NATIONAL SCIENCE FOUNDATION

U.S. DEPARTMENT OF ENERGY | Office of Science

VANDERBILT UNIVERSITY

Systematic uncertainties in high energy nuclear physics measurements

Raghav Kunnawalkam Elayavalli (they/them)
Vanderbilt University
raghavke.me

July 8th, 2024

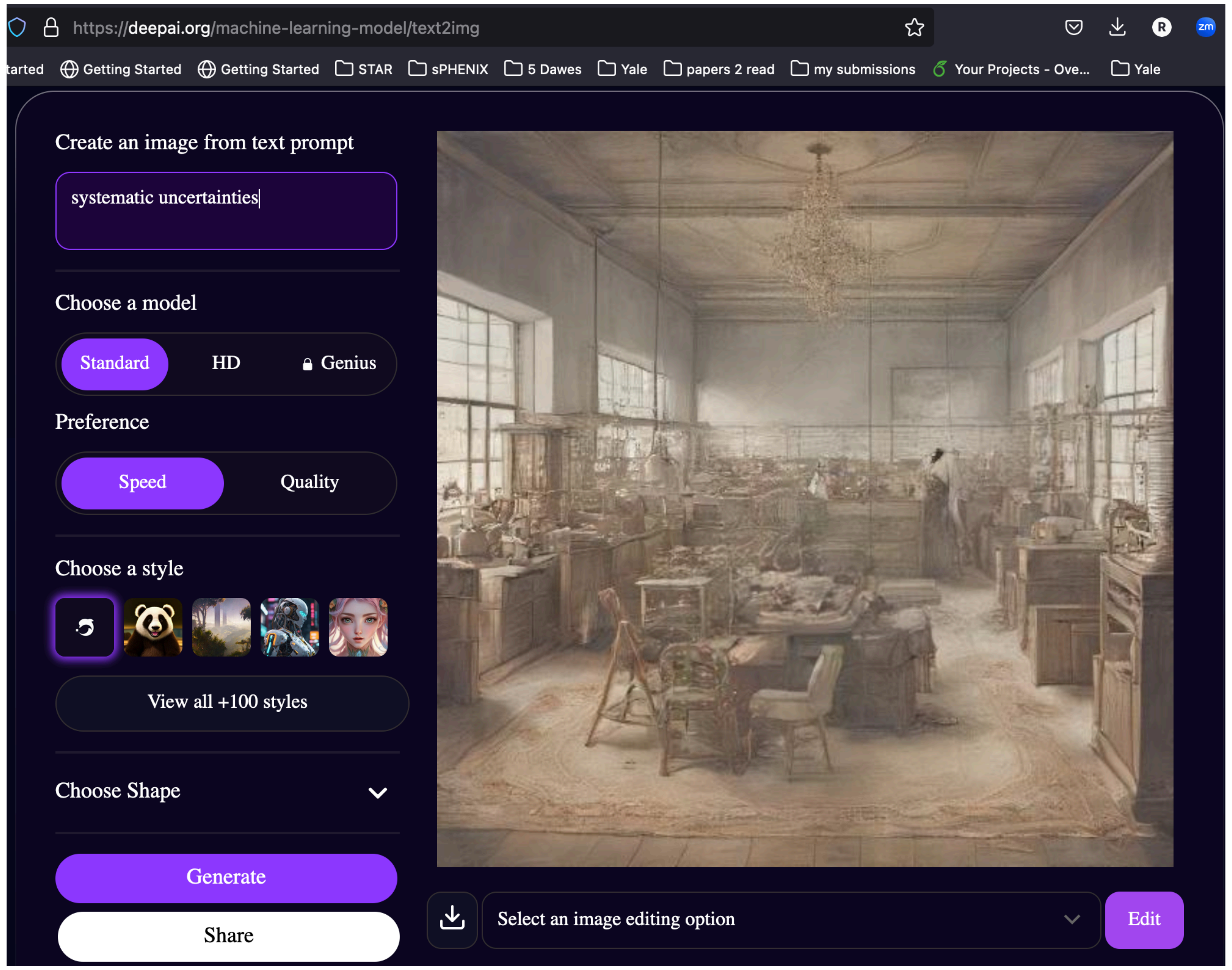
INT WORKSHOP INT-24-88W

Inverse Problems and Uncertainty Quantification in Nuclear Physics

July 8, 2024 - July 12, 2024

Food for thought

- Why should we talk about systematic errors now?
- Systematic errors/uncertainties for measurements in
  (pp and) heavy ion collisions
- How has our uncertainties evolved over 2 decades?
- Detector vs observable specific sources of errors
- Where do we go now? Corrections in the future!

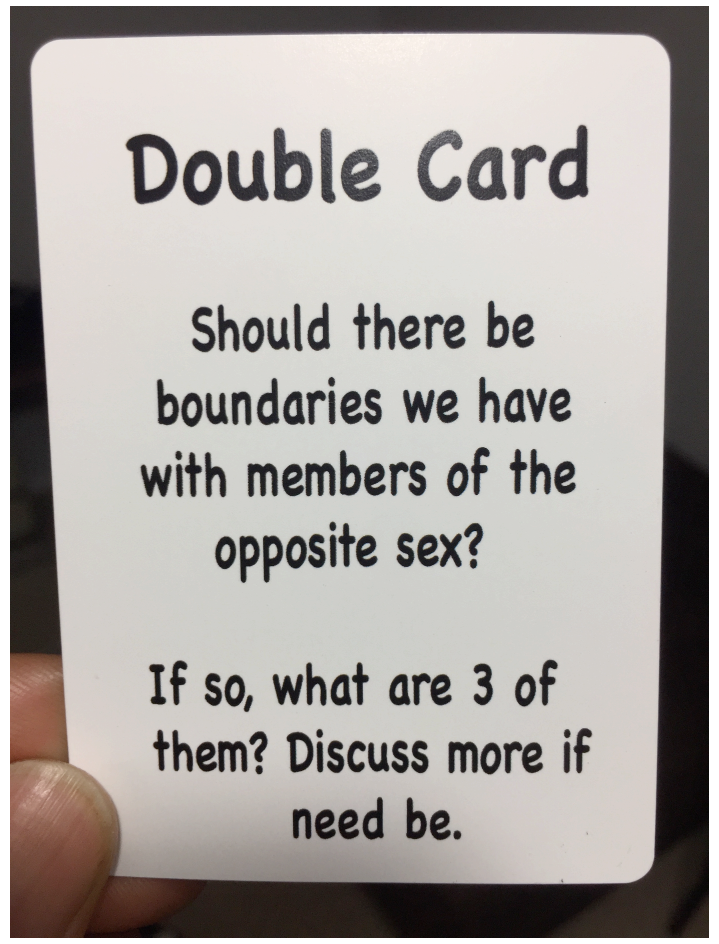


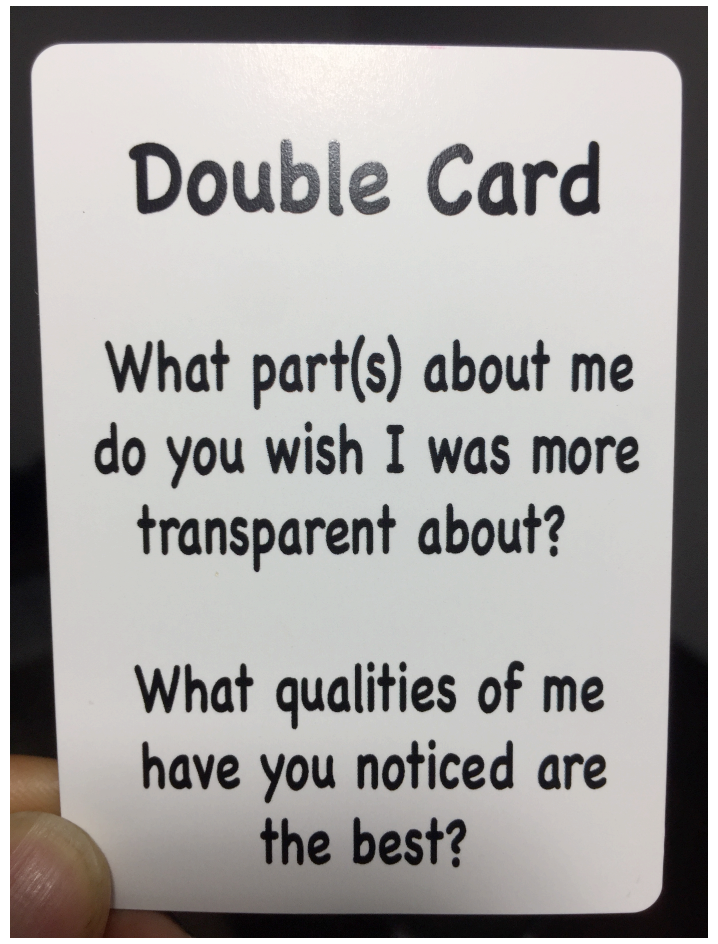
Thanks to - Roger Barlow (U Huddersfield), Mateus F. Carneiro (BNL),
Exp collaborations at LHC and RHIC, Dennis Perepelitsa (U Boulder), Yi Luna Chen (Vanderbilt)

