

Collision Course

Particle Physics meets Machine Learning

Jesse Thaler



University of Maryland Physics Colloquium — October 13, 2020

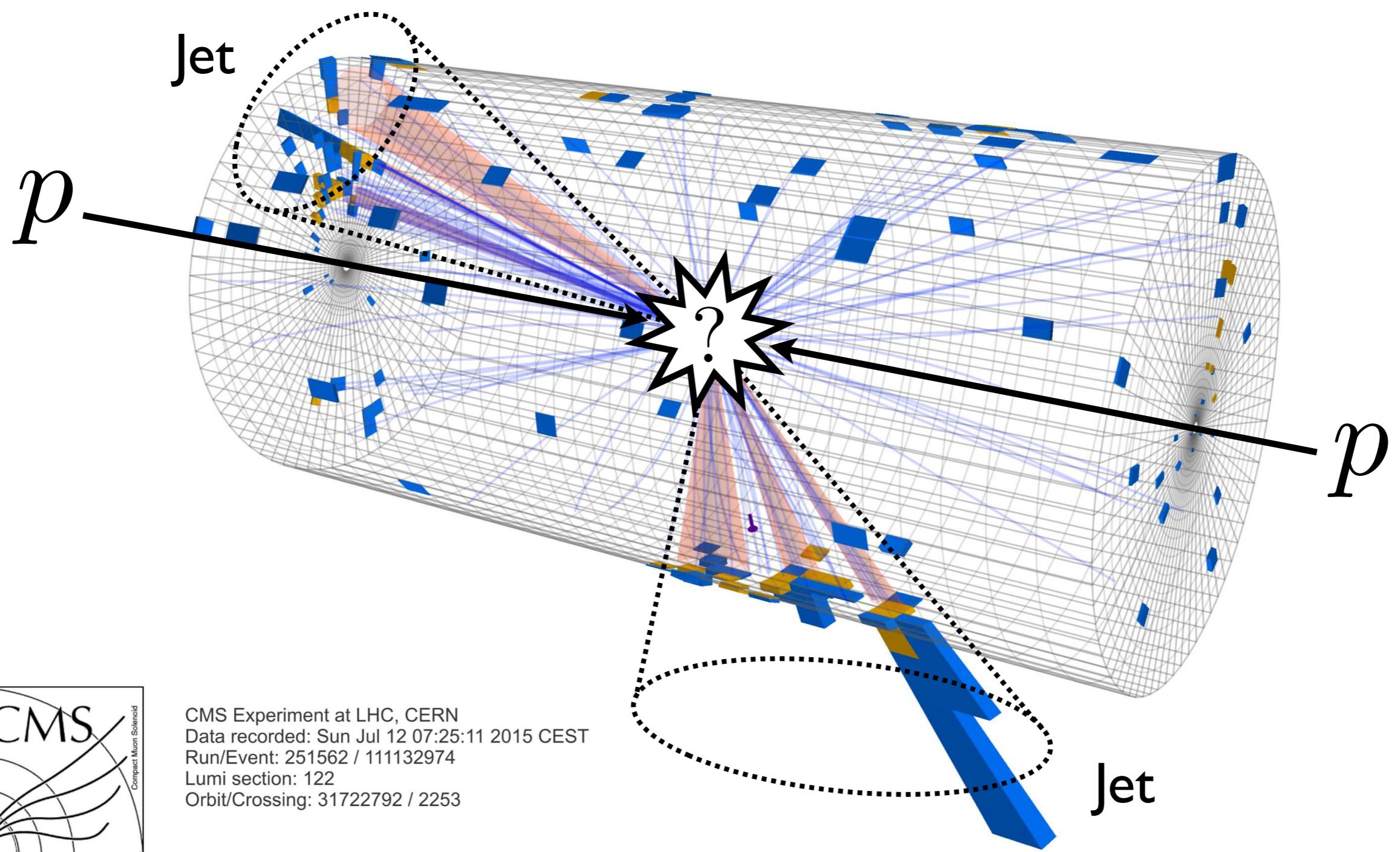
Wearing my New Hat...

The NSF AI Institute for
Artificial Intelligence and Fundamental Interactions



Postdoctoral fellowship opportunity (Oct 20 deadline):
<http://iaifi.org/fellows.html>

“Collision Course”

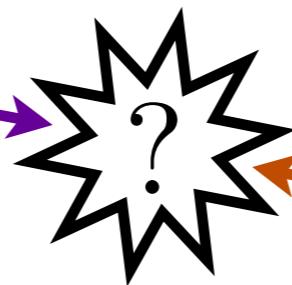


“Collision Course”

“Theoretical Physics for Machine Learning”
Aspen Center for Physics, January 2019

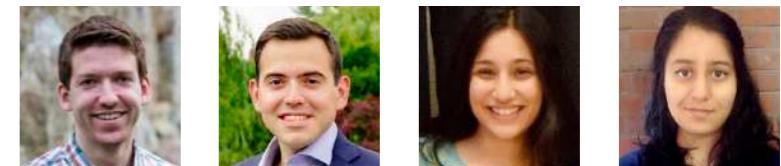
Theoretical
(High Energy)
Physics

Mathematics,
Statistics,
Computer Science



*New insights into particle physics**
*facilitated by advances in machine learning**
(and vice versa?)

Today's Talk: Two Anecdotes

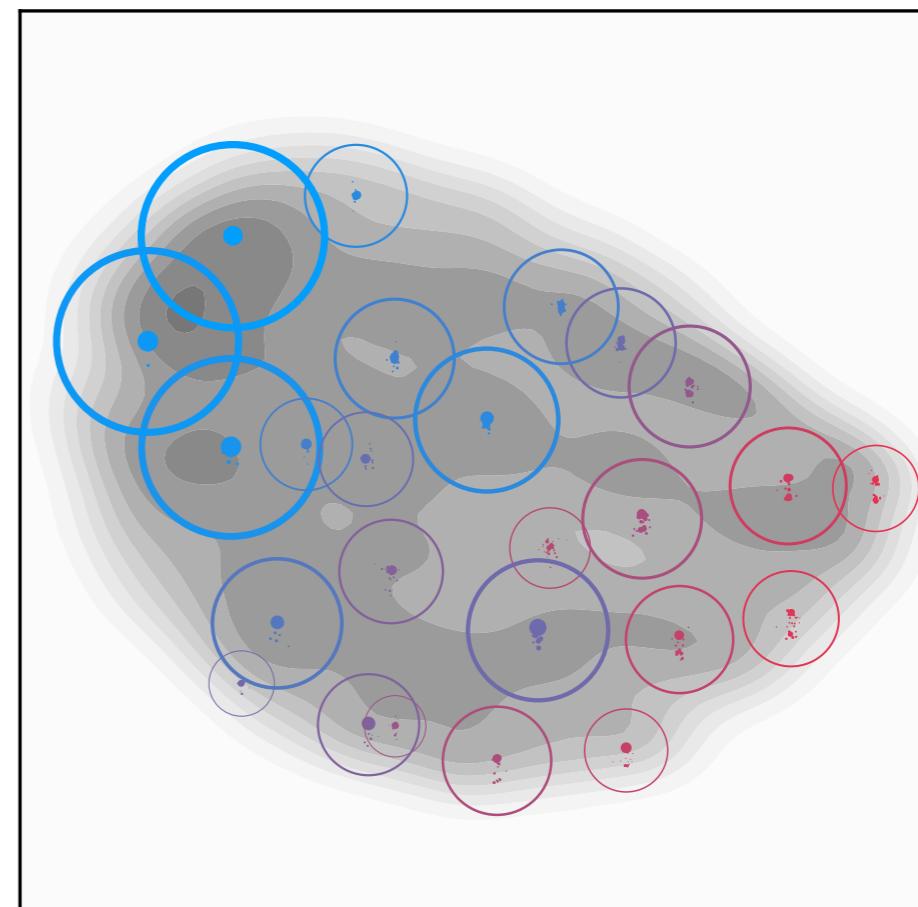


Teaching a Machine to
“Think Like a Physicist”



[Komiske, Metodiev, JDT, [JHEP 2019](#)]

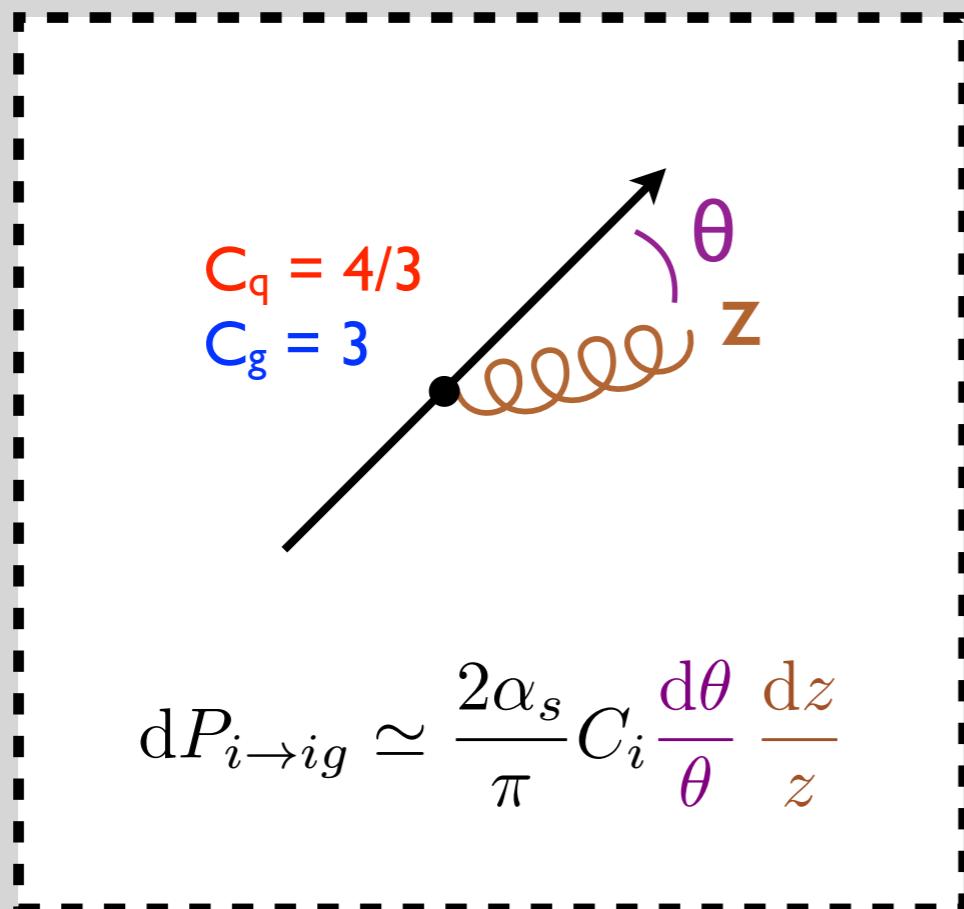
Letting Collider Data
Speak for Itself



[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
based on Komiske, Metodiev, JDT, [PRL 2019](#)]

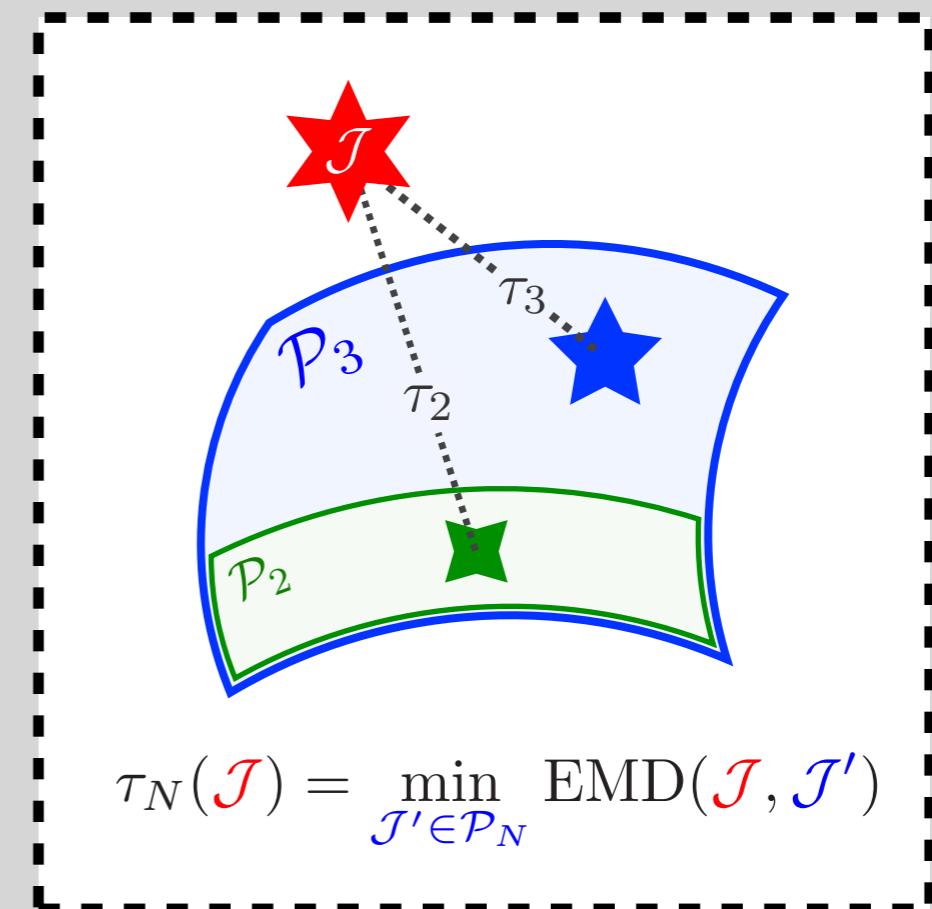
*Data analysis strategies motivated by the
symmetries and structures of particle physics*

Exploiting a Core Prediction of QCD



[Altarelli, Parisi, [NPB 1977](#)]

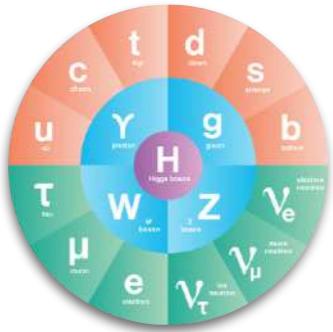
Nested Singularities of Gauge Theories



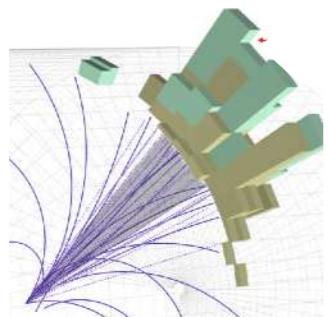
[Stewart, Tackmann, Waalewijn, [PRL 2010](#);
JDT, Van Tilburg, [JHEP 2011](#), [JHEP 2012](#);
rephrased via Komiske, Metodiev, JDT, [JHEP 2020](#)]

New perspectives on key theoretical concepts

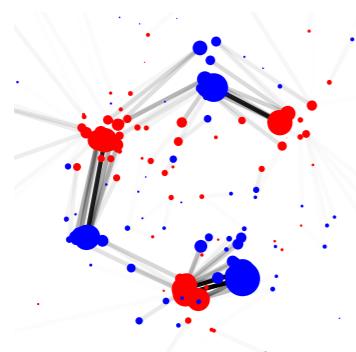
Outline



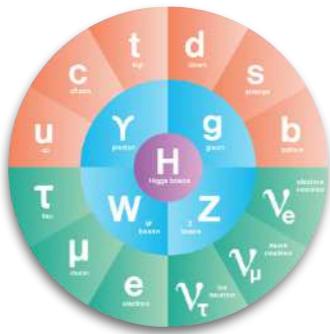
Particle Physics 101



What is a Collider Event?



When are Collider Events Similar?



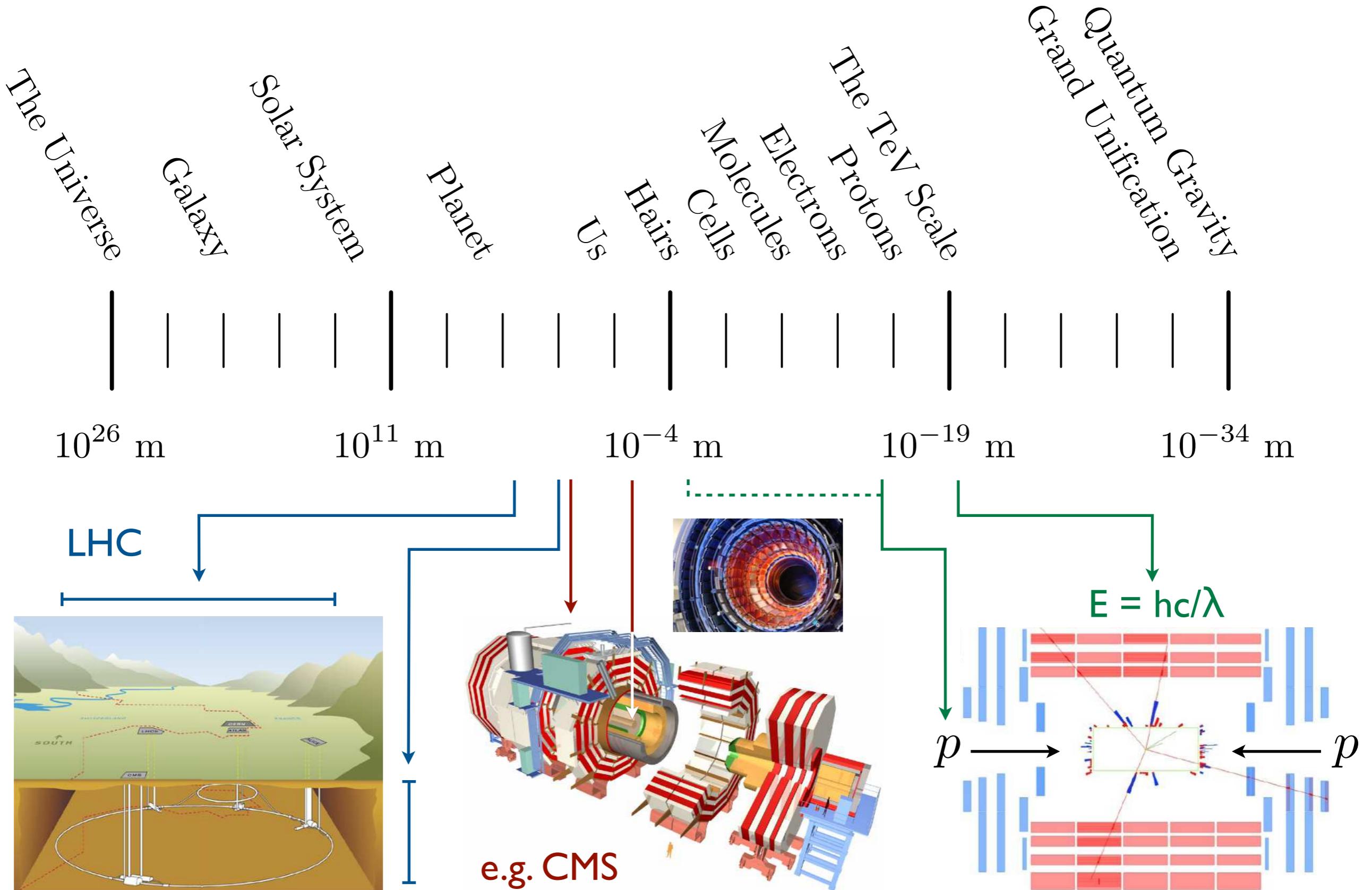
Particle Physics 101



What is a Collider Event?



When are Collider Events Similar?



Mesons

$\pi^\pm \pi^0 \eta K^\pm K^0 \eta' D^\pm D^0 D_s^\pm \eta_c B^\pm B^0 B_s^0 \eta_b \dots$
 $\rho^\pm \rho^0 \omega K^{*\pm} K^{*0} \phi D^{*\pm} D^{*0} D_s^{*\pm} J/\psi B^{*\pm} B^{*0} B_s^{*0} \Upsilon \dots$

Baryons

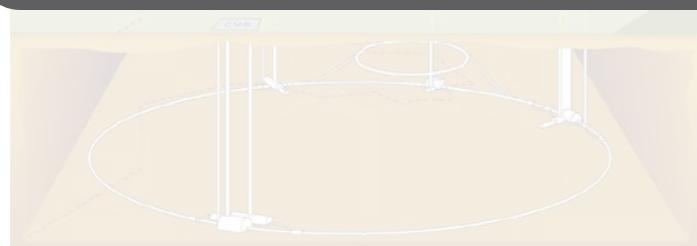
$p n \Lambda^0 \Sigma^+ \Sigma^0 \Sigma^- \Xi^0 \Xi^- \dots$
 $\Delta^{++} \Delta^+ \Delta^0 \Delta^- \Sigma^{*+} \Sigma^{*0} \Sigma^{*-} \Xi^{*0} \Xi^{*-} \Omega^- \dots$

Tetraquarks

$X(3872) Y(4260) Z(4430) \dots$

Pentaquarks

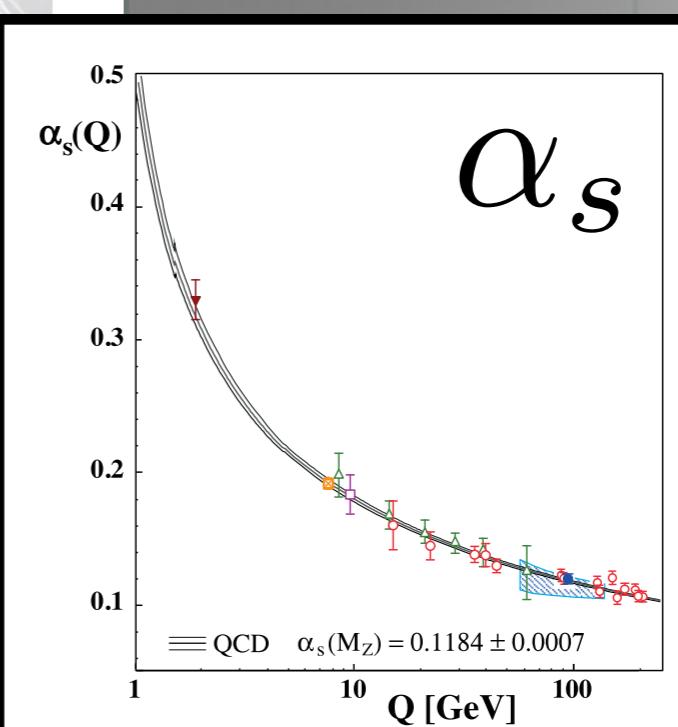
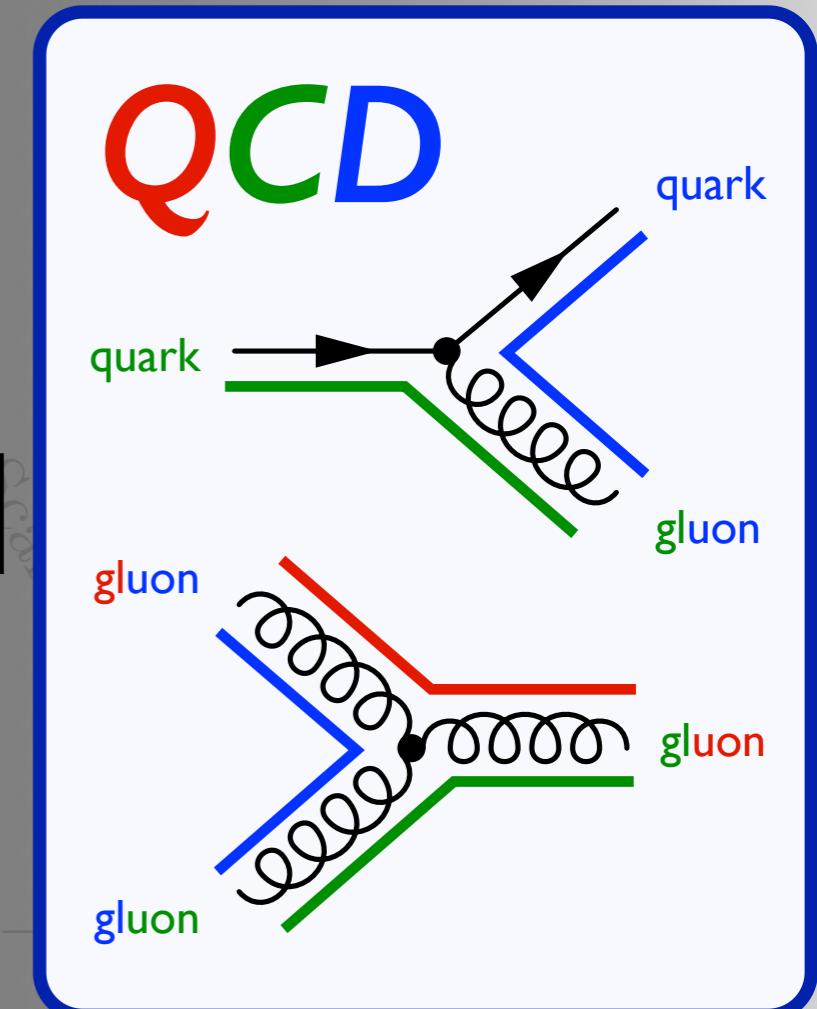
$P_c^+(4450) \dots$



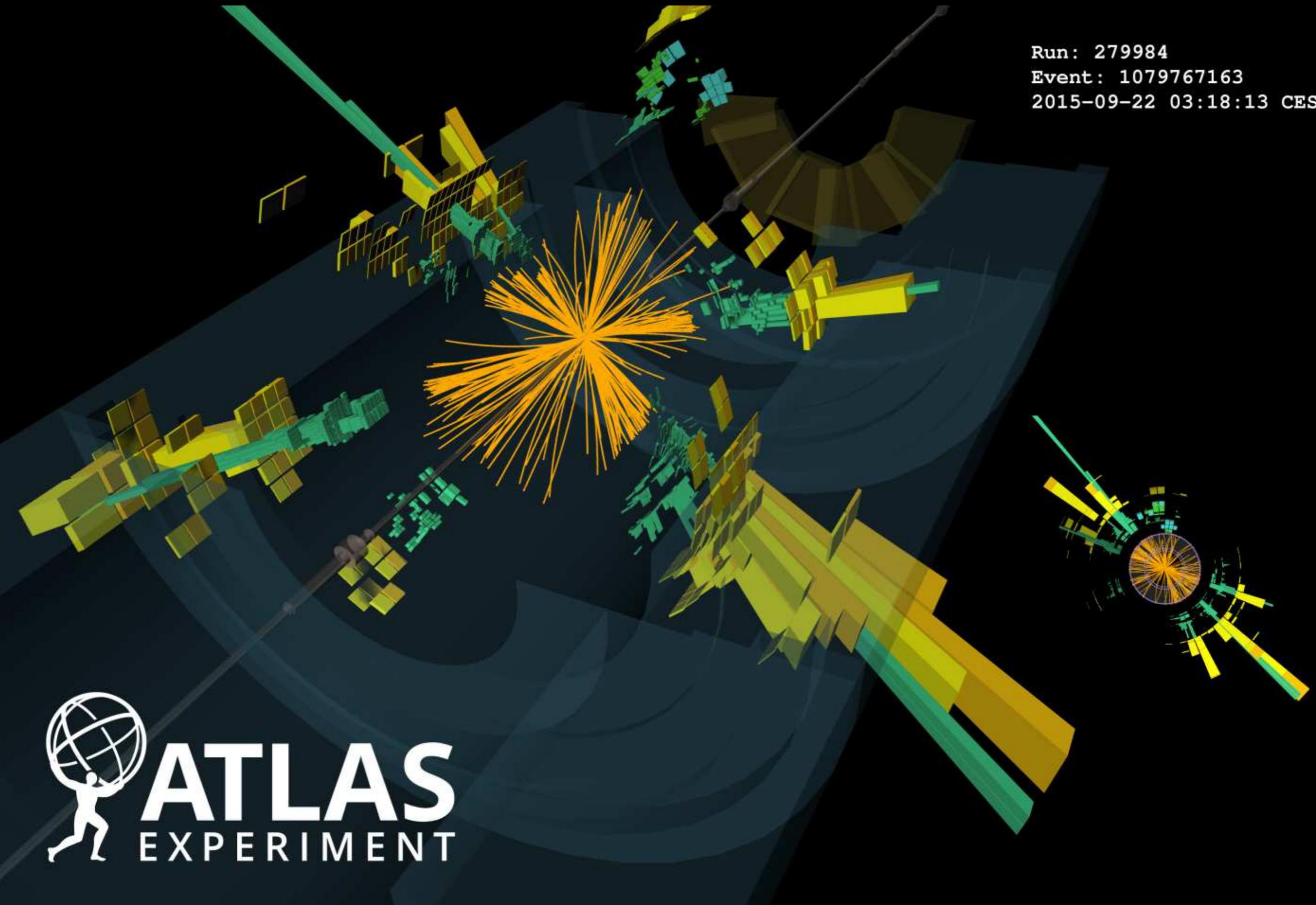
|

e.g. CMS

Jets

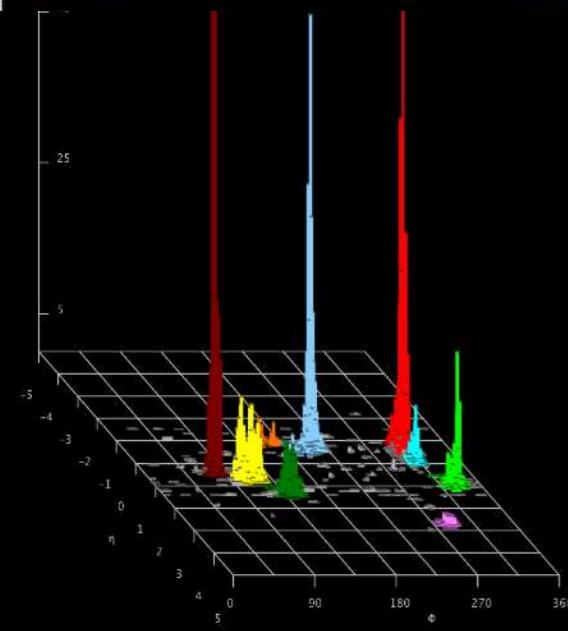
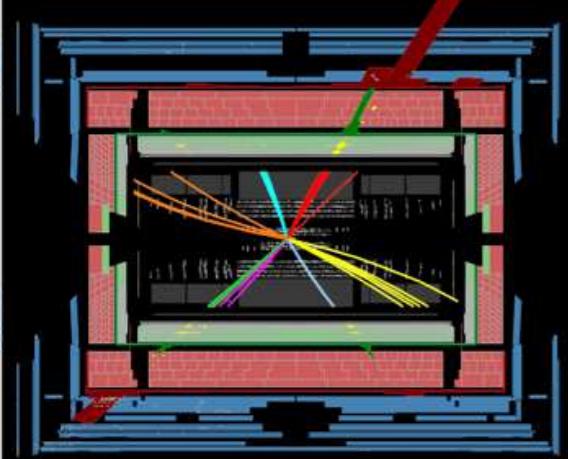


