

OmniFold

Improved Unfolding with Deep Learning

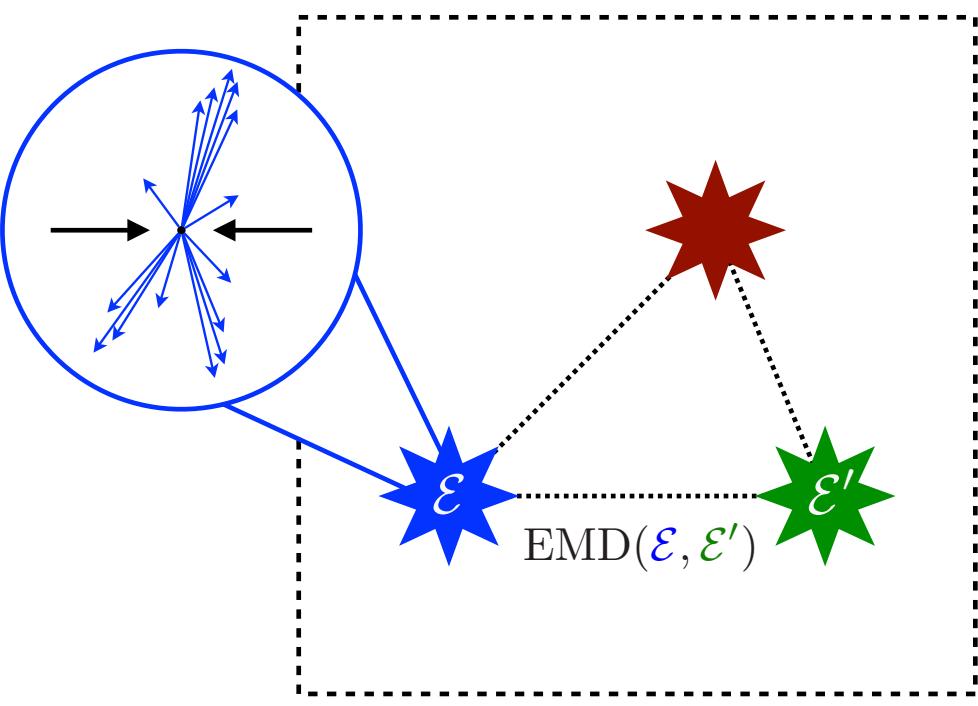
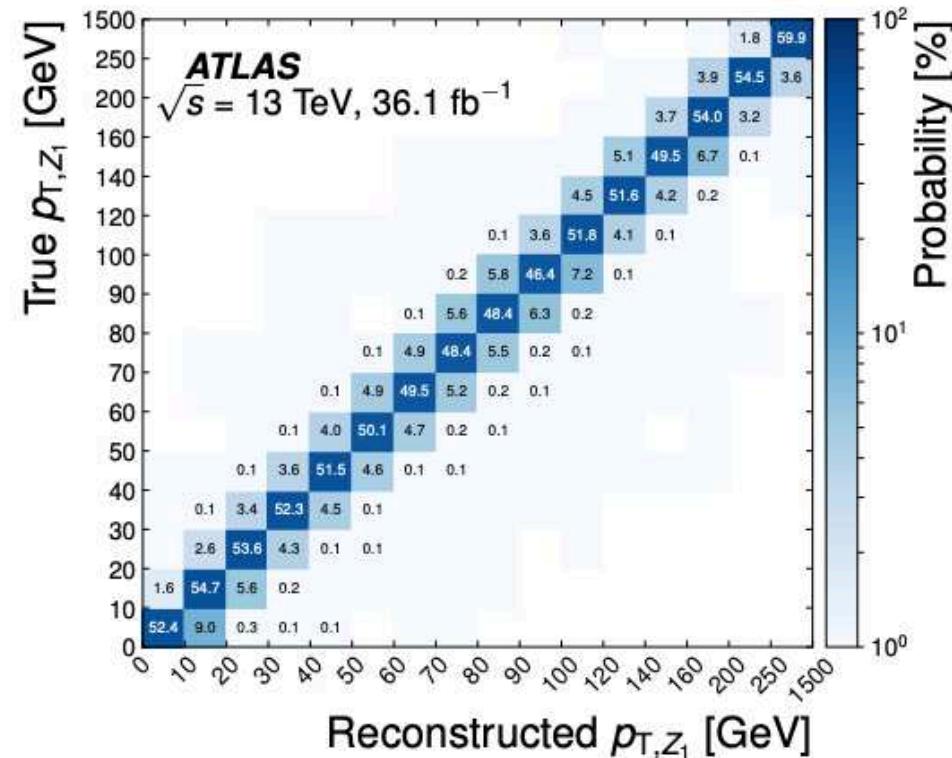
Patrick T. Komiske III

Massachusetts Institute of Technology
Center for Theoretical Physics

Based on work with Anders Andreassen, Eric Metodiev, Ben Nachman, and Jesse Thaler
[1911.09107 \(PRL\)](#)

LHC-Electroweak/Jets Working Group

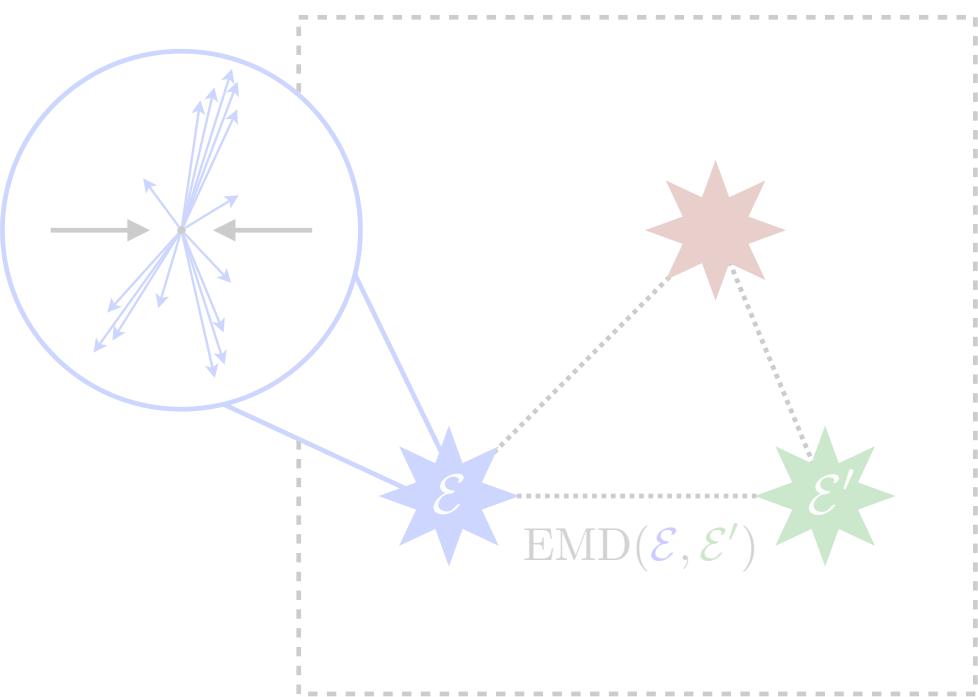
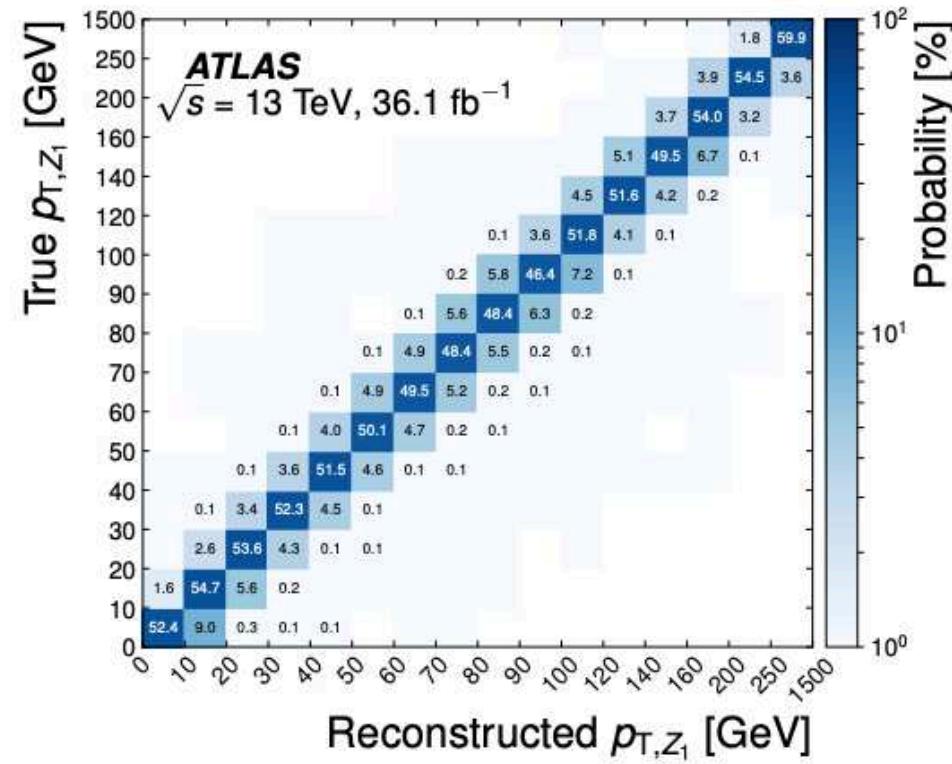
February 8, 2021



Unfolding Setup

OmniFold

Unfolding Beyond Observables



Unfolding Setup

OmniFold

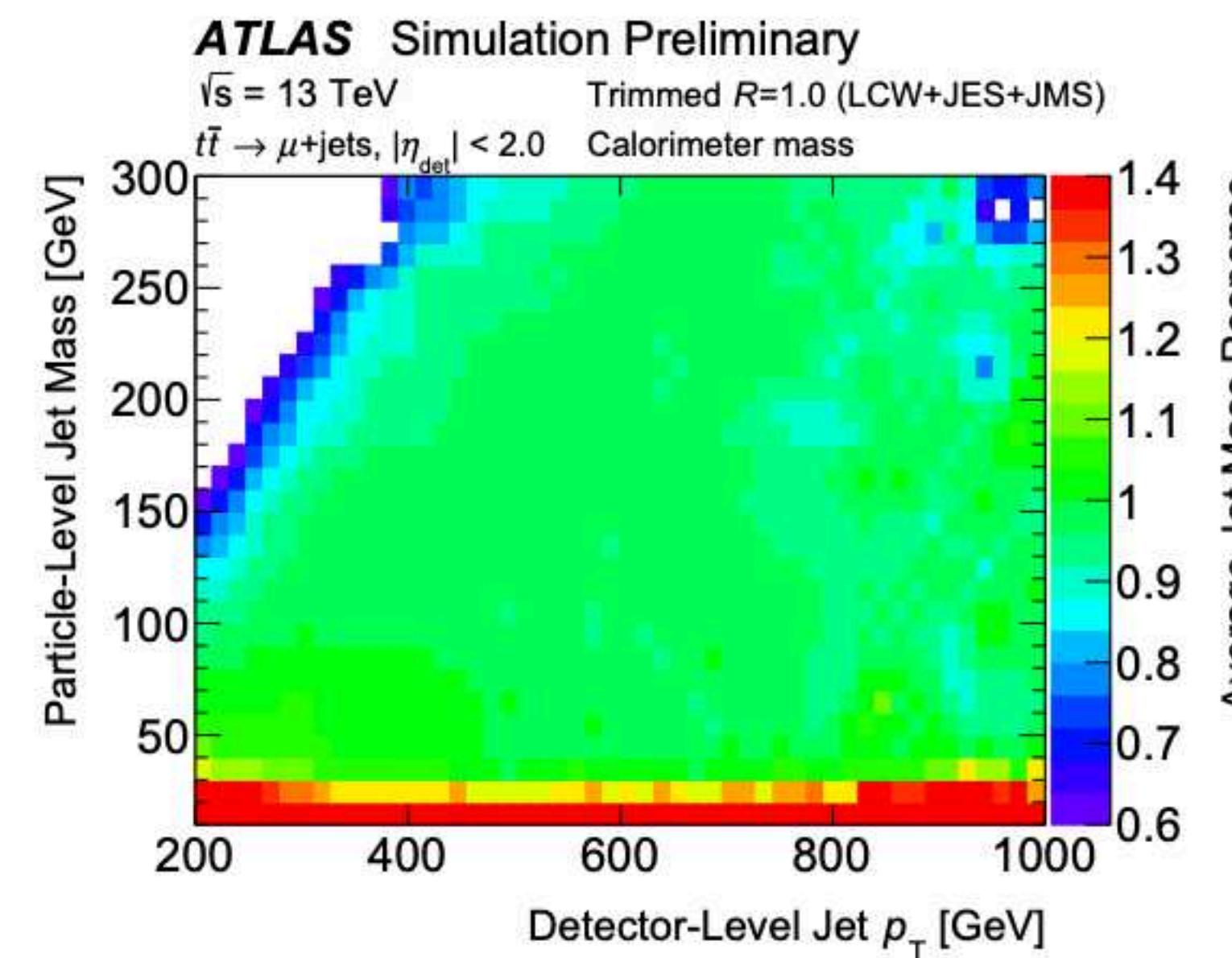
Unfolding Beyond Observables

Correcting for Detector Effects

- Detectors introduce (potentially correlated) smearing and biasing that must be corrected in any measurement
- Material interactions and detector geometry modeled with sophisticated (i.e. expensive) simulation software (e.g. GEANT4)

Forward folding simulates given truth-level events
and calculates detector-level quantities

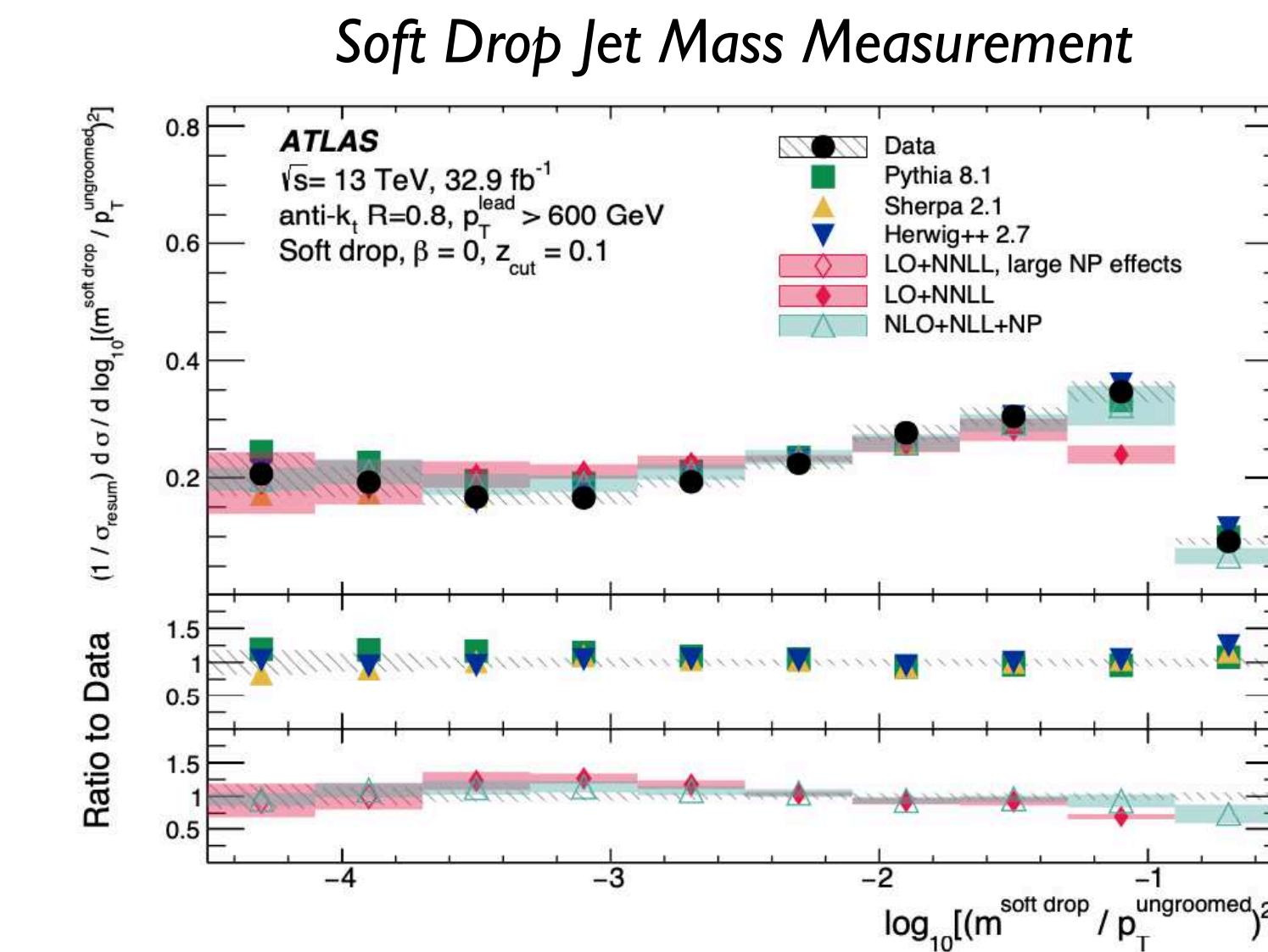
[ATLAS-CONF-2020-022]



Detector response varies according to jet mass and p_T
Explicitly depends on specific detector geometry

Unfolding estimates truth-level quantities given experimental
data and information about detector response

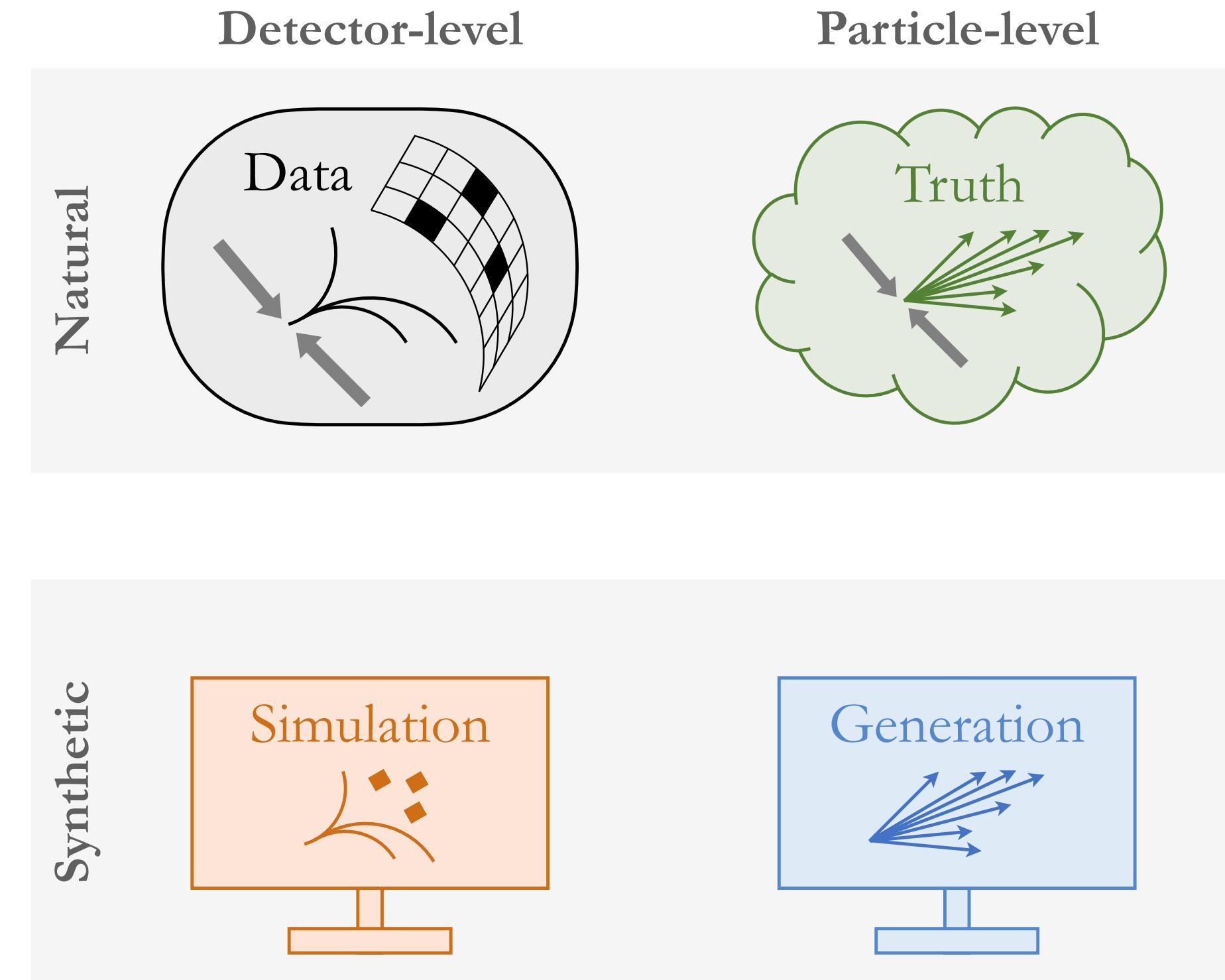
[ATLAS, PRL 2018]



Comparison to precision theory possible in detector-independent manner
Measurement can be used by anyone, no need for detailed experimental information

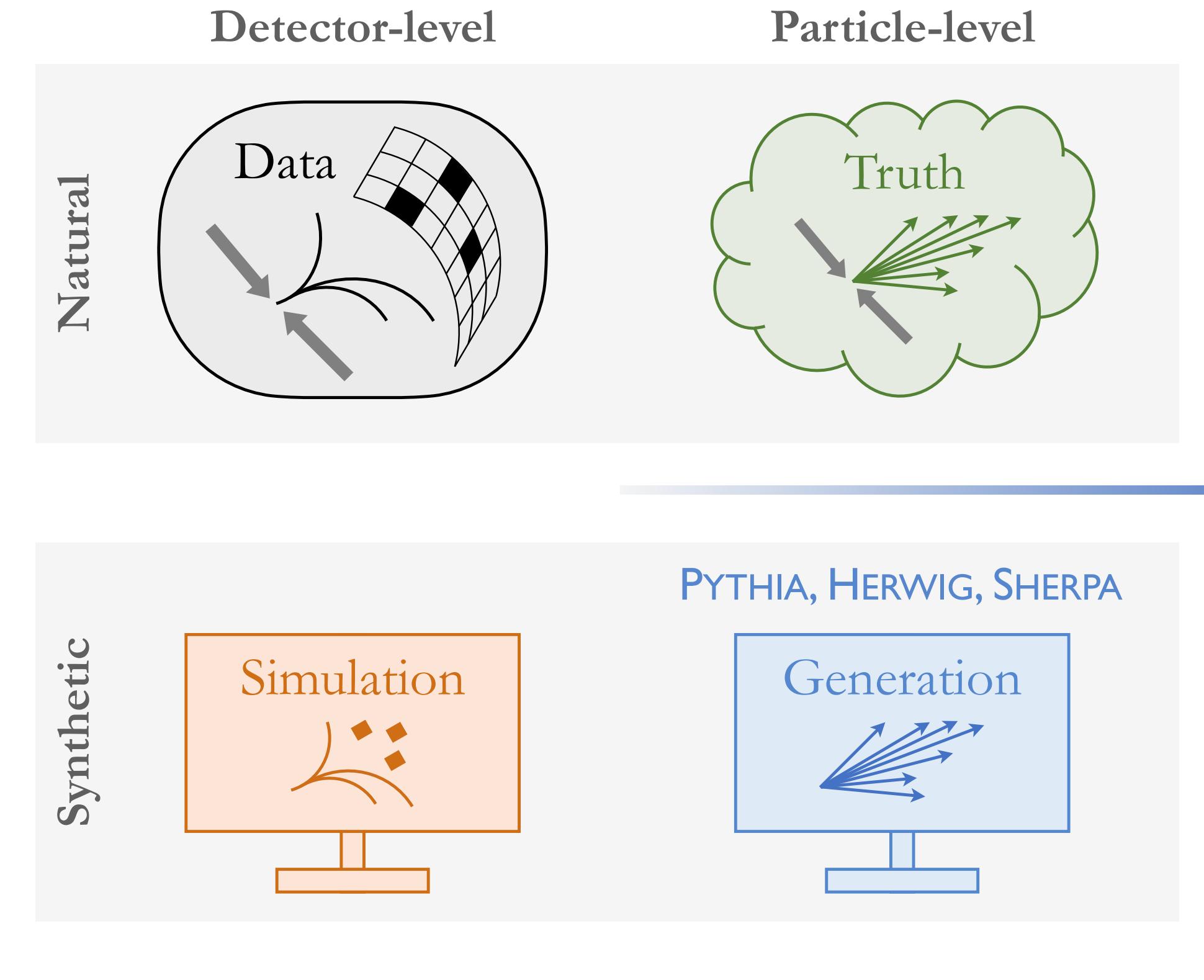
Unfolding Setup

Measurements are affected by *detector effects* of finite resolution and limited acceptance



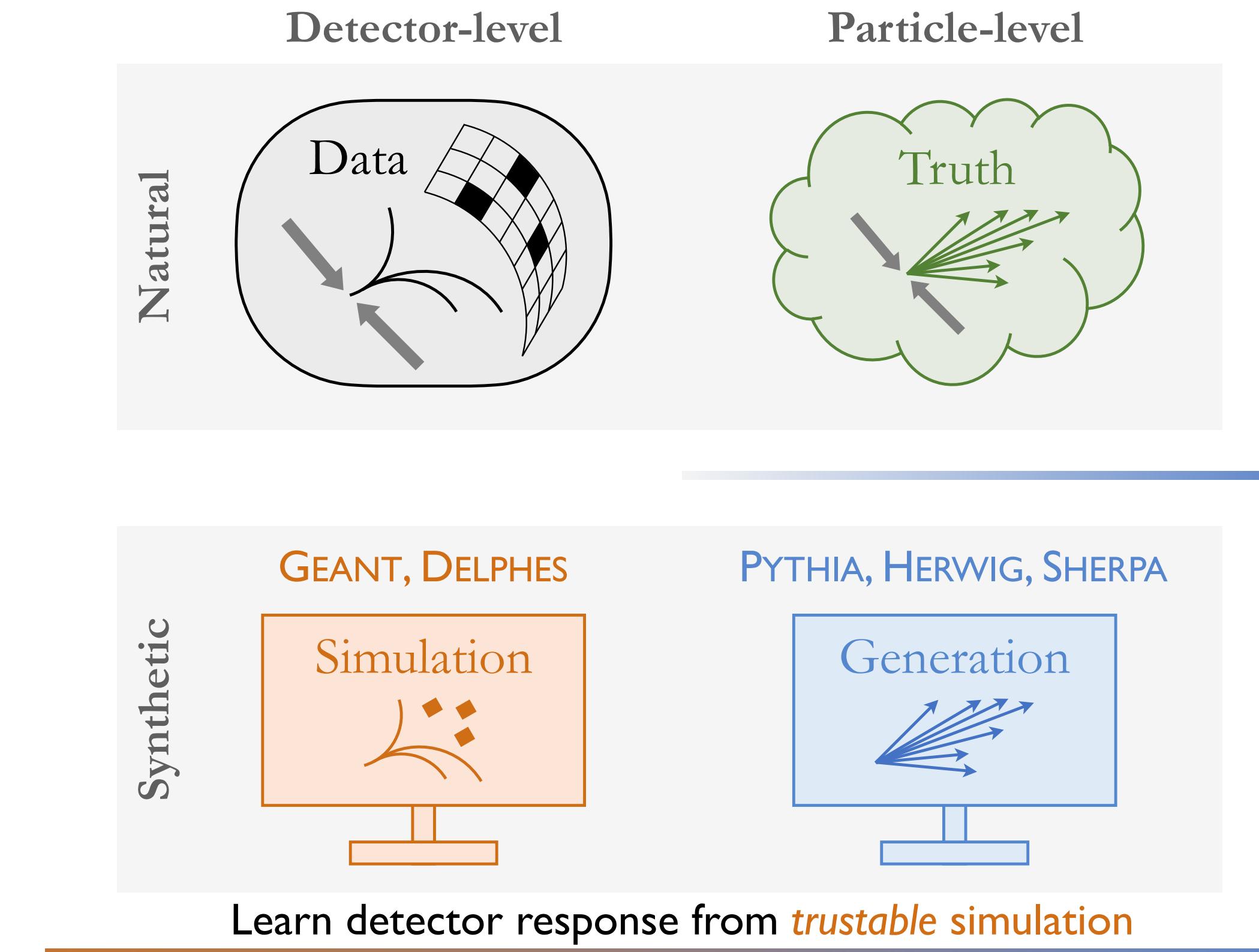
Unfolding Setup

Measurements are affected by *detector effects* of finite resolution and limited acceptance



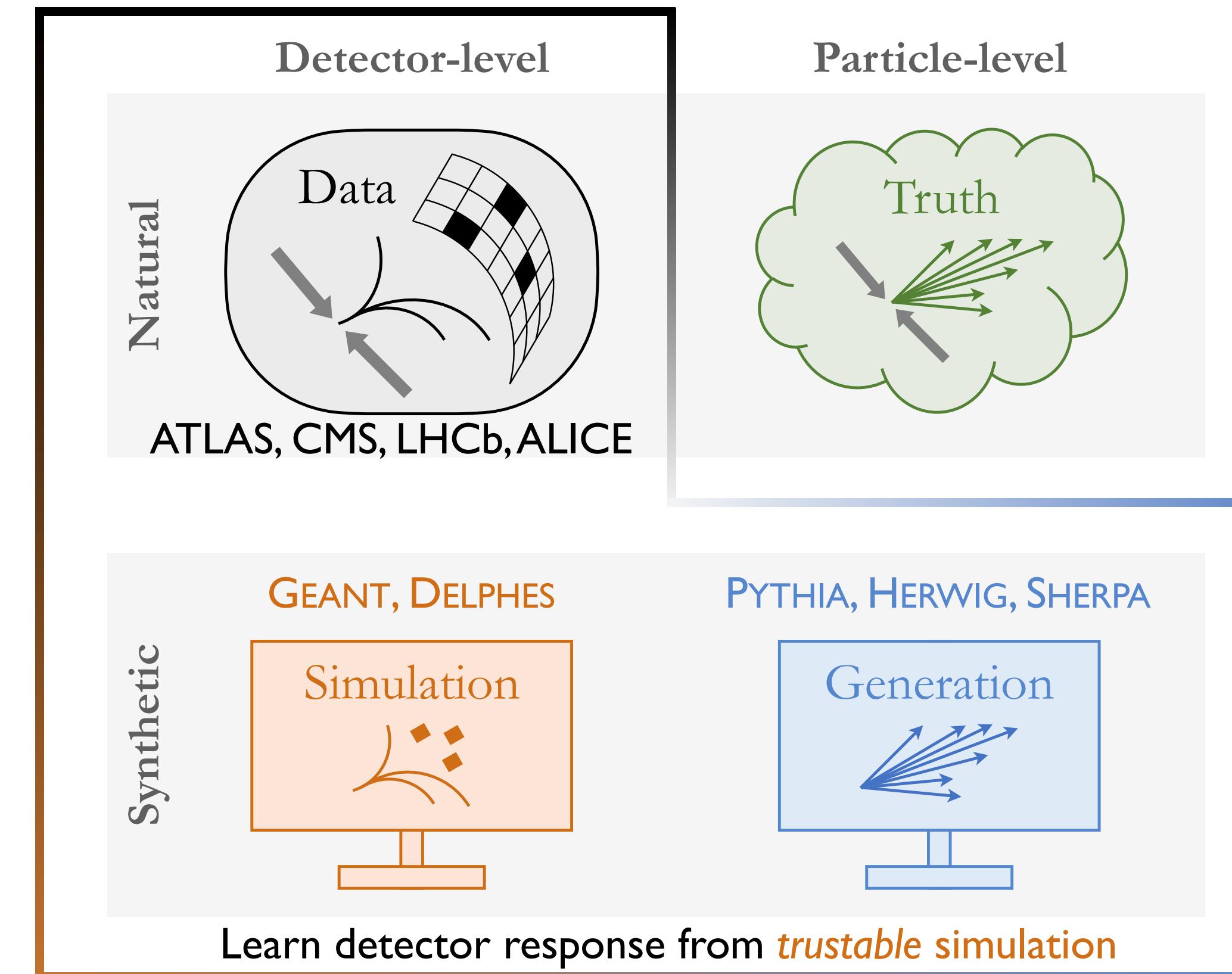
Unfolding Setup

Measurements are affected by *detector effects* of finite resolution and limited acceptance



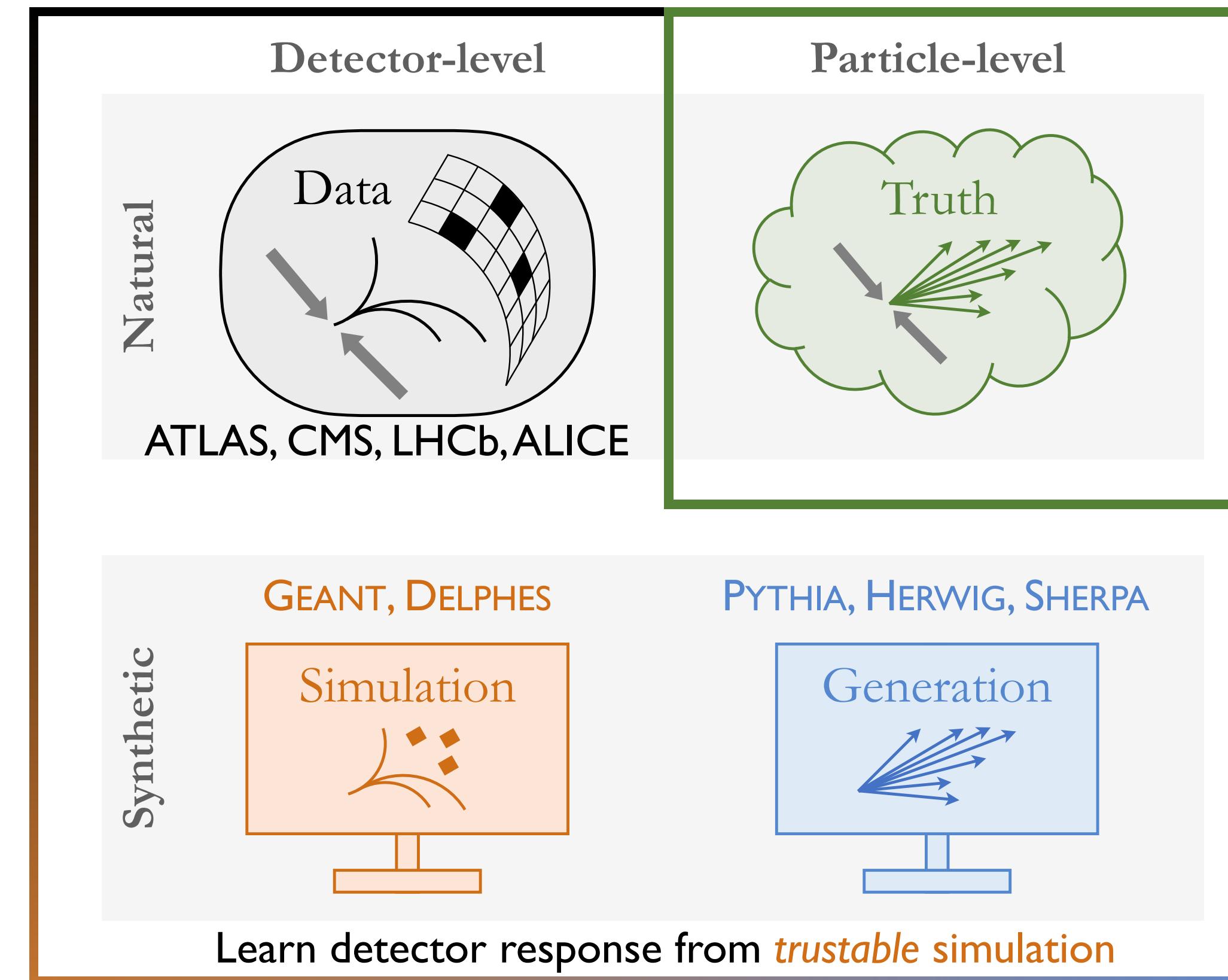
Unfolding Setup

Measurements are affected by *detector effects* of finite resolution and limited acceptance



Unfolding Setup

Measurements are affected by *detector effects* of finite resolution and limited acceptance



Truth-level measurements can be compared across experiments and to *theoretical calculations*

Goal of *unfolding* is to learn a generative *particle-level* model that reproduces the data

Challenges with Traditional Unfolding

Previous methods are inherently binned

- Binning fixed ahead of time, cannot be changed later
- Performance of method sensitive to binning

Limited number of observables

- Binning induces curse of dimensionality

Response matrix depends on auxiliary features

- Detector-level quantity may not capture full detector effect

Challenges with Traditional Unfolding

Previous methods are inherently binned

- Binning fixed ahead of time, cannot be changed later
- Performance of method sensitive to binning

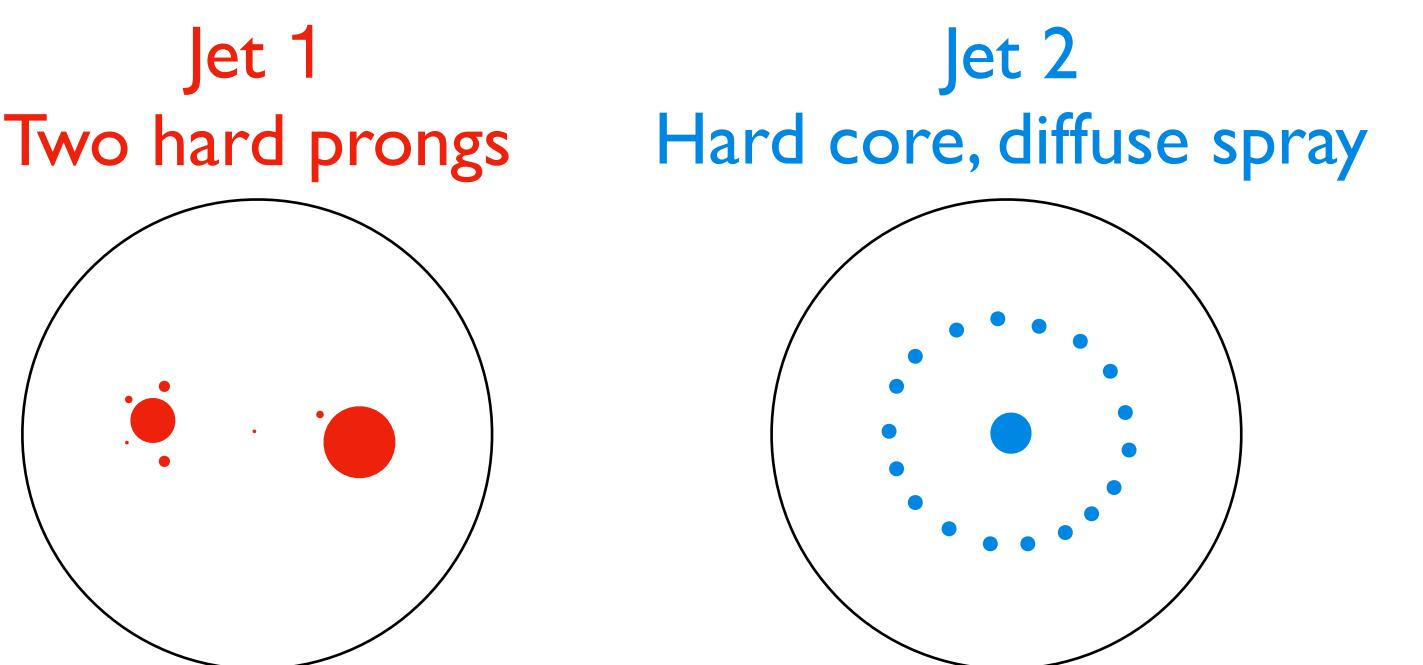
Limited number of observables

- Binning induces curse of dimensionality

Response matrix depends on auxiliary features

- Detector-level quantity may not capture full detector effect

Example – Two jets acquiring the same mass in different ways



Challenges with Traditional Unfolding

Previous methods are inherently binned

Binning fixed ahead of time, cannot be changed later

Performance of method sensitive to binning

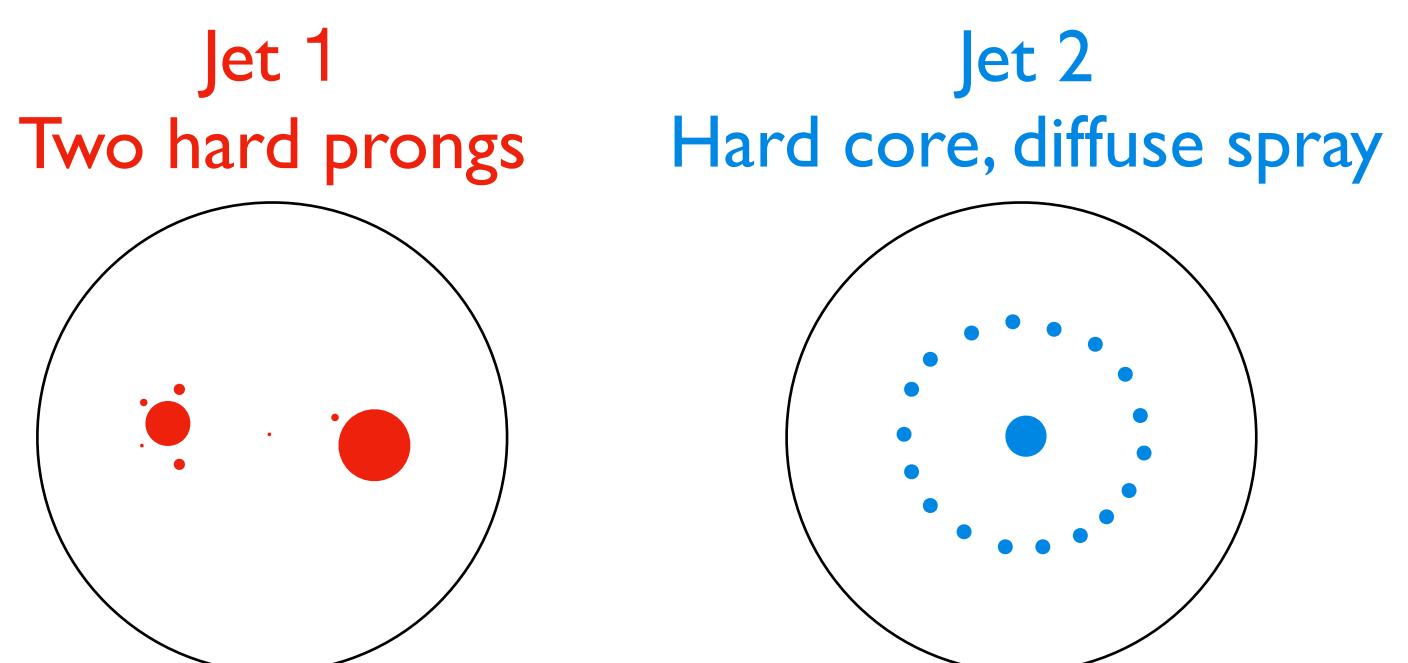
Limited number of observables

Binning induces curse of dimensionality

Response matrix depends on auxiliary features

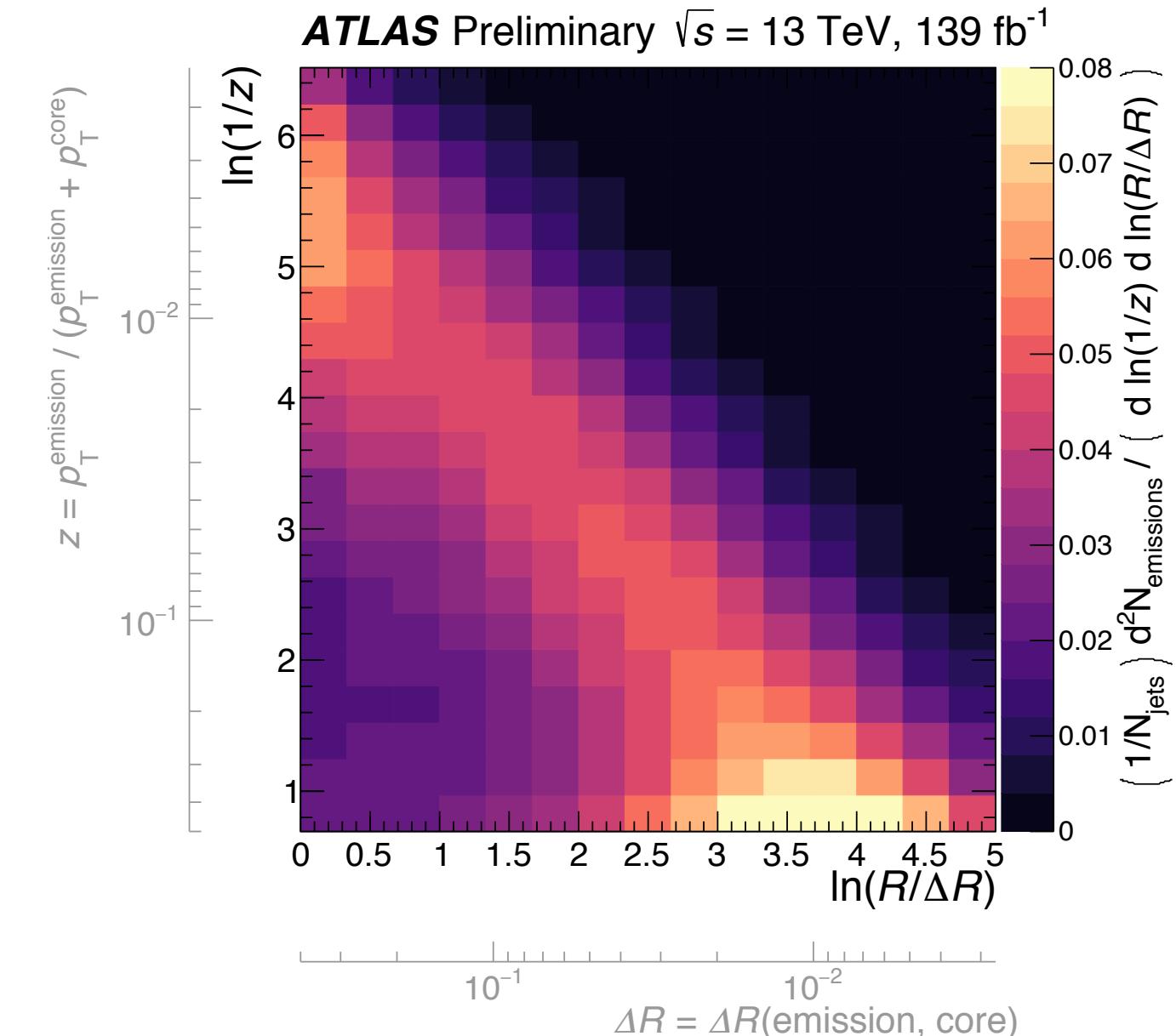
Detector-level quantity may not capture full detector effect

Example – Two jets acquiring the same mass in different ways



Example with IBU

ATLAS State-of-the-art Lund Plane Measurement
[PRL 2020]



21 x 15 bins in $\ln(1/z) \times \ln(R/\Delta R)$

- Must redo unfolding for other binnings e.g. finer/coarser, k_T (diagonal) binning, etc.

Limited to two observables

- $21^2 \times 15^2$ elements in response matrix R
- Going differential in n bins of p_T would multiply size of R by n^2

Traditional Unfolding

Iterated Bayesian Unfolding (IBU)

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]

Maximum likelihood, histogram-based unfolding method for a small number of observables

Choose observable(s) and binning at **detector-level** and **particle-level**

measured distribution: $m_i = \Pr(\text{measure } i)$

true distribution: $t_j^{(0)} = \Pr(\text{truth is } j)$

Calculate **response matrix** R_{ij} from **generated/simulated** pairs of events

$$R_{ij} = \Pr(\text{measure } i \mid \text{truth is } j)$$

Calculate new particle-level distribution using Bayes' theorem

$$\begin{aligned} t_j^{(n)} &= \sum_i \Pr(\text{truth}_{n-1} \text{ is } j \mid \text{measure } i) \times \Pr(\text{measure } i) \\ &= \sum_i \frac{R_{ij} t_j^{(n-1)}}{\sum_k R_{ik} t_k^{(n-1)}} \times m_i \end{aligned}$$

Iterate procedure to remove dependence on prior

Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

Iterated Bayesian Unfolding (IBU)

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]

Maximum likelihood, histogram-based unfolding method for a small number of observables

Choose observable(s) and binning at **detector-level** and **particle-level**

measured distribution: $m_i = \Pr(\text{measure } i)$

true distribution: $t_j^{(0)} = \Pr(\text{truth is } j)$

Calculate **response matrix** R_{ij} from **generated/simulated pairs of events**

$R_{ij} = \Pr(\text{measure } i \mid \text{truth is } j)$

Calculate new particle-level distribution using Bayes' theorem

$$\begin{aligned} t_j^{(n)} &= \sum_i \Pr(\text{truth}_{n-1} \text{ is } j \mid \text{measure } i) \times \Pr(\text{measure } i) \\ &= \sum_i \frac{R_{ij} t_j^{(n-1)}}{\sum_k R_{ik} t_k^{(n-1)}} \times m_i \end{aligned}$$

Iterate procedure to remove dependence on prior

Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\left(\frac{3}{4} \quad \frac{1}{4}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

After one iteration

Iterated Bayesian Unfolding (IBU)

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]

Maximum likelihood, histogram-based unfolding method for a small number of observables

Choose observable(s) and binning at **detector-level** and **particle-level**

measured distribution: $m_i = \Pr(\text{measure } i)$

true distribution: $t_j^{(0)} = \Pr(\text{truth is } j)$

Calculate **response matrix** R_{ij} from **generated/simulated pairs of events**

$R_{ij} = \Pr(\text{measure } i \mid \text{truth is } j)$

Calculate new particle-level distribution using Bayes' theorem

$$\begin{aligned} t_j^{(n)} &= \sum_i \Pr(\text{truth}_{n-1} \text{ is } j \mid \text{measure } i) \times \Pr(\text{measure } i) \\ &= \sum_i \frac{R_{ij} t_j^{(n-1)}}{\sum_k R_{ik} t_k^{(n-1)}} \times m_i \end{aligned}$$

Iterate procedure to remove dependence on prior

Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly

Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\left(\frac{3}{4} \quad \frac{1}{4}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix}_{1/2} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

Iterated Bayesian Unfolding (IBU)

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]

Maximum likelihood, histogram-based unfolding method for a small number of observables

Choose observable(s) and binning at **detector-level** and **particle-level**

measured distribution: $m_i = \Pr(\text{measure } i)$

true distribution: $t_j^{(0)} = \Pr(\text{truth is } j)$

Calculate response matrix R_{ij} from **generated/simulated pairs of events**

$R_{ij} = \Pr(\text{measure } i \mid \text{truth is } j)$

Calculate new particle-level distribution using Bayes' theorem

$$\begin{aligned} t_j^{(n)} &= \sum_i \Pr(\text{truth}_{n-1} \text{ is } j \mid \text{measure } i) \times \Pr(\text{measure } i) \\ &= \sum_i \frac{R_{ij} t_j^{(n-1)}}{\sum_k R_{ik} t_k^{(n-1)}} \times m_i \end{aligned}$$

Iterate procedure to remove dependence on prior

Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\left(\frac{3}{4} \quad \frac{1}{4}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

$$\vdots$$

$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\left(\frac{n+2}{2(n+1)} \quad \frac{n}{2(n+1)}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the n^{th} iteration

Correct truth distribution obtained as $n \rightarrow \infty$

IBU as Reweighting

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]

Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j \quad m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i \quad R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\left(\frac{3}{4} \quad \frac{1}{4}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

⋮

$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\left(\frac{n+2}{2(n+1)} \quad \frac{n}{2(n+1)}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

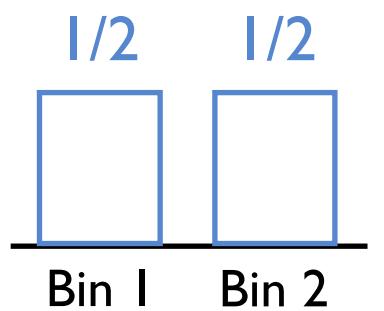
At the nth iteration

Correct truth distribution obtained as $n \rightarrow \infty$

IBU as Reweighting

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]

prior



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

Uniform prior

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

Bins are measured equally

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\left(\frac{3}{4} \quad \frac{1}{4}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/2} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

:

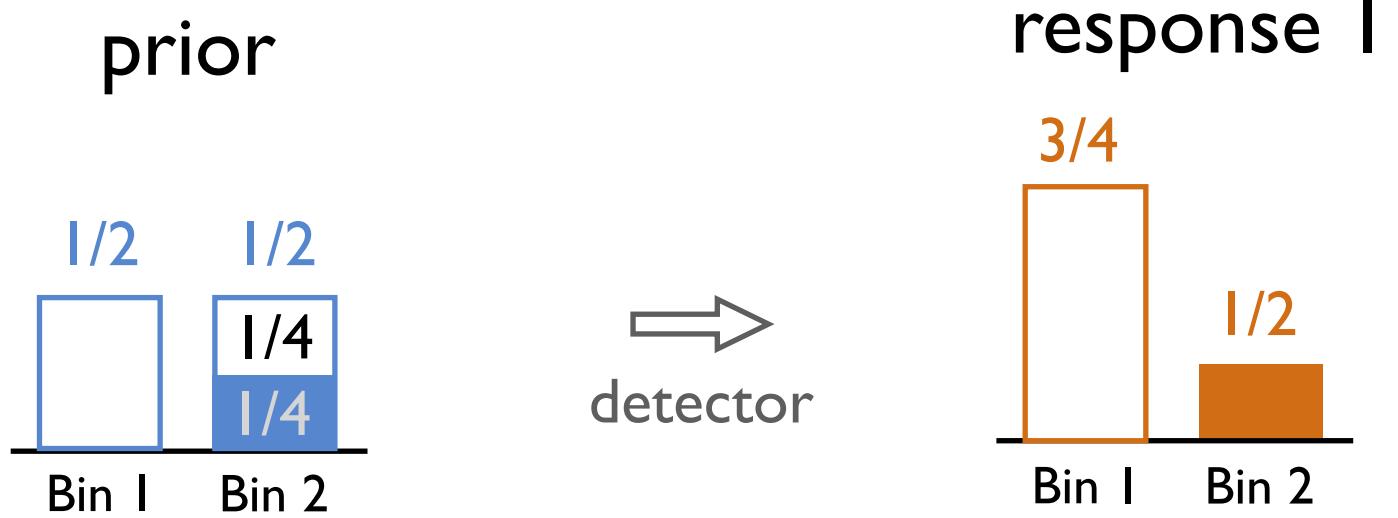
$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\left(\frac{n+2}{2(n+1)} \quad \frac{n}{2(n+1)}\right)_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the nth iteration

Correct truth distribution
obtained as $n \rightarrow \infty$

IBU as Reweighting

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

Uniform prior

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

Bins are measured equally

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/2} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

:

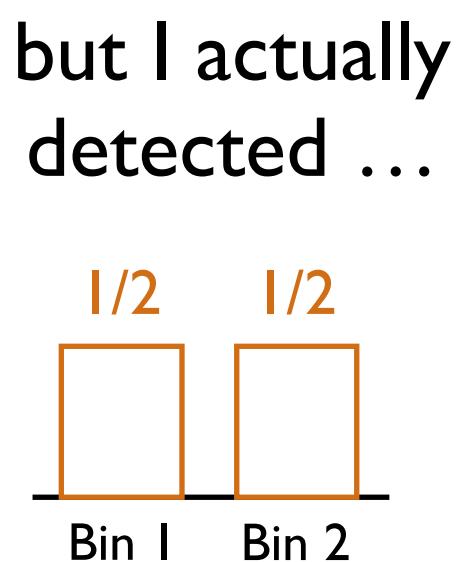
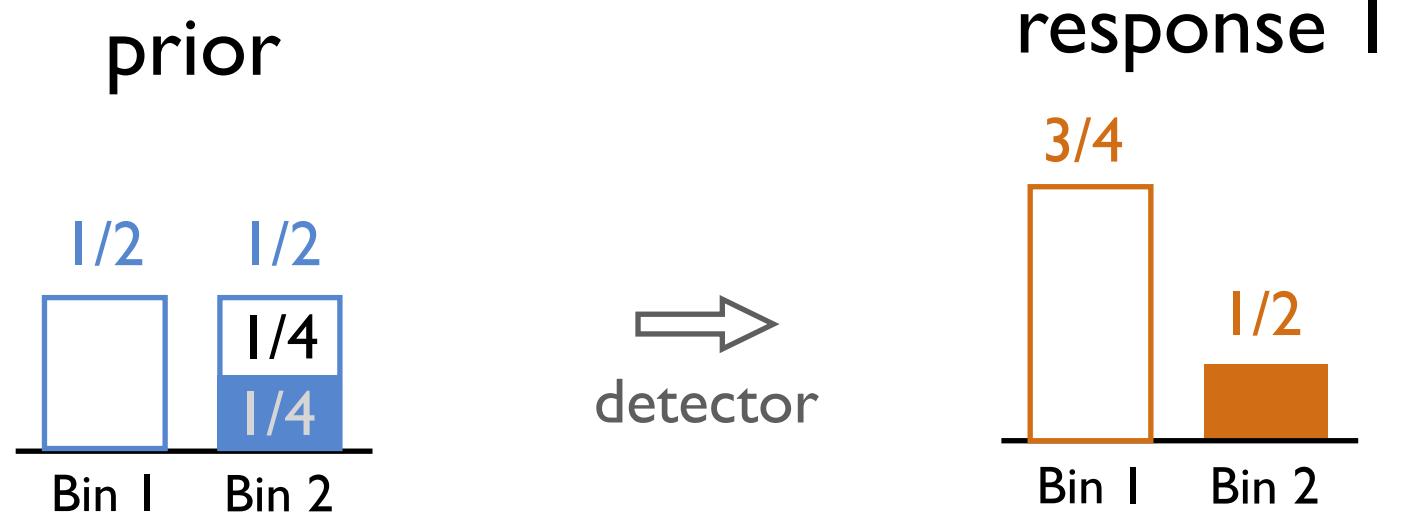
$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the nth iteration

Correct truth distribution
obtained as $n \rightarrow \infty$

IBU as Reweighting

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

⋮

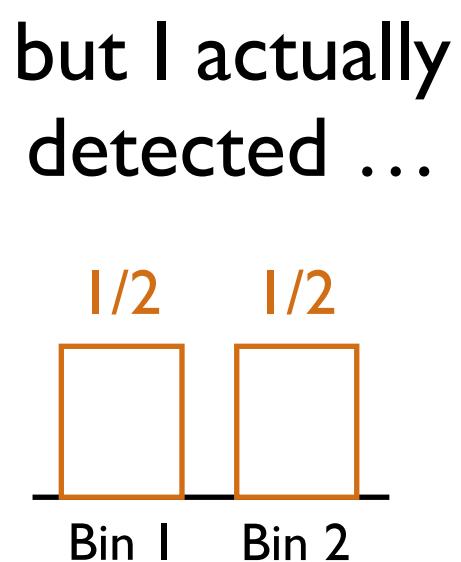
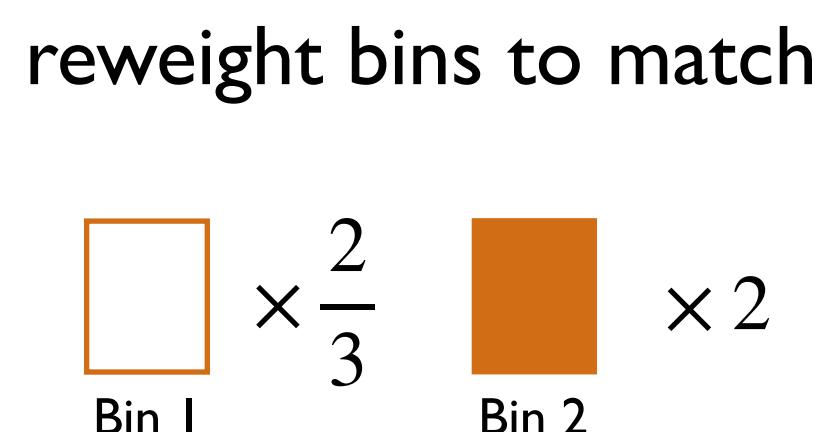
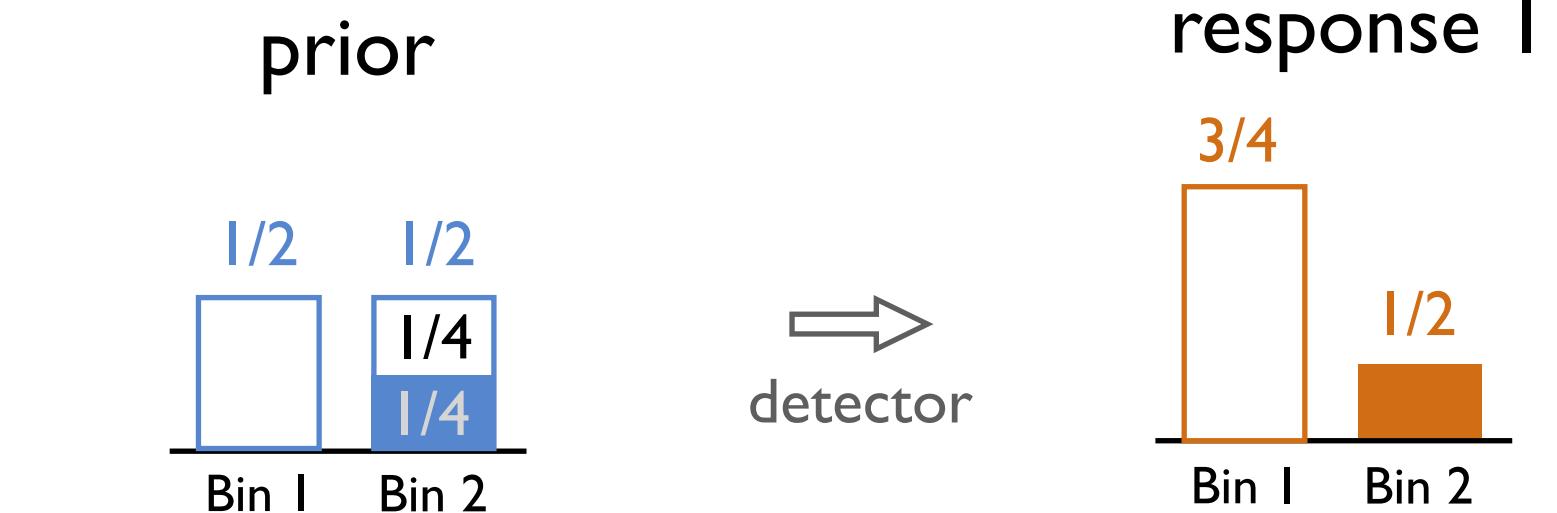
$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the nth iteration

Correct truth distribution
obtained as $n \rightarrow \infty$

IBU as Reweighting

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix}_j = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

⋮

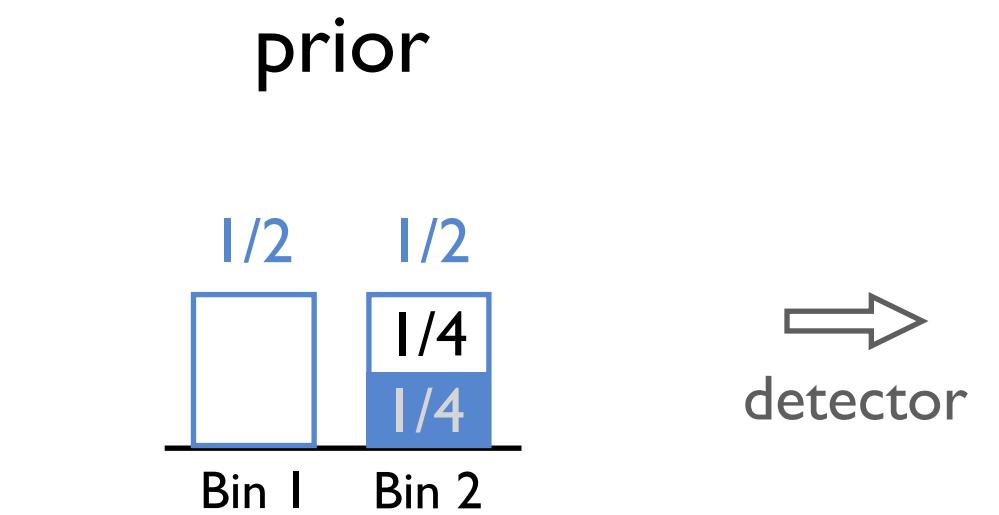
$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the n^{th} iteration

Correct truth distribution
obtained as $n \rightarrow \infty$

IBU as Reweighting

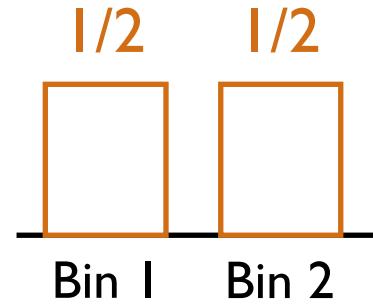
[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



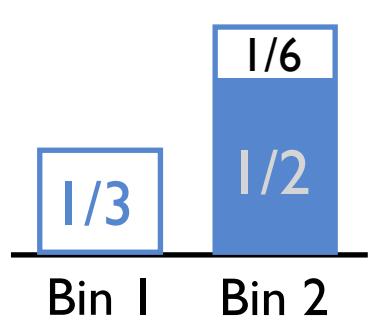
Reweight bins to match

$$\boxed{\text{Bin 1}} \times \frac{2}{3} \quad \boxed{\text{Bin 2}} \times 2$$

but I actually detected ...



