A Closed-Form, Differentiable Metric on Events

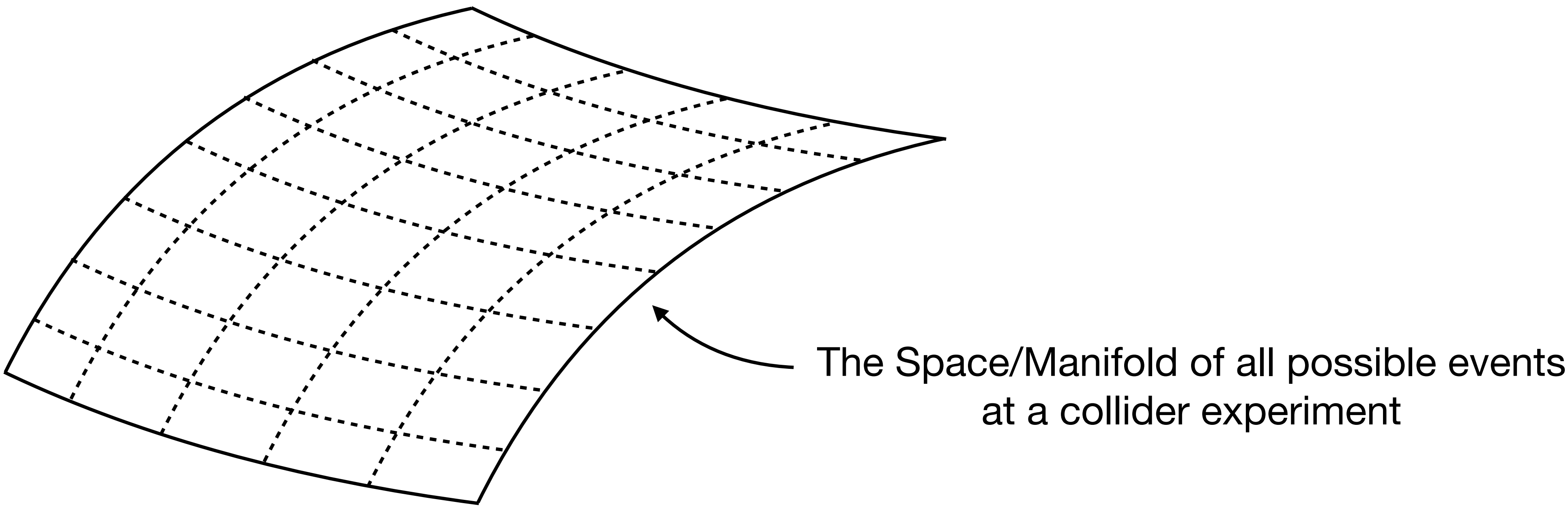
Andrew Larkoski

+ *Jesse Thaler, 2305.03751*

+ *Rikab Gambhir, in progress*

AI/ML Workshop, RHIC/AGS Users Meeting

Metric on the Space of Collider Events



Natural Questions:

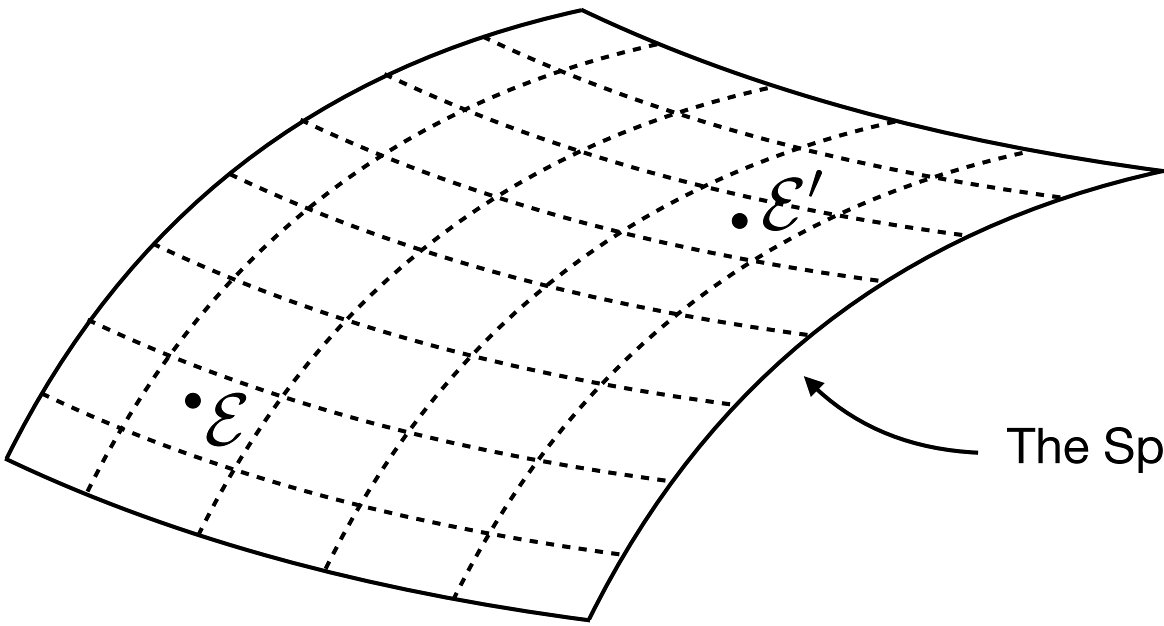
What is the dimensionality of this space?

What is its topology?

What is its volume?

Where are different events located?

Metric on the Space of Collider Events



The Space/Manifold of all possible events
at a collider experiment

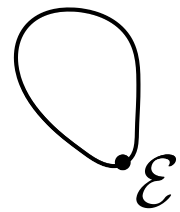
Primitive Question:

What is the distance between events $\mathcal{E}$ and $\mathcal{E}'$?

