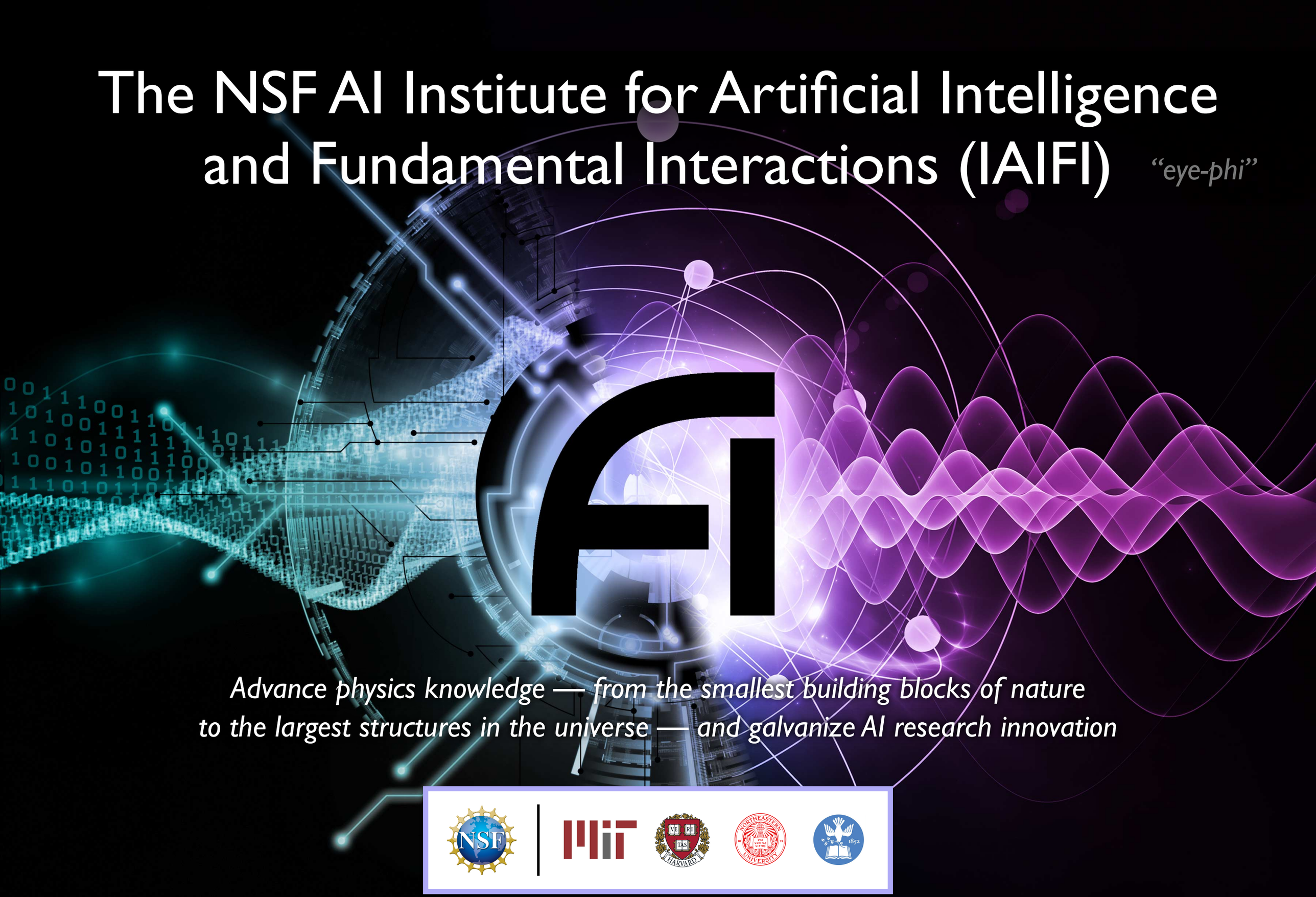
# The Hidden Geometry of Particle Collisions

Jesse Thaler

Theoretical HEP Seminar, Technical University of Munich — April 29, 2021

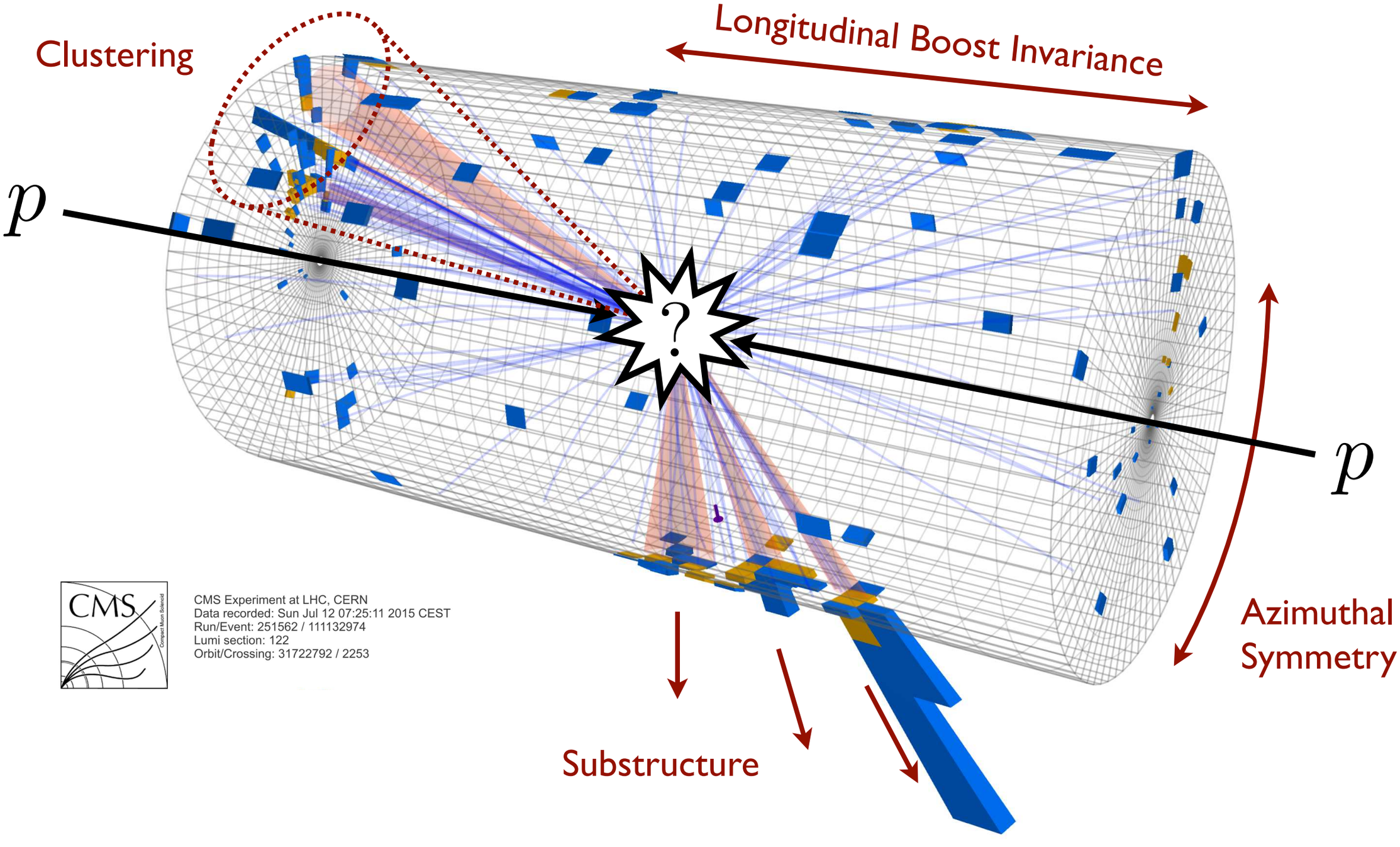
# The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI) *"eye-phi"*



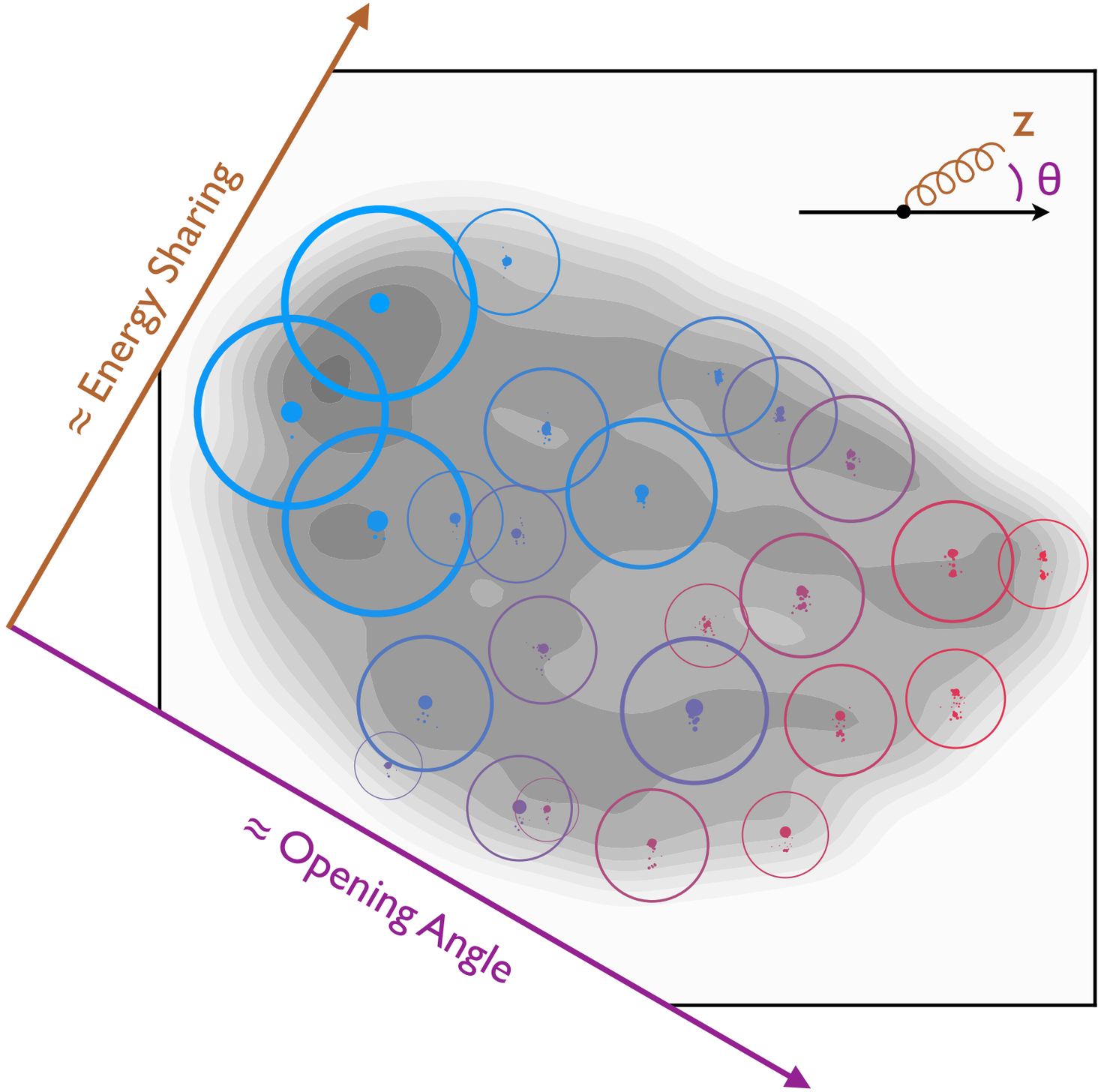

*Advance physics knowledge — from the smallest building blocks of nature to the largest structures in the universe — and galvanize AI research innovation*

[http://iaifi.org, MIT News Announcement]

# The Manifest Geometry of One Collision



# The Emergent Geometry of Many Collisions



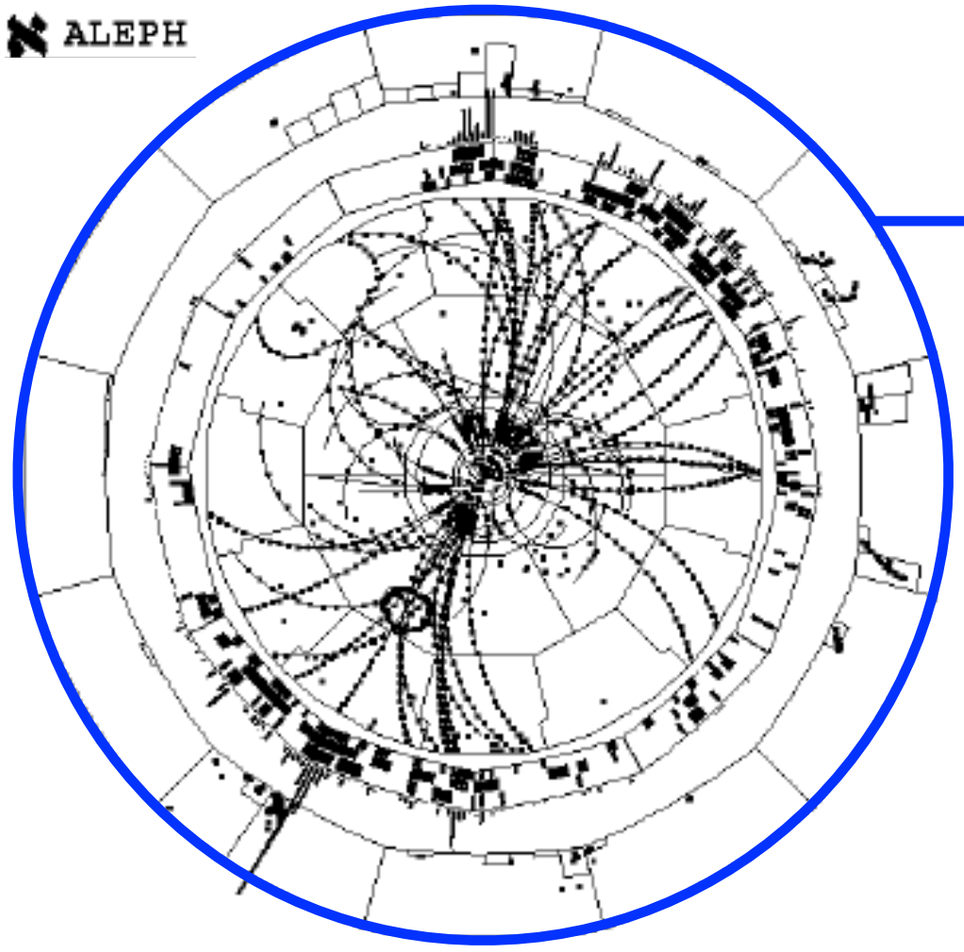




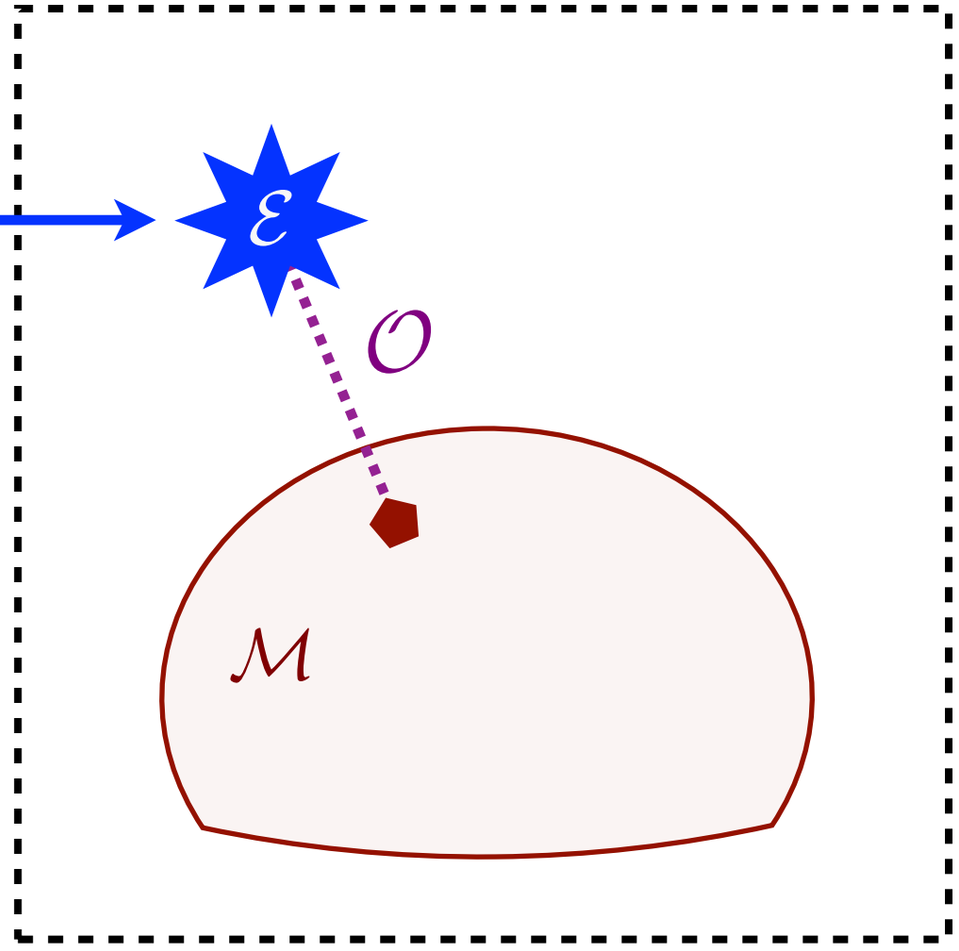
[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020;
based on Komiske, Metodiev, JDT, PRL 2019; using EnergyFlow and CMS Open Data]

# The Hidden Geometry of Particle Collisions

*E.g. Classic QCD Event Shapes*




*One Electron-Positron Event*

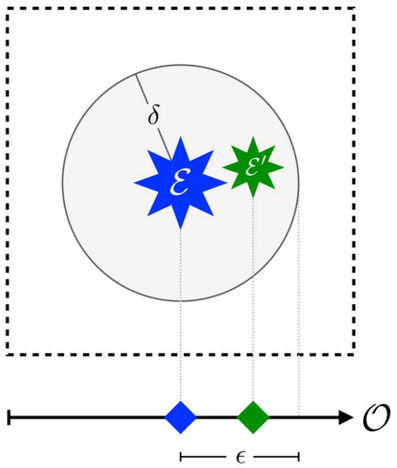



*Distance to a Manifold in Event Space*

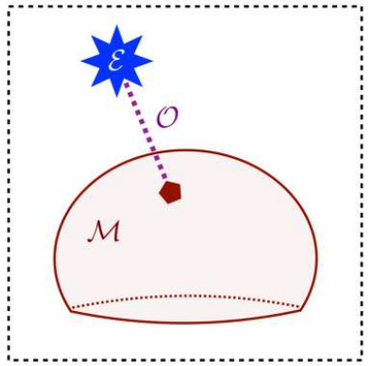
[Komiske, Metodiev, JDT, JHEP 2020]
[Brandt, Peyrou, Sosnowski, Wroblewski, PL 1964; Farhi, PRL 1977]


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

**IRC Safety** is smoothness in the space of events

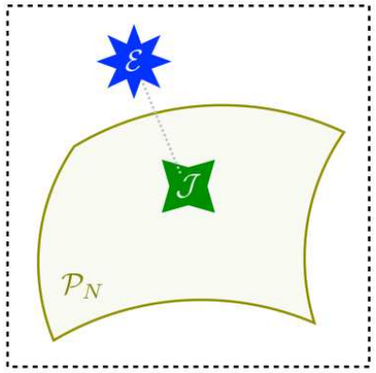



**Event shapes** are distances from events to manifolds.



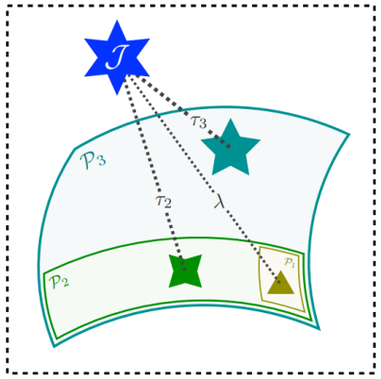

$$\mathcal{O}(\mathcal{E}) = \min_{\mathcal{E}' \in \mathcal{M}} \mathrm{EMD}_{\beta,R}(\mathcal{E}, \mathcal{E}')$$

**Jets** are projections to few-particle manifolds.



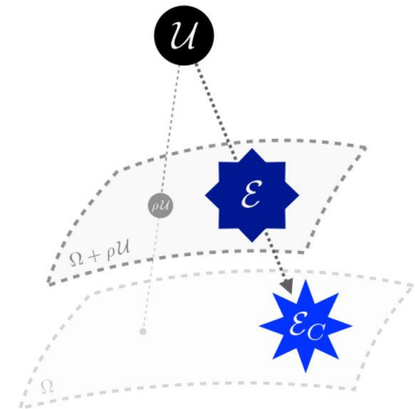

$$\mathcal{J} = \underset{\mathcal{E}' \in \mathcal{P}_N}{\mathrm{argmin}} \, \mathrm{EMD}_{\beta,R}(\mathcal{E}, \mathcal{E}')$$

**Substructure** resolves emissions within the jet.




$$\tau(\mathcal{J}) = \min_{\mathcal{E}' \in \mathcal{P}_N} \mathrm{EMD}_{\beta}(\mathcal{J}, \mathcal{E}').$$

**Pileup mitigation** moves away from uniform radiation.




$$\mathcal{E}_C = \arg\min_{\mathcal{E}'} \mathrm{EMD}(\mathcal{E}, \mathcal{E}' + \rho\,\mathcal{U}).$$

**Taming infinities**

1962-1964
Infrared Safety
[Kinoshita, JMP 1962]
[Lee, Nauenberg, PR 1964]

**Event Shapes**

1977
Thrust, Sphericity
[Farhi, PRL 1977]
[Georgi, Machacek, PRL 1977]

**Jet Algorithms**

1993
$k_T$ jet clustering
[Ellis, Soper, PRD 1993]
[Catani, Dokshitzer, Seymour, Webber, NPB 1993]

1997-1998
C/A jet clustering
[Wobisch, Wengler, 1998]
[Doskhitzer, Leder, Moretti,Webber, JHEP 1997]

**Jet Substructure**

2010-2015
N-(sub)jettiness, XCone
[Stewart, Tackmann, Waalewijn, PRL 2010]
[Thaler, Van Tilburg, JHEP 2011]
[Stewart, Tackmann, Thaler, Vermilion, Wilkason, JHEP 2015]

**Pileup**

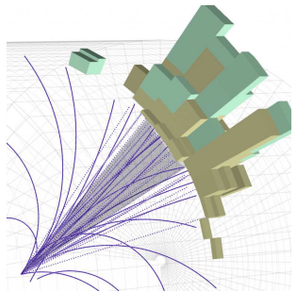
2014-2019
Constituent Subtraction
[Berta, Spousta, Miller, Leitner, JHEP 2014]
[Berta, Masetti, Miller, Spousta, JHEP 2019]

1960 — 2020

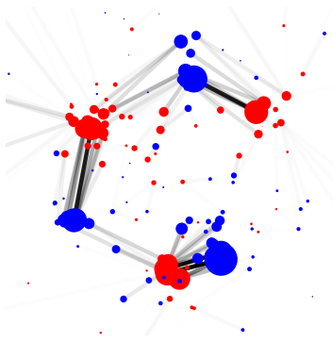
**And many more!**

[timeline from Eric Metodiev]

