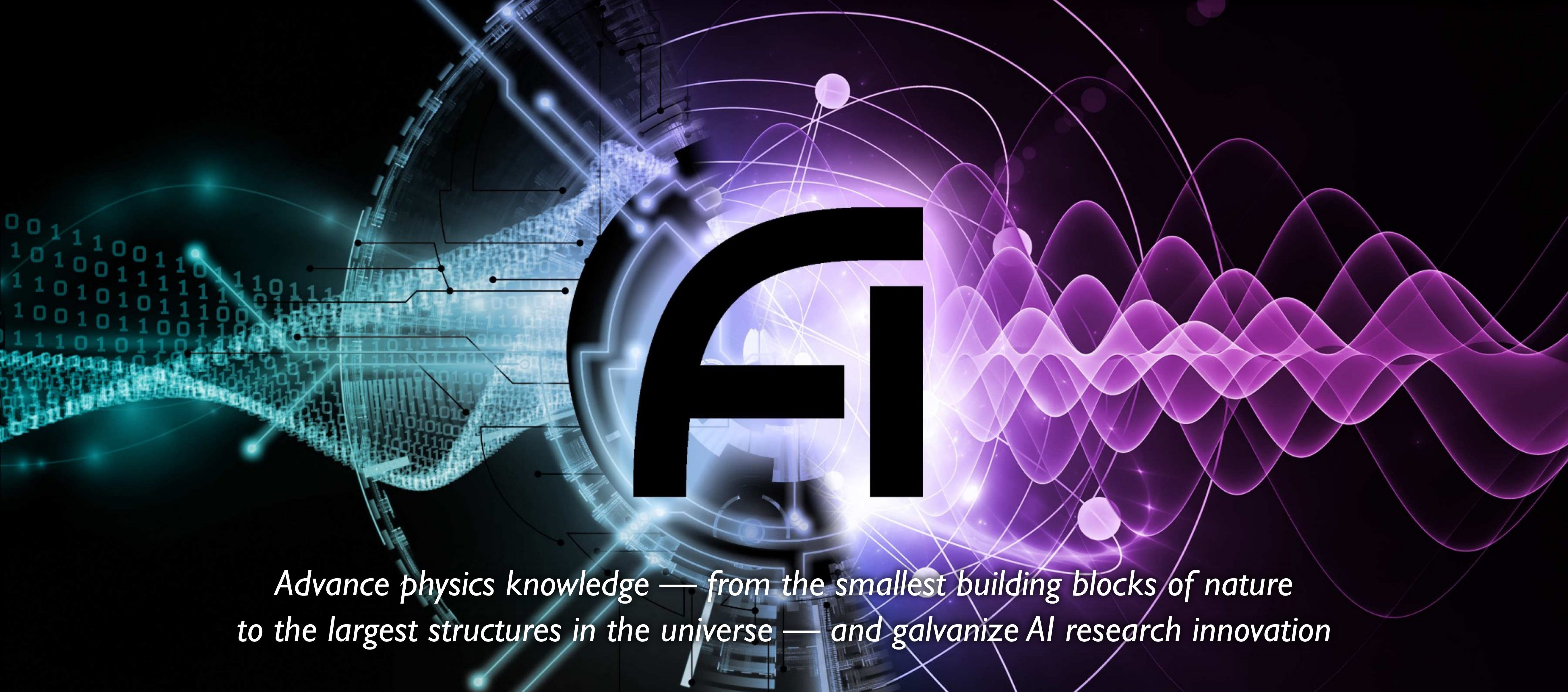
Machine Learning for Fundamental Physics

Jesse Thaler

MIT

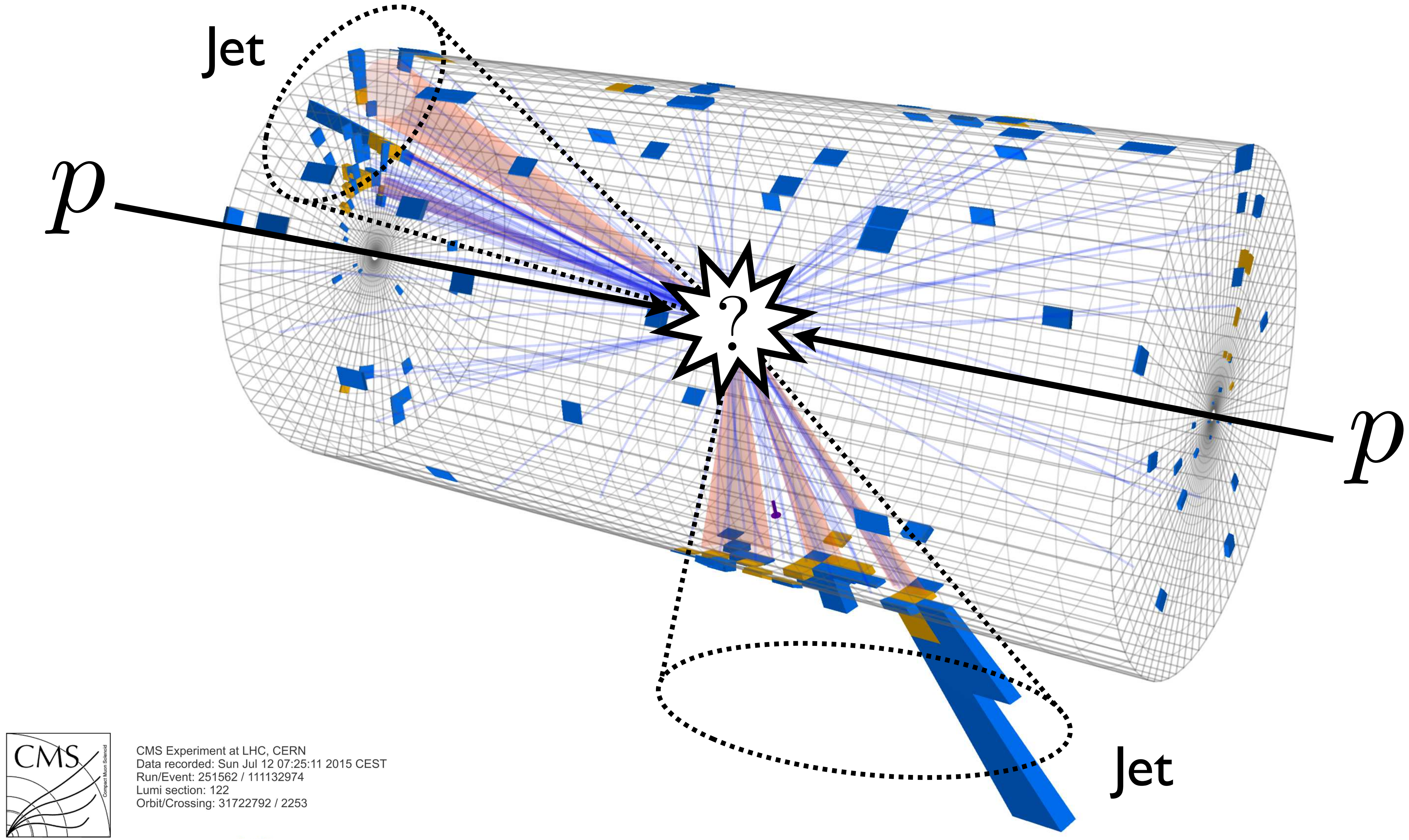
HKUST Jockey Club Institute for Advanced Study, HEP 2021 — January 21, 2021

The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI) *“eye-phi”*



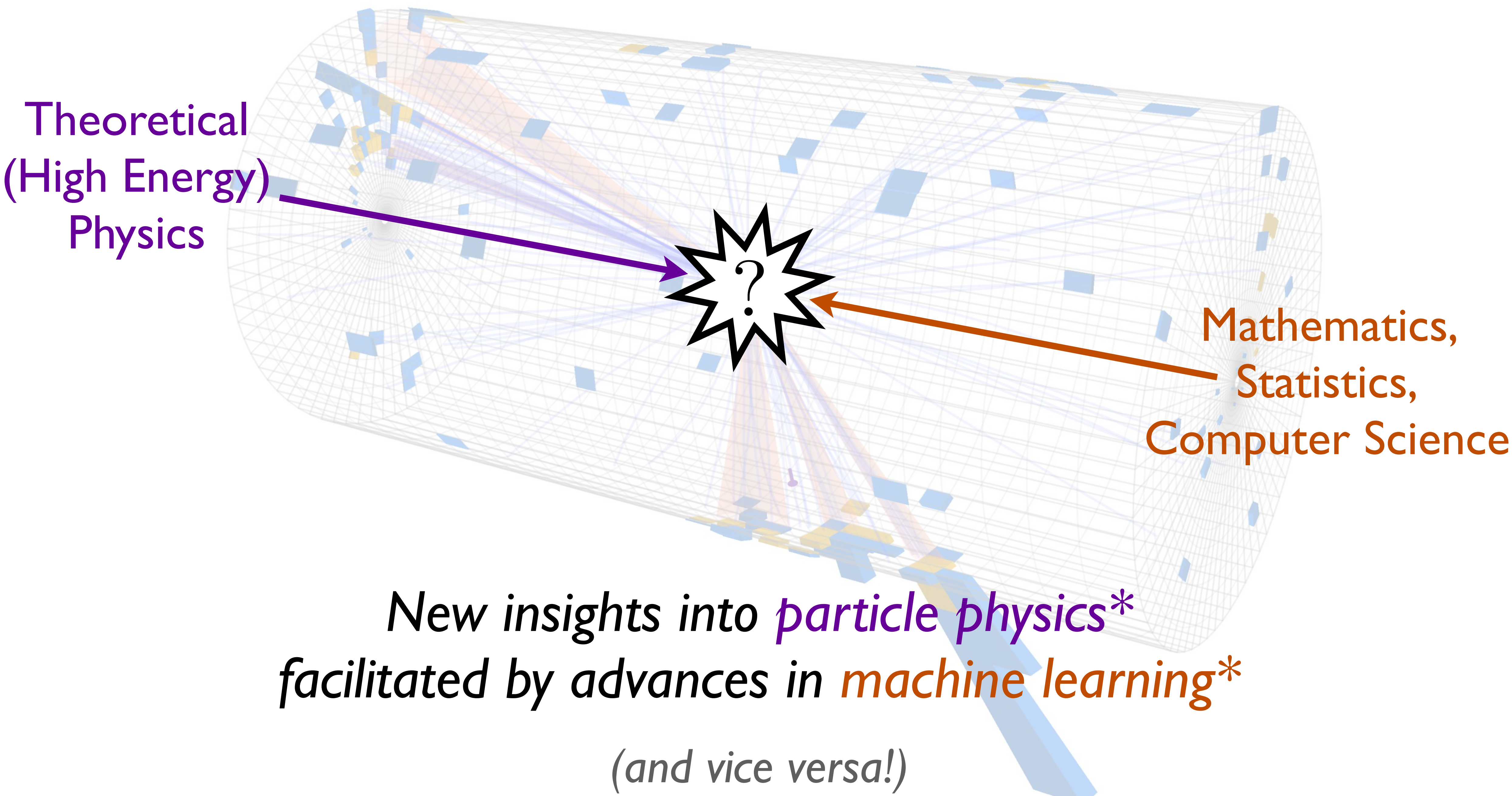
Advance physics knowledge — from the smallest building blocks of nature to the largest structures in the universe — and galvanize AI research innovation

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


Jet
p
?
p
Jet
CMS
Compact Muon Solenoid
CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253

"Collision Course"

"Theoretical Physics for Machine Learning"
Aspen Center for Physics, January 2019

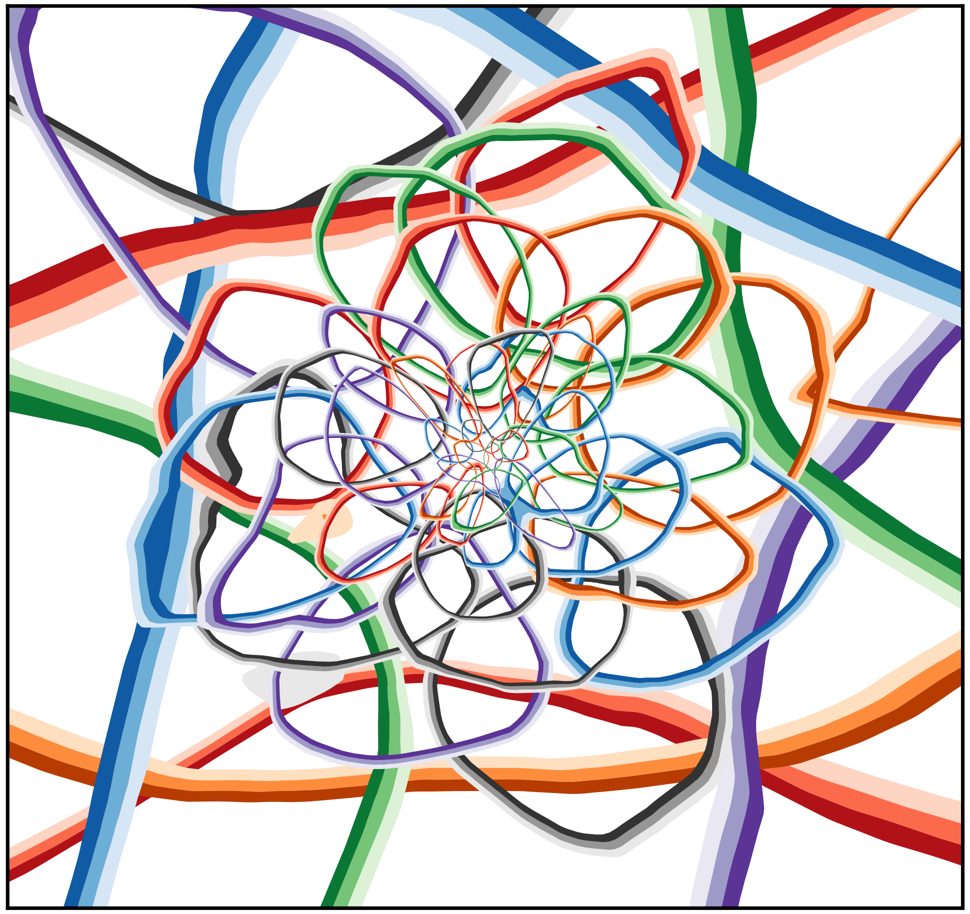


*New insights into particle physics** *facilitated by advances in machine learning**

(and vice versa!)

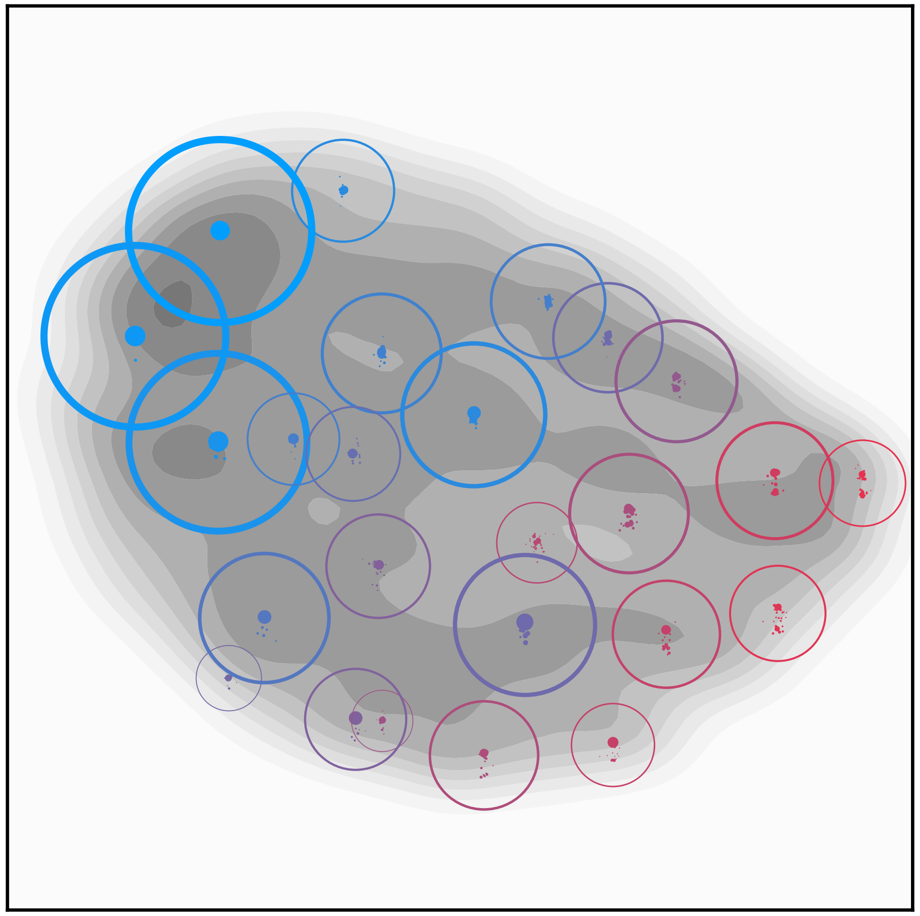
Two Anecdotes

Teaching a Machine to
"Think Like a Physicist"



[Komiske, Metodiev, JDT, JHEP 2019]

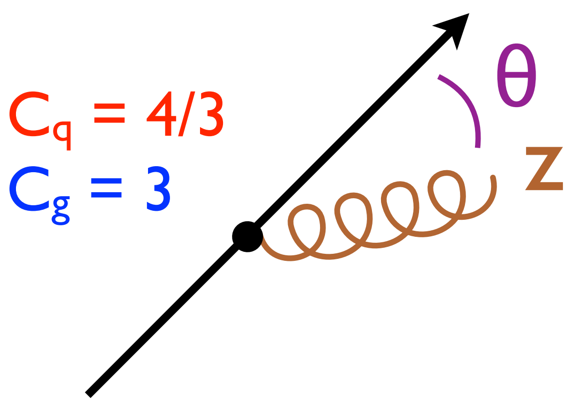
Letting Collider Data
Speak for Itself



[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

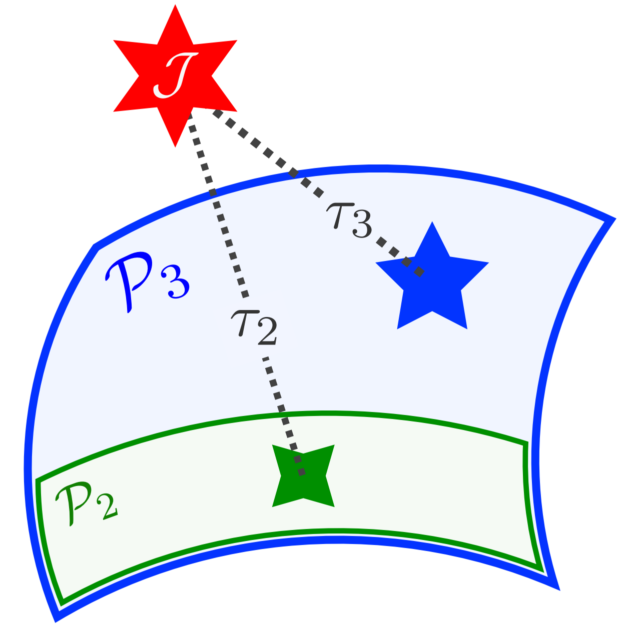
Exploiting a Core Prediction of QCD



$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{\mathrm{d}\theta}{\theta} \frac{\mathrm{d}z}{z}$$

[Altarelli, Parisi, NPB 1977]

Nested Singularities of Gauge Theories

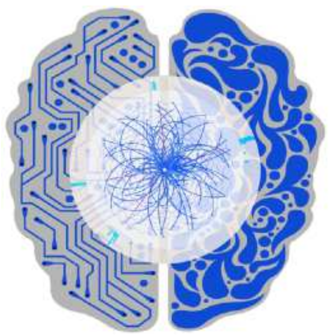


$$\tau_N(\mathcal{J}) = \min_{\mathcal{J}' \in \mathcal{P}_N} \mathrm{EMD}(\mathcal{J}, \mathcal{J}')$$

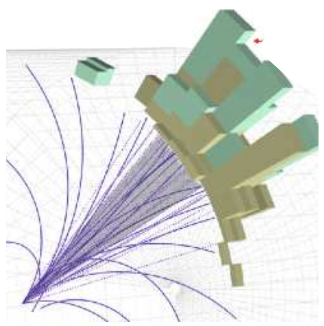
[Stewart, Tackmann, Waalewijn, PRL 2010;
JDT, Van Tilburg, JHEP 2011, JHEP 2012;
rephrased via Komiske, Metodiev, JDT, JHEP 2020]

New perspectives on *key theoretical concepts*

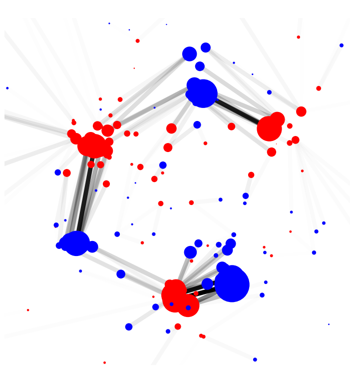
Outline



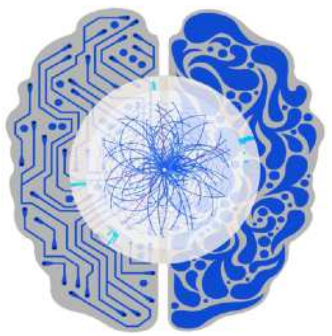
Rise of the Machines?



What is a Collider Event?



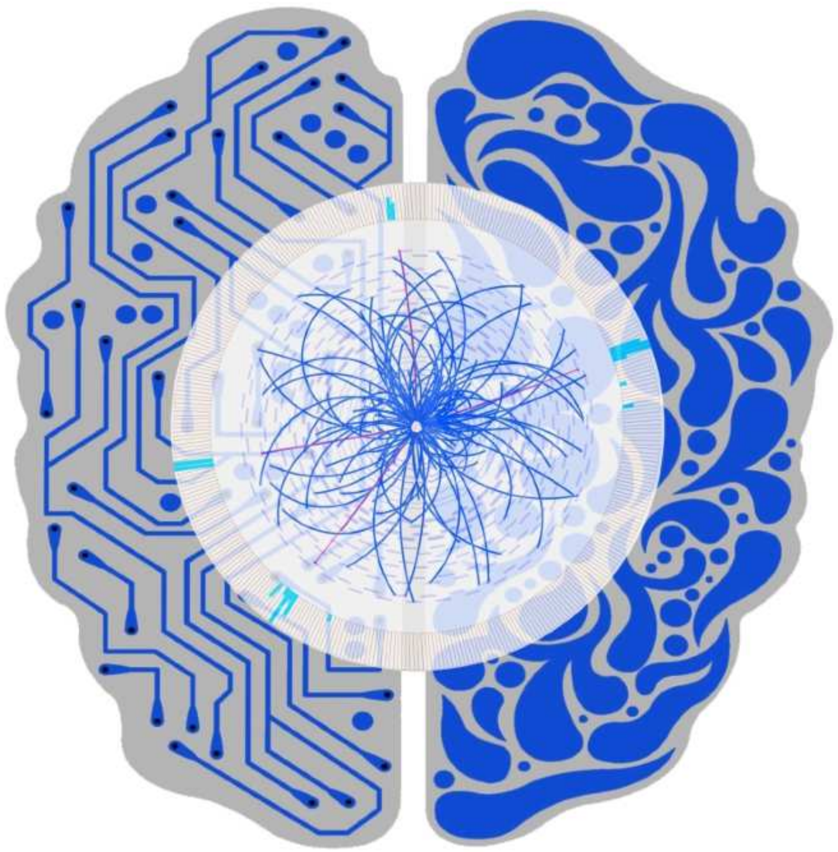
When are Collider Events Similar?



Rise of the Machines?

What is a Collider Event?

When are Collider Events Similar?



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?

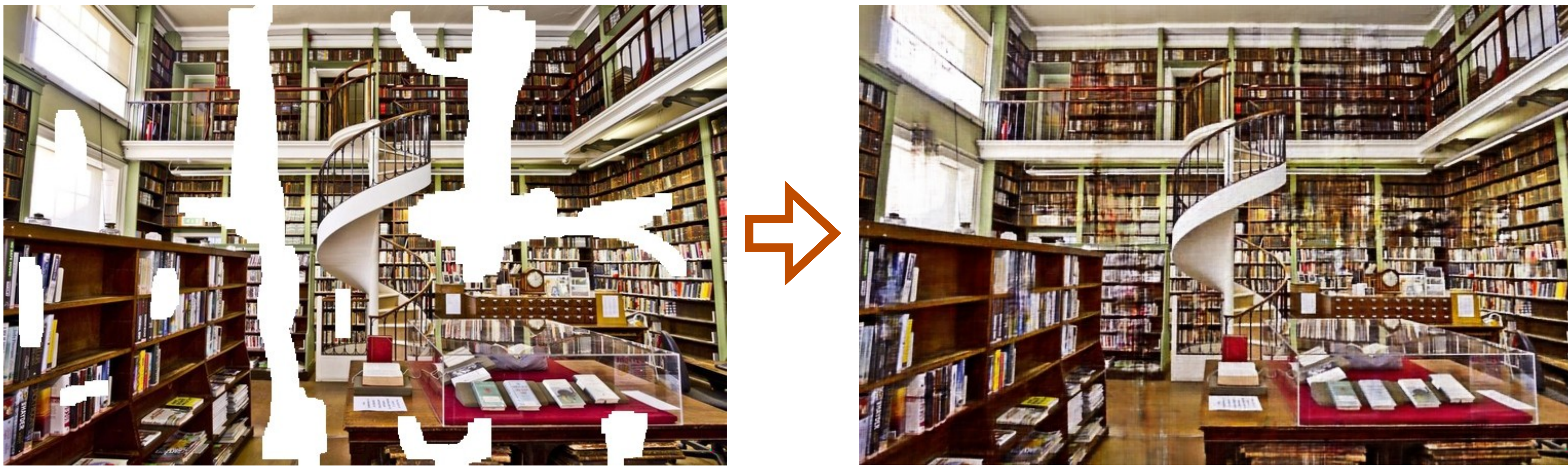
AI^2: Ab Initio Artificial Intelligence

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

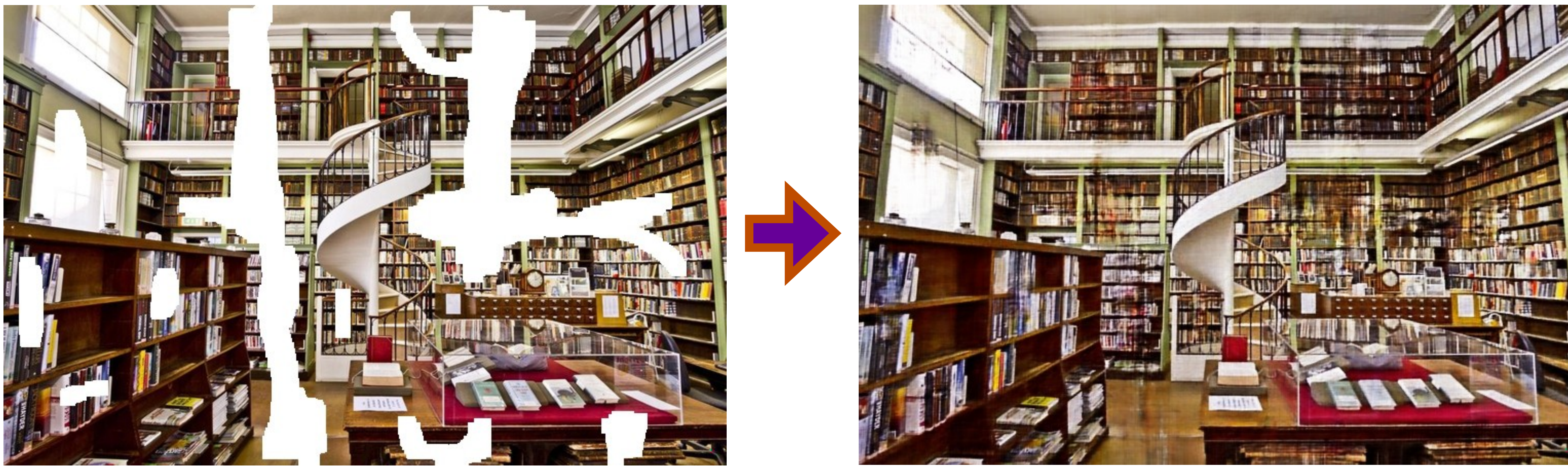


increased computational power and large data sets

[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]

ML Targets for Collider Theory

Master formula for colliders:

$$\sigma_{\text{obs}} \simeq \frac{1}{2E_{\text{CM}}^2} \sum_{n=2}^{\infty} \int d\Phi_n \, |\mathcal{M}_{AB\to 12\ldots n}|^2 \, f_{\text{obs}}(\Phi_n)$$

cross section phase space amplitude observable

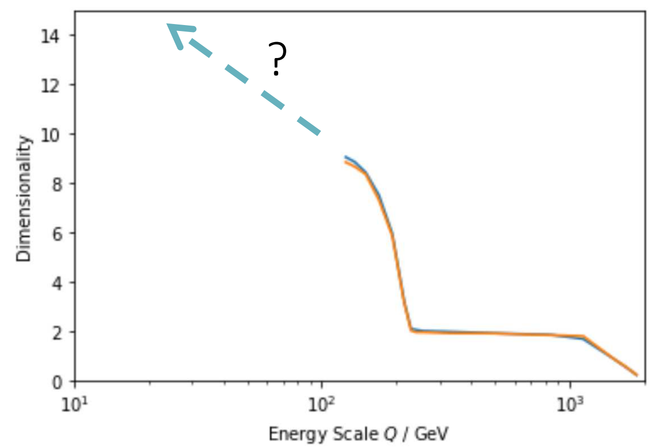
Exciting progress on multiple fronts from theoretical (and experimental) HEP communities!

