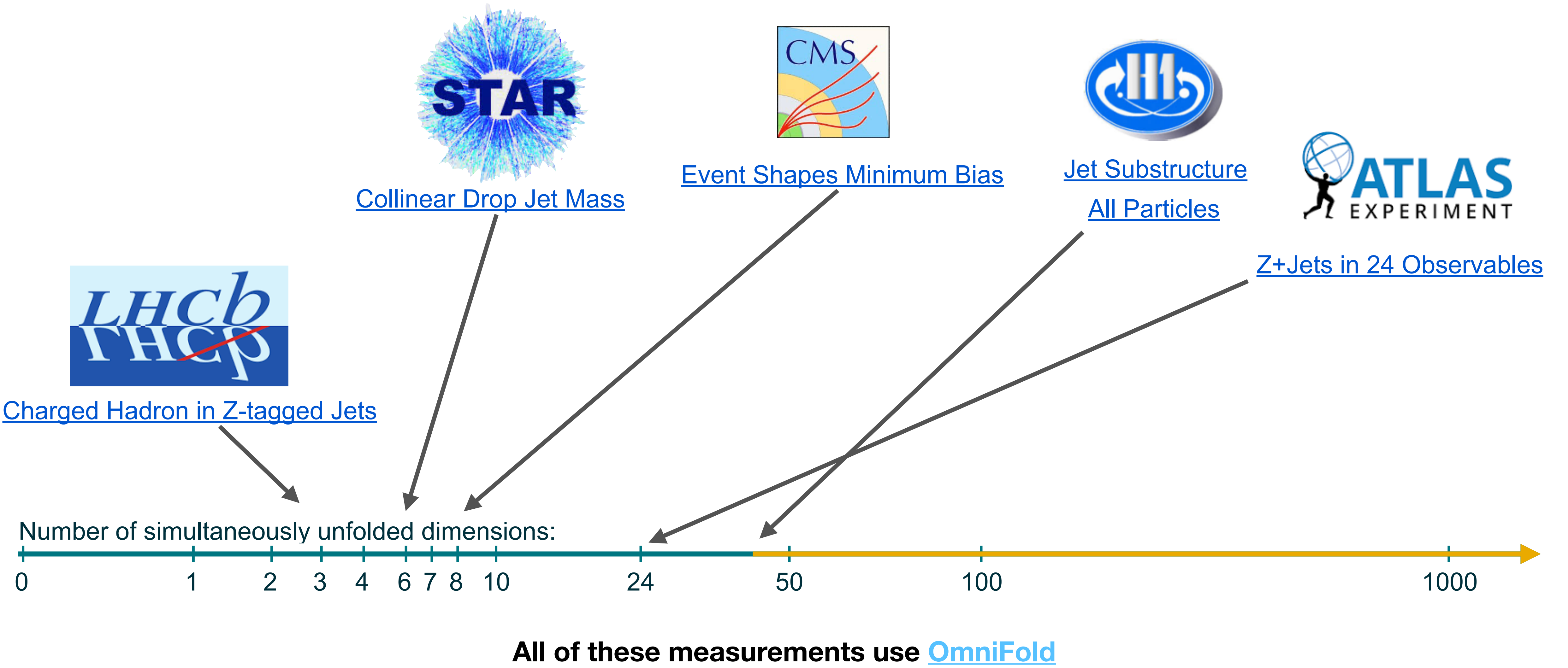




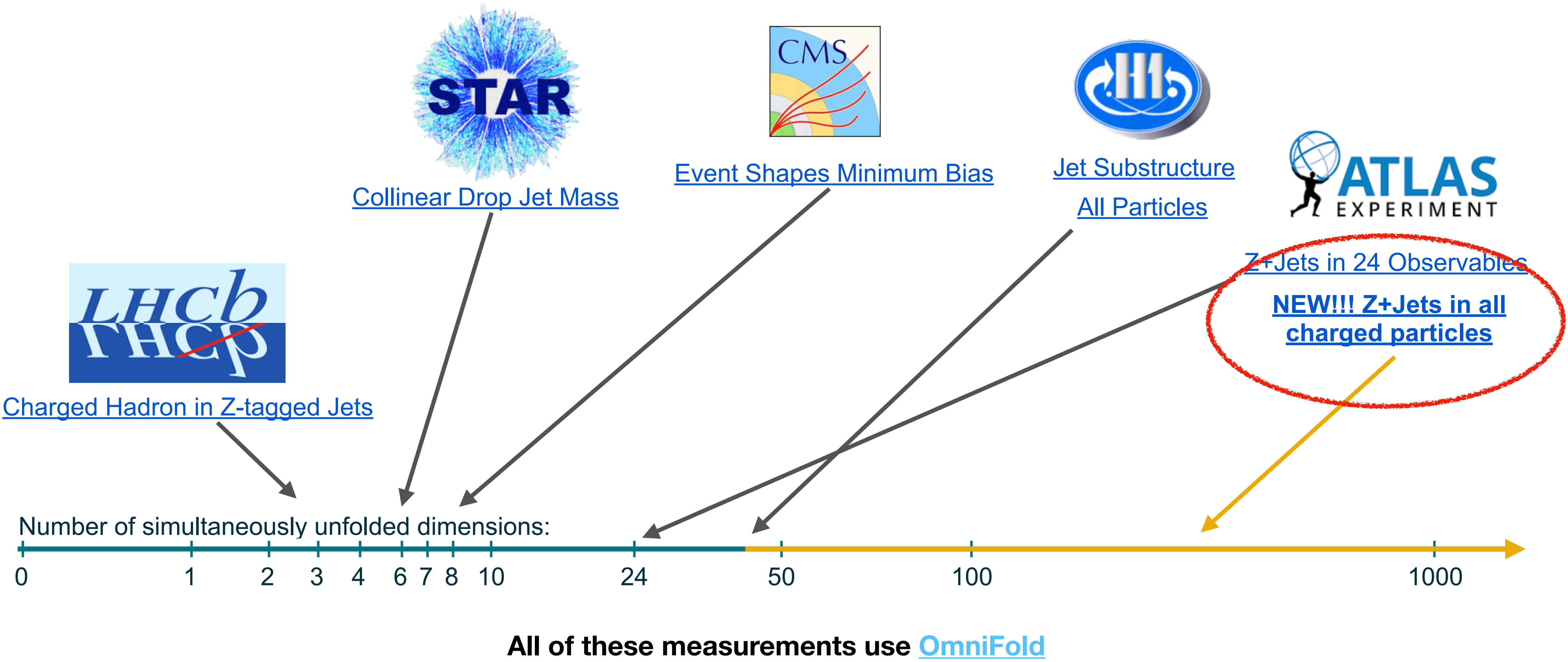
A full phase space measurement of Z +jets production

Kevin Greif, for the ATLAS Collaboration
MLJets, September 17th 2026

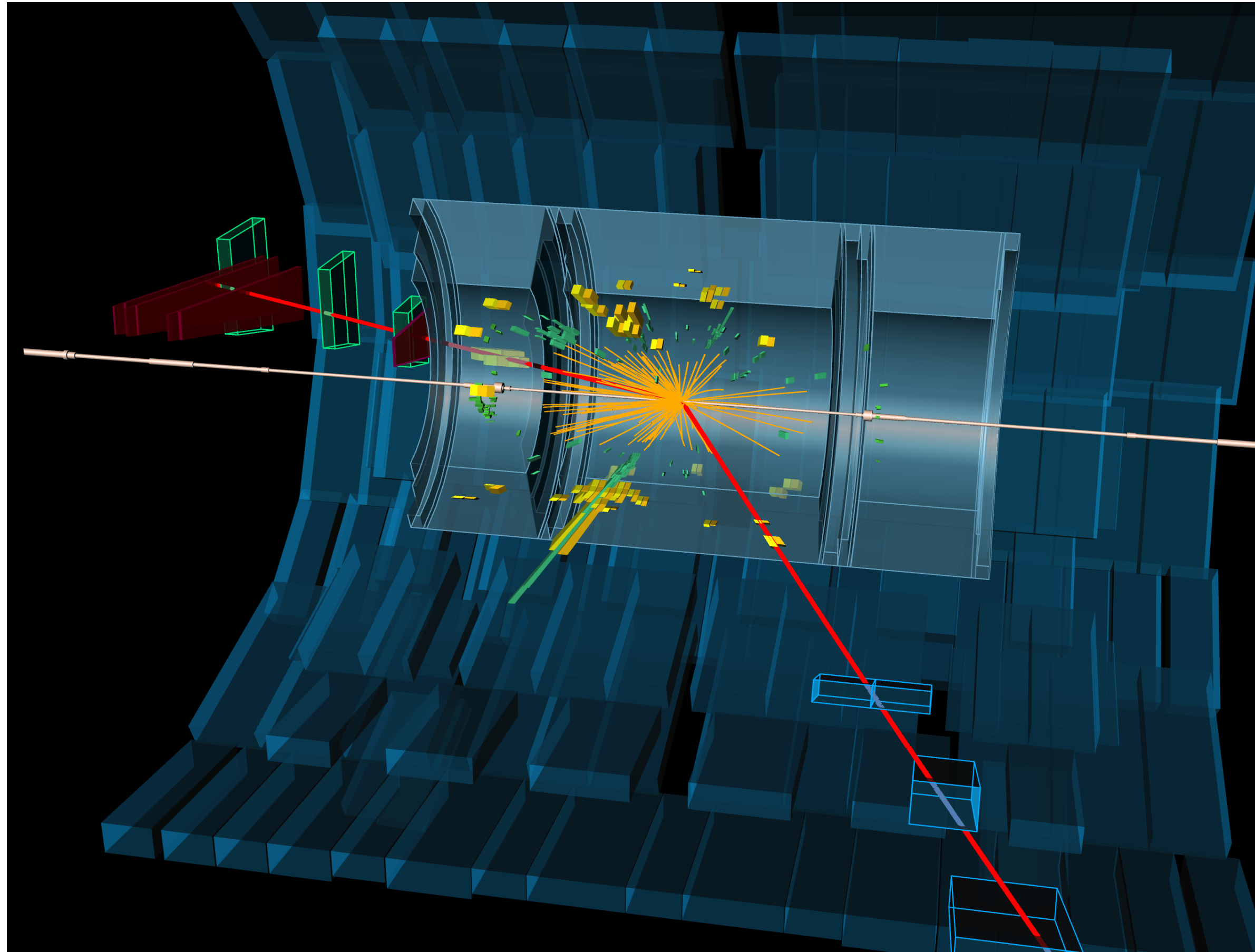
Machine learning based unfolding doing physics



Machine learning based unfolding doing physics



$Z (\rightarrow \mu\mu) + \text{jets}$ production at the LHC



1. Z boson decay

- Muons can be used for triggering and event selection
- Has been extensively studied in pp collisions
- Inclusive cross section is known to NNLO

2. Hadronic recoil

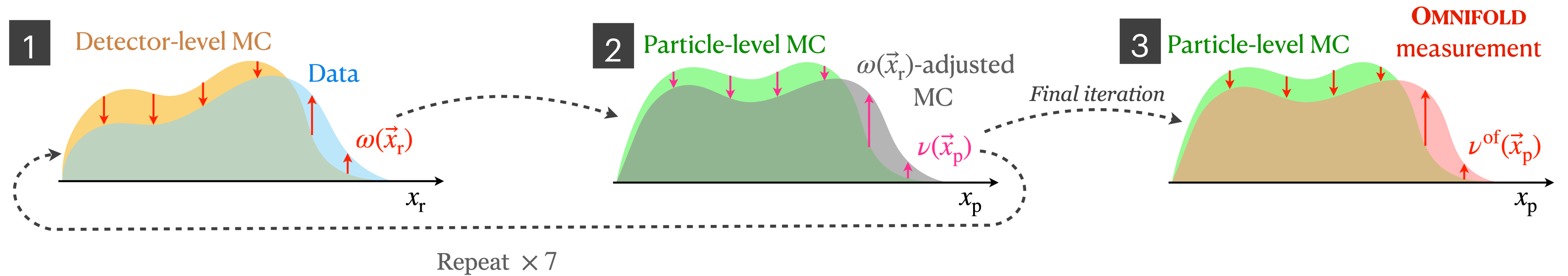
- **Excellent laboratory for studying QCD and jet substructure**
- Unbiased sample of largely quark initiated jets

3. Underlying event

- Relatively unstudied, but contains a lot of interesting physics
- Useful for tuning event generators, modelling can be terrible!

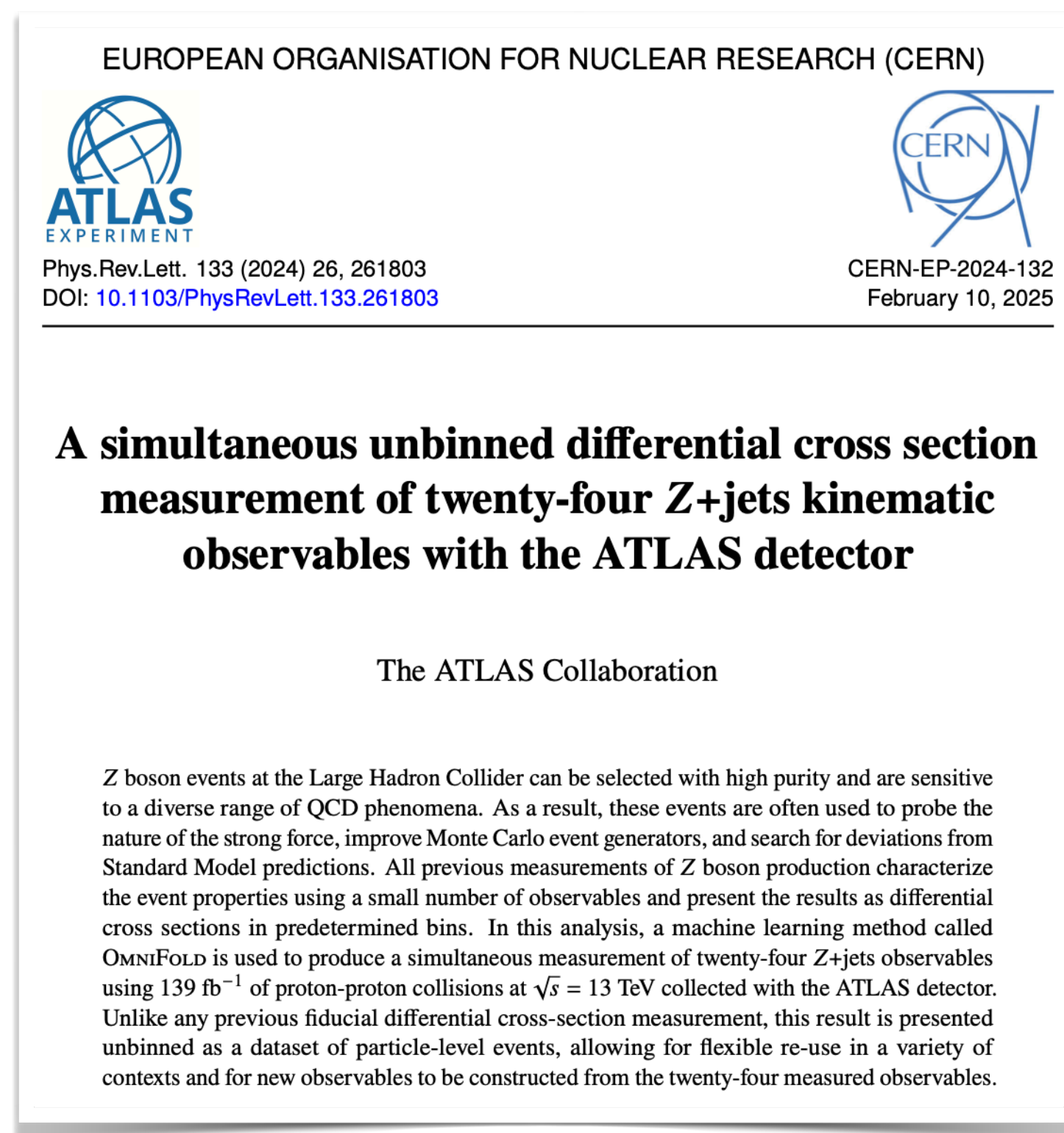
“Full phase space measurement” means we will study all of this at once, through ML based unfolding