Reminder: Three Properties of a Metic

"Identity of Indiscernibles"

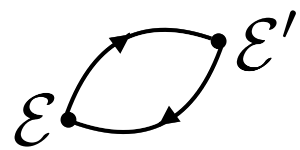
$$d(\mathcal{E},\mathcal{E}) = 0$$



Shortest distance from an event to itself is 0

"Symmetry" (and non-negativity)

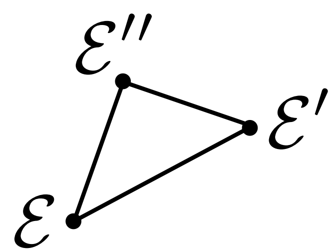
$$d(\mathcal{E},\mathcal{E}') = d(\mathcal{E}',\mathcal{E}) \geq 0$$



Distance from one event to another equals opposite

"Triangle Inequality"

$$d(\mathcal{E},\mathcal{E}') \leq d(\mathcal{E},\mathcal{E}'') + d(\mathcal{E}',\mathcal{E}'')$$



The most direct route is the shortest

First Metric: Energy Mover's Distance (EMD)

Events as Point-Clouds on Detector/Celestial Sphere

$$\mathrm{EMD}(\mathcal{E},\mathcal{E}') = \min_{\{f_{ij}\}} \sum_{ij} f_{ij}\frac{\theta_{ij}}{R} + \left|\sum_i E_i - \sum_j E'_j\right|$$

$$f_{ij} \geq 0, \quad \sum_j f_{ij} \leq E_i, \quad \sum_i f_{ij} \leq E'_j, \quad \sum_{ij} f_{ij} = E_{\min},$$

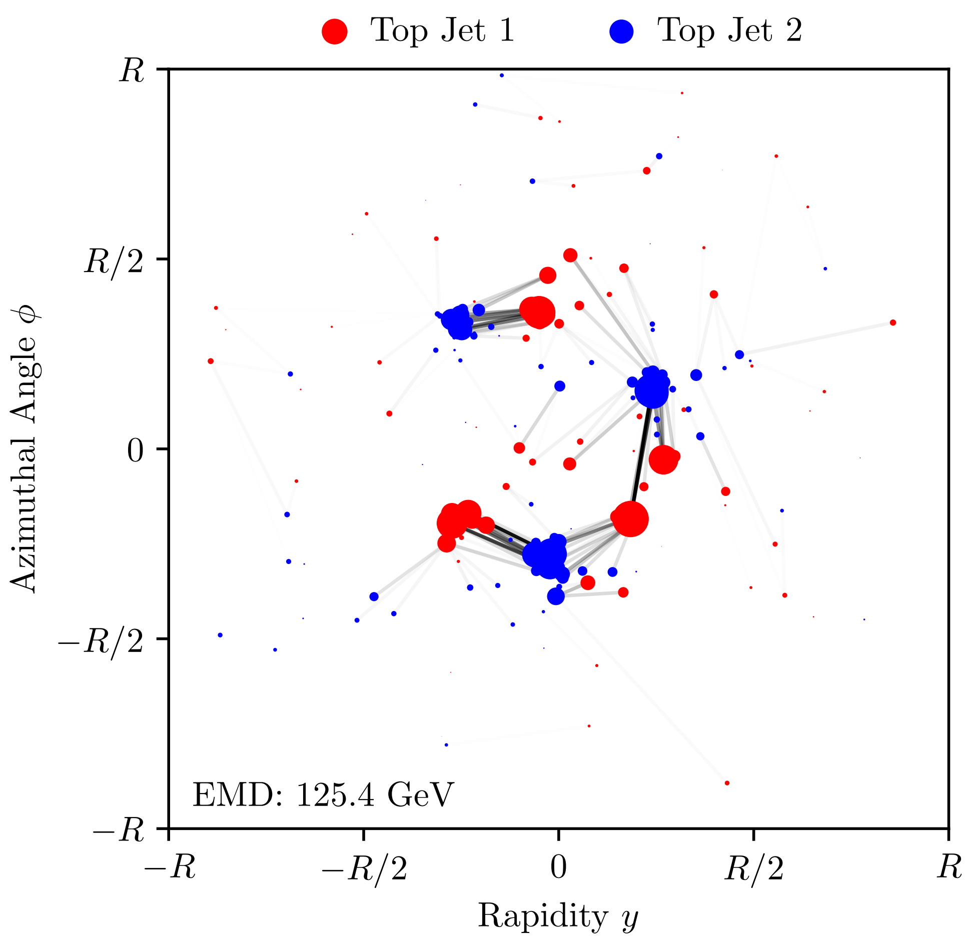
Metric is minimal transportation cost to rearrange one event into another

Monge's "Earth Mover's Distance" applied to colliders

Monge, 1781

Cédric Villani won Fields Medal in 2010 for Optimal Transport

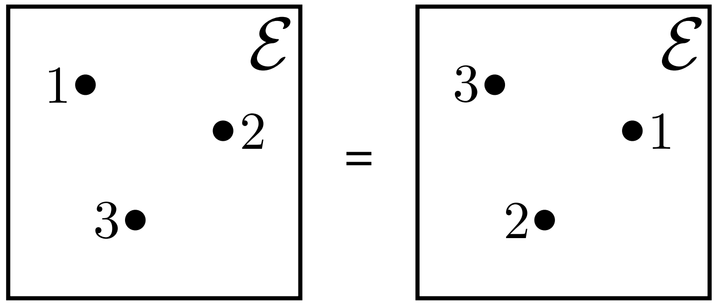
C. Villani, "Optimal Transport: Old and New" (2009)



1902.02346

Three Issues with EMD and other metrics

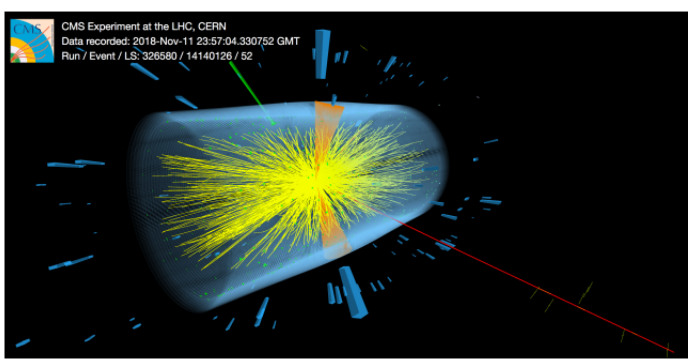
Must Minimize Over All Labelings to Ensure Particle Permutation Invariance



