Some applications of optimal transport in LHC physics analysis

2021/05/05

M. LeBlanc (CERN)

Overview

- Several recent phenomenological notes which propose applications of optimal transport techniques (“Earth/Energy Mover’s Distance” / “Wasserstein-2 Metric”) in collider physics analysis:
  - Proposal, QCD fractal correlation dimension, jet tagging: https://arxiv.org/abs/1902.02346 P. Komiske, E. Metodiev, J. Thaler
  - Pileup mitigation, jet tagging, other applications: https://arxiv.org/abs/2004.04159 P. Komiske, E. Metodiev, J. Thaler
  - QCD event shapes in pp and e+e- collisions: https://arxiv.org/abs/2004.06125 C. Cesarotti, J. Thaler
- I have been trying out these techniques within ATLAS …

Energy-Mover's Distance

The "work" required to rearrange one collision event into another.

Plus a cost to create or destroy energy.

Difference in radiation pattern

Difference in total energy

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

β : angular weighting factor

R : tradeoff between moving energy and creating it

Infrared and collinear safe notion of distance!

Deeply related to the event "energy flow"

$$\mathcal{E}(\hat{n}) = \lim_{r\to\infty} r^2 \int_0^{\infty} dt\, \hat{n}_i T^{0i}(t, r\hat{n})$$

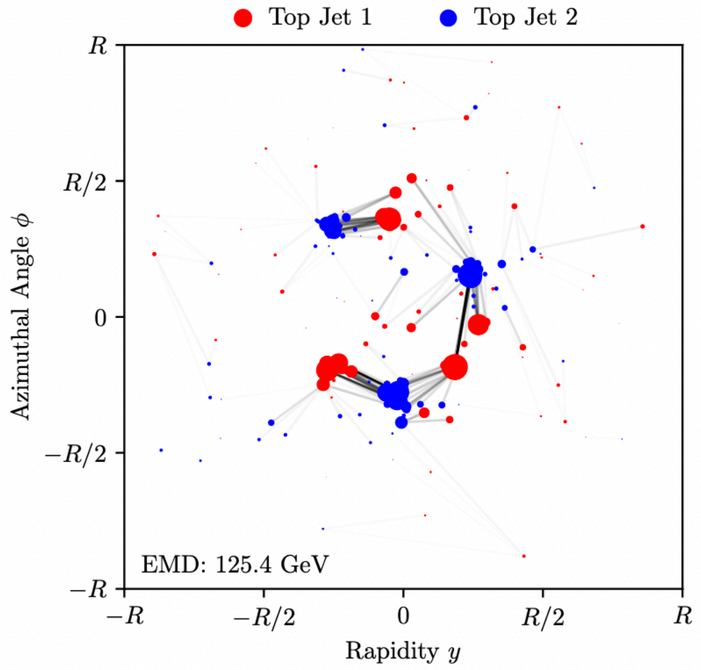
[Sveshnikov, Tkachov, PLB, 9512370]
[Tkachov, IJMP, 9601308]

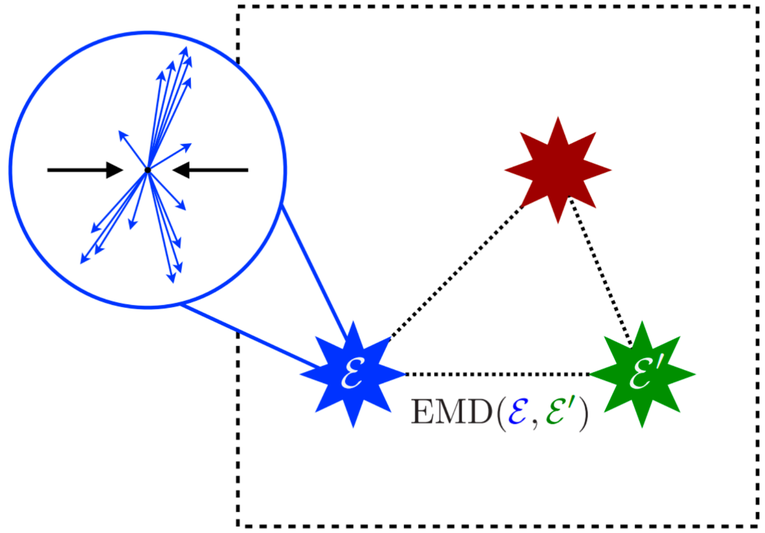
Based on the Earth Mover's or Wasserstein Distance

[Peleg, Werman, Rom, PAMI,1989]
[Rubner, Tomasi, Guibas, IJCV, 2000]

Optimal Transport Problem

python optimal transport library





Slide adapted from E. Metodiev (MIT)

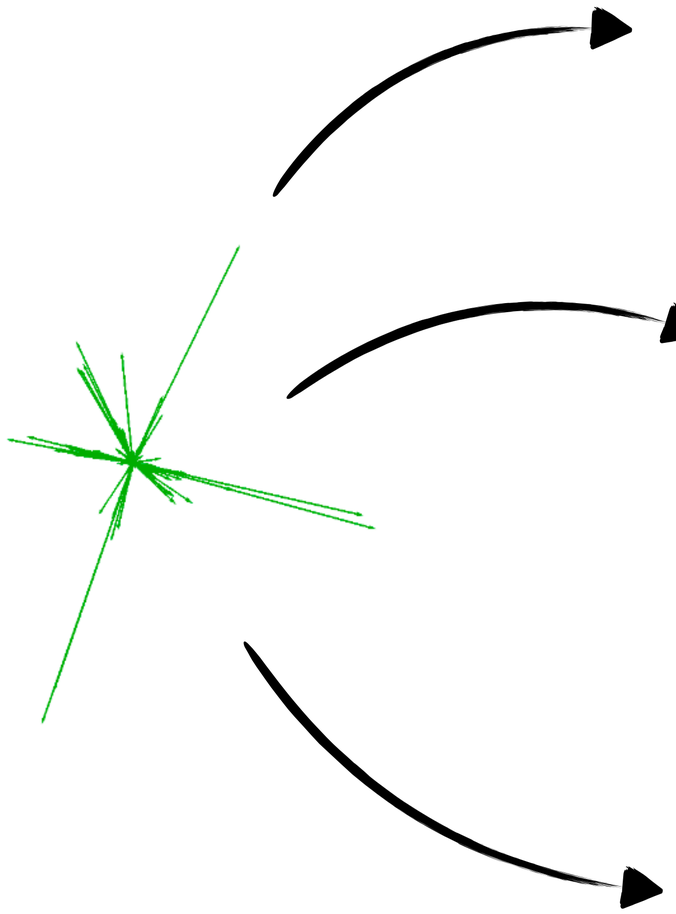
Event isotropy

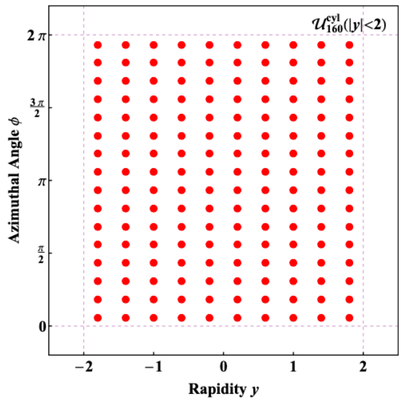
- Defined as the dimensionless distance between a collider event E and a uniform radiation pattern U of the same energy:

  I(E) = EMD(U,E)

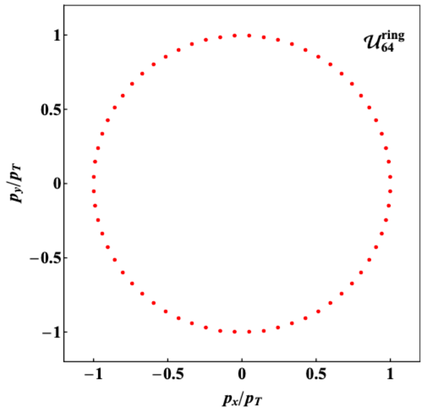
- I(E) exhibits a larger dynamic range for high-multiplicity events than traditional event shapes like thrust, sphericity, etc.

- *Just one calculation per-event!*

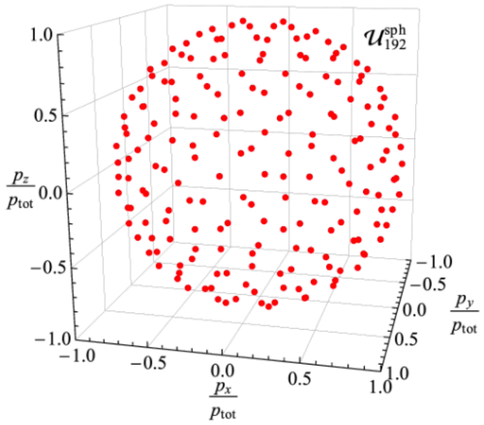




Cylindrical (pp)



Circular (pp)



Spherical (e+e-)

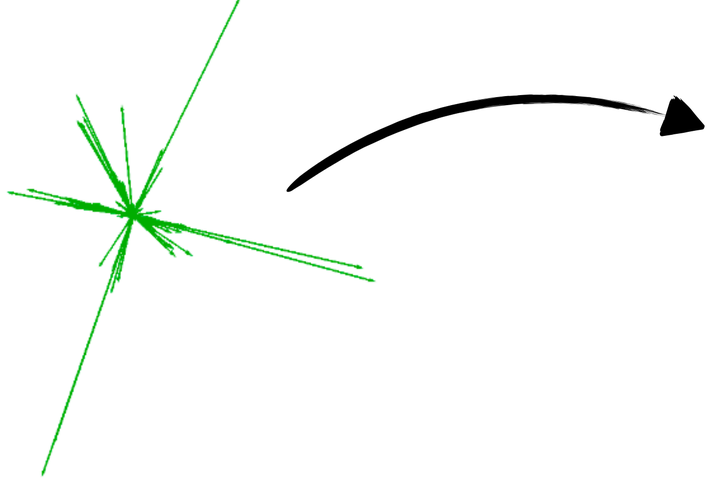
Event isotropy

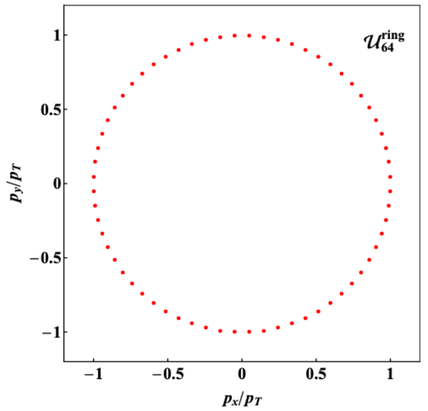
- Defined as the dimensionless distance between a collider event E and a uniform radiation pattern U of the same energy:

  $I(E) = EMD(U,E)$

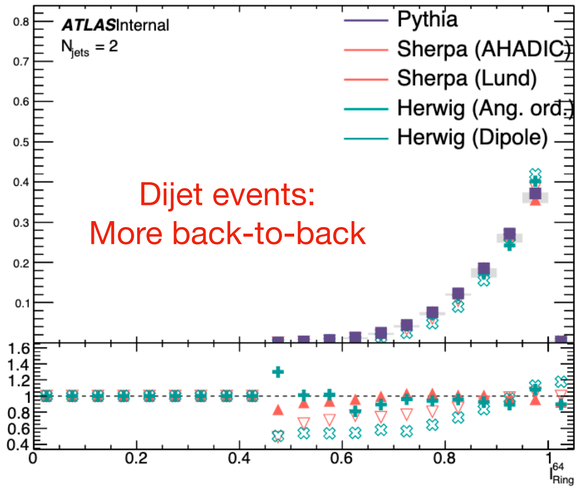
- I(E) exhibits a larger dynamic range for high-multiplicity events than traditional event shapes like thrust, sphericity, etc.

