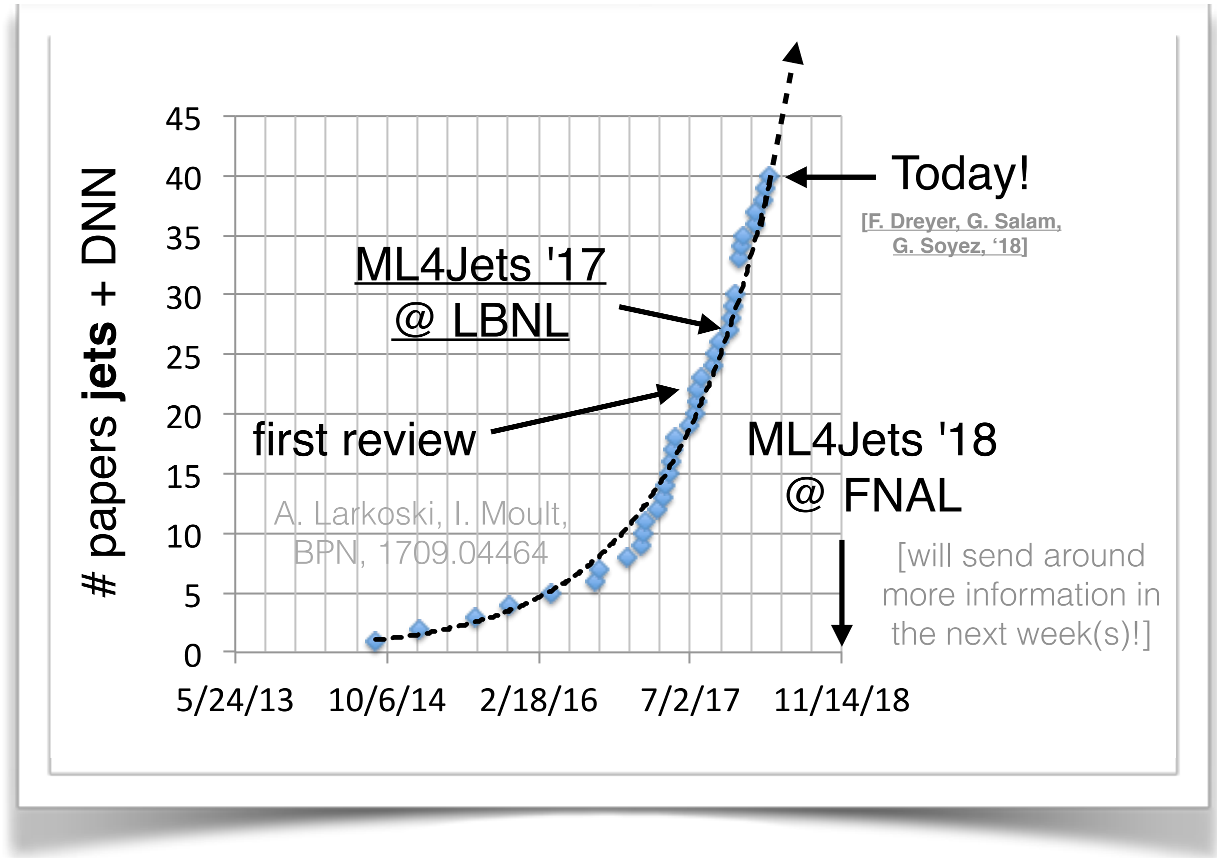
Deep Sets for Particle Jets

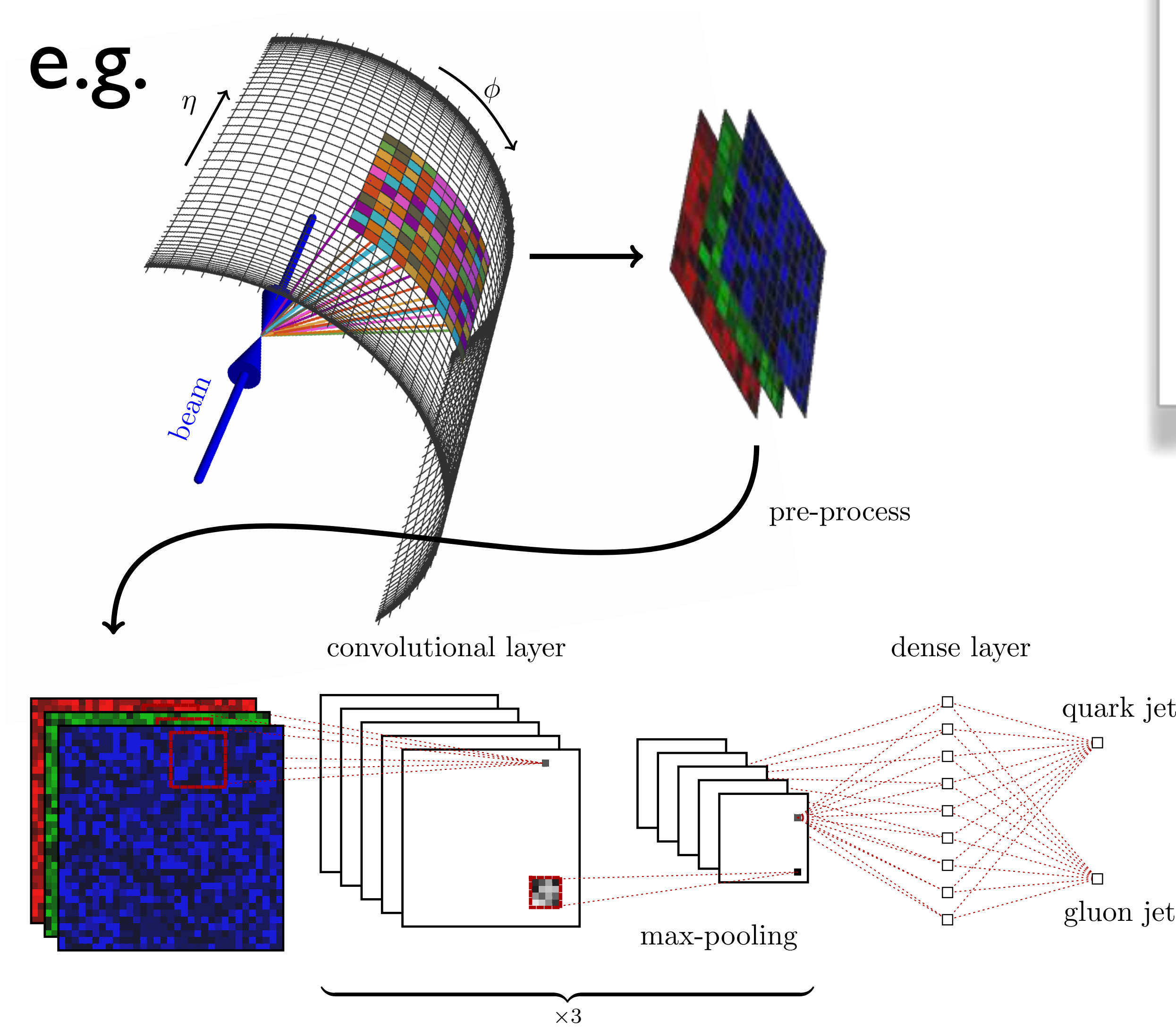
Jesse Thaler

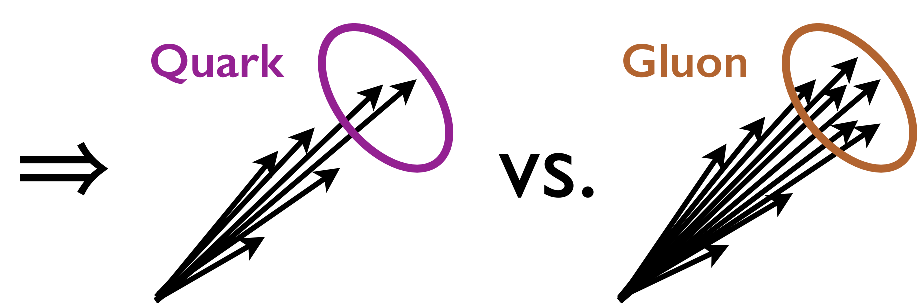
NYU CCPP Seminar — October 17, 2018

The Rise of Machine Learning for Jets



e.g.