Progress in other areas of fundamental physics as well,
e.g. cosmology, dark matter, string theory, nuclear theory, …

[apologies for focus on research from my group in this talk; see HEPML-LivingReview for extensive bibliography]

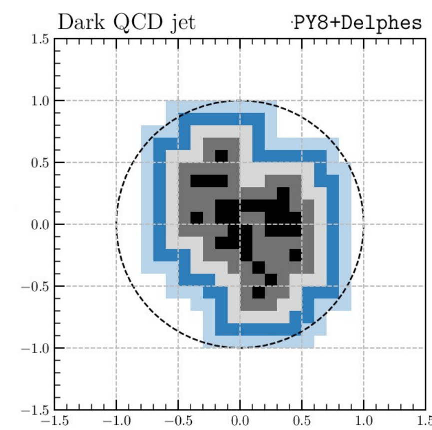
Mini-Workshop: Machine Learning and Open Data

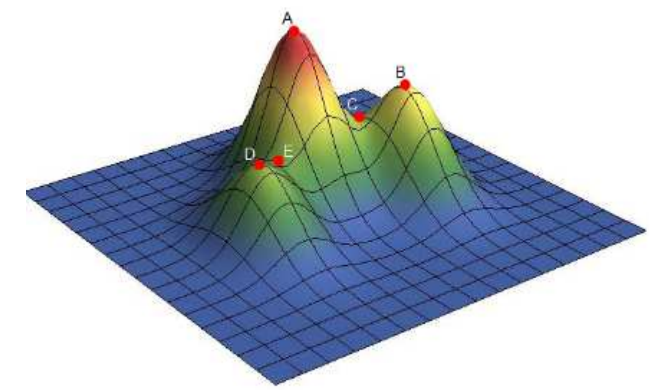
Dimensionality via Autoencoding



[Talk by Jack Collins]

Jet Classification via Topology

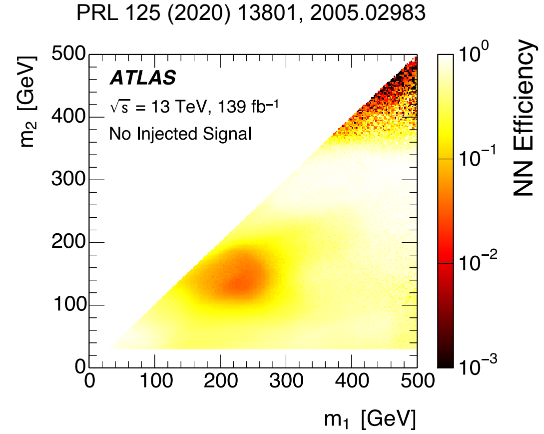




[Talks by Mihoko Nojiri & Lingfeng Li]

Anomaly Detection

ATLAS Collaboration
PRL 125 (2020) 13801, 2005.02983



[Talks by David Shih & Ben Nachman]

CMS Open Data for BSM Searches

Selection	Data	Signal BM
MET primary	4.3×10^7	-
$p_T^{j1} > 150$ GeV, $E_T^{\text{miss}} > 150$ GeV	1.4×10^6	830
One displaced vertex ($N_{vtx,tk} \geq 2$)	3.7×10^5	310
One displaced vertex ($N_{vtx,tk} \geq 3$)	4.7×10^4	240
One displaced vertex ($N_{vtx,tk} \geq 4$, default)	5.5×10^3	140
Two displaced vertices	76	9.8
$p_T^{j1} > 300$ GeV, $E_T^{\text{miss}} > 300$ GeV	1	3.0
Two displaced vertices with vertex $H_T < 40$	0	3.0

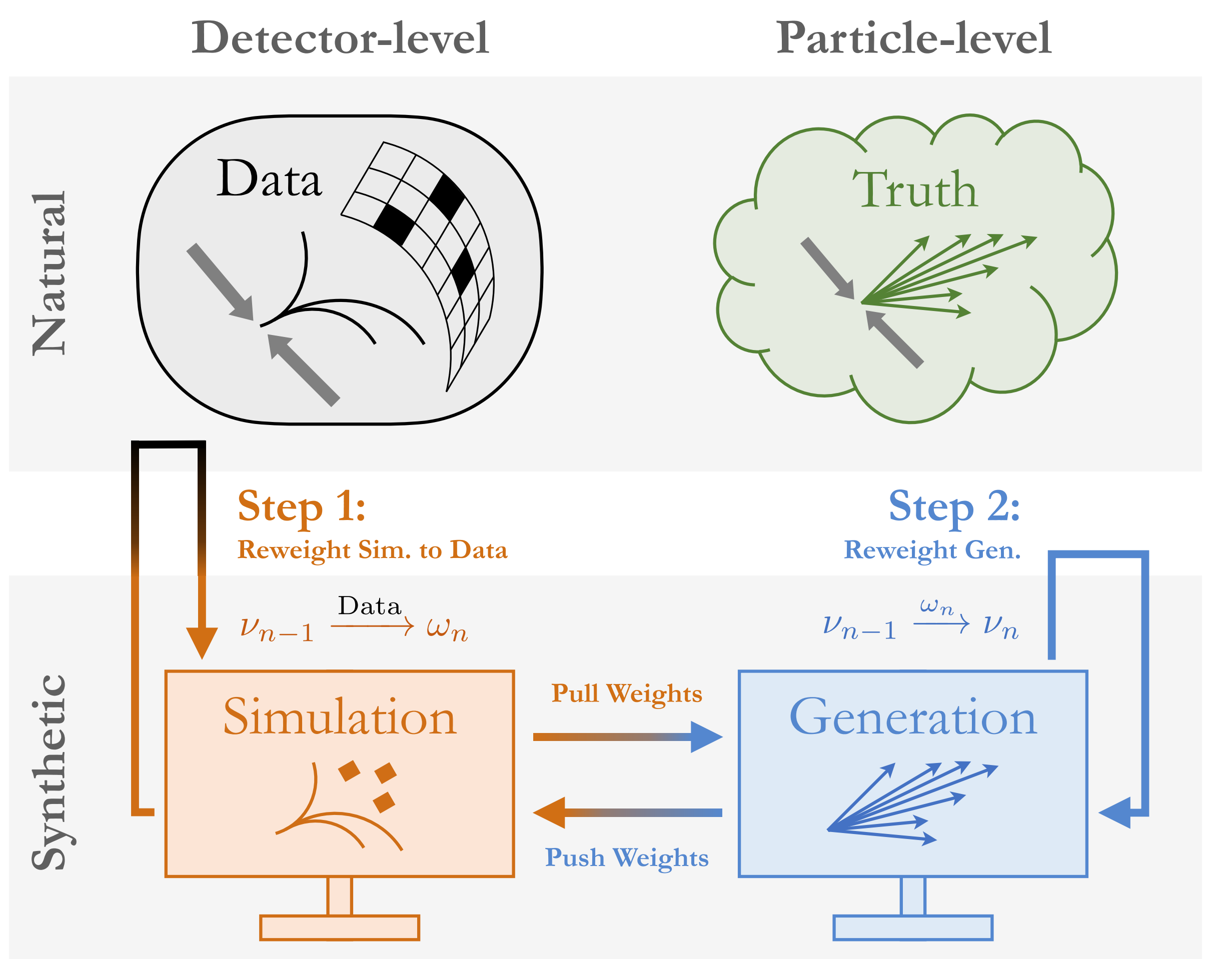
$$m_{\tilde{t}_1} = 360 \text{ GeV}, \ \Delta m = 20 \text{ GeV}$$

[Talk by Haipeng An]

Detector Unfolding

OmniFold

Multi-dimensional unbinned detector corrections via iterated binary classification

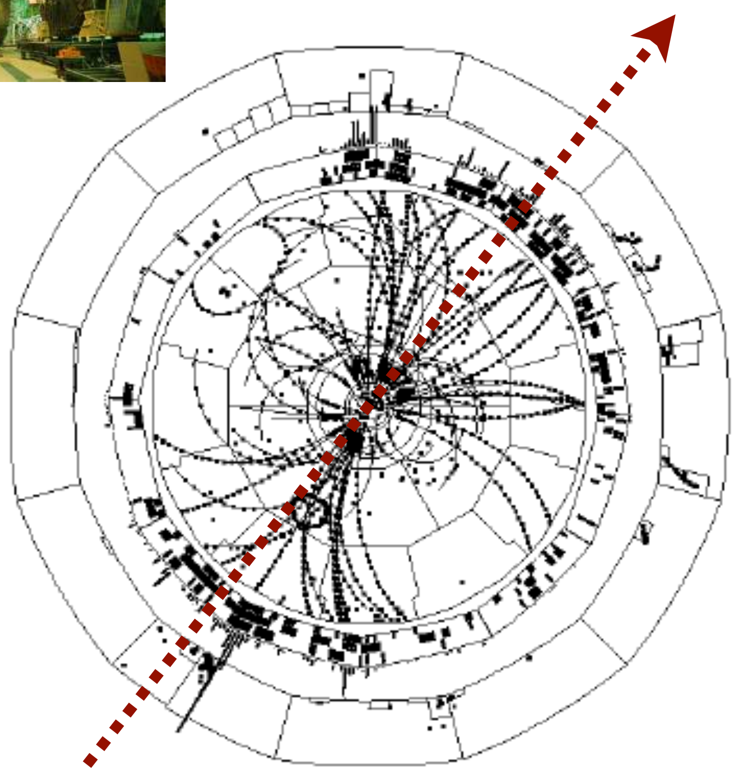


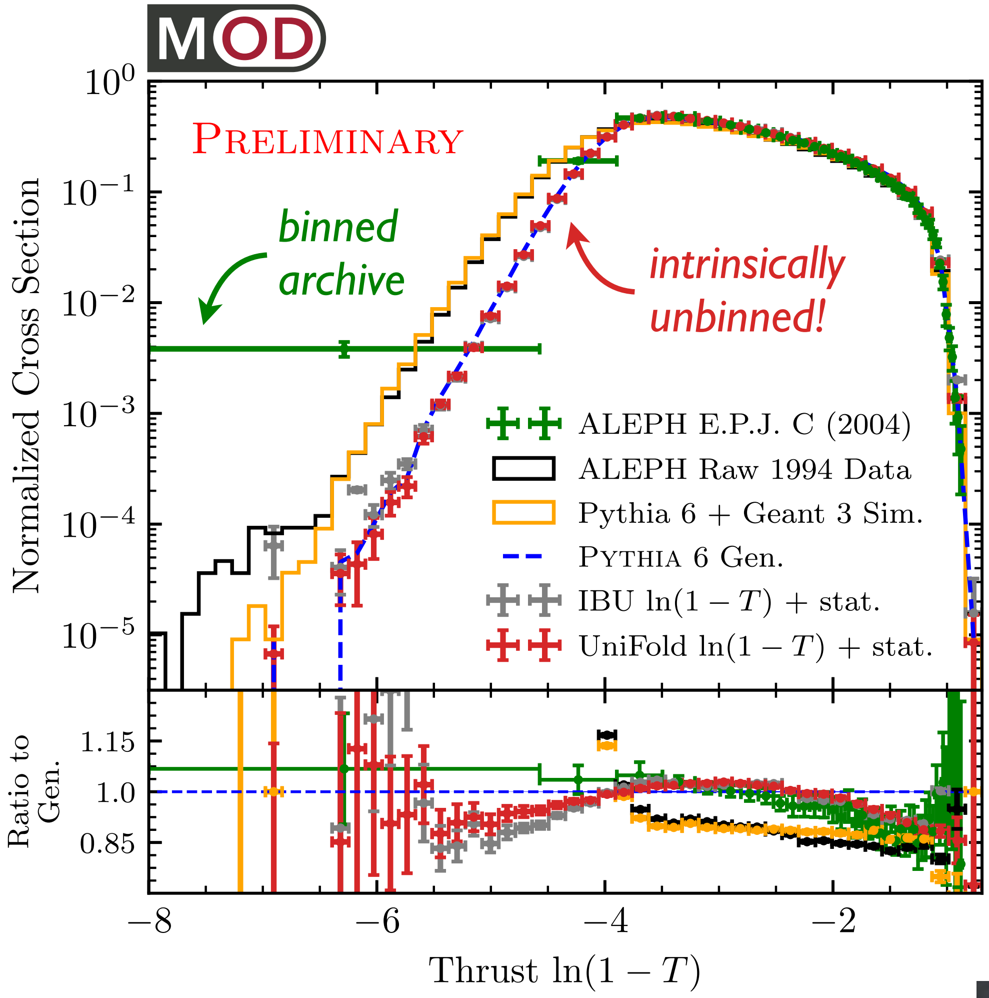
Use ML
to compute
likelihood ratio

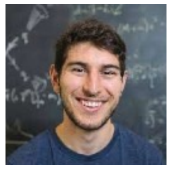
[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020]

Detector Unfolding

Back to the Future with ALEPH Archival Data

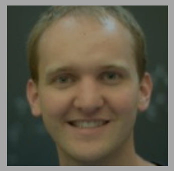


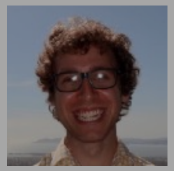




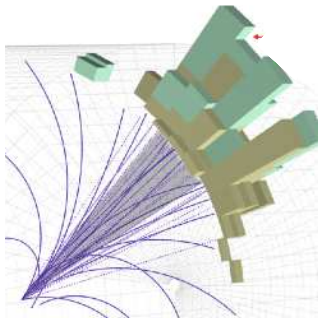
[talk by Badea, ICHEP 2020; cf. ALEPH, EPJC 2004]
[see also Badea, Baty, Chang, Innocenti, Maggi, McGinn, Peters, Sheng, JDT, Lee, PRL 2019]

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





Rise of the Machines?

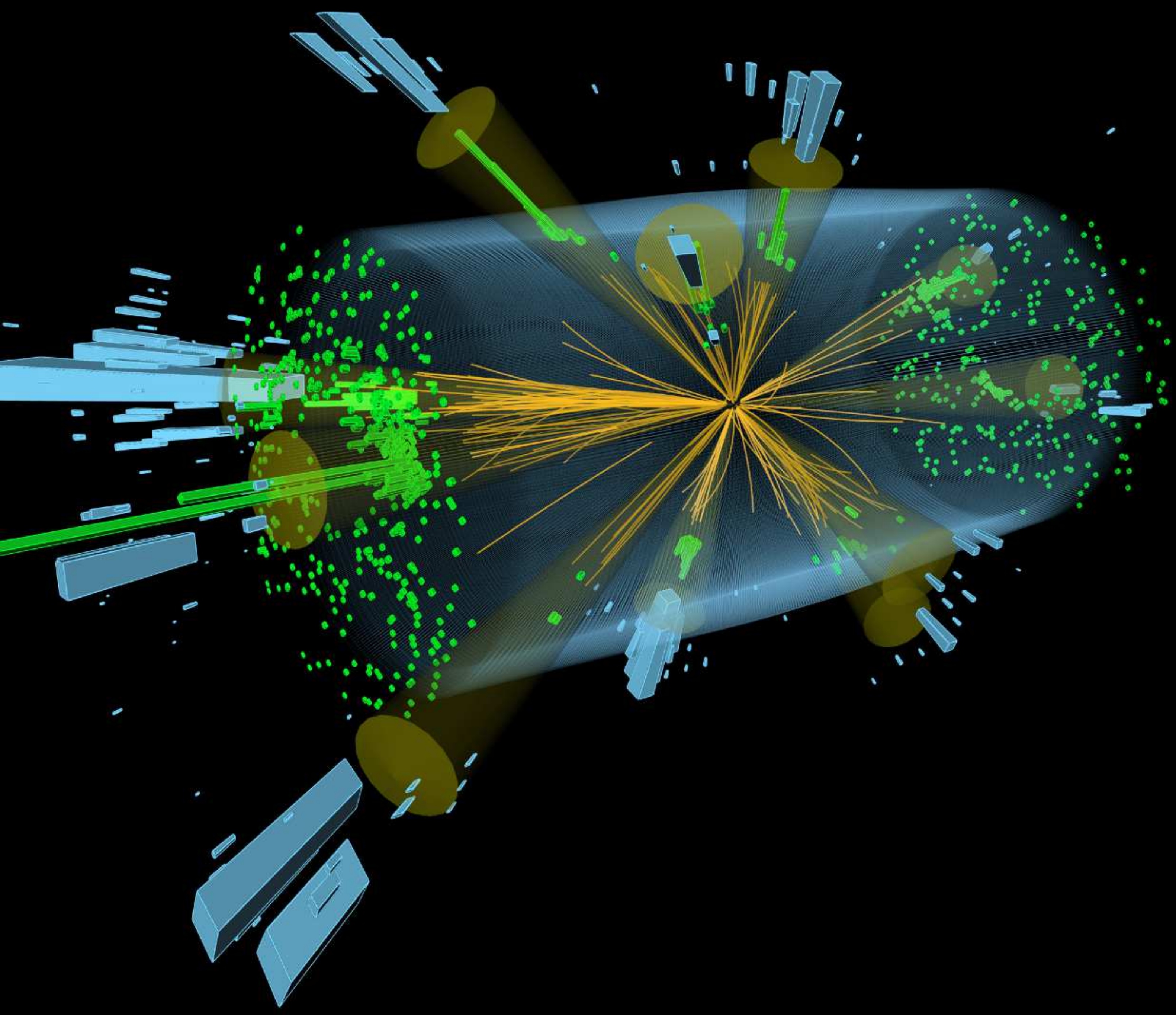


What is a Collider Event?

When are Collider Events Similar?

Collider Event

Collection of points in (momentum) space



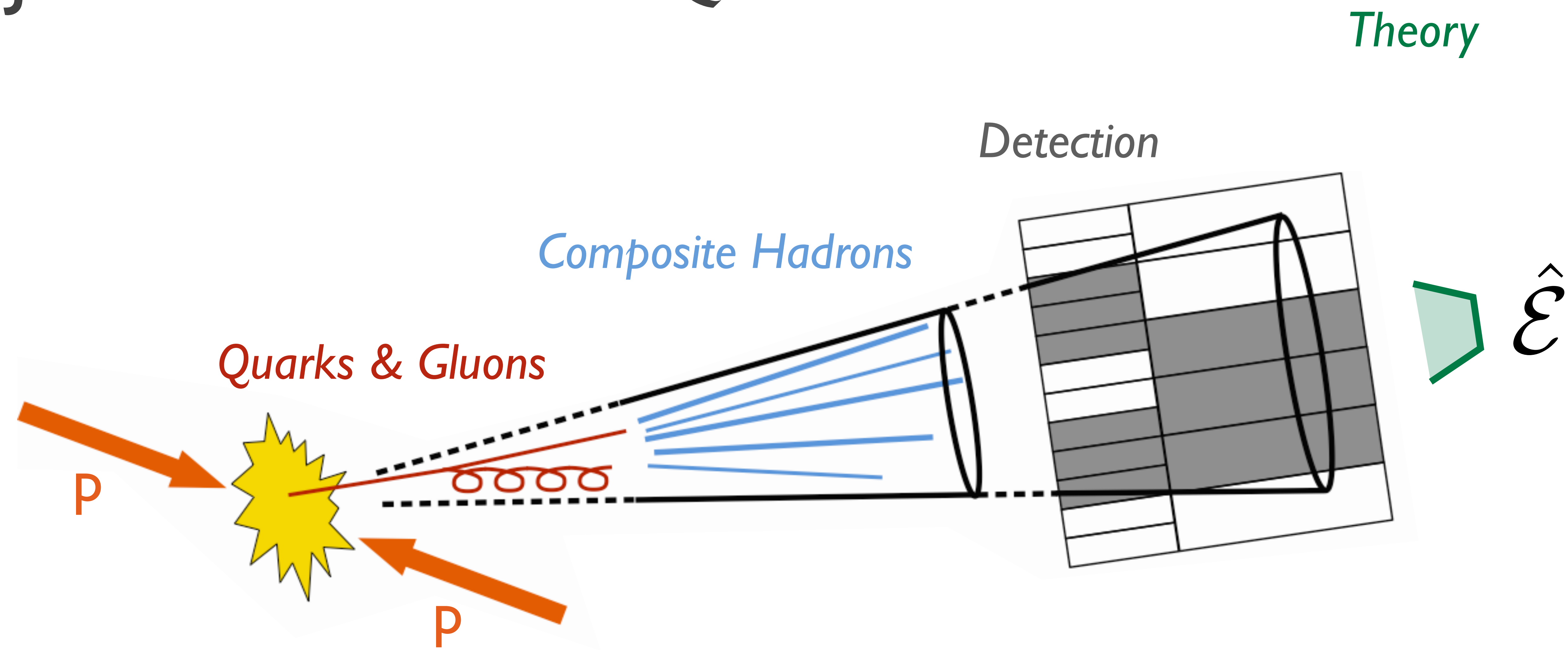
T	E	H	M			
	●			γ	photon	elementary
●	●			$e^{\pm}$	electron	elementary
●	•	•	●	$\mu^{\pm}$	muon	elementary
●	•	●		$\pi^{\pm}$	pion	composite
●	•	●		$K^{\pm}$	kaon	composite
	•	●		K^0_L	K-long	composite
●	•	●		$p/\overline{p}$	proton	composite
	•	●		$n/\overline{n}$	neutron	composite

Point Cloud

Collection of points in (position) space

[Popular Science, 2013]

Jet Formation from QCD



Energy Flow:

Robust to hadronization and detector effects
Well-defined for massless gauge theories

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

[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]

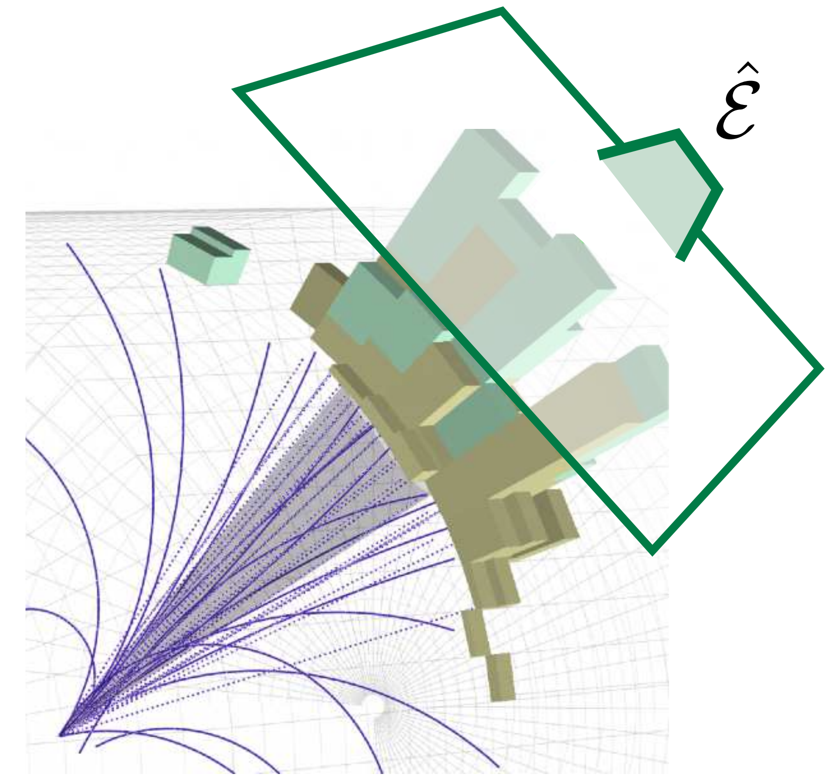
From Energy Flow to Weighted Point Clouds

- Energy-Weighted Directions

$$\vec{p} = \{E, \underbrace{\hat{n}_x, \hat{n}_y, \hat{n}_z}_{\text{Direction}}\}$$

Energy

(suppressing "unsafe" charge/flavor information)

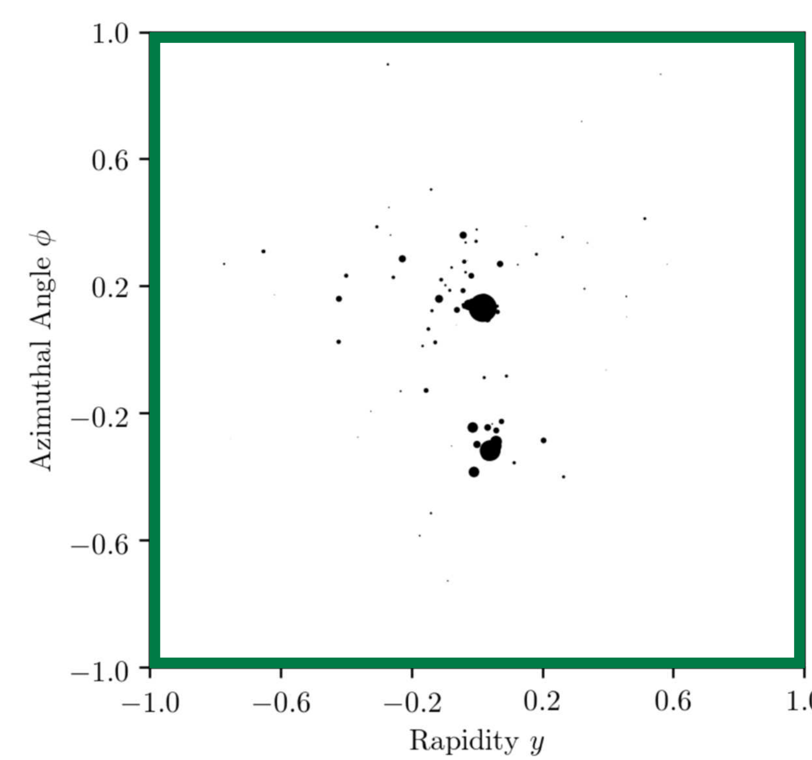


- Equivalently: Energy Density

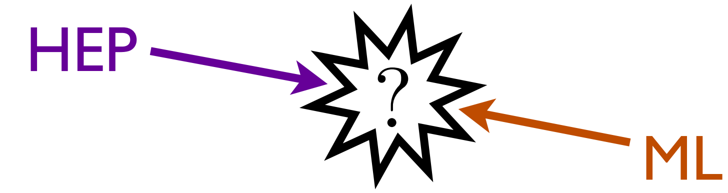
$$\rho(\hat{n}) = \sum_{i \in \mathcal{I}} E_i \, \delta^{(2)}(\hat{n} - \hat{n}_i)$$

Energy

Direction



“Thinking” Like a Physicist



Symmetry: $\mathcal{J} = \{\underbrace{\vec{p}_1, \vec{p}_2, \vec{p}_3, \ldots, \vec{p}_N}_{\text{Unordered, Variable Length Set (QM!)}}\}$

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

Energy weighting (QFT!)

Properly specified problem (i.e. loss function) ⇨

Many (labeled) examples $\{\mathcal{J}\}$ ⇨

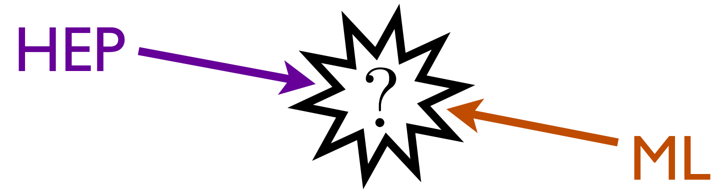
The Machine

⇨ **Solution** $S(\mathcal{J})$

⇨ **Verification** $V(\mathcal{J})$

Check that answer is physically sensible

Energy Flow Networks



Architecture designed around symmetries and interpretability

Permutation invariant

Linear weights (i.e. safe)

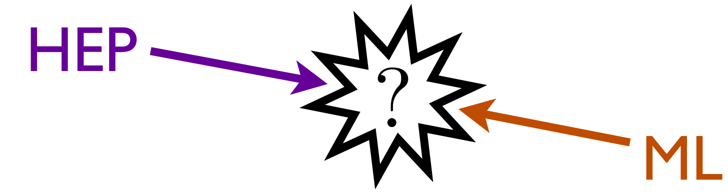
$$S(\mathcal{J}) = F(V_1, V_2, \ldots, V_\ell) \qquad V_a(\mathcal{J}) = \sum_{i \in \mathcal{J}} E_i \, \Phi_a(\hat{n}_i)$$

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, arXiv 2020]

Energy Flow Networks

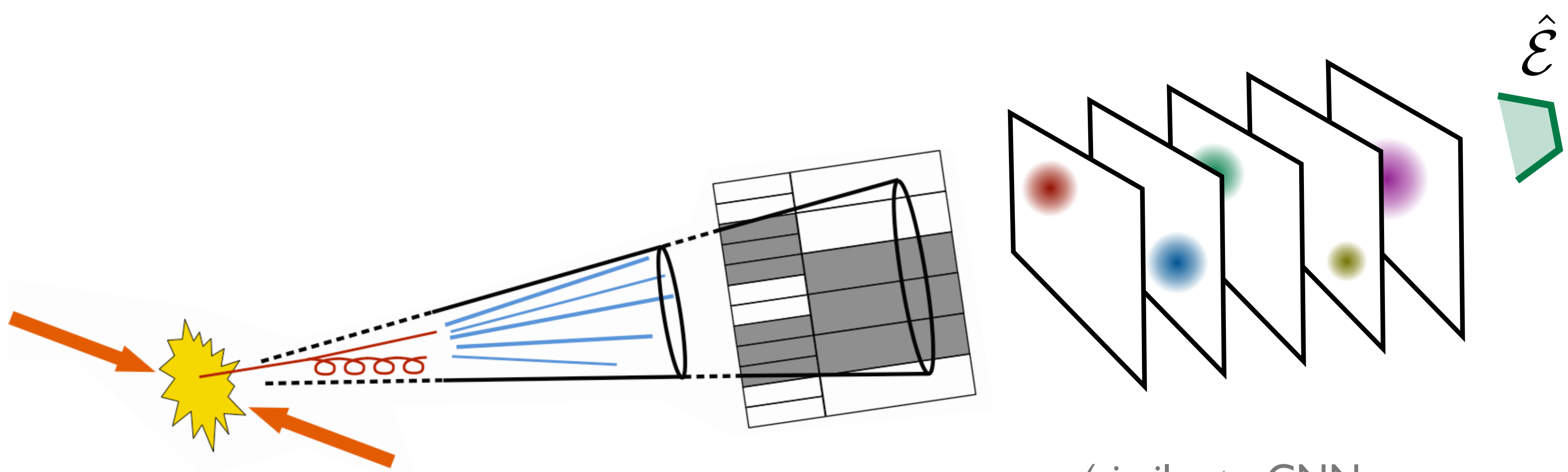


Architecture designed around symmetries and ***interpretability***

Latent space of dim ℓ

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

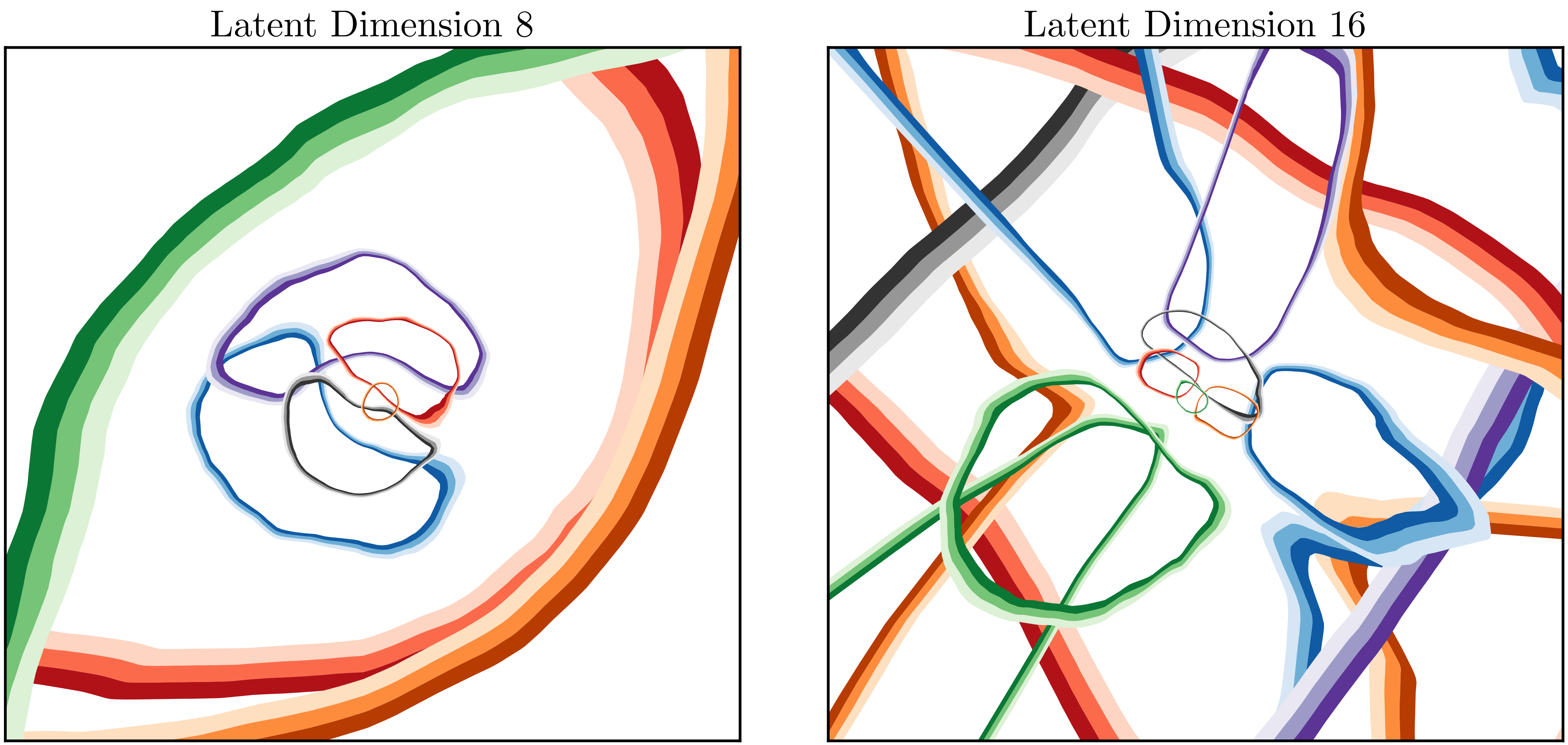
Easy to plot!



