Heavy-Ion model studies with Bayesian analysis

Yi Chen (Vanderbilt U.)

Jul. 12, 2024. INT 24-88W Workshop

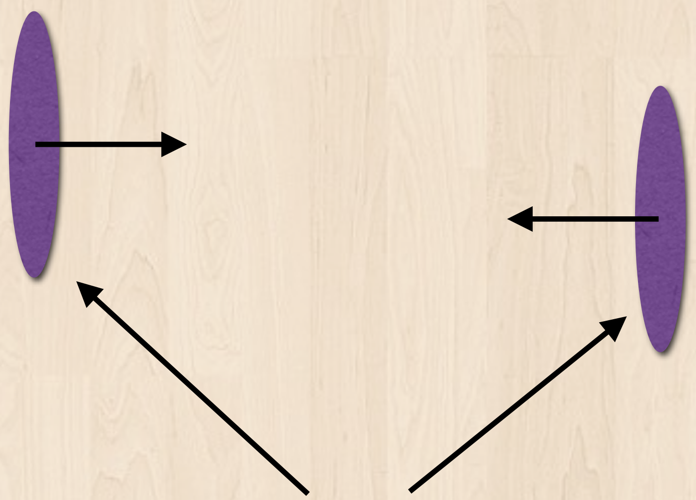
with the JETSCAPE collaboration

Manuscript in preparation 2407.XXXX

The Vanderbilt HENP work is supported by DOE-NP

Setting The Stage: Heavy-Ion Collisions

Ultrarelativistic heavy-ion collisions

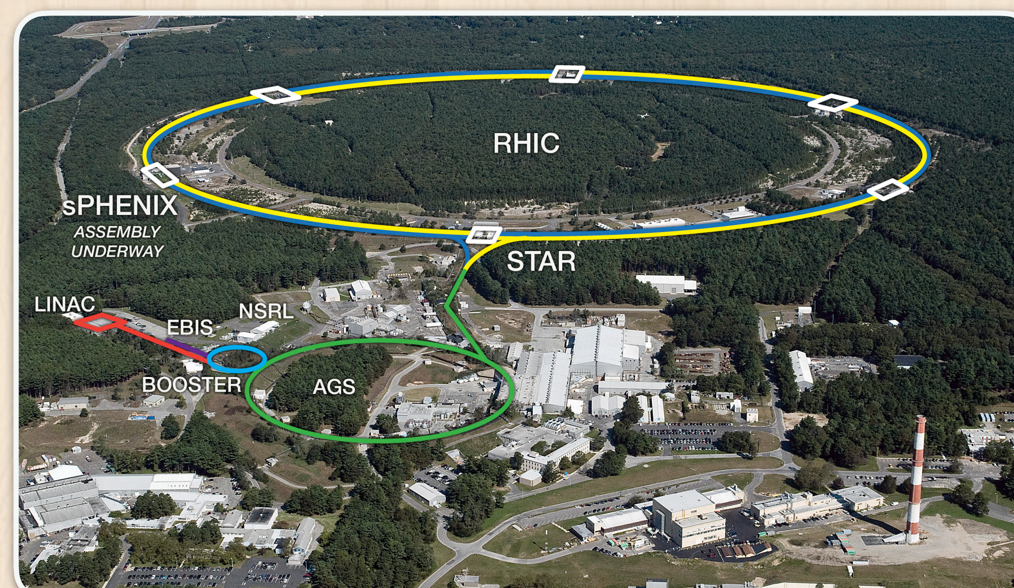


e.g. Pb/Au/Xe/… ion

Accelerate heavy ions to extreme speed and collide

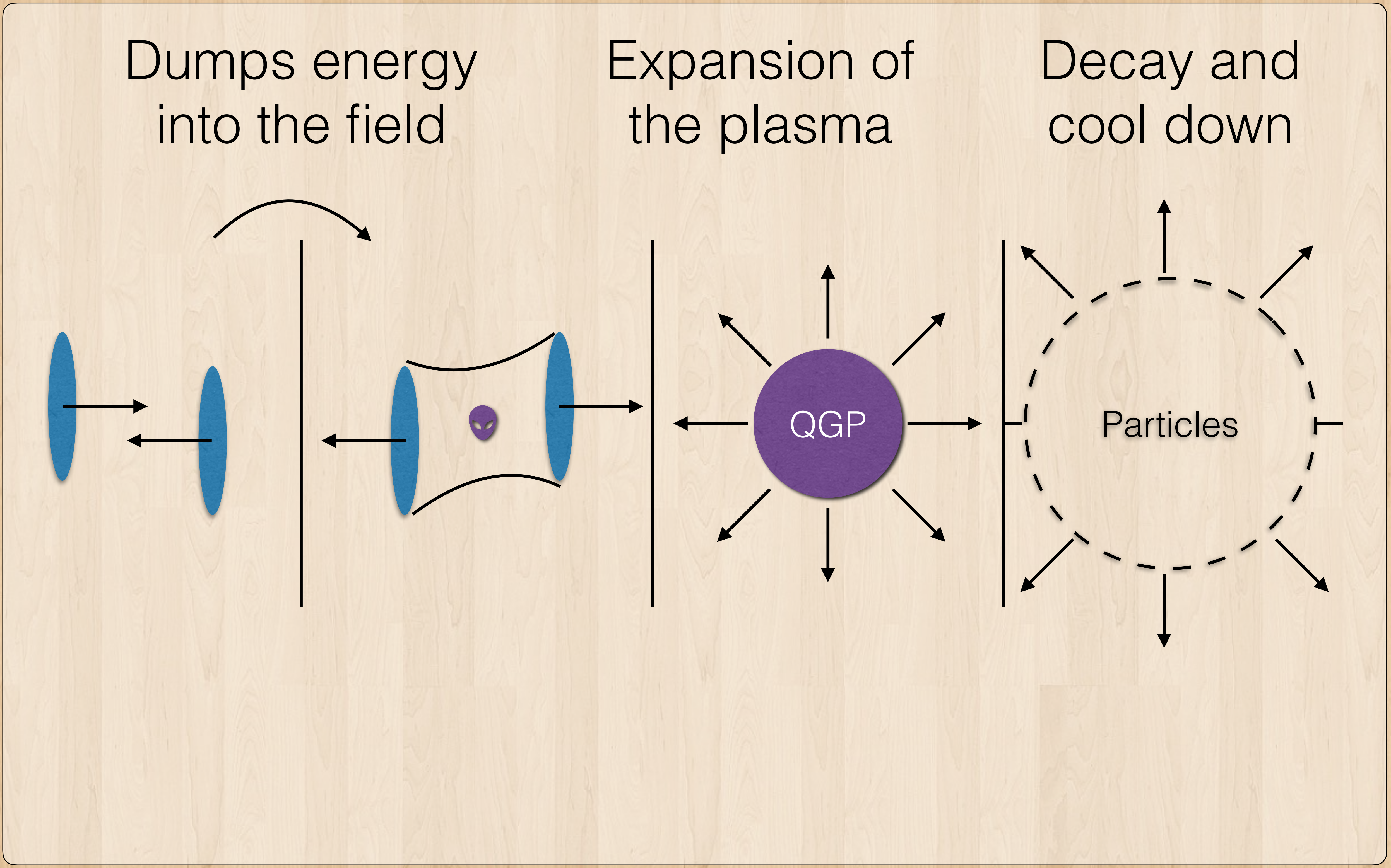
> 99.99999% speed of light (Lorentz γ up to ~2700)

LHC, CERN, Geneva



RHIC, BNL, New York

What happens after collision?



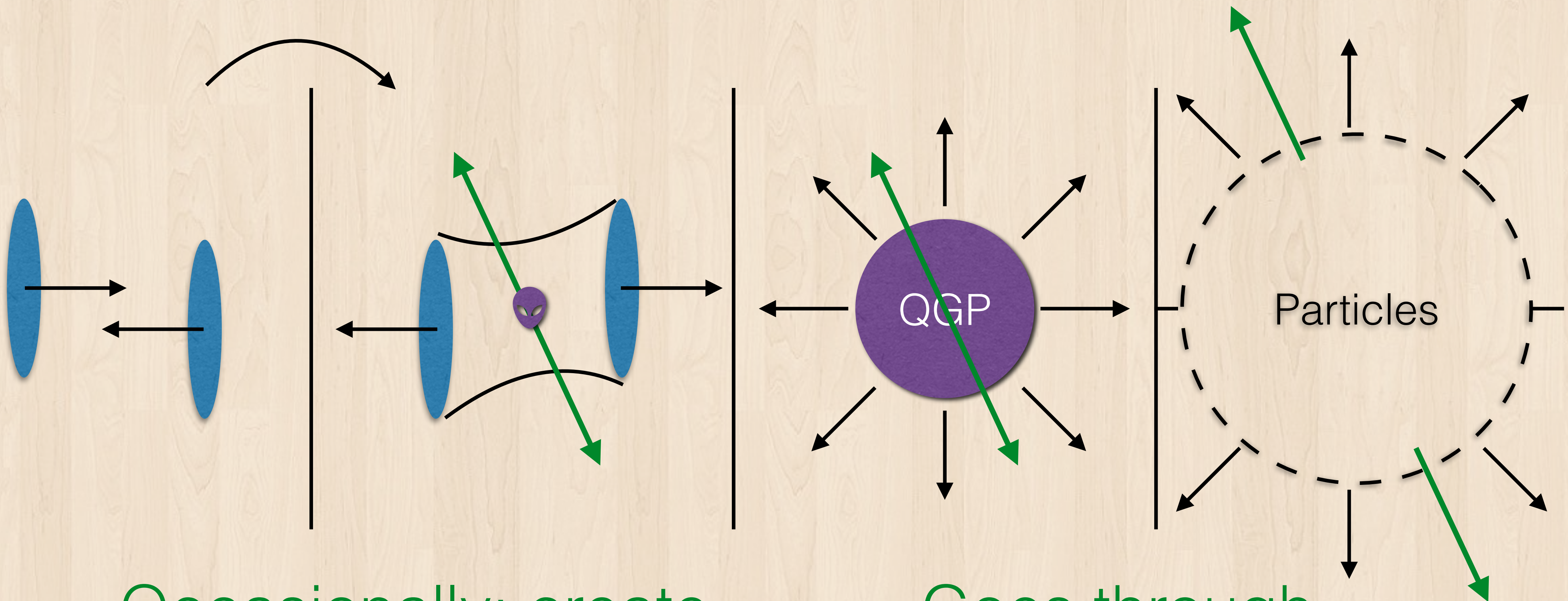
QGP = quark gluon plasma

What happens after collision?

