Analyzing Jet Substructure with Energy Flow

Elementary Particle Physics Journal Club

Eric M. Metodiev
Center for Theoretical Physics
Massachusetts Institute of Technology

Joint work with Patrick Komiske and Jesse Thaler

[1712.07124]  [1810.05165]  [19xx.xxxxx]
April 26, 2019
Jets from the Standard Model

++ = Mass from QCD Radiation

- t: 173 GeV++ (≈ 70%)
- H: 125 GeV (≈ 60%)
- W/Z: 80/91 GeV (≈ 70%)
- b: 4.2 GeV++
- c: 1.3 GeV++
- u,d,s: 100 MeV++
- g: 0++
Infrared (IR) safety – observable is unchanged under addition of a soft particle:

\[ S(\ldots) = S(\varepsilon\ldots) \]

Collinear (C) safety – observable is unchanged under collinear splitting of a particle:

\[ S(\ldots) = S(\ldots\lambda) \quad \forall \lambda \in [0,1] \]

IRC safety guarantees that the soft and collinear divergences of a QFT cancel at each order in perturbation theory (KLN theorem).

Divergences in QCD splitting function:

\[ dP_{i\rightarrow ig} \approx \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z} \]

\[ C_q = C_F = 4/3 \]

\[ C_g = C_A = 3 \]

IRC-safe observables probe hard structure while being insensitive to low energy or small angle modifications.
Outline

Energy Flow Polynomials
A basis of jet substructure observables

Energy Flow Moments
Tensor moments of the radiation pattern

Energy Flow Networks
ML architecture designed to learn from events
Analyzing Jet Substructure with Energy Flow

Outline

Energy Flow Polynomials
A basis of jet substructure observables

Energy Flow Moments
Tensor moments of the radiation pattern

Energy Flow Networks
ML architecture designed to learn from events
Expanding an Arbitrary IRC-safe Observable

Arbitrary IRC-safe observable: $S(p_1^\mu, ..., p_M^\mu)$

- **Energy expansion**: Approximate $S$ with polynomials of $Z_{ij}$
  - **IR safety**: $S$ is unchanged under addition of soft particle
  - **C safety**: $S$ is unchanged under collinear splitting of a particle
  - **Relabeling symmetry**: Particle index is arbitrary

Energy correlator parametrized by angular function $f$

$$
\sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M} Z_{i_1} \cdots Z_{i_N} f(\hat{p}_{i_1}, ..., \hat{p}_{i_N})
$$

[F. Tkachov, hep-ph/9601308]
Expanding an Arbitrary IRC-safe Observable

Arbitrary IRC-safe observable: \( S(p_1^\mu, \ldots, p_M^\mu) \)

- **Energy expansion**: Approximate \( S \) with polynomials of \( z_{ij} \)
  - **IR safety**: \( S \) is unchanged under addition of soft particle
  - **C safety**: \( S \) is unchanged under collinear splitting of a particle
  - **Relabeling symmetry**: Particle index is arbitrary

Energy correlator parametrized by angular function \( f \)

\[
\sum_{i_1=1}^{M} \ldots \sum_{i_N=1}^{M} z_{i_1} \ldots z_{i_N} f(\hat{p}_{i_1}, \ldots, \hat{p}_{i_N})
\]

Energy correlators linearly span IRC-safe observables

- **Angular expansion**: Approximate \( f \) with polynomials in \( \theta_{ij} \)
- **Simplify**: Identify unique analytic structure that emerge

[F.Tkachov, hep-ph/9601308]
Expanding an Arbitrary IRC-safe Observable

Arbitrary IRC-safe observable: \( S(p_1^\mu, \ldots, p_M^\mu) \)

- **Energy expansion:** Approximate \( S \) with polynomials of \( z_{ij} \)
  - **IR safety:** \( S \) is unchanged under addition of soft particle
  - **C safety:** \( S \) is unchanged under collinear splitting of a particle
  - **Relabeling symmetry:** Particle index is arbitrary