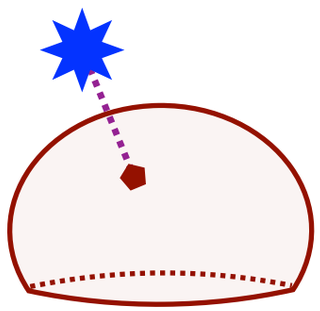
# Outline



## What is a Collider Event?



## When are Events Similar?

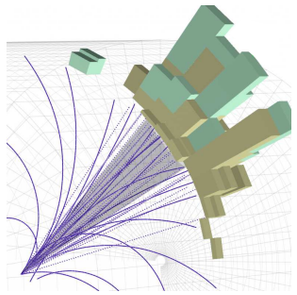


## What can be Geometrized?

# Pause

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

*Anticipating a coda…*



# What is a Collider Event?



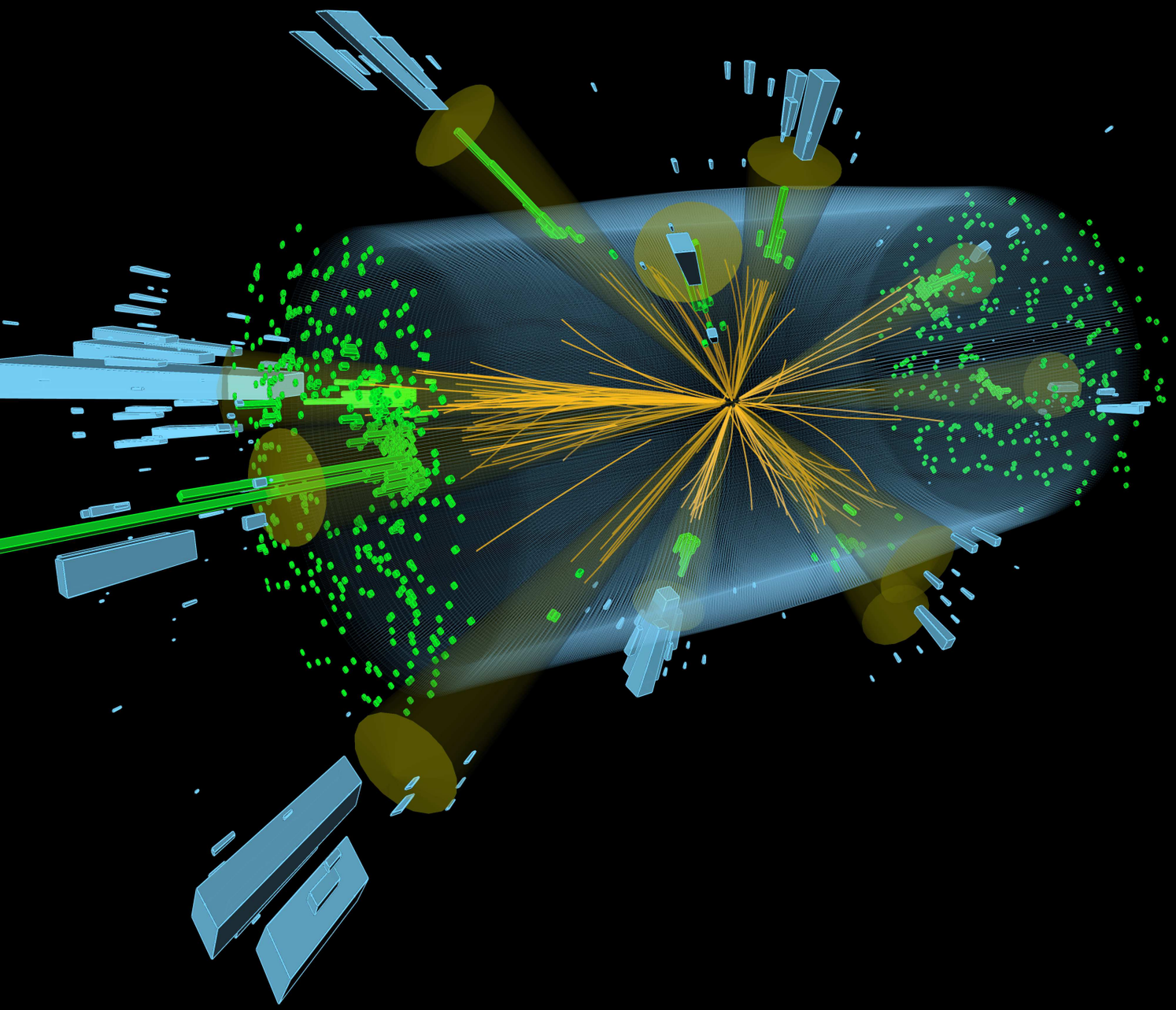
## When are Events Similar?



## What can be Geometrized?

# *Collider Event*

Collection of points in momentum space

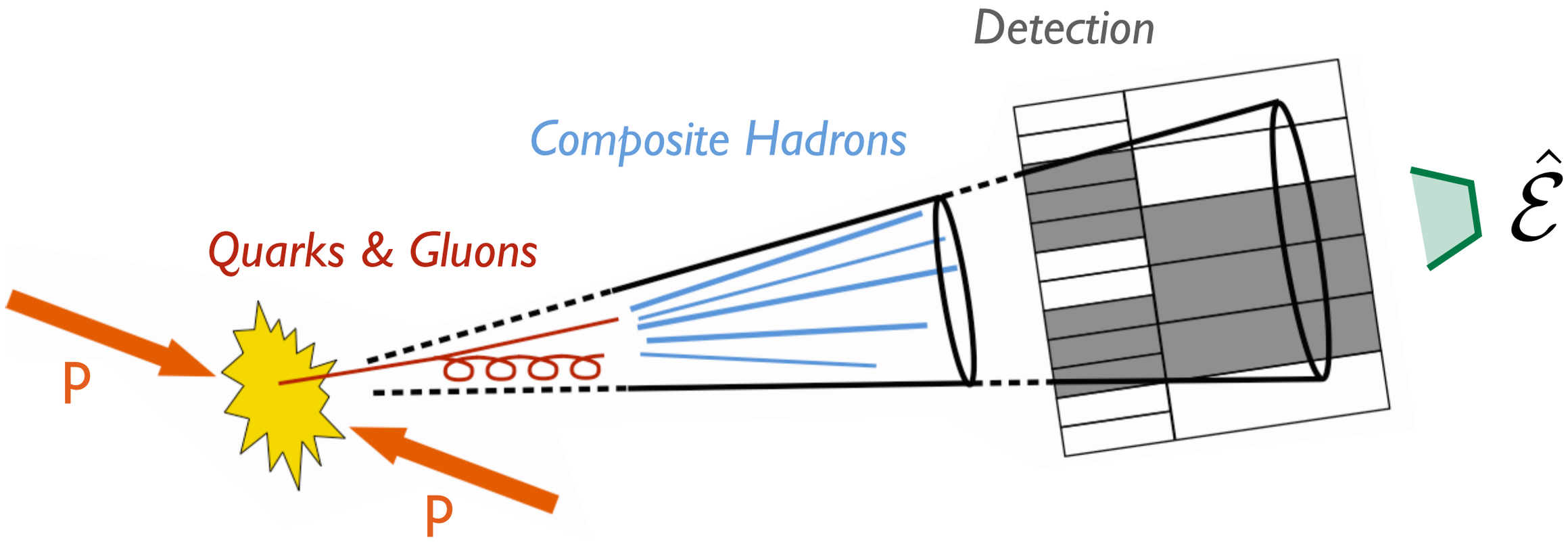


T E H M

| | | |
|---|---|---|
| $\gamma$ | photon | elementary |
| $e^{\pm}$ | electron | elementary |
| $\mu^{\pm}$ | muon | elementary |
| $\pi^{\pm}$ | pion | composite |
| $K^{\pm}$ | kaon | composite |
| $K^0_L$ | K-long | composite |
| $p/\overline{p}$ | proton | composite |
| $n/\overline{n}$ | neutron | composite |

# Jet Formation from QCD

*Theory*




## *Energy Flow:*

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

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

[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]

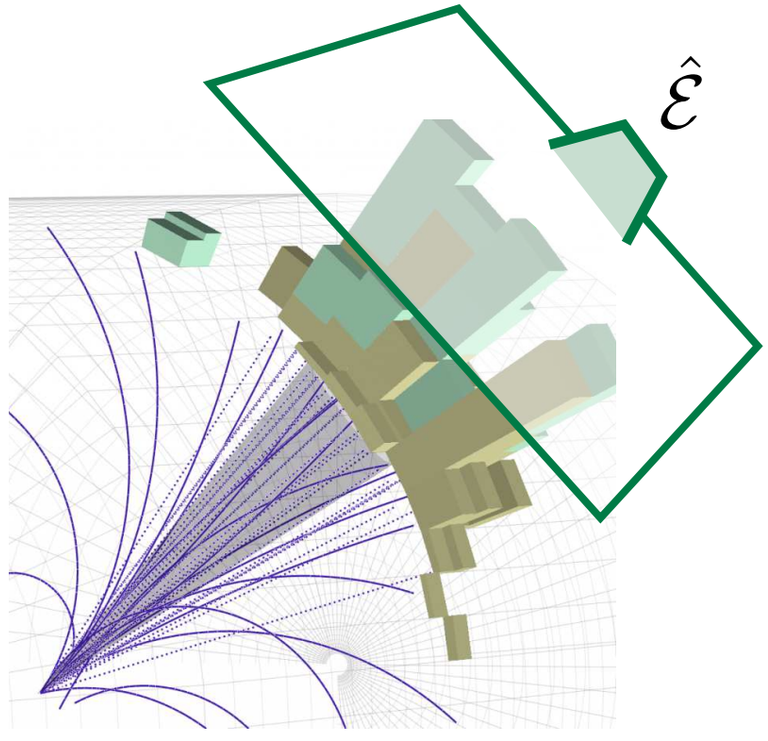
# Jets as Weighted Point Clouds

- Energy-Weighted Directions

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

Energy — Direction

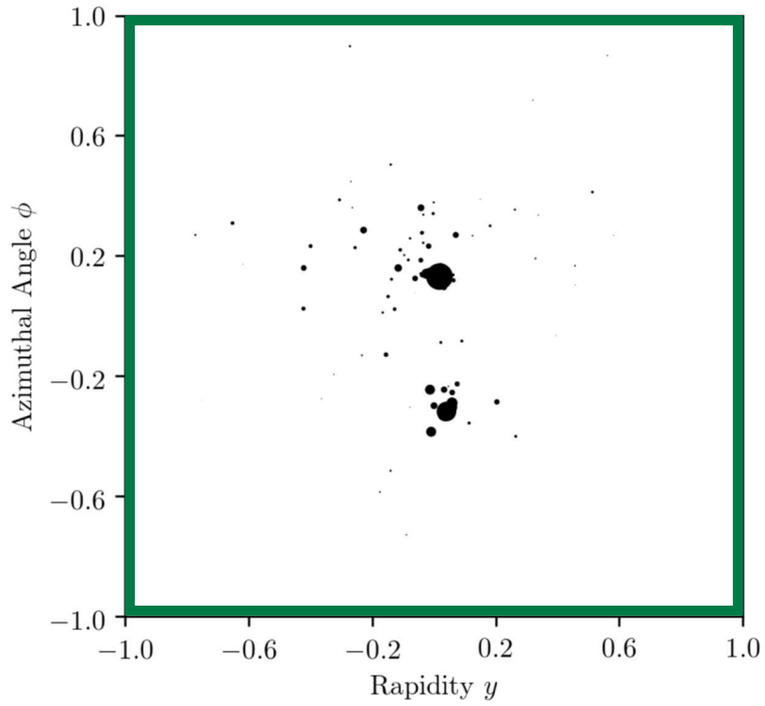
(suppressing "unsafe" charge/flavor information)




- Equivalently: Energy Density

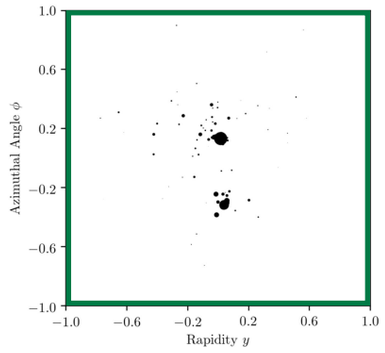
$$\rho(\hat{n}) = \sum_{i \in \mathcal{J}} E_i \, \delta^{(2)}(\hat{n} - \hat{n}_i)$$

Energy — Direction



# Pause

*What is a Collider Event?*

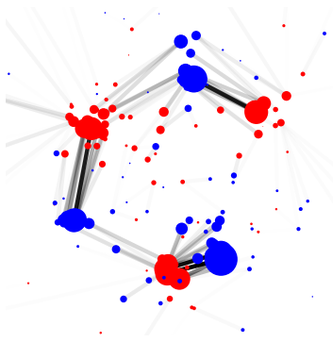



## "Calo" Energy Density

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

(see backup for relevance to QCD & ML)

# What is a Collider Event?



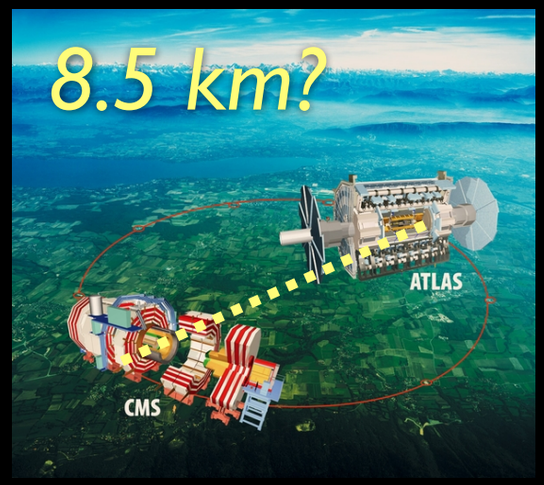
# When are Events Similar?



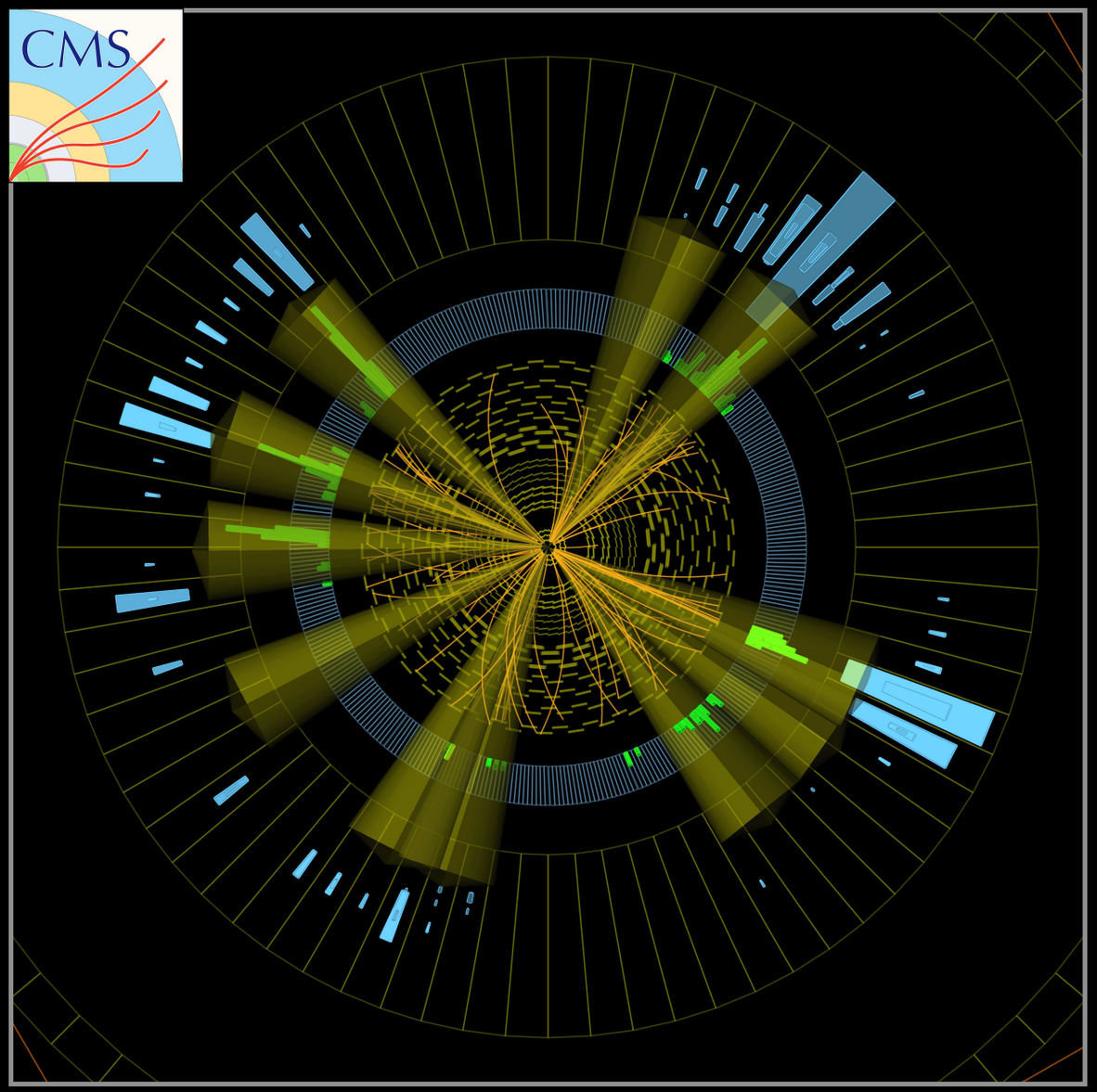
# What can be Geometrized?

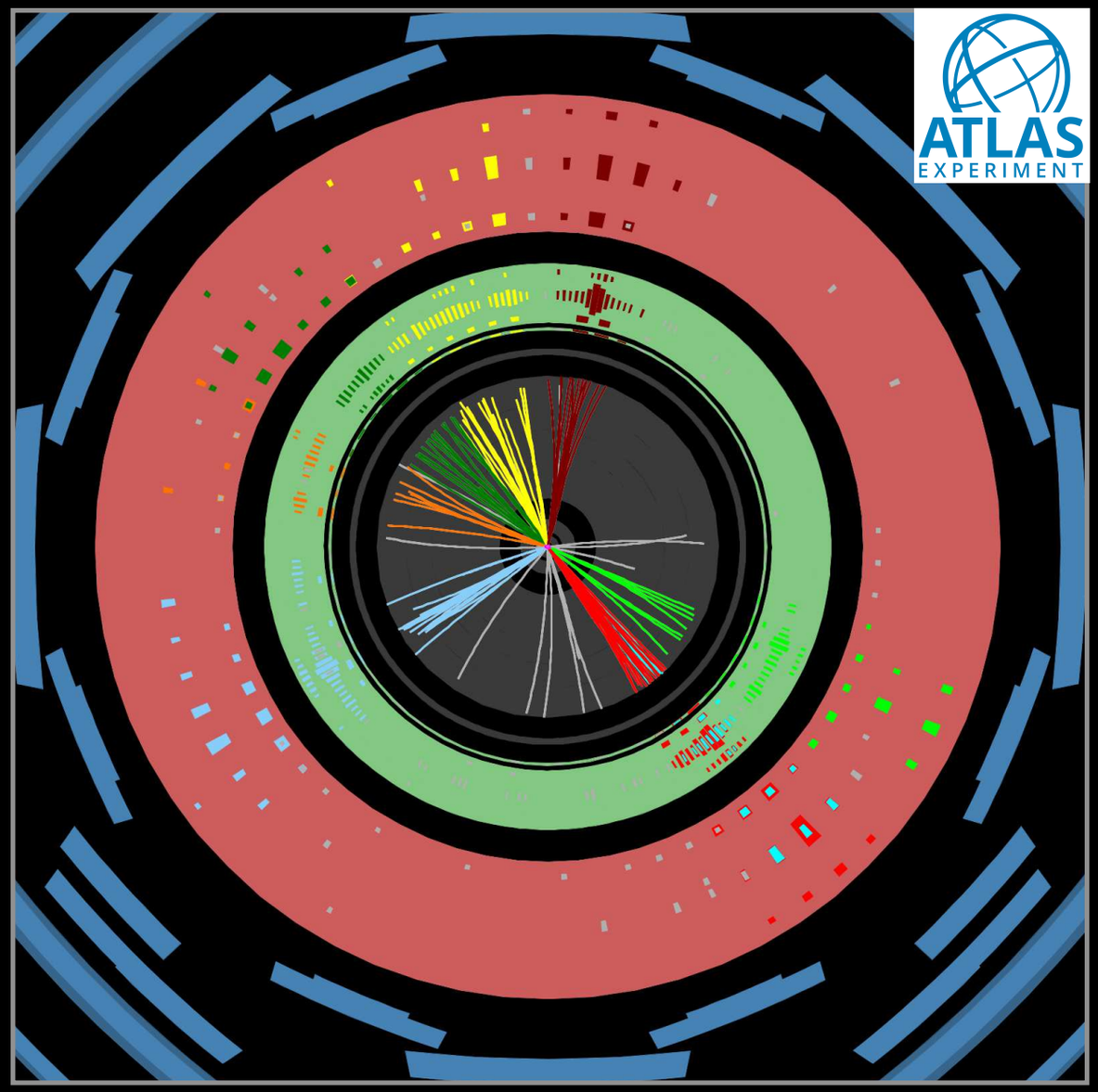
# *Two Collider Events*

Two collections of points in (momentum) space




*How "close" are these?*



# The Earth Mover's Distance

## Optimal Transport:

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

Minimum "work" (stuff x distance) to make one distribution look like another distribution

Déblai

Remblai

[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Vaserštein, 1969; Wikipedia]

# The Earth Mover's Distance

Optimal Transport:

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

Minimum "work" (stuff x distance) to make one distribution look like another distribution

Additive Weight:

Ground Distance:

$\Rightarrow$ Distance Between Distributions


[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Vaseršteĭn, 1969; Wikipedia]

# The Earth Mover's Distance

Optimal Transport:

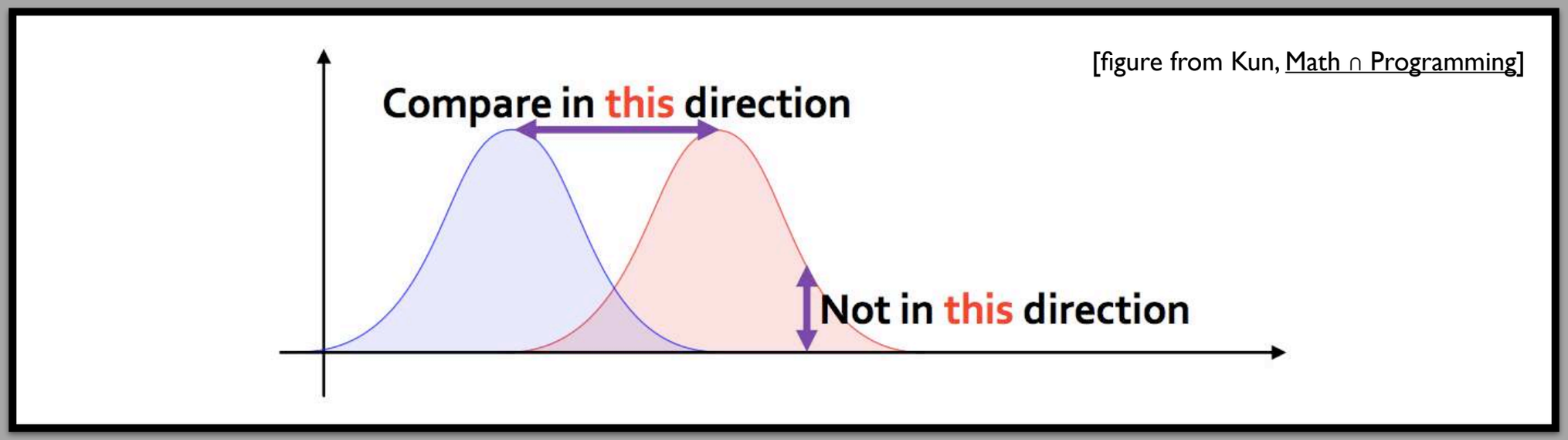
[Peleg, Werman, Rom, IEEE 1989;
Rubner, Tomasi, Guibas, ICCV 1998, ICJV 2000;
Pele, Werman, ECCV 2008; Pele Taskar, GSI 2013]

Minimum "work" (stuff x distance) to make one distribution look like another distribution

