

Exploring The (Metric) Space of Collider Events

CERN Particle and Astro-Particle Physics Seminar

Eric M. Metodiev

Center for Theoretical Physics

Massachusetts Institute of Technology

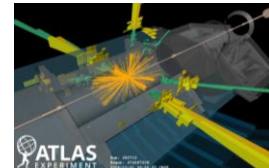
Joint work with Patrick Komiske and Jesse Thaler

[\[1902.02346\]](#)

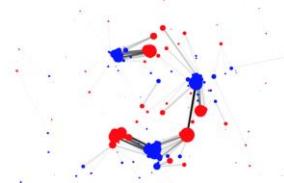
February 22, 2019

Outline

Part I Introduction



When are two events similar?

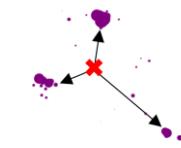


The Energy Mover's Distance

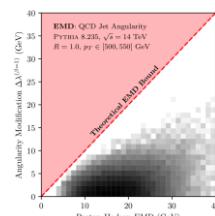


Movie Time

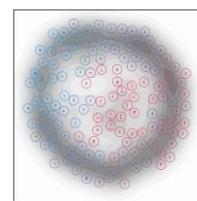
Part II Applications



Observables



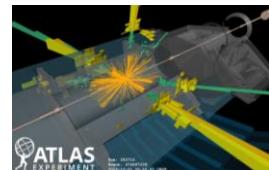
Quantifying event modifications



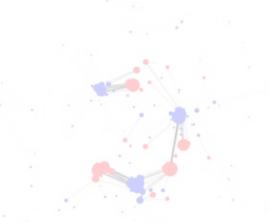
Exploring the Space of Events

Outline

Part I Introduction



When are two events similar?



The Energy Mover's Distance

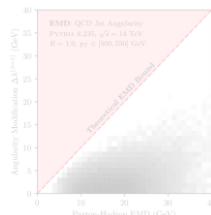


Movie Time

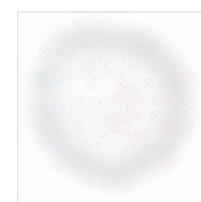
Part II Applications



Observables

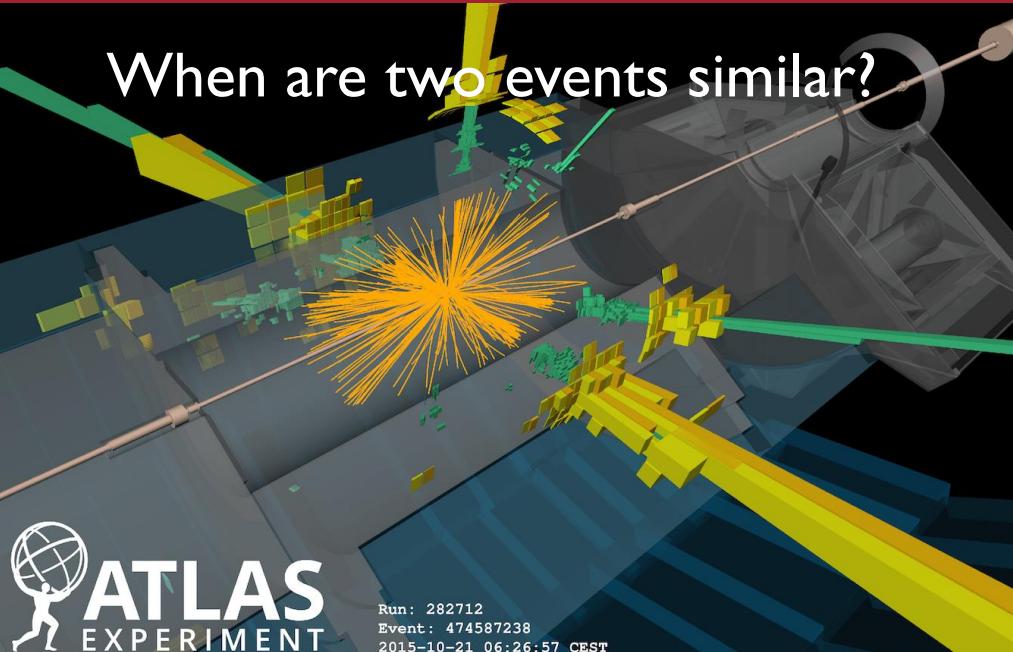


Quantifying event modifications

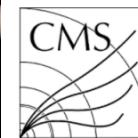


Exploring the Space of Events

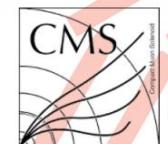
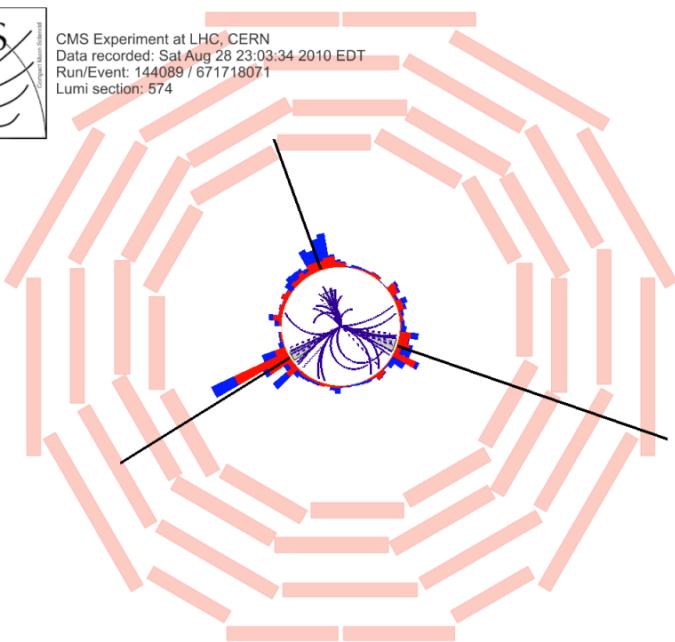
When are two events similar?



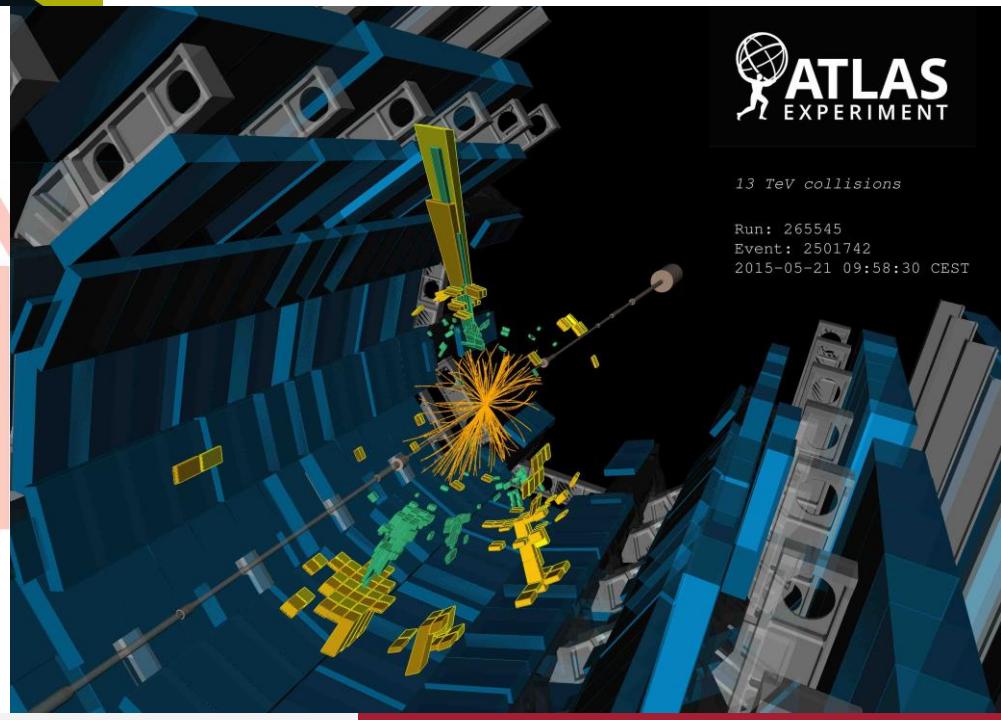
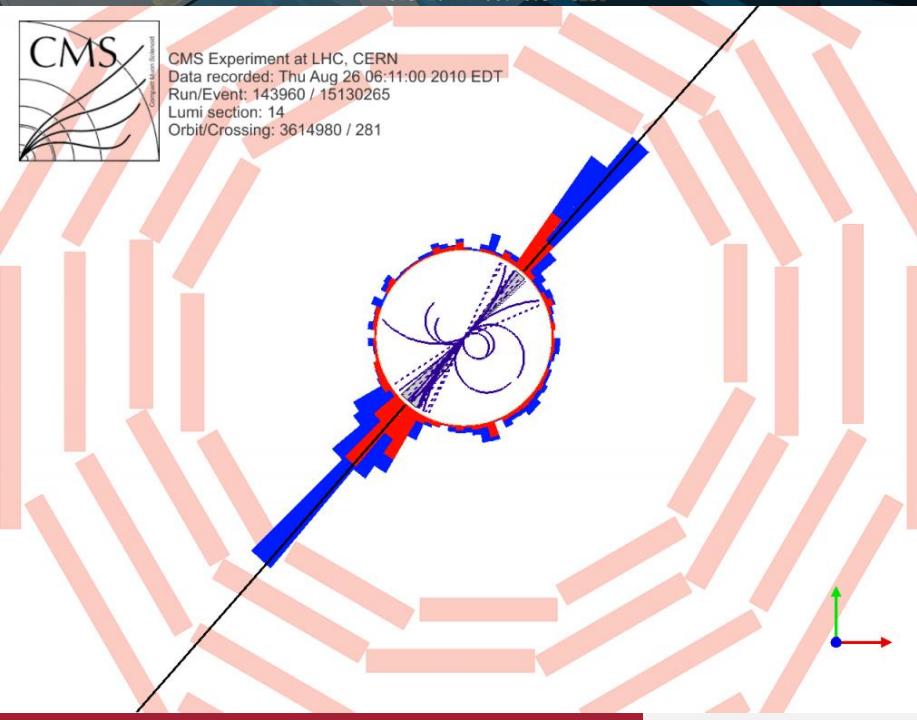
Run: 282712
Event: 474587238
2015-10-21 06:26:57 CEST



CMS Experiment at LHC, CERN
Data recorded: Sat Aug 28 23:03:34 2010 EDT
Run/Event: 144089 / 671718071
Lumi section: 574



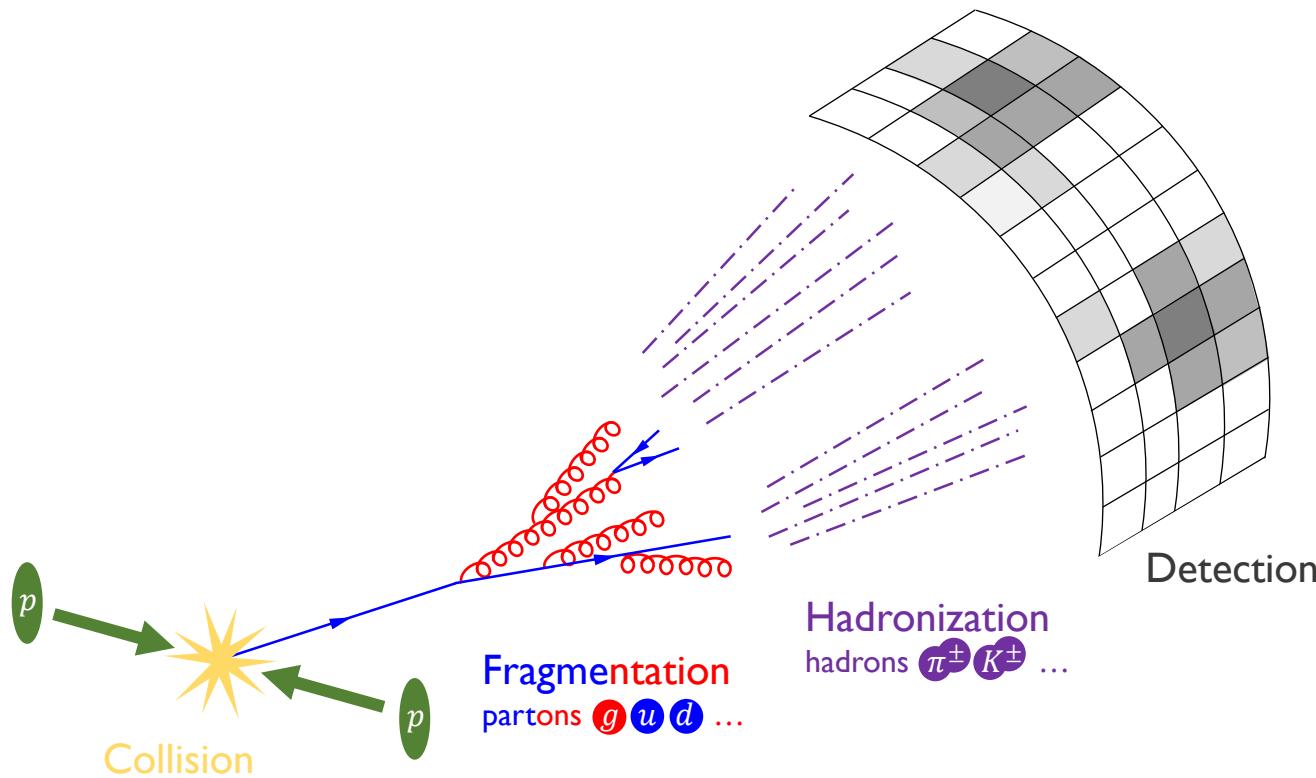
CMS Experiment at LHC, CERN
Data recorded: Thu Aug 26 06:11:00 2010 EDT
Run/Event: 143960 / 15130265
Lumi section: 14
Orbit/Crossing: 3614980 / 281



13 TeV collisions
Run: 265545
Event: 2501742
2015-05-21 09:58:30 CEST

When are two collider events similar?

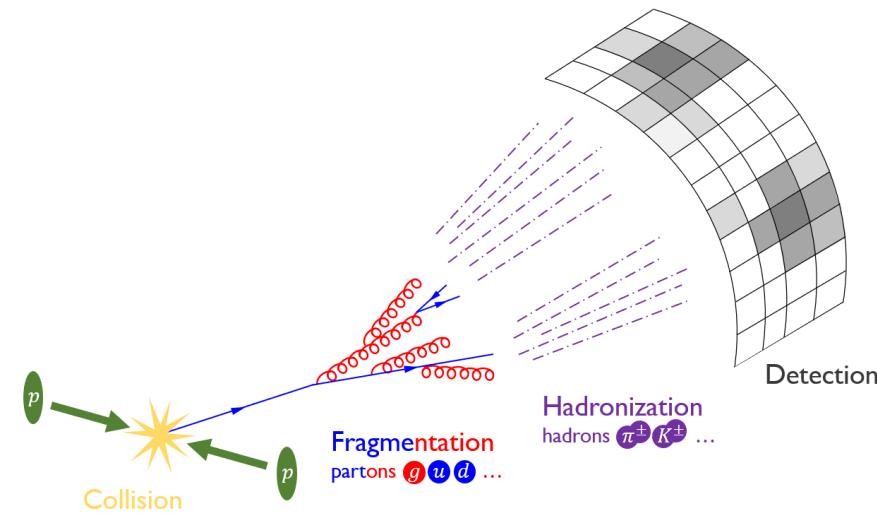
How an event gets its shape



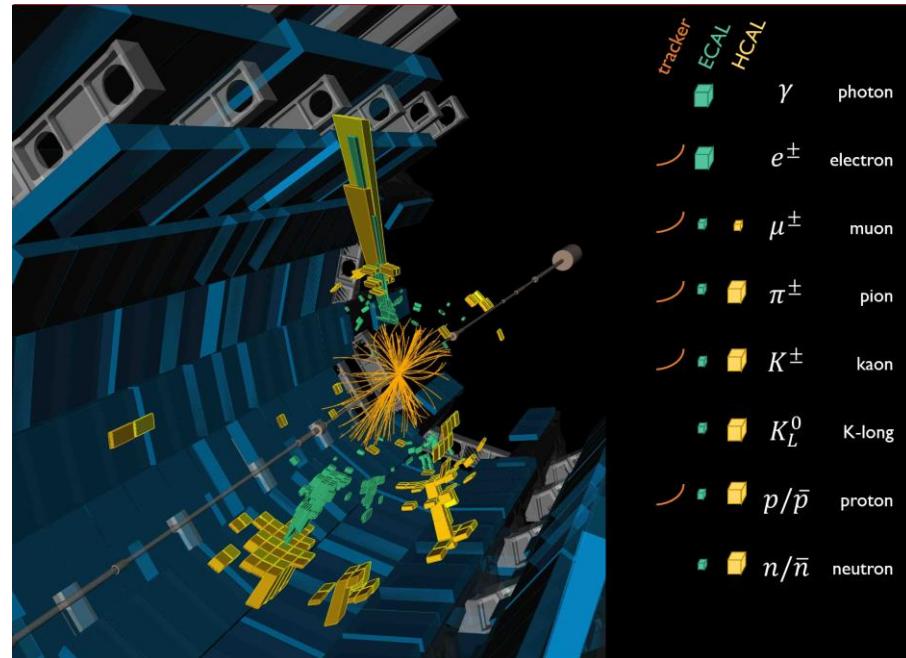
When are two collider events similar?