Pull reweighting back to truth level



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

:

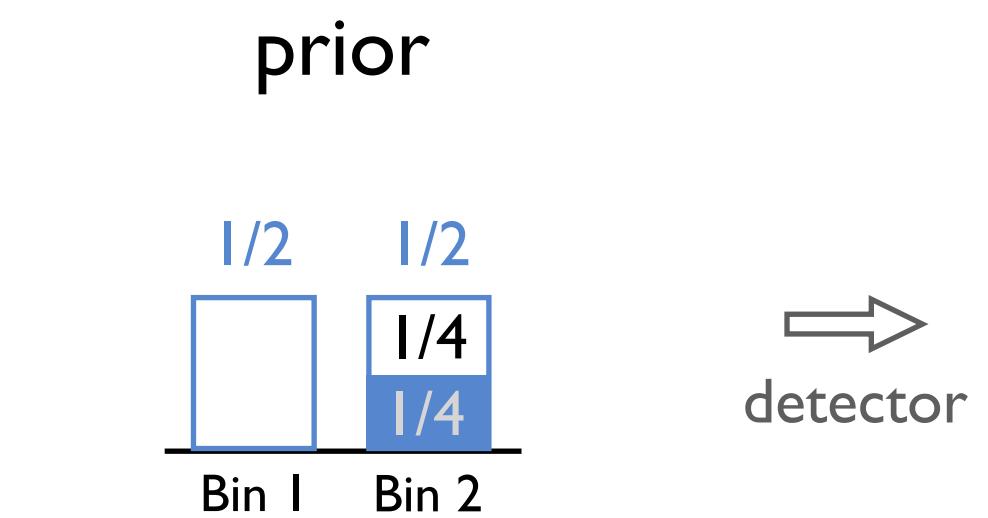
$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the n^{th} iteration

Correct truth distribution obtained as $n \rightarrow \infty$

IBU as Reweighting

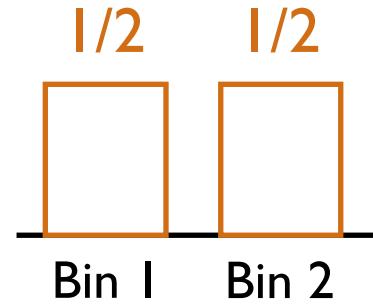
[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



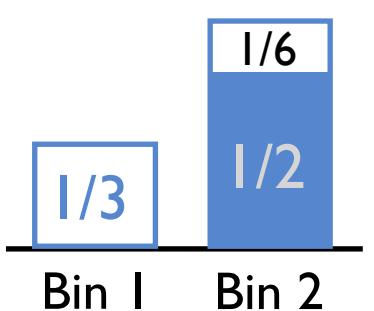
Reweight bins to match

$$\boxed{\text{Bin 1}} \times \frac{2}{3} \quad \boxed{\text{Bin 2}} \times 2$$

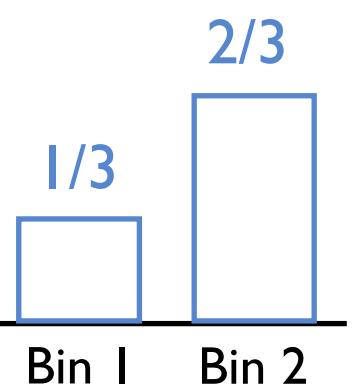
but I actually detected ...



Pull reweighting back to truth level



New estimate of truth



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

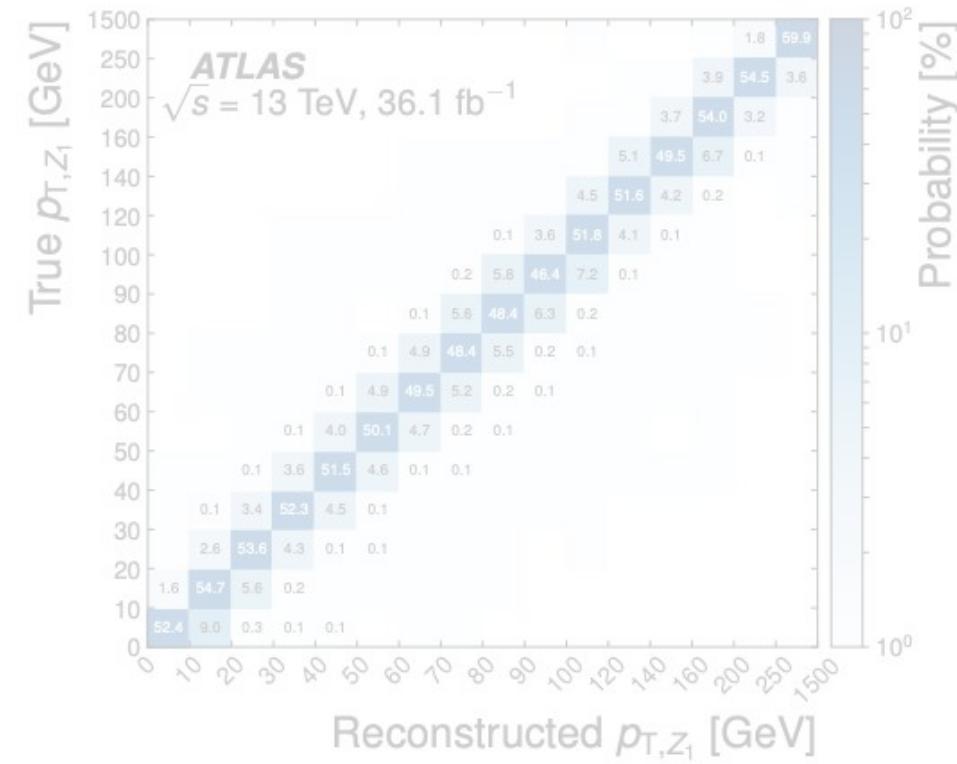
After one iteration

:

$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the n^{th} iteration

Correct truth distribution obtained as $n \rightarrow \infty$

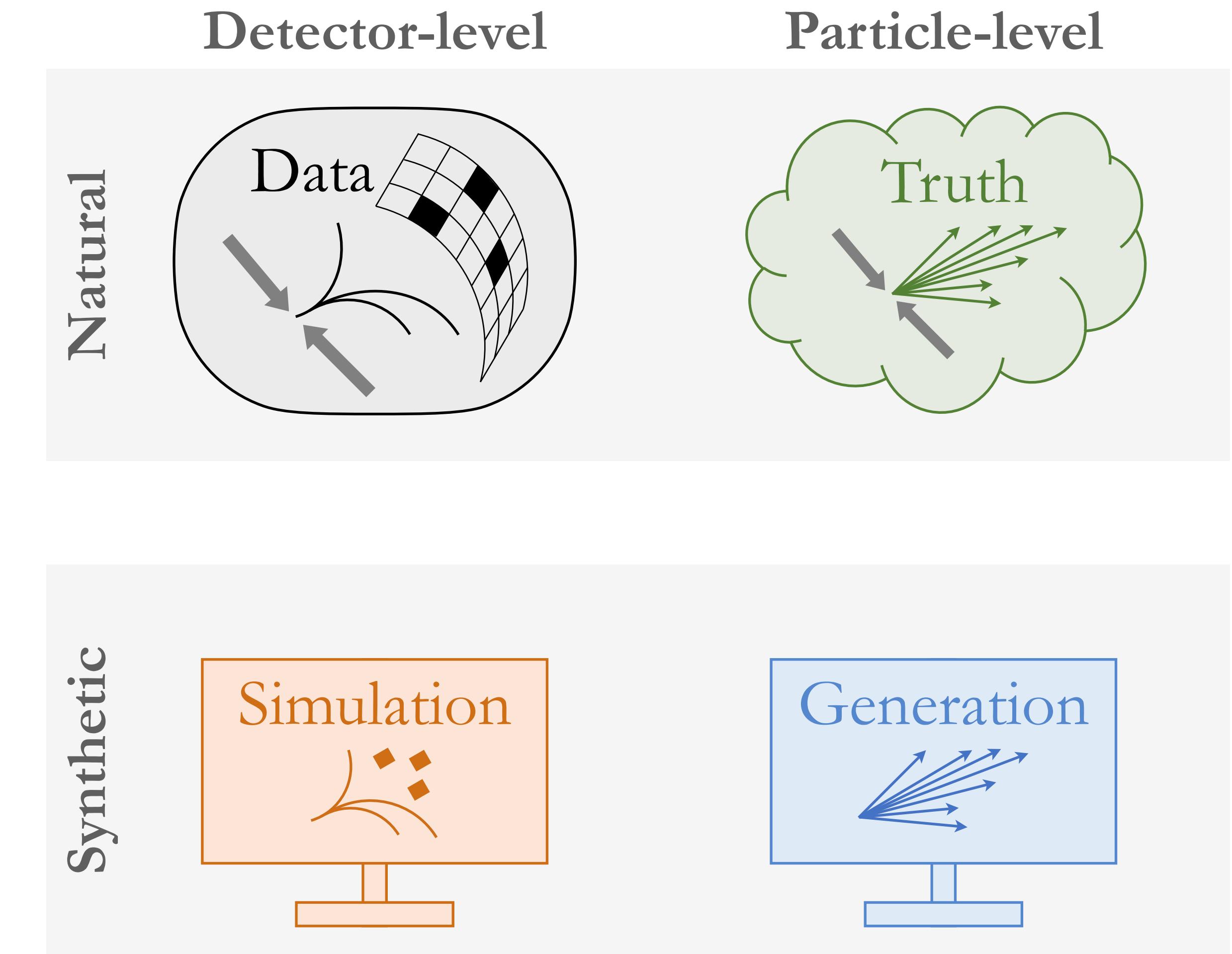


OmniFold Algorithm – Schematic



[Andreassen, PTK, Metodiev, Nachman, Thaler, [PRL 2020](#)]

OmniFold weights particle-level *Gen* to be consistent
with *Data* once passed through the detector



OmniFold Algorithm – Schematic



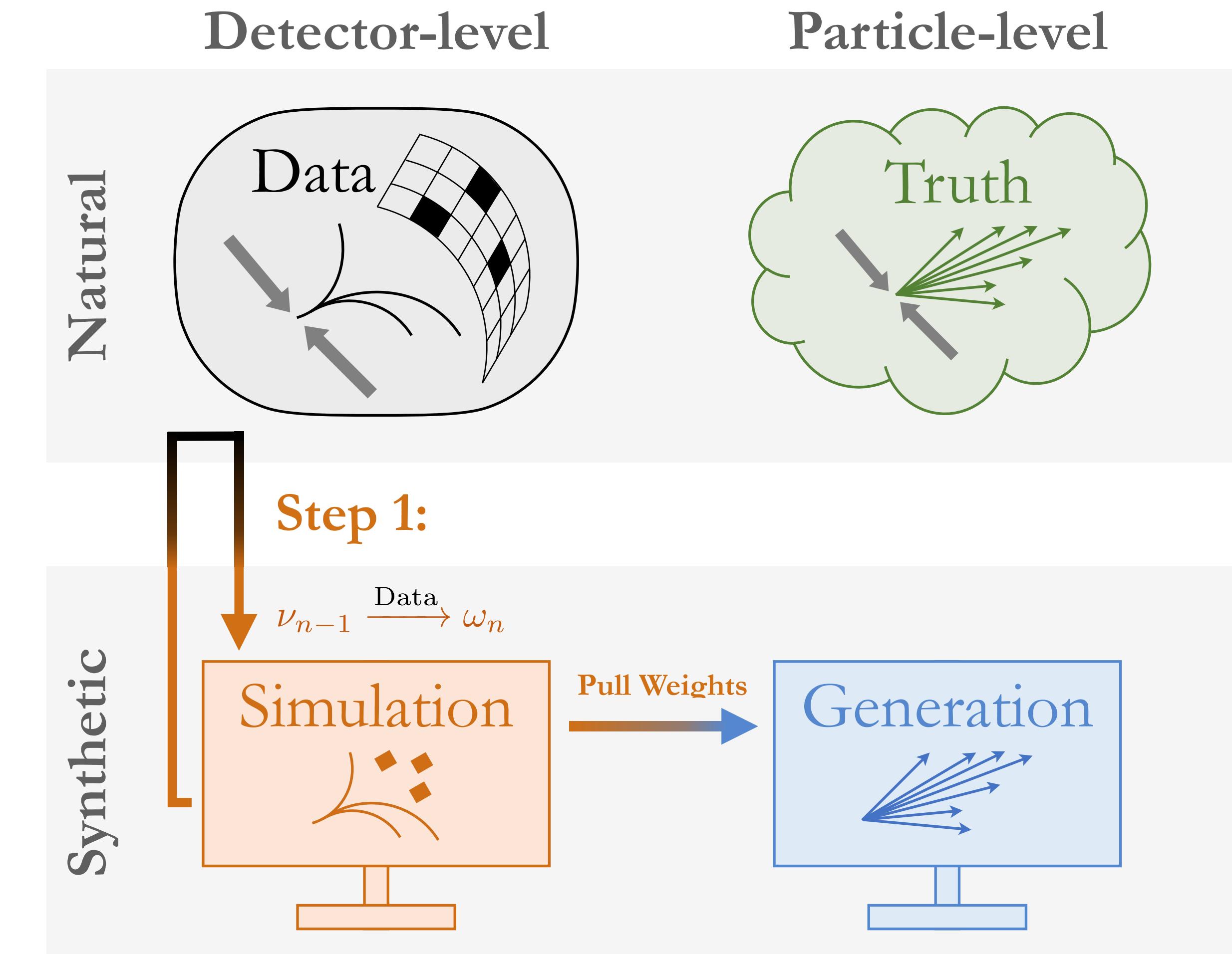
[Andreassen, PTK, Metodiev, Nachman, Thaler, [PRL 2020](#)]

OmniFold weights particle-level Gen to be consistent with Data once passed through the detector

Step 1

- Reweights Sim_{n-1} to data
- Pulls weights back to particle-level Gen_{n-1}

Incorporates the response matrix



OmniFold Algorithm – Schematic



[Andreassen, PTK, Metodiev, Nachman, Thaler, [PRL 2020](#)]

OmniFold weights particle-level Gen to be consistent with Data once passed through the detector

Step 1

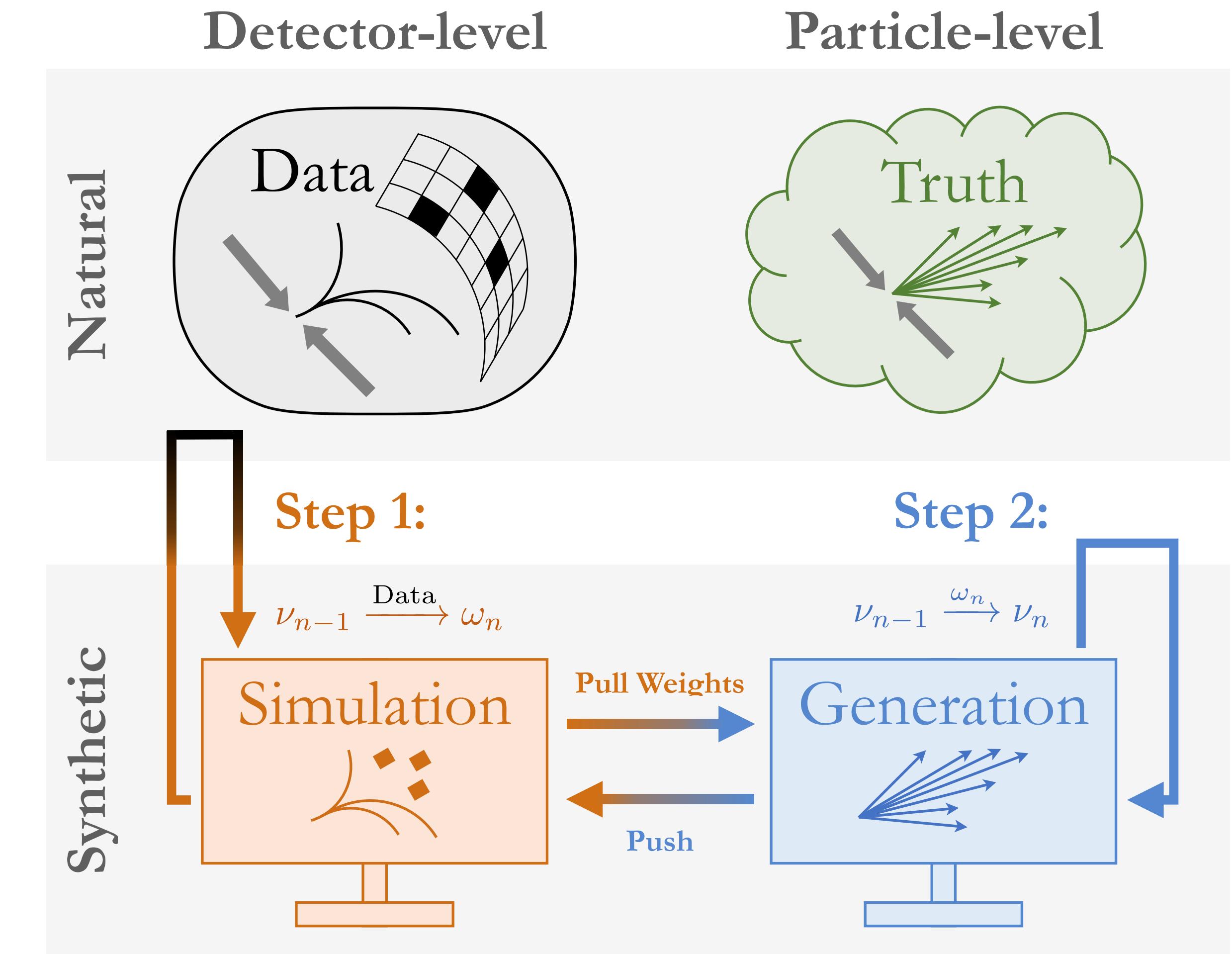
- Reweights Sim_{n-1} to data
- Pulls weights back to particle-level Gen_{n-1}

Incorporates the response matrix

Step 2

- Reweights Gen_{n-1} to (step 1)-weighted gen_{n-1}
- Pushes weights to detector-level Sim_n

Constructs valid particle-level function by averaging gen-level weights



Unfolding via Likelihood Reweighting

Likelihood ratio is optimal binary classifier by Neyman-Pearson lemma

$$L[(w, X), (w', X')](x) = \frac{p_{(w, X)}(x)}{p_{(w', X')}(x)}$$

L – likelihood ratio

w – weights

X – phase space

x – element of X

p – probability density

Unfolding via Likelihood Reweighting

Likelihood ratio is optimal binary classifier by Neyman-Pearson lemma

$$L[(w, X), (w', X')](x) = \frac{p_{(w, X)}(x)}{p_{(w', X')}(x)}$$

L – likelihood ratio

w – weights

X – phase space

x – element of X

p – probability density

Model output of a well-trained classifier accesses likelihood ratio

$$\text{Model}[(w, X), (w', X')](x) \simeq \frac{L[(w, X), (w', X')](x)}{1 + L[(w, X), (w', X')](x)} \quad \text{Assuming softmax output}$$

[Cranmer, Pavez, Louppe, [1506.02169](#); Andreassen, Nachman, [PRD 2020](#)]

Unfolding via Likelihood Reweighting

Likelihood ratio is optimal binary classifier by Neyman-Pearson lemma

$$L[(w, X), (w', X')](x) = \frac{p_{(w, X)}(x)}{p_{(w', X')}(x)}$$

L – likelihood ratio

w – weights

X – phase space

x – element of X

p – probability density

Model output of a well-trained classifier accesses likelihood ratio

$$\text{Model}[(w, X), (w', X')](x) \simeq \frac{L[(w, X), (w', X')](x)}{1 + L[(w, X), (w', X')](x)} \quad \text{Assuming softmax output}$$

[Cranmer, Pavez, Louppe, [1506.02169](#); Andreassen, Nachman, [PRD 2020](#)]

OmniFold repeatedly reweights one weighted sample (A) to another (B)

$$w_{A'}(x) = w_A(x) \times \frac{\text{Model}[(w_B, B), (w_A, A)](x)}{1 - \text{Model}[(w_B, B), (w_A, A)](x)} \quad A' \text{ is statistically indistinguishable from } B$$

Likelihood reweighting benefits from architectural improvements

OmniFold Algorithm – Equations



[Andreassen, PTK, Metodiev, Nachman, Thaler, [PRL 2020](#)]

Inputs

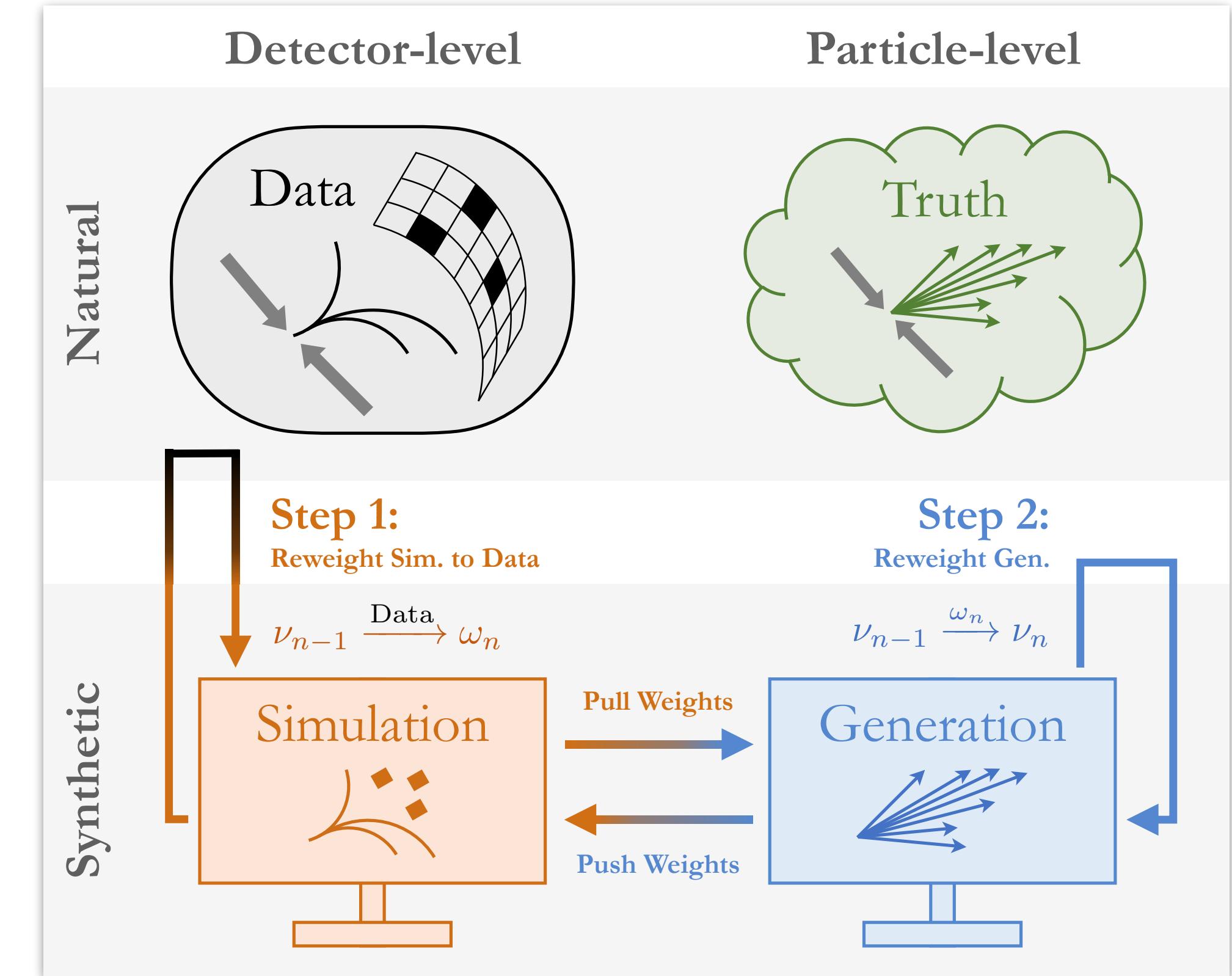
(t, m) – pairs of **Gen** and **Sim** events
 $\nu_0(t)$ – initial particle-level weights for **Gen**
– Data

Results of Steps 1 and 2

$\nu_n(t)$ – particle-level weights for **Gen**, n^{th} iteration
 $\omega_n(m)$ – detector-level weights for **Sim**, n^{th} iteration

Pulling/Pushing Weights

$\omega_n^{\text{pull}}(t) = \omega_n(m)$ – pulling ω_n back to particle-level
 $\nu_n^{\text{push}}(m) = \nu_n(t)$ – pushing ν_n to detector-level



OmniFold Algorithm – Equations



[Andreassen, PTK, Metodiev, Nachman, Thaler, [PRL 2020](#)]

Inputs

(t, m) – pairs of **Gen** and **Sim** events
 $\nu_0(t)$ – initial particle-level weights for **Gen**
– Data

Results of Steps 1 and 2

$\nu_n(t)$ – particle-level weights for **Gen**, n^{th} iteration
 $\omega_n(m)$ – detector-level weights for **Sim**, n^{th} iteration

Pulling/Pushing Weights

$\omega_n^{\text{pull}}(t) = \omega_n(m)$ – pulling ω_n back to particle-level
 $\nu_n^{\text{push}}(m) = \nu_n(t)$ – pushing ν_n to detector-level

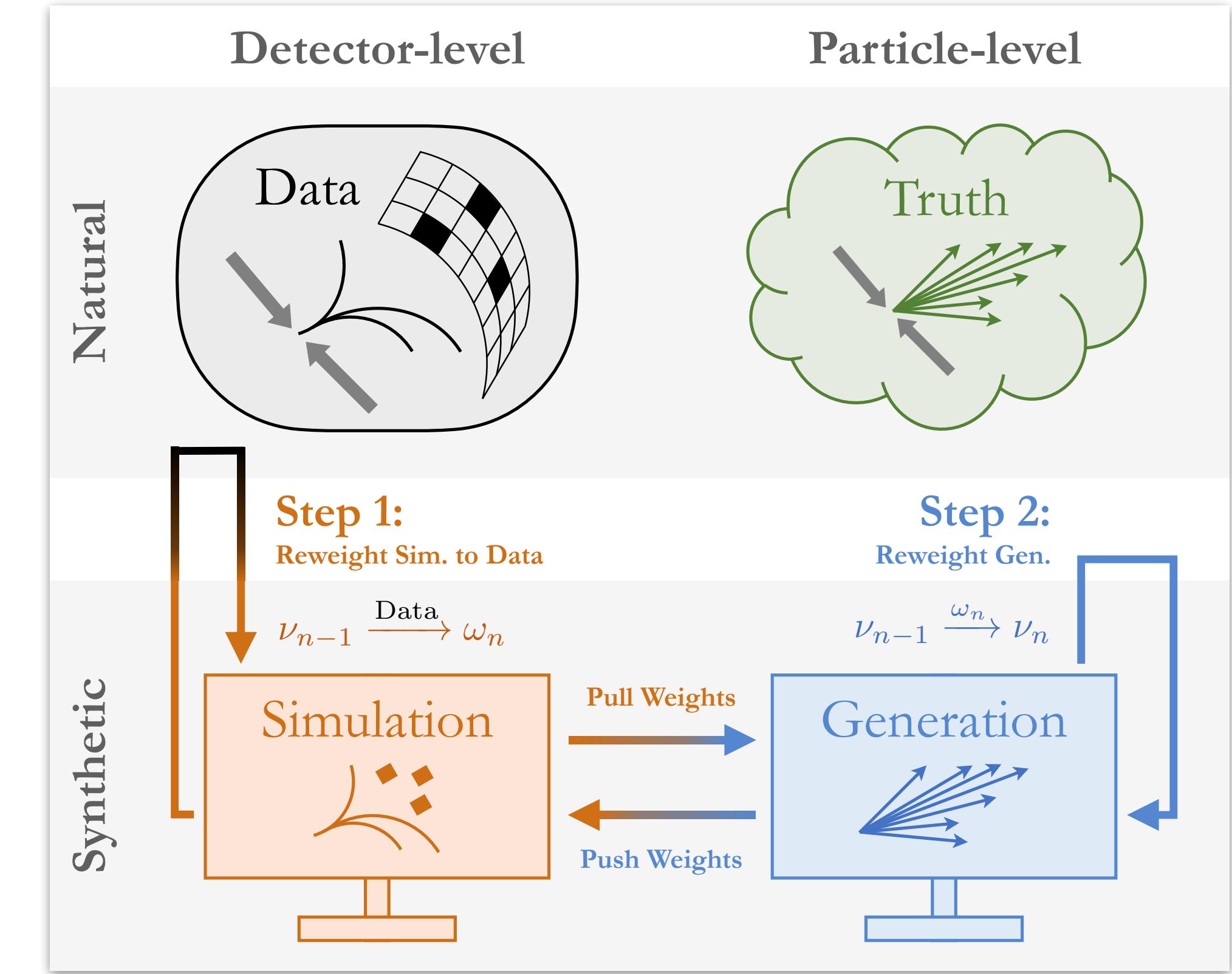
OmniFold

$$\text{Step 1} - \omega_n(m) = \nu_{n-1}^{\text{push}} \times L[(1, \text{Data}), (\nu_{n-1}^{\text{push}}, \text{Sim})](m)$$

$$\text{Step 2} - \nu_n(t) = \nu_{n-1}(t) \times L[(\omega_n^{\text{pull}}, \text{Gen}), (\nu_{n-1}, \text{Gen})](t)$$

Unfold any* observable $p_{\text{Gen}}(t)$ using universal weights $\nu_n(t)$

$$p_{\text{unfolded}}^{(n)}(t) = \nu_n(t) \times p_{\text{Gen}}(t)$$



OmniFold Algorithm – Equations



[Andreassen, PTK, Metodiev, Nachman, Thaler, [PRL 2020](#)]

Inputs

(t, m) – pairs of **Gen** and **Sim** events
 $\nu_0(t)$ – initial particle-level weights for **Gen**
– Data

Results of Steps 1 and 2

$\nu_n(t)$ – particle-level weights for **Gen**, n^{th} iteration
 $\omega_n(m)$ – detector-level weights for **Sim**, n^{th} iteration

Pulling/Pushing Weights

$\omega_n^{\text{pull}}(t) = \omega_n(m)$ – pulling ω_n back to particle-level
 $\nu_n^{\text{push}}(m) = \nu_n(t)$ – pushing ν_n to detector-level

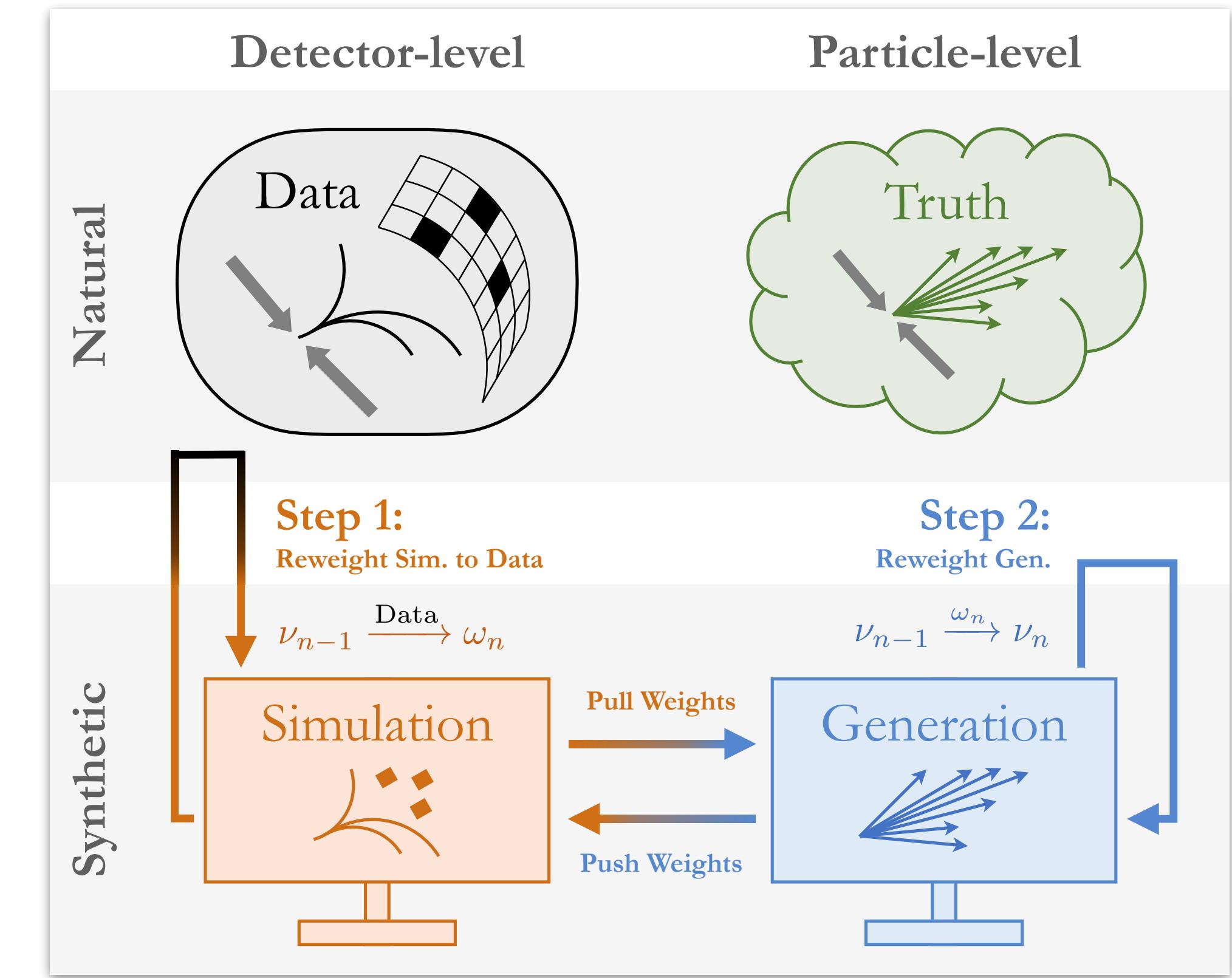
OmniFold

$$\text{Step 1} - \omega_n(m) = \nu_{n-1}^{\text{push}} \times L[(1, \text{Data}), (\nu_{n-1}^{\text{push}}, \text{Sim})](m)$$

$$\text{Step 2} - \nu_n(t) = \nu_{n-1}(t) \times L[(\omega_n^{\text{pull}}, \text{Gen}), (\nu_{n-1}, \text{Gen})](t)$$

Unfold any* observable $p_{\text{Gen}}(t)$ using universal weights $\nu_n(t)$

$$p_{\text{unfolded}}^{(n)}(t) = \nu_n(t) \times p_{\text{Gen}}(t)$$



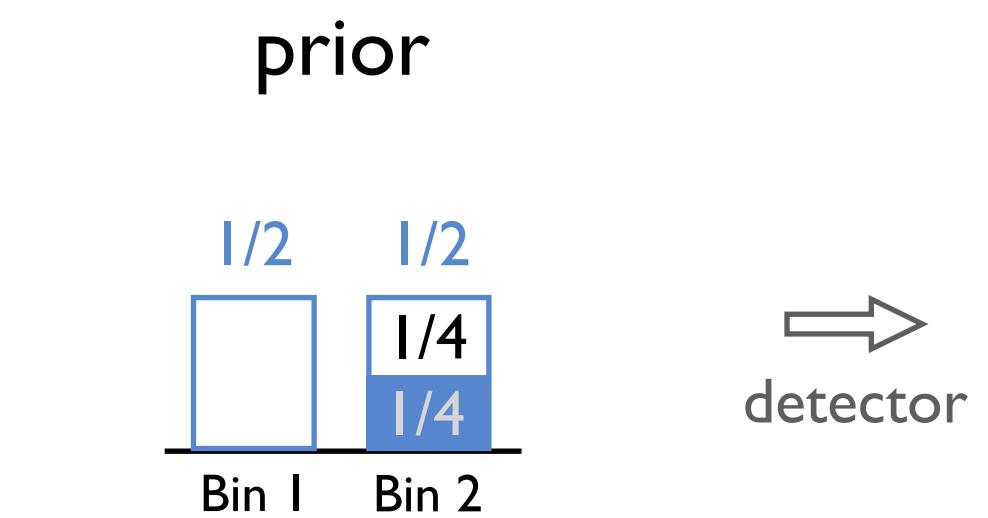
OmniFold is continuous IBU!

After first iteration, with $\nu_0(t) = 1$:

$$\nu_1(t)p_{\text{Gen}}(t) = \int dm p_{\text{Gen}|\text{Sim}}(t|m) p_{\text{Data}}(m)$$

IBU as Reweighting

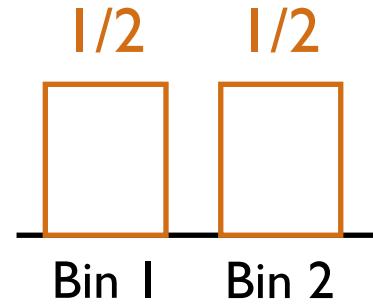
[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



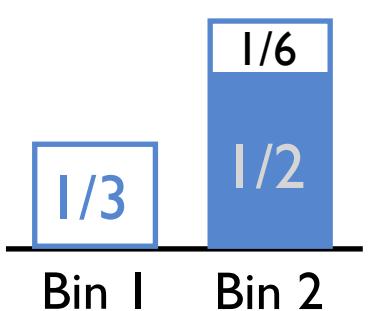
Reweight bins to match

$$\boxed{\text{Bin 1}} \times \frac{2}{3} \quad \boxed{\text{Bin 2}} \times 2$$

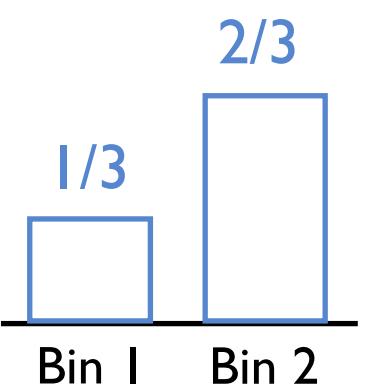
but I actually detected ...



Pull reweighting back to truth level



New estimate of truth



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

:

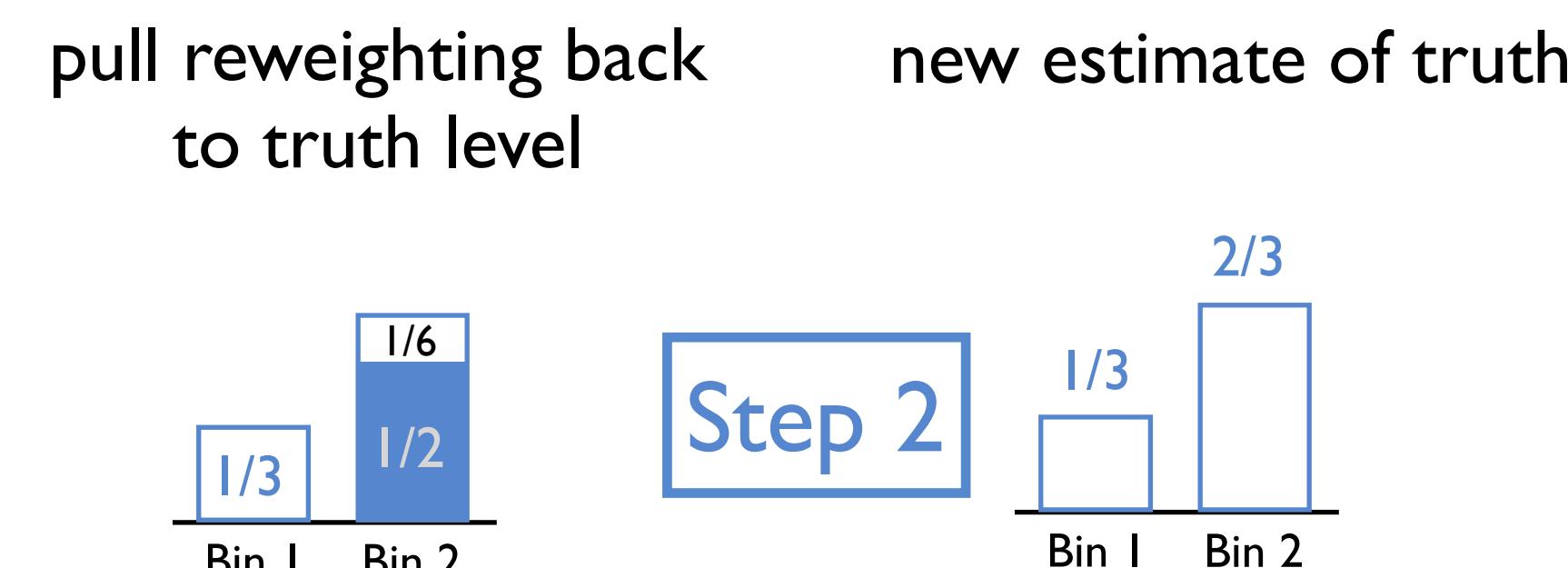
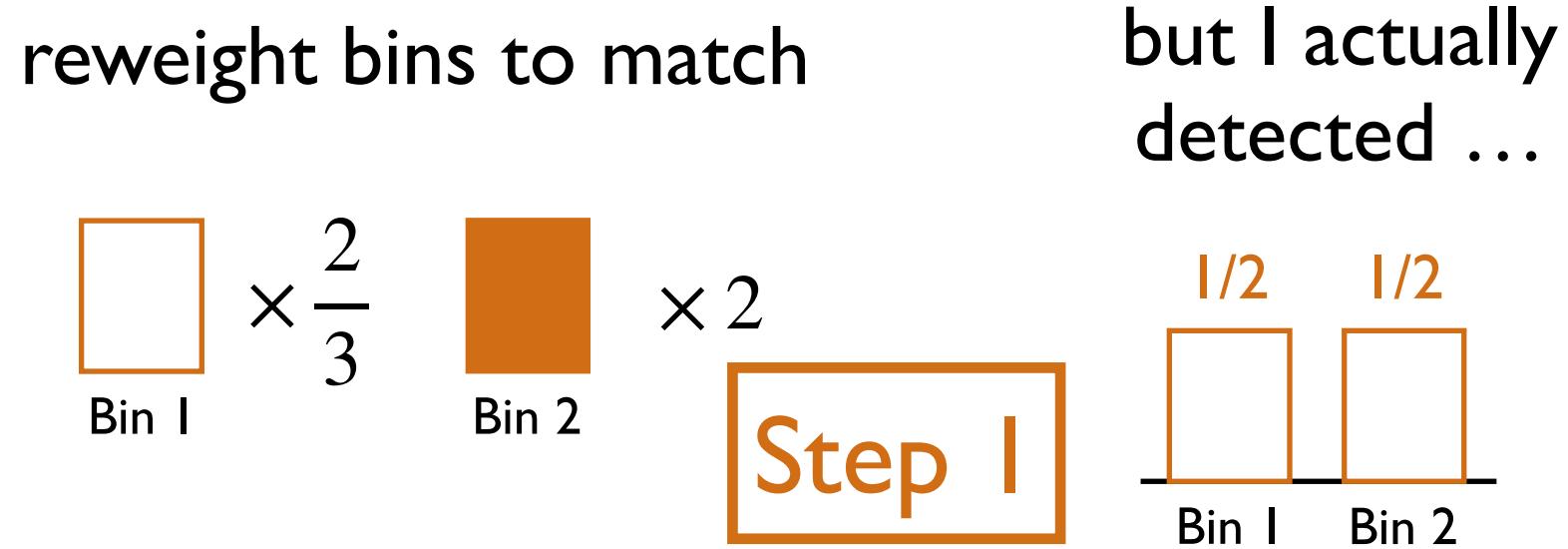
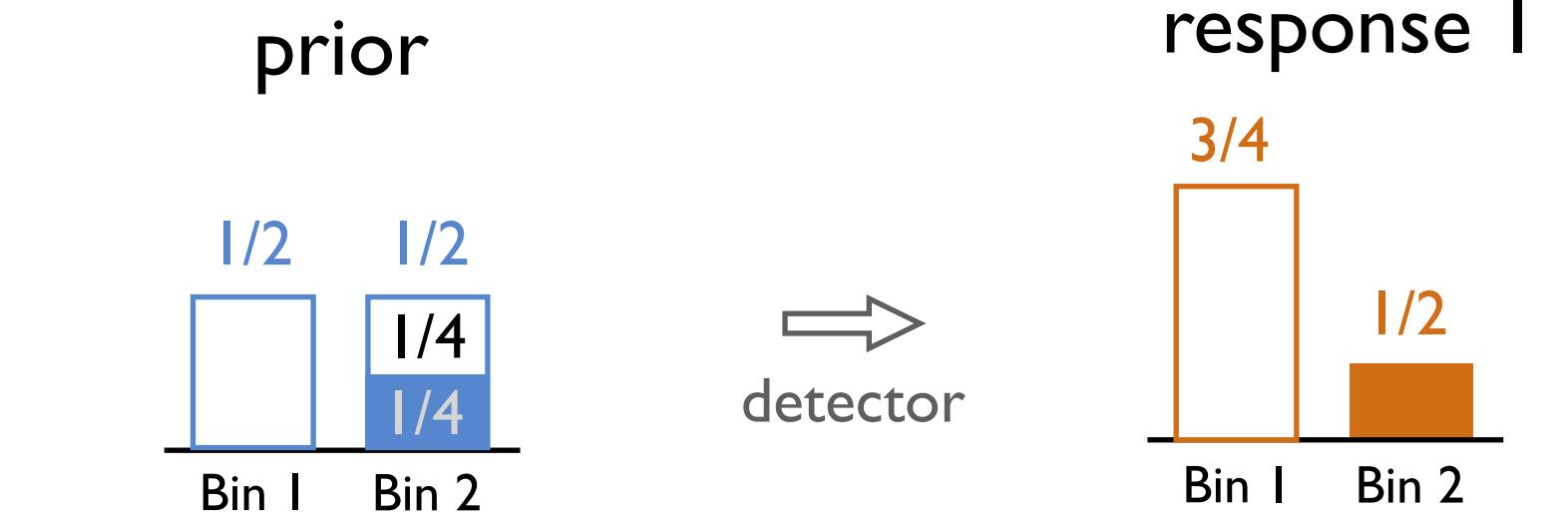
$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the n^{th} iteration

Correct truth distribution obtained as $n \rightarrow \infty$

IBU as Reweighting

[Richardson, JOSA 1972; Lucy, AJ 1974; D'Agostini, NIMPA 1995]



Consider a situation with two particle-level bins and two detector-level bins

$$t_j^{(0)} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_j$$

$$m_i = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i$$

$$R_{ij} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}_{ij}$$

Uniform prior

Bins are measured equally

Bin 1 reconstructed perfectly
Bin 2 reconstructed equally

$$t_j^{(1)} = \sum_i \frac{\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{2} + 0}_{1/3} \\ \underbrace{\frac{1}{2} \times \frac{4}{3} \times \frac{1}{4} + \frac{1}{2} \times 4 \times \frac{1}{4}}_{1/6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}_j$$

After one iteration

$$t_j^{(n)} = \sum_i \frac{\begin{pmatrix} \frac{1}{n+1} & \frac{n}{2(n+1)} \\ 0 & \frac{n}{2(n+1)} \end{pmatrix}_{ij}}{\begin{pmatrix} \frac{n+2}{2(n+1)} & \frac{n}{2(n+1)} \end{pmatrix}_i} \times \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_i = \begin{pmatrix} \frac{1}{n+2} \\ \frac{n+1}{n+2} \end{pmatrix}_j \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}_j$$

At the nth iteration

Constructing High-Dimensional Classifiers

How to represent jets to a machine learning architecture?

