

Collision Course

Particle Physics meets Machine Learning

Jesse Thaler



Physics Colloquium, U.C. San Diego — May 20, 2021

The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI)

“eye-phi”

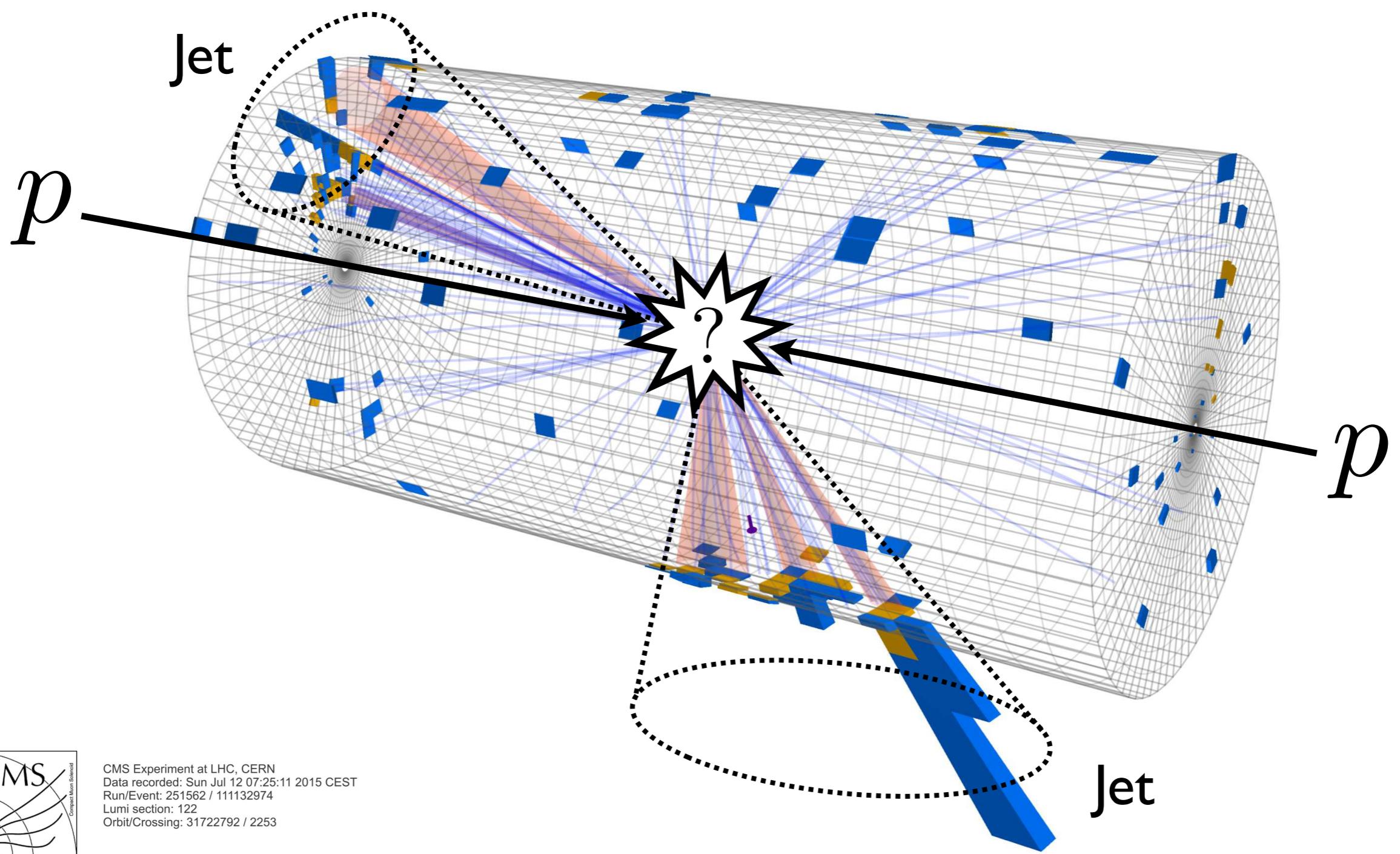


*Advance physics knowledge — from the smallest building blocks of nature
to the largest structures in the universe — and galvanize AI research innovation*



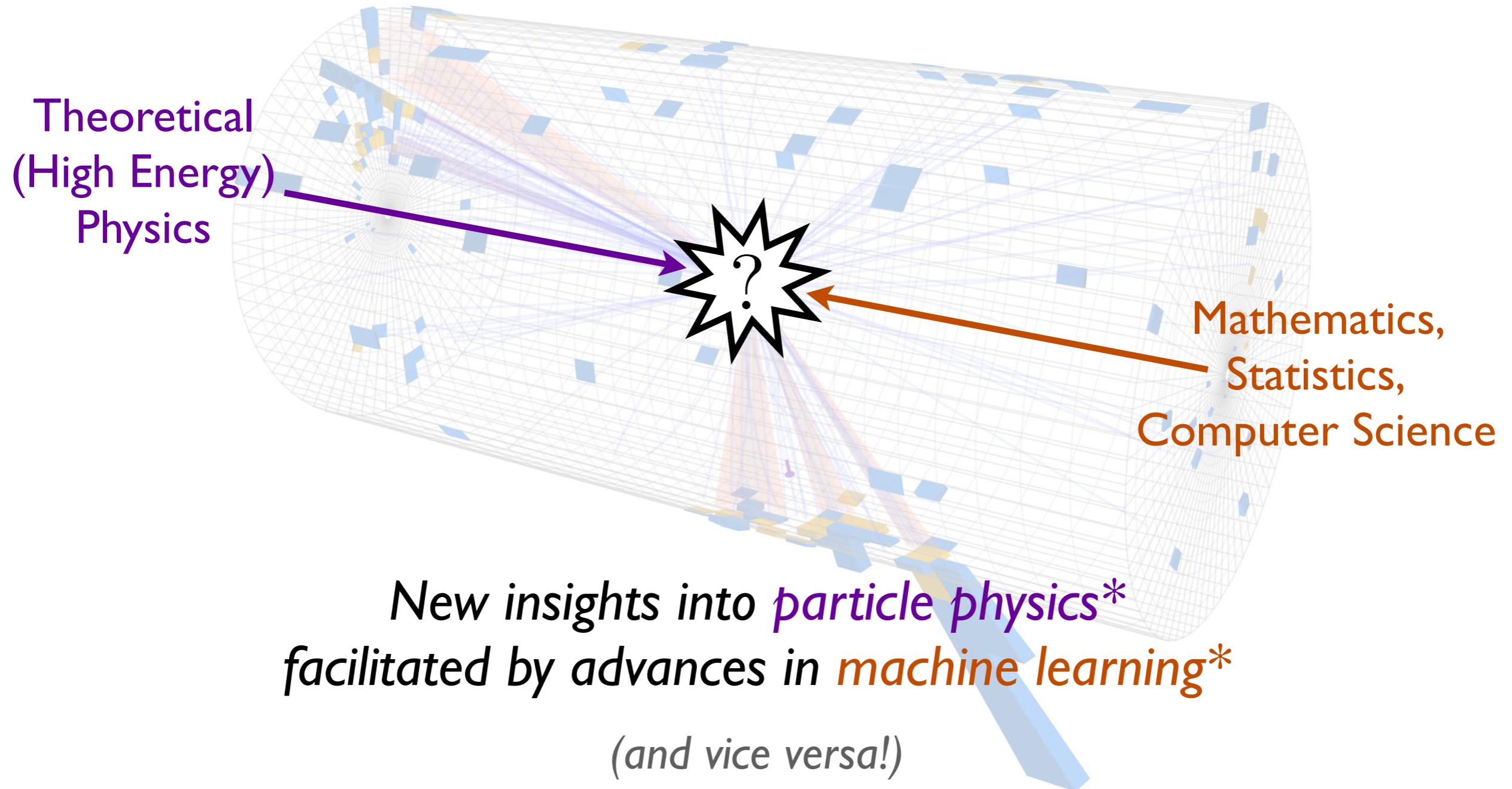
[<http://iaifi.org>, MIT News Announcement]

“Collision Course”



“Collision Course”

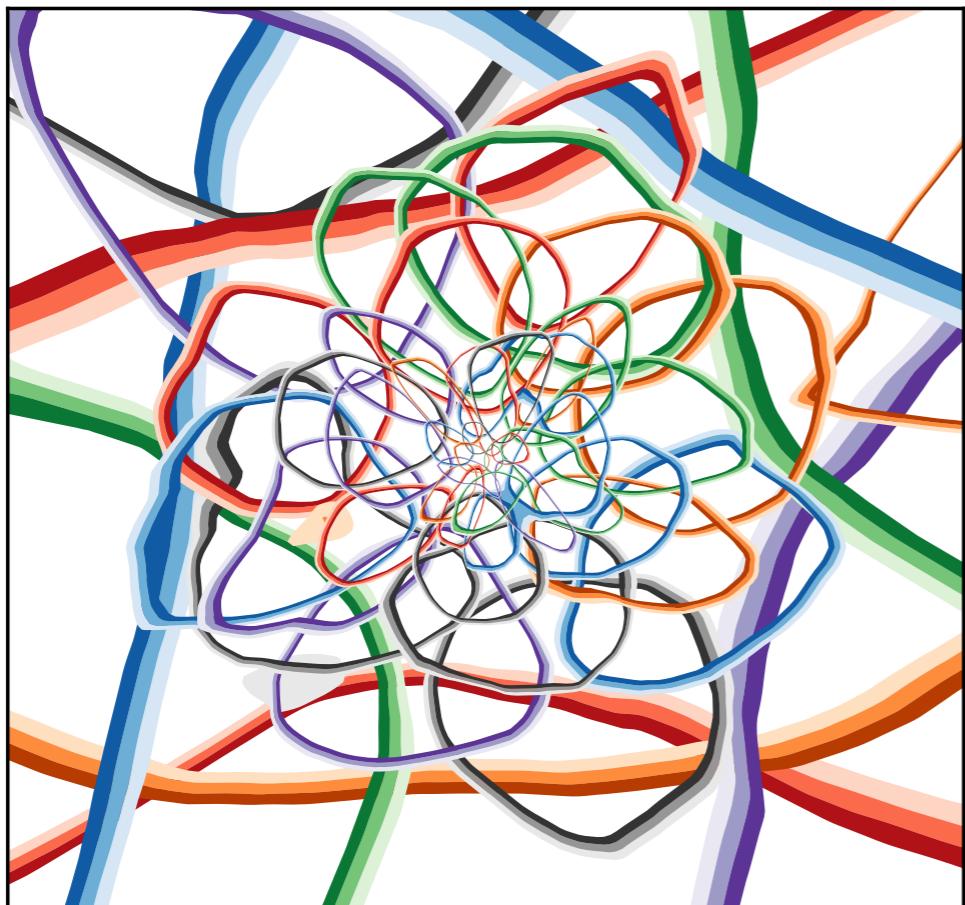
“Theoretical Physics for Machine Learning”
Aspen Center for Physics, January 2019



Two Anecdotes

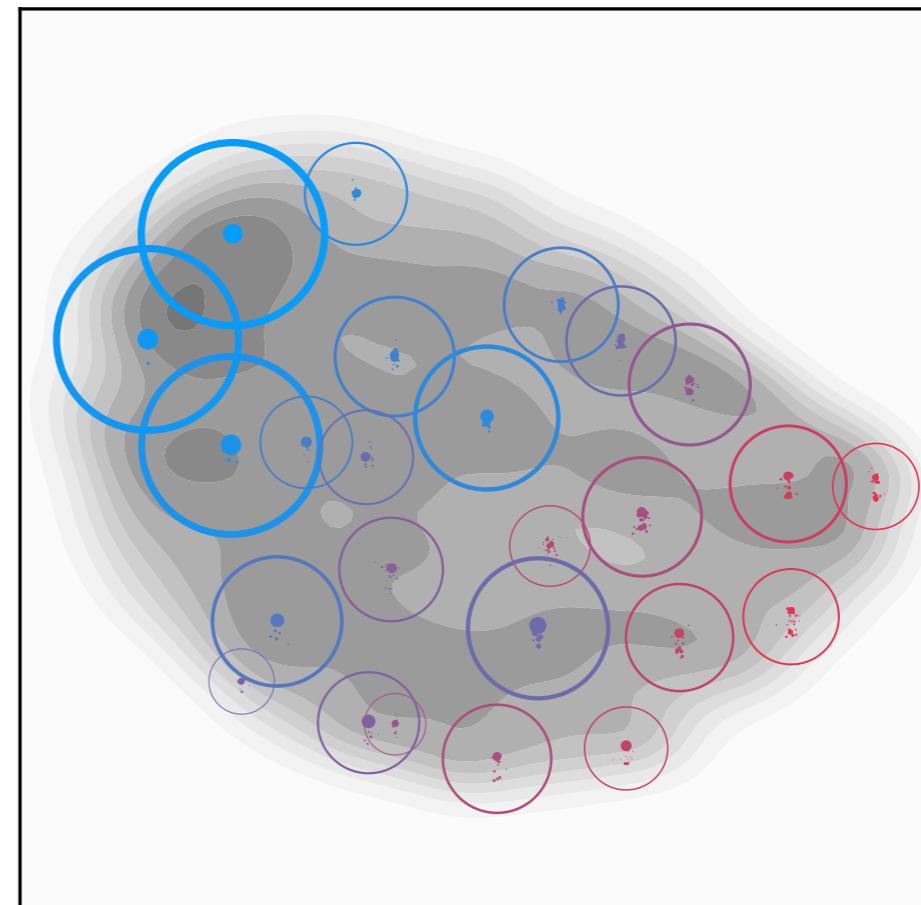


Teaching a Machine to “Think Like a Physicist”



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

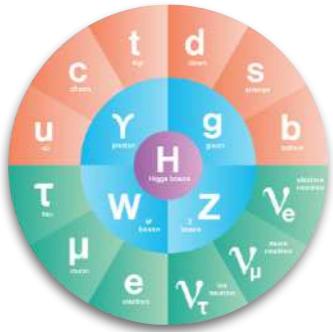
Letting Collider Data Speak for Itself



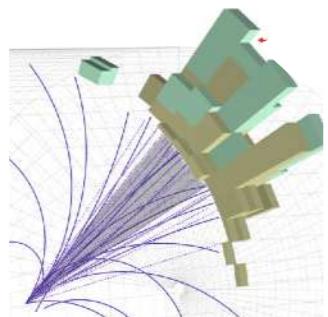
[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
based on Komiske, Metodiev, JDT, [PRL 2019](#)]

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

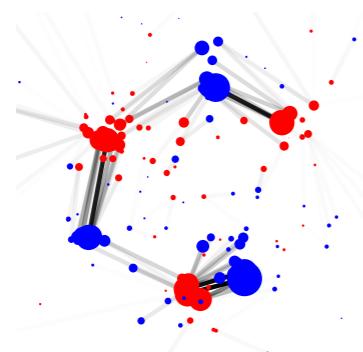
Outline



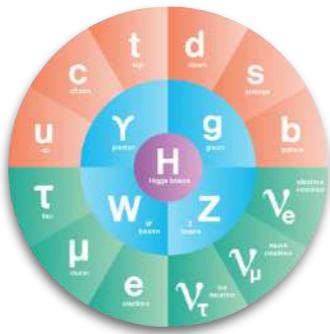
Particle Physics 101



What is a Collider Event?



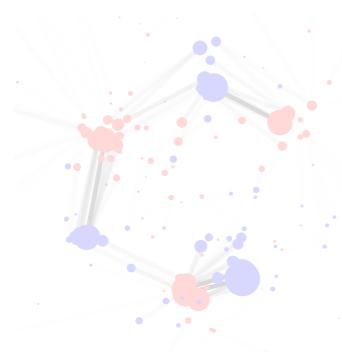
When are Collider Events Similar?



Particle Physics 101

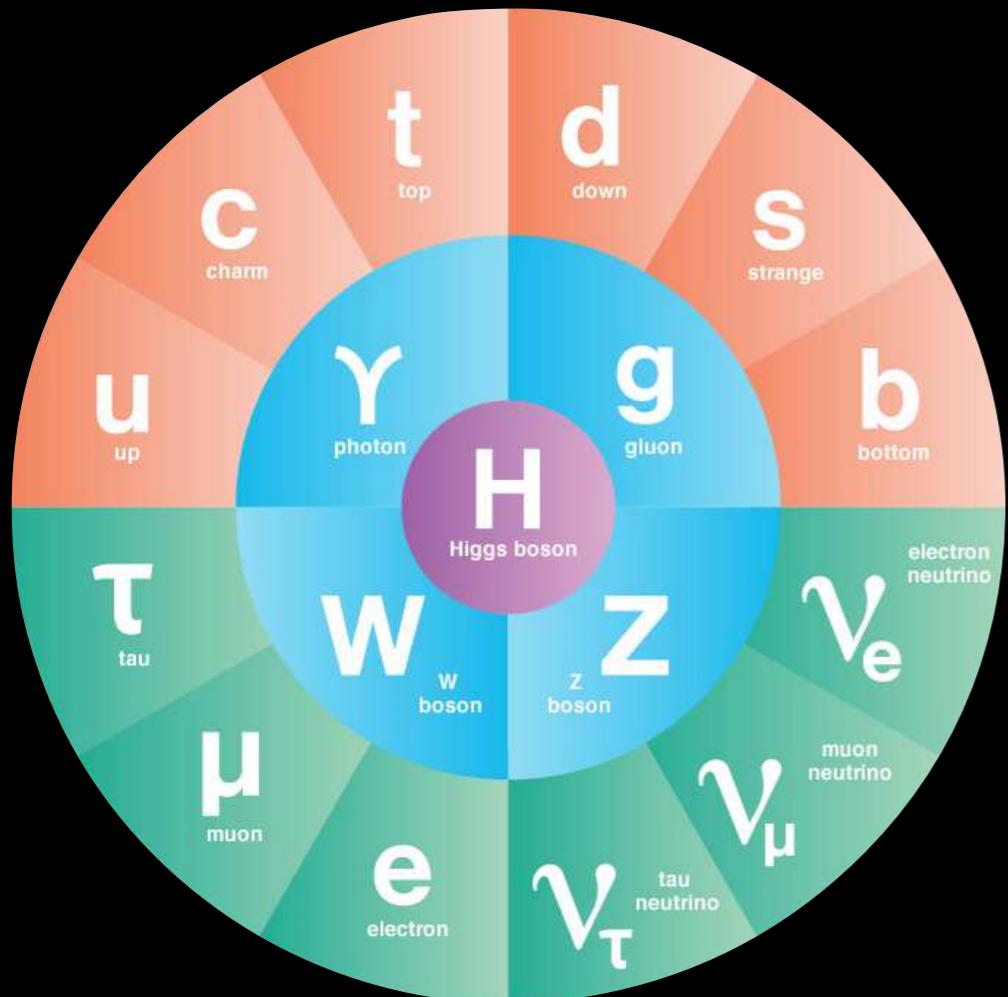


What is a Collider Event?

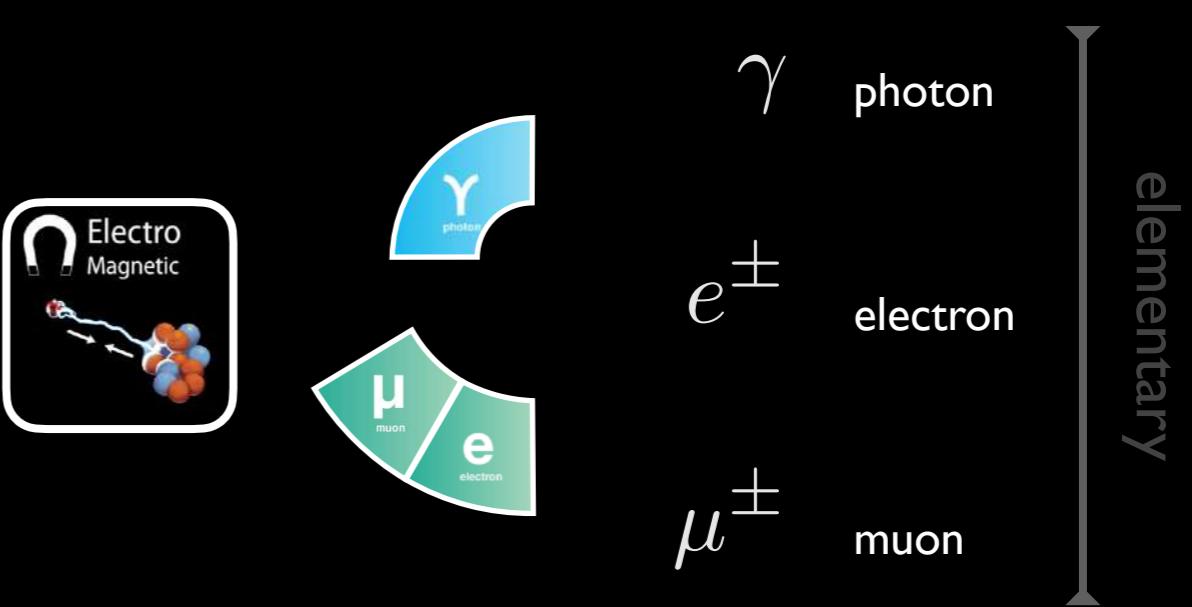
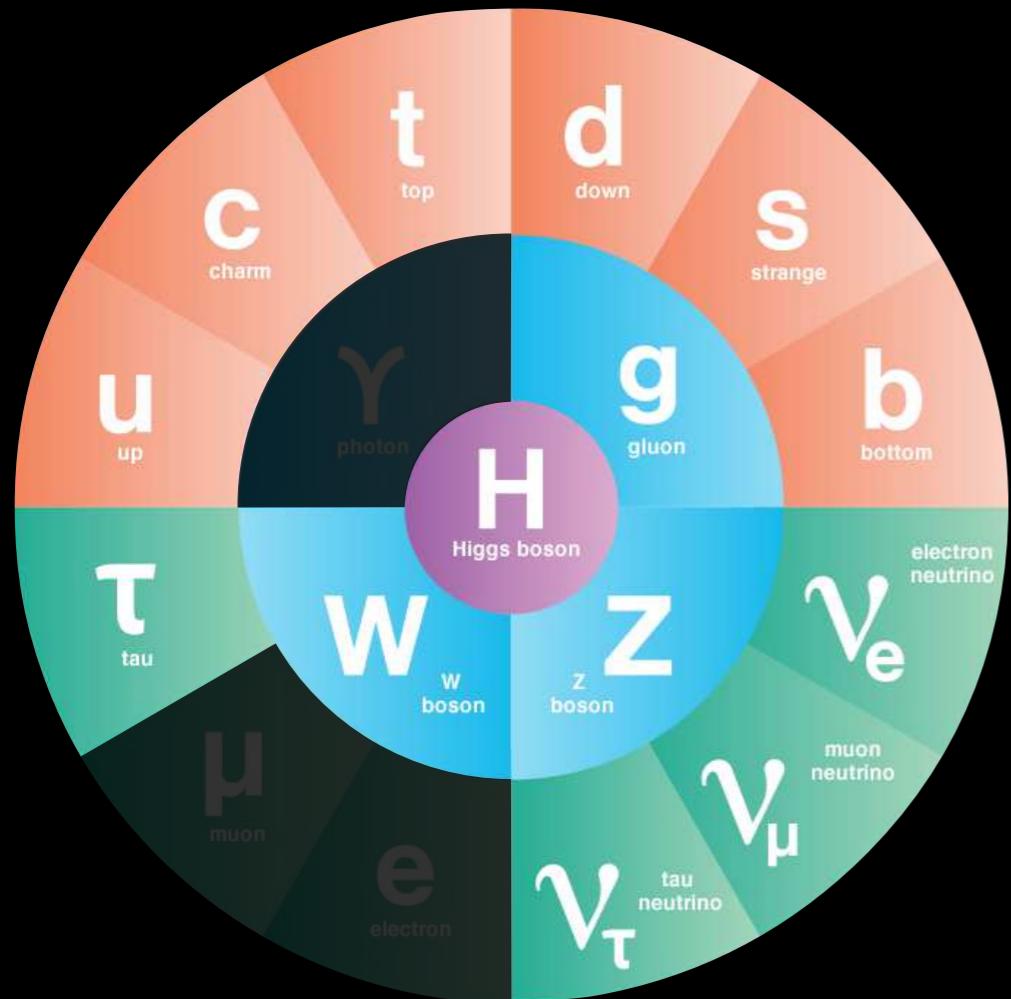


When are Collider Events Similar?

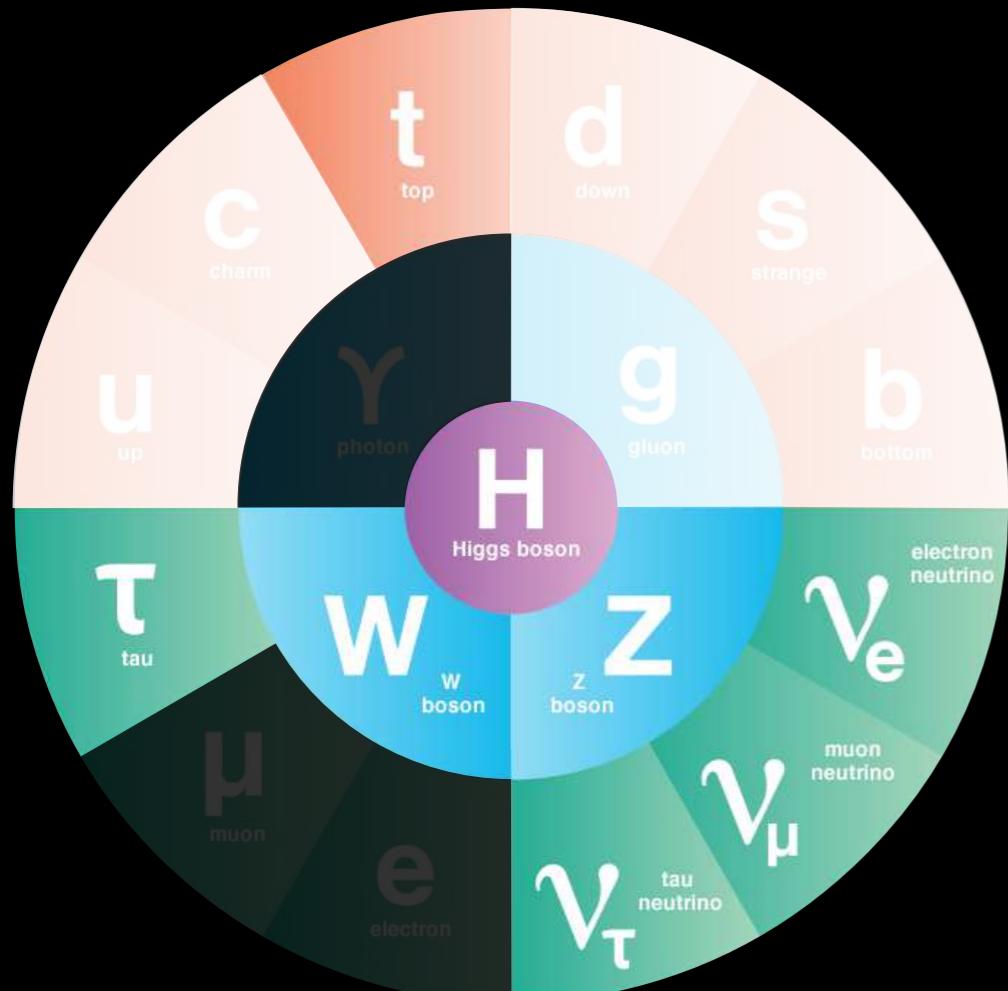
Particle Physics 101



Particle Physics 101

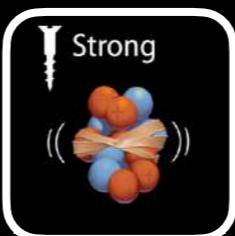


Particle Physics 101



QCD Confinement

**Quarks
&
Gluons**



γ photon

e^+ electron

μ^+ muon

π^+ pion

K^+ kaon

K-long

K_L^0

proton

p/\bar{p}

neutron

n/\bar{n}

elementary

composite

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}

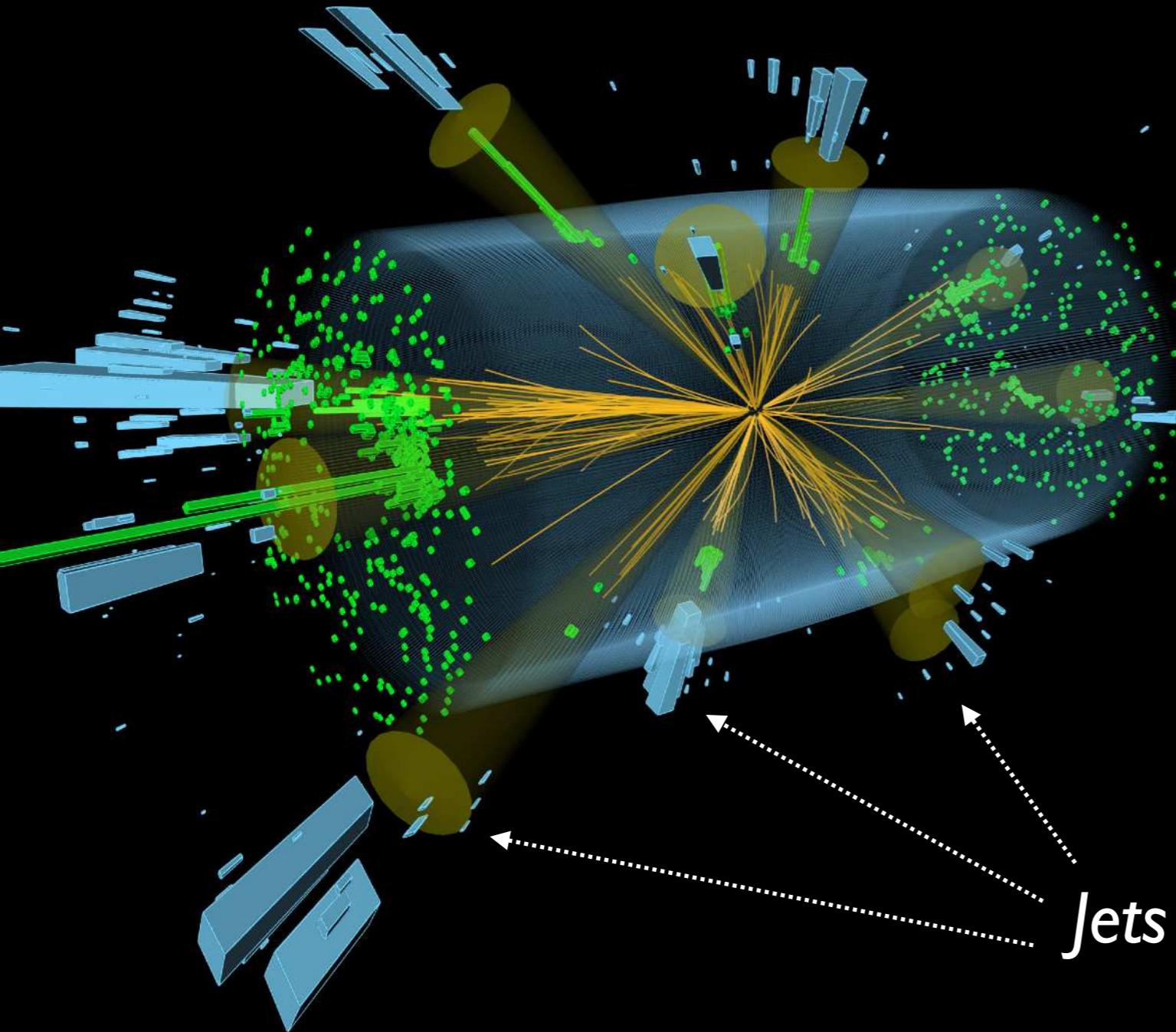
neutron

elementary

composite

Collider Event

Every 25 nanoseconds at the LHC



T E H M



γ

photon



e^+

electron



μ^+

muon



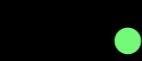
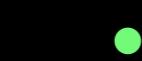
π^+

pion



K^+

kaon



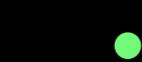
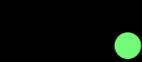
K_L^0

K-long



p/\bar{p}

proton



n/\bar{n}

neutron

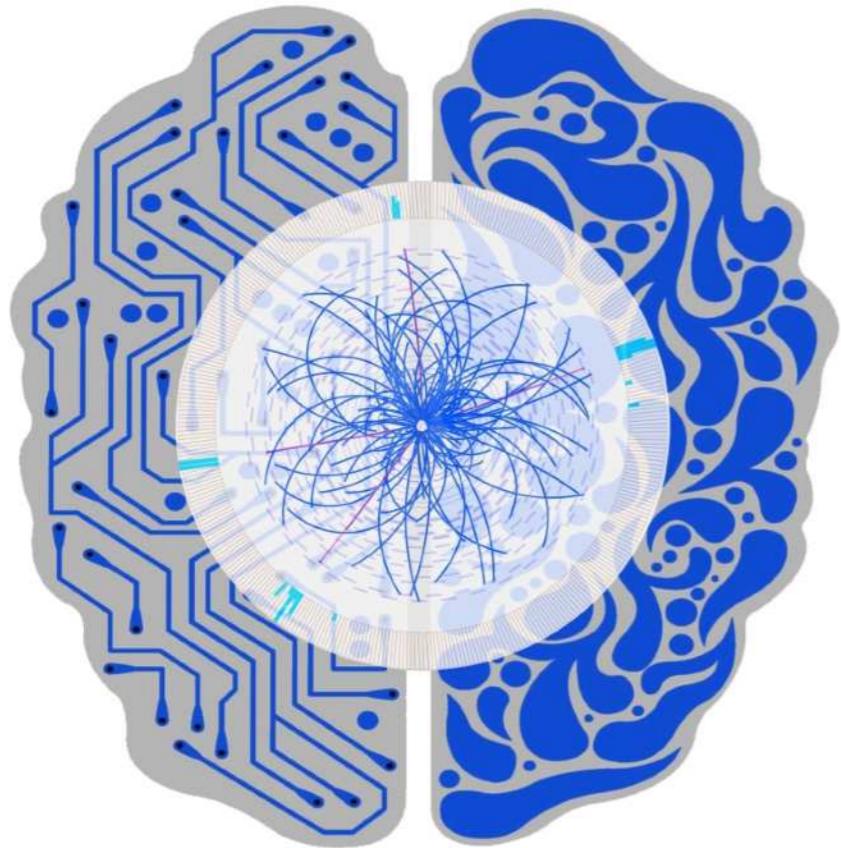
composite

elementary

↓

composite

↓



*Can we teach a machine
to “think” like a physicist?*

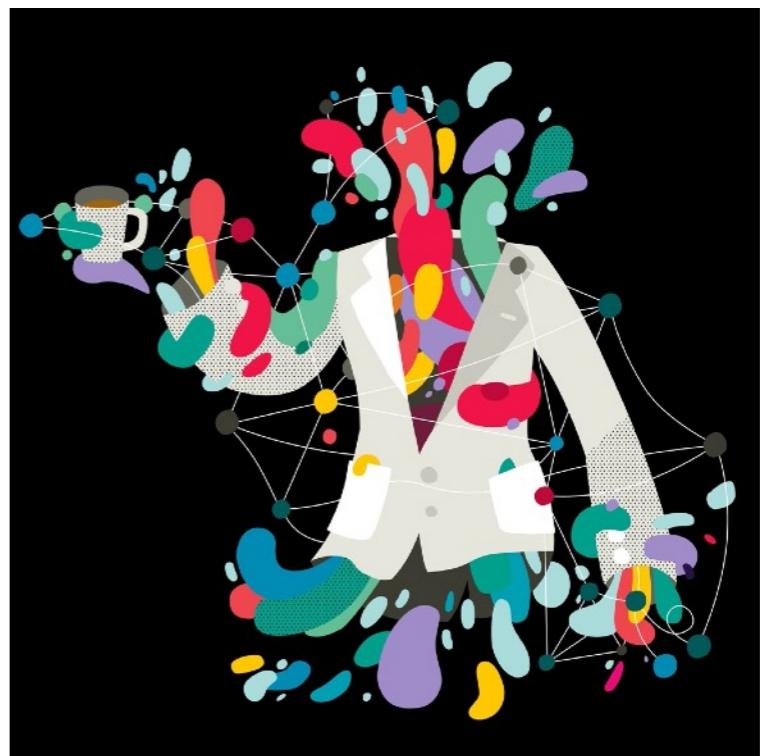
The New York Times



By Dennis Overbye

Nov. 23, 2020

Can a Computer Devise a Theory of Everything?



