Artificial Intelligence for Physics Discovery: Theory Perspective

Jesse Thaler

Associate Professor
Center for Theoretical Physics

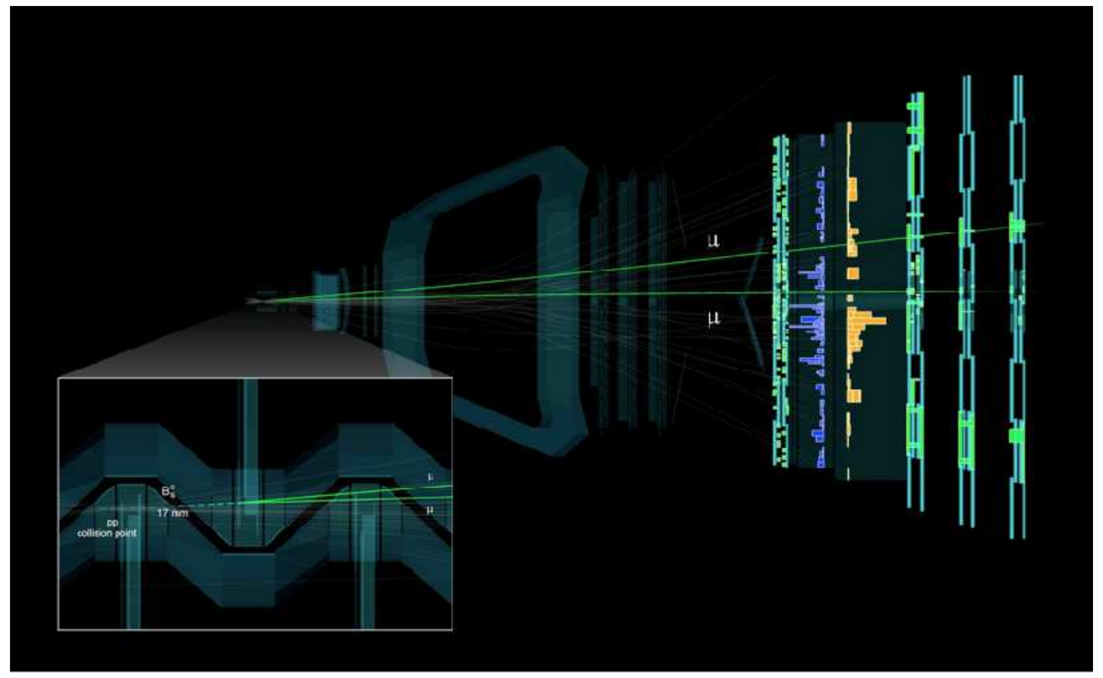
Director, NSF Institute for Artificial Intelligence and Fundamental Interactions (IAIFI)

AAAS Annual Meeting — February 8-11, 2021

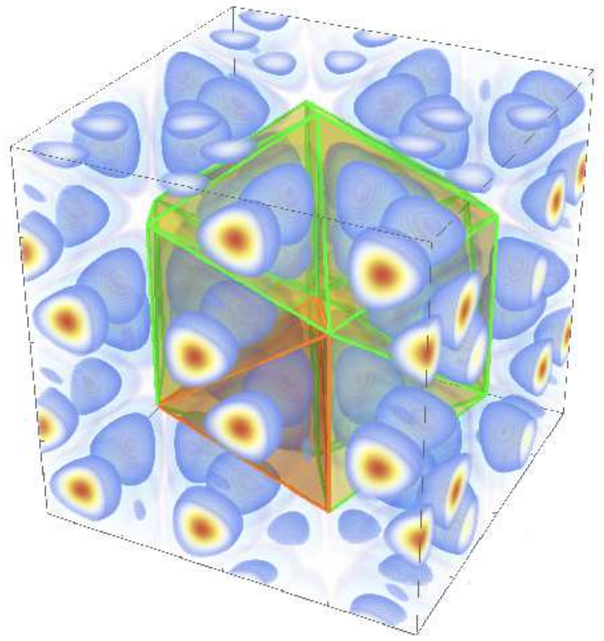
Physics Discovery In and Beyond the Standard Model

