Tools for Precision Jet Substructure Measurements

Unfolding and Tracking

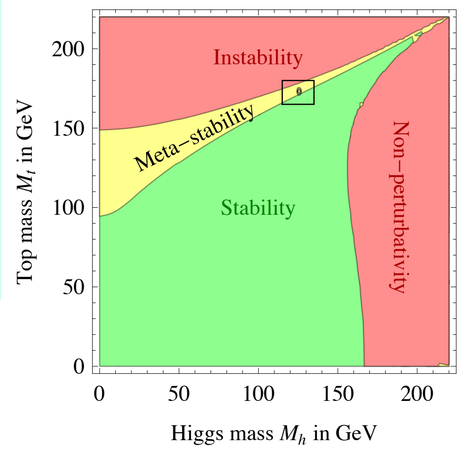
Vinicius Mikuni[1], Ian Moult[2], Ben Nachman[1], **Jingjing (Jing) Pan**[2]

[1] Lawrence Berkeley National Lab
[2] Yale University

Why *Precision* *JSS Measurements*?

"critical point" precision wishlist

- m_H: 0.2 GeV
- m_t: 60 MeV
- α_s: 0.14%

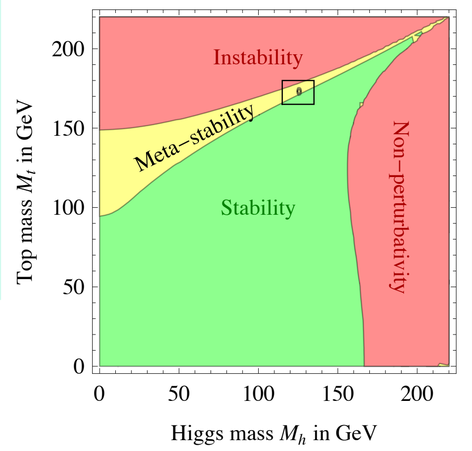


Metastability of the Universe

Why *Precision* *JSS Measurements*?

"critical point" precision wishlist

- m_H: 0.2 GeV
- m_t: 60 MeV
- α_s: 0.14%



Metastability of the Universe

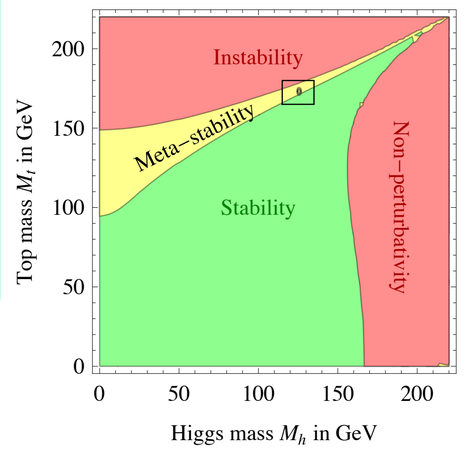
Enhance sensitivity of searches

- **Test** improved MC for jet modeling
- More **robust observables** for tagging (e.g. jet mass)

Why *Precision* *JSS Measurements*?

"critical point" precision wishlist

- m_H: 0.2 GeV
- m_t: 60 MeV
- α_s: 0.14%



- Test improved MC for jet modeling
- More robust observables for tagging (e.g. jet mass)

Metastability of the Universe

Enhance sensitivity of searches

Test the SM and explore emergent properties of QCD

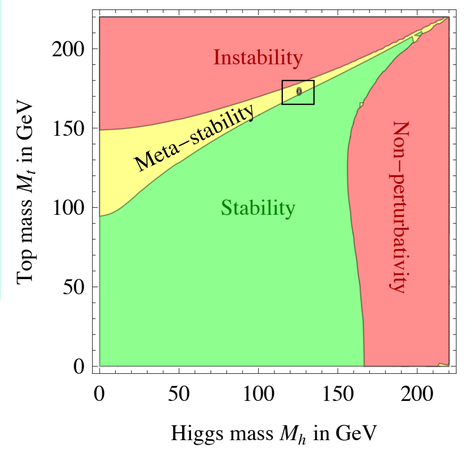
Fundamental subtle questions about QCD

SMEFT / future reinterpretation

Why *Precision* *JSS Measurements*?

"critical point" precision wishlist

- m_H: 0.2 GeV
- m_t: 60 MeV
- α_s: 0.14%



- Test improved MC for jet modeling
- More robust observables for tagging (e.g. jet mass)

Metastability of the Universe

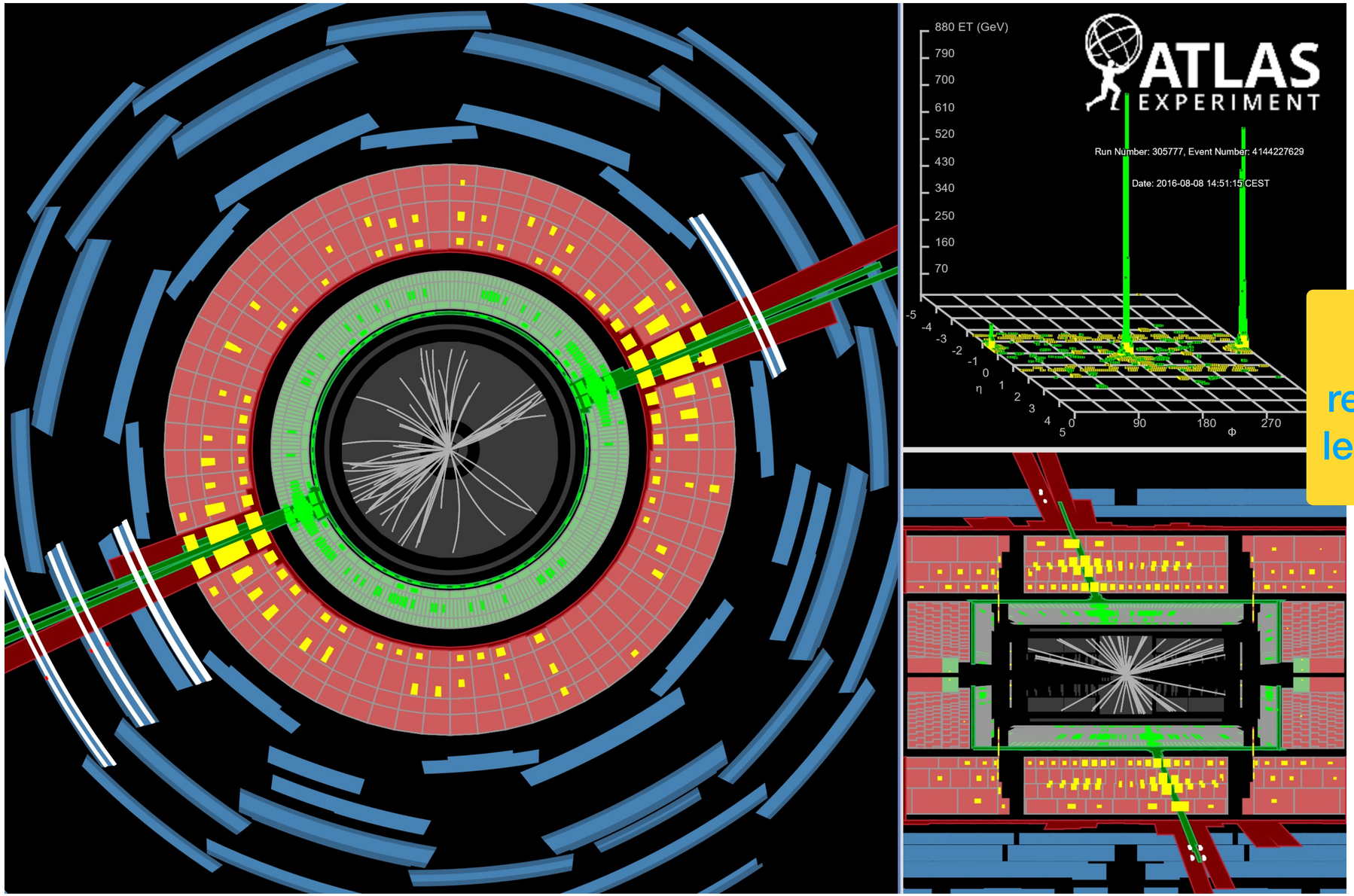
Enhance sensitivity of searches

Test the SM and explore emergent properties of QCD

Fundamental subtle questions about QCD

SMEFT / future reinterpretation

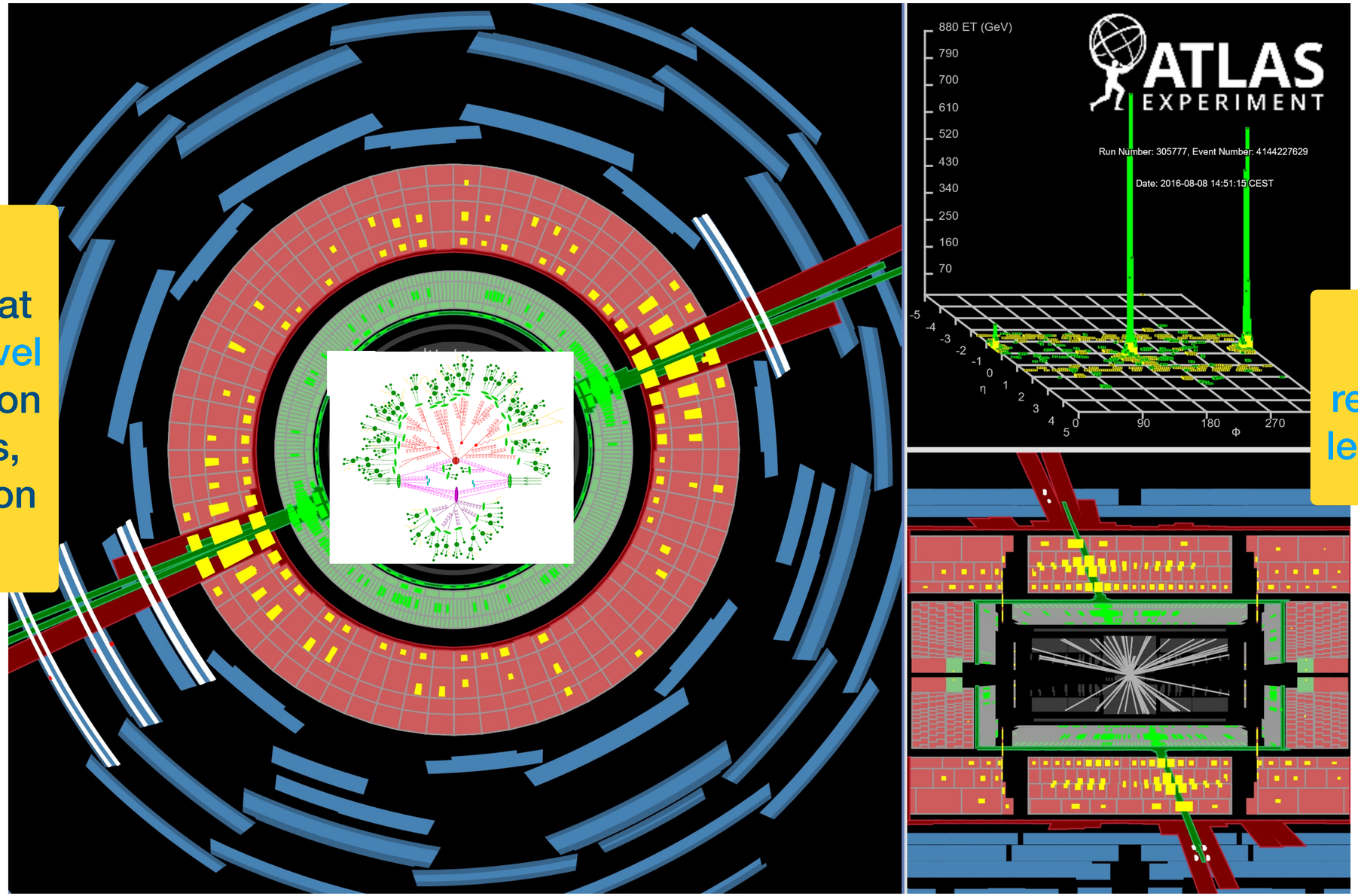
How to Get There?