[e.g. Komiske, Metodiev, Schwartz, 1612.01551; Nachman, Boost 2018 Talk, July 20, 2018; reviews in Larkoski, Moult, Nachman, 1709.04464; Guest, Cranmer, Whiteson, 1806.11484]

My Perspective c. 2016

“Deep Learning” vs. “Deep Thinking”

My Perspective c. 2018

BOOST 2018

10th International Workshop on Boosted Objects Phenomenology, Reconstruction and Searches

“Deep Learning” ~~vs.~~ & “Deep Thinking”

New first-principles studies of QCD facilitated by advances in mathematics, statistics, and computer science

Desired Outcomes ⇔ Algorithms/Observables

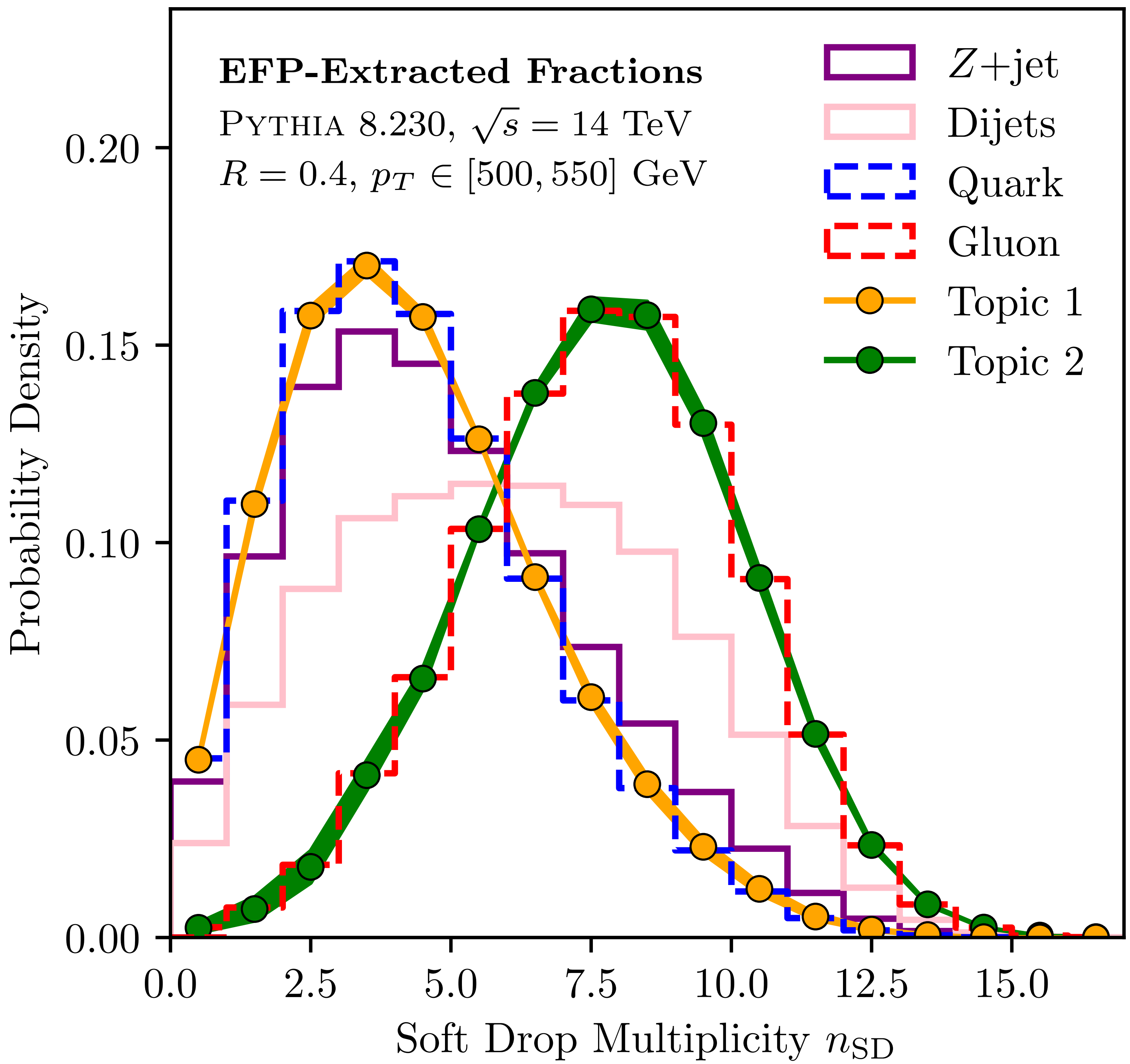
Proximate Reasons for My Conversion

Patrick Komiske

Eric Metodiev

plus Ben Nachman, Kyle Cranmer, Daniel Whiteson, Mike Williams, Matt Schwartz, Dan Roberts, Phiala Shanahan, …

Physics Reasons for My Conversion (1)

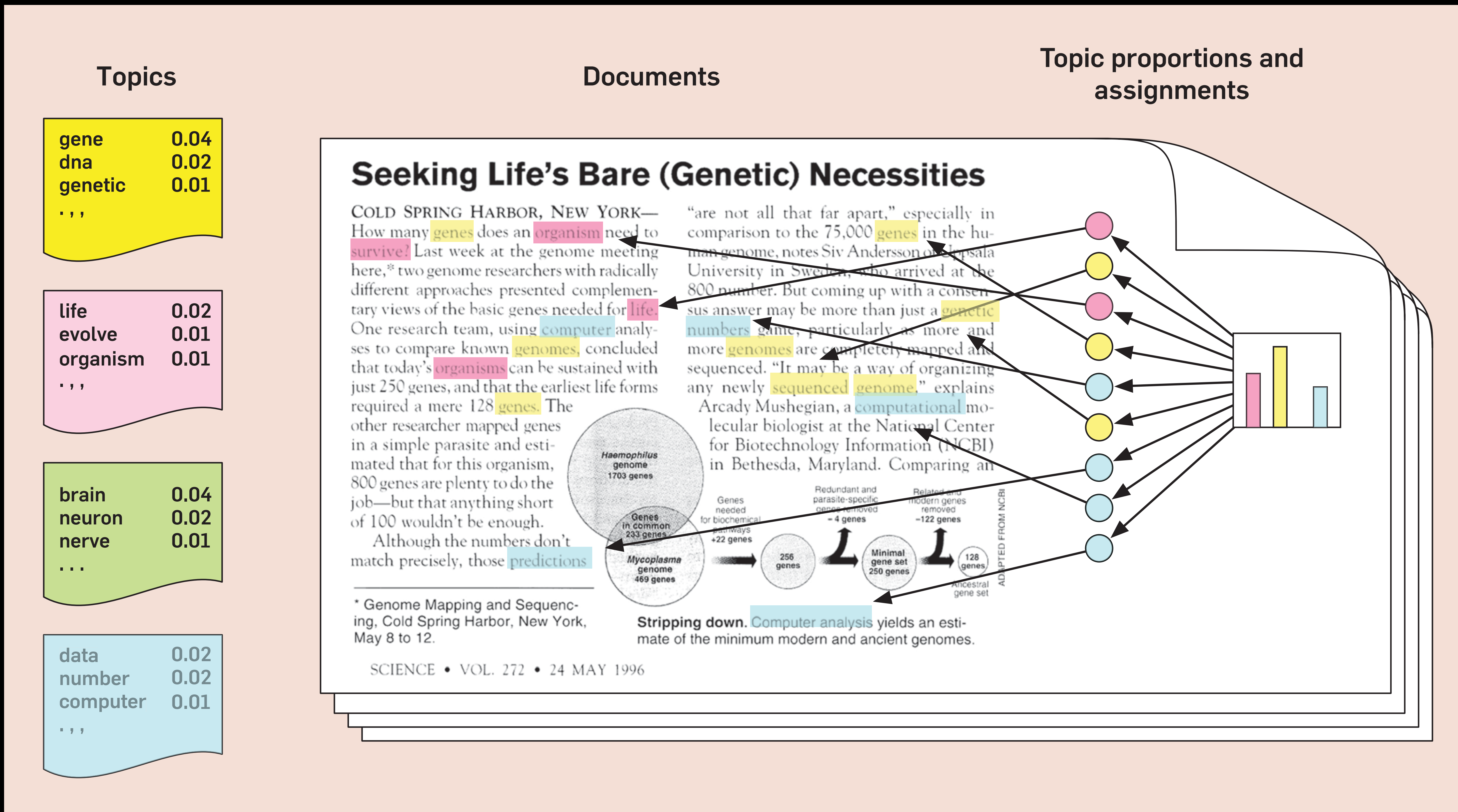


For Offline Discussions

First-principles QCD meets blind source separation with same underlying structure as…

