Collision Course

Particle Physics meets Machine Learning

Jesse Thaler

8.S50 — January 15, 2021

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

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

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

MIT CSAIL

MIT KAVLI INSTITUTE

IDSS

Pulkit Agrawal
Lisa Barsotti
Isaac Chuang
William Detmold
Bill Freeman
Philip Harris
Kerstin Perez
Alexander Rakhlin

Phiala Shanahan
Tracy Slatyer
Marin Soljacic
Justin Solomon
Washington Taylor
Max Tegmark
Jesse Thaler
Mike Williams

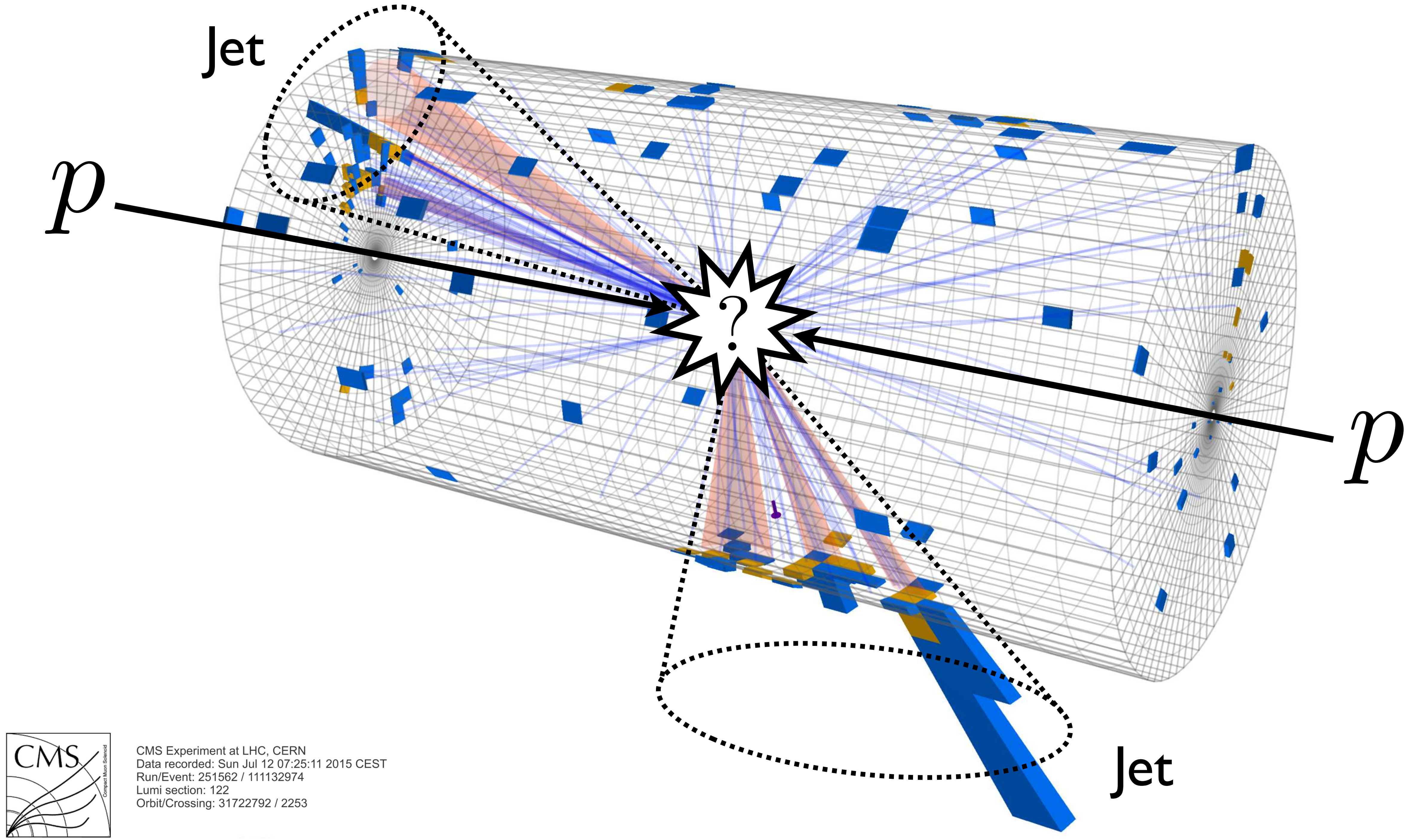
Demba Ba
Edo Berger
Cora Dvorkin
Daniel Eisenstein
Doug Finkbeiner
Matthew Schwartz
Yaron Singer
Todd Zickler

James Halverson
Brent Nelson

Taritree Wongjirad

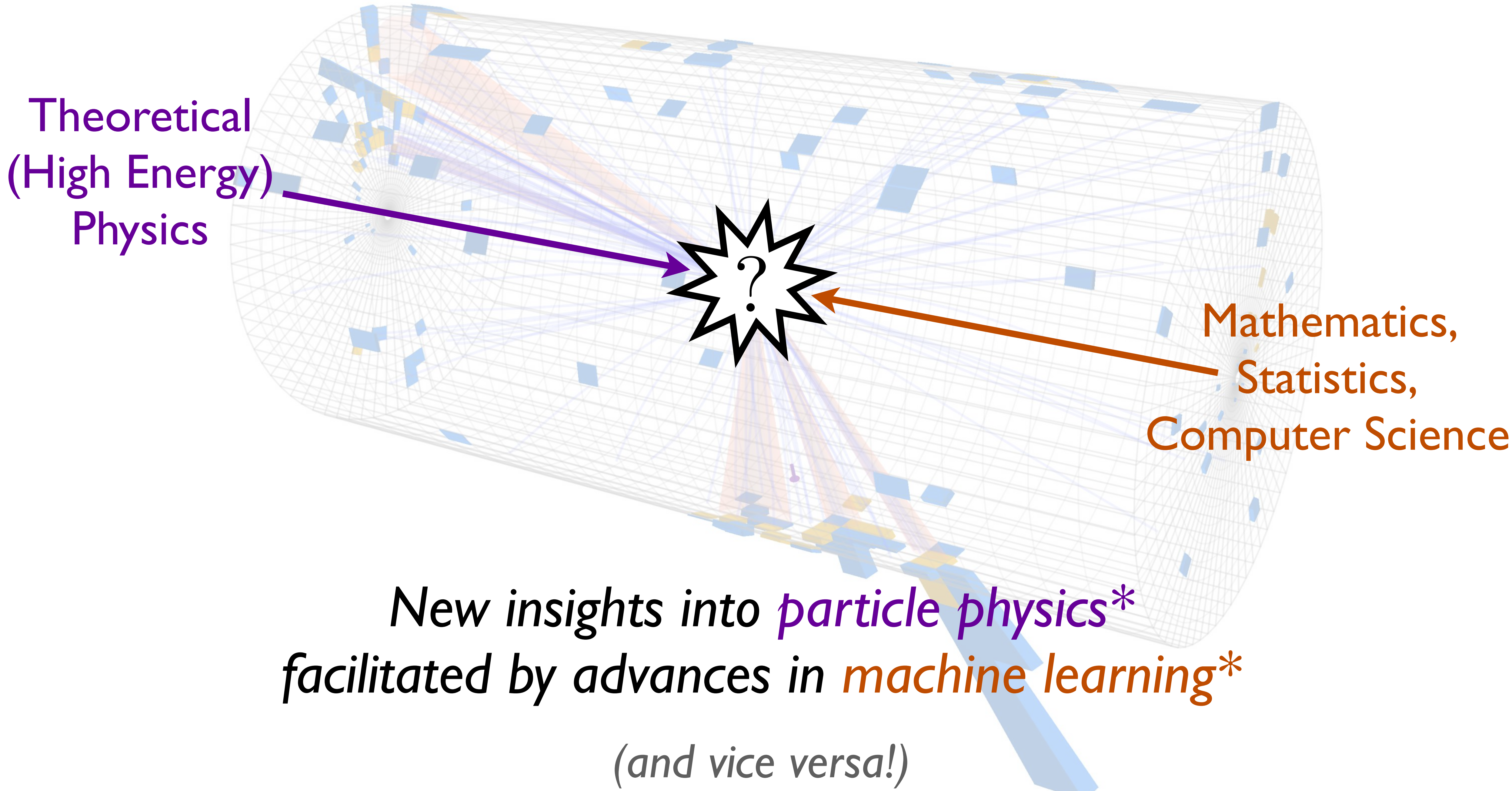
MIT $\subset$ Boston Area: Critical Mass for Transformative Ab Initio AI Research

"Collision Course"



“Collision Course”

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

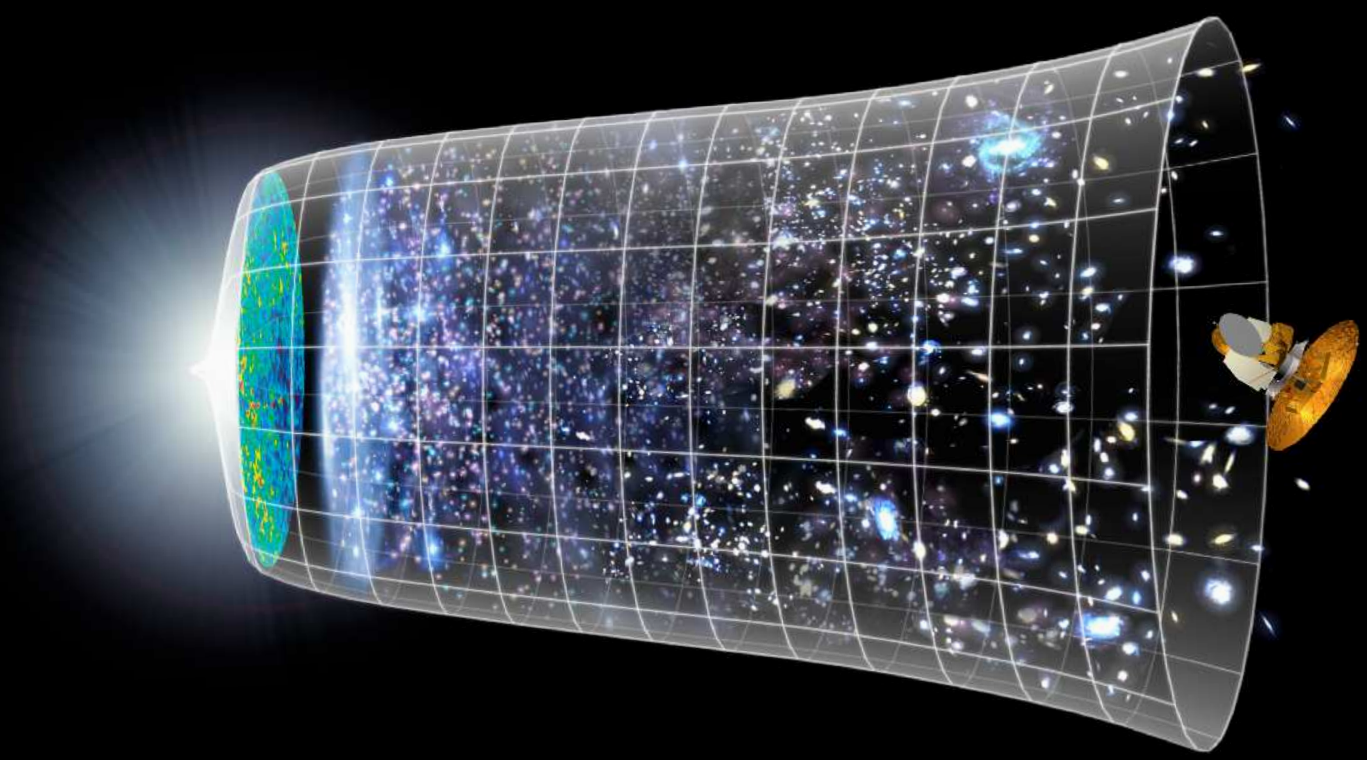


*New insights into particle physics**
*facilitated by advances in machine learning**

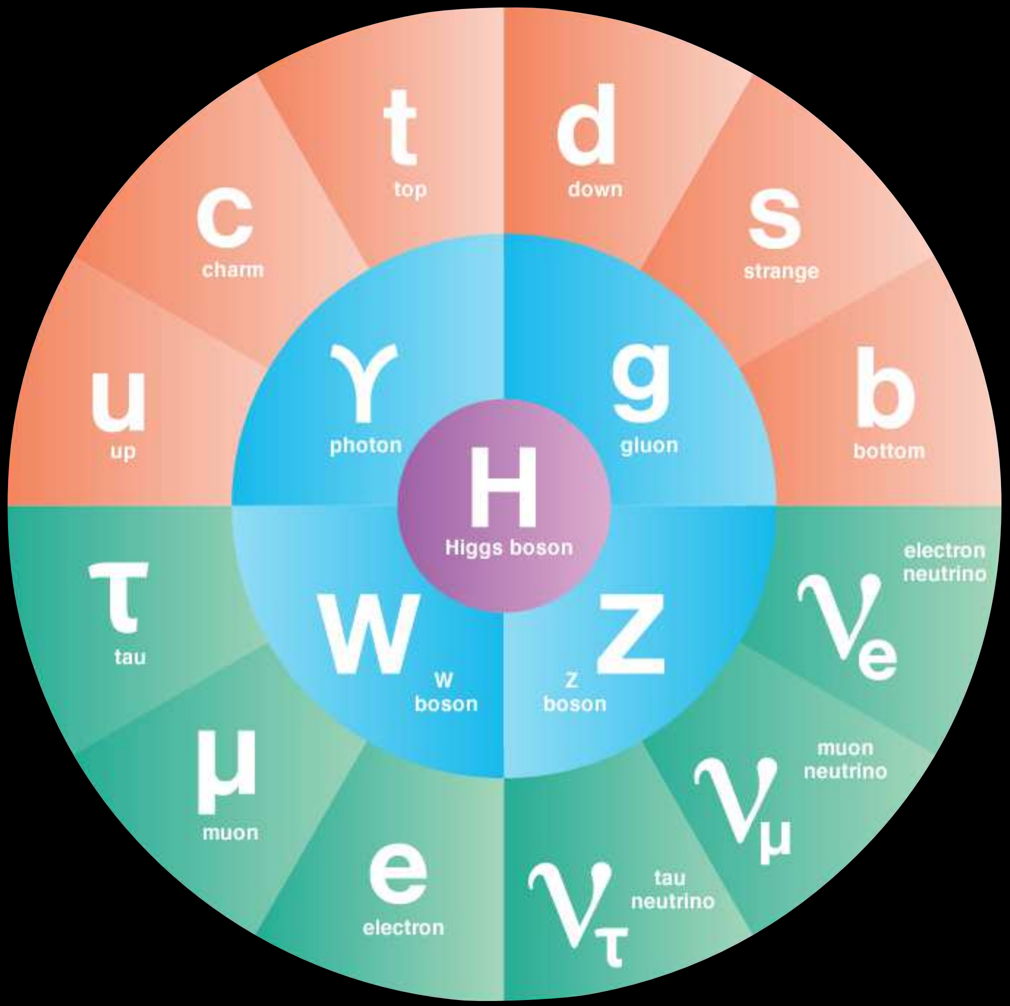
(and vice versa!)