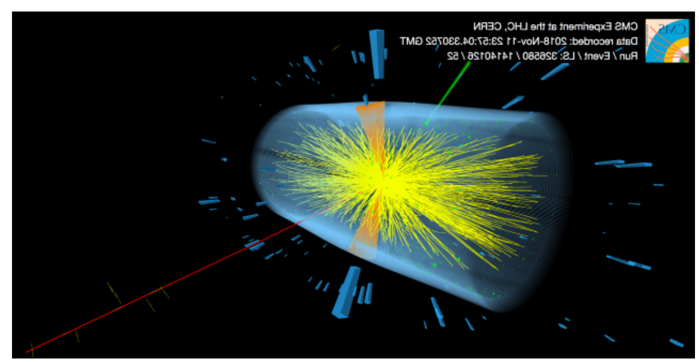
No Closed-Form Expression for Two-Dimensional Optimal Transport

POT: Python Optimal Transport

Non-Zero Distance for Events Related by Isometries

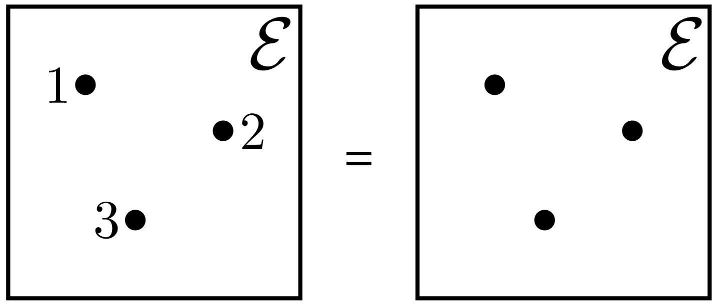


=



New Metric Proposal

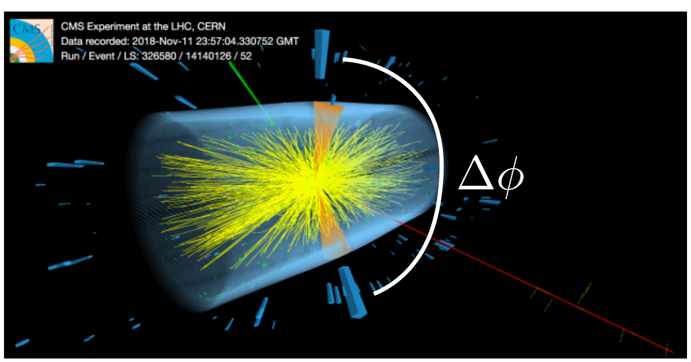
Particle Permutation Invariance Ensured by Not Labeling Particles



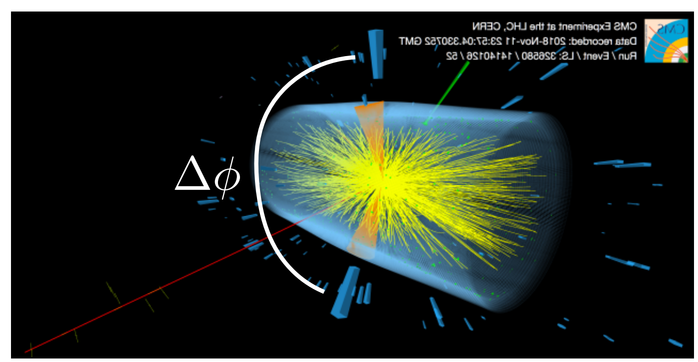
Represent Event as a One-Dimensional Distribution to Use Exact Transport Results

$$d_p(\mathcal{E}, \mathcal{E}') = \int dE^2 \left| S_{\mathcal{E}}^{-1}(E^2) - S_{\mathcal{E}'}^{-1}(E^2) \right|^p$$

Only Use Relative Angles to Ensure Invariance to Isometries



=



Introducing The Spectral EMD

Fundamental Object: “Spectral Function”

name: hep-ph/9601308
first use: Basham, Brown, Ellis, Love 1977

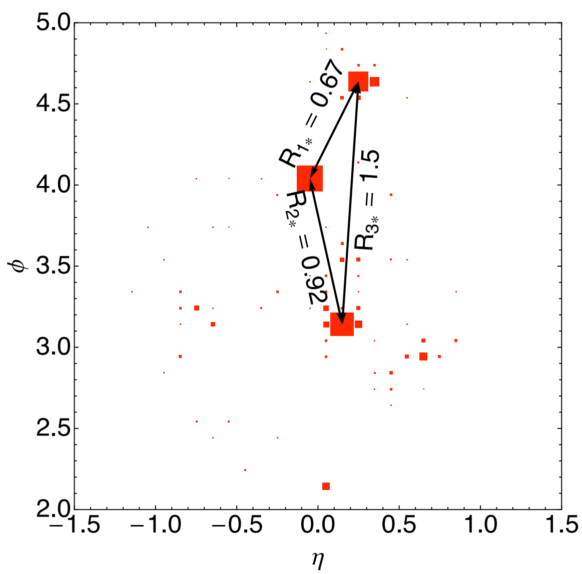
$$s_{\mathcal{E}}(\omega) = \sum_{i,j\in\mathcal{E}} E_i E_j \, \delta(\omega - \omega_{ij})$$