Opportunities for advances via artificial intelligence

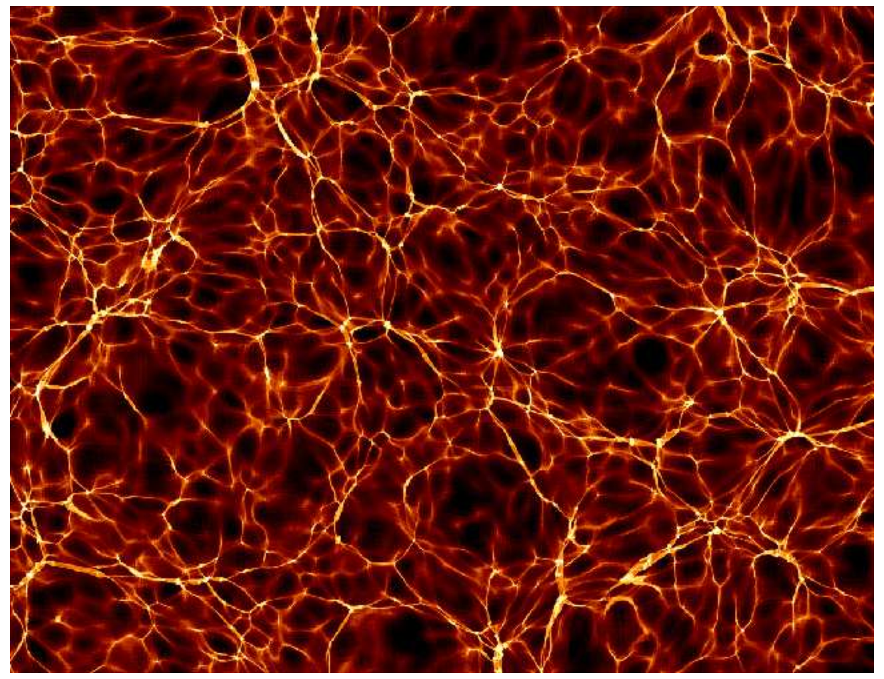
New Particles & Forces



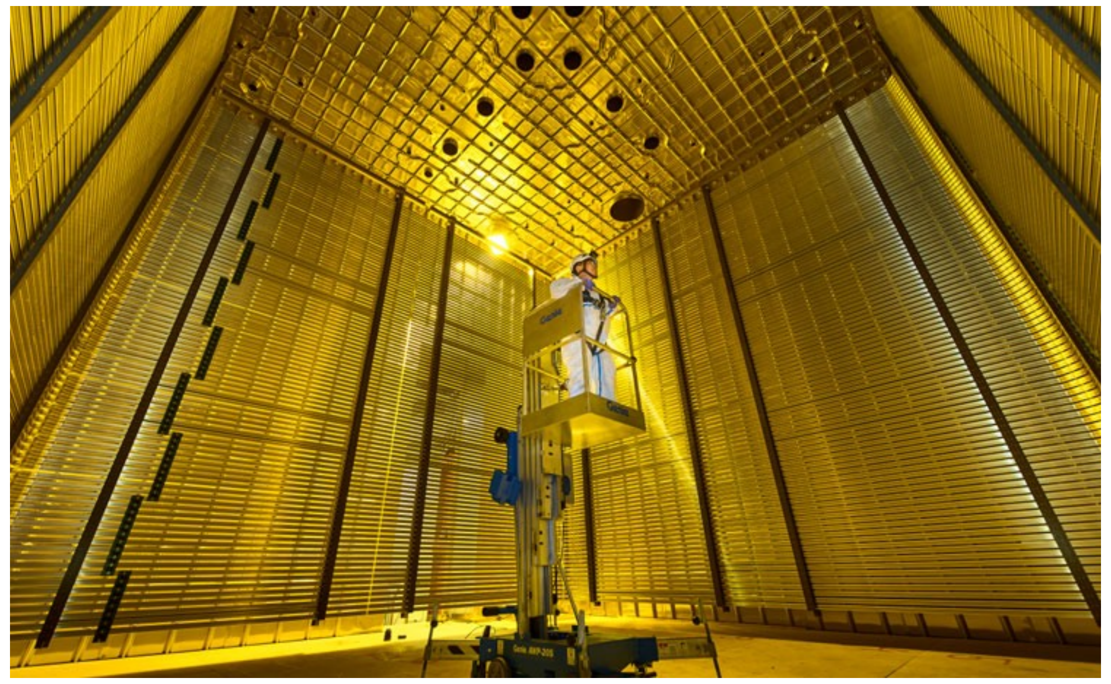
Strong Dynamics



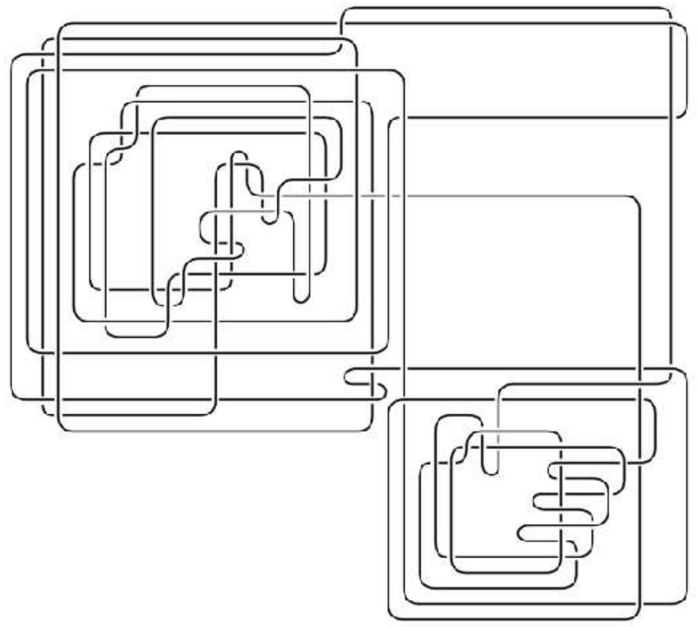
Dark Matter



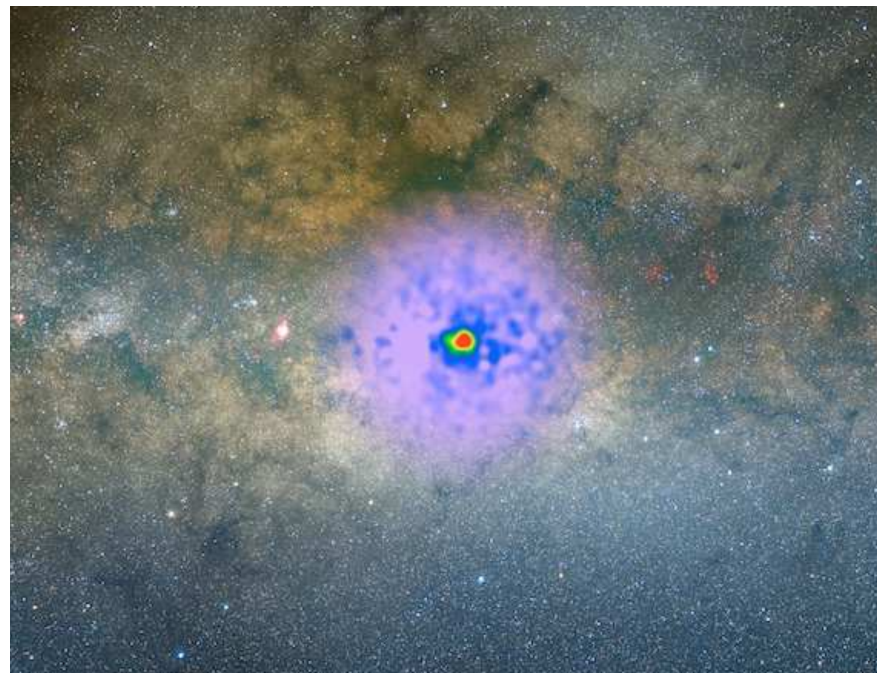
Neutrino Detection



Mathematical Physics

