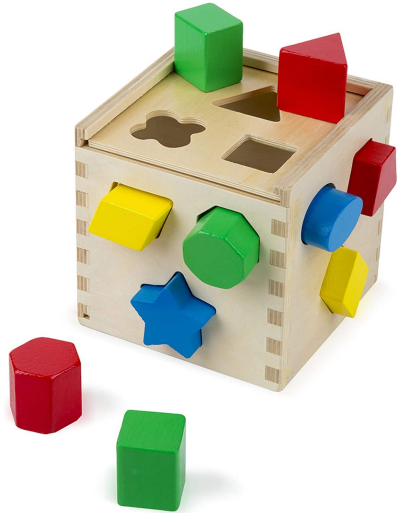











Can you hear the shape of a jet?



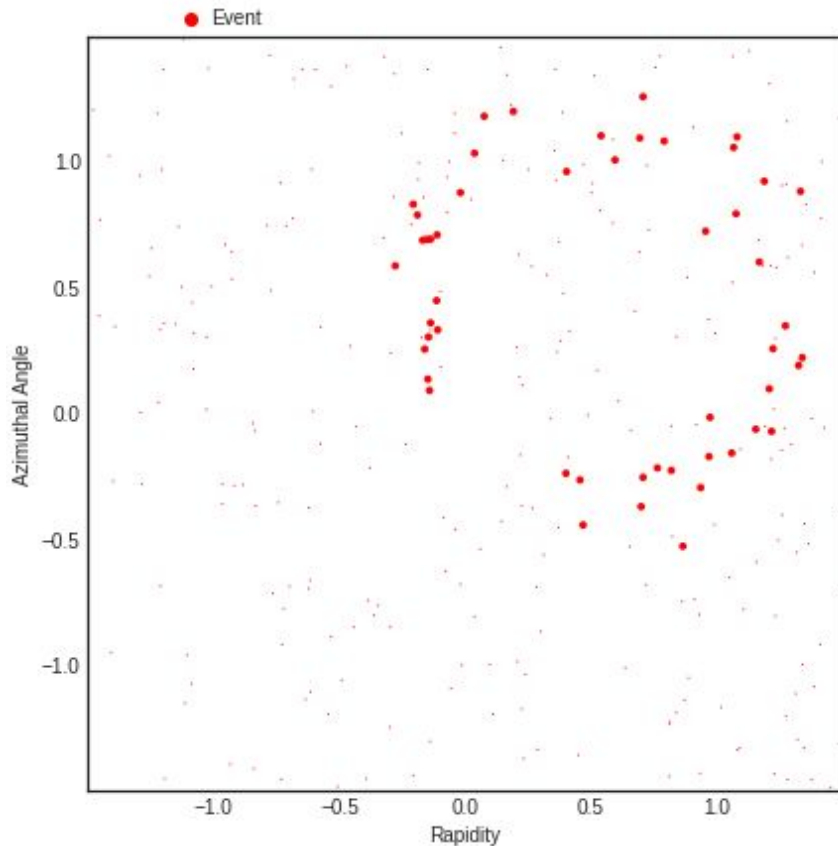
Rikab Gambhir

With Akshunna S. Dogra (  ),
Demba Ba ( ),
Abiy Tasissa ( ),
& Jesse Thaler ( )

Email me questions at rikab@mit.edu!

Based on [Ba, Dogra, **RG**, Tasissa, Thaler, 2211.XXXX] (Coming Soon!)

Fundamental Question: What shape is this?

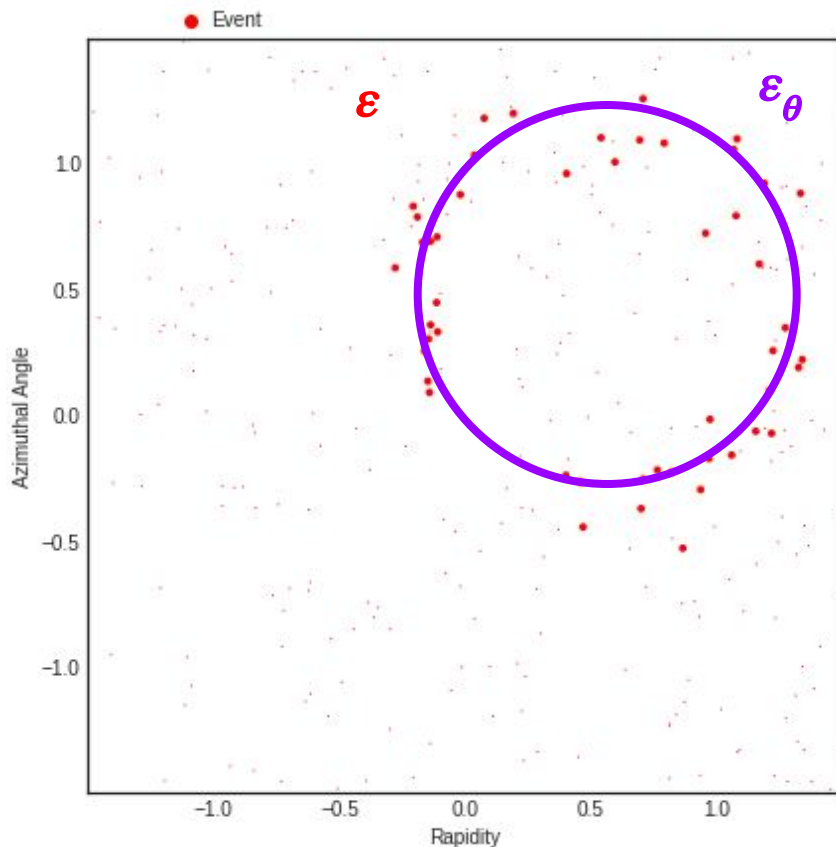


Pictured: (Fake) event that you might have measured at the LHC

Red dots are detector hits on a patch of the LHC cylinder, weighted by energy

Goal: Construct an observable \mathcal{O} that generically answers this question!

Fundamental Question: What shape is this?

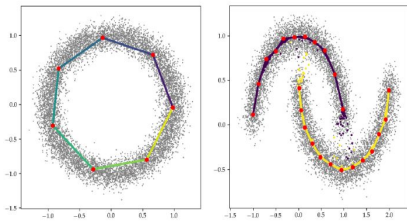


Using the **SHAPER** framework ...

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$
$$\theta = \text{argmin}_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

Circle with radius 0.767, center (0.50, 0.36) and a “circle-ness” value of 0.32.

Yes, you **CAN** hear the shape of a jet!

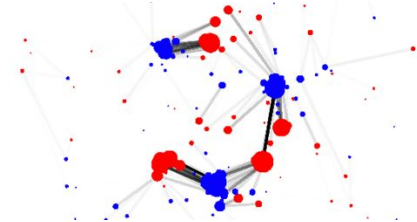


Piecewise-Linear Manifold
Approximation with K-Deep Simplices
(KDS, [2012.02134](#))

Ai

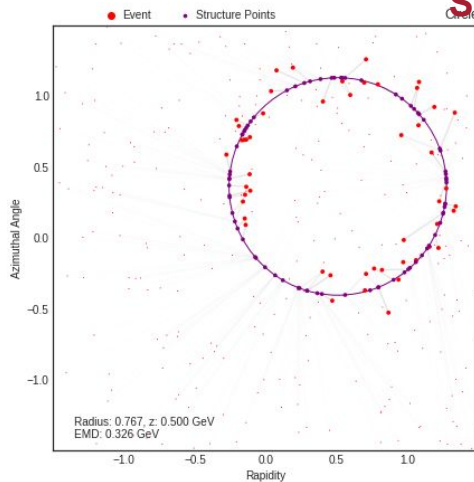
AI

fi



Well-Defined Metric on Particle Collisions
using Energy Mover's Distance (EMD,
[2004.04159](#))

SHAPER: Learning the Shape of Collider Events



$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

$$\theta = \operatorname{argmin}_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

Framework for defining
and calculating useful
observables for collider
physics!