OmniFold in brief

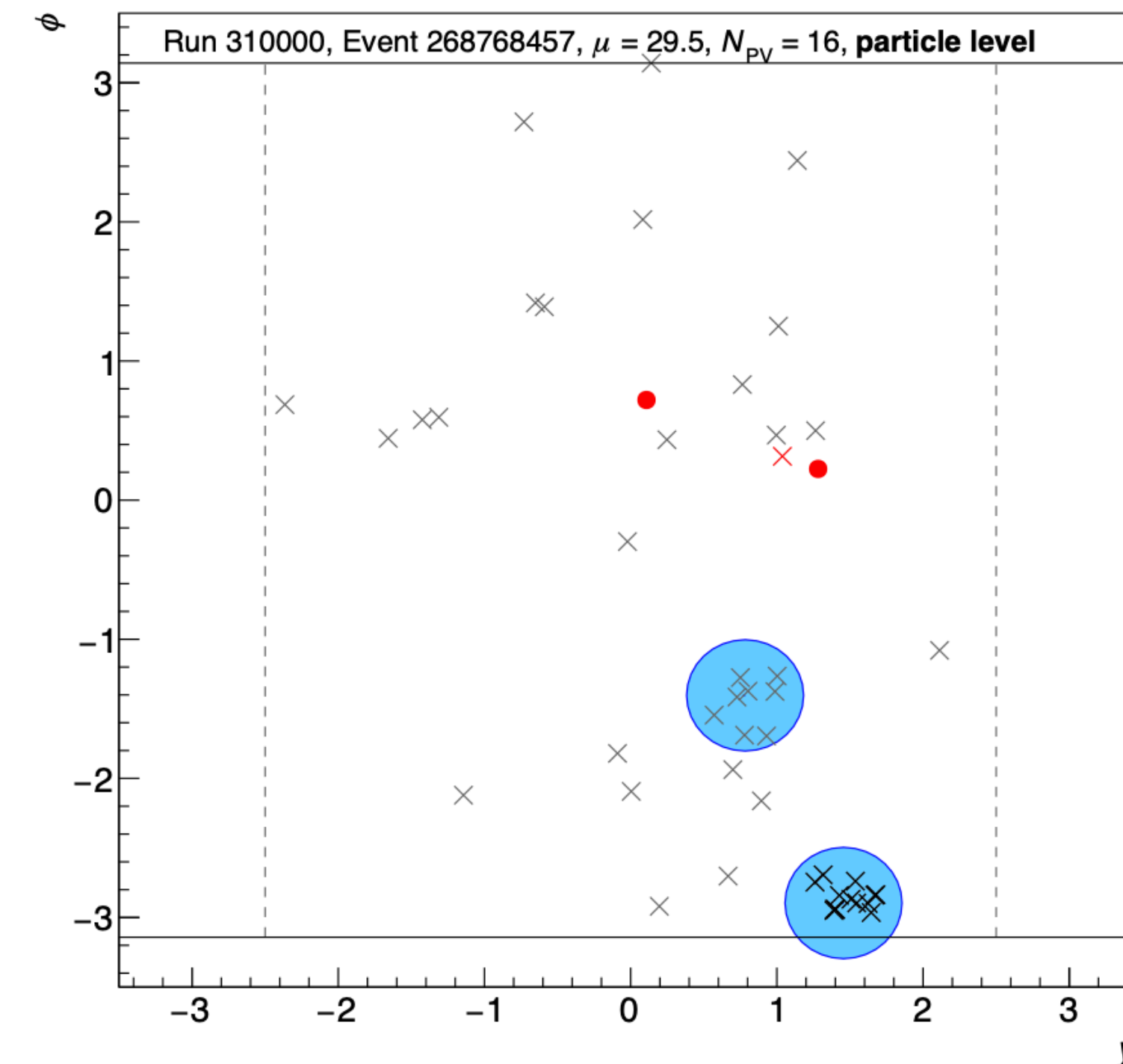
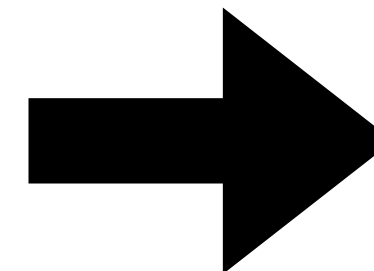


ATLAS Z+jets Omnifold measurements

- Previous measurement: **24 event-level observables**
- New measurement: **full charged-particle phase space, variable dimensional!**



[PRL 133 \(2024\) 261803](https://arxiv.org/abs/2402.05811)



p_T, η, ϕ of all
stable charged
particles are
measured
simultaneously

Mean of ~ 150 simultaneous dimensions

Lessons learned 1: compute is a challenge

$$N = 2 \times 7 \times$$

Run step 1 and step 2

Iterate 7 times

Data stat bootstrap

MC stat bootstrap

$$(10 \times 26 + (100 + 25))$$

Reduce NN initialization uncertainty

Propagate systematic uncertainties

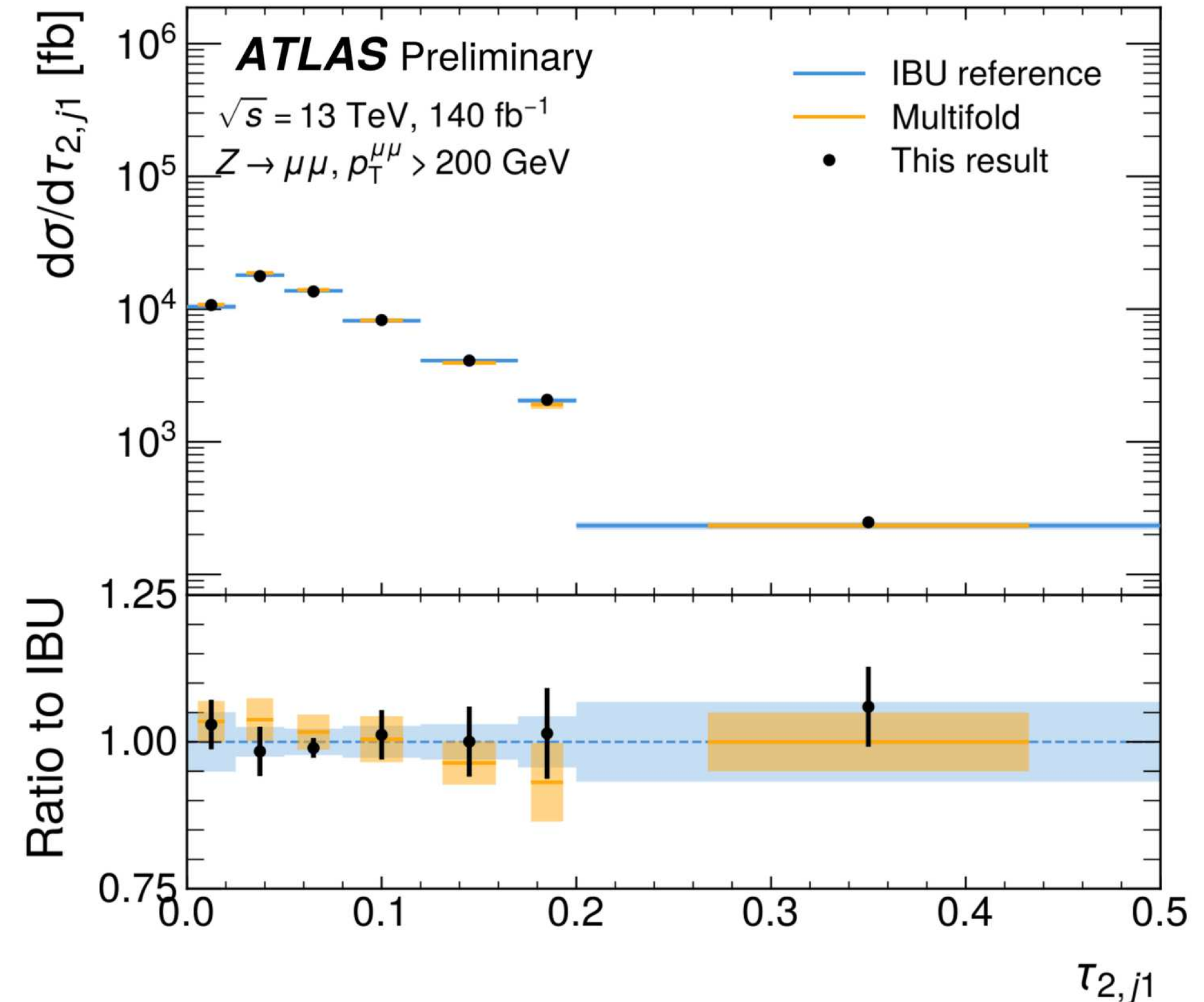
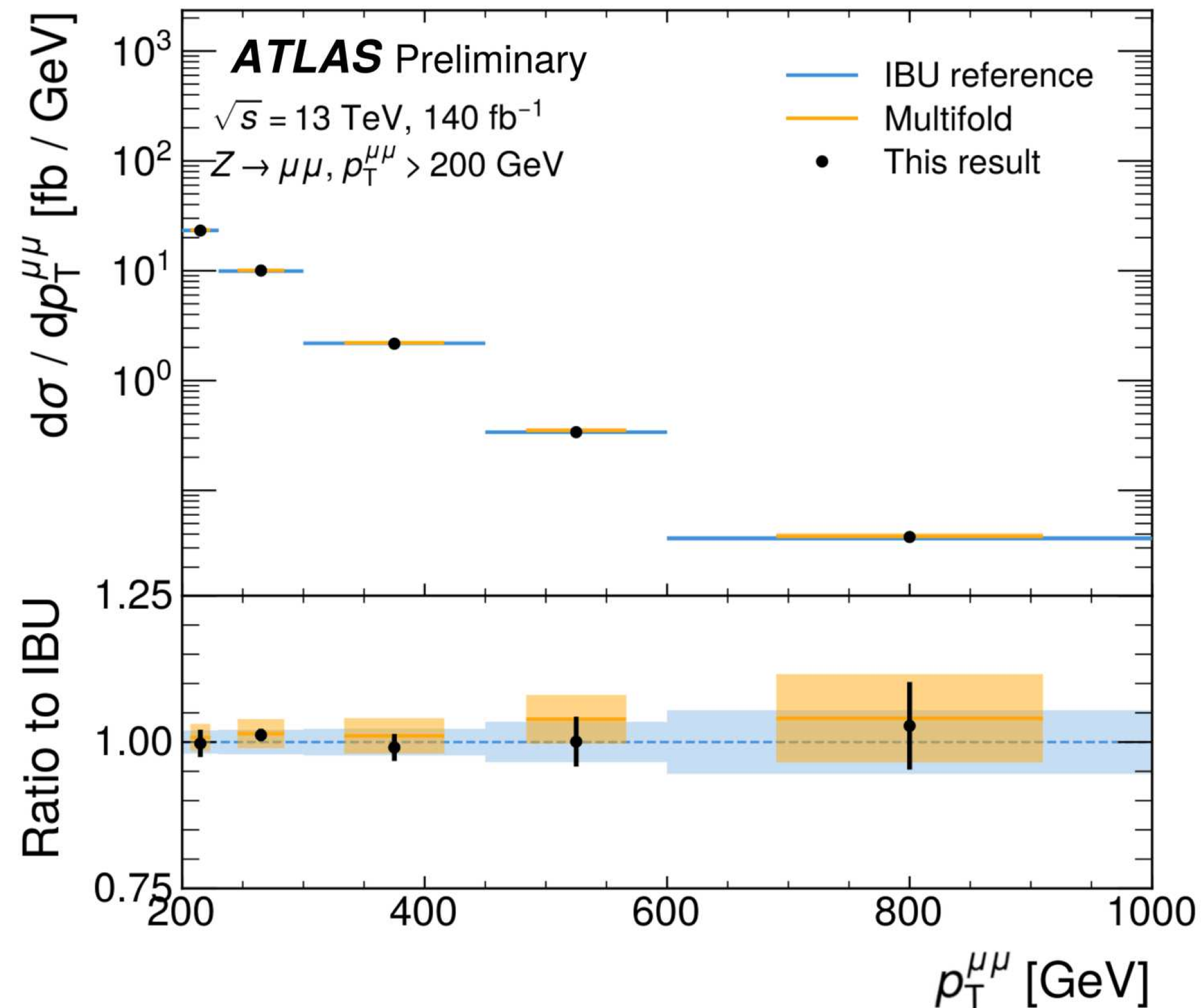
$$= 5390$$