Run: 279984
Event: 1079767163
2015-09-22 03:18:13 CEST

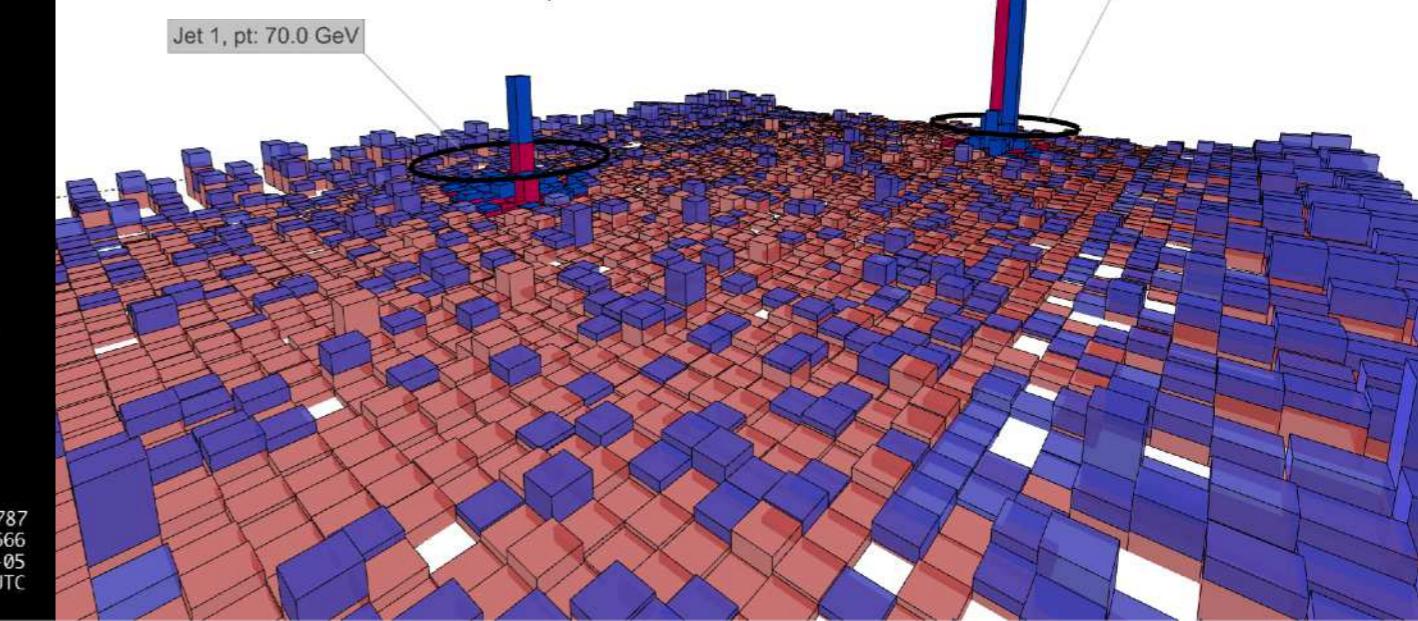
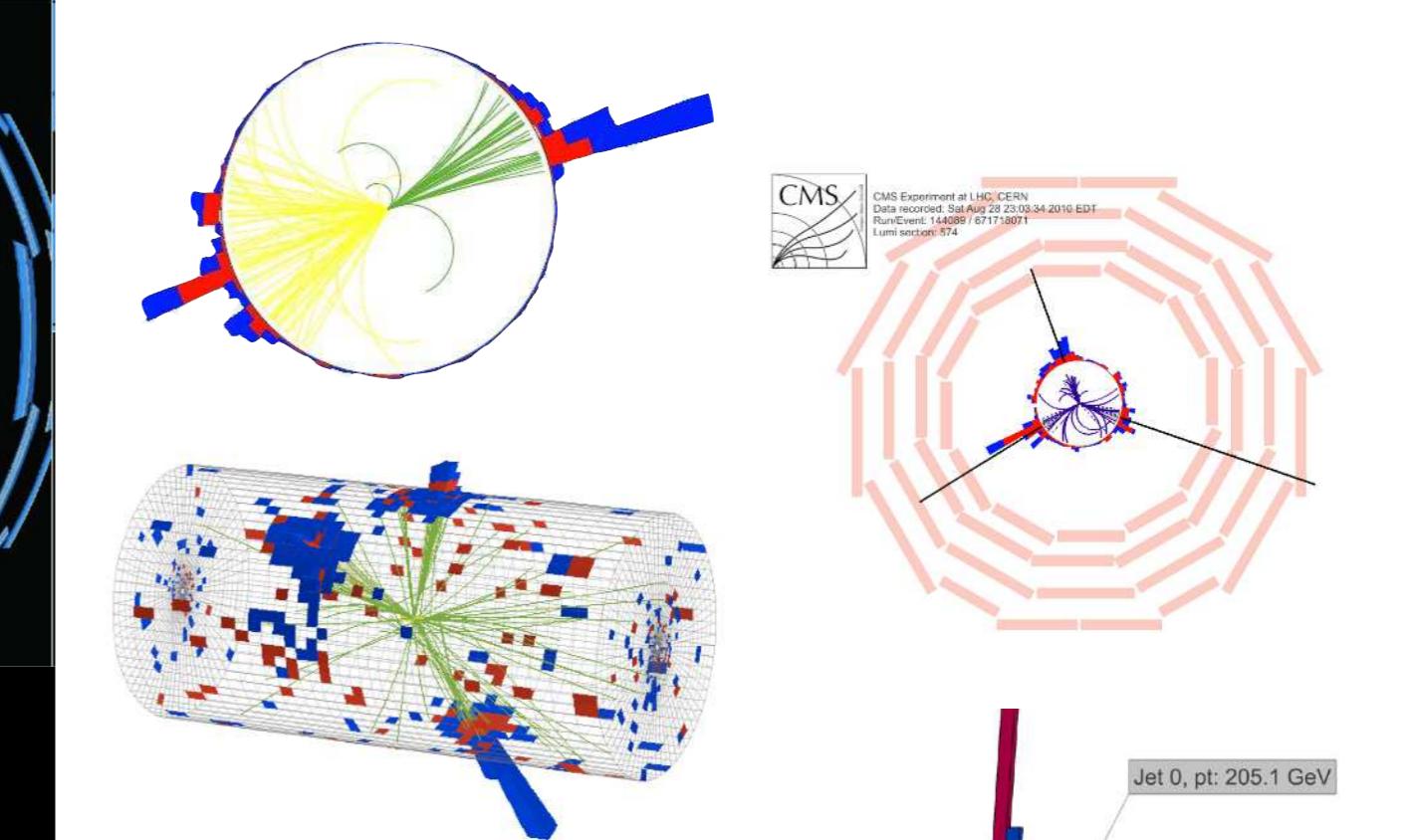
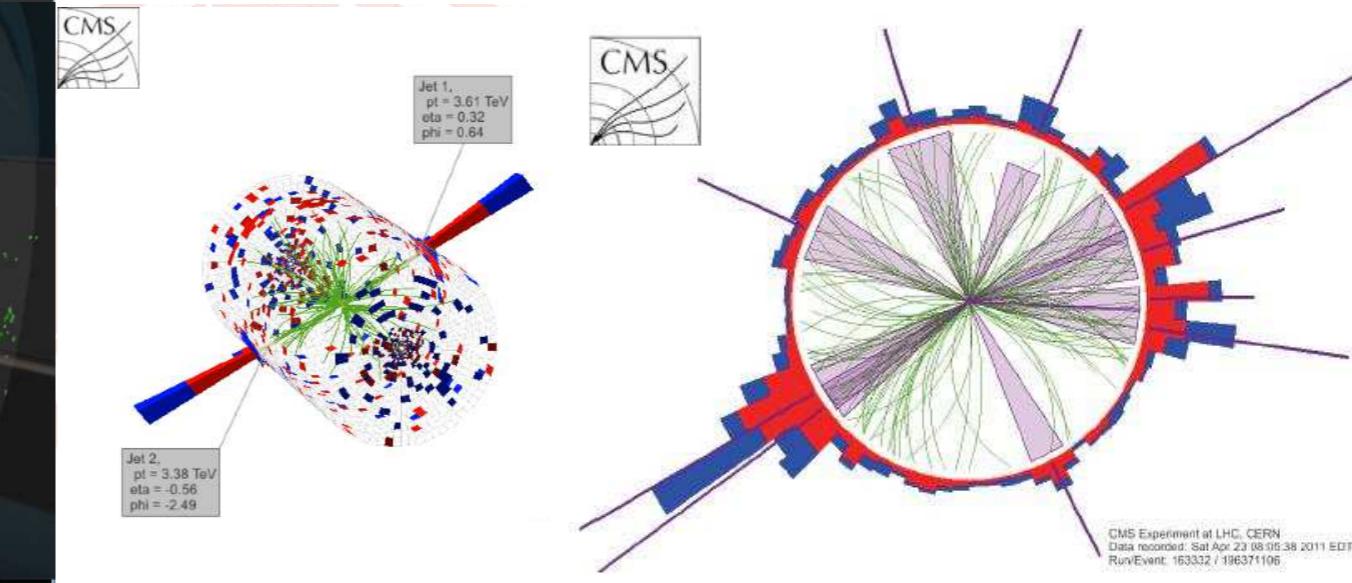
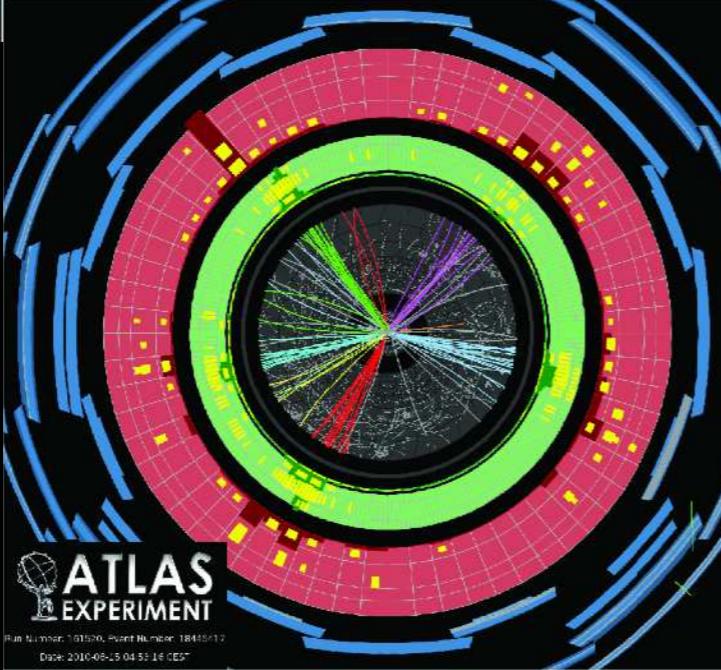
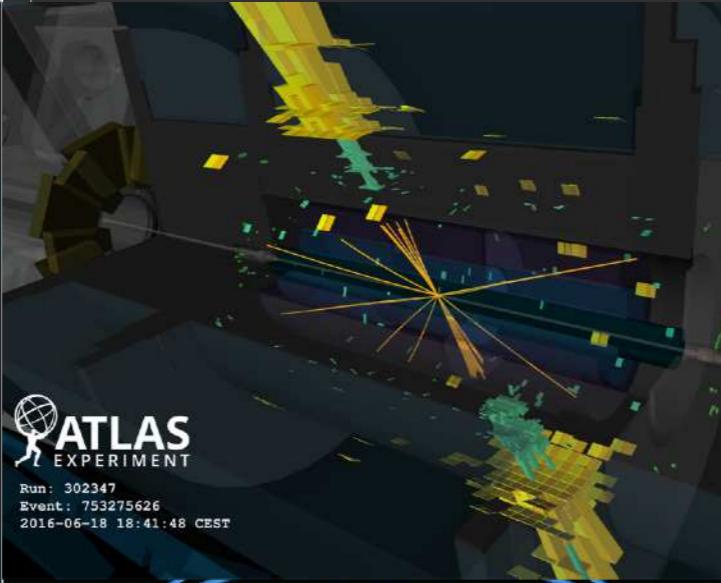
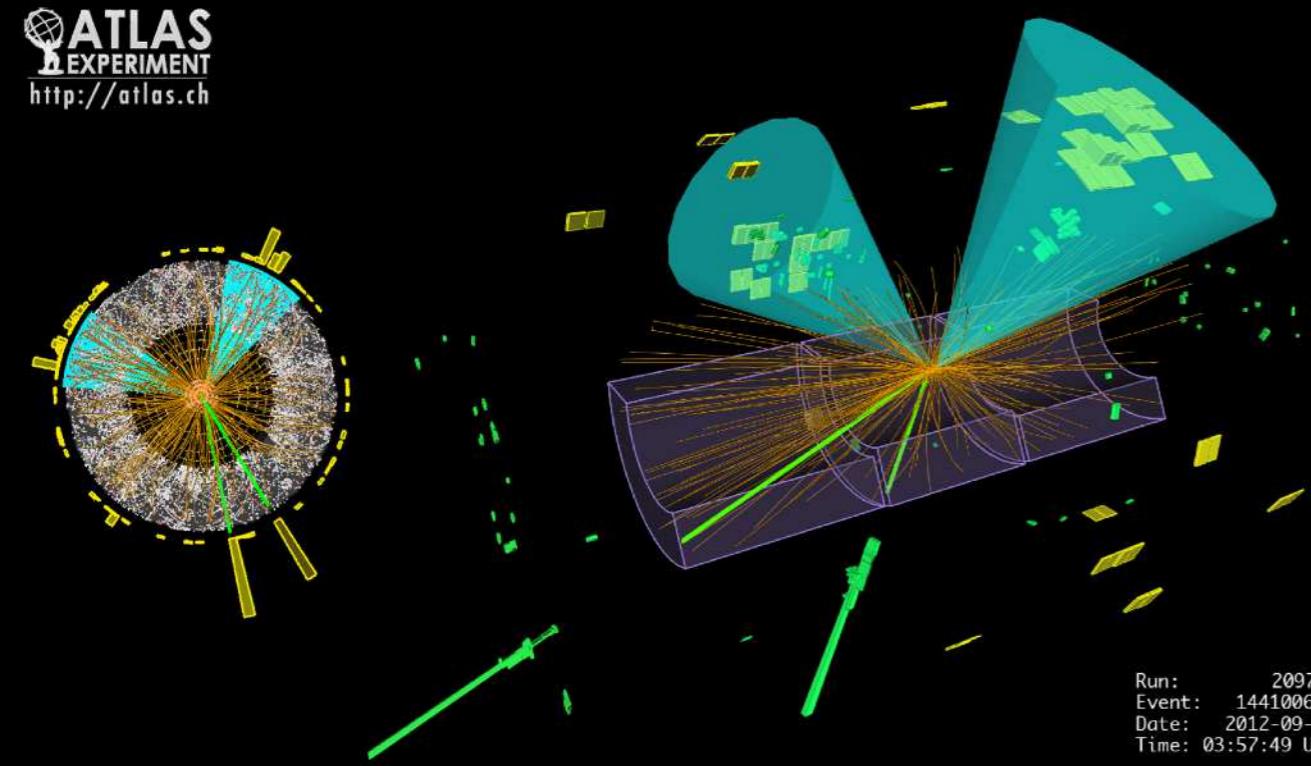


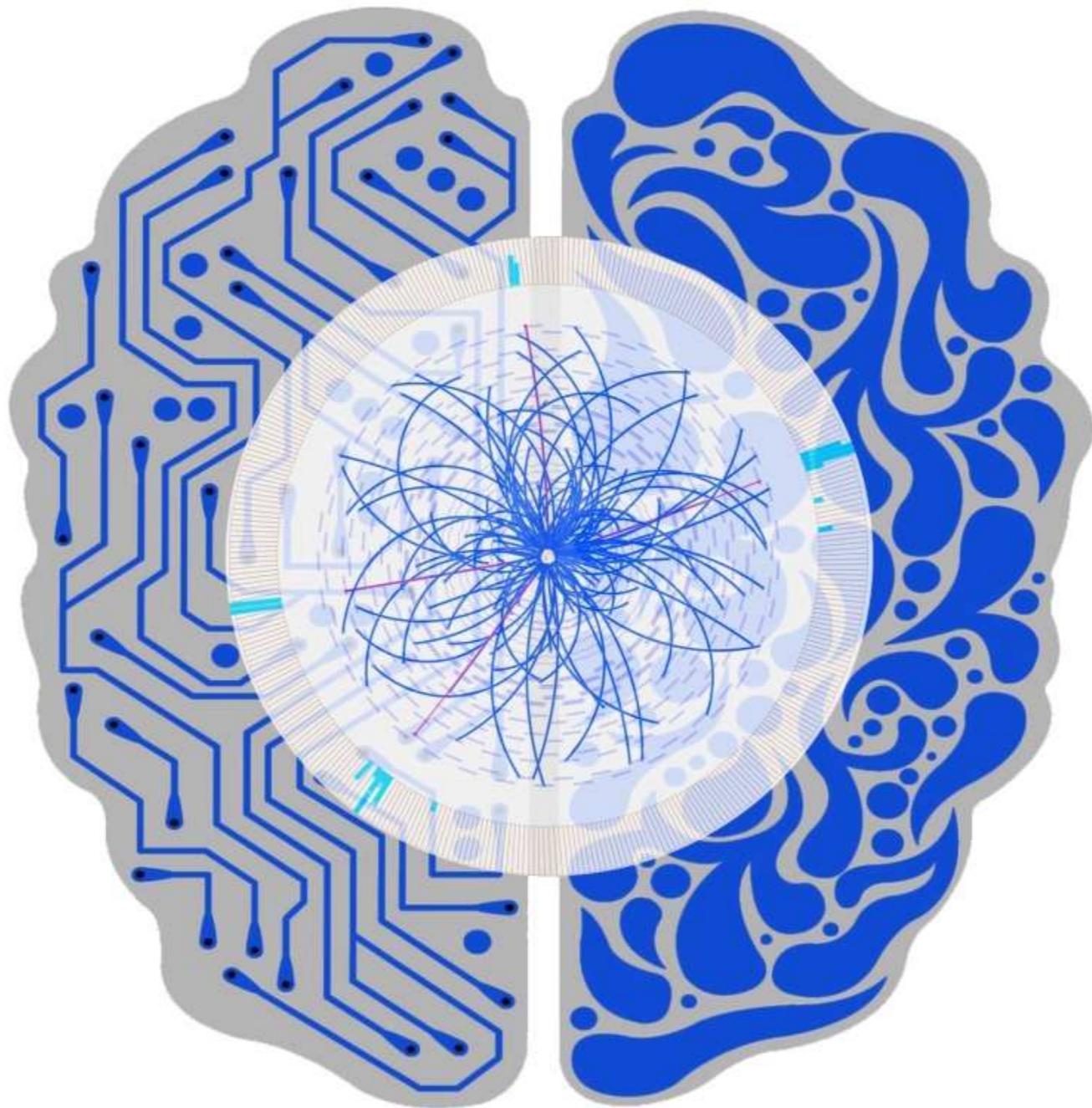
Run Number: 159224, Event Number: 3533152

Date: 2010-07-18 11:05:54 CEST



ATLAS
EXPERIMENT
<http://atlas.ch>

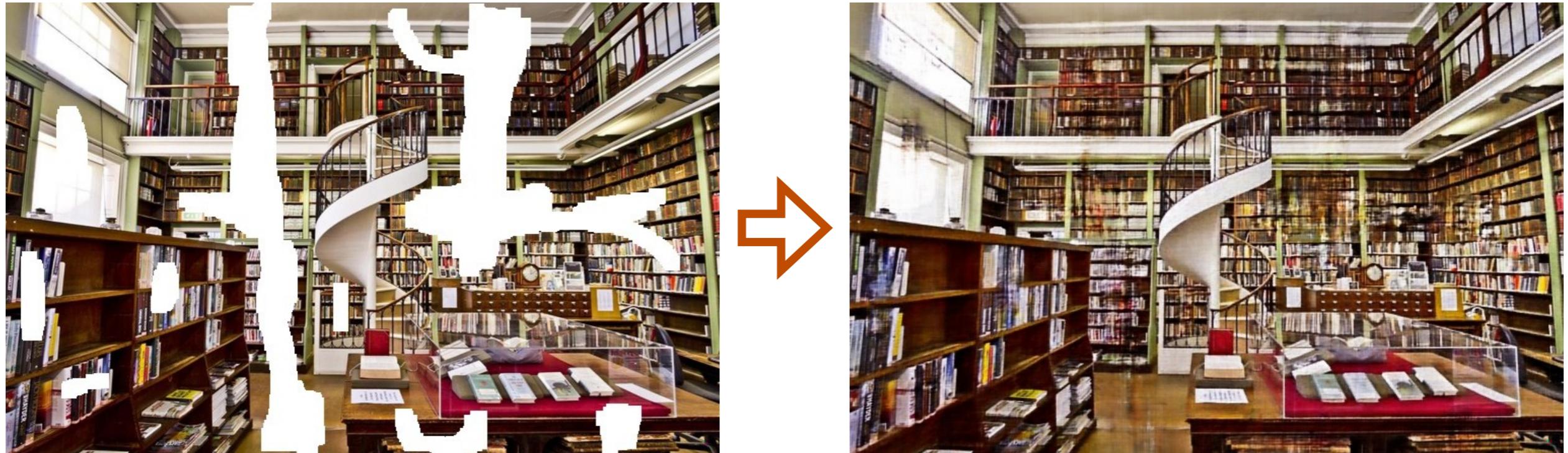




The Rise of Machine Learning

Deep Learning

E.g. Inpainting

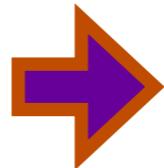


increased computational power and large data sets

[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

Deep Learning meets Deep Thinking

E.g. *Inpainting*



Using randomly initialized neural network (!)