We access reconstruction level quantities

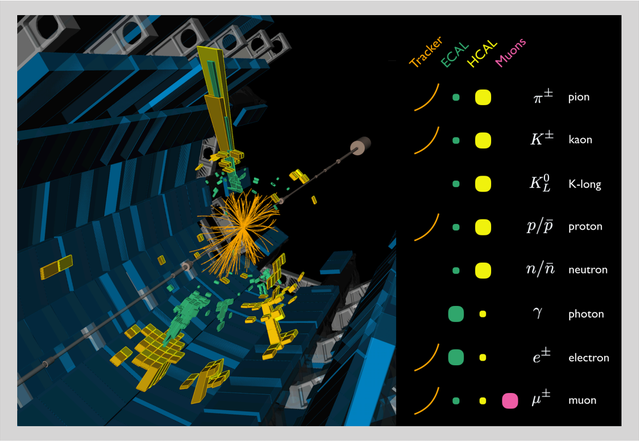
How to Get There?

But need observables at generation level (for comparison with theories, new calibration routines)

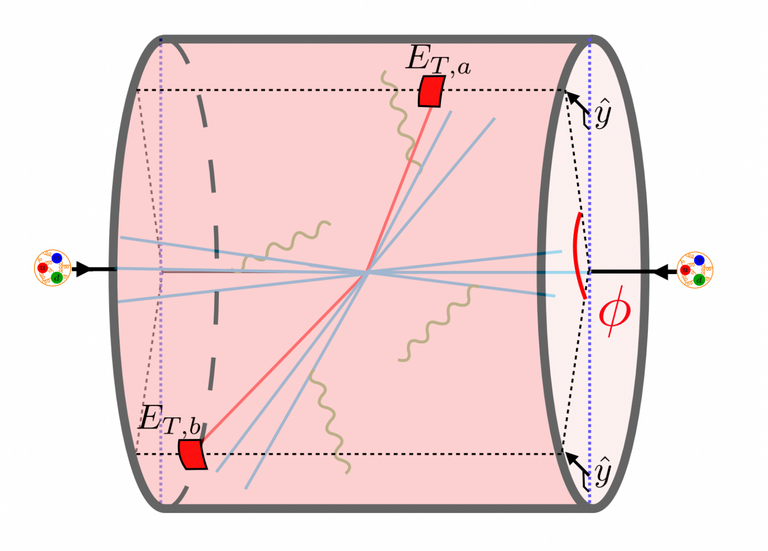


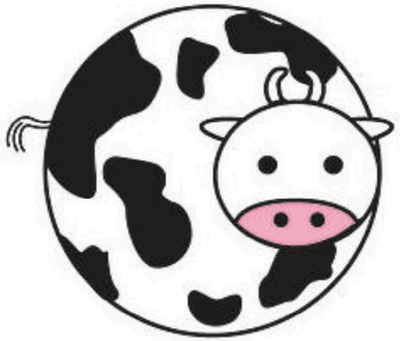
We access reconstruction level quantities

Unfolding: An Inverse Problem



Experiment





Theory

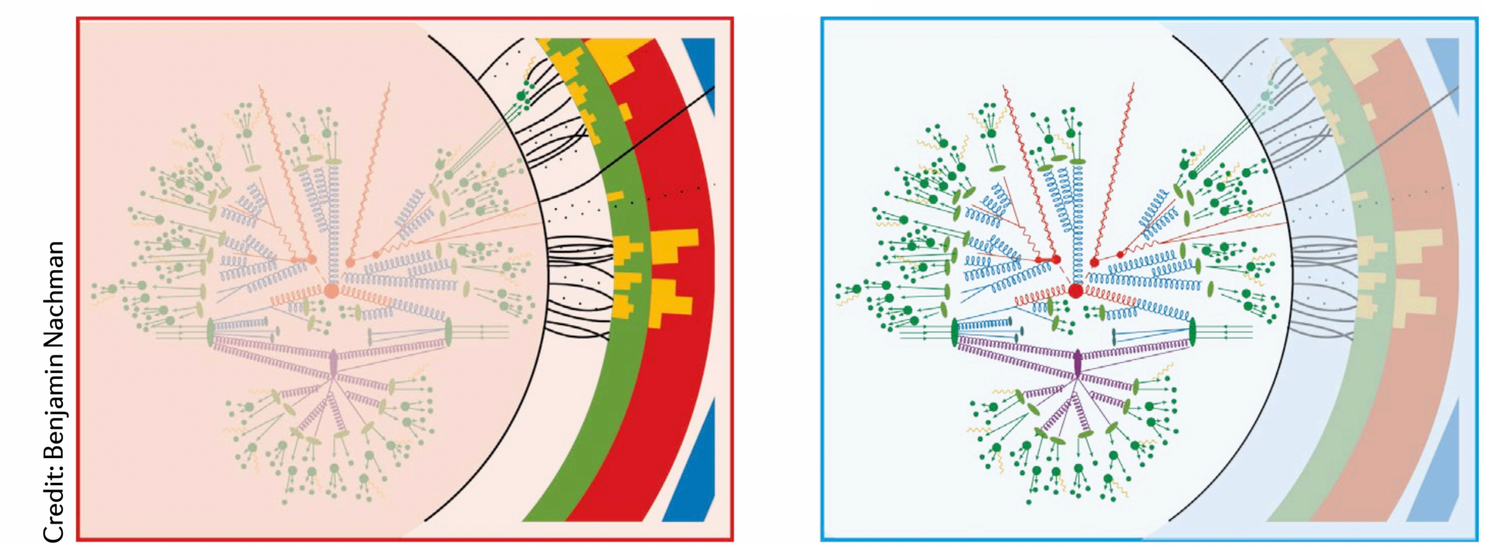
Preserving New Physics while Simultaneously Unfolding All Observables (PRD, 2021)

Outline

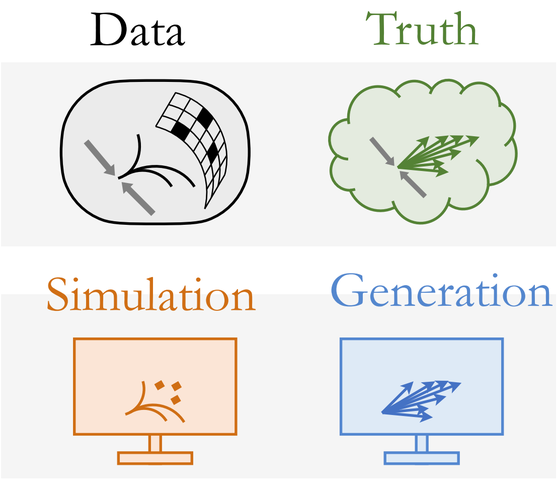
- **Seeing Higher-Order QCD Effects with OmniFold**

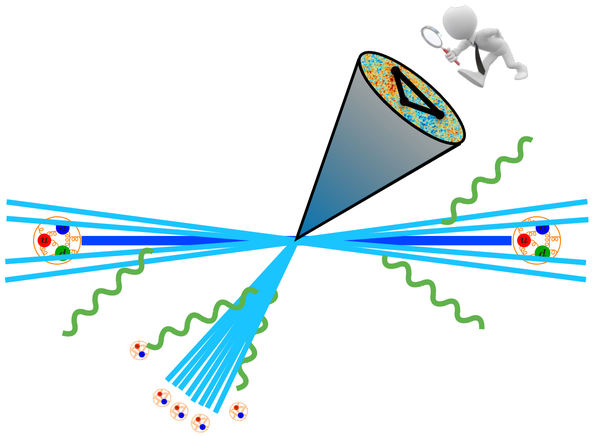
- **All About Precision:**

  Putting JSS on Track(s)



Credit: Benjamin Nachman



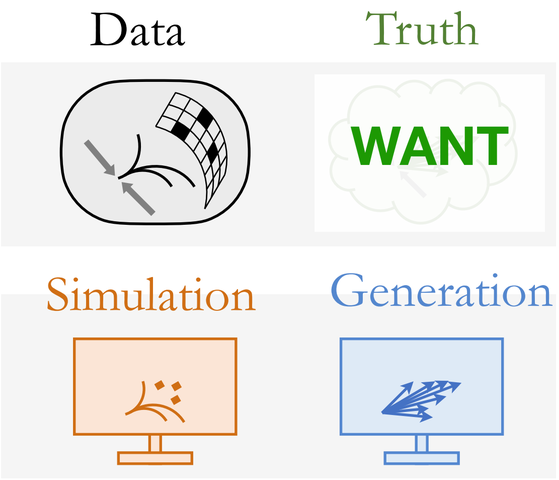


Outline

- **Seeing Higher-Order QCD Effects with OmniFold**

- **All About Precision: Putting JSS on Track(s)**

Credit: Benjamin Nachman



Unfolding

Data inevitably comes with many stochastic defects:

- Acceptance & Efficiency: particles produced not measured
- Detector Noise (“fake”): particles measured not real
- Detector Bias & Resolution
- Combinatorics: detector can change the order of N particles
- Background

Unfolding

Data inevitably comes with many stochastic defects:

- Acceptance & Efficiency: particles produced not measured
- Detector Noise (“fake”): particles measured not real
- Detector Bias & Resolution
- Combinatorics: detector can change the order of N particles
- Background

