

The Hidden Geometry of Particle Collisions

Jesse Thaler



Joint Theory Seminar, UC Davis — November 16, 2020

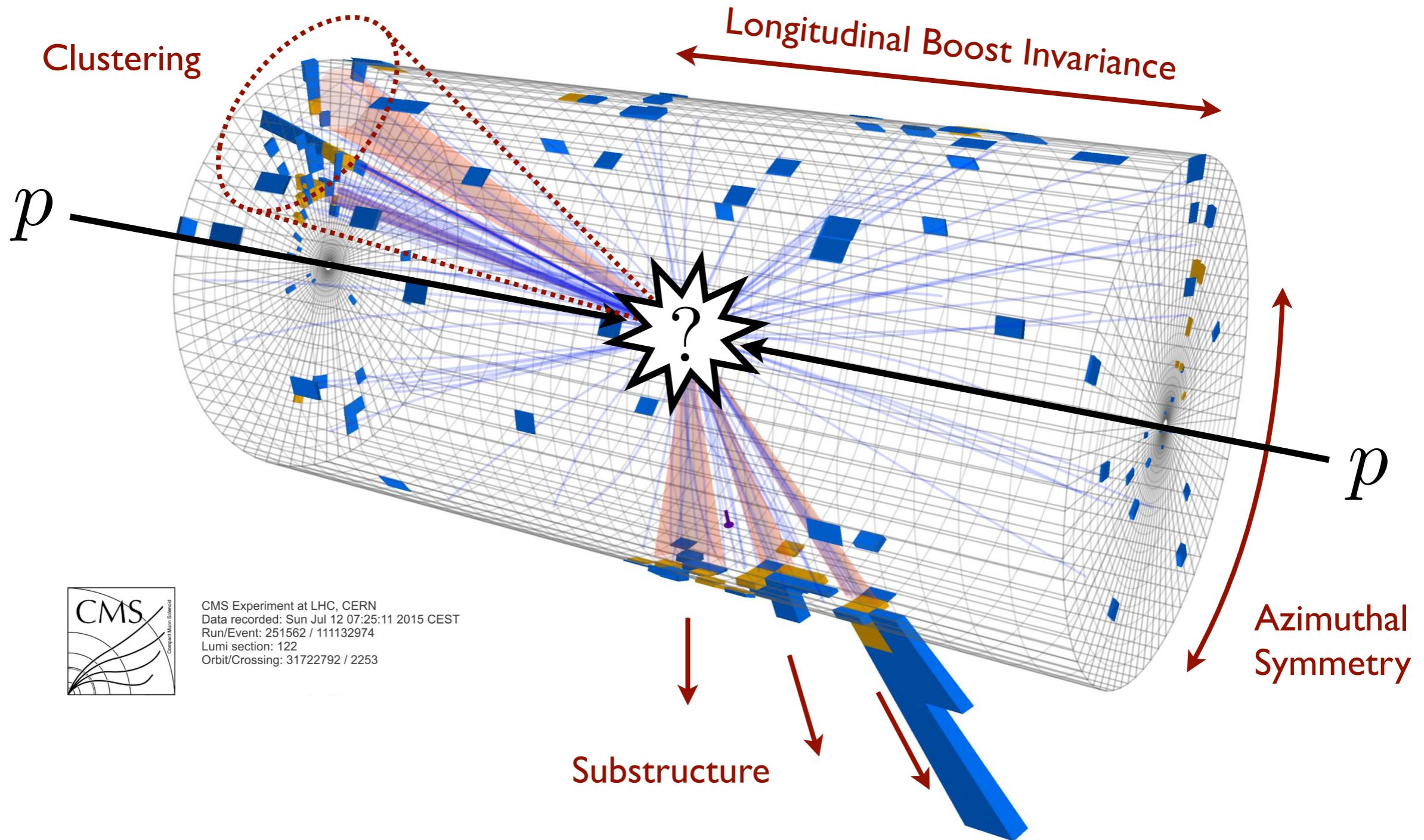
Wearing my New Hat...

The NSF AI Institute for
Artificial Intelligence and Fundamental Interactions



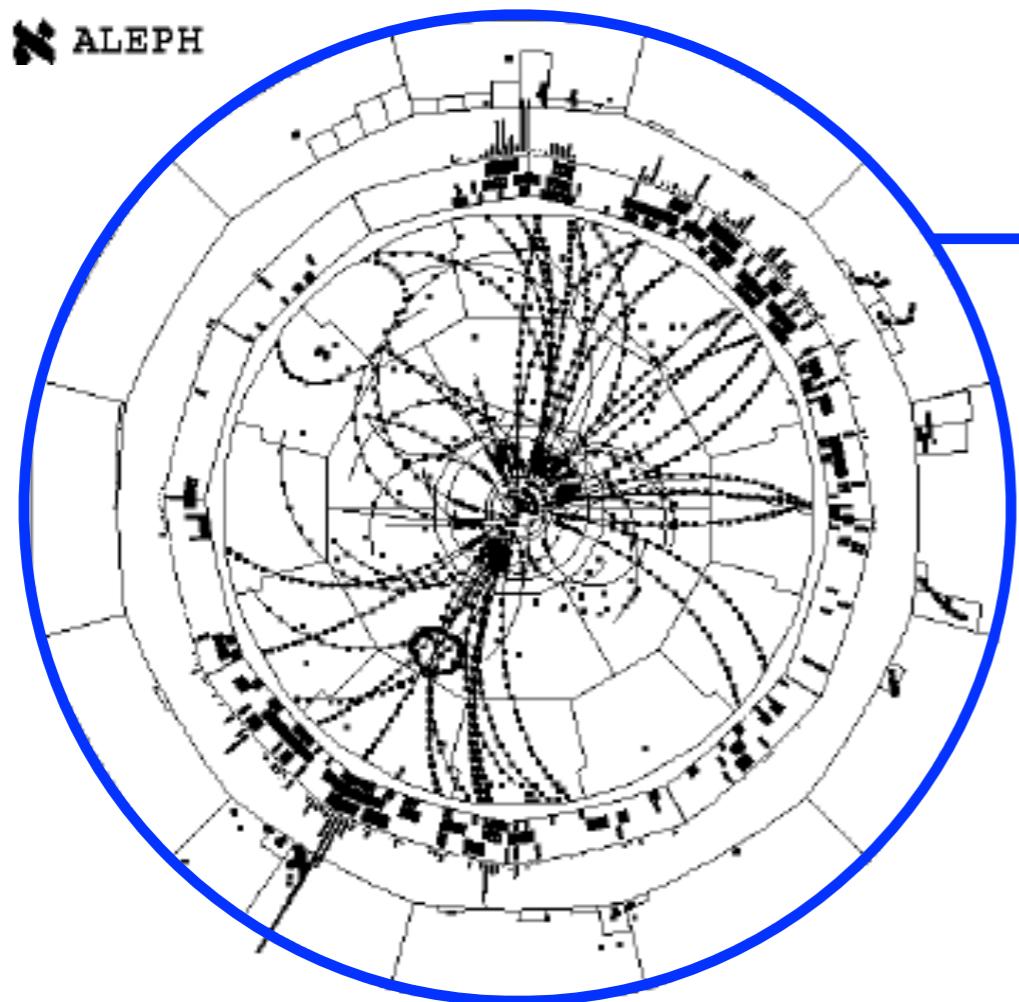
<http://iaifi.org/>

The Manifest Geometry of One Collision

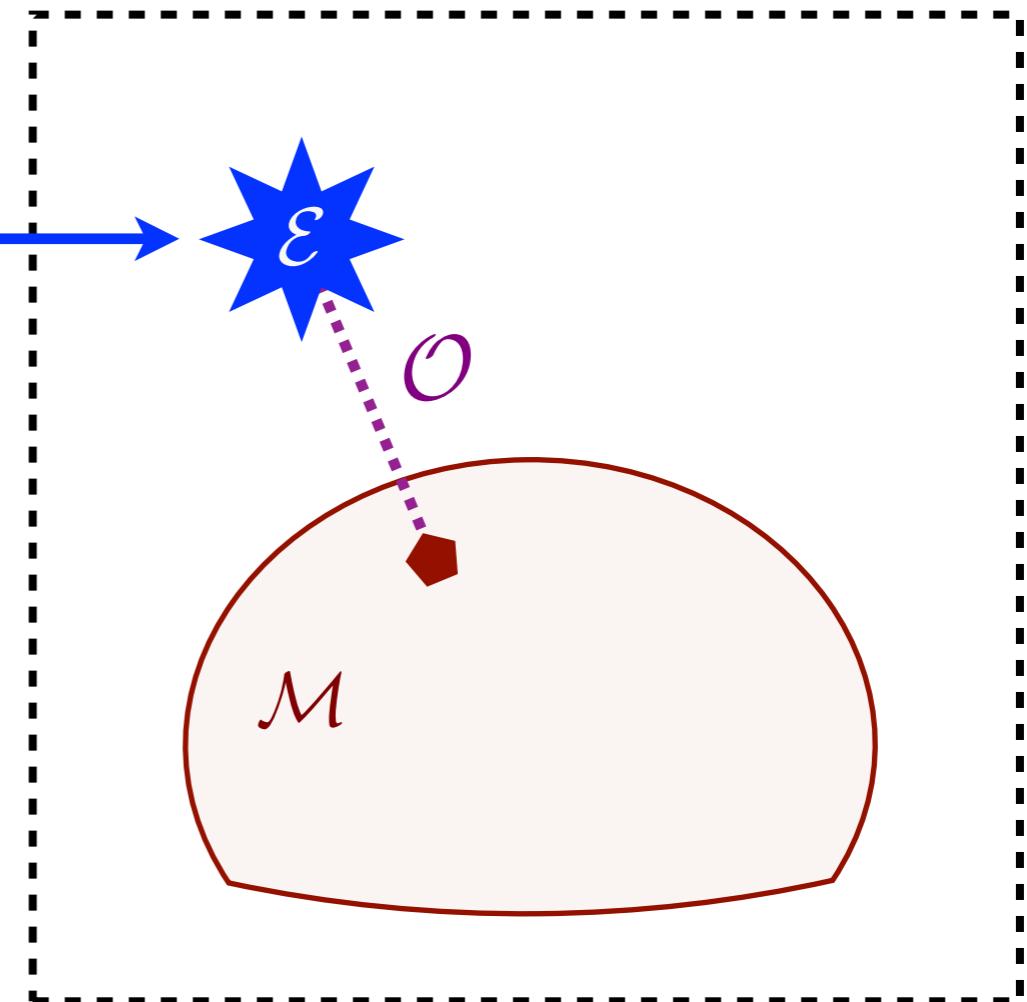


The Hidden Geometry of Particle Collisions

E.g. Classic QCD Event Shapes



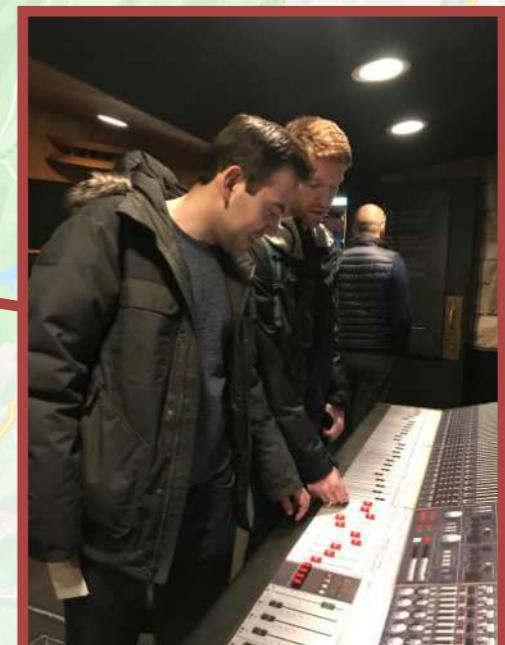
One Electron-Positron Event



Distance to a Manifold in Event Space

[Komiske, Metodiev, JDT, [JHEP 2020](#)
[Brandt, Peyrou, Sosnowski, Wroblewski, [PL 1964](#); Farhi, [PRL 1977](#)]

Scenes from My Sabbatical



Eric Metodiev

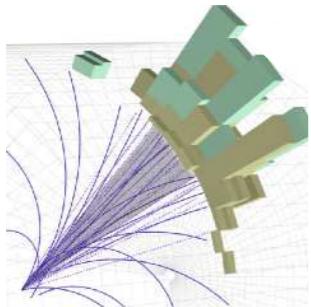


Patrick Komiske

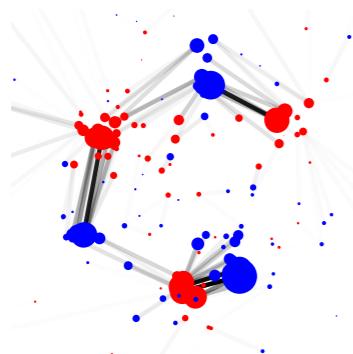
[February 2019;
Simons Sabbatical Fellowship]

SF

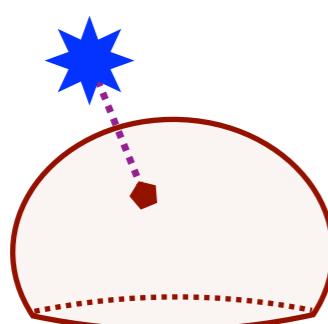
Outline



What is a Collider Event?



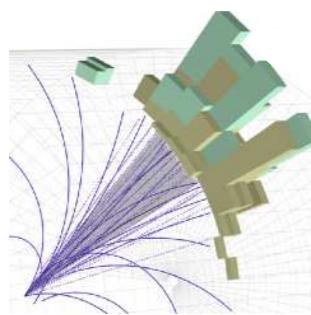
When are Events Similar?



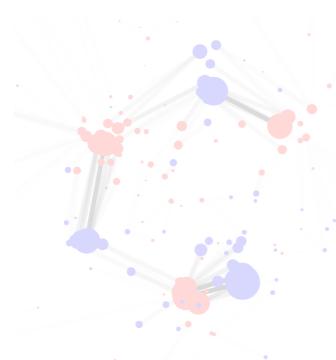
What can be Geometrized?

Pause

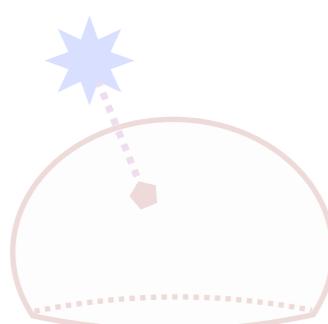
Interrupt me or drop questions in the chat, and I'll try to answer them as I go



What is a Collider Event?



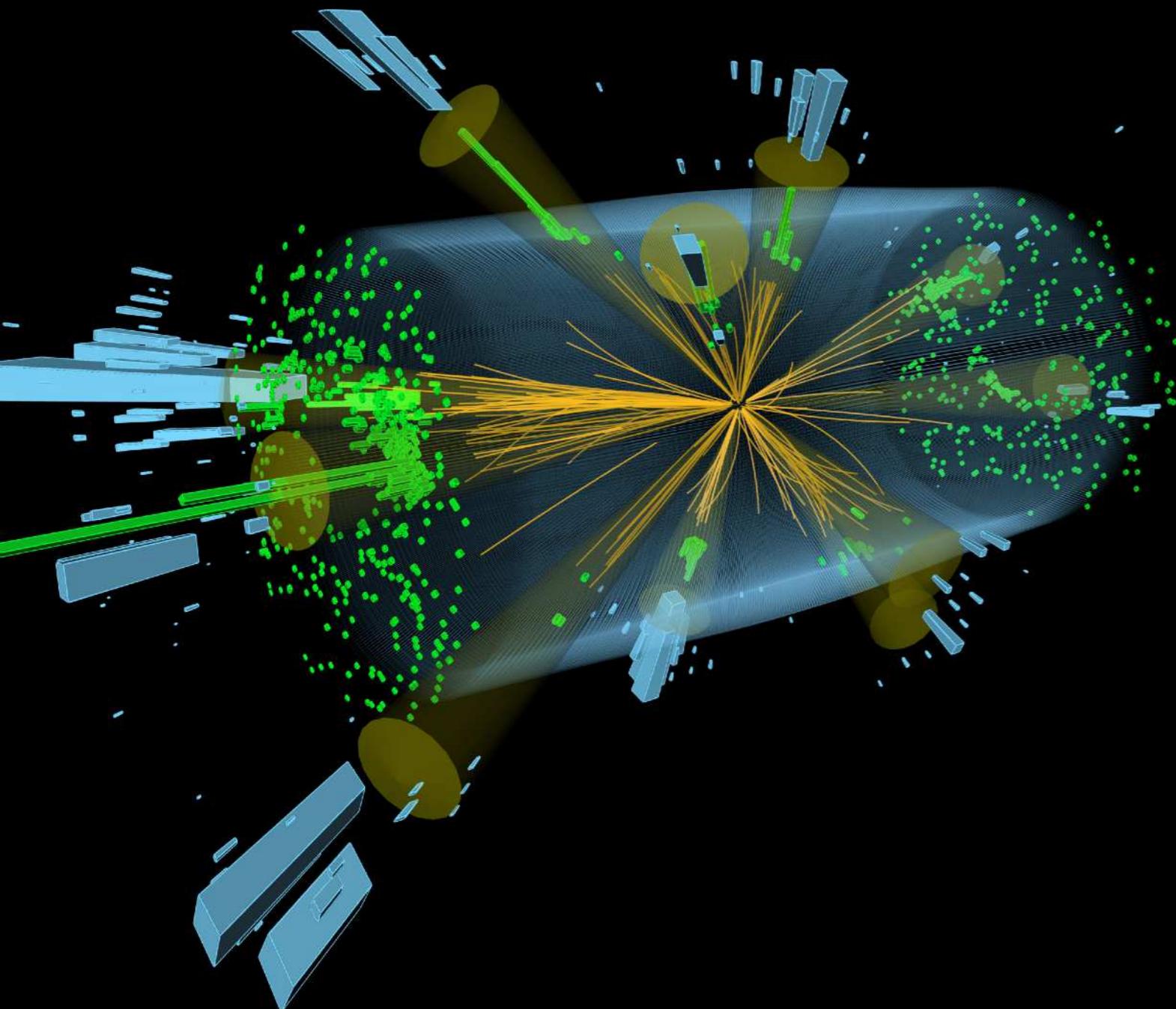
When are Events Similar?



What can be Geometrized?

Collider Event

Collection of points in (momentum) space



T E H M

 γ

photon

 e^+

electron

 μ^+

muon

 π^+

pion

 K^+

kaon

 K_L^0

K-long

 p/\bar{p}

proton

 n/\bar{n}

neutron

elementary

composite

Point Cloud

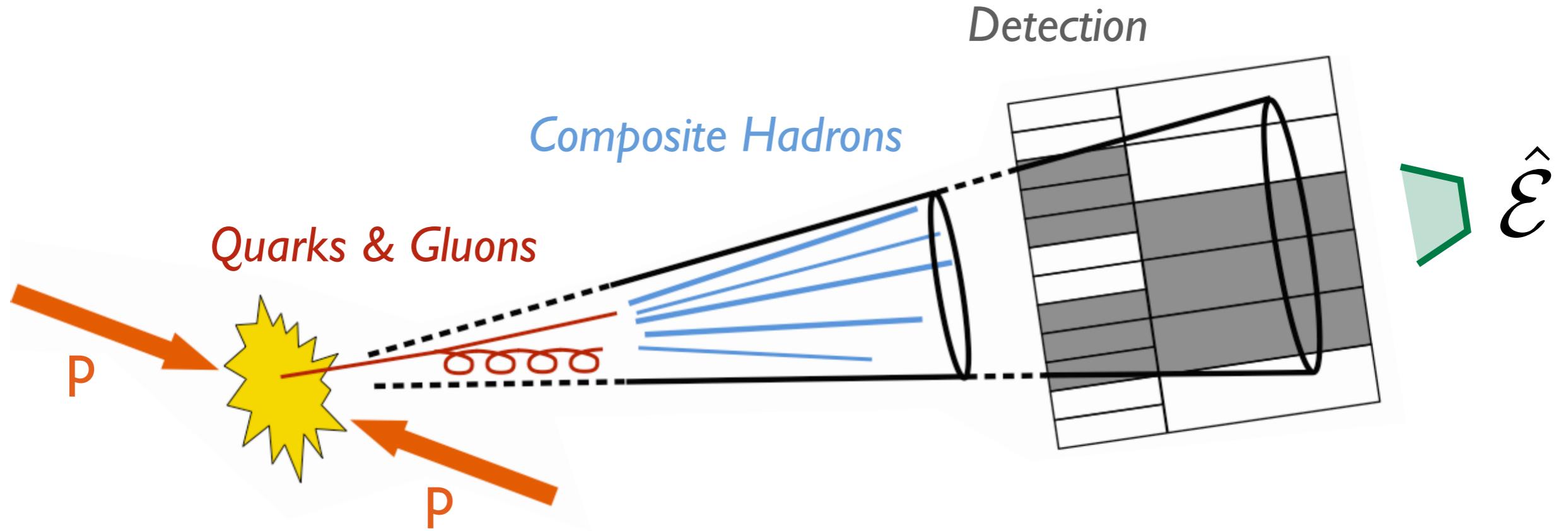
Collection of points in (position) space



[Popular Science, 2013]

Jet Formation from QCD

Theory



Energy Flow:

Robust to hadronization and detector effects
Well-defined for massless gauge theories

$$\hat{\mathcal{E}} \simeq \lim_{t \rightarrow \infty} \hat{n}_i T^{0i}(t, vt\hat{n})$$

[see e.g. Sveshnikov, Tkachov, [PLB 1996](#); Hofman, Maldacena, [JHEP 2008](#); Mateu, Stewart, [JDT, PRD 2013](#); Belitsky, Hohenegger, Korchemsky, Sokatchev, Zhiboedov, [PRL 2014](#); Chen, Moult, Zhang, Zhu, [PRD 2020](#)]

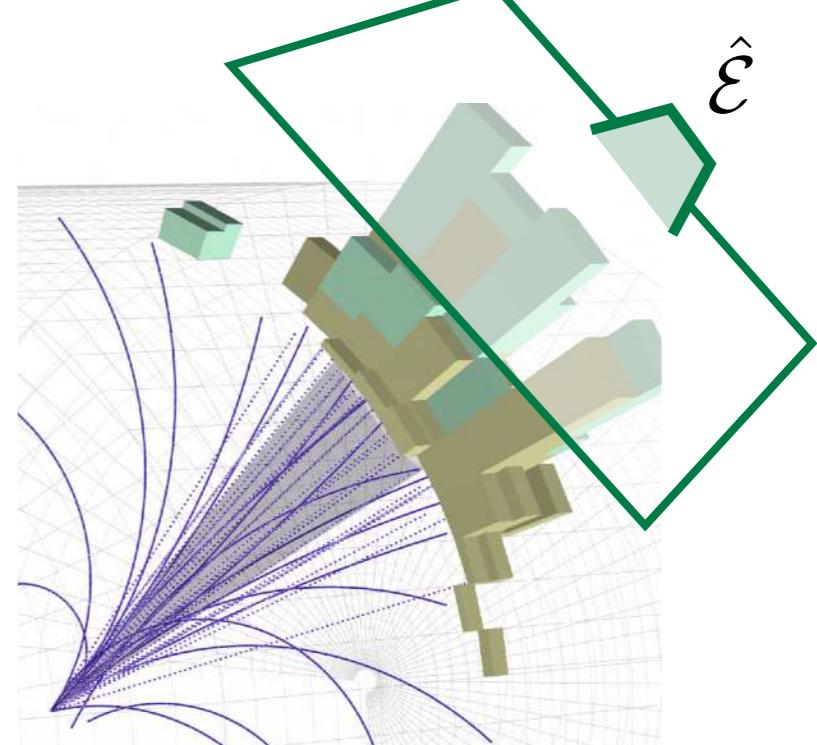
Jets as Weighted Point Clouds

- Energy-Weighted Directions

$$\vec{p} = \{E, \hat{n}_x, \hat{n}_y, \hat{n}_z\}$$

↑ |
Energy Direction

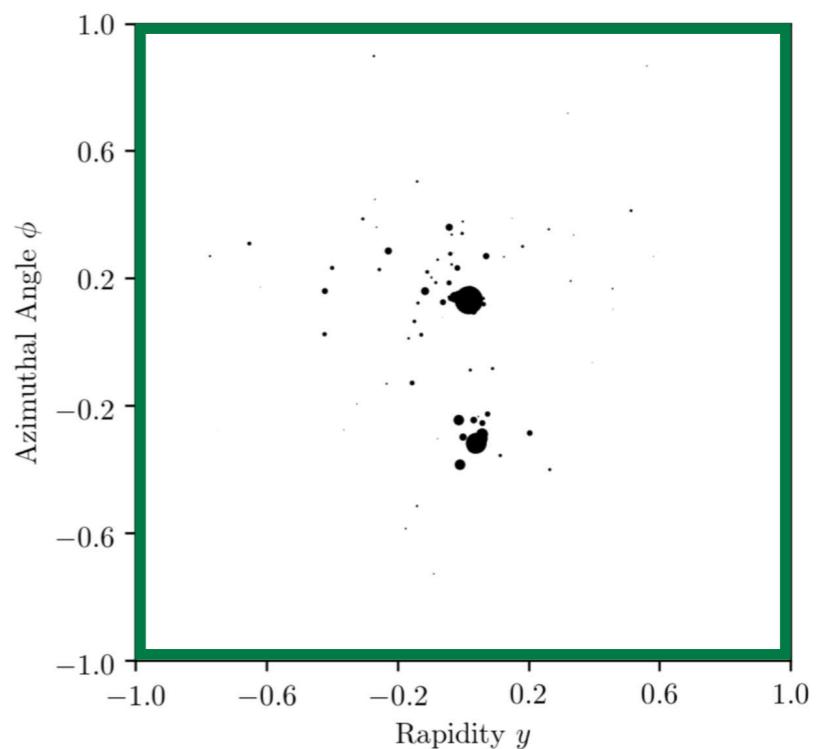
(suppressing “unsafe” charge/flavor information)



- Equivalently: Energy Density

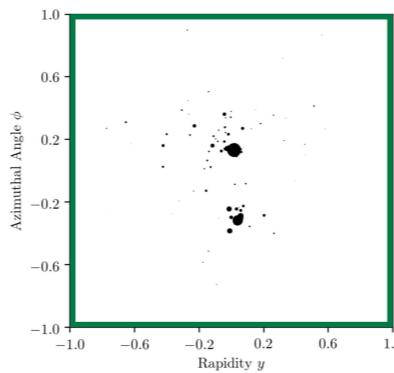
$$\rho(\hat{n}) = \sum_{i \in \mathcal{J}} E_i \delta^{(2)}(\hat{n} - \hat{n}_i)$$

↑ ↑
Energy Direction



Pause

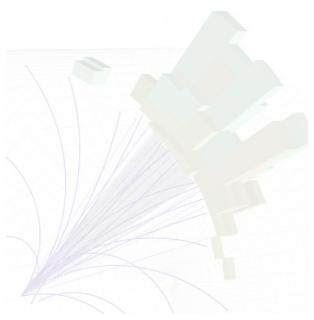
What is a Collider Event?



