

# Cutting Multiparticle Correlators Down to Size

**Patrick T. Komiske III**

Massachusetts Institute of Technology  
Center for Theoretical Physics

*with Eric Metodiev and Jesse Thaler, to appear soon*

BOOST 2019 – MIT

July 24, 2019

# Multiparticle Correlators

Sums of products of energies (transverse momenta) and angles

Definition of energy factor and pairwise angular distance

$$z_i = \frac{p_{Ti}}{\sum_j p_{Tj}} \quad \theta_{ij}^2 = 2n_i^\mu n_{j\mu} = 2 \frac{p_i^\mu}{p_{Ti}} \frac{p_{\mu j}}{p_{Tj}} \underset{\uparrow}{\simeq} (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

central, narrow jet approximation

Graphs represent correlators

vertex  $\leftrightarrow$  energy factor

$$\bullet_j \longleftrightarrow \sum_{i_j=1}^M z_{i_j}$$

edge  $\leftrightarrow$  pairwise angle

$$k \xrightarrow{\hspace{1cm}} l \longleftrightarrow \theta_{i_k i_l}$$

# Multiparticle Correlators

*Ubiquitous observables at the LHC*

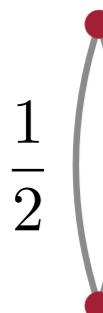
Definition of **energy factor** and pairwise angular distance

$$z_i = \frac{p_{Ti}}{\sum_j p_{Tj}} \quad \theta_{ij}^2 = 2n_i^\mu n_{j\mu} = 2 \frac{p_i^\mu}{p_{Ti}} \frac{p_{j\mu}}{p_{Tj}} \simeq (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

↑  
central, narrow jet approximation

Mass

$$\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M z_i z_j \theta_{ij}^2$$



# Multiparticle Correlators

*Ubiquitous observables at the LHC*

Definition of **energy factor** and pairwise angular distance

$$z_i = \frac{p_{Ti}}{\sum_j p_{Tj}} \quad \theta_{ij}^2 = 2n_i^\mu n_{j\mu} = 2 \frac{p_i^\mu}{p_{Ti}} \frac{p_{j\mu}}{p_{Tj}} \simeq (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

↑  
central, narrow jet approximation

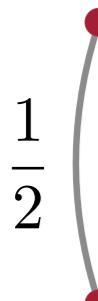
Mass

$$\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M z_i z_j \theta_{ij}^2$$

Energy Correlation Functions (ECFs)

$$\sum_{i_1=1}^M \cdots \sum_{i_N=1}^M z_{i_1} \cdots z_{i_N} \prod_{j < k} \theta_{i_j i_k}^\beta$$

[Larkoski, Salam, Thaler, [JHEP05\(2013\)007](#); Larkoski, Moult, Neill, [JHEP09\(2014\)029](#)]



Used for multi-prong tagging,  
typically in ratios ,  $D_2$ ,  $C_2$ ,  $C_3$ , etc.

Generalized ECFs also useful  
(angular part not monomial)

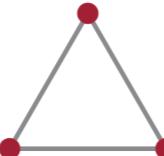
$N = 1:$



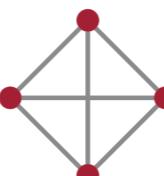
$N = 2:$



$N = 3:$



$N = 4:$



$$C_2 : \frac{\text{triangle area}}{(\text{base length})^2}$$

$$D_2 : \frac{\text{tetrahedron volume}}{(\text{base area})^3}$$

# Multiparticle Correlators

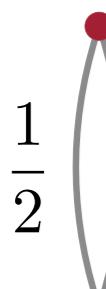
# *Ubiquitous observables at the LHC*

## Definition of energy factor and pairwise angular distance

Mass	Energy Correlation Functions (ECFs)	Energy Flow Polynomials (EFPs)
$\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M z_i z_j \theta_{ij}^2$	$\sum_{i_1=1}^M \cdots \sum_{i_N=1}^M z_{i_1} \cdots z_{i_N} \prod_{j < k} \theta_{i_j i_k}^\beta$	$\sum_{i_1=1}^M \cdots \sum_{i_N=1}^M z_{i_1} \cdots z_{i_N} \prod_{(j,k) \in G} \theta_{i_j i_k}^\beta$

[Larkoski, Salam, Thaler, [1305.0007](#); Larkoski, Moult, Neill, [1409.6298](#)]

[PTK, Metodiev, Thaler, 1712.07124]



Used for multi-prong tagging,  
typically in ratios ,  $D_2$ ,  $C_2$ ,  $C_3$ , etc.

Generalized ECFs also useful  
(angular part not monomial)

N = 1:



N = 2.



N = 3:



$$N = 4.$$



$$C_2 : \frac{1}{(r^2)^2}$$

$$D_2 : \frac{\text{---}}{(\text{---})}$$

## *Linear basis of all IRC-safe observables*

$$\mathcal{O} = \sum_G s_G \text{EFP}_G$$

Degree	Connected Multigraphs
$d = 0$	
$d = 1$	
$d = 2$	
$d = 3$	
$d = 4$	
$d = 5$	

# Multiparticle Correlators

## Ubiquitous observables at the LHC

Definition of **energy factor** and pairwise angular distance

$$z_i = \frac{p_{Ti}}{\sum_j p_{Tj}} \quad \theta_{ij}^2 = 2n_i^\mu n_{j\mu} = 2 \frac{p_i^\mu}{p_{Ti}} \frac{p_{\mu j}}{p_{Tj}} \simeq (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

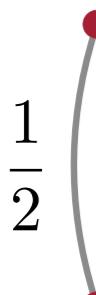
↑  
central, narrow jet approximation

Mass

$$\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M z_i z_j \theta_{ij}^2$$

Energy Correlation Functions (ECFs)

$$\sum_{i_1=1}^M \cdots \sum_{i_N=1}^M z_{i_1} \cdots z_{i_N} \prod_{j < k} \theta_{i_j i_k}^\beta$$



Used for multi-prong tagging,  
typically in ratios ,  $D_2$ ,  $C_2$ ,  $C_3$ , etc.

Generalized ECFs also useful  
(angular part not monomial)

Mass also calculated as

$$\left( \sum_{i=1}^M p_i^\mu \right)^2$$

which is  $O(M)$  to compute

What else is  $O(M)$ ?