An **unordered**, **variable length** collection of particles

Due to quantum-mechanical indistinguishability

Due to probabilistic nature of jet formation

$$J(\{p_1^\mu, \dots, p_M^\mu\}) = J(\{p_{\pi(1)}^\mu, \dots, p_{\pi(M)}^\mu\}), \quad \underbrace{M \geq 1}_{\text{Multiplicity}}, \quad \underbrace{\forall \pi \in S_M}_{\text{Permutations}}$$

p_i^μ represents *all* the particle properties:

- Four-momentum – $(E, p_x, p_y, p_z)_i^\mu$
- Other quantum numbers (e.g. particle id, charge)
- Experimental information (e.g. vertex info, quality criteria, tracking info)

Methods for processing point clouds/jets should respect the appropriate symmetries

Machine Learning for Point Clouds – Deep Sets

A general permutation-symmetric function is additive in a latent space

Deep Sets

[[1703.06114](#)]

**Manzil Zaheer^{1,2}, Satwik Kottur¹, Siamak Ravanbakhsh¹,
Barnabás Póczos¹, Ruslan Salakhutdinov¹, Alexander J Smola^{1,2}**
¹ Carnegie Mellon University ² Amazon Web Services

Deep Sets Theorem [63]. Let $\mathfrak{X} \subset \mathbb{R}^d$ be compact, $X \subset 2^{\mathfrak{X}}$ be the space of sets with bounded cardinality of elements in \mathfrak{X} , and $Y \subset \mathbb{R}$ be a bounded interval. Consider a continuous function $f : X \rightarrow Y$ that is invariant under permutations of its inputs, i.e. $f(x_1, \dots, x_M) = f(x_{\pi(1)}, \dots, x_{\pi(M)})$ for all $x_i \in \mathfrak{X}$ and $\pi \in S_M$. Then there exists a sufficiently large integer ℓ and continuous functions $\Phi : \mathfrak{X} \rightarrow \mathbb{R}^\ell$, $F : \mathbb{R}^\ell \rightarrow Y$ such that the following holds to an arbitrarily good approximation:¹

$$f(\{x_1, \dots, x_M\}) = F \left(\sum_{i=1}^M \Phi(x_i) \right)$$

Machine Learning for Point Clouds – Deep Sets

A general permutation-symmetric function is additive in a latent space

Deep Sets

[1703.06114]

Manzil Zaheer^{1,2}, Satwik Kottur¹, Siamak Ravanbakhsh¹,
Barnabás Póczos¹, Ruslan Salakhutdinov¹, Alexander J Smola^{1,2}
¹ Carnegie Mellon University ² Amazon Web Services

Feature space
Permutation invariance

Deep Sets Theorem [63]. Let $\mathfrak{X} \subset \mathbb{R}^d$ be compact, $X \subset 2^{\mathfrak{X}}$ be the space of sets with bounded cardinality of elements in \mathfrak{X} , and $Y \subset \mathbb{R}$ be a bounded interval. Consider a continuous function $f : X \rightarrow Y$ that is invariant under permutations of its inputs, i.e. $f(x_1, \dots, x_M) = f(x_{\pi(1)}, \dots, x_{\pi(M)})$ for all $x_i \in \mathfrak{X}$ and $\pi \in S_M$. Then there exists a sufficiently large integer ℓ and continuous functions $\Phi : \mathfrak{X} \rightarrow \mathbb{R}^\ell$, $F : \mathbb{R}^\ell \rightarrow Y$ such that the following holds to an arbitrarily good approximation:¹

$$f(\{x_1, \dots, x_M\}) = F \left(\sum_{i=1}^M \Phi(x_i) \right)$$

Variable length

Latent space

General parametrization for a function of sets

Approximating Φ and F with Neural Networks

[PTK, Metodiev, Thaler, JHEP 2019]

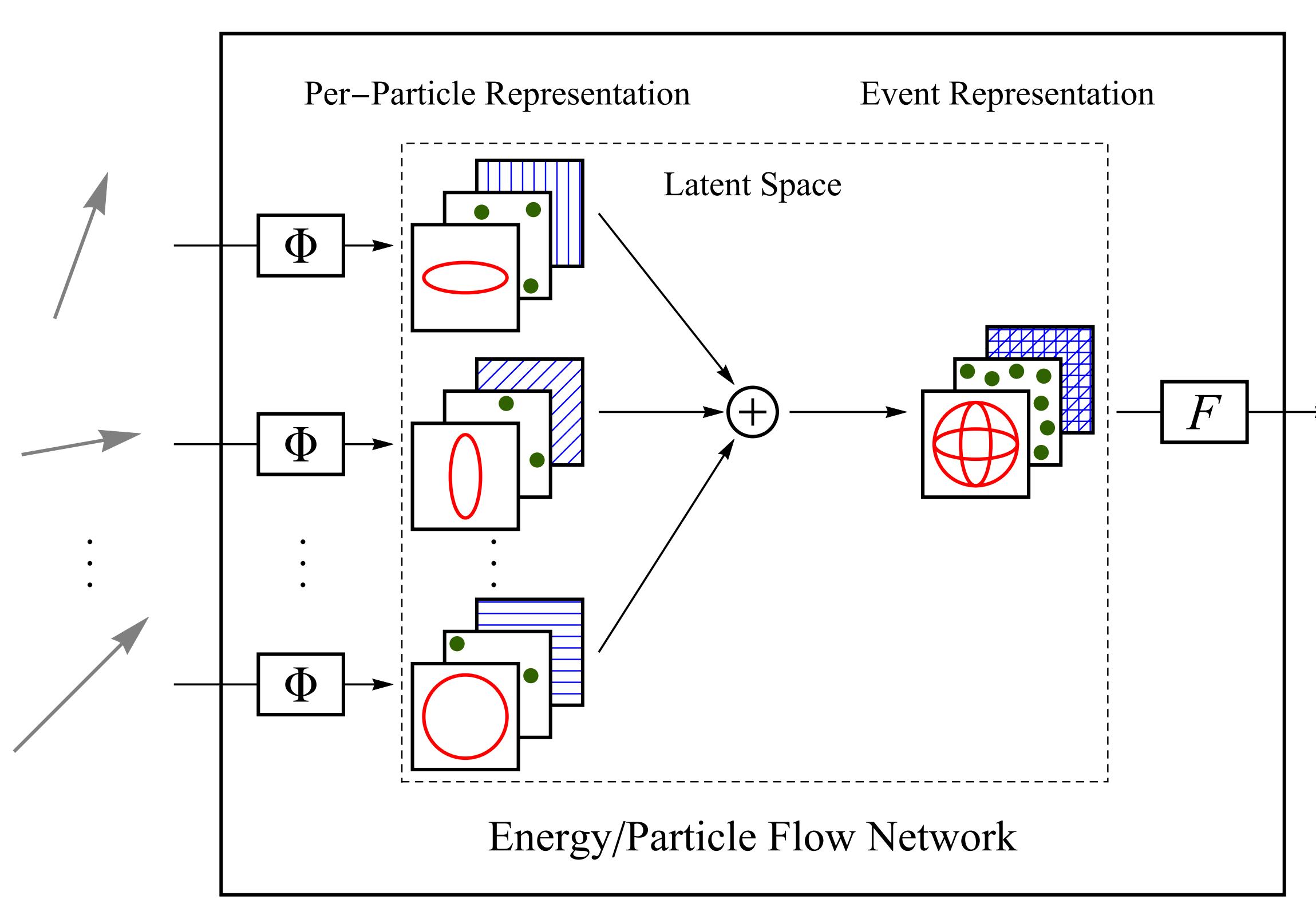
Employ neural networks as arbitrary function approximators

Use fully-connected networks for simplicity

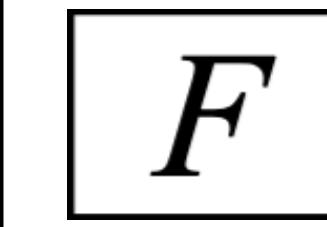
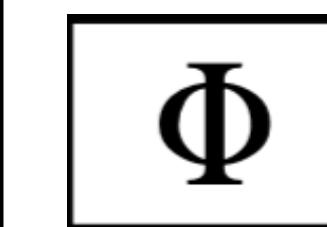
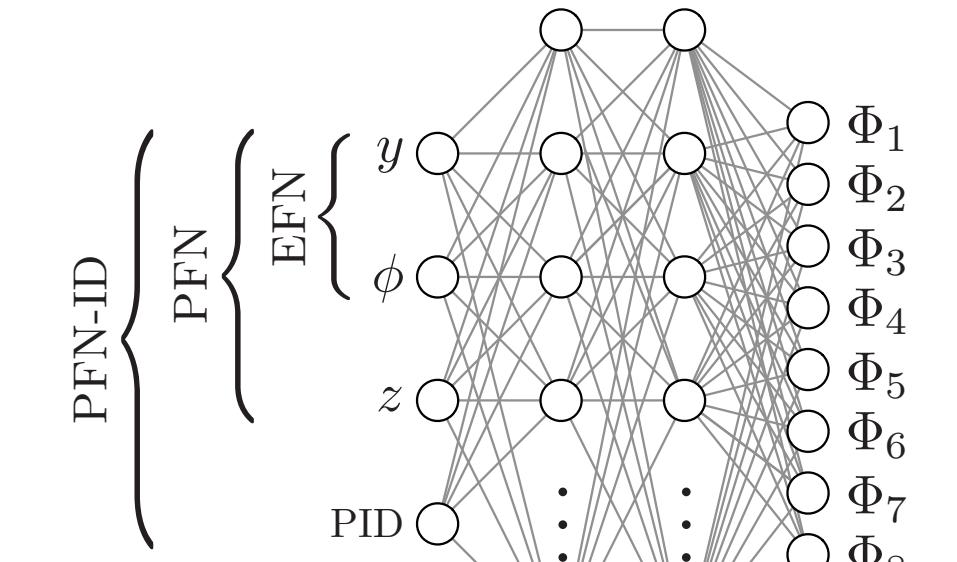
Default sizes – $\Phi: (100, 100, \ell)$, $F: (100, 100, 100)$

Particles

Observable



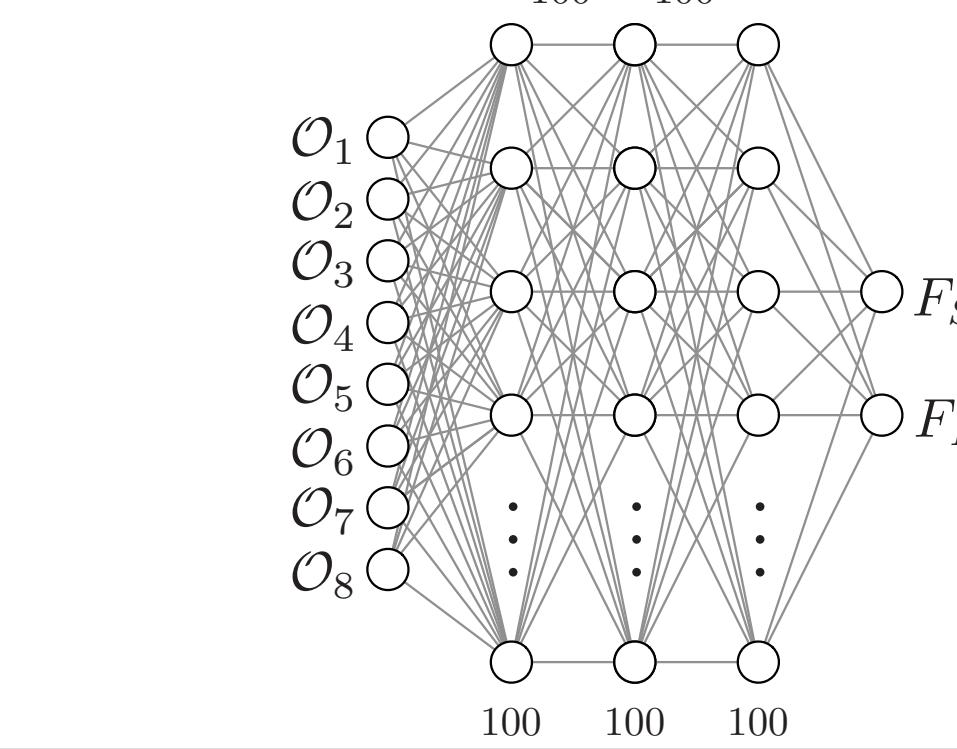
Legend



Particle Flow Network (PFN)

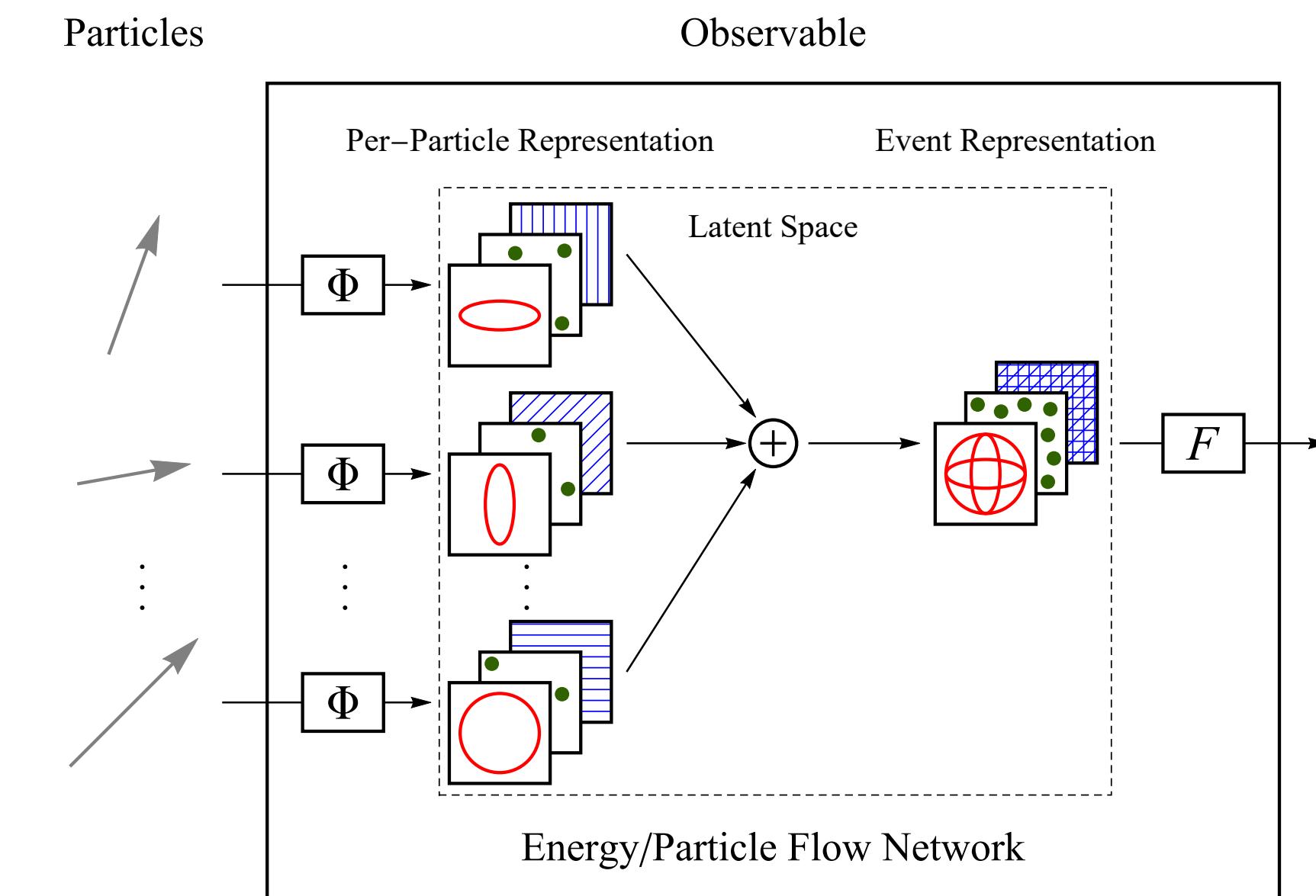
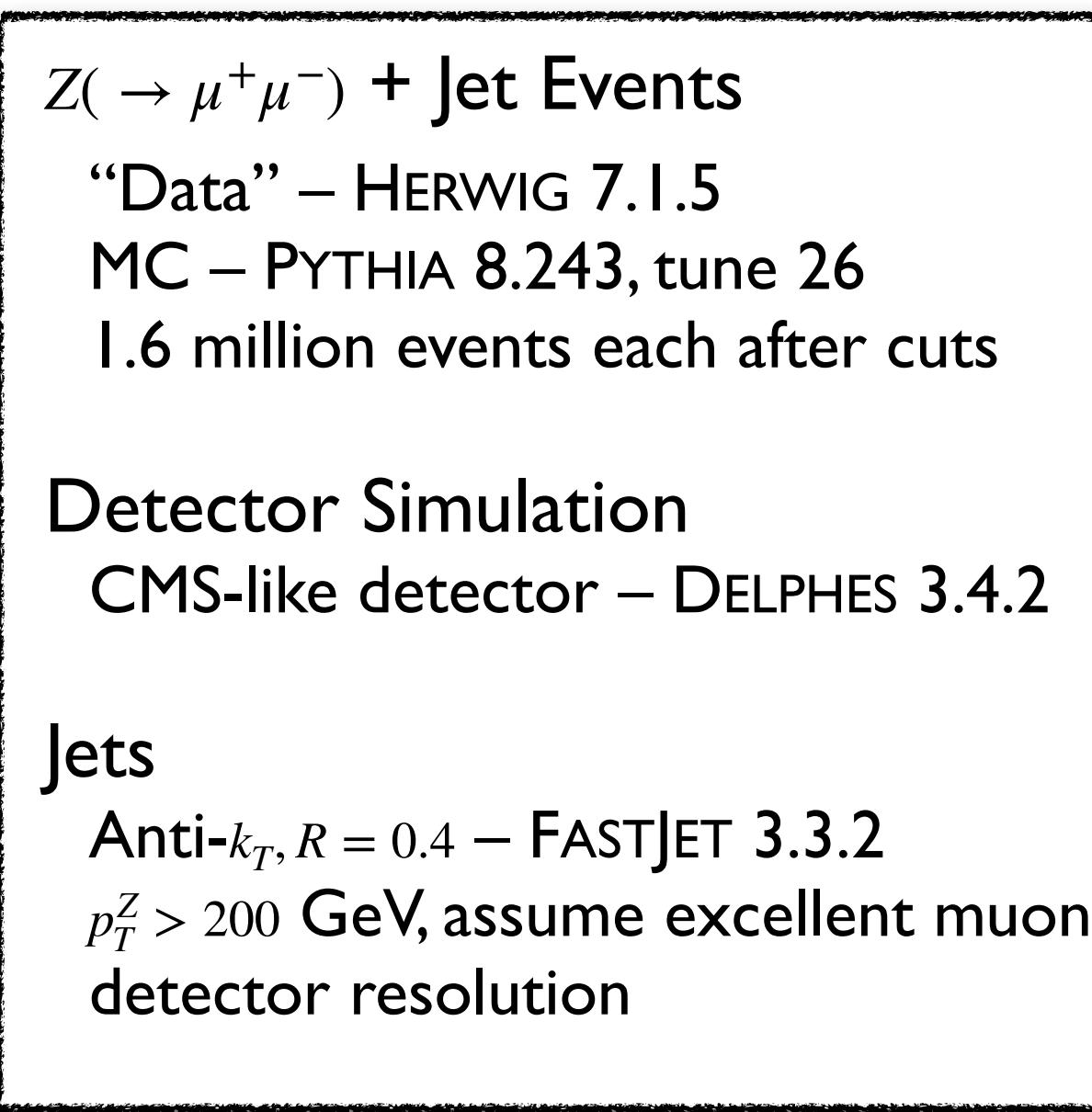
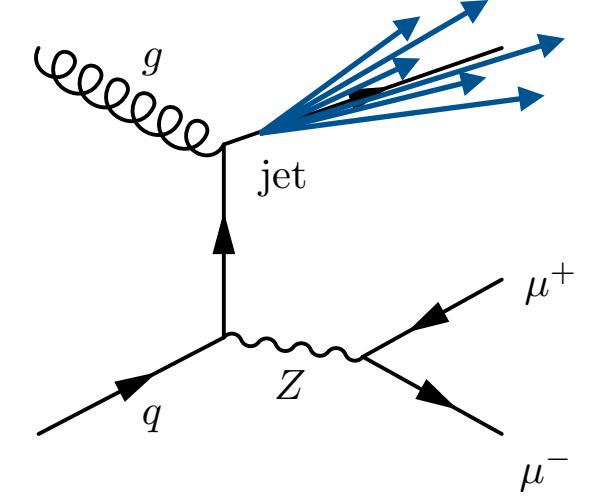
$$\text{PFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \Phi(p_i^\mu) \right)$$

Fully general latent space

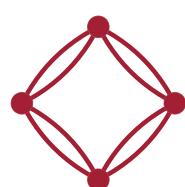


Testing OmniFold

Ingredients for $Z + \text{Jet}$ Case Study



Datasets publicly available
– With two additional Pythia tunes
– Accessible via [EnergyFlow](#)



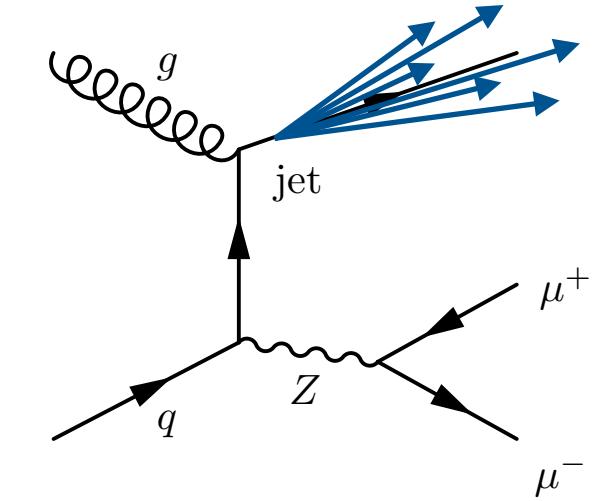
OmniFold Binder Demo



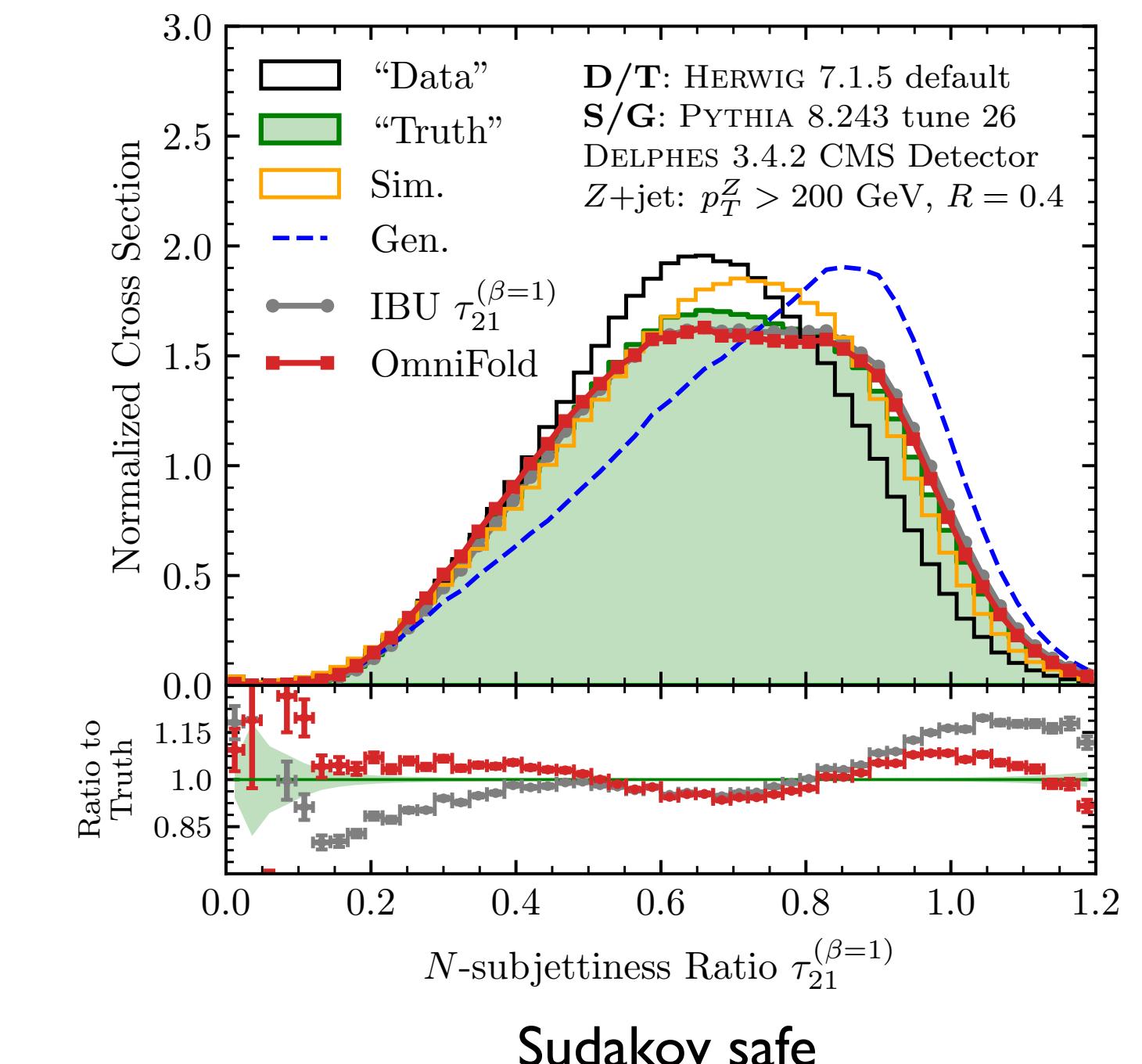
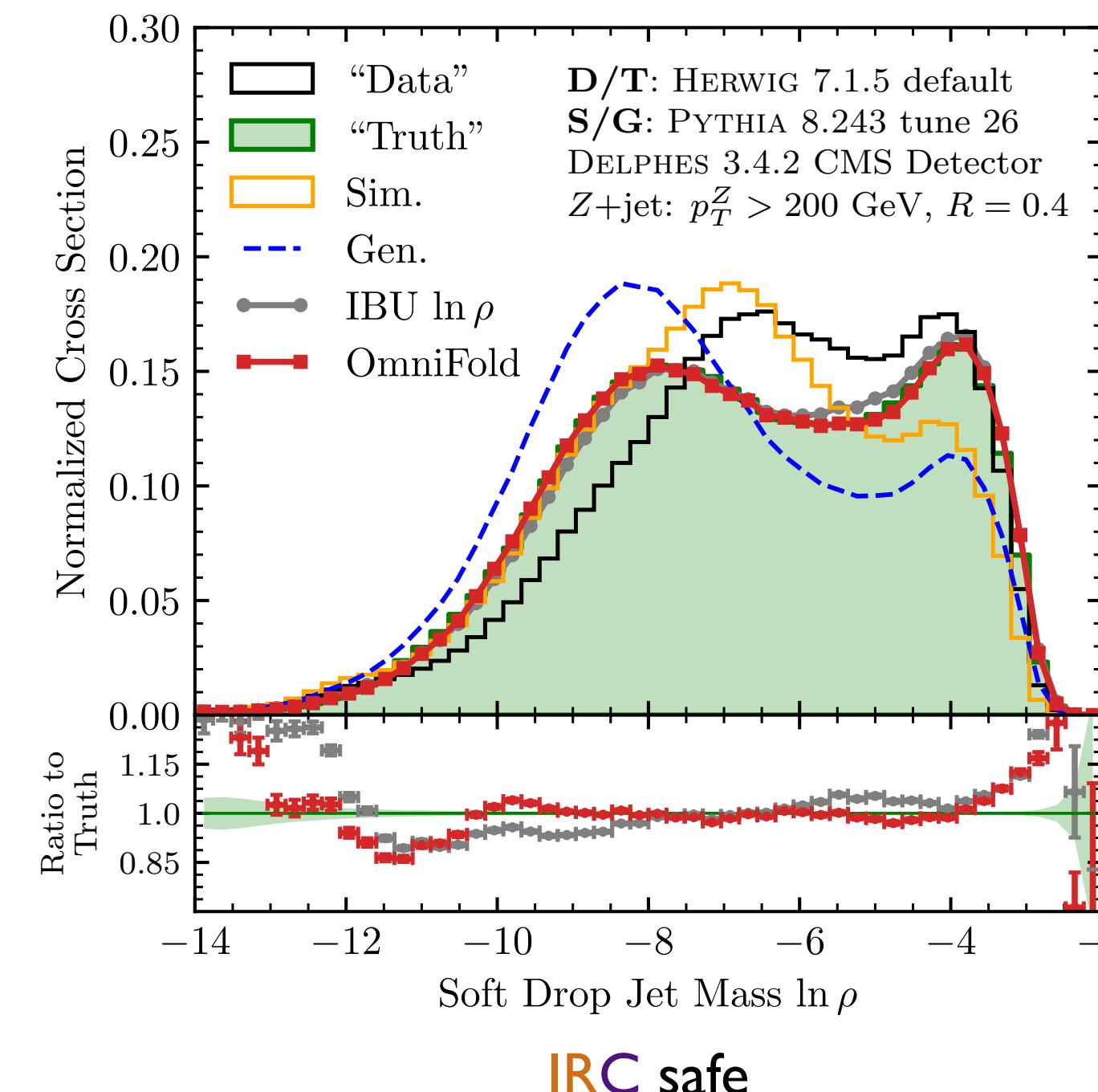
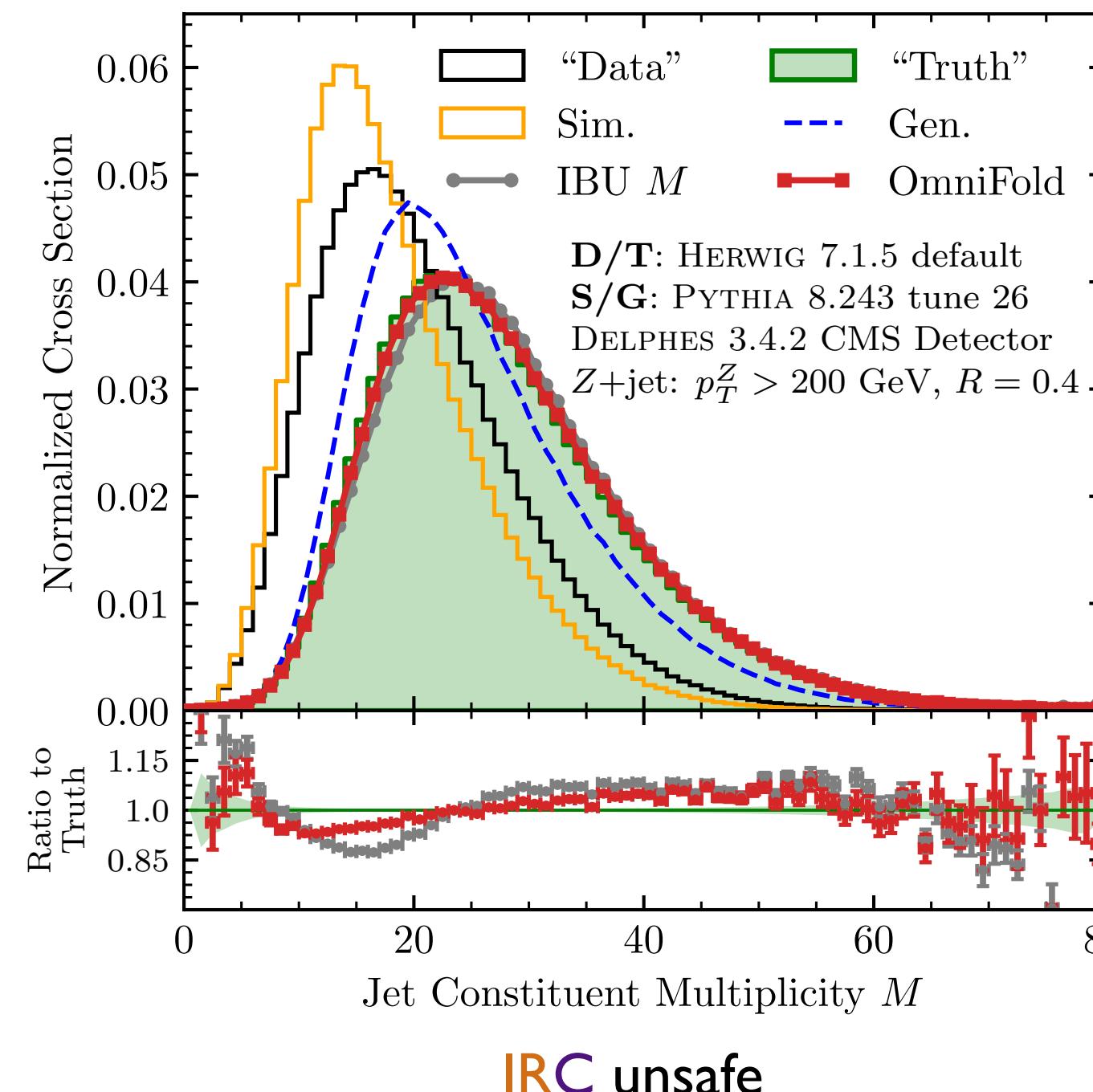
Particle Flow Network (PFN) architecture
processes **full radiation pattern of the event**

- PFN-Ex: $(p_T, y, \phi, \text{PID})$ input features
- Φ : (100, 100, 256) dense layers
- F : (100, 100, 100) dense layers
- ReLU activations, softmax output
- Categorical cross-entropy loss
- 20% validation sample
- 10 epoch patience

OmniFolding Jet Substructure Observables



Single **OmniFold** instantiation vs. repeated applications of IBU



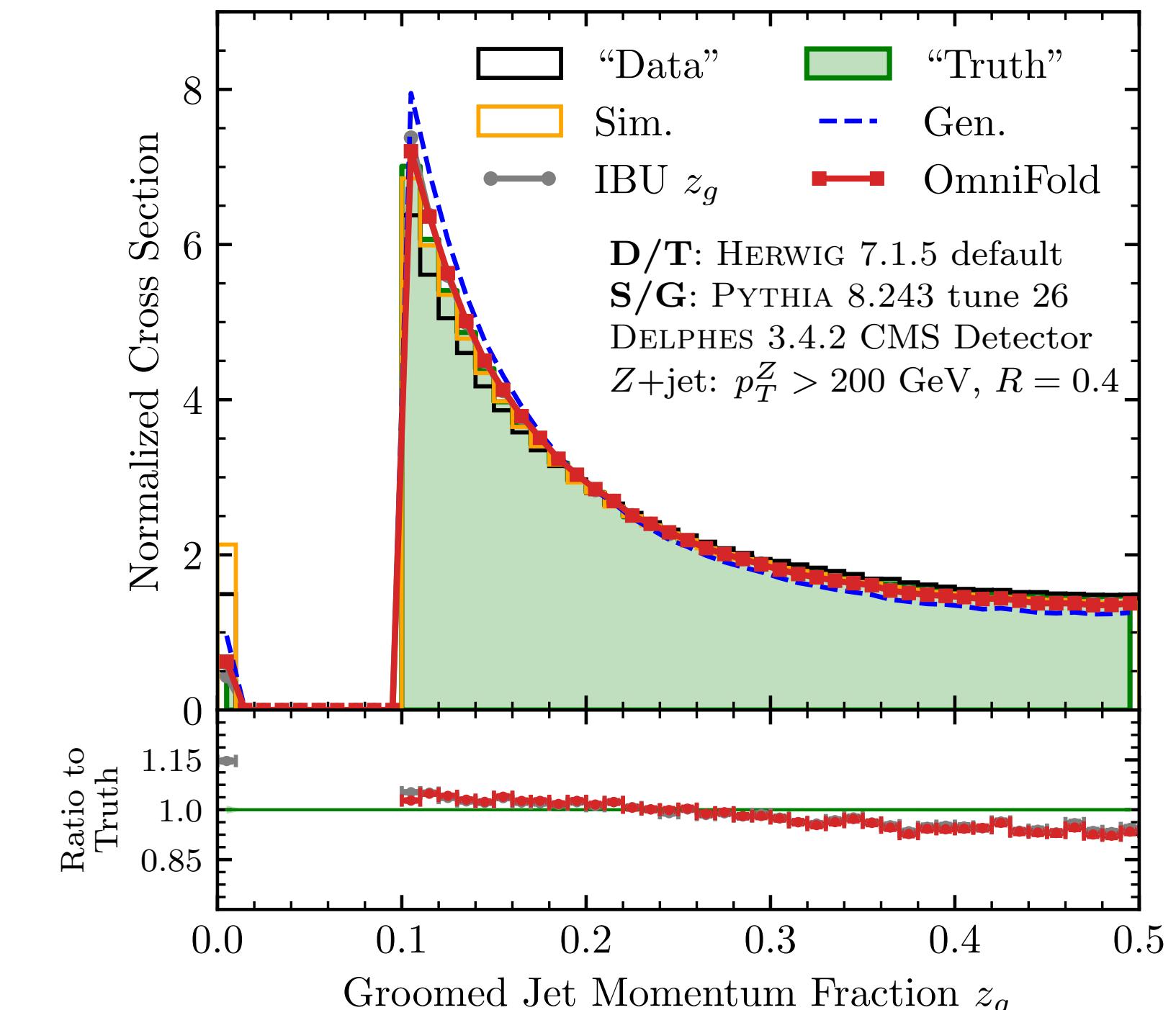
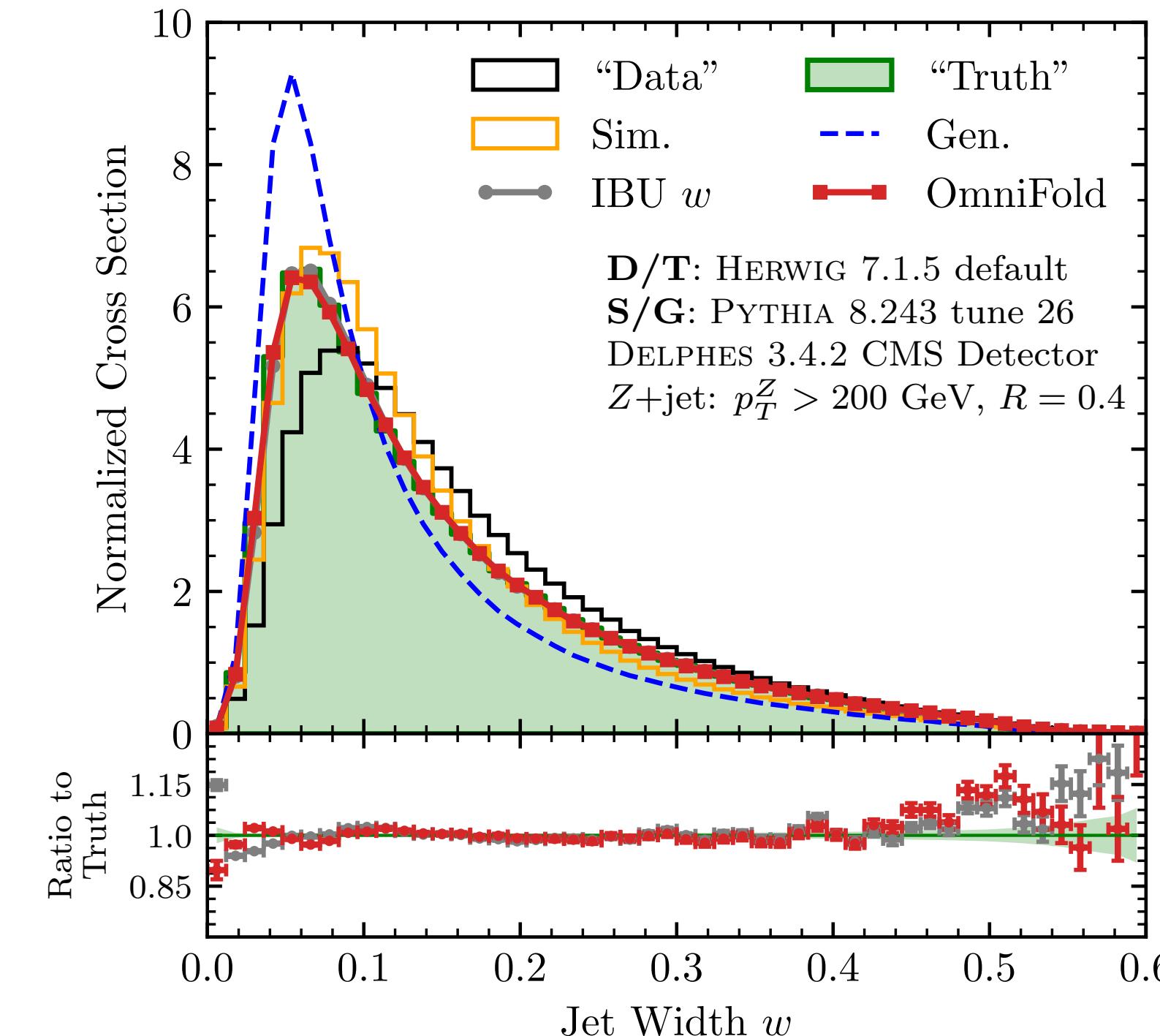
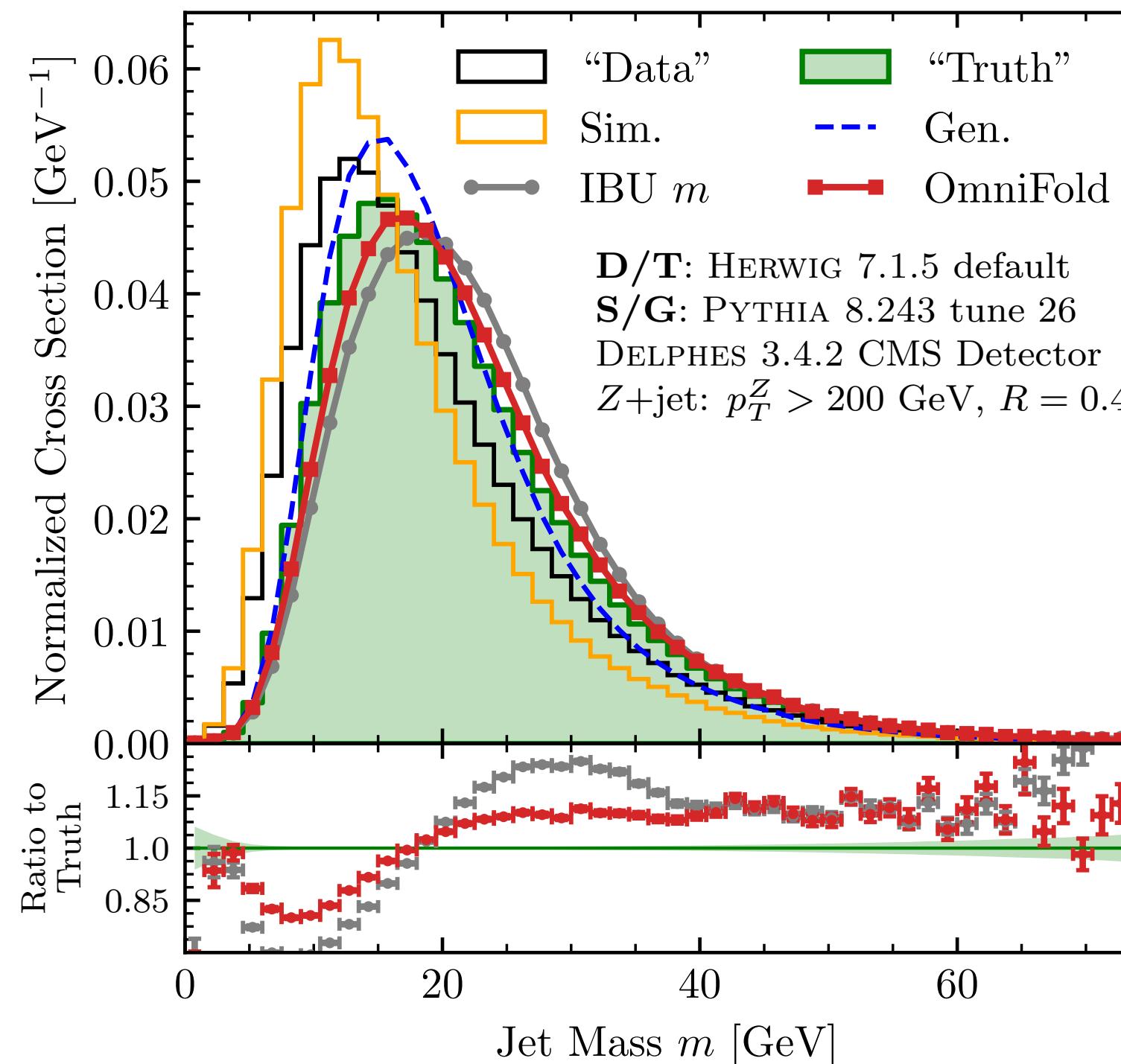
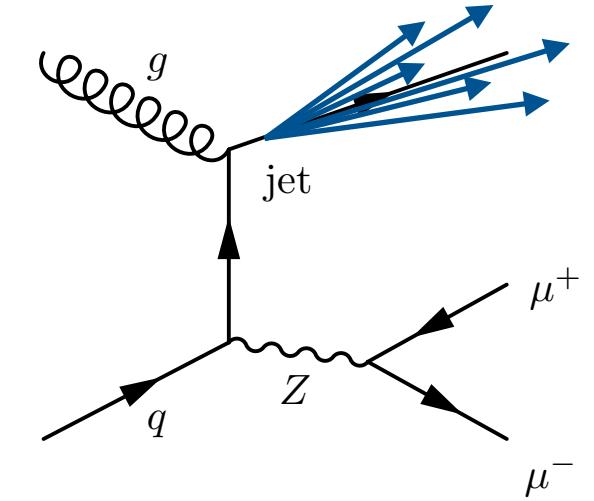
OmniFold equals or outperforms IBU

Five unfolding iterations in all cases

Statistical uncertainties on prior shown in ratio

(See [backup](#) for more on soft drop)

Additional OmniFolded Distributions



Jet mass affected by particle masses

$$m_J^2 = \left(\sum_{i \in \text{jet}} p_i^\mu \right)^2$$

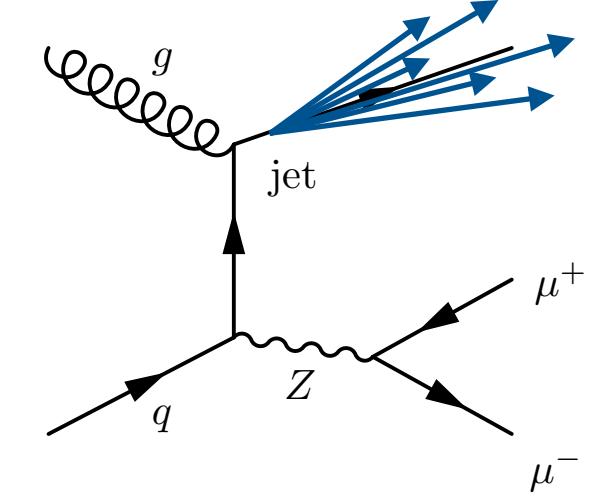
IRC-safe observables easier to unfold

$$w = \frac{1}{\sum_j p_{Tj}} \sum_i p_{Ti} \sqrt{(y_i - y_J)^2 + (\phi_i - \phi_J)^2}$$

z_g remarkably stable under choice of method

$z_g = p_T$ fraction of first splitting to pass soft drop

OmniFold Results by Event Representation



User is free to choose *event representation* in the OmniFold procedure

OMNIFOLD – full phase space information



MULTIFOLD – multiple observables



UNIFOLD – single observable, essentially unbinned IBU

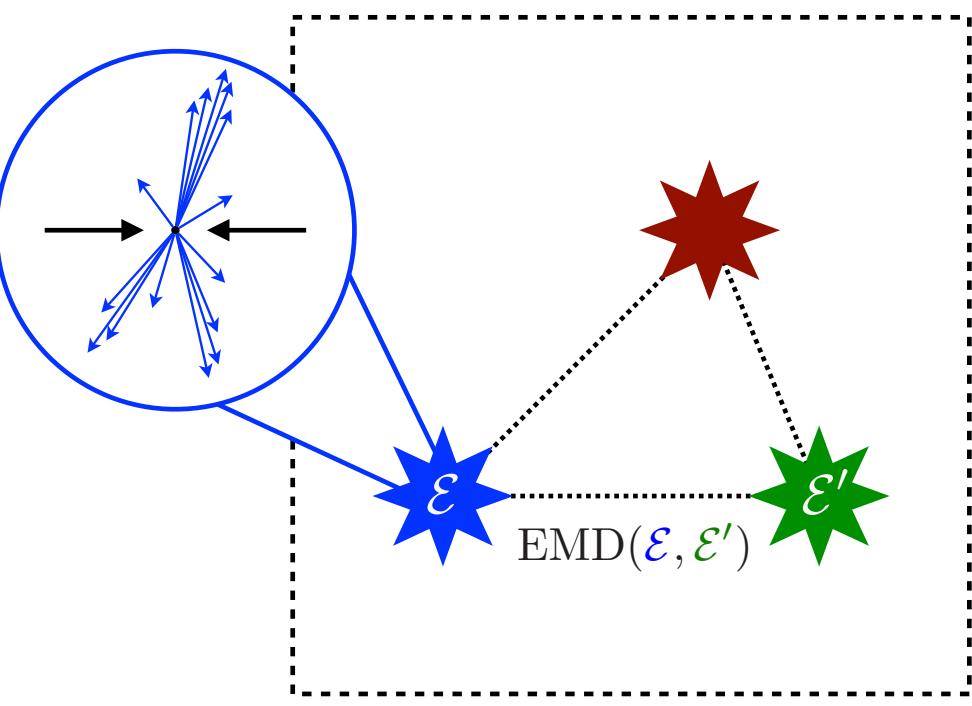
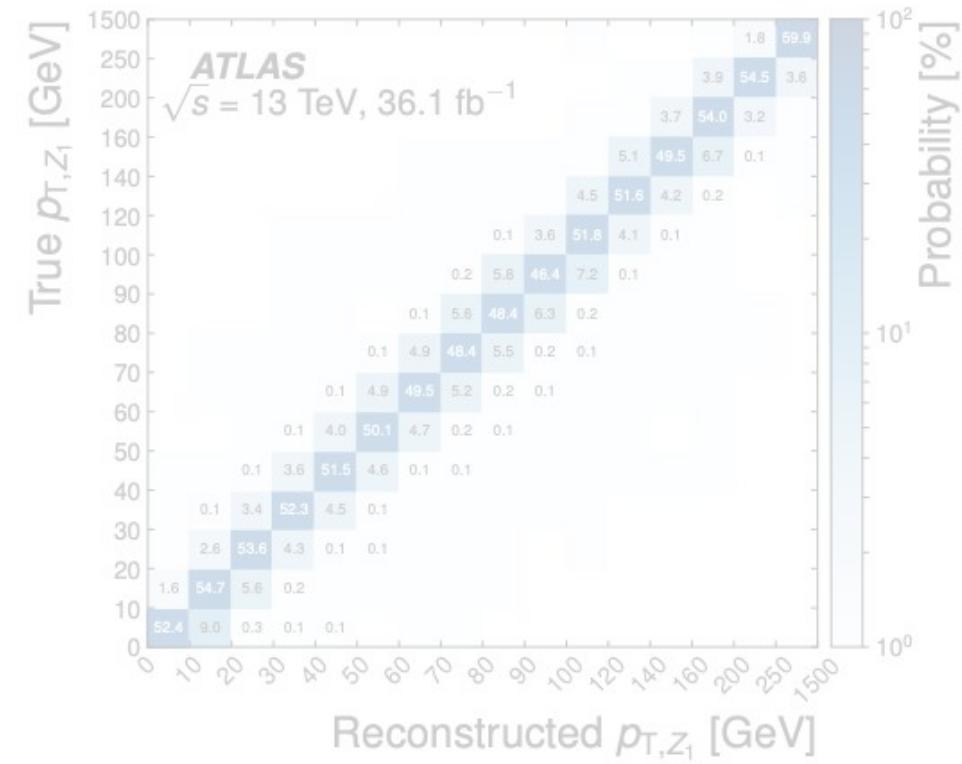
	Observable					
Method	m	M	w	$\ln \rho$	τ_{21}	z_g
OMNIFOLD	2.77	0.33	0.10	0.35	0.53	0.68
MULTIFOLD	3.80	0.89	0.09	0.37	0.26	0.15
UNIFOLD	8.82	1.46	0.15	0.59	1.11	0.59
IBU	9.31	1.51	0.11	0.71	1.10	0.37
Data	24.6	130	15.7	14.2	11.1	3.76
Generation	3.62	15	22.4	19	20.8	3.84

OMNIFOLD/MULTIFOLD *outperforms IBU on all observables!*

$$\Delta(p, q) = \frac{1}{2} \int d\lambda \frac{(p(\lambda) - q(\lambda))^2}{p(\lambda) + q(\lambda)} (\times 10^3)$$

Single **MULTIFOLD** training based on all six observables

UNIFOLD is similar to or
outperforms IBU



Unfolding Setup

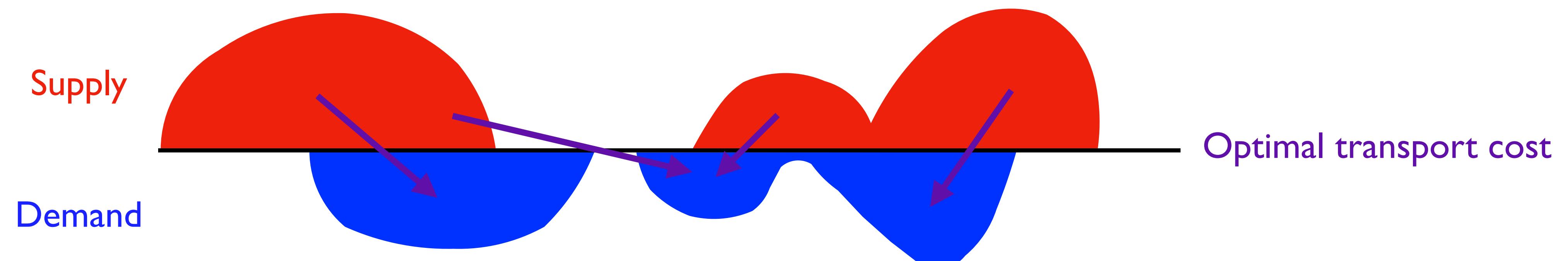
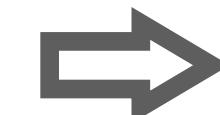
OmniFold

Unfolding Beyond Observables

Optimal Transport in Particle Physics

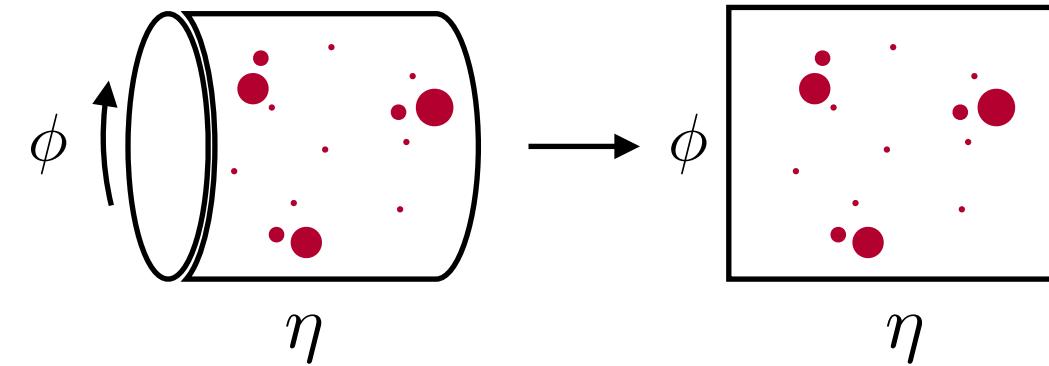
When are Two Distributions Similar?

Optimal transport minimizes the “**work**” (**stuff** x distance) required to transport **supply** to **demand**



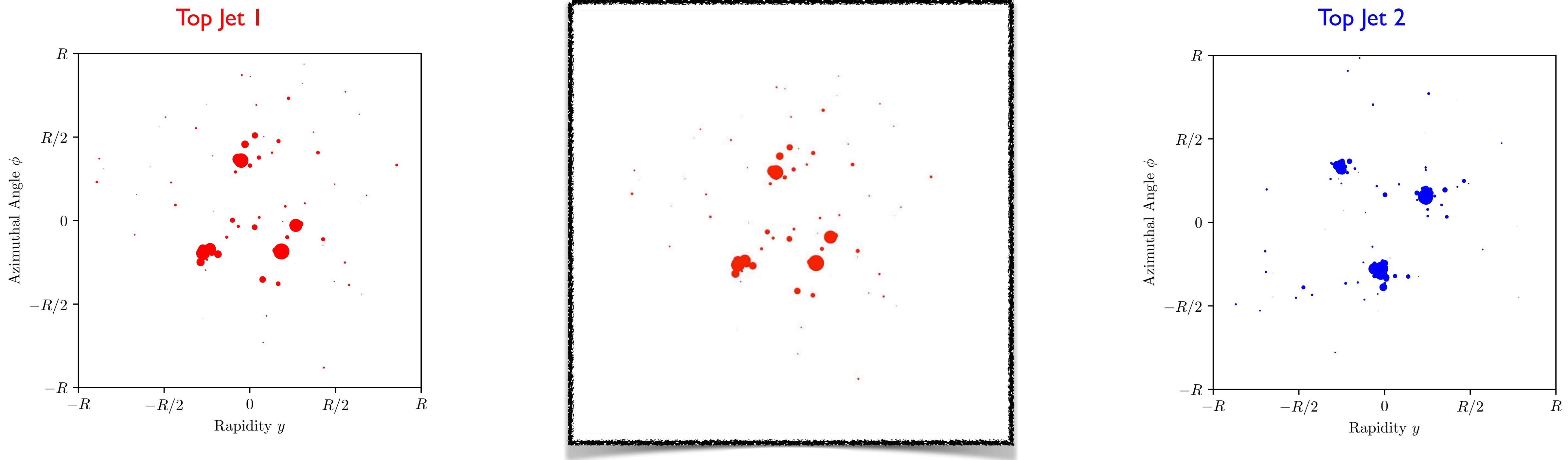
[Monge, 1781; Vaserštejn, 1969; Peleg, Werman, Rom, [IEEE 1989](#); Rubner, Tomasi, Guibas, [ICCV 1998](#), [ICCV 2000](#); Pele, Werman, [ECCV 2008](#); Pele, Taskar, [GSI 2013](#)]

When are Two Events similar?



[PTK, Metodiev, Thaler, PRL 2019]

Optimal transport minimizes the “**work**” (**stuff** x distance) required to transport **supply** to **demand**



$$\mathcal{E}(\hat{n}) = \sum_{i=1}^M E_i \delta(\hat{n} - \hat{n}_i)$$

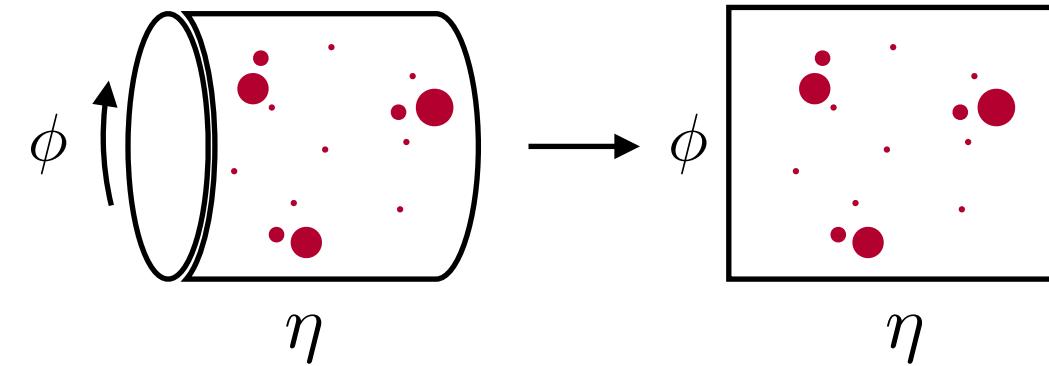
Provides a **metric** on **normalized distributions** in a space with a ground distance measure

↳ symmetric, non-negative, triangle inequality, zero iff identical

$$\theta_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta \phi_{ij}^2}$$

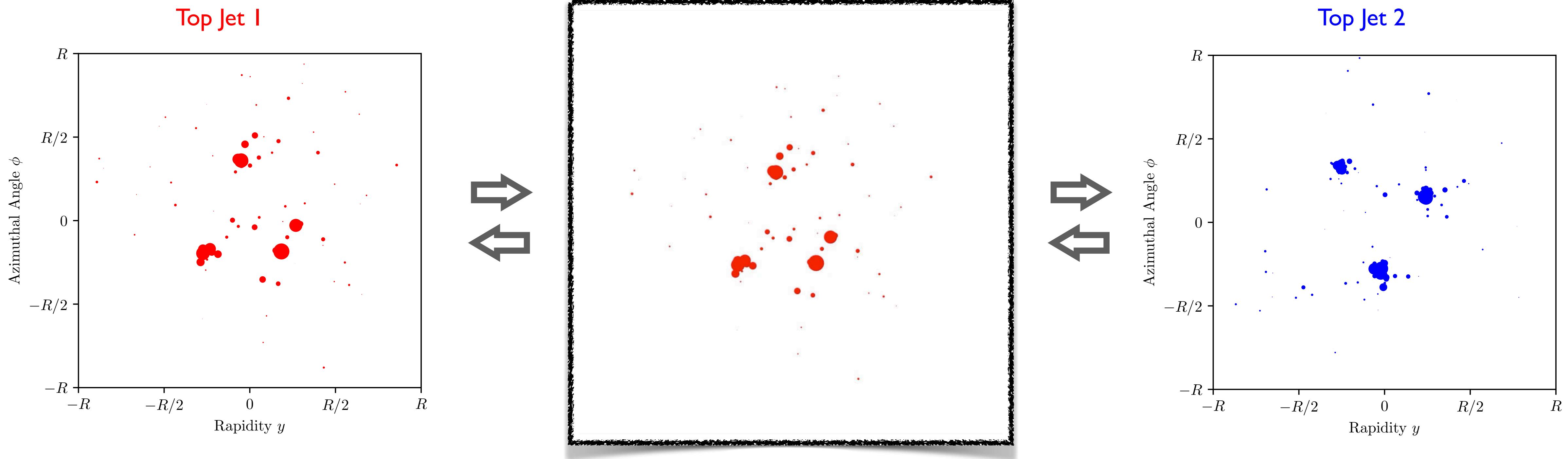
[Peleg, Werman, Rom, IEEE 1989; Rubner, Tomasi, Guibas, ICCV 1998, ICJV 2000; Pele, Werman, ECCV 2008; Pele, Taskar, GSI 2013]

When are Two Events similar?