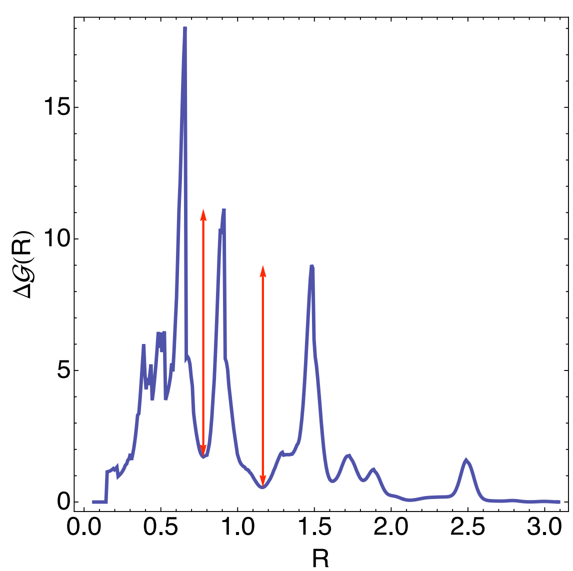
sum runs over all *i*,*j* (including *i* = *j*)

weighted by product of energies

angle between particles *i* and *j*

Previously Used in Jet Substructure for Identification of Dominant Structure





1104.1646

Introducing The Spectral EMD

Fundamental Object: “Spectral Function”

name: hep-ph/9601308
first use: Basham, Brown, Ellis, Love 1977

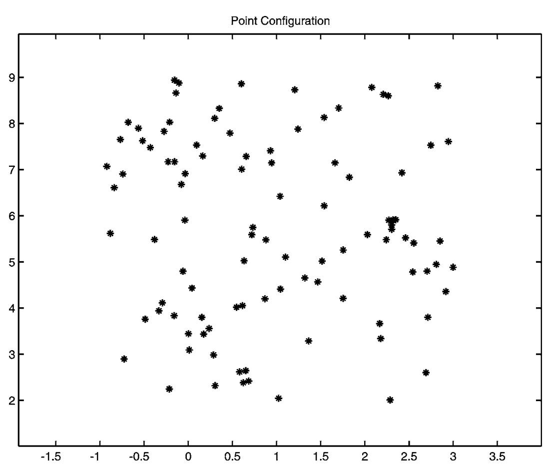
$$s_{\mathcal{E}}(\omega) = \sum_{i,j \in \mathcal{E}} E_i E_j \, \delta(\omega - \omega_{ij})$$

sum runs over all *i*,*j* (including *i* = *j*)

weighted by product of energies

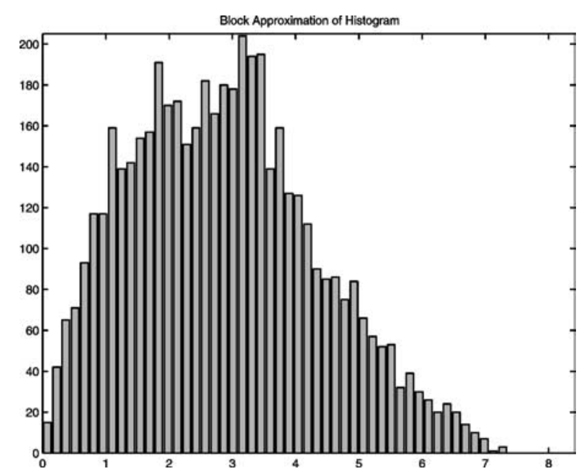
angle between particles *i* and *j*

Distribution of Particles on Celestial Sphere can be Reconstructed
(with probability 1, up to isometries)



Theorem 2.6

Boutin, Kemper 2003



Introducing The Spectral EMD

Fundamental Object: "Spectral Function"

name: hep-ph/9601308
first use: Basham, Brown, Ellis, Love 1977

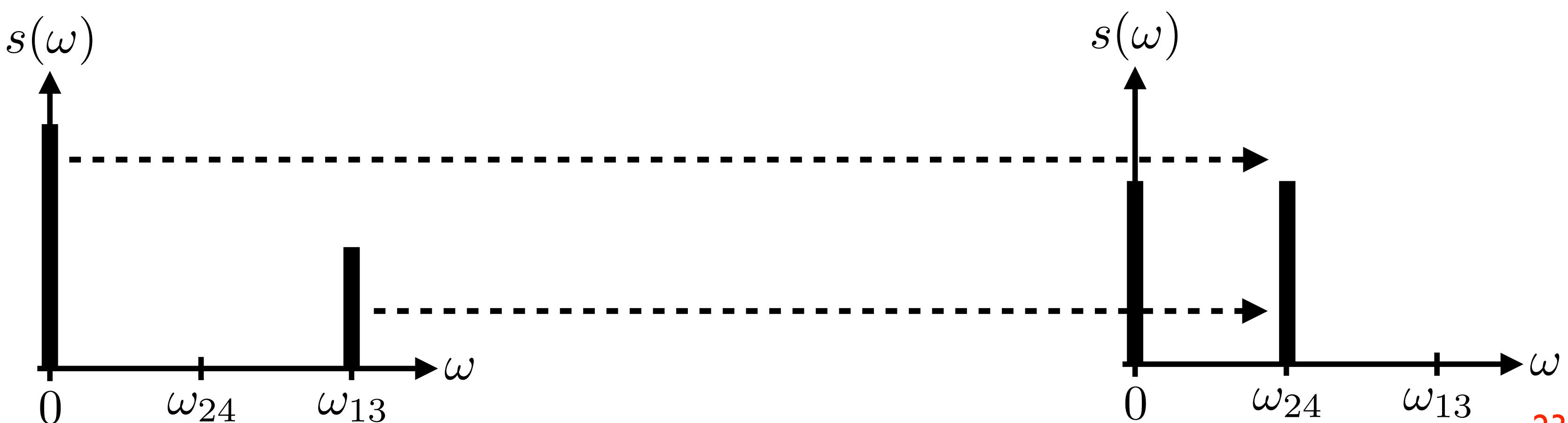
$$s_{\mathcal{E}}(\omega) = \sum_{i,j \in \mathcal{E}} E_i E_j \, \delta(\omega - \omega_{ij})$$

