Artificial Intelligence and Fundamental Physics

Jesse Thaler

LISHEP 2021, Virtual Brazil — July 6, 2021 — jthaler@mit.edu

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]

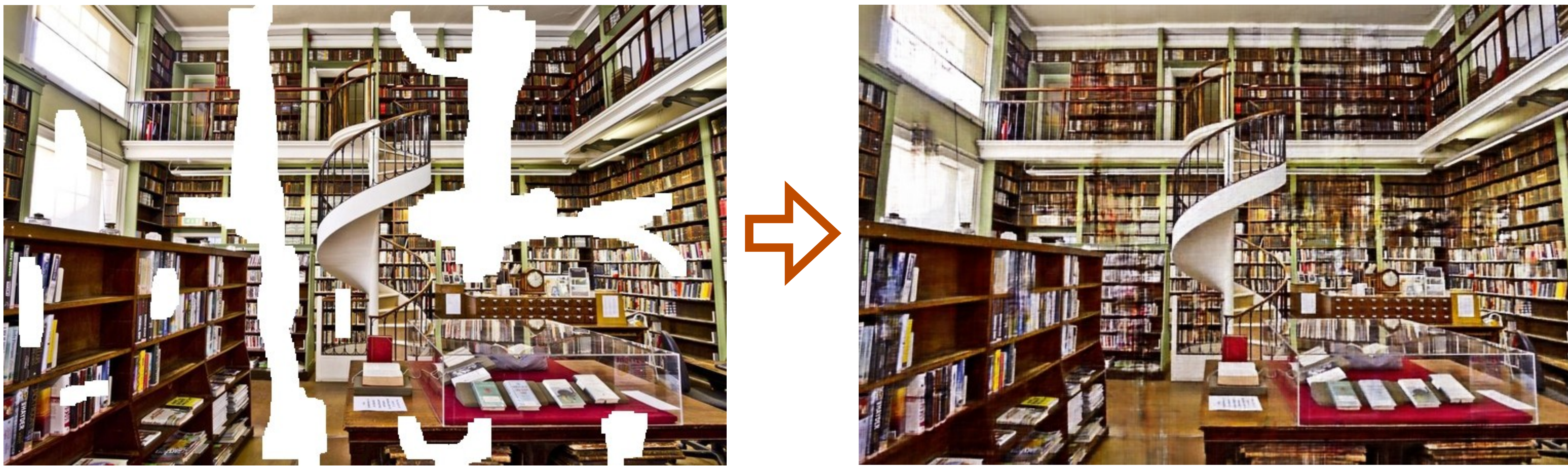
AI^2: Ab Initio Artificial Intelligence

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

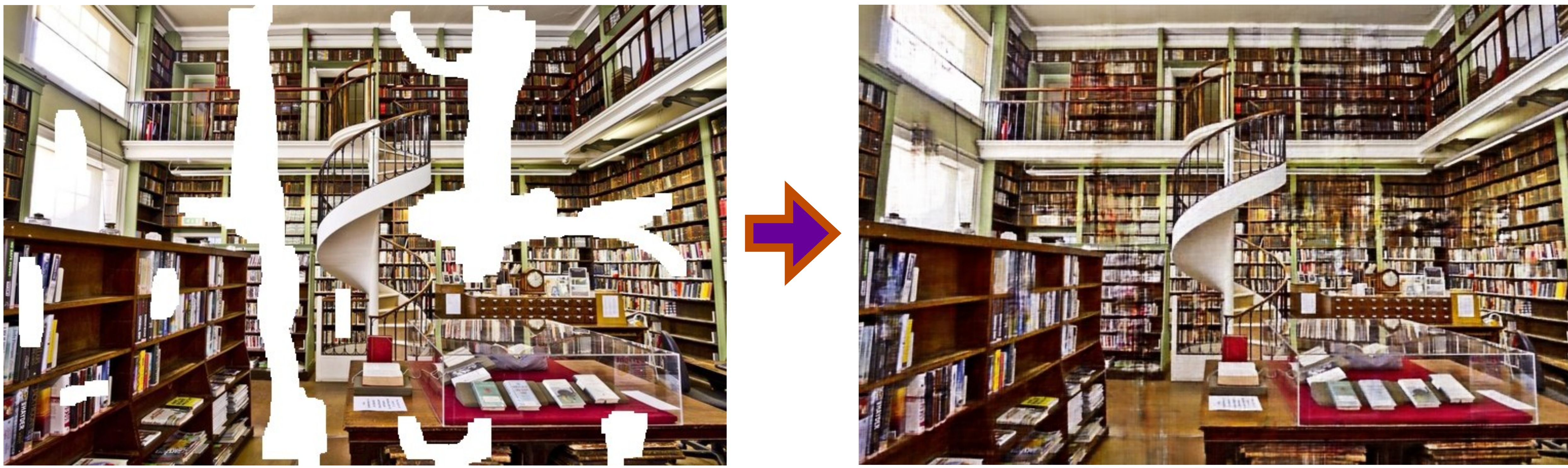


increased computational power and large data sets

[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)

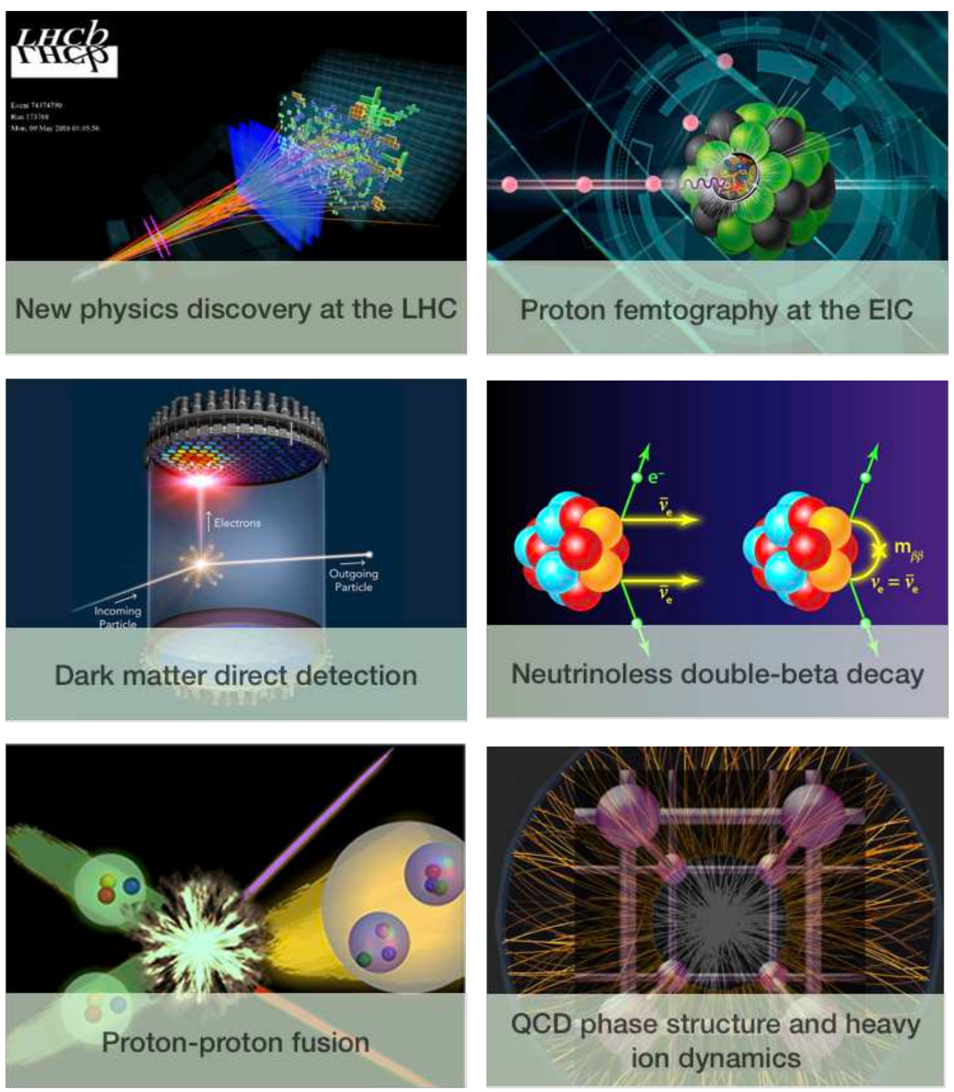
[Ulyanov, Vedaldi, Lempitsky, CVPR 2018]

AI² for Theoretical Physics

see backup for *Experimental Physics* and *Foundational AI*

E.g. Lattice Field Theory for Nuclear/Particle Physics

Equations governing the strong nuclear force are known, but precision computations are extremely demanding (>10% of open supercomputing in US)



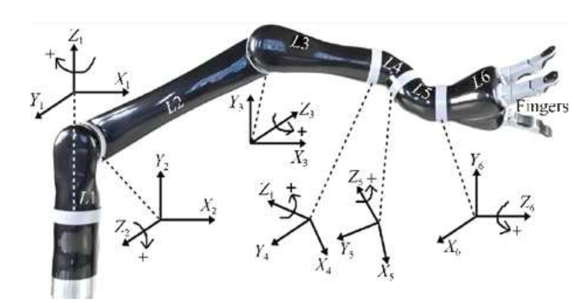
Industry collaboration to develop custom AI tools

Google DeepMind Massachusetts Institute of Technology

Custom generative models based on normalizing flows achieve 1000-fold acceleration while preserving symmetries & guaranteeing exactness

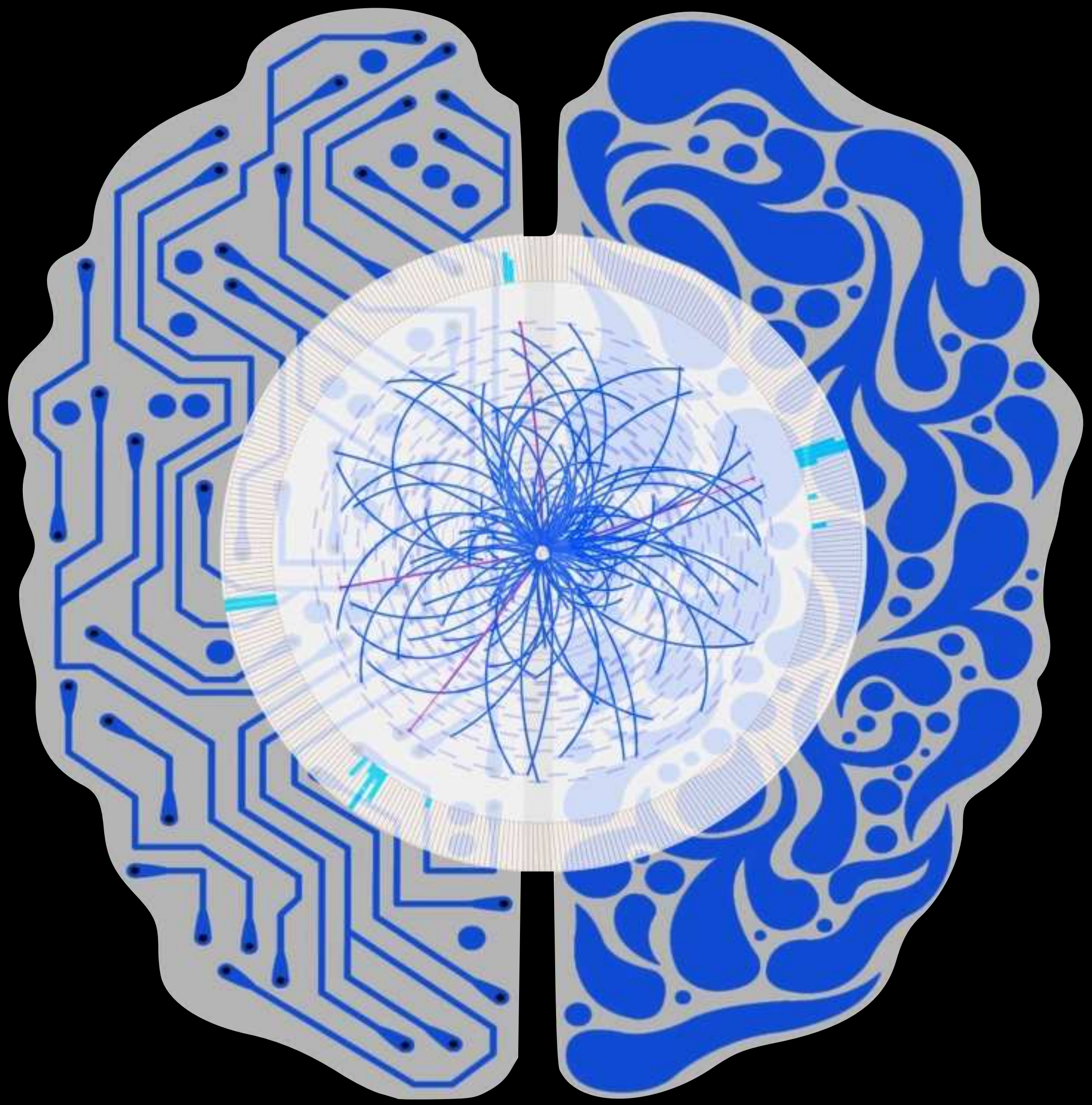
Tools designed for physics find interdisciplinary applications

Robotics



[Kanwar, Albergo, Boyda, Cranmer, Hackett, Racanière, Rezende, **Shanahan**, PRL 2020]

The Lens of Machine Learning



What formalisms are needed to leverage ML for HEP?

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*

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

[see e.g. Cranmer, Pavez, Louppe, arXiv 2015; D'Agnolo, Wulzer, PRD 2019;
simulation-based inference in Cranmer, Brehmer, Louppe, PNAS 2020;
relation to f-divergences in Nguyen, Wainwright, Jordan, AoS 2009; Nachman, Thaler, PRD 2021]

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 focus from solving problems to specifying problems

[see e.g. Cranmer, Pavez, Louppe, arXiv 2015; D'Agnolo, Wulzer, PRD 2019;
simulation-based inference in Cranmer, Brehmer, Louppe, PNAS 2020;
relation to f-divergences in Nguyen, Wainwright, Jordan, AoS 2009; Nachman, Thaler, PRD 2021]

Machine Learning Requirements

If you have in hand…

Well-specified loss
Reliable training data
Learnable function

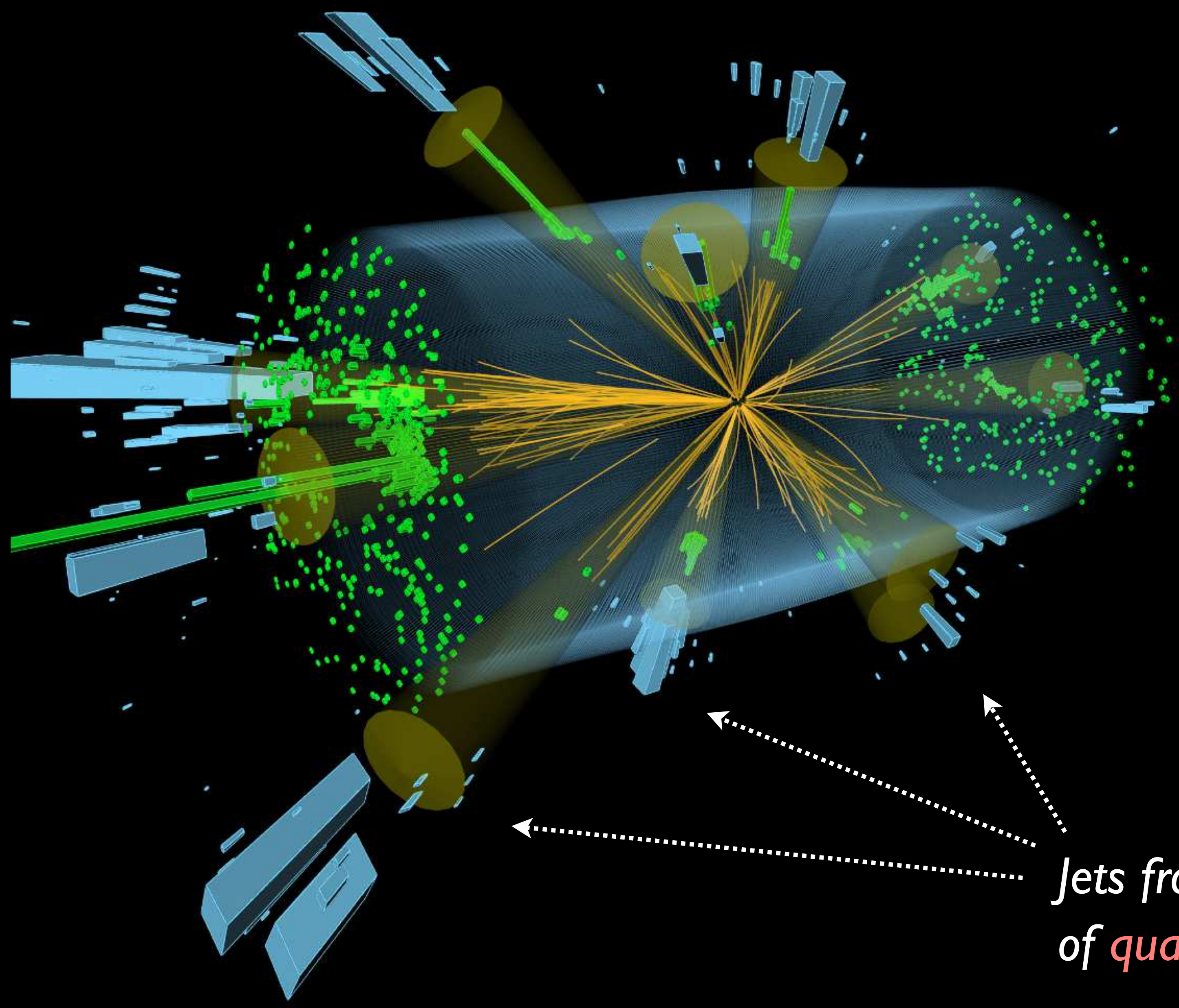
…then you can leverage ML!

Many HEP tasks can be phrased in this language

Physics input essential for robust usage of these tools

[see HEPML-LivingReview for extensive bibliography]

Machine Learning for High-Energy Physics

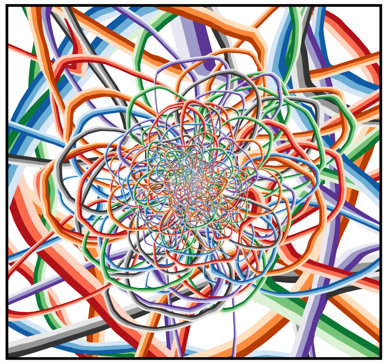


Jets from fragmentation of quarks and gluons

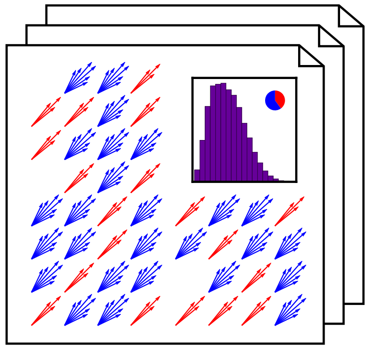
See more in talks by Heather, Javier, and Anima!

From Curmudgeon to Evangelist

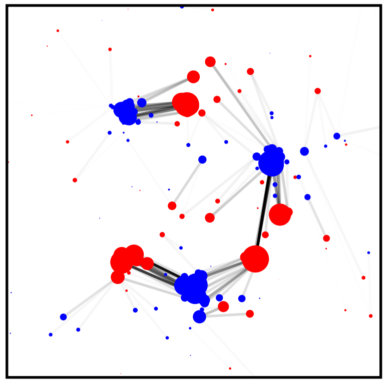
What have been helpful guides in pursing ML ⇔ HEP?



