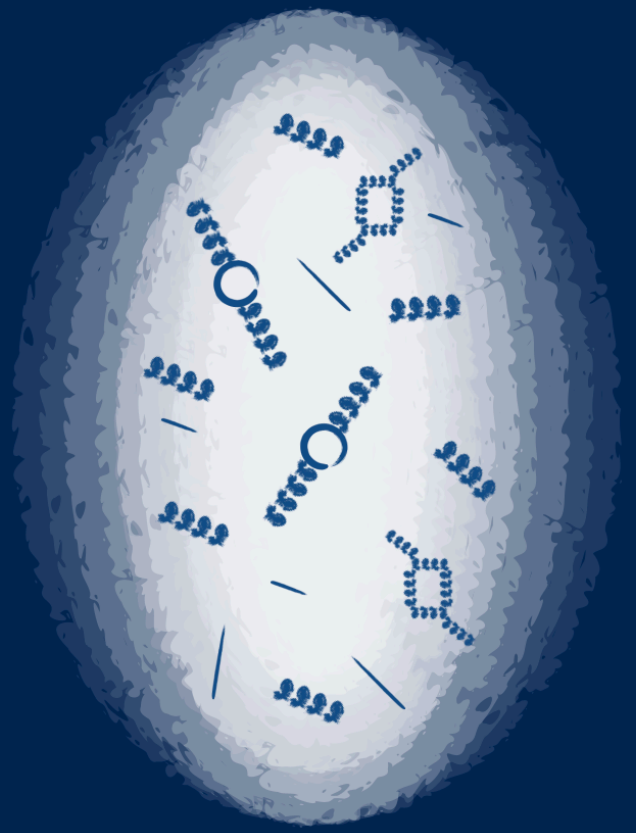
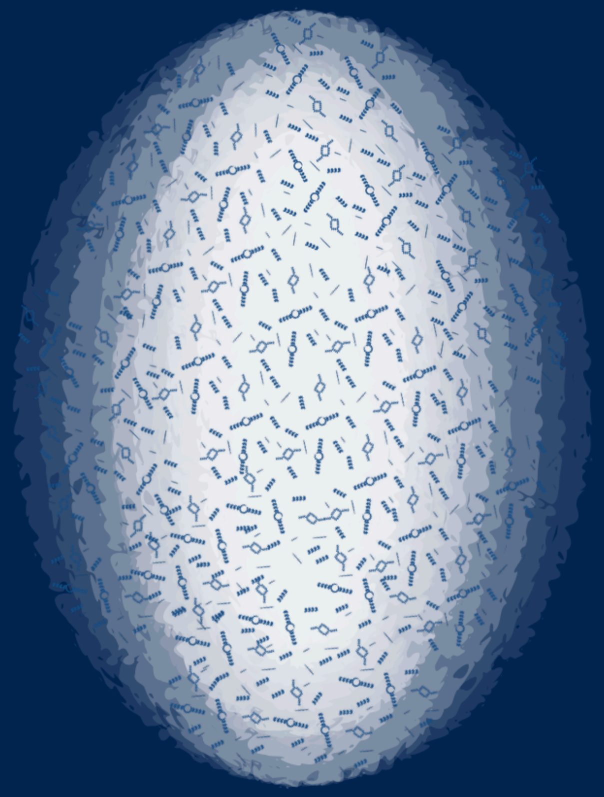
Novel tools for data analysis



James Mulligan

UC Berkeley and LBNL



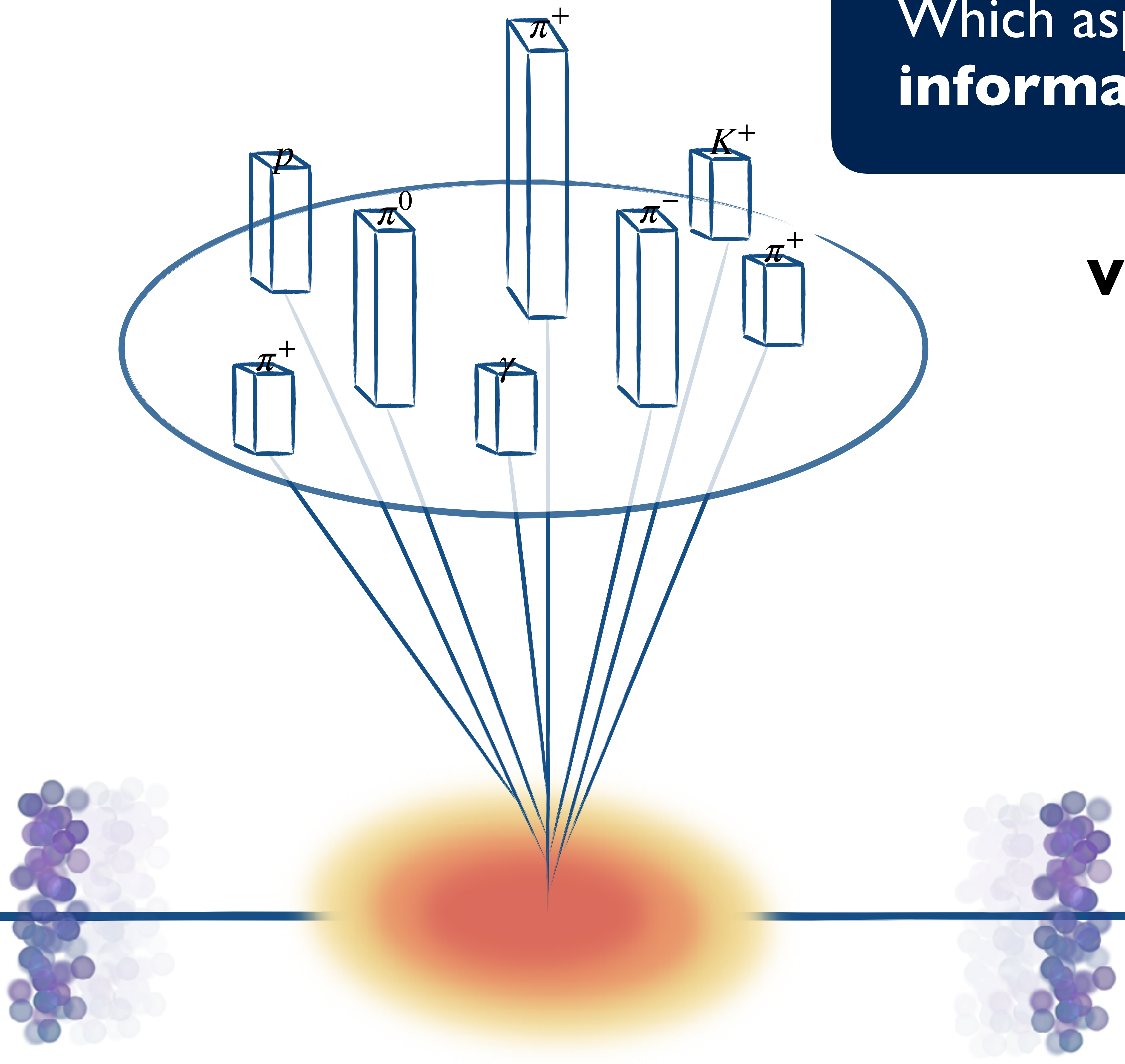
Quark Matter 2023

Houston, Texas

Sep 4, 2023

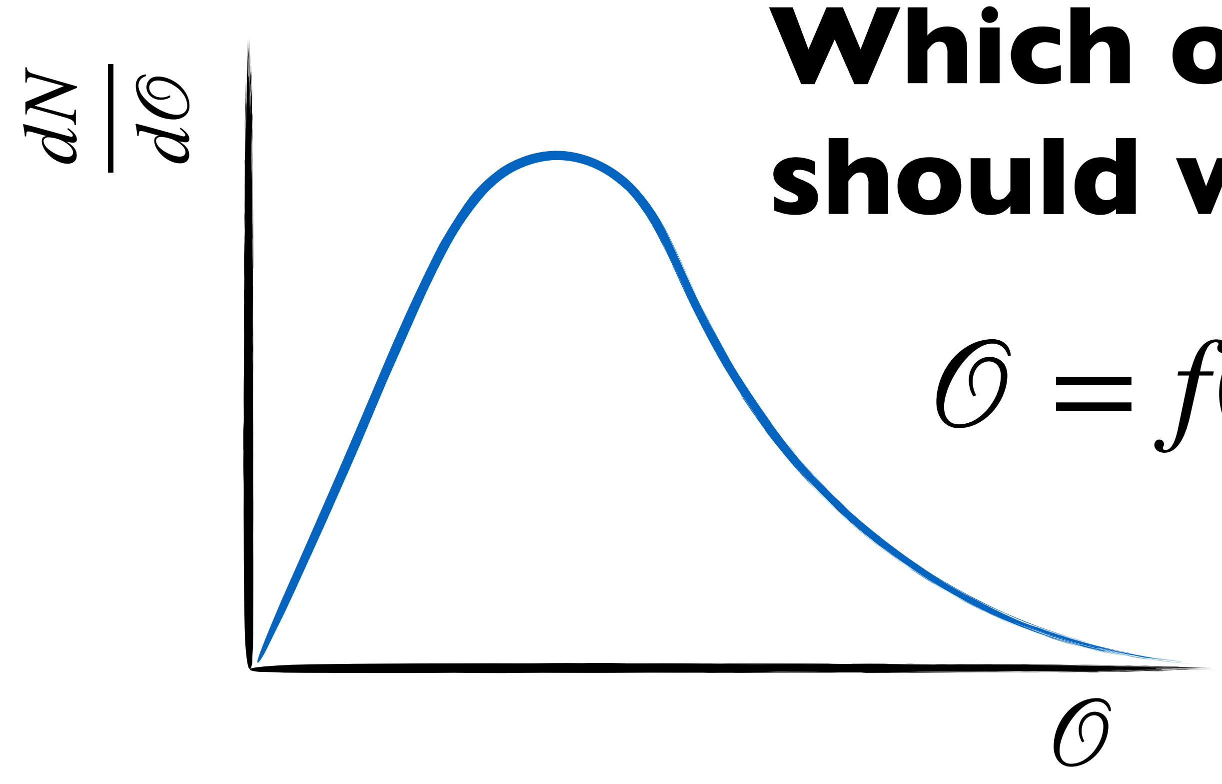
Which aspects of collider events contain **useful information** about emergent properties of QCD?

Which aspects of collider events contain **useful information** about emergent properties of QCD?



Vast phase space

- $\mathcal{O}(10 - 10^4)$ correlated particles per event
- Typically: 1D projection over ensemble



Which observables should we measure?

$$\mathcal{O} = f(\mathbf{p_0}, \ldots, \mathbf{p_n})$$

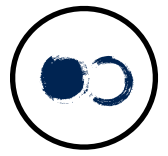
Types of observables

Single objects

Identified hadron

Jet

Quarkonium

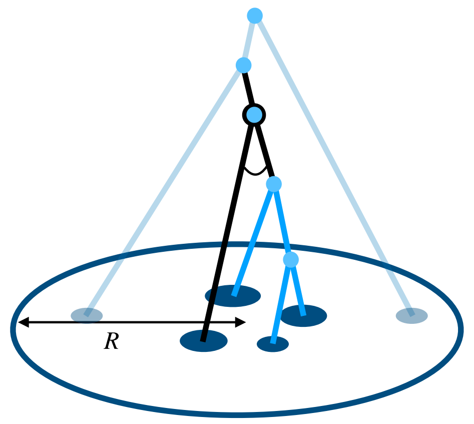


...

These are “straightforward”

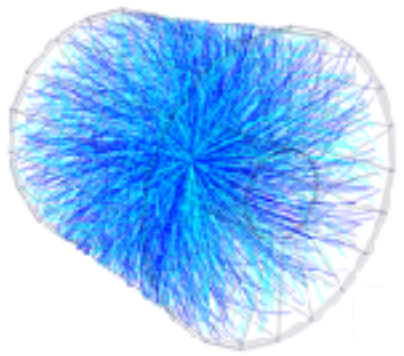
Many-body observables

Jet substructure



HF correlations

Event-wide correlations



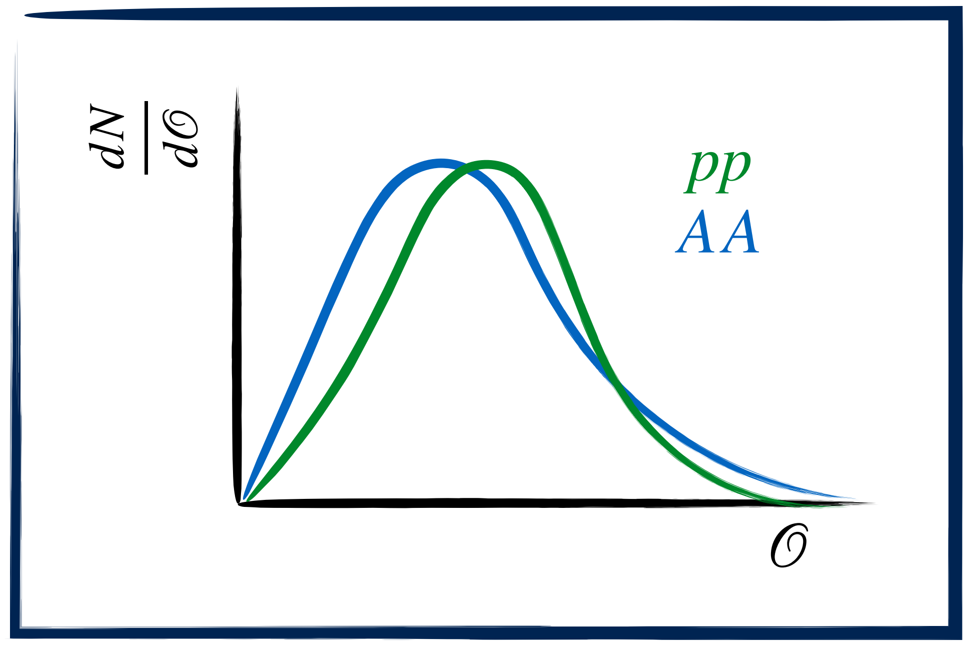
...

Contain the most information

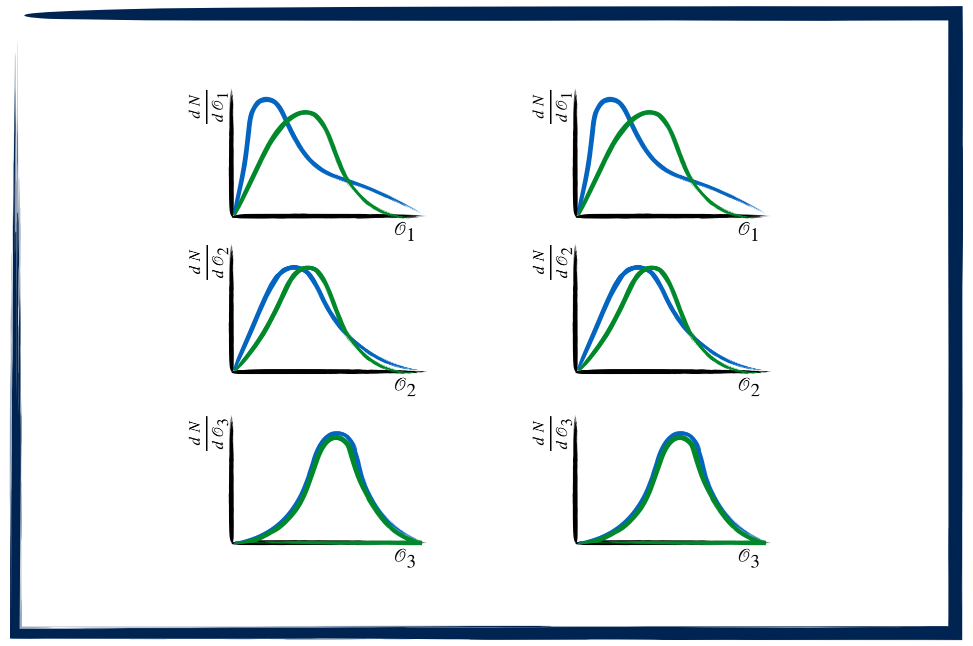
What is our strategy of how to leverage the full available information?

