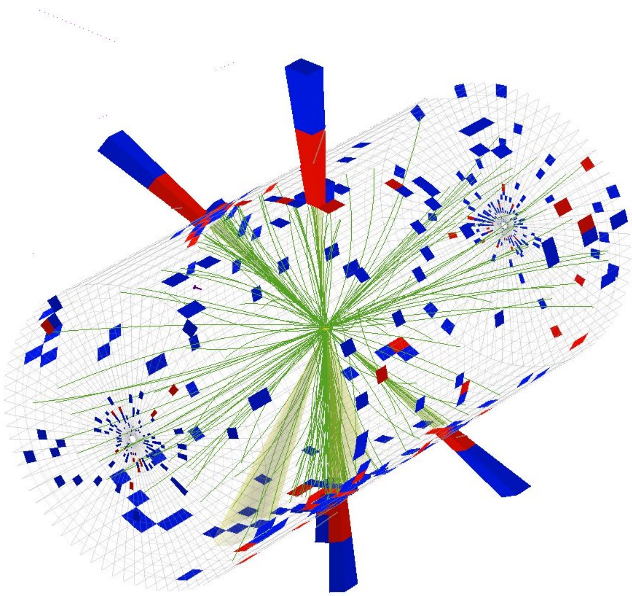
Can You Hear the Shape of a Jet?

Rikab Gambhir

Email me questions at rikab@mit.edu!
Based on [Ba, Dogra, **RG**, Tasissa, Thaler, 2302.12266]
Download with *`pip install pyshaper`*

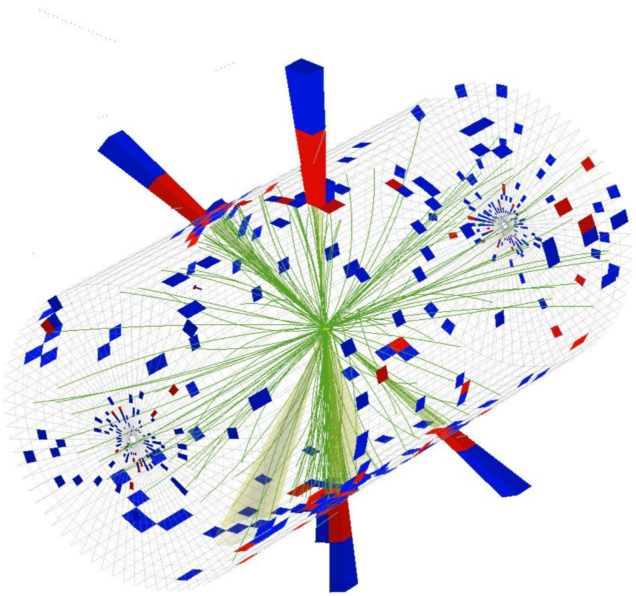
How do we **characterize** collider data?



Complicated collider data – events live in extremely high dimensional LIPS, plus additional quantum numbers!

Can we represent this data in a way that is easier to understand, both experimentally *and* theoretically?

How do we **characterize** collider data?



Complicated collider data – events live in extremely high dimensional LIPS, plus additional quantum numbers!

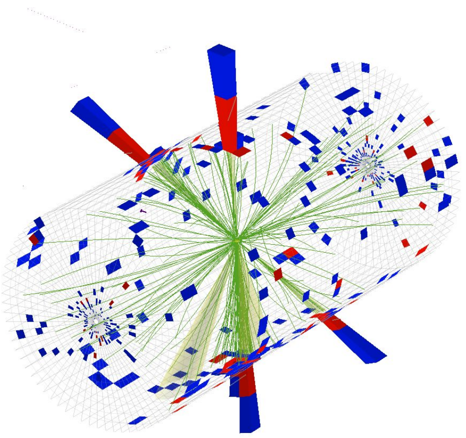
Can we represent this data in a way that is easier to understand, both experimentally *and* theoretically?

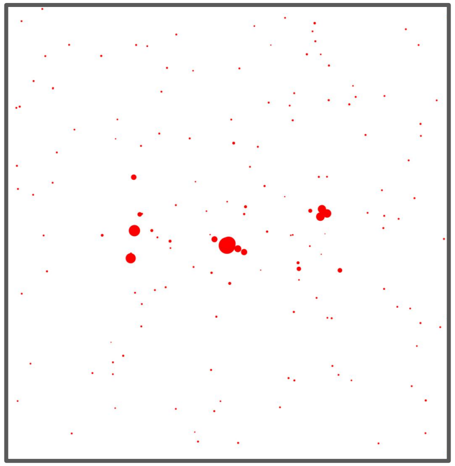
Can you hear the shape of a jet?*

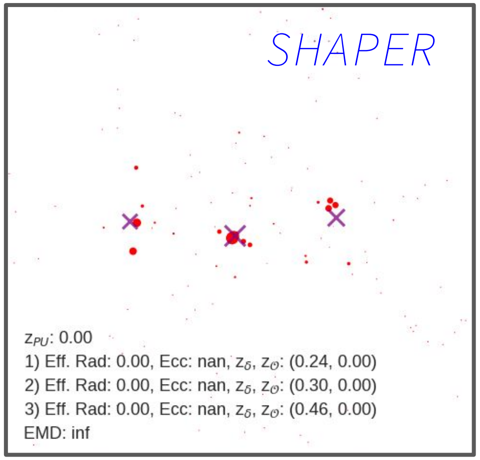
*Explaining this title is mostly outside the scope of this talk

[Ba, Dogra, **RG**, Tasissa, Thaler, 2302.12266]

How do we **characterize** collider data?







Collider Data → **Energy Flow** → **Shapes**

Main Lessons:

1. The **Energy Flow** is a *continuous* embedding of events.
2. Continuity is *IRC-Safety* – theoretically and experimentally robust!
3. Use **Shapes** to summarize energy flows *faithfully*, preserving geometry.

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

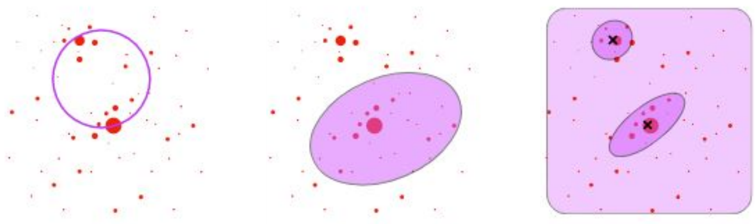
How do we **characterize** collider data?

The Unreasonable Effectiveness of the **Wasserstein** Metric

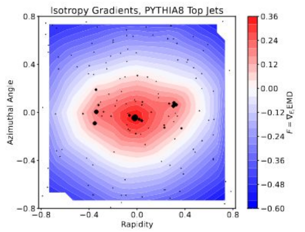
$$\mathrm{EMD}^{(\beta,R)}(\mathcal{E},\mathcal{E}') = \min_{\pi\in\mathcal{M}(\mathcal{X}\times\mathcal{X})}\left[\frac{1}{\beta R^\beta}\left\langle \pi, d(x,y)^\beta\right\rangle\right] + |\Delta E_{\mathrm{tot}}|$$

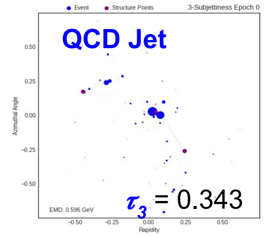
$$\pi(\mathcal{X},Y)\leq\mathcal{E}'(Y),\quad \pi(X,\mathcal{X})\leq\mathcal{E}(X),\quad \pi(\mathcal{X},\mathcal{X}) = \min(E_{\mathrm{tot}},E'_{\mathrm{tot}})$$

Hearing the **Shape** of Jets



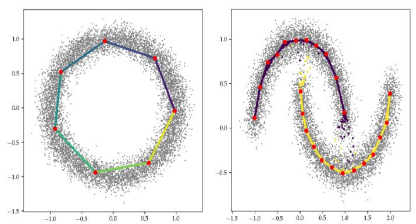
The *SHAPER* Framework



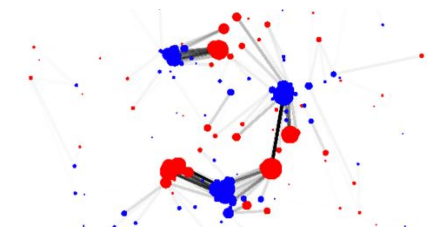


Yes, you *CAN* hear the shape of a jet!

NSF AI Institute for Artificial Intelligence & Fundamental Interactions

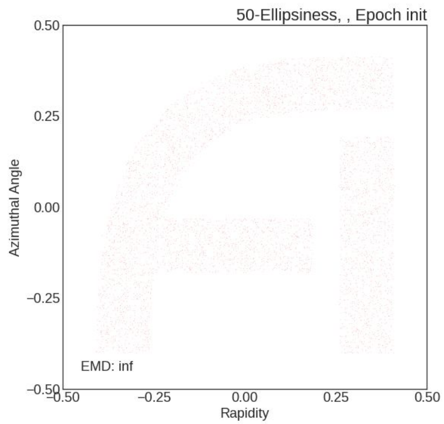


Piecewise-Linear Manifold Approximation with K-Deep Simplices (KDS, 2012.02134)



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!

Section 1

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

The Wasserstein Metric

Collider Physics

Eugene Wigner

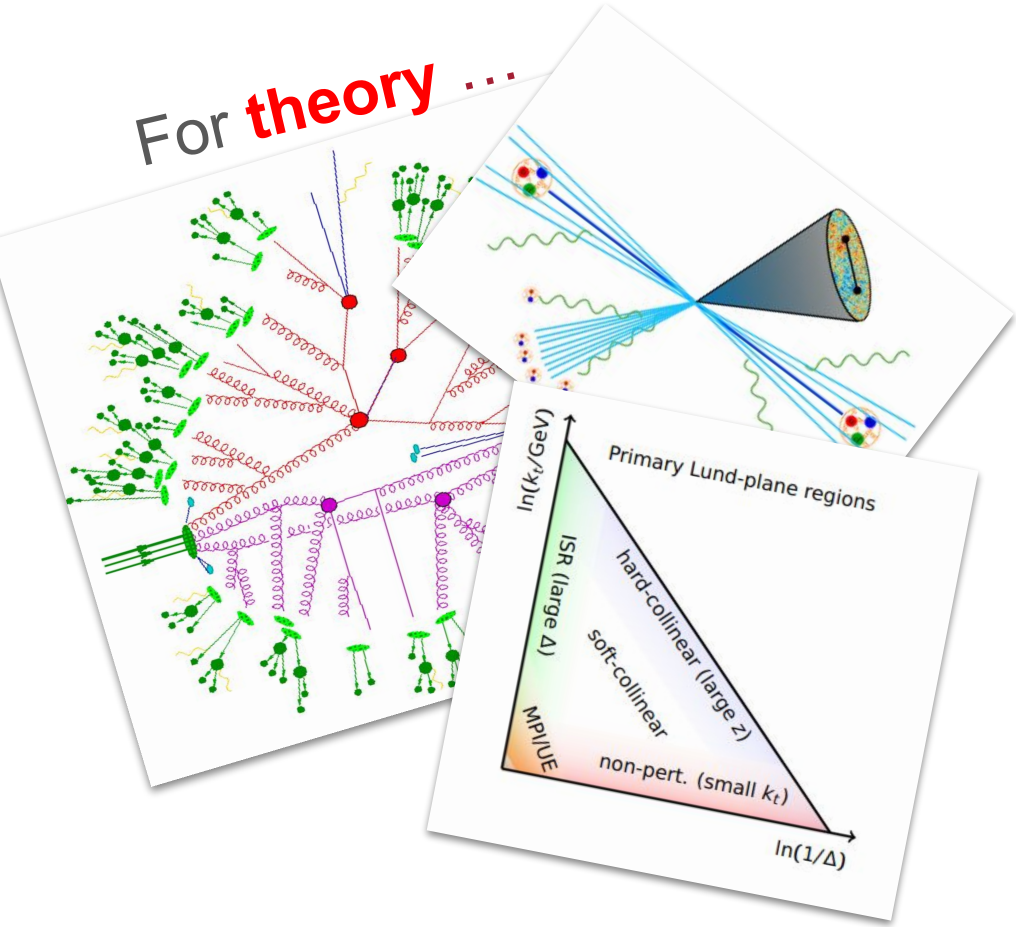
Mathematics, rightly v[illegible]s not only truth, [illegible] supreme beauty cold and austere, like that of sculpture, w[illegible] appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest 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