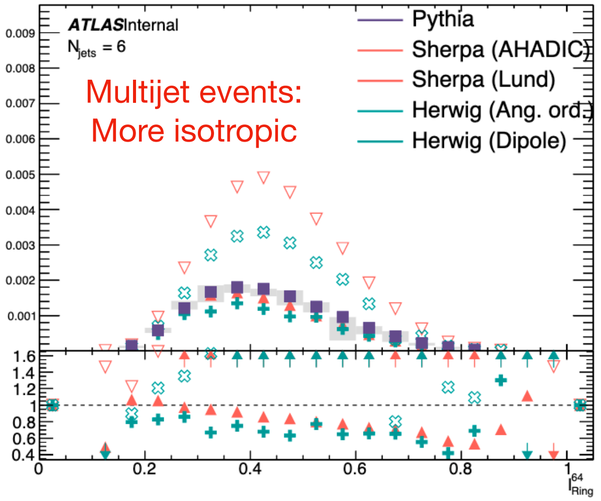
- *Just one calculation per-event!*





Circular (pp)



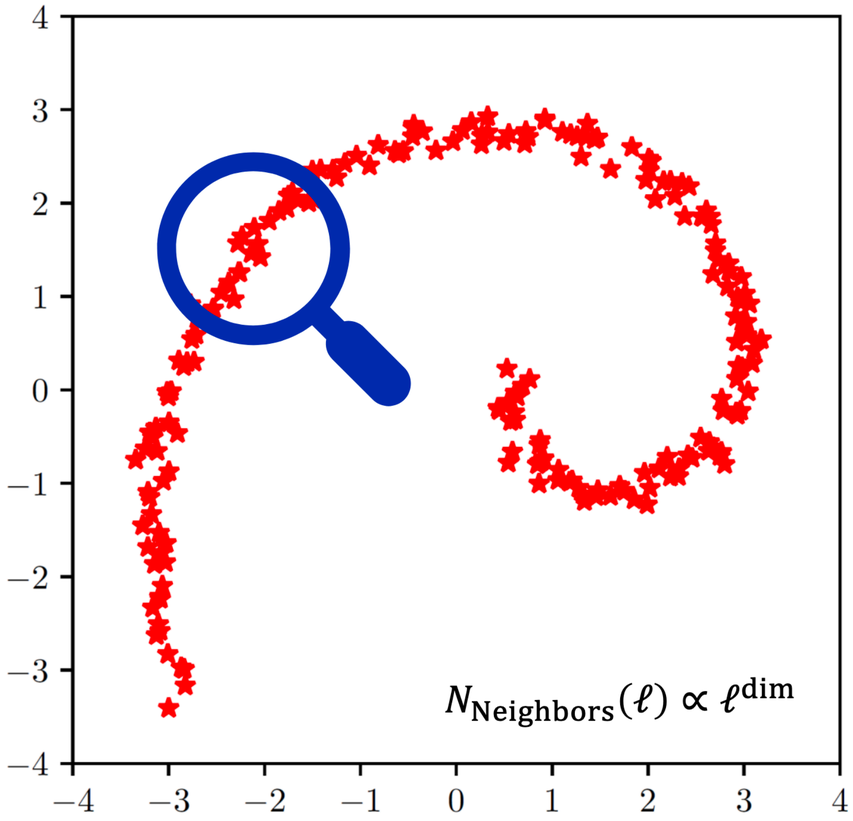


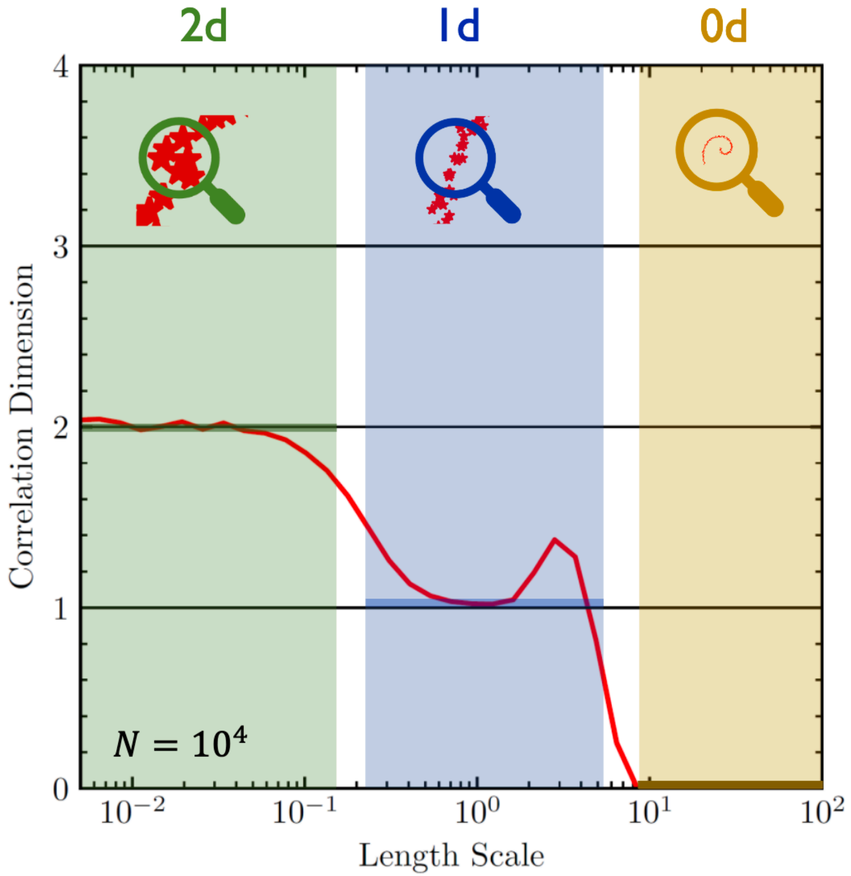
* *n.b.* very preliminary plots — some problems in Sherpa / Herwig normalization

Fractal Correlation Dimension

$$\dim(\ell) = \ell \frac{\partial}{\partial \ell} \ln \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta[\mathrm{d}(x_i, x_j) < \ell]$$

[Grassberger, Procaccia, PRL, 1983] [Kegl, NeurIPS, 2002]

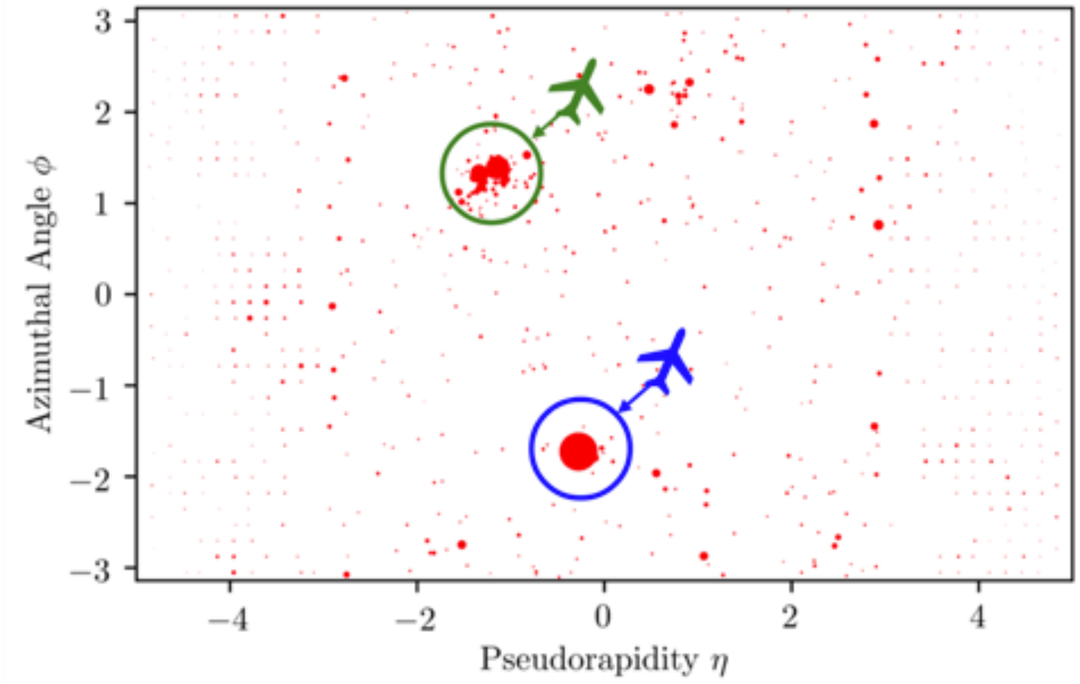


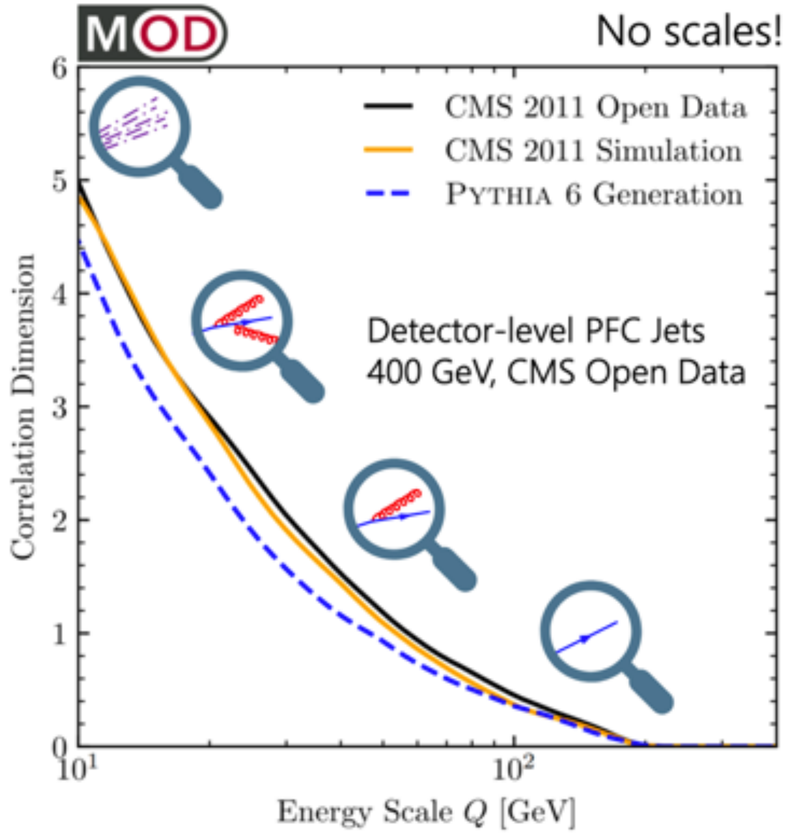


A spectrum of the dataset at a glance.

Fractal Correlation Dimension

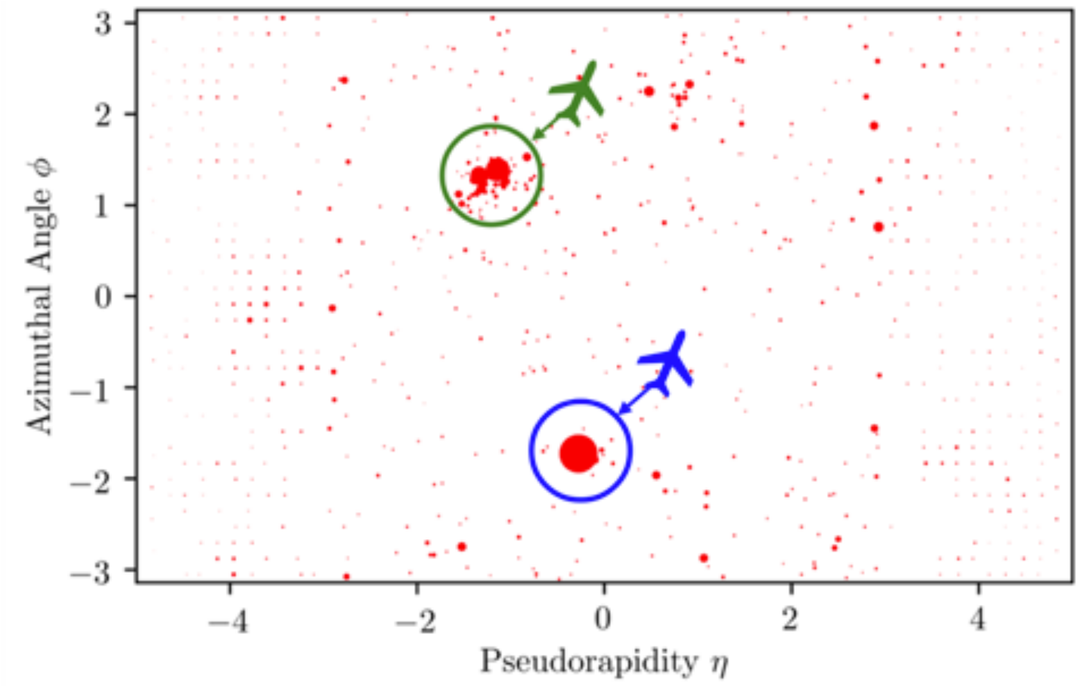
$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln\sum_{i=1}^{N}\sum_{j=1}^{N}\Theta[\mathrm{EMD}(\mathcal{E}_i,\mathcal{E}_j) < Q]$$