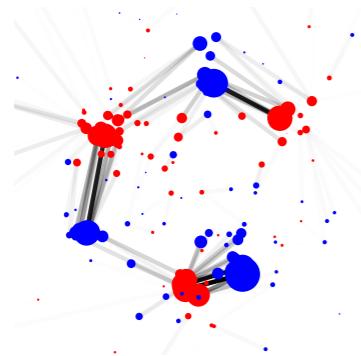
“Calo” Energy Density

$$\mathcal{E}(\hat{n}) = \sum_i E_i \delta(\hat{n} - \hat{n}_i)$$

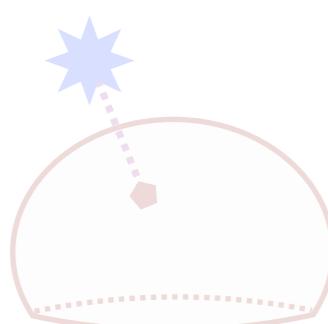
(see backup for relevance to ML and QCD)



What is a Collider Event?



When are Events Similar?

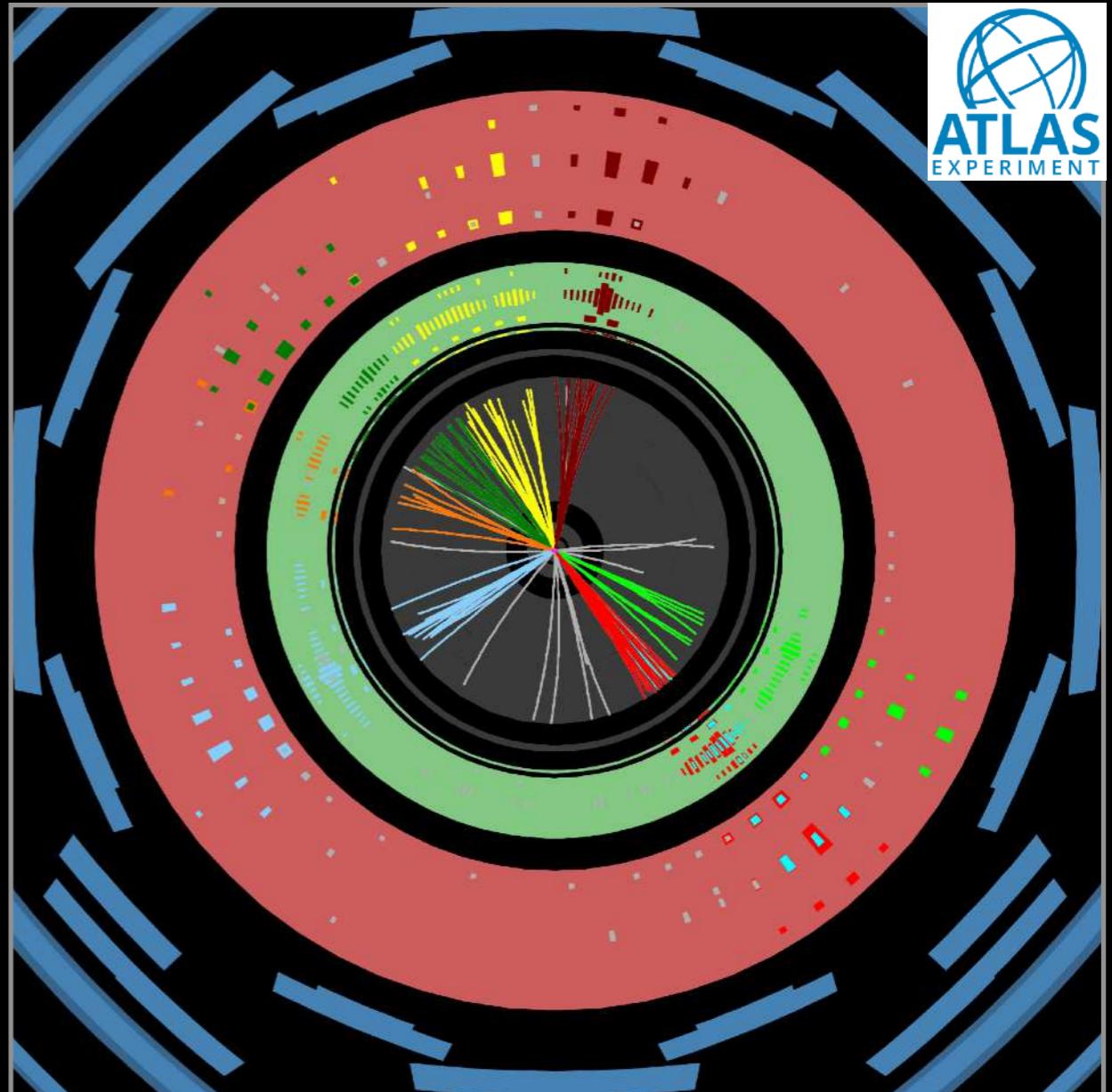
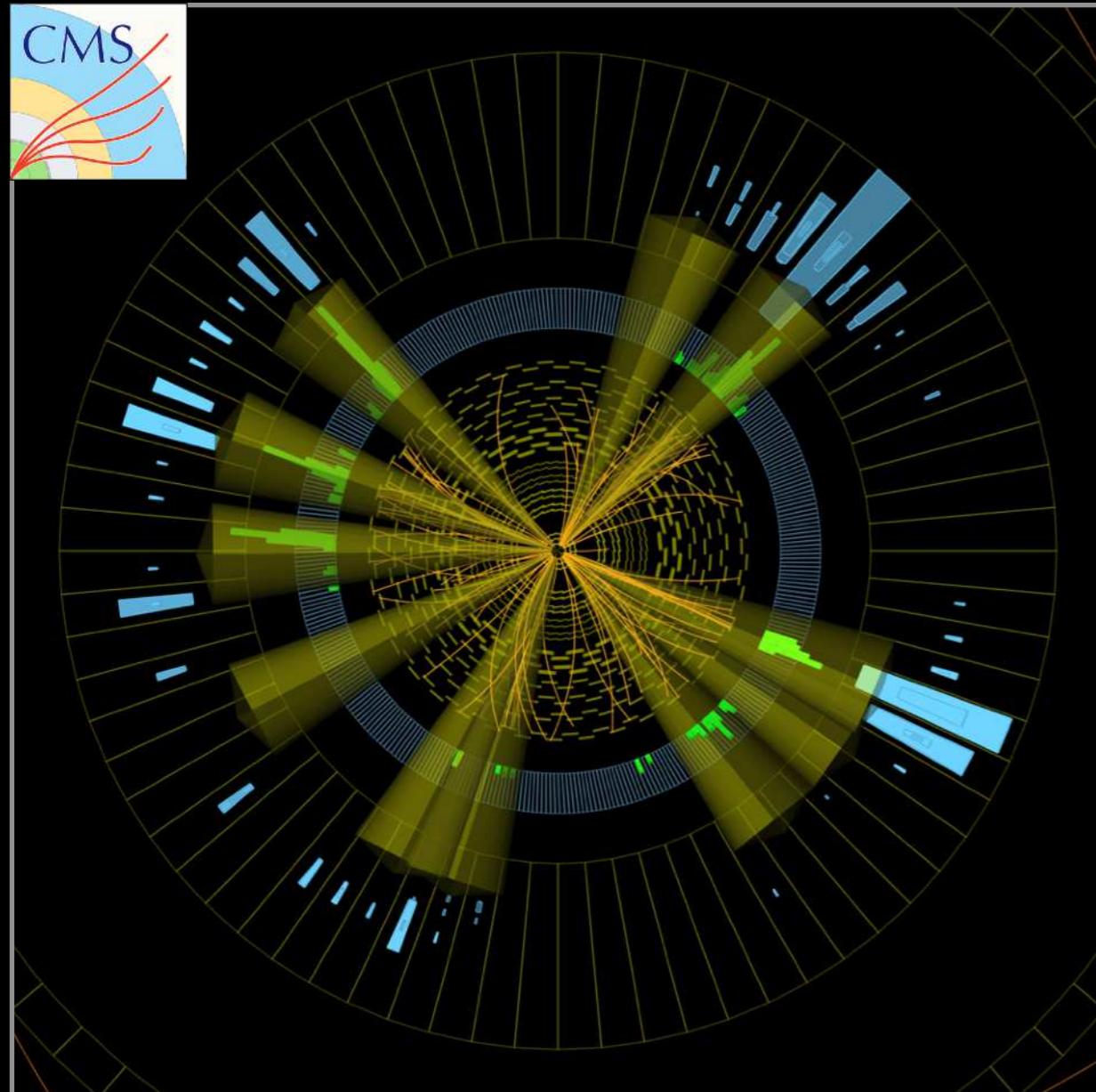


What can be Geometrized?

Two Collider Events

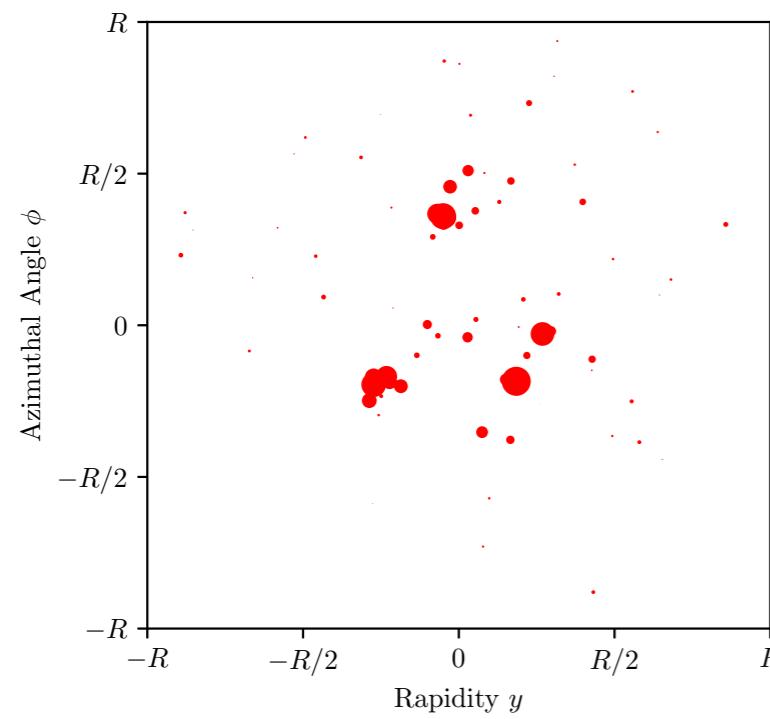
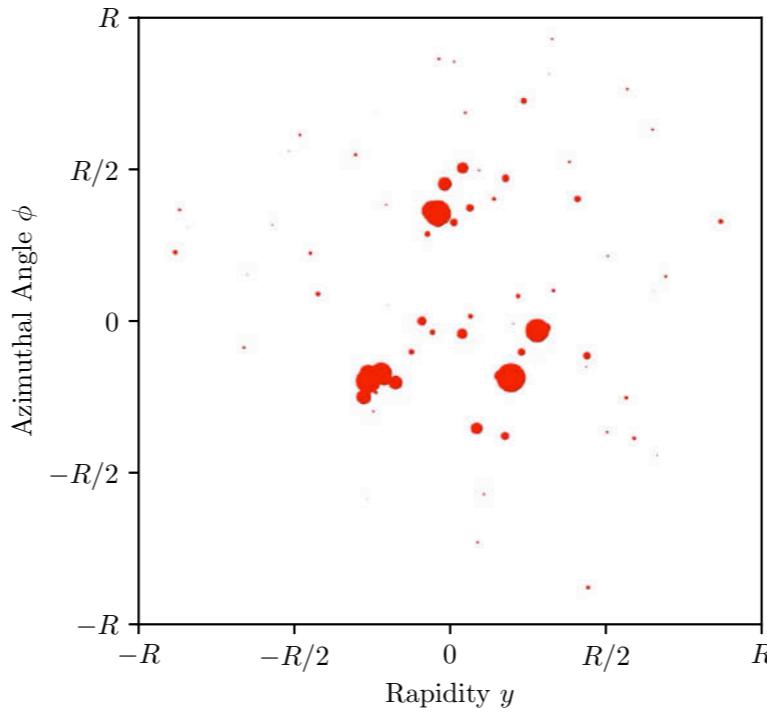
Two collections of points in (momentum) space

How “close” are these?

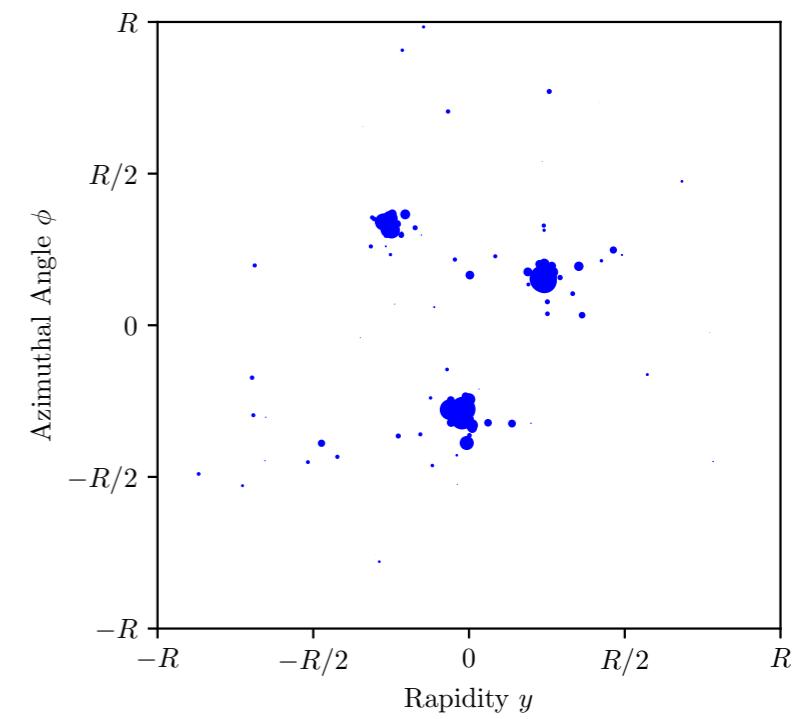


Similarity of Two Energy Flows?

$$\mathcal{E}(\hat{n}) = \sum_i E_i \delta(\hat{n} - \hat{n}_i)$$



Optimal Transport:
Earth Mover's Distance
a.k.a. l -Wasserstein metric



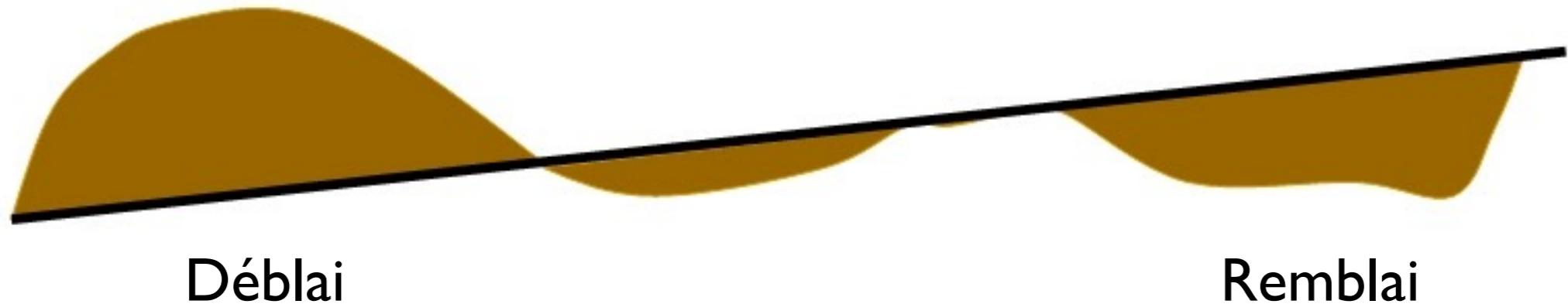
[Komiske, Metodiev, JDT, PRL 2019; code at Komiske, Metodiev, JDT, [energyflow.network](#)]

The Earth Mover's Distance

Optimal Transport:

[Peleg, Werman, Rom, [IEEE 1989](#);
Rubner, Tomasi, Guibas, [ICCV 1998](#), [ICCV 2000](#);
Pele, Werman, [ECCV 2008](#); Pele Taskar, [GSI 2013](#)]

Minimum “work” (stuff \times distance) to make one distribution look like another distribution



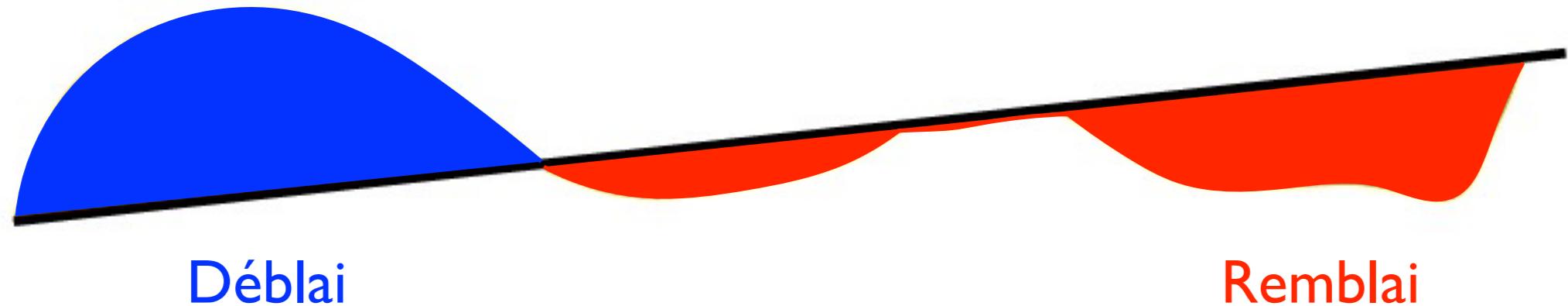
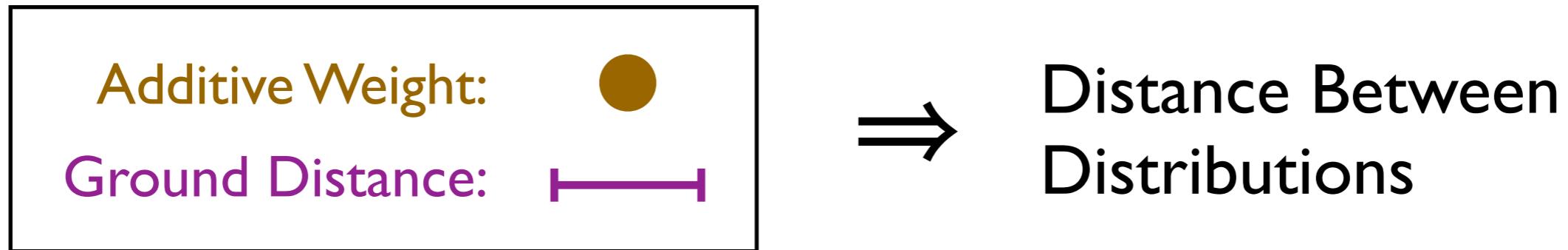
[h/t Niles-Weed, [ML4Jets 2020](#); Monge, 1781; Vaserštejn, 1969; [Wikipedia](#)]

The Earth Mover's Distance

Optimal Transport:

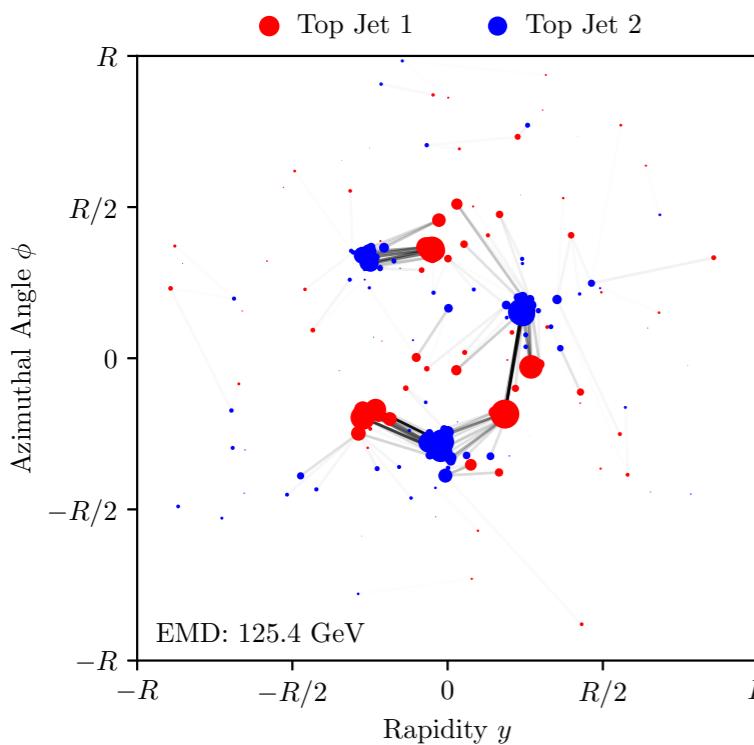
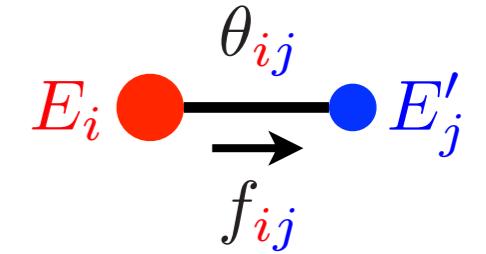
[Peleg, Werman, Rom, [IEEE 1989](#);
Rubner, Tomasi, Guibas, [ICCV 1998](#), [ICCV 2000](#);
Pele, Werman, [ECCV 2008](#); Pele Taskar, [GSI 2013](#)]

Minimum “work” (**stuff** × **distance**) to make
one distribution look like **another distribution**



[h/t Niles-Weed, [ML4Jets 2020](#); Monge, 1781; Vaserštejn, 1969; [Wikipedia](#)]

The Energy Mover's Distance

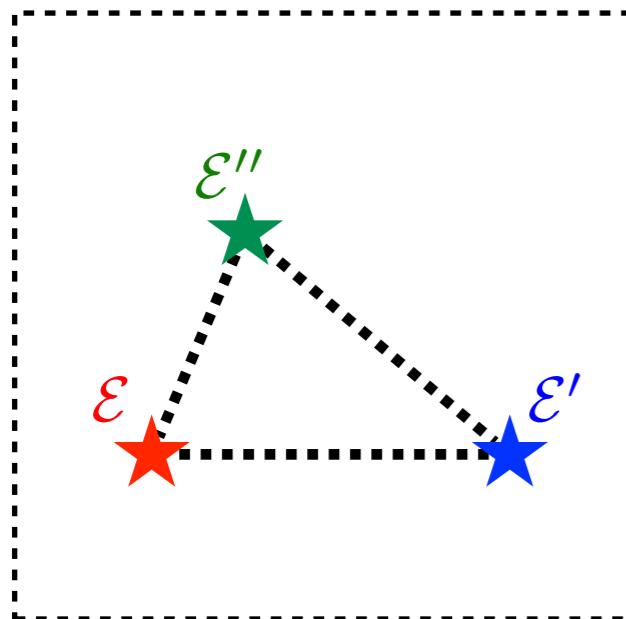


Optimal transport between energy flows...