Additive Weight: ●

Ground Distance: ⟷

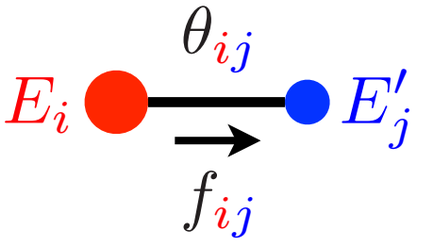
$\Rightarrow$ Distance Between Distributions

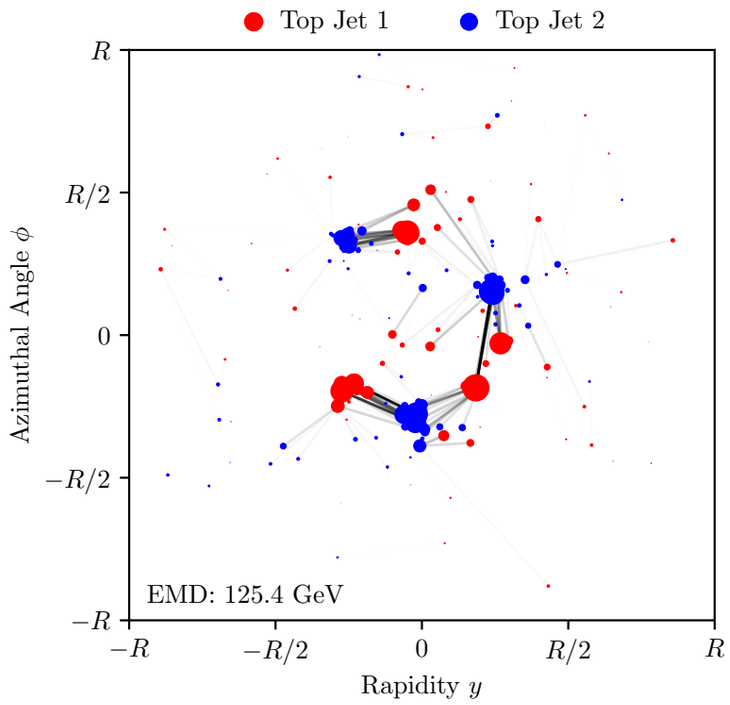


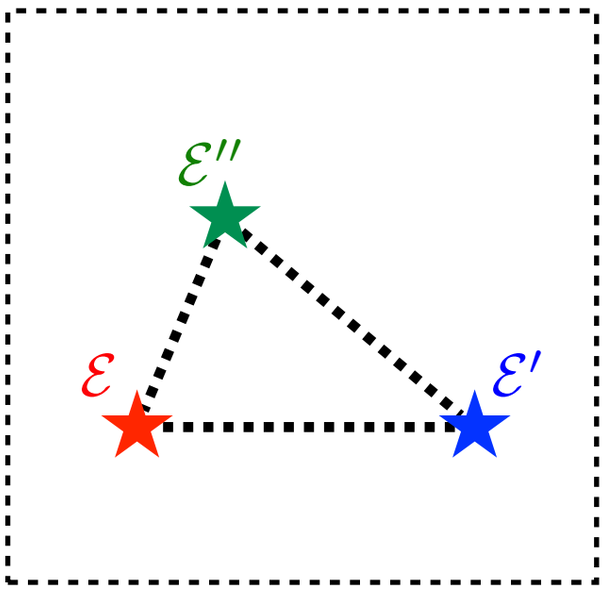

[figure from Kun, Math ∩ Programming]

[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Vaserštein, 1969; Wikipedia]

# The Energy Mover's Distance




## Optimal transport between energy flows…

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

in GeV

Cost to move energy

Cost to create energy

## …defines a metric on the space of events

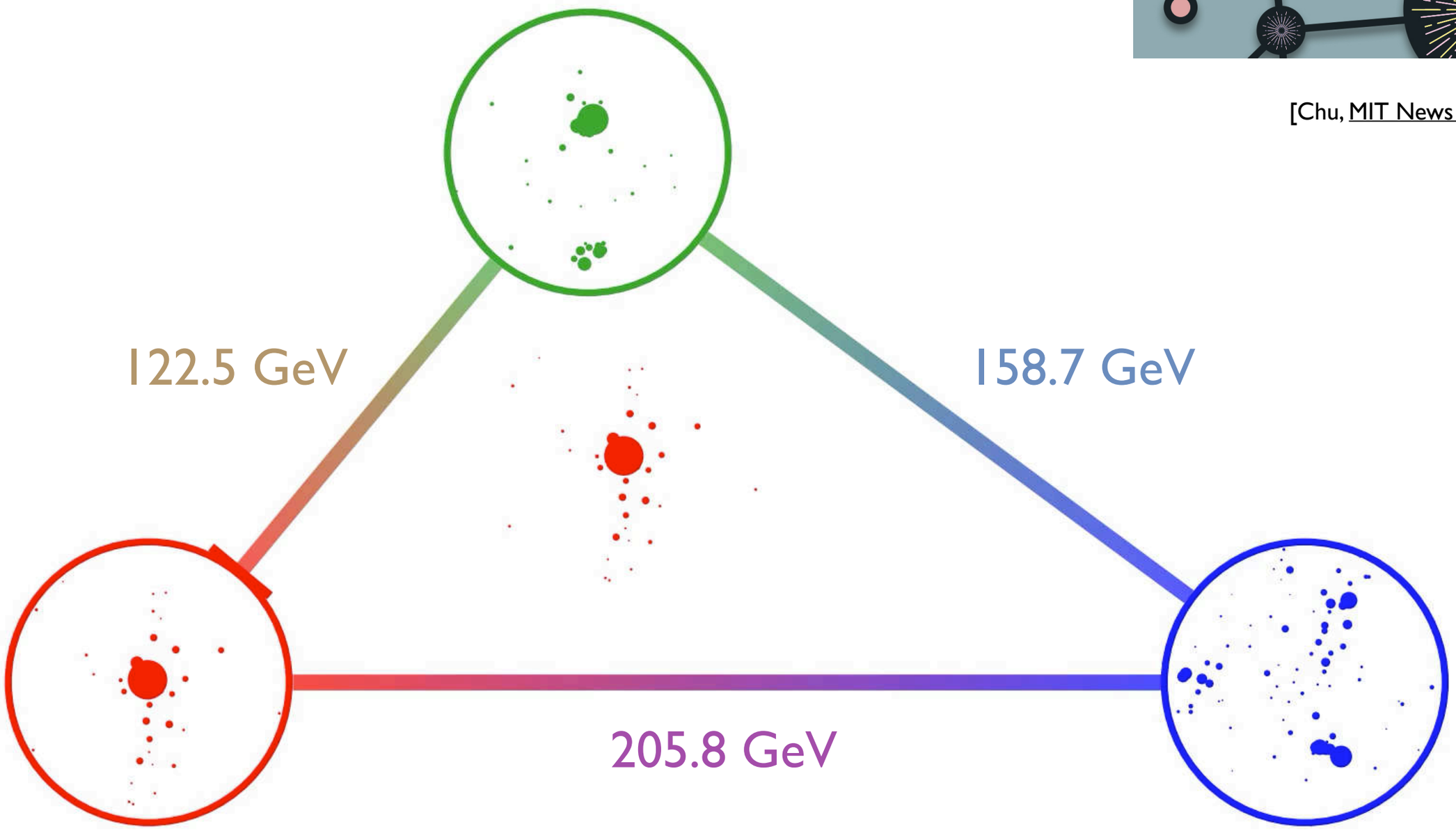
$$0 \leq \mathrm{EMD}(\mathcal{E}, \mathcal{E}') \leq \mathrm{EMD}(\mathcal{E}, \mathcal{E}'') + \mathrm{EMD}(\mathcal{E}', \mathcal{E}'')$$

(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

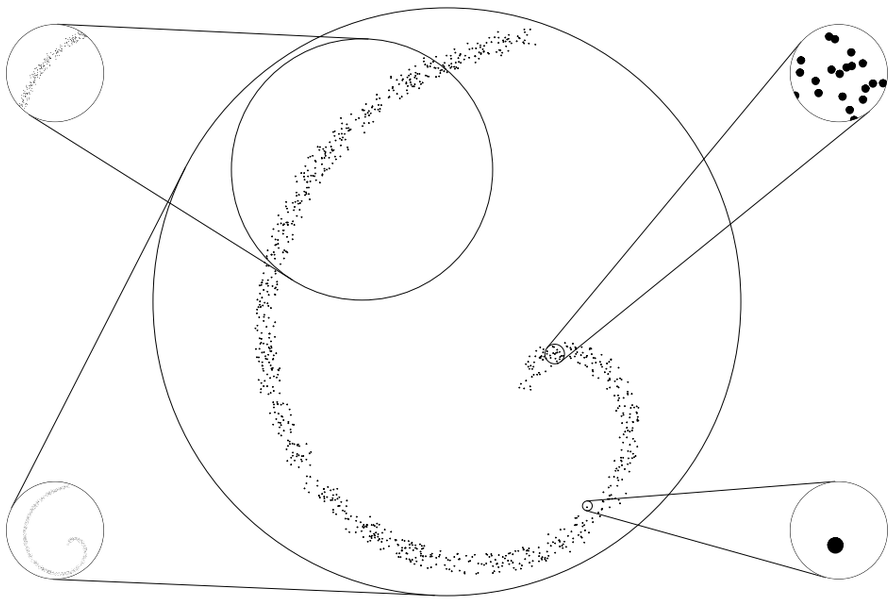
$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

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

[Grassberger, Procaccia, PRL 1983; Kégl, NIPS 2002]

dim≃1

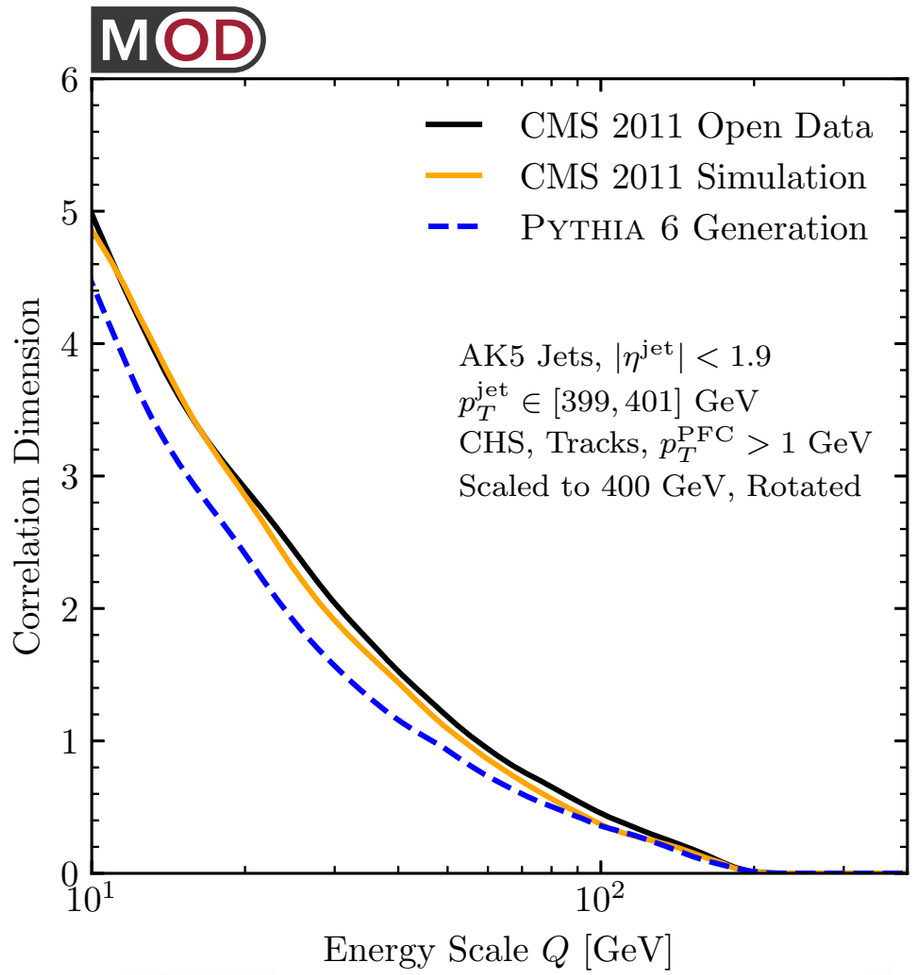
dim≃2

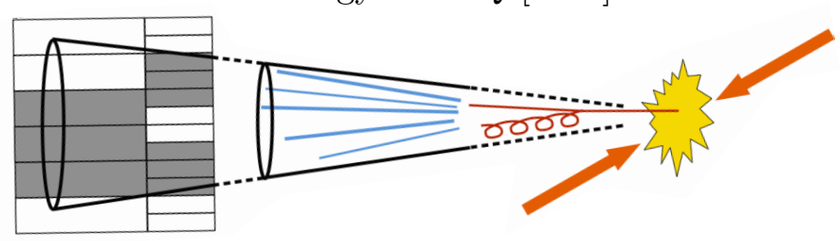


dim≃2

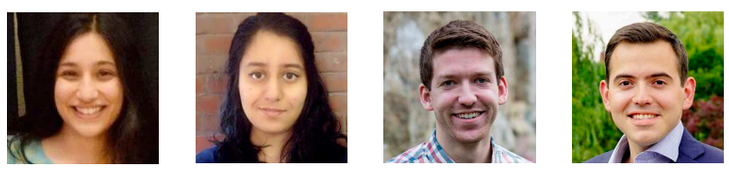
(eventually 0)

dim≃0



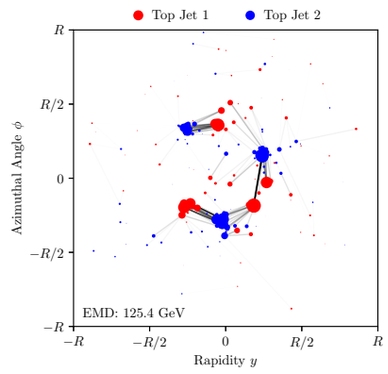



[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020; using CMS Open Data]



# Pause

*When are Events Similar?*




## Small Energy Mover's Distance

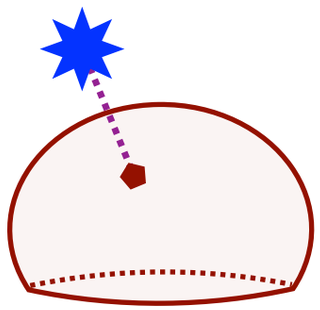
$$\mathrm{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_i \sum_j f_{ij} \frac{\theta_{ij}}{R} + \left| \sum_i E_i - \sum_j E'_j \right|$$

(see backup for more applications)

What is a Collider Event?



When are Events Similar?

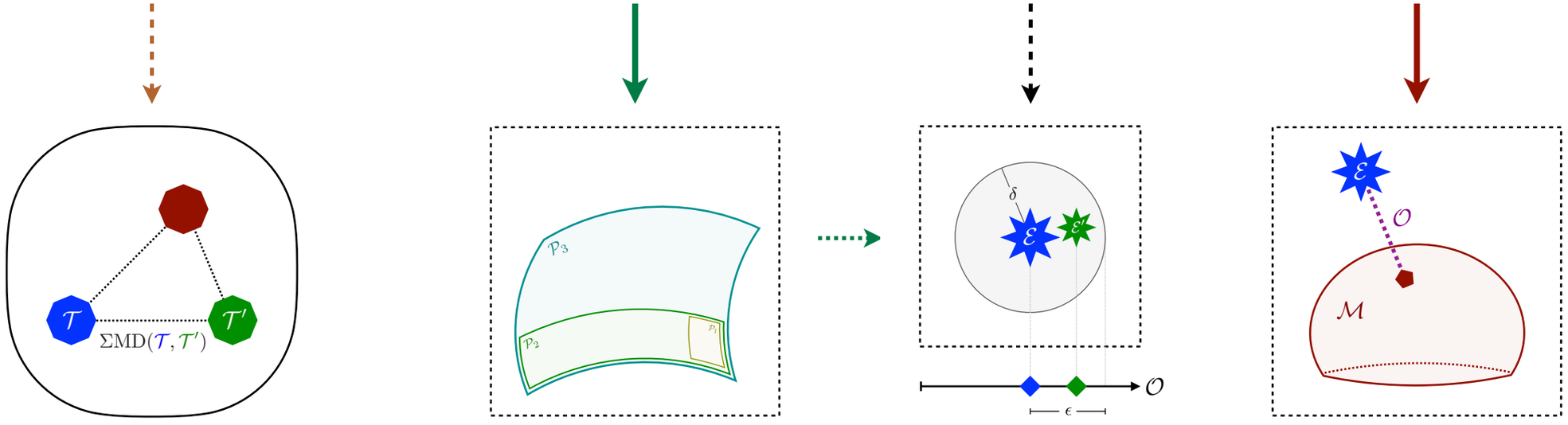


What can be Geometrized?

# Master Formula for Collider Physics

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

cross section | phase space | amplitude | observable

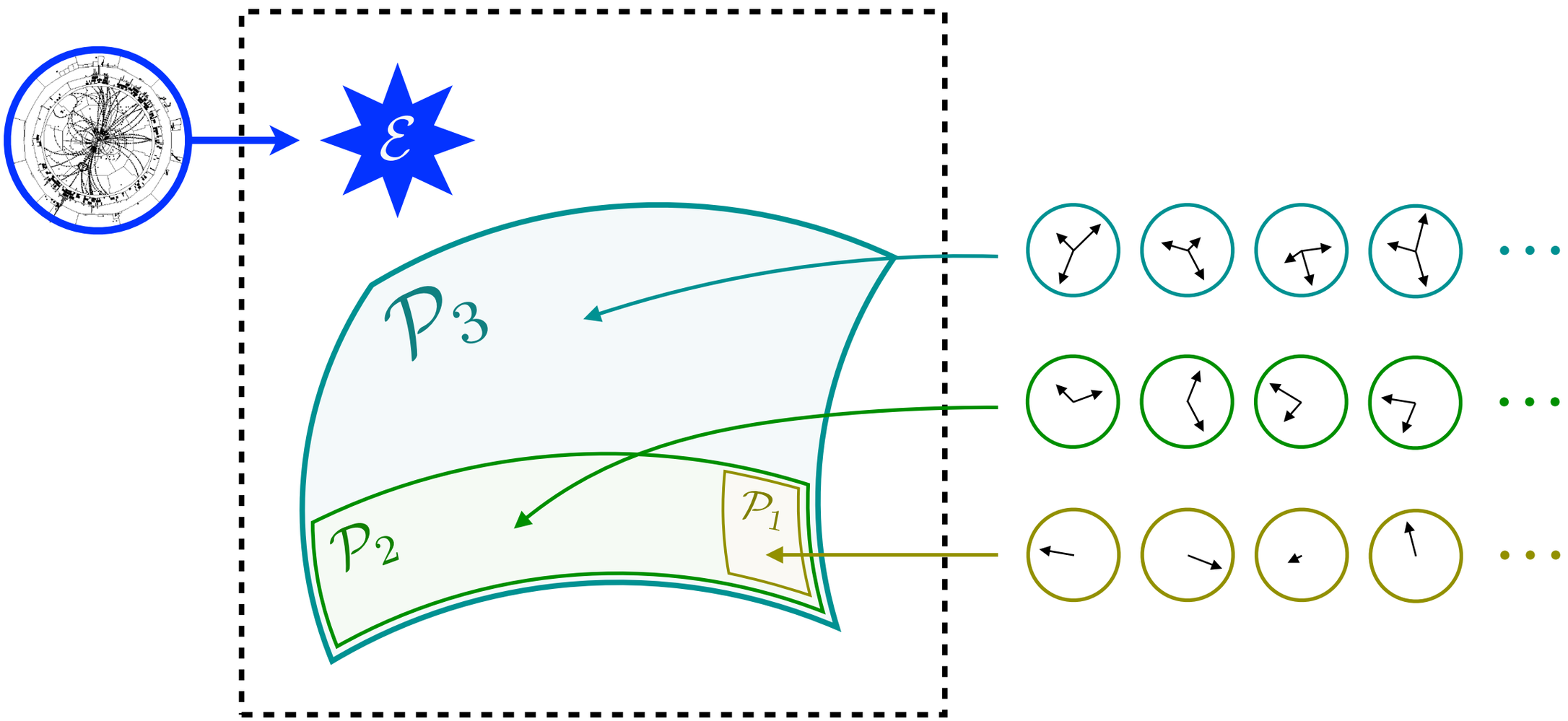



[Komiske, Metodiev, JDT, JHEP 2020]

# Introducing N-particle Manifolds

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

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



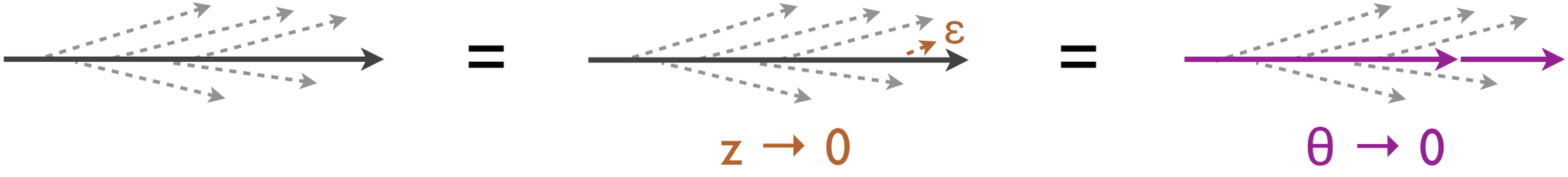

$\mathcal{P}_N \supset \mathcal{P}_{N-1} \supset \ldots \supset \mathcal{P}_2 \supset \mathcal{P}_1$ by soft/collinear limits

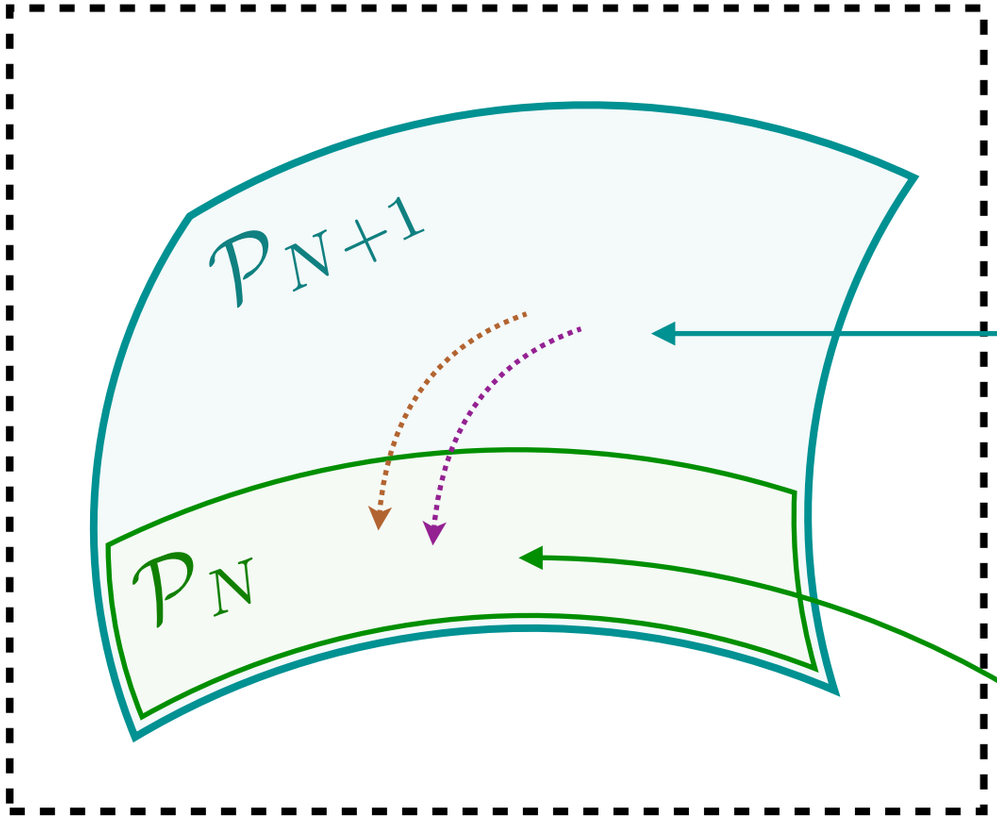
[see related discussion in Larkoski, Melia, PRD 2020]

# When are Two Events the Same?

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

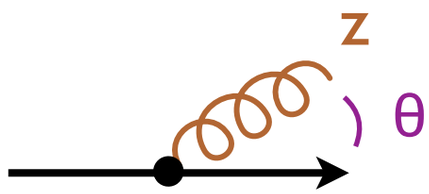
*Energy Flow unchanged by infinitesimal soft/collinear emissions*




Infrared divergences "live" together!