AI²: Ab Initio Artificial Intelligence

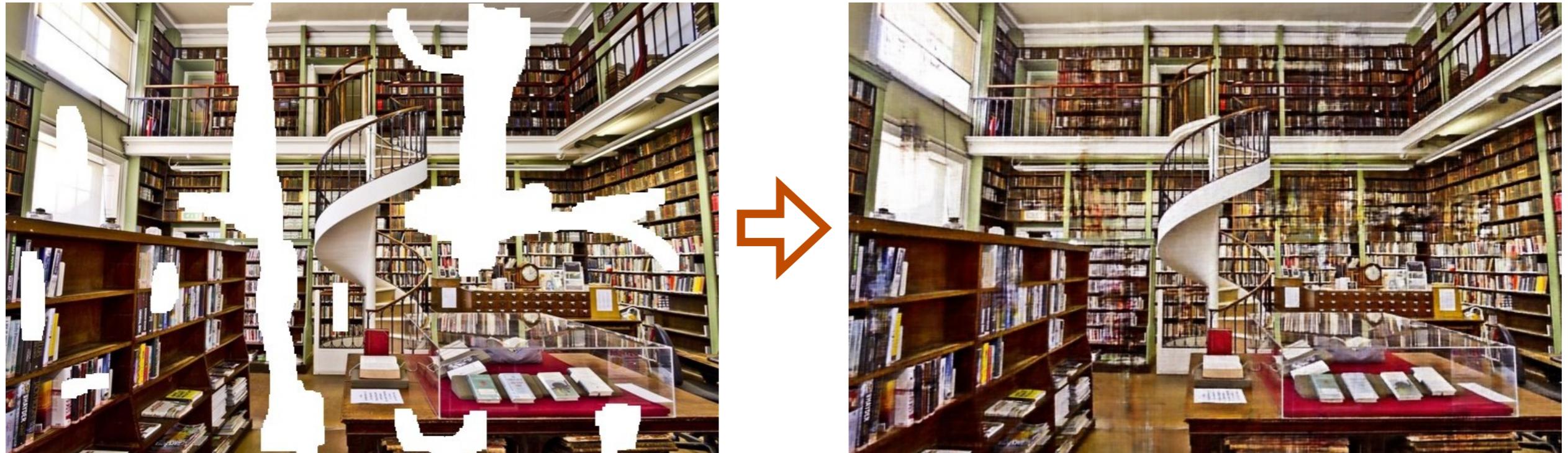


*Machine learning that incorporates
first principles, best practices, and domain knowledge
from fundamental physics*

*Symmetries, conservation laws, scaling relations, limiting behaviors, locality, causality,
unitarity, gauge invariance, entropy, least action, factorization, unit tests,
exactness, systematic uncertainties, reproducibility, verifiability, ...*

Deep Learning

E.g. Inpainting

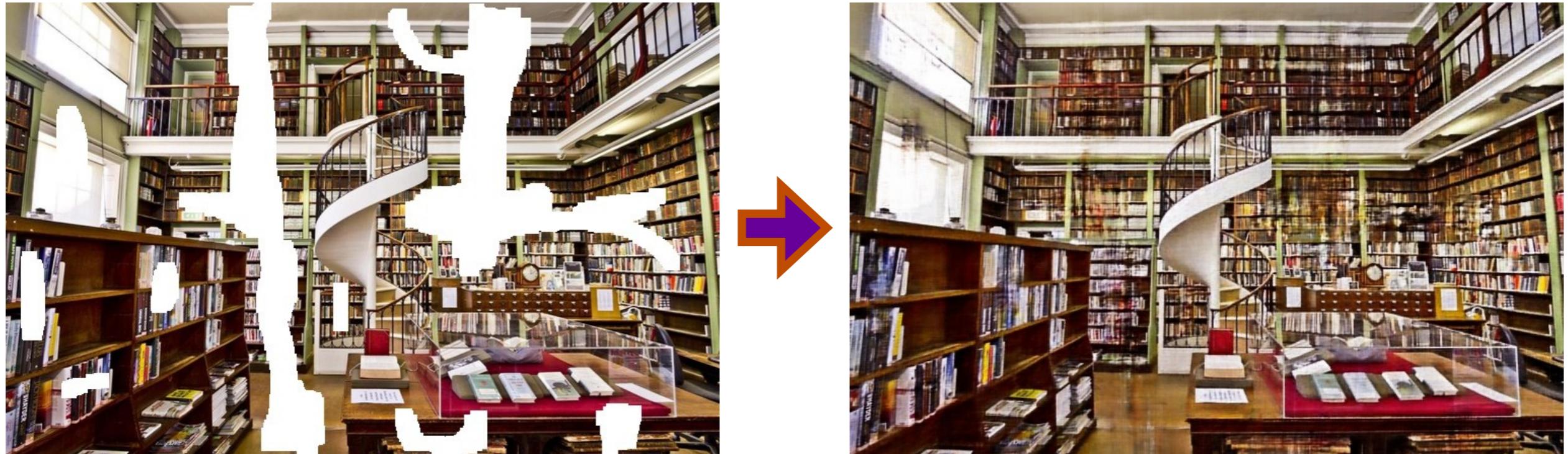


increased computational power and large data sets

[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

Deep Learning meets Deep Thinking

E.g. *Inpainting*



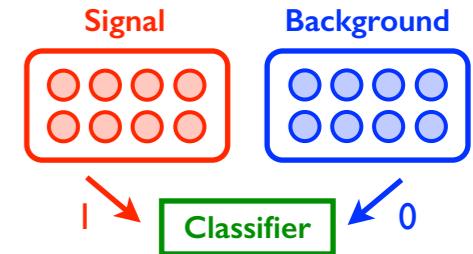
Using randomly initialized neural network (!)

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

[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

E.g. Quark/Gluon Classification

“Hello, World!” of Jet Physics



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

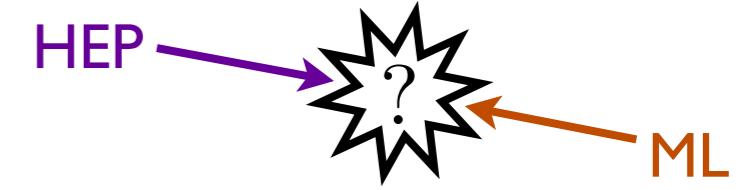
$$h(\text{Quark}) = 1$$
$$h(\text{Gluon}) = 0$$

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

(Neyman-Pearson lemma)

[see e.g. Gras, Höche, Kar, Larkoski, Lönnblad, Plätzer, Siódmok, Skands, Soyez, JDT, [JHEP 2017](#); Komiske, Metodiev, Schwartz, [JHEP 2017](#); Komiske, Metodiev, JDT, [JHEP 2018](#)]

E.g. Search for Supersymmetry

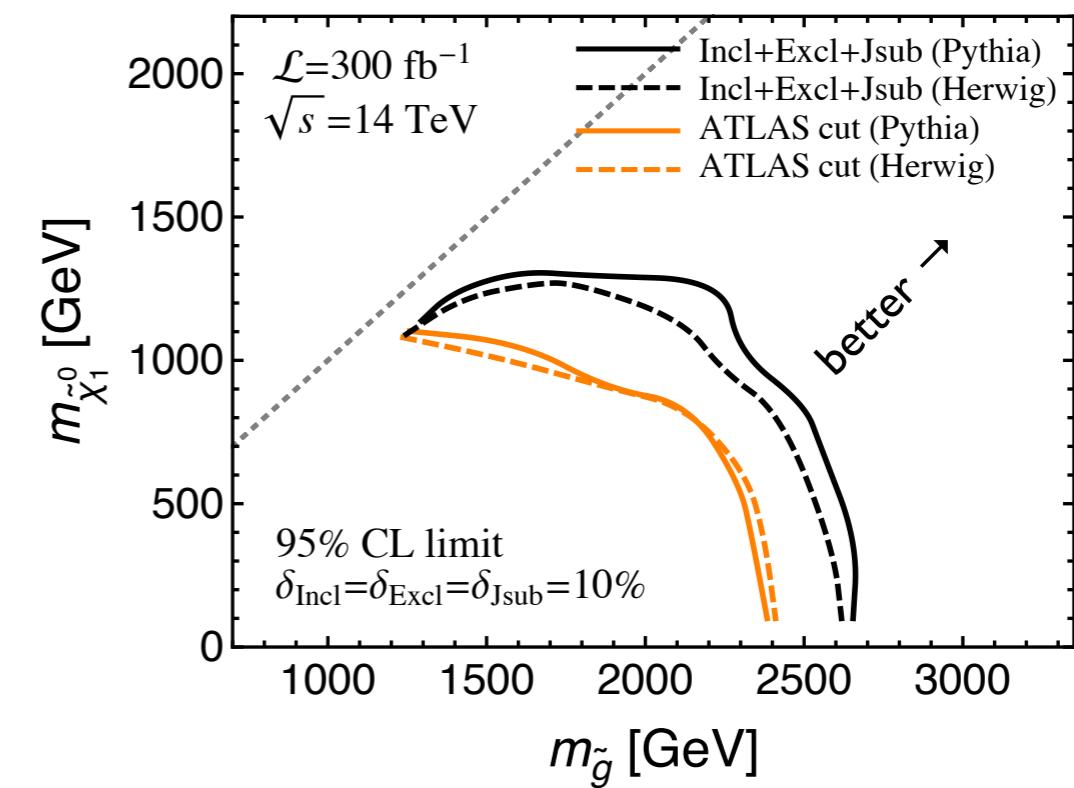
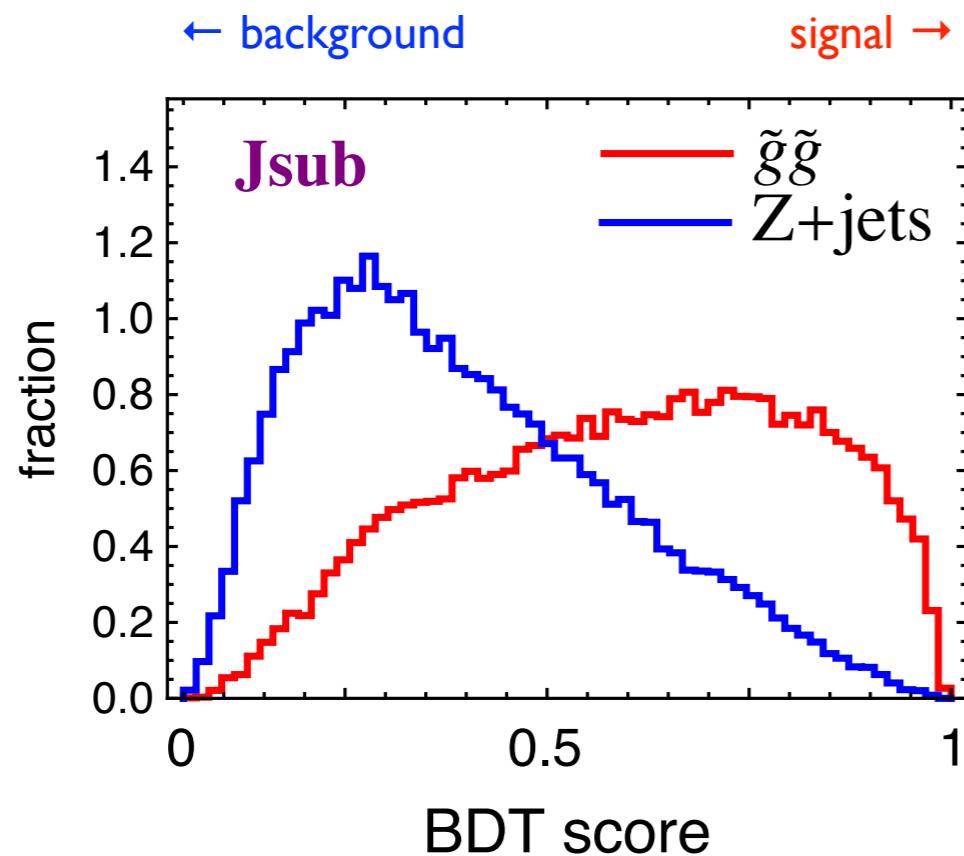
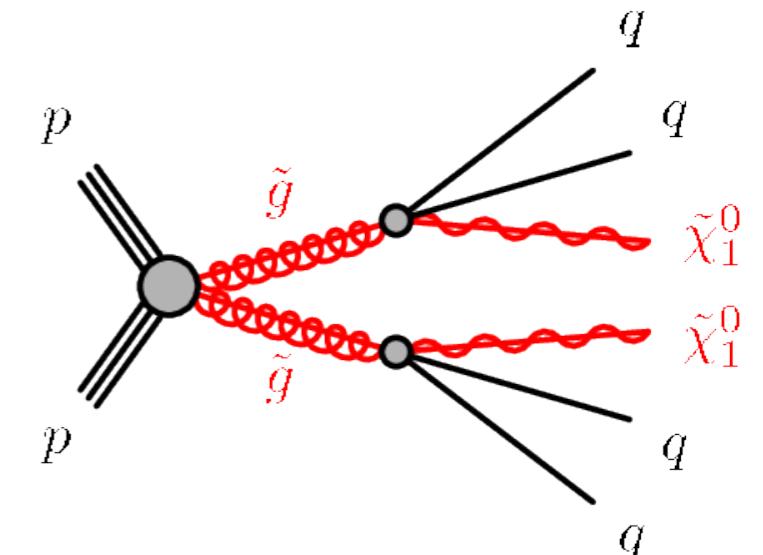


Classifier: Boosted decision tree (for each of 4 jets)

Inputs: Jet mass, width, track multiplicity

Signal: Quark enriched

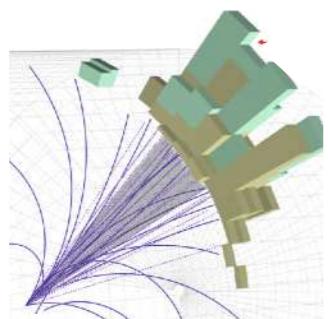
Background: Gluon enriched



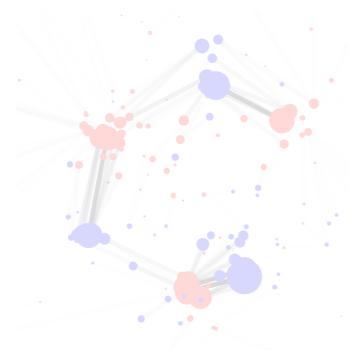
[Bhattacherjee, Mukhopadhyay, Nojiri, Sakakie, Webber, [JHEP 2017](#)]



Particle Physics 101



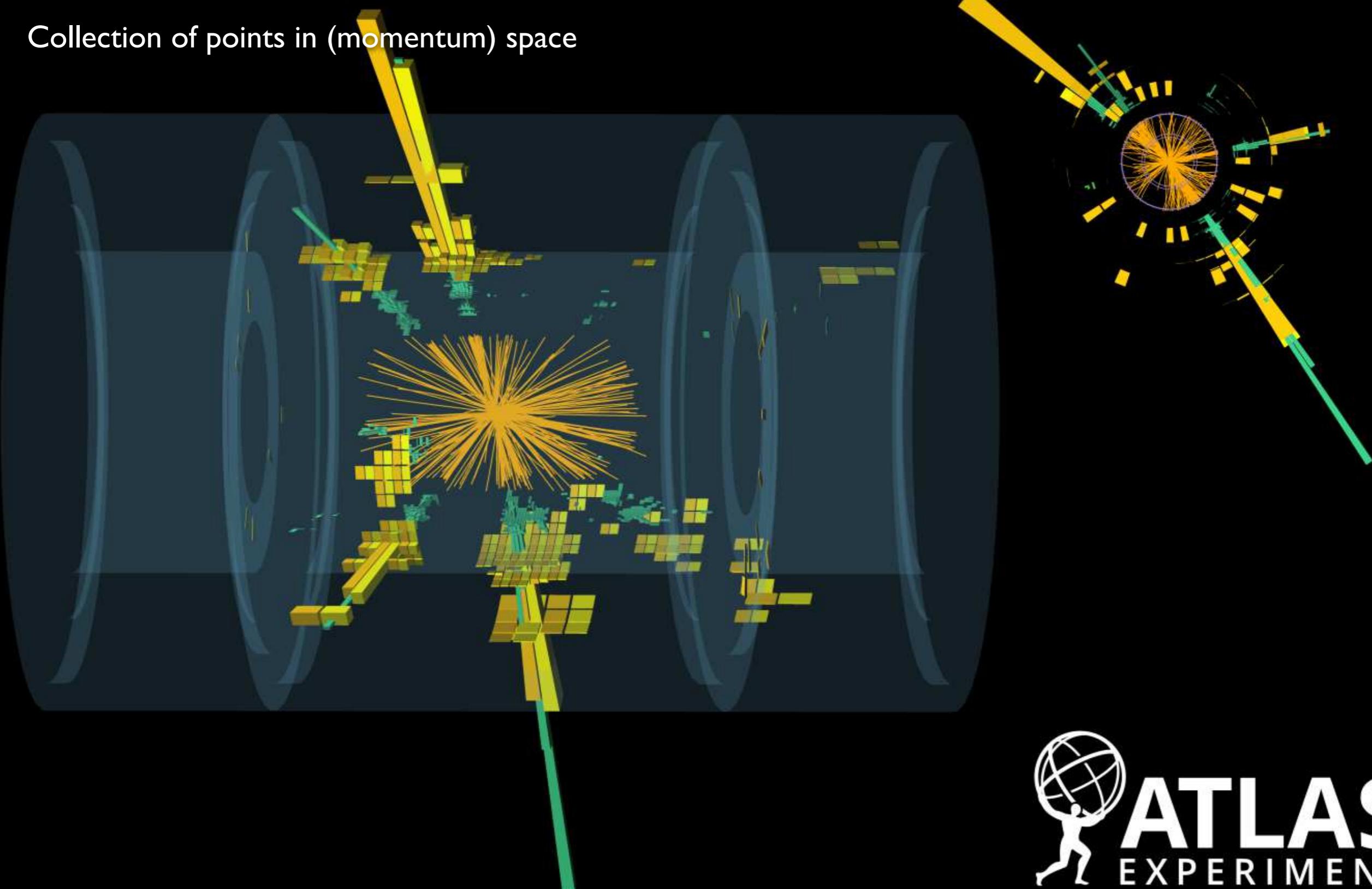
What is a Collider Event?



When are Collider Events Similar?

Collider Event

Collection of points in (momentum) space



Point Cloud

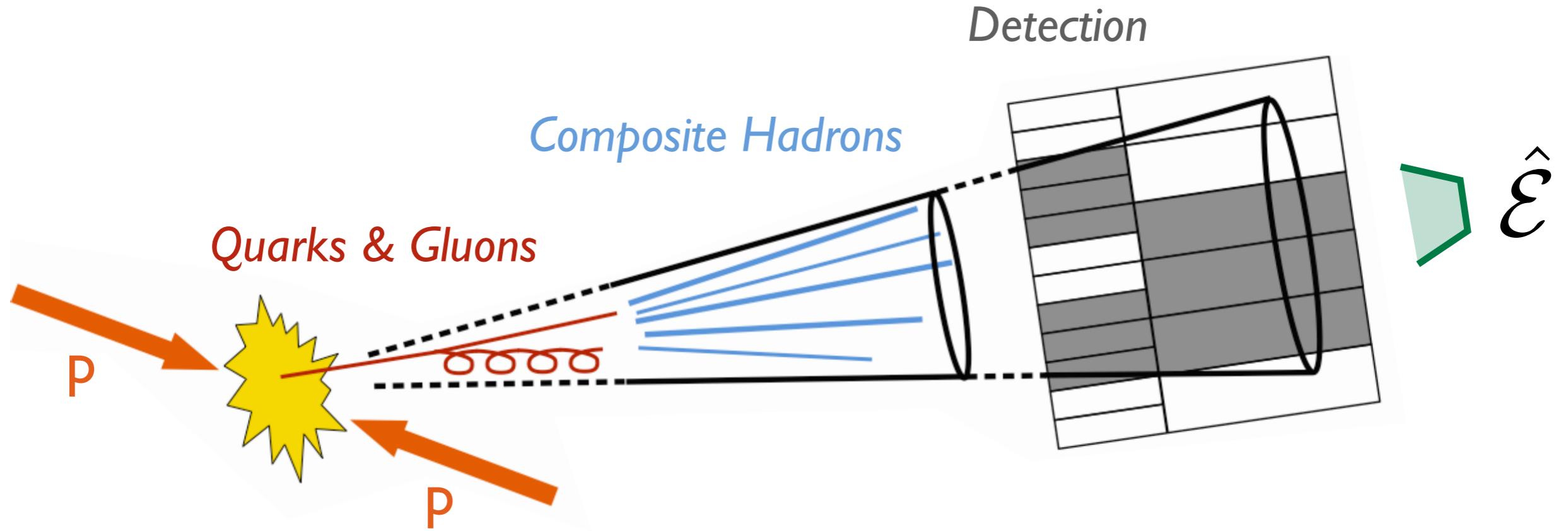
Collection of points in (position) space



[Popular Science, 2013]

Jet Formation from QCD

Theory



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](#)]

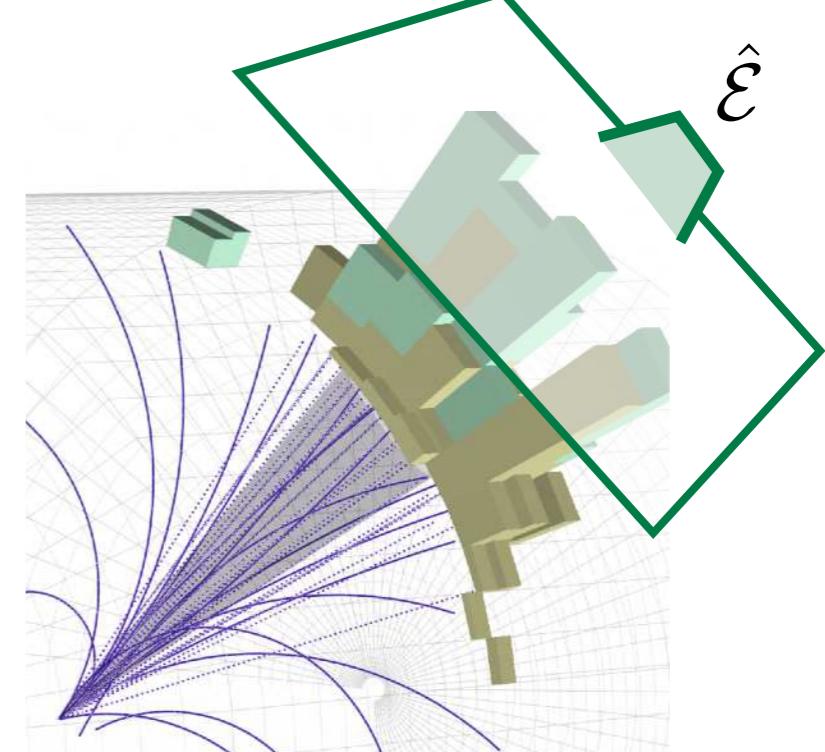
From Energy Flow to Weighted Point Clouds

- Energy-Weighted Directions

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

↑
Energy |
 Direction

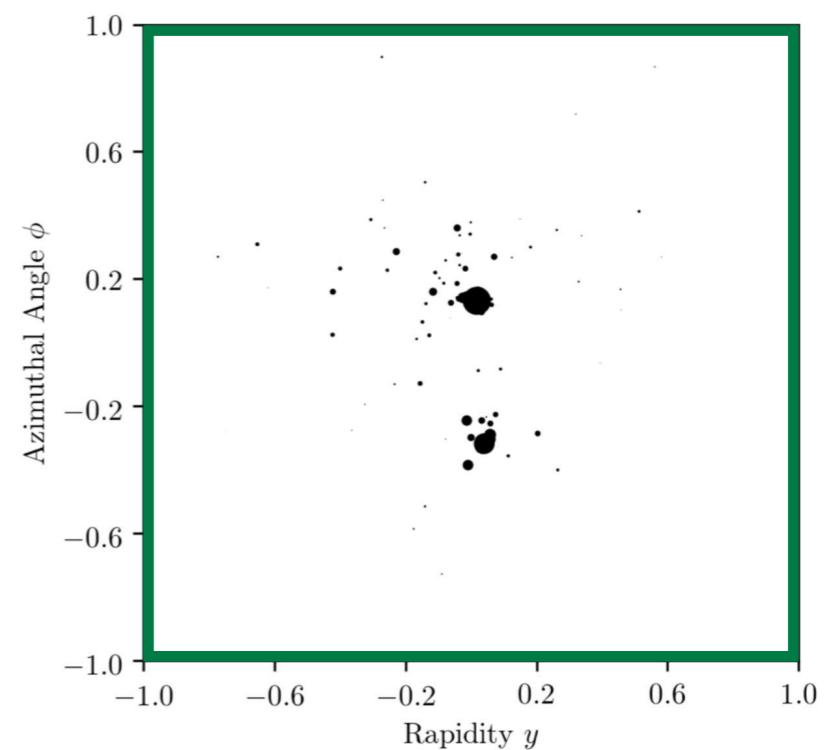
(suppressing “unsafe” charge/flavor information)



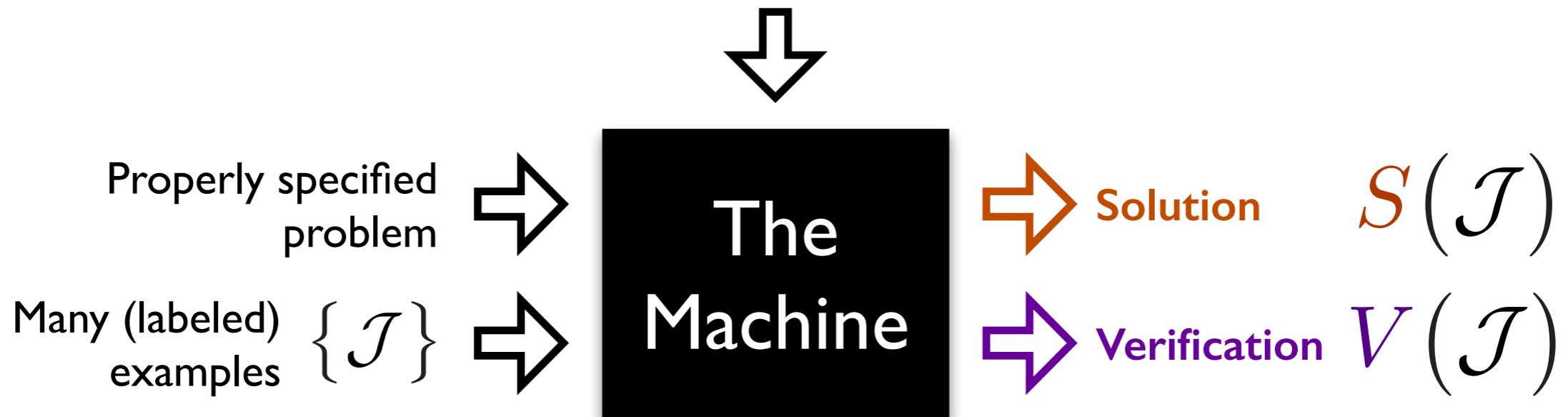
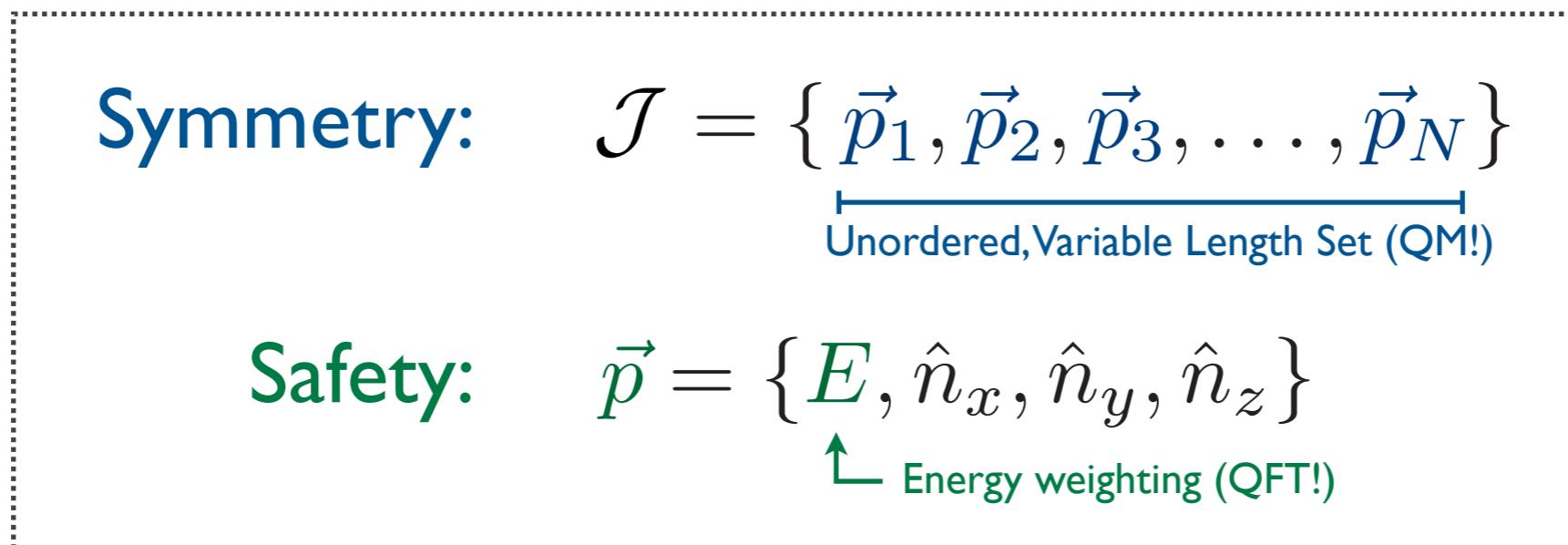
- Equivalently: Energy Density

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

↑
Energy ↑
 Direction



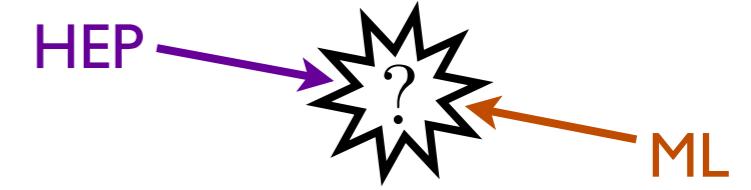
“Thinking” Like a Physicist



*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 **Linear weights (i.e. safe)**

↑ ↓

Parametrized with **Neural Networks**

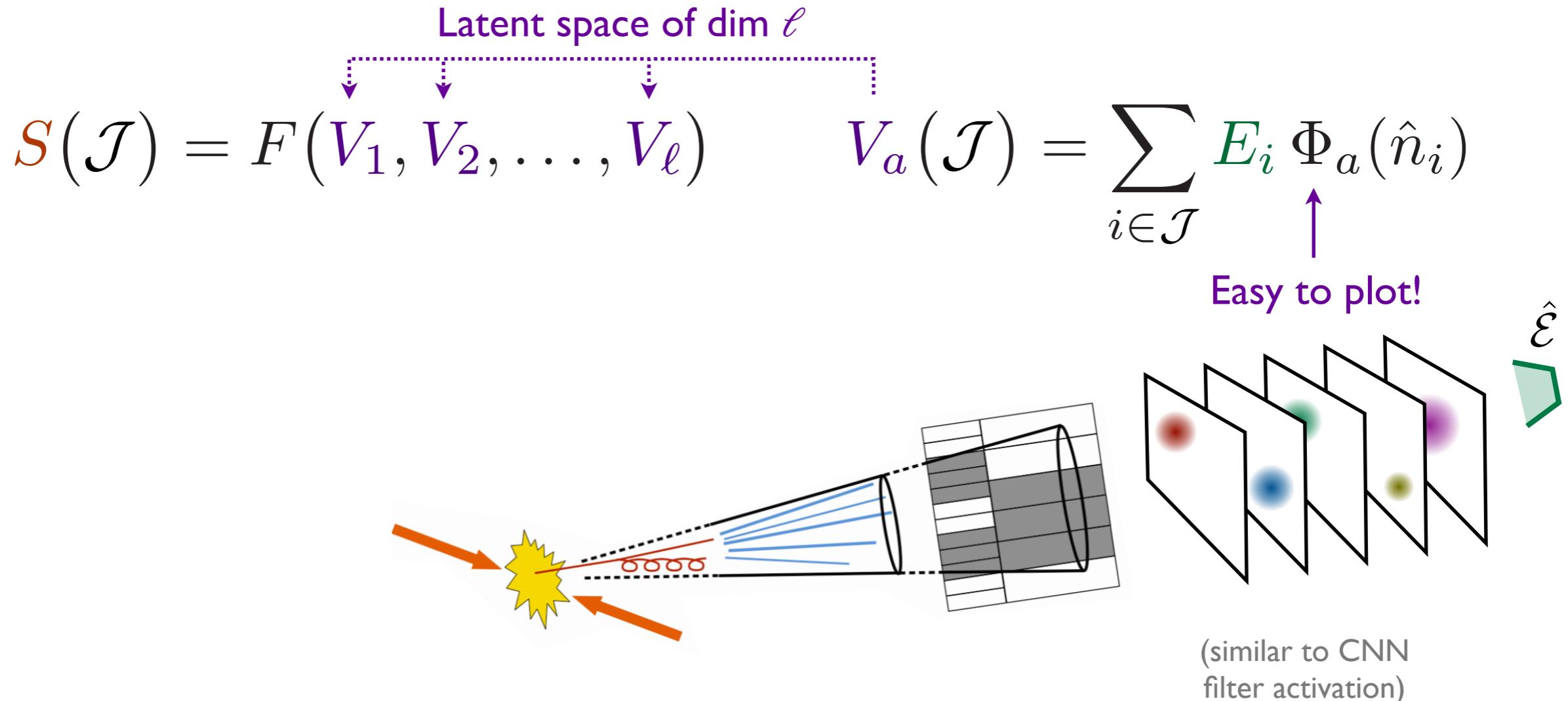
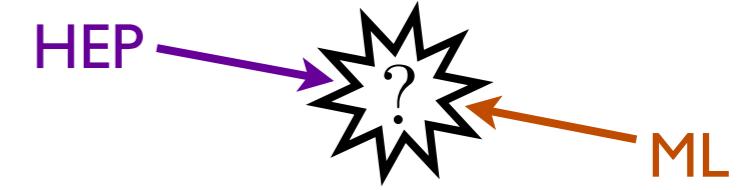
Provably describes any **safe** observable (!)
Excellent jet classification performance*

[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](#); other set-based architecture in Qu, Gouskos, [PRD 2020](#); Mikuni, Canelli, [EPJP 2020](#); Dolan, Ore, [PRD 2021](#); graph-based approach in Moreno, Cerri, Duarte, Newman, Nguyen, Periwal, Pierini, Serikova, Spiropulu, Vlimant, [EPJC 2020](#); Lorentz-equivariant approach in Bogatskiy, Anderson, Offermann, Roussi, Miller, Kondor, [arXiv 2020](#); histogram pooling in Cranmer, Kreisch, Pisani, Villaescusa-Navarro, Spergel, Ho, [ICLR SimDL 2021](#)]



Energy Flow Networks

Architecture designed around symmetries and *interpretability*

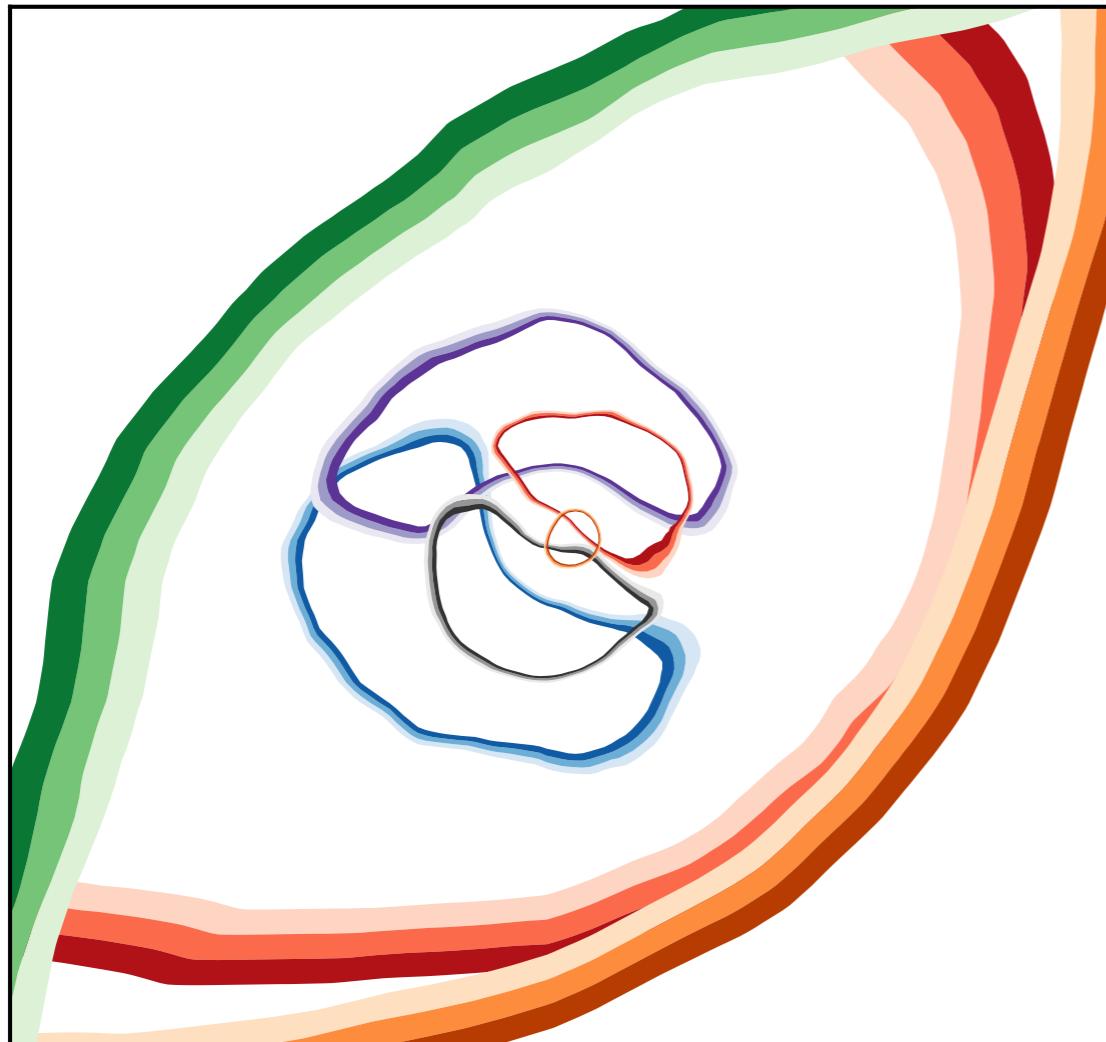


[Komiske, Metodiev, JDT, [JHEP 2019](#); see also Komiske, Metodiev, JDT, [JHEP 2018](#); code at [energyflow.network](#); special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, [NIPS 2017](#); other set-based architecture in Qu, Gouskos, [PRD 2020](#); Mikuni, Canelli, [EPJP 2020](#); Dolan, Ore, [PRD 2021](#); graph-based approach in Moreno, Cerri, Duarte, Newman, Nguyen, Periwal, Pierini, Serikova, Spiropulu, Vlimant, [EPJC 2020](#); Lorentz-equivariant approach in Bogatskiy, Anderson, Offermann, Roussi, Miller, Kondor, [arXiv 2020](#); histogram pooling in Cranmer, Kreisch, Pisani, Villaescusa-Navarro, Spergel, Ho, [ICLR SimDL 2021](#)]

