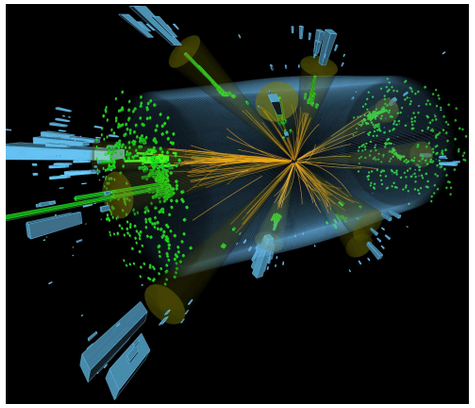
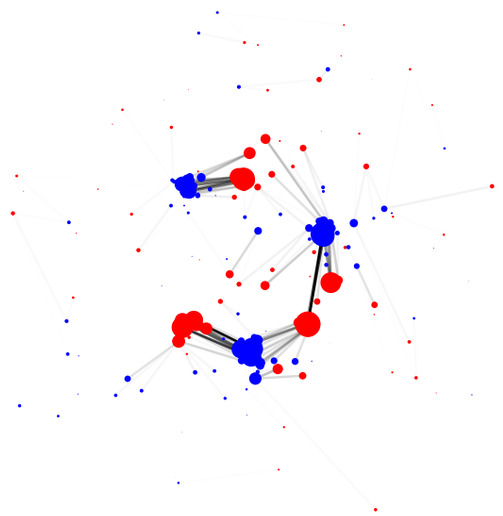
The (Metric) Space of Collider Events

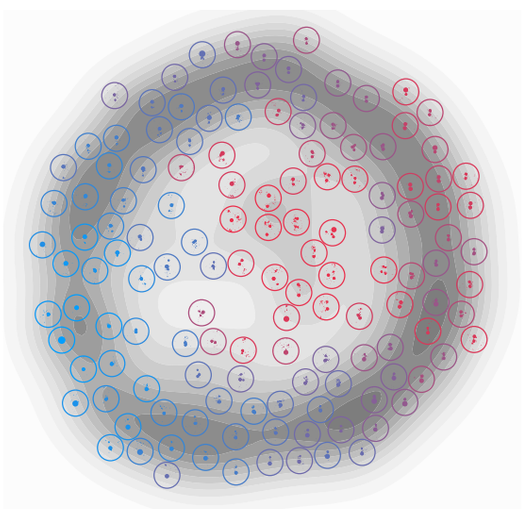
Patrick T. Komiske III
Massachusetts Institute of Technology
Center for Theoretical Physics

with Eric Metodiev and Jesse Thaler, 1902.02346

Elementary Particle Theory Seminar – Maryland Center for Fundamental Physics
University of Maryland, College Park
March 25, 2019



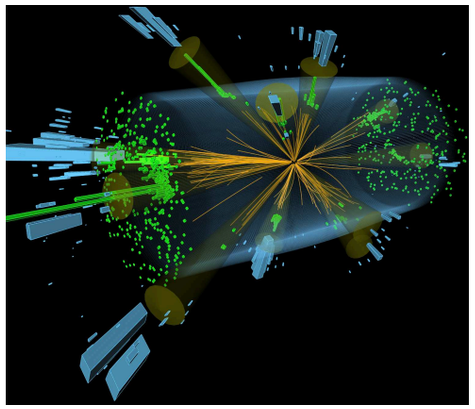




Collider Event Foundations

The Energy Mover's Distance

Particle Physics Applications



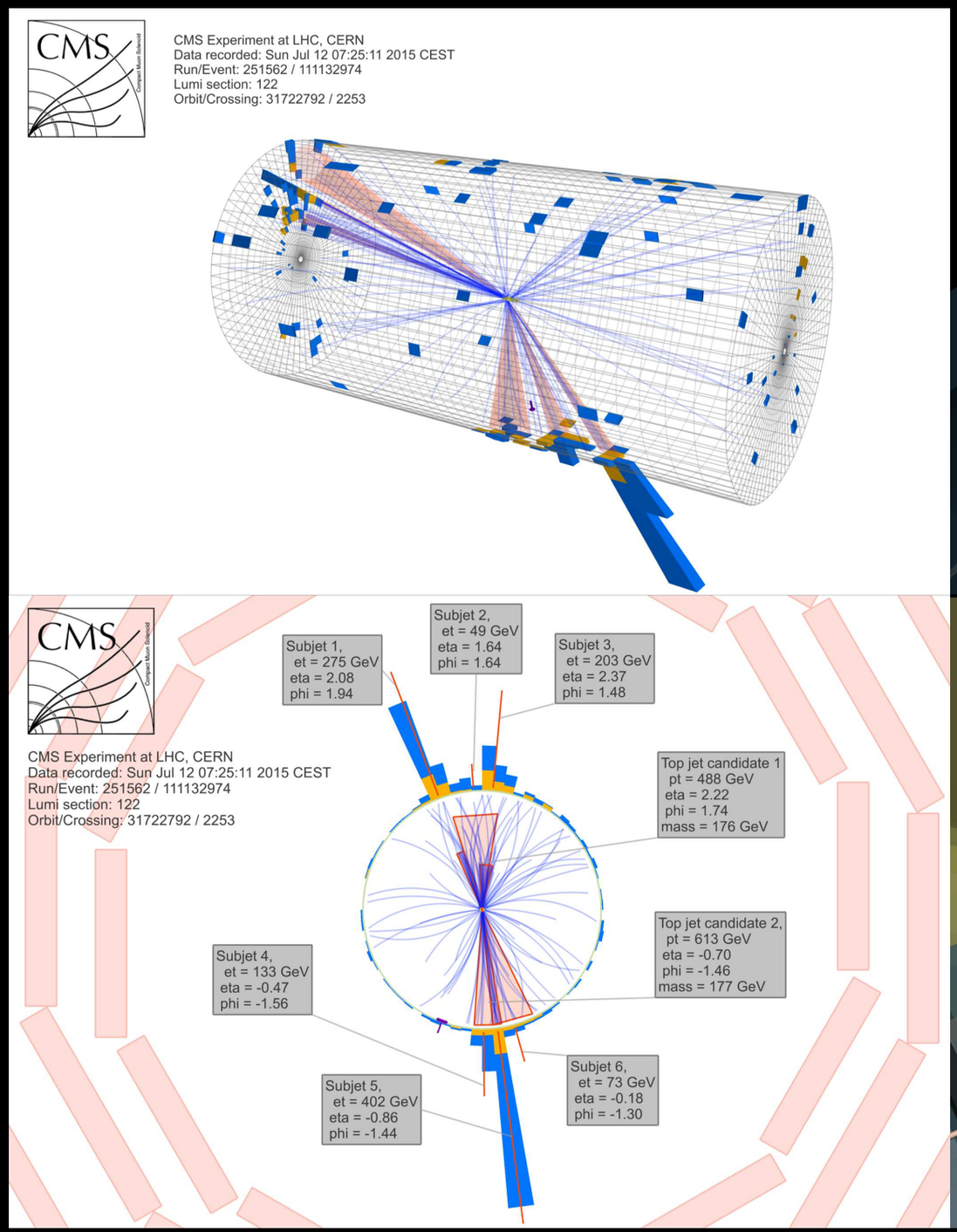
Collider Event Foundations

The Energy Mover's Distance

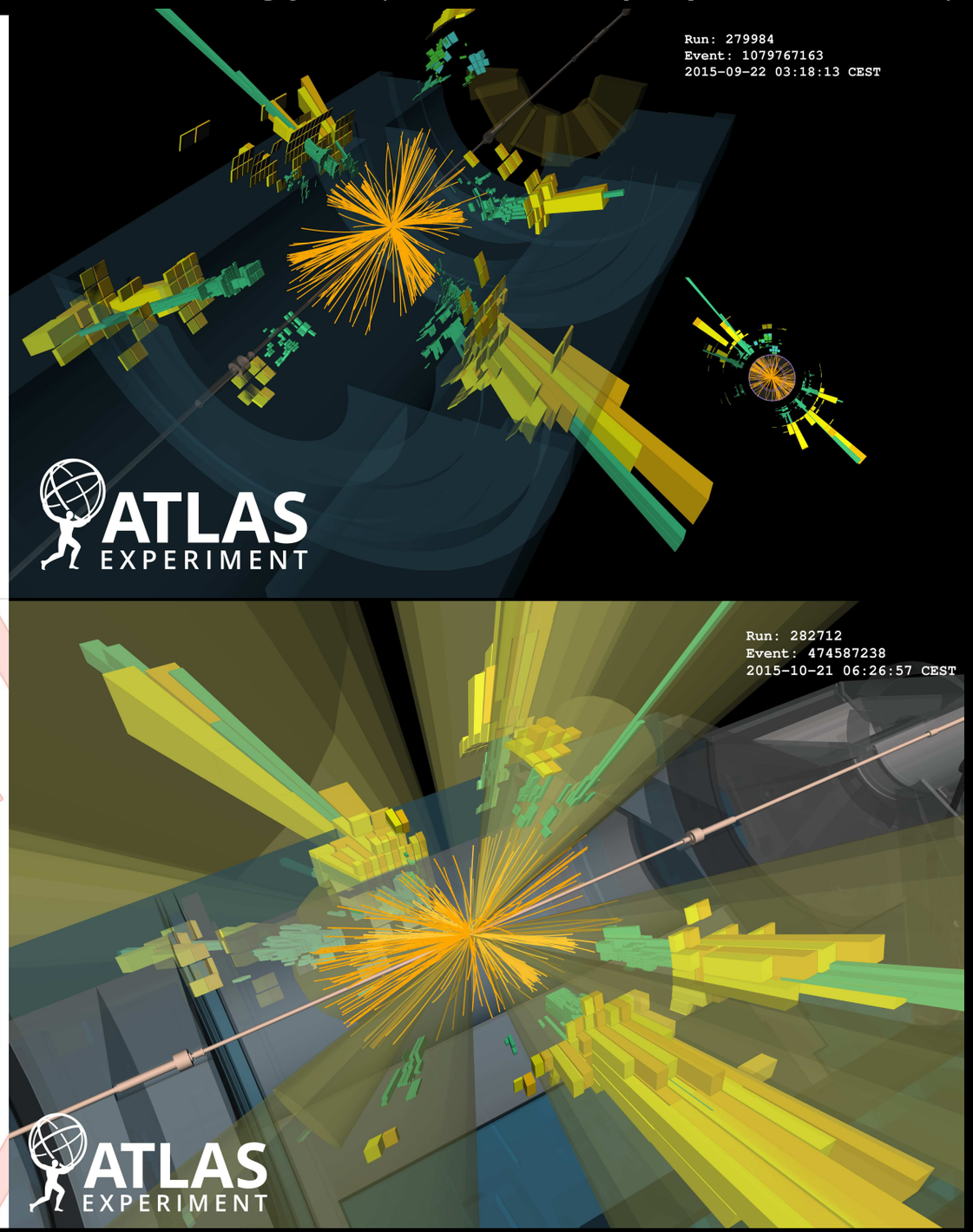
Particle Physics Applications

Fascinating Event Topologies at the LHC

New physics searches involve complicated final states including jets (collimated sprays of hadrons)



CMS hadronic $t\bar{t}$ event



ATLAS high jet multiplicity events

Jet Formation in Theory

Hard collision

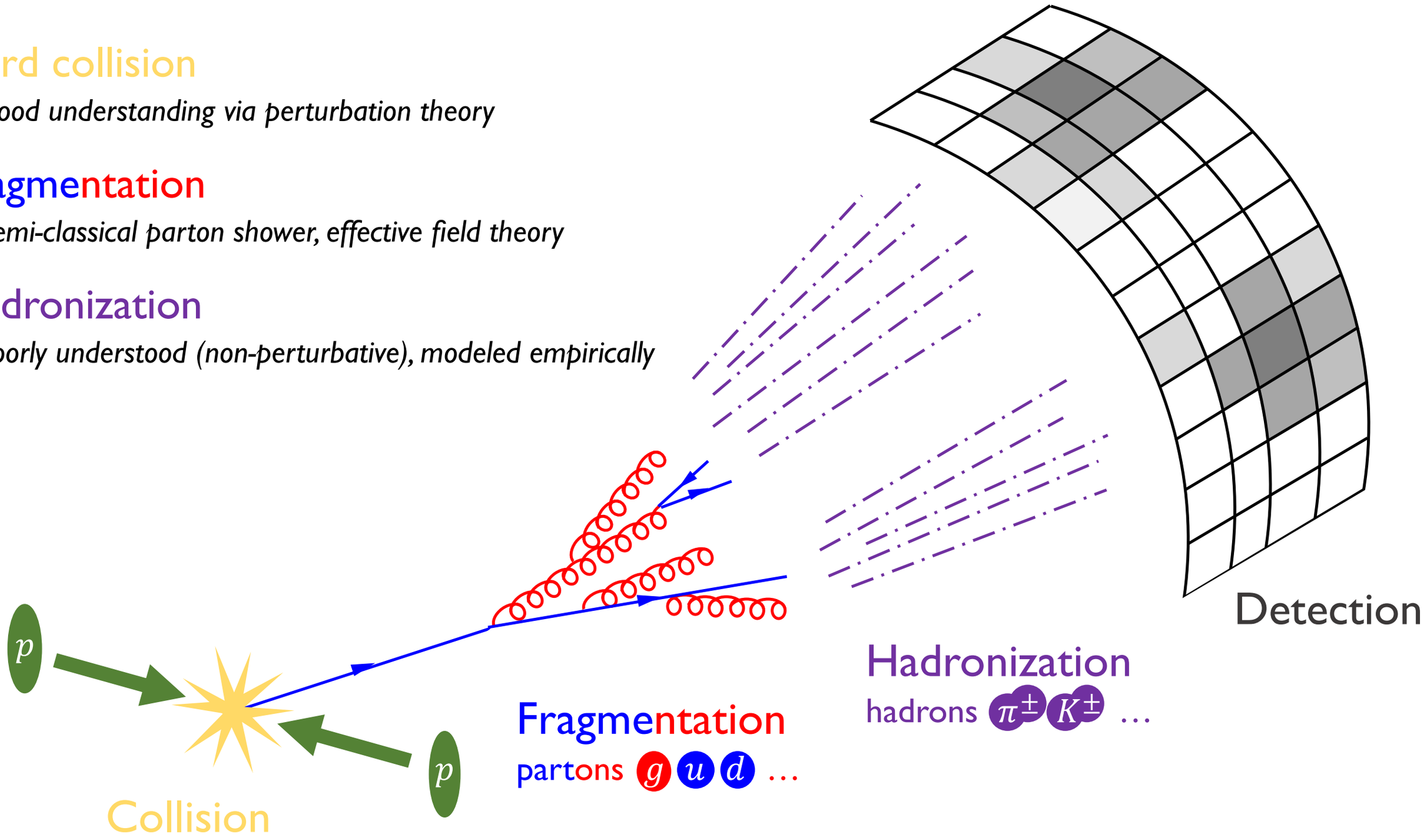
Good understanding via perturbation theory

Fragmentation

Semi-classical parton shower, effective field theory

Hadronization

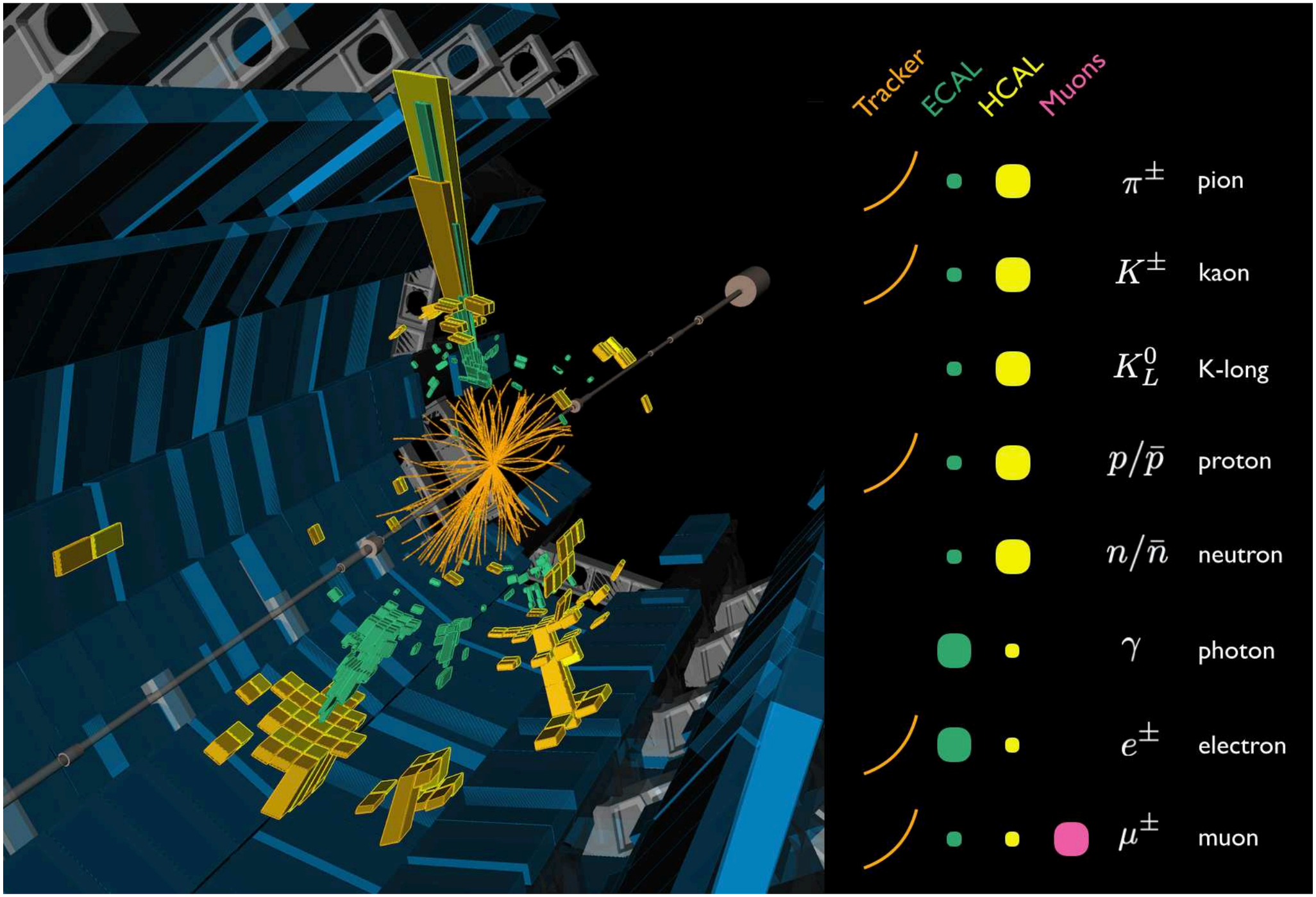
Poorly understood (non-perturbative), modeled empirically



Cartoon of jet formation as a multi-scale process

Diagram by Eric Metodiev

Jet Detection in Experiment



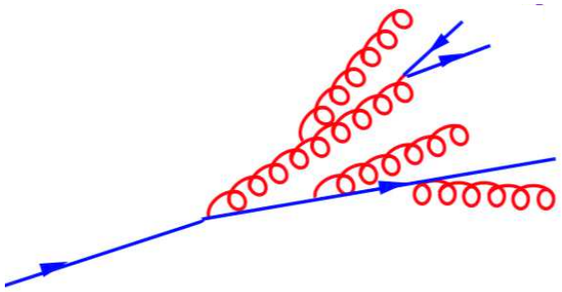
What information is both theoretically and experimentally robust?

Energy Flow

Events, Theoretically

$$|\mathcal{E}\rangle = |(p_1^\mu, \vec{q}_1); (p_2^\mu, \vec{q}_2); \ldots\rangle$$

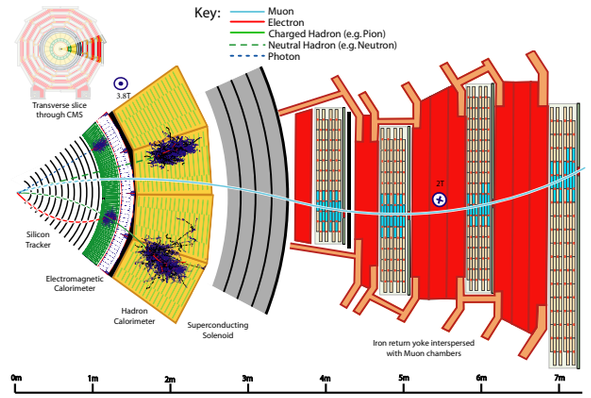
quantum state?



parton branching history?

Events, Experimentally

O(10 million) electrical signals?



set PF candidates?

The energy flow (distribution of energy) is robust to fragmentation, hadronization, detector effects

Energy Flow ⟷ Infrared and Collinear Safe Information

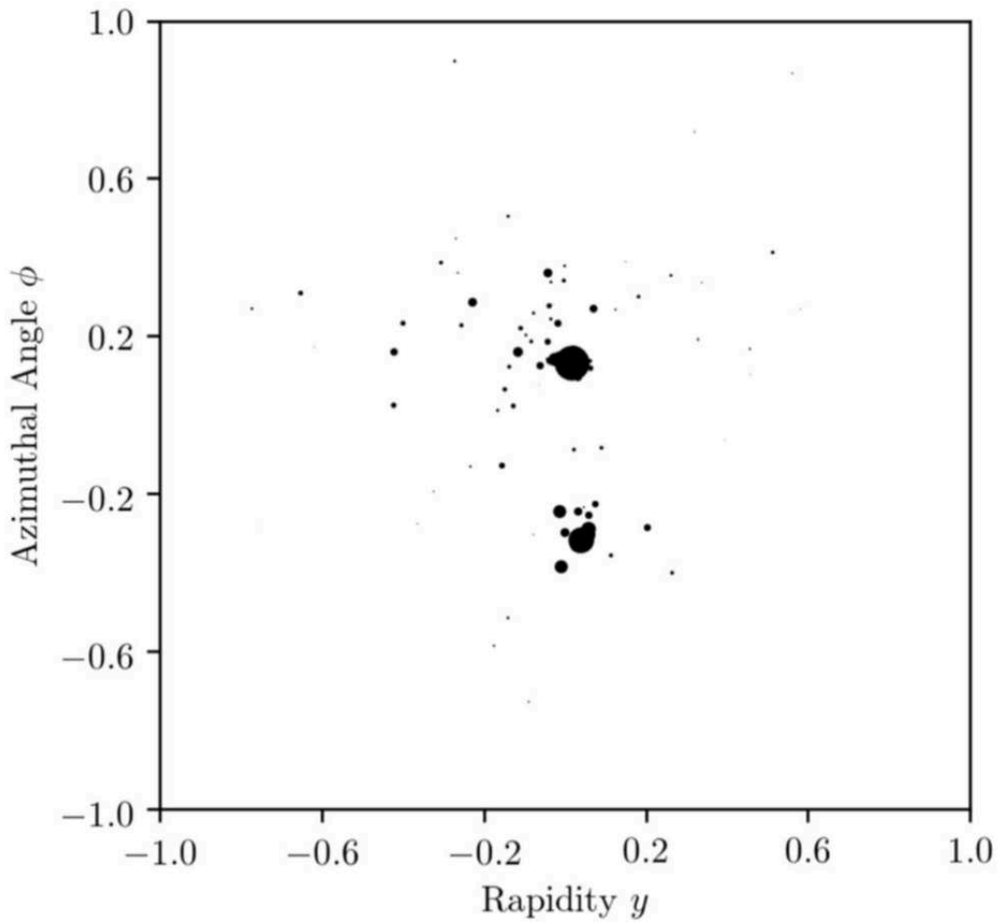
Energy Flow

Events, Theoretically | Events, Experimentally

$|\mathcal{E}\rangle$

T… n,

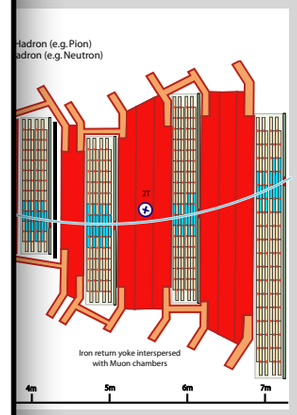
Energy Flow Representation of Jets



$$\rho(\hat{p}) = \sum_{i=1}^{M} E_i \delta(\hat{p} - \hat{p}_i)$$

Energy Flow Distribution | Energy (p_T) | Direction (y,φ)

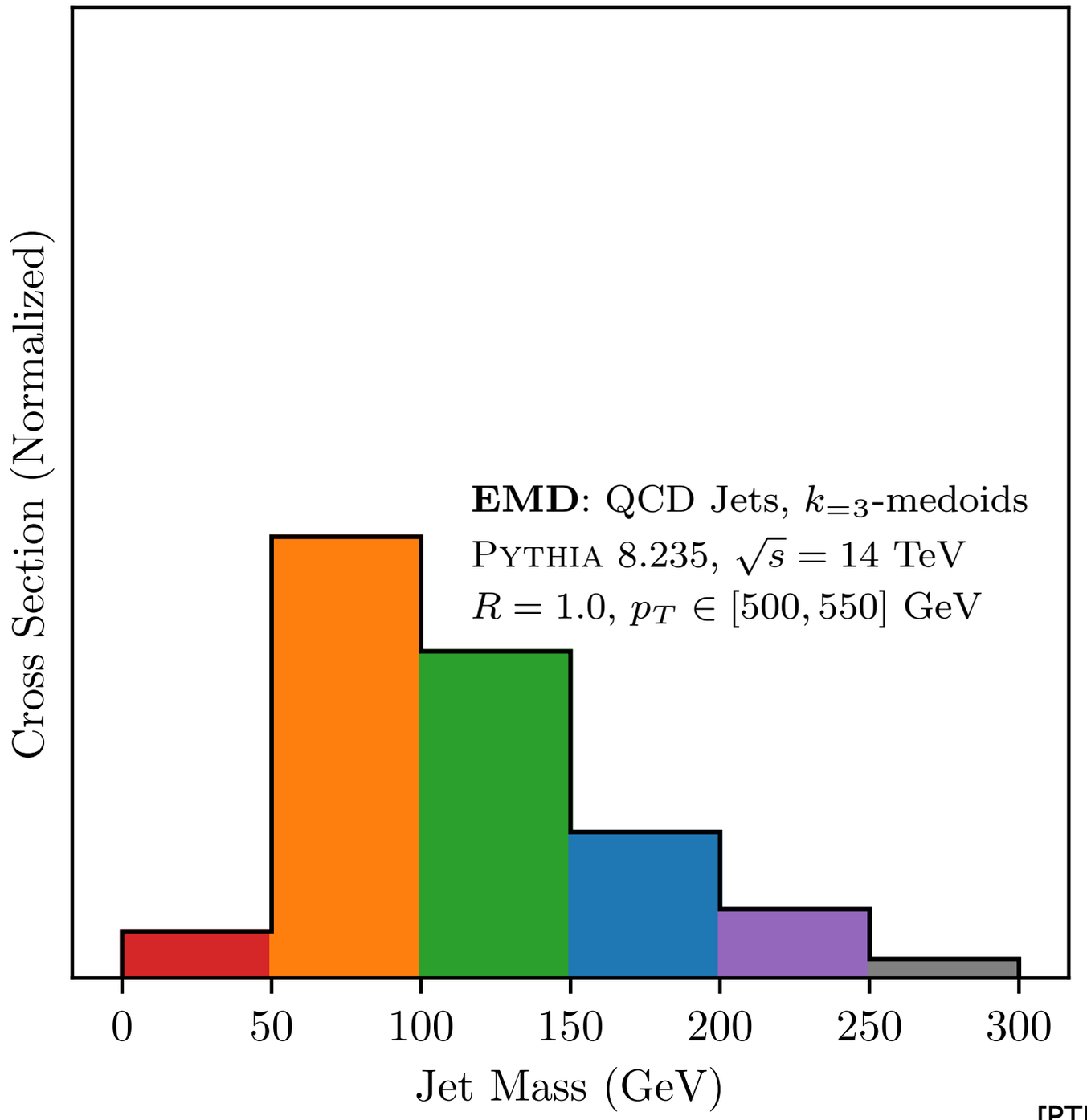
IRC safe
No flavor/charge information
Unpixelized geometry



didates?