Pillars of Fundamental Physics

Big Bang Cosmology



Standard Model



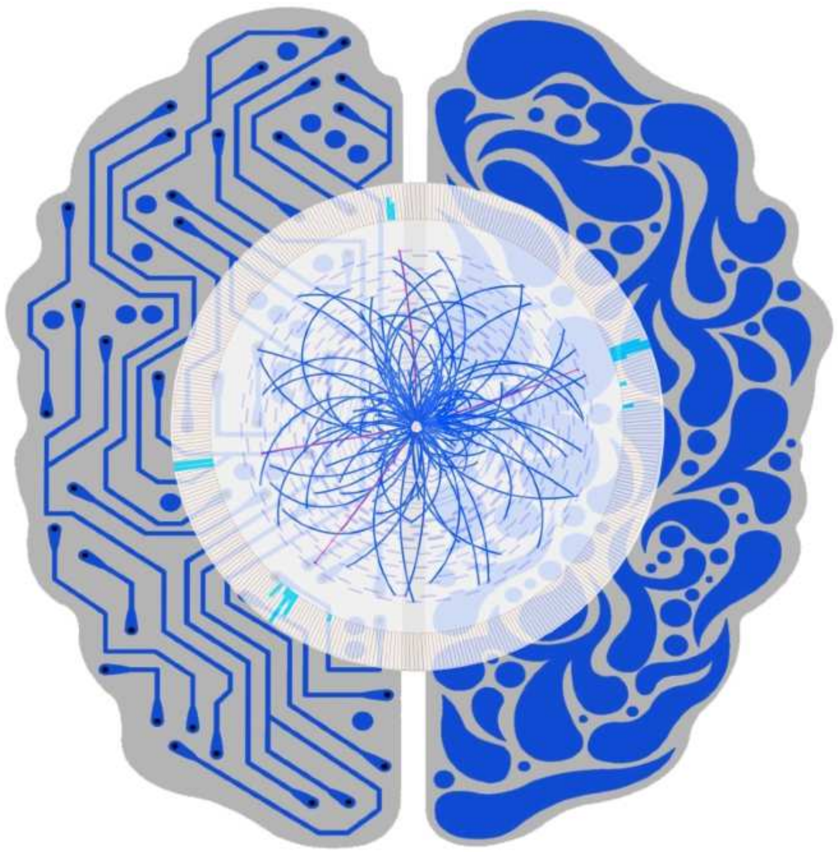
Pillars of Fundamental Physics

Big Bang Cosmology

Standard Model

Triumphs of Human Intelligence

My research focus



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?

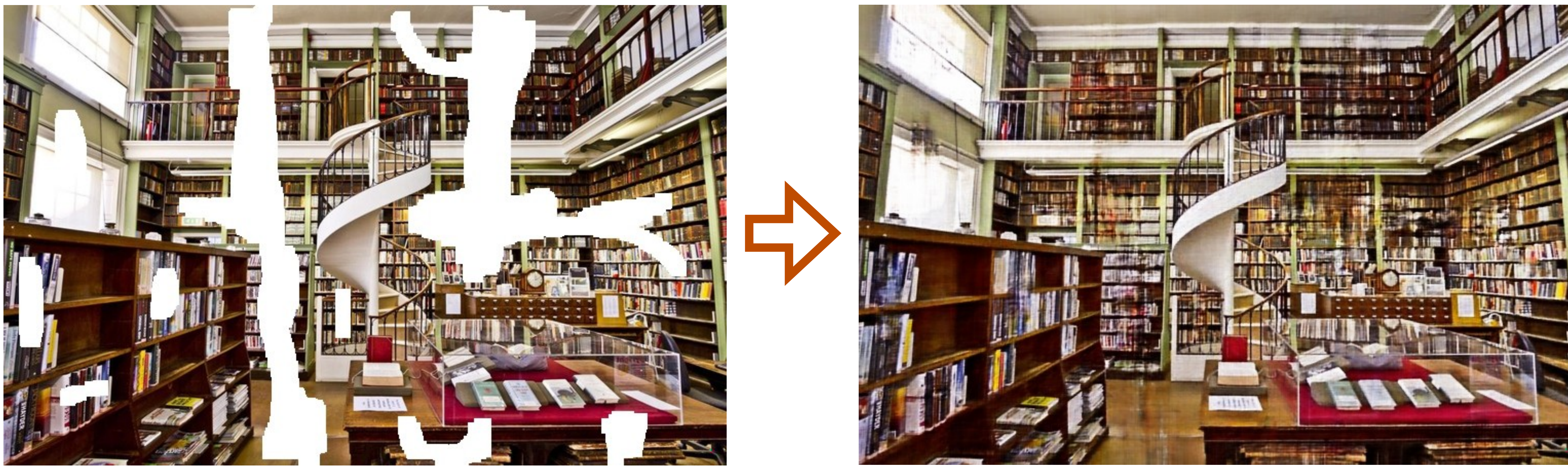
AI^2: Ab Initio Artificial Intelligence

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

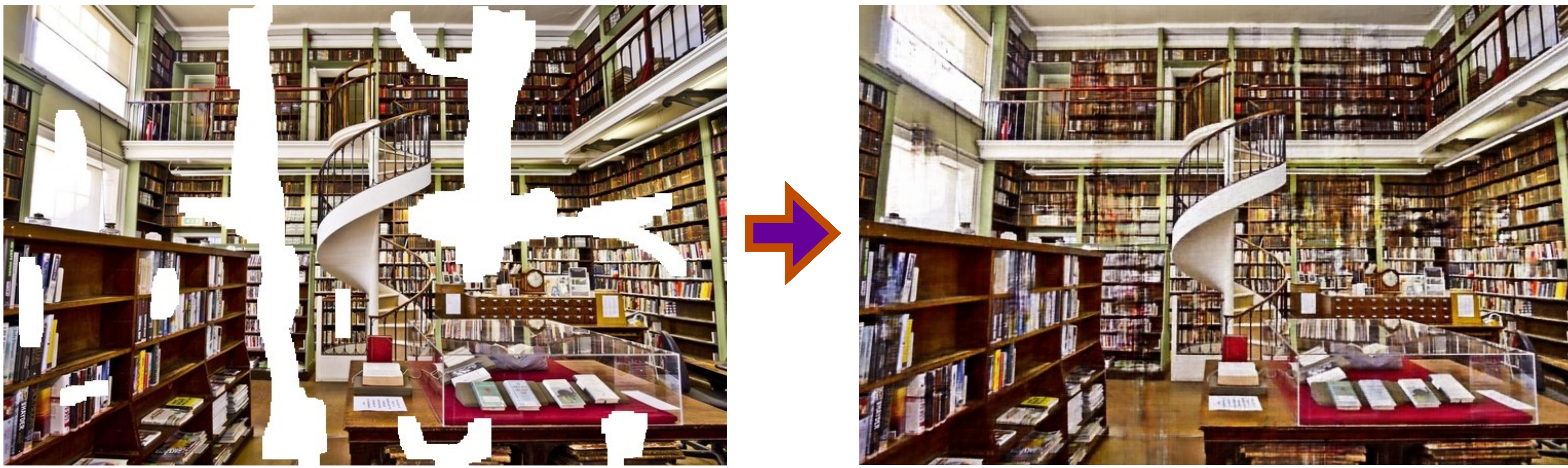


increased computational power and large data sets

[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

Deep Learning meets Deep Thinking

E.g. Inpainting



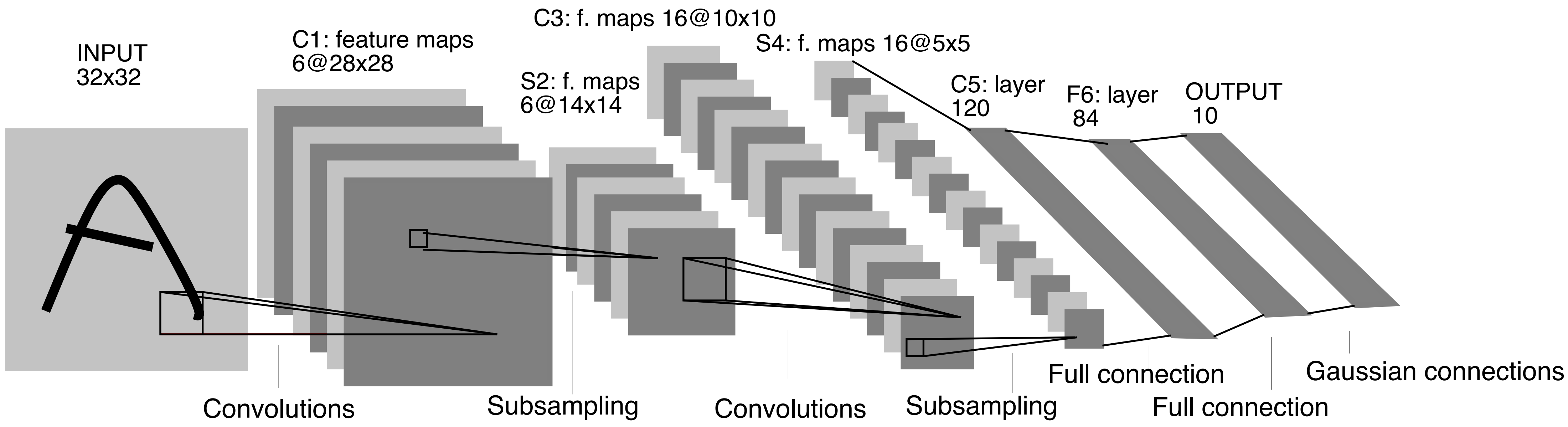
Using randomly initialized neural network (!)

Progress made by understanding the structure of problems

(not just increased computational power and large data sets)

[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

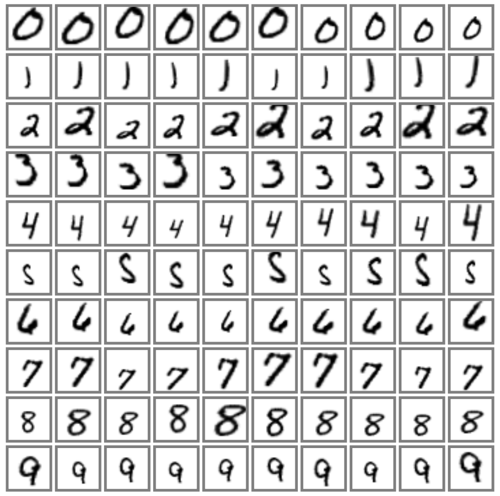
Symmetries of Convolutional NNs



Explores different length scales

Enforces Translation Equivariance

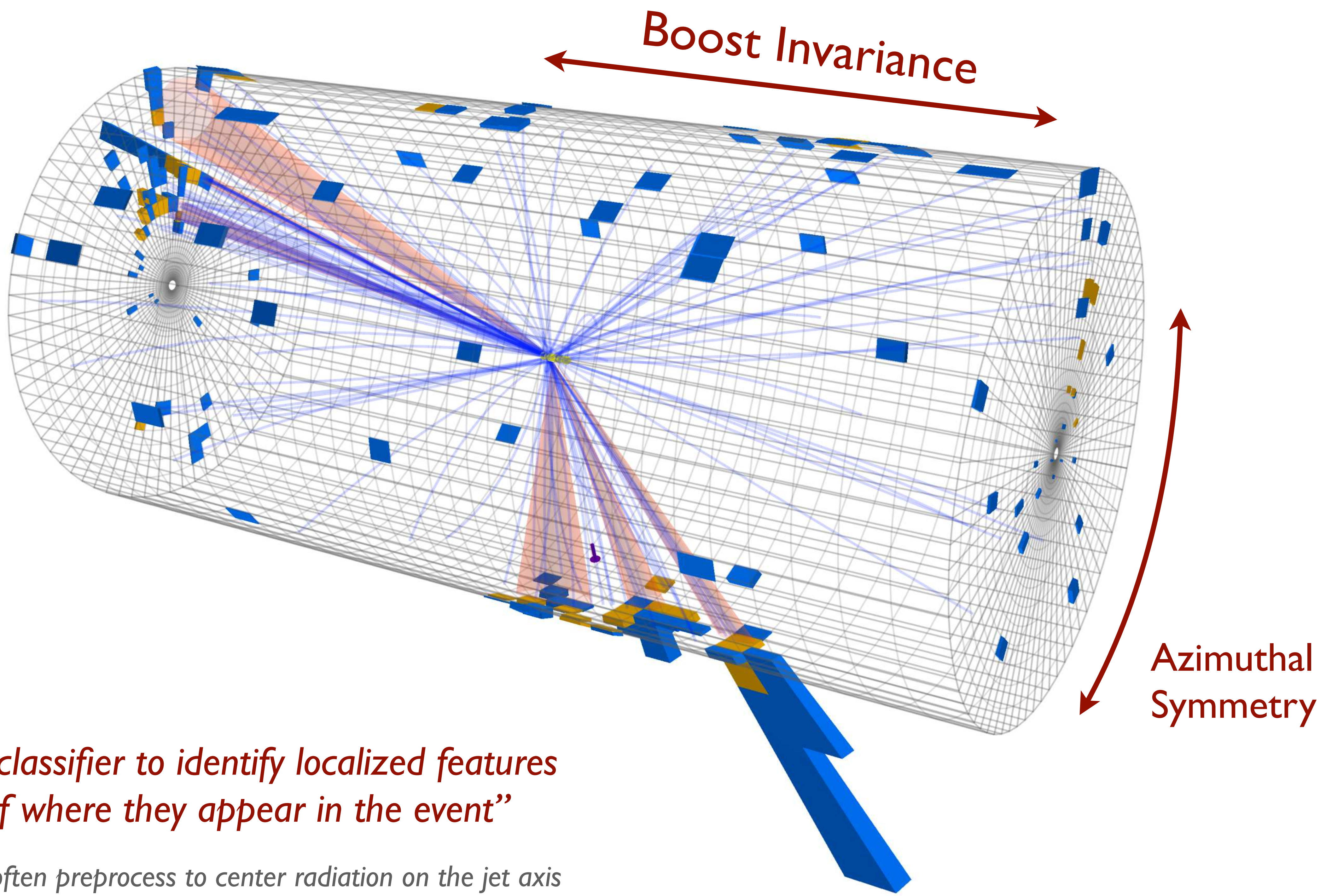
= Covariance for a Physicist



"I want my classifier to identify localized features regardless of where they appear in the image"

[image from LeCun, Bottou, Bengio, Haffner, 1998]

