Combining Energy Correlators with Machine Learning

Katherine Fraser

Department of Physics
Harvard University
kfraser@g.harvard.edu

ML4Jets 2023

[240X.XXXX with A. Bhattacharya, M. Schwartz]

Outline

1. Review of Energy Correlators
2. Machine Learning Energy Correlator Space
    a. A Supervised Approach
    b. Normalizing Flows

Outline

1. Review of Energy Correlators
2. Machine Learning Energy Correlator Space
   a. A Supervised Approach
   b. Normalizing Flows

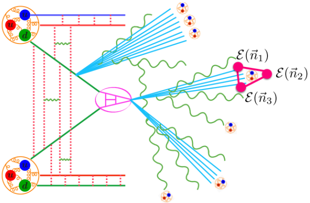
Defining Energy Correlators

Energy flow operators are a theoretical definition of calorimeter cells.

$$\varepsilon(\hat{n}) = \int_0^\infty dt \lim_{r\to\infty} r^2 n^i T_{0i}(t, r\vec{n})$$

Direction

Energy Momentum Tensor

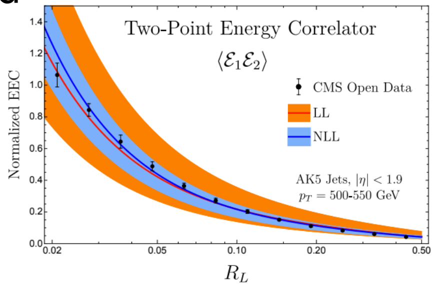


Energy correlators are correlation functions of these operators. They are not defined per event since each event gives multiple contributions.

[2004.11381: Hoffman, Maldecena; 2004.11381: Chen et al]

Energy Correlators for Jet Substructure

- Energy correlators in the collinear limit can be used to measure jet substructure.
- They have advantages over traditional jet observables, including being insensitive to soft radiation.



- In massless QCD jets, the correlation functions obey power-law scaling, but other scales also imprint in energy correlators.

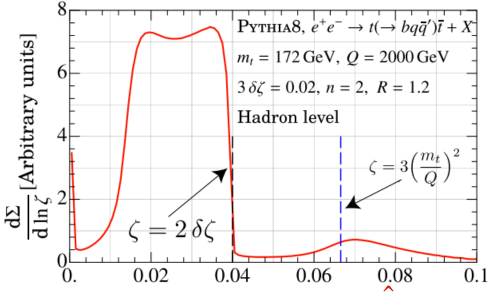
[2205.03414: Lee et al]
[2201.07800: Komiske et al]

Application: Top Mass using E3Cs

Sum over triplets

$$E3C(n) = \sum_{i<j<k} \int d\sigma \frac{E_i^n E_j^n E_k^n}{Q^{3n}} \delta_{\text{shape}}$$

Choose shape to collapse to 1d



$\longleftarrow$ $\hat{\zeta}_{ij} = (1 - \cos(\theta_{ij}))/2$

$$\delta_{\text{equilateral}} = \delta(3\zeta - \Sigma\zeta_{pair}) \prod \theta(\delta\zeta - |\zeta_{lm} - \zeta_{mn}|)$$

[2201.08393: Holguin et al]

Why Add ML?

To measure the Top mass, higher order ECs were collapsed down to one dimension.

The choice of equilateral triangles for this observable was convenient but arbitrary.

Can we avoid making these choices?

ML uses more of the information! We can:

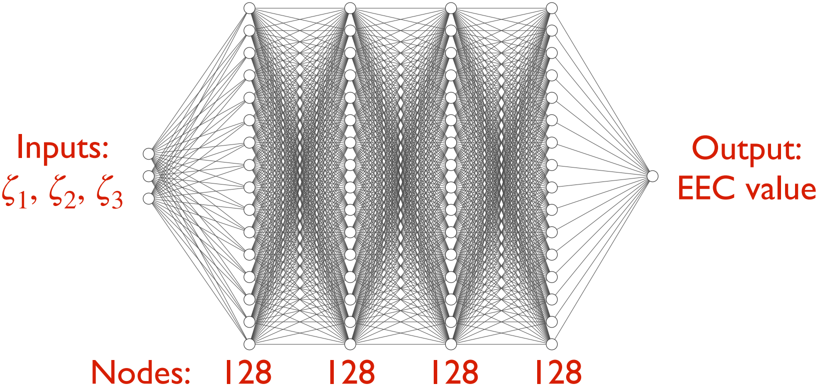
- Learn the full distribution & avoid shape choice
- Easily study multiple different shapes

