Deconvolving Detector Effects for Distribution Moments

Krish Desai, Benjamin Nachman, Jesse Thaler UC Berkeley, LBNL, MIT krish.desai@berkeley.edu

Abstract

- Deconvolving detector distortions – critical step to **compare cross section measurements** with **theoretical predictions**
- Most approaches require binning
- Theoretical predictions often at the level of moments.
- New approach to **directly unfold distribution moments** as a function of other observables **without discretizing** data
- **Moment Unfolding** uses a**modified GAN** architecture
- **Demonstrate the performance** of this approach using jet substructure measurements in collider physics.

Background

- **Unfolding (deconvolution):** correcting detector distortions in experimental data – necessary for accurately comparing data between experiments and with theoretical predictions.
- Typically, entire spectra unfolded – moments computed afterward.
- **Current approaches** discretize support–then unfold histogram.
- This binning procedure introduces **discretization artifacts.**
- Unfold without binning – generic solution to unfolding entire spectra – may compromise precision for small set of moments.
- Dedicated machine learning-based unfolding method to directly unfold the observable moments.
- GAN learns reweighting function inspired by the Boltzman equation – parameters identified with observable moments
- Noniterative in contrast to methods like `Omnifold` (*Andreassen et al*, Phys. Rev. Lett. 124, 182001 – 2020)