[PTK, Metodiev, Thaler, PRL 2019]

Optimal transport minimizes the “**work**” (**stuff** x distance) required to transport **supply** to **demand**



$$\mathcal{E}(\hat{n}) = \sum_{i=1}^M E_i \delta(\hat{n} - \hat{n}_i)$$

Provides a **metric** on **normalized distributions** in a space with a ground distance measure

↳ symmetric, non-negative, triangle inequality, zero iff identical

$$\theta_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta \phi_{ij}^2}$$

[Peleg, Werman, Rom, IEEE 1989; Rubner, Tomasi, Guibas, ICCV 1998, ICJV 2000; Pele, Werman, ECCV 2008; Pele, Taskar, GSI 2013]

The Energy Mover's Distance (EMD)

[PTK, Metodiev, Thaler, PRL 2019]

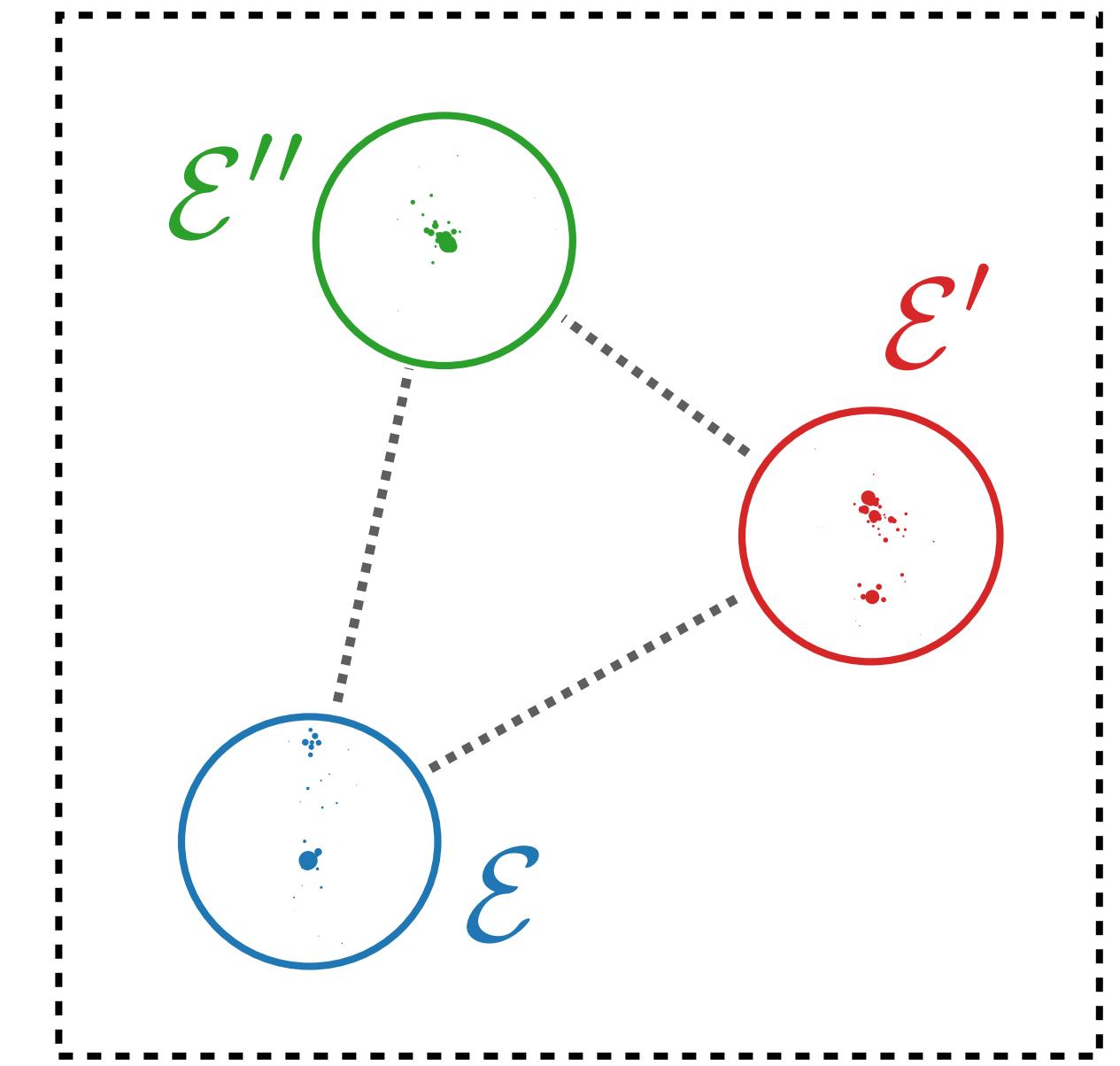
EMD between energy flows defines a metric on the space of events

$$\text{EMD}_{\beta,R}(\mathcal{E}, \mathcal{E}') = \min_{\{f_{ij} \geq 0\}} \sum_i \sum_j f_{ij} \left(\frac{\theta_{ij}}{R} \right)^\beta + \left| \sum_i E_i - \sum_j E'_j \right|$$

Cost of optimal transport Cost of energy creation

$$\sum_j f_{ij} \leq E_i, \quad \sum_i f_{ij} \leq E'_j, \quad \sum_{ij} f_{ij} = \min \left(\sum_i E_i, \sum_j E'_j \right)$$

Capacity constraints to ensure proper transport



R : controls cost of transporting energy vs. destroying/creating it

β : angular weighting exponent

Triangle inequality satisfied for $R \geq d_{\max}/2$

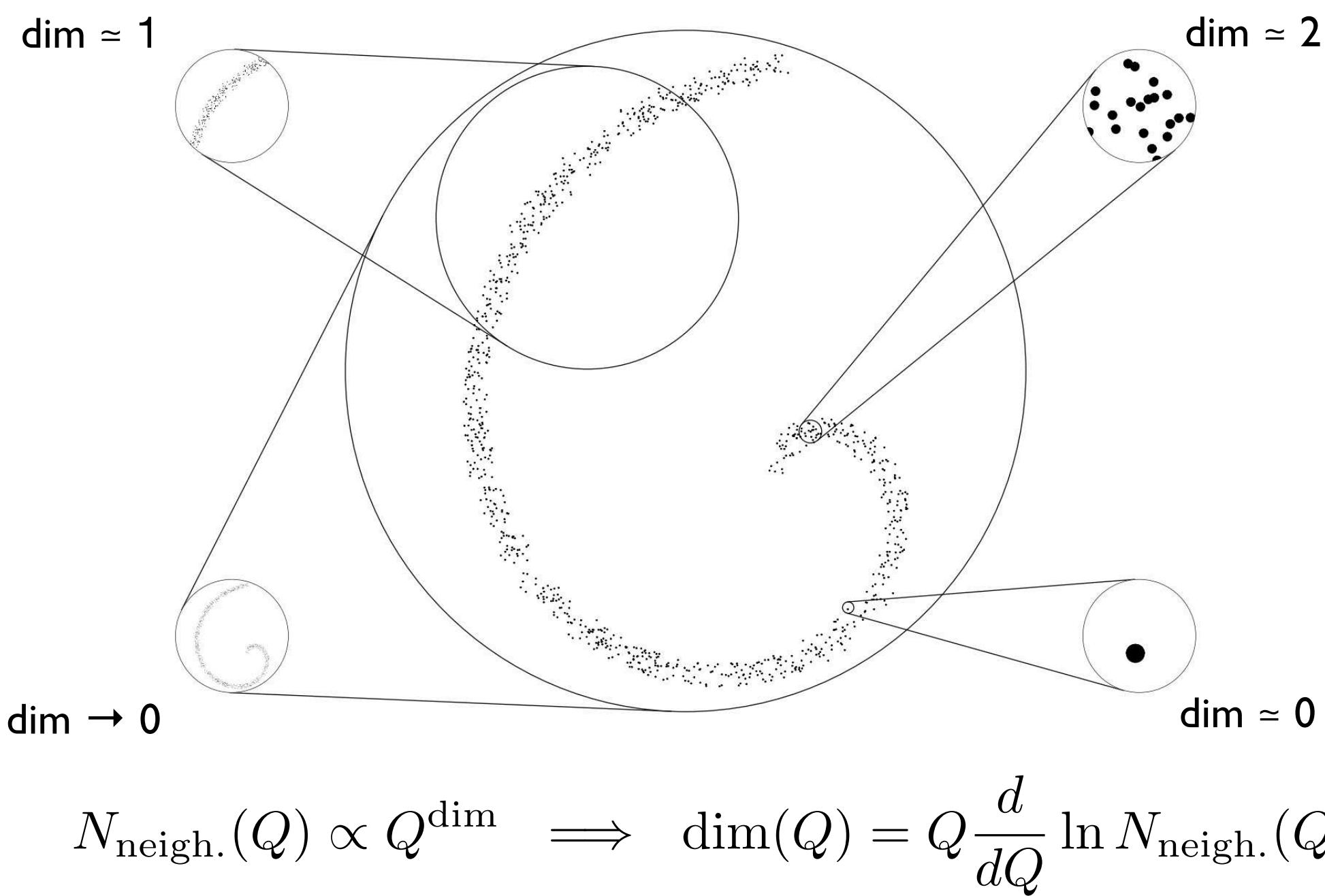
$$0 \leq \text{EMD}(\mathcal{E}, \mathcal{E}') \leq \text{EMD}(\mathcal{E}, \mathcal{E}'') + \text{EMD}(\mathcal{E}'', \mathcal{E}')$$

i.e. $R \geq$ jet radius for conical jets

Quantifying Event-Space Manifolds

[PTK, Metodiev, Thaler, PRL 2019]

Correlation dimension: how does the # of elements within a ball of size Q change?



Correlation dimension lessons:

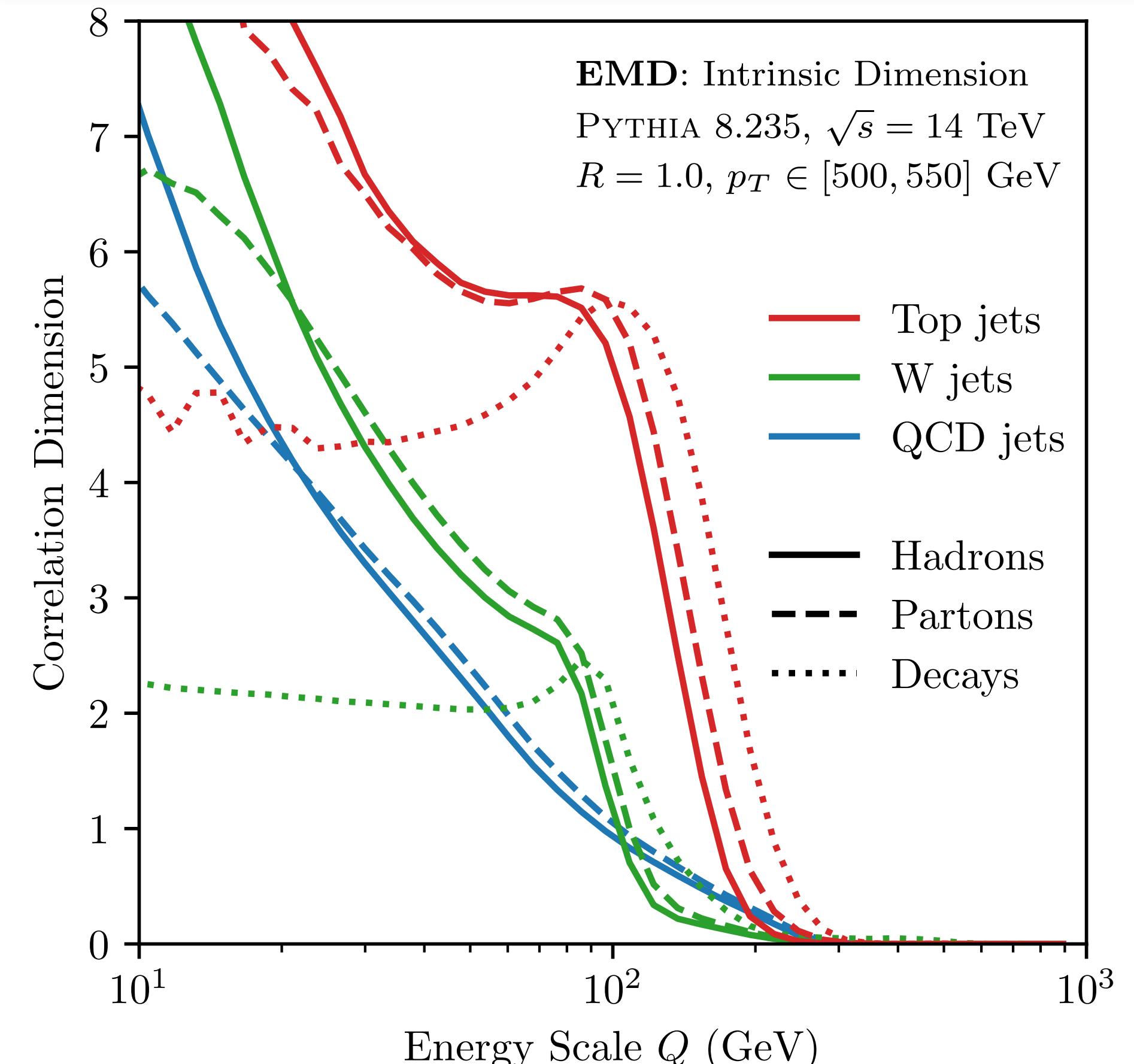
Decays are "constant" dim. at low Q

Complexity hierarchy: QCD < W < Top

Fragmentation increases dim. at smaller scales

Hadronization important around 20-30 GeV

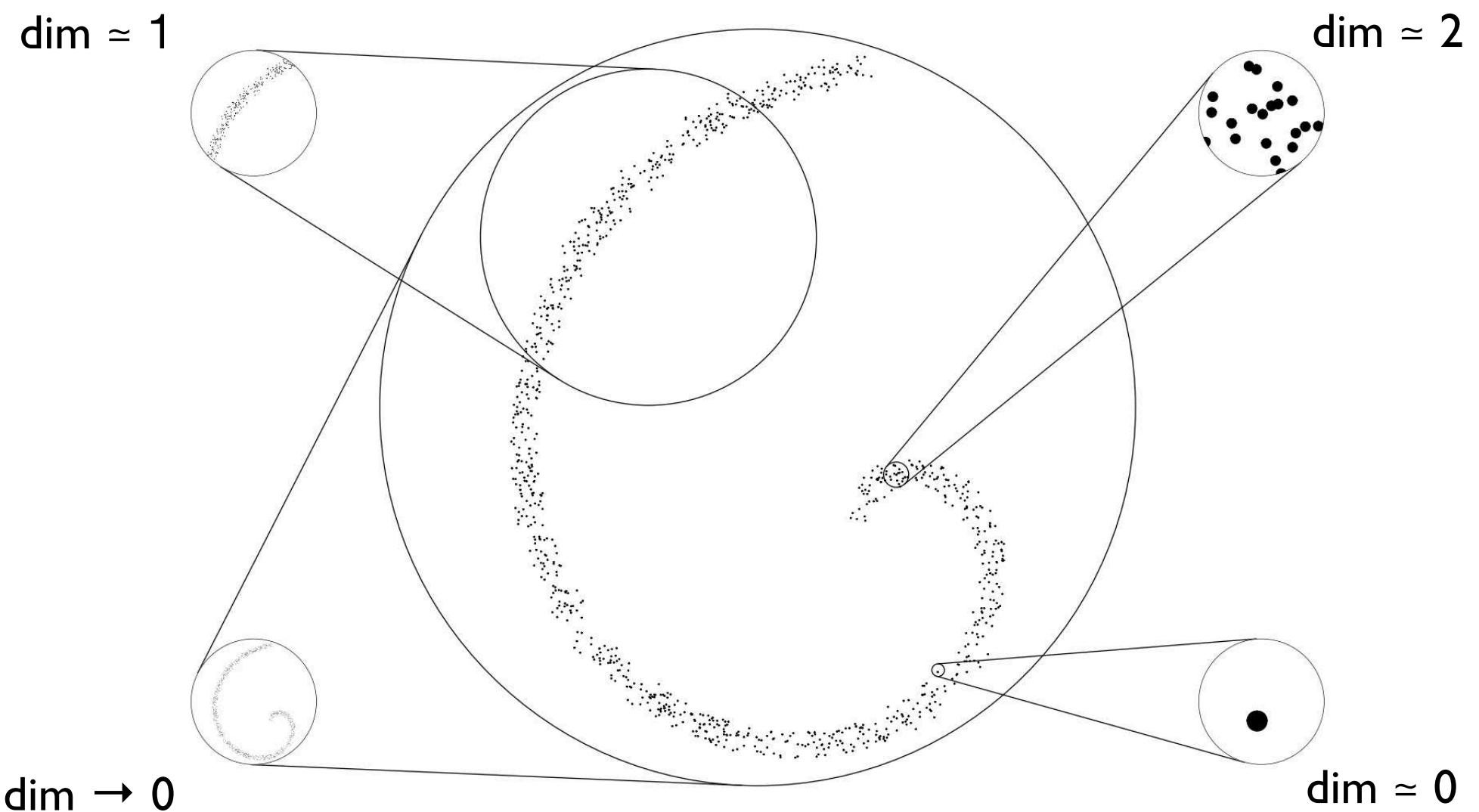
$$\text{dim}(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$



[Grassberger, Procaccia, PRL 1983]

Unfolding Beyond Observables

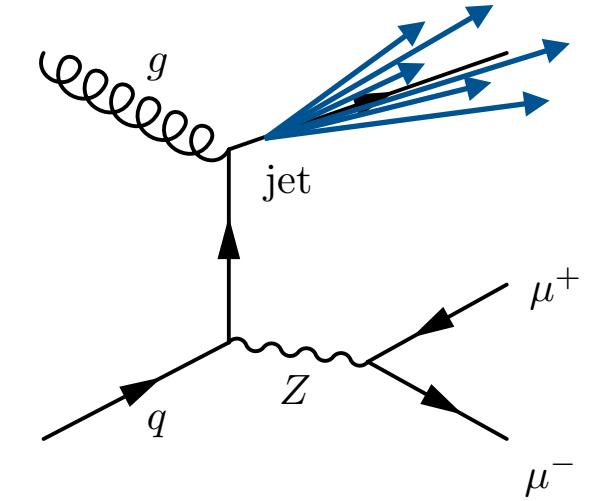
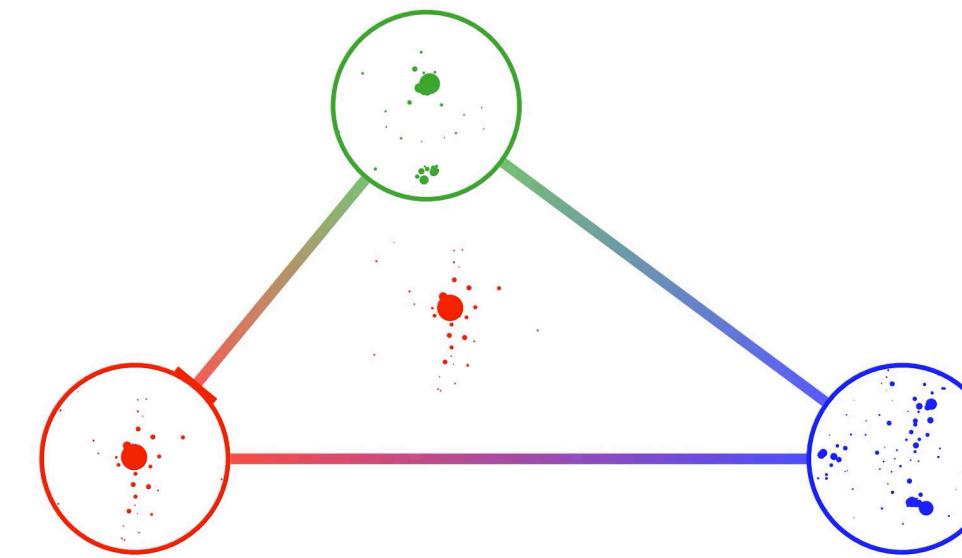
Correlation dimension: how does the # of elements within a ball of size Q change?



$$N_{\text{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q) = Q \frac{d}{dQ} \ln N_{\text{neigh.}}(Q)$$

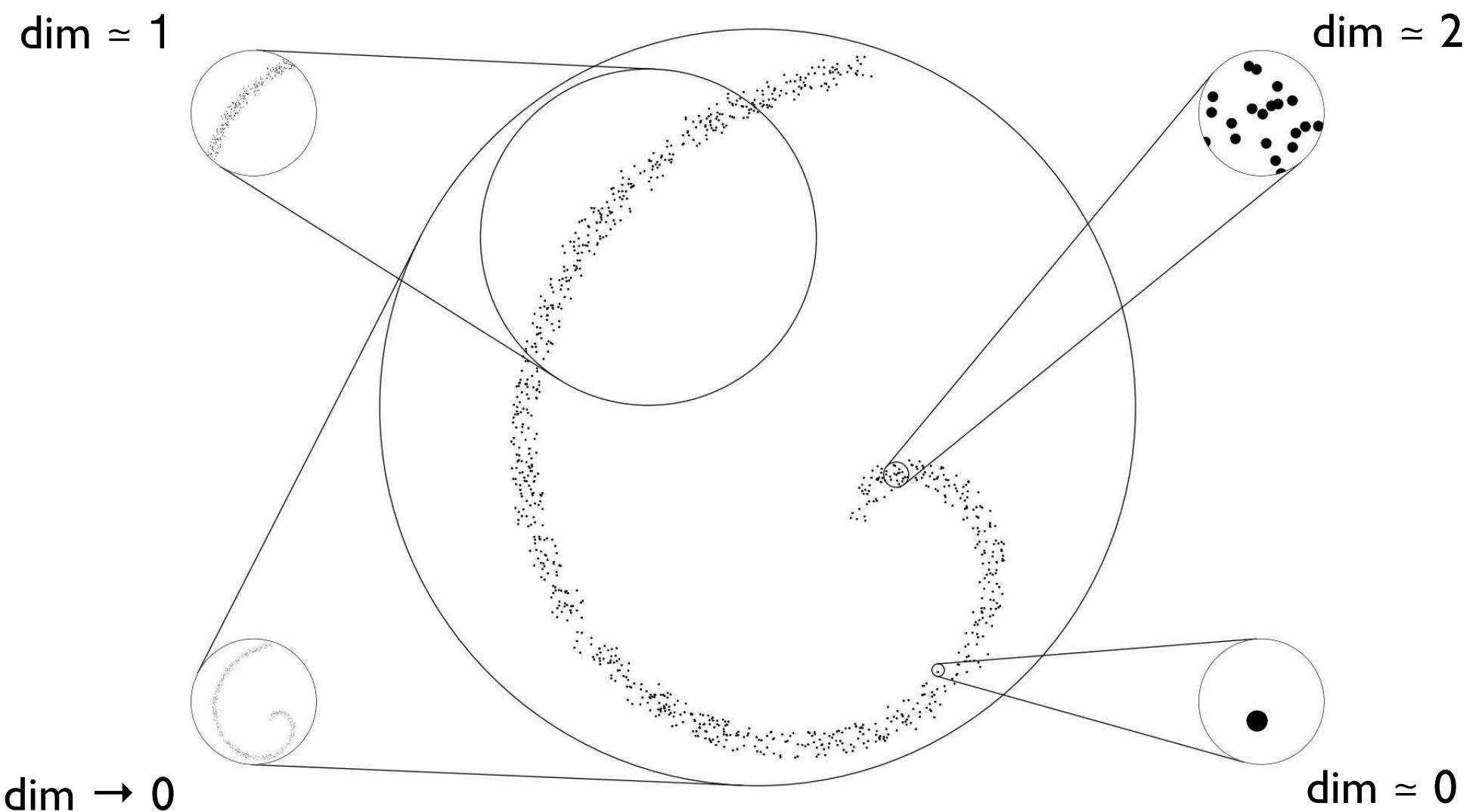
$$\dim(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \boxed{w_i w'_j} \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$

Weighted events naturally accommodated



Unfolding Beyond Observables

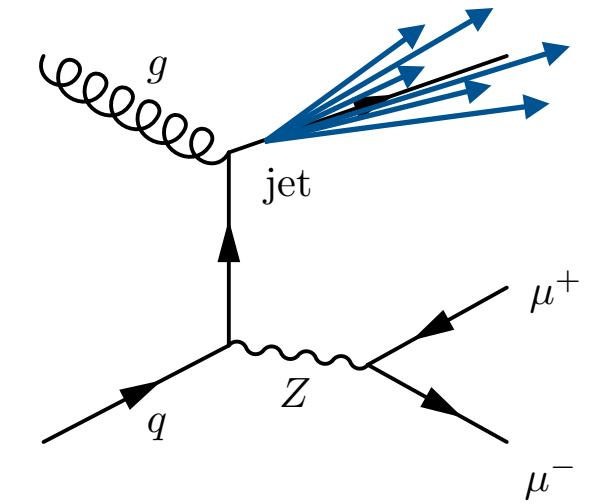
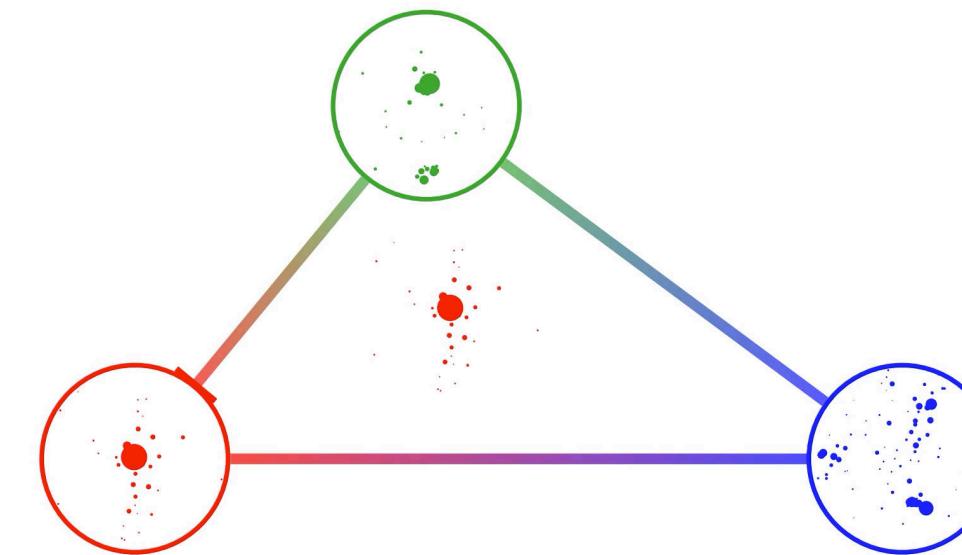
Correlation dimension: how does the # of elements within a ball of size Q change?



$$N_{\text{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q) = Q \frac{d}{dQ} \ln N_{\text{neigh.}}(Q)$$

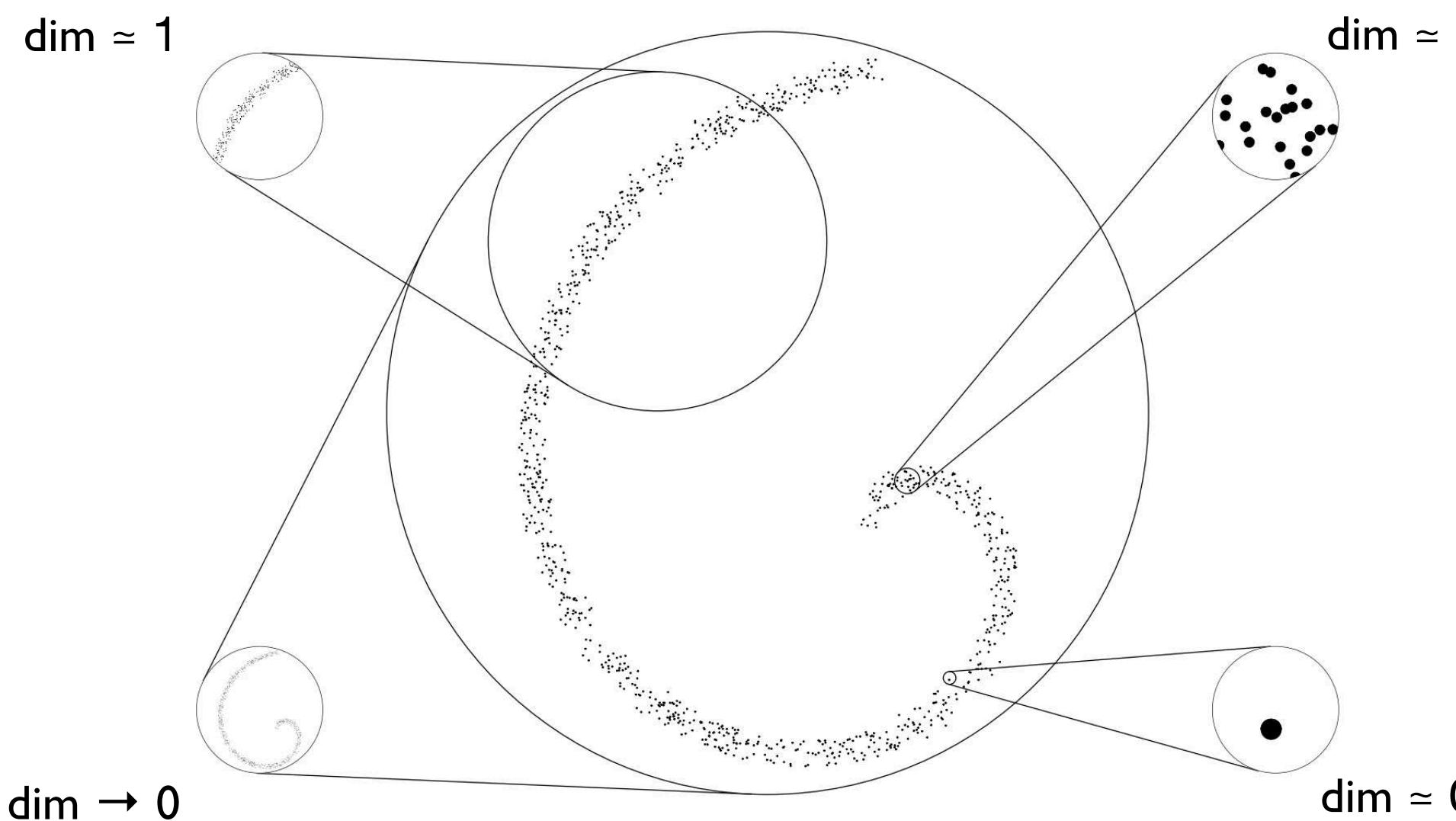
$$\dim(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \boxed{w_i w'_j} \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$

Weighted events naturally accommodated



Unfolding Beyond Observables

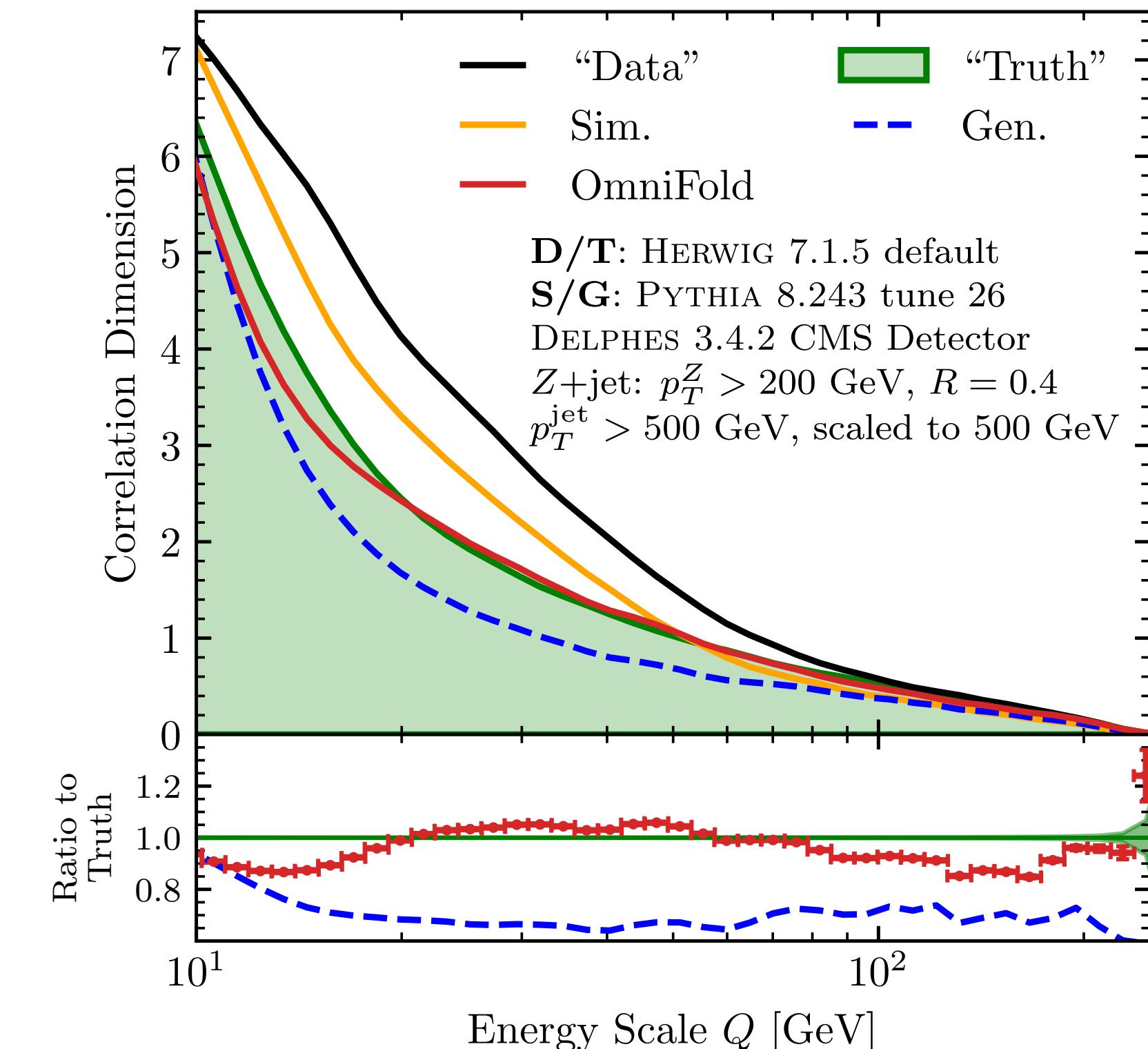
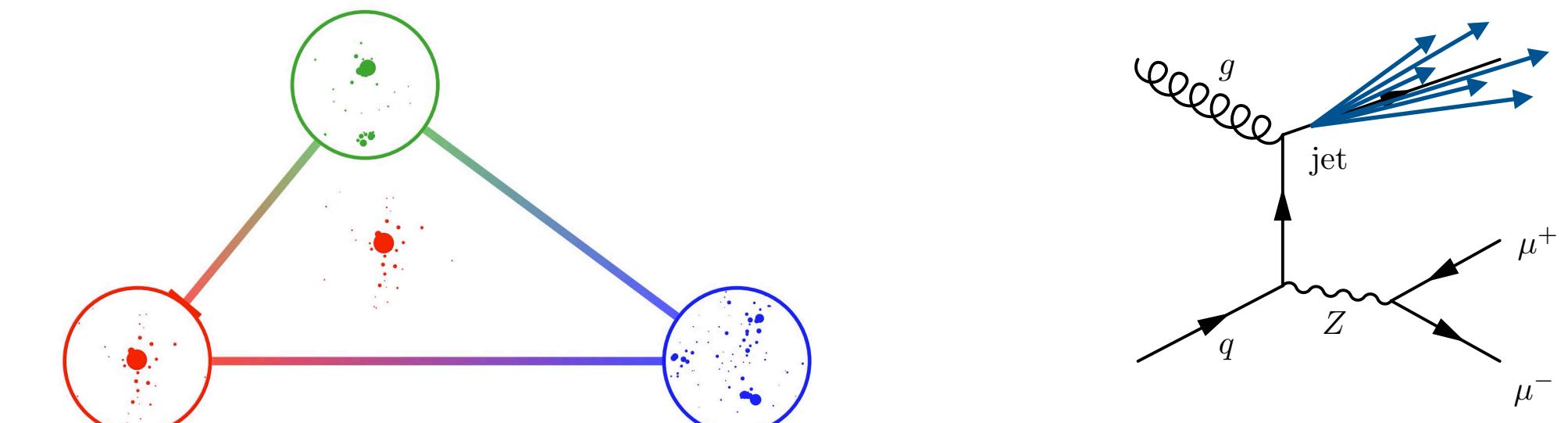
Correlation dimension: how does the # of elements within a ball of size Q change?



$$N_{\text{neigh.}}(Q) \propto Q^{\text{dim}} \implies \text{dim}(Q) = Q \frac{d}{dQ} \ln N_{\text{neigh.}}(Q)$$

$$\text{dim}(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \boxed{w_i w'_j} \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$

Weighted events naturally accommodated



Same **OmniFold** training can unfold a complicated function of pairs of events!

Larger detector effects and loss of stats seen at low Q

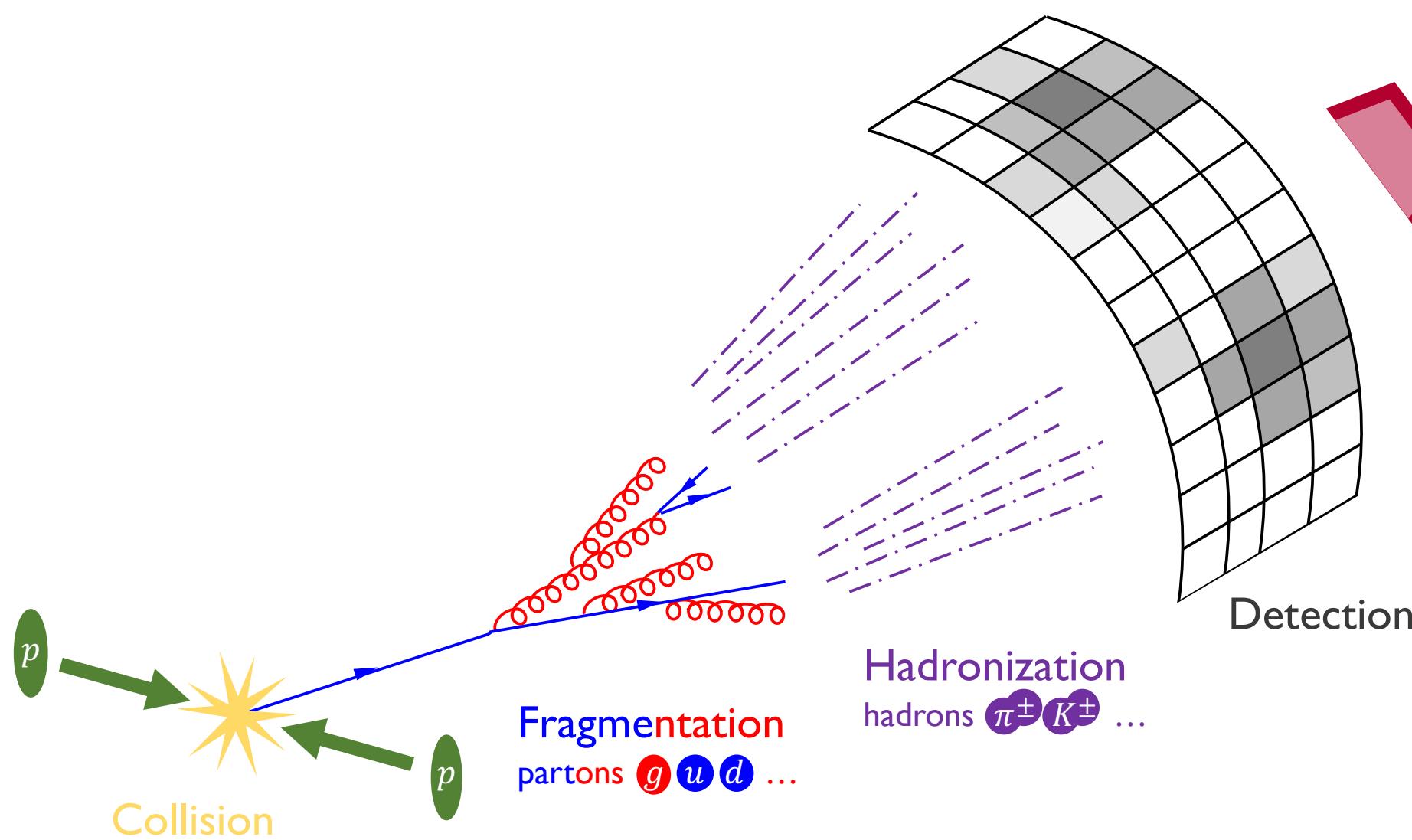
Beyond Observables via Weighted Cross Sections

Standard observables

Calculates a single number for each jet/event
and study the distribution of values

Weighted cross section

Calculate a distributional quantity per event
and study the mean distribution



$$\mathcal{E}(\hat{n}) = \int_0^{\infty} dt \lim_{r \rightarrow \infty} r^2 n^i T_{0i}(t, r\hat{n})$$

Stress-energy flow

Beyond Observables via Weighted Cross Sections

Standard observables

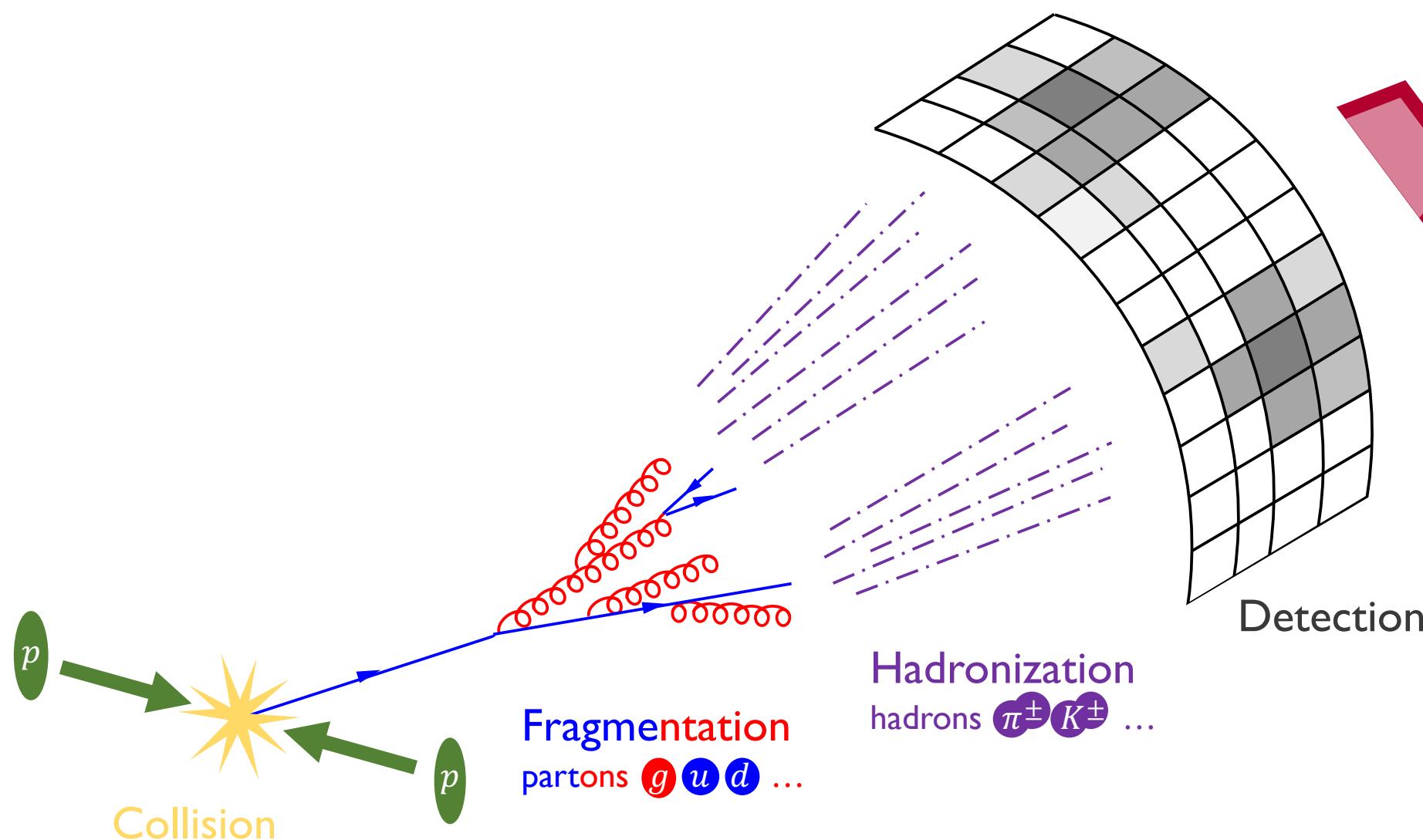
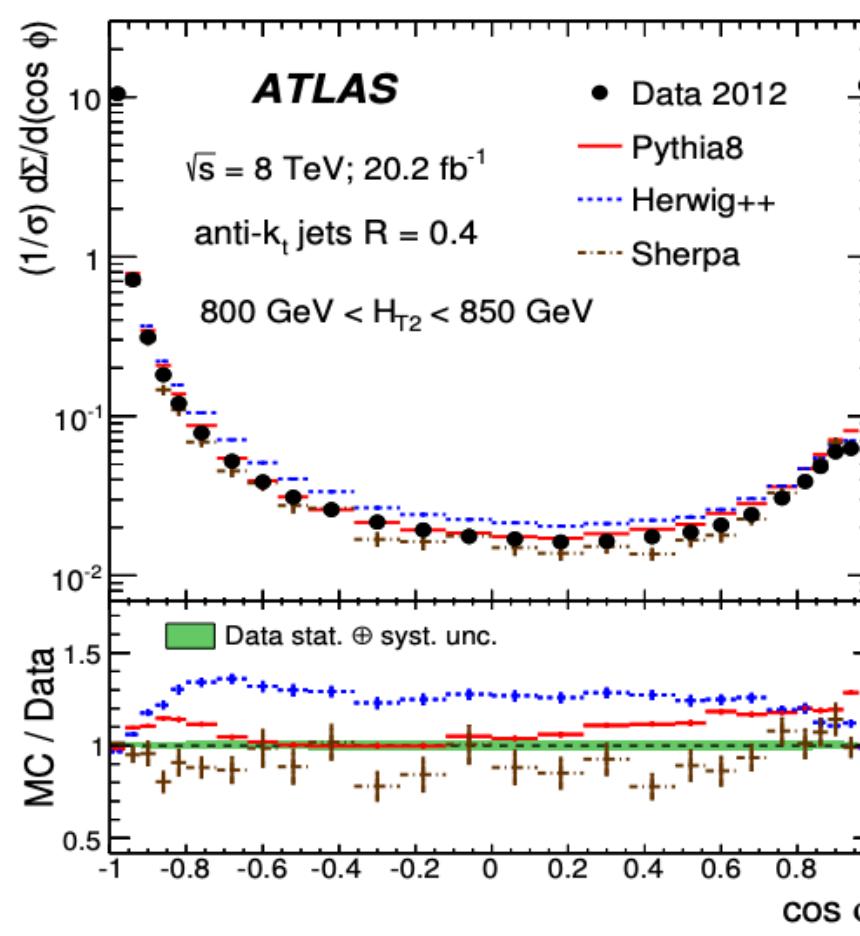
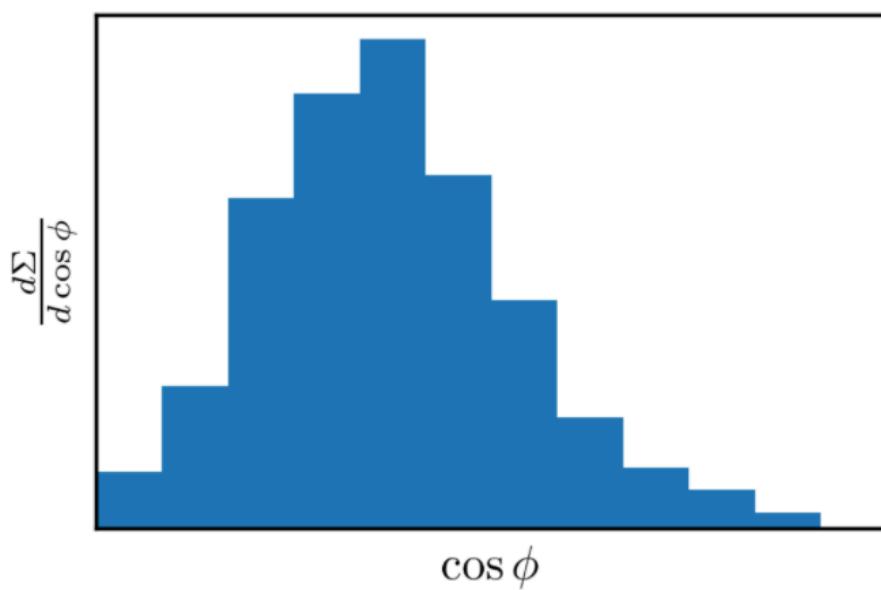
Calculates a single number for each jet/event
and study the distribution of values

Weighted cross section

Calculate a distributional quantity per event
and study the mean distribution

e.g. energy-energy correlator (EEC)

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



$$\mathcal{E}(\hat{n}) = \int_0^\infty dt \lim_{r \rightarrow \infty} r^2 n^i T_{0i}(t, r\hat{n})$$

Stress-energy flow

Beyond Observables via Weighted Cross Sections

Standard observables

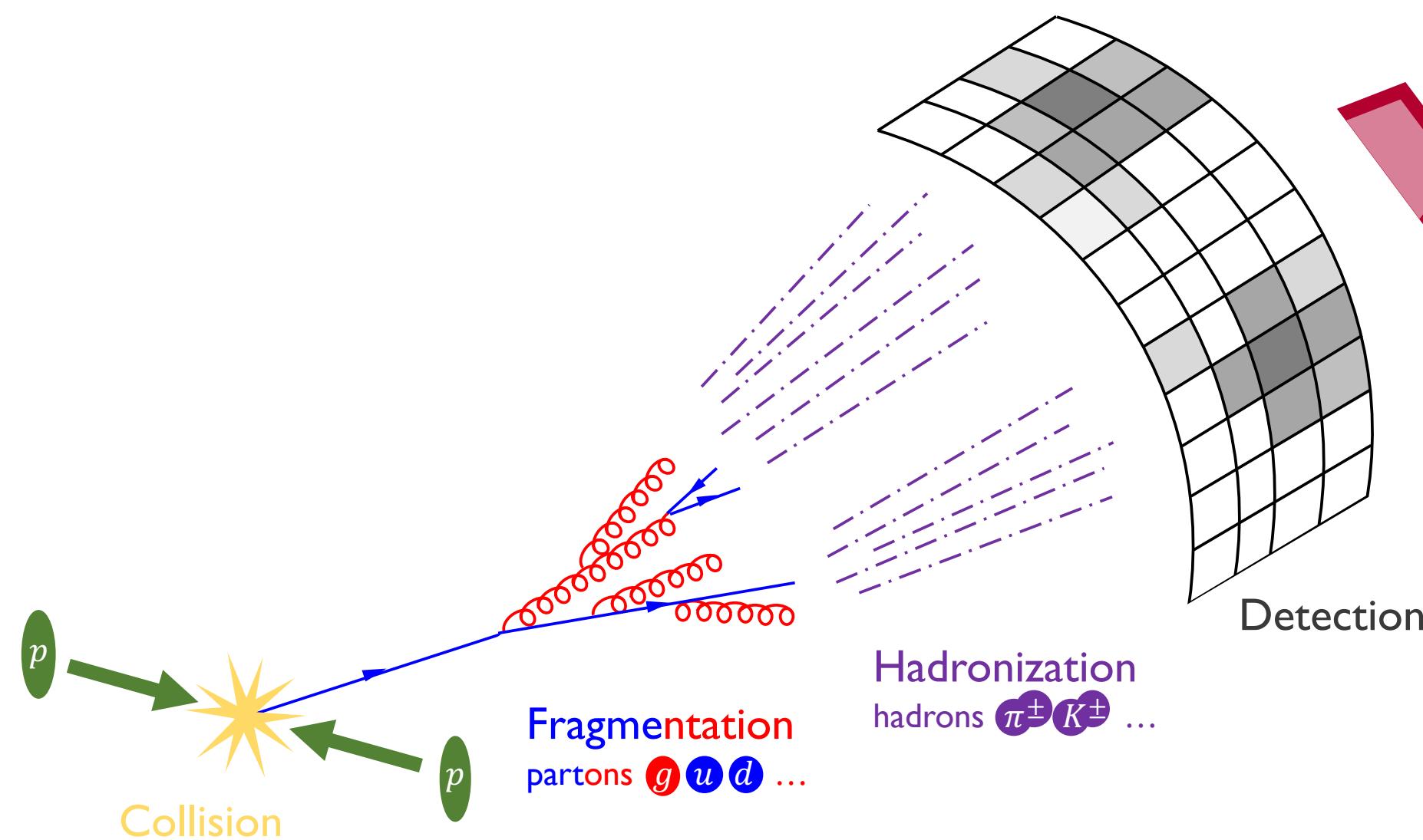
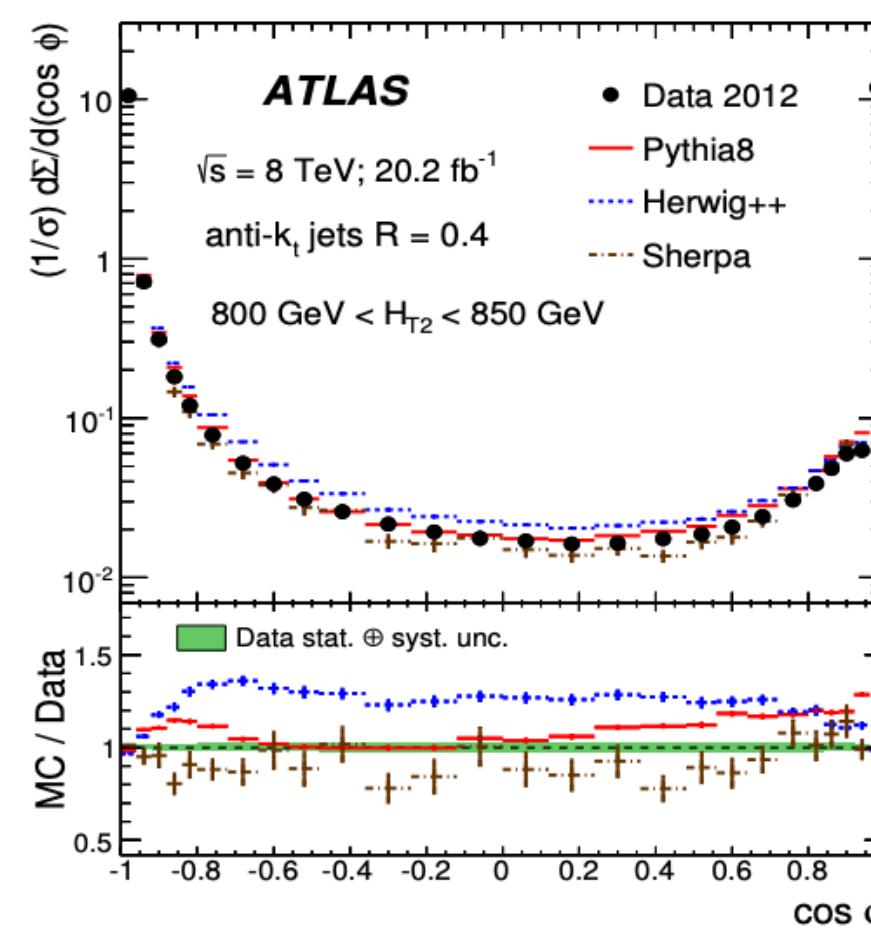
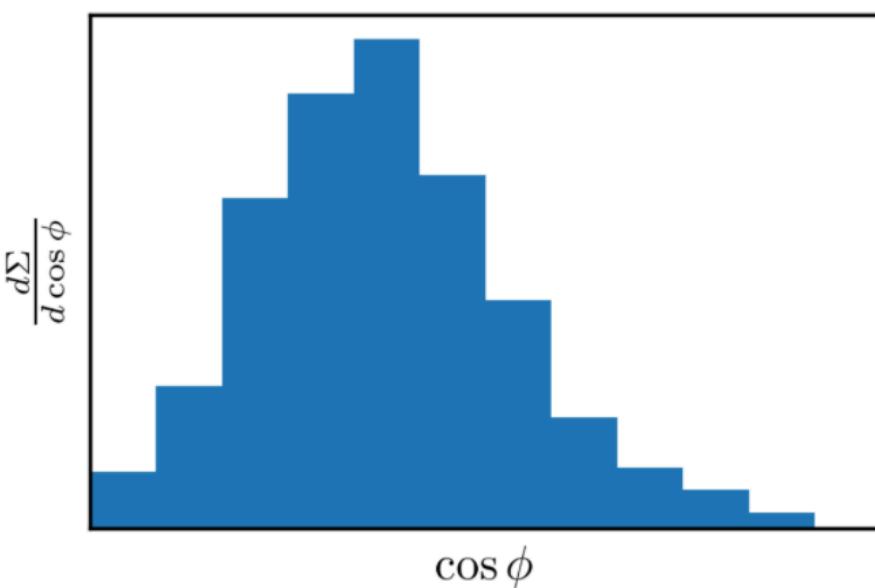
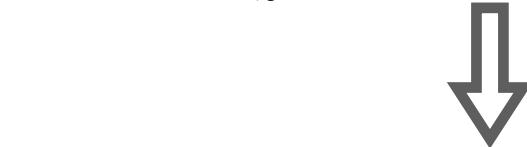
Calculates a single number for each jet/event
and study the distribution of values

Weighted cross section

Calculate a distributional quantity per event
and study the mean distribution

e.g. energy-energy correlator (EEC)

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

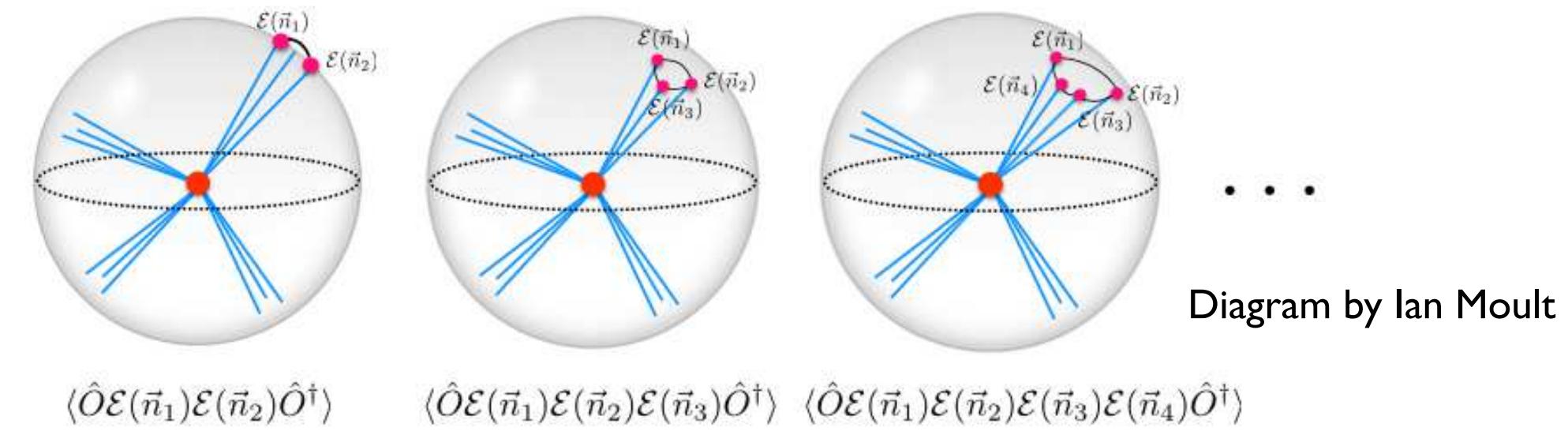


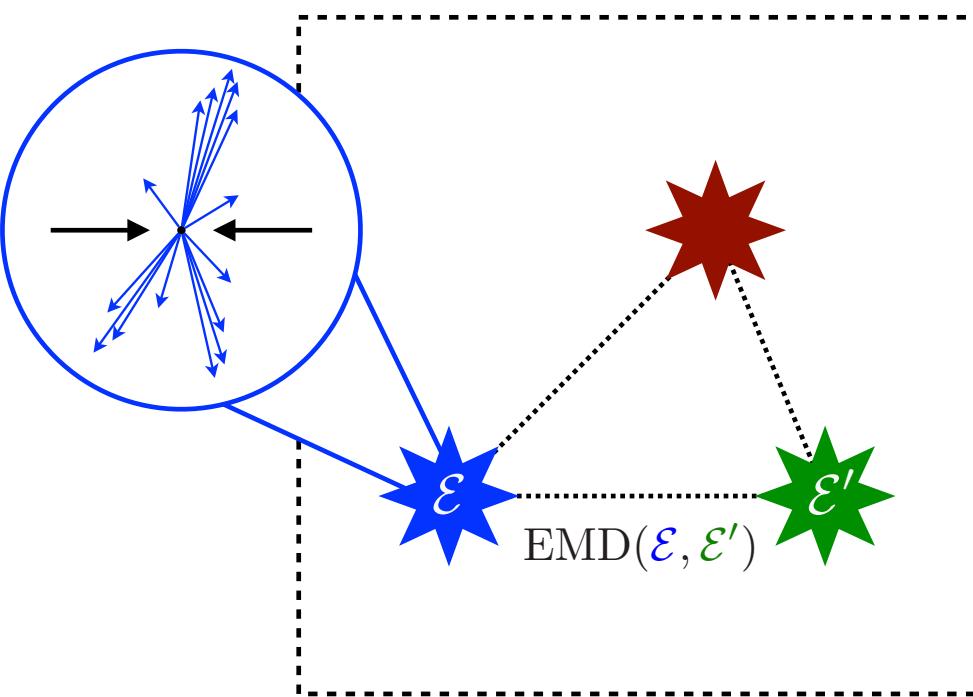
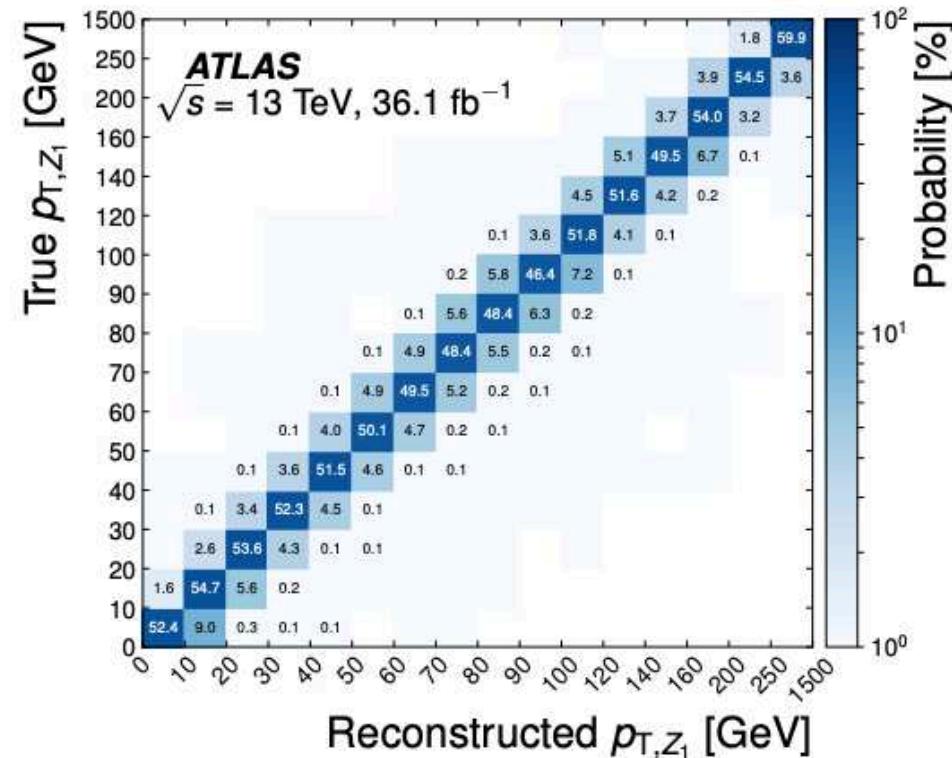
$$\mathcal{E}(\hat{n}) = \int_0^\infty dt \lim_{r \rightarrow \infty} r^2 n^i T_{0i}(t, r\hat{n})$$

Stress-energy flow

$$\frac{1}{\sigma_{\text{tot}}} \frac{d\sigma}{d\hat{n}_1 \cdots d\hat{n}_N} = \frac{\langle \mathcal{O}\mathcal{E}(\hat{n}_1) \cdots \mathcal{E}(\hat{n}_N)\mathcal{O}^\dagger \rangle}{\langle \mathcal{O}\mathcal{O}^\dagger \rangle}$$

Correlations of energy flow operators can be directly studied!





