Investigating the Use of Omnifold in Unfolding Jet Transverse Momenta

By Simon Browning of Abilene Christian University
for the STAR Collaboration

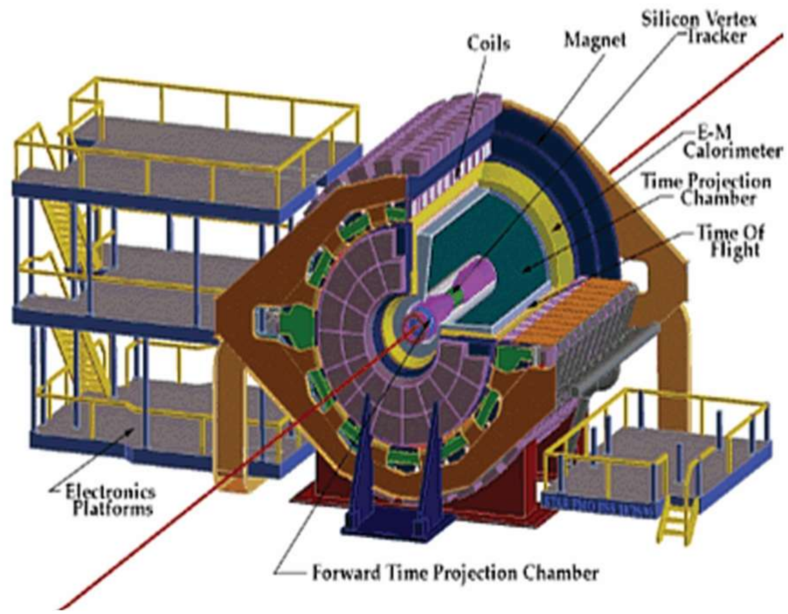
This research was supported by US DOE MENP
Grant DE-FG02-03ER41243.

Abstract - Summary

- Detector effects distort information, requiring data to be “unfolded”.
- Correcting for effects becomes challenging with steeply-falling spectra.
- Iterative Bayesian Unfolding (IBU) is well-established but requires data to be binned.
- Omnifold is a new method that uses machine learning to unfold data event-by-event.
- Investigate Omnifold’s ability to recover true distributions.

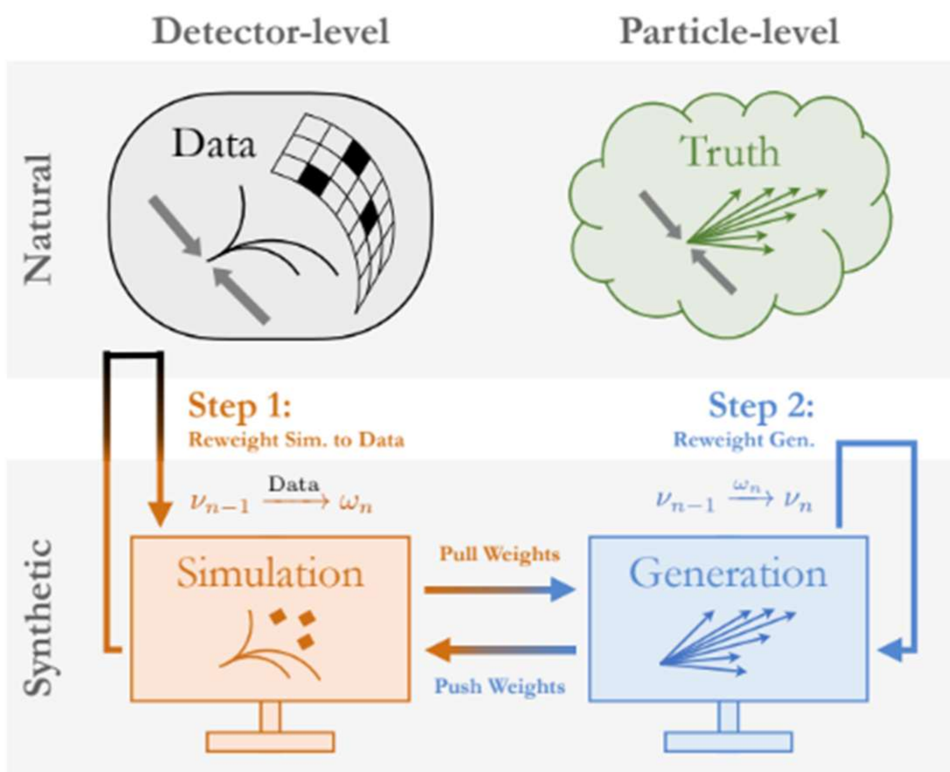
The STAR Experiment

- Studies the strong interaction between subatomic particles.
- Housed at BNL
- Focuses on studying proton spin and the quark-gluon plasma (QGP).
- Jet analysis requires the unfolding of multiple observables simultaneously.
- Difficult to do with traditional/fixed-bin unfolding techniques.