Method

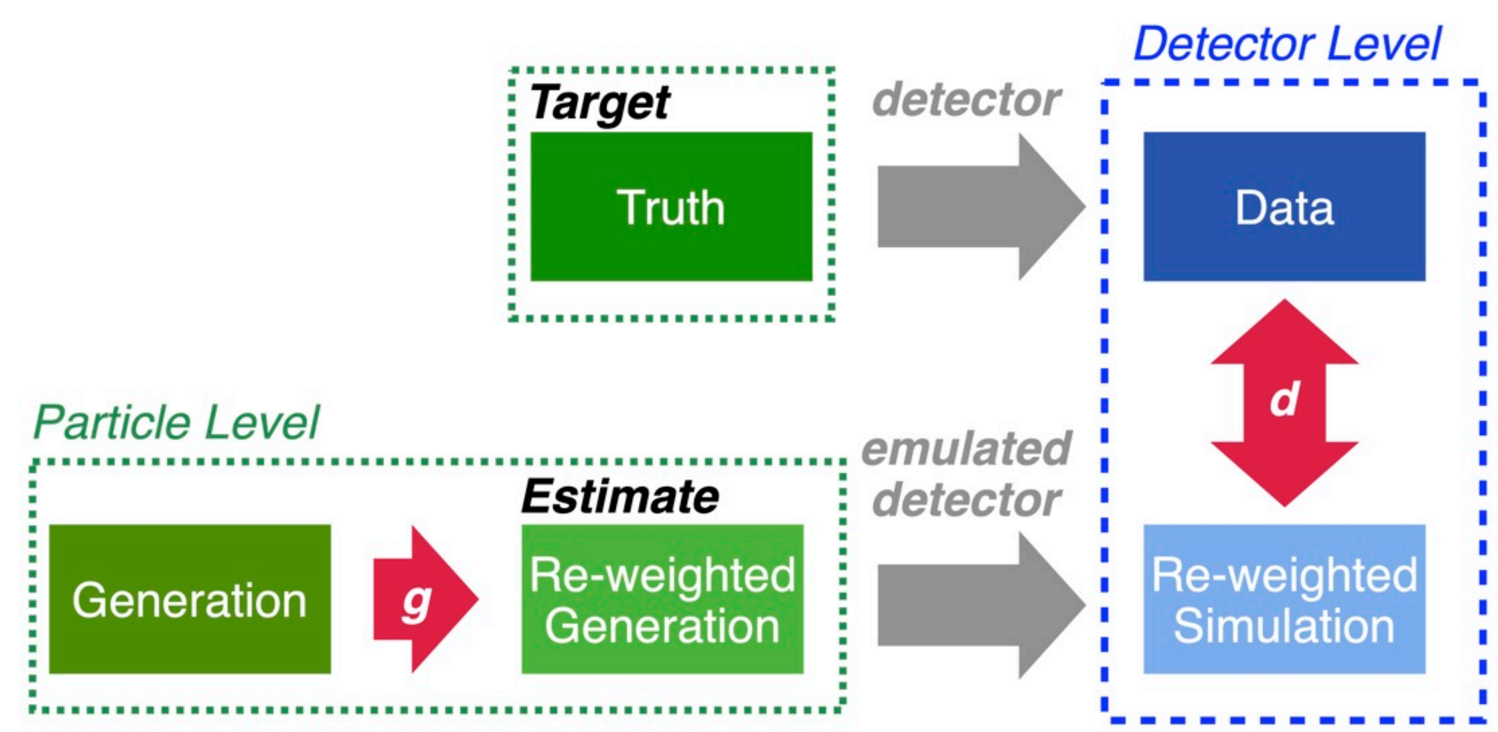
Neural Networks

Generator: $g(x) = e^{\lambda_1 x + \cdots + \lambda_n x^n}$ **reweights** generation to truth. n trainable parameters λ_i learn n moments
Discriminator: $d : \mathbb{R} \to [0,1]$ distinguishes **reweighted simulation** from **data**

- Maxwell–Boltzmann distribution maximizes entropy while holding mean energy constant
- Moment unfolding maximises BCE loss holding moments constant

Data sets

- Simulation (X_S) : detector level simulation
- Generation (X_G) : particle level simulation
- Data (X_D) : detector level data
- Truth (X_T) : particle level data



Machine Learning Implementation

Loss Function

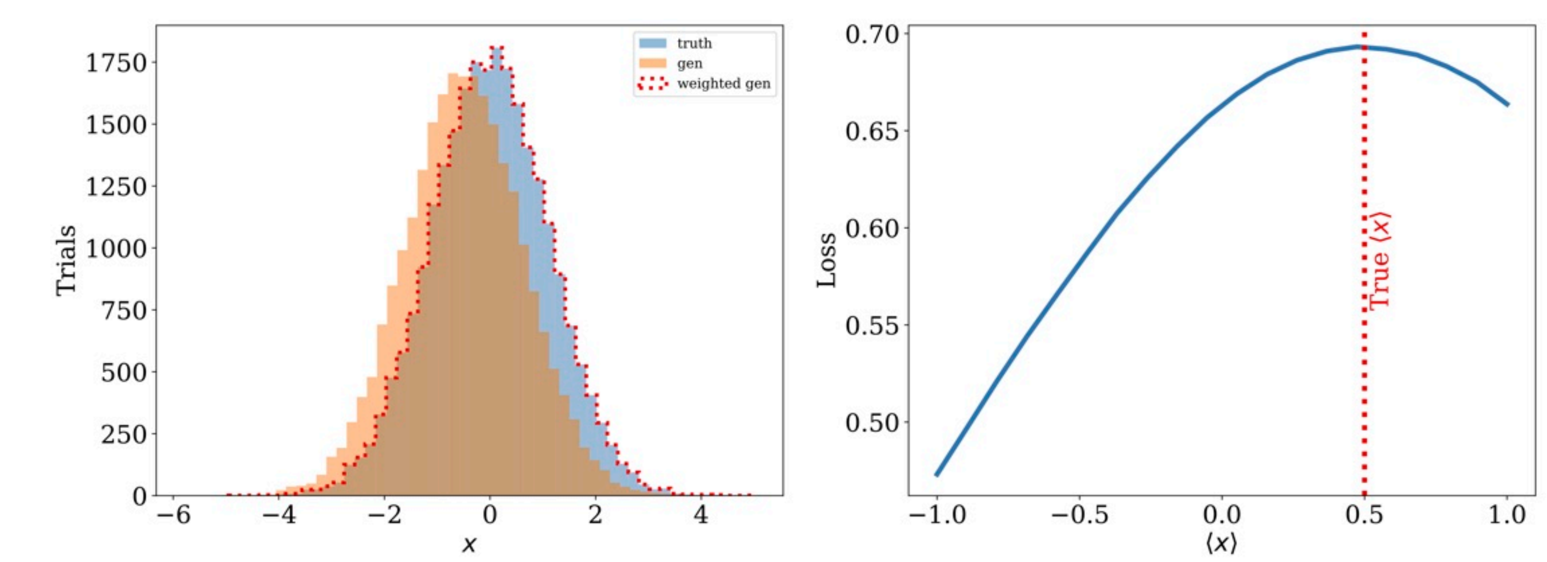
$$L[g,d] = -\frac{1}{N}\sum_{x_D}\left[\log\left(d(x_D)\right) + \sum_{(x_G,x_S)} g(x_G)\log\left(1-d(x_S)\right)\right]$$

- All neural networks implemented using the Keras high-level API with TensorFlow2 backend, optimized with Adam.
- Discriminator d parametrized with three hidden layers, 50 nodes per layer. Intermediate layers → ReLU, last layer → sigmoid.

