Moment Unfolding with Deep Learning

Krish Desai

University of California, Berkeley/Lawrence Berkeley National Laboratory

in collaboration with Benjamin Nachman (LBNL) and Jesse Thaler (MIT)

France-Berkeley PHYSTAT Conference on Unfolding

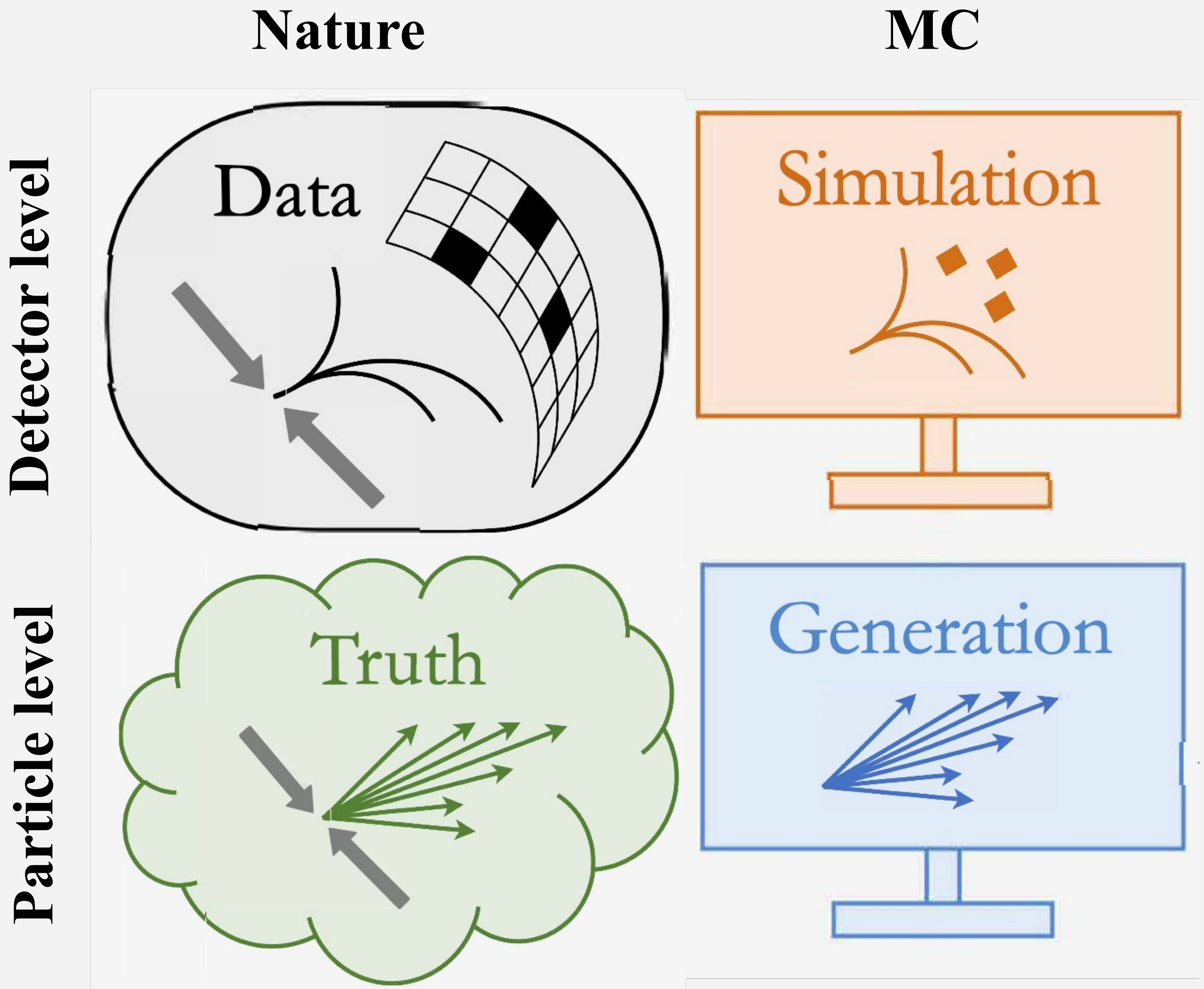
Goal: Unfold Moments – Remove detector distortions from means, variances etc. of jet observables

- Why?

- Moments known theoretically more precisely than full distributions
  - E.g. Wouter J. Waalewijn 1209.3019

$$n^{th} moment = \langle x^n \rangle = \sum_i p_i x_i^n$$

Datasets



Nature
MC
Detector level
Particle level
Data
Simulation
Truth
Generation

Existing methods and challenges

- Binned Methods e.g. IBU + various corrections
  - Cause artifacts when computing moments[2]

Average Jet Charge	Jet p_{T} Range [100 GeV]									
Systematic Uncertainty [%]	[0.5,1]	[1,2]	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,10]	[10,12]	[12,15]
Total Jet Energy Scale	$^{+39}_{-52}$	$^{+13}_{-3}$	$^{+2.2}_{-2.9}$	$^{+3.4}_{-0.9}$	$^{+0.4}_{-1.3}$	$^{+0.9}_{-2.4}$	$^{+1.0}_{-0.8}$	$^{+0.8}_{-1.1}$	$^{+0.2}_{-0.6}$	$^{+1.0}_{-1.3}$
Jet Energy Resolution	$^{+54.4}_{-54.4}$	$^{+4.4}_{-4.4}$	$^{+1.6}_{-1.6}$	$^{+0.9}_{-0.9}$	$^{+0.1}_{-0.1}$	$^{+0.6}_{-0.6}$	$^{+0.2}_{-0.2}$	$^{+0.1}_{-0.1}$	$^{+0.0}_{-0.0}$	$^{+0.1}_{-0.1}$
Charged Energy Loss	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{+1.5}_{-0.0}$	$^{+1.2}_{-0.0}$	$^{+1.4}_{-0.0}$	$^{+1.1}_{-0.0}$	$^{+1.2}_{-0.0}$	$^{+2.2}_{-0.0}$
Other Tracking	$^{+5.2}_{-2.8}$	$^{+2.7}_{-0.4}$	$^{+0.4}_{-0.4}$	$^{+0.8}_{-0.4}$	$^{+1.1}_{-0.3}$	$^{+0.5}_{-0.5}$	$^{+0.7}_{-0.7}$	$^{+0.9}_{-1.2}$	$^{+1.1}_{-1.8}$	$^{+0.9}_{-2.0}$
Track Multiplicity	$^{+0.0}_{-1.1}$	$^{+0.5}_{-0.0}$	$^{+0.0}_{-0.1}$	$^{+0.0}_{-0.8}$	$^{+0.0}_{-0.4}$	$^{+0.0}_{-0.8}$	$^{+0.0}_{-1.5}$	$^{+0.0}_{-2.3}$	$^{+0.0}_{-3.3}$	$^{+0.0}_{-4.6}$
Correction Factors	$^{+51}_{-51}$	$^{+11}_{-11}$	$^{+1.3}_{-1.3}$	$^{+0.7}_{-0.7}$	$^{+1.1}_{-1.1}$	$^{+0.8}_{-0.8}$	$^{+0.2}_{-0.2}$	$^{+0.3}_{-0.3}$	$^{+0.1}_{-0.1}$	$^{+0.0}_{-0.0}$
Unfolding Procedure	$^{+22.3}_{-22.3}$	$^{+1.7}_{-1.7}$	$^{+1.0}_{-1.0}$	$^{+0.3}_{-0.3}$	$^{+0.1}_{-0.1}$	$^{+1.3}_{-1.3}$	$^{+1.6}_{-1.6}$	$^{+1.9}_{-1.9}$	$^{+0.0}_{-0.0}$	$^{+3.3}_{-3.3}$
Total Systematic	$^{+87}_{-94}$	$^{+18}_{-12}$	$^{+3.1}_{-3.7}$	$^{+3.7}_{-1.8}$	$^{+2.2}_{-1.8}$	$^{+2.3}_{-3.0}$	$^{+2.5}_{-2.4}$	$^{+2.5}_{-3.4}$	$^{+1.6}_{-3.8}$	$^{+4.2}_{-6.1}$
Data Statistics	170	26	2.9	1.4	0.7	0.9	1.2	2.6	5.7	7.3
Total Uncertainty	$^{+191}_{-194}$	$^{+32}_{-29}$	$^{+4.3}_{-4.7}$	$^{+3.9}_{-2.2}$	$^{+2.3}_{-1.9}$	$^{+2.4}_{-3.2}$	$^{+2.7}_{-2.7}$	$^{+3.6}_{-4.3}$	$^{+6.0}_{-6.8}$	$^{+8.4}_{-9.5}$
Measured Value [e]	0.005	0.011	0.029	0.042	0.054	0.065	0.080	0.101	0.110	0.143