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

↑
in GeV

Cost to move energy Cost to create energy



...defines a metric on the space of events

$$0 \leq \text{EMD}(\mathcal{E}, \mathcal{E}') \leq \text{EMD}(\mathcal{E}, \mathcal{E}'') + \text{EMD}(\mathcal{E}', \mathcal{E}'')$$

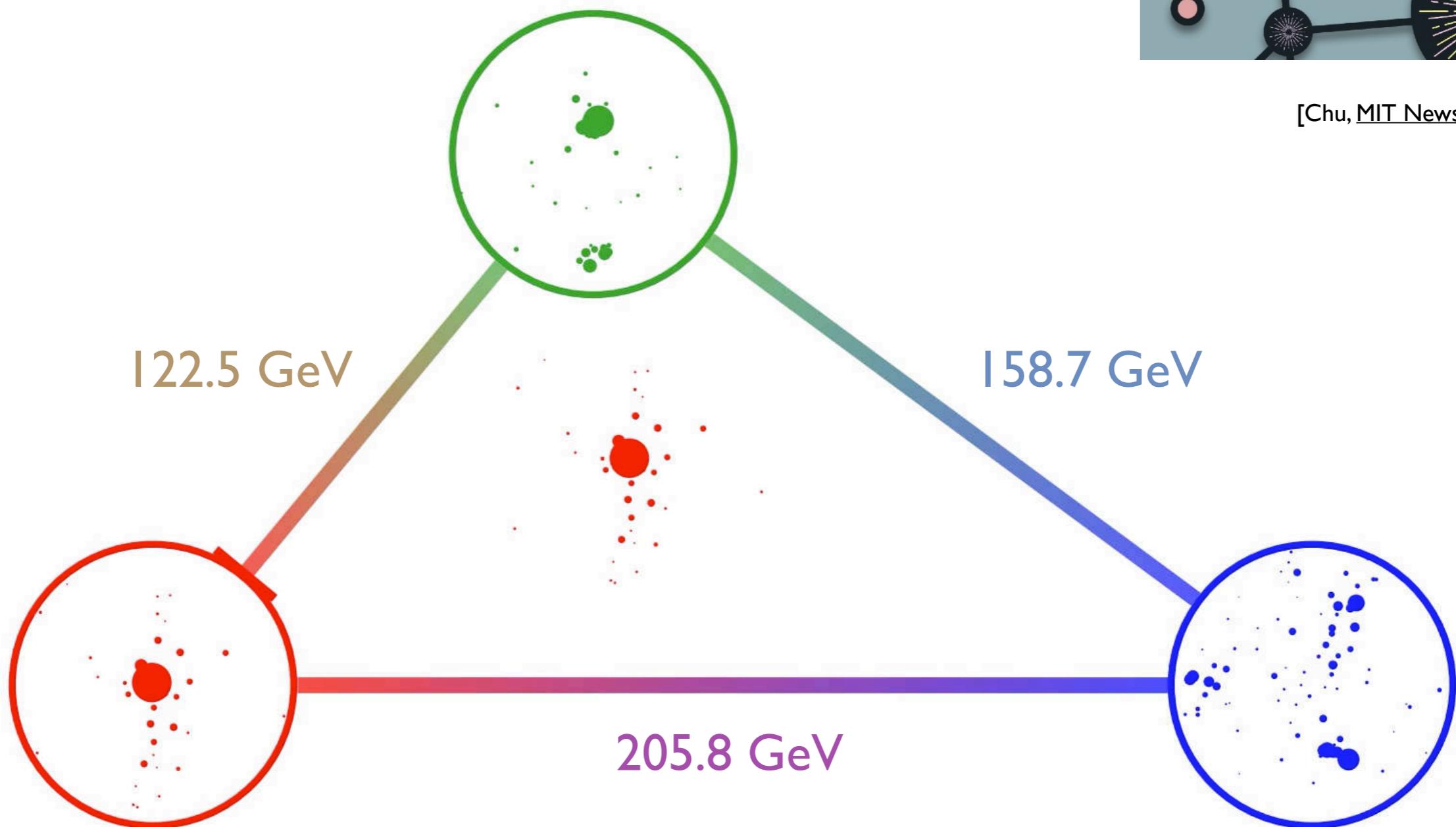
(assuming $R \geq \theta_{\max}/2$, i.e. $R \geq$ jet radius for conical jets)

[Komiske, Metodiev, JDT, PRL 2019;
 see also Pele, Werman, ECCV 2008; Pele, Taskar, GSI 2013;
 [see flavored variant in Crispim Romão, Castro, Milhano, Pedro, Vale, arXiv 2020]
 [see computational speed up in Cai, Cheng, Craig, Craig, arXiv 2020]

Similarity of Three Energy Flows?



[Chu, MIT News July 2019]



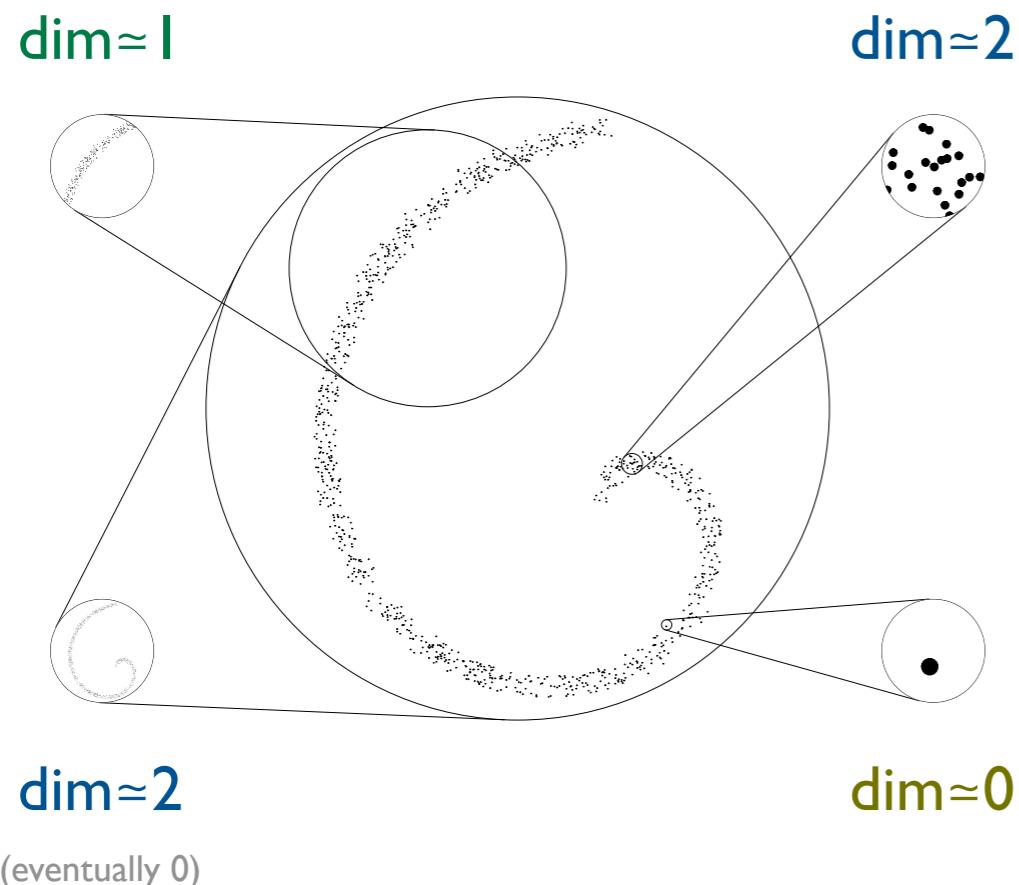
[Komiske, Metodiev, JDT, PRL 2019; code at Komiske, Metodiev, JDT, [energyflow.network](#); see alternative graph network approach in Mullin, Pacey, Parker, White, Williams, arXiv 2019]

Dimensionality of Space of Jets

$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

$$\Rightarrow \dim(r) \sim r \frac{\partial}{\partial r} \ln N_{\text{neighbors}}(r)$$

[Grassberger, Procaccia, [PRL 1983](#); Kégl, [NIPS 2002](#)]



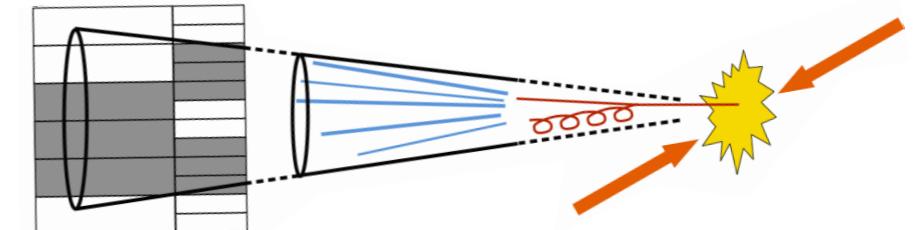
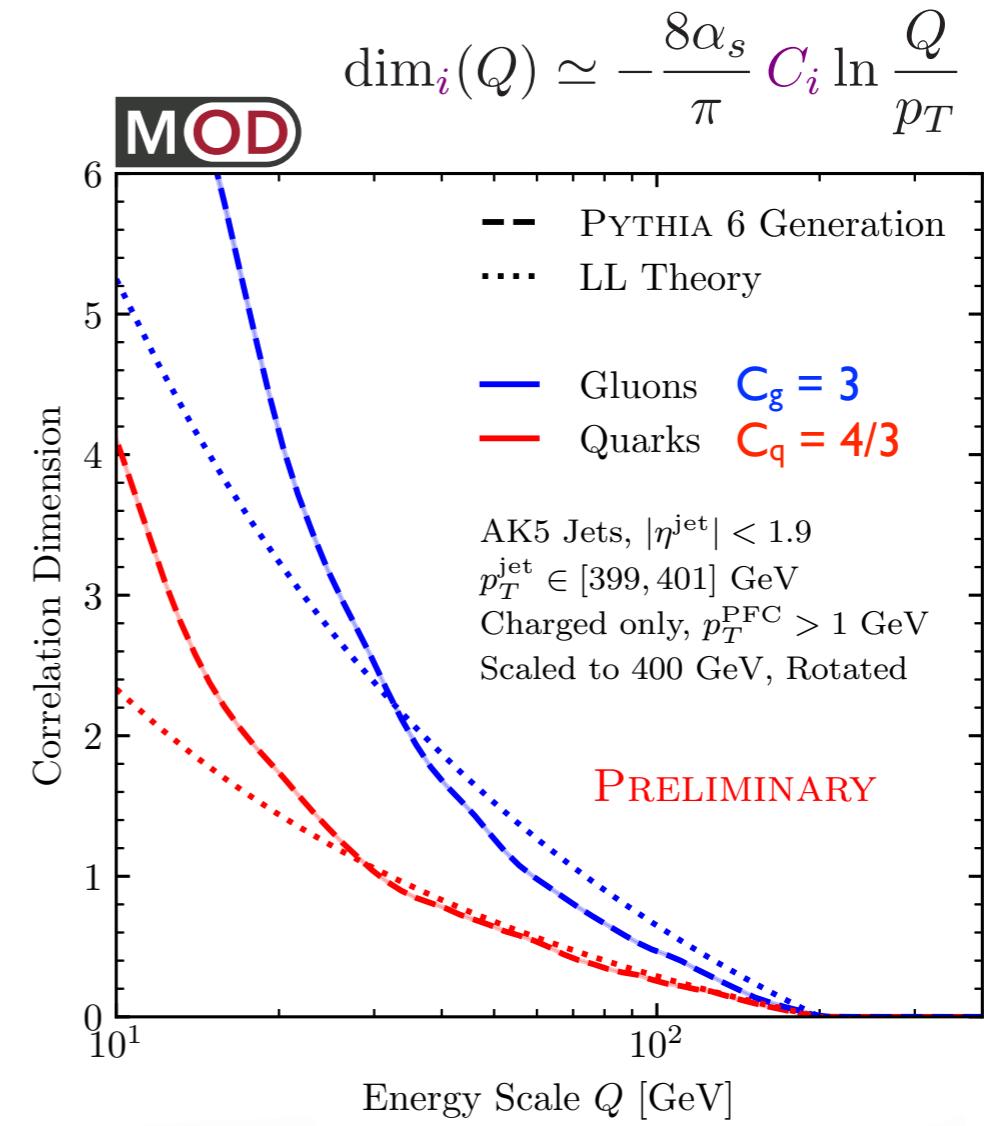
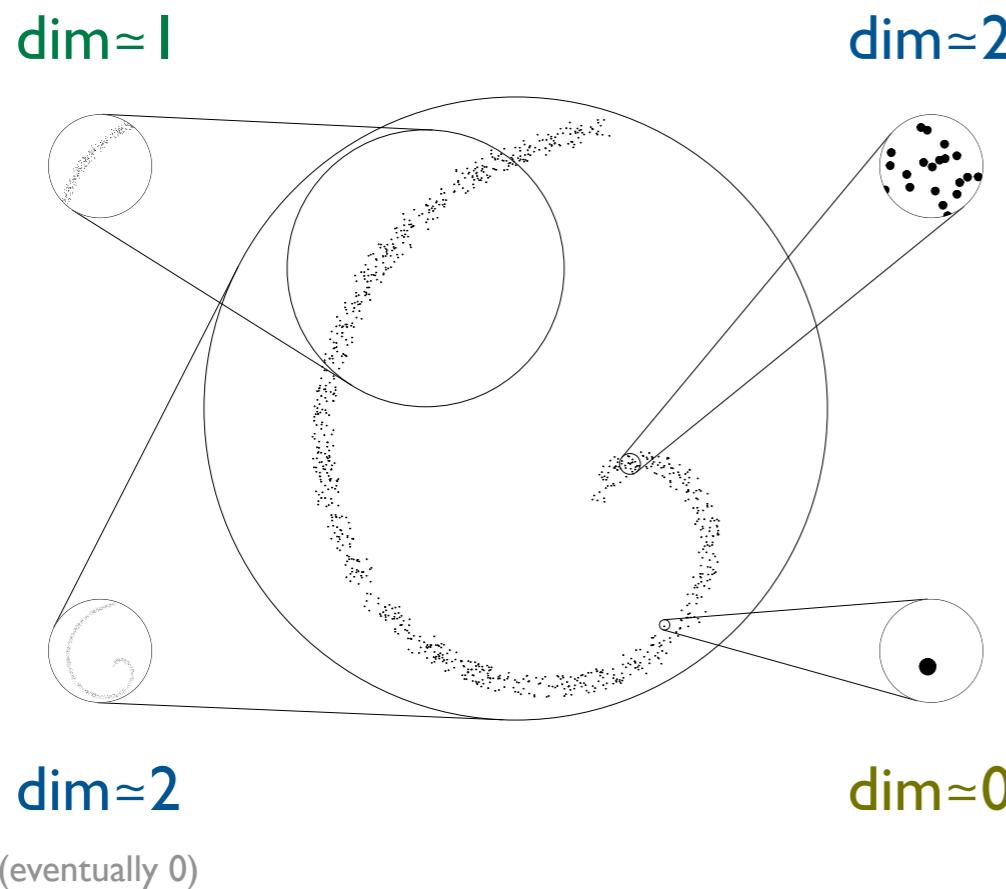
Dimensionality of Space of Jets



$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

$$\Rightarrow \dim(r) \sim r \frac{\partial}{\partial r} \ln N_{\text{neighbors}}(r)$$

[Grassberger, Procaccia, [PRL 1983](#); Kégl, [NIPS 2002](#)]



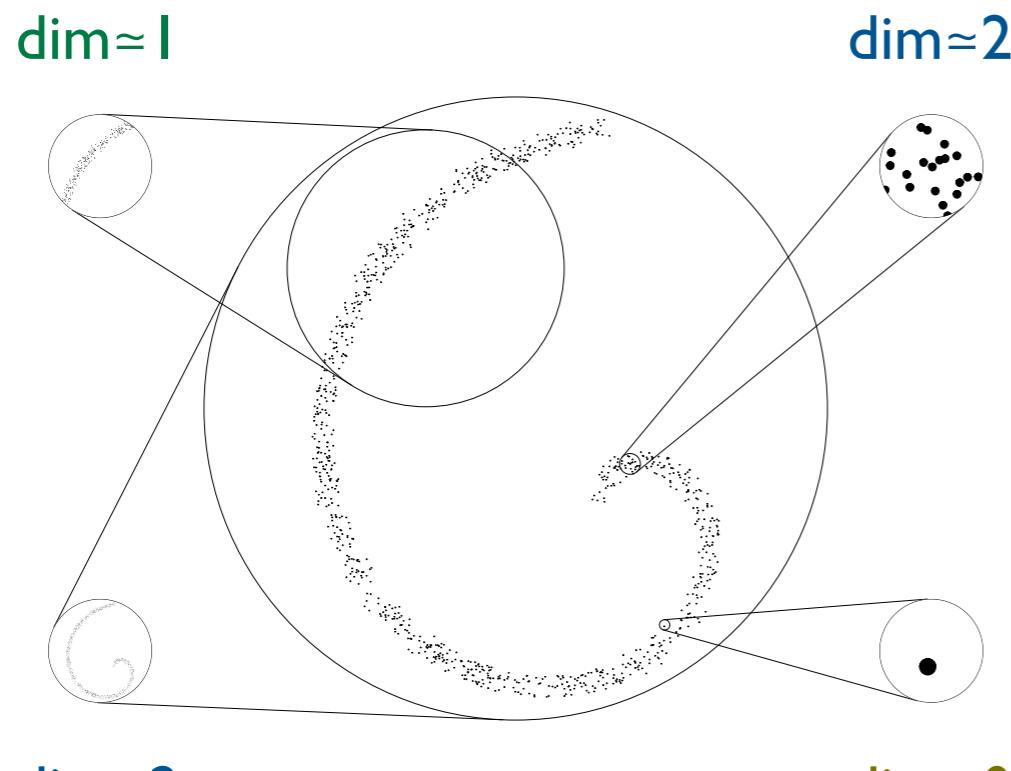
Dimensionality of Space of Jets



$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

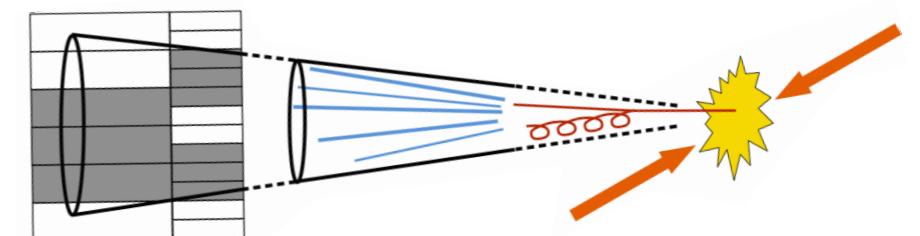
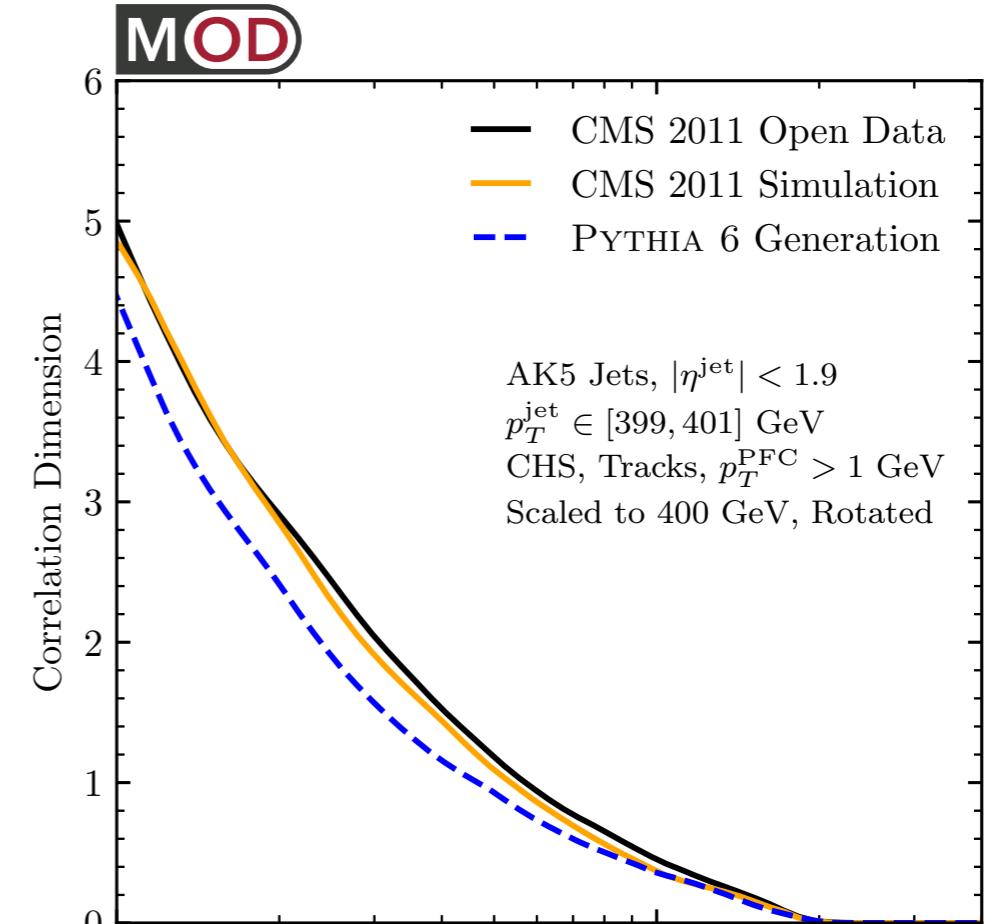
$$\Rightarrow \dim(r) \sim r \frac{\partial}{\partial r} \ln N_{\text{neighbors}}(r)$$

[Grassberger, Procaccia, [PRL 1983](#); Kégl, [NIPS 2002](#)]



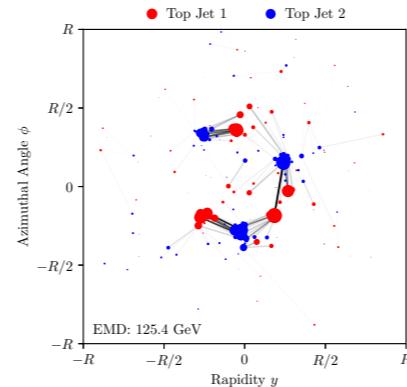
(eventually 0)

[Komiske, Mastandrea, Metodiev, Naik, [JDT, PRD 2020](#);
using [CMS Open Data](#)]



Pause

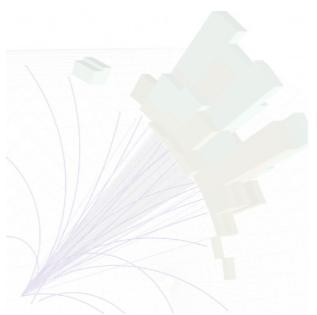
When are Events Similar?



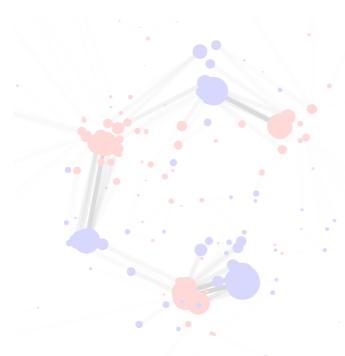
Small Energy Mover's Distance

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

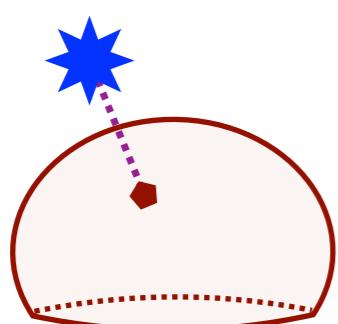
(see backup for more applications)



What is a Collider Event?



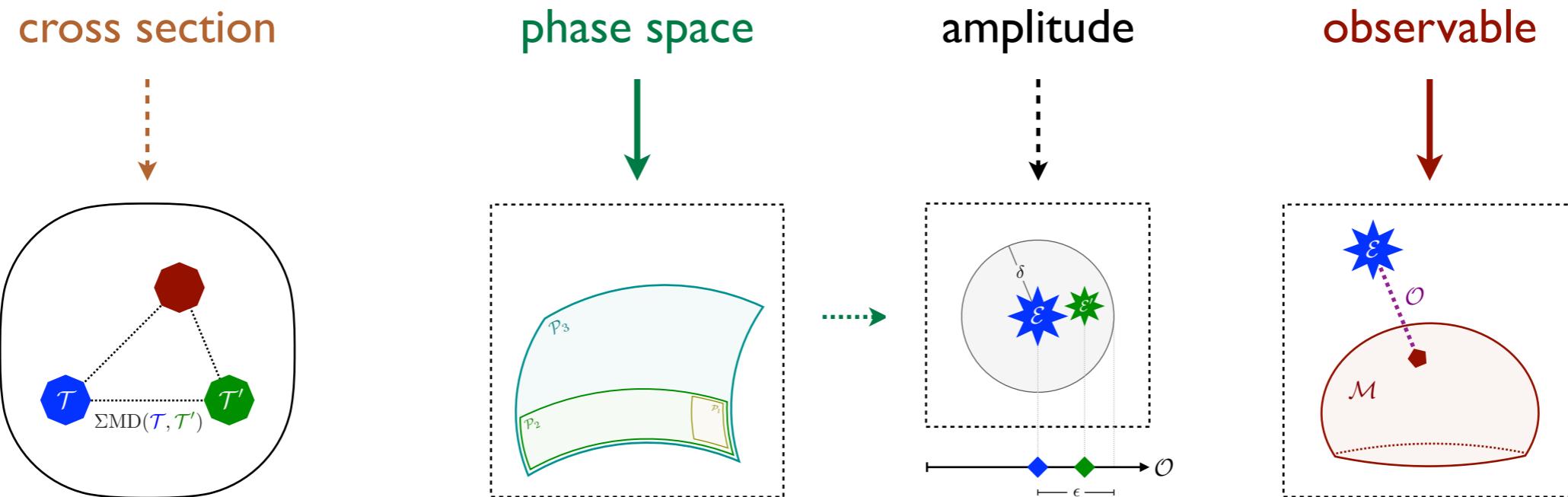
When are Events Similar?



What can be Geometrized?