(similar to CNN filter activation)

[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, arXiv 2020]

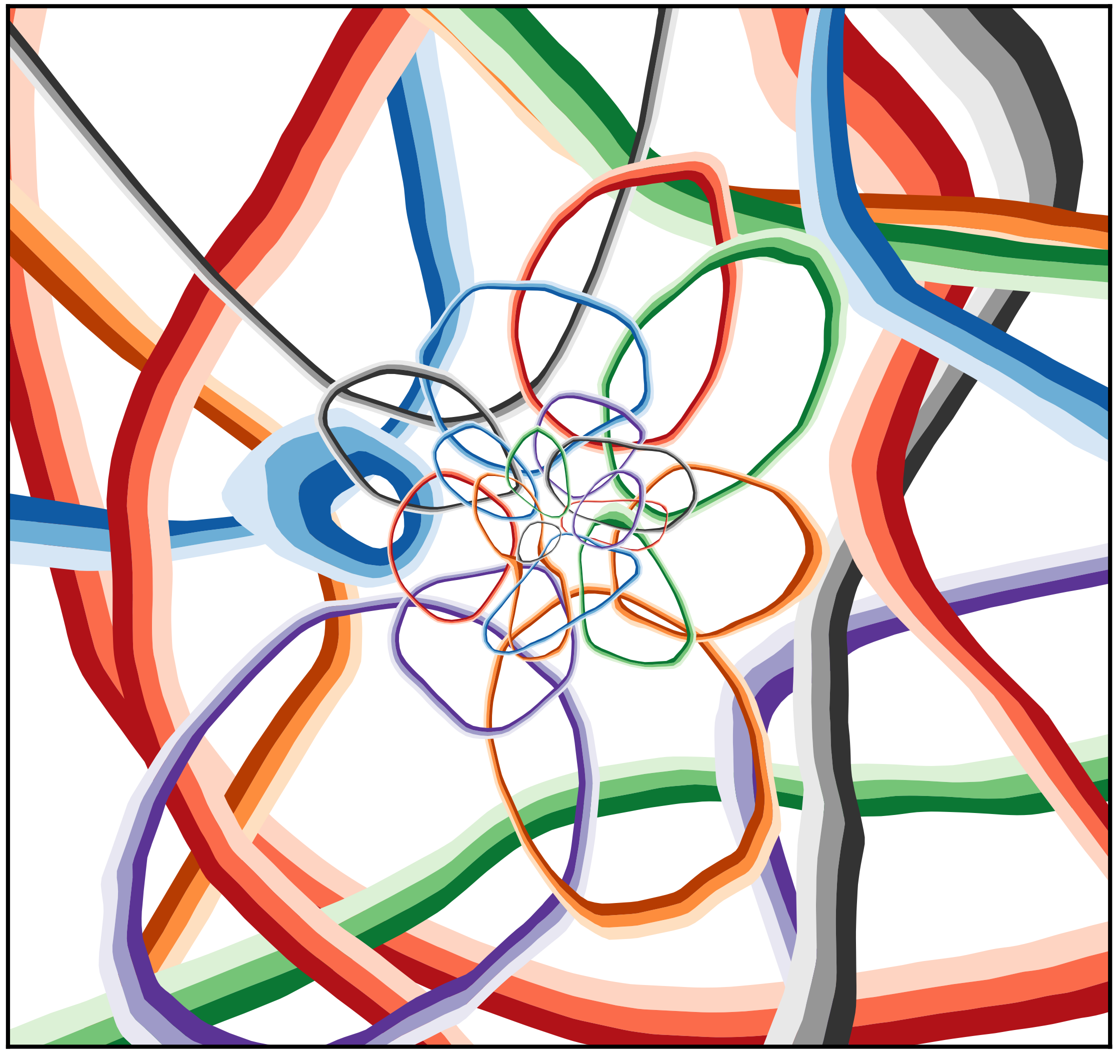

Psychedelic Network Visualization



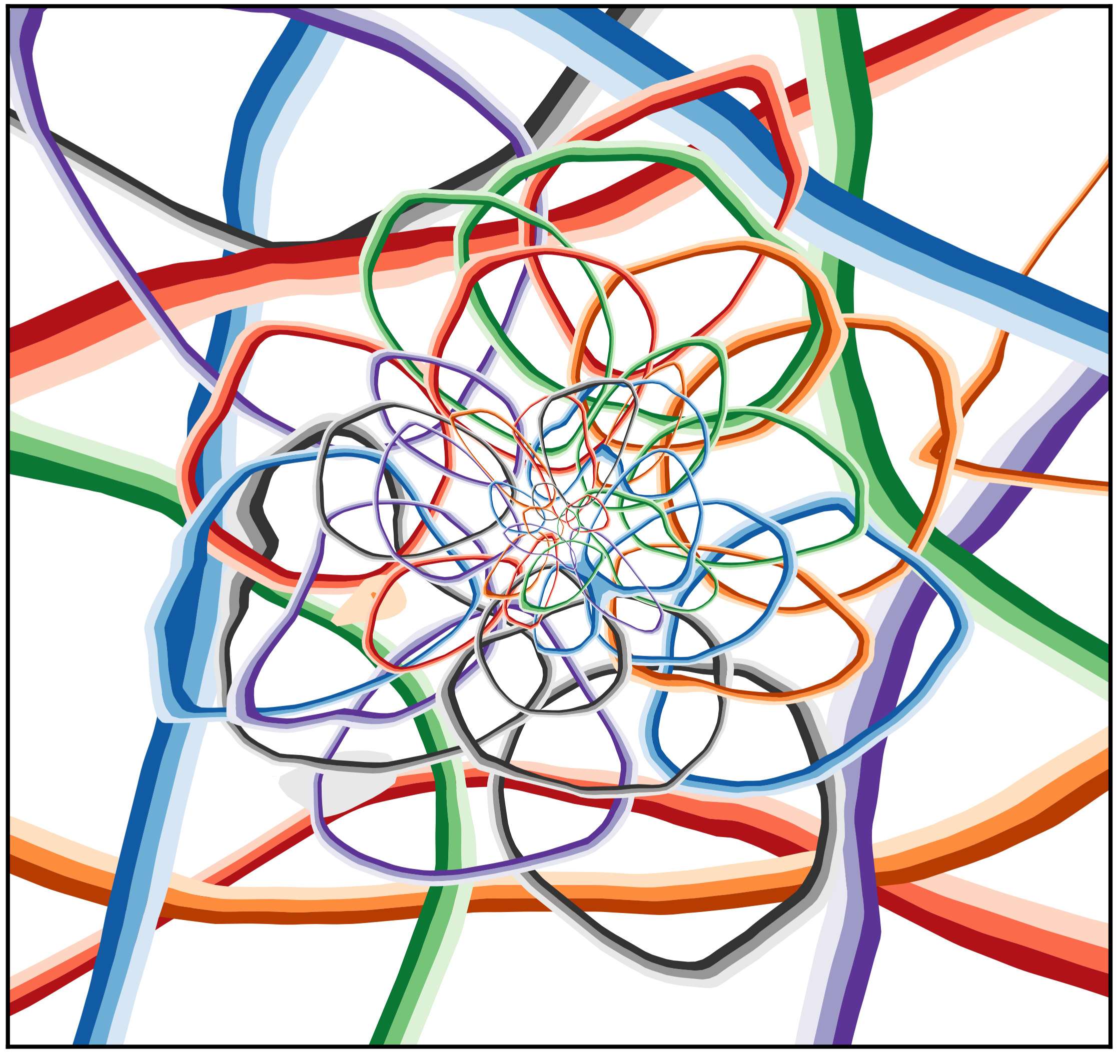
For the case of quark vs. gluon classification

Psychedelic Network Visualization

Latent Dimension 32

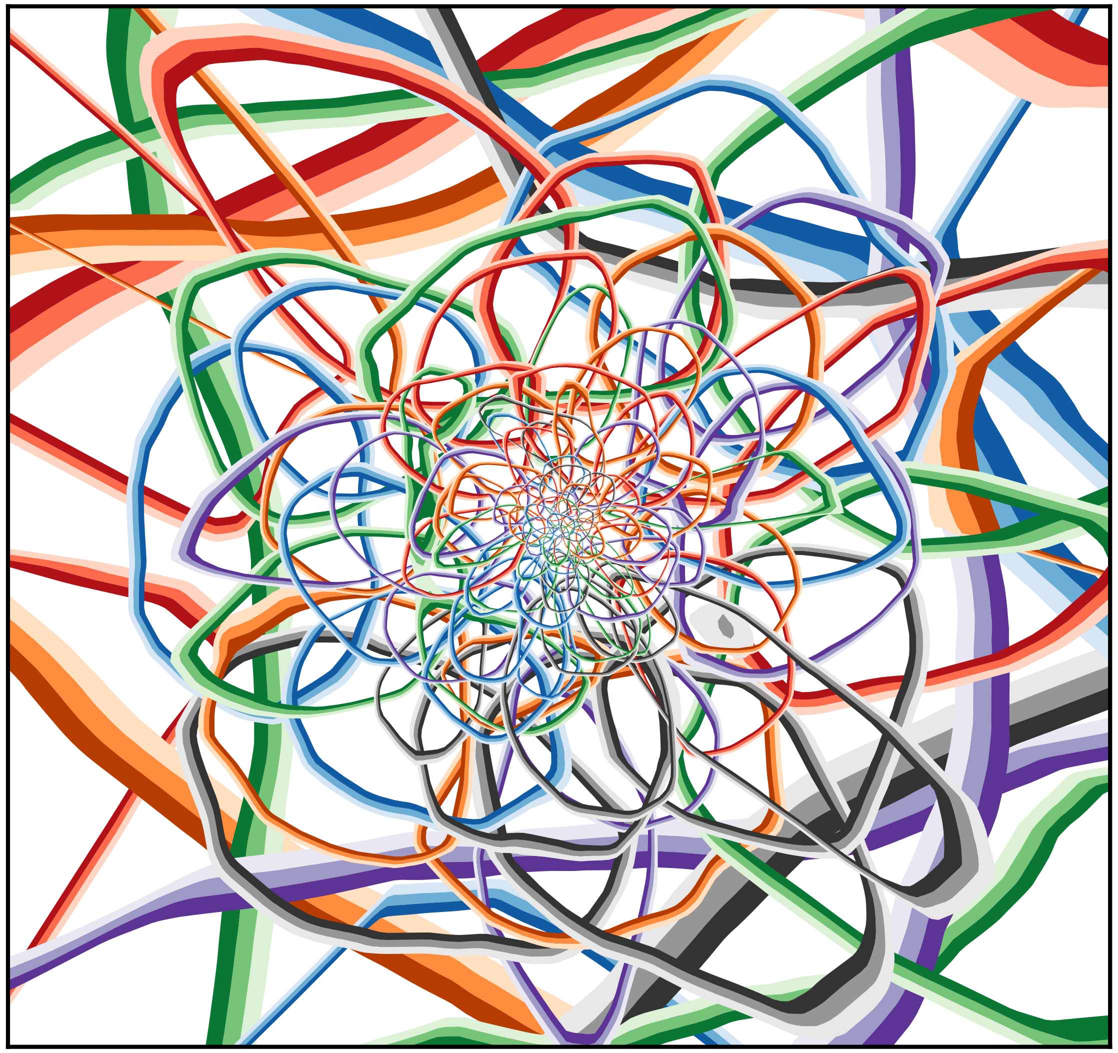


Latent Dimension 64

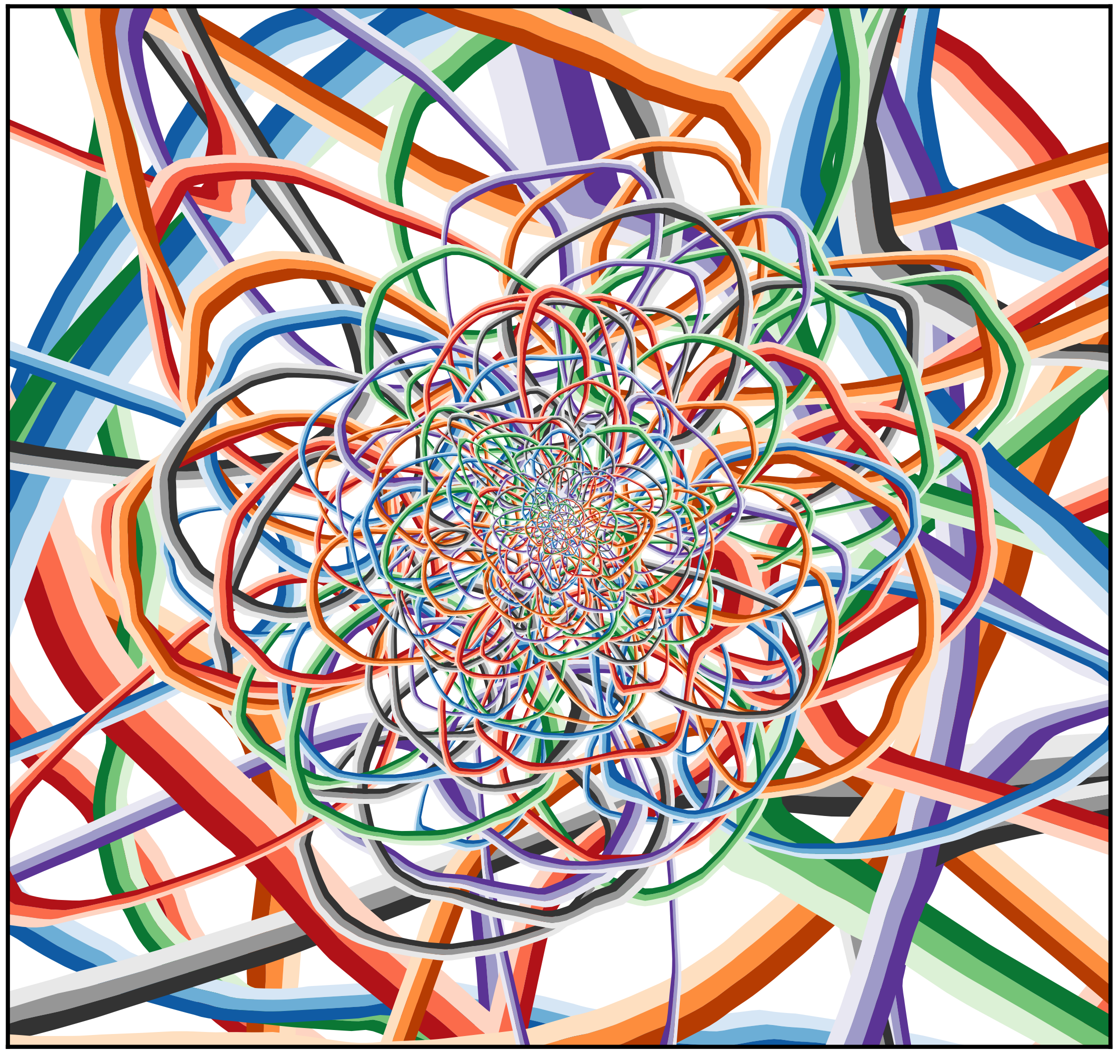


Psychedelic Network Visualization

Latent Dimension 128

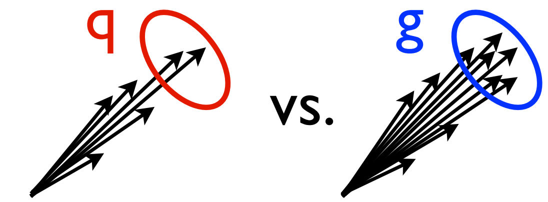


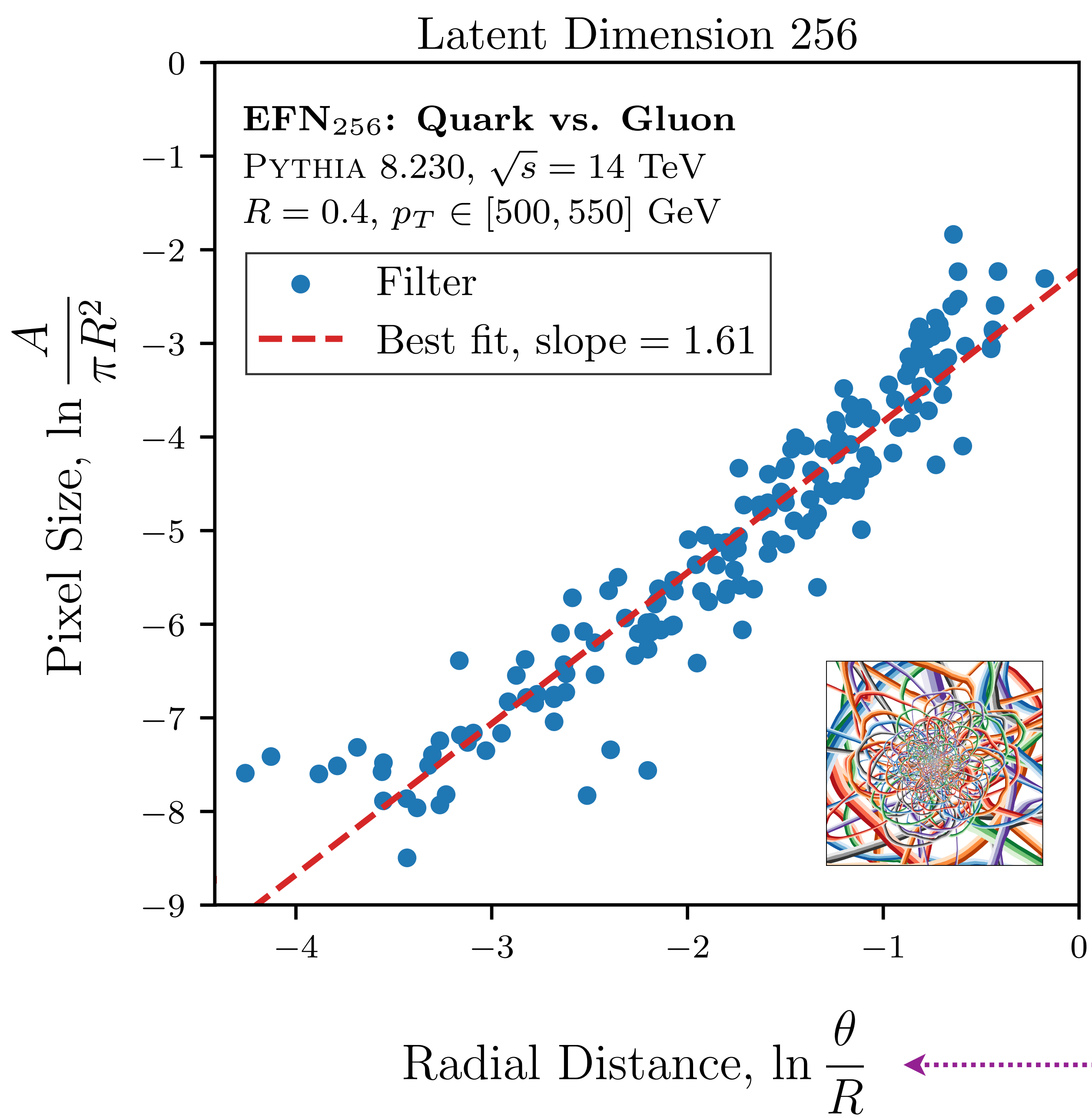
Latent Dimension 256

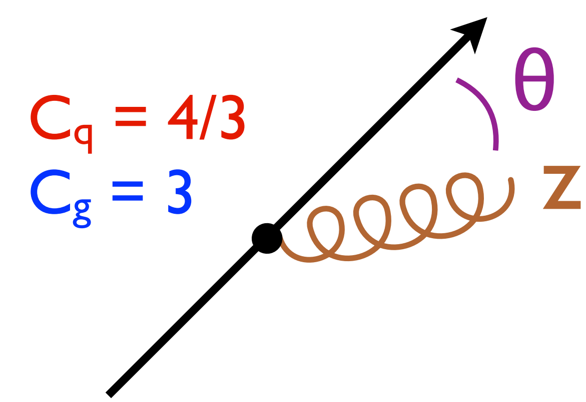


Singularity structure of QCD!

Machine Learning Collinear QCD







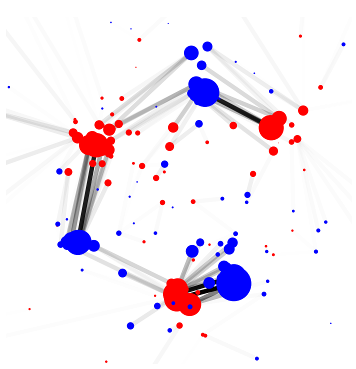
$$dP_{i \to ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$

Collinear Soft

[Komiske, Metodiev, JDT, JHEP 2019]

Rise of the Machines?

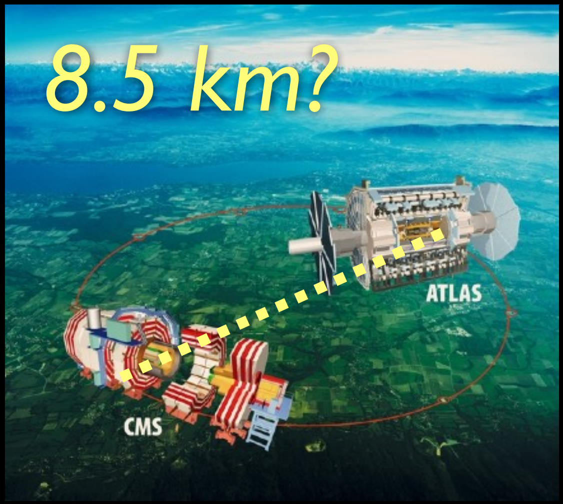
What is a Collider Event?



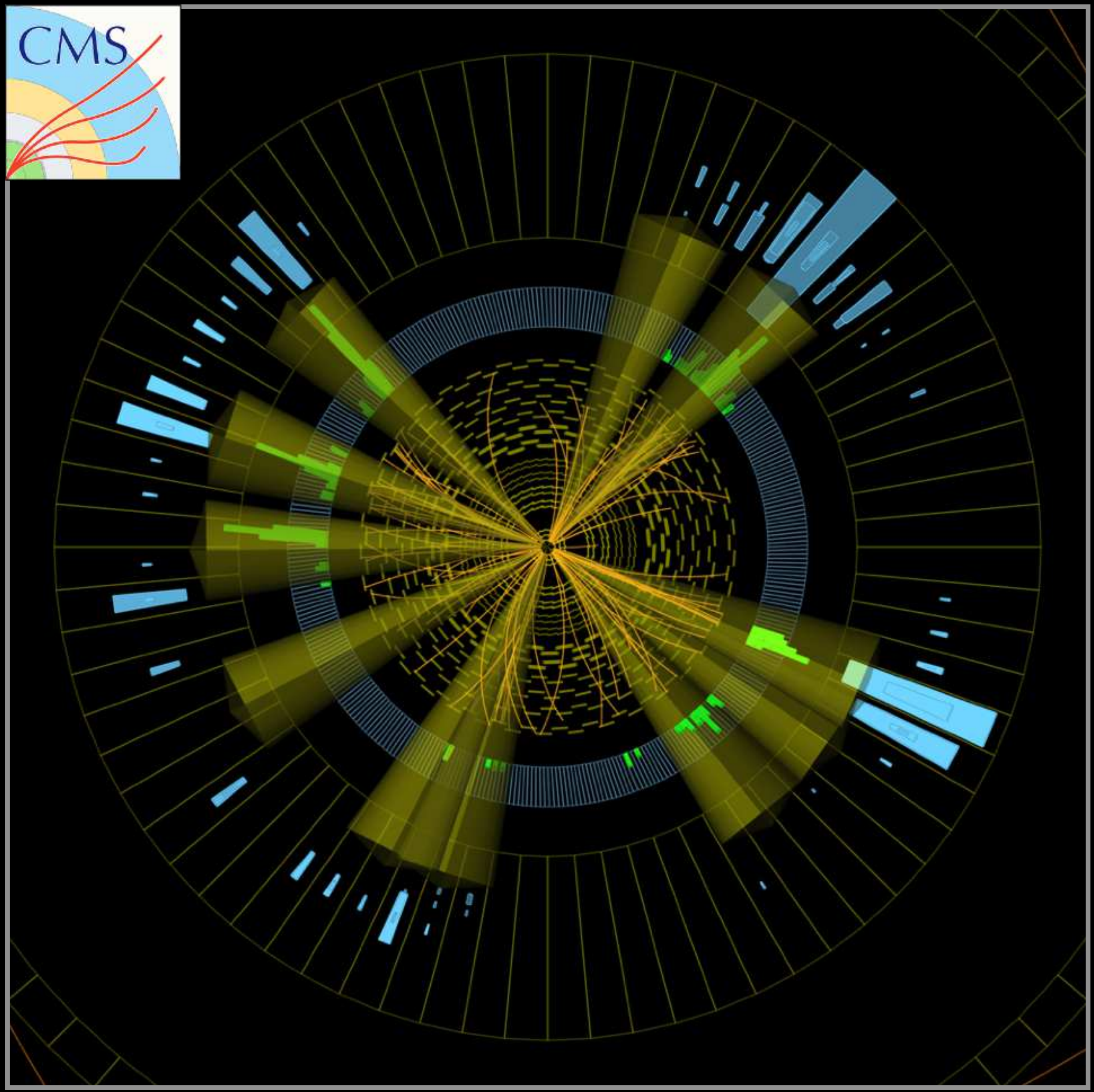
When are Collider Events Similar?