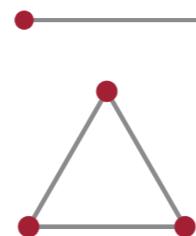
$N = 1$ :



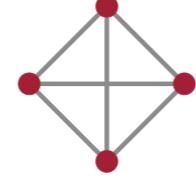
$N = 2$ :



$N = 3$ :



$N = 4$ :



$$C_2 : \frac{1}{(\bullet - \bullet)^2}$$

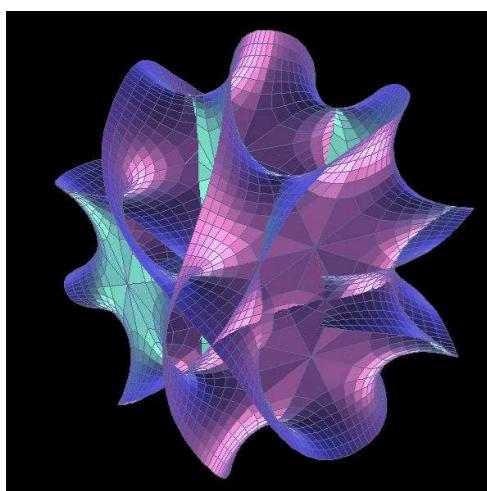
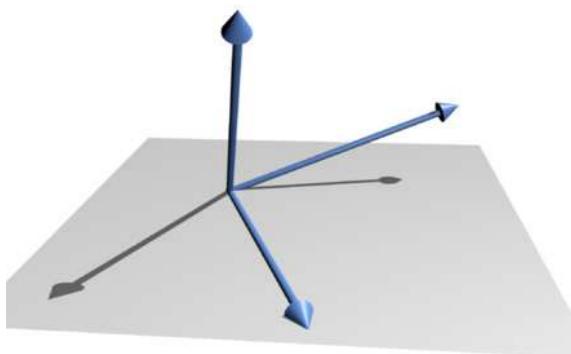
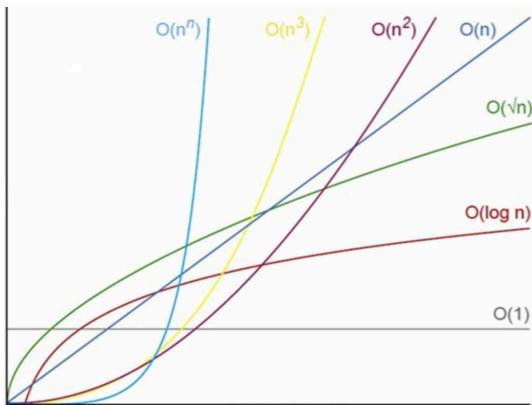
$$D_2 : \frac{1}{(\bullet - \bullet)^3}$$

Linear basis of all IRC-safe observables

$$\mathcal{O} = \sum_G s_G \text{EFP}_G$$

Degree	Connected Multigraphs
$d = 0$	.
$d = 1$	•
$d = 2$	•—•, •—•—•, •—•—•—•
$d = 3$	•—•—•, •—•—•—•, •—•—•—•—•, •—•—•—•—•—•, •—•—•—•—•—•—•
$d = 4$	•—•—•—•, •—•—•—•—•, •—•—•—•—•—•, •—•—•—•—•—•—•, •—•—•—•—•—•—•—•, •—•—•—•—•—•—•—•—•
$d = 5$	•—•—•—•—•, •—•—•—•—•—•, •—•—•—•—•—•—•, •—•—•—•—•—•—•—•—•, •—•—•—•—•—•—•—•—•—•, •—•—•—•—•—•—•—•—•—•—•

# Outline



## Computational Complexity

Multiparticle correlators are  $\mathcal{O}(M^N)$  to compute in general  
*Many can actually be computed in  $\mathcal{O}(M)$*

Experiment

## Linear Tensor Identities

Multiparticle correlators exhibit mysterious linear redundancies  
*All redundancies understood via cutting graphs*

## Counting Superstring Amplitudes

Counting independent kinematic polynomials difficult  
*Immediate enumeration through multigraphs*

Theory

# Computational Complexity – BOOST 2018

*Naive computation complexity of an energy correlator is  $\mathcal{O}(M^N)$*

EnergyCorrelator fjcontrib solution:

# Computational Complexity – BOOST 2018

*Naive computation complexity of an energy correlator is  $\mathcal{O}(M^N)$*

EnergyCorrelator fjcontrib solution:

```
// if N > 5, then throw error
if (_N > 5) {
    throw Error("EnergyCorrelator is only hard coded for N = 0,1,2,3,4,5");
}
```

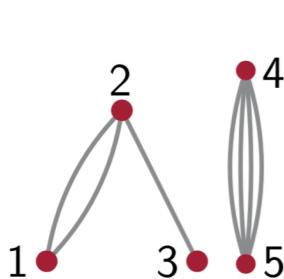
# Computational Complexity – BOOST 2018

*Naive computation complexity of an energy correlator is  $\mathcal{O}(M^N)$*

EnergyCorrelator fjcontrib solution:

```
// if N > 5, then throw error
if (_N > 5) {
    throw Error("EnergyCorrelator is only hard coded for N = 0,1,2,3,4,5");
}
```

Variable elimination (VE) algorithm:  $\mathcal{O}(M^\chi)$ ,  $\chi \lesssim N$



Disconnected is product of connected

$$= \left( \sum_{i_1=1}^M \sum_{i_2=1}^M \sum_{i_3=1}^M z_{i_1} z_{i_2} z_{i_3} \theta_{i_1 i_2}^2 \theta_{i_2 i_3} \right) \left( \sum_{i_4=1}^M \sum_{i_5=1}^M z_{i_4} z_{i_5} \theta_{i_4 i_5}^4 \right)$$

VE find clever parentheses placement to minimize computation



$$\begin{aligned} &= \underbrace{\sum_{i_1=1}^M \sum_{i_2=1}^M \sum_{i_3=1}^M \sum_{i_4=1}^M \sum_{i_5=1}^M \sum_{i_6=1}^M \sum_{i_7=1}^M \sum_{i_8=1}^M z_{i_1} z_{i_2} z_{i_3} z_{i_4} z_{i_5} z_{i_6} z_{i_7} z_{i_8}}_{\mathcal{O}(M^8)} \prod_{j=2}^7 \theta_{i_1 i_j} \\ &= \underbrace{\sum_{i_1=1}^M z_{i_1} \left( \sum_{i_2=1}^M z_{i_2} \theta_{i_1 i_2} \right)^7}_{\mathcal{O}(M^2)} \end{aligned}$$

All tree graphs become  $\mathcal{O}(M^2)$

$\chi = N$  iff  $G$  is complete graph, ECFs still slow

[PTK, Metodiev, Thaler, [1712.07124](#)]

Welcome to EnergyFlow

EnergyFlow is a Python package containing a suite of particle physics tools. Originally designed to compute Energy Flow Polynomials (EFPs), as of version [0.10.0](#) the package expanded to include implementations of Energy Flow Networks (EFNs) and Particle Flow Networks (PFNs). As of version [0.11.0](#), functions for facilitating the computation of the Energy Mover's Distance (EMD) on particle physics events are included. To summarize the main features:

- Energy Flow Polynomials:** EFPs are a collection of jet substructure observables which form a complete linear basis of IRC-safe observables. EnergyFlow provides tools to compute EFPs on events for several energy and angular measures as well as custom measures.
- Energy Flow Networks:** EFNs are infrared- and collinear-safe models designed for learning from collider events as unordered, variable-length sets of particles. EnergyFlow contains customizable Keras implementations of EFNs.
- Particle Flow Networks:** PFNs are general models designed for learning from collider events as unordered, variable-length sets of particles, based on the Deep Sets framework. EnergyFlow contains customizable Keras implementations of PFNs.
- Energy Mover's Distance:** The EMD is a common metric between probability distributions that has been adapted for use as a metric between collider events. EnergyFlow contains code to

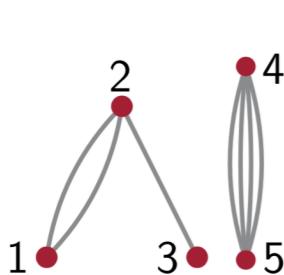
<https://energyflow.network>

# Computational Complexity – BOOST 2018

## BOOST 2019

*Can we do better – perhaps  $\mathcal{O}(M)$  as for mass?*

Variable elimination (VE) algorithm:  $\mathcal{O}(M^\chi)$ ,  $\chi \lesssim N$



Disconnected is product of connected

$$= \left( \sum_{i_1=1}^M \sum_{i_2=1}^M \sum_{i_3=1}^M z_{i_1} z_{i_2} z_{i_3} \theta_{i_1 i_2}^2 \theta_{i_2 i_3} \right) \left( \sum_{i_4=1}^M \sum_{i_5=1}^M z_{i_4} z_{i_5} \theta_{i_4 i_5}^4 \right)$$

VE find clever parentheses placement to minimize computation

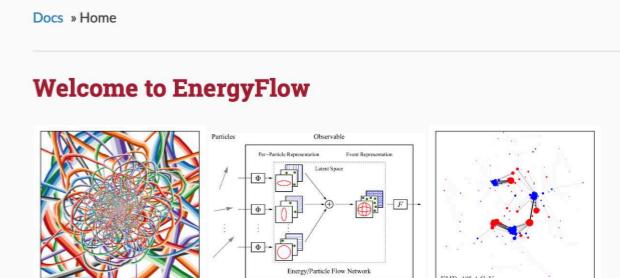


$$\begin{aligned} &= \underbrace{\sum_{i_1=1}^M \sum_{i_2=1}^M \sum_{i_3=1}^M \sum_{i_4=1}^M \sum_{i_5=1}^M \sum_{i_6=1}^M \sum_{i_7=1}^M \sum_{i_8=1}^M z_{i_1} z_{i_2} z_{i_3} z_{i_4} z_{i_5} z_{i_6} z_{i_7} z_{i_8}}_{\mathcal{O}(M^8)} \prod_{j=2}^7 \theta_{i_1 i_j} \\ &= \underbrace{\sum_{i_1=1}^M z_{i_1} \left( \sum_{i_2=1}^M z_{i_2} \theta_{i_1 i_2} \right)^7}_{\mathcal{O}(M^2)} \end{aligned}$$

All tree graphs become  $\mathcal{O}(M^2)$

$\chi = N$  iff  $G$  is complete graph, ECFs still slow

[PTK, Metodiev, Thaler, [1712.07124](#)]



EnergyFlow is a Python package containing a suite of particle physics tools. Originally designed to compute Energy Flow Polynomials (EFPs), as of version [0.10.0](#) the package expanded to include implementations of Energy Flow Networks (EFNs) and Particle Flow Networks (PFNs). As of version [0.11.0](#), functions for facilitating the computation of the Energy Mover's Distance (EMD) on particle physics events are included. To summarize the main features:

- **Energy Flow Polynomials:** EFPs are a collection of jet substructure observables which form a complete linear basis of IRC-safe observables. EnergyFlow provides tools to compute EFPs on events for several energy and angular measures as well as custom measures.
- **Energy Flow Networks:** EFNs are infrared- and collinear-safe models designed for learning from collider events as unsorted, variable-length sets of particles. EnergyFlow contains customizable Keras implementations of EFNs.
- **Particle Flow Networks:** PFNs are general models designed for learning from collider events as unsorted, variable-length sets of particles, based on the Deep Sets framework. EnergyFlow contains customizable Keras implementations of PFNs.
- **Energy Mover's Distance:** The EMD is a common metric between probability distributions that has been adapted for use as a metric between collider events. EnergyFlow contains code to

<https://energyflow.network>

# Energy Flow Moments (EFMs)

[PTK, Metodiev, Thaler, to appear soon]

$$\theta_{ij} = \sqrt{2n_i^\mu n_{j\mu}} \quad \beta = 2 \text{ removes square root}$$

Factors of  $n_i^\mu$  can be organized in optimal way

EFM<sub>v</sub> is a totally symmetric little group tensor

$$\mathcal{I}^{\mu_1 \dots \mu_v} = \sum_{i=1}^M z_i n_i^{\mu_1} \dots n_i^{\mu_v}$$

$v$	0	1	2	3	4	5	6
$n_{\text{components}}^{(d=4)}$	1	4	10	20	35	56	84

# Energy Flow Moments (EFMs)

[PTK, Metodiev, Thaler, to appear soon]

$$\theta_{ij} = \sqrt{2n_i^\mu n_{j\mu}} \quad \beta = 2 \text{ removes square root}$$

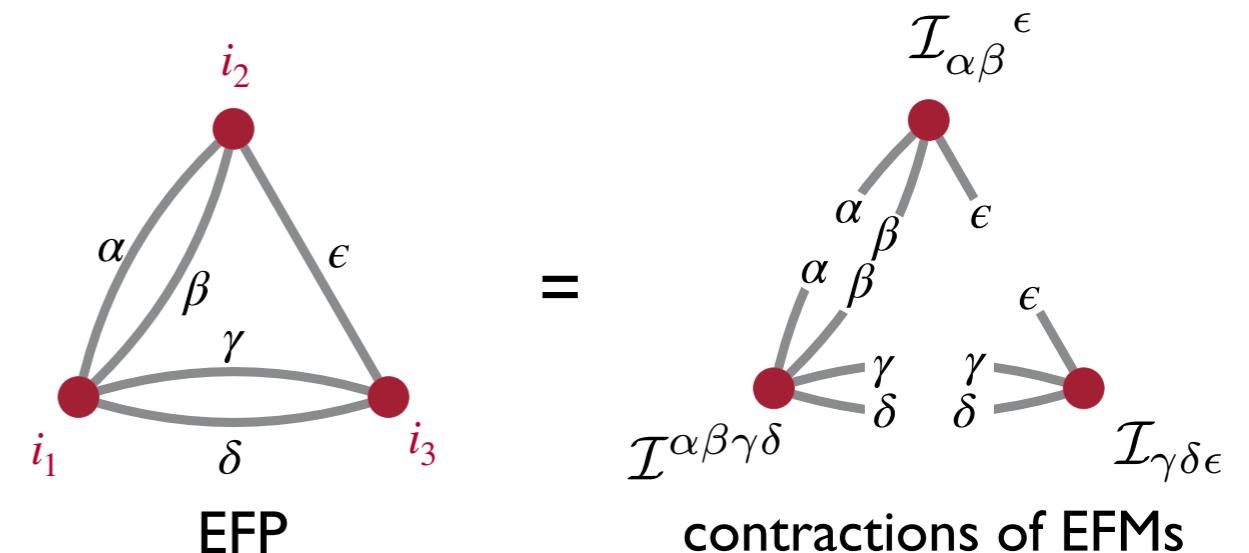
Factors of  $n_i^\mu$  can be organized in optimal way