~~THE UNREASONABLE EFFECTIVENESS OF MATHEMATICS IN THE NATURAL SCIENCES~~

The Wasserstein Metric

Collider Physics

Eugene Wigner

Mathematics, rightly viewed, possesses not only truth, but supreme beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show. The true spirit of delight, the exaltation, the sense of being more than Man, which is the touchstone of the highest excellence, is to be found in mathematics as surely as in poetry.

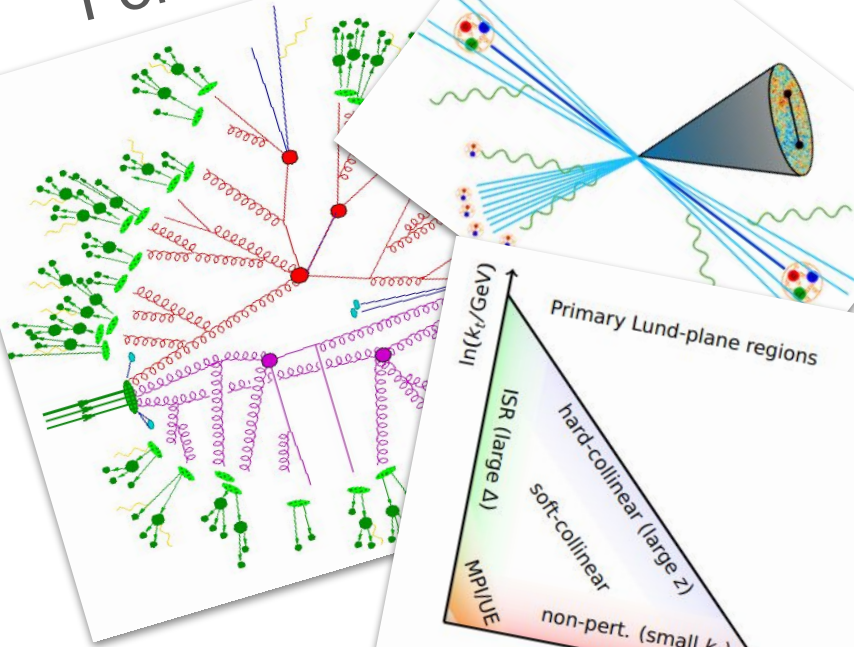
- BERTRAND RUSSELL, Study of Mathematics

Robust Observables

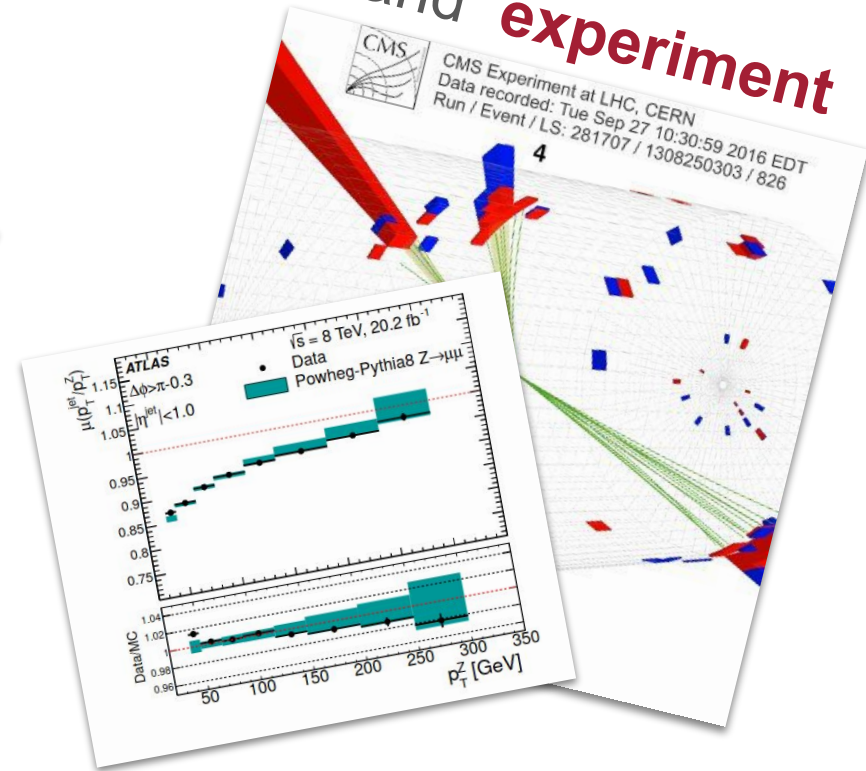
Images from [Bothmann et. al., 1905.09127;
Lee, Męcaj, Moutl, 2205.03414;
Dreyer, Salam, Soyez, 1807.04758
CMS, 1810.10069;
ATLAS, 1703.10485]

We want *Robust Observables*!

For **theory** ...



and **experiment**



Robust Observables

We want *Robust Observables!*

For **theory** ...

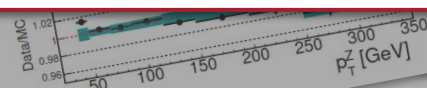
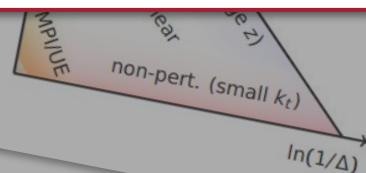
Worry about:

- Perturbativity
- Hadronization
- Choice of Shower
- Interpretability
- ... and more

and **experiment**

Worry about:

- Finite Resolution
- Particle Reconstruction
- Differences between detectors
- ... and more



Robust Observables

Images from [Bothmann et. al., 1905.09127;
Lee, Meza, Mout, 2205.03414;
Scheel, 1807.04758
Scheel, 1810.10069;
Scheel, 1703.10485]

See Ian Mout's Talk from Tuesday!

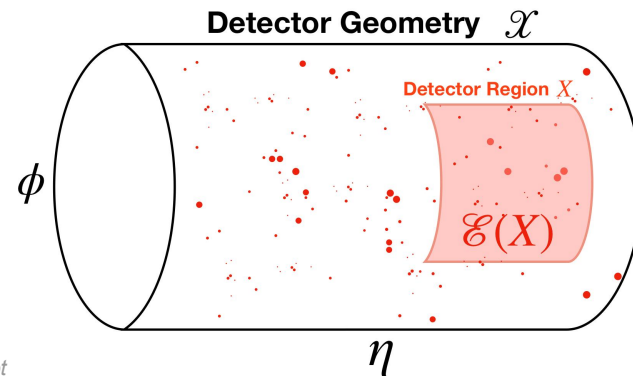
We want *Robust Observables*!