Outline

1. Review of Energy Correlators
2. Machine Learning Energy Correlator Space
    a. A Supervised Approach
    b. Normalizing Flows

The Supervised Approach

The minimal way to learn the zeta histograms is interpolating the 3d histogram with a dense network



Training the Supervised Approach

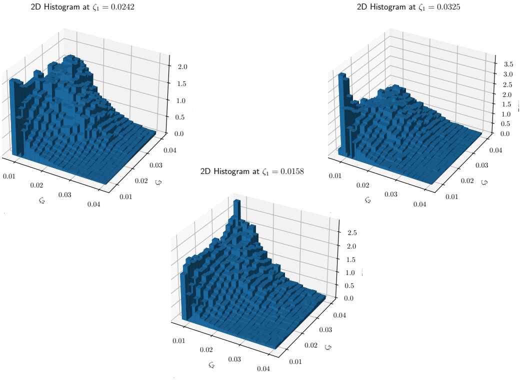
Inputs:

$\zeta_1, \zeta_2, \zeta_3$ values sampled from pythia data

```
tensor([[0.6863, 0.4768, 0.1022],
        [0.6863, 0.1870, 0.3579],
        [0.6863, 0.5797, 0.0271],
        [0.6863, 0.1903, 0.3486],
        [0.6863, 0.0221, 0.5742],
        [0.6863, 0.0680, 0.5477],
        [0.6863, 0.4893, 0.0440],
        [0.6863, 0.3960, 0.1274],
        [0.6863, 0.0815, 0.4047],
        [0.6863, 0.3630, 0.1059],
        [0.6863, 0.3014, 0.1514],
        [0.6863, 0.2518, 0.1899],
        [0.6863, 0.2869, 0.1627],
        [0.6863, 0.2914, 0.1572],
        [0.6863, 0.2883, 0.1606],
        [0.6863, 0.2826, 0.1638],
        [0.6863, 0.2781, 0.1693],
        [0.6863, 0.2814, 0.1658],
        [0.6863, 0.2764, 0.1701],
        [0.6863, 0.2666, 0.1779],
        [0.6863, 0.2697, 0.1750],
        [0.6863, 0.2742, 0.1713],
        [0.6863, 0.2681, 0.1771],
        [0.6863, 0.2728, 0.1728],
        [0.6863, 0.2709, 0.1745],
        [0.6863, 0.2708, 0.1745],
        [0.6863, 0.2391, 0.2028],
        [0.6863, 0.1837, 0.2608],
        [0.6863, 0.2449, 0.1964],
        [0.6863, 0.2216, 0.2199],
        [0.6863, 0.2304, 0.2097],
        [0.6863, 0.1872, 0.2552],
        [0.6863, 0.2194, 0.2207],
        [0.6863, 0.2107, 0.2294],
        [0.6863, 0.1804, 0.2625],
```

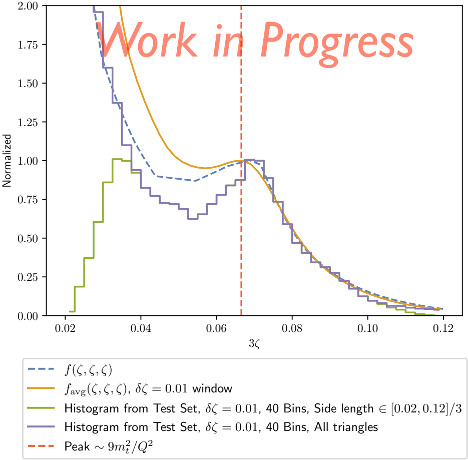
⋮

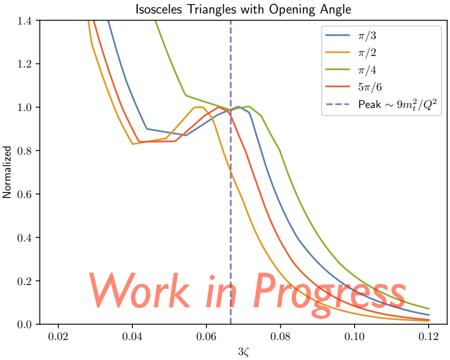
Output: EEC value at $\zeta_1, \zeta_2, \zeta_3$ for resampled histograms



To obtain the truth function for each batch, randomly sample half of 100k events and use the bin values of that histogram.

