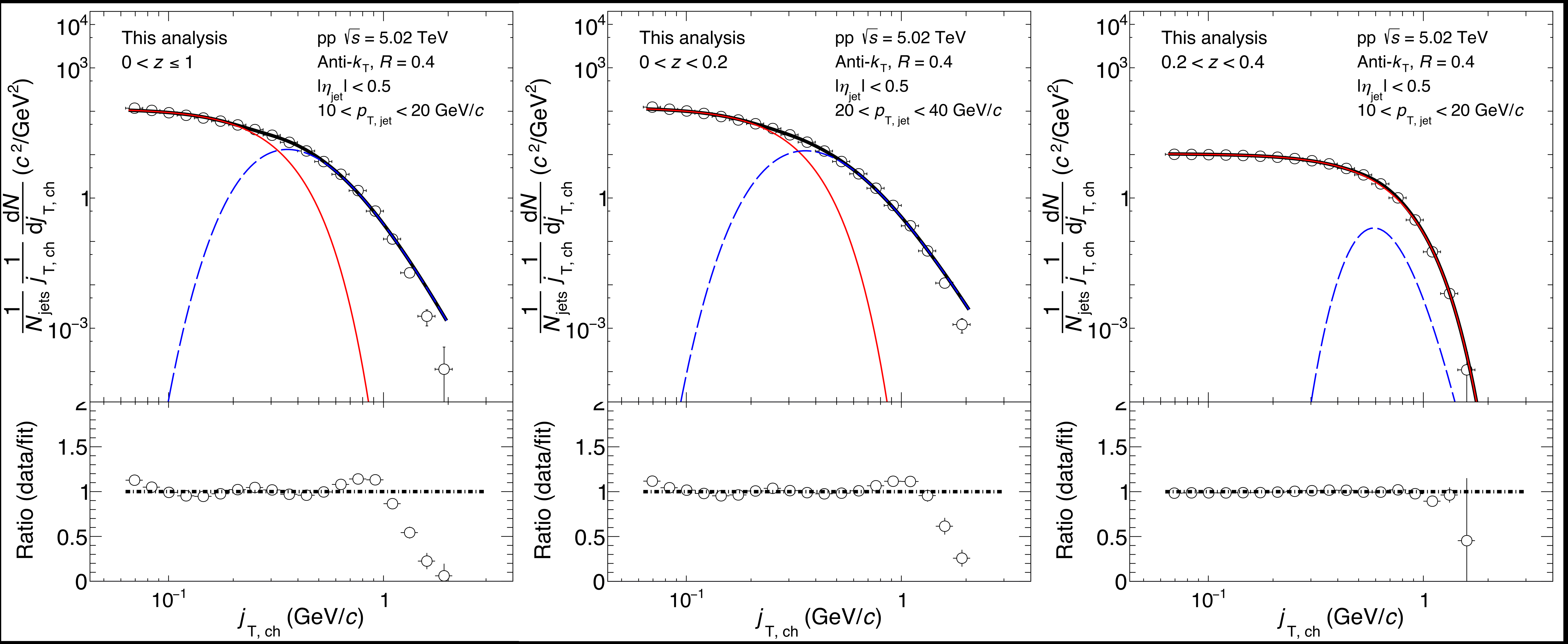


- Two-component fitting was applied to the data, confirming effective performance across various z bins
- However, issues were encountered with fitting in low $p_{\mathrm{T,jet}}$ or low z regions in data and Leading-Order MC
- Considering adjusting the fitting range and j_{T} range for data points depending on the kinematic range