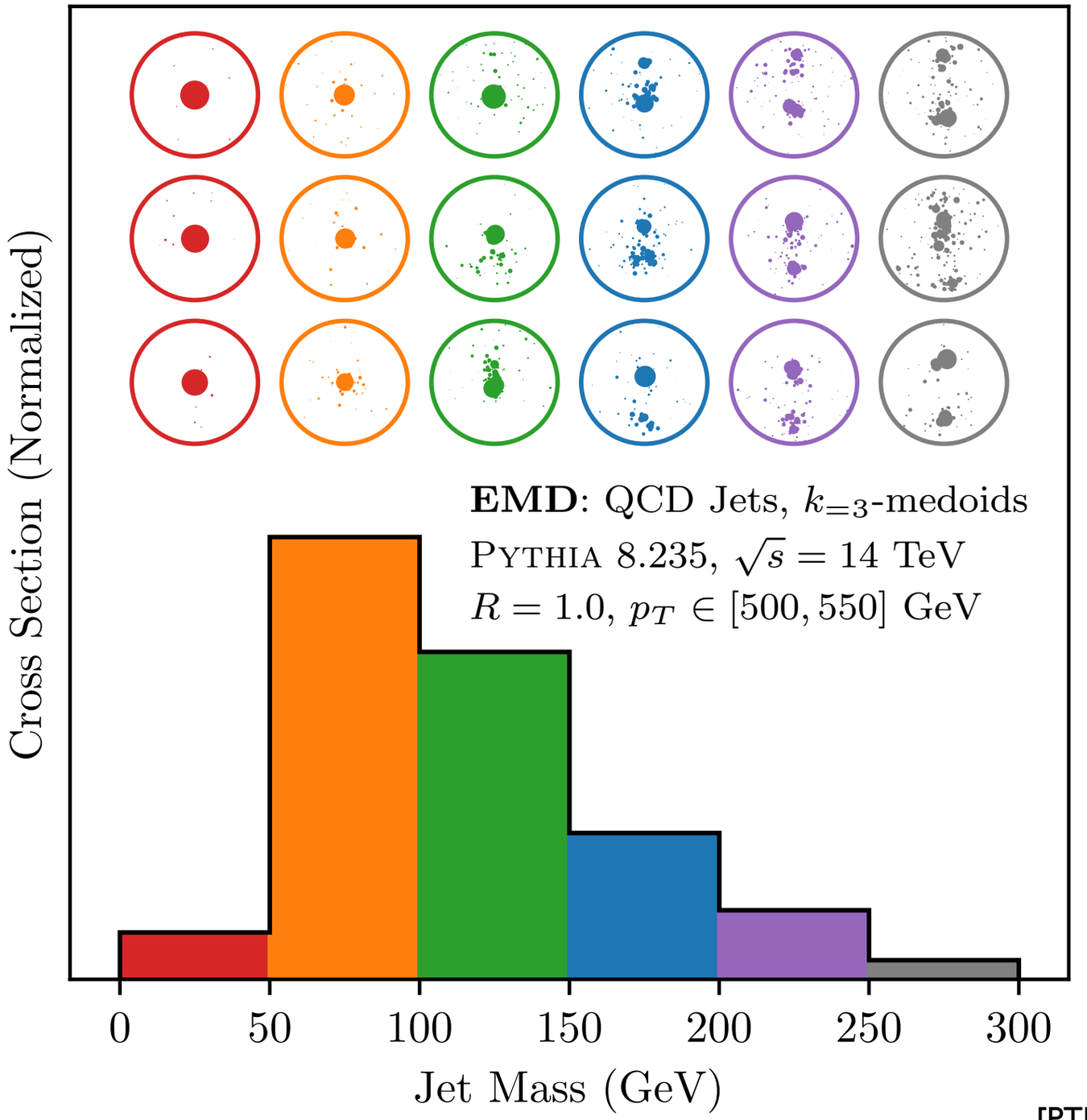
Energy Flow ⟷ Infrared and Collinear Safe Information

Particle Physics Histograms



[PTK, Metodiev, Thaler, 1902.02346]

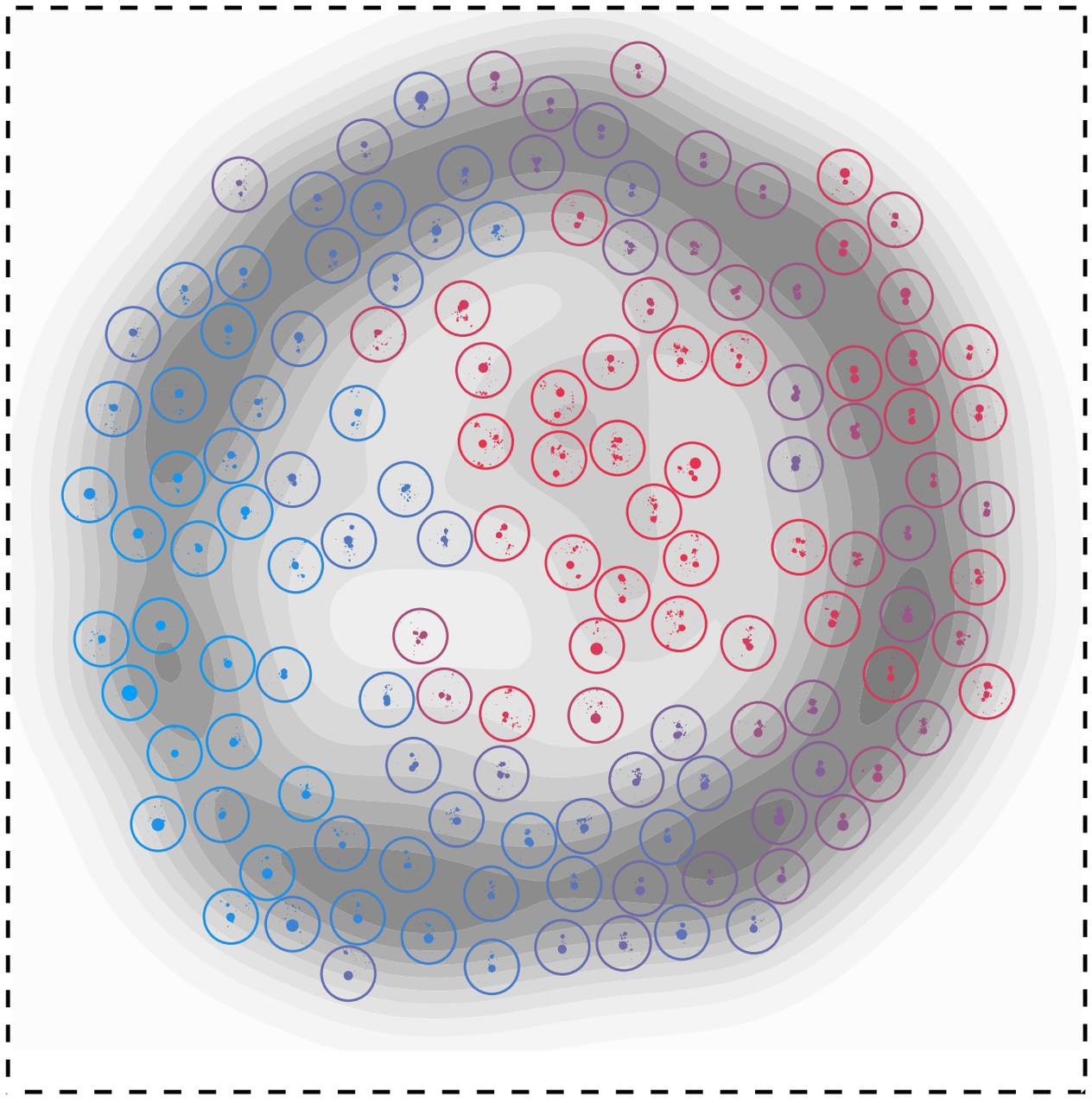
Particle Physics Histograms



Three "most" representative jets in each bin

[PTK, Metodiev, Thaler, 1902.02346]

Boosted *W* Jets

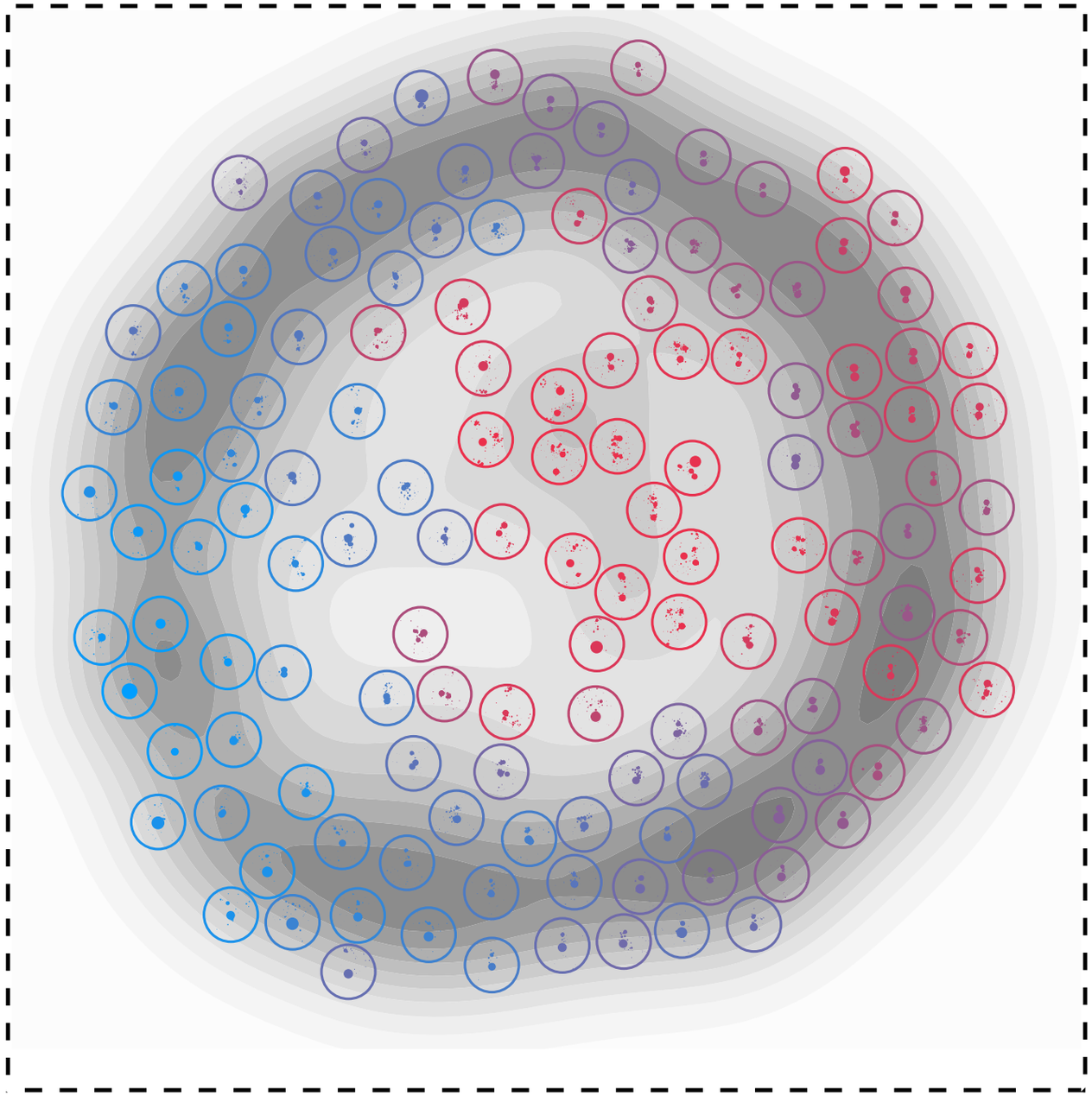


Abstract space of W jets

[PTK, Metodiev, Thaler, 1902.02346]

Boosted *W* Jets

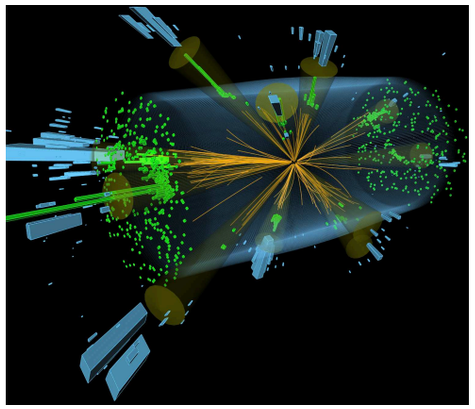
Gray contours represent the density of jets

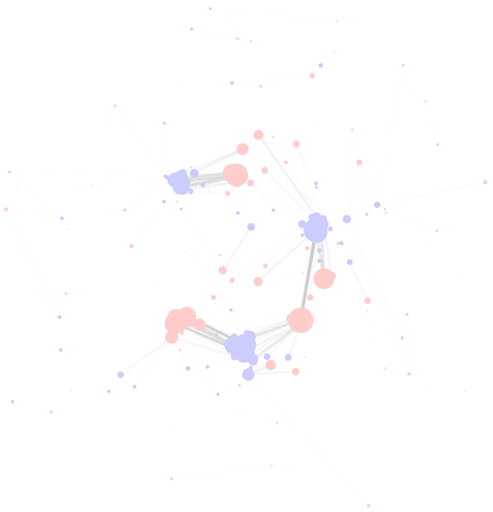


Each circle is a particular *W* jet

Abstract space of W jets

[PTK, Metodiev, Thaler, 1902.02346]



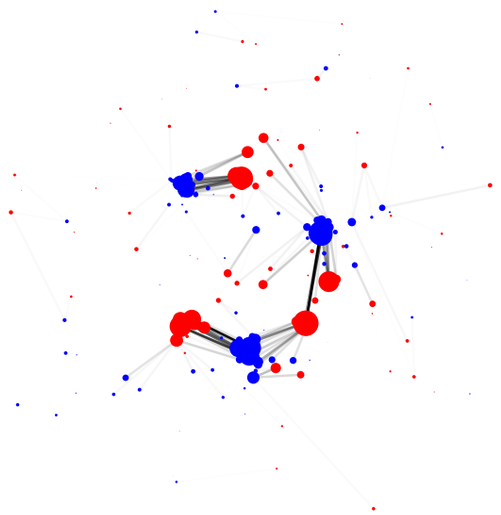


Collider Event Foundations

IRC-safe energy flow is theoretically and experimentally robust

The Energy Mover's Distance

Particle Physics Applications



Collider Event Foundations

IRC-safe energy flow is theoretically and experimentally robust

The Energy Mover's Distance

Particle Physics Applications

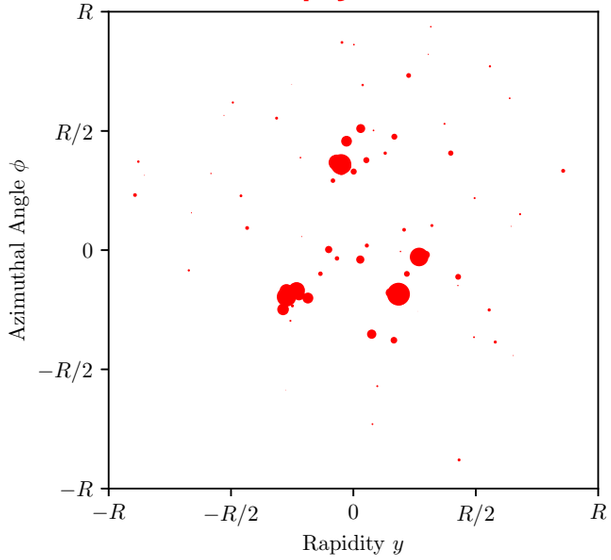
The Earth Mover's Distance

A metric on normalized distributions in a space with a ground distance measure

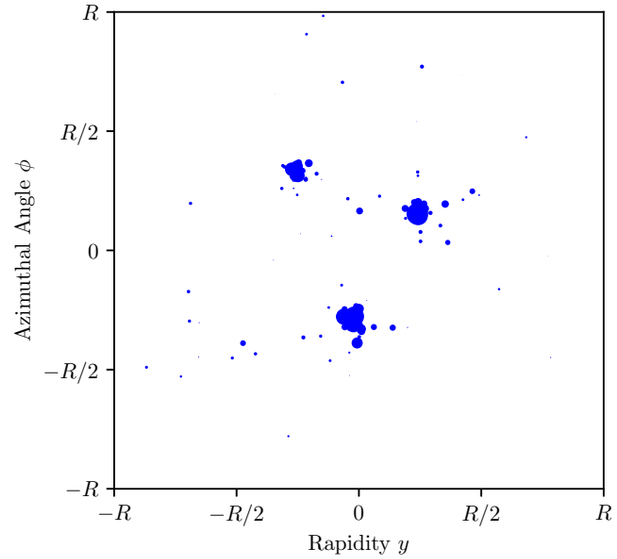
↳ symmetric, non-negative, triangle inequality, zero iff identical

The minimum "work" (stuff x distance) required to transport supply to demand

Top Jet 1



Top Jet 2



Related to *optimal transport* theory – commonly used as a metric on the space of images

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

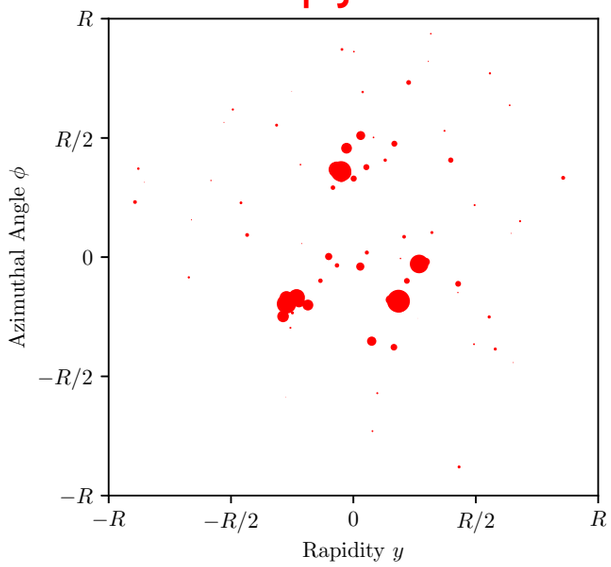
The Earth Mover's Distance

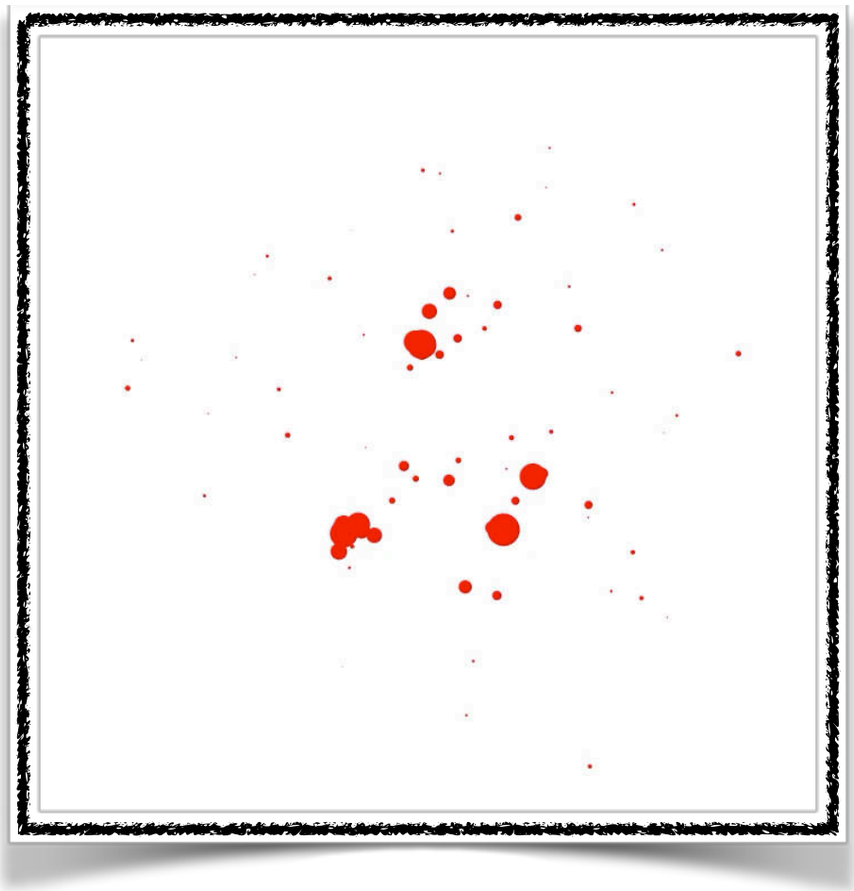
A metric on normalized distributions in a space with a ground distance measure

↳ symmetric, non-negative, triangle inequality, zero iff identical

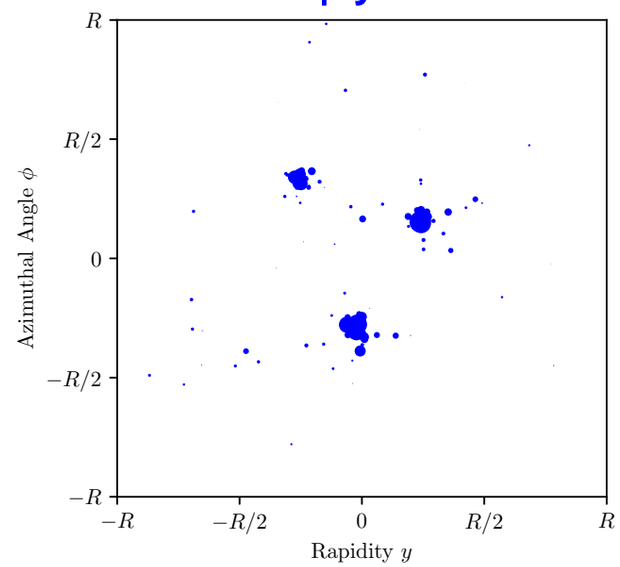
The minimum "work" (stuff x distance) required to transport supply to demand

Top Jet 1





Top Jet 2



Related to *optimal transport* theory – commonly used as a metric on the space of images

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

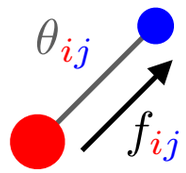
The Energy Mover's Distance

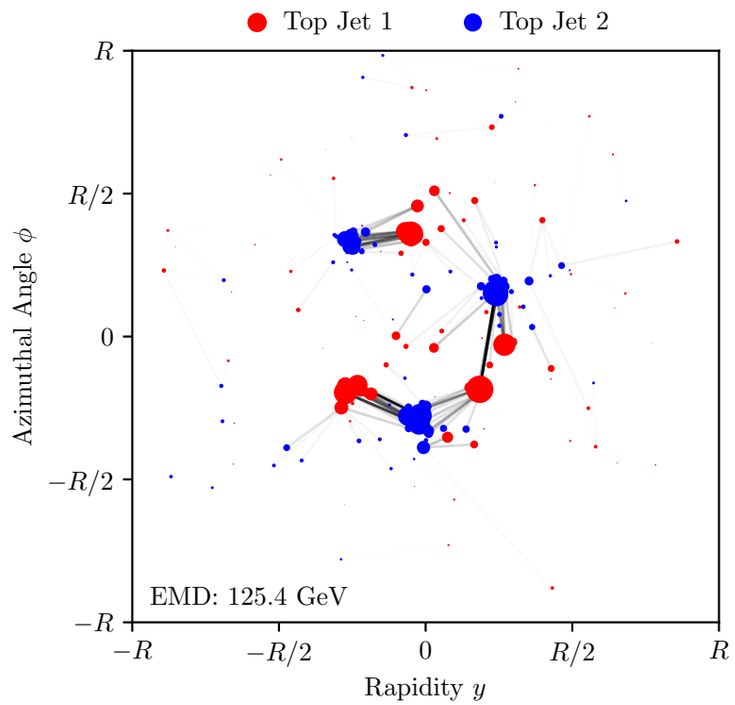
[PTK, Metodiev, Thaler, 1902.02346]