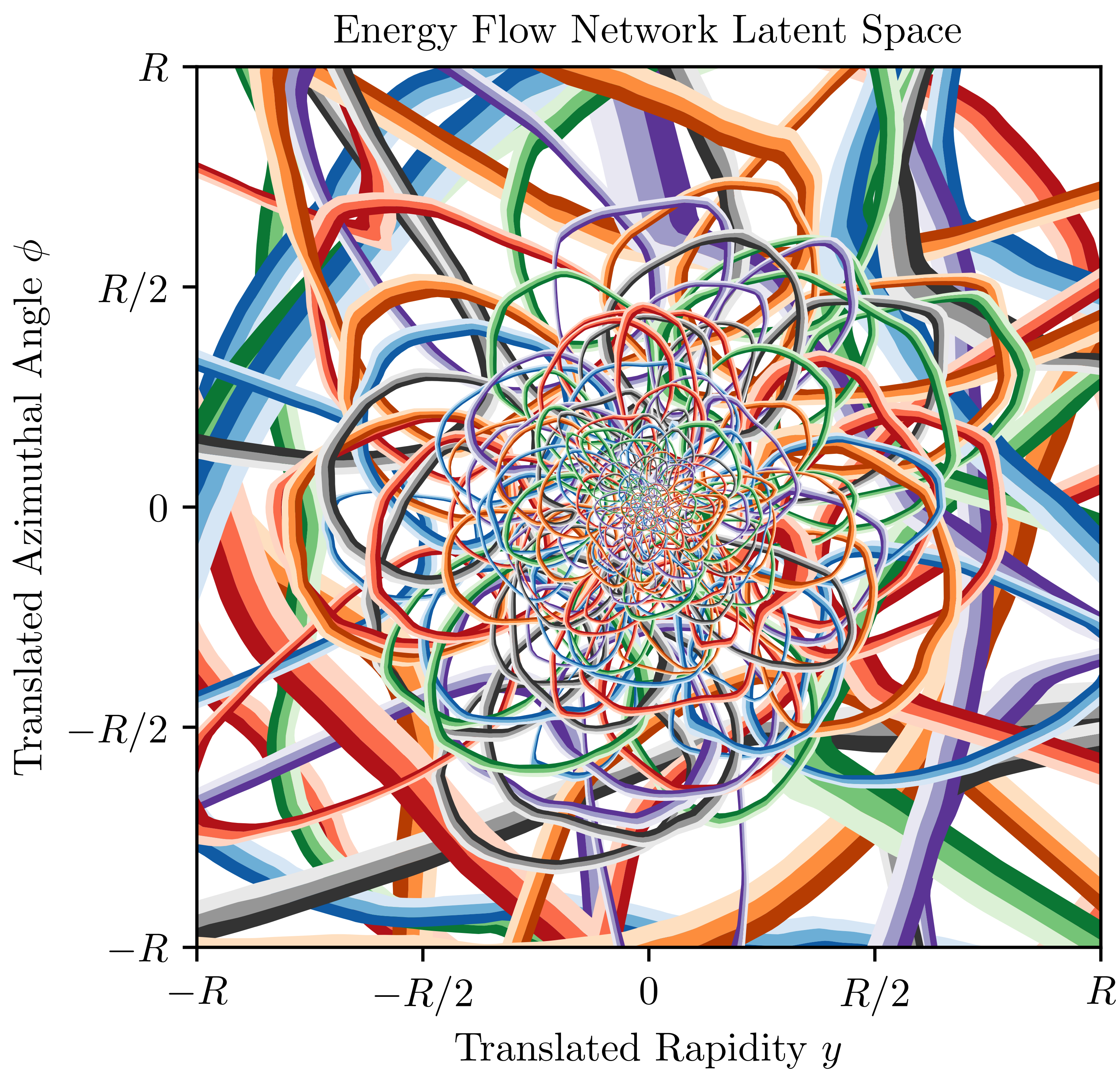
[Komiske, Metodiev, JDT, 1809.01140;
see also Metodiev, JDT, 1802.00008; Metodiev, Nachman, JDT, 1708.02949]

Topic Modeling



[Blei, 2012]

Physics Reasons for My Conversion (2)



Today's Talk

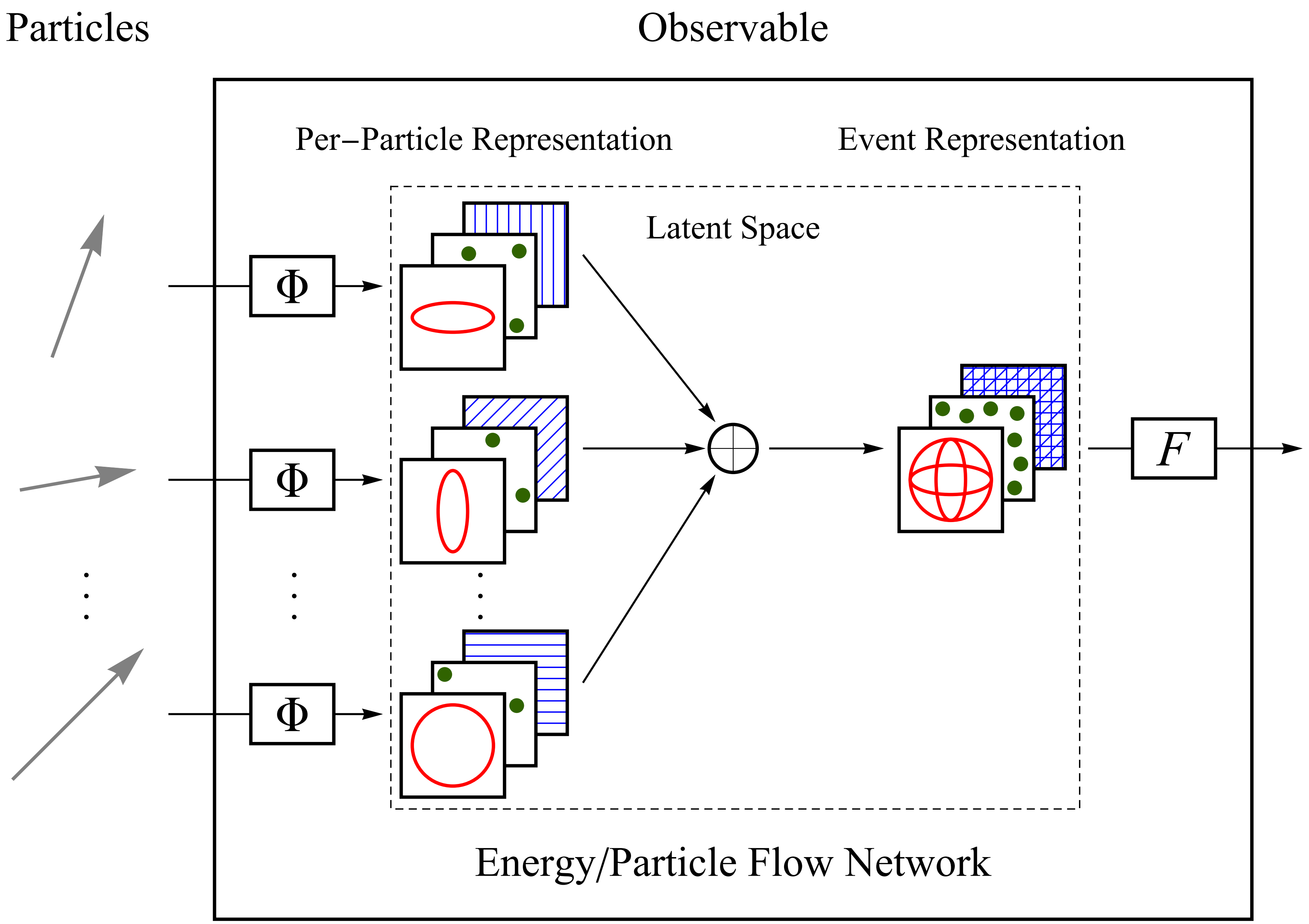
First-principles QCD meets neural networks with same underlying symmetries as…

[Komiske, Metodiev, JDT, 1810.05165; see also Komiske, Metodiev, JDT, 1712.07124]

Point Clouds

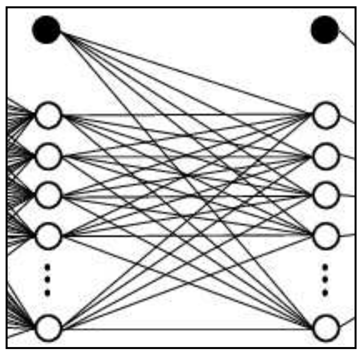
[Popular Science, 2013]

Introducing Energy Flow Networks

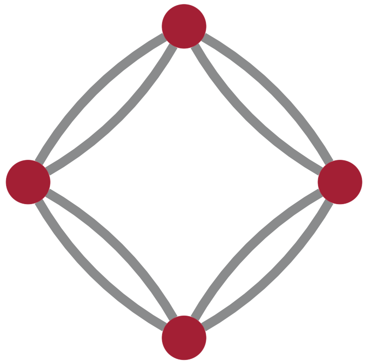


[Komiske, Metodiev, JDT, 1810.05165]