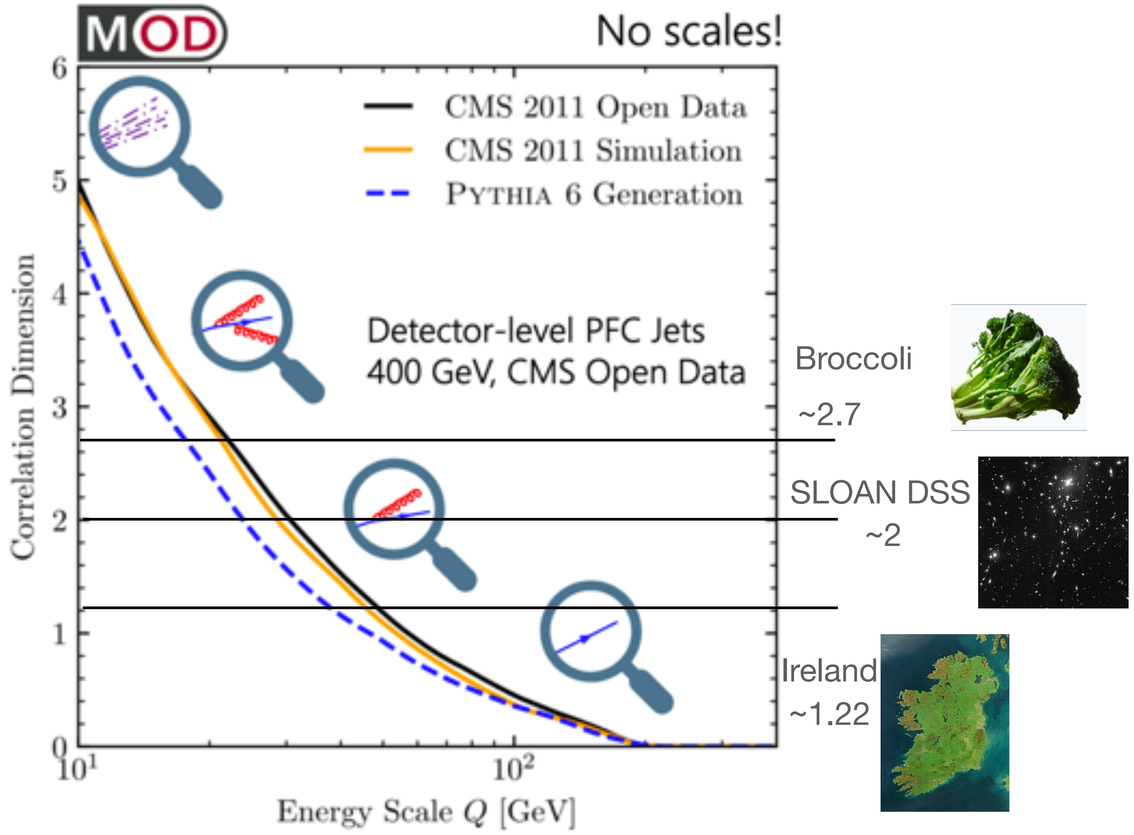




Fractal Correlation Dimension

$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln\sum_{i=1}^{N}\sum_{j=1}^{N}\Theta[\mathrm{EMD}(\mathcal{E}_i,\mathcal{E}_j) < Q]$$





Summary

- Activities within ATLAS mainly within the SM Jet & Photon Physics subgroup:
- Isotropy event shape measurement using jets was kicked off and exists on glance.
  - Natural extension to future track-based event shape measurement, but higher particle multiplicity means longer time to process each event!
  - Already struggling … time scales with particle multiplicity as $N^3 \log^2(N)$.
- Fractal correlation dimension study part of Omnifold analysis, looking at hadronic recoil in Z+jets events.
  - You can probably imagine how long it takes to calculate something defined in all pairs of events in our Z+jets samples (a lot of accounting).

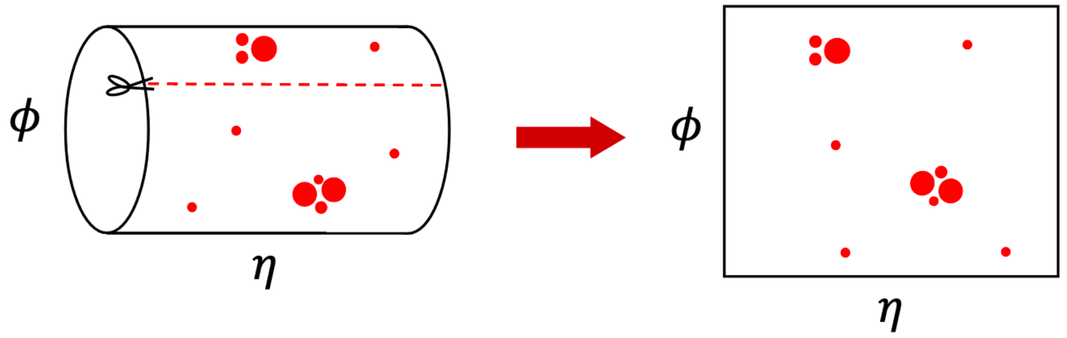
Technical remarks

- Both studies rely on existing analysis inputs:
  - Event shape inputs ~ 5 TB (multijets), on /eos at /eos/atlas/atlascerngroupdisk/phys-sm/JetPhoton/R32v18/
  - Omnifold inputs ~ 16 GB (Z+jets, but pair-wise calculation), on /eos at /eos/home-l/lmiller/ZjetDataFiles/GridRunFeb20/slimmedSamples/
- I require a stable system with quick I/O and many CPUs in order to speed up these workflows, as they are already set up to be multi-threaded.
  - Large amounts of RAM are not so critical (I think).
  - I have become a bit worried that these studies will be rather expensive to run on the cloud, due to the type of VM I would be interested in : it's possible that there are more cost-effective options to try out first.

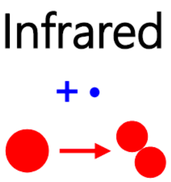
Backup slides (from E. Metodiev)
Taken from
https://indico.cern.ch/event/906711/contributions/3851596/attachments/2031795/3424378/ATLAS2020_Metodiev.pdf

See also J. Thaler's CERN-TH colloquium on this topic:
https://indico.cern.ch/event/888504/

When are two collisions similar?

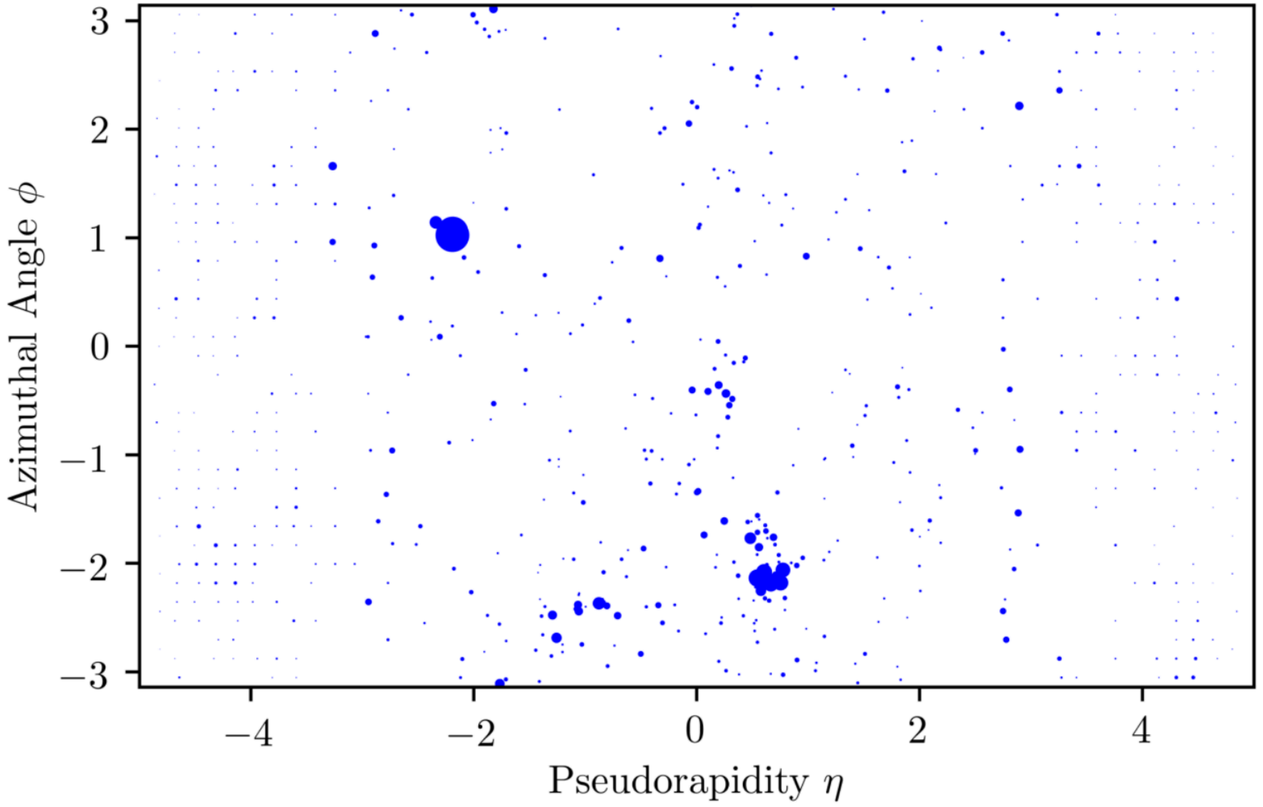


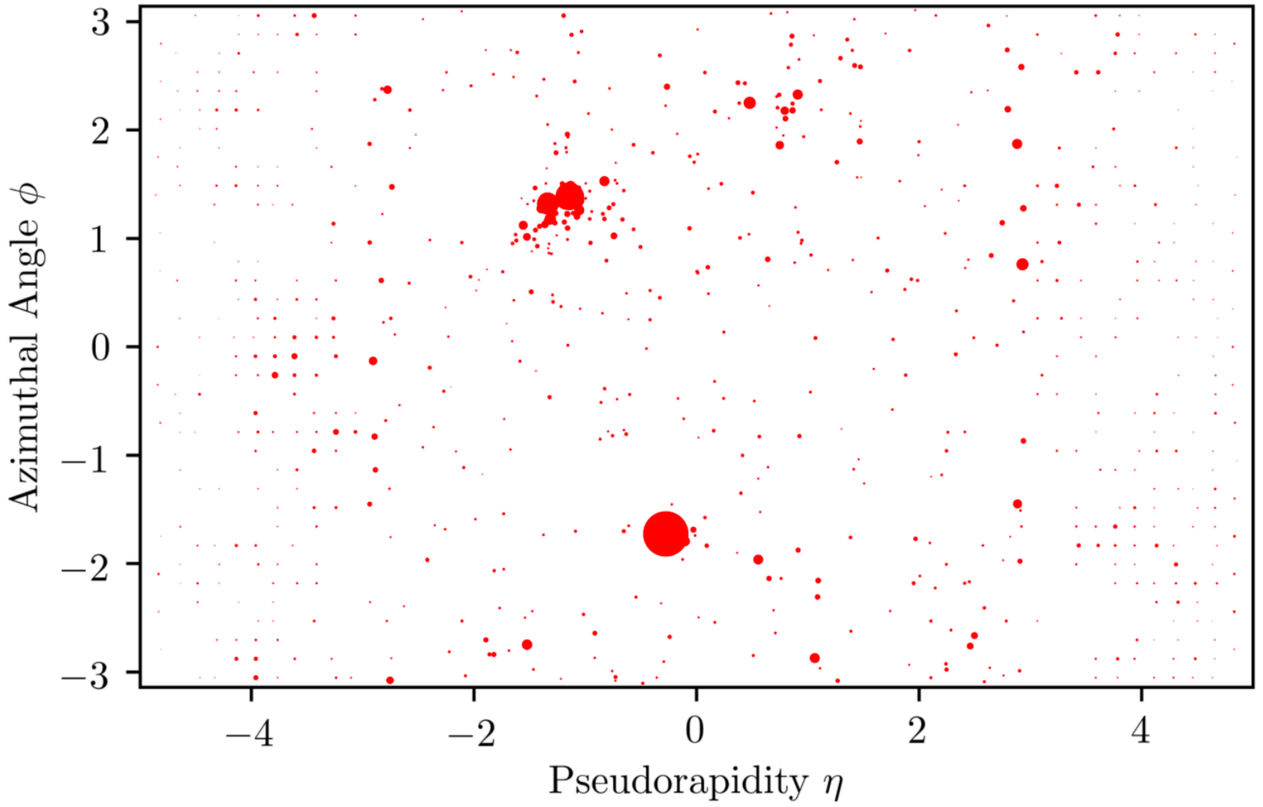
Infrared and Collinear Safety says distance must be invariant under:



Addition of zero-energy particles

Collinear splitting of one particle into two



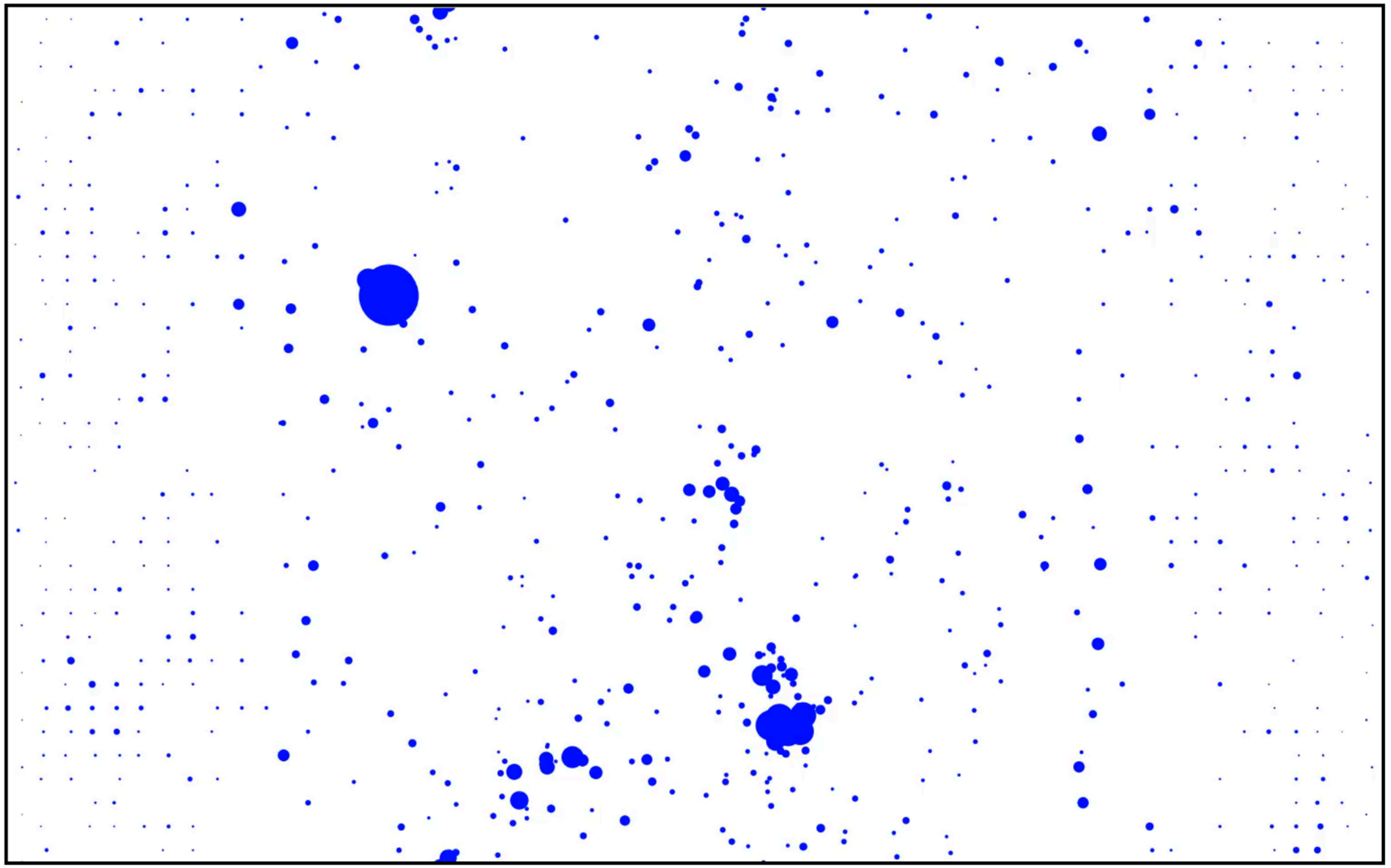


Dijet events from 2011 CMS Open Data – Particle Flow Candidates.

When are two collisions similar?

The Energy Mover's Distance (EMD)

[Komiske, EMM, Thaler, PRL, 1902.02346]



The "work" required to rearrange one collision event into another.

Plus a cost to create or destroy energy.

Infrared and collinear safe notion of distance!

Deeply related to the event "energy flow"

$$\mathcal{E}(\hat{n}) = \lim_{r\to\infty} r^2 \int_0^\infty dt\, \hat{n}_i T^{0i}(t, r\hat{n})$$

[Sveshnikov, Tkachov, PLB, 9512370]
[Tkachov, IJMP, 9601308]

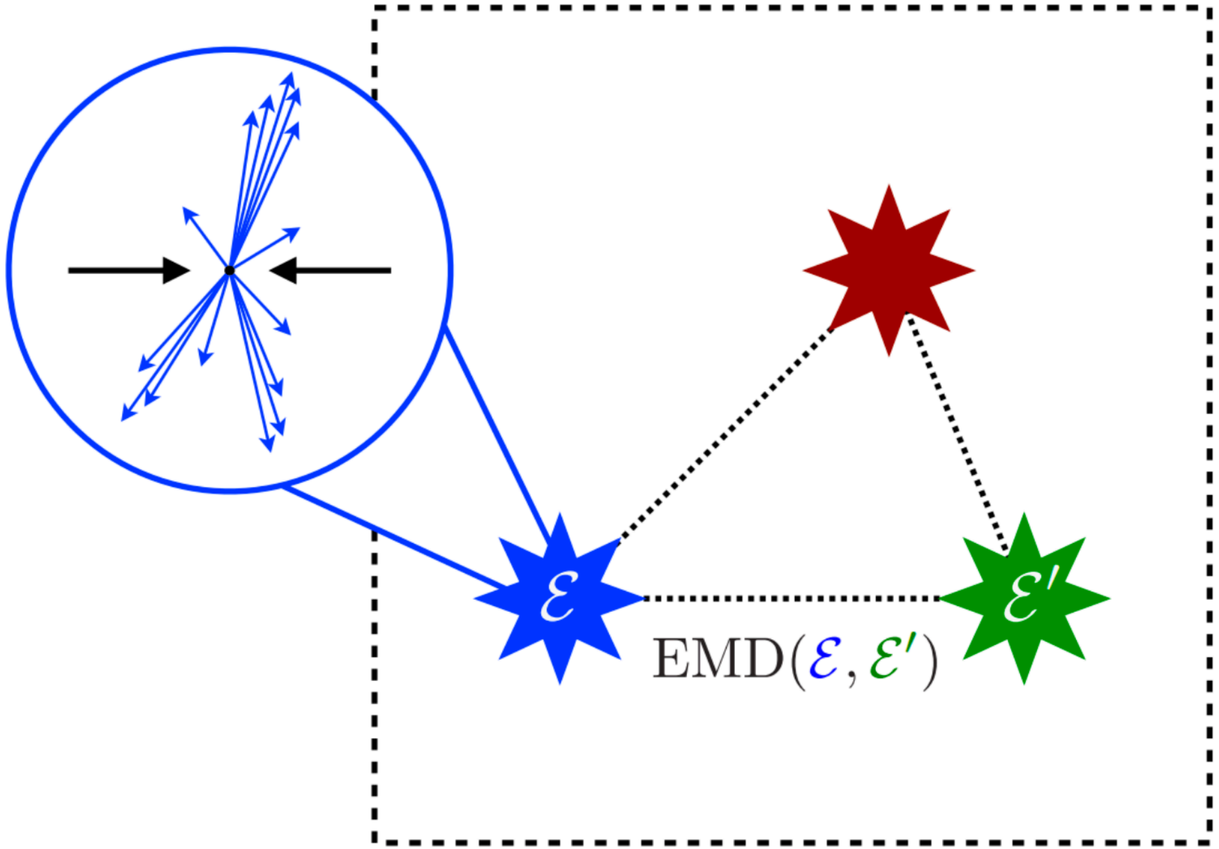
Based on the Earth Mover's or Wasserstein Distance

[Peleg, Werman, Rom, PAMI,1989]
[Rubner, Tomasi, Guibas, IJCV, 2000]

Optimal Transport Problem

python optimal transport library

The Space of Collider Events

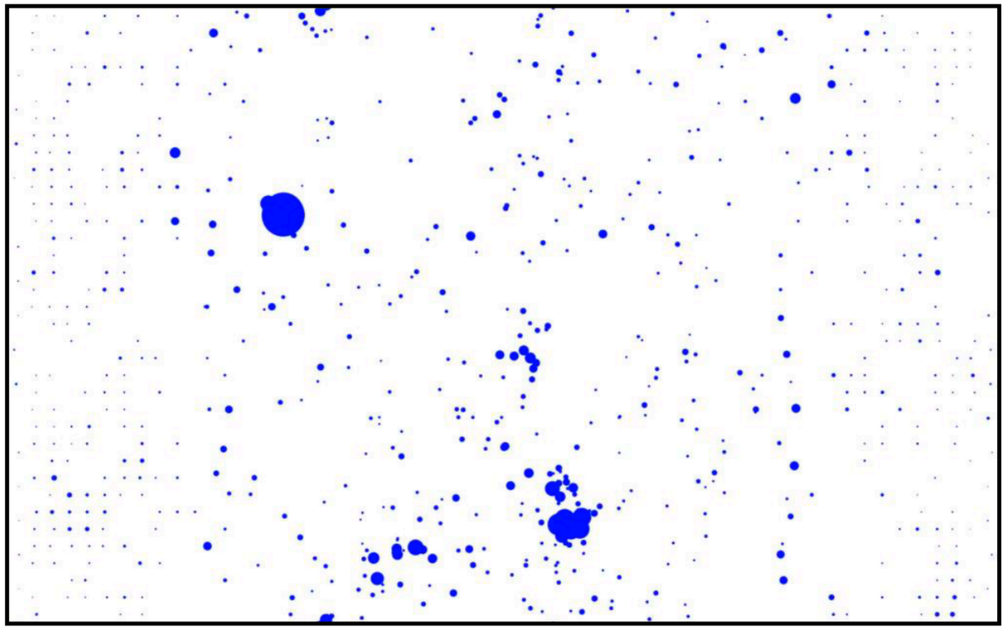


Difference in radiation pattern

Difference in total energy

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

β : angular weighting factor
R : tradeoff between moving energy and creating it



Enabling New Directions: The Fractal Dimension of QCD

A new probe of the fractal nature of QCD.

Goes beyond an observable, $\mathcal{O}(\varepsilon)$

"How much information is in a jet?"