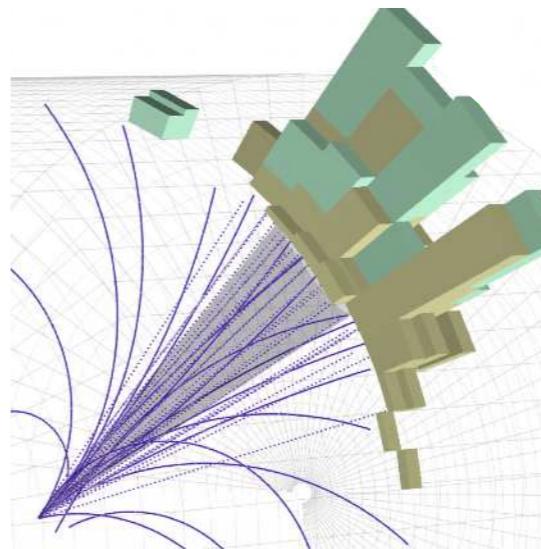
Progress made by understanding the structure of problems
(not just increased computational power and large data sets)

[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

Cartoon of Machine Learning

"ML4Jets"
NYU, January 2020

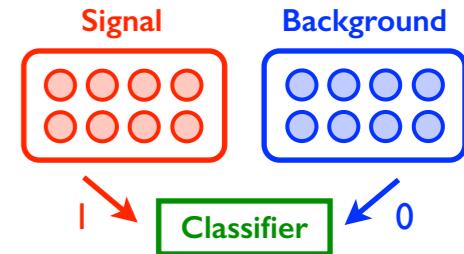
For this talk: \mathcal{J} = jet



E.g.: **Problem** = Minimize loss function
Solution = Multi-layer neural network
Strategy = Stochastic gradient descent

Quark/Gluon Classification

“Hello, World!” of Jet Physics

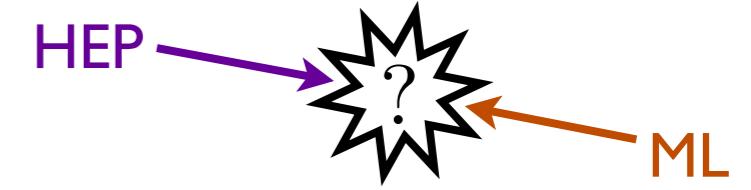


Find $h\left(\begin{array}{c} \nearrow \\ \nearrow \\ \nearrow \\ \nearrow \\ \nearrow \end{array}\right)$ such that

$$h(\text{Quark}) = 1$$
$$h(\text{Gluon}) = 0$$

Best you can do: $h(\mathcal{J}) = \frac{p(\mathcal{J}|Q)}{p(\mathcal{J}|Q) + p(\mathcal{J}|G)}$
(Neyman-Pearson lemma)

E.g. Search for Supersymmetry

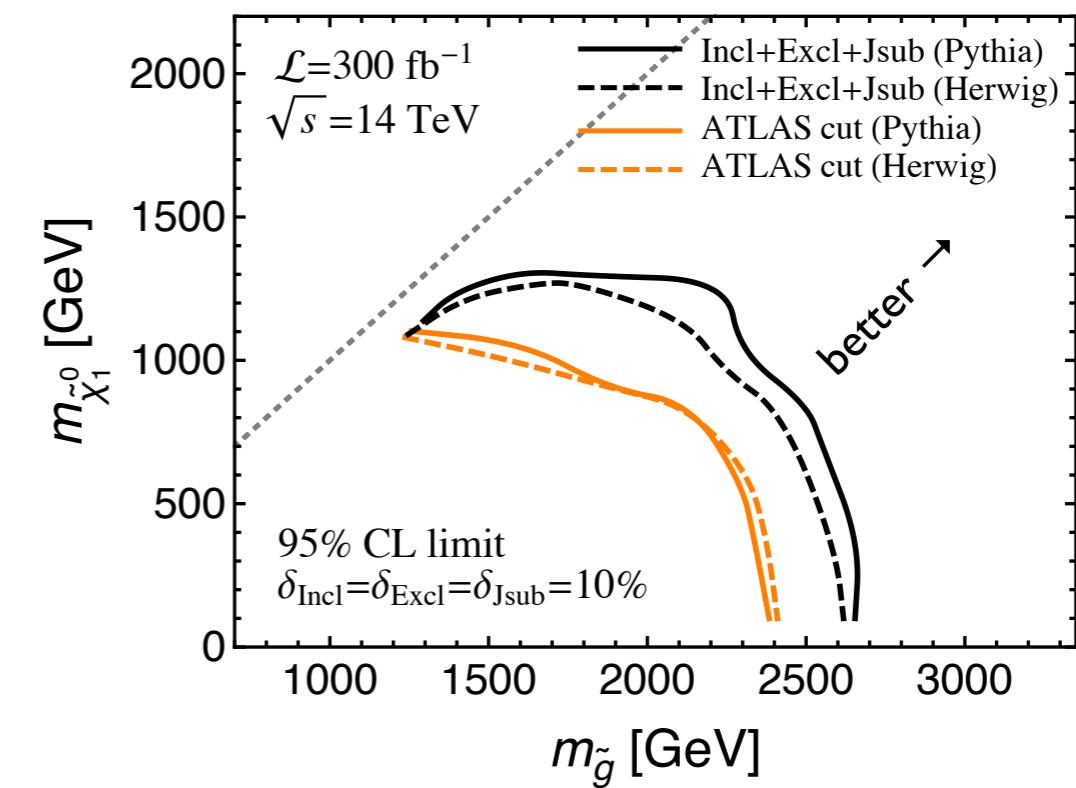
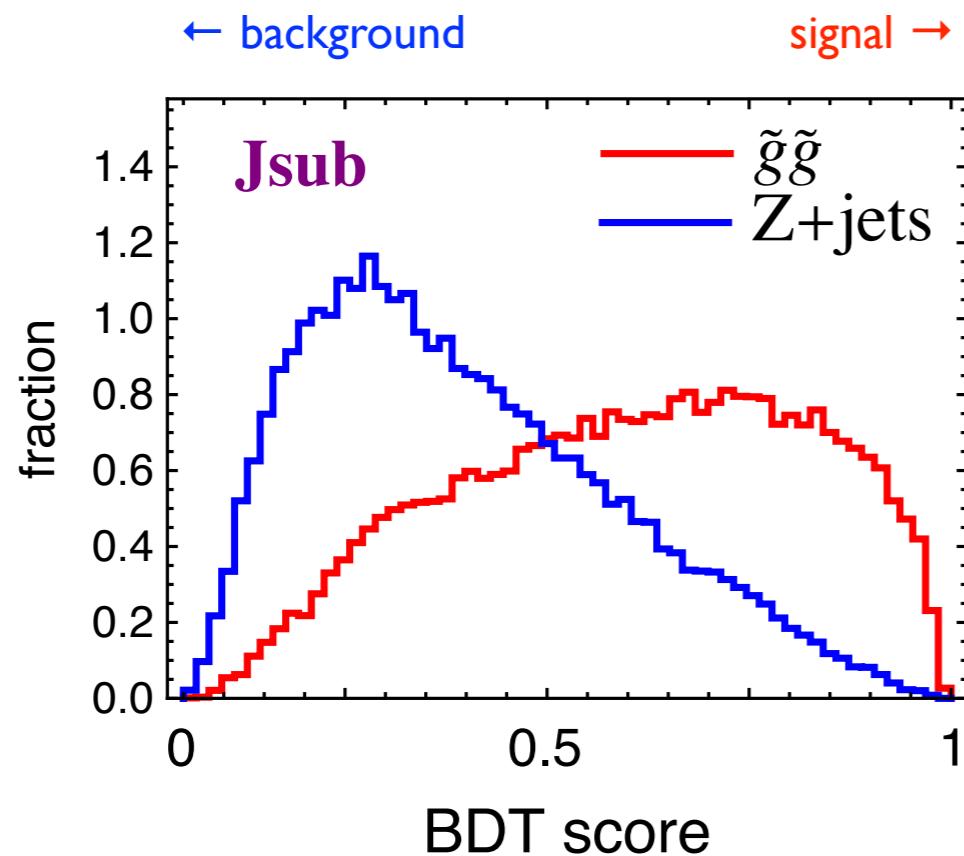
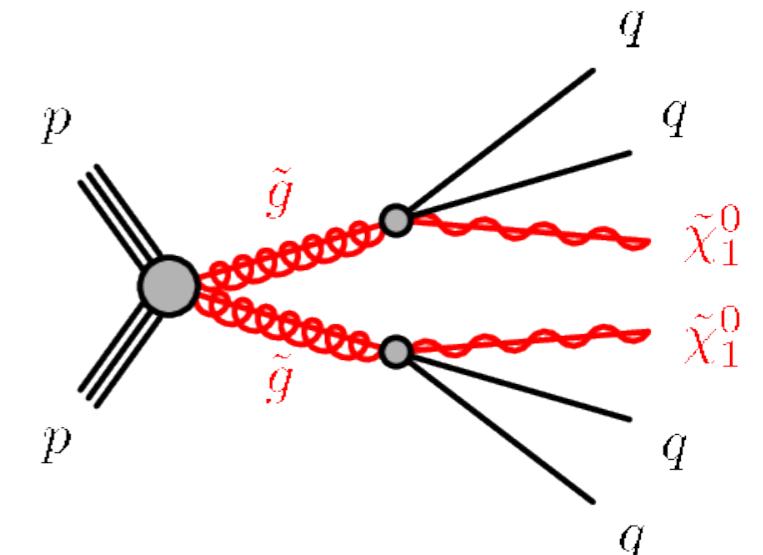


Classifier: Boosted decision tree (for each of 4 jets)

Inputs: Jet mass, width, track multiplicity

Signal: Quark enriched

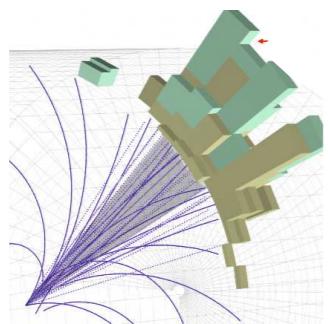
Background: Gluon enriched



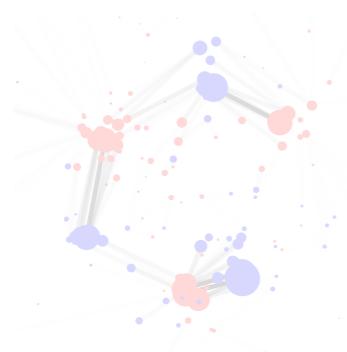
[Bhattacherjee, Mukhopadhyay, Nojiri, Sakakie, Webber, [JHEP 2017](#)]



Particle Physics 101



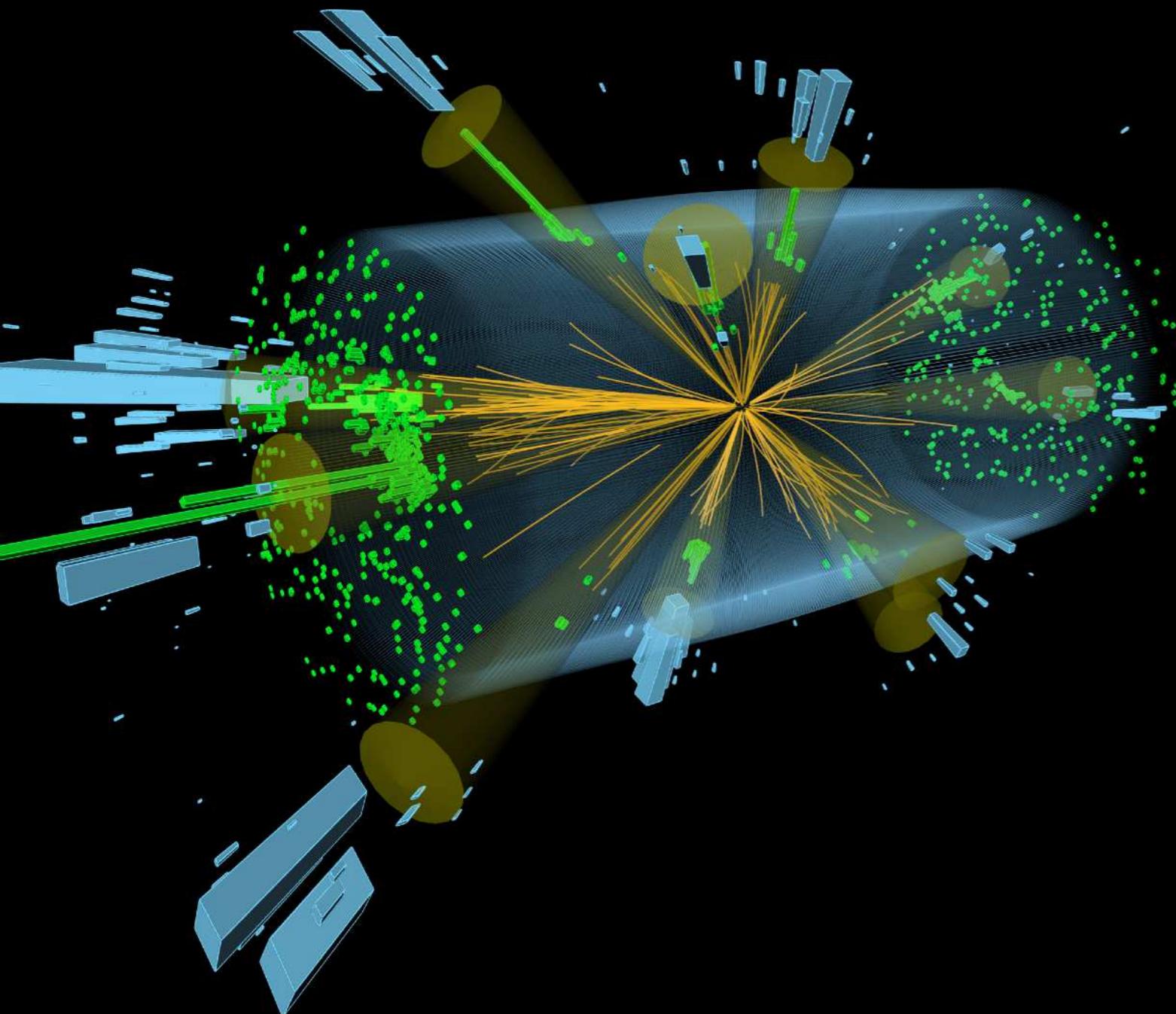
What is a Collider Event?



When are Collider Events Similar?

Collider Event

Collection of points in (momentum) space



T E H M

 γ

photon

 e^+

electron

 μ^+

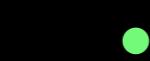
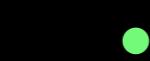
muon

 π^+

pion

 K^+

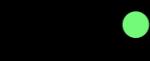
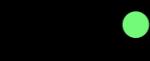
kaon

 K_L^0

K-long

 p/\bar{p}

proton

 n/\bar{n}

neutron

elementary

composite

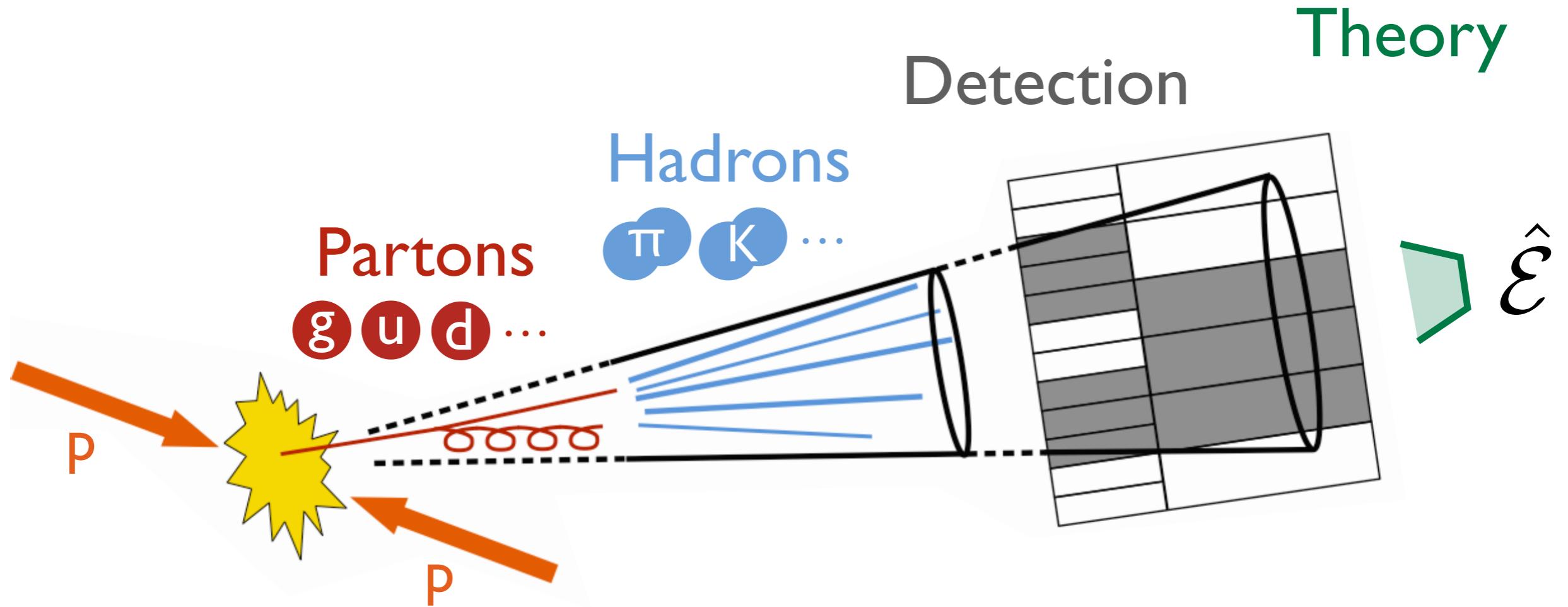
Point Cloud

Collection of points in (position) space



[Popular Science, 2013]

Dynamics of Jet Formation



Energy Flow:

Robust to hadronization and detector effects

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

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

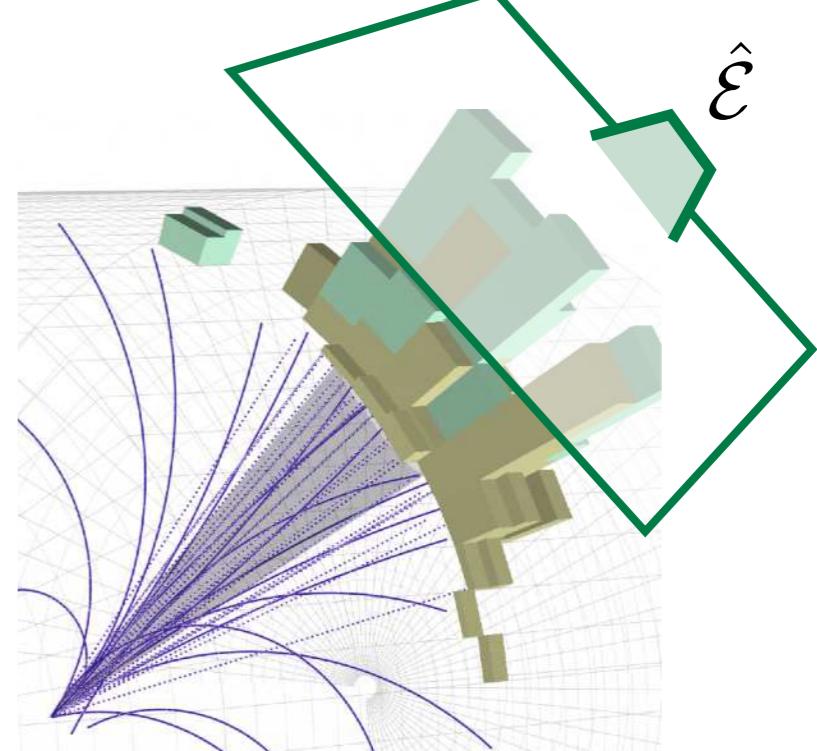
Jets as Weighted Point Clouds

- Energy-Weighted Directions

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

↑ |
Energy Direction

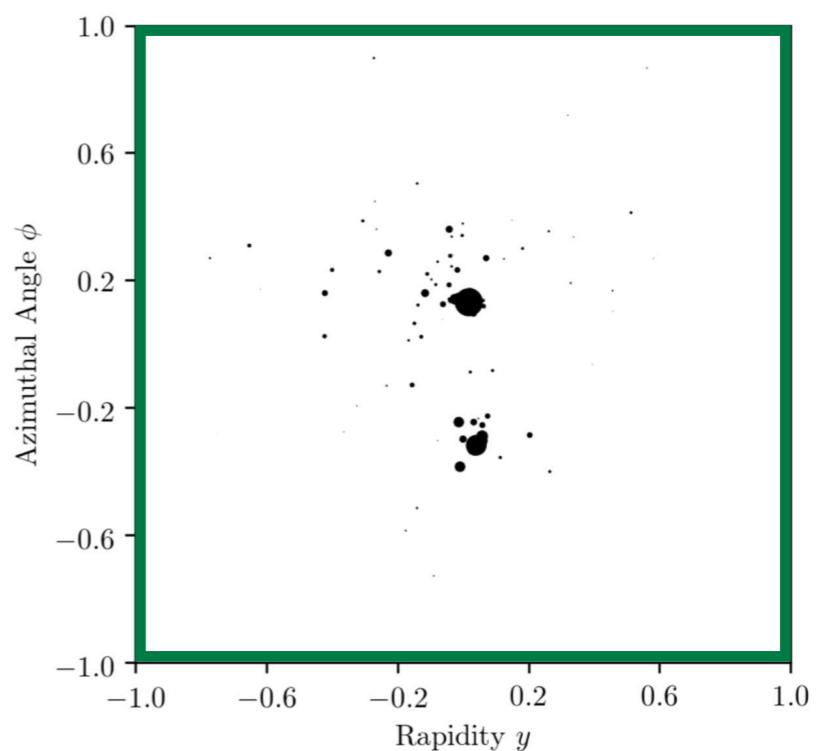
(suppressing “unsafe” charge/flavor information)



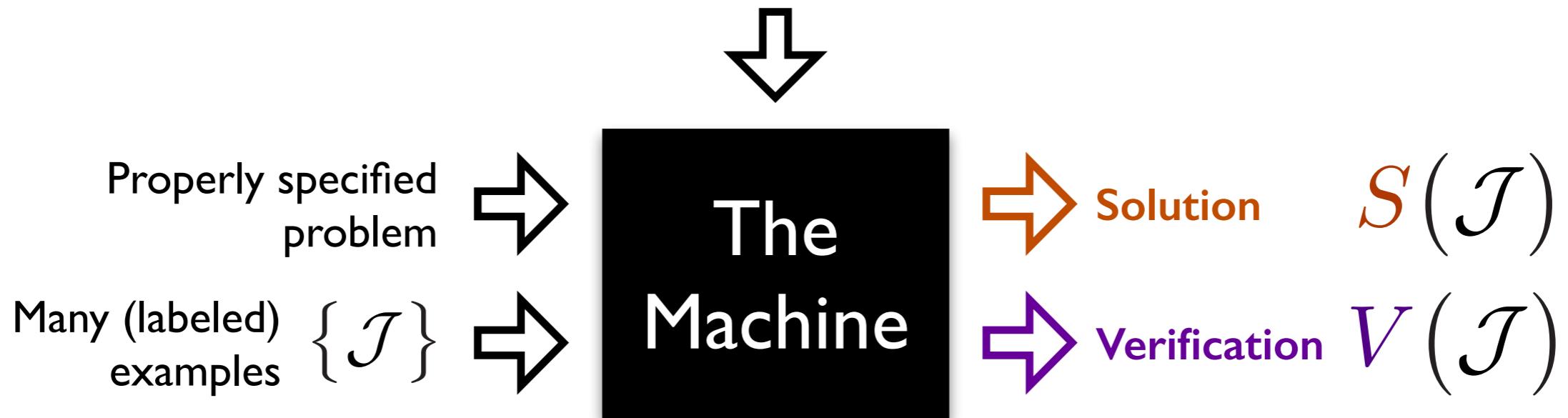
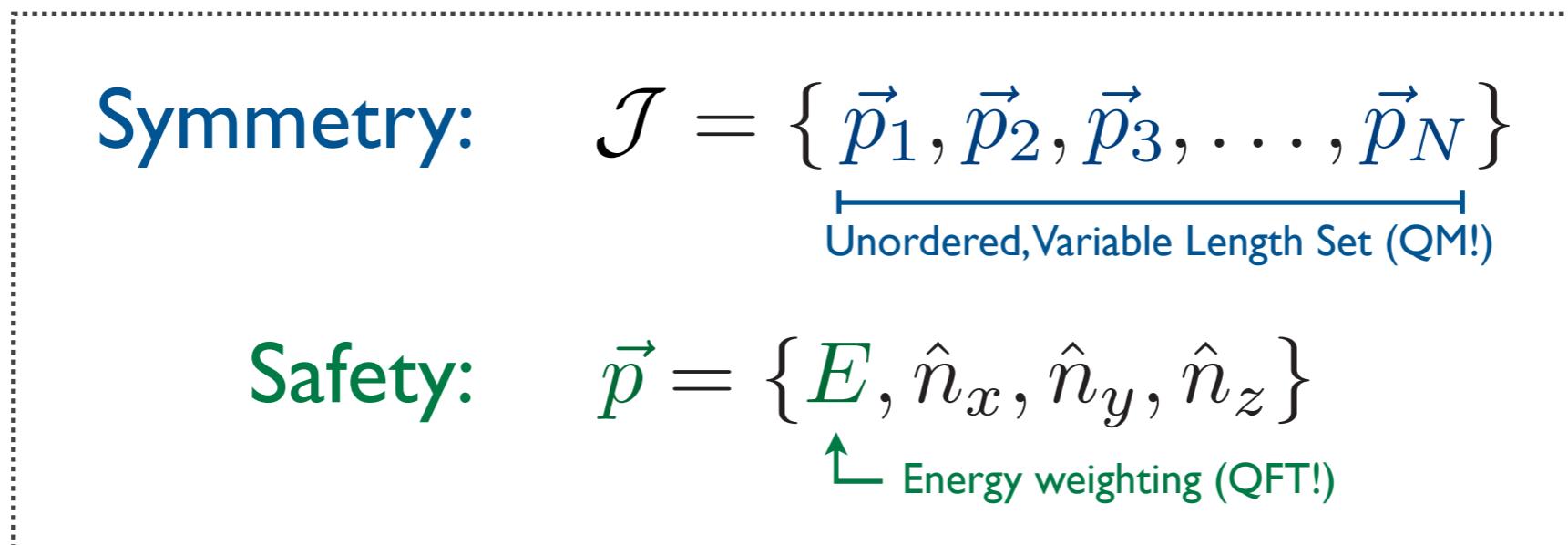
- Equivalently: Energy Density

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

↑ ↑
Energy Direction



“Thinking” Like a Physicist



*Check that answer
is physically sensible*

Energy Flow Networks

Architecture designed around **symmetries** and **interpretability**



$$S(\mathcal{J}) = F(V_1, V_2, \dots, V_\ell)$$
$$V_a(\mathcal{J}) = \sum_{i \in \mathcal{J}} E_i \Phi_a(\hat{n}_i)$$

Permutation invariant \downarrow Linear weights (i.e. safe) \downarrow

Parametrized with Neural Networks

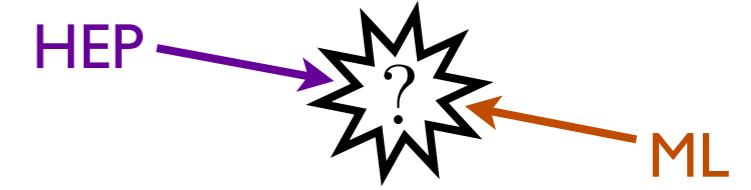
Provably describes any **safe** observable (!)*
Excellent jet classification performance

[Komiske, Metodiev, JDT, [JHEP 2019](#); see also Komiske, Metodiev, JDT, [JHEP 2018](#); code at [energyflow.network](#); special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, [NIPS 2017](#)]



Energy Flow Networks

Architecture designed around symmetries and *interpretability*



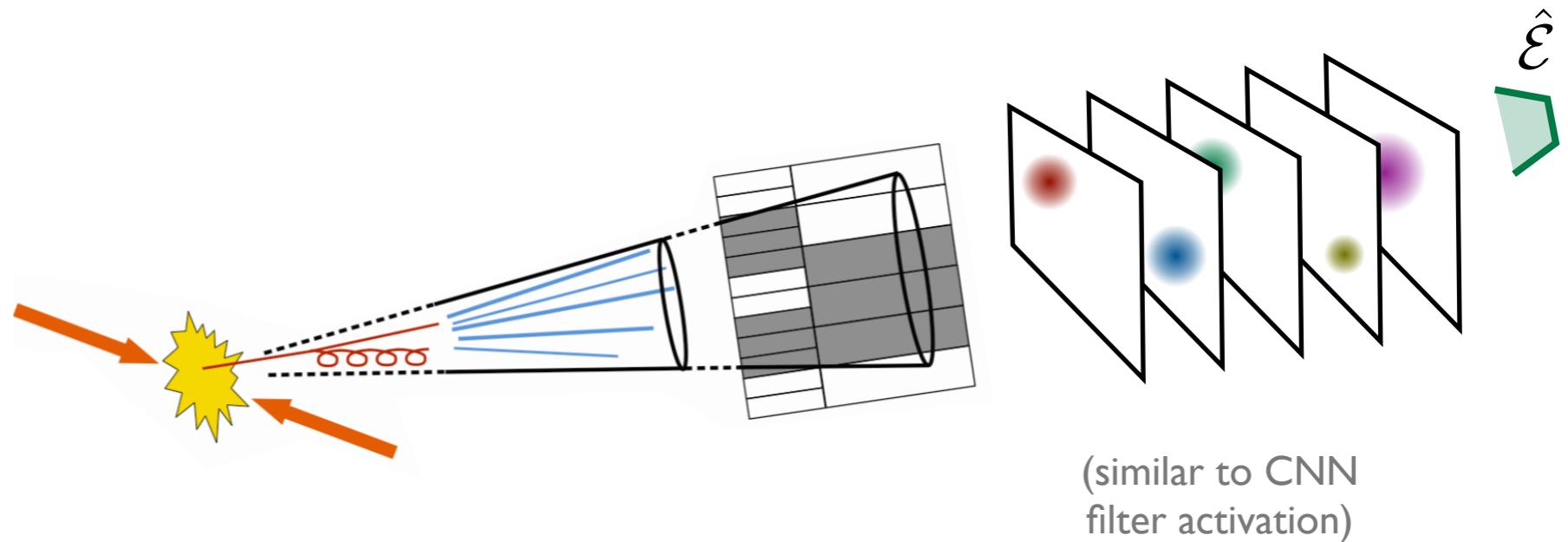
$$S(\mathcal{J}) = F(V_1, V_2, \dots, V_\ell)$$

Latent space of dim ℓ

$V_a(\mathcal{J}) = \sum_{i \in \mathcal{J}} E_i \Phi_a(\hat{n}_i)$

Can visualize if ℓ is small

Easy to plot!