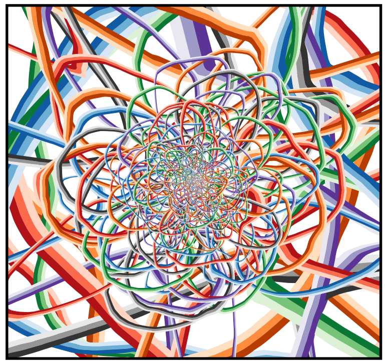
Can theoretical structures be encoded directly?



Can strategy be defined on physical quantities?



Can we leverage unsupervised machine learning?



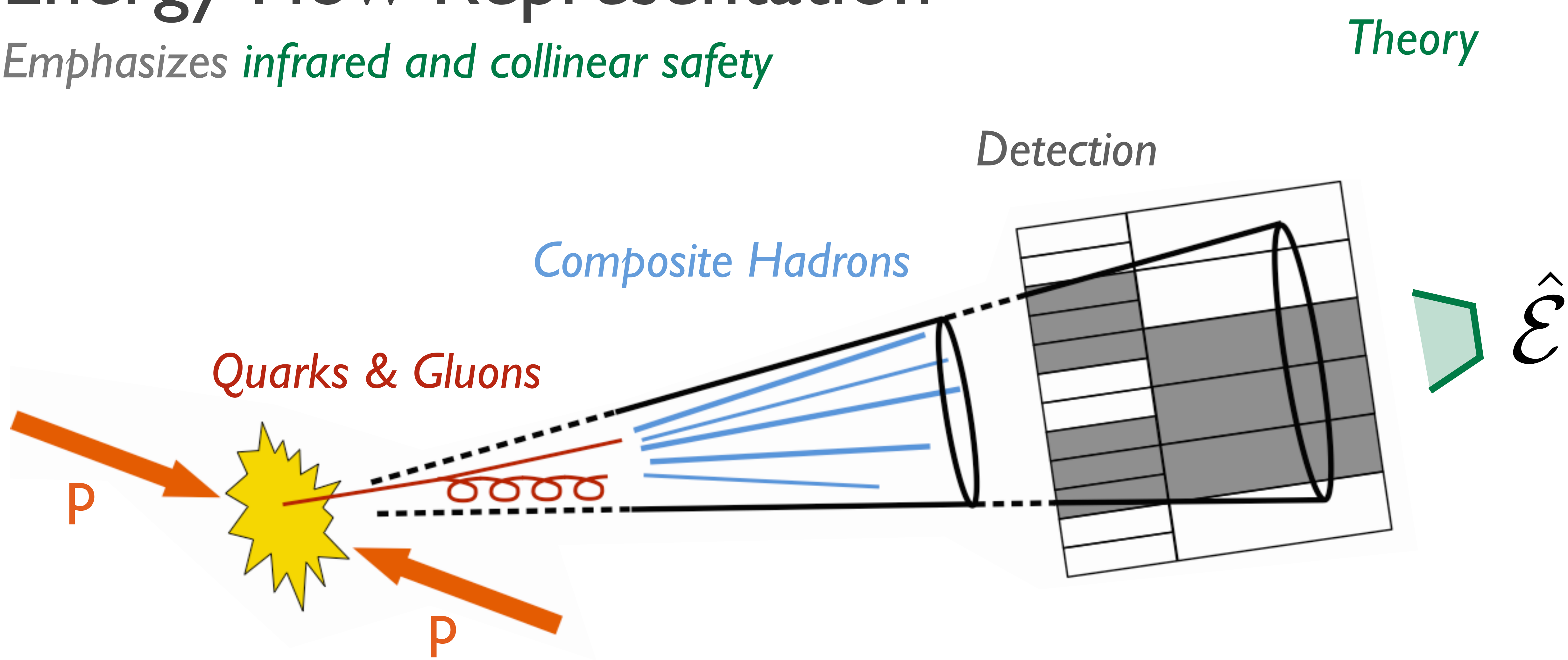
Can theoretical structures be encoded directly?

Can strategy be defined on physical quantities?

Can we leverage unsupervised machine learning?

Energy Flow Representation

Emphasizes ***infrared and collinear safety***



Energy Flow:

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

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

[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]
[complementary perspective on IRC unsafe information in Chakraborty, Lim, Nojiri, Takeuchi, JHEP 2020]

Energy Flow Networks

Architecture designed around symmetries and interpretability

Permutation invariant

Linear weights (i.e. safe)

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

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]

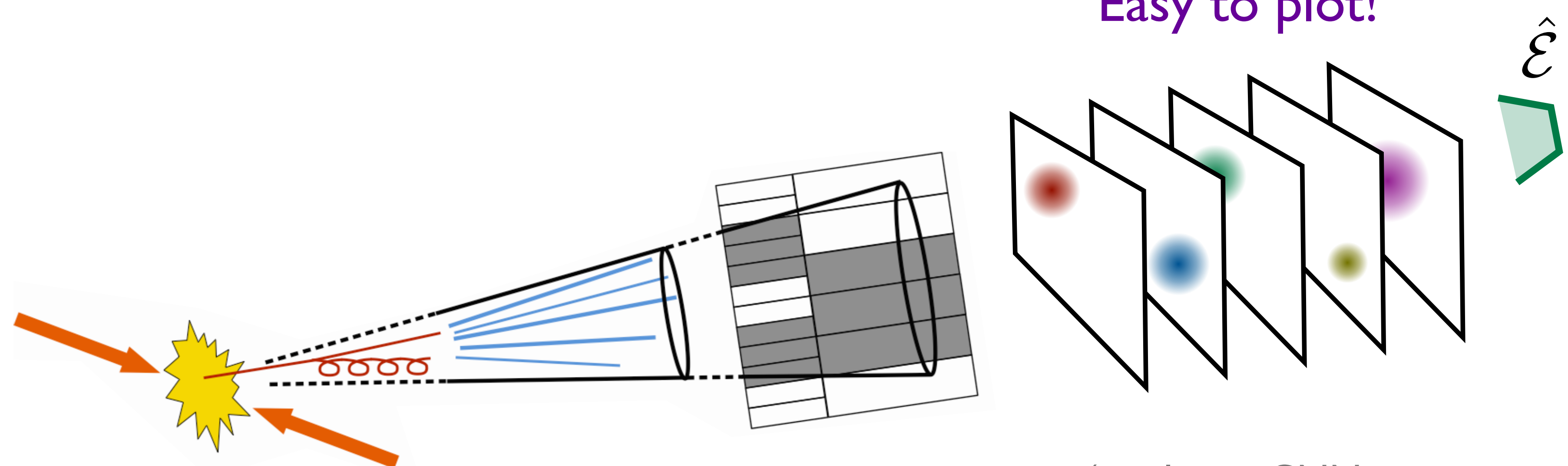
Energy Flow Networks

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

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)

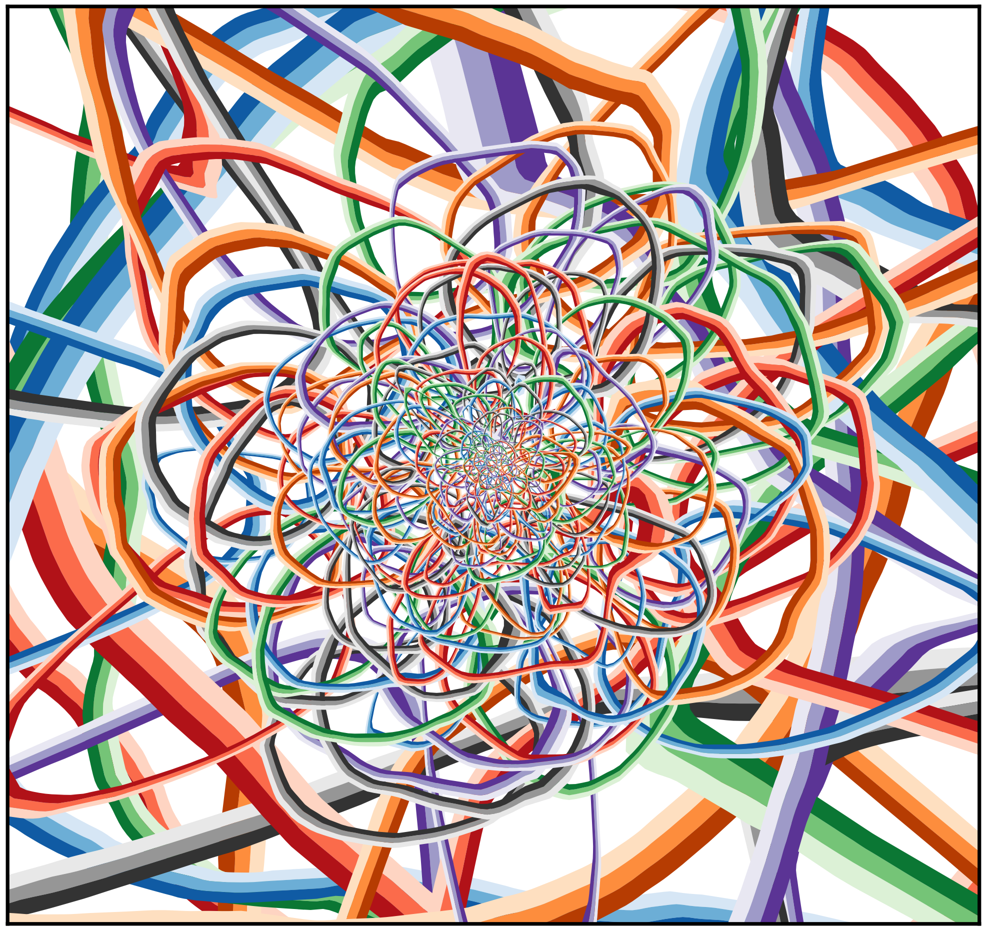
[Komiske, Metodiev, JDT, JHEP 2019; see also Komiske, Metodiev, JDT, JHEP 2018; code at energyflow.network;
special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, NIPS 2017;
other set-based architecture in Qu, Gouskos, PRD 2020; Mikuni, Canelli, EPJP 2020; Dolan, Ore, PRD 2021;
graph-based 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***

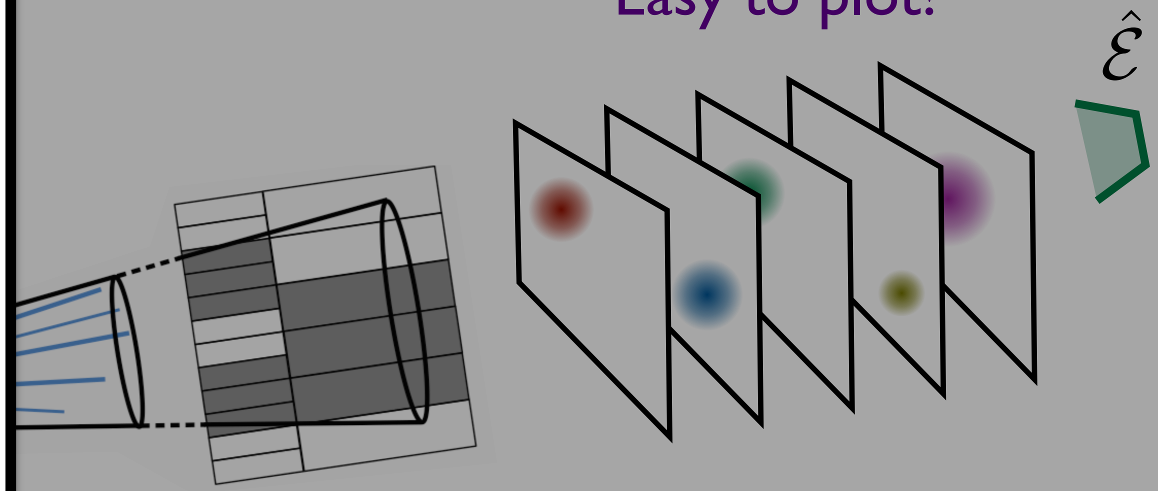
Psychedelic Network Visualization

Latent Dimension 256



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

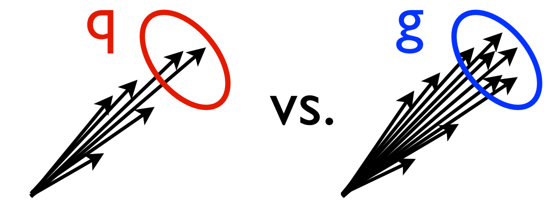
Easy to plot!

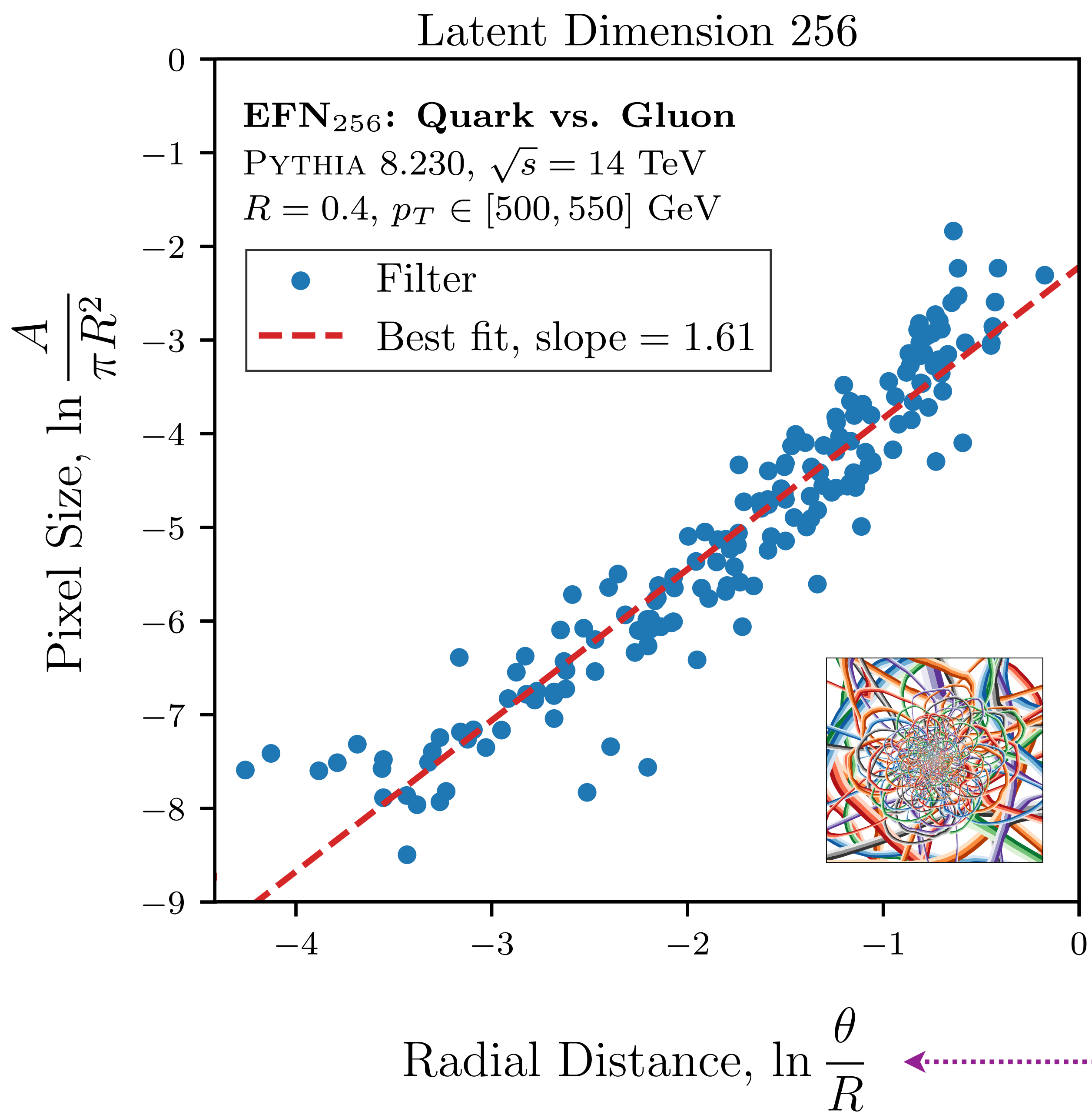


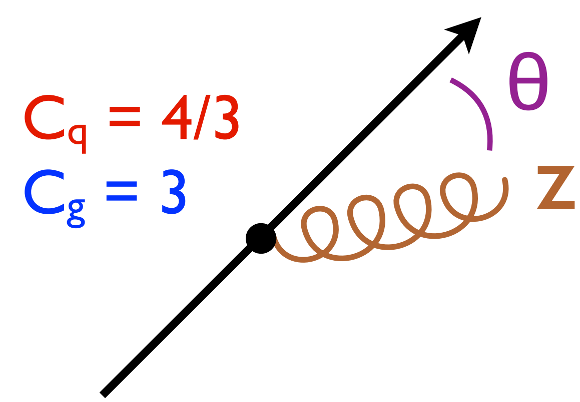
(similar to CNN filter activation)

[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, 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]

Machine Learning Collinear QCD

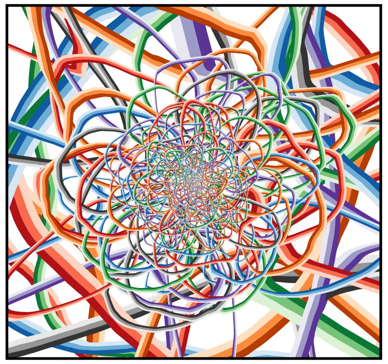






$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \underbrace{\frac{\mathrm{d}\theta}{\theta}}_{\text{Collinear}} \underbrace{\frac{\mathrm{d}z}{z}}_{\text{Soft}}$$

[Komiske, Metodiev, JDT, JHEP 2019]



Can theoretical structures be encoded directly?

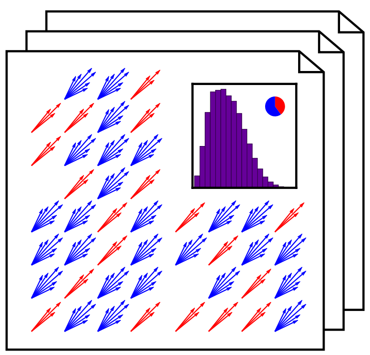
Energy Flow Networks ⇔ IRC Safety + Permutations

Can strategy be defined on physical quantities?

Can we leverage unsupervised machine learning?

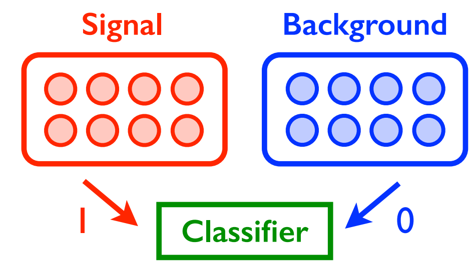
Can theoretical structures be encoded directly?

Energy Flow Networks ⇔ IRC Safety + Permutations



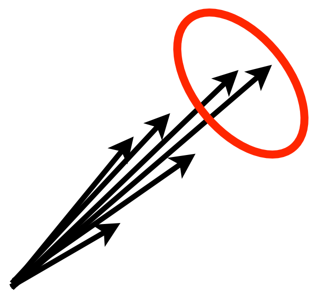
Can strategy be defined on physical quantities?

Can we leverage unsupervised machine learning?