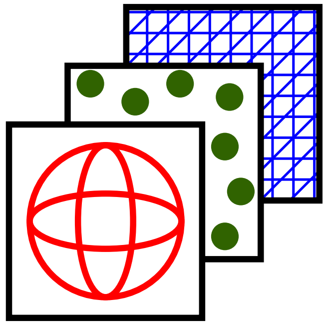
Outline



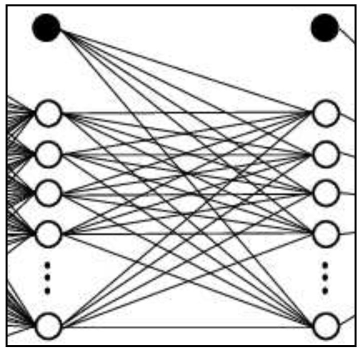
Into the Network



Symmetries & Safety



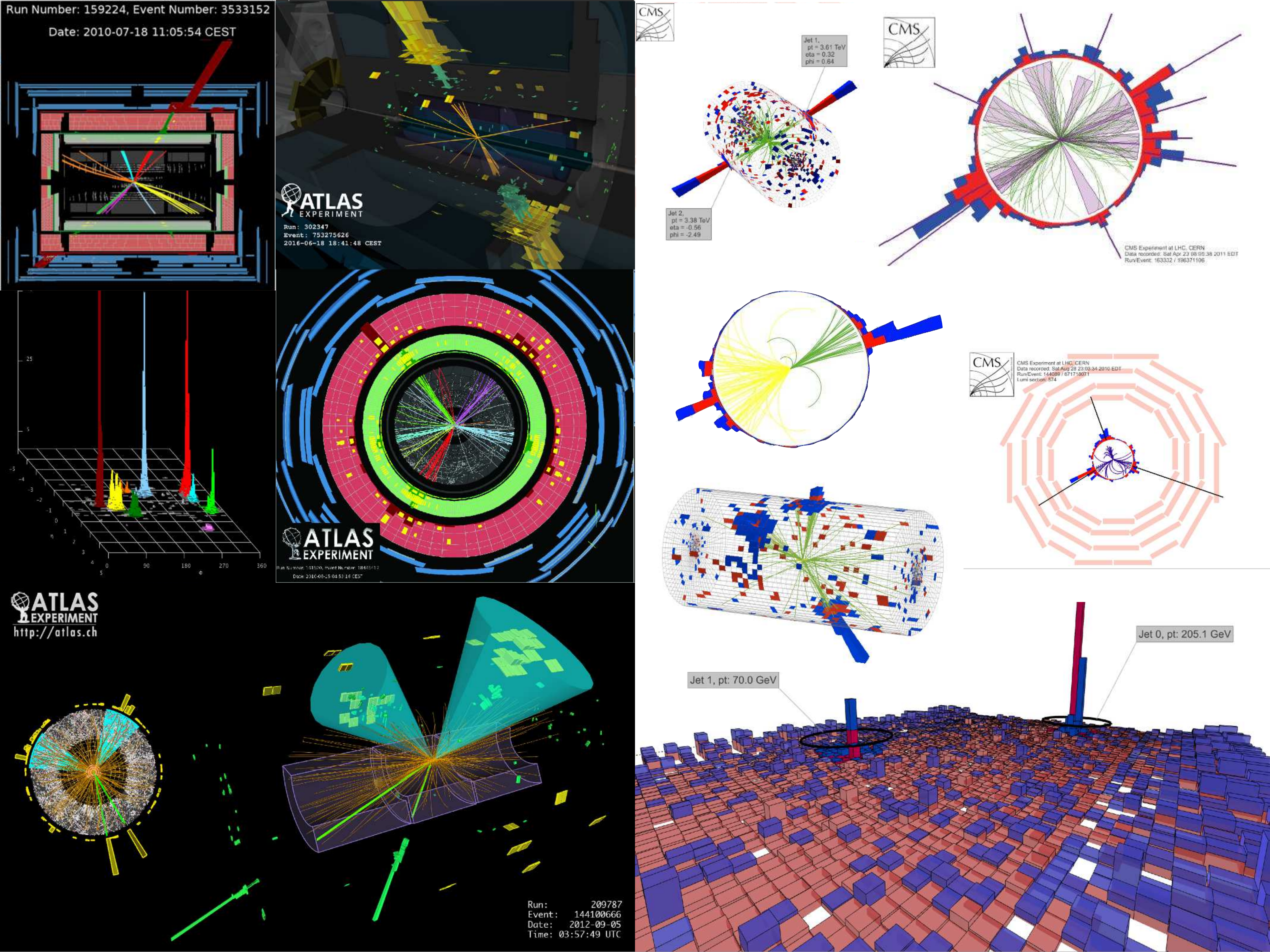
Deep Sets for Particle Jets



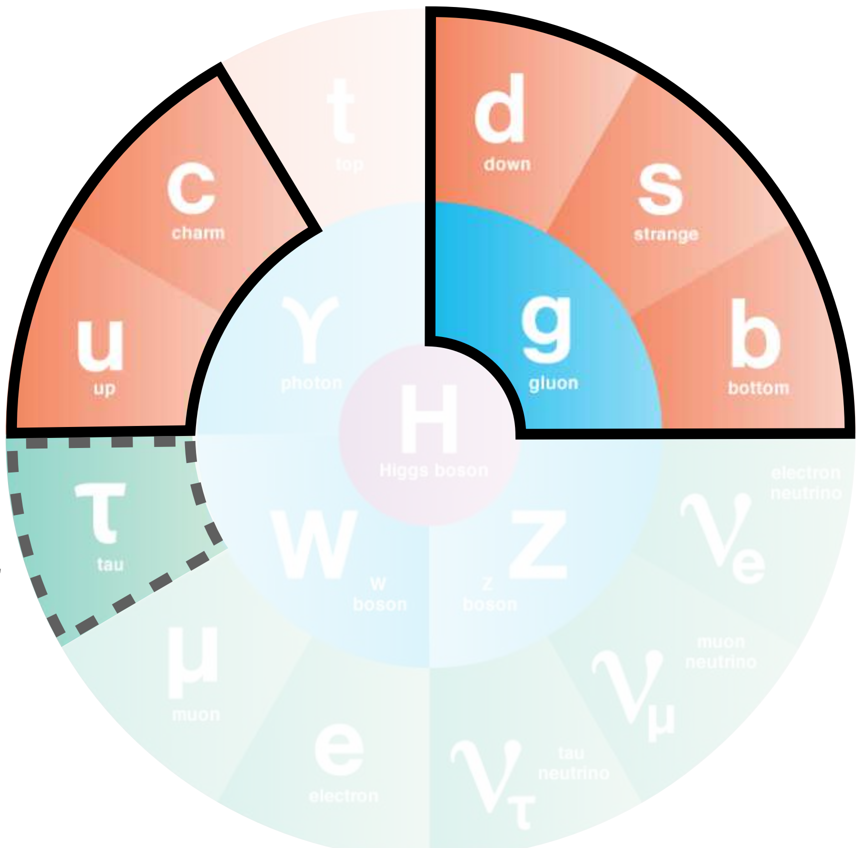
Into the Network

Symmetries & Safety

Deep Sets for Particle Jets

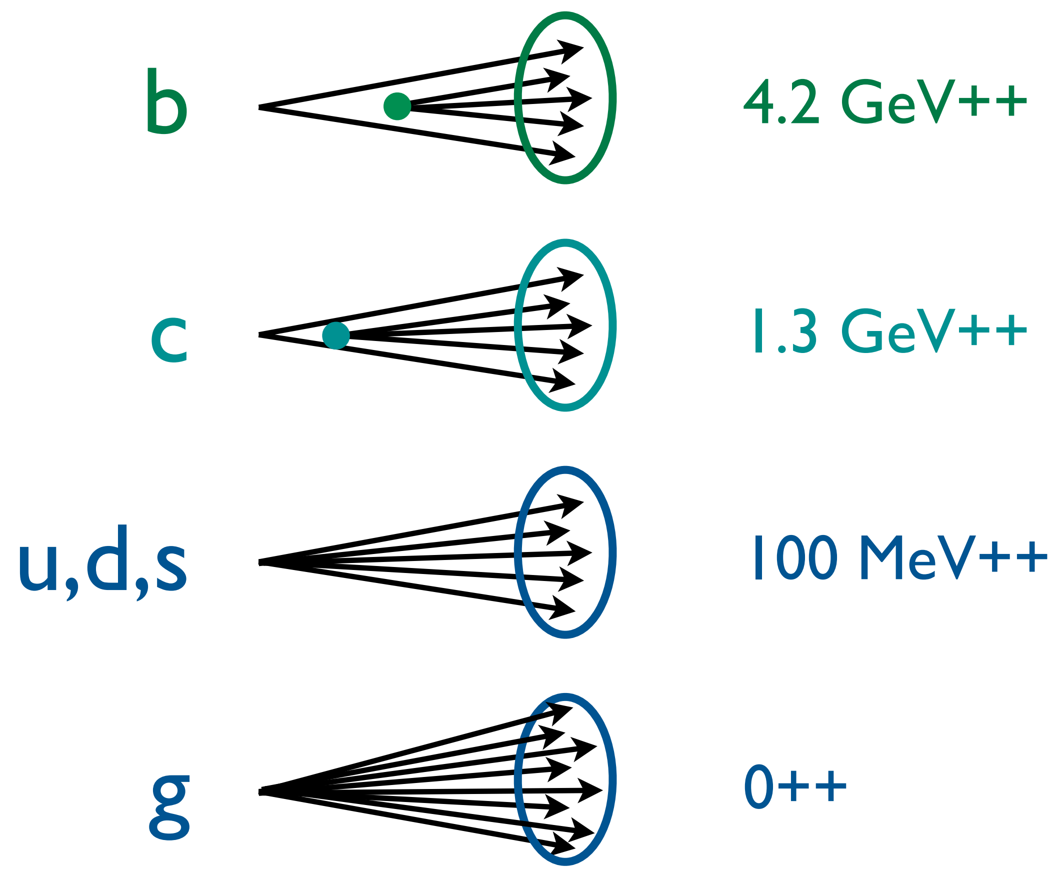


Run Number: 159224, Event Number: 3533152
Date: 2010-07-18 11:05:54 CEST
ATLAS EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST
CMS
Jet 1,
pt = 3.61 TeV
eta = 0.32
phi = 0.64
Jet 2,
pt = 3.38 TeV
eta = -0.56
phi = -2.49
CMS Experiment at LHC, CERN
ATLAS EXPERIMENT
ATLAS EXPERIMENT
http://atlas.ch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