https://www.thegamecrafter.com/games/same-page

Let's make sure we're on the

Same Page





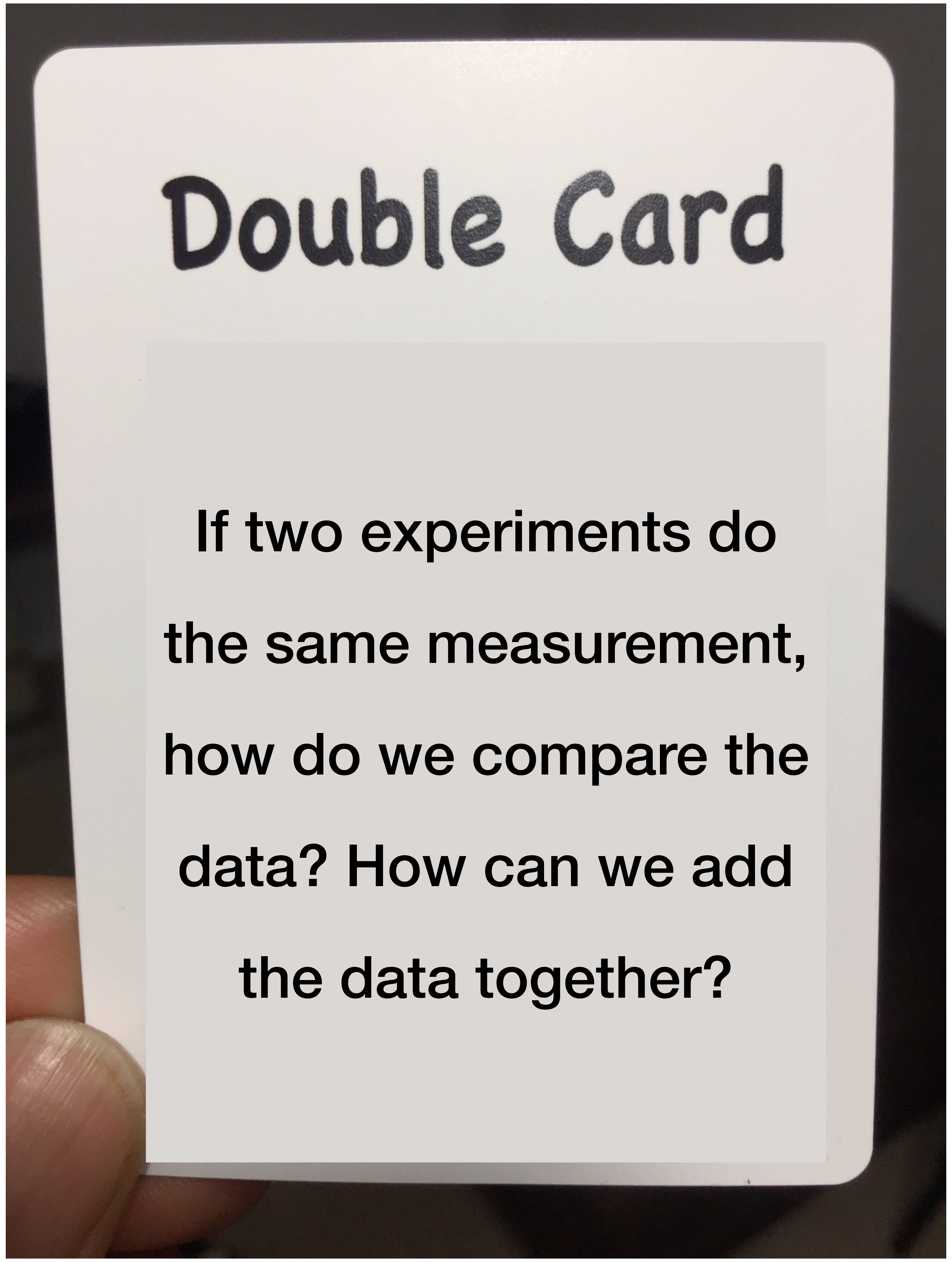
"Dating" Edition

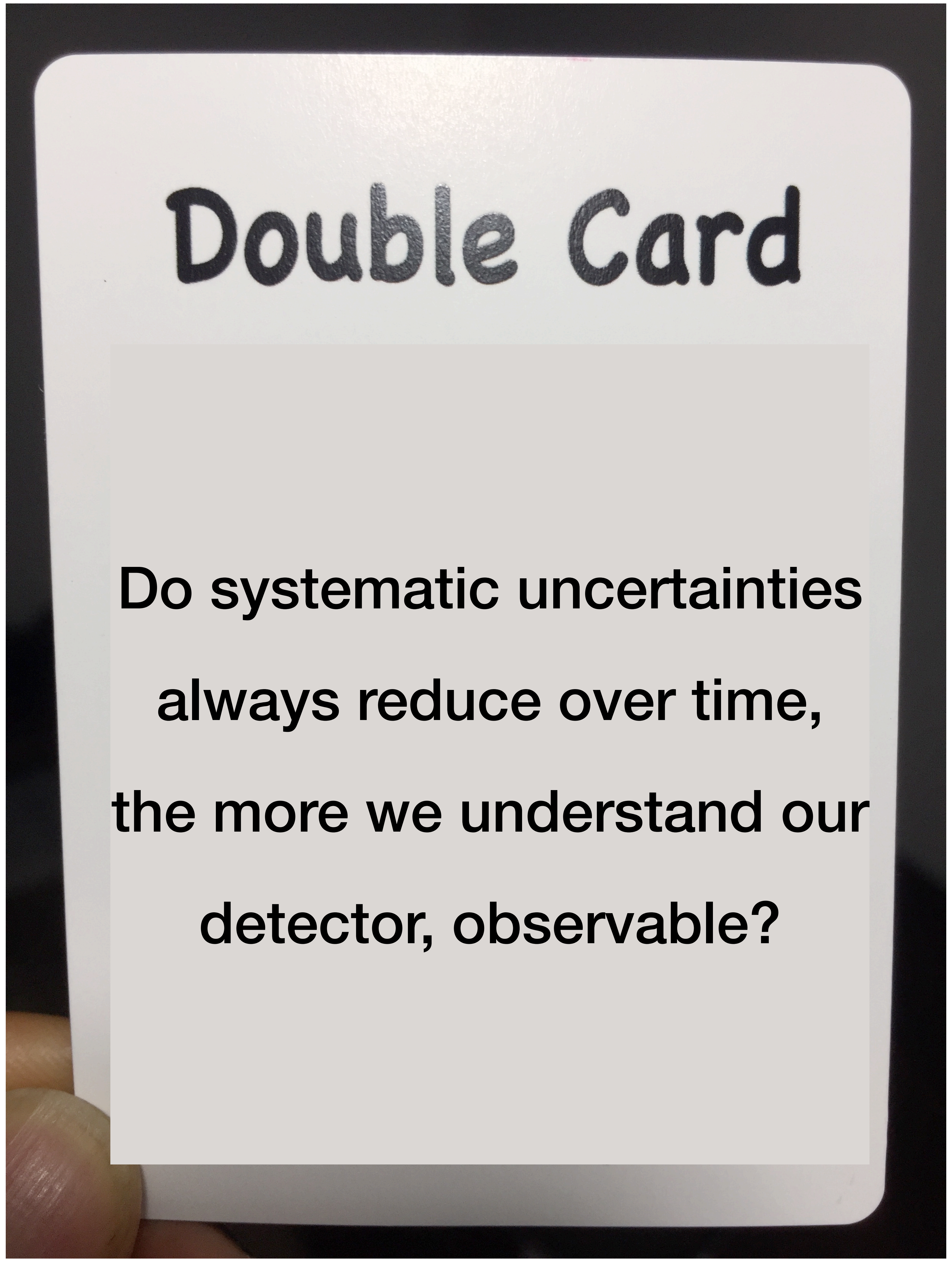
Buy for $23.99

Let's make sure we're on the Same Page

"Systematic Uncertainty" Edition

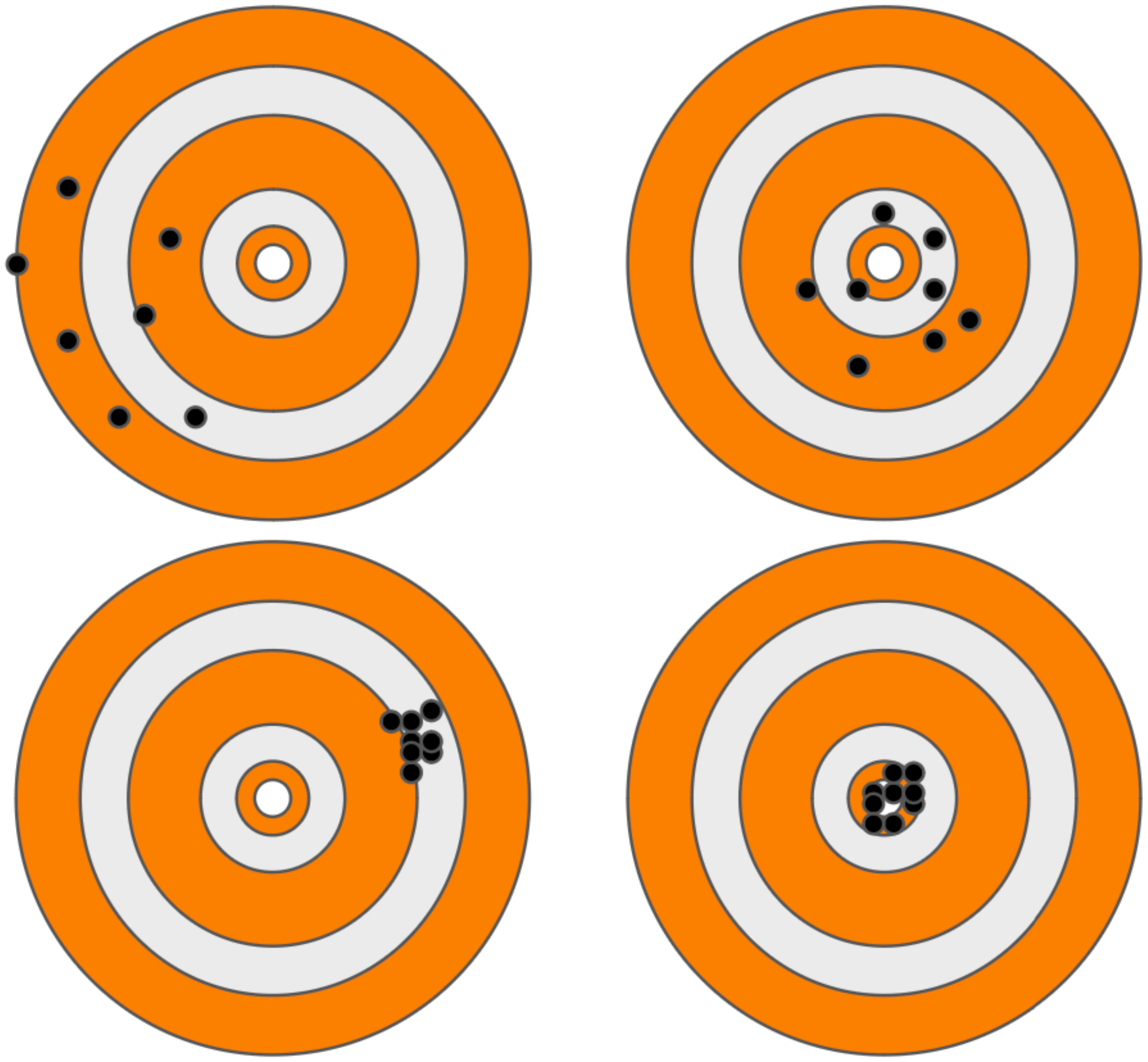
Buy for $0.00


Double Card
If two experiments do the same measurement, how do we compare the data? How can we add the data together?

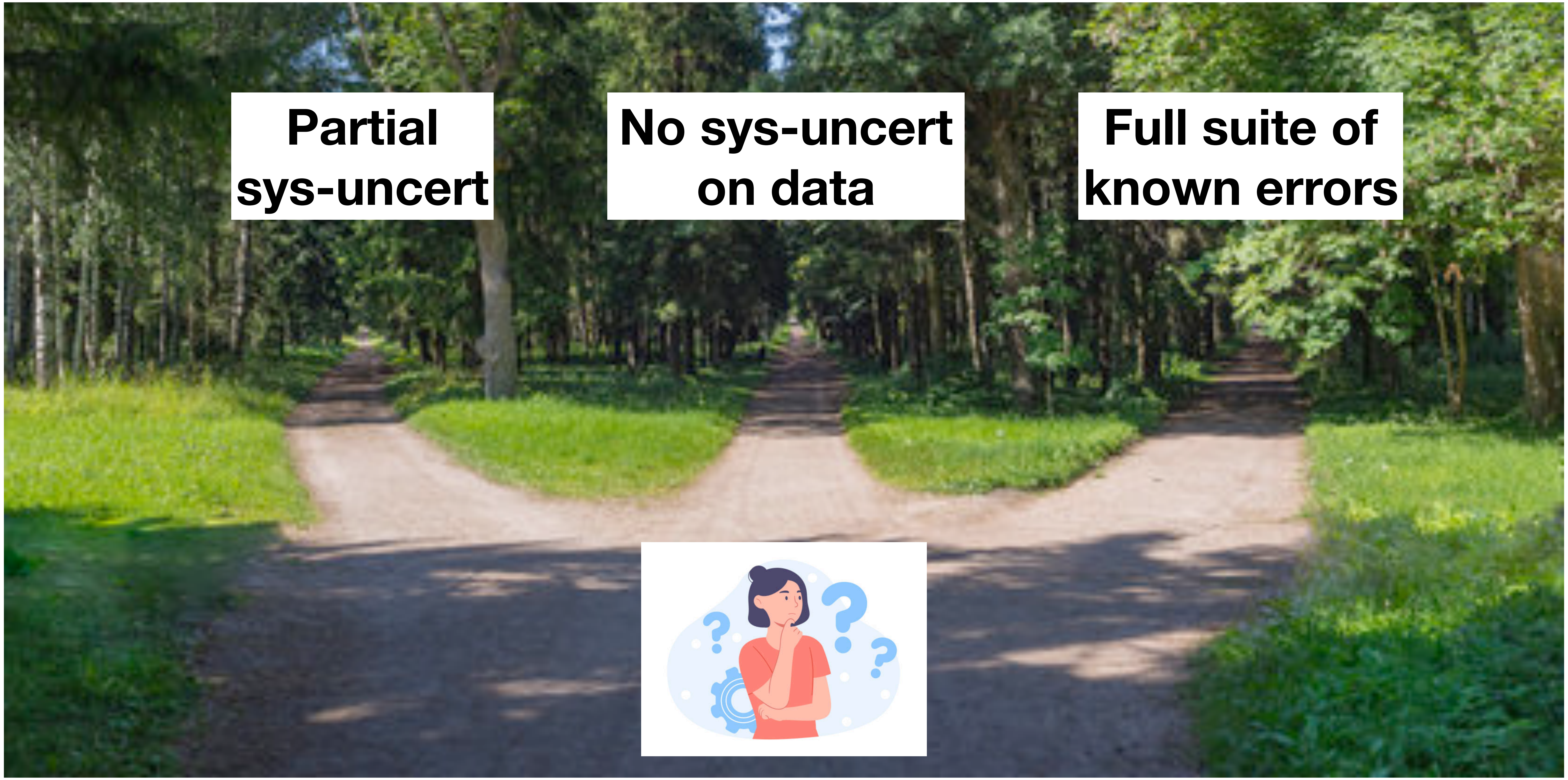

Double Card
Do systematic uncertainties always reduce over time, the more we understand our detector, observable?

Systematic errors

- Circles are our detector - hits are something we measure/observe
- These are four separate hit patterns from the same set of observables
- Interplay evident between precision and accuracy!
- What changed between the 4 of them? Are any random noise? Anything we can correct?



Three ways to proceed here



Circa 2001 AD (first from RHIC)

Elliptic Flow in $Au + Au$ Collisions at $\sqrt{s_{NN}} = 130\,\mathrm{GeV}$

K. H. Ackermann *et al.* (STAR Collaboration)

Phys. Rev. Lett. **86**, 402 – Published 15 January 2001

Fig. 3 shows v_2 as a function of centrality of the collision. Although this figure was made with the subevents chosen as in Fig. 2, the same results within errors were obtained with the other correlation methods. Restricting the primary vertex z position to reduce TPC acceptance edge effects also made no difference. From the results of the study of non-flow contributions by different subevent selections and the maximum magnitudes of the first and higher-order harmonics, we estimate a systematic error for v_2 of about 0.005, with somewhat smaller uncertainty for the mid-centralities where the resolution of the event plane is high. The systematic errors are not included in the figures.

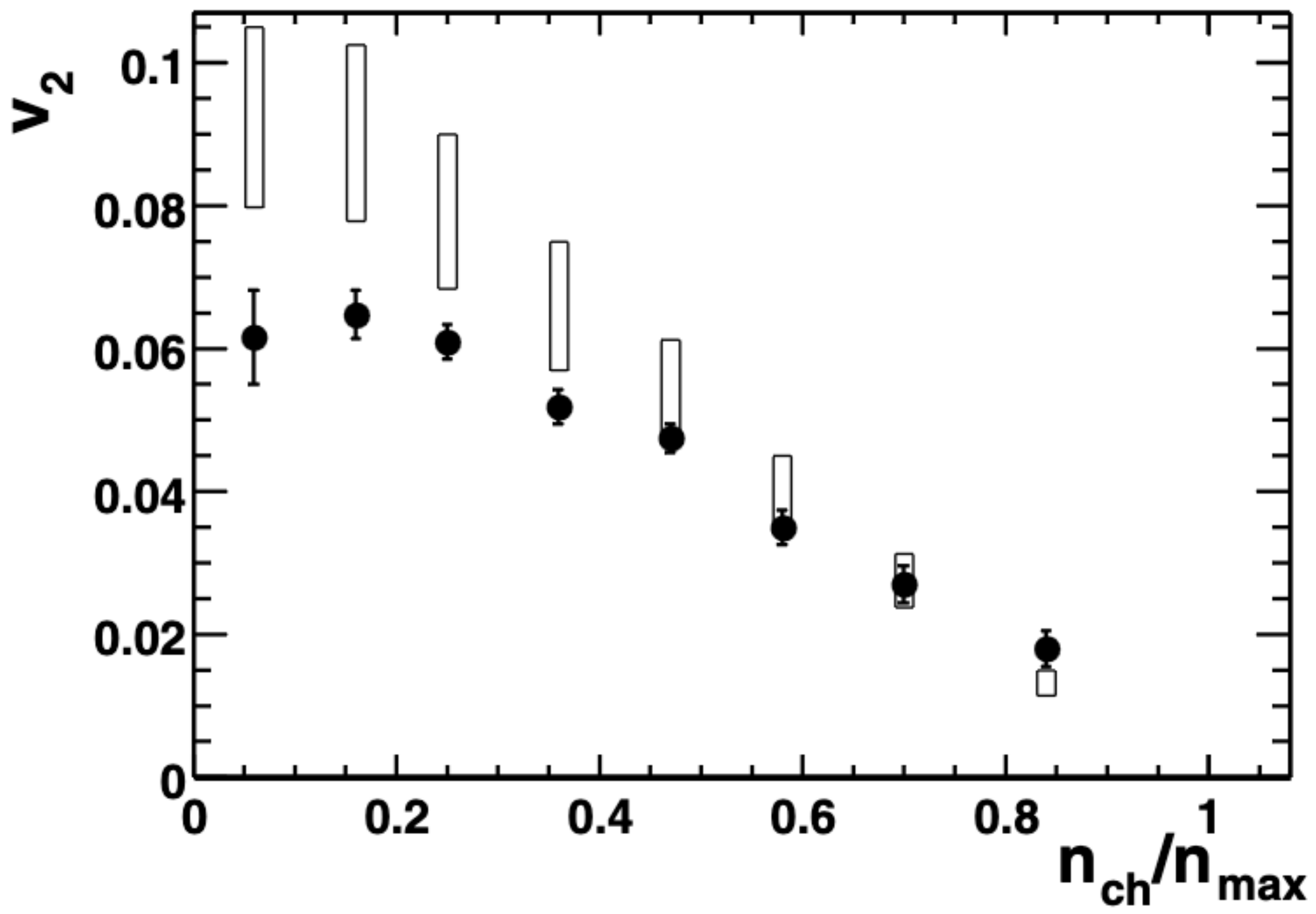


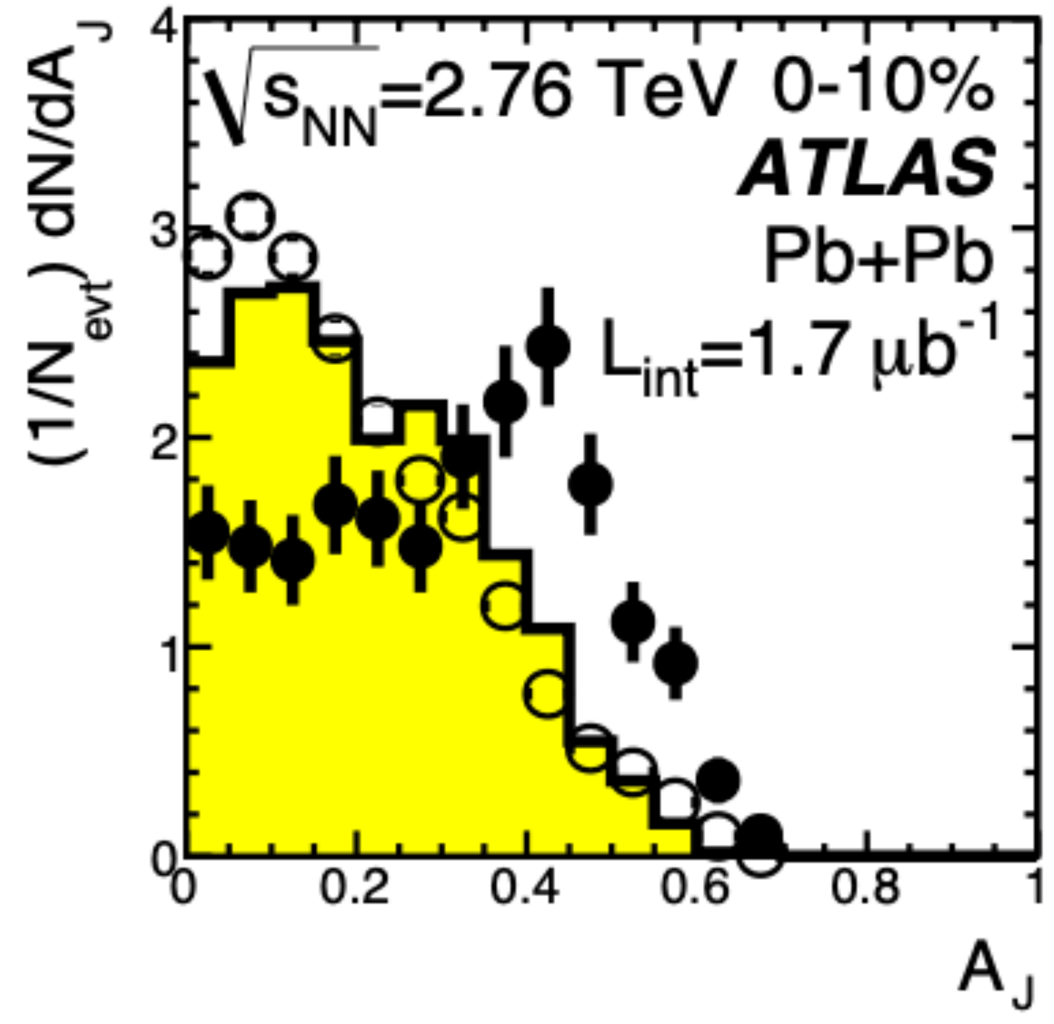
FIG. 3. Elliptic flow (solid points) as a function of centrality defined as n_{ch}/n_{max}. The open rectangles show a range of values expected for v_2 in the hydrodynamic limit, scaled from ϵ, the initial space eccentricity of the overlap region. The lower edges correspond to ϵ multiplied by 0.19 and the upper edges to ϵ multiplied by 0.25.