[2] ATLAS Collaboration 1509.05190

Existing methods and challenges

- Discriminative methods (e.g. OmniFold style methods)
  - General – attempt to unfold entire distributions rather than moments
  - Iterative – involves a large number of NNs and regulating by hand

- Generative methods (e.g. via normalizing flows)
  - Non-perturbative – attempt to learn scratch rather than perturb simulation

Moment Unfolding

- Novel re-weighting based GAN Like Method
- Deconvolves moments w/o binning
- Circumvents difficulty of computing PDF
- Circumvents binning artefacts
- Direct comparison with computations

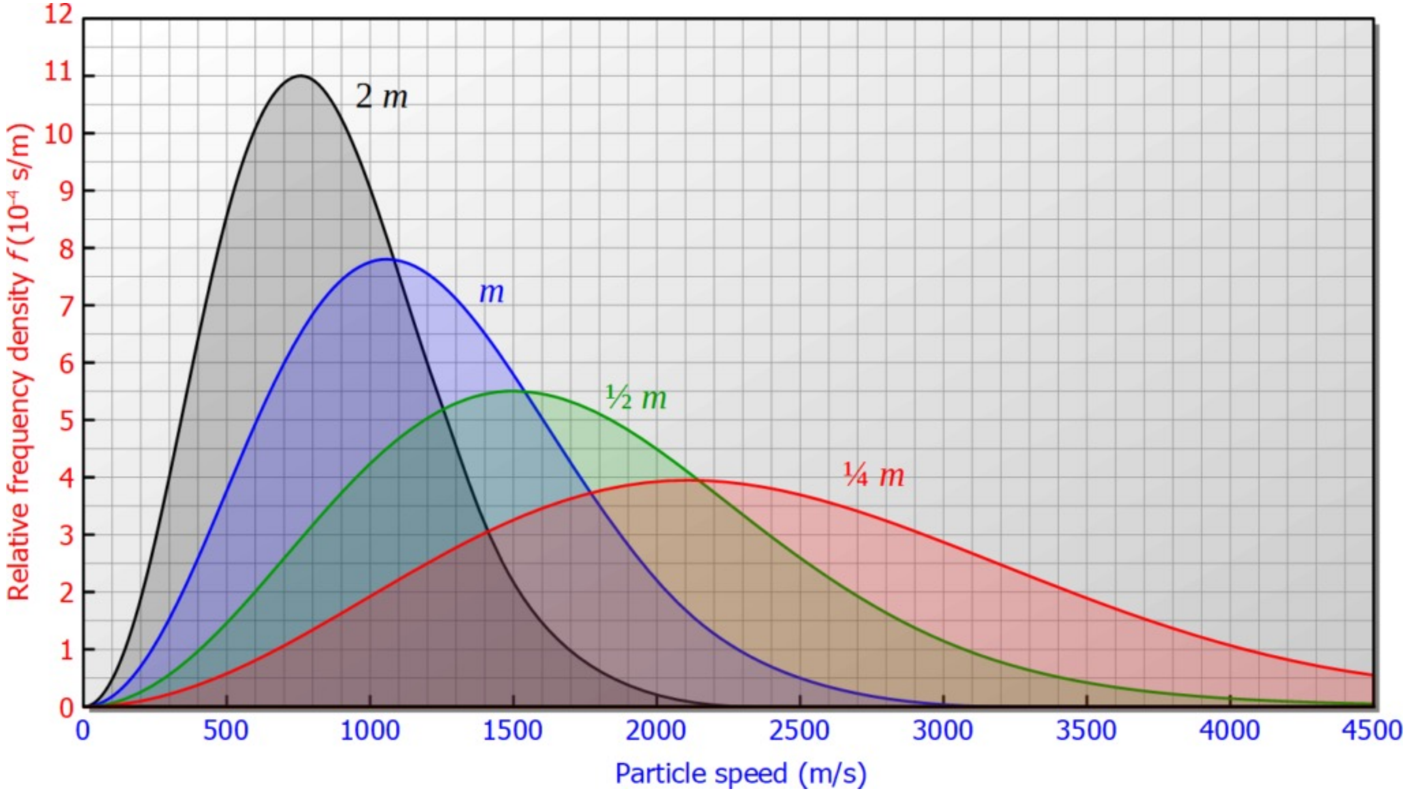
Advantages

- Pointwise reweighting based on matching moment by moment
- Non-iterative and perturbative
- Light, fast, precise

Inspiration

Maxwell–Boltzmann Distribution:

Generates weights using the inverse temperature to reweight the particle energies to maximize entropy holding mean energy constant



$$p_i = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}$$

Inspiration

Moment Unfolding:

Generates weights to reweight the events to maximize binary cross entropy loss while holding moments constant

$$g(z;\vec{\beta}) = \frac{e^{-\sum_{a=1}^{n} \beta_a z^a}}{\sum_{z \in \mathrm{Gen}} e^{-\sum_{a=1}^{n} \beta_a z^a}}$$

$$L[d,g] = -\sum_{x \in \mathrm{Data}} \log d(x) - \sum_{(z,x) \in \mathrm{Gen} \times \mathrm{Sim}} g(z;\vec{\beta}) \log(1 - d(x))$$