Energy correlator parametrized by angular function \( f \)

\[
\sum_{i_1=1}^M \ldots \sum_{i_N=1}^M z_{i_1} \ldots z_{i_N} f(\hat{p}_{i_1}, \ldots, \hat{p}_{i_N})
\]

Energy correlators linearly span IRC-safe observables

- **Angular expansion:** Approximate \( f \) with polynomials in \( \theta_{ij} \)
- **Simplify:** Identify unique analytic structure that emerge

Obtain linear spanning basis of Energy Flow Polynomials, "EFPs":

\[
S \simeq \sum_{g \in G} s_G \text{EFP}_G, \quad \text{EFP}_G \equiv \sum_{i_1=1}^M \ldots \sum_{i_N=1}^M z_{i_1} \ldots z_{i_N} \prod_{(k,\ell) \in G} \theta_{i_k i_\ell}
\]
Anatomy of an Energy Flow Polynomial:

In equations: \[ \text{EFP}_G = \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \cdots \sum_{i_N=1}^{M} \prod_{(k,l) \in G} \theta_{i_k i_l} \]

\[ z_{i_1} z_{i_2} \cdots z_{i_N} \]
Anatomy of an Energy Flow Polynomial:

In equations:

$$EFP_G = \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \cdots \sum_{i_N=1}^{M} z_{i_1} z_{i_2} \cdots z_{i_N} \prod_{(k,l) \in G} \theta_{i_k i_l}$$

In words:

- **Correlator**: Sum over all $N$-tuples of particle in the event
- **Energies**: Product of the $N$ energy fractions
- **Angles**: One $\theta_{i_k i_l}$ for each edge in $(k, l) \in G$
Anatomy of an Energy Flow Polynomial:

\[
\text{EFP}_G = \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \ldots \sum_{i_N=1}^{M} z_{i_1} z_{i_2} \ldots z_{i_N} \prod_{(k,l) \in G} \theta_{i_k i_l}
\]

**In equations:**

**In words:**

- **Correlator**
- **Energies**
- **Angles**

Sum over all \(N\)-tuples of particle in the event

Product of the \(N\) energy fractions

One \(\theta_{i_k i_l}\) for each edge in \((k, l) \in G\)

**In pictures:**

- \(j\) \(\rightarrow\) \(z_{ij}\)
- \(k\) \(\cdots\) \(l\) \(\rightarrow\) \(\theta_{i_k i_l}\)

(e.g.)

\[
\sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \sum_{i_3=1}^{M} \sum_{i_4=1}^{M} z_{i_1} z_{i_2} z_{i_3} z_{i_4} \theta_{i_1 i_2} \theta_{i_2 i_3} \theta_{i_3 i_4} \theta_{i_2 i_4}^2
\]

(any index labelling works)
Organization of the basis

EFPs are truncated by angular degree $d$, the order of the angular expansion.

Finite number at each order in $d$
All prime EFPs up to $d=5$

Exactly 1000 EFPs up to degree $d=7$

Image files for all of the prime EFP multigraphs up to $d = 7$ are available [here](#).
Familiar Jet Substructure Observables as EFPs

Scaled Jet Mass:
\[ \frac{m_j^2}{p_{Tj}^2} = \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} z_{i_1} z_{i_2} (\cosh \Delta y_{i_1i_2} - \cos \Delta \phi_{i_1i_2}) = \frac{1}{2} + \ldots \]

Jet Angularities:
\[ \lambda^{(\alpha)} = \sum_{i} z_{i}^{\alpha} \theta_{i}^{\alpha} \]
\[ \lambda^{(4)} = -\frac{3}{4} \]
\[ \lambda^{(6)} = -\frac{3}{2} + \frac{5}{8} \]

Energy Correlation Functions (ECFs):
\[ e_{N}^{(\beta)} = \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \ldots \sum_{i_N=1}^{M} z_{i_1} z_{i_2} \ldots z_{i_N} \prod_{k<l \in \{1, \ldots, N\}} \theta_{i_ki_l}^{\beta} \]

[Eric M. Metodiev, MIT]