Dumps energy into the field

Expansion of the plasma

Decay and cool down

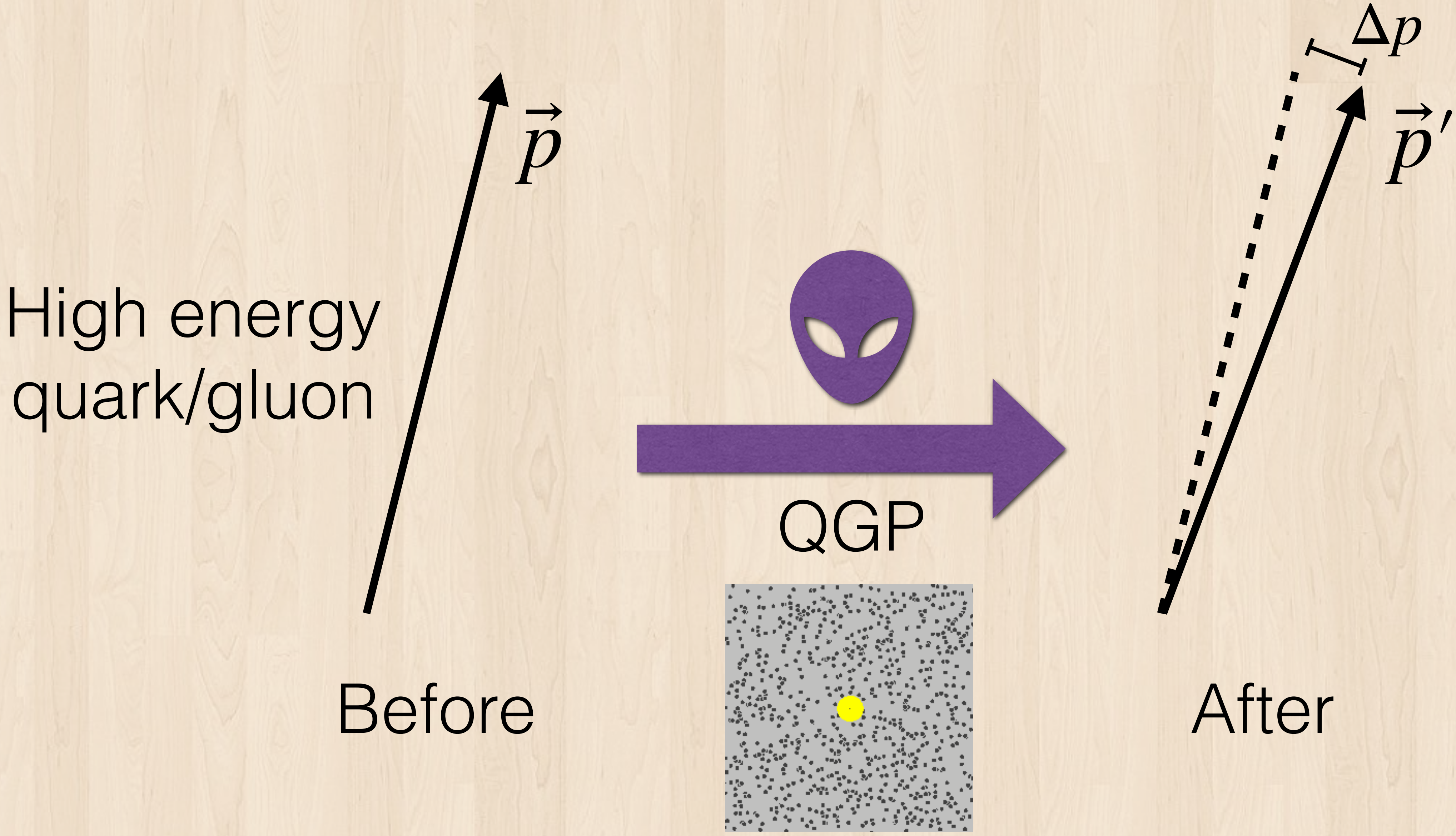


Occasionally: create high energy particles

Goes through QGP and interact

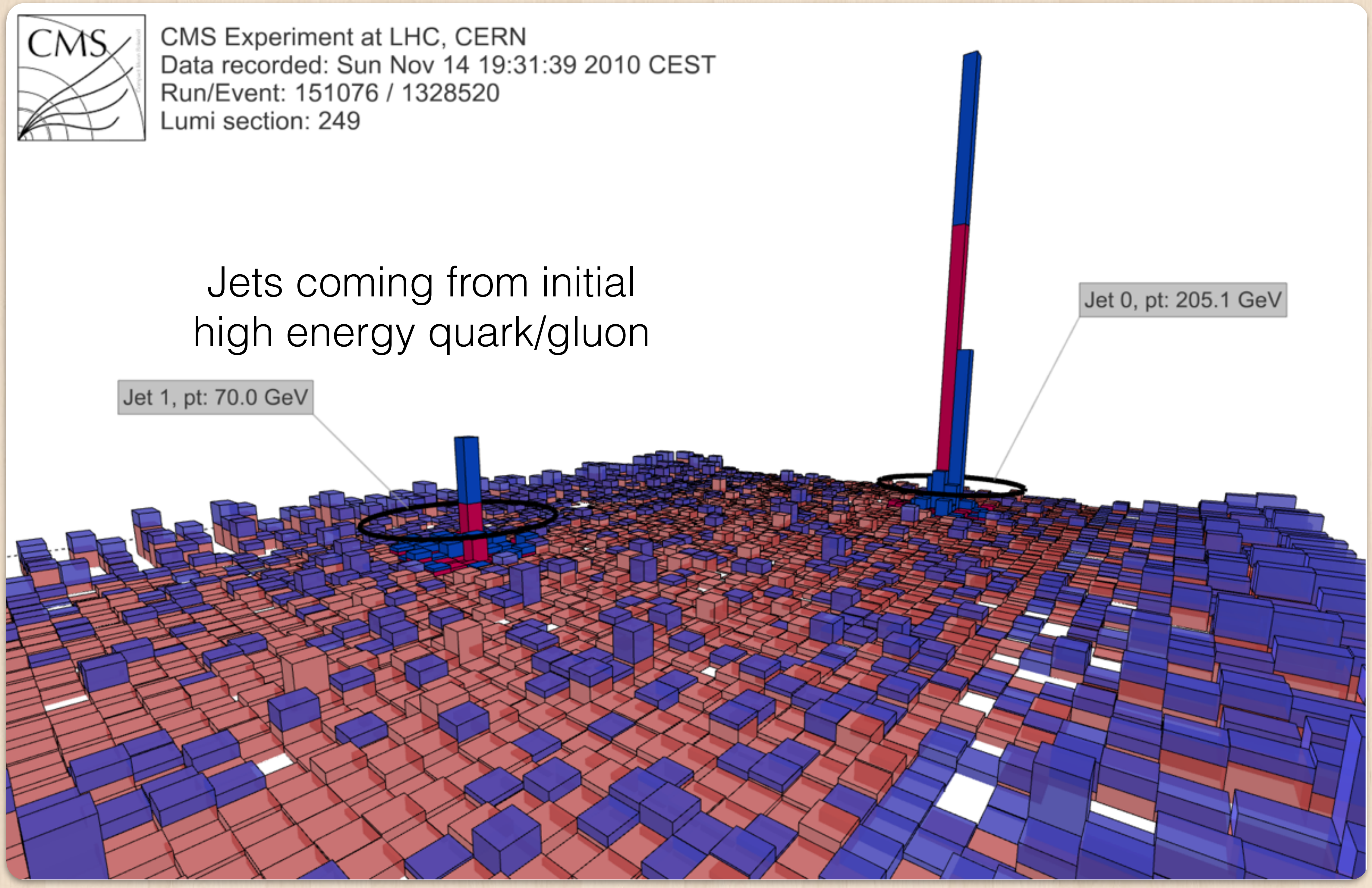
e.g. Jet quenching effect

Transport coefficient $\hat{q}$

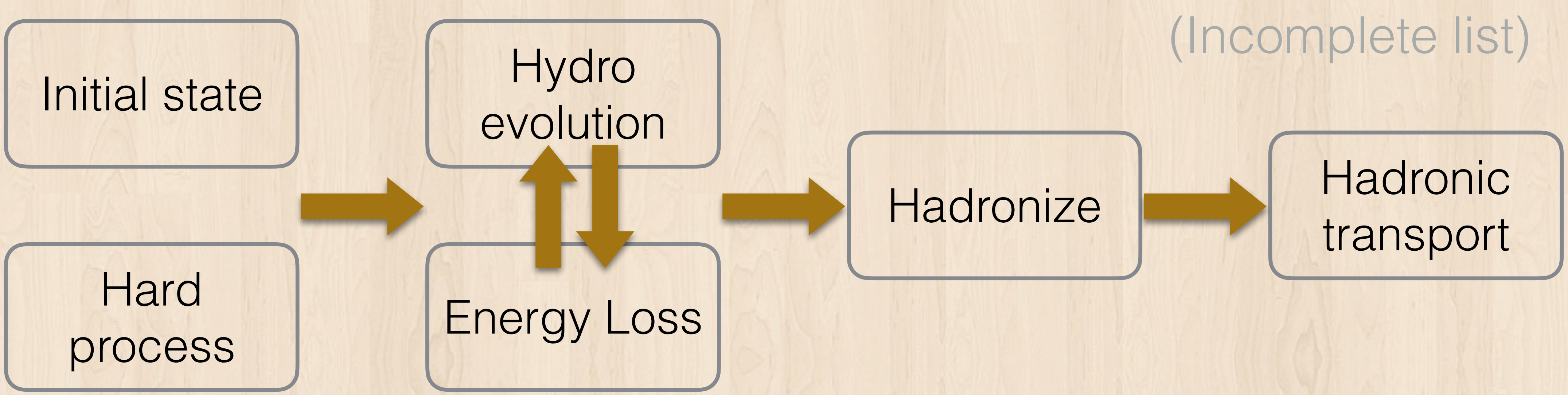


$\hat{q}$ characterizes the size of the Δp^2 after traveling some length in QGP

Example collision event



Current approach to modeling

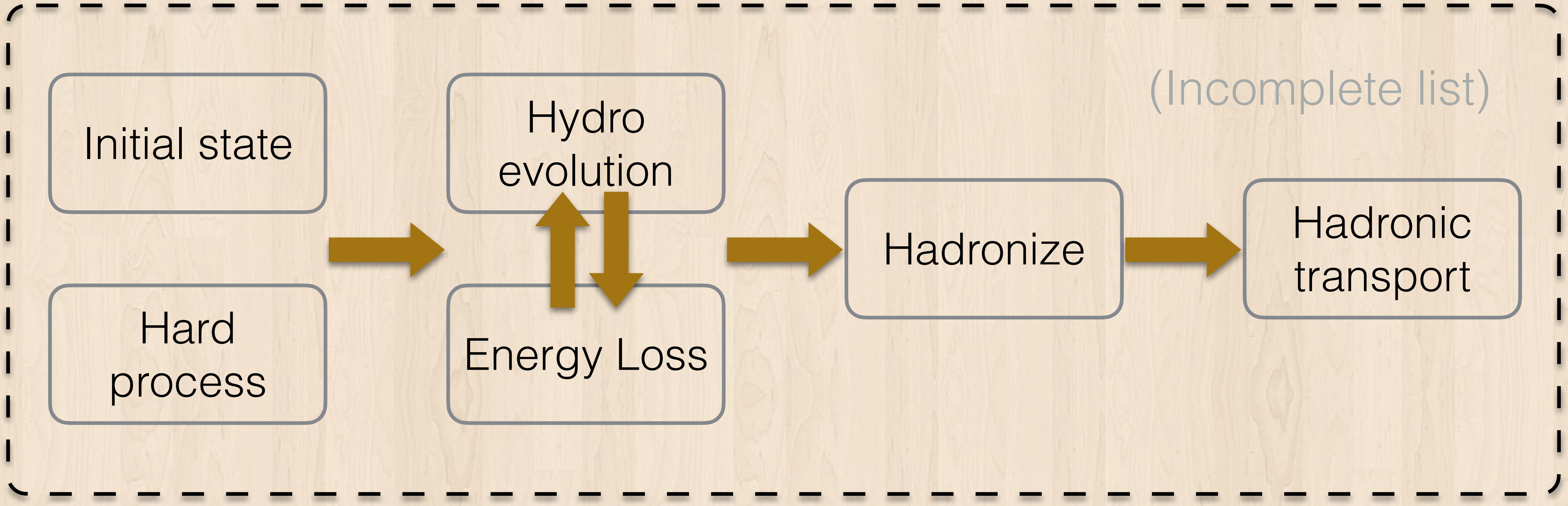


A LOT of parameters needed to specify the whole thing

Both in each block and the interface between blocks

Usually **different code bases**

Code framework



JETSCAPE framework:

- Modular design
- Unified block interface
- Easily extensible
- Easy to run (Docker image, etc)

The Analysis

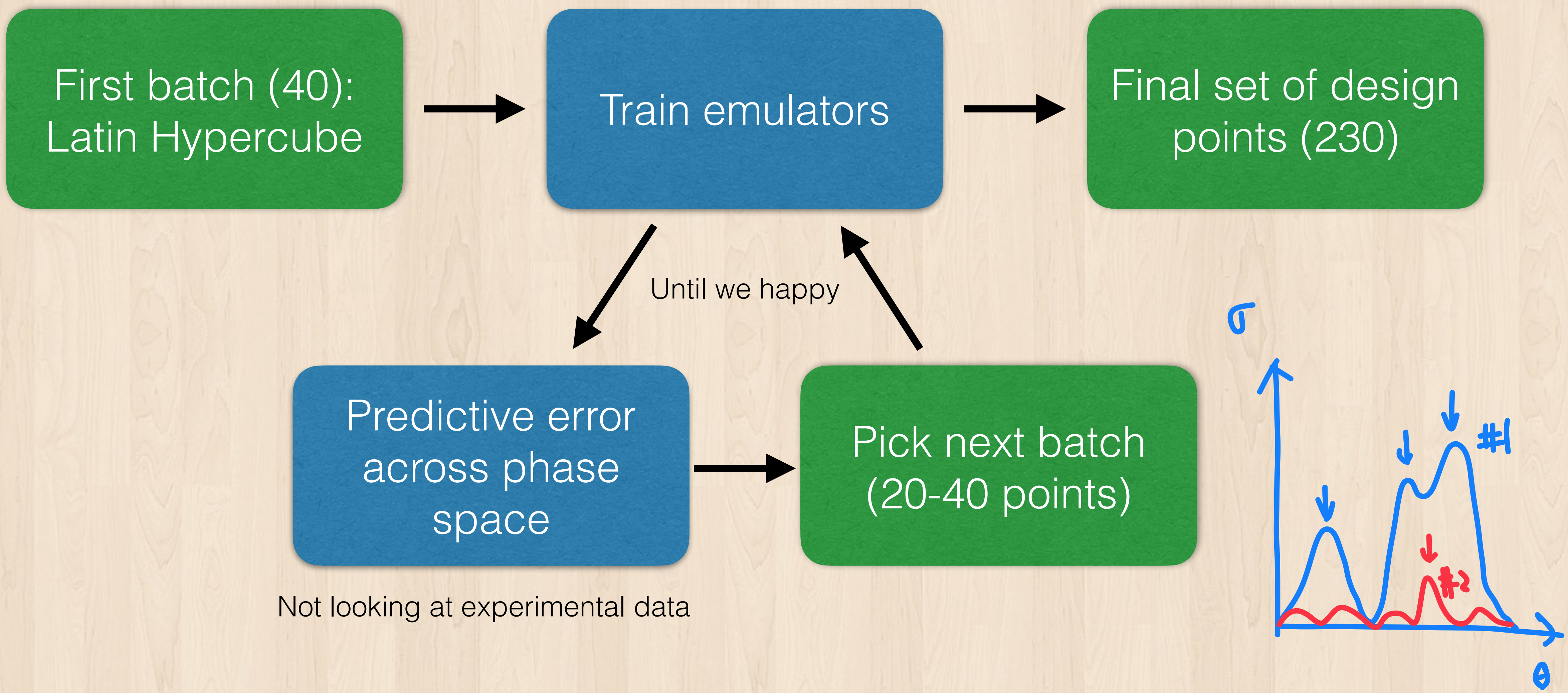
The problem we want to solve

- **Extract $\hat{q}$**
- Look at **jet and hadron suppression** data
    - Particles go through QGP and lose energy
    - Amount of suppression → amount of interaction
    - Amount of interaction → $\hat{q}$

Choice of datasets

- We adopt an agnostic approach: **all qualified dataset** by a cutoff time (Feb 2022) are included
  - "Qualified" = right category and in target phase space and possible to compare rigorously
- Different collision systems (AuAu, PbPb) across three CM energies (200 GeV, 2.76 TeV, 5.02 TeV)
- In total **729 data points** used, jump up from previous iteration of analysis of similar nature
- We do our best to **reproduce covariance matrix** (more later)
  - Reported uncertainty sources + guesses from the rest

Active learning design points



Prioritize reducing predictive error across the full space

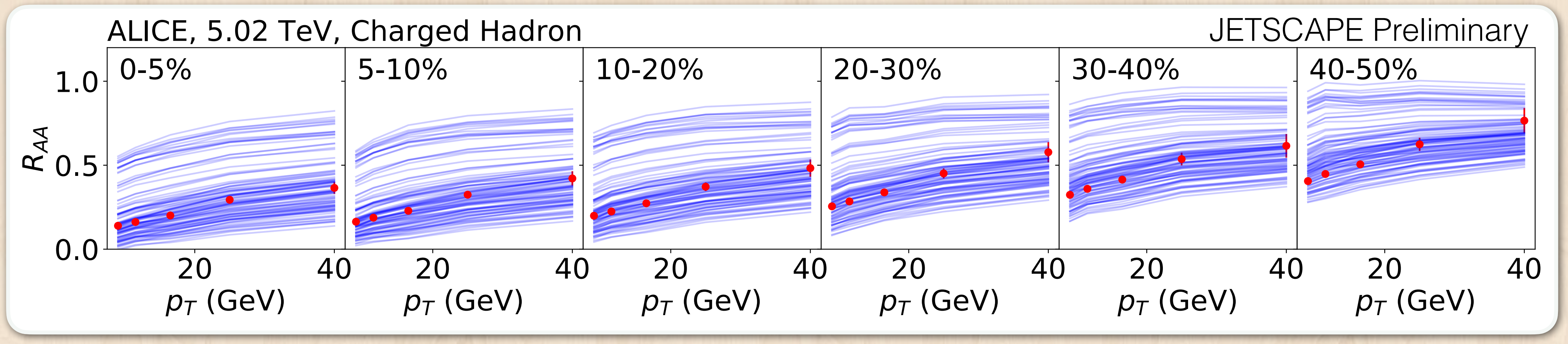
Computing resources

- Effort in computing during 2022
  - O(10M) CPU hours in total
  - Lots of lessons learned — unified submission interface across multiple HPC systems, data curation including all systematic uncertainties, iteration on design points, file I/O logistics, etc.
- Calculated many more observables than are used in this iteration → fast turnaround for next analyses

So we run the analysis…

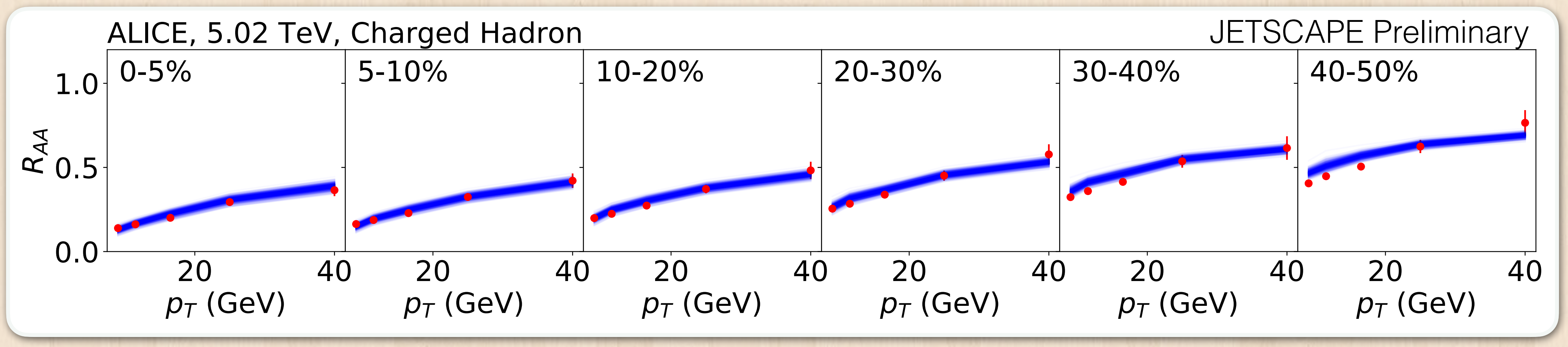
R_{AA} ~ amount of suppression

Data
Calculation



Analysis

Data
Best Fit



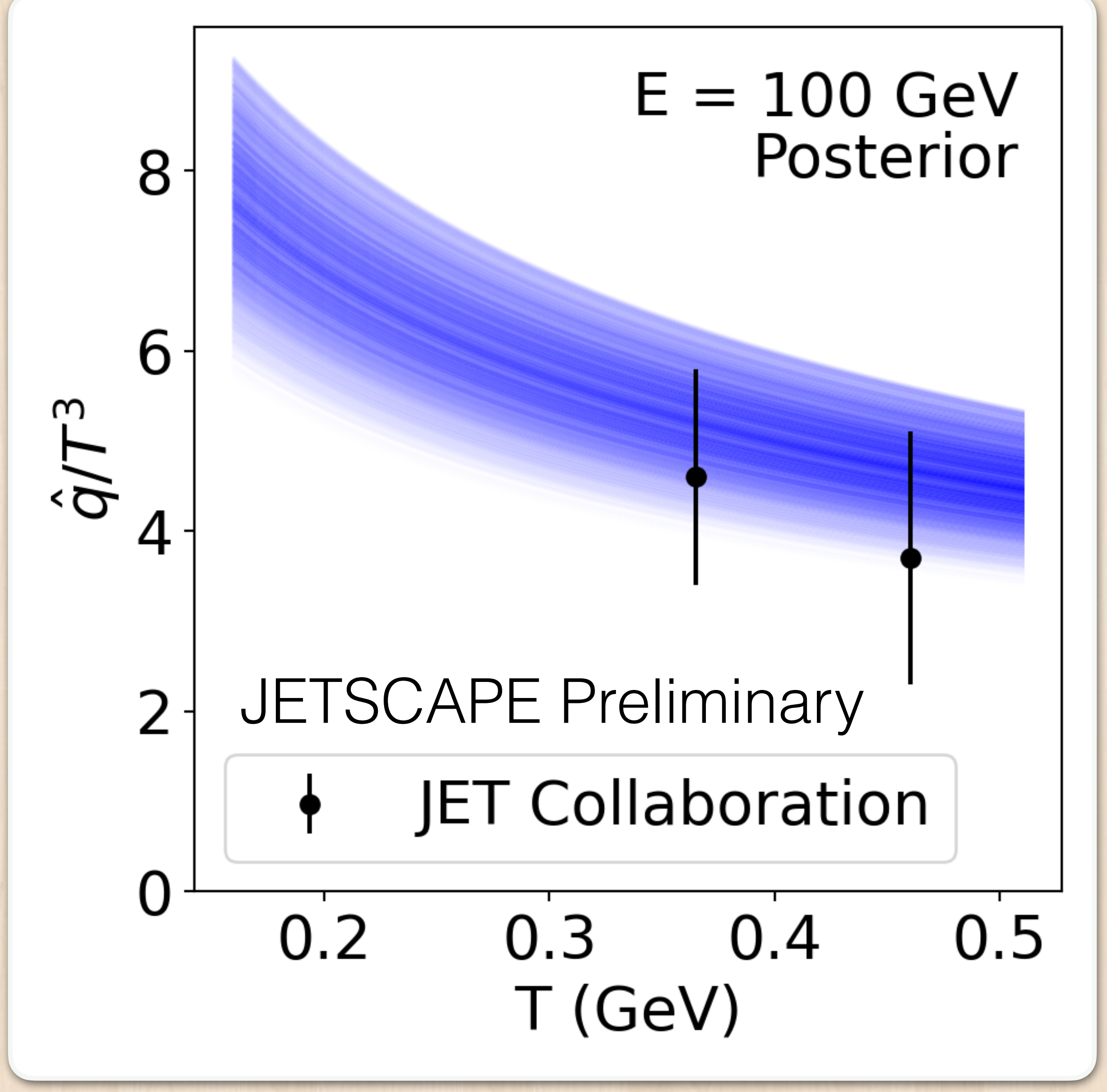
arXiv 2407.XXXX

Extracted $\hat{q}$

Compatible with JET collaboration result

All good?

Let's look closer…

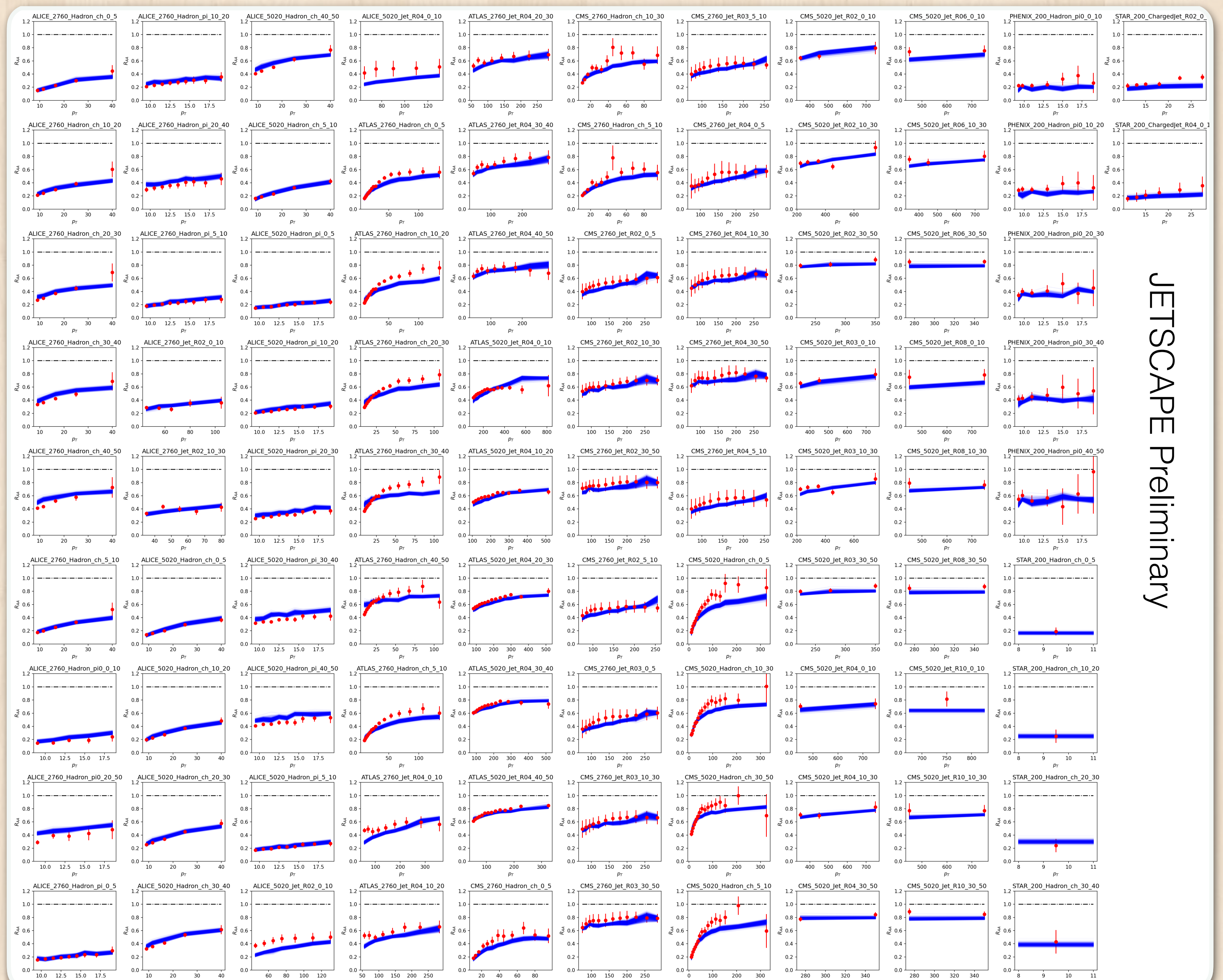


arXiv 2407.XXXX

Phys. Rev. C 90, 014909 (2014)

Posterior observables

(Don't stare too closely, we have zoomed in version in the next pages)



Data

Best fit

Overall reasonable agreement is observed

Tension for some measurements?