Unfolding Setup

- Unfolding corrects distributions for detector effects using detector simulation in a *prior independent* manner
- Traditional unfolding (e.g. IBU) is based on histograms and limited to a few observables due to the curse of dimensionality

OmniFold

- Maximum likelihood solution to the unfolding problem (like IBU)
- Likelihood-free inference uses high-dim. classifiers to capture the full phase-space
- Learns a single particle-level weighting function that unfolds all observables

Unfolding Beyond Observables

- Non-per-event quantities can be of broad interest in particle physics
- Traditional unfolding is challenged due to reliance on low-dimensional histograms
- Can be unfolded with OmniFold as easily and naturally as any other quantity

OmniFold Etymology

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

The Seasons played around his knees
Like Children round a sire –
Grandfather of the Days is He
Of Dawn, the Ancestor –

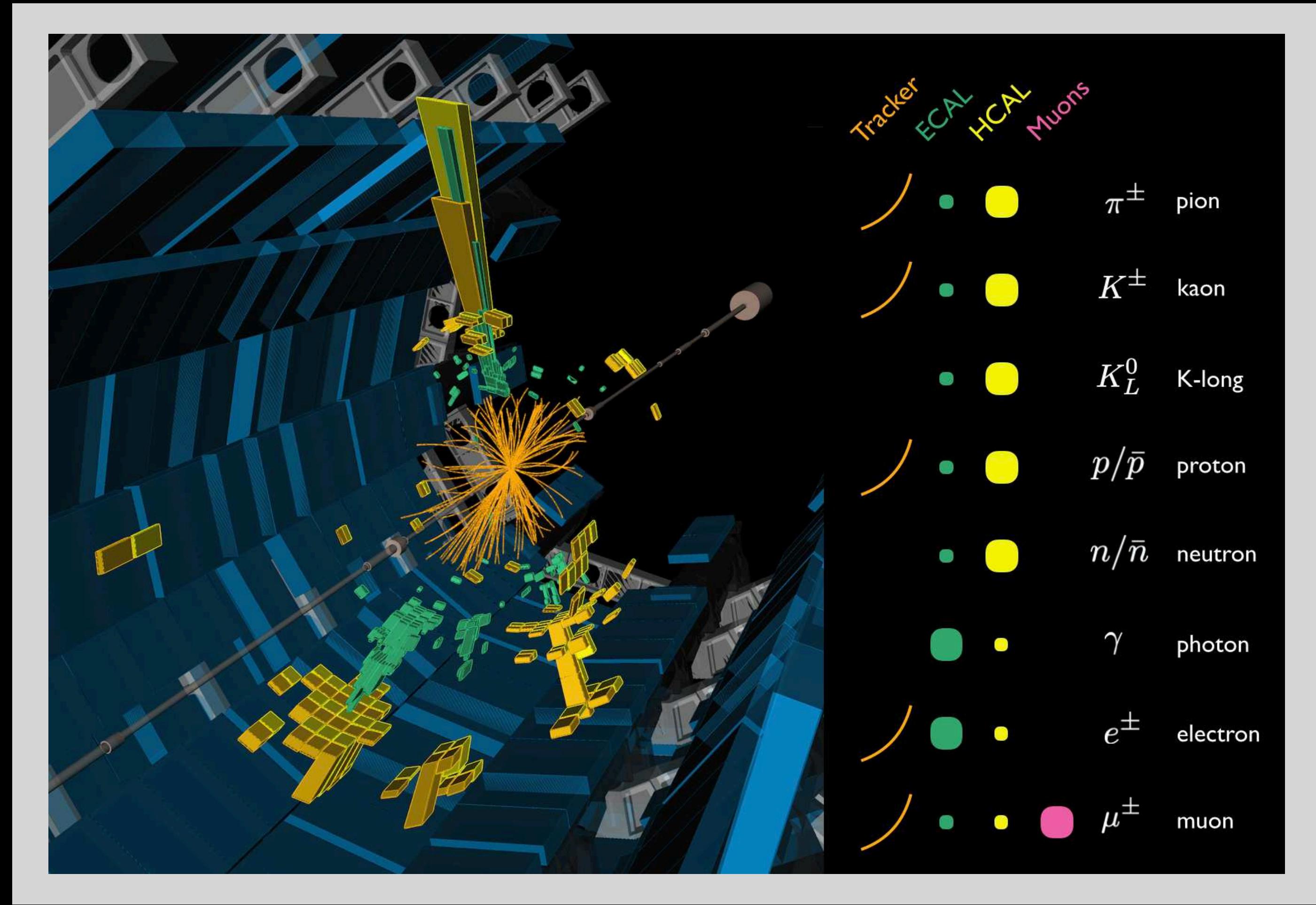
Emily Dickinson, #975



Additional Slides

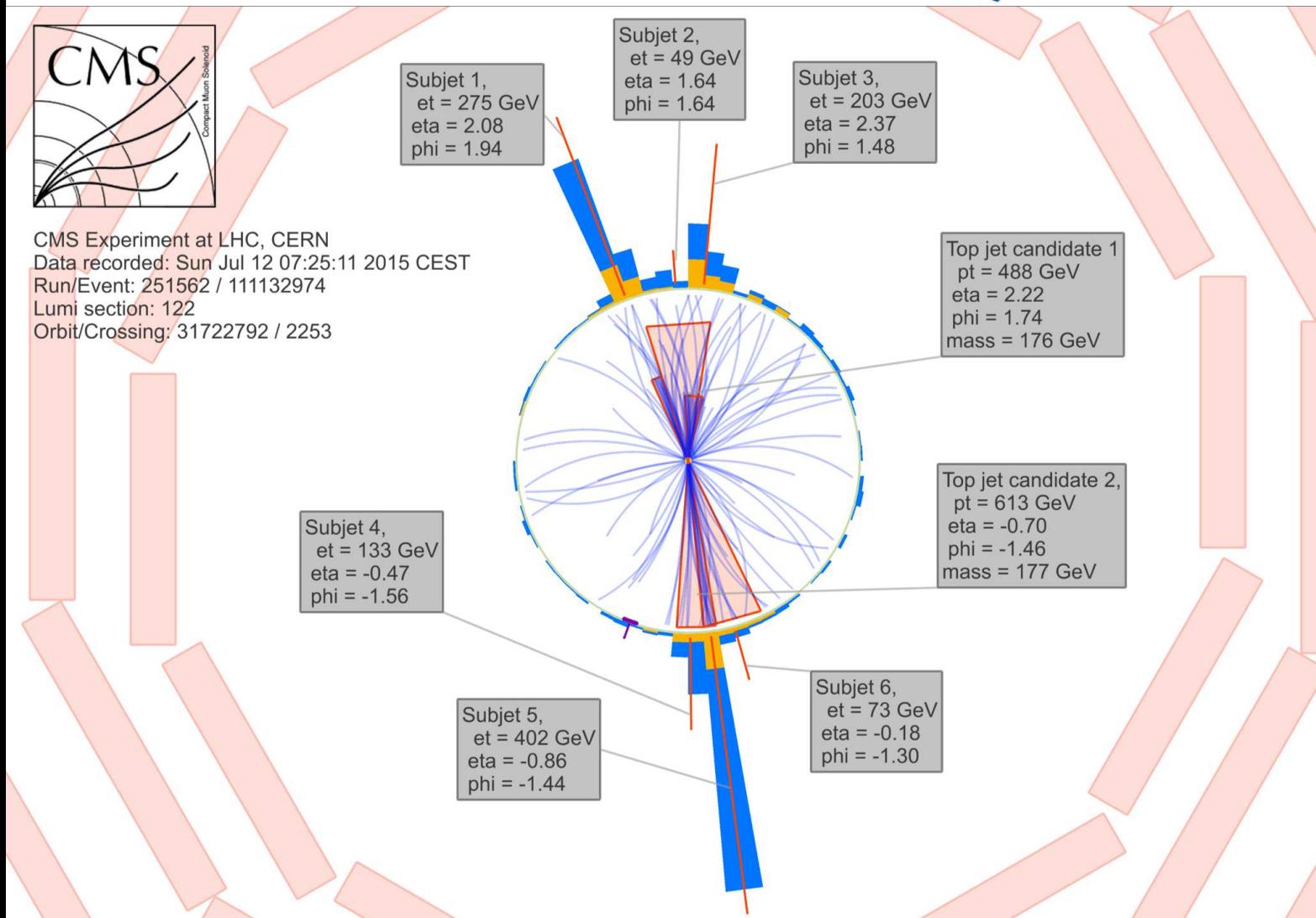
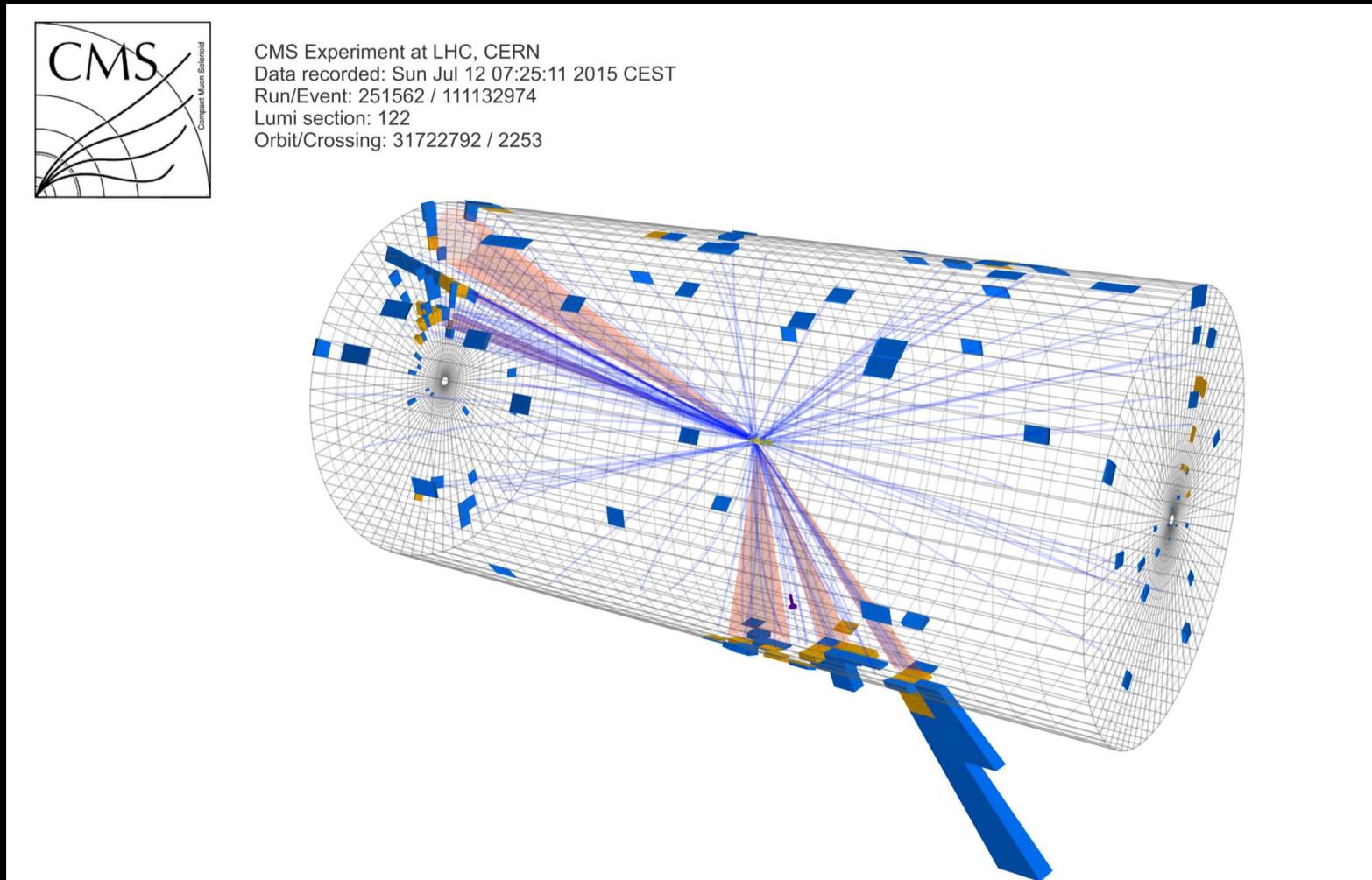
Particle Collider Events

High-energy collisions produce final state particles with **energy**, **direction**, **charge**, **flavor**, and **other quantum numbers**

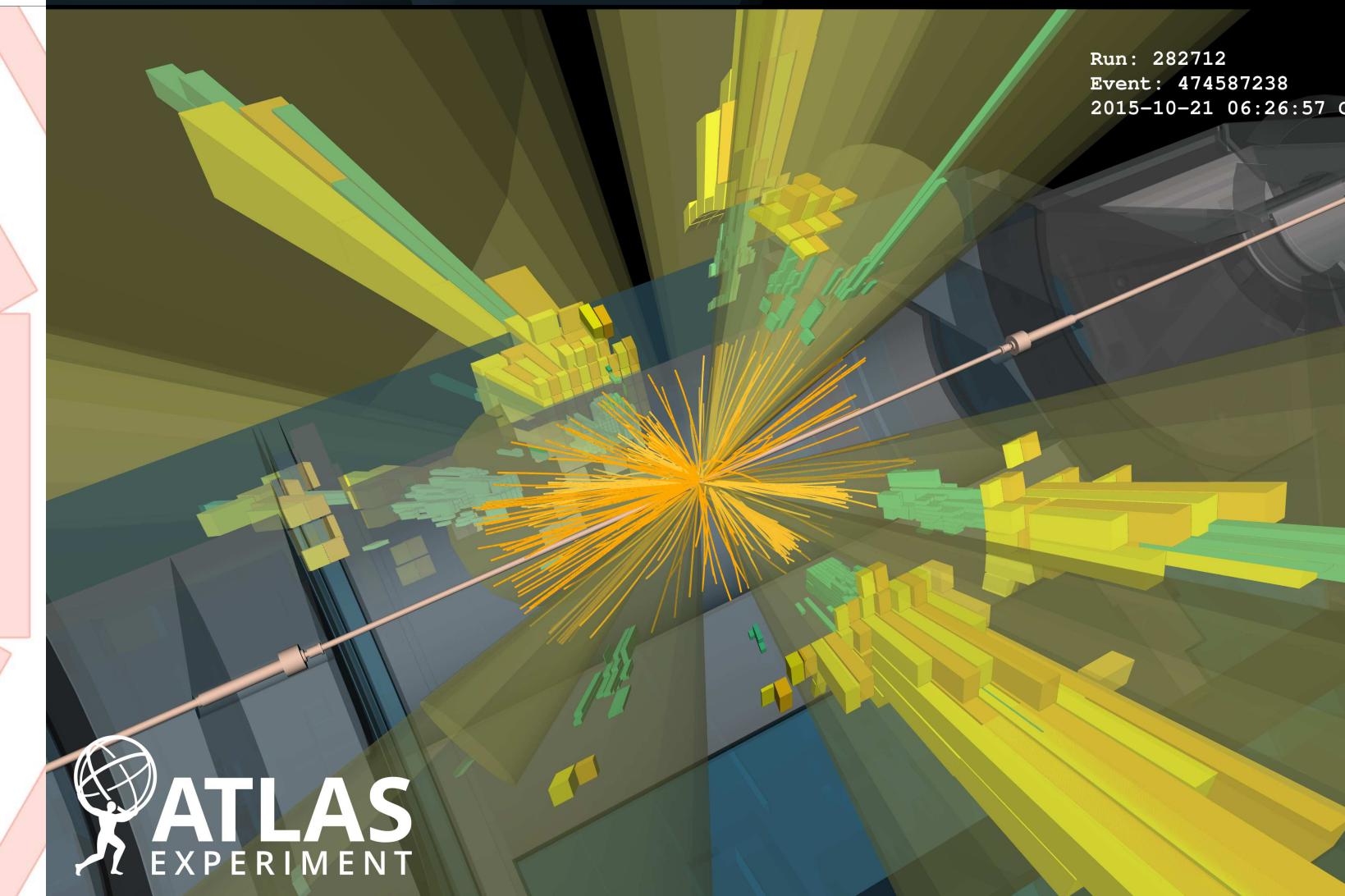
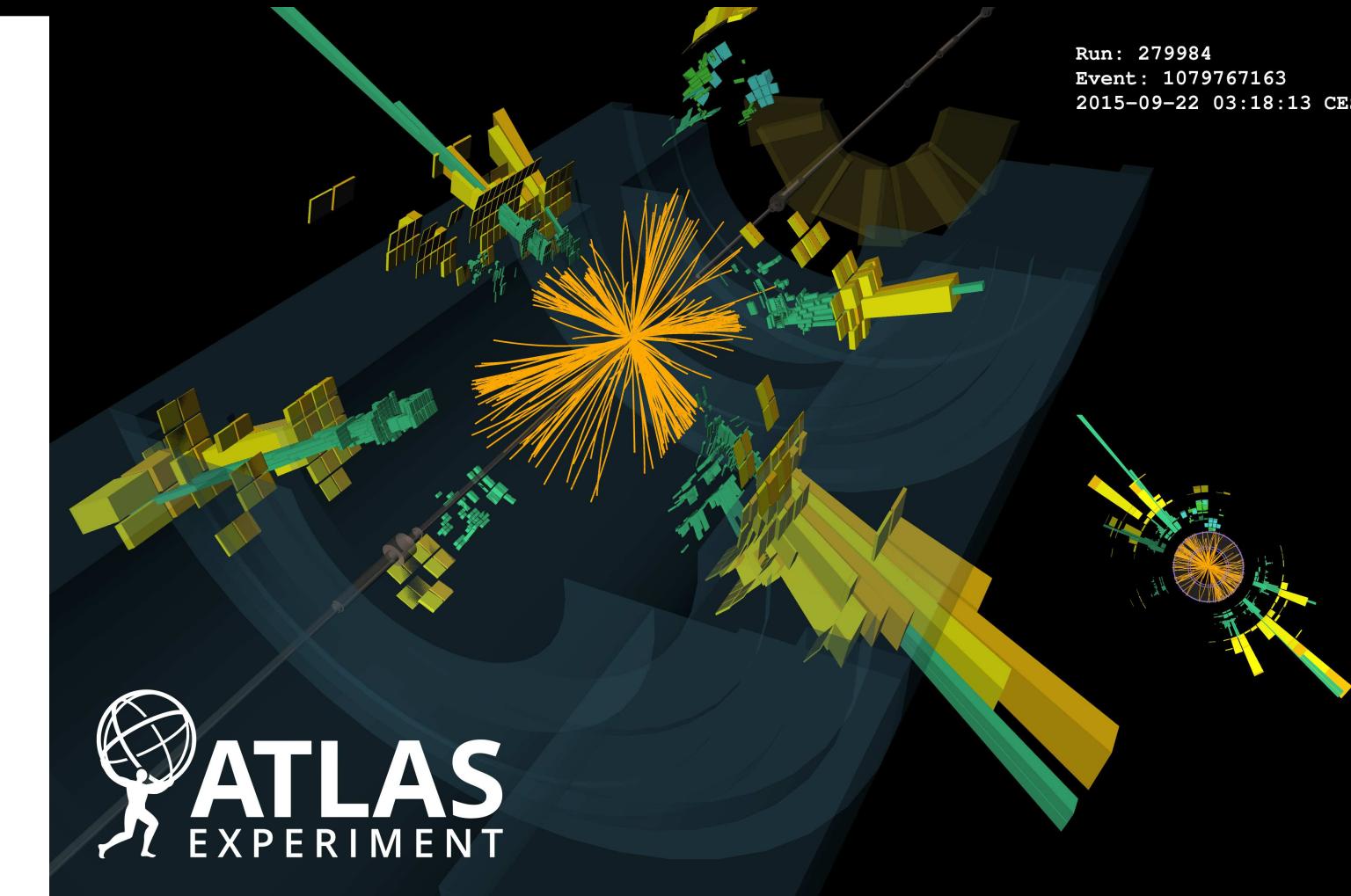


Jets at the Large Hadron Collider

CMS hadronic $t\bar{t}$ event

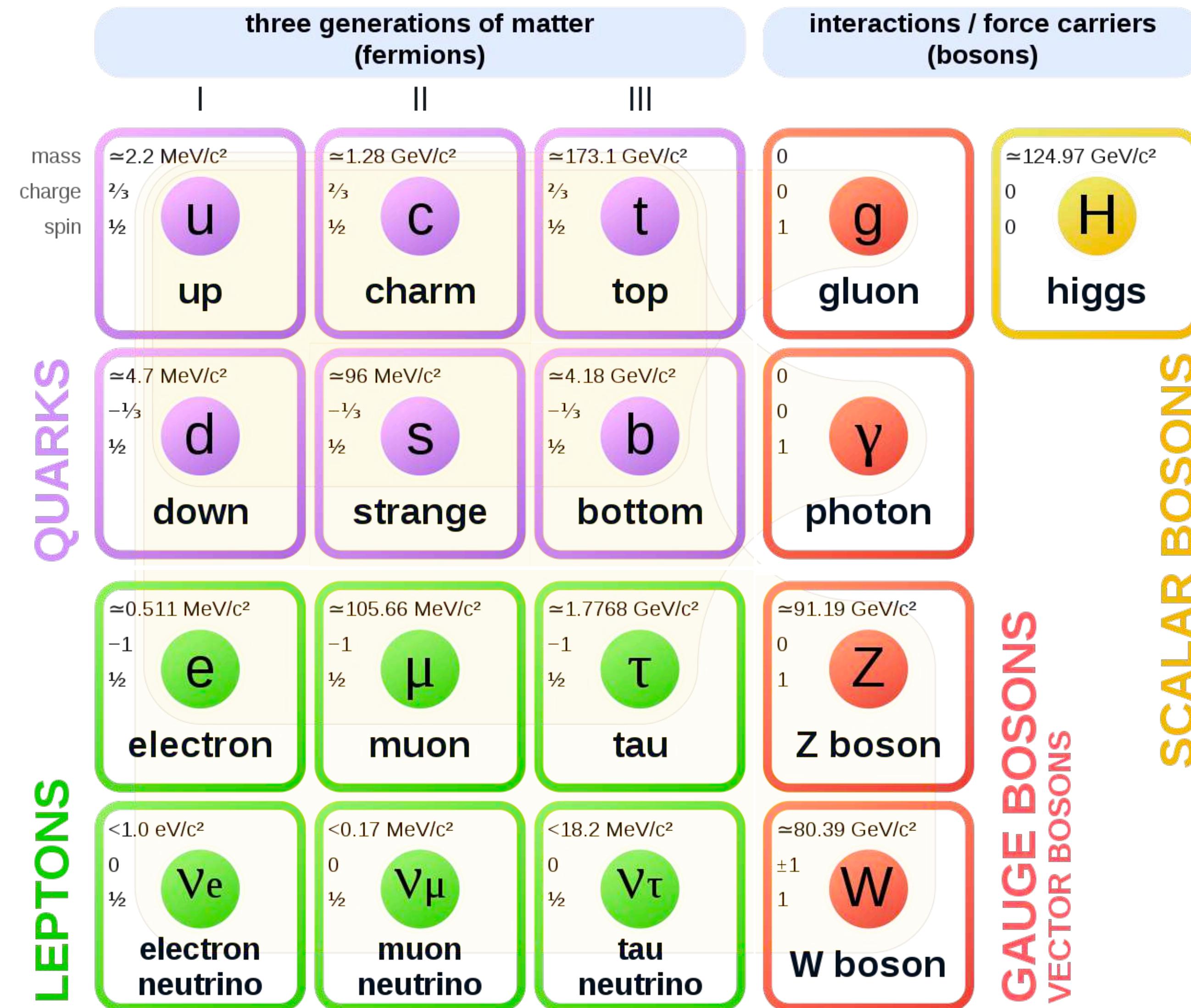


ATLAS high jet multiplicity events

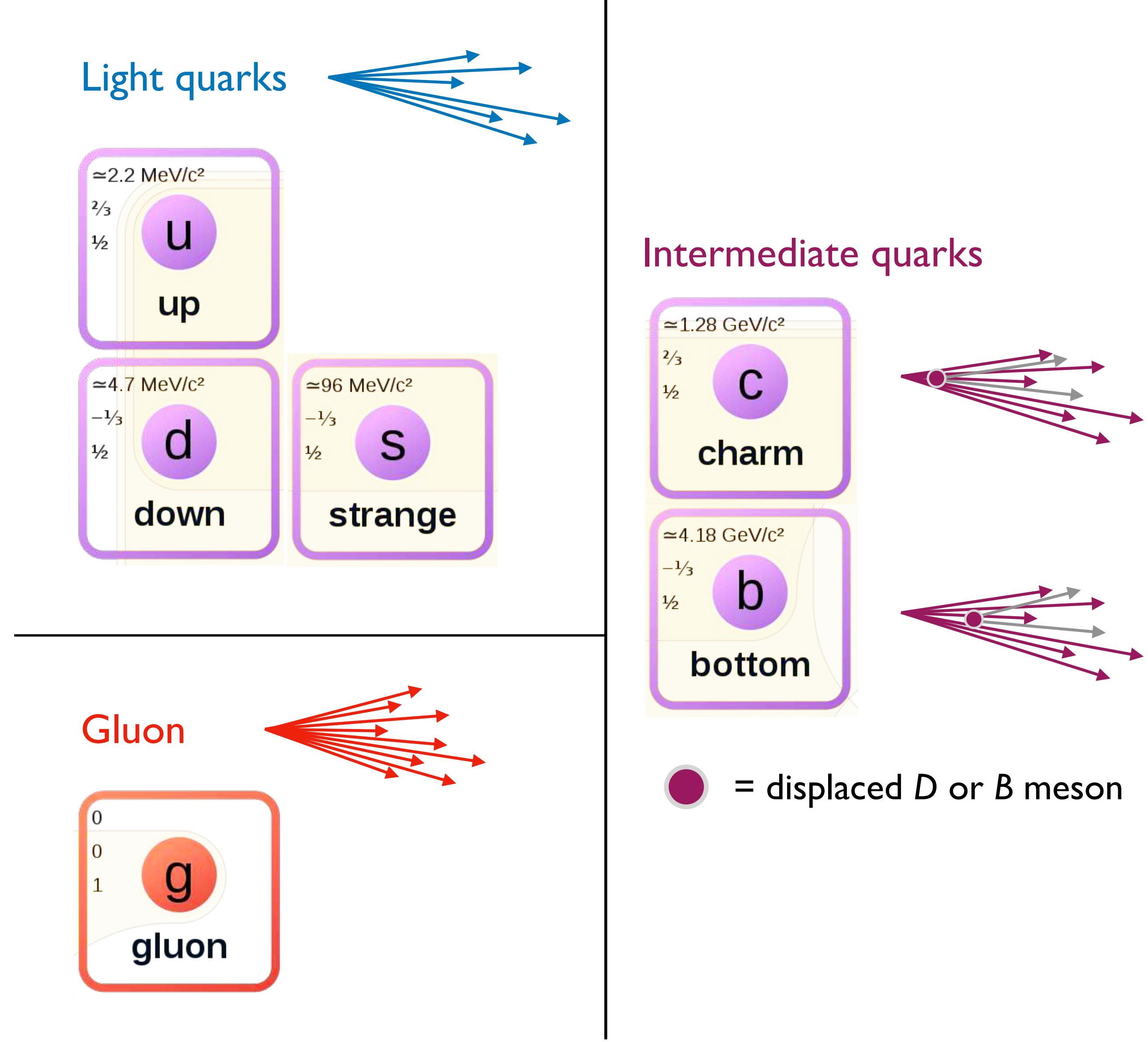


Jet substructure techniques enabled by fantastic detector resolution and reconstruction

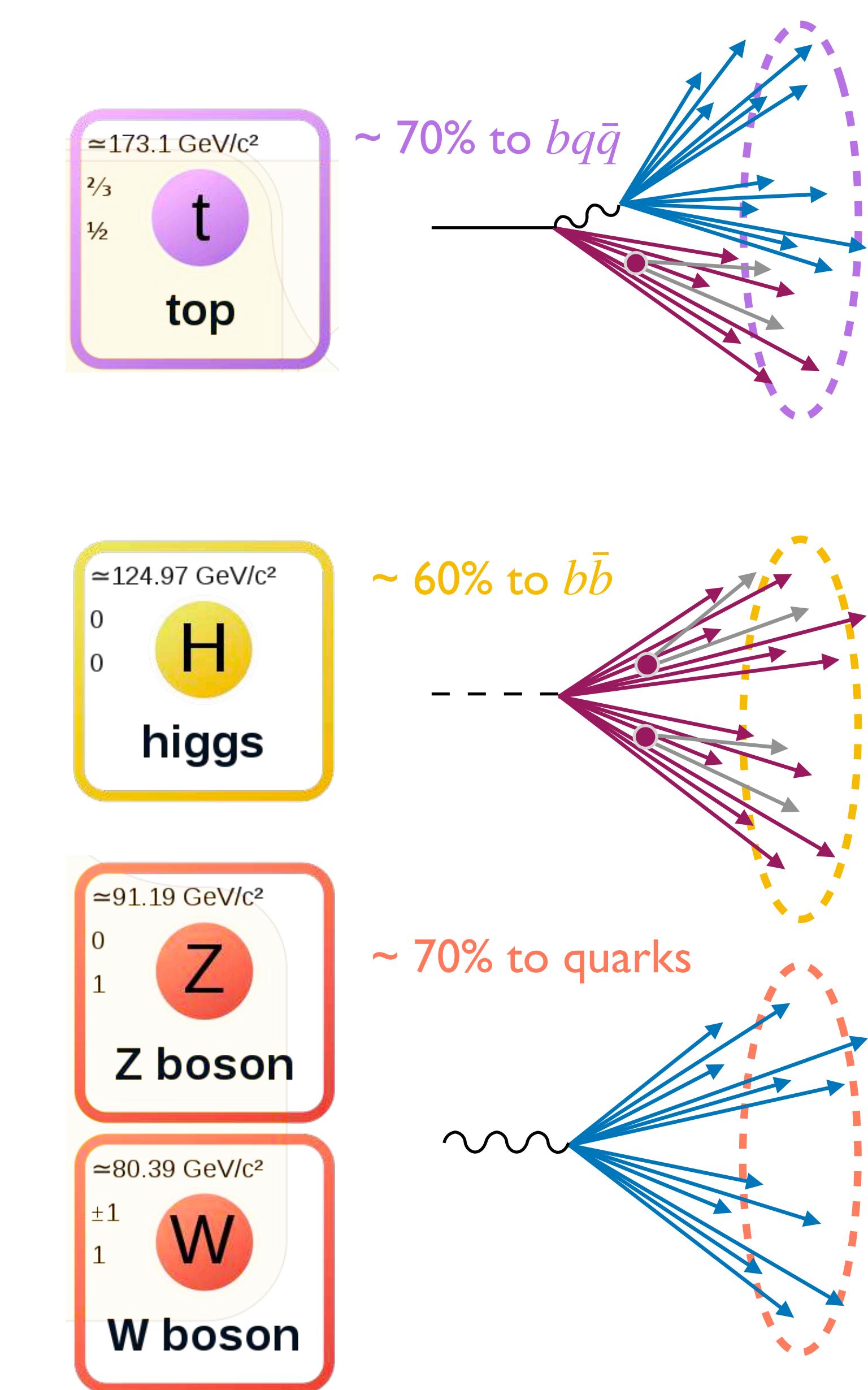
Standard Model of Particle Physics



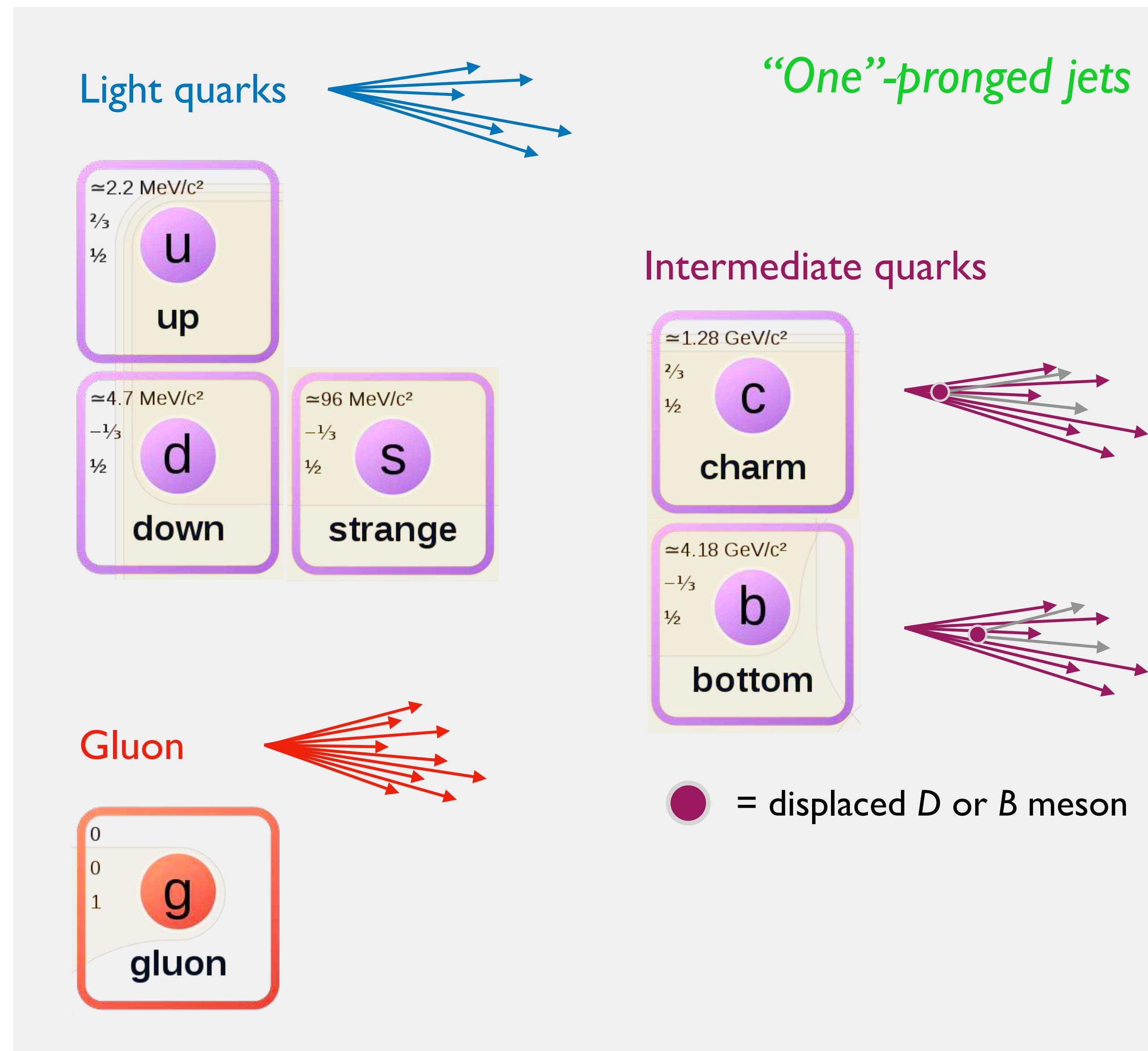
Standard Model of Particle Physics – as Jets



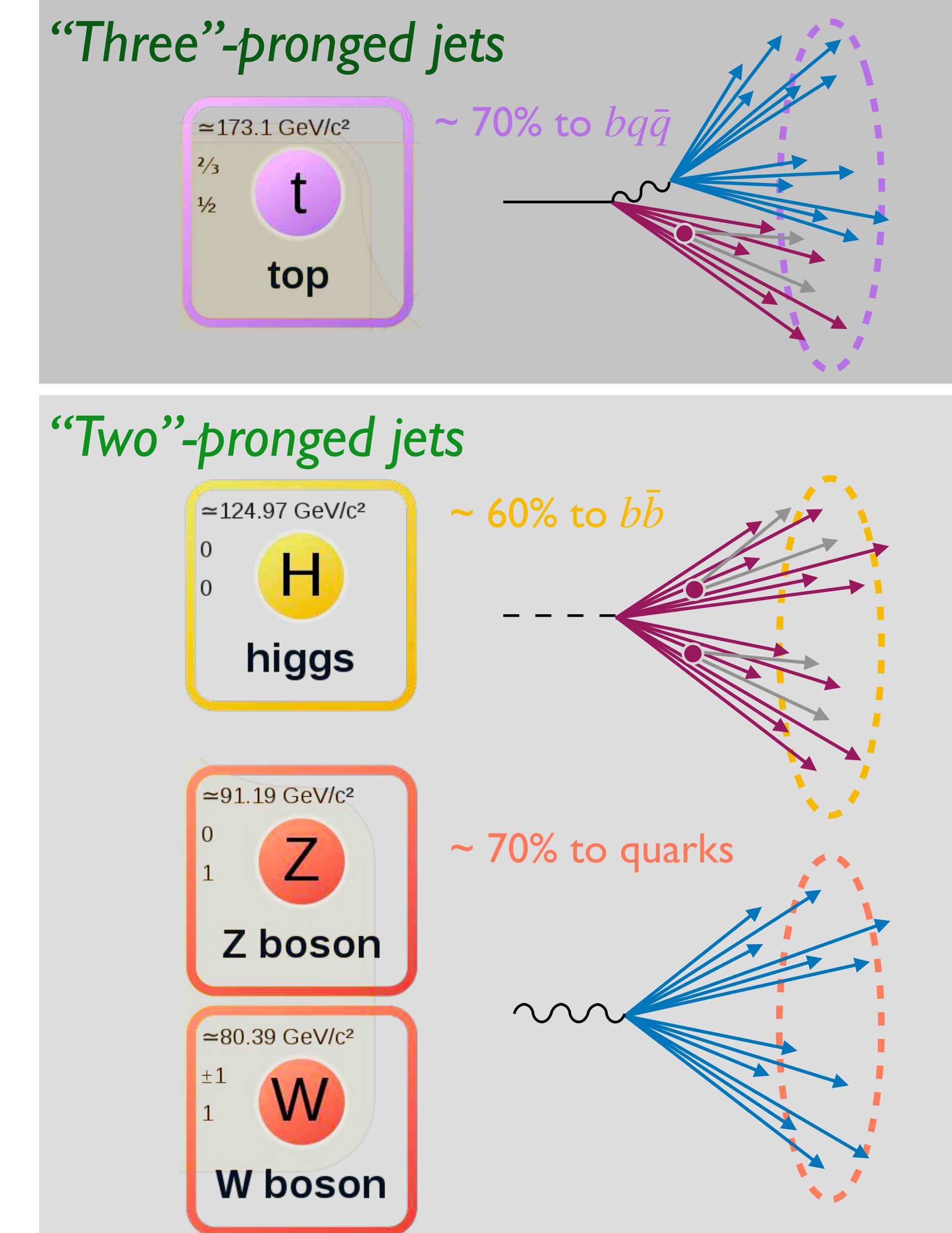
Heavy particles that further decay



Standard Model of Particle Physics – as Jets

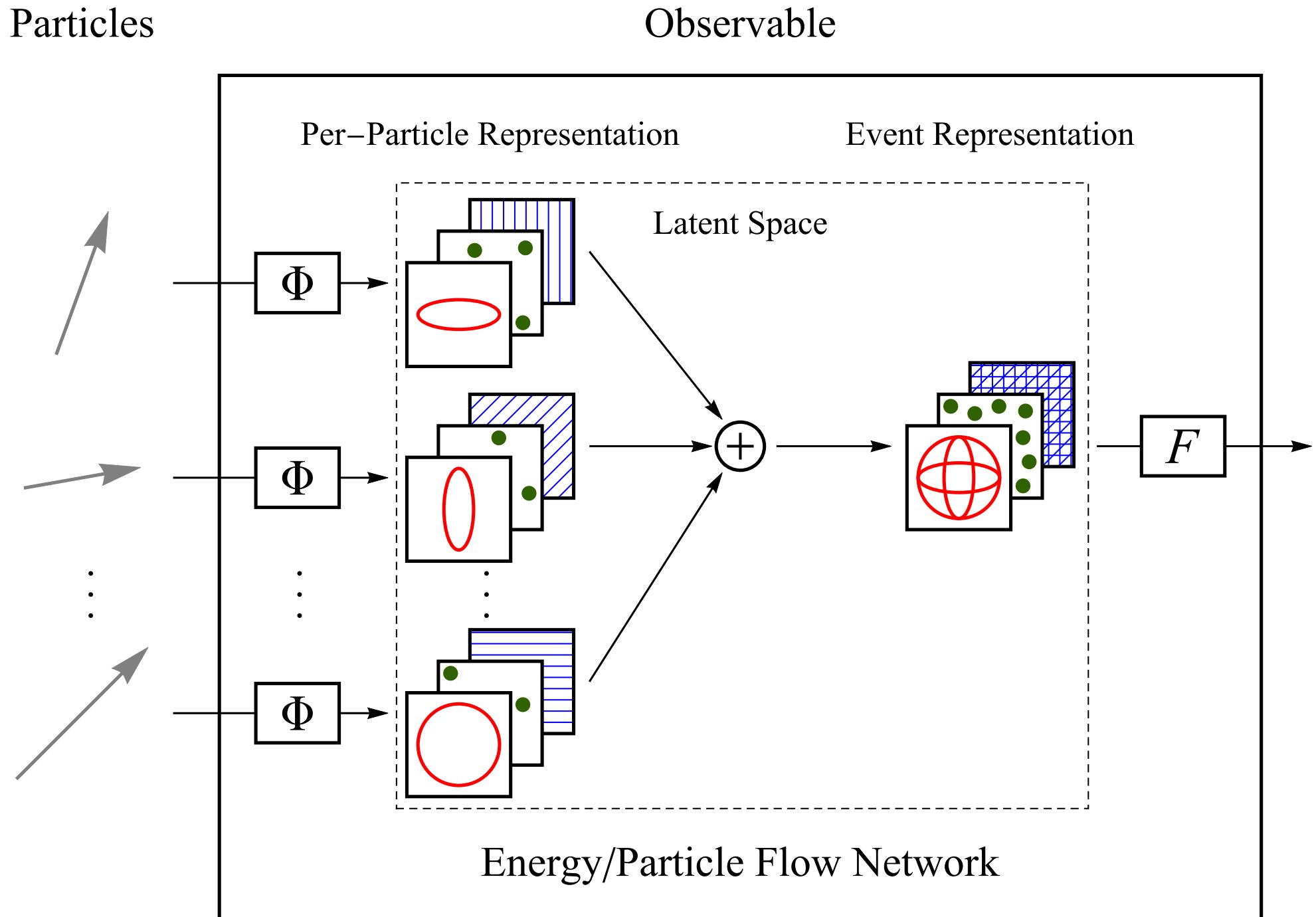


Heavy particles that further decay



Energy/Particle Flow Networks (EFNs/PFNs)

[Zaheer, Kottur, Ravanbakhsh, Póczos, Salakhutdinov, Smola, [I703.06114](#);
 PTK, Metodiev, Thaler, [I810.05165](#);
[EnergyFlow Python Package](#)]



Particle Flow Network (PFN)

$$\text{PFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \Phi(p_i^\mu) \right)$$

Fully general latent space

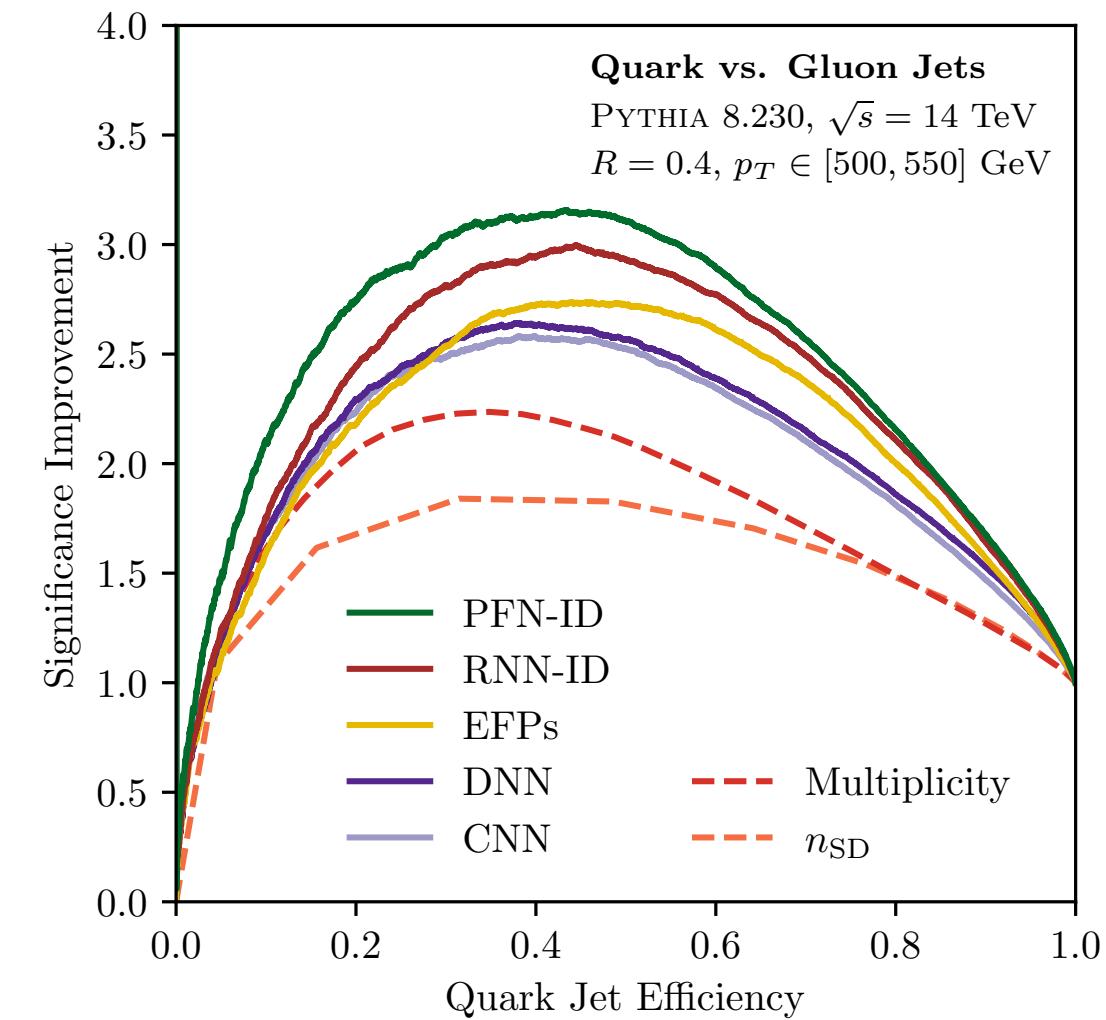
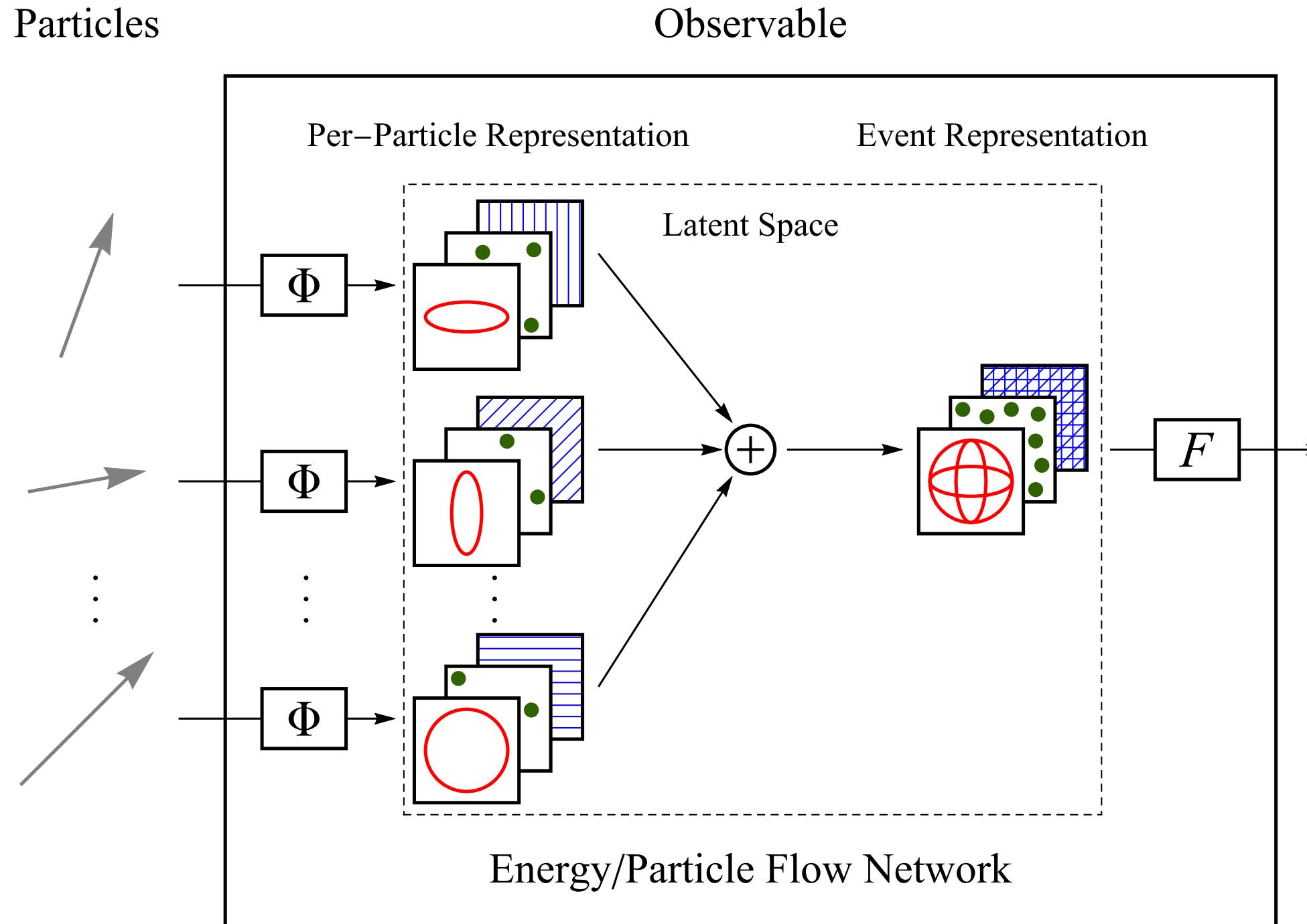
Energy Flow Network (EFN)

$$\text{EFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \textcolor{brown}{z}_i \Phi(\hat{p}_i) \right)$$

IRC-safe latent space

Energy/Particle Flow Networks (EFNs/PFNs)

[Zaheer, Kottur, Ravanbakhsh, Póczos, Salakhutdinov, Smola, [I703.06114](#);
 PTK, Metodiev, Thaler, [I810.05165](#);
[EnergyFlow Python Package](#)]



Improved performance (and training)
 compared to RNN and CNN

Particle Flow Network (PFN)

$$\text{PFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \Phi(p_i^\mu) \right)$$