Analyzing Jet Substructure with Energy Flow

and many more…
Jet Tagging Performance – **Quark vs. Gluon Jets**

ROC curves for *quark* vs. *gluon* jet tagging

![ROC curves for quark vs. gluon jet tagging](image)

- **N-subjettiness**: [J. Thaler, K. Van Tilburg, 1011.2268, 1108.2701]
- **N-subjettiness basis**: [K. Datta, A. Larkoski, 1704.08249]
- **QG CNNs**: [P. Komiske, EMM, M. Schwartz, 1612.01551]
- **ML/NN review**: [A. Larkoski, I. Moult, B. Nachman, 1709.04464]

Linear classification with EFPs is comparable to modern machine learning techniques
High $d$ EFPs are important
Convergence by $d \leq 7$

High $N$ EFPs are important
### Top Tagging Community Comparison

Community comparison of top tagging methods:

<table>
<thead>
<tr>
<th>Method</th>
<th>AUC</th>
<th>Acc</th>
<th>#Param</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN [16]</td>
<td>0.981</td>
<td>0.930</td>
<td>610k</td>
</tr>
<tr>
<td>ResNeXt [30]</td>
<td>0.984</td>
<td>0.936</td>
<td>1.46M</td>
</tr>
<tr>
<td>TopoDNN [18]</td>
<td>0.972</td>
<td>0.916</td>
<td>59k</td>
</tr>
<tr>
<td>Multi-body N-subjettiness 6 [24]</td>
<td>0.979</td>
<td>0.922</td>
<td>57k</td>
</tr>
<tr>
<td>Multi-body N-subjettiness 8 [24]</td>
<td>0.981</td>
<td>0.929</td>
<td>58k</td>
</tr>
<tr>
<td>TreeNiN [43]</td>
<td>0.982</td>
<td>0.933</td>
<td>34k</td>
</tr>
<tr>
<td>P-CNN</td>
<td>0.980</td>
<td>0.930</td>
<td>348k</td>
</tr>
<tr>
<td>ParticleNet [47]</td>
<td>0.985</td>
<td>0.938</td>
<td>498k</td>
</tr>
<tr>
<td>LBN [19]</td>
<td>0.981</td>
<td>0.931</td>
<td>705k</td>
</tr>
<tr>
<td>LoLa [22]</td>
<td>0.980</td>
<td>0.929</td>
<td>127k</td>
</tr>
<tr>
<td>Energy Flow Polynomials [21]</td>
<td>0.980</td>
<td>0.932</td>
<td>1k</td>
</tr>
<tr>
<td>Energy Flow Network [23]</td>
<td>0.979</td>
<td>0.927</td>
<td>82k</td>
</tr>
<tr>
<td>Particle Flow Network [23]</td>
<td>0.982</td>
<td>0.932</td>
<td>82k</td>
</tr>
<tr>
<td>GoaT</td>
<td>0.985</td>
<td>0.939</td>
<td>35k</td>
</tr>
</tbody>
</table>

[1902.09914]
Outline

Energy Flow Polynomials
A basis of jet substructure observables

Energy Flow Moments
Tensor moments of the radiation pattern

Energy Flow Networks
ML architecture designed to learn from events
Energy Flow Moments

\[ m^2 = \sum_{i=1}^{M} \sum_{j=1}^{M} E_i E_j n_i^\mu n_j^\mu \]

\[ = O(M^2) \]

Take:

\[ \theta_{ij} = n_i^\mu n_j^\mu = (1 - \hat{n}_i \cdot \hat{n}_j) \]

\[ = O(M) \]
Energy Flow Moments

\[ m^2 = \sum_{i=1}^{M} \sum_{j=1}^{M} E_i E_j n_i^\mu n_j^\mu = \left( \sum_{i=1}^{M} E_i n_i^\mu \right) \left( \sum_{j=1}^{M} E_j n_j^\mu \right) \]

\[ O(M^2) \]

\[ O(M) \]