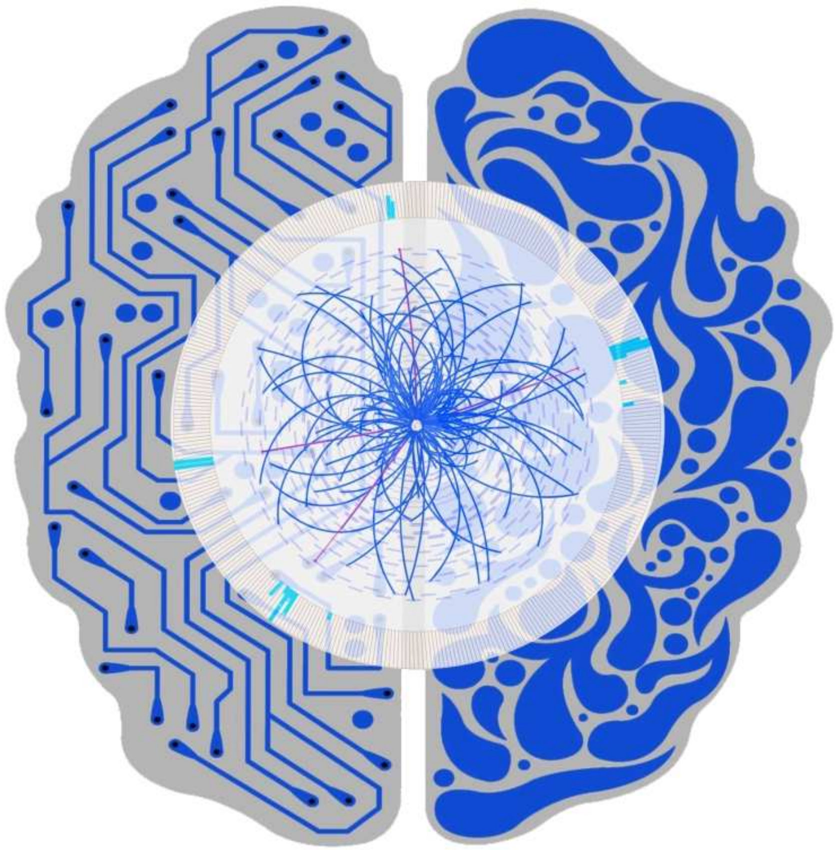
Symmetries of Collision Events



"I want my classifier to identify localized features regardless of where they appear in the event"

In practice, we often preprocess to center radiation on the jet axis

[image from CMS, 2015]



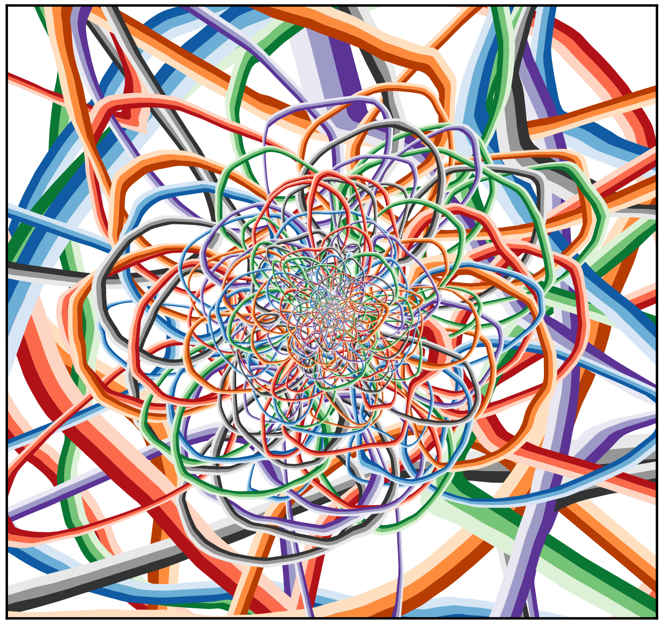
Can we encode known structures of particle physics data into a neural network architecture?

AI²: Ab Initio Artificial Intelligence

Convolutional Neural Networks ⇔ Translational Equivariance

⇒ Momentum Conservation

Energy Flow Networks ⇔ Identical Particles (QM)
Infrared/Collinear Safety (QFT)



[Komiske, Metodiev, JDT, JHEP 2019]

AI — Powerful strategy to analyze LHC collisions

× AI — Efficient neural network for point clouds

= AI² — Cross-cutting research across disciplines

(

Likelihood Ratio Trick

Key example of simulation-based inference

Many particle physics 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: $$\arg\min_{f(x)} L = \frac{p(x)}{q(x)}$$ *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 particle physics 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: $$\frac{\partial \mathcal{L}}{\partial f} = 0$$

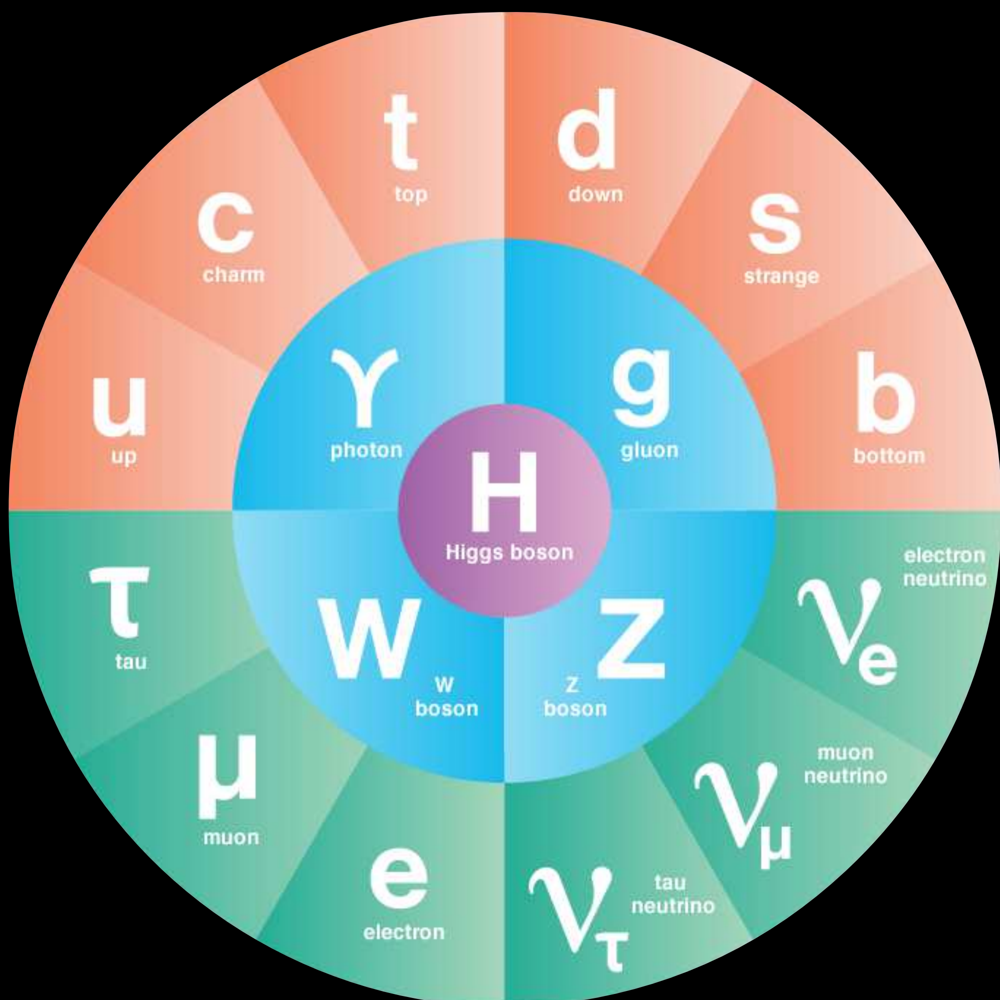
Solution: $$f(x) = \frac{p(x)}{q(x)}$$

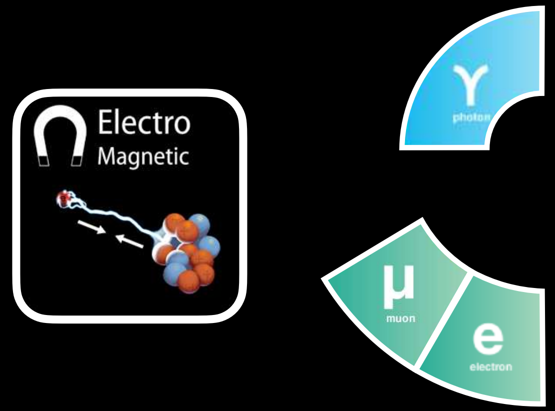
Requires shift in theoretical focus from solving problems to specifying problems

[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]

)