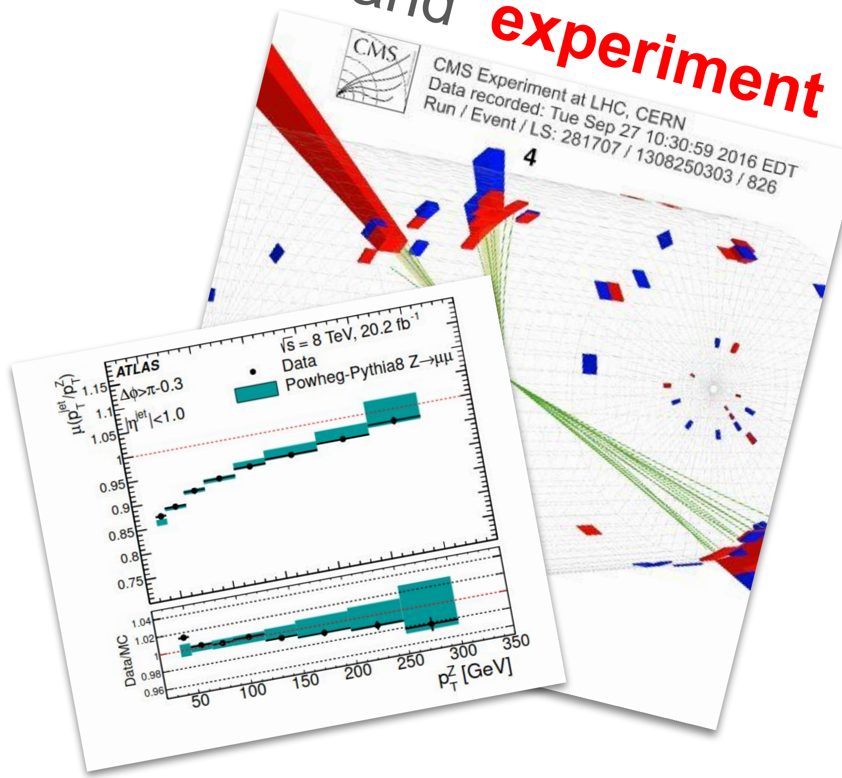
Images from [Bothmann et. al., 1905.09127;
Lee, Meçaj, Moult, 2205.03414;
Dreyer, Salam, Soyez, 1807.04758
CMS, 1810.10069;
ATLAS, 1703.10485]

We want ***Robust Observables***!

For **theory** …



and **experiment**

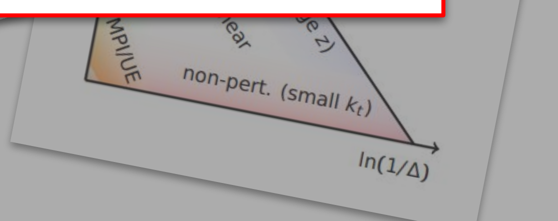


We want ***Robust Observables***!

For theory …

Worry about:

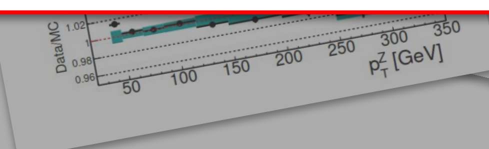
- Perturbativity
- Hadronization
- Choice of Shower
- Interpretability
- … and more



and experiment

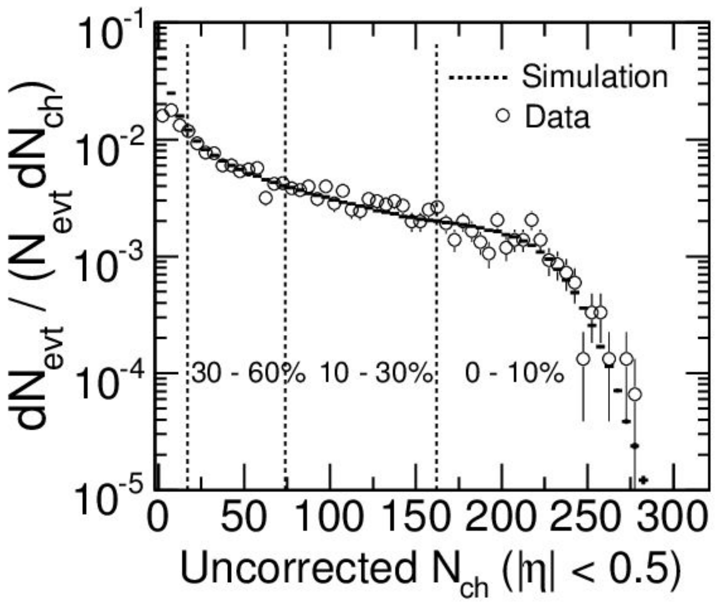
Worry about:

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



The IRC Divergence

We want to predict* the collider output (or at least, some of it).



Lets try calculating the number of particles in a collision!

$$\frac{d\sigma}{dN} = \sum_{N'=0}^{\infty} \sum_{\text{QM\#'s}} \int d\Pi_{N'} \, |\mathcal{M}_{p_1 \ldots p_{N'}, \text{QM\#'s}}|^2 \, \delta(N - N')$$

$$\xrightarrow{N=3} \frac{2\alpha_s C_F}{2\pi} \log\left(\frac{1}{\Lambda_{IR}}\right) \log\left(\frac{1}{\theta_{\text{cut}}}\right) \quad ✗$$

Lesson: You can't predict everything – the **IRC** (infrared and collinear) divergence spoils it!**

*By "predict" in this talk, I mean in perturbative first-principles QCD, with no additional empirical models

**This IR-divergence structure is a generic feature of 4D Yang-Mills. Cannot be cured by renormalization

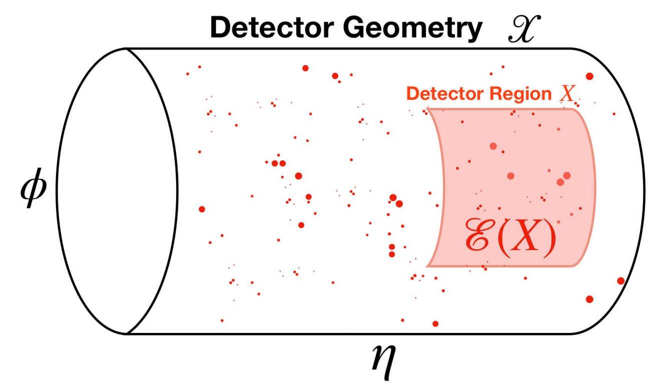
The Energy Flow

What *is* IRC-safe? Embed event data into the **energy flow**:

Detector Coordinate (η, ϕ)

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

Energy Fraction E_i / E_{Tot}



The **energy flow** contains *ALL* IRC-safe information about an event!

- Claim 1: The energy flow is a continuous embedding of events.
- Claim 2: An observable $\mathcal{O}(\mathcal{E})$ is IRC-safe if and only if it is continuous with respect to the *weak* topology* on energy flows.
- Claim 3: The Wasserstein metric is the only *faithful* distance-preserving metric on energy flows.

*What happened to charge, flavor, mass, and other particle information? Ask me!

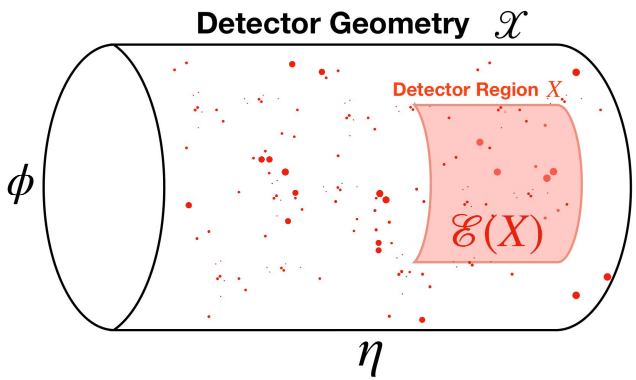
**In this talk, will treat “jets” and “events” as the same thing – the difference here is unimportant

Mathematical Details

Formally, an energy flow is a **measure** on the detector boundary

$$\mathcal{E}(X) = \int_X dx\, \mathcal{E}(x)$$

"How much total energy did I see in the detector region *X*?"



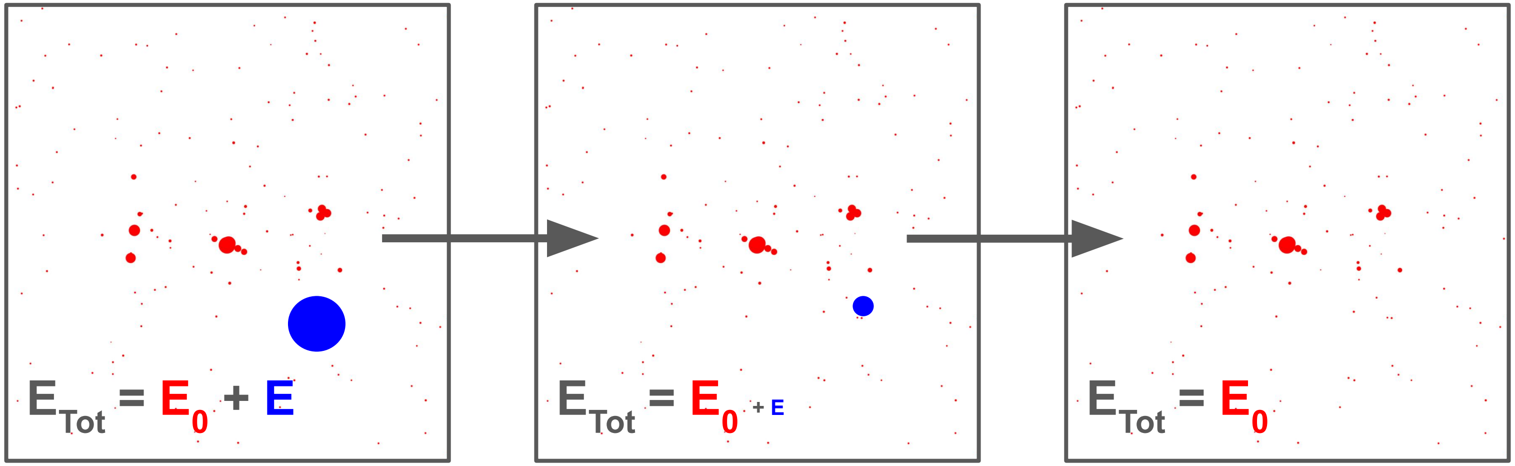
The fundamental operation on measures is **integration**:

$$\langle \mathcal{E}, \phi \rangle \equiv \int_{\mathcal{X}} dx\, \mathcal{E}(x)\, \phi(x)$$

"What is the energy-weighted expectation value of ϕ?"