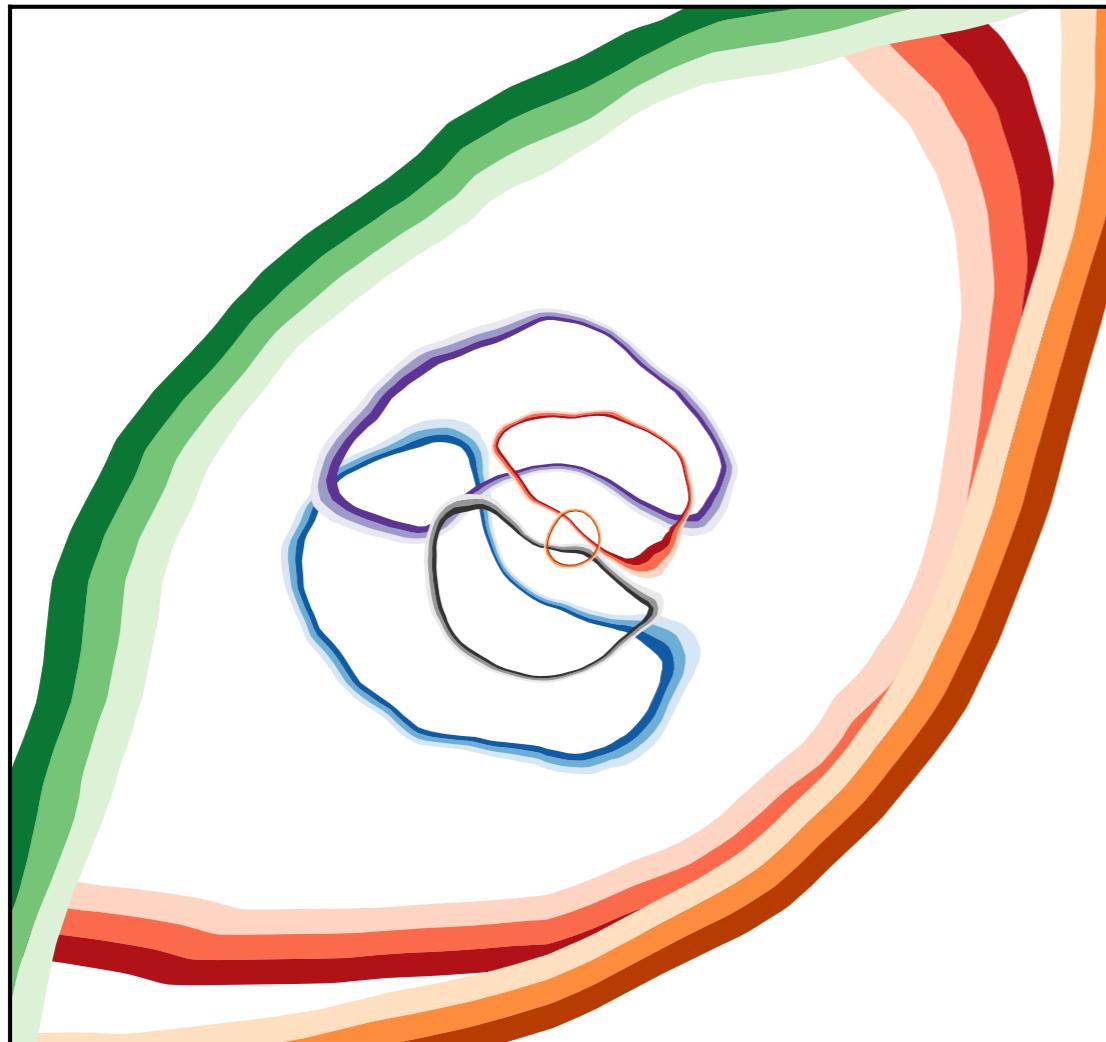


[Komiske, Metodiev, JDT, [JHEP 2019](#); see also Komiske, Metodiev, JDT, [JHEP 2018](#); code at [energyflow.network](#); special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, [NIPS 2017](#)]

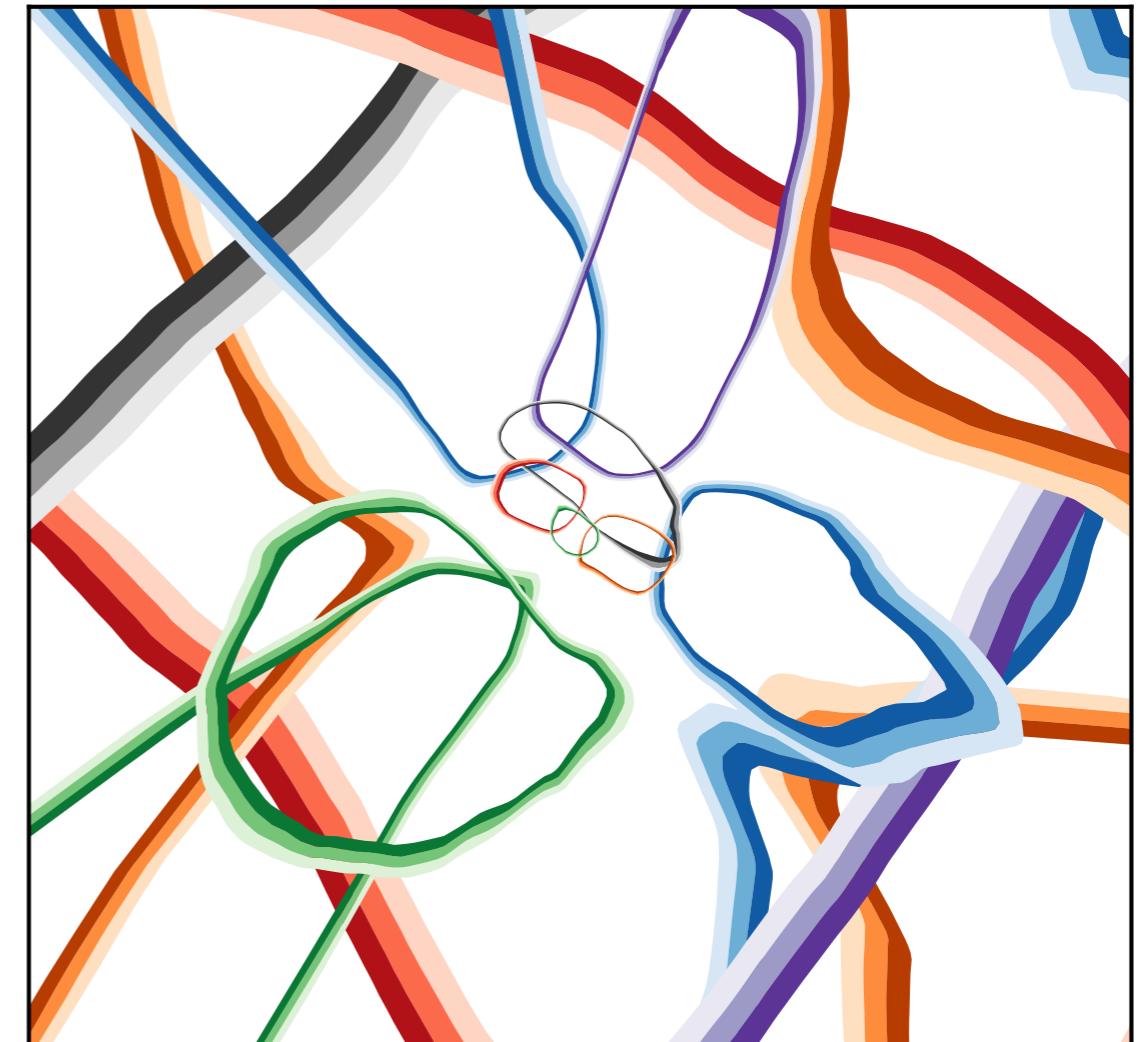


Psychedelic Network Visualization

Latent Dimension 8



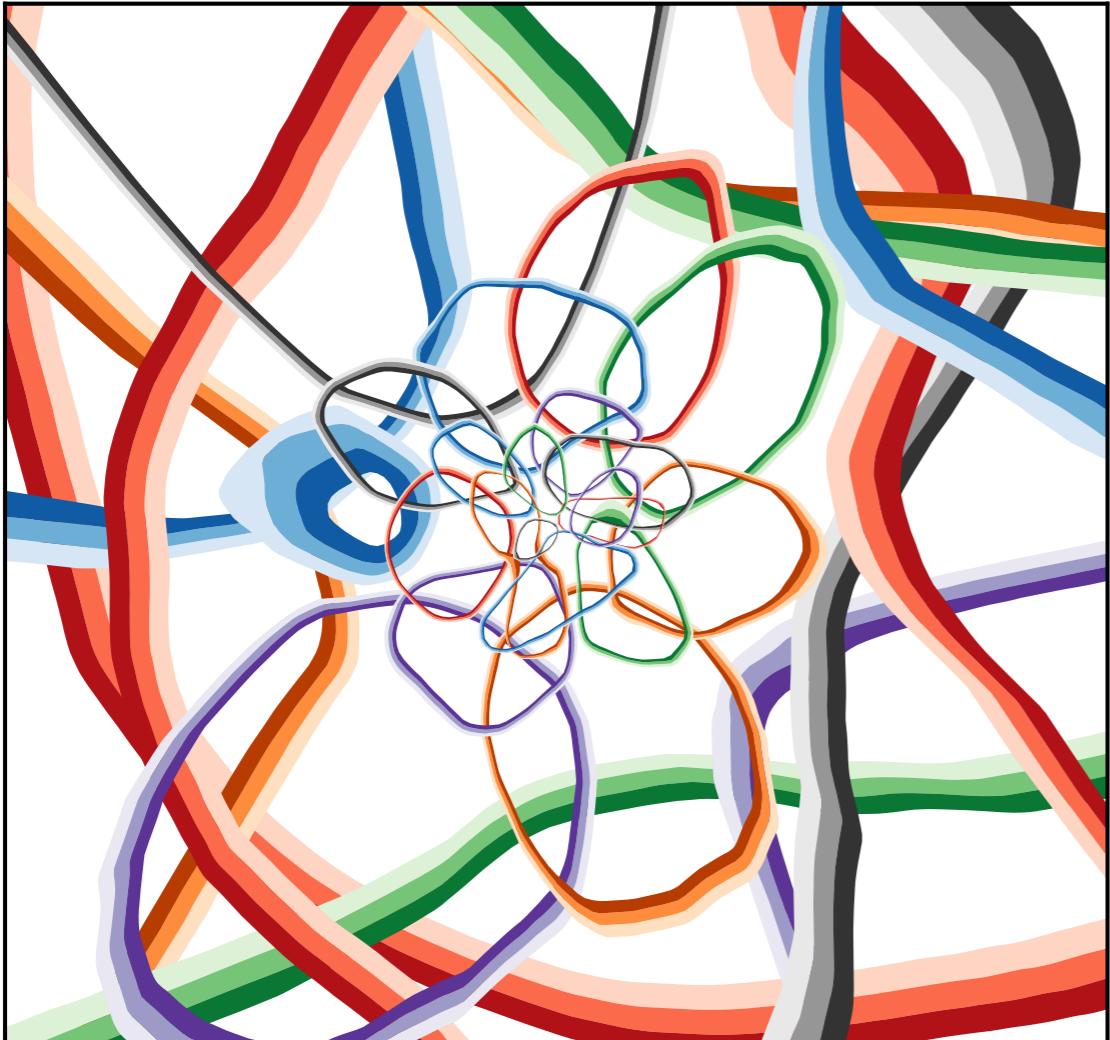
Latent Dimension 16



For the case of **quark** vs. **gluon** classification

Psychedelic Network Visualization

Latent Dimension 32

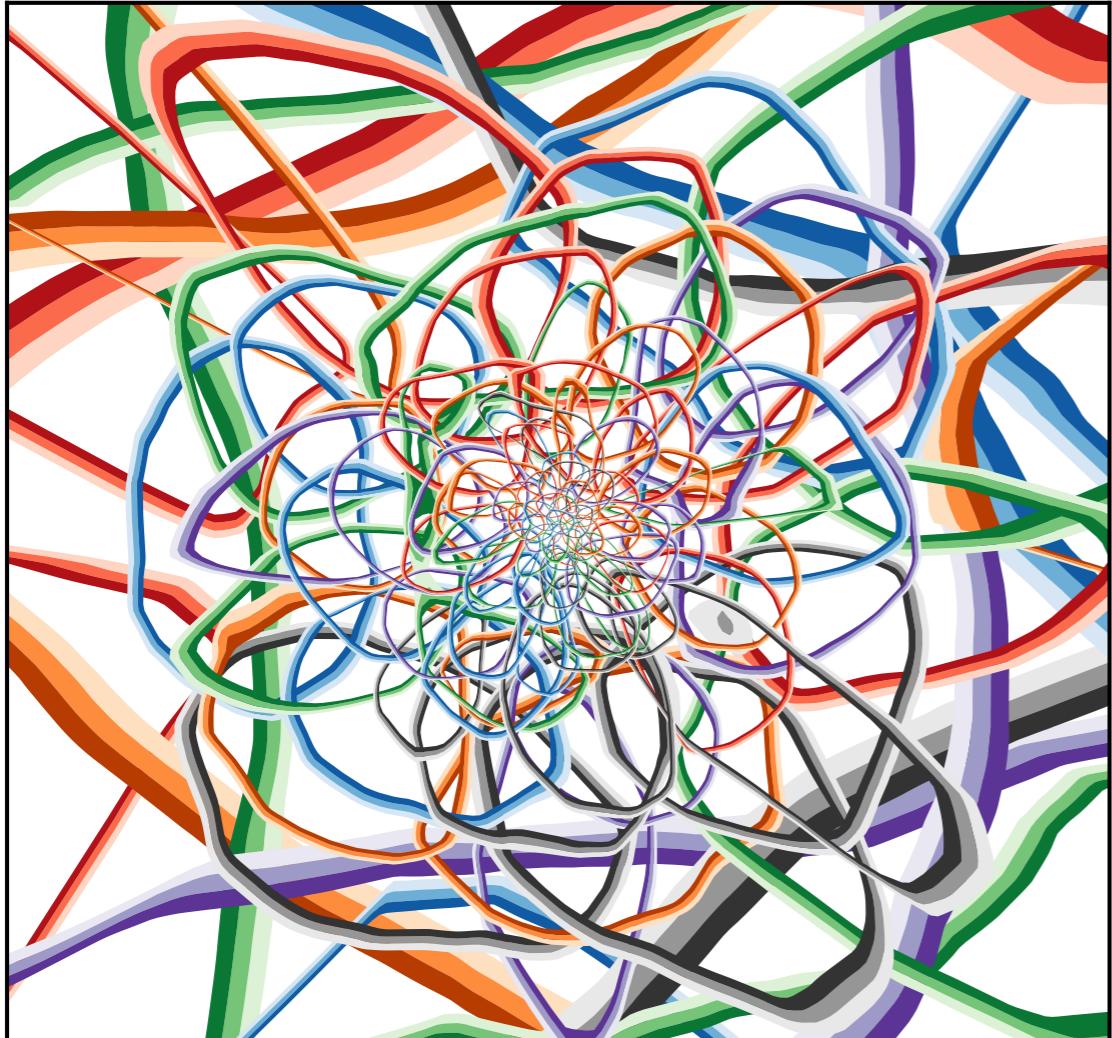


Latent Dimension 64

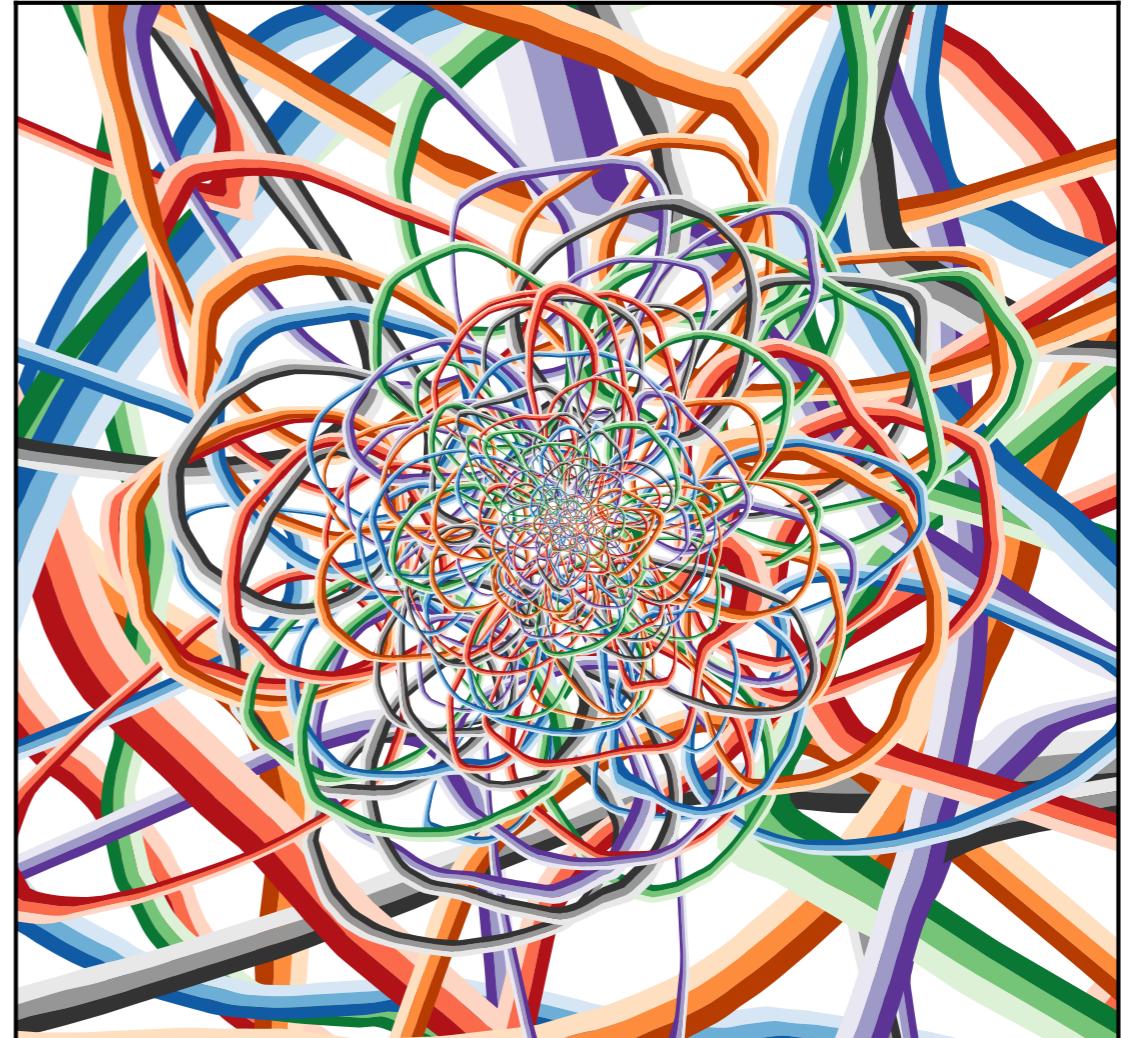


Psychedelic Network Visualization

Latent Dimension 128

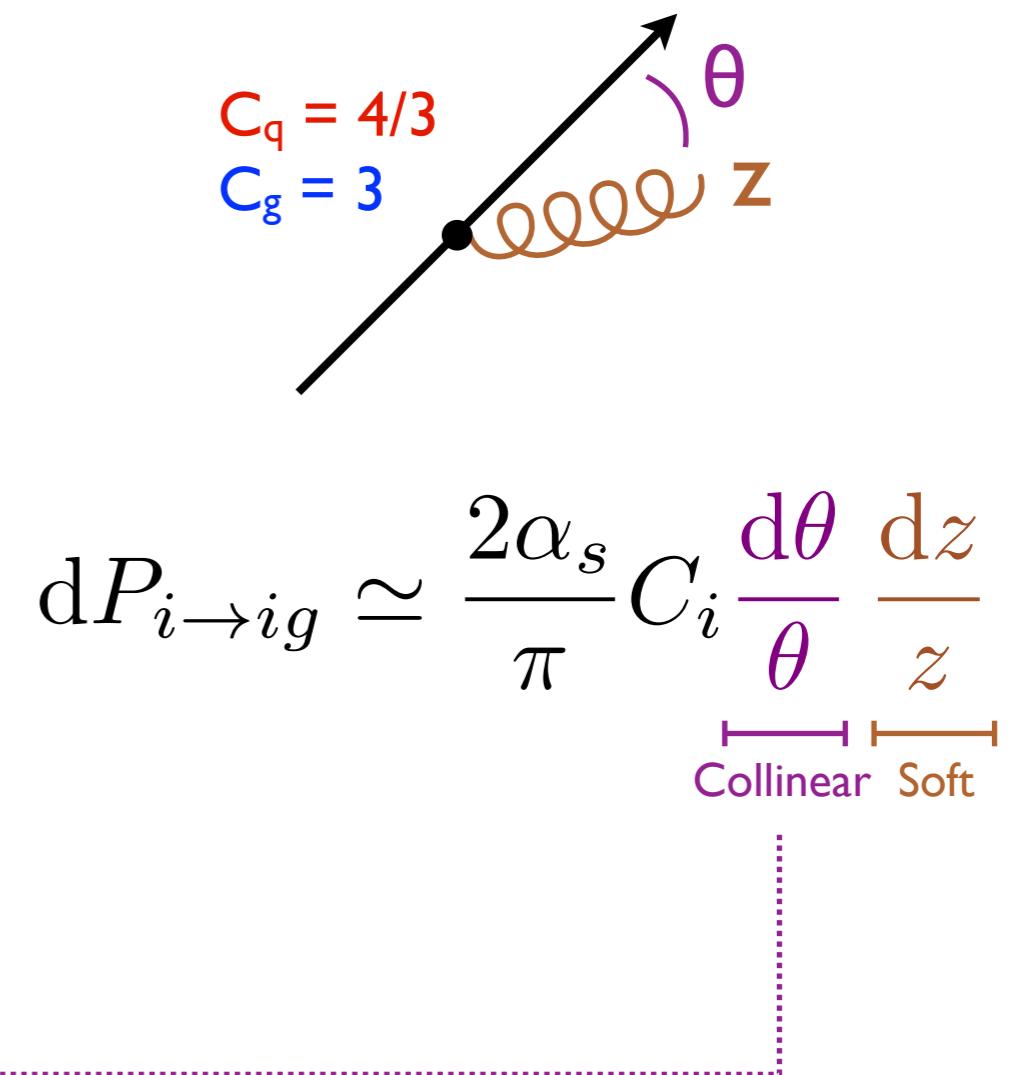
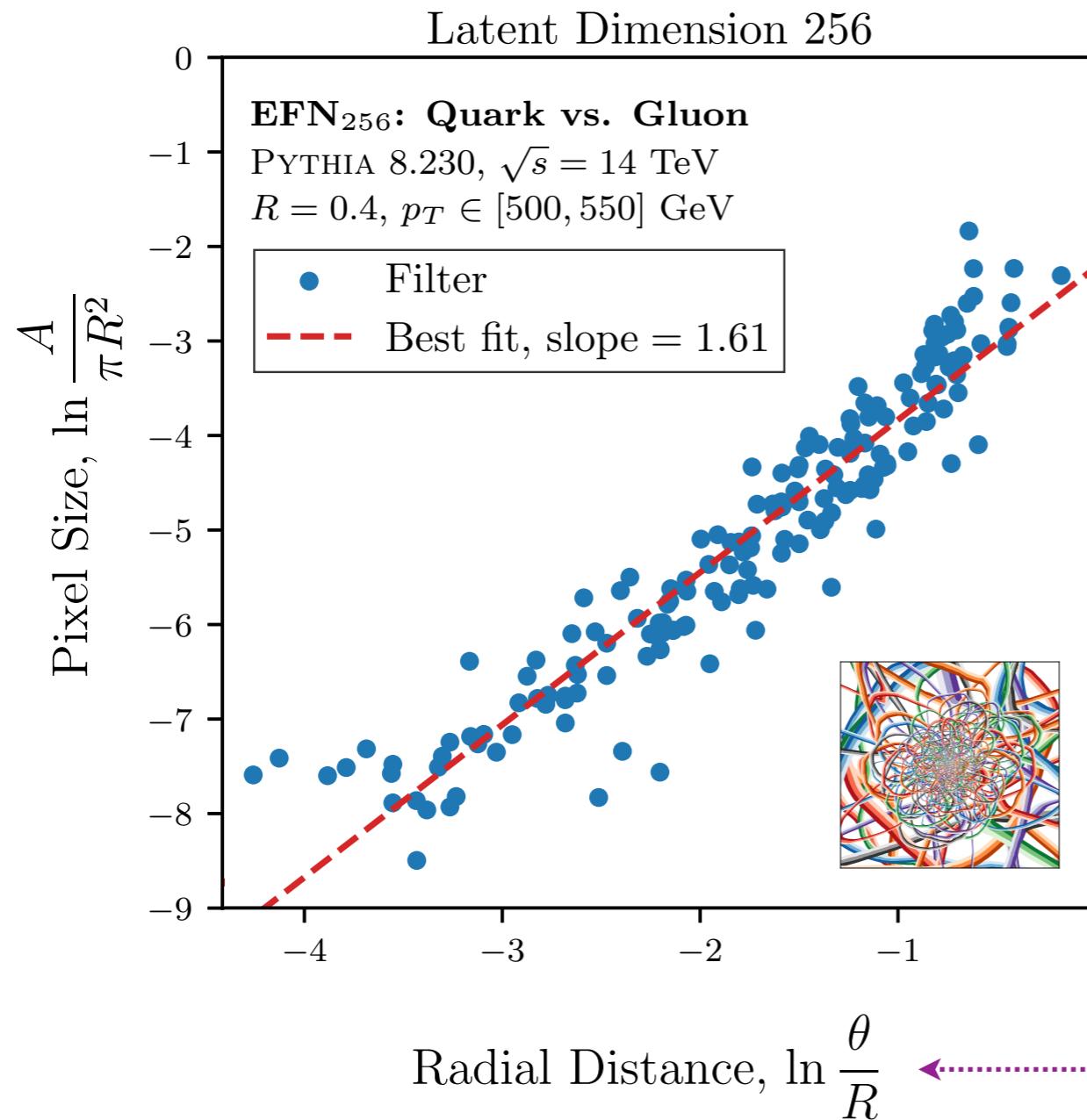
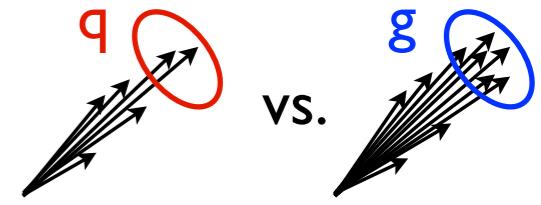


Latent Dimension 256



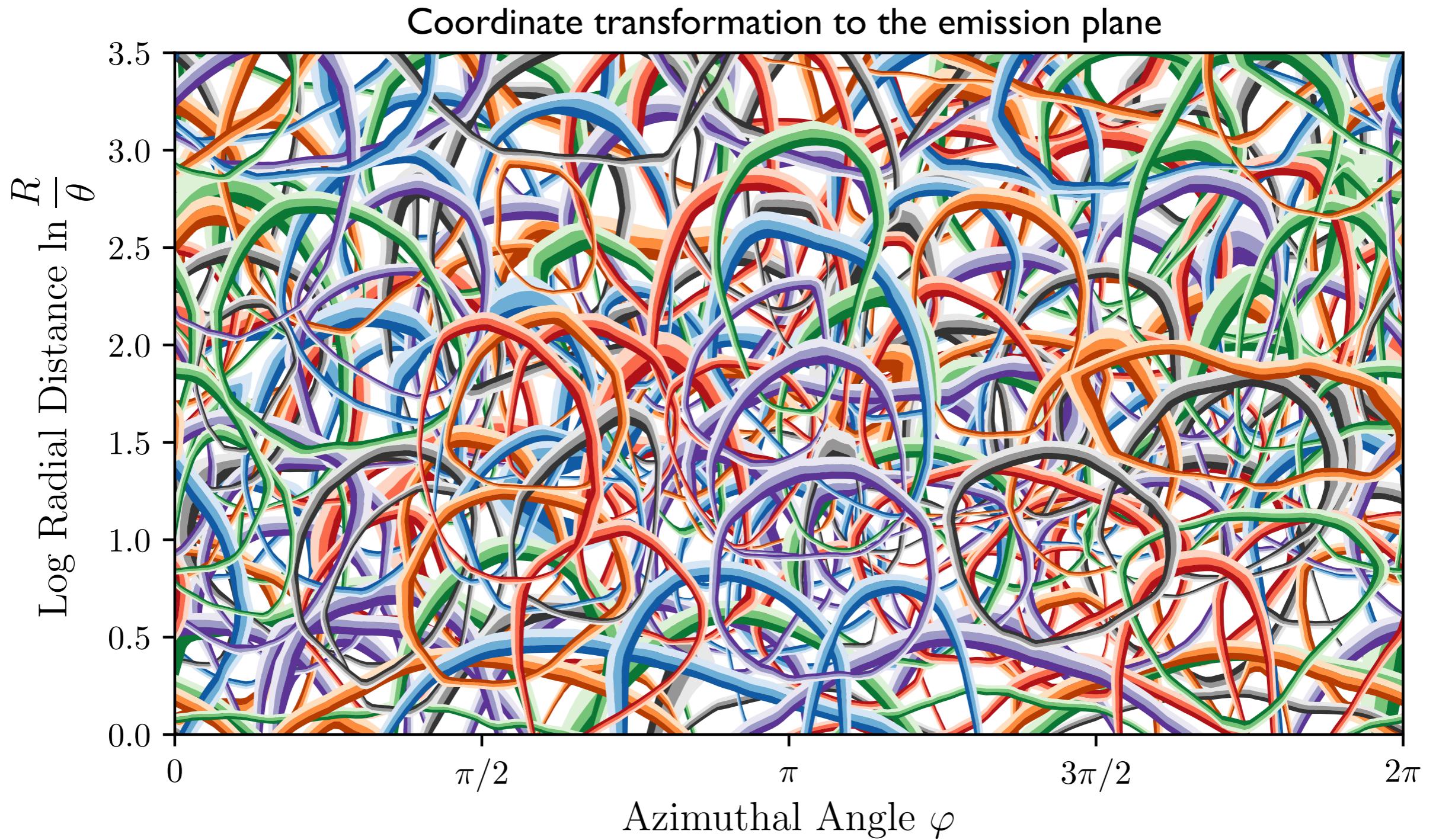
Singularity structure of QCD!