Particle Physics 101





Quarks

&

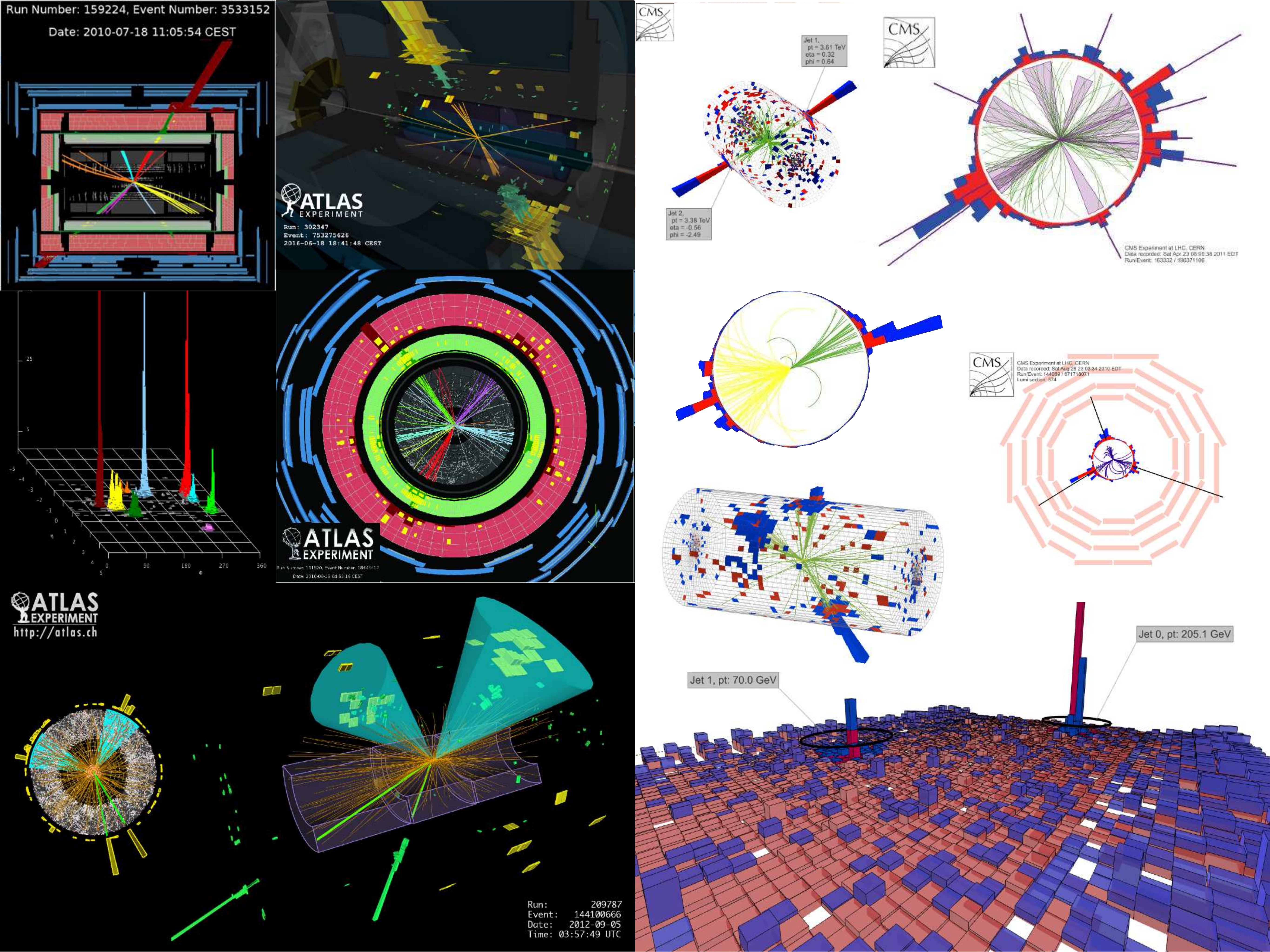
Gluons

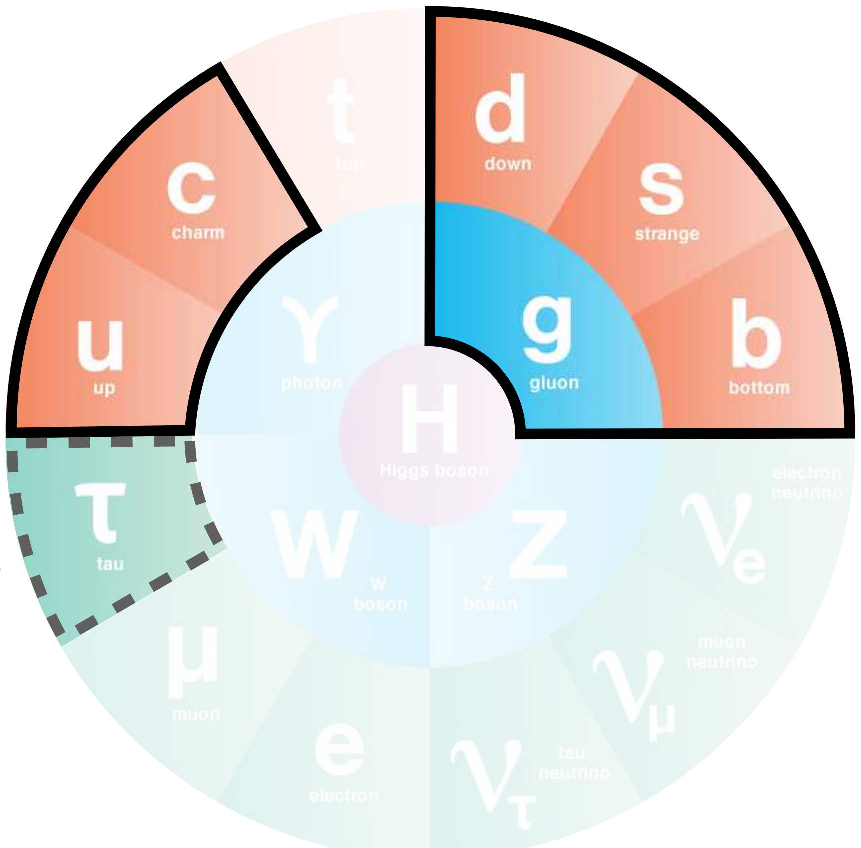
elementary

γ	photon
$e^\pm$	electron
$\mu^\pm$	muon

composite

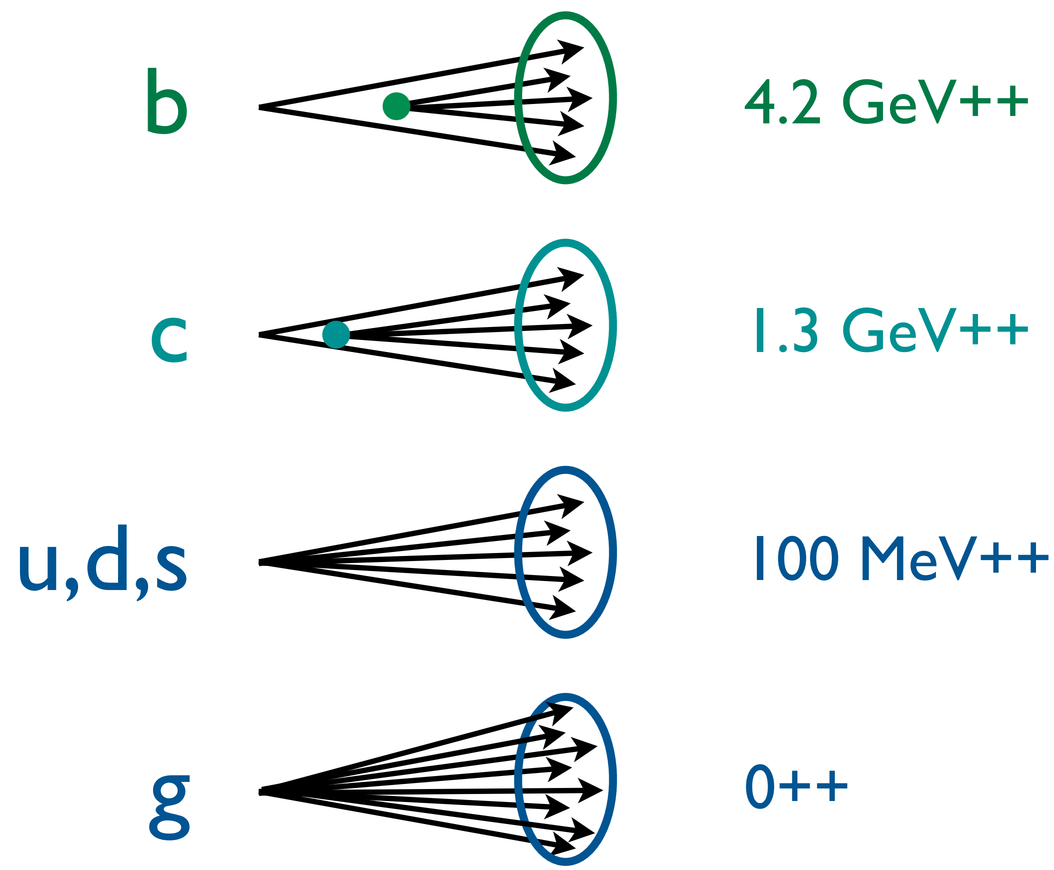
$\pi^\pm$	pion
$K^\pm$	kaon
K_L^0	K-long
$p/\overline{p}$	proton
$n/\overline{n}$	neutron


Run Number: 159224, Event Number: 3533152
Date: 2010-07-18 11:05:54 CEST
ATLAS EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST
CMS
Jet 1,
pt = 3.61 TeV
eta = 0.32
phi = 0.64
Jet 2,
pt = 3.38 TeV
eta = -0.56
phi = -2.49
CMS
ATLAS EXPERIMENT
CMS
ATLAS EXPERIMENT
http://atlas.ch
Run: 209787
Event: 144100666
Date: 2012-09-05
Time: 03:57:49 UTC
Jet 1, pt: 70.0 GeV
Jet 0, pt: 205.1 GeV

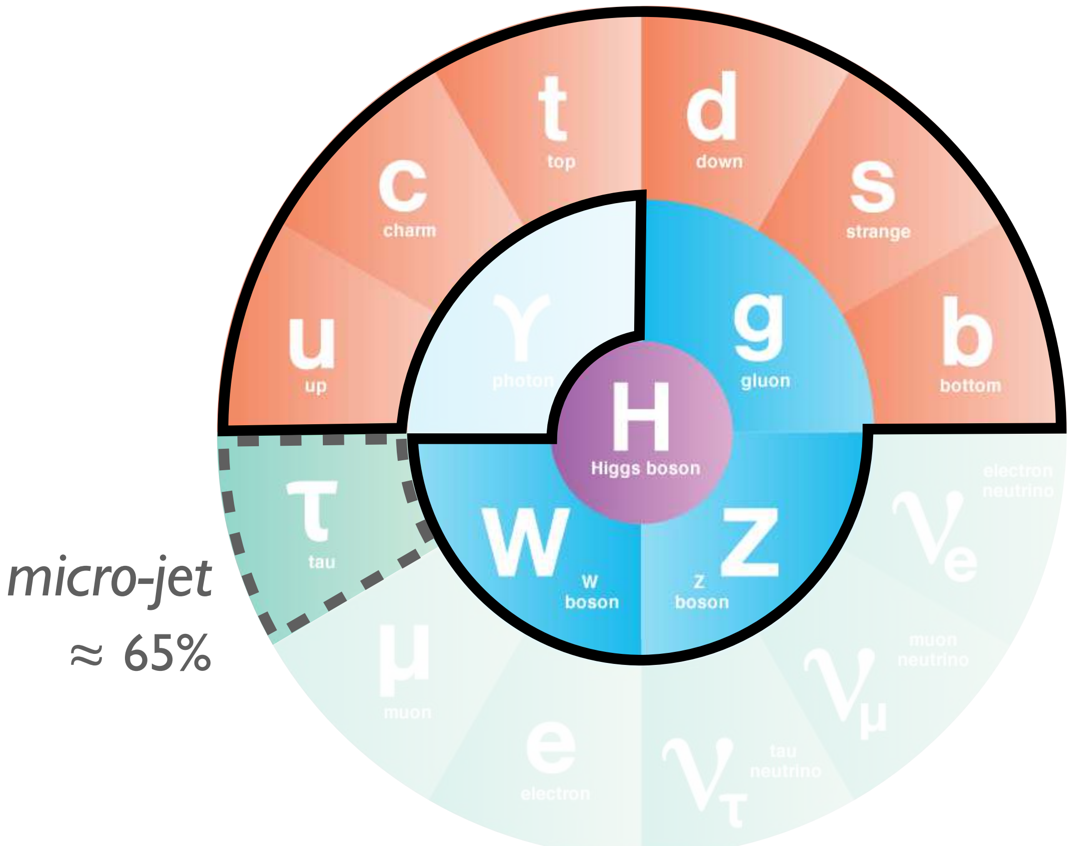


micro-jet
≈ 65%

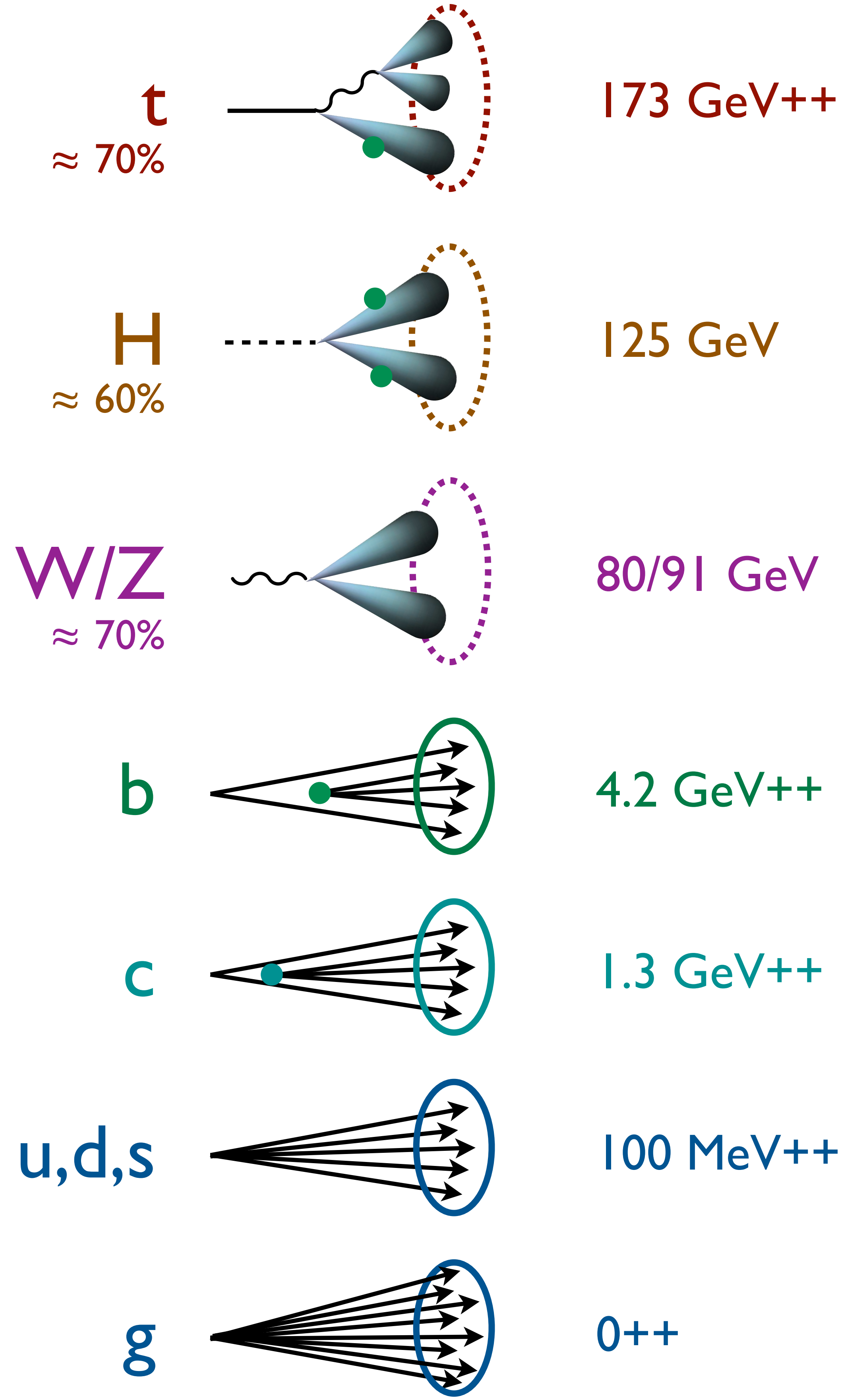
Jets from the Standard Model



++ = Mass from QCD Radiation



Jets from the Standard Model



++ = Mass from QCD Radiation

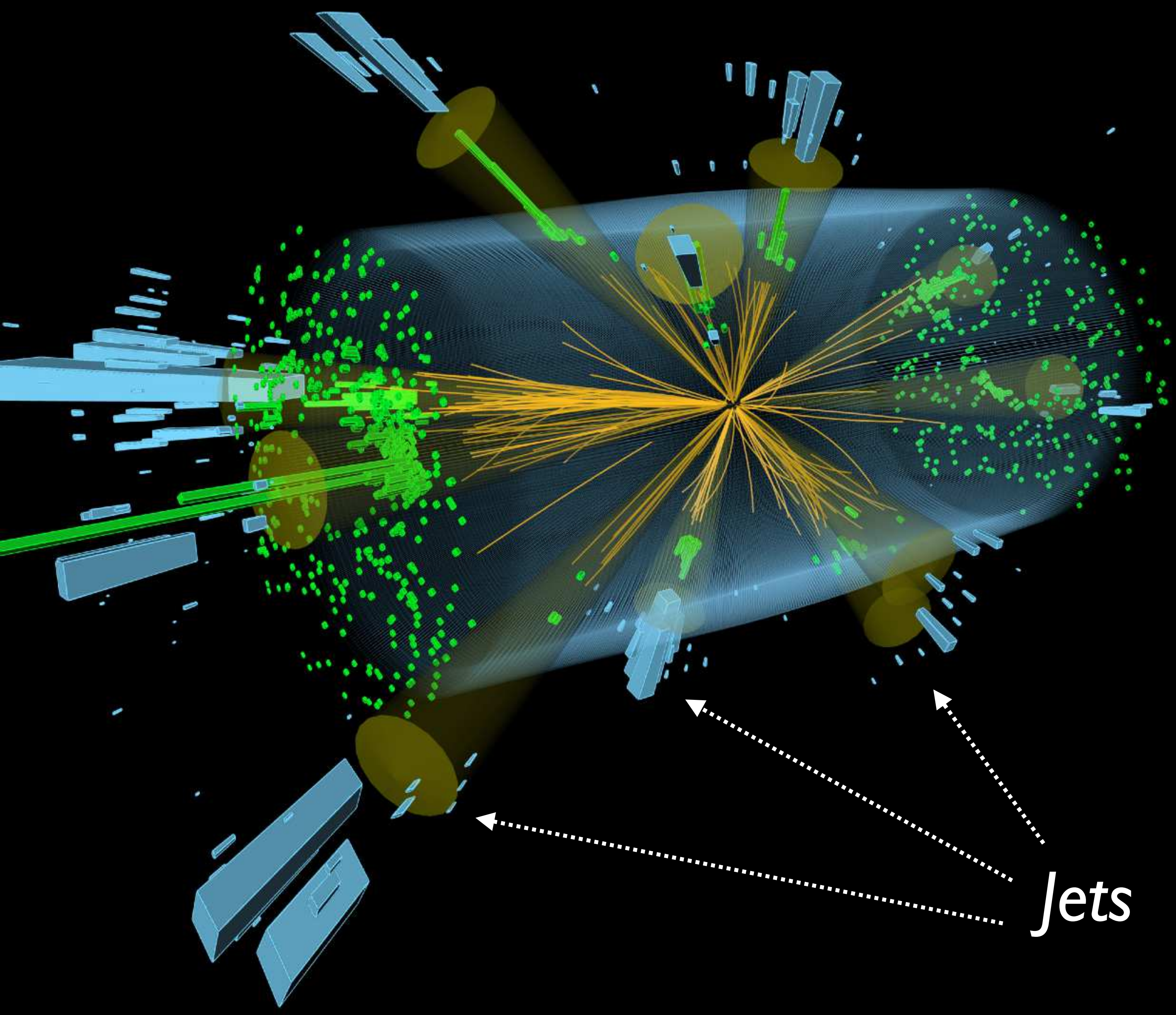
T	E	H	M		
	●			γ	photon
●	●			$e^{\pm}$	electron
●	•	•	●	$\mu^{\pm}$	muon
●	•	●		$\pi^{\pm}$	pion
●	•	●		$K^{\pm}$	kaon
	•	●		K^0_L	K-long
●	•	●		$p/\overline{p}$	proton
	•	●		$n/\overline{n}$	neutron