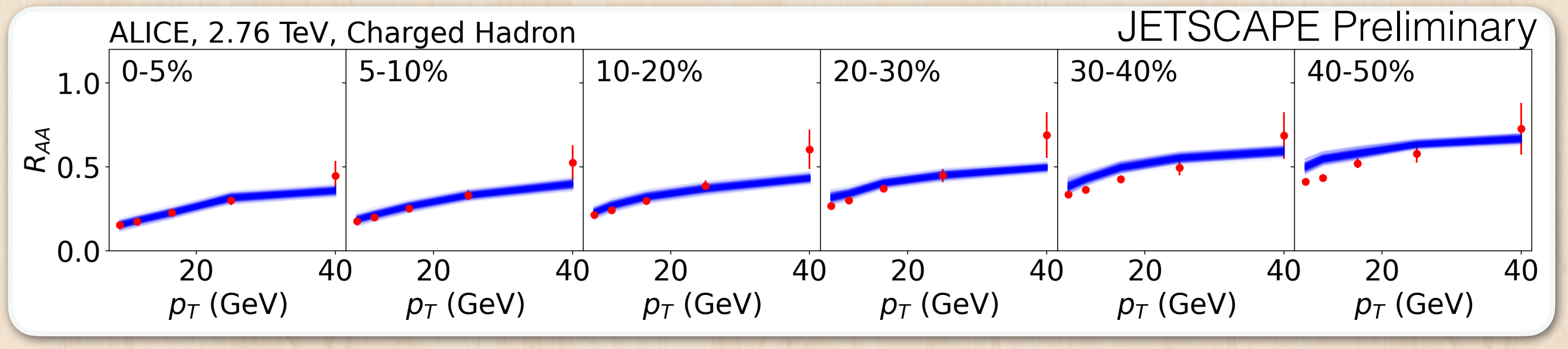
arXiv 2407.XXXX

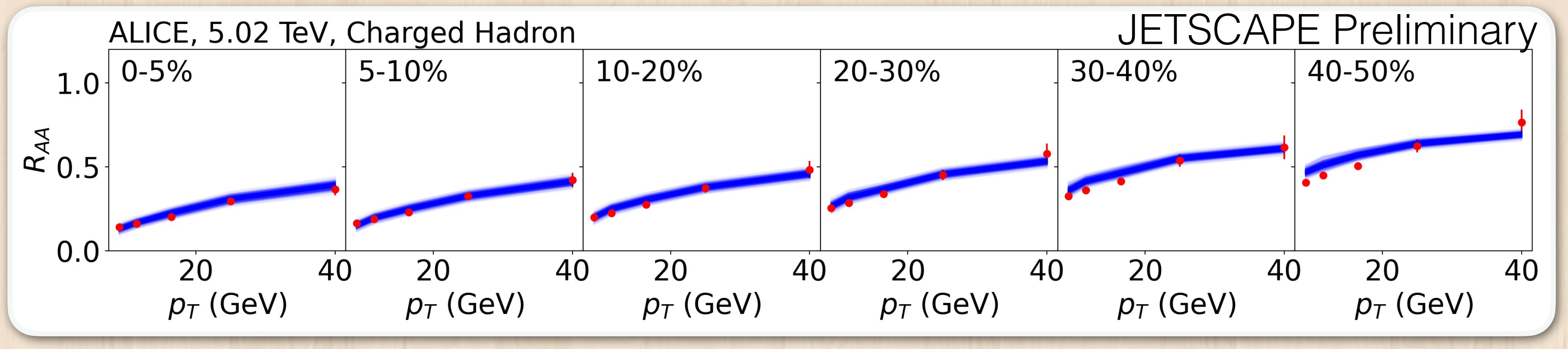
Looking closer — hadrons

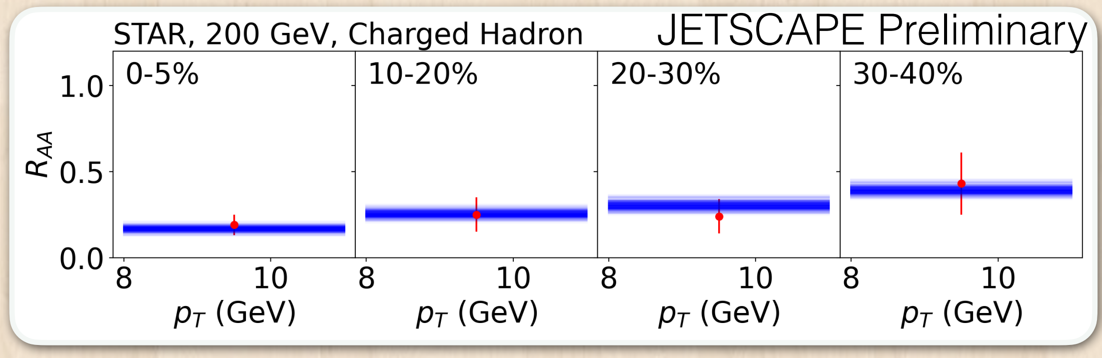
Generally great agreement at lower p_T

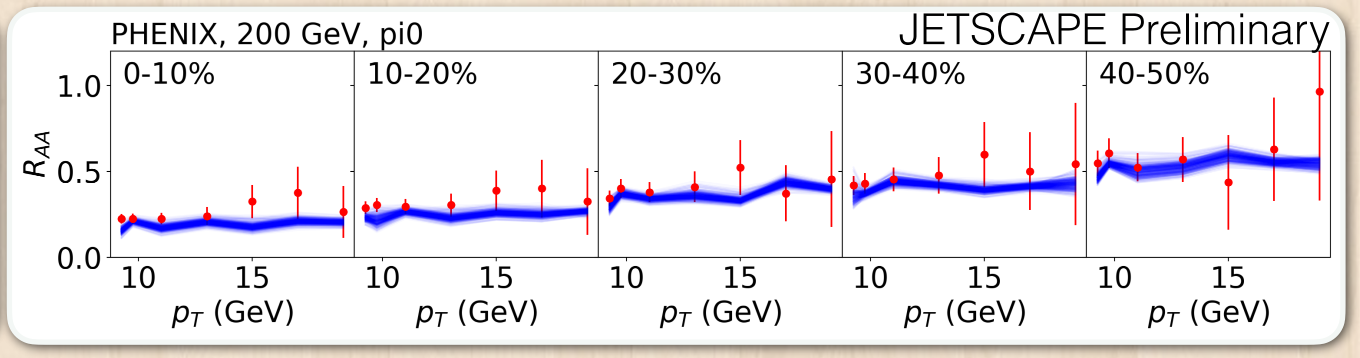
No large difference across experiments

R_{AA} ~ amount of suppression



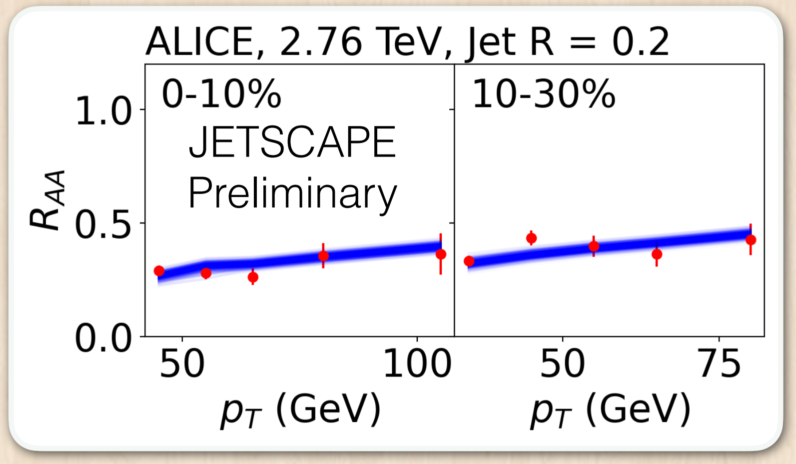




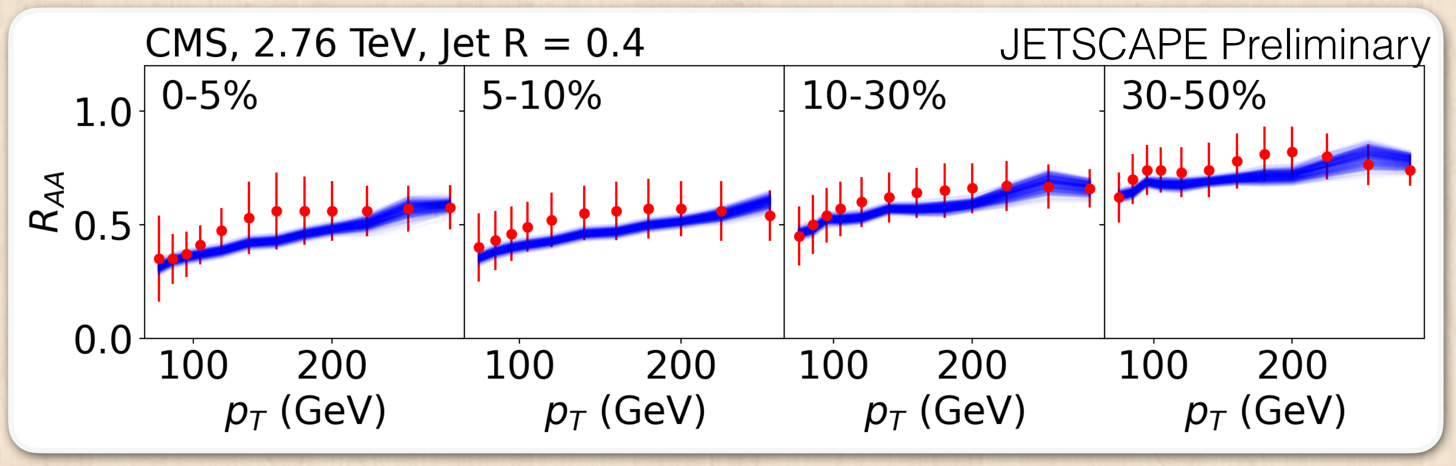


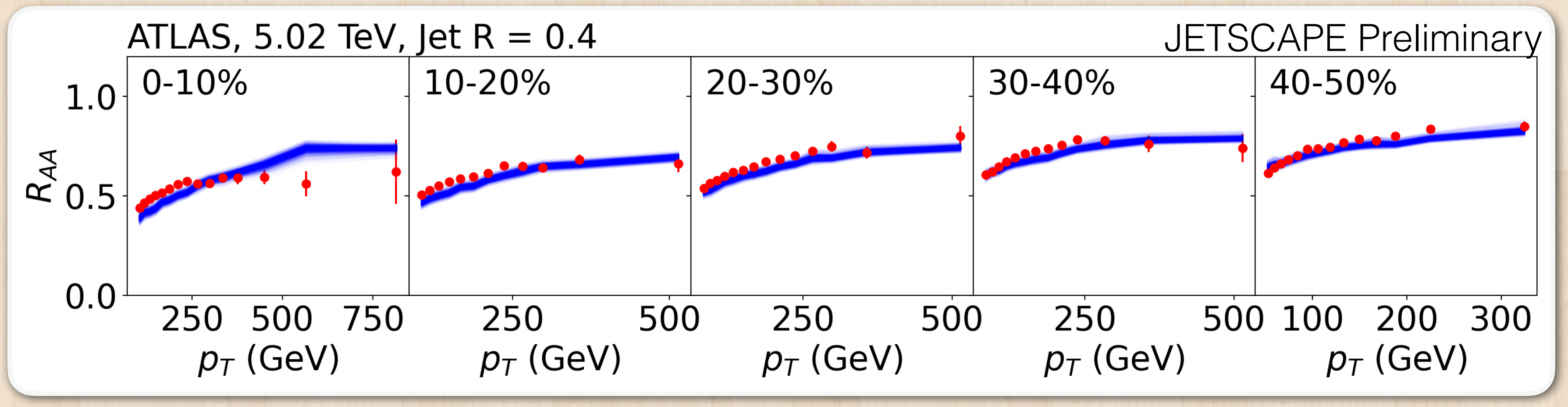
arXiv 2407.XXXX

Looking closer — jets



Also generally good agreement





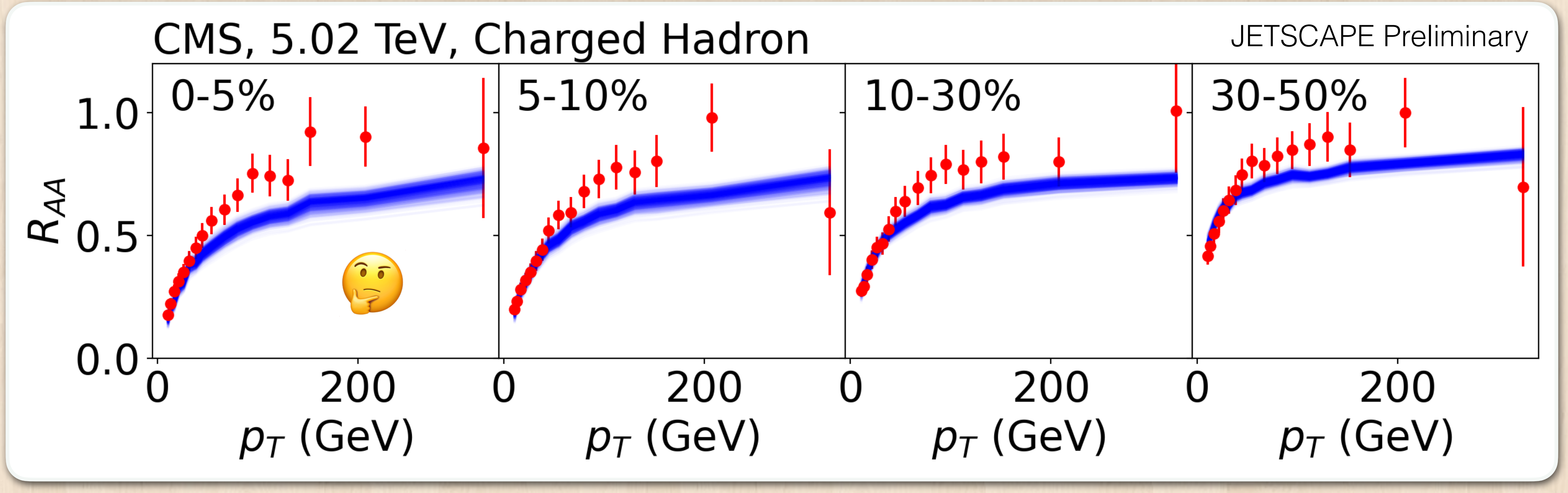
Hmm?

arXiv 2407.XXXX

Looking even closer — hadrons

Things deviate a bit going to higher p_T

Uncertainty smallest at lower $p_T \rightarrow$ drives result



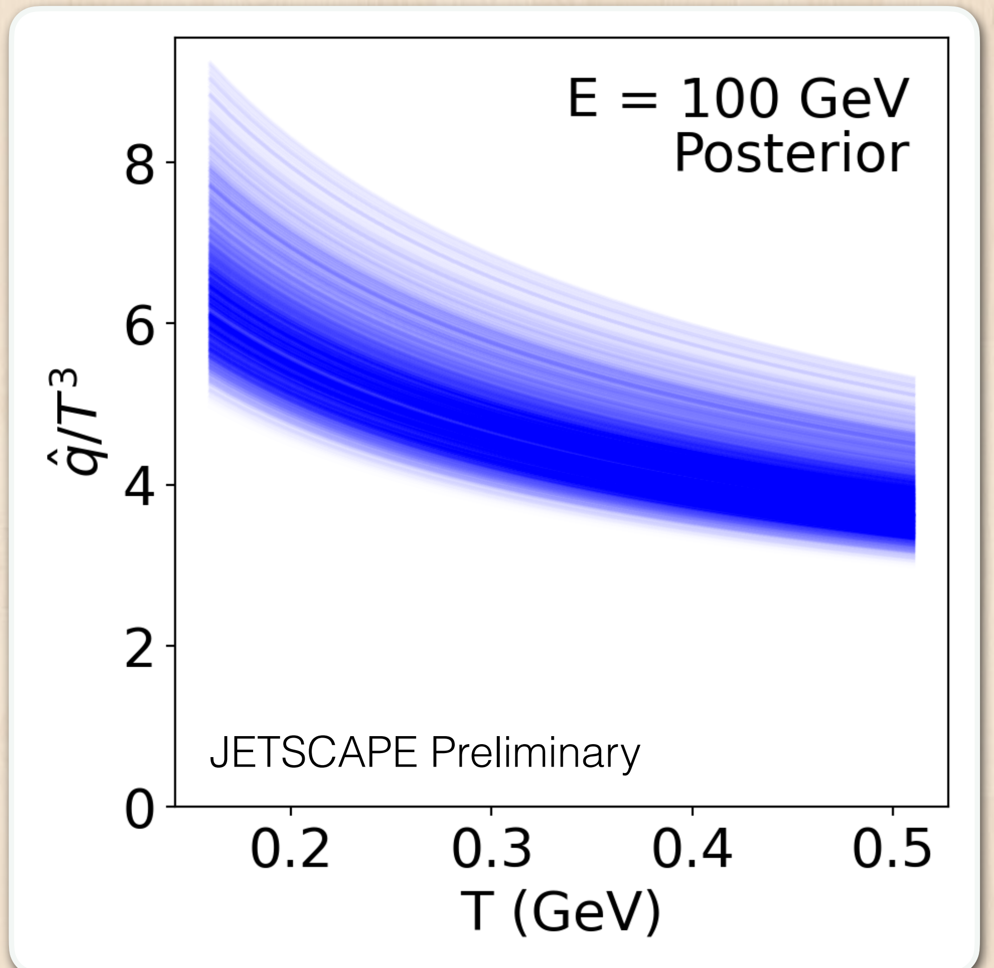
How can we gain more insight?

How can we gain more insight?