Circa 2011 AD (first from LHC)

Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the ATLAS Detector at the LHC

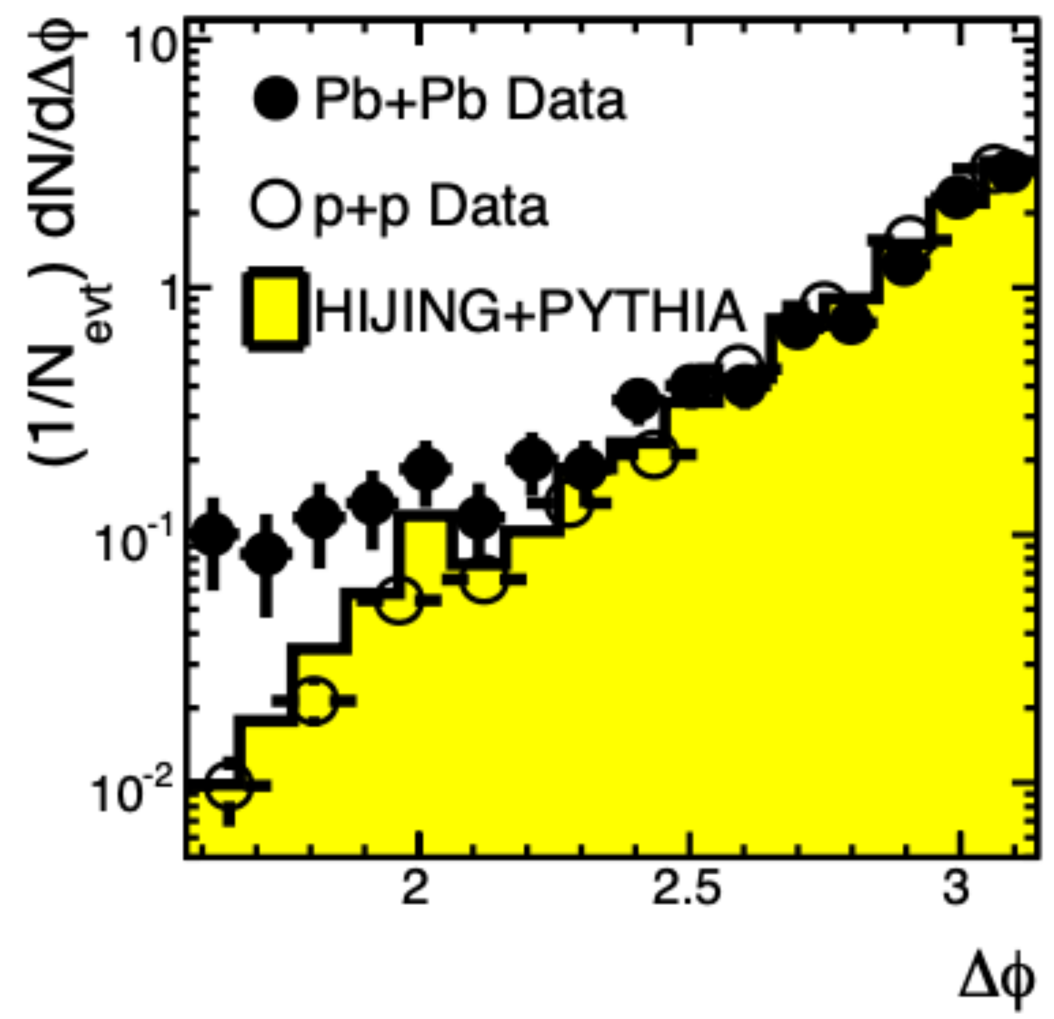
G. Aad *et al.* (ATLAS Collaboration)
Phys. Rev. Lett. **105**, 252303 – Published 13 December 2010

Physics See Viewpoint: A "Little Bang" arrives at the LHC



- Fully raw-data measurement!

- There are *NO* mentions of systematic uncertainties in the entire publication



fine-grained, longitudinally segmented electromagnetic and hadronic calorimeters. The transverse energies of dijets in opposite hemispheres are observed to become systematically more unbalanced with increasing event centrality leading to a large number of events which contain highly asymmetric dijets.

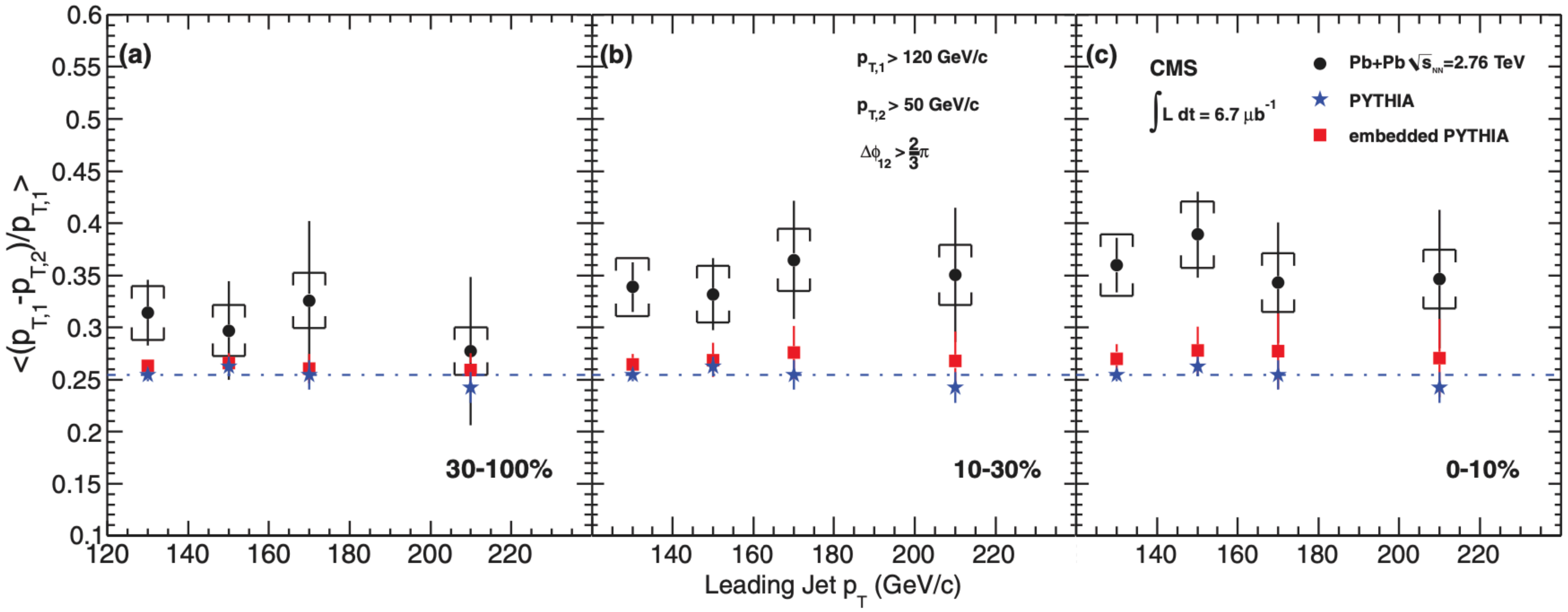
Some amount of systematics…

PHYSICAL REVIEW C **84**, 024906 (2011)

Observation and studies of jet quenching in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV

S. Chatrchyan *et al.**
(CMS Collaboration)
(Received 10 February 2011; published 12 August 2011)

tively. The dominant contribution to the systematic uncertainty comes from the observed p_T dependence of the residual jet energy correction in PbPb events (6% out of a total systematic uncertainty of 8%). The jet energy resolution and underlying event subtraction uncertainties contribute ∼4% each.



- Has some uncertainties, data is somewhat corrected, but still it is not directly comparable to theory

Another (recent) example of the mid-way

PHYSICAL REVIEW C **105**, 044906 (2022)

Editors' Suggestion

STAR Collaboration

Differential measurements of jet substructure and partonic energy loss in Au + Au collisions at $\sqrt{s_{NN}} = 200$ GeV

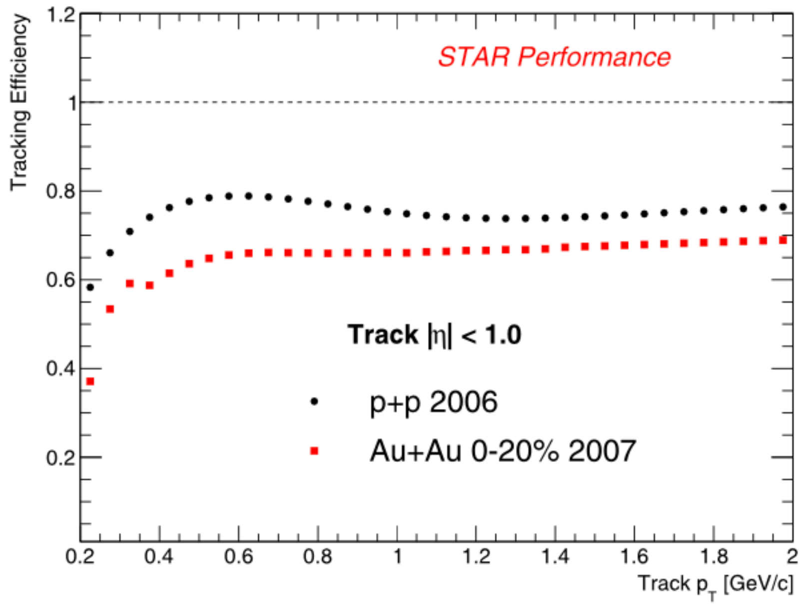


FIG. 11. TPC tracking efficiencies for the 2006 $p + p$ and 2007 Au + Au datasets utilized in the embedding studies for tracks within $|\eta| < 1.0$.

APPENDIX: DETECTOR EFFECTS AND COMPARISONS

The Au + Au data in this publication are compared to an embedded reference at the detector level which presents the measurement without any correction for detector effects. We therefore provide the relevant performance parameters for the STAR detector, mainly the TPC and the BEMC. This enables predictions of MC models or theoretical calculations to be directly applied to the detector-level data. For charged-particle tracks in the TPC, the tracking efficiency is shown in Fig. 11 as a function of the track p_T for particles at midrapidity ($|\eta| < 1.0$). The red and black markers show the efficiencies for $p + p$ and Au + Au 0–20% events taken during 2006 and 2007, respectively. The tracking efficiency is also assumed to be flat as a function of track momentum for $2.0 < p_T < 30$ GeV/c for both datasets. The TPC also produces a momentum smearing which is modeled by

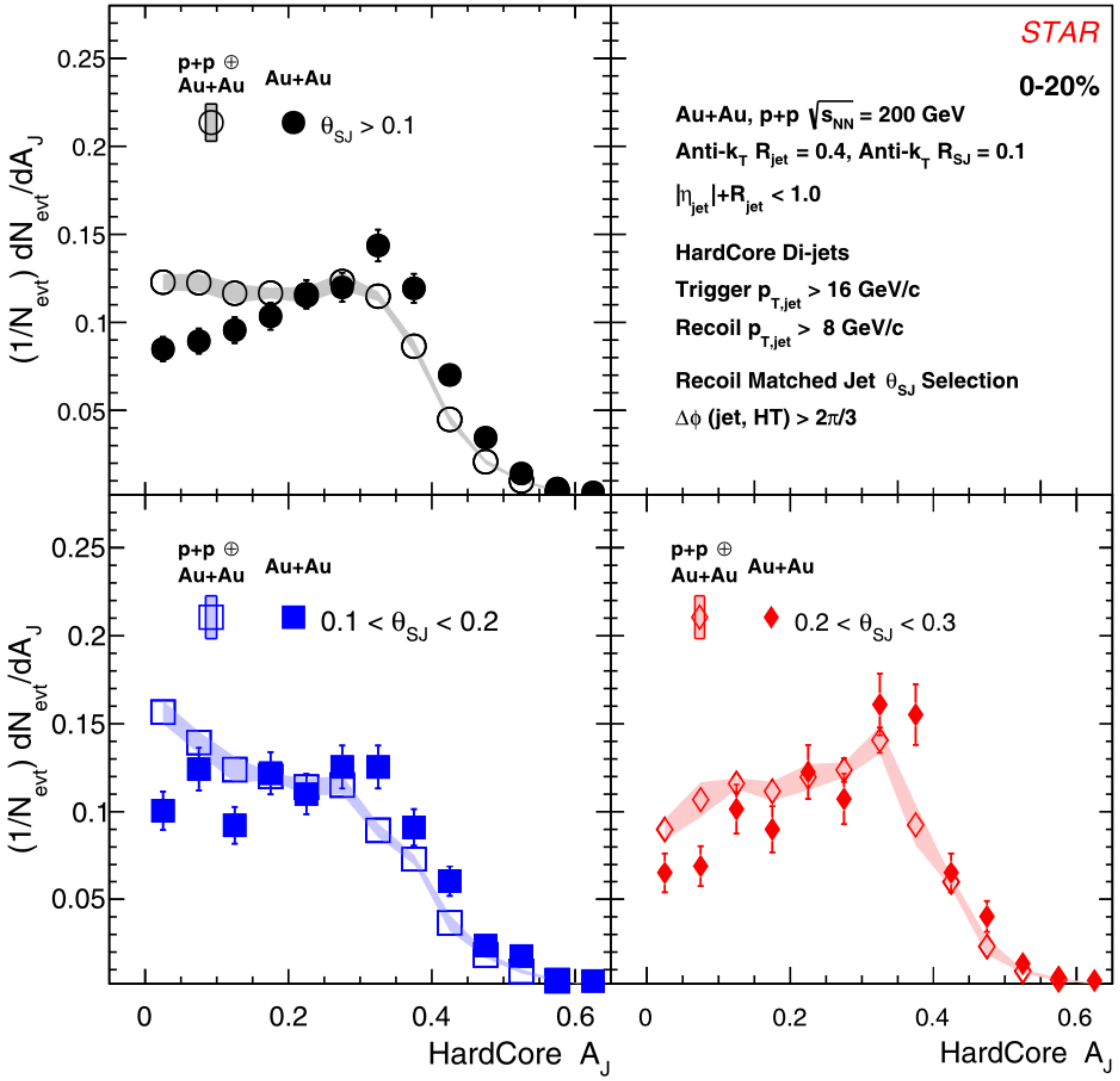
$$\sigma = -0.026 + 0.02 p_T^{\text{true}} + 0.003 (p_T^{\text{true}})^2, \quad \text{(A1)}$$

taken to be the same for both $p + p$ and Au + Au collisions.

The BEMC has a spatial segmentation of 0.05×0.05 in (η, ϕ) with an energy resolution of $\sigma(E_T) = 14\%/\sqrt{E_T}$ [94]. The hadronic correction procedure described at the beginning of Sec. II ensures that the energy deposited by charged particles in the BEMC is not double counted, such that $\sigma(E_T)$ estimates the error in the neutral energy of a jet.

In addition to the preceding detector effects, the impact of the heavy-ion underlying event on the jet momentum and substructure observables should be taken into account for direct comparison with the data presented here. These effects for the HardCore and Matched jet momenta are presented in the supplemental material of an earlier publication [28]. The left panels of Figs. 4 and 7 of this reference show the effect of the heavy-ion underlying event on the substructure observables.

- Data is detector level, reference is corrected to match our AA data with systematic uncertainties - physics extraction is apples to apples
- BUT data is fundamentally not corrected! Theory/MC needs to smear their predictions according to the detector and then one can compare. Why?



State of the art in pp-land (ATLAS)

A simultaneous unbinned differential cross section measurement of twenty-four Z+jets kinematic observables with the ATLAS detector

ATLAS Collaboration • Georges Aad (Marseille, CPPM) Show All(2919)

May 30, 2024

40 pages

e-Print: 2405.20041 [hep-ex]

Report number: CERN-EP-2024-132

Systematic uncertainties are split into 25 components that are each treated as independent. Experimental sources of uncertainty include systematic bias due to: the muon efficiency and calibration [67]; track reconstruction [68]; pileup modeling; and, the luminosity measurement [85]. Theoretical uncertainties are evaluated for variations of PDF and α_s choices [71], QCD scales [71], and the generator tune [86]. An uncertainty ("unfolding prior") for the imperfect particle-level shape of the initial MC sample is assessed by reweighing the nominal MC sample at particle level such that it approximately agrees with data for the 24 observables. This reweighing function is constructed using a sequence of one-dimensional Gaussian-kernel

Four types of stochastic uncertainties are assessed: statistical uncertainties on the data and the MC training sets are each assessed by bootstrapping [87] (100 and 25 weights, respectively); an uncertainty due to the NN stability is calculated from the standard error on the median of weights of the 100 individual NNs; and an additional uncertainty is assigned for the limited statistics of the nominal event dataset. Overall, the total uncertainty in most bins chosen to illustrate the final measurement is between 3% and 5%, but can grow as large as 15% in tails of distributions. The unfolding uncertainty from the unfolding prior and hidden variables tends to be the dominant contributor for many observables, in particular for the eight jet substructure variables (see Figure 1(d)).