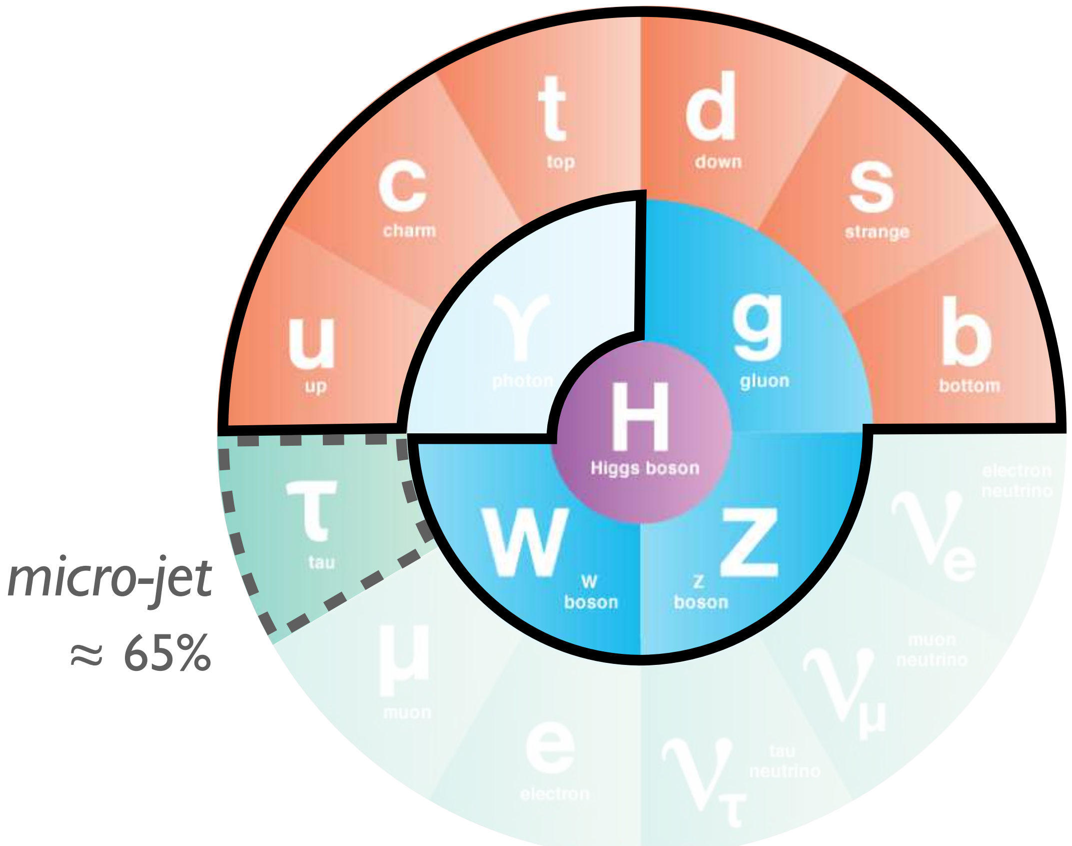


micro-jet
≈ 65%

Jets from the Standard Model

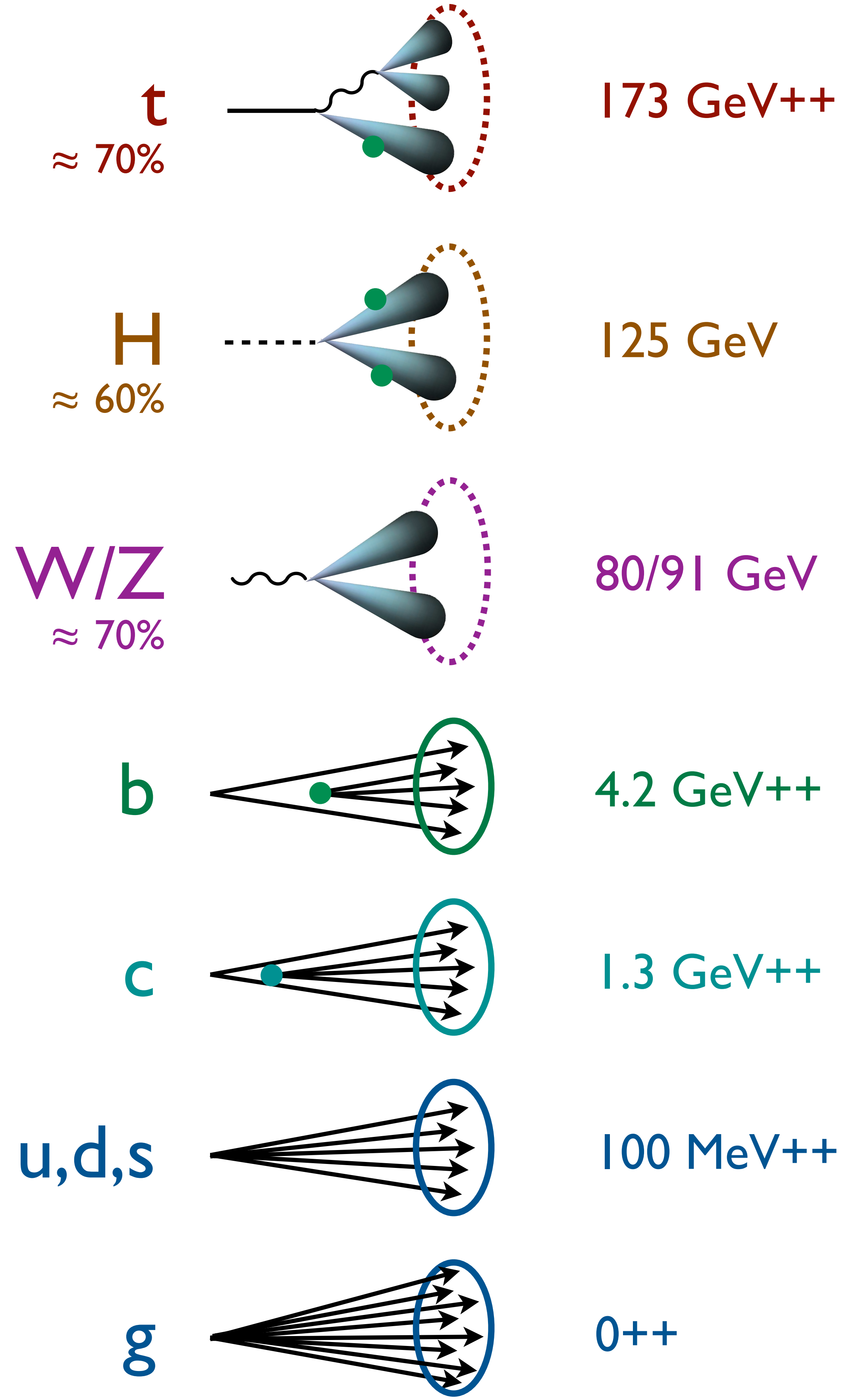


++ = Mass from QCD Radiation



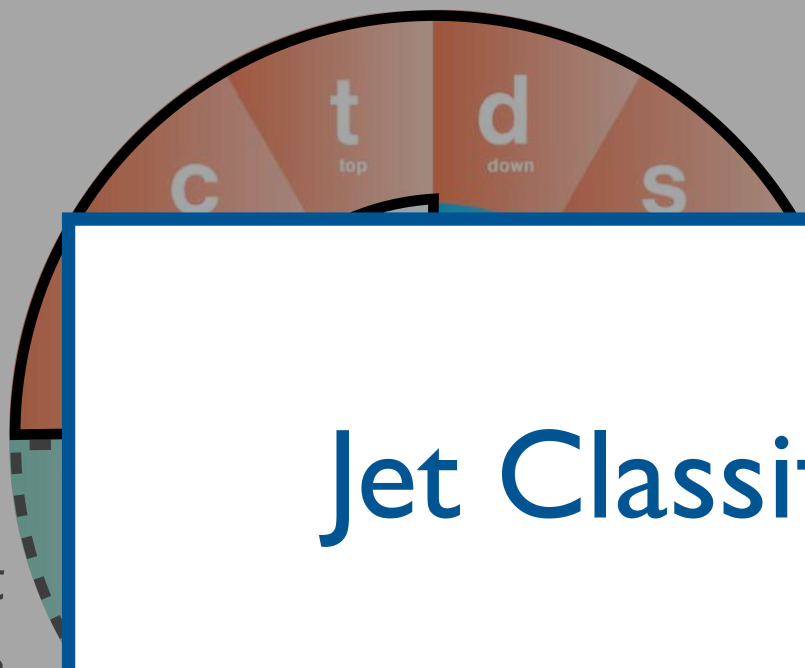
Jets from the Standard Model

++ = Mass from QCD Radiation

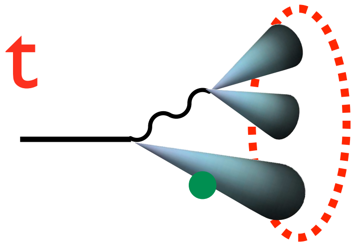


Jet Classification is “Solved”

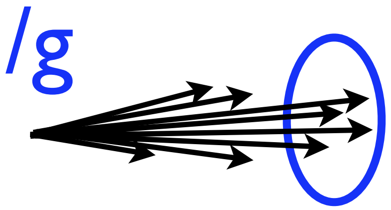
(assuming well-defined jet categories and trustable training samples)