A *collider event* is...

Theoretically: very complicated



Experimentally: very complicated



However:

The *energy flow* (distribution of energy) is the information that is robust to:
fragmentation, hadronization, detector effects, ...

[\[N.A. Sveshnikov, F.V. Tkachov, 9512370\]](#)

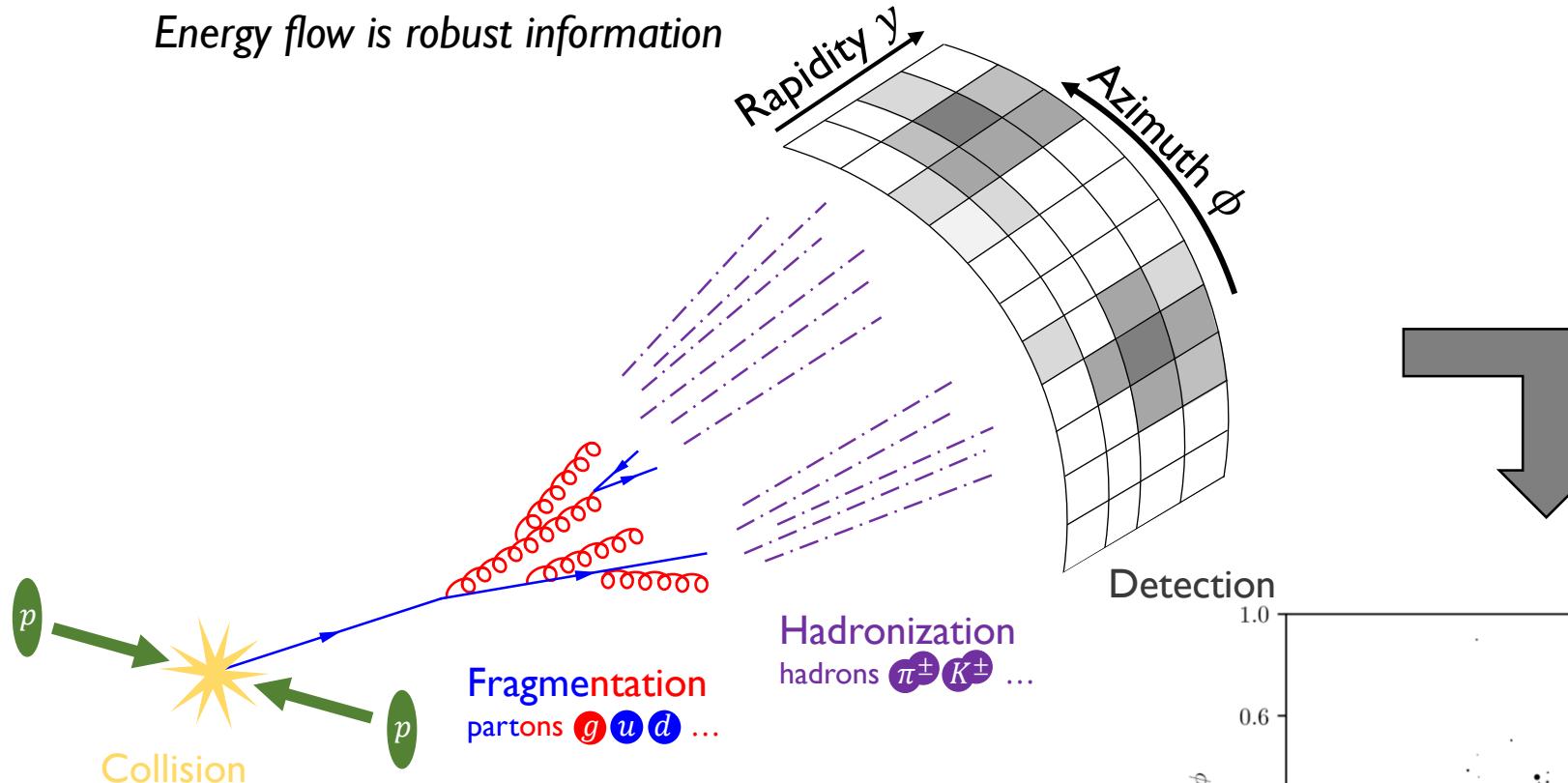
[\[F.V. Tkachov, 9601308\]](#)

[\[P.S. Cherzor, N.A. Sveshnikov, 9710349\]](#)

Energy flow \Leftrightarrow Infrared and Collinear (IRC) Safe information

When are two collider events similar?

Energy flow is robust information

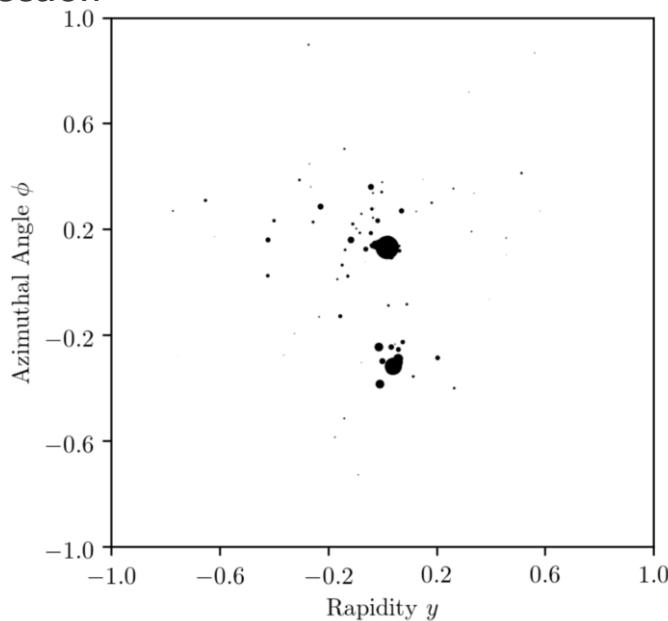


Treat events as distributions of energy:

Ignoring particle flavor, charge...

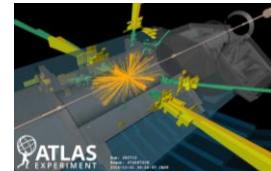
$$\sum_{i=1}^M E_i \delta(\hat{p}_i)$$

↑
energy ↑
 direction



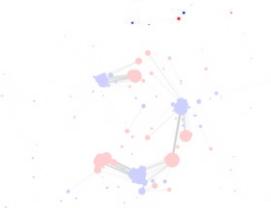
Outline

Part I Introduction



When are two events similar?

When they have similar distributions of energy



The Energy Mover's Distance

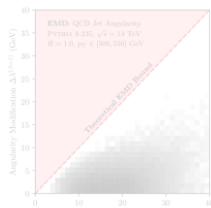


Movie Time

Part II Applications



Observables



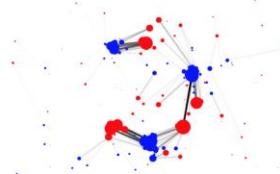
Quantifying event modifications



Exploring the Space of Events

Outline

Part I Introduction



When are two events similar?

When they have similar distributions of energy

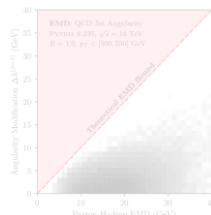
The Energy Mover's Distance

Movie Time

Part II Applications



Observables



Quantifying event modifications



Exploring the Space of Events

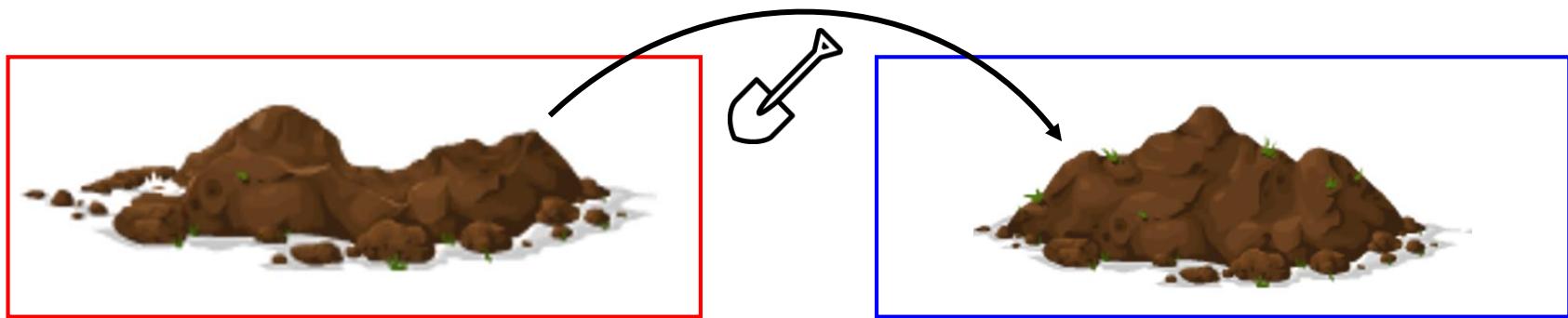
The Energy Mover's Distance

Review: *The Earth Mover's Distance*

Earth Mover's Distance: the minimum “work” ($\text{stuff} \times \text{distance}$) to rearrange one pile of dirt into another

[S. Peleg, M. Werman, H. Rom]

[Y. Rubner, C. Tomasi, and L.J. Guibas]



Metric on the space of (normalized) distributions: *symmetric, non-negative, triangle inequality*

Distributions are close in EMD \Leftrightarrow their expectation values are close.

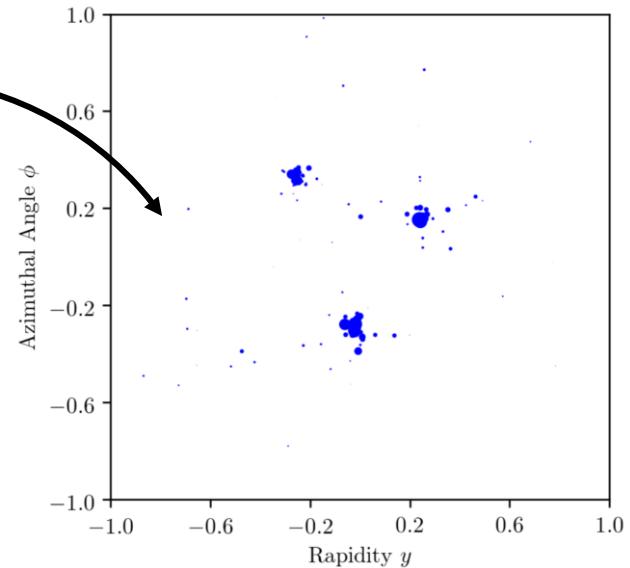
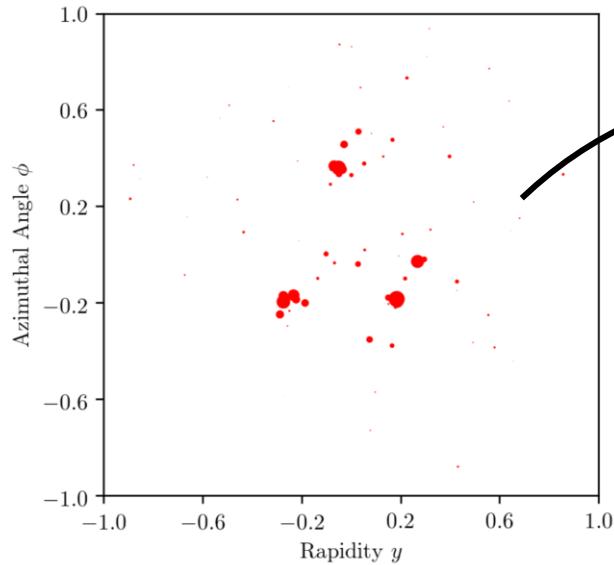
Also known as the 1-Wasserstein metric.

The Energy Mover's Distance

From Earth to Energy

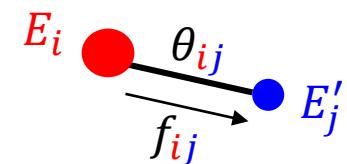
Energy Mover's Distance: the minimum “work” (**energy** \times angle) to rearrange one event (pile of energy) into another

[P.T. Komiske, EMM, J. Thaler, 1902.02346]



$$\text{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_{i=1}^M \sum_{j=1}^{M'} f_{ij} \frac{\theta_{ij}}{R}$$

Difference in
radiation pattern

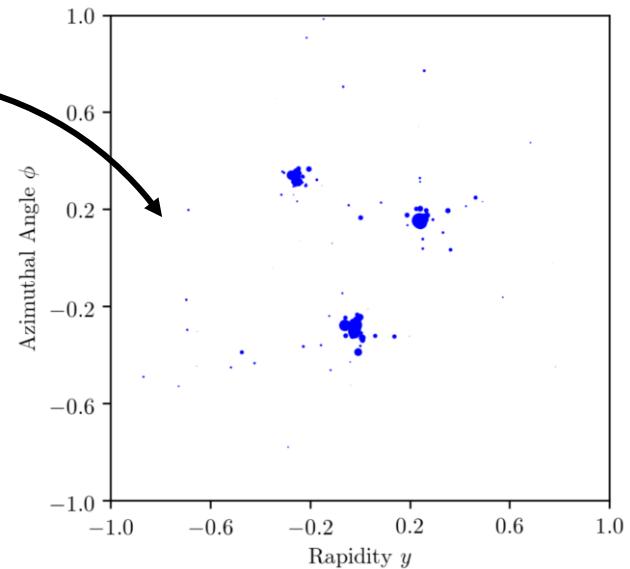
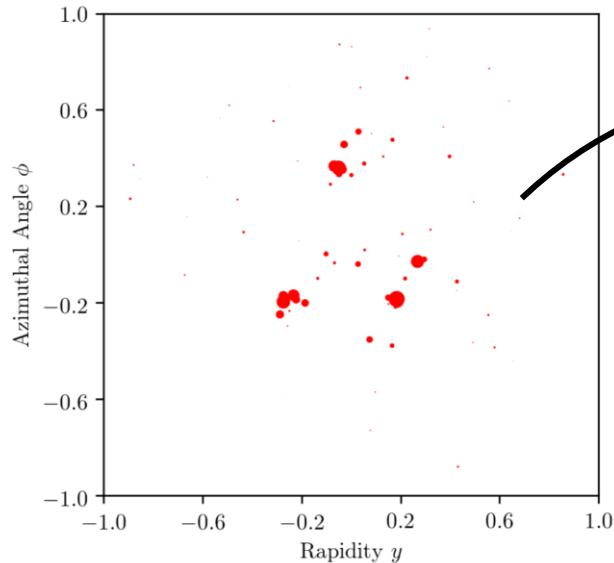


The Energy Mover's Distance

From Earth to Energy

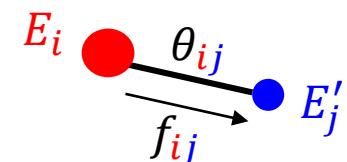
Energy Mover's Distance: the minimum “work” (**energy** \times angle) to rearrange one event (pile of energy) into another

[P.T. Komiske, EMM, J. Thaler, 1902.02346]



$$\text{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_{i=1}^M \sum_{j=1}^{M'} f_{ij} \frac{\theta_{ij}}{R} + \left| \sum_{i=1}^M E_i - \sum_{j=1}^{M'} E'_j \right|$$

Difference in radiation pattern Difference in total energy

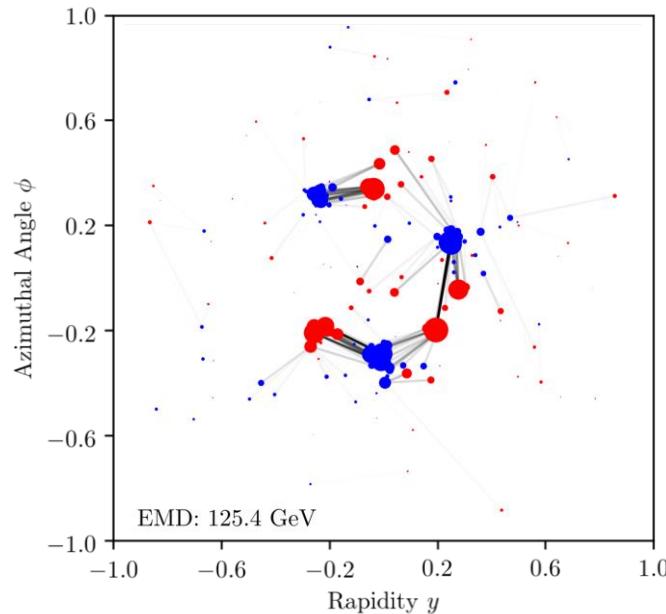


The Energy Mover's Distance

From Earth to Energy

Energy Mover's Distance: the minimum “work” (**energy** \times angle) to rearrange one event (pile of energy) into another

[P.T. Komiske, EMM, J. Thaler, 1902.02346]



EMD has dimensions of energy

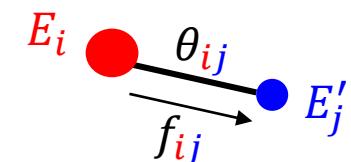
True metric as long as $R \geq \frac{1}{2} \theta_{\max}$
 $R \geq$ the jet radius, for conical jets

Solvable via Optimal Transport problem.