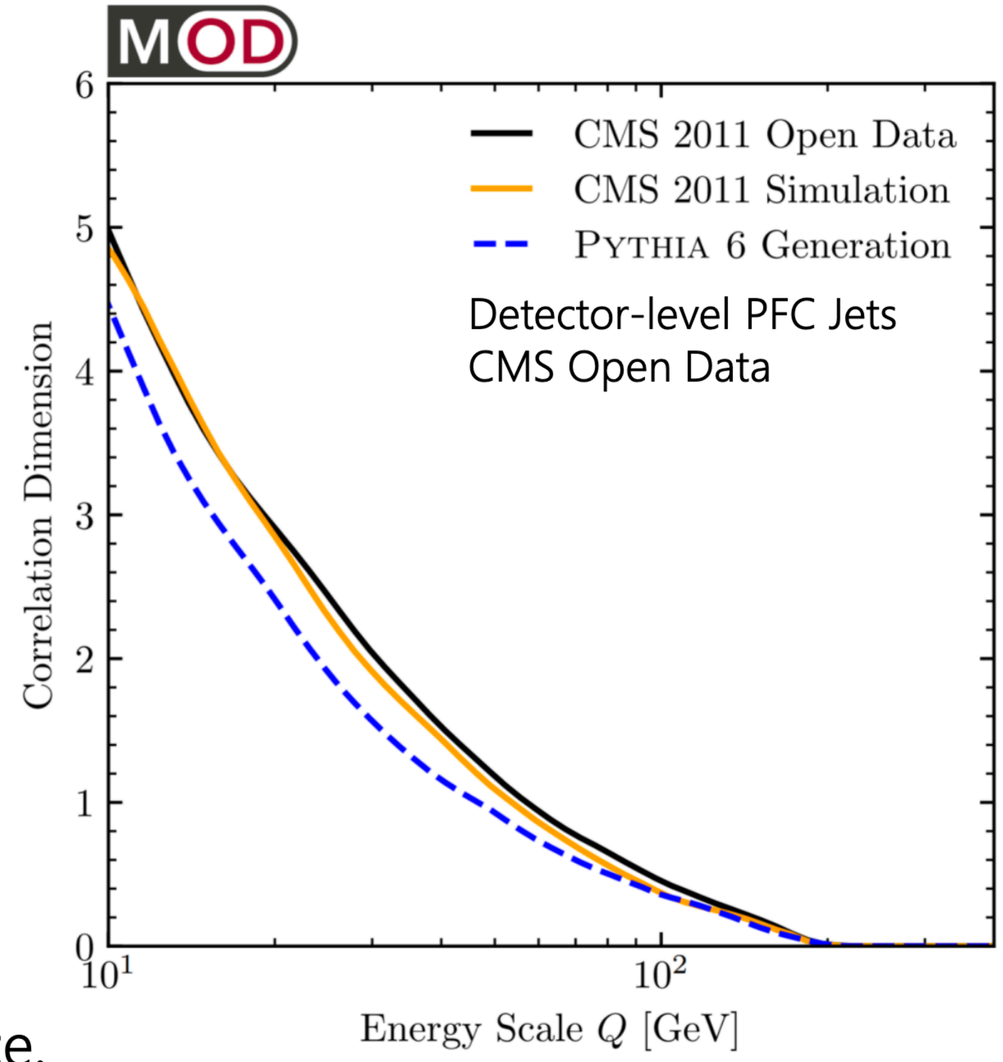
[Datta, Larkoski, JHEP, 1704.08249]

"How many particles do I resolve at this energy scale?"

[Larkoski, EMM, JHEP, 1906.01639]

P.S. Related to the event-event correlators of Theory Space.

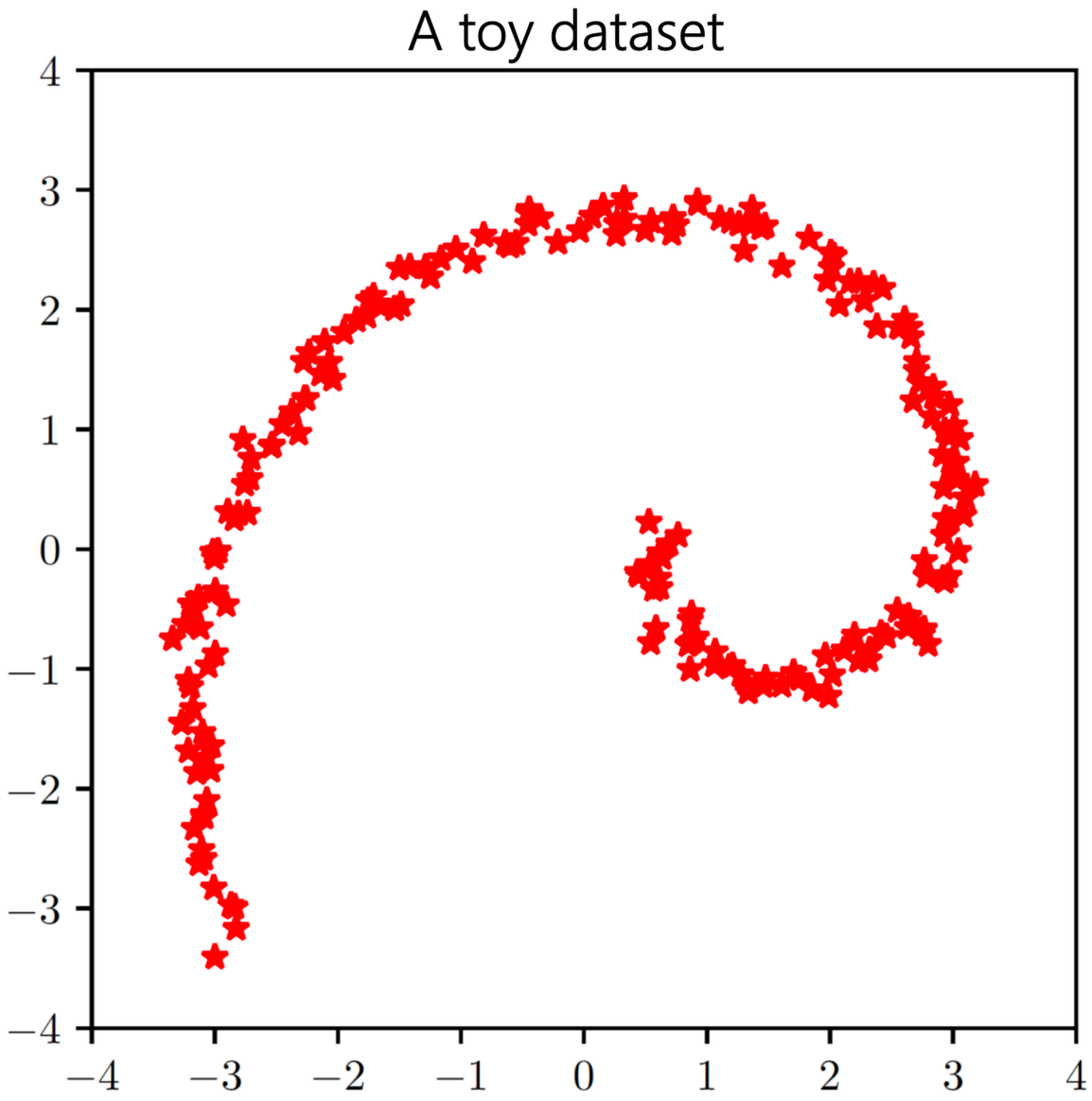
[Komiske, EMM, Thaler, 2004.04159]



[Komiske, EMM, Thaler, PRL, 1902.02346]

[Komiske, Mastandrea, EMM, Naik, Thaler, PRD, 1908.08542]

Enabling New Directions: The Fractal Dimension of QCD



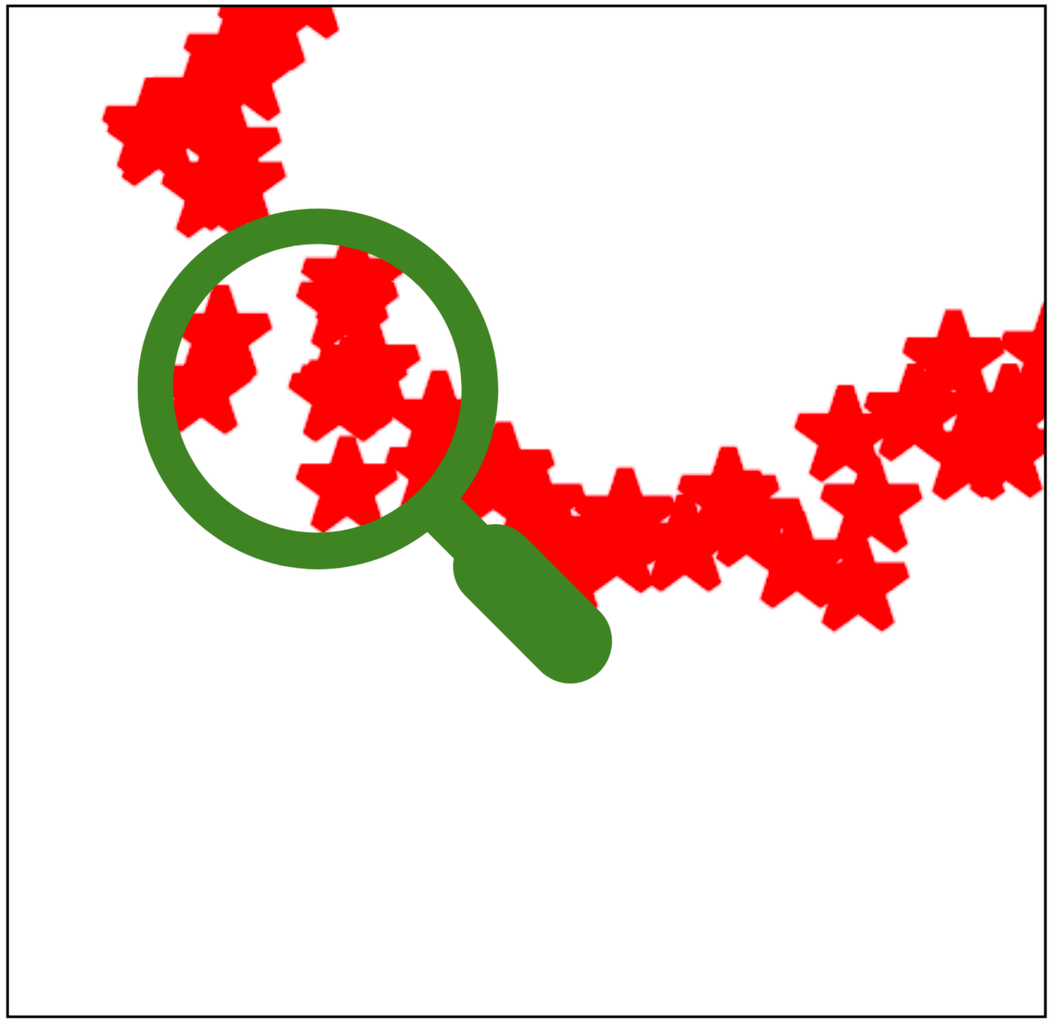
Questions

What are the scales in the system?

What is its dimensionality or complexity?

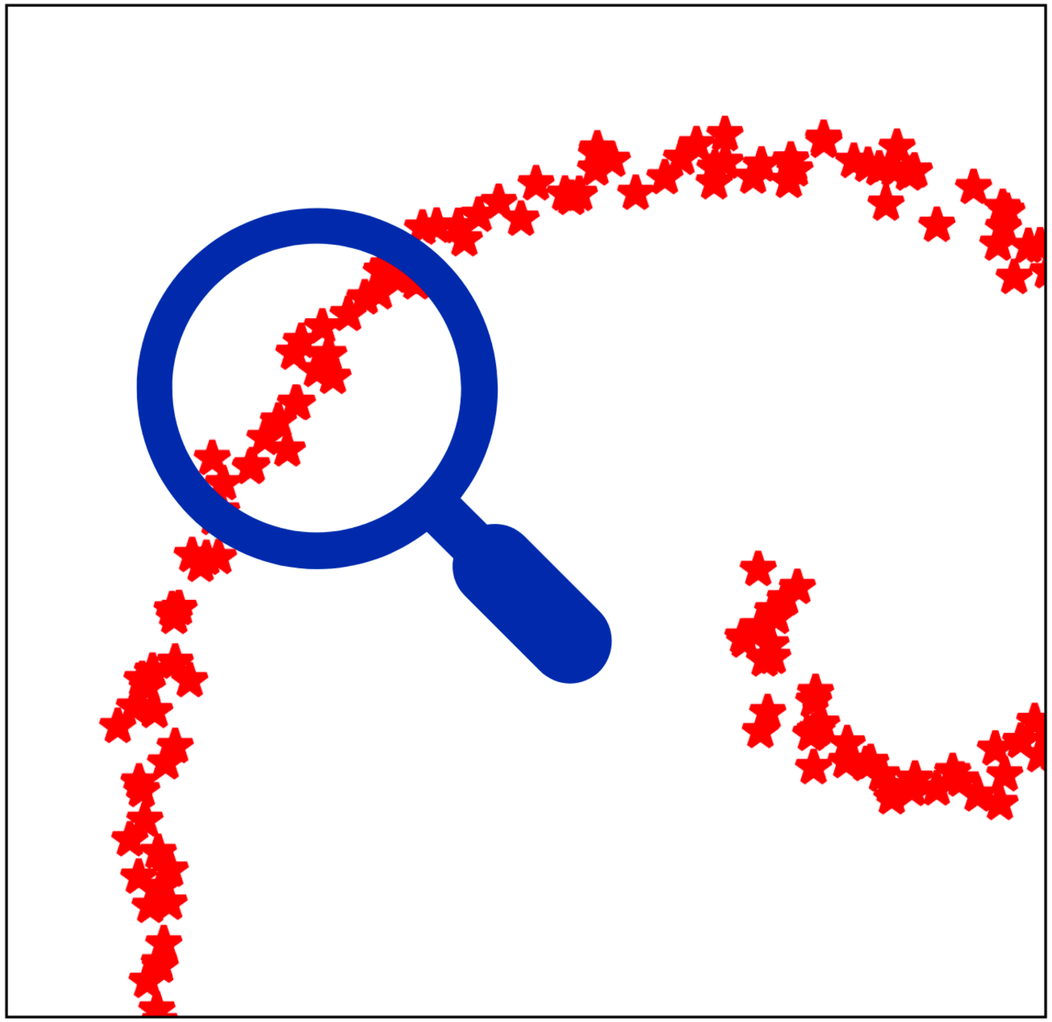
How do I characterize it?