sum runs over all *i*,*j* (including *i* = *j*)

weighted by product of energies

angle between particles *i* and *j*

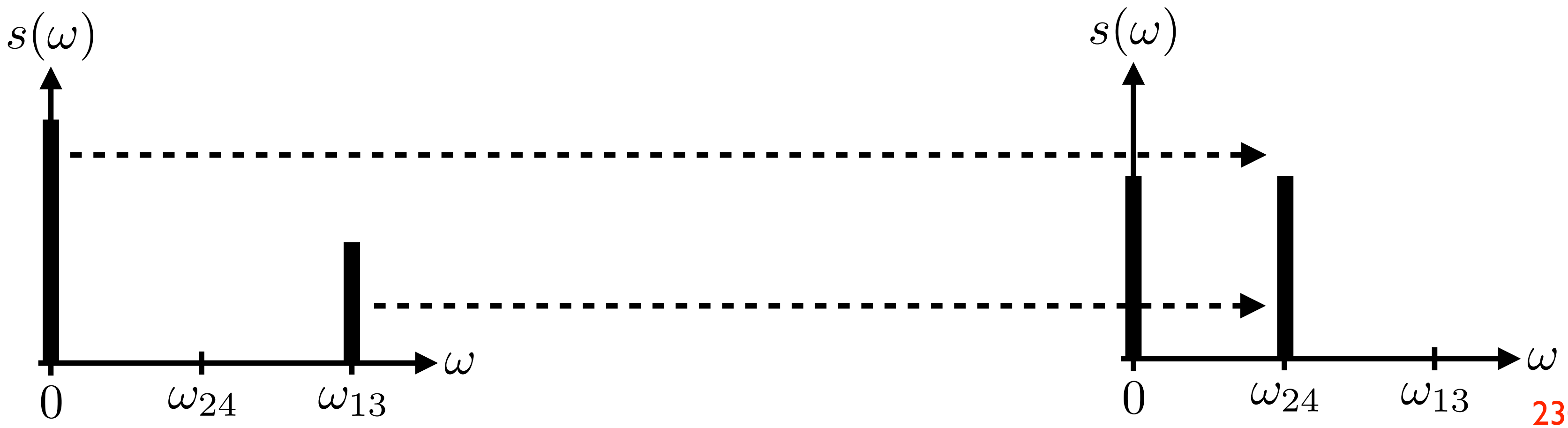
Distance Between Events is then Cost to Rearrange Spectral Functions



2305.03751

One-Dimensional Optimal Transport

Distance Between Events is then Cost to Rearrange Spectral Functions



2305.03751

"*p*-Wasserstein" Metric

$$d_p(\mathcal{E}, \mathcal{E}') = \int dE^2 \left| S_{\mathcal{E}}^{-1}(E^2) - S_{\mathcal{E}'}^{-1}(E^2) \right|^p$$

Kantorovich 1939
Wasserstein 1969
Dobrushin 1970

Only a one-dimensional integral to evaluate!

$S^{-1}(E^2)$ = Inverse Cumulative Spectral Function

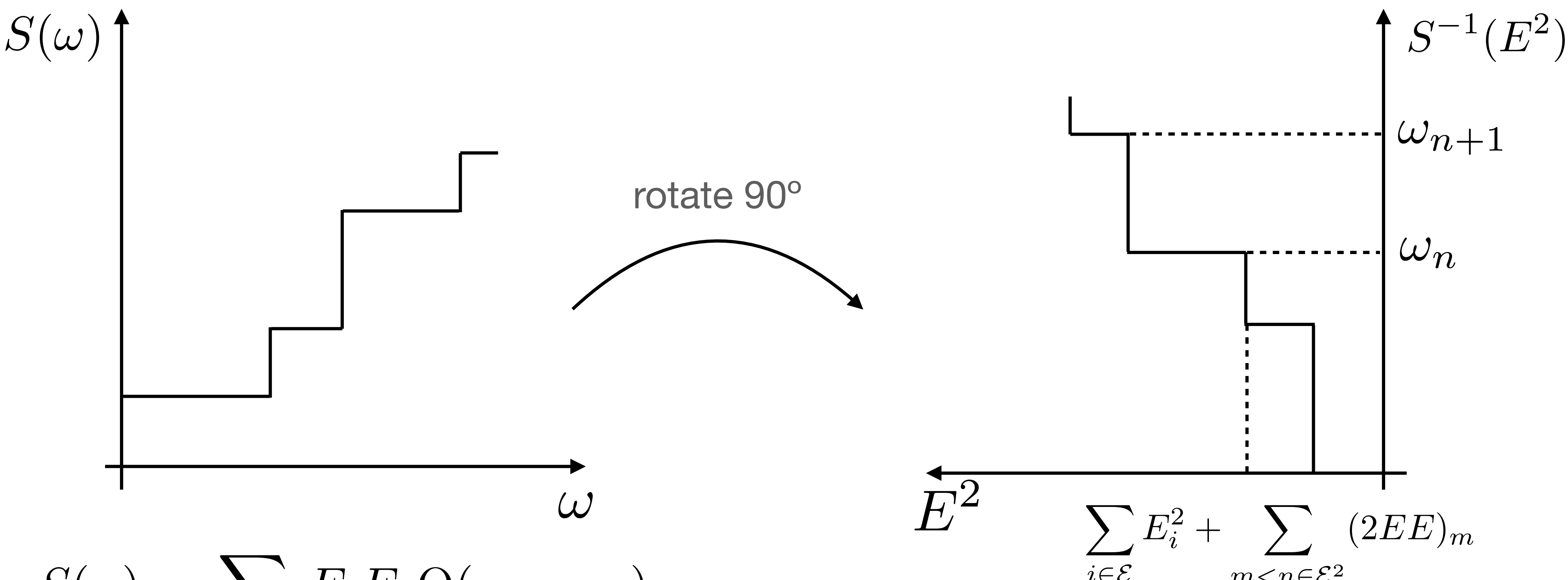
One-Dimensional Optimal Transport

“*p*-Wasserstein” Metric

$$d_p(\mathcal{E},\mathcal{E}') = \int dE^2 \left| S_{\mathcal{E}}^{-1}(E^2) - S_{\mathcal{E}'}^{-1}(E^2) \right|^p$$