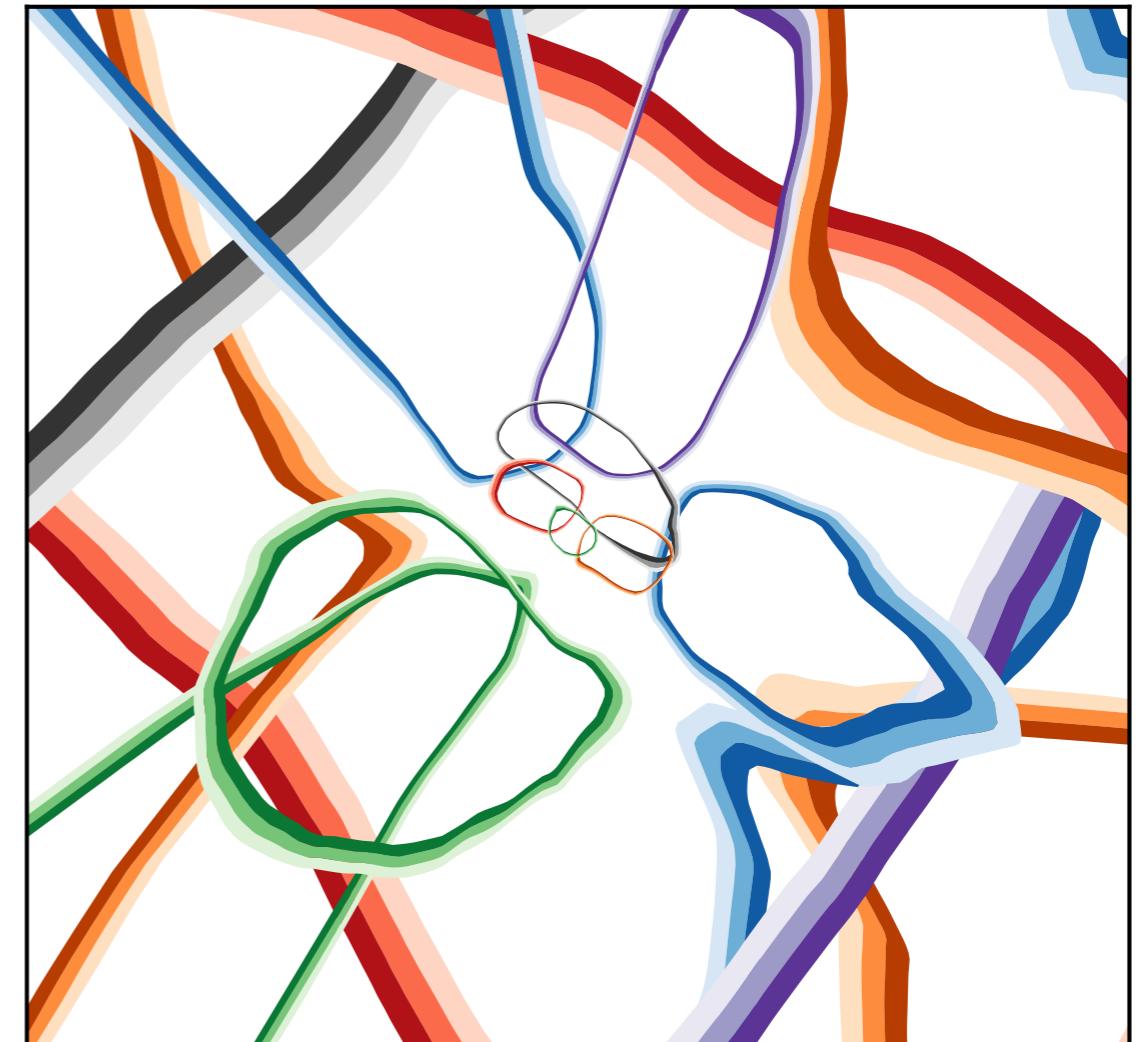


Psychedelic Network Visualization

Latent Dimension 8



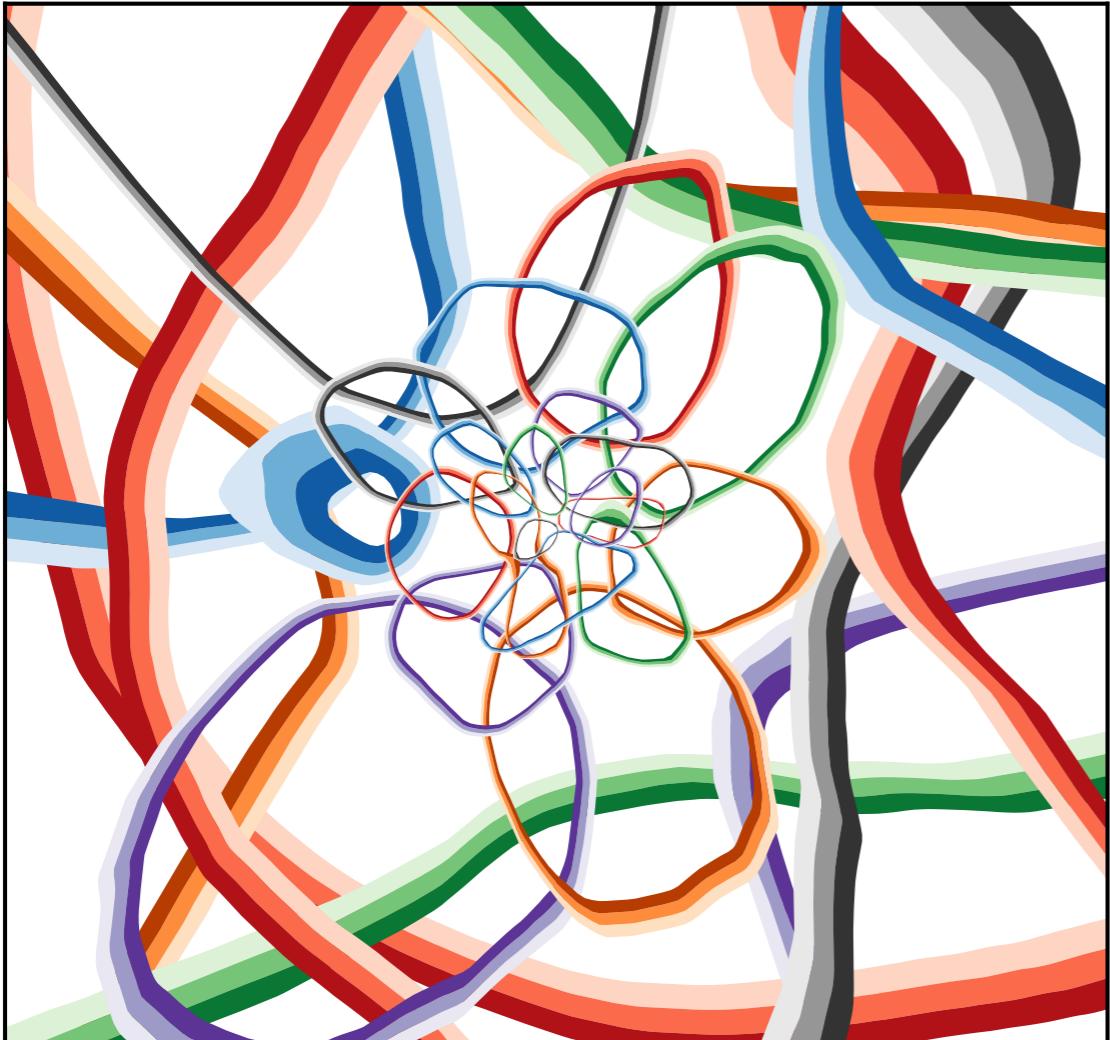
Latent Dimension 16



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

Psychedelic Network Visualization

Latent Dimension 32

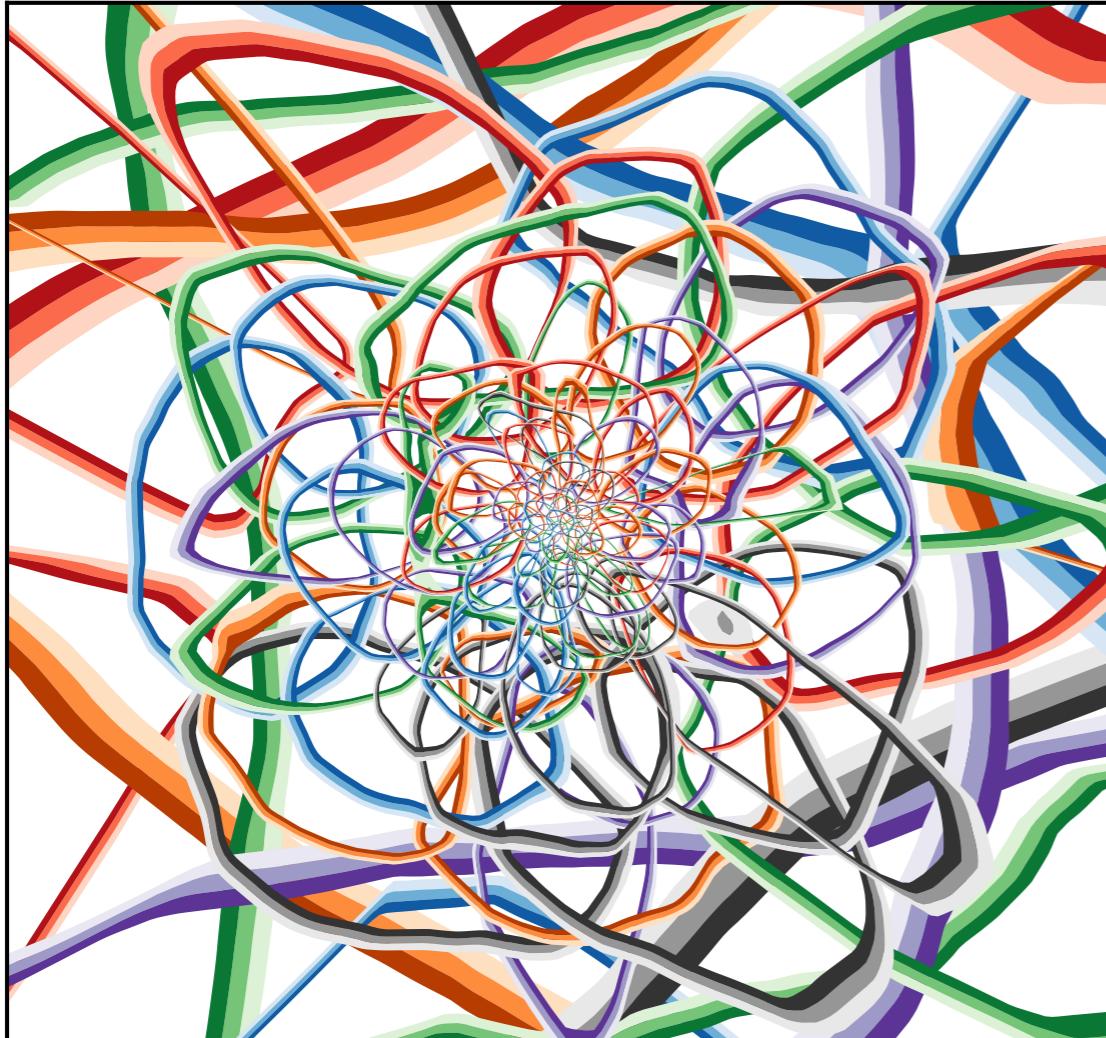


Latent Dimension 64

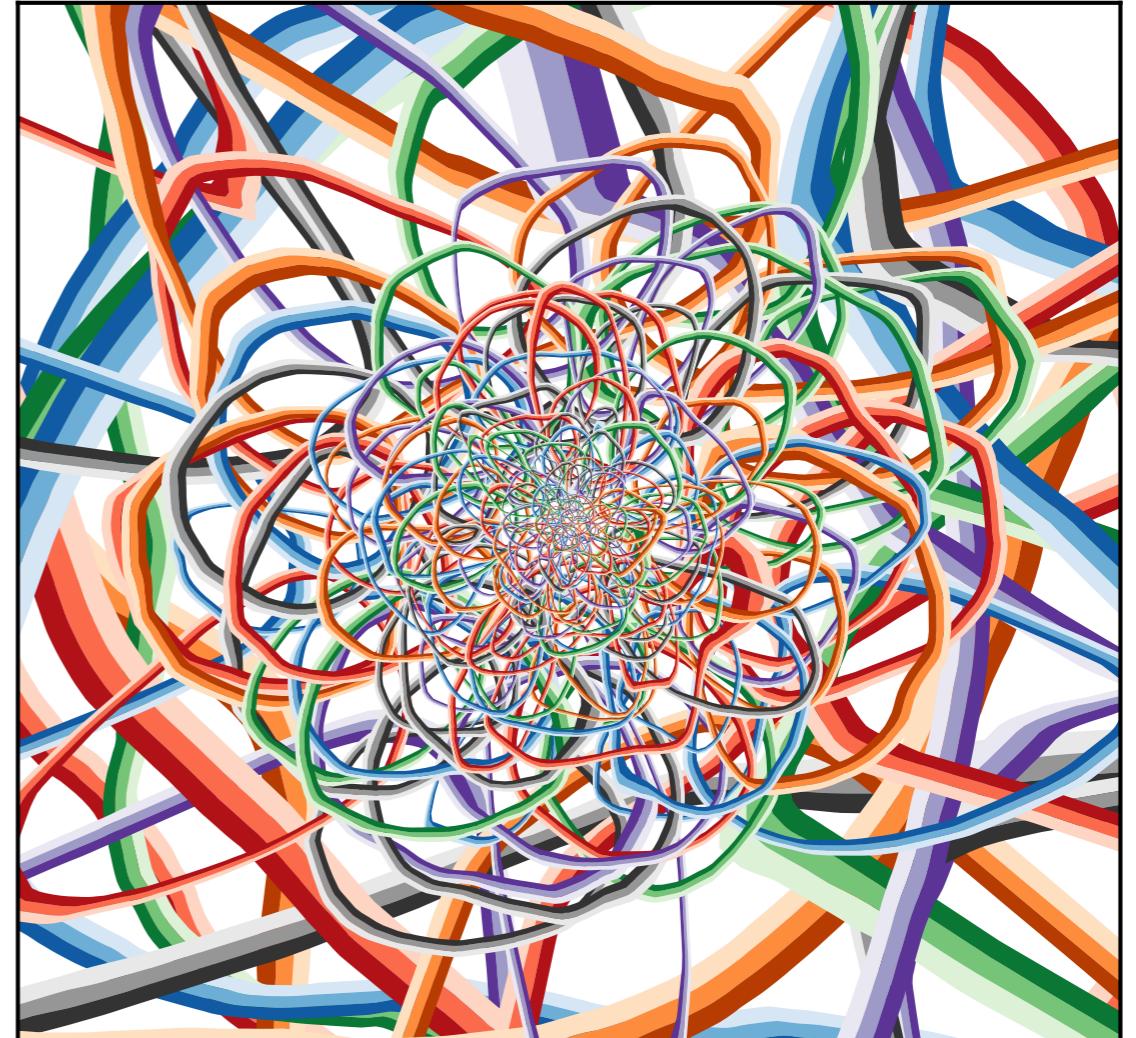


Psychedelic Network Visualization

Latent Dimension 128

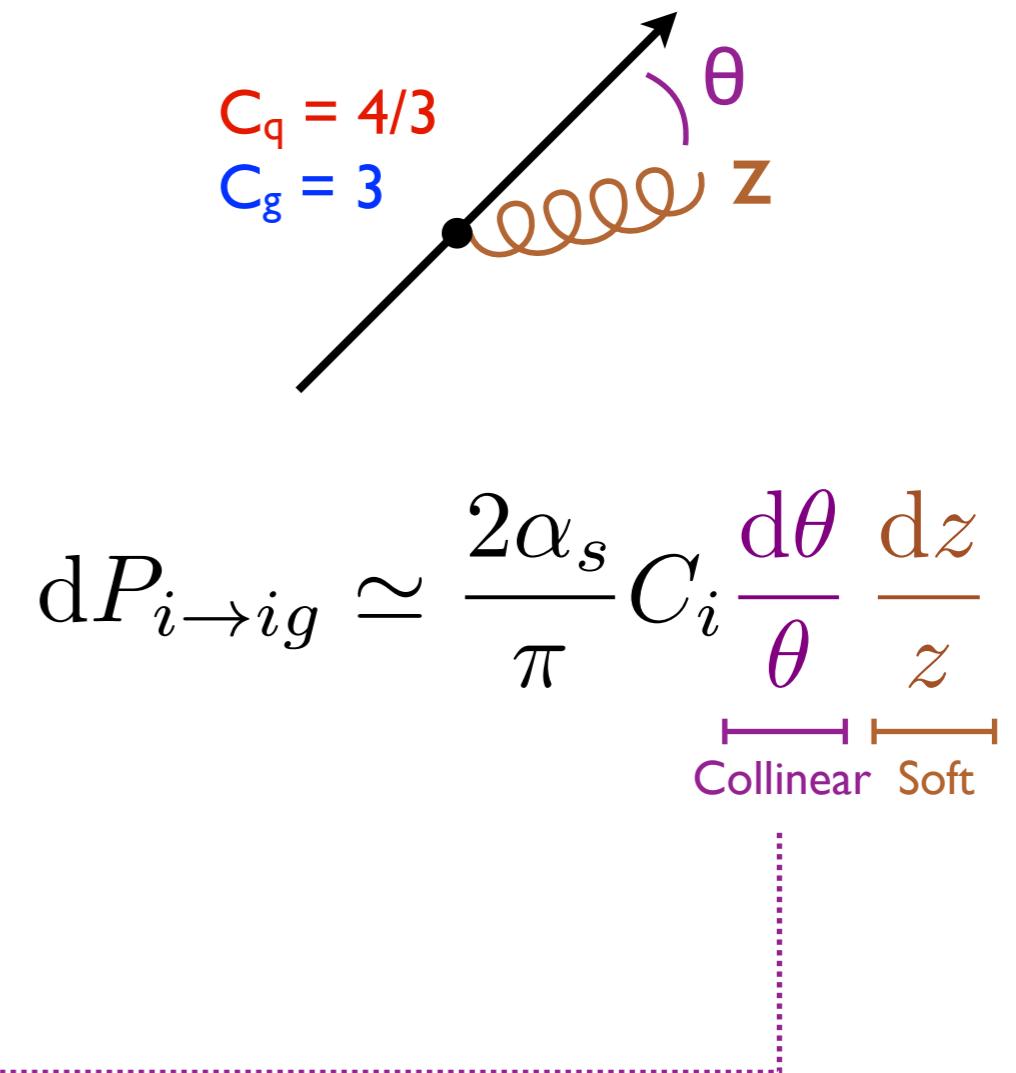
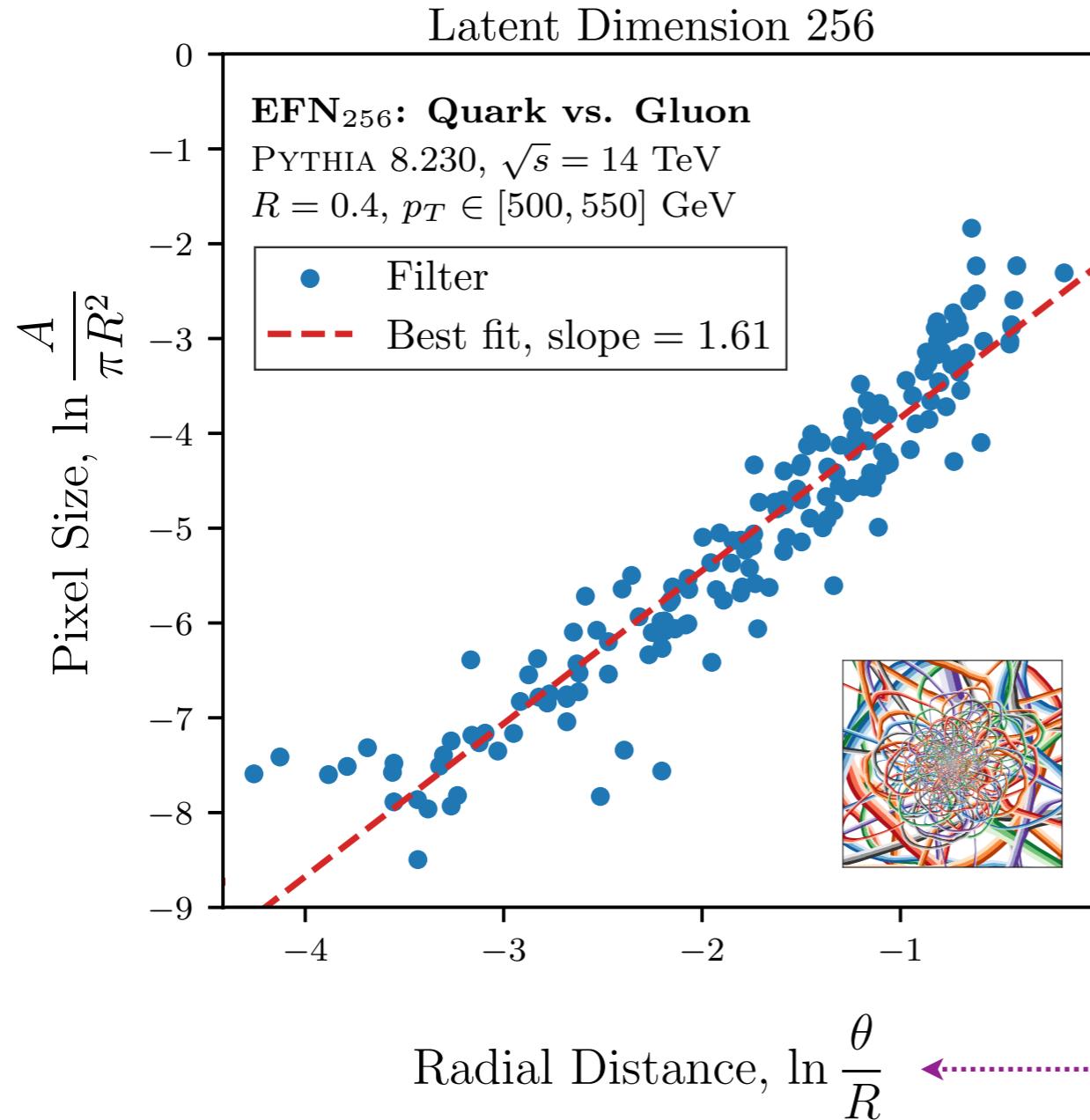


Latent Dimension 256



Singularity structure of QCD!

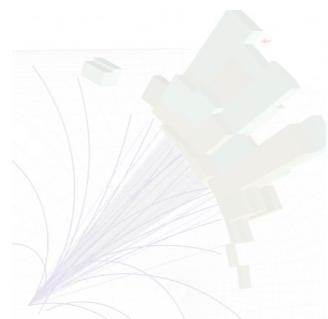
Machine Learning Collinear QCD



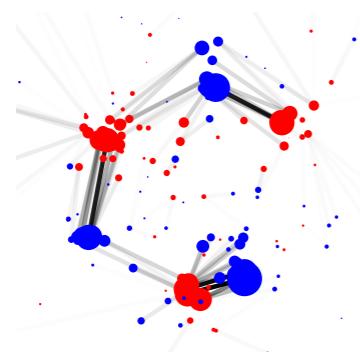
[Komiske, Metodiev, JDT, JHEP 2019]



Particle Physics 101



What is a Collider Event?

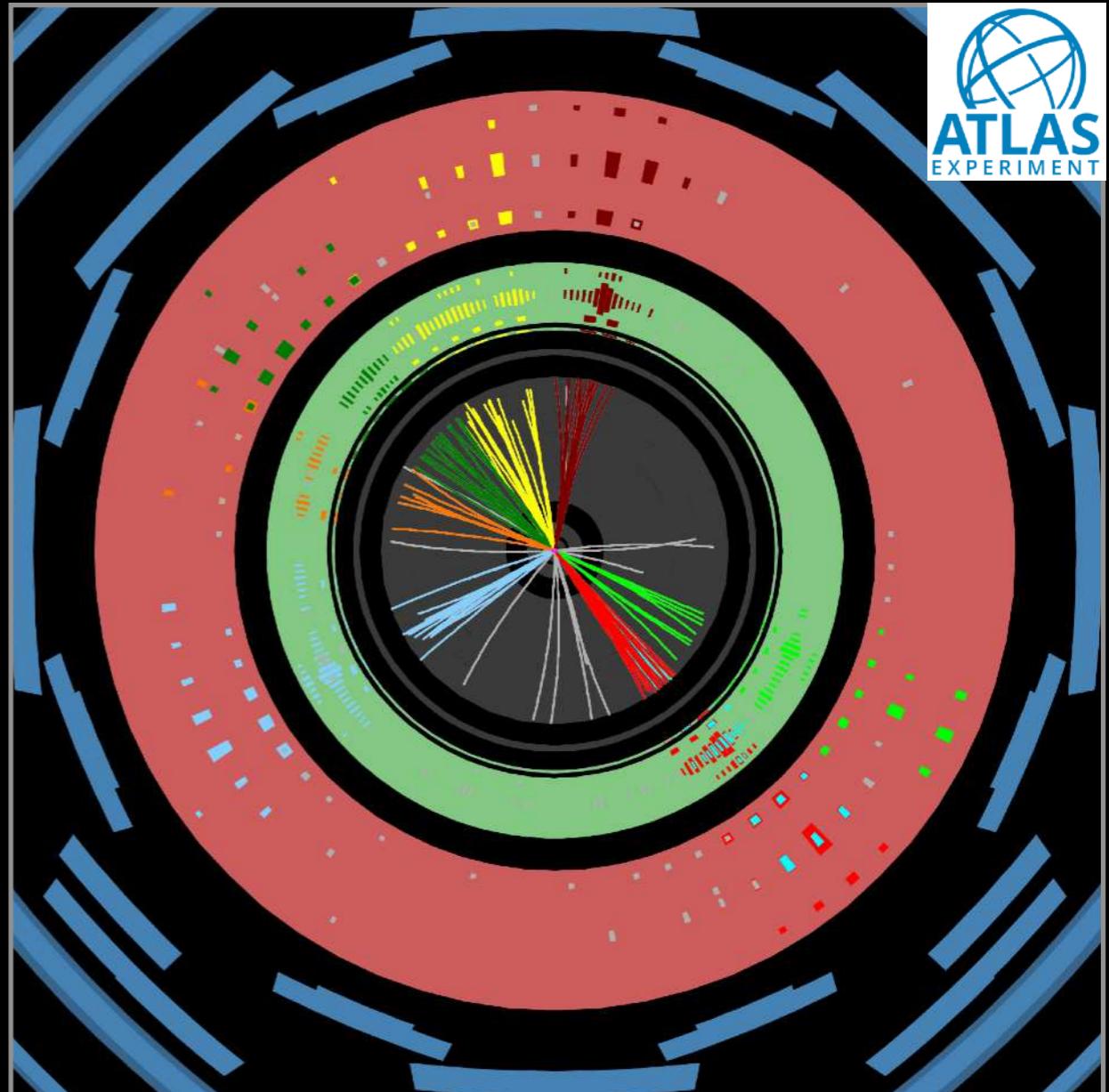
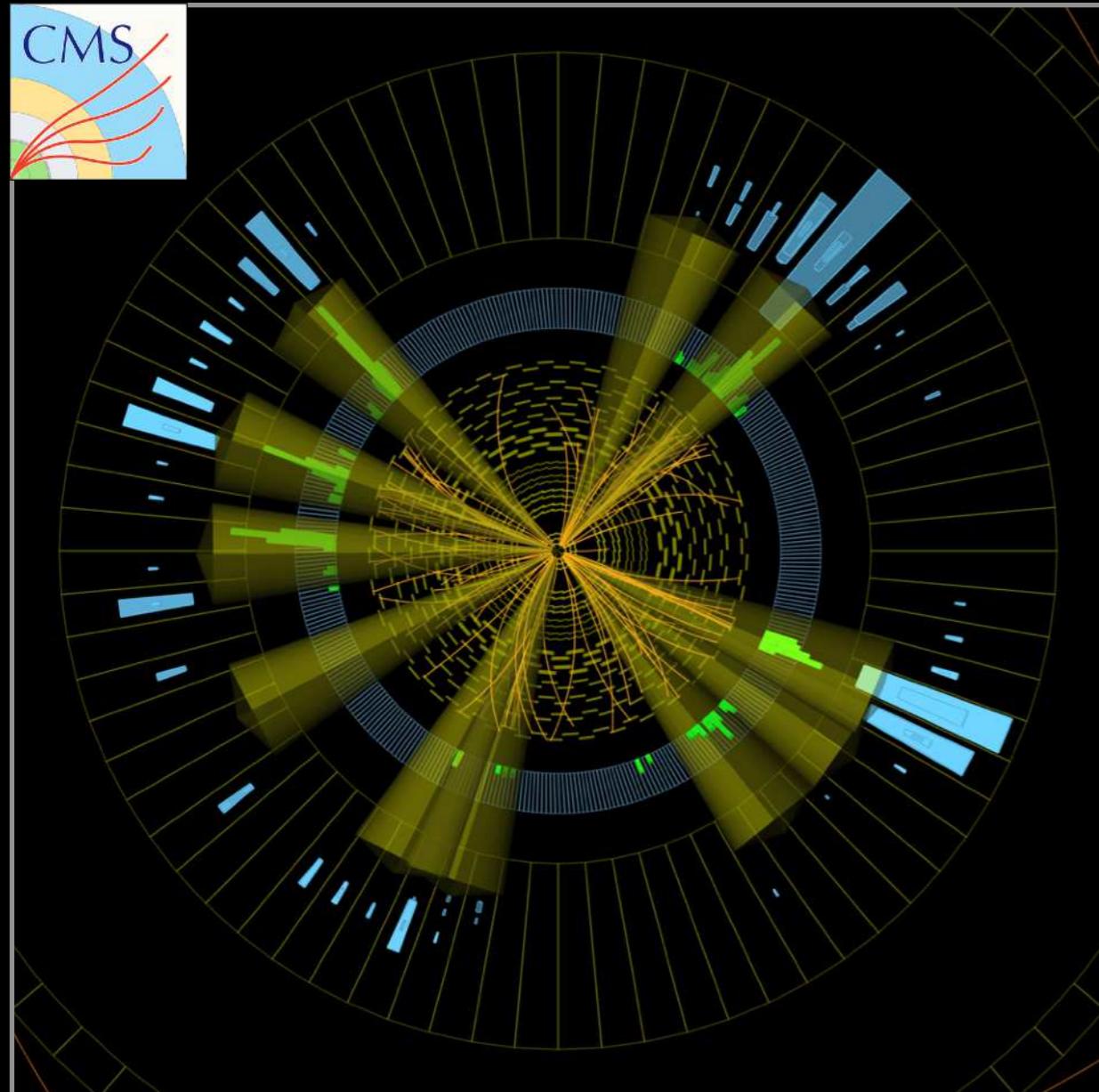


When are Collider Events Similar?

Two Collider Events

Two collections of points in (momentum) space

How “close” are these?

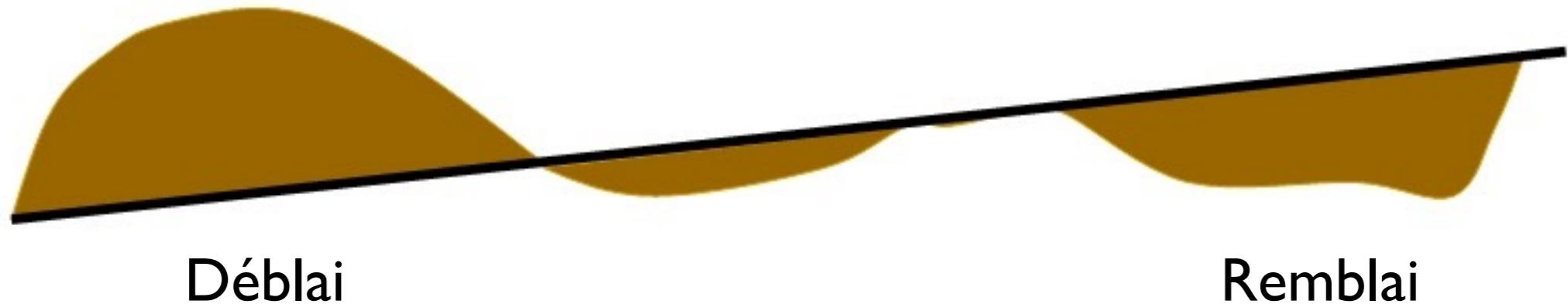


The Earth Mover's Distance

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



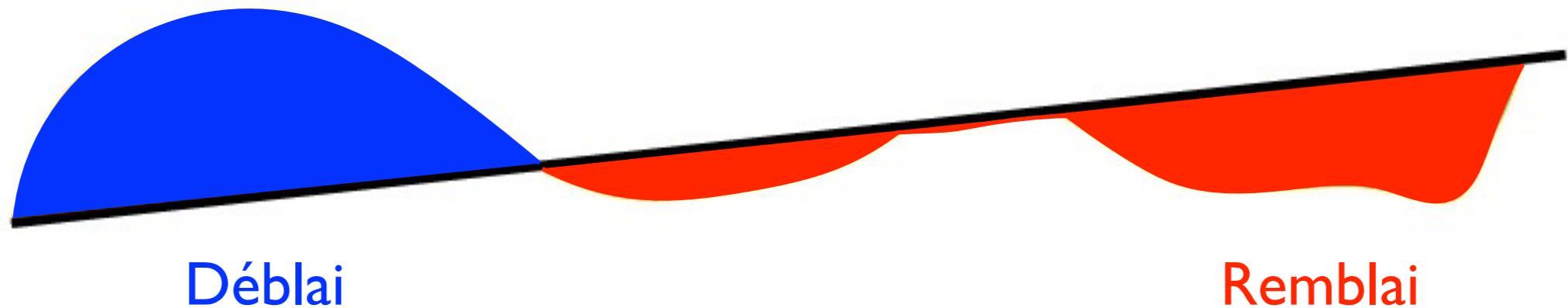
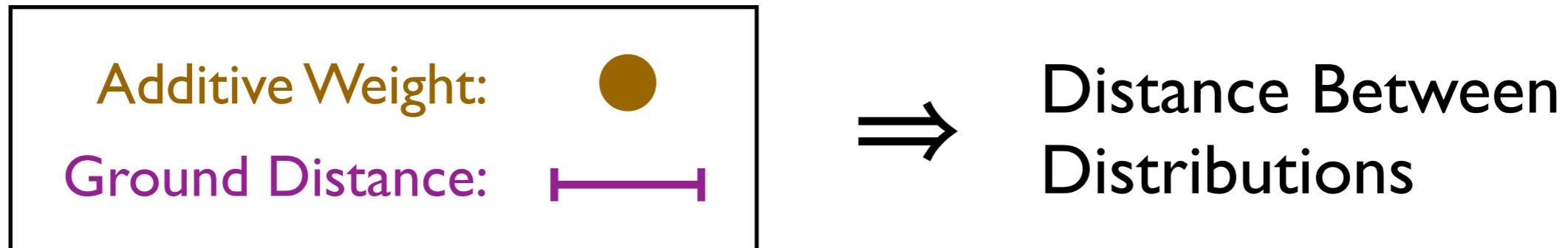
[h/t Niles-Weed, [ML4Jets 2020](#); Monge, 1781; Vaserštejn, 1969; [Wikipedia](#)]

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**



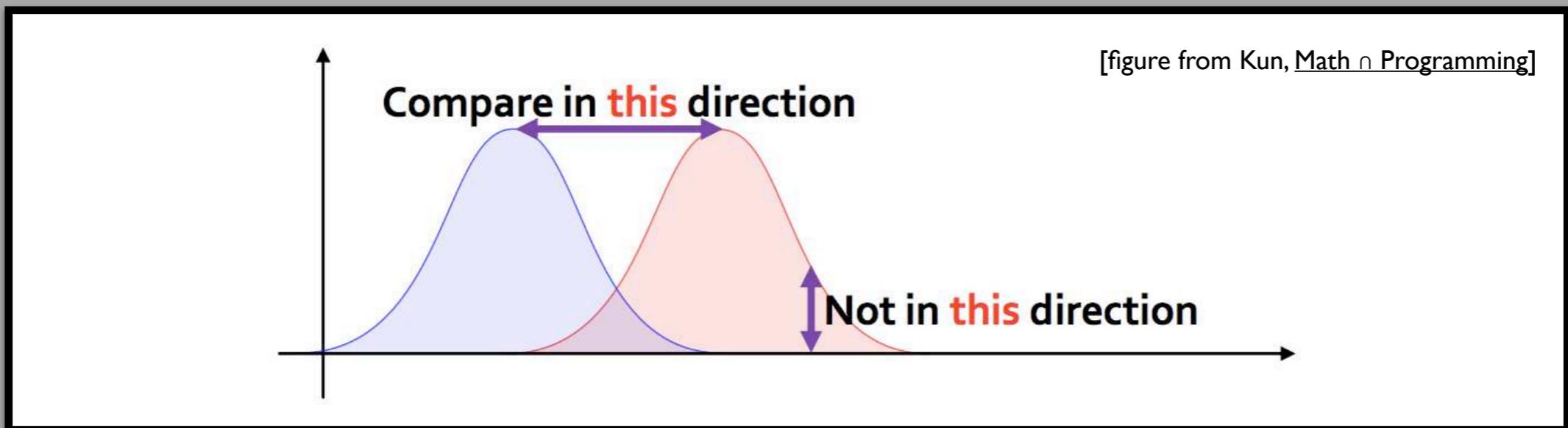
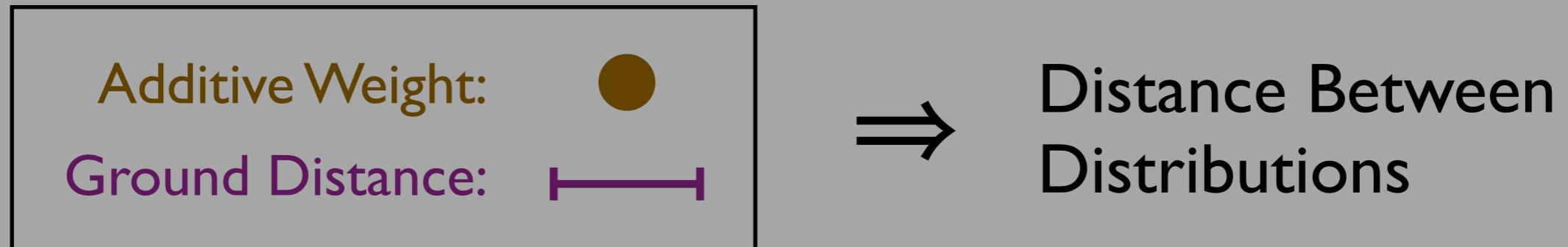
[h/t Niles-Weed, [ML4Jets 2020](#); Monge, 1781; Vaserštejn, 1969; [Wikipedia](#)]

The Earth Mover's Distance

Optimal Transport:

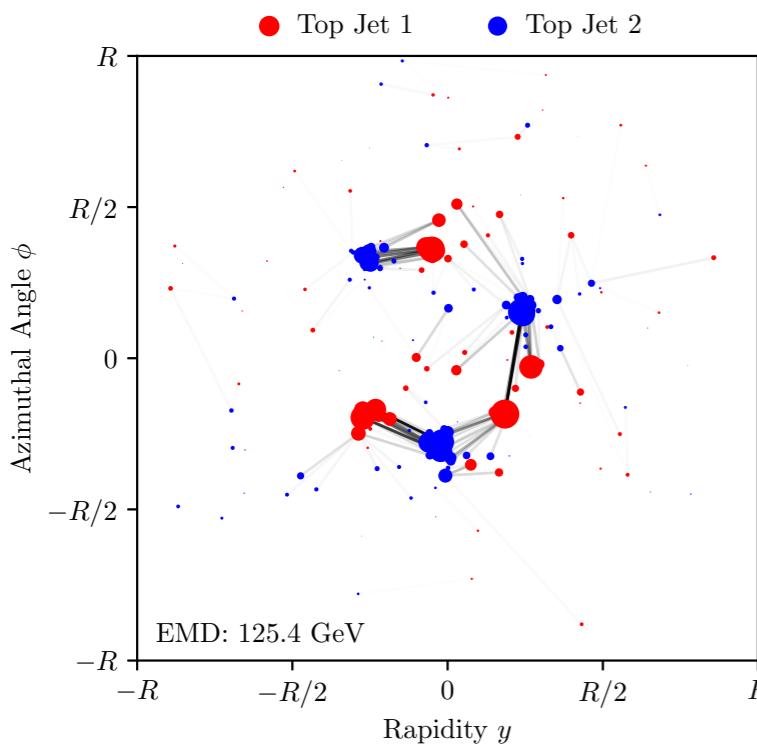
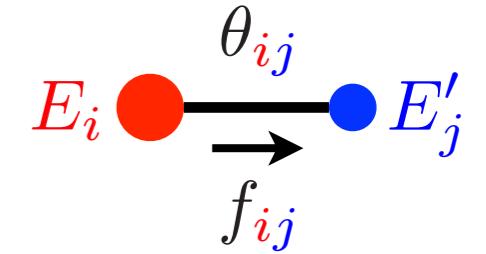
[Peleg, Werman, Rom, [IEEE 1989](#);
Rubner, Tomasi, Guibas, [ICCV 1998](#), [ICCV 2000](#);
Pele, Werman, [ECCV 2008](#); Pele Taskar, [GSI 2013](#)]

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 Energy Mover's Distance



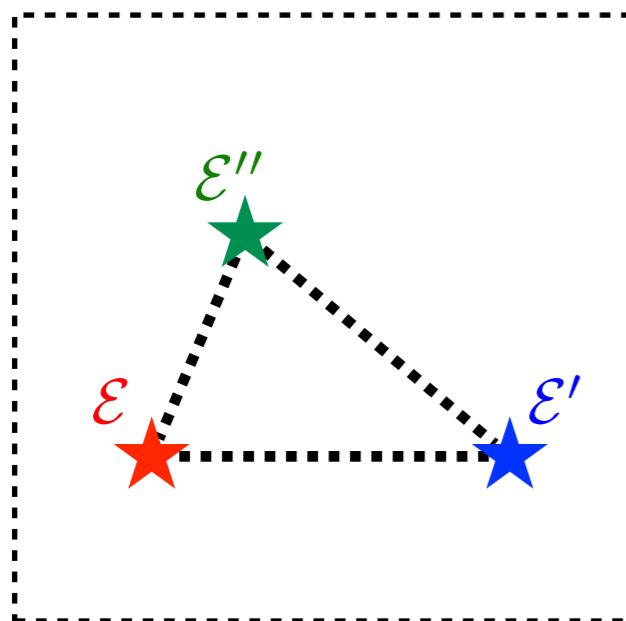
Optimal transport between energy flows...