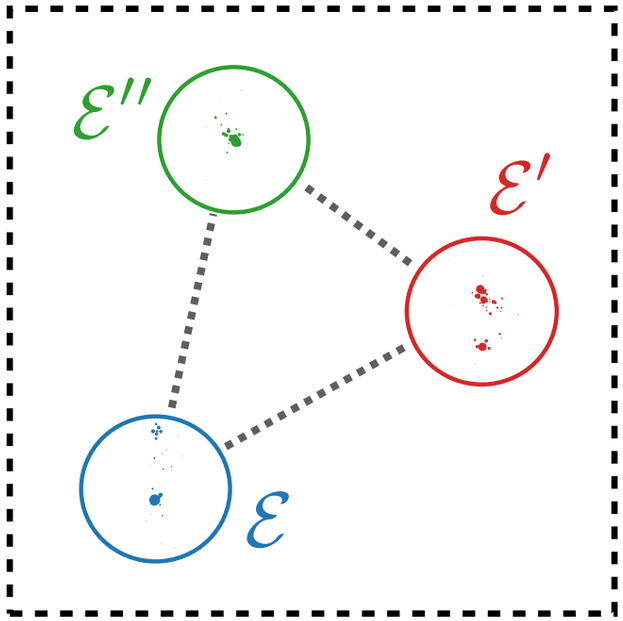
EMD between energy flows defines a metric on the space of events

$$\mathrm{EMD}(\mathcal{E}, \mathcal{E}') = \underbrace{\min_{\{f_{ij} \geq 0\}} \sum_i \sum_j f_{ij} \frac{\theta_{ij}}{R}}_{\text{Cost of optimal transport}} + \underbrace{\left| \sum_i E_i - \sum_j E'_j \right|}_{\text{Cost of energy creation}}$$

$$\sum_j f_{ij} \leq E_i, \quad \sum_i f_{ij} \leq E'_j, \quad \sum_{ij} f_{ij} = \min\left(\sum_i E_i, \sum_j E'_j\right)$$





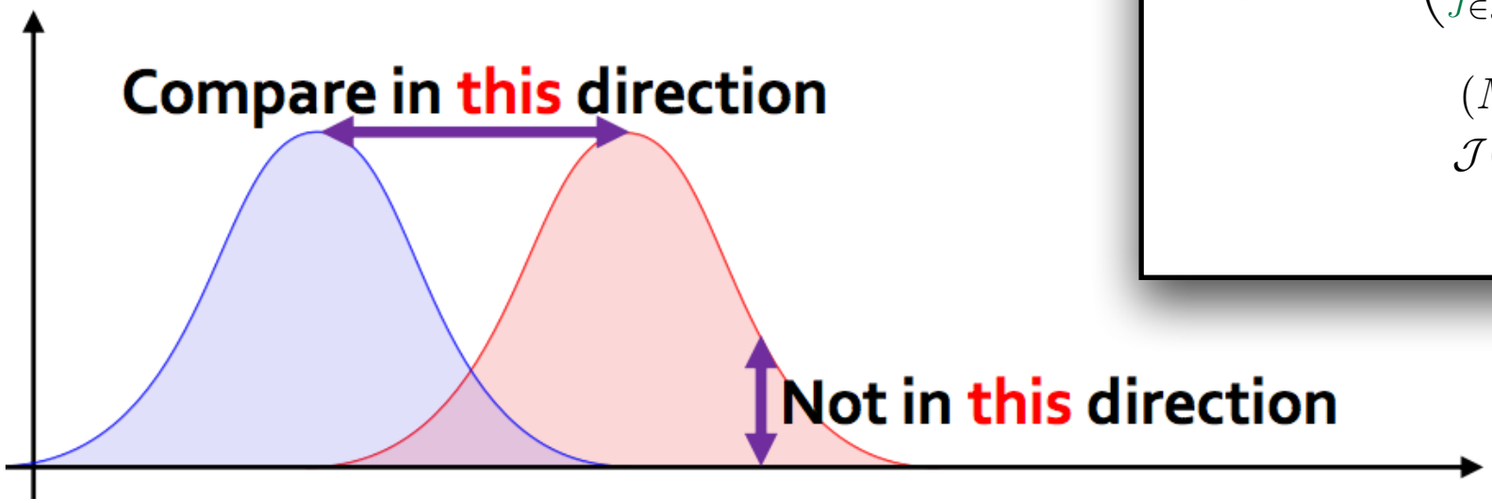


Triangle inequality satisfied for $R \geq d_{\max}/2$

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

Mathematics of the Earth Mover's Distance

p-Wasserstein distance is a metric on probability distributions



[figure from Kun, Math ∩ Programming]

$$W_p(\mu,\nu) \equiv \left(\inf_{J \in \mathcal{J}(\mu,\nu)} \int_{M\times M} d(x,y)^p \, \mathrm{d}J(x,y) \right)^{1/p}$$

$(M,\, d)$, metric space

$\mathcal{J}(\mu,\nu)$, space of joint distributions with marginals $\mu,\, \nu$

Earth mover's distance is 1-Wasserstein metric on discrete distributions

Recent use in Machine Learning

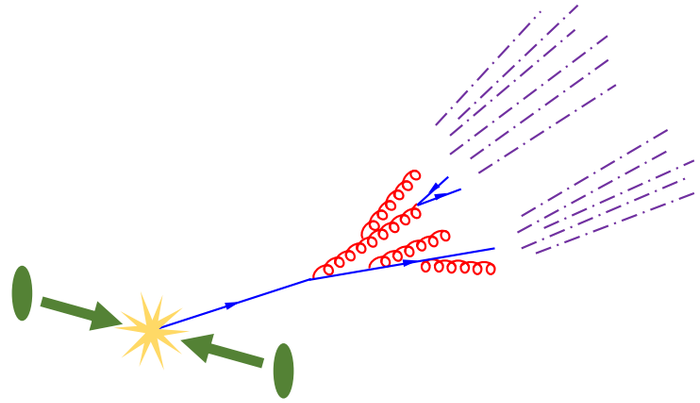
Wasserstein Generative Adversarial Networks

[Arjovsky, Chintala, Bottou, 1701.07875;
in particle physics:
- Erdmann, Geiger, Glombitza, Schmidt, 1802.03325
- Erdmann, Glombitza, Quast, 1807.01954]

Wasserstein(-Wasserstein) Autoencoders

[Tolstikhin, Bousquet, Gelly, Shoelkopf, 1711.01558]
[Zhang, Gao, Jiao, Liu, Wang, Yang, 1902.09323]