network fine-tunings required for measurement



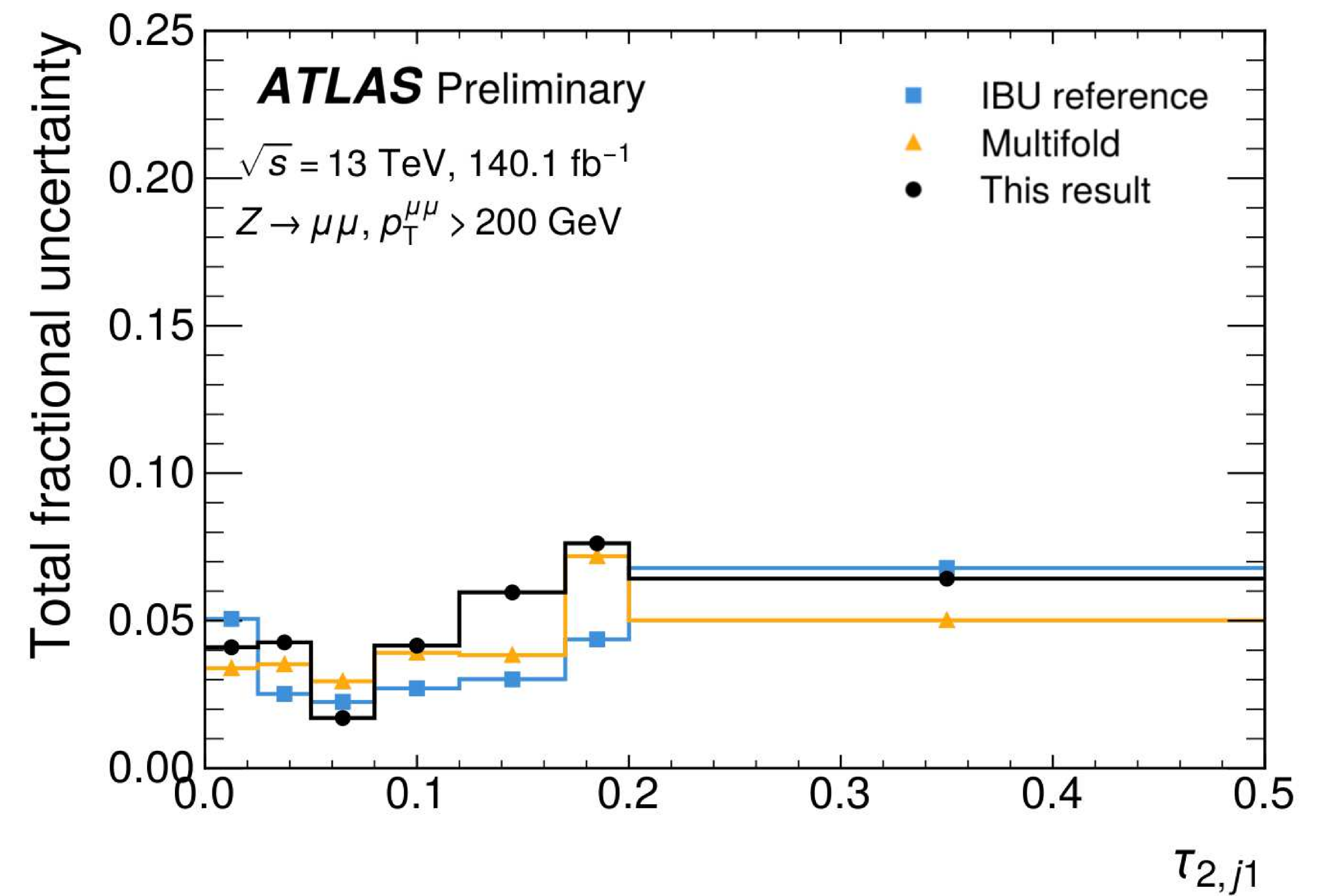
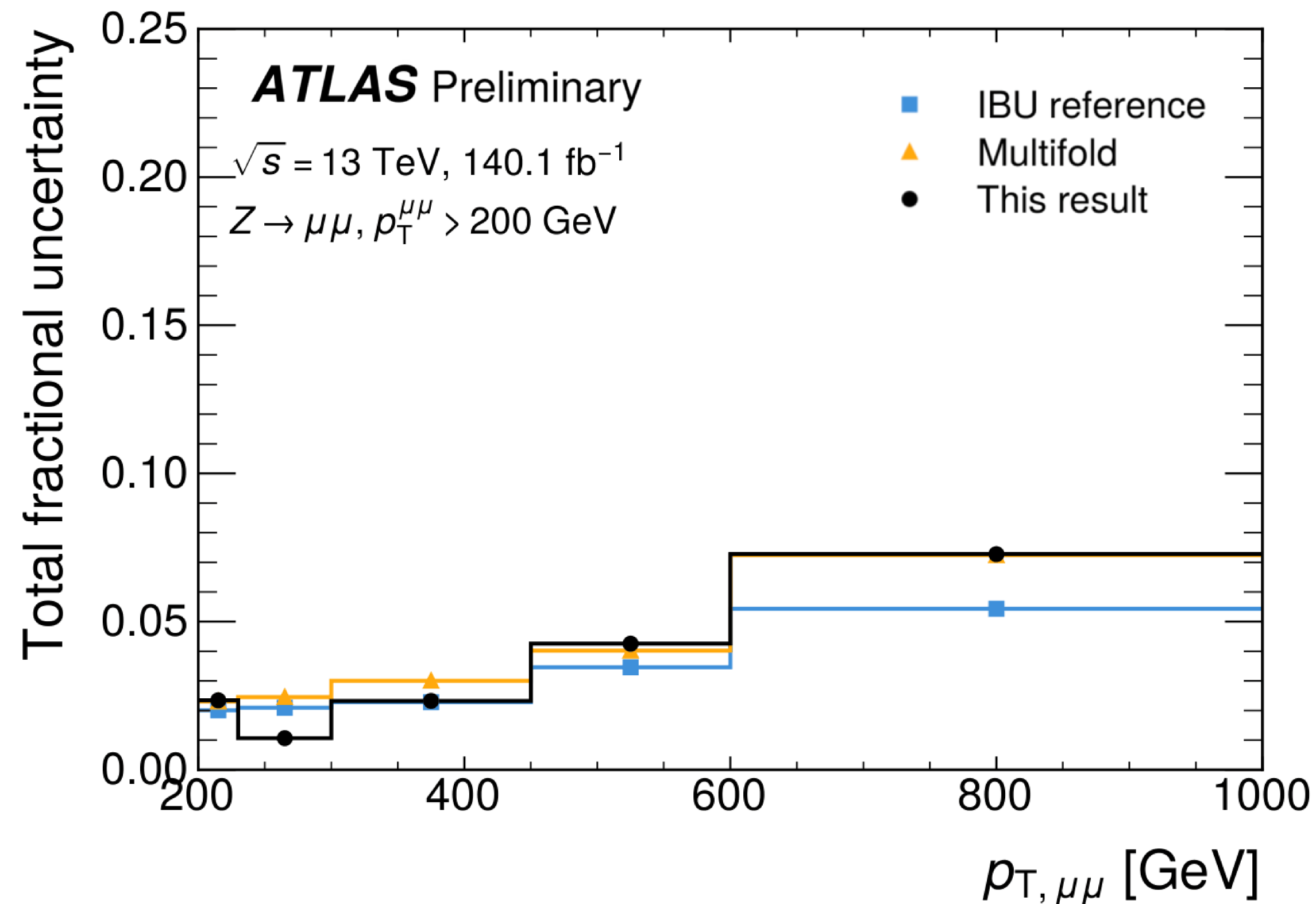
Lessons learned 2: uncertainties are comparable

- Compared this measurement to traditional binned measurements and previous round 24-dimensional measurement (multifold)

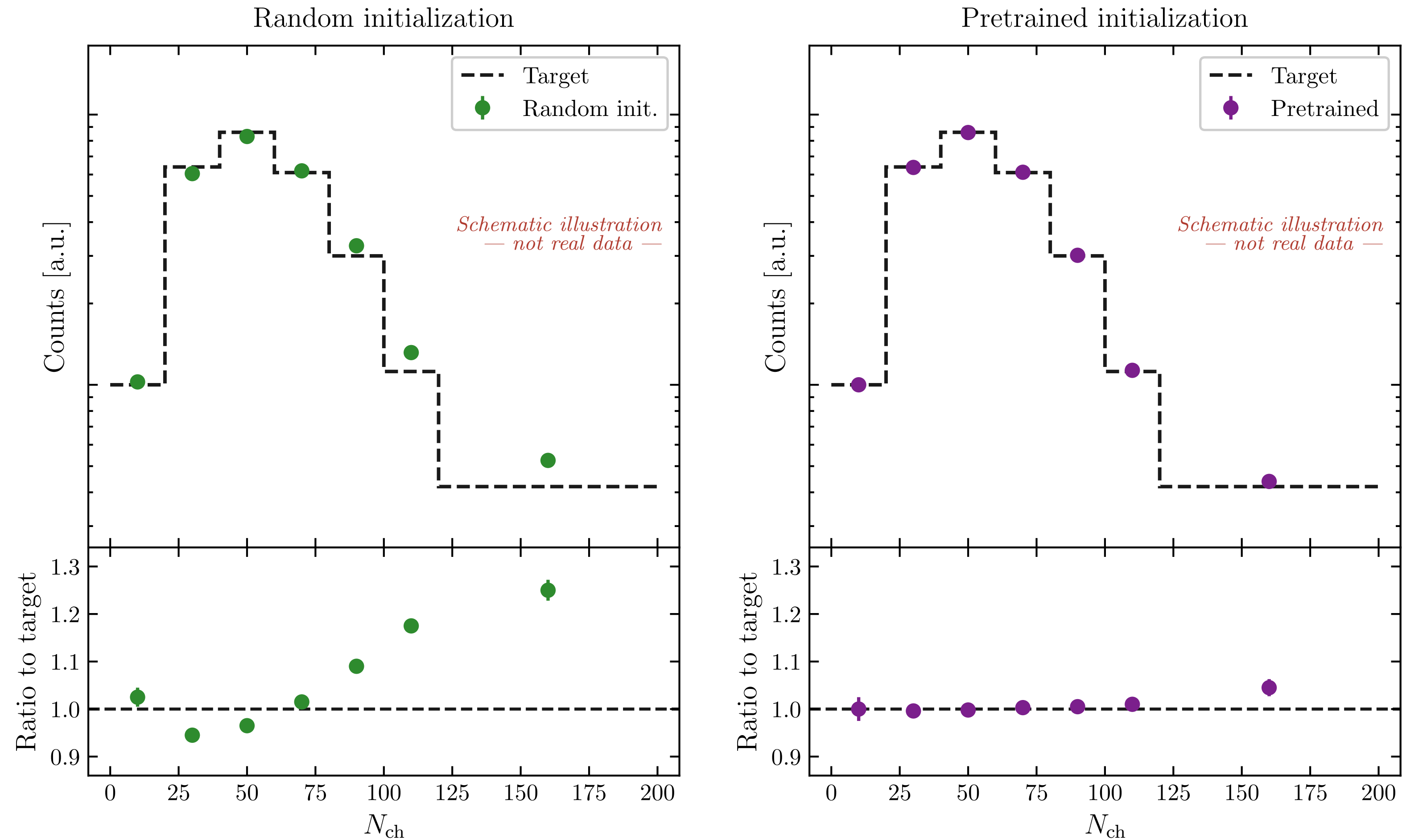


Lessons learned 2: uncertainties are comparable

- Compared this measurement to traditional binned measurements and previous round 24-dimensional measurement (multifold)



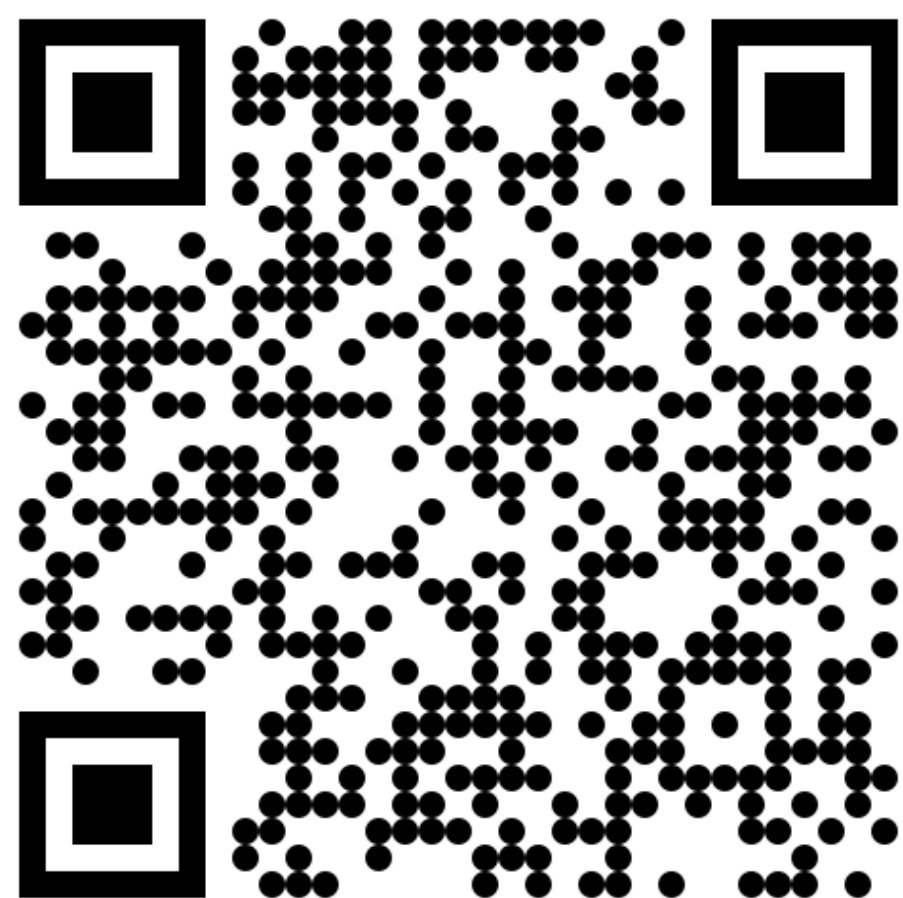
Lessons learned 3: pretraining is essential



Only 247k events enter signal region: pretraining networks on simulation drastically improves accuracy!
Essential context for work on foundation models: data constrained likelihood ratio estimation should be benchmarks

Public data and results

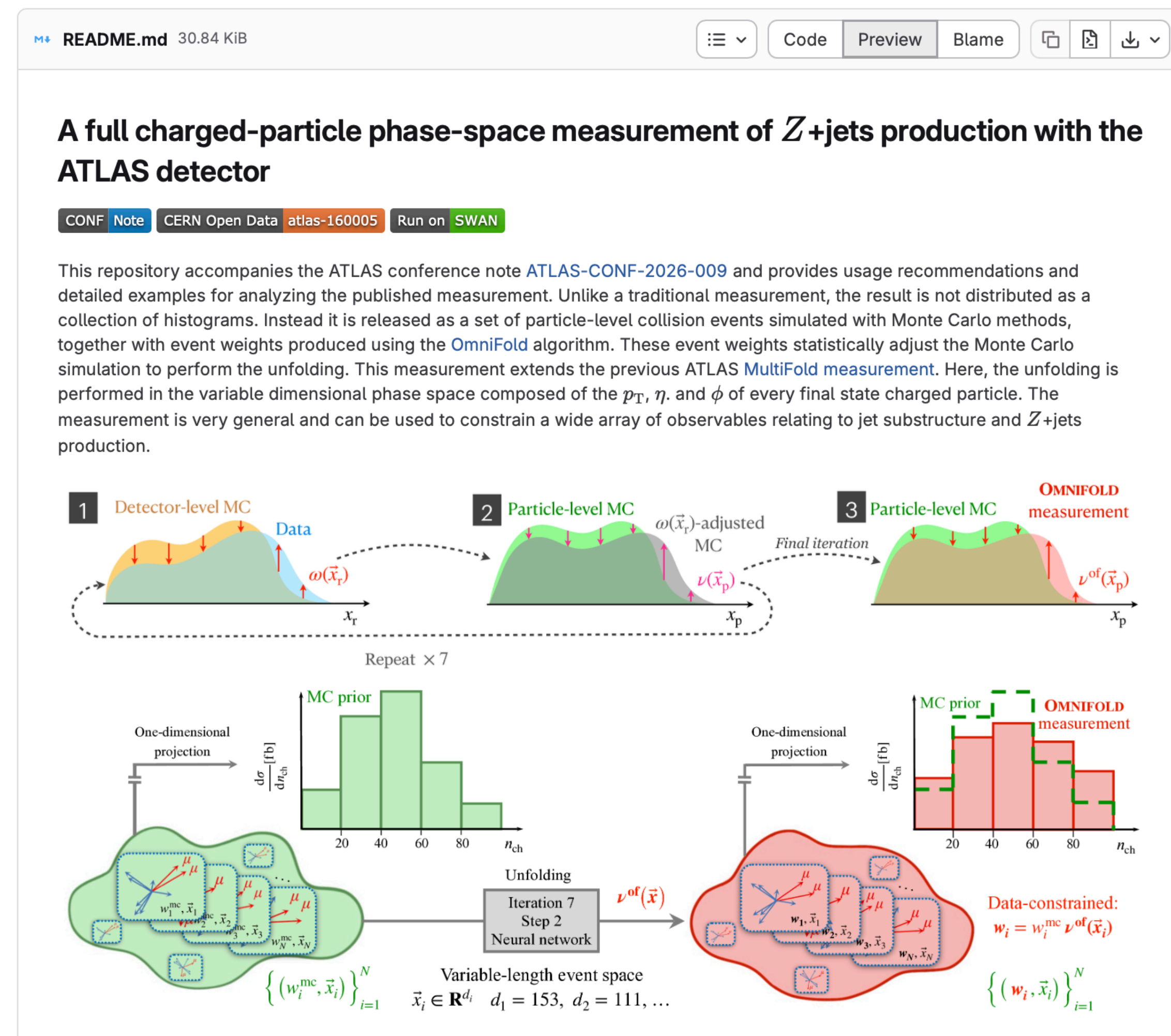
- Measurement is natively un-binned and high-dimensional
- Publish measurement is particle-level events and weights
 - Dataset, not histogram
- **Downstream analysis encouraged!**