Mathematical Details – Topology

Definition (The **Weak* Topology**): A sequence of measures converges if all of their expectation values converge, as real numbers.

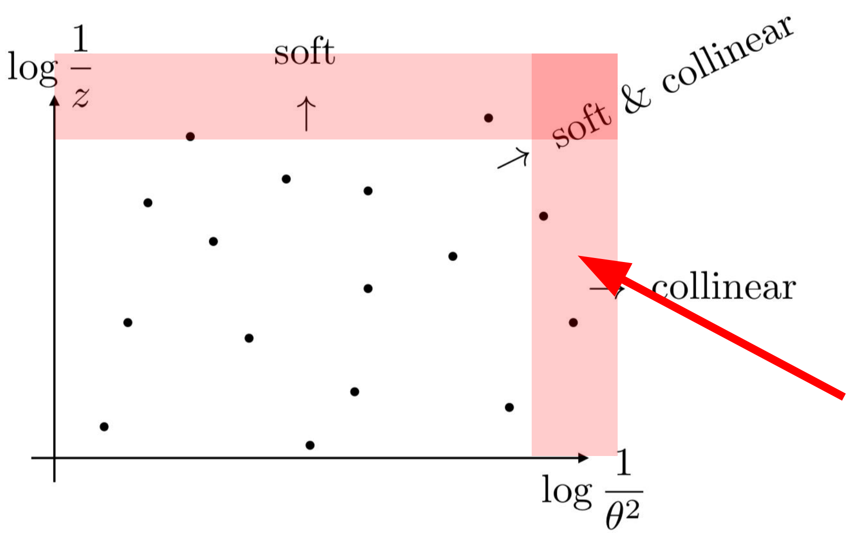


Definition (**Weak* Continuity**): An observable $\mathcal{O}(\mathcal{E})$ is continuous with respect to energy flows if, for any sequence of measures $\mathcal{E}_n$ that converges to $\mathcal{E}$, the sequence of real numbers $\mathcal{O}(\mathcal{E}_n)$ converges to $\mathcal{O}(\mathcal{E})$.

Topology ⇔ IRC-Safety

Two ways to change the expectation values of an energy flow:

1. Change a particle's energy slightly, or add a low-energy particle - **IR**
2. Move a particle's position slightly, or split particles in two - **C**



An observable $\mathcal{O}$ is continuous if it changes only slightly under the above perturbations.

The regions of phase space causing IRC divergences is suppressed — $\mathcal{O}$ is **IRC-Safe**!

Proved in [Ba, Dogra, **RG**, Tasissa, Thaler, 2302.12266]
Ask me later for proof details!

Mathematical Details - Geometry

When are two events similar? We need a metric to compare!

Properties we want:

1. … is non-negative, non-degenerate, symmetric and finite.
2. … is weak* continuous (IRC-safe)
3. … lifts the detector metric **faithfully**

Explained shortly!

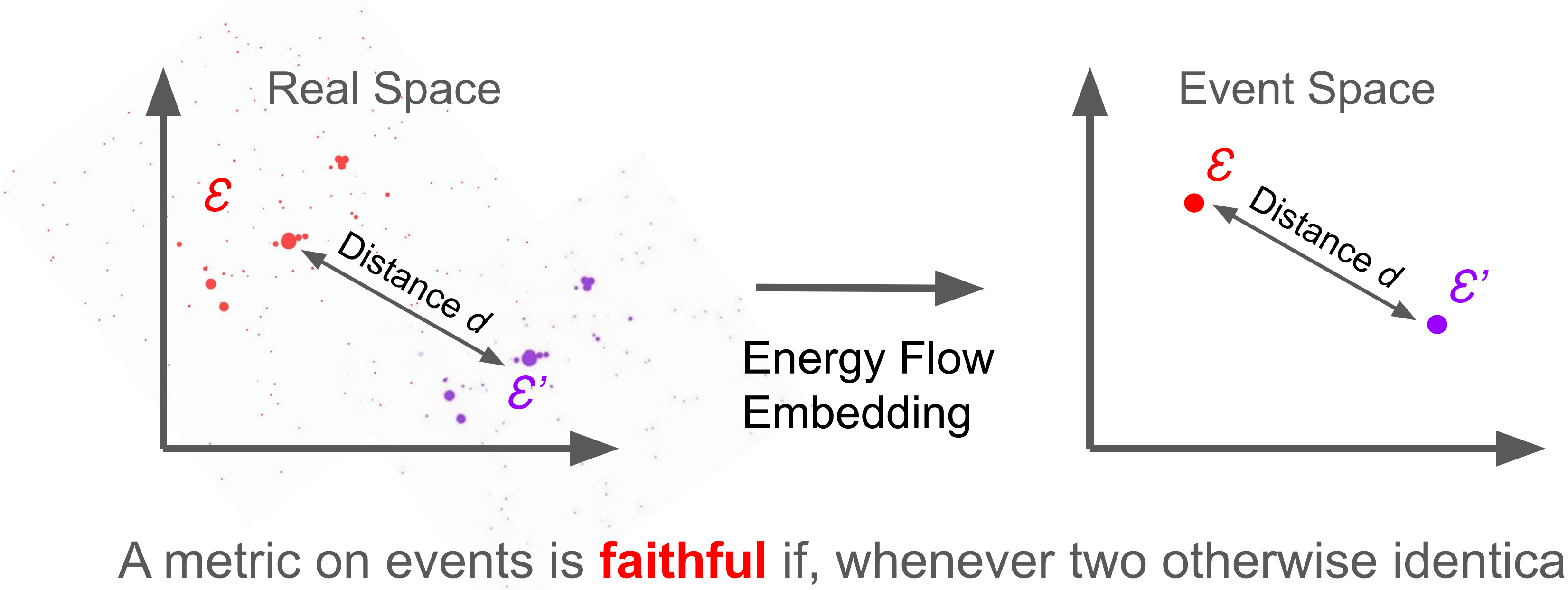
The only* metric on distributions satisfying the above is the **Wasserstein Metric**:

$$\text{EMD}^{(\beta,R)}(\mathcal{E},\mathcal{E}') = \min_{\pi\in\mathcal{M}(\mathcal{X}\times\mathcal{X})}\left[\frac{1}{\beta R^\beta}\left\langle \pi, d(x,y)^\beta \right\rangle\right] + |\Delta E_{\text{tot}}|$$

$$\pi(\mathcal{X},Y) \leq \mathcal{E}'(Y), \quad \pi(X,\mathcal{X}) \leq \mathcal{E}(X), \quad \pi(\mathcal{X},\mathcal{X}) = \min(E_{\text{tot}}, E'_{\text{tot}})$$

*There exist other metrics on distributions that are faithful only for very specific real-space distance norms, but we want them all!

The Importance of Being **Faithful***



A metric on events is **faithful** if, whenever two otherwise identical events ℰ and ℰ' are separated in real space by a distance *d*, the distance between the events is also *d*. Or *d* to a constant power

Only the **Wasserstein Metric** does this! → Can use to build **event** and **jet shapes** (old and new)!

Faithfulness also ensures very nice **numerical properties**, including no vanishing or exploding gradients.

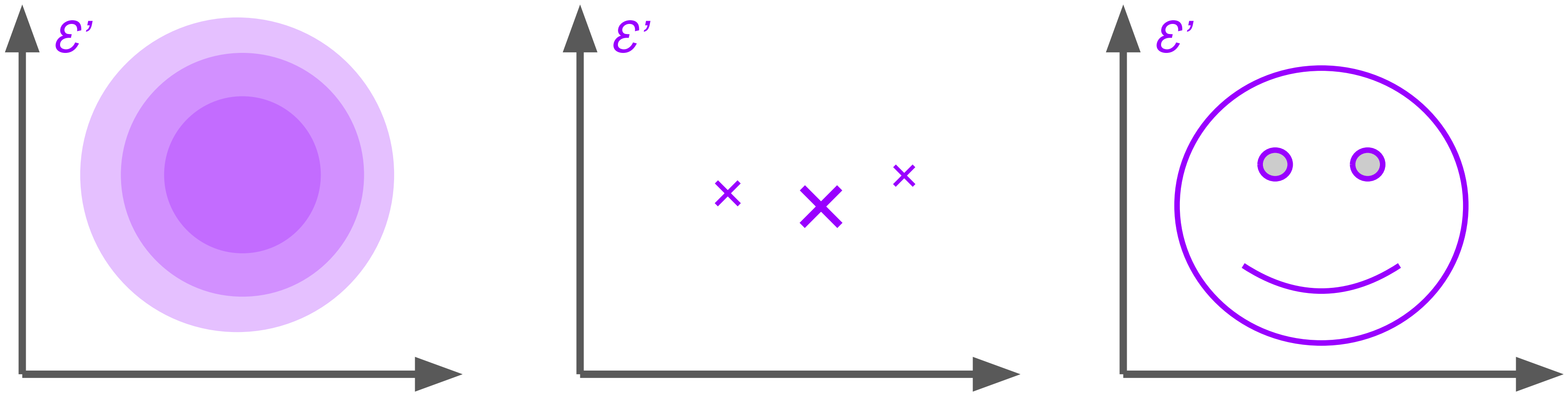
Section 2

Hearing the Shape of Jets

Shapes as Energy Flows

Energy flows don't have to be real events – they can be *any* energy distribution in detector space, or **shape**.

Can make anything you want! Even continuous or complicated shapes.
(Or, something you calculate in perturbative QCD)



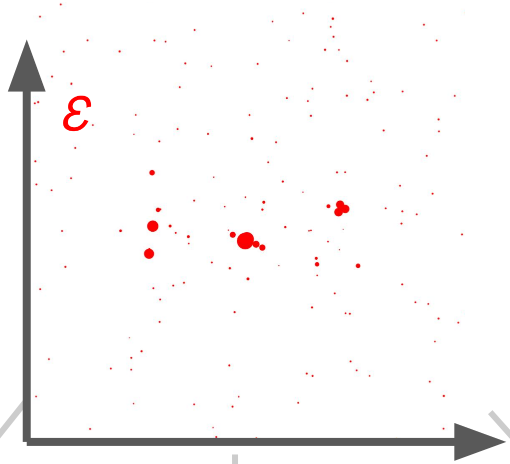
Shape = 2D Gaussian

Shape = 3 Points

Shape = Smile

Shapiness

The EMD between a real event or jet $\mathcal{E}$ and idealized shape $\mathcal{E}'$ is the [shape]iness of $\mathcal{E}$ – a well defined IRC-safe observable!