EFM<sub>v</sub> is a totally symmetric little group tensor

$$\mathcal{I}^{\mu_1 \dots \mu_v} = \sum_{i=1}^M z_i n_i^{\mu_1} \dots n_i^{\mu_v}$$

$v$	0	1	2	3	4	5	6
$n_{\text{components}}^{(d=4)}$	1	4	10	20	35	56	84

EFMs result from cutting edges of EFP graph



# Energy Flow Moments (EFMs)

[PTK, Metodiev, Thaler, to appear soon]

$$\theta_{ij} = \sqrt{2n_i^\mu n_{j\mu}} \quad \beta = 2 \text{ removes square root}$$

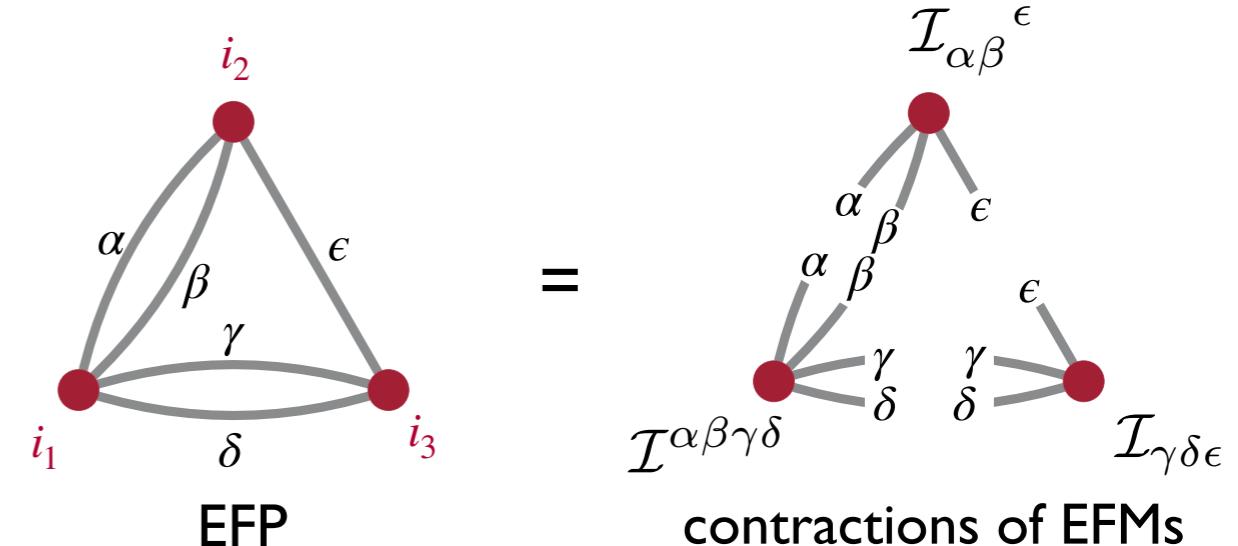
Factors of  $n_i^\mu$  can be organized in optimal way

EFM<sub>v</sub> is a totally symmetric little group tensor

$$\mathcal{I}^{\mu_1 \dots \mu_v} = \sum_{i=1}^M z_i n_i^{\mu_1} \dots n_i^{\mu_v}$$

$v$	0	1	2	3	4	5	6
$n_{\text{components}}^{(d=4)}$	1	4	10	20	35	56	84

EFMs result from cutting edges of EFP graph



$$\begin{aligned}
 &= \sum_{i_1=1}^M \sum_{i_2=1}^M \sum_{i_3=1}^M z_{i_1} z_{i_2} z_{i_3} \theta_{i_1 i_2}^2 \theta_{i_1 i_3}^2 \theta_{i_2 i_3} \\
 &= 2^5 \underbrace{\left( \sum_{i_1=1}^M z_{i_1} n_{i_1}^\alpha n_{i_1}^\beta n_{i_1}^\gamma n_{i_1}^\delta \right)}_{\mathcal{I}^{\alpha\beta\gamma\delta}} \underbrace{\left( \sum_{i_2=1}^M z_{i_2} n_{i_2}^\alpha n_{i_2}^\beta n_{i_2}^\epsilon \right)}_{\mathcal{I}_{\alpha\beta}^\epsilon} \underbrace{\left( \sum_{i_3=1}^M z_{i_3} n_{i_3}^\gamma n_{i_3}^\delta n_{i_3}^\epsilon \right)}_{\mathcal{I}_{\gamma\delta\epsilon}}
 \end{aligned}$$

Naively  $\mathcal{O}(M^3)$  EFP shown to be  $\mathcal{O}(M)$

# Energy Flow Moments (EFMs)

[PTK, Metodiev, Thaler, to appear soon]

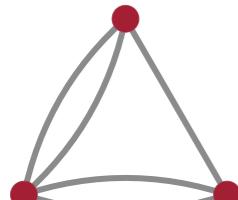
$$\theta_{ij} = \sqrt{2n_i^\mu n_{j\mu}} \quad \beta = 2 \text{ removes square root}$$

Factors of  $n_i^\mu$  can be organized in optimal way

EFM<sub>v</sub> is a totally symmetric little group tensor

$$\mathcal{I}^{\mu_1 \dots \mu_v} = \sum_{i=1}^M z_i n_i^{\mu_1} \dots n_i^{\mu_v}$$