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

↑
in GeV

Cost to move energy

Cost to create energy



...defines a metric on the space of events

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

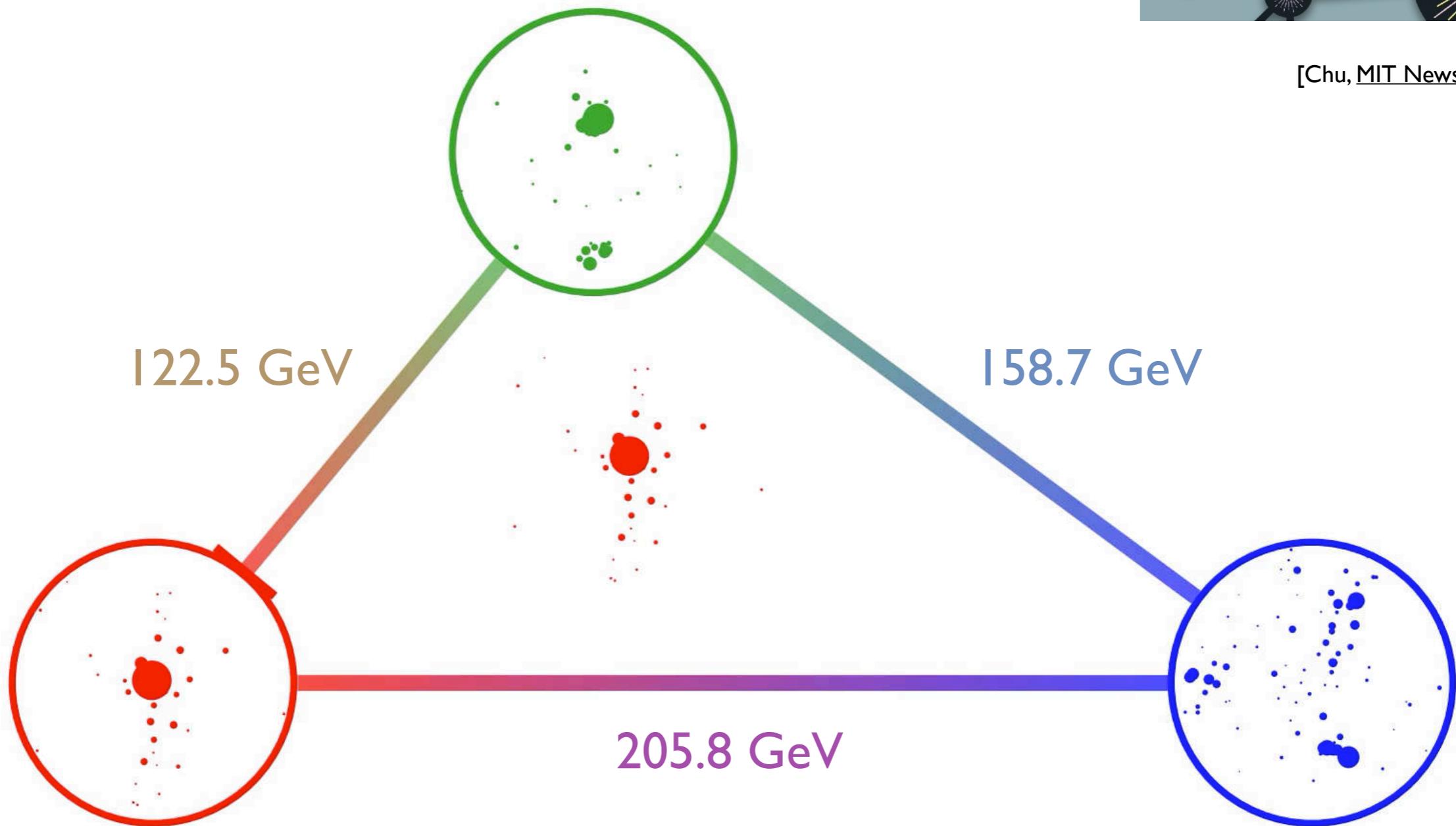
(assuming $R \geq \theta_{\max}/2$, i.e. $R \geq$ jet radius for conical jets)

[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, [EPJC 2021](#)]
[see computational speed up in Cai, Cheng, Craig, [PRD 2020](#)]

Triangulating the Space of Jets



[Chu, MIT News July 2019]



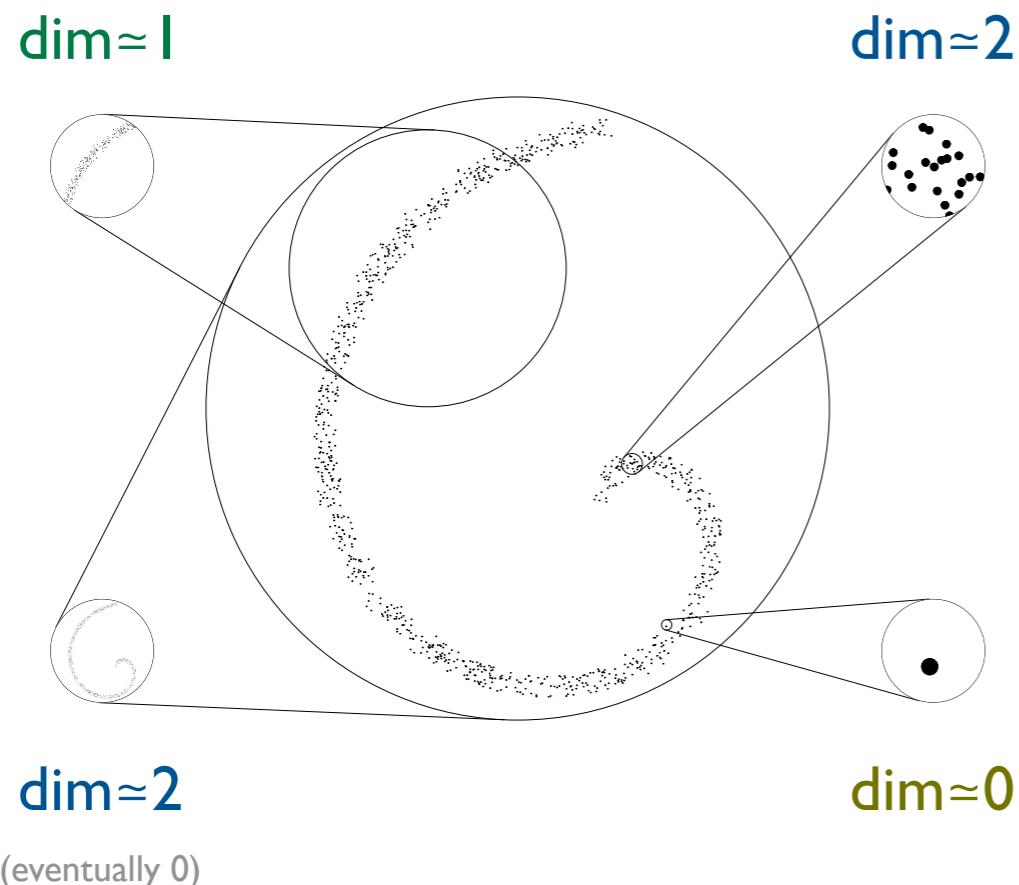
[Komiske, Metodiev, JDT, [PRL 2019](#); code at Komiske, Metodiev, JDT, [energyflow.network](#);
see alternative graph network approach in Mullin, Pacey, Parker, White, Williams, [JHEP 2021](#)]

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](#)]



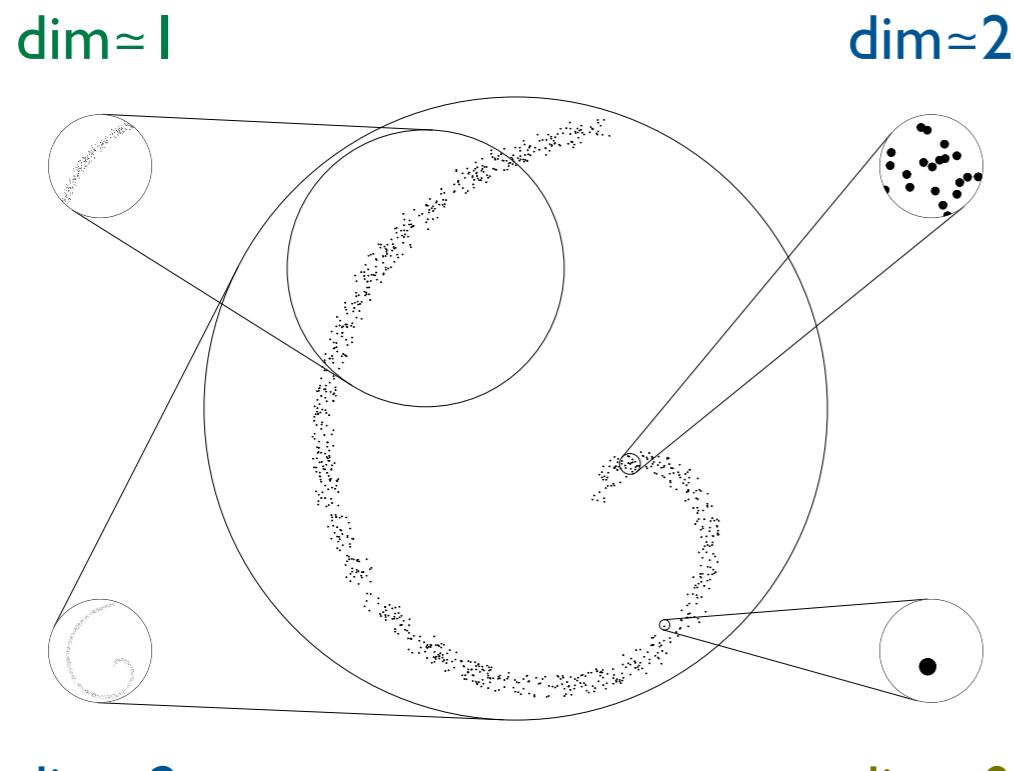
Dimensionality of Space of Jets



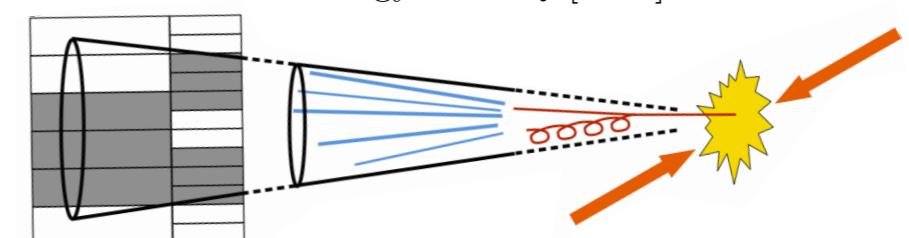
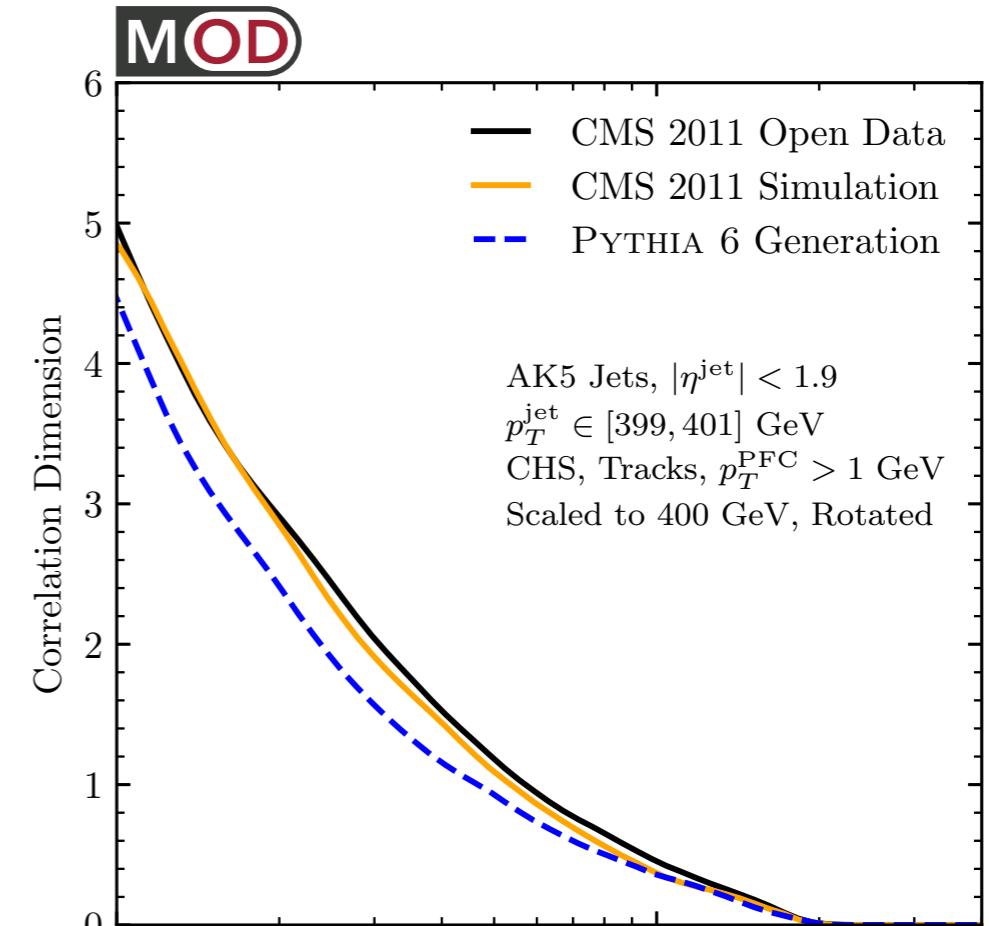
$$N_{\text{neighbors}}(r) \sim r^{\dim}$$

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

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



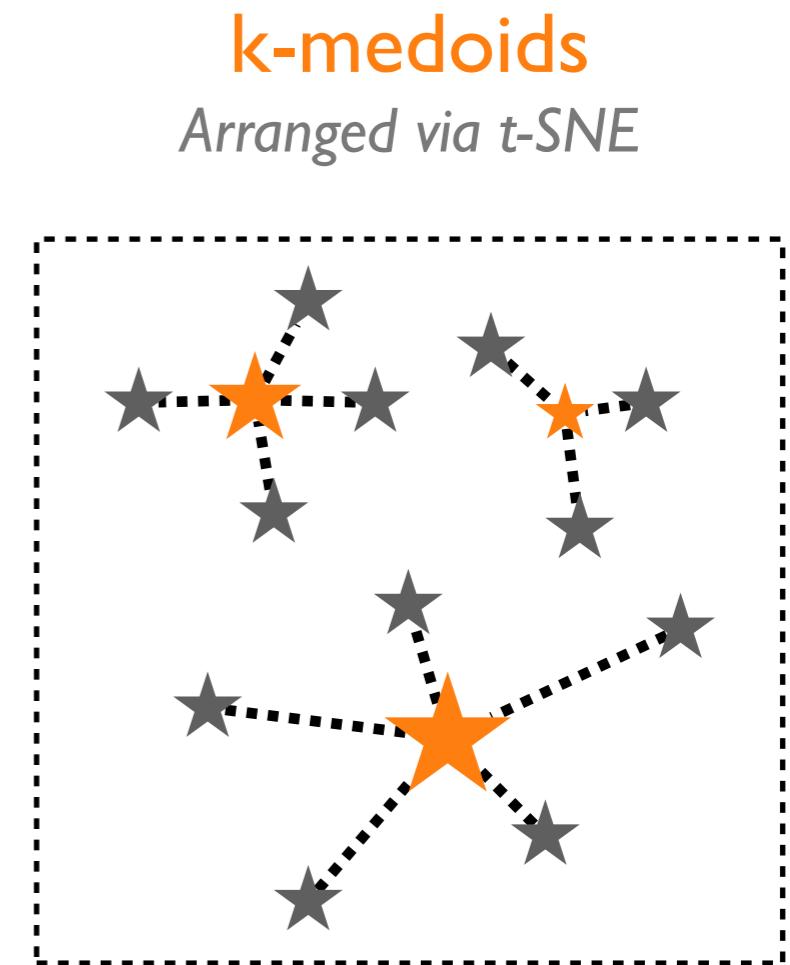
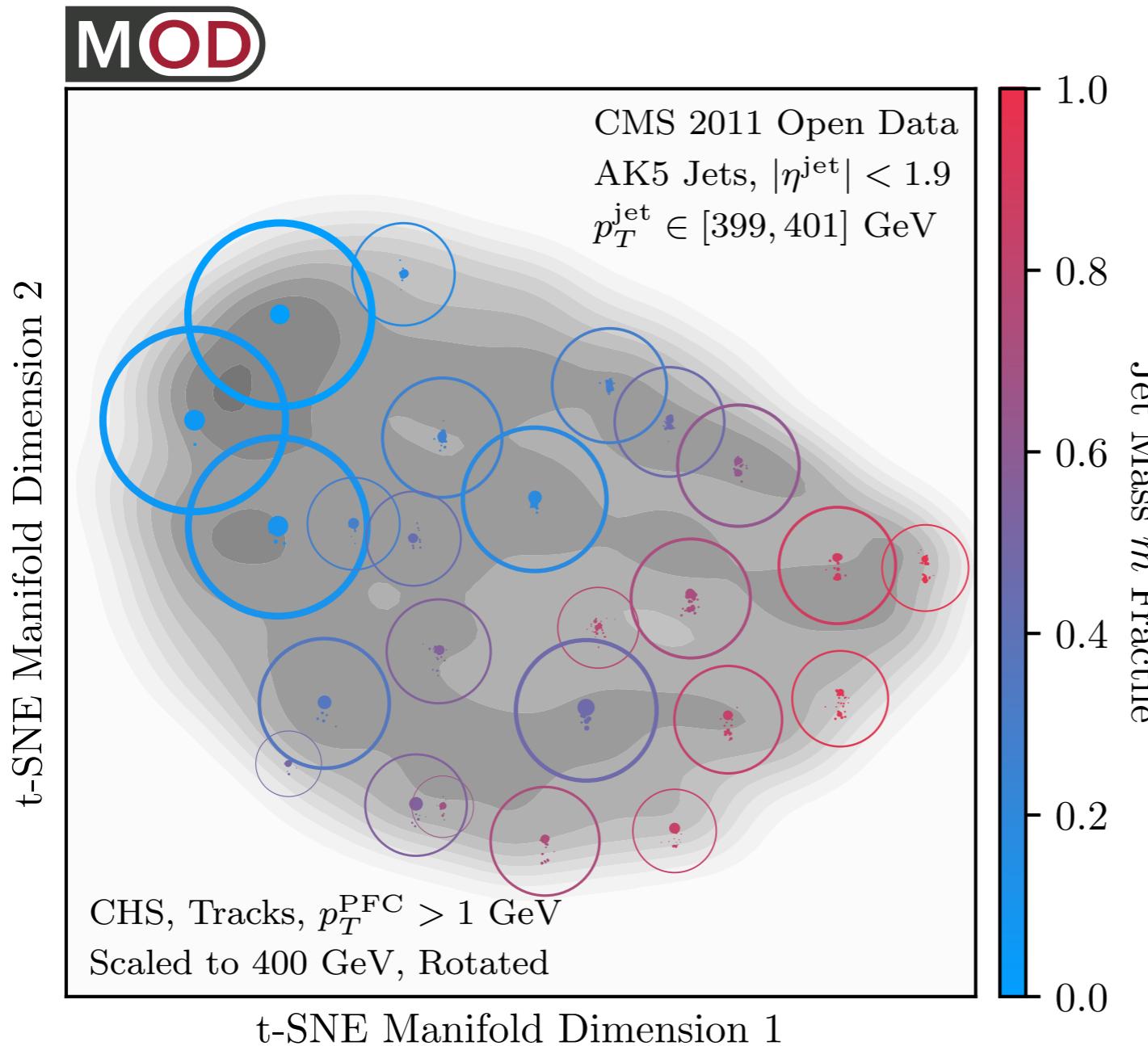
(eventually 0)



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



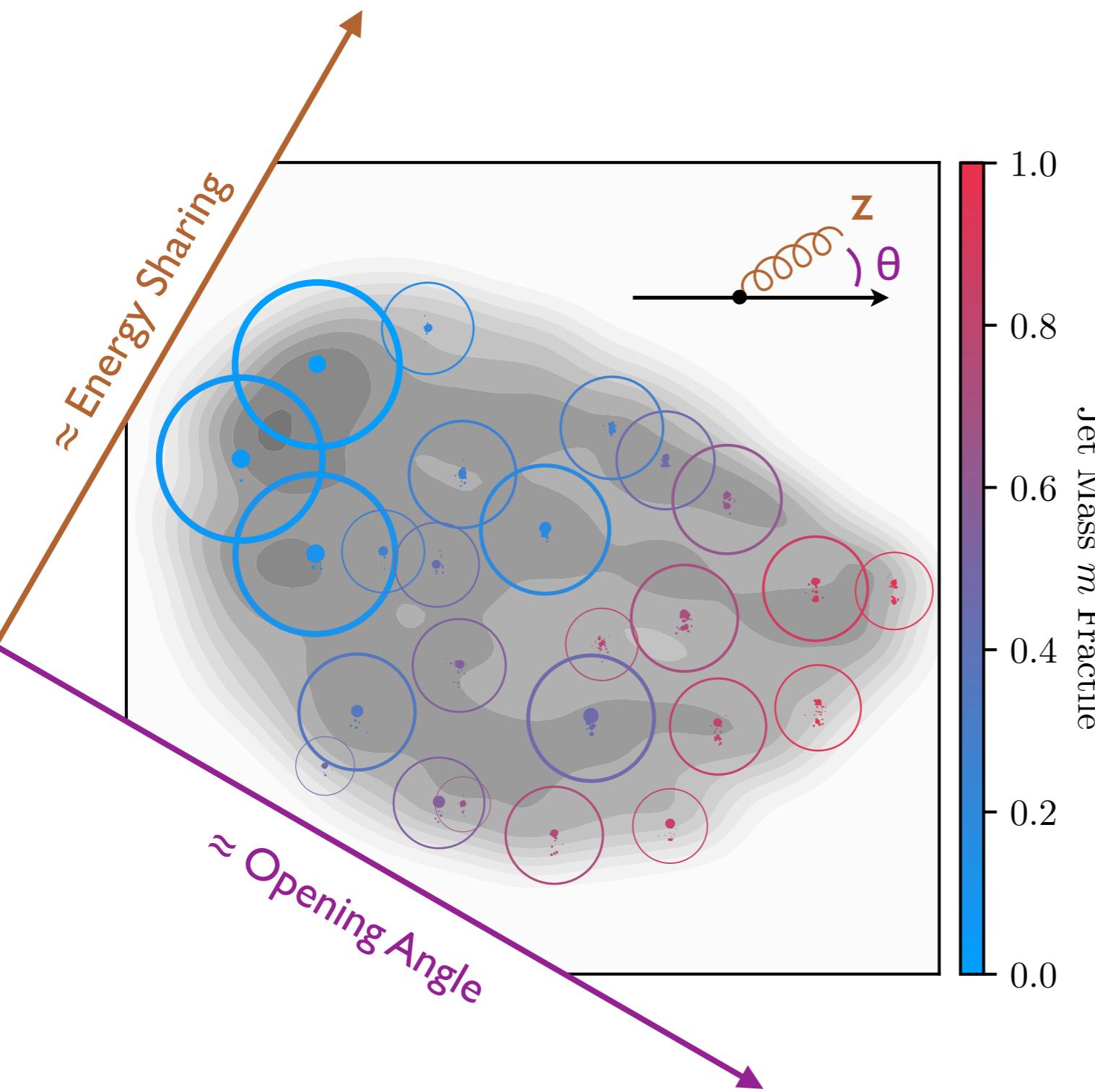
Most Representative Jets



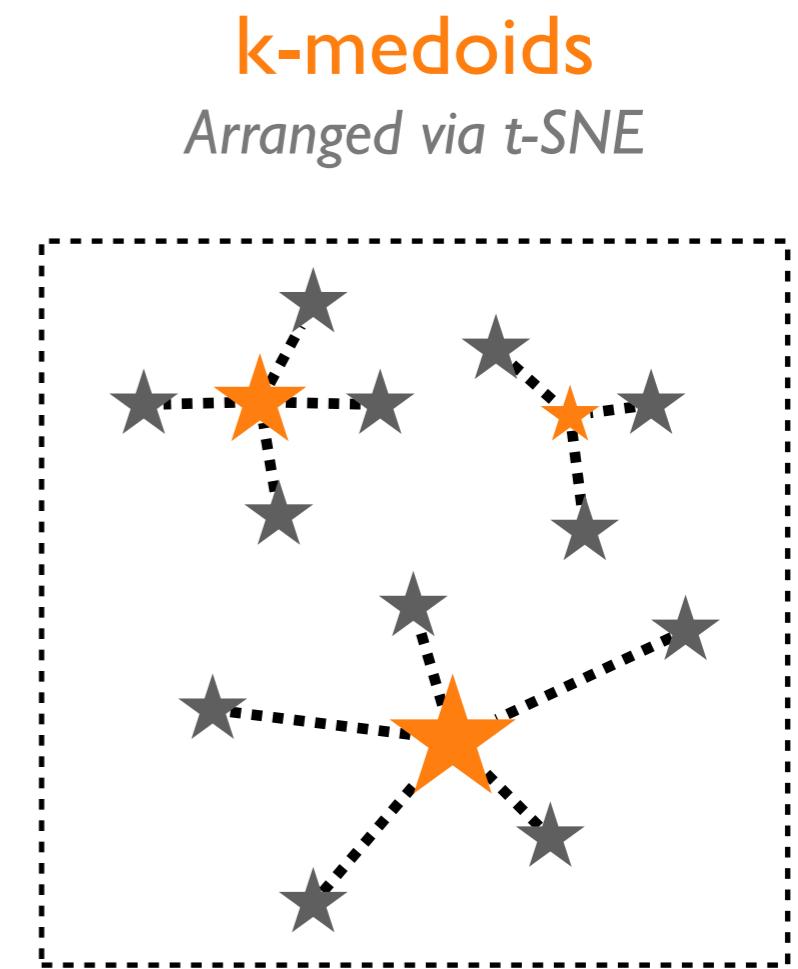
[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
using van der Maaten, Hinton, [JMLR 2008](#)]



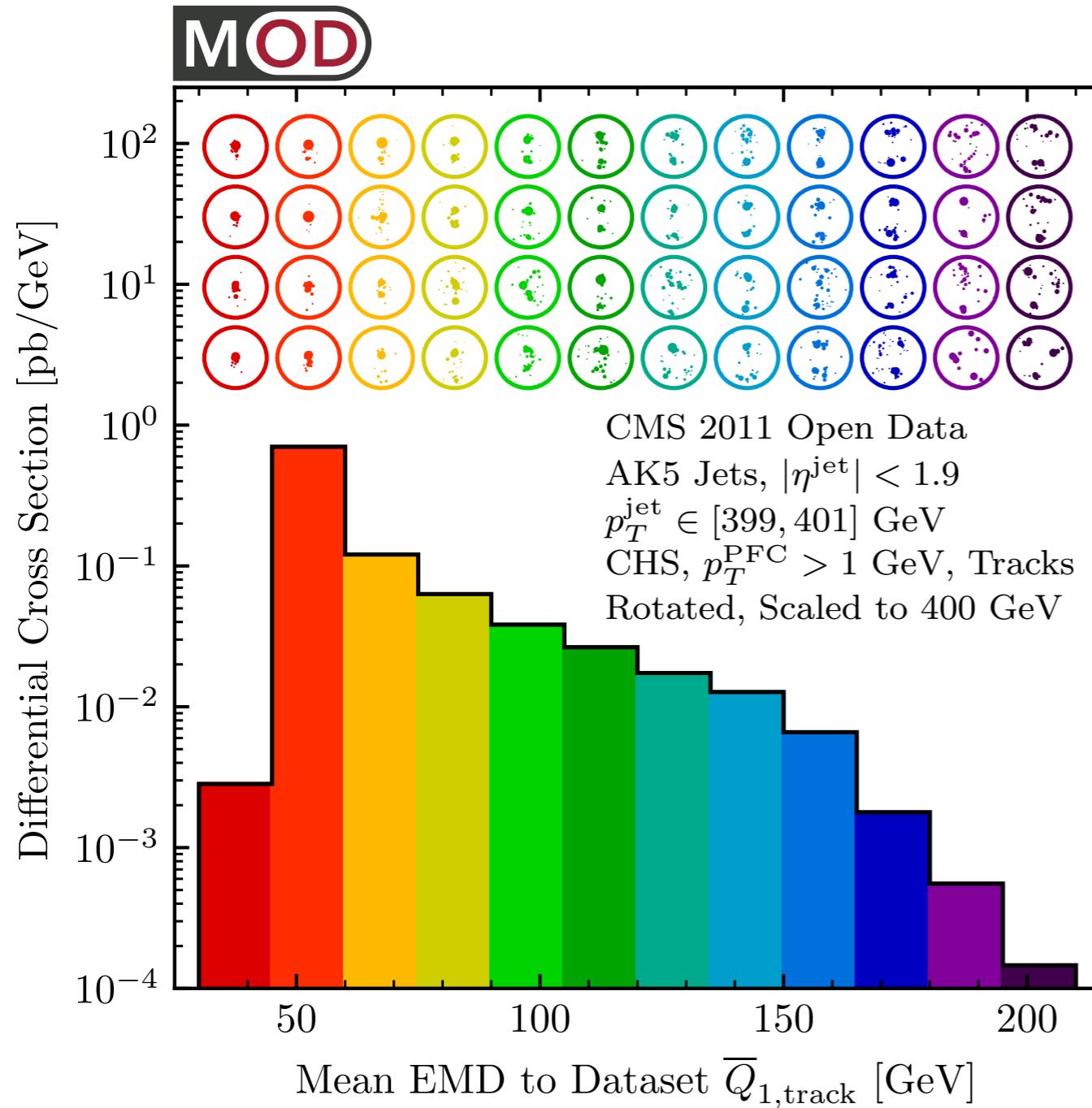
Most Representative Jets



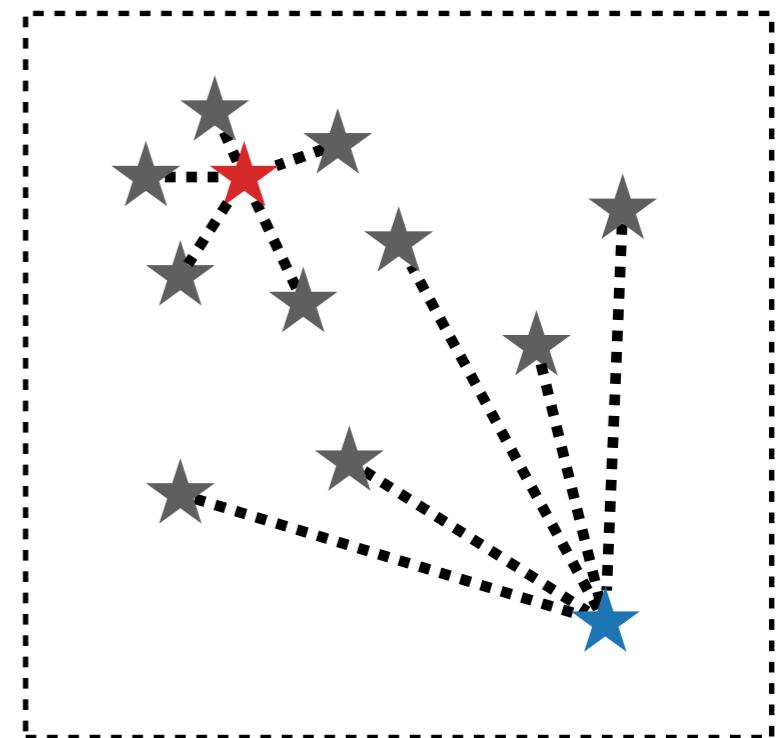
[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
using van der Maaten, Hinton, [JMLR 2008](#)]



Least Representative Jets



New Physics?
Or tails of QCD?