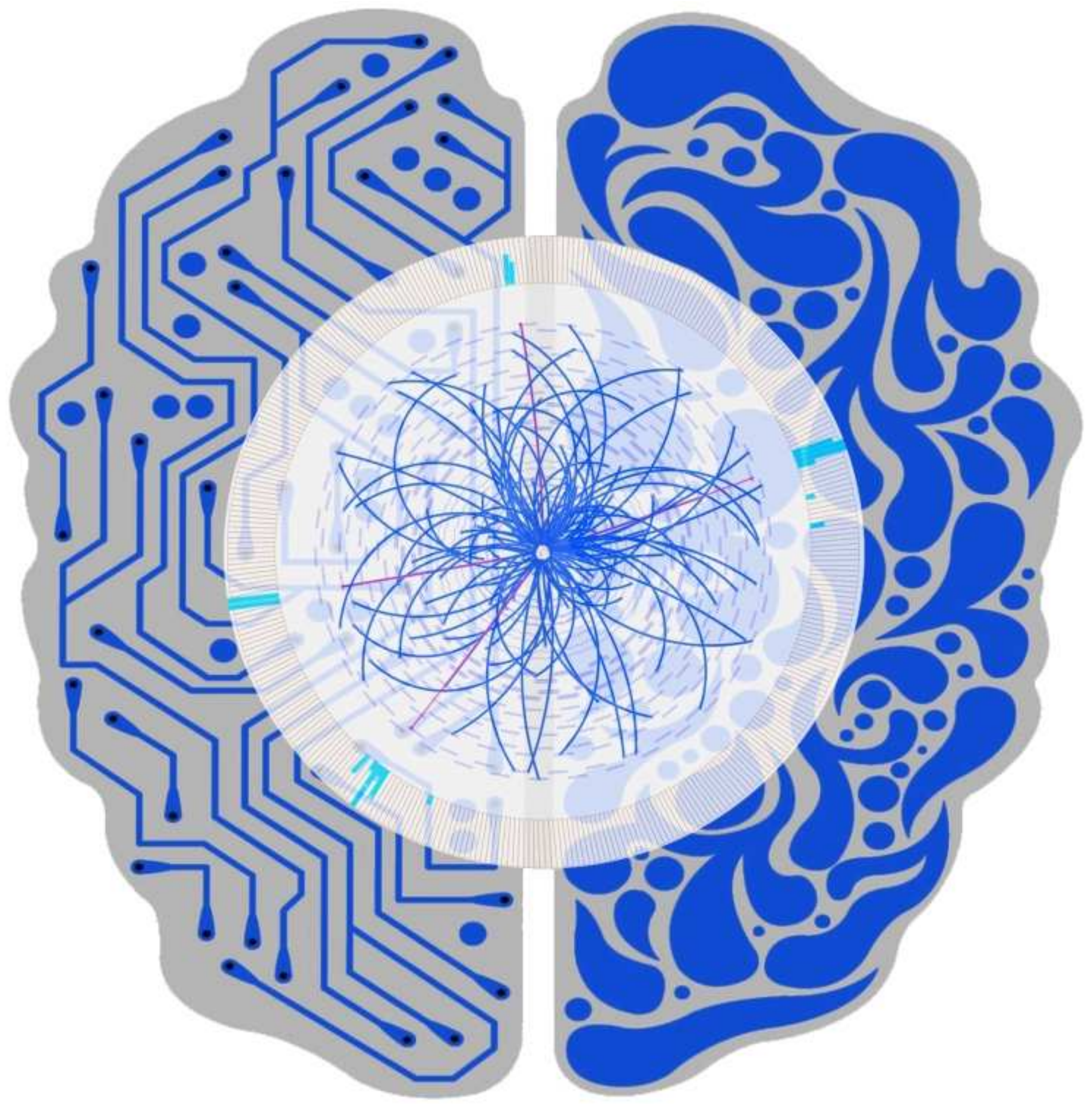


Astrophysics & Cosmology



…

Theory Perspective

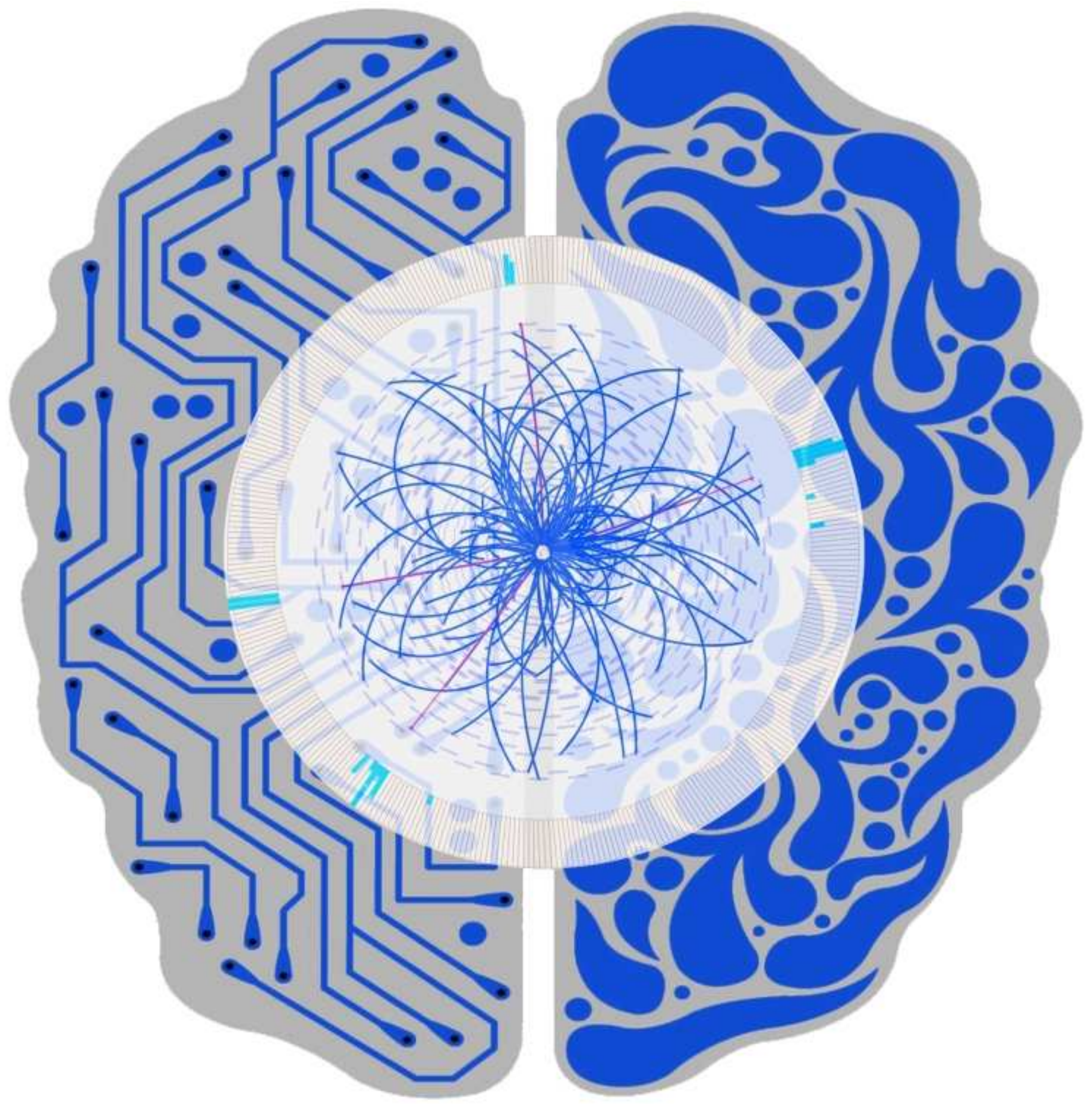


Can a machine "think"
like a physicist?


The New York Times
By Dennis Overbye
Nov. 23, 2020

Can a Computer Devise a Theory of Everything?

Theory Perspective



Can a machine "think" like a physicist?