Fully general latent space

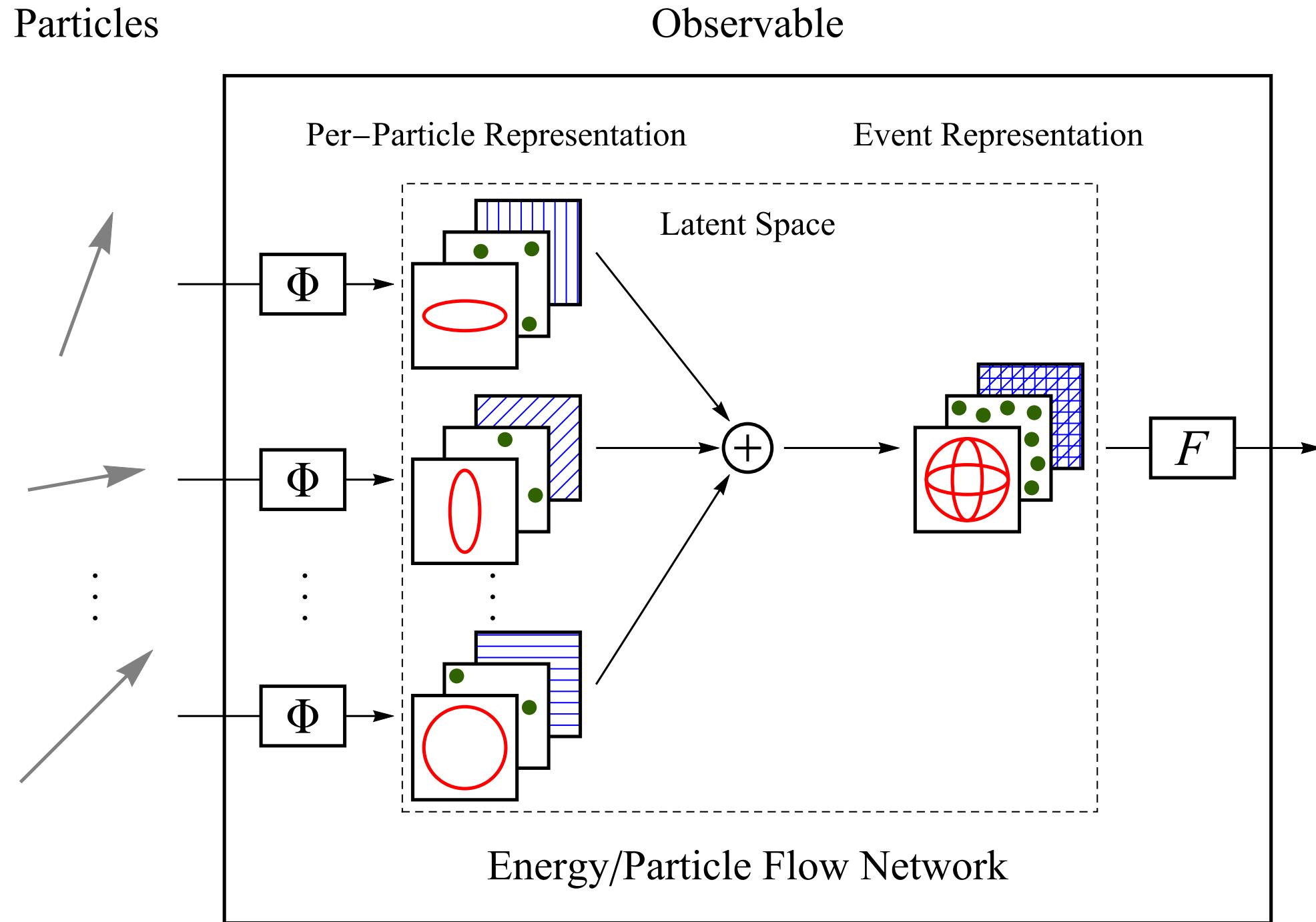
Energy Flow Network (EFN)

$$\text{EFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \textcolor{brown}{z}_i \Phi(\hat{p}_i) \right)$$

IRC-safe latent space

Energy/Particle Flow Networks (EFNs/PFNs)

[Zaheer, Kottur, Ravanbakhsh, Póczos, Salakhutdinov, Smola, [I703.06114](#);
 PTK, Metodiev, Thaler, [I810.05165](#);
[EnergyFlow Python Package](#)]



Particle Flow Network (PFN)

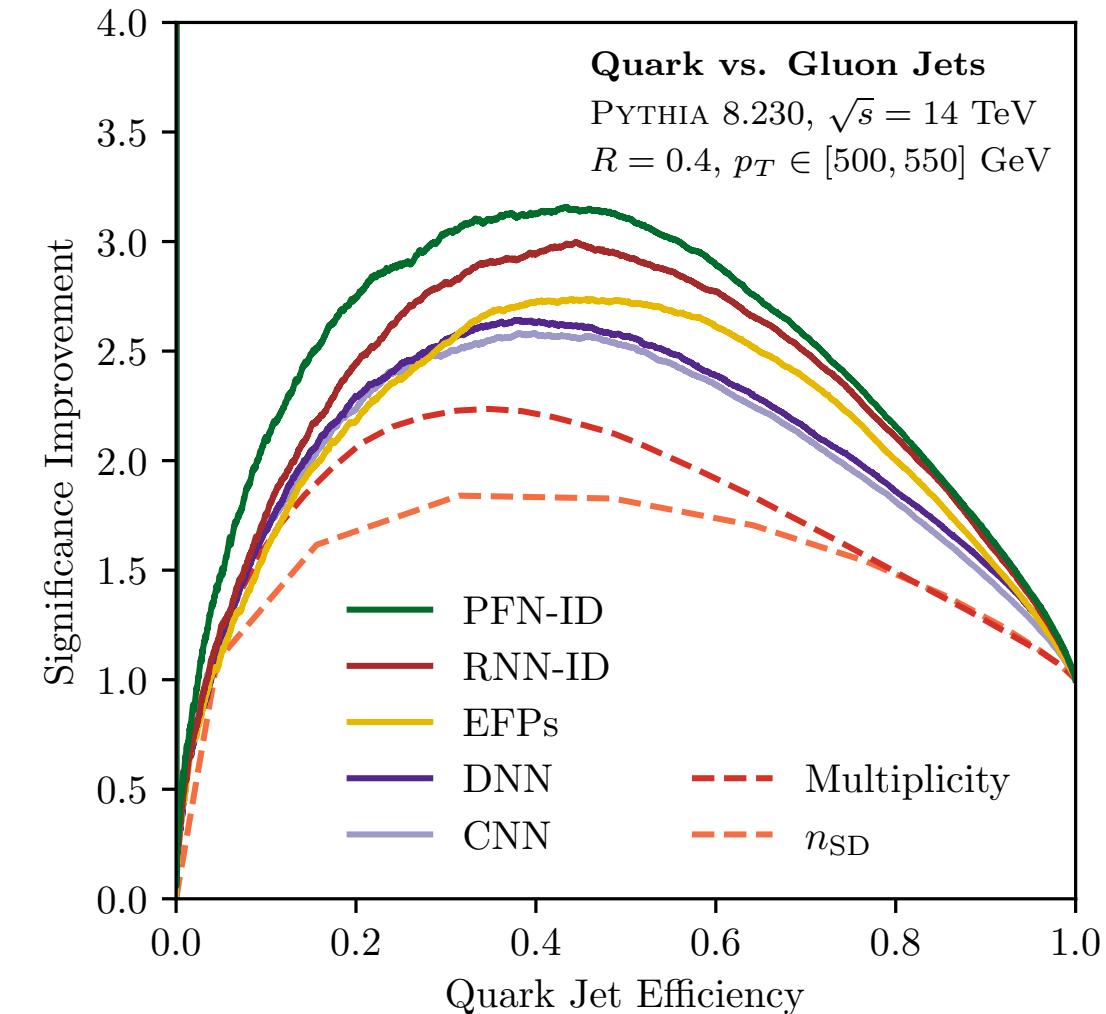
$$\text{PFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \Phi(p_i^\mu) \right)$$

Fully general latent space

Energy Flow Network (EFN)

$$\text{EFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \textcolor{brown}{z}_i \Phi(\hat{p}_i) \right)$$

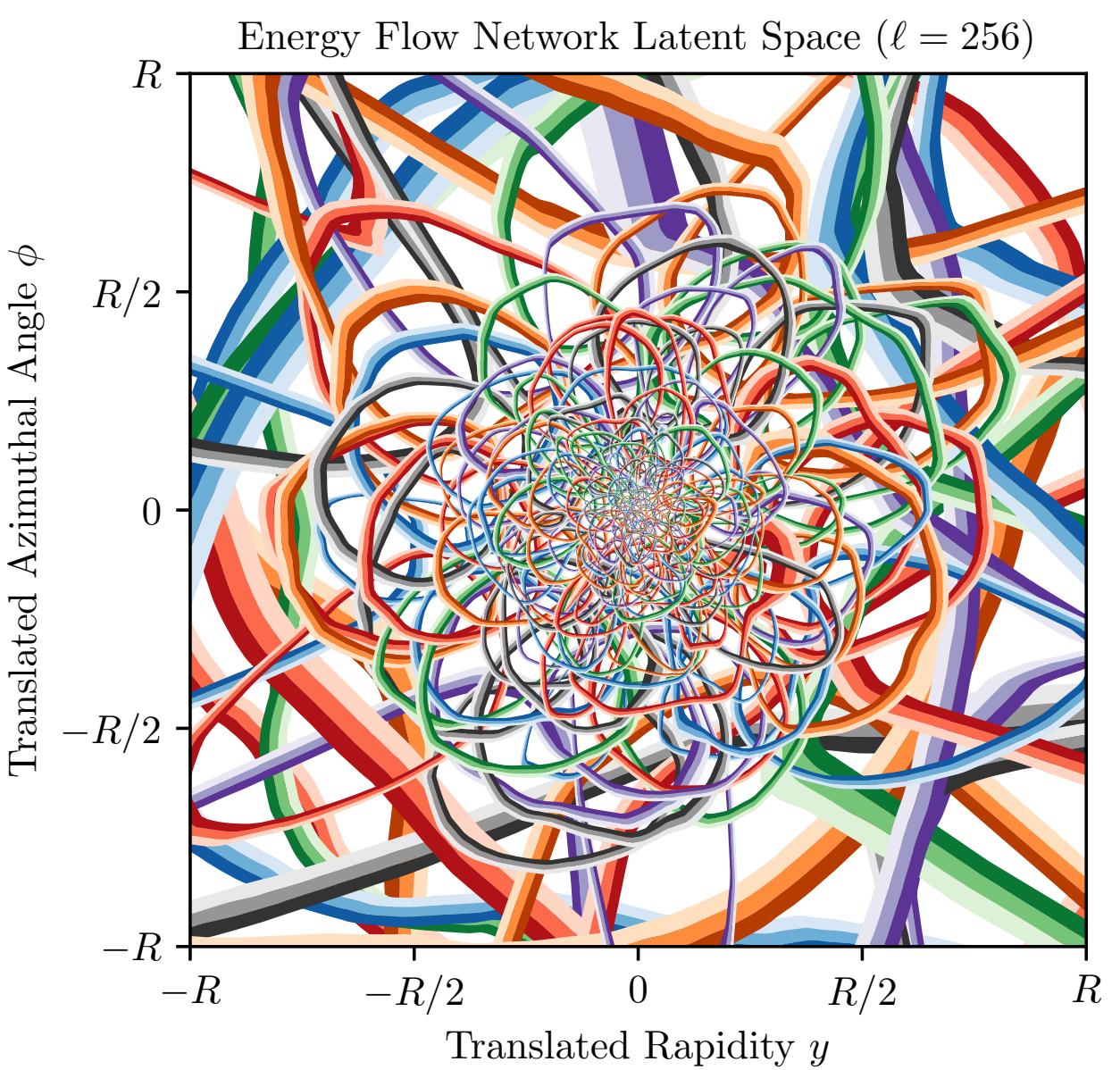
IRC-safe latent space



Improved performance (and training)
 compared to RNN and CNN

Latent space visualization reveals
 what the network has learned

Dynamic pixel sizing related to
 collinear singularity of QCD!



Quark vs. Gluon: Classification Performance

PFN-ID: Full particle flavor info

$$(\gamma, \pi^\pm, K^\pm, K_L, p, \bar{p}, n, \bar{n}, e^\pm, \mu^\pm)$$

PFN-Ex: Experimentally accessible info

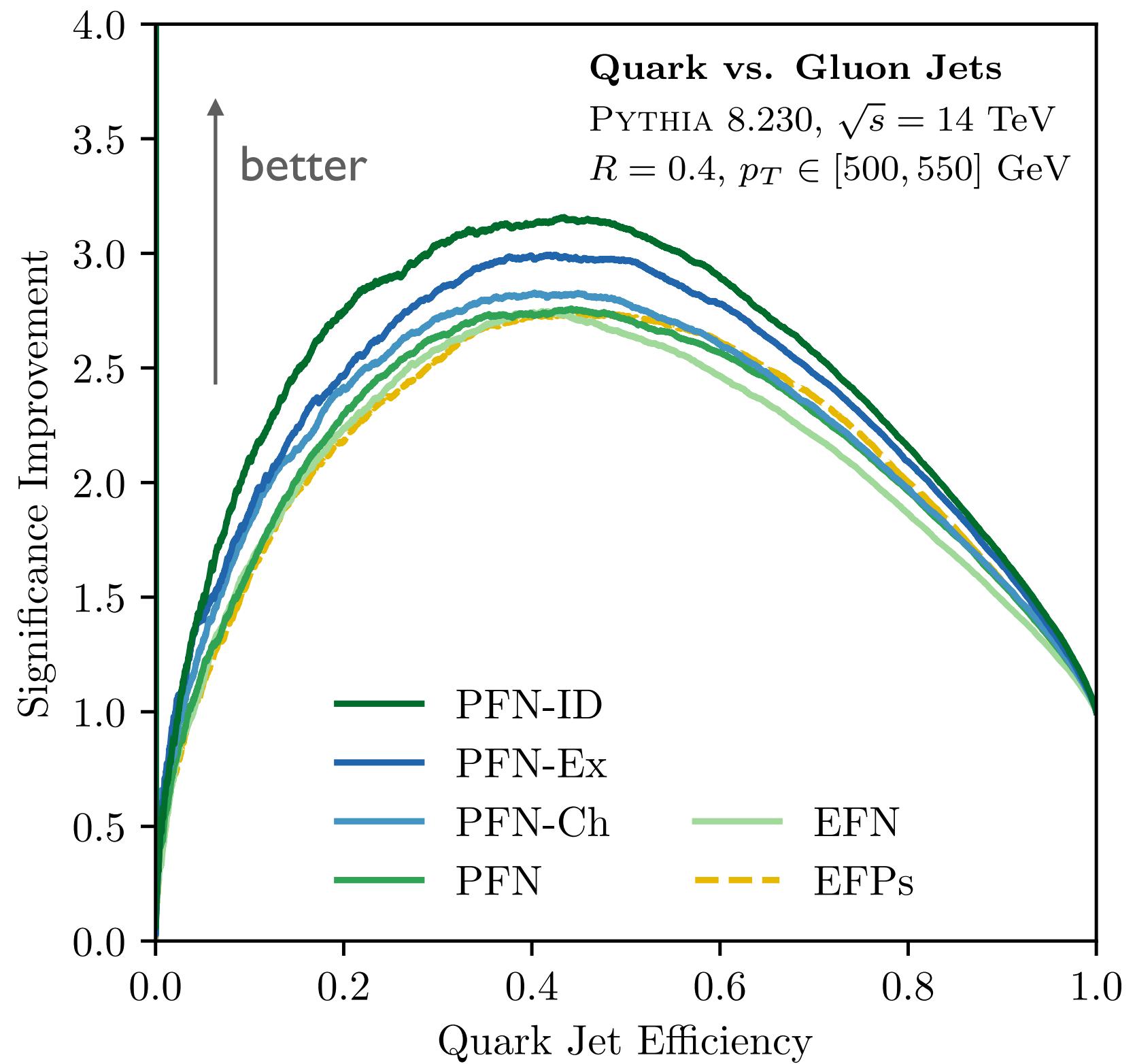
$$(\gamma, h^{\pm,0}, e^\pm, \mu^\pm)$$

PFN-Ch: Particle charge info

$$(+, 0, -)$$

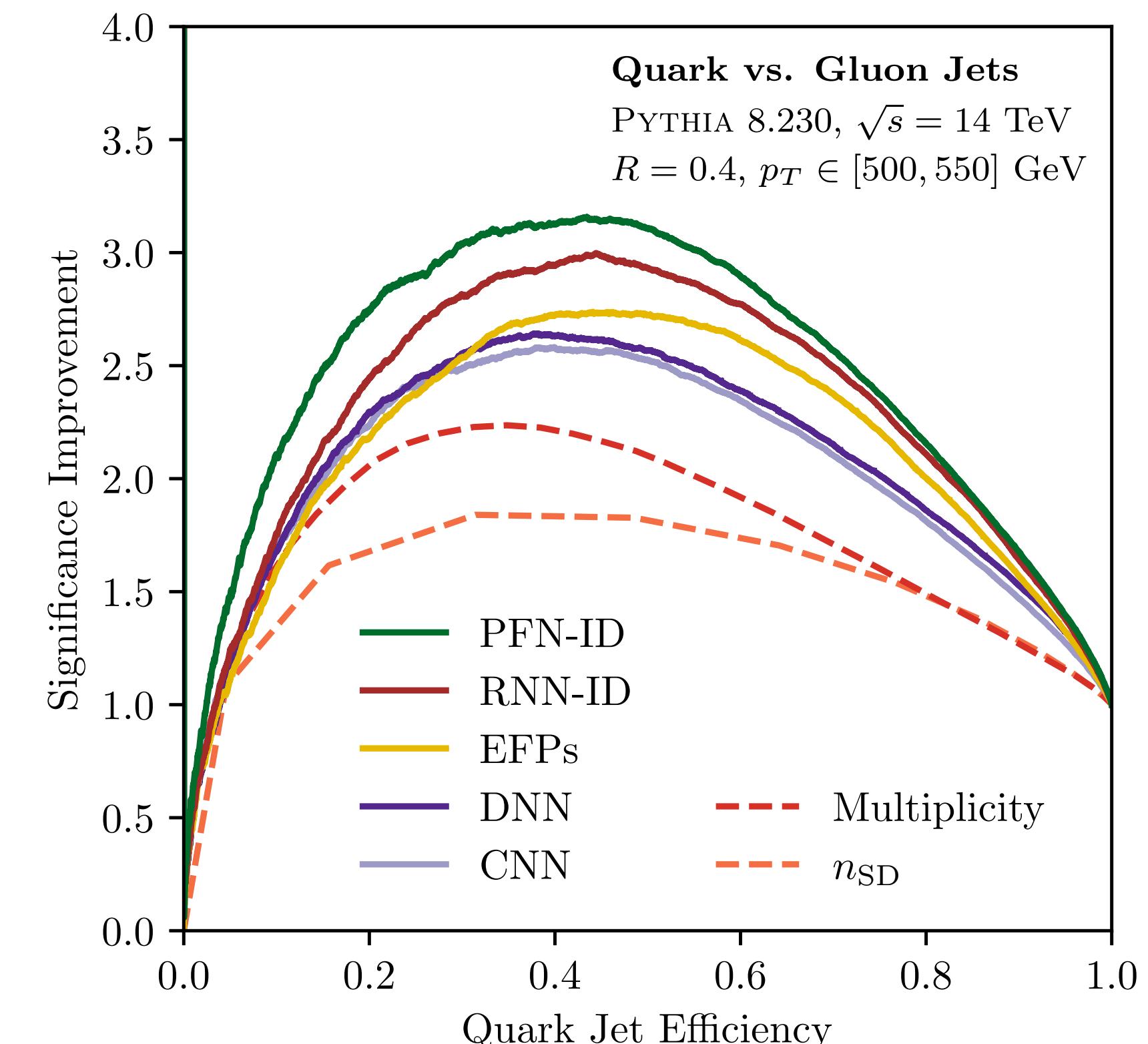
PFN: No particle type info, arbitrary energy dependence

EFN: **IRC**-safe latent space



Latent space dimension $\ell = 256$

EFPs are comparable to EFN



PFN-ID better than RNN-ID

Quark vs. Gluon: Latent Dimension Sweep

PFN-ID: Full particle flavor info

$$(\gamma, \pi^\pm, K^\pm, K_L, p, \bar{p}, n, \bar{n}, e^\pm, \mu^\pm)$$

PFN-Ex: Experimentally accessible info

$$(\gamma, h^{\pm,0}, e^\pm, \mu^\pm)$$

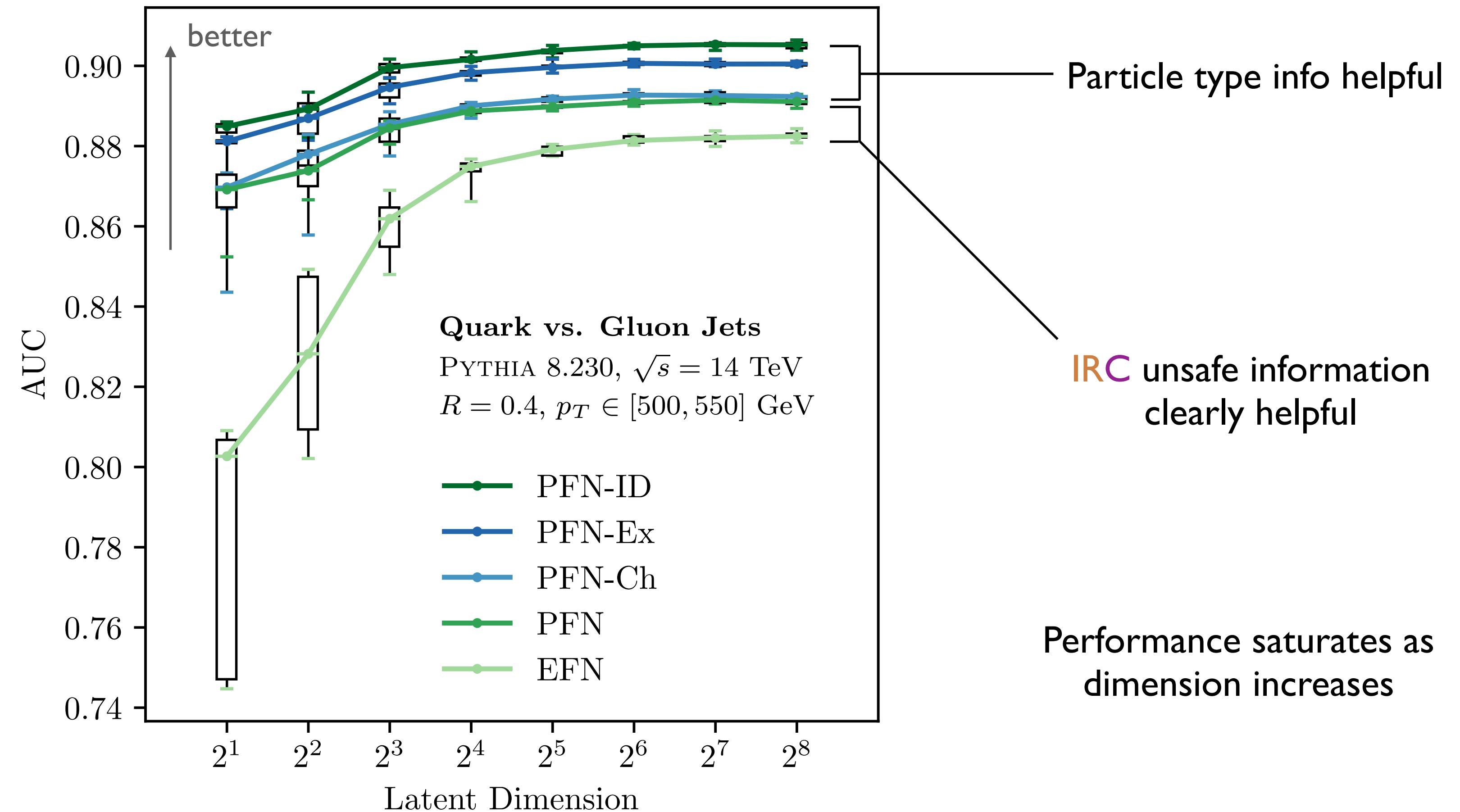
PFN-Ch: Particle charge info

$$(+, 0, -)$$

PFN: No particle type info, arbitrary

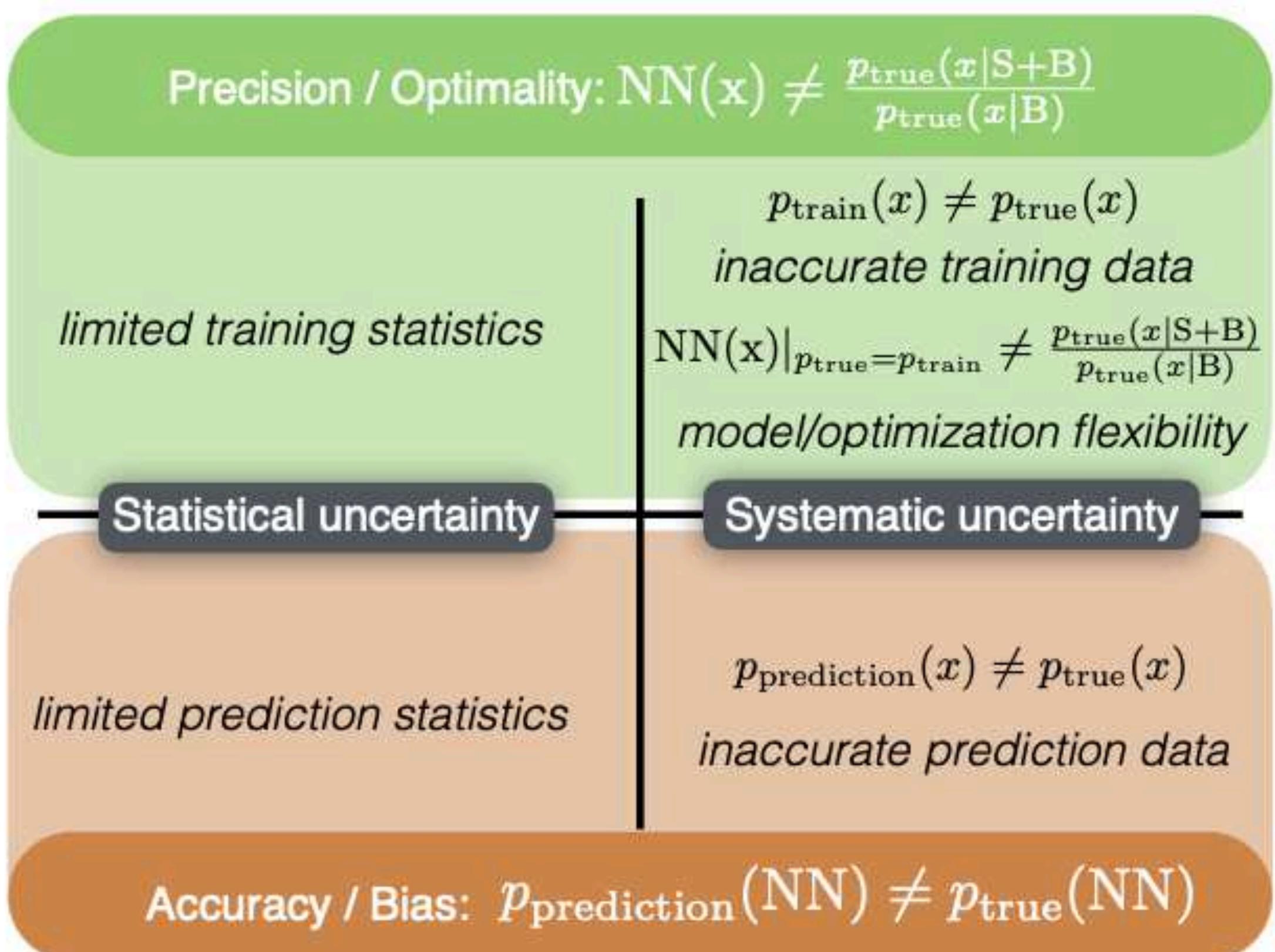
energy dependence

EFN: **IRC**-safe latent space



Dealing with Uncertainties

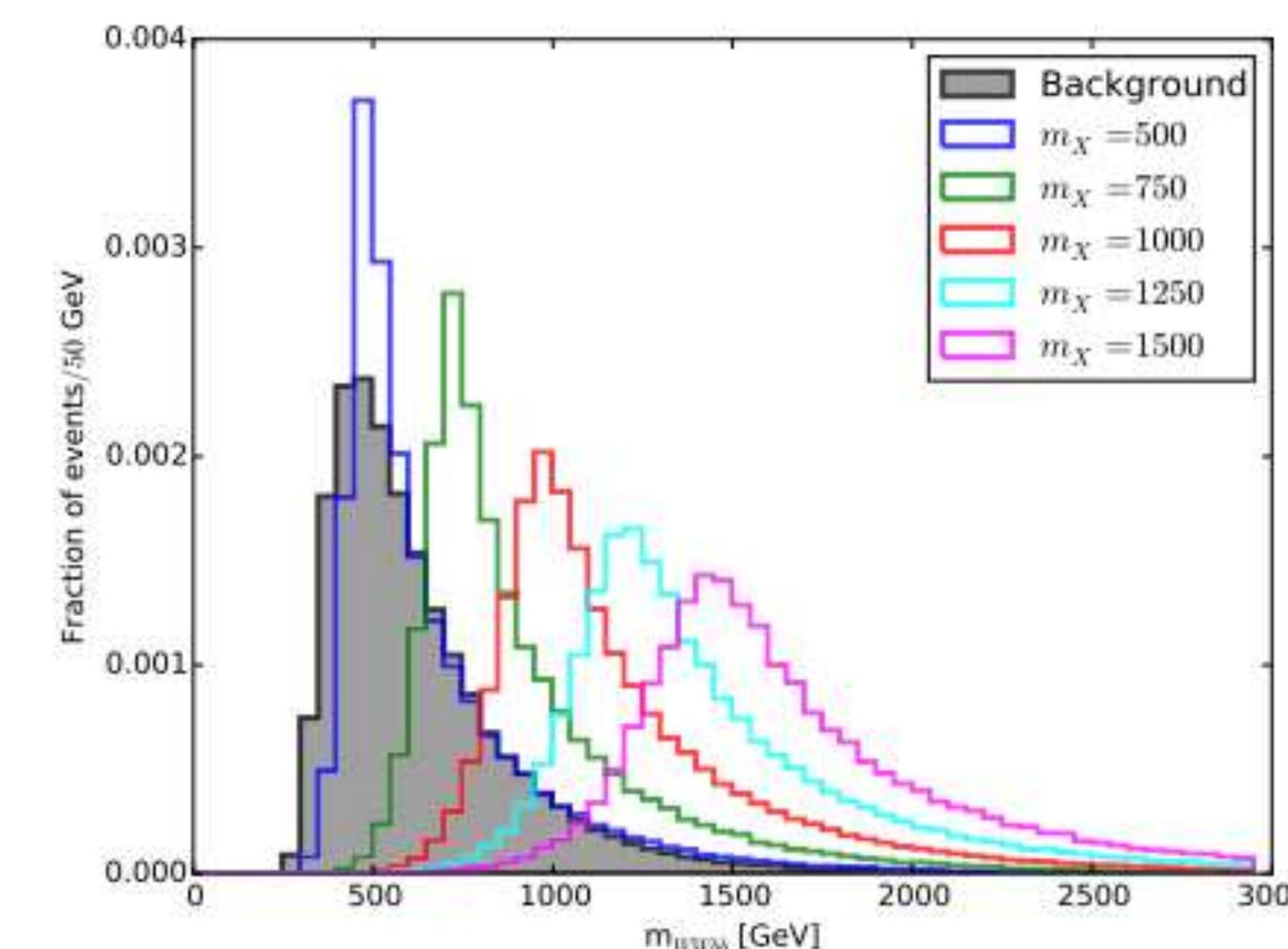
Sources of uncertainty in a statistical analysis



[Nachman, [1909.03081](#)]

Parametrized models could enable efficient profiling to handle systematic uncertainties

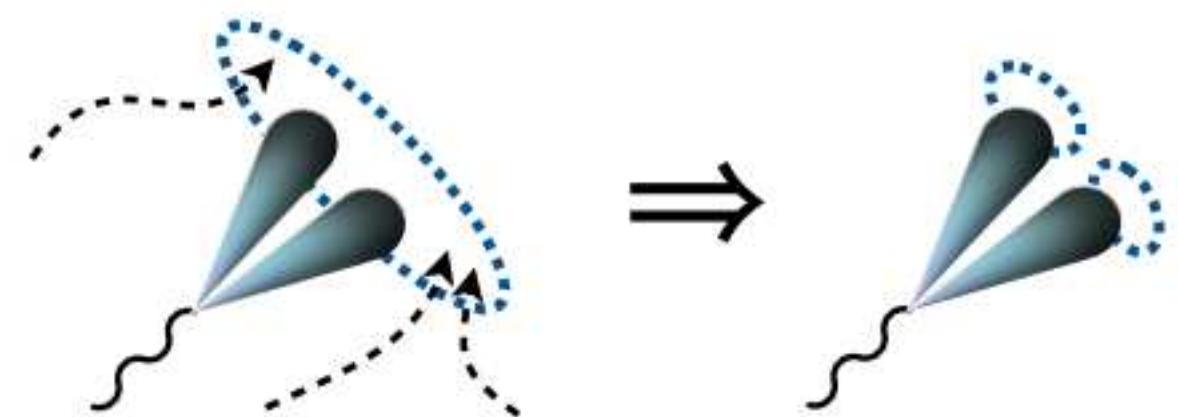
[similar to Baldi, Cranmer, Faust, Sadowski, Whiteson, [1601.07913](#)]



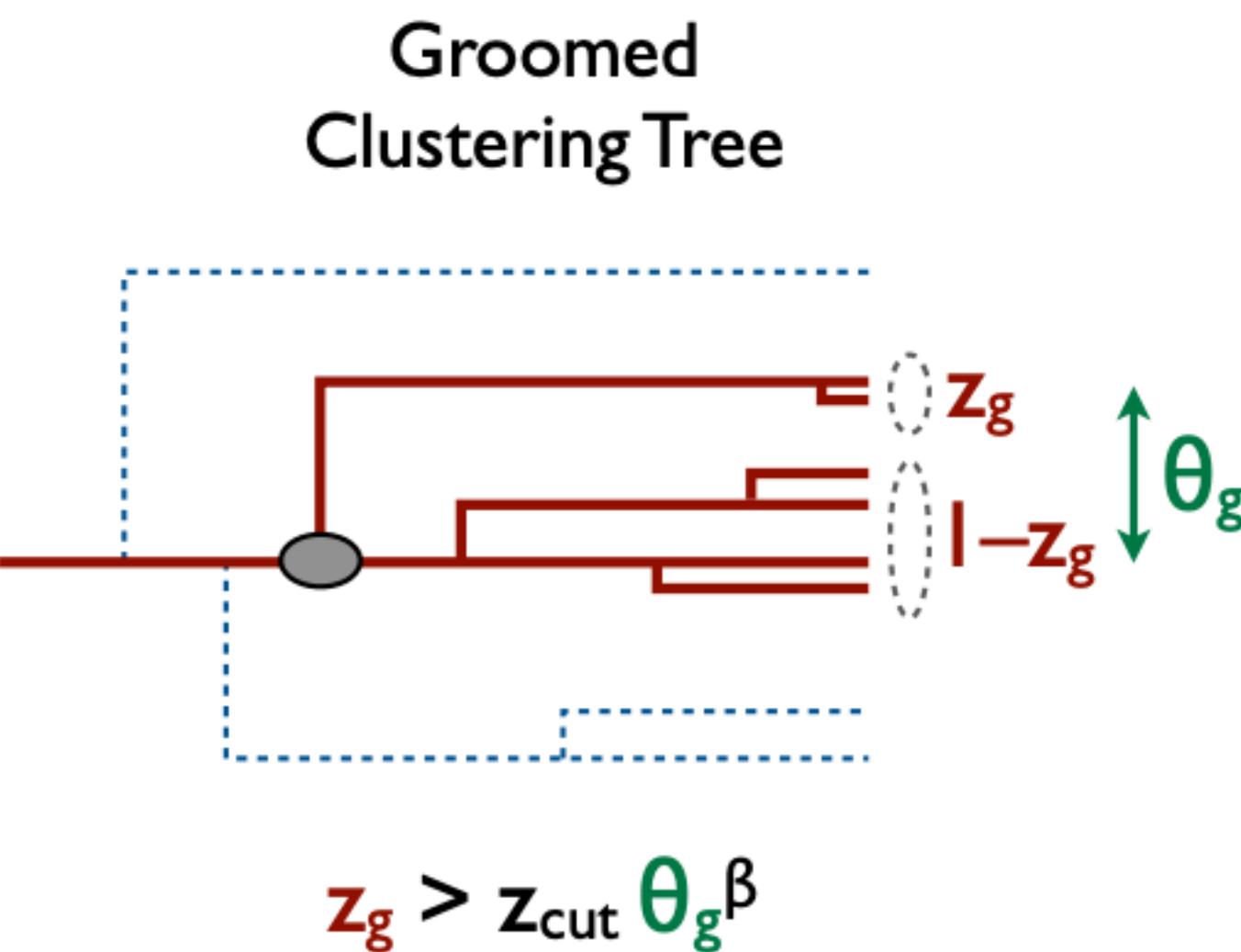
Training a neural network on several different signal masses and allowing it to interpolate between them

Soft Drop

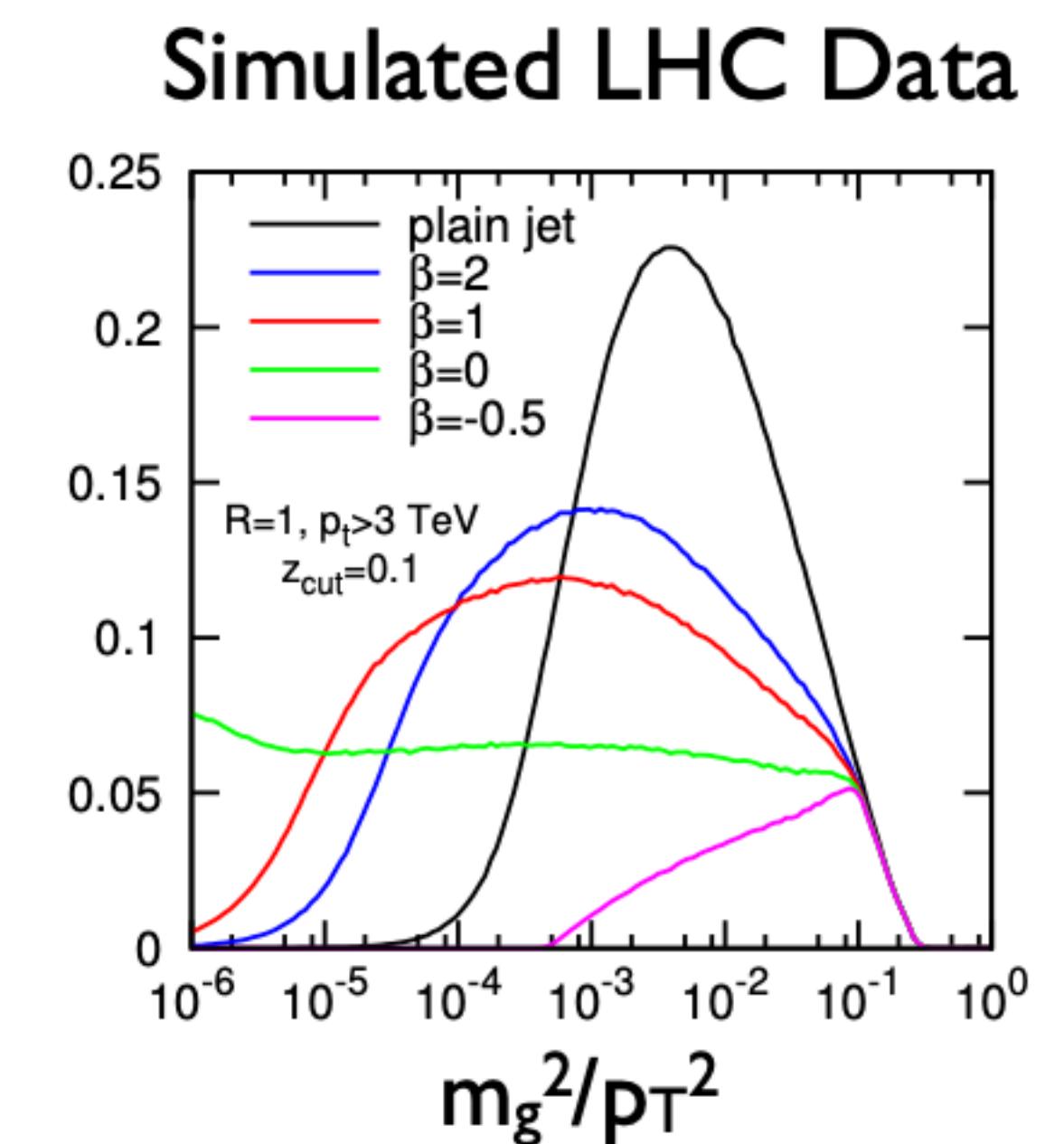
Contaminating radiation in jets
necessitates grooming



Soft Drop algorithm



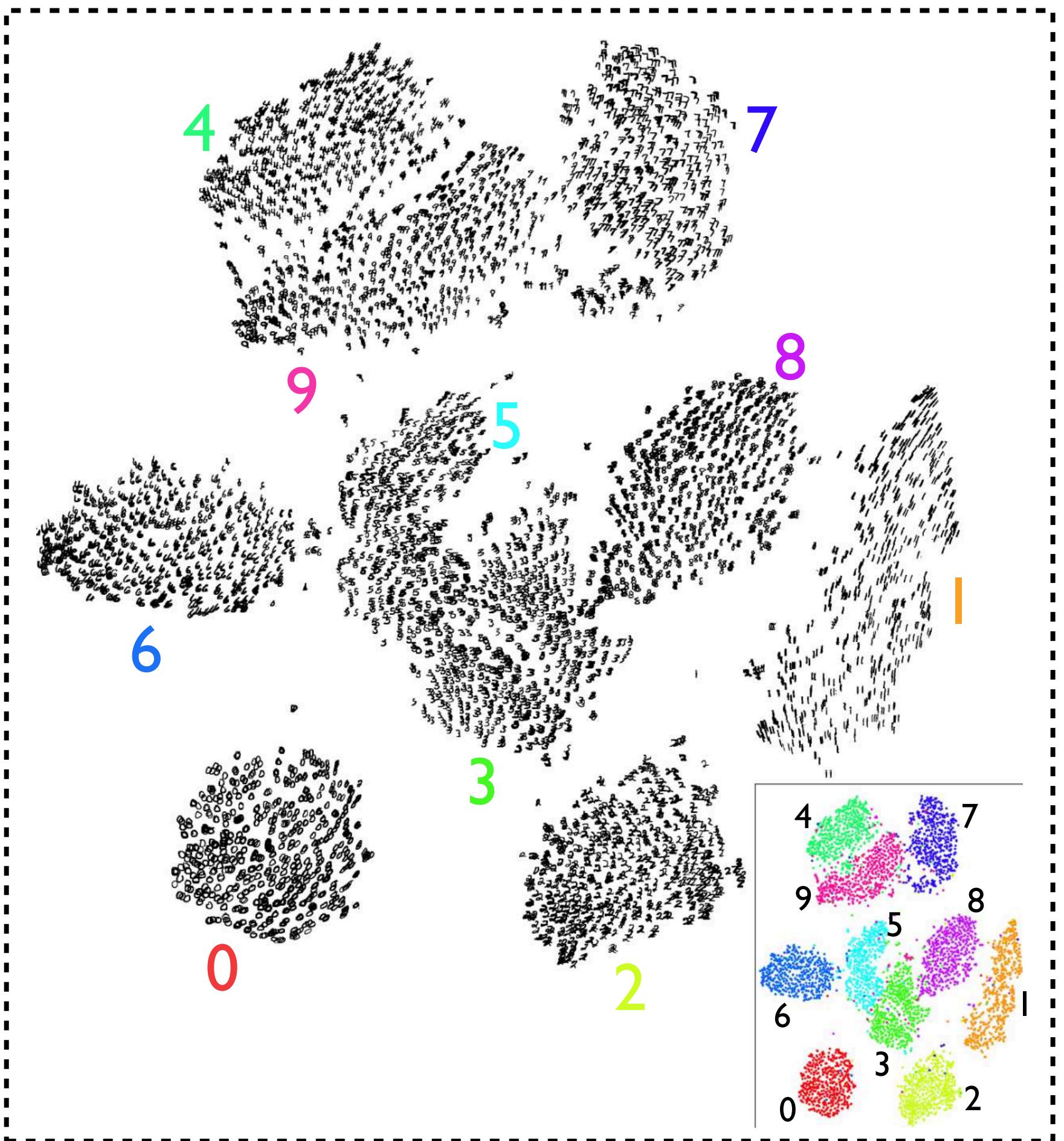
Calculating mass on SD jets



Diagrams by Jesse Thaler

Visualizing Geometry in the Space of Events

t-Distributed Stochastic Neighbor Embedding (t-SNE)
MNIST handwritten digits

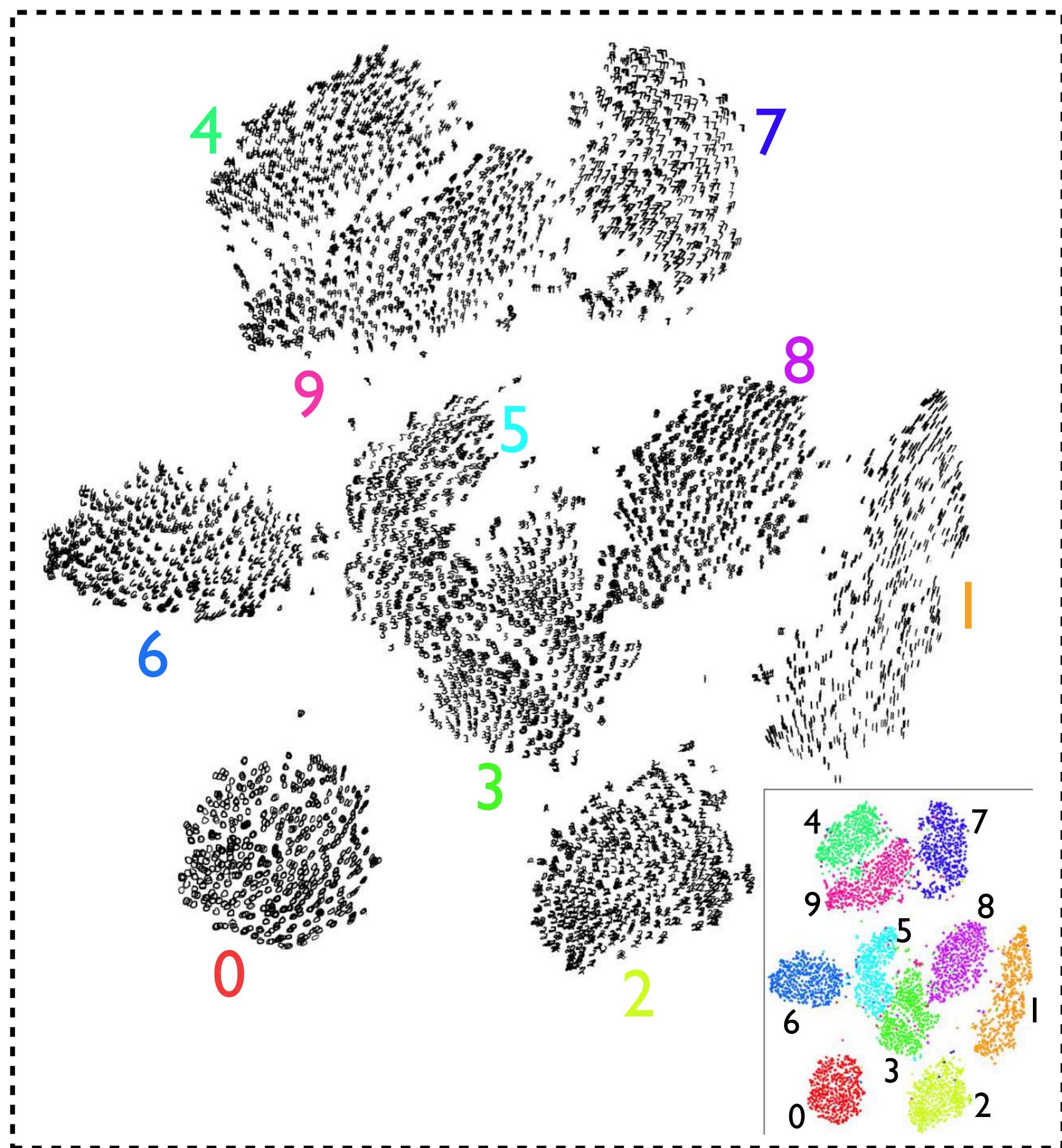


[L. van der Maaten, G. Hinton, JMLR 2008]

Visualizing Geometry in the Space of Events

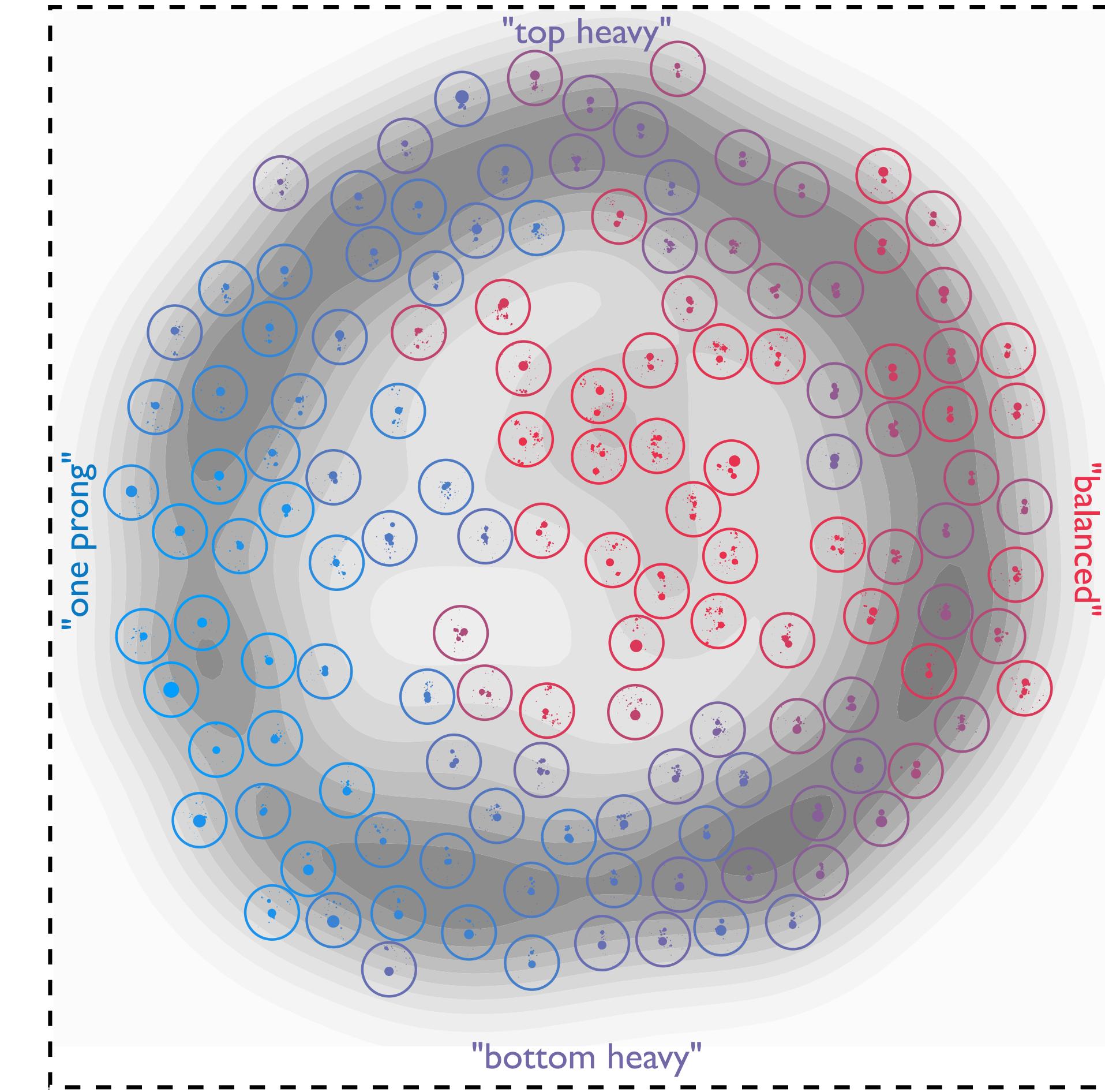
[PTK, Metodiev, Thaler, PRL 2019]

t-Distributed Stochastic Neighbor Embedding (t-SNE)
MNIST handwritten digits

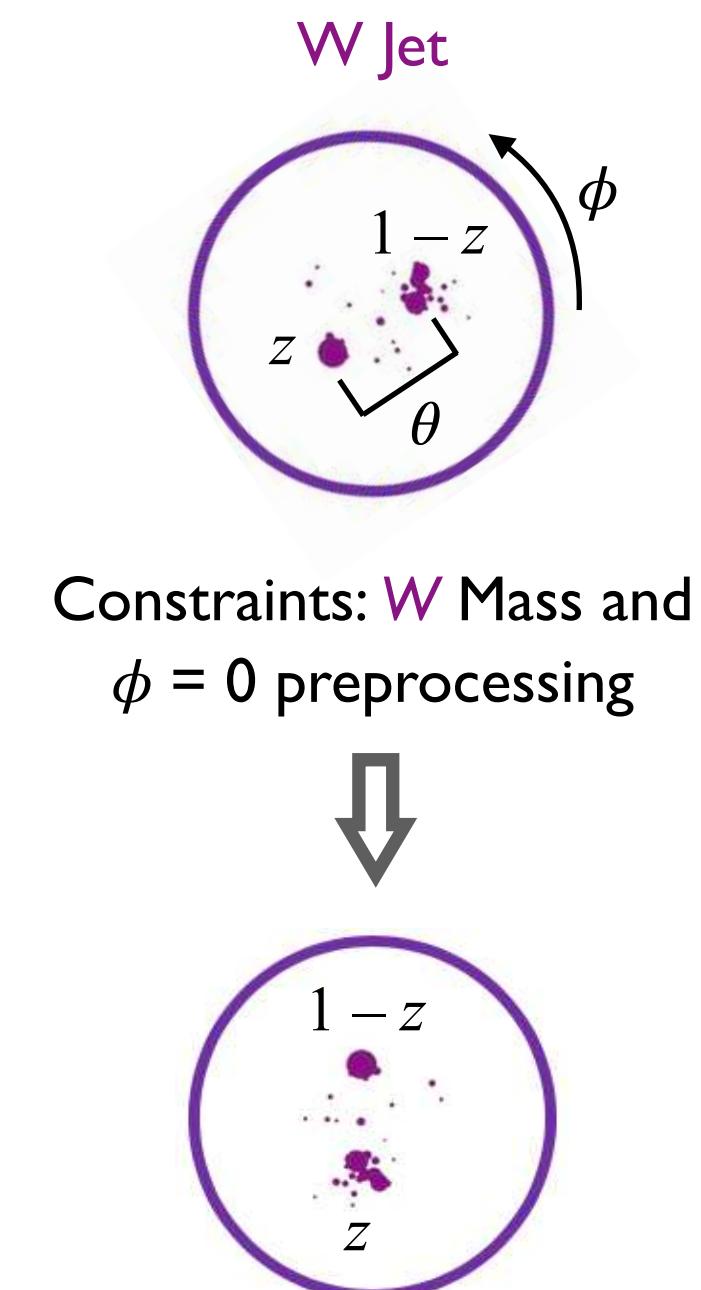


[L. van der Maaten, G. Hinton, JMLR 2008]

Geometric space of W jets

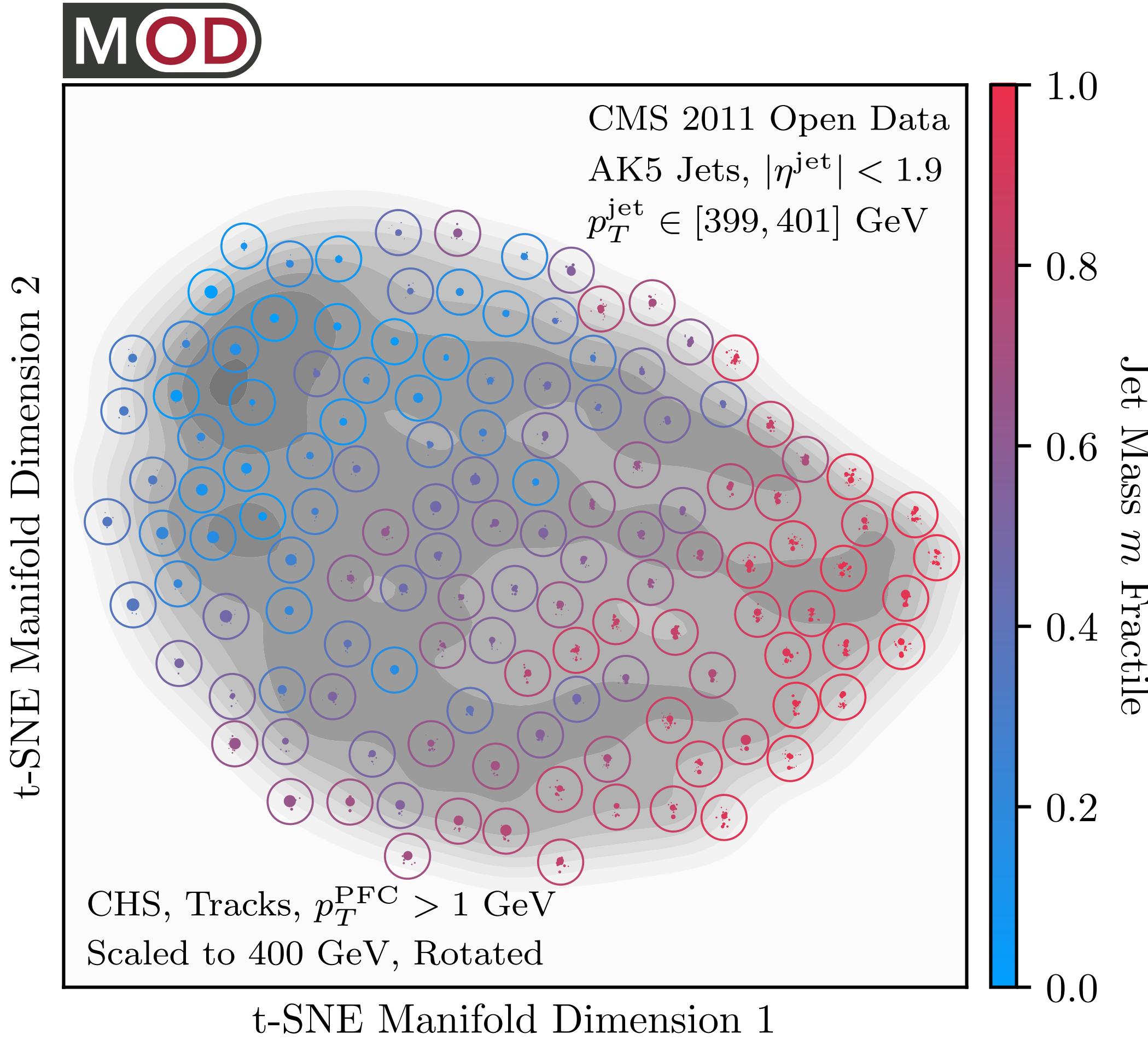


Gray contours represent the density of jets
Each circle is a particular W jet



Visualizing Geometry in CMS Open Data

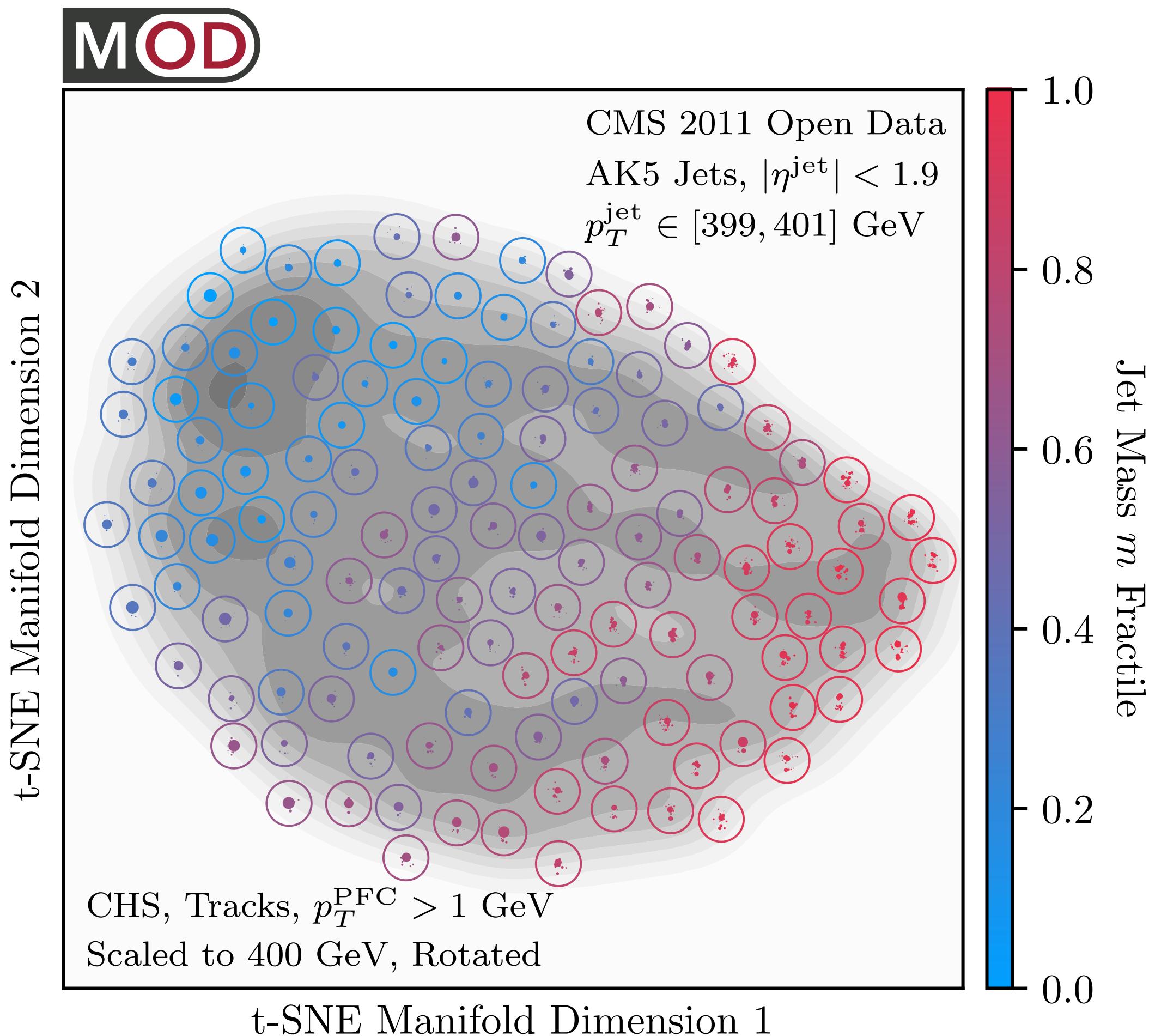
[PTK, Mastandrea, Metodiev, Naik, Thaler, [PRD 2019](#); code and datasets at [energyflow.network](#)]