$v$	0	1	2	3	4	5	6
$n_{\text{components}}^{(d=4)}$	1	4	10	20	35	56	84

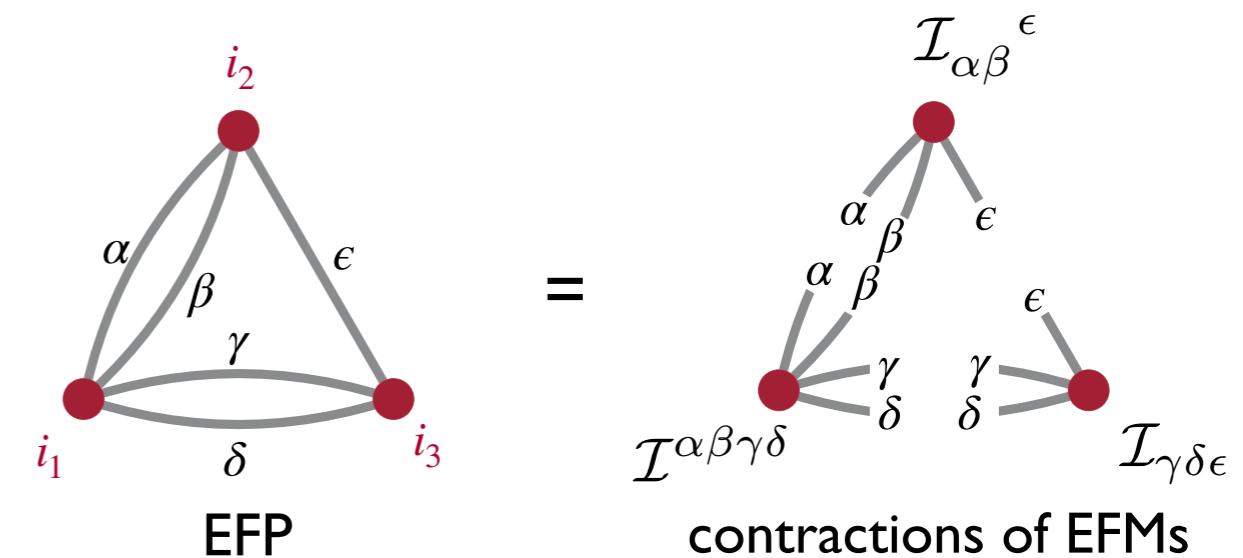


$$= \sum_{i_1=1}^M \sum_{i_2=1}^M \sum_{i_3=1}^M z_{i_1} z_{i_2} z_{i_3} \theta_{i_1 i_2}^2 \theta_{i_1 i_3}^2 \theta_{i_2 i_3}^2$$

$$= 2^5 \underbrace{\left( \sum_{i_1=1}^M z_{i_1} n_{i_1}^\alpha n_{i_1}^\beta n_{i_1}^\gamma n_{i_1}^\delta \right)}_{\mathcal{I}^{\alpha\beta\gamma\delta}} \underbrace{\left( \sum_{i_2=1}^M z_{i_2} n_{i_2}^\epsilon \right)}_{\mathcal{I}_{\alpha\beta}^\epsilon} \underbrace{\left( \sum_{i_3=1}^M z_{i_3} n_{i_3}^\epsilon n_{i_3}^\gamma n_{i_3}^\delta n_{i_3}^\epsilon \right)}_{\mathcal{I}_{\gamma\delta\epsilon}}$$

Naively  $\mathcal{O}(M^3)$  EFP shown to be  $\mathcal{O}(M)$

EFMs result from cutting edges of EFP graph



All  $\beta = 2$  EFPs are  $\mathcal{O}(M)$

$\text{ECF}_N^{(\beta=2)}$  are all  $\mathcal{O}(M)$   
 $D_2^{(\beta=2)}, C_2^{(\beta=2)}$  are  $\mathcal{O}(M)$

# Linear Tensor Identities

Linear redundancies among EFPs are troublesome

Studying coefficients of linear fit difficult

$$\mathcal{O} = \sum_G s_G \text{EFP}_G$$

Examples of redundancies

in 3 or fewer spacetime dimensions

$$0 = 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$
$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 12 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 4 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - 3 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$

in 4 or fewer spacetime dimensions

$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 5 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$

# Linear Tensor Identities

Linear redundancies among EFPs are troublesome

Studying coefficients of linear fit difficult

$$\mathcal{O} = \sum_G s_G \text{EFP}_G$$

Examples of redundancies

in 3 or fewer spacetime dimensions

$$0 = 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 12 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 4 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - 3 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$

in 4 or fewer spacetime dimensions

$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 5 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$

## Tensor Identity Recipe

Consider tensor over  $n$  dimensional vector space

Antisymmetrize  $m > n$  indices

Result is zero because any assignment of  $n$  possible values to  $m$  slots has a repetition

$$T_{b_1 \dots b_\ell [c_1 \dots c_m]}^{a_1 \dots a_k} = 0$$

Bonus: all tensor identities up to ones governed by existing symmetries take above form  
[Sneddon, *Journal of Mathematical Physics*]

# Linear Tensor Identities

Linear redundancies among EFPs are troublesome

Studying coefficients of linear fit difficult

$$\mathcal{O} = \sum_G s_G \text{EFP}_G$$

Examples of redundancies

in 3 or fewer spacetime dimensions

$$0 = 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} - \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \quad \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \iff 0 = \mathcal{I}_{[\alpha}^{\beta} \mathcal{I}_{\beta}^{\gamma} \mathcal{I}_{\gamma}^{\delta} \mathcal{I}_{\delta]}^{\alpha}$$

$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} - 12 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} + 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \quad \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 4 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} - 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} - 3 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \quad \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \iff 0 = \mathcal{I}_{[\alpha} \mathcal{I}_{\beta}^{\alpha} \mathcal{I}_{\gamma}^{\beta} \mathcal{I}_{\delta]}^{\gamma} \mathcal{I}^{\delta}$$

in 4 or fewer spacetime dimensions

$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} - 5 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - \end{array} + \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \quad \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \iff 0 = \mathcal{I}_{[\alpha}^{\beta} \mathcal{I}_{\beta}^{\gamma} \mathcal{I}_{\gamma}^{\delta} \mathcal{I}_{\delta]}^{\epsilon} \mathcal{I}_{\epsilon}^{\alpha}$$

## Tensor Identity Recipe

Consider tensor over  $n$  dimensional vector space

Antisymmetrize  $m > n$  indices

Result is zero because any assignment of  $n$  possible values to  $m$  slots has a repetition

$$T_{b_1 \dots b_\ell [c_1 \dots c_m]}^{a_1 \dots a_k} = 0$$

Bonus: all tensor identities up to ones governed by existing symmetries take above form

[Sneddon, *Journal of Mathematical Physics*]

# Linear Tensor Identities

Linear redundancies among EFPs are troublesome

Studying coefficients of linear fit difficult

$$\mathcal{O} = \sum_G s_G \text{EFP}_G$$

Examples of redundancies

in 3 or fewer spacetime dimensions

$$0 = 2 \begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet - \end{array} - \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \iff 0 = \mathcal{I}_{[\alpha}^{\beta} \mathcal{I}_{\beta}^{\gamma} \mathcal{I}_{\gamma}^{\delta} \mathcal{I}_{\delta]}^{\alpha}$$