The New York Times — By Dennis Overbye — Nov. 23, 2020

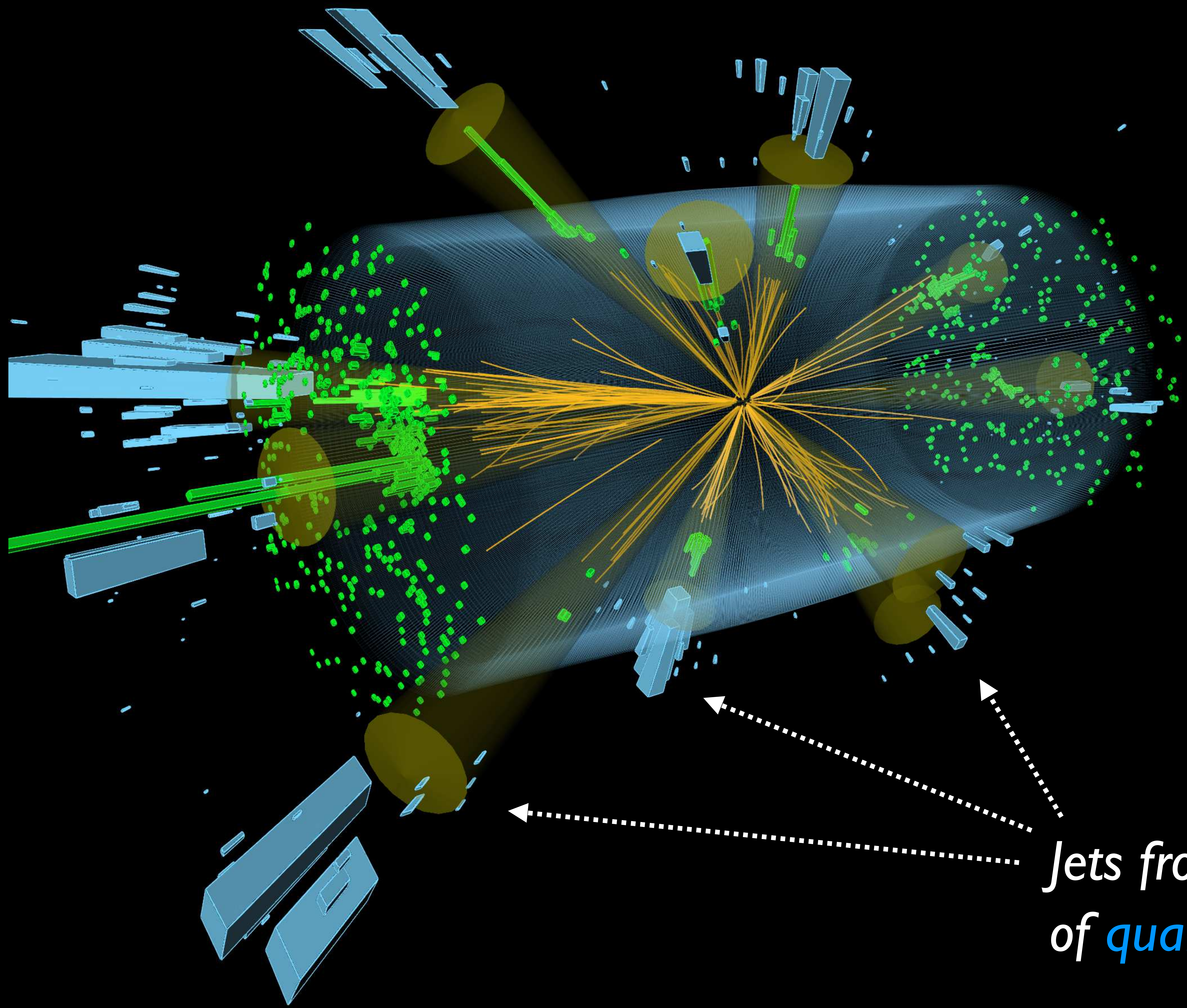
Can a Computer Devise a Theory of Everything?

Deep Learning meets Deep Thinking

Artificial intelligence based on first principles & best practices of fundamental physics

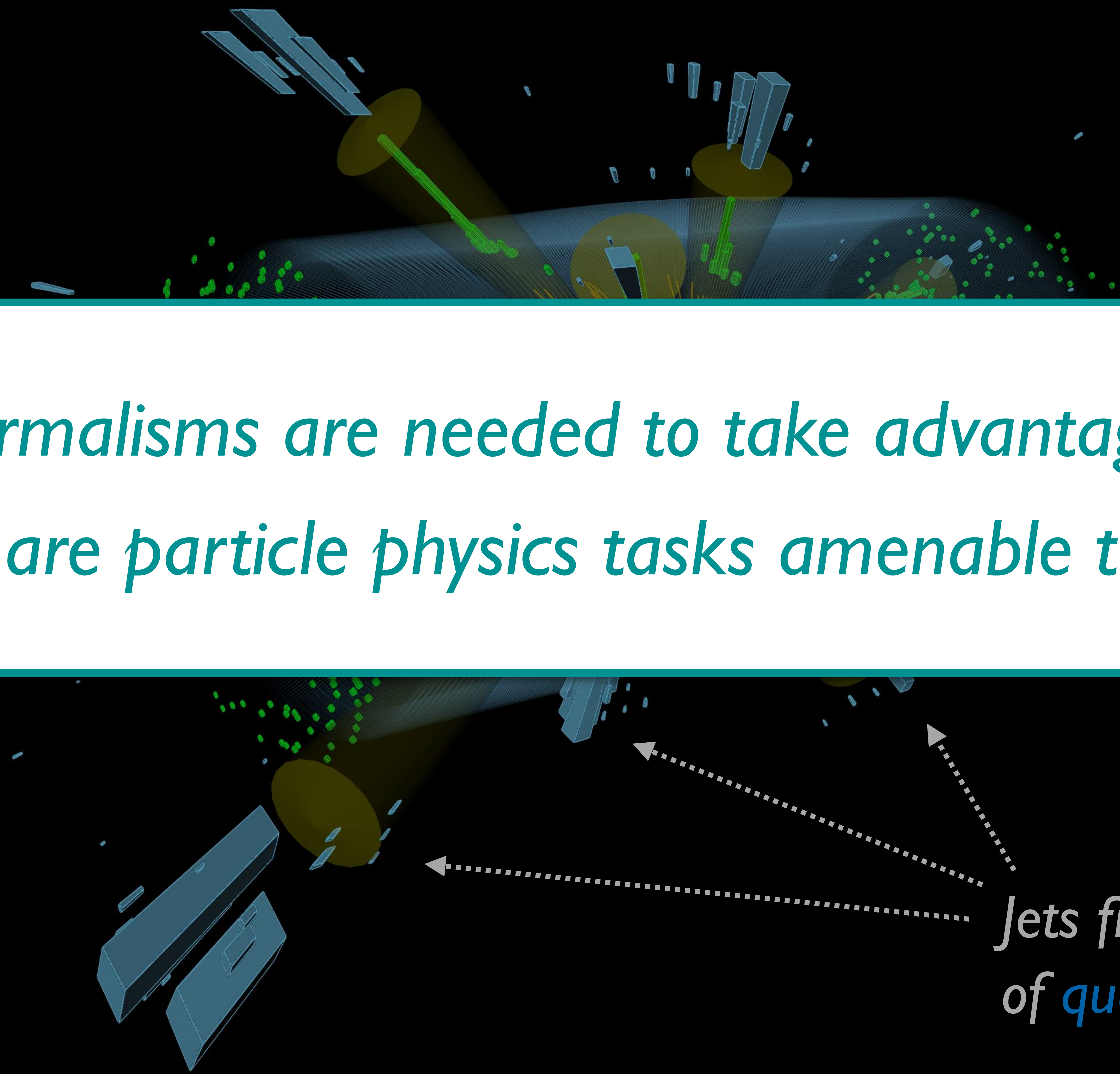
[see NSF Institute for Artificial Intelligence and Fundamental Interactions, http://iaifi.org/]