Data



Code



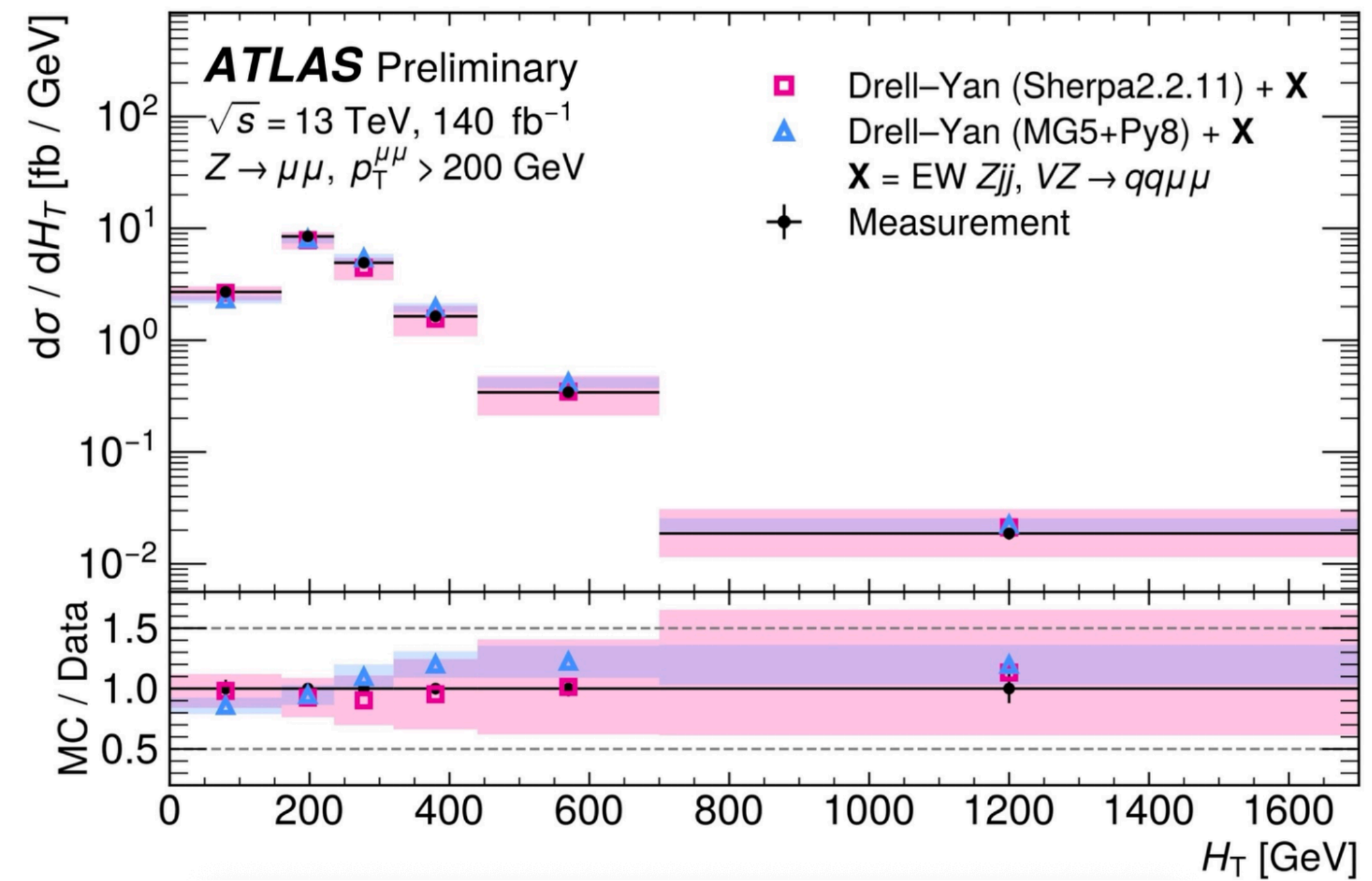
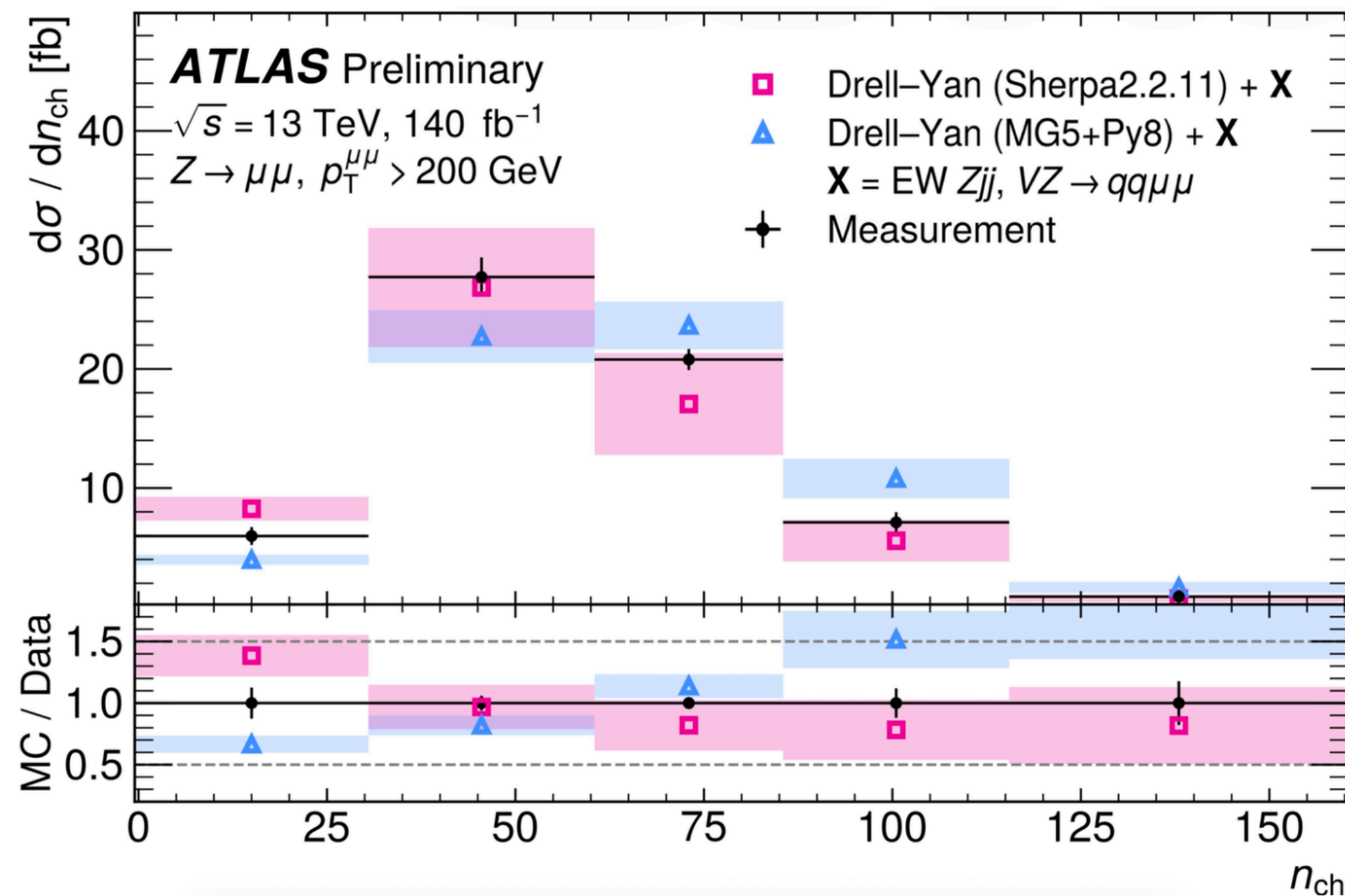
Standard event-level observables

- Number of charged hadrons (N_{ch}) and H_{T} are IRC unsafe observables
- Measurement is **simultaneous and un-binned**. Key abilities:

Re-binning on the fly

2D Histograms

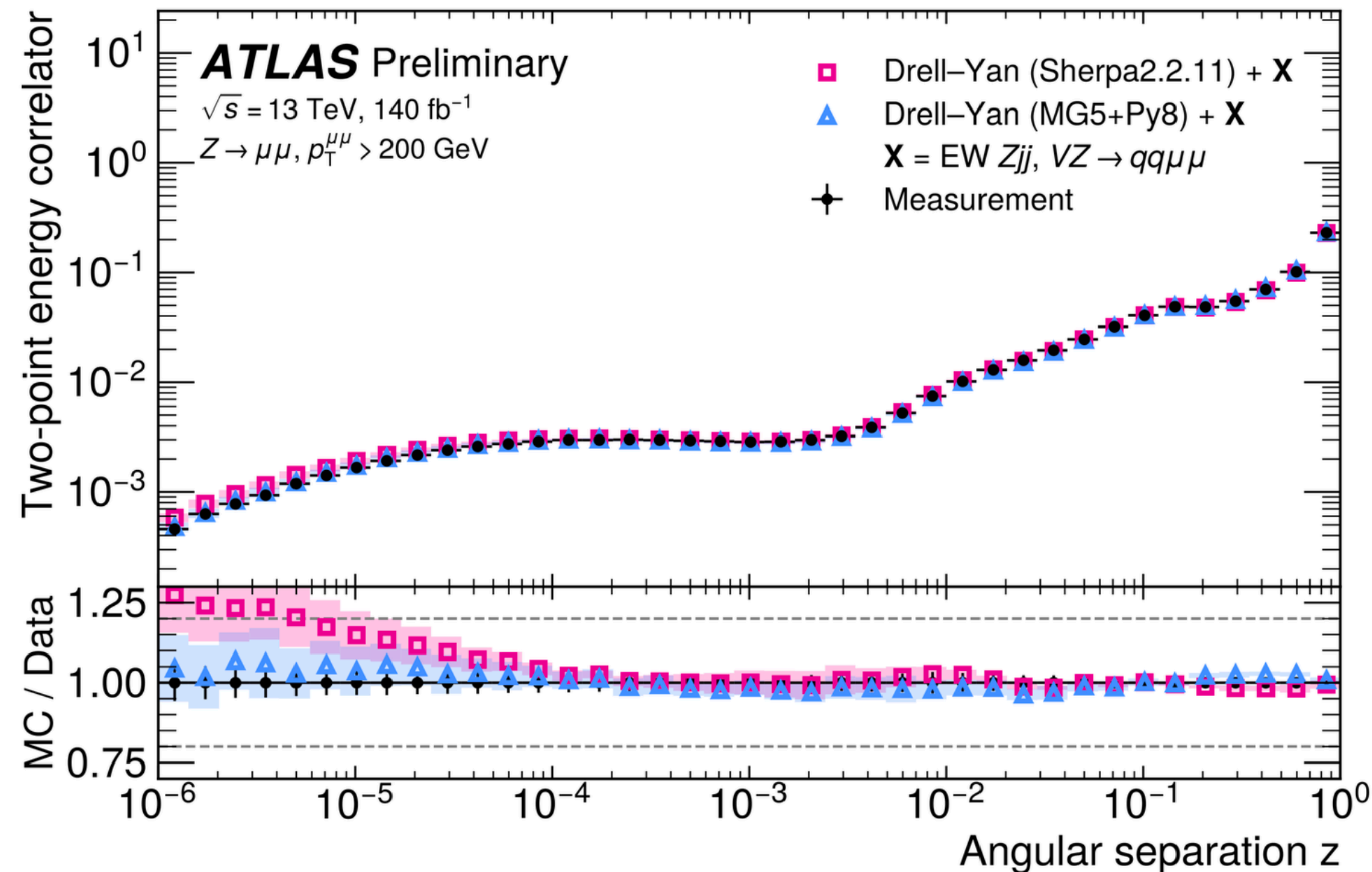
Phase space cuts



Two point energy correlator measurement

$$EEC(z) = \frac{1}{\sigma} \sum_{i,j} \int d\sigma \frac{E_i E_j}{Q^2} \delta(z - \frac{1 - |\cos \theta_{ij}|}{2}),$$

- Two-point energy correlator (EEC) measured with **all charged particles**
- Difficult to unfold with binned methods!
 - Omnifold handles two-point and higher order correlators naturally



Not covered due to time:

1. Jet p_T profile measurement with different clustering algorithms
2. Intrinsic dimensionality of jet data



Conclusions

- First full phase space measurement to be performed at the LHC: **unimaginable 10 years ago!**
- Compute and training stats are significant bottlenecks
 - Somewhat alleviated with pretraining
- Hopefully many more such measurements in the future

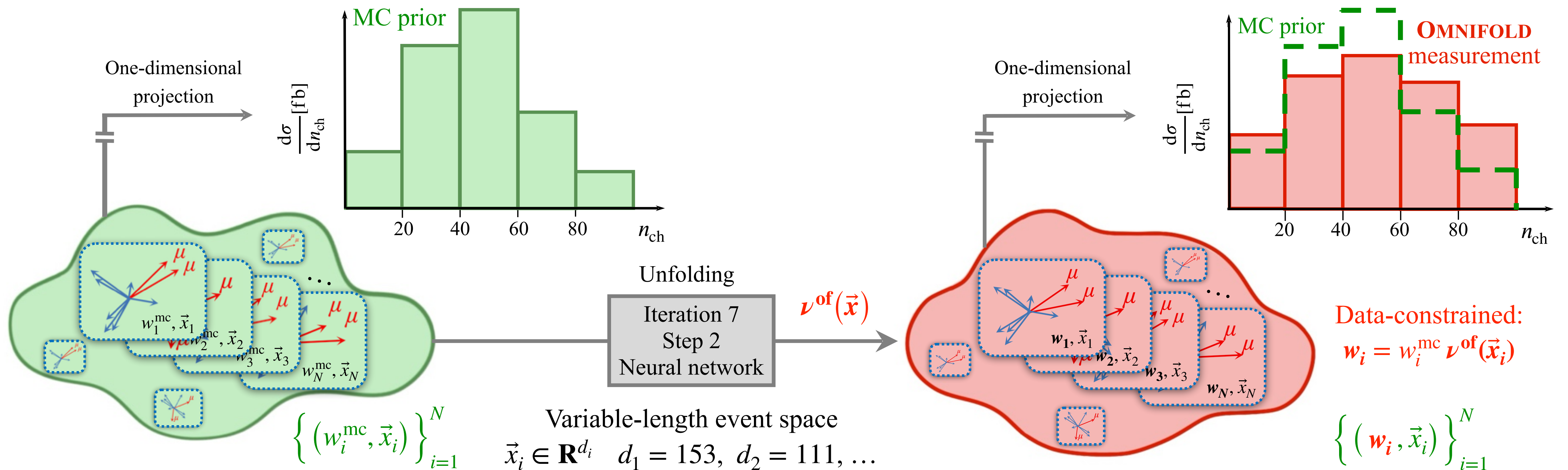


"The Mountain sat upon the Plain
In his tremendous Chair –
His observation omnifold,
His inquest everywhere –"