~1ms to compute EMD for two jets with 100 particles.

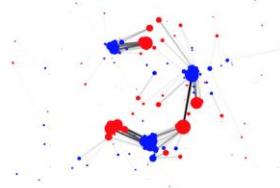
$$\text{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f\}} \sum_{i=1}^M \sum_{j=1}^{M'} f_{ij} \frac{\theta_{ij}}{R} + \left| \sum_{i=1}^M E_i - \sum_{j=1}^{M'} E'_j \right|$$

Difference in radiation pattern Difference in total energy



Outline

Part I Introduction



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

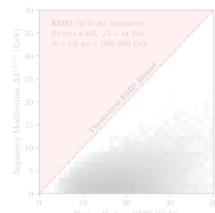
Work to rearrange one event into another.

Movie Time

Part II Applications



Observables



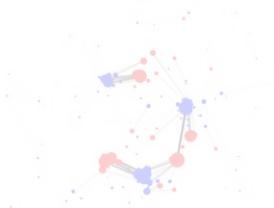
Quantifying event modifications



Exploring the Space of Events

Outline

Part I Introduction



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

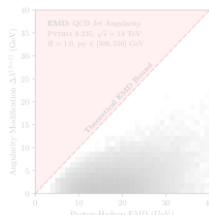
Work to rearrange one event into another.

Movie Time

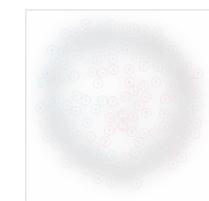
Part II Applications



Observables



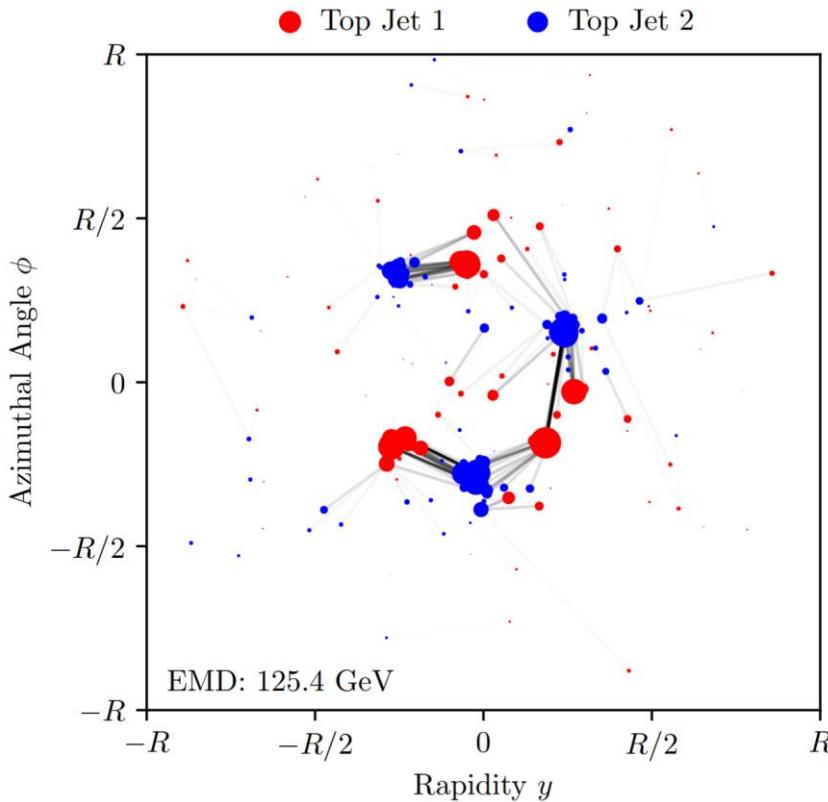
Quantifying event modifications



Exploring the Space of Events

Movie Time: Visualizing the EMD

Taking a walk in the space of events

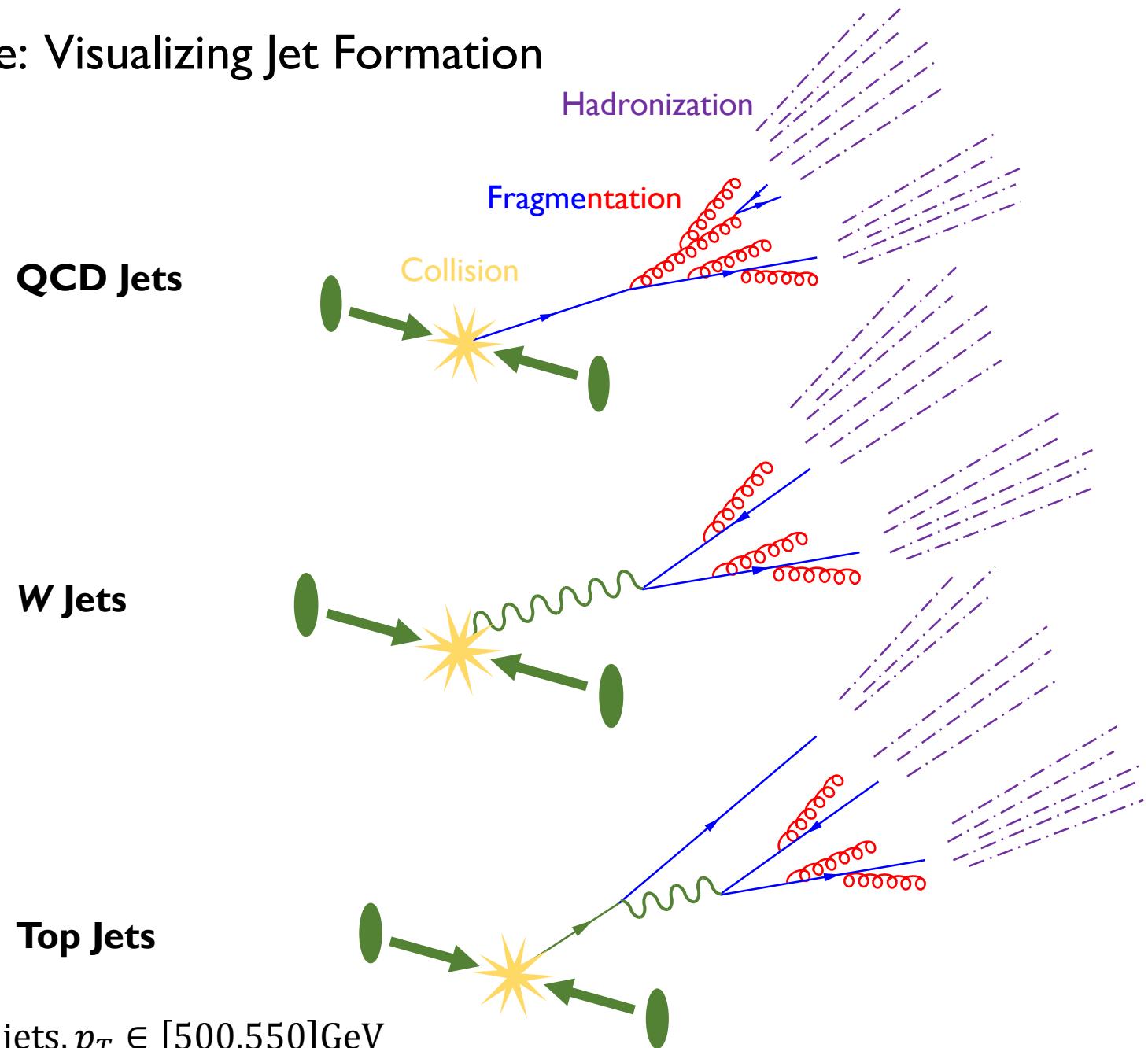


EMD is the cost of an optimal transport problem.

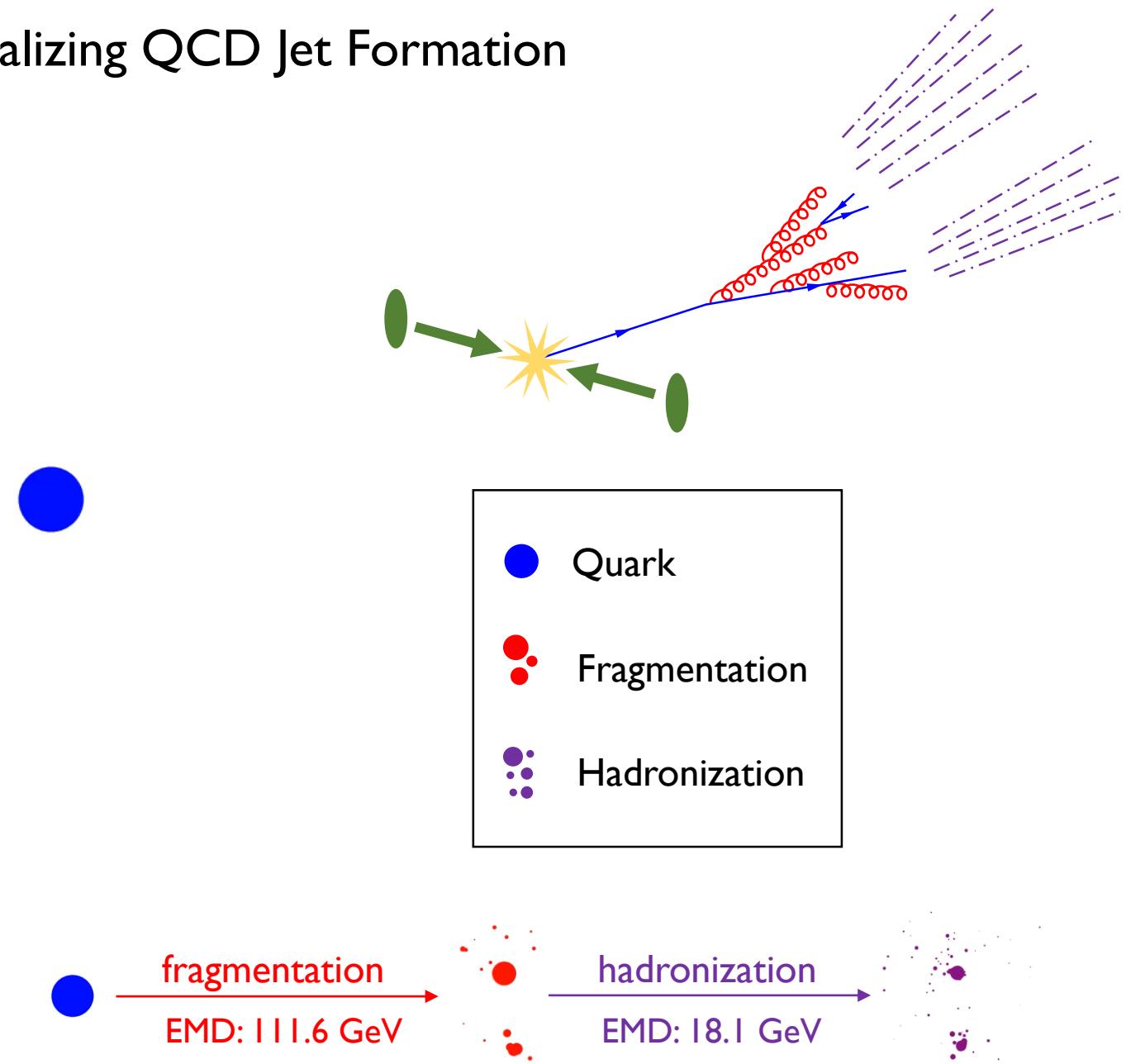
We also get the *shortest path* between the events.

Interpolate along path to visualize!

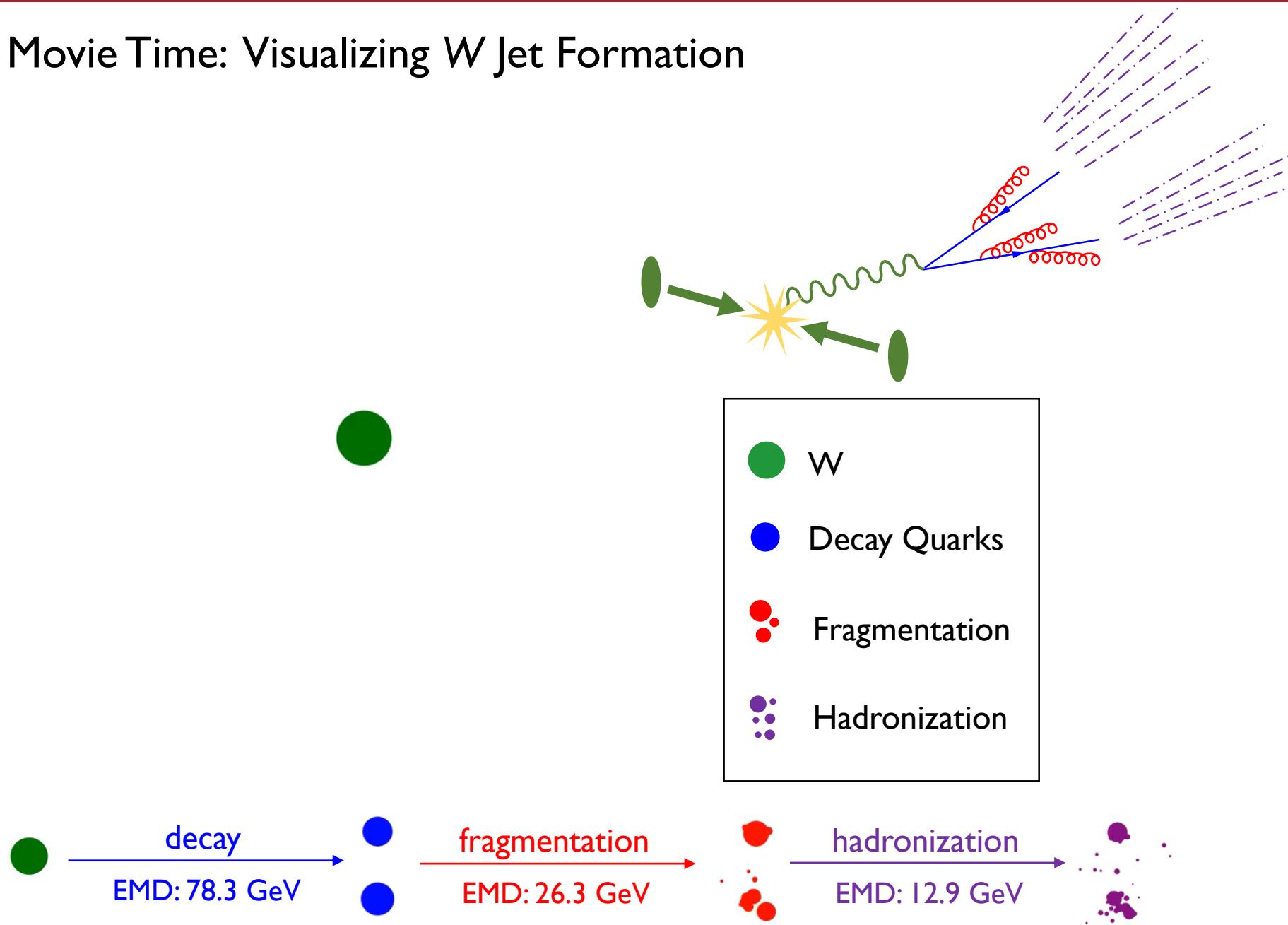
Movie Time: Visualizing Jet Formation



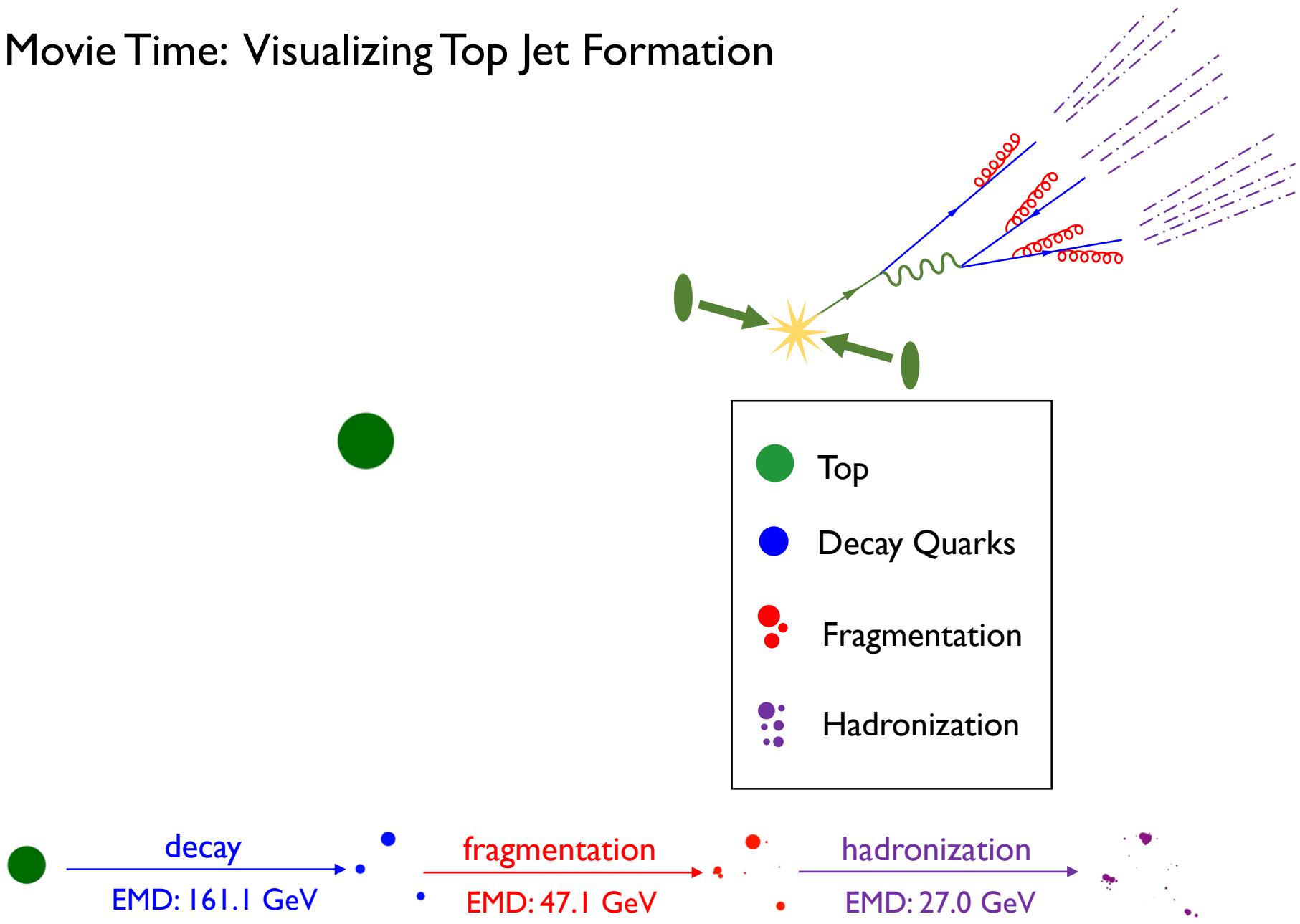
Movie Time: Visualizing QCD Jet Formation