Build observables out of *Energy Flows*!

$$\mathcal{E}(\vec{y}) = \sum_i z_i \delta(\vec{y} - \vec{y}_i)$$

Detector Coordinate (η, ϕ)

Energy Fraction E_j / E_{Tot}



- Captures all the IRC-Safe / kinematically accessible calorimeter information in an event.
- Generalizes events - *any* probability distribution of energy could be an event! Events can be **real** or **idealized**

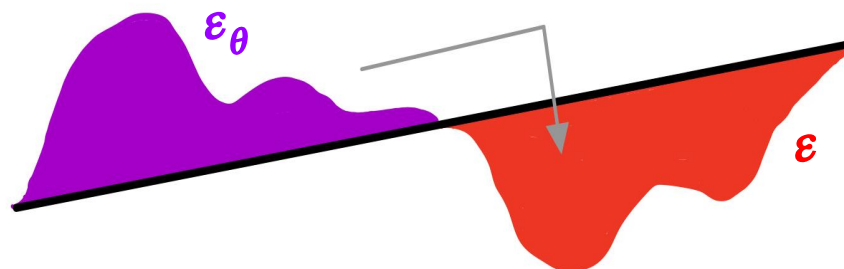
Observables and Wasserstein

It can be shown that *any* observable on events, that* ...

1. ... is non-negative and finite
2. ... is IRC-safe
3. ... is translationally invariant
4. ... is invariant to particle labeling
5. ... respects the detector metric *faithfully***

... can be written as an optimization of the **Wasserstein Metric (Earth/Energy Mover's Distance)** between the real event and a manifold of idealized energy flows

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$
$$\theta = \operatorname{argmin}_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

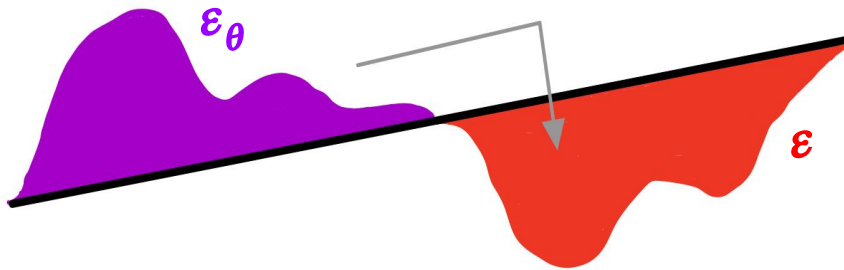


EMD = Work done to move “dirt” optimally

*Ask me for more details on this offline!

** Preserves distances between *extended* objects, not just points

Observables and Wasserstein



EMD = Work done to move “dirt” optimally

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_\theta \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$

$$\theta = \operatorname{argmin}_{\mathcal{E}'_\theta \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$

- 4. ... is invariant to particle labeling
- 5. ... respects the detector metric *faithfully*

$$\text{EMD}^{(\beta, R)}(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geq 0} \left[\frac{1}{\beta R^\beta} \sum_{i=1}^M \sum_{j=1}^N \pi_{ij} d_{ij}^\beta \right] + |\Delta E_{\text{tot}}|$$

$$\sum_{i=1}^M \pi_{ij} \leq E'_j, \quad \sum_{j=1}^N \pi_{ij} \leq E_i \quad \text{and} \quad \sum_{j=1}^N \pi_{ij} = \min(E_{\text{tot}}, E'_{\text{tot}})$$

Ask me for more details on this later!



Hearing Jets

Effect of Jet Aircraft Noise on Hearing
A. COLES & J.J. KNIGHT
184, 1803-1804 (1959) | Cite this article
Metrics
4 Citations | Metrics
hearing of per

Observable \Leftrightarrow Manifolds

Many existing observables have this form!

Observables \Leftrightarrow Manifold of Shapes

- N -subjettiness \Leftrightarrow Manifold of N -point events
- N -jettiness \Leftrightarrow Manifold of N -point events with floating total energy
- Thrust \Leftrightarrow Manifold of back-to-back point events
- Event Isotropy \Leftrightarrow Uniform distribution
- ... and more!

All of the form “How much like **[shape]** does my **event** look like?”

We generalize this to build more observables!

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

$$\theta = \underset{\mathcal{E}'_{\theta} \in \mathcal{M}}{\text{argmin}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

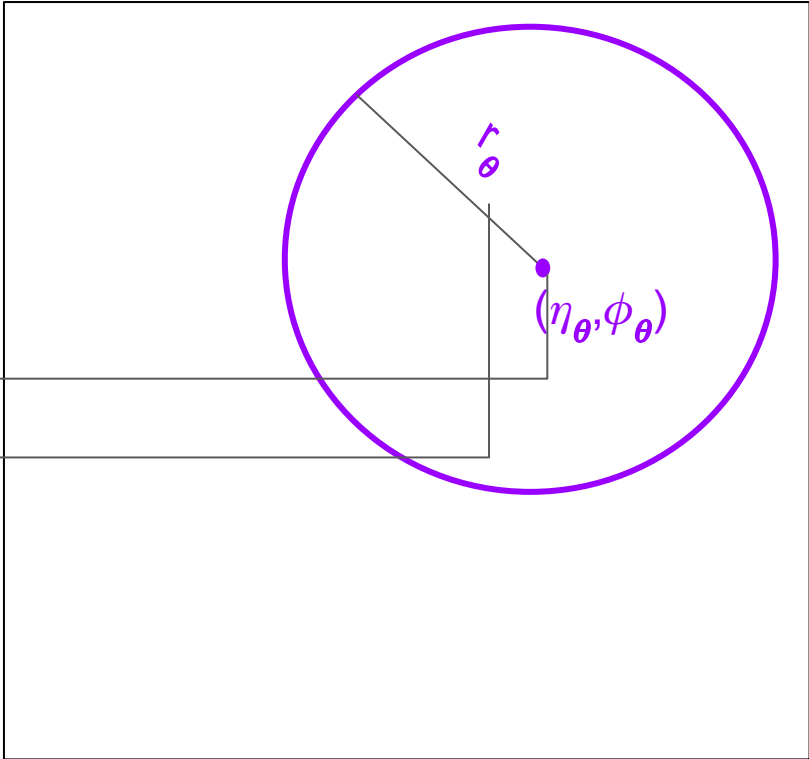
Shape Observables

Let's go further! Consider *any* manifold of parameterized energy flows

Define a *shape* as any parameterized energy flow

For example, for a circle with parameterized radius and center*

$$\mathcal{E}(\vec{y}) = \begin{cases} \frac{1}{2\pi r_\theta} & |\vec{y} - \vec{y}_\theta| = r_\theta \\ 0 & |\vec{y} - \vec{y}_\theta| \neq r_\theta \end{cases}$$



Can then sample points (particles) from this idealized distribution!

Parameterized circle written as an energy flow

*Uniform prior by choice for simplicity. In principle, we can pick any parameterized normalized distribution.

Shape Observables

Given any manifold of parameterized energy flows (probability distributions) \mathcal{M} , representing shapes, we define the **shape observable** \mathcal{O} and **shape parameters** θ on an event \mathcal{E} as:

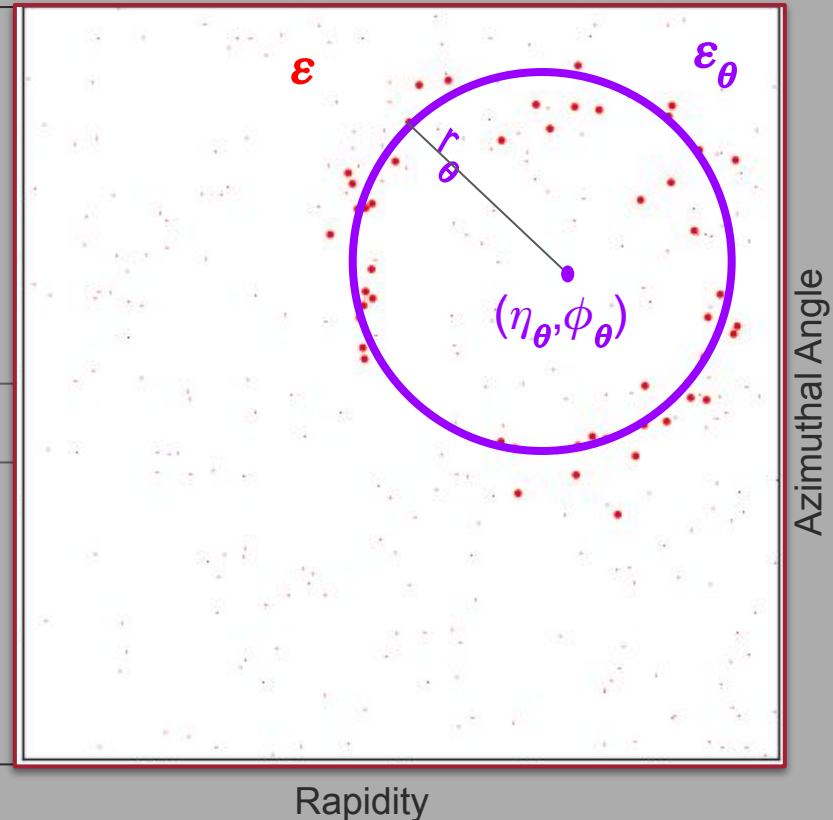
$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

$$\theta = \operatorname{argmin}_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

Shape observables are the generalization of event and jet shapes like N-(sub)jettiness, thrust, and isotropy!

[Ba, Dogra, **RG**, Tasissa, Thaler, 2211.XXXX]

parameterized energy flows



Parameterized circle written as an energy flow

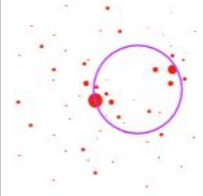
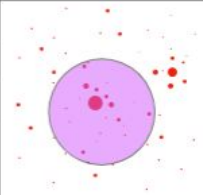
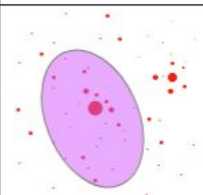
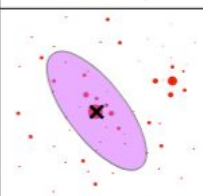
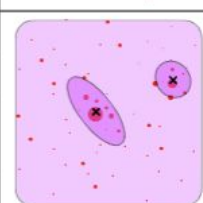
*Uniform prior by choice for simplicity. In principle, we can pick any parameterized normalized distribution.

Hearing Shapes

$\mathcal{O}_{\mathcal{M}}(\mathcal{E})$ answers: “How much like a shape in \mathcal{M} does my event \mathcal{E} look like?”

$\theta_{\mathcal{M}}(\mathcal{E})$ answers: “Which shape in \mathcal{M} does my event \mathcal{E} look like?”

Can define complex manifolds to probe increasingly subtle geometric structure, and even old combine shape observables to create new composite ones!

Shape	Specification	Illustration
Ringiness \mathcal{O}_R	Manifold of Rings $\mathcal{E}_{x_0, R_0}(x) = \frac{1}{2\pi R_0}$ for $ x - x_0 = R_0$ $x_0 = \text{Center}, R_0 = \text{Radius}$	
Diskiness \mathcal{O}_D	Manifold of Disks $\mathcal{E}_{x_0, R_0}(x) = \frac{1}{\pi R_0^2}$ for $ x - x_0 \leq R_0$ $x_0 = \text{Center}, R_0 = \text{Radius}$	
Ellipsiness \mathcal{O}_E	Manifold of Ellipses $\mathcal{E}_{x_0, a, b, \varphi}(x) = \frac{1}{\pi ab}$ for $x \in \text{Ellipse}_{x_0, a, b, \varphi}$ $x_0 = \text{Center}, a, b = \text{Semi-axes}, \varphi = \text{Tilt}$	
(Ellipse Plus Point)iness	Composite Shape $\mathcal{O}_E \oplus \tau_1$ Fixed to same center x_0	
N-(Ellipse Plus Point)iness Plus Pileup	Composite Shape $N \times (\mathcal{O}_E \oplus \tau_1) \oplus \mathcal{I}$	

Some examples of shapes you can define!

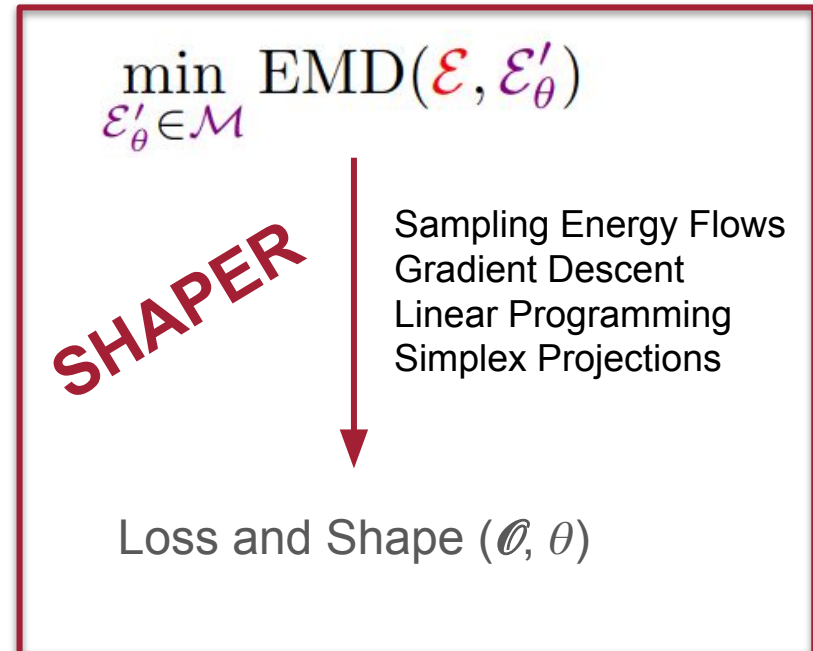


The SHAPER Framework

SHAPER

Shape-**H**unting **A**lgorithm using
Parameterized **E**nergy
Reconstruction

- Framework for defining and building IRC-safe observables using parameterized objects
- Easy to programmatically define new observables by specifying parameterization, or by combining shapes
- Returns EMD distance and optimal shape parameters



Estimating Wasserstein

We need a *differentiable, fast* approximation to the EMD for our minimizations

Sinkhorn Divergence: A strictly convex approximation to EMD! Kantorovich potential formalism::

$$\mathcal{O}(\mathcal{E}) = \min_{\mathcal{E}_\theta \in \mathcal{M}} [S_\epsilon(\mathcal{E}, \mathcal{E}')] \text{ and } \theta(\mathcal{E}) = \operatorname{argmin}_{\mathcal{E}_\theta \in \mathcal{M}} [S_\epsilon(\mathcal{E}, \mathcal{E}')] , \quad \text{where}$$

$$S_\epsilon(\mathcal{E}, \mathcal{E}') = \operatorname{OT}_\epsilon(\mathcal{E}, \mathcal{E}_\theta) - \frac{1}{2} \operatorname{OT}_\epsilon(\mathcal{E}, \mathcal{E}) - \frac{1}{2} \operatorname{OT}_\epsilon(\mathcal{E}_\theta, \mathcal{E}_\theta), \quad \text{and}$$

$$\operatorname{OT}_\epsilon(\mathcal{E}, \mathcal{E}') = \max_{f, g: \mathcal{X} \rightarrow \mathbb{R}} \left[\sum_{i=1}^M E_i f(x_i) + \sum_{j=1}^N E'_j g(y_j) - \epsilon^\beta \log \left(\sum_{ij} E_i E'_j \left(e^{\frac{1}{\epsilon^\beta} (f(x_i) + g(y_j)) - \frac{d(x_i, y_j)^\beta}{R^\beta}} \right) \right) \right]$$

Can take gradients with respect to the entire event – very useful!