Enabling New Directions: The Fractal Dimension of QCD



Small scales: Two-dimensional plane

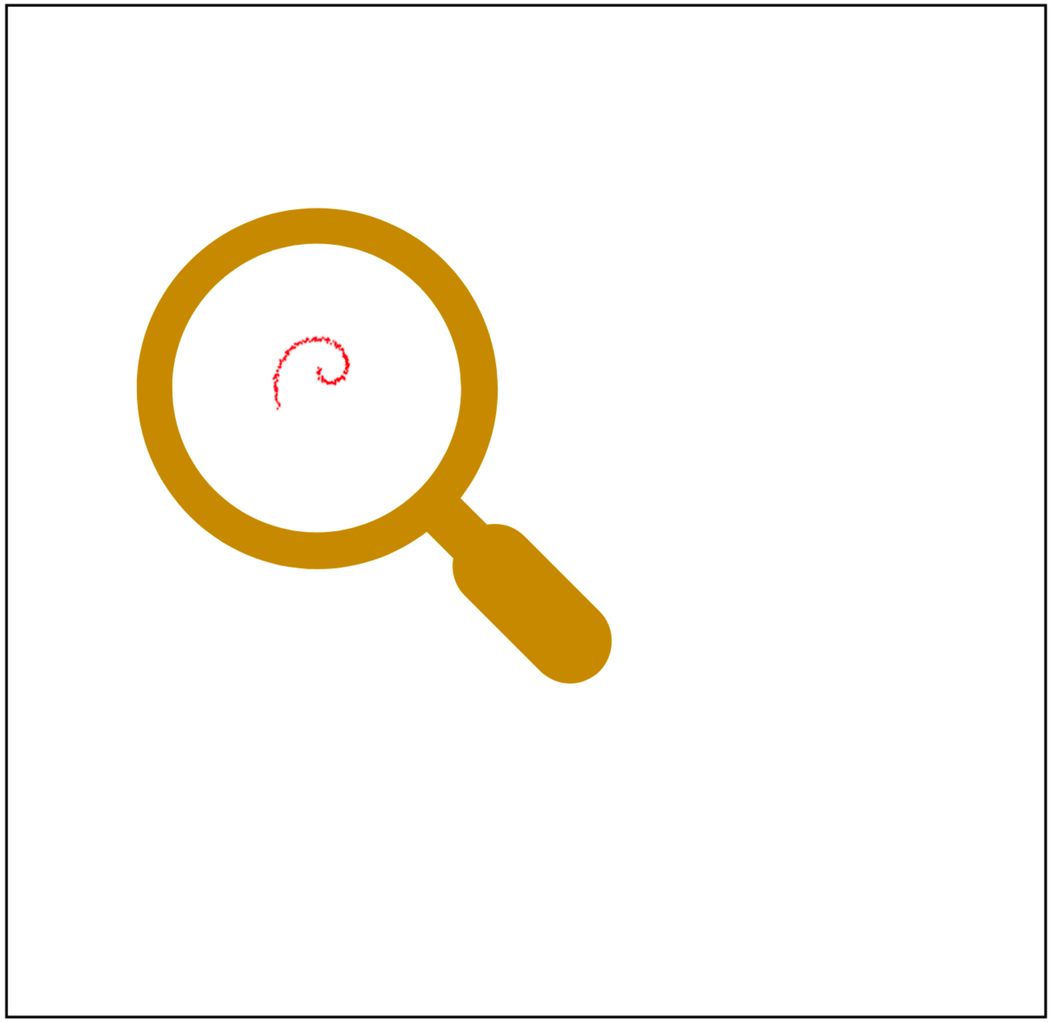
Enabling New Directions: The Fractal Dimension of QCD



Small scales: Two-dimensional plane

Medium scales: One-dimensional line

Enabling New Directions: The Fractal Dimension of QCD



Small scales: Two-dimensional plane

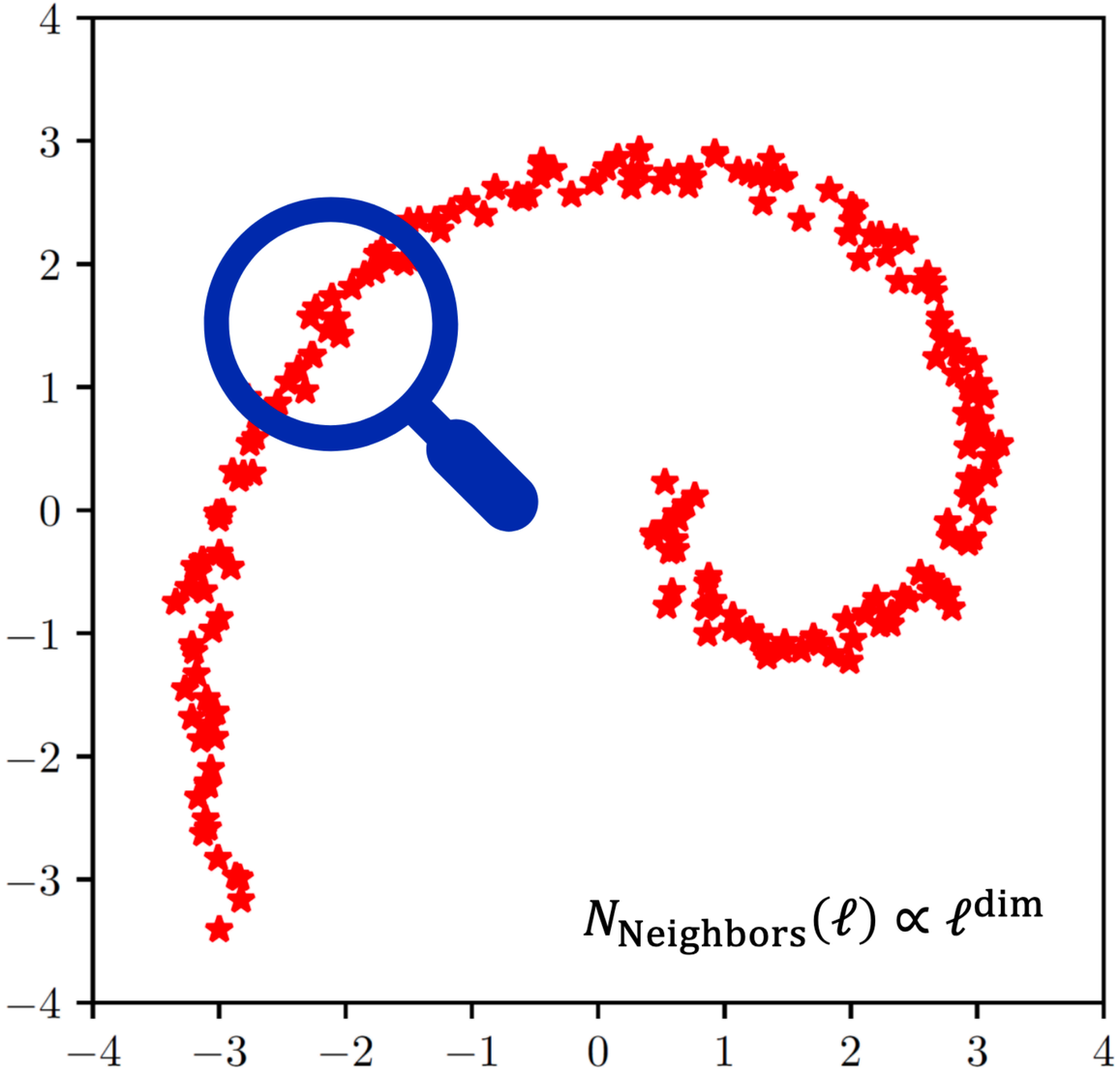
Medium scales: One-dimensional line

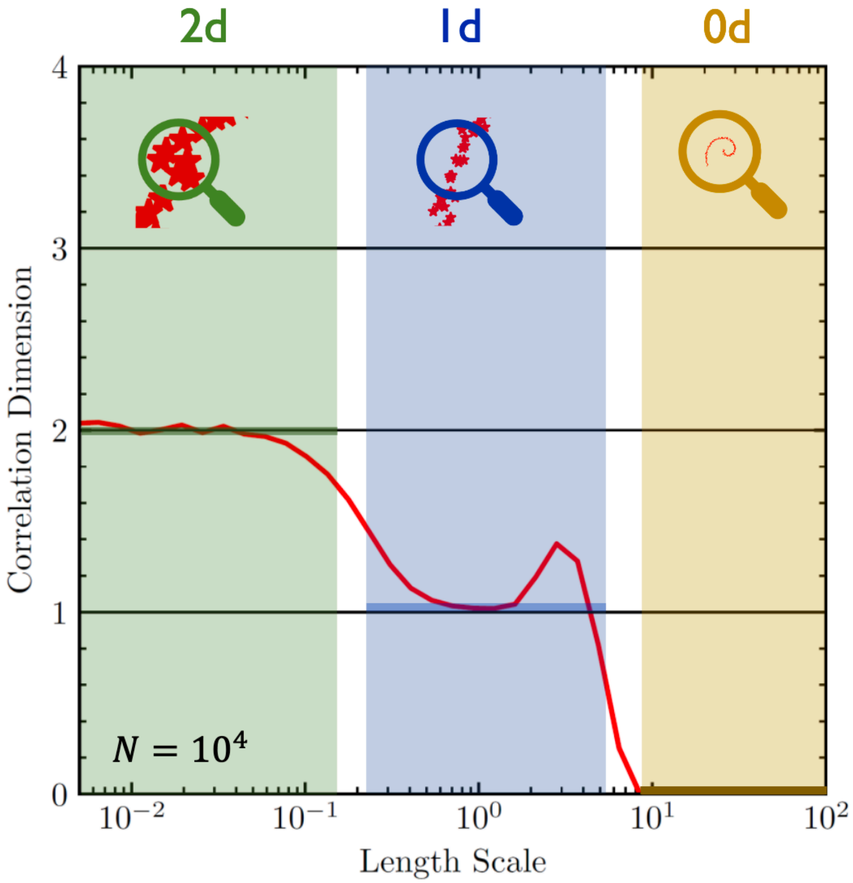
Large scales: Zero-dimensional point

Enabling New Directions

$$\dim(\ell) = \ell \frac{\partial}{\partial \ell} \ln \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta[\mathrm{d}(x_i, x_j) < \ell]$$

[Grassberger, Procaccia, PRL, 1983] [Kegl, NeurIPS, 2002]