Two-component fit

Data

Monash

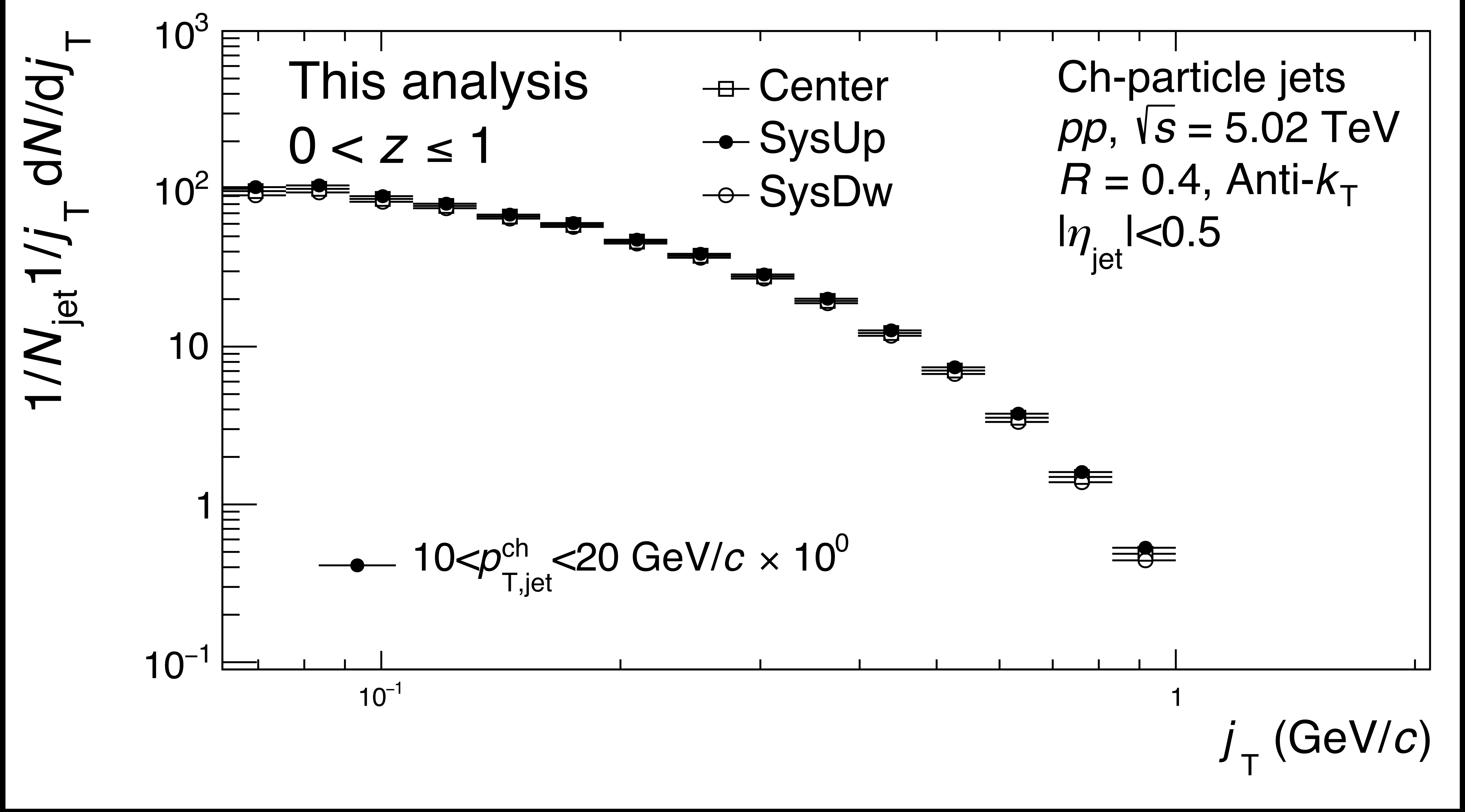


- Two-component fitting was applied to the data, confirming effective performance across various z bins
- However, issues were encountered with fitting in low $p_{\mathrm{T,jet}}$ or low z regions in data and Leading-Order MC
- Considering adjusting the j_{T} range for fitting data points depending on the kinematic range