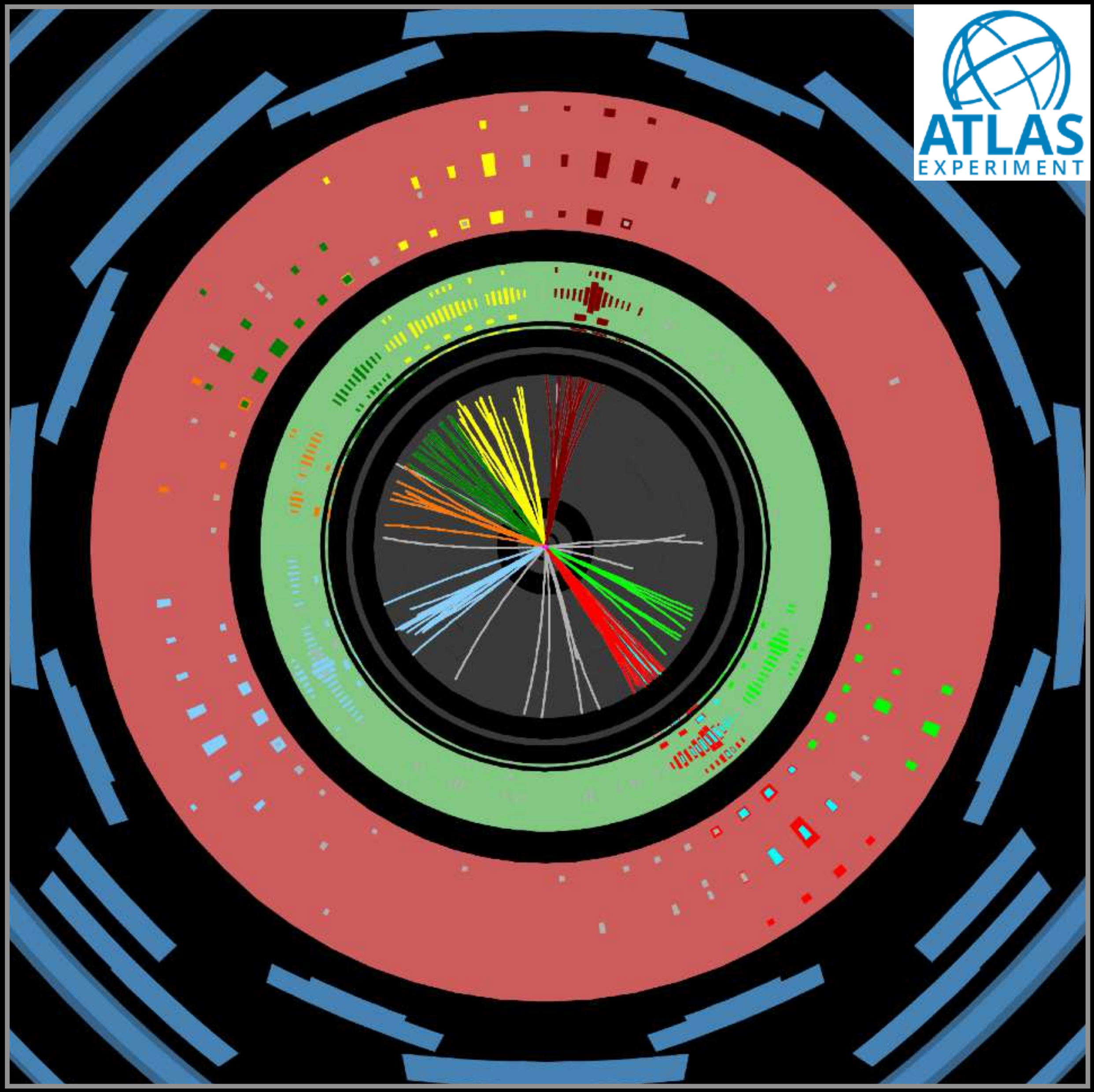
Two Collider Events

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, ICJV 2000;
Pele, Werman, ECCV 2008; Pele Taskar, GSI 2013]

Minimum "work" (stuff x distance) to make one distribution look like another distribution



Déblai

Remblai

[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Vaseršteĭn, 1969; Wikipedia]

The Earth Mover's Distance

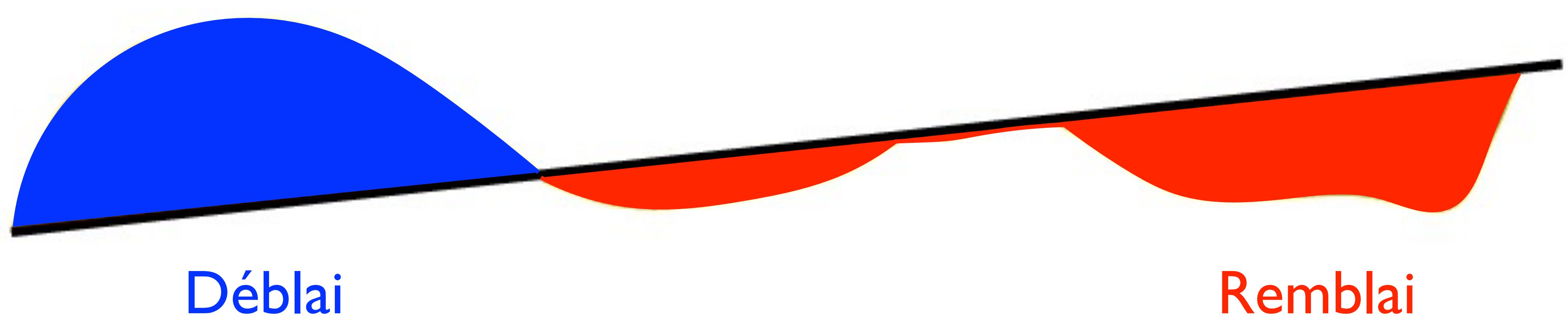
Optimal Transport: Minimum "work" (stuff x distance) to make one distribution look like another distribution

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

Additive Weight:

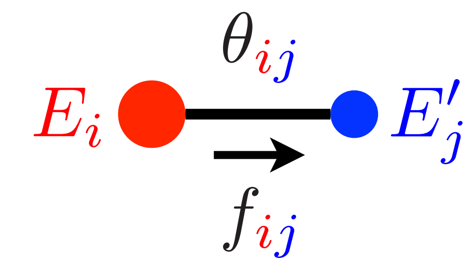
Ground Distance:

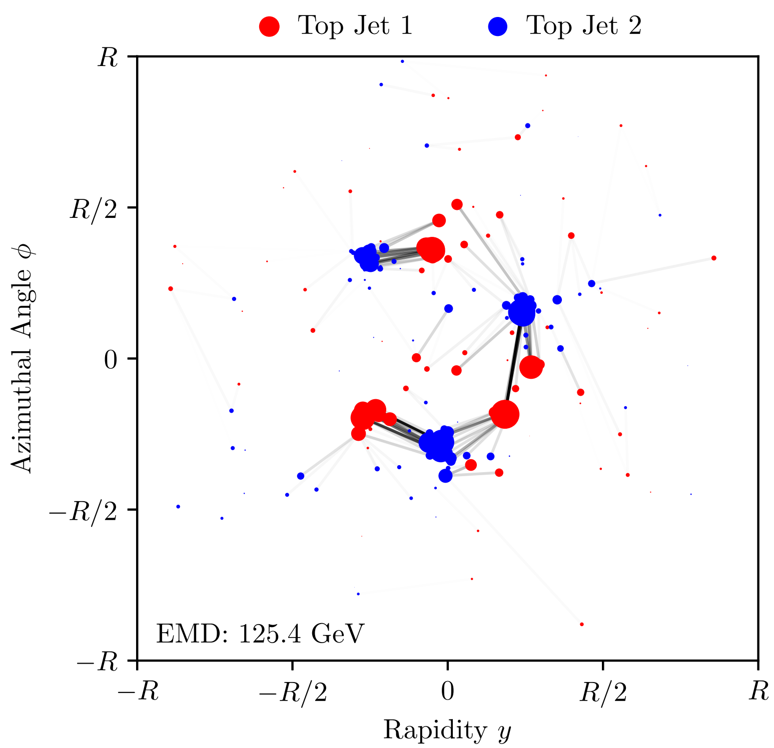
⇒ Distance Between Distributions



[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Vaseršteĭn, 1969; Wikipedia]

The Energy Mover's Distance

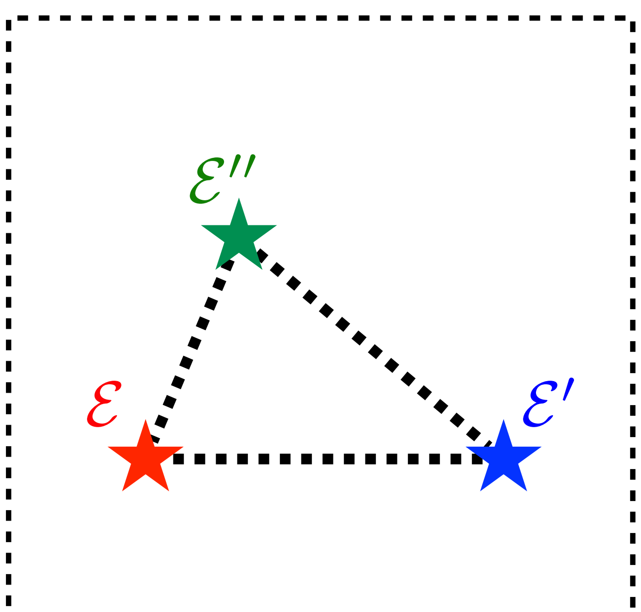




Optimal transport between energy flows…

$$\mathrm{EMD}(\mathcal{E}, \mathcal{E}') = \underbrace{\min_{\{f\}} \sum_i \sum_j f_{ij} \frac{\theta_{ij}}{R}}_{\text{Cost to move energy}} + \underbrace{\left| \sum_i E_i - \sum_j E'_j \right|}_{\text{Cost to create energy}}$$

in GeV



…defines a metric on the space of events

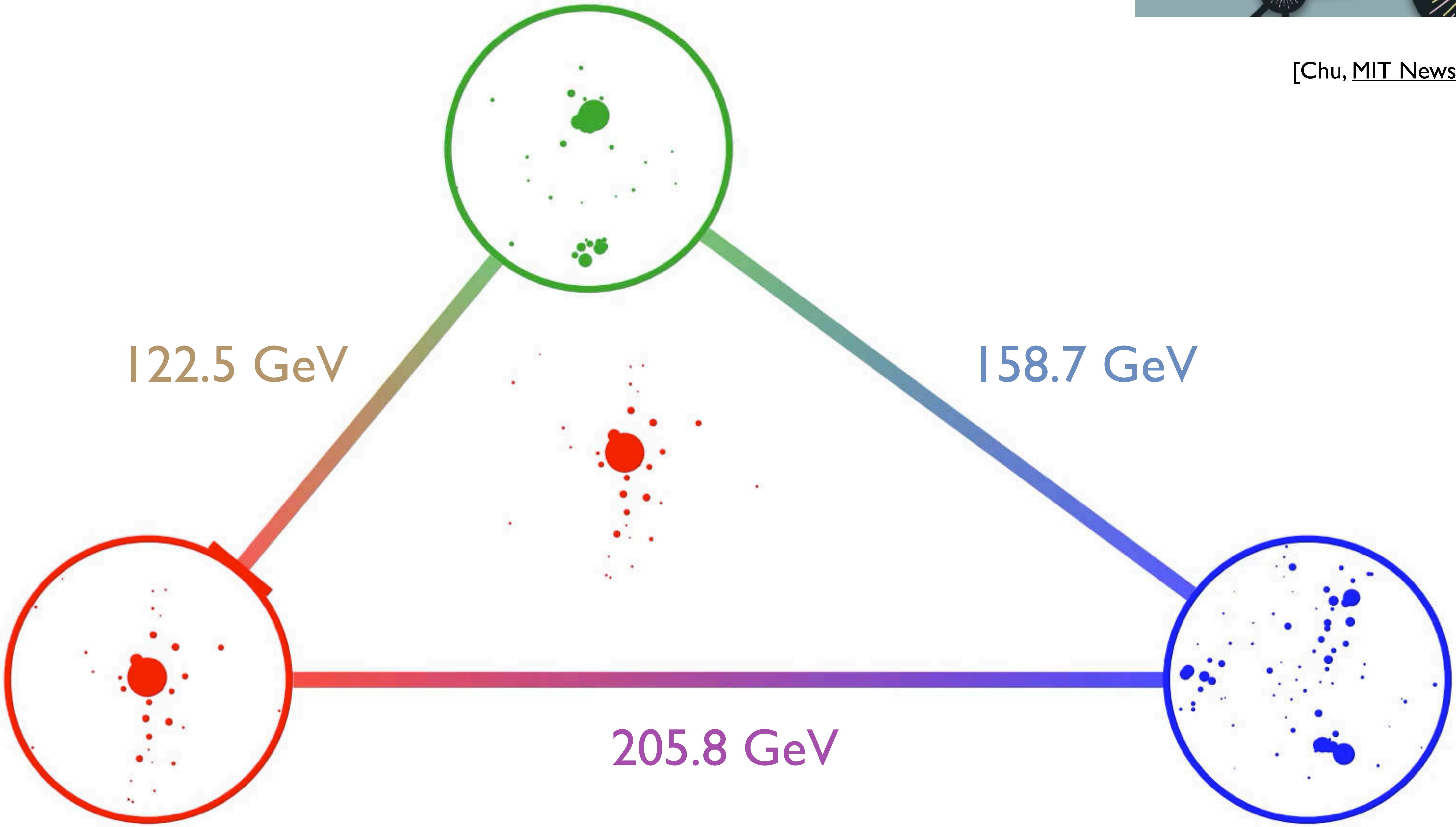
$$0 \leq \mathrm{EMD}(\mathcal{E}, \mathcal{E}') \leq \mathrm{EMD}(\mathcal{E}, \mathcal{E}'') + \mathrm{EMD}(\mathcal{E}', \mathcal{E}'')$$

(assuming R ≥ $\theta_{max}/2$, i.e. R ≥ 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]
[see computational speed up in Cai, Cheng, Craig, Craig, arXiv 2020]

Triangulating the Space of Jets

[Chu, MIT News July 2019]



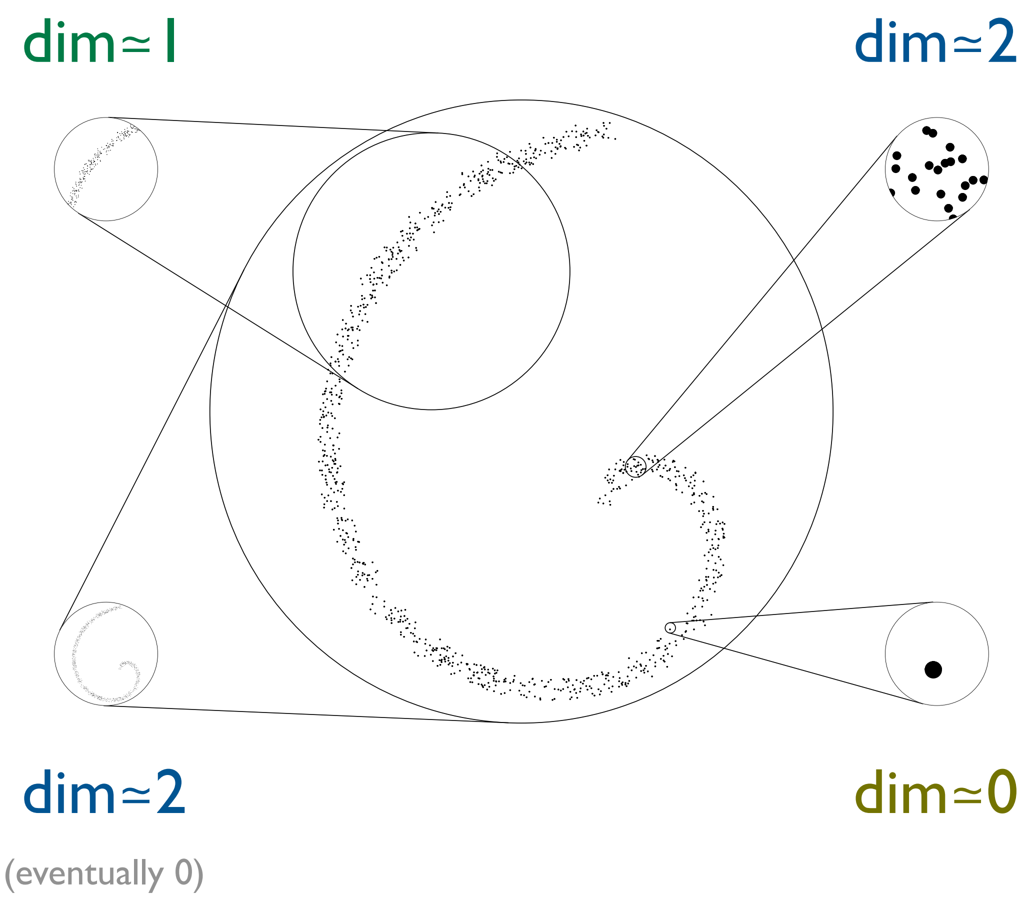
[Komiske, Metodiev, JDT, PRL 2019; code at Komiske, Metodiev, JDT, energyflow.network;
see alternative graph network approach in Mullin, Pacey, Parker, White, Williams, arXiv 2019]

Dimensionality of Space of Jets

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

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

[Grassberger, Procaccia, PRL 1983; Kégl, NIPS 2002]

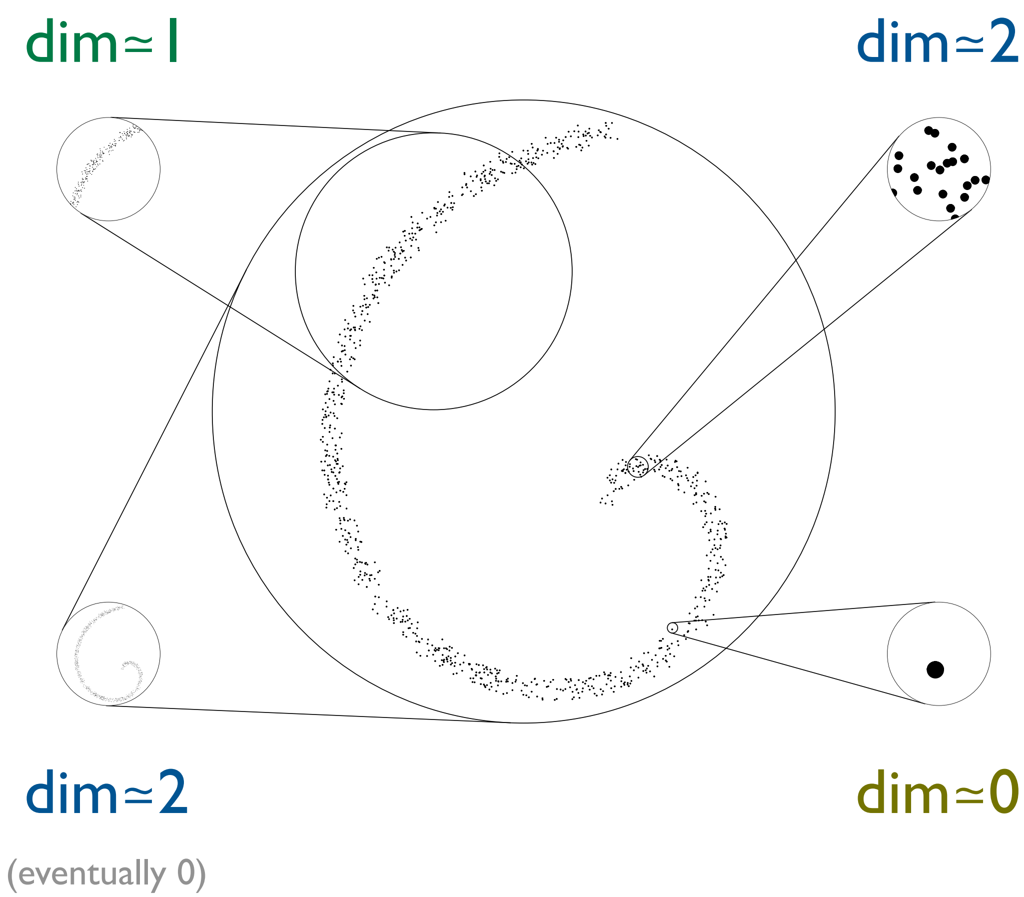


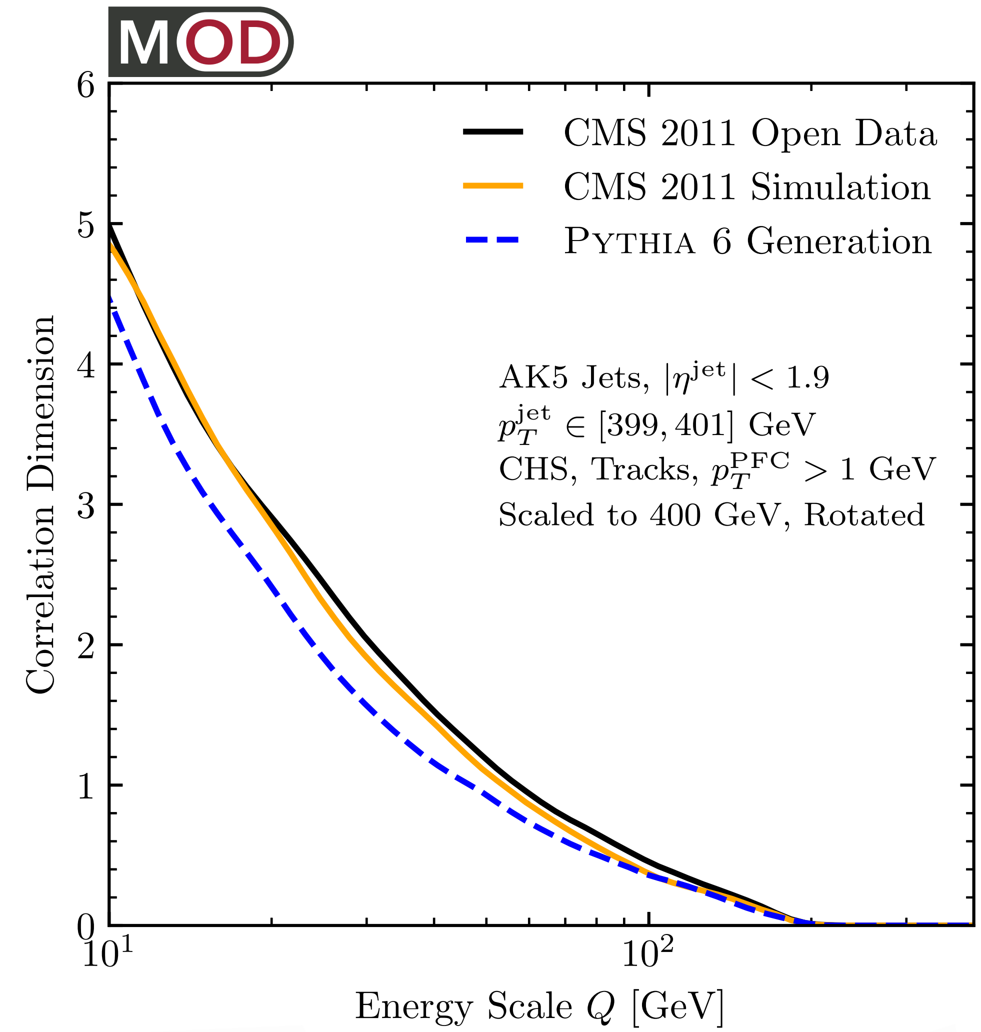
Dimensionality of Space of Jets

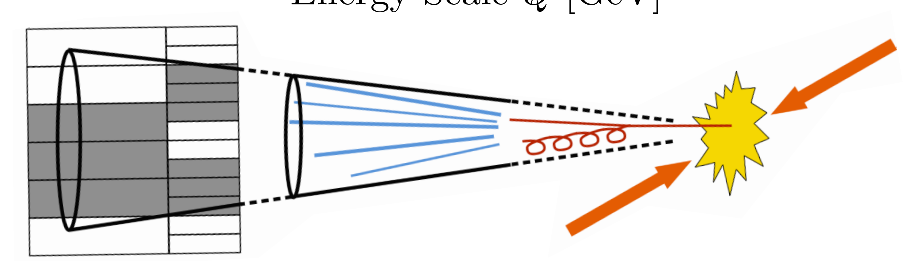
$$N_{\text{neighbors}}(r) \sim r^{\text{dim}}$$

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

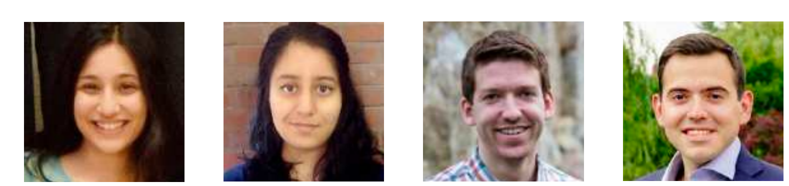
[Grassberger, Procaccia, PRL 1983; Kégl, NIPS 2002]



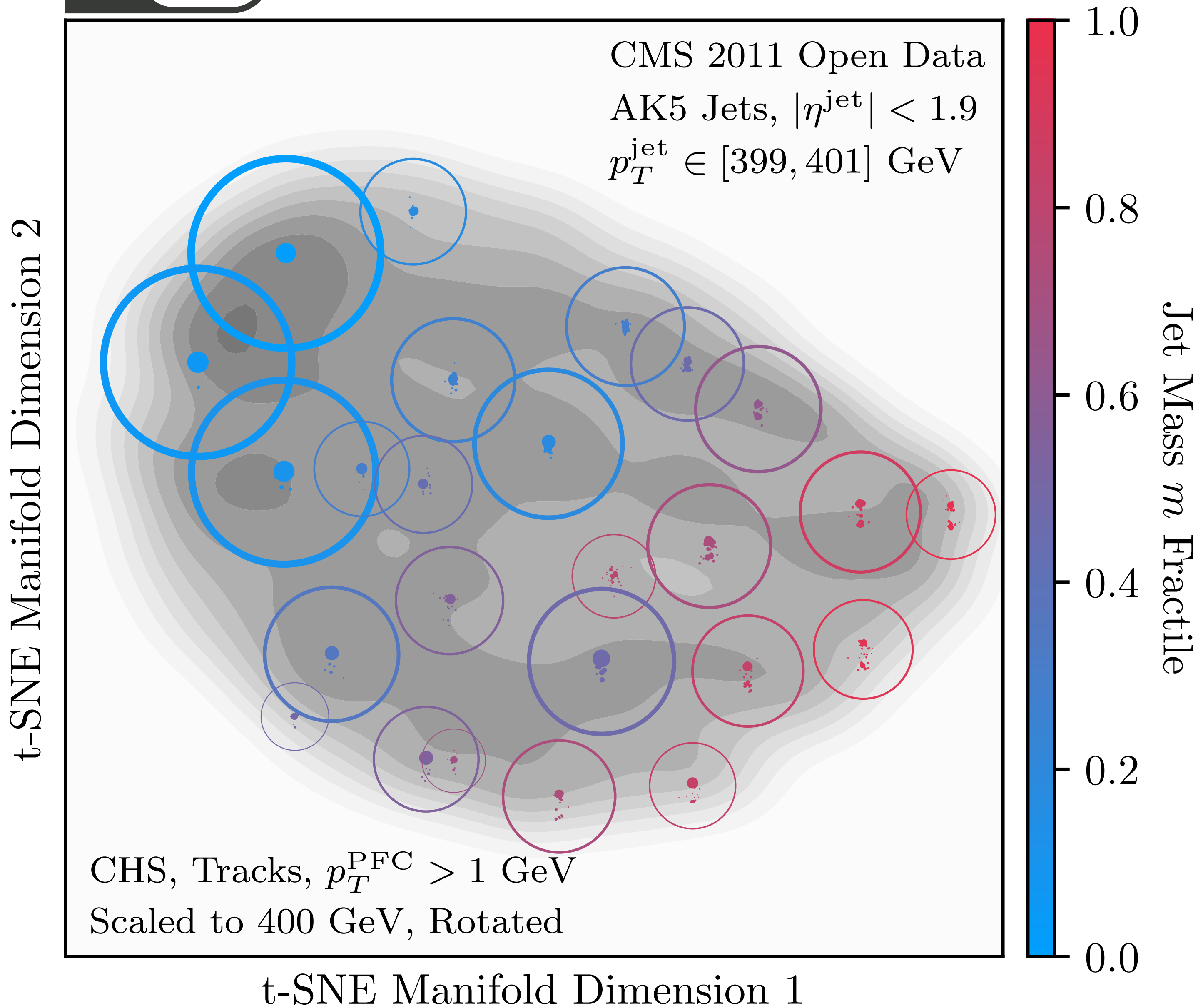




[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020;
using CMS Open Data]