See Ouail Kitouni's talk later for more!

Algorithm 3.4: Symmetric Sinkhorn algorithm, with debiasing

Parameters: Cost function $C : (x_i, y_j) \in \mathcal{X} \times \mathcal{X} \mapsto C(x_i, y_j) \in \mathbb{R}$,
Temperature $\epsilon > 0$.

Input: Positive measures $\alpha = \sum_{i=1}^N \alpha_i \delta_{x_i}$ and $\beta = \sum_{j=1}^M \beta_j \delta_{y_j}$ with the same mass.

```

1:  $f_i^{\beta \rightarrow \alpha}, g_j^{\alpha \rightarrow \beta}, f_i^{\alpha \leftrightarrow \alpha}, g_j^{\beta \leftrightarrow \beta} \leftarrow \mathbf{0}_{\mathbb{R}^N}, \mathbf{0}_{\mathbb{R}^M}, \mathbf{0}_{\mathbb{R}^N}, \mathbf{0}_{\mathbb{R}^M}$  ▷ Dual vectors.
2: repeat ▷ The four lines below are executed simultaneously.
3:    $f_i^{\beta \rightarrow \alpha} \leftarrow \frac{1}{2} f_i^{\beta \rightarrow \alpha} + \frac{1}{2} \min_{y \sim \beta, \epsilon} [C(x_i, y) - g^{\alpha \rightarrow \beta}(y)]$ , ▷  $\alpha \leftarrow \beta$ 
    $g_j^{\alpha \rightarrow \beta} \leftarrow \frac{1}{2} g_j^{\alpha \rightarrow \beta} + \frac{1}{2} \min_{x \sim \alpha, \epsilon} [C(x, y_j) - f^{\beta \rightarrow \alpha}(x)]$ , ▷  $\beta \leftarrow \alpha$ 
    $f_i^{\alpha \leftrightarrow \alpha} \leftarrow \frac{1}{2} f_i^{\alpha \leftrightarrow \alpha} + \frac{1}{2} \min_{x \sim \alpha, \epsilon} [C(x_i, x) - f^{\alpha \leftrightarrow \alpha}(x)]$ , ▷  $\alpha \leftarrow \alpha$ 
    $g_j^{\beta \leftrightarrow \beta} \leftarrow \frac{1}{2} g_j^{\beta \leftrightarrow \beta} + \frac{1}{2} \min_{y \sim \beta, \epsilon} [C(y, y_j) - g^{\beta \leftrightarrow \beta}(y)]$ . ▷  $\beta \leftarrow \beta$ 
4: until convergence up to a set tolerance. ▷ Monitor the updates on the potentials.
5: return  $f_i^{\beta \rightarrow \alpha} - f_i^{\alpha \leftrightarrow \alpha}, g_j^{\alpha \rightarrow \beta} - g_j^{\beta \leftrightarrow \beta}$  ▷ Debaised dual potentials  $F(x_i)$  and  $G(y_j)$ .

```

Implemented using the [KerOps+GeomLoss](#) Python Package!

The SHAPER Algorithm

$$\min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

To estimate a shape observable ...

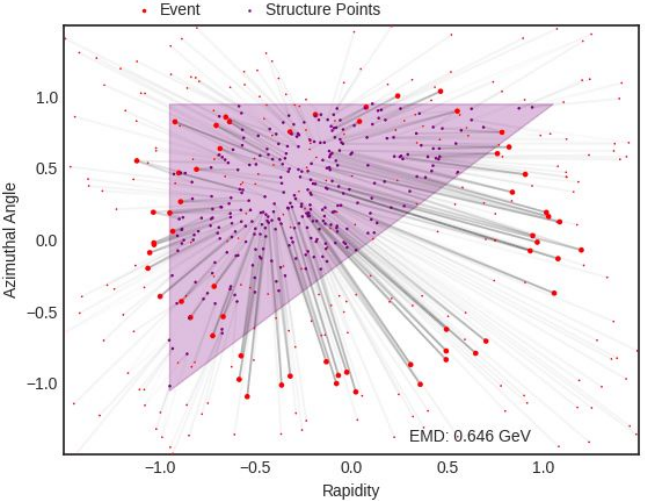
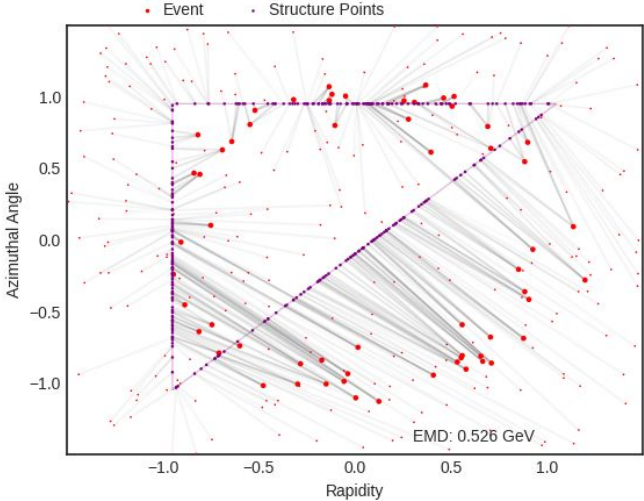
1. Define a parameterized distribution that can be sampled \Leftrightarrow **Manifold** \mathcal{M}
2. Initialize the parameters θ in an IRC-safe way (Usually k_T)
3. Use the Sinkhorn Algorithm to estimate the EMD between your **event** \mathcal{E} and **shape** \mathcal{E}_{θ}
4. Calculate the gradients of the EMD using the Kantorovich potentials
5. Use the gradients to update θ (using ADAM or another optimizer)
6. Repeat 3-5 until convergence
7. Return the loss $\mathcal{O}_{\mathcal{M}}$ and the optimal parameters $\theta_{\mathcal{M}}$

SHAPER



Loss and Shape (\mathcal{O}, θ)

Fun Animations

How **triangle-y** is an event? (Boundary or filled in)?