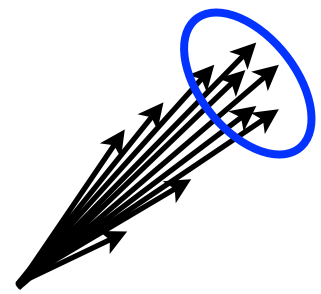
Quark/Gluon Classification

"Hello, World!" of Jet Physics

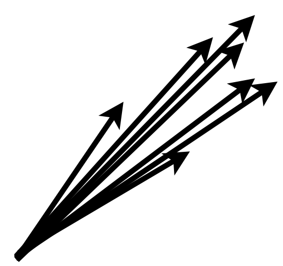


Quark
$C_q = 4/3$

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: $h(\mathcal{J}) = \left(1 + \dfrac{p(\mathcal{J}|\mathrm{G})}{p(\mathcal{J}|\mathrm{Q})}\right)^{-1}$

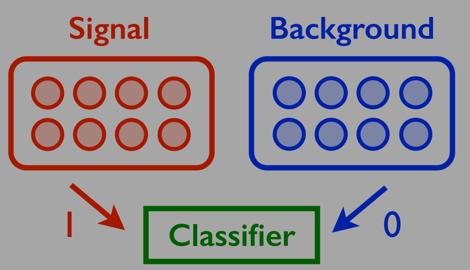
(Neyman-Pearson lemma)

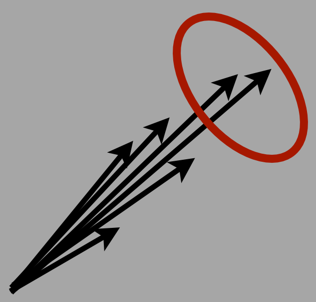
Likelihood ratio yields optimal binary classifier (and vice versa)

[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]

Quark/Gluon Classification

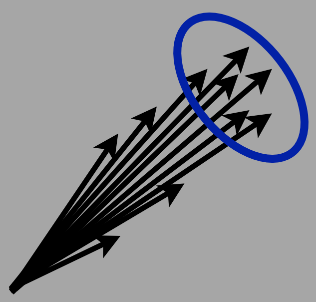
"Hello, World!" of Jet Physics





Quark
$C_q = 4/3$

vs.



Gluon
$C_g = 3 = 9/3$

What do you mean by "quark" and "gluon"?

Jets are clusters of colorless hadrons!

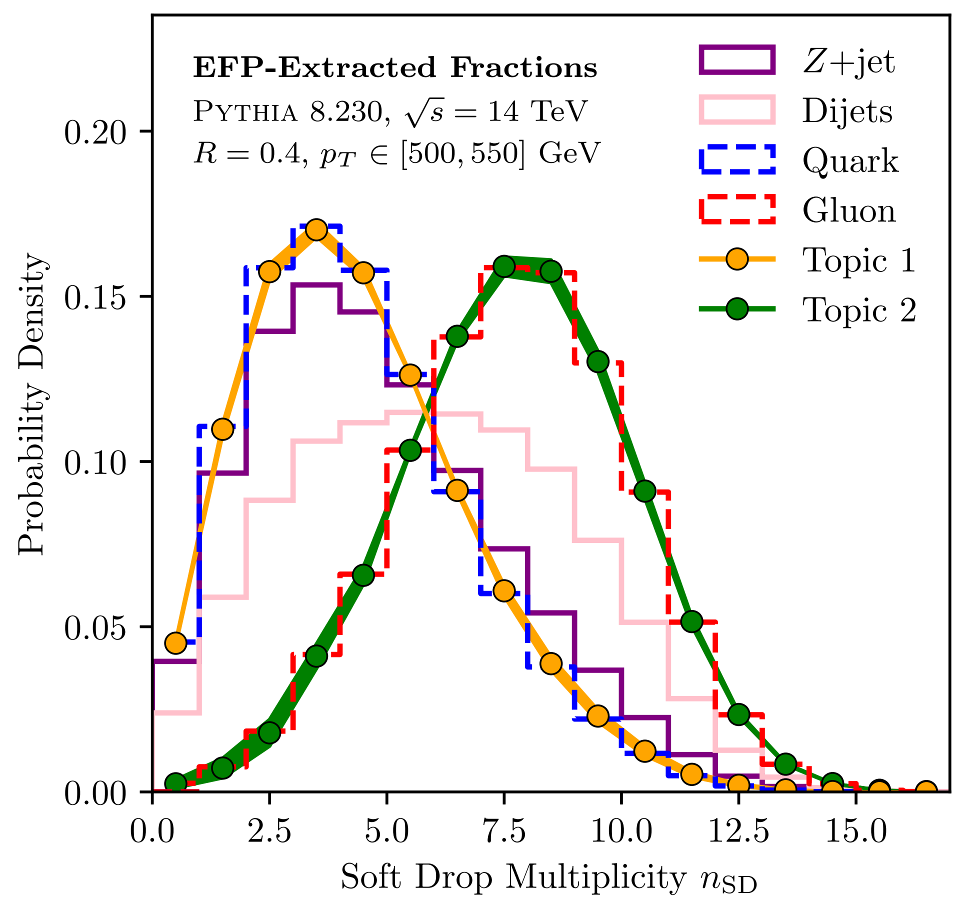
Parton shower "truth" is but a (useful) fiction!

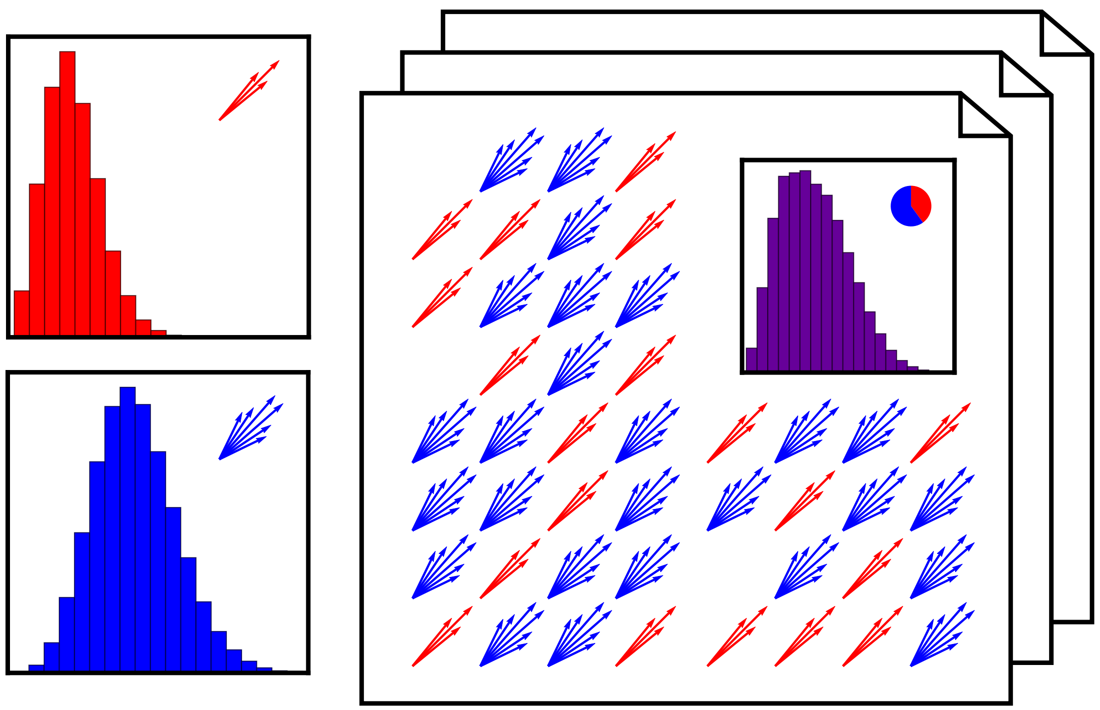
Likelihood ratio yields optimal binary classifier (and vice versa)

[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]

Topic Modeling to Disentangle Jet Categories

While you can't unambiguously label individual jets, you can extract quark and gluon distributions from hadron-level measurements



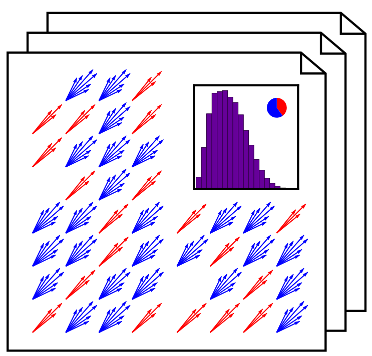


Key concept from natural language processing: "anchor words"

[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]

Can theoretical structures be encoded directly?

Energy Flow Networks ⇔ IRC Safety + Permutations



Can strategy be defined on physical quantities?

Jet Topics ⇔ Hadron-Level Approach to QCD Partons

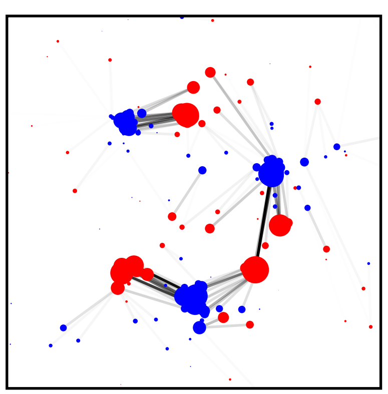
Can we leverage unsupervised machine learning?

Can theoretical structures be encoded directly?

Energy Flow Networks ⇔ IRC Safety + Permutations

Can strategy be defined on physical quantities?

Jet Topics ⇔ Hadron-Level Approach to QCD Partons



Can we leverage unsupervised machine learning?

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]

Déblai Remblai

[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Kantorovich, 1939; Vaseršteĭn, 1969; Wikipedia]

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

Additive Weight:

Ground Distance:

$\Rightarrow$ Distance Between Distributions

Déblai

Remblai

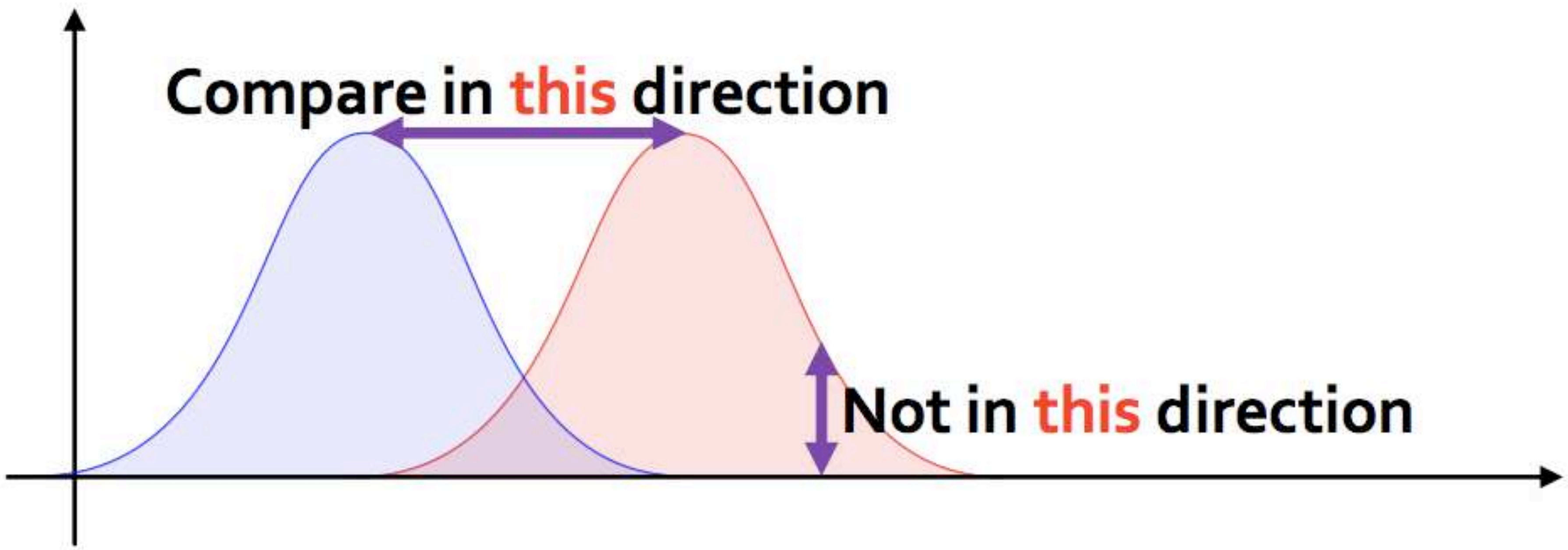
[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Kantorovich, 1939; Vaseršteĭn, 1969; Wikipedia]

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]

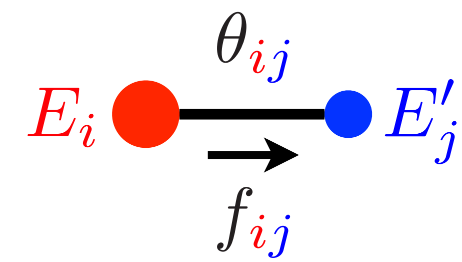
"Horizontal" comparison (EMD) yields better dynamic range than "vertical" comparison (e.g. KL)

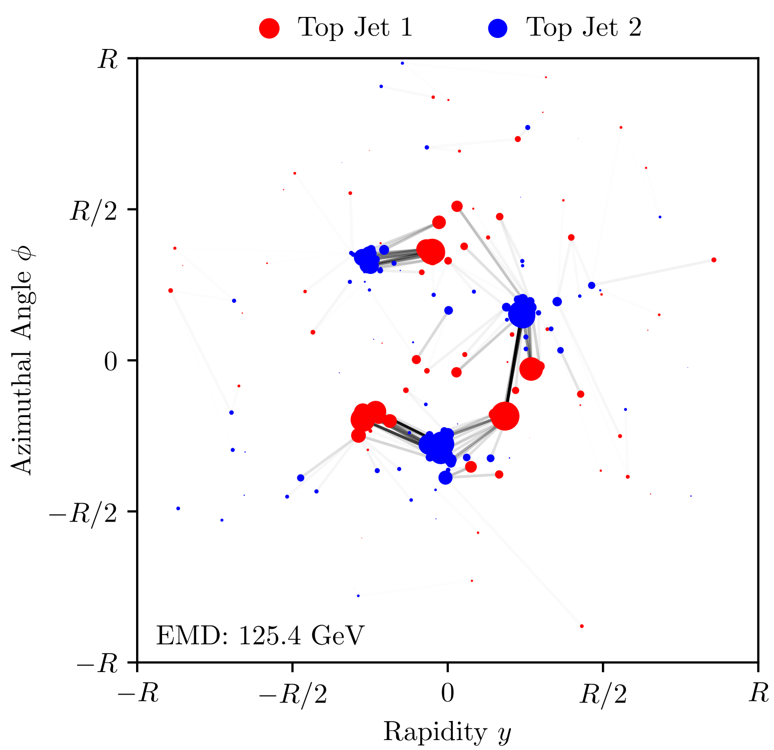


[figure from Kun, Math ∩ Programming]

[h/t Niles-Weed, ML4Jets 2020; Monge, 1781; Kantorovich, 1939; Vaseršteĭn, 1969; Wikipedia]

The Energy Mover's Distance

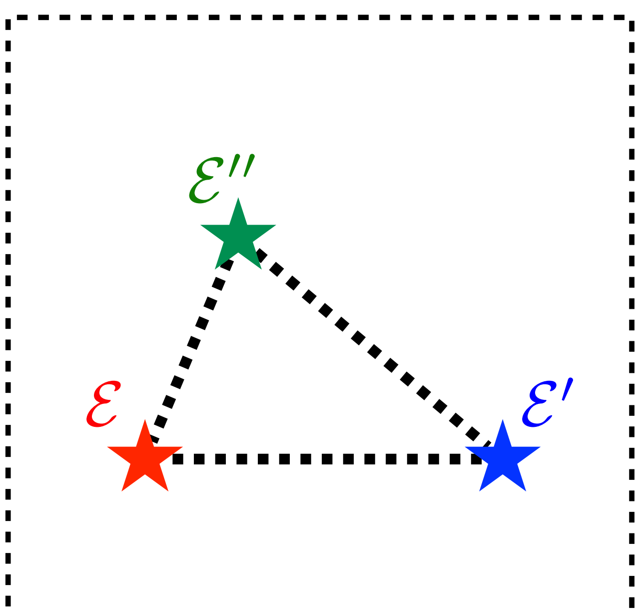




Optimal transport between energy flows…

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

in GeV



…defines a metric on the space of events

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

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

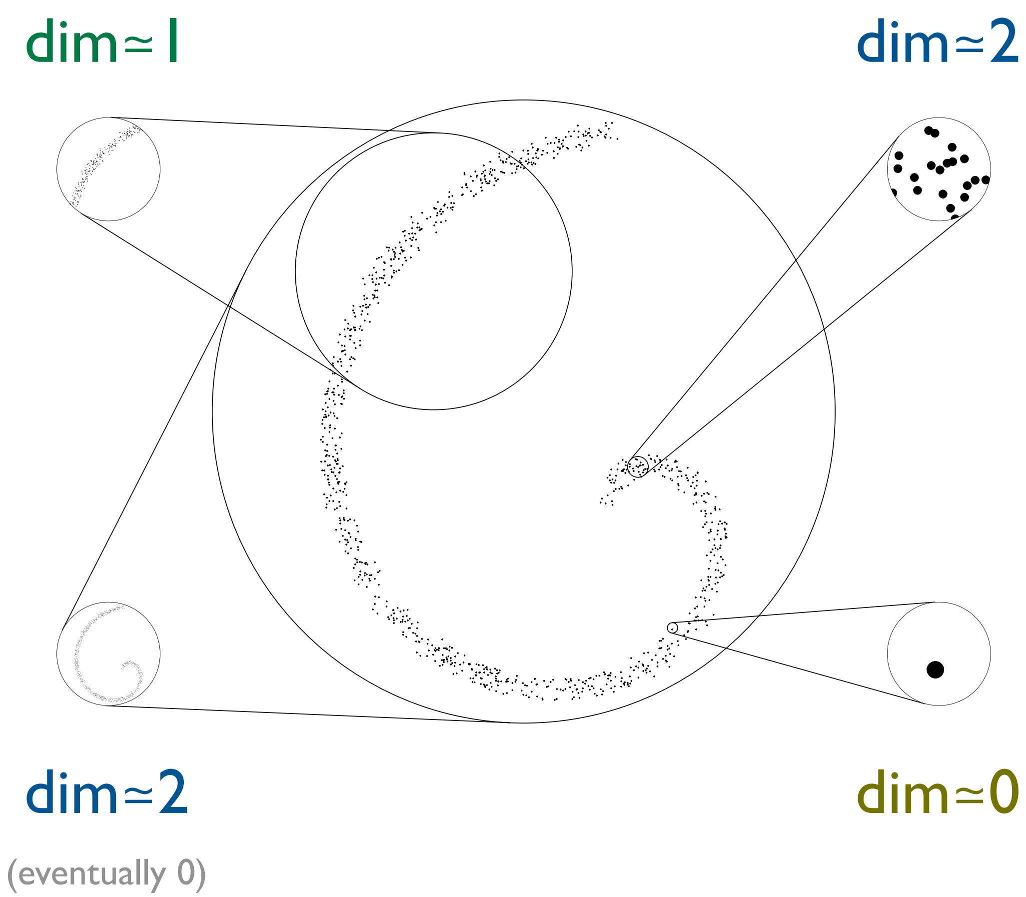
[Komiske, Metodiev, JDT, PRL 2019; see also Pele, Werman, ECCV 2008; Pele, Taskar, GSI 2013]
[see flavored variant in Crispim Romão, Castro, Milhano, Pedro, Vale, EPJC 2021]
[see computational speed up in Cai, Cheng, Craig, Craig, PRD 2020]