Putting the AI in Altarelli-Parisi



[Komiske, Metodiev, JDT, JHEP 2019]

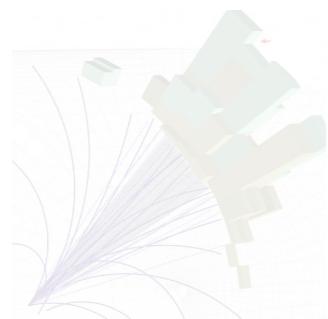
Ready for the Hirshhorn?



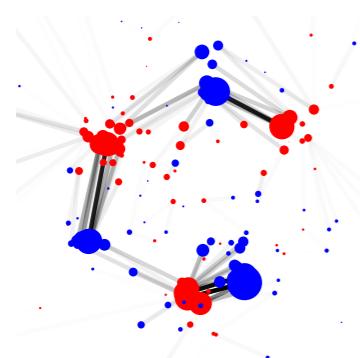
[Komiske, Metodiev, JDT, [JHEP 2019](#); see also Dreyer, Salam, Soyez, [JHEP 2018](#)]



Particle Physics 101



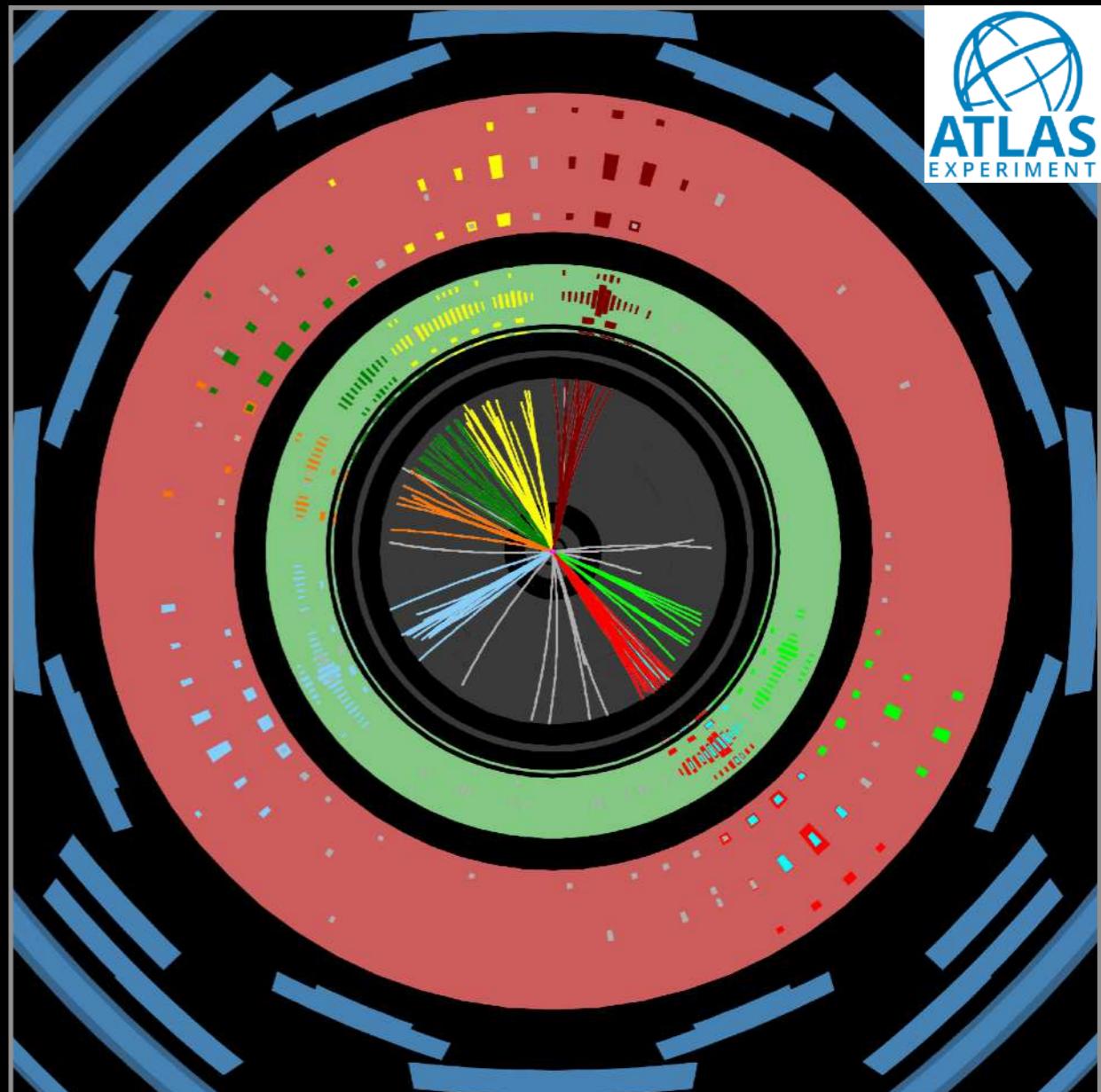
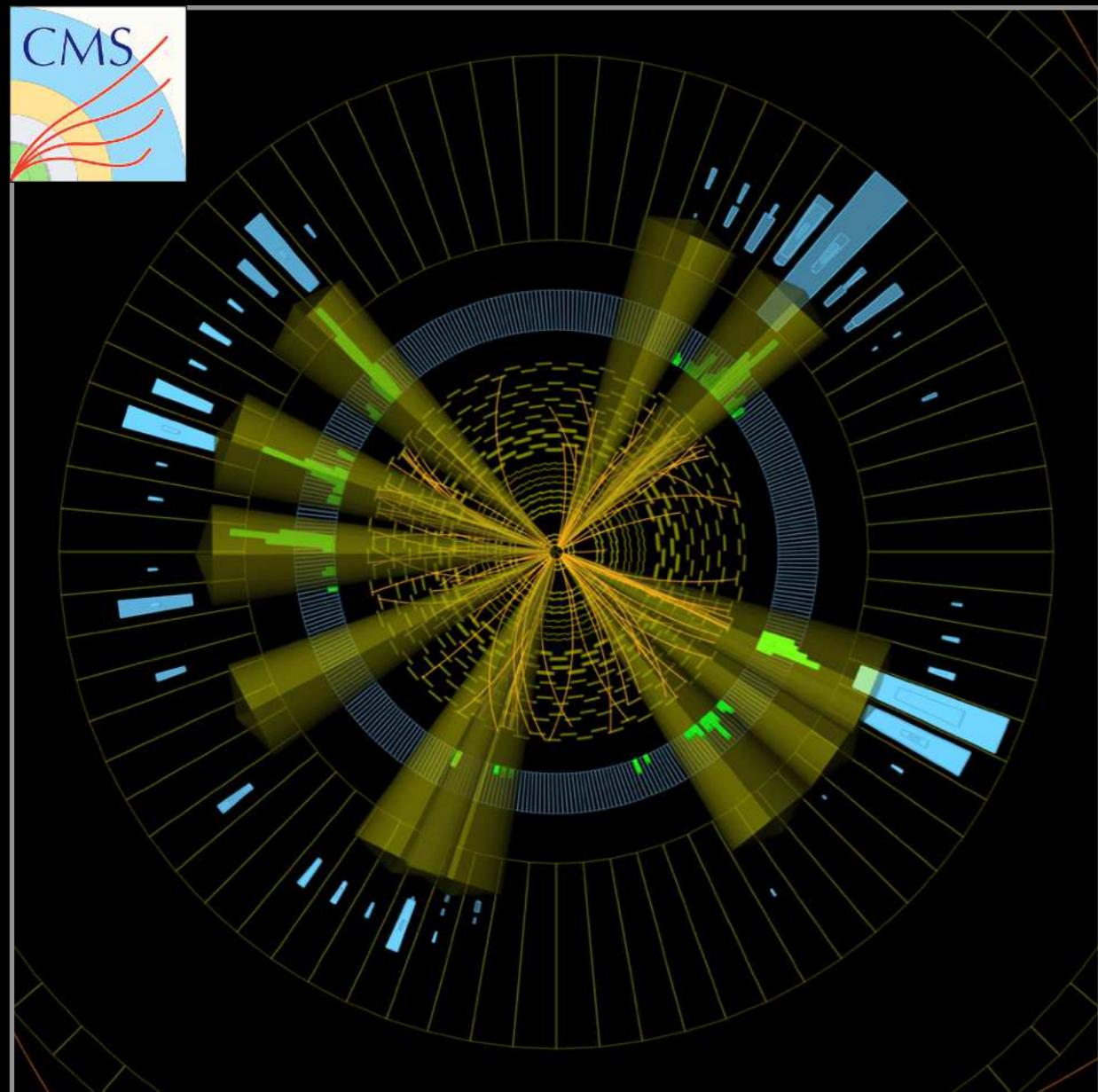
What is a Collider Event?



When are Collider Events Similar?

Two Collider Events

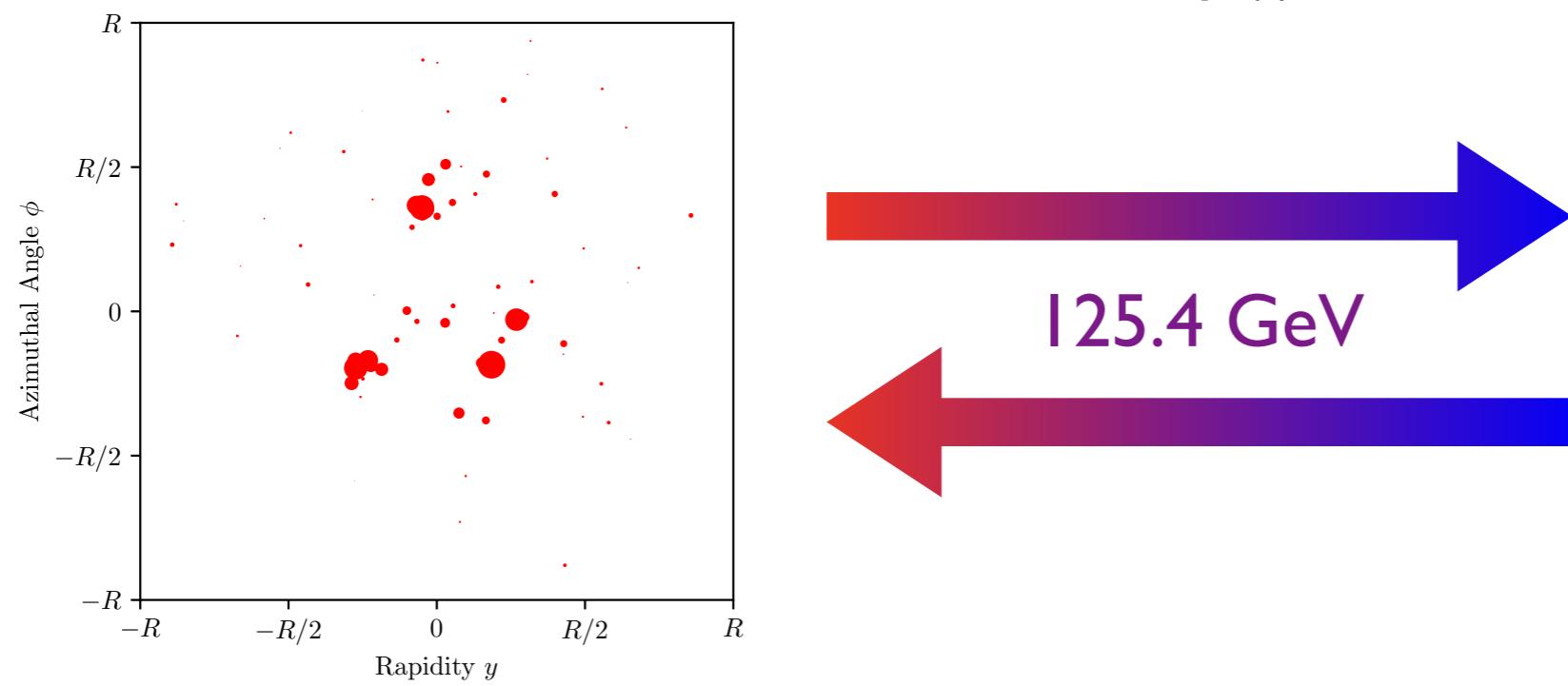
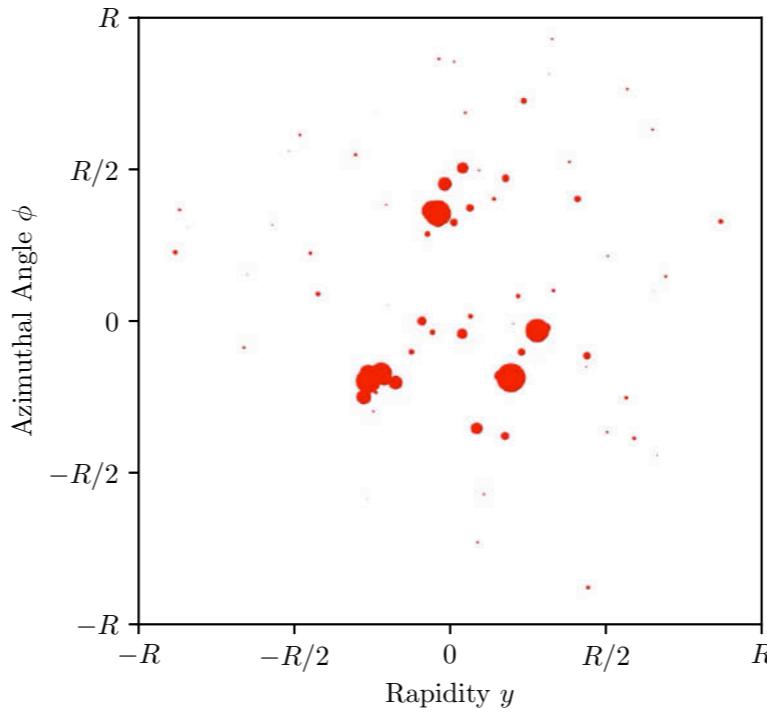
Two collections of points in (momentum) space



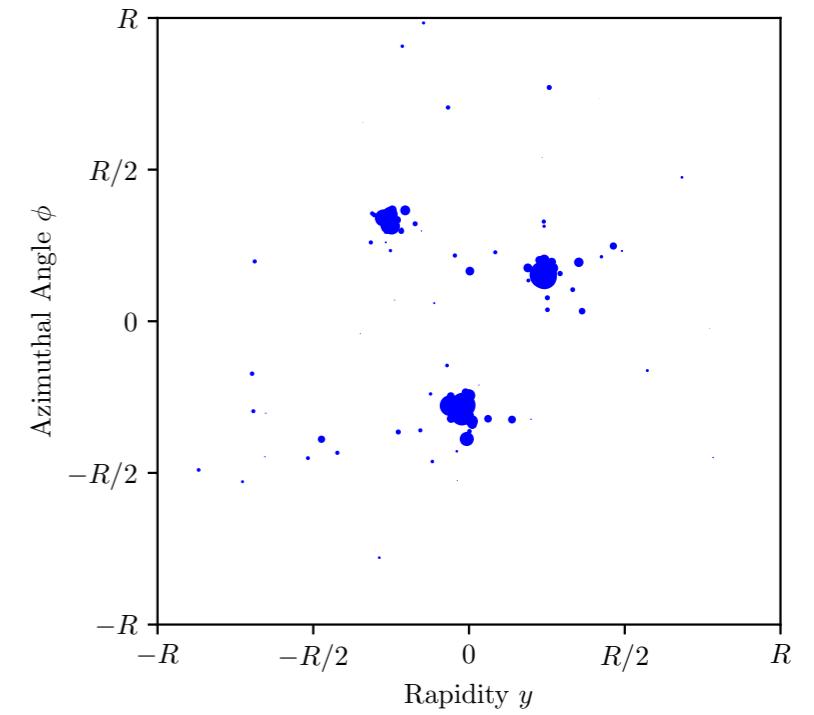
How “close” are these? (8.5 km?)

Similarity of Two Energy Flows?

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



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



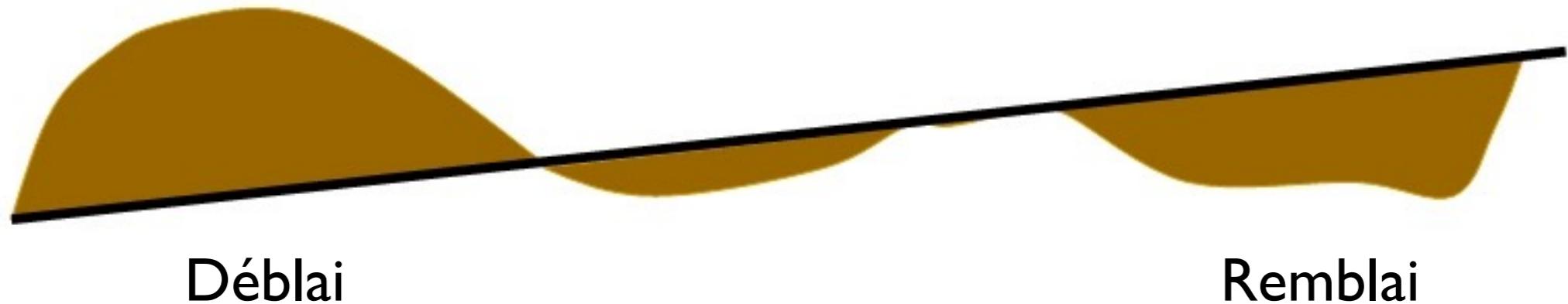
[Komiske, Metodiev, JDT, PRL 2019; code at Komiske, Metodiev, JDT, [energyflow.network](#)]

The Earth Mover's Distance

Optimal Transport:

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

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



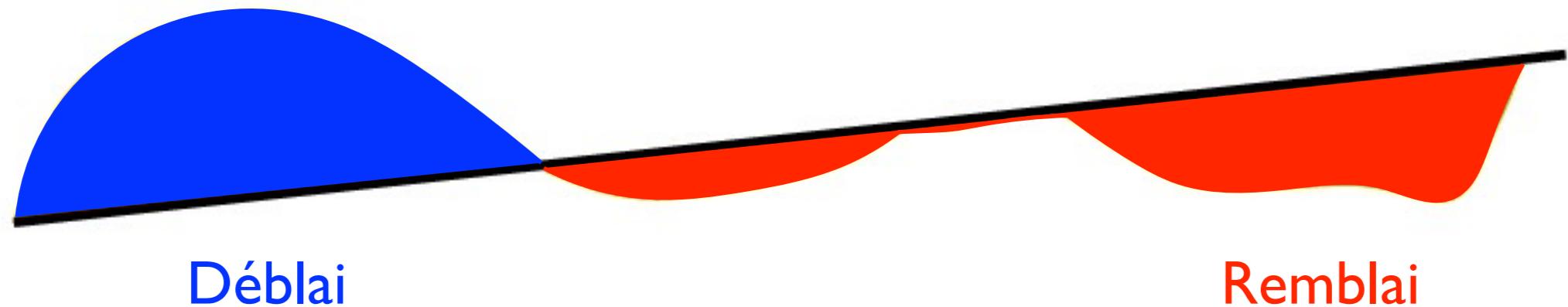
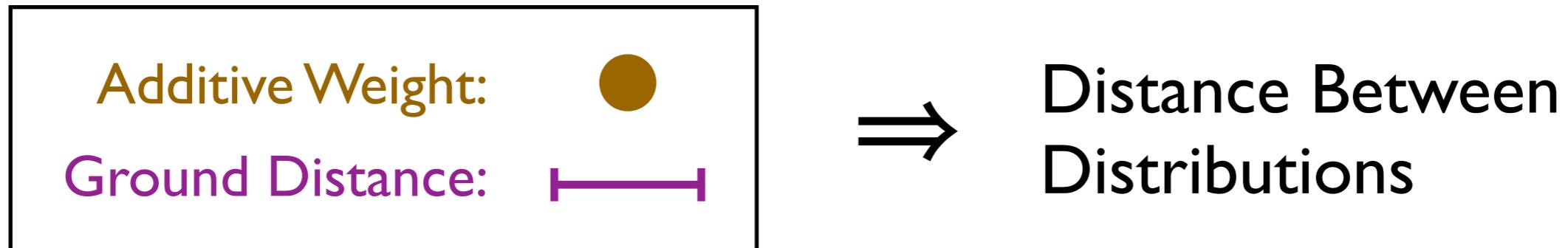
[h/t Niles-Weed, [ML4Jets 2020](#); Monge, 1781; Vaserštejn, 1969; [Wikipedia](#)]

The Earth Mover's Distance

Optimal Transport:

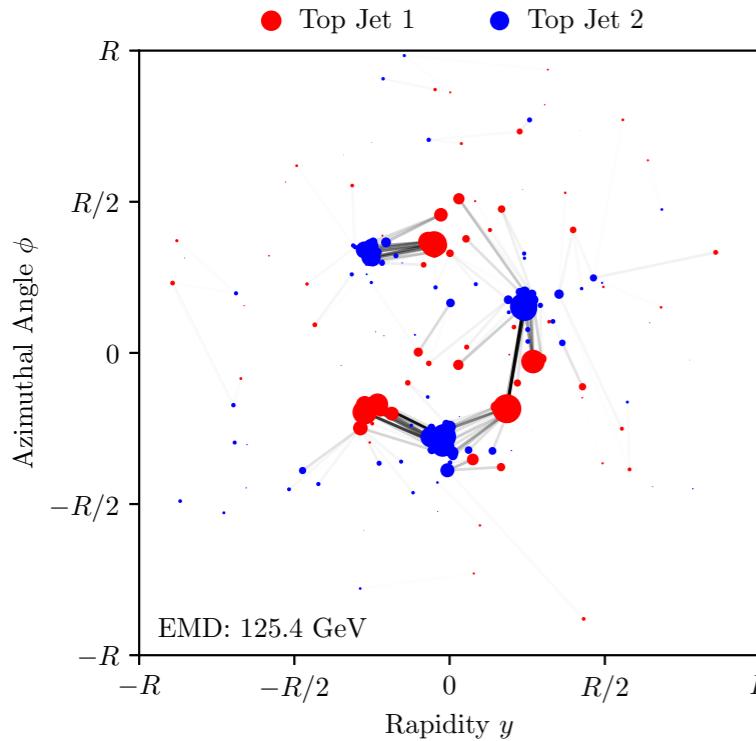
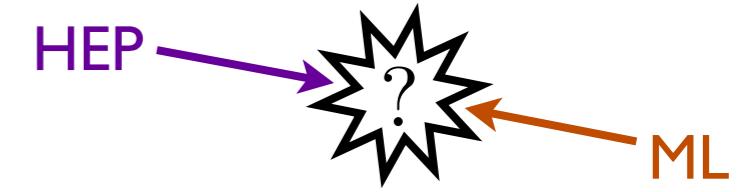
[Peleg, Werman, Rom, [IEEE 1989](#);
Rubner, Tomasi, Guibas, [ICCV 1998](#), [ICCV 2000](#);
Pele, Werman, [ECCV 2008](#); Pele Taskar, [GSI 2013](#)]

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



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

The Energy Mover's Distance

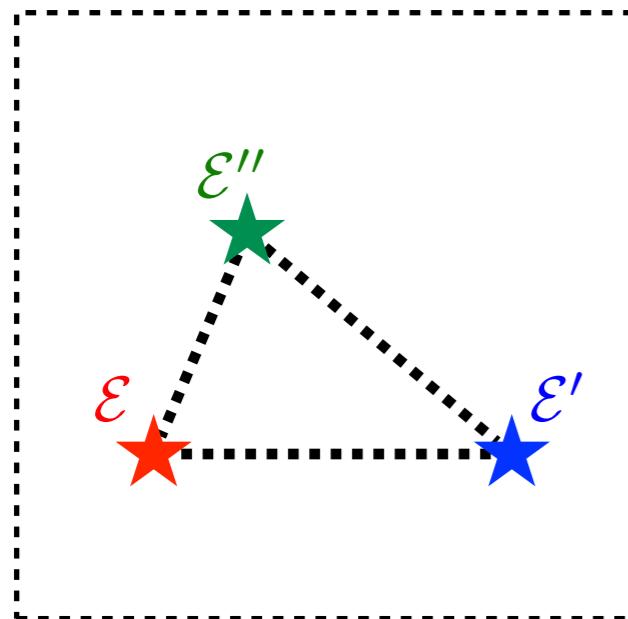
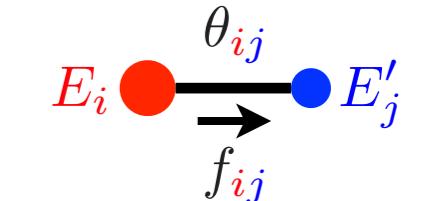


Optimal transport of energy flows...