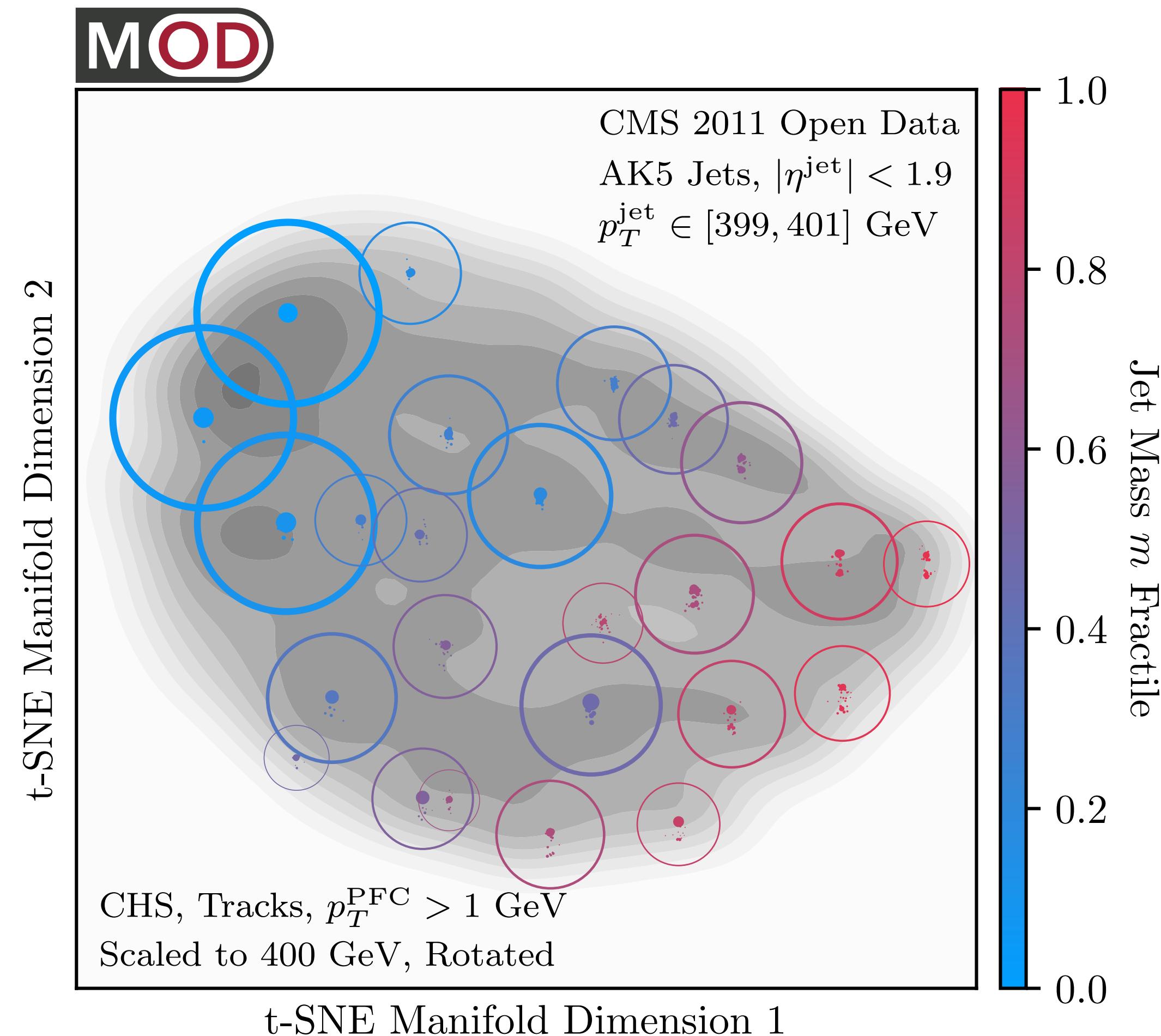
Example jets sprinkled throughout

Visualizing Geometry in CMS Open Data

[PTK, Mastandrea, Metodiev, Naik, Thaler, [PRD 2019](#); code and datasets at [energyflow.network](#)]



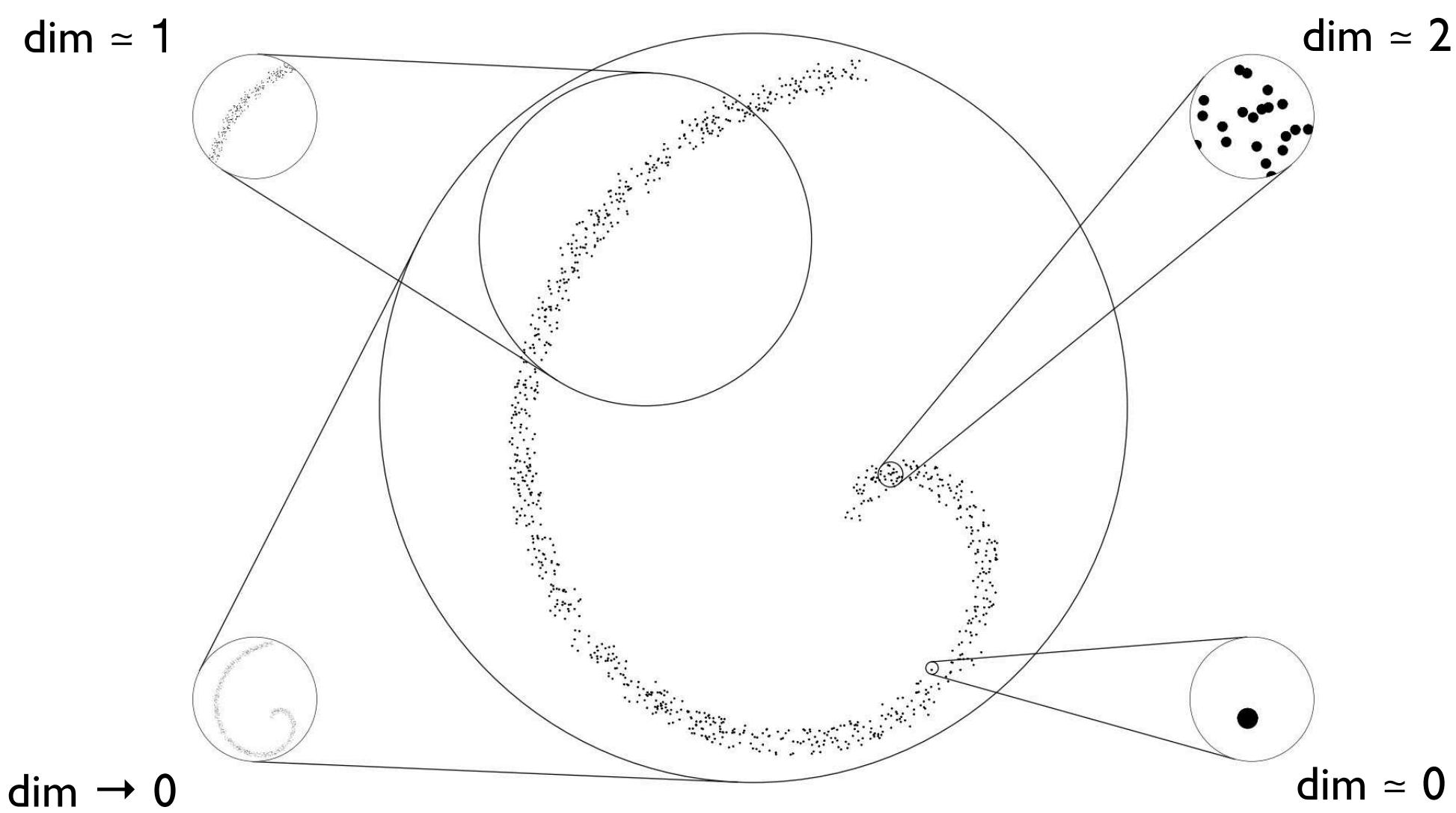
Example jets sprinkled throughout



25 most representative jets ("medoids")
Size is proportional to number of jets associated to that medoid

Quantifying Event-Space Manifolds

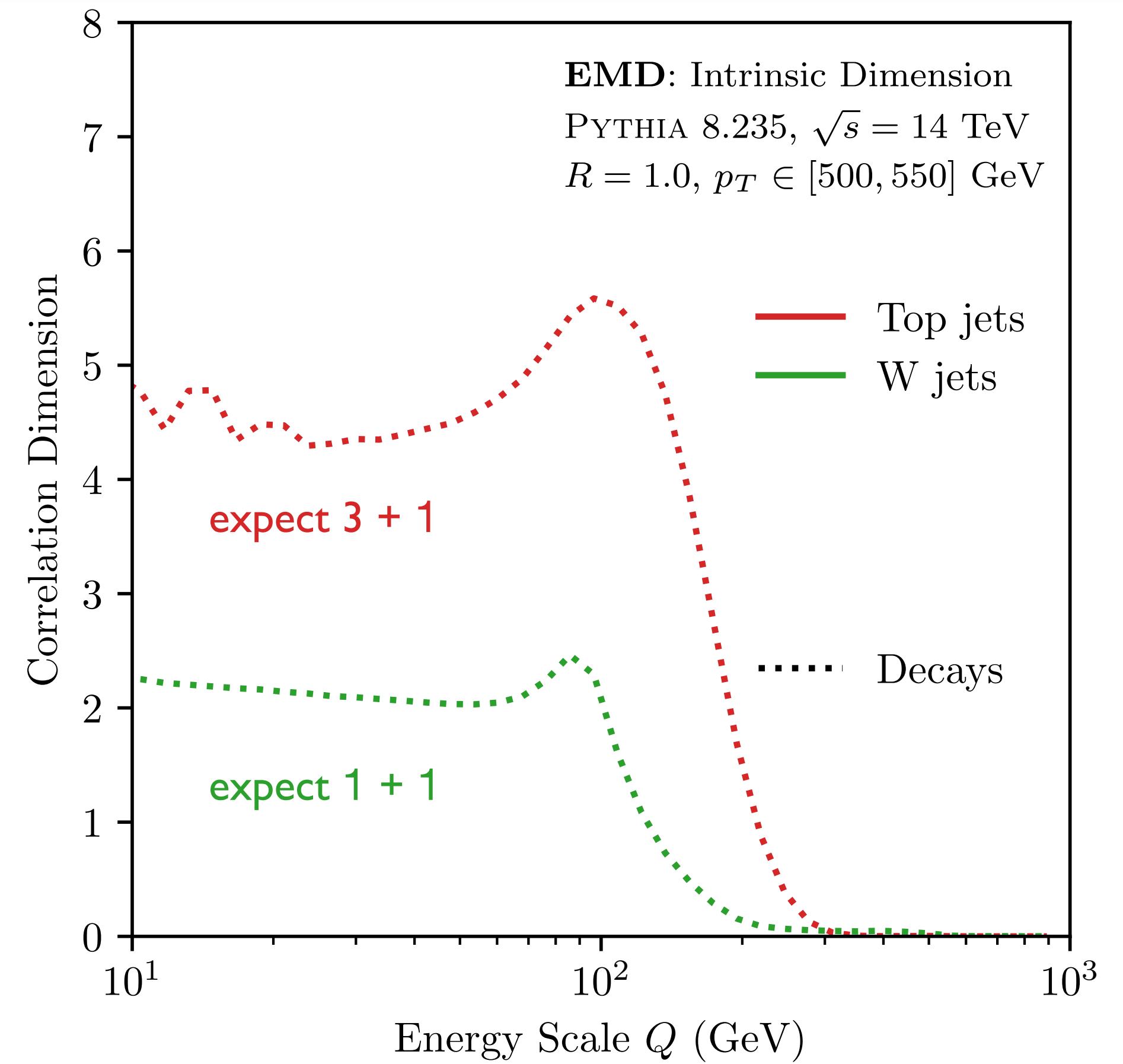
Correlation dimension: how does the # of elements within a ball of size Q change?



$$N_{\text{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q) = Q \frac{d}{dQ} \ln N_{\text{neigh.}}(Q)$$

Correlation dimension lessons:
Decays are "constant" dim. at low Q

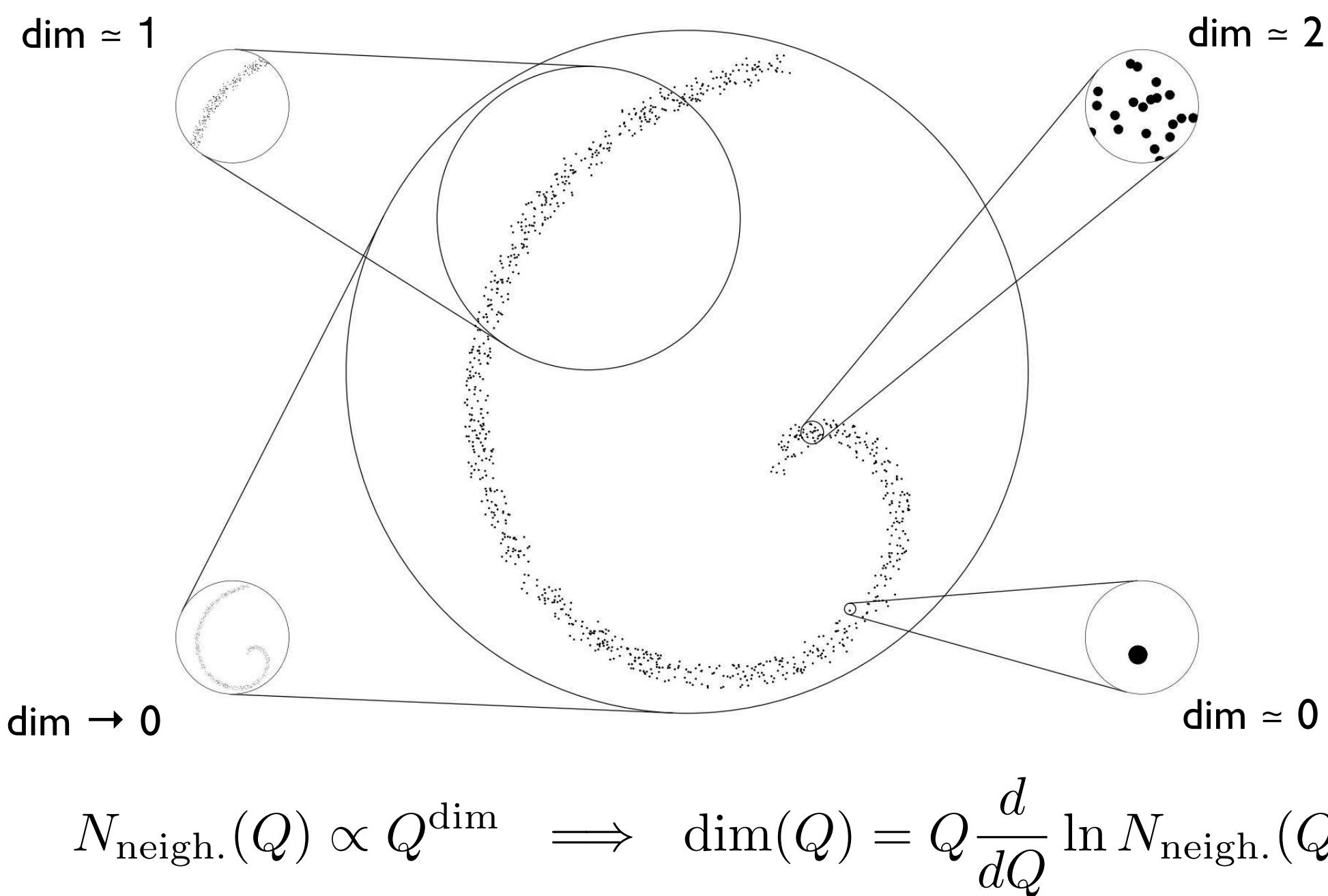
$$\dim(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$



[Grassberger, Procaccia, [PRL 1983](#); PTK, Metodiev, Thaler, [PRL 2019](#)]

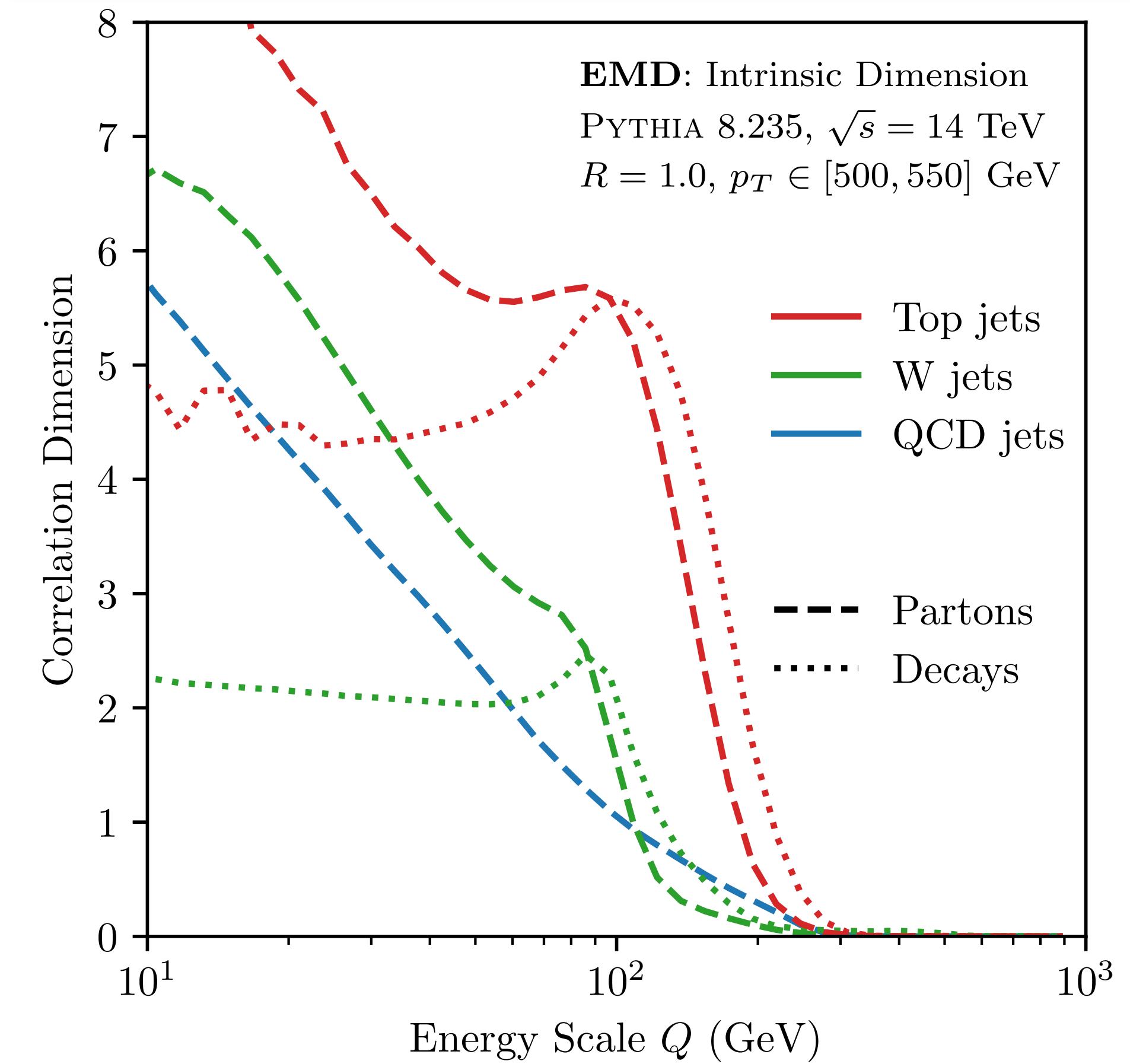
Quantifying Event-Space Manifolds

Correlation dimension: how does the # of elements within a ball of size Q change?



Correlation dimension lessons:
 Decays are "constant" dim. at low Q
 Complexity hierarchy: QCD < W < Top
 Fragmentation increases dim. at smaller scales

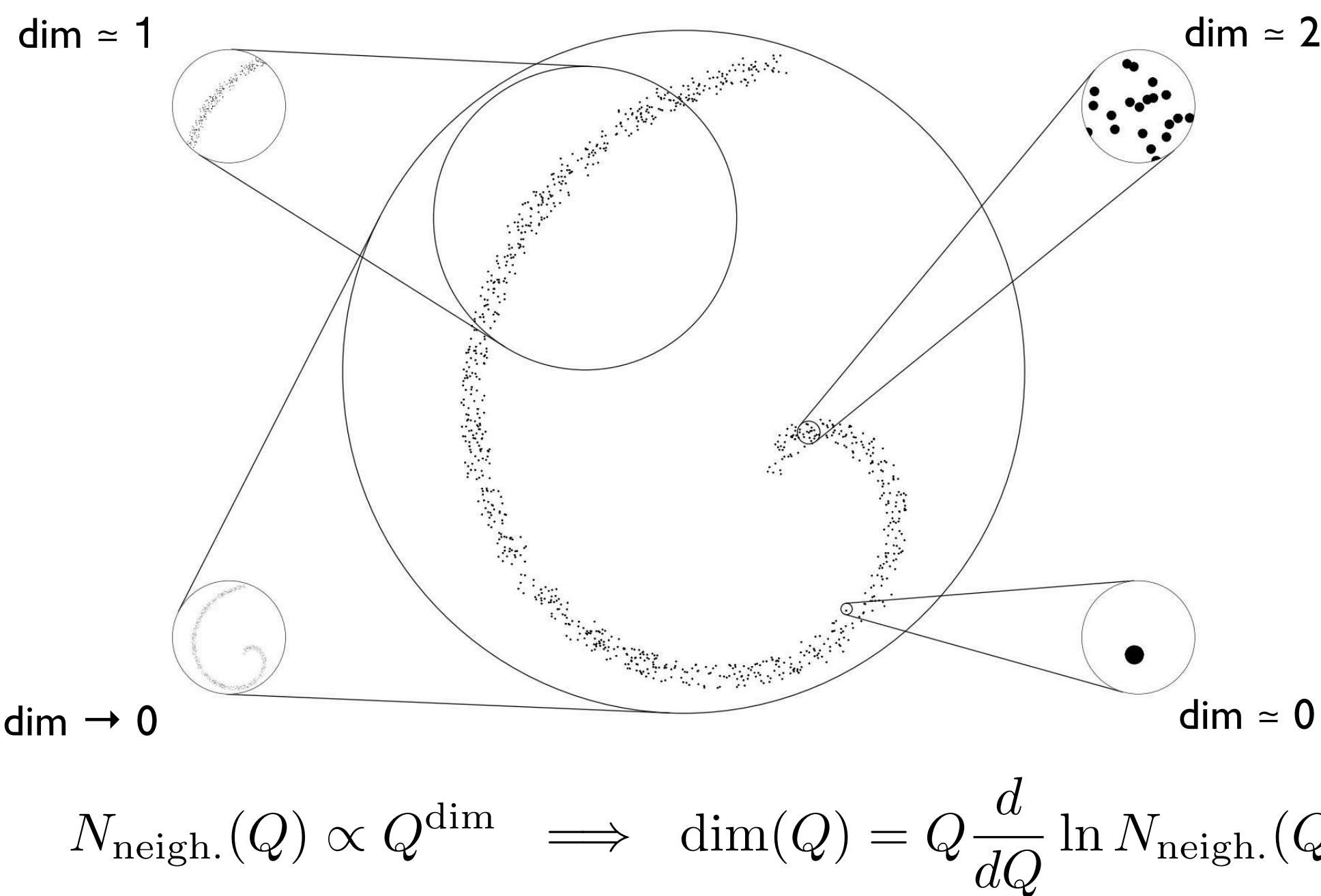
$$\text{dim}(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$



[Grassberger, Procaccia, [PRL 1983](#); PTK, Metodiev, Thaler, [PRL 2019](#)]

Quantifying Event-Space Manifolds

Correlation dimension: how does the # of elements within a ball of size Q change?



Correlation dimension lessons:

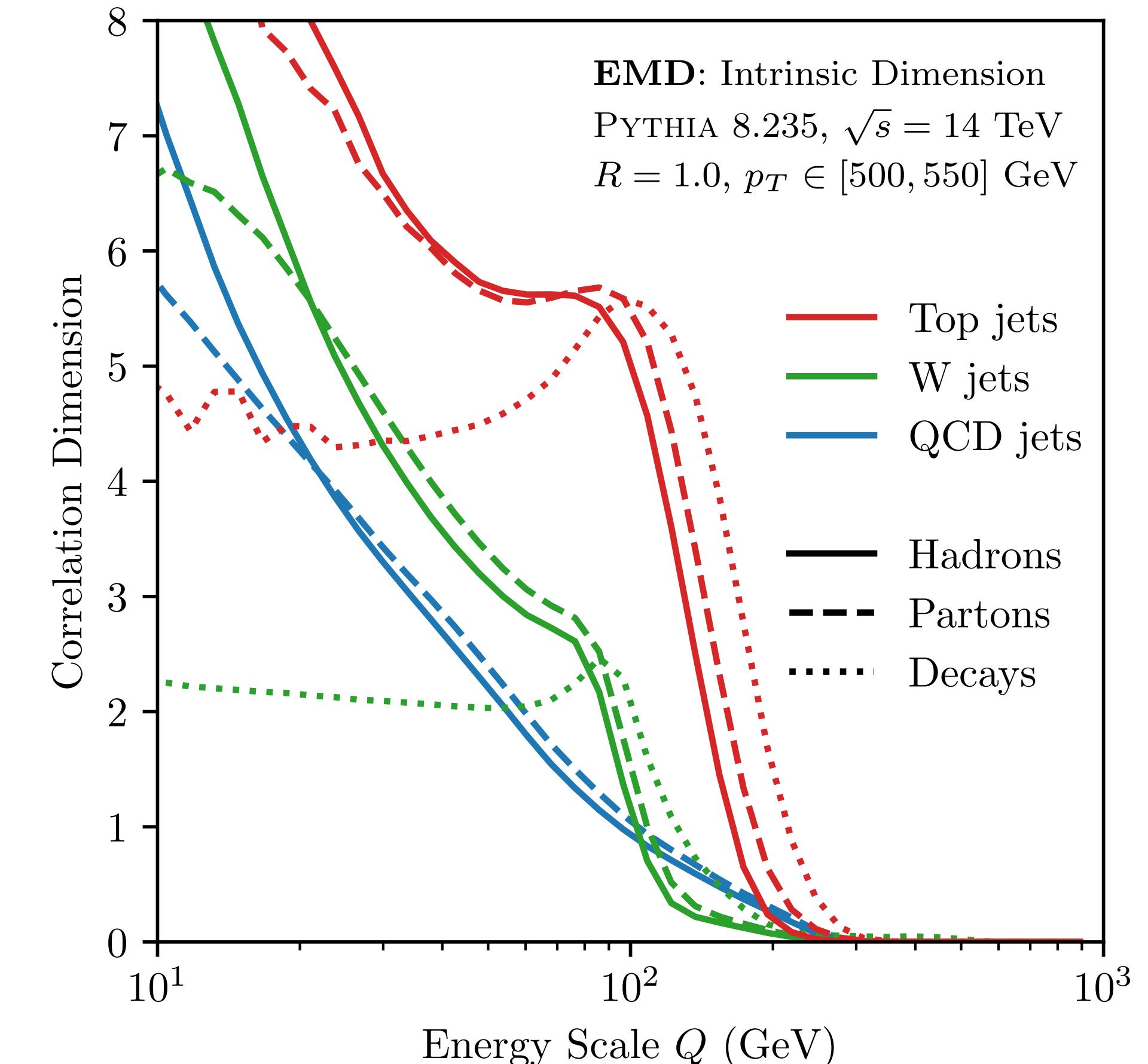
Decays are "constant" dim. at low Q

Complexity hierarchy: QCD < W < Top

Fragmentation increases dim. at smaller scales

Hadronization important around 20-30 GeV

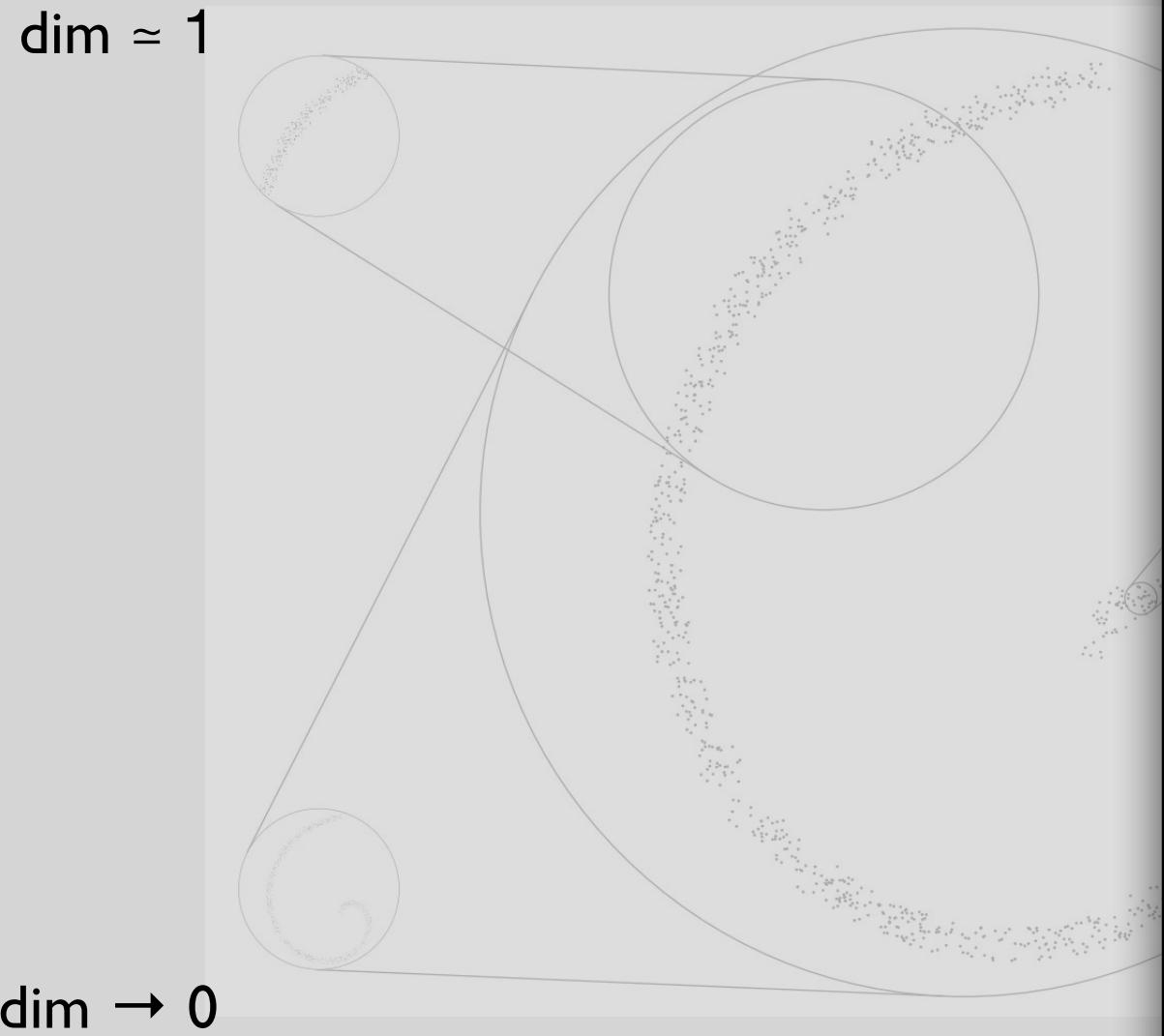
$$\text{dim}(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$



[Grassberger, Procaccia, [PRL 1983](#); PTK, Metodiev, Thaler, [PRL 2019](#)]

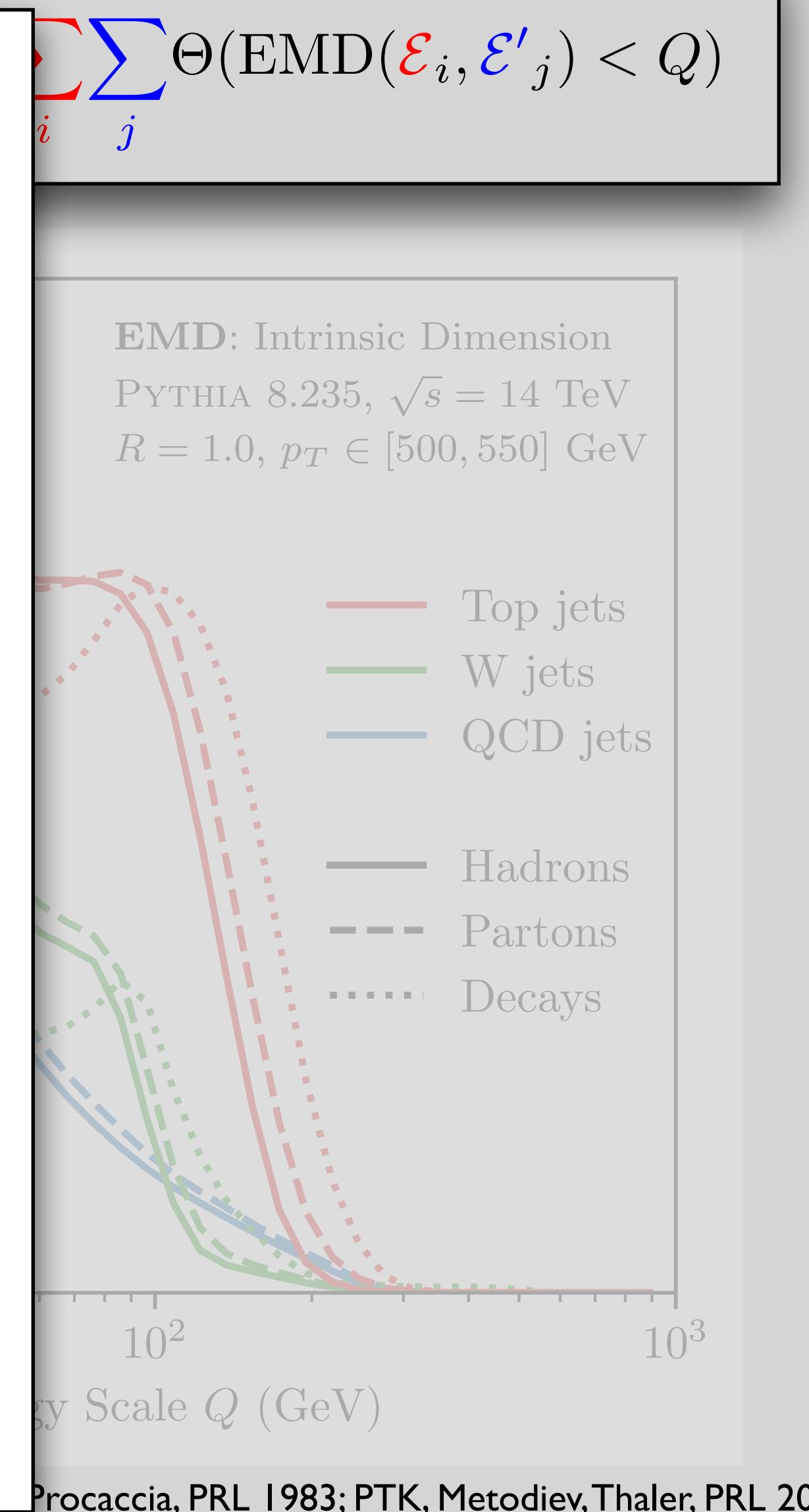
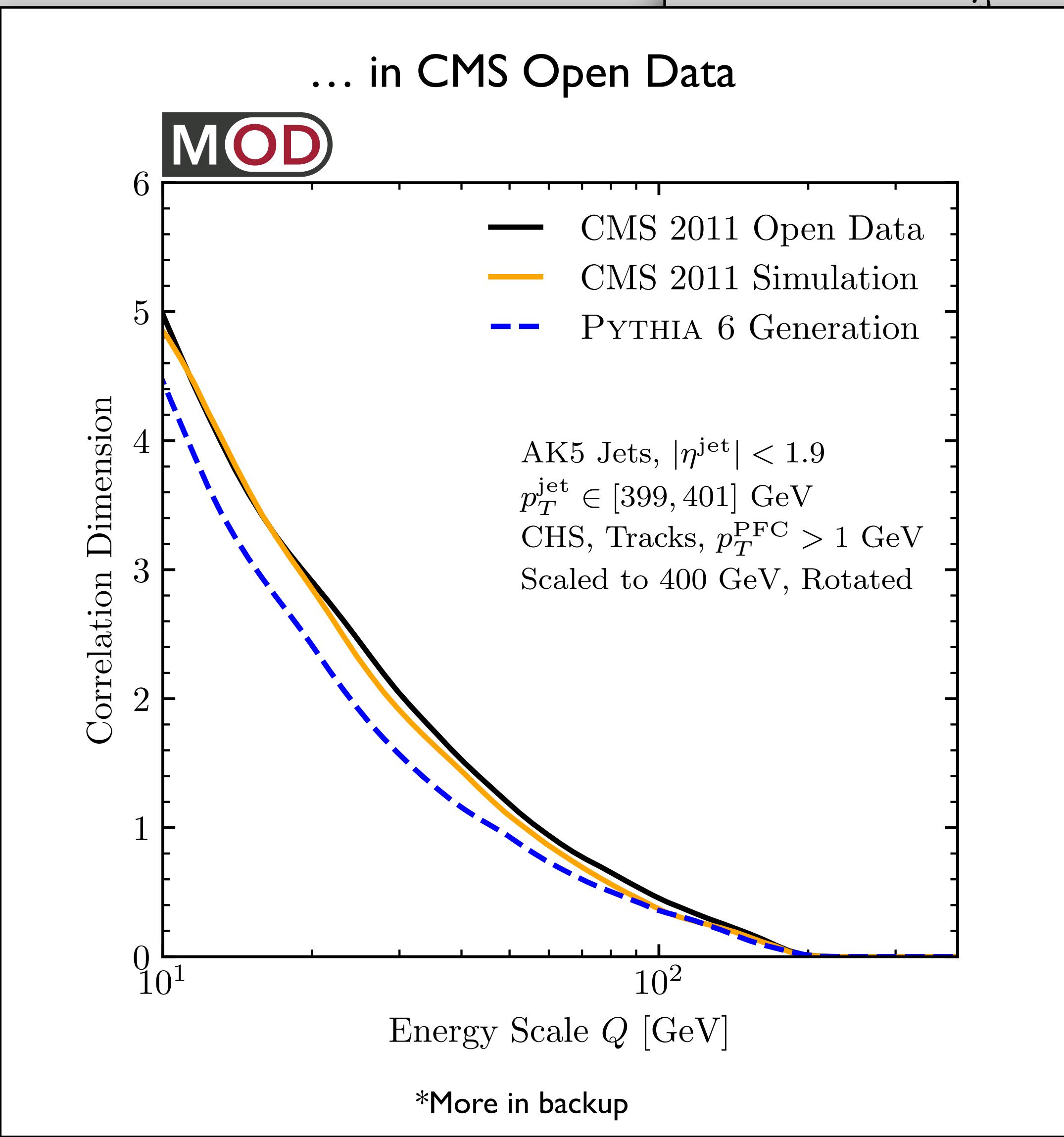
Quantifying Event-Space Manifolds

Correlation dimension: how many elements within a ball of size Q



$$N_{\text{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q)$$

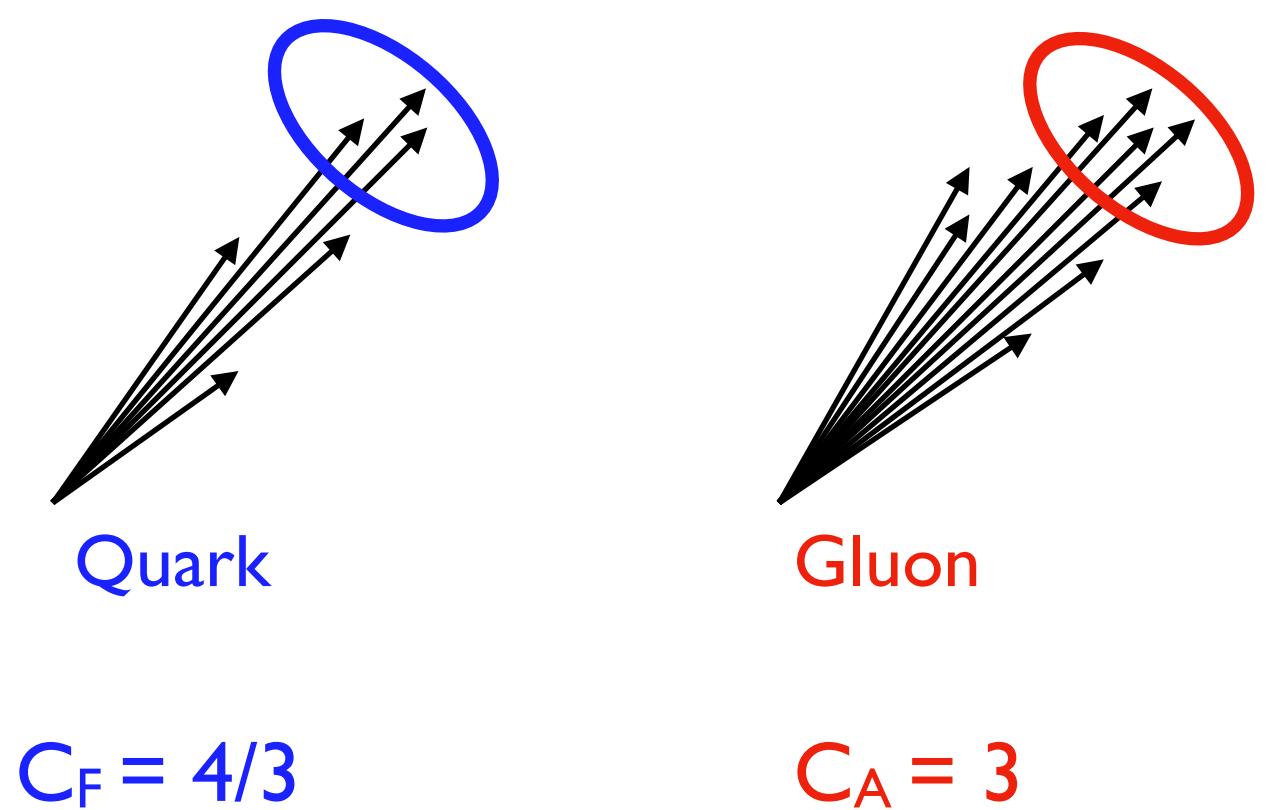
Correlation dimension
Decays are "constant" dimension
Complexity hierarchy: Q
Fragmentation increases dimension
Hadronization important



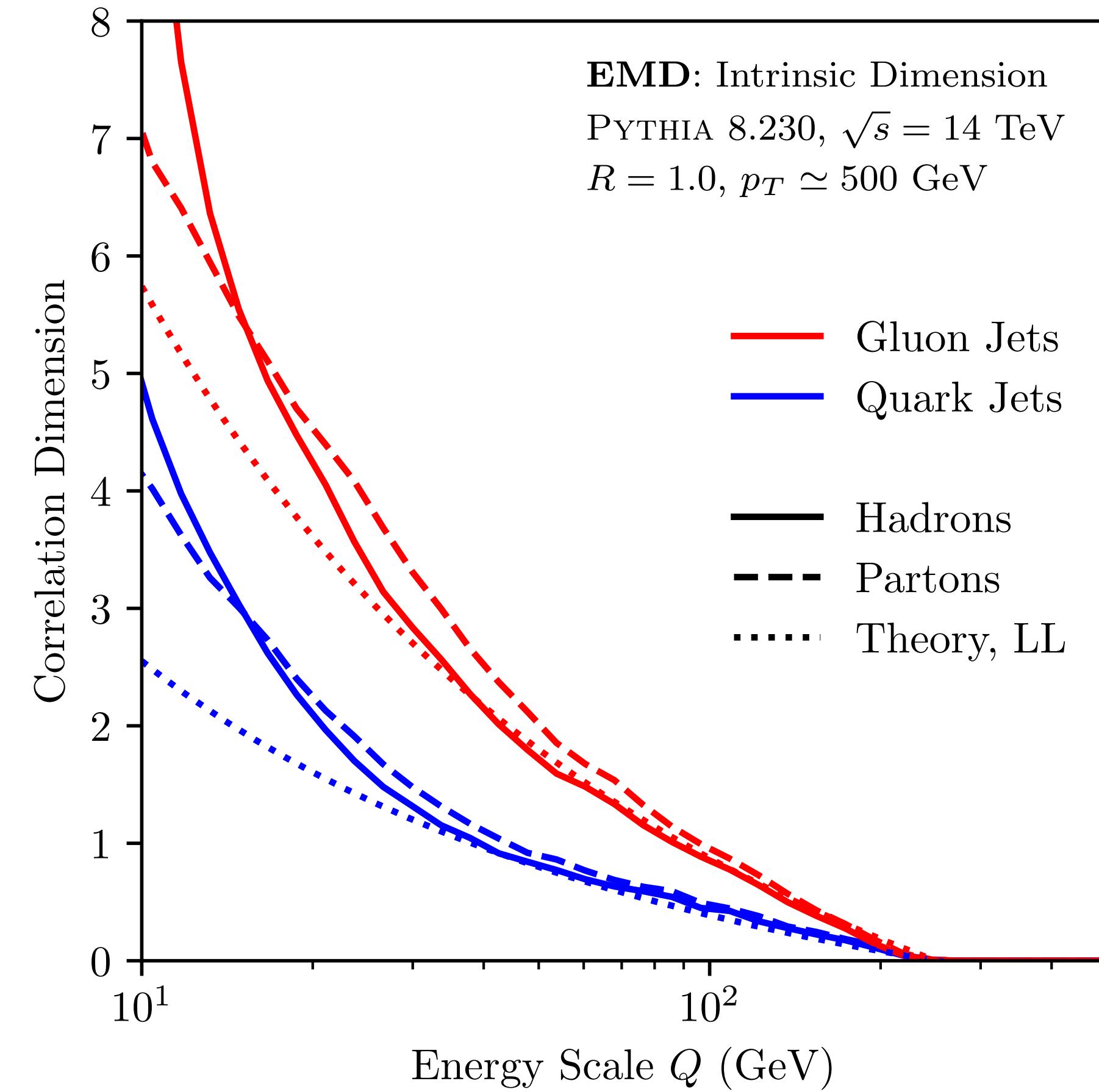
Quark and Gluon Correlation Dimensions

Leading log (single emission) calculation:

$$\dim_i(Q) \simeq -\frac{8\alpha_s}{\pi} C_i \ln \frac{Q}{p_T/2}$$



$$\dim(Q) = Q \frac{\partial}{\partial Q} \ln \sum_i \sum_j \Theta(\text{EMD}(\mathcal{E}_i, \mathcal{E}'_j) < Q)$$



[PTK, Metodiev, Thaler, to appear soon]

Defining IRC Safety Precisely

[Sterman, Weinberg, [PRL 1997](#); Sterman, [PRD 1978](#); Banfi, Salam, Zanderighi, [JHEP 2005](#)]

Infrared and collinear safety is a proxy for perturbative calculability of an observable

Defining IRC Safety Precisely

[Sterman, Weinberg, [PRL 1997](#); Sterman, [PRD 1978](#); Banfi, Salam, Zanderighi, [JHEP 2005](#)]

Infrared and collinear safety is a proxy for perturbative calculability of an observable

Exact **IRC** invariance

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \mathcal{O}(0p_0^\mu, p_1^\mu, \dots, p_M^\mu)$$

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \mathcal{O}(\lambda p_1^\mu, (1 - \lambda)p_1^\mu, \dots, p_M^\mu)$$

Guarantees observable is well-defined on **energy** flows

Allows for pathological observables, e.g. pseudo-multiplicity

Defining IRC Safety Precisely

[Sterman, Weinberg, [PRL 1997](#); Sterman, [PRD 1978](#); Banfi, Salam, Zanderighi, [JHEP 2005](#)]

Infrared and collinear safety is a proxy for perturbative calculability of an observable

Exact **IRC** invariance

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \mathcal{O}(0p_0^\mu, p_1^\mu, \dots, p_M^\mu)$$

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \mathcal{O}(\lambda p_1^\mu, (1 - \lambda)p_1^\mu, \dots, p_M^\mu)$$

Guarantees observable is well-defined on **energy** flows

Allows for pathological observables, e.g. pseudo-multiplicity

Smooth **IRC** invariance

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \lim_{\epsilon \rightarrow 0} \mathcal{O}(\epsilon p_0^\mu, p_1^\mu, \dots, p_M^\mu)$$

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \lim_{p_0^\mu \rightarrow p_1^\mu} \mathcal{O}(\lambda p_0^\mu, (1 - \lambda)p_1^\mu, \dots, p_M^\mu)$$

Eliminates common observables with hard boundaries

Defining IRC Safety Precisely

[Sterman, Weinberg, [PRL 1997](#); Sterman, [PRD 1978](#); Banfi, Salam, Zanderighi, [JHEP 2005](#)]

Infrared and collinear safety is a proxy for perturbative calculability of an observable

Exact **IRC** invariance

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \mathcal{O}(0p_0^\mu, p_1^\mu, \dots, p_M^\mu)$$

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \mathcal{O}(\lambda p_1^\mu, (1 - \lambda)p_1^\mu, \dots, p_M^\mu)$$

Guarantees observable is well-defined on **energy** flows

Allows for pathological observables, e.g. pseudo-multiplicity

Smooth **IRC** invariance

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \lim_{\epsilon \rightarrow 0} \mathcal{O}(\epsilon p_0^\mu, p_1^\mu, \dots, p_M^\mu)$$

$$\mathcal{O}(p_1^\mu, \dots, p_M^\mu) = \lim_{p_0^\mu \rightarrow p_1^\mu} \mathcal{O}(\lambda p_0^\mu, (1 - \lambda)p_1^\mu, \dots, p_M^\mu)$$

Eliminates common observables with hard boundaries

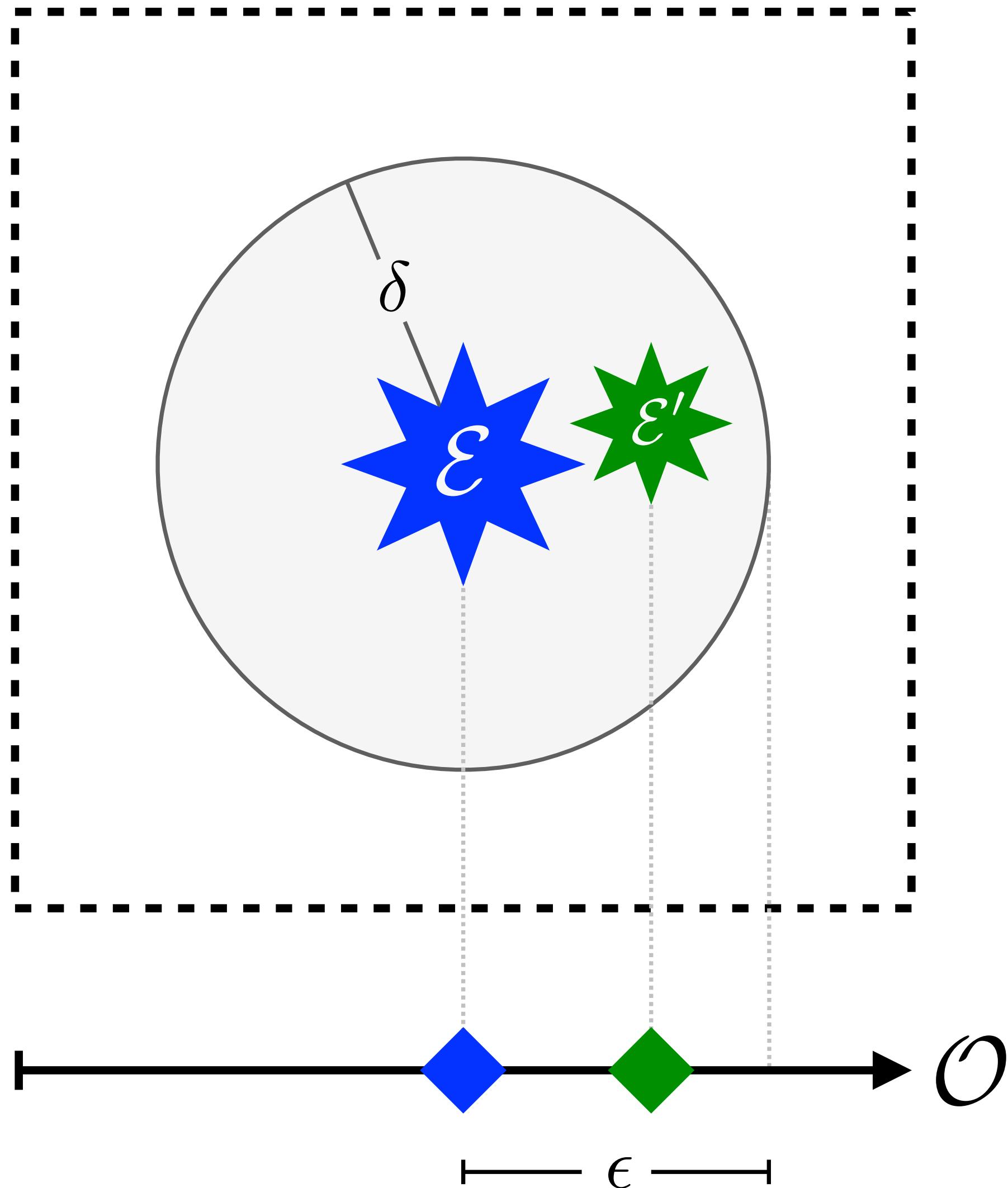
All Observables	Comments
Multiplicity ($\sum_i 1$)	IR unsafe and C unsafe
Momentum Dispersion [65] ($\sum_i E_i^2$)	IR safe but C unsafe
Sphericity Tensor [66] ($\sum_i p_i^\mu p_i^\nu$)	IR safe but C unsafe
Number of Non-Zero Calorimeter Deposits	C safe but IR unsafe

Defined on Energy Flows	
Pseudo-Multiplicity ($\min\{N \mid \mathcal{T}_N = 0\}$)	Robust to exact IR or C emissions

Infrared & Collinear Safe	
Jet Energy ($\sum_i E_i$)	
Heavy Jet Mass [67]	
Soft-Dropped Jet Mass [38, 68]	
Calorimeter Activity [69] (N_{95})	
Disc. at jet boundary	
Disc. at hemisphere boundary	
Disc. at grooming threshold	
Disc. at cell boundary	

More EMD Geometry – Continuity in the Space of Events

[PTK, Metodiev, Thaler, 2004.04 | 59]