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

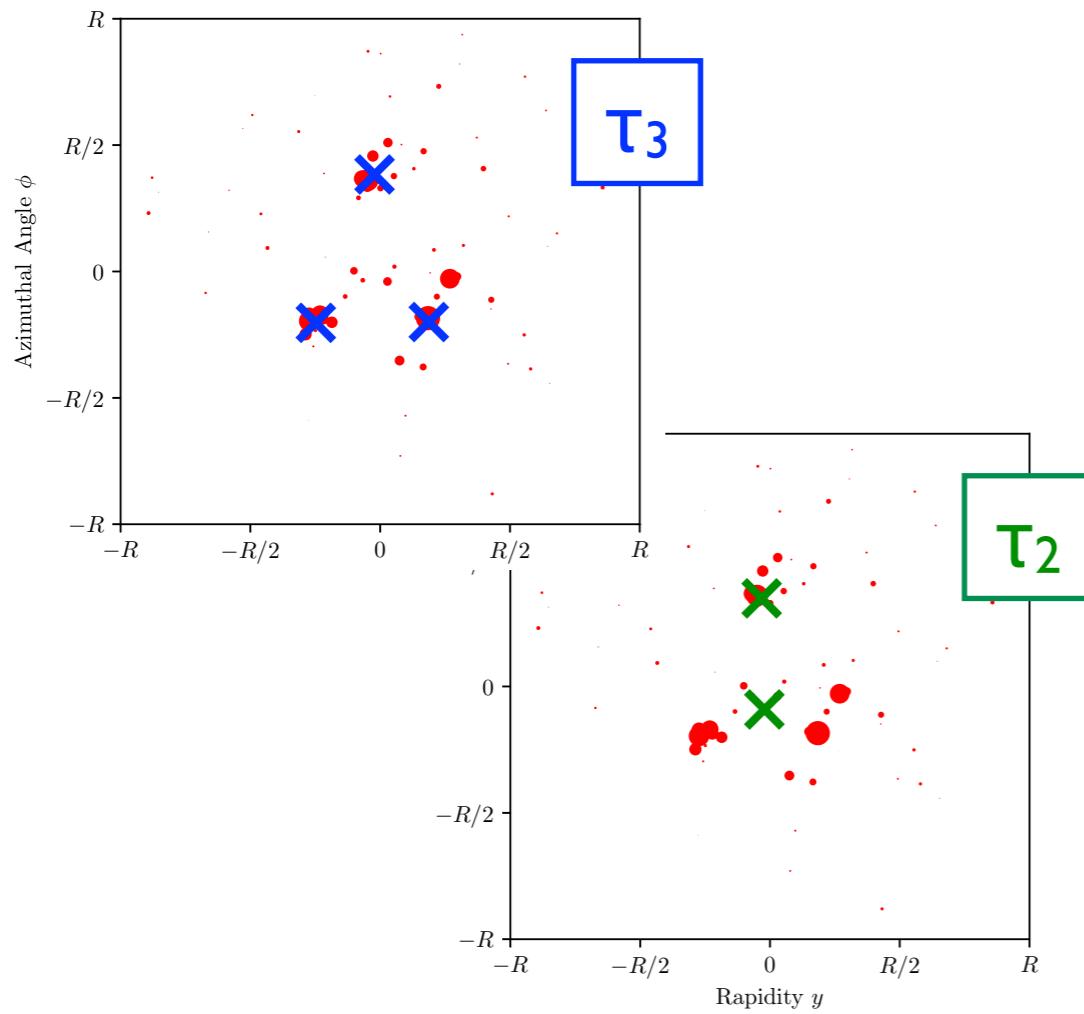


*“Sounds cool, but didn’t you already
do something like this a decade ago?”*

N-subjettiness

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

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



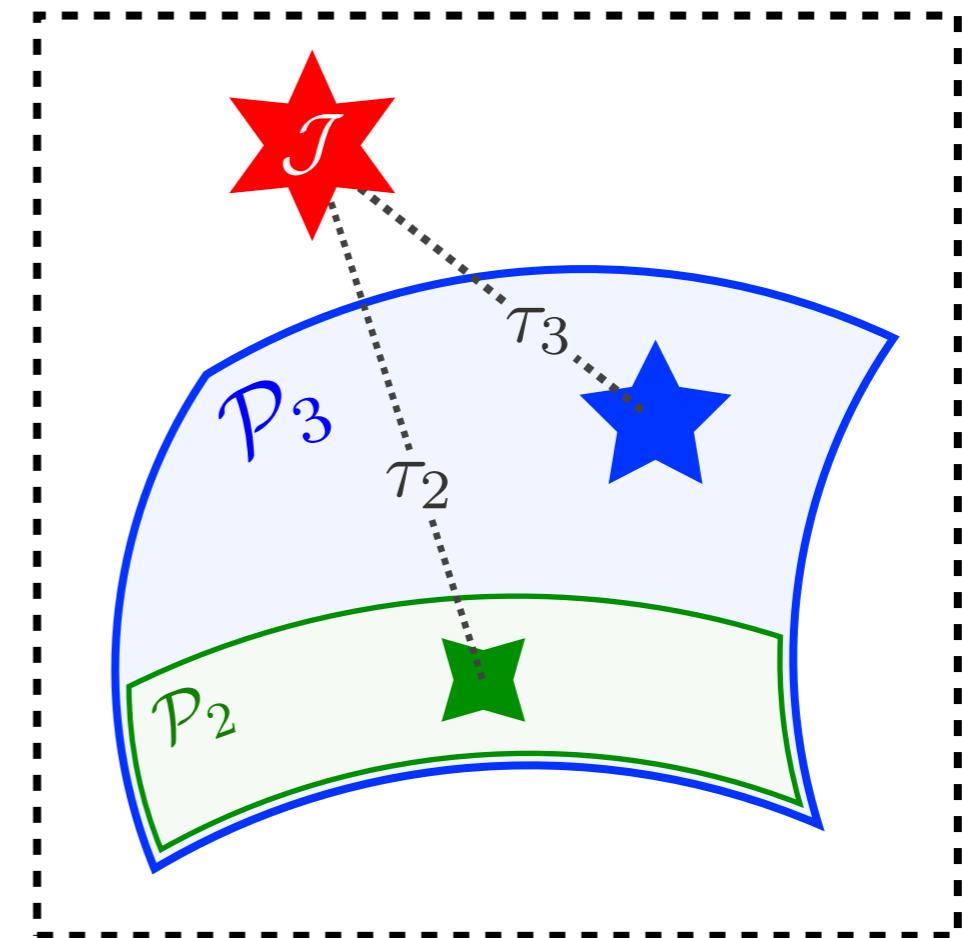
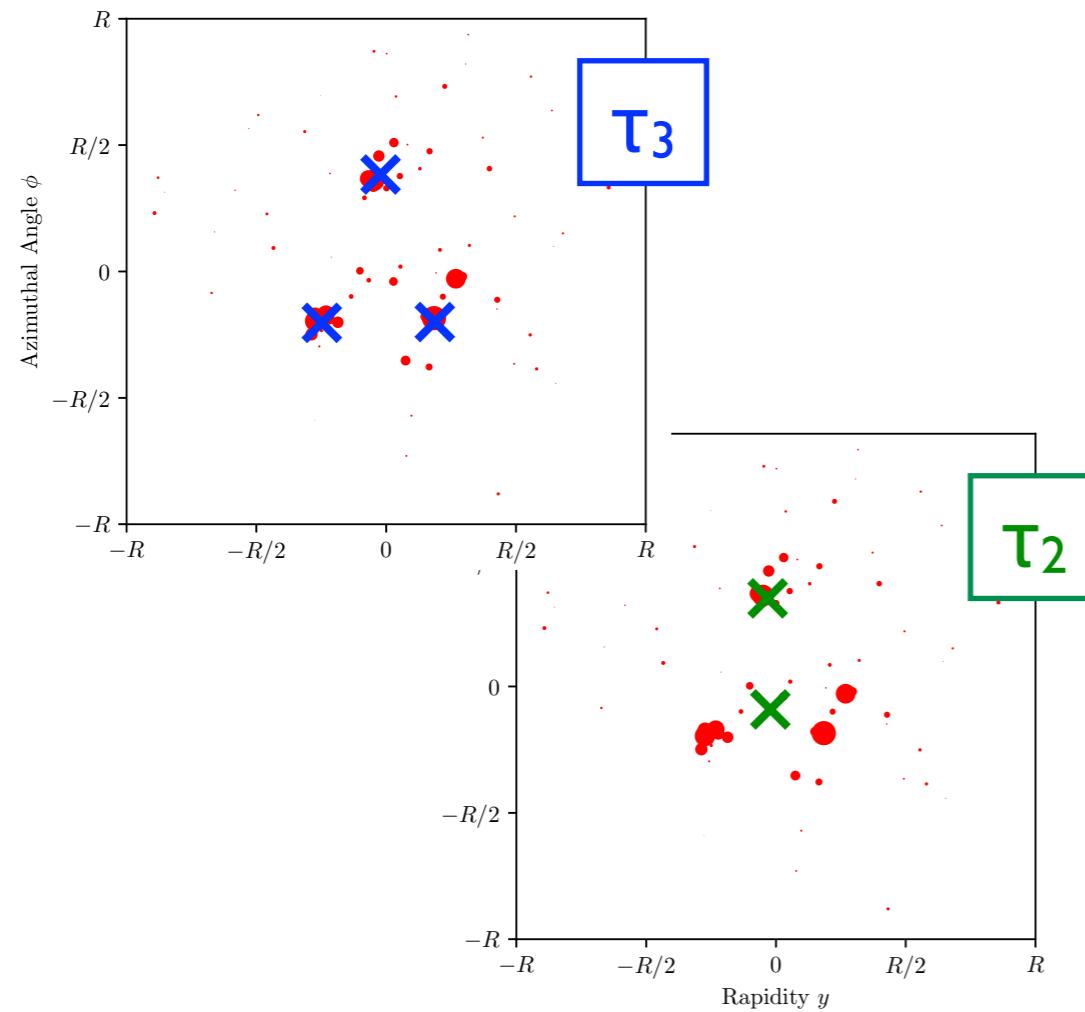
[JDT, Van Tilburg, [JHEP 2011](#), [JHEP 2012](#);
based on Brandt, Dahmen, [ZPC 1979](#); Stewart, Tackmann, Waalewijn, [PRL 2010](#)]



N-subjettiness = Point to Manifold EMD

...is secretly an optimal transport problem

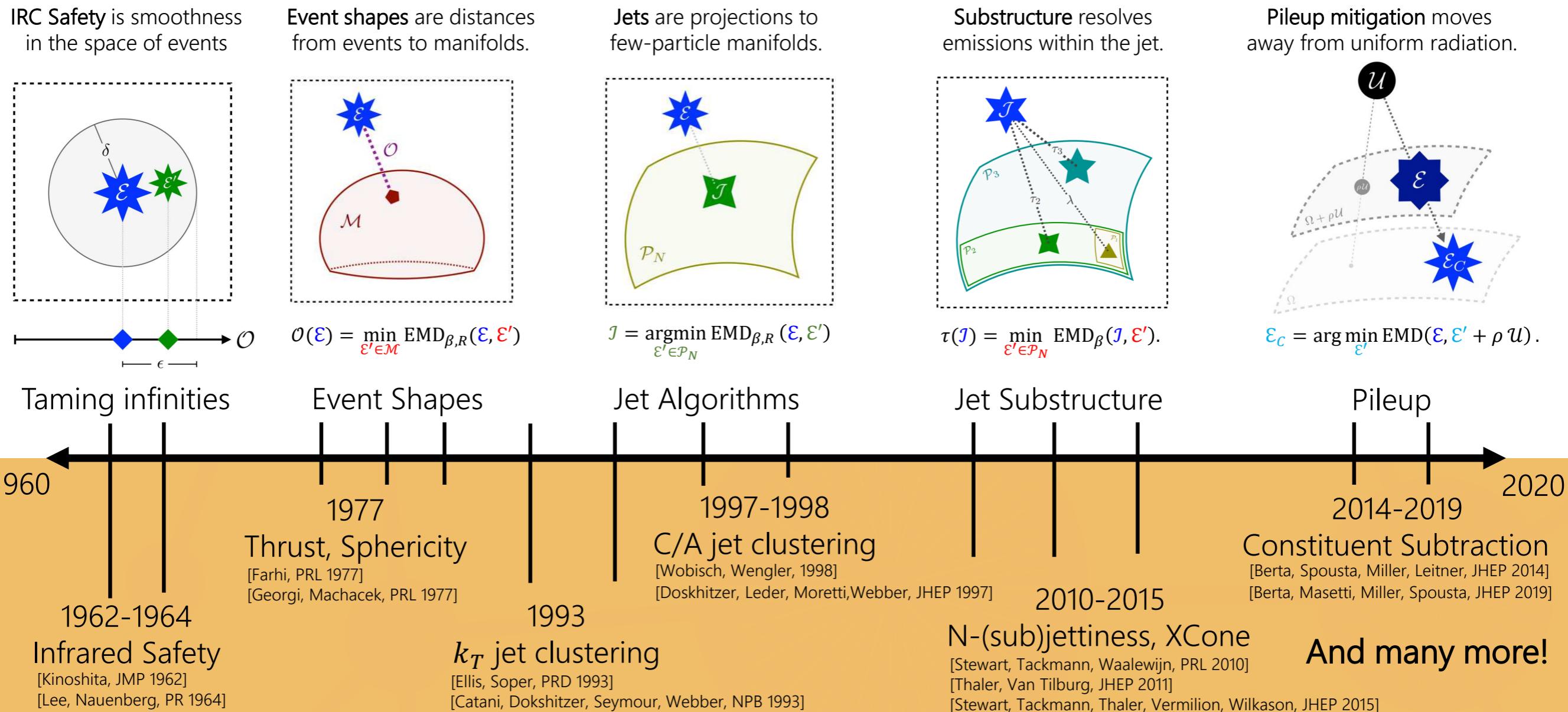
$$\tau_N(\mathcal{J}) = \min_{\mathcal{J}' \in \mathcal{P}_N} \text{EMD}(\mathcal{J}, \mathcal{J}')$$



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

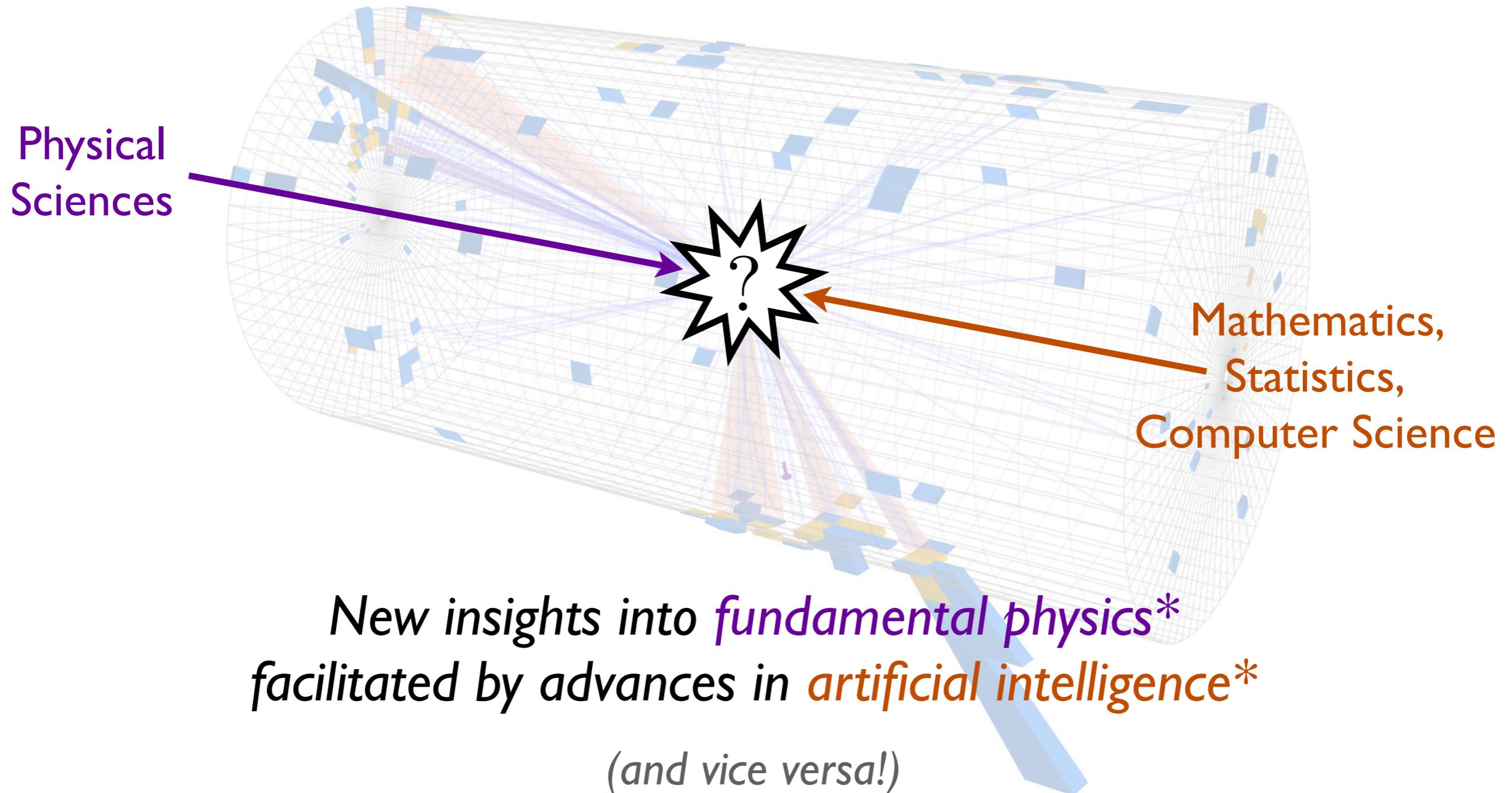


Six Decades of Collider Physics Translated into a New Geometric Language!



[Komiske, Metodiev, JDT, JHEP 2020; timeline by Metodiev]

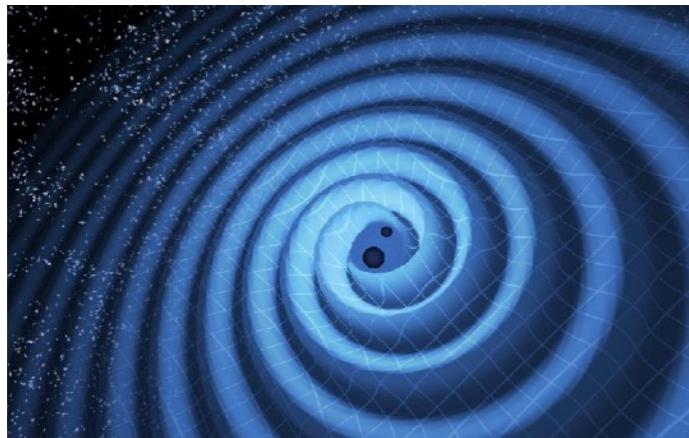
“Collision Course”



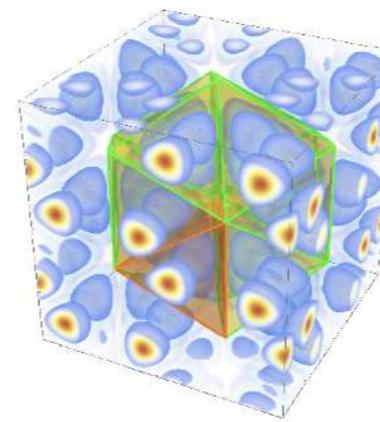
Artificial Intelligence \leftrightarrow Fundamental Interactions



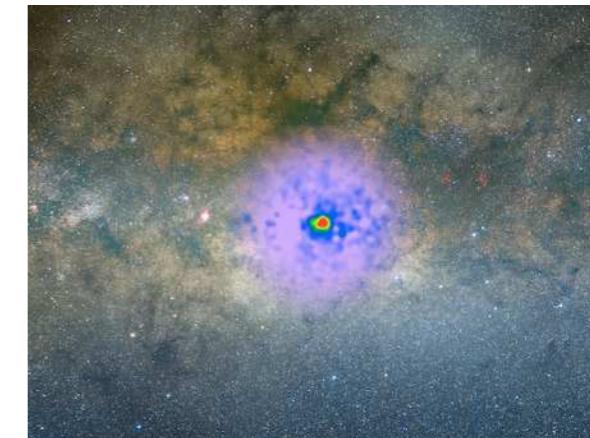
Gravitational Waves



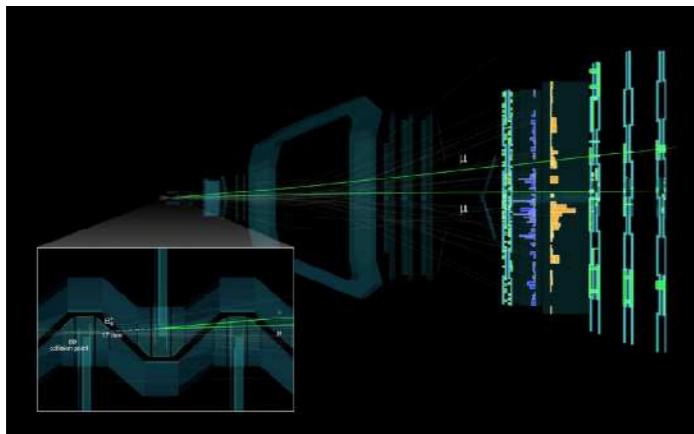
Nuclear Physics



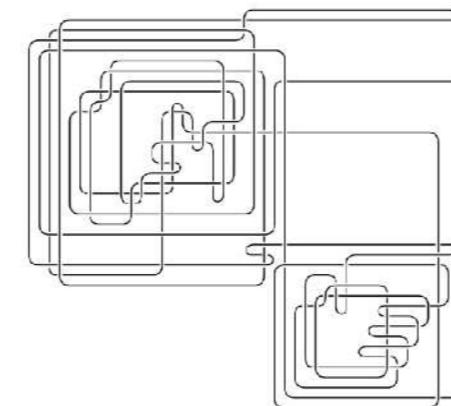
Astrophysics



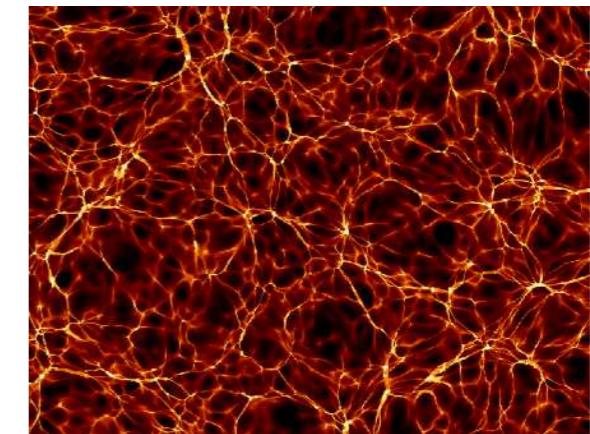
Particle Colliders



Mathematical Physics



Dark Matter

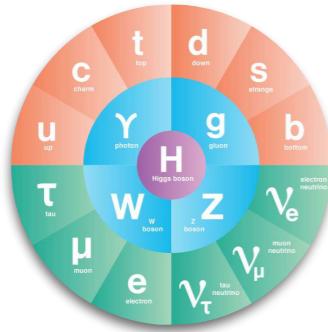
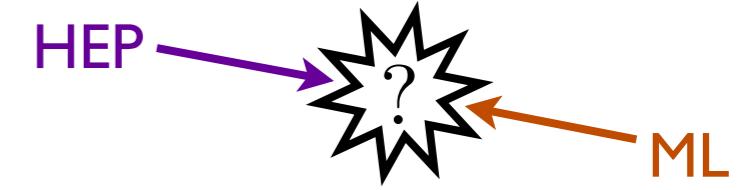


...

Machine learning that incorporates first principles, best practices, and domain knowledge from fundamental physics

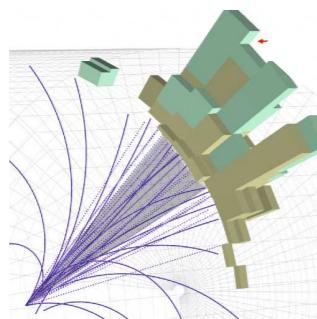
[<http://iaifi.org>]

Summary



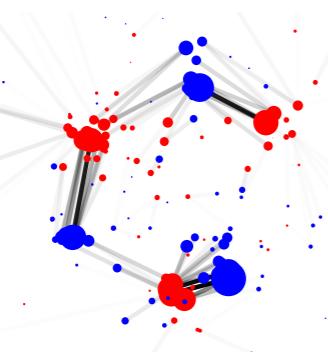
Particle Physics 101

*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*

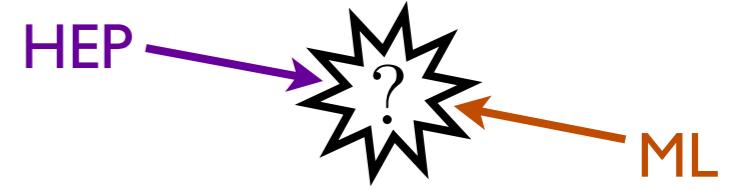


When are Collider Events Similar?

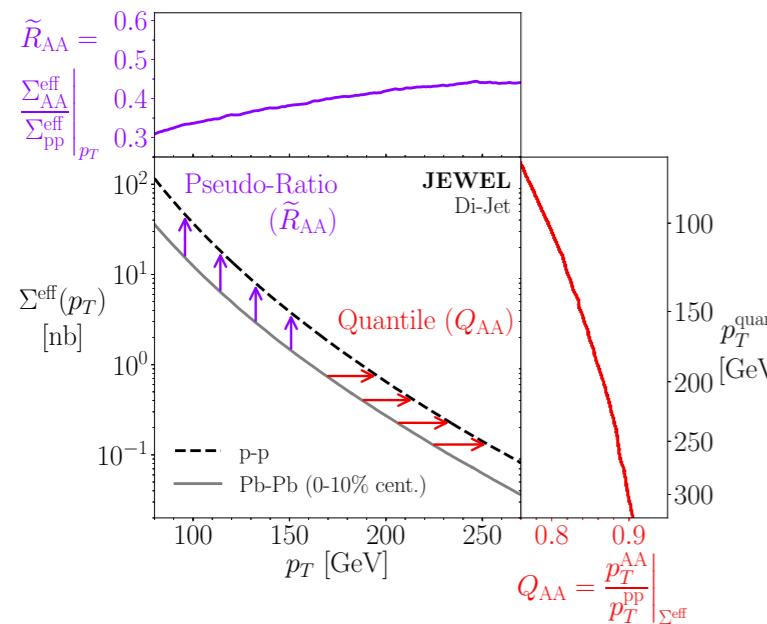
*When their energy flows are similar
Inspires unsupervised learning strategies based on event geometry*

Backup Slides

More Collisions

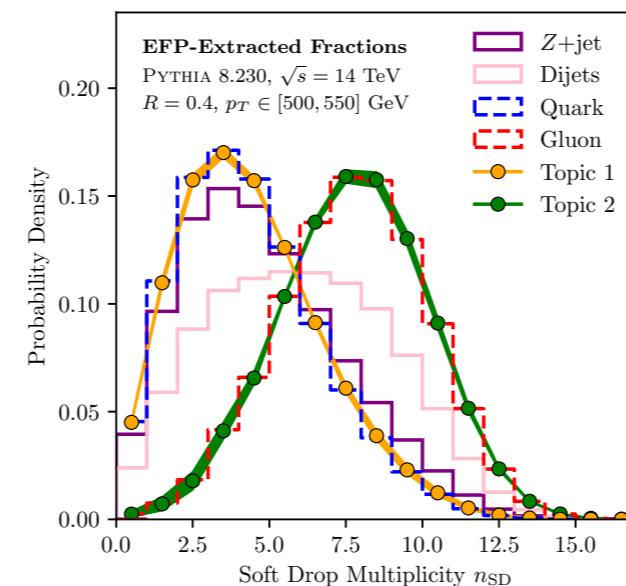


Jet Quenching via Optimal Transport



[Brewer, Milhano, JDT, PRL 2019]

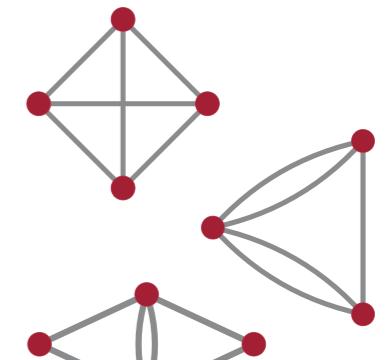
Quark/Gluon Definitions via Blind Source Separation



[Komiske, Metodiev, JDT, JHEP 2018;
Brewer, JDT, Turner; PRD 2021]

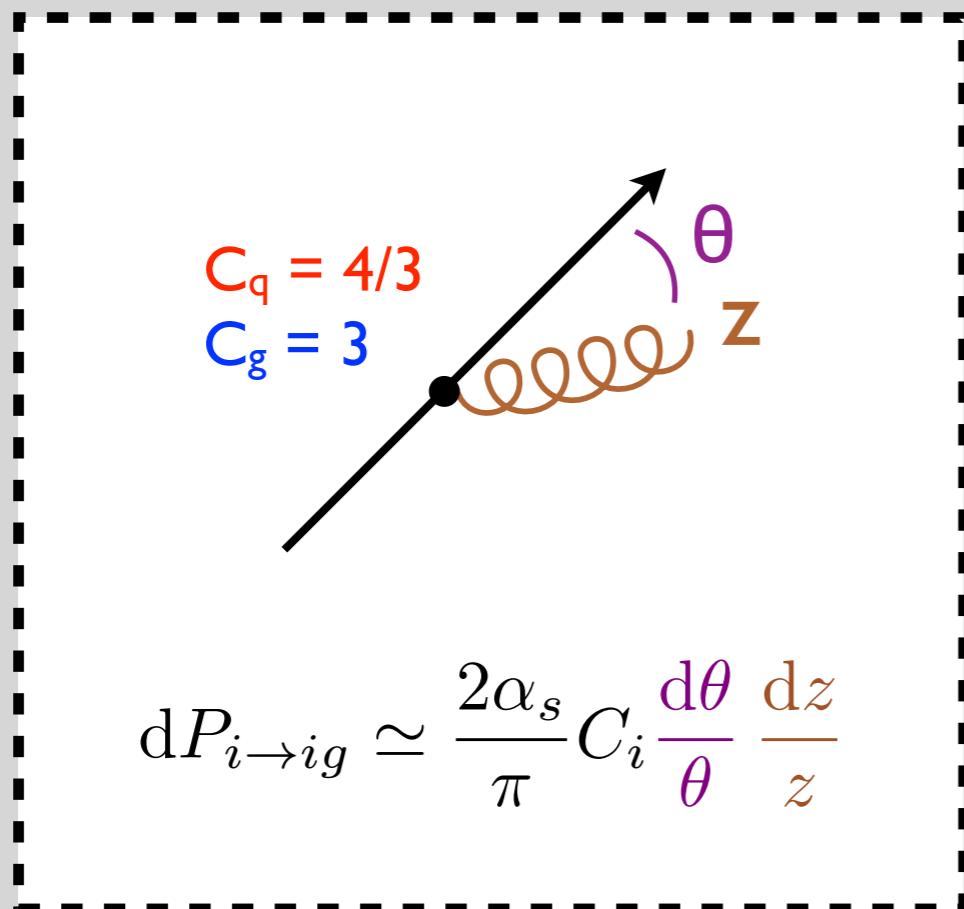
Kinematic Decomposition via Graph Theory

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



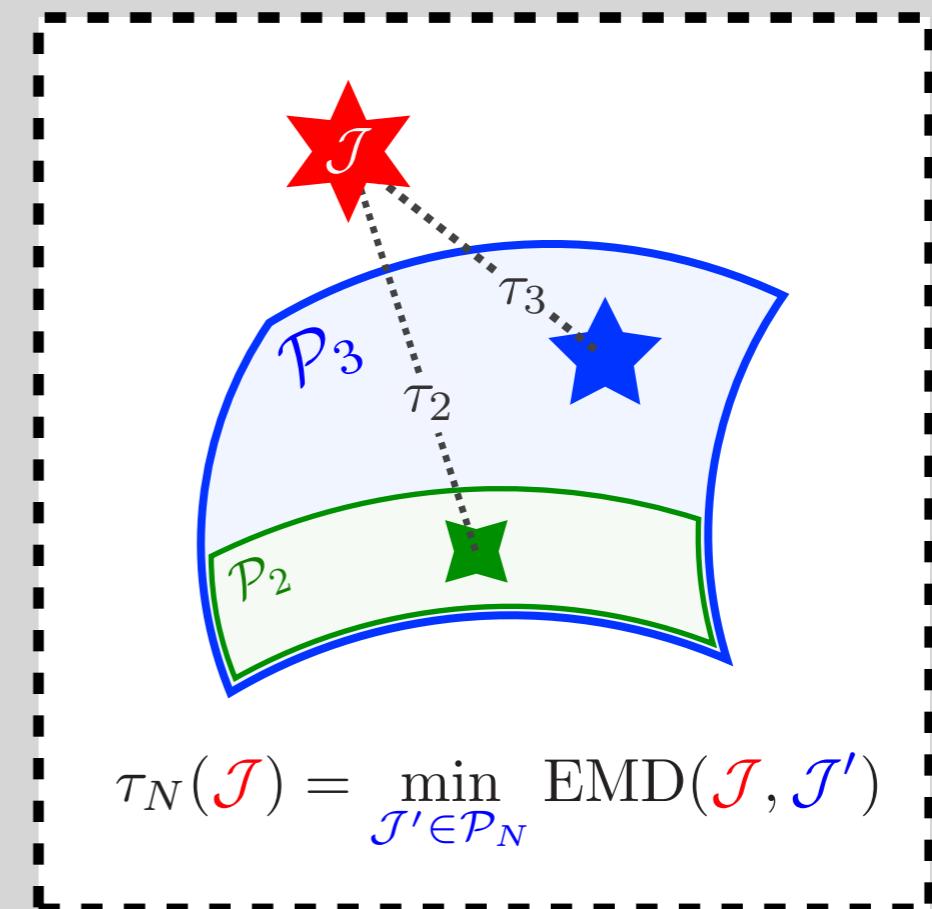
[Komiske, Metodiev, JDT,
JHEP 2018, PRD 2020]

Exploiting a Core Prediction of QCD



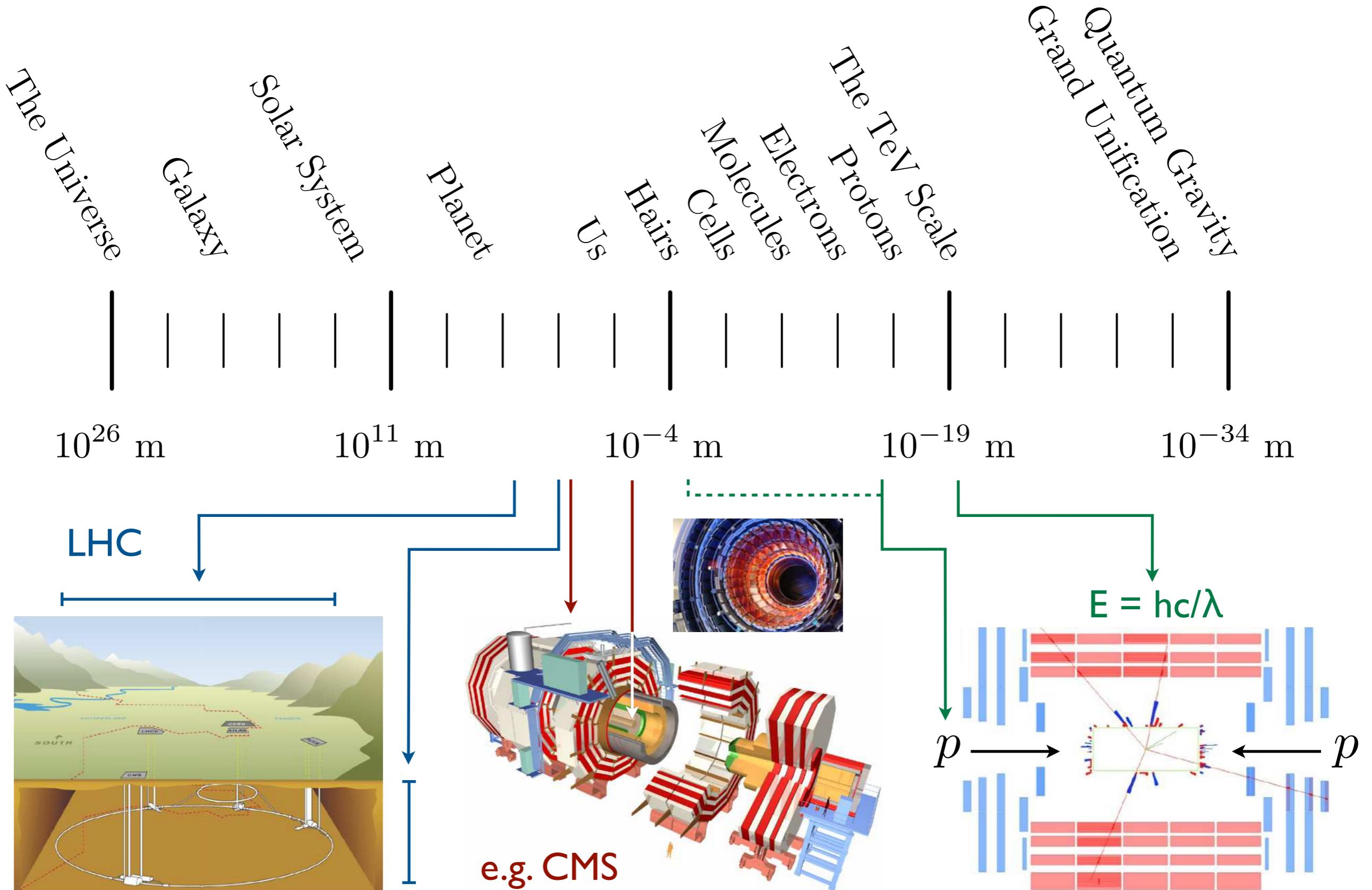
[Altarelli, Parisi, [NPB 1977](#)]

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



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

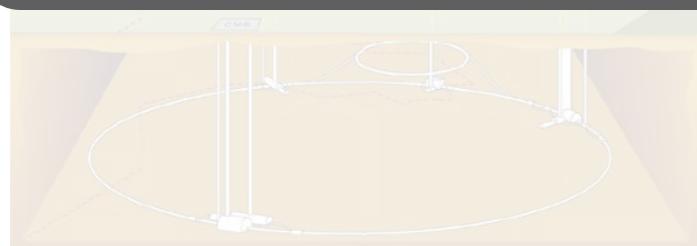
$p n \Lambda^0 \Sigma^+ \Sigma^0 \Sigma^- \Xi^0 \Xi^- \dots$
 $\Delta^{++} \Delta^+ \Delta^0 \Delta^- \Sigma^{*+} \Sigma^{*0} \Sigma^{*-} \Xi^{*0} \Xi^{*-} \Omega^- \dots$

Tetraquarks

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

Pentaquarks

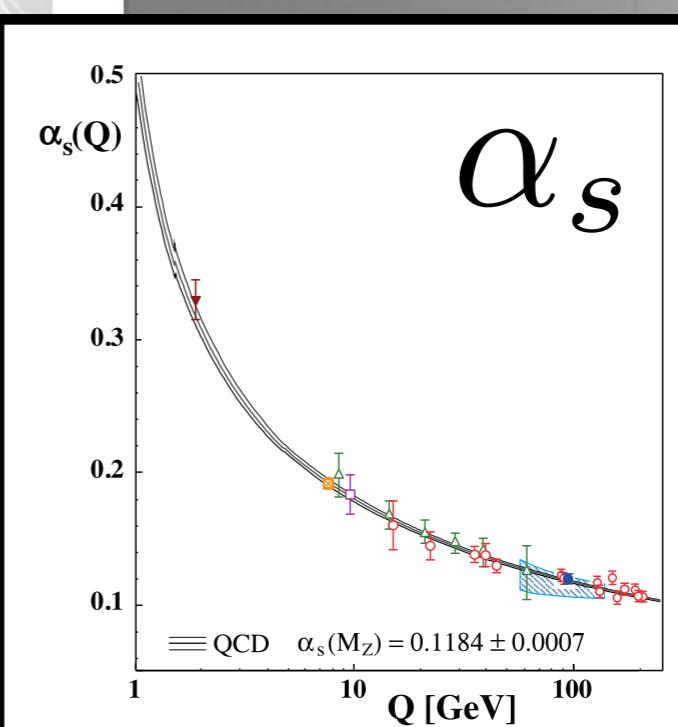
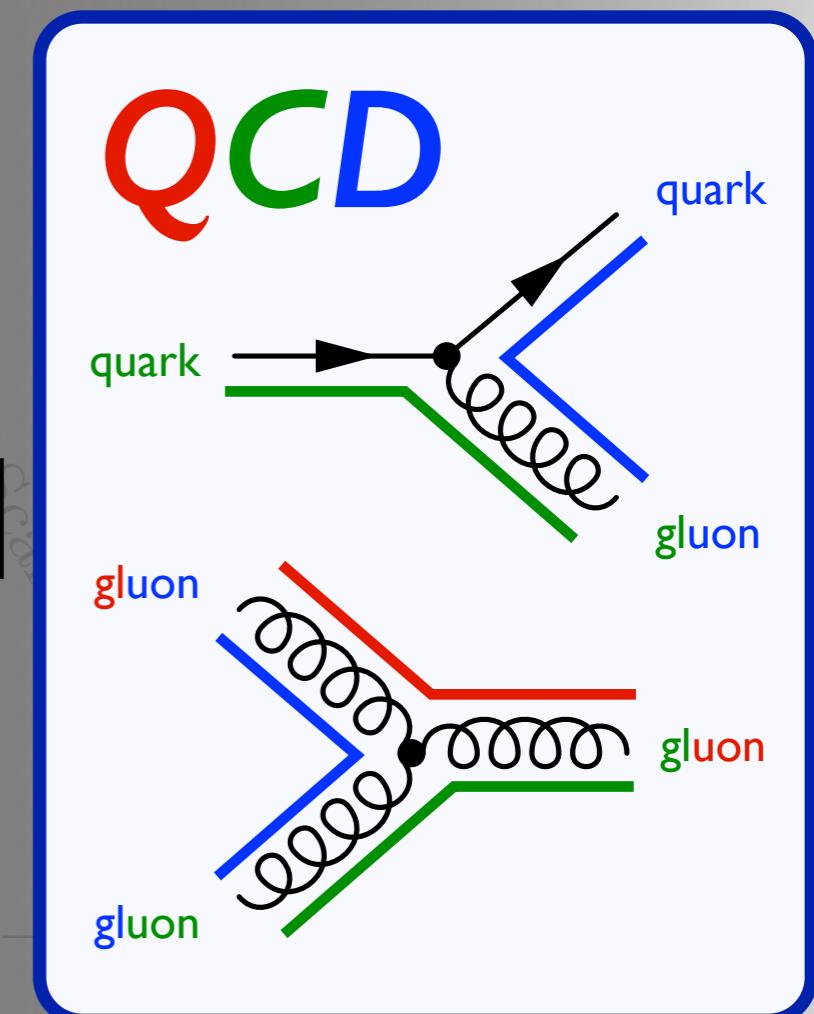
$P_c^+(4450) \dots$