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]

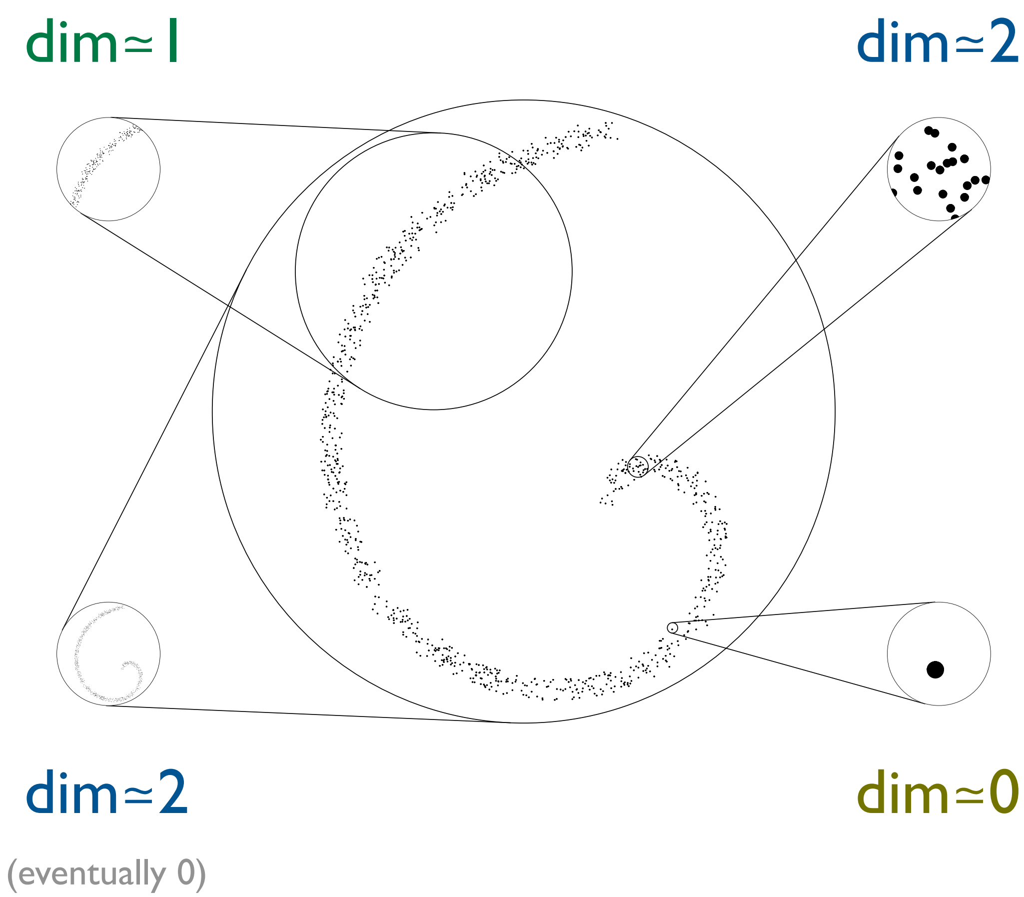


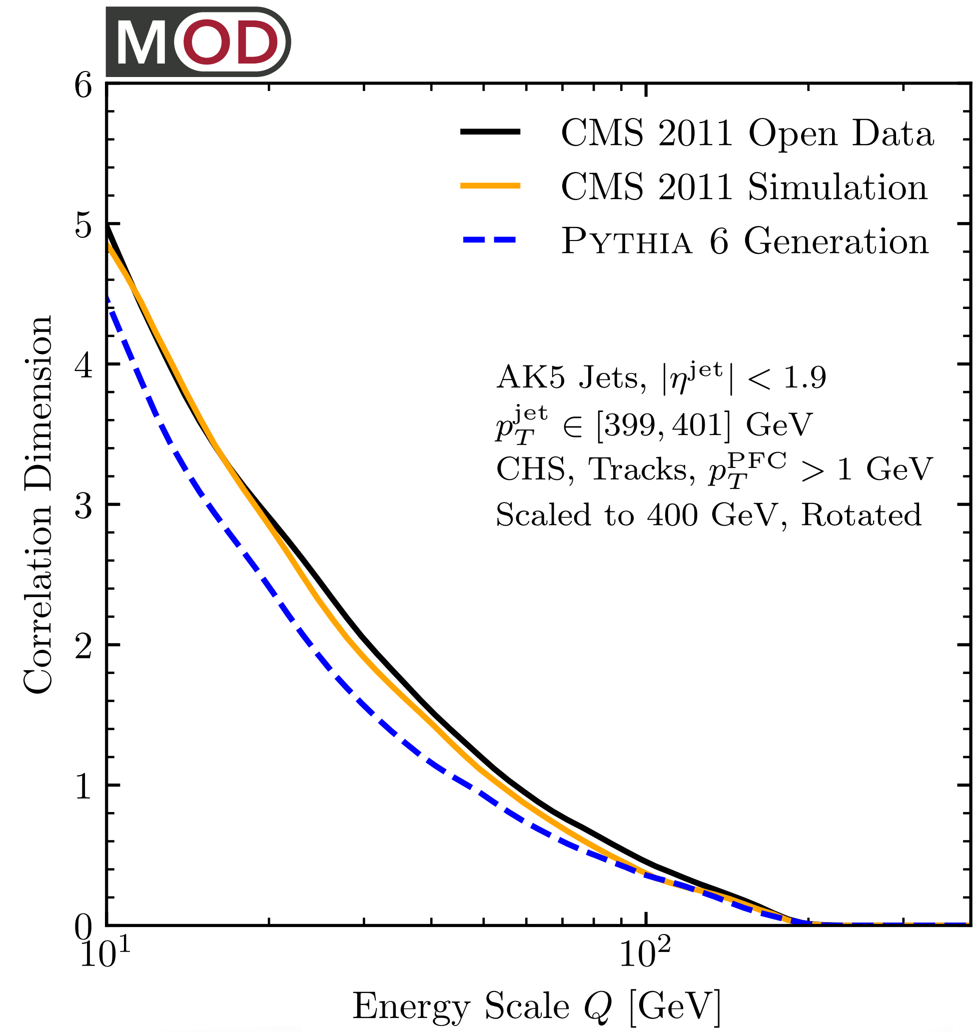
Dimensionality of Space of Jets

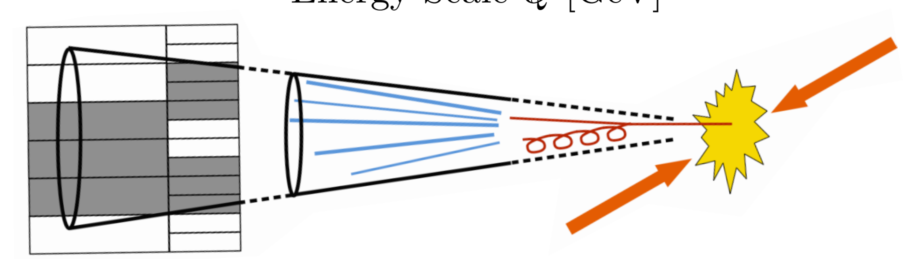
$$N_{\text{neighbors}}(r) \sim r^{\text{dim}}$$

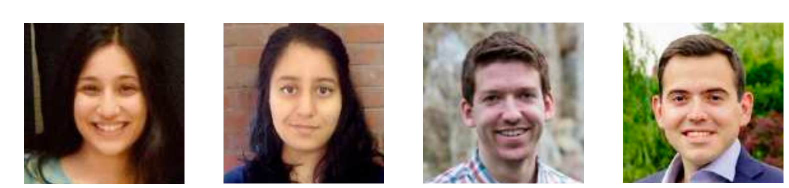
$$\Rightarrow \quad \text{dim}(r) \sim r \frac{\partial}{\partial r} \ln N_{\text{neighbors}}(r)$$

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





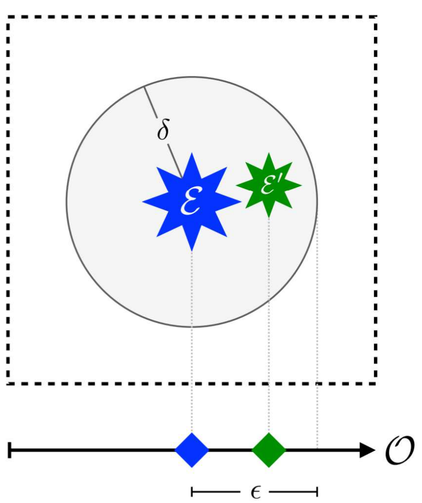




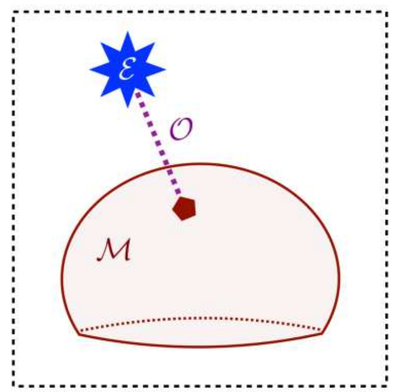
[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020;
using CMS Open Data]

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

IRC Safety is smoothness in the space of events

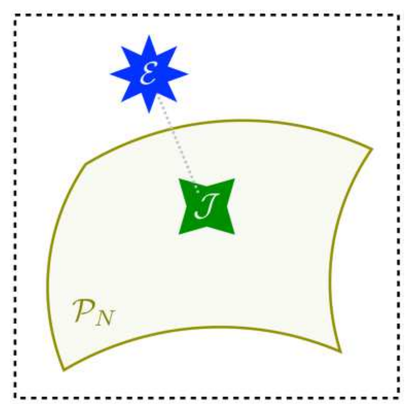


Event shapes are distances from events to manifolds.



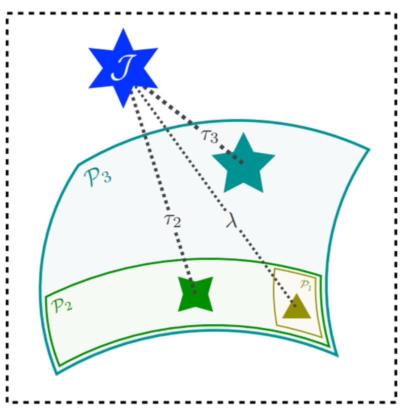
$$\mathcal{O}(\mathcal{E}) = \min_{\mathcal{E}'\in\mathcal{M}} \mathrm{EMD}_{\beta,R}(\mathcal{E},\mathcal{E}')$$

Jets are projections to few-particle manifolds.



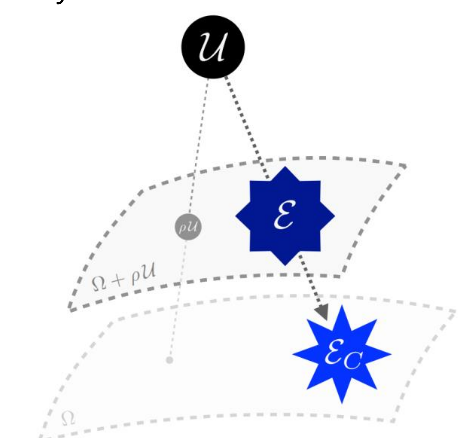
$$\mathcal{J} = \operatorname{argmin}_{\mathcal{E}'\in\mathcal{P}_N} \mathrm{EMD}_{\beta,R}(\mathcal{E},\mathcal{E}')$$

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!

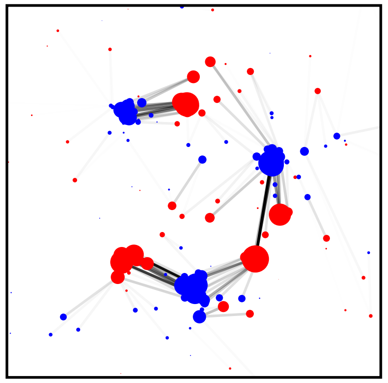
[timeline from Eric Metodiev]

Can theoretical structures be encoded directly?

Energy Flow Networks ⇔ IRC Safety + Permutations

Can strategy be defined on physical quantities?

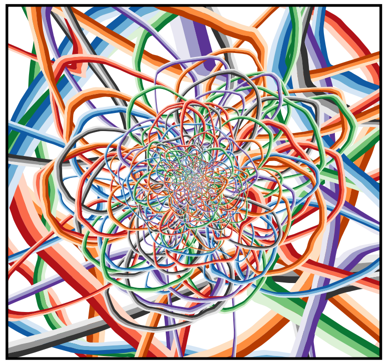
Jet Topics ⇔ Hadron-Level Approach to QCD Partons



Can we leverage unsupervised machine learning?

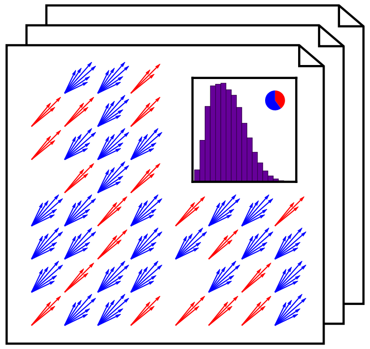
Energy Mover's Distance ⇔ Geometric Strategies for Collider Physics

Artificial Intelligence and Fundamental Physics



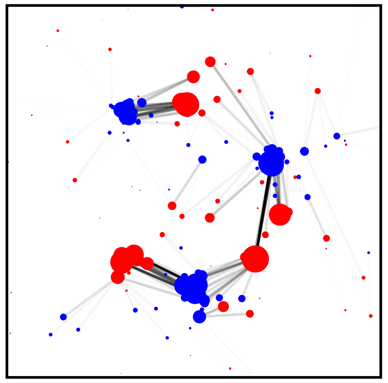
Can theoretical structures be encoded directly?

Energy Flow Networks ⇔ IRC Safety + Permutations



Can strategy be defined on physical quantities?

Jet Topics ⇔ Hadron-Level Approach to QCD Partons



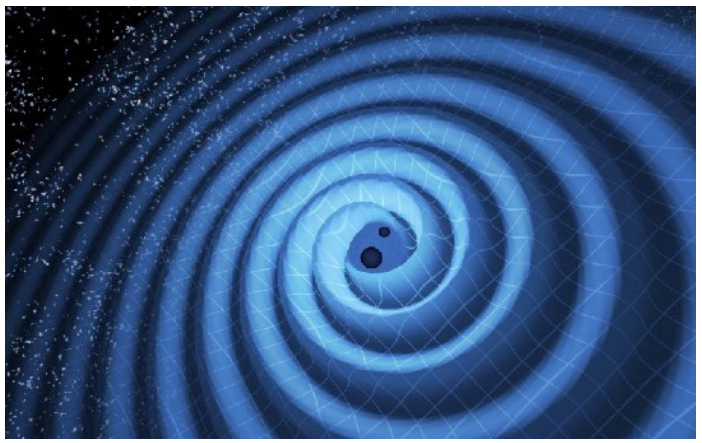
Can we leverage unsupervised machine learning?

Energy Mover's Distance ⇔ Geometric Strategies for Collider Physics

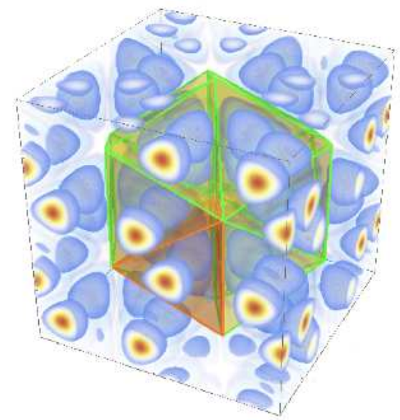
Physics insights essential for developing these tools

Artificial Intelligence ⇔ Fundamental Interactions

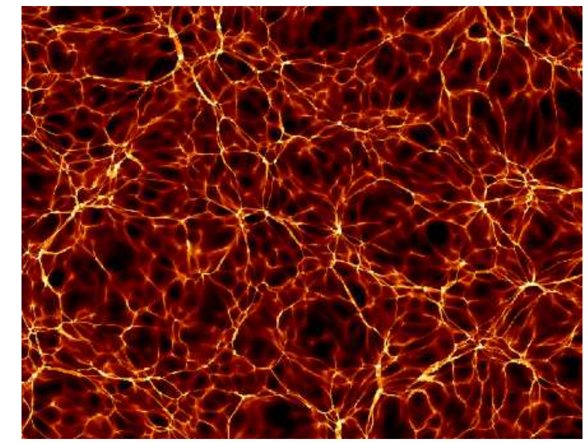
Gravitational Waves



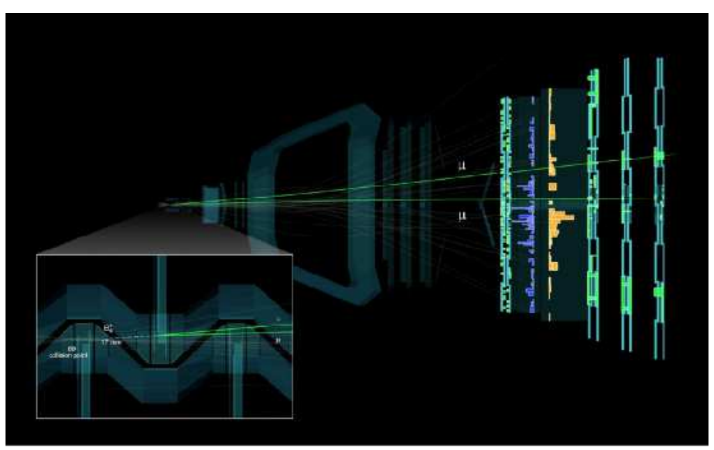
Nuclear Physics



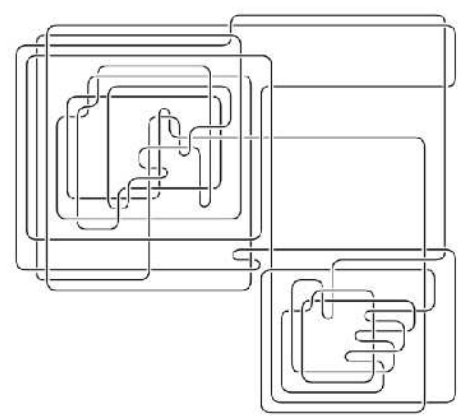
Dark Matter



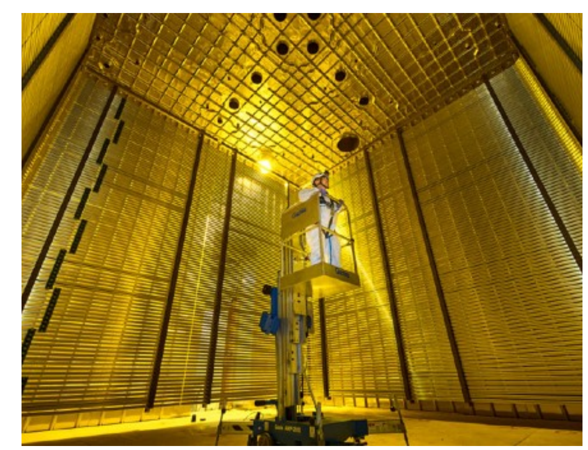
Particle Colliders



Mathematical Physics



Neutrino Detection

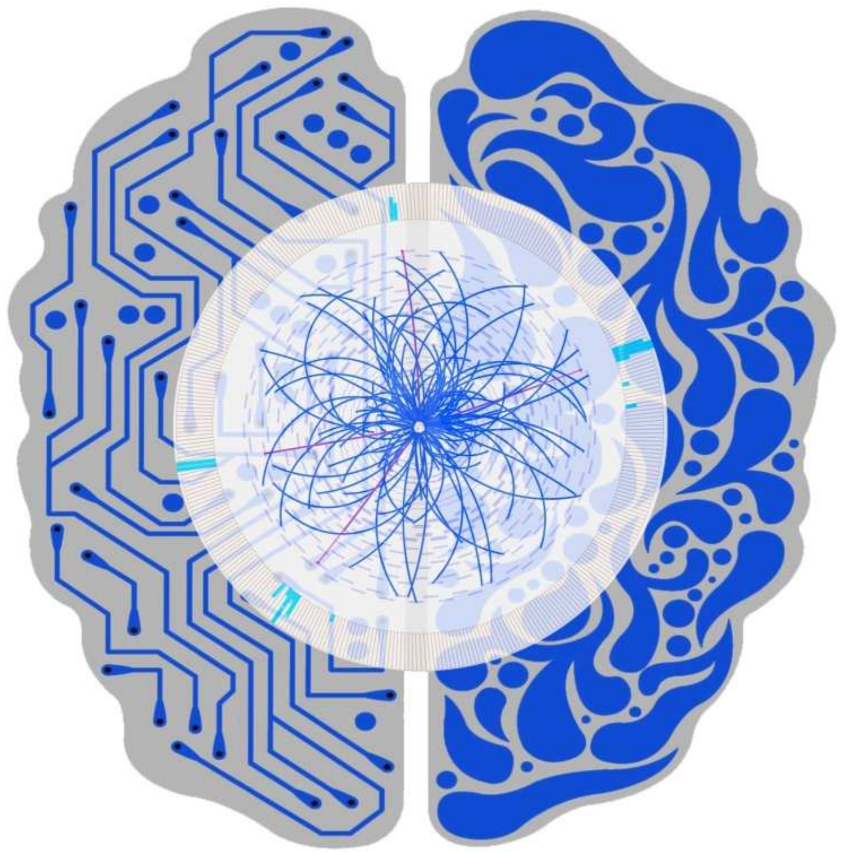


…

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

[http://iaifi.org]

Backup Slides



Can we teach a machine to "think" like a physicist?