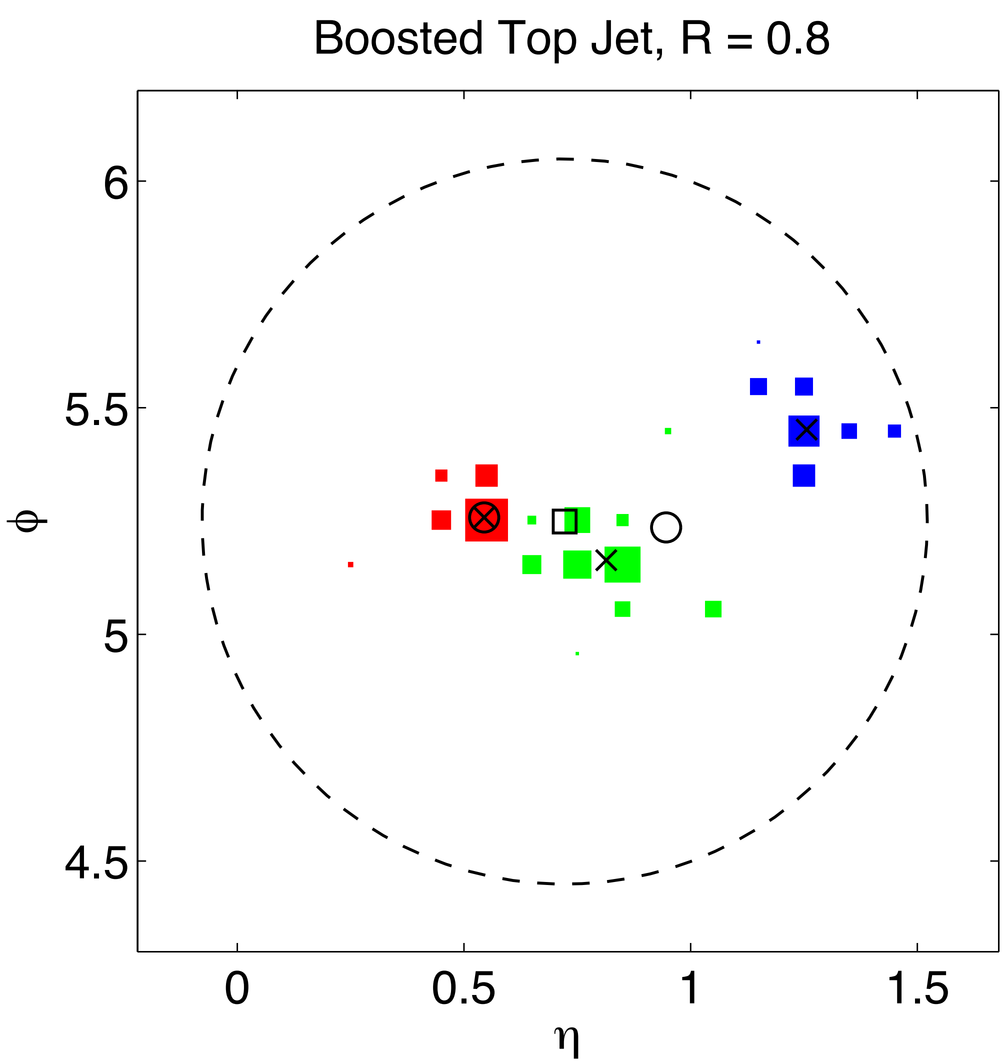


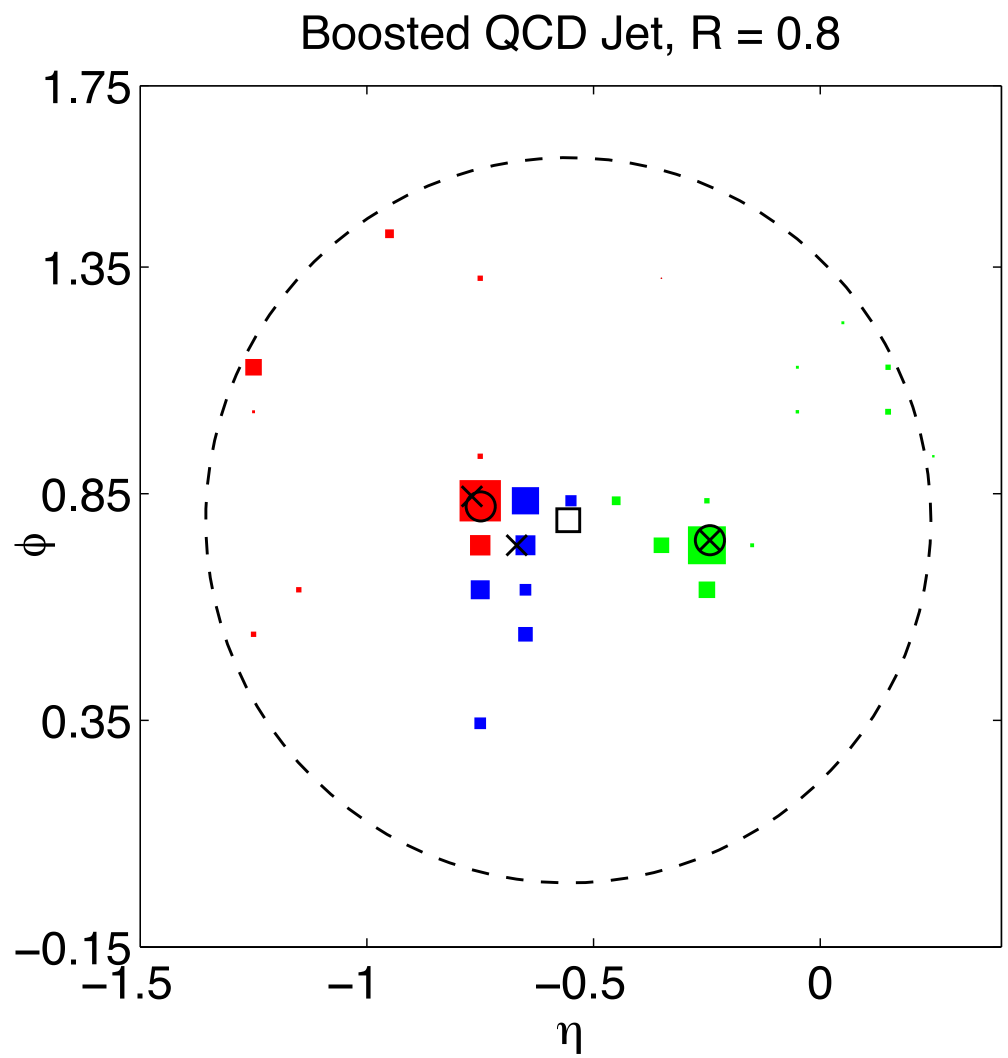
++ = Mass from QCD Radiation



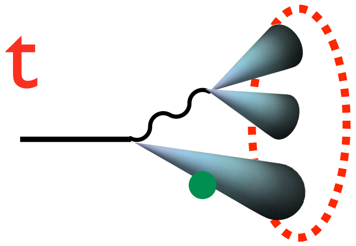
3-Prong vs. 1-Prong



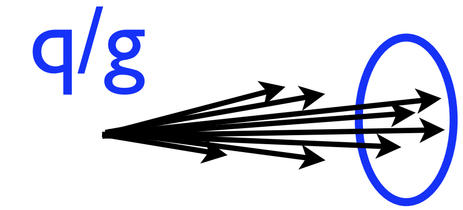


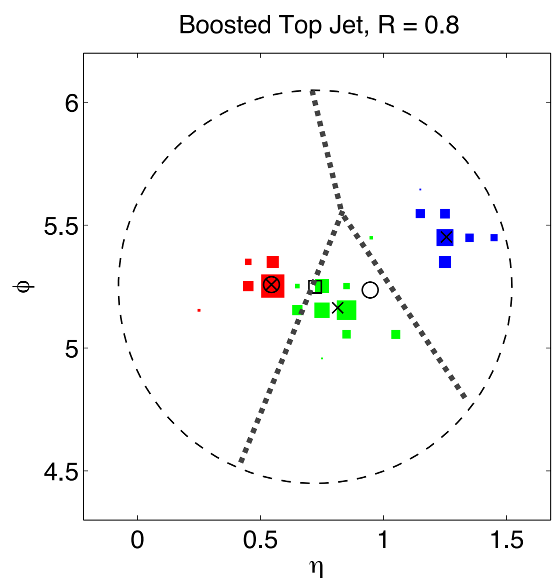


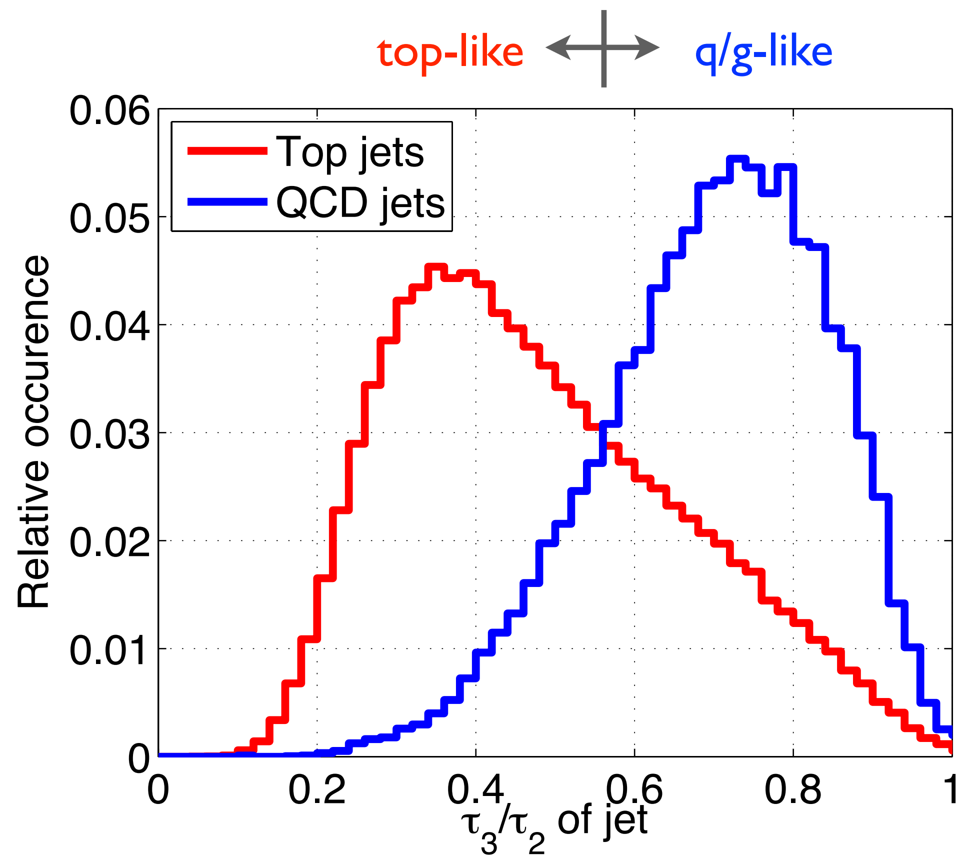
If your eyes can do it…