elementary

composite

Collider Event

Collection of points in (momentum) space



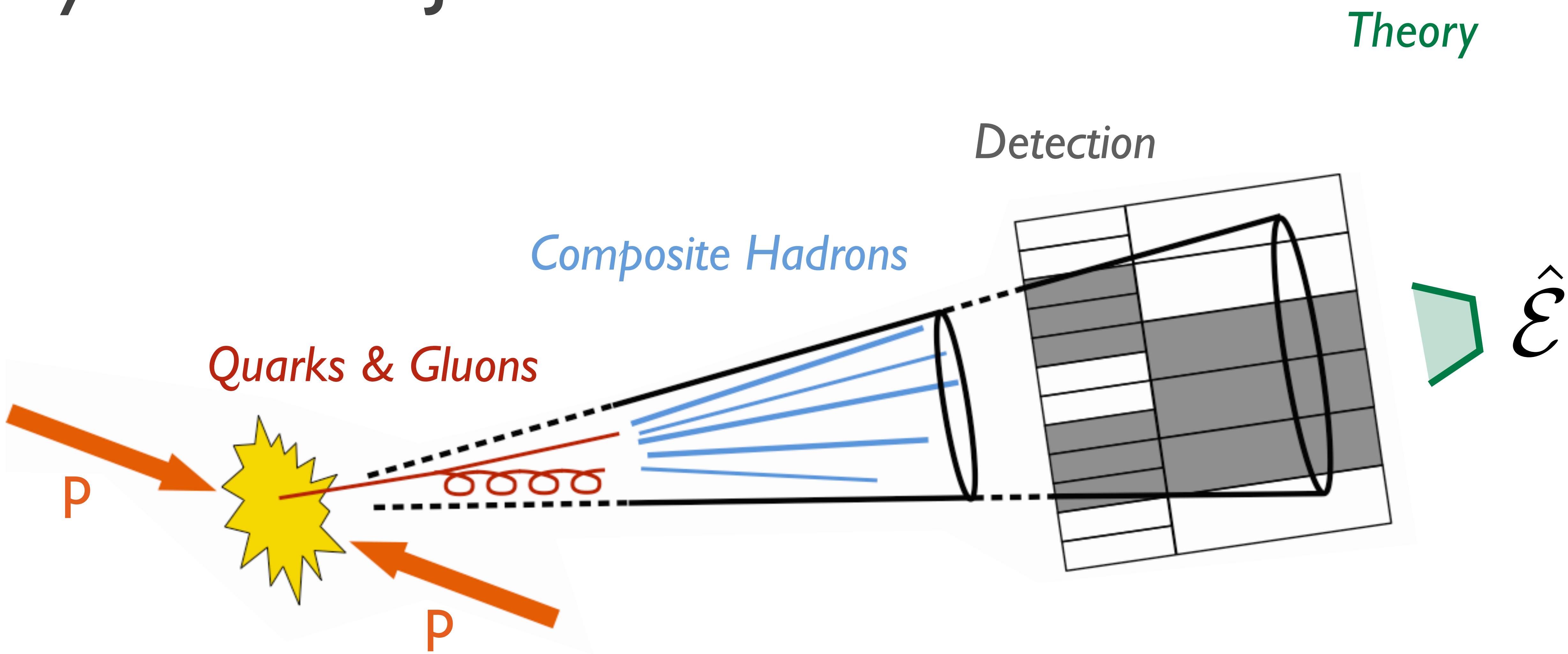
T	E	H	M			
	●			γ	photon	elementary
●	●			$e^{\pm}$	electron	elementary
●	●	●	●	$\mu^{\pm}$	muon	elementary
●	●	●		$\pi^{\pm}$	pion	composite
●	●	●		$K^{\pm}$	kaon	composite
	●	●		K_L^0	K-long	composite
●	●	●		$p/\overline{p}$	proton	composite
	●	●		$n/\overline{n}$	neutron	composite

Point Cloud

Collection of points in (position) space

[Popular Science, 2013]

Dynamics of Jet Formation



Energy Flow:

Robust to hadronization and detector effects

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

[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]

Patrick Komiske

Eric Metodiev

Principles of Fundamental Physics

Robustness of Energy Flow

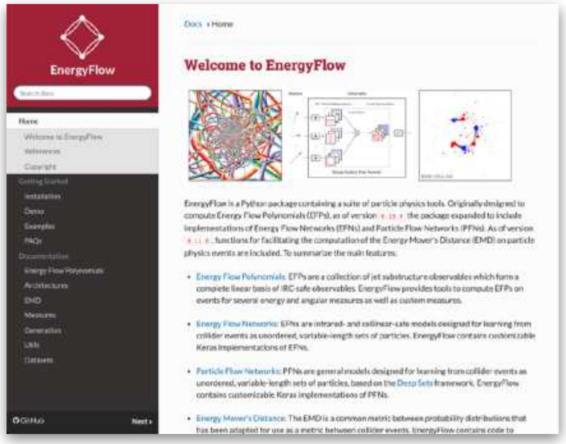
[Komiske, Metodiev, JDT, JHEP 2018]

?

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]

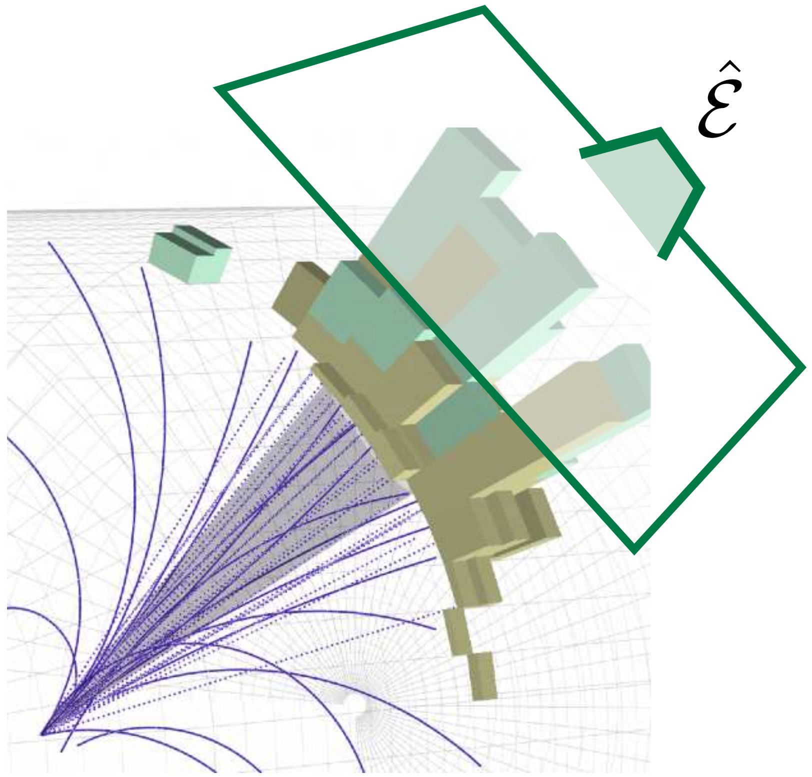
Jets as Weighted Point Clouds

- Energy-Weighted Directions

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

Energy Direction

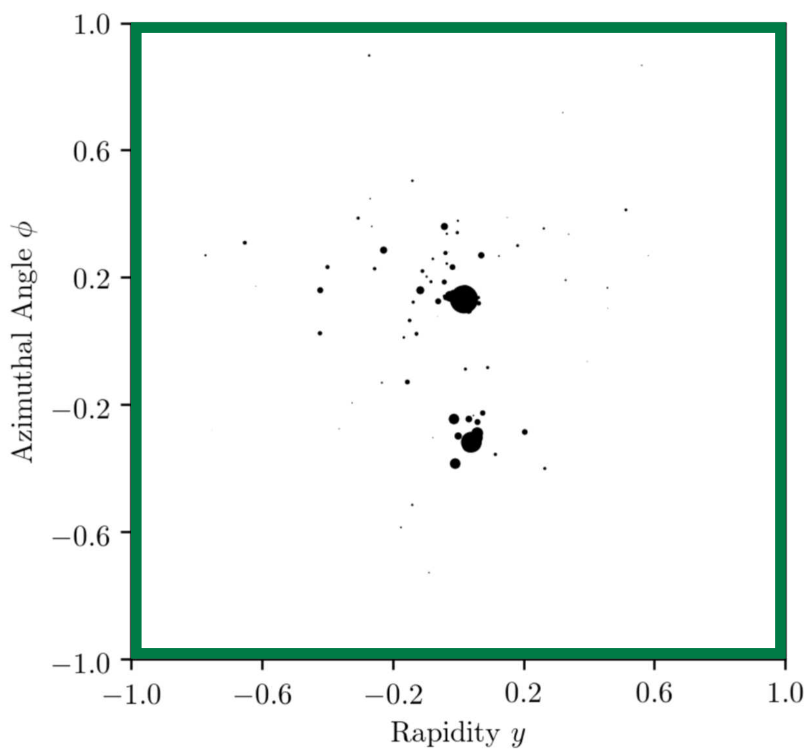
(suppressing "unsafe" charge/flavor information)



- Equivalently: Energy Density

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

Energy Direction



“Thinking” Like a Physicist

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

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

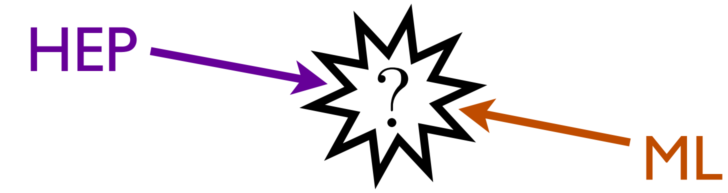
The Machine

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)

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

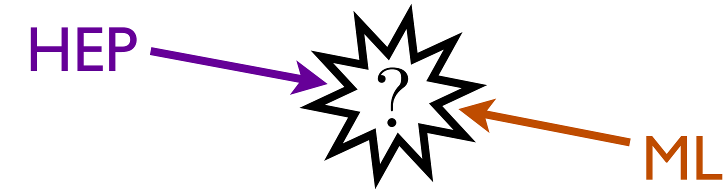
Parametrized with Neural Networks

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]

Energy Flow Networks

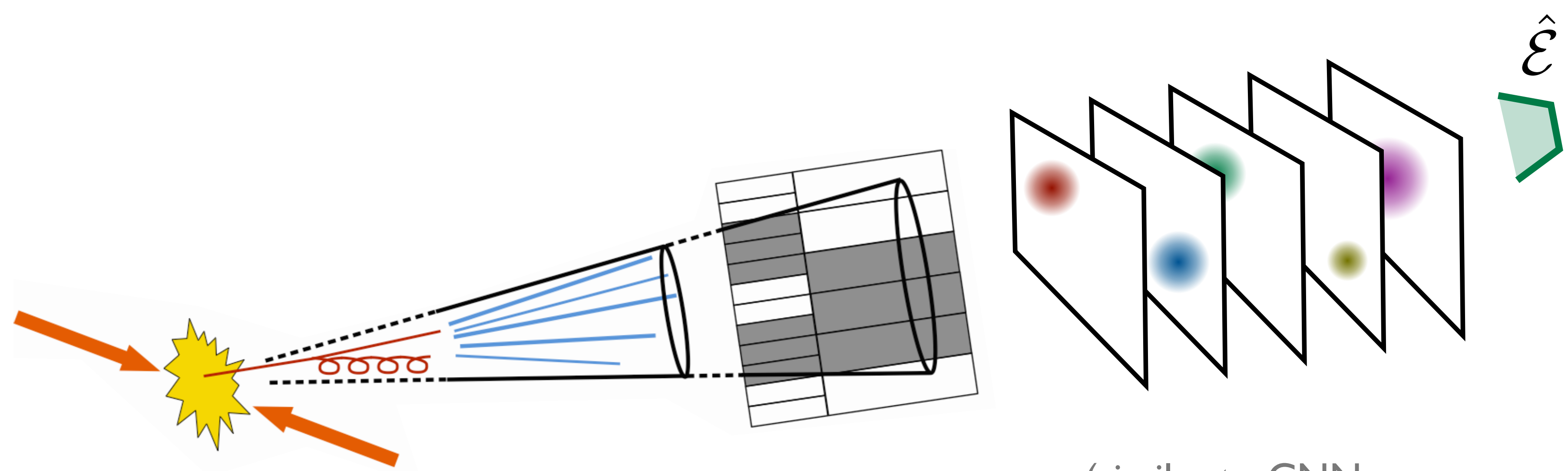
*Architecture designed around symmetries and **interpretability***



Latent space of dim ℓ

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

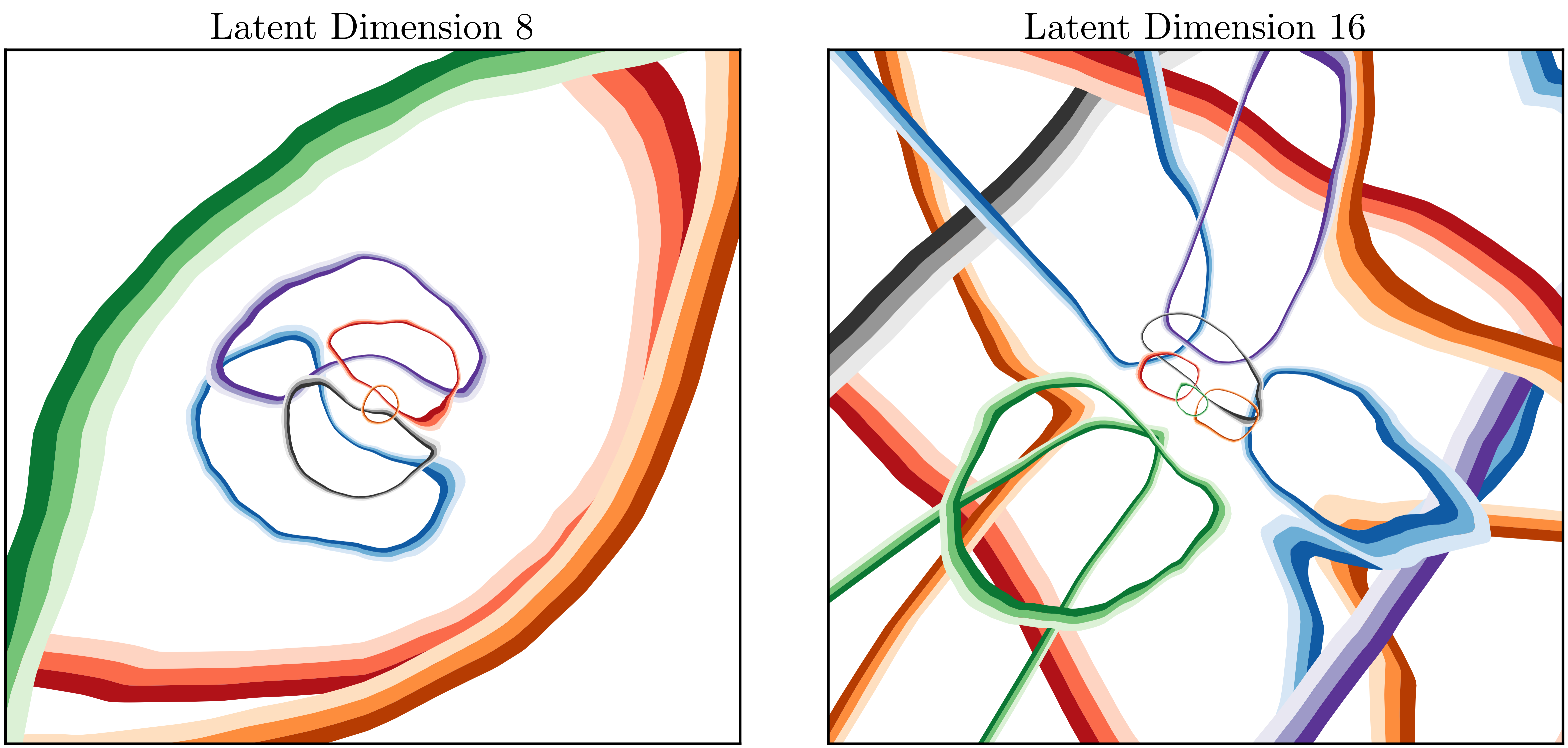
Easy to plot!