Take: \( \theta_{ij} = n_i^\mu n_j^\mu \)

\[ I^{\mu_1 \mu_2 \cdots \mu_v} = \sum_{i=1}^{M} E_i n_i^{\mu_1} n_i^{\mu_2} \cdots n_i^{\mu_v} = 1 \cdots \]

[See also J. Donogue, F. Low, S-Y. Pi, 1979]
**Energy Flow Moments**

\[ m^2 = \sum_{i=1}^{M} \sum_{j=1}^{M} E_i E_j n_i^\mu n_j^\mu = \left( \sum_{i=1}^{M} E_i n_i^\mu \right) \left( \sum_{j=1}^{M} E_j n_j^\mu \right) \]

\( O(M^2) \)

\[ I^{\mu_1 \mu_2 \cdots \mu_v} = \sum_{i=1}^{M} E_i n_i^{\mu_1} n_i^{\mu_2} \cdots n_i^{\mu_v} = \]

\[ 1 \cdots v \]

\[ = \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \sum_{i_3=1}^{M} E_{i_1} E_{i_2} E_{i_3} \theta_{i_1 i_2}^2 \theta_{i_1 i_3}^2 \theta_{i_2 i_3}^2 \]

\[ = I^{\mu \nu \rho \sigma} I^{\tau}_{\mu \nu} I^{\rho \sigma \tau} \]

[See also J. Donogue, F. Low, S-Y. Pi, 1979]

Take: \( \theta_{ij} = n_i^\mu n_j^\mu \)
Energy Flow Moments

\[
I^{\mu_1 \mu_2 \cdots \mu_v} = \sum_{i=1}^{M} E_i n_i^{\mu_1} n_i^{\mu_2} \cdots n_i^{\mu_v} = \cdots
\]

Understand EFP relations:

\[
5! \times \mathcal{I}_{[\mu_1 \mu_2 \mu_3 \mu_4 \mu_5]}^{\mu_1 \mu_2} \mathcal{I}_{\mu_3 \mu_4 \mu_5}^{\mu_1} = 6 \times \left( \begin{array}{c}
\text{pentagon} \\
\end{array} \right) - 5 \times \left( \begin{array}{c}
\text{triangle} \\
\end{array} \right) = 0
\]

via Cayley-Hamilton and 3+1 dimensions

Counting symmetric kinematic polynomials:

THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES®

A226919 Asymptotic number of completely symmetric polynomials of degree n up to momentum conservation in the limit as the number of particles increases.

<table>
<thead>
<tr>
<th>Edges d</th>
<th>Loopless, Leafless Multigraphs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Connected</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>68</td>
</tr>
<tr>
<td>8</td>
<td>217</td>
</tr>
<tr>
<td>9</td>
<td>718</td>
</tr>
<tr>
<td>10</td>
<td>2,553</td>
</tr>
<tr>
<td>11</td>
<td>9,574</td>
</tr>
<tr>
<td>12</td>
<td>38,005</td>
</tr>
<tr>
<td>13</td>
<td>157,306</td>
</tr>
<tr>
<td>14</td>
<td>679,682</td>
</tr>
<tr>
<td>15</td>
<td>3,047,699</td>
</tr>
<tr>
<td>16</td>
<td>14,150,278</td>
</tr>
</tbody>
</table>

1, 0, 1, 2, 5, 11, 34, 87, 279 (list; graph; refs; listen; history; text; internal format)

OFFSET 0,4

COMMENTS The values for n >= 6 are only conjectural.

Table of n, a(n) for n=0..8.

Analyzing Jet Substructure with Energy Flow

Outline

- Energy Flow Polynomials
  *A basis of jet substructure observables*

- Energy Flow Moments
  *Tensor moments of the radiation pattern*

- Energy Flow Networks
  *ML architecture designed to learn from events*
Towards Machine Learning

A generic IRC-safe observable $O$ can be written via moments as:

$$O = F \left( \sum_{i=1}^{M} E_i n_i^{\mu_1} n_i^{\mu_2} \cdots n_i^{\mu_N} \right)$$

Idea: Let angular structure be generic function:

$$\Phi: \mathbb{R}^2 \to \mathbb{R}^\ell$$

$$F: \mathbb{R}^\ell \to \mathbb{R}$$

Approximate $\Phi$ and $F$ with neural networks to learn an observable.