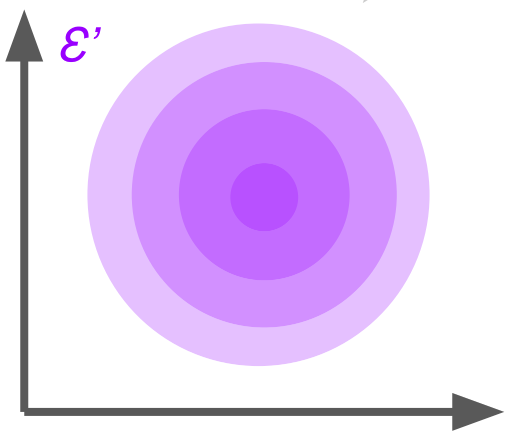


Answers the question: "How much like the shape $\mathcal{E}'$ is my event $\mathcal{E}$?"

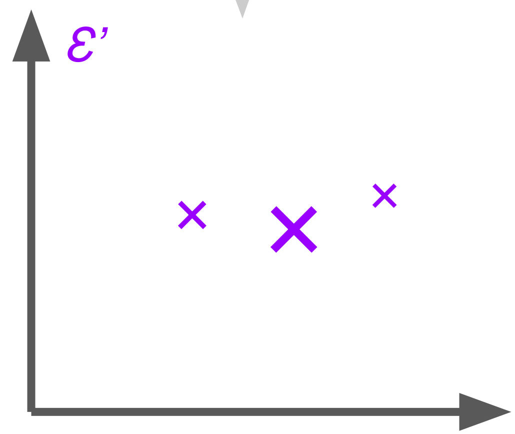
Med EMD($\mathcal{E}$, $\mathcal{E}'$)
"Gausiness"

Low EMD($\mathcal{E}$, $\mathcal{E}'$)
"3-Pointiness"
AKA "3-Subjettiness"

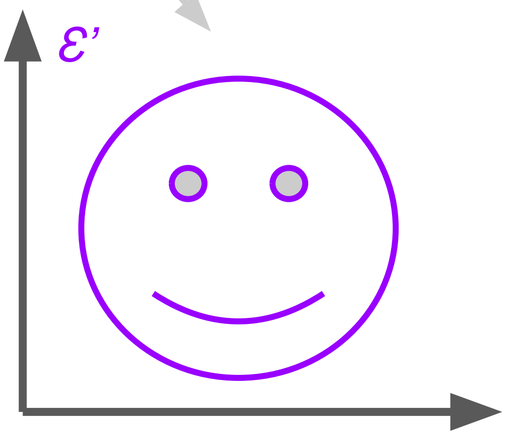
High EMD($\mathcal{E}$, $\mathcal{E}'$)
"Smileyness"



Shape = 2D Gaussian



Shape = 3 Points



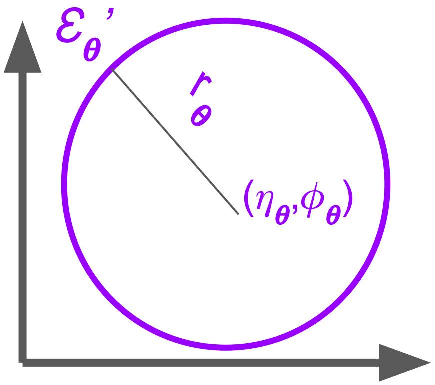
Shape = Smile

Mathematical Details - Shapiness

Rather than a single shape, consider a **parameterized manifold** $\mathcal{M}$ of **energy flows**.

e.g. The manifold of uniform circle energy flows:

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



Then, for an event $\mathcal{E}$, define the **shapiness** $\mathcal{O}_\mathcal{M}$ and **shape parameters** $\theta_\mathcal{M}$, given by:

$$\mathcal{O}_\mathcal{M}(\mathcal{E}) \equiv \min_{\mathcal{E}_\theta \in \mathcal{M}} \text{EMD}^{(\beta,R)}(\mathcal{E}, \mathcal{E}_\theta)$$

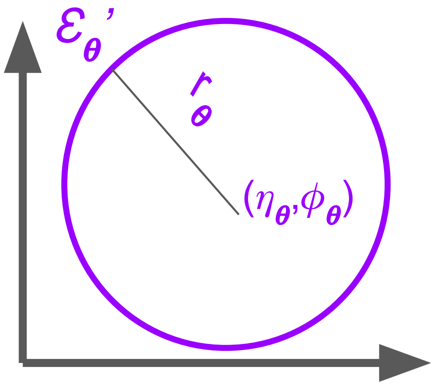
$$\theta_\mathcal{M}(\mathcal{E}) \equiv \underset{\mathcal{E}_\theta \in \mathcal{M}}{\text{argmin}}\, \text{EMD}^{(\beta,R)}(\mathcal{E}, \mathcal{E}_\theta)$$

Mathematical Details - Shapiness

Rather than a single shape, consider a **parameterized manifold** $\mathcal{M}$ of **energy flows.**

e.g. The manifold of uniform circle energy flows:

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



Then, for an event $\mathcal{E}$, define the **shapiness** $\mathcal{O}_\mathcal{M}$ and **shape parameters** $\theta_\mathcal{M}$, given by:

$$\mathcal{O}_\mathcal{M}(\mathcal{E}) \equiv \min_{\mathcal{E}_\theta \in \mathcal{M}} \text{EMD}^{(\beta,R)}(\mathcal{E}, \mathcal{E}_\theta)$$

$$\theta_\mathcal{M}(\mathcal{E}) \equiv \underset{\mathcal{E}_\theta \in \mathcal{M}}{\text{argmin}}\, \text{EMD}^{(\beta,R)}(\mathcal{E}, \mathcal{E}_\theta)$$

[P. Komiske, E. Metodiev, and J. Thaler, 2004.04159;
J. Thaler, and K. Van Tilburg, 1011.2268;
I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, 1004.2489.;
S. Brandt, C. Peyrou, R. Sosnowski and A. Wroblewski, PRL 12 (1964) 57-61;
C. Cesarotti, and J. Thaler, 2004.06125]

Observables ⇔ Manifolds of Shapes

Observables can be specified solely by defining a **manifold of shapes**:

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) \equiv \min_{\mathcal{E}_\theta \in \mathcal{M}} \text{EMD}^{(\beta,R)}(\mathcal{E}, \mathcal{E}_\theta),$$

$$\theta_{\mathcal{M}}(\mathcal{E}) \equiv \underset{\mathcal{E}_\theta \in \mathcal{M}}{\text{argmin}}\, \text{EMD}^{(\beta,R)}(\mathcal{E}, \mathcal{E}_\theta),$$

Many well-known observables* already have this form!

Observable	Manifold of Shapes
N-Subjettiness	Manifold of *N*-point events
N-Jettiness	Manifold of *N*-point events with floating total energy
Thrust	Manifold of back-to-back point events
Event / Jet Isotropy	Manifold of the single uniform event

… and more!

All of the form "How much like **[shape]** does my **event** look like?"
Generalize to *any* shape.

*These observables are usually called event shapes or jet shapes in the literature – we are making this literal!

Hearing Jets Ring

(and Disk, and Ellipse)

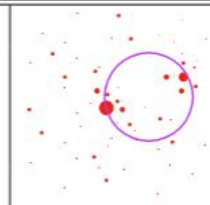
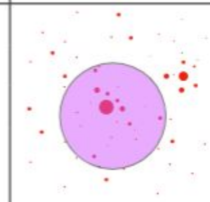
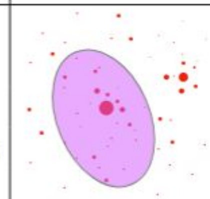
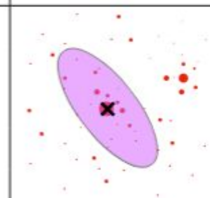
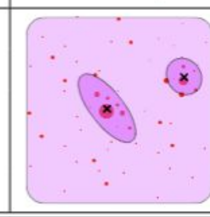
$\mathcal{O}_{\mathcal{M}}(\mathcal{E})$ answers: “How much like a shape in $\mathcal{M}$ does my event $\mathcal{E}$ look like?”

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

Can define complex manifolds to probe increasingly subtle geometric structure!

Disk + δ-function   **Ring** + δ-function

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

Shape	Specification	Illustration
Ringiness $\mathcal{O}_R$	**Manifold of Rings** $\mathcal{E}_{x_0,R_0}(x) = \frac{1}{2\pi R_0}$ for $\lvert x - x_0\rvert = R_0$; x_0 = Center, R_0 = Radius	
Diskiness $\mathcal{O}_D$	**Manifold of Disks** $\mathcal{E}_{x_0,R_0}(x) = \frac{1}{\pi R_0^2}$ for $\lvert x - x_0\rvert \leq R_0$; x_0 = Center, R_0 = 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 = Center, a, b = Semi-axes, φ = 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 new shapes you can define!

Some Nice Properties

Shapes are an *infinite class* of observables, generalizing the *N*-subjettiness, and shape parameters generalize cone-type jet algorithms.

- **IRC-Safety:** Shapiness is IRC-safe and is, *in principle*, numerically calculable in perturbation theory. It's not clear if shape parameters are also always IRC-safe, but they seem to be
- **Monotonicity:** If manifold 1 contains manifold 2, then the observable corresponding to manifold 1 will be less than or equal to the observable corresponding to manifold 2.
- **Closure:** The shapiness will be 0 if and only if the event is contained within the manifold already.
- **Approximation Bounds:** The EMD can be used to bound Lipschitz-additive observables.
- **Upper Bounds:** If the detector is bounded by a maximum real space distance, then that is the maximum value for any shapiness.

Next: How do we actually **calculate** these?

Section 3

The *SHAPER* Framework

Download with *pip install pyshaper*

The *SHAPER** Framework

Shape-**H**unting **A**lgorithm using **P**arameterized **E**nergy **R**econstruction

- 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 shapiness and optimal shape parameters

```python
# SHAPER
from pyshaper.CommonObservables import buildCommmonObservables
from pyshaper.Observables import Observable
from pyshaper.Shaper import Shaper

# Use Pre-built Observables (N-subjets, rings, disks, ellipses)
observables, pointers = buildCommmonObservables(N = 3, beta = 1, R = 0.8)

# Make new observables by defining energy probability distributions
def uniform_sampler(N, param_dict):
    points = torch.FloatTensor(N, 2).uniform_(-0.8, 0.8).to(device)
    zs = torch.ones((N,)).to(device) / N
    return (points, zs)

observables["Isotropy"] = Observable({}, uniform_sampler, beta = 1, R = 0.8)

# Run SHAPER on data
shaper = Shaper(observables, device = "cpu")
shaper.to(device)
emds, params = shaper.calculate(dataset)

# Done!
```

Example usage of *pySHAPER*, a python implementation of *SHAPER*.

*Thanks to Sam for helping come up with this name.