Results

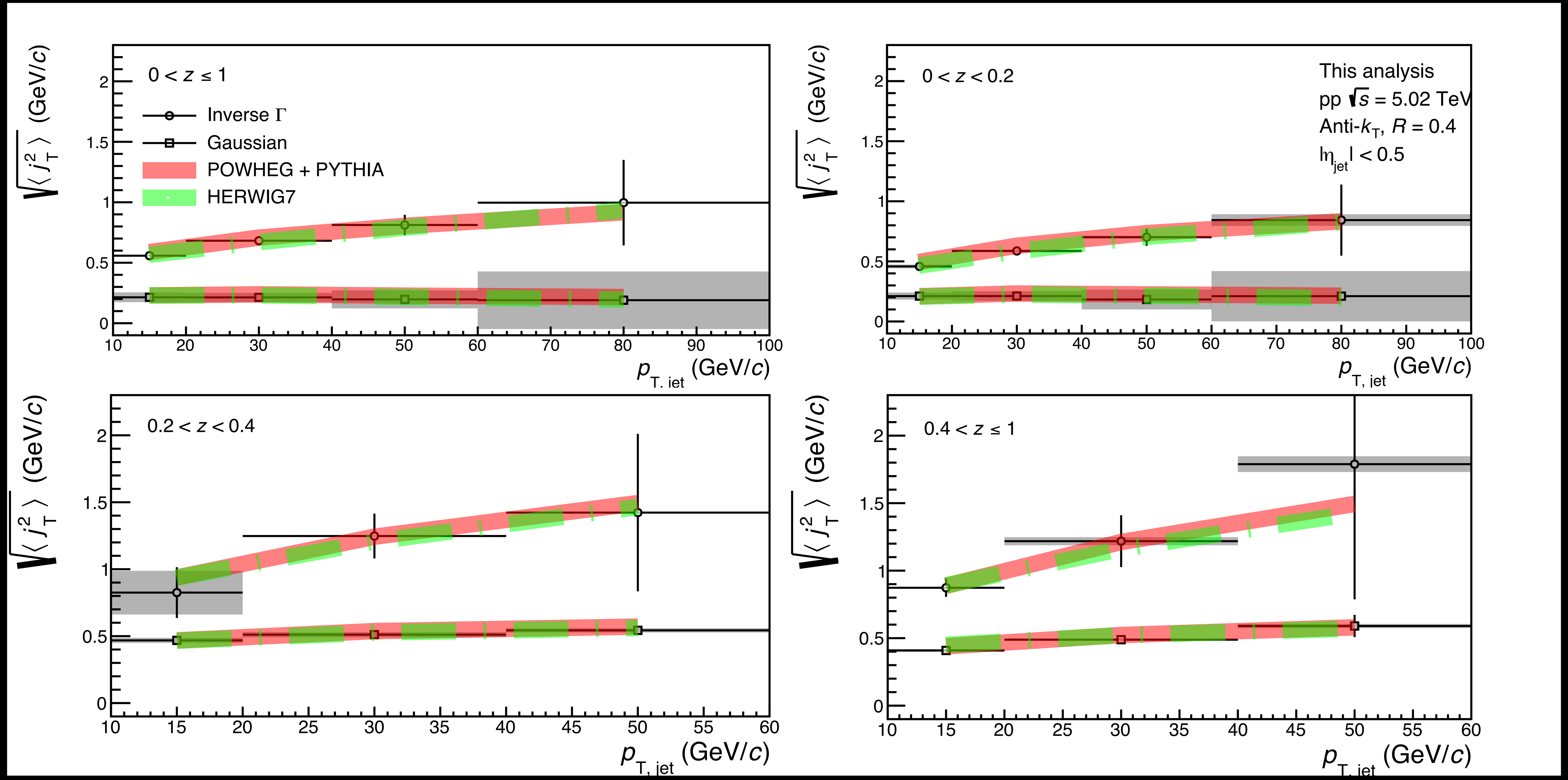
Systematic errors

- Drew RMS plots without PYTHIA8
- Systematic errors for the RMS distribution are implemented
  - New j_T distributions created by applying total systematic error to center points, scaling up and down
  - RMS values extracted from the scaled-up, scaled-down, and original j_T distributions
  - Differences in RMS values were calculated to measure systematic error