Generalize beyond IRC safety?

Energy Flow Network
IRC safe

$$\text{EFN} = F \left( \sum_{i=1}^{M} E_i \Phi(\hat{n}_i) \right)$$

Particle Flow Network
IRC unsafe

$$\text{PFN} = F \left( \sum_{i=1}^{M} \Phi(E_i, n_i^{\mu}, \cdots) \right)$$

Many observables are easily interpreted in EFN and PFN language.

[P. Komiske, EMM, J. Thaler, 1810.05165]
Deep Sets

All permutation symmetric functions have an **additivity** similar to EFMs

**Theorem 7** Let \( f : [0, 1]^M \to \mathbb{R} \) be a permutation invariant continuous function iff it has the representation

\[
f(x_1, \ldots, x_M) = \rho \left( \sum_{m=1}^{M} \phi(x_m) \right)
\]

**Proof sketch:** Stone-Weierstrass theorem with elementary symmetric polynomials
Point Cloud


A frame from a Luminar LIDAR system
Point Cloud


[See also H. Qu, L. Gouskos, 1902.08570]  
An LHC event from the CMS Detector
Strategies to Process Jets

Images

Observables

e.g. [L. de Oliveira, et al., 1511.05190]
[K. Datta, A. Larkoski, 1704.08249]
[P.T. Komiske, EMM, J. Thaler, 1712.07124]
[P.T. Komiske, EMM, M.D. Schwartz, 1612.01551]
[M. Andrews, et al., 1902.08276]

Sequences

e.g. [T. Cheng, 1711.02633]
[K. Datta, A. Larkoski, 1704.08249]
[P.T. Komiske, EMM, J. Thaler, 1712.07124]
[G. Louppe, et al., 1702.00748]

Point Clouds

...
Classification Performance – Quark vs. Gluon Jets

PFN-ID compares favorably to other architectures and observables
EFN Latent Dimension Sweep – Quark vs. Gluon Jets

Performance saturates as latent dimension increases

IRC-unsafe information helpful

Adding particle type information helpful

PFN-ID: Full particle flavor info
($\pi^\pm$, $K^\pm$, $p$, $\bar{p}$, $n$, $\bar{n}$, $\gamma$, $K_L$, $e^\pm$, $\mu^\pm$)

PFN-Ex: Experimentally accessible info
($h^{\pm,0}$, $\gamma$, $e^\pm$, $\mu^\pm$)

PFN-Ch: Particle charge info

PFN: Four momentum information

EFN: IRC-safe latent space
What is being learned?

\[
\text{EFN} \left( \{ p_1^\mu, \ldots, p_M^\mu \} \right) = F \left( \sum_{i=1}^{M} E_i \overline{\Phi}(\hat{n}_i) \right) \quad \hat{n}_i = (y_i, \phi_i)
\]

Manifestly IRC-safe latent space

Calorimeter Images as EFN Filters

Radiation Moments as EFN Filters
What is being learned?

Learned EFN Filters
Simultaneous Visualization Strategy

Analyzing Jet Substructure with Energy Flow
Psychedelic Visualizations

Translated Azimuthal Angle $\phi$

Latent Dimension 16

Latent Dimension 32

Translated Rapidity $y$
Psychedelic Visualizations

Translated Azimuthal Angle $\phi$

Latent Dimension 64

Latent Dimension 128

Translated Rapidity $y$
Psychedelic Visualizations

Latent Dimension 256

Translated Azimuthal Angle $\phi$

Translated Rapidity $y$

$R$

$R/2$

$R/2$

$-R/2$

$-R$

$-R$

$0$

$0$
Psychedelic Visualizations

**Latent Dimension 256**

**EFN\textsubscript{256}: Quark vs. Gluon**

PYTHIA 8.230, $\sqrt{s} = 14$ TeV

$R = 0.4$, $p_T \in [500, 550]$ GeV

- Blue: Filter
- Red: Best fit, slope = 1.61

Emissions from quark or gluon are distributed according to:

$$dP_{i\to ig} \approx \frac{2\alpha_s C_i}{\pi} \frac{d\theta}{\theta} \frac{dz}{z}$$