Challenge: correct all the above accurately & precisely

→ preserve higher order effects → test improvements of jet modeling

A Timely Example: Testing Improvements of Jet Modeling

α_s extraction from EEC

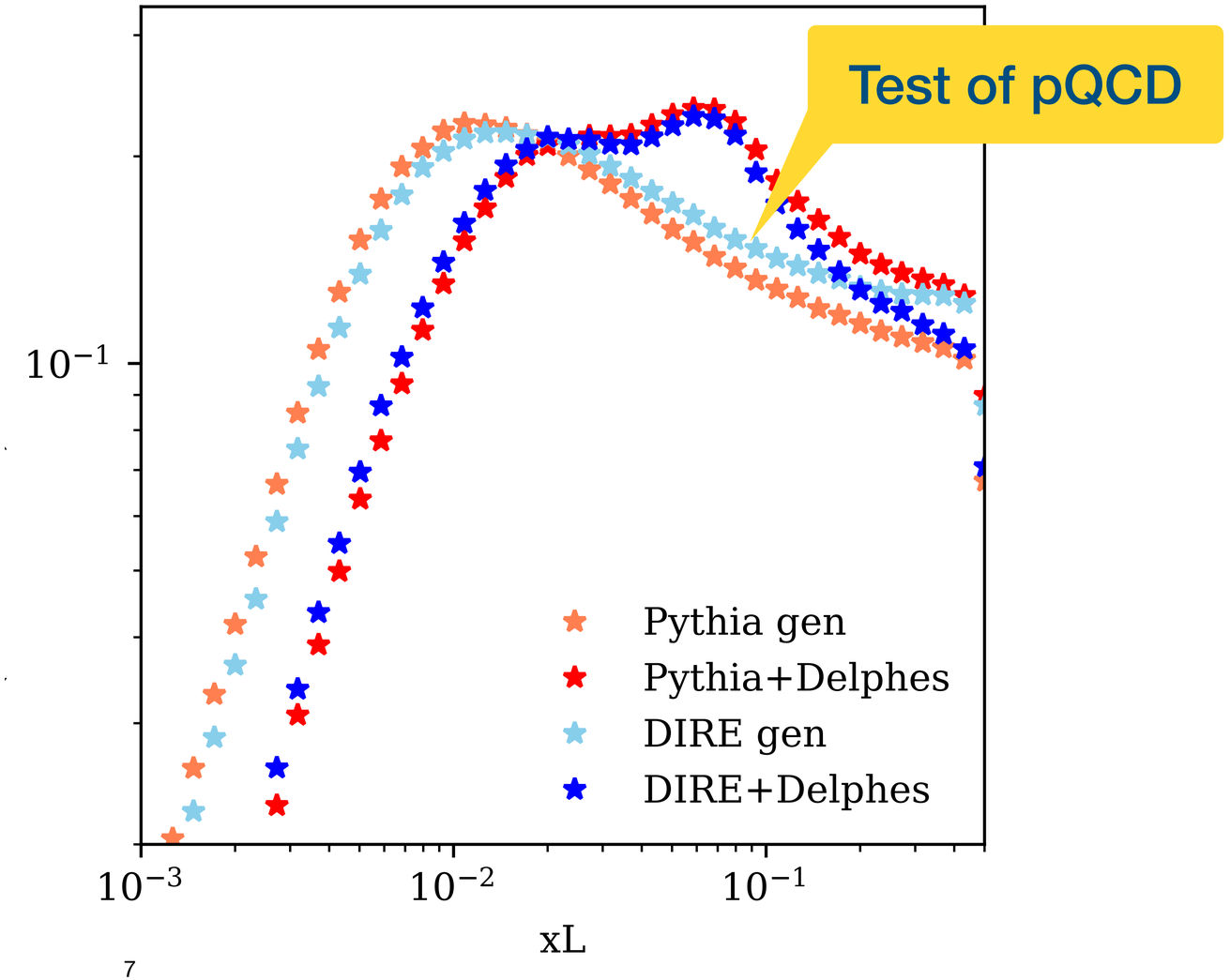
DIRE: **NLO DGLAP** added in the parton shower

(Collinear) Energy Correlators: **field-theoretic, broad appeal** expect better modeling by DIRE

$$\langle \mathcal{E}\mathcal{E} \rangle = \sum_{ij} \int \mathrm{d}x'_L \, \frac{p_{\mathrm{T,i}} p_{\mathrm{T,j}}}{p^2_{\mathrm{T,jet}}} \, \delta(x'_L - x_L)$$

(~4.4 % **NLO effects** in the slope proportional to α_s)

DIRE vs. Pythia at generation-level vs. after Delphes



Testing Improvements of Jet Modeling

Setup

Jet definition: AK4, Energy Flow algorithm from Delphes 3.5.0

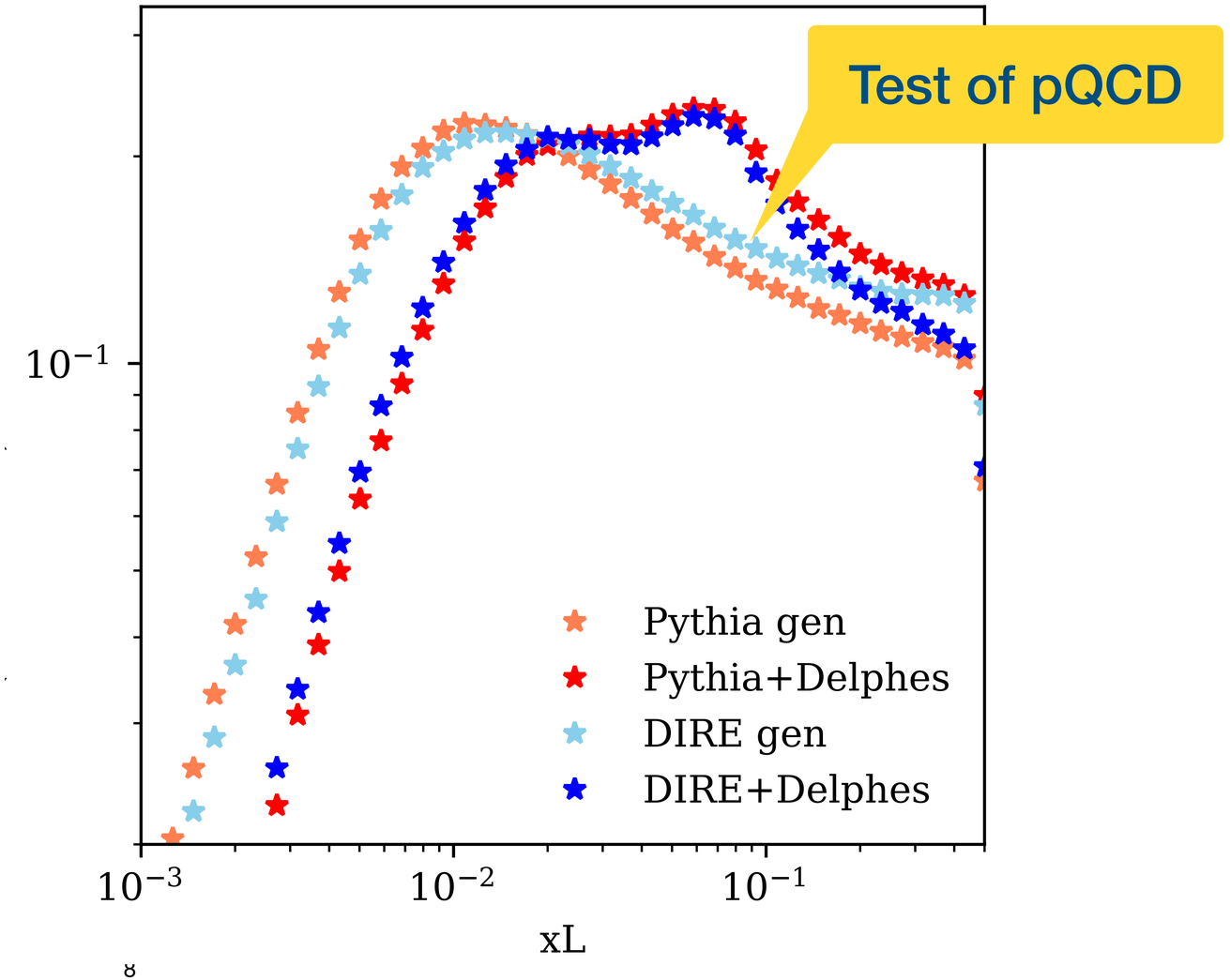
Selection: $p_{\mathrm{T,jet}} \in [500,\ 550]$ GeV, $\eta < 2.5$, the leading jet from Z+jets

Input: 4-vectors ($p_{\mathrm{T}}, \eta, \phi$, PID) of all the hadrons within the jet

Neural network architecture: PFN, Point Cloud Transformer (PCT)

https://arxiv.org/abs/2012.09688

DIRE vs. Pythia at generation-level vs. after Delphes



[A. Andreassen et al., PRL 2020]