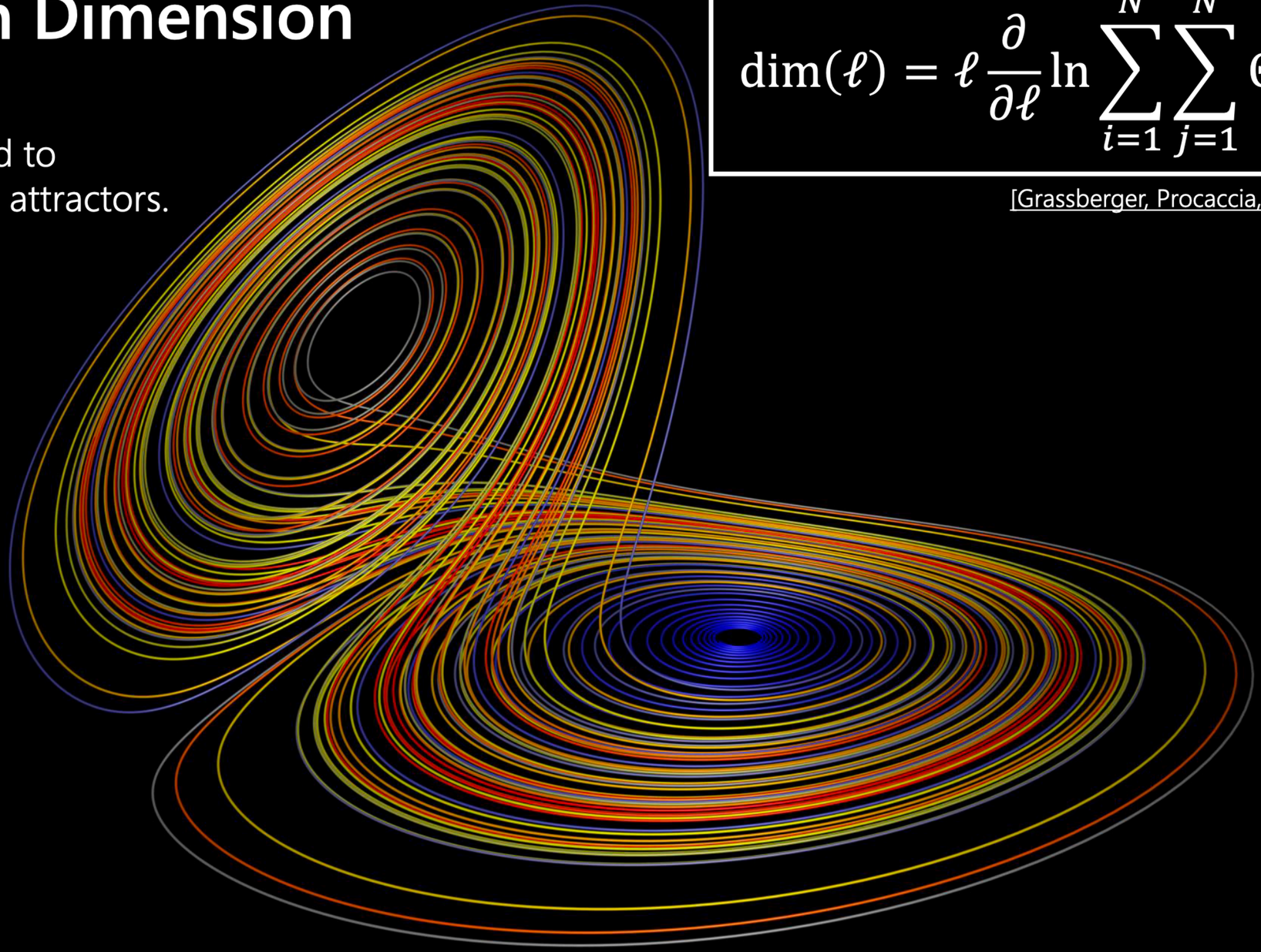
A spectrum of the dataset at a glance.

Correlation Dimension

Originally introduced to characterize strange attractors.

$$\dim(\ell) = \ell \frac{\partial}{\partial \ell} \ln \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta[\mathrm{d}(x_i, x_j) < \ell]$$

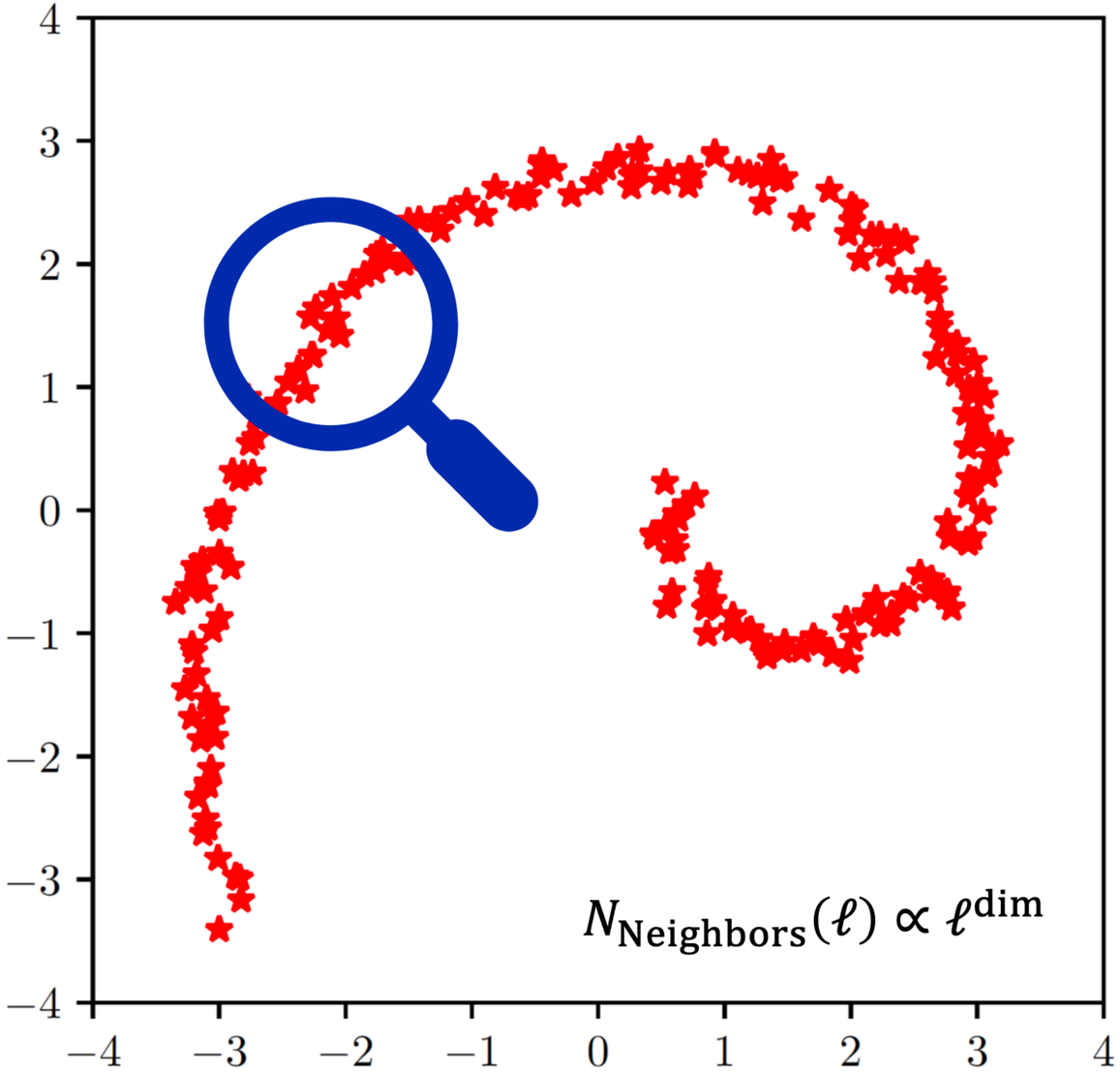
[Grassberger, Procaccia, PRL, 1983] [Kegl, NeurIPS, 2002]

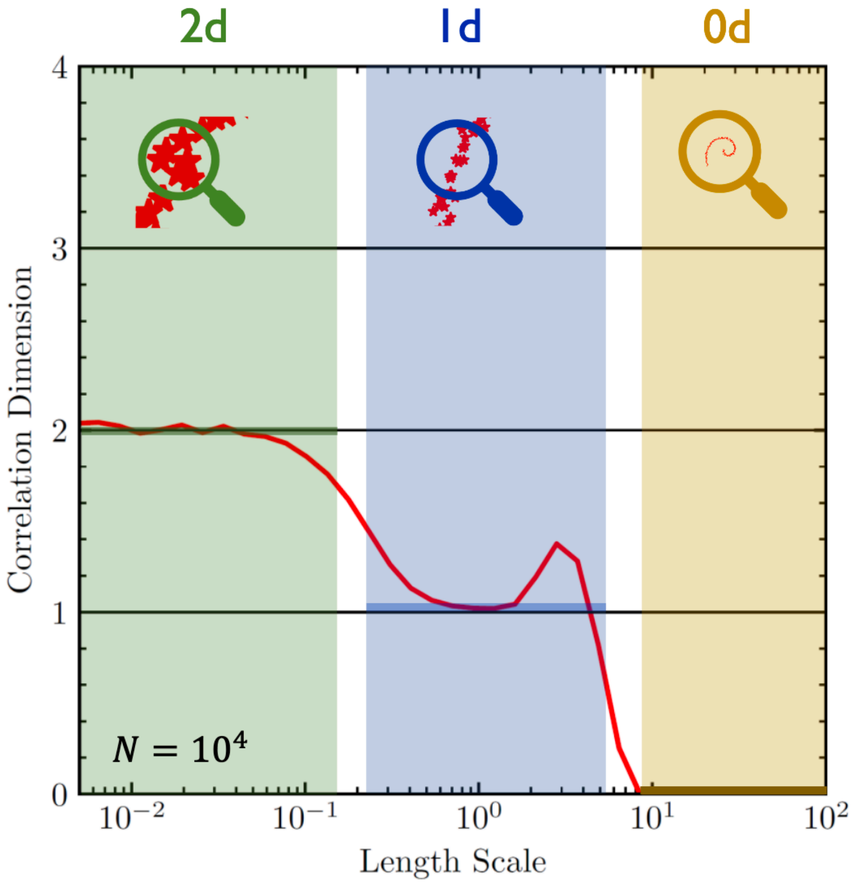


Enabling New Directions

$$\dim(\ell) = \ell \frac{\partial}{\partial \ell} \ln \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta[\mathrm{d}(x_i, x_j) < \ell]$$

[Grassberger, Procaccia, PRL, 1983] [Kegl, NeurIPS, 2002]