[J. Feydy, tel.archives-ouvertes.fr/tel-02945979;
B. Charlier, J. Feydy, J. Alexis Glaunès F. D. Collin, G. Durif, JMLR:v22:20-275;
J. Feydy, T. Séjourné, F. X. Vialard, S. Amari, A. Trouvé, G. Peyré, 1810.08278;
O. Kitouni, N. Nolte, M Williams, 2209.15624]

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:

Algorithm 3.4: Symmetric Sinkhorn algorithm, with debiasing

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

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

Implemented using the KerOps+GeomLoss Python Package!

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

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

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

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

See “Finding NEEMo” for alternatives!

Algorithm

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

SHAPER

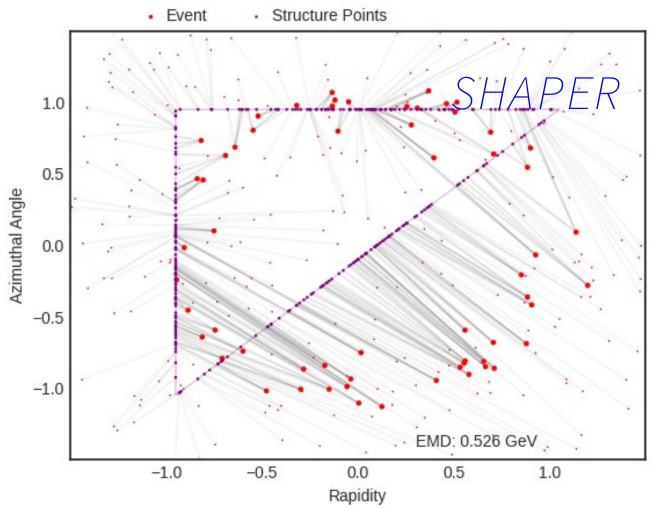
To estimate a shape observable …

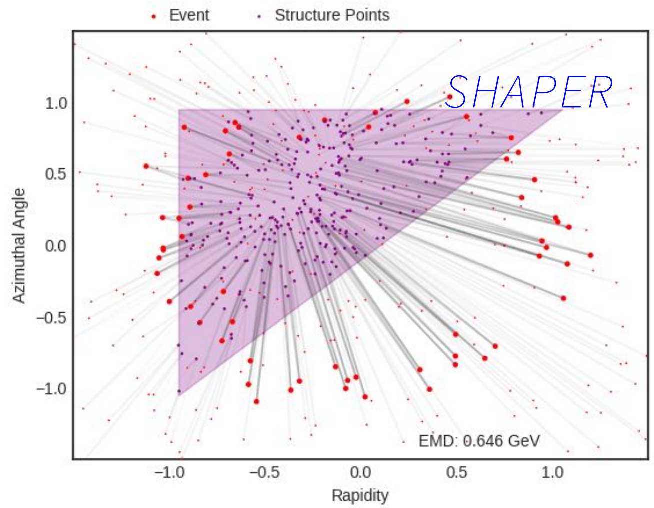
1. Define a parameterized distribution that can be sampled ⇔ **Manifold** $\mathcal{M}$
2. Initialize the parameters $\boldsymbol{\theta}$ in an IRC-safe way (Usually k_T)
3. Use the Sinkhorn Algorithm to estimate the EMD between your **event** $\boldsymbol{\mathcal{E}}$ and **shape** $\boldsymbol{\mathcal{E}_\theta}$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

EMD and Shape ($\mathcal{O}$, θ)

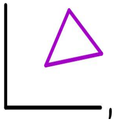
Fun **Animations**

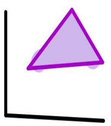
How **triangle-y** is an event? (Boundary or filled in)?





Red: Event ε (Not real)
Purple: Shape ε_θ
Grey: Matrix π_{ij} connecting *particles* and *triangle*

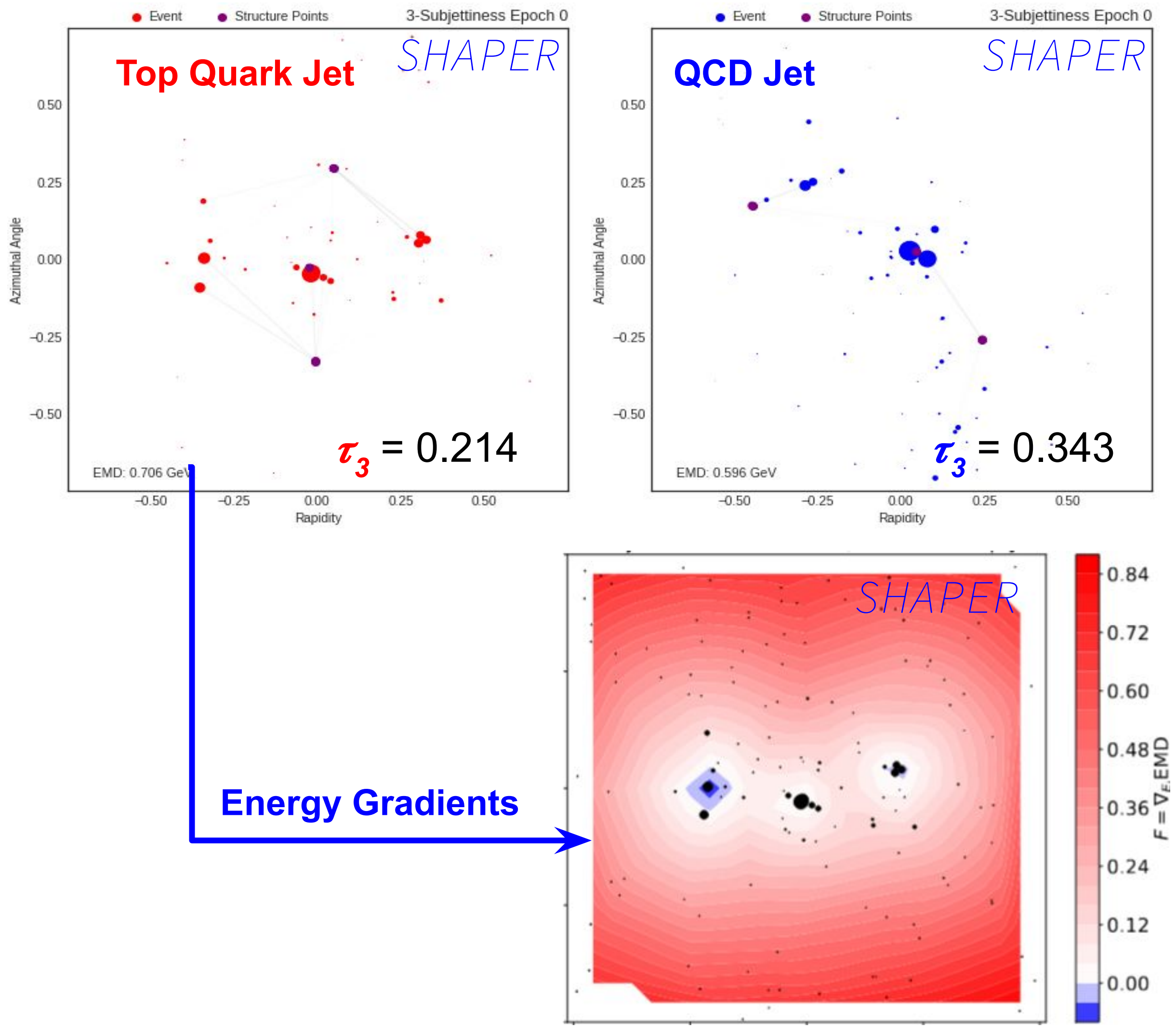
Left: ε_θ =  EMD = 0.245

Right: ε_θ =  EMD = 0.279

N-Subjettiness

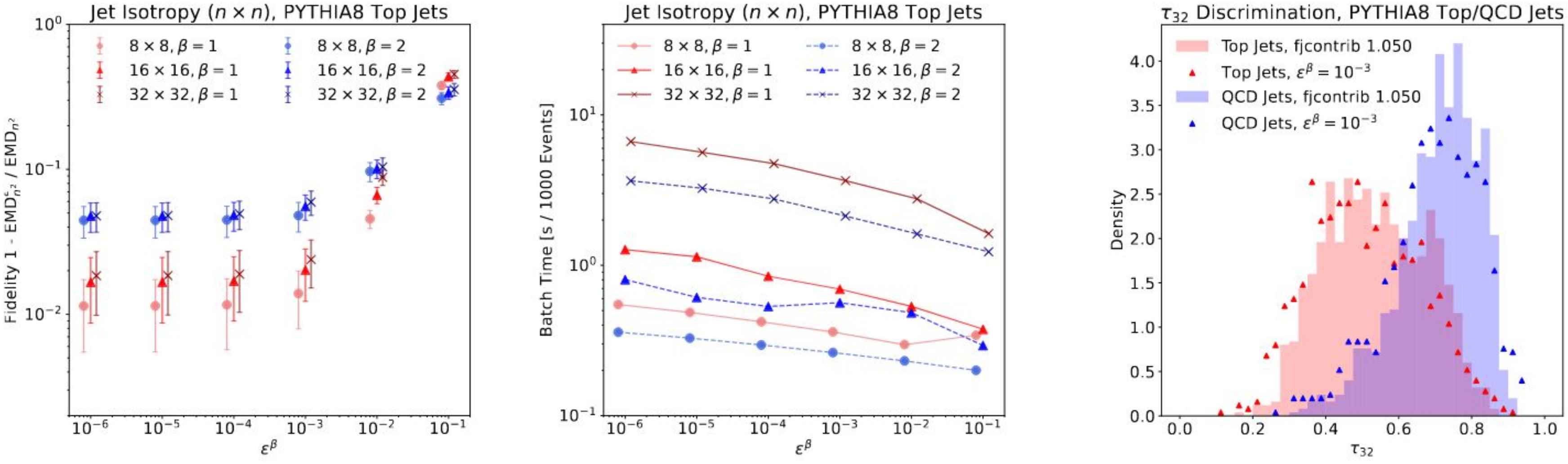
Easy to compute your favorite classic jet observables!

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



Details - Fidelity and Performance

Comparison of *pySHAPER* implementation to other techniques for **Jet Isotropy** and ***N*-Subjettiness**:



It's pretty good!