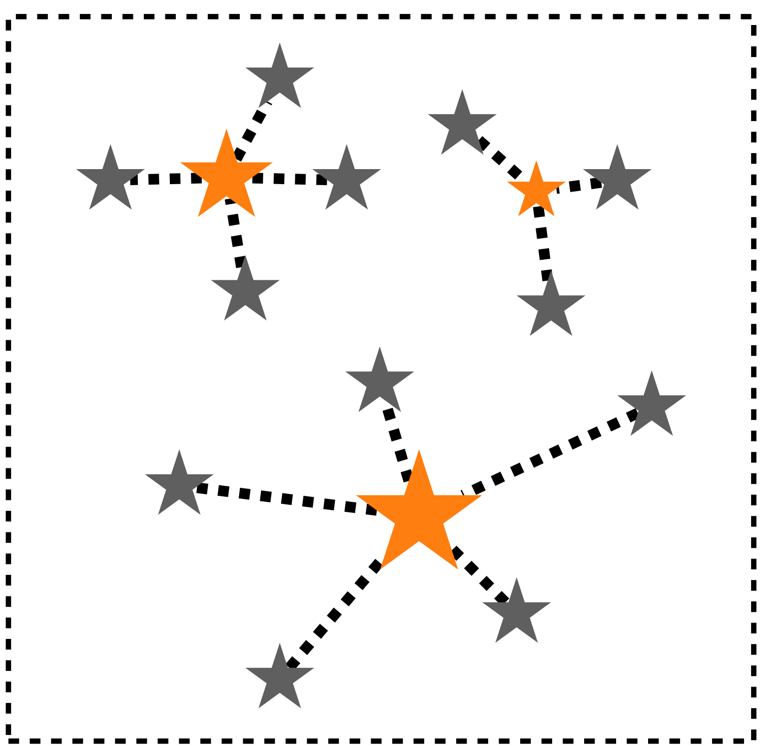


Most Representative Jets



k-medoids

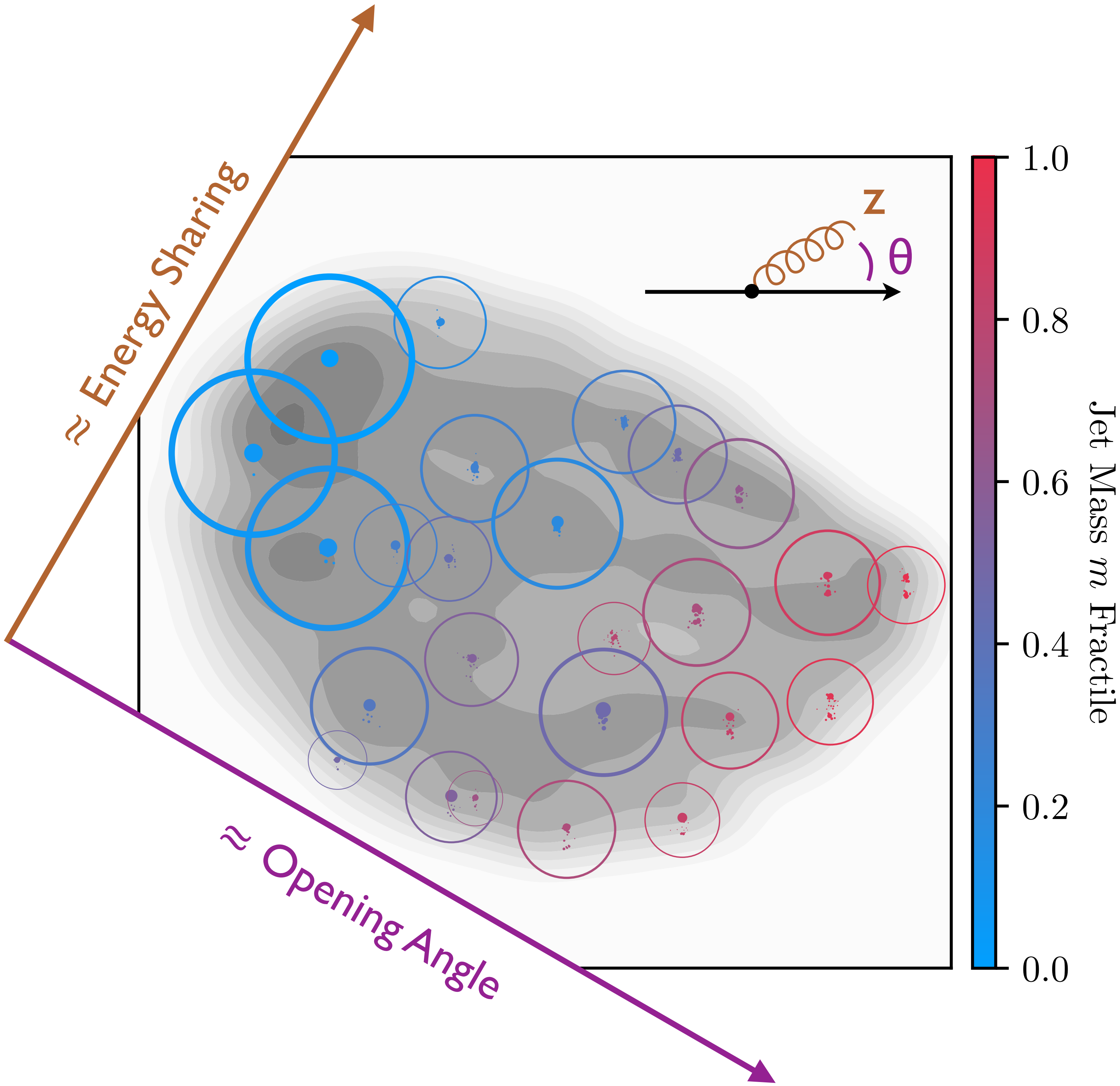
Arranged via t-SNE



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

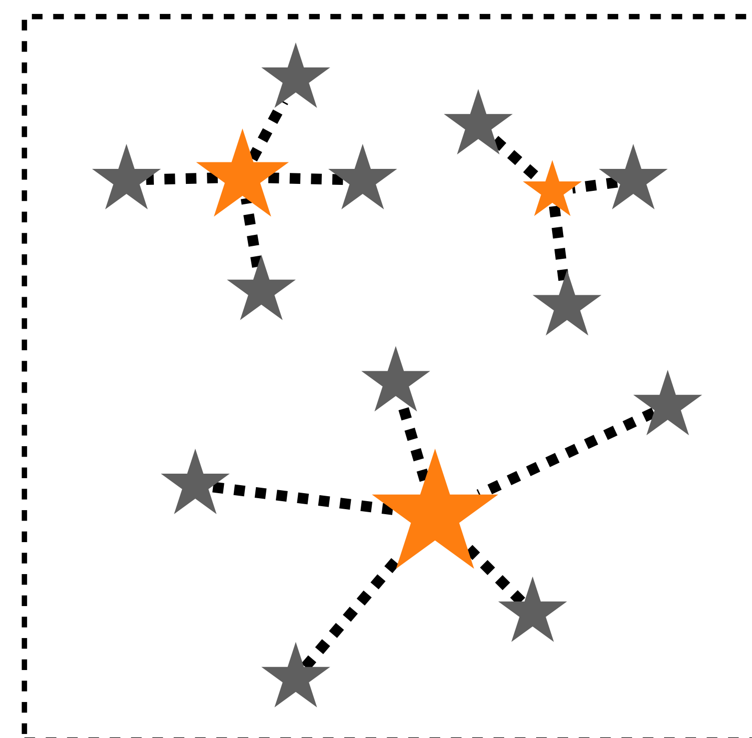
   

Most Representative Jets



k-medoids

Arranged via t-SNE

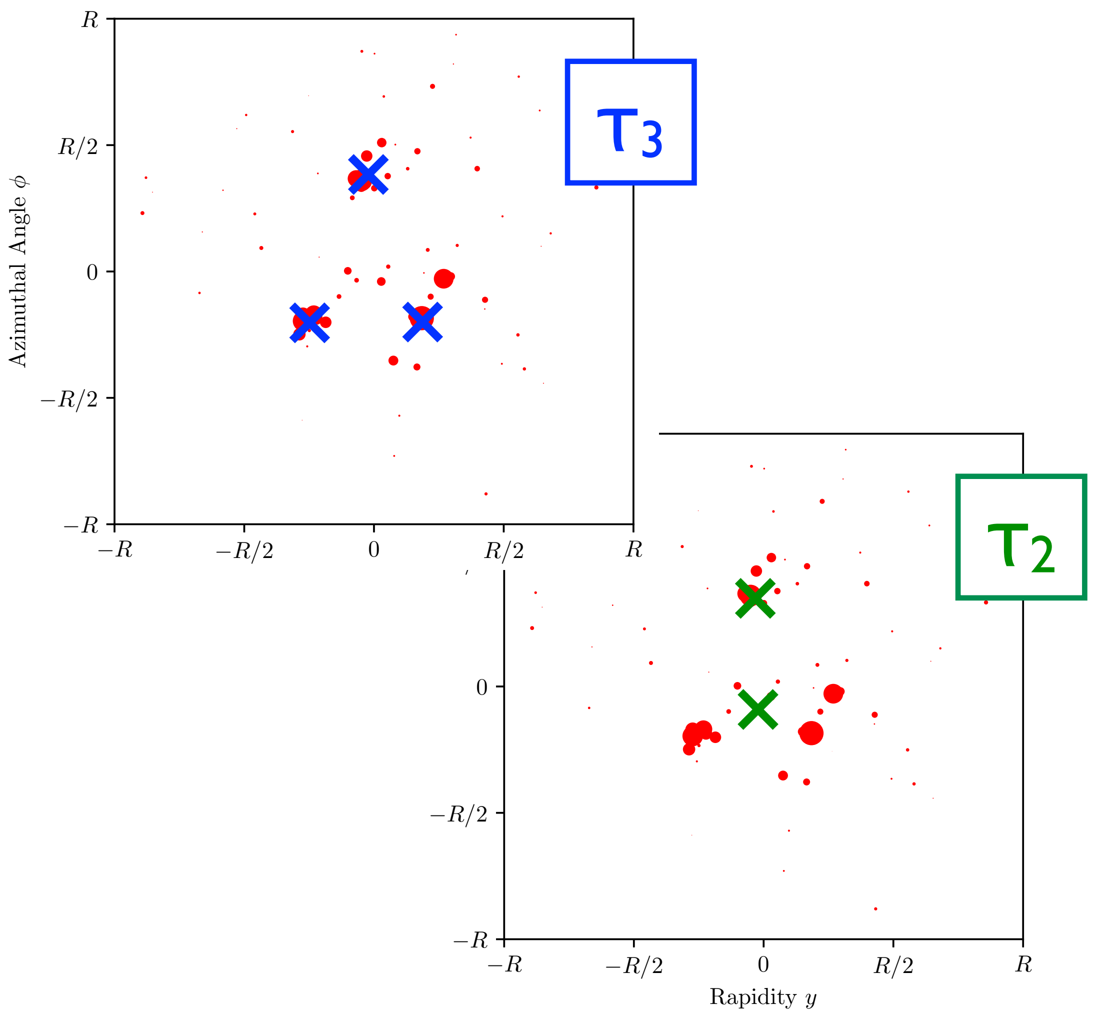


[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020;
using van der Maaten, Hinton, JMLR 2008]

N-subjettiness

Ubiquitous jet substructure observable used for almost a decade…

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

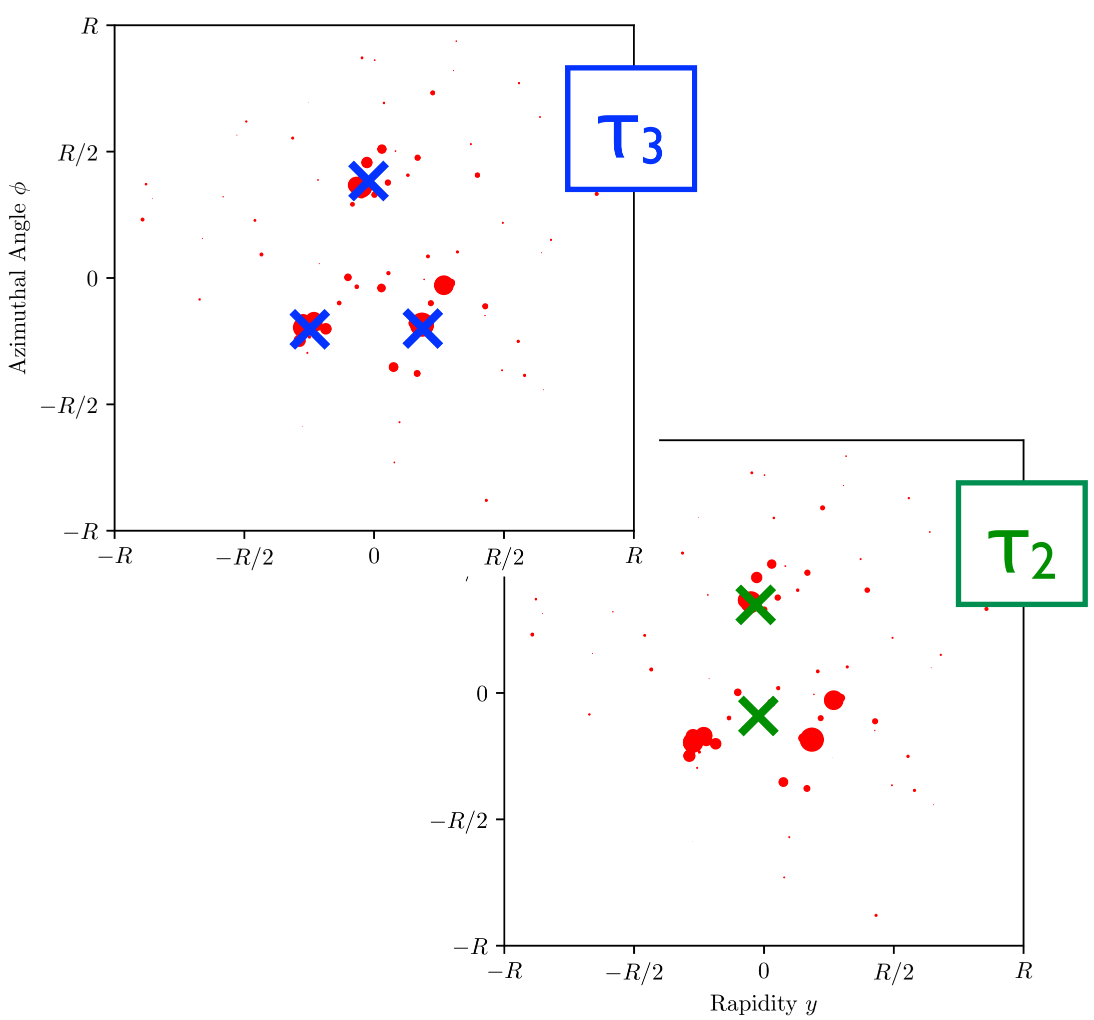


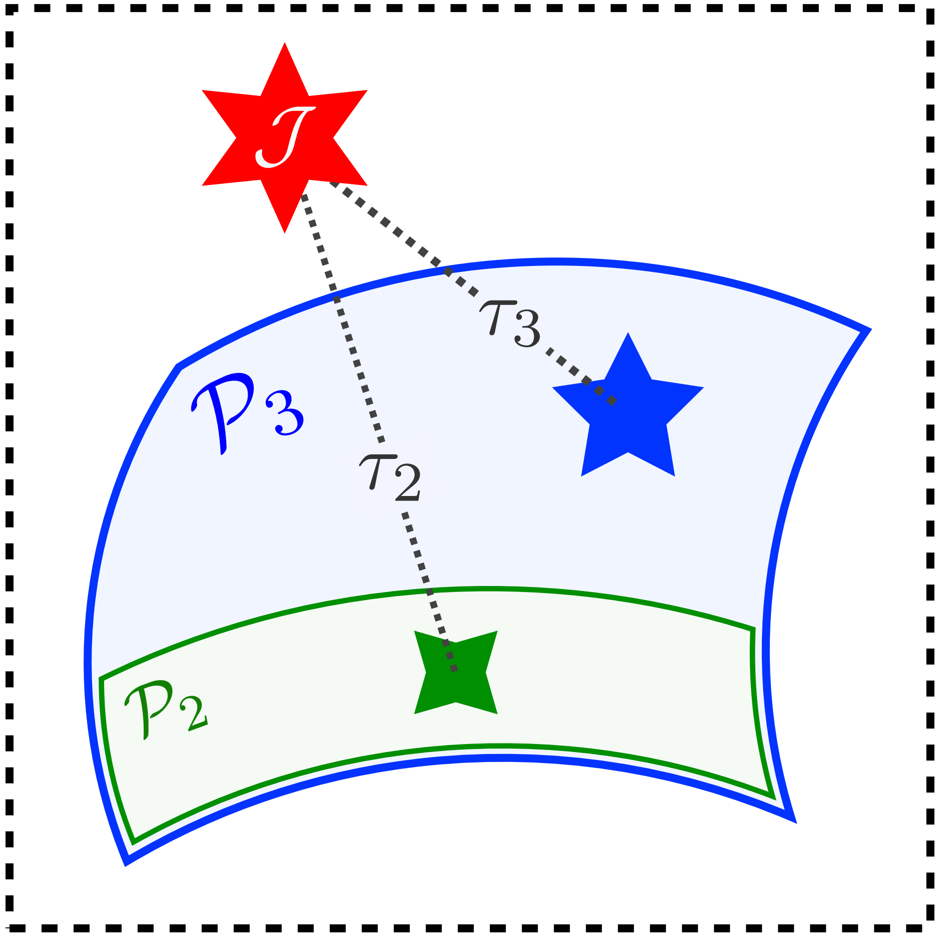
[JDT, Van Tilburg, JHEP 2011, JHEP 2012;
based on Brandt, Dahmen, ZPC 1979; Stewart, Tackmann, Waalewijn, PRL 2010]

N-subjettiness = Point to Manifold EMD

…is secretly an optimal transport problem

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

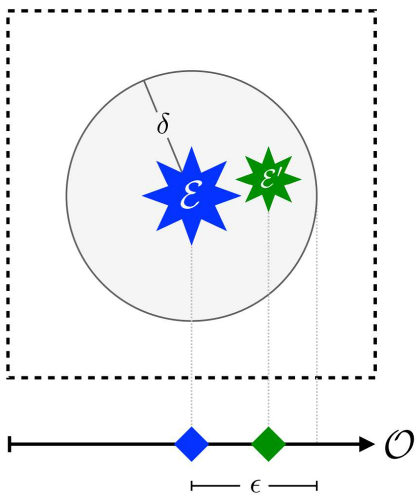




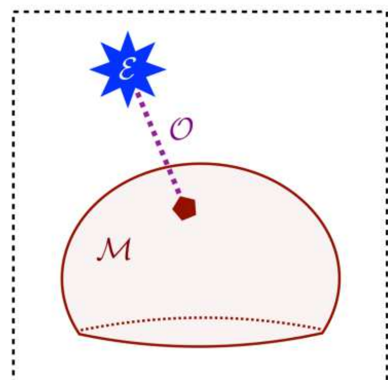
[JDT, Van Tilburg, JHEP 2011, JHEP 2012;
rephrased via Komiske, Metodiev, JDT, JHEP 2020; see opposite limit in Cesarotti, JDT, JHEP 2020]

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

IRC Safety is smoothness in the space of events

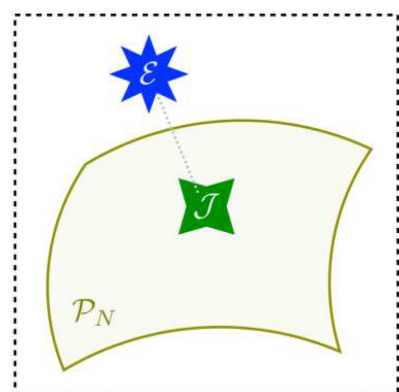


Event shapes are distances from events to manifolds.



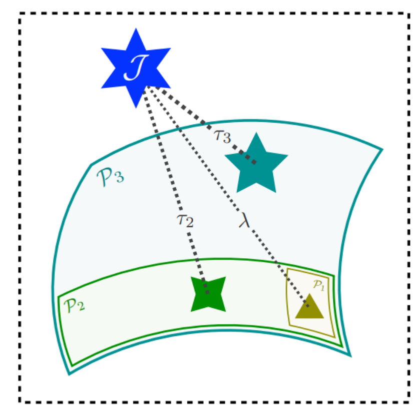
$\mathcal{O}(\mathcal{E}) = \min_{\mathcal{E}'\in\mathcal{M}} \mathrm{EMD}_{\beta,R}(\mathcal{E},\mathcal{E}')$

Jets are projections to few-particle manifolds.



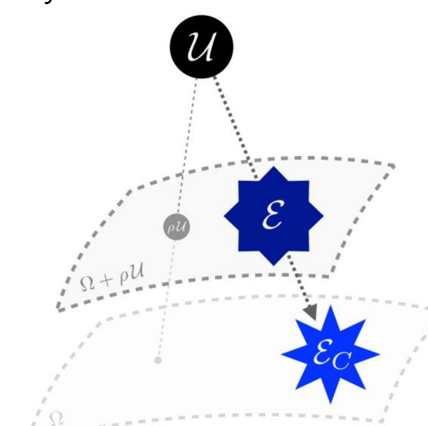
$\mathcal{J} = \operatorname{argmin}_{\mathcal{E}'\in\mathcal{P}_N} \mathrm{EMD}_{\beta,R}(\mathcal{E},\mathcal{E}')$

Substructure resolves emissions within the jet.



$\tau(\mathcal{J}) = \min_{\mathcal{E}'\in\mathcal{P}_N} \mathrm{EMD}_{\beta}(\mathcal{J},\mathcal{E}')$.

Pileup mitigation moves away from uniform radiation.



$\mathcal{E}_C = \arg\min_{\mathcal{E}'} \mathrm{EMD}(\mathcal{E},\mathcal{E}' + \rho\,\mathcal{U})$.

Taming infinities | Event Shapes | Jet Algorithms | Jet Substructure | Pileup

1960 — 2020

1962-1964
Infrared Safety
[Kinoshita, JMP 1962]
[Lee, Nauenberg, PR 1964]

1977
Thrust, Sphericity
[Farhi, PRL 1977]
[Georgi, Machacek, PRL 1977]

1993
k_T jet clustering
[Ellis, Soper, PRD 1993]
[Catani, Dokshitzer, Seymour, Webber, NPB 1993]

1997-1998
C/A jet clustering
[Wobisch, Wengler, 1998]
[Doskhitzer, Leder, Moretti,Webber, JHEP 1997]

2010-2015
N-(sub)jettiness, XCone
[Stewart, Tackmann, Waalewijn, PRL 2010]
[Thaler, Van Tilburg, JHEP 2011]
[Stewart, Tackmann, Thaler, Vermilion, Wilkason, JHEP 2015]

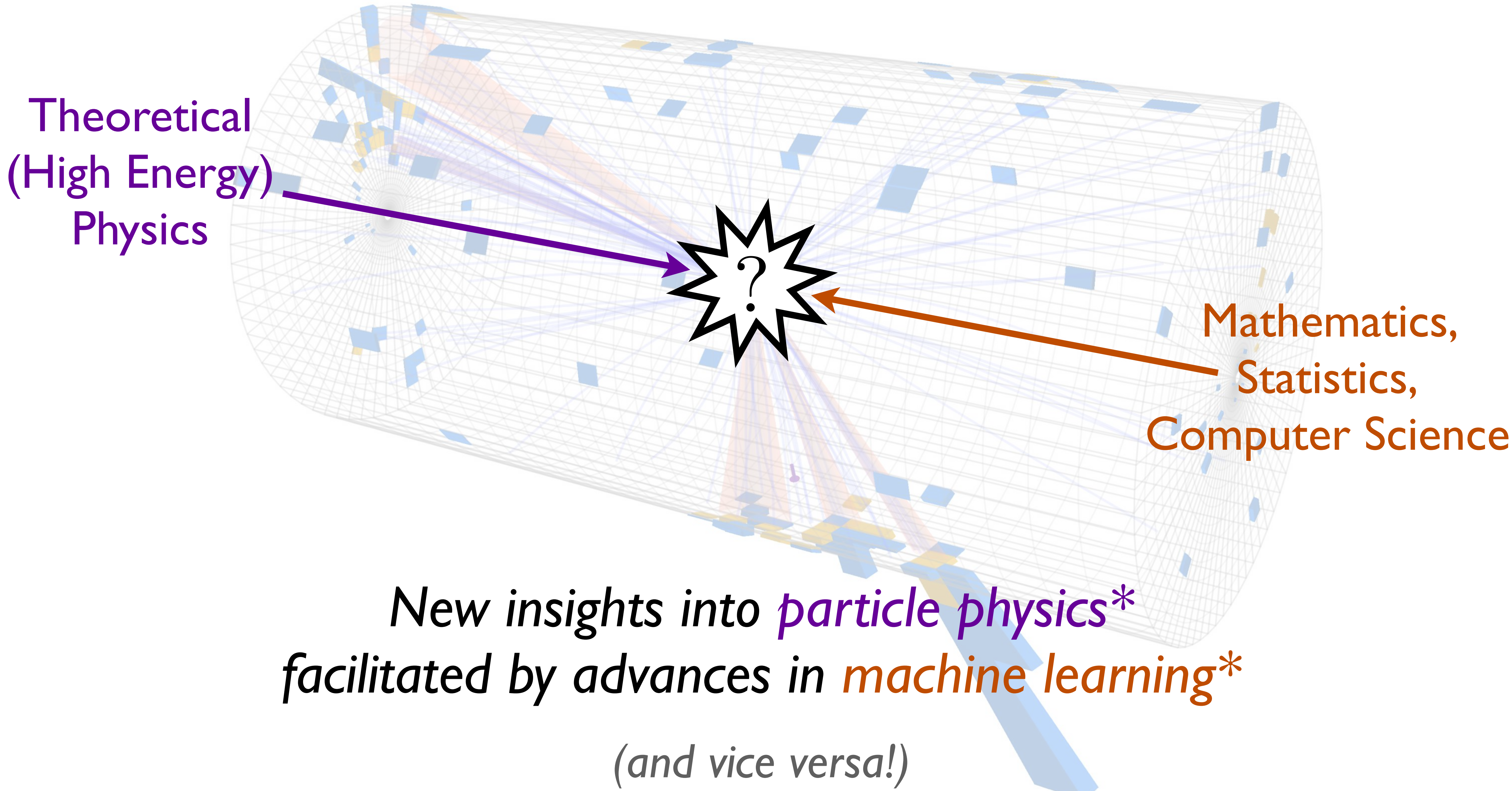
2014-2019
Constituent Subtraction
[Berta, Spousta, Miller, Leitner, JHEP 2014]
[Berta, Masetti, Miller, Spousta, JHEP 2019]

And many more!

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

"Collision Course"

"Theoretical Physics for Machine Learning"
Aspen Center for Physics, January 2019

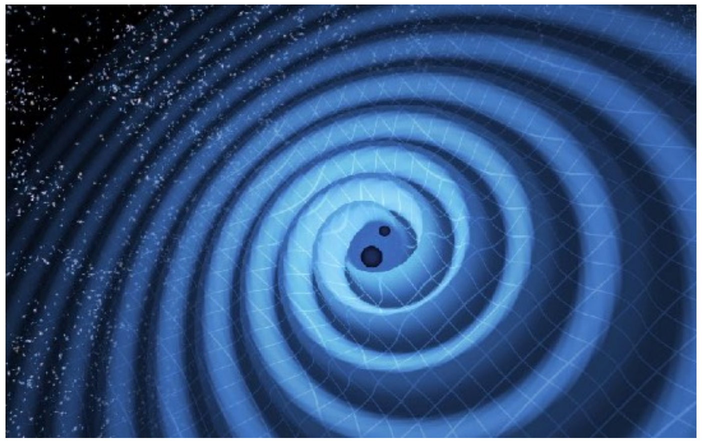


*New insights into particle physics**
*facilitated by advances in machine learning**

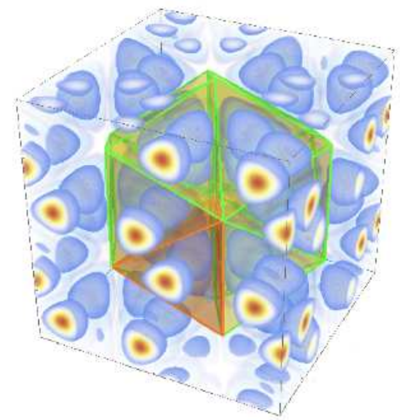
(and vice versa!)

Artificial Intelligence ⇔ Fundamental Interactions

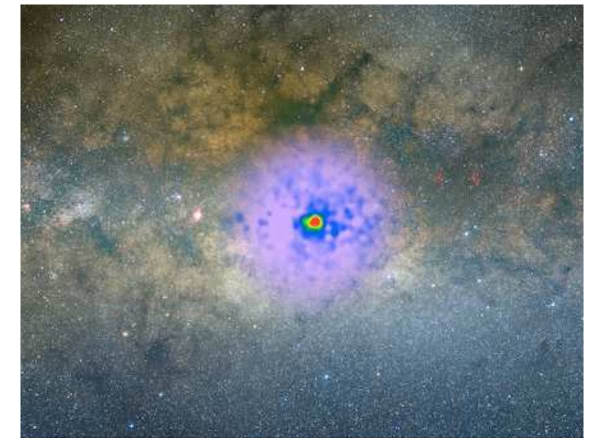
Gravitational Waves



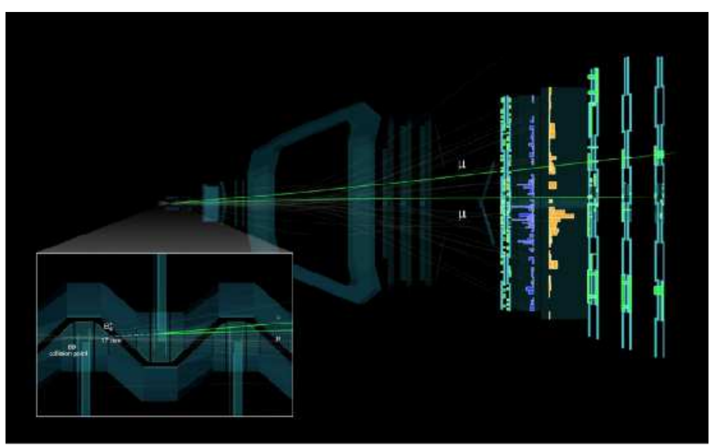
Nuclear Physics



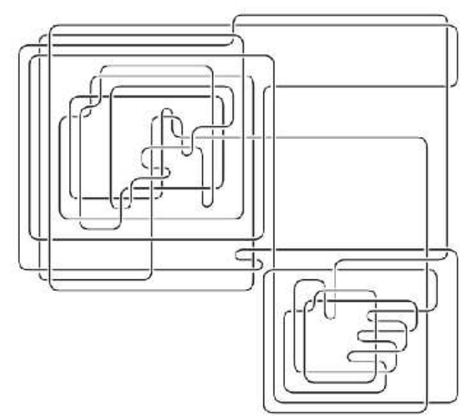
Astrophysics



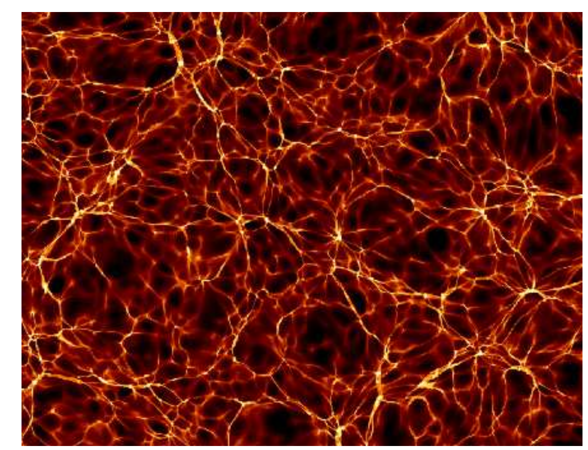
Particle Colliders



Mathematical Physics



Dark Matter

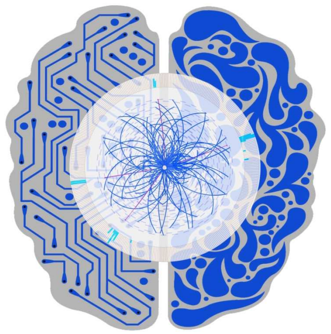


…

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

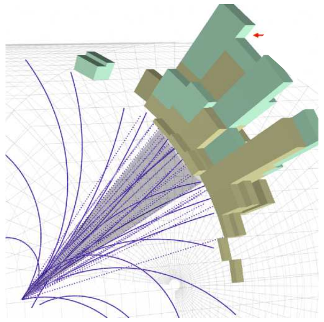
[http://iaifi.org]

Summary



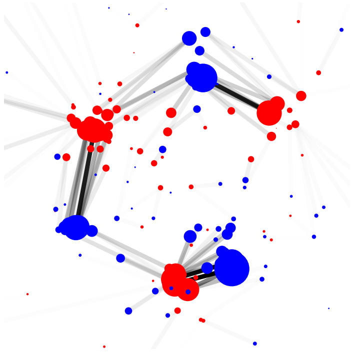
Rise of the Machines?