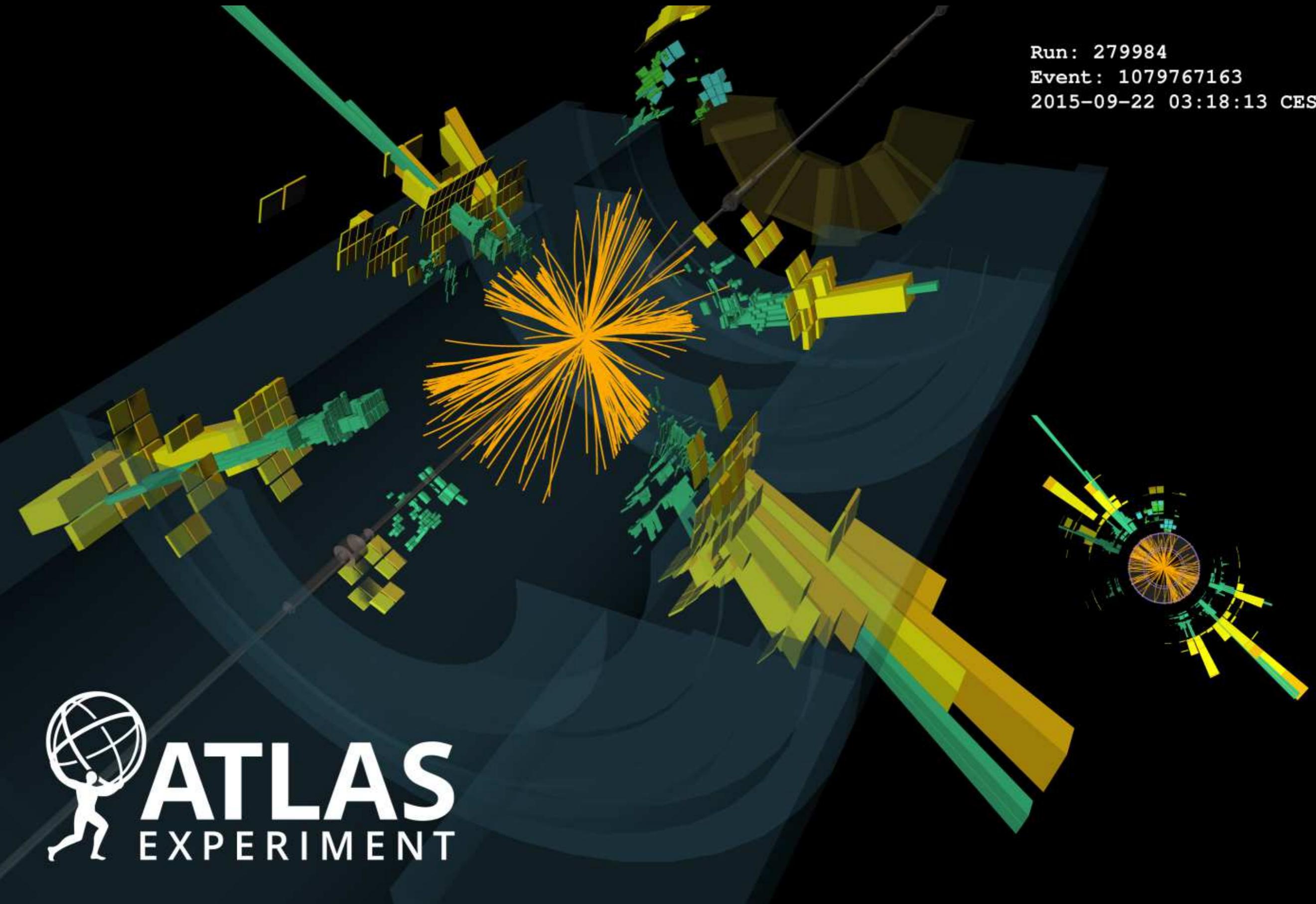
|

e.g. CMS

Jets

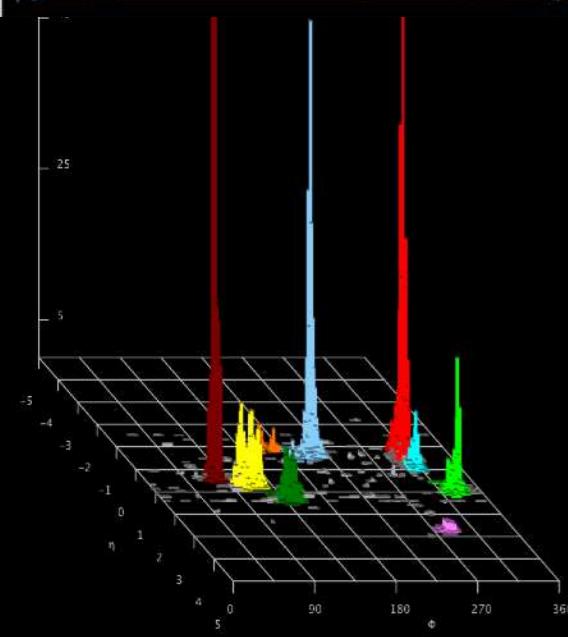
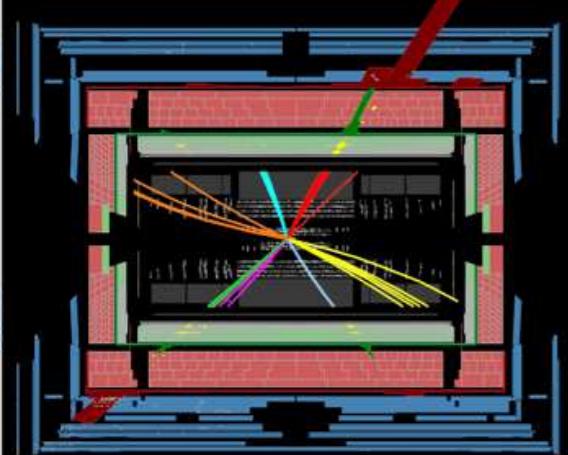


Run: 279984
Event: 1079767163
2015-09-22 03:18:13 CEST

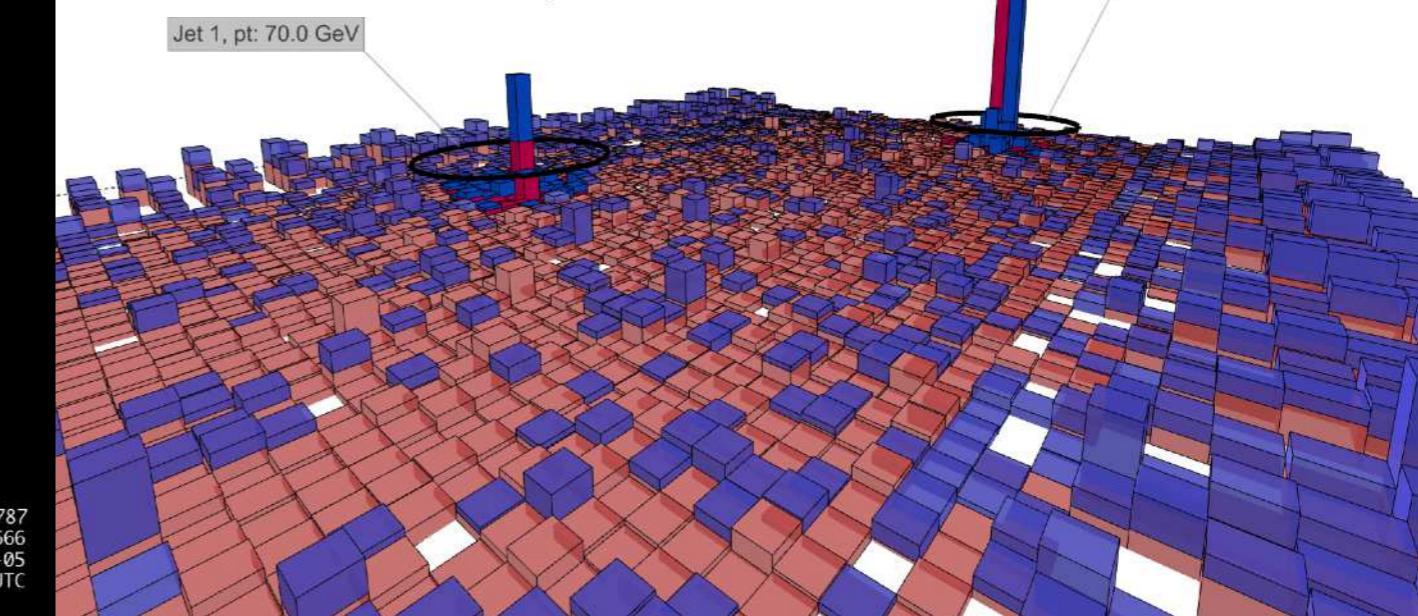
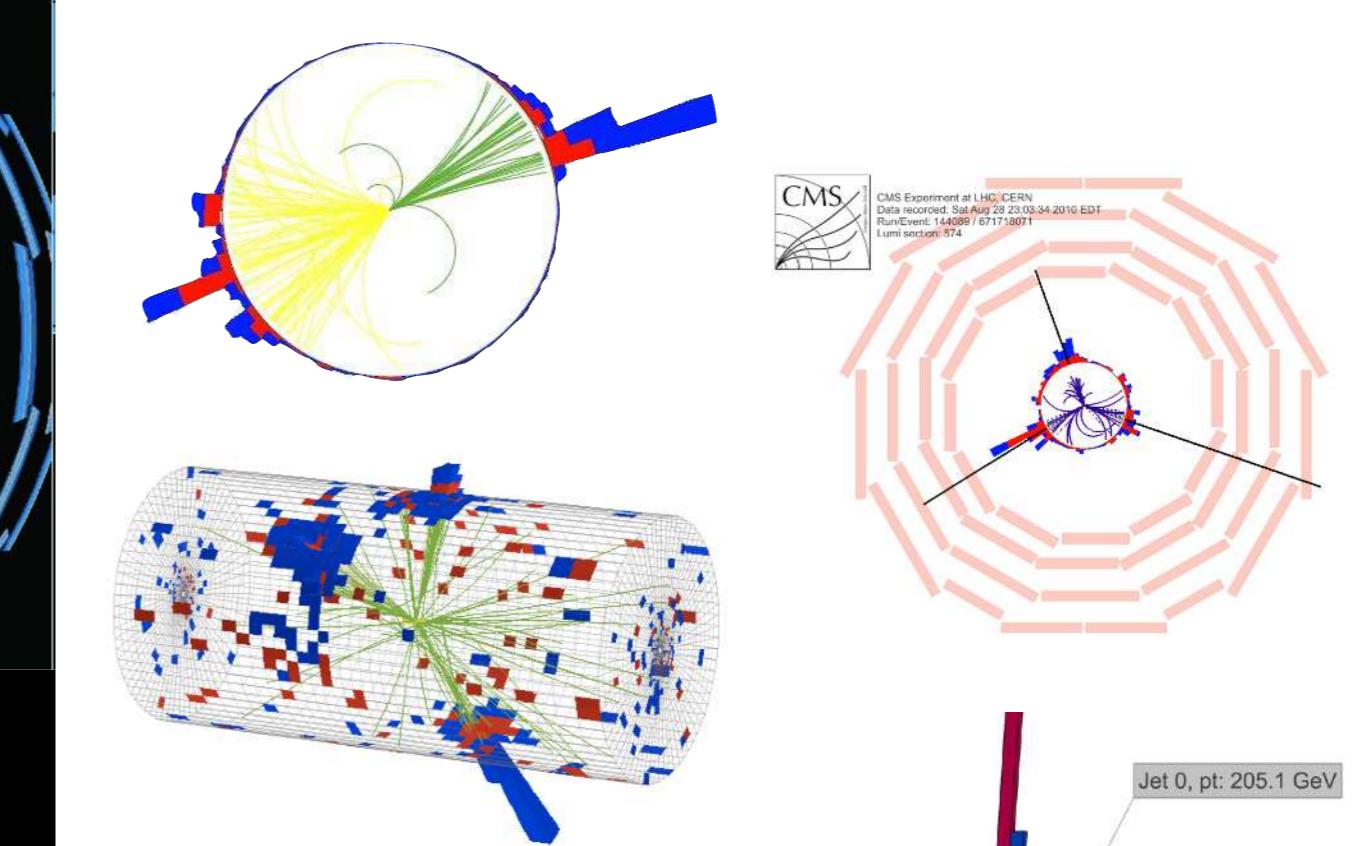
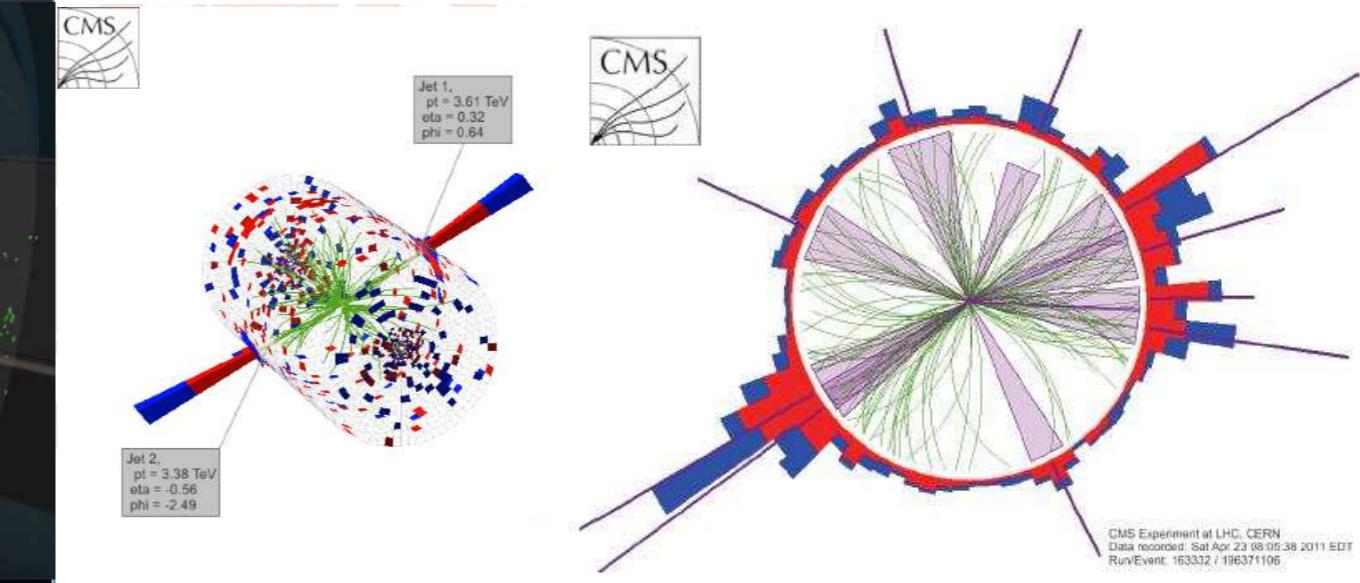
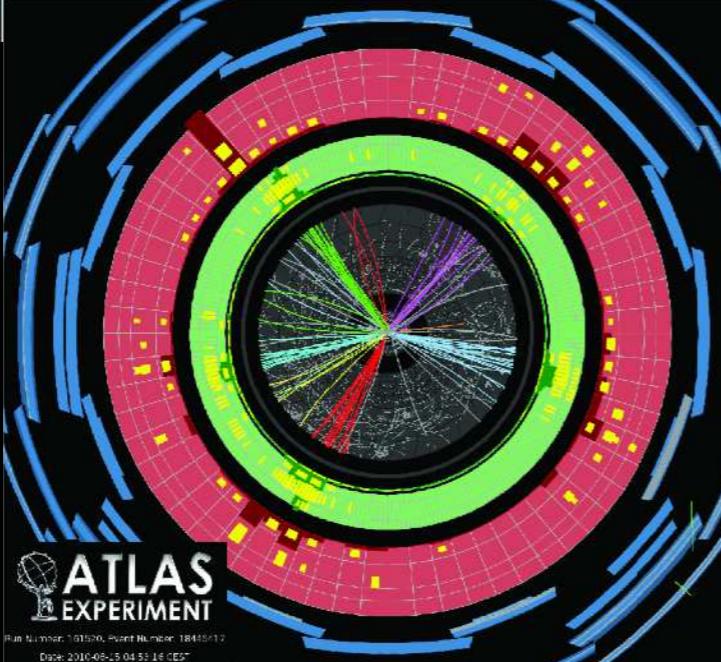
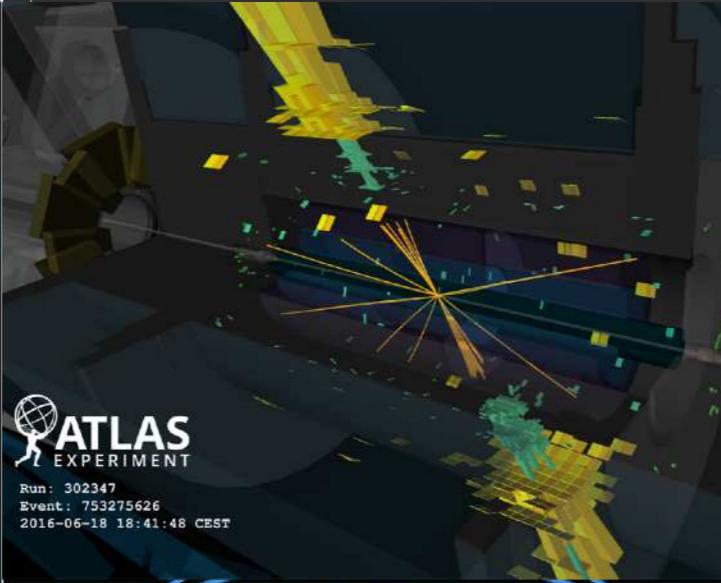
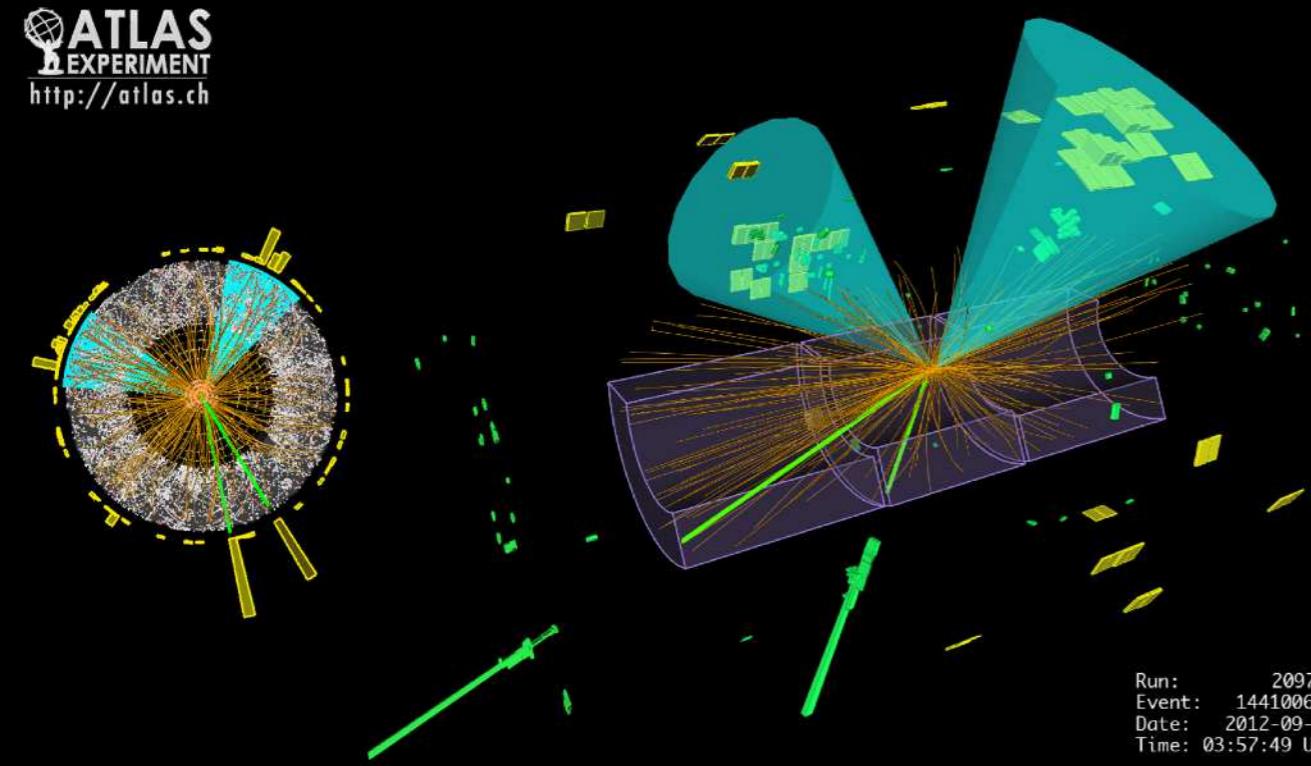


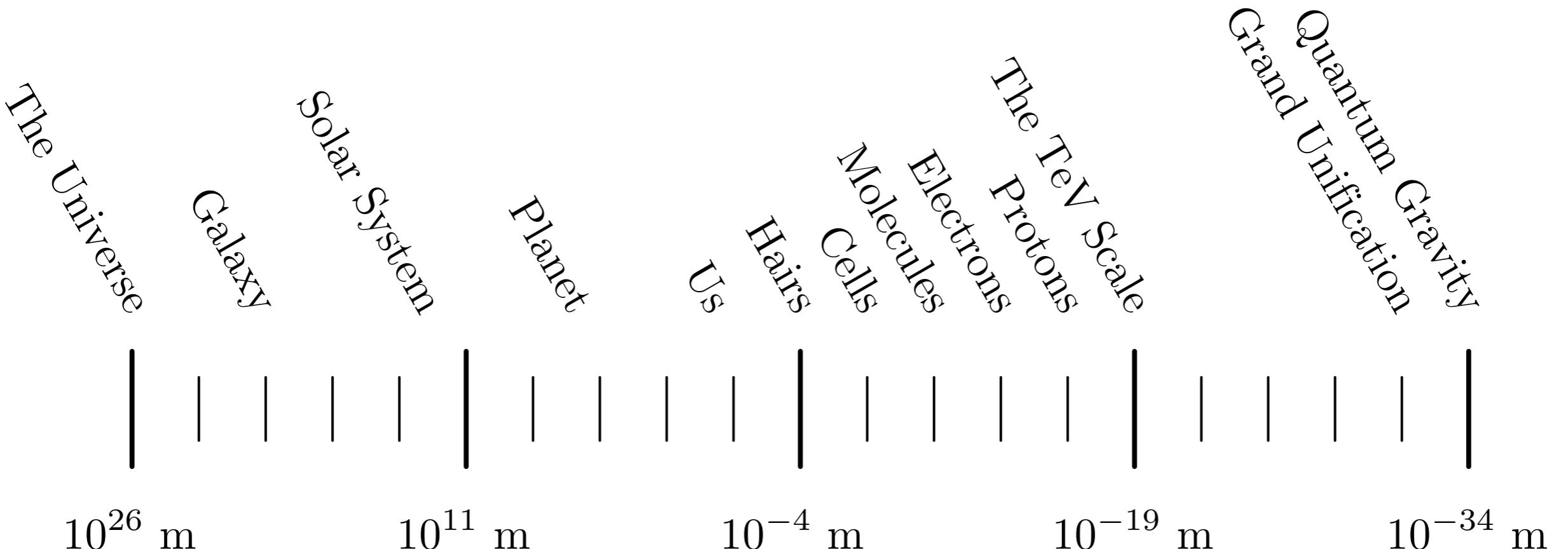
Run Number: 159224, Event Number: 3533152

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

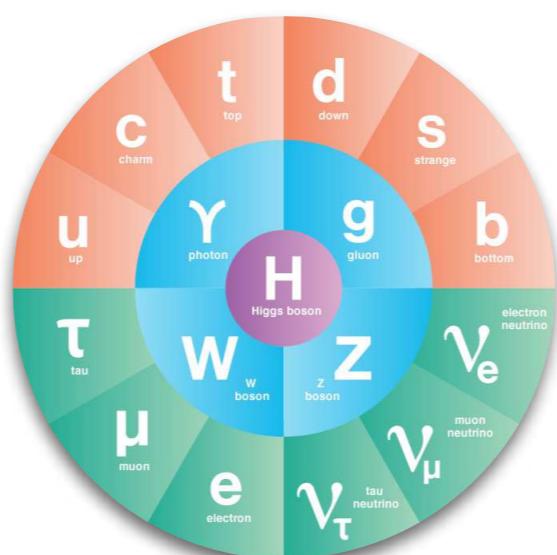


ATLAS
EXPERIMENT
<http://atlas.ch>



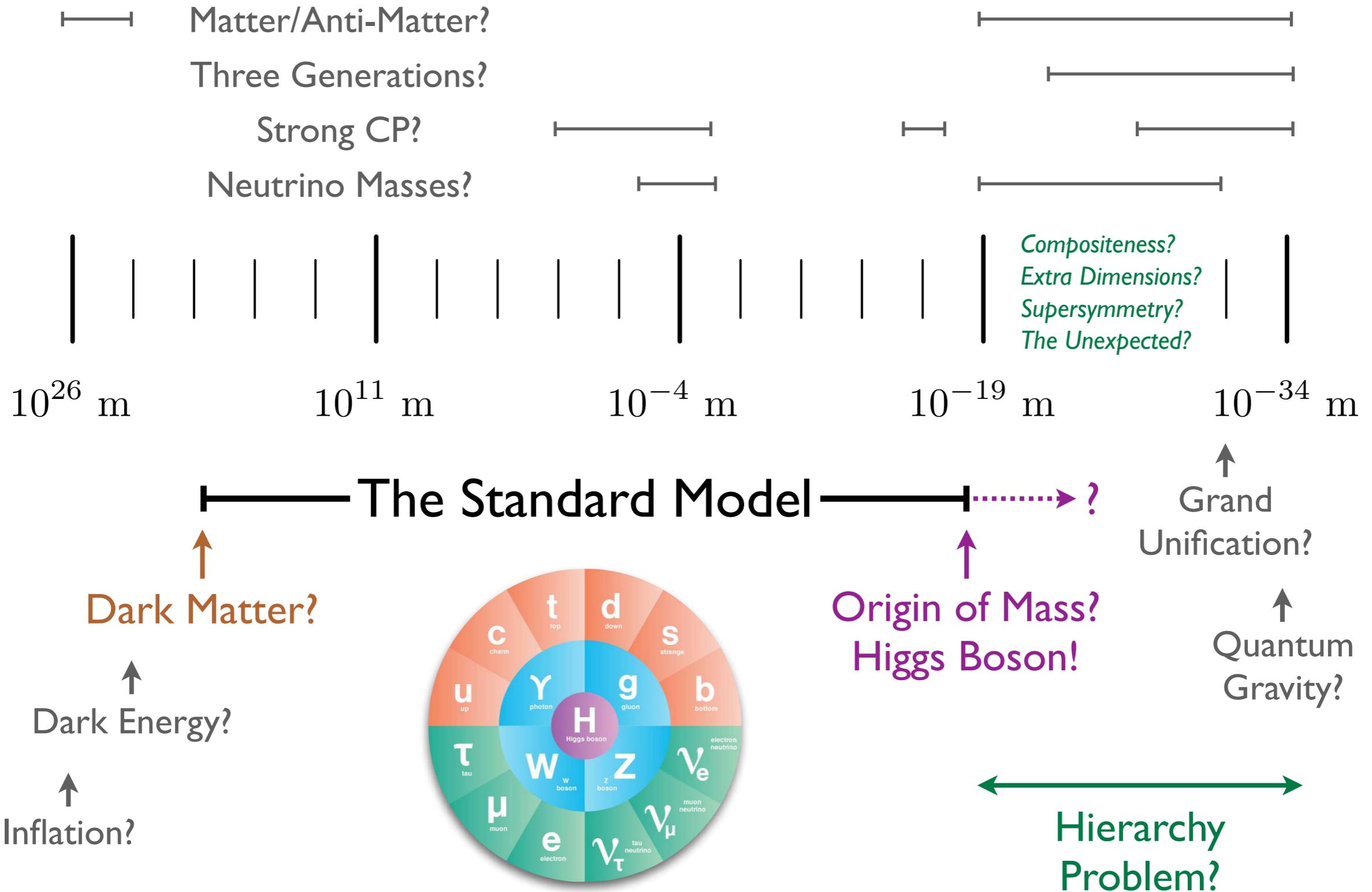


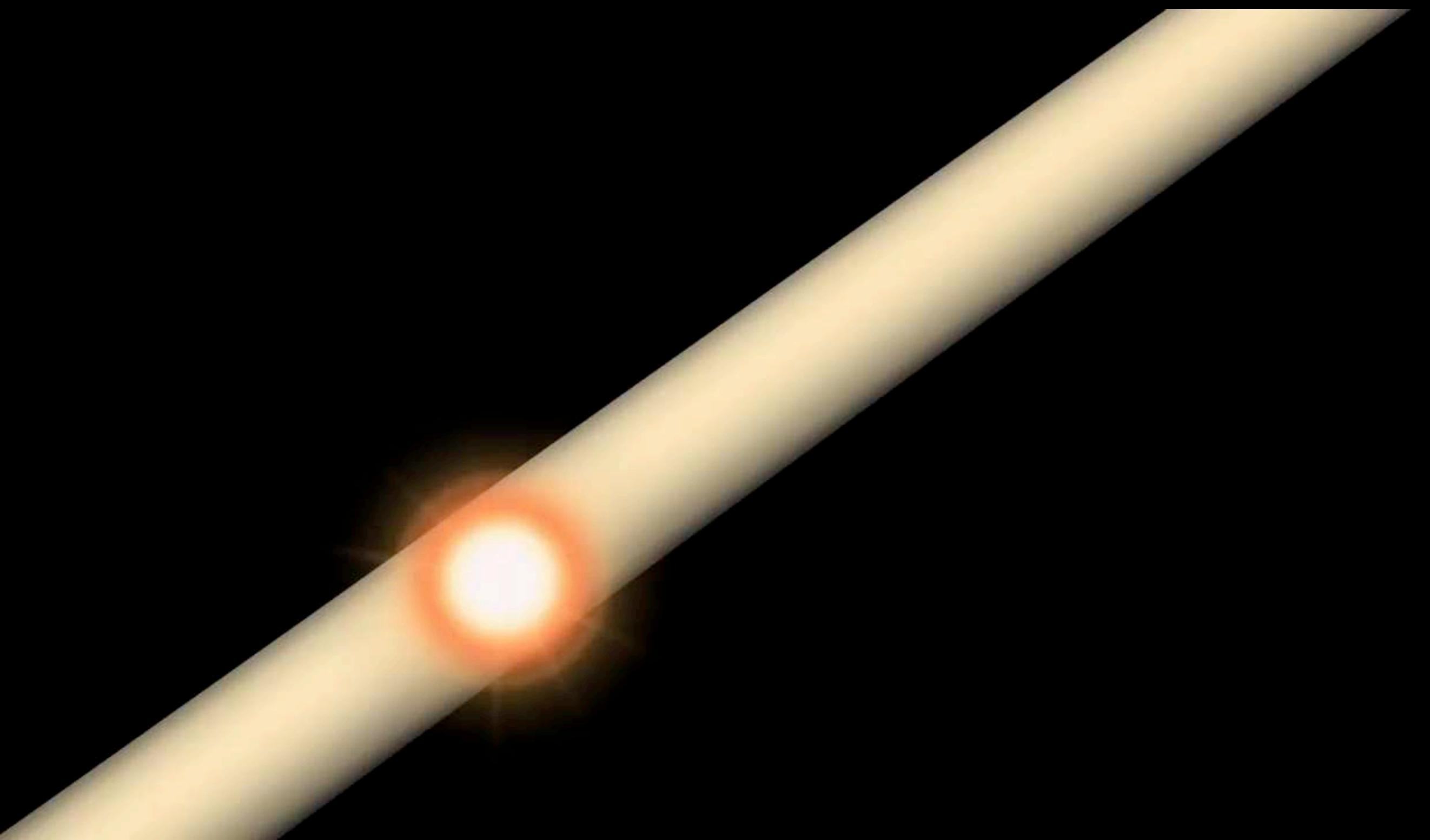
The Standard Model



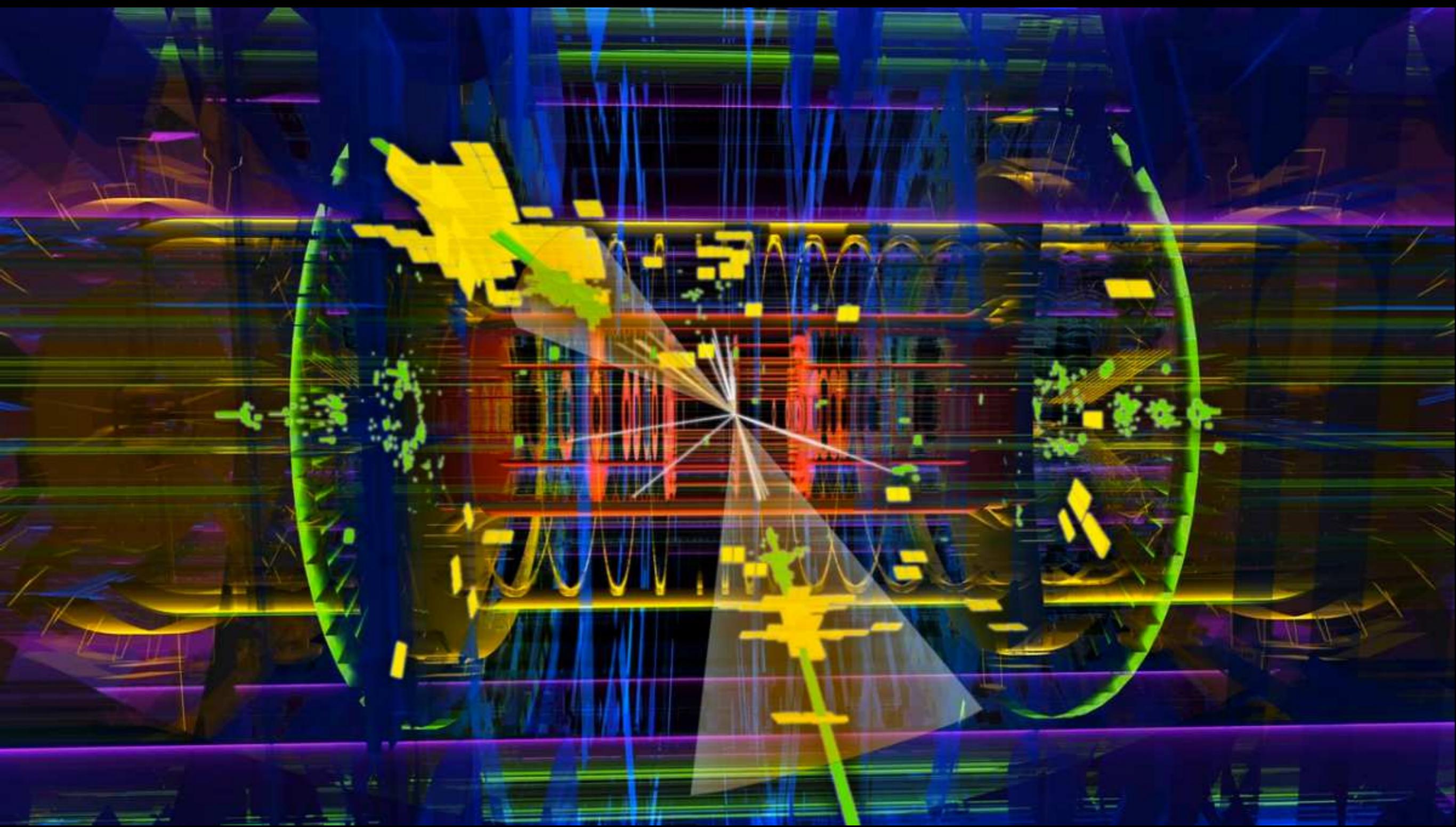
Puzzles in Particle Physics

SnowMass2021





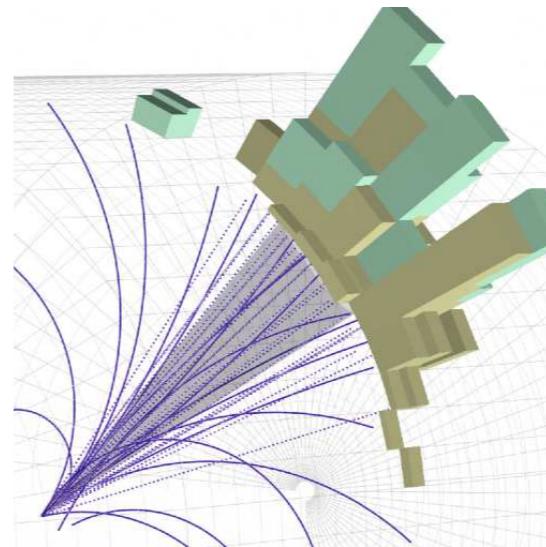




Cartoon of Machine Learning

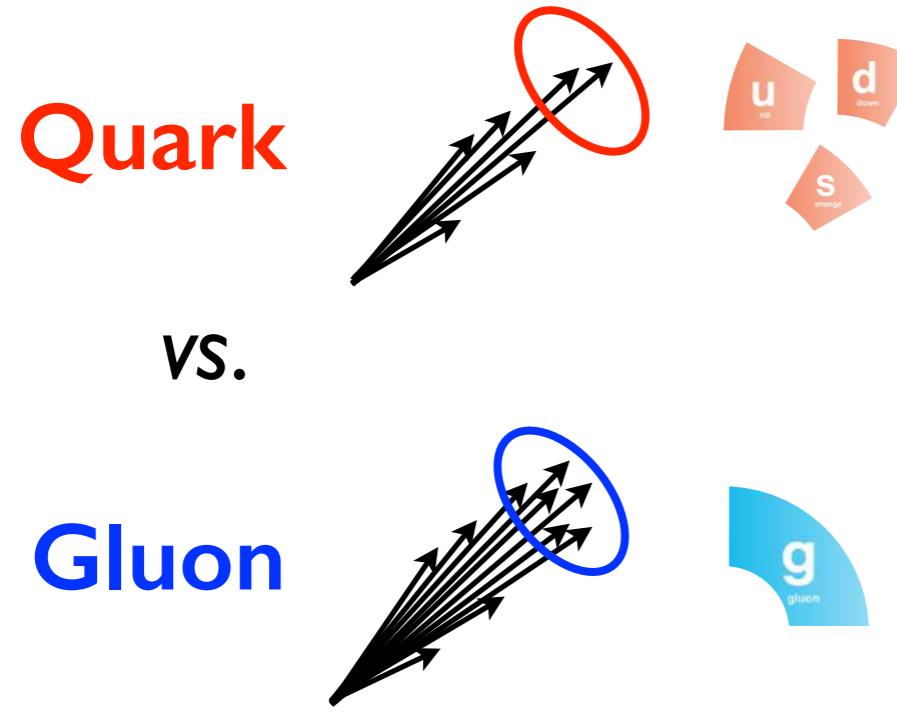
“ML4Jets”
NYU, January 2020

For this talk: \mathcal{J} = jet

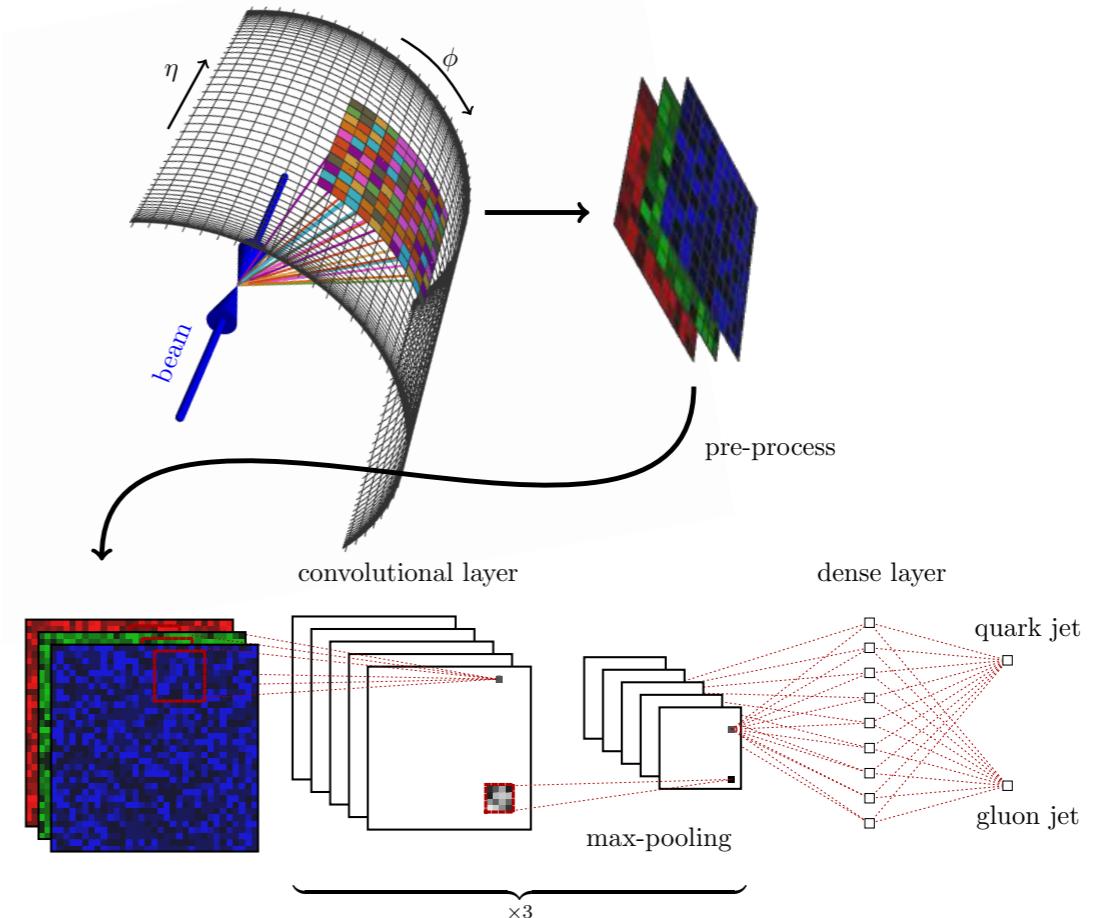


From Curmudgeon...