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

↑
in GeV

Cost to move energy Cost to create energy



...defines a metric on the space of events

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

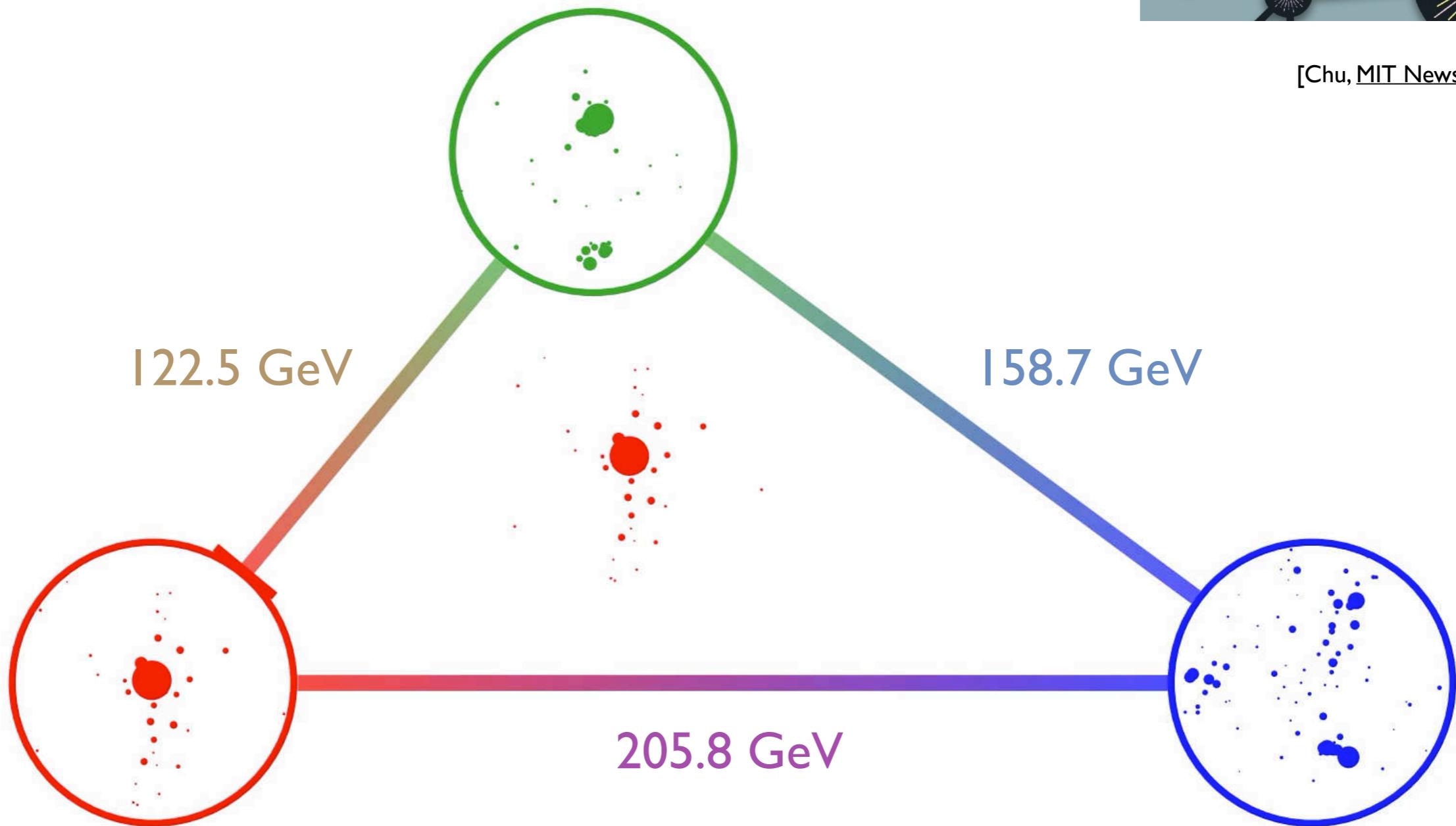
(assuming $R \geq \theta_{\max}/2$, i.e. $R \geq$ jet radius for conical jets)

[Komiske, Metodiev, JDT, [PRL 2019](#);
see also Pele, Werman, [ECCV 2008](#); Pele, Taskar, [GSI 2013](#);
see flavored variant in Crispim Romão, Castro, Milhano, Pedro, Vale, [arXiv 2020](#)]

Triangulating the Space of Jets



[Chu, MIT News July 2019]



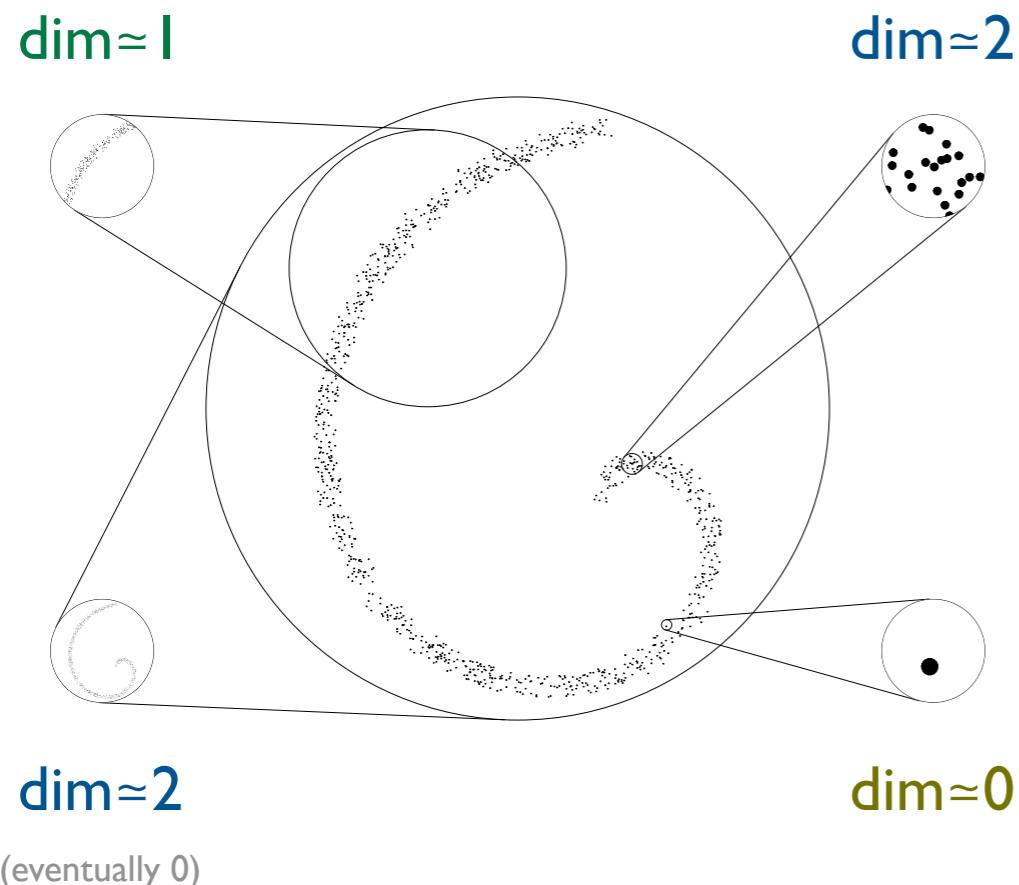
[Komiske, Metodiev, JDT, PRL 2019; code at Komiske, Metodiev, JDT, [energyflow.network](#)]

Dimensionality of Space of Jets

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

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

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



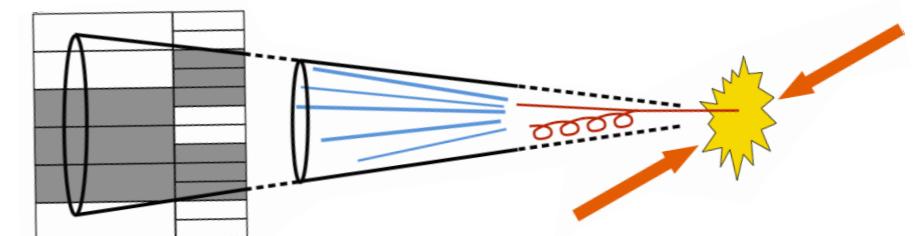
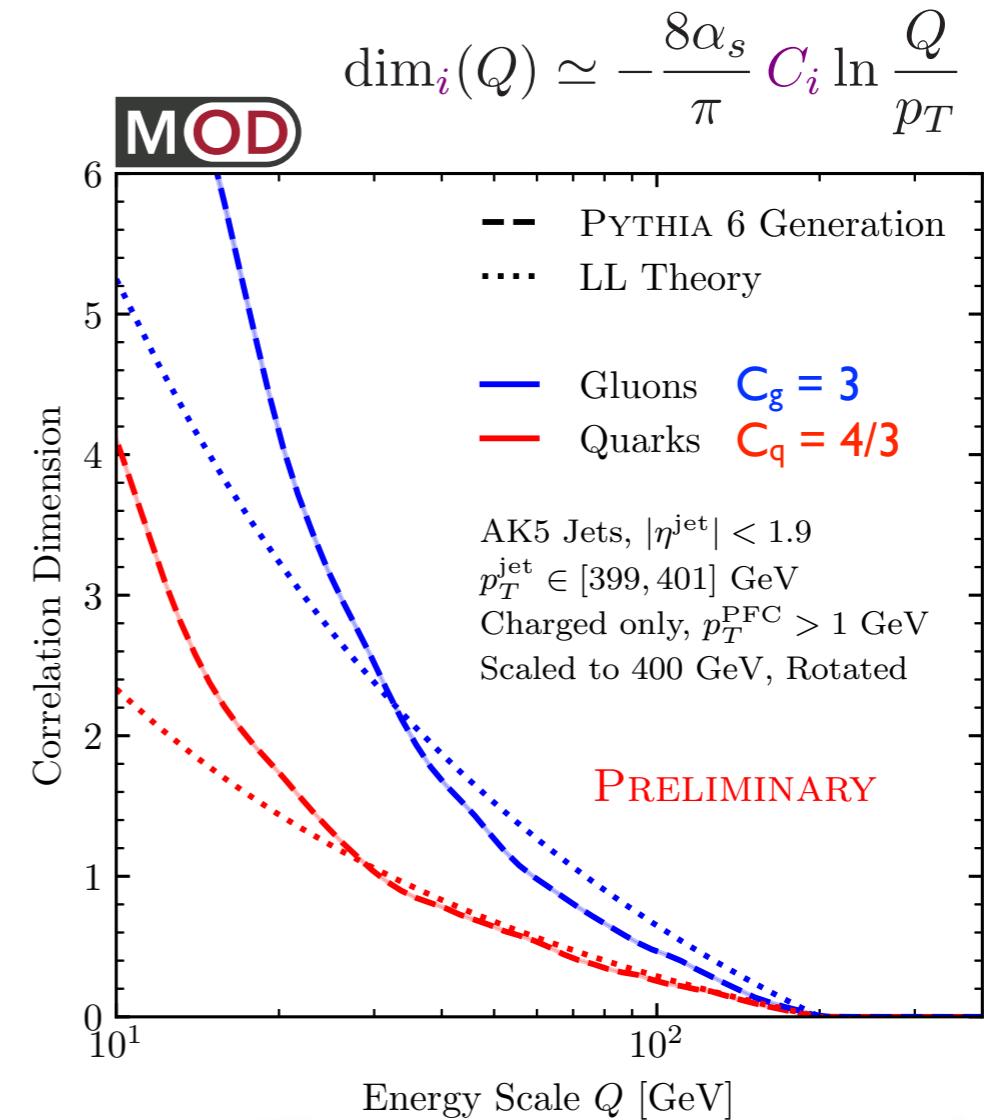
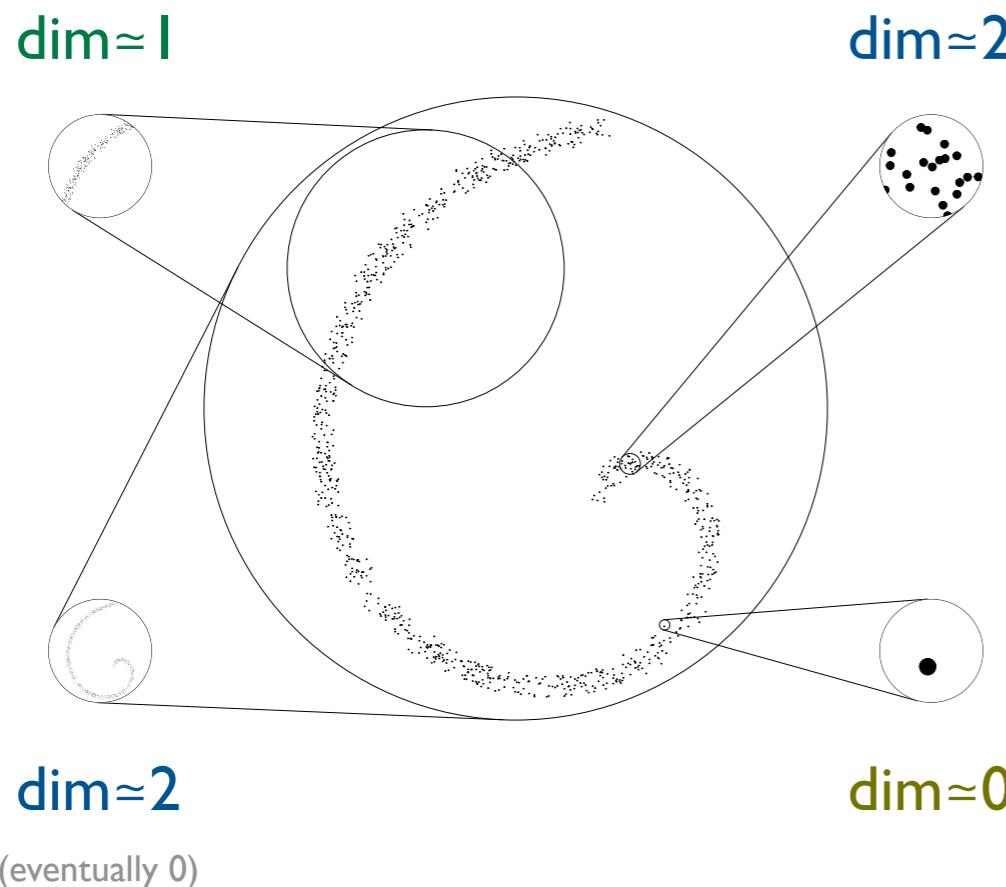
Dimensionality of Space of Jets



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

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

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



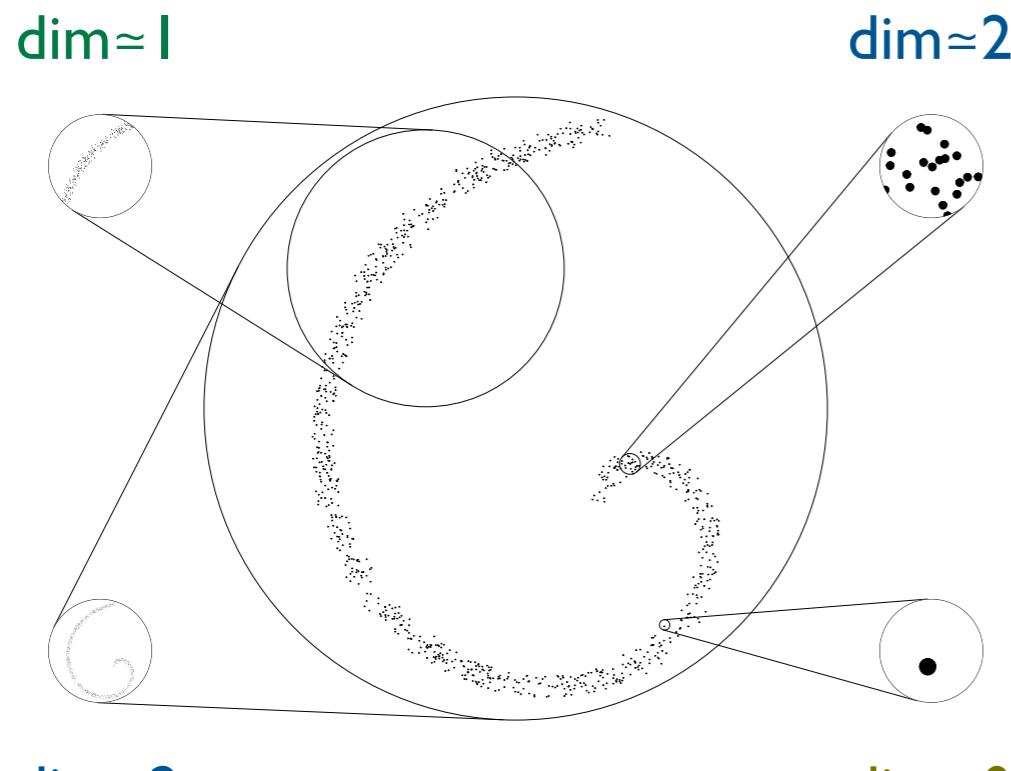
Dimensionality of Space of Jets



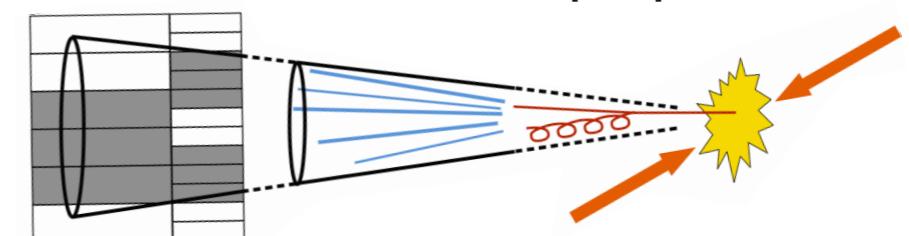
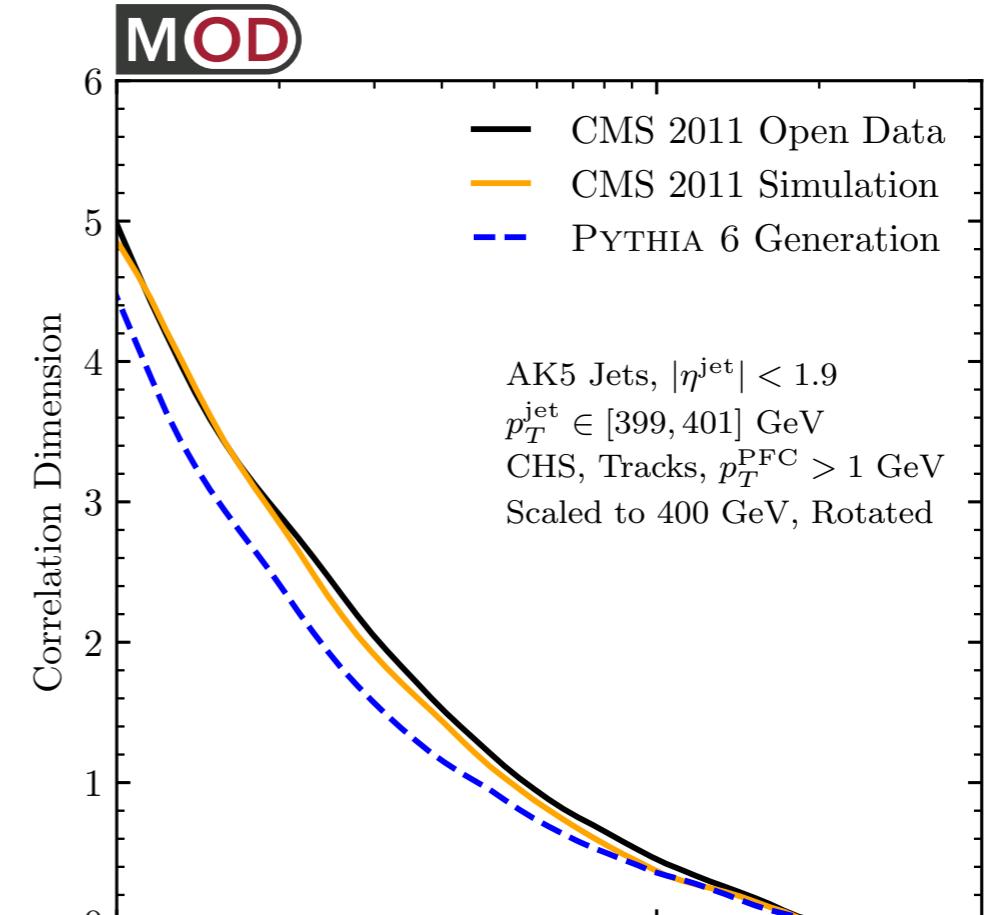
$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

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

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



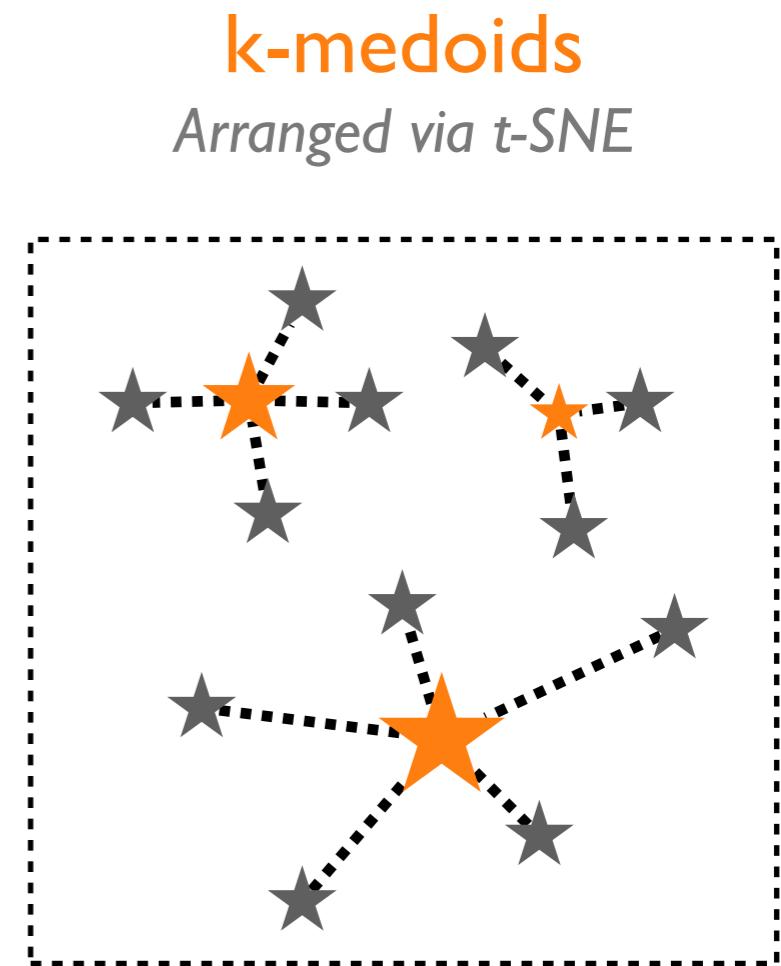
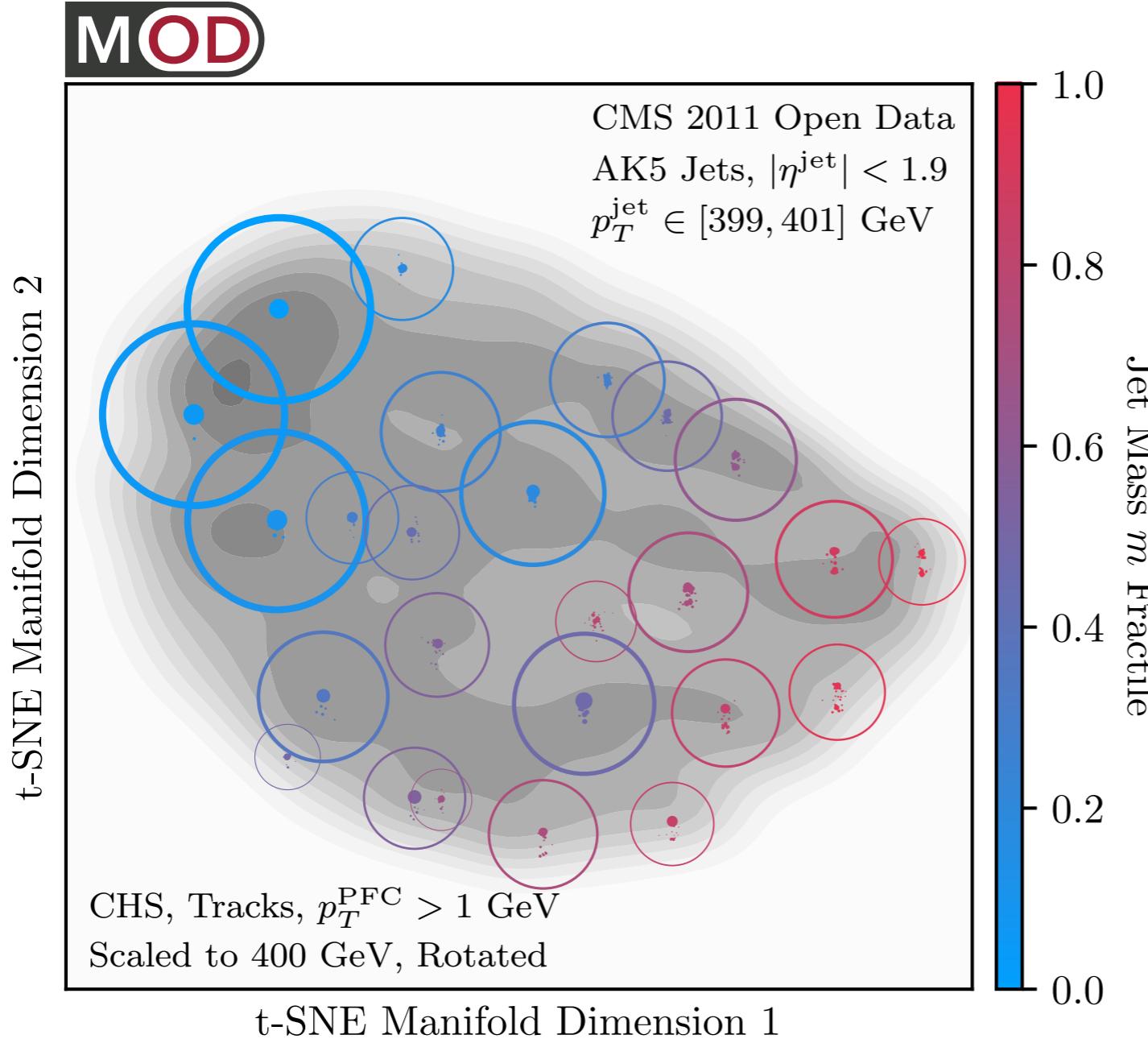
(eventually 0)



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

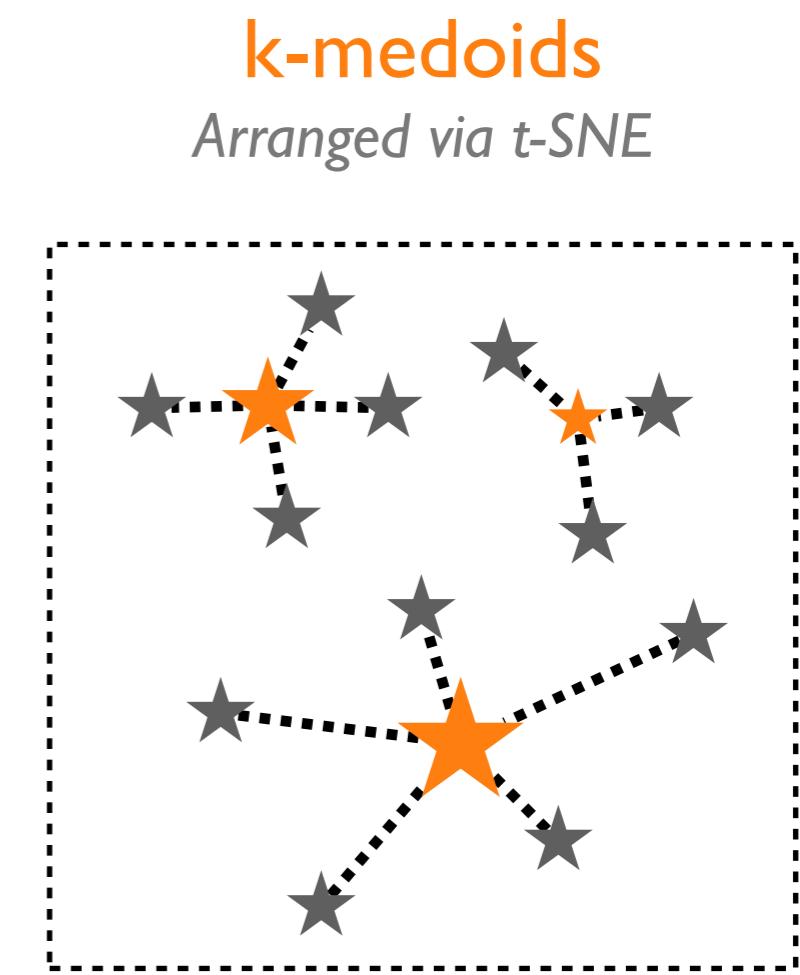
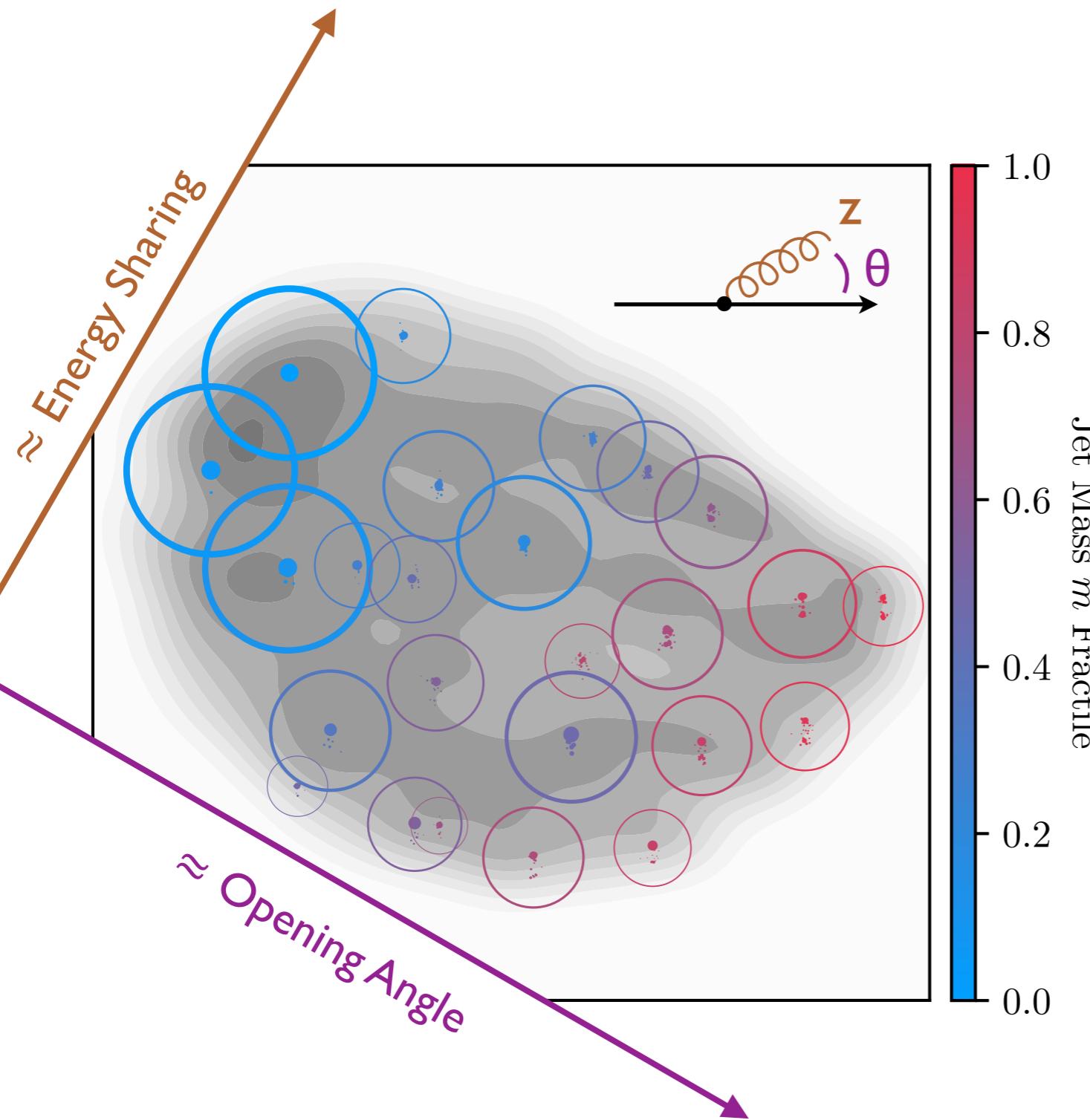