Movie Time: Visualizing W Jet Formation

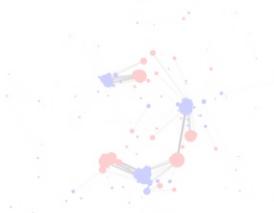


Movie Time: Visualizing Top Jet Formation



Outline

Part I Introduction



When are two events similar?

When they have similar distributions of energy

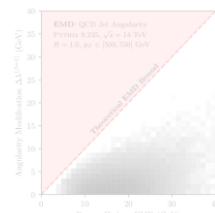
The Energy Mover's Distance

Work to rearrange one event into another.

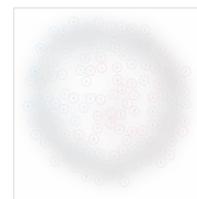
Movie Time

Visualize energy movement and jet formation.

Observables



Part II Applications

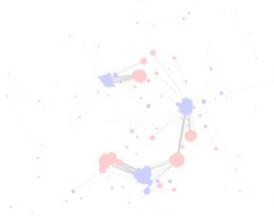


Quantifying event modifications

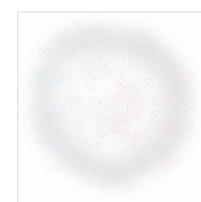
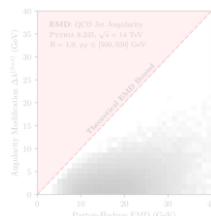
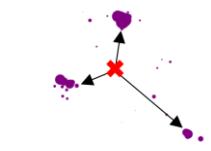
Exploring the Space of Events

Outline

Part I Introduction



Part II Applications



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

Work to rearrange one event into another.

Movie Time

Visualize energy movement and jet formation.

Observables

Quantifying event modifications

Exploring the Space of Events

Observables

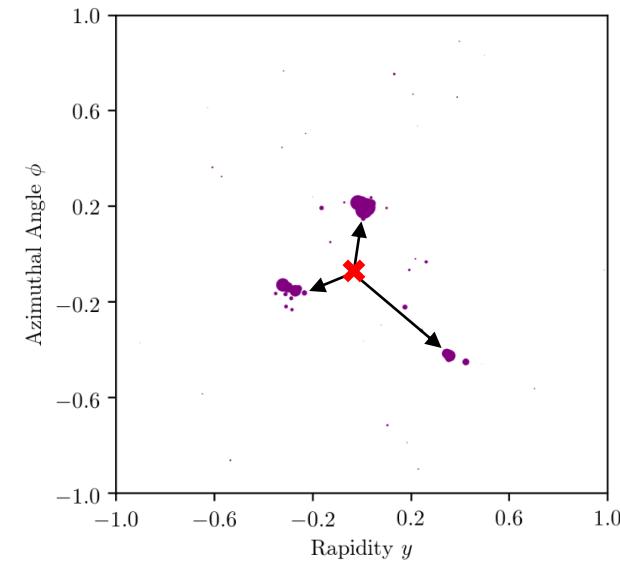
N -subjettiness:

$$\tau_N^{(\beta)} = \sum_{i=1}^M E_i \min_{N \text{ axes}} \{\theta_{1,k}^\beta, \theta_{2,k}^\beta, \dots, \theta_{N,k}^\beta\}$$

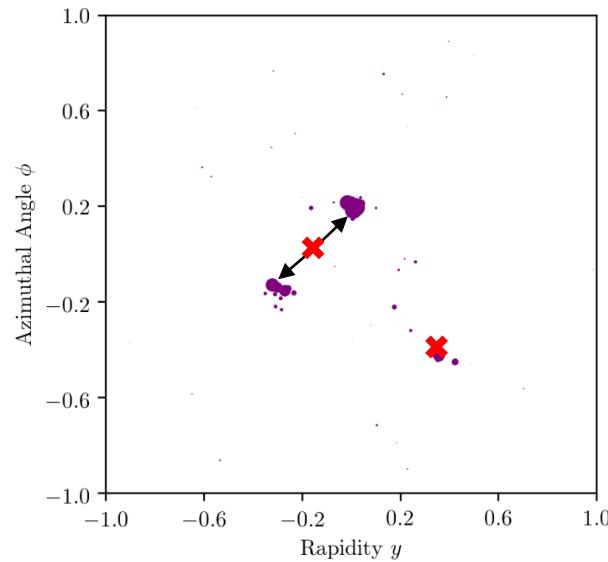
measures how well jet energy is aligned into N subjets

[\[J. Thaler, K. Van Tilburg, 1011.2268\]](#)
[\[J. Thaler, K. Van Tilburg, 1108.2701\]](#)

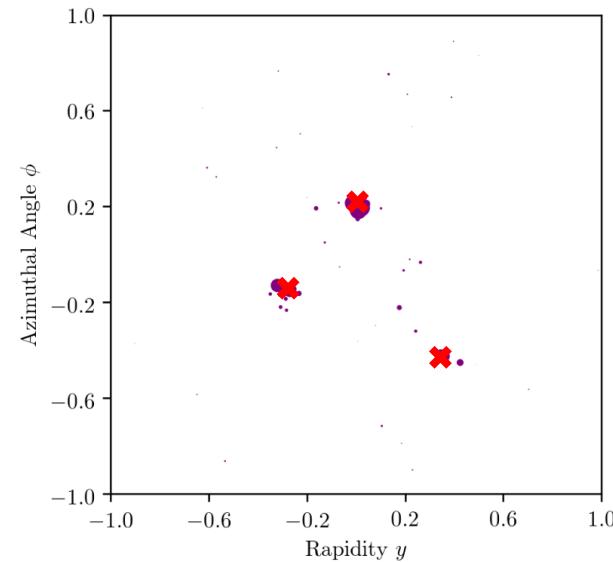
$\tau_1/E \gg 0$



$\tau_1/E > \tau_2/E \gg 0$



$\tau_3/E \simeq 0$



Observables

N -subjettiness:

$$\tau_N^{(\beta)} = \sum_{i=1}^M E_i \min_{N \text{ axes}} \{\theta_{1,k}^\beta, \theta_{2,k}^\beta, \dots, \theta_{N,k}^\beta\}$$

measures how well jet energy is aligned into N subjets

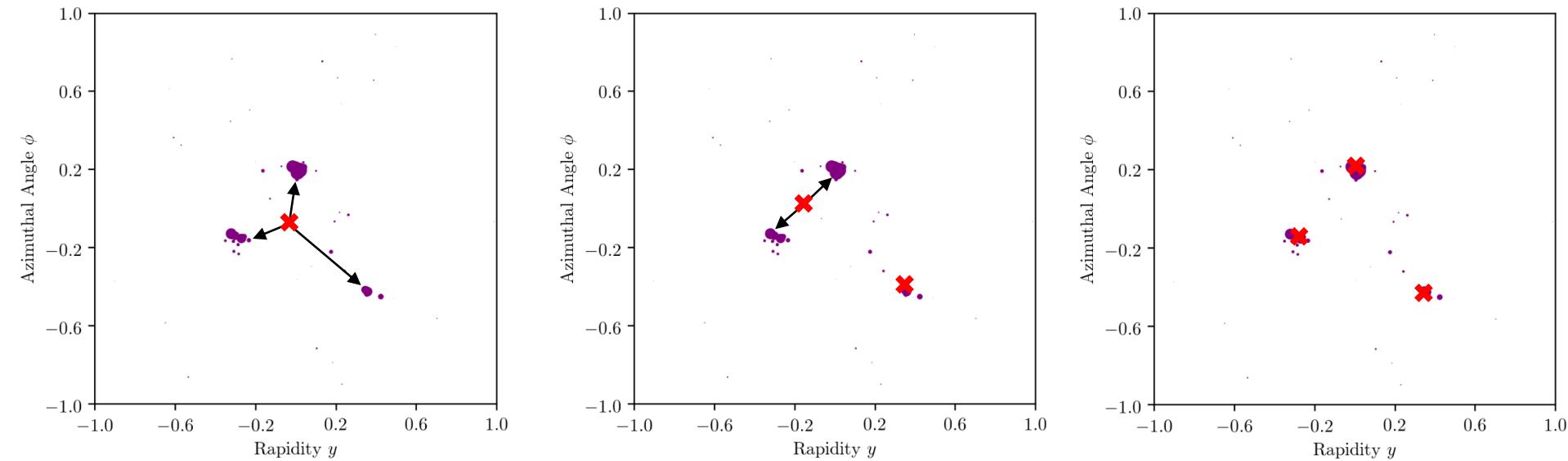
[\[J. Thaler, K. Van Tilburg, 1011.2268\]](#)

[\[J. Thaler, K. Van Tilburg, 1108.2701\]](#)

$$\tau_1/E \gg 0$$

$$\tau_1/E > \tau_2/E \gg 0$$

$$\tau_3/E \simeq 0$$



N -subjettiness is the EMD between the event and the closest N -particle event.

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

$\beta \neq 1$ corresponds to other p -Wasserstein distances with $p = \beta$.

Observables

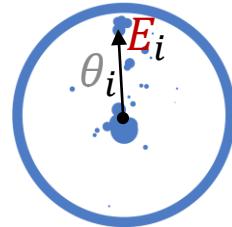
Getting quantitative

Take any additive IRC-safe observable: $\mathcal{O}(\mathcal{E}) = \sum_{i=1}^M E_i \Phi(\hat{p}_i)$

e.g. jet angularities: $\lambda^{(\beta)} = \sum_{i=1}^M E_i \theta_i^\beta$

[\[C. Berger, T. Kucs, and G. Sterman, 0303051\]](#)

[\[A. Larkoski, J. Thaler, and W. Waalewijn, 1408.3122\]](#)



Observables

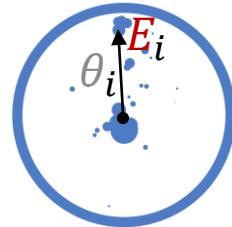
Getting quantitative

Take any additive IRC-safe observable: $\mathcal{O}(\mathcal{E}) = \sum_{i=1}^M E_i \Phi(\hat{p}_i)$

e.g. jet angularities: $\lambda^{(\beta)} = \sum_{i=1}^M E_i \theta_i^\beta$

[C. Berger, T. Kucs, and G. Sterman, 0303051]

[A. Larkoski, J. Thaler, and W. Waalewijn, 1408.3122]



Via the Kantorovich-Rubinstein dual formulation of EMD:

Earth Mover's
Distance

$$\text{EMD}(\mathcal{E}, \mathcal{E}') \geq \frac{1}{RL} \left| \sum_{i=1}^M E_i \Phi(\hat{p}_i) - \sum_{j=1}^{M'} E'_j \Phi(\hat{p}'_j) \right| = \frac{1}{RL} |\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')|$$

Difference in
observable values

“Lipschitz constant” of Φ
i.e. bound on its derivative

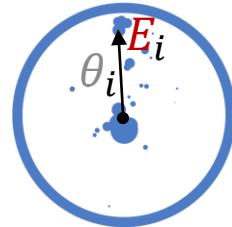
Observables

Getting quantitative

Take any additive IRC-safe observable: $\mathcal{O}(\mathcal{E}) = \sum_{i=1}^M E_i \Phi(\hat{p}_i)$

e.g. jet angularities: $\lambda^{(\beta)} = \sum_{i=1}^M E_i \theta_i^\beta$

[C. Berger, T. Kucs, and G. Sterman, 0303051]
[\[A. Larkoski, J. Thaler, and W. Waalewijn, 1408.3122\]](#)



Via the Kantorovich-Rubinstein dual formulation of EMD:

Earth Mover's

Distance

$$\text{EMD}(\mathcal{E}, \mathcal{E}') \geq \frac{1}{RL} \left| \sum_{i=1}^M E_i \Phi(\hat{p}_i) - \sum_{j=1}^{M'} E'_j \Phi(\hat{p}'_j) \right| = \frac{1}{RL} |\mathcal{O}(\mathcal{E}) - \mathcal{O}(\mathcal{E}')|$$

Difference in
observable values

“Lipschitz constant” of Φ
i.e. bound on its derivative

For $\beta \geq 1$ jet angularities, $L = \beta/R$ over the jet cone, so:

$$|\lambda^{(\beta)}(\mathcal{E}) - \lambda^{(\beta)}(\mathcal{E}')| \leq \beta \text{EMD}(\mathcal{E}, \mathcal{E}')$$

The EMD provides a robust upper bound to any modifications of these observables.

Observables

Key idea: Energy-weighted angular structures contain all the IRC-safe information.

$$\frac{1}{RL} \left| \sum_{i=1}^M E_i \Phi(\hat{p}_i) - \sum_{j=1}^{M'} E'_j \Phi(\hat{p}_i) \right| \leq \text{EMD}(\mathcal{E}, \mathcal{E}')$$

Theorem: Any infrared and collinear safe observable \mathcal{O} can be approximated arbitrarily well as:

$$\mathcal{O}(p_1, \dots, p_M) = F \left(\sum_{i=1}^M E_i \vec{\Phi}(\hat{p}_i) \right)$$

for some $\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^\ell$ and $F: \mathbb{R}^\ell \rightarrow \mathbb{R}$ and sufficiently large ℓ .

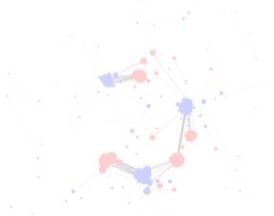
[\[M. Zaheer, et al., 1703.06114\]](#)

[\[P.T. Komiske, EMM, J. Thaler, 1810.05165\]](#)

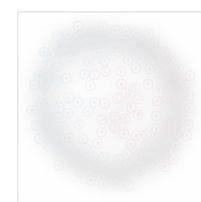
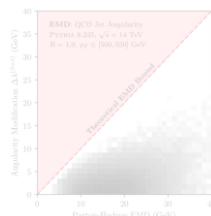
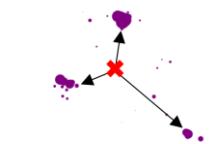
Events close in EMD are close in all infrared and collinear safe information!

Outline

Part I Introduction



Part II Applications



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

Work to rearrange one event into another.

Movie Time

Visualize energy movement and jet formation.

Observables

Conceptually rich connections to EMD.

Quantifying event modifications

Exploring the Space of Events

Outline

Part I Introduction



When are two events similar?
When they have similar distributions of energy

The Energy Mover's Distance

Work to rearrange one event into another.

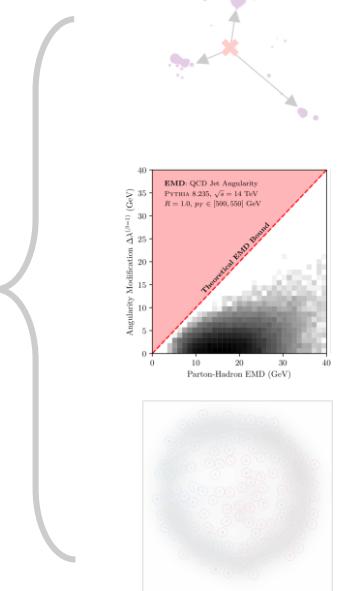
Movie Time

Visualize energy movement and jet formation.