Critical points of L

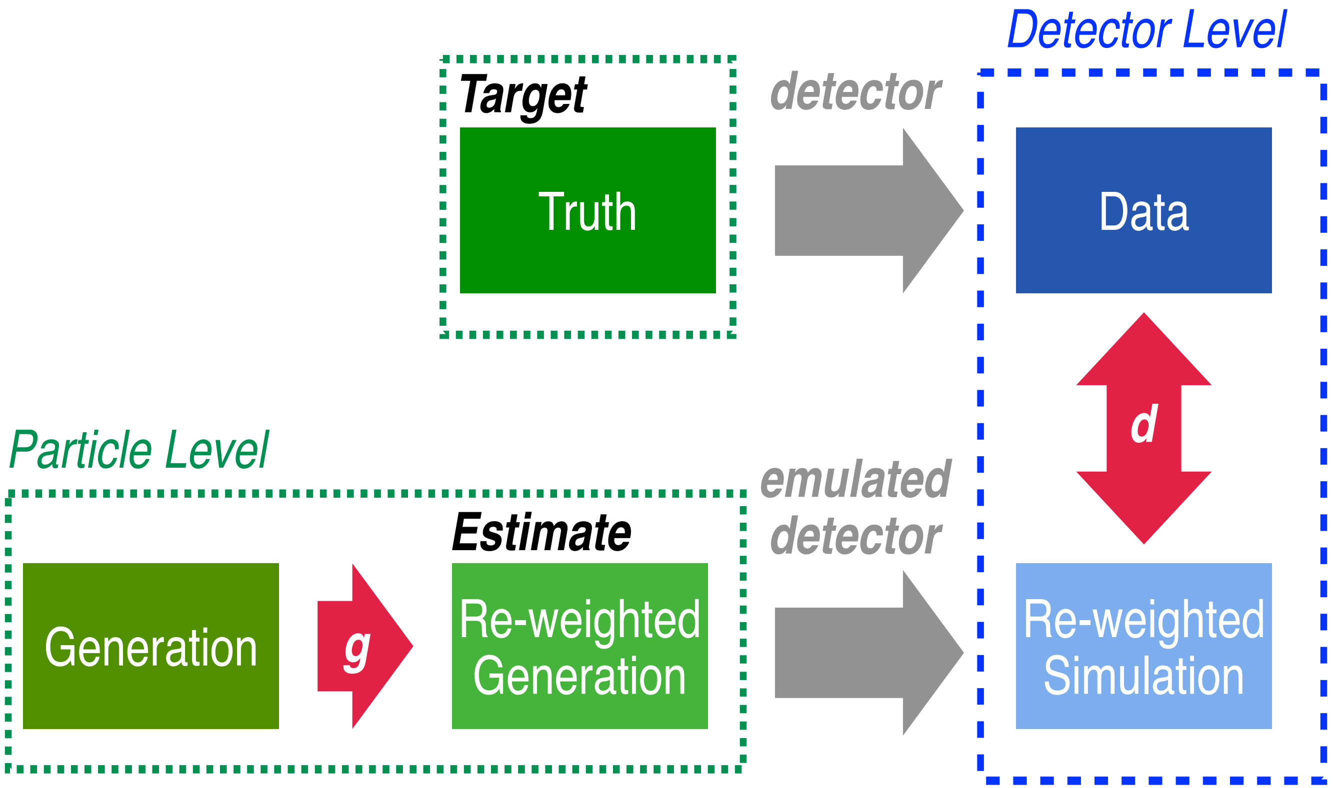
Let r(x | z) be the detector response function. Then

$$\frac{\delta L}{\delta d} = 0 \implies d(x) = \frac{p_{\mathrm{Data}}(x)}{p_{\mathrm{Data}}(x) \; + \int g(z)\, q_{\mathrm{Gen.}}(z)\, r(\, x \mid z)\, dz}$$

and

$$\frac{\delta L}{\delta g} = 0 \implies \forall k \in \{1, \ldots n\}\; \mathbb{E}_{\mathrm{Truth}}[Z^k] = \mathbb{E}_{\mathrm{Gen.}}[g(Z) Z^k] + \mathcal{O}(\mathrm{Std}(r))$$

Model



GAN Like Model

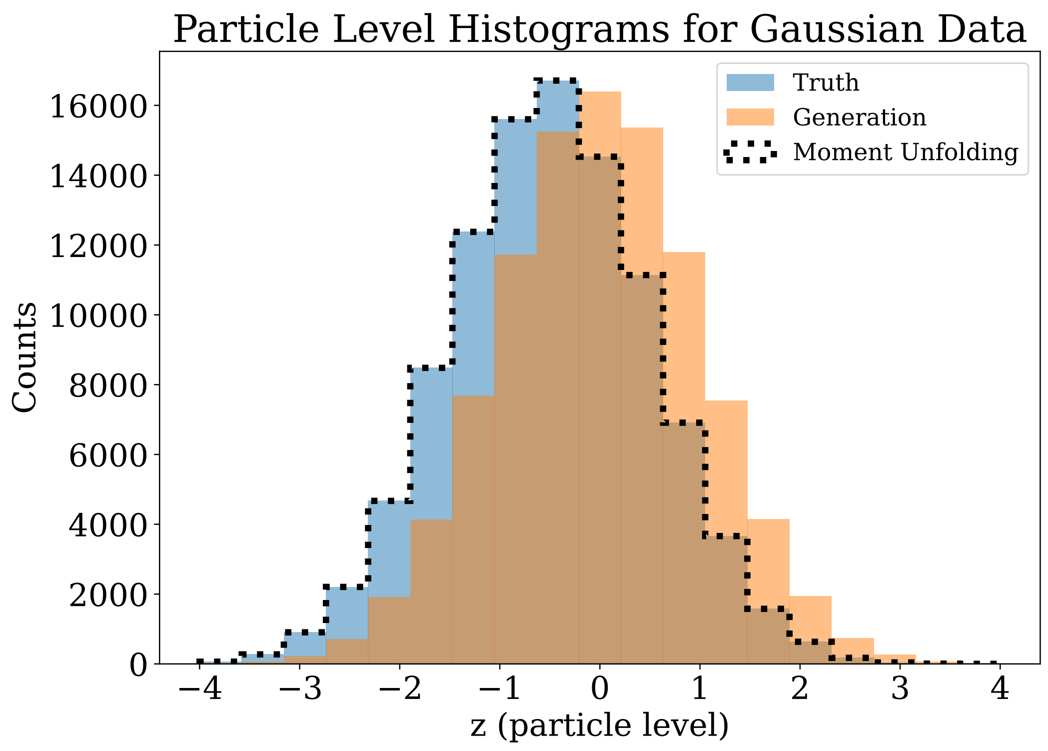
Generator:

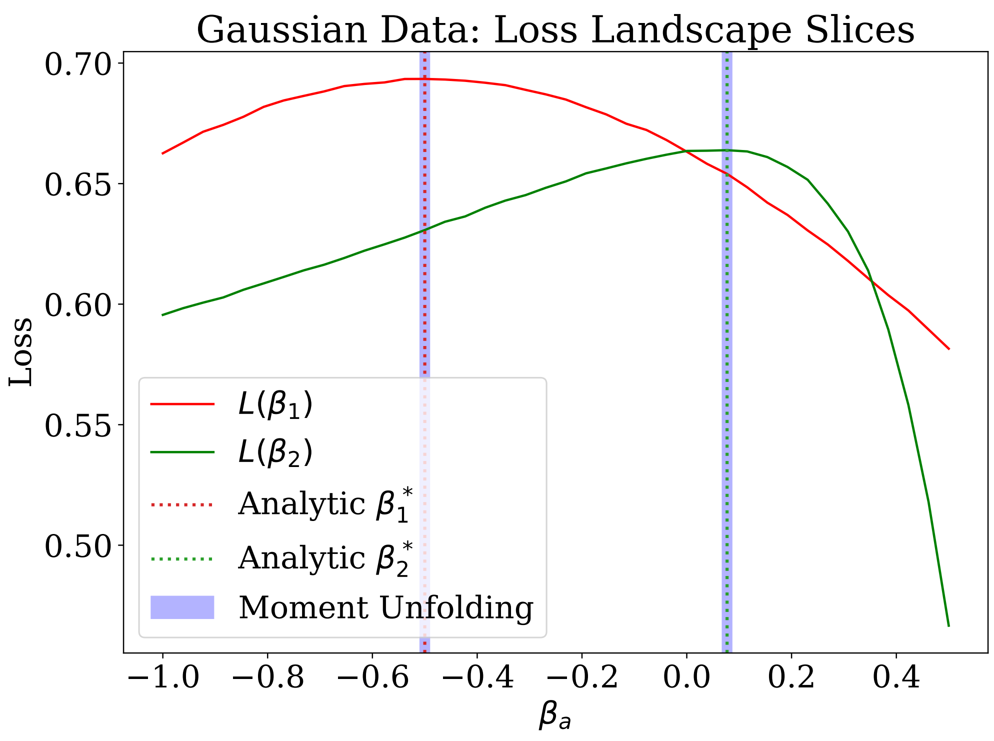
Reweighter function

$$g(z;\vec{\beta}) = \frac{1}{P}e^{-\beta_1 z - \ldots - \beta_n z^n}$$

Way fewer parameters than a traditional GAN

Gaussian Example





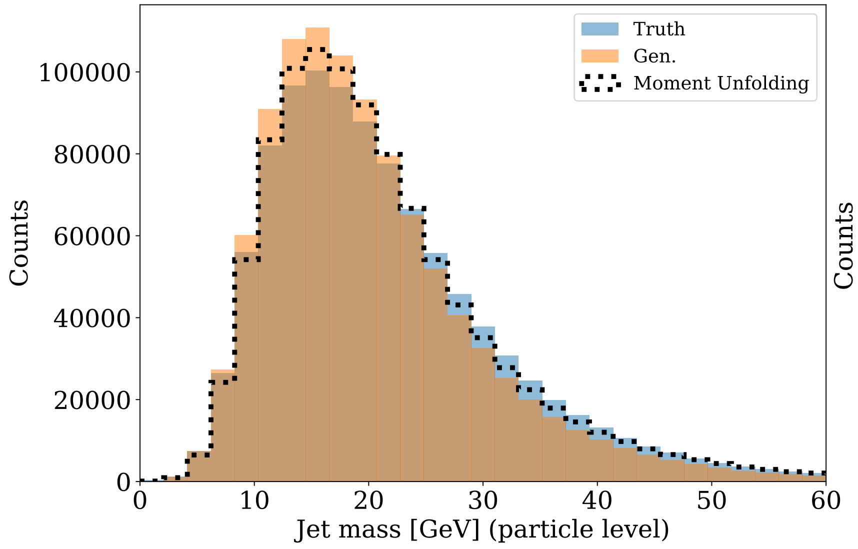
Truth: N(0, 1), Gen: N(0.5, 1), Distortions: N(0, 0.8)

Collider examples

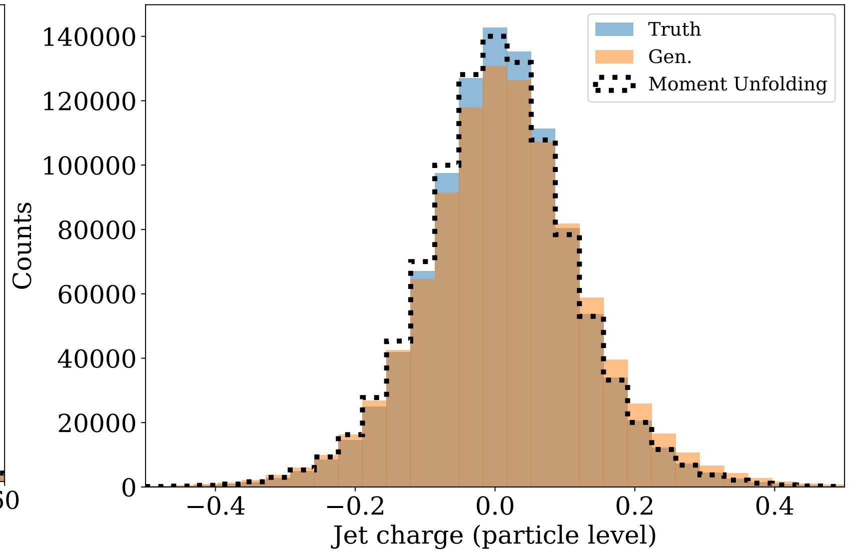
- Datasets (Z Jets, p_T > 200 GeV)
  - Detector Effects - Delphes 3.4.2
  - Truth/Data – Pythia 8.243 with tune 26 + Delphes
  - Generation/Simulation – Herwig 7.1.5 with the default tune + Delphes