Three approaches to learn about the QGP

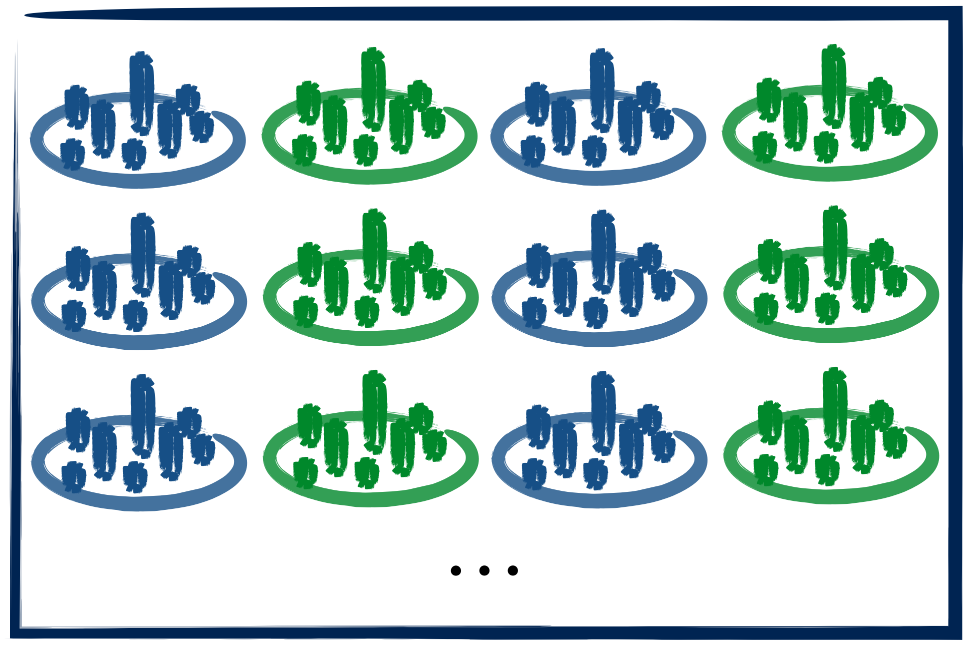
1. Single observables



2. Sets of observables

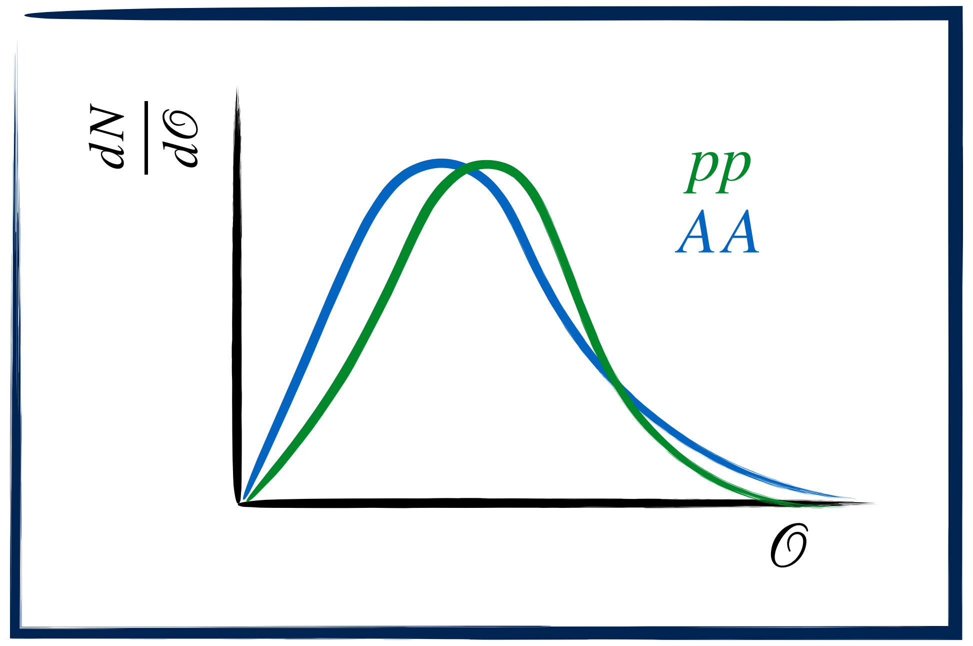


3. Particle-level information



Three approaches to learn about the QGP

1. Single observables



2. Sets of observables

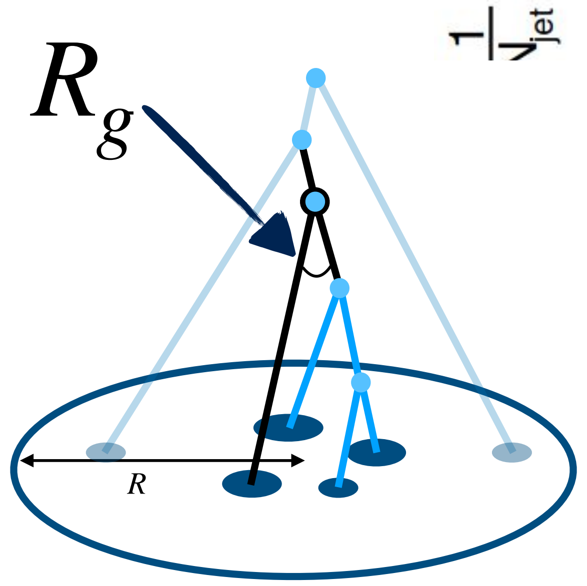
3. Particle-level information

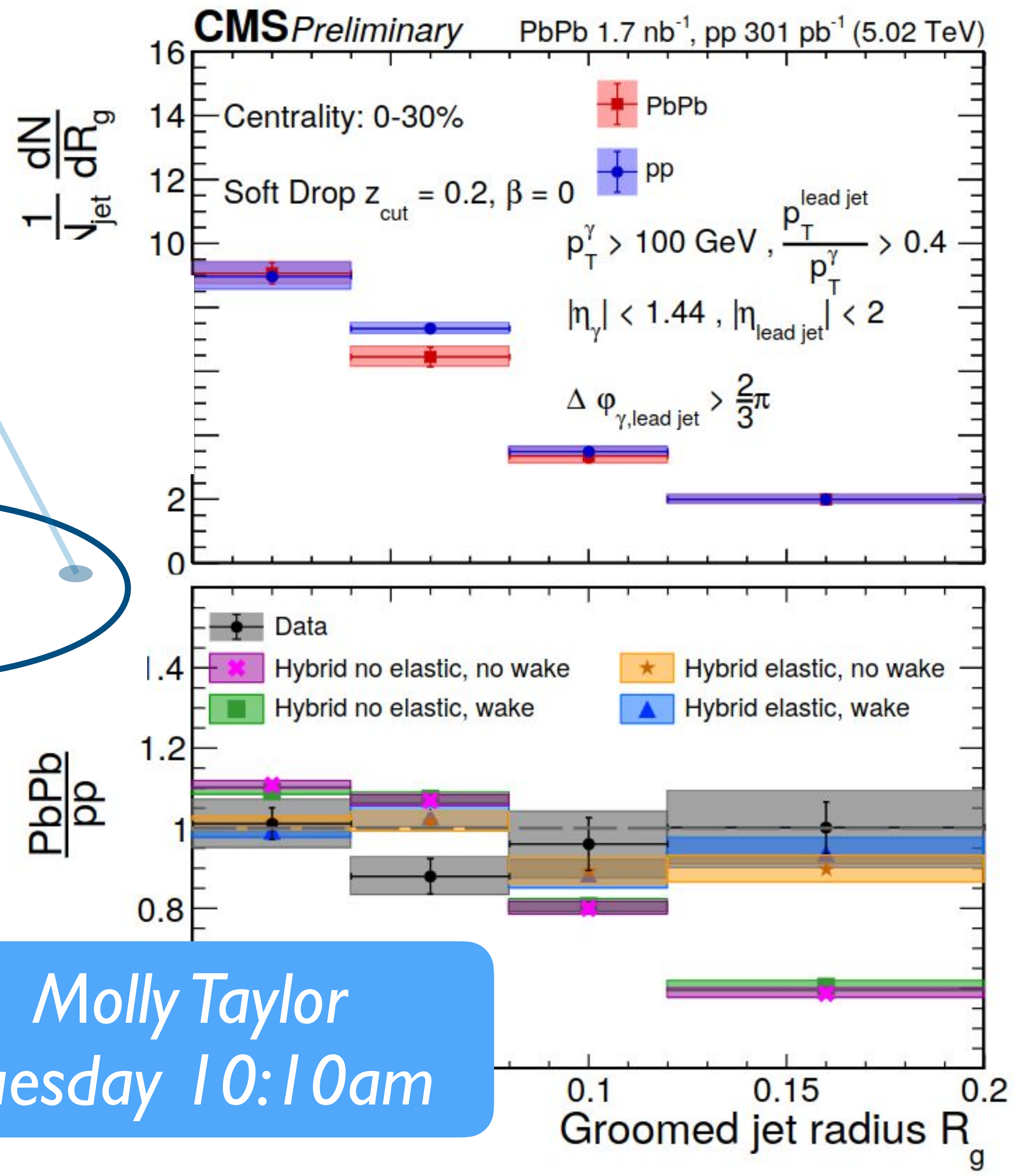
Experimental and theoretical building blocks

New channels

Example: CMS γ-jet R_g

Provide handle on confounding QCD mechanisms



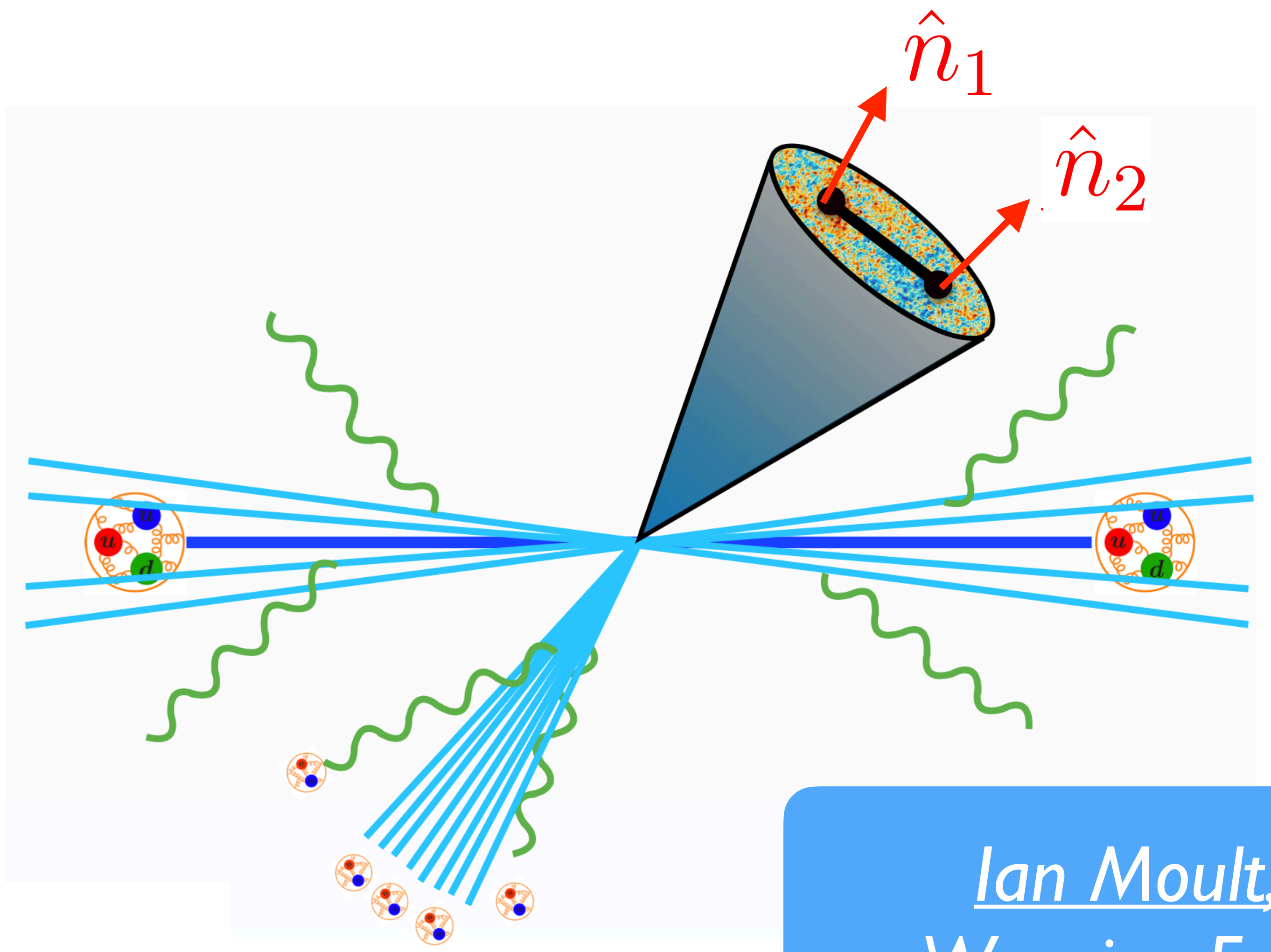


Molly Taylor
Tuesday 10:10am

New insights

Example: Energy correlators

Novel non-perturbative observables that may be calculable in QCD



Ian Moult, Tues 9:10am
Wenqing Fan, Wed 8:50am
Carlota Andres, Wed 10:10am

Connecting experiment and theory

Observables must be corrected for detector and background effects

Example: High-dimensional corrections

ML-based unfolding algorithms may allow us to directly explore high-dimensional observables

Andreassen et al. PRL 124 182001 (2020)

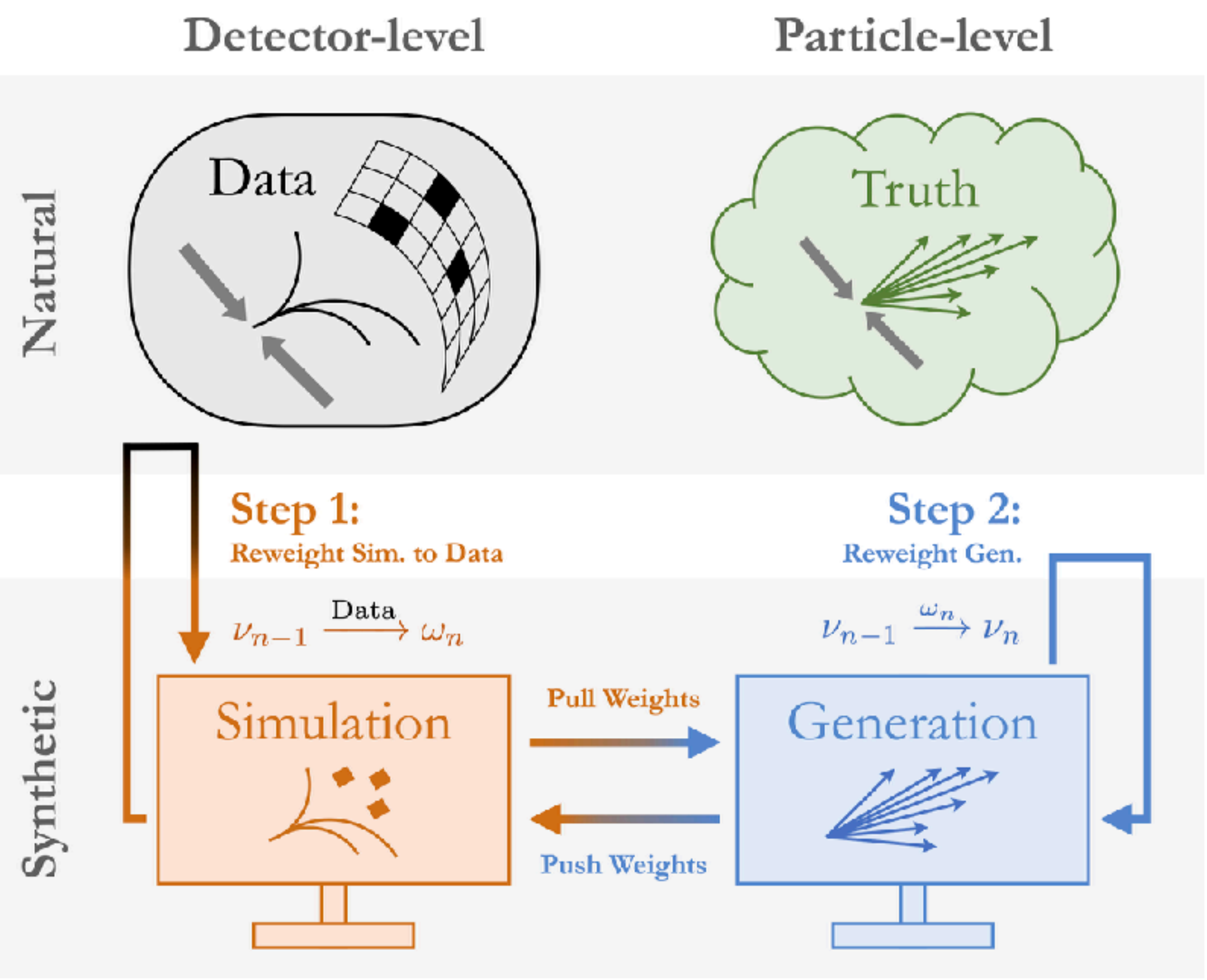
Can this be successfully implemented in heavy-ion environment?