$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 12 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet + 6 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + 4 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 2 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet - 3 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} \iff 0 = \mathcal{I}_{[\alpha} \mathcal{I}_{\beta}^{\alpha} \mathcal{I}_{\gamma}^{\beta} \mathcal{I}_{\delta]}^{\gamma} \mathcal{I}^{\delta}$$

in 4 or fewer spacetime dimensions

$$0 = 6 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \quad \diagup \\ \bullet - 5 \end{array} \begin{array}{c} \bullet \\ \diagup \quad \diagdown \quad \diagup \\ \bullet \end{array} \iff 0 = \mathcal{I}_{[\alpha}^{\beta} \mathcal{I}_{\beta}^{\gamma} \mathcal{I}_{\gamma}^{\delta} \mathcal{I}_{\delta}^{\epsilon} \mathcal{I}_{\epsilon]}^{\alpha}$$

## Tensor Identity Recipe

Consider tensor over  $n$  dimensional vector space

Antisymmetrize  $m > n$  indices

Result is zero because any assignment of  $n$  possible values to  $m$  slots has a repetition

$$T_{b_1 \dots b_\ell [c_1 \dots c_m]}^{a_1 \dots a_k} = 0$$

Bonus: all tensor identities up to ones governed by existing symmetries take above form

[Sneddon, *Journal of Mathematical Physics*]

Other types of identities – e.g. when  $M$  is small

$$M \leq 2$$

$$0 = \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}, \quad 0 = 2 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$

$$0 = \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} + \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array} - \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \end{array}$$

Could be useful in a partonic calculation, more in backup

# Counting Superstring Amplitudes

Constructing a basis of amplitudes – how large is it?

[Boels, [I304.7918](#); OEIS [A226919](#)]

non-isomorphic multigraph



*Q: What is the number of symmetric polynomials of degree  $d$  in kinematic variables  $s_{ij} = p_i \cdot p_j$  up to momentum conservation?*



$$\theta_{ij}^2 = 2n_i \cdot n_j$$

$$0 = \sum_{i=1}^M p_i^\mu = \mathcal{I}^\mu =$$


**A:** Same as the number of non-isomorphic multigraphs with no leaves (vertices of valency one)

# Counting Superstring Amplitudes

Constructing a basis of amplitudes – how large is it?

[Boels, [I304.7918](#); OEIS [A226919](#)]

non-isomorphic multigraph



*Q: What is the number of symmetric polynomials of degree  $d$  in kinematic variables  $s_{ij} = p_i \cdot p_j$  up to momentum conservation?*



$$\theta_{ij}^2 = 2n_i \cdot n_j$$

$$0 = \sum_{i=1}^M p_i^\mu = \mathcal{I}^\mu =$$



**A:** Same as the number of non-isomorphic multigraphs with no leaves (vertices of valency one)

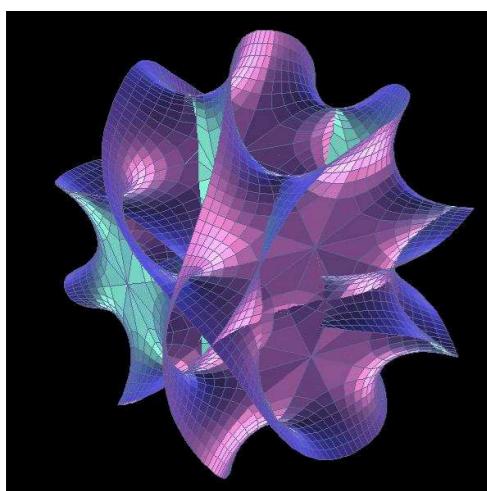
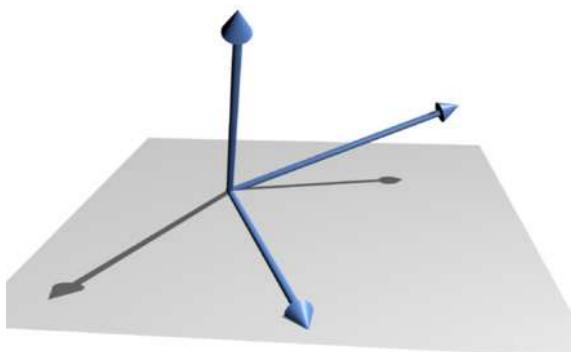
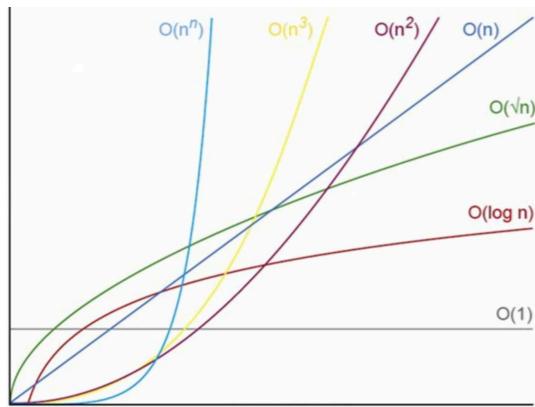
New OEIS Entries!  
[A307317](#), [A307316](#)

[PTK, Metodiev, Thaler, to appear soon]

Edges $d$	Leafless Multigraphs	
	Connected A307317	All A307316
1	0	0
2	1	1
3	2	2
4	4	5
5	9	11
6	26	34
7	68	87
8	217	279
9	<b>718</b>	<b>897</b>
10	<b>2 553</b>	<b>3 129</b>
11	<b>9 574</b>	<b>11 458</b>
12	<b>38 005</b>	<b>44 576</b>
13	<b>157 306</b>	<b>181 071</b>
14	<b>679 682</b>	<b>770 237</b>
15	<b>3 047 699</b>	<b>3 407 332</b>
16	<b>14 150 278</b>	<b>15 641 159</b>

Bolded values previously unknown

# Summary



## Computational Complexity

Multiparticle correlators are  $\mathcal{O}(M^N)$  to compute in general

$\beta = 2$  EFPs can be computed in  $\mathcal{O}(M)$

Why not use  $D_2^{(\beta=2)}$ ? Performance in backup

Experiment

## Linear Tensor Identities

Multiparticle correlators exhibit mysterious linear redundancies

All redundancies understood via cutting graphs  
and applying master antisymmetrization identity

## Counting Superstring Amplitudes

Counting independent kinematic polynomials difficult

Immediate enumeration through multigraphs  
and new OEIS sequences!