Cumulative Spectral Function

Inverse Cumulative Spectral Function



$$S(\omega) = \sum_{i,j\in\mathcal{E}} E_i E_j \Theta(\omega - \omega_{ij})$$

One-Dimensional Optimal Transport

"*p*-Wasserstein" Metric

$$d_p(\mathcal{E},\mathcal{E}') = \int dE^2 \left| S_{\mathcal{E}}^{-1}(E^2) - S_{\mathcal{E}'}^{-1}(E^2) \right|^p$$

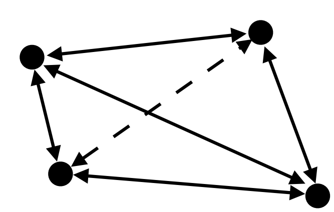
Inverse Cumulative Spectral Function in Gory Detail:

$$S^{-1}(E^2) = \sum_{\substack{n\in\mathcal{E}^2 \\ \omega_n<\omega_{n+1}}} \omega_n \, \Theta\left(\sum_{i\in\mathcal{E}} E_i^2 + \sum_{m\leq n\in\mathcal{E}^2} (2EE)_m - E^2 \right) \Theta\left(E^2 - \sum_{i\in\mathcal{E}} E_i^2 - \sum_{m<n\in\mathcal{E}^2} (2EE)_m \right)$$

Need to sort pairwise angles in increasing value

Explicitly permutation and isometry invariant

Overcomplete event information: order-N^2 pairwise angles,
but only need order-N to triangulate the plane (Euler characteristic)



Closed-Form Expression for 2-Wasserstein Spectral Metric

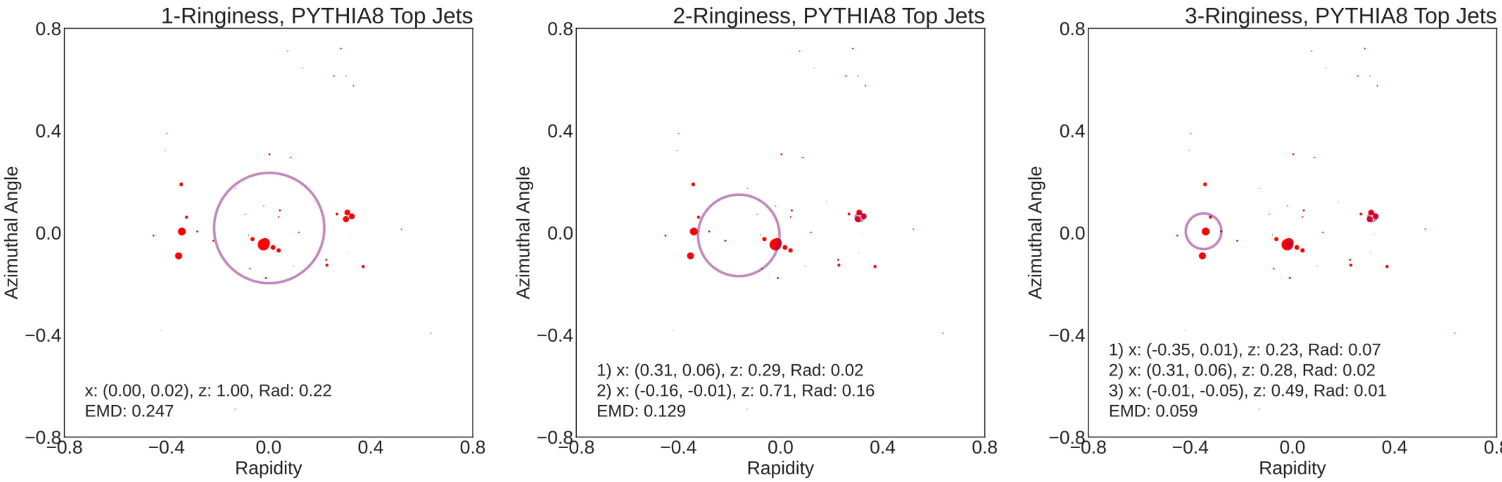
$$
\begin{aligned}
d_2(\mathcal{E}_1,\mathcal{E}_2) = &\sum_{i<j\in\mathcal{E}_1} 2E_iE_j\omega_{ij}^2 + \sum_{i<j\in\mathcal{E}_2} 2E_iE_j\omega_{ij}^2 \\
&- 2\sum_{\substack{n\in\mathcal{E}_1^2,\, l\in\mathcal{E}_2^2\\ \omega_n<\omega_{n+1}\\ \omega_l<\omega_{l+1}}} \omega_n\omega_l \left(\min\left[\sum_{n\le m\in\mathcal{E}_1^2}(2EE)_m, \sum_{l\le k\in\mathcal{E}_2^2}(2EE)_k\right] - \max\left[\sum_{n<m\in\mathcal{E}_1^2}(2EE)_m, \sum_{l<k\in\mathcal{E}_2^2}(2EE)_k\right]\right) \\
&\times \Theta\left(\min\left[\sum_{n\le m\in\mathcal{E}_1^2}(2EE)_m, \sum_{l\le k\in\mathcal{E}_2^2}(2EE)_k\right] - \max\left[\sum_{n<m\in\mathcal{E}_1^2}(2EE)_m, \sum_{l<k\in\mathcal{E}_2^2}(2EE)_k\right]\right)
\end{aligned}
$$

Computational cost is due to ordering ~N^2 angles

Memory costs may be significant for high-multiplicities:
~40 kB to record 4-vectors of 1000 particles
~10 MB to record spectral function of 1000 particles