Jet classification via image recognition



Multi-channel convolutional neural networks



[e.g. Komiske, Metodiev, Schwartz, [JHEP 2017](#);
cf. Gras, Höche, Kar, Larkoski, Lönnblad, Plätzer, Siódak, Skands, Soyez, [JHEP 2017](#)]



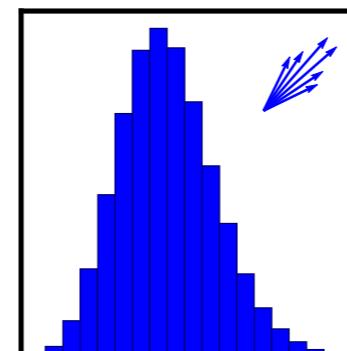
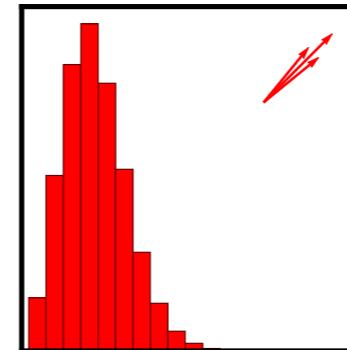
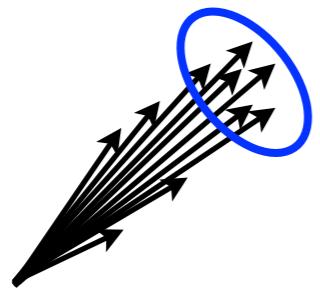
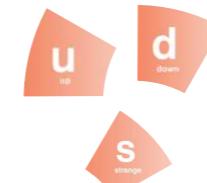
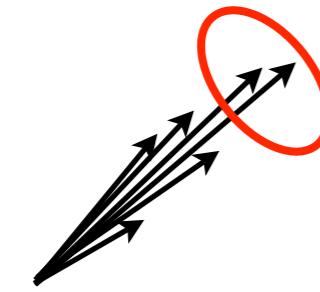
...to Evangelist

Jet flavor definitions via natural language processing

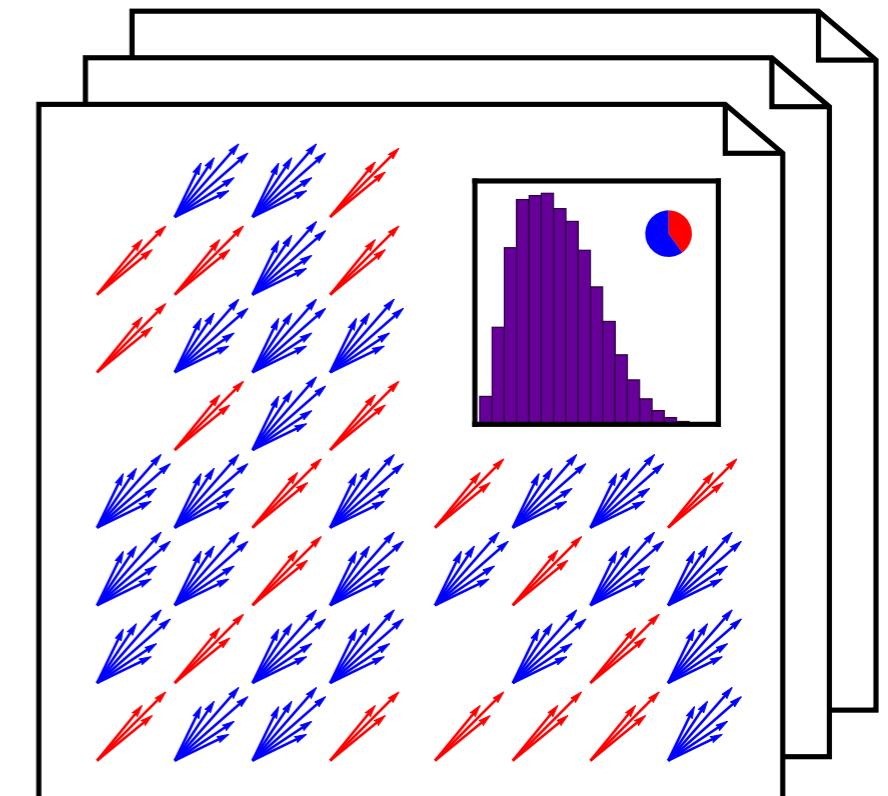
Quark

vs.

Gluon



Topic Modeling / Blind Source Separation



[Komiske, Metodiev, JDT, [JHEP 2018](#);
using Metodiev, Nachman, JDT, [JHEP 2017](#); Metodiev, JDT, [PRL 2018](#); Komiske, Metodiev, JDT, [JHEP 2019](#)]



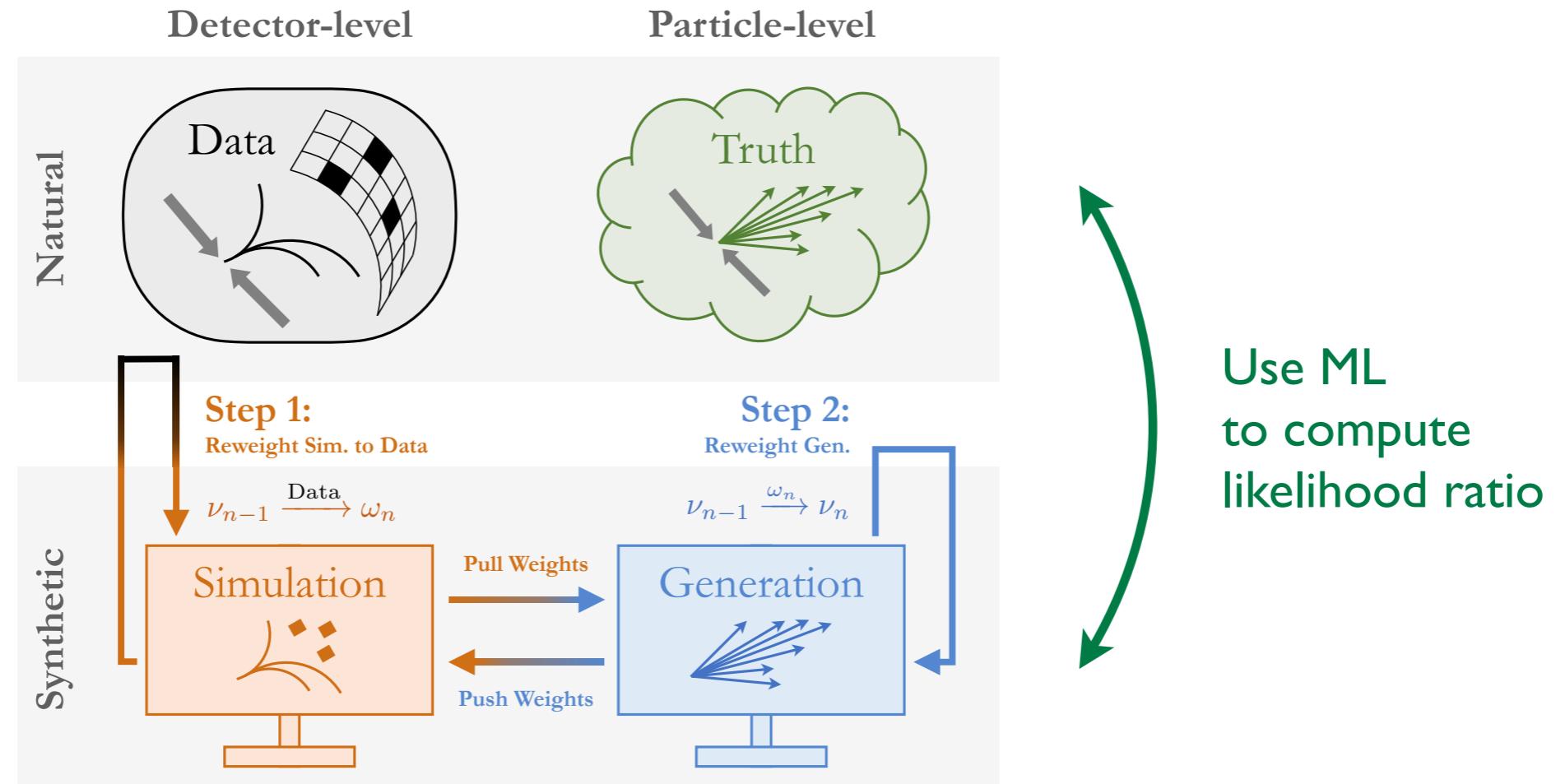
Detector Unfolding



OmniFold



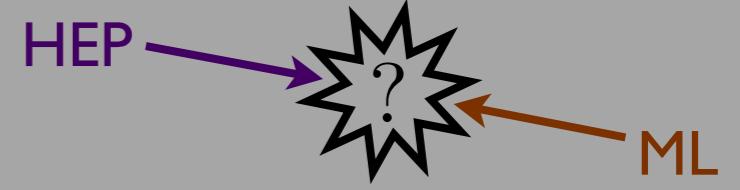
*Multi-dimensional unbinned detector corrections
via iterated binary classification*



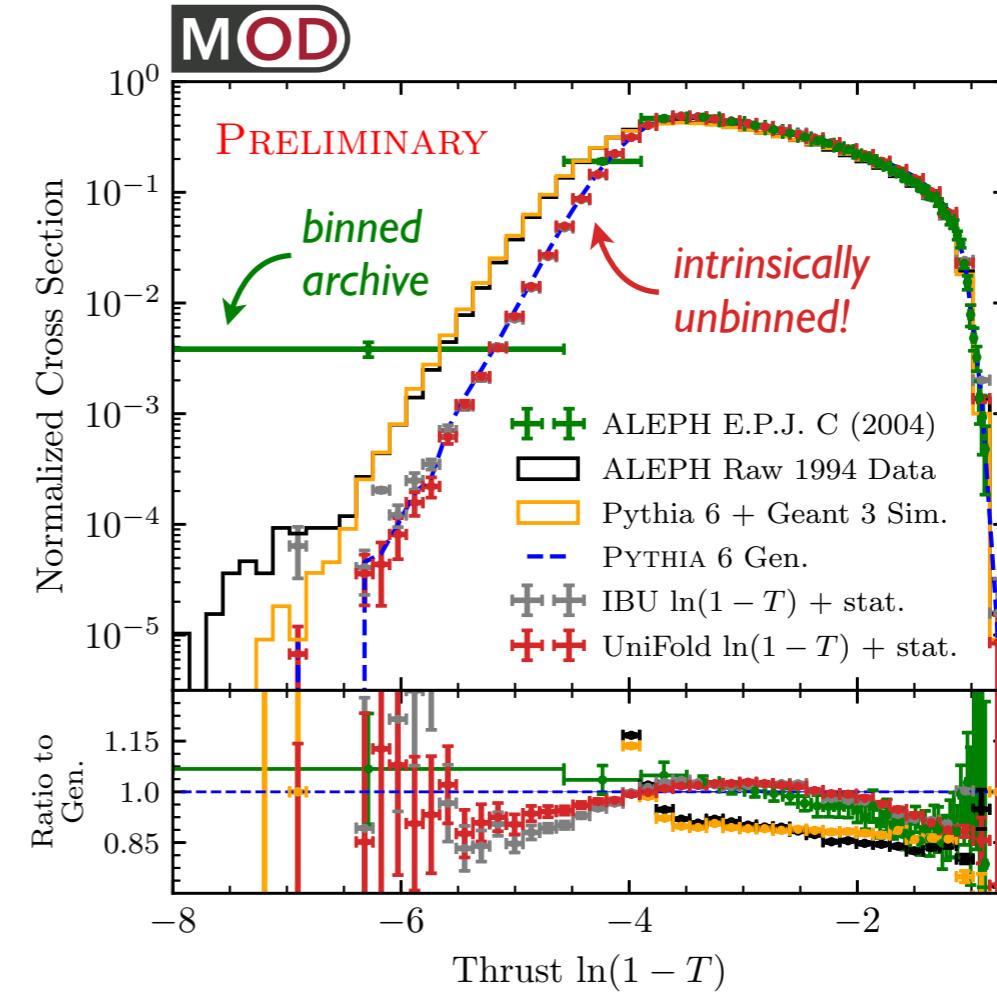
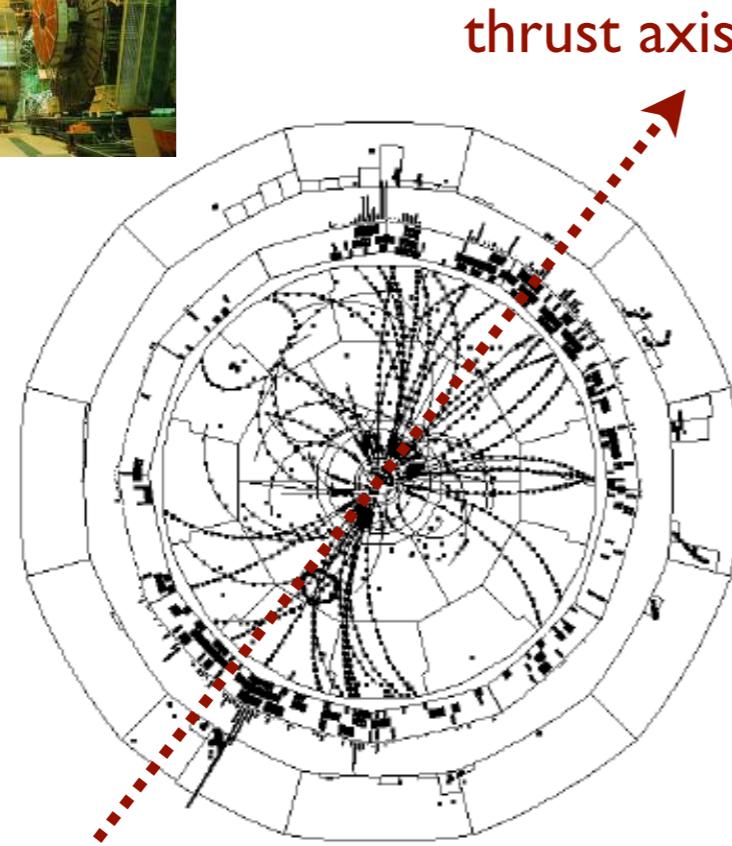
[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020]



Detector Unfolding



Back to the Future with ALEPH Archival Data



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



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



Likelihood Ratio Trick

Many HEP problems can be expressed in this form!

Key example of *simulation-based inference*

Goal: Estimate $p(x)$ / $q(x)$

Training Data: Finite samples P and Q

Learnable Function: $f(x)$ parametrized by, e.g., neural networks

Loss Function(al): $L = -\langle \log f(x) \rangle_P + \langle f(x) - 1 \rangle_Q$

Asymptotically: $\arg \min_{f(x)} L = \frac{p(x)}{q(x)}$ *Likelihood ratio*

$-\min_{f(x)} L = \int dx p(x) \log \frac{p(x)}{q(x)}$ *Kullback–Leibler divergence*

[see e.g. D’Agnolo, Wulzer, [PRD 2019](#); simulation-based inference in Cranmer, Brehmer, Louppe, [PNAS 2020](#); relation to f-divergences in Nguyen, Wainwright, Jordan, [AoS 2009](#)]

Likelihood Ratio Trick

Many HEP problems can be expressed in this form!

Key example of *simulation-based inference*

Asymptotically, same structure as **Lagrangian mechanics!**

Action: $L = \int dx \mathcal{L}(x)$

Lagrangian: $\mathcal{L}(x) = -p(x) \log f(x) + q(x)(f(x) - 1)$

Euler-Lagrange: $\frac{\partial \mathcal{L}}{\partial f} = 0$

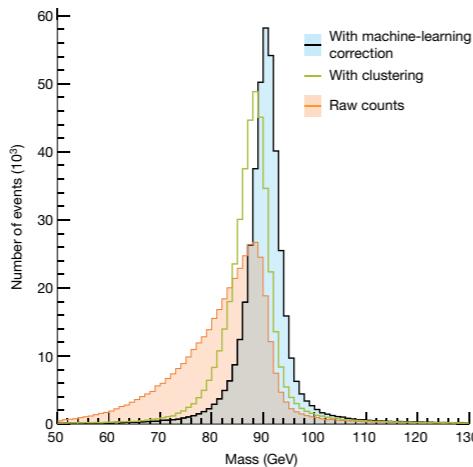
Solution: $f(x) = \frac{p(x)}{q(x)}$

Requires shift in theoretical focus from solving problems to *specifying problems*

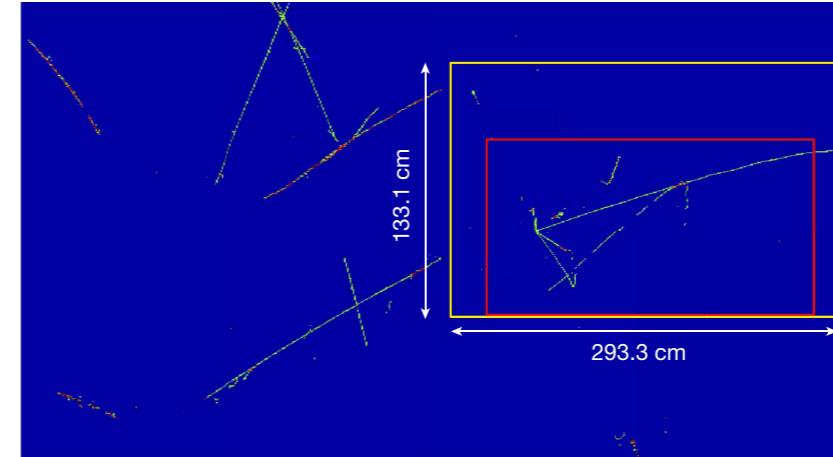
[see e.g. D'Agnolo, Wulzer, [PRD 2019](#); simulation-based inference in Cranmer, Brehmer, Louppe, [PNAS 2020](#); relation to f-divergences in Nguyen, Wainwright, Jordan, [AoS 2009](#)]

Extensive Use of ML in HEP

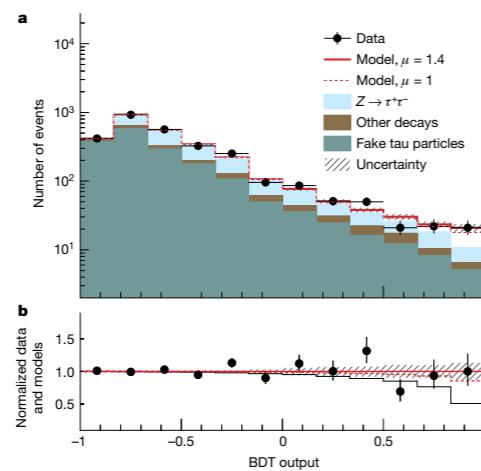
CMS: $Z \rightarrow e^+e^-$ calibration



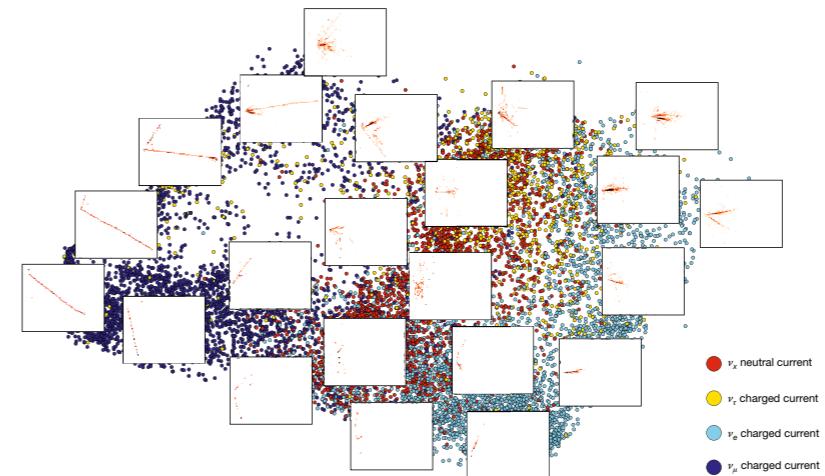
MicroBooNE: Object Identification



ATLAS: $H \rightarrow \mu^+\mu^-$ search



NOvA: Object Classification



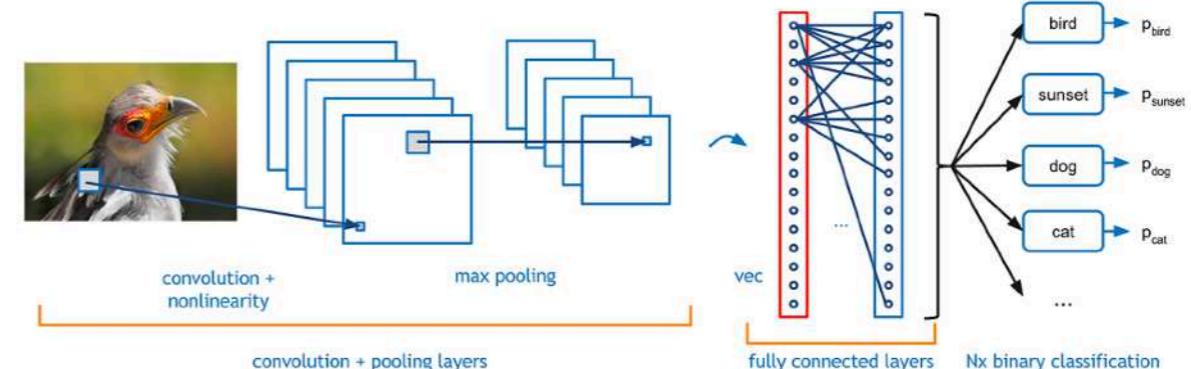
Machine learning is transforming many aspects of society, including fundamental physics research

[Radovic, Williams, Rousseau, Kagan, Bonacorsi, Himmel, Aurisano, Terao, Wongjirad, [Nature 2018](#)]

Off-the-Shelf ML for HEP?

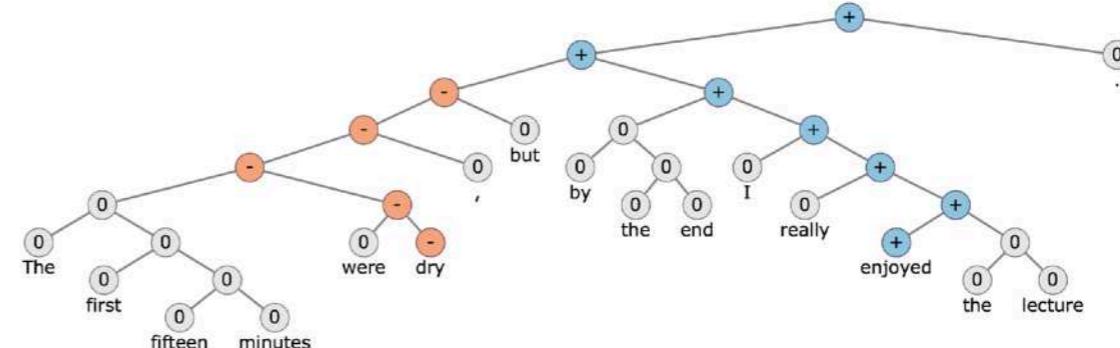
2D Images?

Appropriate for fixed-grid calorimeters,
but less ideal for tracking detectors



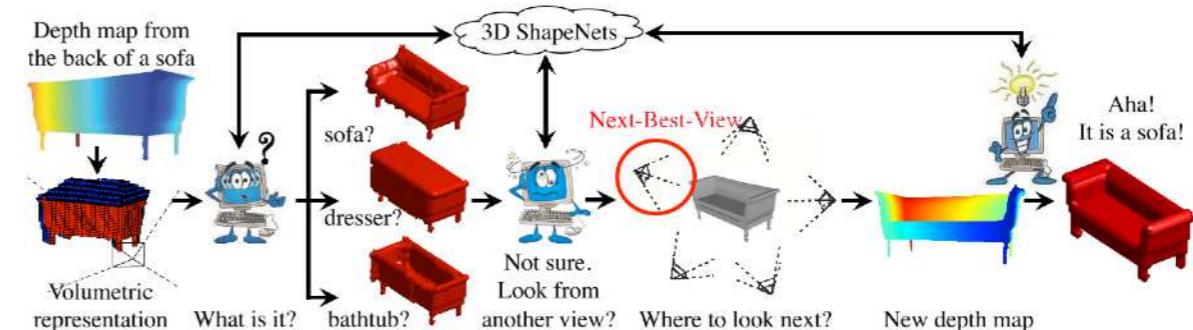
Natural Language?

Clustering can yield “semantic” structure, but
identical particles have no intrinsic ordering



3D Objects?

Much closer to particle physics,
though doesn't capture all symmetries



Principles of Fundamental Physics

Robustness of Energy Flow

[Komiske, Metodiev, JDT, JHEP 2018]



Patrick Komiske



Eric Metodiev



SF



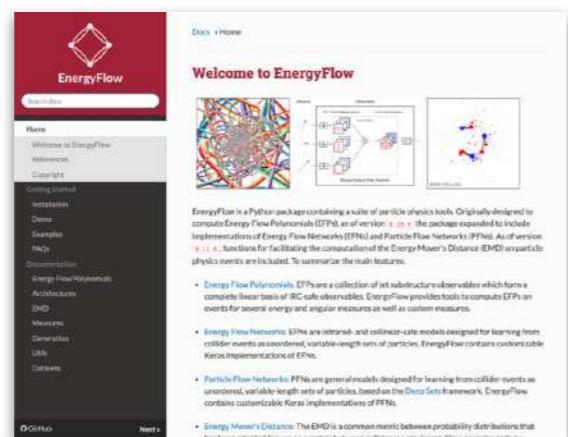
Power of Artificial Intelligence

Point Cloud Learning

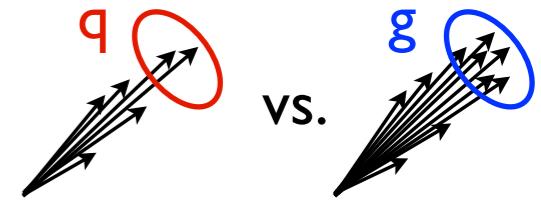
[Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, NIPS 2017]

Energy Flow Networks

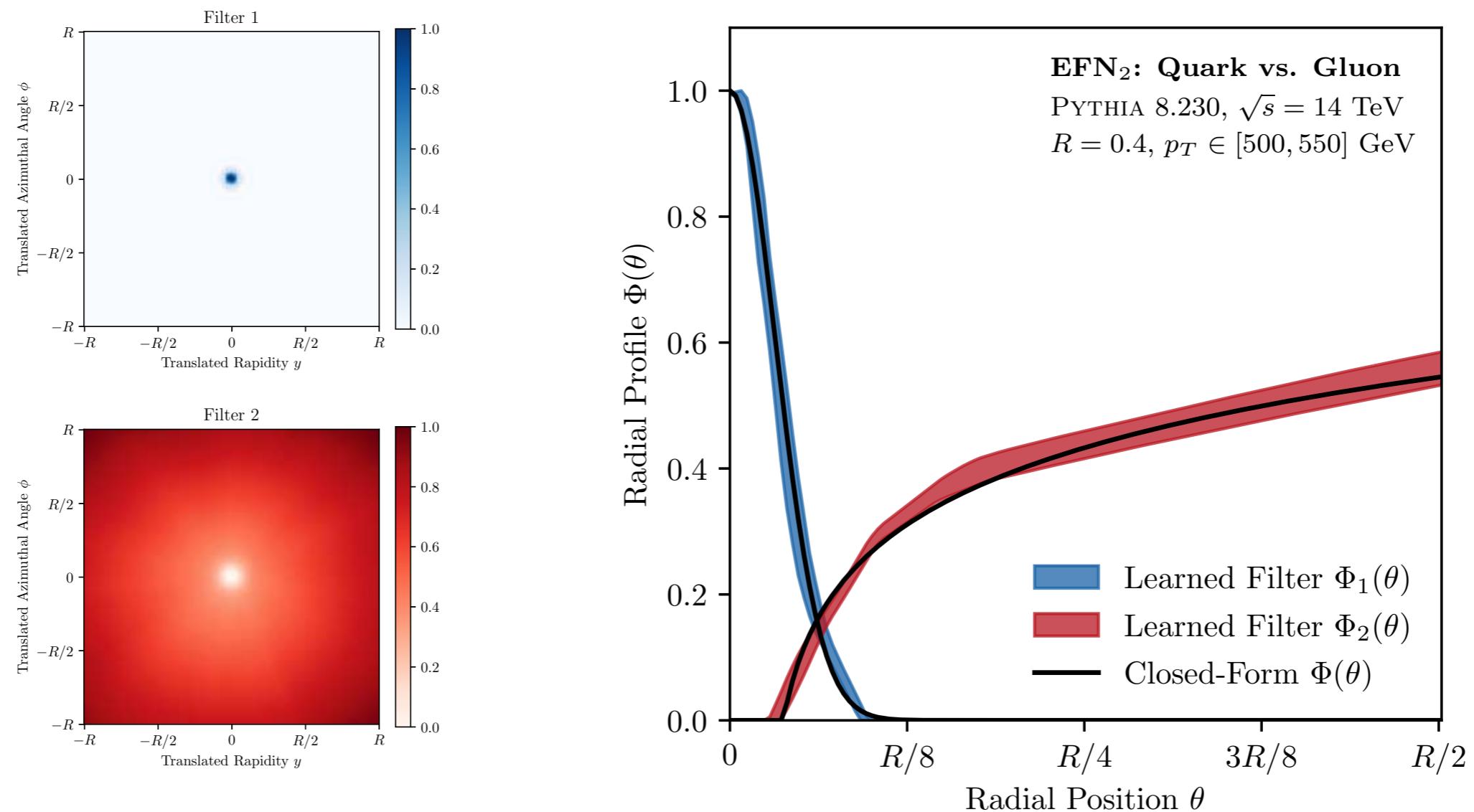
<https://energyflow.network/>
[Komiske, Metodiev, JDT, JHEP 2019]



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]

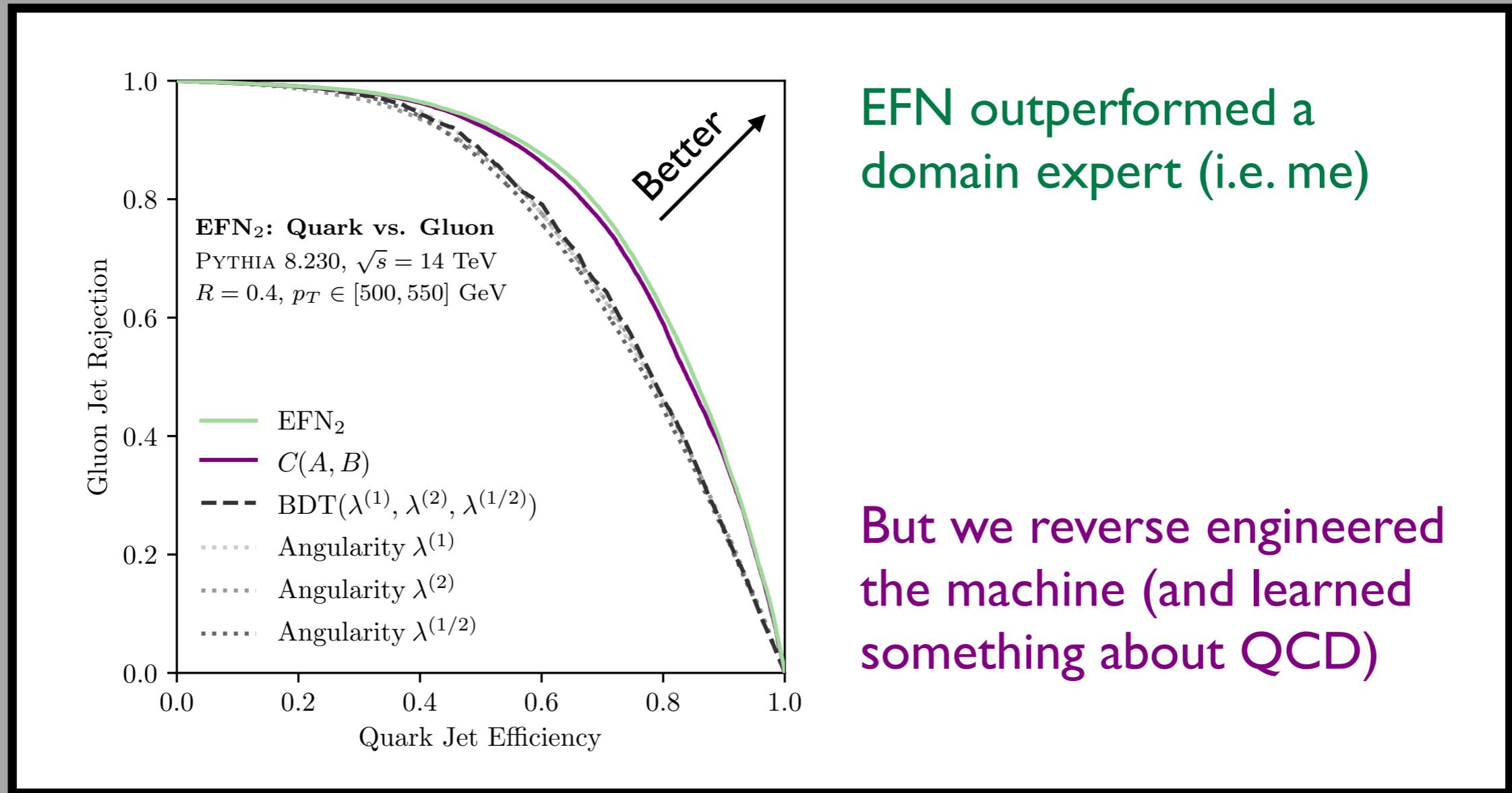
Learning from the Machine



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

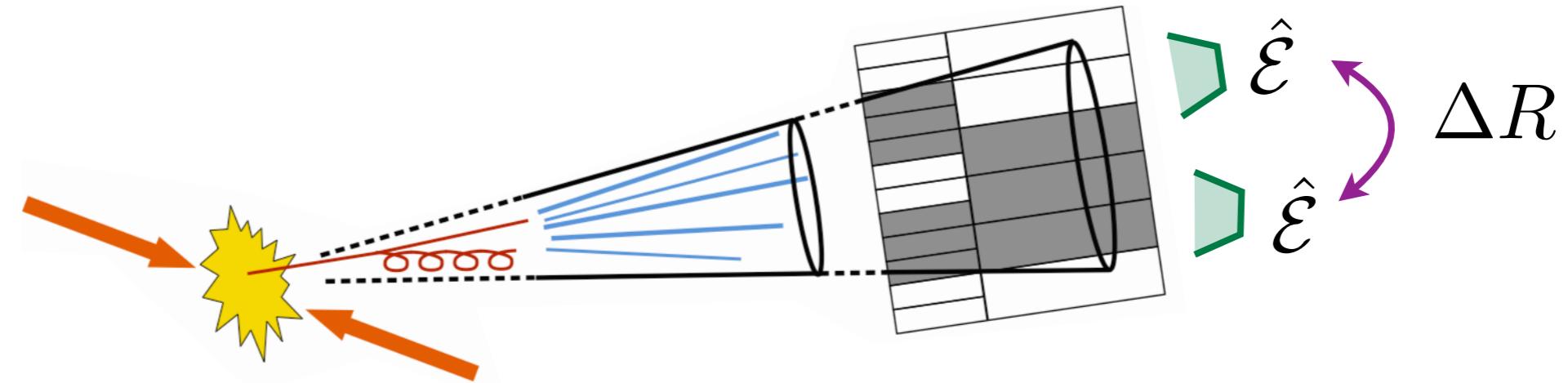
$$\sum_{i \in \text{jet}} z_i f(\theta_i)$$