Observables

Conceptually rich connections to EMD.

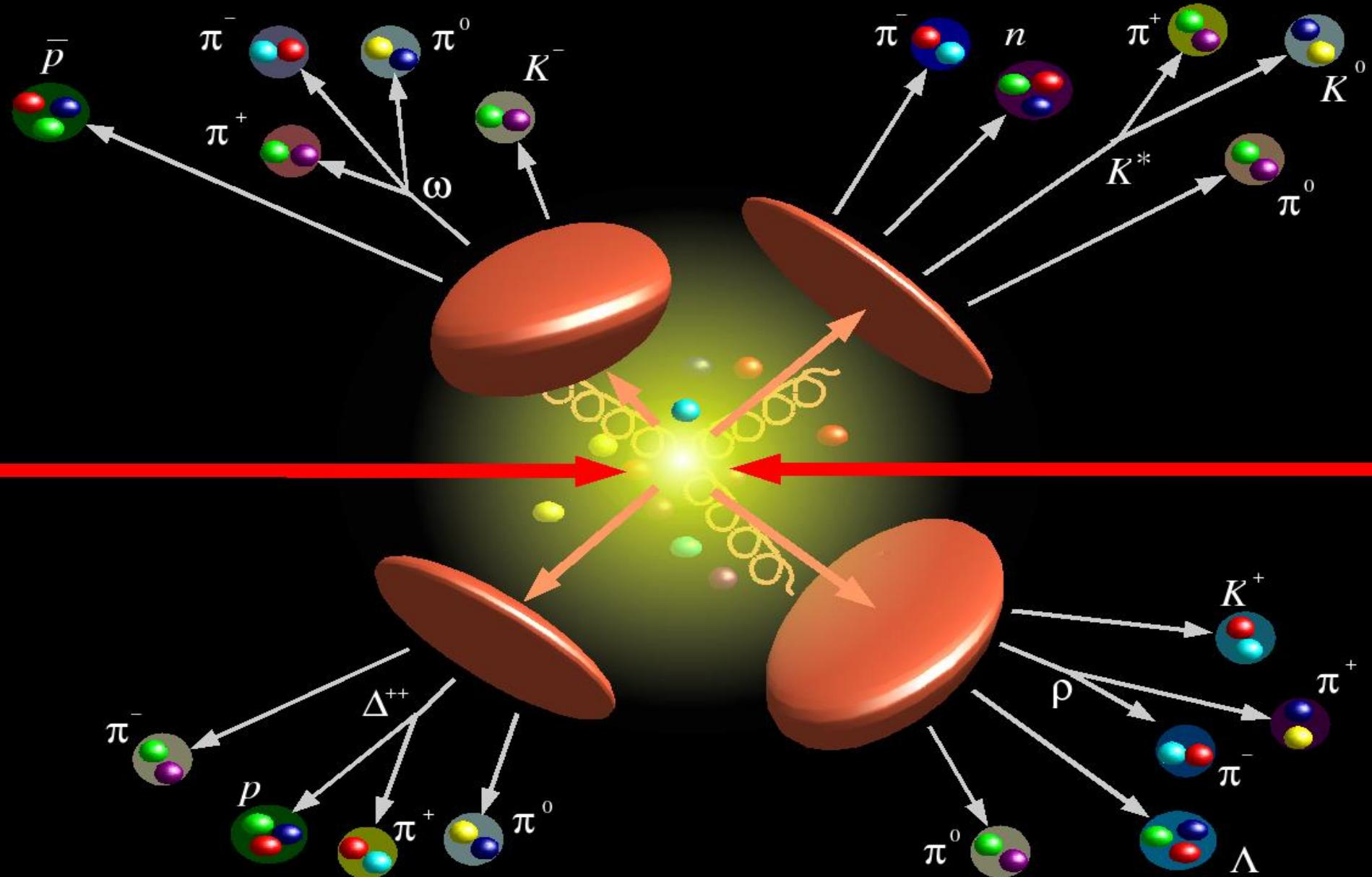
Part II Applications



Quantifying event modifications

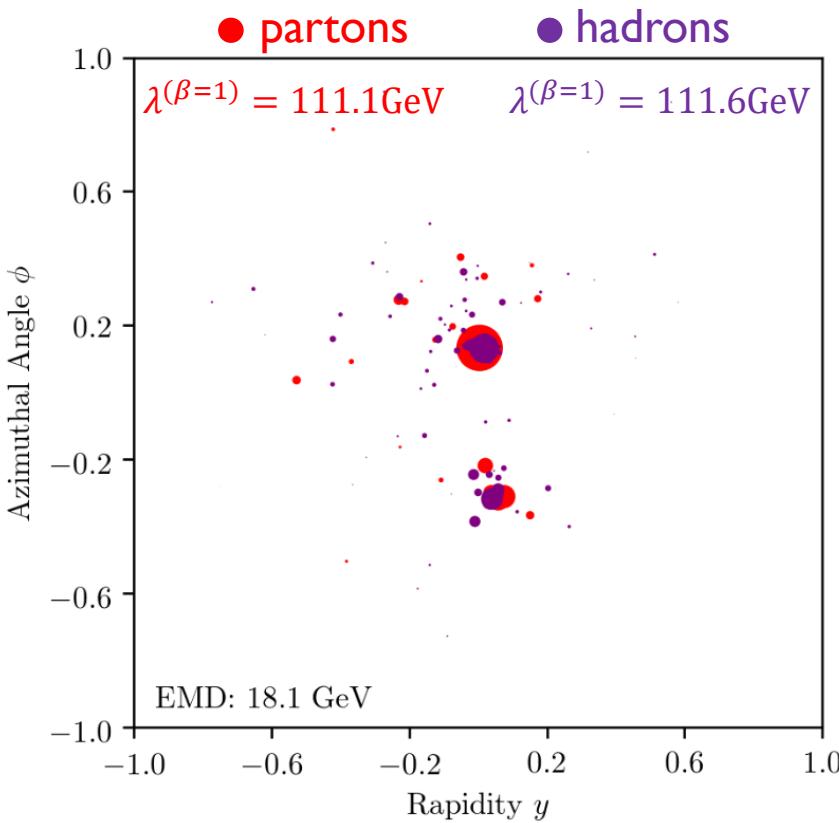
Exploring the Space of Events

Quantifying event modifications: Hadronization



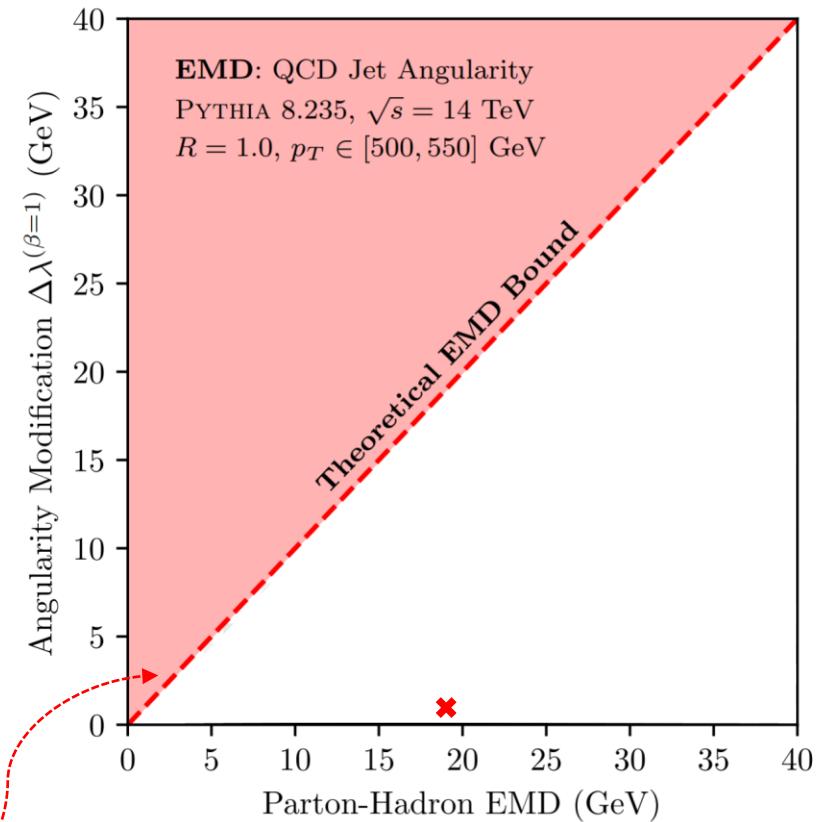
Quantifying event modifications: Hadronization

$$\lambda^{(\beta=1)} = \sum_{i=1}^M E_i \theta_i$$



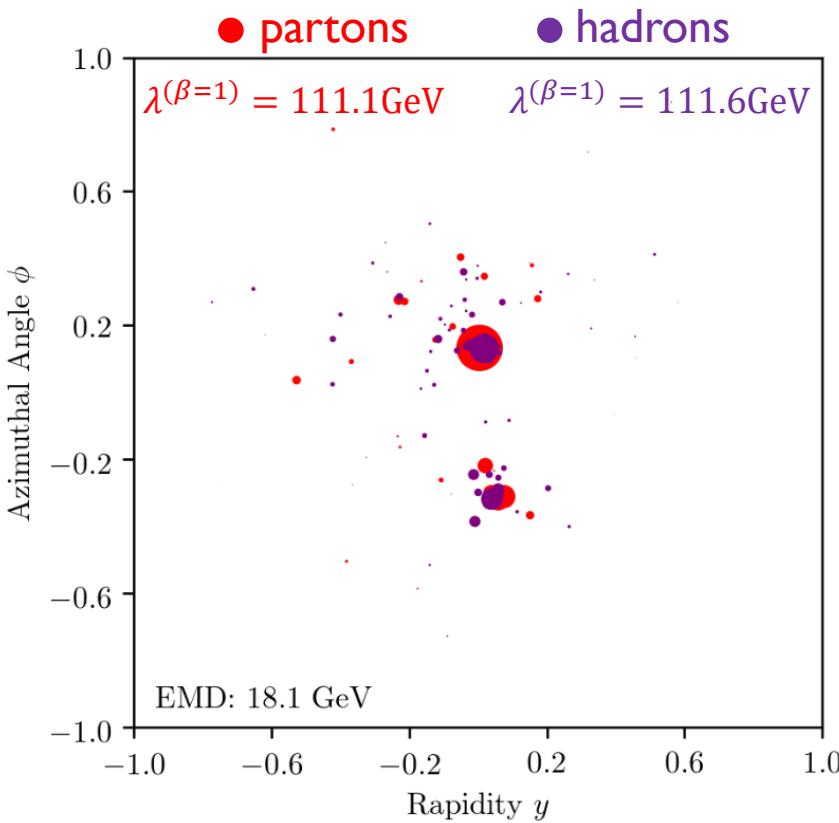
$$\begin{aligned}\mathcal{E} &= \mathcal{E}_{\text{partons}} \\ \mathcal{E}' &= \mathcal{E}_{\text{hadrons}}\end{aligned}$$

$$|\lambda^{(\beta=1)}(\mathcal{E}) - \lambda^{(\beta=1)}(\mathcal{E}')| \leq \text{EMD}(\mathcal{E}, \mathcal{E}')$$



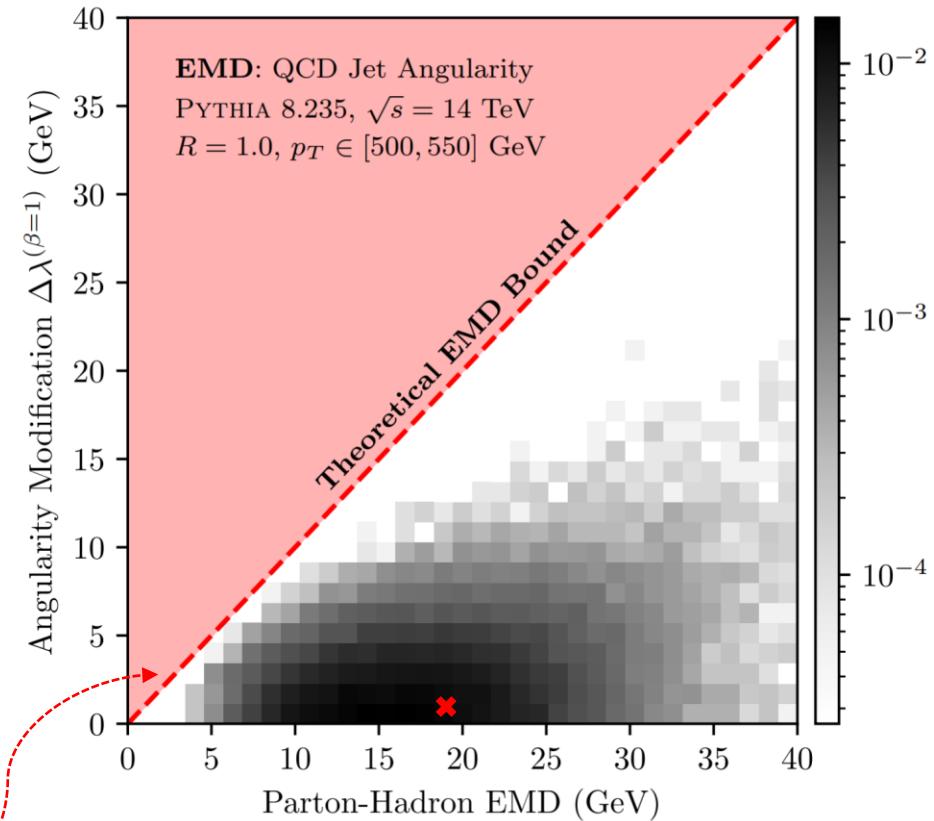
Quantifying event modifications: Hadronization

$$\lambda^{(\beta=1)} = \sum_{i=1}^M E_i \theta_i$$

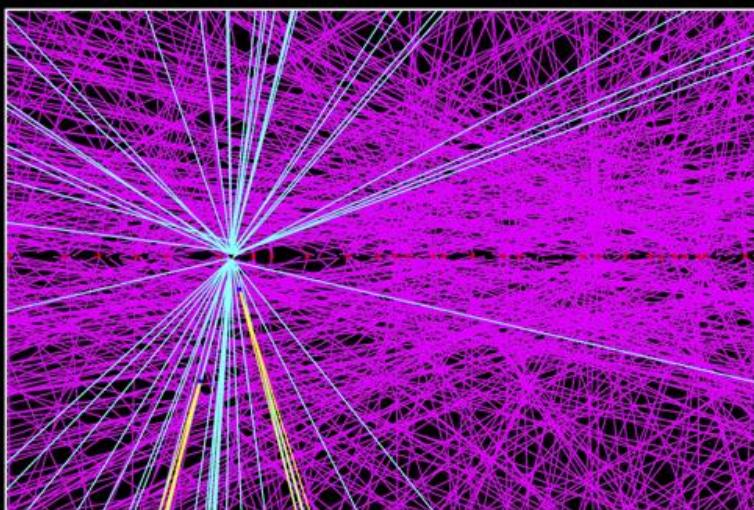


$$\begin{aligned}\mathcal{E} &= \mathcal{E}_{\text{partons}} \\ \mathcal{E}' &= \mathcal{E}_{\text{hadrons}}\end{aligned}$$