The Supervised Approach: Results



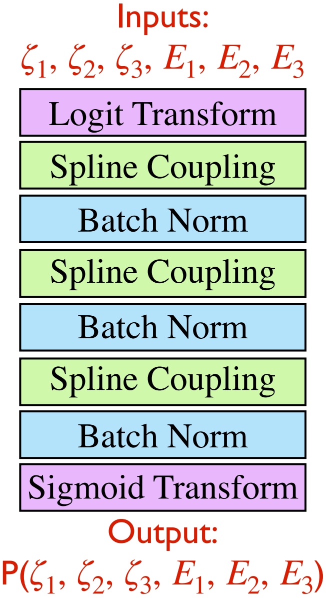


Outline

1. Review of Energy Correlators
2. Machine Learning Energy Correlator Space
    a. A Supervised Approach
    b. Normalizing Flows

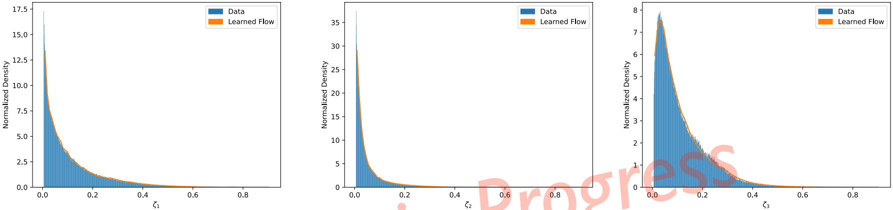
Another approach: Normalizing Flows

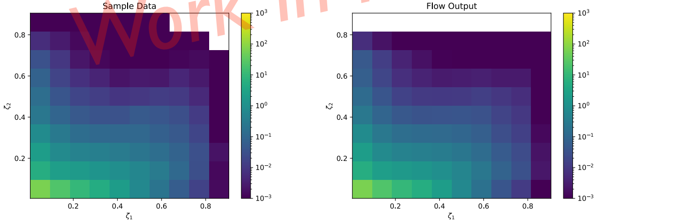
- To avoid binning, do simulation based inference with normalizing flows.
- We use a simple Spline Coupling based flow architecture where the split dimension increases per layer.
- Mask inputs with $\zeta < 0.005$ and train for 20 epochs with Adam.



Normalizing Flow Results I

We can check that the flow learns the ζ distributions, including their correlations:

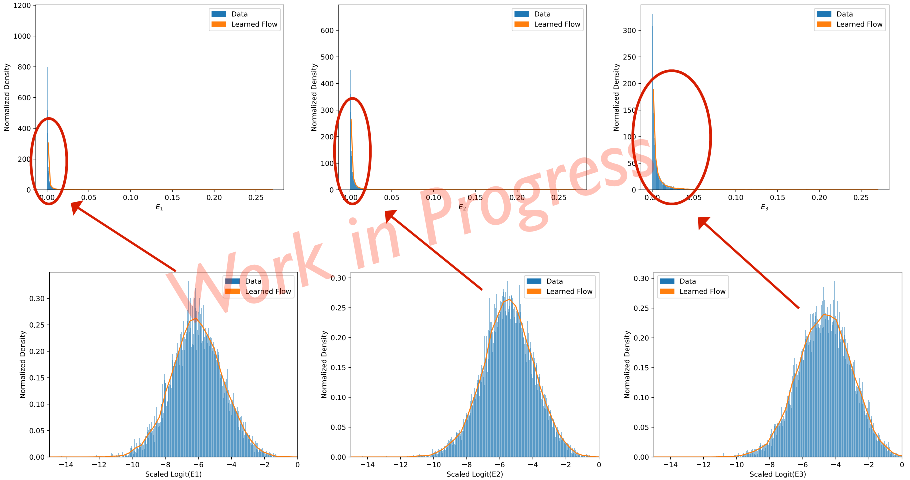




Work in Progress

Normalizing Flow Results II

We can also check the energy distributions:



Current Directions: Normalizing Flows

We have the six dimensional full distribution. Could use for comparing theory to directly data if the flow is trained on the data.

What else can we do with this distribution?

- Integrate over phase space to get different lower dimensional projections, including 1D shapes.
- Use as input to a NN for another task.
- Learn joint distribution with MC parameters, useful for Top Mass.

Summary

1. Machine learning can give us a more complete picture of energy correlator space.

2. Simple dense networks are very efficient but more sophisticated methods like flows also work.

3. Many potential ways to use these distributions, stay tuned for the finished study!