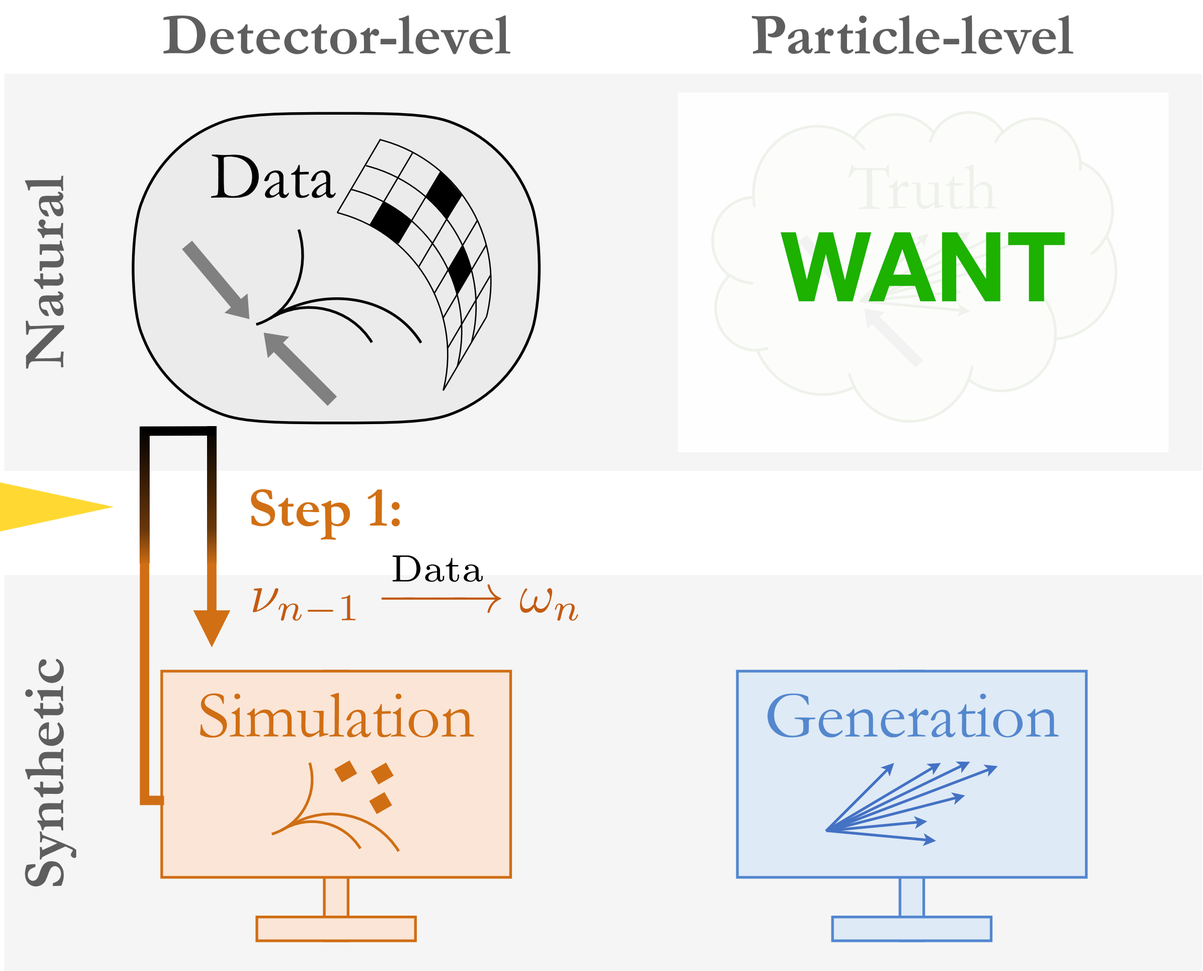
OmniFold

Likelihood-Free Inference

- Learn this likelihood ratio by training a classifier on reco-level MC vs. data:

$$W(\text{reco}) = \frac{p_{\text{data}}(\text{reco})}{p_{\text{MC}}(\text{reco})}$$

- Reweight reco-level MC to match data



[A. Andreassen et al., PRL 2020]

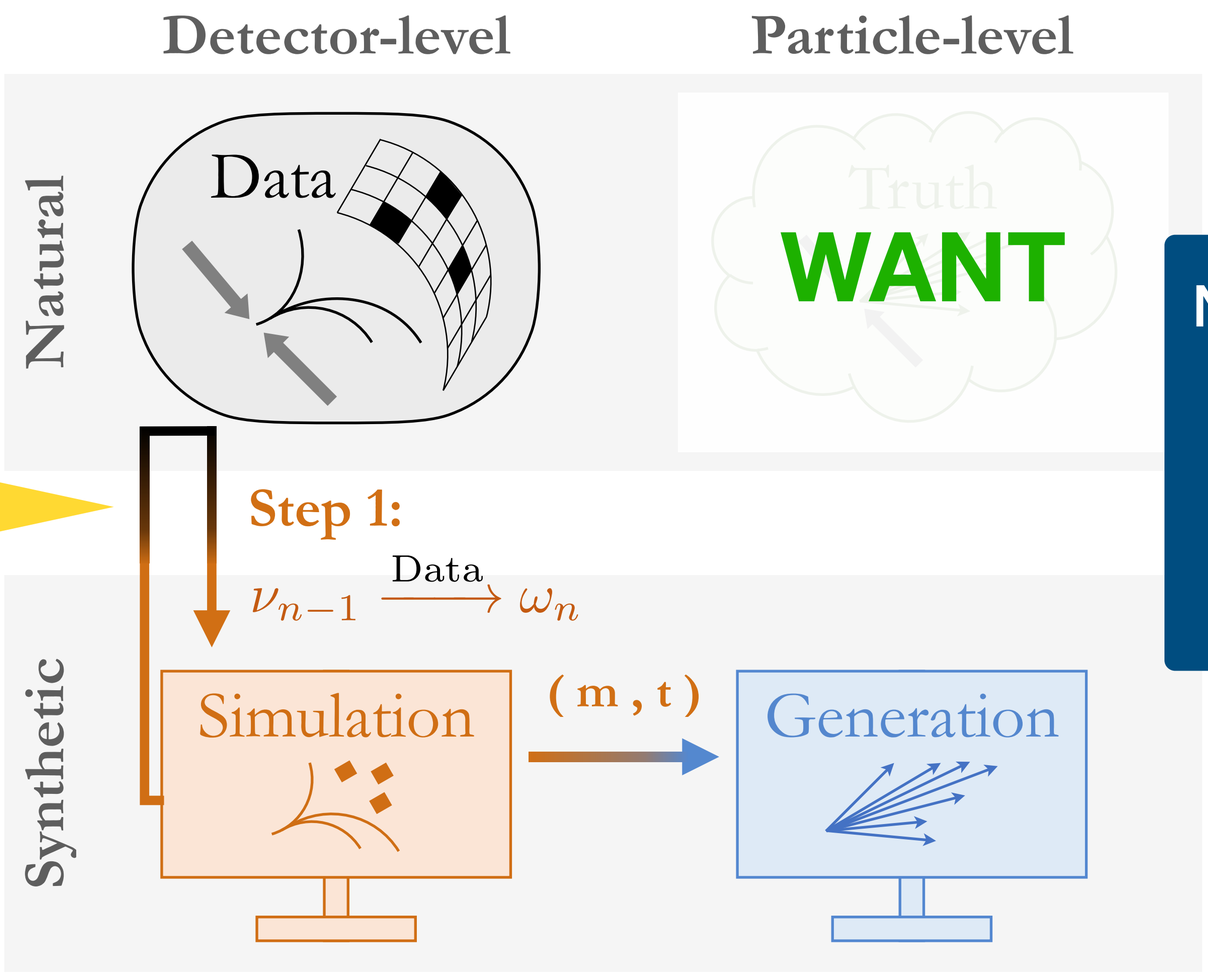
OmniFold

Likelihood-Free Inference

- Learn this likelihood ratio by training a classifier on reco-level MC vs. data:

$$W(\text{reco}) = \frac{p_{\text{data}}(\text{reco})}{p_{\text{MC}}(\text{reco})}$$

- Reweight reco-level MC to match data

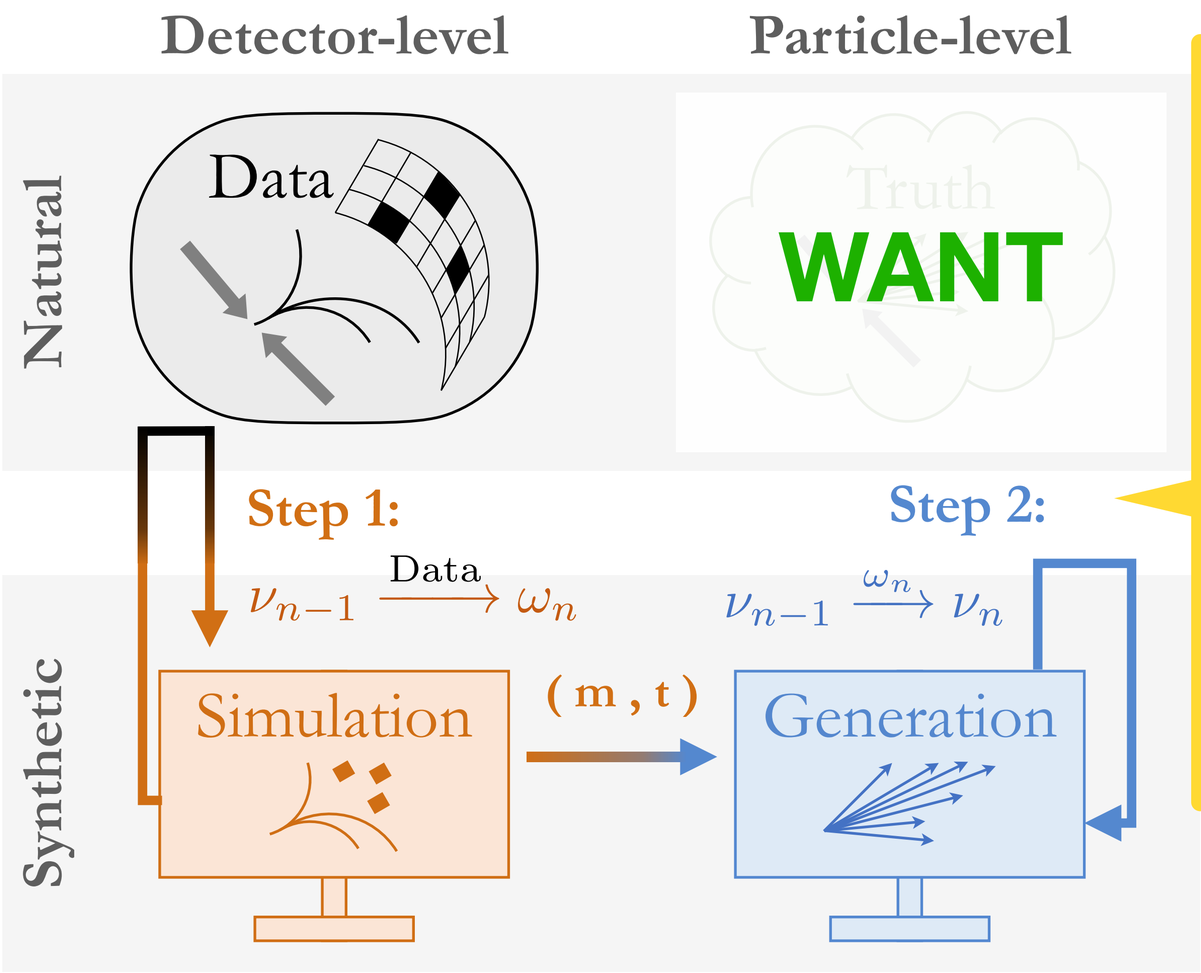


Now with data-shaped reco-level MC → can pull back to gen level and obtain the truth distribution!