Visualizing Jet Formation – QCD Jets



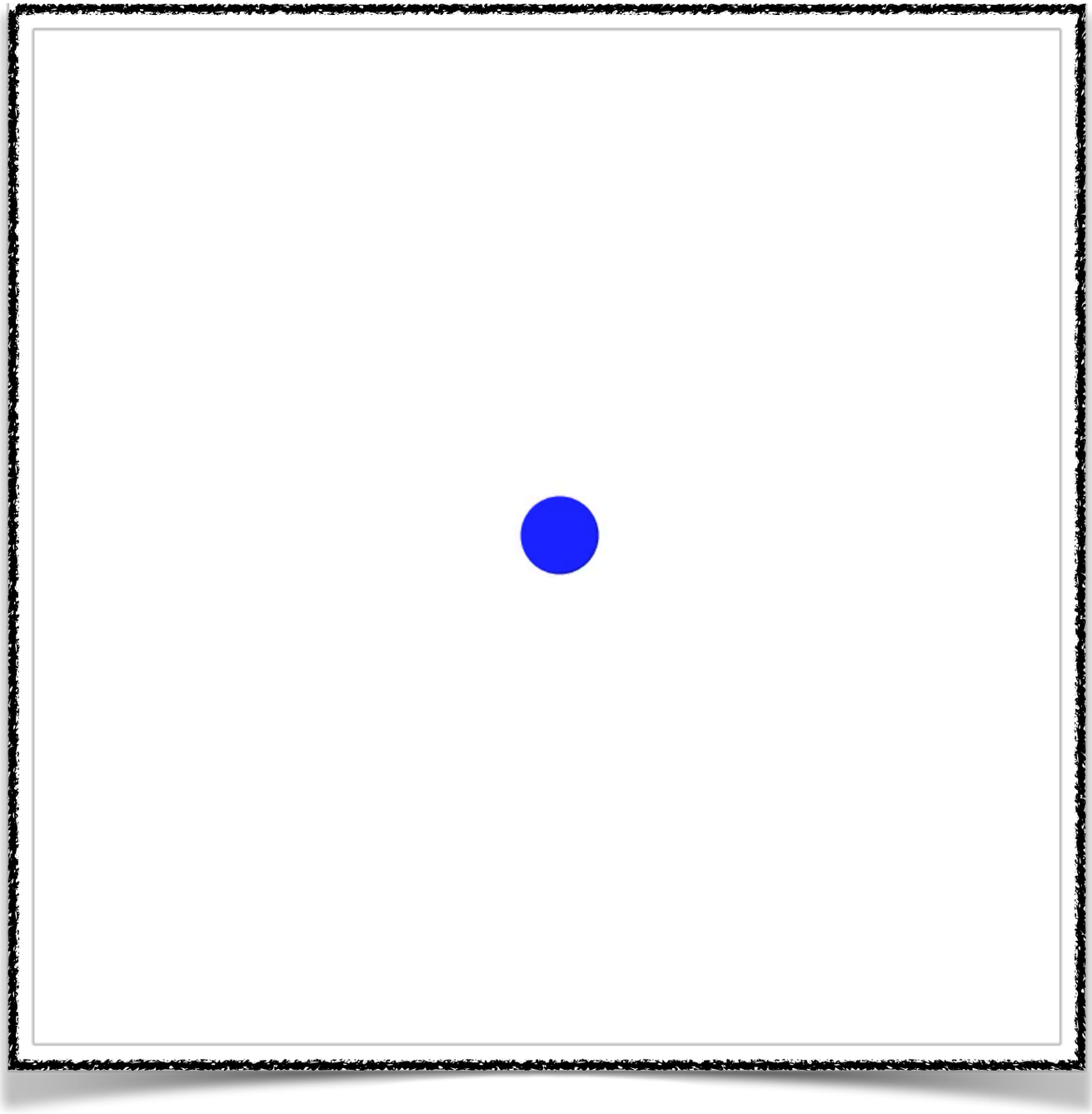
500 GeV

fragmentation

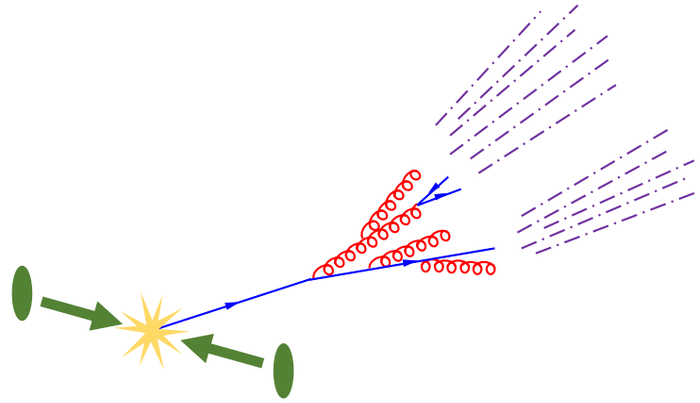
EMD: 111.6 GeV

hadronization

EMD: 18.1 GeV



Visualizing Jet Formation – QCD Jets



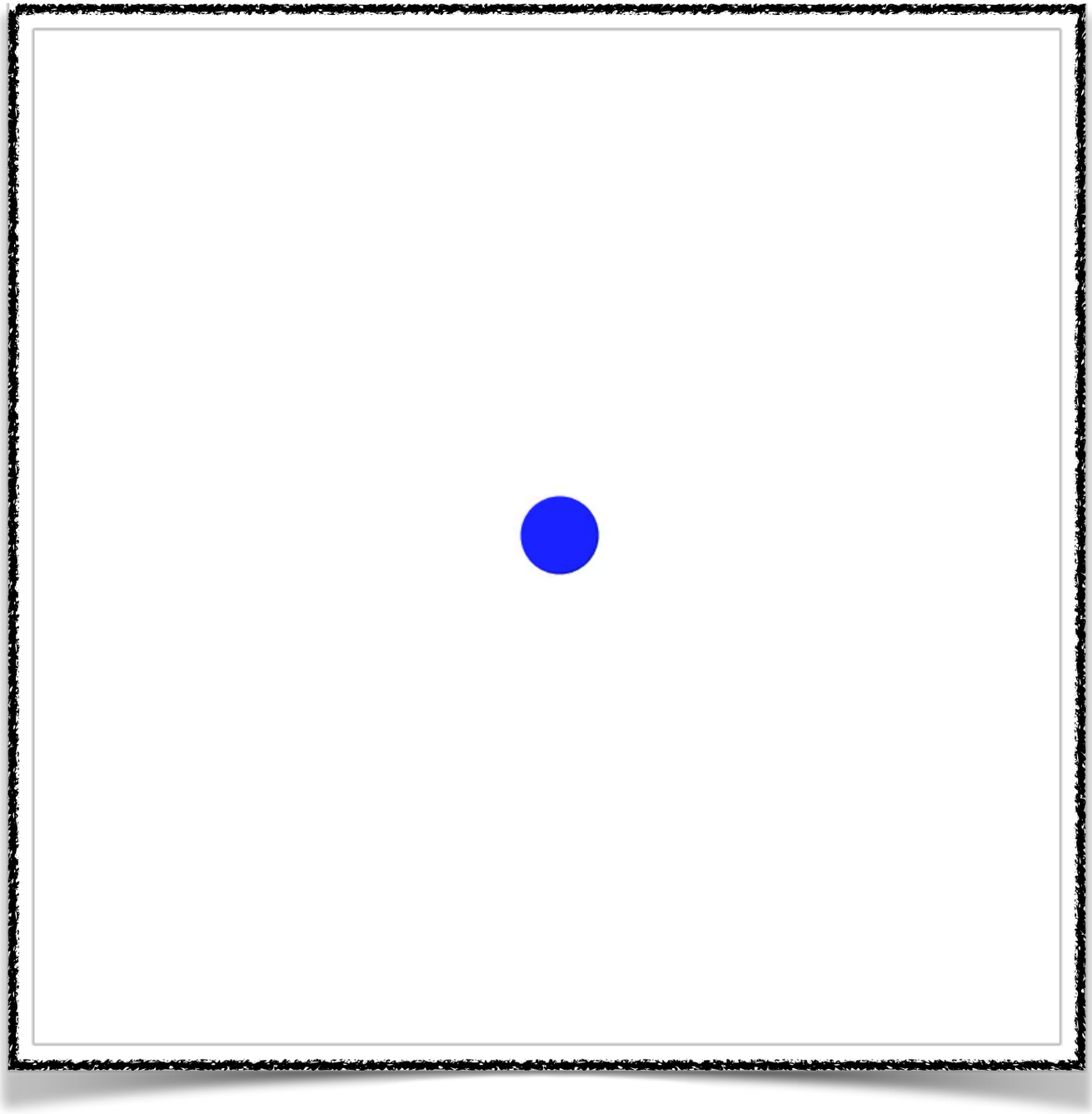
500 GeV

fragmentation

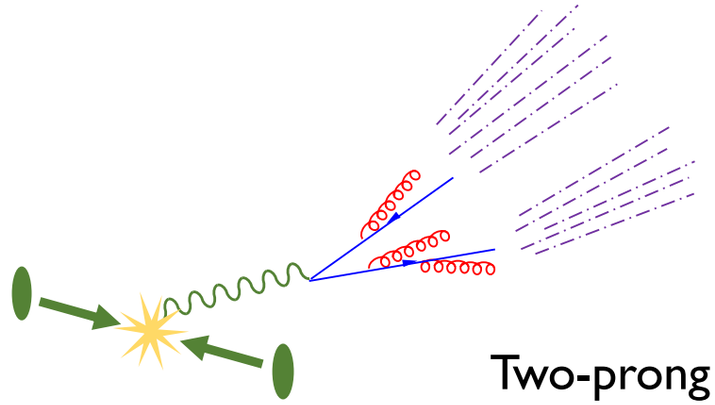
EMD: 111.6 GeV

hadronization

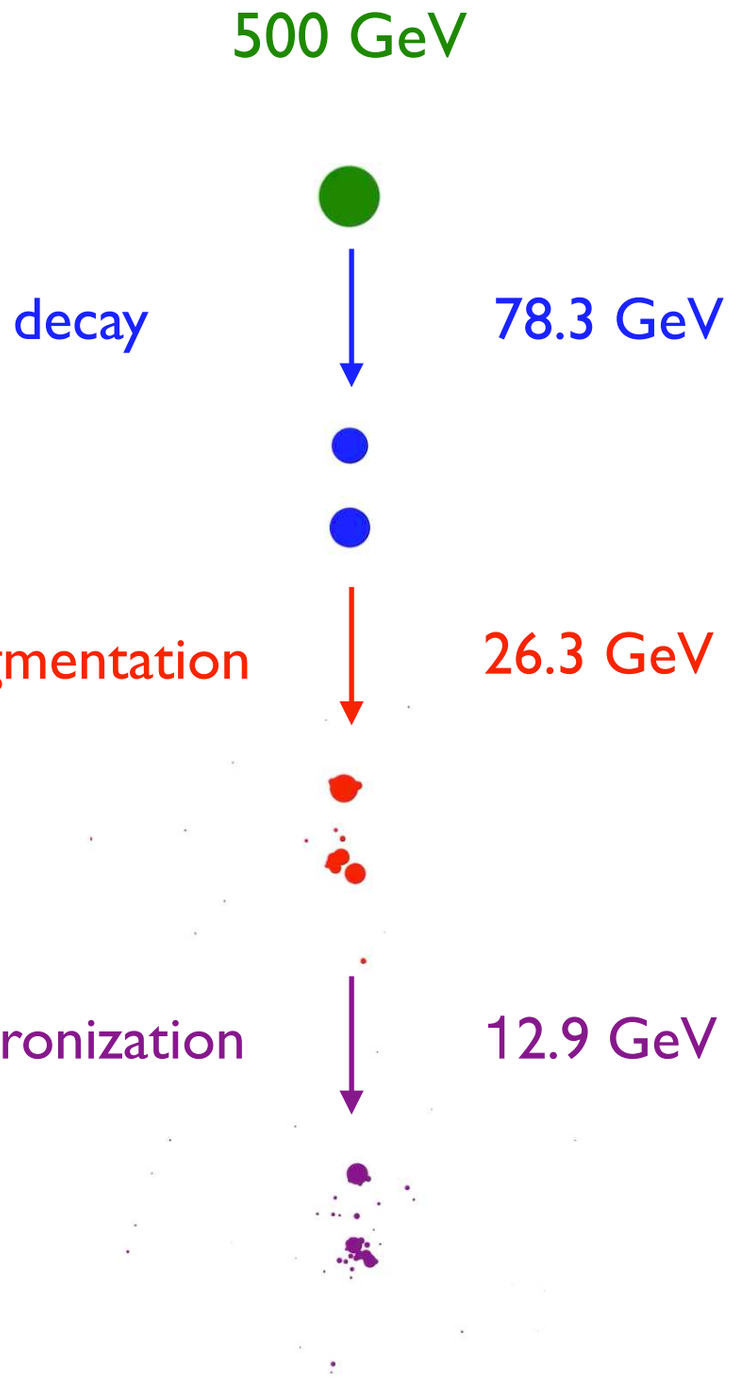
EMD: 18.1 GeV

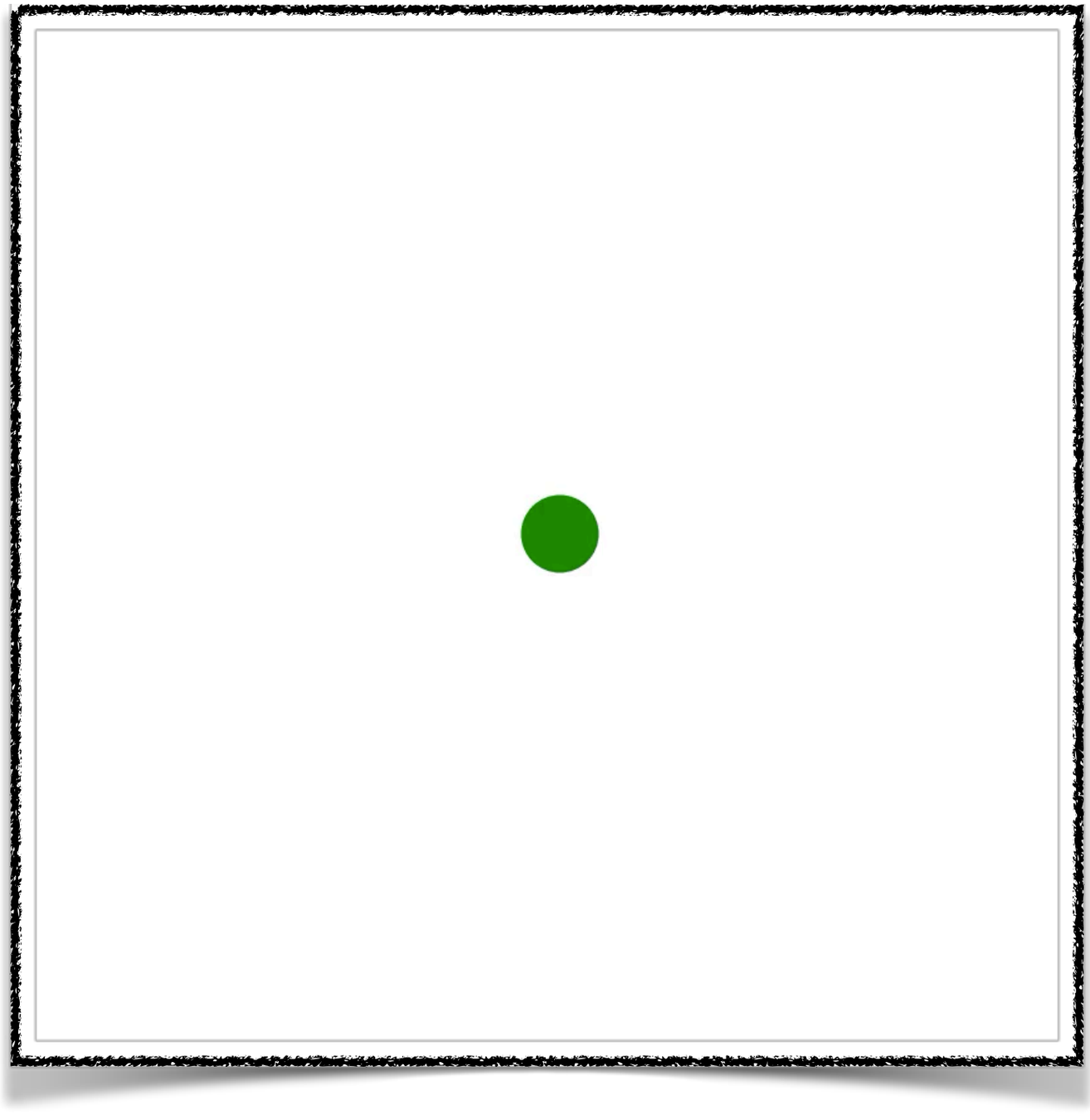


Visualizing Jet Formation – *W* Jets

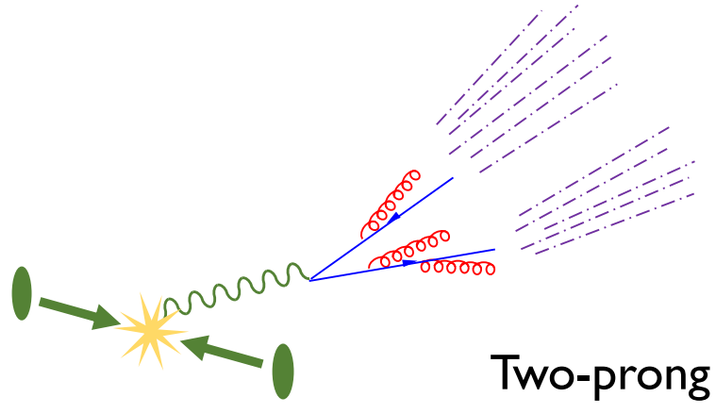


Two-prong decay

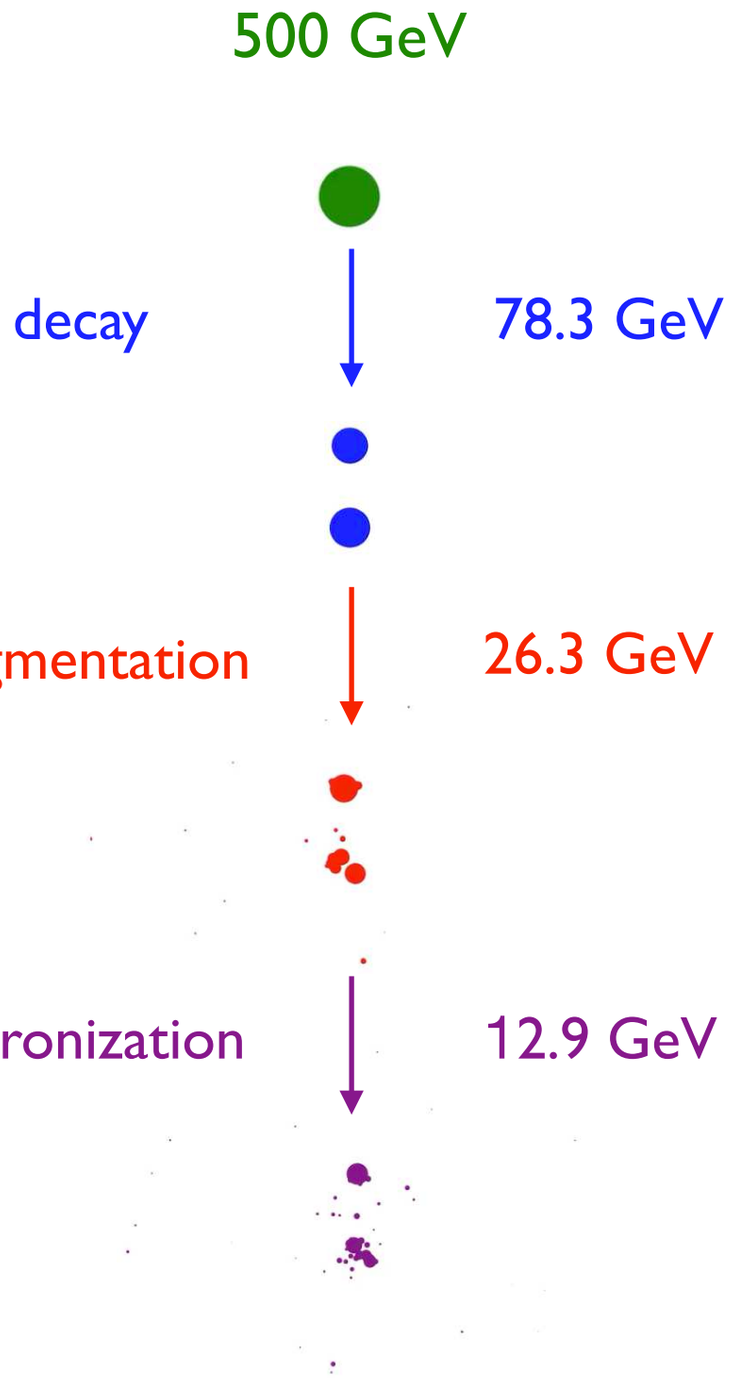


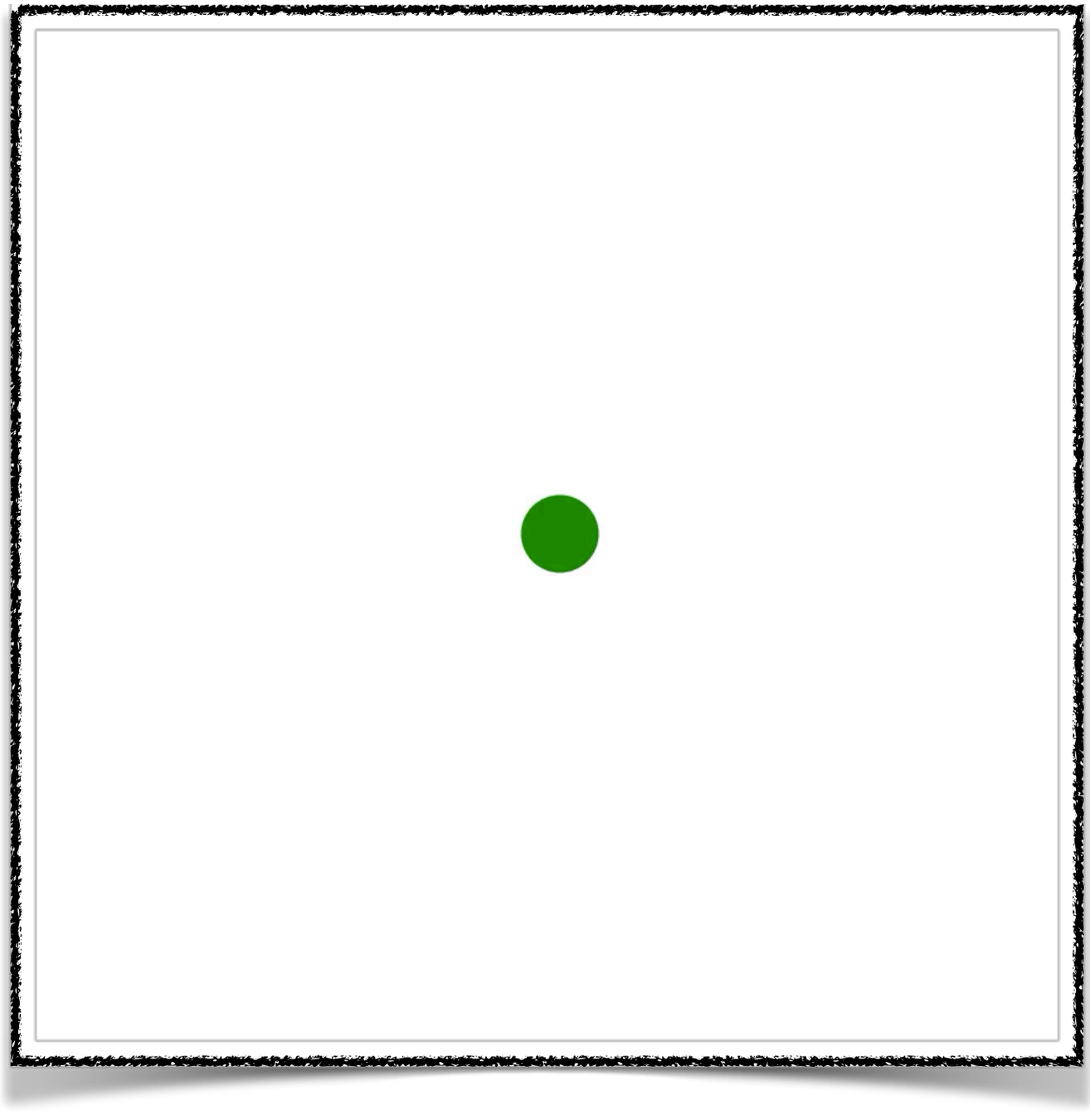


Visualizing Jet Formation – *W* Jets

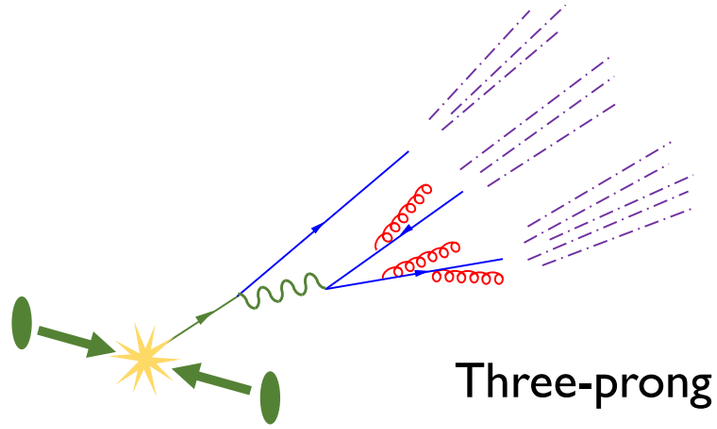


Two-prong decay

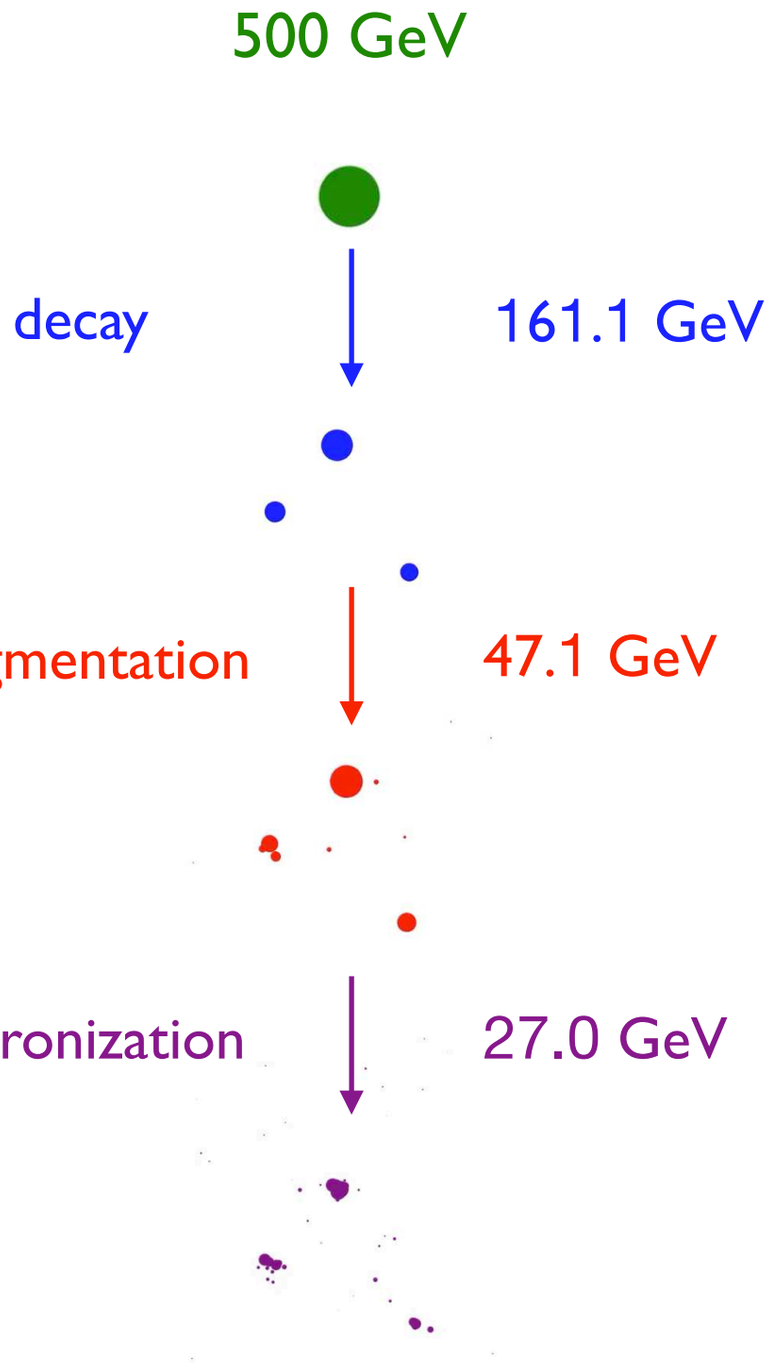


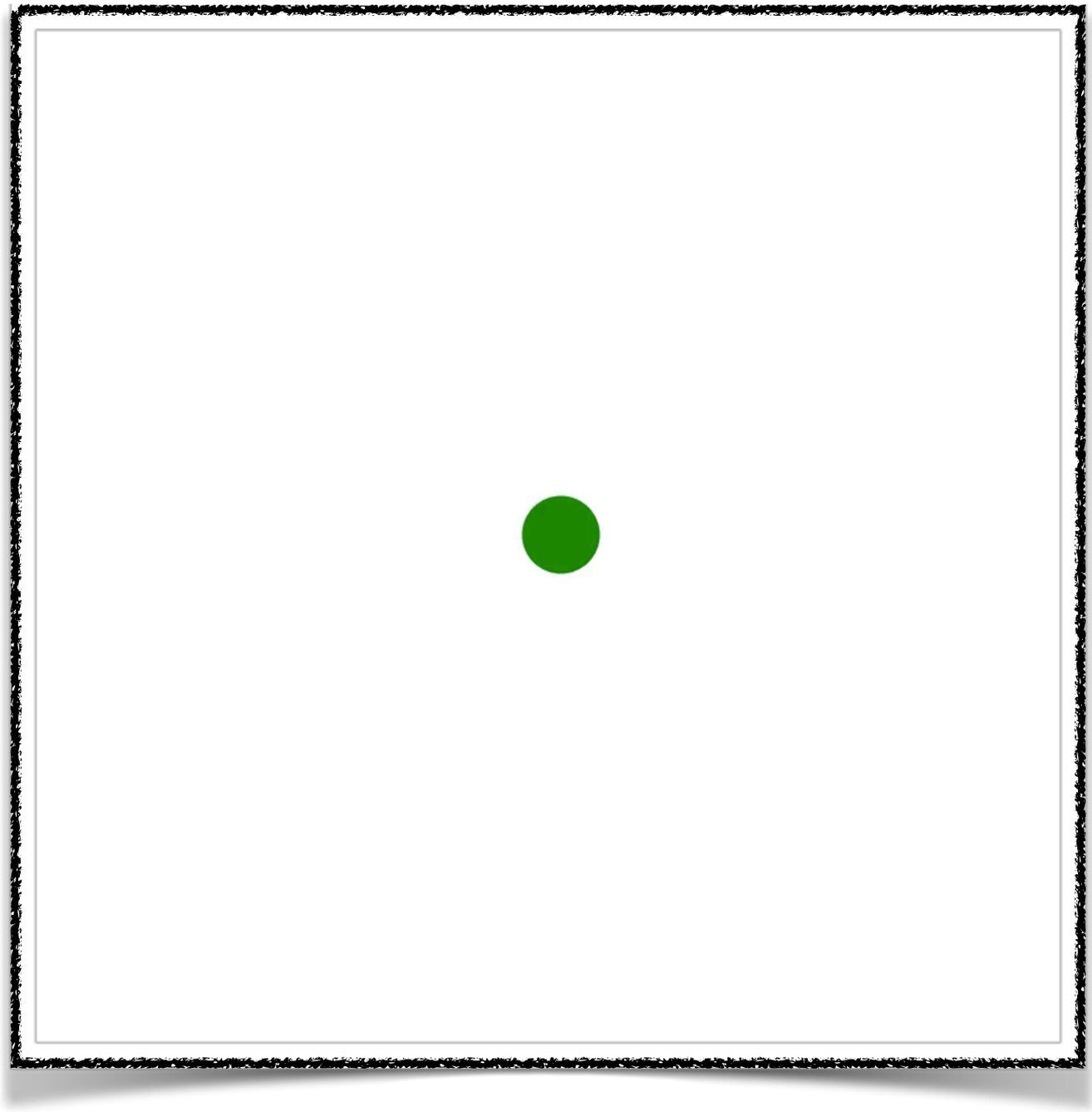


Visualizing Jet Formation – Top Jets

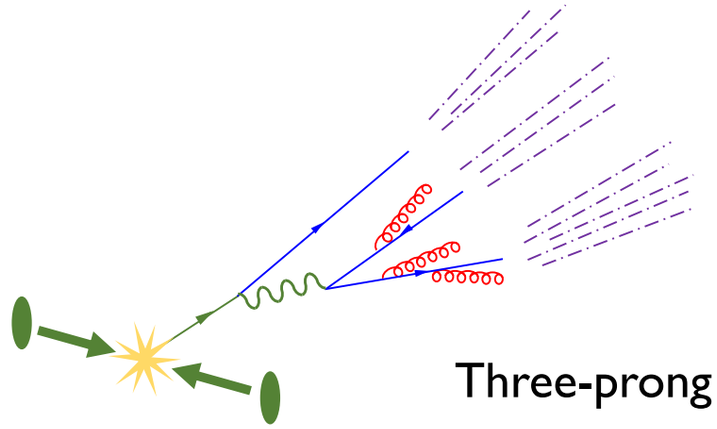


Three-prong decay

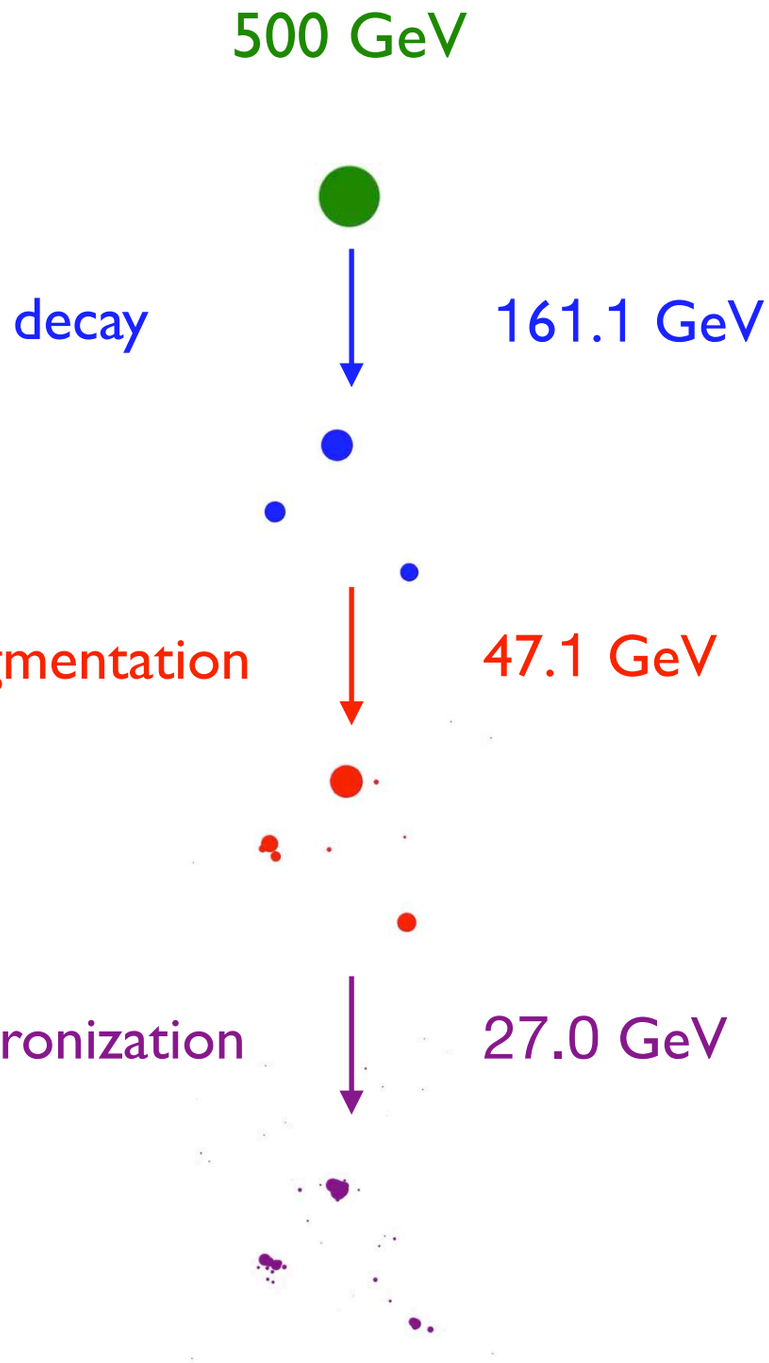


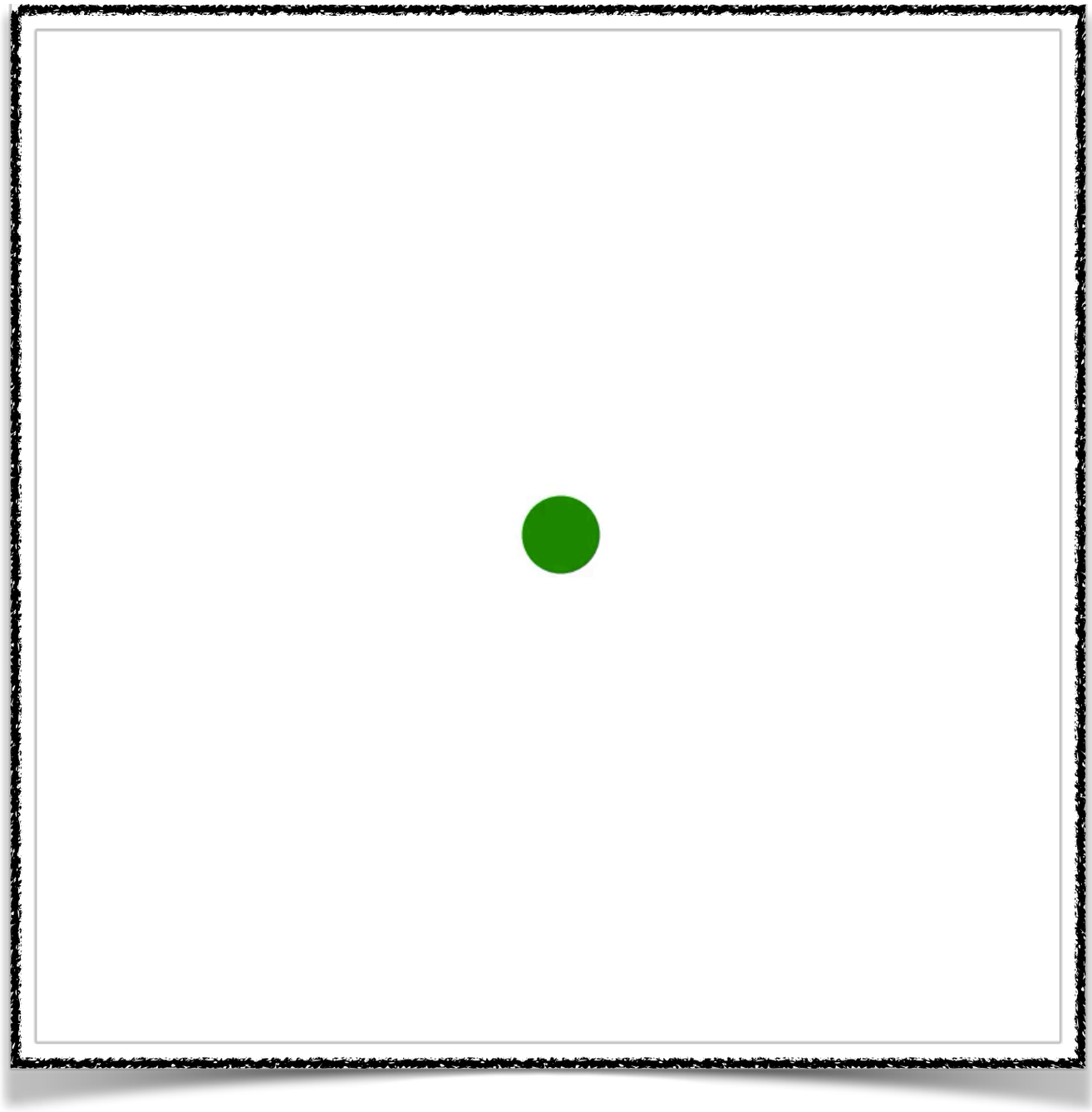


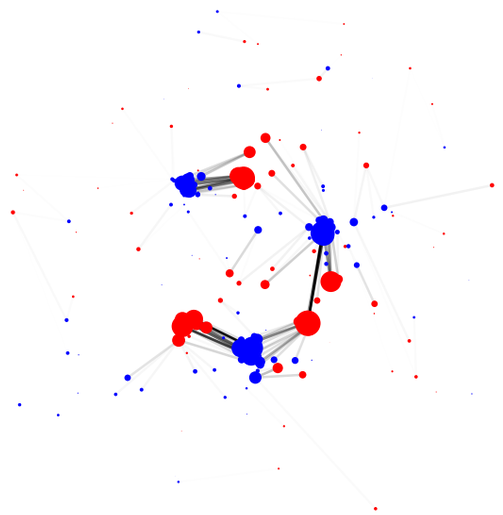
Visualizing Jet Formation – Top Jets



Three-prong decay







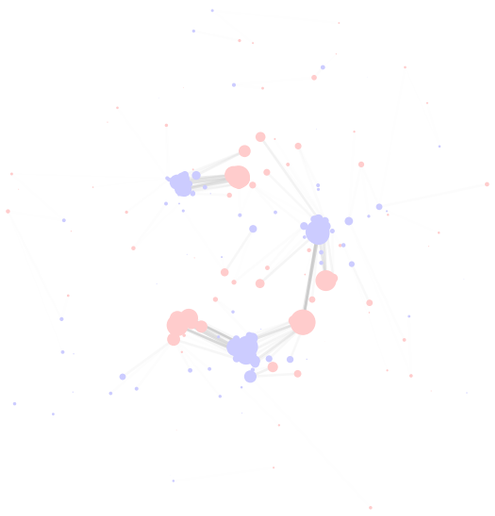
Collider Event Foundations

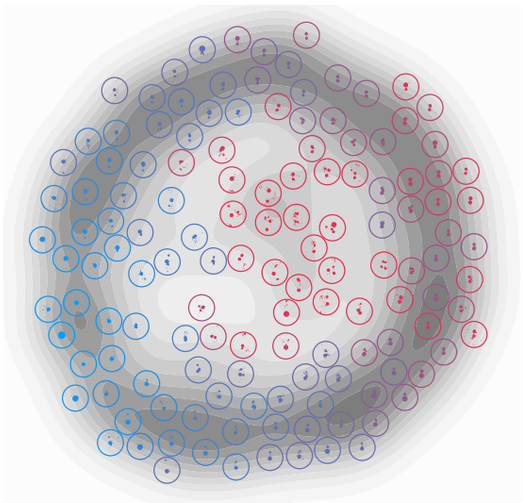
IRC-safe energy flow is theoretically and experimentally robust