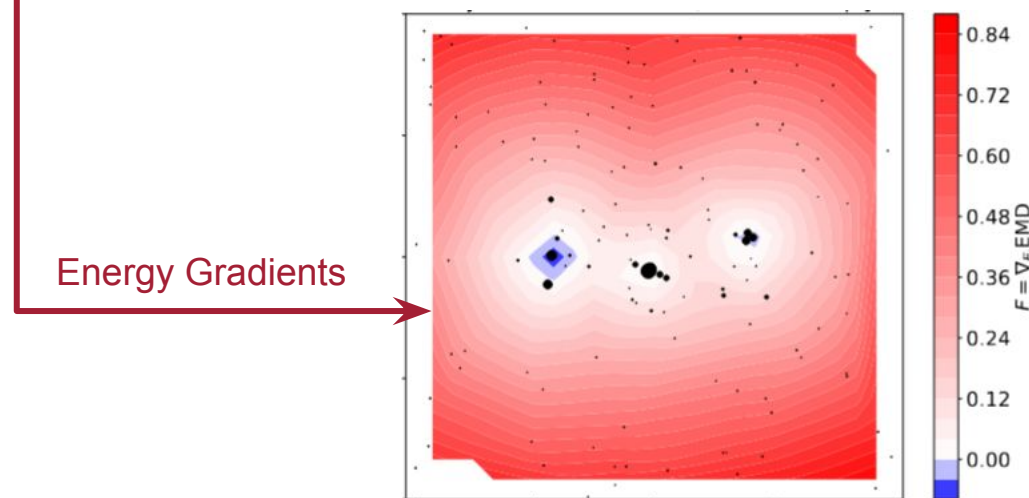
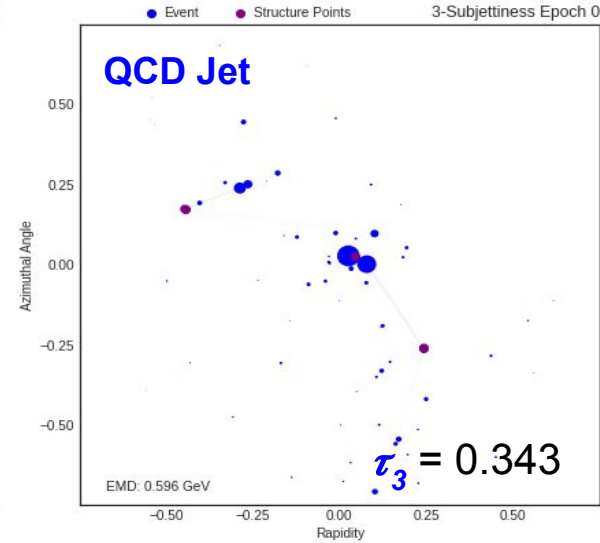
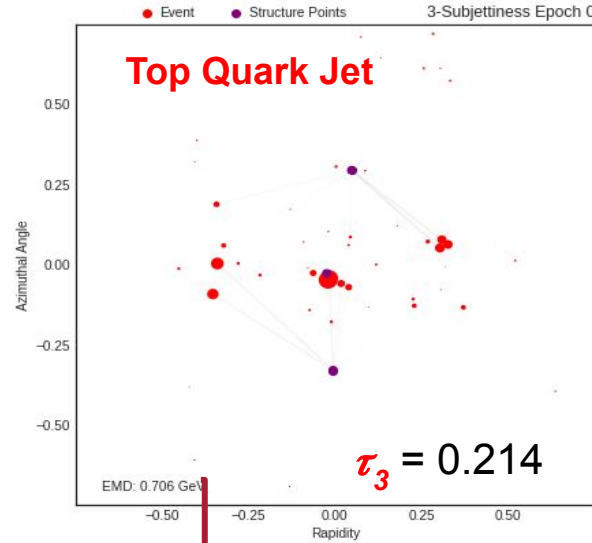
Red: Event
Purple: Shape ε_θ
Grey: Matrix f_{ij} connecting *particles* and *triangle*

Left: $\varepsilon_\theta =$  EMD = 0.245
Right: $\varepsilon_\theta =$  EMD = 0.279

N-Subjettiness

Easy to compute your favorite classic jet observables!

We can even get *gradients* of our observables with respect to the events!



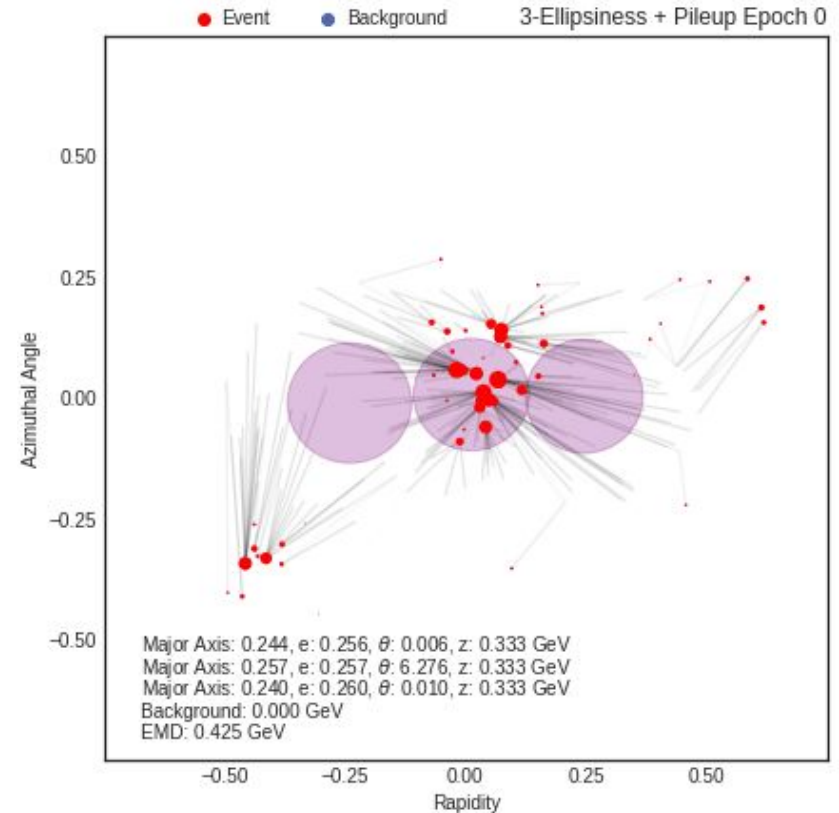
New IRC-Safe Observables

The **SHAPER** framework makes it easy to invent new jet observables!

e.g. **N-Ellipsiness+Pileup** as a jet algorithm.

- Learn jet centers
- Dynamic jet radii (no R hyperparameter)
- Dynamic eccentricities and angles
- Dynamic jet energies
- Learned parameters for discrimination

Can design custom specialized jet algorithms to learn jet substructure!



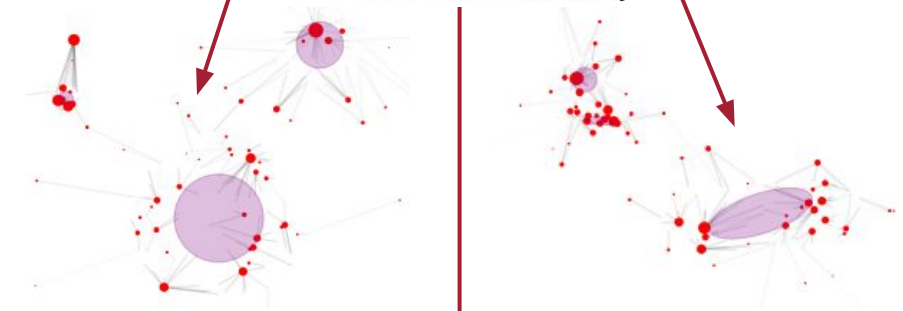
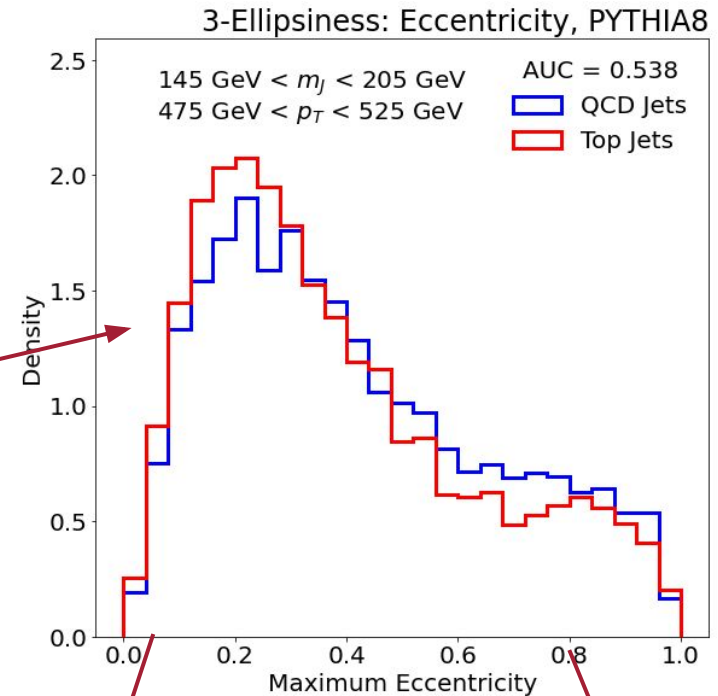
New IRC-Safe Observables

The **SHAPER** framework makes it easy to invent new jet observables!

e.g. **N-Ellipsiness** as a jet algorithm.