[A. Andreassen et al., PRL 2020]

OmniFold

Likelihood-Free Inference



Not yet…
Need 1-to-1 mapping of gen-level kinematics

Learn another likelihood ratio

$$W(\text{gen}) = \frac{p_{\text{rw MC}}(\text{gen})}{p_{\text{MC}}(\text{gen})}$$

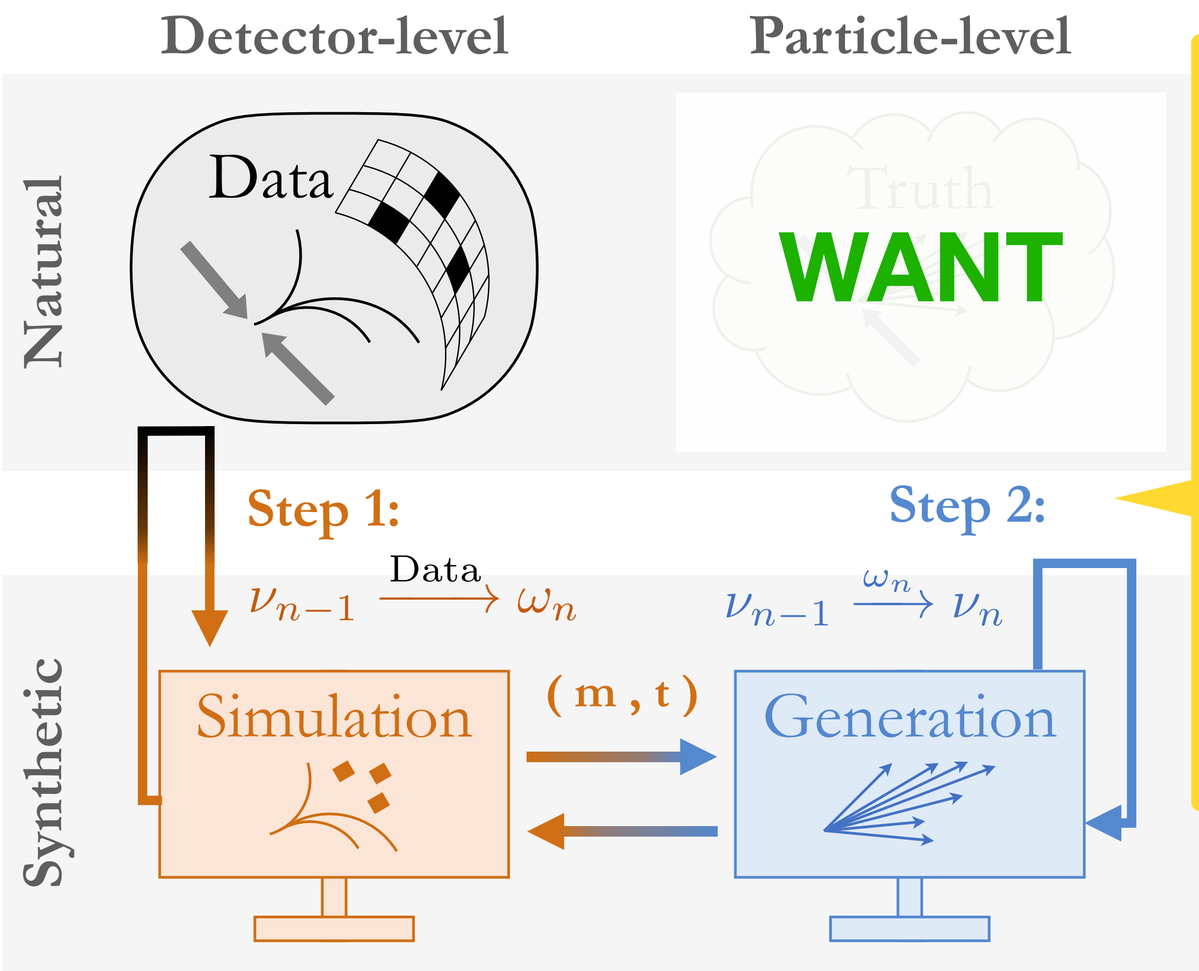
→can then apply to gen-level sample

[A. Andreassen et al., PRL 2020]

OmniFold

Likelihood-Free Inference

Iterate the prior dependence away



Not yet…
Need 1-to-1 mapping of gen-level kinematics

Learn another likelihood ratio

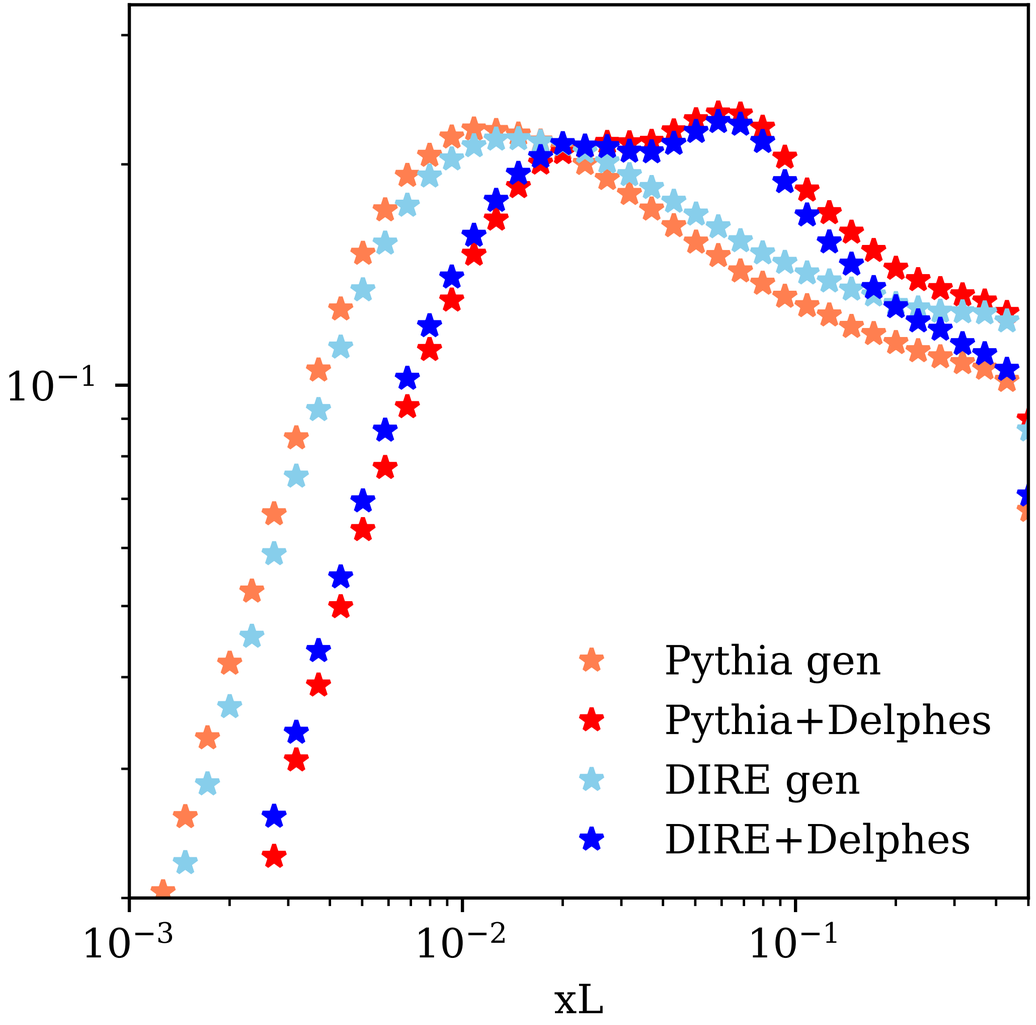
$$W(\text{gen}) = \frac{p_{\text{rw MC}}(\text{gen})}{p_{\text{MC}}(\text{gen})}$$

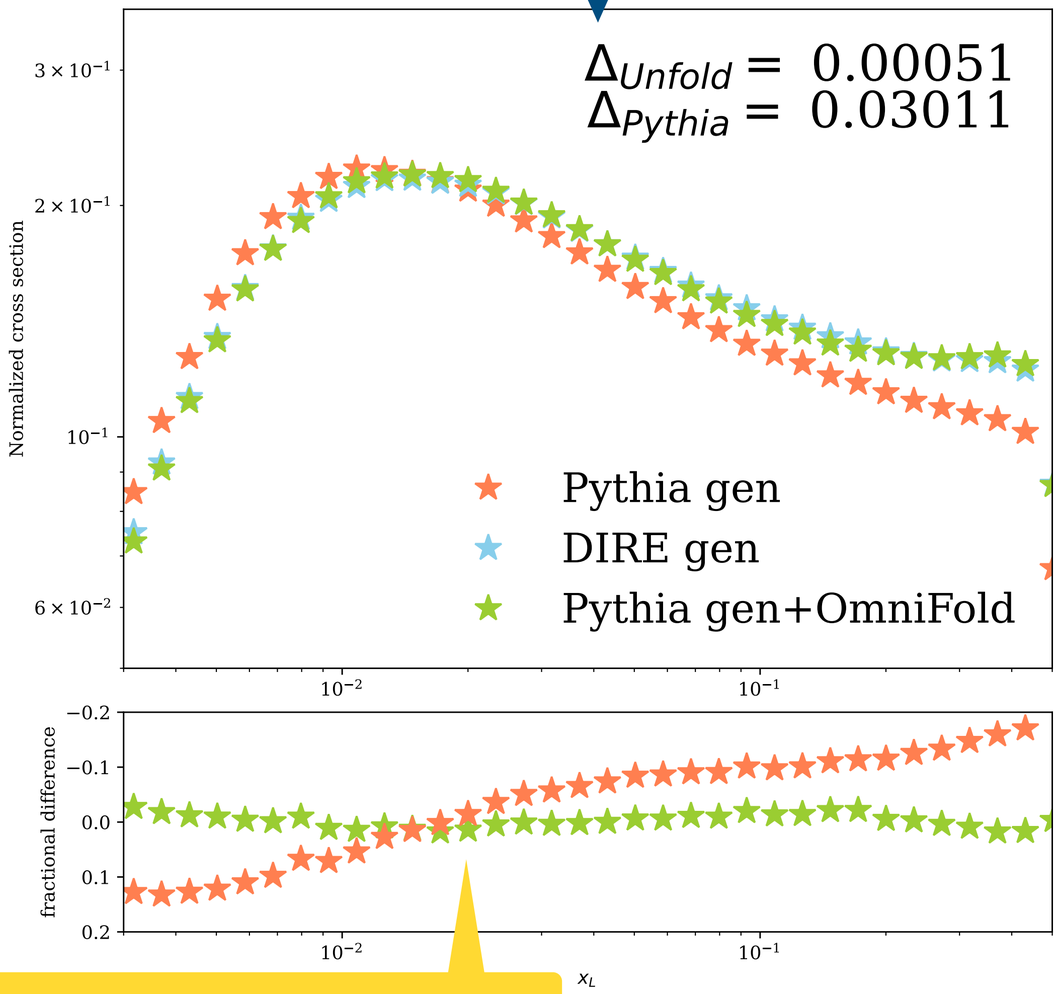
→can then apply to gen-level sample

[Toolkit for Multivariate Data Analysis, A. Hoecker et al.]