The Energy Mover's Distance

Quantifies the difference in energy flow between events

Particle Physics Applications





Collider Event Foundations

IRC-safe energy flow is theoretically and experimentally robust

The Energy Mover's Distance

Quantifies the difference in energy flow between events

Particle Physics Applications

Quantifying Event Modifications

[PTK, Metodiev, Thaler, 1902.02346]

Mathematics

1-Wasserstein metric bounds the difference in expectation values between distributions

⇨

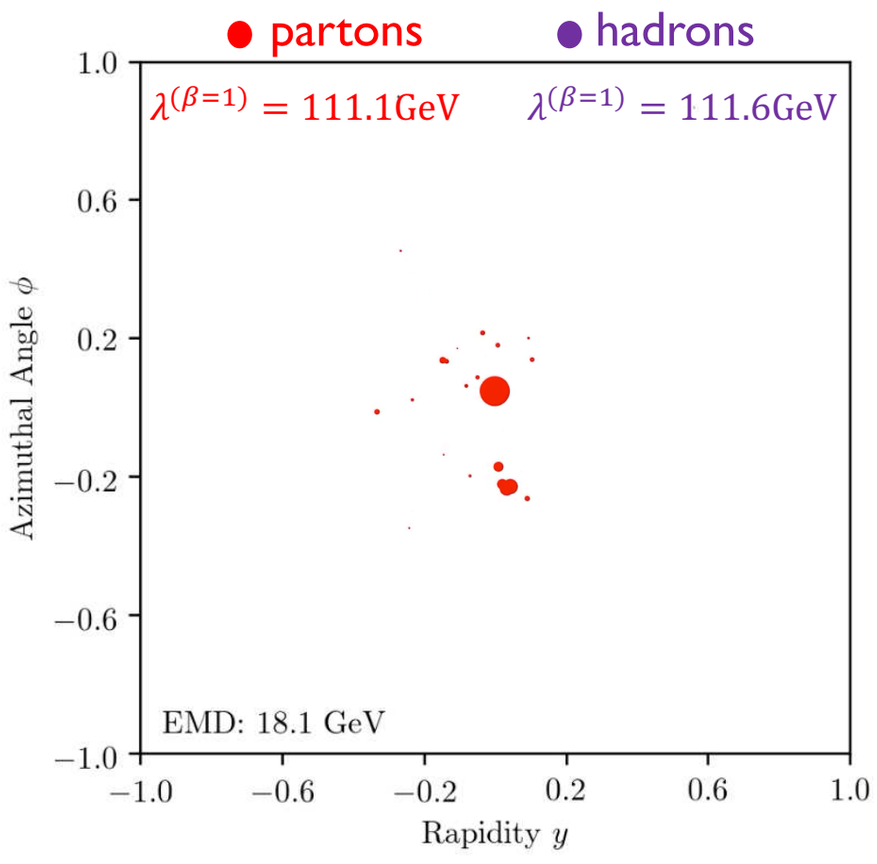
Physics

Events close in EMD are close according to IRC-safe observables

$$\mathrm{EMD}(\mathcal{E},\mathcal{E}') \geq \frac{1}{RL}\left|\sum_i E_i\Phi(\hat{p}_i) - \sum_j E'_j\Phi(\hat{p}'_j)\right| = \frac{1}{RL}\left|\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')\right|$$

via Kantorovich-Rubinstein duality

Additive IRC-safe observable

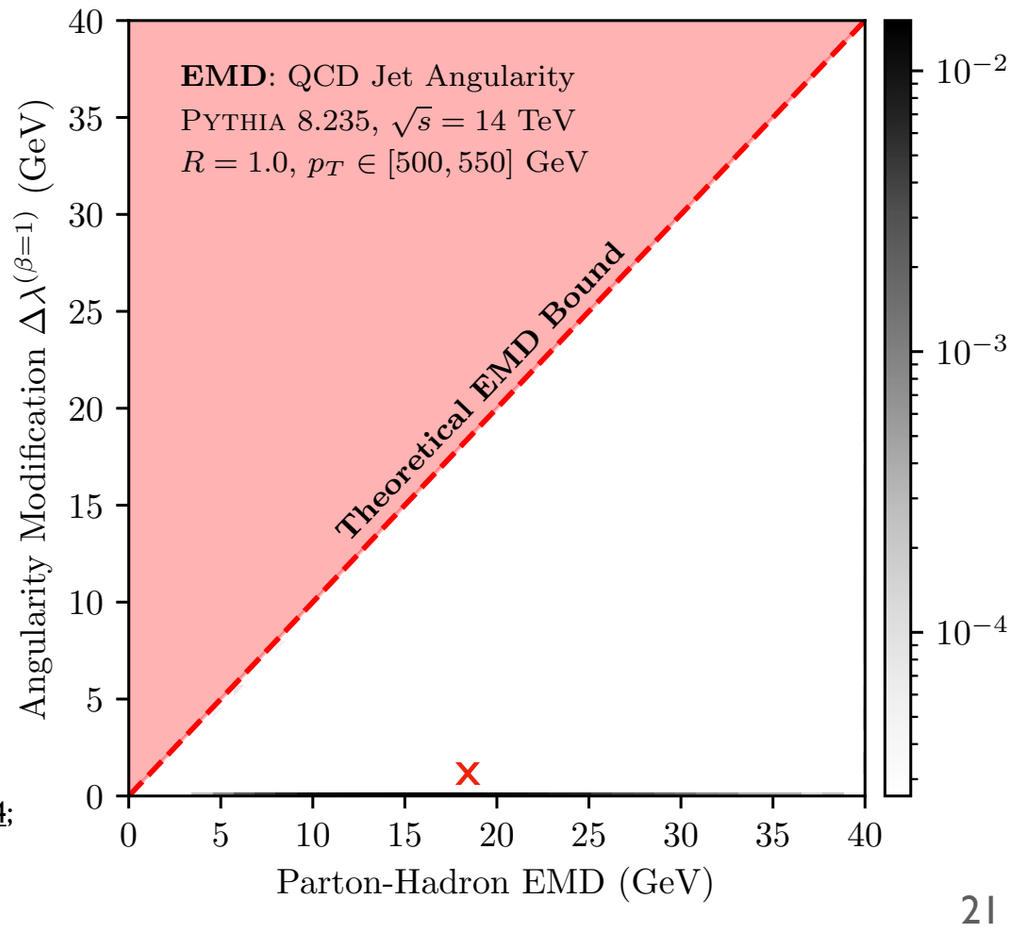


e.g. bounding IRC-safe angularities

$$\lambda^{(\beta)}(\mathcal{E}) = \sum_i E_i\theta_i^\beta$$

Can do the same for pileup, detector effects

[Berger, Kucs, Sterman, hep-ph/0303051; Ellis, Vermilion, Walsh, Hornig, Lee, 1001.0014; Larkoski, Thaler, Waalewijn, 1408.3122]



Quantifying Event Modifications

[PTK, Metodiev, Thaler, 1902.02346]

Mathematics

1-Wasserstein metric bounds the difference in expectation values between distributions

⇨

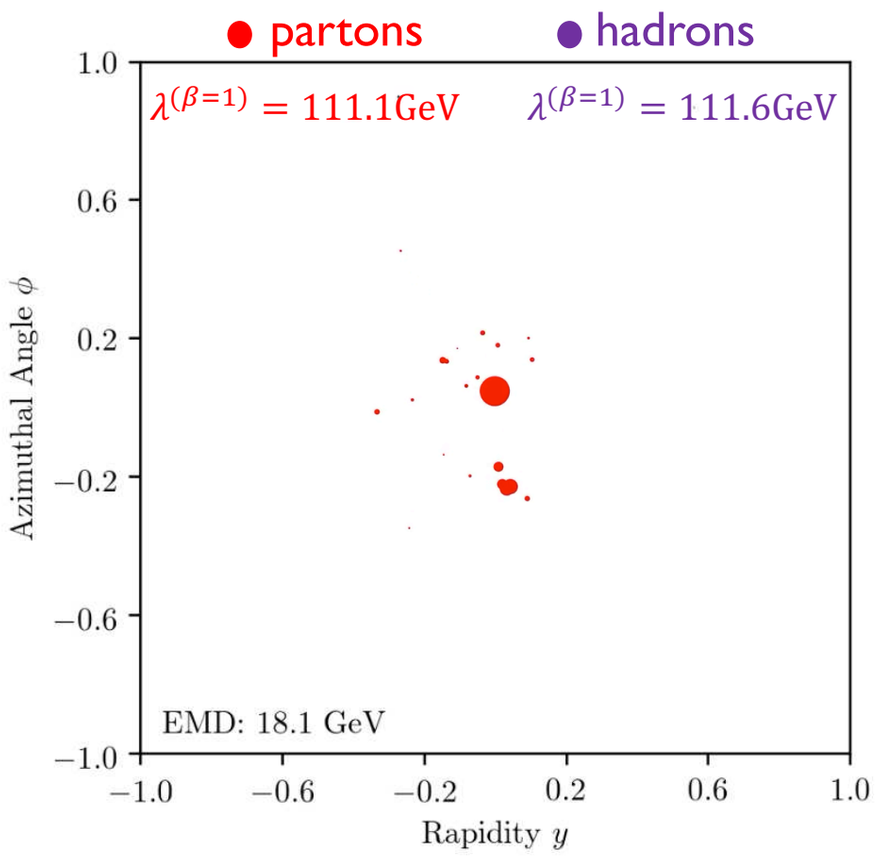
Physics

Events close in EMD are close according to IRC-safe observables

$$\mathrm{EMD}(\mathcal{E},\mathcal{E}') \geq \frac{1}{RL}\left|\sum_i E_i\Phi(\hat{p}_i) - \sum_j E'_j\Phi(\hat{p}'_j)\right| = \frac{1}{RL}\left|\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')\right|$$

via Kantorovich-Rubinstein duality

Additive IRC-safe observable

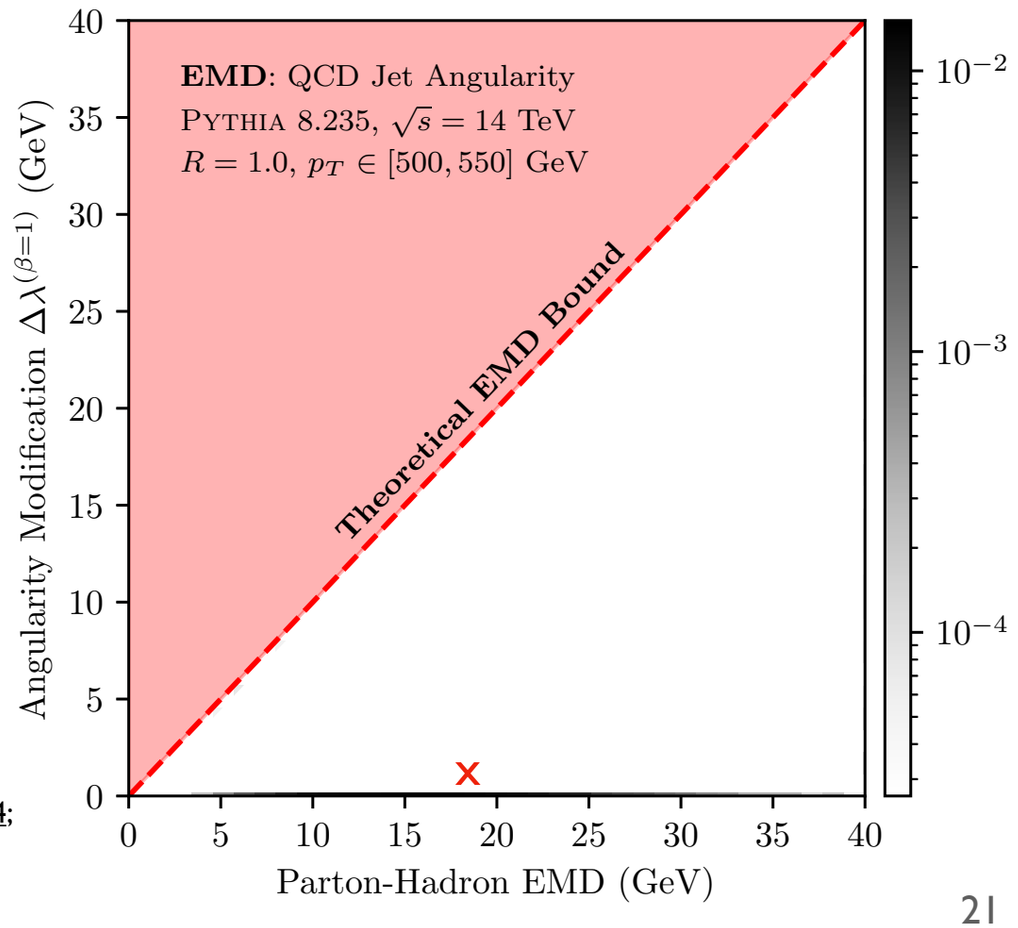


e.g. bounding IRC-safe angularities

$$\lambda^{(\beta)}(\mathcal{E}) = \sum_i E_i\theta_i^\beta$$

Can do the same for pileup, detector effects

[Berger, Kucs, Sterman, hep-ph/0303051; Ellis, Vermilion, Walsh, Hornig, Lee, 1001.0014; Larkoski, Thaler, Waalewijn, 1408.3122]



Quantifying Event Modifications

[PTK, Metodiev, Thaler, 1902.02346]

Mathematics

1-Wasserstein metric bounds the difference in expectation values between distributions

⇨

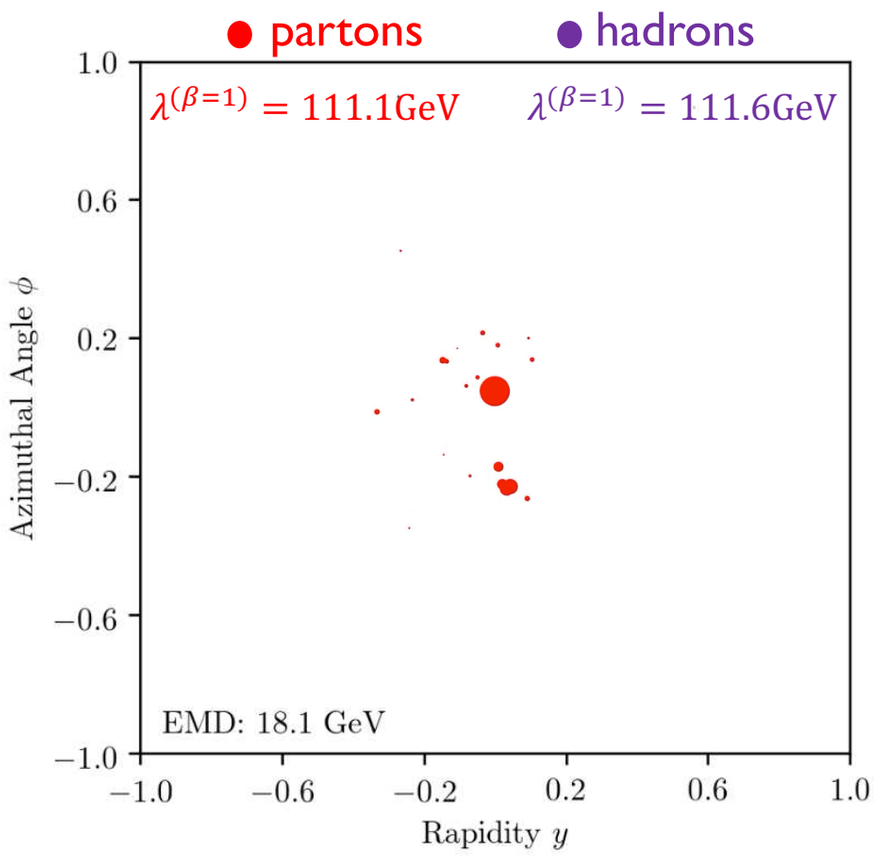
Physics

Events close in EMD are close according to IRC-safe observables

$$\mathrm{EMD}(\mathcal{E},\mathcal{E}') \geq \frac{1}{RL}\left|\sum_i E_i\Phi(\hat{p}_i) - \sum_j E'_j\Phi(\hat{p}'_j)\right| = \frac{1}{RL}\left|\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')\right|$$

via Kantorovich-Rubinstein duality

Additive IRC-safe observable

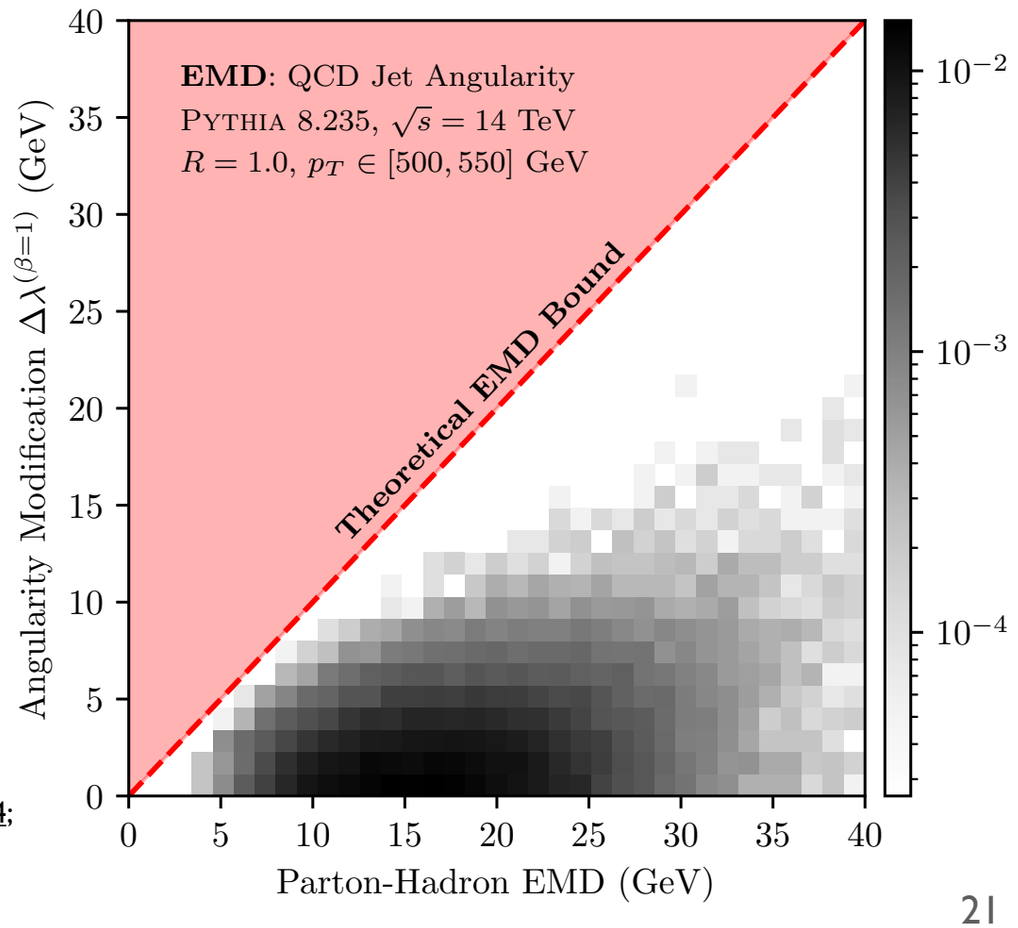


e.g. bounding IRC-safe angularities

$$\lambda^{(\beta)}(\mathcal{E}) = \sum_i E_i\theta_i^\beta$$

Can do the same for pileup, detector effects

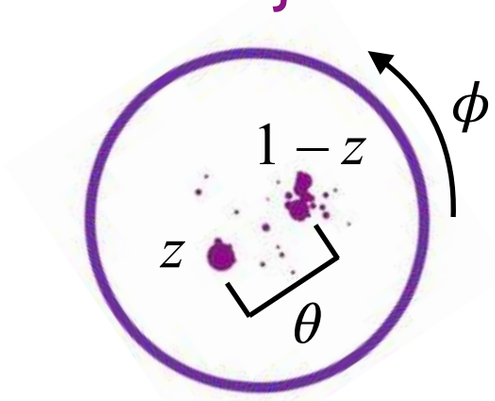
[Berger, Kucs, Sterman, hep-ph/0303051; Ellis, Vermilion, Walsh, Hornig, Lee, 1001.0014; Larkoski, Thaler, Waalewijn, 1408.3122]



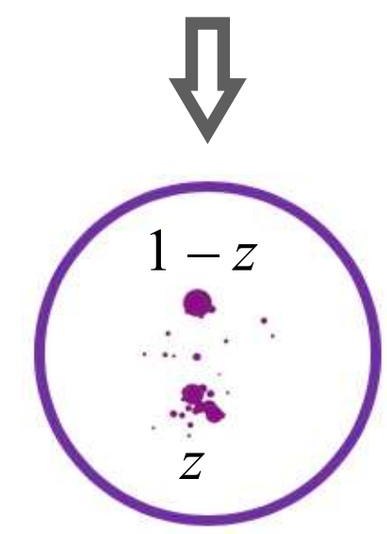
Visualizing the Metric Space of *W* Jets

Embed high-dimension manifold in low-dimensional space?

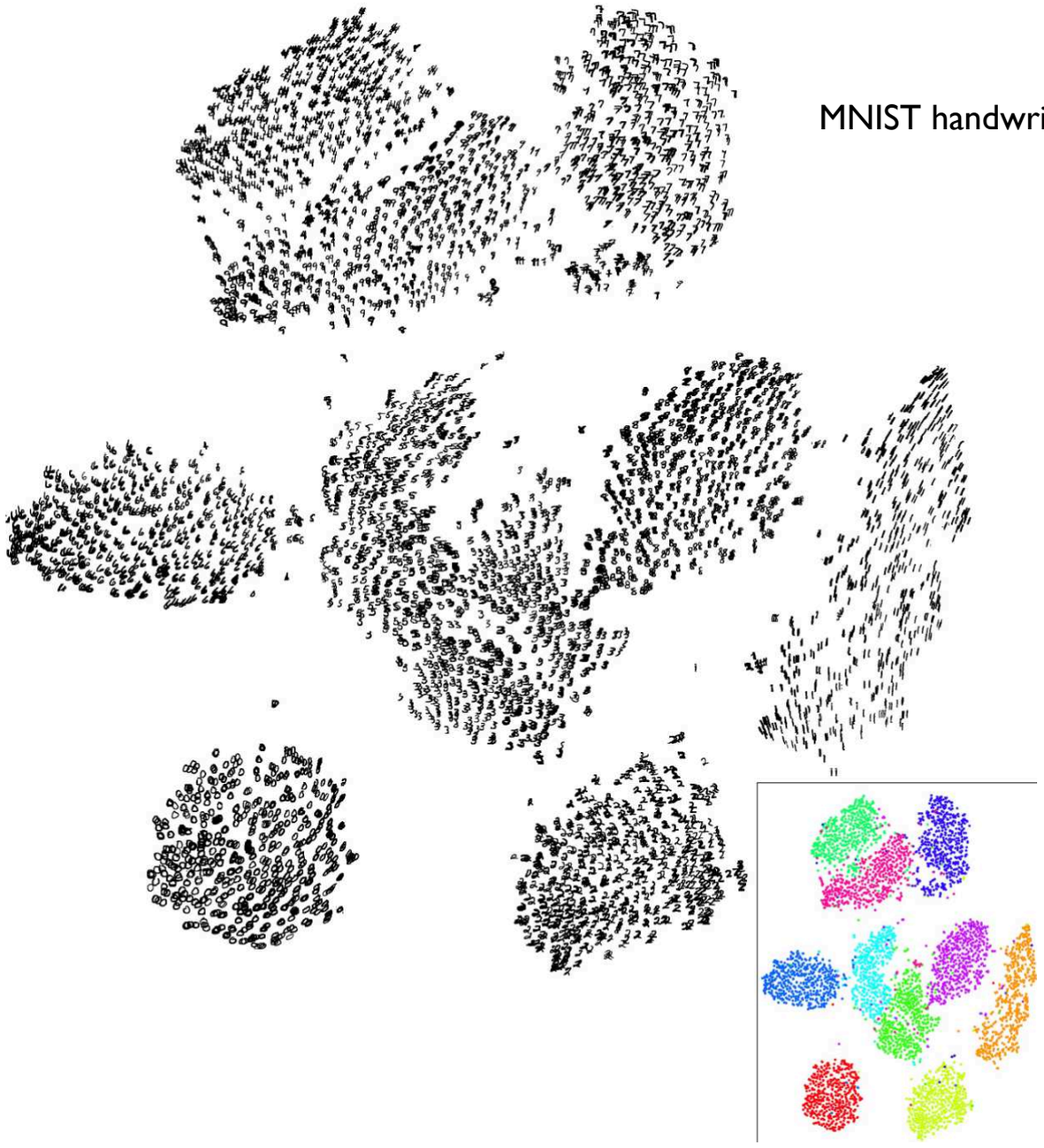
W Jet



Constraints: *W* Mass and ϕ = 0 preprocessing



t-Distributed Stochastic Neighbor Embedding

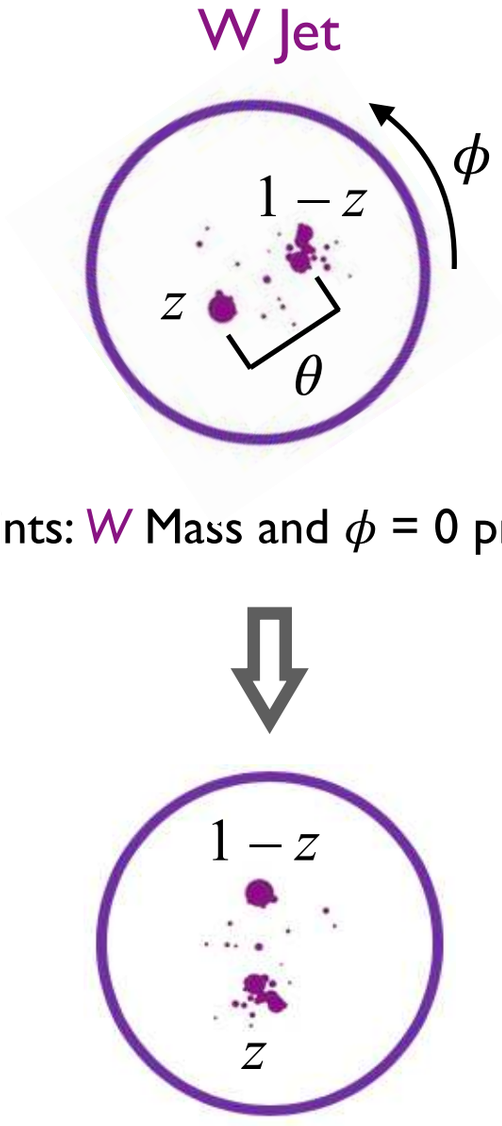


MNIST handwritten digits

[L. van der Maaten, G. Hinton, JMLR 2008]