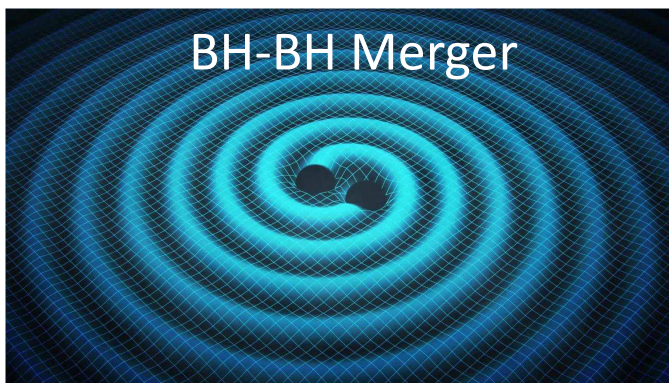

The New York Times
By Dennis Overbye
Nov. 23, 2020

Can a Computer Devise a Theory of Everything?

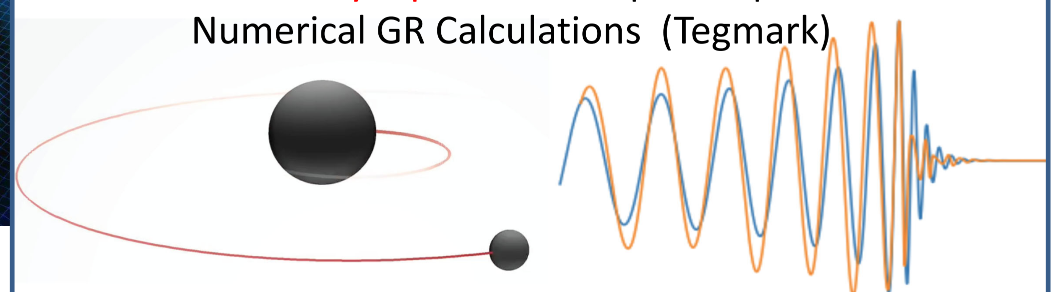
AI² for Experimental Physics

E.g. Gravitational Wave Interferometry at LIGO

Potential to enhance the physics potential of flagship experiments via improved calibrations, better quantification of uncertainties, enhanced interpretability, and sub-microsecond inference

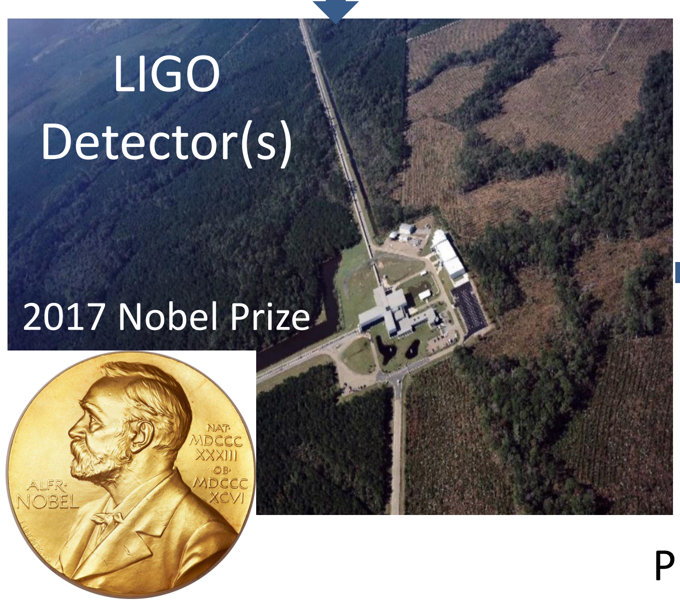


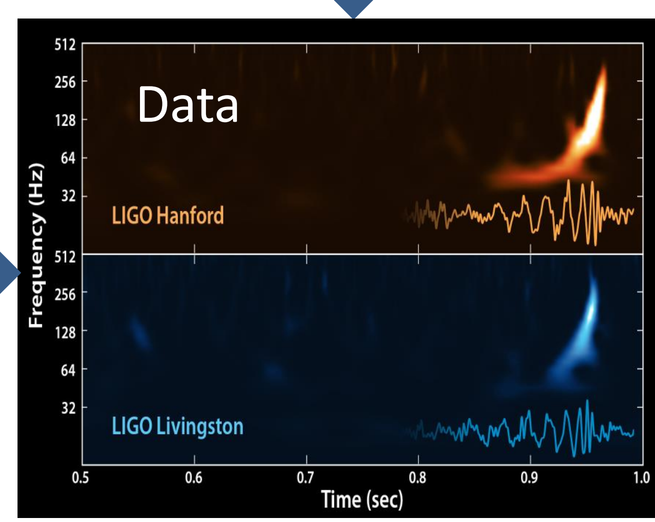
Theory Inputs: AI² to Speed Up Numerical GR Calculations (Tegmark)

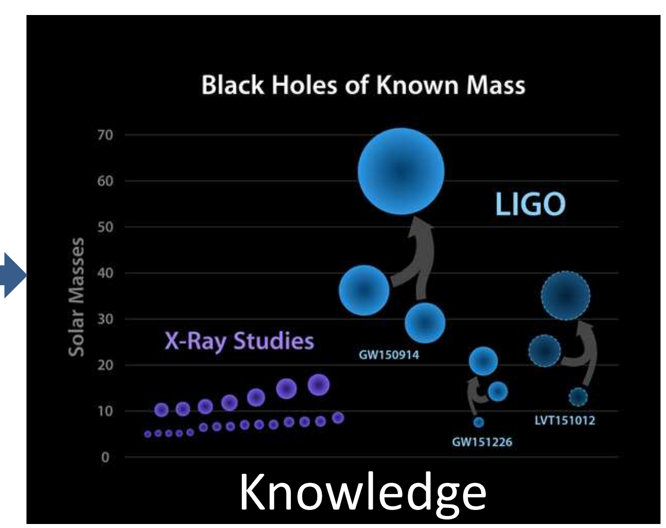


LHC Fast-ML to speed up multi-messenger astrophysics (Harris)

Autoencoders improve LIGO's sensitivity by 20% [2005.06534]







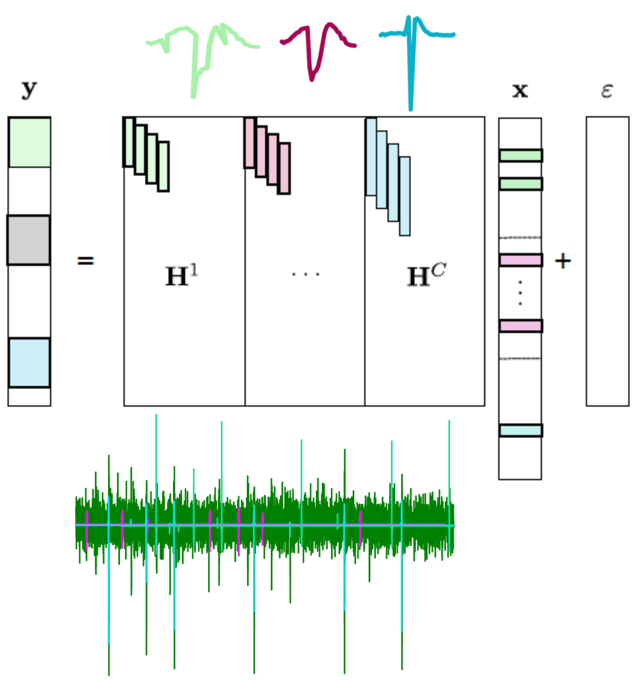
Propose using RL for Noise Reduction (Barsotti, Rakhlin)

AI^2 for Foundational AI

E.g. Deconvolution Across Disciplines

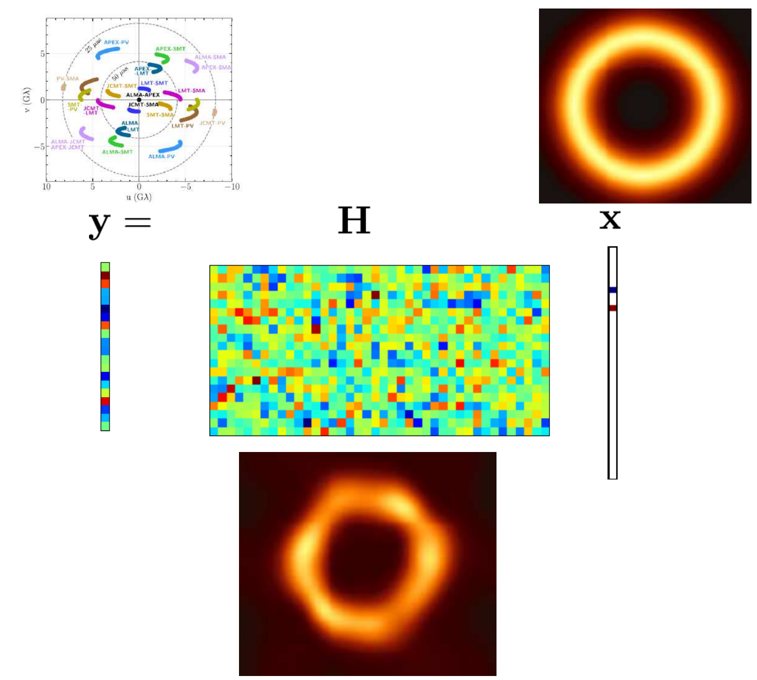
The unique features of physics applications and the power of physics principles offer compelling research opportunities to advance the field of AI research itself

Sparse Coding Networks and Neuronal Source Separation (Ba)



$\Leftrightarrow$

Event Horizon Telescope and Black Hole Imaging (Freeman)

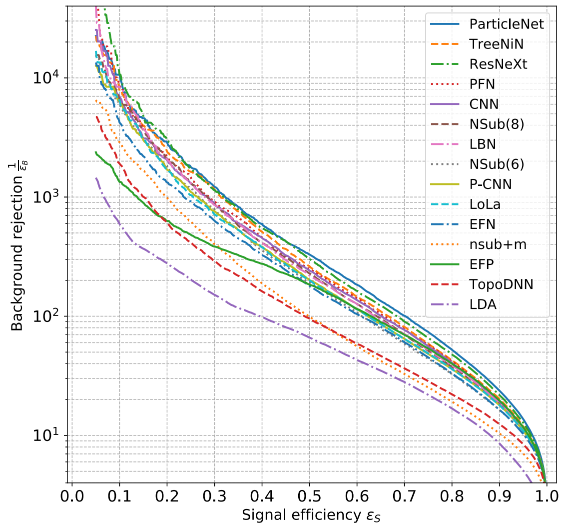


Capitalize on physics priors and interpretability for improved robustness
Leverage tools from physics to explain ability of networks to generalize

Optimization in Collider Physics

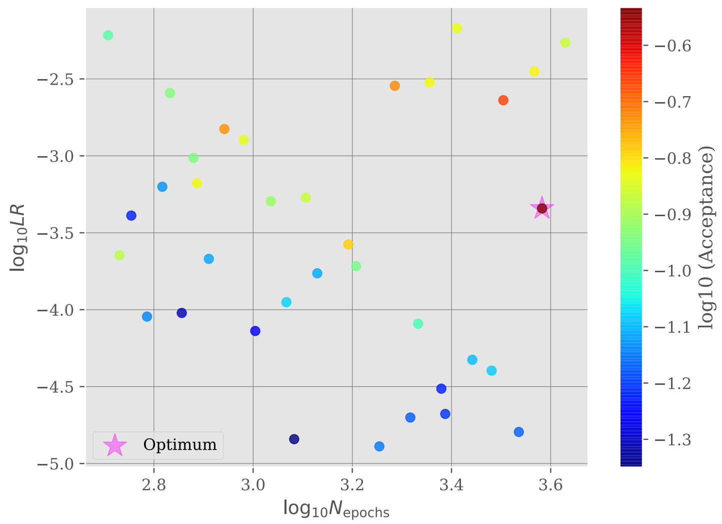
This slide is far from exhaustive

Jet Classification



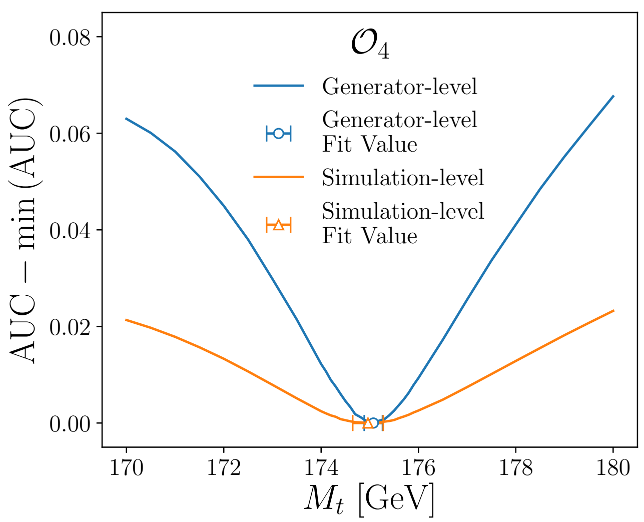
[e.g. Kasieczka, Plehn, et al., SciPost 2019]

Phase Space Integration



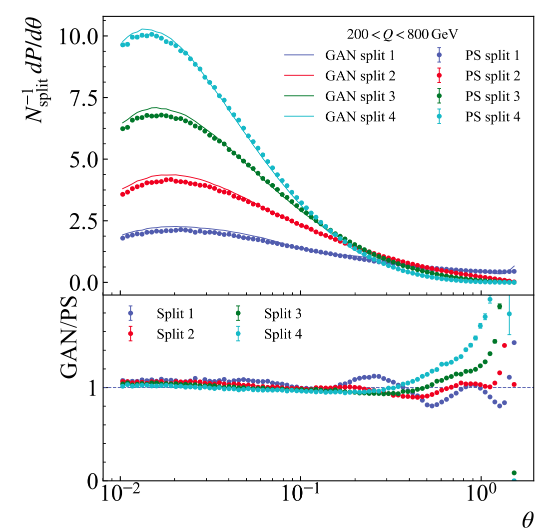
[e.g. Gao, Höche, Isaacson, Krause, Schulz, PRD 2020]

Parameter Estimation



[e.g. Andreassen, Hsu, Nachman, Suaysom, Suresh, PRD 2021]

Parton Shower Modeling



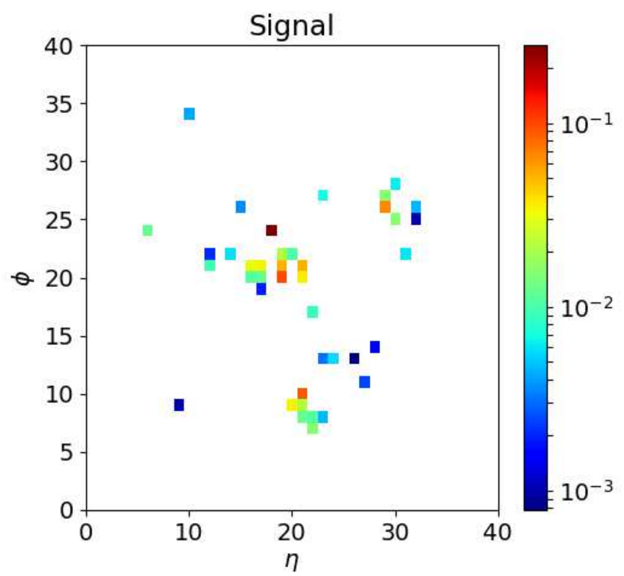
[e.g. Lai, Neill, Płoskoń, Ringer, arXiv 2020]

...