Testing Improvements of Jet Modeling

$$\Delta(p,q) = \int \mathrm{d}\lambda \frac{(p(\lambda) - q(\lambda))^2}{p(\lambda) + q(\lambda)}$$



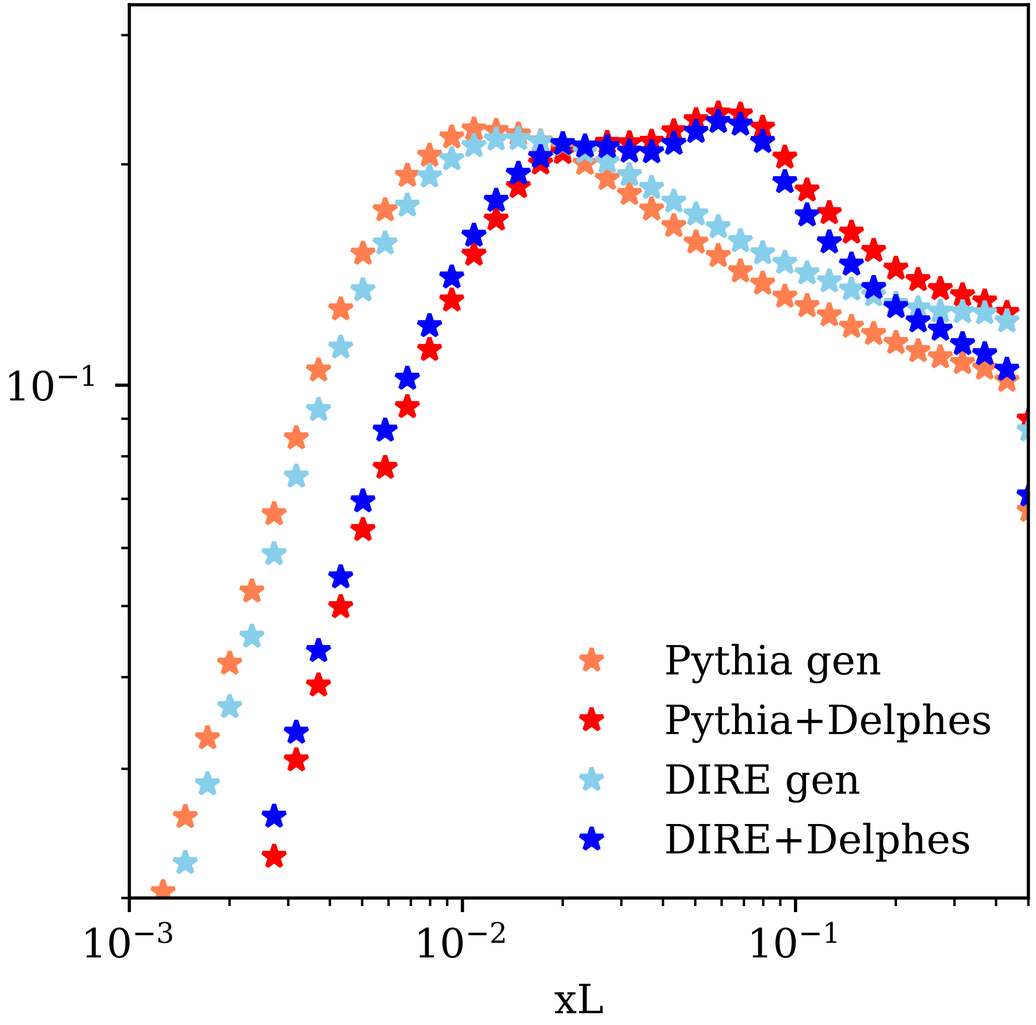


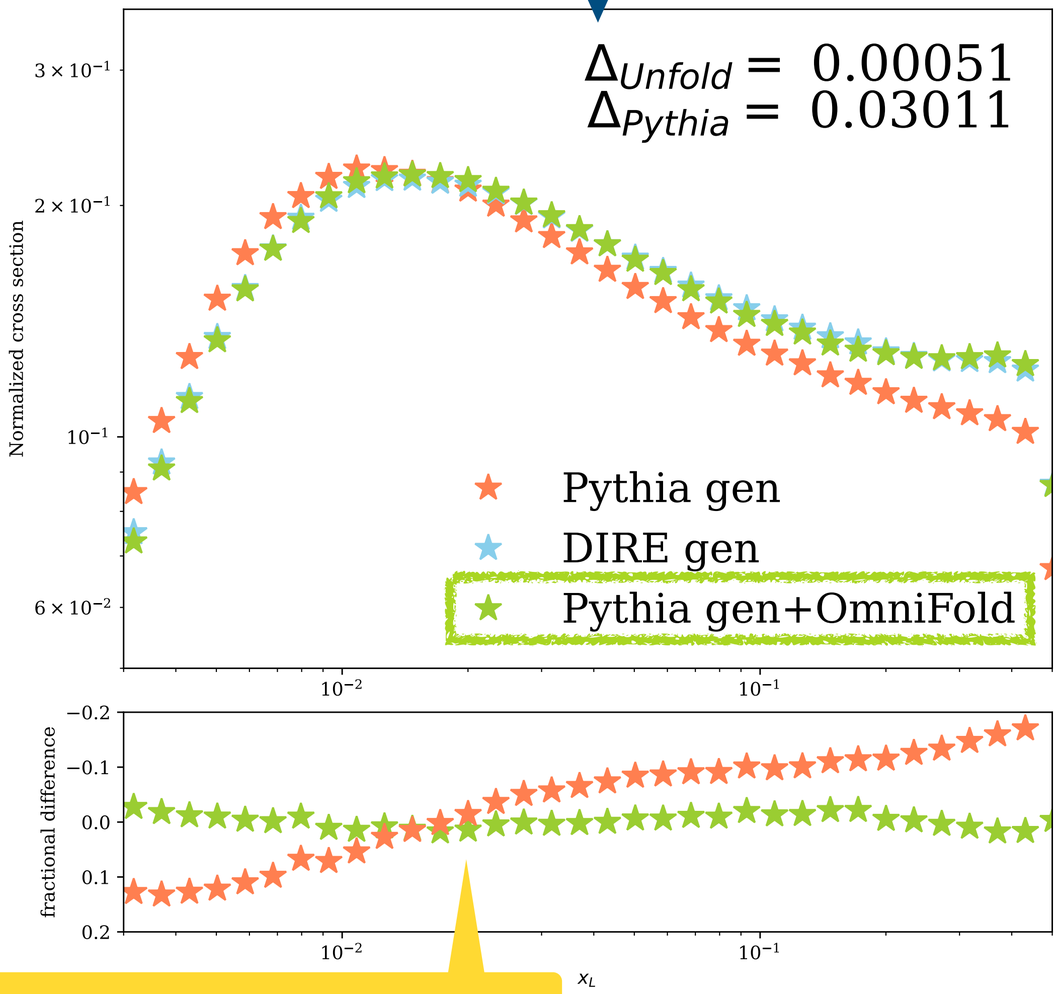
Frac. diff. from DIRE $< 2.2\,\%$

[Toolkit for Multivariate Data Analysis, A. Hoecker et al.]

Testing Improvements of Jet Modeling

$$\Delta(p,q) = \int \mathrm{d}\lambda \frac{(p(\lambda)-q(\lambda))^2}{p(\lambda)+q(\lambda)}$$





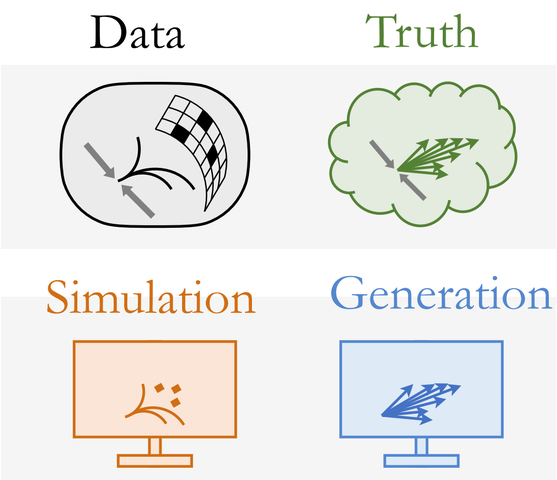
Frac. diff. from DIRE $< 2.2\,\%$

Outline

- Seeing Higher-Order QCD effects with ML Unfolding

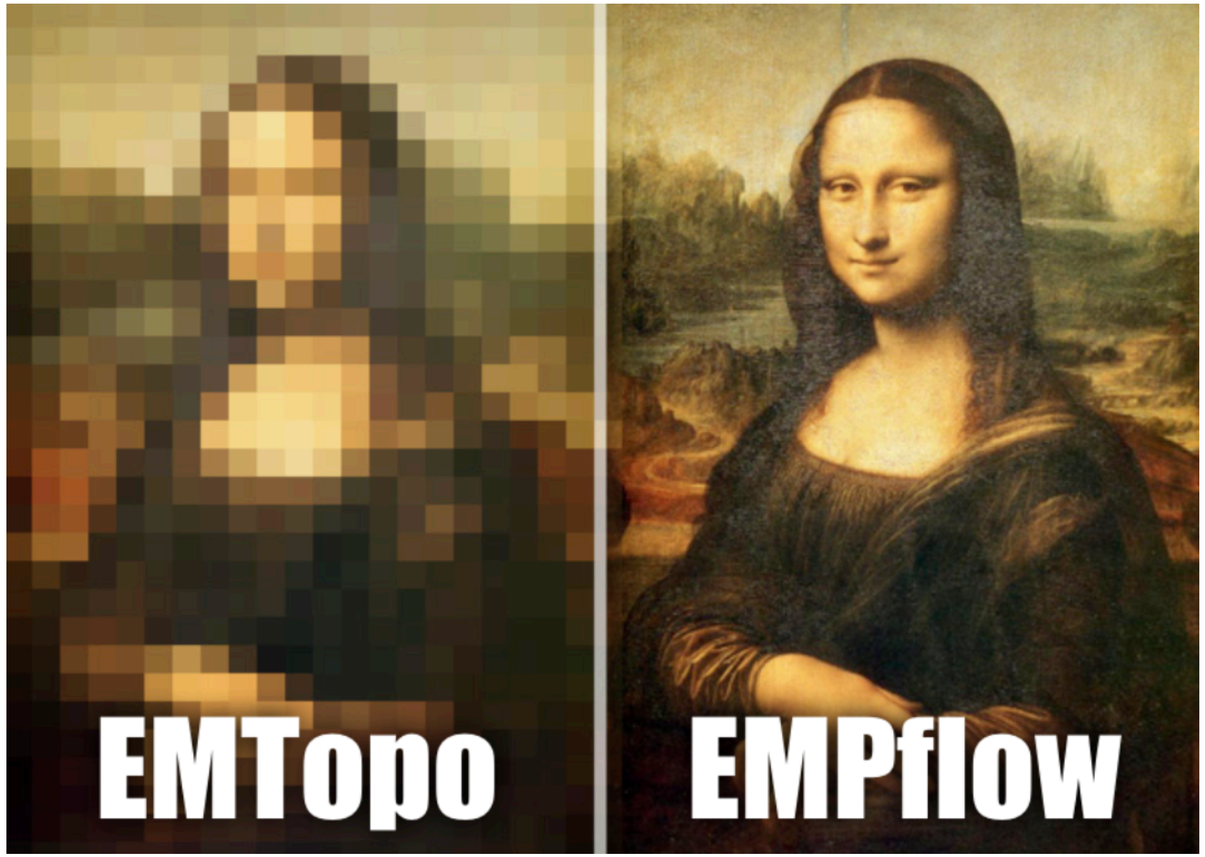
- **All About Precision:**

  Putting JSS on Track(s)