Substructure Variables

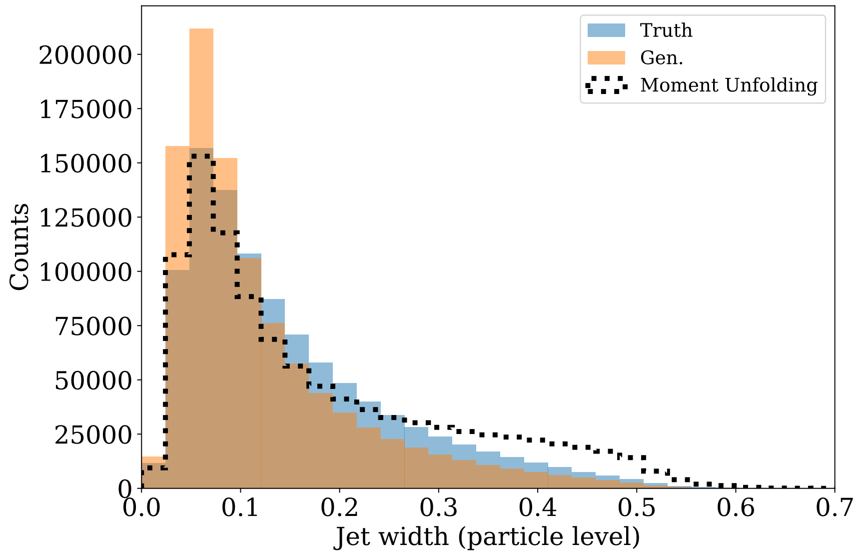
Jet Mass



Jet Charge



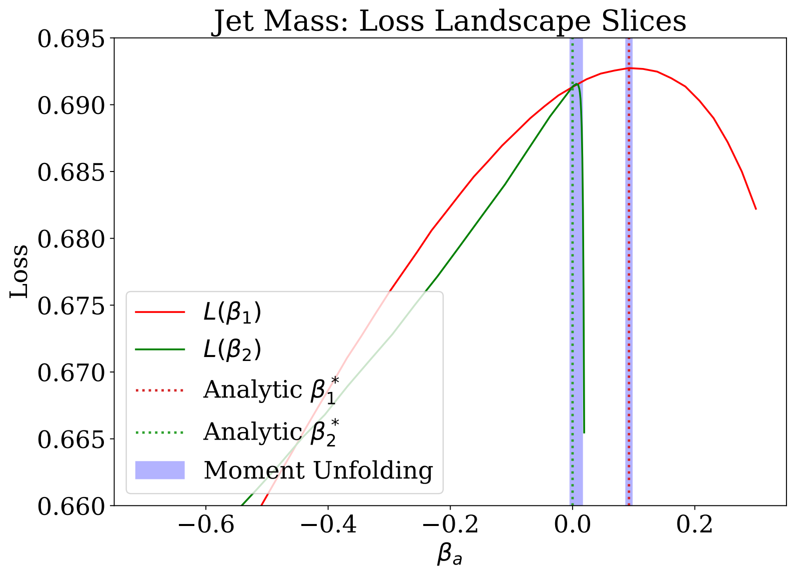
Jet Width

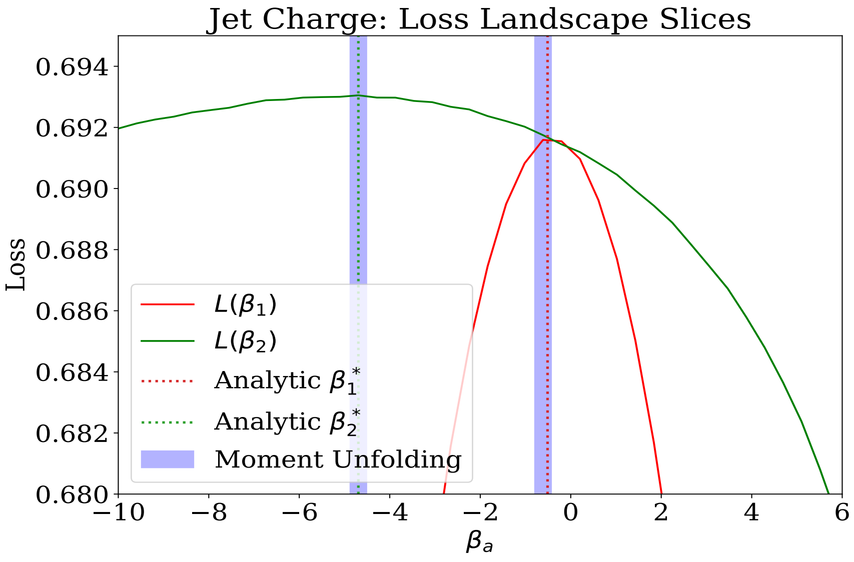


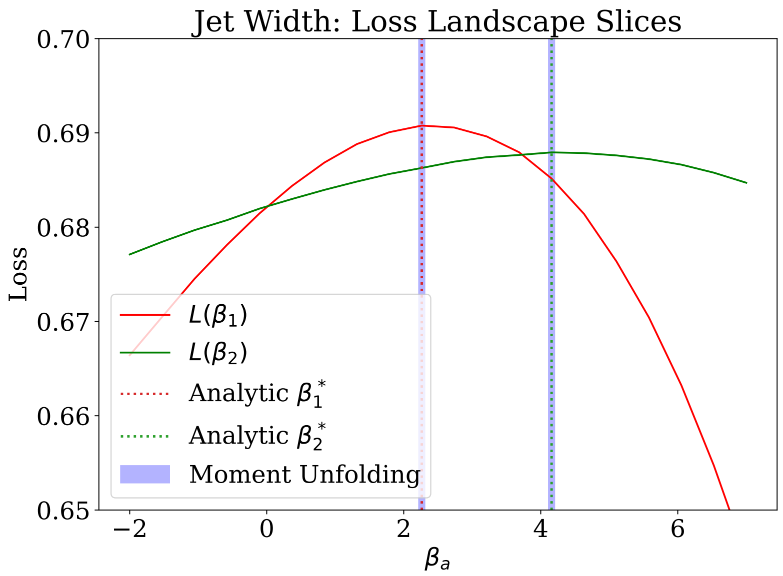
The method looks for weights that will reconstruct the moments, not necessarily the densities

Loss Landscapes

One dimensional slices of $L(\vec{\beta})$ space







Unfolded results

Observable	Truth	Generation	Moment Unfolding
$\langle M \rangle$	$(2.182 \pm 0.030) \times 10^1$	$(2.064 \pm 0.043) \times 10^1$	$(2.173 \pm 0.047) \times 10^1$
$\langle M^2 \rangle$	$(6.049 \pm 0.222) \times 10^2$	$(5.360 \pm 0.350) \times 10^2$	$(6.115 \pm 0.364) \times 10^2$
$\langle Q \rangle$	$(1.006 \pm 0.037) \times 10^{-2}$	$(1.582 \pm 0.038) \times 10^{-2}$	$(1.090 \pm 0.040) \times 10^{-2}$
$\langle Q^2 \rangle$	$(1.216 \pm 0.082) \times 10^{-2}$	$(1.508 \pm 0.074) \times 10^{-2}$	$(1.207 \pm 0.074) \times 10^{-2}$
$\langle W \rangle$	$(1.498 \pm 0.025) \times 10^{-1}$	$(1.231 \pm 0.029) \times 10^{-1}$	$(1.499 \pm 0.029) \times 10^{-1}$
$\langle W^2 \rangle$	$(3.370 \pm 0.113) \times 10^{-2}$	$(2.421 \pm 0.128) \times 10^{-2}$	$(3.374 \pm 0.128) \times 10^{-2}$