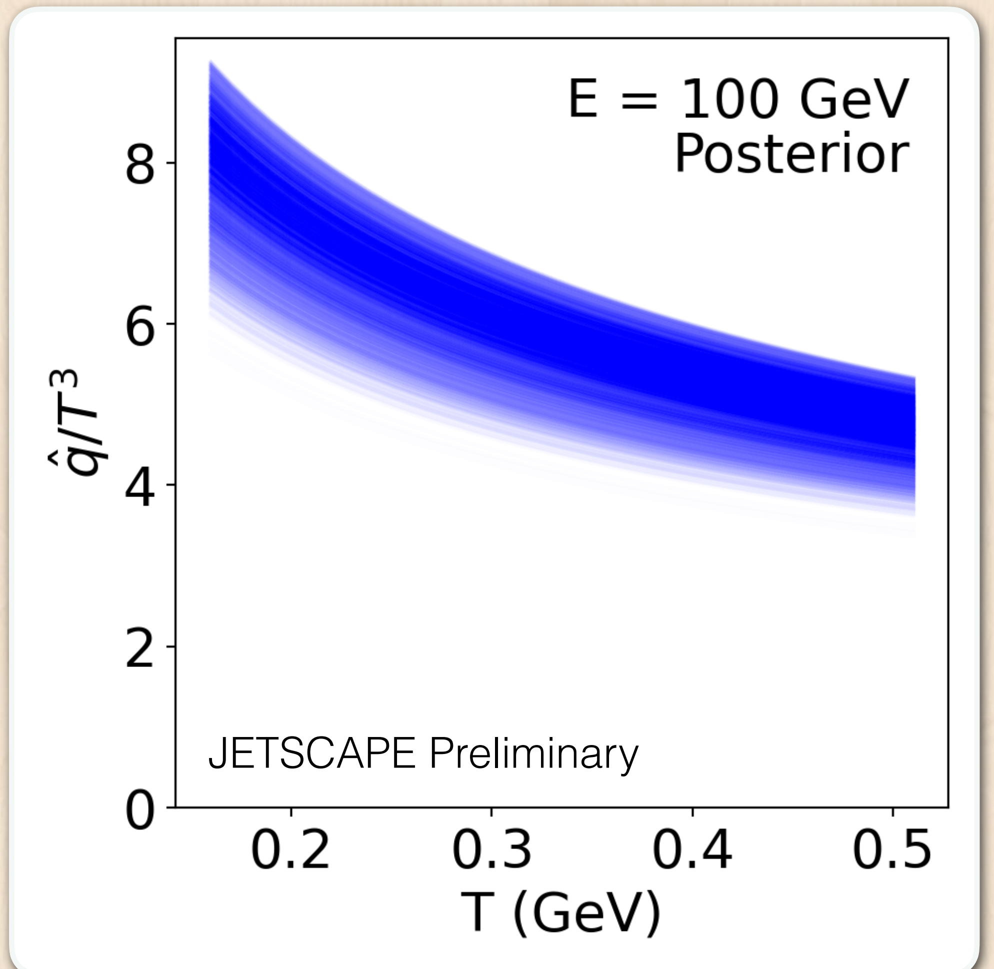
Idea: slice and dice datasets

- **Split datasets in different ways** and perform Bayesian analyses on subsets of data
- Investigate if there is any **systematic problem** with modeling
    - Similar measurements from multiple experiments useful

$\hat{q}$: jets vs hadrons



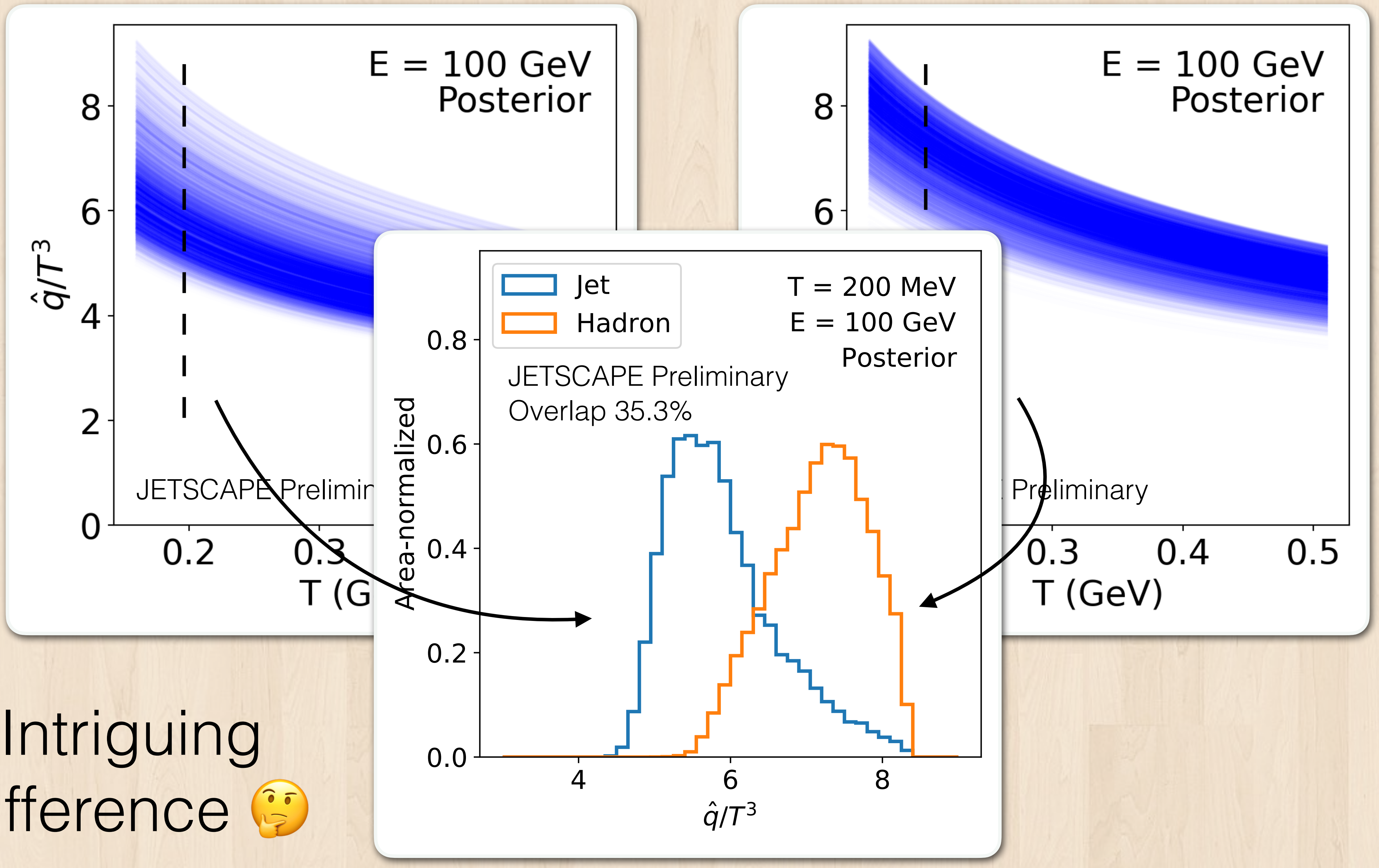
If we do analysis with only jet data



If we do analysis with only hadron data

arXiv 2407.XXXX

$\hat{q}$: jets vs hadrons



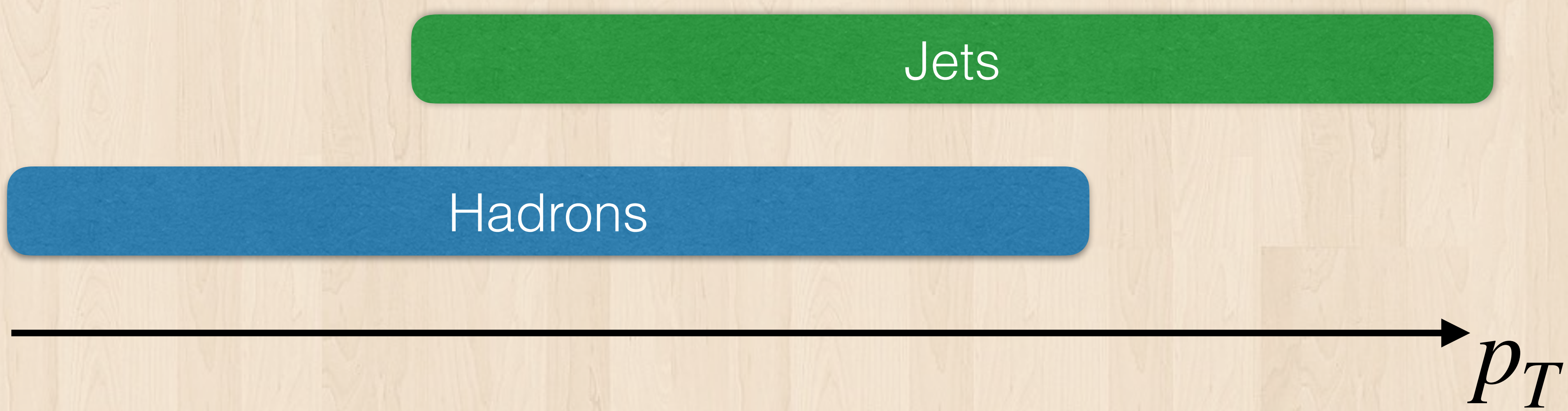
Intriguing difference 🤔

Kinematic ranges

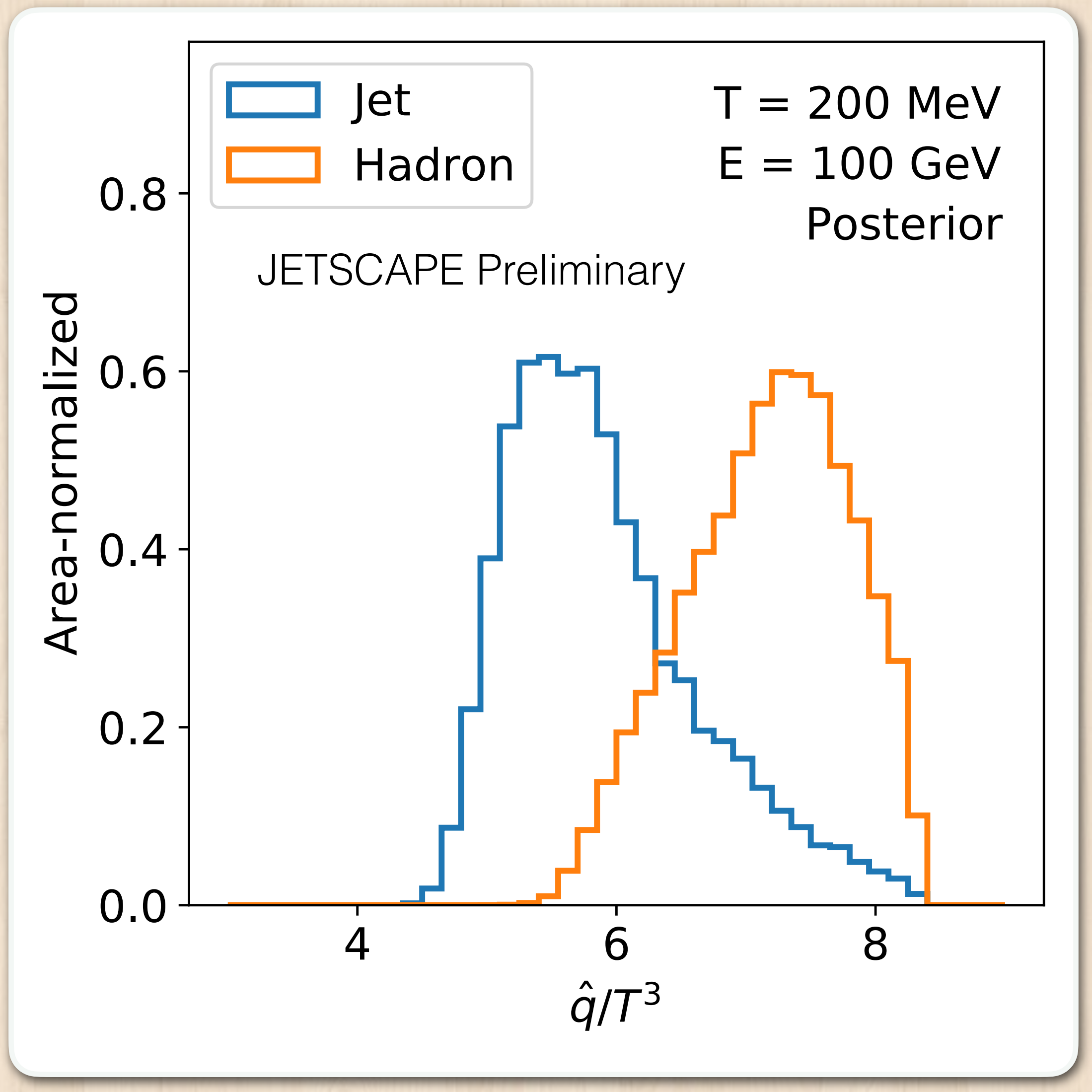
Is the difference we see inherent in the type of observables, or some other sources?

$\hat{q}$ supposedly should be same across observables?