What is Omnifold?

- A machine-learning based algorithm.
- A non-fixed-bin unfolding method.
- Deals with natural and Monte Carlo synthetic data sets.
- Each data set contains detector-level and particle-level data.
- Keras neural network classifier.

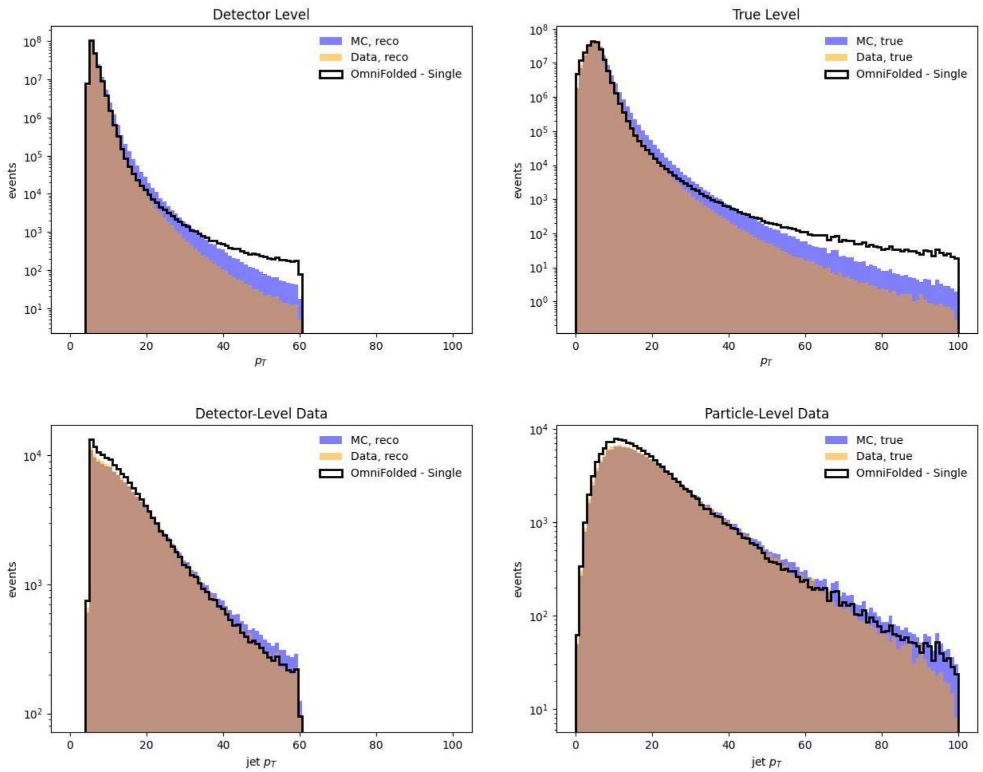


Andreassen, Anders, et al., *Omnifold: A Method to Simultaneously Unfold All Observables*, 2019.