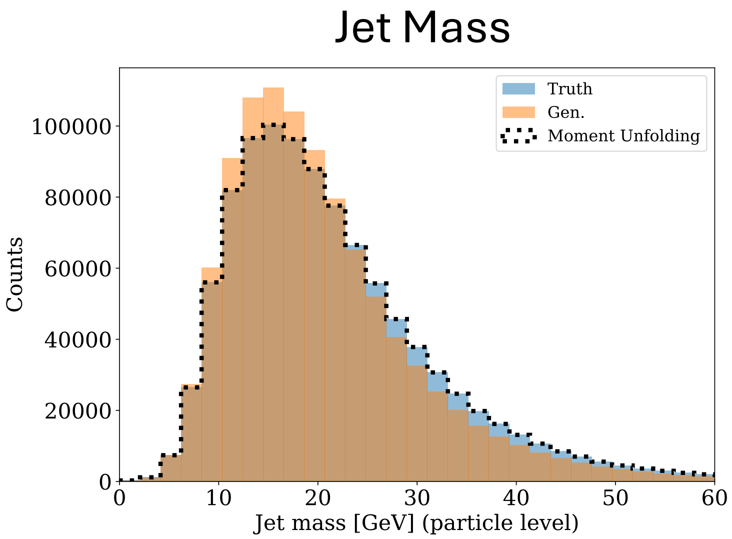
Momentum dependence

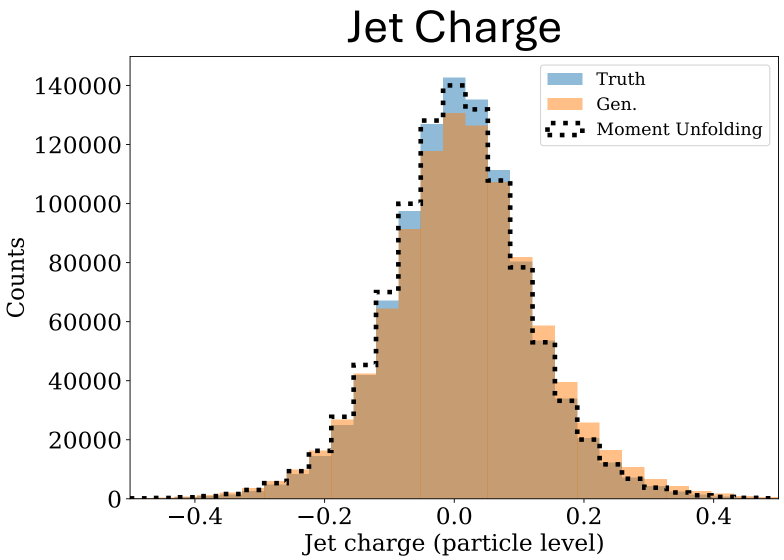
$$g[z, \vec{\beta}(p_T)] = \frac{1}{\mathrm{P}} \mathrm{e}^{-\sum_{a=1}^{n} \beta_a(p_T)\, z^a}$$

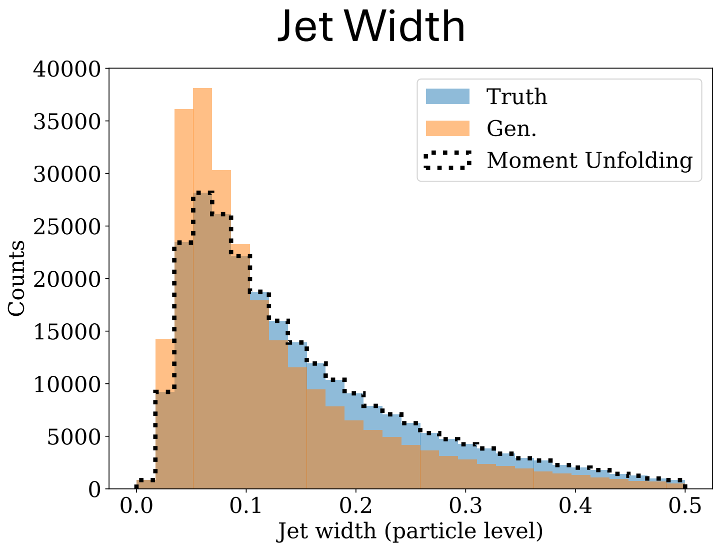
Empirically, the ratio of spectra is approximately linear so we parametrize β_a as linear functions

$$\vec{\beta}(p_T) = \overrightarrow{\beta^{(0)}}(p_T) + p_T\, \overrightarrow{\beta^{(1)}}(p_T)$$