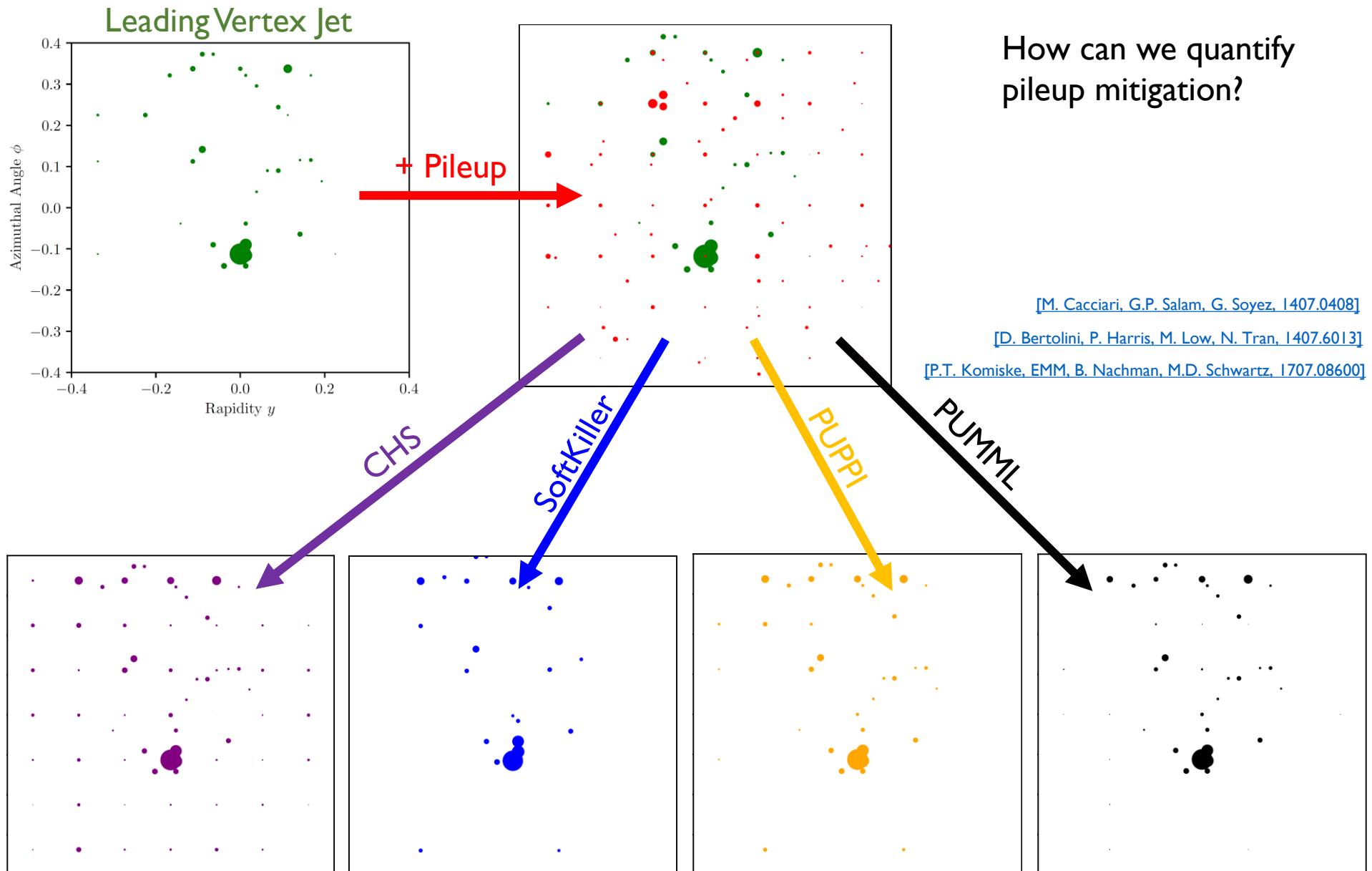
$$|\lambda^{(\beta=1)}(\mathcal{E}) - \lambda^{(\beta=1)}(\mathcal{E}')| \leq \text{EMD}(\mathcal{E}, \mathcal{E}')$$



Quantifying event modifications: Pileup

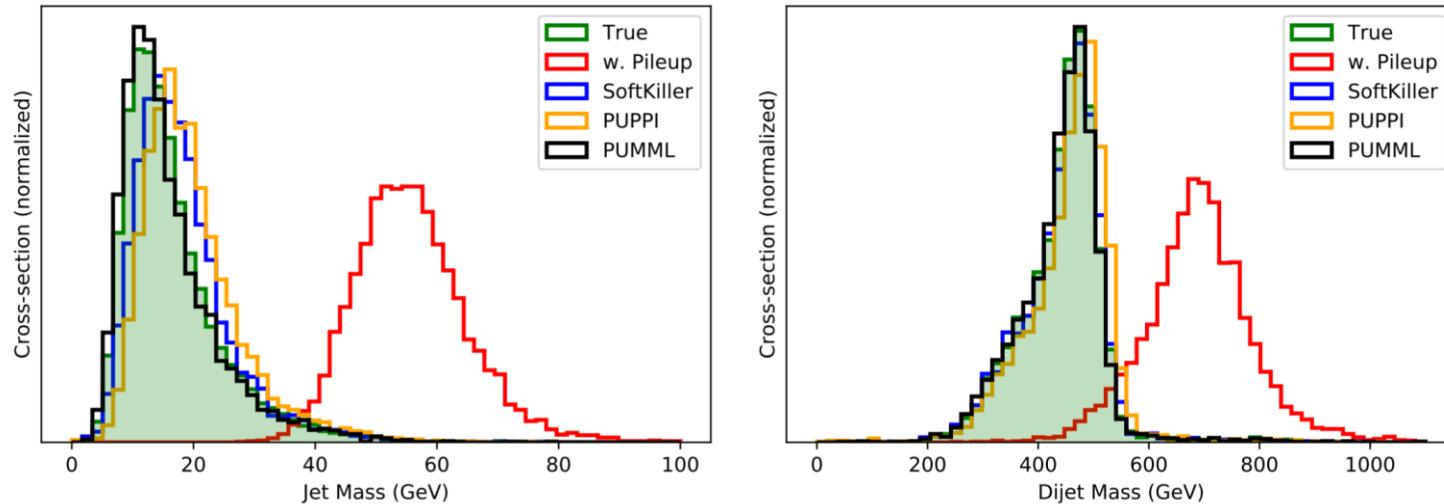


Quantifying event modifications: Pileup



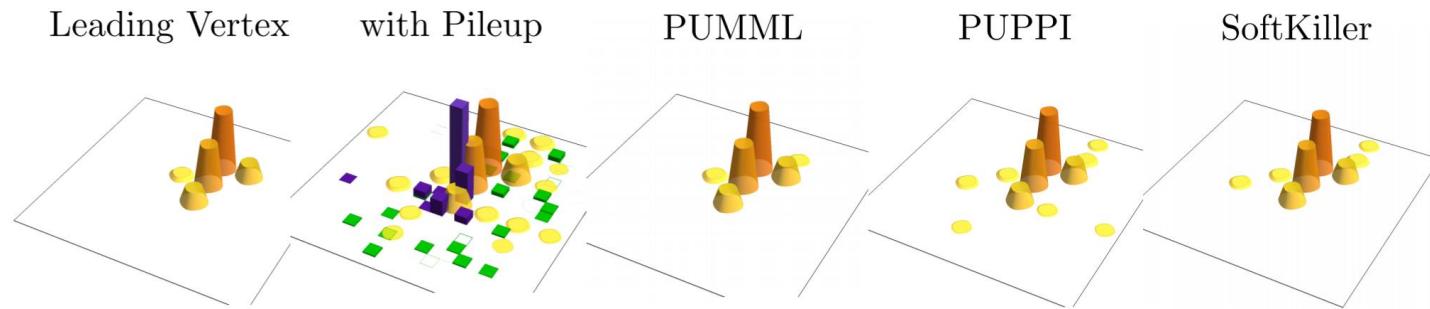
Quantifying event modifications: Pileup

Compare on a collection of observables?



Requires ad hoc choices of observables.

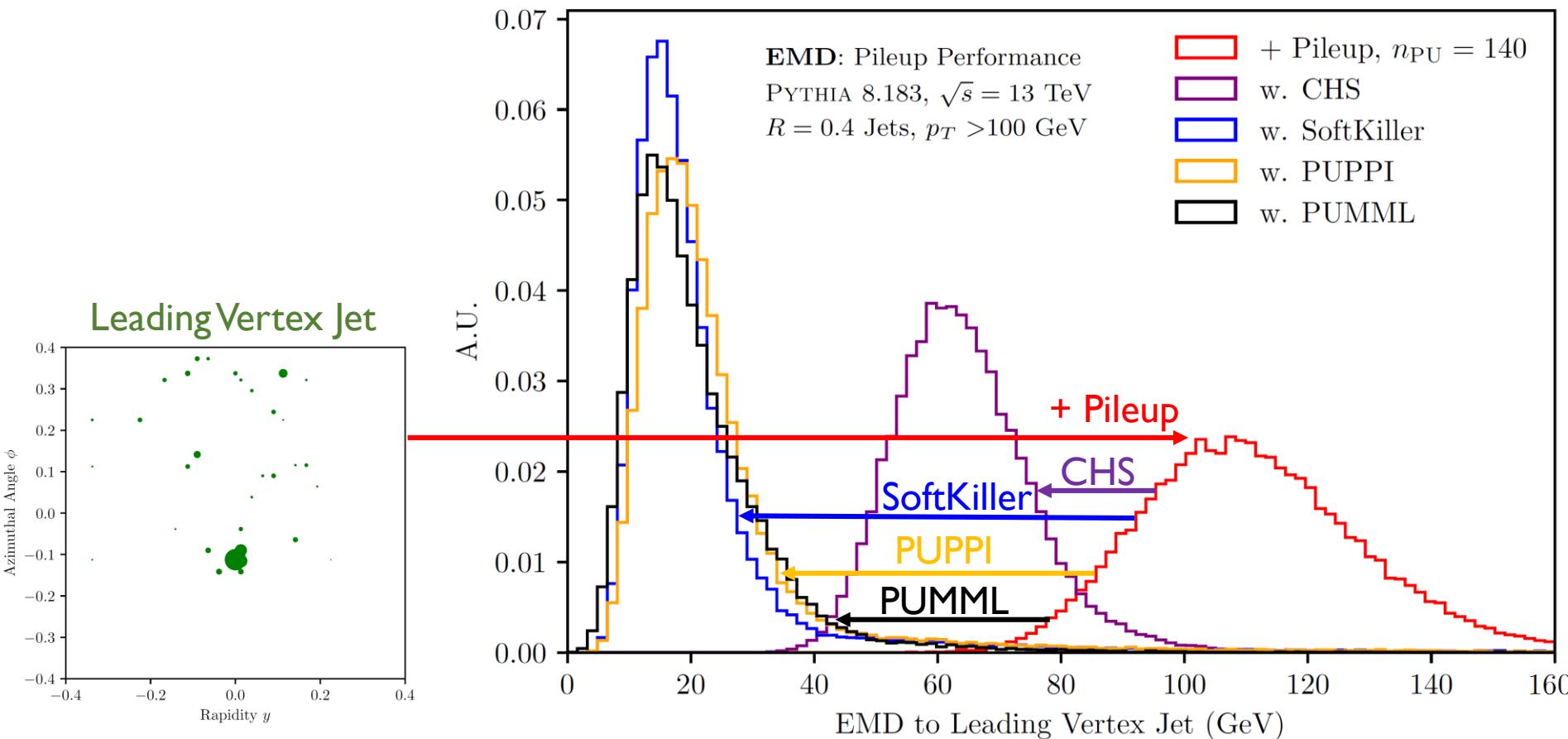
Compare calorimeter images pixel by pixel?



Discontinuous under physically-sensible single-pixel perturbations.
Undesirable behavior with increasing resolution.

Quantifying event modifications: Pileup

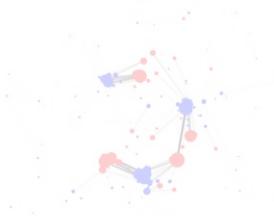
Measure pileup mitigation performance with EMD!



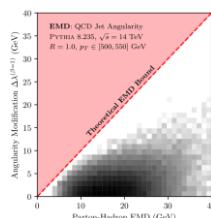
Guarantees performance on IRC safe observables.
Stable under physically-sensible perturbations.
Train to optimize EMD with machine learning?

Outline

Part I Introduction



Part II Applications



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

Work to rearrange one event into another.

Movie Time

Visualize energy movement and jet formation.

Observables

Conceptually rich connections to EMD.

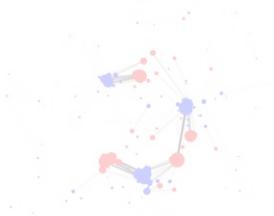
Quantifying event modifications

Hadronization, pileup, detector effects

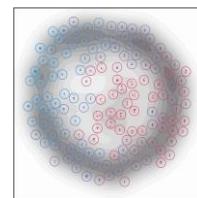
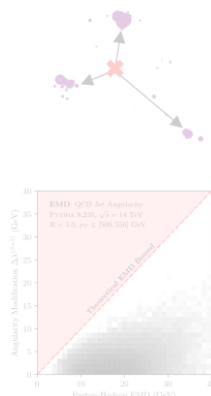
Exploring the Space of Events

Outline

Part I Introduction



Part II Applications



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

Work to rearrange one event into another.

Movie Time

Visualize energy movement and jet formation.

Observables

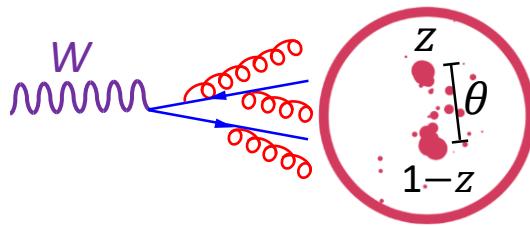
Conceptually rich connections to EMD.

Quantifying event modifications

Hadronization, pileup, detector effects

Exploring the Space of Events

Exploring the Space of Events: W jets



W jets are 2-pronged:

z : Energy Sharing of Prongs

$θ$: Angle between Prongs

$φ$: Azimuthal orientation

Constrained by W mass:

$$z(1 - z)\theta^2 = \frac{p_{\mu J}^2}{p_T^2} = \frac{m_W^2}{p_T^2}$$

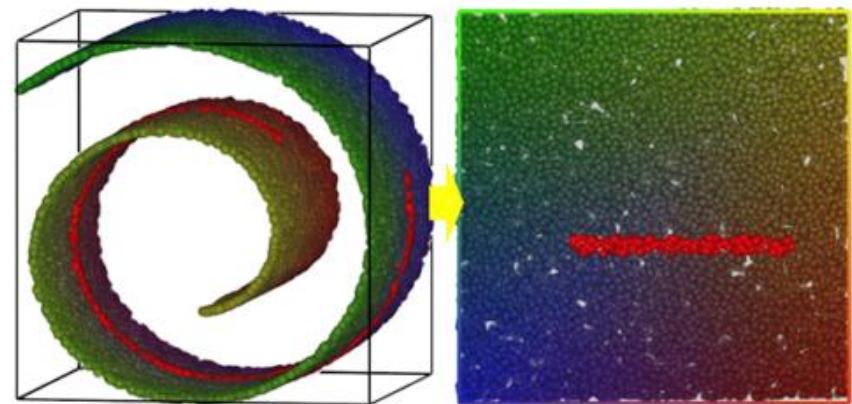
Hence we expect a **two-dimensional** space of W jets.

After $φ$ rotation: **one-dimensional**

Visualize the space of events with t-Distributed Stochastic Neighbor Embedding (t-SNE).

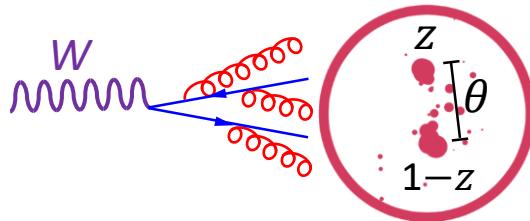
[\[L. van der Maaten, G. Hinton\]](#)

Finds an embedding into a low-dimensional manifold that respects distances.



Src: <http://web-ext.u-aizu.ac.jp/~shigeo/home.html>

Exploring the Space of Events: W jets



W jets are 2-pronged:

z : Energy Sharing of Prongs

θ : Angle between Prongs

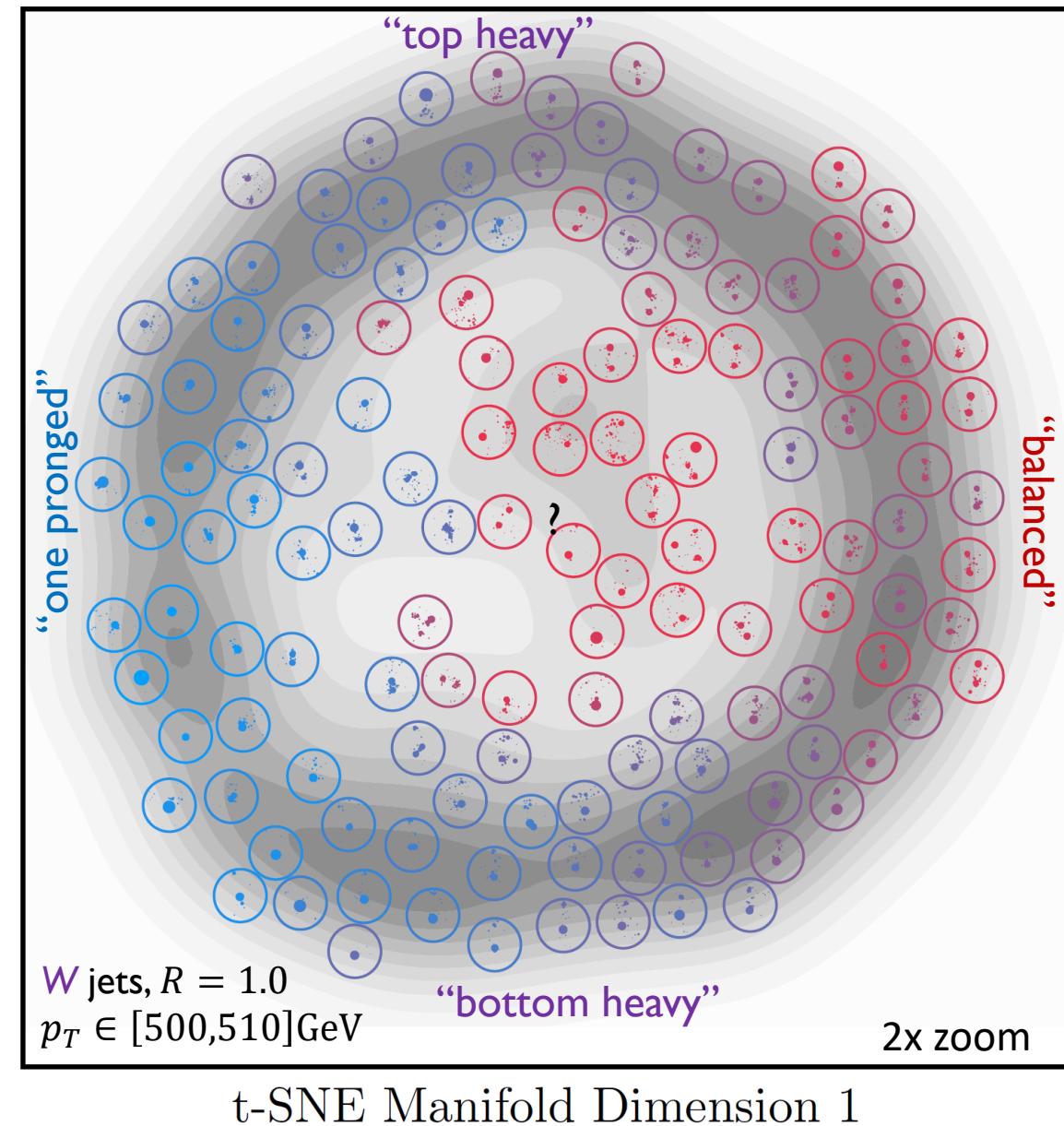
φ : Azimuthal orientation

Constrained by W mass:

$$z(1-z)\theta^2 = \frac{p_{\mu J}^2}{p_T^2} = \frac{m_W^2}{p_T^2}$$

Hence we expect a **two-dimensional** space of W jets.