Jet Representations

Pixelized Image

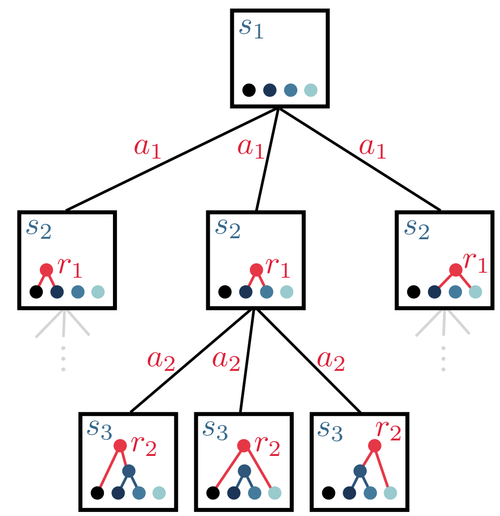
Calorimetry



[review in Kagan, arXiv 2020]

Hierarchical Tree

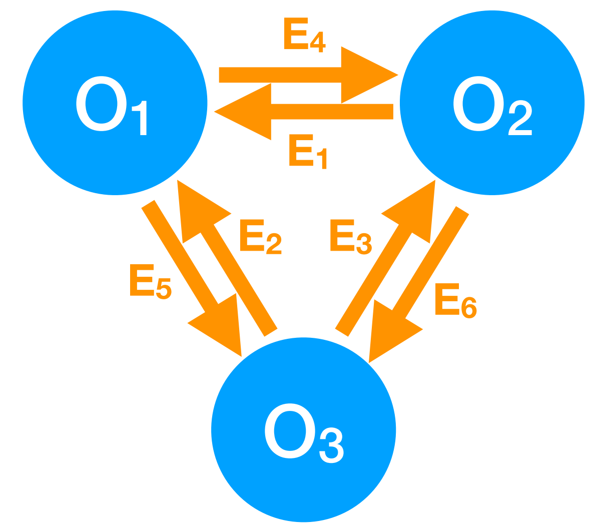
Binary Splittings



[e.g. Brehmer, Macaluso, Pappadopulo, Cranmer, NeurIPS 2020]

Graphs

Pairwise Interactions

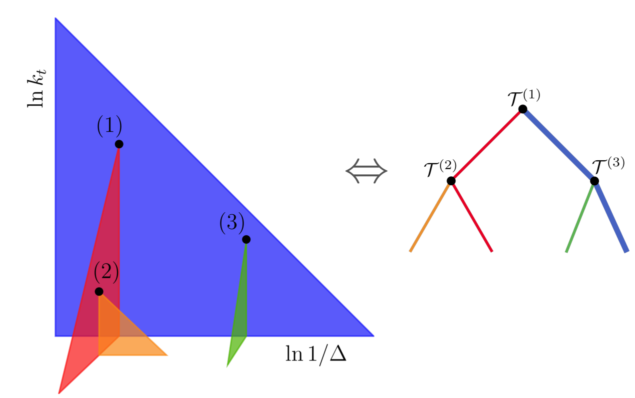


[e.g. Moreno, Cerri, Duarte, Newman, Nguyen, Periwal, Pierini, Serikova, Spiropulu, Vlimant, EPJC 2020]

Imposes implicit theoretical prior; affects choice of network architecture

E.g. recent progress with
Lund Plane + Graph Networks

[Dreyer, Qu, JHEP 2021]



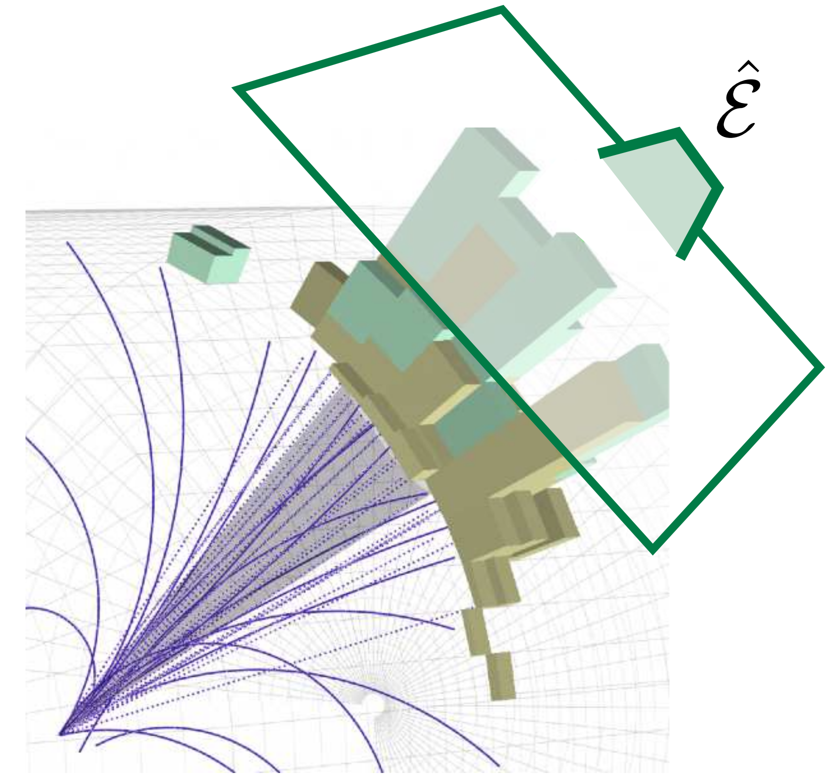
Jets as Weighted Point Clouds

- Energy-Weighted Directions

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

Energy Direction

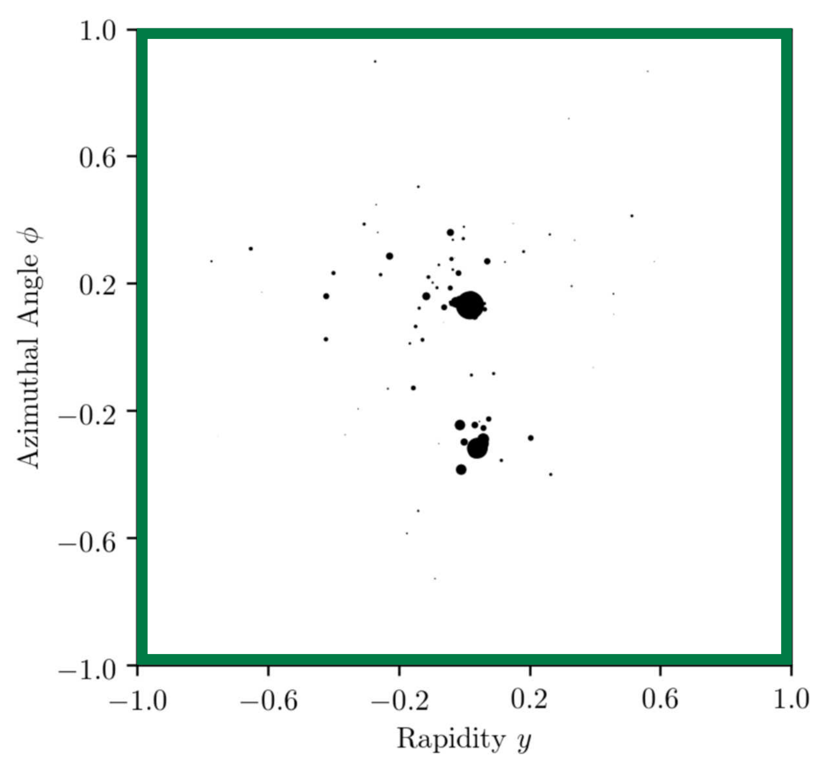
(suppressing "unsafe" charge/flavor information)



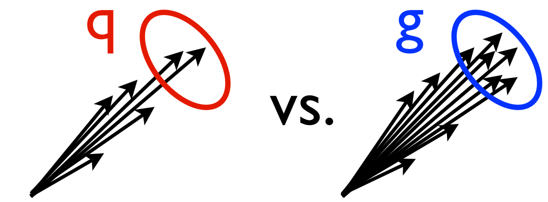
- Equivalently: Energy Density

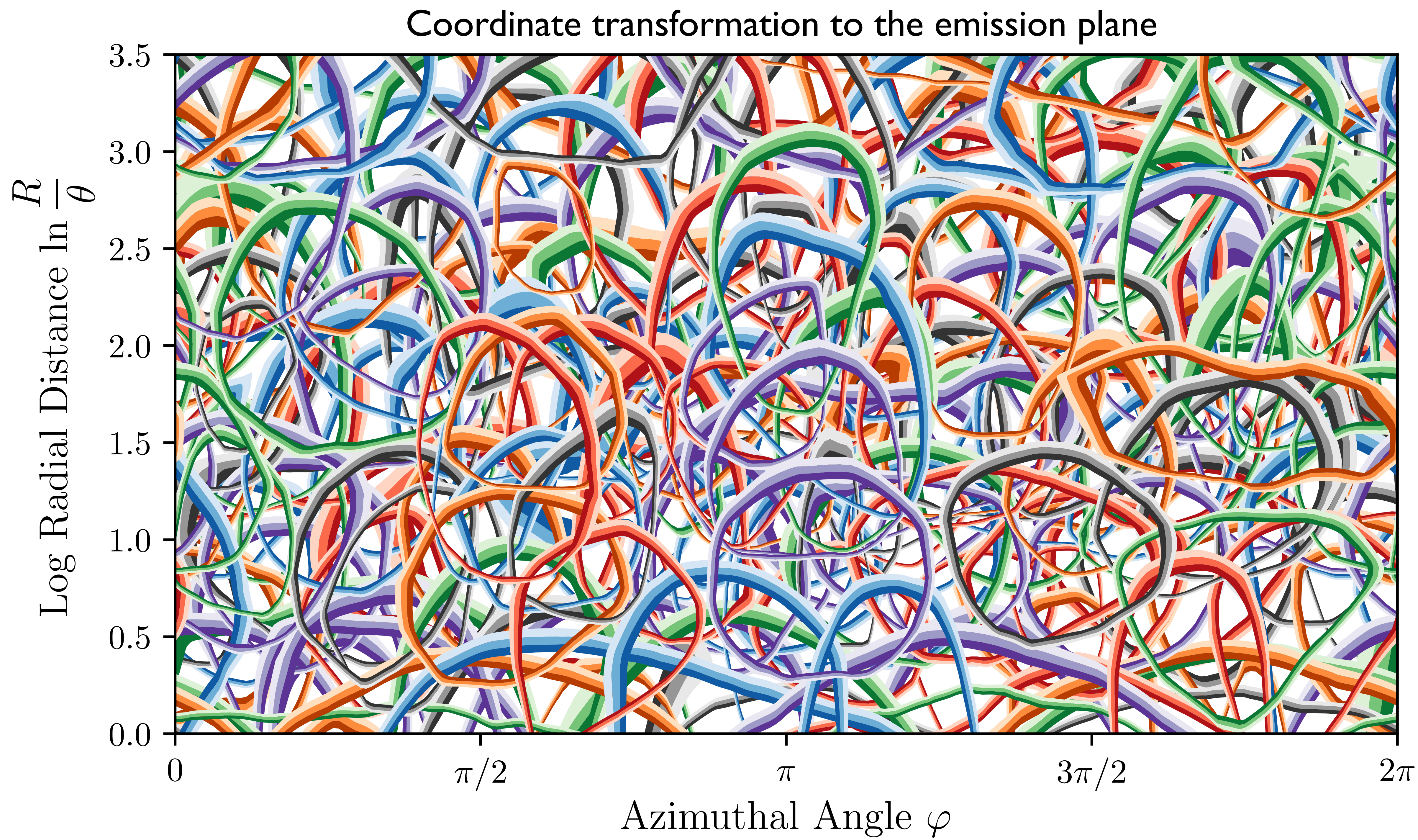
$$\rho(\hat{n}) = \sum_{i \in \mathcal{J}} E_i \, \delta^{(2)}(\hat{n} - \hat{n}_i)$$

Energy Direction



Ready for the MAM?

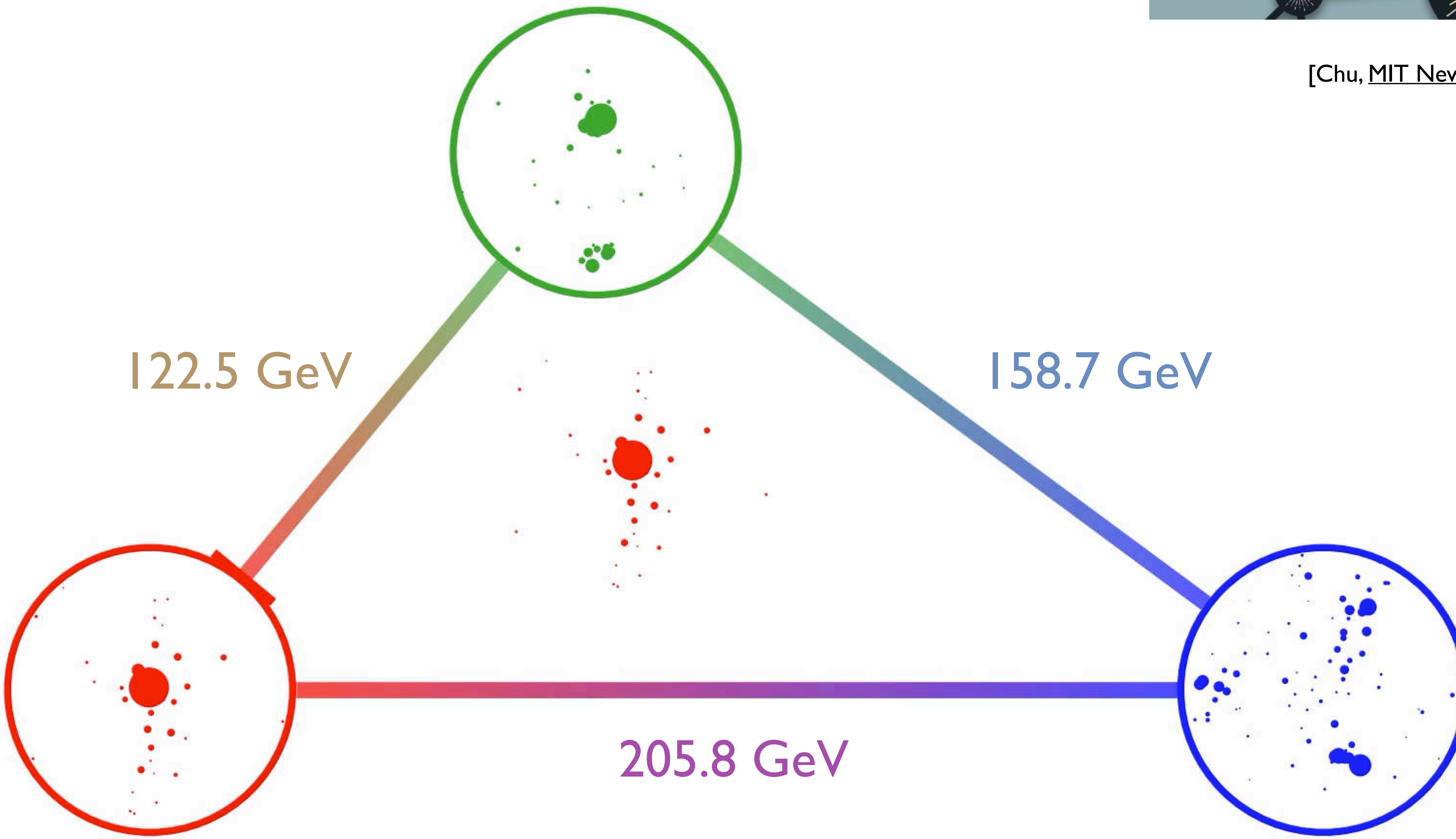




[Komiske, Metodiev, JDT, JHEP 2019; see also Dreyer, Salam, Soyez, JHEP 2018]

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]

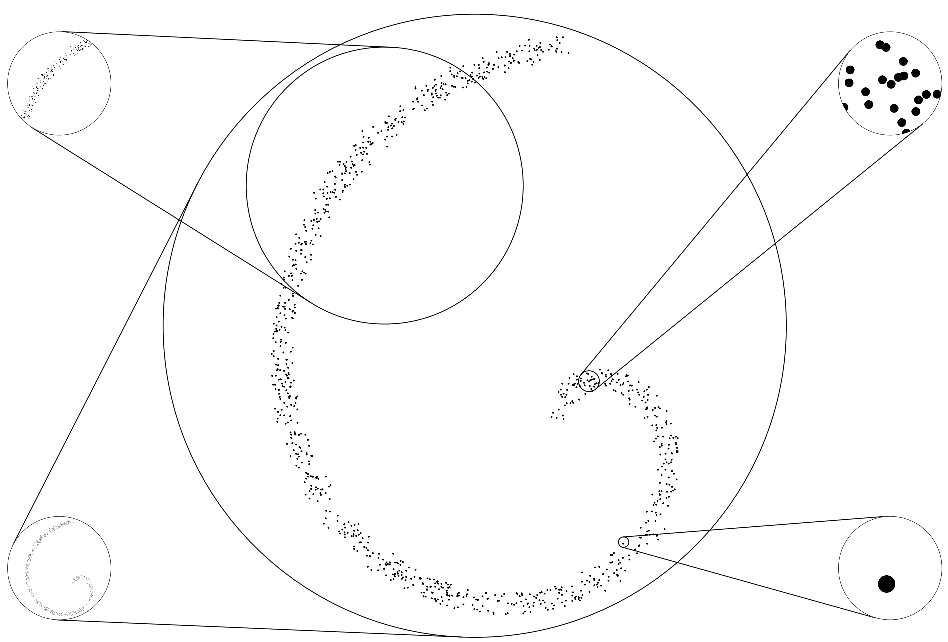
Dimensionality of Space of Jets

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

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

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

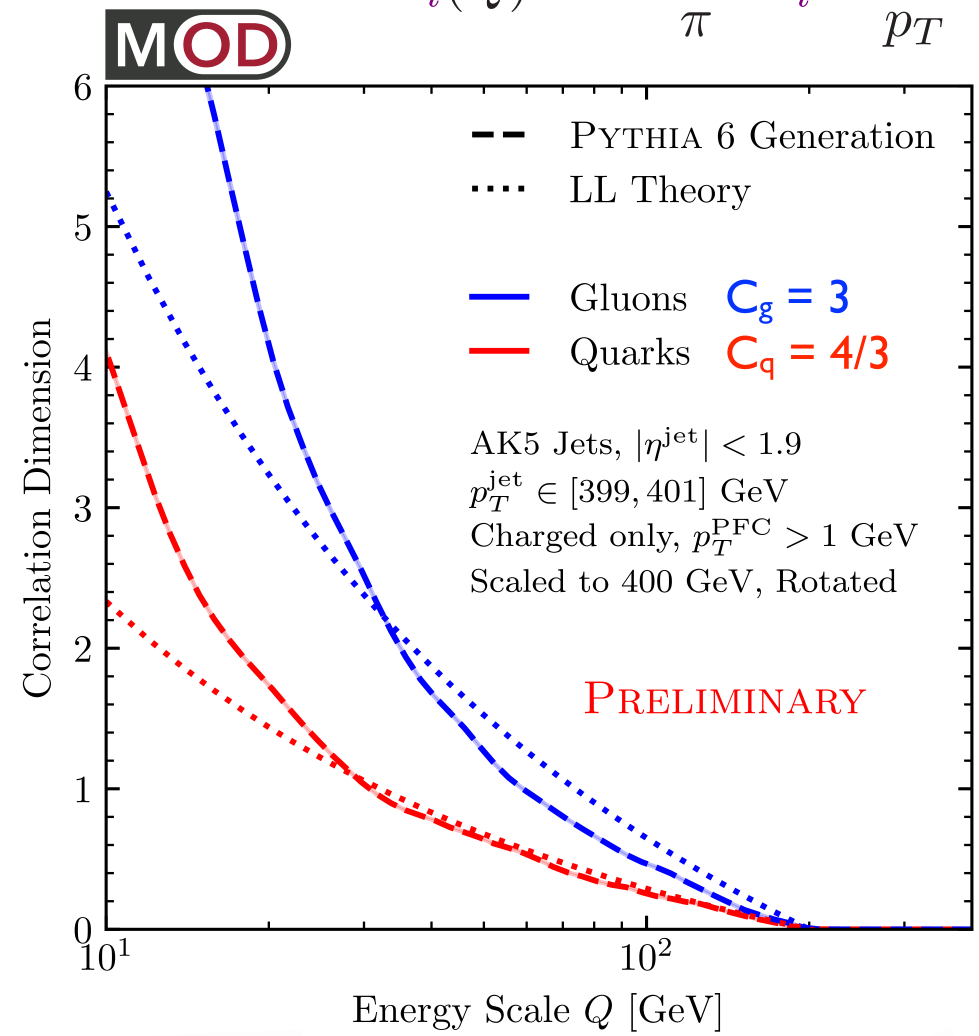
dim≃1 dim≃2

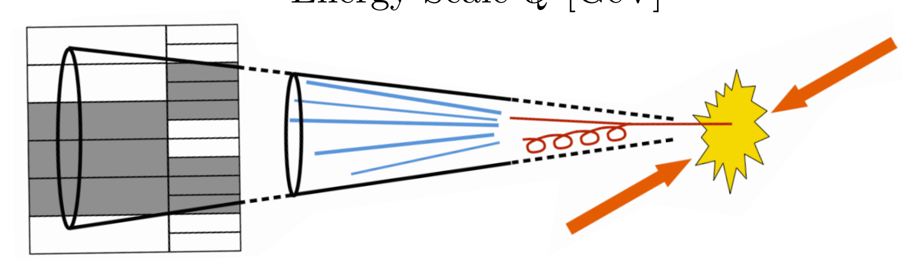


dim≃2 dim≃0

(eventually 0)