**Real:**



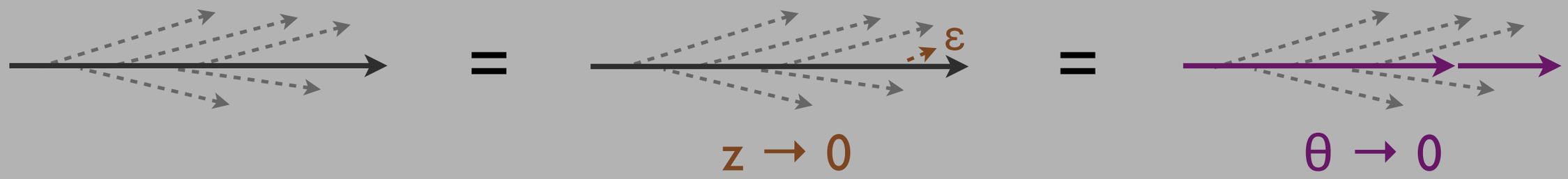

$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \, \frac{\mathrm{d}z}{z} \, \frac{\mathrm{d}\theta}{\theta}$$

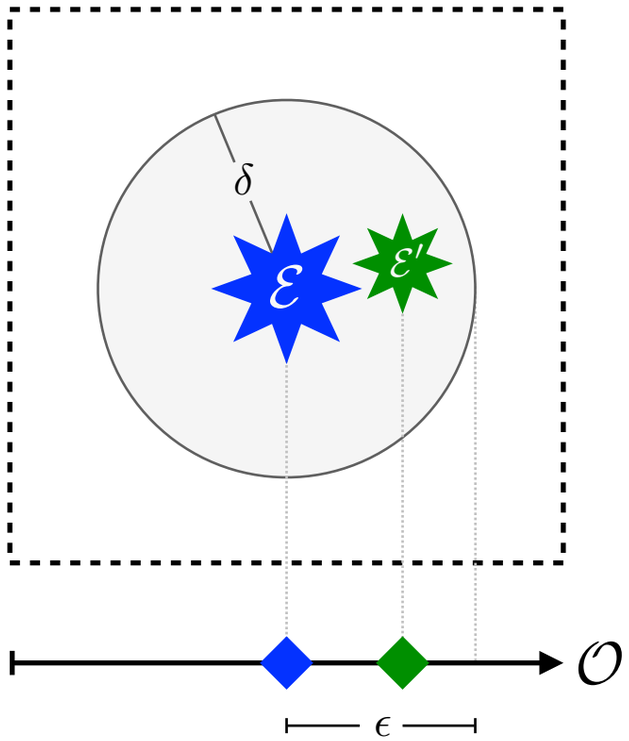
**Virtual:**

# 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 & Collinear Safety

≈ calculable in perturbative quantum field theory

*is** ← (see backup for subtleties)

## Continuity in EMD Space

[Komiske, Metodiev, JDT, JHEP 2020]
[Sterman, Weinberg, PRL 1977; Sterman, PRD 1979]
[see also Banfi, Salam, Zanderighi, JHEP 2005; Larkoski, Marzani, JDT, PRD 2015]

# *EMD seems to define the "natural" geometry for massless gauge theories*

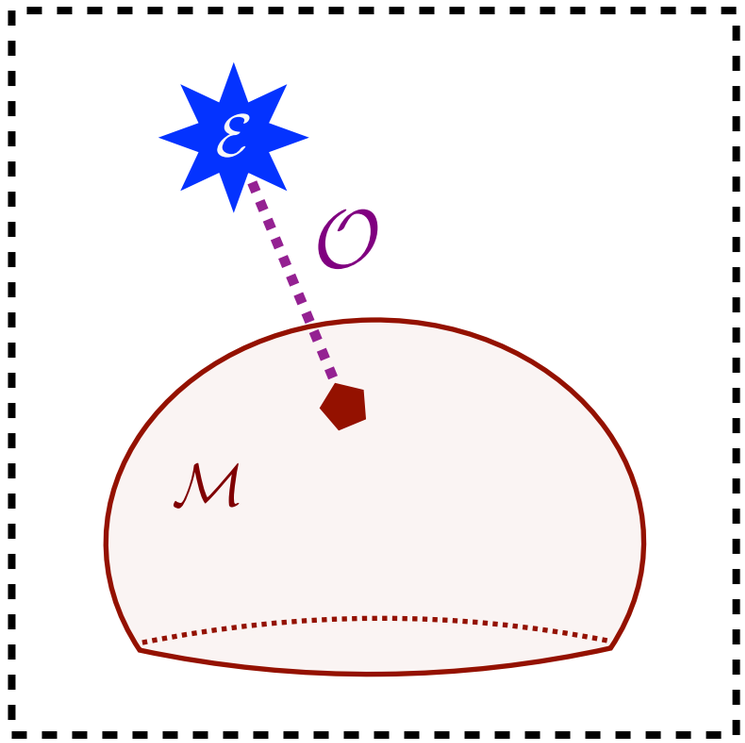
Open question: Can you define $\left|\mathcal{M}_{AB\to 12\ldots n}\right|^2$ directly in this space?

What does it mean to "integrate" in this space?

# Manifolds for Observables

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

One Event




Set of Events

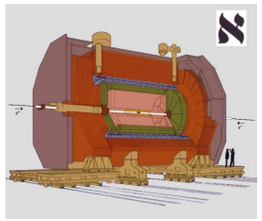
Distance of Closest Approach ⇒ Observable

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

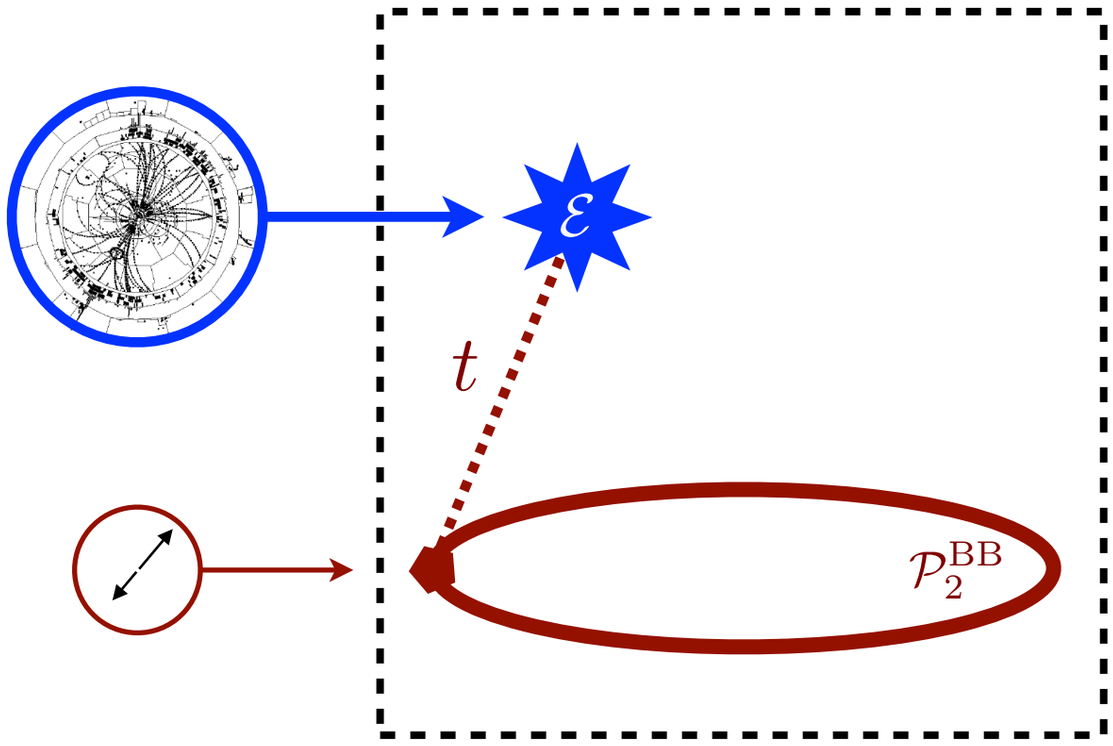
[Komiske, Metodiev, JDT, JHEP 2020]

# E.g. Thrust

*How dijet-like is an event?*



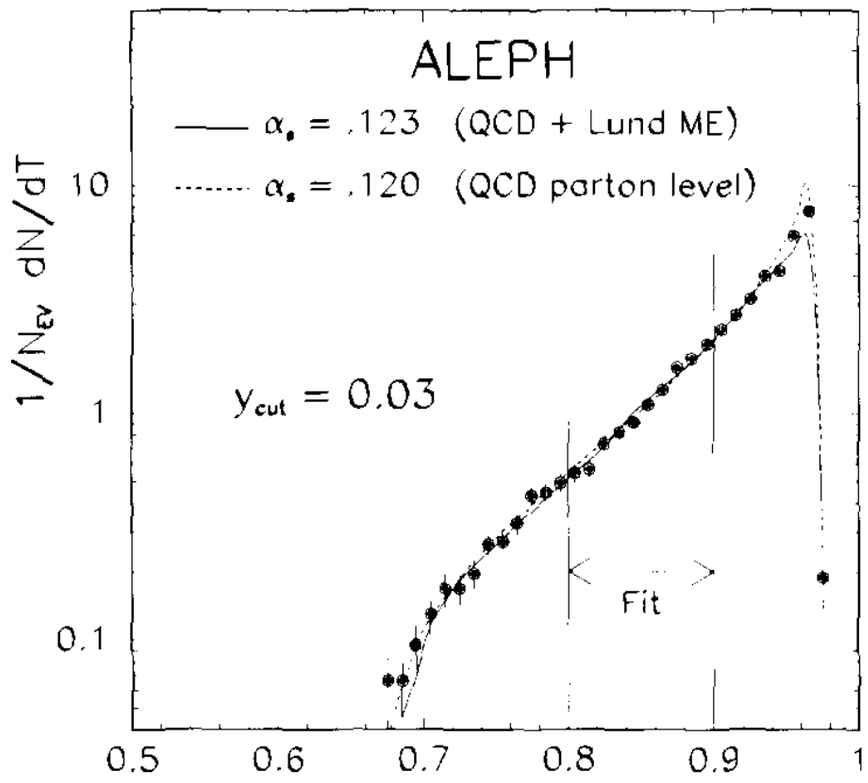
$$t(\mathcal{E}) = \min_{\mathcal{E}' \in \mathcal{P}_2^{\text{BB}}} \text{EMD}_2(\mathcal{E}, \mathcal{E}')$$




All Back-to-Back Two Particle Configurations

$$\mathcal{P}_2^{\text{BB}} = \left\{ \quad \cdots \right\}$$

(using β=2 EMD variant)



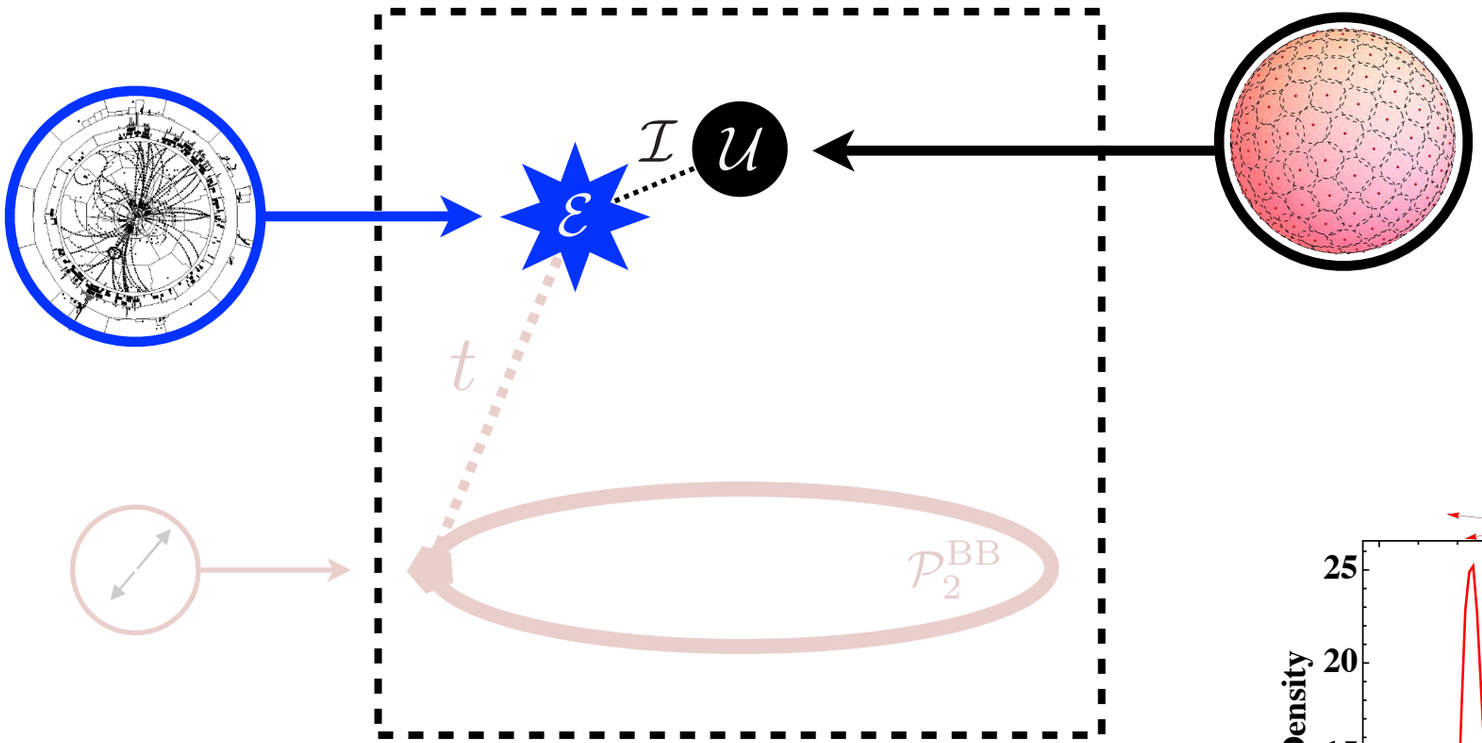

$$1 - \frac{t}{2E_{\text{CM}}}$$

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

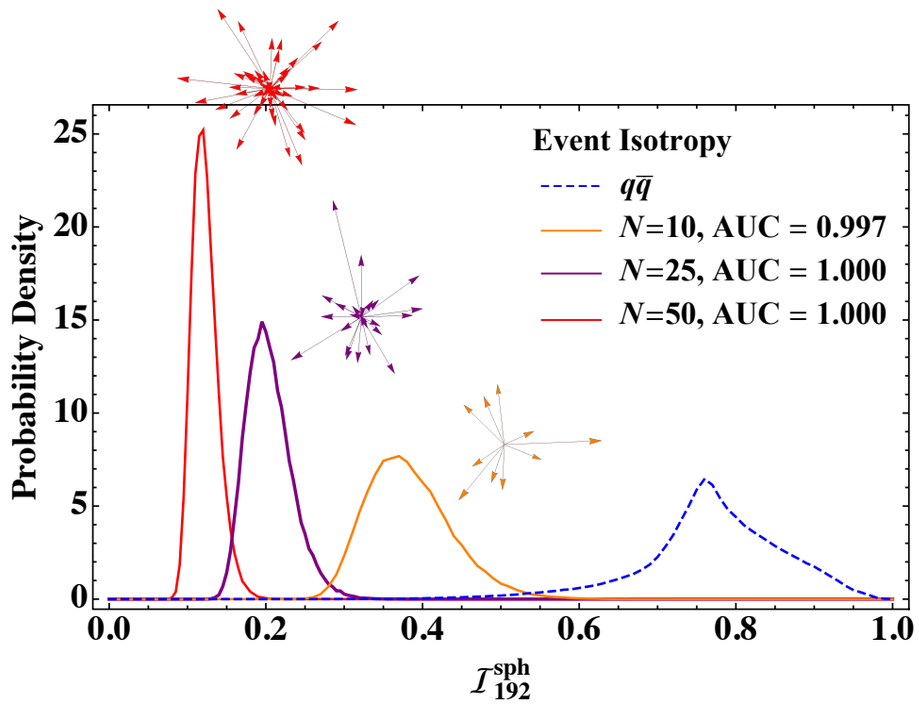
[Komiske, Metodiev, JDT, JHEP 2020]
[Brandt, Peyrou, Sosnowski, Wroblewski, PL 1964; Farhi, PRL 1977; ALEPH, PLB 1991]

# New! Event Isotropy

*How isotropic is an event?*



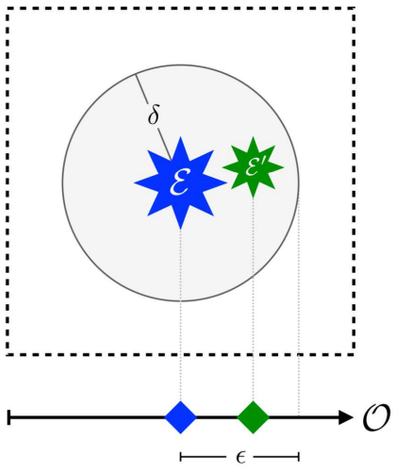

$$\mathcal{I}(\mathcal{E}) = \mathrm{EMD}(\mathcal{E}, \mathcal{U})$$



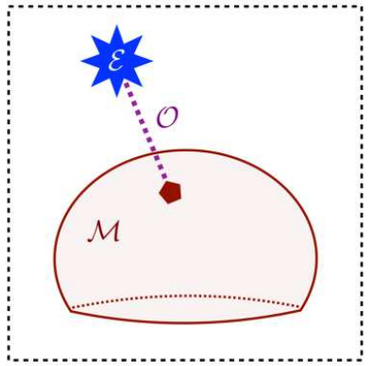

[Cesarotti, JDT, JHEP 2020;
see also Cesarotti, Reece, Strassler, arXiv 2020]

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

**IRC Safety** is smoothness in the space of events

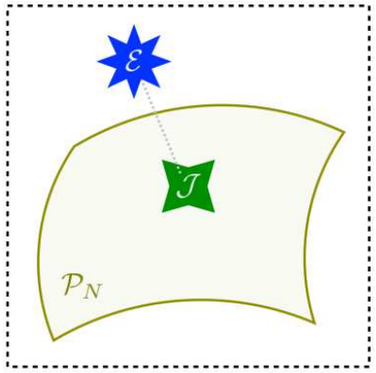



**Event shapes** are distances from events to manifolds.



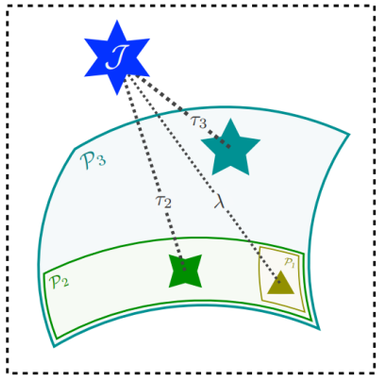

$$\mathcal{O}(\mathcal{E}) = \min_{\mathcal{E}' \in \mathcal{M}} \mathrm{EMD}_{\beta,R}(\mathcal{E}, \mathcal{E}')$$

**Jets** are projections to few-particle manifolds.



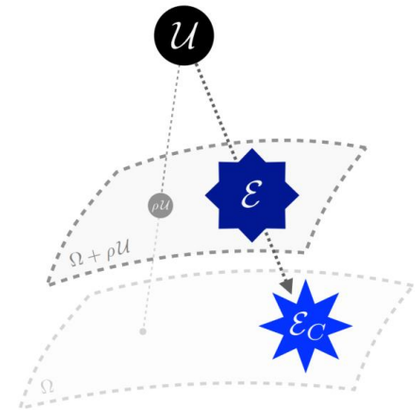

$$\mathcal{J} = \underset{\mathcal{E}' \in \mathcal{P}_N}{\mathrm{argmin}} \, \mathrm{EMD}_{\beta,R}(\mathcal{E}, \mathcal{E}')$$

**Substructure** resolves emissions within the jet.




$$\tau(\mathcal{J}) = \min_{\mathcal{E}' \in \mathcal{P}_N} \mathrm{EMD}_{\beta}(\mathcal{J}, \mathcal{E}').$$

**Pileup mitigation** moves away from uniform radiation.




$$\mathcal{E}_C = \arg\min_{\mathcal{E}'} \mathrm{EMD}(\mathcal{E}, \mathcal{E}' + \rho\,\mathcal{U}).$$

1960

**Taming infinities**

1962-1964
Infrared Safety
[Kinoshita, JMP 1962]
[Lee, Nauenberg, PR 1964]

**Event Shapes**

1977
Thrust, Sphericity
[Farhi, PRL 1977]
[Georgi, Machacek, PRL 1977]

**Jet Algorithms**

1993
$k_T$ jet clustering
[Ellis, Soper, PRD 1993]
[Catani, Dokshitzer, Seymour, Webber, NPB 1993]

1997-1998
C/A jet clustering
[Wobisch, Wengler, 1998]
[Doskhitzer, Leder, Moretti,Webber, JHEP 1997]

**Jet Substructure**

2010-2015
N-(sub)jettiness, XCone
[Stewart, Tackmann, Waalewijn, PRL 2010]
[Thaler, Van Tilburg, JHEP 2011]
[Stewart, Tackmann, Thaler, Vermilion, Wilkason, JHEP 2015]

**Pileup**

2014-2019
Constituent Subtraction
[Berta, Spousta, Miller, Leitner, JHEP 2014]
[Berta, Masetti, Miller, Spousta, JHEP 2019]

2020

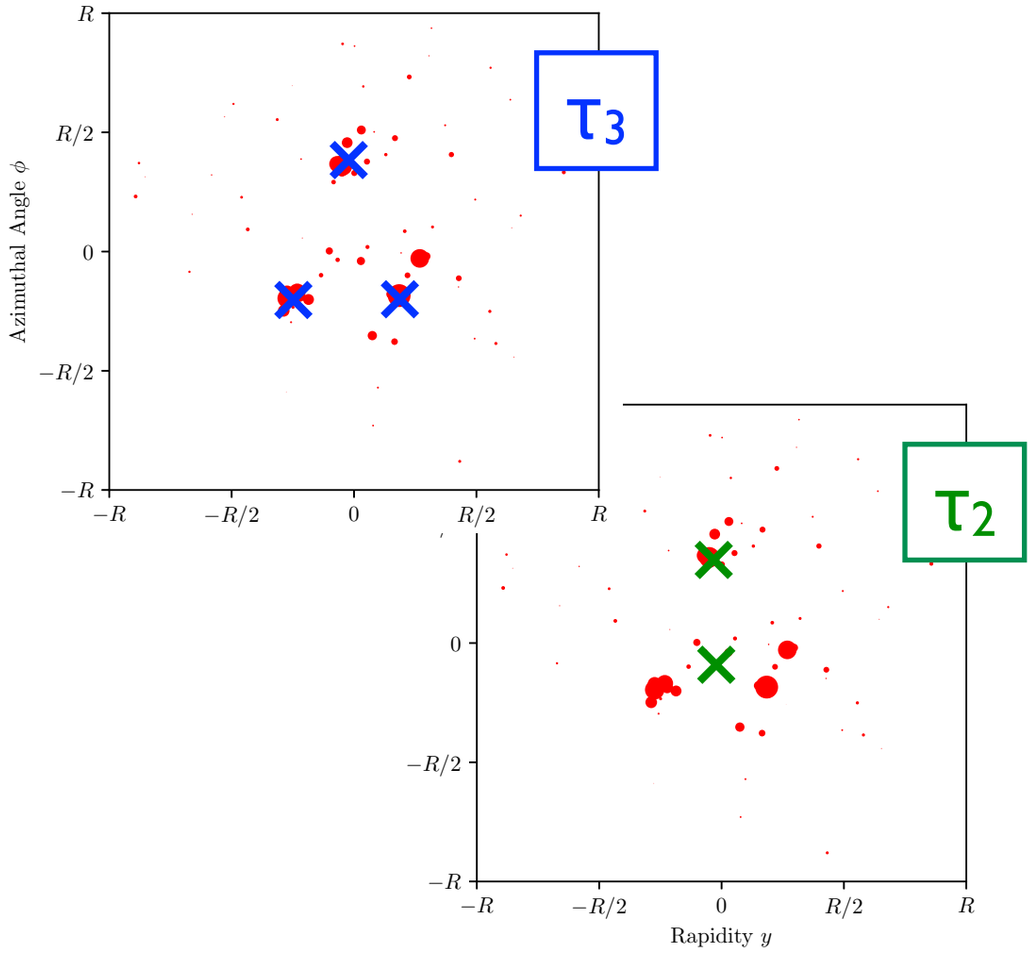
**And many more!**

[timeline from Eric Metodiev]