→ Open question

Youqi Song
Poster Tues 5:30pm

See also:
Novel application of mixed-event background correction

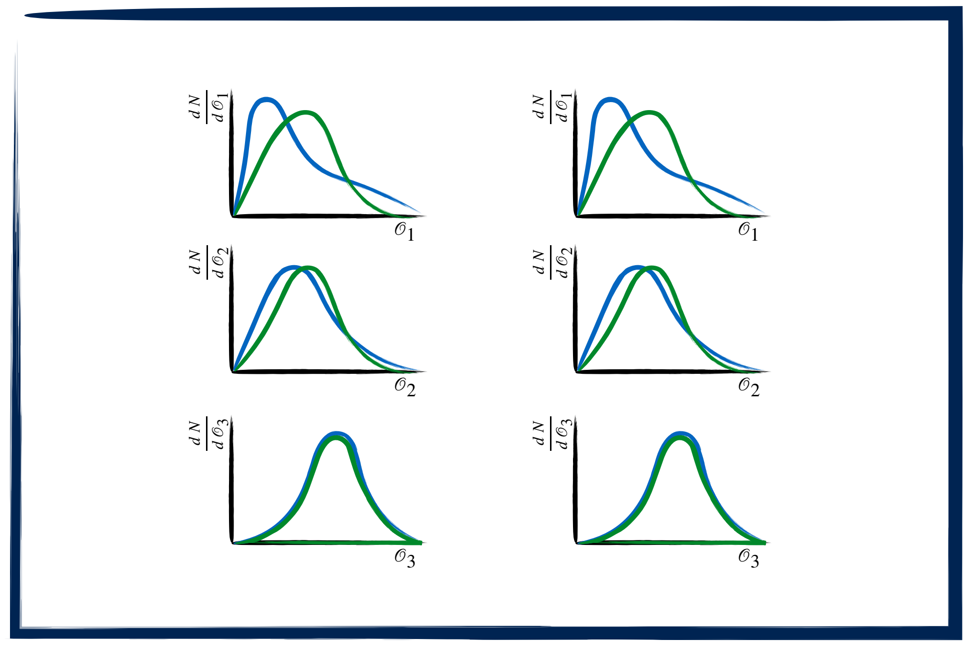
Nadine Grünwald
Tuesday 11:40am



Three approaches to learn about the QGP

1. Single observables

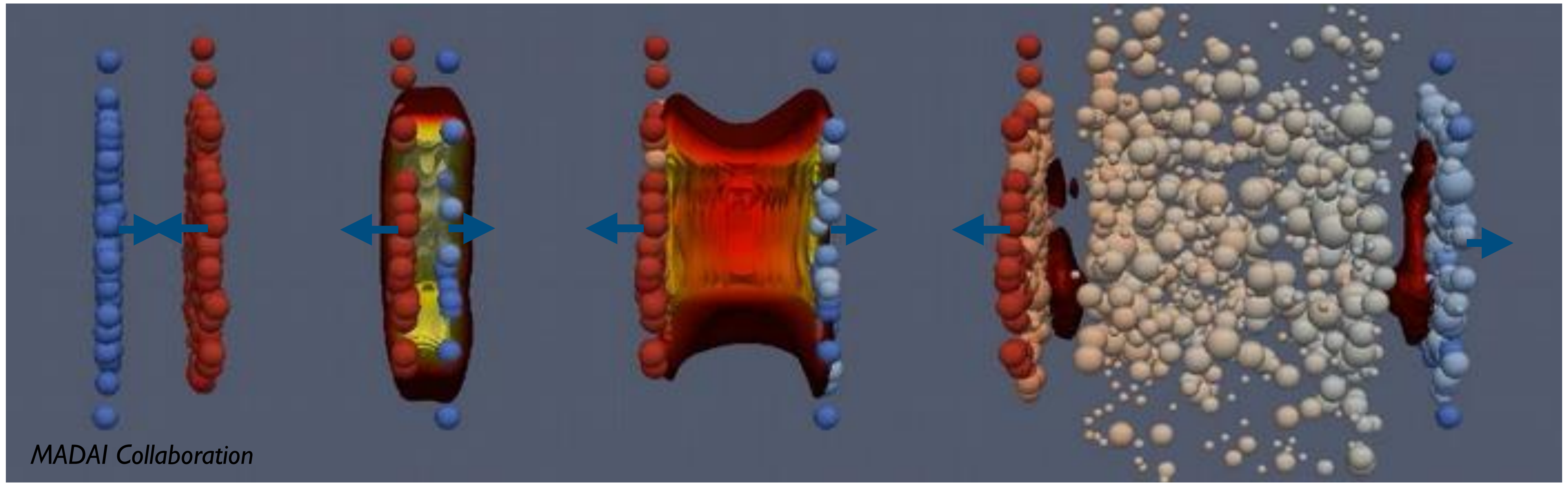
2. Sets of observables



3. Particle-level information

There is no golden observable

Multiple stages of heavy-ion collision — fit + predict



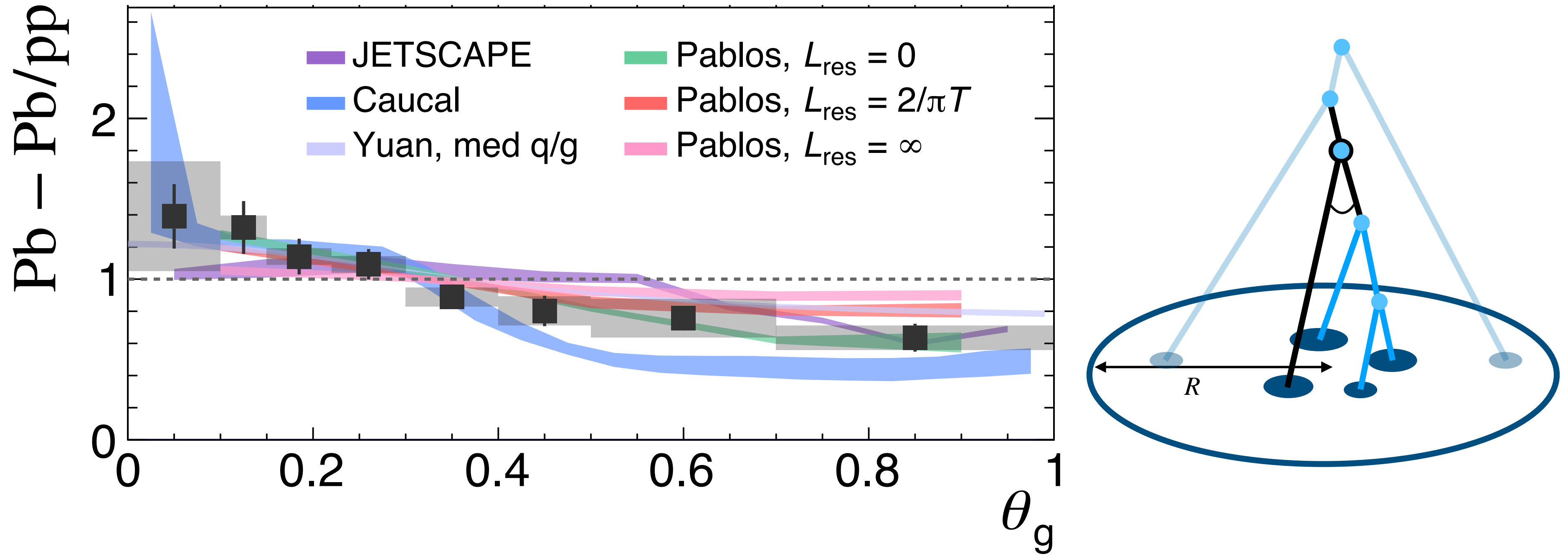
initial state

hydrodynamic evolution
medium response

hadronization
hadronic rescattering

Similar predictions for single observables

ALICE PRL 128 (10), 102001 (2022)



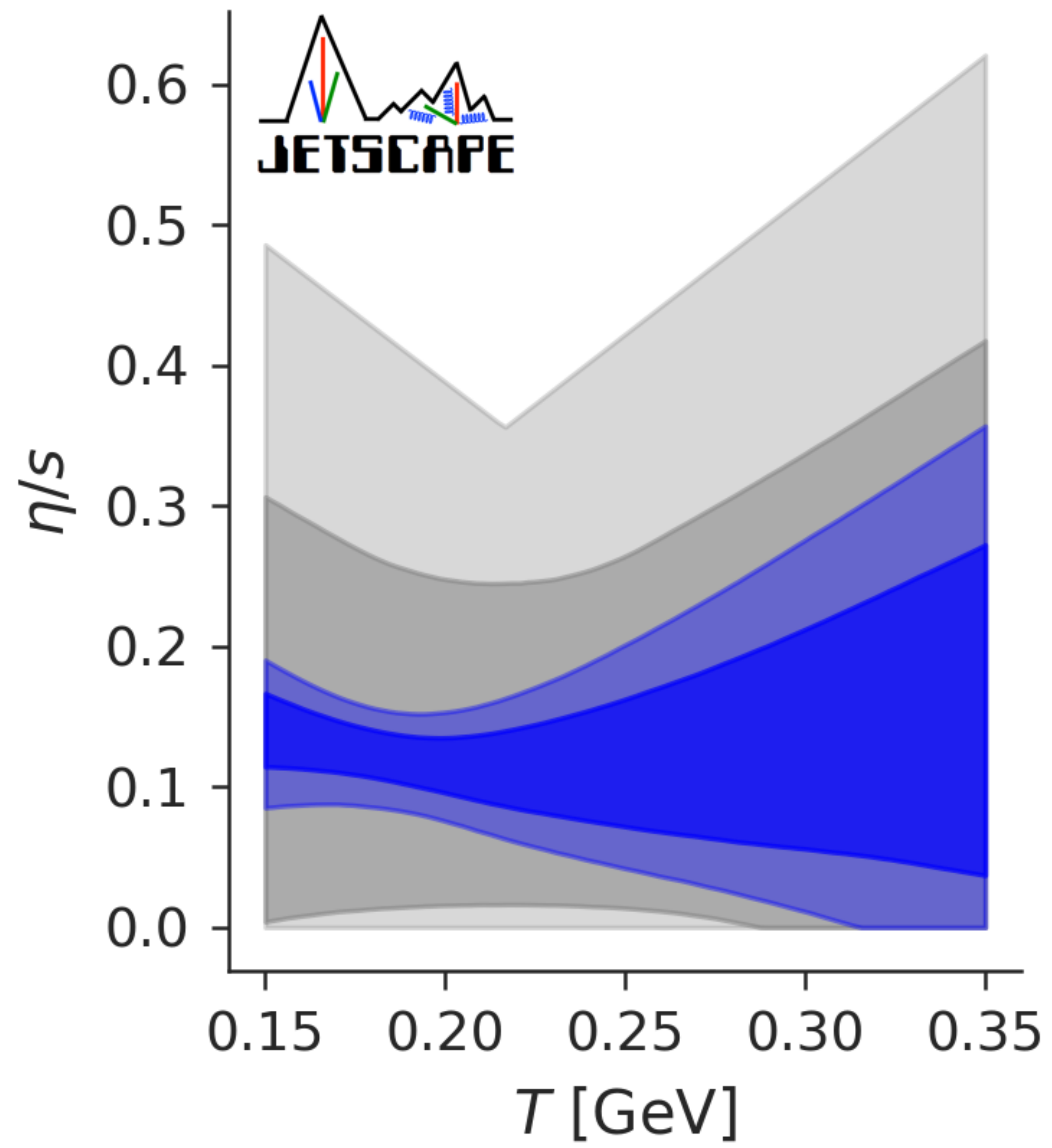
→ Need multiple observables to constrain medium properties

Bayesian inference in heavy-ion collisions

Andi Mankolli, Tues 8:50am
Mayank Singh, Tues 9:50am
Mauricio Teixeira, Tues 9:50am
Chun Shen, Tues 3:30pm
Raymond Ehlers, Tues 3:50pm
Ulrich Heinz, Wed 8:50am
Débora Mroczek, Wed 11:00am
Wenbin Zhao, Wed 3:40pm
…

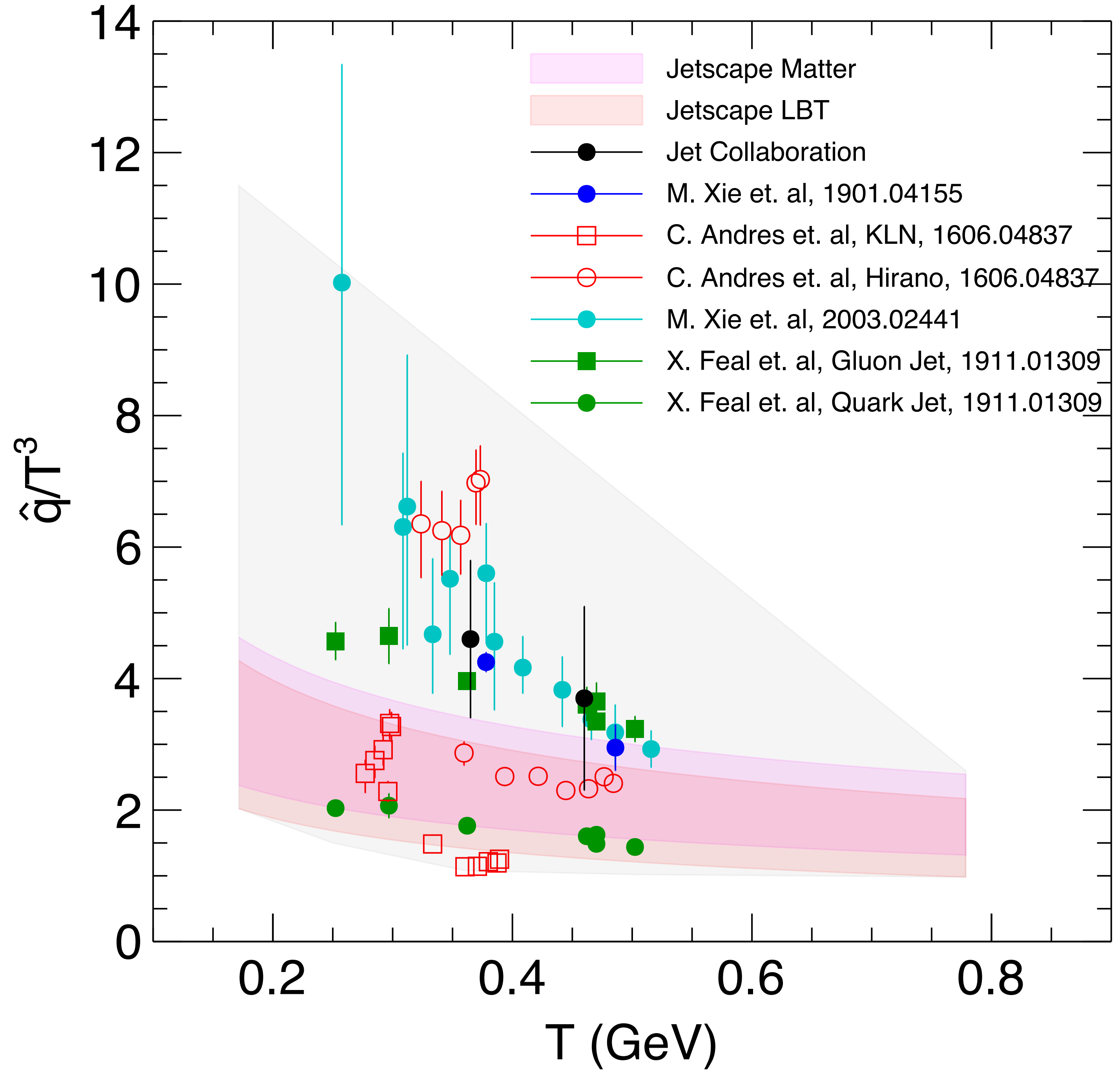
Shear viscosity

JETSCAPE PRL 126, 242301 (2021)



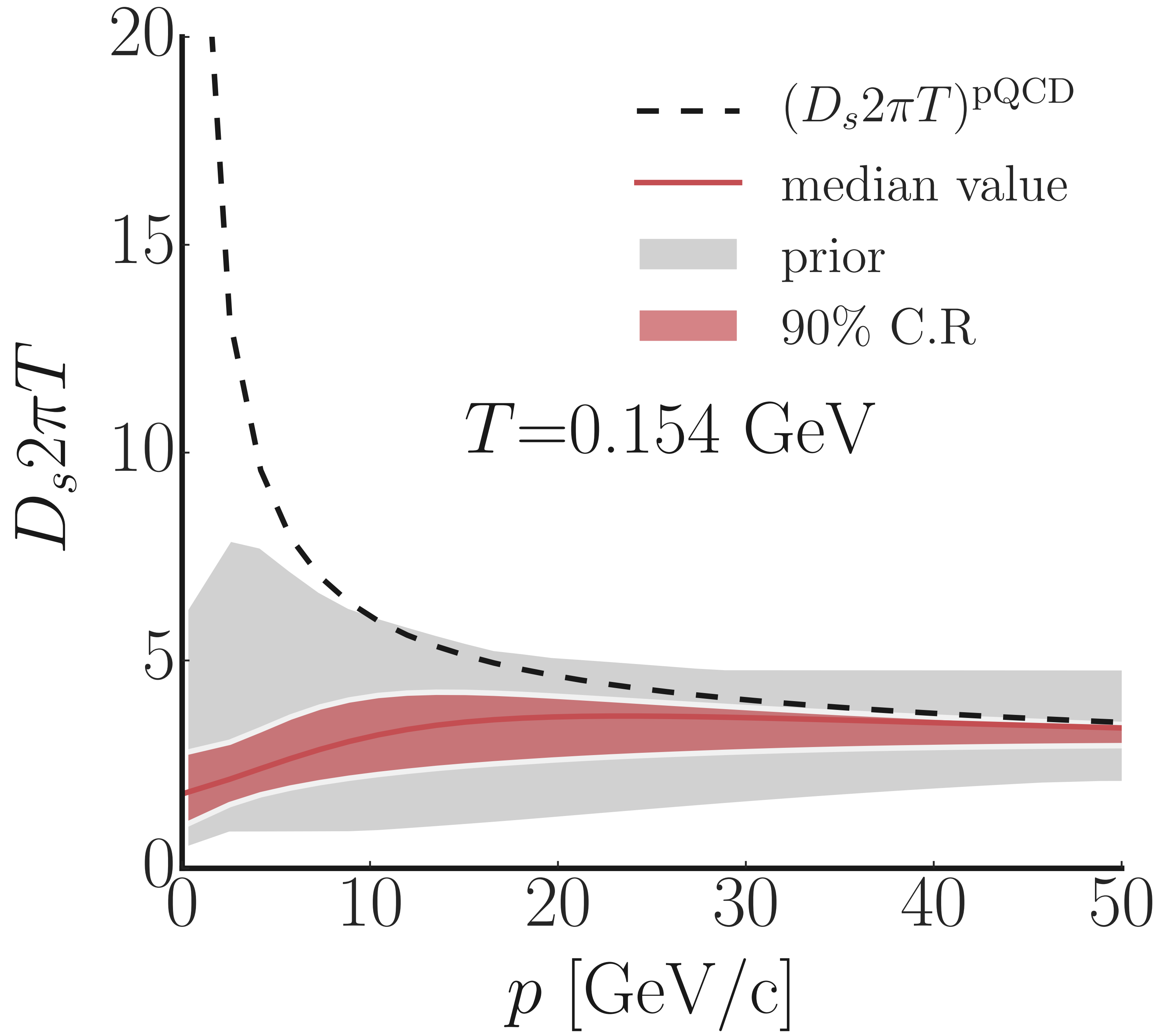
Jet transport

Apolinario, Lee, Winn PPNP 127, 103990 (2022)



HF transport

Xu et al. PRC 97 1, 014907 (2018)



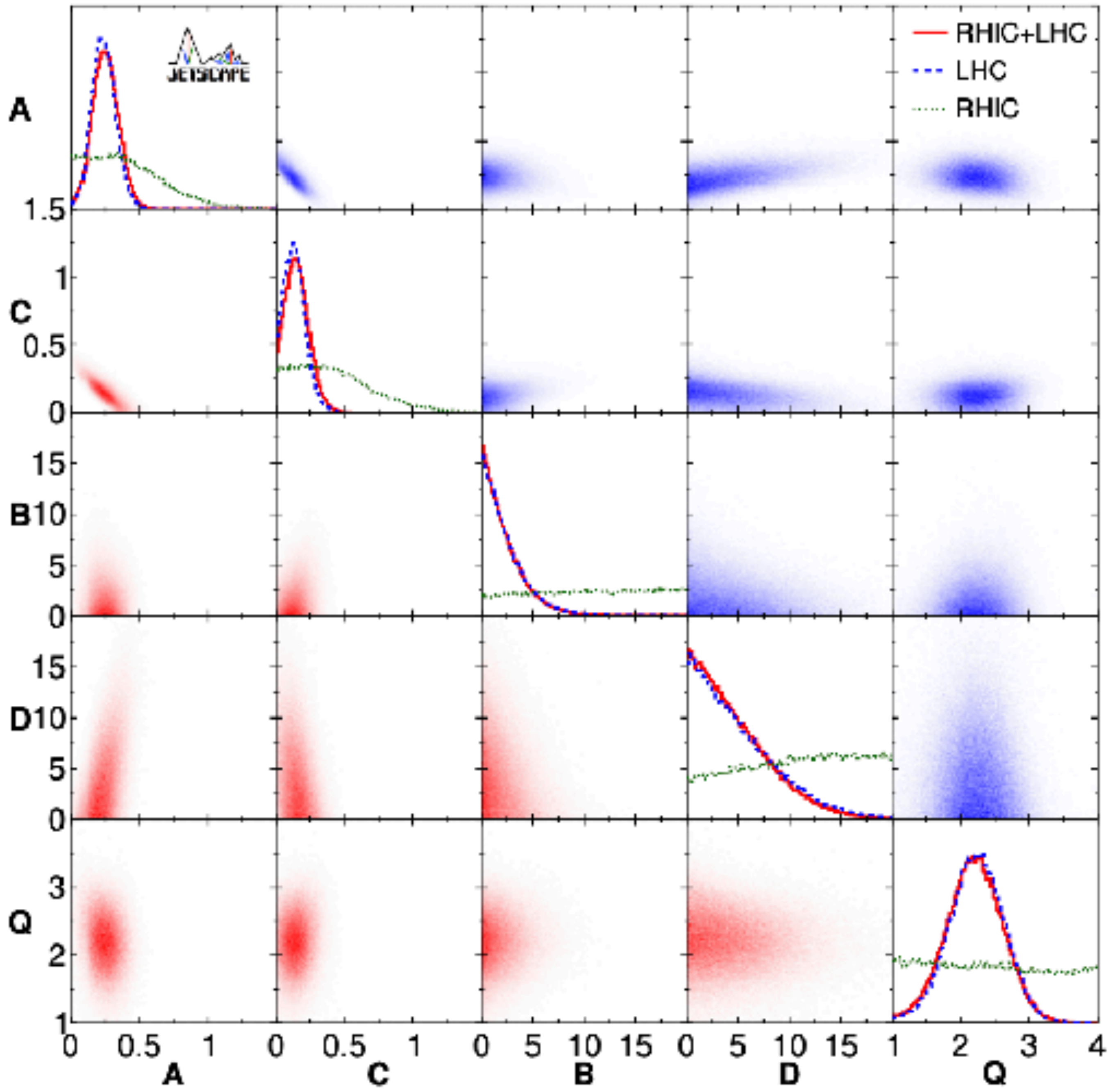
…

Experimental guidance from Bayesian inference

Long-term planning

Example:
Restrict to either RHIC or LHC data

JETSCAPE PRC 104, 024905 (2021)