Most Representative Jets



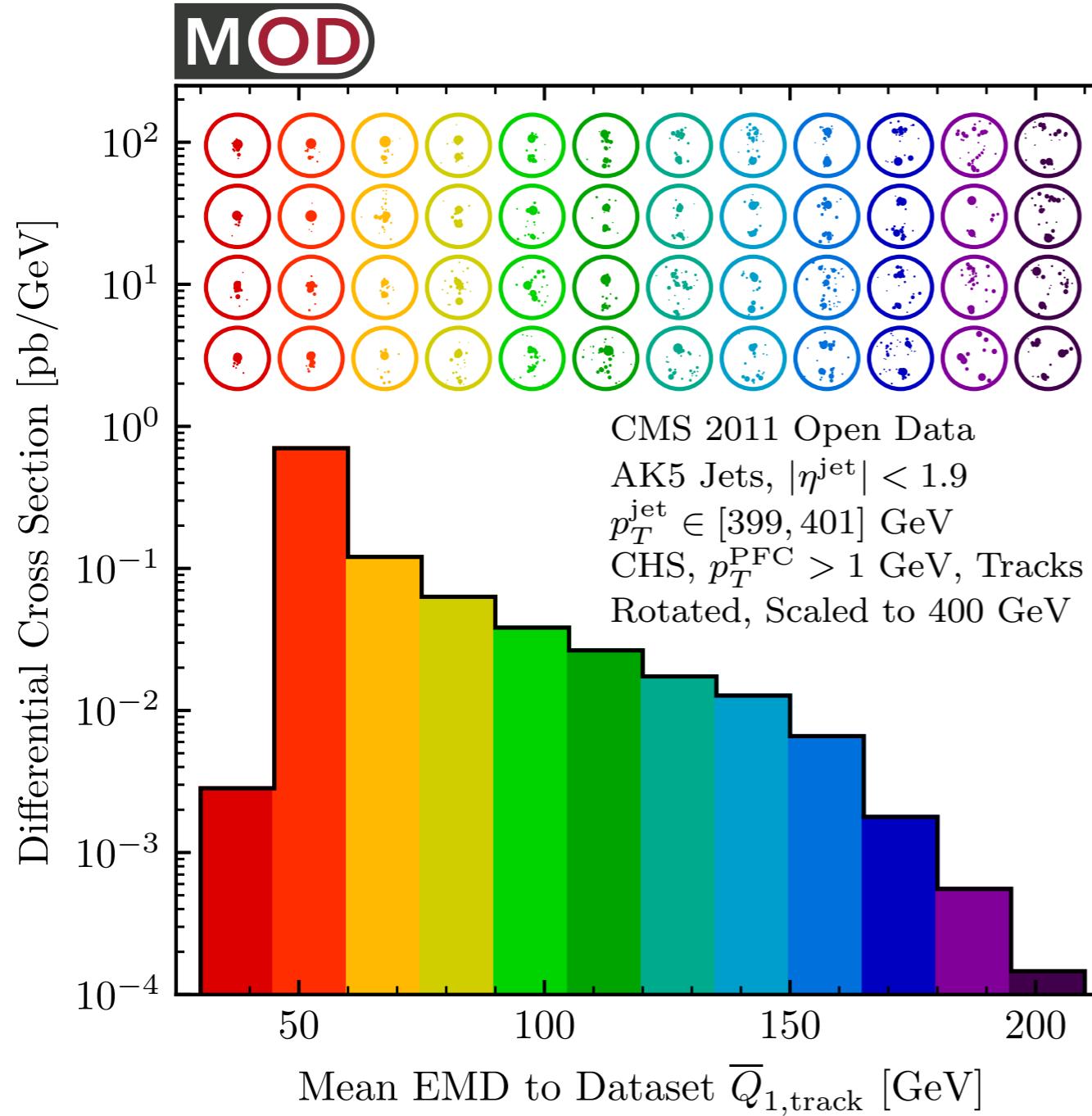
[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020; using van der Maaten, Hinton, JMLR 2008]

Most Representative Jets

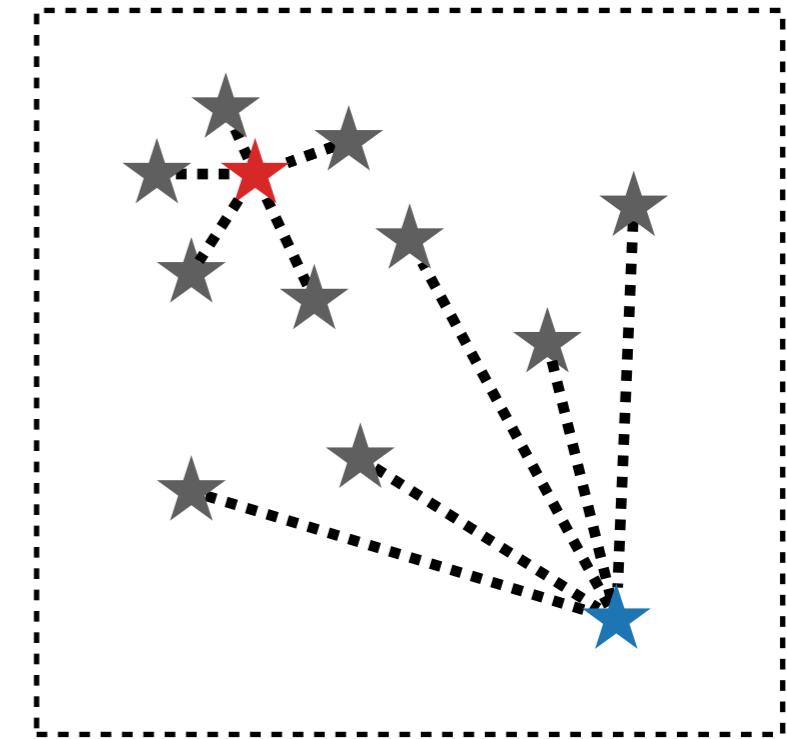


[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020; using van der Maaten, Hinton, JMLR 2008]

Least Representative Jets



New Physics?
Or tails of QCD?

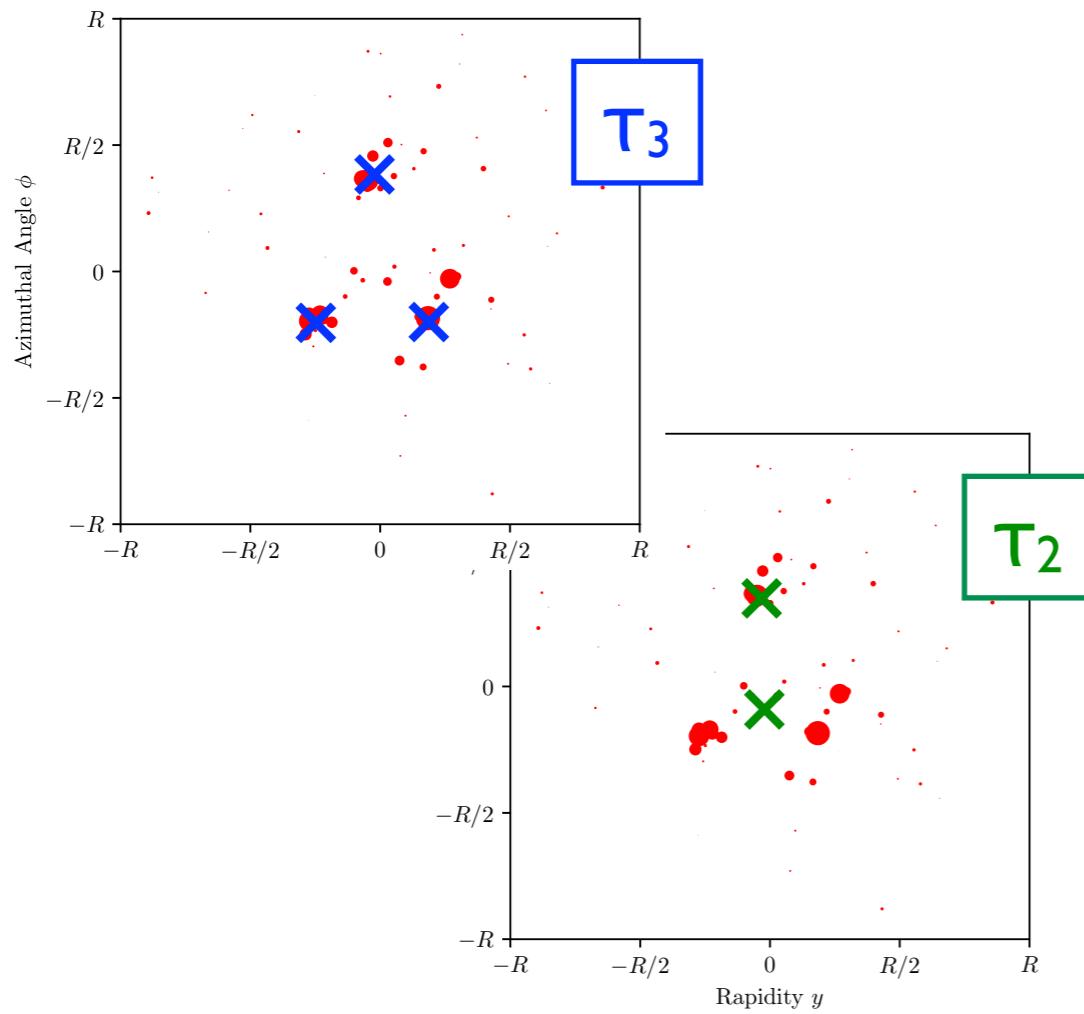


[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#)]

N-subjettiness

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

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

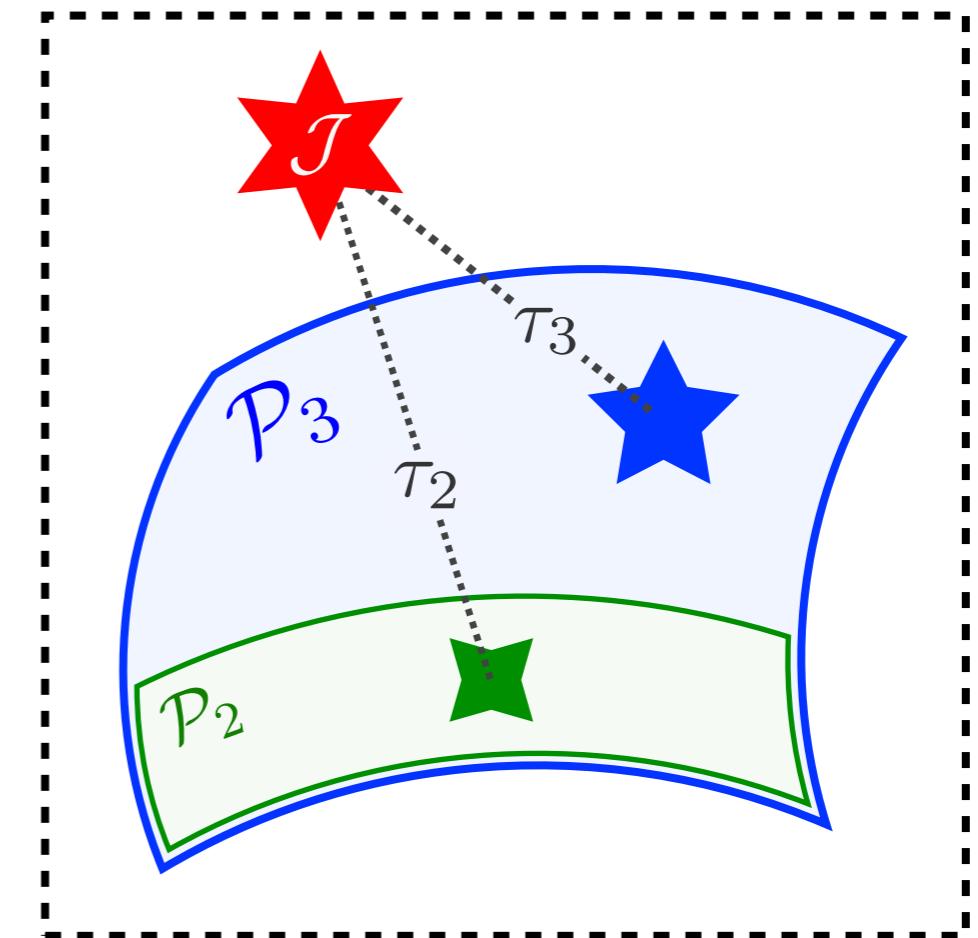
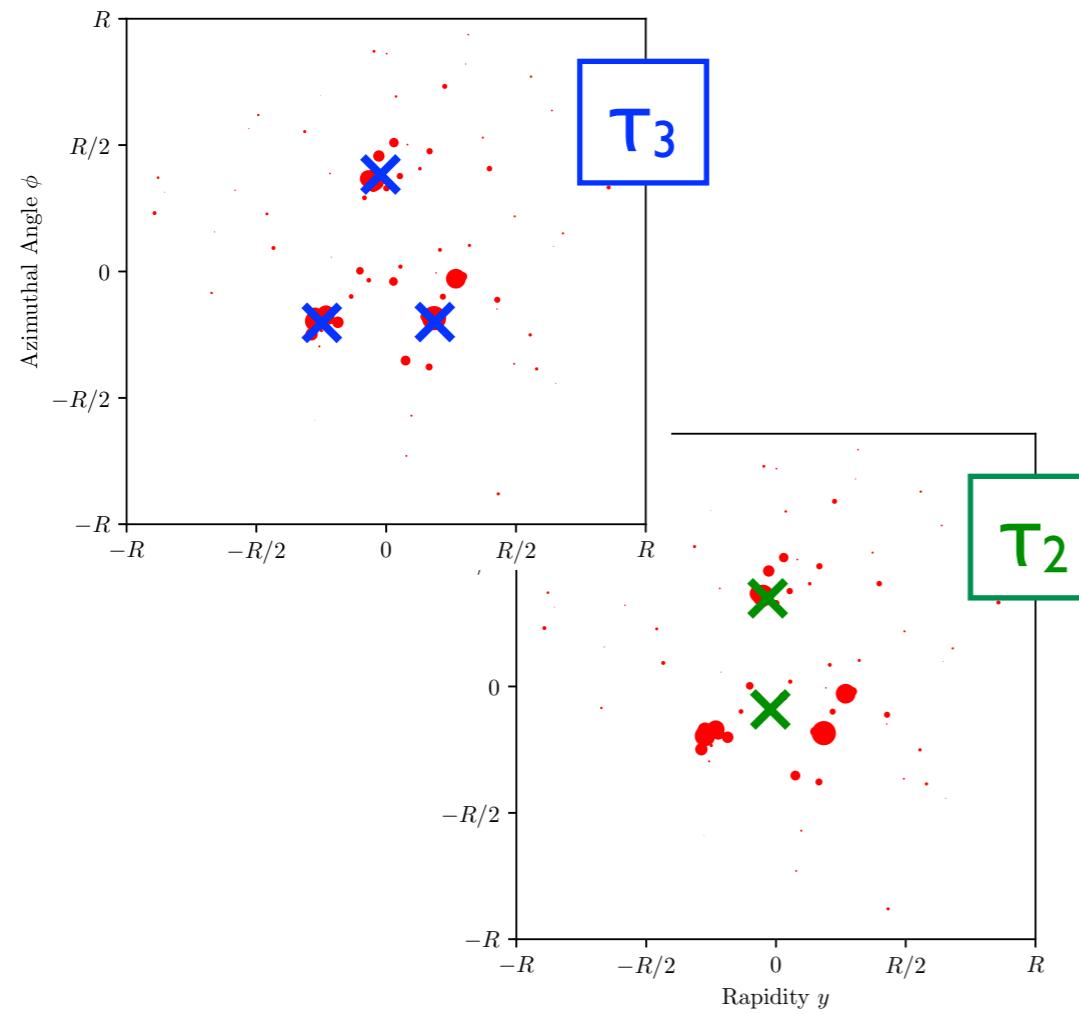


[JDT, Van Tilburg, [JHEP 2011](#), [JHEP 2012](#);
based on Brandt, Dahmen, [ZPC 1979](#); Stewart, Tackmann, Waalewijn, [PRL 2010](#)]

N-subjettiness = Point to Manifold EMD

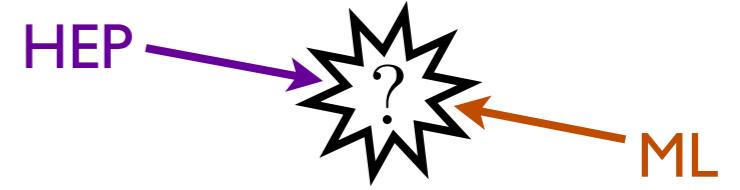
...is secretly an optimal transport problem

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

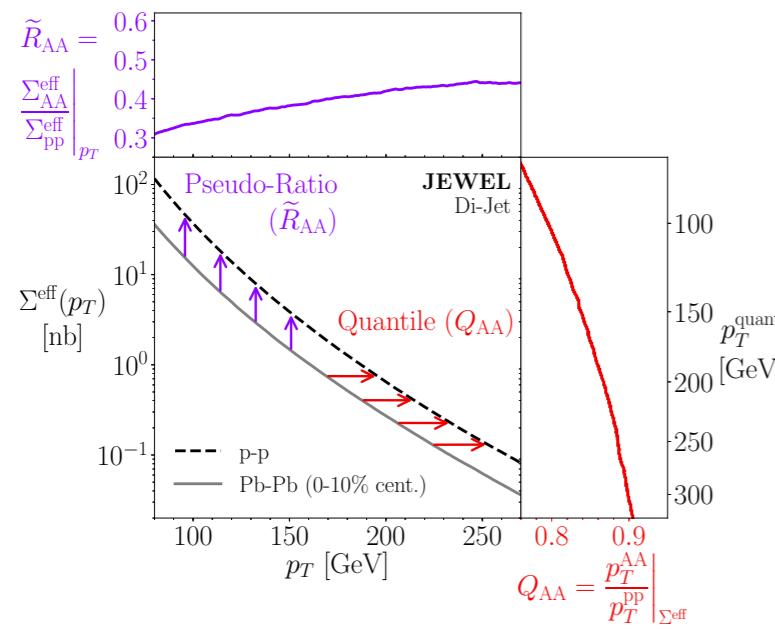


[JDT, Van Tilburg, [JHEP 2011](#), [JHEP 2012](#);
rephrased via Komiske, Metodiev, JDT, [JHEP 2020](#); see opposite limit in Cesarotti, JDT, [JHEP 2020](#)]

More Collisions

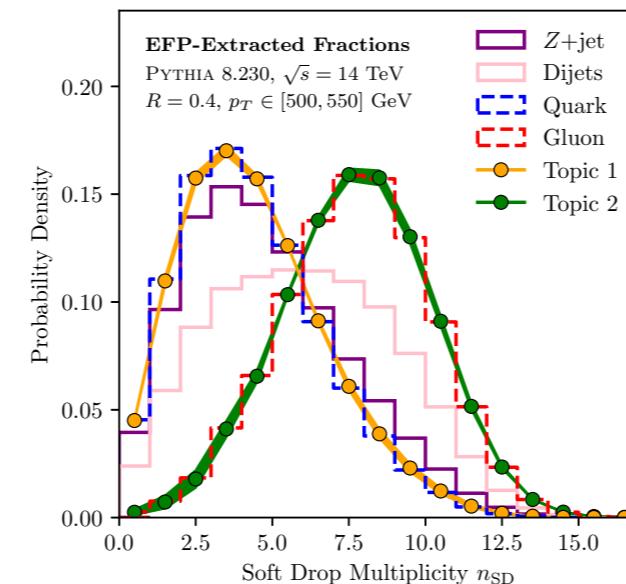


Jet Quenching via Optimal Transport



[Brewer, Milhano, JDT, PRL 2019]

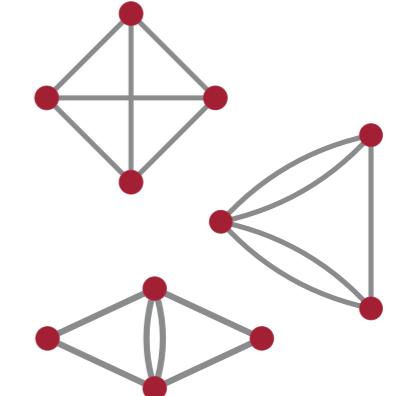
Quark/Gluon Definitions via Blind Source Separation



[Komiske, Metodiev, JDT, JHEP 2018;
Brewer, JDT, Turner; arXiv 2020]

Kinematic Decomposition via Graph Theory

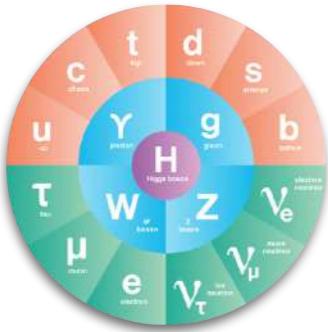
Edges d	Leafless Multigraphs		
	Connected	All	A307316
1	0	0	
2	1	1	
3	2	2	
4	4	5	
5	9	11	
6	26	34	
7	68	87	
8	217	279	
9	718	897	
10	2 553	3 129	
11	9 574	11 458	
12	38 005	44 576	
13	157 306	181 071	
14	679 682	770 237	
15	3 047 699	3 407 332	
16	14 150 278	15 641 159	



[Komiske, Metodiev, JDT,
JHEP 2018, PRD 2020]

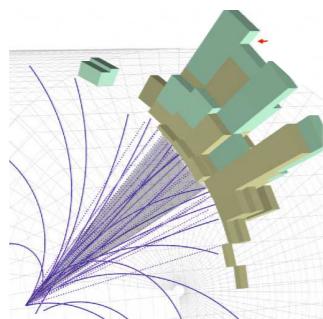
New insights into particle physics*
facilitated by advances in machine learning*

Summary



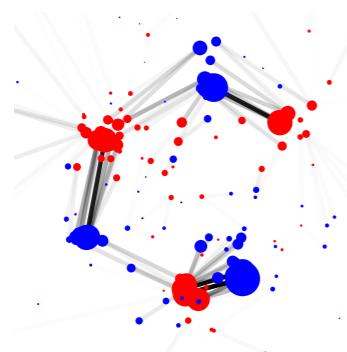
Particle Physics 101

*High-energy collisions can yield insights into fundamental physics
Machine learning offers powerful tools to analyze collision debris*



What is a Collider Event?

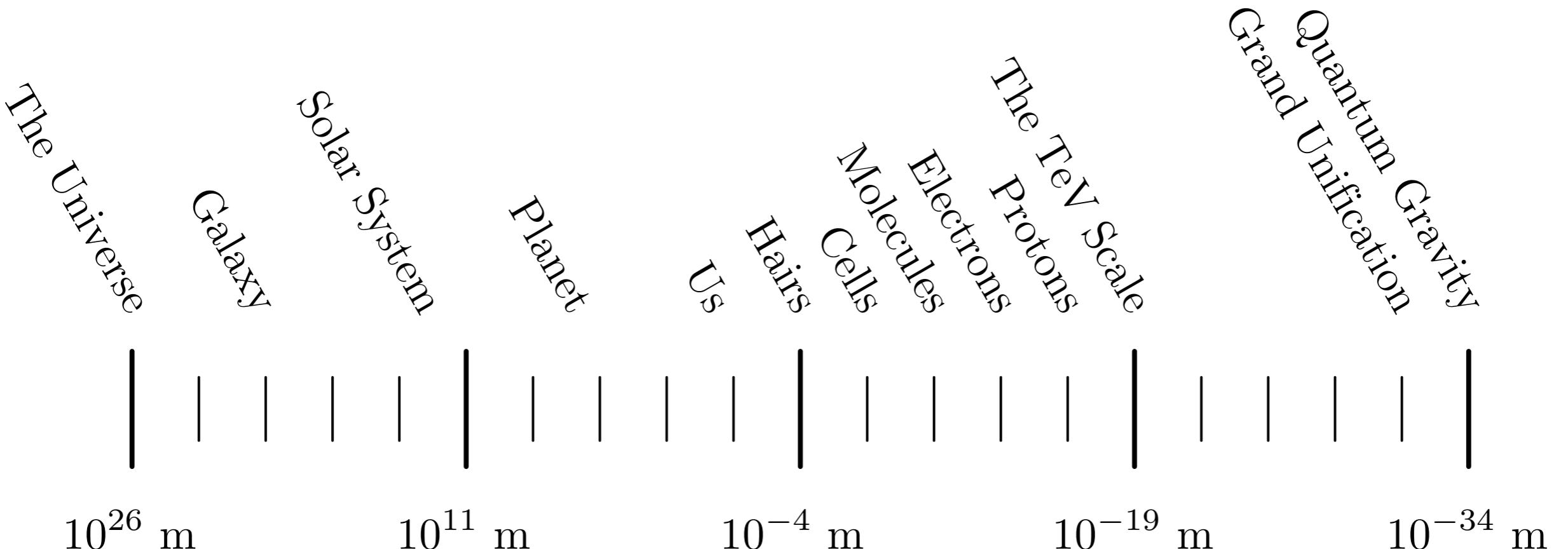
*Unordered set of particles describing energy flow of jets
Inspires network architectures designed for symmetry and safety*



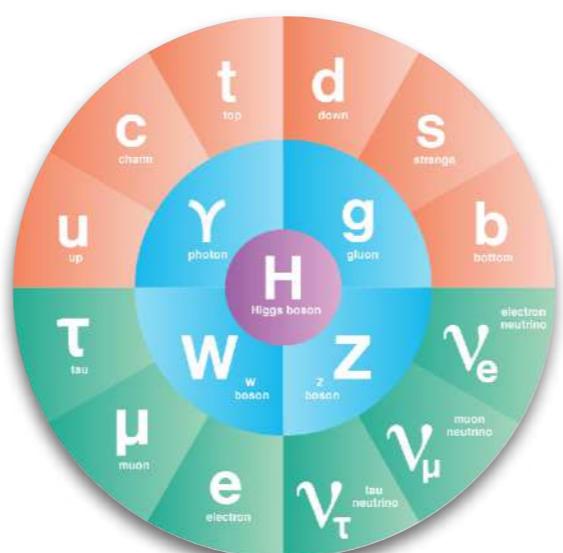
When are Collider Events Similar?