Learning from Particle Collisions



Jets from fragmentation of quarks and gluons

Learning from Particle Collisions



“What formalisms are needed to take advantage of ML?”

“Why are particle physics tasks amenable to ML?”

Jets from fragmentation of quarks and gluons

Machine Learning Foundations

E.g.: Likelihood ratio trick

"What formalisms are needed?"
"Why are tasks amenable to ML?"

Goal: Estimate likelihood ratio 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: $$\underset{f(x)}{\arg\min}\, L = \frac{p(x)}{q(x)}$$

$$-\min_{f(x)} L = \int dx\, p(x) \log \frac{p(x)}{q(x)}$$

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]

Machine Learning Foundations

"What formalisms are needed?"
"Why are tasks amenable to ML?"

E.g.: Likelihood ratio trick

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

Foundation of Deep Learning is indeed Deep Thinking

[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]

Machine Learning Requirements

If you have…

Well-specified loss
Reliable training data
Learnable function

…then you can leverage ML!

Many particle physics tasks can be translated into this language

Highlight via 3 examples from my research with:

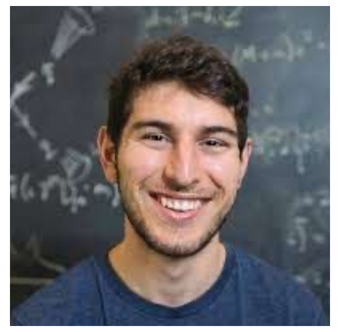

Patrick Komiske


Eric Metodiev

+


Ben Nachman


Anders Andreassen


Anthony Badea

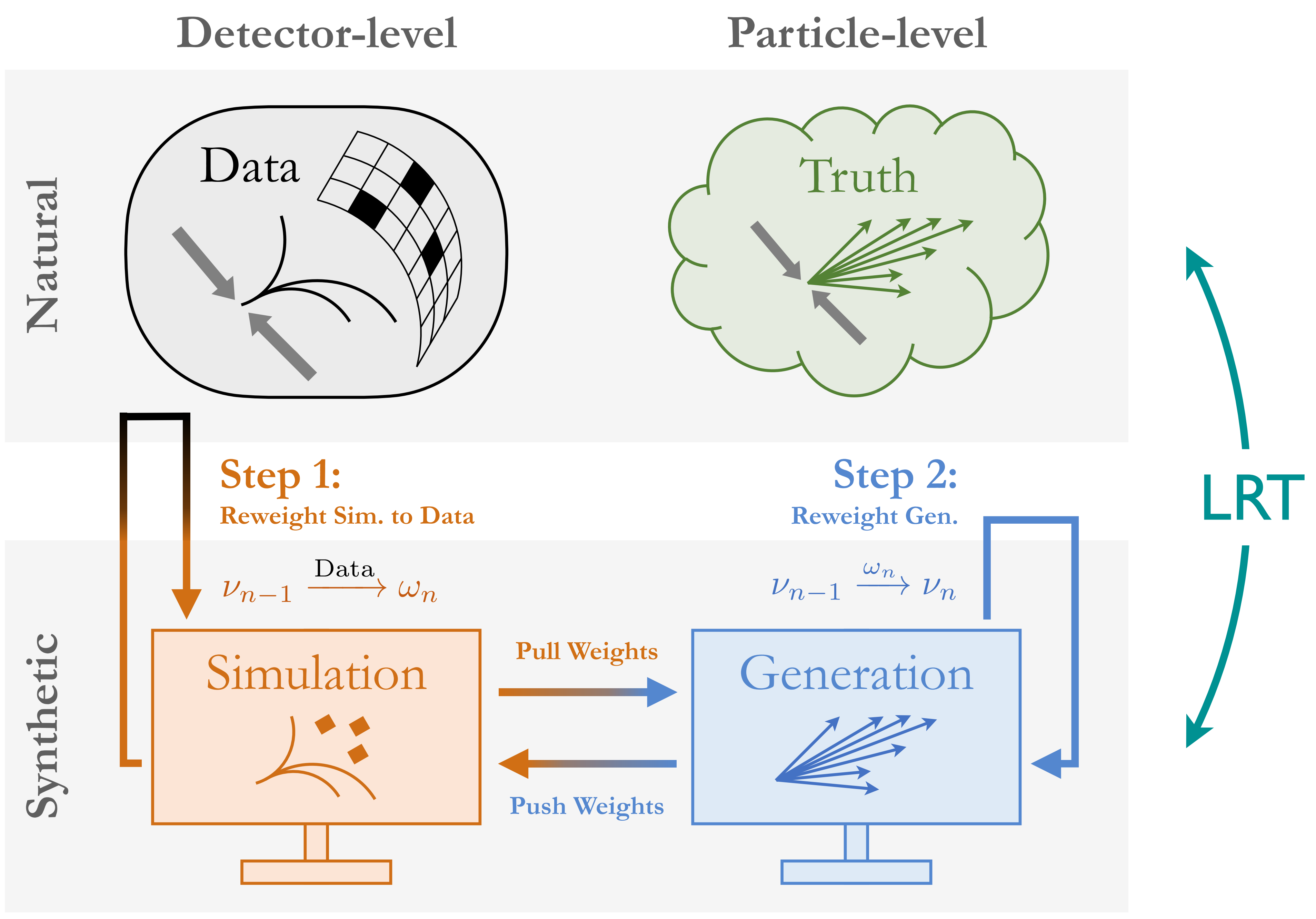
[see HEPML-LivingReview for extensive bibliography]