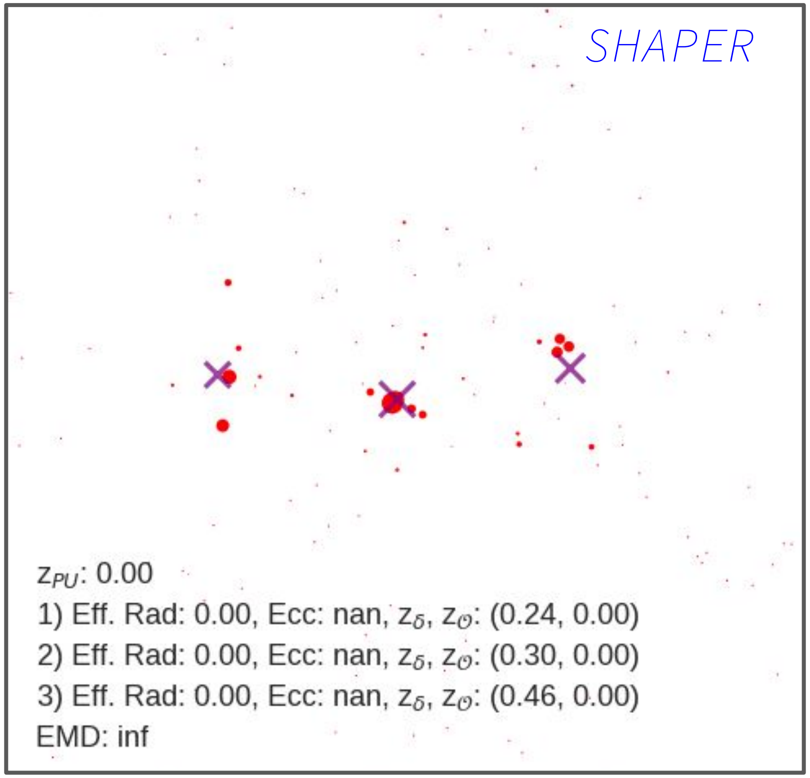
[see also L., B. Nachman, A. Schwartzman, C. Stansbury, 1509.02216;
see also B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, C. Wanotayaroj, 1407.2922;
see also M. Cacciari, G. Salam, 0707.1378]

New IRC-Safe Observables

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

e.g. ***N-(*Ellipse+Point)iness+Pileup** as a jet algorithm:

- Learn jet centers + collinear radiation
- Dynamic jet radii (no *R* parameter!)
- Dynamic eccentricities and angles
- Dynamic jet energies
- Dynamic pileup (no z_{cut} parameter!!!)
- Learned parameters for discrimination



Can design custom specialized jet algorithms to learn jet substructure!

[see also L., B. Nachman, A. Schwartzman, C. Stansbury, 1509.02216;
see also B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, C. Wanotayaroj, 1407.2922;
see also M. Cacciari, G. Salam, 0707.1378]

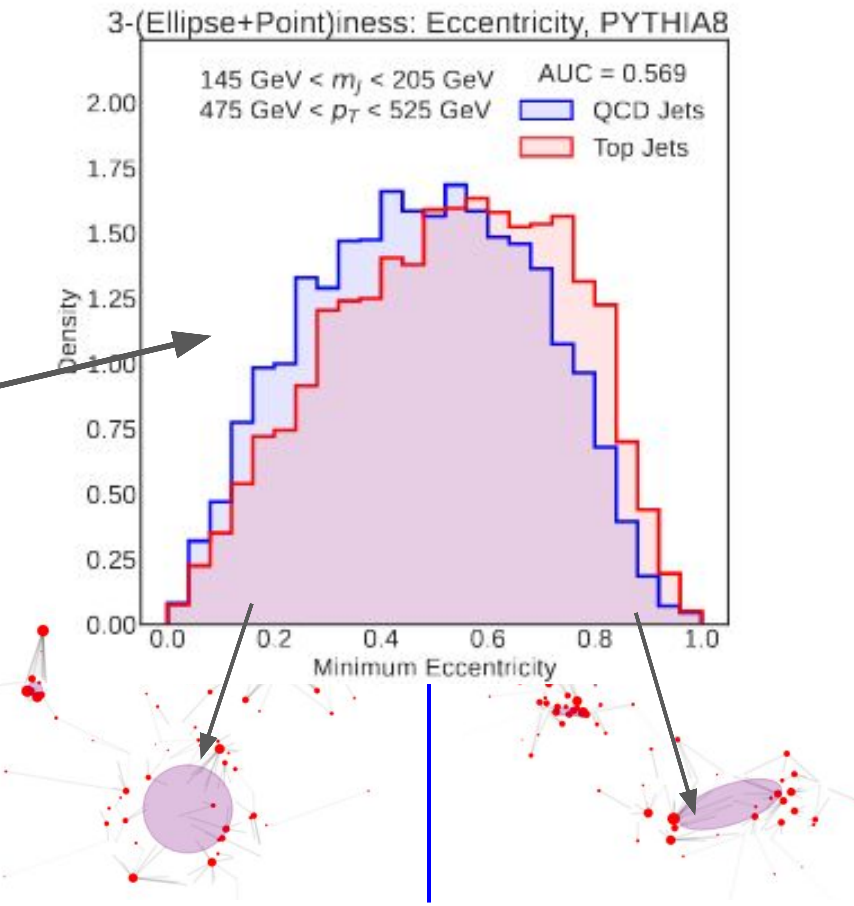
New IRC-Safe Observables

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

e.g. ***N-(*Ellipse+Point)iness+Pileup** as a jet algorithm:

- Learn jet centers + collinear radiation
- Dynamic jet radii (no *R* parameter!)
- Dynamic **eccentricities** and angles
- Dynamic jet energies
- Dynamic pileup (no z_{cut} parameter!!!)
- Learned parameters for discrimination

Can design custom specialized jet algorithms to learn jet substructure!



Low Eccentricity (.001)   High Eccentricity (.972)

[see also L., B. Nachman, A. Schwartzman, C. Stansbury, 1509.02216;
see also B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, C. Wanotayaroj, 1407.2922;
see also M. Cacciari, G. Salam, 0707.1378]

New IRC-Safe Observables

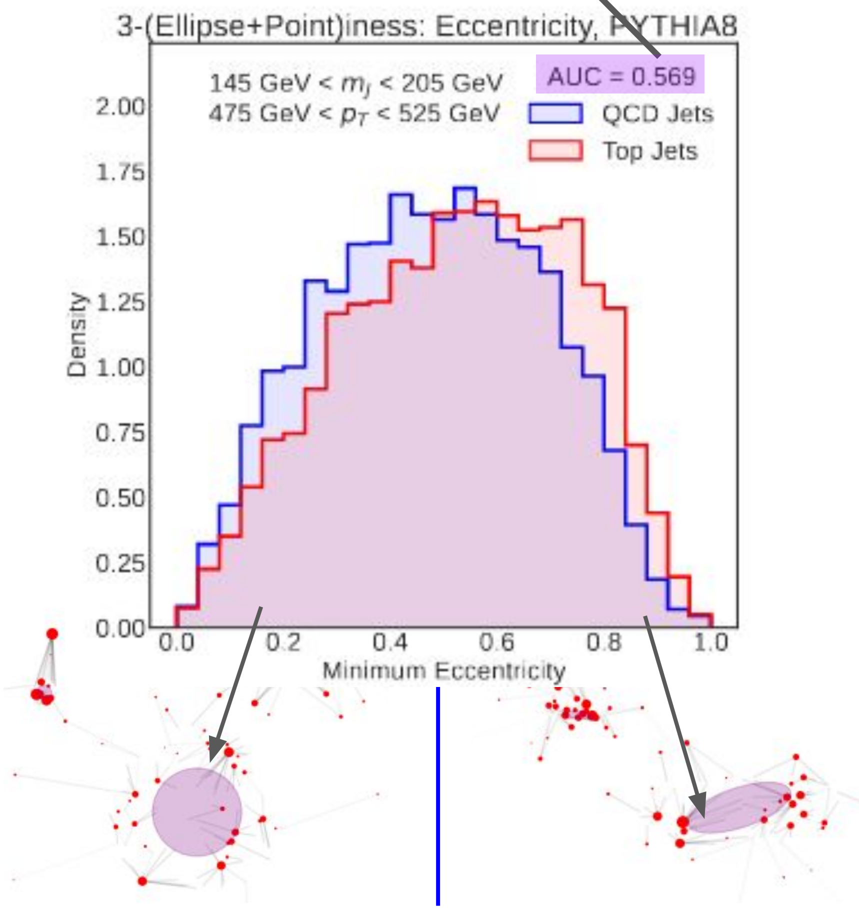
Eccentricity can distinguish between top/QCD jets? Nontrivial result!

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

e.g. ***N*-(Ellipse+Point)iness+Pileup** as a jet algorithm:

- Learn jet centers + collinear radiation
- Dynamic jet radii (no *R* parameter!)
- Dynamic **eccentricities** and angles
- Dynamic jet energies
- Dynamic pileup (no z_{cut} parameter!!!)
- Learned parameters for discrimination

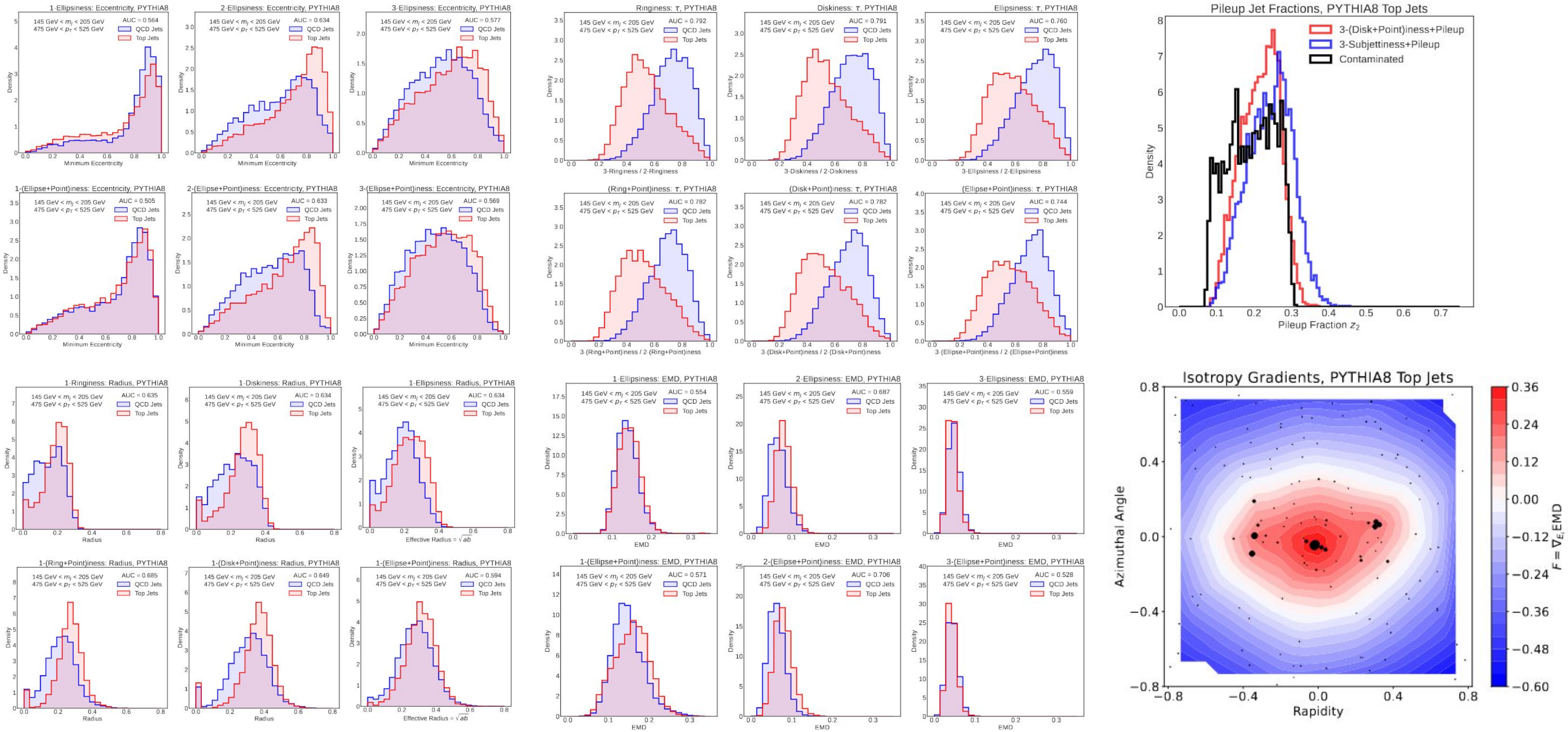
Can design custom specialized jet algorithms to learn jet substructure!