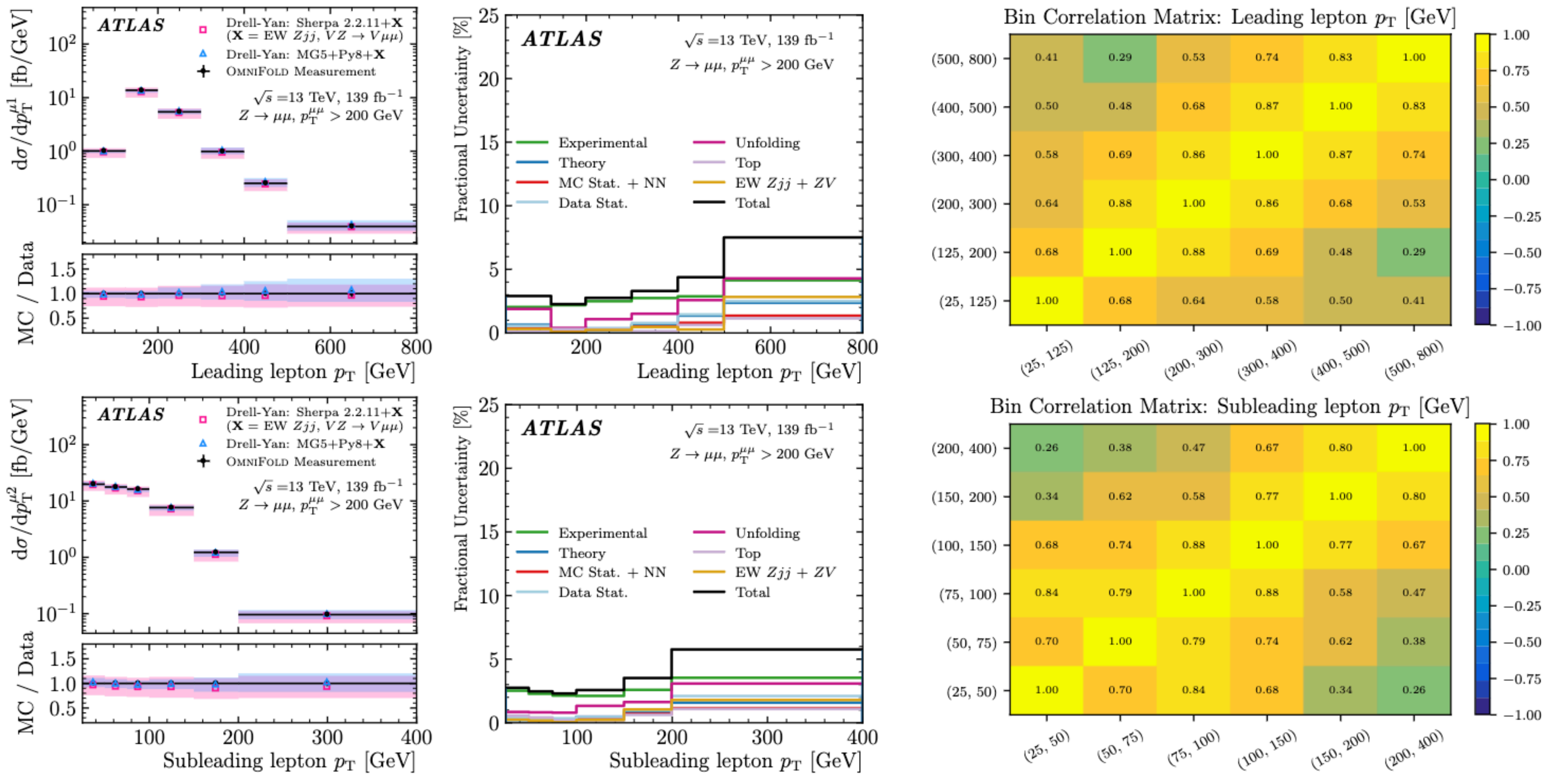


Figure 5: Measured differential cross sections along with predictions from MadGraph and Sherpa 2.2.11 (left), the total measurement uncertainty and its breakdown into sources (middle), and the associated correlation matrix (right) for $p_T^{\mu1}$ (top) and $p_T^{\mu2}$ (bottom).

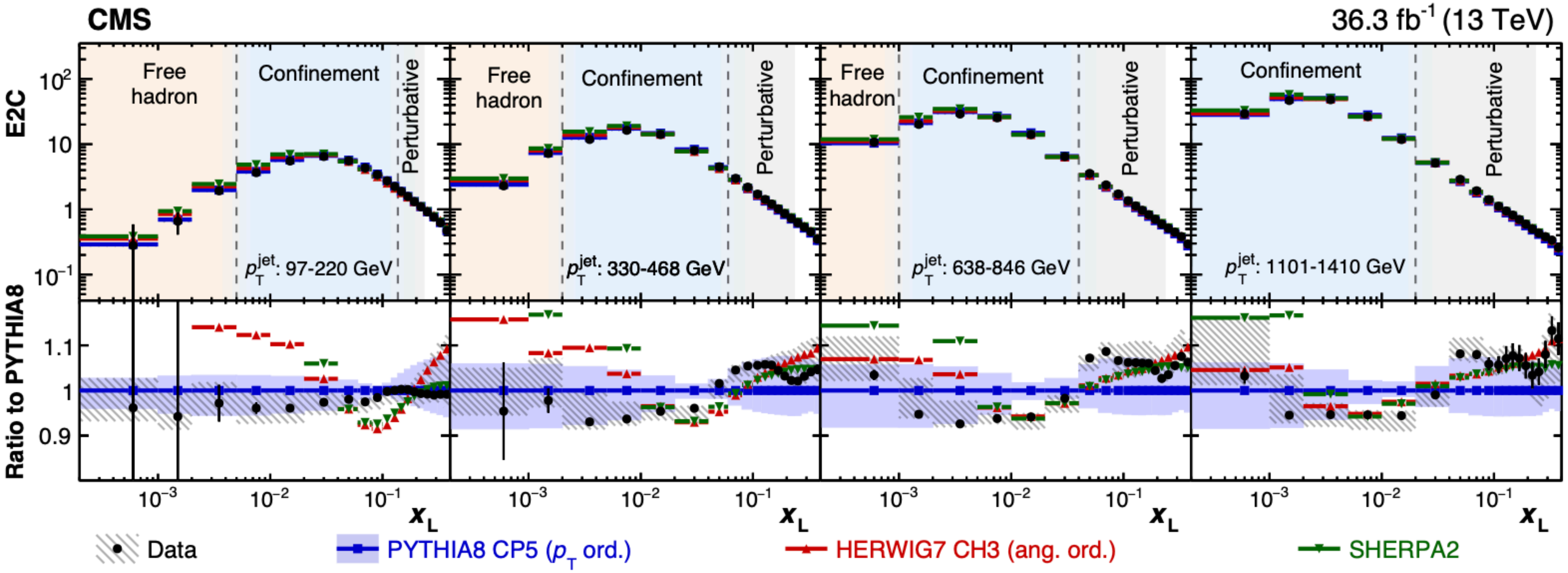
State of the art in pp-land (CMS)

Measurement of energy correlators inside jets and determination of the strong coupling $\alpha_S(m_Z)$

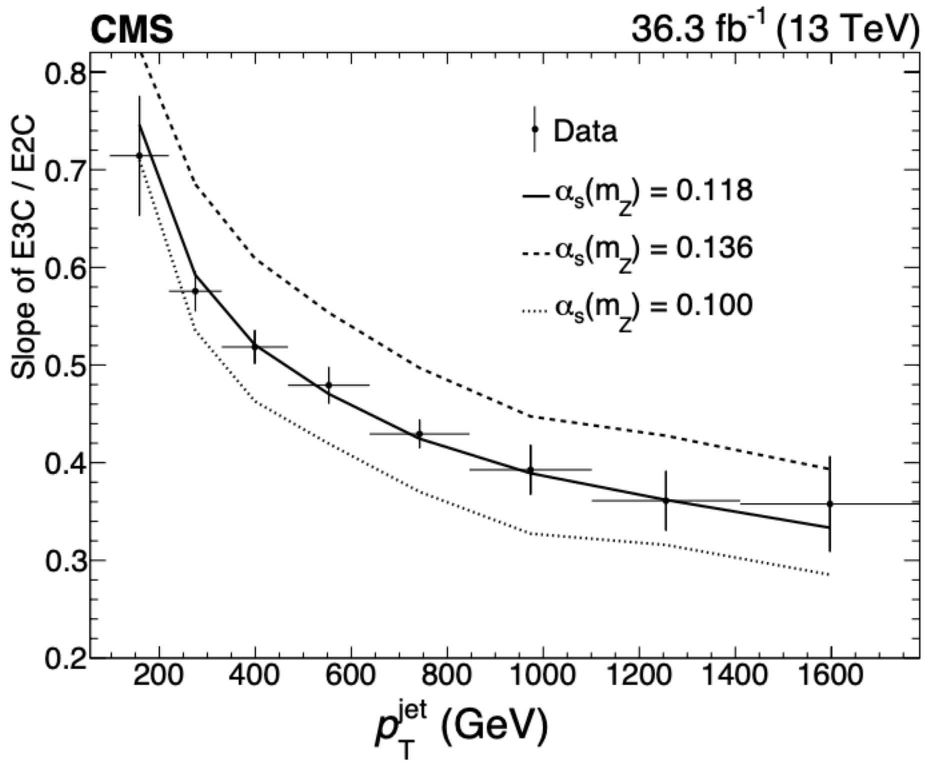
https://arxiv.org/abs/2402.13864

The CMS Collaboration*

interaction: confinement and asymptotic freedom. By comparing the ratio of the two measured distributions with theoretical calculations that resum collinear emissions at approximate next-to-next-to-leading logarithmic accuracy matched to a next-to-leading order calculation, the strong coupling is determined at the Z boson mass: $\alpha_S(m_Z) = 0.1229^{+0.0040}_{-0.0050}$, the most precise $\alpha_S(m_Z)$ value obtained using jet substructure observables.



The energy scale uncertainty of the jet constituents, 3% for photons, 5% for neutral particles, and 1% for charged particles, affects the energy of these particles and, consequently, the overall jet p_T. The track reconstruction efficiency uncertainty, 3%, reflects the mismodeling of the efficiency to reconstruct charged-particle tracks in the dense core of the jets [53]. The MC event generators differ in the PS and hadronization modeling, in the tuning of parameters, and in the fixed-order matrix calculation for the hard scattering. These differences are not covered by the renormalization scale uncertainty in the PS and by varying the UE models in PYTHIA8. Therefore, we use the differences between the results obtained with the baseline PYTHIA8 MC and the other simulated samples to evaluate one-sided uncertainties in the MC modeling. The (unfolded) particle-level data distributions are recomputed with the responses corresponding to the variations mentioned above, to establish their uncertainty. The variations in MC modeling contribute the largest source of uncertainty, 2–10% depending on x_L and p_T, followed by the neutral particle energy scale, which contributes an uncertainty of 1–2%.



Detector specific vs observable specific

- Tracking systems
  - Reconstruction efficiency
  - Momentum resolution
  - Fitting/performance (secondaries/decays)
- Calorimeters
  - Energy resolution
  - Physics size restriction
- PID
  - dE/dx selections etc…

- Jets
  - Clusterizer finding efficiency
  - Corrections, unfolding
- Resonances
  - Invariant mass resolution
  - Two track resolution
- Flow
  - Convolution of tracking pointing and energy resolution
  - Focuses more on counts, less sensitive overall

A common way to correct in data

- Start with your favorite experimental observable
- It needs to be corrected for detector effects
- Say you have a Monte Carlo generator A - TRUTH (gen) and a GEANT (reco) sample
- You split your A sample into two statistically independent datasets - A1 and A2
- You match the reco to gen - derive your corrections from A1, apply to A2 - this ratio is what we typically call 'closure' - this better be 1…
- Then you apply the corrections you got from A to your data - you then have a corrected result!

Are we done…?

Slight issue here…

- What if your reco A DOES NOT match raw (uncorrected) data?
- Now, say you have two generators - A and B - different physics implementations (PYTHIA vs HERWIG for example)
- You run them through your detector and get reco A and reco B.
- What if reco A and reco B are indistinguishable from each other?
- What happens if truth A and truth B are similar in one observable BUT different in another observable - how to assign an uncertainty to the data that inherently theory dependent

Correction procedure - Non closure

- There is precedent for this!

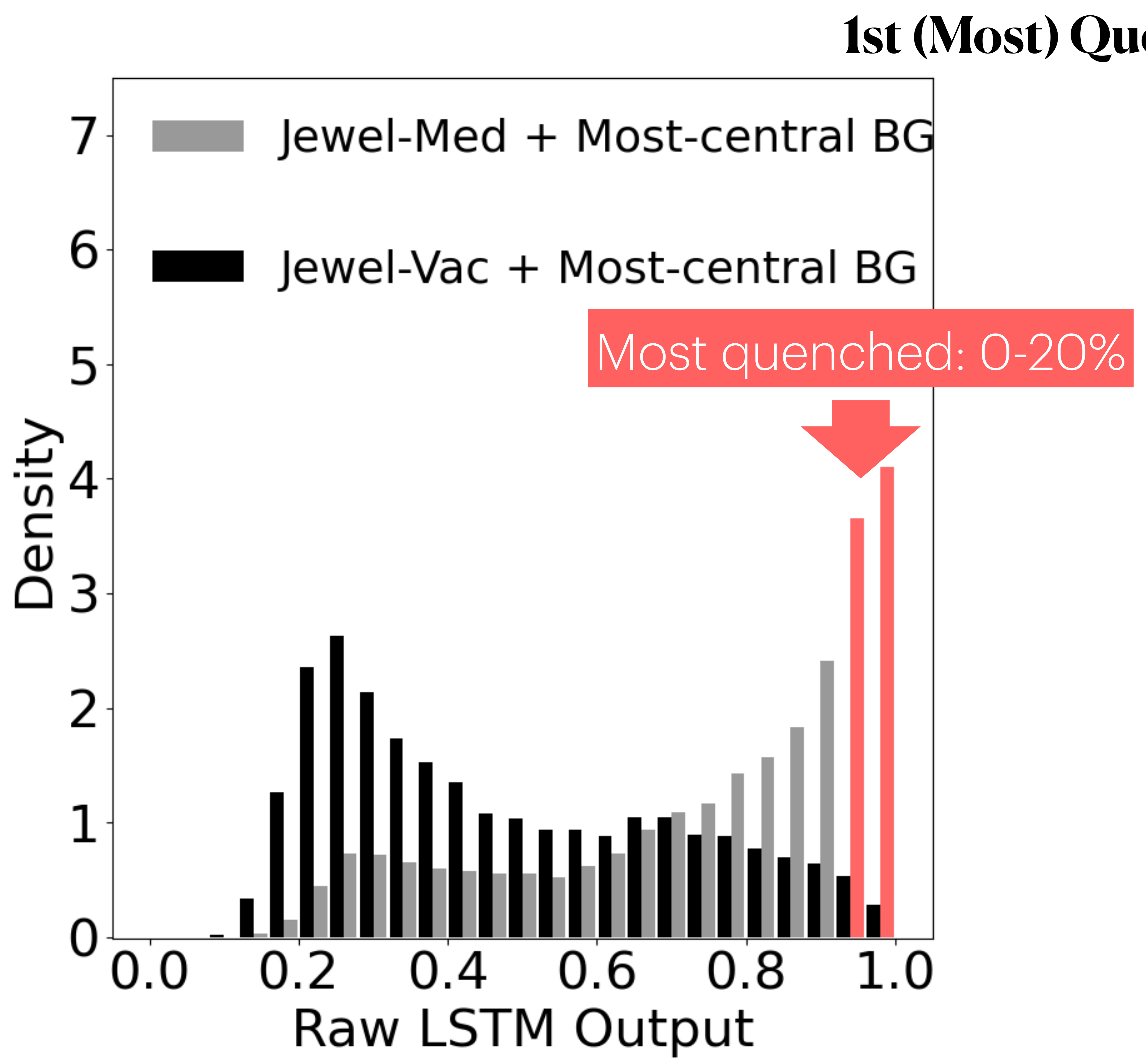
- If you start with 4, your detector changes it to 5, you add a correction of 20%, you go back to 4.2 - the difference between 4 and 4.2 is an uncertainty…