(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]

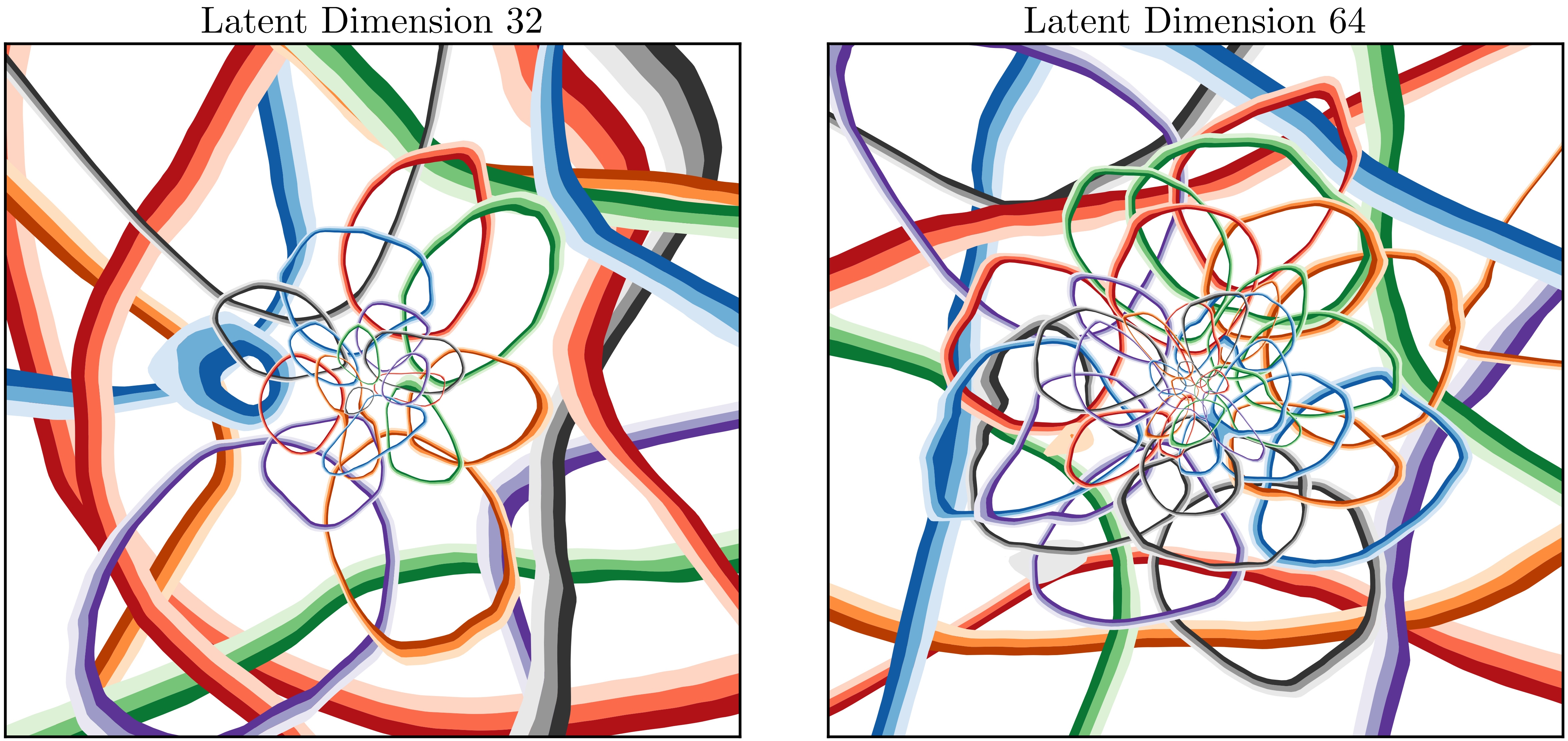
Psychedelic Network Visualization



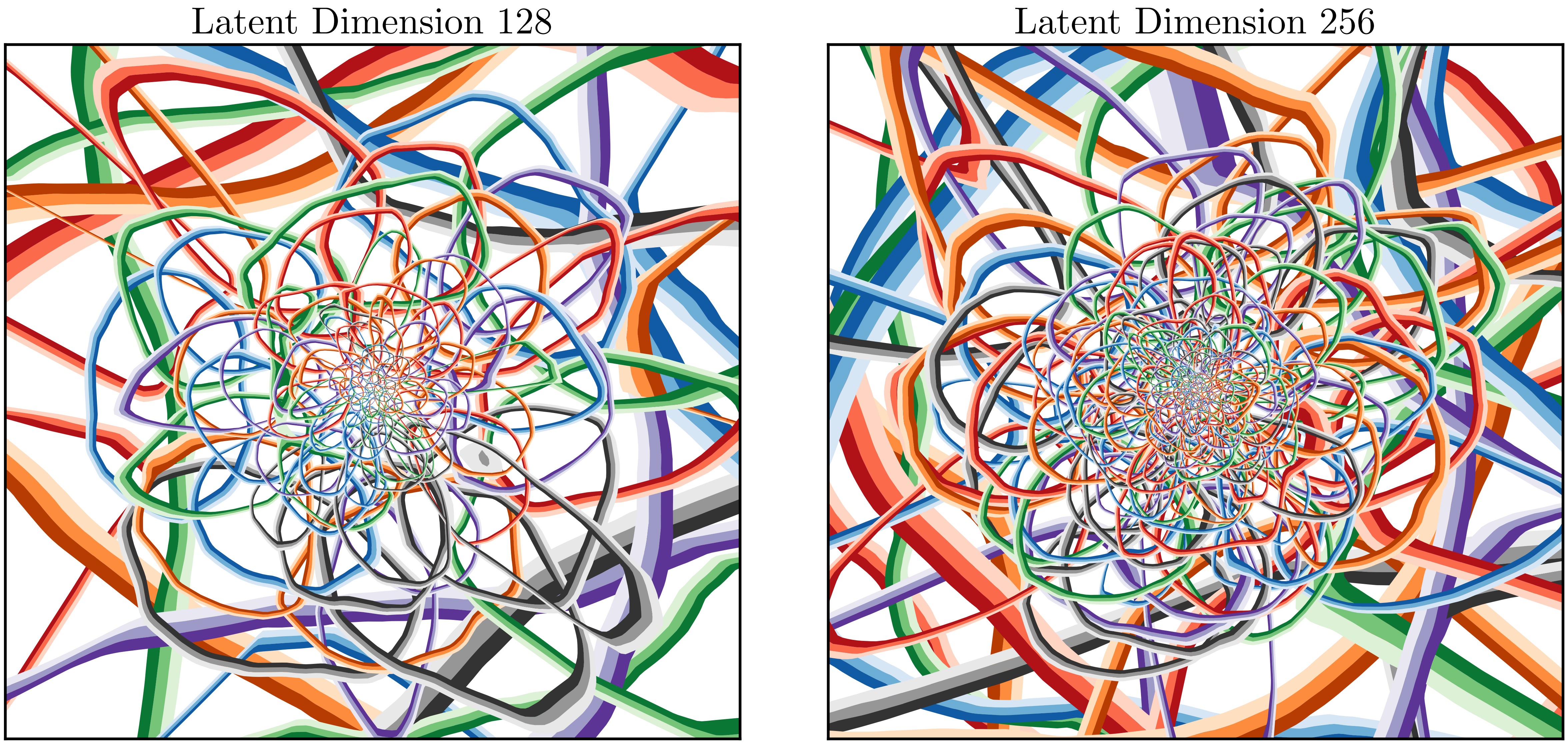
"Hello, World!" of Jets: Quark 

vs. Gluon 

Psychedelic Network Visualization

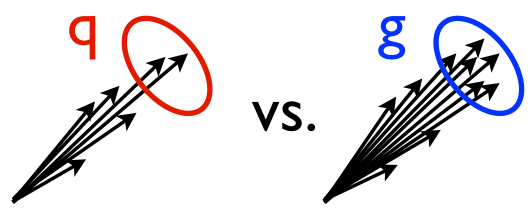


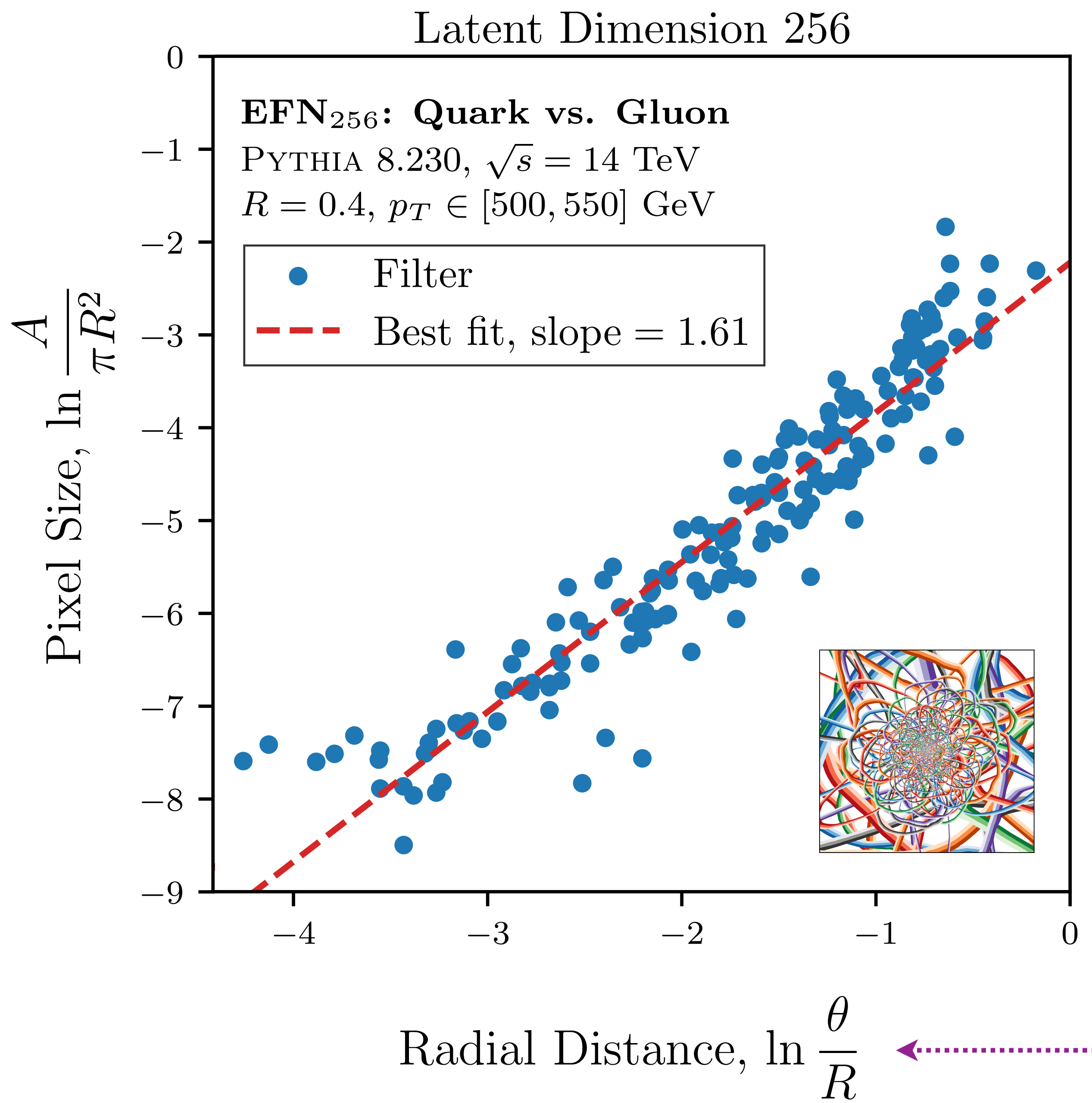
Psychedelic Network Visualization

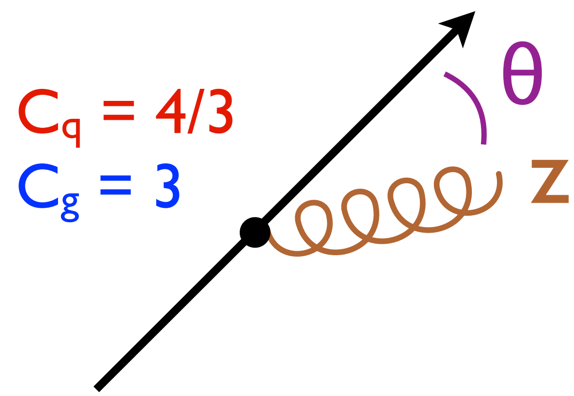


Fractal structure of the strong force!