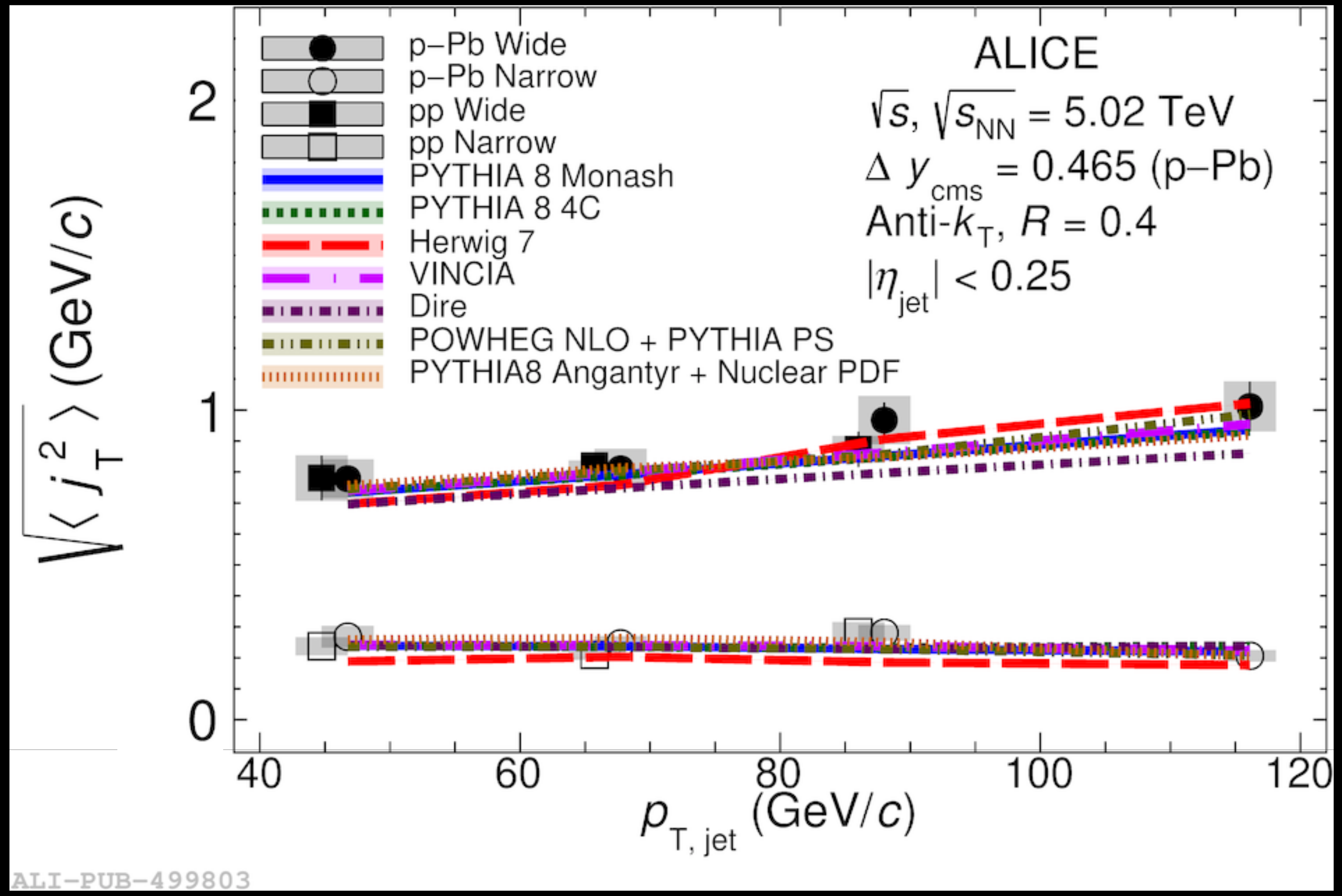


Results

RMS j_T - $p_{T,jet}$



- Drew RMS plots without PYTHIA8
- Systematic errors for the RMS distribution are implemented
  - New j_T distributions created by applying total systematic error to center points, scaling up and down
  - RMS values extracted from the scaled-up, scaled-down, and original j_T distributions
  - Differences in RMS values were calculated to measure systematic error