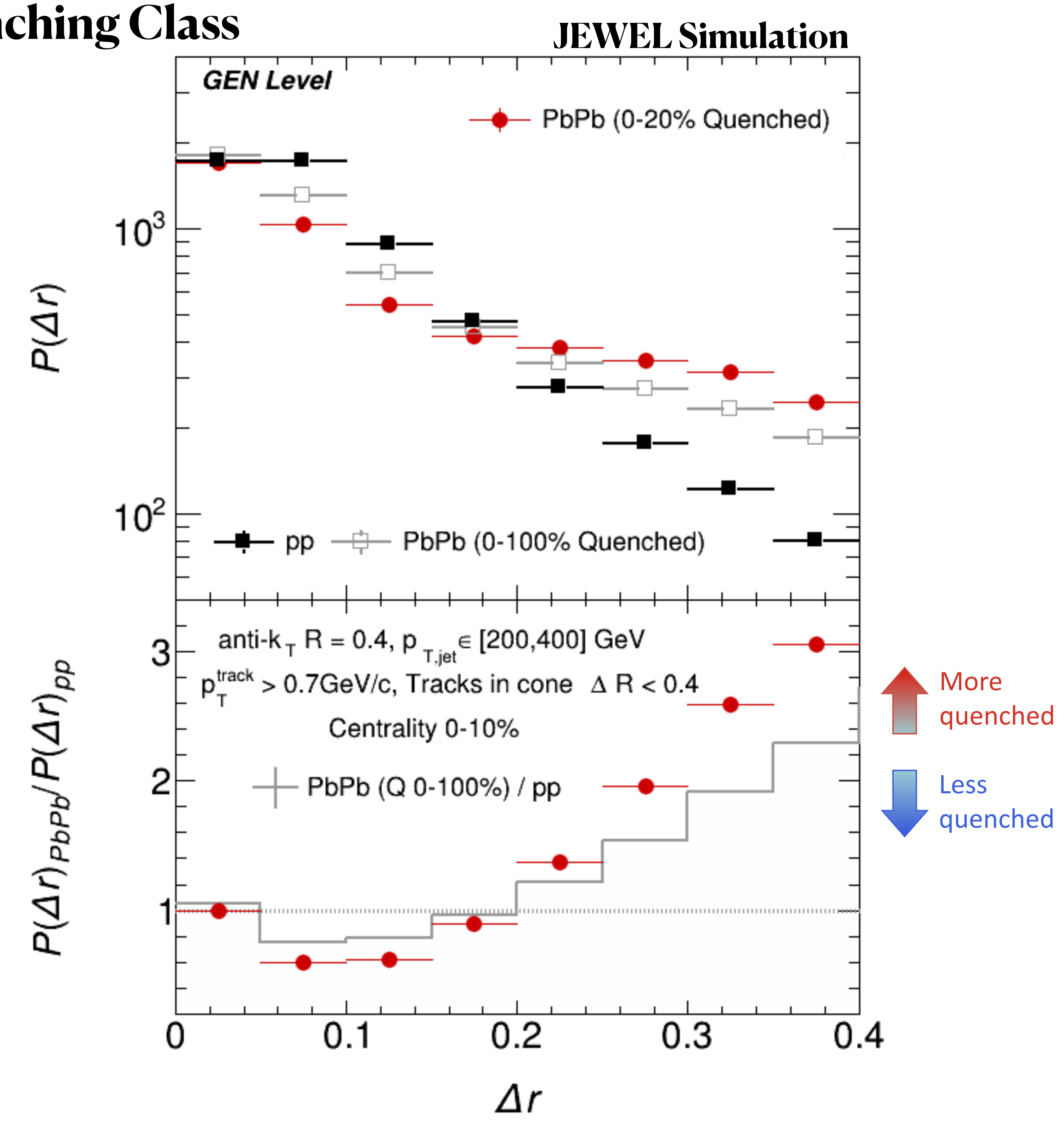
Experimentalist

- Slowly starting to be the industry standard

- Jets are an interesting examples for this - different jets are different

Lets ask a simple question



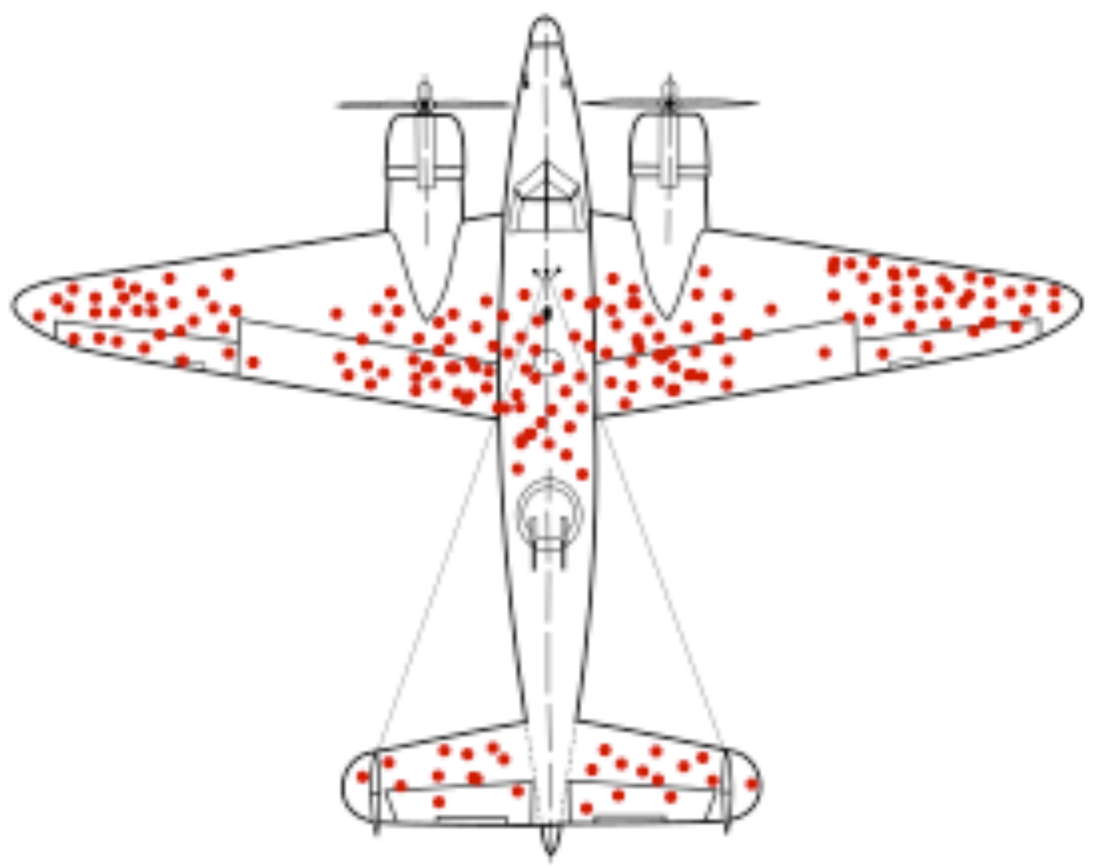


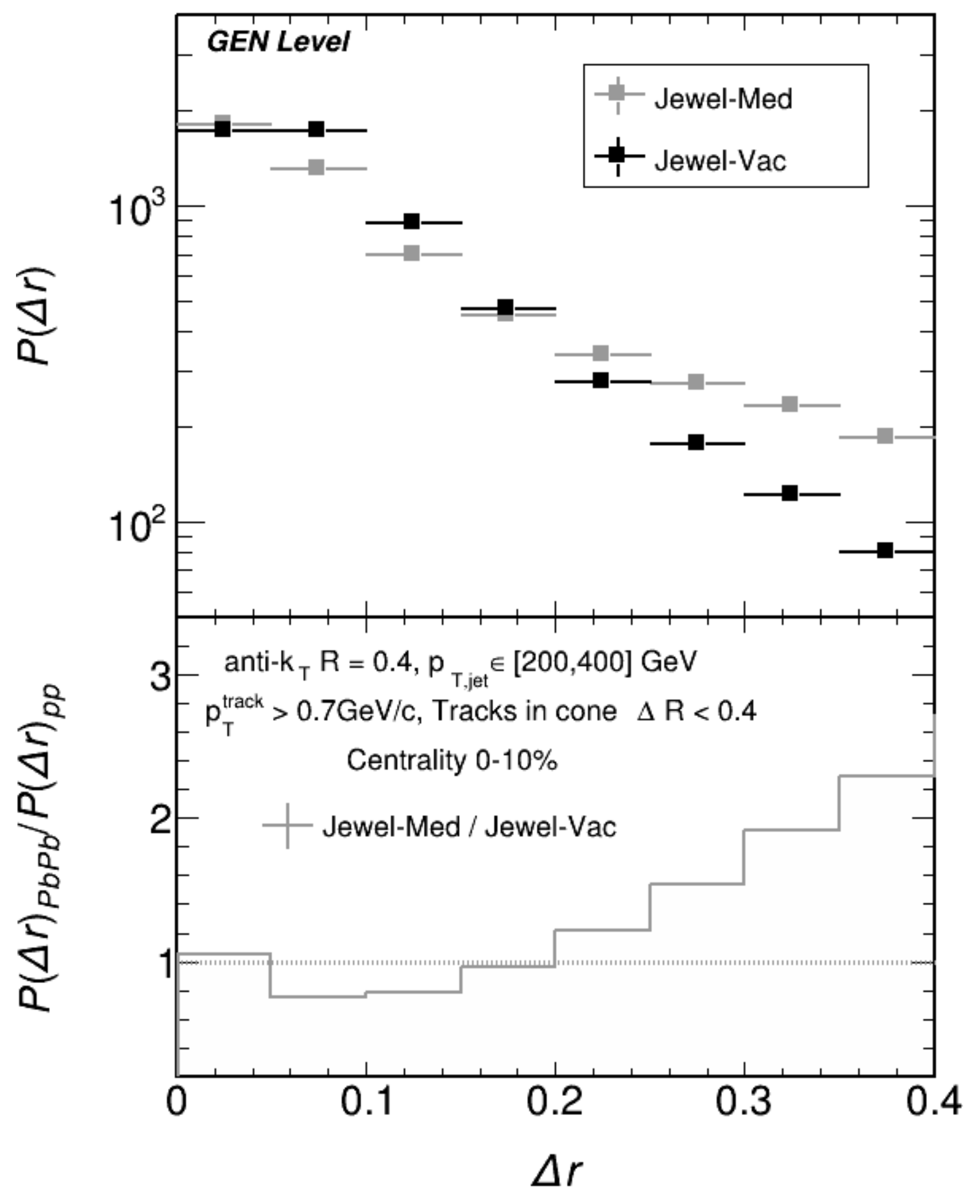
Quenchness: The LSTM output for each medium jet.
If the value is closer to 1, then the jet is more quenched.
Vice versa.

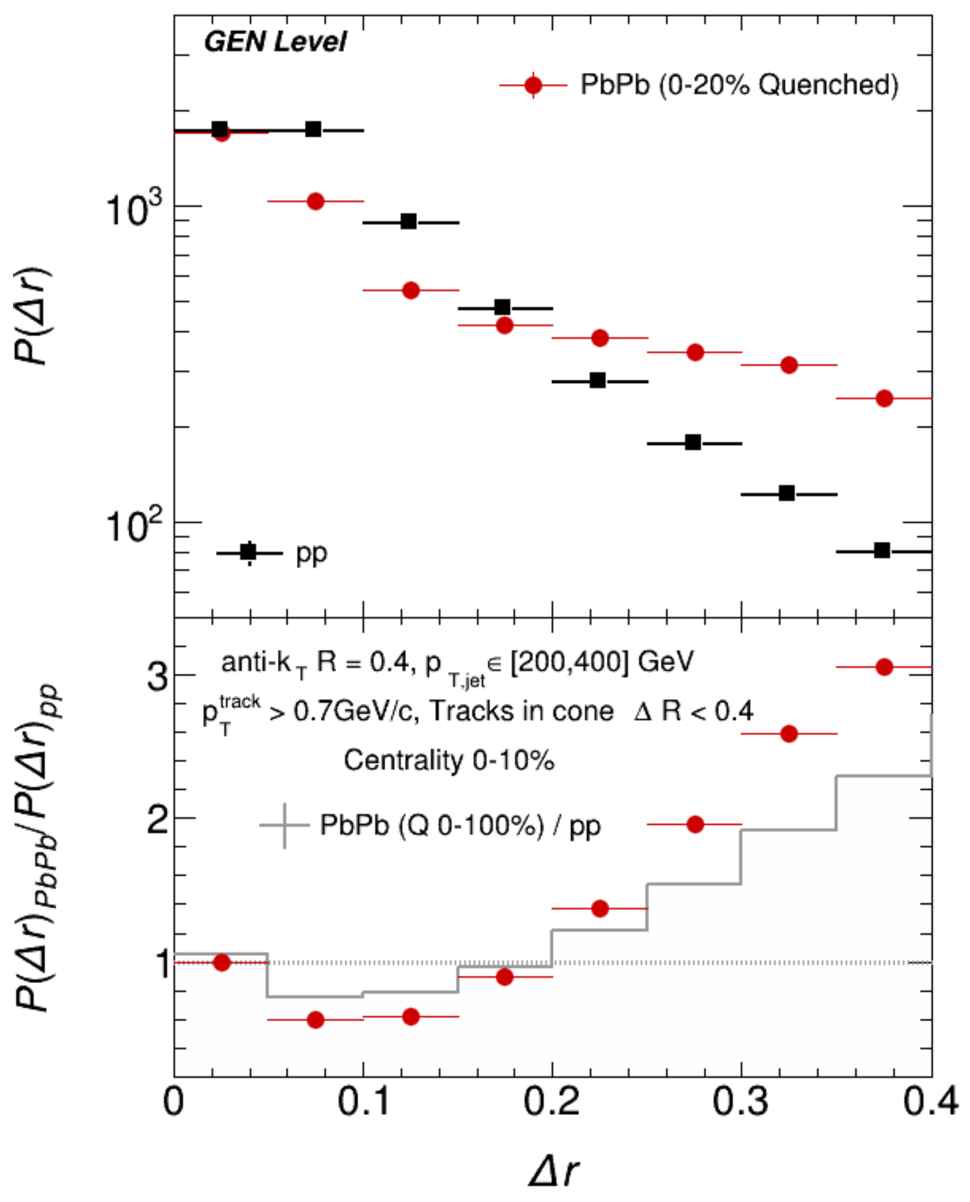
Yilun Wu (Vanderbilt) RHIC/AGS Users Meeting 2024

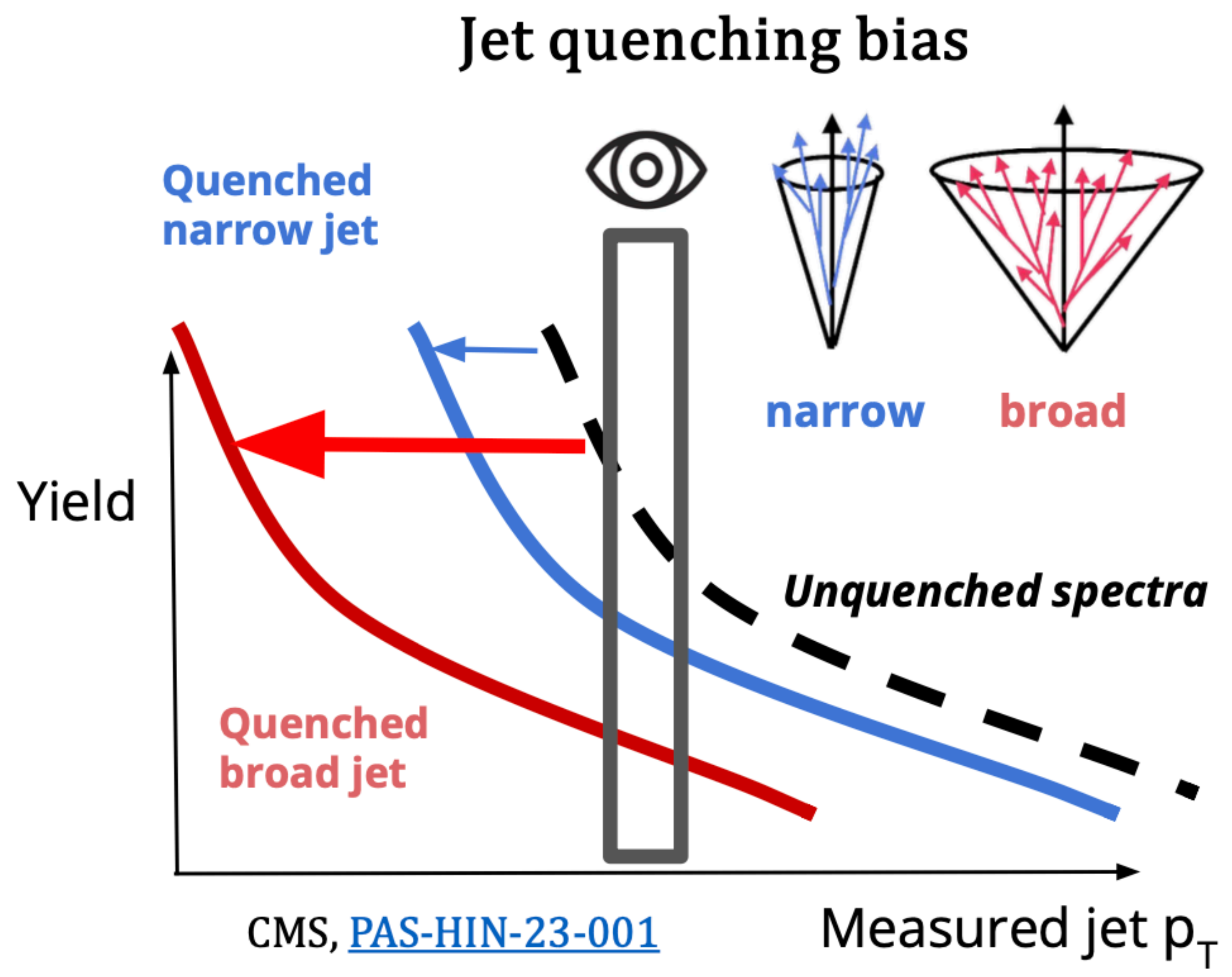
Be very careful!