# N-subjettiness

*Ubiquitous jet substructure observable used for almost a decade…*

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

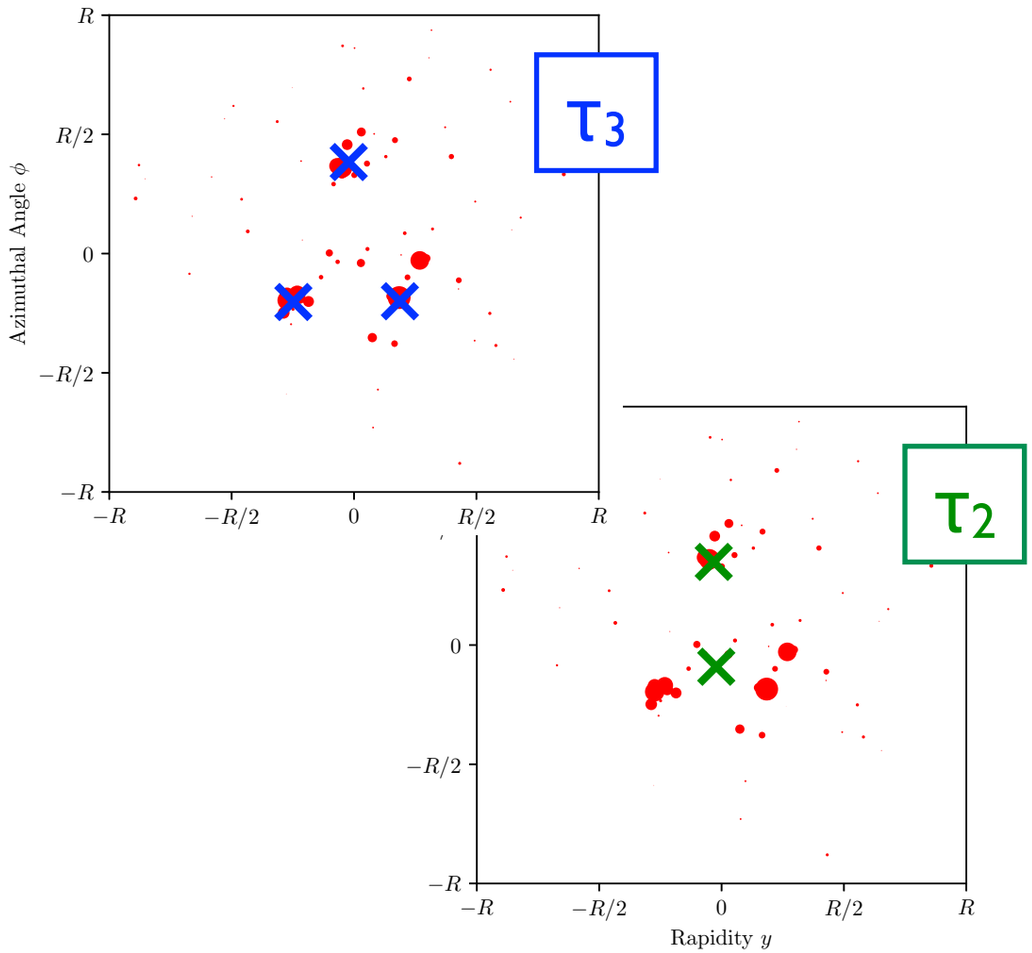


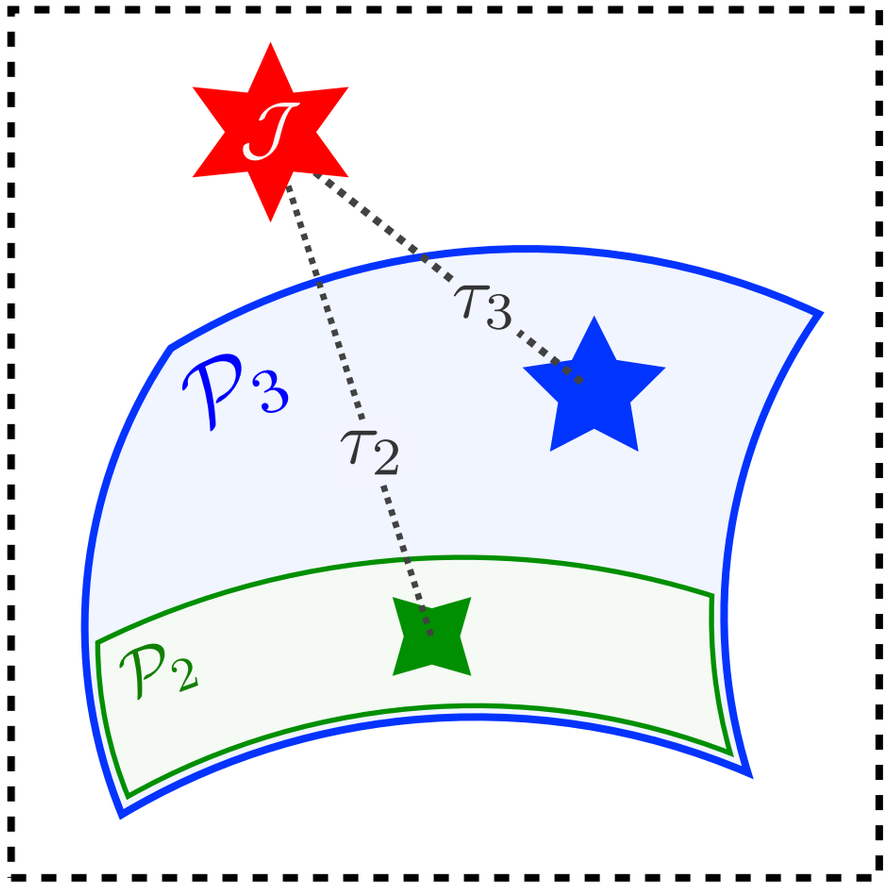

[JDT, Van Tilburg, JHEP 2011, JHEP 2012;
based on Brandt, Dahmen, ZPC 1979; Stewart, Tackmann, Waalewijn, PRL 2010]

# N-subjettiness = Point to Manifold EMD

*…is secretly an optimal transport problem*

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



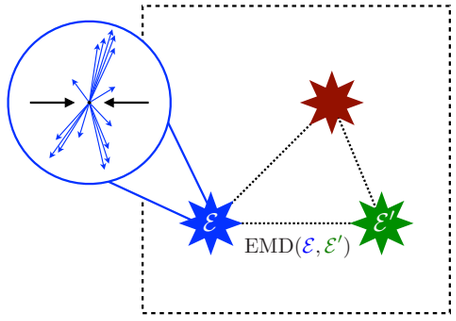




[JDT, Van Tilburg, JHEP 2011, JHEP 2012;
rephrased in the language of Komiske, Metodiev, JDT, PRL 2019]



# Pause

*What can be Geometrized?*

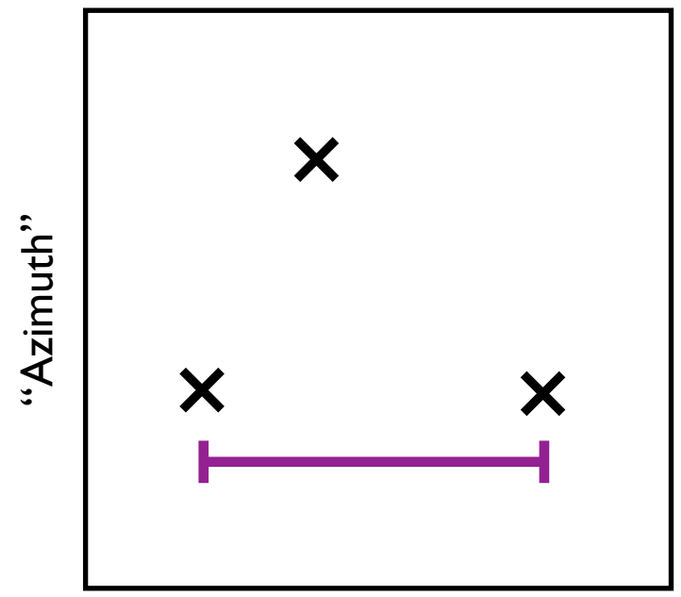



IRC Safety, Observables,
Jet Algorithms, Pileup Mitigation

*And now, the grand finale…*

*How far down does this rabbit hole go?*

# Direction Space




× = Direction

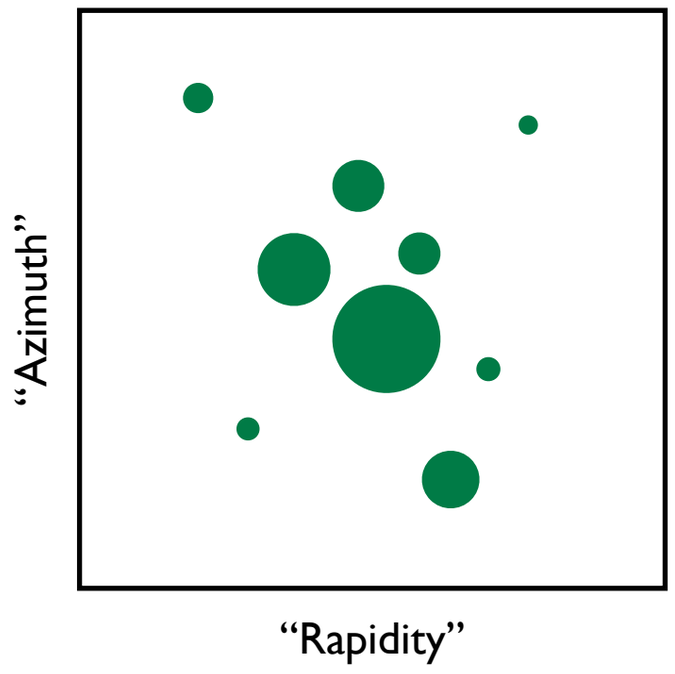
$$n_i^\mu = \frac{p_i^\mu}{E_i} = (1, \hat{n})^\mu$$

⟷ = Angular Distance

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

(for massless particles)

# Direction Space Distribution




= Weighted Direction

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

$$w_i = E_i$$

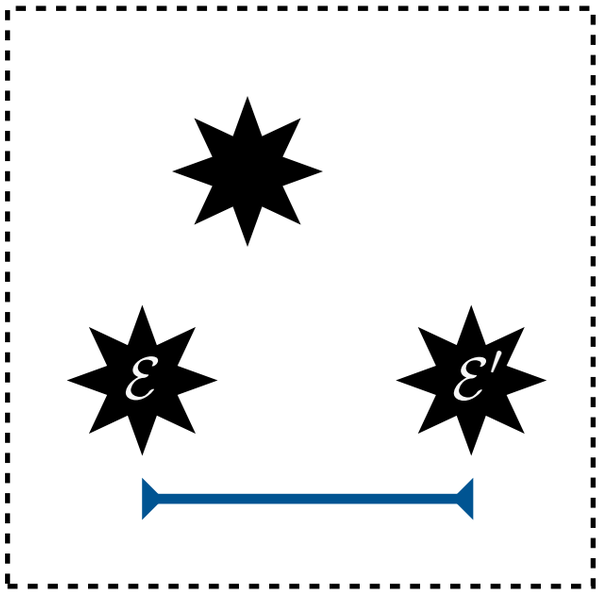
= Angular Distance

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

(for massless particles)

= Event

# Event Space




= Event

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

= EMD

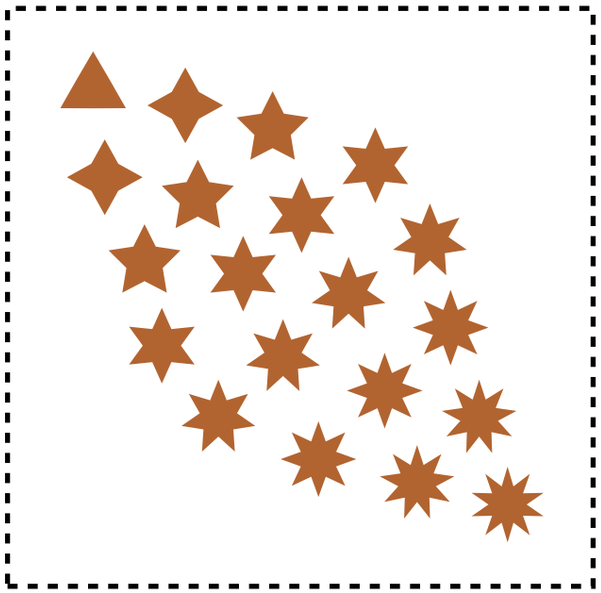
Energy
Mover's Distance

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

(for equal total energy)

[Komiske, Metodiev, JDT, PRL 2019]

# Event Space Distribution



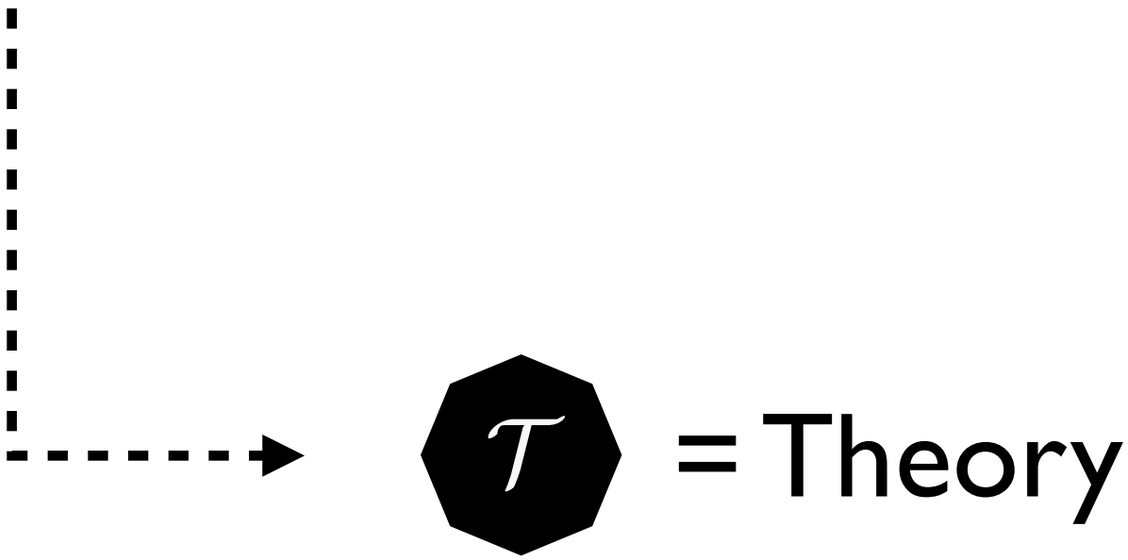
= Weighted Event $\quad \mathcal{E}(\hat{n}) = \sum_i E_i \, \delta(\hat{n} - \hat{n}_i)$

$w_a = \sigma_a$

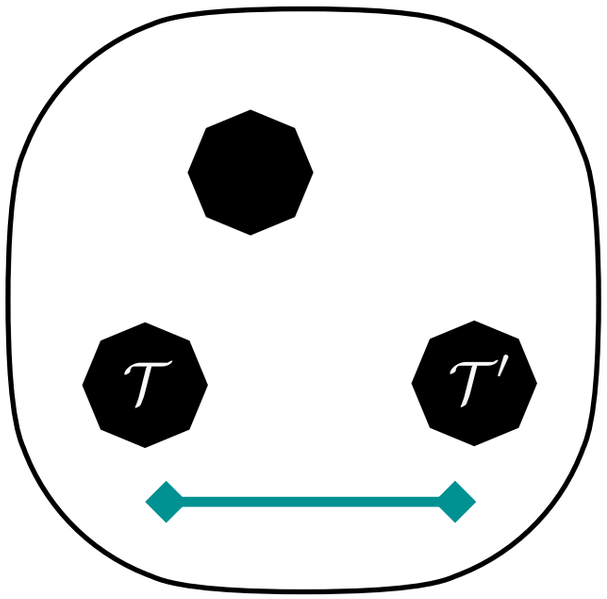
= EMD $\quad \mathrm{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_i \sum_j f_{ij} \theta_{ij}$

Energy
Mover's Distance

(for equal total energy)



# Theory Space




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

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

Cross-Section Mover's Distance

(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

jet

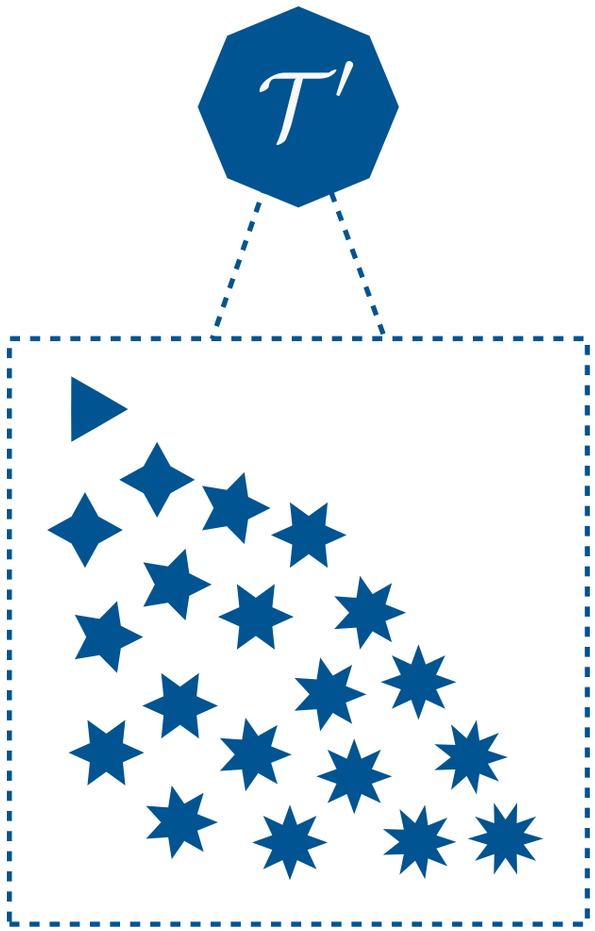
Pb

QGP

Pb

quenched jet

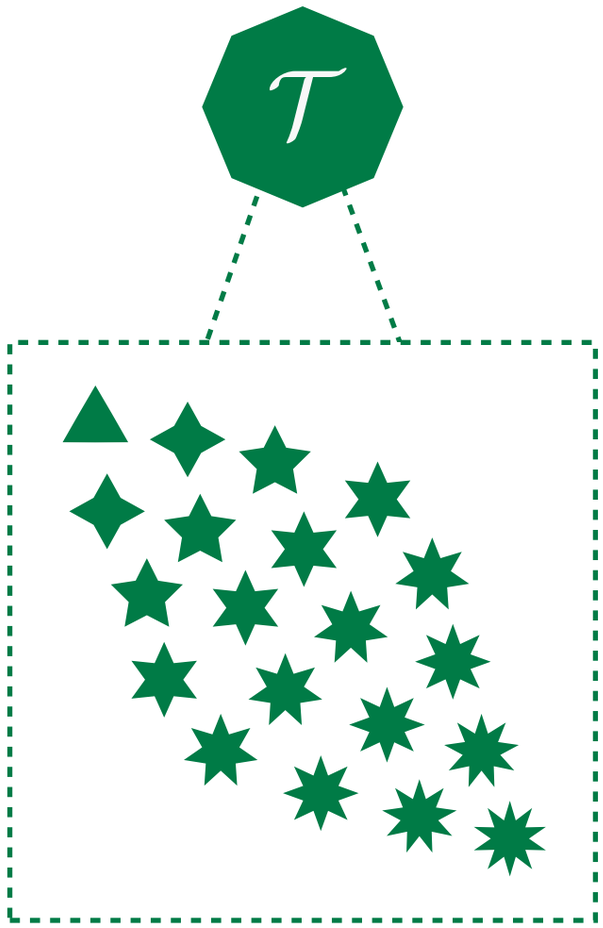
Theory Prime:
In-Medium QCD




ΣMD

⇔

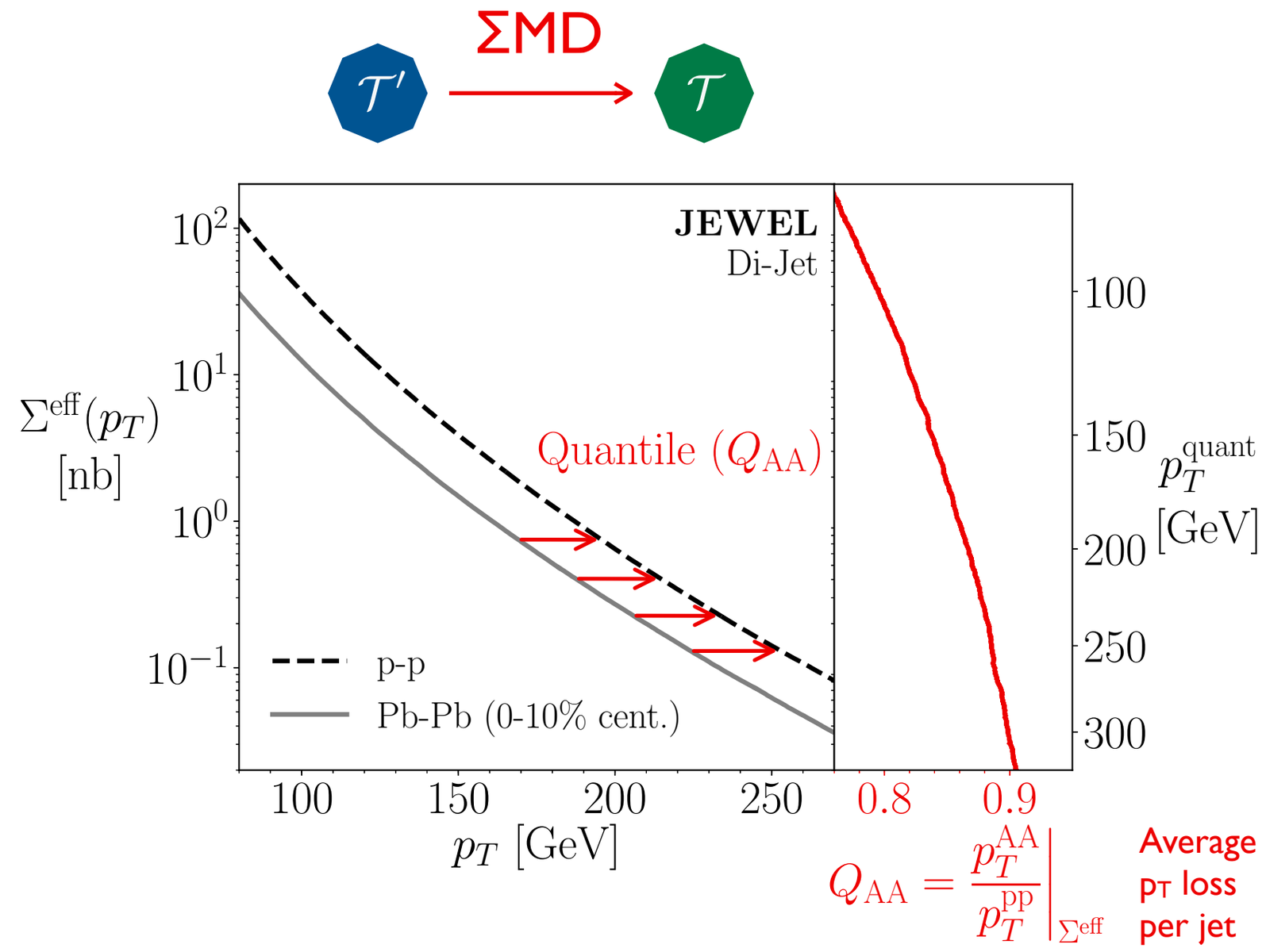
Theory:
Vacuum QCD




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

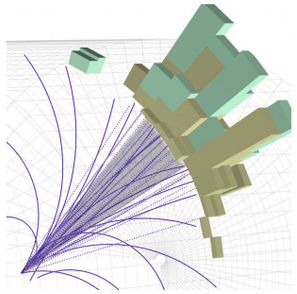
# Jet Quenching via Quantile Matching

*Equivalent to following a geodesic in theory space (!)*




[Brewer, Milhano, JDT, PRL 2019]