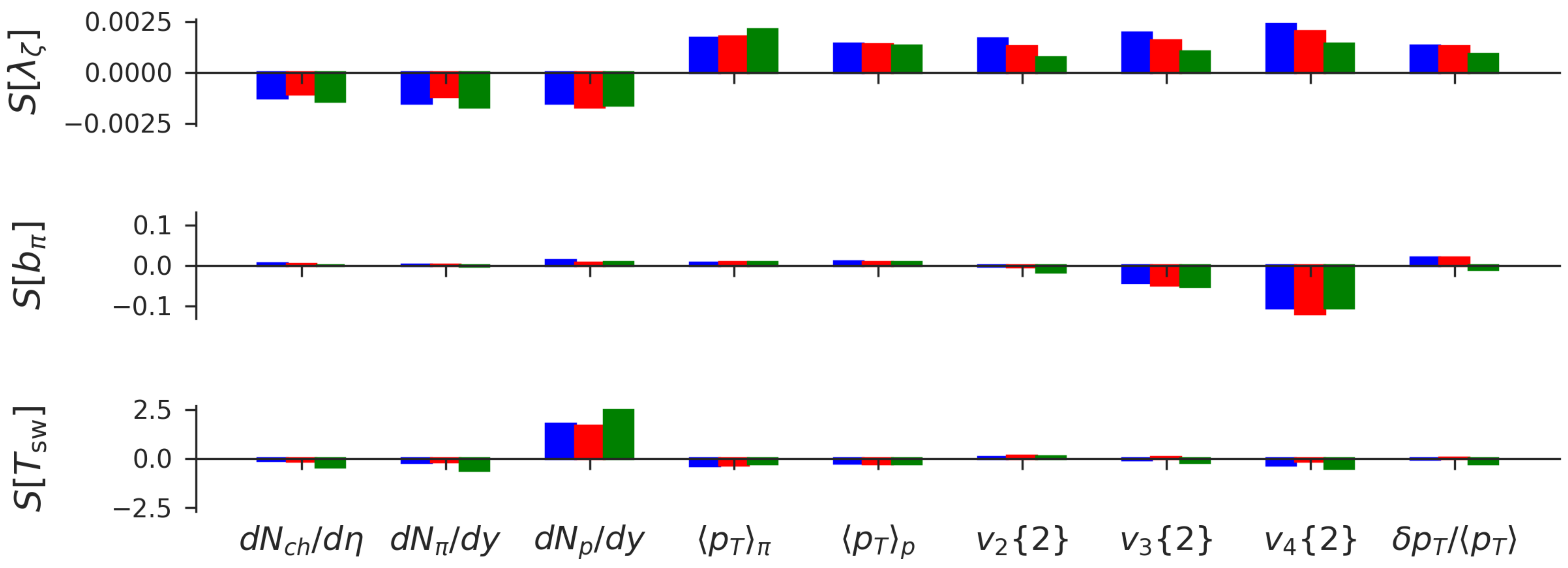


Fit dominated by LHC data

Parameter sensitivity

Quantify impact of a model parameter on measured observables

JETSCAPE PRC 103, 054904 (2021)



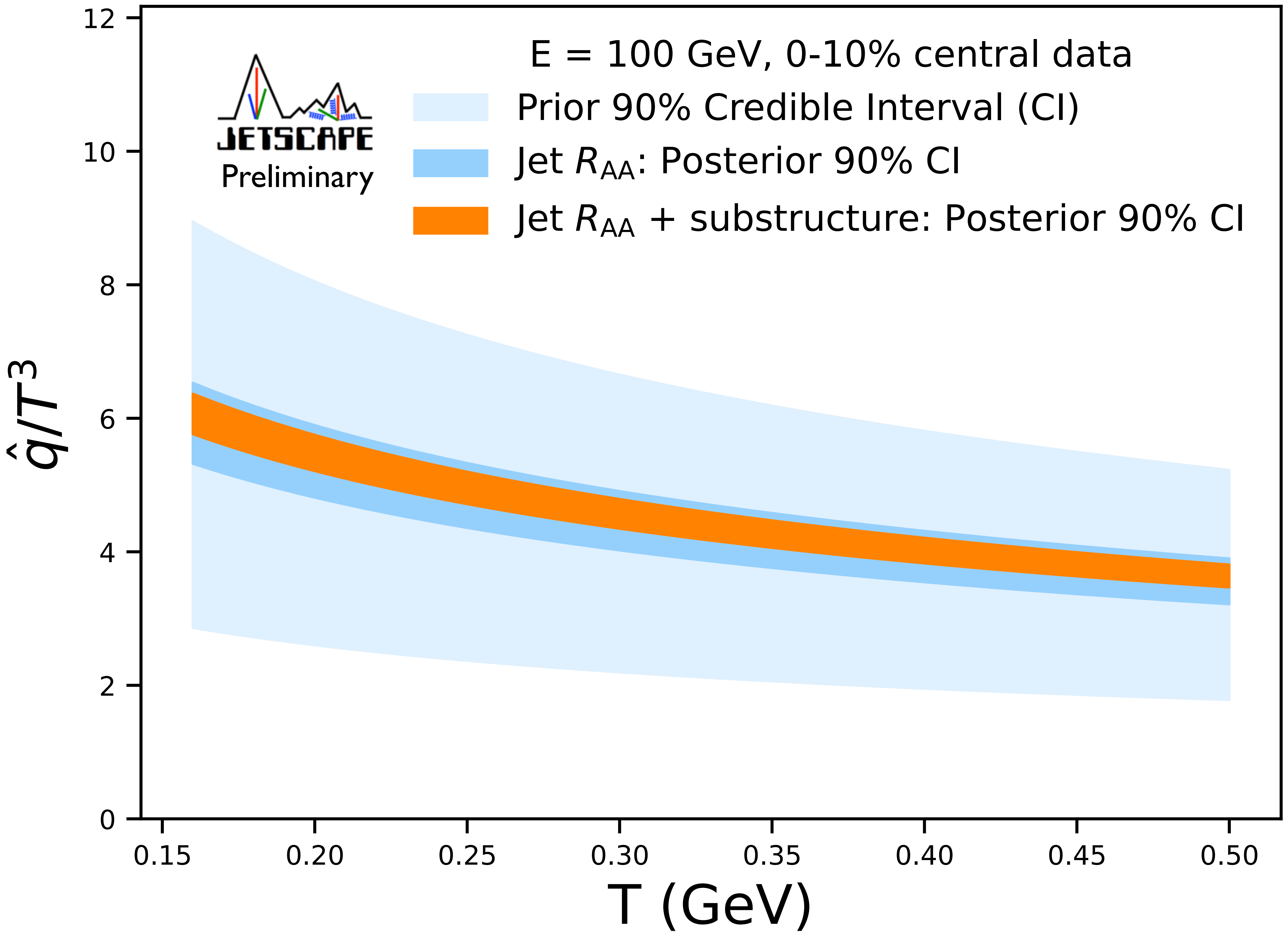
See also:
Lai arXiv 1810.00835
Sangaline, Pratt PRC 93, 024908 (2016)

Model-dependent guidance on where to focus experimental effort

Sensitivity of transport coefficients to observables

Jet substructure constrains $\hat{q}$

Raymond Ehlers
Tuesday 3:50pm

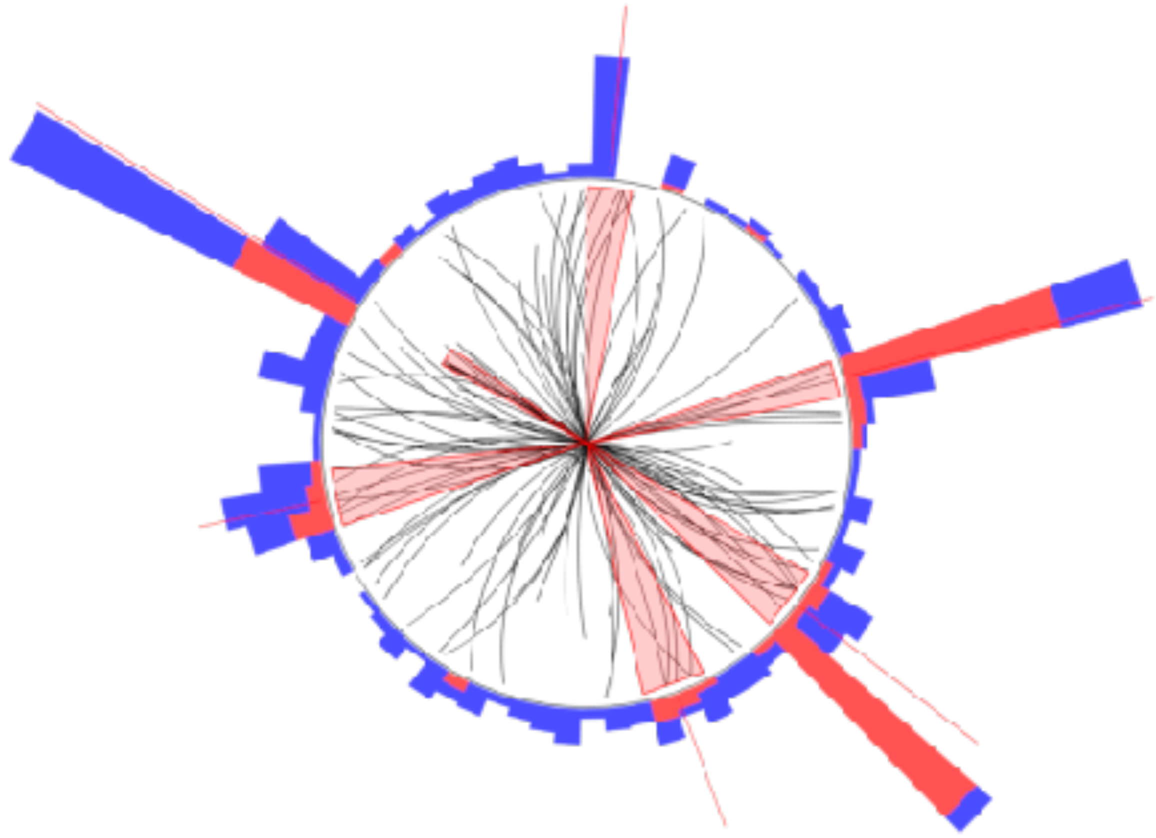


Starting point: **systematic, iterative selection of observables** to constrain QGP properties