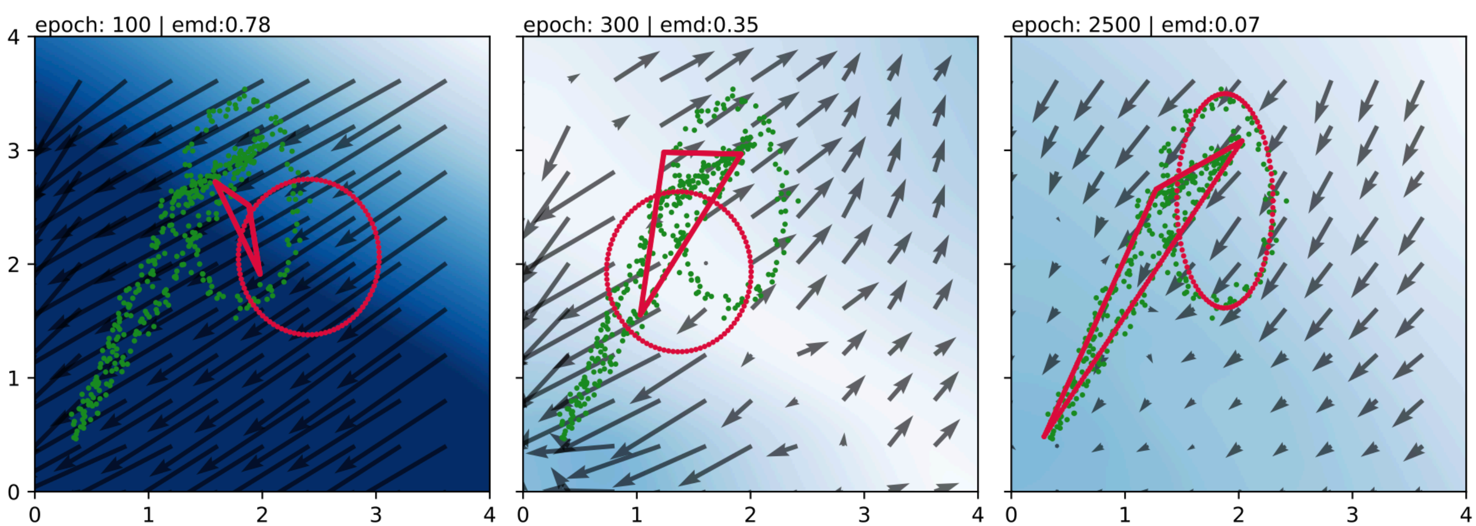
May be ways to compress spectral function information
c.f., spectral function versus Fourier transform

Comparison with Other Differentiable Metric Proposals Based on EMD



SHAPER: 2302.12266

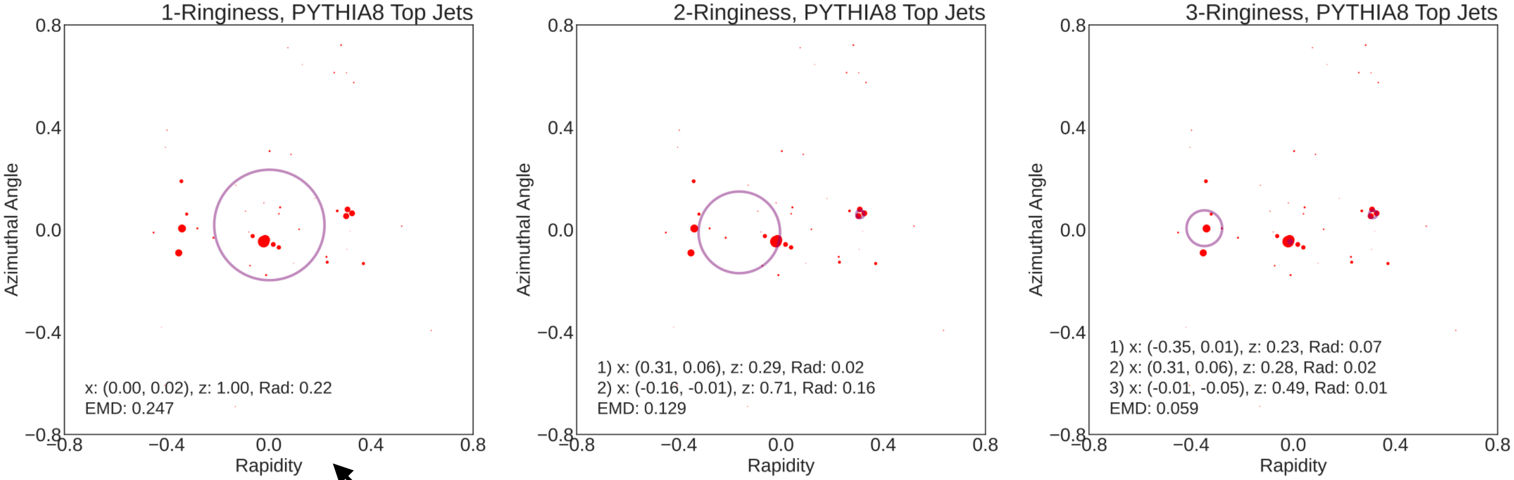
Model complex shapes by large numbers of points and then use POT



NEEMo: 2209.15624
EMD CNN: 2306.04712

Construct a new neural network for geometric fitting of particles in event

Comparison with Other Differentiable Metric Proposals Based on EMD



SHAPER: 2302.12266

"1-Ringiness" with SHAPER/EMD

Closed-Form "1-Ringiness" with Spectral EMD:

$$d_2(\mathcal{E}, \mathcal{E}_{\text{ring}}) = \sum_{i<j\in\mathcal{E}} 2E_iE_j\omega_{ij}^2 + 2E_{\text{tot}}^2R^2$$
$$- \frac{8R}{\pi}E_{\text{tot}}^2 \sum_{\substack{n\in\mathcal{E}^2 \\ \omega_n<\omega_{n+1}}} \omega_n \left[\sin\left(\frac{\pi}{2E_{\text{tot}}^2}\sum_{m\geq n}(2EE)_m\right) - \sin\left(\frac{\pi}{2E_{\text{tot}}^2}\sum_{m>n}(2EE)_m\right)\right]$$