Momentum Dependence

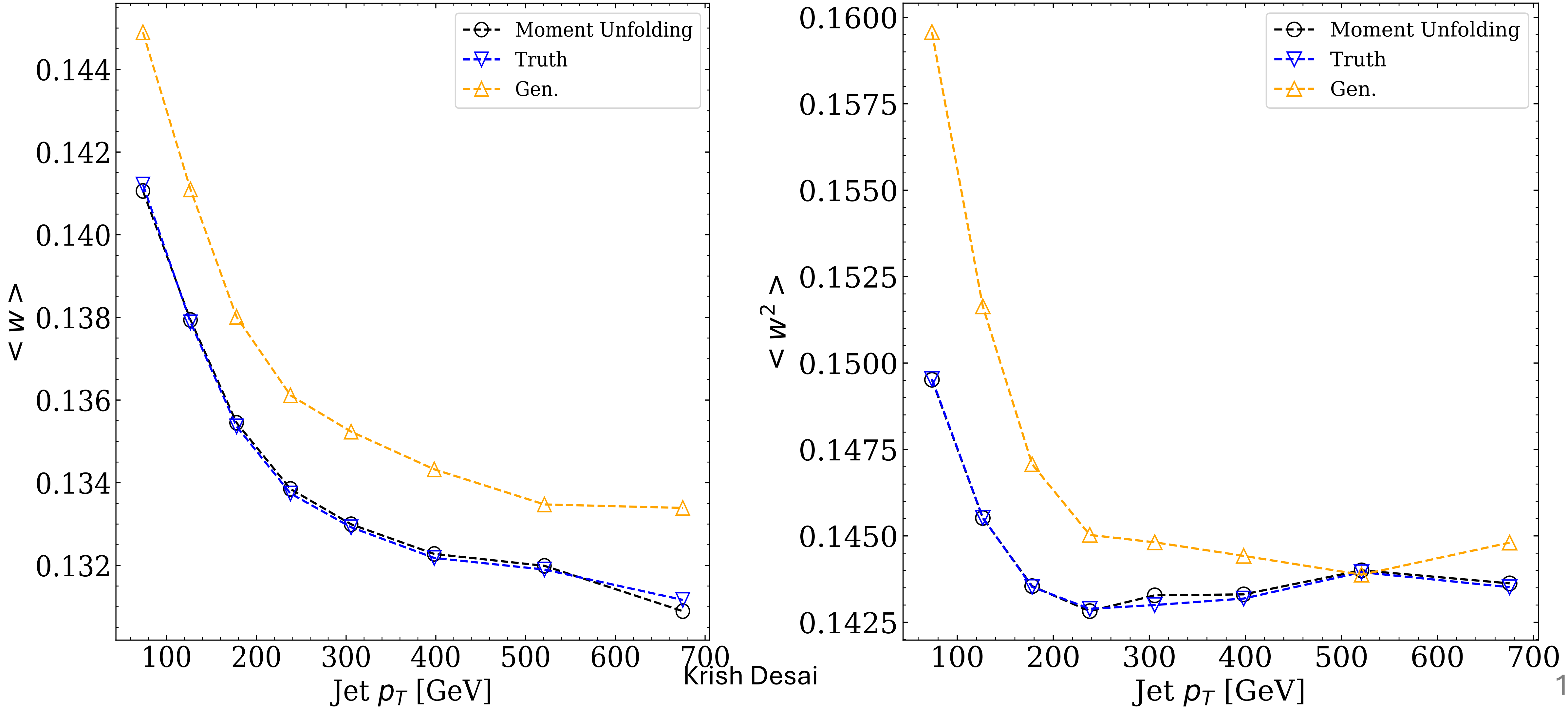




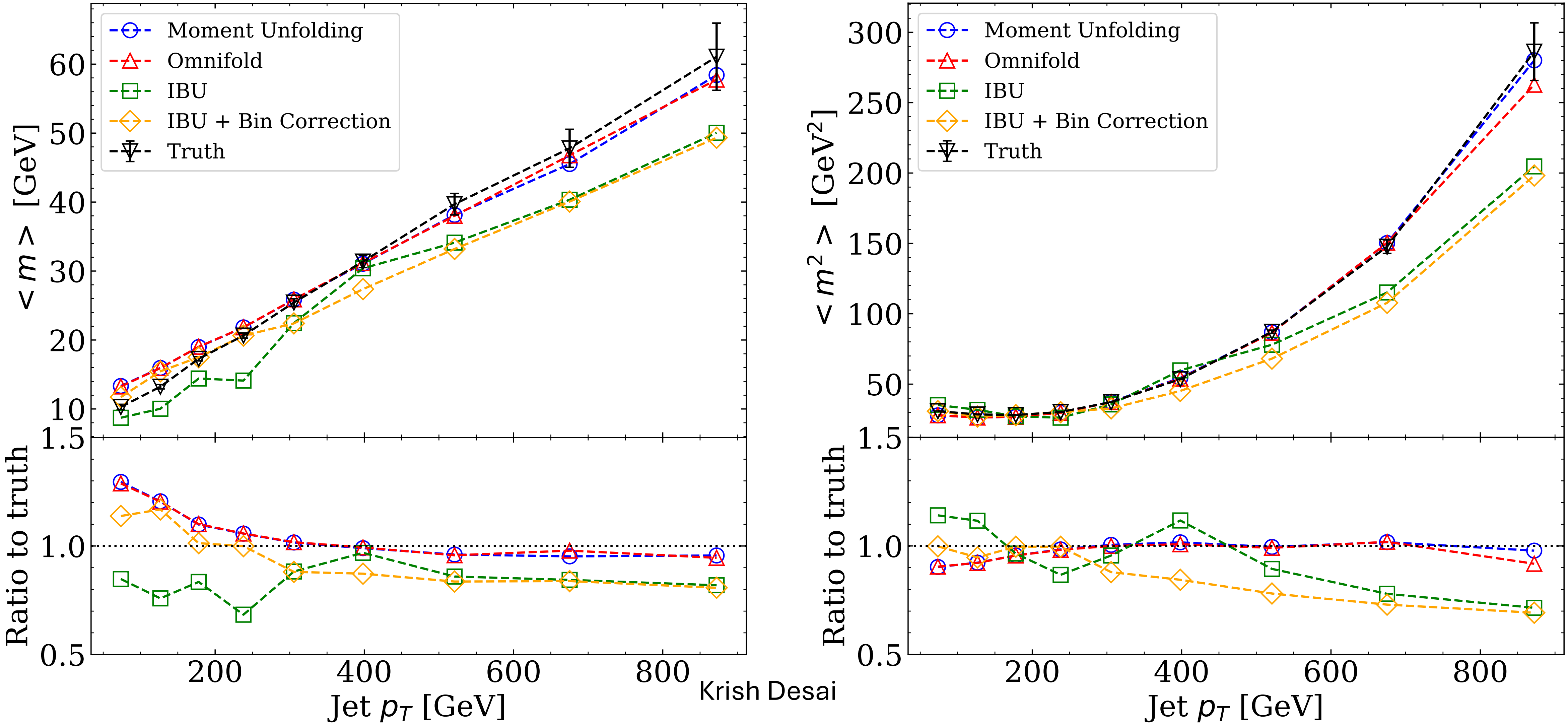


- Substructure variables unfolded conditional on jet p_T but plotted inclusively
- Note greater accuracy of unfolding

Jet variables as a function of p_T



Comparison to other methods



Krish Desai

Bonus: Unfold entire distributions?

Infinite Moment Unfolding

$$g(x) = e^{\lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots} = e^{f(x)}$$

Success:

- Analytically can be proven to converge to MLE

Problems:

- Renormalization
- Generator stabilization

Conclusion and Outlook

Introduced novel reweighting-based GAN-like model to unfold moments without binning artifacts.

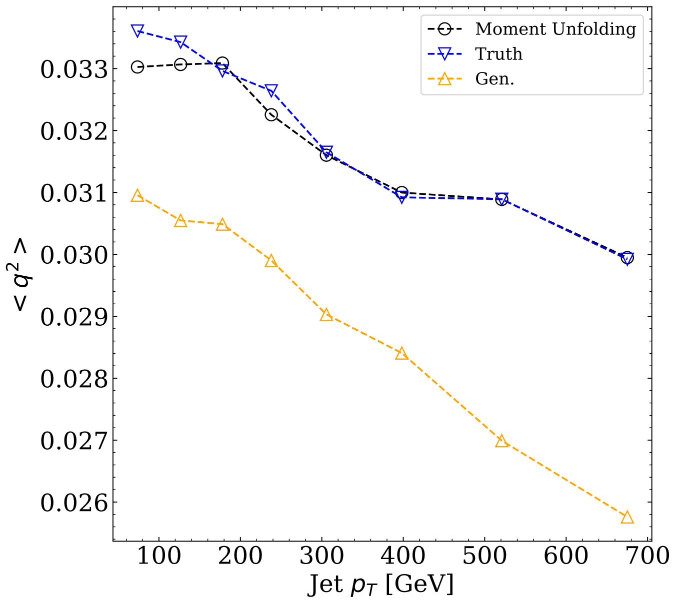
Successfully unfolded moments of jet substructure variables

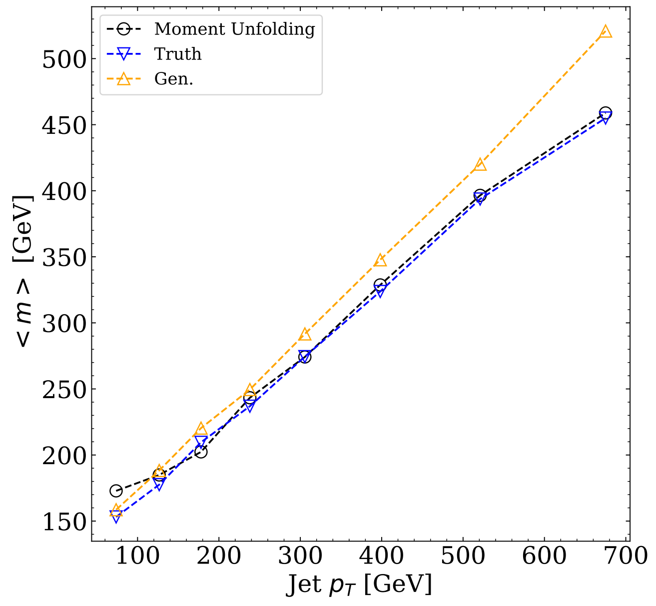
Precisely computed momentum dependence of moments