Note: Very few published jet observables can be directly compared to theory

- This analysis: $D(z), z_g, \theta_g$

Universality

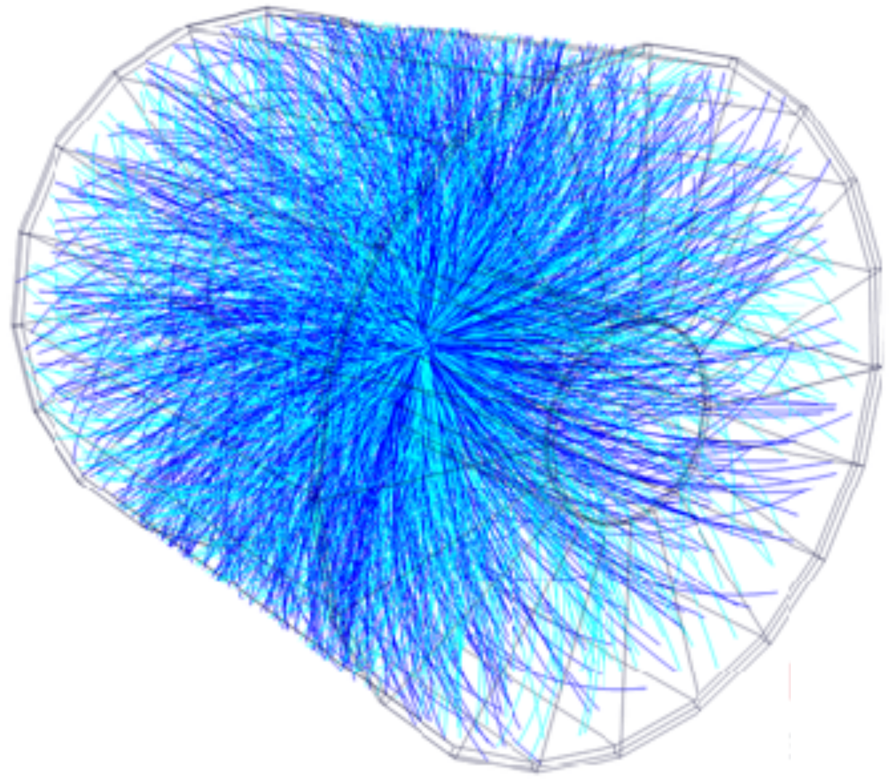


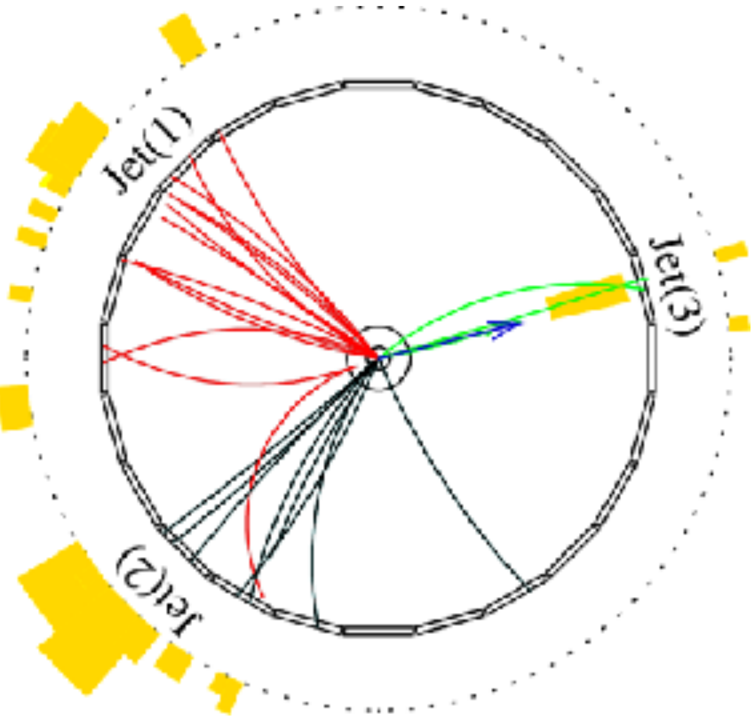
pp

High energies

AA

High densities



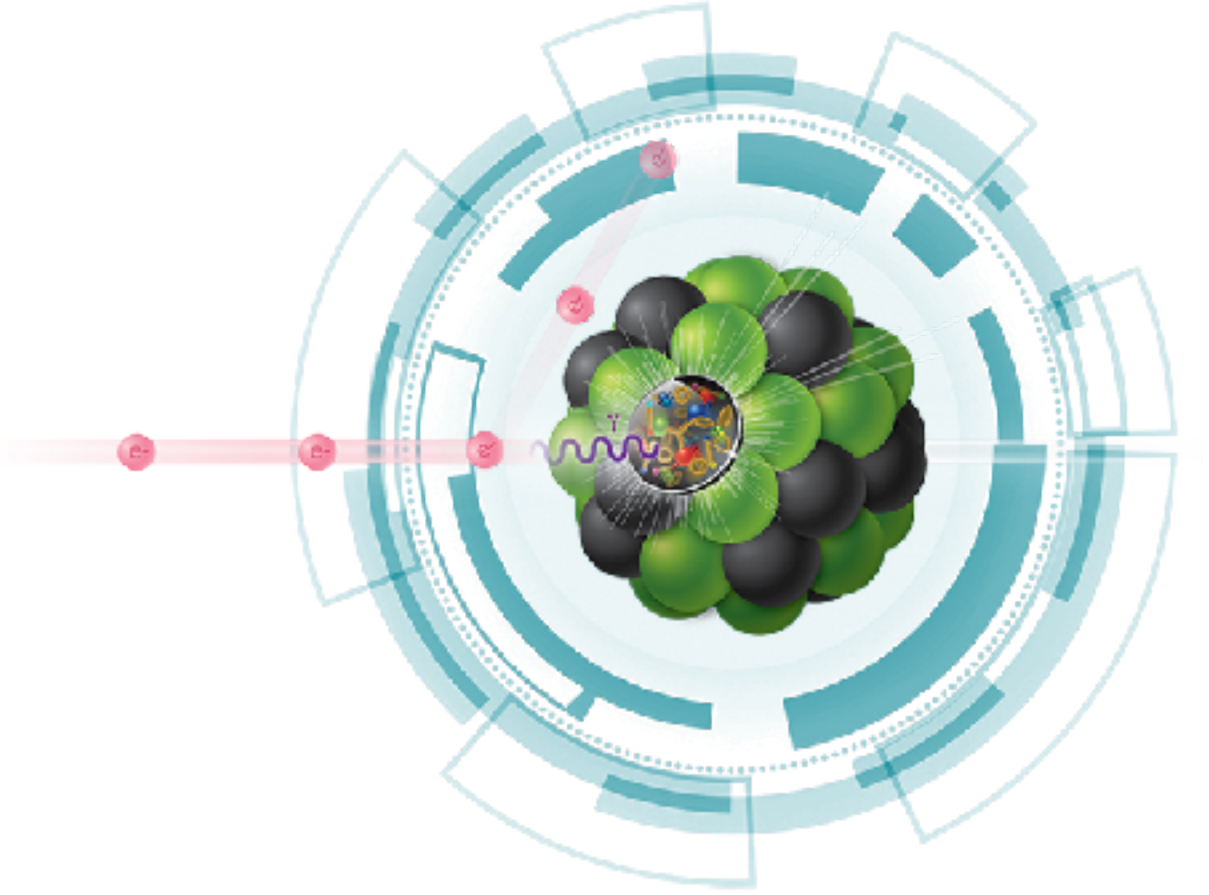


e^+e^-

Elementary beams

ep, eA

Precision probes



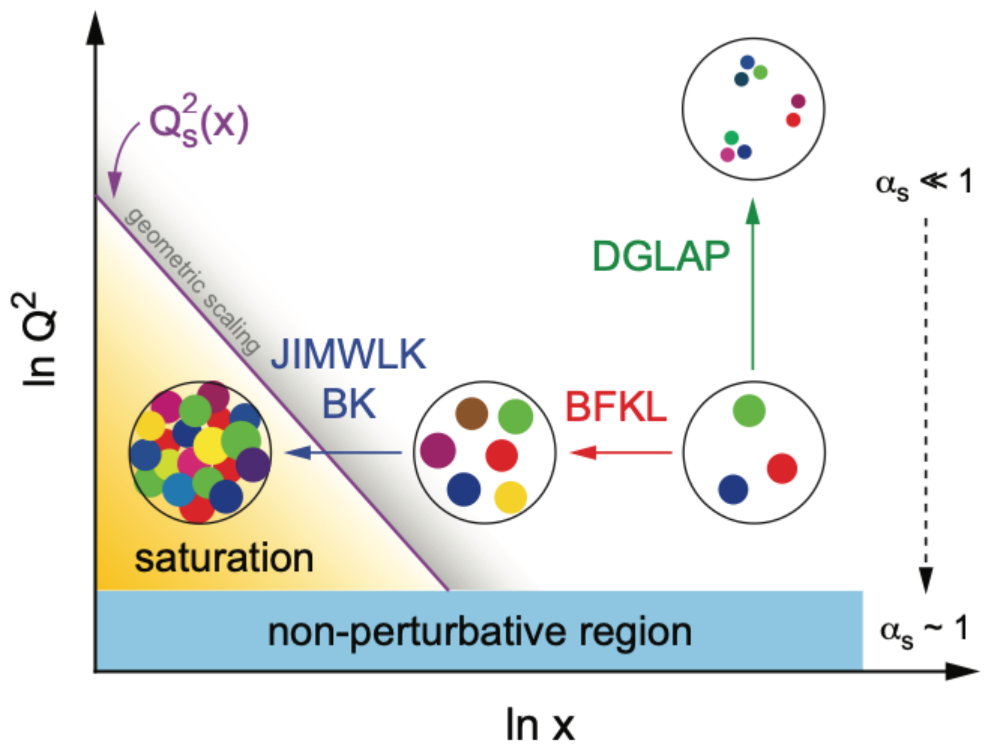
<u>Abundant</u>, <u>complex</u>, <u>complementary</u> data sets

Universality

Small-x evolution

Complementarity:

- Forward pA at LHC
- eA DIS at EIC



Physics of the ALICE Forward Calorimeter

Florian Jonas, Wed 11:40am

Factorization

Example: universality of jet fragmentation in QGP

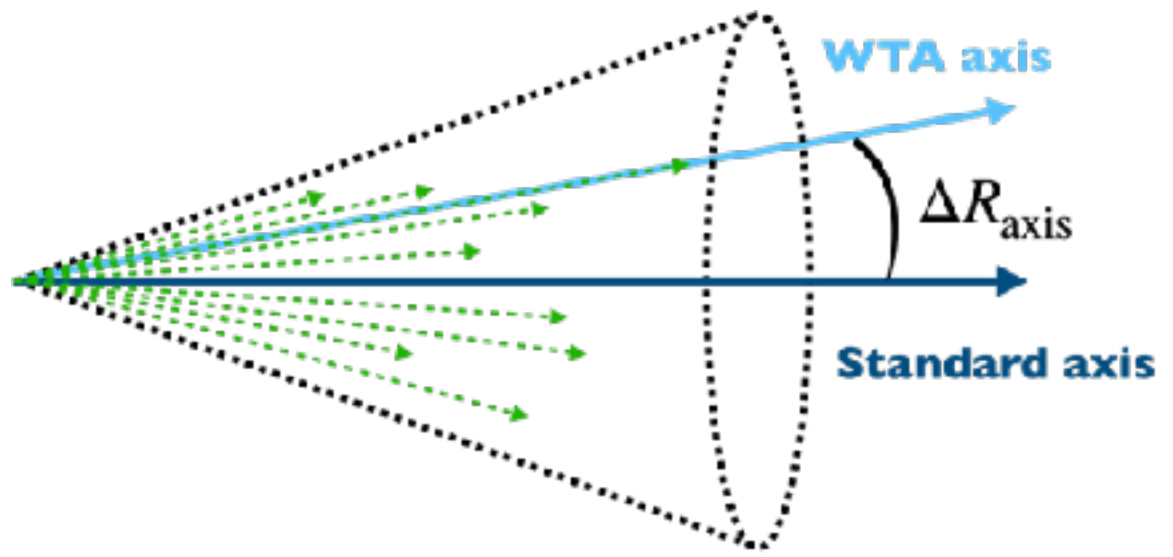
ALICE JHEP 05 245 (2023)
Qiu et al. PRL 122 (2019) 25

$$J_{r,\mathrm{med}}(z) = J_{\mathrm{med}}(z)$$

parton → subjet   parton → jet

Non-perturbative physics

Example: universality of CSS kernel (TMD fragmentation)



Drell-Yan vs. jet substructure

ALICE JHEP 07 (2023) 201

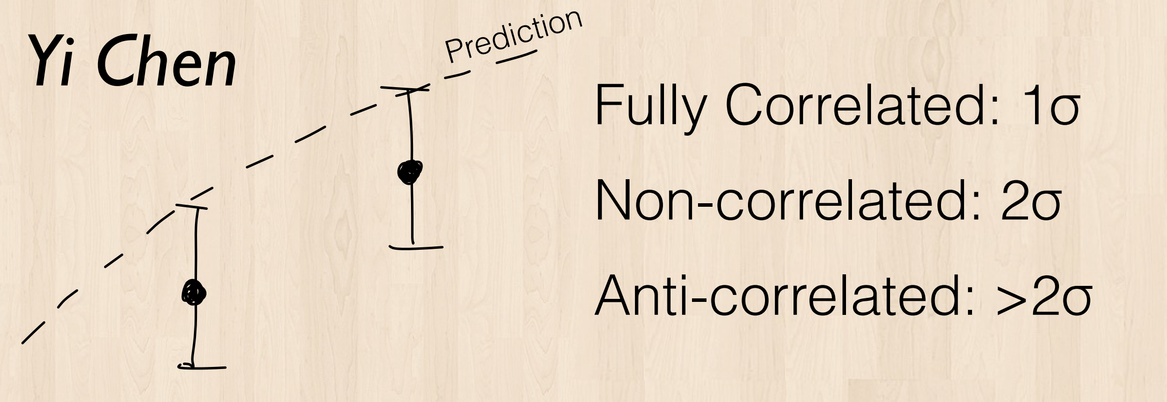
…

Statistical inference: rigorous universality tests across observables and collision systems

Current bottlenecks in Bayesian inference

Experiment

Report uncertainty correlations



Report **signed** unc. breakdowns in HEPData (or cov. matrix) [Example]

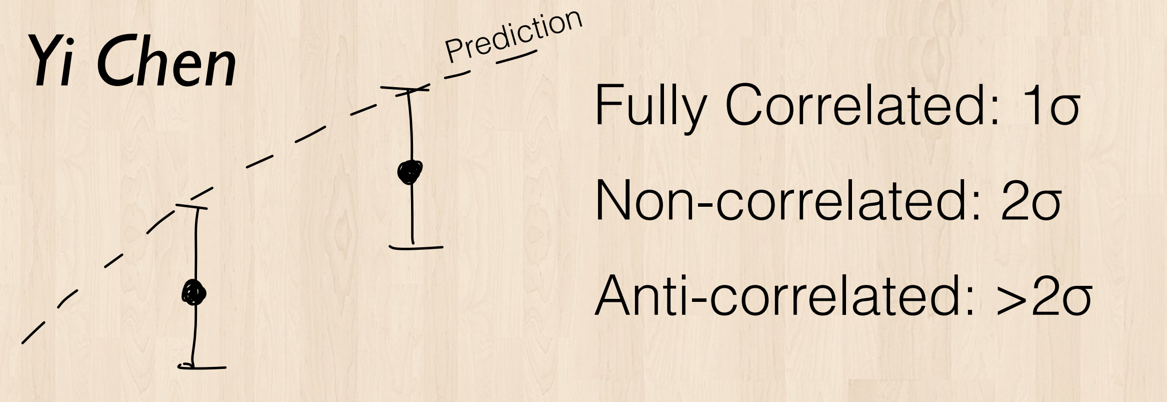
Easy but crucial — experiments should require this

Theory

Current bottlenecks in Bayesian inference

Experiment

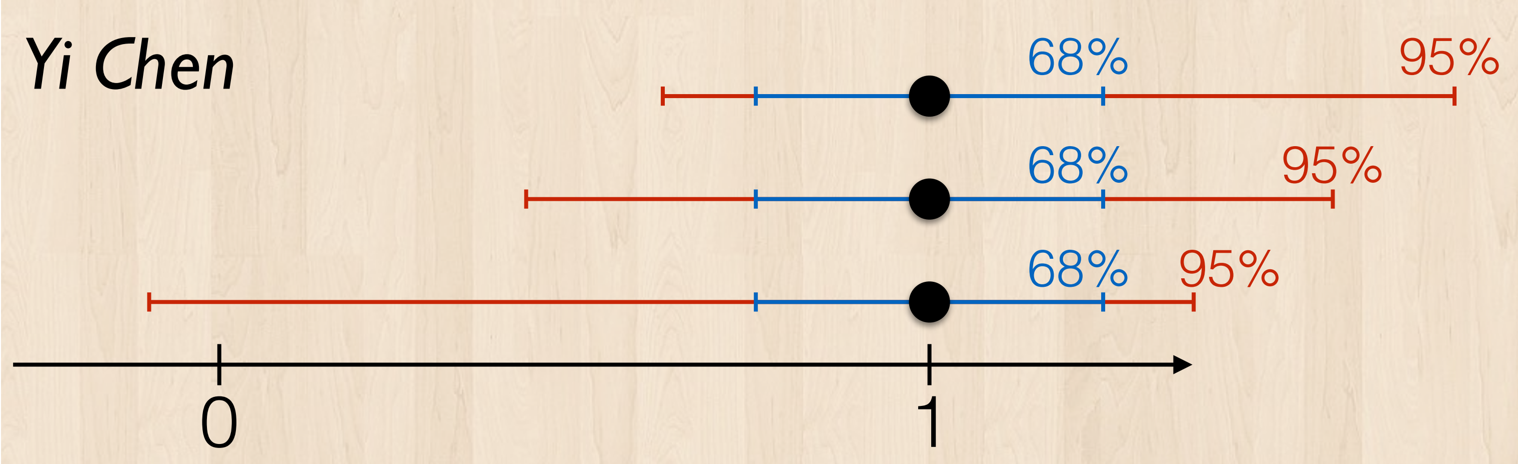
Report uncertainty correlations



Report **signed** unc. breakdowns in HEPData (or cov. matrix) [Example](#)

Easy but crucial — experiments should require this

Characterize shape of likelihood



Ideally: report more than 1σ interval

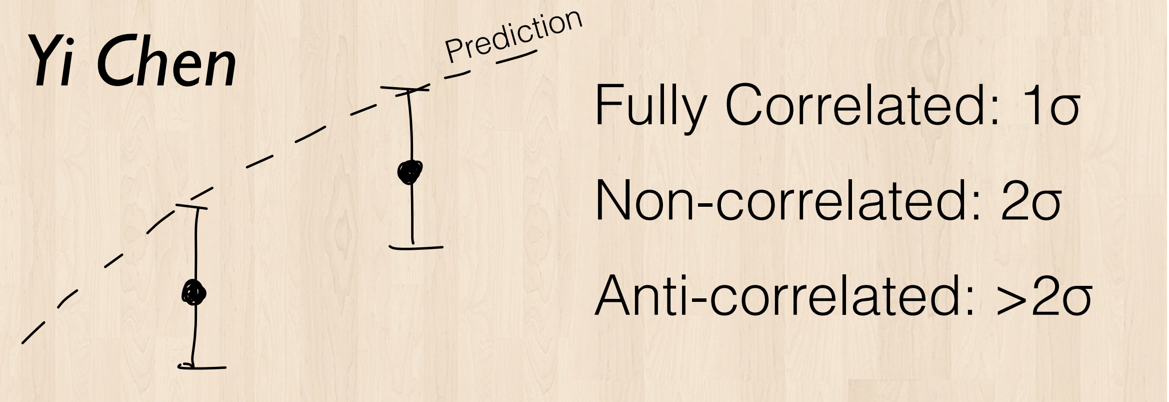
Should explore this for key observables

Theory

Current bottlenecks in Bayesian inference

Experiment

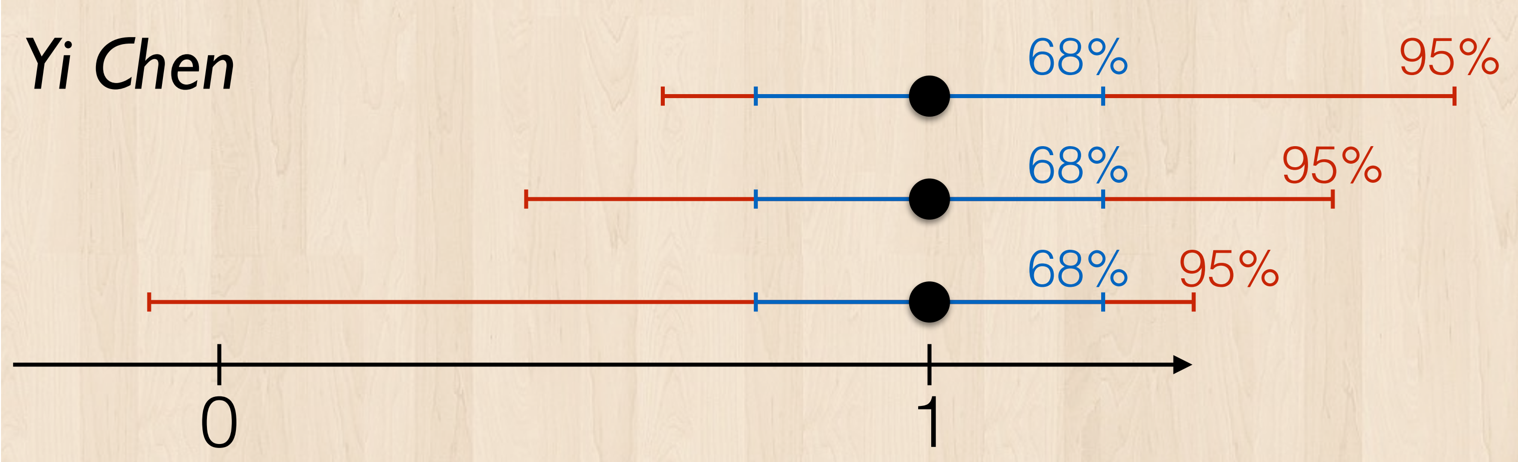
Report uncertainty correlations



Report **signed** unc. breakdowns in HEPData (or cov. matrix) [Example]

Easy but crucial — experiments should require this

Characterize shape of likelihood



Ideally: report more than 1σ interval

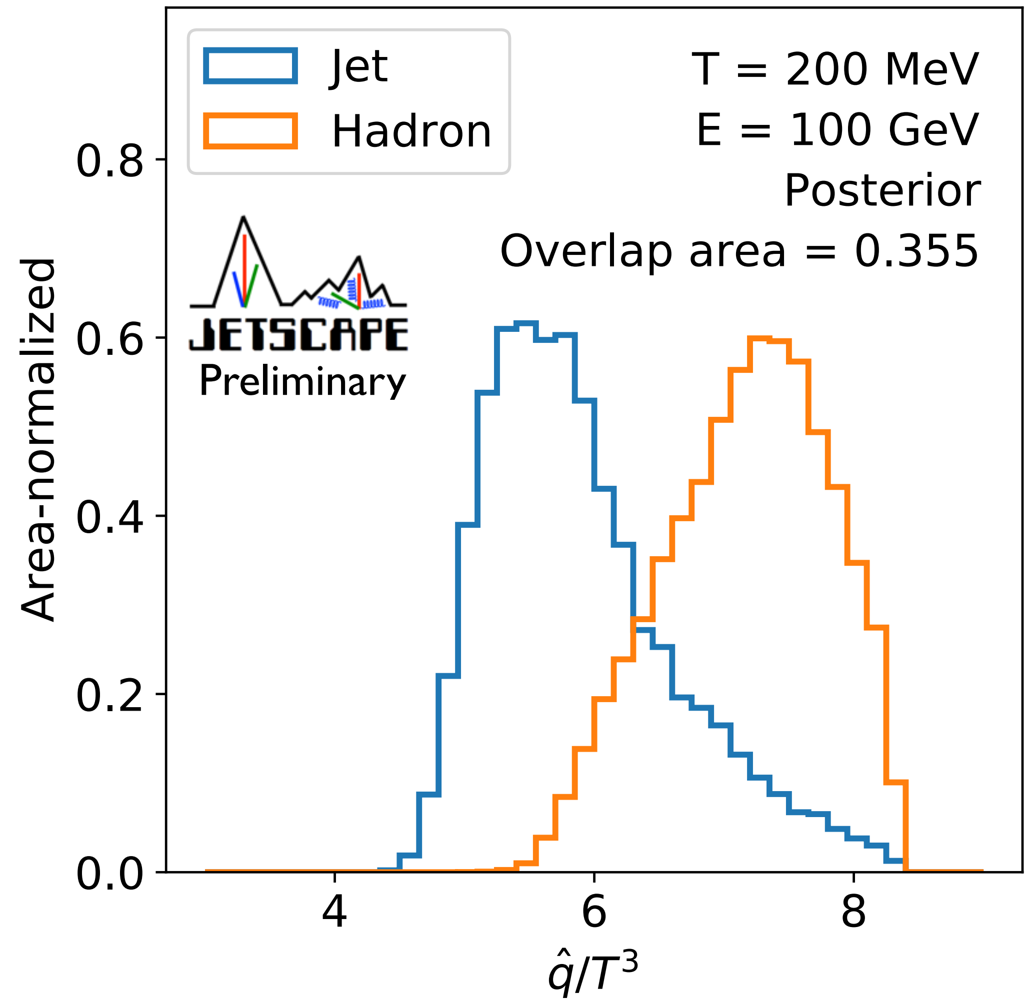
Should explore this for key observables

Theory

Model uncertainty

Goal: Model-independent QGP properties

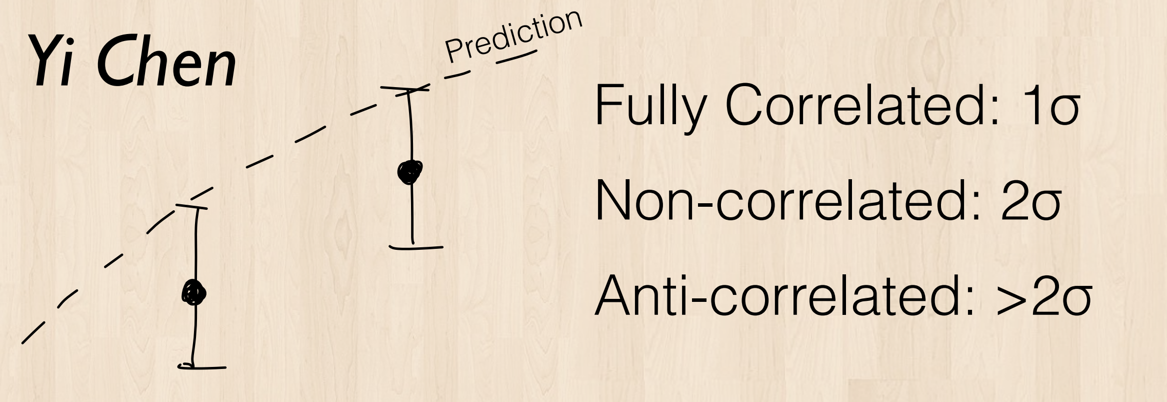
Requires: quantifiable model uncertainties