Master Formula for Collider Physics

$$\sigma_{\text{obs}} \simeq \frac{1}{2E_{\text{CM}}^2} \sum_{n=2}^{\infty} \int d\Phi_n |\mathcal{M}_{AB \rightarrow 12\dots n}|^2 f_{\text{obs}}(\Phi_n)$$

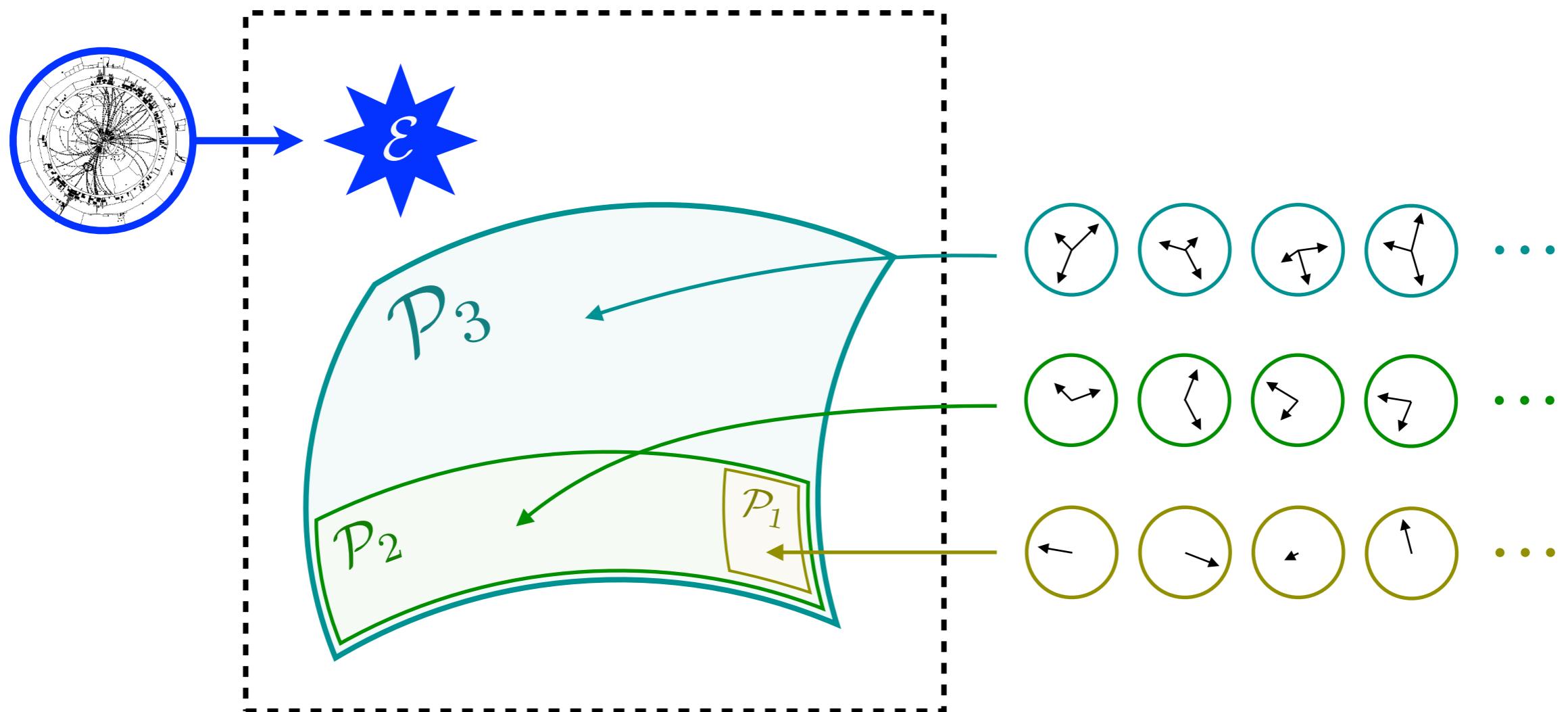


[Komiske, Metodiev, JDT, [JHEP 2020](#)]

Introducing N-particle Manifolds

$$\sum_{n=2}^{\infty} \int d\Phi_n$$

\mathcal{P}_N = set of all N-particle configurations

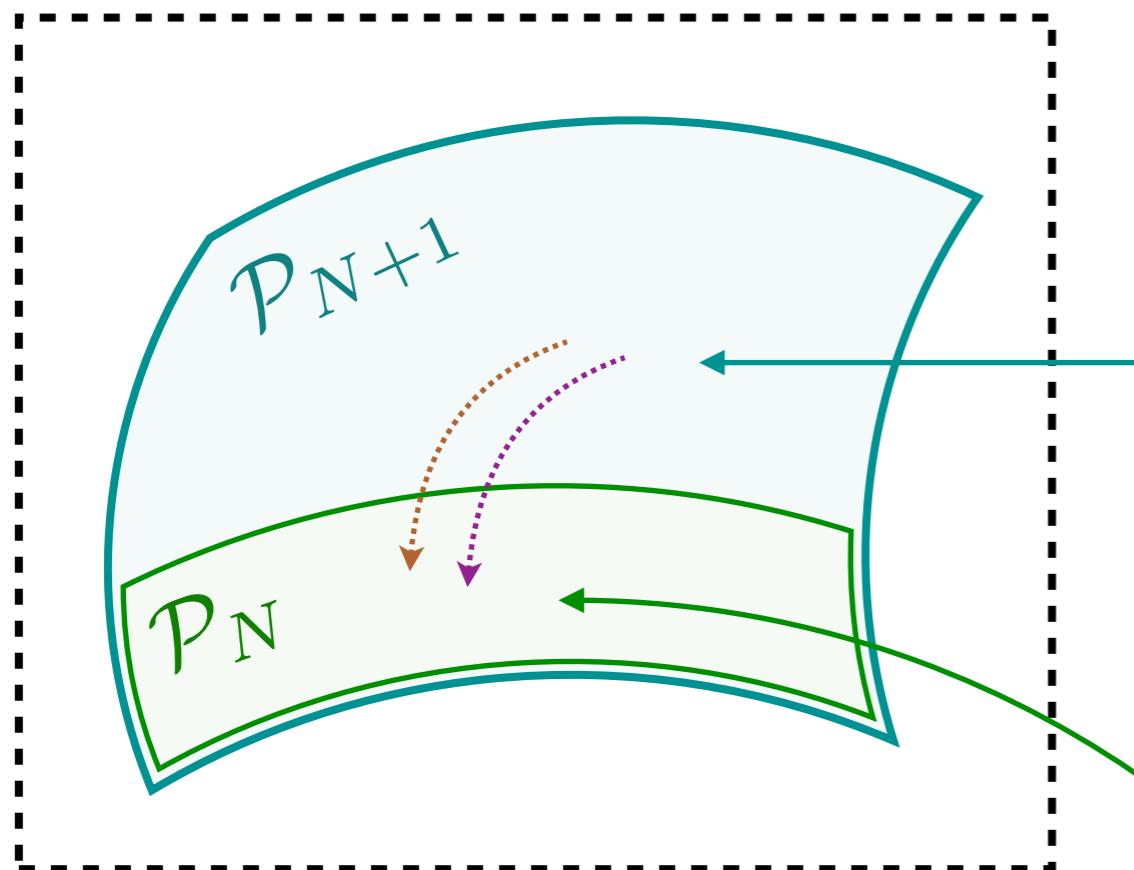
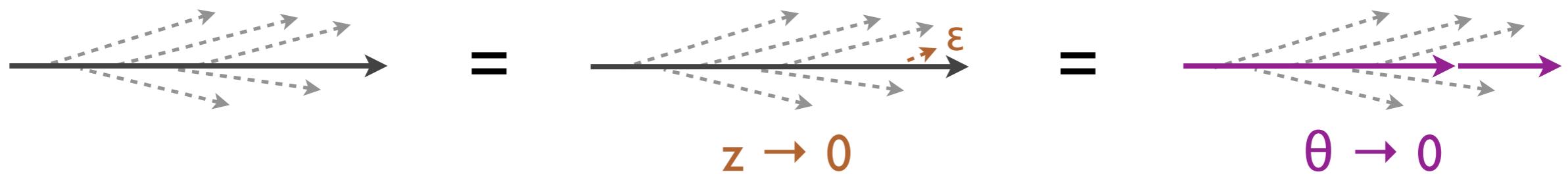


$\mathcal{P}_N \supset \mathcal{P}_{N-1} \supset \dots \supset \mathcal{P}_2 \supset \mathcal{P}_1$ by **soft/collinear limits**

When are Two Events the Same?

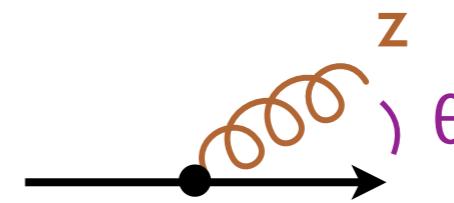
$$\mathcal{E}(\hat{n}) = \sum_i E_i \delta(\hat{n} - \hat{n}_i)$$

Energy Flow unchanged by infinitesimal soft/collinear emissions



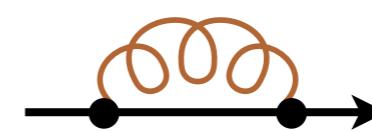
Infrared divergences “live” together!

Real:



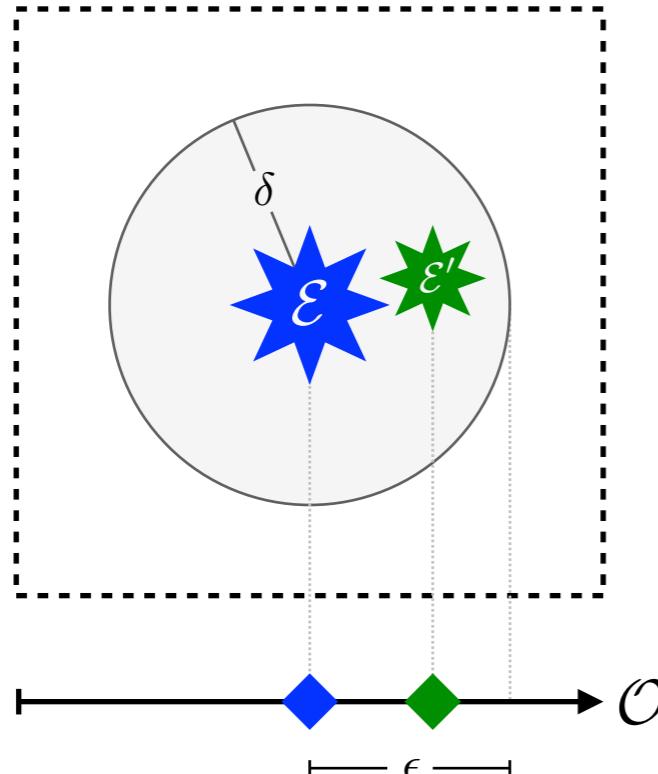
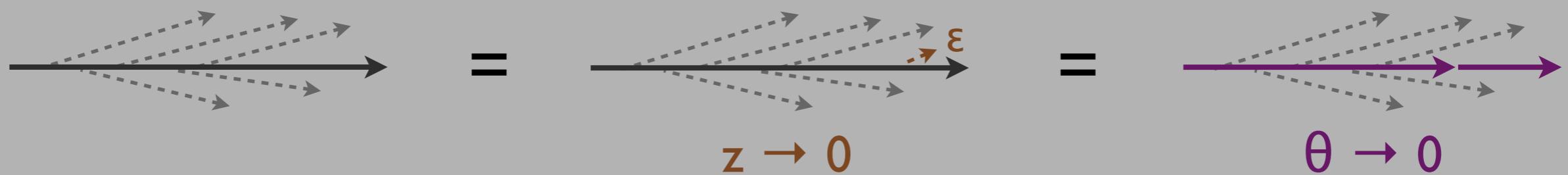
$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{dz}{z} \frac{d\theta}{\theta}$$

Virtual:



When are Two Events the Same?

Energy Flow unchanged by infinitesimal soft/collinear emissions



Infrared & Collinear Safety

≈ calculable in perturbative quantum field theory

*is**

Continuity in EMD Space

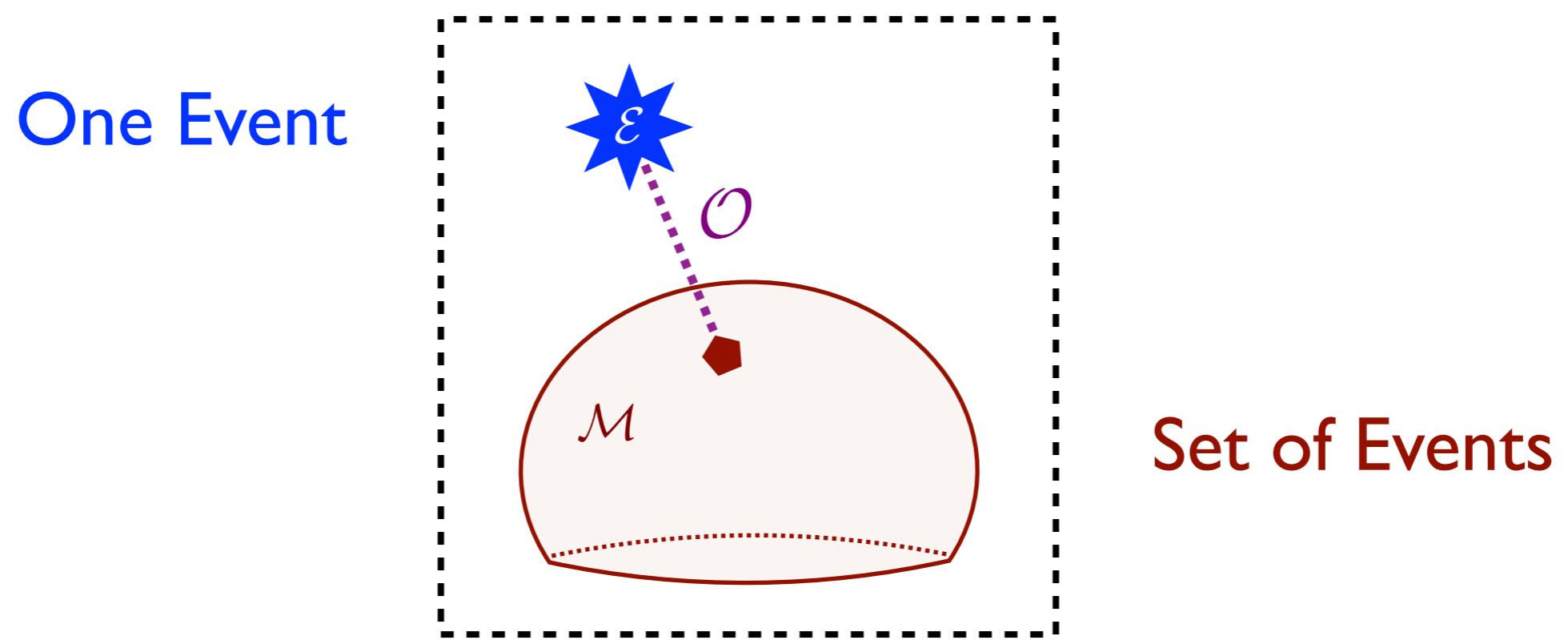
[Komiske, Metodiev, JDT, [JHEP 2020](#)
[Sterman, Weinberg, [PRL 1977](#); Sterman, [PRD 1979](#)
[see also Banfi, Salam, Zanderighi, [JHEP 2005](#)]

*EMD seems to define the “natural”
geometry for massless gauge theories*

Open question: Can you define $|\mathcal{M}_{AB \rightarrow 12\dots n}|^2$ directly in this space?

Manifolds for Observables

$$f_{\text{obs}}(\Phi_n)$$



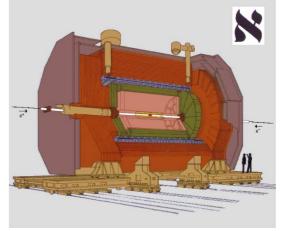
Distance of Closest Approach \Rightarrow Observable

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

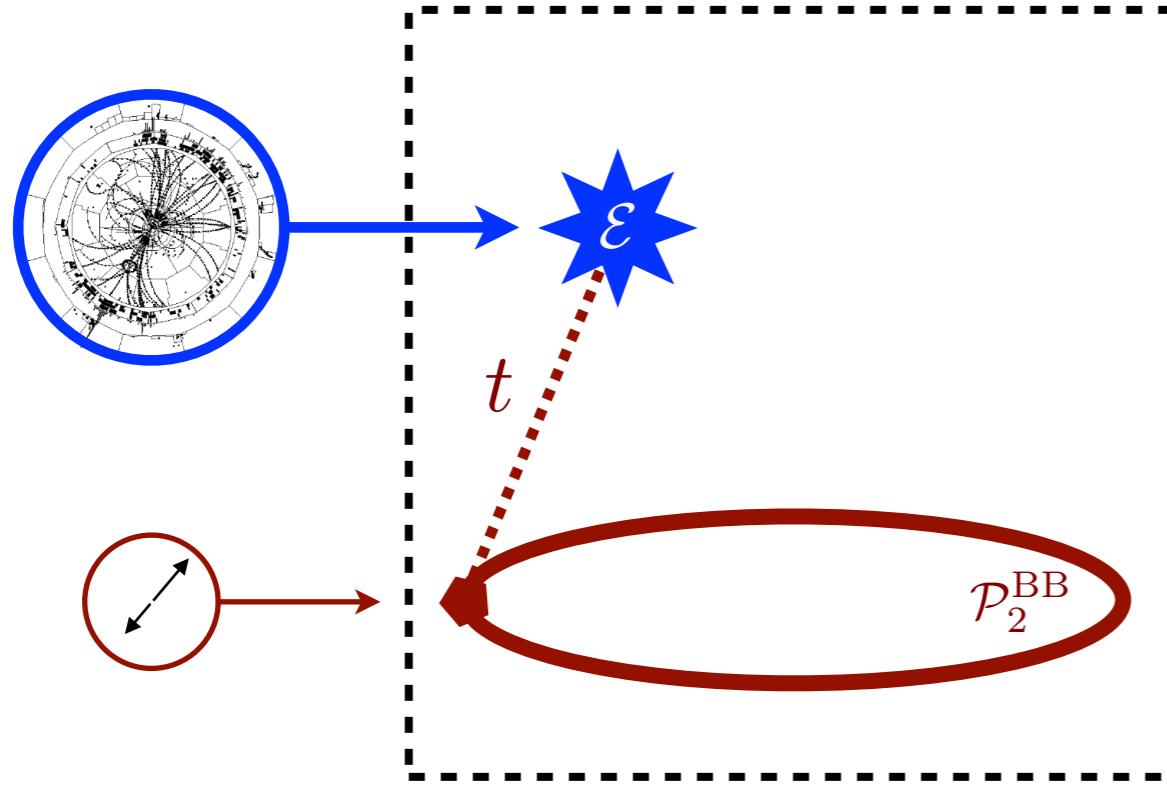
[Komiske, Metodiev, JDT, [JHEP 2020](#)]

E.g. Thrust

How dijet-like is an event?



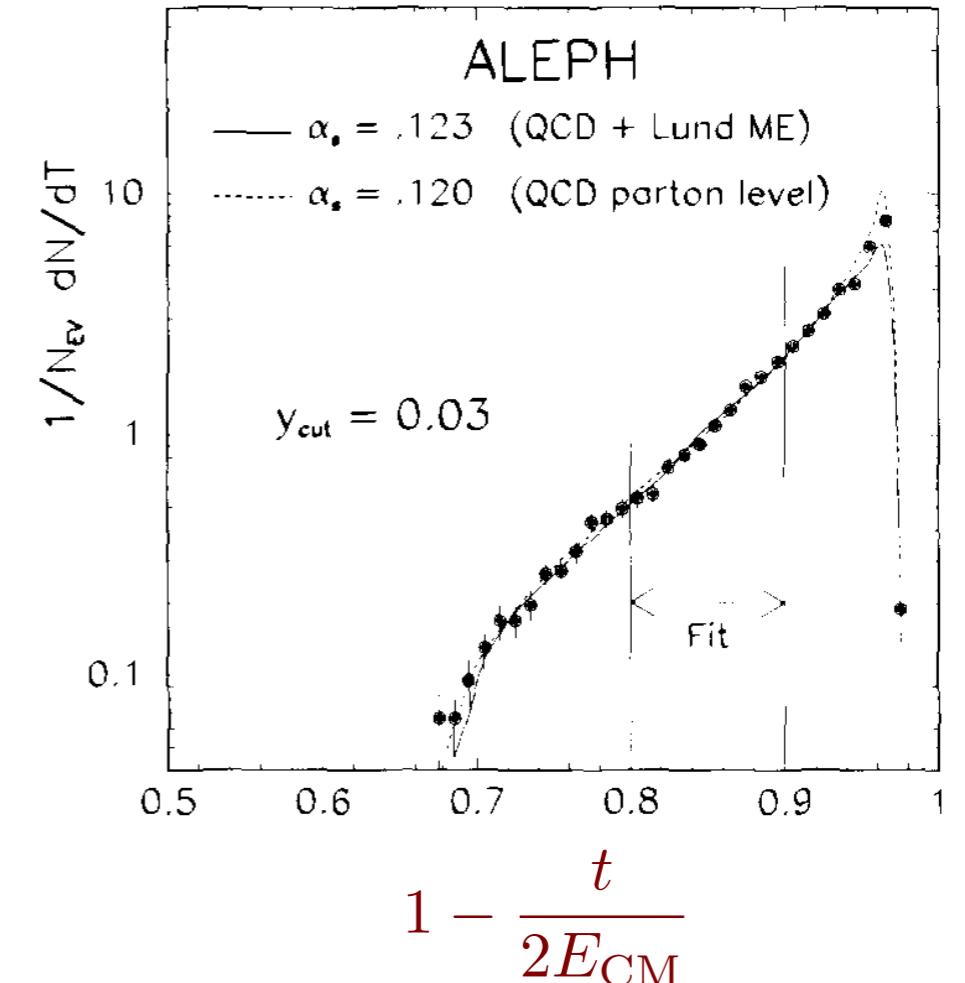
$$t(\mathcal{E}) = \min_{\mathcal{E}' \in \mathcal{P}_2^{\text{BB}}} \text{EMD}_2(\mathcal{E}, \mathcal{E}')$$



All Back-to-Back Two Particle Configurations

$$\mathcal{P}_2^{\text{BB}} = \left\{ \begin{array}{c} \text{red circle with internal arrow} \\ \text{red circle with internal double-headed arrow} \\ \text{red circle with internal diagonal arrow} \\ \text{red circle with internal horizontal arrow} \\ \dots \end{array} \right\}$$

(using $\beta=2$ EMD variant)



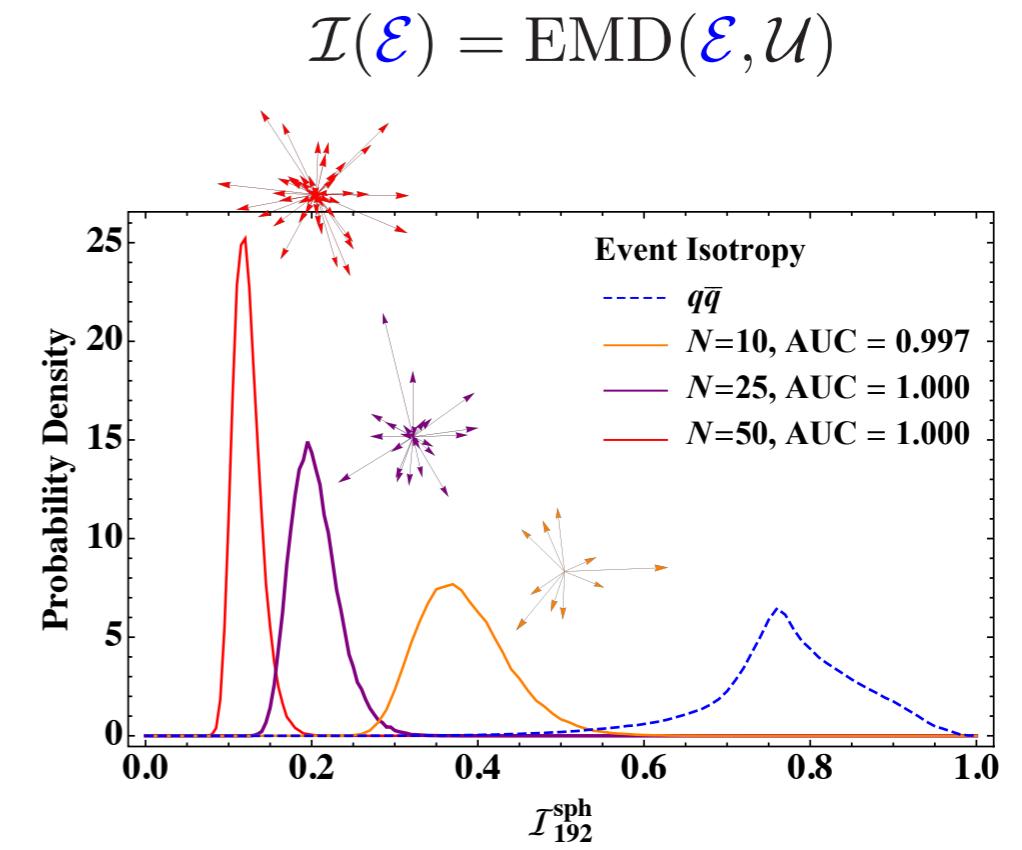
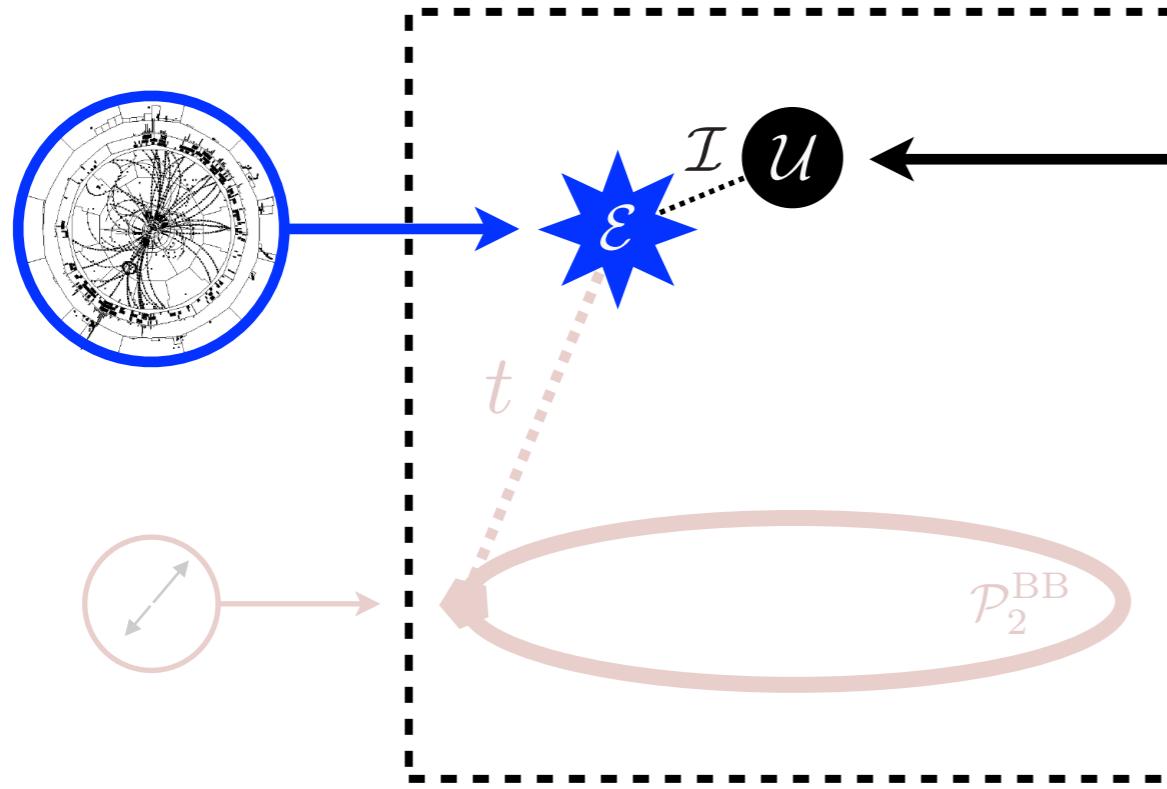
$$1 - \frac{t}{2E_{\text{CM}}}$$

$$\text{cf. } T(\mathcal{E}) = \max_{\hat{n}} \frac{\sum_i |\vec{p}_i \cdot \hat{n}|}{\sum_j |\vec{p}_j|}$$

[Komiske, Metodiev, JDT, JHEP 2020]
 [Brandt, Peyrou, Sosnowski, Wroblewski, PL 1964; Farhi, PRL 1977; ALEPH, PLB 1991]

New! Event Isotropy

How isotropic is an event?