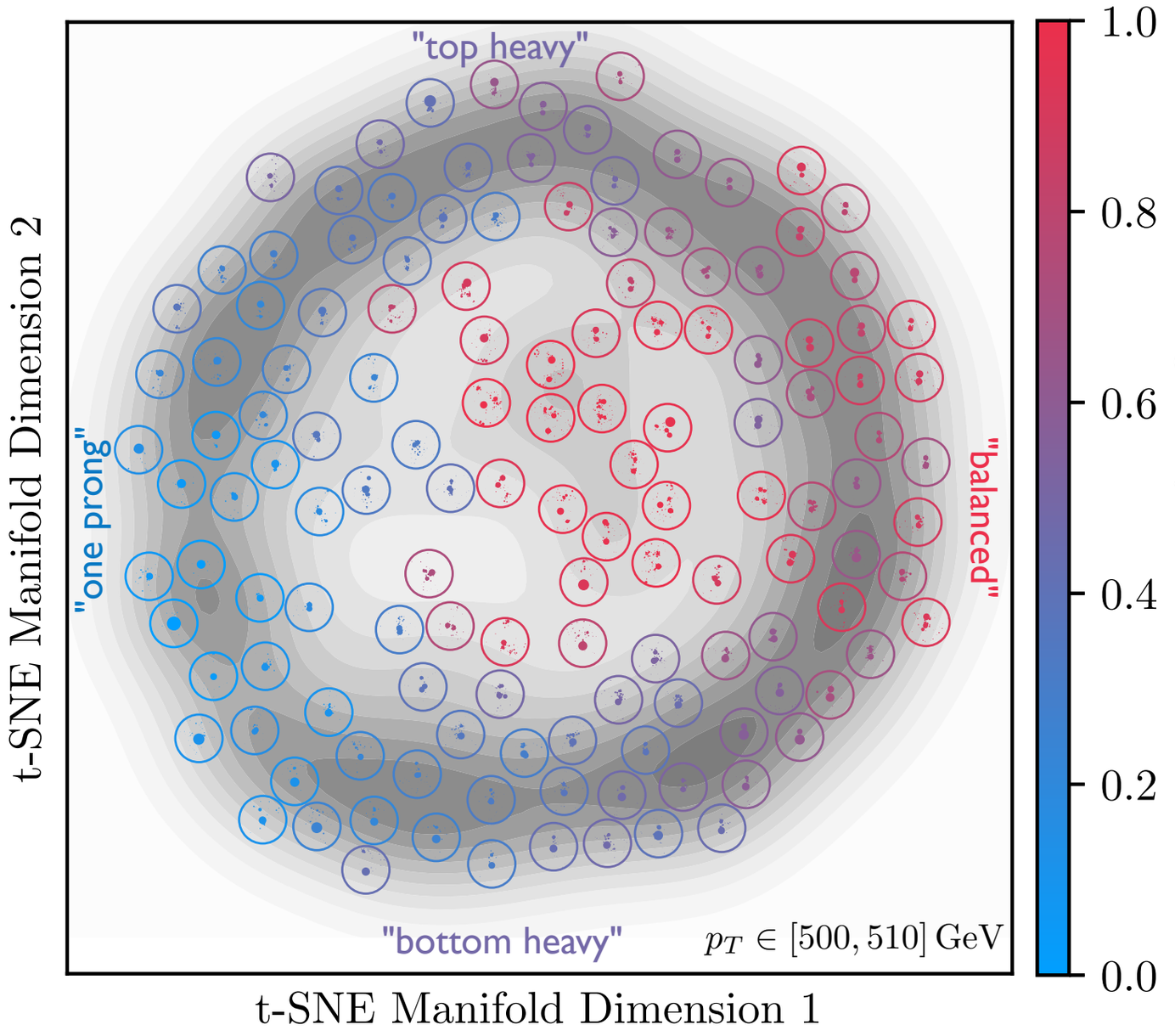
Visualizing the Metric Space of *W* Jets

Embed high-dimension manifold in low-dimensional space?



Constraints: *W* Mass and ϕ = 0 preprocessing

t-Distributed Stochastic Neighbor Embedding

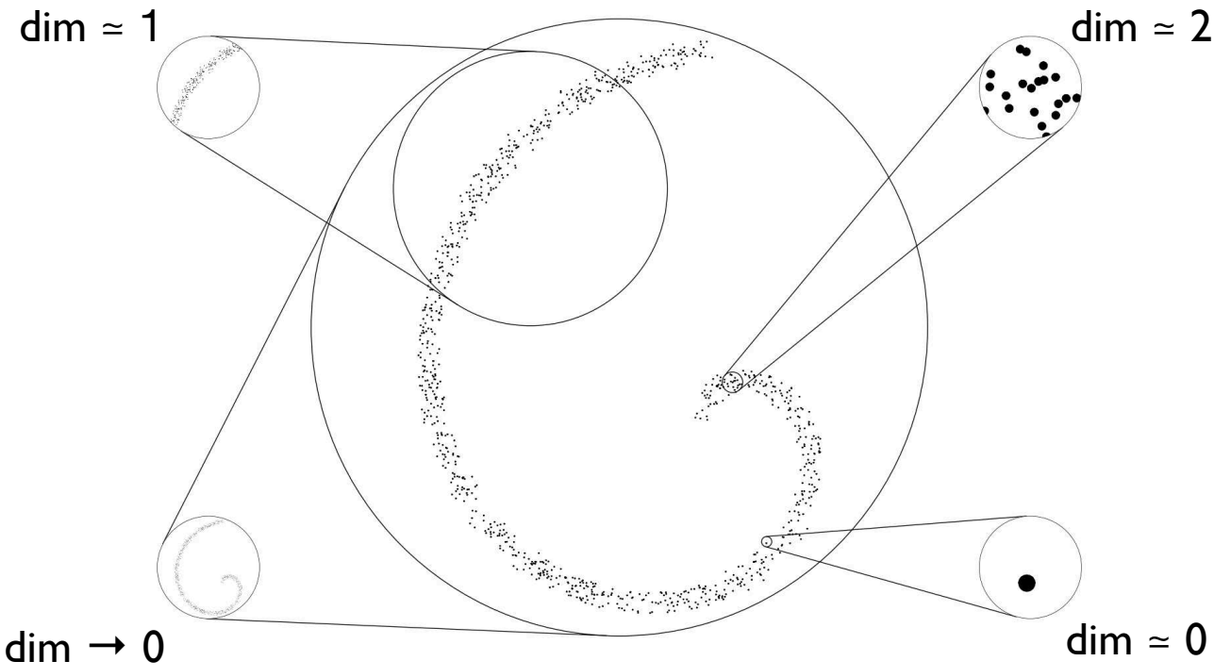


[L. van der Maaten, G. Hinton, JMLR 2008]

Manifold Dimensions of Event Space

Correlation dimension: how does the # of elements within a ball of size Q change?

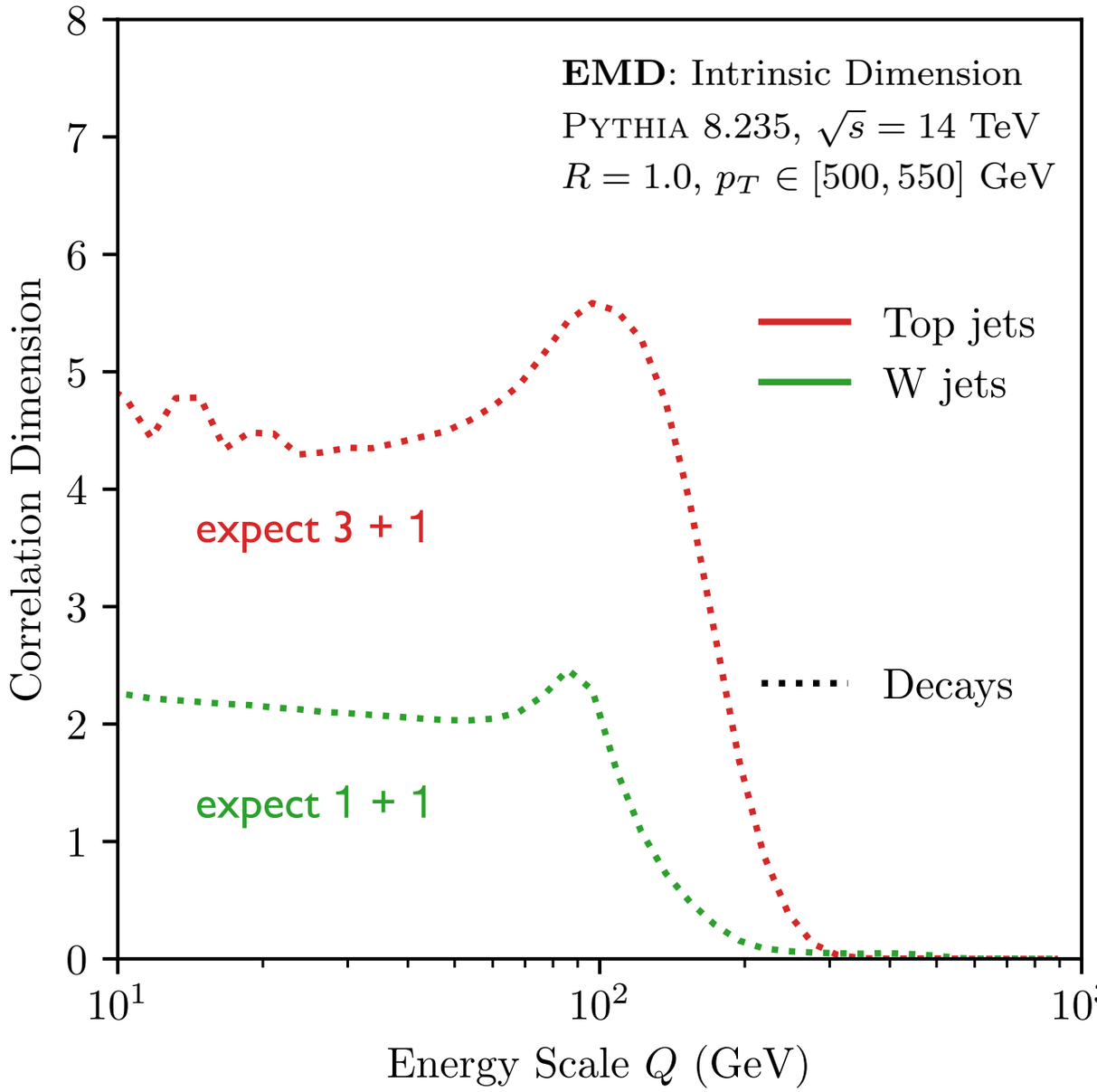
$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln\sum_i\sum_j\Theta(\mathrm{EMD}(\mathcal{E}_i,\mathcal{E'}_j) < Q)$$



$$N_{\mathrm{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q) = Q\frac{d}{dQ}\ln N_{\mathrm{neigh.}}(Q)$$

Correlation dimension lessons:

Decays are "constant" dim. at low Q

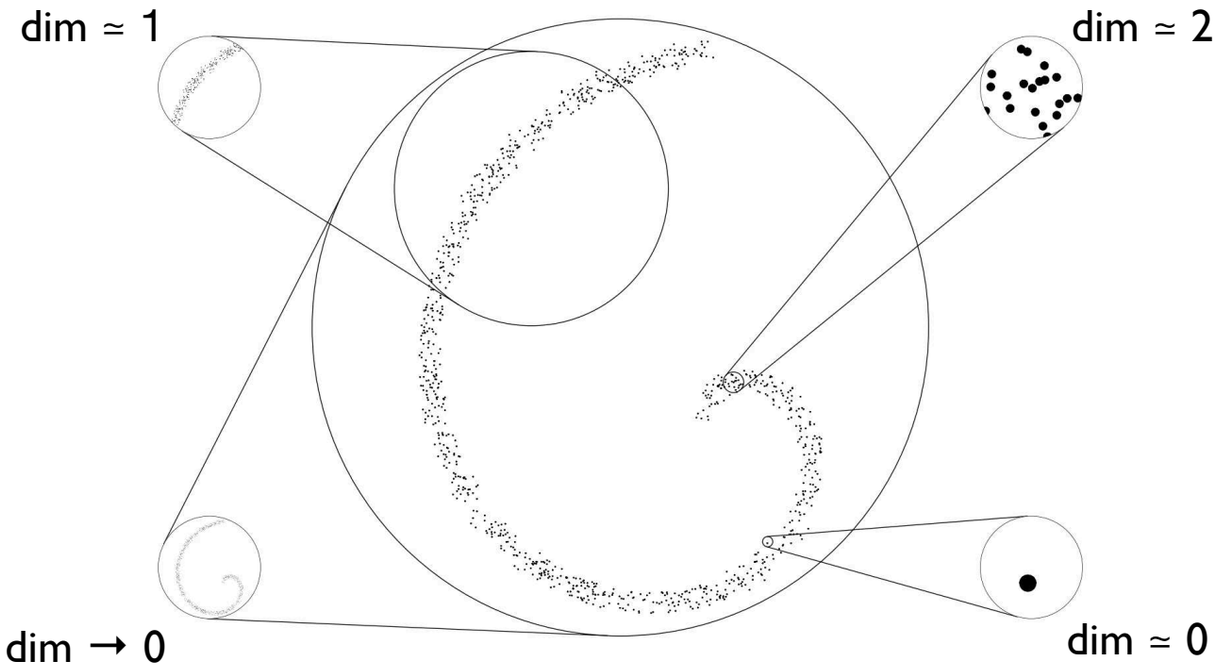


[Grassberger, Procaccia, PRL 1983; PTK, Metodiev, Thaler, 1902.02346]

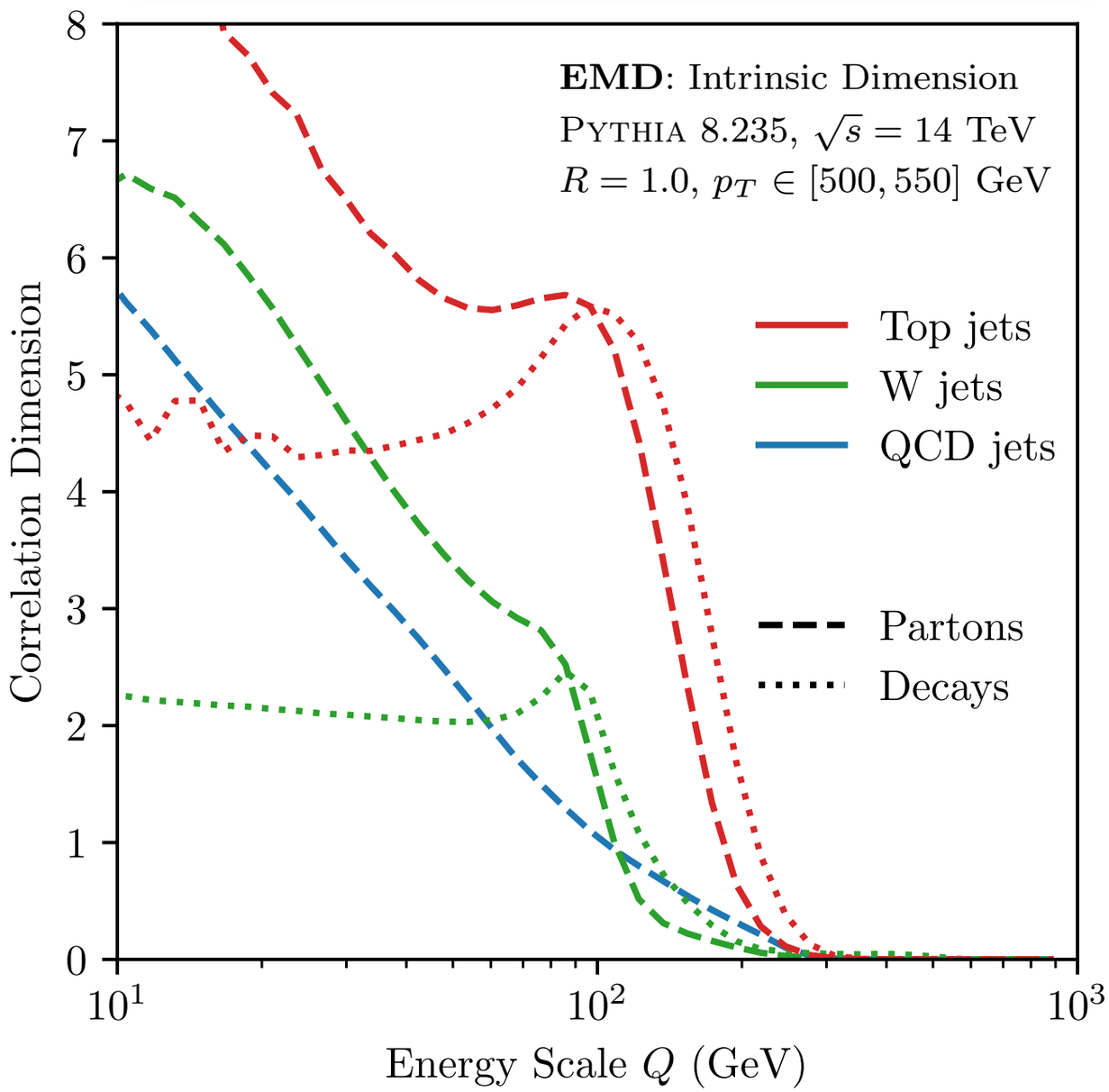
Manifold Dimensions of Event Space

Correlation dimension: how does the # of elements within a ball of size Q change?

$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln\sum_i\sum_j\Theta(\mathrm{EMD}(\mathcal{E}_i,\mathcal{E'}_j) < Q)$$



$$N_{\mathrm{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q) = Q\frac{d}{dQ}\ln N_{\mathrm{neigh.}}(Q)$$



Correlation dimension lessons:

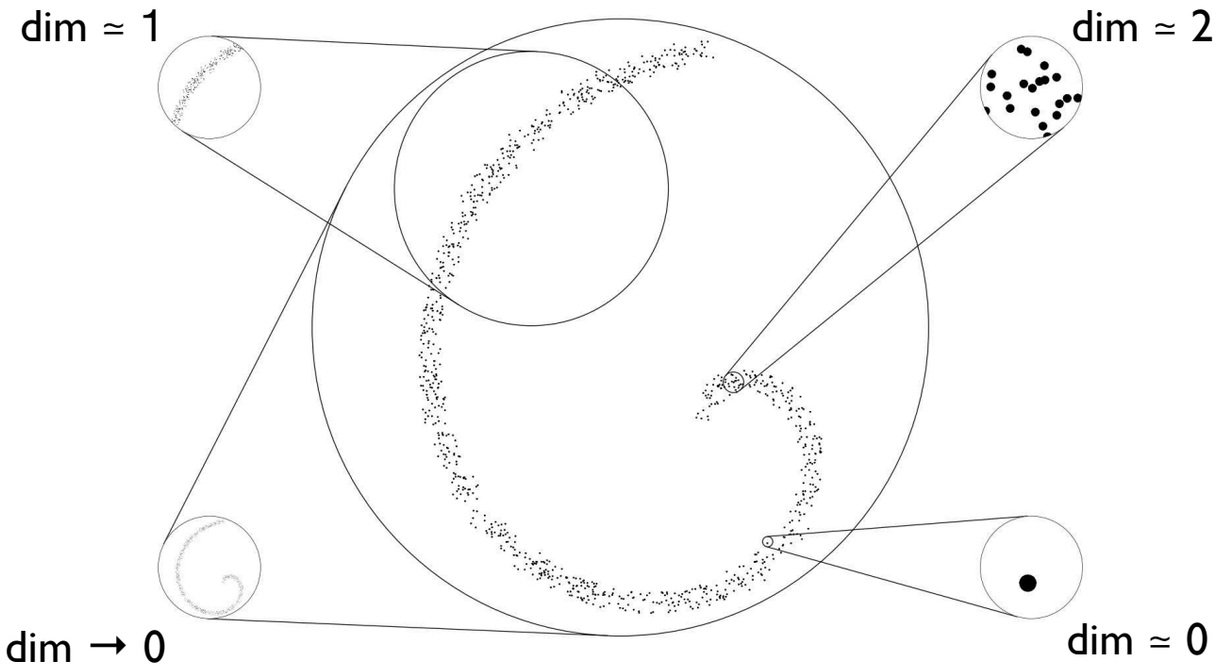
- Decays are "constant" dim. at low Q
- Complexity hierarchy: QCD < W < Top
- Fragmentation increases dim. at smaller scales

[Grassberger, Procaccia, PRL 1983; PTK, Metodiev, Thaler, 1902.02346]

Manifold Dimensions of Event Space

Correlation dimension: how does the # of elements within a ball of size Q change?

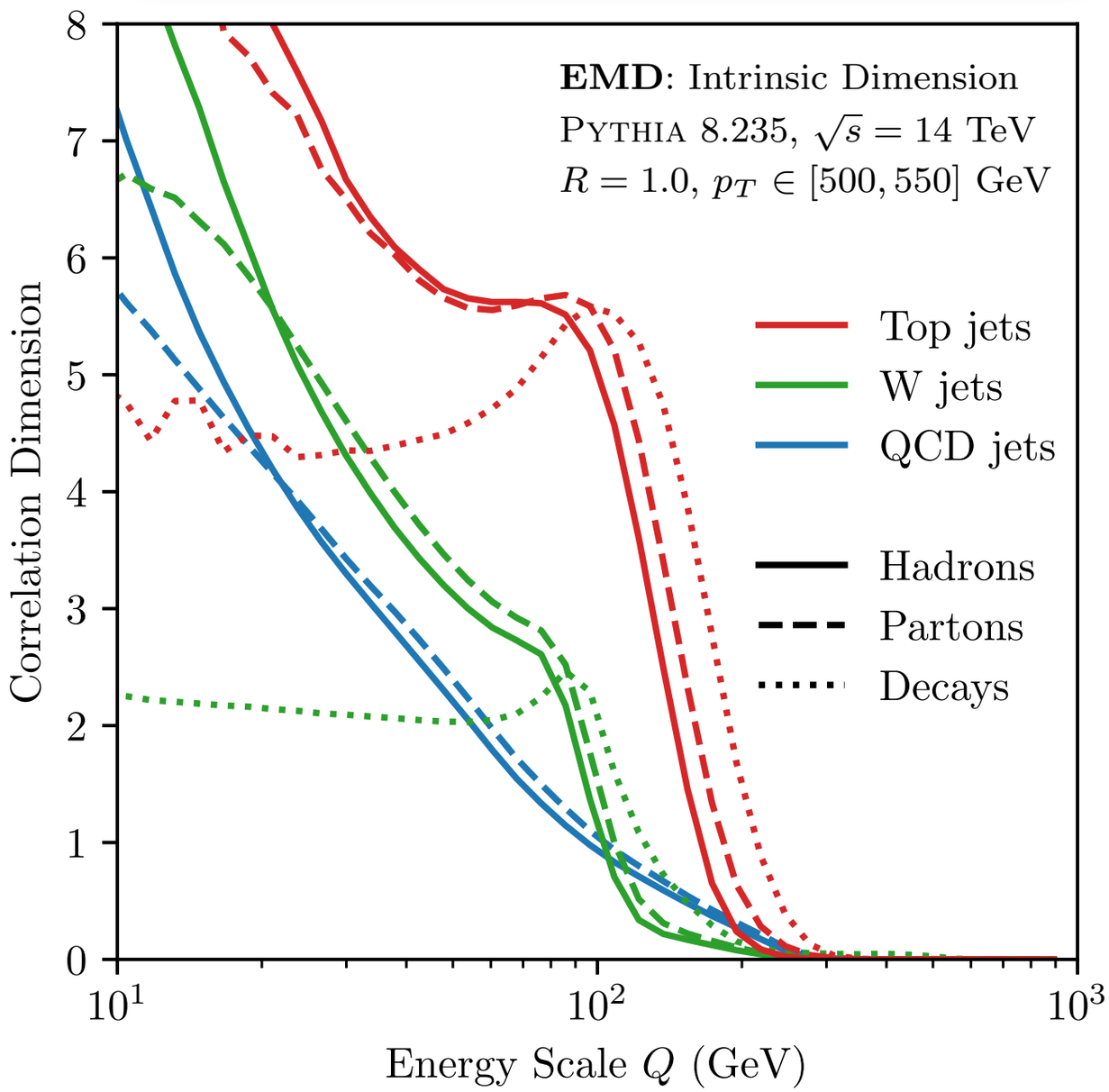
$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln\sum_i\sum_j\Theta(\mathrm{EMD}(\mathcal{E}_i,\mathcal{E'}_j) < Q)$$



$$N_{\mathrm{neigh.}}(Q) \propto Q^{\dim} \implies \dim(Q) = Q\frac{d}{dQ}\ln N_{\mathrm{neigh.}}(Q)$$