Current bottlenecks in Bayesian inference

Experiment

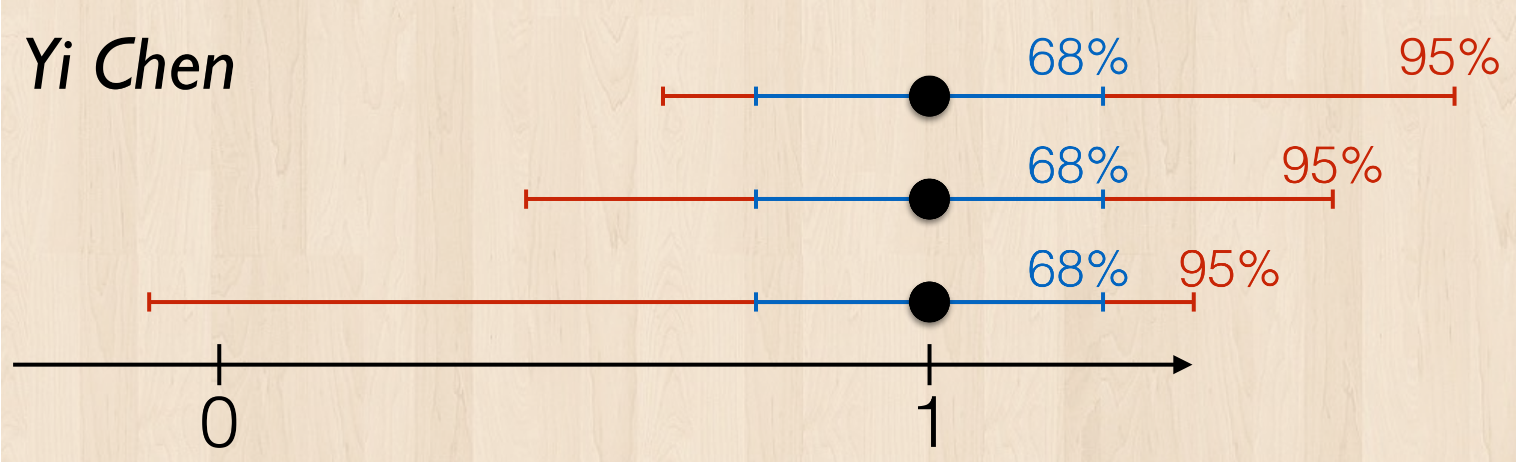
Report uncertainty correlations



Report **signed** unc. breakdowns in HEPData (or cov. matrix) [Example]

Easy but crucial — experiments should require this

Characterize shape of likelihood



Ideally: report more than 1σ interval

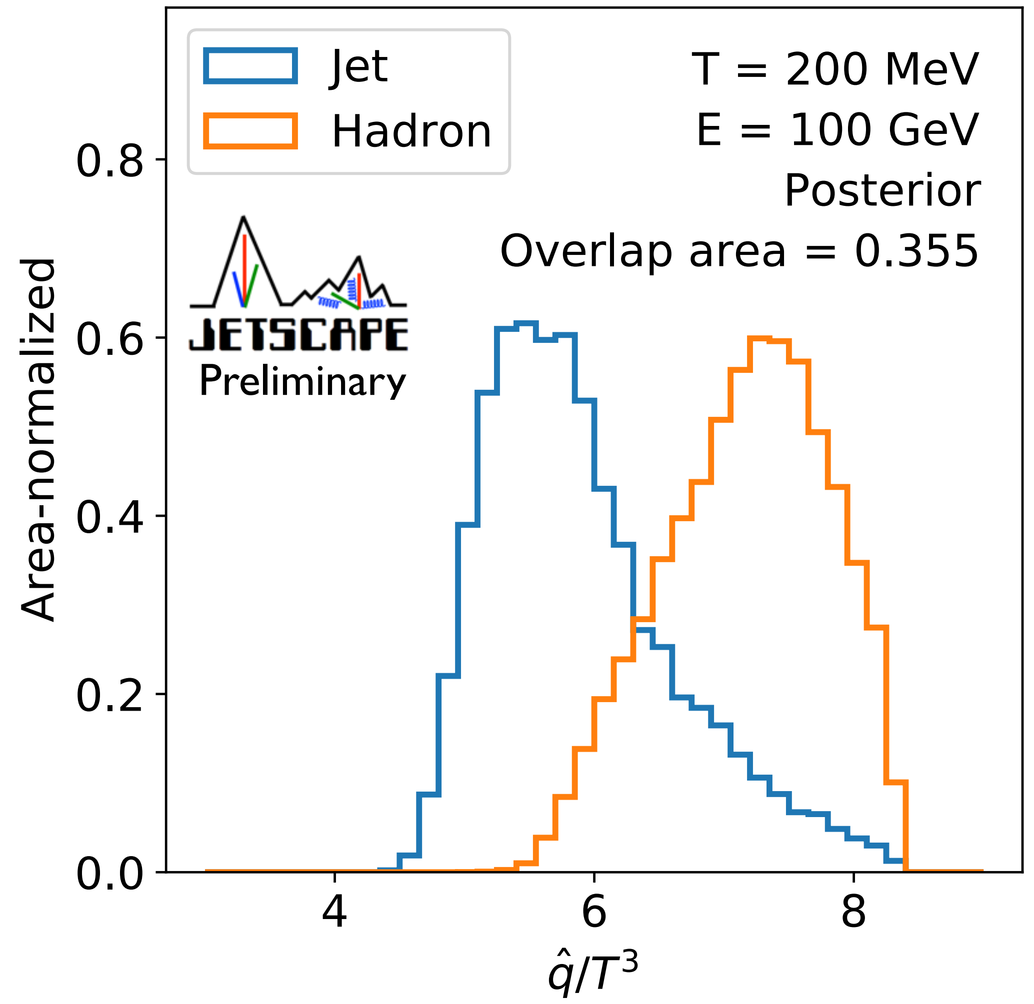
Should explore this for key observables

Theory

Model uncertainty

Goal: Model-independent QGP properties

Requires: quantifiable model uncertainties



Long-term: compare extracted quantity to calculated quantity

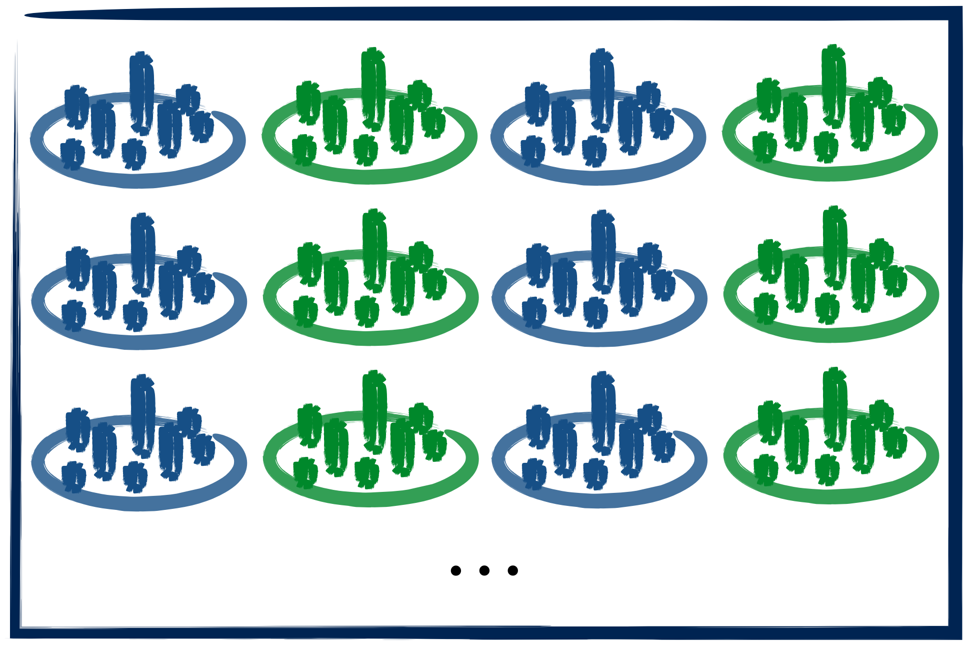
[Attila Pasztor, Mon 4:00pm] [Savage/Shanahan, Fri 12:00/3:30pm] …

Three approaches to learn about the QGP

1. Single observables

2. Sets of observables

3. Particle-level information



Learning from data

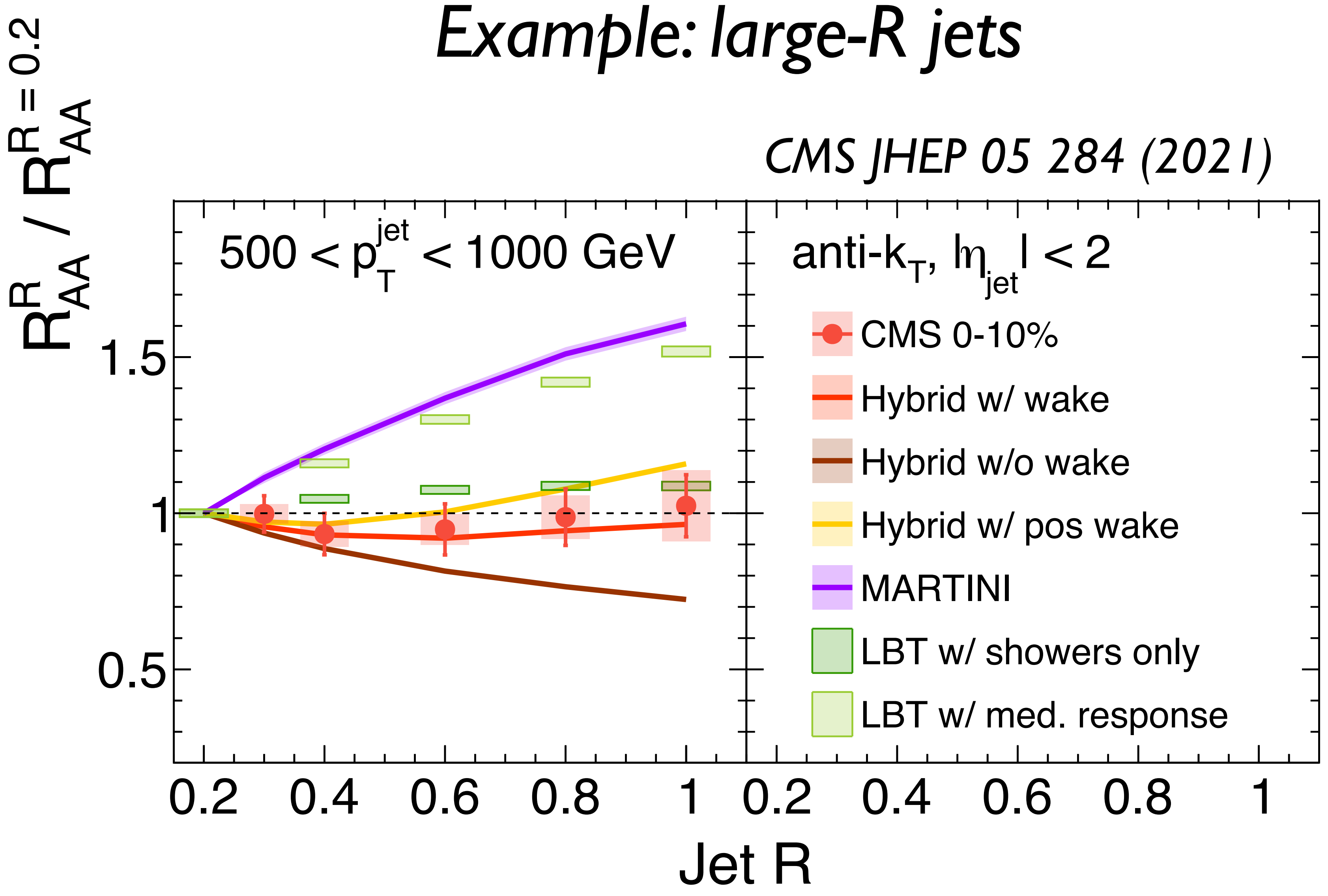
Can we guide the experimental program in a *model-independent* way?

Models are imperfect

- Real-time quantum dynamics
- Non-perturbative processes
- Parton shower approximations
- …

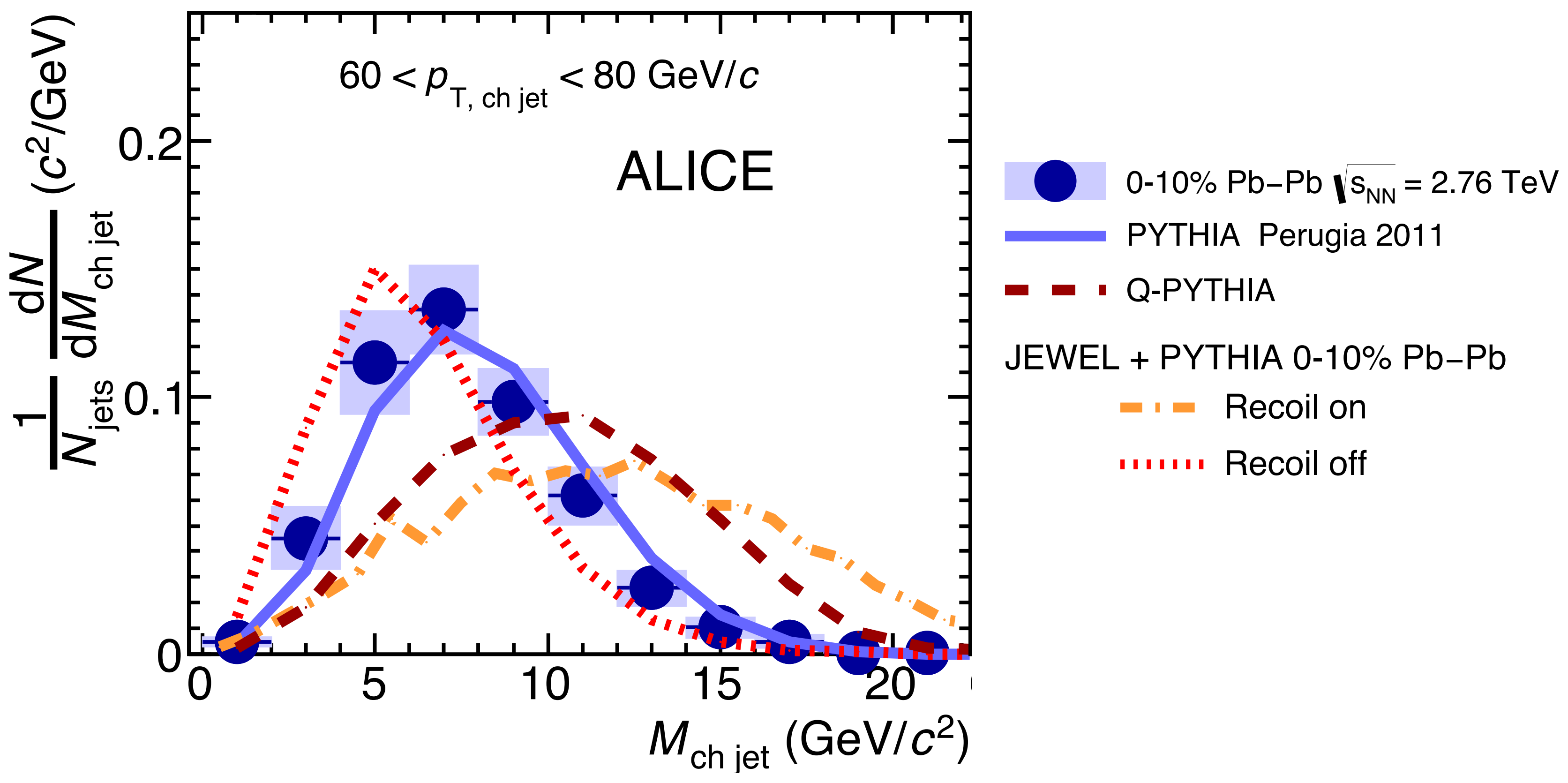
Example: large-R jets

CMS JHEP 05 284 (2021)



Example: jet mass

ALICE PLB 776 249 (2018)