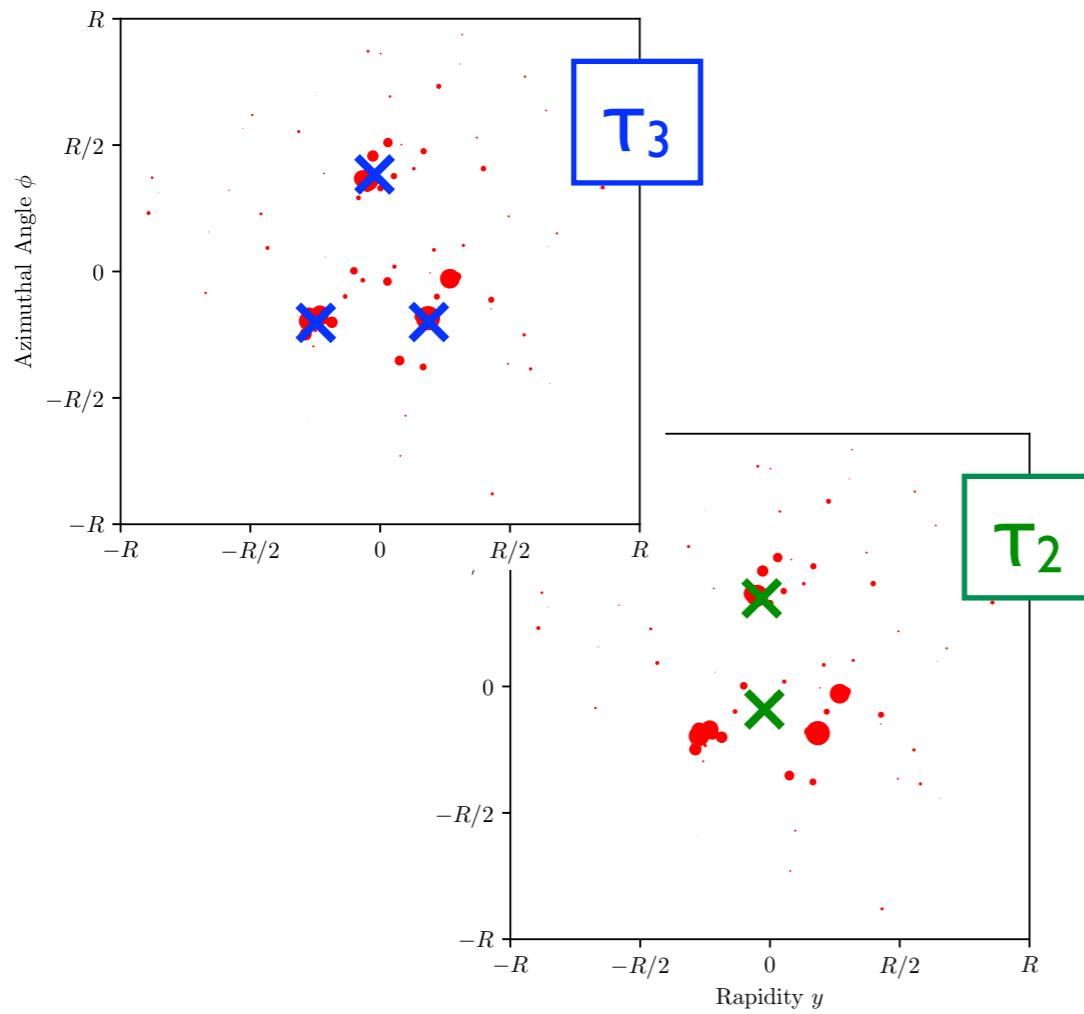


[Cesarotti, JDT, [JHEP 2020](#);
see also Cesarotti, Reece, Strassler, [arXiv 2020](#)]

N-subjettiness

Ubiquitous jet substructure observable used for almost a decade...

$$\tau_N(\mathcal{J}) = \min_{N \text{ axes}} \sum_i E_i \min \{\theta_{1,i}, \theta_{2,i}, \dots, \theta_{N,i}\}$$

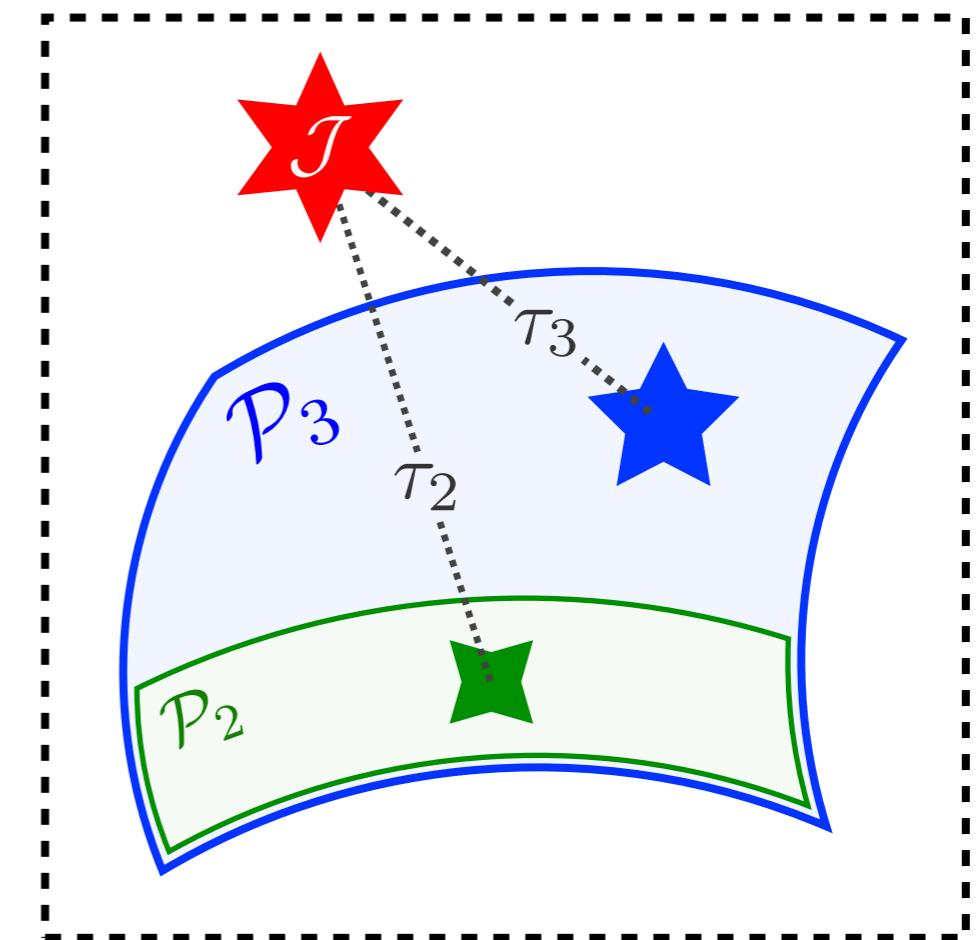
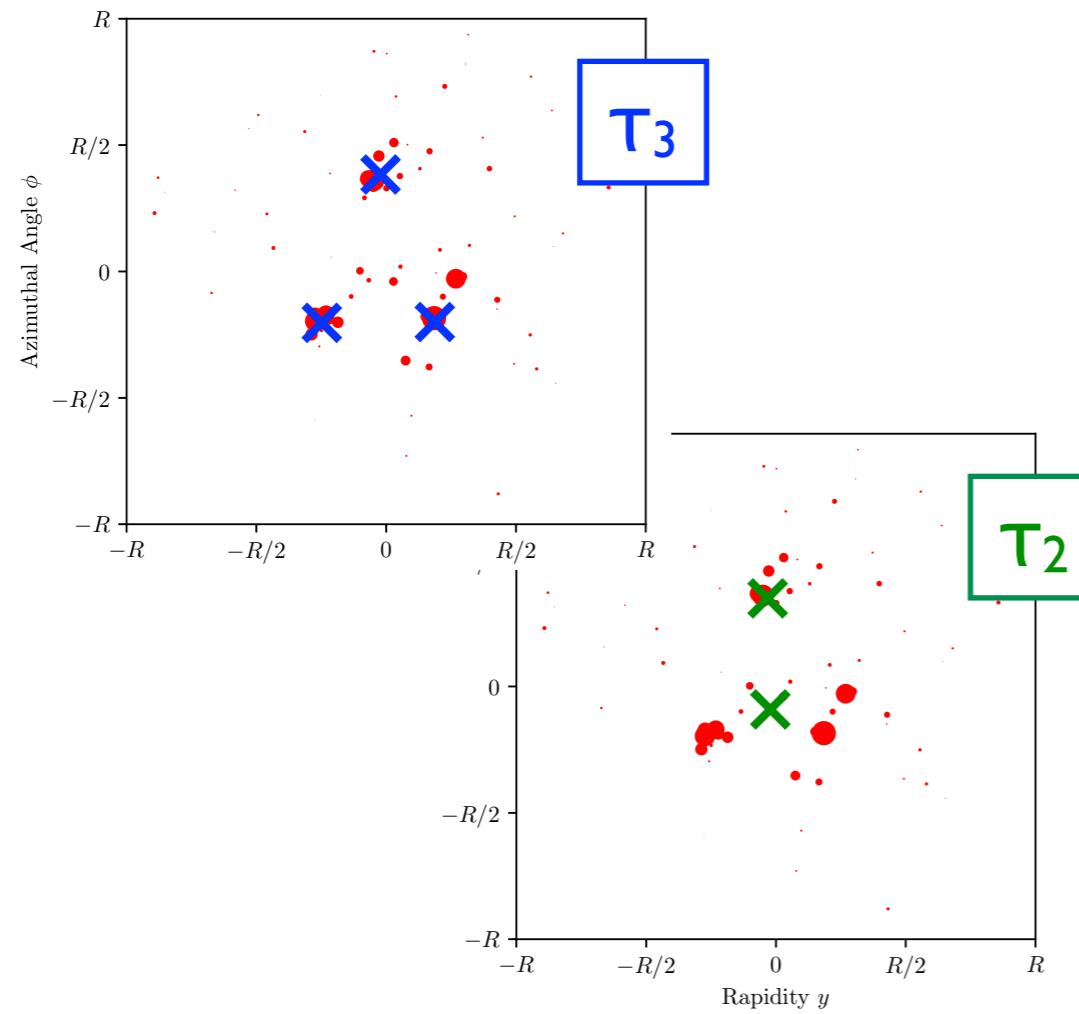


[JDT, Van Tilburg, [JHEP 2011](#), [JHEP 2012](#);
based on Brandt, Dahmen, [ZPC 1979](#); Stewart, Tackmann, Waalewijn, [PRL 2010](#)]

N-subjettiness = Point to Manifold EMD

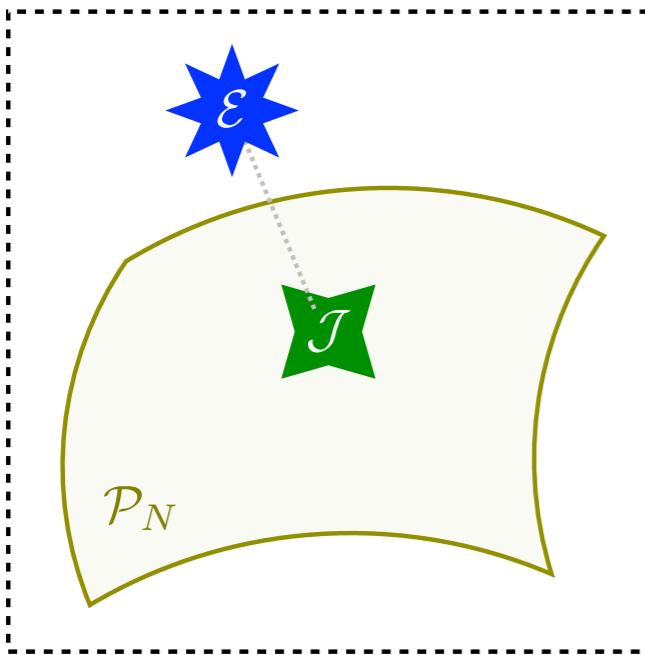
...is secretly an optimal transport problem

$$\tau_N(\mathcal{J}) = \min_{\mathcal{J}' \in \mathcal{P}_N} \text{EMD}(\mathcal{J}, \mathcal{J}')$$



[JDT, Van Tilburg, [JHEP 2011](#), [JHEP 2012](#);
rephrased in the language of Komiske, Metodiev, JDT, [PRL 2019](#)]

More Fun with N-particle Manifolds



N-jettiness

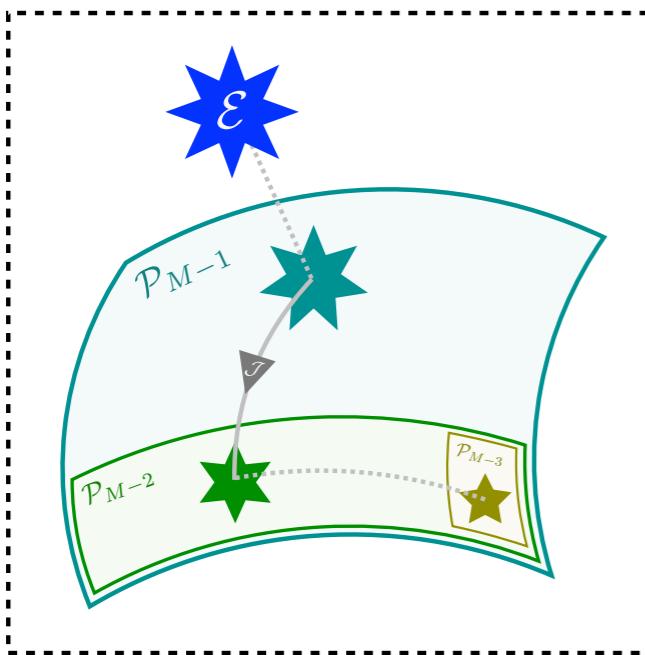
Distance of closest approach to N -particle manifold

[Brandt, Dahmen, [ZPC 1979](#); Stewart, Tackmann, Waalewijn, [PRL 2010](#)]

Exclusive Cone Jet Finding

Point of closest approach on N -particle manifold

[Stewart, Tackmann, JDT, Vermilion, Wilkason, [JHEP 2015](#)]



Sequential Jet Recombination

Iteratively stepping between various N -particle manifolds

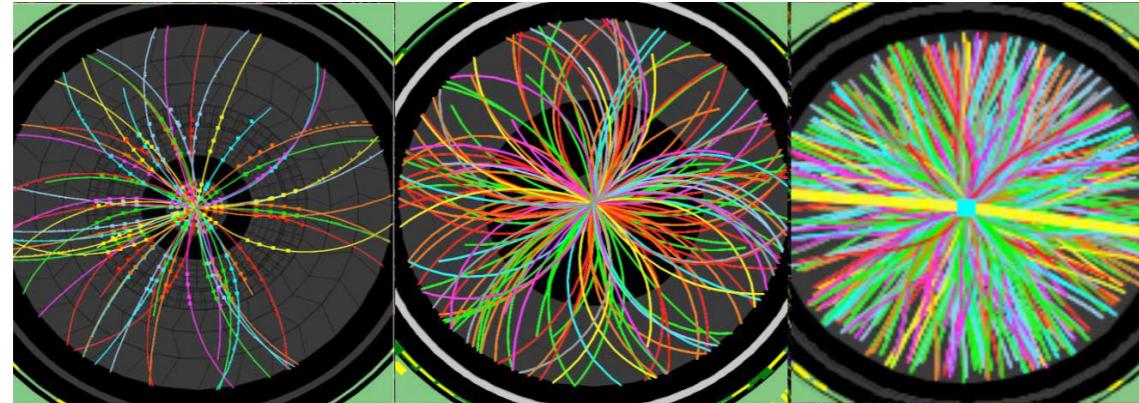
[Catani, Dokshitzer, Seymour, Webber, [NPB 1993](#); Ellis, Soper, [PRD 1993](#)]

[Dokshitzer, Leder, Moretti, Webber, [JHEP 1997](#); Wobisch, Wengler, [arXiv 1999](#)]

[Butterworth, Couchman, Cox, Waugh, [CPC 2003](#); Larkoski, Neill, JDT, [JHEP 2014](#)]

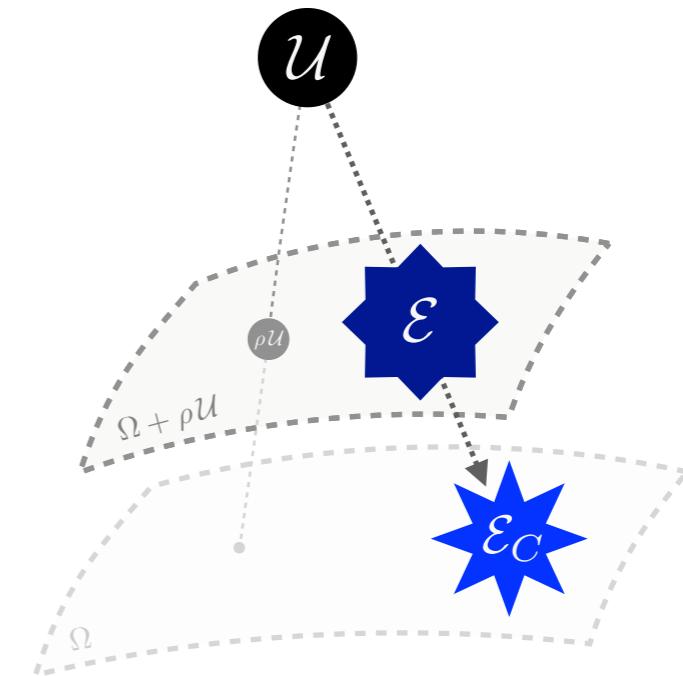
[Komiske, Metodiev, JDT, [JHEP 2020](#)]

Pileup Mitigation



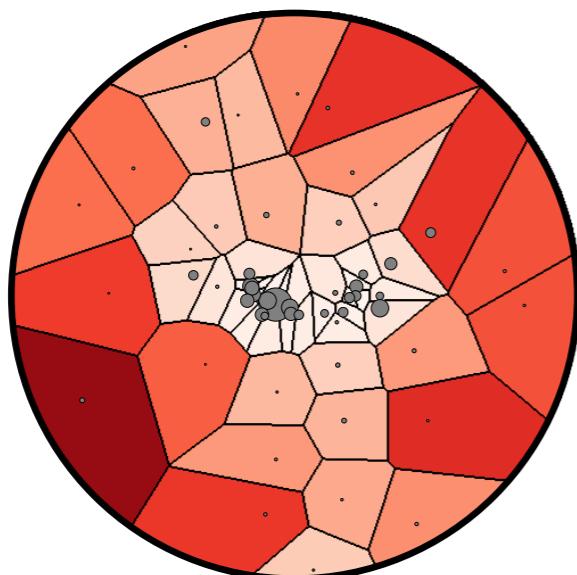
[see review in Soyez, PR 2019]

Uniform event contamination from overlapping proton-proton collisions



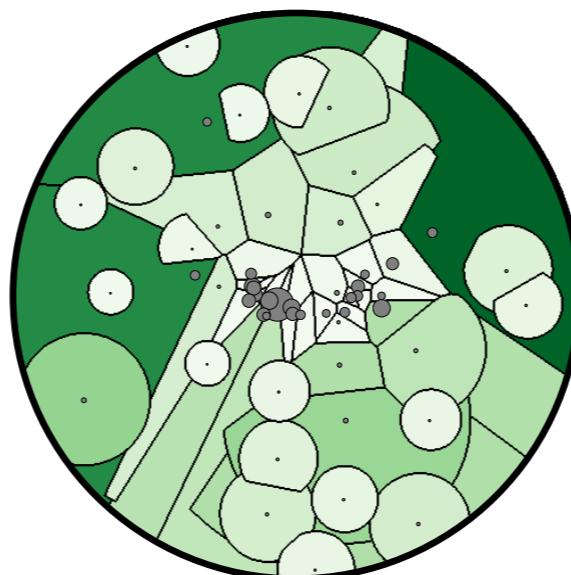
Pileup Mitigation:
“Move away” from uniform event

Voronoi



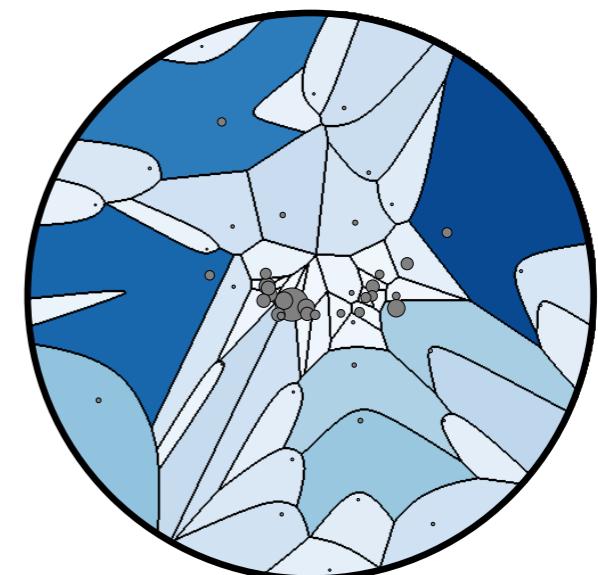
[Cacciari, Salam, Soyez, JHEP 2008]

Constituent Subtraction



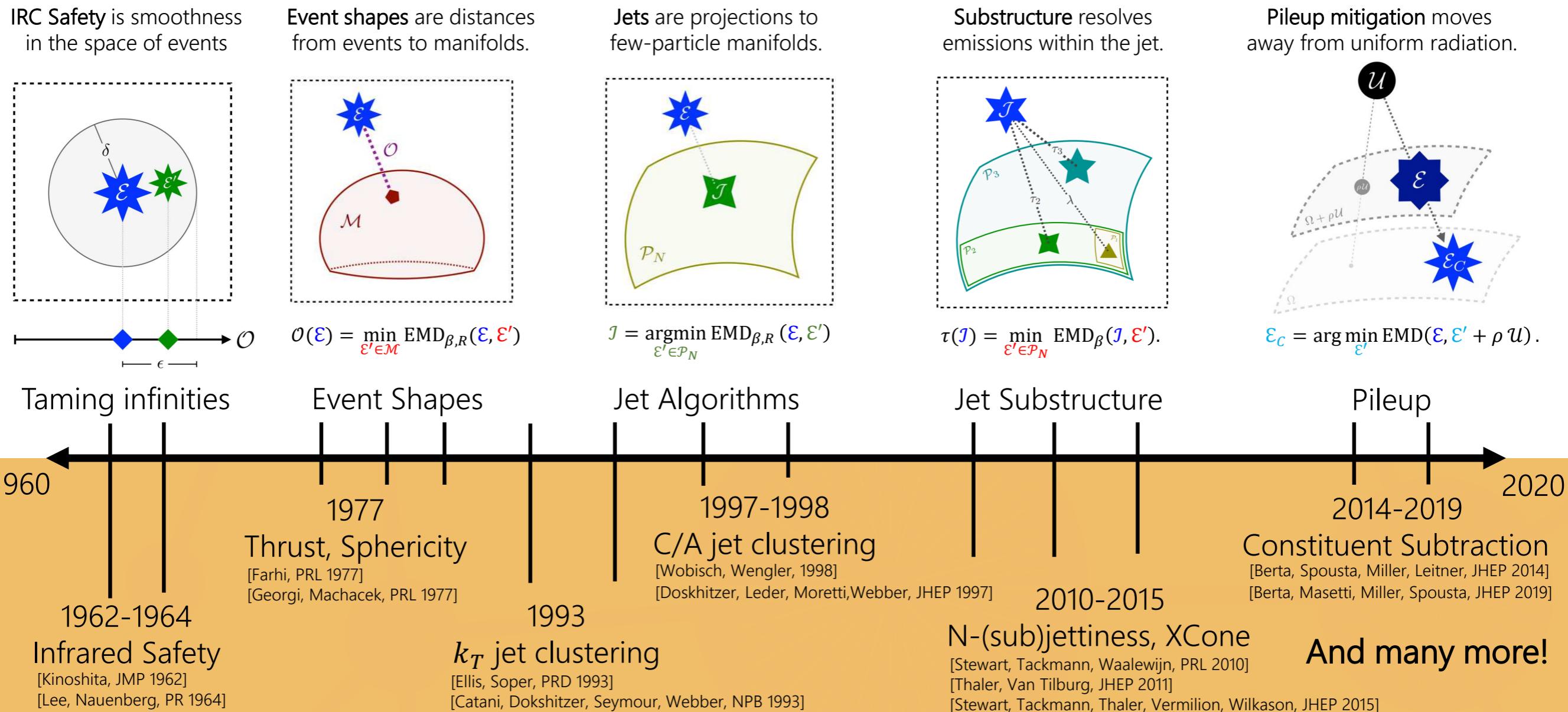
[Berta, Spousta, Miller, Leitner, JHEP 2014]

Apollonius



[Komiske, Metodiev, JDT, JHEP 2020]

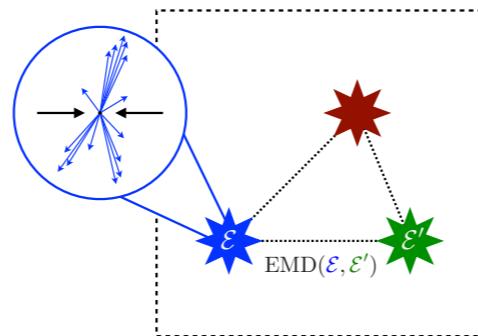
Six Decades of Collider Physics Translated into a New Geometric Language!