Can easily analytically optimize for ring radius *R*

Summary

By exploiting the information in the spectral function, we can construct a metric on events that is invariant to isometries and expressed in closed form

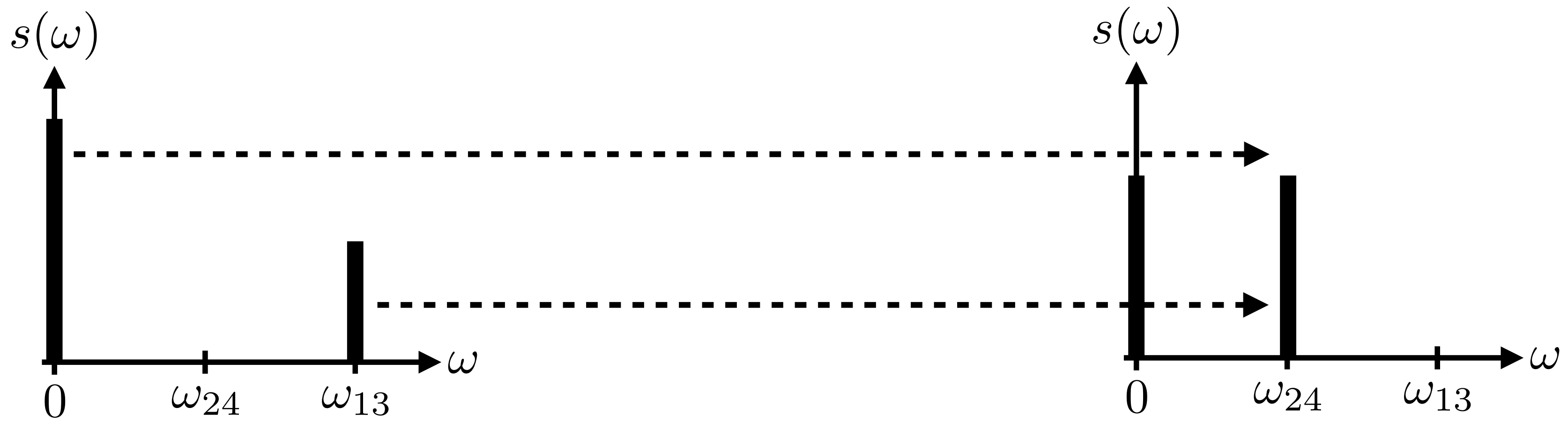
Preliminary studies also indicate that the EMD is significantly more computationally intensive that the spectral metric (at least a factor of N slower)

Simple, differentiable, closed form expressions for the spectral metric enable first-principles theory calculations that are impossible with the EMD

Look forward to more applications and public code in the near future!

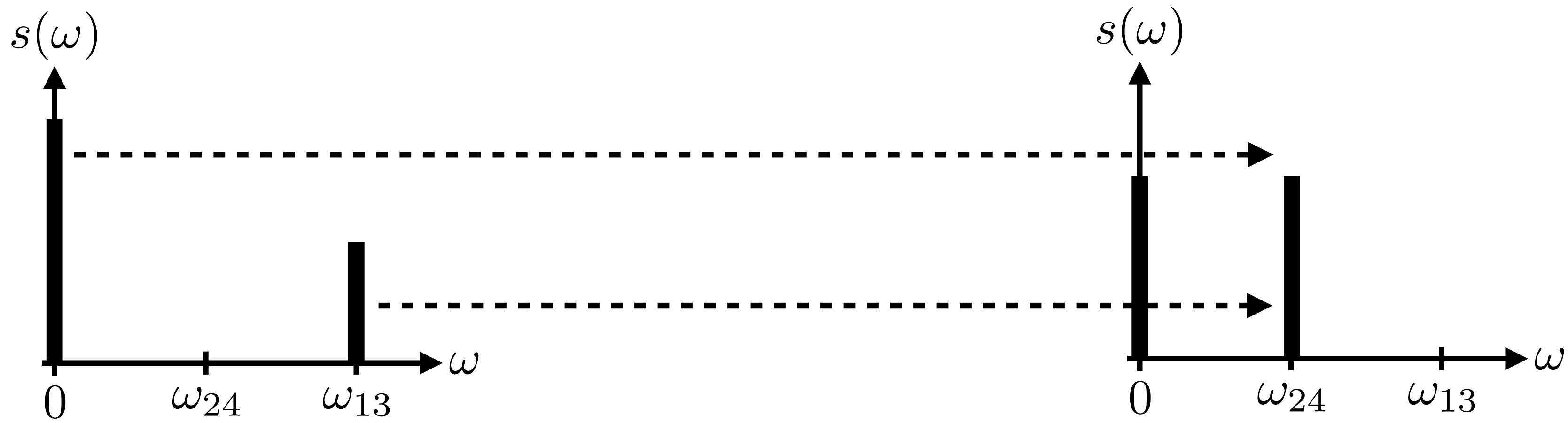
Bonus

Example: 2-Wasserstein Metric for Jets with 2 Particles Each



$$d_2(\mathcal{E}_1, \mathcal{E}_2) = 2E_1E_3\omega_{13}^2 + 2E_2E_4\omega_{24}^2 - \ 4\min[E_1E_3, E_2E_4]\omega_{13}\omega_{24}$$

Example: 2-Wasserstein Metric for Jets with 2 Particles Each



$$d_2(\mathcal{E}_1, \mathcal{E}_2) = \boxed{2E_1E_3\omega_{13}^2 + 2E_2E_4\omega_{24}^2} - 4\min[E_1E_3, E_2E_4]\omega_{13}\omega_{24}$$

sum of jets' two-point energy correlation functions

Differentiable in all pairwise angles: extremization is easy

Non-analytic in energies: extremization is a discrete partitioning problem