Scaling of Strong Interactions



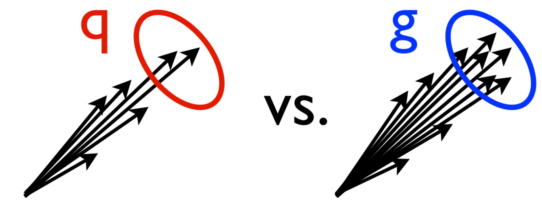


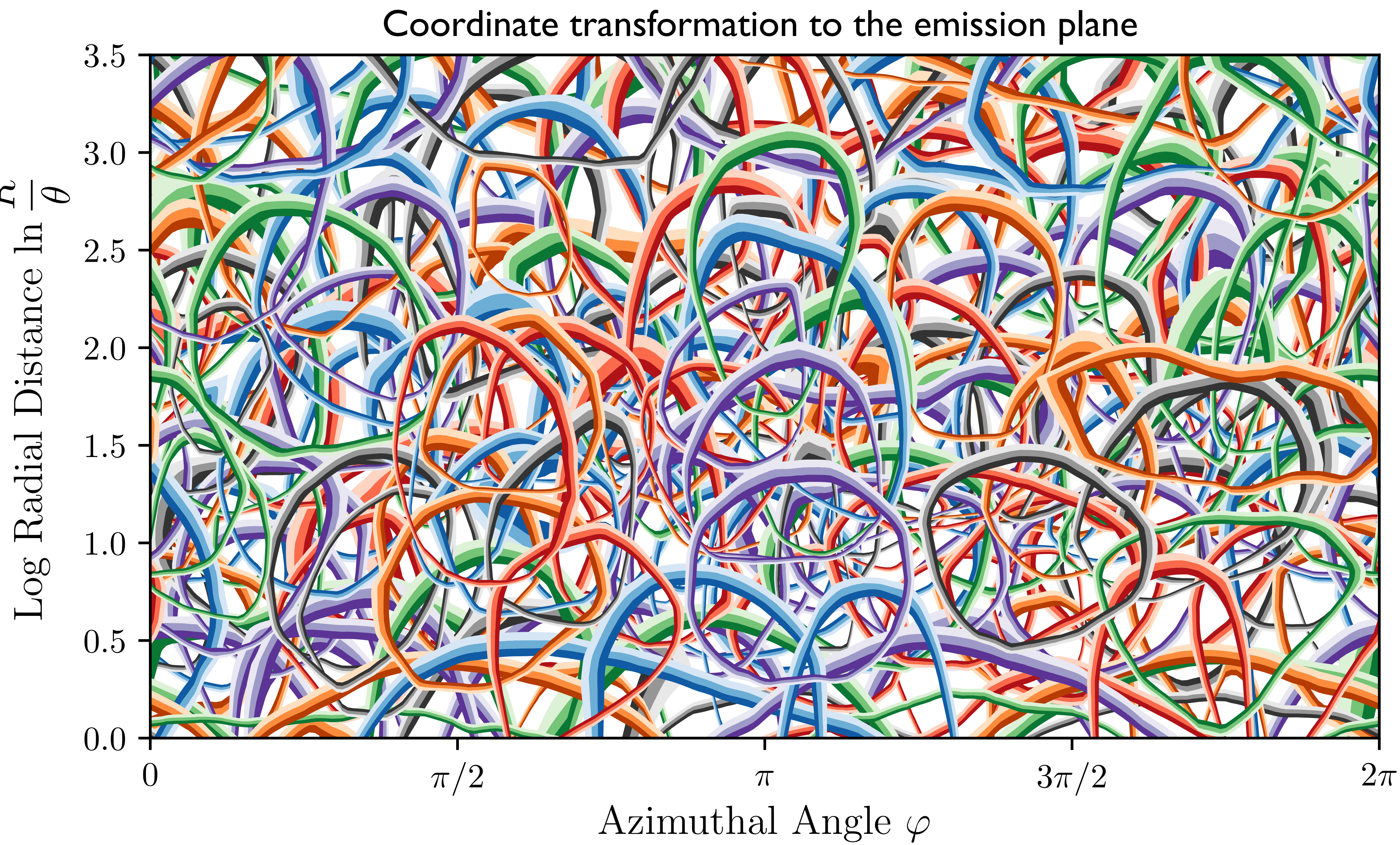


$$dP_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \underbrace{\frac{d\theta}{\theta}}_{\text{Collinear}} \underbrace{\frac{dz}{z}}_{\text{Soft}}$$

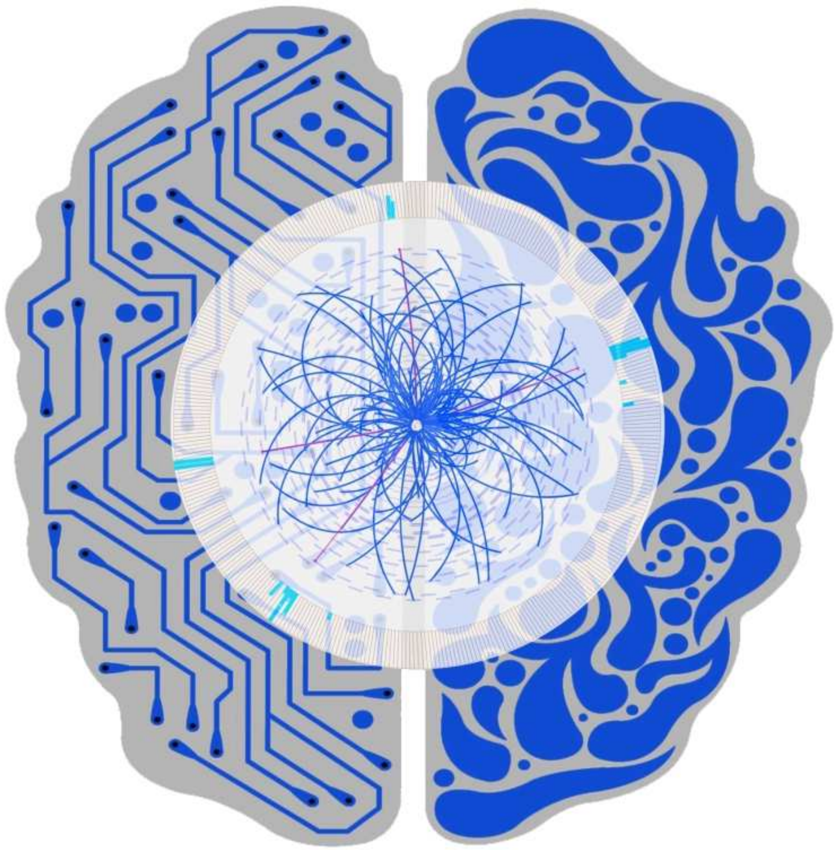
[Komiske, Metodiev, JDT, JHEP 2019]

Ready for the ICA?

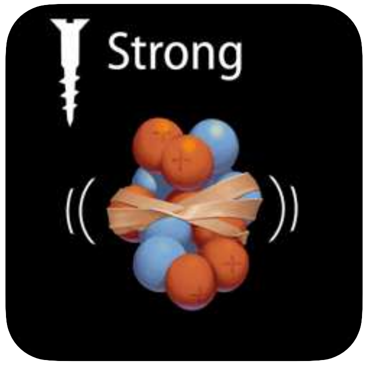




[Komiske, Metodiev, JDT, JHEP 2019; see also Dreyer, Salam, Soyez, JHEP 2018]



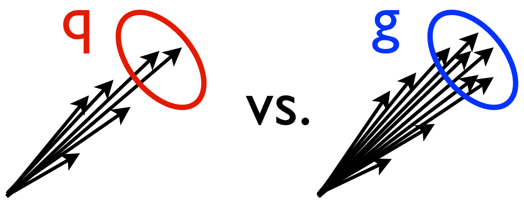
We taught a machine to "think" like a physicist…



…and it learned fractal structure of strong force!

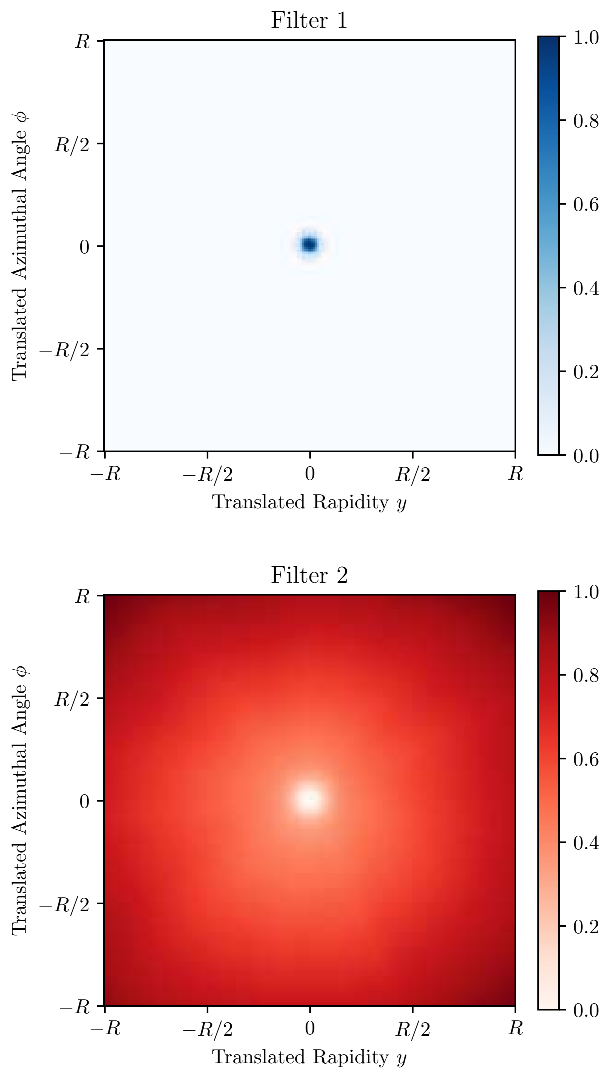
"Ok, but did you really learn something you didn't already know?"

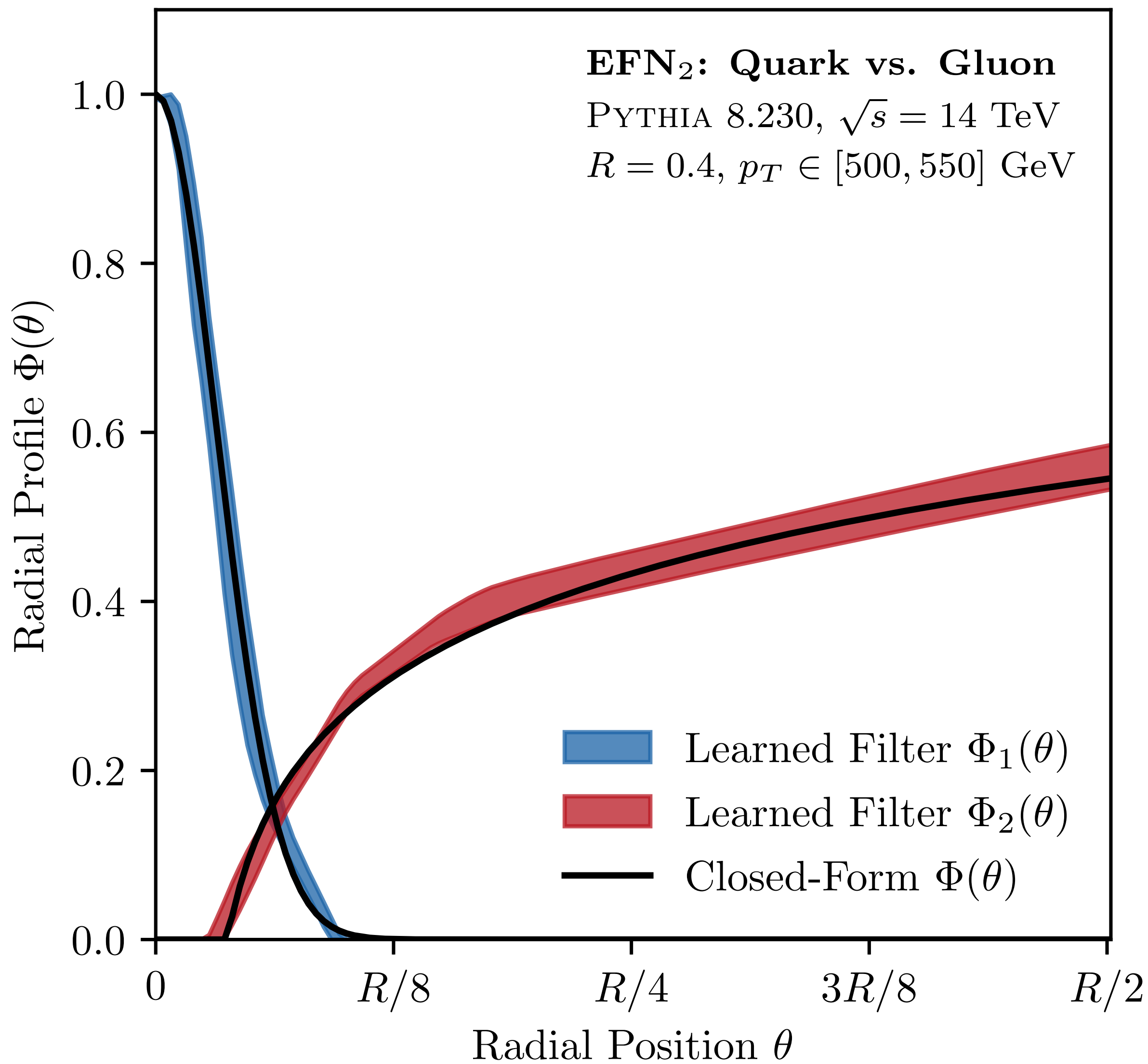
Learning from the Machine



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

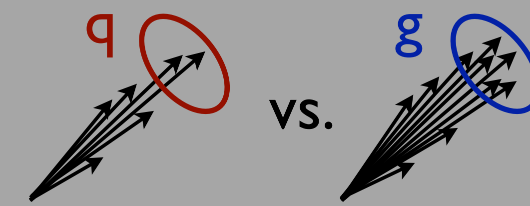
cf. Angularities:
$f(\theta) = \theta^\beta$





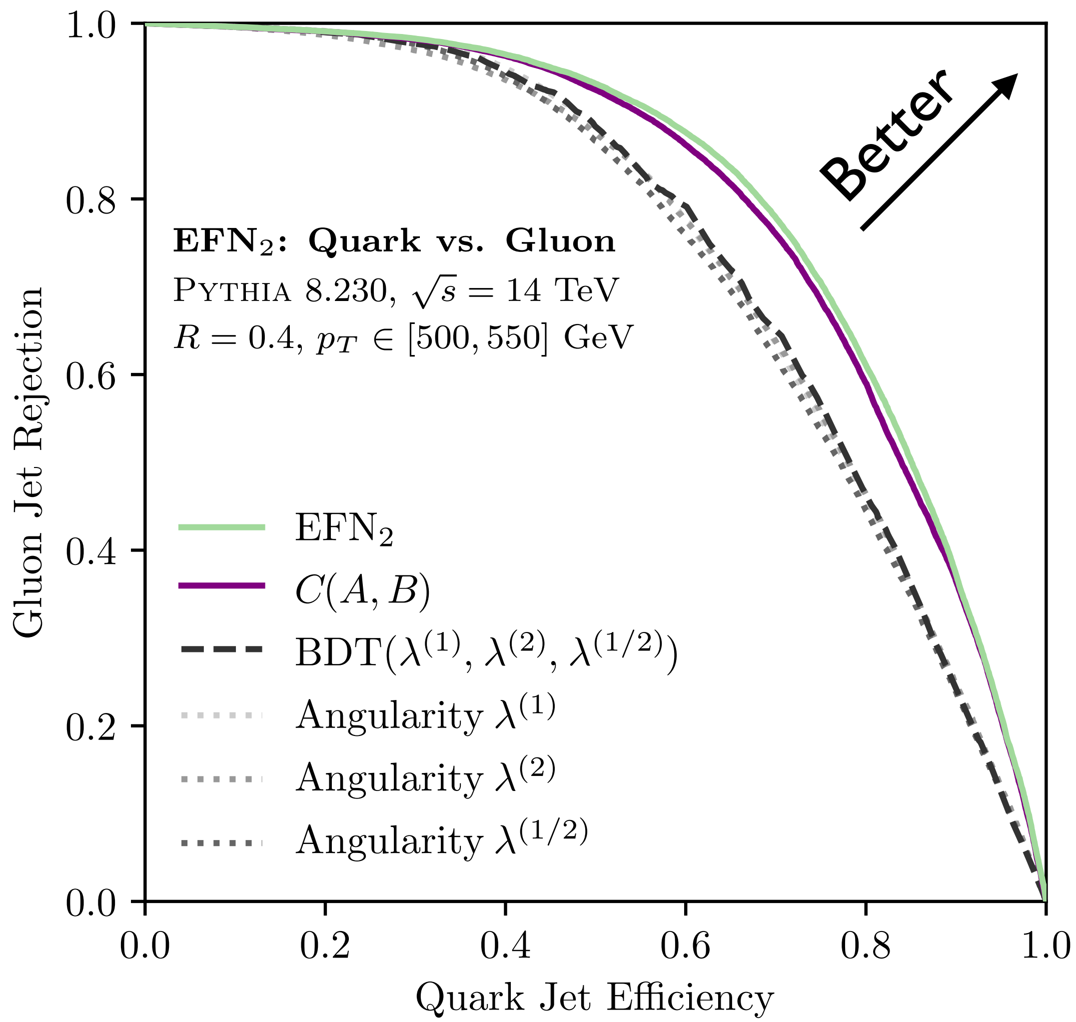
[Komiske, Metodiev, JDT, JHEP 2019;
cf. Larkoski, JDT, Waalewijn, JHEP 2014; using Berger, Kucs, Sterman, PRD 2003; Ellis, Vermilion, Walsh, Hornig, Lee, JHEP 2010]