[timeline from Eric Metodiev]

Pause

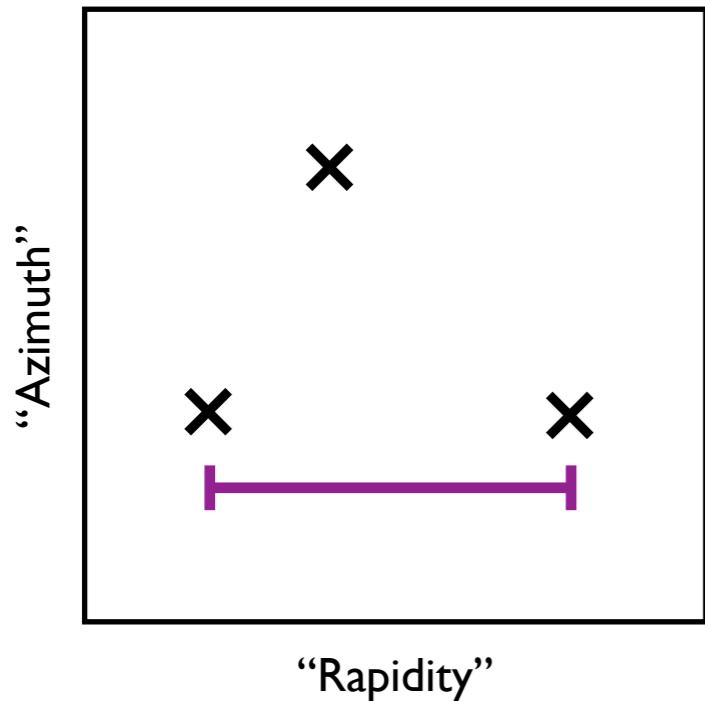
What can be Geometrized?



IRC Safety, Observables,
Jet Algorithms, Pileup Mitigation

How far down does this rabbit hole go?

Direction Space



x = Direction

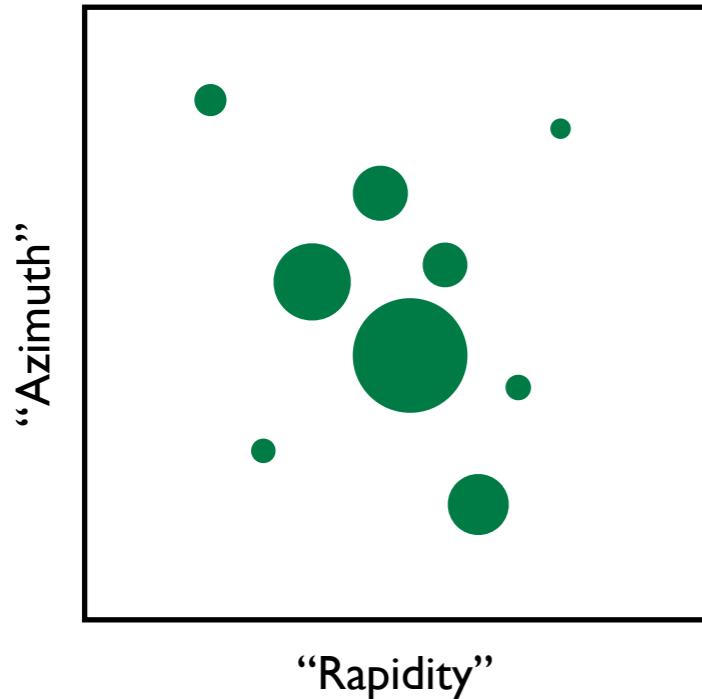
= Angular Distance

$$n_i^\mu = \frac{p_i^\mu}{E_i} = (1, \hat{n})^\mu$$

$$\theta_{ij} = \sqrt{2n_i^\mu n_{j\mu}}$$

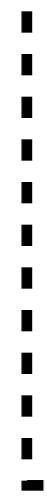
(for massless particles)

Direction Space Distribution



● = Weighted Direction

— = Angular Distance



★ = Event

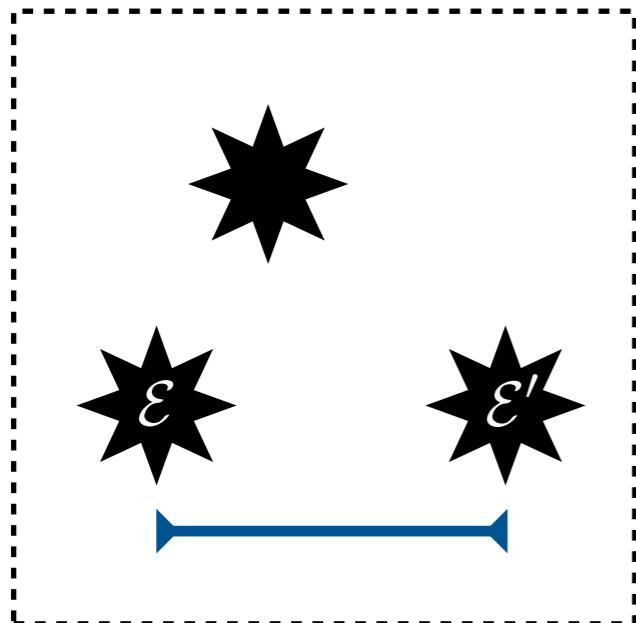
$$n_i^\mu = \frac{p_i^\mu}{E_i} = (1, \hat{n})^\mu$$

$$w_i = E_i$$

$$\theta_{ij} = \sqrt{2n_i^\mu n_{j\mu}}$$

(for massless particles)

Event Space



★ = Event
↔ = EMD
Energy Mover's Distance

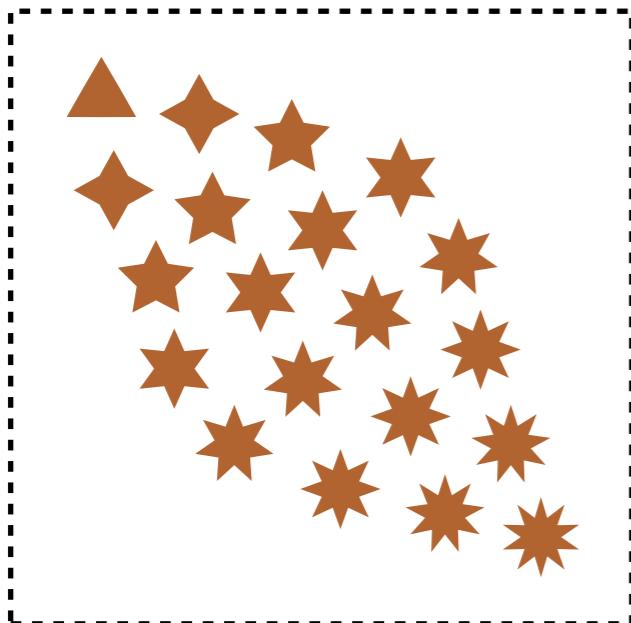
$$\mathcal{E}(\hat{n}) = \sum_i E_i \delta(\hat{n} - \hat{n}_i)$$

$$\text{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_i \sum_j f_{ij} \theta_{ij}$$

(for equal total energy)

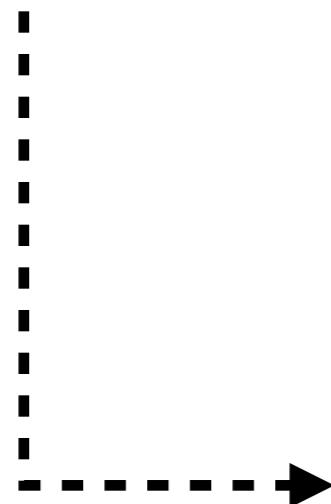
[Komiske, Metodiev, JDT, PRL 2019]

Event Space Distribution



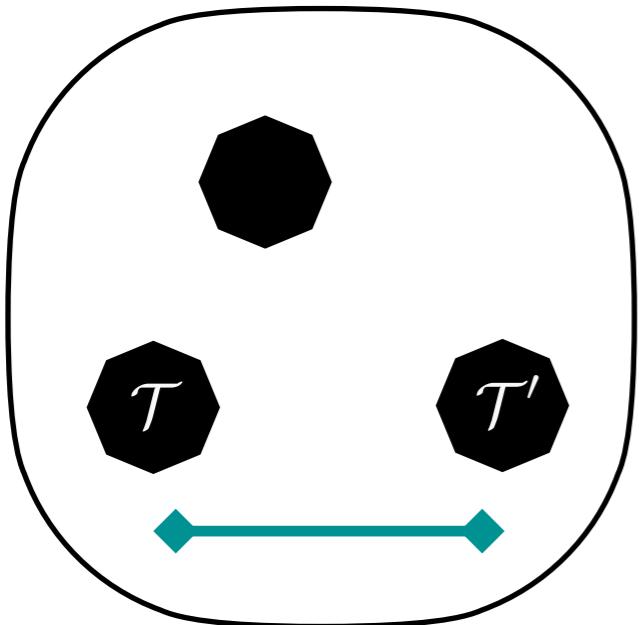
= **Weighted Event** $\mathcal{E}(\hat{n}) = \sum_i E_i \delta(\hat{n} - \hat{n}_i)$
 $w_a = \sigma_a$

= **EMD**
Energy
Mover's Distance $\text{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_i \sum_j f_{ij} \theta_{ij}$
(for equal total energy)



= **Theory**

Theory Space



● = Theory

↔ = ΣMD
Cross-Section
Mover's Distance

$$\mathcal{T}(\mathcal{E}) = \sum_a \sigma_a \delta(\mathcal{E} - \mathcal{E}_a)$$

$$\Sigma\text{MD}(\mathcal{T}, \mathcal{T}') = \min_{\{\mathcal{F}\}} \sum_a \sum_b \mathcal{F}_{ab} \text{EMD}(\mathcal{E}_a, \mathcal{E}'_b)$$

(for equal total xsec)

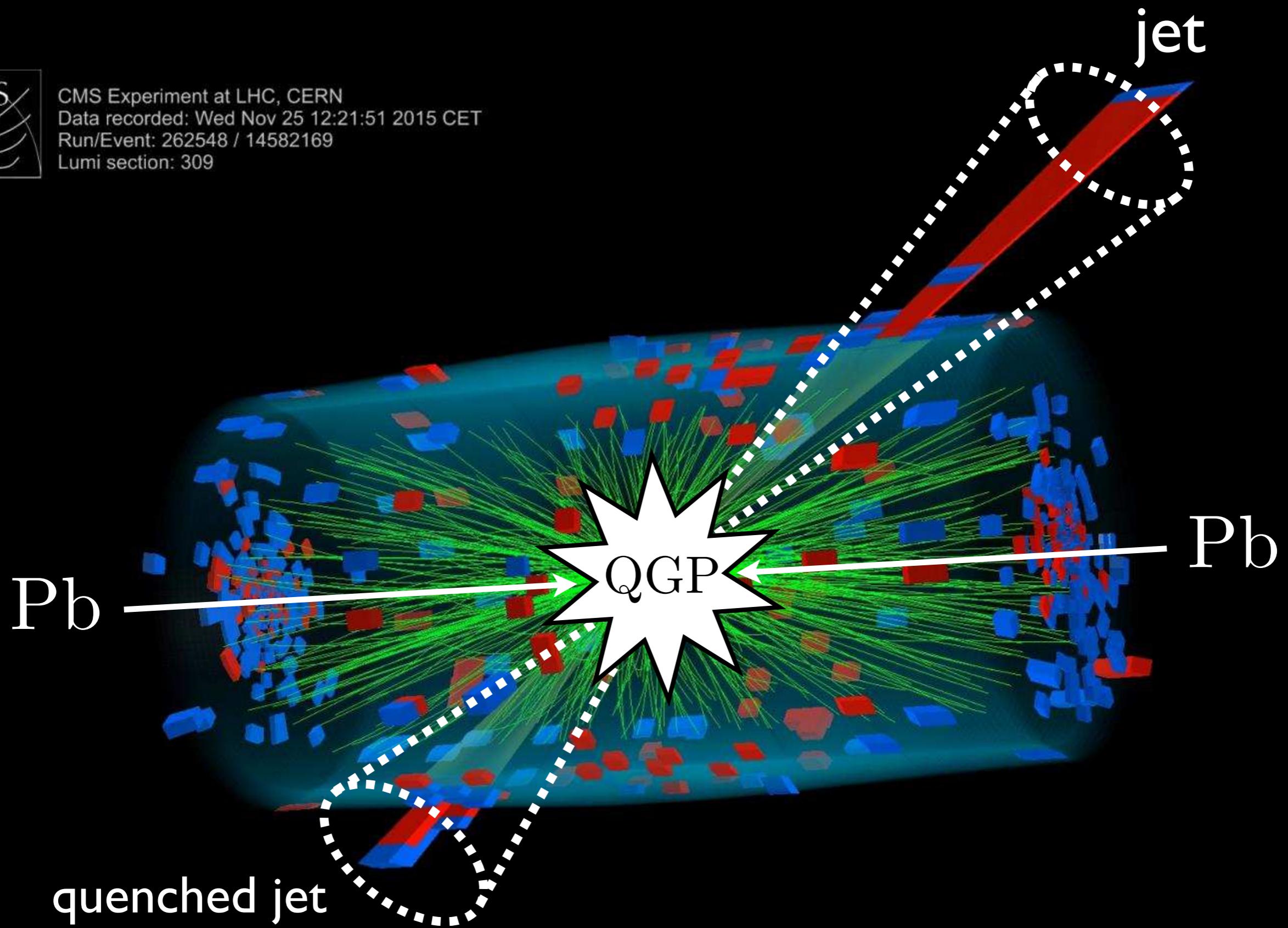
A distance between theories!

(e.g. EMD : N-jettiness :: ΣMD : k-eventiness)

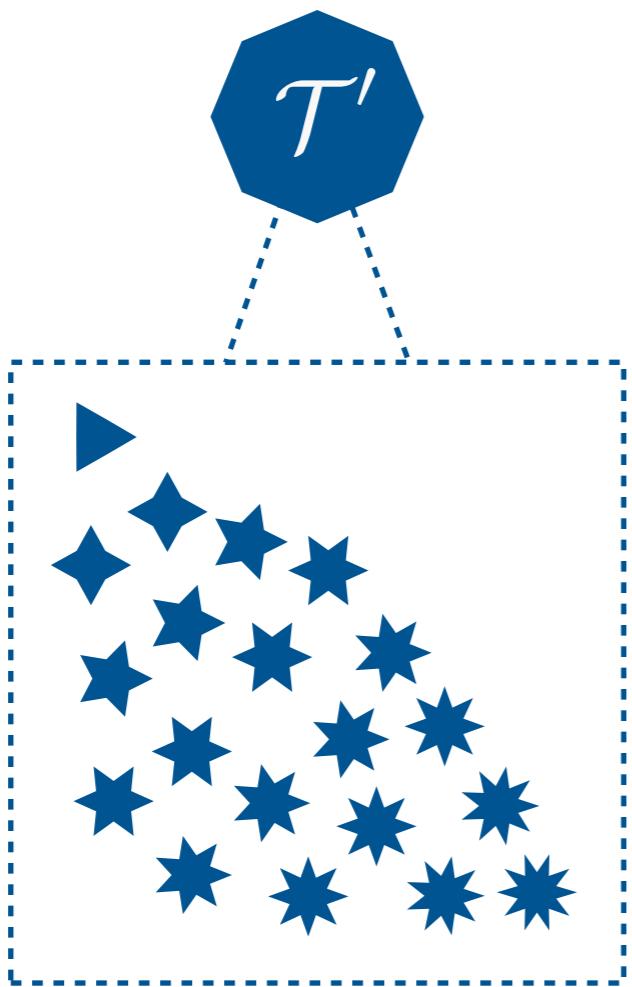
[Komiske, Metodiev, JDT, [JHEP 2020](#)]



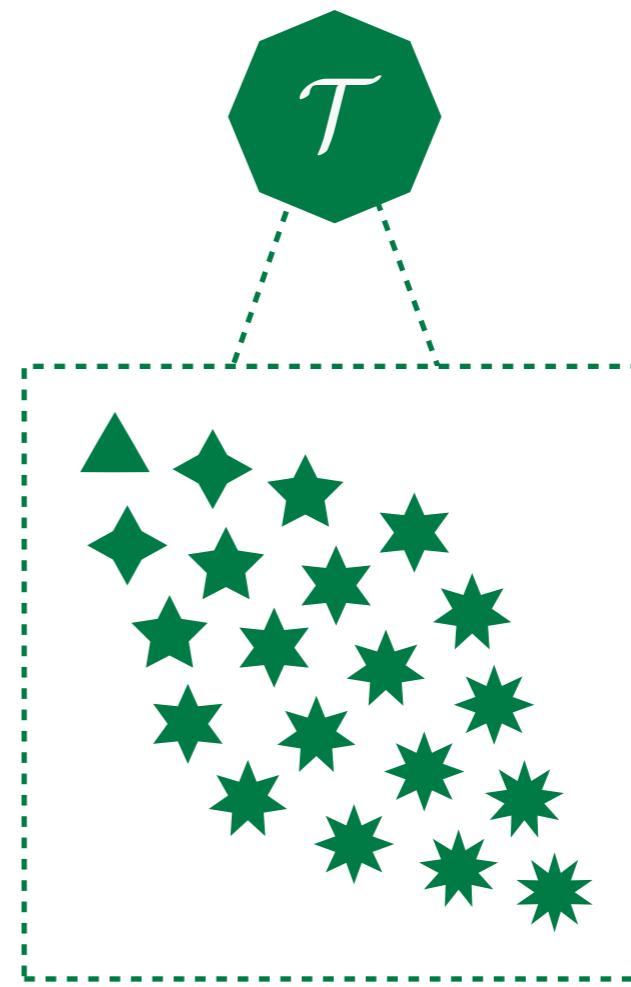
CMS Experiment at LHC, CERN
Data recorded: Wed Nov 25 12:21:51 2015 CET
Run/Event: 262548 / 14582169
Lumi section: 309



Theory Prime: In-Medium QCD



Theory: Vacuum QCD

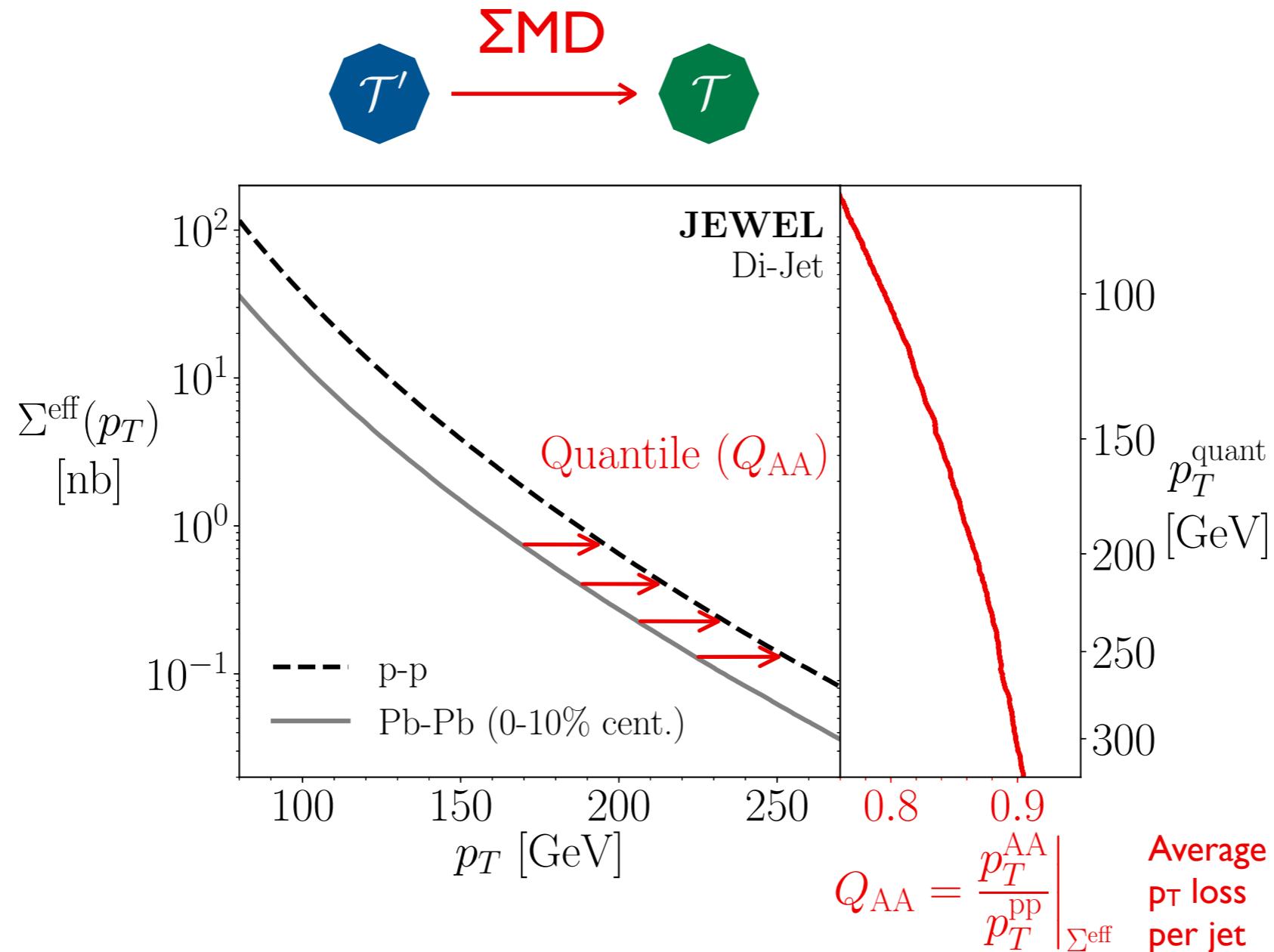


ΣMD
 \iff

*Optimal transportation plan defines mapping
between in-medium jets and vacuum jets!*

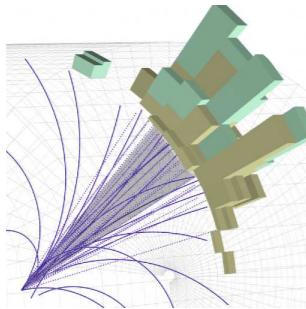
Jet Quenching via Quantile Matching

Equivalent to following a geodesic in theory space (!)