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





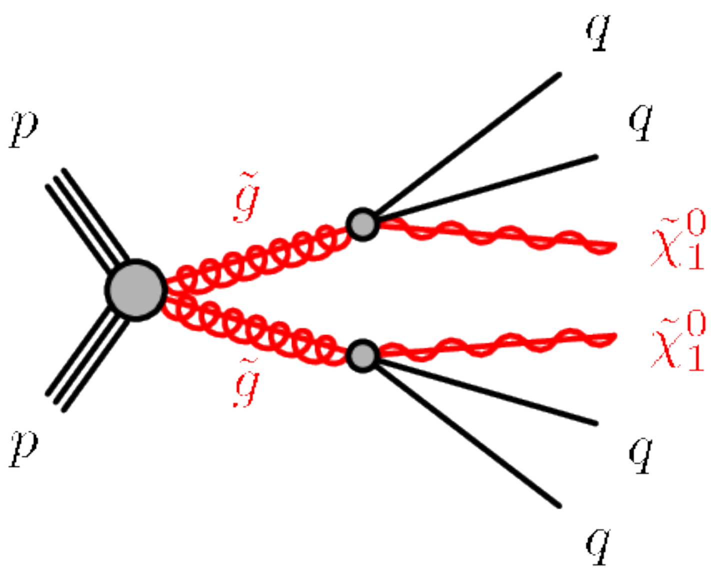
E.g. Search for Supersymmetry

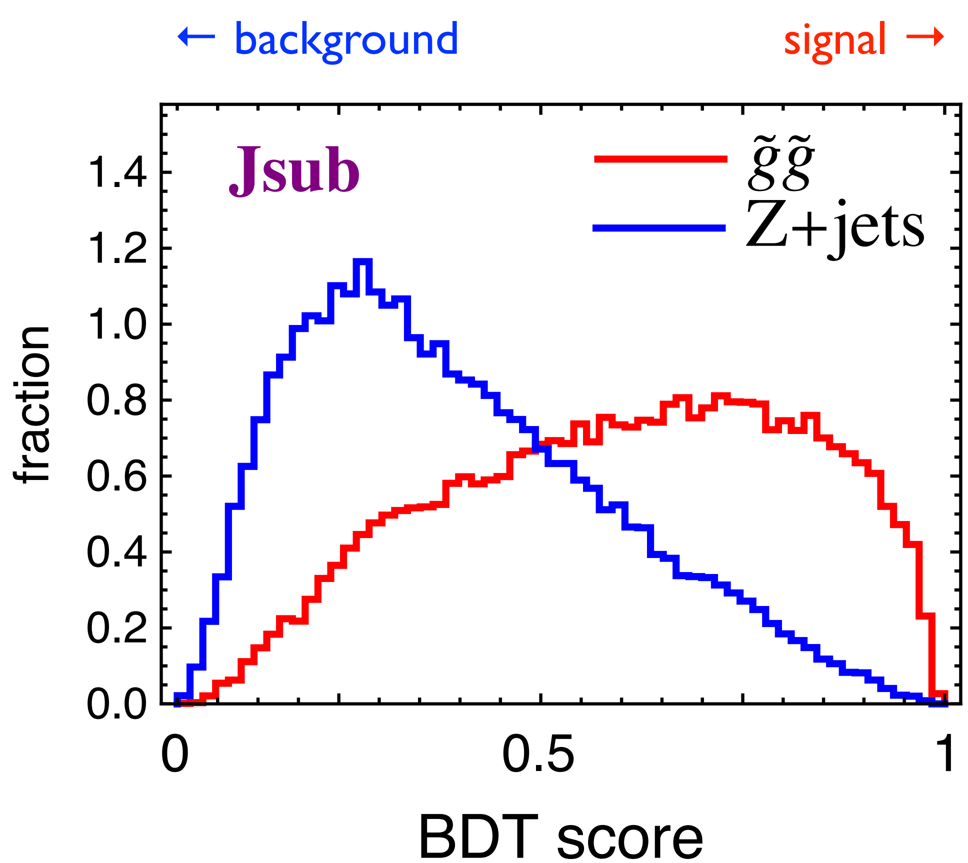
Classifier: Boosted decision tree (for each of 4 jets)

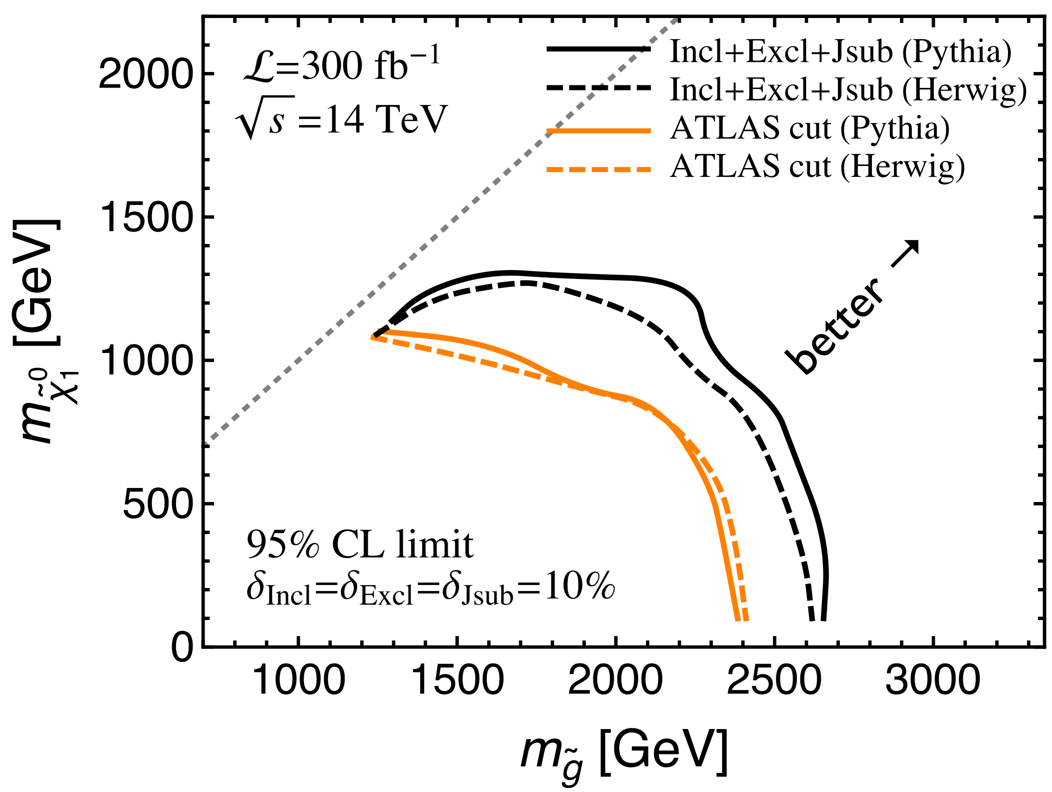
Inputs: Jet mass, width, track multiplicity

Signal: **Quark enriched**

Background: **Gluon enriched**





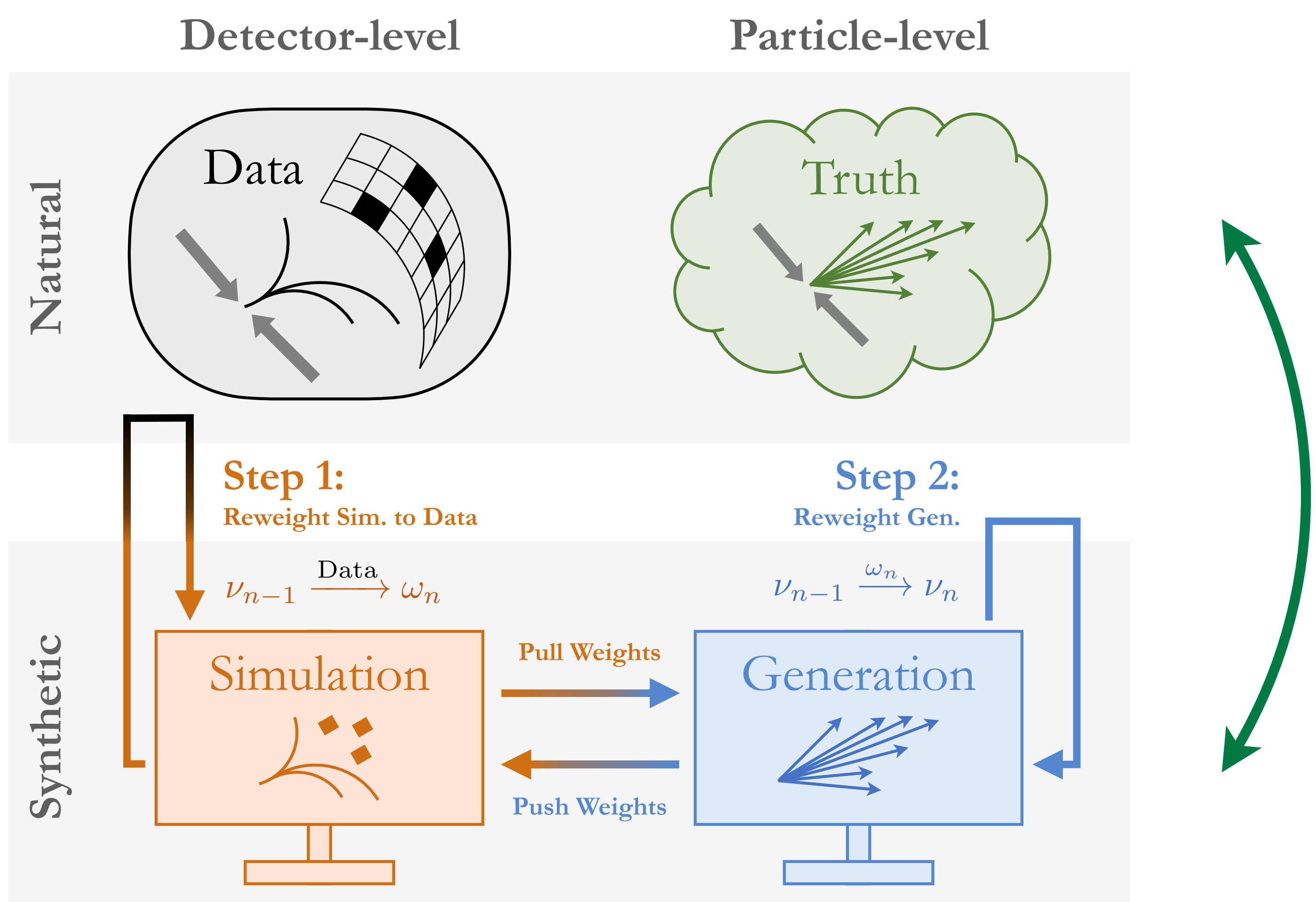


[Bhattacherjee, Mukhopadhyay, Nojiri, Sakakie, Webber, JHEP 2017]

E.g. Detector Unfolding

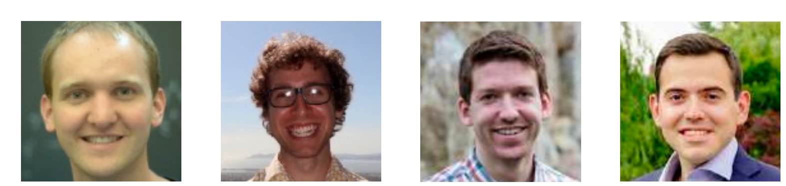
OmniFold

Multi-dimensional unbinned detector corrections via iterated application of likelihood ratio trick



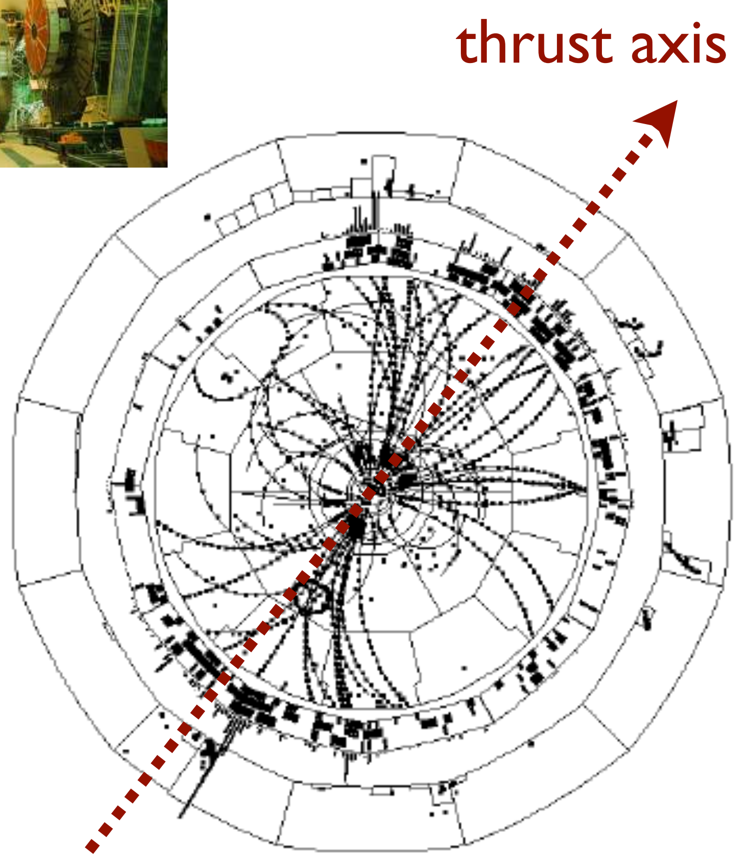
Use ML to compute reweighting factors

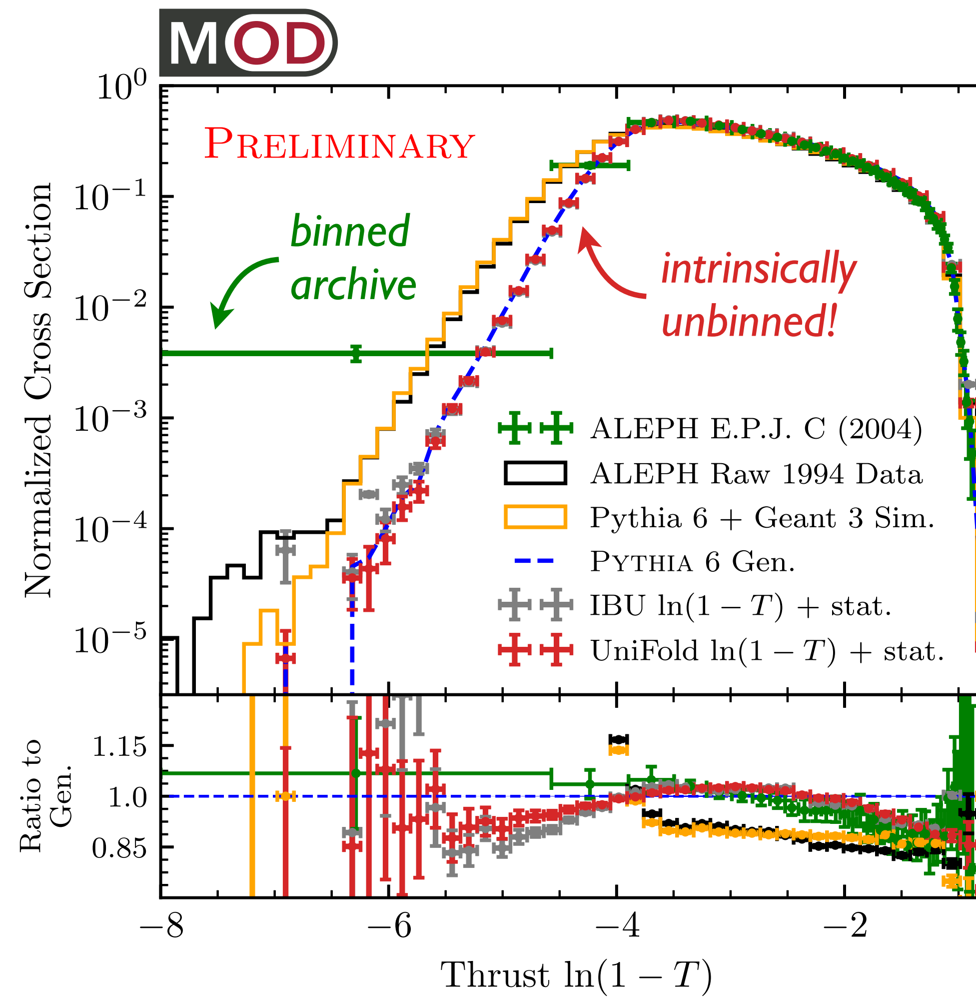
[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020]

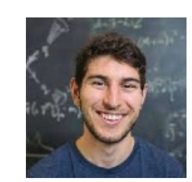


E.g. Detector Unfolding

Back to the Future with ALEPH Archival Data

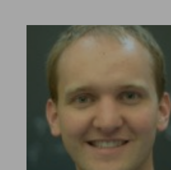
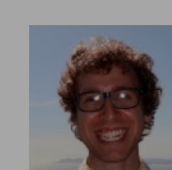




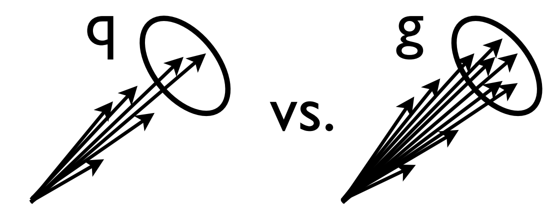


[talk by Badea, ICHEP 2020; cf. ALEPH, EPJC 2004]
[see also Badea, Baty, Chang, Innocenti, Maggi, McGinn, Peters, Sheng, JDT, Lee, PRL 2019]

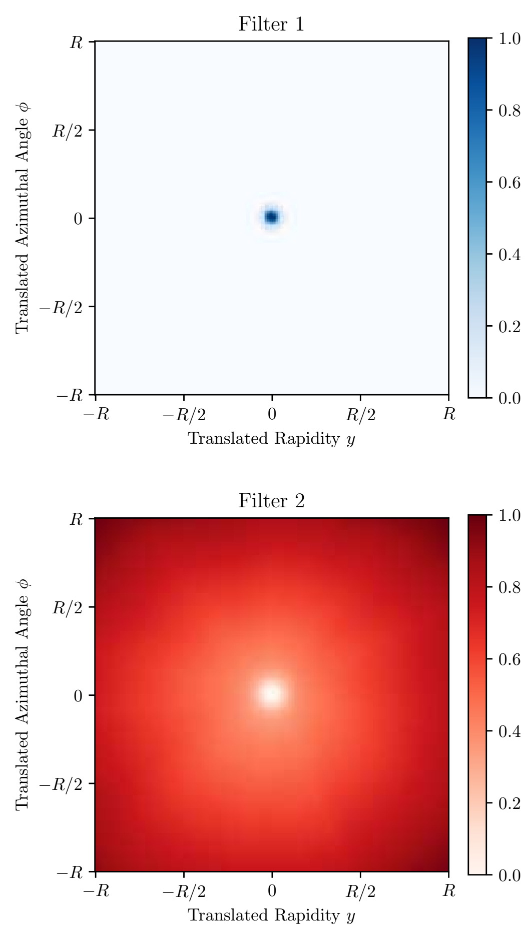
[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020]