# Rewriting General EFP as Contraction of EFMs

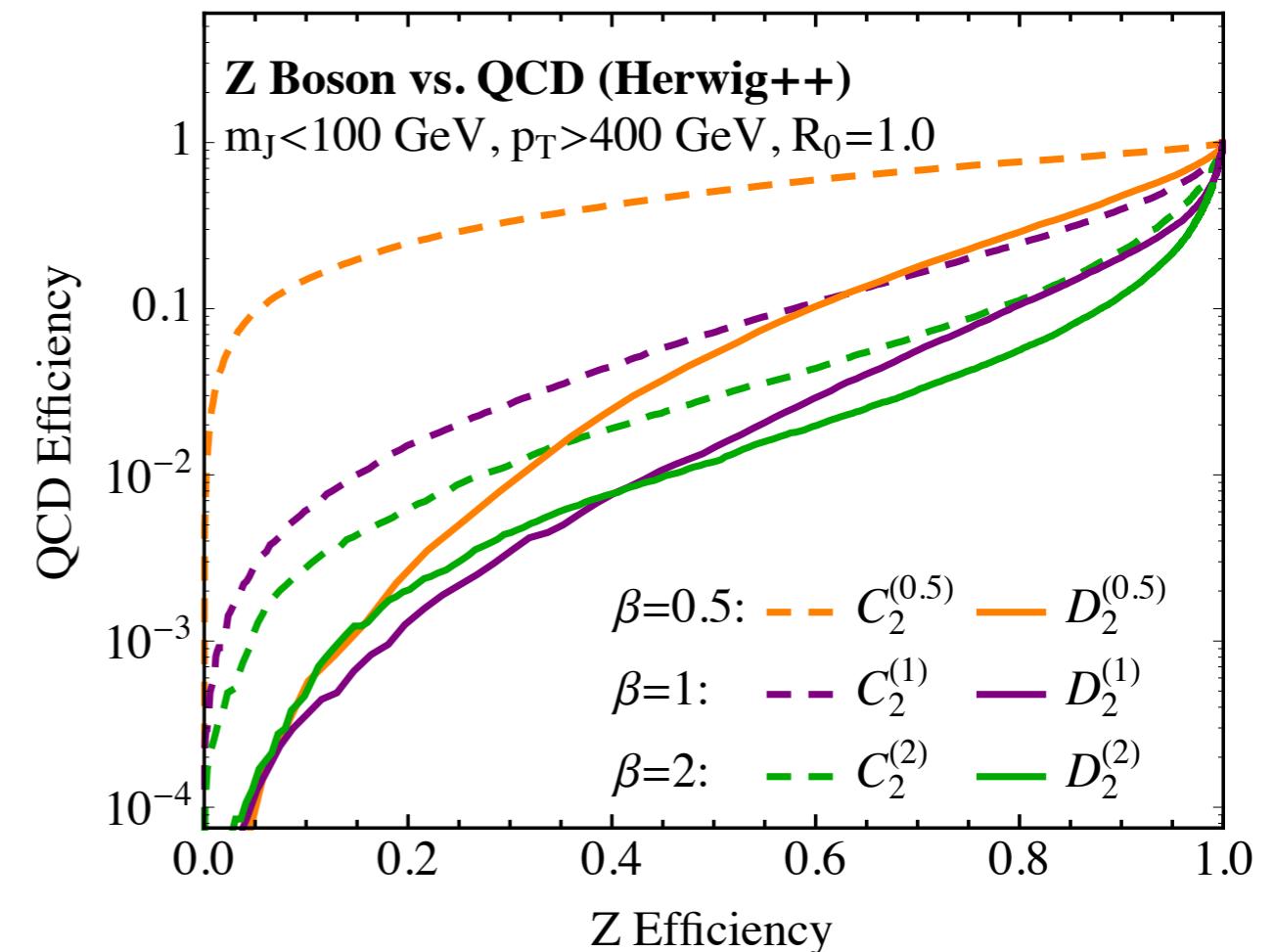
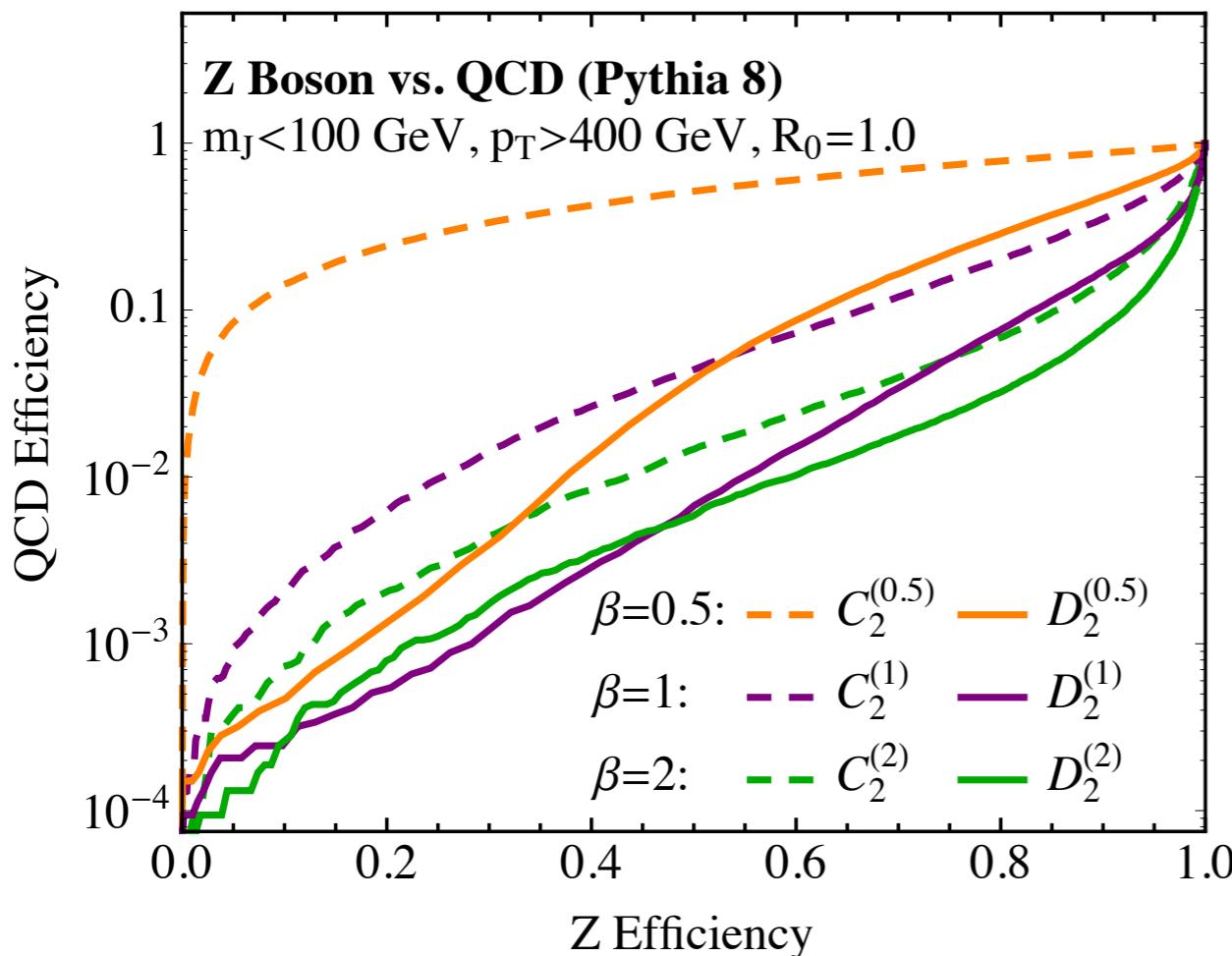
$$\begin{aligned}
 \text{EFP}_G &= \sum_{i_1=1}^M \cdots \sum_{i_N=1}^M z_{i_1} \cdots z_{i_N} \prod_{(k,\ell) \in G} 2\eta_{\mu\nu} n_{i_k}^\mu n_{i_\ell}^\nu \\
 &= \left( \prod_{j=1}^N \sum_{i_j=1}^M z_{i_j} n_{i_j}^{\mu_1^j} n_{i_j}^{\mu_2^j} \cdots n_{i_j}^{\mu_{v_j}^j} \right) \prod_{(k,\ell) \in G} 2\eta_{\mu_{A_{k\ell}}^k \mu_{A_{\ell k}}^\ell} \\
 &= \underbrace{\left( \prod_{j=1}^N \mathcal{I}^{\mu_1^j \mu_2^j \cdots \mu_{v_j}^j} \right)}_{\text{EFMs}} \underbrace{\prod_{(k,\ell) \in G} 2\eta_{\mu_{A_{k\ell}}^k \mu_{A_{\ell k}}^\ell}}_{\text{Contraction of edges}}
 \end{aligned}$$