After φ rotation: **one-dimensional**



Exploring the Space of Jets: Correlation Dimension

VOLUME 50, NUMBER 5

PHYSICAL REVIEW LETTERS

31 JANUARY 1983

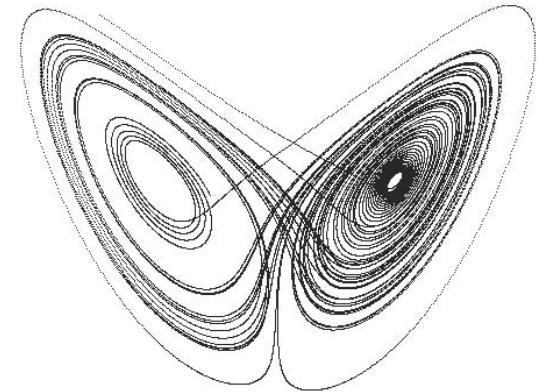
Characterization of Strange Attractors

Peter Grassberger^(a) and Itamar Procaccia

Chemical Physics Department, Weizmann Institute of Science, Rehovot 76100, Israel

(Received 7 September 1982)

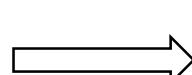
A new measure of strange attractors is introduced which offers a practical algorithm to determine their character from the time series of a single observable. The relation of this new measure to fractal dimension and information-theoretic entropy is discussed.



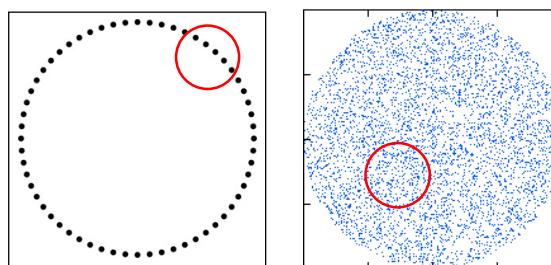
Intuition:

$$N_{\text{neighboring}}(r) \propto r^{\dim}$$

points



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



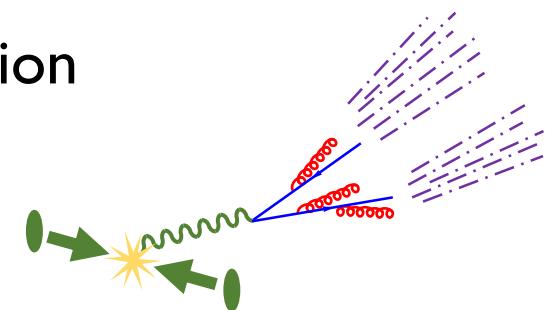
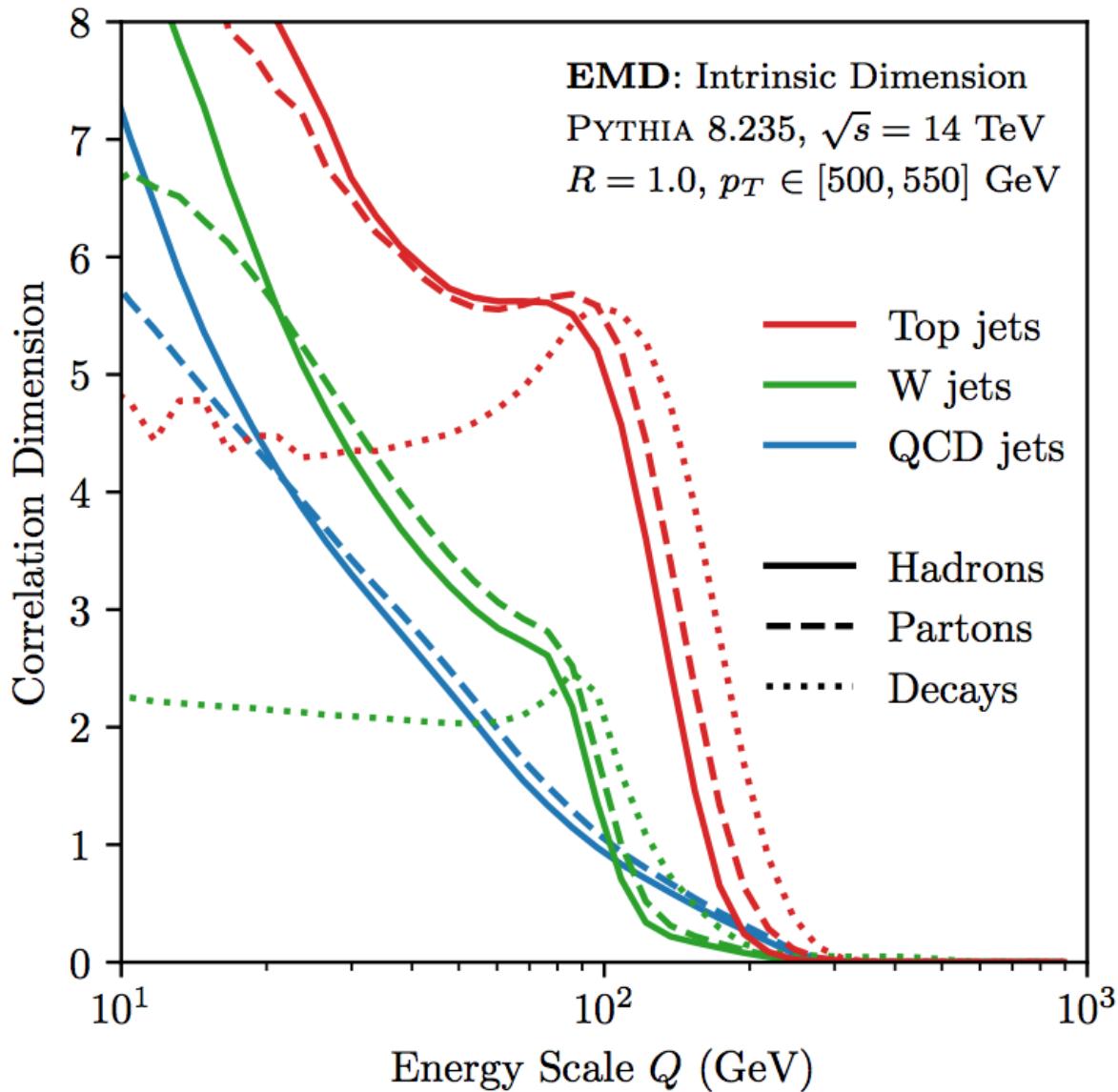
Correlation dimension:

$$\dim(Q) = Q \frac{\partial}{\partial Q} \ln \sum_{i=1}^N \sum_{j=1}^N \Theta[\text{EMD}(\varepsilon_i, \varepsilon_j) < Q]$$

Energy scale Q
dependence

Count neighbors in
ball of radius Q

Exploring the Space of Jets: Correlation Dimension



QCD jets are simplest.

W jets are more complicated.

Top jets are most complex.

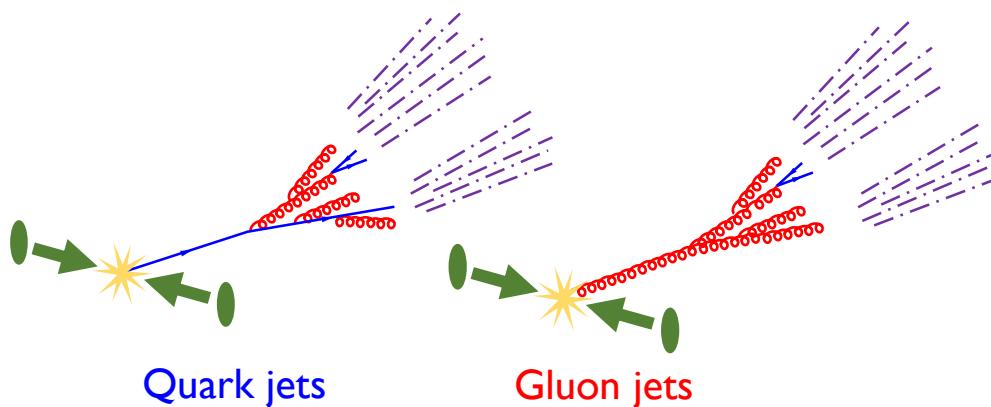
“Decays” have \sim constant dimension.

Fragmentation becomes more complex at lower energy scales.

Hadronization becomes relevant at scales around 20 GeV.

Can we understand this analytically?

Exploring the Space of Jets: Correlation Dimension

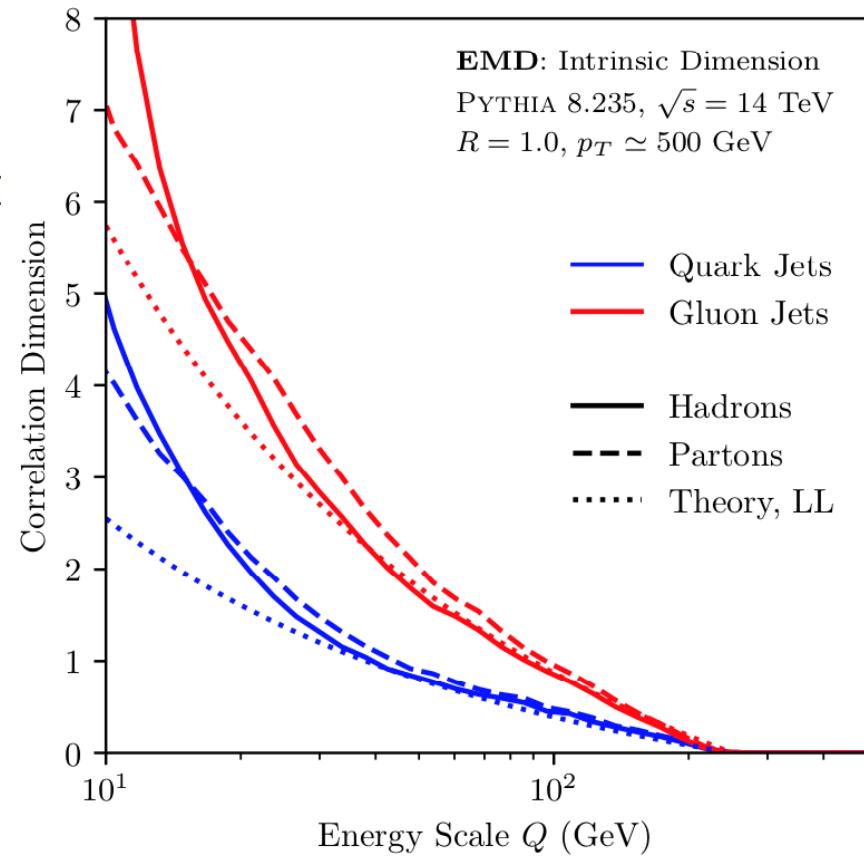


$$\text{At LL: } \dim_{q/g}(Q) = -\frac{8\alpha_s C_{q/g}}{\pi} \ln \frac{Q}{p_T/2}$$

$$+ \text{1-loop running of } \alpha_s$$

$$C_q = C_F = \frac{4}{3}$$

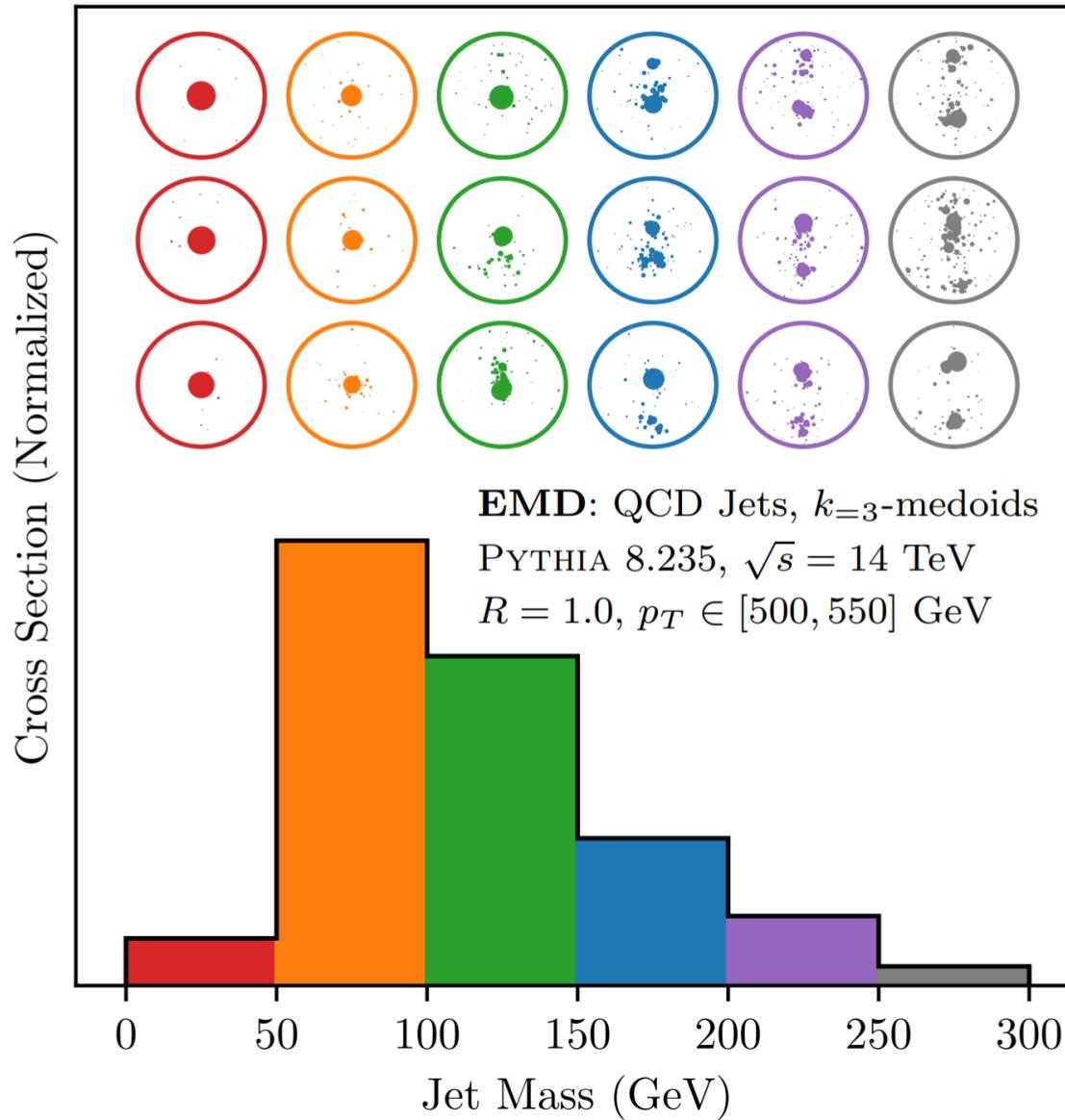
$$C_g = C_A = 3$$



Dimension blows up at low energies.

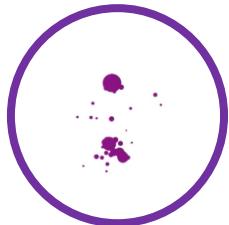
Jets are “more than fractal”?