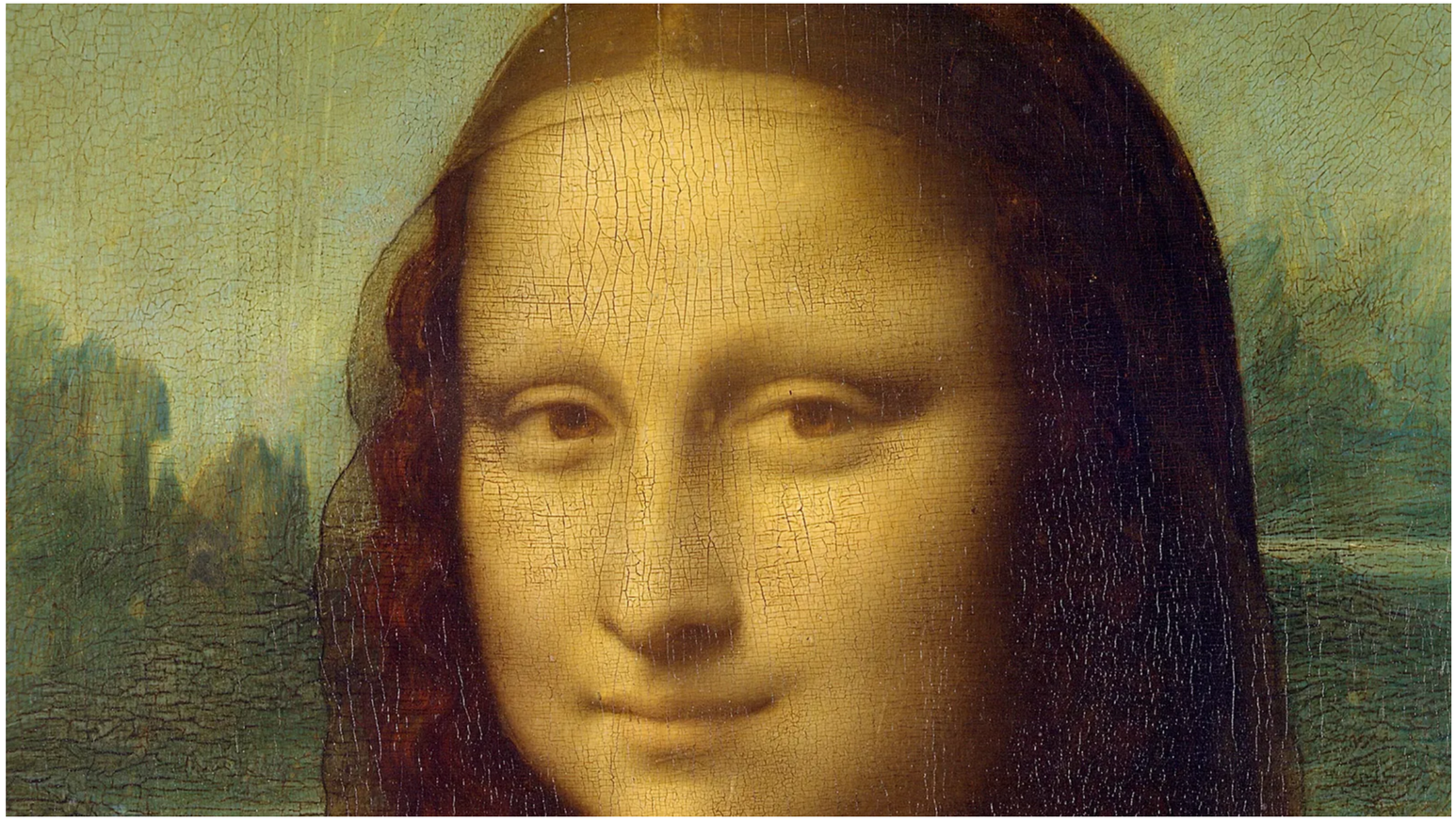
Putting JSS on Track(s)

All About Precision



Putting JSS on Track(s)

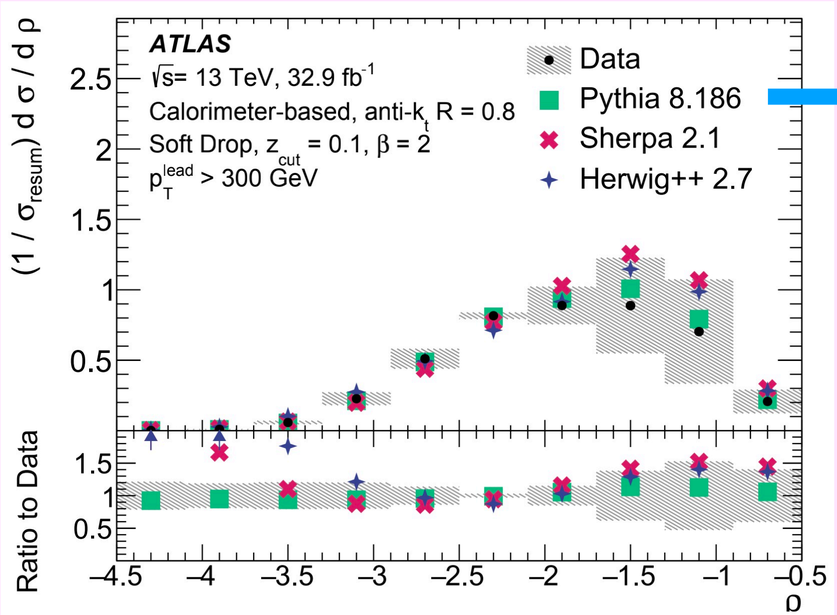
All About Precision

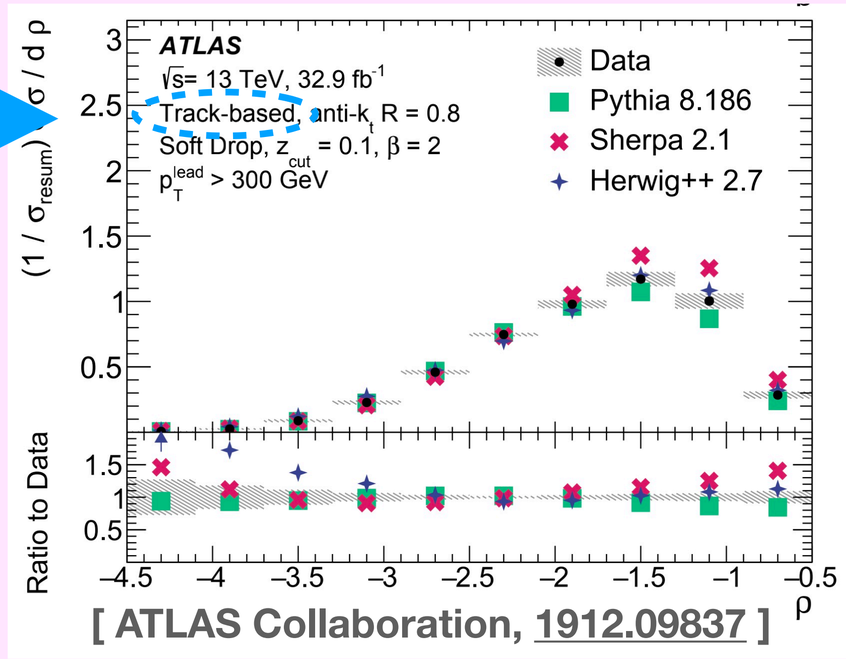


Putting JSS on Track(s)

All About Precision

Much better precision with only charged particles!





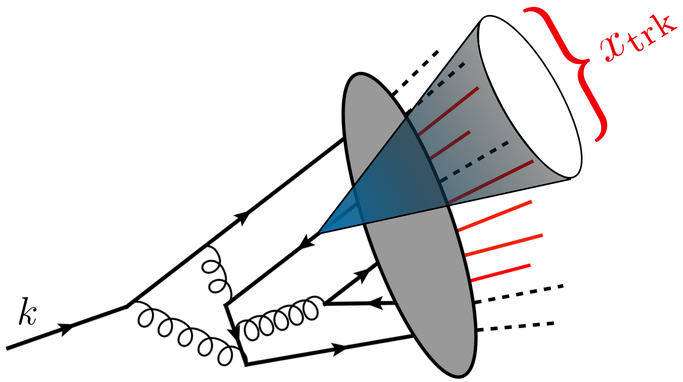
[ATLAS Collaboration, 1912.09837]

Putting JSS on Track(s)

Through the Magic of Track Functions

Momentum fraction of a hard parton converted into **charged hadrons**

- Describes also **correlations** amongst hadrons



Measure $x_{\rm trk}$ → Track functions

$$T_g(z) \propto \delta(z_1 + z_2 - z)$$

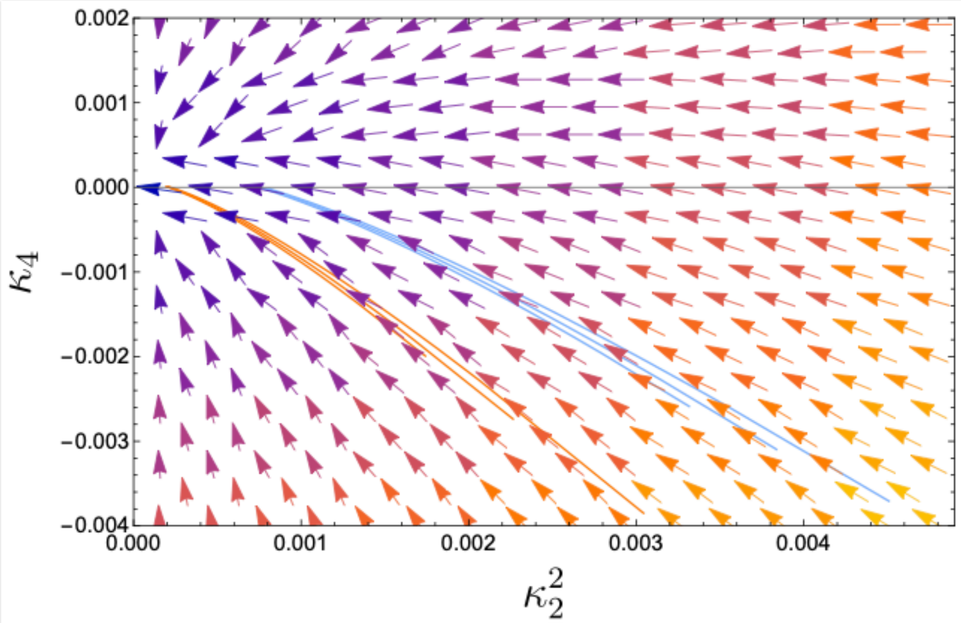
$$D_{g\to\pi^+}(z) \propto \delta(z_1 - z) + \delta(z_2 - z)$$