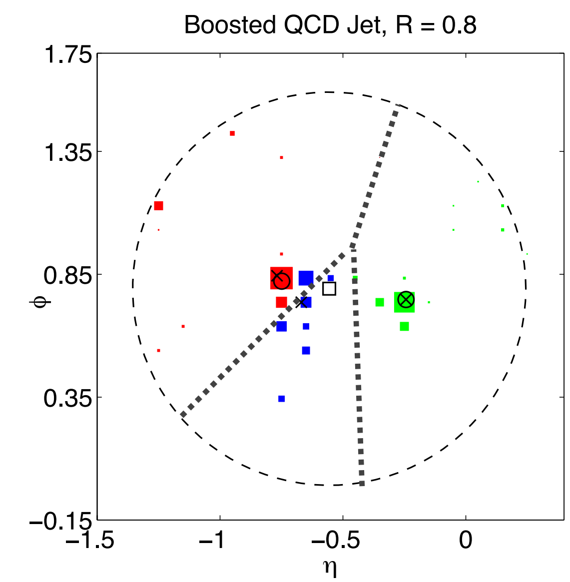


3-Prong vs. 1-Prong





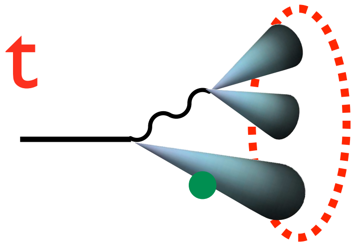




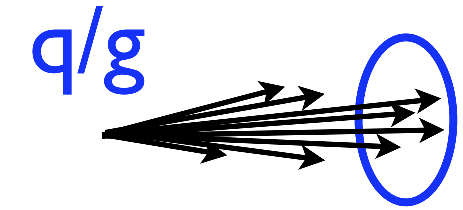
N-subjettiness

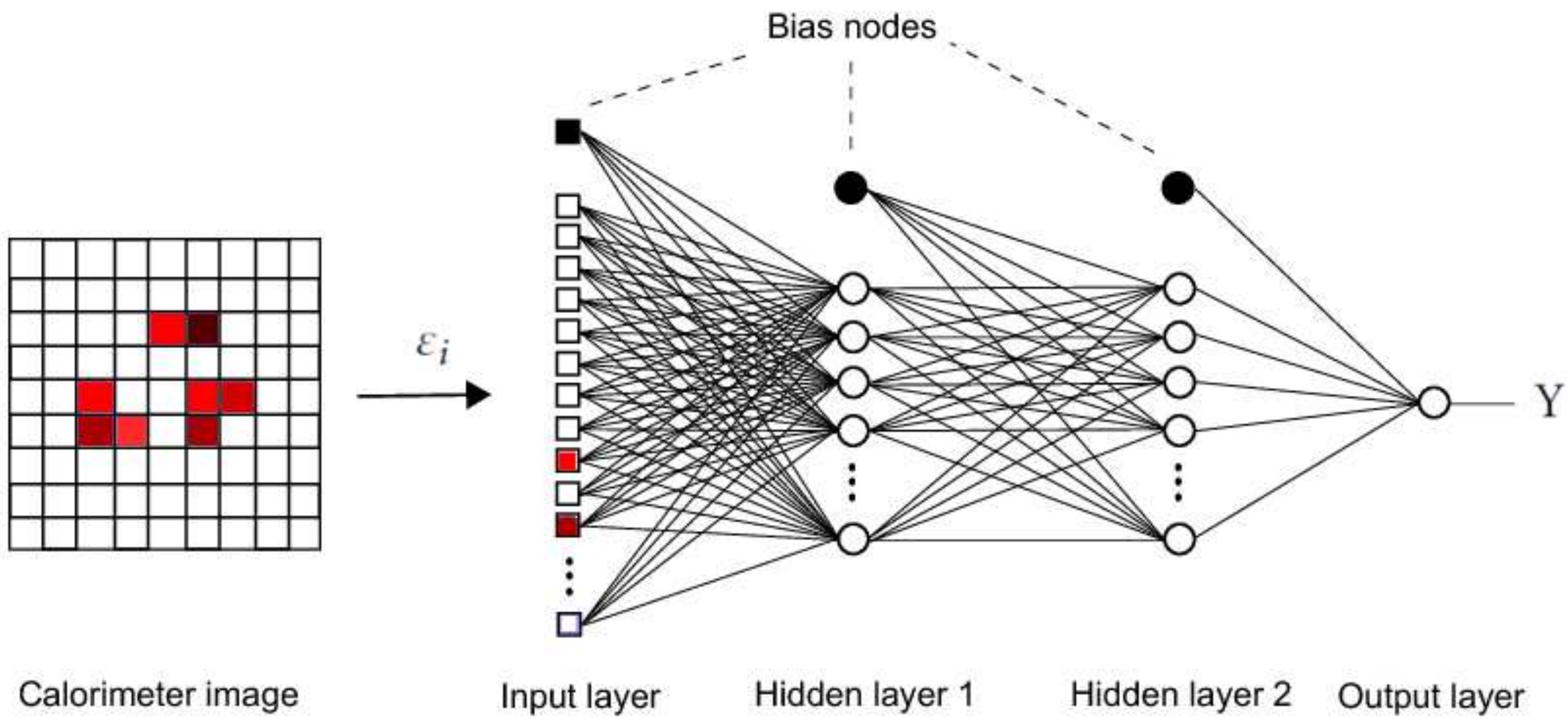
"Deep Thinking": $$\tau_N = \sum_k p_{T,k} \min \{\Delta R_{k,1}, \Delta R_{k,2}, \ldots, \Delta R_{k,N}\}$$