[Brewer, Milhano, JDT, PRL 2019]

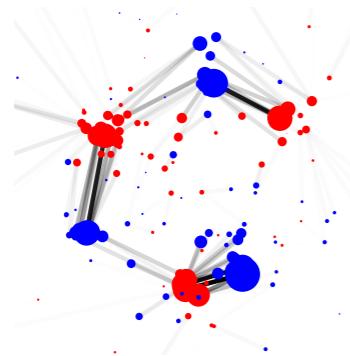
Summary



What is a Collider Event?

An unordered set of particles that describes the energy flow away from the collision point

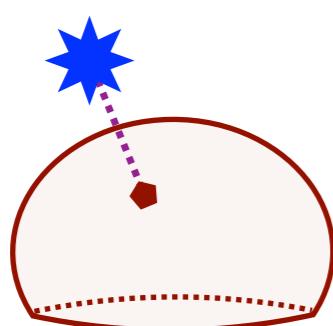
(ask me about ML & QCD)



When are Events Similar?

When they are close in the geometric space triangulated by the energy mover's distance

(ask me more about EMD)



What can be Geometrized?

Many concepts/techniques in quantum field theory and collider physics from the last half century

Fin

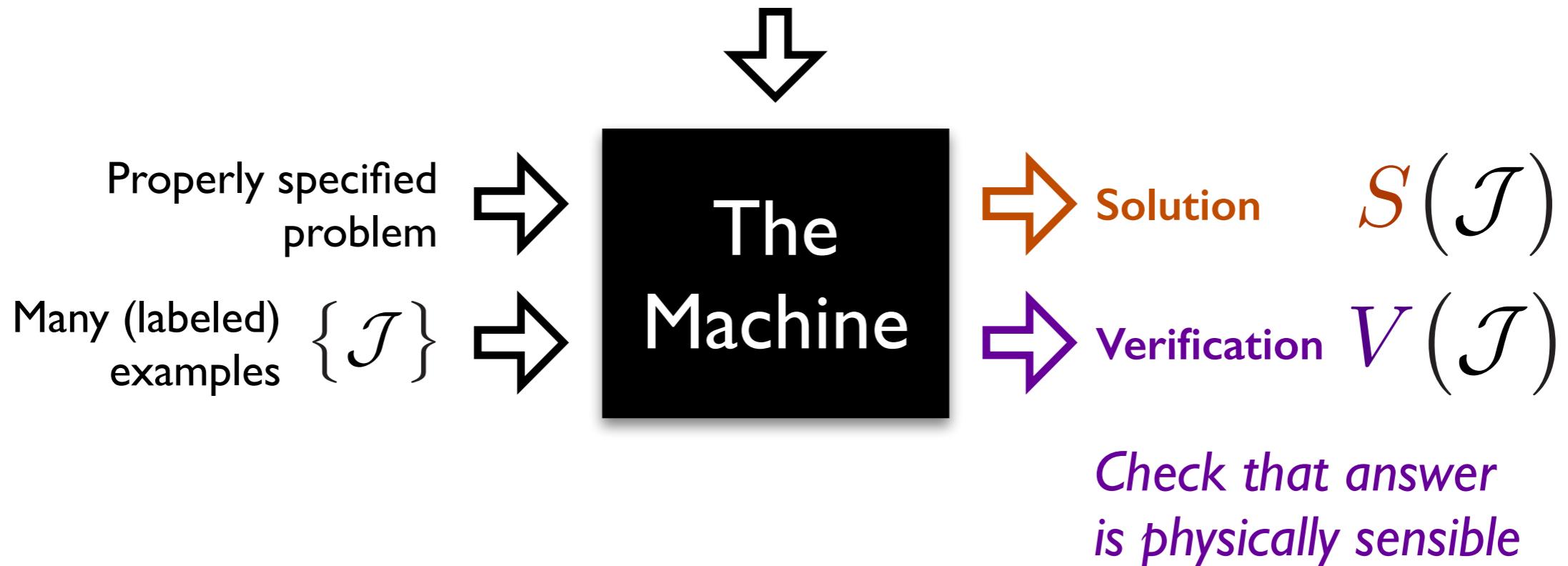
Backup Slides

Aside: Machine Learning for Jets

“ML4Jets”
NYU, January 2020

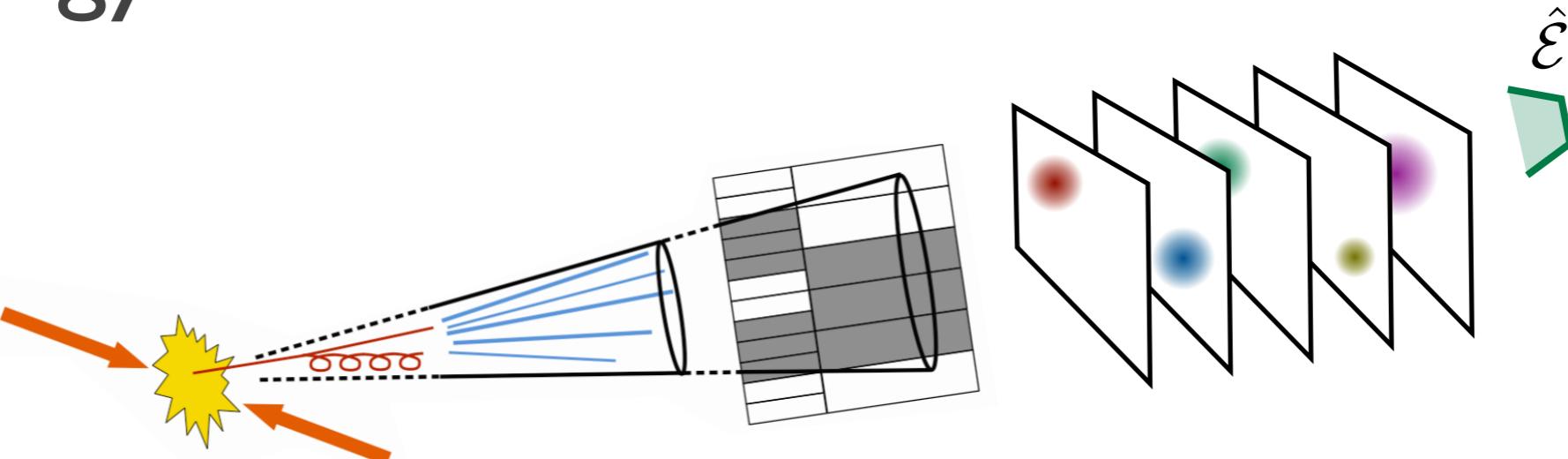
Symmetry: $\mathcal{J} = \{ \vec{p}_1, \vec{p}_2, \vec{p}_3, \dots, \vec{p}_N \}$
Unordered, Variable Length Set (QM!)

Safety: $\vec{p} = \{ E, \hat{n}_x, \hat{n}_y, \hat{n}_z \}$
Energy weighting (QFT!)

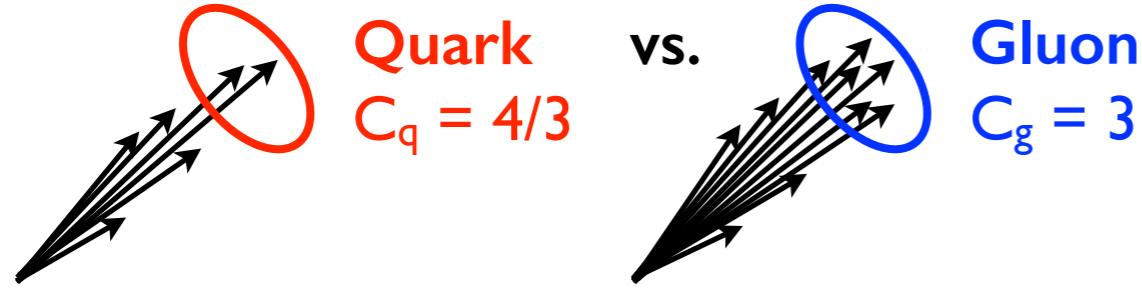


*Check that answer
is physically sensible*

E.g. Energy Flow Networks

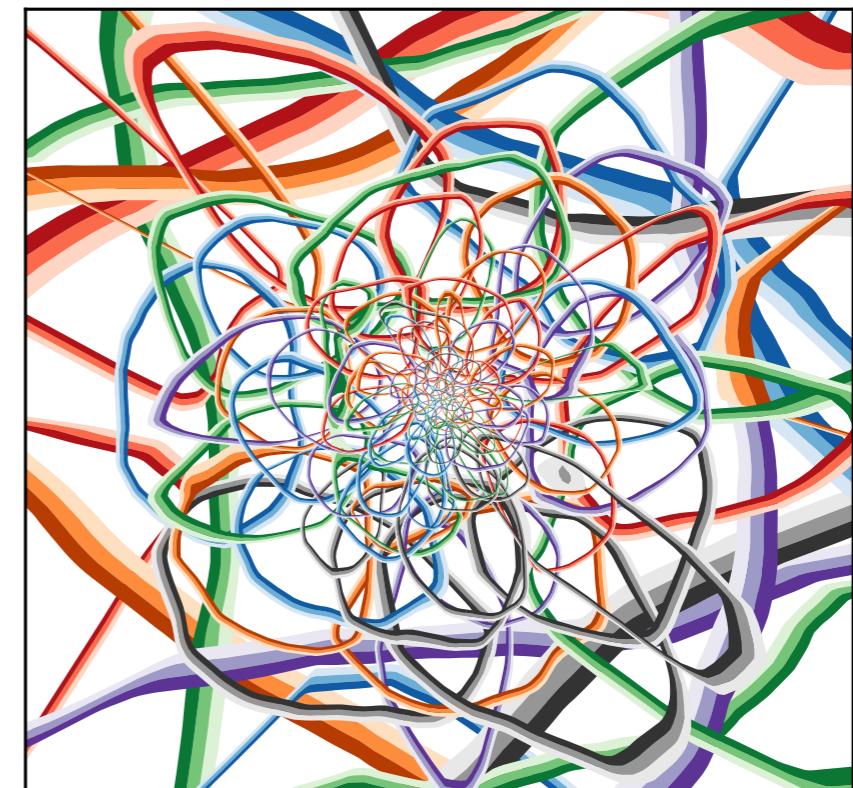
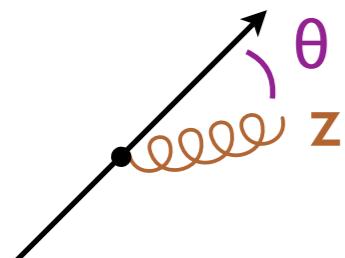


Learning QCD singularities!



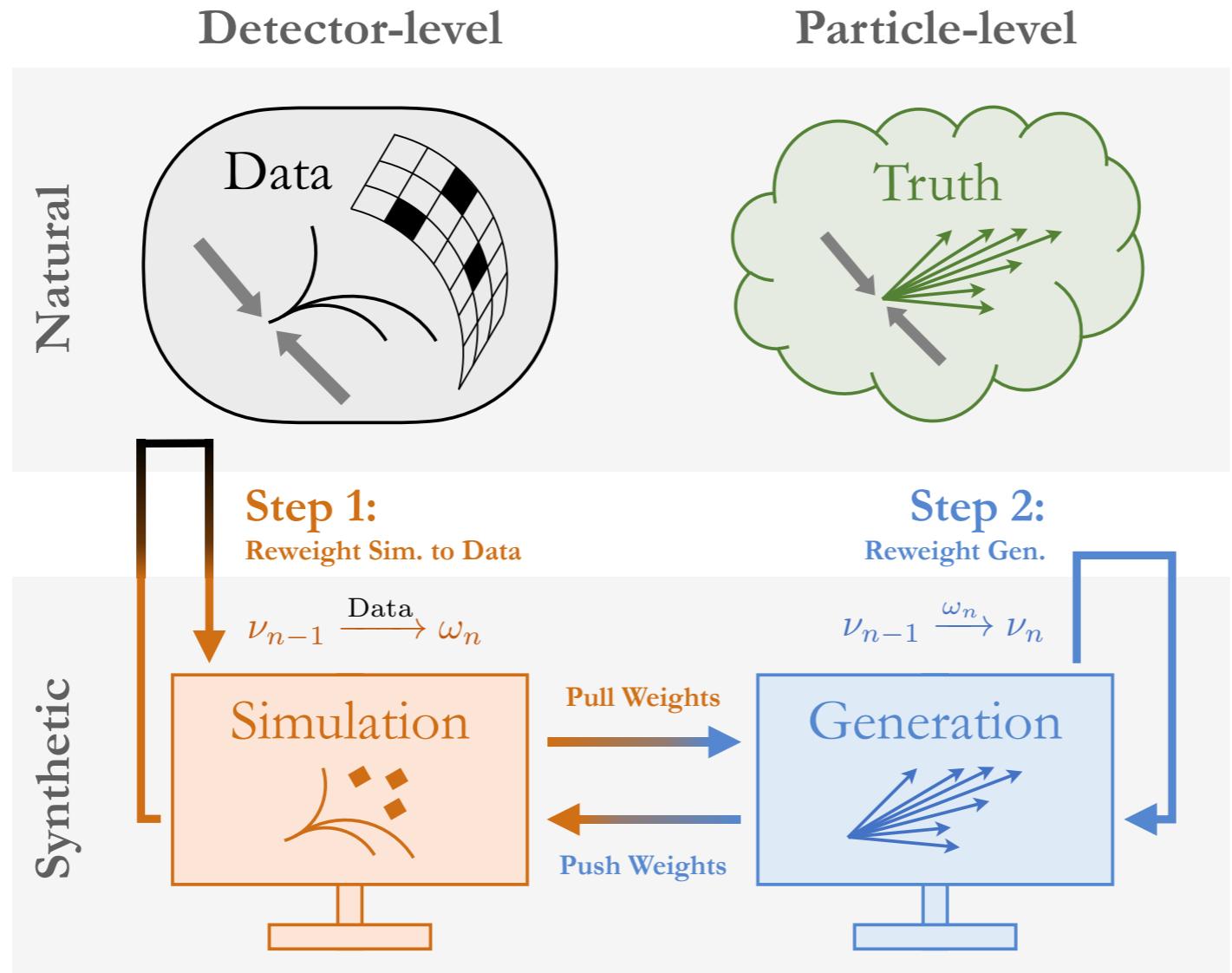
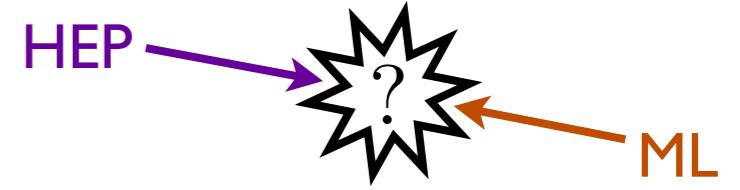
AP splitting probability:

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



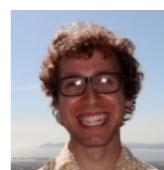
[Komiske, Metodiev, JDT, [JHEP 2019](#); see also Komiske, Metodiev, JDT, [JHEP 2018](#), [PRD 2020](#); special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, [NIPS 2017](#)]

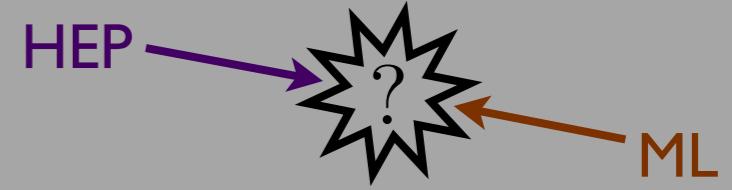
OmniFold



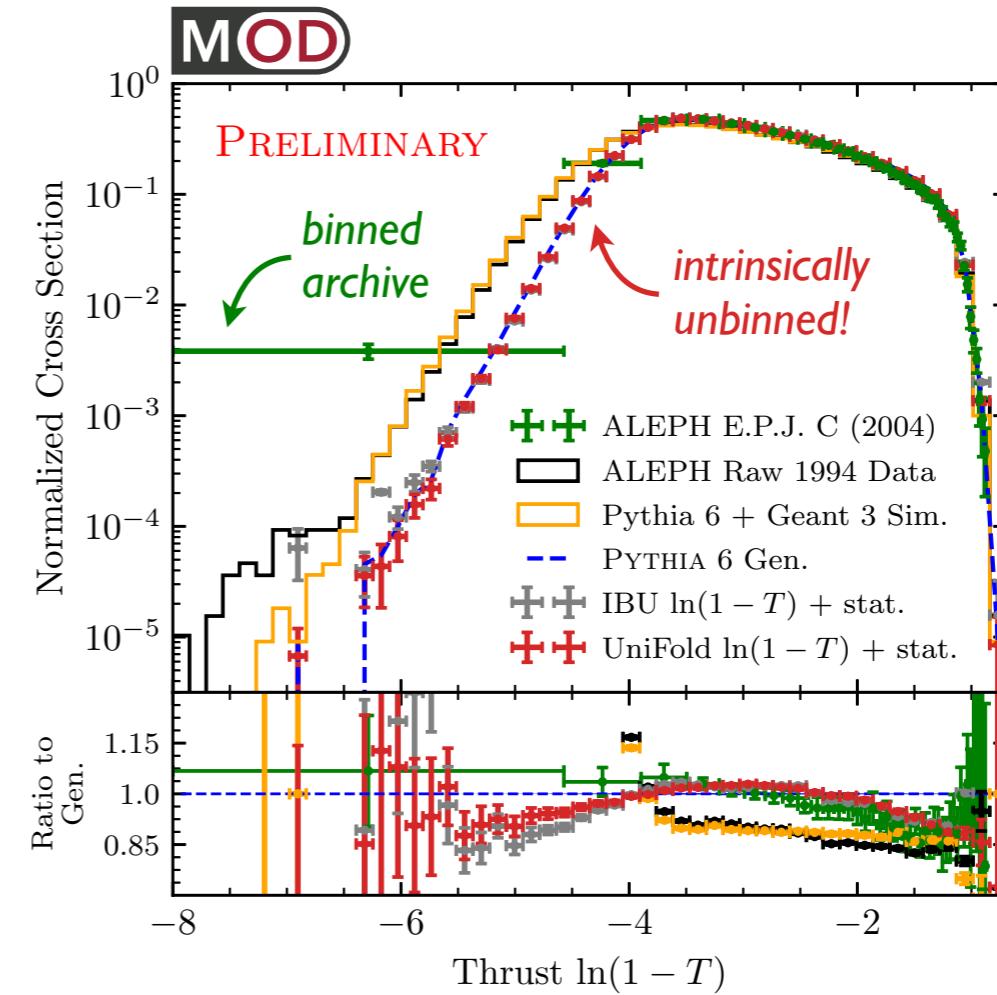
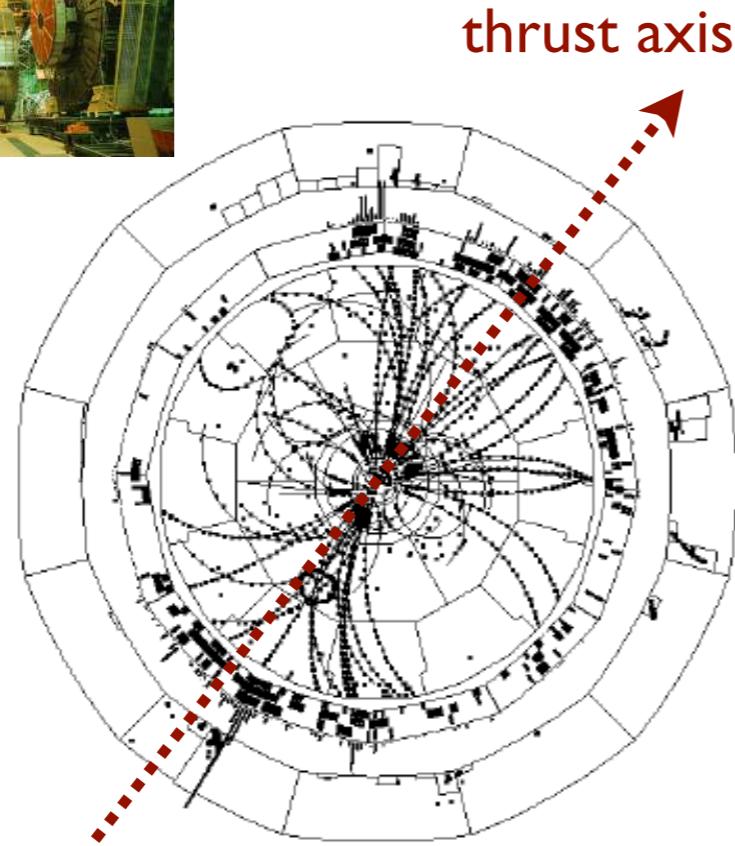
Multi-dimensional unbinned detector corrections via iterated binary classification

[Andreassen, Komiske, Metodiev, Nachman, JDT, [PRL 2020](#)]

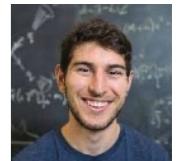




Back to the Future with ALEPH Archival Data



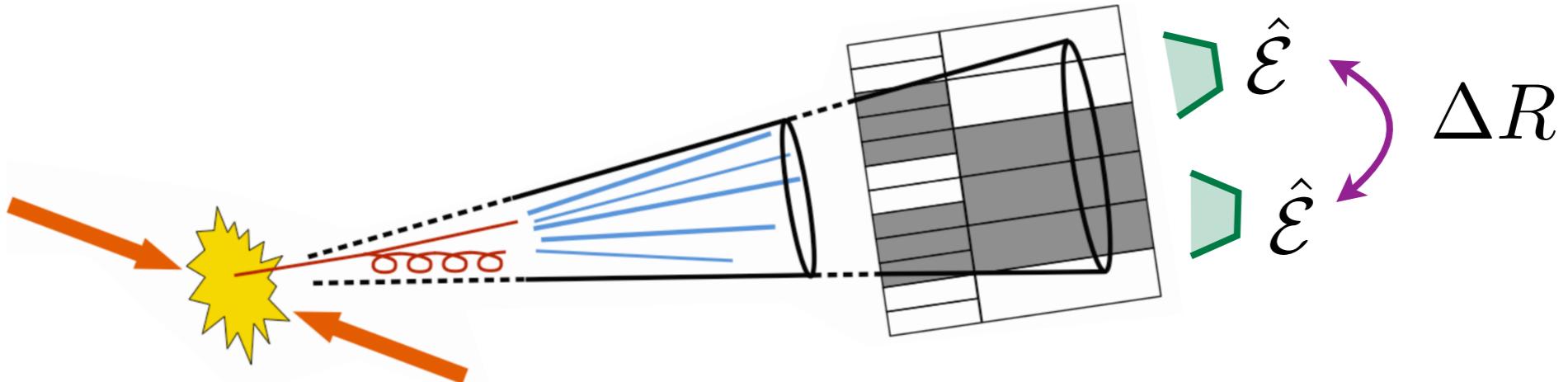
[talk by Badea, [ICHEP 2020](#); cf. ALEPH, [EPJC 2004](#)
 [see also Badea, Baty, Chang, Innocenti, Maggi, McGinn, Peters, Sheng, JDT, Lee, [PRL 2019](#)]



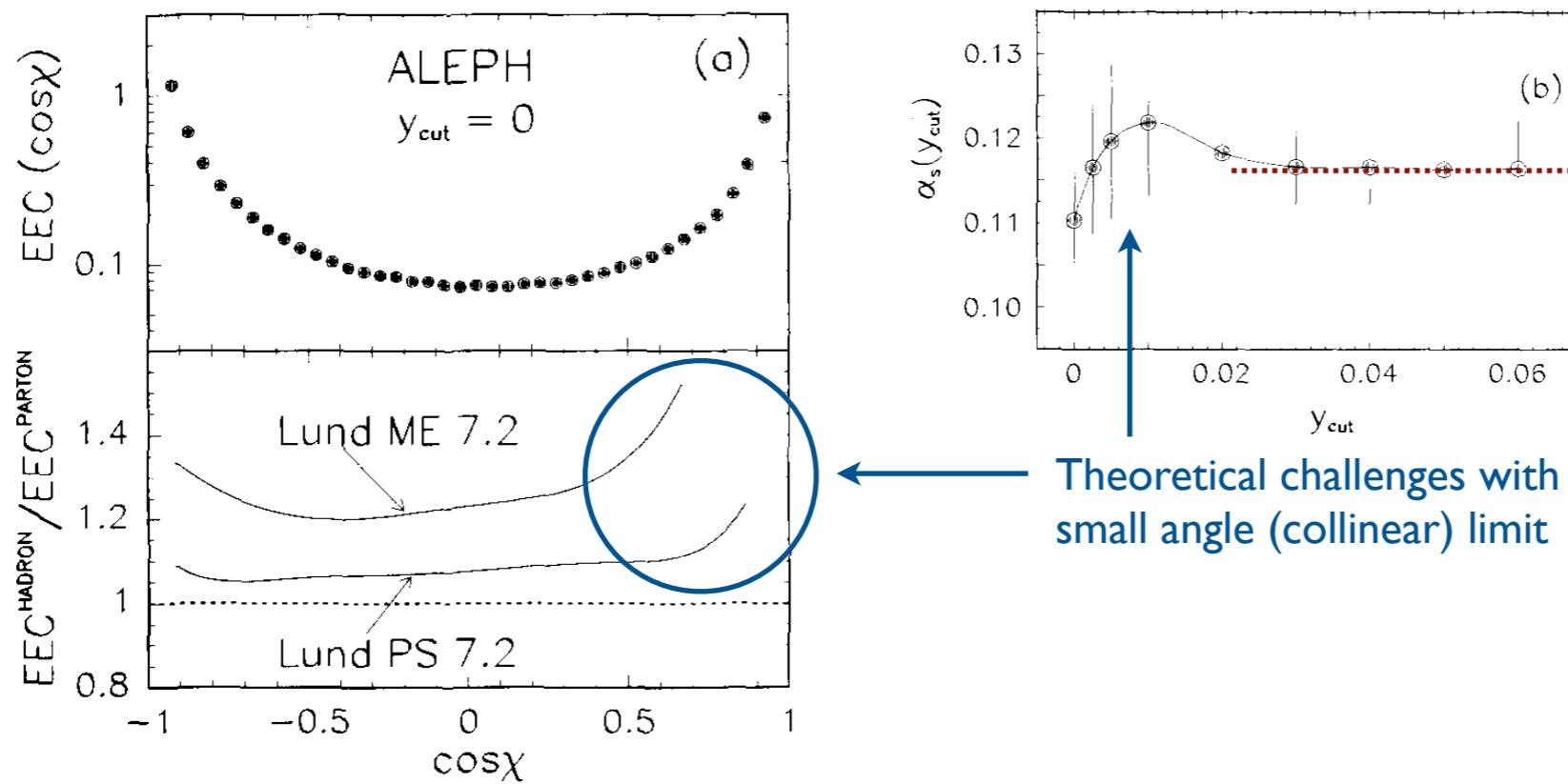
[Andreassen, Komiske, Metodiev, Nachman, JDT, [PRL 2020](#)]