Low Eccentricity (.001)

High Eccentricity (.972)

New IRC-Safe Observables



… **Lots** of extractable information!

Automatic Grooming with Shapes

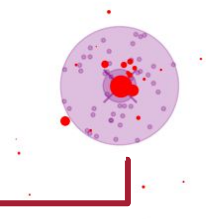
Use shapes to approximate events and extract masses – model pileup with a uniform background with floating weight!

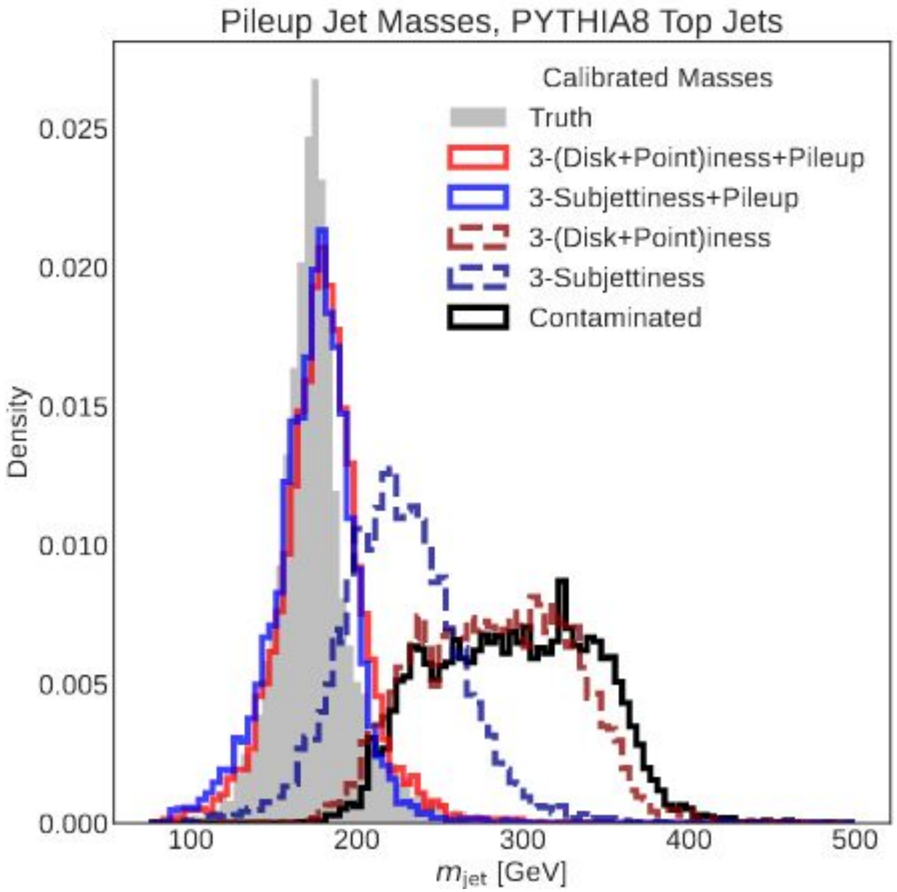
No external hyperparameters, unlike softdrop. Only need to assume pileup is uniform!

Contaminate top jets with 5-30% extra energy spread uniformly in an 0.8x0.8 plane

Consider 4 shapes:

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

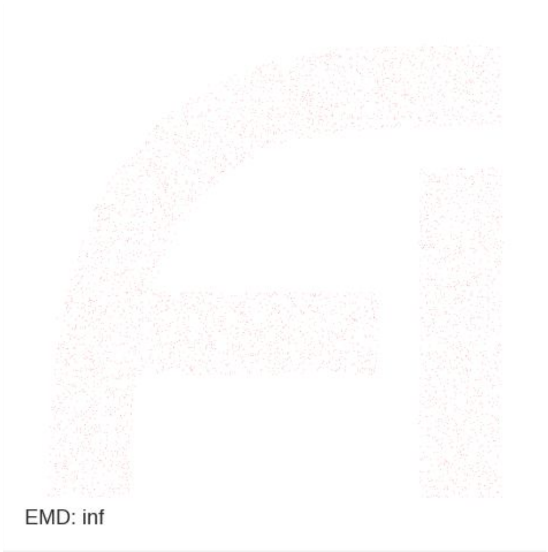




Can also consider ellipses instead of disks – only marginally better performance

Some Last Fun **Animations**

The **50-** and **100-Ellipsinesses** of some (probably fake) collider events



… Can you hear the shape of these “jets”?

Outlook.

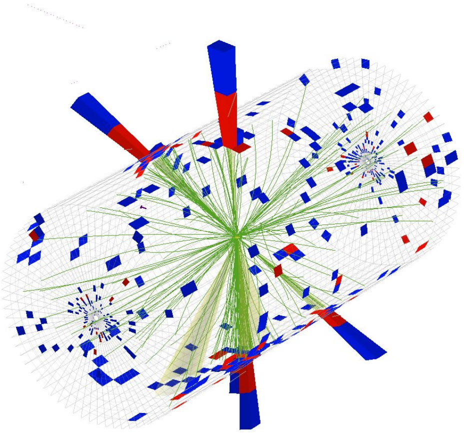
Future Thinking

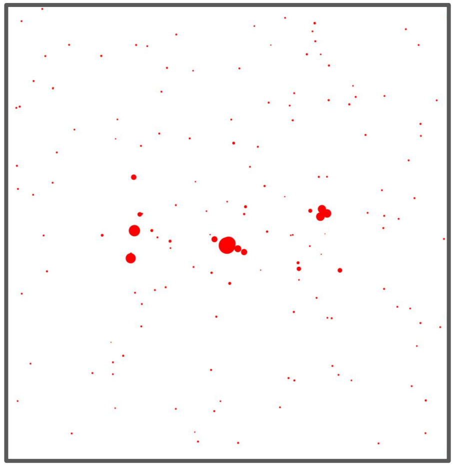
Some things I am thinking about still…

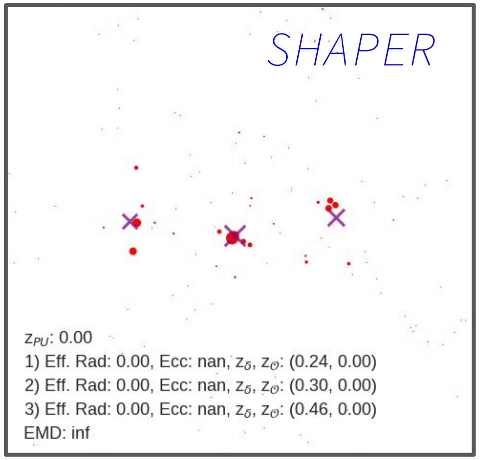
- Actual perturbative calculations
  - I claimed shapes are all *in principle* calculable — how about in practice?
  - Is there a generic way to resum this class of observables?
  - Some shapes (e.g thrust, *N*-subjettiness) are "easier" and faster to compute than generic shapes — why?
- Conjecture: Positivity is Perturbativity
  - Why doesn't a similar structure seem to work for charge, flavor, etc? The culprit seems to be that unlike energy, most quantum numbers aren't positive semidefinite!
- Gradients
  - Potentially useful for experimentalists – error propagation and unfolding for observables?
  - Can be used to define pileup susceptibility?
- Which shapes?
  - I've defined *every* shape – which ones should we actually use?
  - Annulusness for charm quark dead cones? Endcappiness for proton remnants? So many!

Conclusion

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





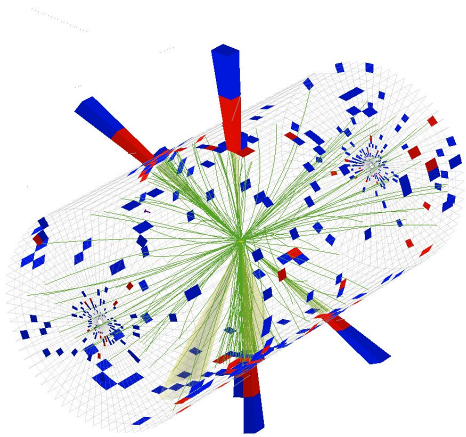


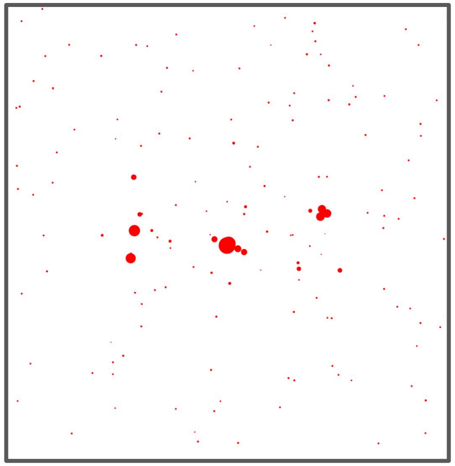
Collider Data → **Energy Flow** → **Shapes**