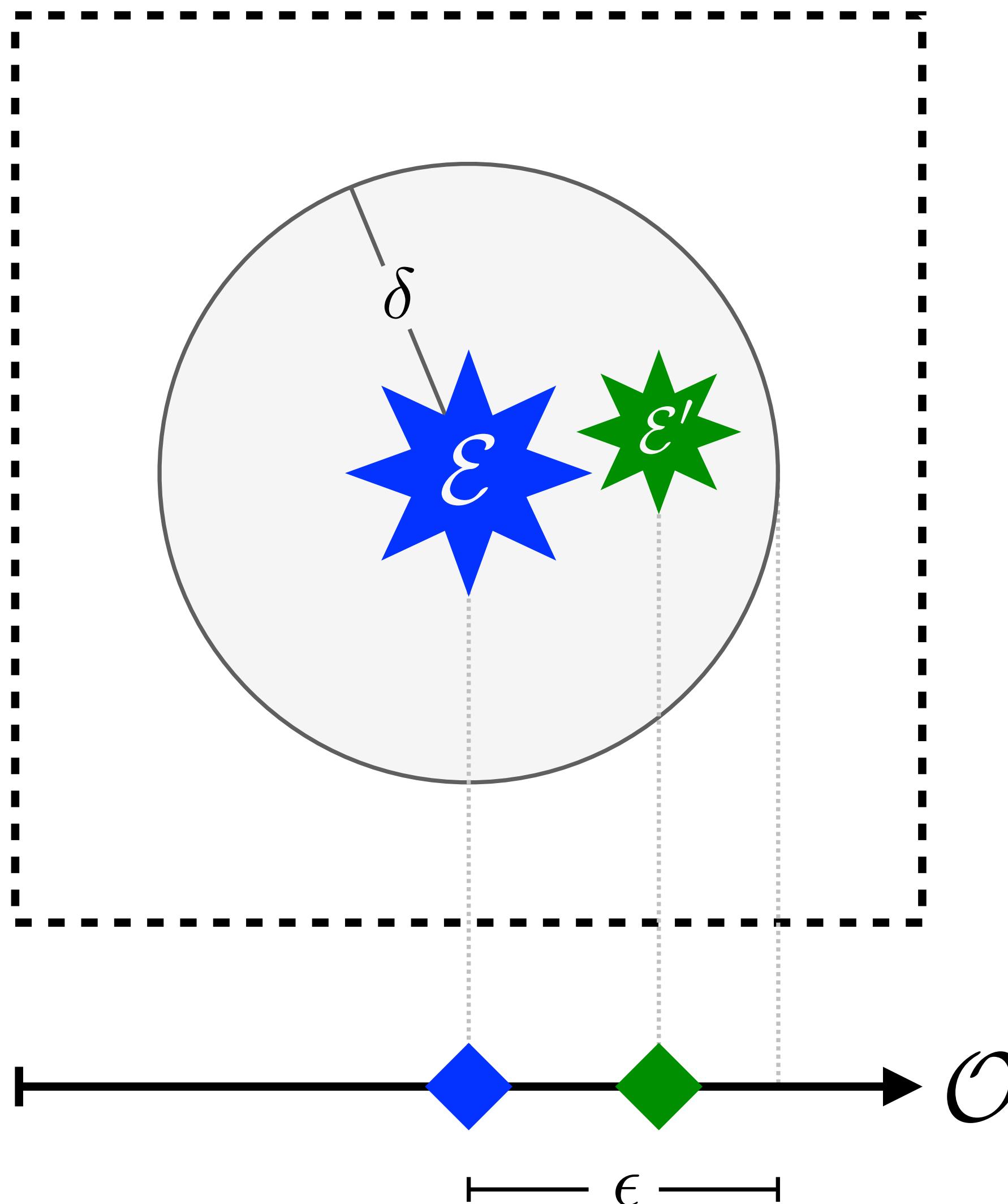
Classic $\epsilon - \delta$ definition of continuity in a metric space

An observable \mathcal{O} is **EMD continuous** at an event \mathcal{E} if, for any $\epsilon > 0$, there exists a $\delta > 0$ such that for all events \mathcal{E}' :

$$\text{EMD}(\mathcal{E}, \mathcal{E}') < \delta \implies |\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')| < \epsilon.$$

More EMD Geometry – Continuity in the Space of Events

[PTK, Metodiev, Thaler, 2004.04159]



Classic $\epsilon - \delta$ definition of continuity in a metric space

An observable \mathcal{O} is **EMD continuous** at an event \mathcal{E} if, for any $\epsilon > 0$, there exists a $\delta > 0$ such that for all events \mathcal{E}' :

$$\text{EMD}(\mathcal{E}, \mathcal{E}') < \delta \implies |\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')| < \epsilon.$$

Towards a geometric definition of **IRC Safety**

IRC Safety = EMD Continuity*

*on all but a negligible set[‡] of events

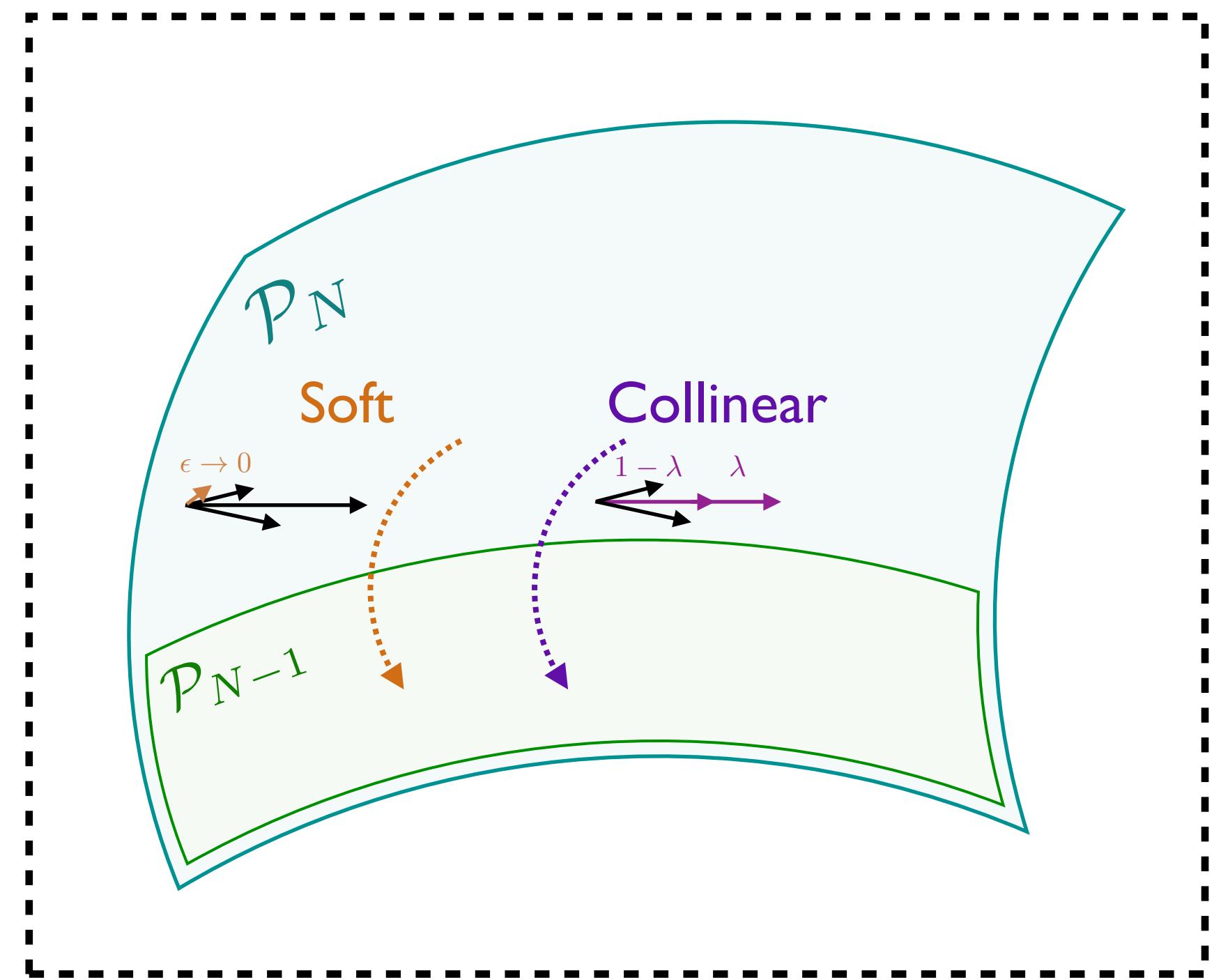
[‡]a negligible set is one that contains no positive-radius EMD-ball

⋮

Perturbation Theory in the Space of Events

[PTK, Metodiev, Thaler, 2004.04.159]

Infrared singularities of massless
gauge theories appear on each P_N



Perturbation Theory in the Space of Events

[PTK, Metodiev, Thaler, 2004.04.159]

Sudakov safety

[Larkoski, Thaler, JHEP 2014; Larkoski, Marzani, Thaler, PRD 2015]

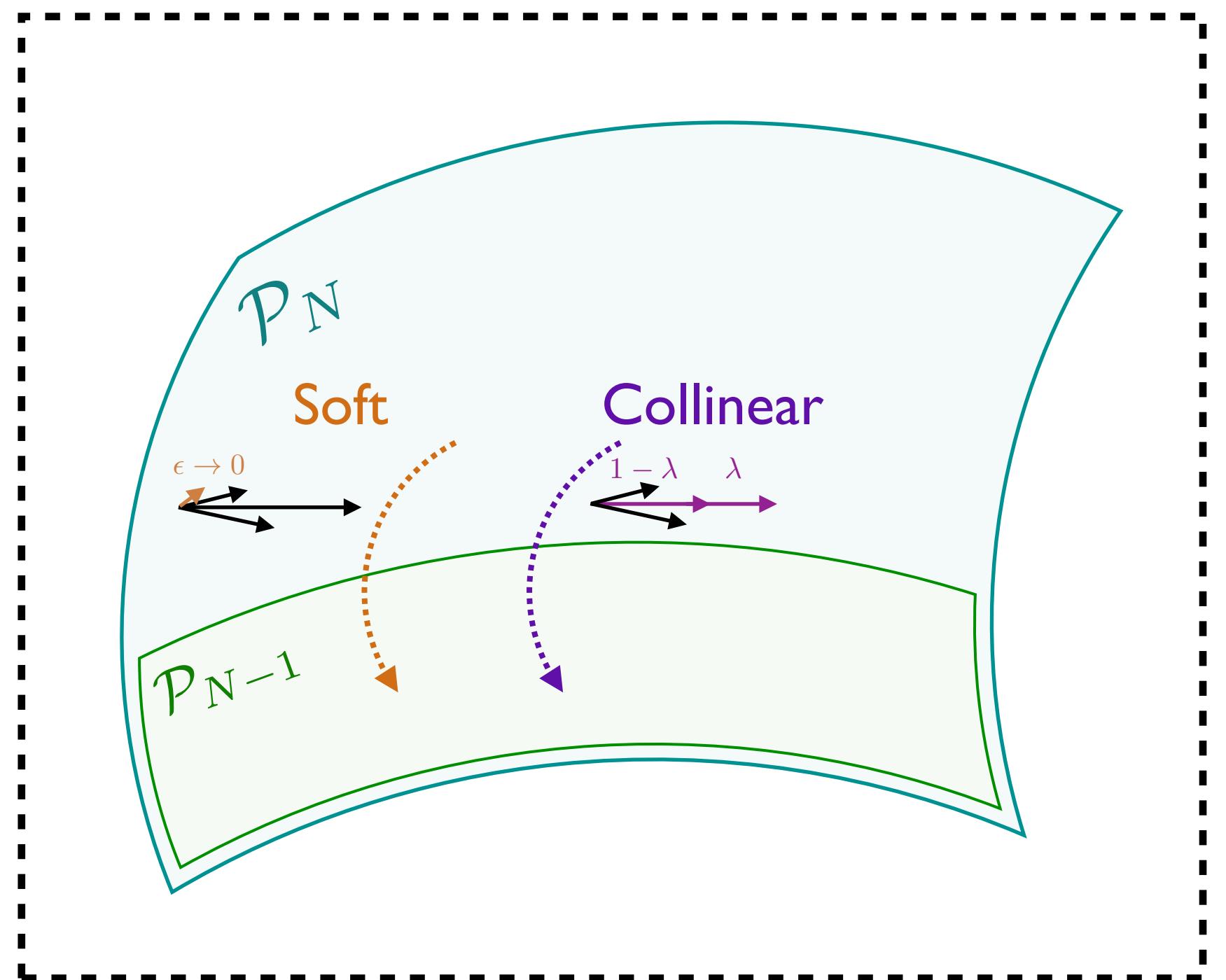
Some observables have discontinuities on P_N for some N

A resummed **IRC-safe companion** can mitigate the divergences

$$p(\mathcal{O}_{\text{Sudakov}}) = \int d\mathcal{O}_{\text{Comp.}} p(\mathcal{O}_{\text{Sudakov}} | \mathcal{O}_{\text{Comp.}}) p(\mathcal{O}_{\text{Comp.}})$$

Event geometry suggests N -(sub)jettiness as universal companion

Infrared singularities of massless gauge theories appear on each P_N



Perturbation Theory in the Space of Events

[PTK, Metodiev, Thaler, 2004.04.159]

Sudakov safety

[Larkoski, Thaler, JHEP 2014; Larkoski, Marzani, Thaler, PRD 2015]

Some observables have discontinuities on P_N for some N

A resummed **IRC-safe companion** can mitigate the divergences

$$p(\mathcal{O}_{\text{Sudakov}}) = \int d\mathcal{O}_{\text{Comp.}} p(\mathcal{O}_{\text{Sudakov}} | \mathcal{O}_{\text{Comp.}}) p(\mathcal{O}_{\text{Comp.}})$$

Event geometry suggests N -(sub)jettiness as universal companion

Fixed-order calculability

[Sterman, PRD 1979; Banfi, Salam, Zanderighi, JHEP 2005]

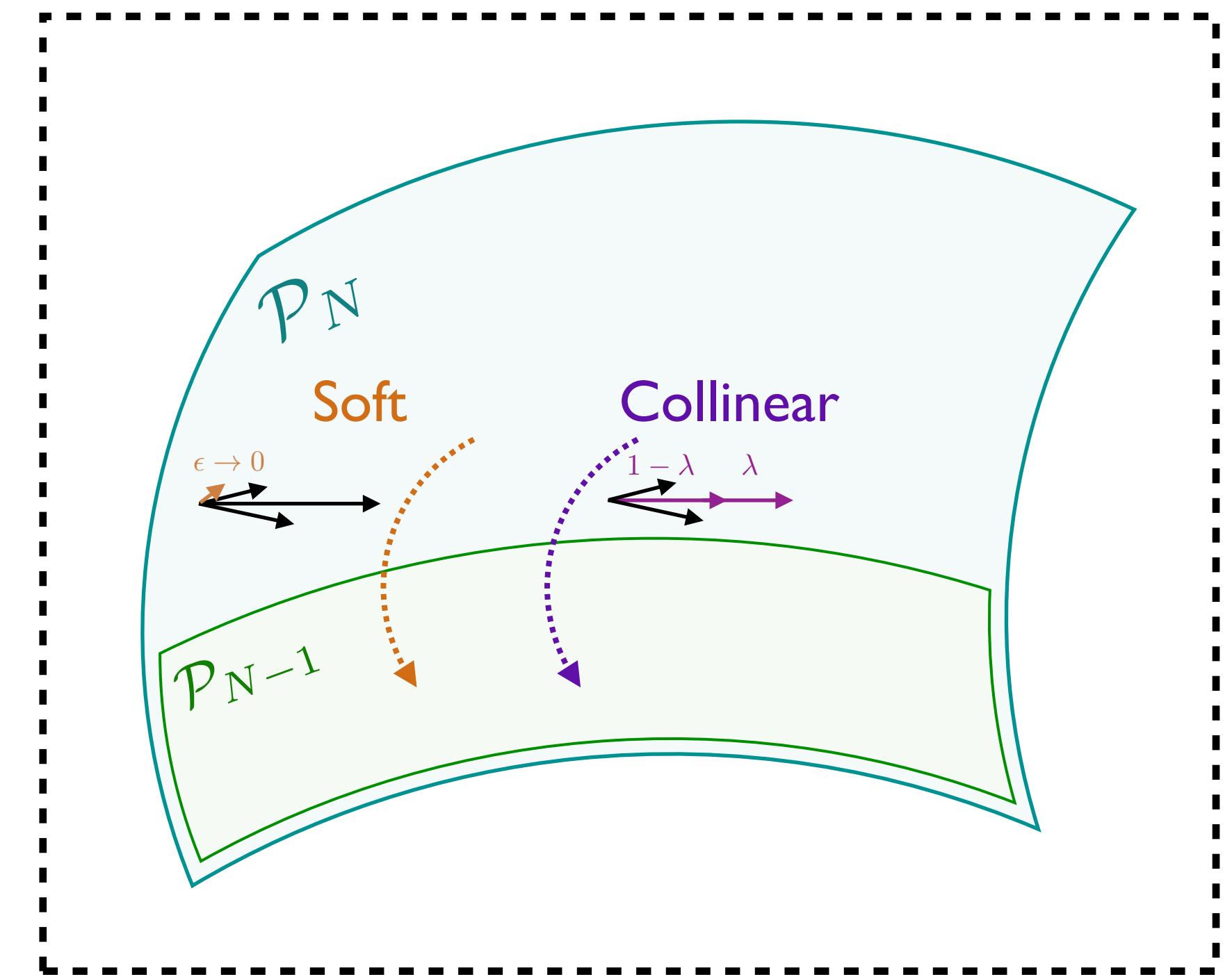
Is a statement of integrability on each P_N

EMD continuity must be upgraded to EMD-Hölder continuity on each P_N

$$\lim_{\mathcal{E} \rightarrow \mathcal{E}'} \frac{\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')}{\text{EMD}(\mathcal{E}, \mathcal{E}')^c} = 0, \quad c > 0$$

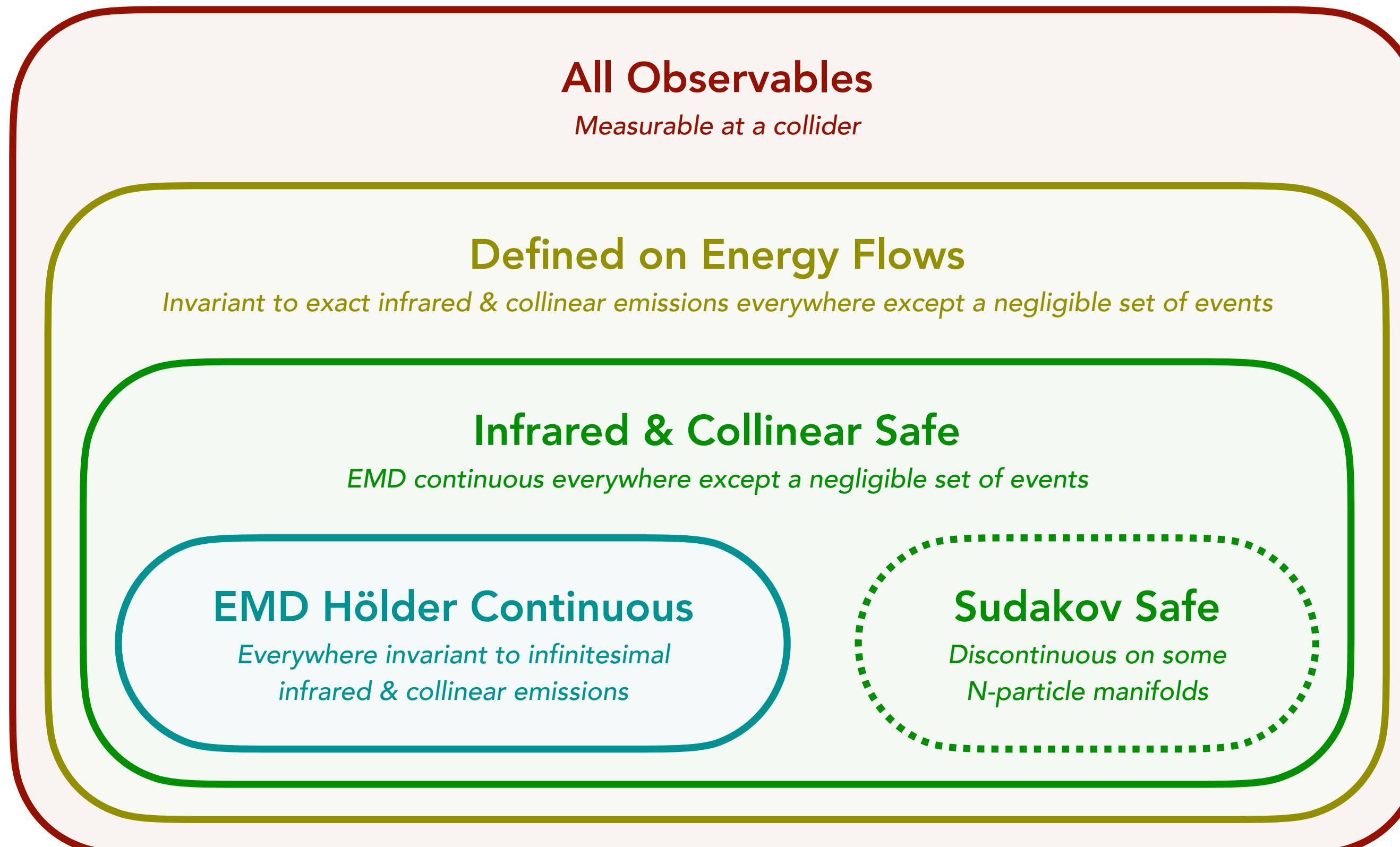
Example: $V(\mathcal{E}) = \mathcal{T}_2(\mathcal{E}) \left(1 + \frac{1}{\ln E(\mathcal{E})/\mathcal{T}_3(\mathcal{E})} \right)$ is EMD continuous but not EMD Hölder continuous (it is Sudakov safe)

Infrared singularities of massless gauge theories appear on each P_N



Hierarchy of IRC Safety Definitions

[PTK, Metodiev, Thaler, 2004.04.159]



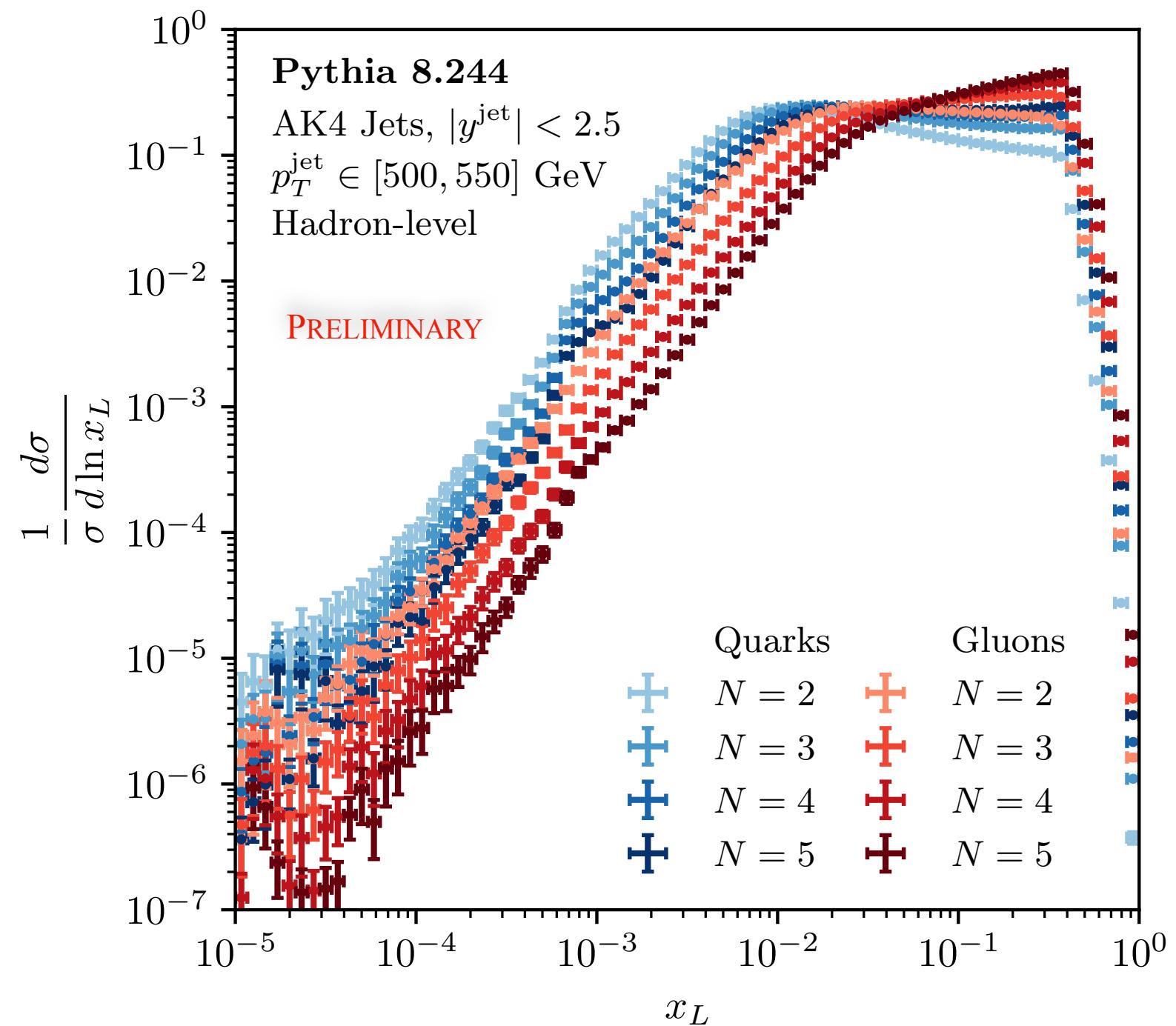
All Observables	Comments
Multiplicity ($\sum_i 1$)	IR unsafe and C unsafe
Momentum Dispersion [65] ($\sum_i E_i^2$)	IR safe but C unsafe
Sphericity Tensor [66] ($\sum_i p_i^\mu p_i^\nu$)	IR safe but C unsafe
Number of Non-Zero Calorimeter Deposits	C safe but IR unsafe
Defined on Energy Flows	
Pseudo-Multiplicity ($\min\{N \mid \mathcal{T}_N = 0\}$)	Robust to exact IR or C emissions
Infrared & Collinear Safe	
Jet Energy ($\sum_i E_i$)	Disc. at jet boundary
Heavy Jet Mass [67]	Disc. at hemisphere boundary
Soft-Dropped Jet Mass [38, 68]	Disc. at grooming threshold
Calorimeter Activity [69] (N_{95})	Disc. at cell boundary
Sudakov Safe	
Groomed Momentum Fraction [39] (z_g)	Disc. on 1-particle manifold
Jet Angularity Ratios [37]	Disc. on 1-particle manifold
N -subjettiness Ratios [47, 48] (τ_{N+1}/τ_N)	Disc. on N -particle manifold
V parameter [36] (Eq. (2.11))	Hölder disc. on 3-particle manifold
EMD Hölder Continuous Everywhere	
Thrust [40, 41]	
Sphericity [42]	
Angularities [70]	
N -jettiness [44] (\mathcal{T}_N)	
C parameter [71–74]	Resummation beneficial at $C = \frac{3}{4}$
Linear Sphericity [72] ($\sum_i E_i n_i^\mu n_i^\nu$)	
Energy Correlators [36, 75–77]	
Energy Flow Polynomials [15, 17]	

Energy-Energy Correlators – Projection to Longest Side

[PTK, Moult, Thaler, Zhu, to appear soon]

Integrate out shape dependence but keep overall size dependence

$$\frac{d\Sigma^{[N]}}{dx_L} = \sum_n \sum_{1 \leq i_1 \leq \dots \leq i_N \leq n} \int d\sigma_n \frac{E_{i_1} \cdots E_{i_N}}{Q^N} \delta(x_L - \max_{1 \leq j < k \leq N} \{\theta_{i_j i_k}\})$$

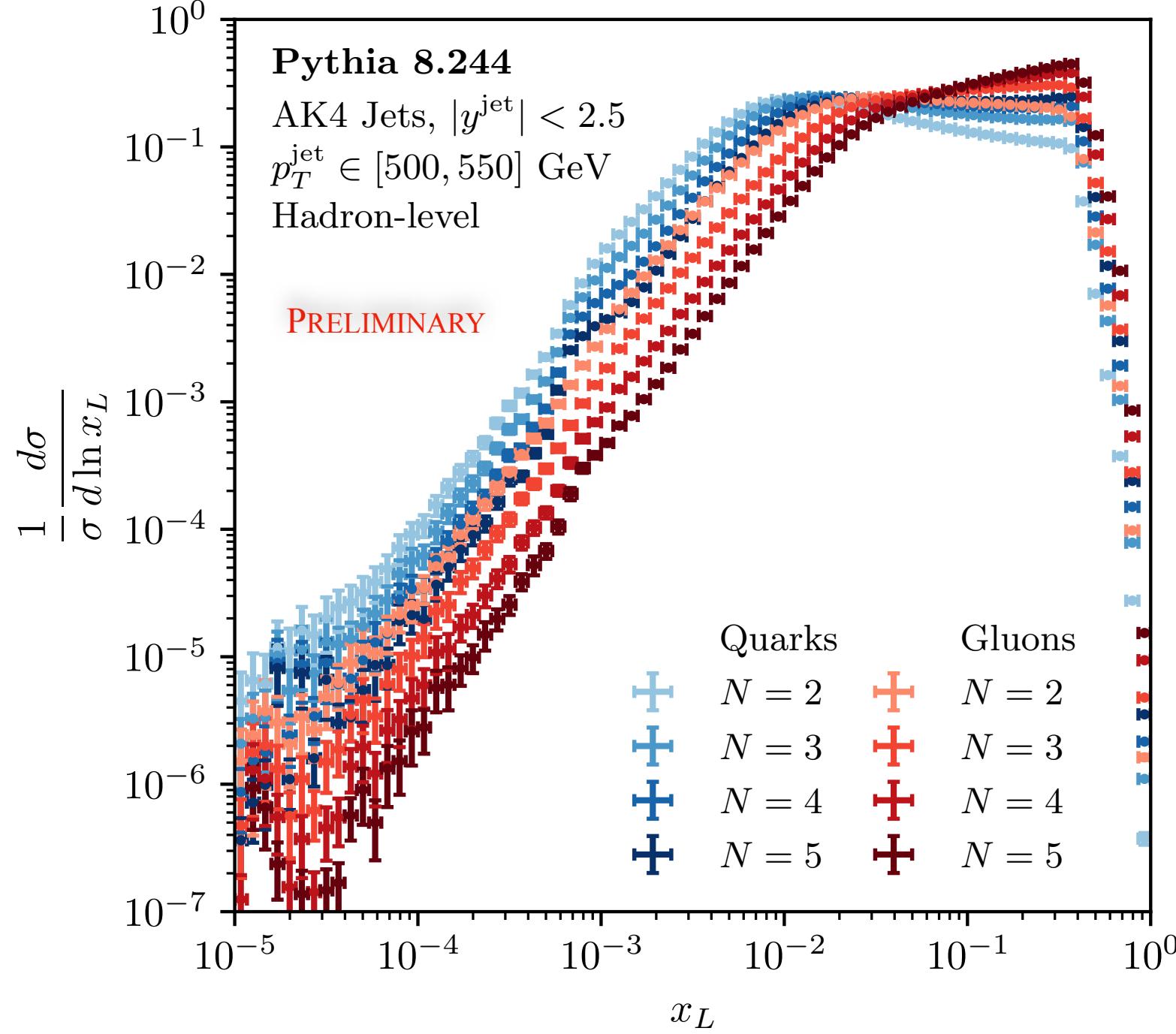


Energy-Energy Correlators – Projection to Longest Side

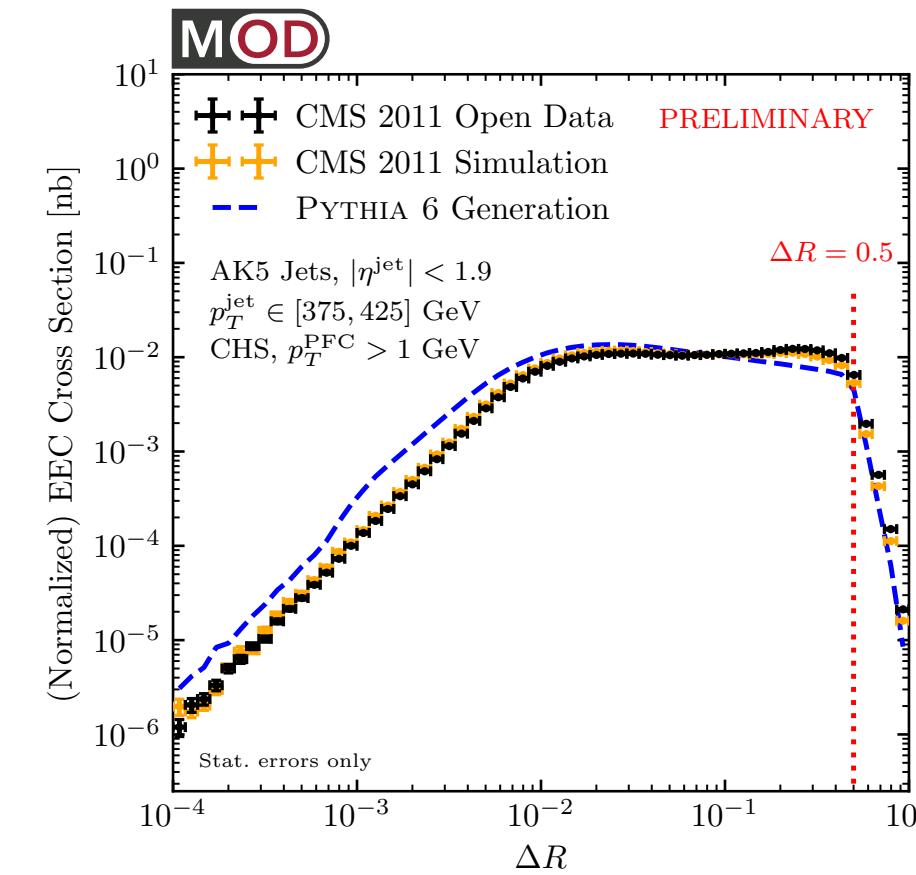
[PTK, Moult, Thaler, Zhu, to appear soon]

Integrate out shape dependence but keep overall size dependence

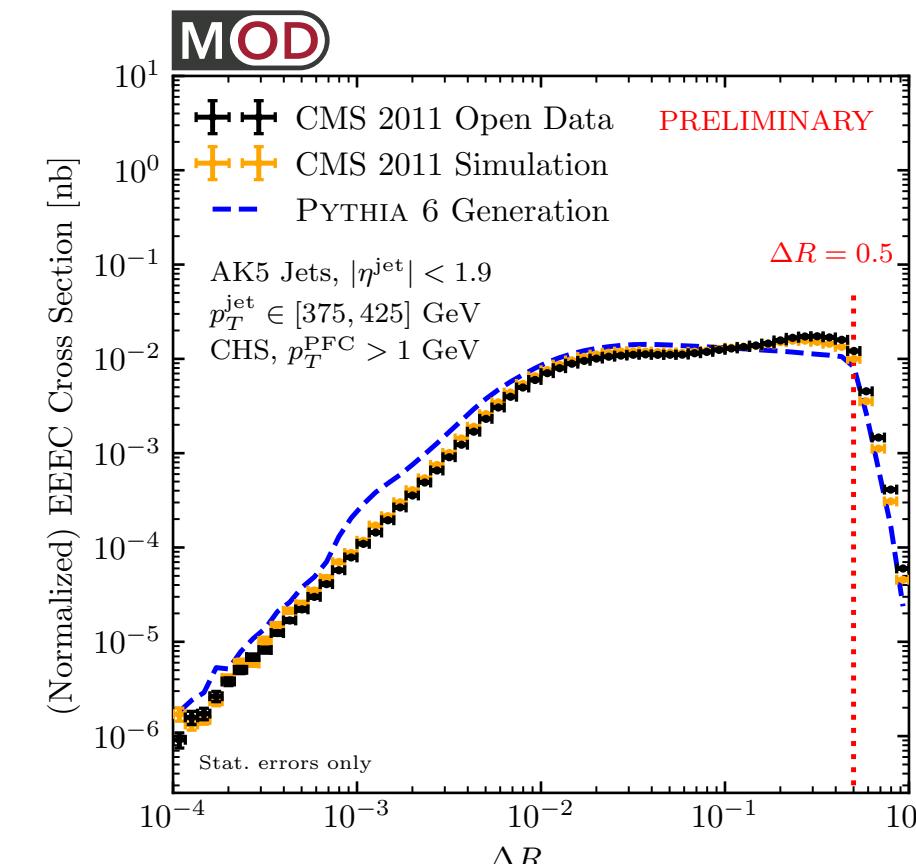
$$\frac{d\Sigma[N]}{dx_L} = \sum_n \sum_{1 \leq i_1 \leq \dots \leq i_N \leq n} \int d\sigma_n \frac{E_{i_1} \cdots E_{i_N}}{Q^N} \delta(x_L - \max_{1 \leq j < k \leq N} \{\theta_{i_j i_k}\})$$



$N = 2$



$N = 3$



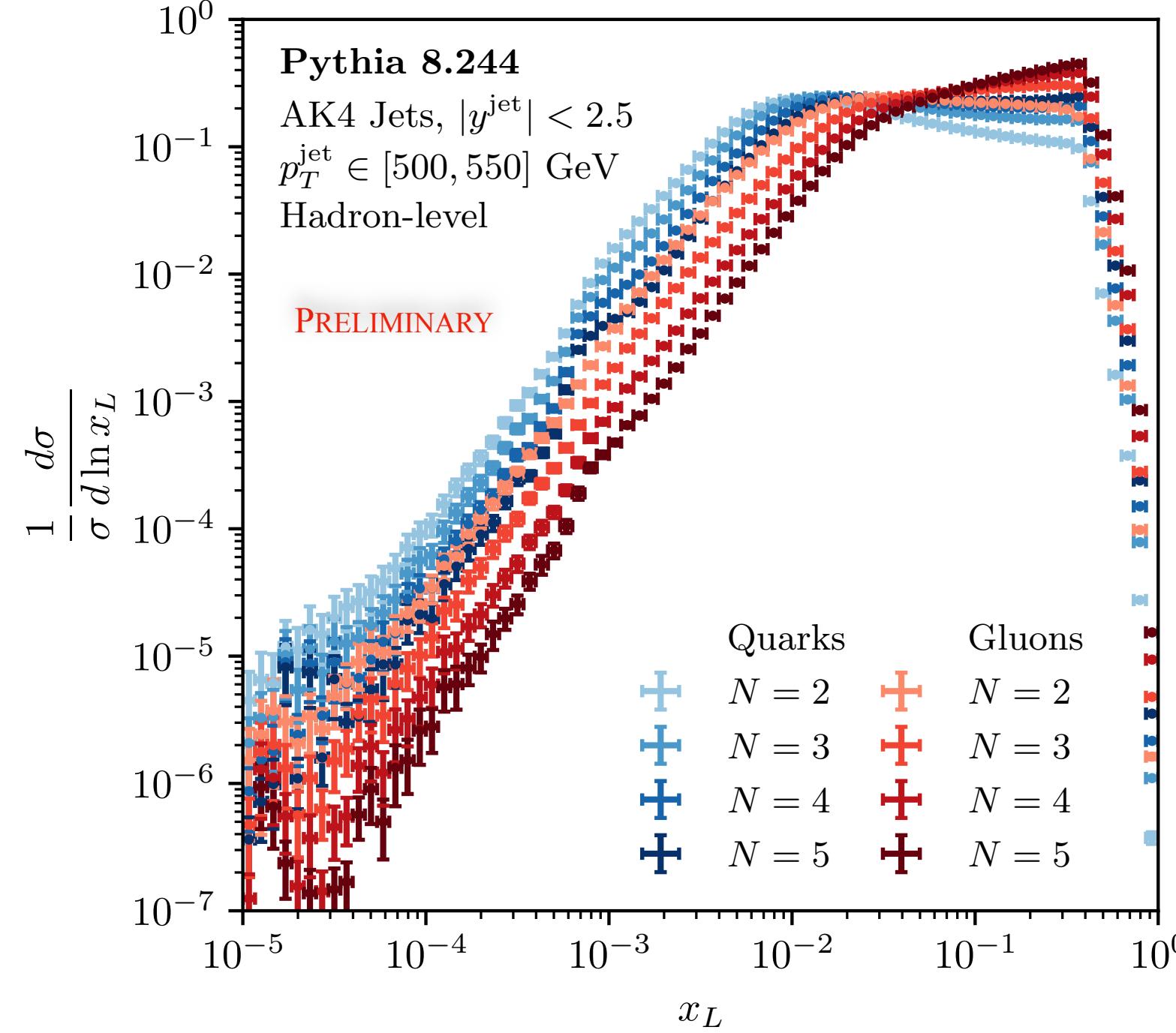
Energy-Energy Correlators – Projection to Longest Side

[PTK, Moult, Thaler, Zhu, to appear soon]

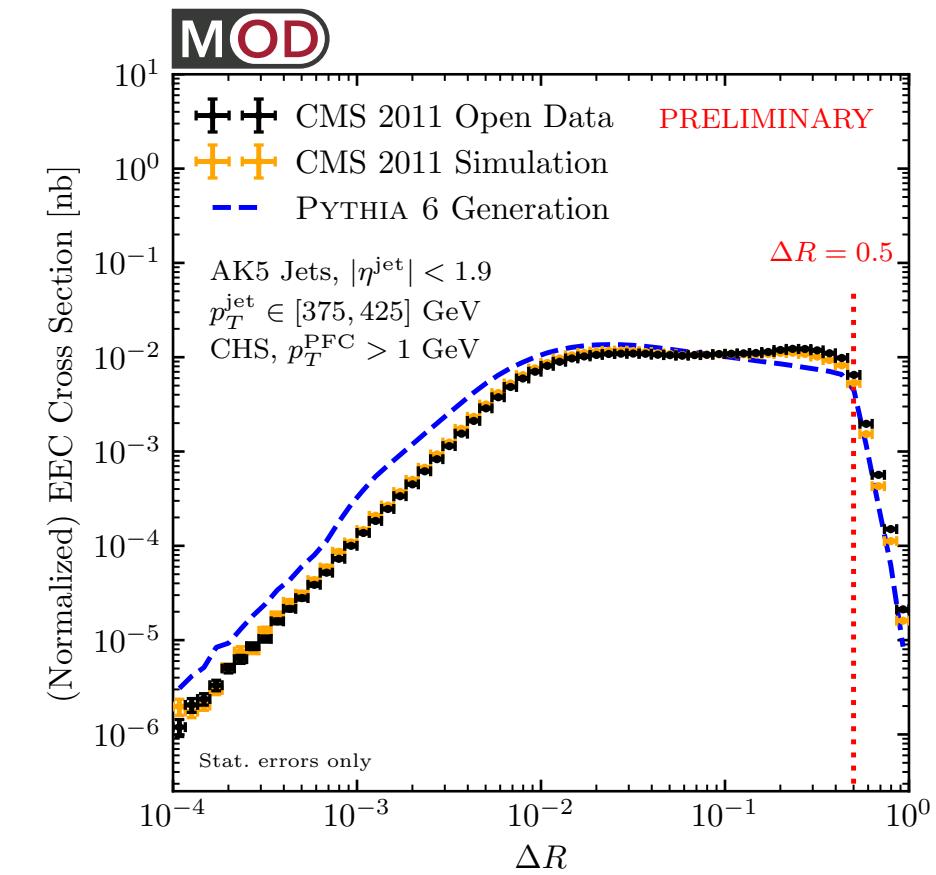
Integrate out shape dependence but keep overall size dependence

$$\frac{d\Sigma[N]}{dx_L} = \sum_n \sum_{1 \leq i_1 \leq \dots \leq i_N \leq n} \int d\sigma_n \frac{E_{i_1} \cdots E_{i_N}}{Q^N} \delta(x_L - \max_{1 \leq j < k \leq N} \{\theta_{i_j i_k}\})$$

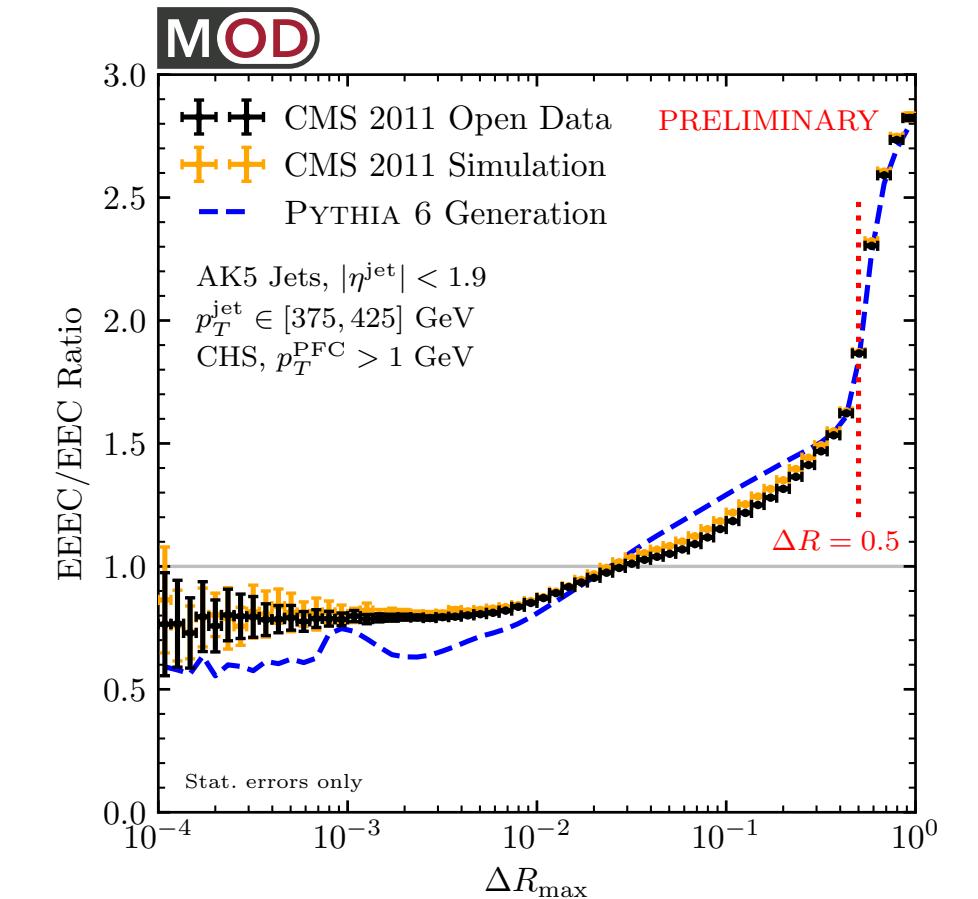
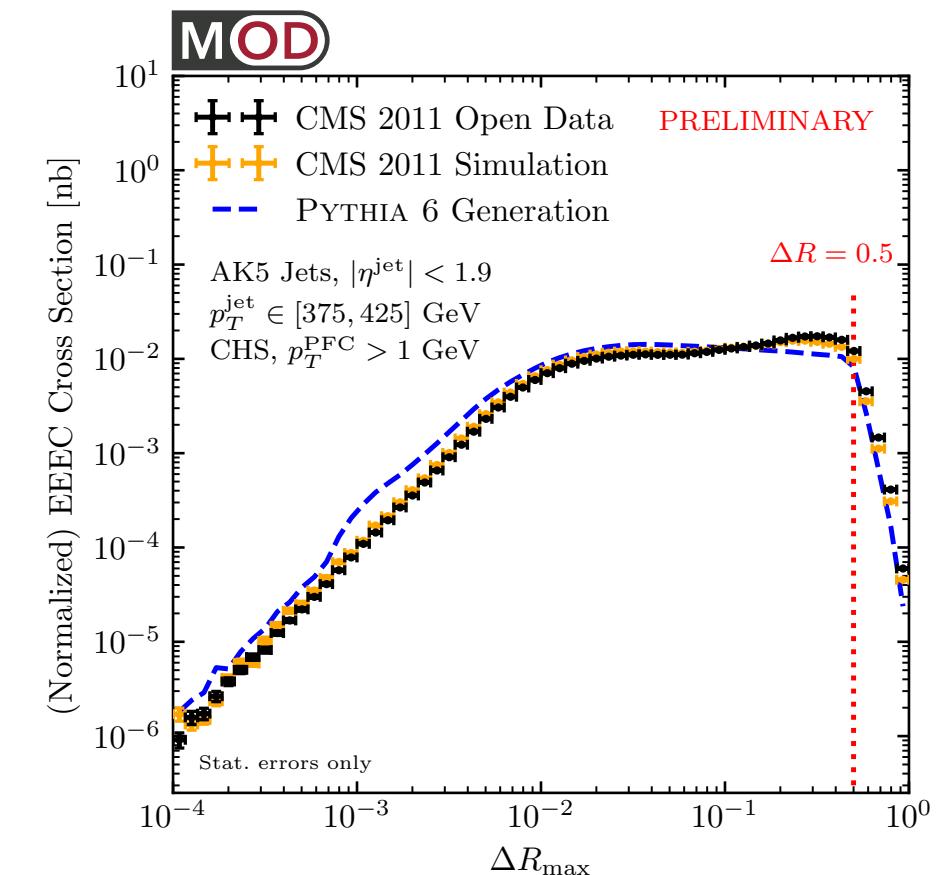
EEEC/EEC Ratio



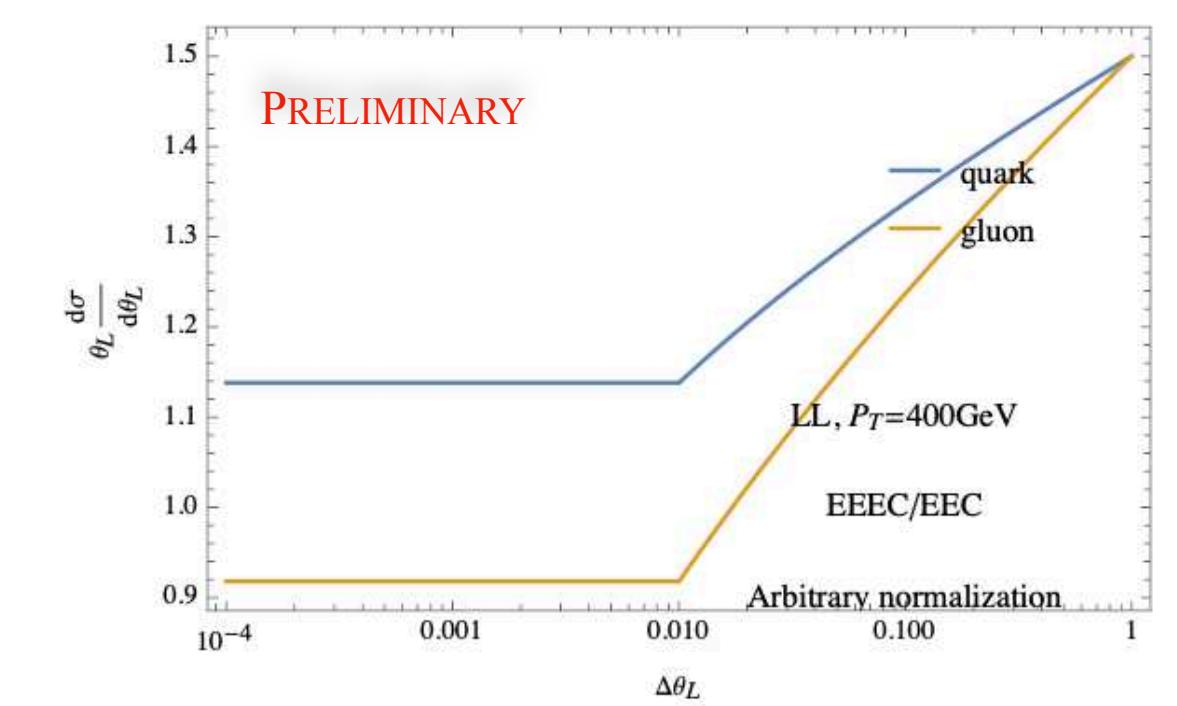
$N = 2$



$N = 3$



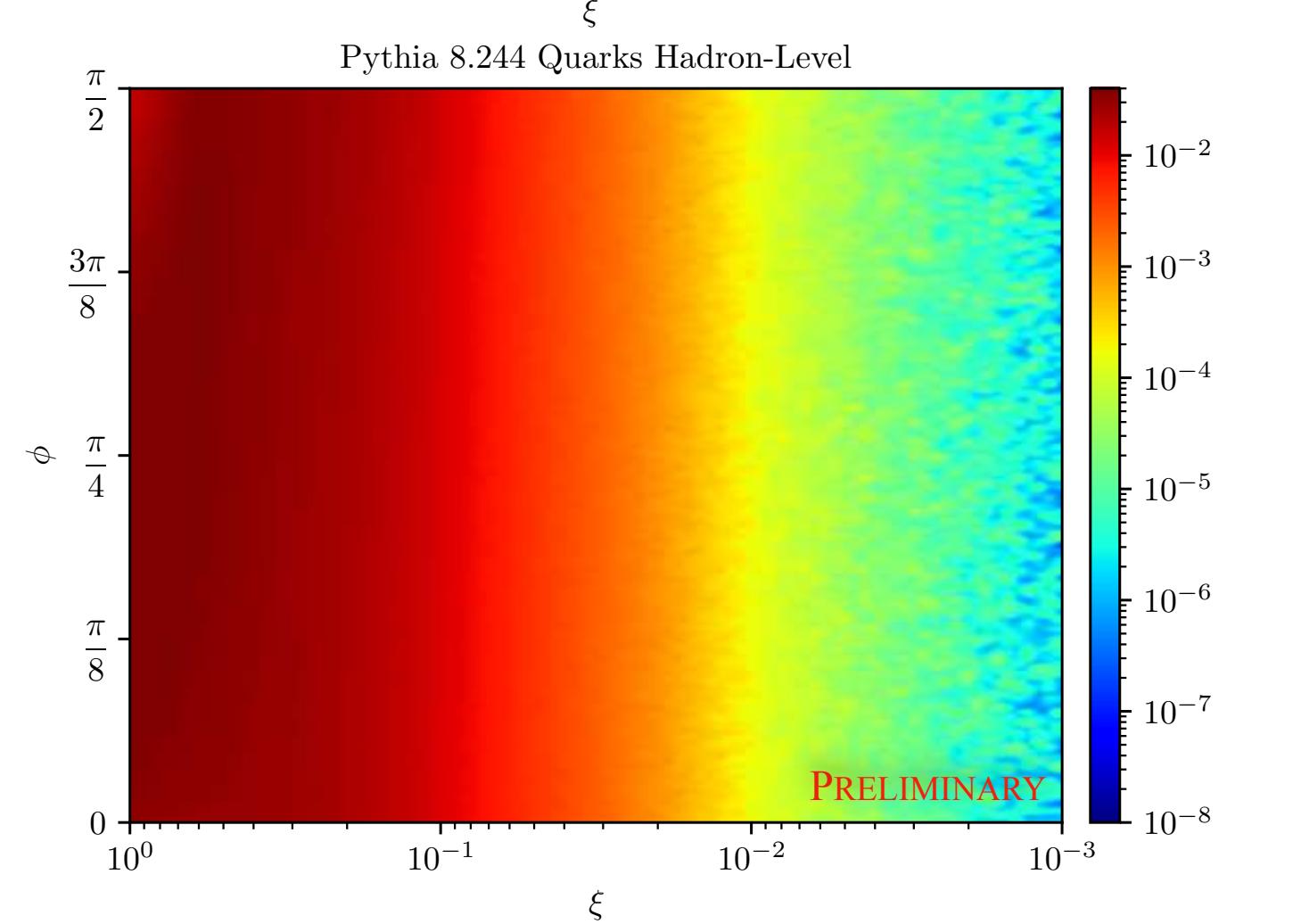
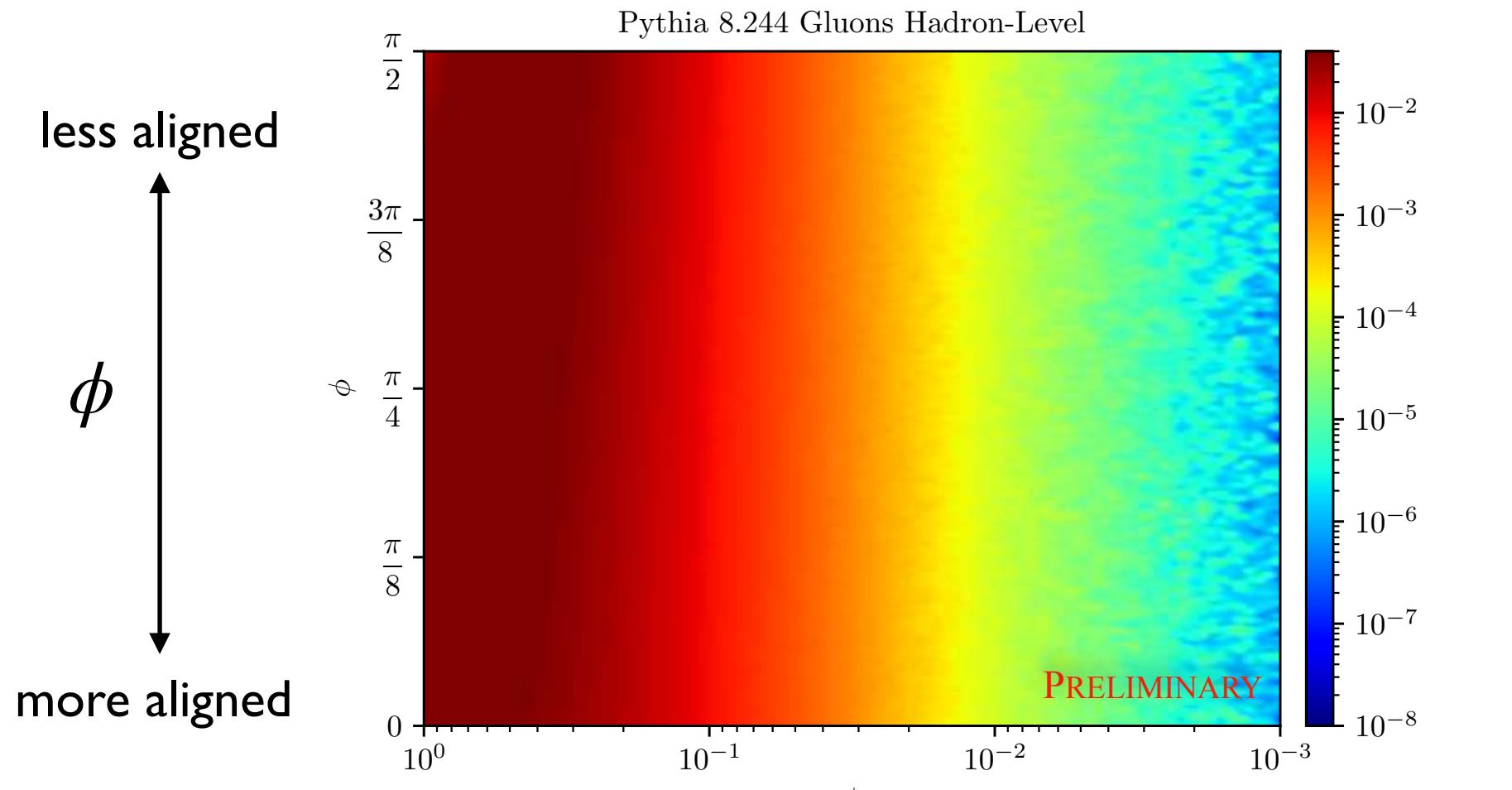
LL prediction of ratio



EEEC – Full Shape Dependence

[PTK, Moult, Thaler, Zhu, to appear soon]

For $x_L \sim 0.01$



less collinear \longleftrightarrow more collinear

EEEC – Full Shape Dependence

[PTK, Moult, Thaler, Zhu, to appear soon]