Energy-Energy Correlators

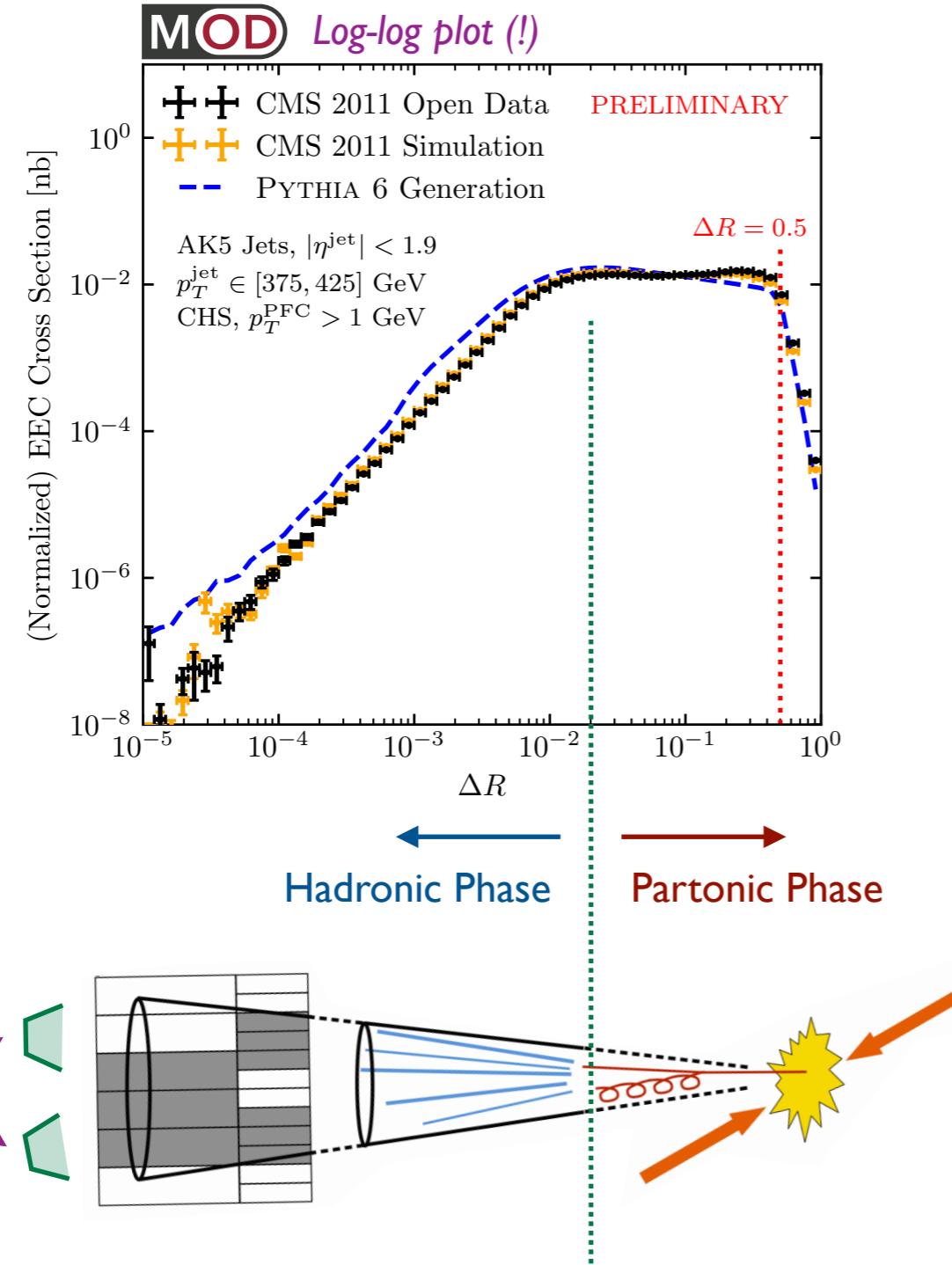


A long history in probing collinear dynamics of QCD



[Basham, Brown, Ellis, Love, [PRL 1978](#); ALEPH, [PLB 1991](#); see Chen, Moult, Zhang, Zhu, [PRD 2020](#)]

QCD Phase Transition in Jets?



Behind the scenes at **BOSTON 2019**

Jesse Thaler

EE(E)C?

To: Ian Moult

July 24, 2019 at 2:01 PM

JT

Hi Ian,

I appreciate the provocative lunch! At some point, we should talk about what you might want to see measured on the CMS Open Data with respect to EE(E)C.

Cheers,
Jesse

Jesse Thaler

Preliminary EEC plot

To: Ian Moult, Cc: Patrick Komiske

July 24, 2019 at 10:16 PM

JT

Hi Ian,

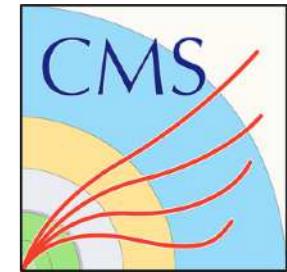
Below is a preliminary EEC plot on the CMS Open Data, made by Patrick. (Strictly speaking, this is the pT-pT-Correlator...)

We do this within single jets in the pT range specified in the plot, and the dashed vertical red line is the jet radius (beyond which things don't really make sense).

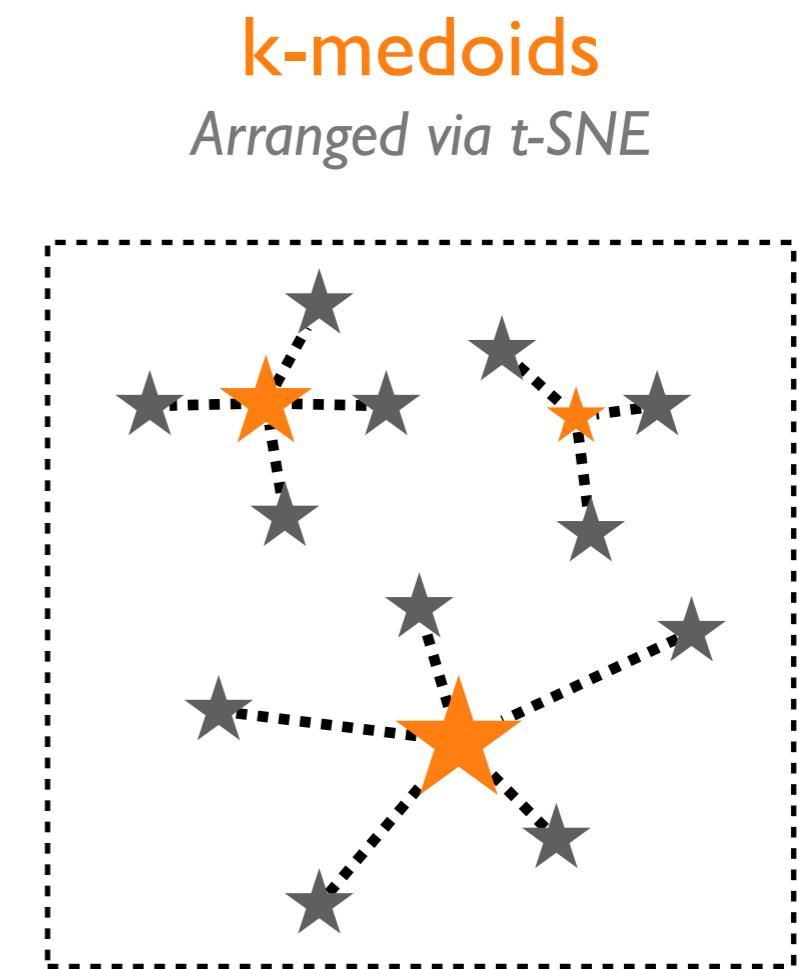
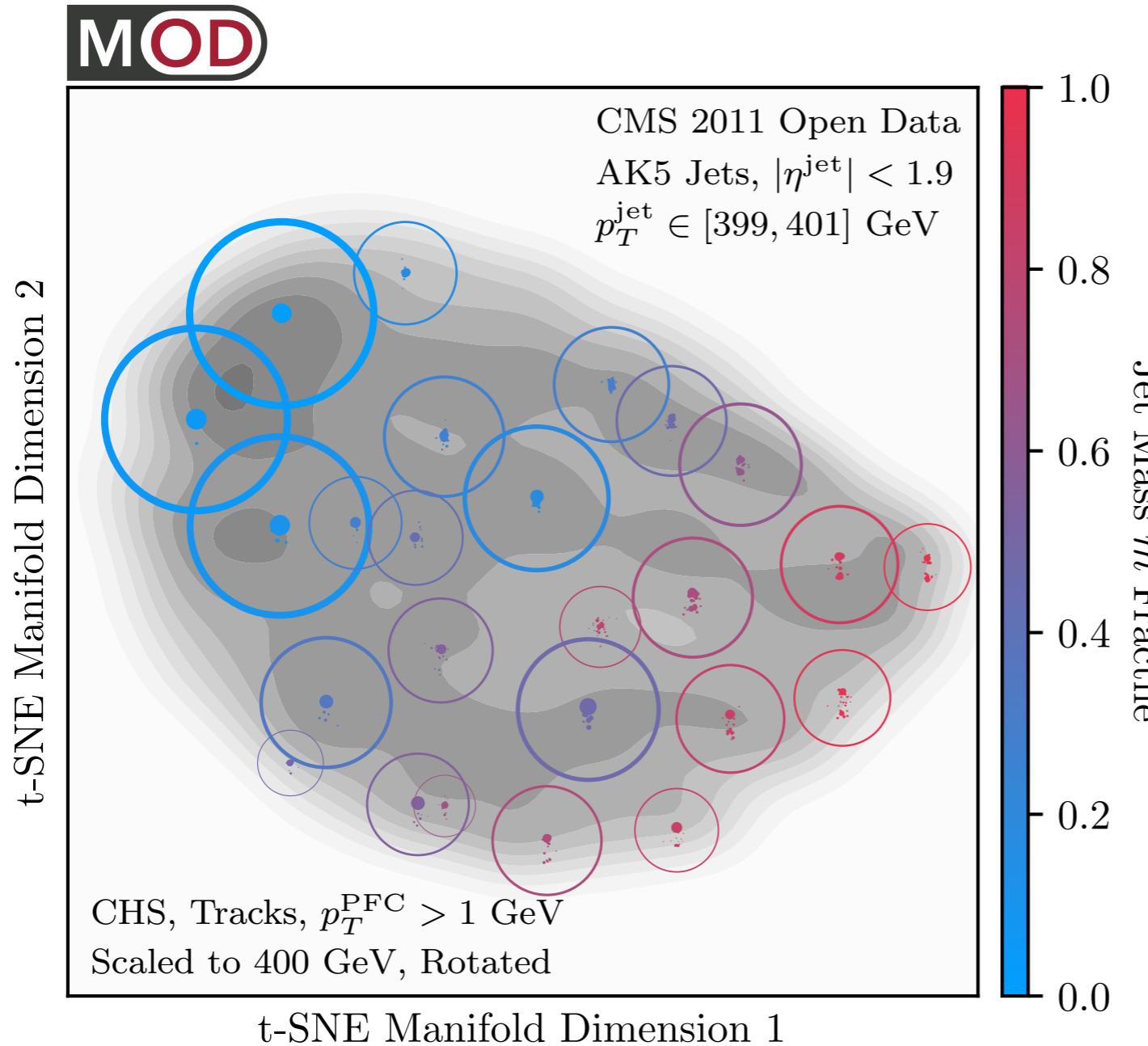
[Komiske, Moult, JDT, et al., in progress; see talks by Moult, [BOOST 2019](#), [BOOST 2020](#)]



Most Representative Jets



[<http://opendata.cern.ch/>]

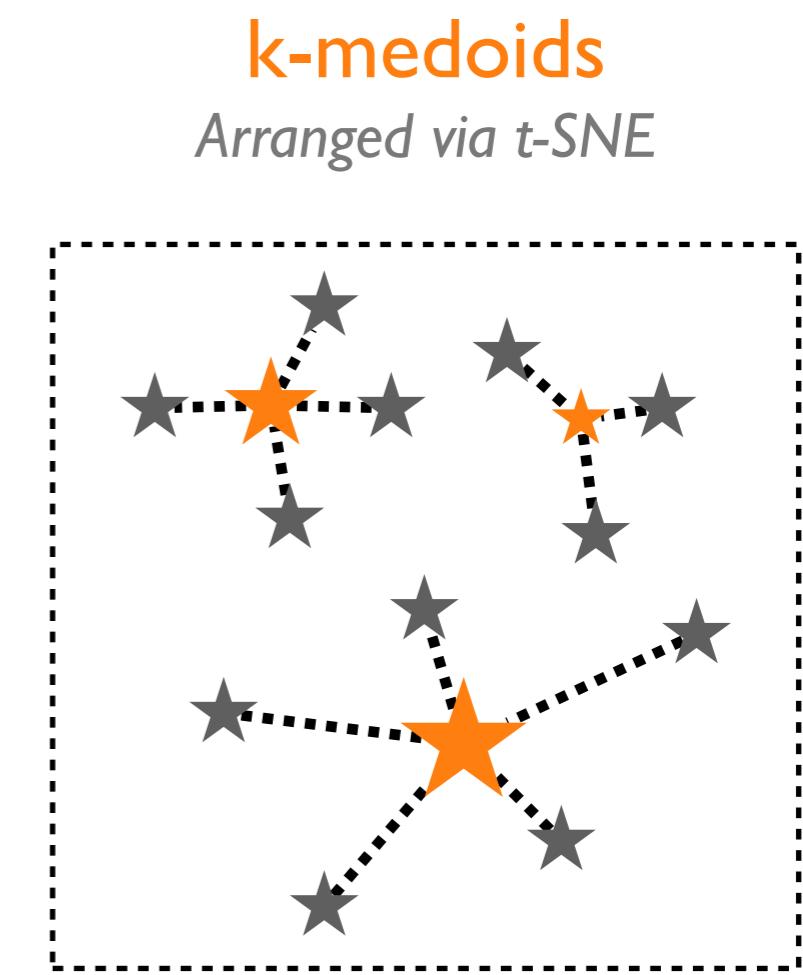
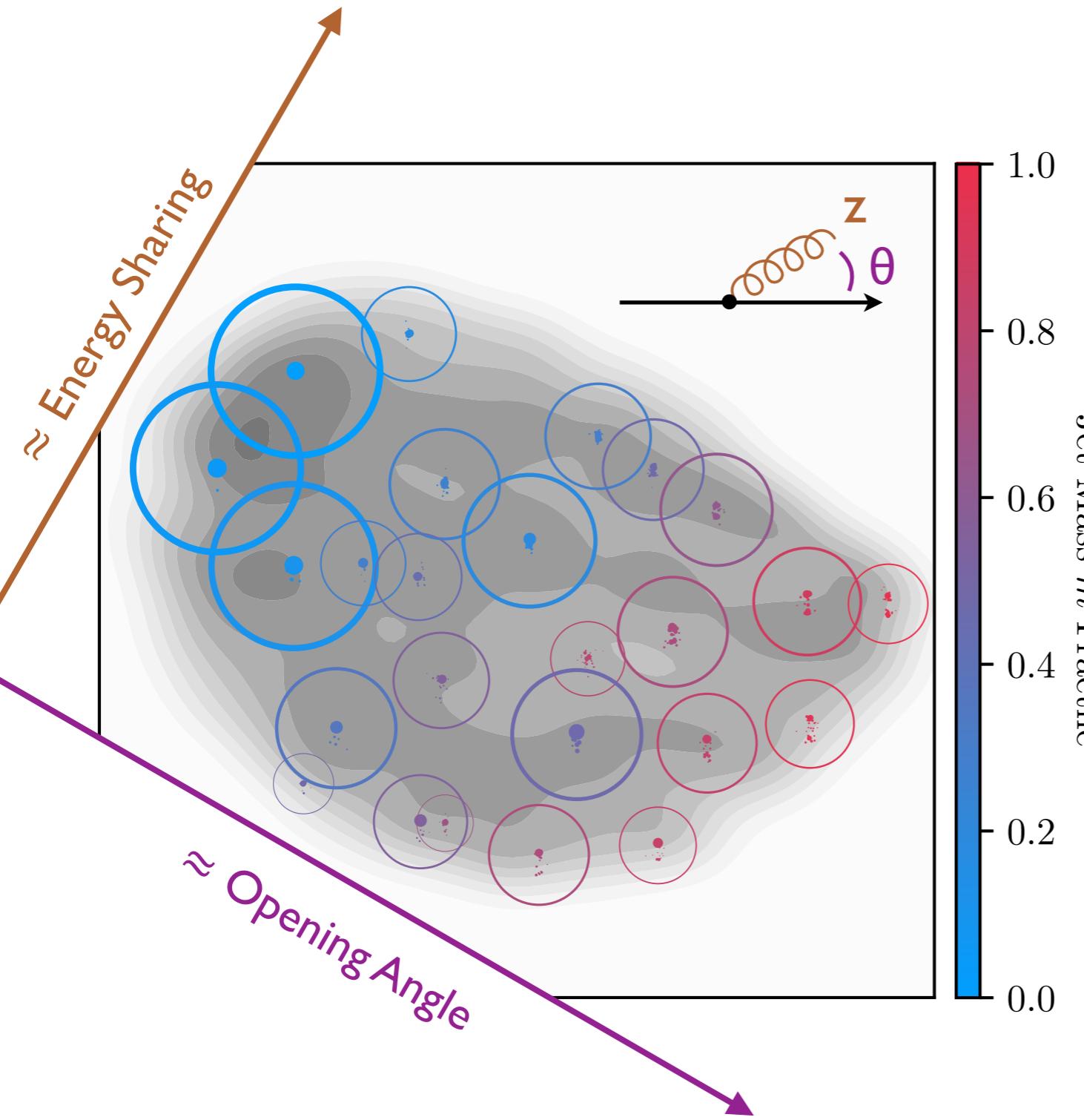


[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#); using van der Maaten, Hinton, [JMLR 2008](#)]

Most Representative Jets

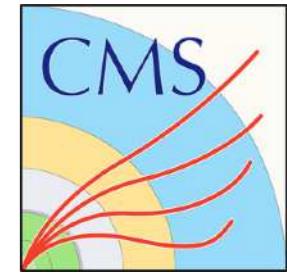


[<http://opendata.cern.ch/>]

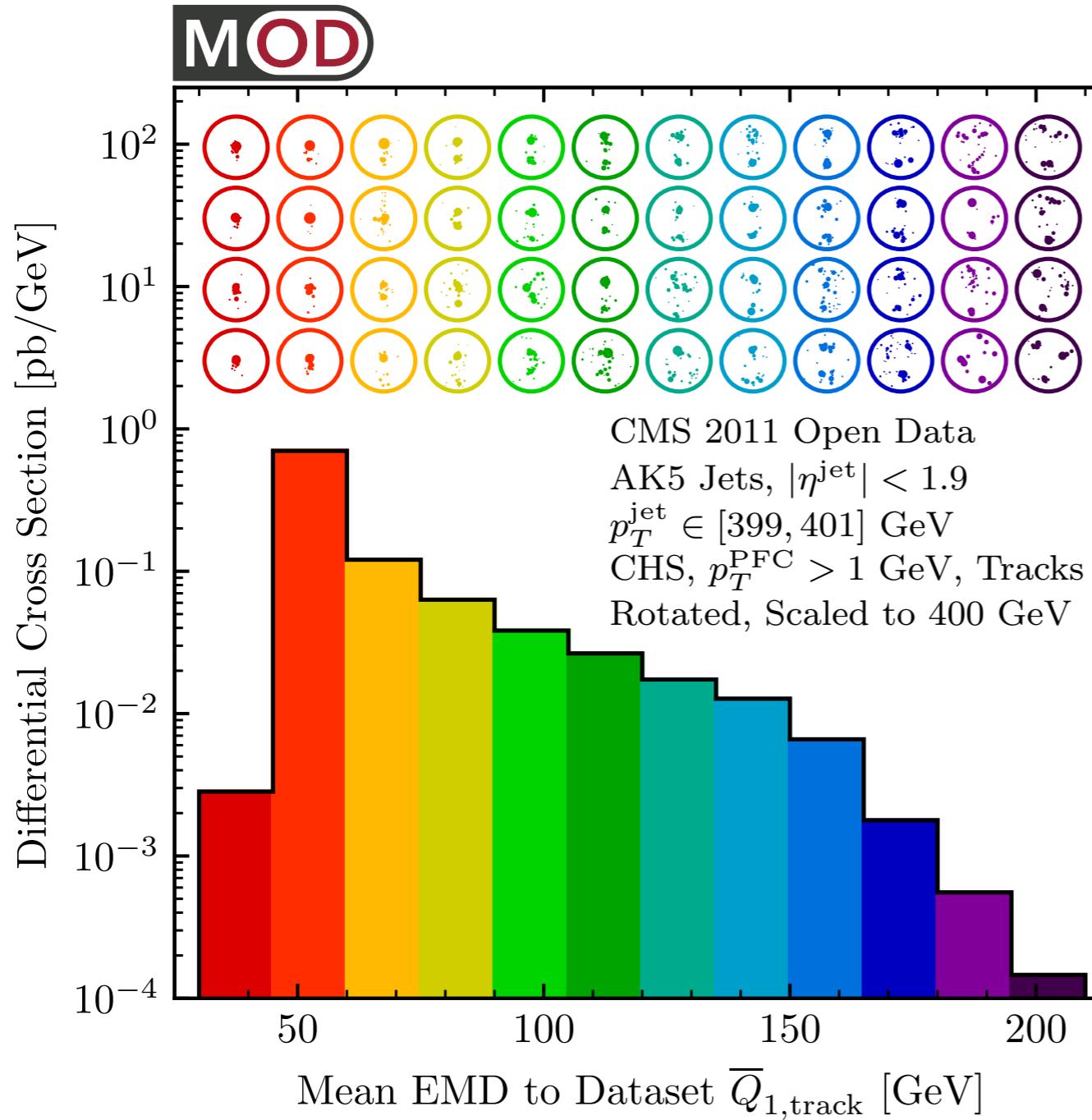


[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020; using van der Maaten, Hinton, JMLR 2008]

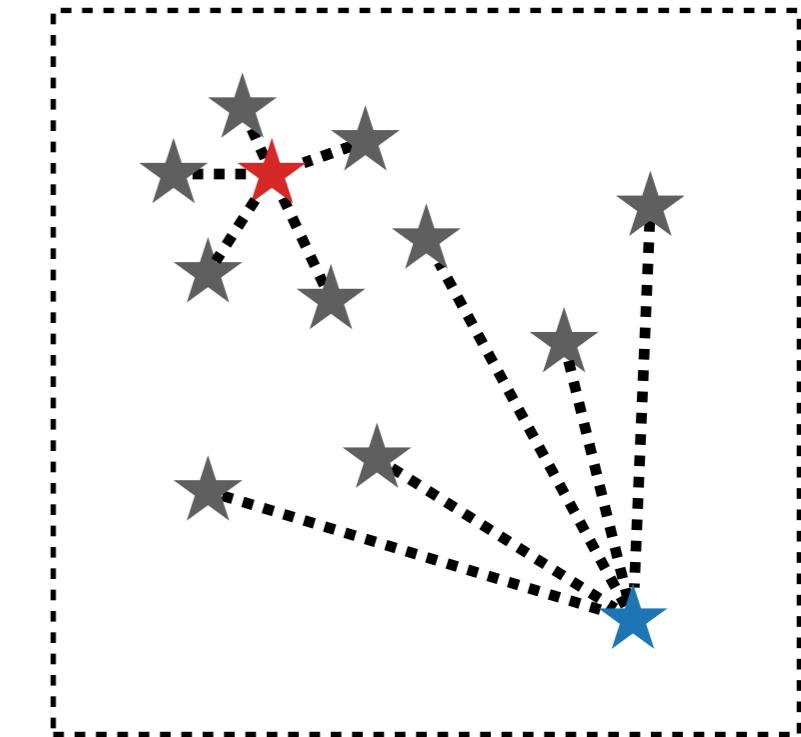
Least Representative Jets



[<http://opendata.cern.ch/>]



New Physics?
Or tails of QCD?



[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#)]

Additional Travel Documentation