Detector Unfolding

Reweighting strategy exploiting likelihood ratio trick

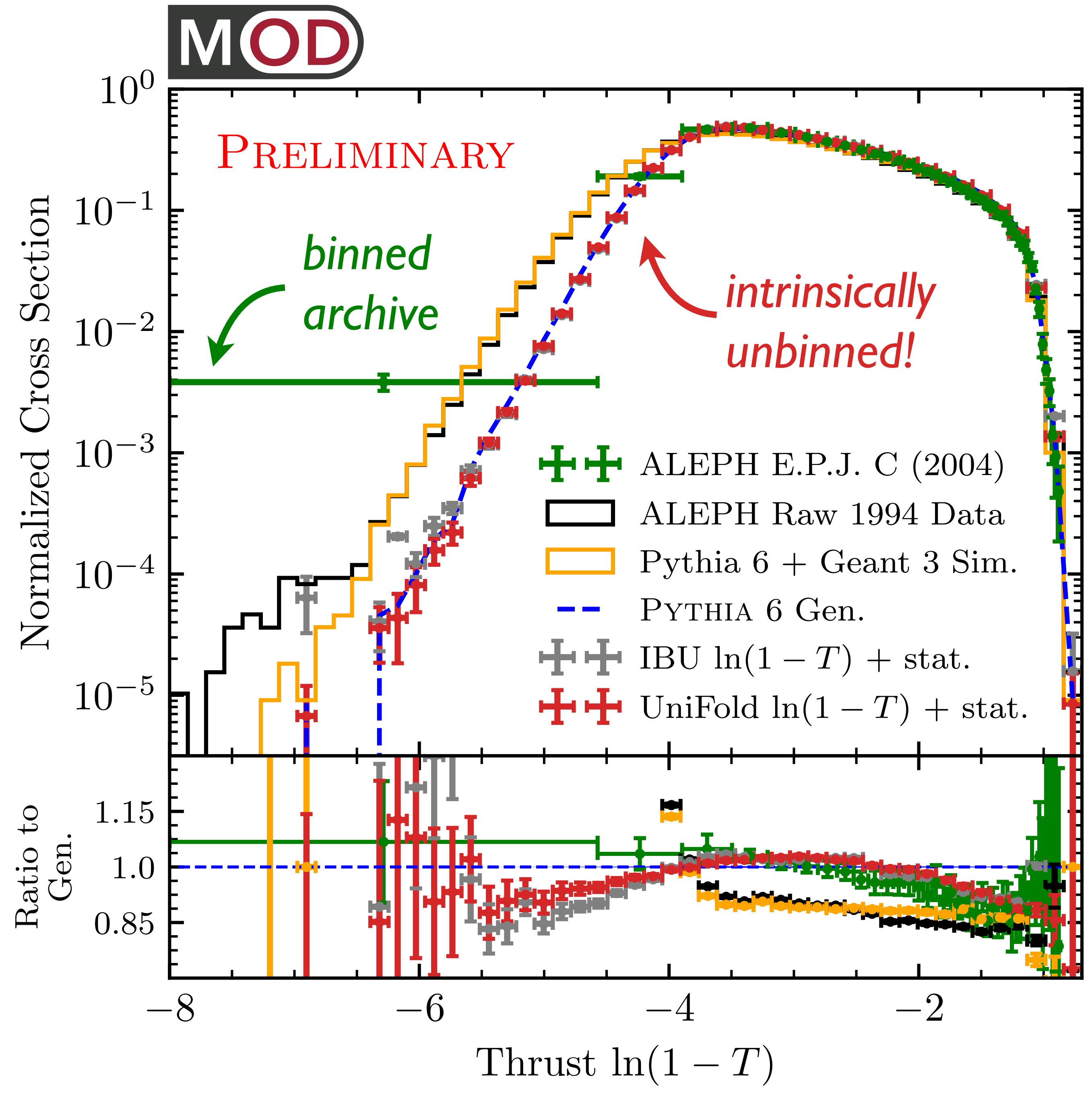
Well-specified loss
Reliable training data
Learnable function

OmniFold



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

Multi-dimensional unbinned detector corrections via iterated binary classification

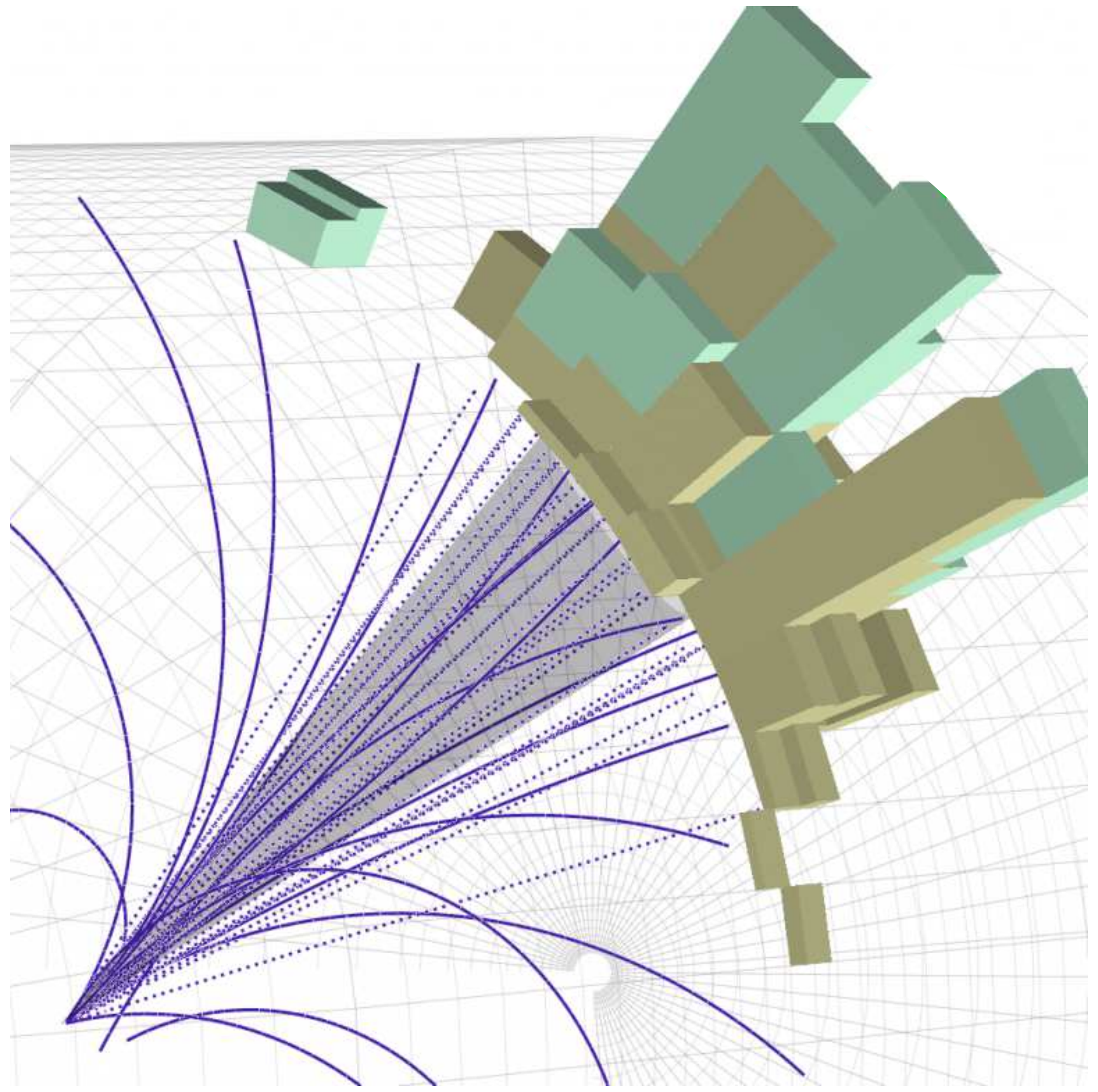


[talk by Badea, ICHEP 2020; cf. ALEPH, EPJC 2004]