Exploring the Space of Events: k -medoids



Exploring the Space of Events: Jet Classification

Classify W jets vs. QCD jets



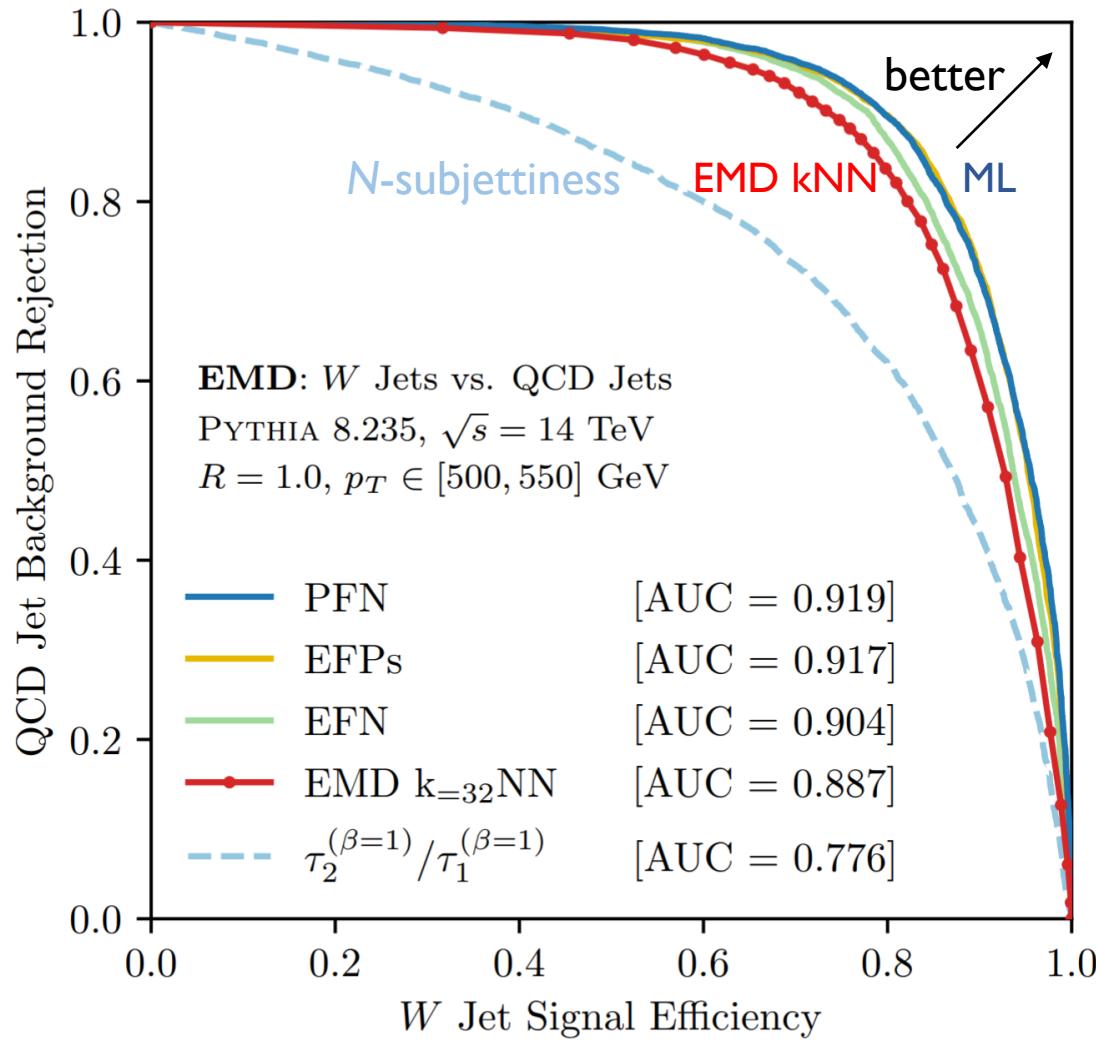
vs.



Look at a jet's nearest neighbors (kNN) to predict its class.

Optimal IRC-safe classifier with enough data.

Nearing performance of ML.



Exploring the Space of Events

Use EMD as a measure of event similarity

Unsupervised clustering algorithms can
be used to cluster events

Jets are clusters of particles
???? are clusters of jets

VP Tree: $O(\log(N))$ neighbor query time

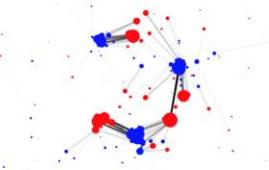
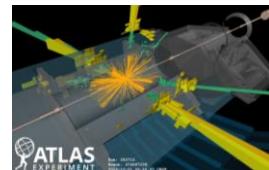
Much more to explore.

Vantage Point (VP) Tree

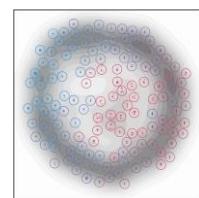
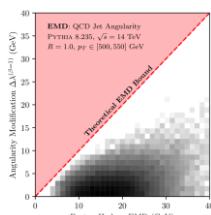
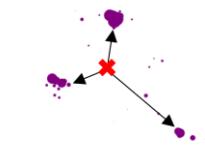


Summary

Part I Introduction



Part II Applications



When are two events similar?

When they have similar distributions of energy

The Energy Mover's Distance

Work to rearrange one event into another.

Movie Time

Visualize energy movement and jet formation.

Observables

Conceptually rich connections to EMD.

Quantifying event modifications

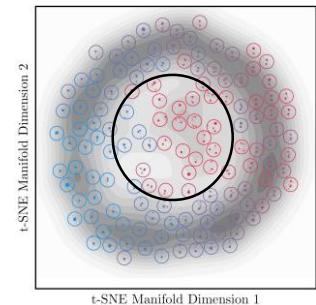
Hadronization, pileup, detector effects

Exploring the Space of Events

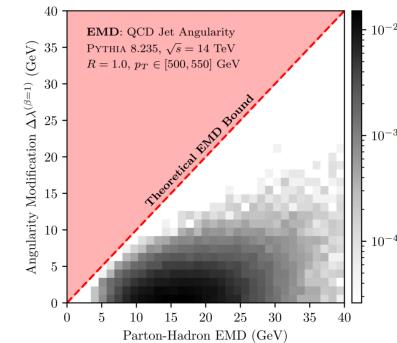
Unlock new ideas and techniques with EMD

Going Beyond

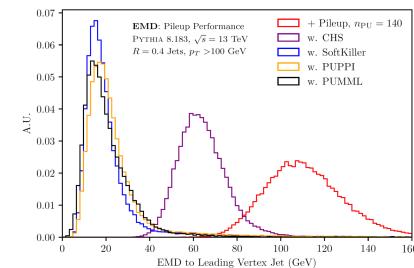
- Model (in)dependent anomaly detection?



- Sharpen the parton-hadron duality of energy flow?



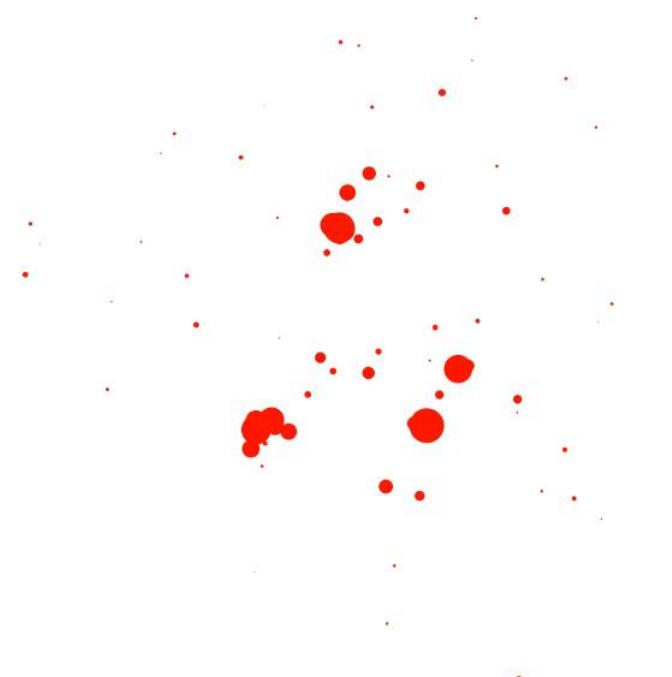
- Train ML models to optimize EMD directly?



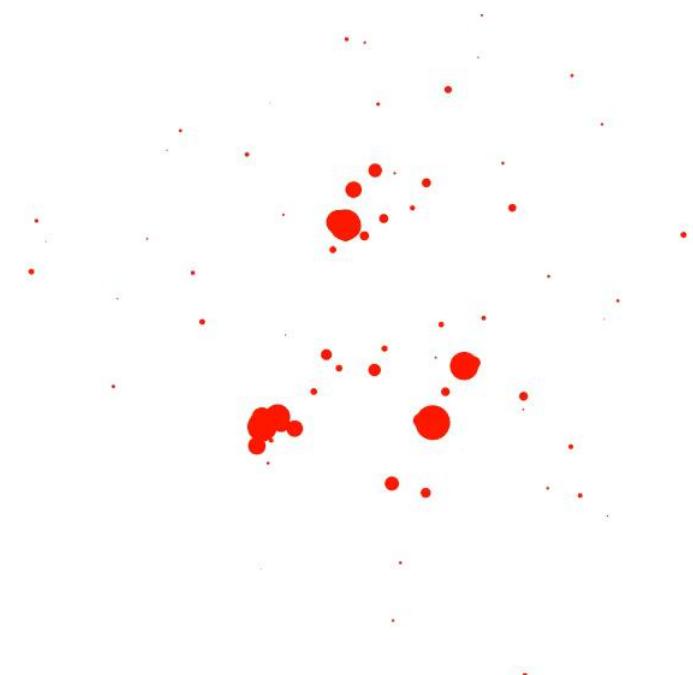
- Include flavor information?

The End

Thank you!



Extra Slides



Exploring the Space of Jets: Correlation Dimension

Sketch of leading log (one emission) calculation:

$$\dim_{q/g}(Q) = Q \frac{\partial}{\partial Q} \ln \sum_{i=1}^N \sum_{j=1}^N \Theta[\text{EMD}(\varepsilon_i, \varepsilon_j) < Q]$$

$$= Q \frac{\partial}{\partial Q} \ln \Pr [\text{EMD} < Q]$$

$$= Q \frac{\partial}{\partial Q} \ln \Pr [\lambda^{(\beta=1)} < Q; C_{q/g} \rightarrow 2 C_{q/g}]$$

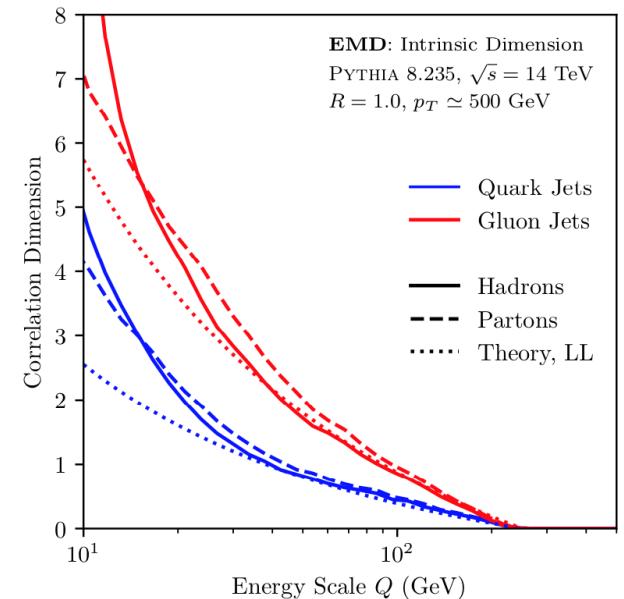
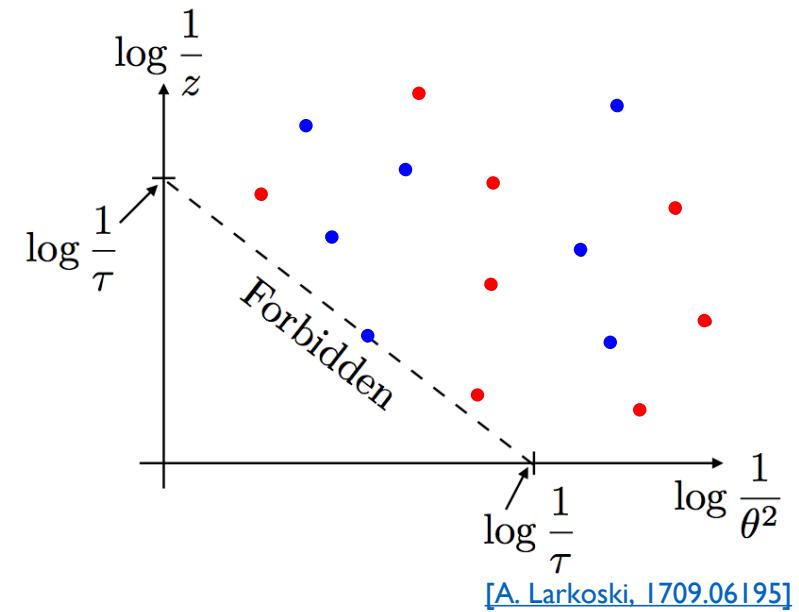
$$= Q \frac{\partial}{\partial Q} \ln \exp \left(- \frac{4\alpha_s C_{q/g}}{\pi} \ln^2 \frac{Q}{p_T/2} \right)$$

$$= - \frac{8\alpha_s C_{q/g}}{\pi} \ln \frac{Q}{p_T/2}$$

+ 1-loop running of α_s

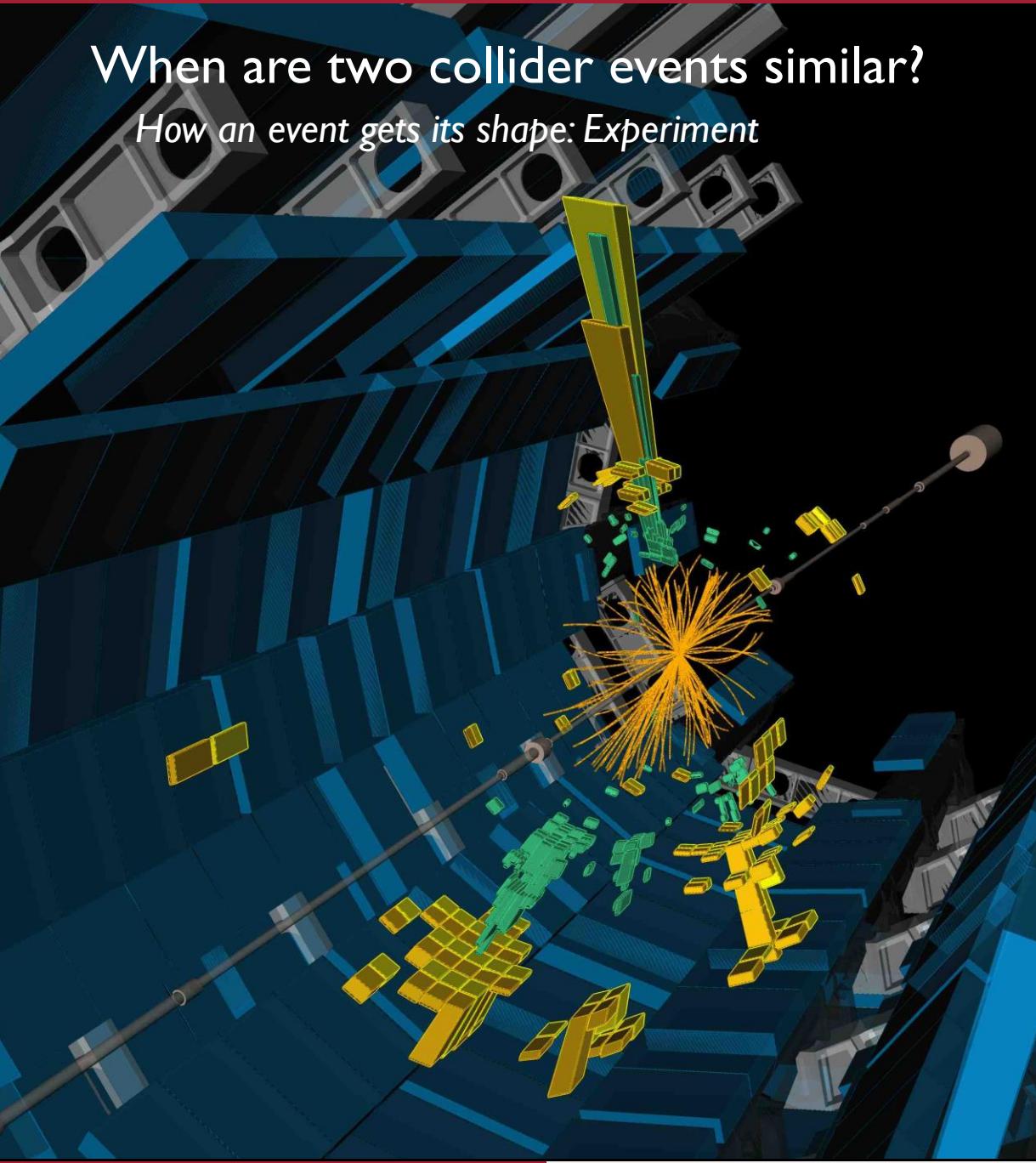
$$C_q = C_F = \frac{4}{3}$$

$$C_g = C_A = 3$$



When are two collider events similar?

How an event gets its shape: Experiment



tracker	ECAL	HCal	
			γ photon
			e^\pm electron
			μ^\pm muon
			π^\pm pion
			K^\pm kaon
			K_L^0 K-long
			p/\bar{p} proton
			n/\bar{n} neutron

Pileup Mitigation with PUMML