One potential candidate: kinematic range

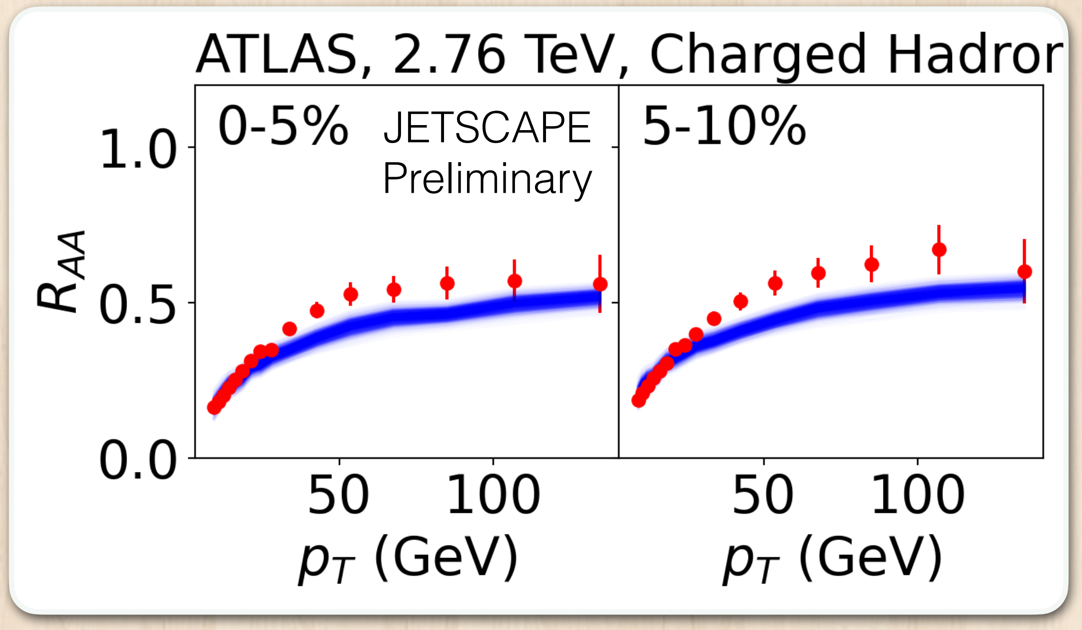


Hadrons, high vs low



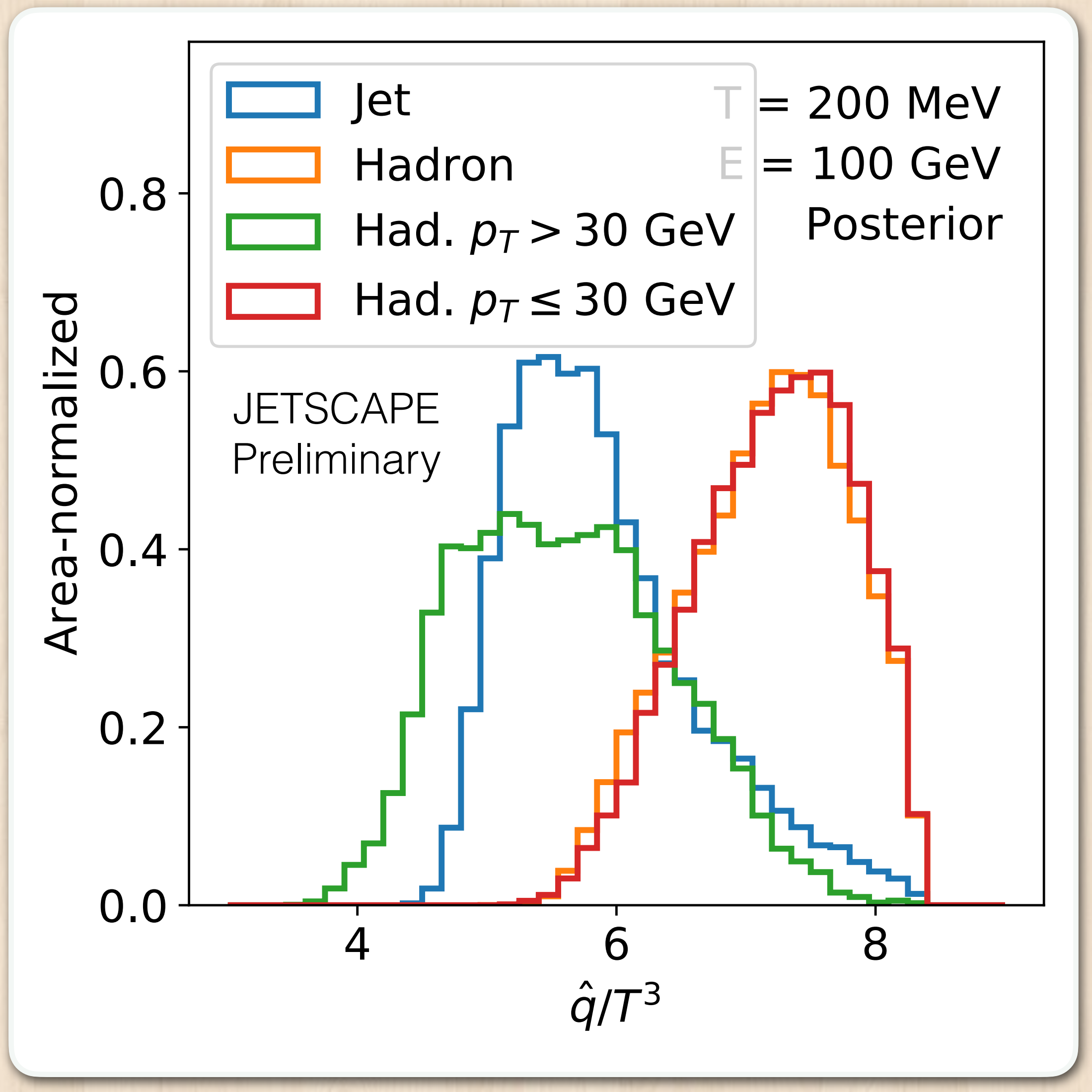
Full p_T range

R_{AA} ~ amount of suppression



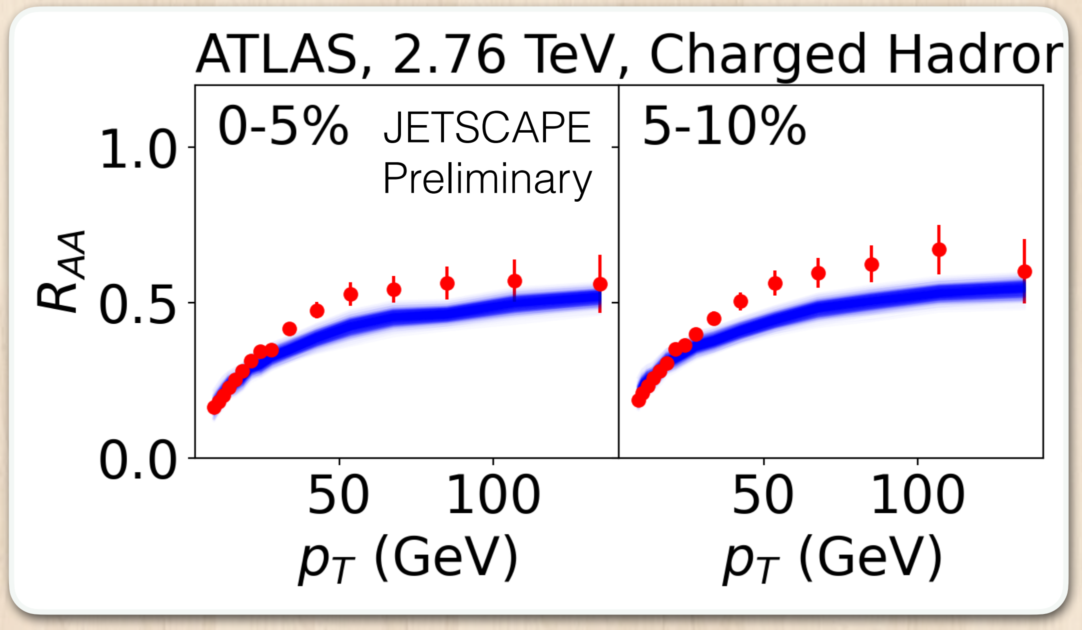
arXiv 2407.XXXX

Hadrons, high vs low

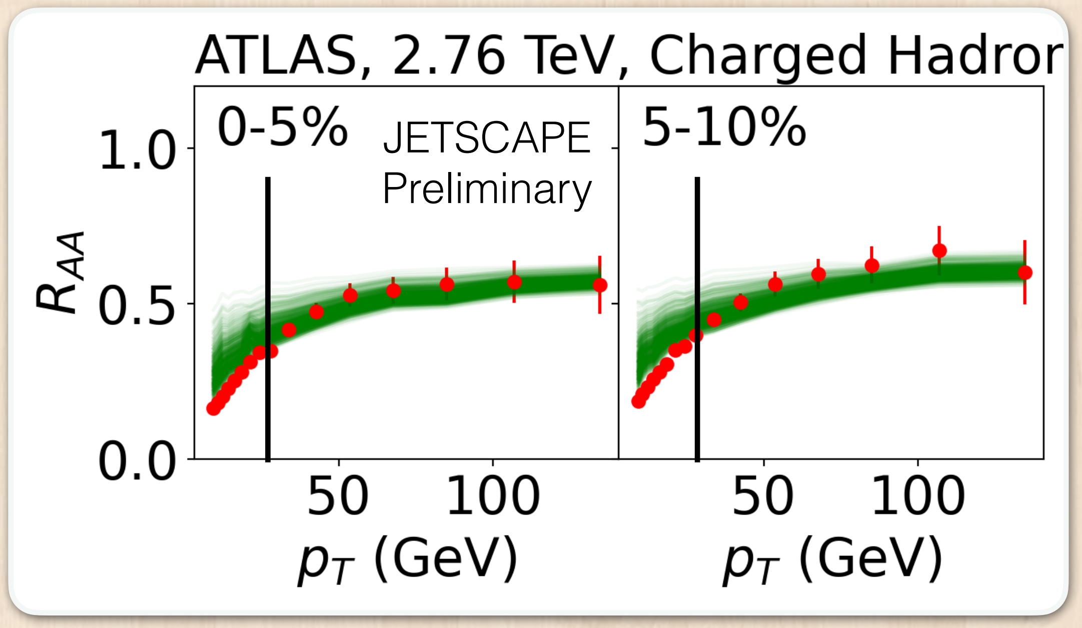


Full p_T range

R_{AA} ~ amount of suppression



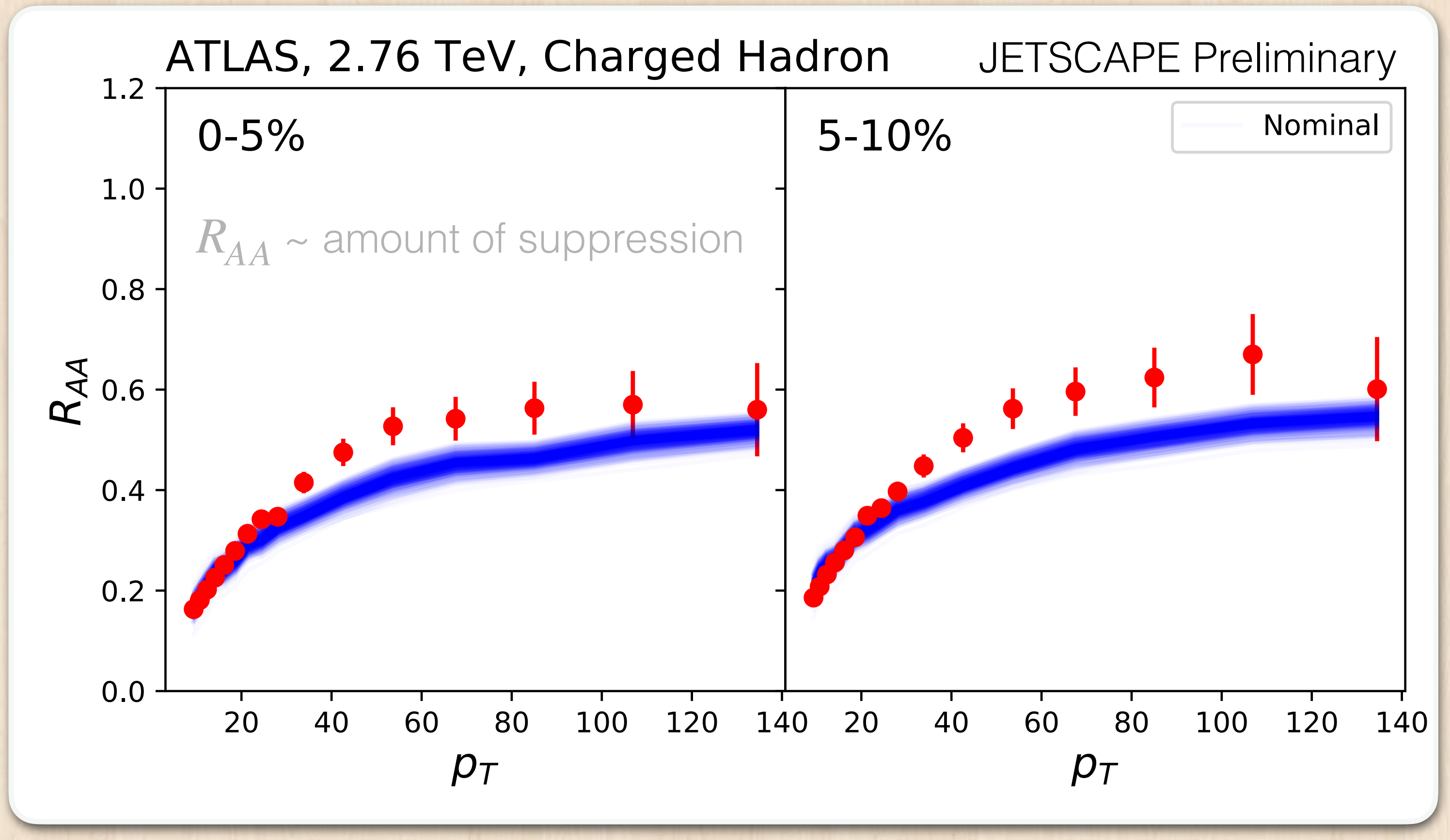
Only hadrons $p_T > 30$ GeV

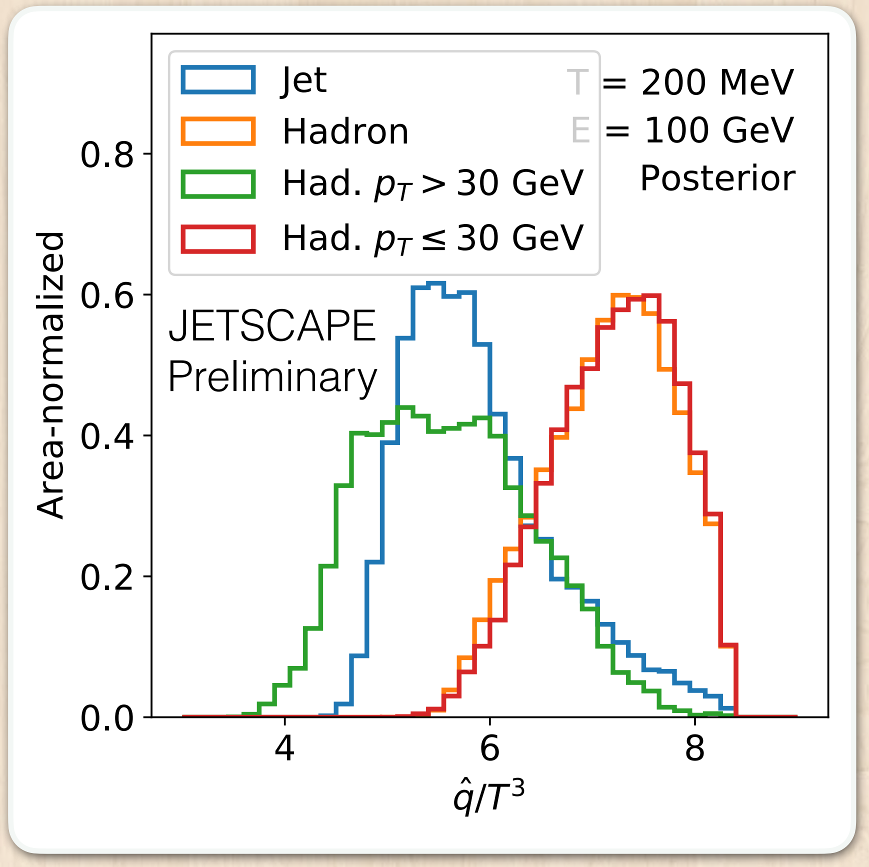


By eye green seems “better”!

arXiv 2407.XXXX

So what happened?





Low p_T part dominates: small experimental uncertainty

High p_T part in line with jet data

Points clearly to phase space for model improvement

→ question of “model uncertainty”

arXiv 2407.XXXX

Implications

- We can **scrutinize** the specific **model** used in this round of simulations in great detail
  - Low vs high p_T, central vs peripheral, jet vs hadron, different radii jet, and so on
  - Future: would be nice to do this with **more models**
- Isolate regions of interest
- Important feedback to models
- Points to interesting question: **model uncertainties**?

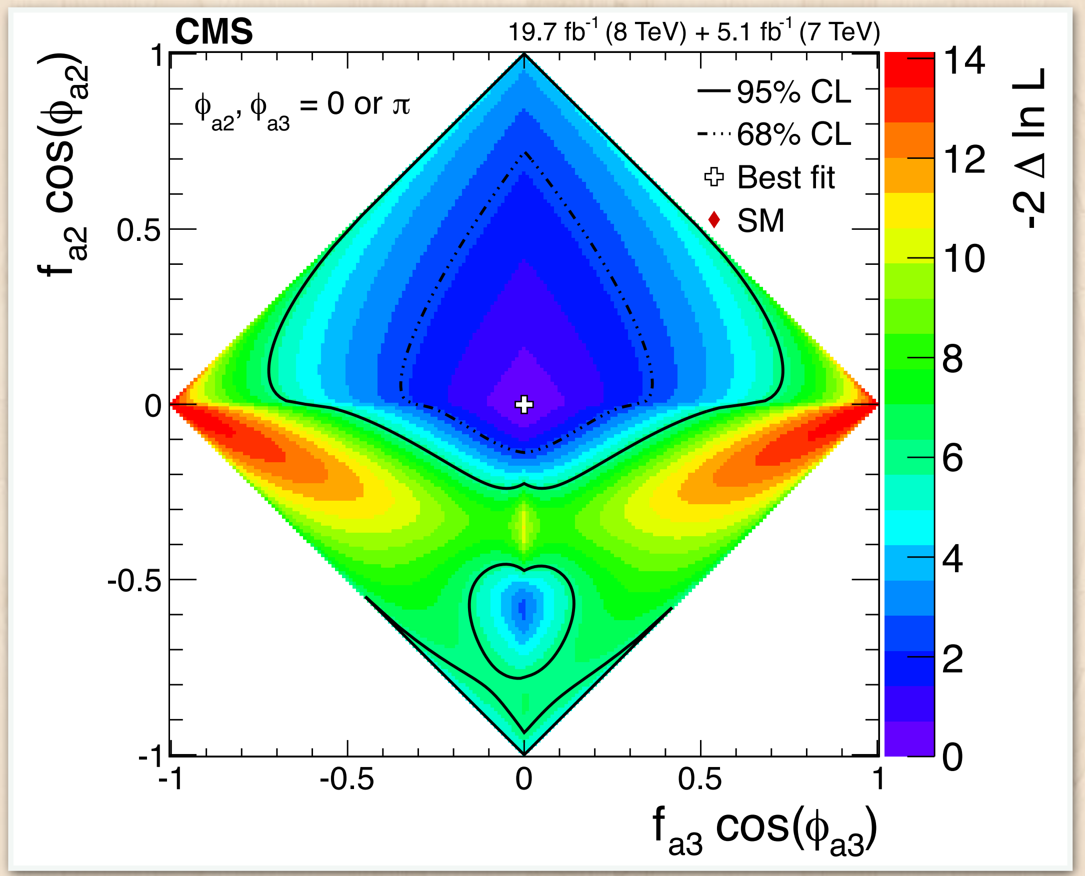
Discussions: Uncertainties

What does uncertainty mean?

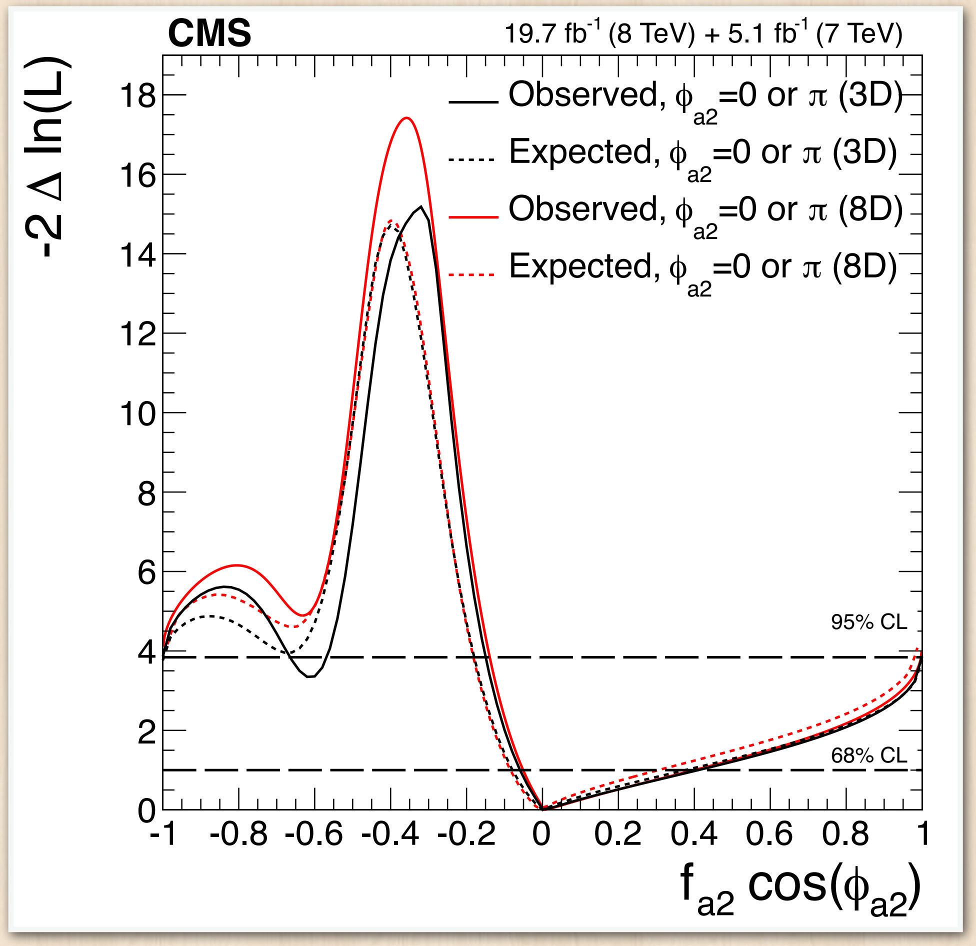
- In experiments there are always some distributions behind the scene (likelihoods, Bayesian posterior, etc)
  - They tell you something about the “true” value
- “Uncertainty” is then some sort of width or range quoted from these distributions
  - There is no universal prescription from measurement to measurement (especially systematics)

Example from Higgs measurement

$$\mathscr{L} \sim c\left(HZ_\mu Z^\mu + a_2 HZ_{\mu\nu}Z^{\mu\nu} + a_3 HZ_{\mu\nu}\tilde{Z}^{\mu\nu}\right) + \ldots$$