Disentangling Jet Categories

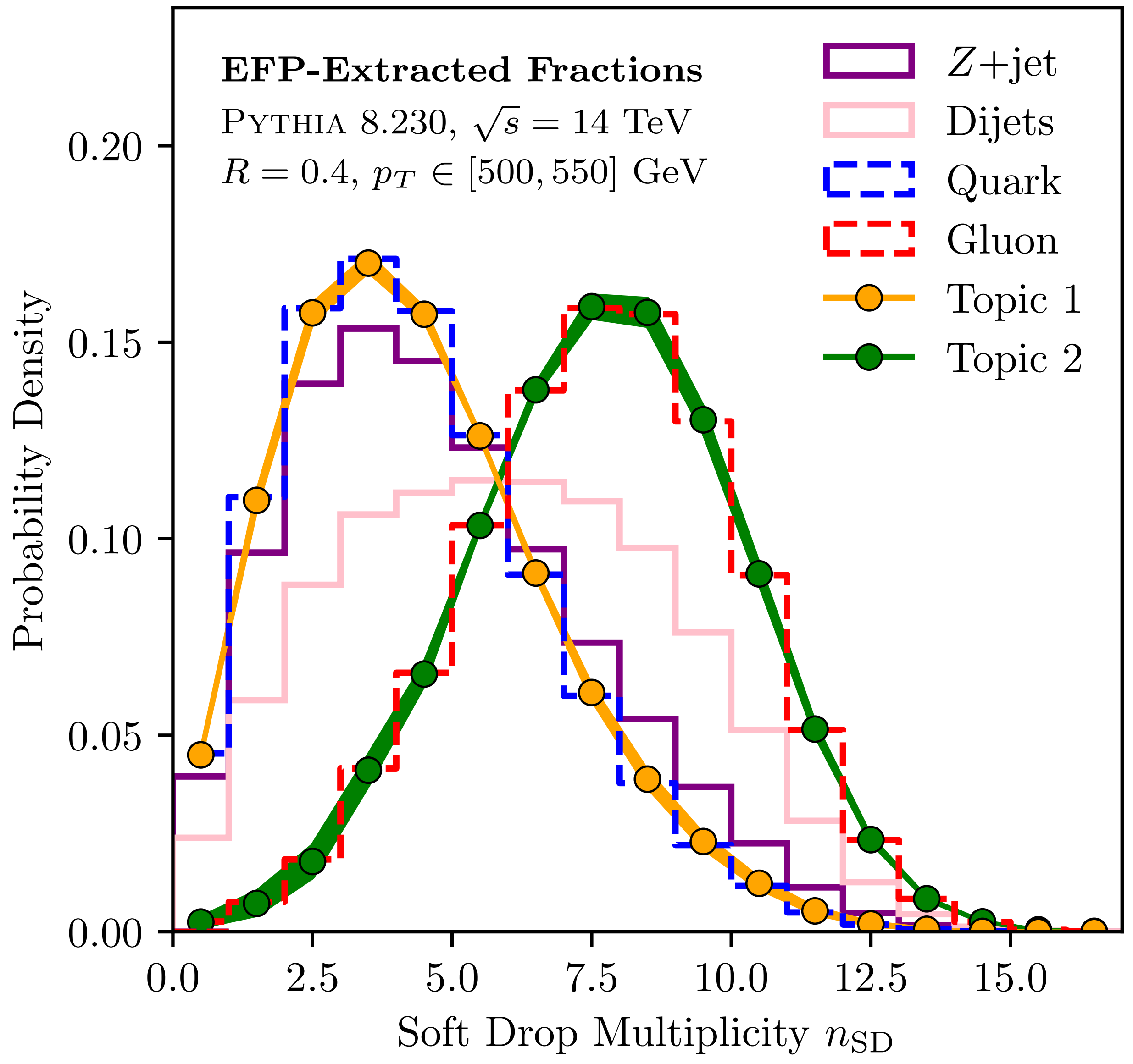
Separating quarks from gluons in real collider data

Does this jet arise from a quark or gluon?



Addressing this question with synthetic data encounters large theoretical uncertainties

While you can't label individual jets…



…you can extract quark and gluon distributions from mixed samples of real collider data

[Komiske, Metodiev, JDT, JHEP 2018; using Metodiev, Nachman, JDT, JHEP 2017; Metodiev, JDT, PRL 2018]
see also Blanchard, Flaska, Handy, Pozzi, Scott, PLMR 2013; Katz-Samuels, Blanchard, Scott, JMLR 2016]