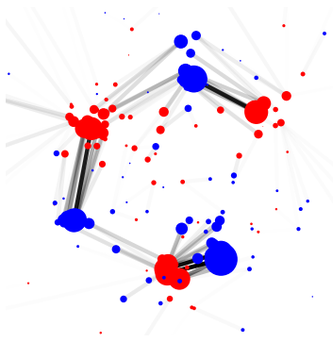
# Summary



## What is a Collider Event?

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

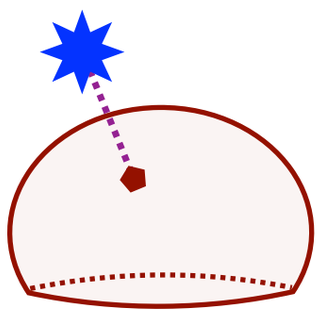
*(ask me about QCD/ML relevance)*



## 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*

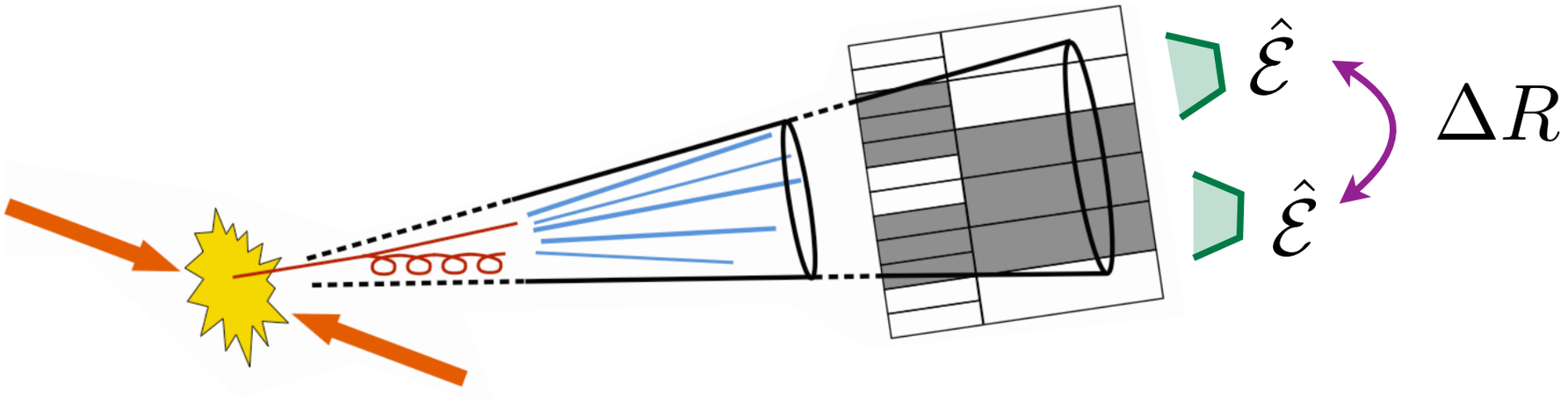
*(ask me more about safety)*

# Fin

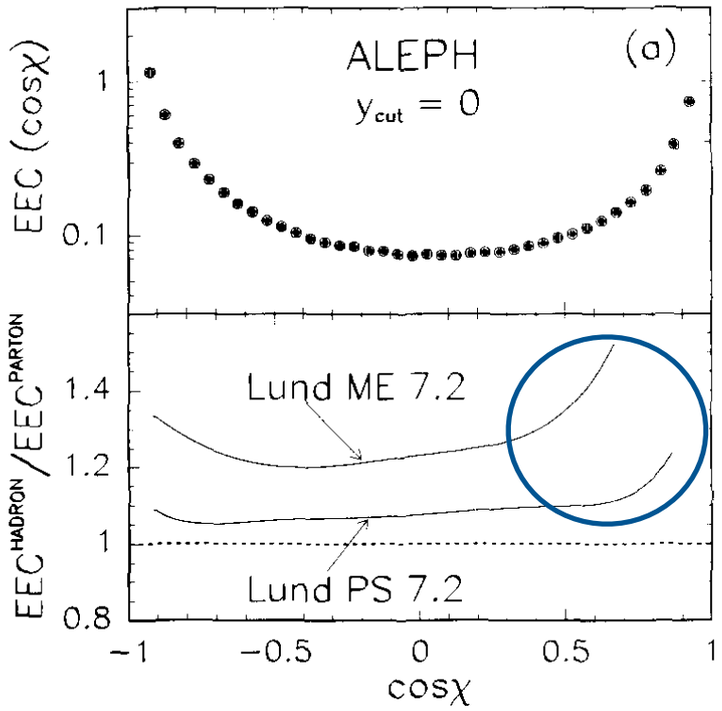
*Whew!*

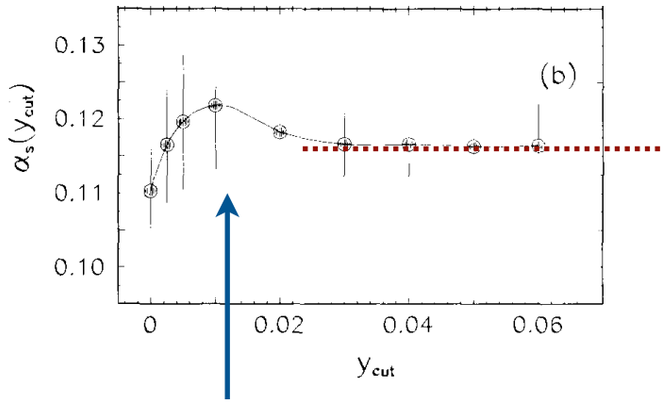
# *Backup Slides*

# Energy-Energy Correlators




A long history in probing collinear dynamics of QCD



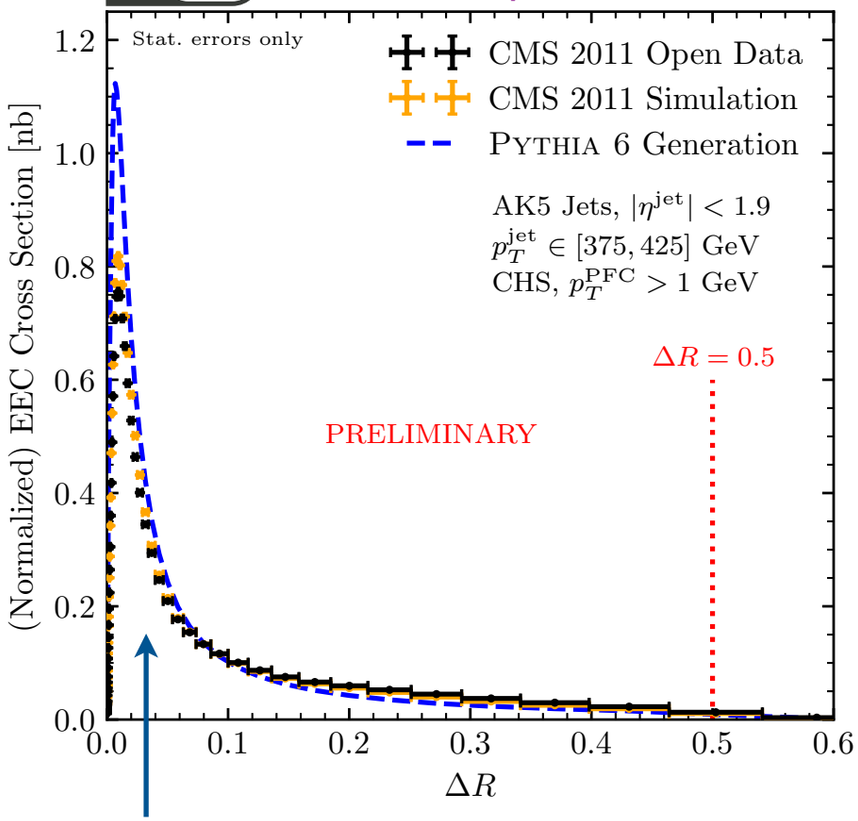

Extracting the strong coupling constant

Theoretical challenges with small angle (collinear) limit

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

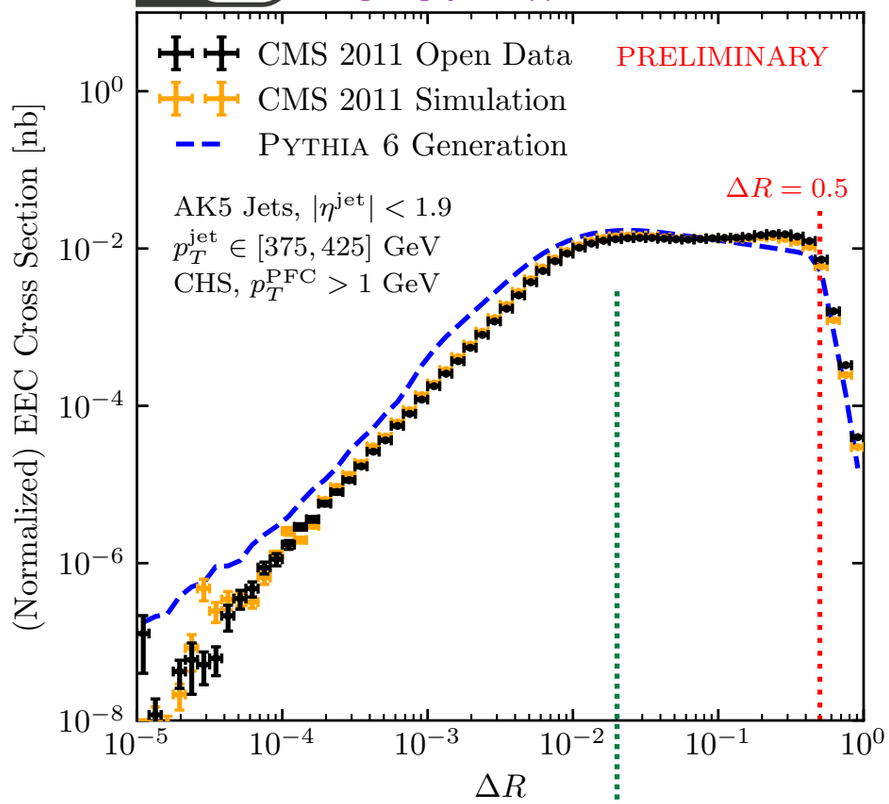
# QCD Phase Transition in Jets?

MOD *Linear-linear plot*

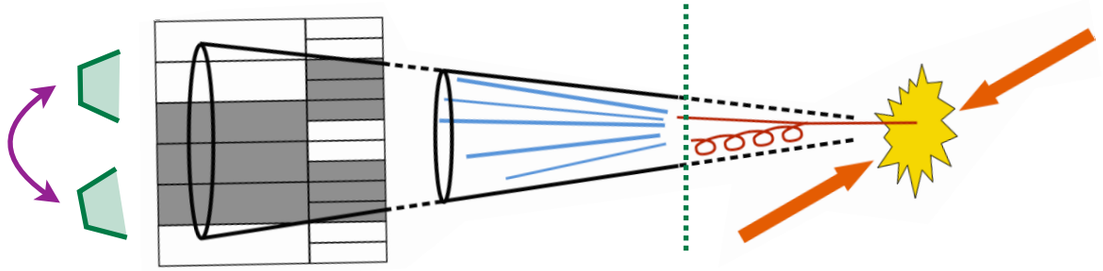



Are we learning something about small angle limit of QCD?

MOD *Log-log plot (!)*



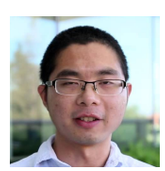

Hadronic Phase — Partonic Phase



First Jet EEC Plot from the LHC (!)

[Komiske, Moult, JDT, Zhu, in progress; see talks by Moult, BOOST 2019, BOOST 2020]





# *Point Cloud*

Collection of points in position space

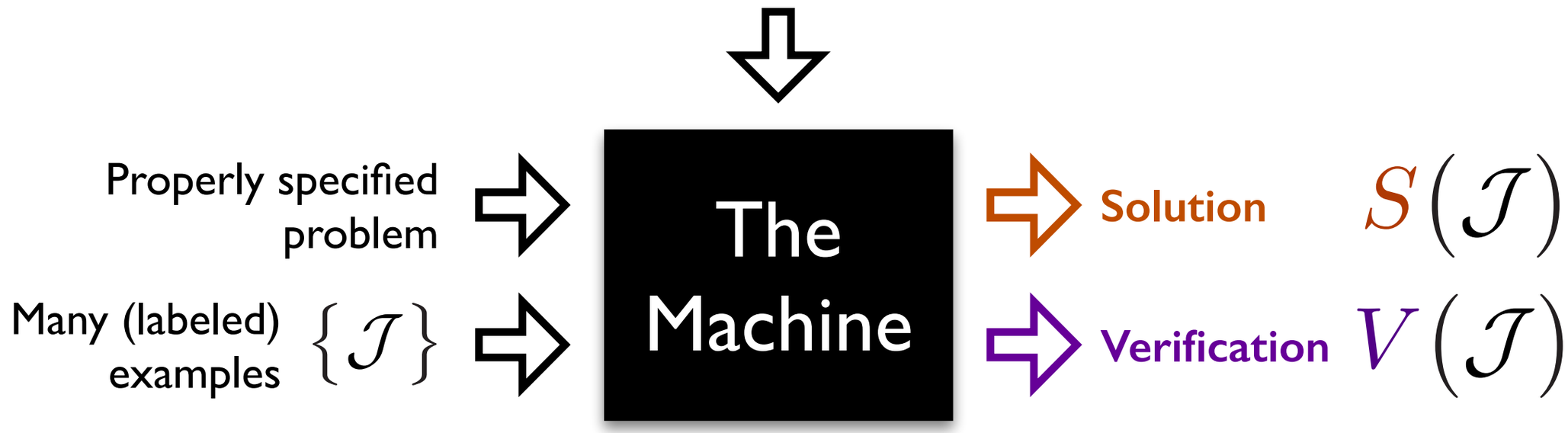
[Popular Science, 2013]

# Aside: Machine Learning for Jets

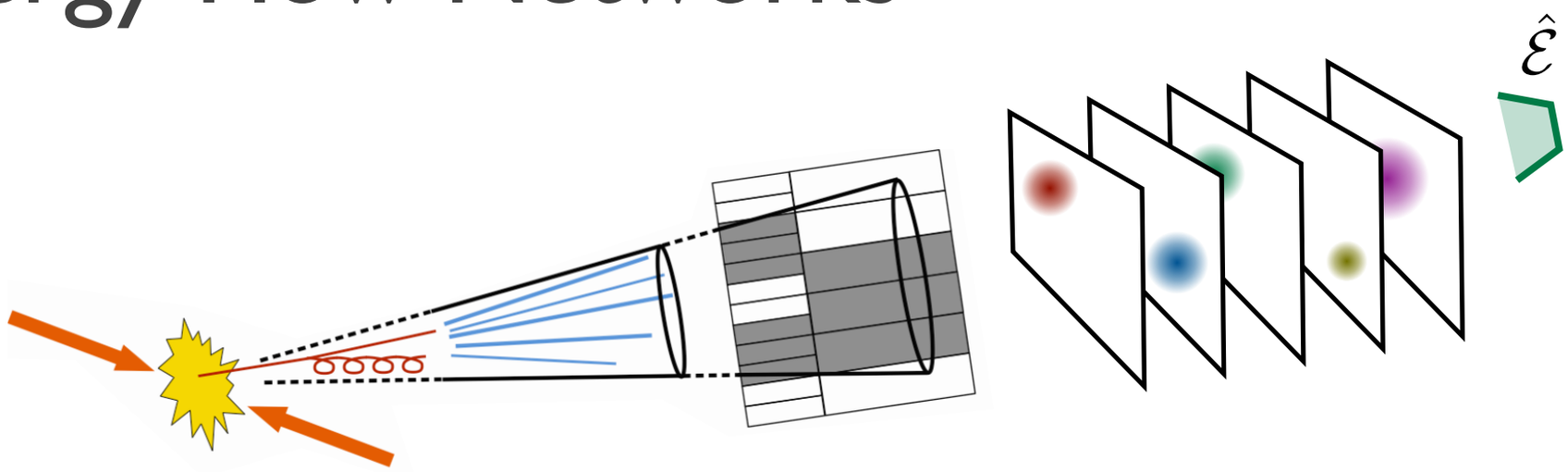


**Symmetry:** $\mathcal{J} = \{\underbrace{\vec{p}_1, \vec{p}_2, \vec{p}_3, \ldots, \vec{p}_N}_{\text{Unordered, Variable Length Set (QM!)}}\}$

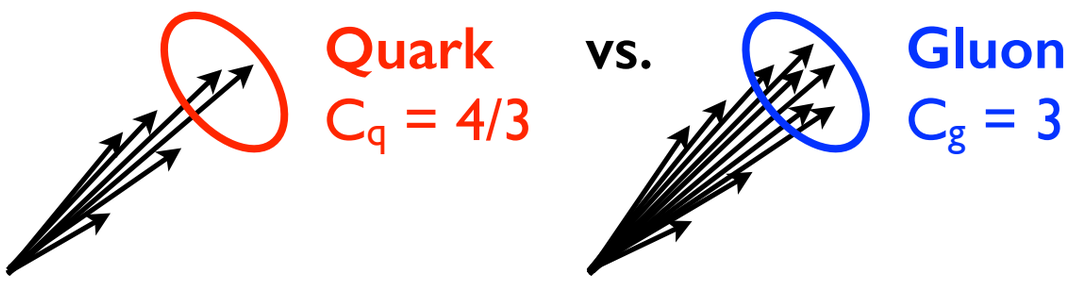
**Safety:** $\vec{p} = \{\underbrace{E}_{\text{Energy weighting (QFT!)}}, \hat{n}_x, \hat{n}_y, \hat{n}_z\}$

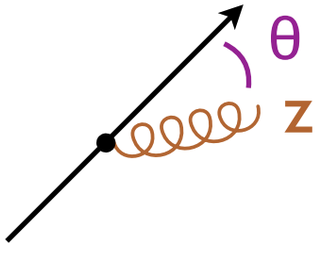


# E.g. Energy Flow Networks



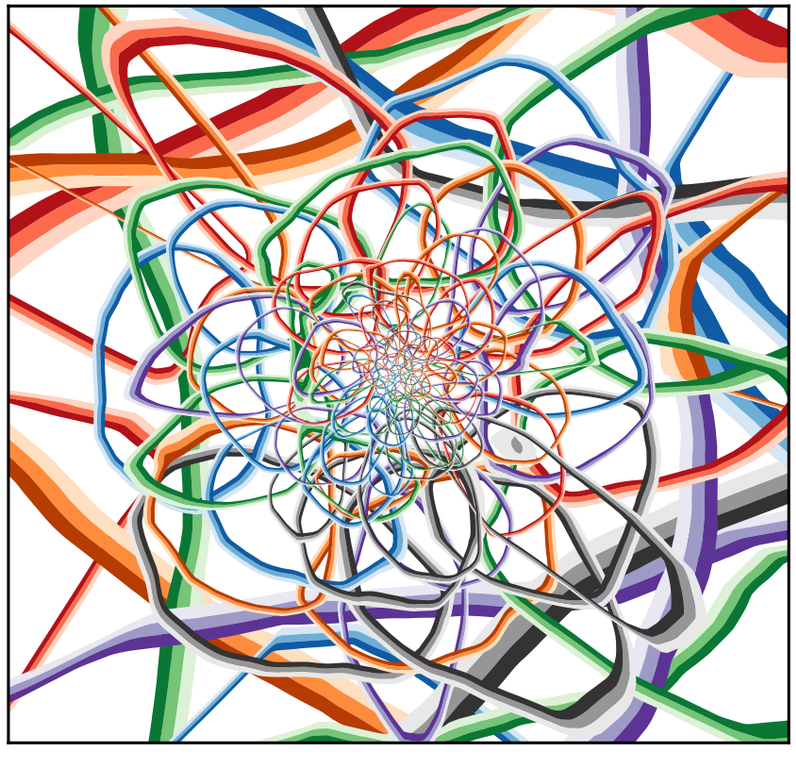

*Learning QCD singularities!*



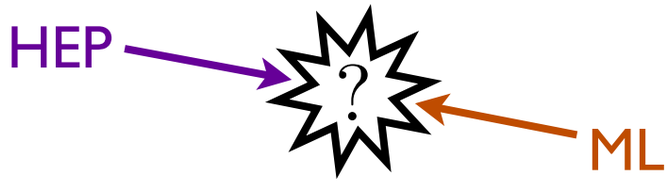

AP splitting probability:

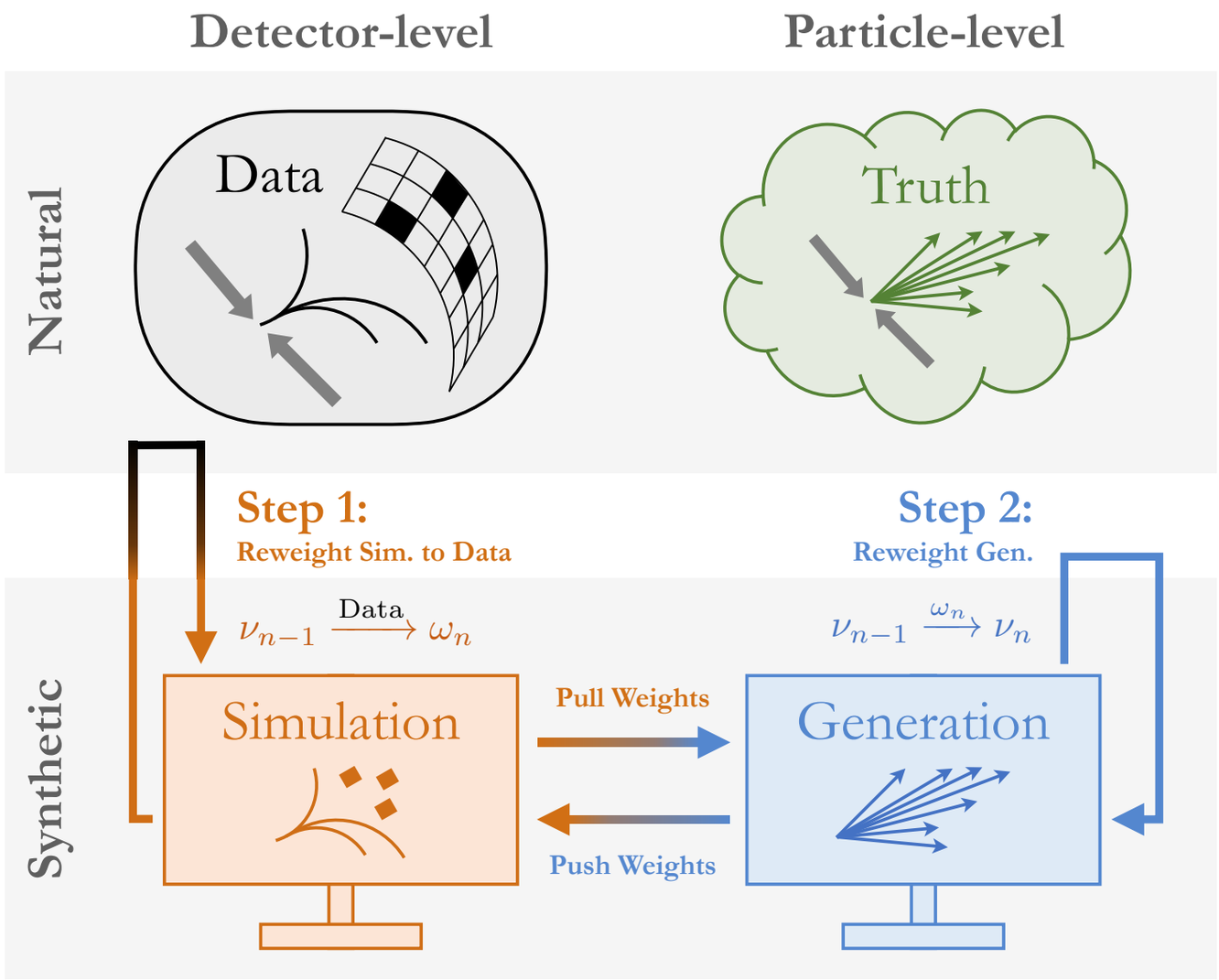
$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{\mathrm{d}\theta}{\theta} \frac{\mathrm{d}z}{z}$$



[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



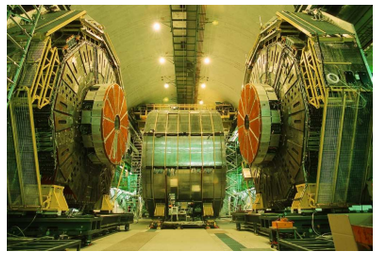

Using, e.g., Energy Flow Networks

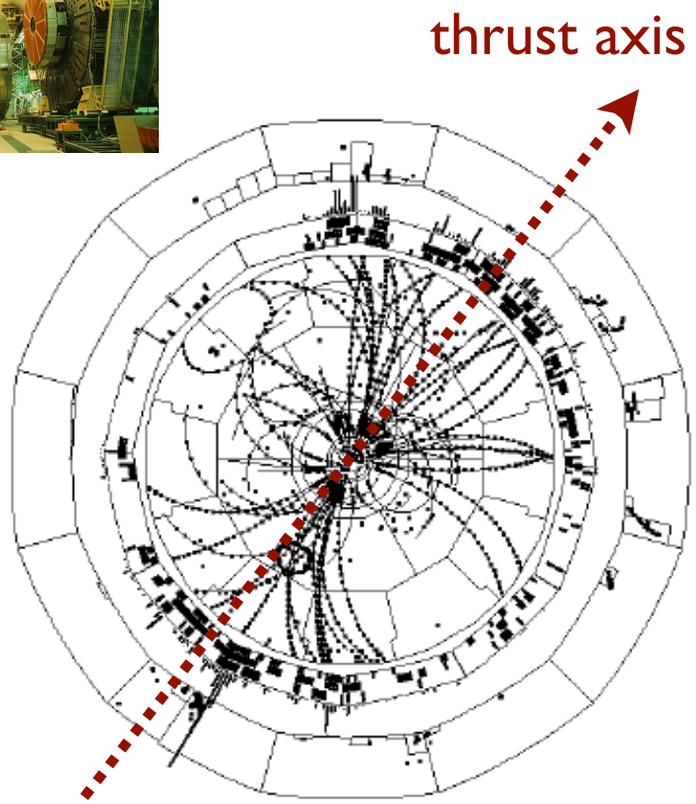
*Multi-dimensional unbinned detector corrections via iterated binary classification*

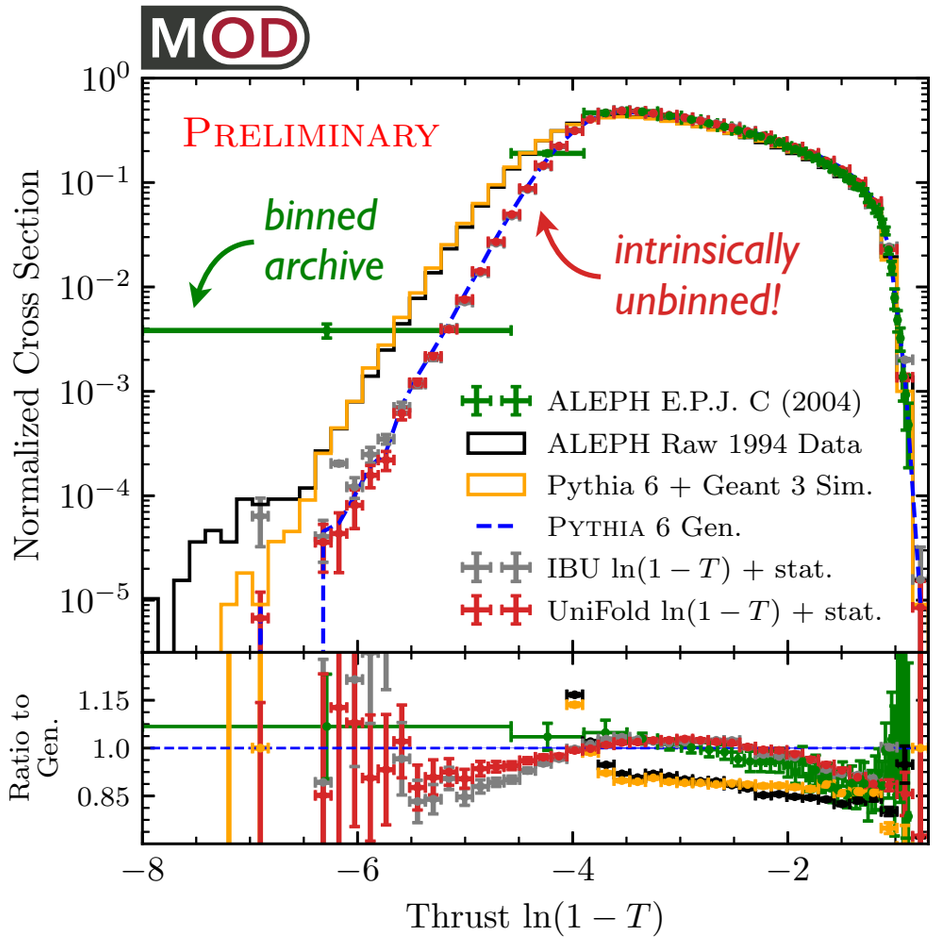
[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020]

# OmniFold

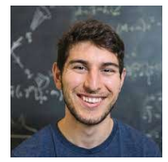

## Back to the Future with ALEPH Archival Data



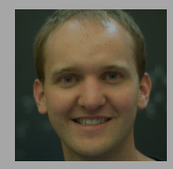



[talk by Badea, ICHEP 2020; cf. ALEPH, EPJC 2004]
[see also Badea, Baty, Chang, Innocenti, Maggi, McGinn, Peters, Sheng, JDT, Lee, PRL 2019]

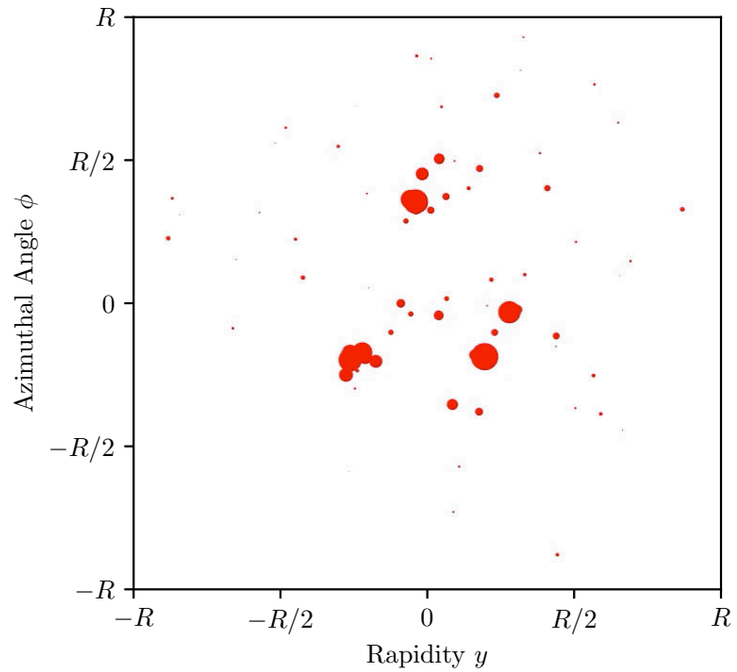


[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020]



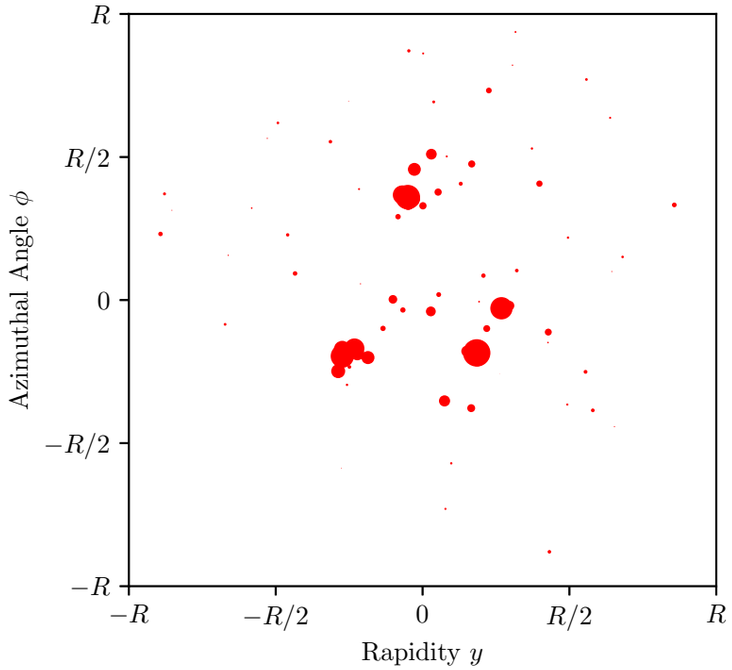

# Similarity of Two Energy Flows?

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

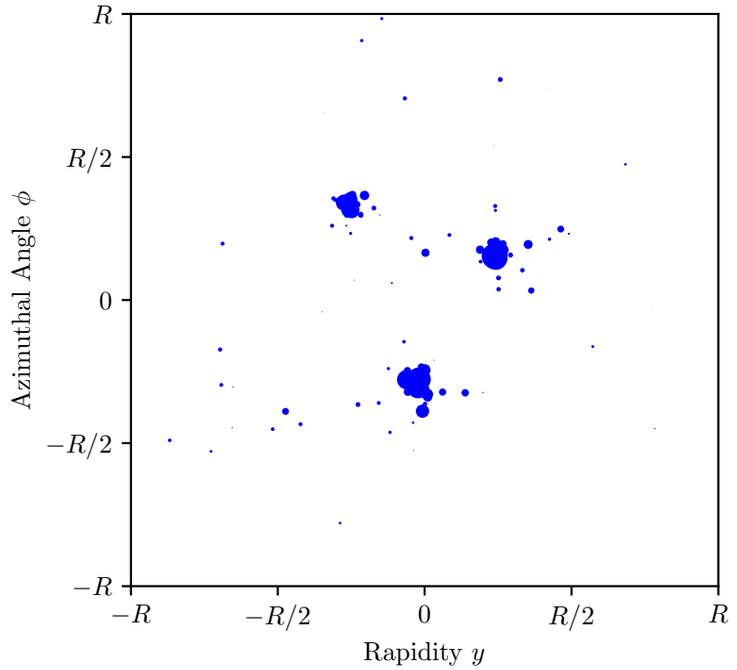



## Optimal Transport:

*Earth Mover's Distance*
*a.k.a. 1-Wasserstein metric*




125.4 GeV




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

# Dimensionality of Space of Jets

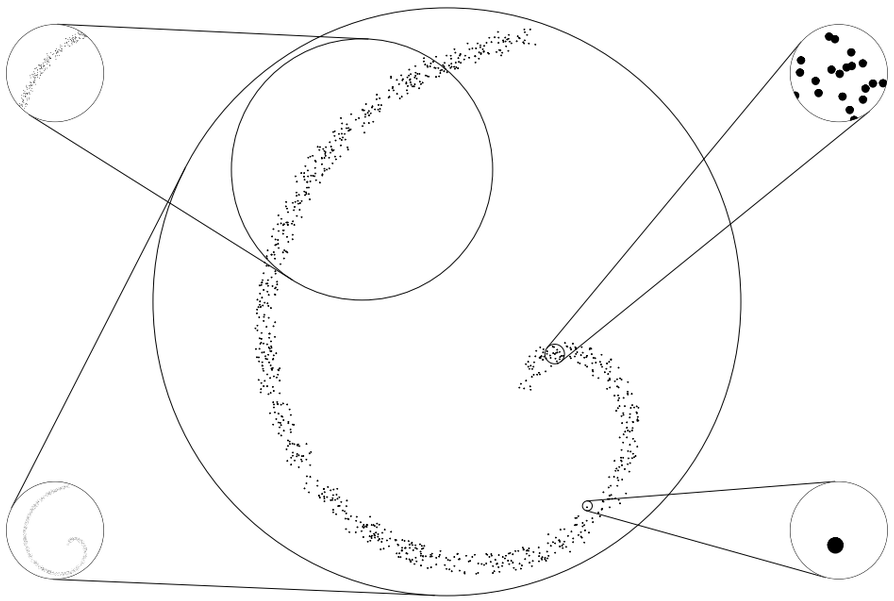

$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

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

[Grassberger, Procaccia, PRL 1983; Kégl, NIPS 2002]

dim≃1

dim≃2

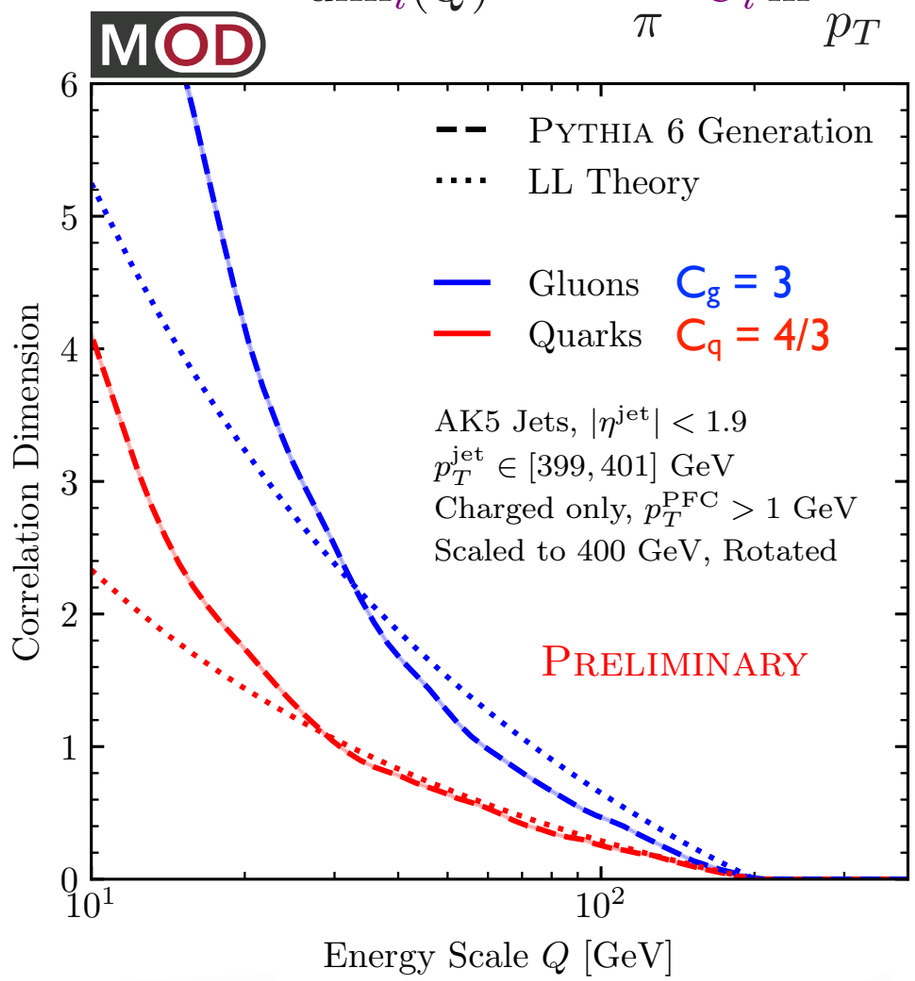


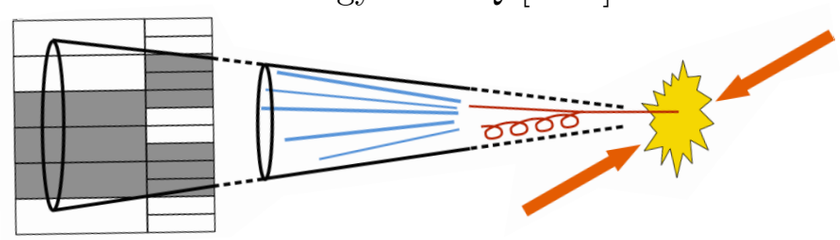
dim≃2

(eventually 0)

dim≃0

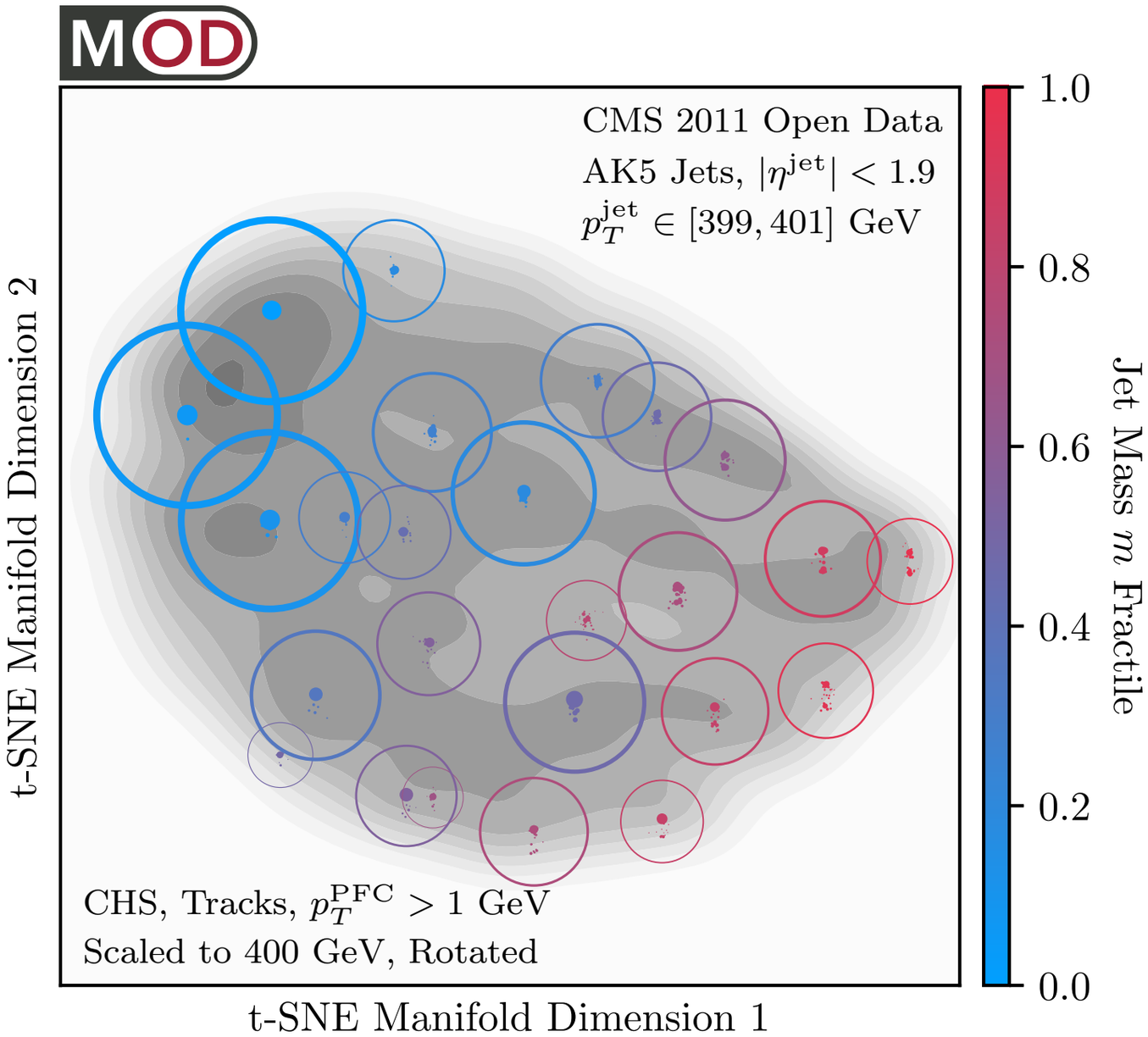
$$\dim_i(Q) \simeq -\frac{8\alpha_s}{\pi} C_i \ln \frac{Q}{p_T}$$