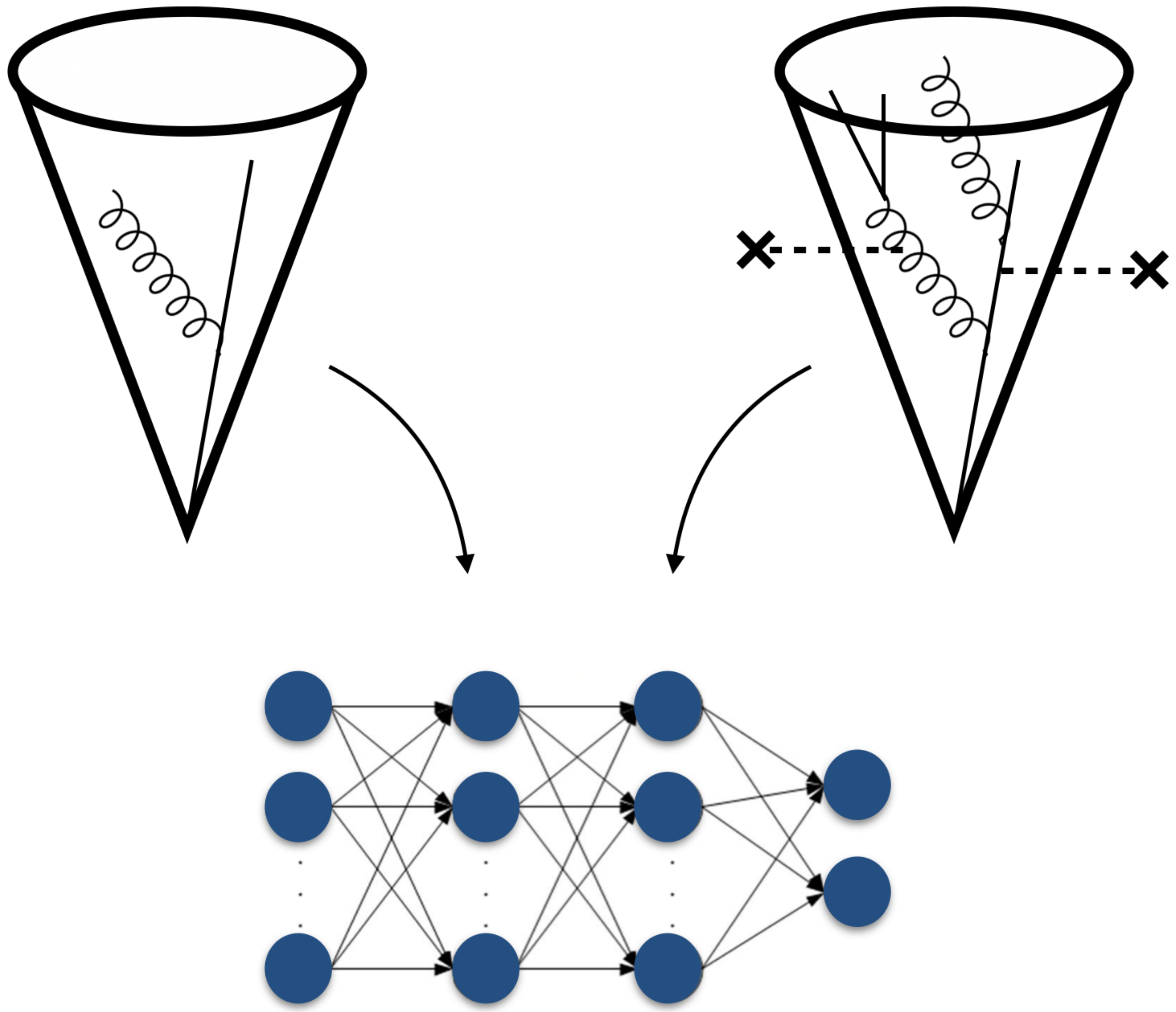
Learning from data

Chien, Elayavalli 1803.03589
Du, Pablos, Tywoniuk JHEP 03 206 (2021)
Apolinário et al. JHEP 11 219 (2021)
Lai et al. JHEP 10, 011 (2022)
Liu et al. JHEP 04 (2023) 140

The physics of many QGP signatures is encoded in the difference between ensembles of proton-proton and heavy-ion events

Learn a function that encodes the differences between proton-proton and heavy-ion events

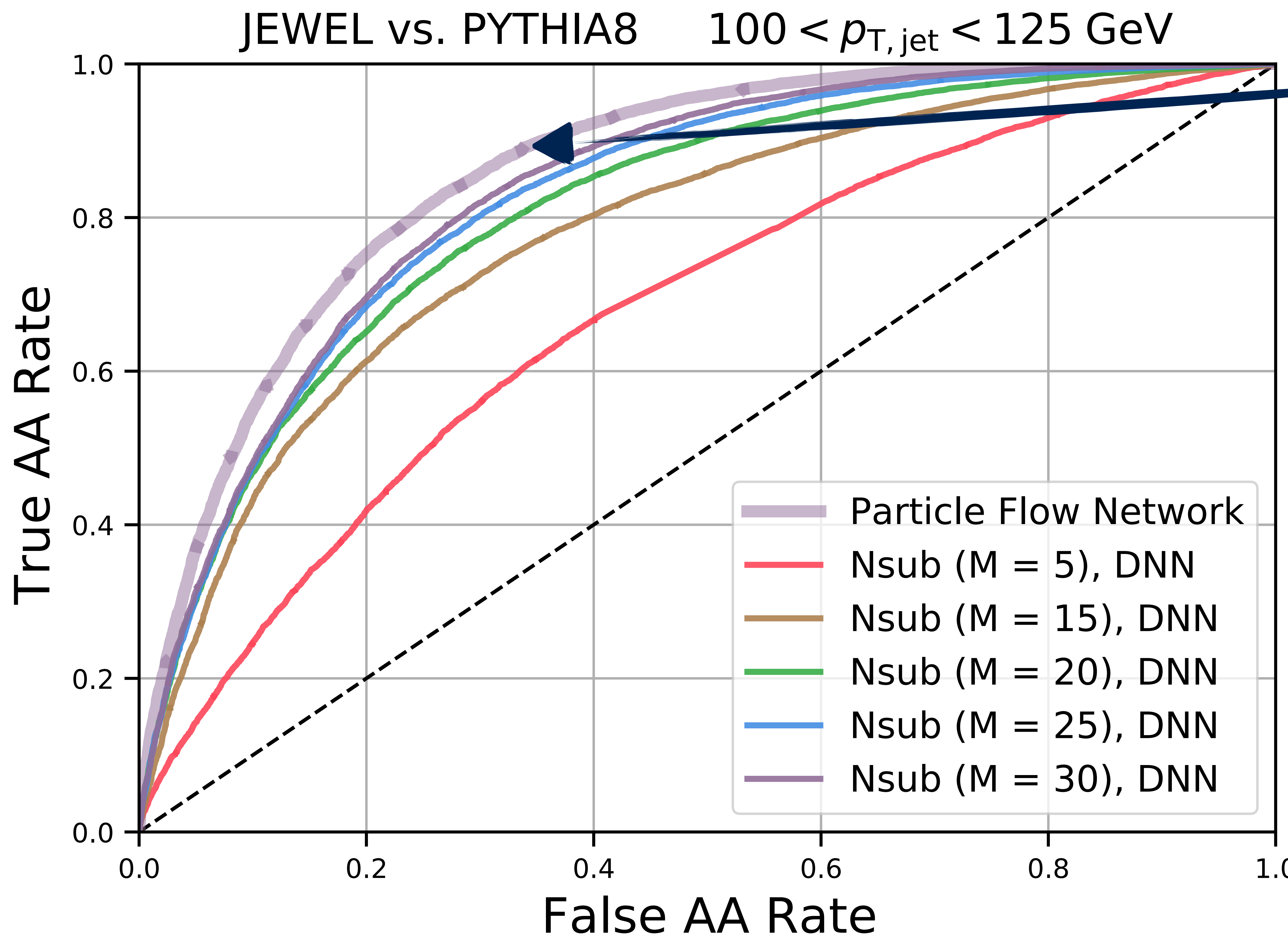
→ ML classifier trained **directly on experimental data**



Goal: Use ML to discriminate pp from AA events in a way that is **interpretable**

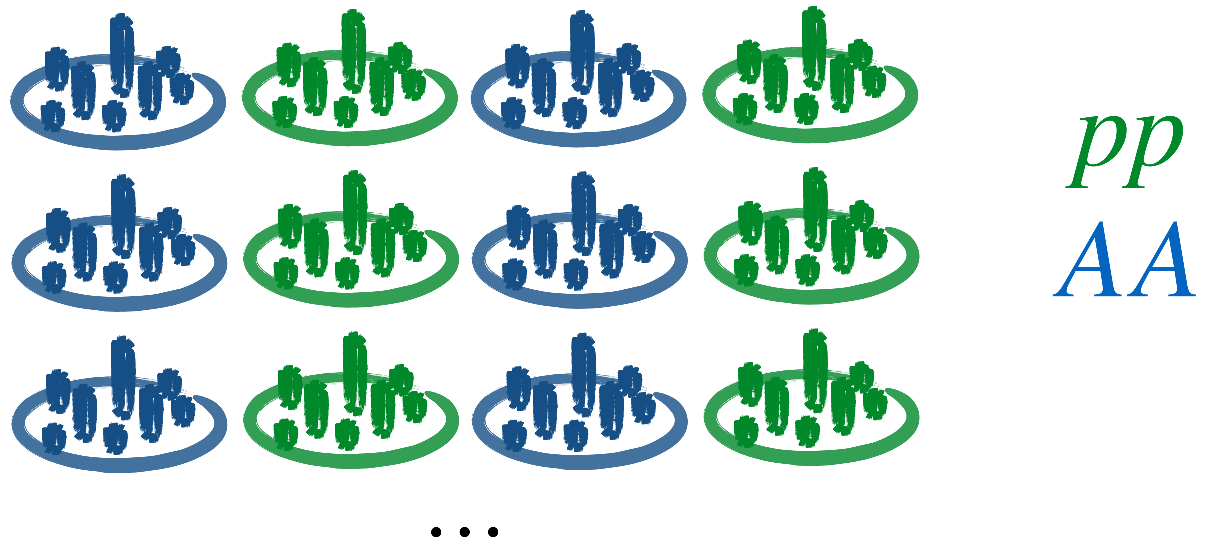
The information content of jet quenching

Lai, Mulligan, Płoskoń, Ringer JHEP 10, 011 (2022)



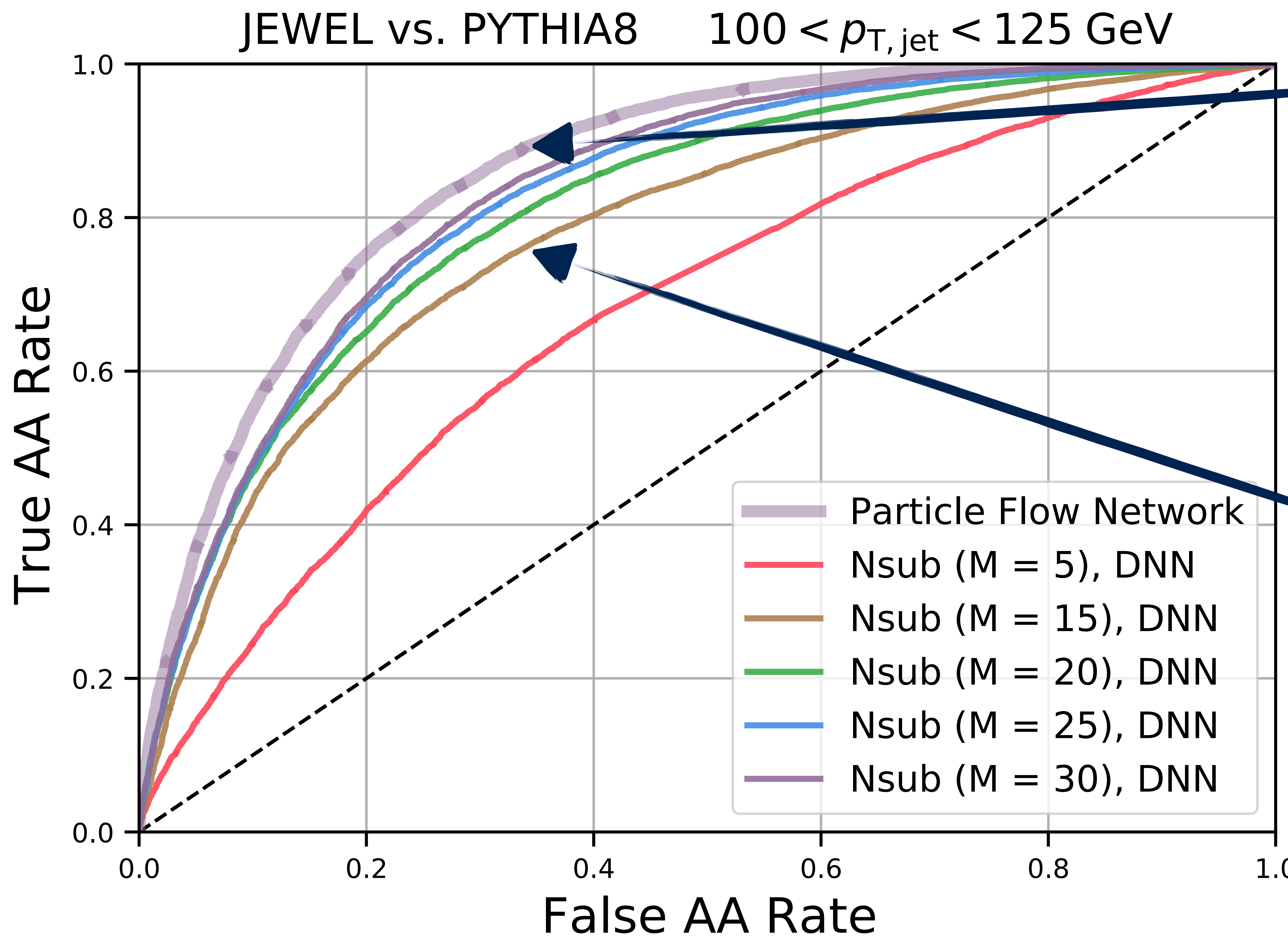
"Optimal" classifier

- Input: four-vectors of all jet particles



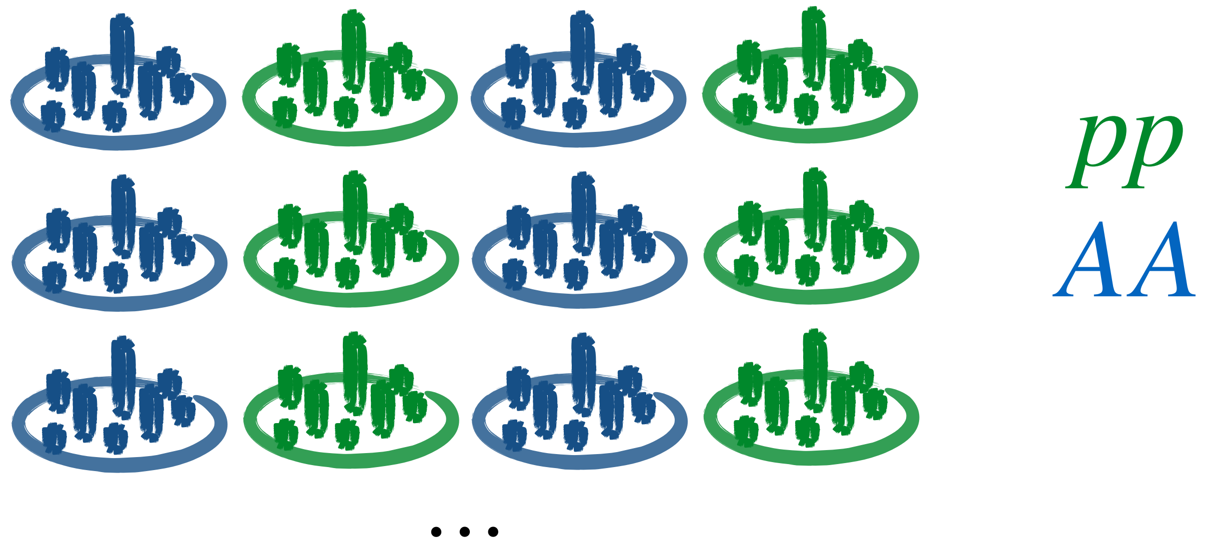
The information content of jet quenching

Lai, Mulligan, Płoskoń, Ringer JHEP 10, 011 (2022)

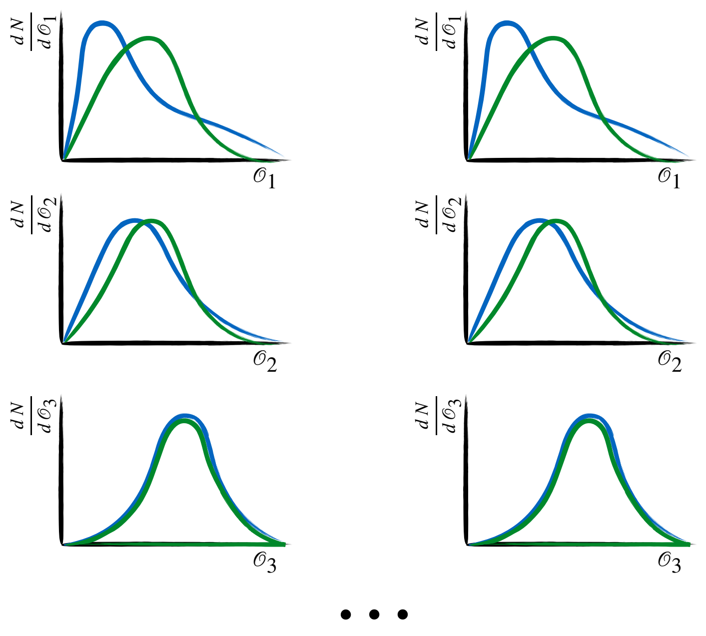


"Optimal" classifier

- Input: four-vectors of all jet particles



How many high-level observables do we need to measure?