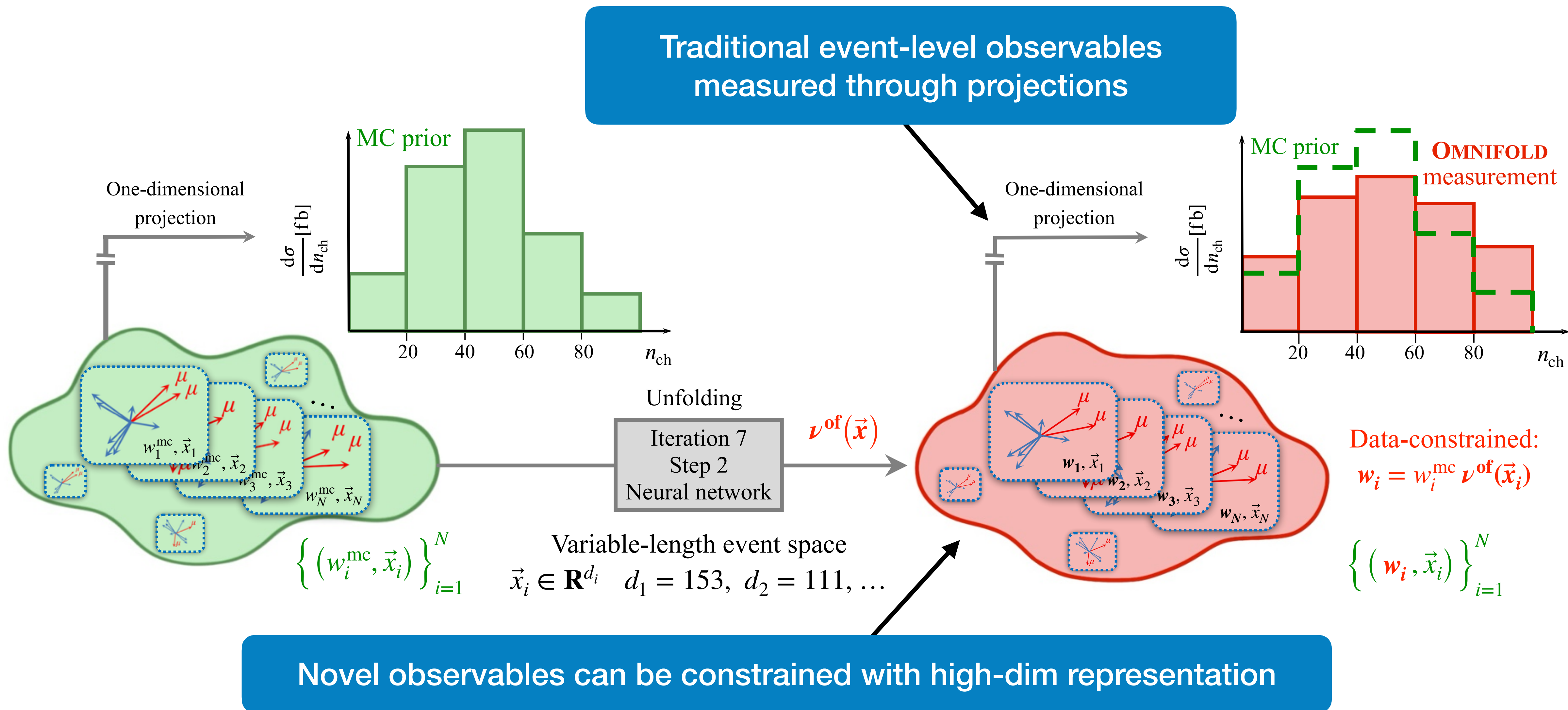
Emily Dickinson, 975

Backup

Final product: particle-level MC and event weights

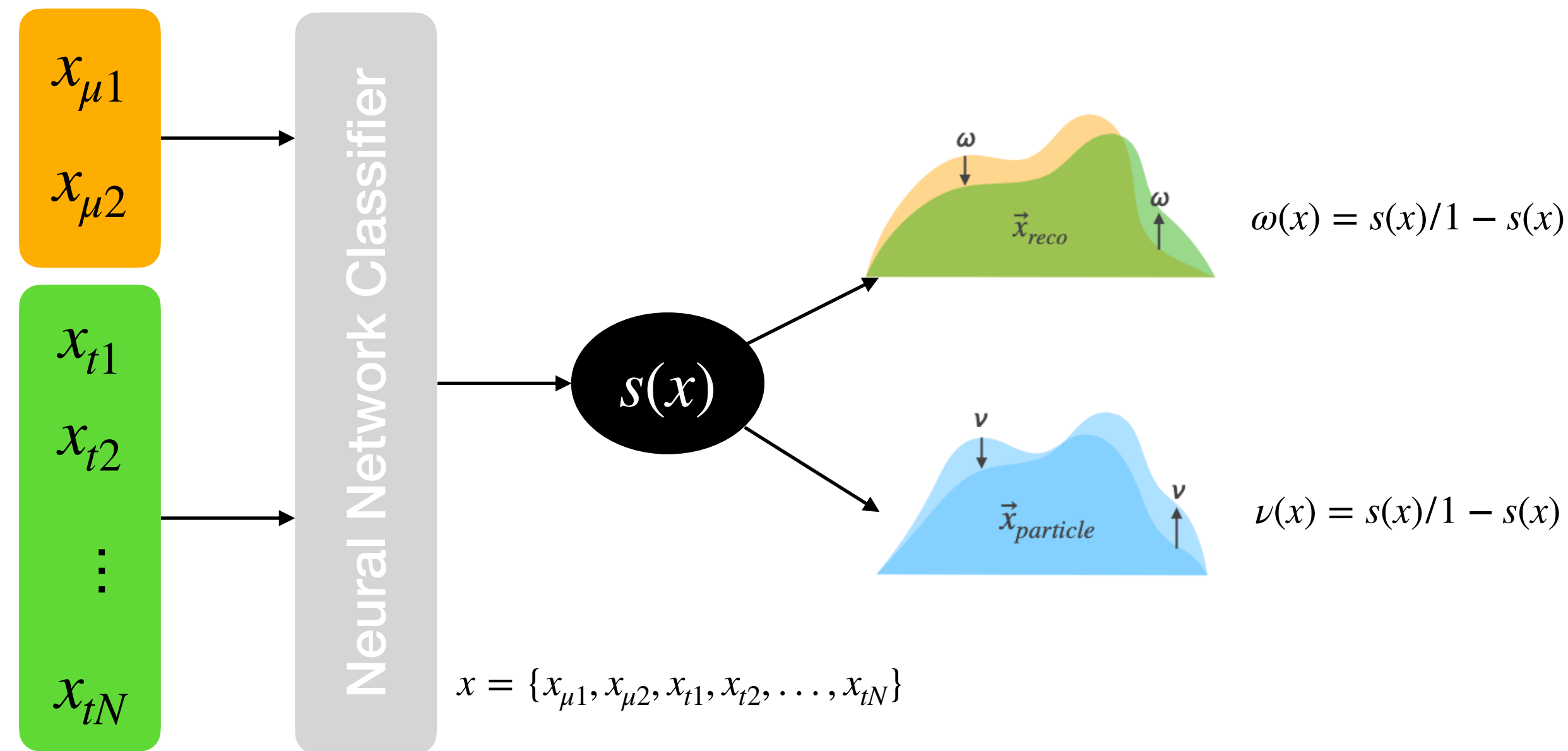


Final product: particle-level MC and event weights



How to measure variable dimensional cross sections

- Need to estimate $\mathcal{O}(100)$ dimensional likelihood ratios
- Utilize transformer neural networks: [ParticleTransformer](#)



Network Inputs

Particles:

- Raw kinematics
 - $(\log p_T, \eta, \phi, m)$
- Jet indices, expressed as one-hot encodings

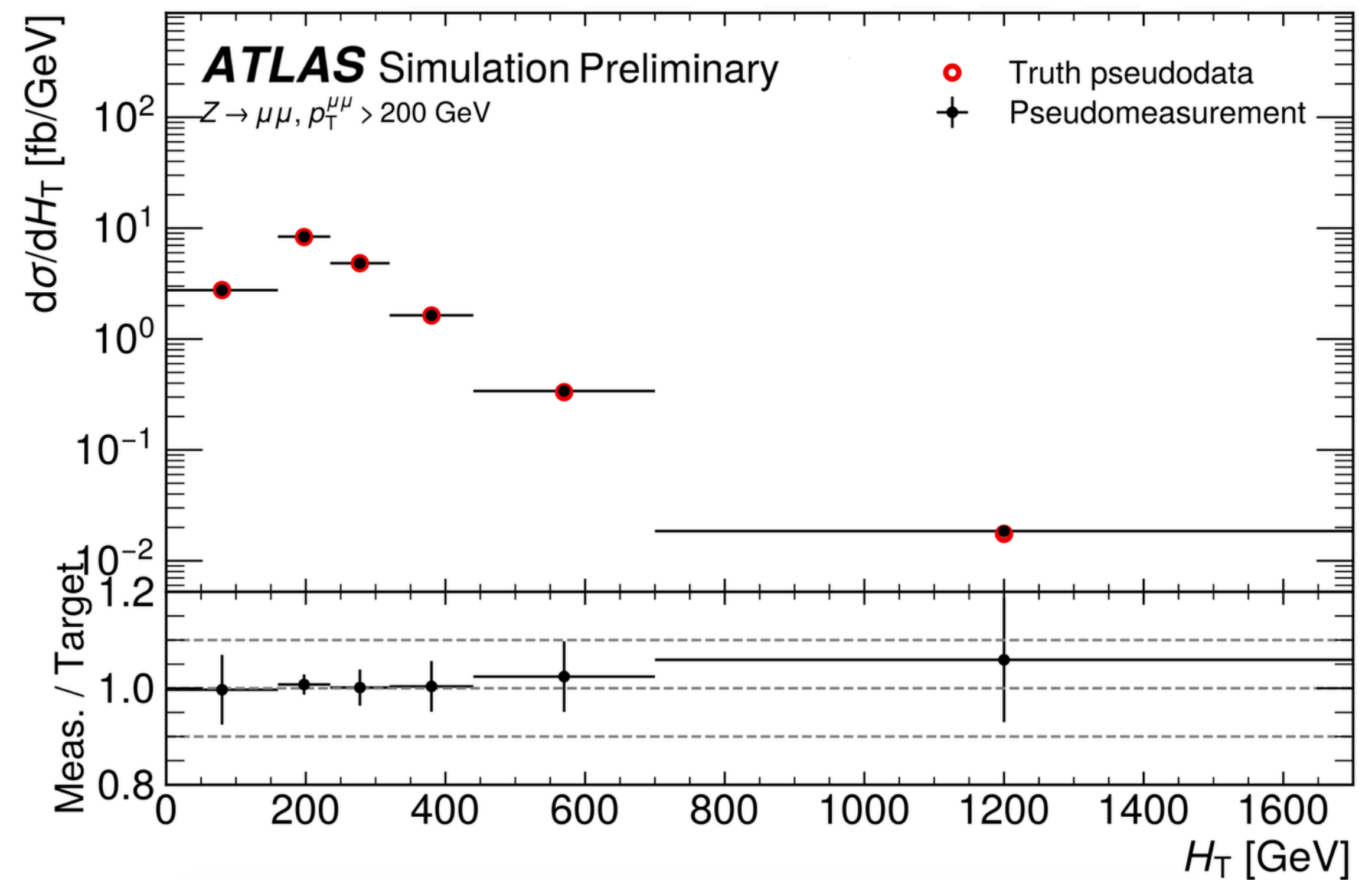
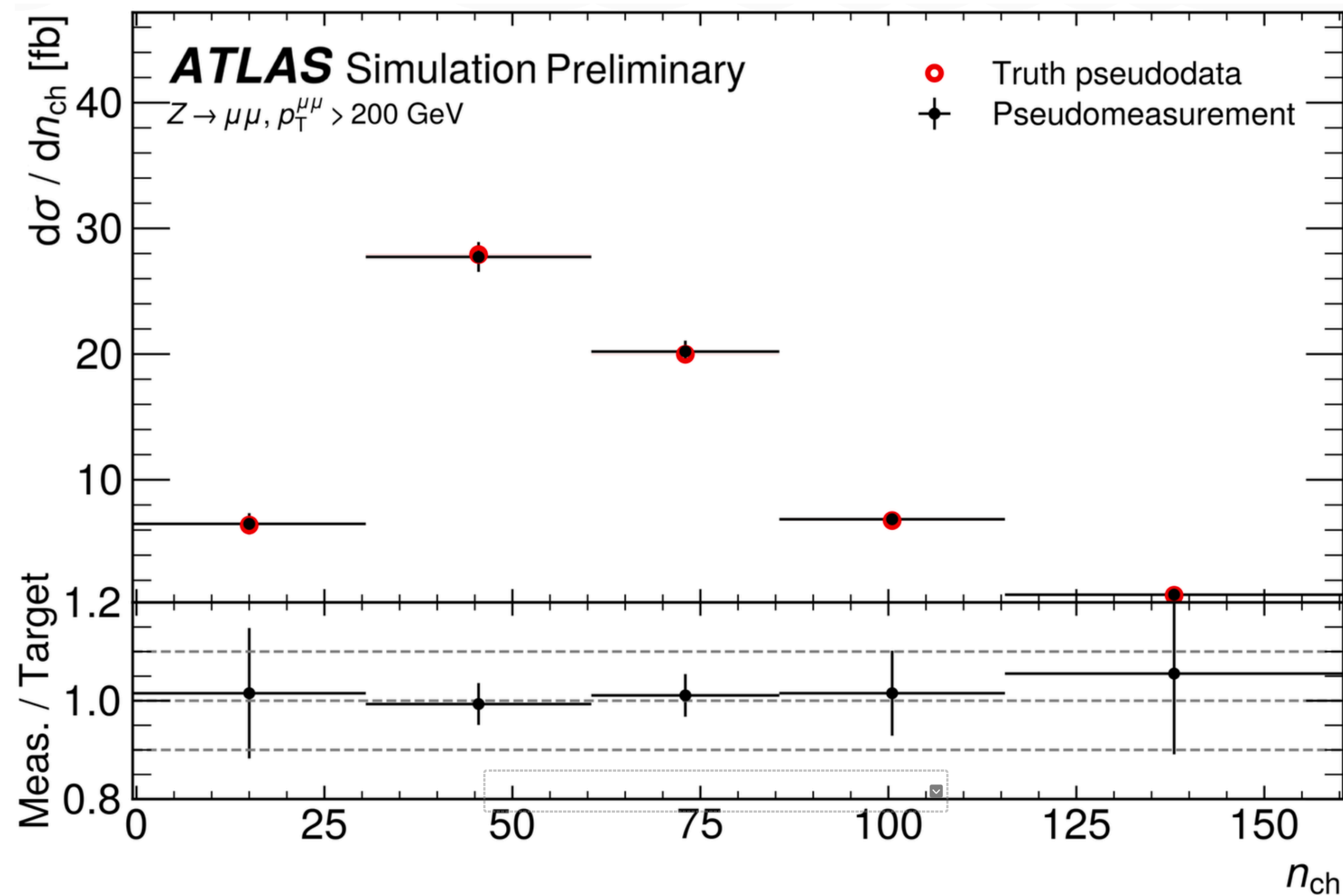
Interactions:

- $\Delta = \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2}$
- $k_T = \min(p_{T,a}, p_{T,b})\Delta$
- $z = \min(p_{T,a}, p_{T,b})/(p_{T,a} + p_{T,b})$
- m^2

Large scale *pretraining* of neural networks on MC simulation drastically improves the unfolding accuracy

Simulation based validation

- How to validate measurement of an arbitrary observable?
- Publish auxiliary **pseudodata** measurement with known particle-level target
- Bin measurement and target in the same observable and calculate χ^2



Simulation based validation

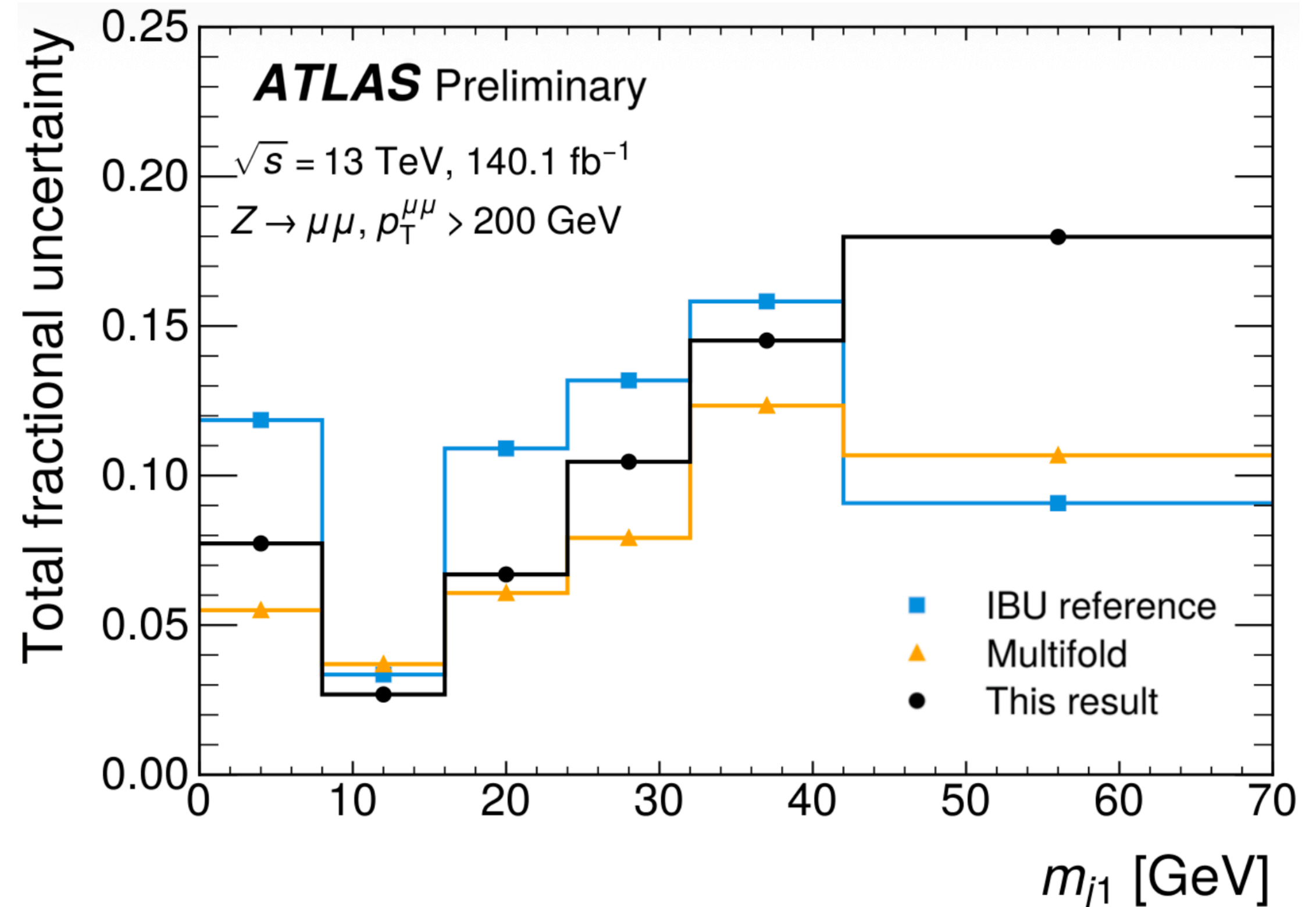
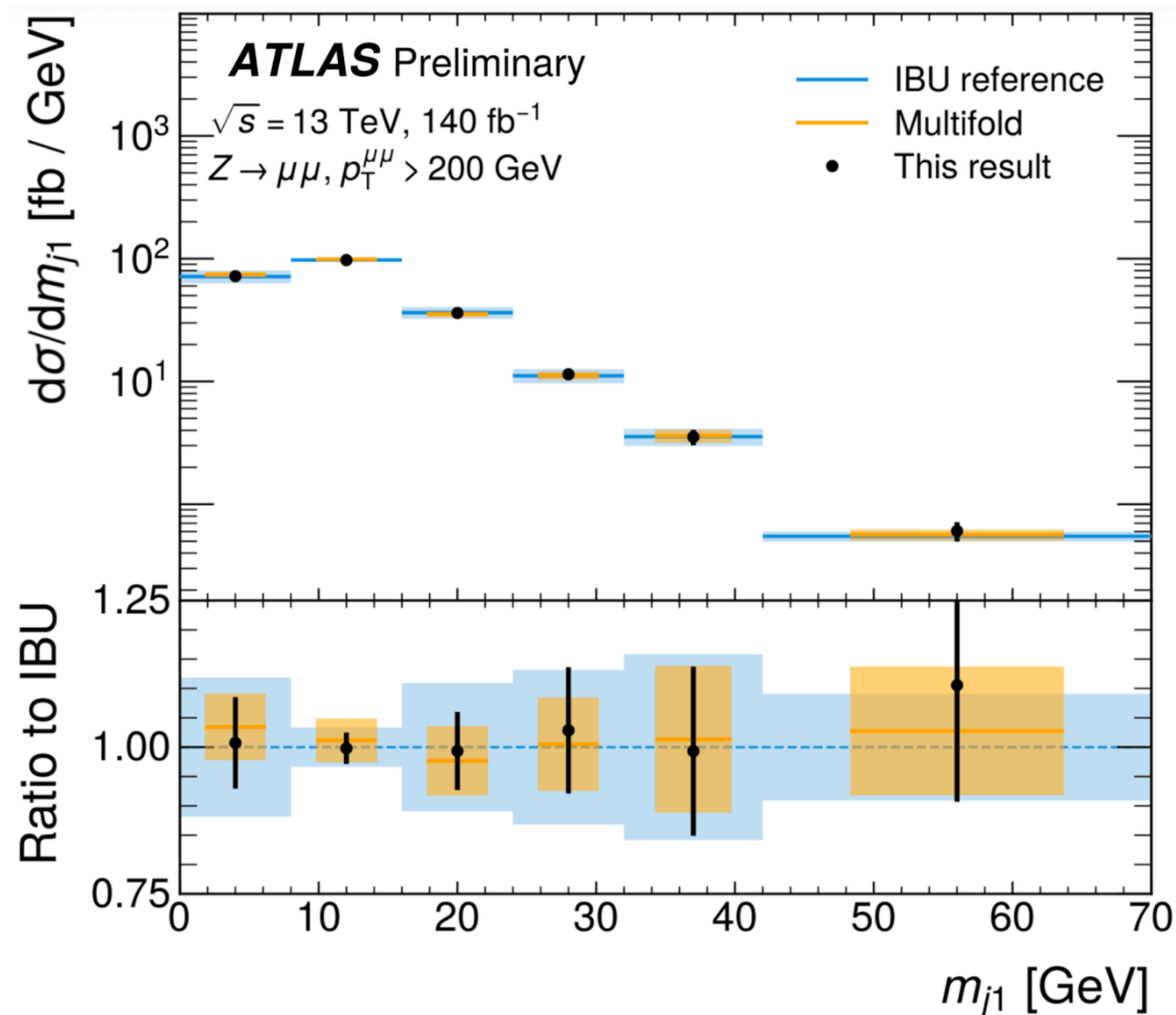
- How to validate measurement of an arbitrary observable?
- Publish auxiliary **pseudodata** measurement with known particle-level target
- Bin measurement and target in the same observable and calculate χ^2

Observable	n_{dof}	χ^2	p
H_{T}	6	6.21	0.400
n_{ch}	5	2.41	0.790
$n_{\text{ch,j1}}$	6	10.4	0.110
$n_{\text{ch,j2}}$	5	4.72	0.450
$\eta_{\mu 1}$	14	11.1	0.678
$\eta_{\mu 2}$	14	4.34	0.993
m_{j1}	6	15.6	0.016
m_{j2}	4	9.23	0.056
$p_{\text{T},\mu 1}$	6	4.57	0.600
$p_{\text{T},\mu 2}$	6	8.20	0.224
$p_{\text{T},\mu\mu}$	5	4.43	0.490
$p_{\text{T,j1}}$	6	11.3	0.081
$p_{\text{T,j2}}$	4	2.12	0.713

Observable	n_{dof}	χ^2	p
$\phi_{\mu 1}$	16	13.5	0.634
$\phi_{\mu 2}$	16	17.3	0.364
ϕ_{j1}	16	7.02	0.973
ϕ_{j2}	16	7.77	0.955
$\tau_{1,\text{j1}}$	7	3.32	0.854
$\tau_{1,\text{j2}}$	5	6.97	0.223
$\tau_{2,\text{j1}}$	7	11.9	0.105
$\tau_{2,\text{j2}}$	5	5.40	0.369
$\tau_{3,\text{j1}}$	4	1.87	0.760
$\tau_{3,\text{j2}}$	4	1.92	0.750
$y_{\mu\mu}$	14	15.7	0.334
y_{j1}	18	14.0	0.727
y_{j2}	10	5.69	0.840

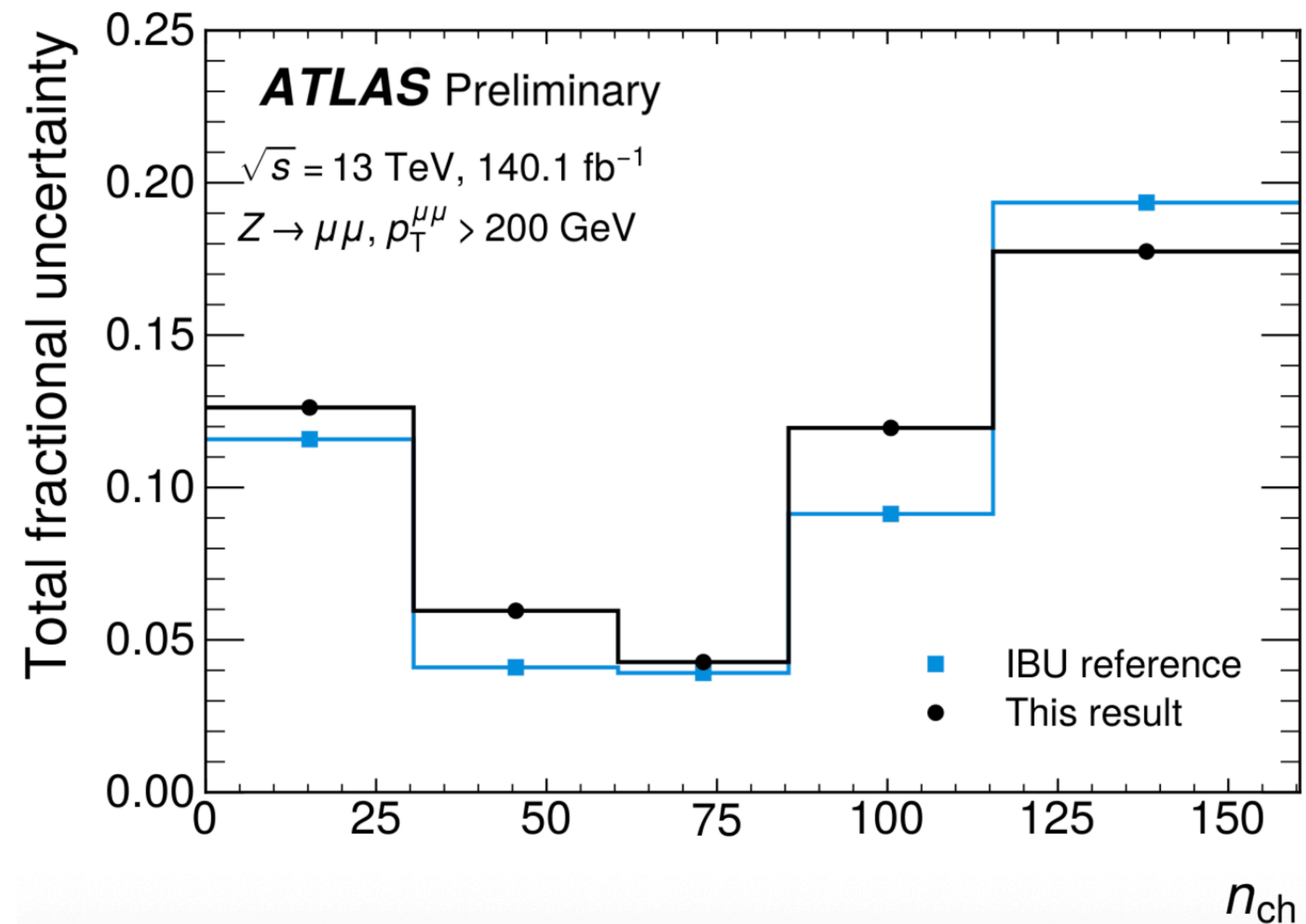
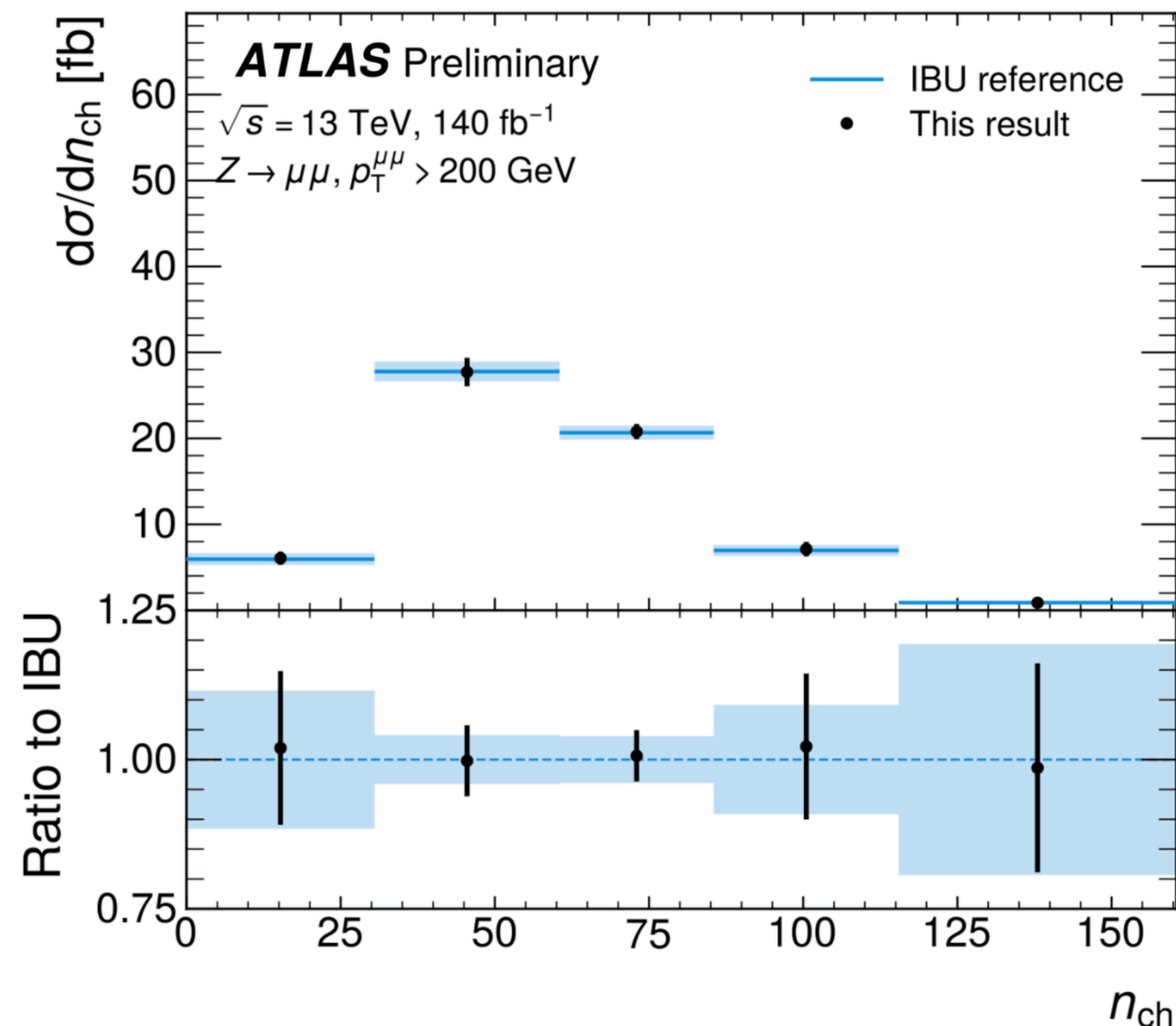
Comparison to baseline measurements