Size of CP-odd HZZ term



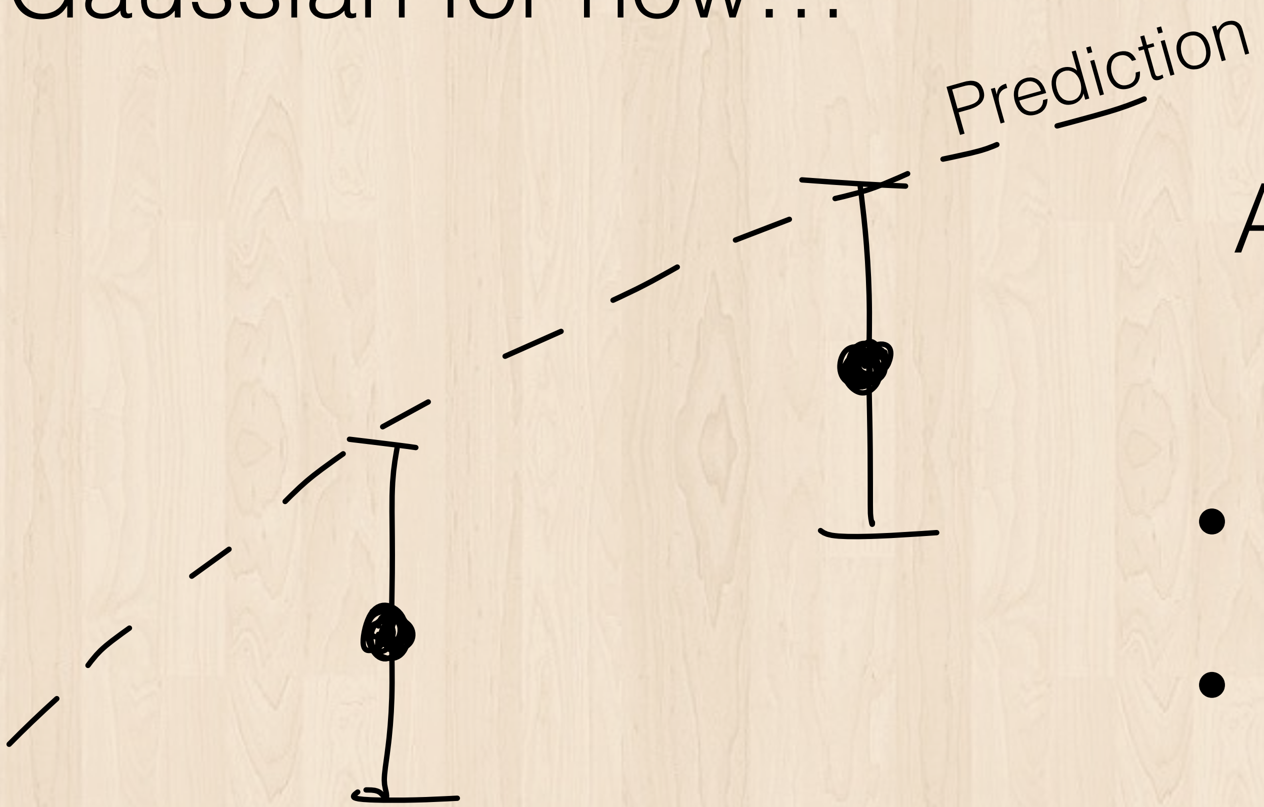
Size of higher order CP-even HZZ term

Side note: this is an inverse problem with an interesting non-Bayesian approach

Phys.Rev. D92 (2015) 012004

Data uncertainty correlation

Suppose everything is Gaussian for now…



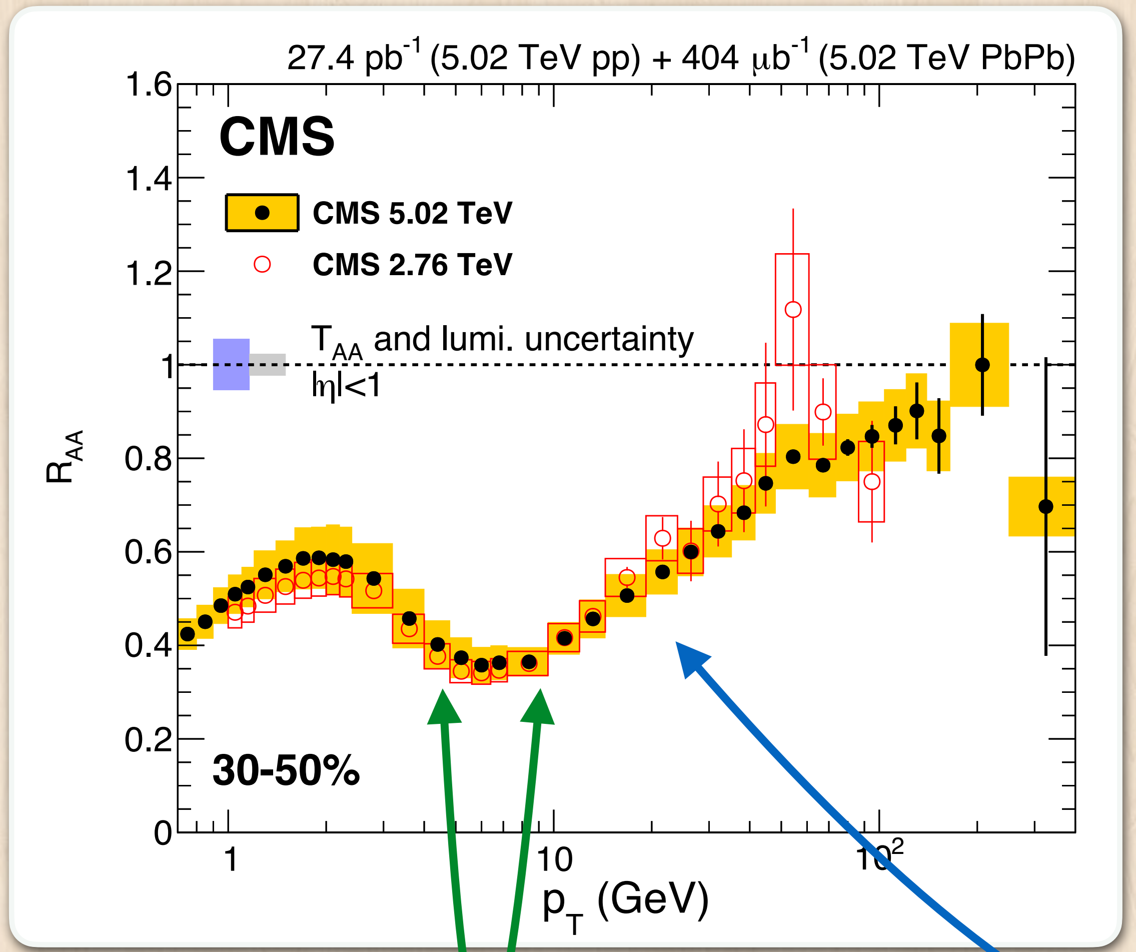
Correlation is key!

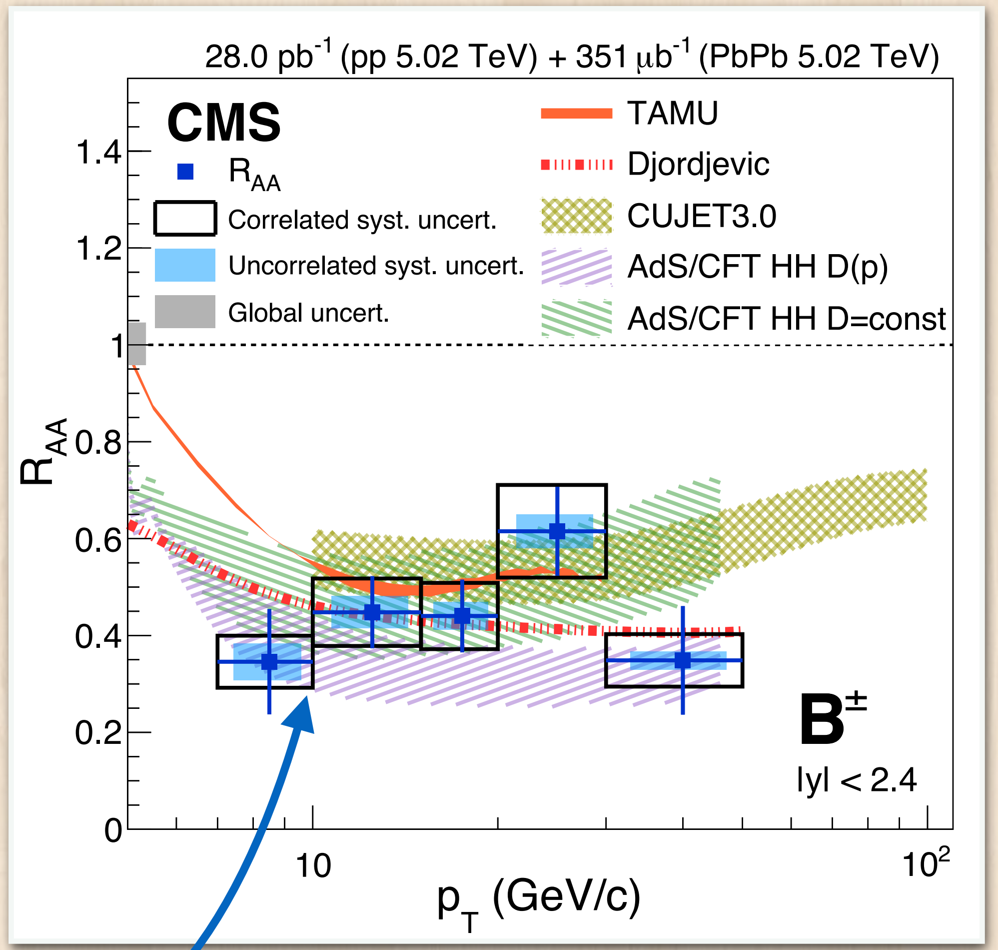
Agreement depends on uncertainty correlation

- Fully Correlated: "1σ"
- Non-correlated: "2σ"
- Anti-correlated: ">2σ"

Faithfully capturing the correlation is **crucial**

Different types of correlations





Correlations within measurement

Correlations across measurements

Most are *not* reported by experiments

JHEP 04 (2017) 039

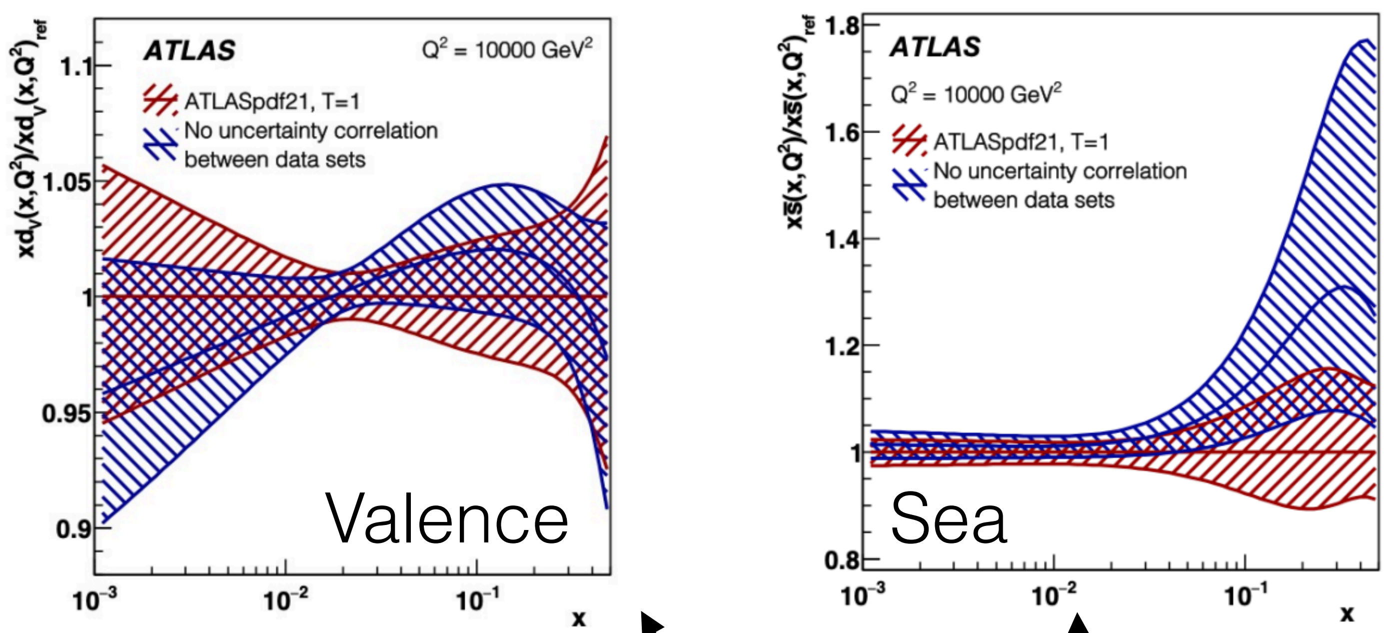
Phys. Rev. Lett. 119 (2017) 152301

This is parton distribution function

Cautionary tale from PDF analysis

Effect of correlation **across measurements**

Impact of the Correlation Between Data Sets



When the correlations of the systematic uncertainties between V+jets, ttbar, inclusive jets are not applied, substantial difference wrt the nominal PDFs is observed at 10,000 GeV^2, a scale relevant for precision LHC physics

Ratio to nominal

ICHEP 2022, Bologna Italy, July 6-13, 2022 Zhiqing Zhang, IJCLab, Orsay 7/12

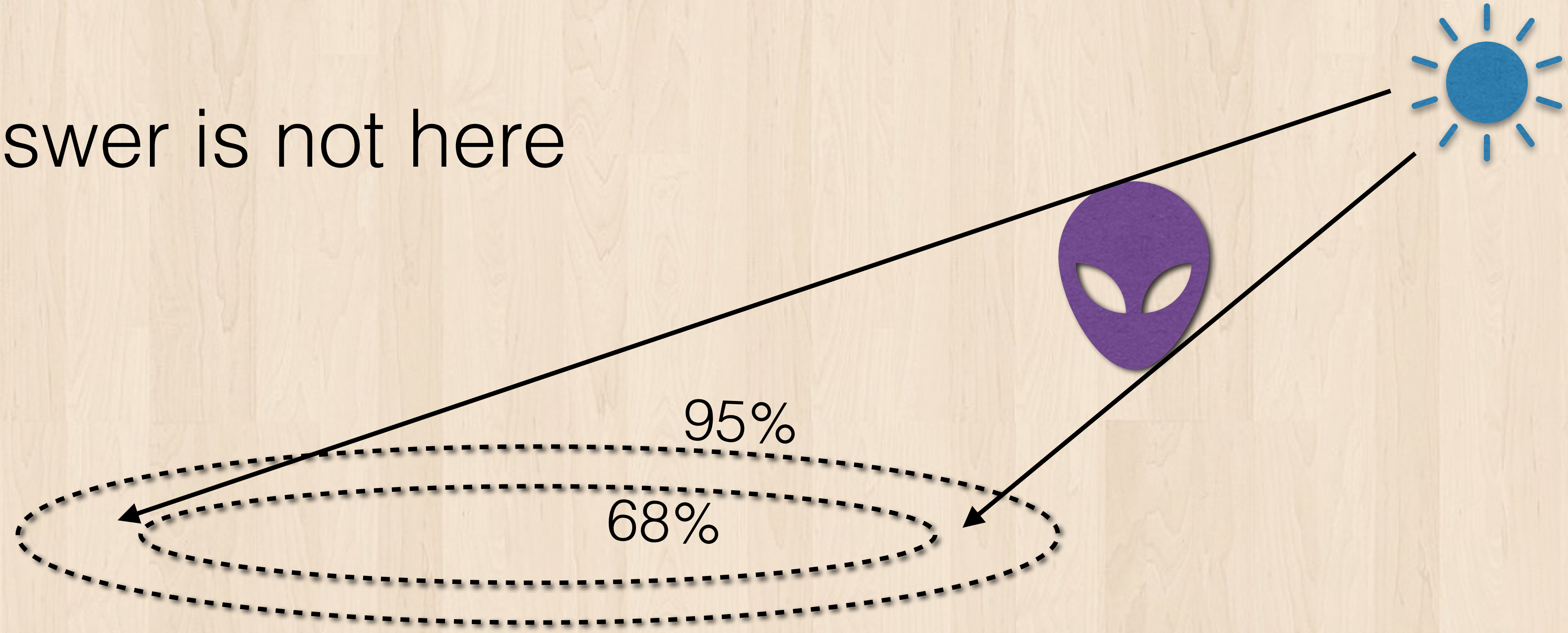
Significantly different depending on the correlation

What about model uncertainty?

- Ideally: some distribution that encodes where the "true" value should lie

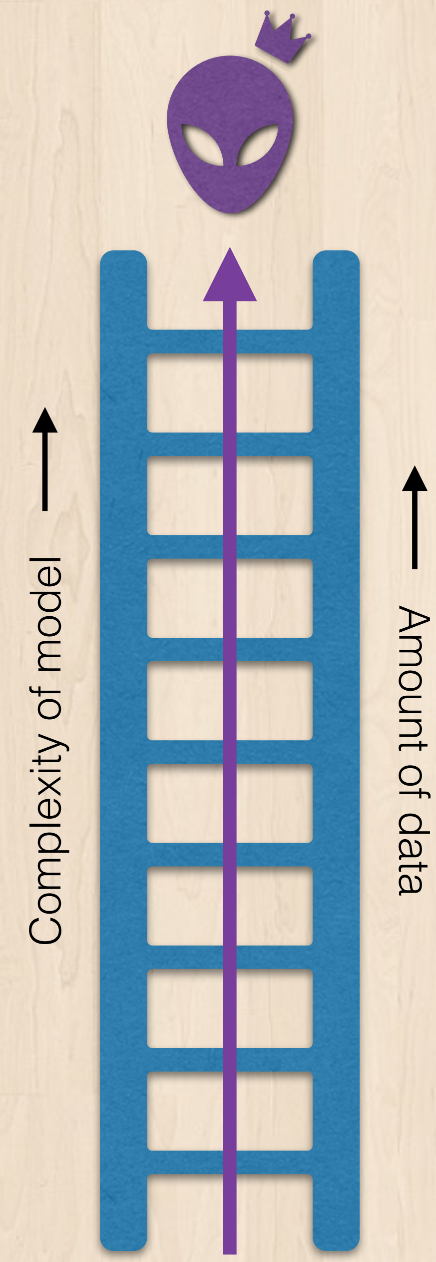
- Bayesian parameter extraction: best parameters within a predefined model space

  - Full answer is not here



Improving model alongside data

- Constant **improvements to the model** needed to get closer to the truth
- Even though we used many measurements, there are many other potential measurement types to study
    - More information on **uncertainties** from experiments will be nice
- Lots of interesting things to explore



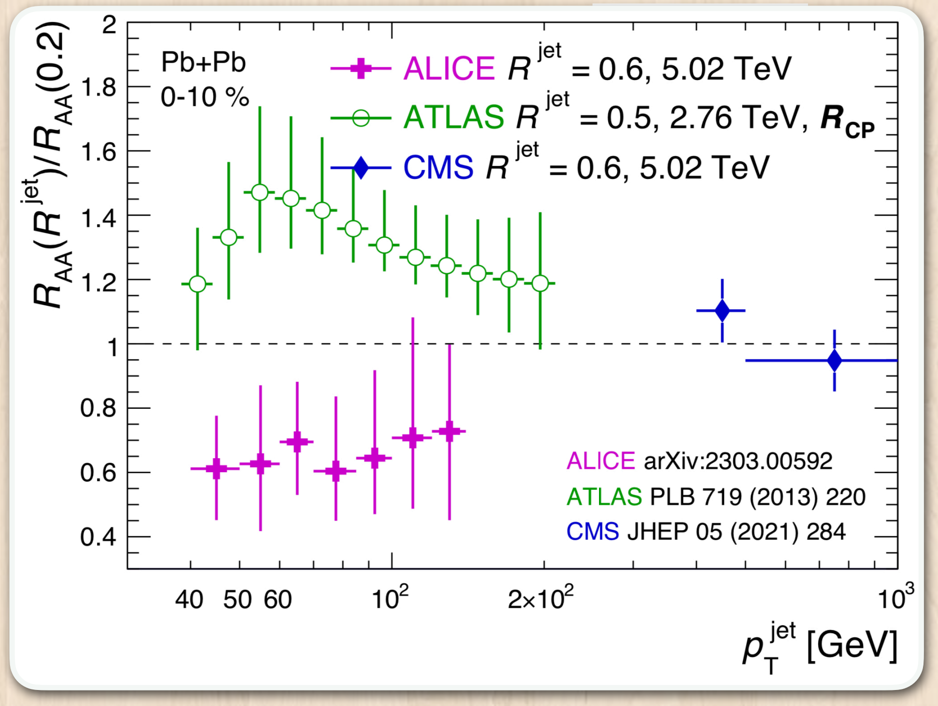
Summary

(Near-) future prospects

We also calculated huge number of **other jet-related observables**

Move one step at a time and **sequentially include more observables** → stay tuned for many new results in the near future!