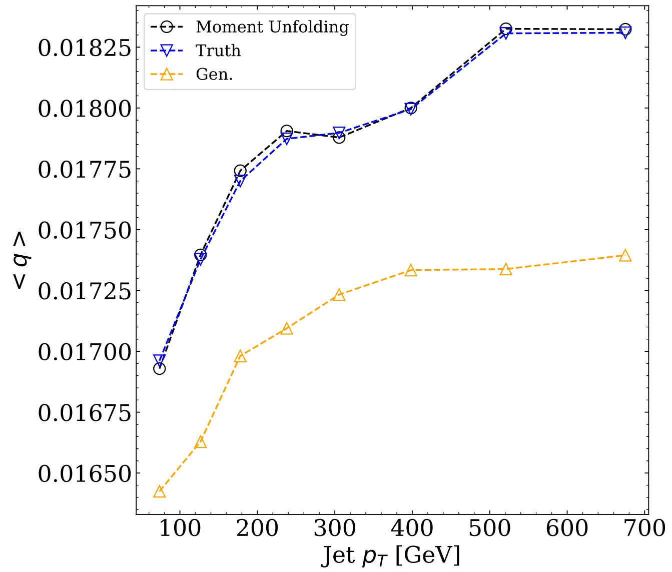
Infinite Unfolding and Renormalization

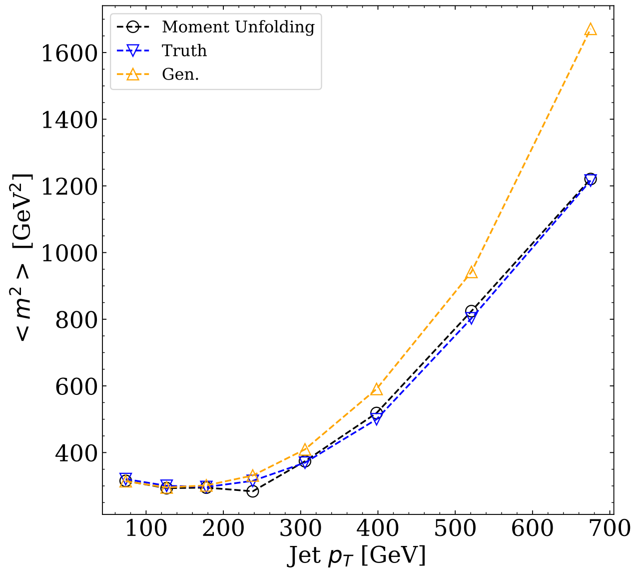
Questions?

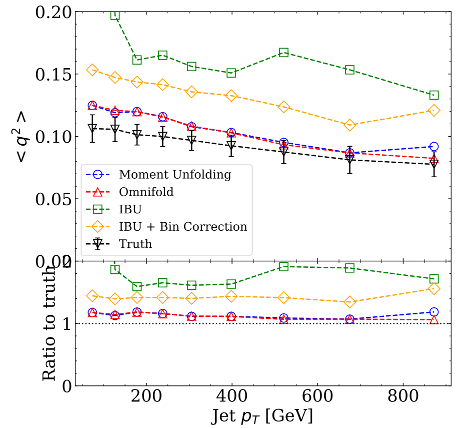
Nachman Machine Learning Group

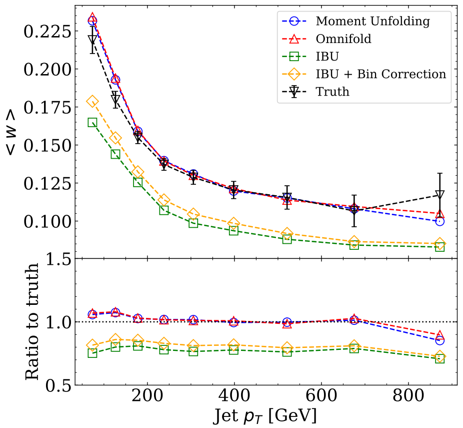

Moment Unfolding
Truth
Gen.
$< q^2 >$
Jet p_T [GeV]

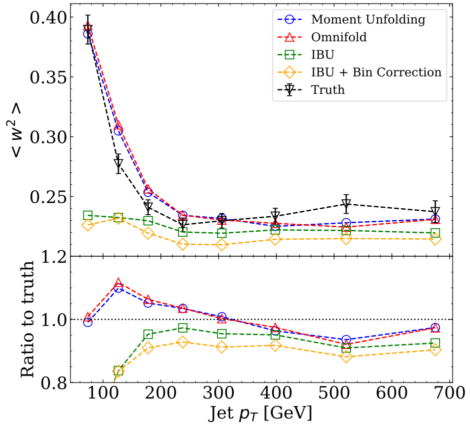

Moment Unfolding
Truth
Gen.
$< m >$ [GeV]
Jet p_T [GeV]

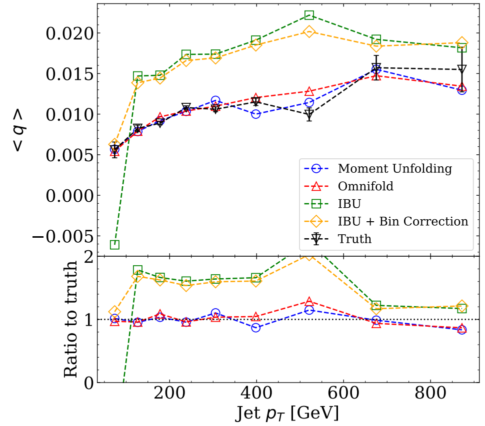

Moment Unfolding
Truth
Gen.
$< q >$
Jet p_T [GeV]


Moment Unfolding
Truth
Gen.
$< m^2 >$ [GeV^2]
Jet p_T [GeV]


0.20
0.15
0.10
0.05
0.00
< q^2 >
Moment Unfolding
Omnifold
IBU
IBU + Bin Correction
Truth
Ratio to truth
1
0
200
400
600
800
Jet p_T [GeV]


0.225
0.200
0.175
0.150
0.125
0.100
< w >
Moment Unfolding
Omnifold
IBU
IBU + Bin Correction
Truth
1.5
1.0
0.5
Ratio to truth
200
400
600
800
Jet p_T [GeV]


0.40
0.35
0.30
0.25
< w^2 >
Moment Unfolding
Omnifold
IBU
IBU + Bin Correction
Truth
1.2
1.0
0.8
Ratio to truth
100
200
300
400
500
600
700
Jet p_T [GeV]


0.020
0.015
0.010
0.005
0.000
−0.005
< q >
Moment Unfolding
Omnifold
IBU
IBU + Bin Correction
Truth
2
1
0
Ratio to truth
200
400
600
800
Jet p_T [GeV]