- Compare to auxiliary measurements using traditional binned methods
- Omnifold can essentially reproduce accuracy of binned measurements!
 - Unfolding uncertainties (e.g. hidden variable) can be larger

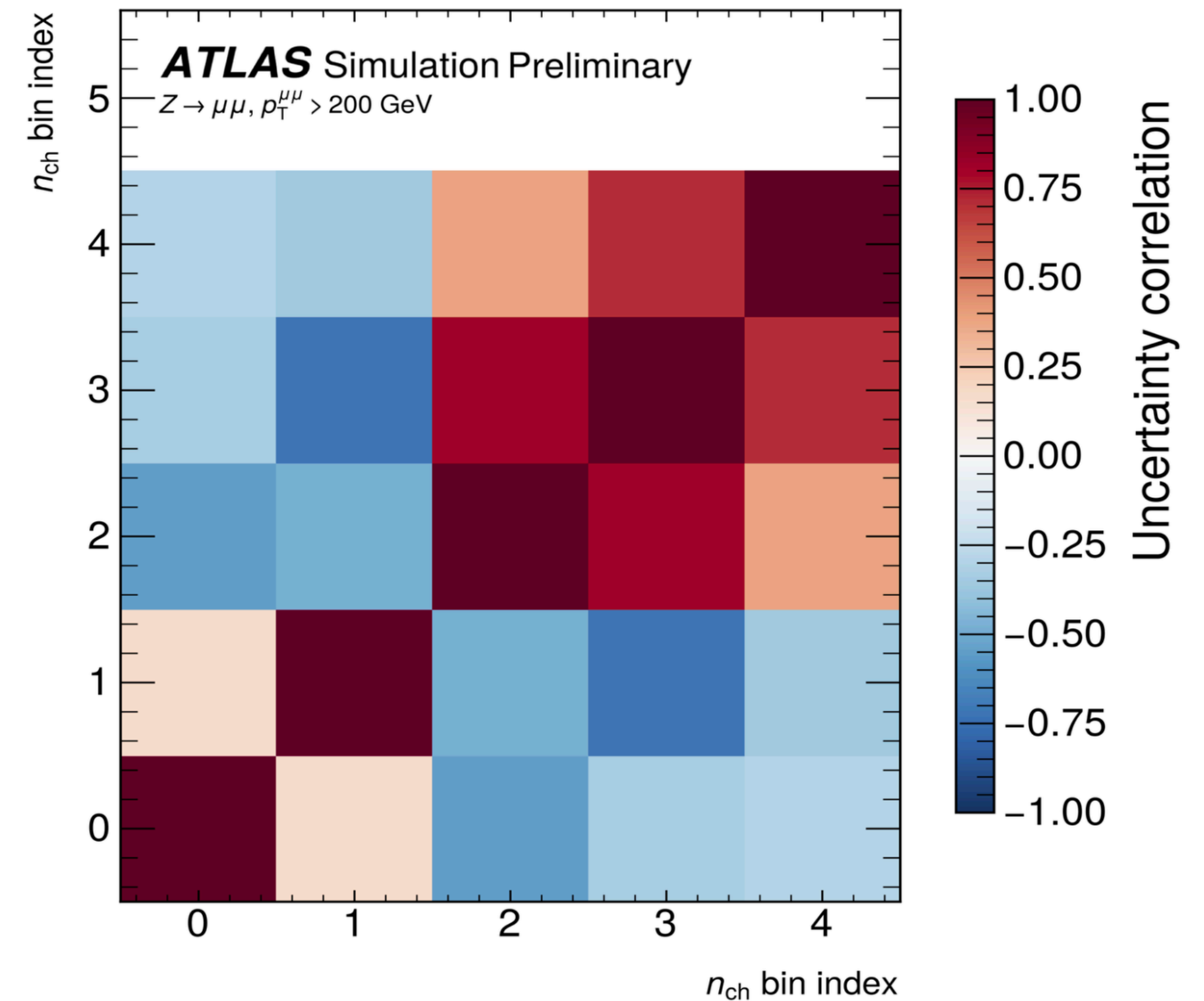
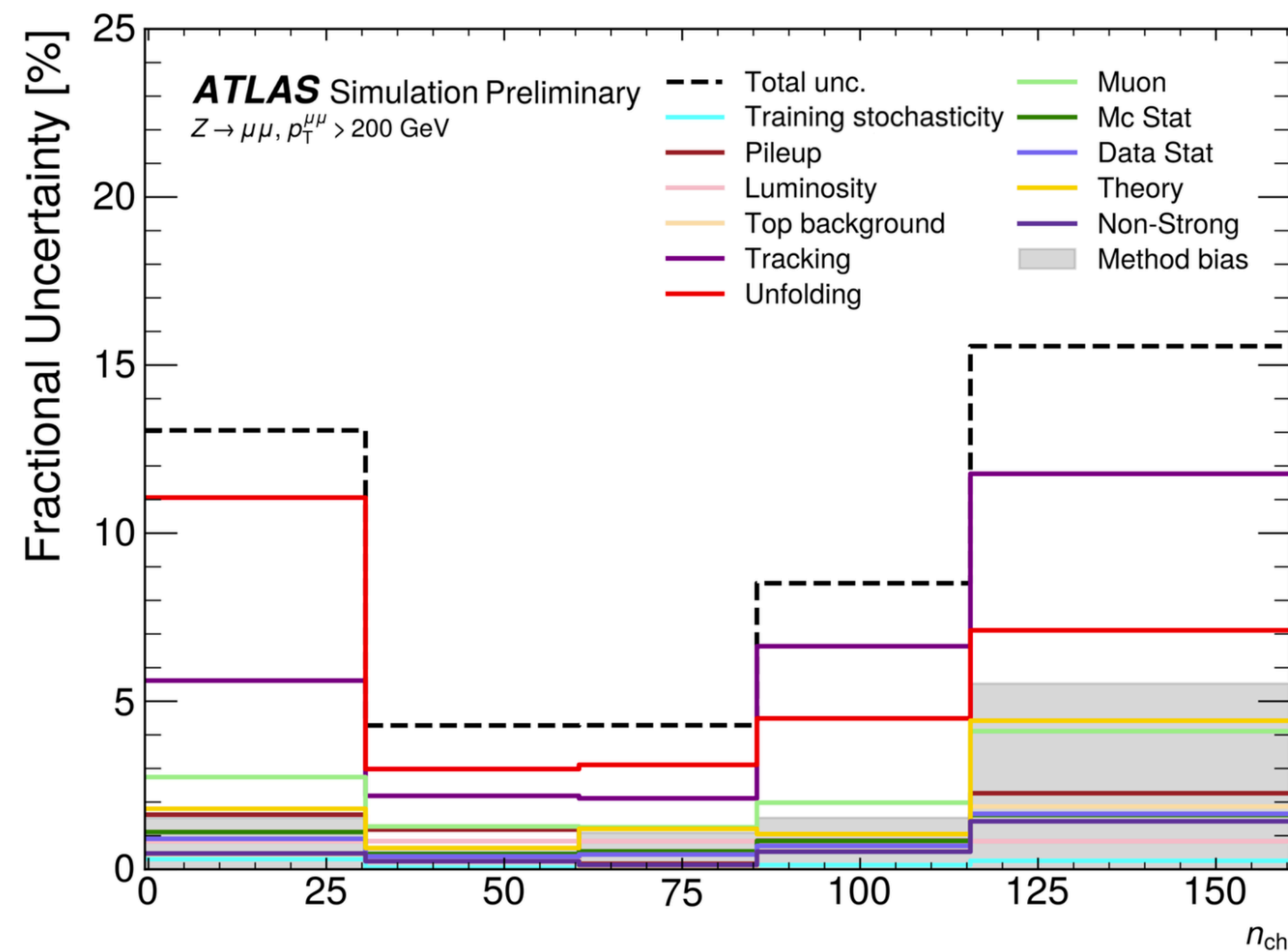


Comparison to baseline measurements

- Compare to auxiliary measurements using traditional binned methods
- Omnifold can essentially reproduce accuracy of binned measurements!
 - Unfolding uncertainties (e.g. hidden variable) can be larger

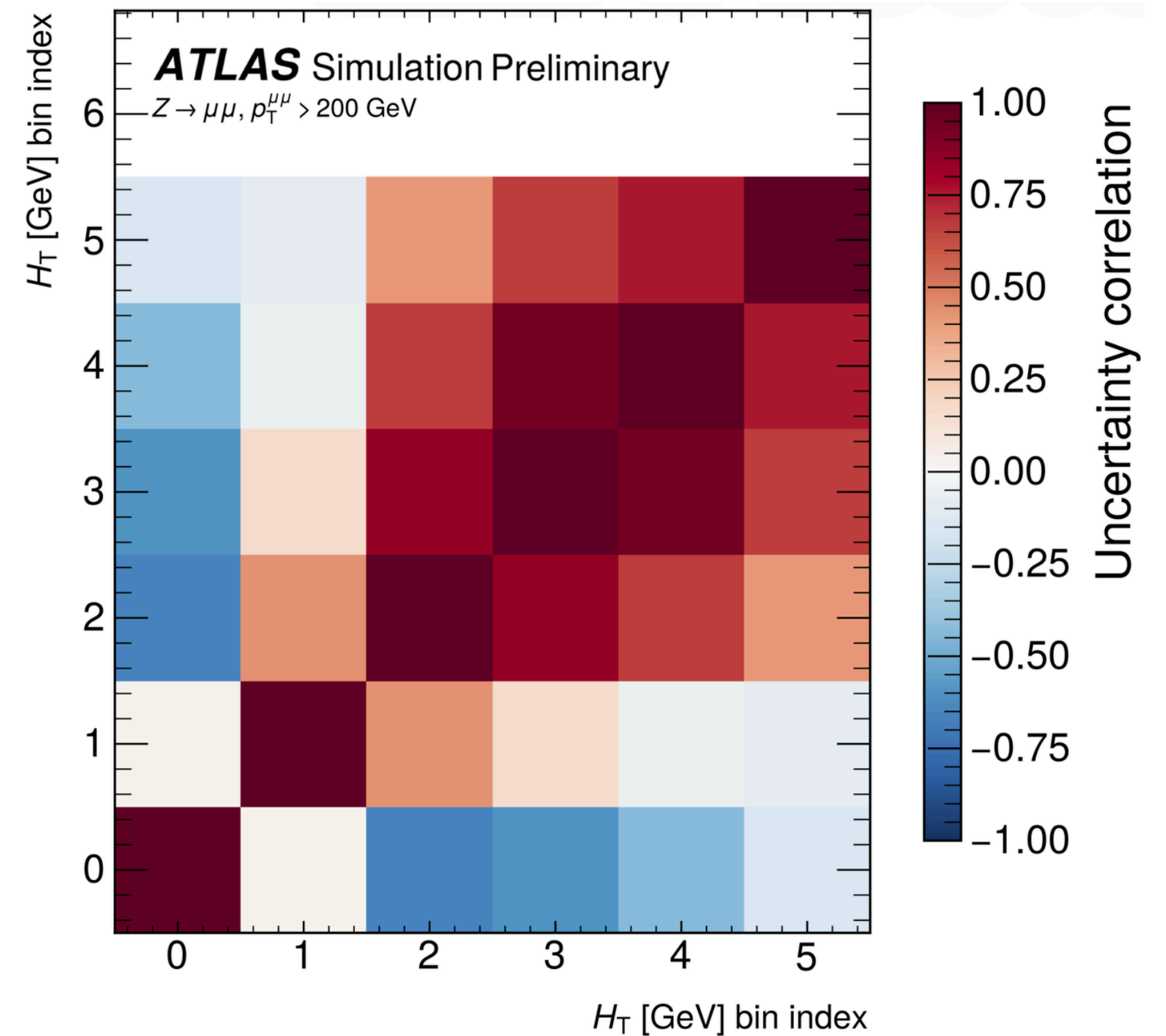
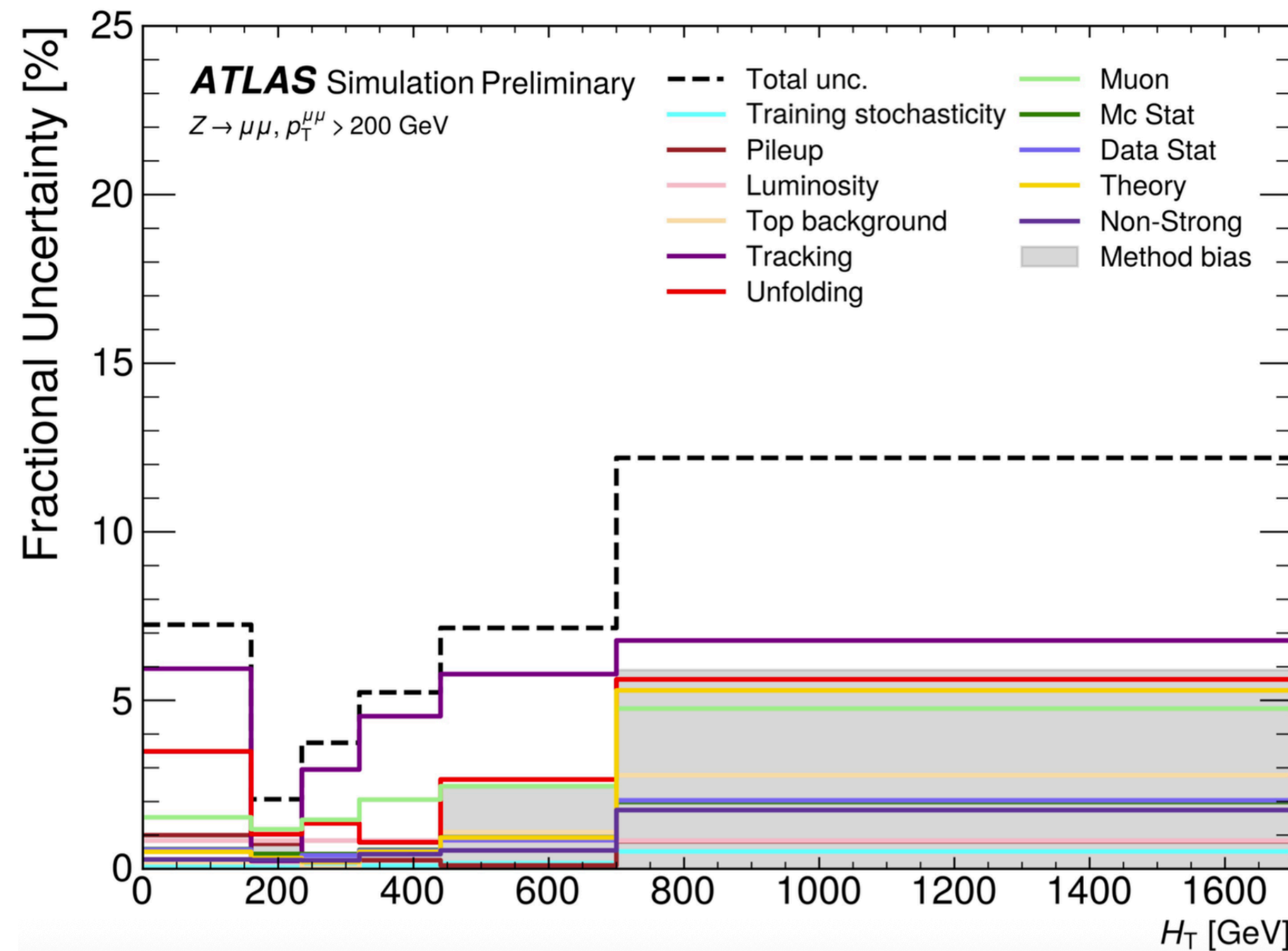