- Signal and detector effects are similar…
- Survivor bias is a real thing for some class of measurements





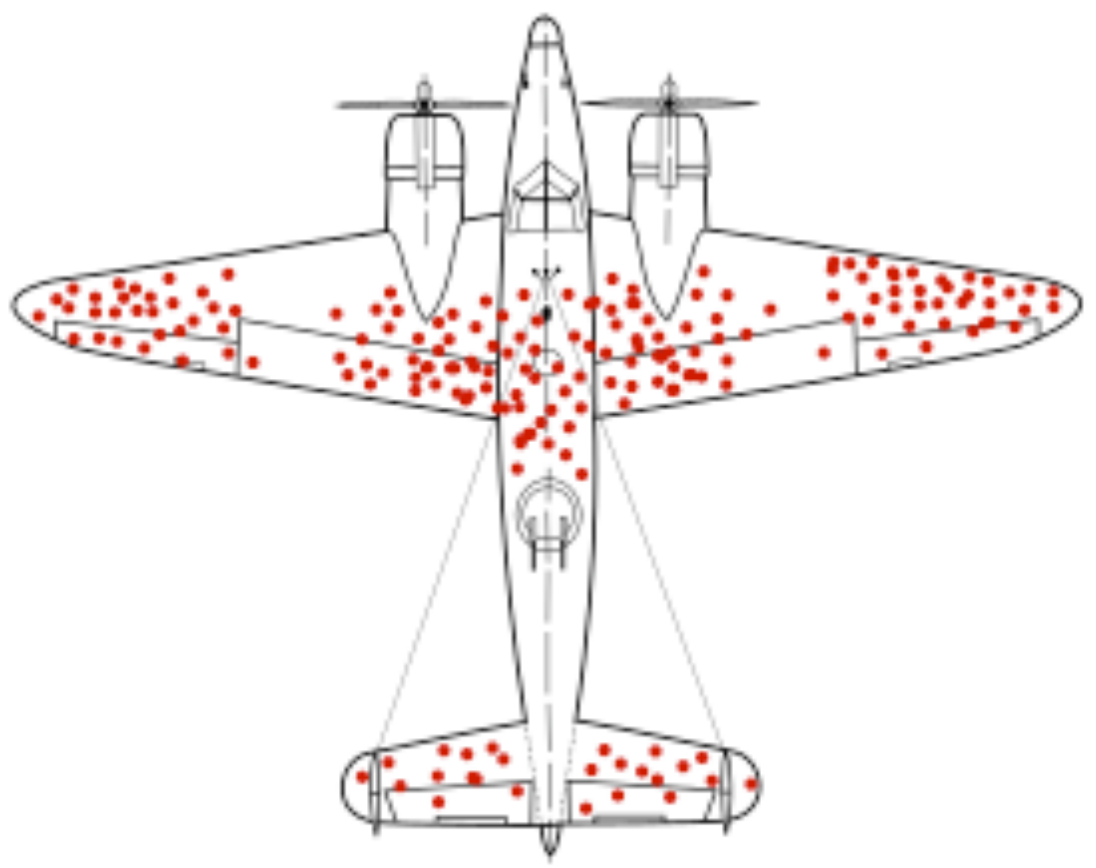


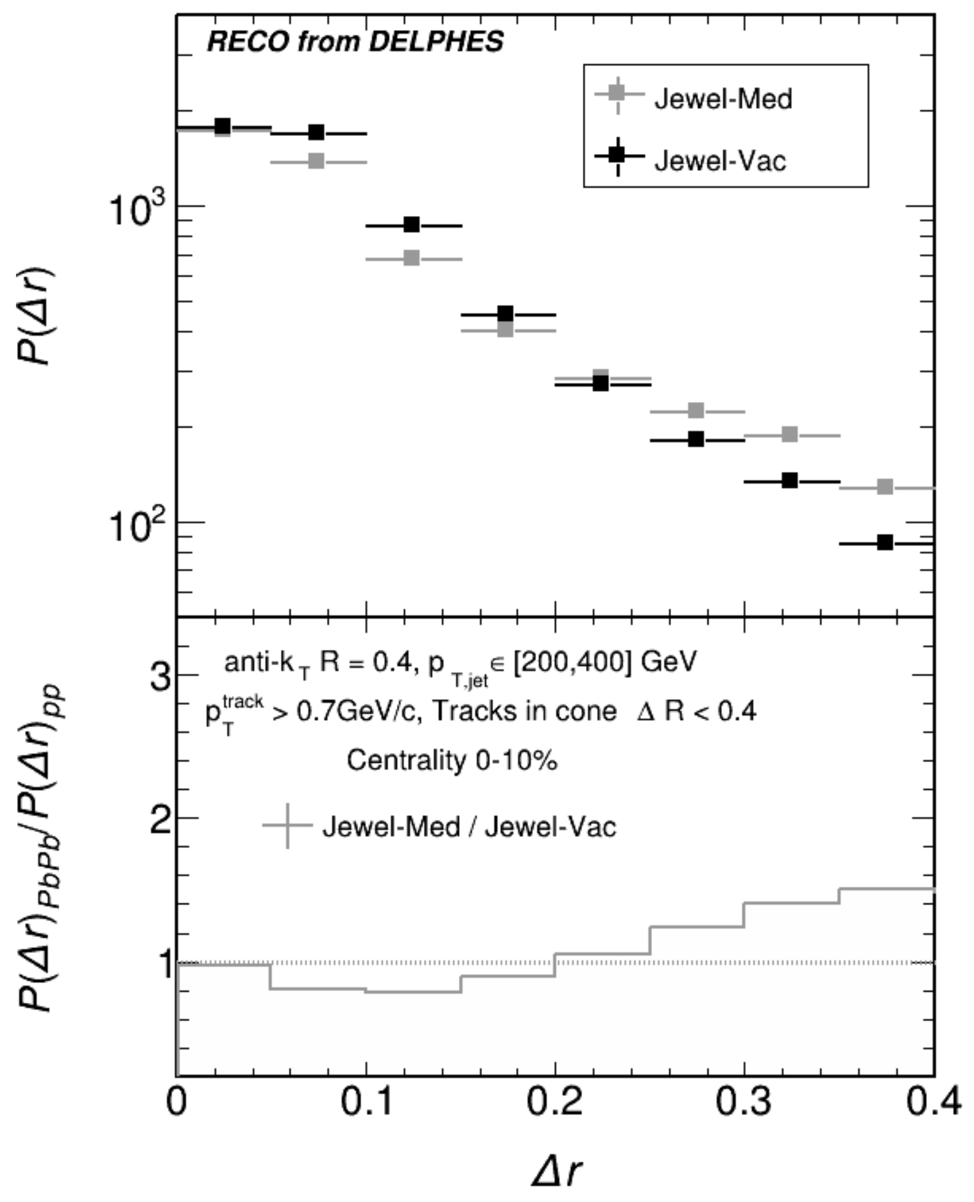


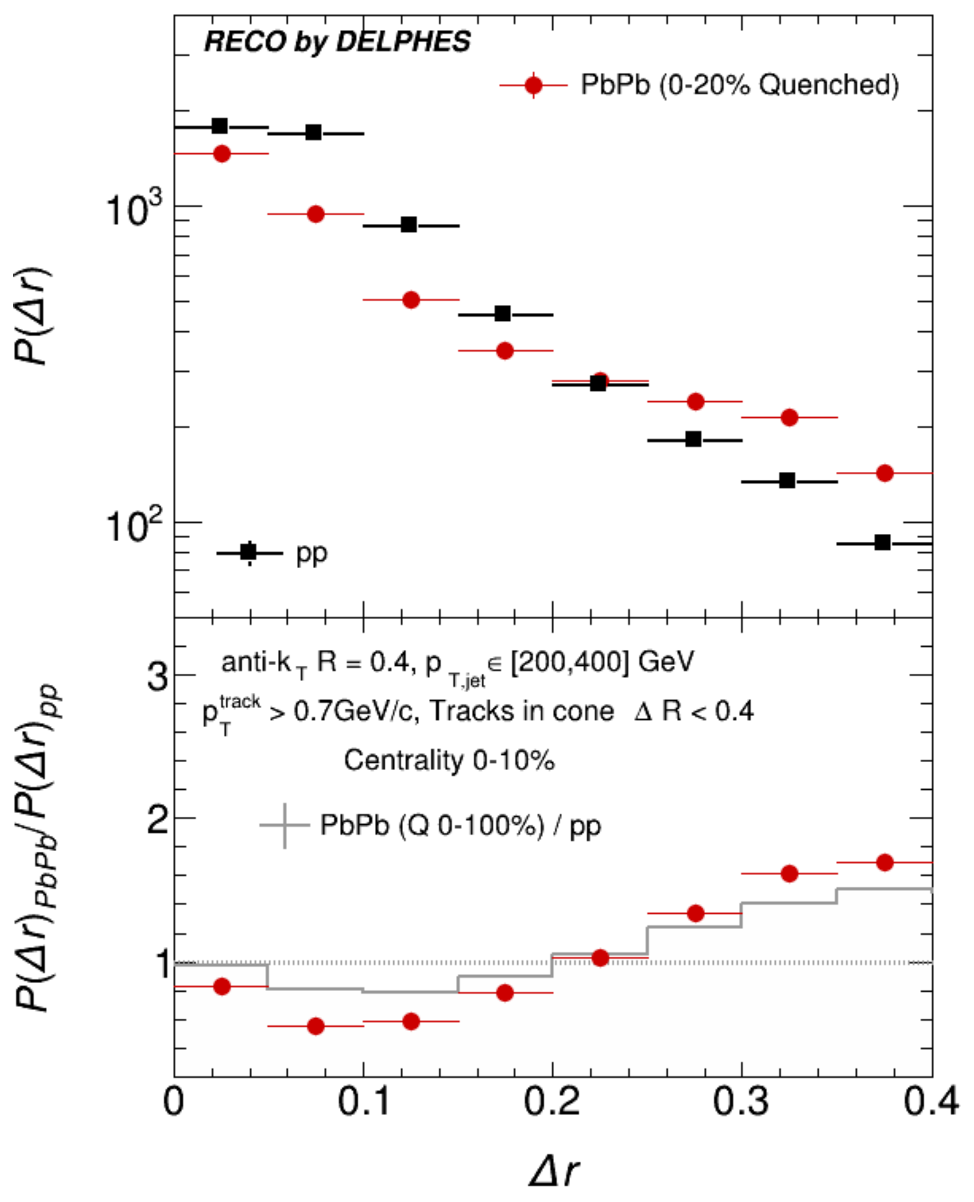
Yilun Wu (Vanderbilt) RHIC/AGS Users Meeting 2024

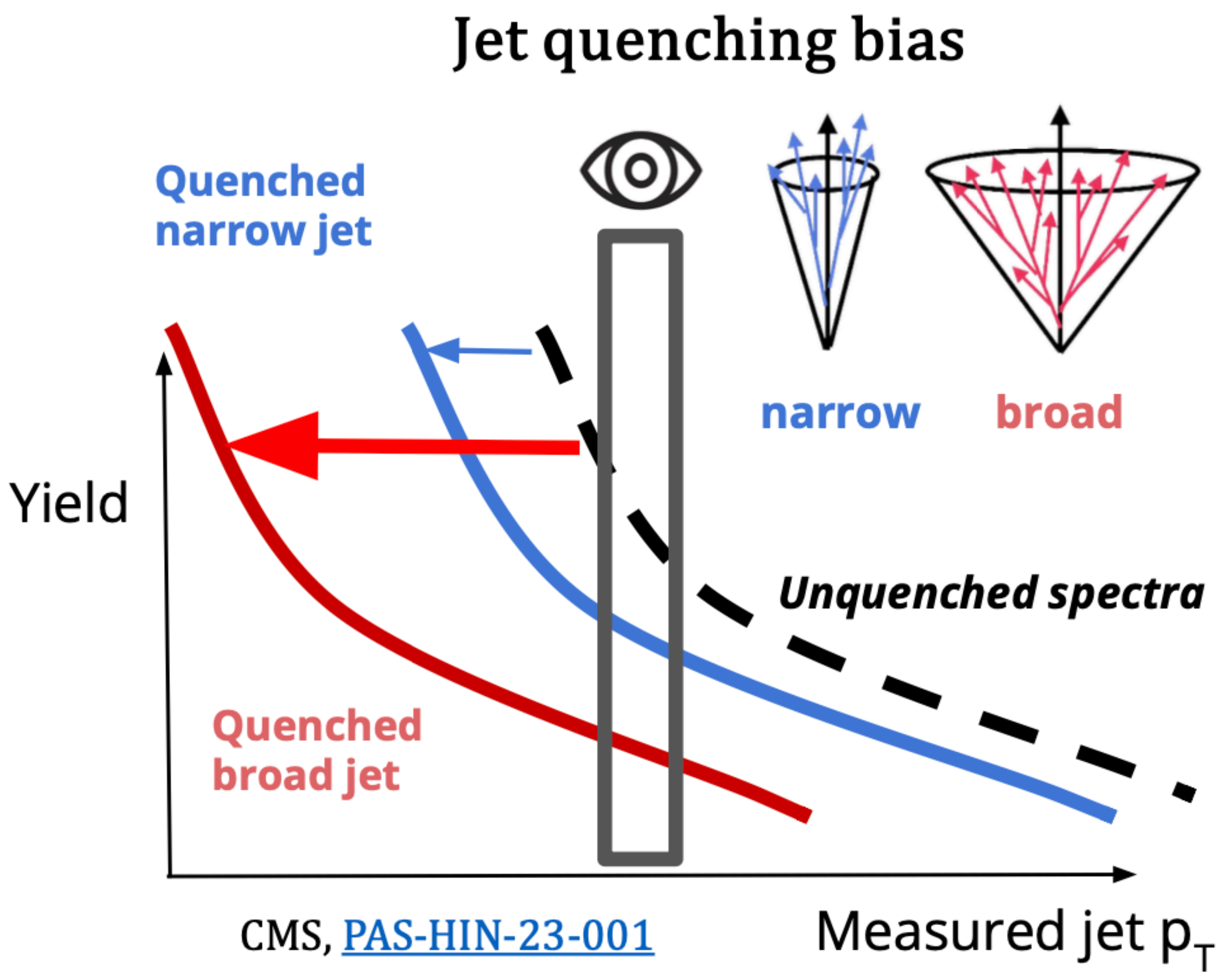
Be very careful!