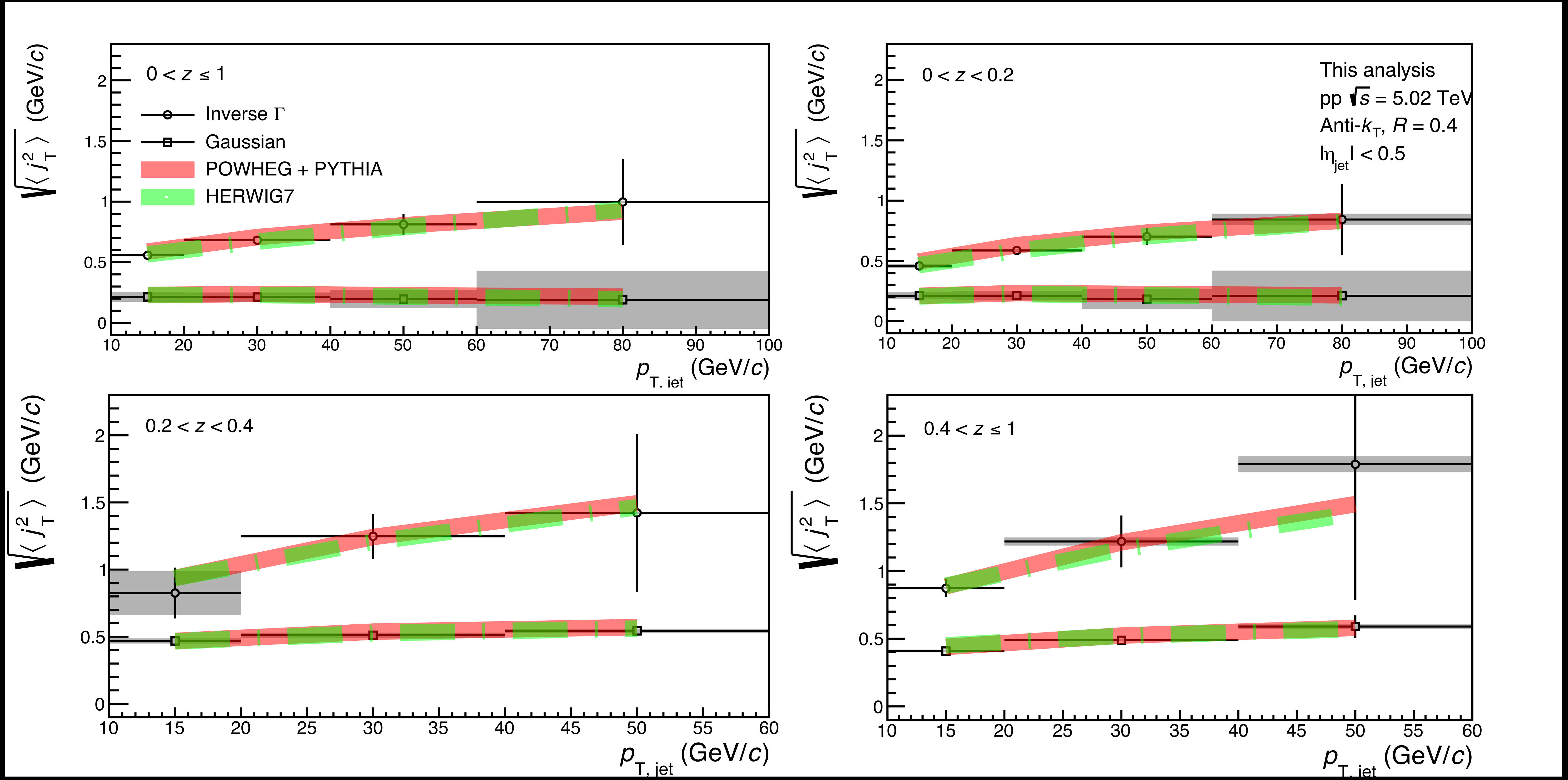


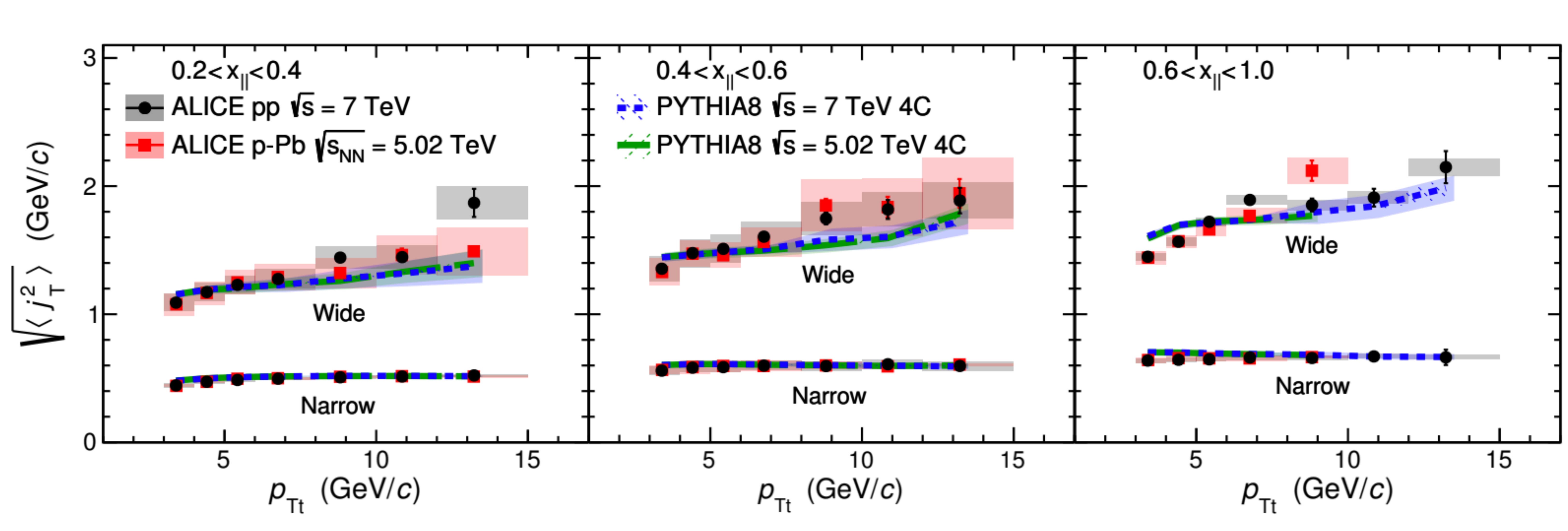
JHEP09(2021)211

Results

RMS j_{T} - $p_{\mathrm{T,jet}}$

- Drew RMS plots without PYTHIA8
- Systematic errors for the RMS distribution are implemented
  - New j_{T} distributions created by applying total systematic error to center points, scaling up and down
  - RMS values extracted from the scaled-up, scaled-down, and original j_{T} distributions
  - Differences in RMS values were calculated to measure systematic error