Machine learning offers powerful tools to analyze collision debris
Progress towards the fusion of deep learning and "deep thinking"



What is a Collider Event?

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



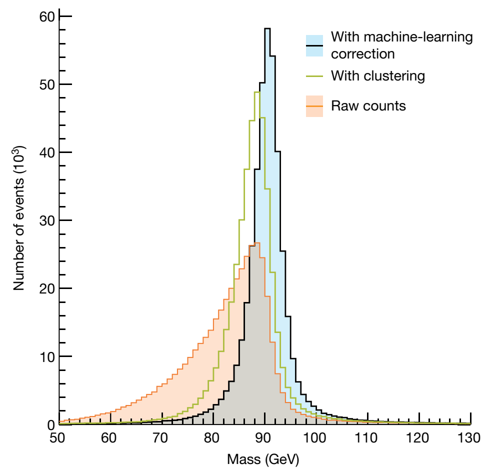
When are Collider Events Similar?

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

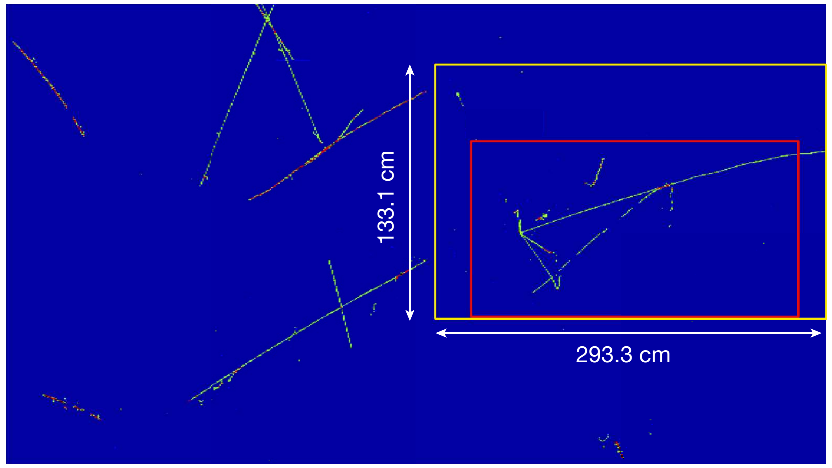
Backup Slides

Extensive Use of ML in HEP

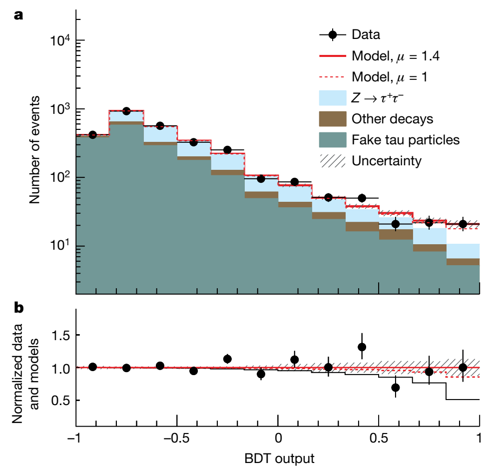
CMS: $Z \to e^+e^-$ calibration



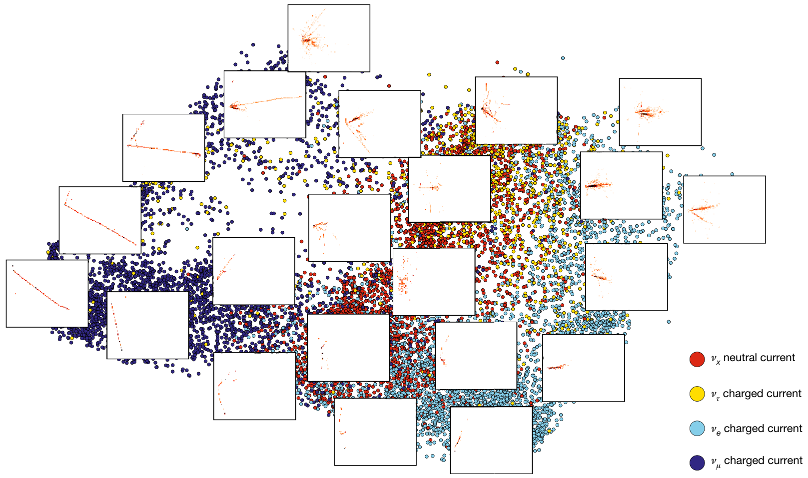
MicroBooNE: Object Identification



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



NOvA: Object Classification



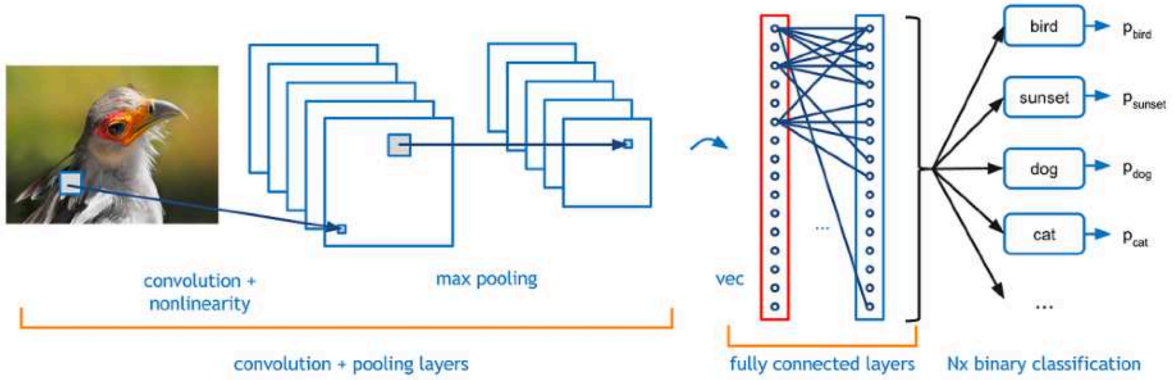
Machine learning is transforming many aspects of society, including fundamental physics research

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

Off-the-Shelf ML for HEP?

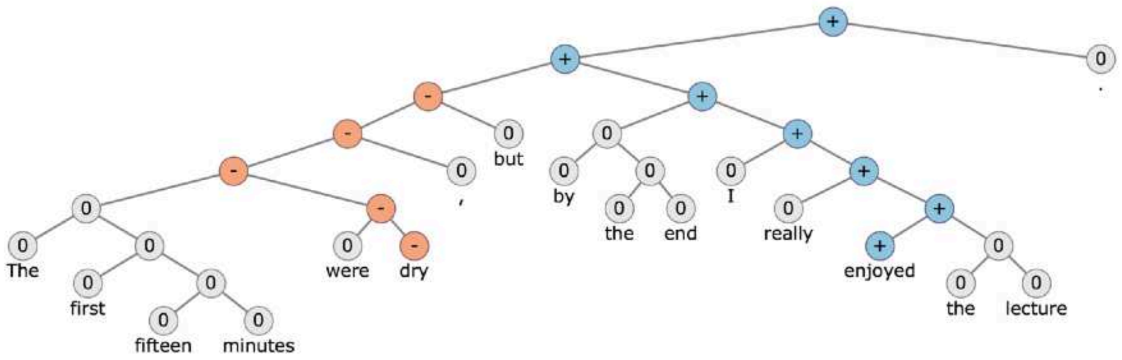
2D Images?

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



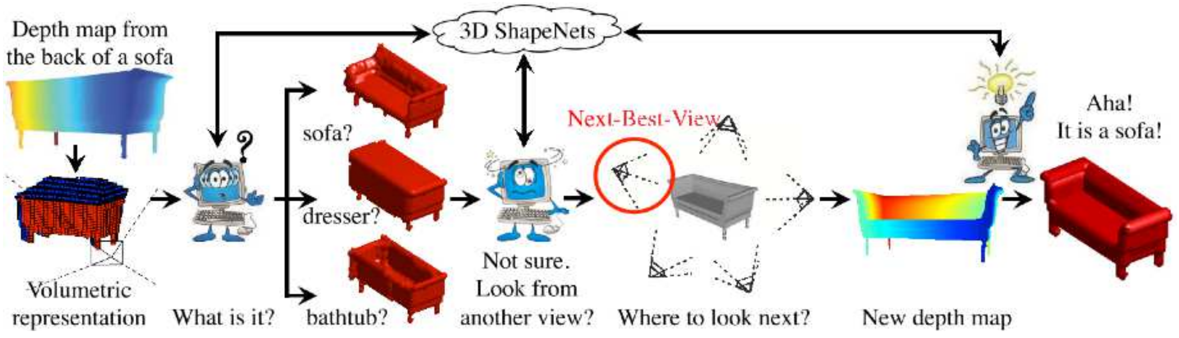
Natural Language?

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



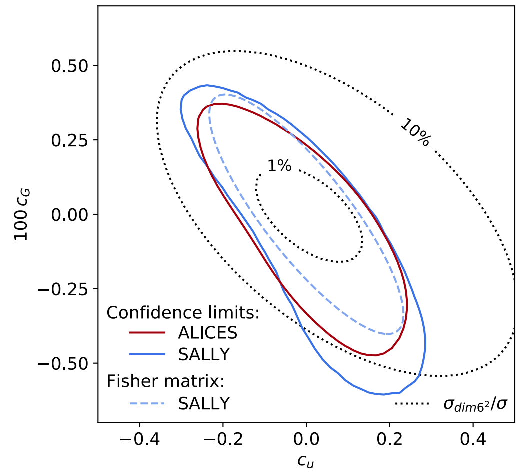
3D Objects?

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



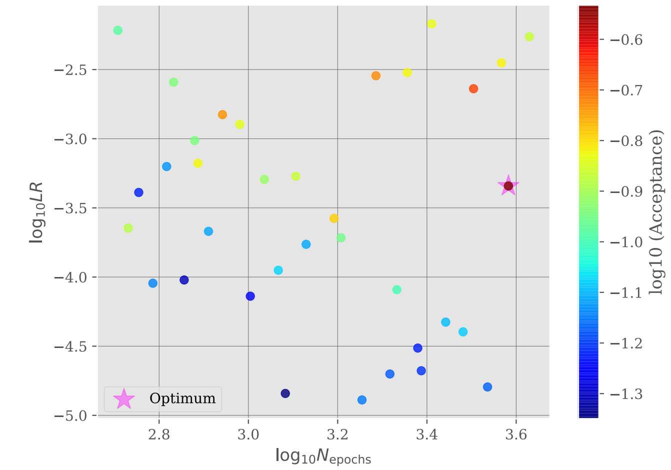
ML Targets for Collider Theory

e.g. Parameter Inference



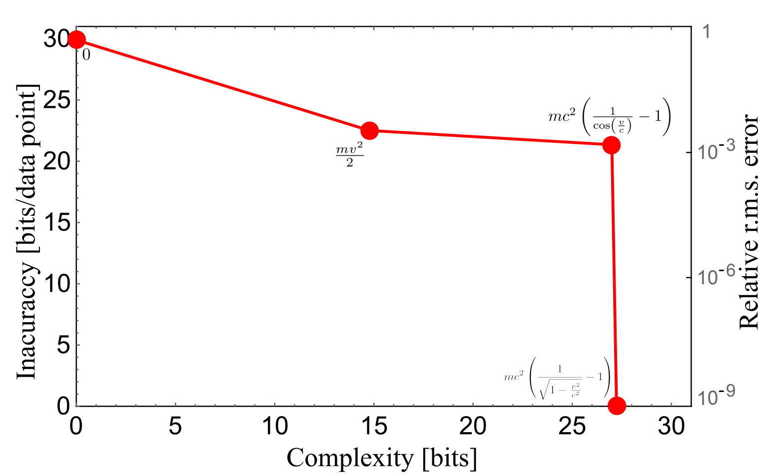
[Brehmer, Kling, Espejo, Cranmer, CSBS 2020]

e.g. Normalizing Flows



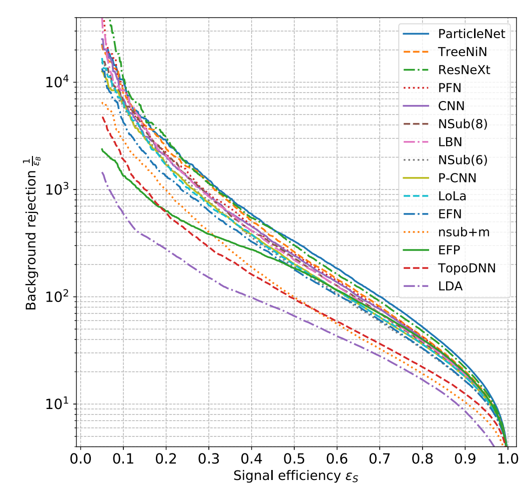
[Gao, Höche, Isaacson, Krause, Schulz, PRD 2020]

e.g. Symbolic Regression



[Udrescu, Tan, Feng, Neto, Wu, Tegmark, NeurIPS 2020]

e.g. Jet Classification



[Kasieczka, Plehn, et al., SciPost 2019]

[apologies for focus on research from my group in this talk; see HEPML-LivingReview for extensive bibliography]

Likelihood Ratio Trick

Key example of simulation-based inference

Many HEP problems can be expressed in this form!

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

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

$$-\min_{f(x)} L = \int \mathrm{d}x\, p(x) \log \frac{p(x)}{q(x)}$$ *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: $$L = \int \mathrm{d}x\, \mathcal{L}(x)$$

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

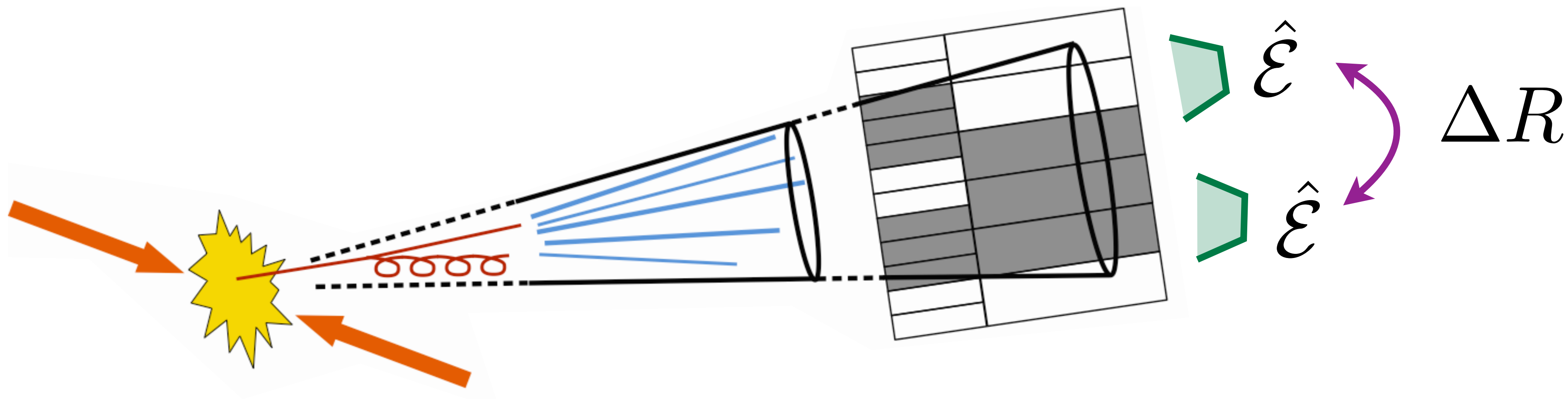
Euler-Lagrange: $$\frac{\partial \mathcal{L}}{\partial f} = 0$$

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

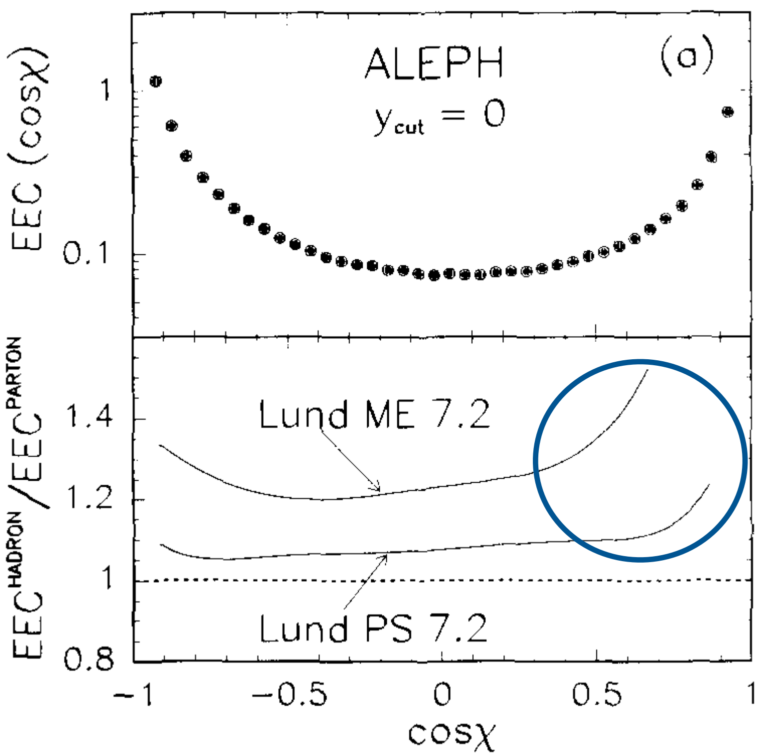
Requires shift in theoretical focus from solving problems to specifying problems

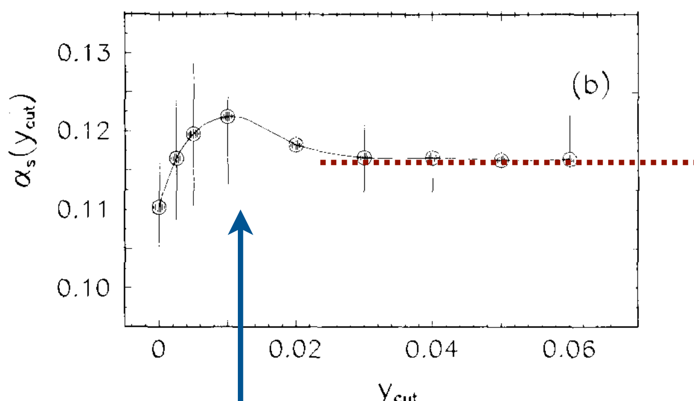
[see e.g. D'Agnolo, Wulzer, PRD 2019; simulation-based inference in Cranmer, Brehmer, Louppe, PNAS 2020; relation to f-divergences in Nguyen, Wainwright, Jordan, AoS 2009]

Energy-Energy Correlators



A long history in probing collinear dynamics of QCD



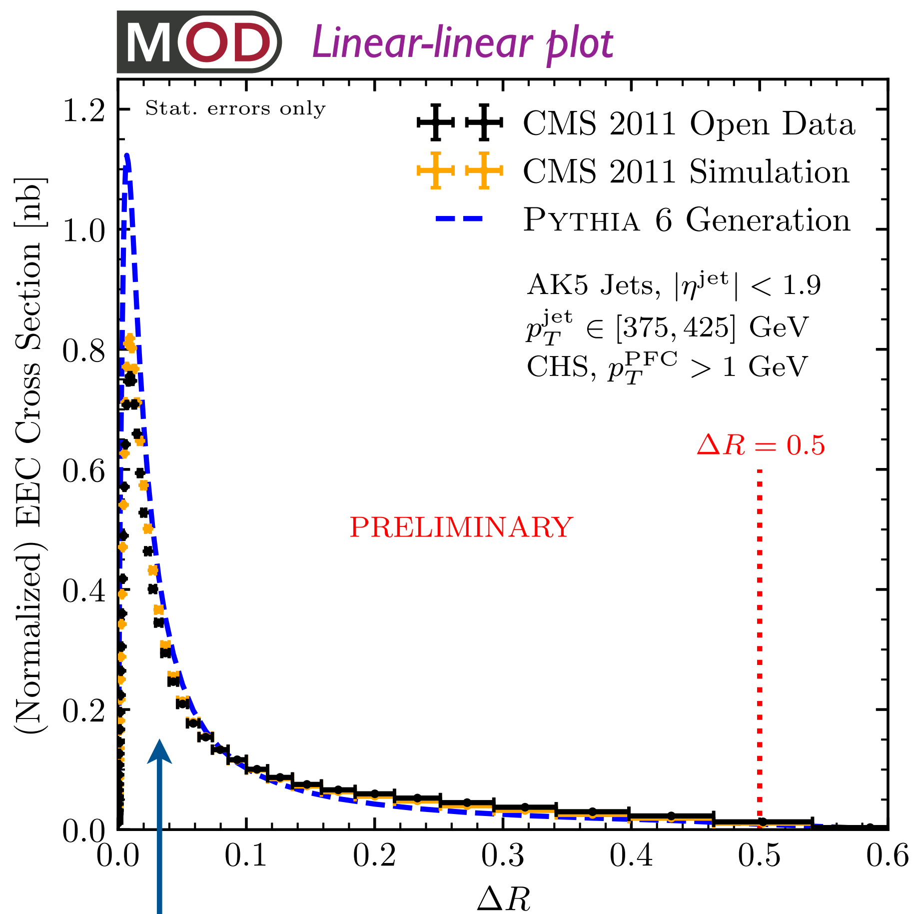


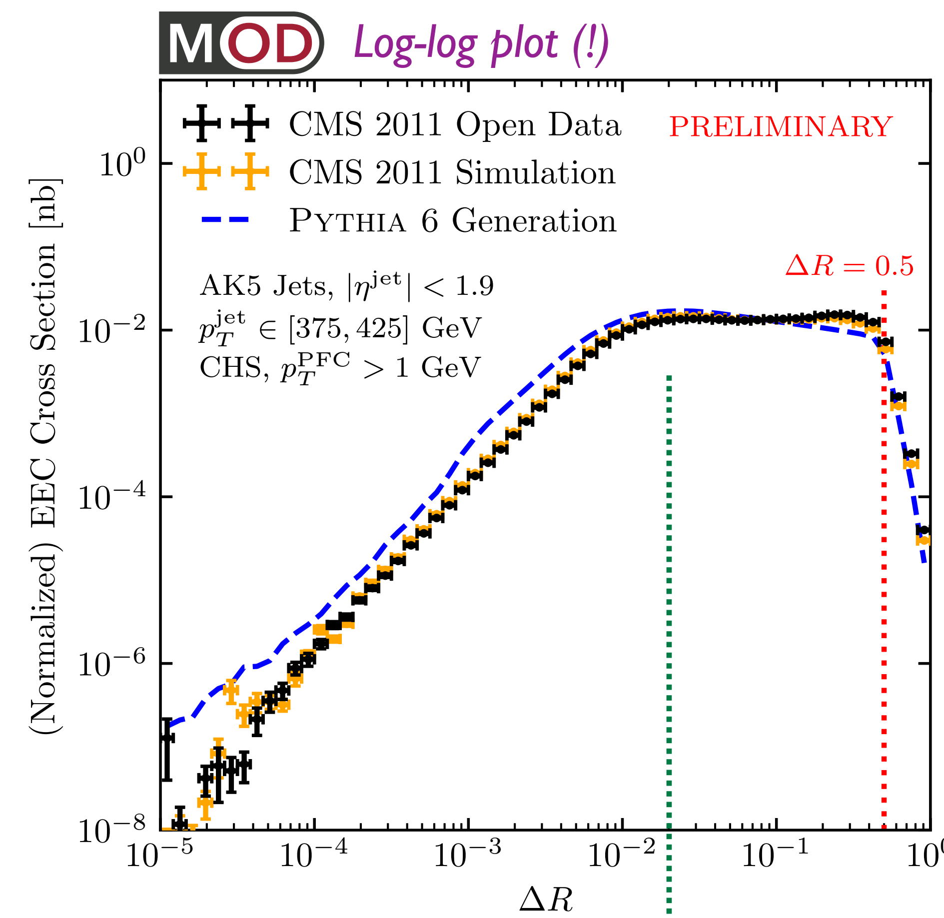
Extracting the strong coupling constant

Theoretical challenges with small angle (collinear) limit

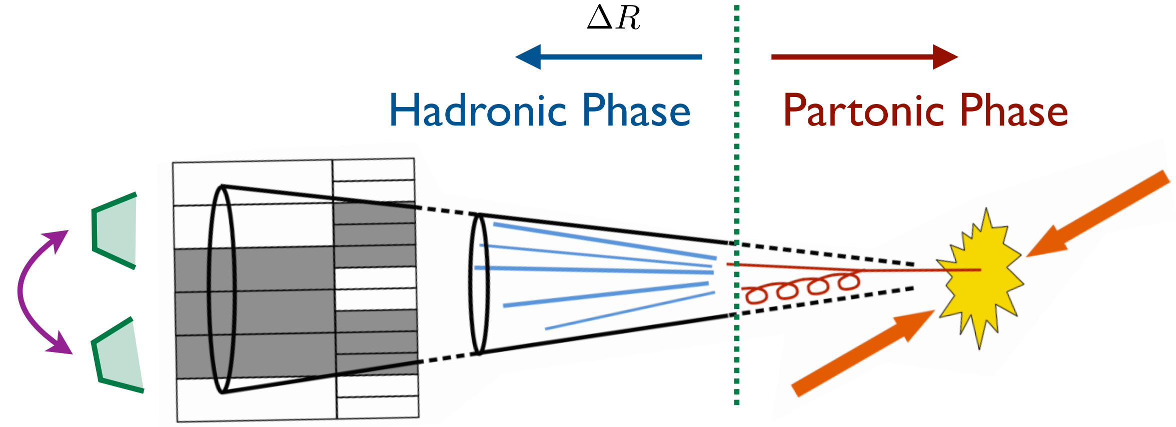
[Basham, Brown, Ellis, Love, PRL 1978; ALEPH, PLB 1991; see Chen, Moult, Zhang, Zhu, PRD 2020]

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]

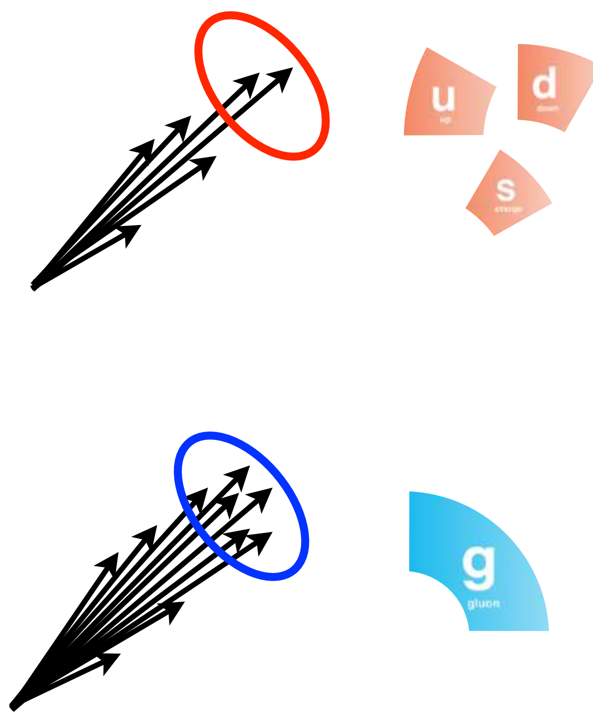
From Curmudgeon…

Jet classification via image recognition

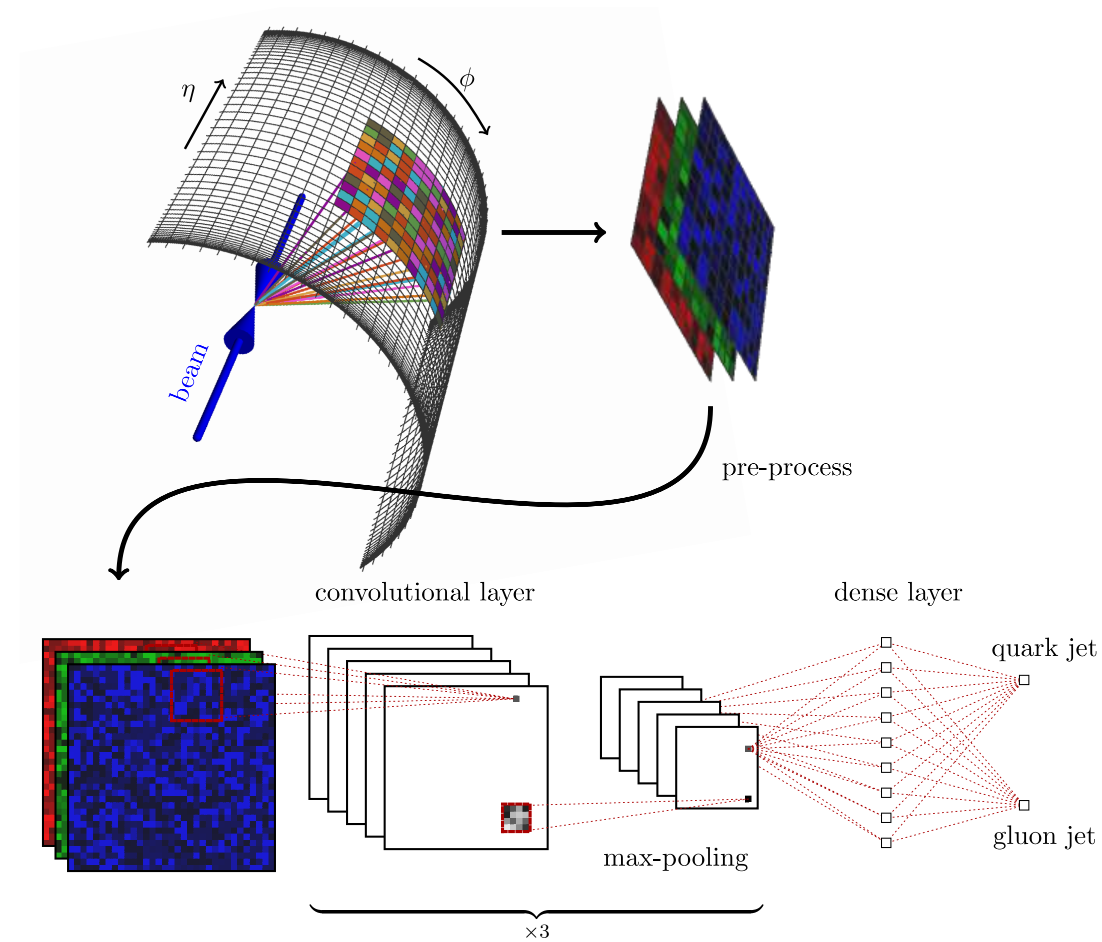
Quark

vs.