Uncertainties and covariance matrices



See [public codebase and user manual](#) for many more examples

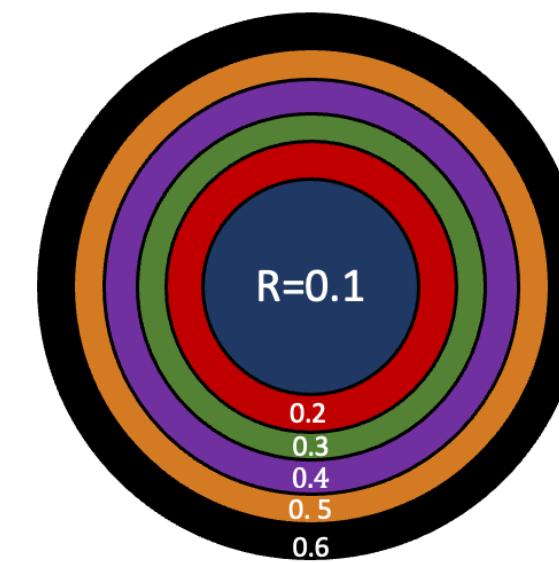
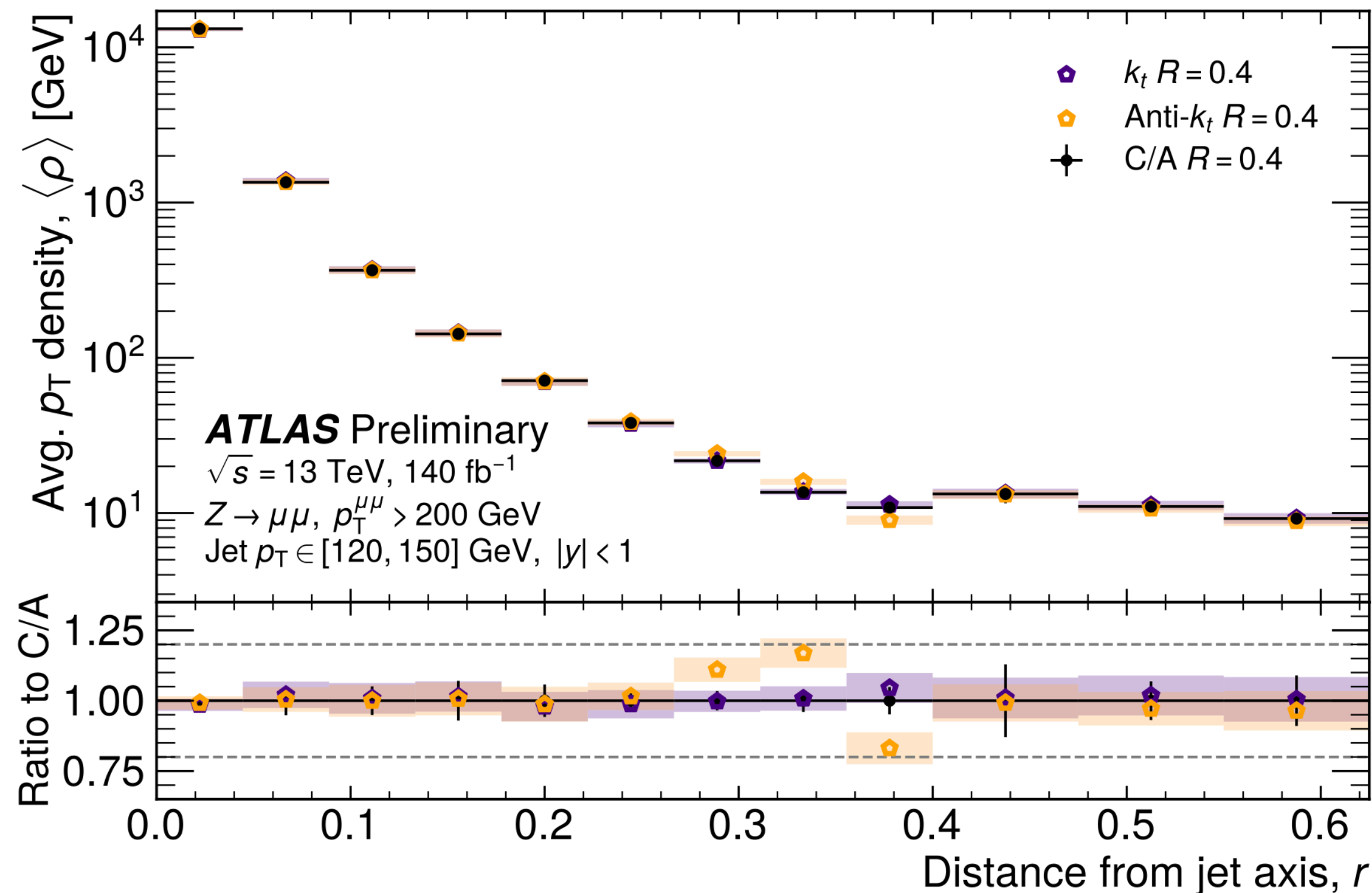
Uncertainties and covariance matrices



See [public codebase and user manual](#) for many more examples

Jet p_T profile measurement

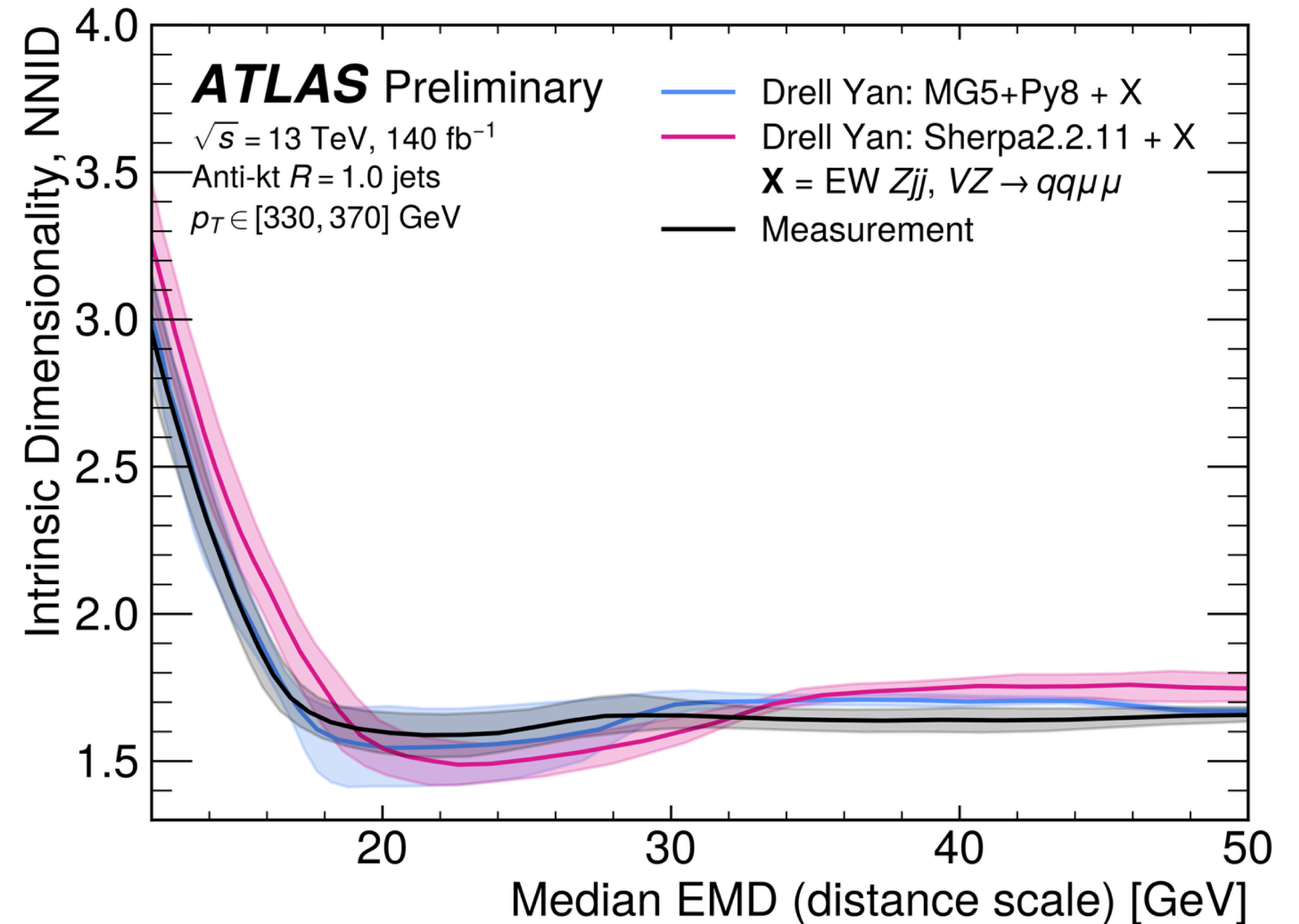
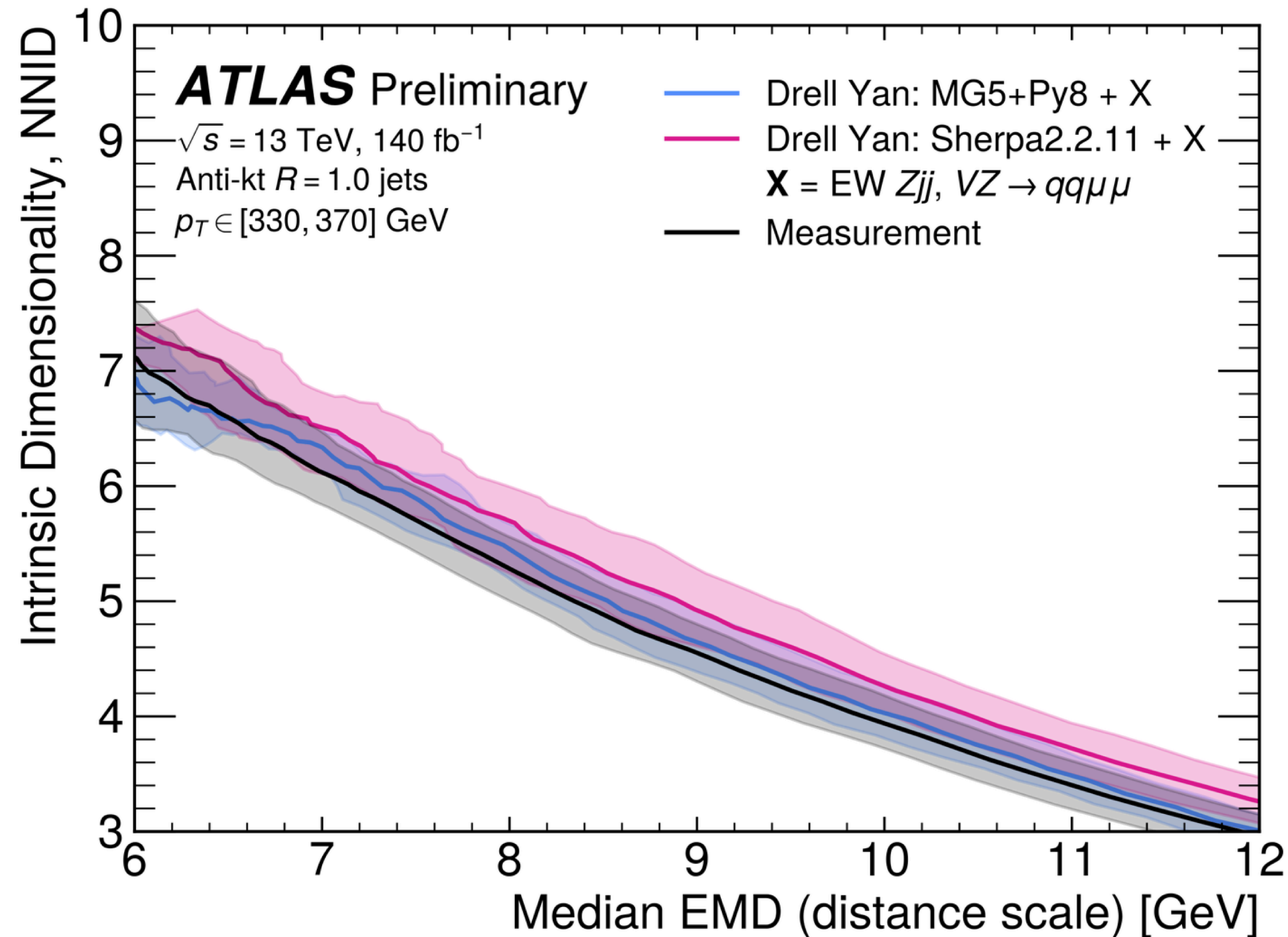
- Can **cluster jets** with unbinned event sample
- Possible to quickly measure an arbitrary jet shape observable, and compare across clustering algorithms



$$\langle \rho(r) \rangle = \frac{1}{N_{\text{jets}}} \sum_{\text{jets}} \frac{p_T(r - \Delta r/2, r + \Delta r/2)}{A_{\text{annulus}}}$$

Anti- k_t jet clustering is greedy
 Jet centered on high p_T tracks
 Boundary depleted

NNID Measurement



First measurement of the intrinsic dimensionality of collider data!
These plots are **impossible** to produce with binned unfolding methods