Correlation dimension lessons:

- Decays are "constant" dim. at low Q
- Complexity hierarchy: QCD < W < Top
- Fragmentation increases dim. at smaller scales
- Hadronization important around 20-30 GeV



[Grassberger, Procaccia, PRL 1983; PTK, Metodiev, Thaler, 1902.02346]

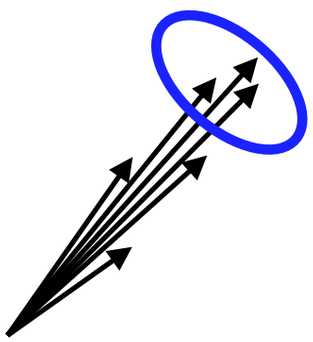
Quark and Gluon Correlation Dimensions

$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln \sum_i \sum_j \Theta(\mathrm{EMD}(\mathcal{E}_i, \mathcal{E'}_j) < Q)$$

Leading log calculation:

$$\dim_i(Q) \simeq -\frac{8\alpha_s}{\pi} C_i \ln \frac{Q}{p_T/2}$$

color factor

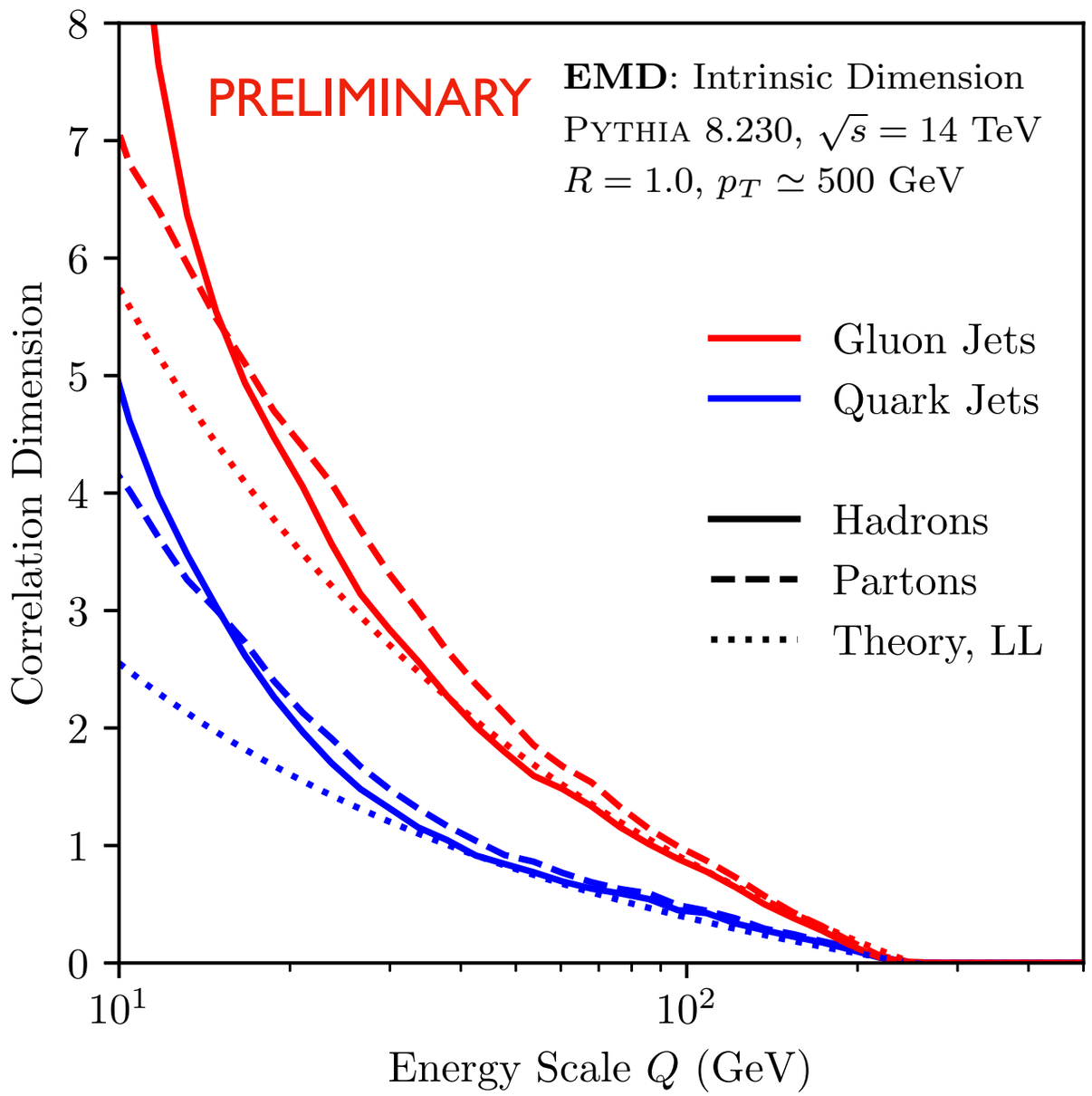


Quark

Gluon

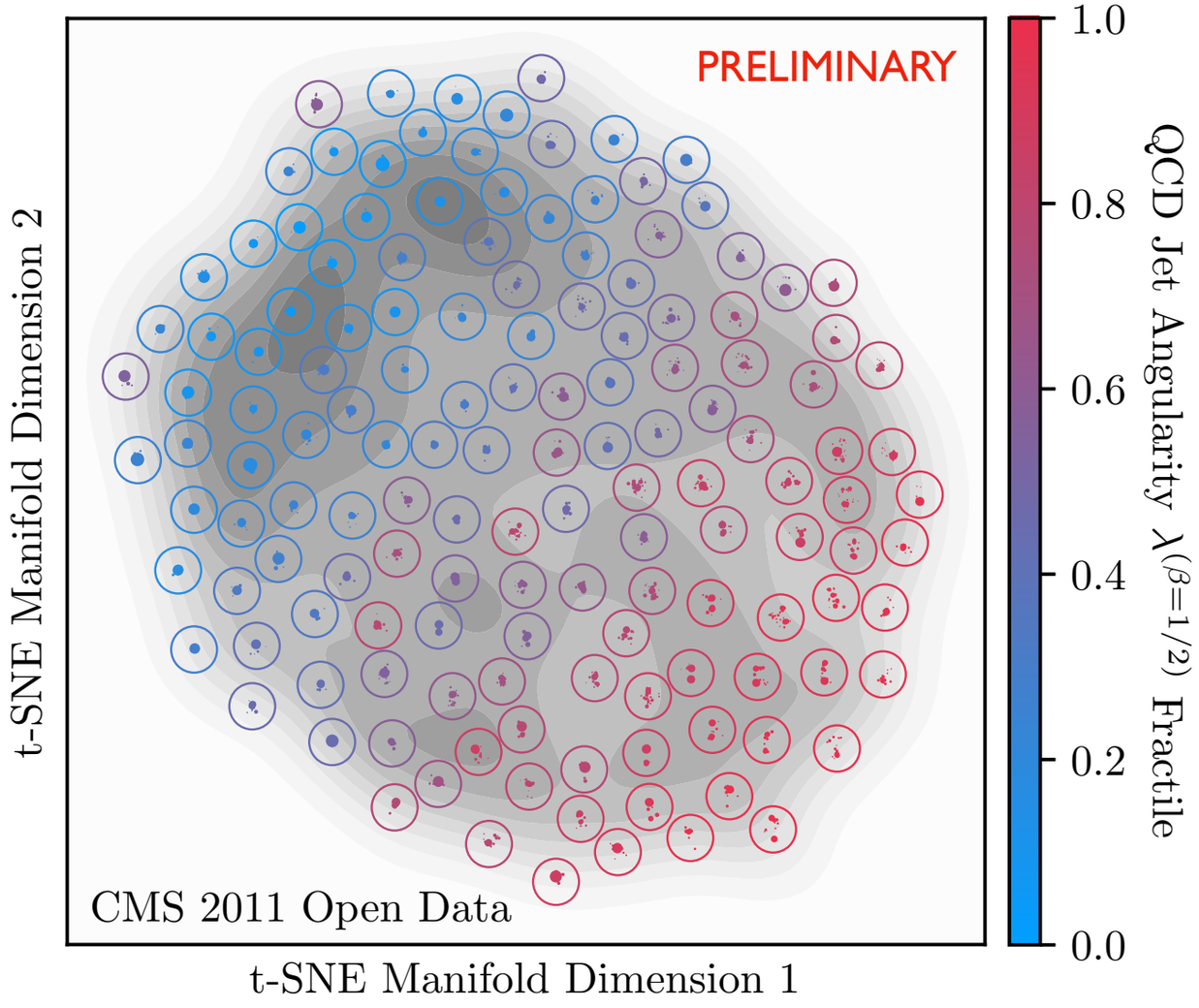
$C_F = 4/3$

$C_A = 3$

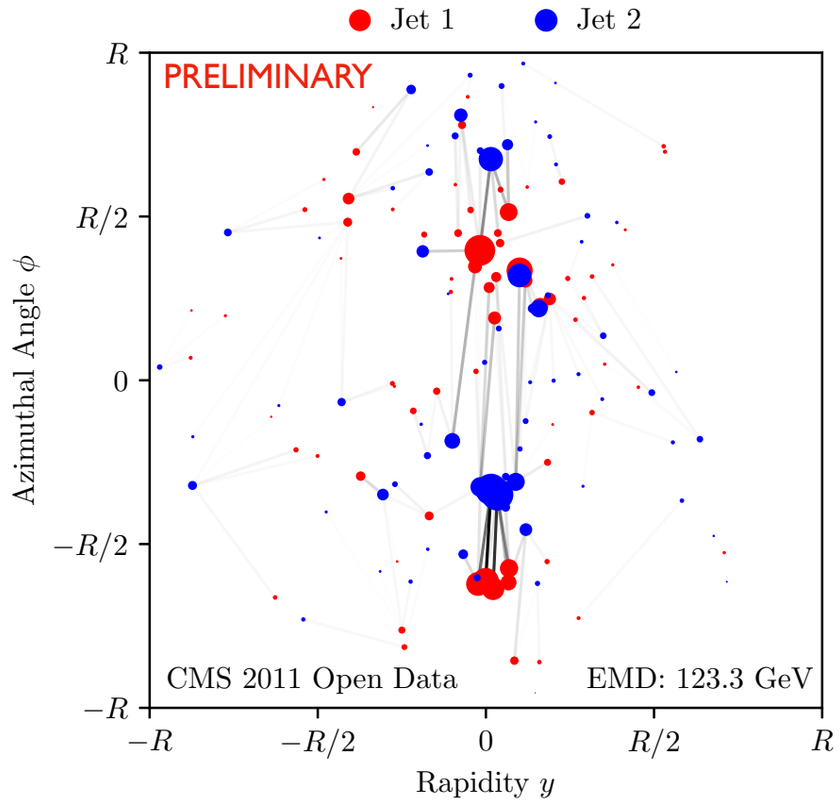


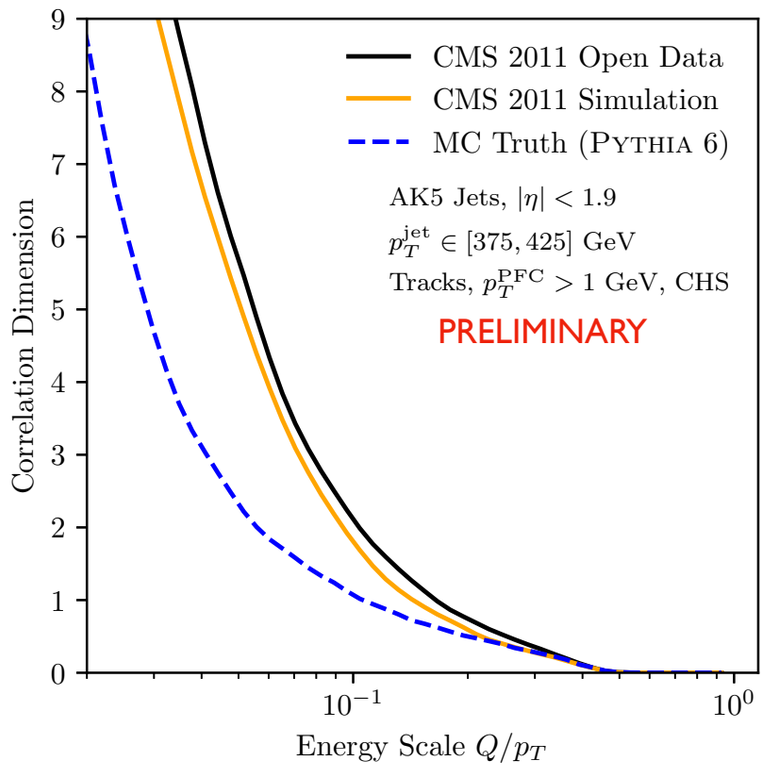
Visualizing Jets with CMS Open Data

Application of EMD techniques to the jet primary dataset in CMS 2011 Open Data



[Mastandrea, Naik, PTK, Metodiev, Thaler, *in progress*]

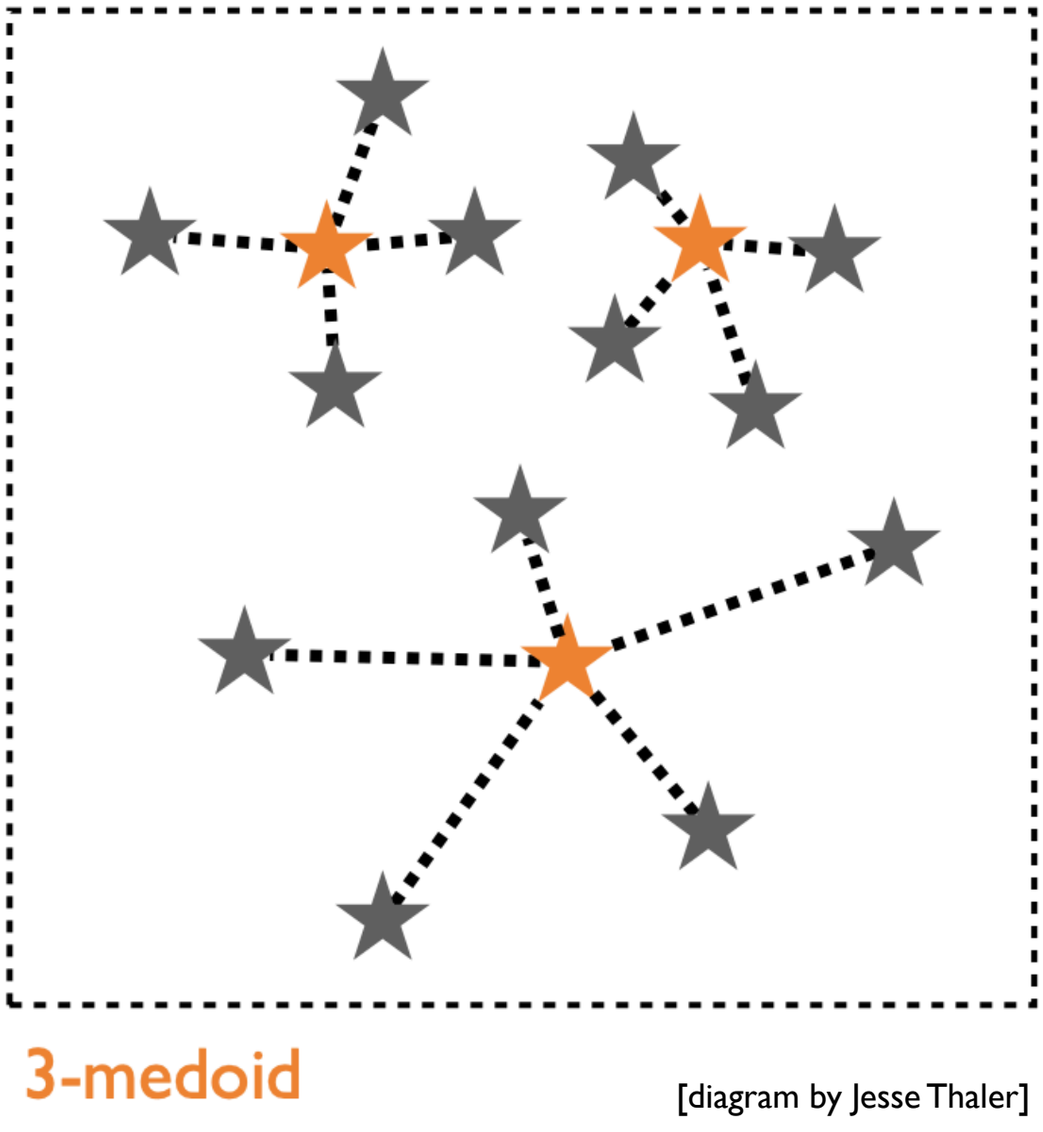




Identifying Representative Jets

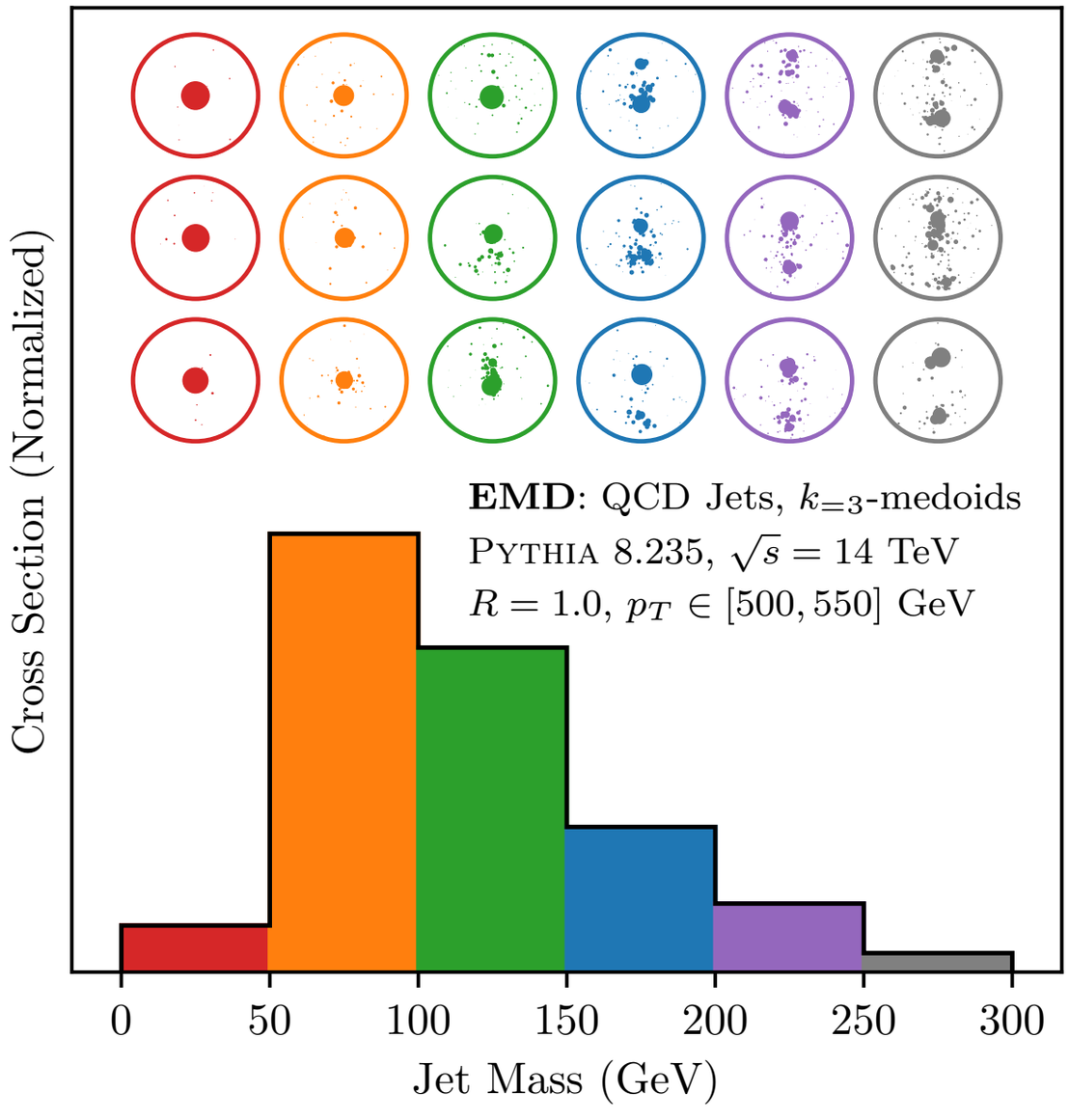
medoid: element selected to best represent a set of elements

k-medoids: k clusters to minimize total distance of points to medoids



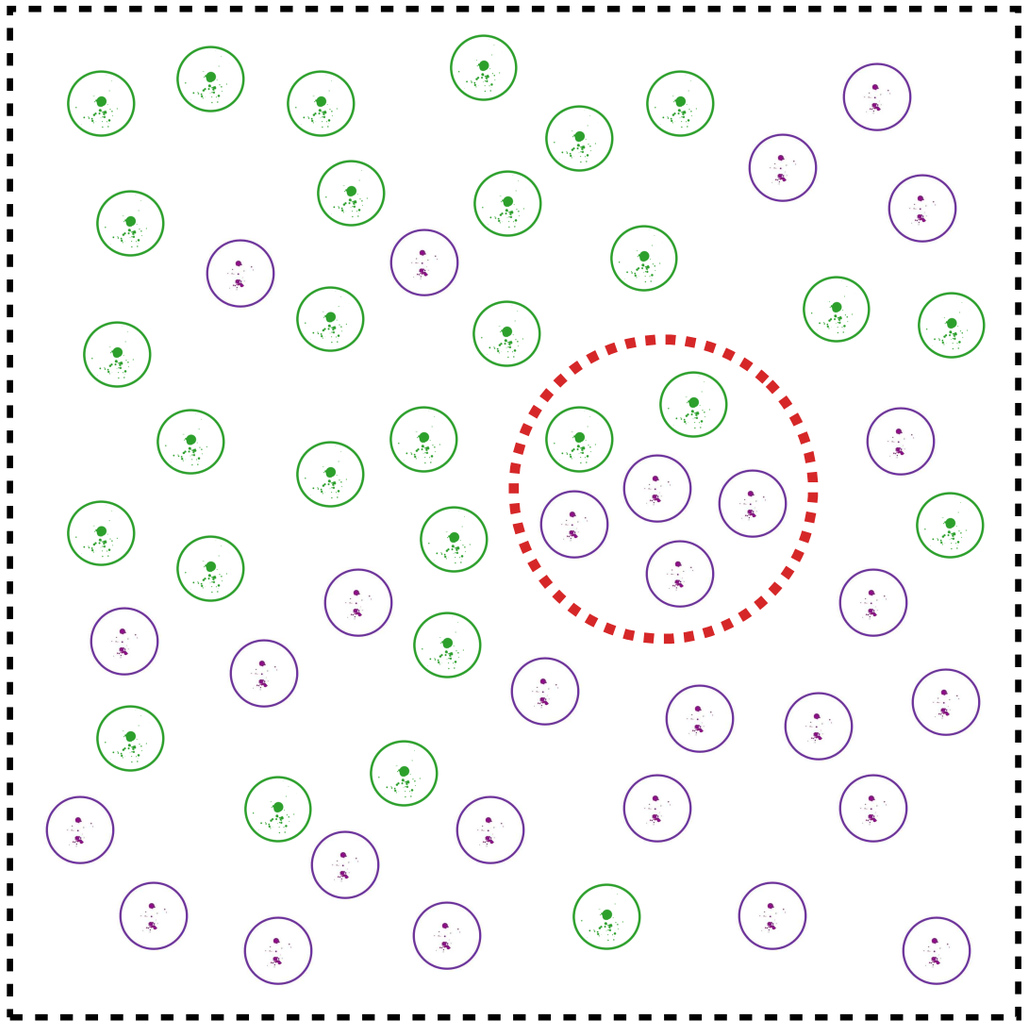
3-medoid

[diagram by Jesse Thaler]