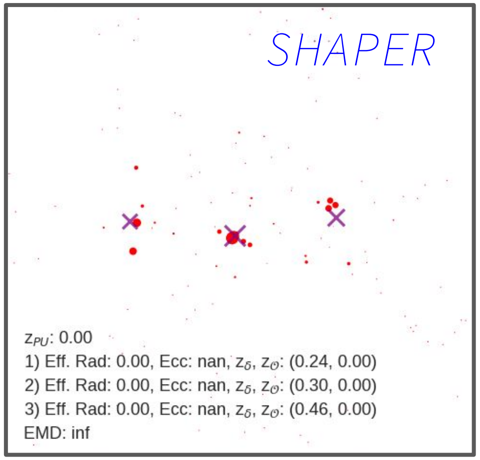
Download with *pip install pyshaper*

Conclusion

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





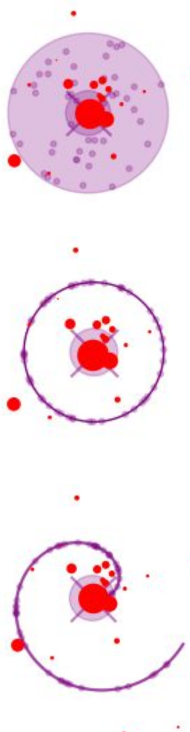


Collider Data → **Energy Flow** → **Shapes**

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

Appendices

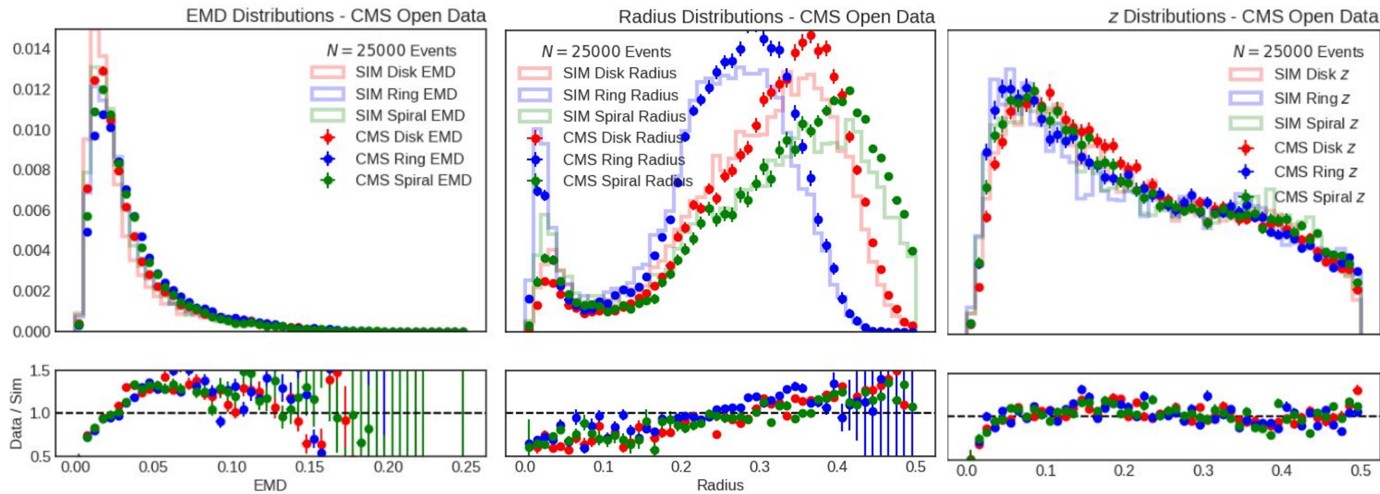
Observables on CMS OpenData



Disk + δ-function

Ring + δ-function

Spiral + δ-function



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

Building SHAPER

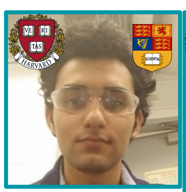
Key Component: The Loss function! Step 3: Synthesis

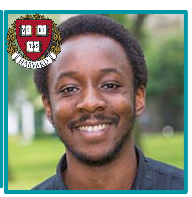
$$\mathcal{L}_R(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geqslant 0} \left[\sum_{i=1}^{M} \sum_{j=1}^{M'} \pi_{ij} \frac{|x_i - x'_j|}{R} \right] + \left| \sum_{i=1}^{M} z_i - \sum_{j=1}^{M'} z'_j \right|,$$

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

Dogra

Ba





K-Deep Simplices, Dictionary Learning, & Manifold Learning

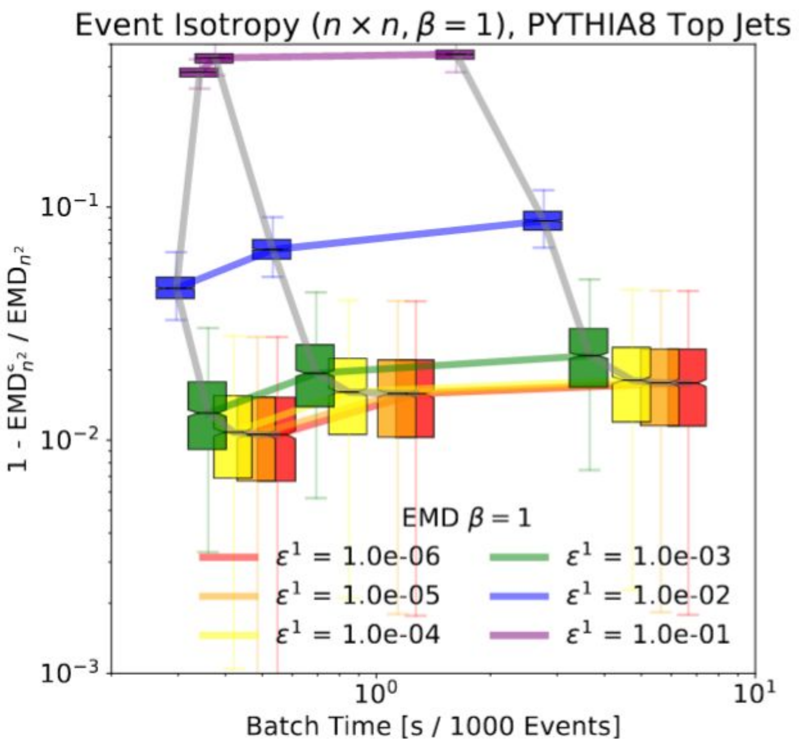
+

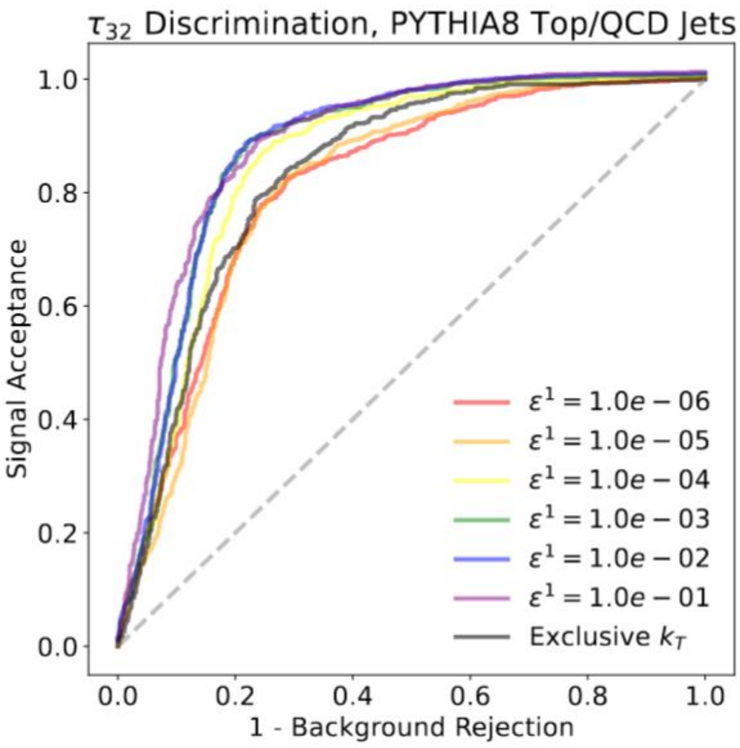
IRC Safety, Unclustered Radiation, & Wasserstein Geometry

Gambhir

Thaler

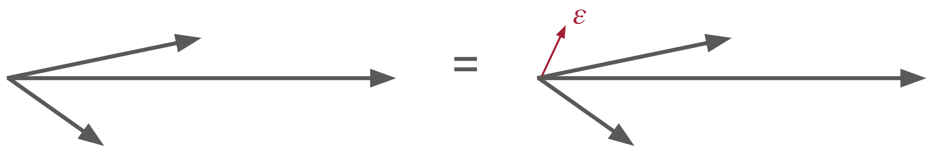
Performance Benchmarks



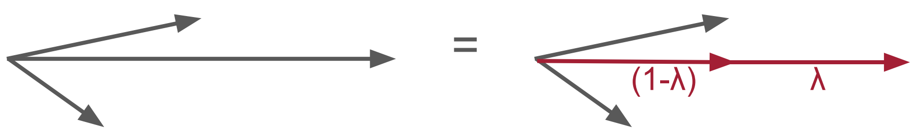


IRC Safety

Infrared Safety: An observable is unchanged under a soft emission



Collinear Safety: An observable is unchanged under a collinear splitting



[P. Tankala, A. Tasissa, J. M. Murphy, D. Ba, 2012.02134;
see also F.Dornaika, L.Weng, DOI: 0.1007/s13042-019-01035-z;
see also S. Roweis and L.Saul, DOI: 10.1126/science.290.5500.2323]

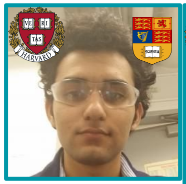
Building SHAPER

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

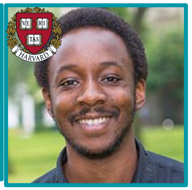
$$\mathcal{L}_R(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geqslant 0} \left[\sum_{i=1}^{M} \sum_{j=1}^{M'} \pi_{ij} \frac{|x_i - x'_j|}{R} \right],$$

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

Dogra



Ba



K-Deep Simplices, Dictionary Learning, & Manifold Learning

[P. Komiske, E. Metodiev, J. Thaler, 1902.02346;
see also T. Cai, J. Cheng, K. Craig, N. Craig, 2111.03670;
see also C. Zhang, Y. Cai, G. Lin, C. Shen, 2003.06777;
see also L. Hou, C. Yu, D. Samaras, 1611.05916;
see also M. Arjovsky, S. Chintala, L. Bottou, 1701.07875]

Building SHAPER

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

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

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

Dogra Ba

K-Deep Simplices,
Dictionary Learning, &
Manifold Learning

IRC Safety,
Unclustered Radiation, &
Wasserstein Geometry

Gambhir Thaler

Observables and Wasserstein

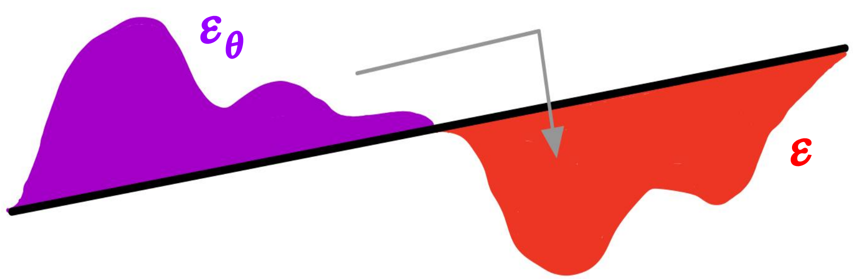
It can be shown that *any* observable on events, that* …

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

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

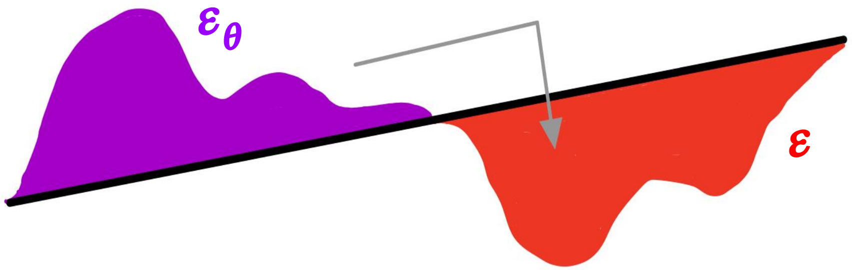
$$\theta = \underset{\mathcal{E}'_\theta \in \mathcal{M}}{\text{argmin}}\, \text{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$



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

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

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

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

… c

Was

Dist

mar

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

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

Ask me for more details on this after!