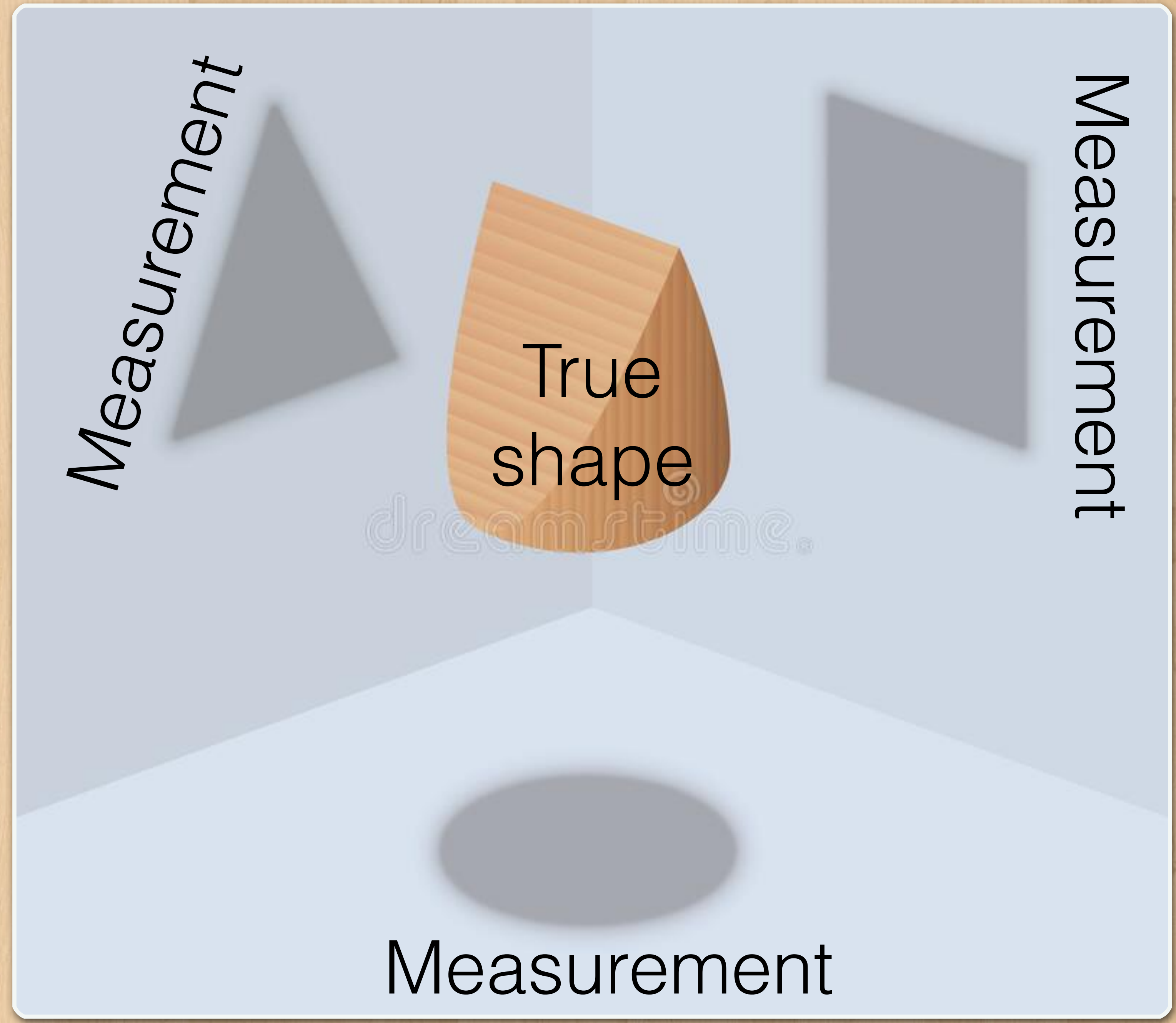
Plot taken from Y. Go, HP 2023



Explore the model + experimental landscape

Summary

- We performed an updated analysis on $\hat{q}$ **extraction** using a lot more data compared to previous iteration
- Bayesian analysis is useful as a **tool for model studies** → inform model design and improvement
  - A way forward to sort through the proliferation of models in high energy HI collisions
- **Experimental uncertainties**: we should advocate to experiment groups to release more information

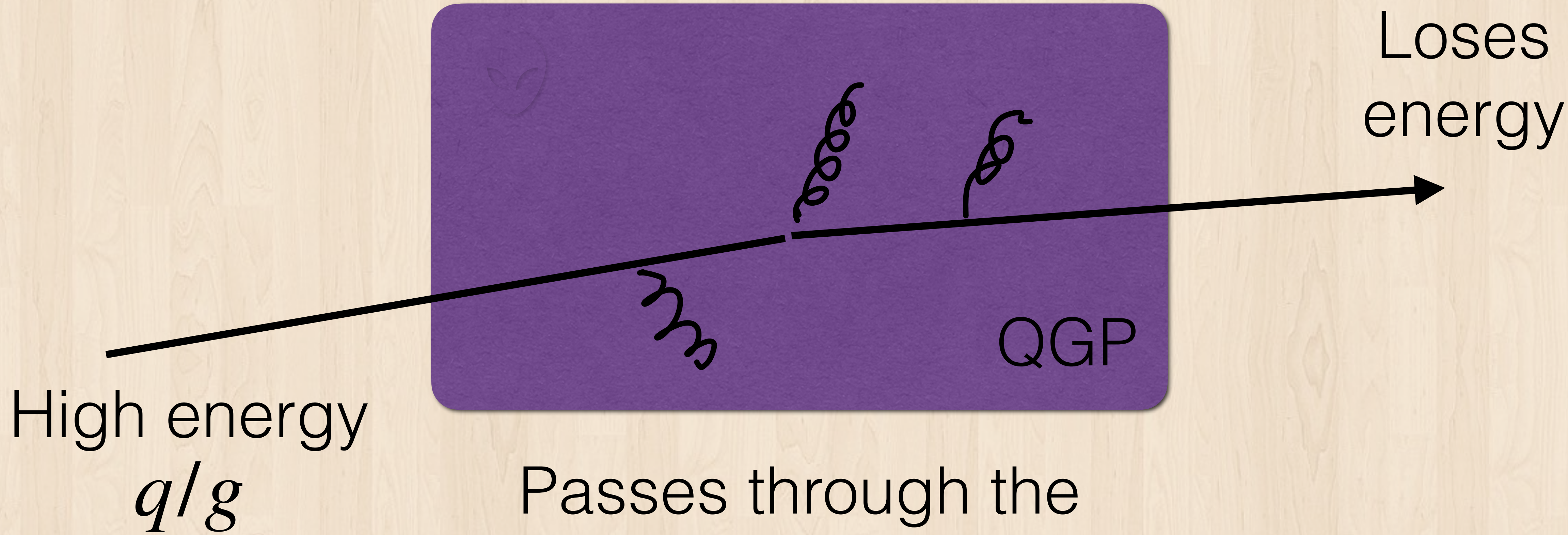

Measurement
True shape
Measurement
Measurement

Backup Slides Ahead

Computing resources

- Hydrodynamic evolution takes nontrivial time to run
  - We use pre-generated hydrodynamic profiles and propagate jets on top of them
  - But they are significant in size and we want to distribute to the computing nodes → logistics…
- Need O(20k-30k) core-hours per design point to match experimental precision

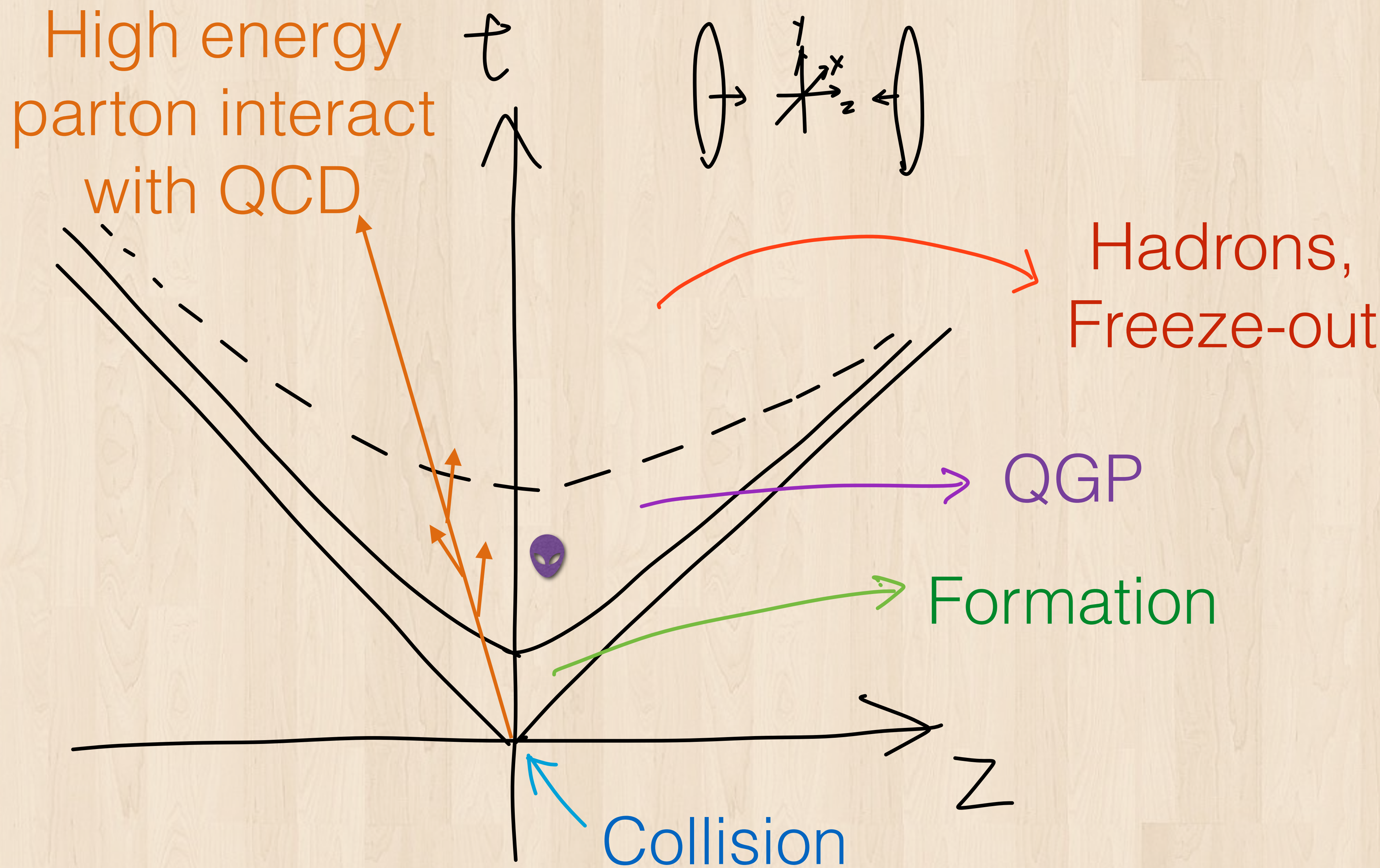
Jet quenching



Passes through the hot QCD fireball and interacts with it

We want to study the "strength" of this interaction

Simplified space-time diagram



How to extract parameter with this complex system?

Parametrization of $\hat{q}$

$$\hat{q}(E,T,Q) = \underline{\hat{q}^{run}_{HTL}} \times \underline{f(Q^2)}$$

$$\underline{\hat{q}^{run}_{HTL}} = \alpha_{s,fix} \times \alpha_s(\mu^2)\ C_a \frac{50.484}{\pi} T^3 \ln\left(\frac{\mu^2}{6\pi T^2 \alpha_{s,fix}}\right)$$

Inspired from exponential "PDF": $f_{QGP}(x) \sim e^{-c_3 x}$

$$\underline{f(Q^2)} \equiv N_0 \frac{\exp\left(c_3\left(1-\frac{Q^2}{2EM}\right)\right) - 1}{1+c_1 \ln(Q^2/\Lambda^2_{QCD})+c_2 \ln(Q^2/\Lambda^2_{QCD})}\Bigg|_{Q^2 \geq Q_0^2}$$

Set by $f(Q_0^2) = 1$

Other parameters
Q_0: virtuality switch to LBT
τ_0: start time

2204.01163, see also previous talk

Parametrization of $\hat{q}$

$$\hat{q}(E,T,Q) = \underline{\hat{q}^{run}_{HTL}} \times \underline{f(Q^2)}$$

$$\underline{\hat{q}^{run}_{HTL}} = \alpha_{s,fix} \times \alpha_s(\mu^2)\; C_a \frac{50.484}{\pi} T^3 \ln\left(\frac{\mu^2}{6\pi T^2 \alpha_{s,fix}}\right)$$

Inspired from exponential “PDF”: $f_{QGP}(x) \sim e^{-c_3 x}$

$$\underline{f(Q^2)} \equiv N_0 \left.\frac{\exp\left(c_3\left(1 - \frac{Q^2}{2EM}\right)\right) - 1}{1 + c_1 \ln(Q^2/\Lambda^2_{QCD}) + c_2 \ln(Q^2/\Lambda^2_{QCD})}\right|_{Q^2 \geq Q_0^2}$$

Set by $f(Q_0^2) = 1$

: parameters (6 in total)