Uniform pixelization in emission plane implies:

$$\frac{d\theta}{\theta} d\phi = \theta^{-2} dy d\phi$$

$$\ln \frac{A}{\pi R^2} = 2 \ln \frac{\theta}{R}$$
Psychedelic Visualizations

Latent Dimension 256

Log Radial Distance $\ln \frac{R}{\theta}$

Azimuthal Angle $\varphi$
Extracting New Analytic Observables

$\ell = 2$ latent space dimension has radially symmetric filters:

$$\mathcal{O}_1 = \sum_{i=1}^{M} z_i \Phi_1(\theta_i)$$

$$\mathcal{O}_2 = \sum_{i=1}^{M} z_i \Phi_2(\theta_i)$$
Extracting New Analytic Observables

\( \ell = 2 \) latent space dimension has radially symmetric filters:

\[
\mathcal{O}_1 = \sum_{i=1}^{M} z_i \Phi_1(\theta_i) \\
\mathcal{O}_2 = \sum_{i=1}^{M} z_i \Phi_2(\theta_i)
\]

Fit functions of the following forms:

\[
A_{r_0} = \sum_i z_i e^{-\theta_i^2/r_0^2}, \quad B_{r_1, \beta} = \sum_i z_i \ln(1 + \beta(\theta_i - r_1)) \Theta(\theta_i - r_1).
\]

\( A \) and \( B \) separate collinear and wide-angle regions of phase space, unlike traditional angularities which mix them.
Extracting New Analytic Observables

Can also visualize $F$ in the two-dimensional $(\mathcal{O}_1, \mathcal{O}_2)$ phase space

**Learned**

![Learned Observables](image)

Extract analytic form for $F$ as distance from a point:

$$C(A, B) = (A - a_0)^2 + (B - b_0)^2$$
Analyzing Jet Substructure with Energy Flow

Extracting New Analytic Observables

Can also visualize $F$ in the two-dimensional $(\mathcal{O}_1, \mathcal{O}_2)$ phase space.

Extract analytic form for $F$ as distance from a point:

$$C(A, B) = (A - a_0)^2 + (B - b_0)^2$$

Extracted $C, A, B$ do a good job of reproducing the learned $\mathcal{O}_1, \mathcal{O}_2, F$. 
Benchmarking New Analytic Observables

Extracted $C(A, B)$ performs nearly as well as $\text{EFN}_2$

Multivariate combination (BDT) of three other angularities does not do as well

Successfully reverse engineered what the machine learned
Analyzing Jet Substructure with Energy Flow

Outline

Energy Flow Polynomials
A basis of jet substructure observables

Energy Flow Moments
Tensor moments of the radiation pattern

Energy Flow Networks
ML architecture designed to learn from events
The End

Thank you!
Extra Slides
Connection with the Stress-Energy Operator

At the heart is the Energy Flow Operator:

\[
\hat{\mathcal{E}}(\hat{n}, v) = \lim_{t \to \infty} \hat{n}_i T^{0i}(t, vt\hat{n})
\]

in the \( \hat{n} \) direction at velocity \( v \)

Progress has been made in computing correlations of \( \hat{\mathcal{E}}(\hat{n}, v) \) in conformal field theory

[Eric M. Metodiev, MIT]

IRC-safe observables are built out of energy correlators:

\[
C_f = \sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M} Z_{i_1} \cdots Z_{i_N} f(\hat{p}_{i_1}, \cdots, \hat{p}_{i_N})
\]

Rigid energy structure

Arbitrary angular function \( f \)
Multigraph/EFP Correspondence

\[ \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \sum_{i_3=1}^{M} \sum_{i_4=1}^{M} z_{i_1} z_{i_2} z_{i_3} z_{i_4} \theta_{i_1 i_2} \theta_{i_2 i_3} \theta_{i_3 i_4} \theta_{i_2 i_4} \]