Gluon



Multi-channel convolutional neural networks



[e.g. Komiske, Metodiev, Schwartz, JHEP 2017;
cf. Gras, Höche, Kar, Larkoski, Lönnblad, Plätzer, Siódmok, Skands, Soyez, JDT, JHEP 2017]

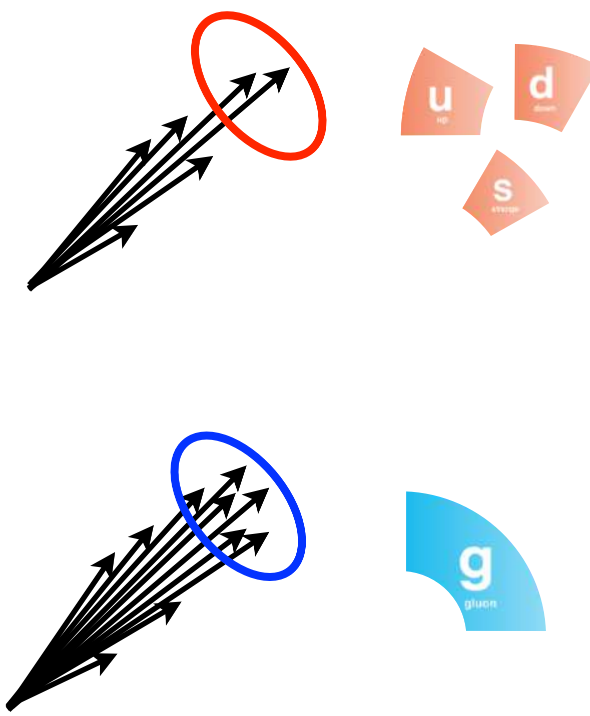
…to Evangelist

Jet flavor definitions via natural language processing

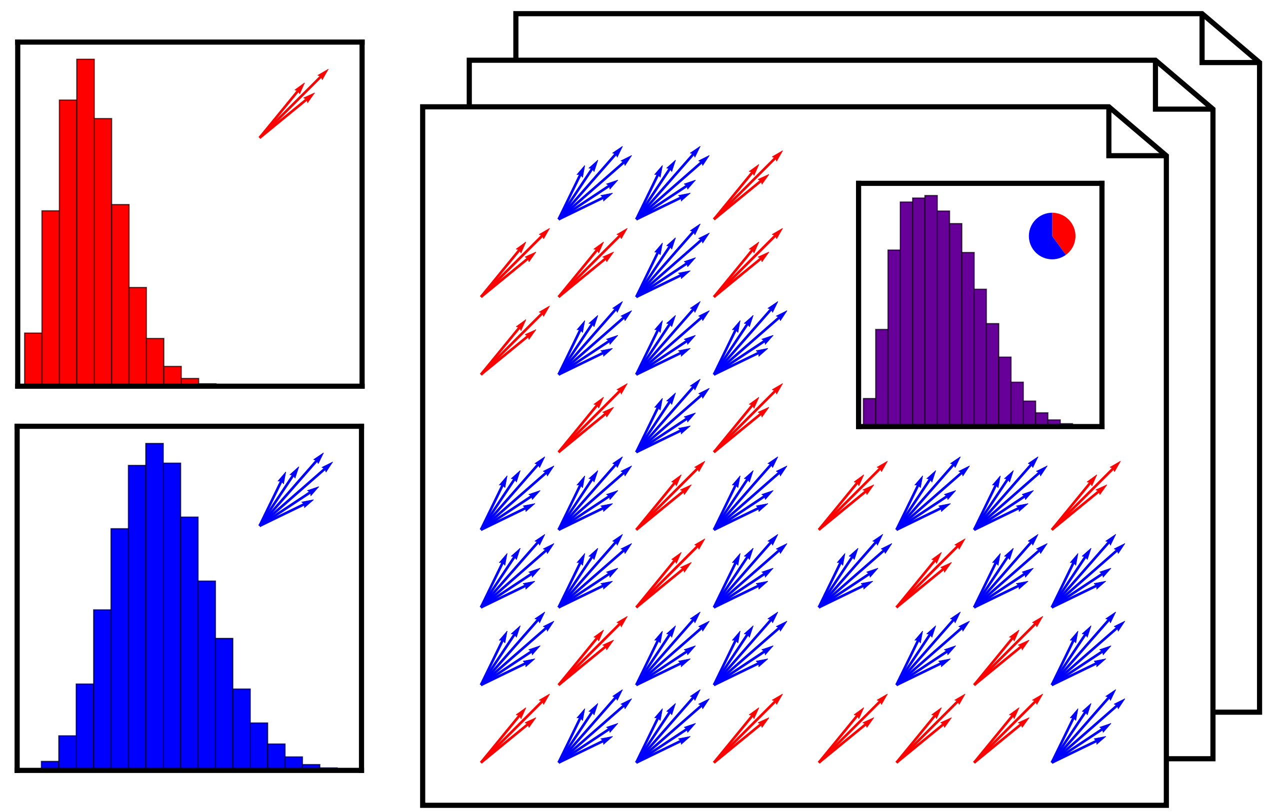
Quark

vs.

Gluon



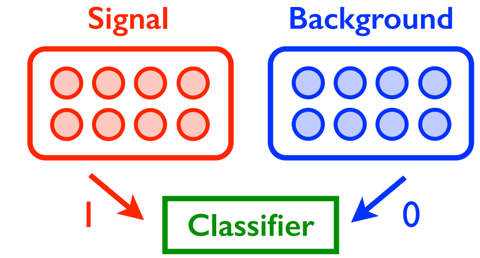
Topic Modeling / Blind Source Separation

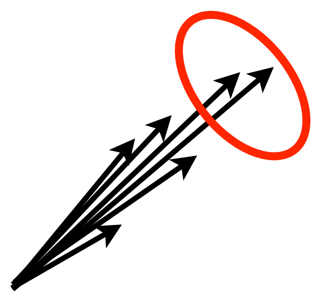
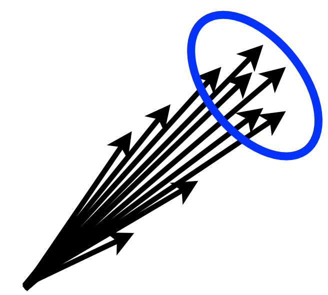


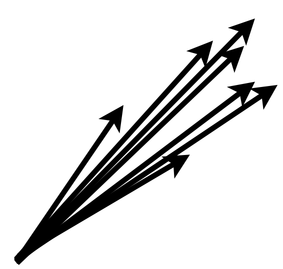
[Komiske, Metodiev, JDT, JHEP 2018;
using Metodiev, Nachman, JDT, JHEP 2017; Metodiev, JDT, PRL 2018; Komiske, Metodiev, JDT, JHEP 2019]

E.g. Quark/Gluon Classification

"Hello, World!" of Jet Physics



e.g.  **Quark** C_q = 4/3 vs.  **Gluon** C_g = 3 = 9/3

Find $h\Big($  $\Big)$ such that

$$h(\textbf{Quark}) = 1$$

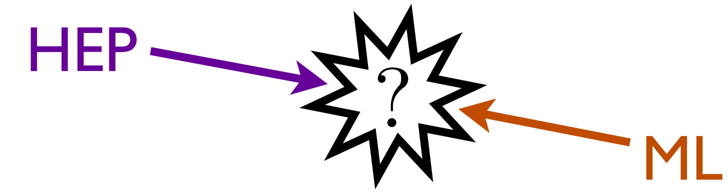
$$h(\textbf{Gluon}) = 0$$

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

(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

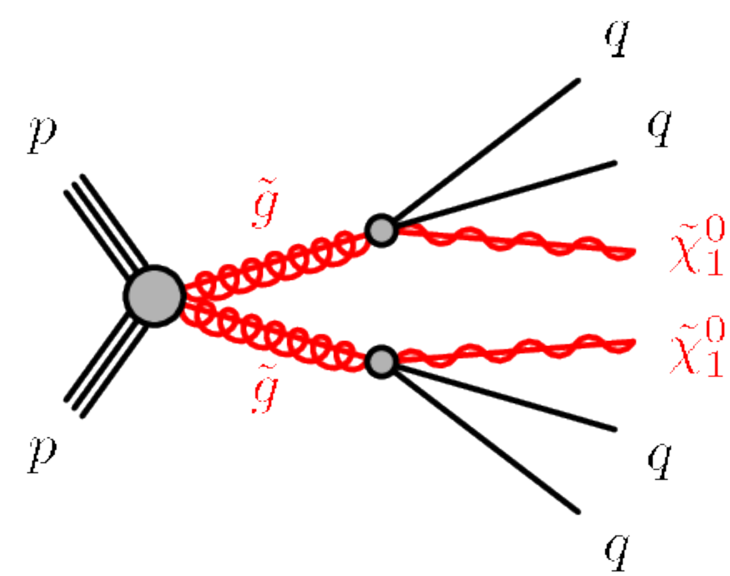


Classifier: Boosted decision tree (for each of 4 jets)

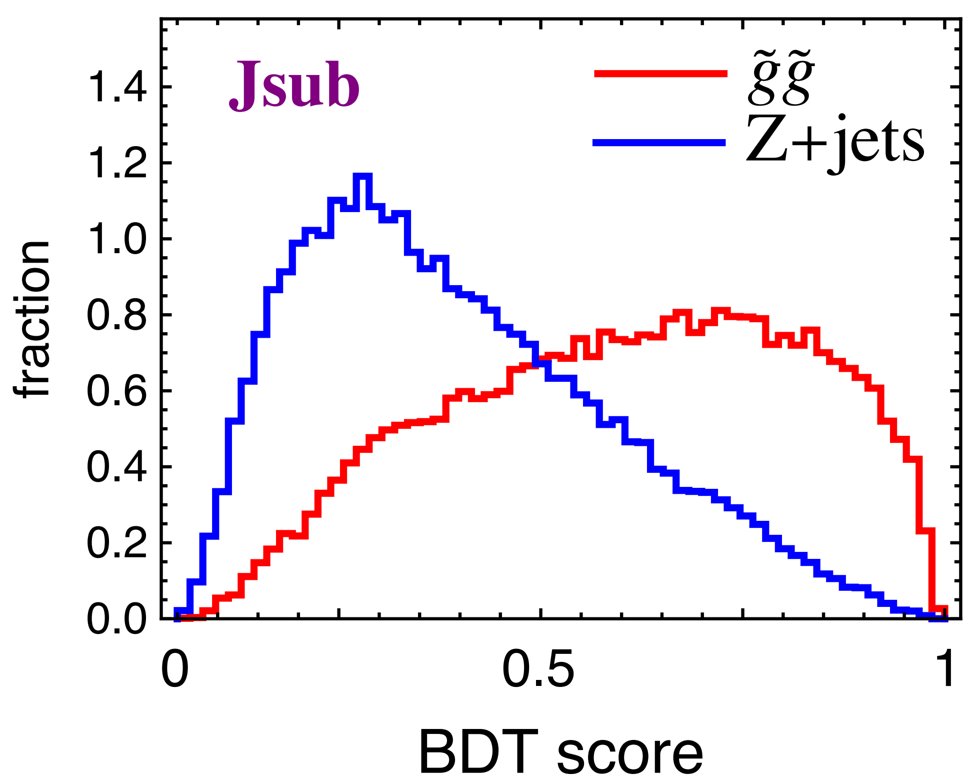
Inputs: Jet mass, width, track multiplicity

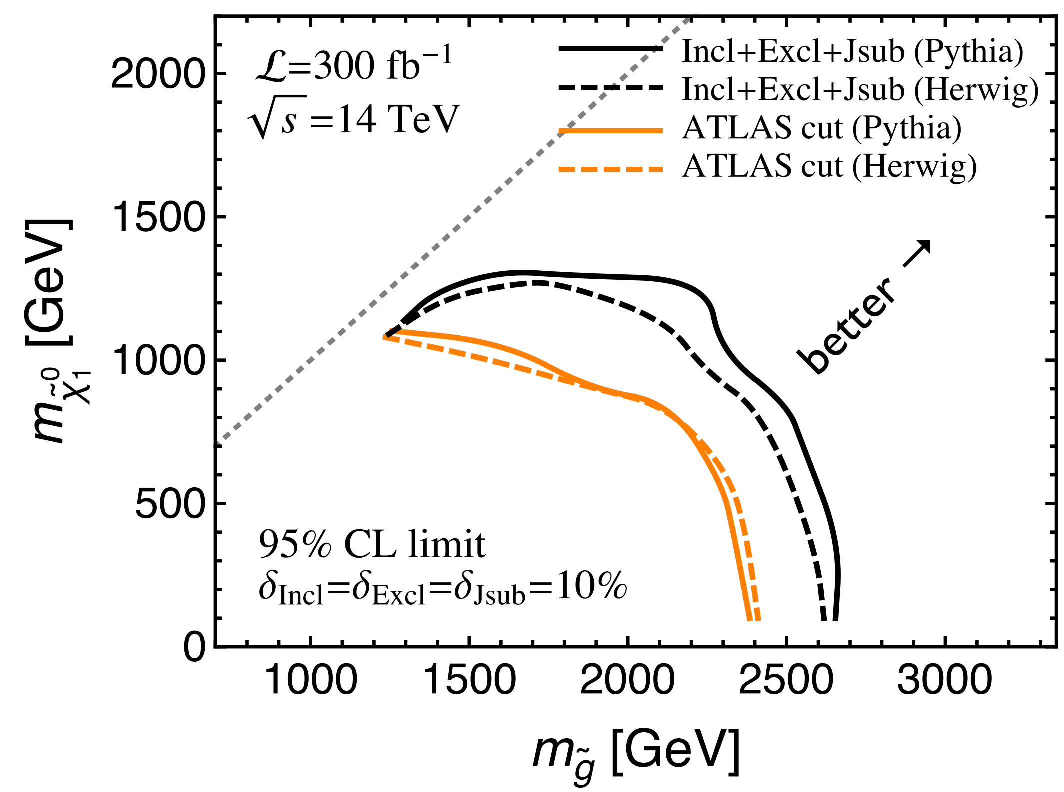
Signal: **Quark enriched**

Background: **Gluon enriched**



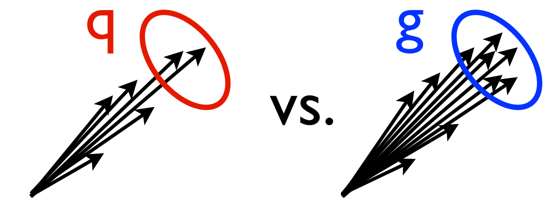
← background signal →





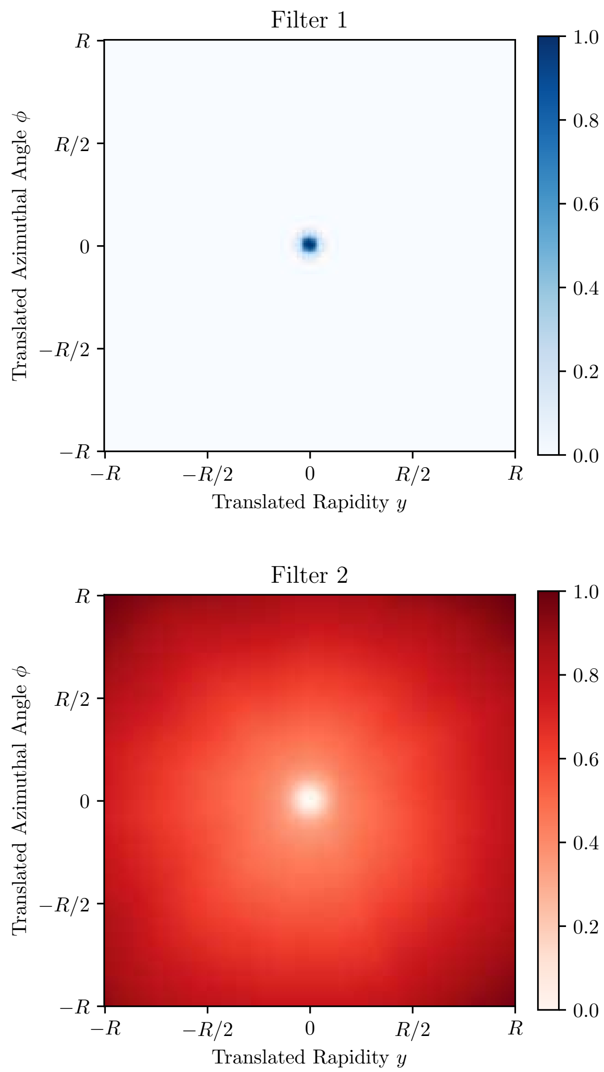
[Bhattacherjee, Mukhopadhyay, Nojiri, Sakakie, Webber, JHEP 2017]

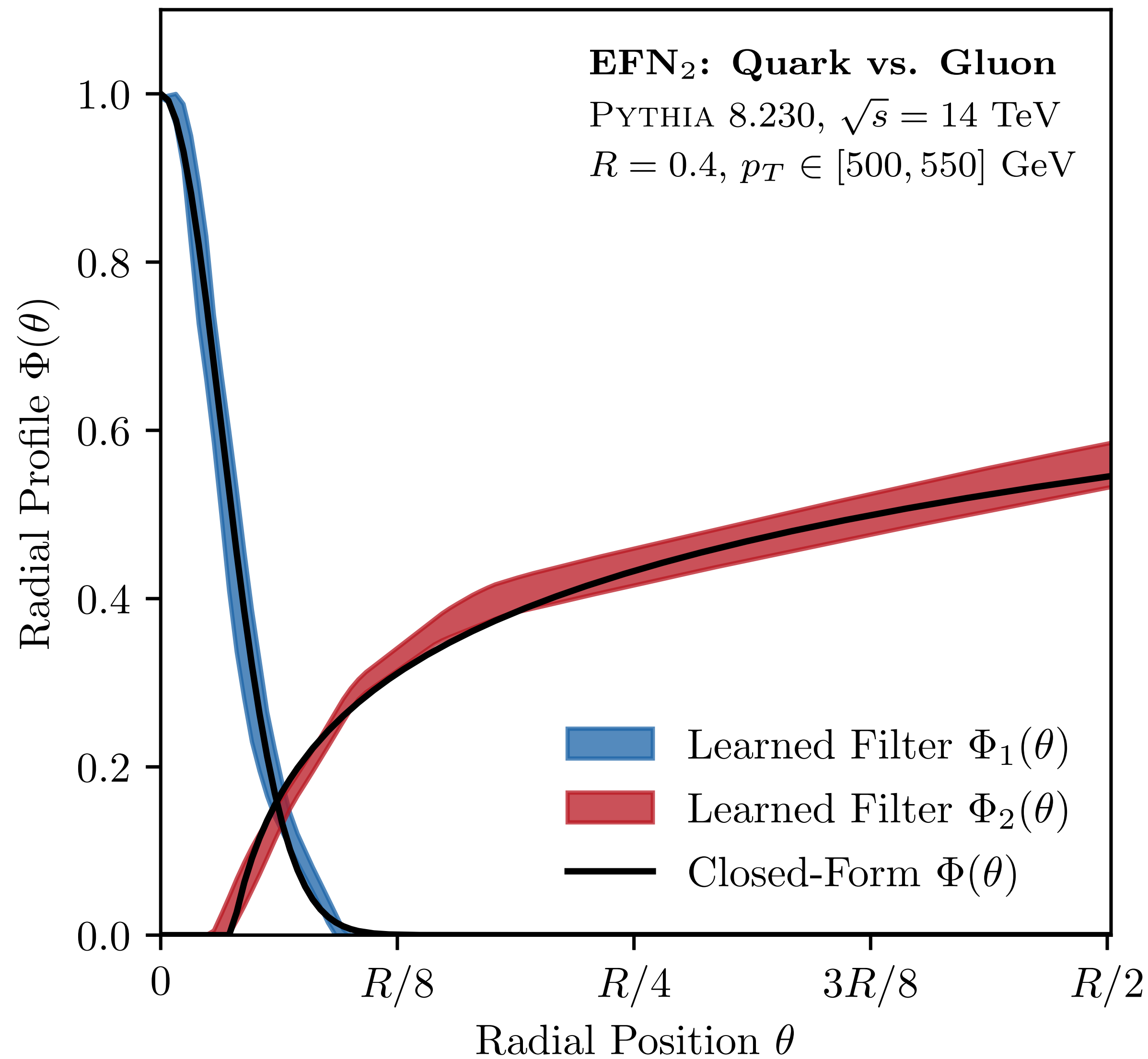
Learning from the Machine



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

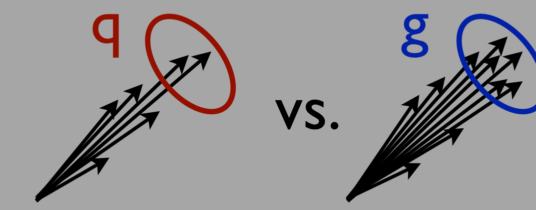
cf. Angularities:
$f(\theta) = \theta^\beta$





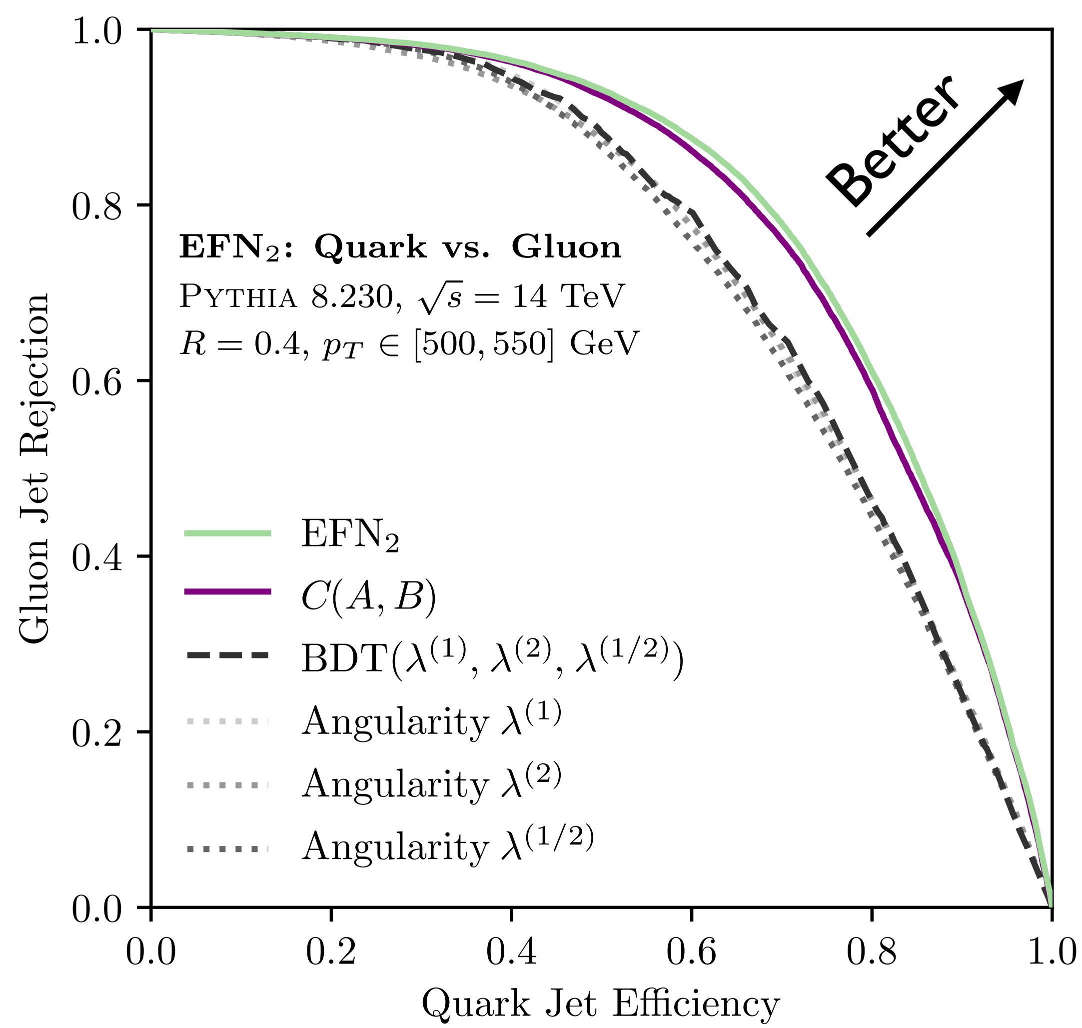
[Komiske, Metodiev, JDT, JHEP 2019;
cf. Larkoski, JDT, Waalewijn, JHEP 2014; using Berger, Kucs, Sterman, PRD 2003; Ellis, Vermilion, Walsh, Hornig, Lee, JHEP 2010]

Learning from the Machine



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

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



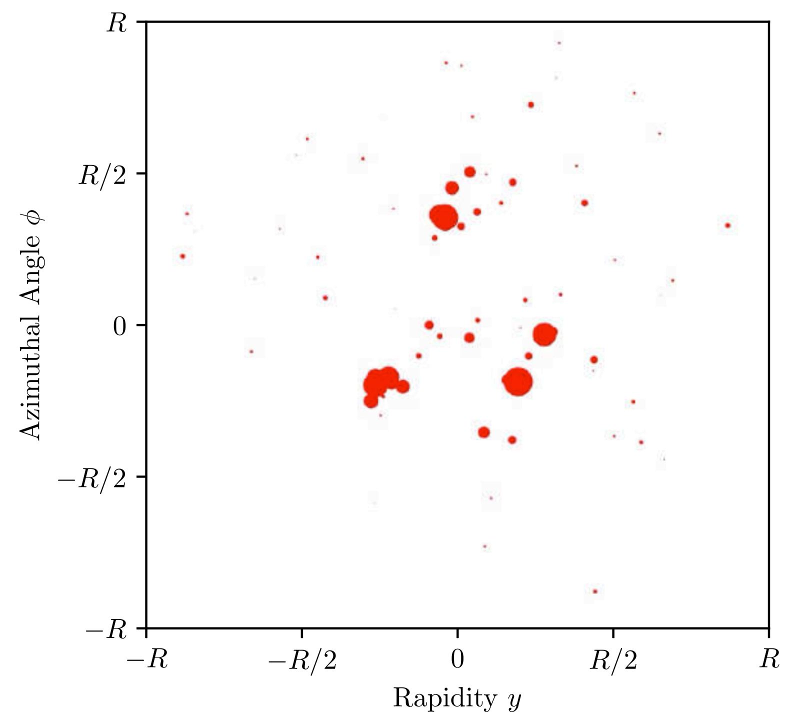
EFN outperformed a domain expert (i.e. me)

But we reverse engineered the machine (and learned something about QCD)

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

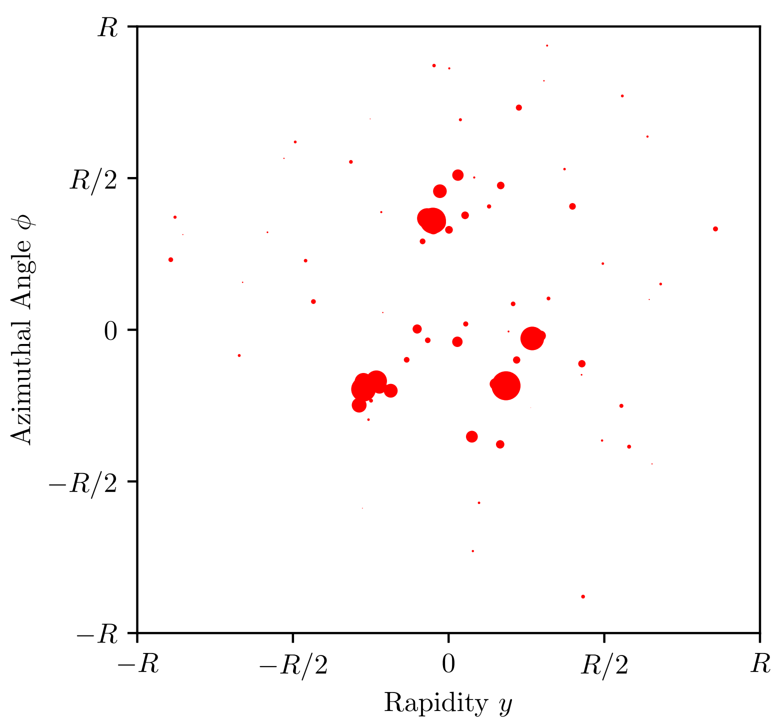
Similarity of Two Energy Flows?