- Signal and detector effects are similar…
- Survivor bias is a real thing for some class of measurements



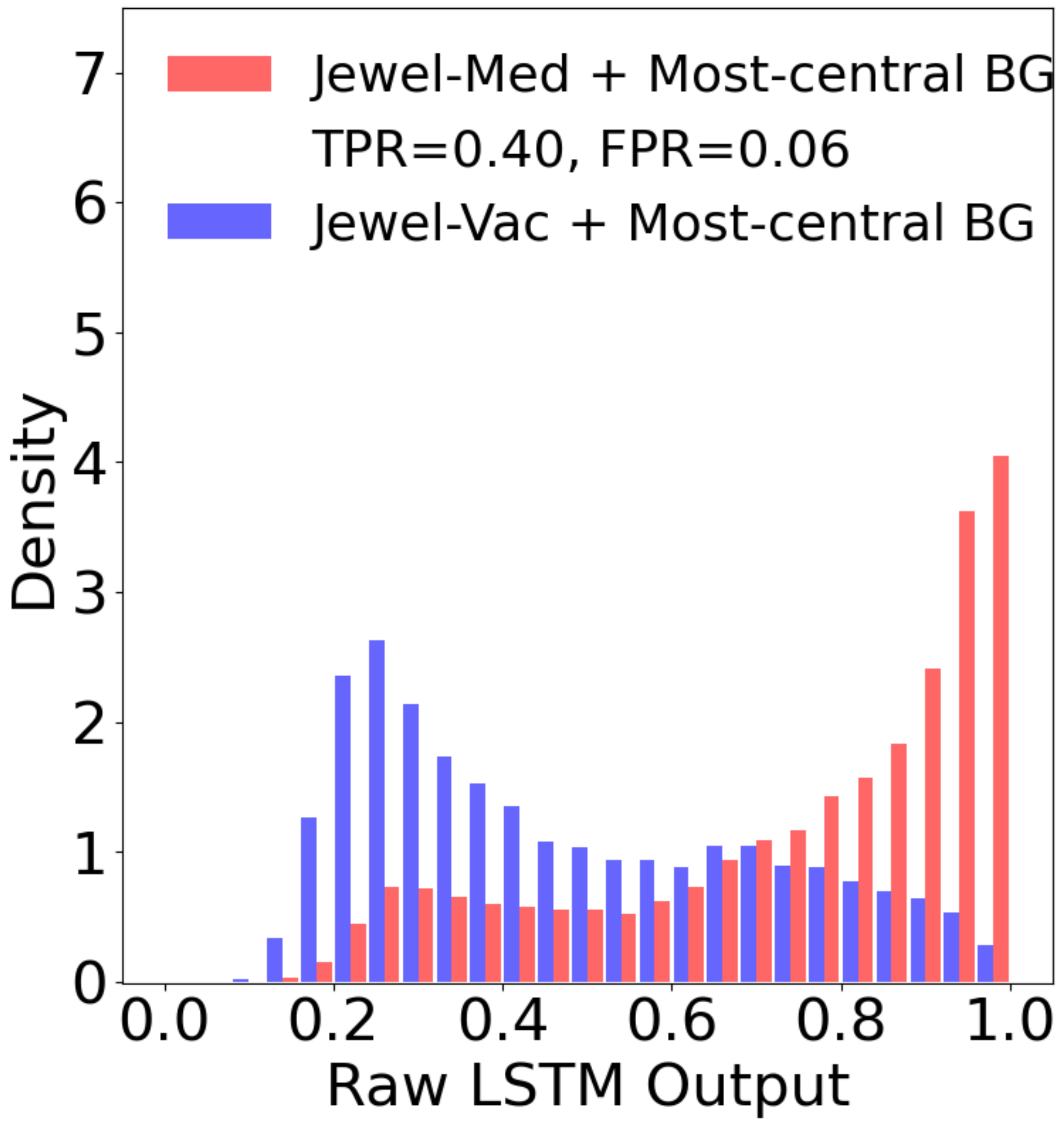




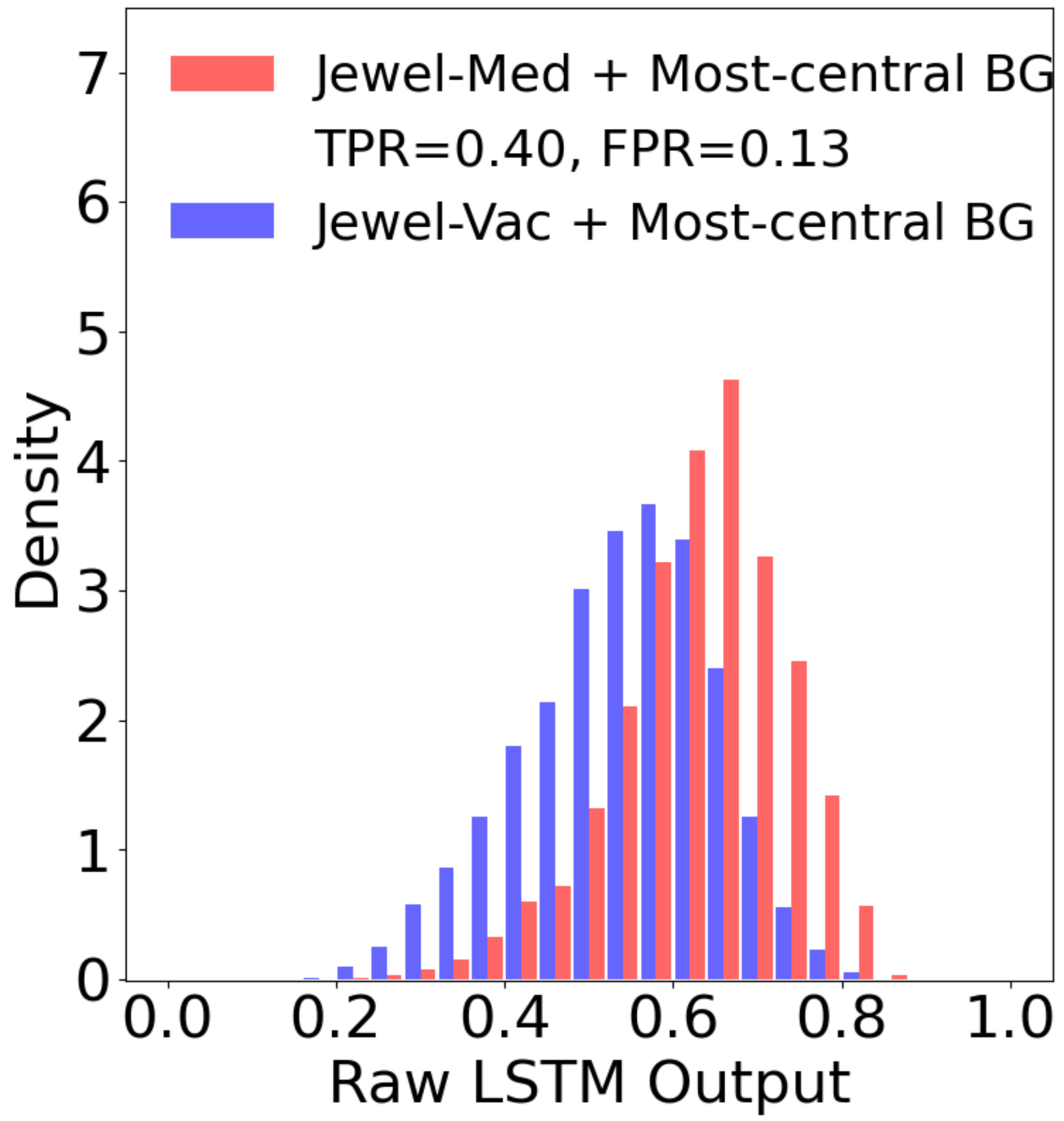


Yilun Wu (Vanderbilt) RHIC/AGS Users Meeting 2024

GEN Level jet training



Detector effects smear the differences between medium jets and vacuum jets

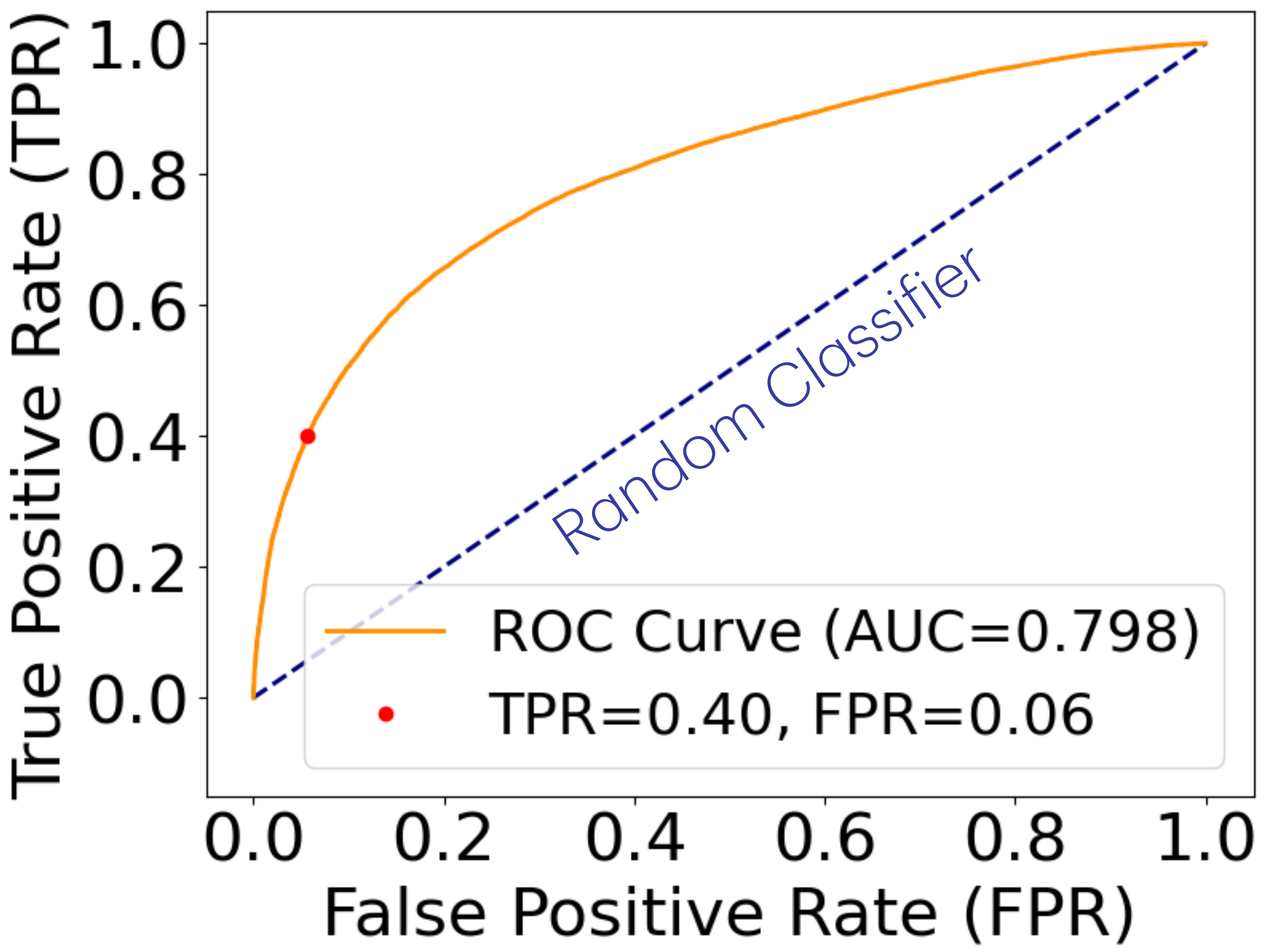


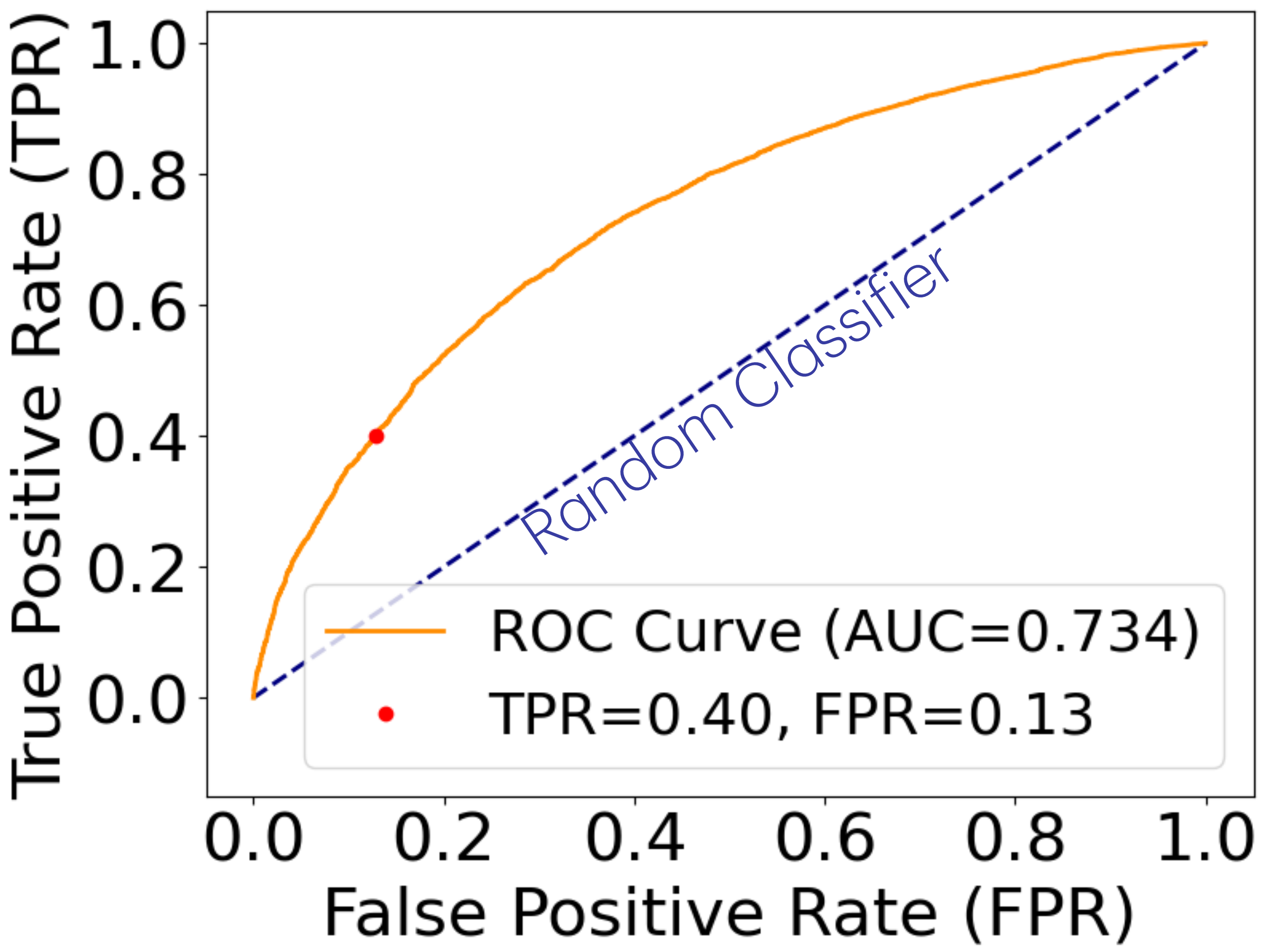
RECO jet training

EFlow Candidates from DELPHES:

1) Combine the Tracker + Calorimeters

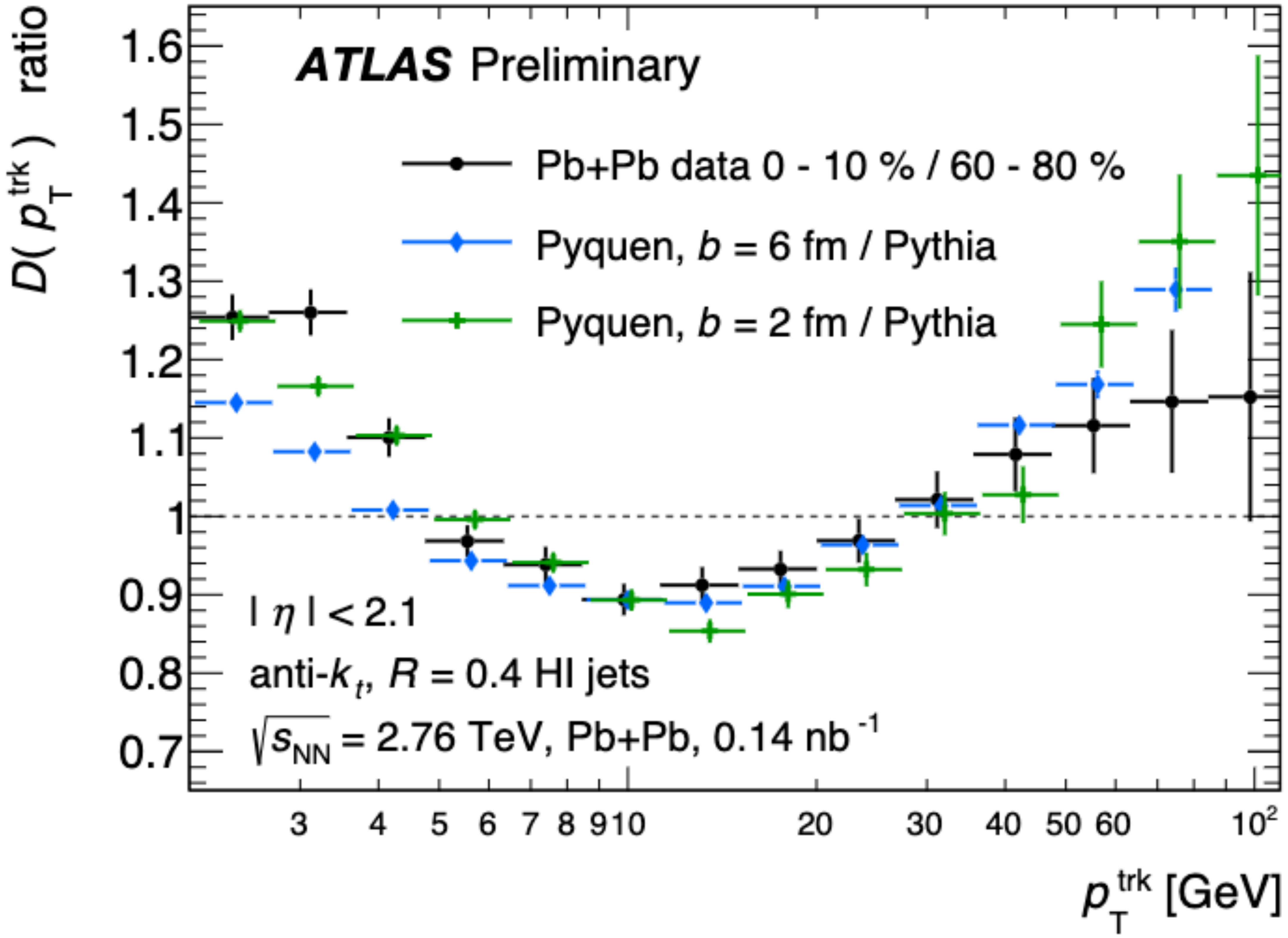
2) Comparable to CMS Particle Flow Candidates

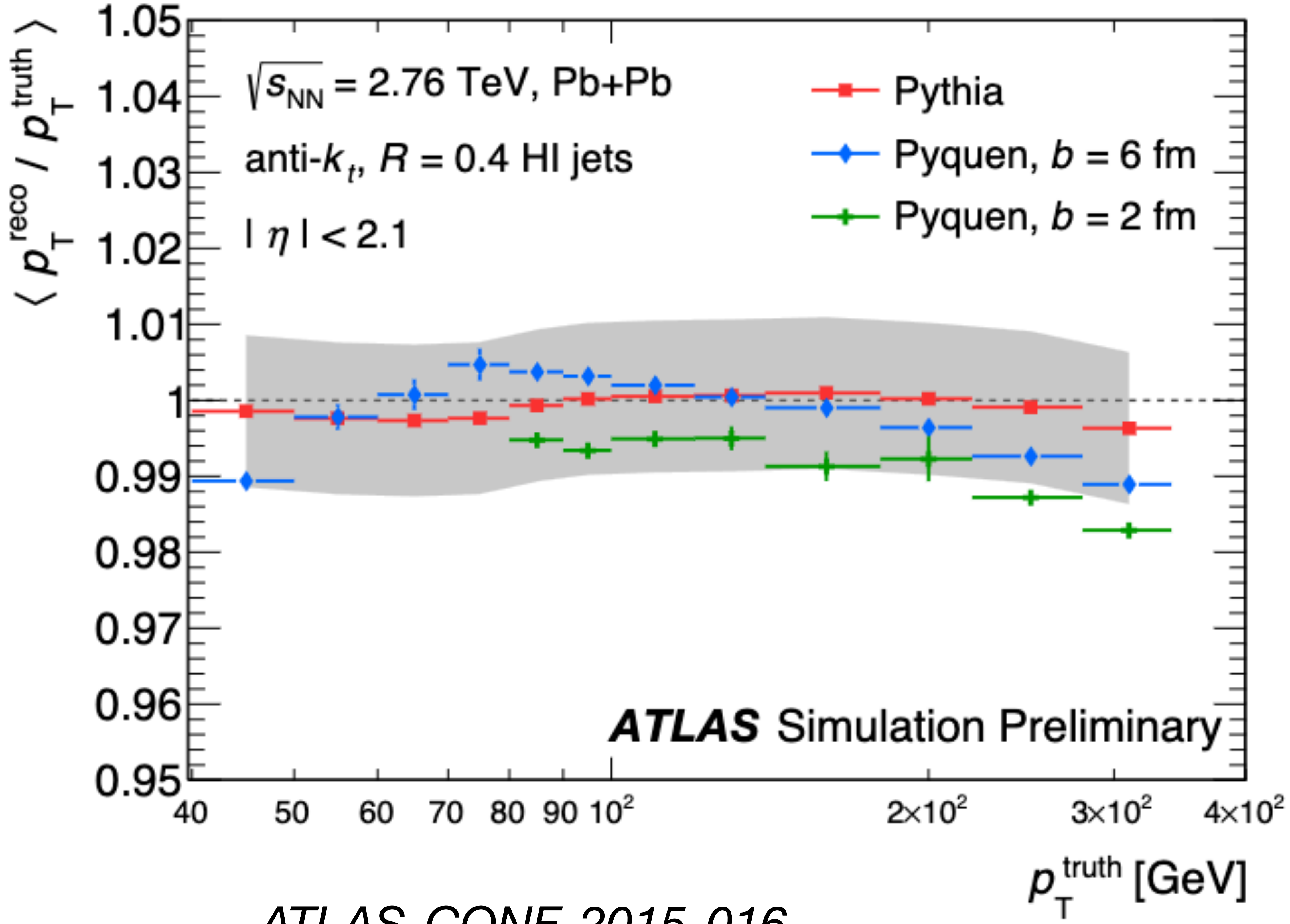




Uncertainties due to unfolding corrections

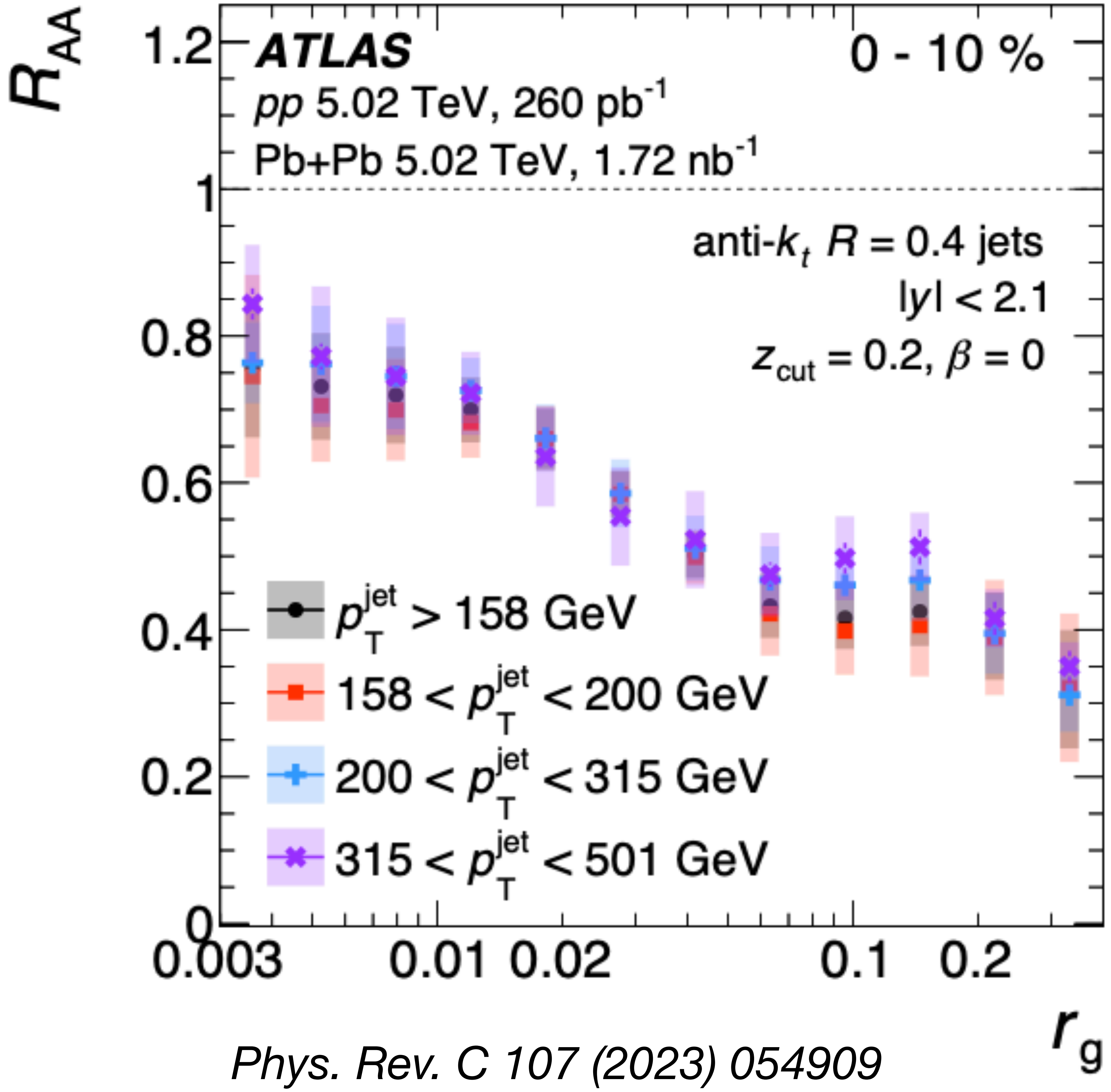
- Impact of jet energy loss on the jet energy corrections and substructure
- What about other models…? Is looking at the fragmentation function enough?





ATLAS-CONF-2015-016

What does this translate to?



Phys. Rev. C 107 (2023) 054909

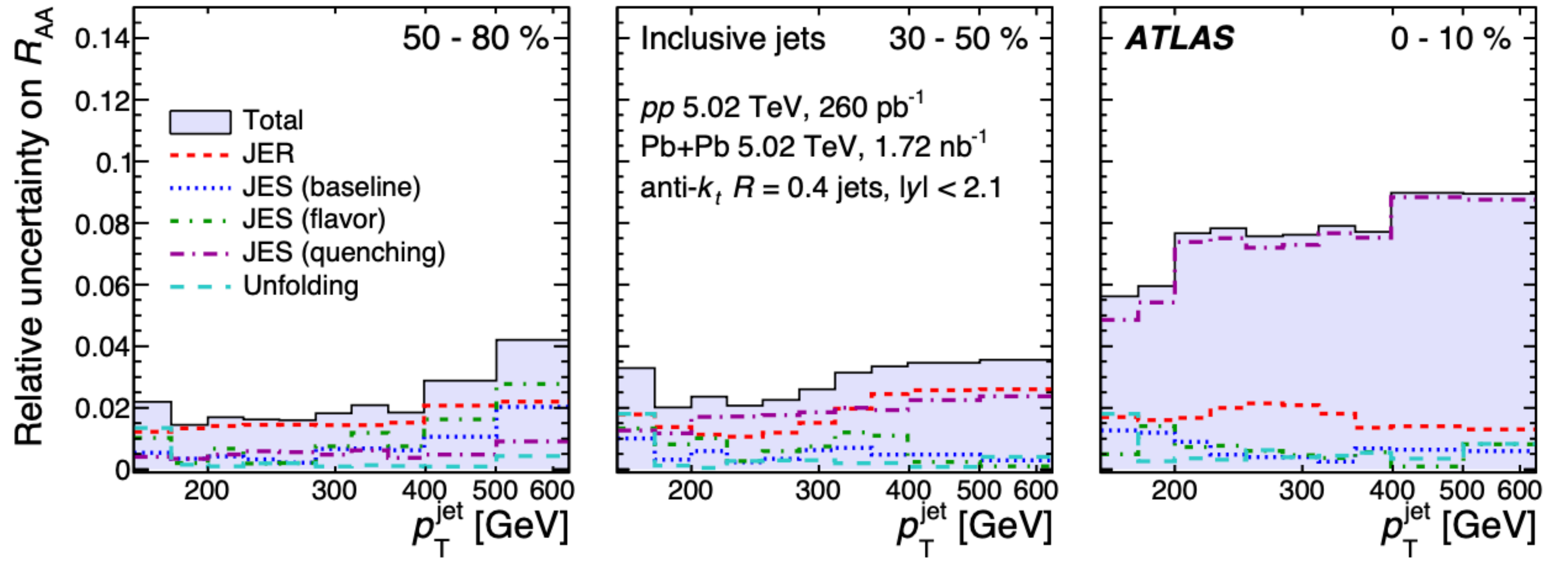


Figure 4: The relative systematic uncertainties in the R_{AA} measurements as a function of p_T^{jet} in different centrality intervals of Pb+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The legend applies to all of the panels.

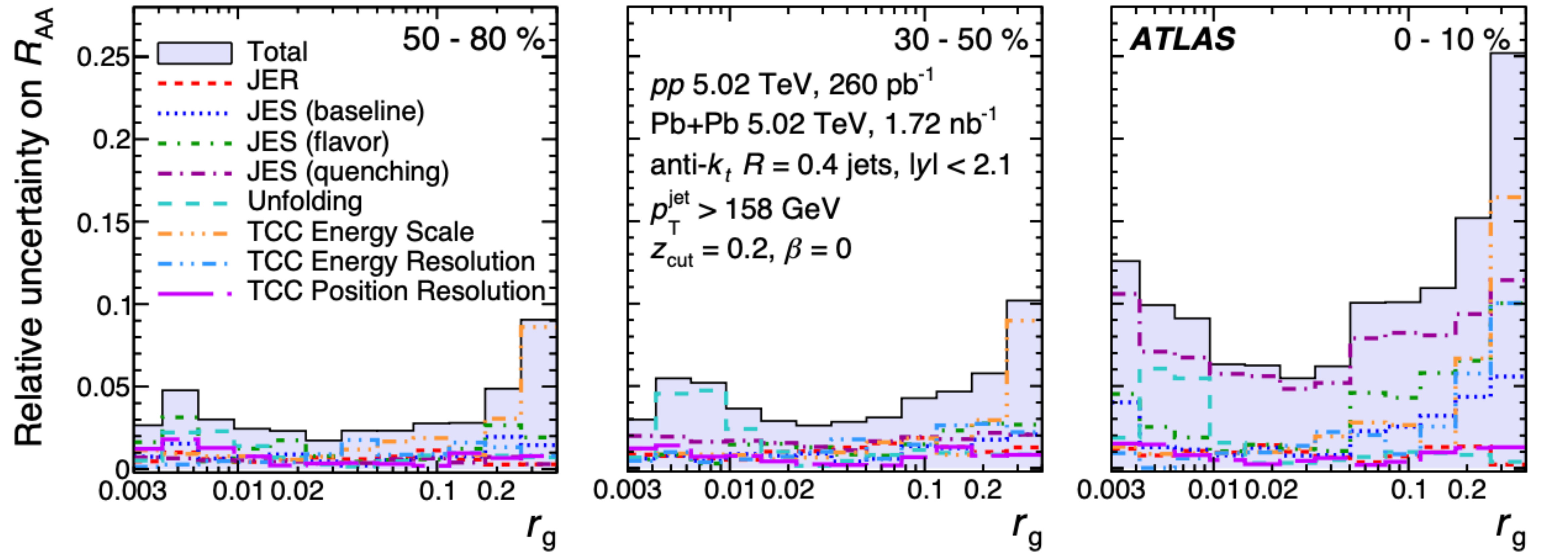


Figure 5: The relative systematic uncertainties in the R_{AA} measurements as a function of r_g in different centrality intervals of Pb+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV shown for soft-drop parameters $z_{cut} = 0.2$ and $\beta = 0$. The legend applies to all of the panels.

Interplay of signal, background, detector effects in the observable of interest

- What if the quantity and the specific kinematic range you are looking at is extremely affected by the modeling of the MC
- Add to that the effect of the energy loss and detector effects
- For jets in particular - this is a convoluted process which makes it challenging!
- Parton Shower $\oplus$ energy loss $\oplus$ hadronization $\oplus$ detector effects
- Is any of these sources factorable?
- What about the data from the last decade that used an unquenched simulation to correct simulated data…?