See also:

Lu et al., *JHEP* 08 046 (2022)

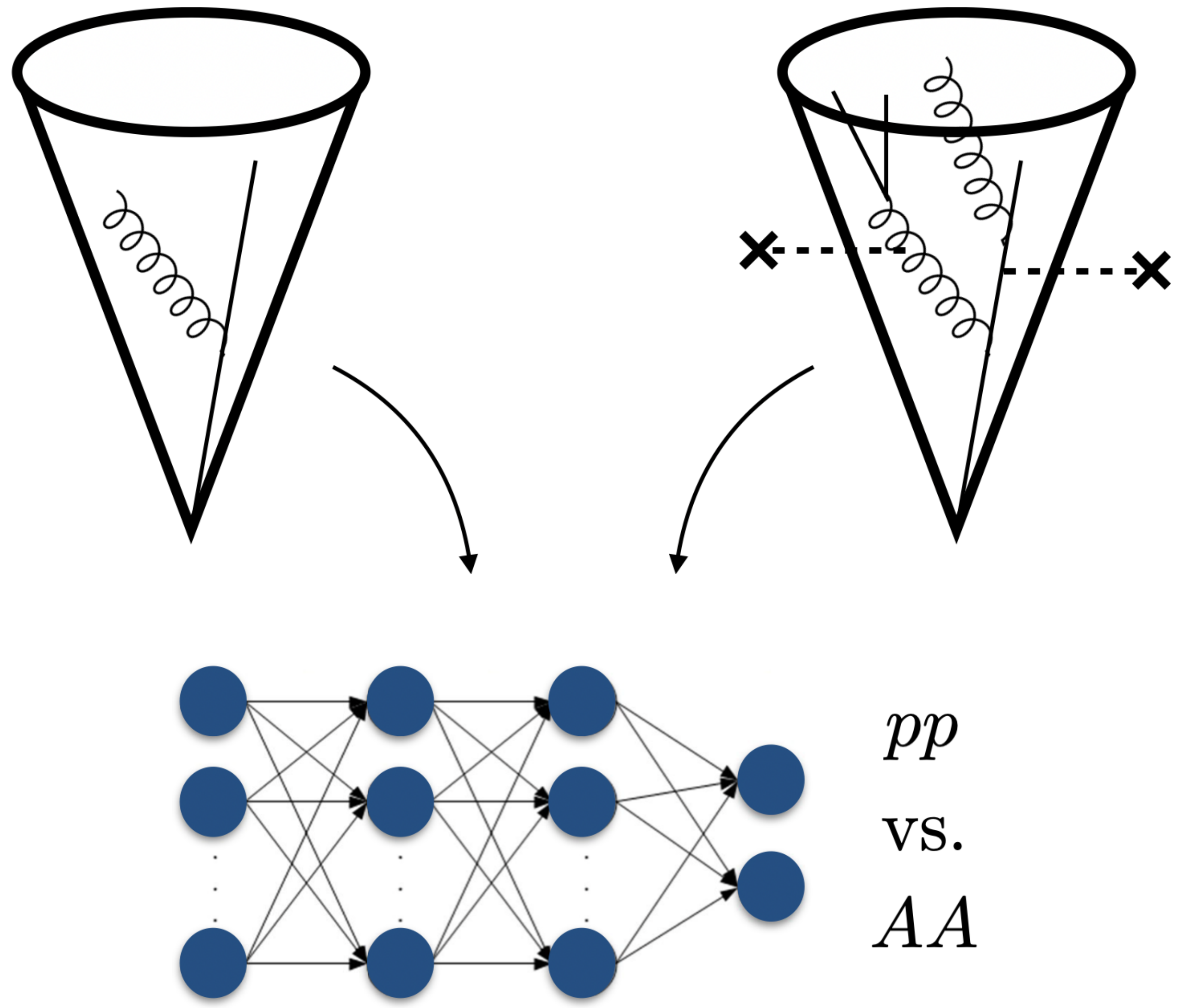
Romão et al. 2304.07196

Systematic approach: how many observables does one need to measure?

Experimental guidance from ML

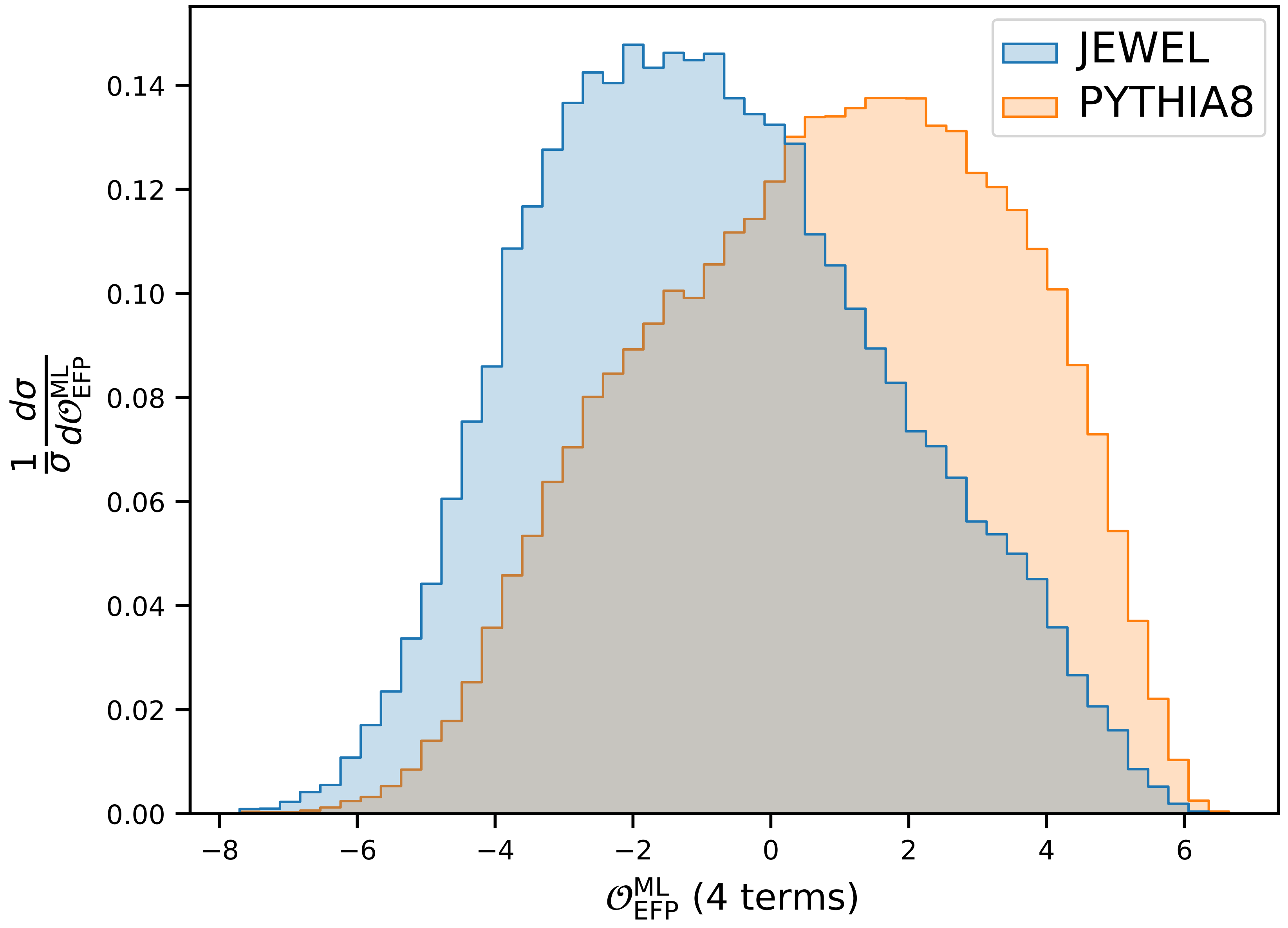
Lai, Mulligan, Płoskoń, Ringer JHEP 10, 011 (2022)

Design the **most strongly modified observable** that is **theoretically calculable**



$$\max_{\theta} \left| \frac{d\sigma_{AA}}{d\sigma_{pp}}(\theta) - 1 \right|$$

Symbolic regression



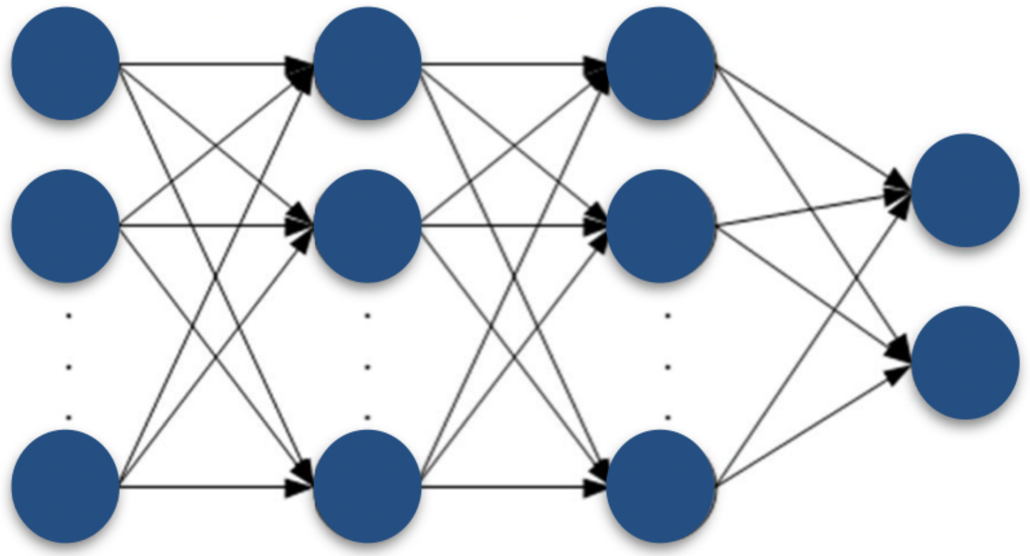
First step in a new paradigm: data-driven design of complete set of calculable observables

Complementary to Bayesian approach

Data-driven guidance across QCD systems

Lee, Mulligan, Płoskoń, Ringer, Yuan JHEP 03 (2023) 085

Cold nuclear matter



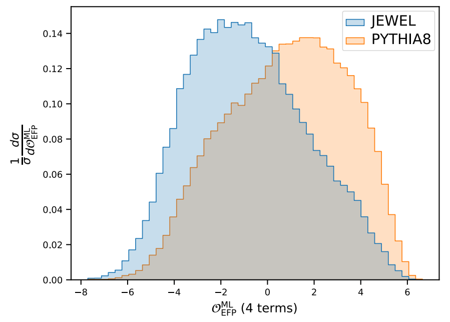
pA

vs.

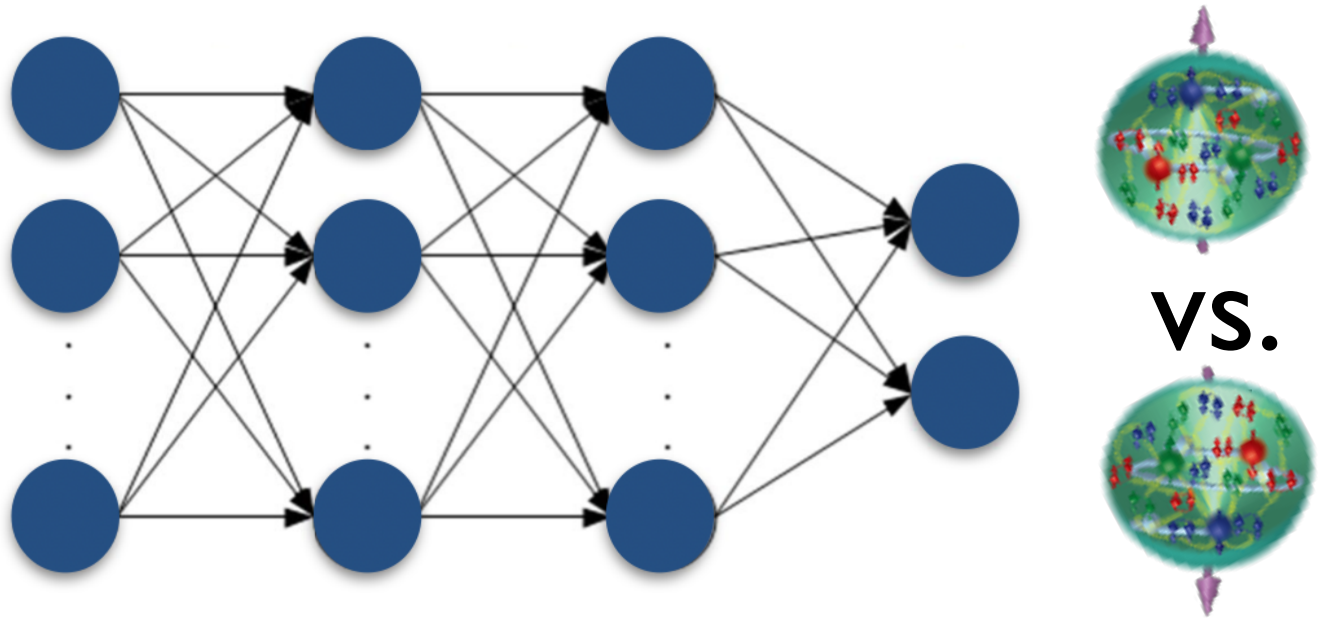
pp

Data-driven bound on jet modification in small systems

$$\max_{\theta} \left| \frac{d\sigma_{pA}}{d\sigma_{pp}}(\theta) - 1 \right| \longrightarrow$$



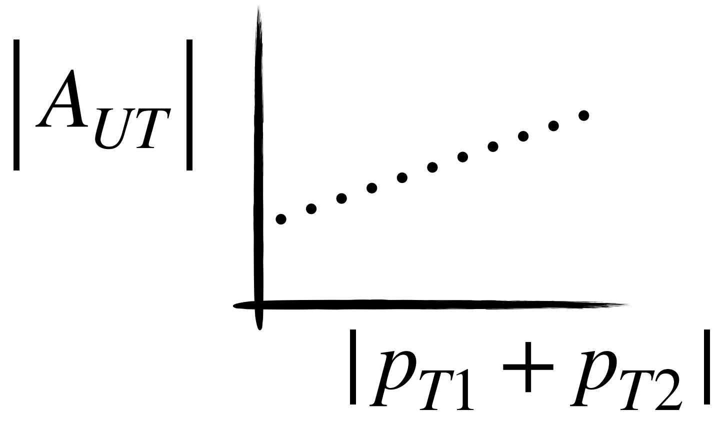
Spin physics



Data-driven observable design to maximize TSSAs — hadron structure

$$A_{UT} = \frac{\mathrm{d}\sigma^{\uparrow} - \mathrm{d}\sigma^{\downarrow}}{\mathrm{d}\sigma^{\uparrow} + \mathrm{d}\sigma^{\downarrow}}$$

$$\max_{\theta} \, |A_{UT}(\theta)| \longrightarrow$$



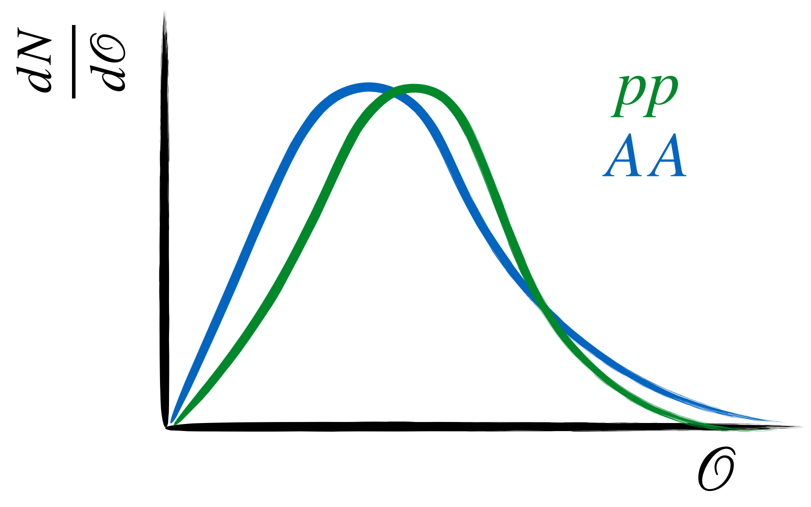
Can apply these directly on experimental data today at RHIC and LHC — and the future EIC

Summary

We have vast freedom in what we choose to measure at colliders in order to elucidate emergent behaviors of QCD — are we fully exploiting the data sets we have in hand?

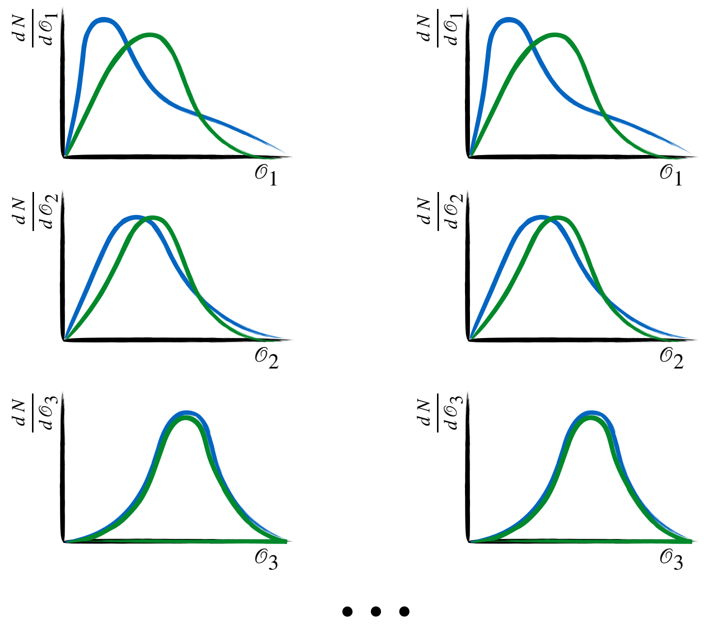
Single observables

Experimental and theoretical building blocks



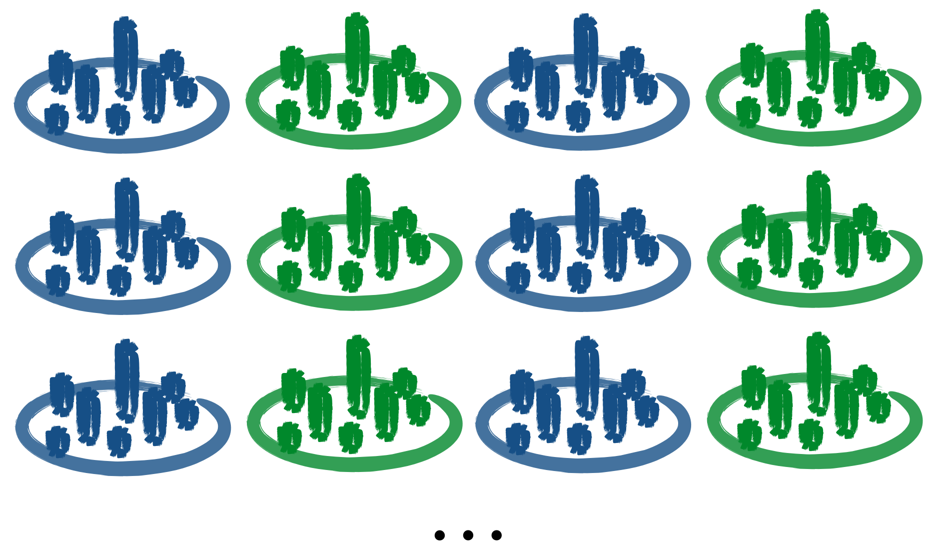
Sets of observables

Rigorous, model-dependent inference of QCD properties



Particle information

Model-independent learning directly from experimental data



A new era: systematic, iterative design of sets of experimental analyses