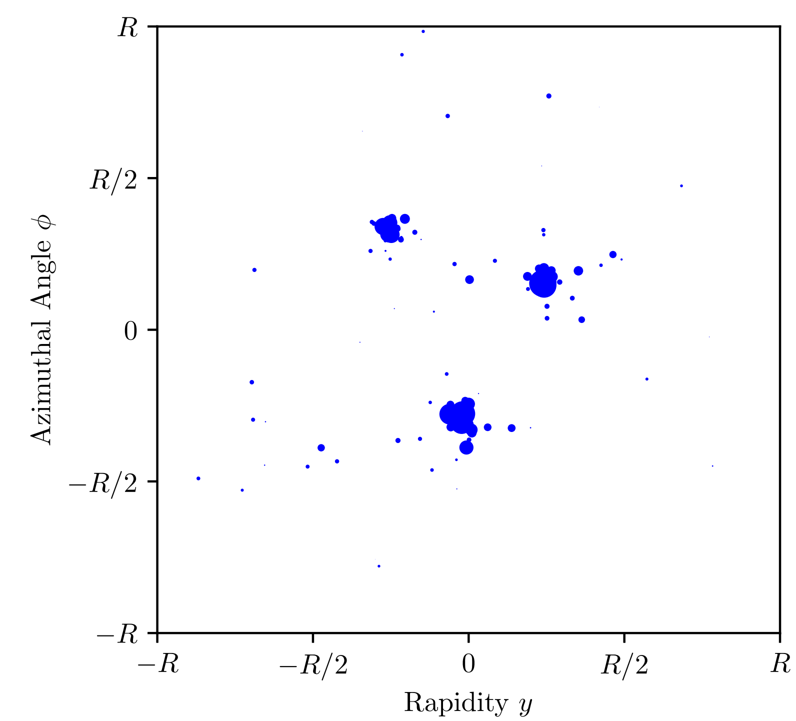
$$\mathcal{E}(\hat{n}) = \sum_i E_i \, \delta(\hat{n} - \hat{n}_i)$$



Optimal Transport:

Earth Mover's Distance
a.k.a. 1-Wasserstein metric





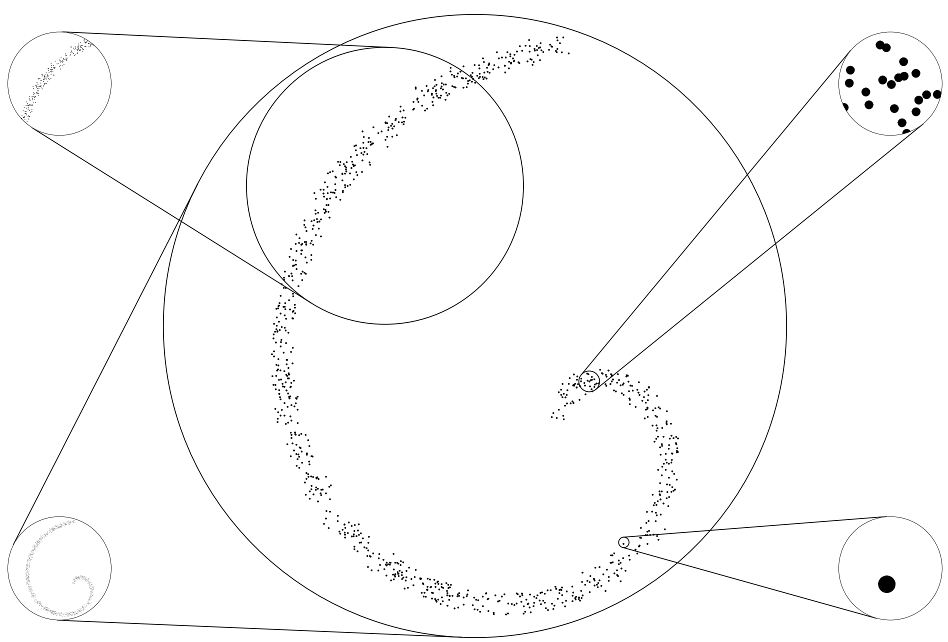
[Komiske, Metodiev, JDT, PRL 2019; code at Komiske, Metodiev, JDT, energyflow.network]

Dimensionality of Space of Jets

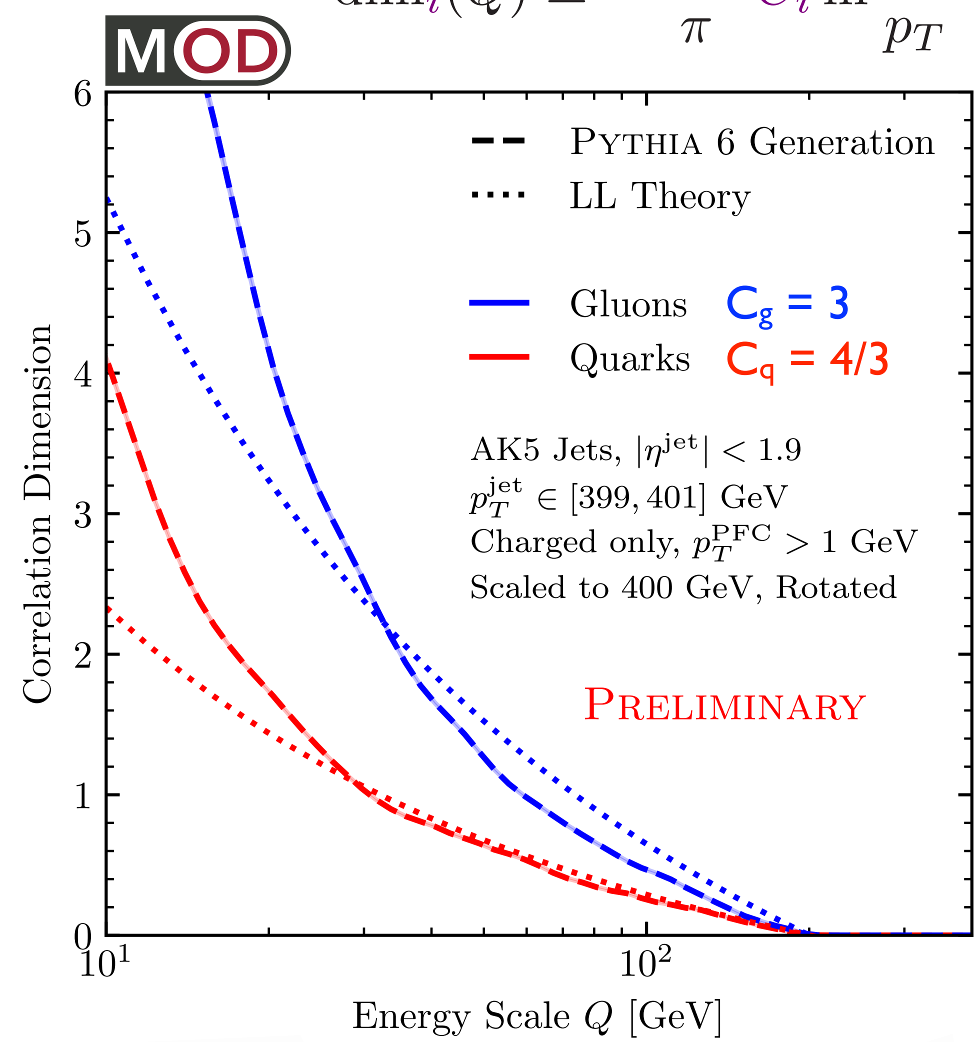
$$N_{\text{neighbors}}(r) \sim r^{\text{dim}}$$

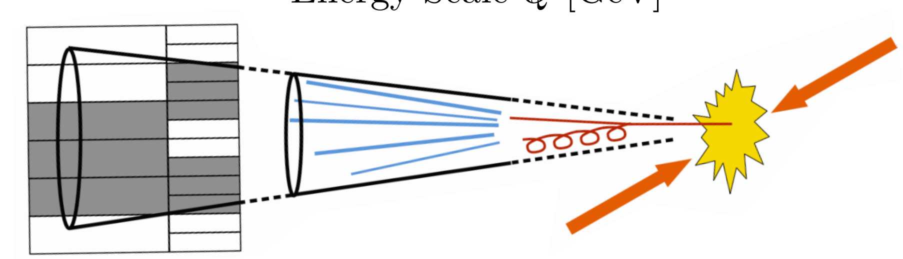
$$\Rightarrow \quad \text{dim}(r) \sim r \frac{\partial}{\partial r} \ln N_{\text{neighbors}}(r)$$

[Grassberger, Procaccia, PRL 1983; Kégl, NIPS 2002]

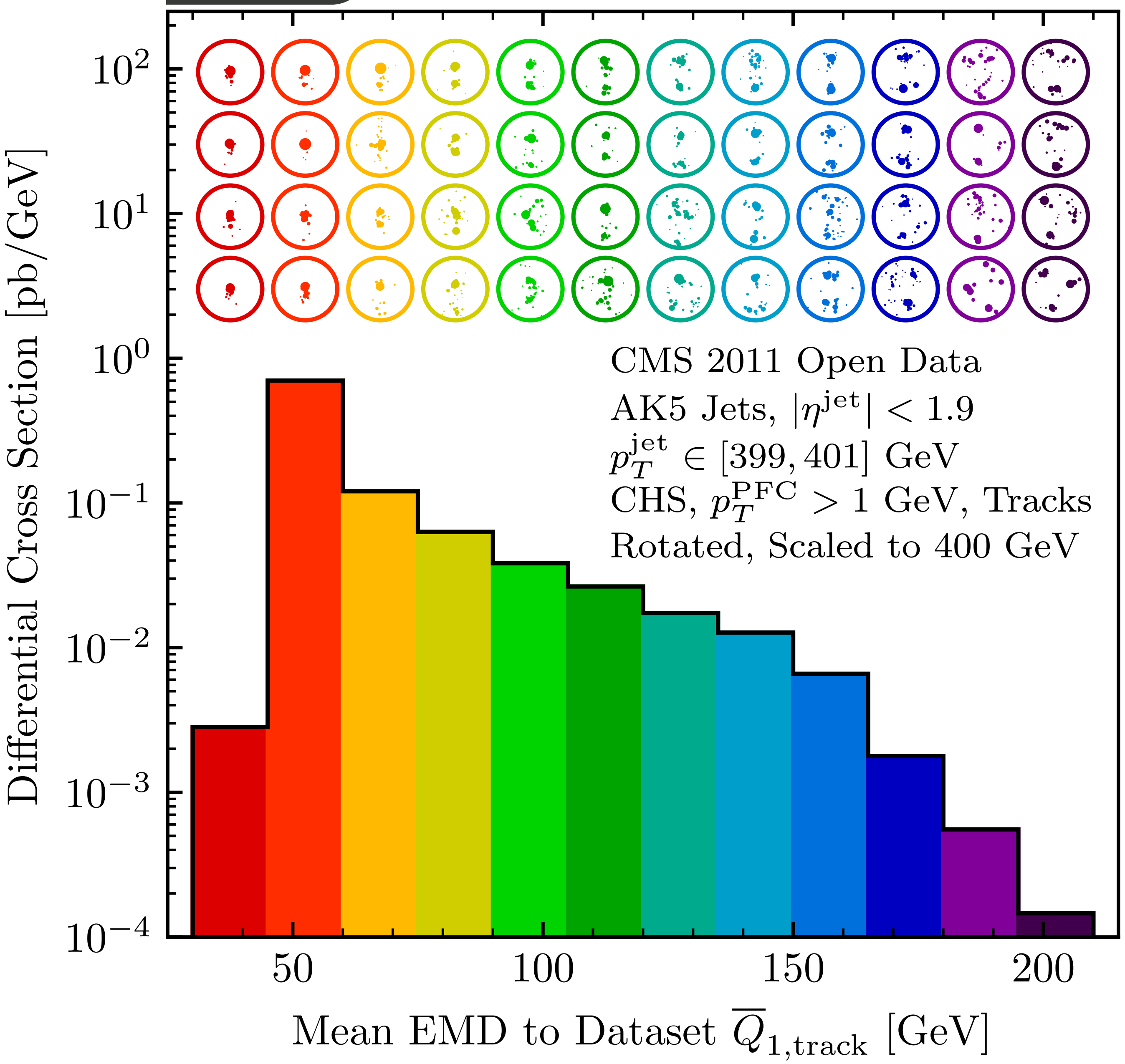


$$\text{dim}_i(Q) \simeq -\frac{8\alpha_s}{\pi} C_i \ln \frac{Q}{p_T}$$



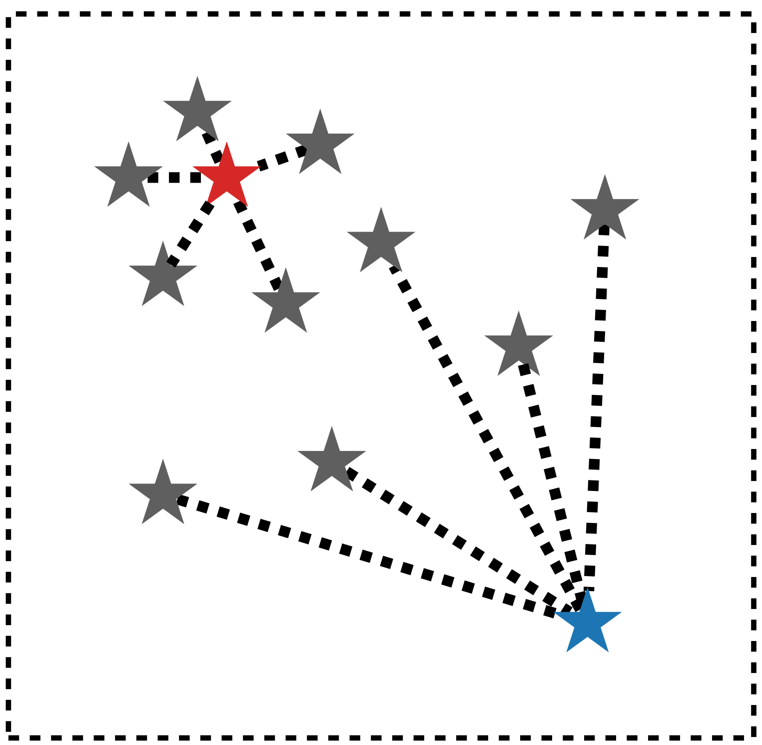


Least Representative Jets



New Physics?

Or tails of QCD?



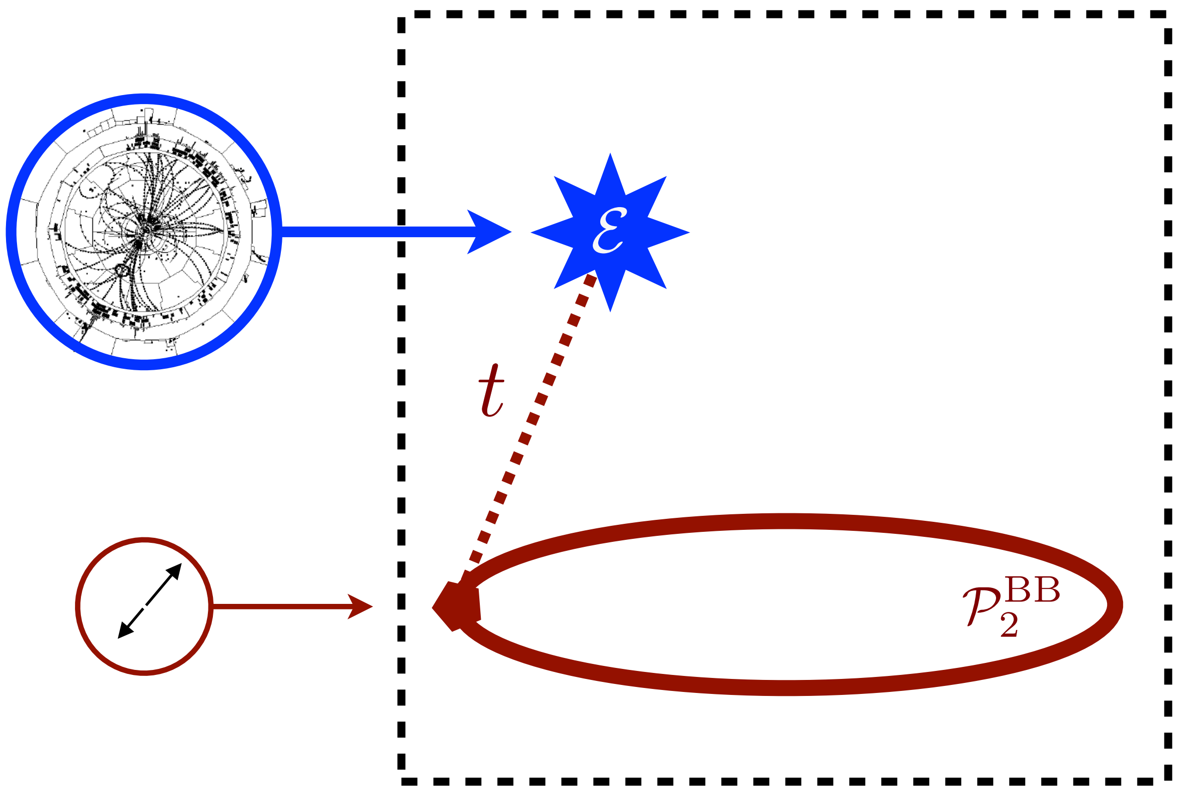



[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020]

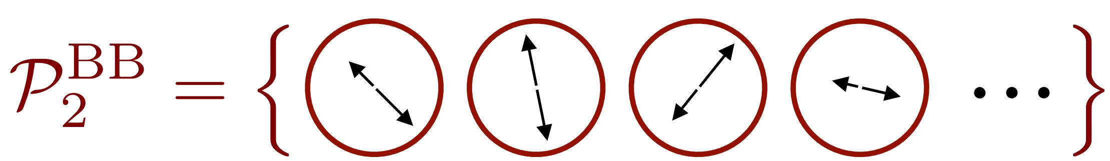
Back to the Future with Thrust

How dijet-like is an event?

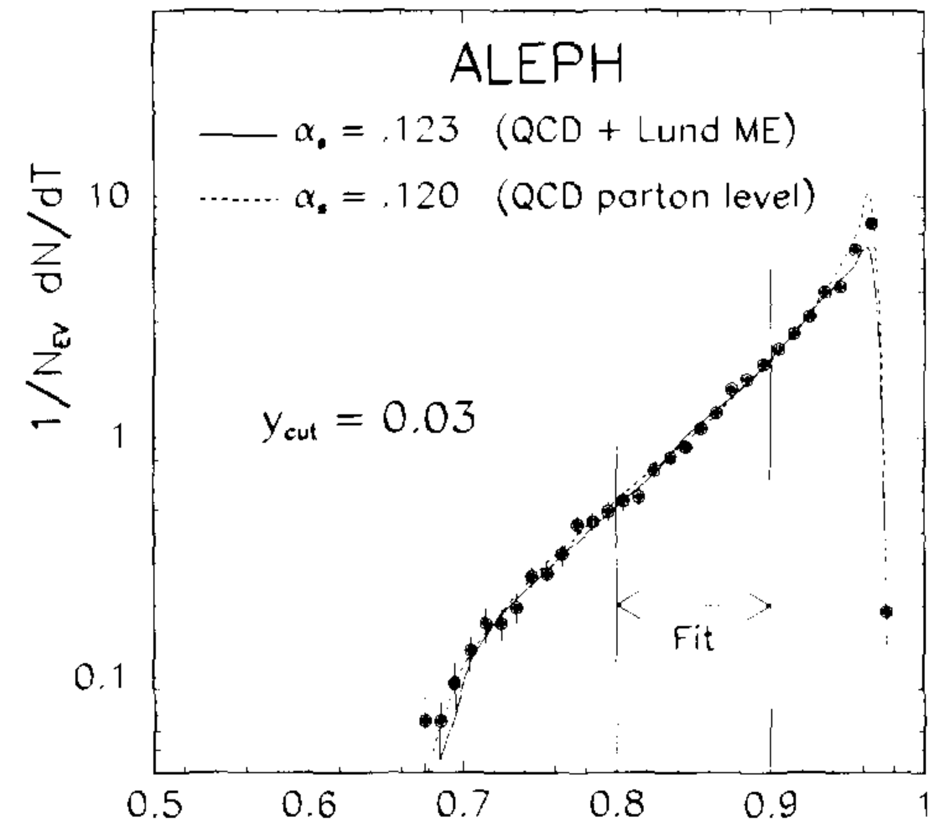
$$t(\mathcal{E}) = \min_{\mathcal{E}' \in \mathcal{P}_2^{\rm BB}} \mathrm{EMD}_2(\mathcal{E}, \mathcal{E}')$$



All Back-to-Back Two Particle Configurations



(using β=2 EMD variant)



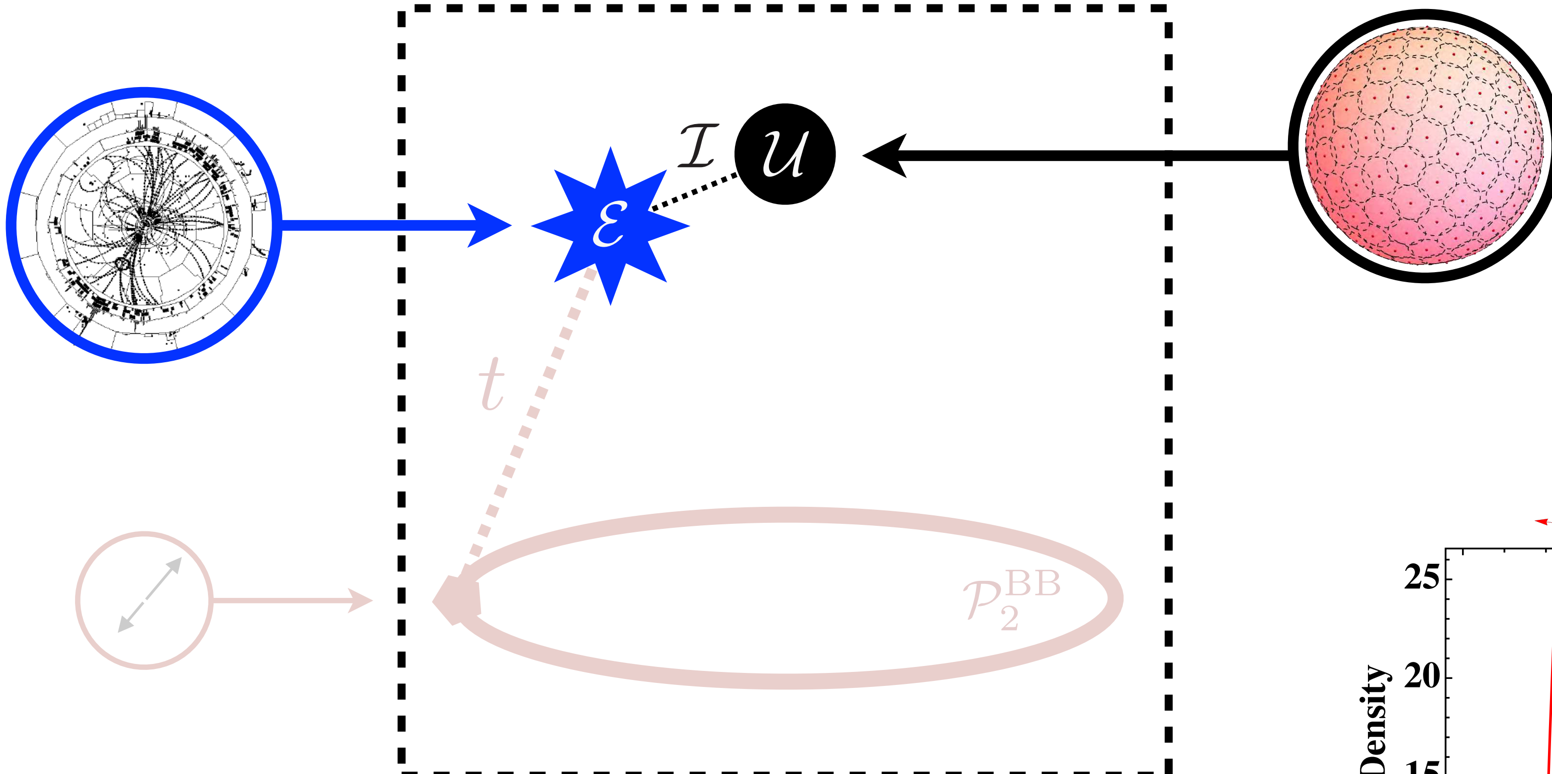
$$1 - \frac{t}{2E_{\rm CM}}$$

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

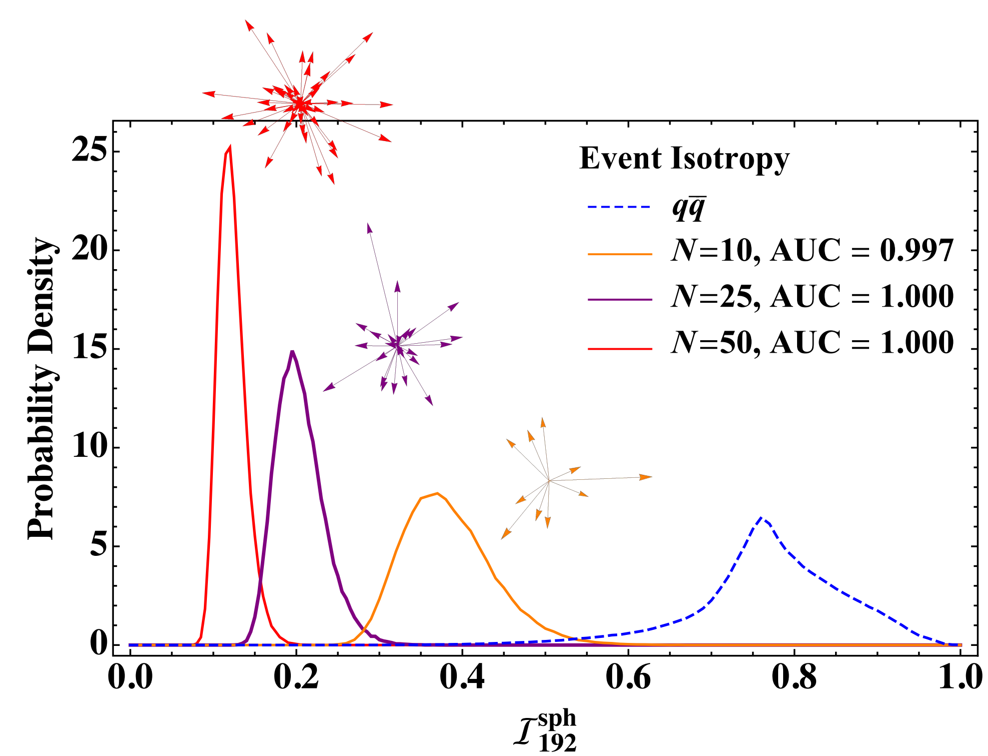
[Komiske, Metodiev, JDT, JHEP 2020]
[Brandt, Peyrou, Sosnowski, Wroblewski, PL 1964; Farhi, PRL 1977; ALEPH, PLB 1991]

Event Isotropy from Collider Geometry

How uniform is an event?

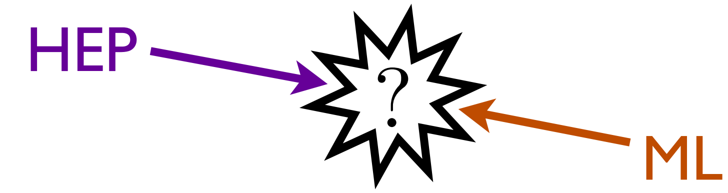


$$\mathcal{I}(\mathcal{E}) = \text{EMD}(\mathcal{E}, \mathcal{U})$$

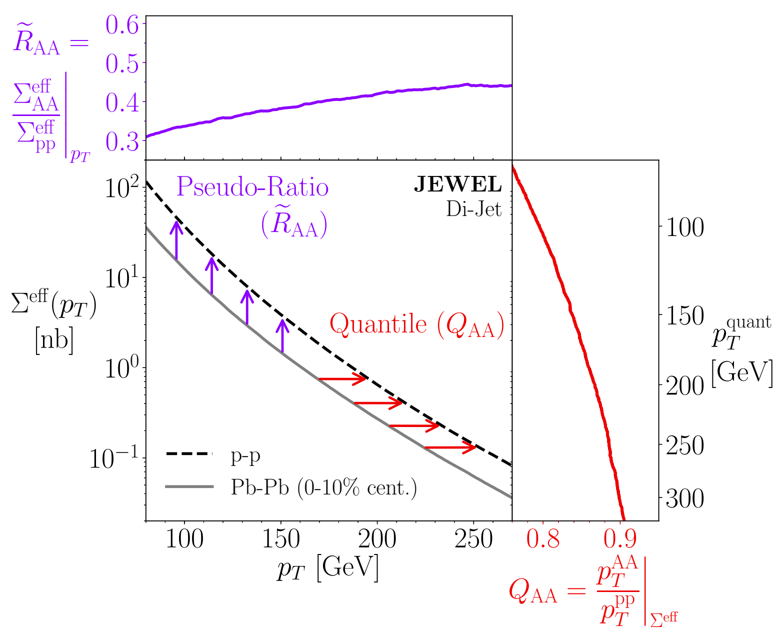


[Cesarotti, JDT, JHEP 2020;
see also Cesarotti, Reece, Strassler, arXiv 2020]

More Collisions

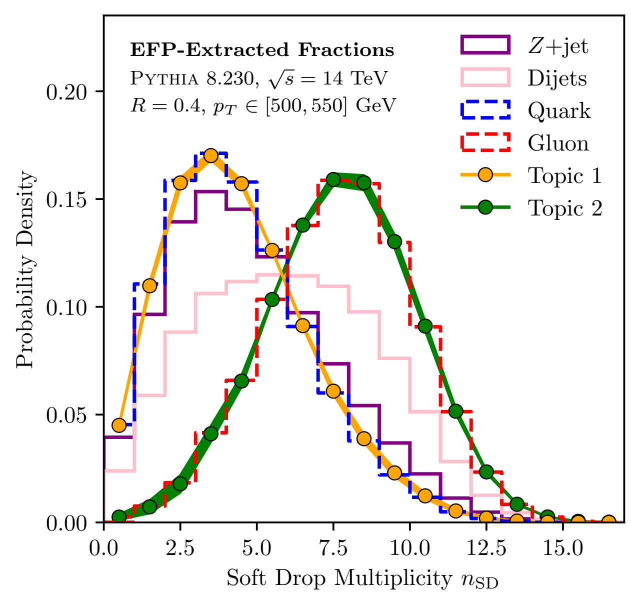


Jet Quenching
via Optimal Transport



[Brewer, Milhano, JDT, PRL 2019]

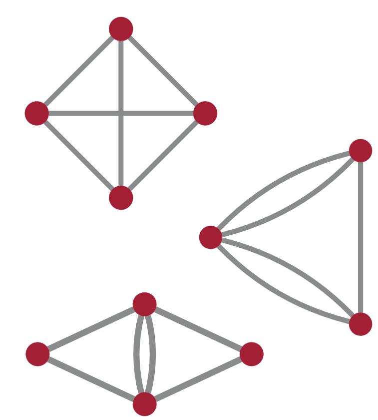
Quark/Gluon Definitions
via Blind Source Separation



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

Kinematic Decomposition
via Graph Theory

	Leafless Multigraphs	
	Connected	All
Edges d	A307317	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]