- Learn jet centers
- Dynamic jet radii (no R hyperparameter)
- Dynamic **eccentricities** and angles
- Dynamic jet energies
- Learned parameters for discrimination

Can design custom specialized jet algorithms to learn jet substructure!



Low Max Eccentricity (.001)

High Max Eccentricity (.972)

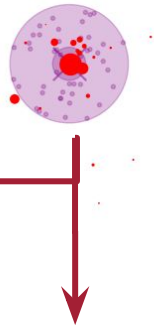
Grooming with Shapes

Use shapes to approximate events and extract masses - model **pileup** with a uniform background with floating weight!

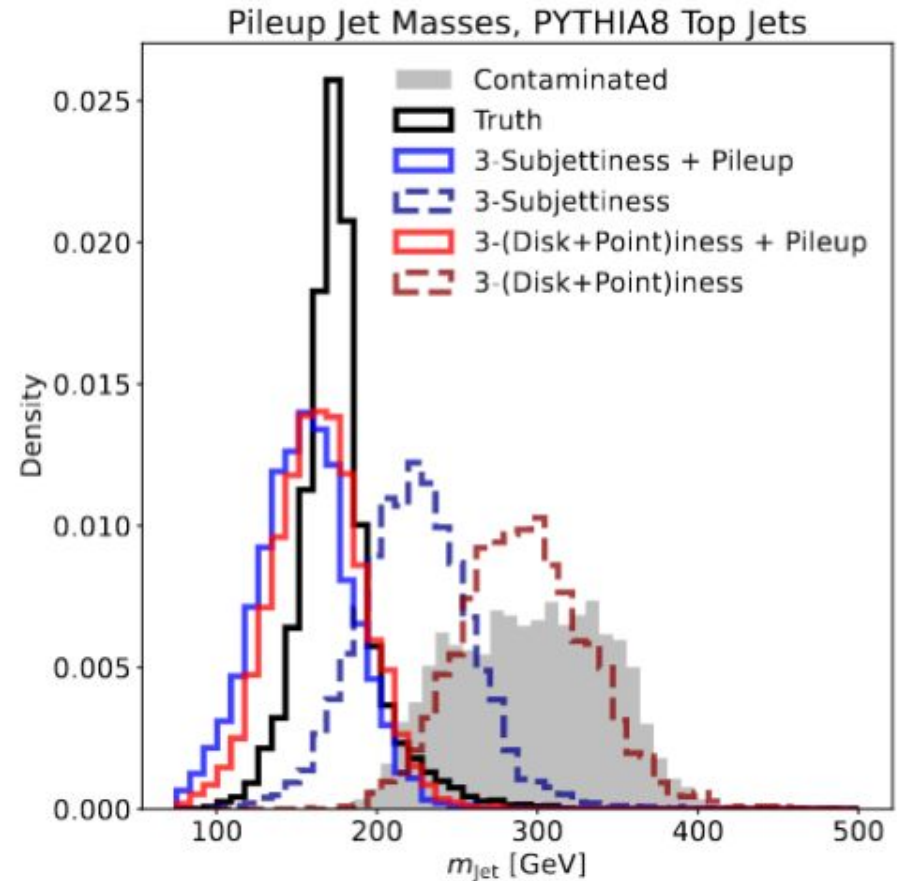
Contaminate top jets with 5-30% extra energy spread uniformly in an 0.8×0.8 plane

Consider 4 shapes:

- 3-Subjettiness
- 3-Subjettiness + Uniform
- 3-(Disk+Point)iness
- 3-(Disk+Point)iness + Uniform



Probing **collinear** (δ -function) and **soft** (shape) structure!



Outlook

- The Wasserstein metric is **the** natural language for jet observables, based on IRC-safety and geometry!
- **SHAPER** is a machine learning framework for calculating generalized observables programmatically!
- Playground for defining and building custom observables and jet algorithms!

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

$$\theta = \operatorname{argmin}_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

Outlook

- The Wasserstein metric is **the** natural language for jet observables, based on IRC-safety and geometry!
- **SHAPER** is a machine learning framework for calculating generalized observables programmatically!
- Playground for defining and building custom observables and jet algorithms!

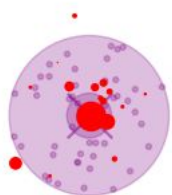
$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$
$$\theta = \operatorname{argmin}_{\mathcal{E}'_{\theta} \in \mathcal{M}} \text{EMD}(\mathcal{E}, \mathcal{E}'_{\theta})$$

Using **SHAPER**, you **CAN** hear the shape of a jet!

More questions? Email me at rikab@mit.edu
Appearing on arXiv soon! (Plus code!)