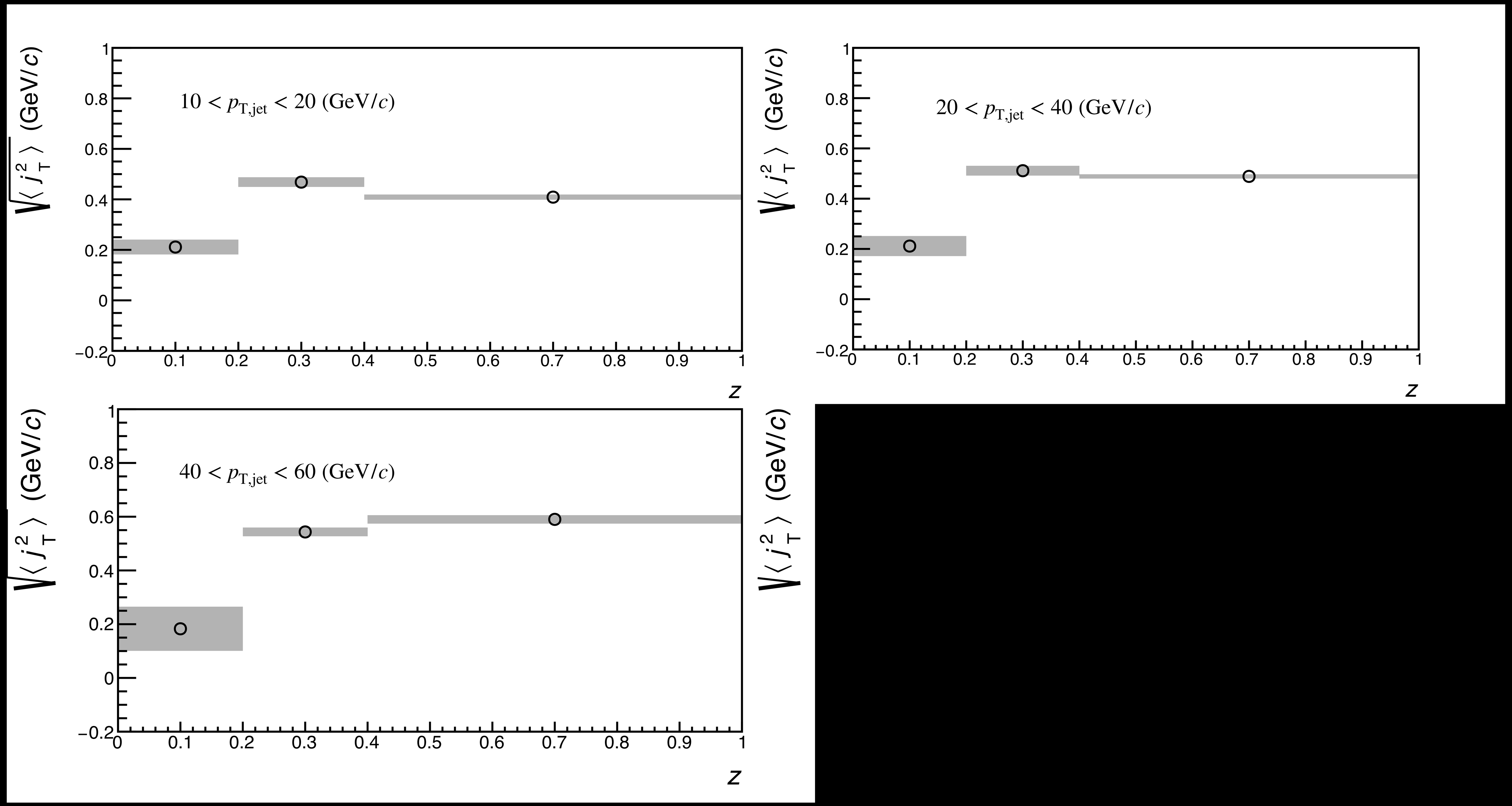




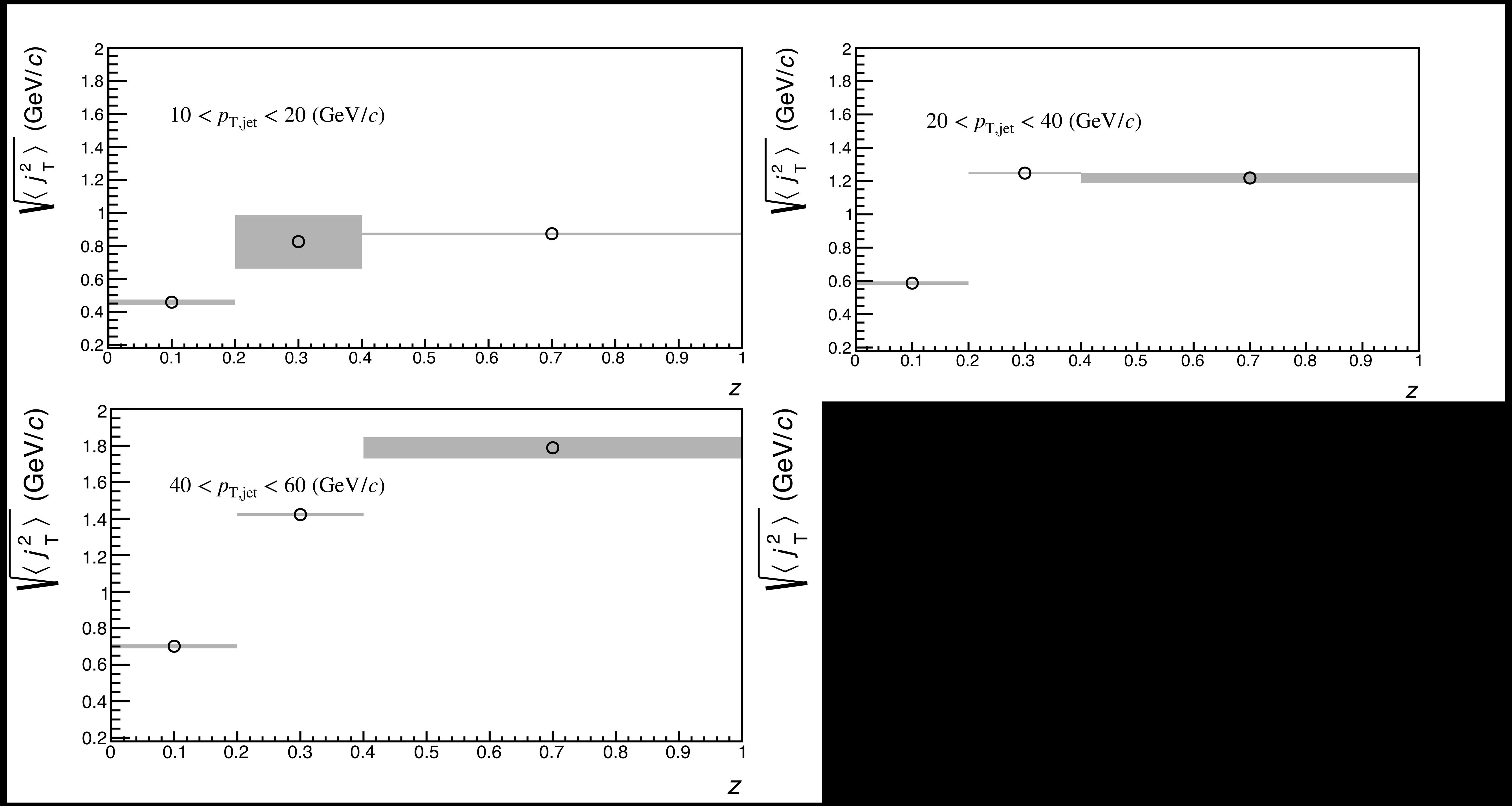
JHEP03(2019)169

Results

RMS j_T - z



Gaussian component



Inverse Γ component

- Observed the RMS j_T values from each fitting as a function of z
- Also planning to calculate the j_T yields from each fitting