Gaussian Case Study

Gaussian distributions

- $X_T \sim \mathcal{N}(0.5, 1)$
- $X_G \sim \mathcal{N}(0, 1)$
- detector effects $Z \sim \mathcal{N}(0, 5)$

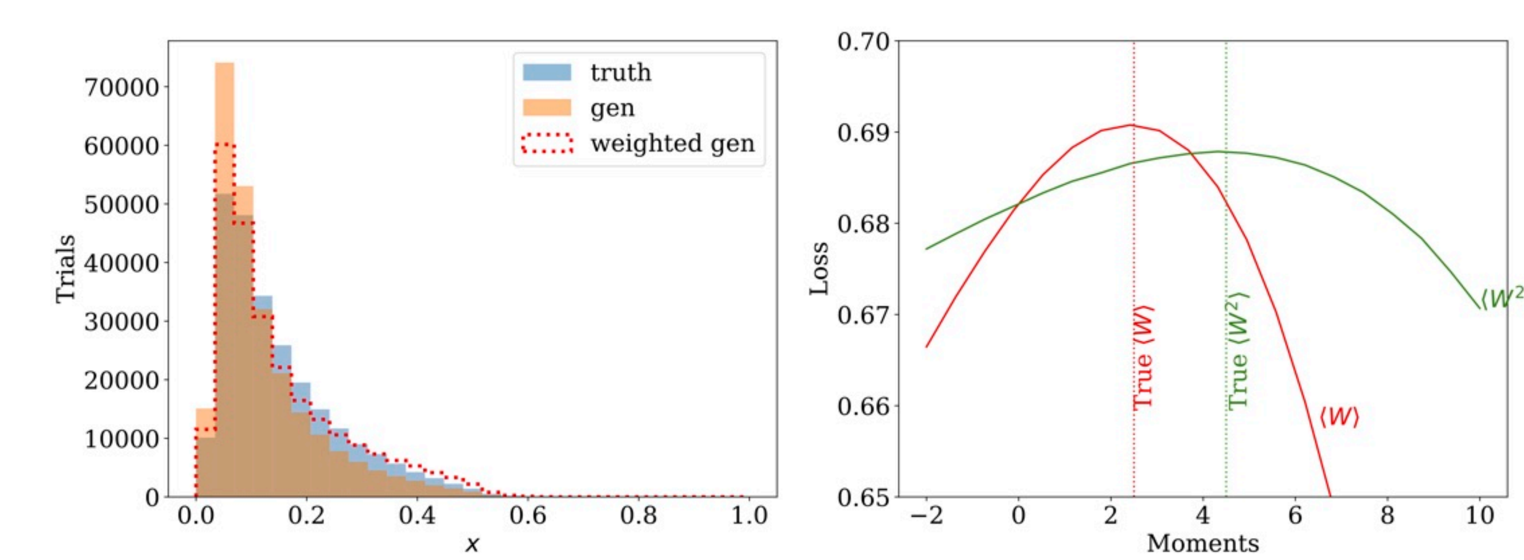


Jet Substructure

Jet width

$$w = \frac{1}{p_{T,jet}}\sum_{i\in \text{jet}} p_{T,i}\Delta R_i$$

- $p_{T,i}$ = transverse momentum
- ΔR_i = angular distance from jet axis
- Example focusses on jet width. Method can be applied to other jet observables like charge, mass, etc.



- First two moments of the reweighted generation match truth well
- Full distributions not statistically identical
- This is because higher moments are relevant and are not the same between truth and generation.

Conclusions and Outlook

- Proposed Moment Unfolding as a novel, flexible, unbinned, and non-iterative reweighting–based deconvolution method
- Showed promising results when applied to both Gaussian datasets as well as detector data from the LHC
- Future work – whether this method could be used to unfold infinitely many moments, i.e. entire probability density
- Important questions about partition function normalization, stability, and overlapping support