Learning from the Machine



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

cf. Angularities: f(θ) = θ^β

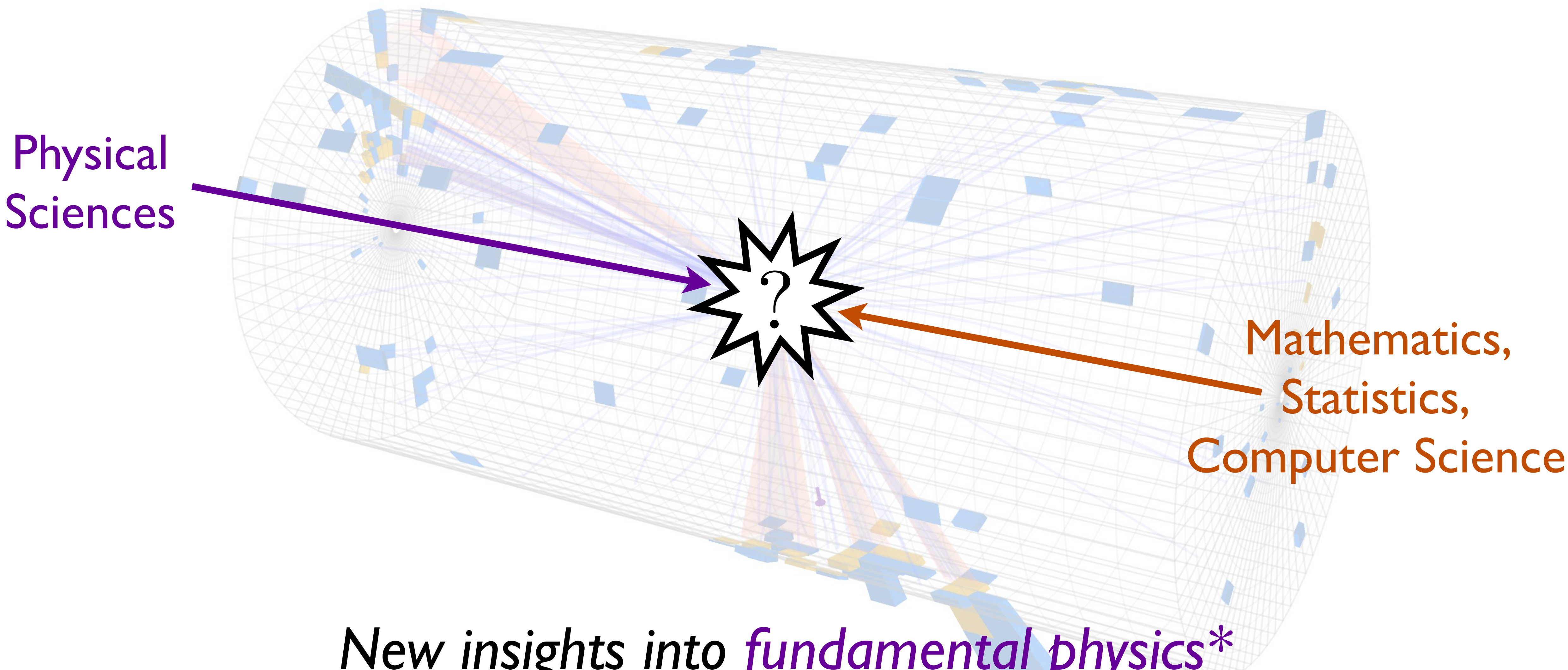


EFN outperformed a domain expert (i.e. me)

But we could reverse engineer the machine

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

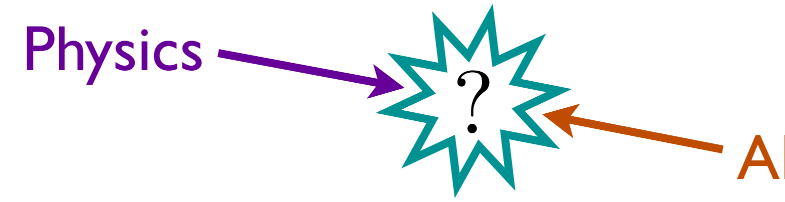
“Collision Course”



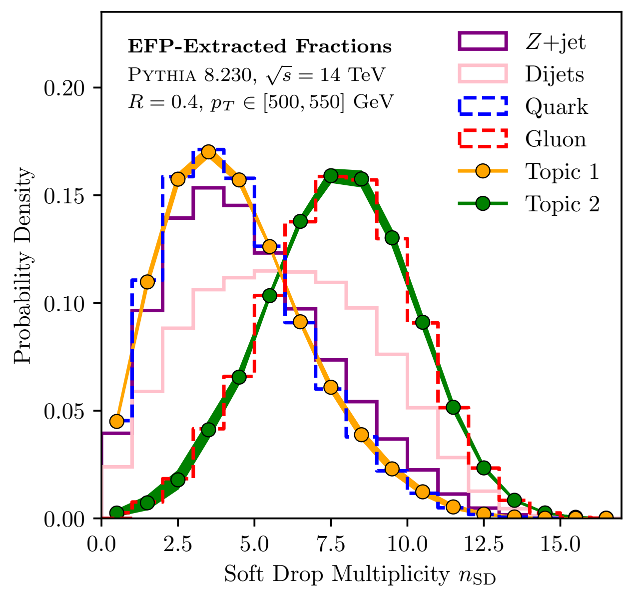
*New insights into fundamental physics**
*facilitated by advances in artificial intelligence**

(and vice versa!)

More Collisions

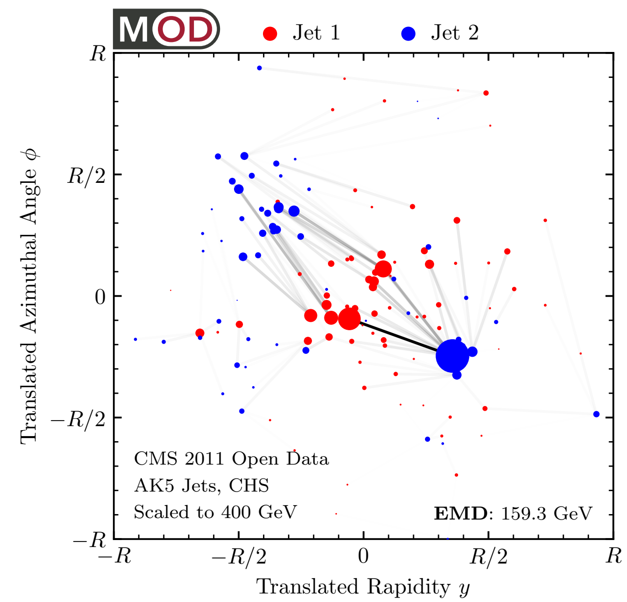


Quark/Gluon Definitions via Blind Source Separation



[Metodiev, JDT, PRL 2018;
Komiske, Metodiev, JDT, JHEP 2018]

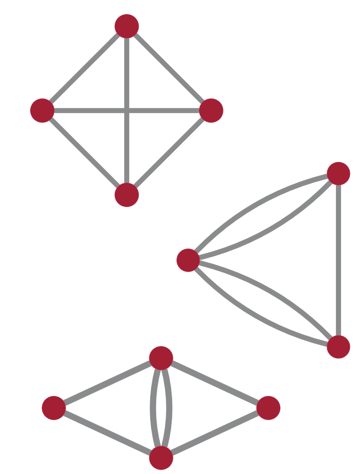
Half-Century of Collider Physics via Optimal Transport (!)



[Komiske, Metodiev, JDT, PRL 2019, JHEP 2020;
Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020]

Kinematic Decomposition via Graph Theory

	Leafless Multigraphs	
Edges d	Connected A307317	All A307316
1	0	0
2	1	1
3	2	2
4	4	5
5	9	11
6	26	34
7	68	87
8	217	279
9	**718**	**897**
10	**2 553**	**3 129**
11	**9 574**	**11 458**
12	**38 005**	**44 576**
13	**157 306**	**181 071**
14	**679 682**	**770 237**
15	**3 047 699**	**3 407 332**
16	**14 150 278**	**15 641 159**



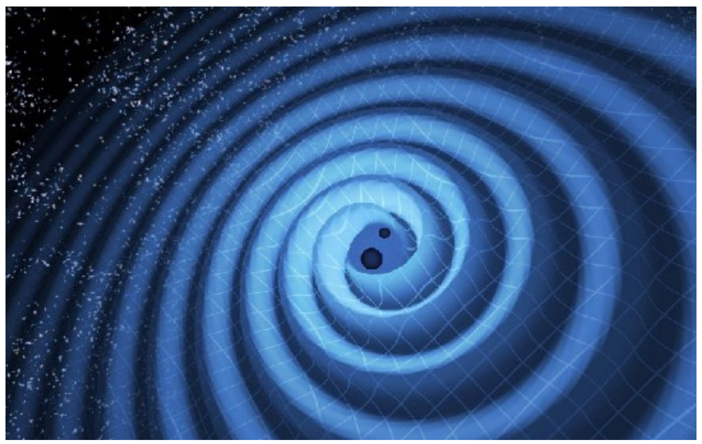
[Komiske, Metodiev, JDT,
JHEP 2018, PRD 2020]

New for 2021!

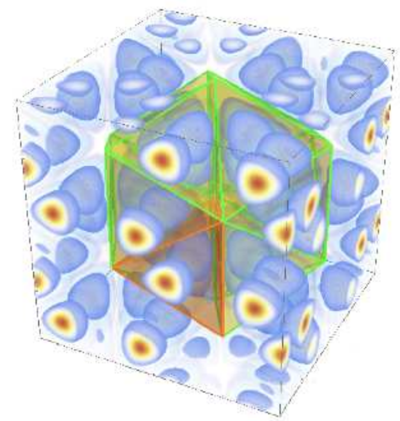
Interdisciplinary PhD in Physics, Statistics & Data Science

Artificial Intelligence ⇔ *Fundamental Interactions*

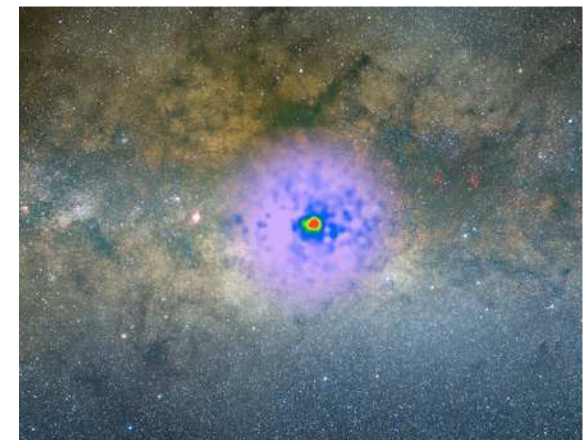
Gravitational Waves



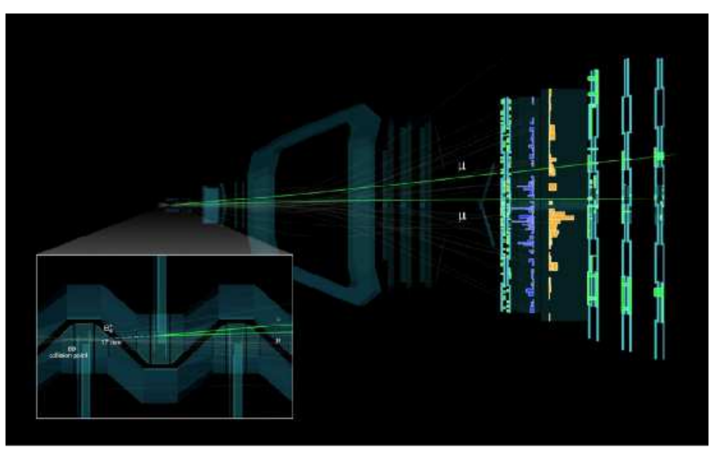
Nuclear Physics



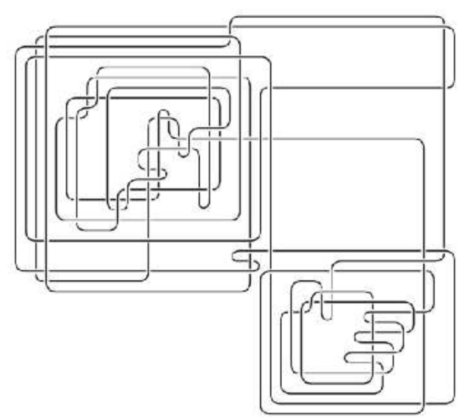
Astrophysics



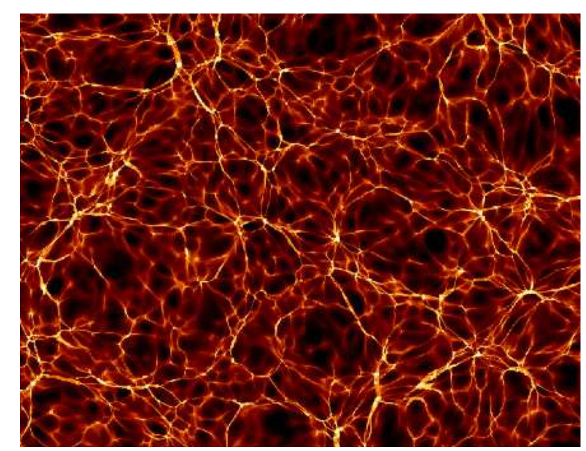
Particle Colliders



Mathematical Physics



Dark Matter



…

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/