Appendices

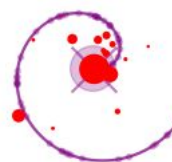
Observables on CMS OpenData



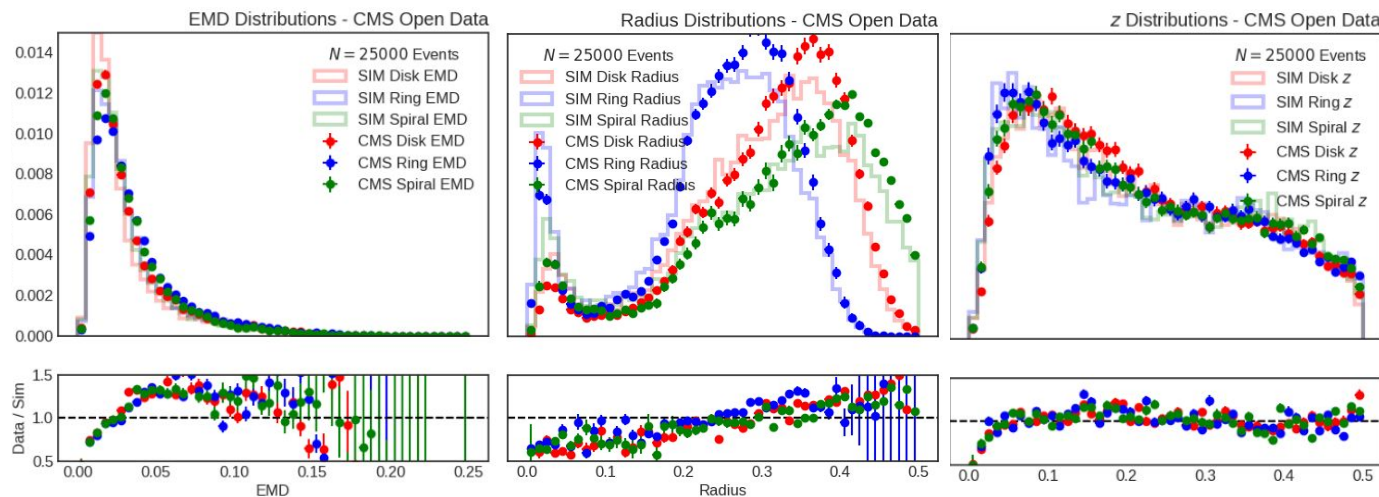
Disk + δ -function



Ring + δ -function

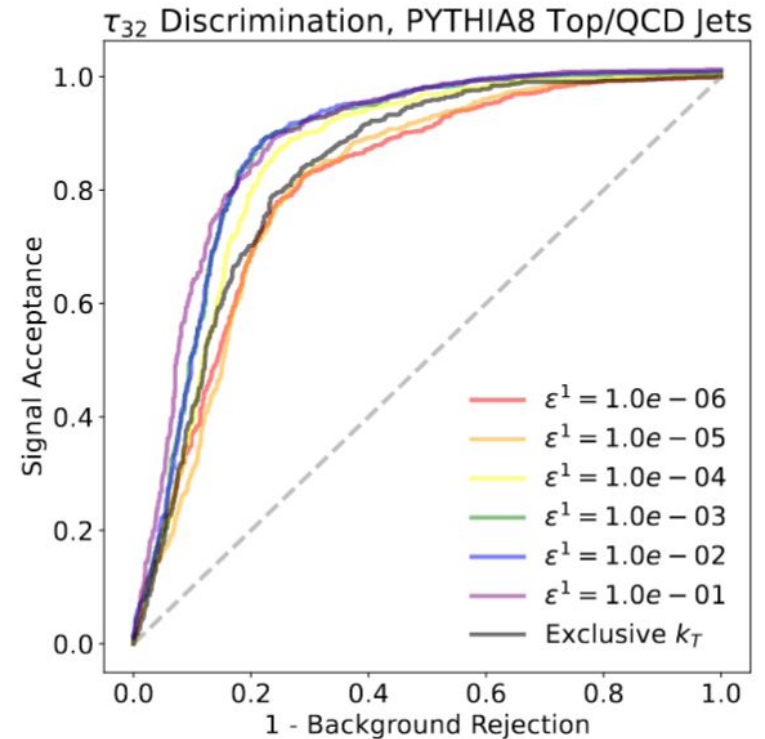
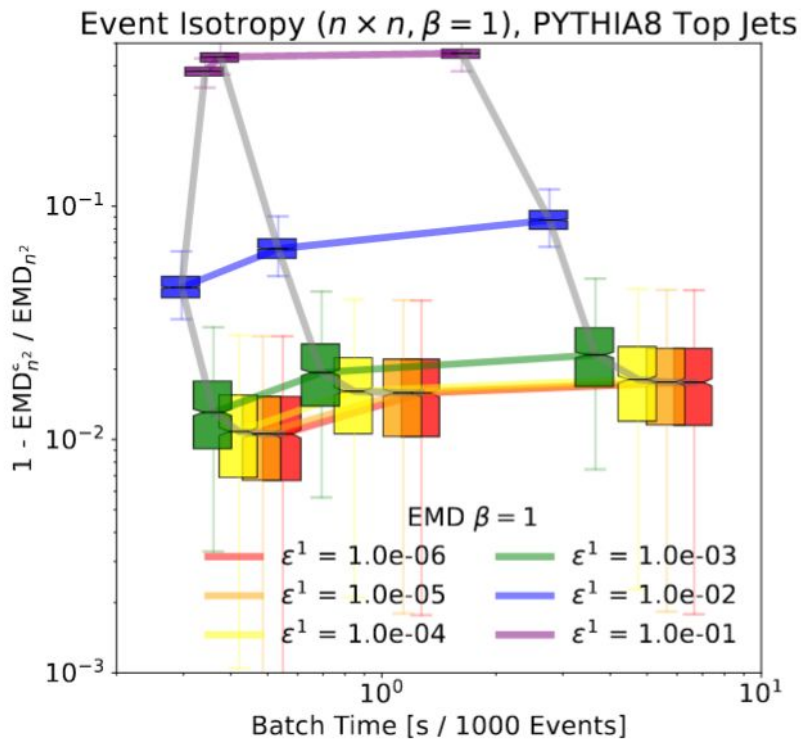


Spiral + δ -function



Probing **collinear** (δ -function) and **soft** (shape) structure!

Performance Benchmarks

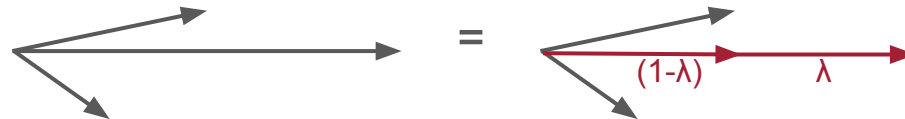


IRC Safety

Infrared Safety: An observable is unchanged under a soft emission



Collinear Safety: An observable is unchanged under a collinear splitting



Building SHAPER

Key Component: The Loss function! Step 1: Manifold Learning

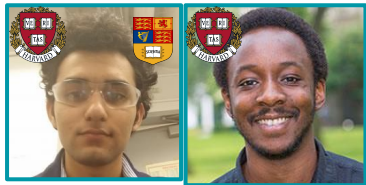
$$\mathcal{L}_R(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geq 0} \left[\sum_{i=1}^M \sum_{j=1}^{M'} \pi_{ij} \frac{|x_i - x'_j|}{R} \right],$$

where $\sum_{i=1}^M \pi_{ij} = 1$

Dogra

Ba

Ai



**K-Deep Simplices,
Dictionary Learning, &
Manifold Learning**

Building SHAPER

Key Component: The Loss function! Step 2: Physical Principles

$$\mathcal{L}_R(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geq 0} \left[\sum_{i=1}^M \sum_{j=1}^{M'} \pi_{ij} \frac{|x_i - x'_j|}{R} \right] + \left| \sum_{i=1}^M z_i - \sum_{j=1}^{M'} z'_j \right|,$$

where $\sum_{i=1}^M \pi_{ij} \leq z'_j$, $\sum_{j=1}^{M'} \pi_{ij} \leq z_i$, $\sum_{i,j} \pi_{ij} = \min \left(\sum_{i=1}^M z_i, \sum_{j=1}^{M'} z'_j \right)$

Dogra

Ba

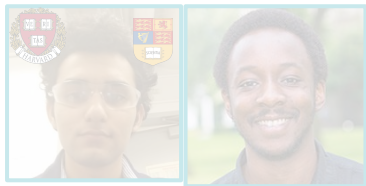
Ai

K-Deep Simplices,
 Dictionary Learning, &
 Manifold Learning

fi

IRC Safety,
 Unclustered Radiation, &
 Wasserstein Geometry

Gambhir Thaler



Building SHAPER

Key Component: The Loss function! Step 3: Synthesis

$$\mathcal{L}_R(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geq 0} \left[\sum_{i=1}^M \sum_{j=1}^{M'} \pi_{ij} \frac{|x_i - x'_j|}{R} \right] + \left| \sum_{i=1}^M z_i - \sum_{j=1}^{M'} z'_j \right|,$$

where $\sum_{i=1}^M \pi_{ij} \leq z'_j$, $\sum_{j=1}^{M'} \pi_{ij} \leq z_i$, $\sum_{i,j}^{M,M'} \pi_{ij} = \min \left(\sum_{i=1}^M z_i, \sum_{j=1}^{M'} z'_j \right)$

Dogra

Ba

Ai

K-Deep Simplices,
Dictionary Learning, &
Manifold Learning

+

fi

IRC Safety,
Unclustered Radiation, &
Wasserstein Geometry

Gambhir Thaler