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

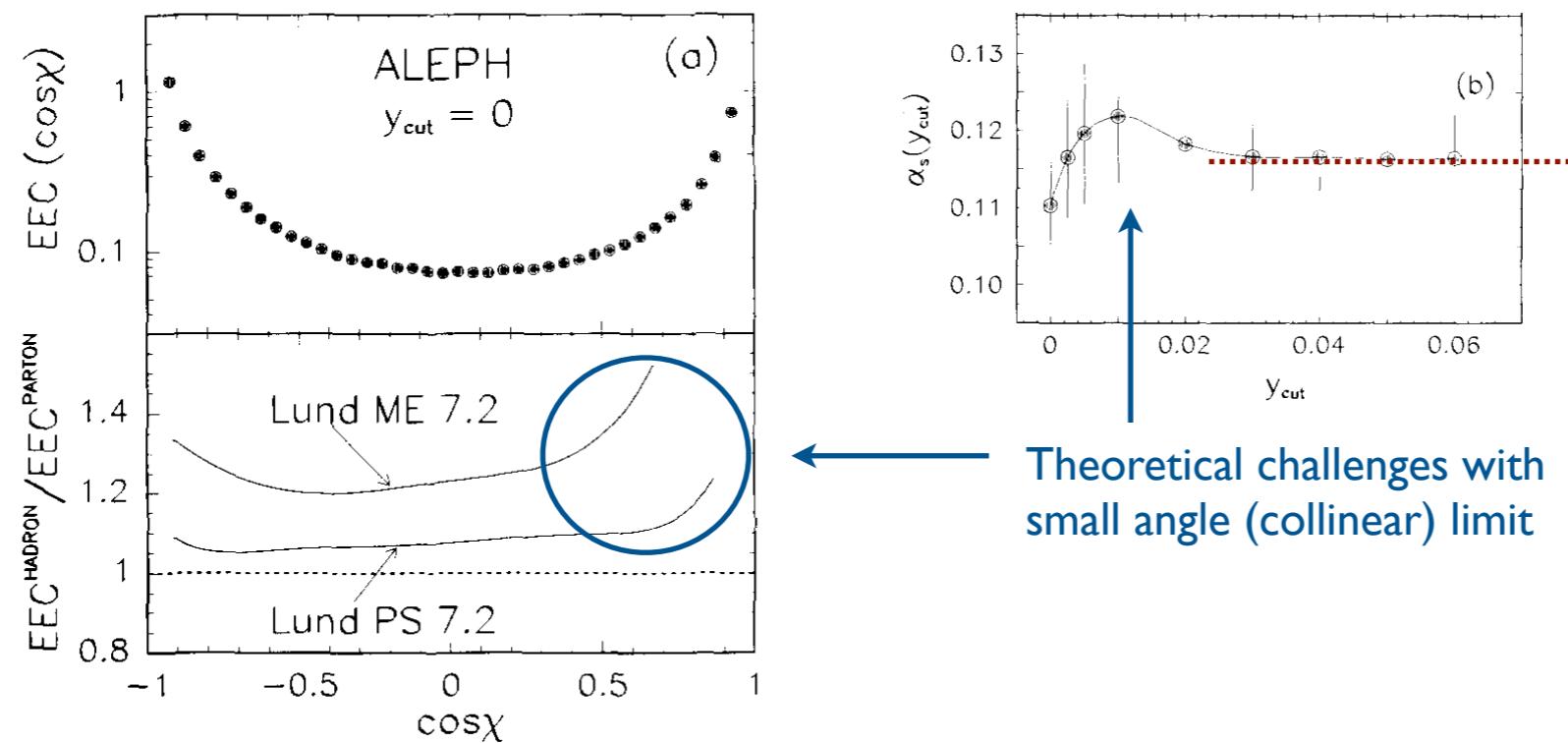


[Komiske, Metodiev, JDT, [JHEP 2019](#);
cf. Larkoski, JDT, Waalewijn, [JHEP 2014](#); using Berger, Kucs, Sterman, [PRD 2003](#); Ellis, Vermilion, Walsh, Hornig, Lee, [JHEP 2010](#)]

Energy-Energy Correlators

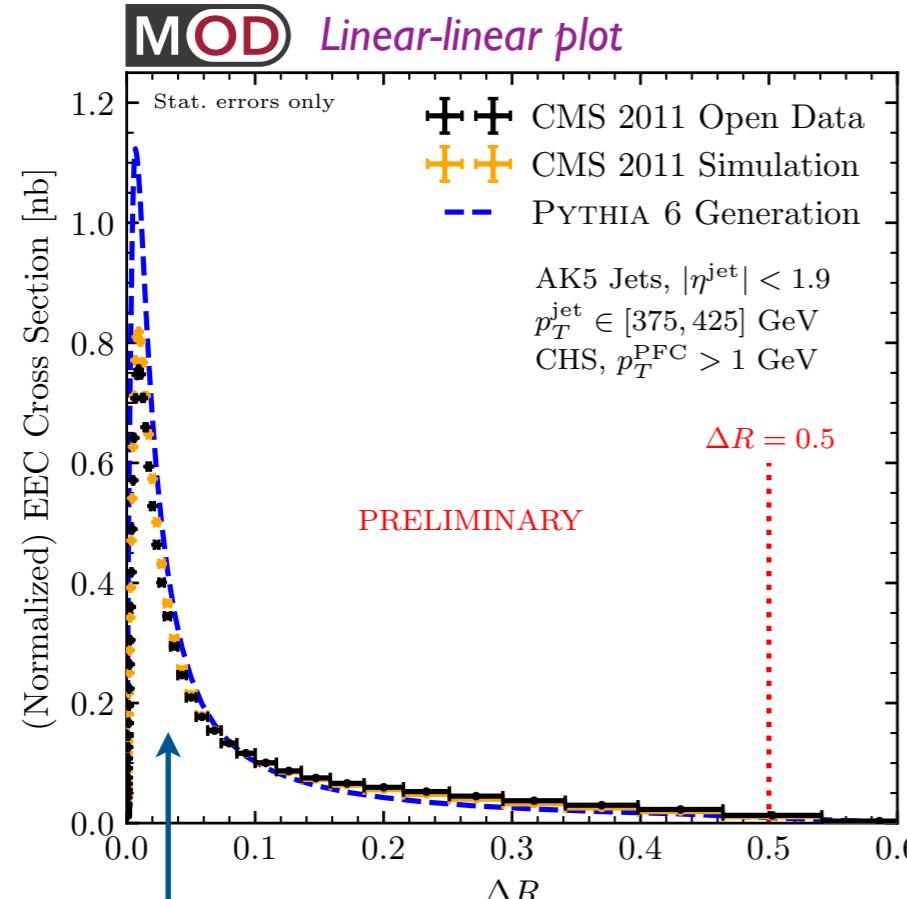


A long history in probing collinear dynamics of QCD



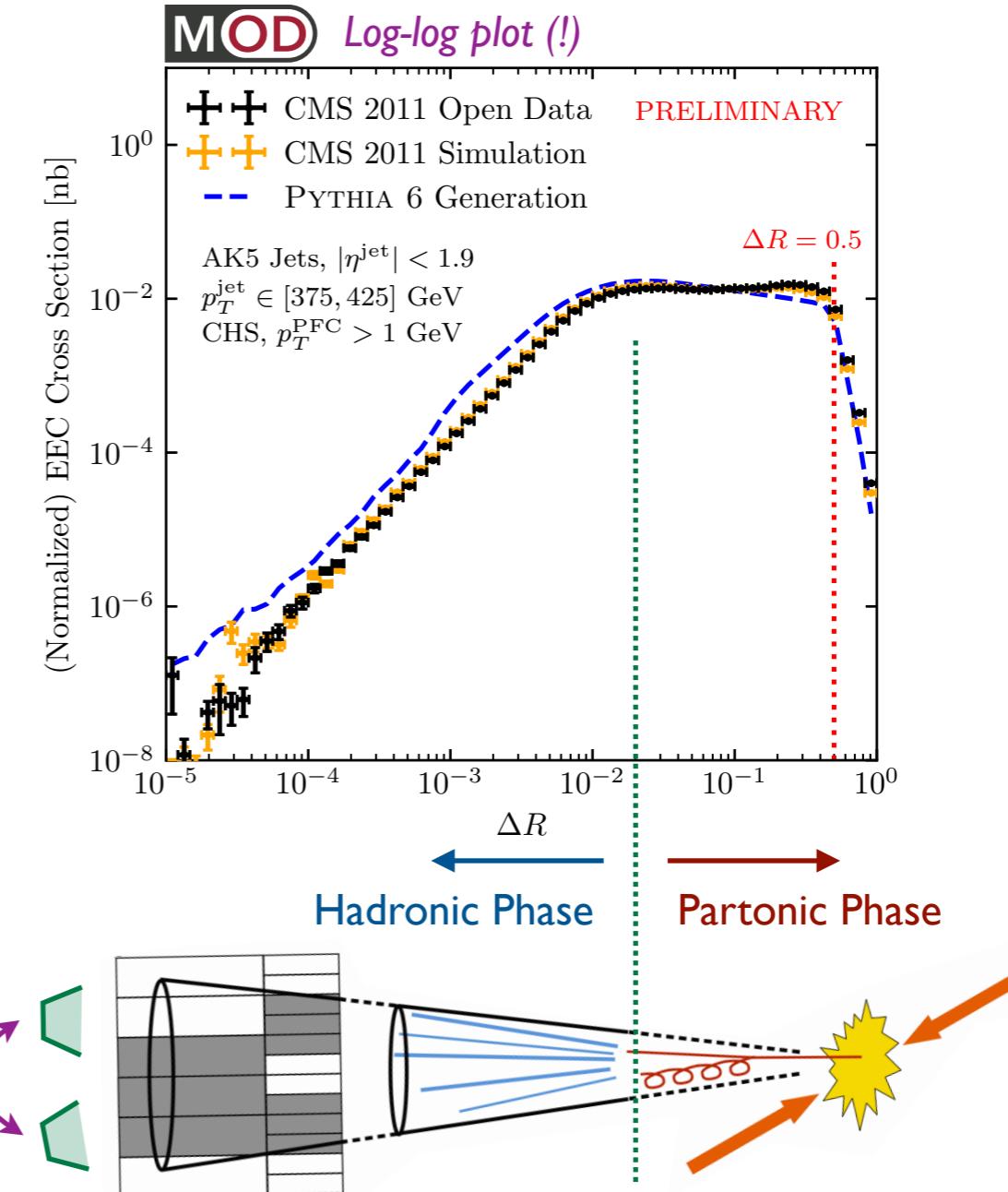
[Basham, Brown, Ellis, Love, [PRL 1978](#); ALEPH, [PLB 1991](#); see Chen, Moult, Zhang, Zhu, [PRD 2020](#)]

QCD Phase Transition in Jets?



Are we learning something about small angle limit of QCD?

First Jet EEC Plot from the LHC (!)

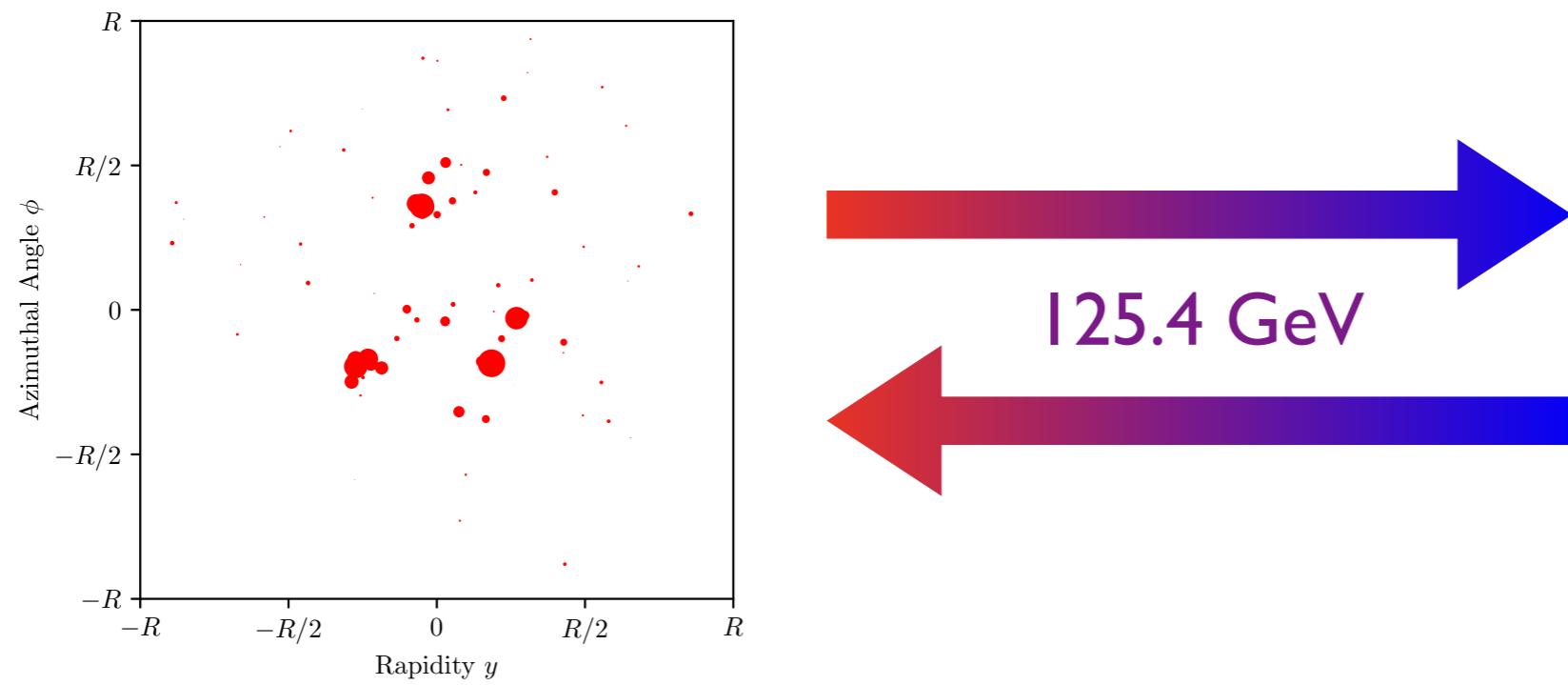
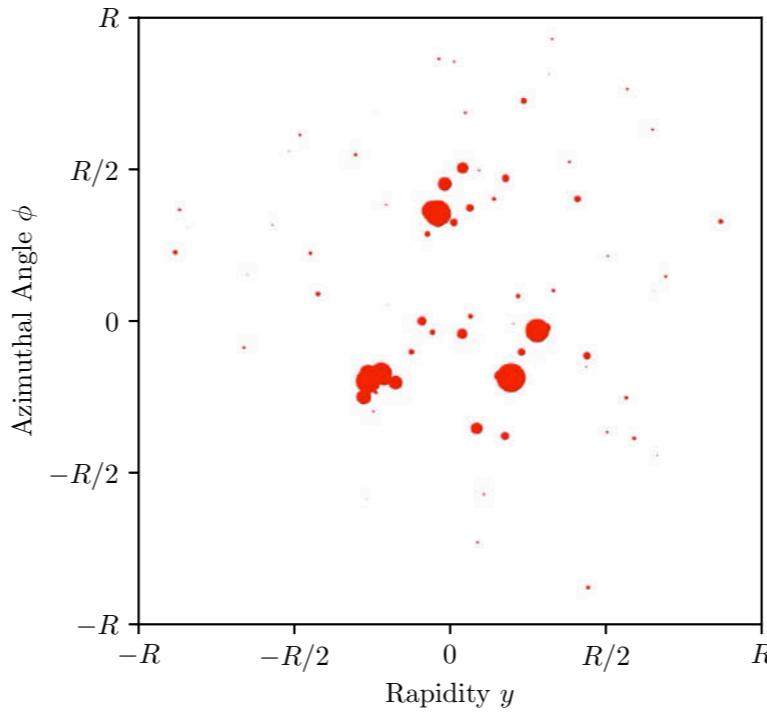


[Komiske, Moult, JDT, Zhu, in progress; see talks by Moult, [BOOST 2019](#), [BOOST 2020](#)]

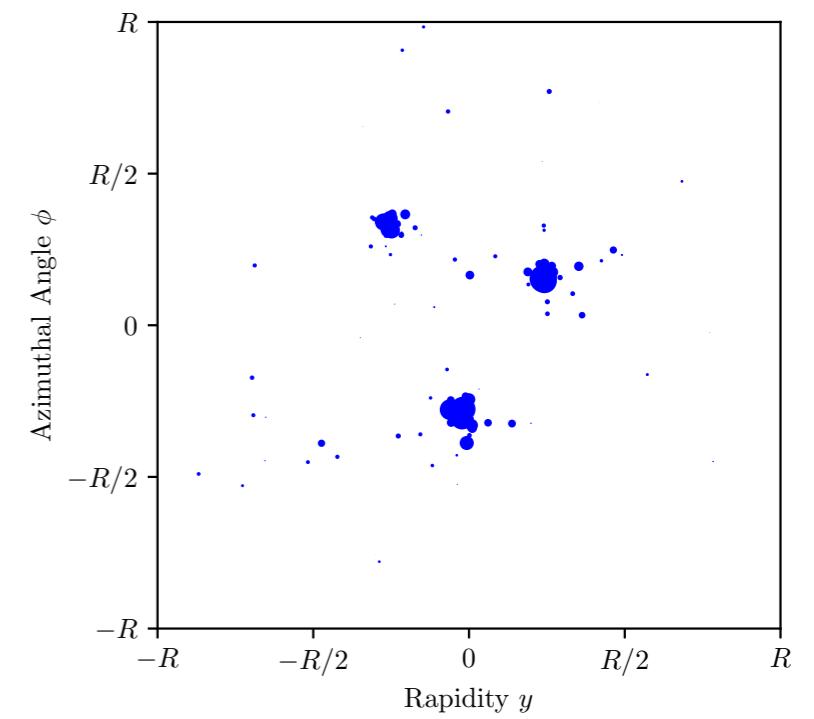


Similarity of Two Energy Flows?

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



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



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

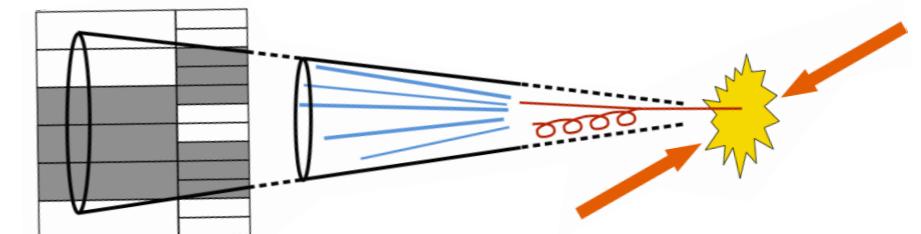
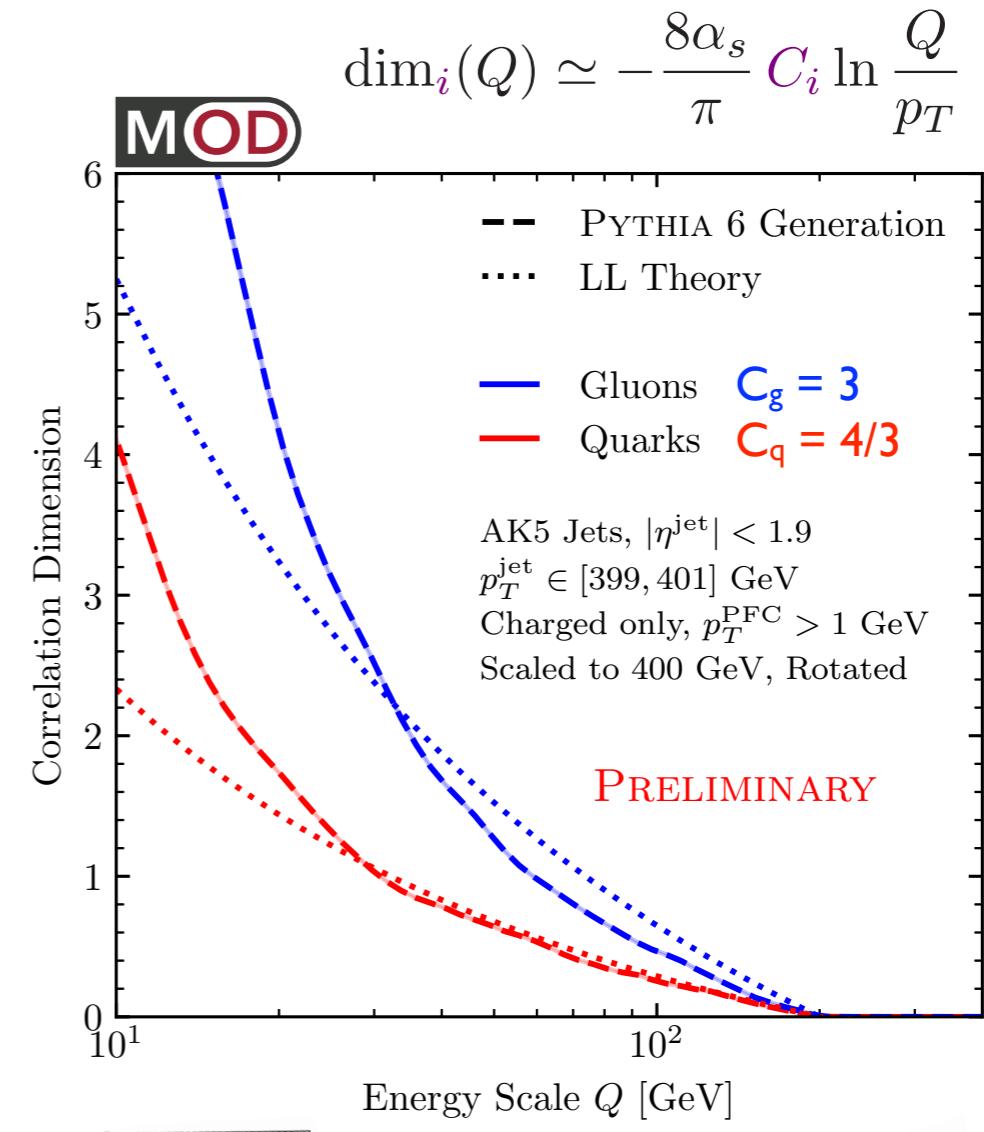
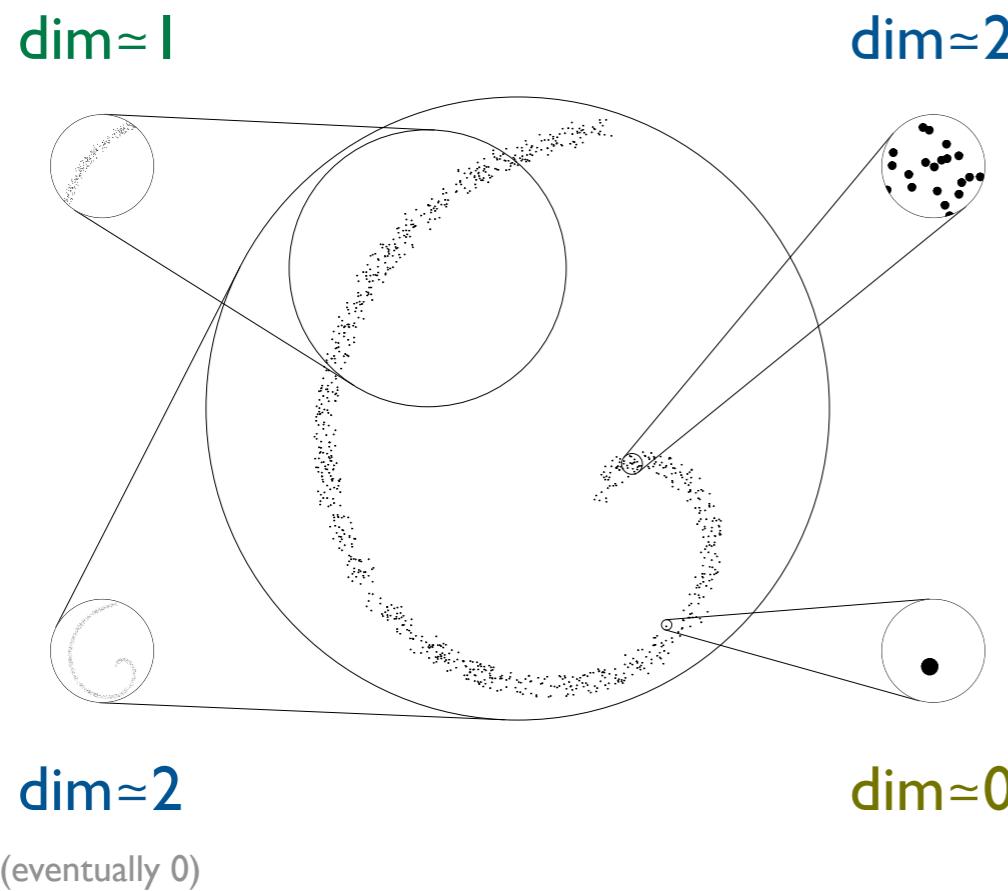
Dimensionality of Space of Jets



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

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

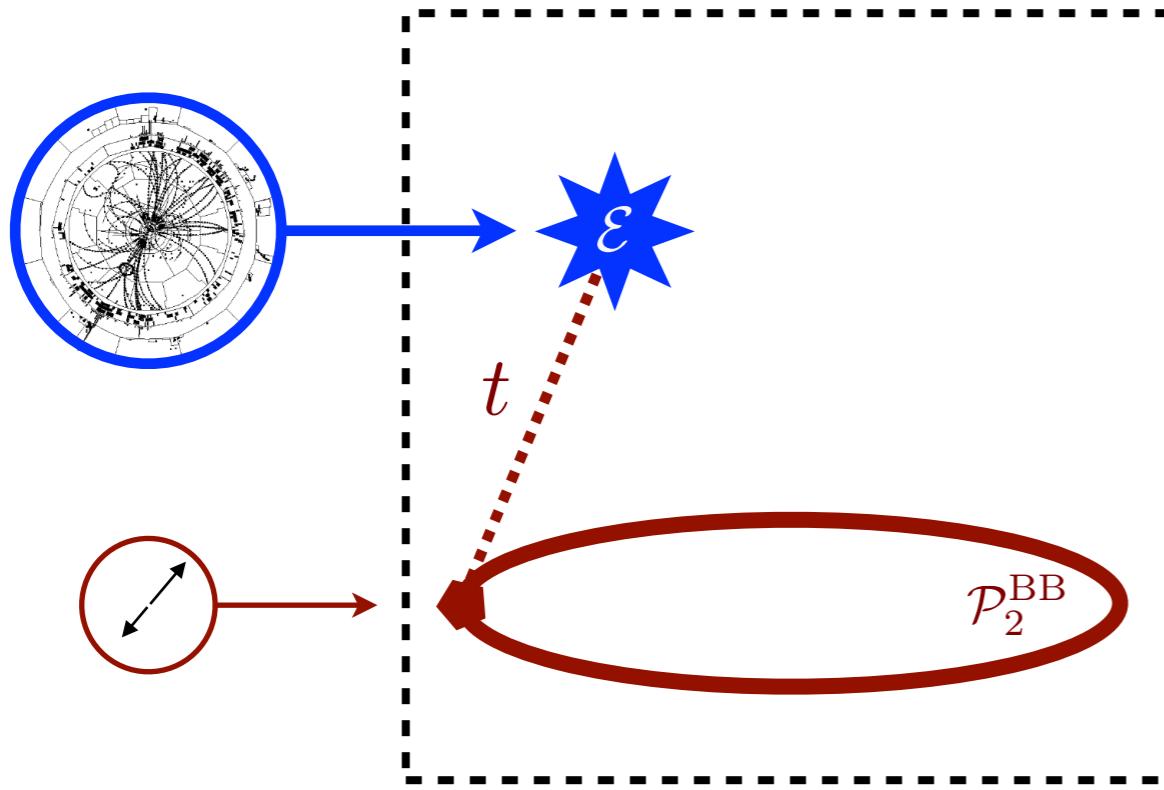
[Grassberger, Procaccia, [PRL 1983](#); Kégl, [NIPS 2002](#)]



Back to the Future with Thrust

How dijet-like is an event?

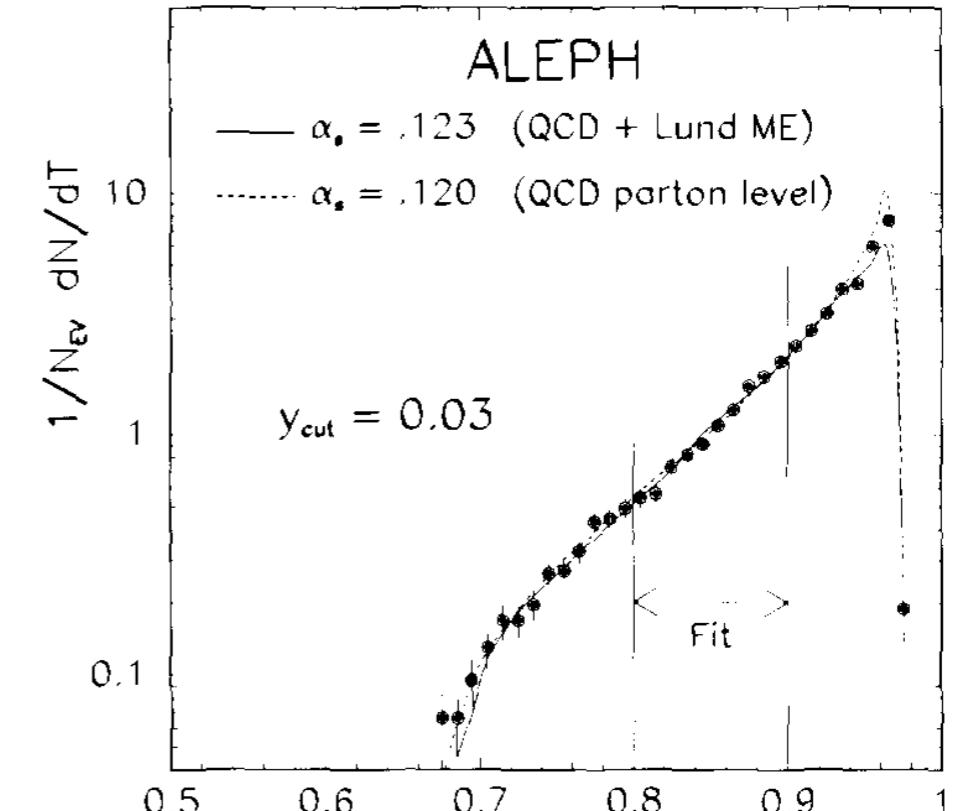
$$t(\mathcal{E}) = \min_{\mathcal{E}' \in \mathcal{P}_2^{\text{BB}}} \text{EMD}_2(\mathcal{E}, \mathcal{E}')$$



All Back-to-Back Two Particle Configurations

$$\mathcal{P}_2^{\text{BB}} = \left\{ \begin{array}{c} \text{red circles with arrows} \\ \dots \end{array} \right\}$$

(using $\beta=2$ EMD variant)



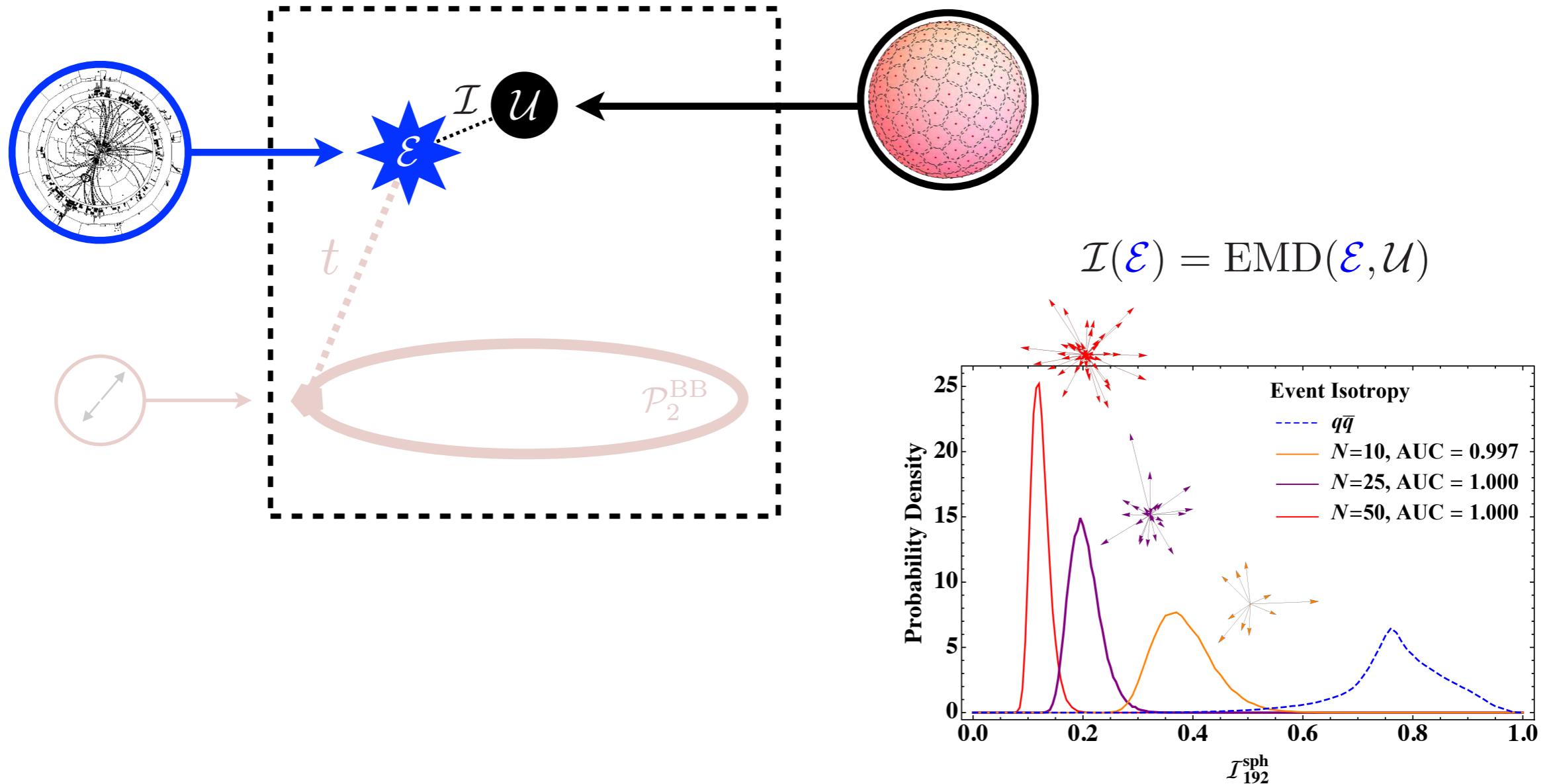
$$1 - \frac{t}{2E_{\text{CM}}}$$

(flipped, linear version of
ALEPH thrust plot from before)

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

Event Isotropy from Collider Geometry

How isotropic is an event?



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

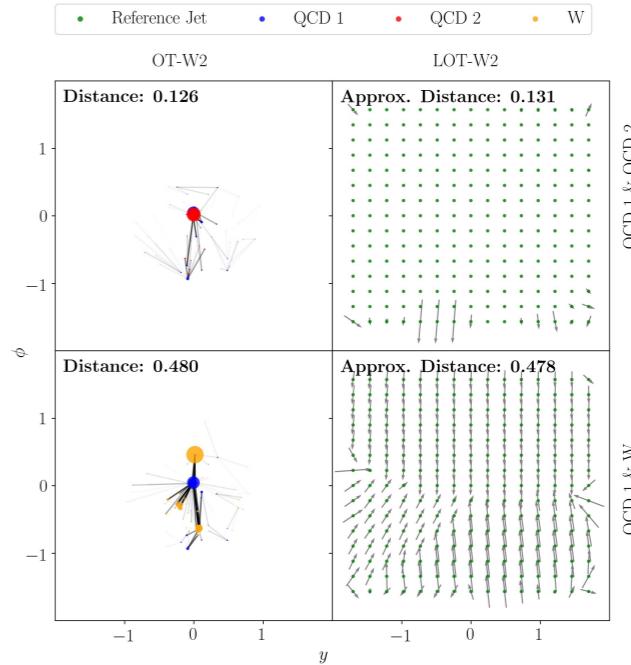


“Sounds cool, but isn’t this computationally daunting?”

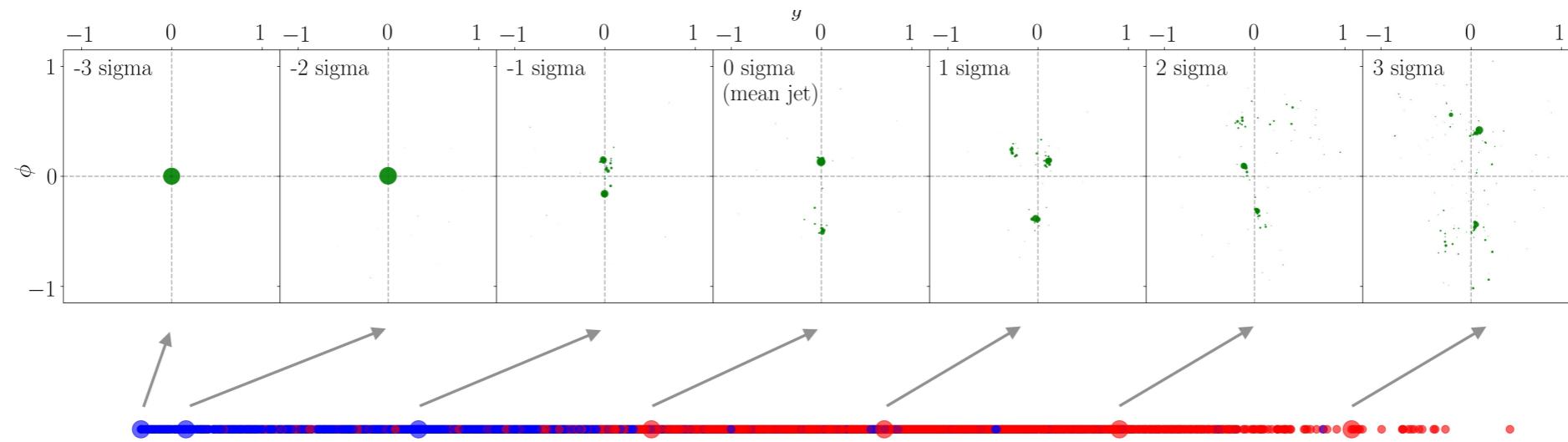
Linearized Optimal Transport

With the help of a reference event, transportation distances* can be efficiently mapped to Euclidean distances

* assuming the 2-Wasserstein measure



Enables coordinate-based techniques like Linear Discriminate Analysis



[Cai, Cheng, Craig, Craig, PRD 2020]