TL;DR – edges of EFP with  $\beta = 2$  can be cut and rearranged into EFMs

# Two-Prong Classification with Varying $\beta$

$\beta = 2$  for both  $D_2$  and  $C_2$  for both Pythia 8 and Herwig++  
works better than  $\beta = 1$  for Z vs. QCD

[Larkoski, Moult, Neill, [1409.6298](#)]



# Additional Linear Identities

All identities fundamentally due to antisymmetrizing over more indices than dimensions

Finite spacetime dimension –  $\beta = 2$  only

$$0 = 2 \begin{array}{c} \text{diamond graph} \\ \text{with 4 red vertices} \end{array} - \begin{array}{c} \text{two ovals} \end{array} \quad d \leq 3$$

$$0 = 6 \begin{array}{c} \text{pentagon graph} \\ \text{with 5 red vertices} \end{array} - 5 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} - \begin{array}{c} \text{oval} \end{array} \quad d \leq 4$$

Finite particle number – cutting open vertices

"hamburger tensors"

$$\mathcal{M}_G = \sqrt{z_{i_1} \cdots z_{i_N}} \prod_{(k,\ell) \in G} \theta_{i_k i_\ell}$$

graph rules

$$\textcolor{red}{\bullet}_i = \sqrt{z_i}, \quad j \text{---} k = \theta_{jk}$$

$$\begin{array}{c} i_1 \\ \backslash \quad / \\ i_2 \quad i_3 \end{array} = \sqrt{z_{i_1} z_{i_2} z_{i_3}} \theta_{i_1 i_2}^3 \theta_{i_1 i_3}$$

Identities come from antisymmetrizing over  $M + 1$  or more vertex indices, works for any  $\theta_{jk}$ !

Euclidean subslicing –  $e^+ e^-$  only

$$e^+ e^- : \quad n_i^\mu = (1, \hat{n})^\mu$$

Presence of 1 means that  $d$  dim. tensors are exactly related to  $d - 1$  dim tensors and hence satisfy more identities

ex. – holds in  $d \leq 4$  for  $e^+ e^-$

$$\begin{aligned} 0 &= 6 \begin{array}{c} \text{diamond graph} \\ \text{with 4 yellow vertices} \end{array} - 16 \begin{array}{c} \text{triangle graph} \\ \text{with 3 yellow vertices} \end{array} - 3 \begin{array}{c} \text{two ovals} \\ \text{with 2 yellow vertices each} \end{array} + 24 \begin{array}{c} \text{oval} \\ \text{with 1 yellow vertex} \end{array} - 16, \\ 0 &= 6 \begin{array}{c} \text{pentagon graph} \\ \text{with 5 yellow vertices} \end{array} - 12 \begin{array}{c} \text{triangle graph} \\ \text{with 3 yellow vertices} \end{array} - 3 \begin{array}{c} \text{two ovals} \\ \text{with 2 yellow vertices each} \end{array} - 2 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} + 12 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} + 6 \begin{array}{c} \text{two ovals} \\ \text{with 2 red vertices each} \end{array} - 8, \\ 0 &= 6 \begin{array}{c} \text{diamond graph} \\ \text{with 4 red vertices} \end{array} + 16 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} - 3 \begin{array}{c} \text{two ovals} \\ \text{with 2 red vertices each} \end{array} - 48 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} + 24 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array}, \\ 0 &= 6 \begin{array}{c} \text{pentagon graph} \\ \text{with 5 red vertices} \end{array} - 12 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} - 3 \begin{array}{c} \text{two ovals} \\ \text{with 2 red vertices each} \end{array} - 2 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} + 4 \begin{array}{c} \text{triangle graph} \\ \text{with 3 red vertices} \end{array} + 6 \begin{array}{c} \text{two ovals} \\ \text{with 2 red vertices each} \end{array}. \end{aligned}$$

# Additional Linear Relation Material

		Lorentzian graphs							
		$d = 0$	$d = 1$	$d = 2$	$d = 3$				
		•				...			
•	1								
	2	-1							
	4	-4	1						
	4	-4	1						
	4	-4		1					
	8	-12	6		-1				
	8	-12	6			-1			
	8	-12	2	4			-1		
	8	-12	6				-1		
	8	-12	4	2				-1	
	8	-12	2	4					-1
	8	-12	2	4					
	8	-12		6					-1
	16	-32	24		-8				1
	16	-32	8	16			-8		
	16	-32	4	20			-4	-4	
	16	-32	12	12	-2		-6		
	16	-32		16	8			-8	
	16	-32		20	4	-2		-2	-4
	16	-32	4	16	4		-4		-4
	16	-32	4	12	8	-2		-4	-2
	16	-32	4	20		-4		-4	
	16	-32	24			-8			
	16	-32	16	8		-2	-4		-2
	16	-32	12	12		-4		-4	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Cutting vertices demonstrates matrix multiplication can be used to calculate some graphs

$$= \text{tr} \left( \begin{array}{c} i_1 \\ \vdots \\ i_2 \end{array} \right)^3$$

$$\sim \mathcal{O}(M^{2.3728639})$$

Works for all angular measures