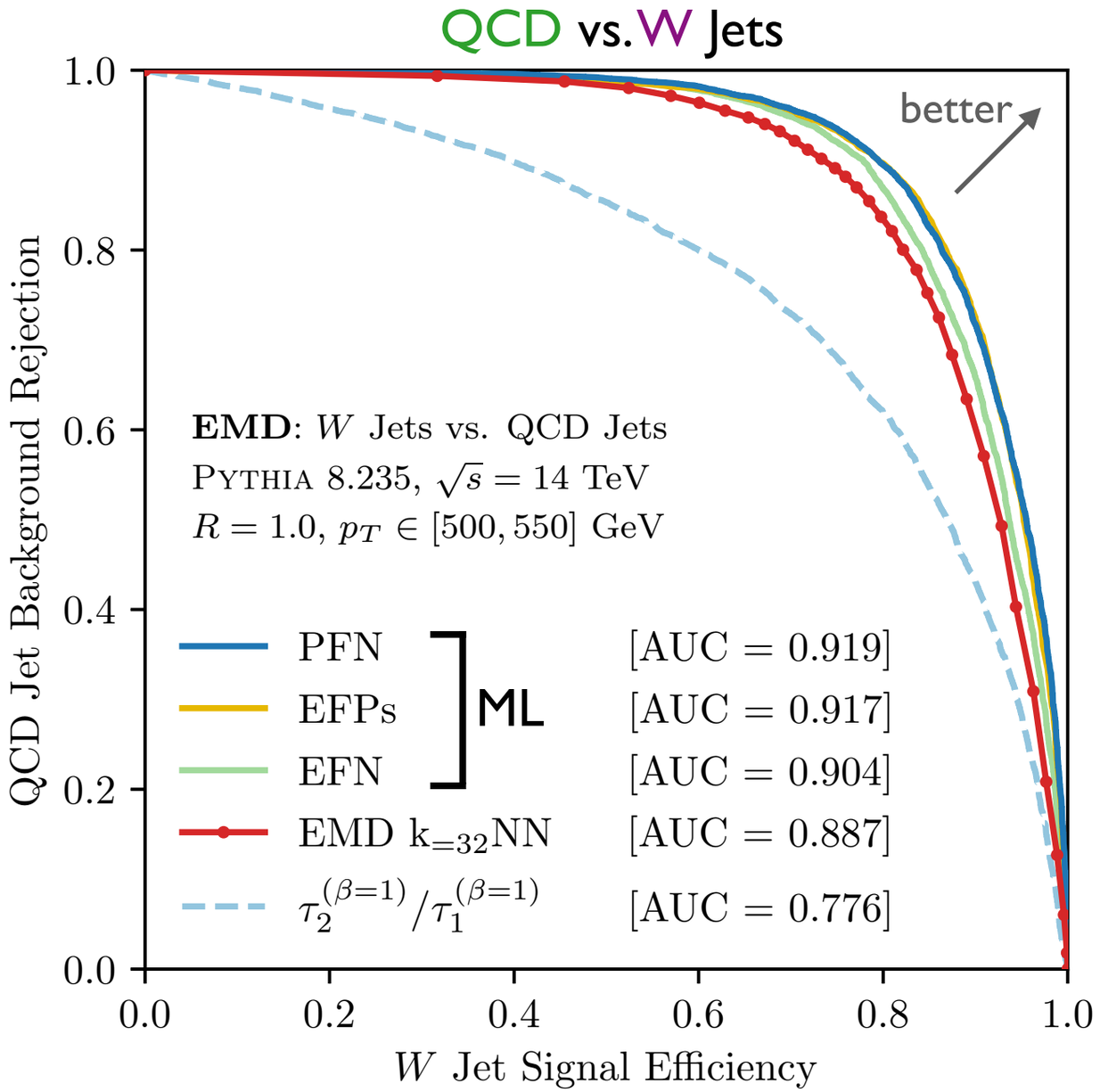


[PTK, Metodiev, Thaler, 1902.02346]

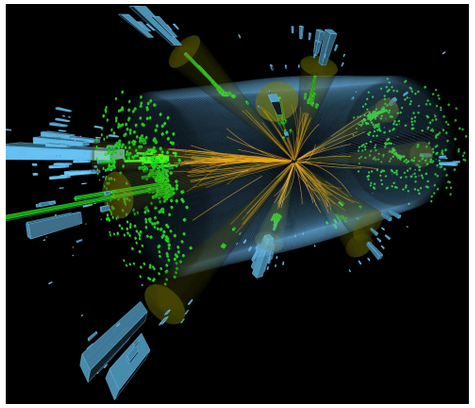
Jet Classification via Nearest-Neighbor Density Estimation

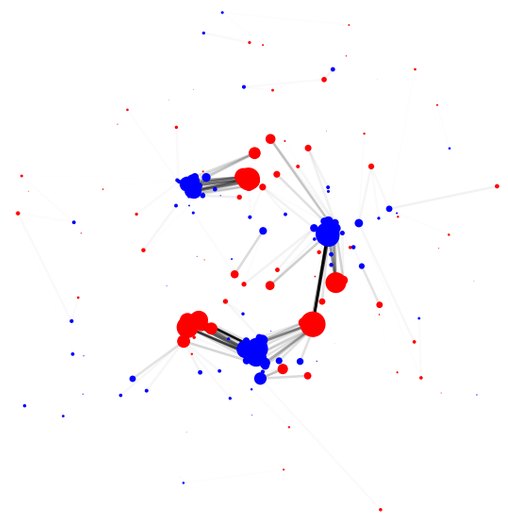


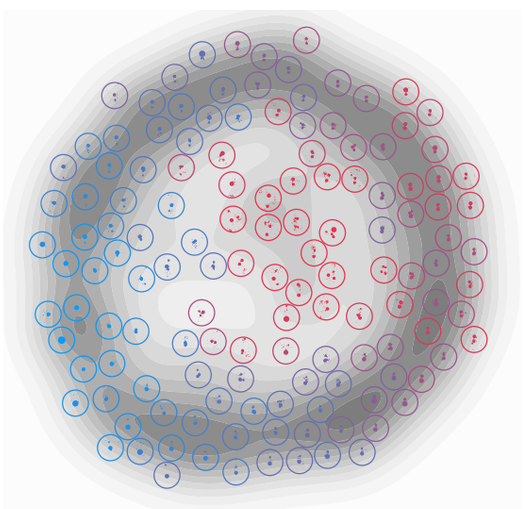
Approaches performance of modern machine learning



[PTK, Metodiev, Thaler, 1902.02346;
comparison to Thaler, Van Tilburg, 1011.2268, 1108.2701; PTK, Metodiev, Thaler, 1712.07124, 1810.05165;]







Collider Event Foundations

IRC-safe energy flow is theoretically and experimentally robust

The Energy Mover's Distance

Quantifies the difference in energy flow between events

Particle Physics Applications

Quantifying modifications, visualizing and exploring event space

Further Directions

Experimental

Quantify (or even mitigate?) pileup/detector effects

Non-parametric density estimates (unfolding?)

Automated data compression (triggering?)

Theoretical

Define new observables with EMD?

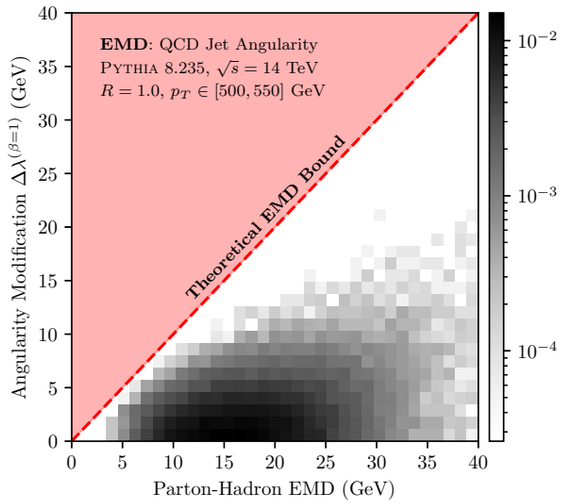
Precision QCD calculations of event space geometry?

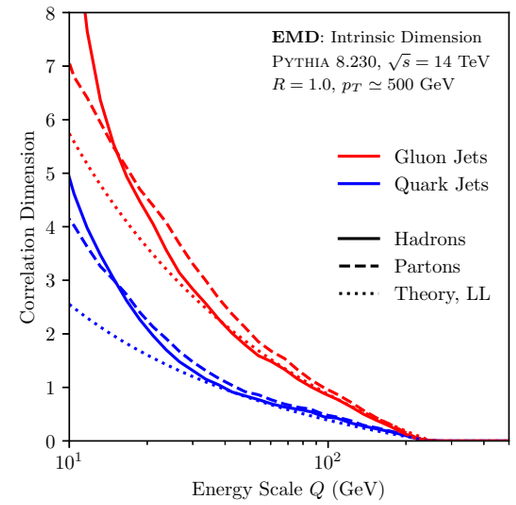
Event Mover's Distance between ensembles?

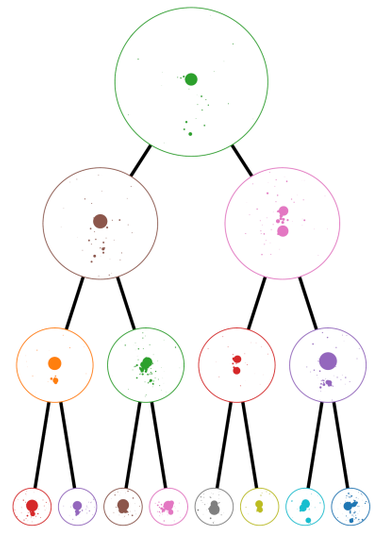
Algorithmic

Loss function for modern ML in particle physics?

Metric trees to turn $O(N^2)$ into $O(N \log N)$?







[BOOST 2019, July 22-26, MIT]

Phenomenology | Reconstruction | Searches | Algorithms | Measurements | Calculations
Modeling | Machine Learning | Pileup Mitigation | Heavy-Ion Collisions | Future Colliders

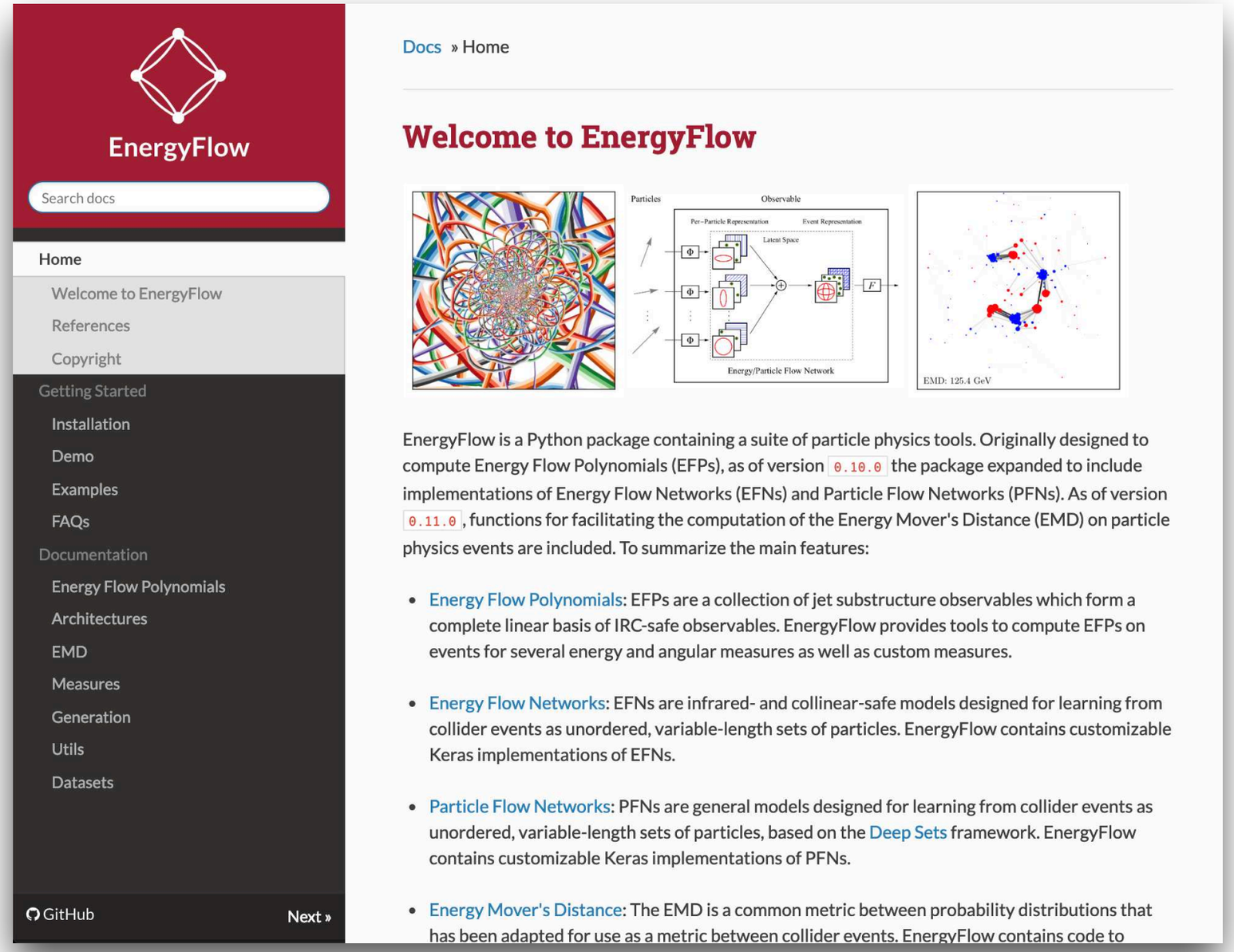
EnergyFlow Python Package

https://energyflow.network

Parallelized EMD calculations via the Python Optimal Transport library

Keras implementations of EFNs, PFNs, DNNs, CNNs, efficient EFP computation

Several detailed examples and demos for common use cases and visualization procedures



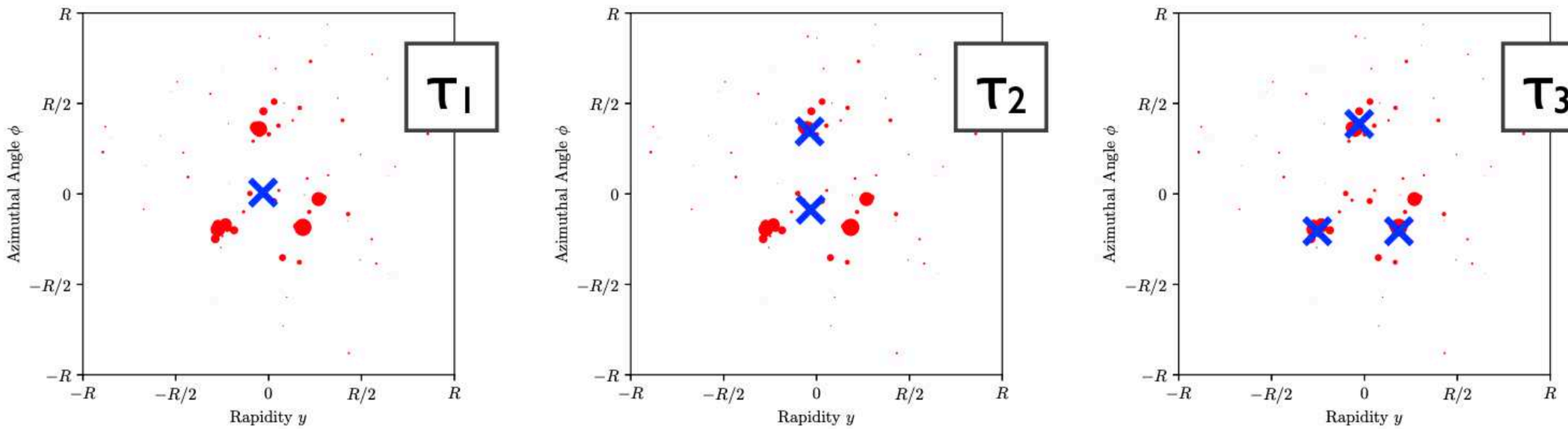
Backup Slides

Connection to N-subjettiness

$$\tau_N^{(\beta)}(\mathcal{E}) = \min_{N \text{ axes}} \sum_i E_i \min \left\{ \theta_{1,i}^{\beta}, \theta_{2,i}^{\beta}, \ldots, \theta_{N,i}^{\beta} \right\}$$

IRC safe

kind of arbitrary



$$\tau_N(\mathcal{E}) = \min_{|\mathcal{E}'|=N} \mathrm{EMD}(\mathcal{E}, \mathcal{E}')$$

for β = 1

Related to p-Wasserstein metric for p = β > 1

very satisfying

[slide from talk by J. Thaler]

[JDT, Van Tilburg, 1011.2268, 1108.2701; based on Brandt, Dahmen, ZPC 1979; Stewart, Tackmann, Waalewijn, 1004.2489]

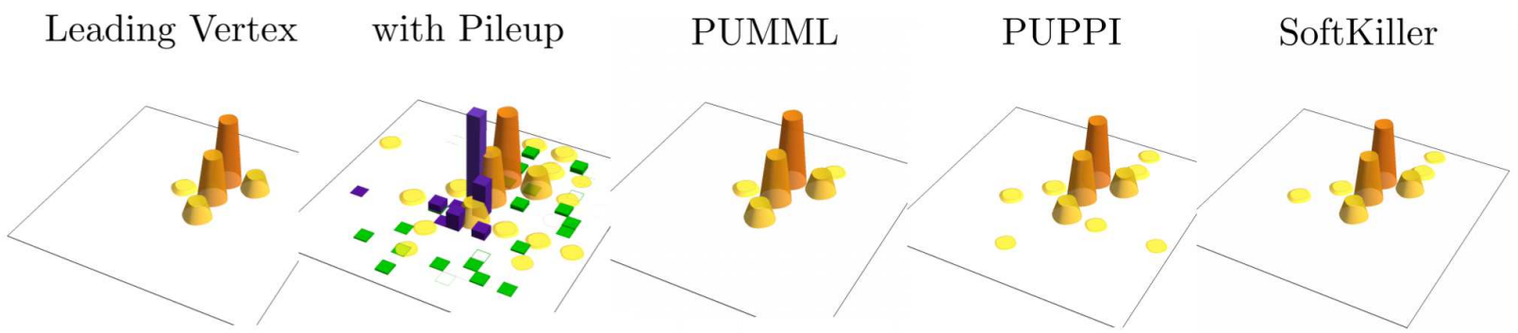
Pileup Removal with Machine Learning (PUMML) and EMD

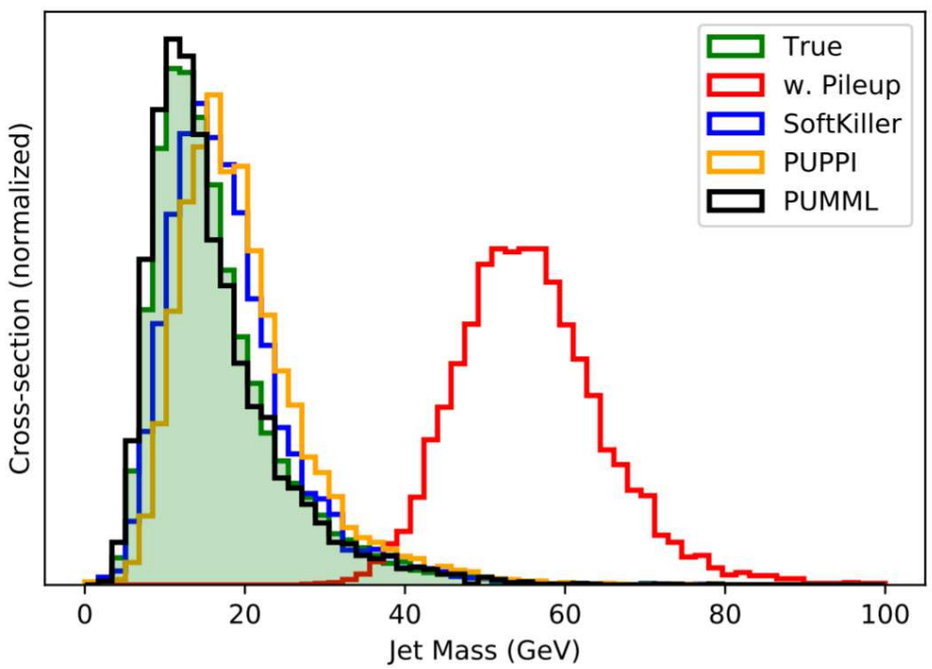
PUMML with jet images

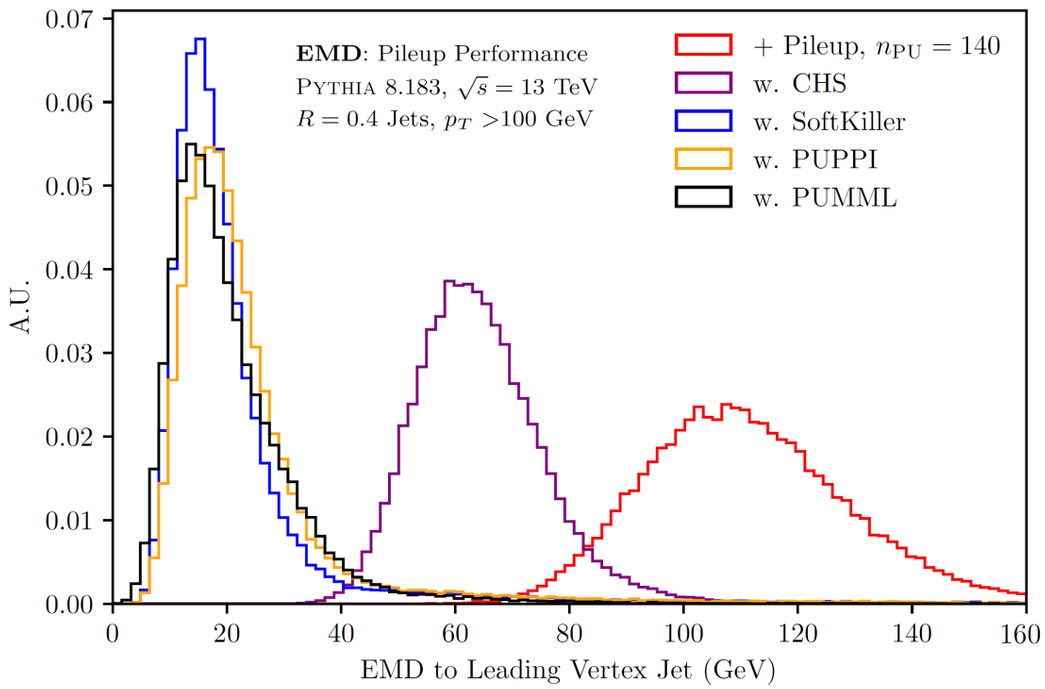
- *pixel-based loss function*
- *compared specific IRC-safe observables*

PUMML with EMD

- *no pixelation*
- *related to all IRC-safe observables*



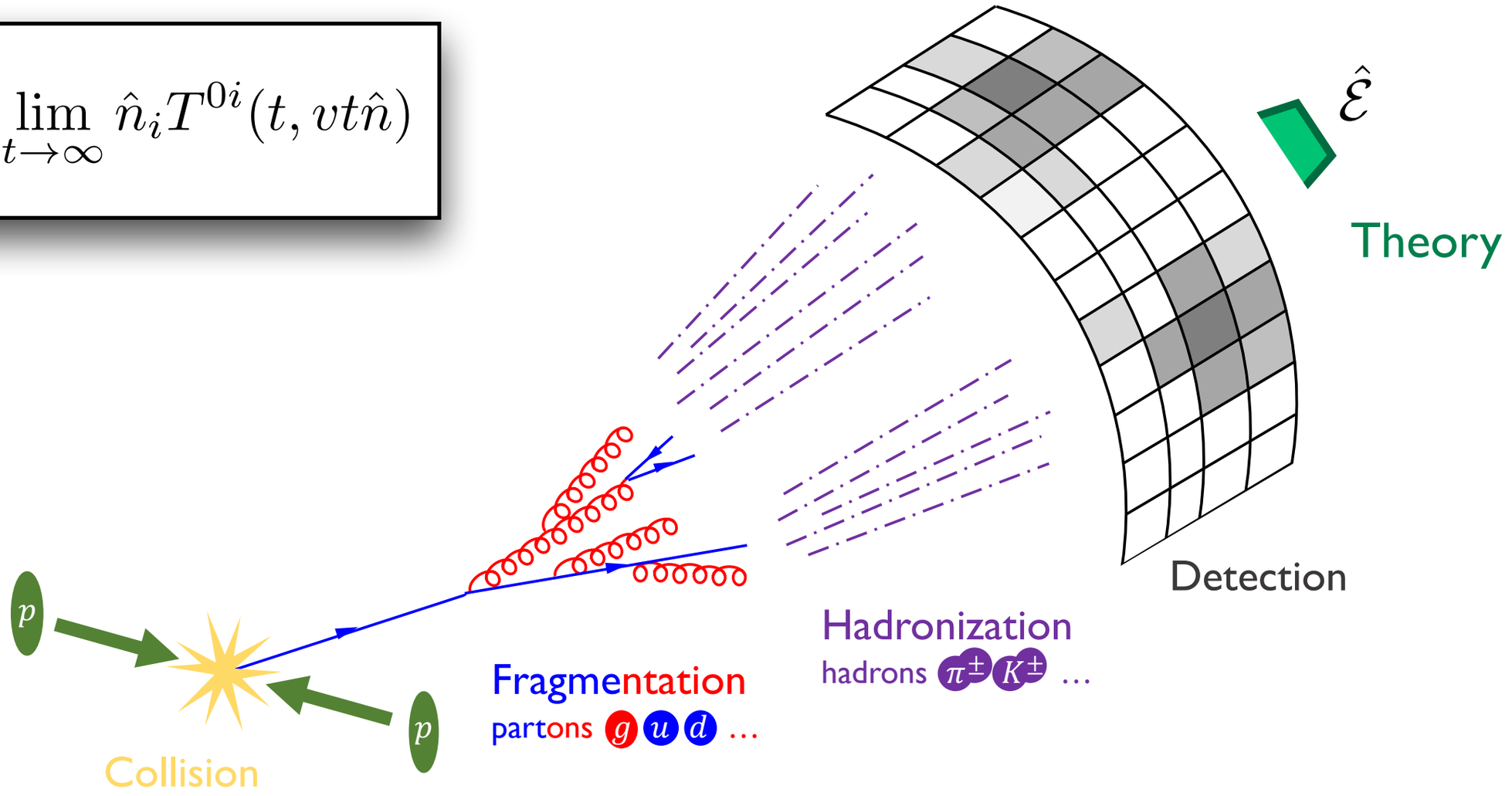




[PTK, Metodiev, Nachman, Schwartz, 1707.08600]

Stress-Energy Flow Operator

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



Stress-energy flow – measure of event/jet structure that is robust to non-perturbative and detector effects

[Sveshnikov, Tkachov, hep-ph/9512370; Hofman, Maldacena, 0803.1467; Mateu, Stewart, Thaler, 1209.3781; PTK, Metodiev, Thaler, 1712.07124, 1810.05165]