- **Multigraph** vs. **EFP**
- **Number of vertices** \( N \) \( \rightarrow \) **N-particle correlator**
- **Number of edges** \( d \) \( \rightarrow \) **Degree of angular monomial**
- **Treewidth + 1** \( \chi \) \( \rightarrow \) **Optimal VE Complexity**

- Connected \( \rightarrow \) Prime
- Disconnected \( \rightarrow \) Composite

E.g. Tree graph EFPs are \( O(M^2) \)!

Surprisingly efficient to compute.

Stay tuned… See P. Komiske’s talk.
Computational Complexity of EFPs

• Like other energy correlators, EFPs are naively $\mathcal{O}(M^N)$
• *Factorability* of summand in EFP formula can speed up computation

\[
\sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \sum_{i_3=1}^{M} \sum_{i_4=1}^{M} \sum_{i_5=1}^{M} \sum_{i_6=1}^{M} \sum_{i_7=1}^{M} \sum_{i_8=1}^{M} z_{i_1} z_{i_2} z_{i_3} z_{i_4} z_{i_5} z_{i_6} z_{i_7} z_{i_8} \theta_{i_1 i_2} \theta_{i_1 i_3} \theta_{i_1 i_4} \theta_{i_1 i_5} \theta_{i_1 i_6} \theta_{i_1 i_7} \theta_{i_1 i_8}
\]

\[
\mathcal{O}(M^8)
\]

\[
\sum_{i_1=1}^{M} z_{i_1} \left( \sum_{i_2=1}^{M} z_{i_2} \theta_{i_1 i_2} \right)^7
\]

\[
\mathcal{O}(M^2)
\]

Other algebraic simplifications are also possible by choosing parentheses wisely.
Energy Flow Networks

Many observables are easily interpreted in EFN language

\[
\text{EFN} \left( \{p_1^\mu, \ldots, p_M^\mu\} \right) = F \left( \sum_{i=1}^{M} E_i \Phi(n_i^\mu) \right)
\]

Manifestly IRC-safe latent space

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

Fully general latent space
Familiar Jet Substructure Observables as EFNs or PFNs

\[
\text{EFN} \left( \{ p_1^\mu, \ldots, p_M^\mu \} \right) = F \left( \sum_{i=1}^{M} z_i \Phi(\hat{p}_i) \right) \quad \text{Manifestly IRC-safe latent space}
\]

\[
\text{PFN} \left( \{ p_1^\mu, \ldots, p_M^\mu \} \right) = F \left( \sum_{i=1}^{M} \Phi(p_i^\mu) \right) \quad \text{Fully general latent space}
\]

<table>
<thead>
<tr>
<th>Observable $\mathcal{O}$</th>
<th>Map $\Phi$</th>
<th>Function $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>$m$</td>
<td>$p^\mu$</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>$M$</td>
<td>1</td>
</tr>
<tr>
<td>Track Mass</td>
<td>$m_{\text{track}}$</td>
<td>$p^\mu \mathbb{I}_{\text{track}}$</td>
</tr>
<tr>
<td>Track Multiplicity</td>
<td>$M_{\text{track}}$</td>
<td>$\mathbb{I}_{\text{track}}$</td>
</tr>
<tr>
<td>Jet Charge [69]</td>
<td>$Q_\kappa$</td>
<td>$(p_T, Q p_T^\kappa)$</td>
</tr>
<tr>
<td>Eventropy [71]</td>
<td>$z \ln z$</td>
<td>$(p_T, p_T \ln p_T)$</td>
</tr>
<tr>
<td>Momentum Dispersion [90]</td>
<td>$p_T^D$</td>
<td>$(p_T, p_T^2)$</td>
</tr>
<tr>
<td>$C$ parameter [91]</td>
<td>$C$</td>
<td>$(</td>
</tr>
</tbody>
</table>

Many observables are easily interpreted in EFN language.