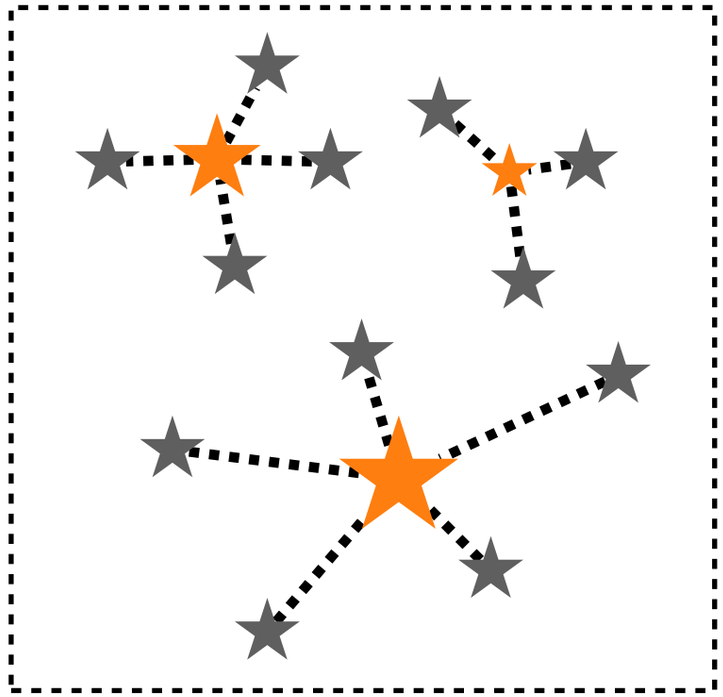
# Most Representative Jets

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




## k-medoids

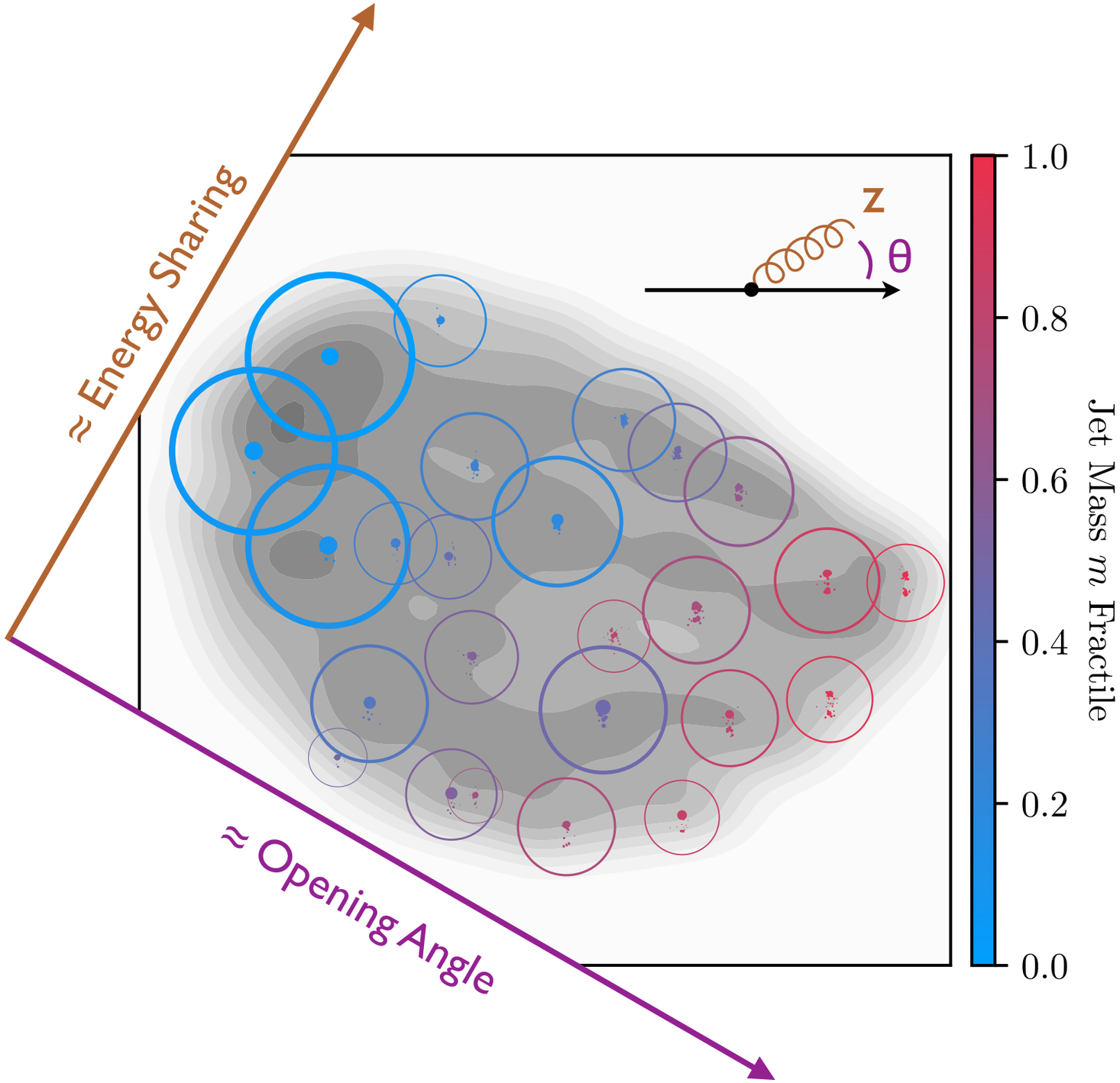
*Arranged via t-SNE*



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

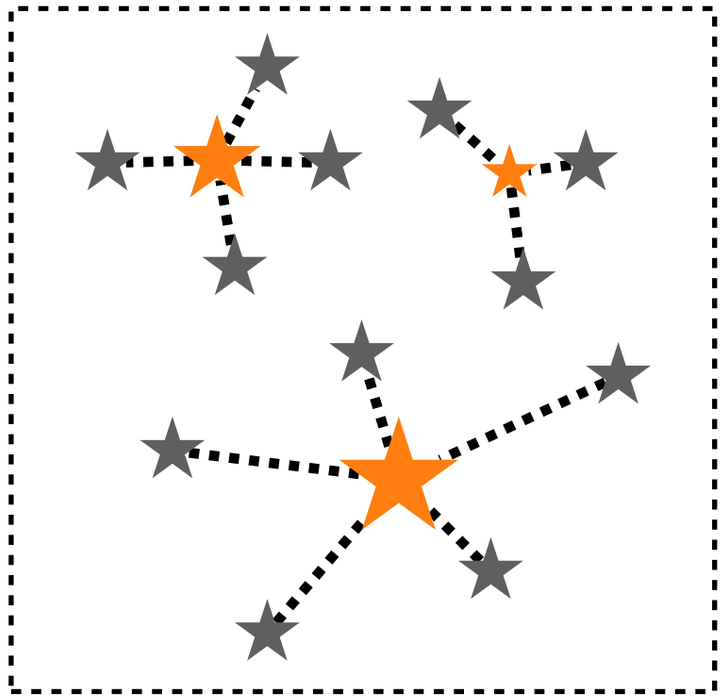
# Most Representative Jets

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




## k-medoids

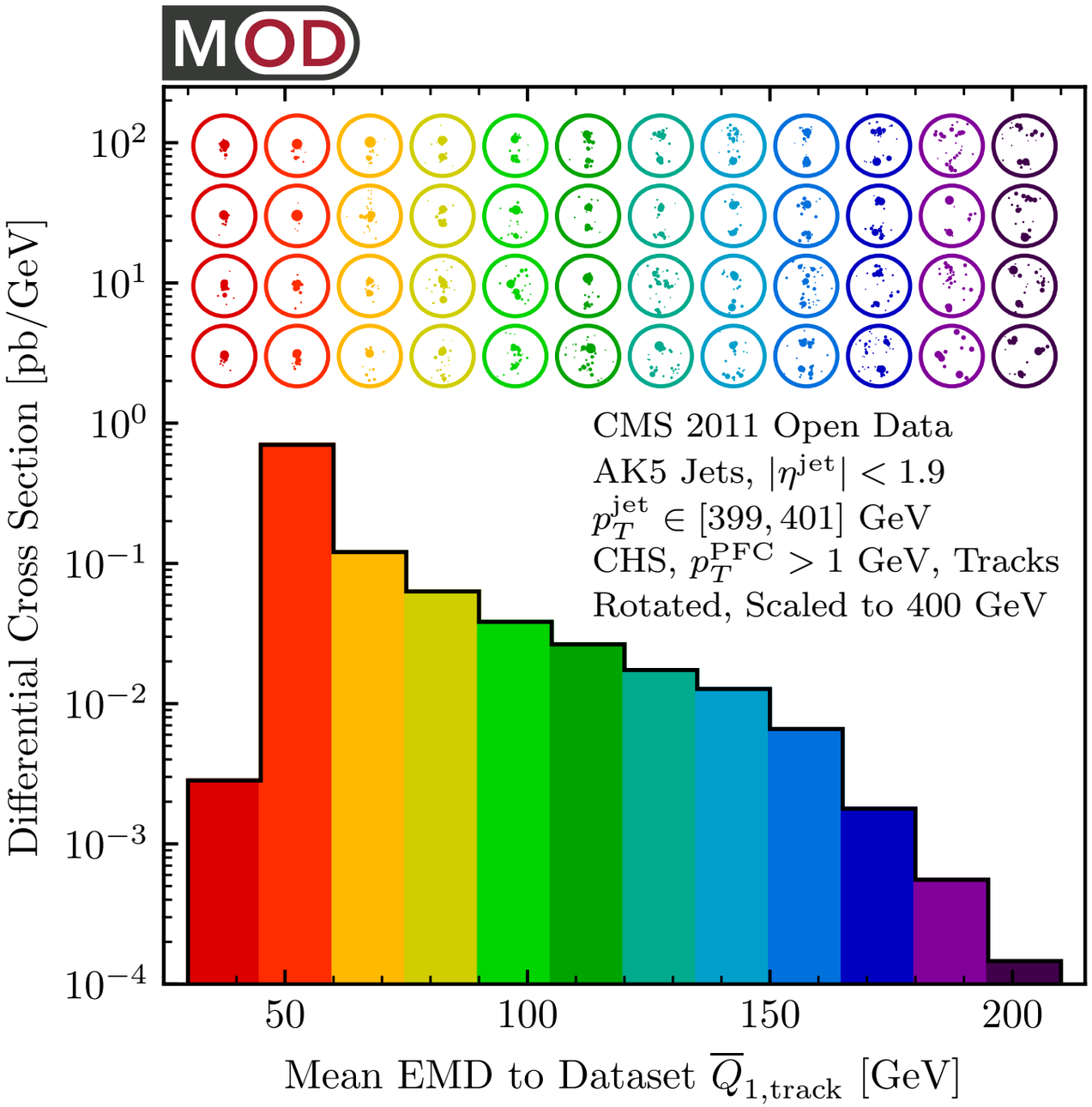
*Arranged via t-SNE*



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

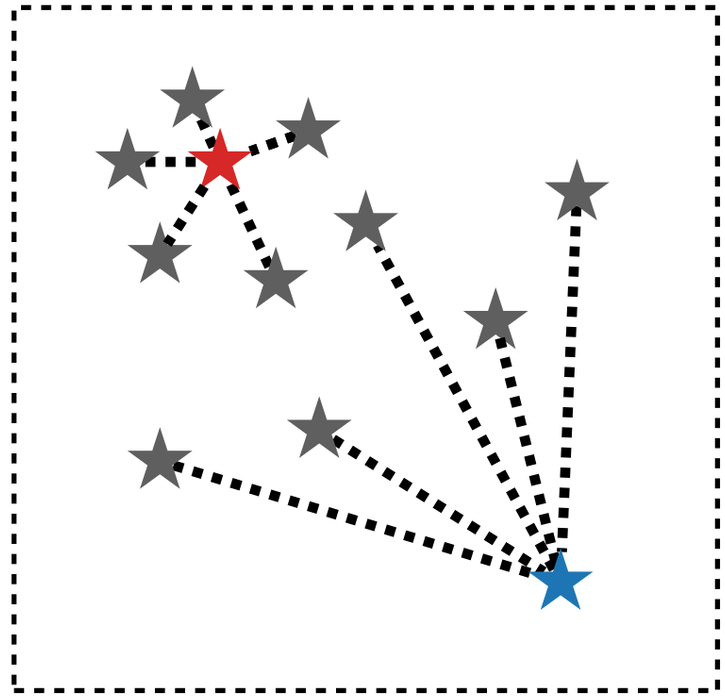
# Least Representative Jets

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



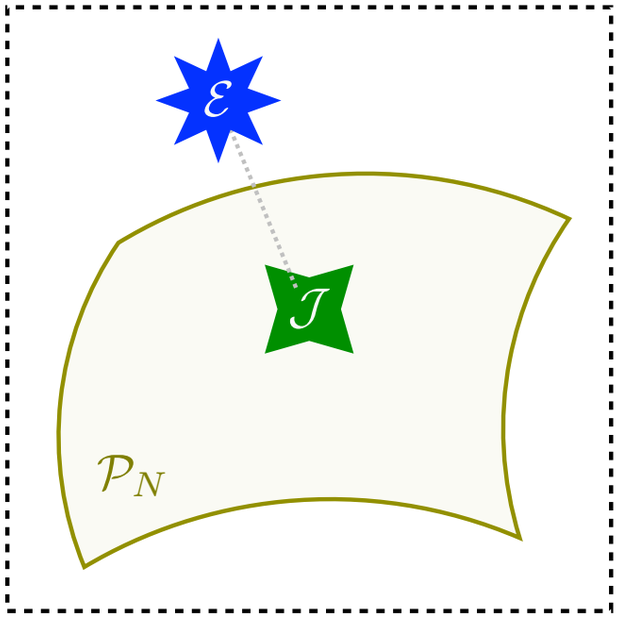

## New Physics?

*Or tails of QCD?*



[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020]

# More Fun with N-particle Manifolds




## N-jettiness

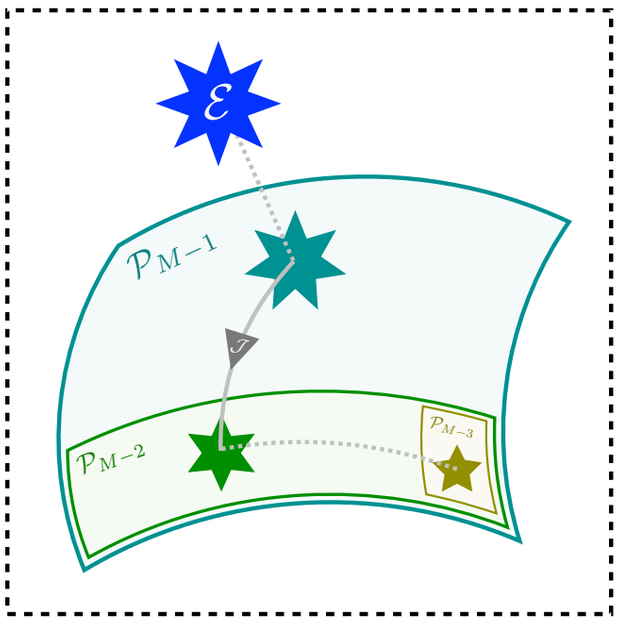
*Distance of closest approach to N-particle manifold*

[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]



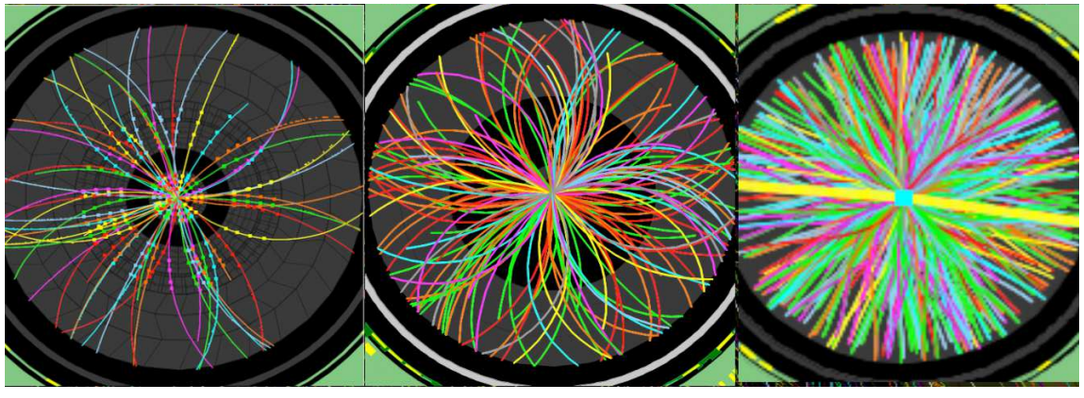

## Sequential Jet Recombination

*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]

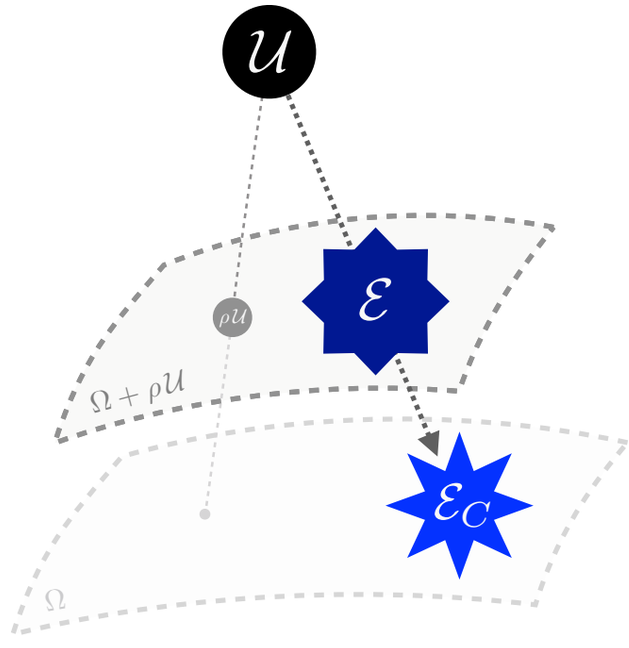
[Komiske, Metodiev, JDT, JHEP 2020]

# Pileup Mitigation



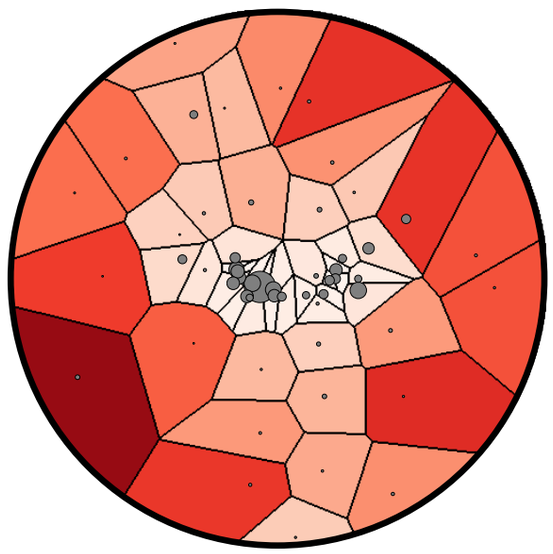
[see review in Soyez, PR 2019]

Uniform event contamination from overlapping proton-proton collisions



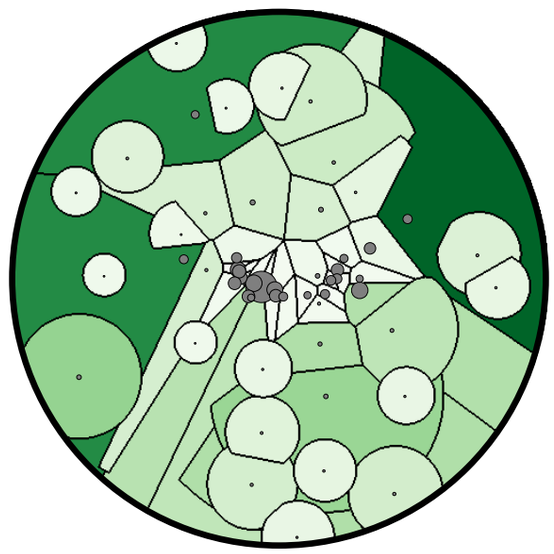

Pileup Mitigation:
"Move away" from uniform event

*Voronoi*



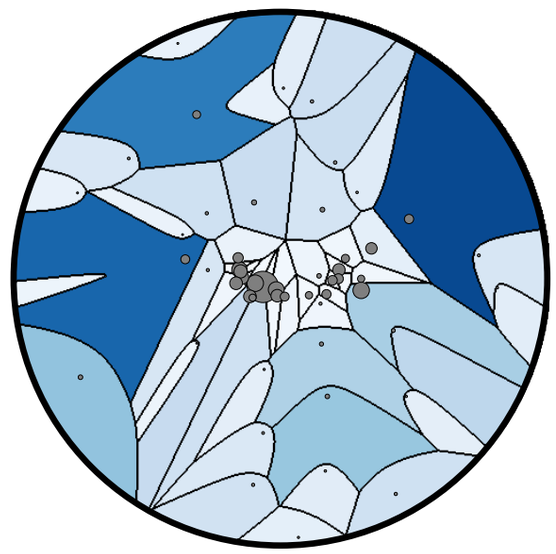
[Cacciari, Salam, Soyez, JHEP 2008]

*Constituent Subtraction*



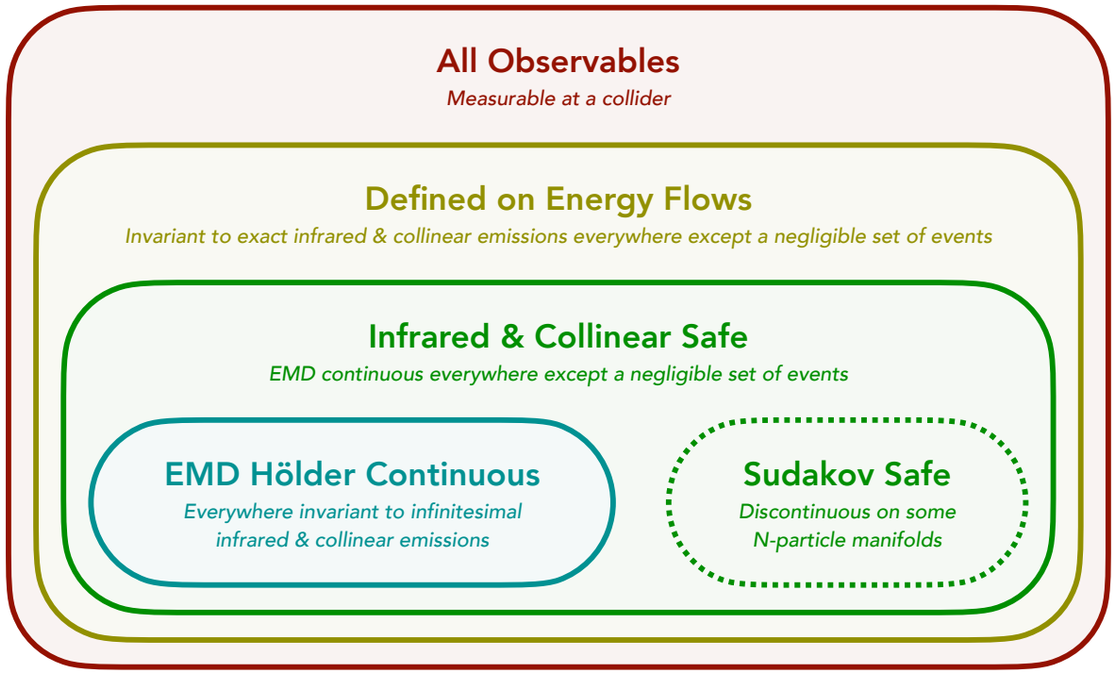
[Berta, Spousta, Miller, Leitner, JHEP 2014]

*Apollonius*



[Komiske, Metodiev, JDT, JHEP 2020]

# Observable Taxonomy




| **All Observables** | Comments |
|---|---|
| Multiplicity ($\sum_i 1$) | IR unsafe and C unsafe |
| Momentum Dispersion [65] ($\sum_i E_i^2$) | IR safe but C unsafe |
| Sphericity Tensor [66] ($\sum_i p_i^\mu p_i^\nu$) | IR safe but C unsafe |
| Number of Non-Zero Calorimeter Deposits | C safe but IR unsafe |
| **Defined on Energy Flows** | |
| Pseudo-Multiplicity ($\min\{N \,\vert\, \mathcal{T}_N = 0\}$) | Robust to exact IR or C emissions |
| **Infrared & Collinear Safe** | |
| Jet Energy ($\sum_i E_i$) | Disc. at jet boundary |
| Heavy Jet Mass [67] | Disc. at hemisphere boundary |
| Soft-Dropped Jet Mass [38, 68] | Disc. at grooming threshold |
| Calorimeter Activity [69] ($N_{95}$) | Disc. at cell boundary |
| *Sudakov Safe* | |
| Groomed Momentum Fraction [39] ($z_g$) | Disc. on 1-particle manifold |
| Jet Angularity Ratios [37] | Disc. on 1-particle manifold |
| $N$-subjettiness Ratios [47, 48] ($\tau_{N+1}/\tau_N$) | Disc. on $N$-particle manifold |
| $V$ parameter [36] (Eq. (2.11)) | Hölder disc. on 3-particle manifold |
| **EMD Hölder Continuous Everywhere** | |
| Thrust [40, 41] | |
| Spherocity [42] | |
| Angularities [70] | |
| $N$-jettiness [44] ($\mathcal{T}_N$) | |
| $C$ parameter [71–74] | Resummation beneficial at $C = \frac{3}{4}$ |
| Linear Sphericity [72] ($\sum_i E_i n_i^\mu n_i^\nu$) | |
| Energy Correlators [36, 75–77] | |
| Energy Flow Polynomials [15, 17] | |

[Komiske, Metodiev, JDT, JHEP 2020; cf. Sterman, PRD 1979; Banfi, Salam, Zanderighi, JHEP 2005; Larkoski, Marzani, JDT, PRD 2015]