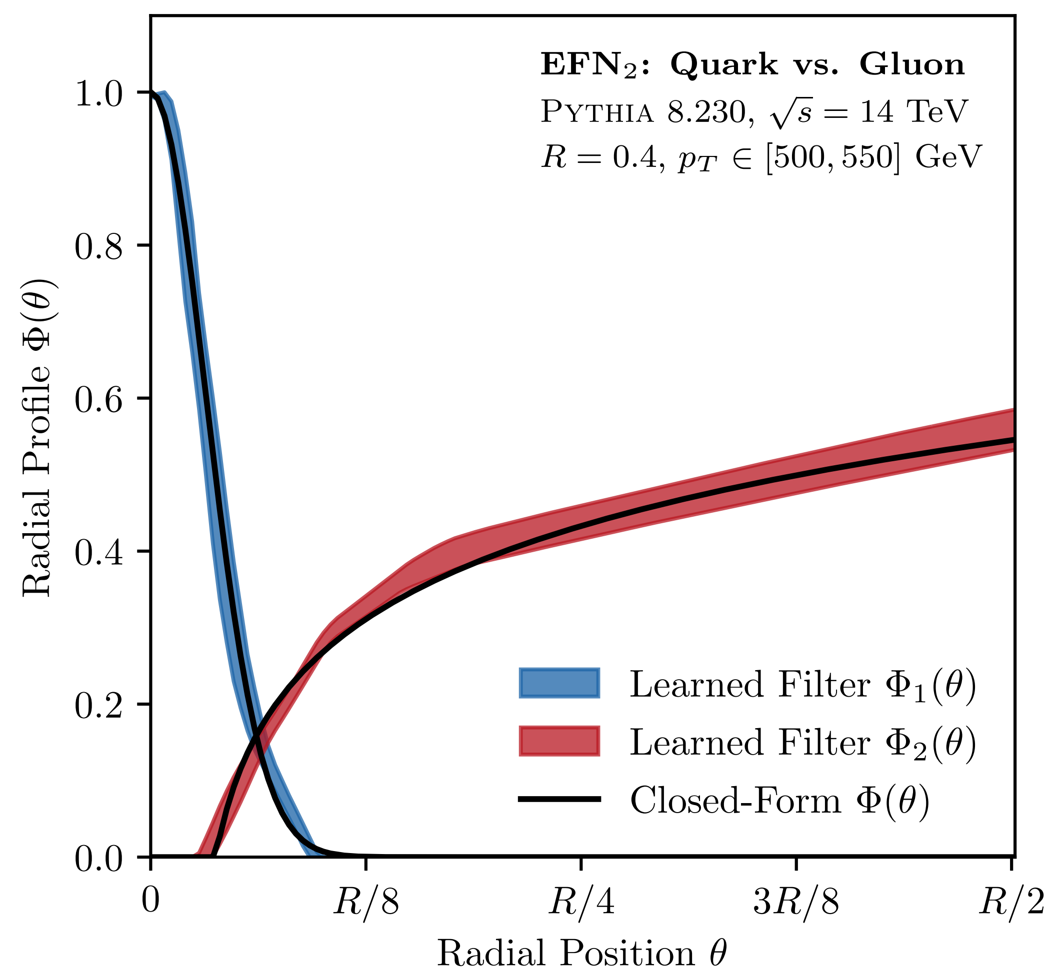



Learning from the Machine



For ℓ = 2, EFN learns radial moments: $\sum_{i \in \text{jet}} z_i f(\theta_i)$ cf. Angularities: f(θ) = θ^β

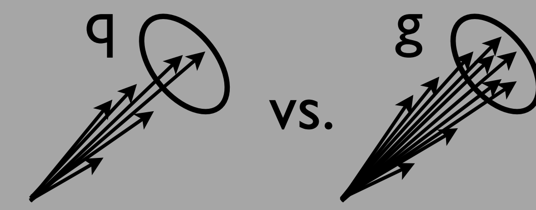




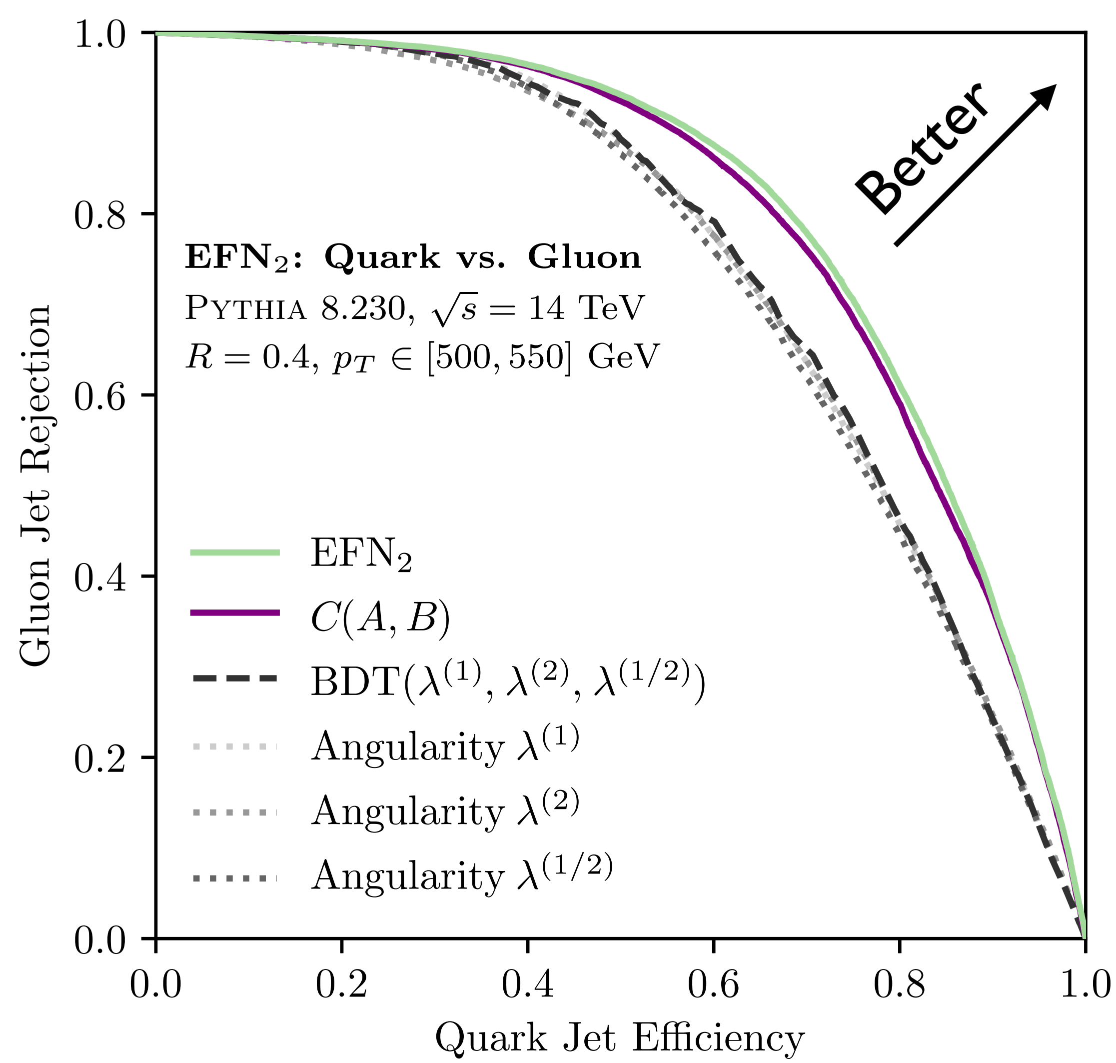
Traditional QCD observables emphasize homogeneous angular scaling
But EFN reveals that likelihood ratio exhibits collinear/wide-angle separation

[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 learns radial moments: $\sum_{i \in \text{jet}} z_i f(\theta_i)$ cf. Angularities: f(θ) = θ^β

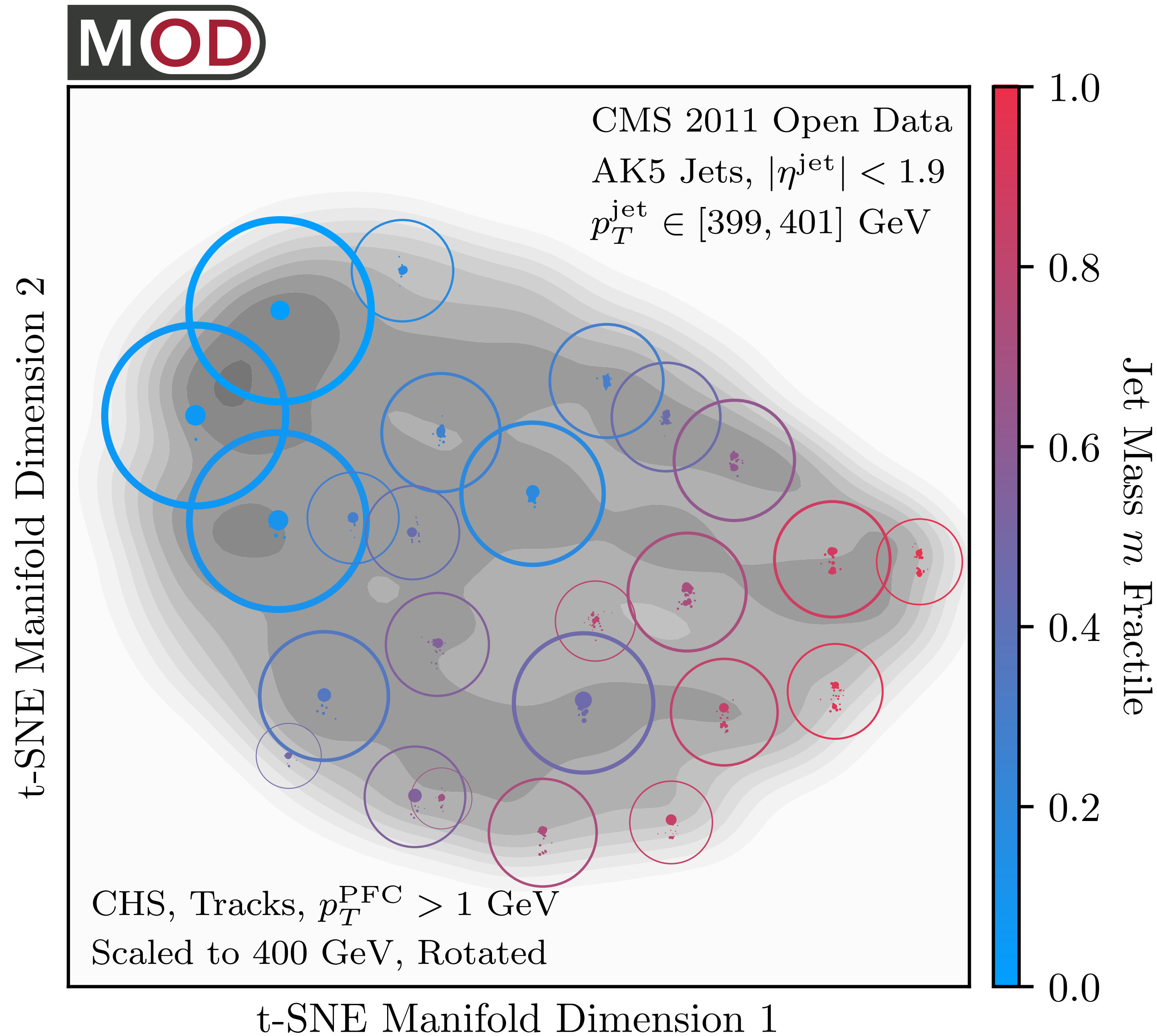


EFN outperformed a domain expert (i.e. me)

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

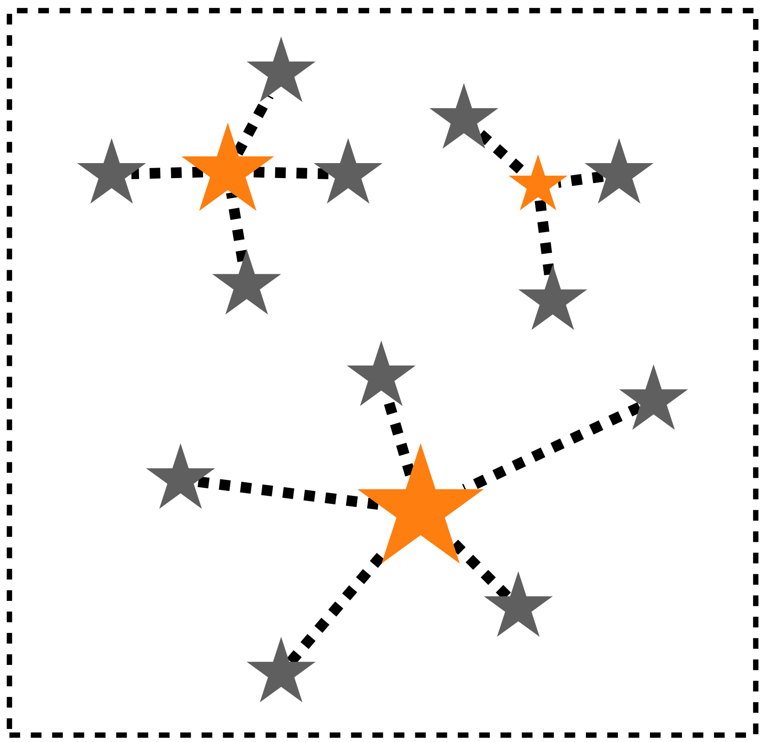
[Komiske, Metodiev, JDT, JHEP 2019;
cf. Larkoski, JDT, Waalewijn, JHEP 2014; using Berger, Kucs, Sterman, PRD 2003; Ellis, Vermilion, Walsh, Hornig, Lee, JHEP 2010]

Most Representative Jets



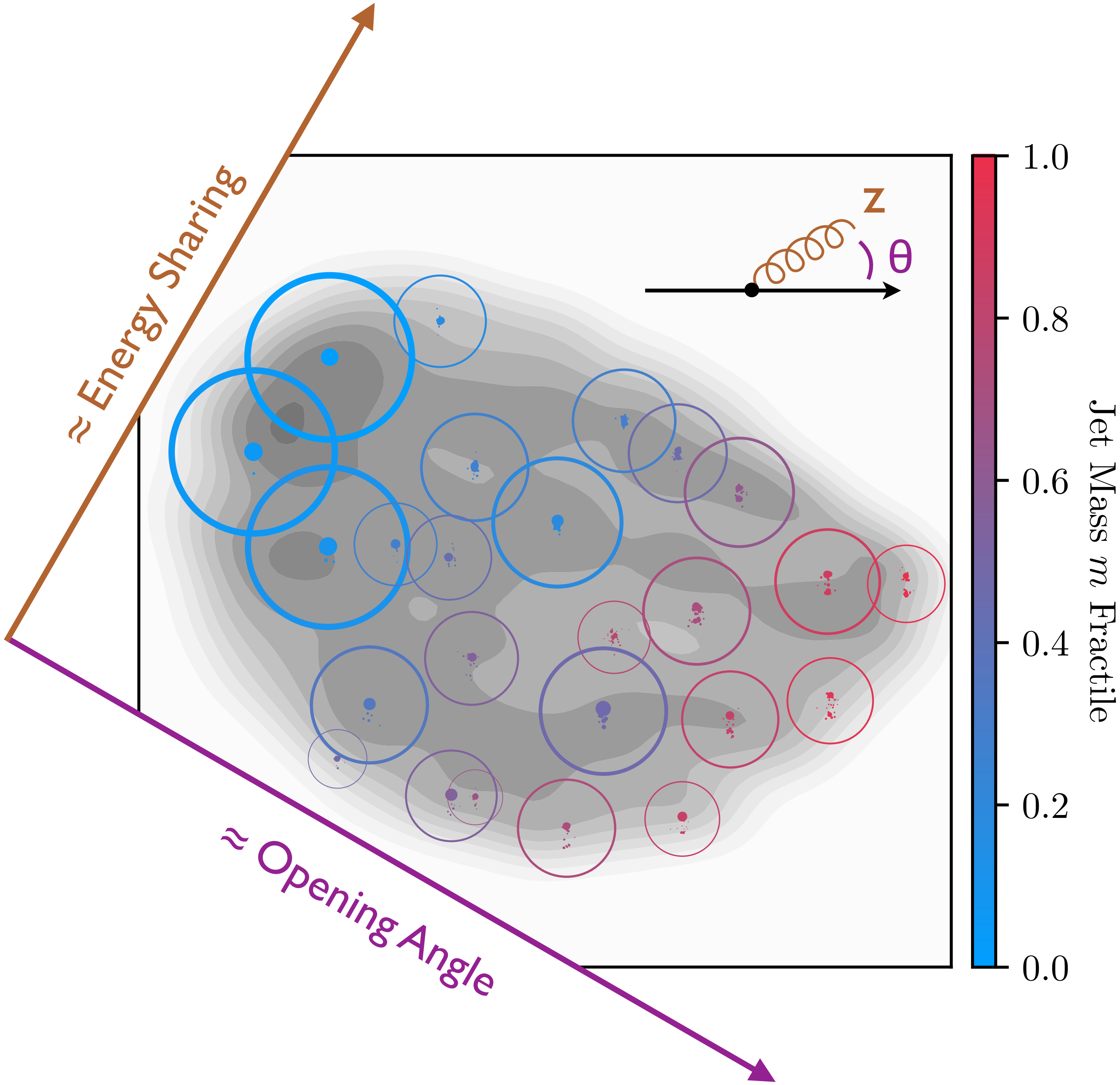
k-medoids

Arranged via t-SNE



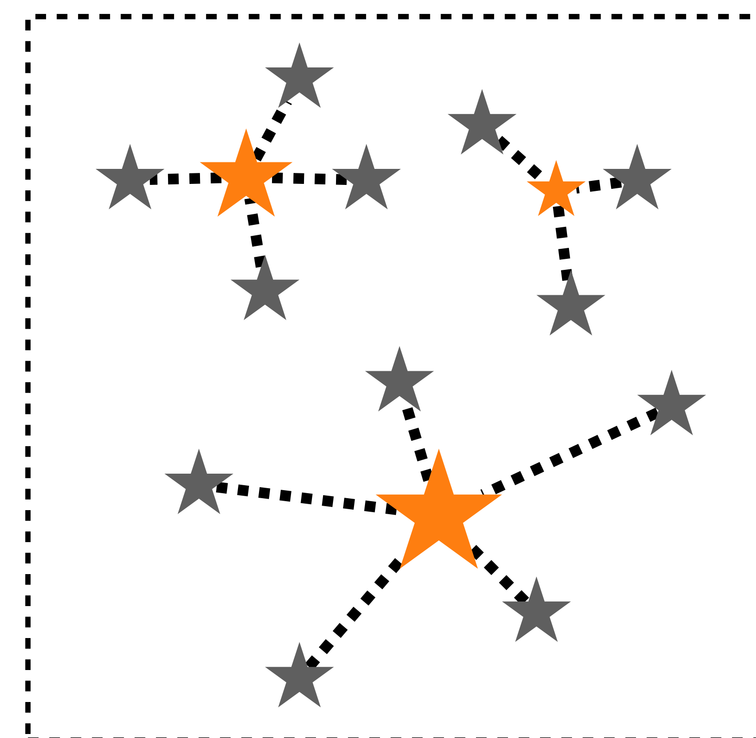
[Komiske, Mastandrea, Metodiev, Naik, JDT, PRD 2020;
using van der Maaten, Hinton, JMLR 2008]

Most Representative Jets



k-medoids

Arranged via t-SNE



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