*When their energy flows are similar
Inspires unsupervised learning strategies based on event geometry*

Backup Slides

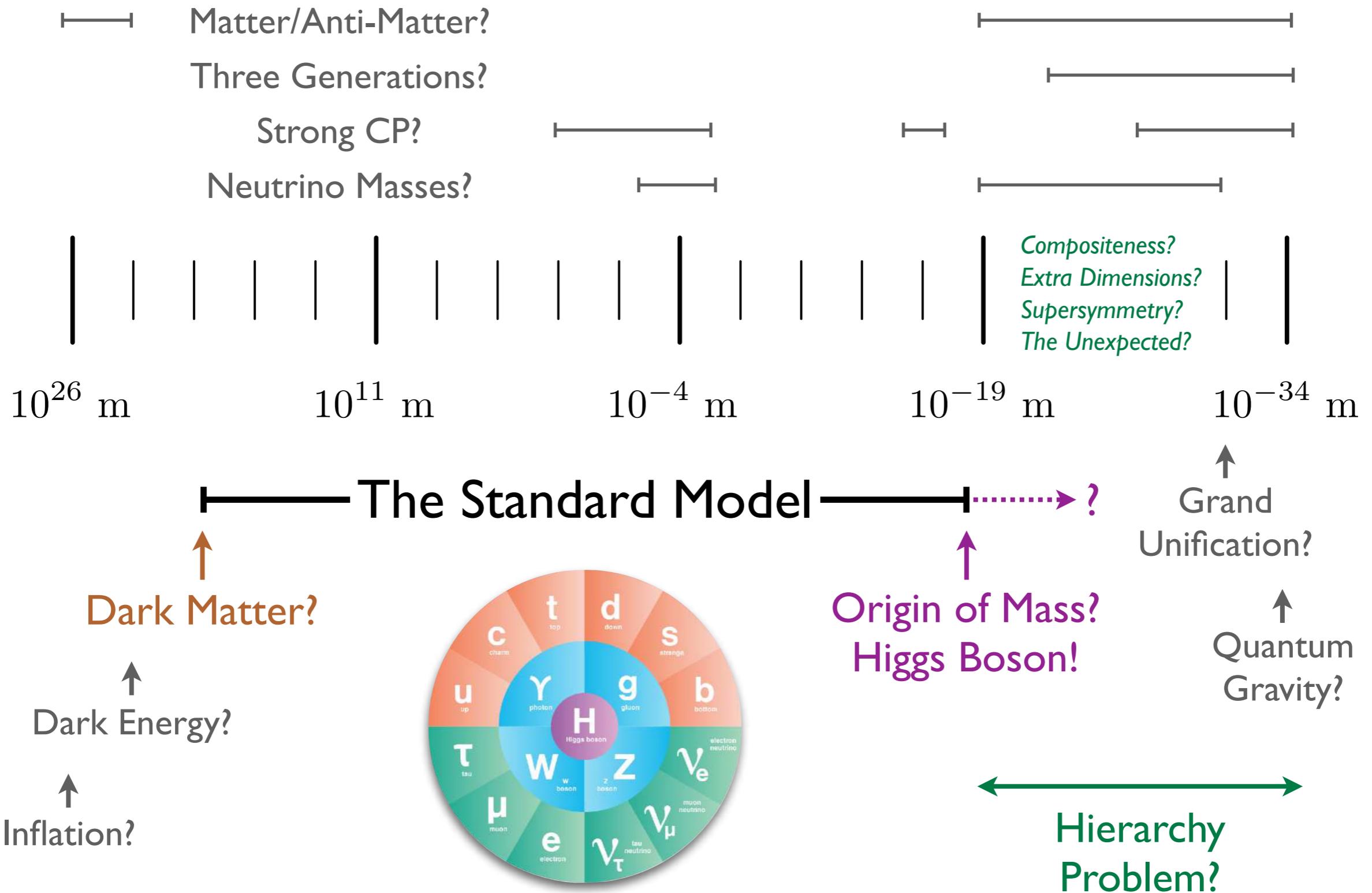


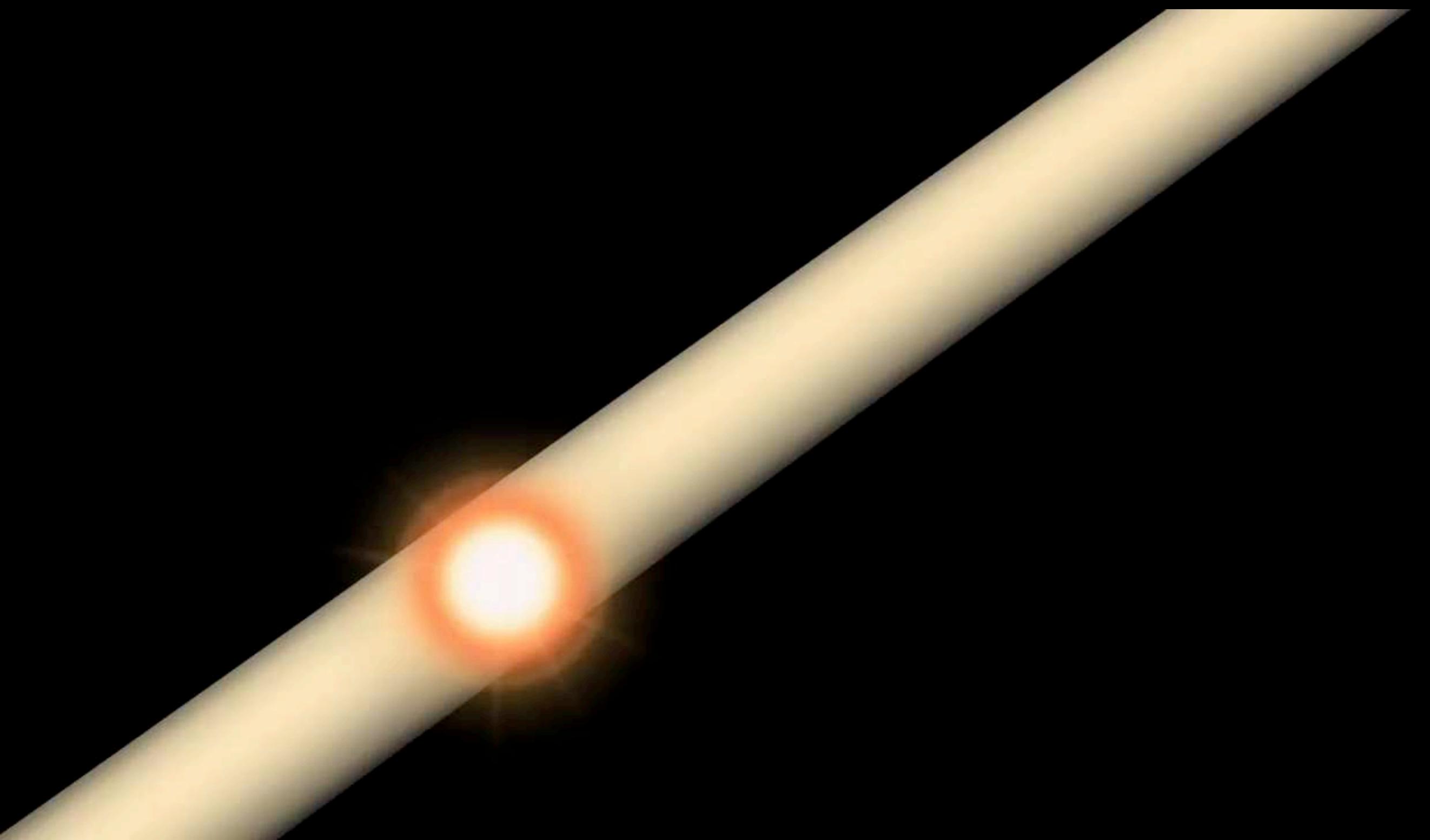
The Standard Model



Puzzles in Particle Physics

SnowMass2021









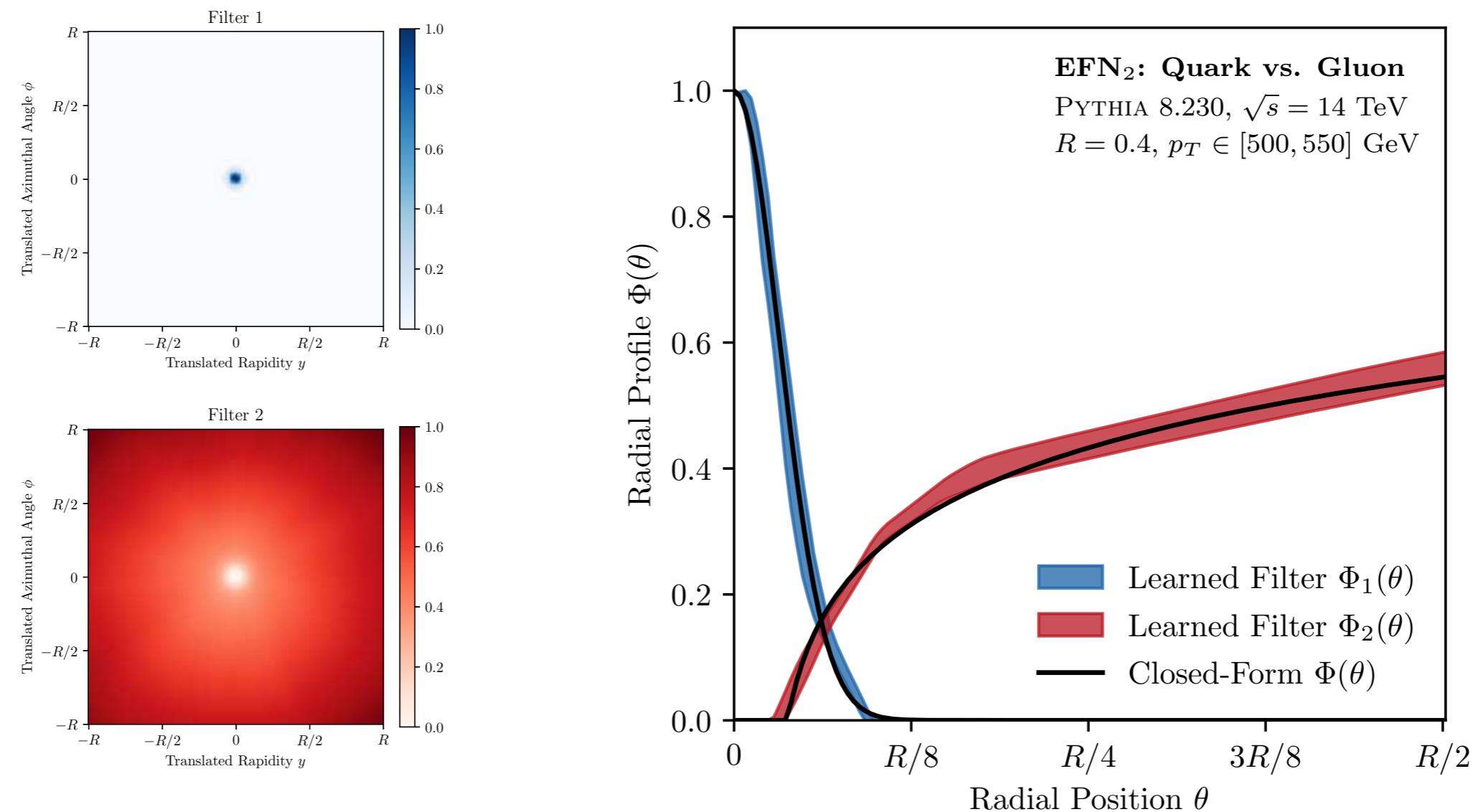
*“Ok, but did you really learn something
you didn’t already know?”*

Learning from the Machine



For $\ell = 2$ EFN, radial moments: $\sum_{i \in \text{jet}} z_i f(\theta_i)$

cf. Angularities:
 $f(\theta) = \theta^\beta$



[Komiske, Metodiev, JDT, JHEP 2019;
cf. Larkoski, JDT, Waalewijn, JHEP 2014; using Berger, Kucs, Sterman, PRD 2003; Ellis, Vermilion, Walsh, Hornig, Lee, JHEP 2010]

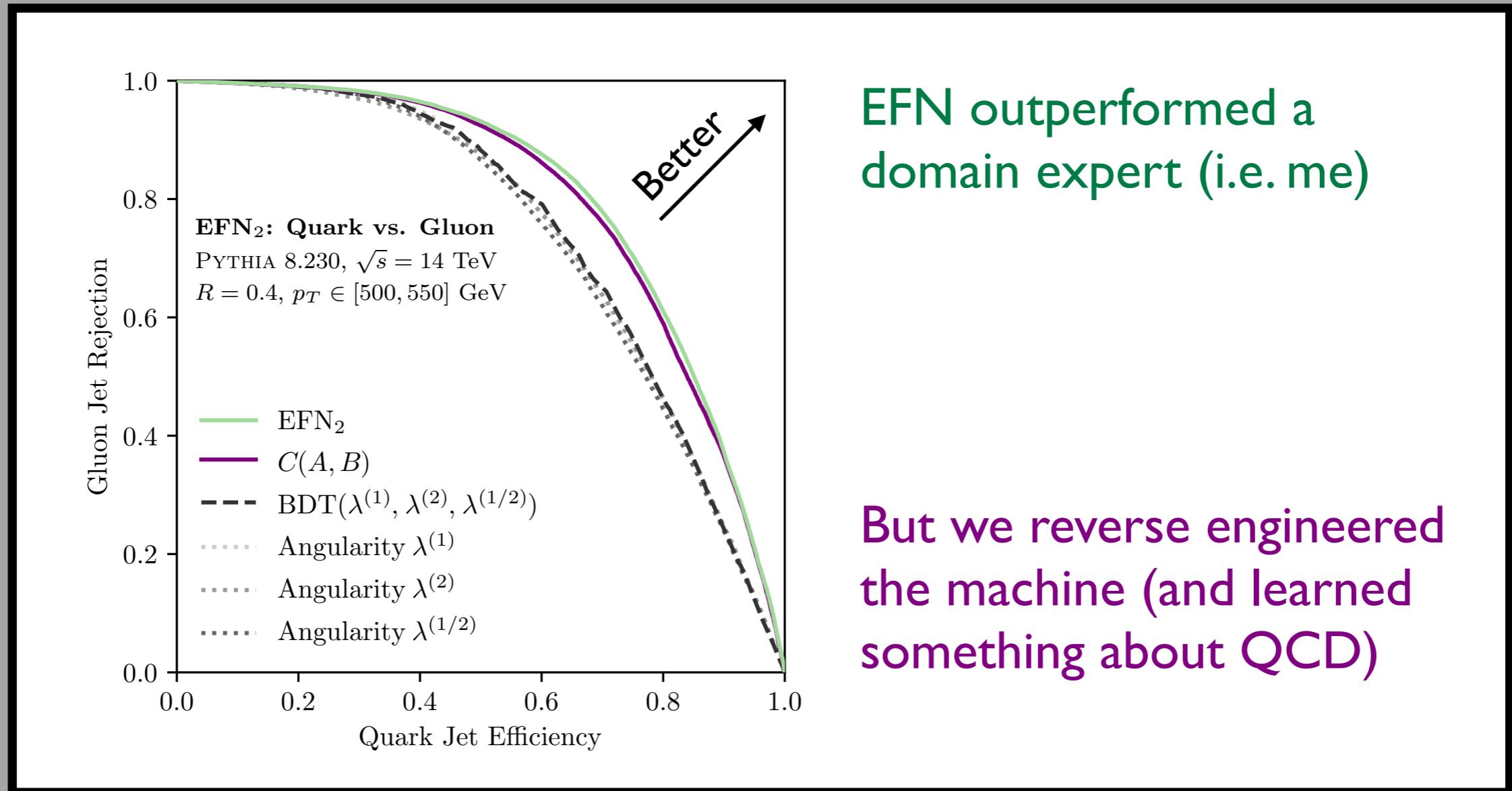
Learning from the Machine



For $\ell = 2$ EFN, radial moments:

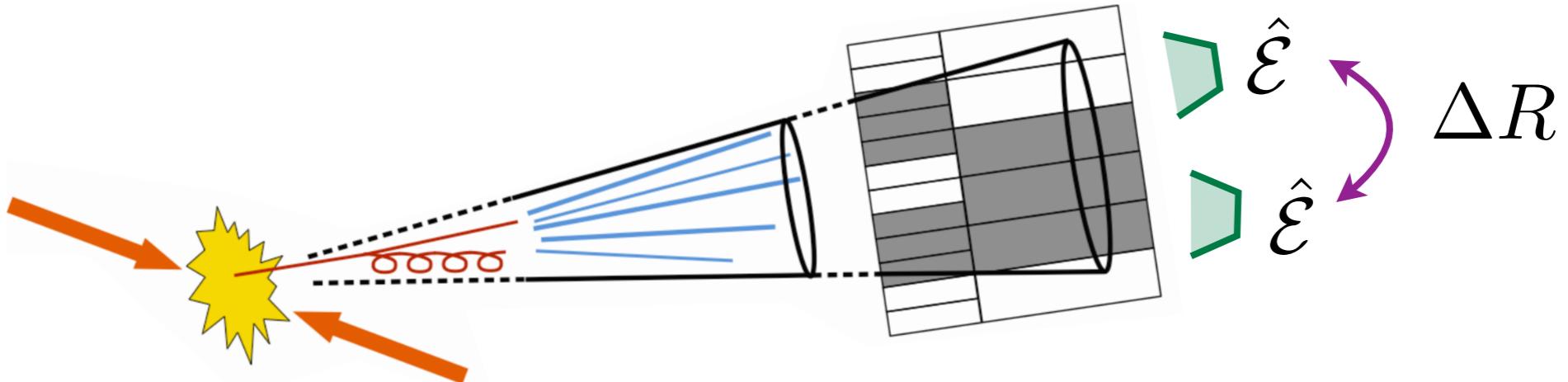
$$\sum_{i \in \text{jet}} z_i f(\theta_i)$$

cf. Angularities:
 $f(\theta) = \theta^\beta$

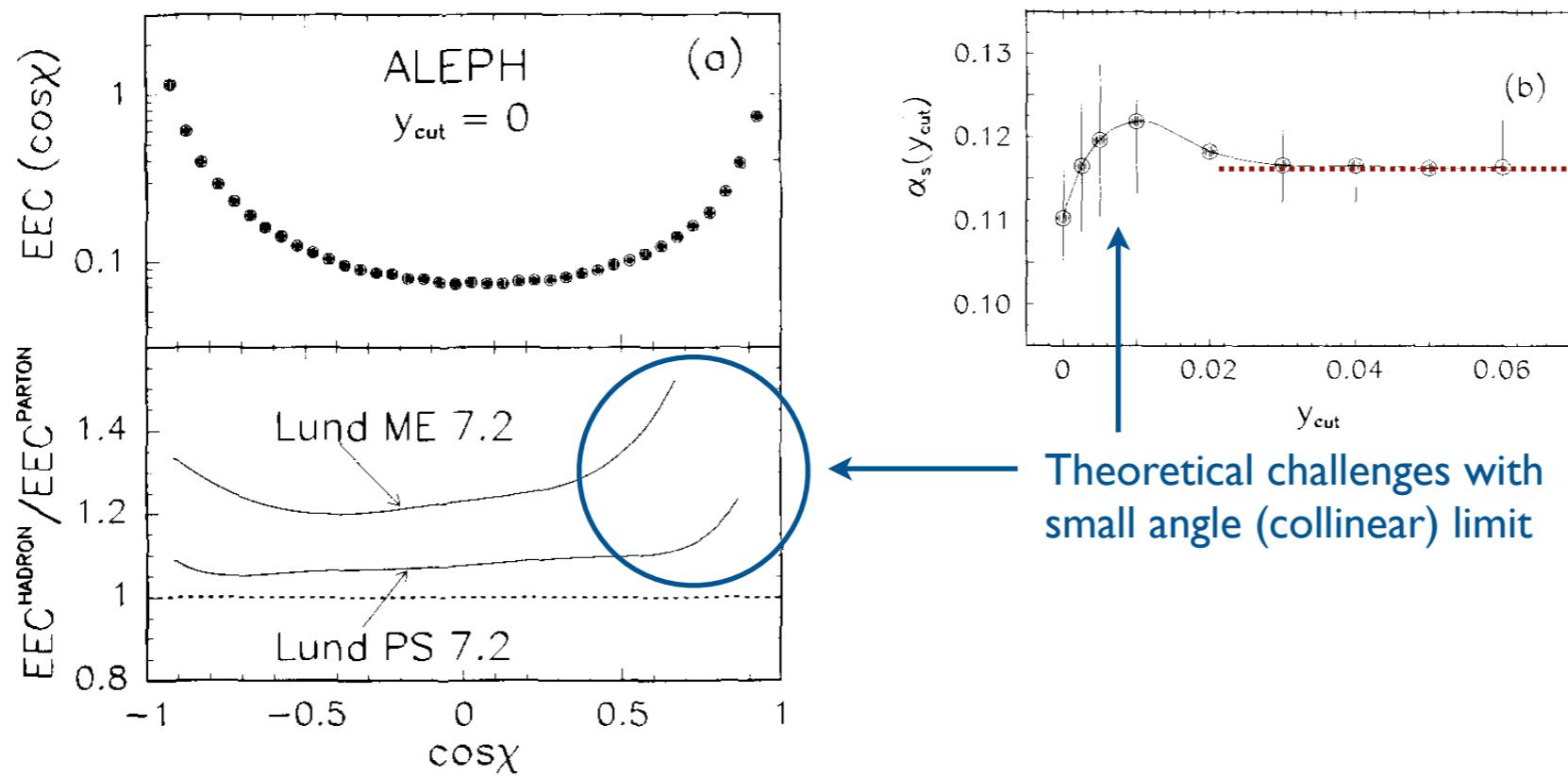


[Komiske, Metodiev, JDT, [JHEP 2019](#);
cf. Larkoski, JDT, Waalewijn, [JHEP 2014](#); using Berger, Kucs, Sterman, [PRD 2003](#); Ellis, Vermilion, Walsh, Hornig, Lee, [JHEP 2010](#)]

Energy-Energy Correlators

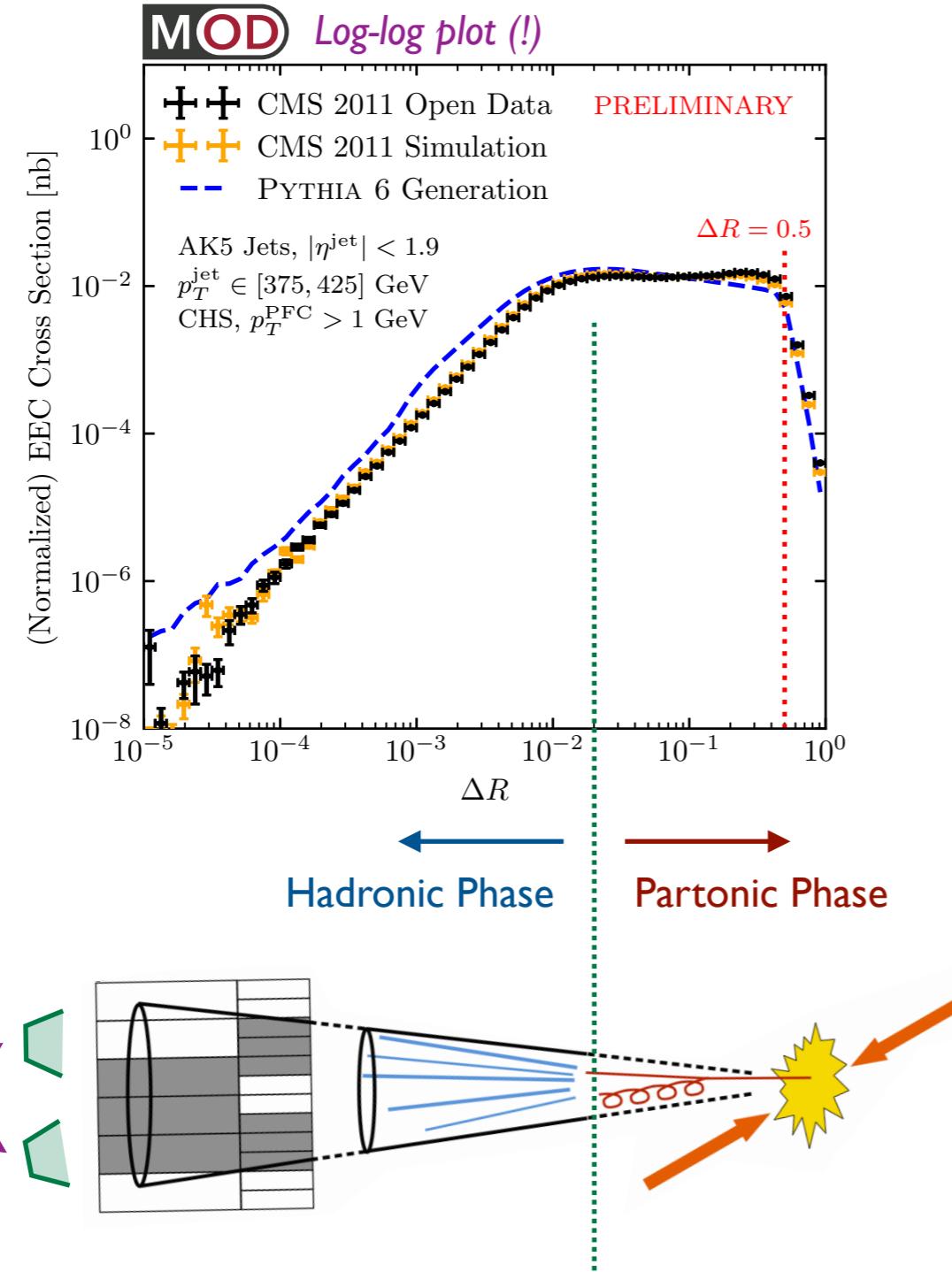


A long history in probing collinear dynamics of QCD

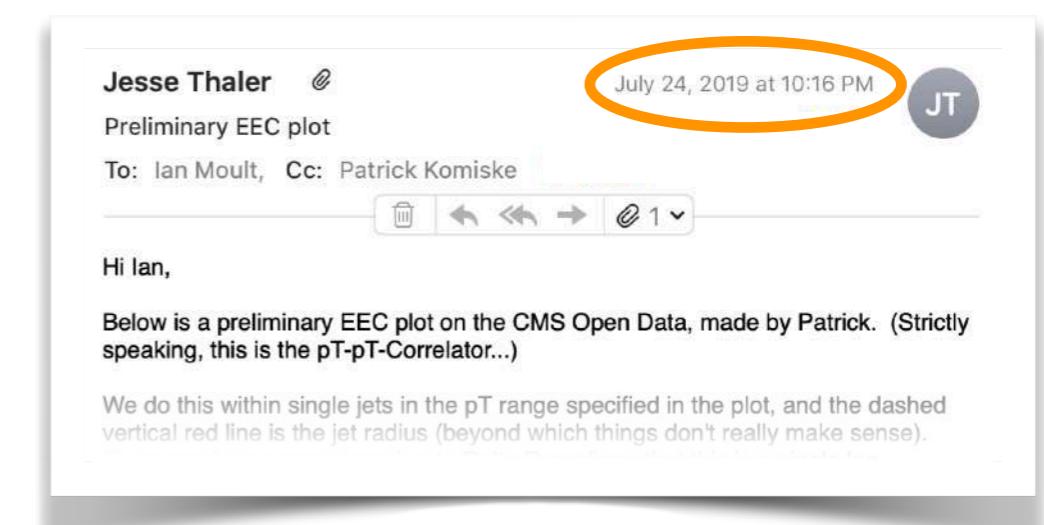
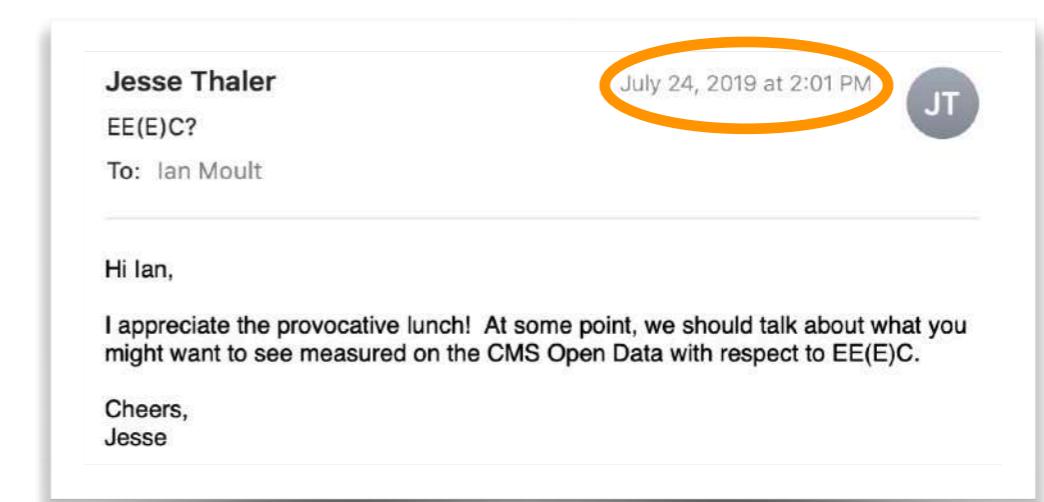


[Basham, Brown, Ellis, Love, [PRL 1978](#); ALEPH, [PLB 1991](#);
see Chen, Moult, Zhang, Zhu, [PRD 2020](#)]

QCD Phase Transition in Jets?



Behind the scenes at **BOSTON 2019**



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