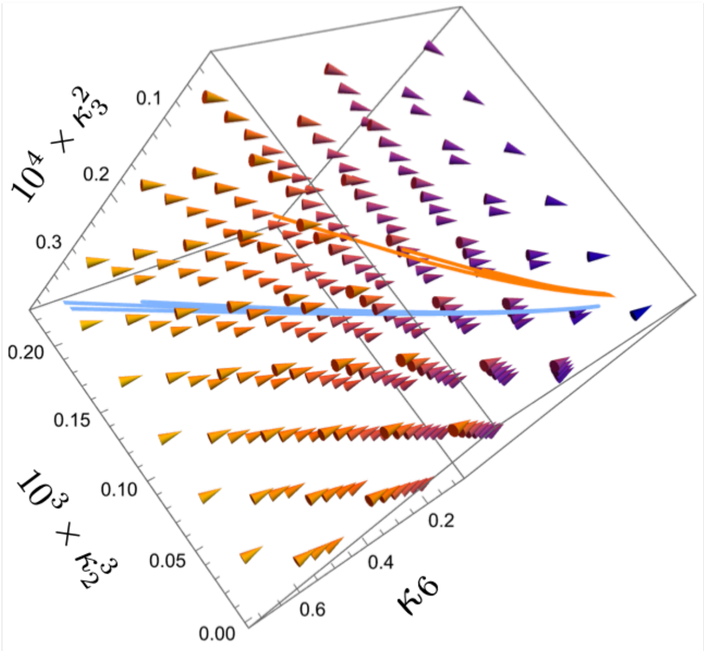
Fragmentation function only describes **single identified hadron**

Track Functions as a Probe of QCD

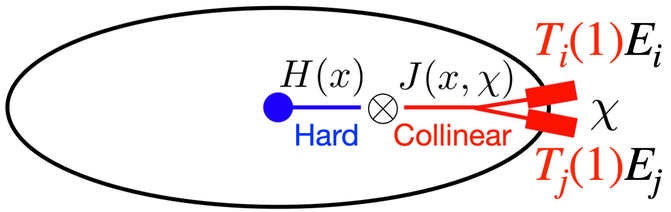
Track functions will play two key roles in Jet Substructure:

- Enable **precision calculations of JSS observables on charged hadrons.**
- **Provide interesting probes of QCD RG flows**

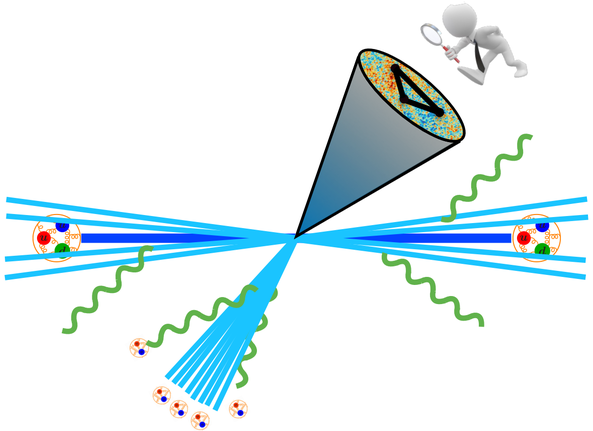




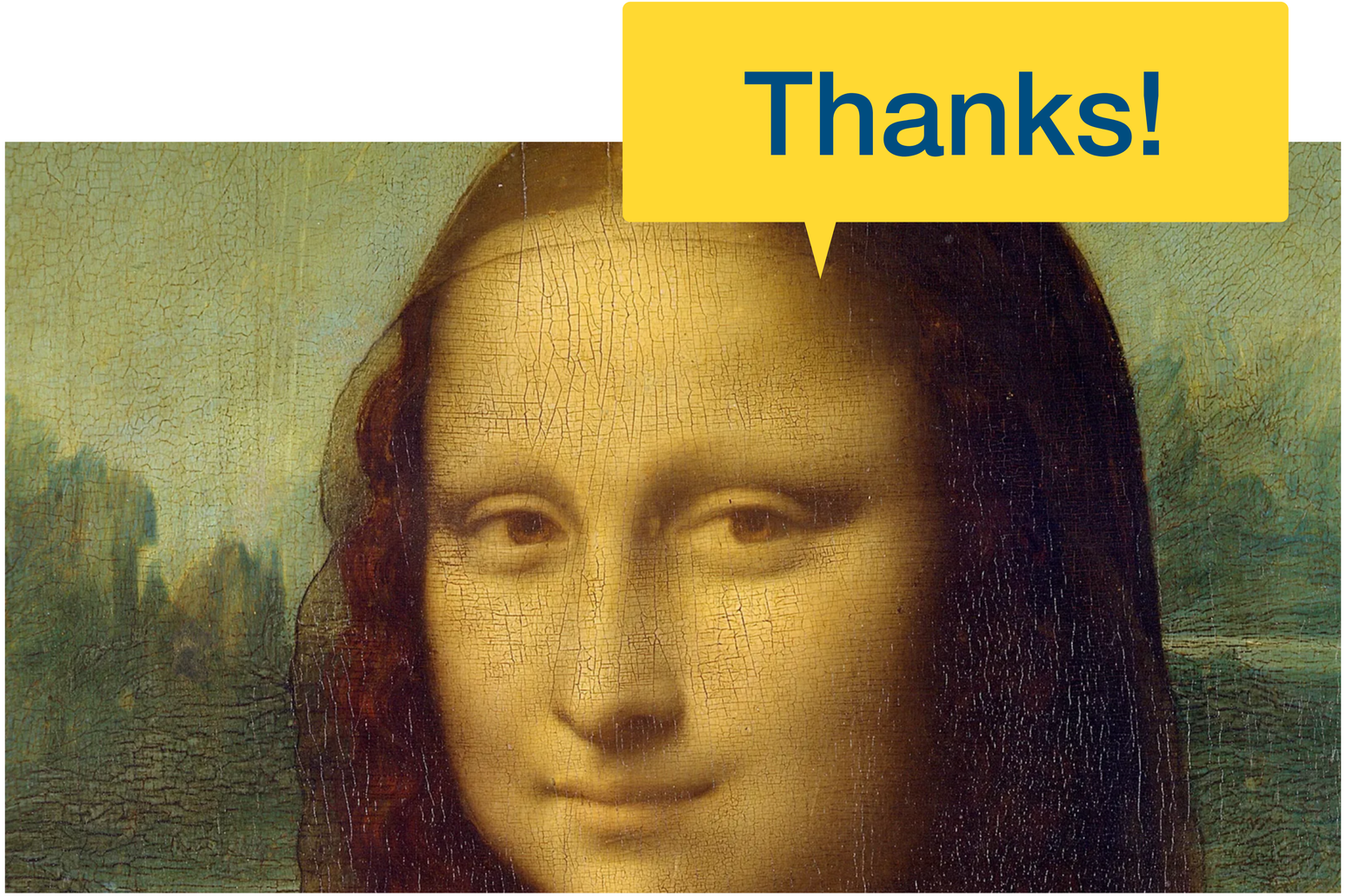
Summary



Progress towards experimental studies of a new class of JSS observables **Energy Correlators**



- **Track functions are an important aspect of a precision program.**
- **We have shown that Omnifold can be sensitive to higher order QCD effects. Bright future for precision measurements!**


Thanks!

Backup

OmniFold

“Likelihood-ratio trick”

- likelihood function hard to approach in high dimensions

- the ratio of two likelihood functions can be approximated by the decision function of a binary classifier; make good use of MC

- can process variable length, unordered input with proper neural network architecture

$$\underset{f}{\operatorname{argmin}}\, L[f] = \frac{p(x \mid \theta_0)}{p(x \mid \theta_1)} = \mathcal{L}(x)$$

$$\begin{aligned}
\frac{f(x)}{1-f(x)} &= \frac{\frac{p(x|\theta_0)}{p(x|\theta_0)+p(x|\theta_1)}}{1-\frac{p(x|\theta_0)}{p(x|\theta_0)+p(x|\theta_1)}} \\
&= \frac{p(x \mid \theta_0)}{\cancel{p(x \mid \theta_0)} + p(x \mid \theta_1) - \cancel{p(x \mid \theta_0)}} \\
&= \frac{p(x \mid \theta_0)}{p(x \mid \theta_1)} = \mathcal{L}(x)
\end{aligned}$$

Neural Networks for Full Phase-space Reweighting and Parameter Tuning

Learning Likelihood Ratios with Neural Network Classifiers

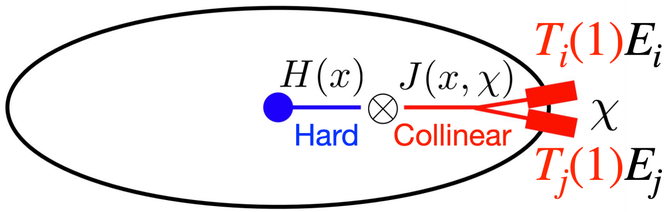
OmniFold

“Likelihood-ratio trick”

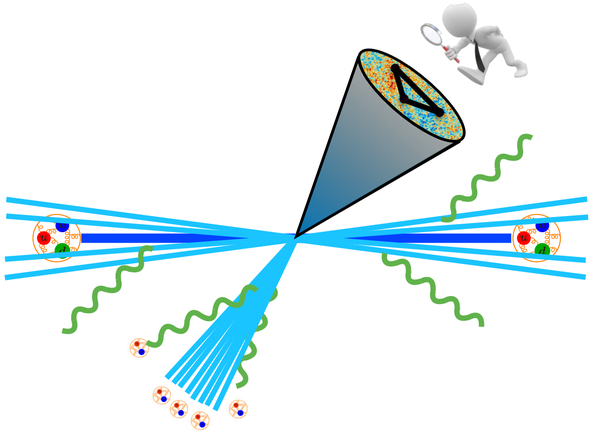
Admits multiple advantages:

- Naturally unbinned & high dimensional
- Variable length, unordered sets → full phase space unfolding
- Converges to maximum likelihood estimate of the truth distribution
- Computationally efficient (by reweighting i.e. learning small correction)
- Beyond per-event observables, readily reinterpretable

Summary



Progress towards experimental studies of a new class of JSS observables **Energy Correlators**



- **First measurement of track functions to come shortly from ATLAS! Stay Tuned!**

- **We have shown that Omnifold can be sensitive to higher order QCD effects. Bright future for precision measurements!**