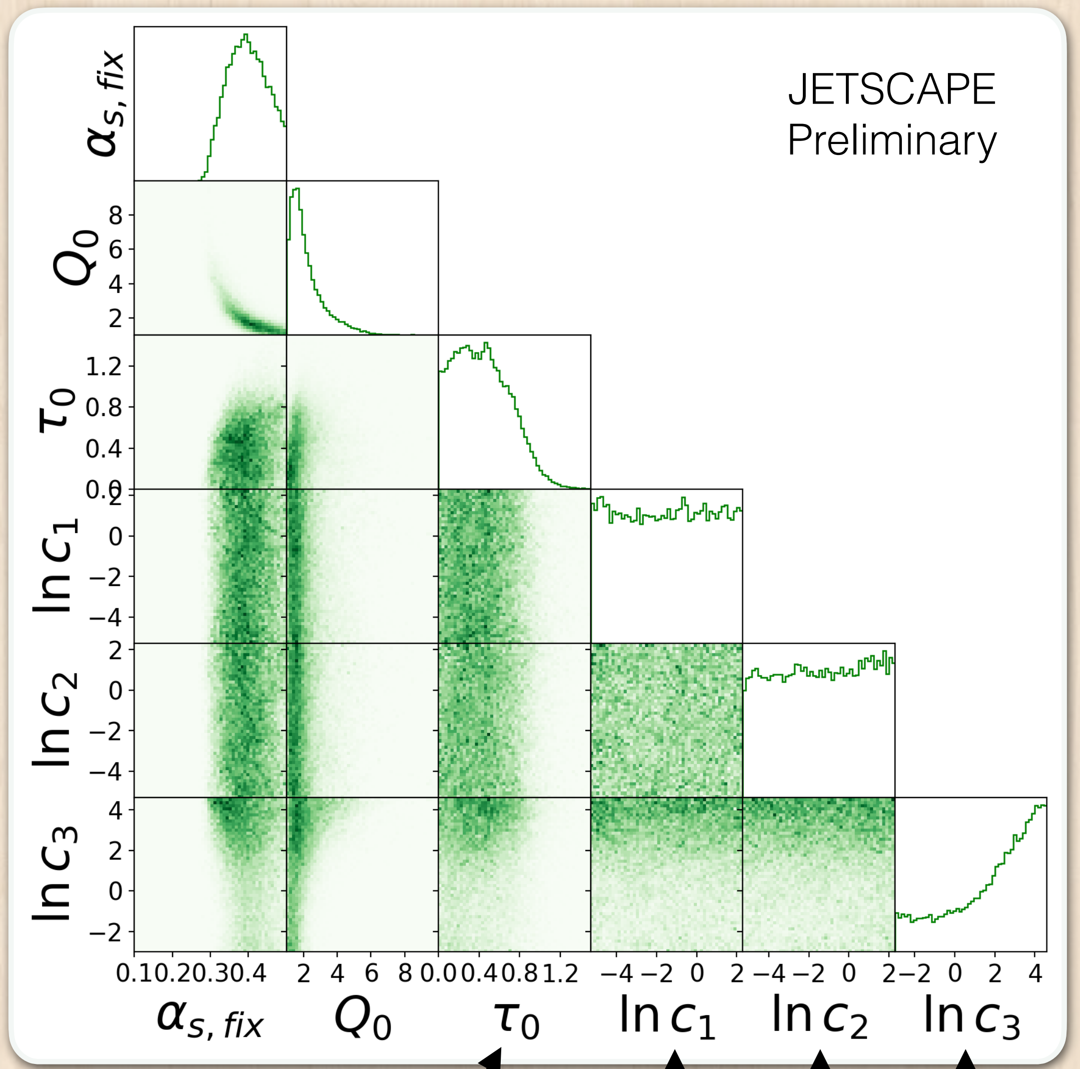
Other parameters

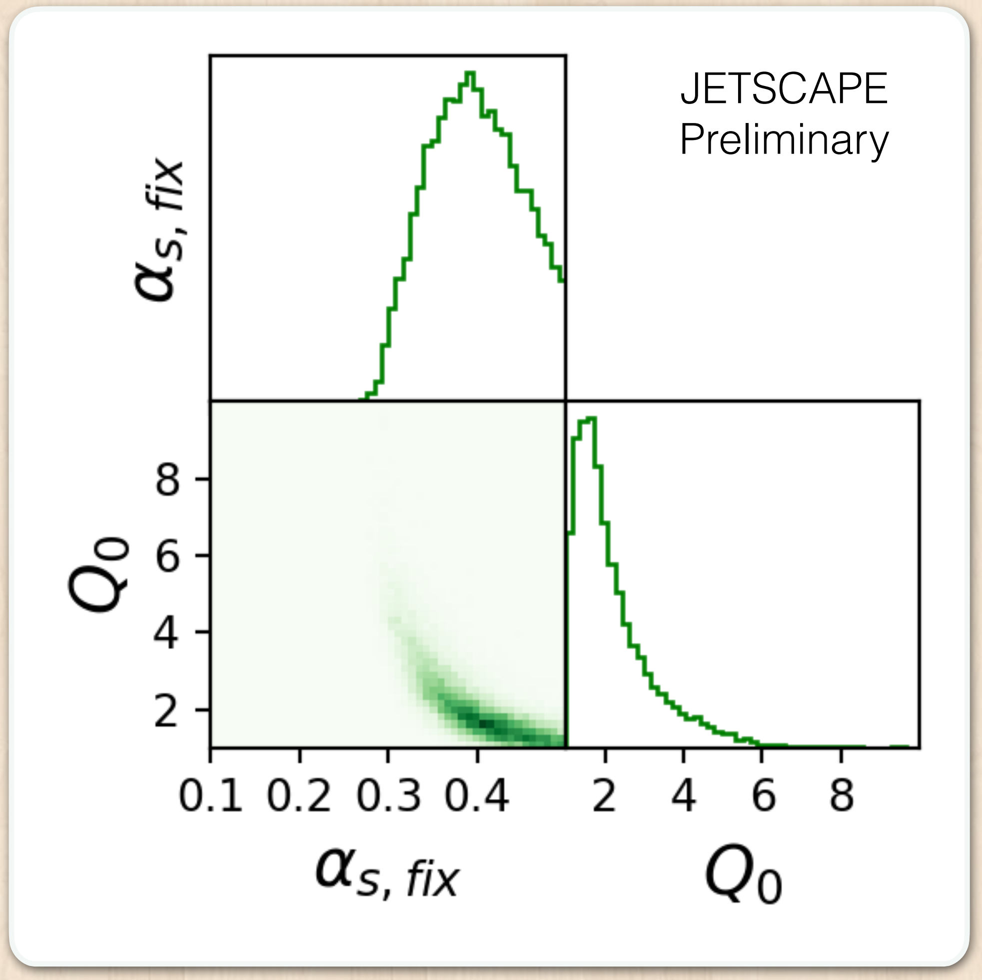
Q_0: virtuality switch to LBT

τ_0: start time

2204.01163, see also previous talk

Posterior distribution





Start time

Virtuality dependent terms

Anti-correlation between $\alpha_{s,fix}$ and Q switch

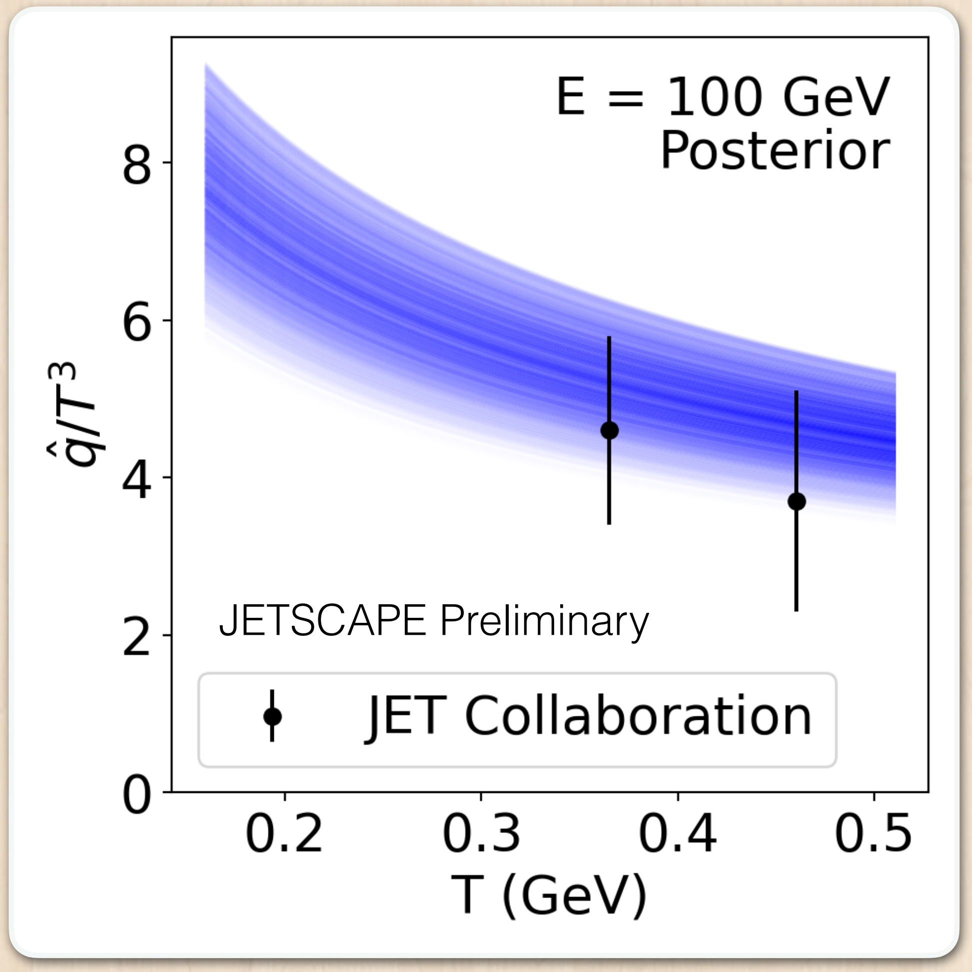
$\propto \hat{q}$

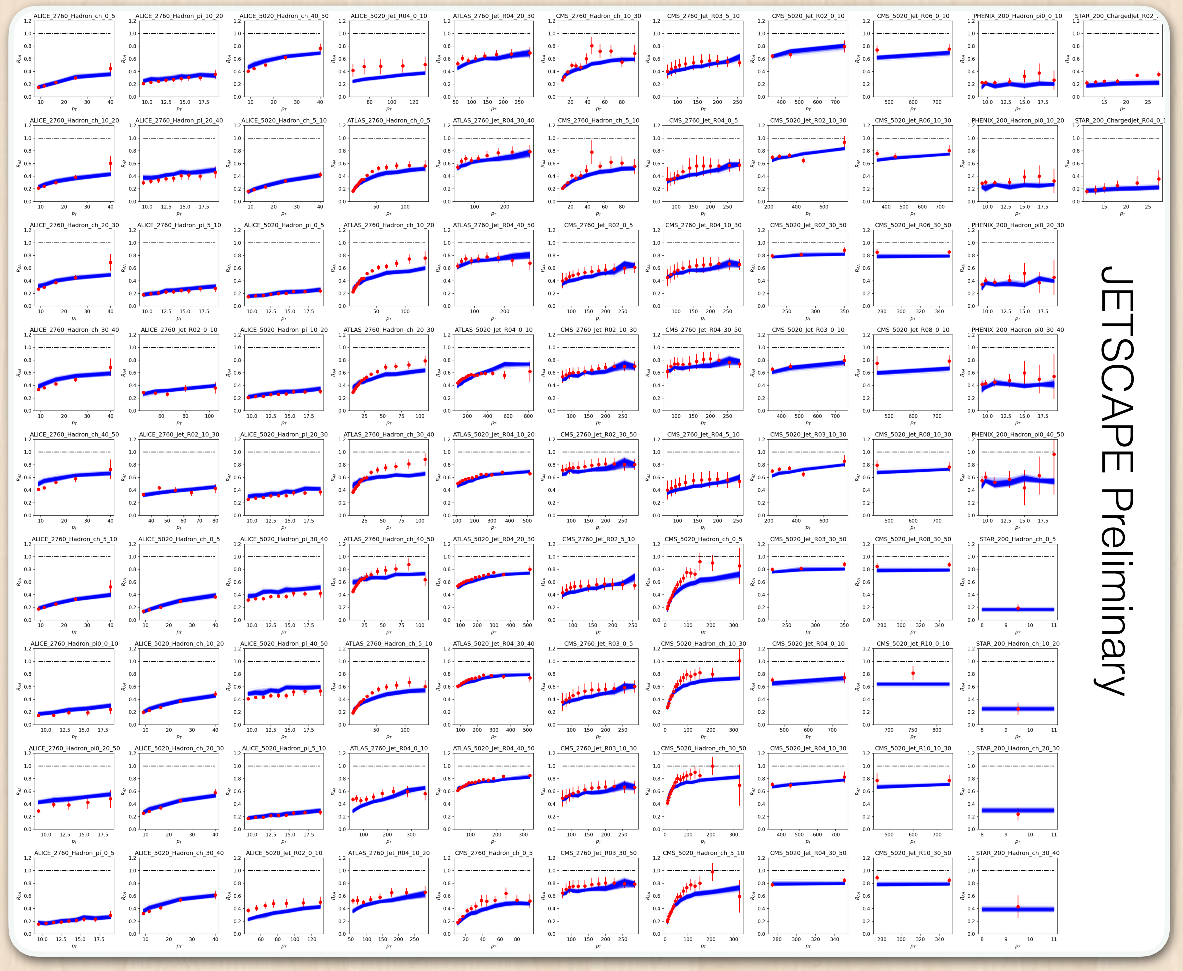
Between MATTER and LBT

New analysis of $\hat{q}$

Included jet R_{AA} into the mix!

General reasonable description of data

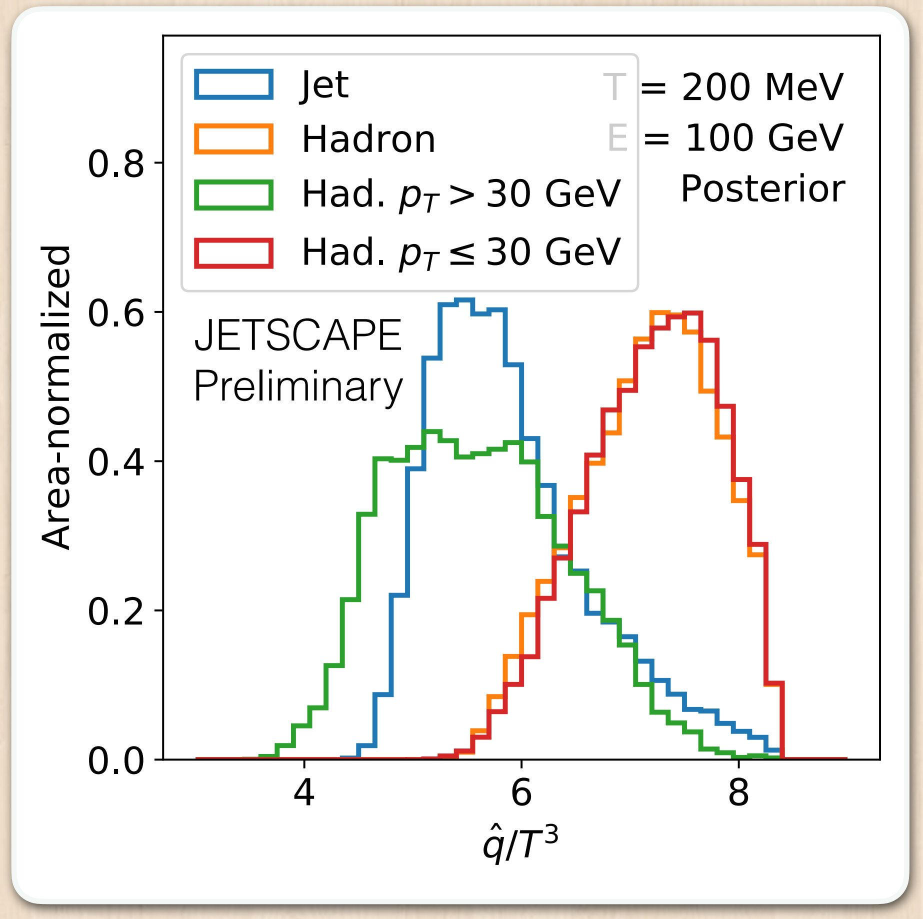




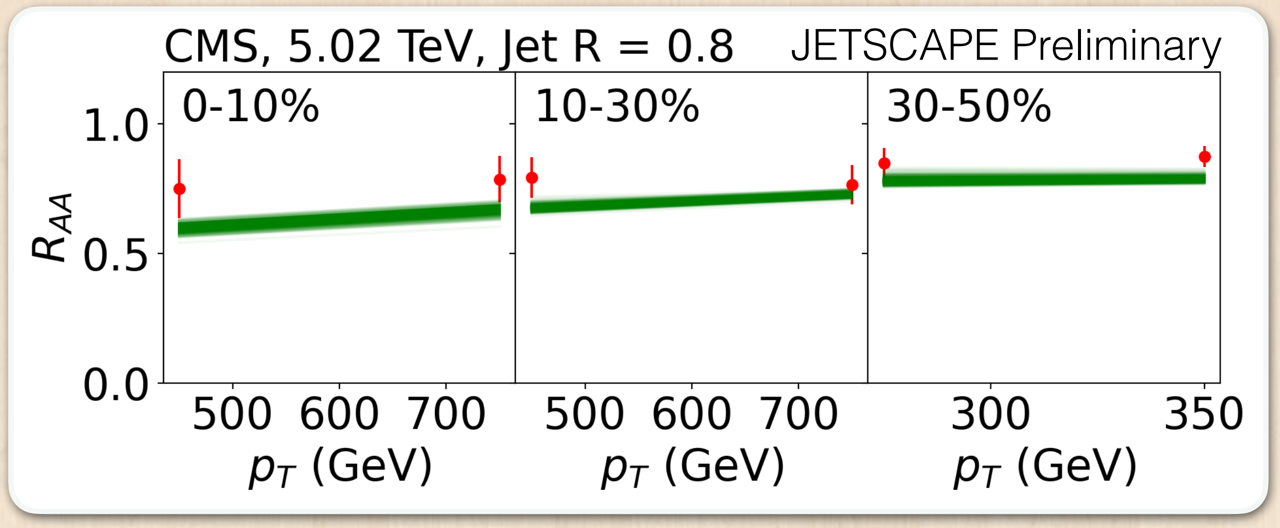
All these impossible without a framework

Endless possibilities

Bayesian analysis: powerful tool for not only **parameter extraction** but also **model studies**



Pinpoint interesting phase space in model



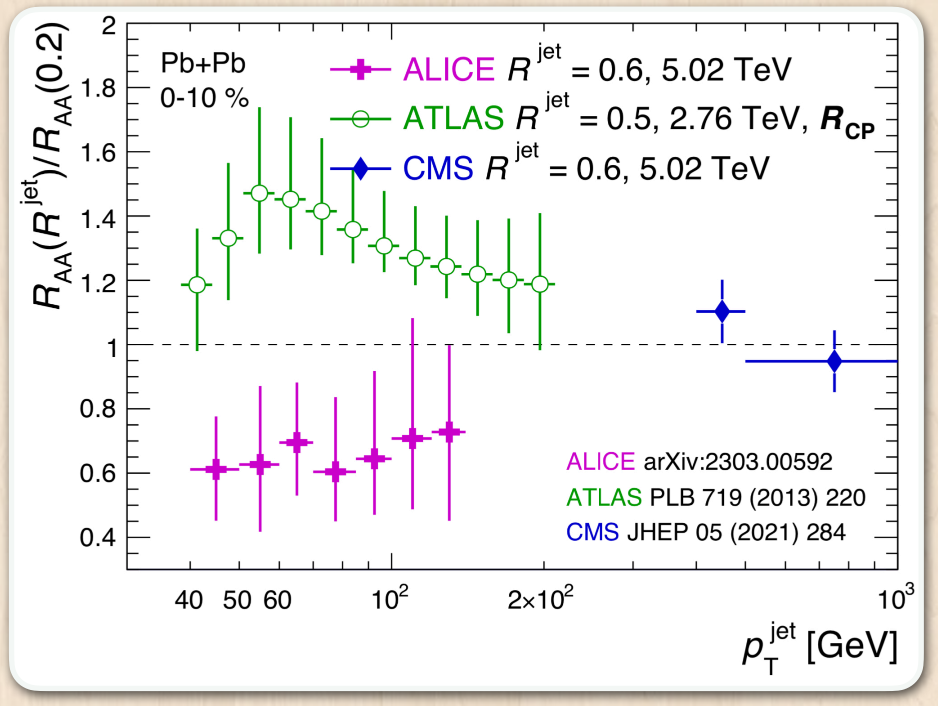
Evaluate how well model does in new observables

Theory uncertainties?

(Near-) future prospects

We also calculated huge number of **other jet-related observables**

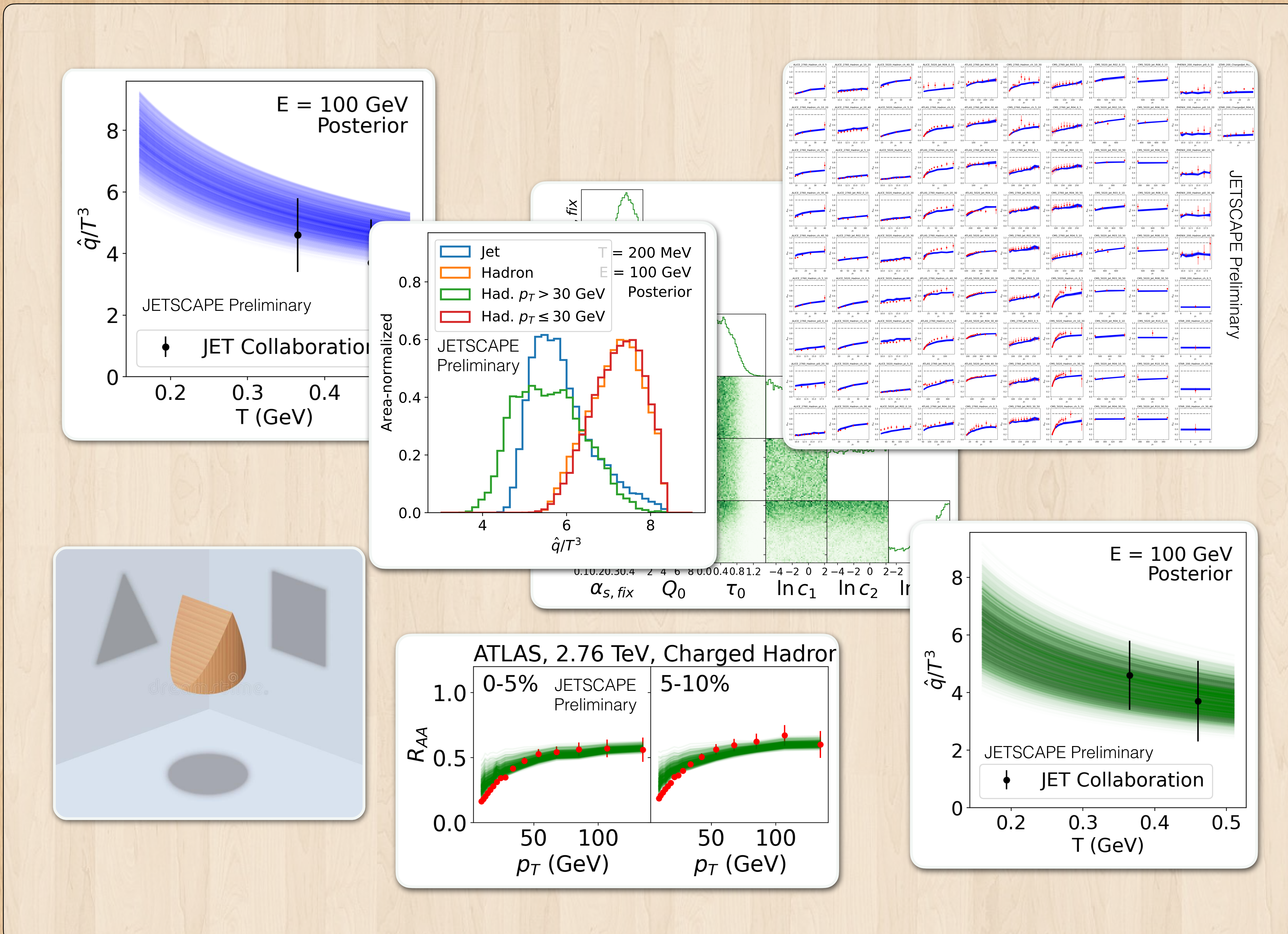
Plot taken from Y. Go, Mon Mar 27



Move one step at a time and **sequentially include more observables** → stay tuned for many new results in the near future!

Important to include ALL eligible data

Ready to explore the theory / experimental landscape


E = 100 GeV
Posterior
JETSCAPE Preliminary
JET Collaboration
T (GeV)
Jet
Hadron
Had. $p_T > 30$ GeV
Had. $p_T \leq 30$ GeV
T = 200 MeV
E = 100 GeV
Posterior
JETSCAPE
Preliminary
Area-normalized
$\hat{q}/T^3$
$\alpha_{s,\,fix}$
Q_0
τ_0
$\ln c_1$
$\ln c_2$
JETSCAPE Preliminary
ATLAS, 2.76 TeV, Charged Hadron
0-5%
5-10%
R_{AA}
p_T (GeV)
E = 100 GeV
Posterior
JETSCAPE Preliminary
JET Collaboration
T (GeV)