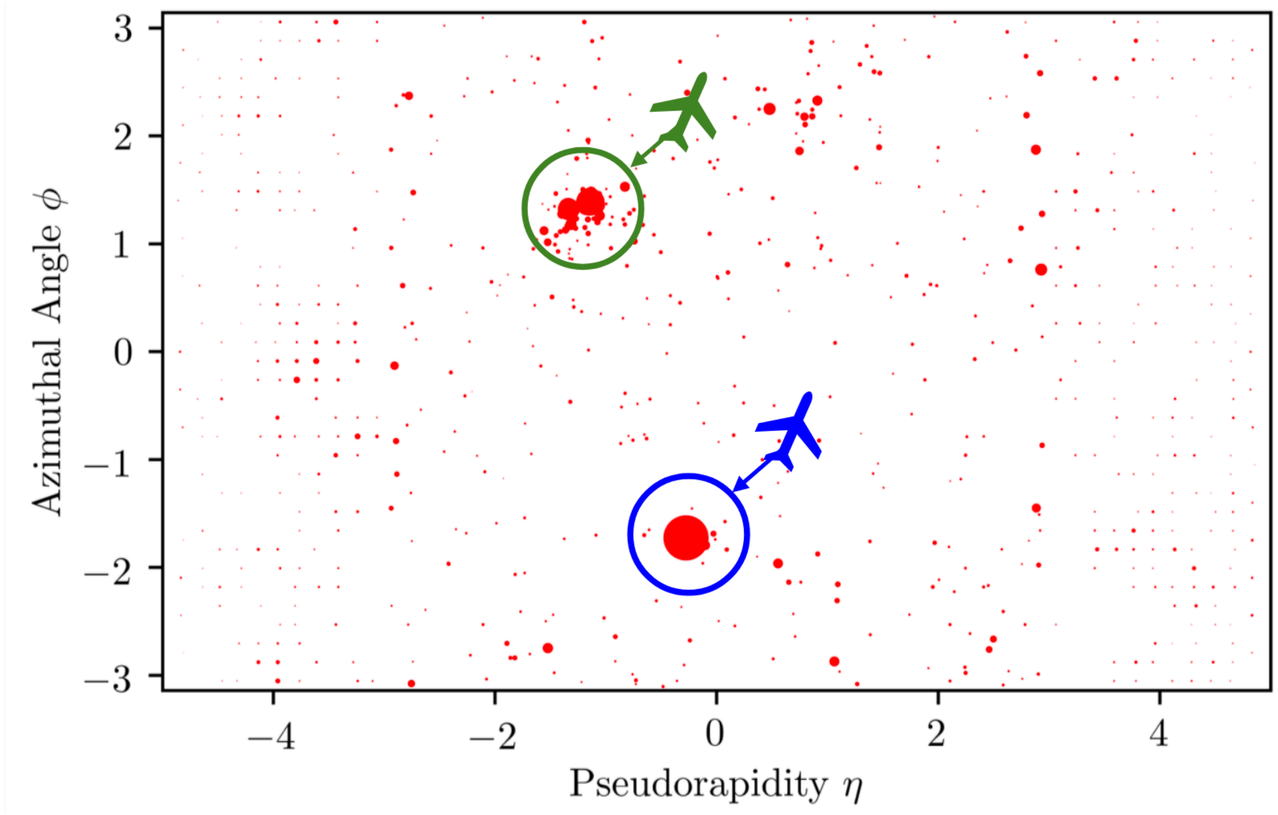




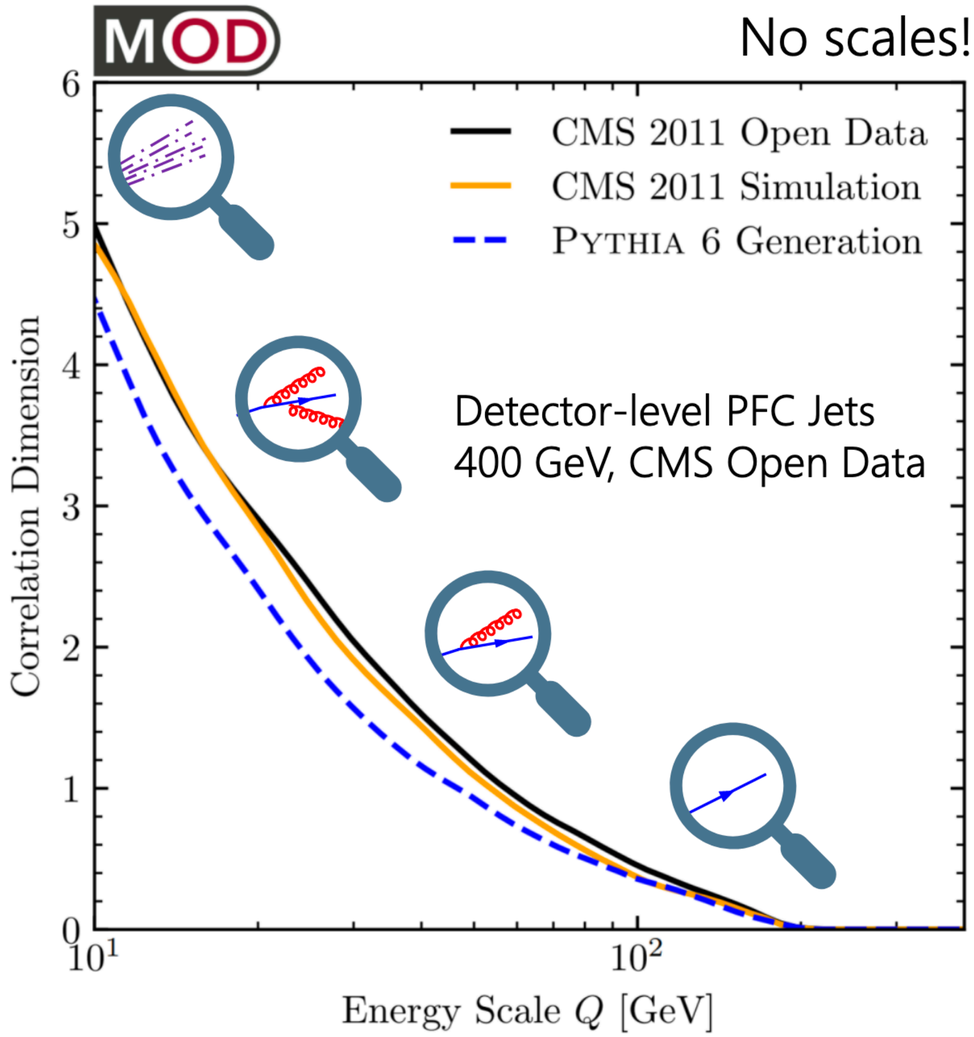
A spectrum of the dataset at a glance.

Enabling New Directions: The Fractal Dimension of QCD

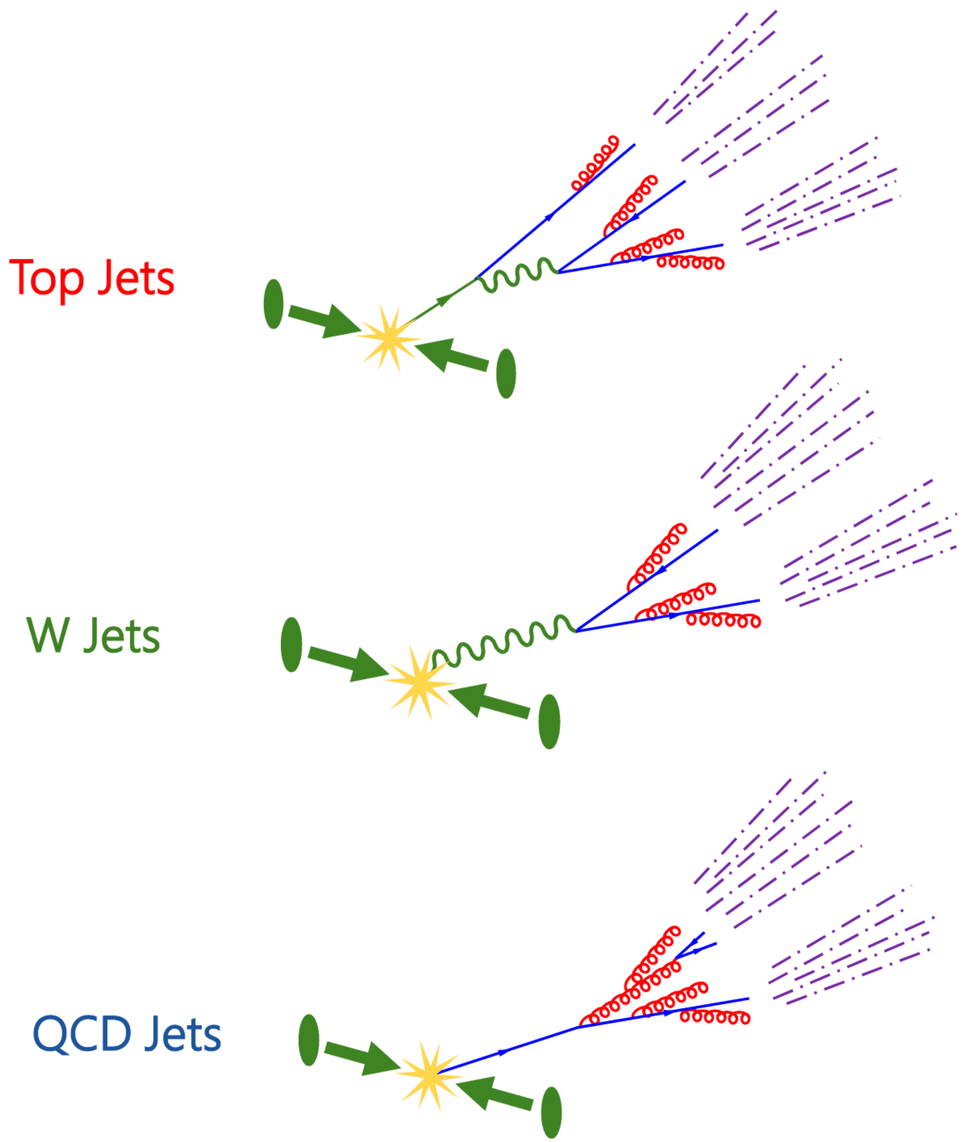
$$\dim(Q) = Q\frac{\partial}{\partial Q}\ln\sum_{i=1}^{N}\sum_{j=1}^{N}\Theta[\mathrm{EMD}(\mathcal{E}_i,\mathcal{E}_j) < Q]$$

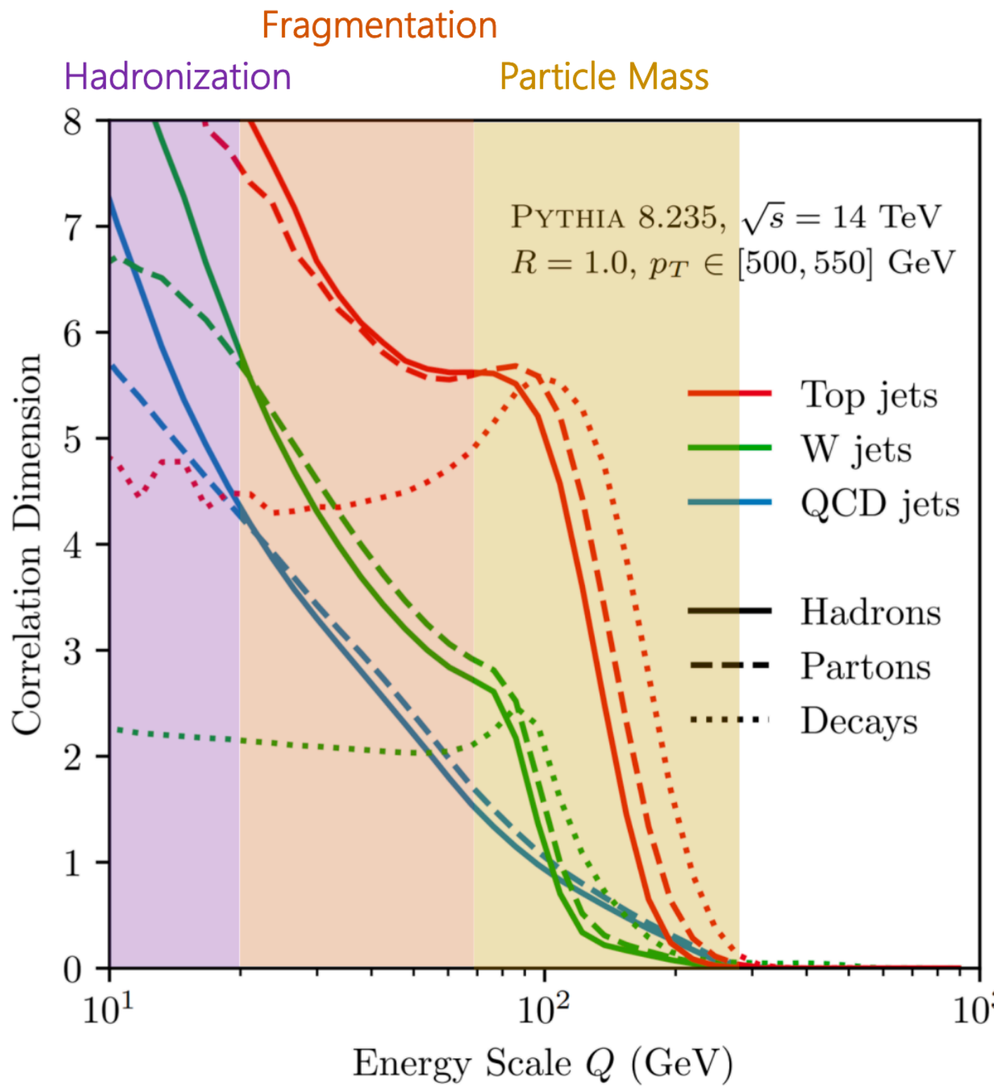


+ centering and rotation



Enabling New Directions: The Fractal Dimension of the SM





[Komiske, EMM, Thaler, PRL, 1902.02346]