Energy Flow Networks

Flexible architecture built around symmetries of particle data

Decompose full solution into latent space of visualizable elements

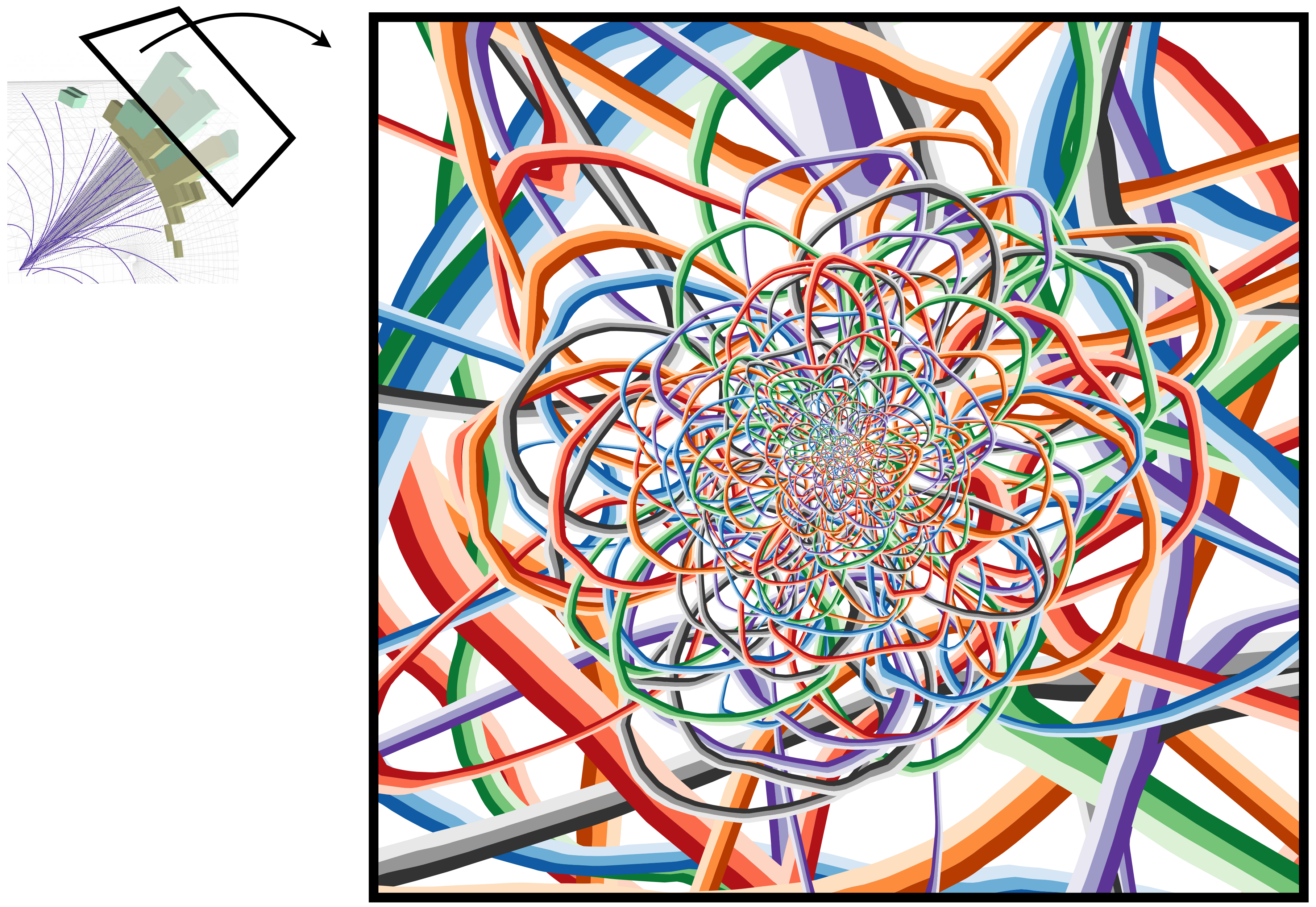
$$S(\mathcal{J}) = F(V_1, V_2, \ldots, V_\ell)$$

Identical particles are indistinguishable (QM)

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

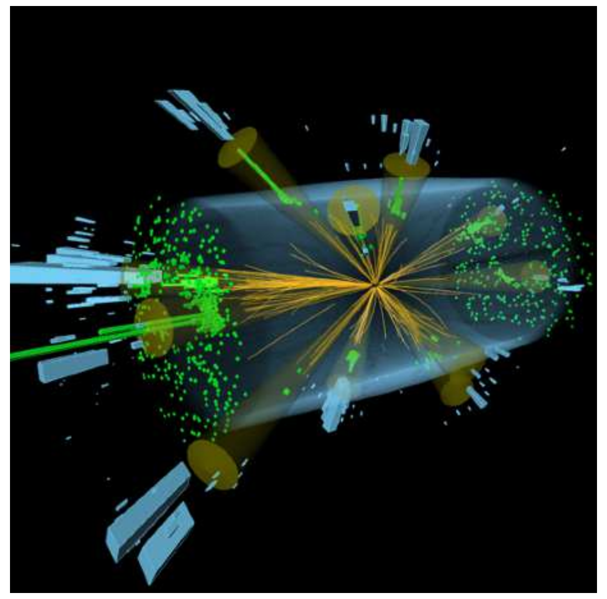
Energy flow is "safe" (QFT)

Network learns to solve jet tasks by exploiting fractal structure of jet formation



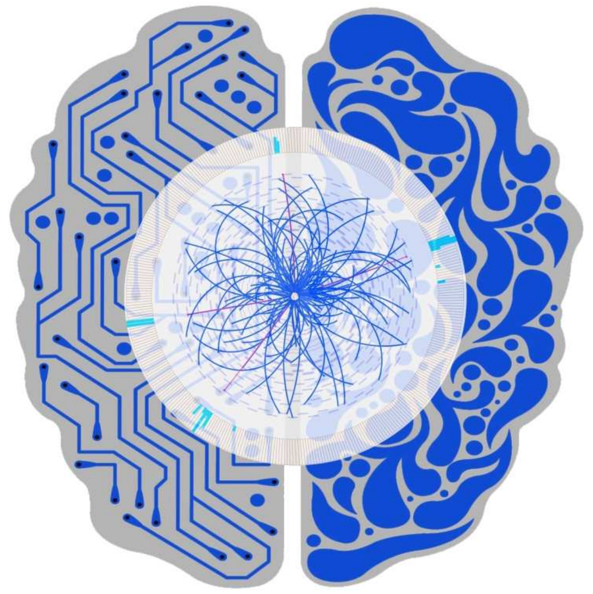
[Komiske, Metodiev, JDT, JHEP 2019; see also Komiske, Metodiev, JDT, JHEP 2018; code at energyflow.network; special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, NeurIPS 2017]

Summary: Artificial Intelligence for Physics Discovery

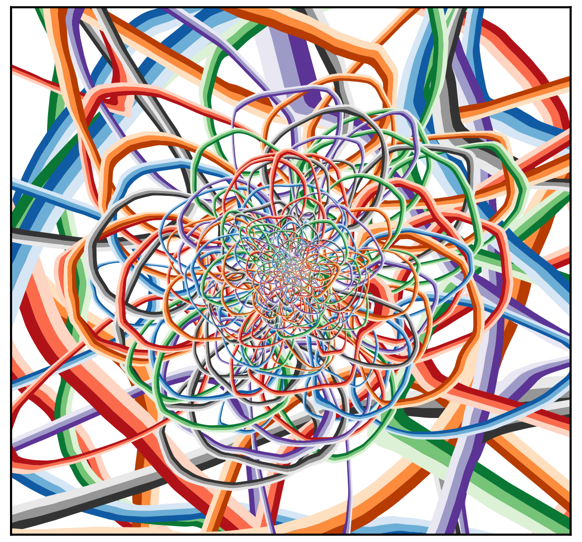


Many opportunities for physics discovery with ML

Requires shift of theoretical perspective
from solving problems to specifying problems



Many physics tasks translatable to well-specified loss, reliable training data, and learnable function



AI techniques are most powerful when they directly incorporate physics principles and best practices

For Discussion

Welcoming Daniel Whiteson into the conversation

"But what has the machine learned?"

"How do we ensure robustness of AI?"

"Can AI replace traditional physics theory?"

Looking forward to a lively discussion of these and other topics in February!