Steeply Falling Spectra

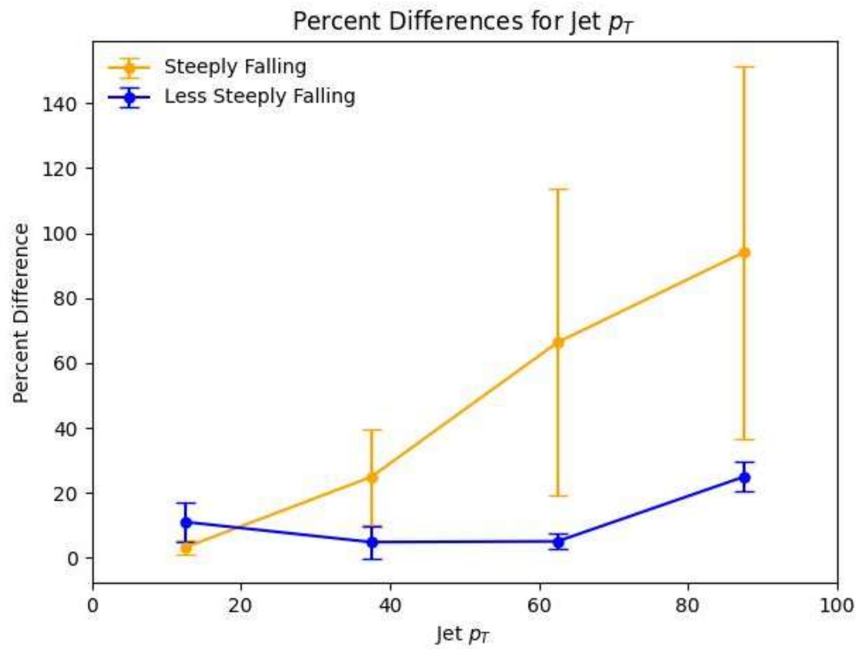
- We seek to generate data that approximates steeply-falling jet p_T spectra.
- Test data generated via fast Monte Carlo based on STAR run 11 detector simulation.
- Generating the low-statistics tails of steeply falling spectra is numerically expensive.
- We generate less steeply-falling data and then reweight to better reproduce observed jet p_T spectra.
- We apply Omnifold to each distribution to test sensitivity to steepness of the spectra.
- Better performance on less steep spectra.

Counts vs. Transverse Jet Momentum [GeV]



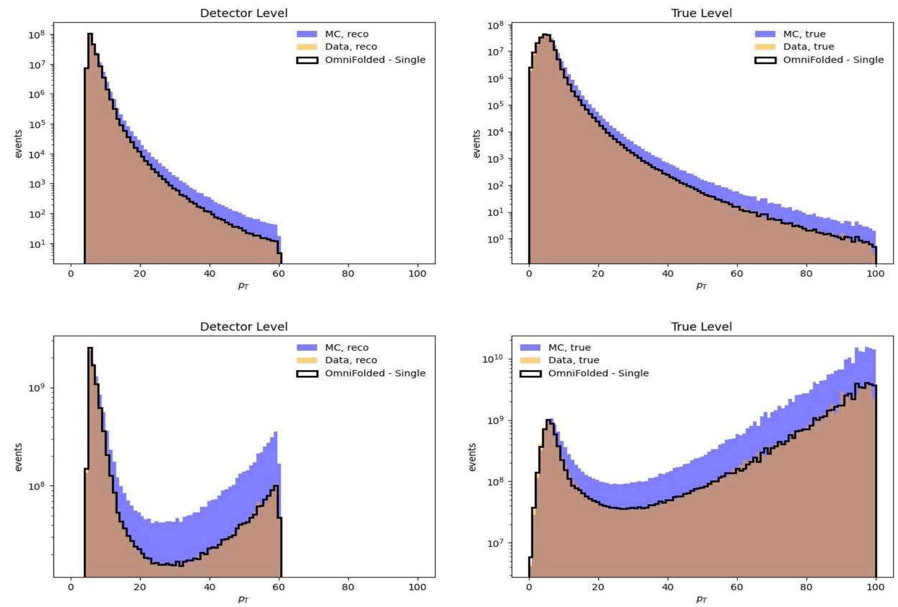
Steeply Falling Spectra

- We add all events in 25 GeV bins and compare the true spectra to the unfolding.
- Averages and standard deviations calculated from 5 tests.
- Consistently fits low momenta well.
  - High statistics/weight.
- Struggles fitting the high momenta tail.
  - Low statistics/weight.
- Better performance with less steep spectra.



Custom Weights

- Omnifold treats low momenta with more importance.
- How can we make Omnifold treat the tail with more importance?
- Give Omnifold event weights modified by a custom weighting function.
- Rescales weights to minimize the discrepancy between high and low momenta.



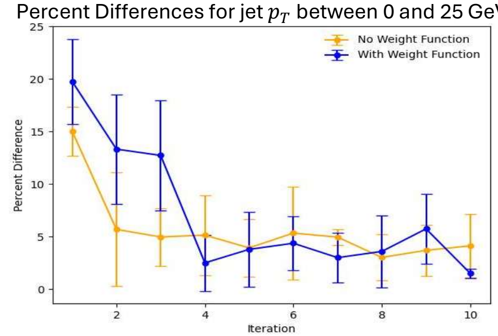
$$Weight_{new} = Weight_{old} \exp(p_T^{0.68})$$

Results

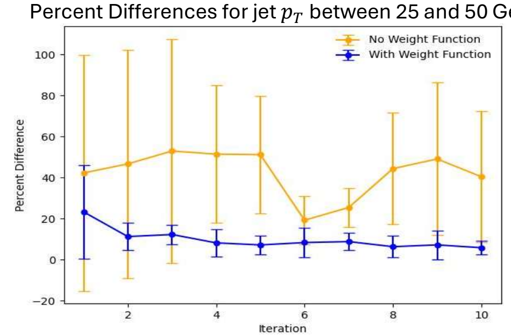
- Omnifold consistently fitted custom-weighted spectra more effectively.
- Lower average percent error and standard deviation than unaltered steeply falling spectra.

Averages and standard deviations calculated from 5 tests.

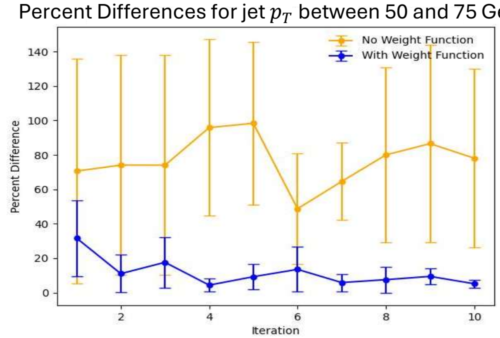
Percent Differences for jet p_T between 0 and 25 GeV



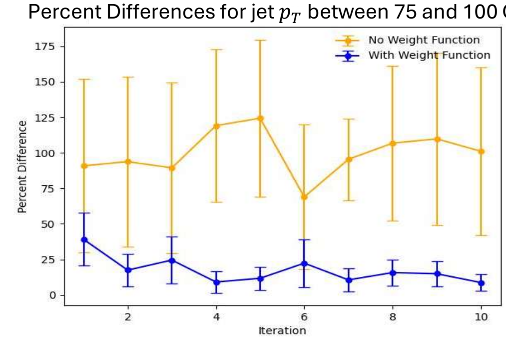
Percent Differences for jet p_T between 25 and 50 GeV



Percent Differences for jet p_T between 50 and 75 GeV



Percent Differences for jet p_T between 75 and 100 GeV



Future Work

Reweighting function must be custom made for each data set.

Find a reliable data fitting routine for steeply falling spectra.

Allows us to solve for the best reweighting function.

Investigation into uncertainty and error that this approach causes.

Acknowledgements

Special thanks to:

- Dr. Joseph Atchison
- Dr. Jim Drachenberg
- The Department of Energy
- The ACU Nuclear Physics Research Group
- The STAR Collaboration

This research was supported by US DOE MENP Grant DE-FG02-03ER41243.

Steeply Falling Spectra

$$\left(\exp(a_0 + a_1 * x) + \frac{a_2}{x^{a_3}}\right)\left(1 - \frac{1}{a_4 * x^{a_5} + 1}\right)$$

Where a_i are fit to simulate Run 11 embedding data, and can be varied to generate differing natural and synthetic distributions. Below are example parameter values for the natural distribution.

$$a_0 = 17.3, a_1 = -8.9\frac{1}{GeV}, a_2 = 2.4 * 10^{13}\ GeV^{a_3}, a_3 = 6.5, a_4 = 7.4 * 10^{-7} GeV^{-a_5}, a_5 = 8.2$$

$$\ln\left(\left(\exp(a_0 + a_1 * x) + \frac{a_2}{x^{a_3}}\right)\left(1 - \frac{1}{a_4 * x^{a_5} + 1}\right)\right)\left(\frac{x * GeV^{1.25}}{x^{2.25} + \Lambda^2}\right)^2$$

Where Λ is chosen in order to simulate Run 11 embedding data before weights are introduced.

$$\Lambda = 22\ \mathrm{Gev}^{1.125}$$