[e.g. JDT, Van Tilburg, 1011.2268, 1108.2701]



3-Prong vs. 1-Prong





“Deep Learning”: BDTs, FLDs, DNNs, CNNs, RNNs, …

[e.g. Almeida, Backović, Cliche, Lee, Perelstein, 1501.05968]

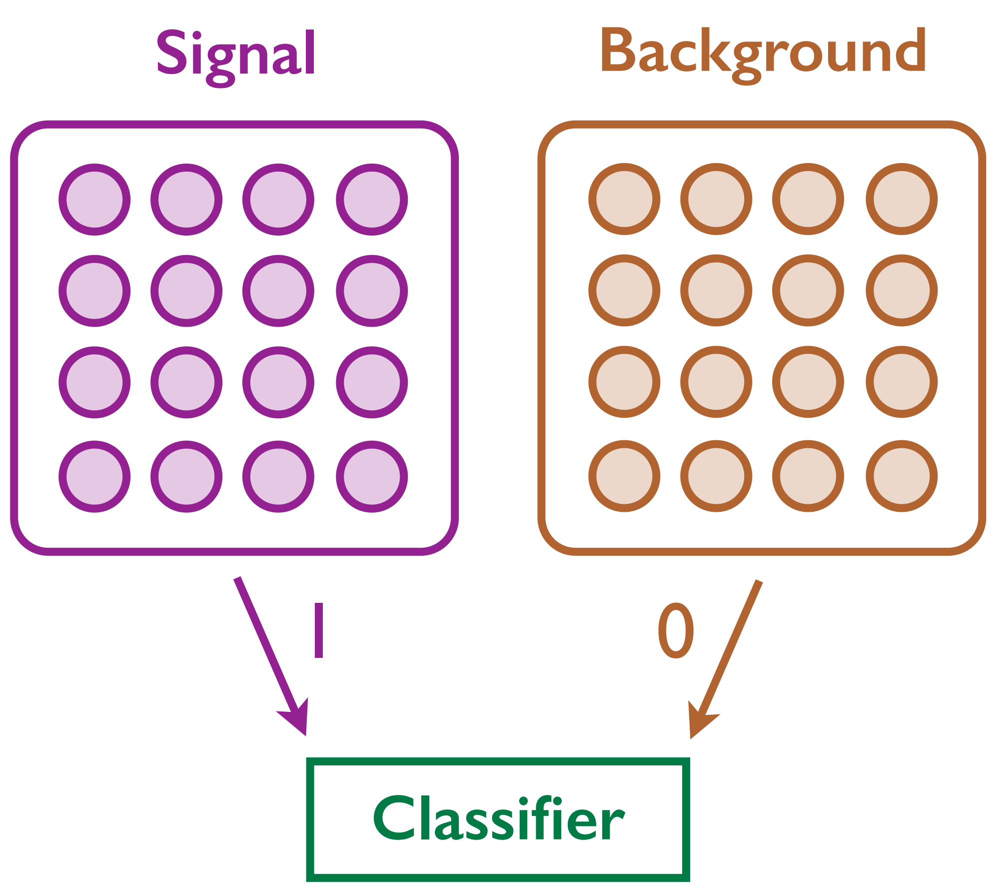
A Cartoon of Machine Learning

For fully-supervised jet classification

(see backup for regression, generation, modeling)

$$\ell_{\mathrm{MSE}} = \Big\langle \big(h(\vec{x}) - 1\big)^2 \Big\rangle_{\text{signal}} + \Big\langle \big(h(\vec{x}) - 0\big)^2 \Big\rangle_{\text{background}}$$

Classifier Inputs



Minimize Loss Function

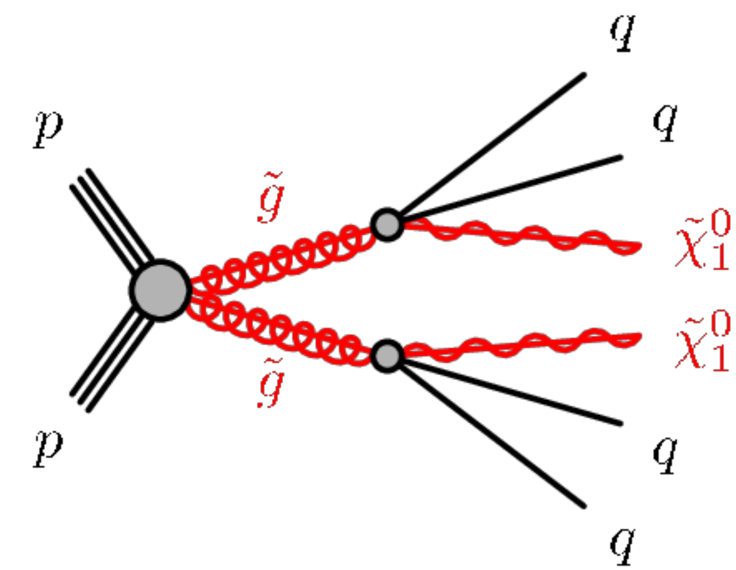
(assuming infinite training sets, and flexible enough functional form)

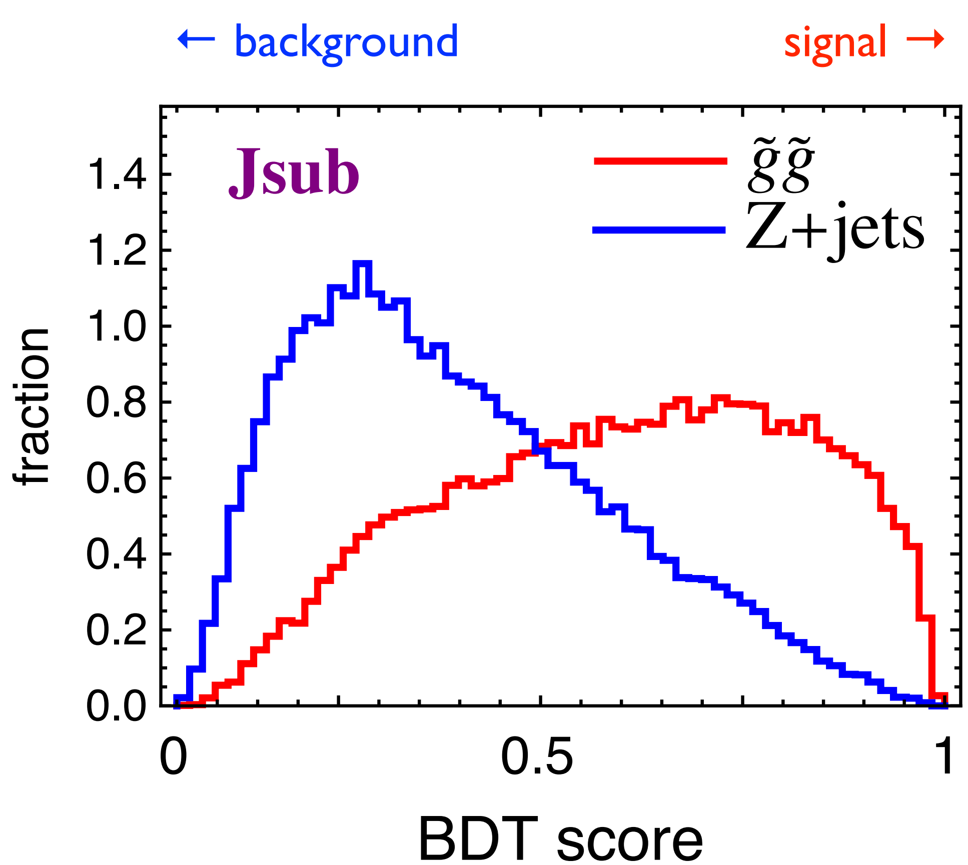
$$h(\vec{x}) = \frac{p_{\mathrm{sig}}(\vec{x})}{p_{\mathrm{sig}}(\vec{x}) + p_{\mathrm{bkgd}}(\vec{x})}$$

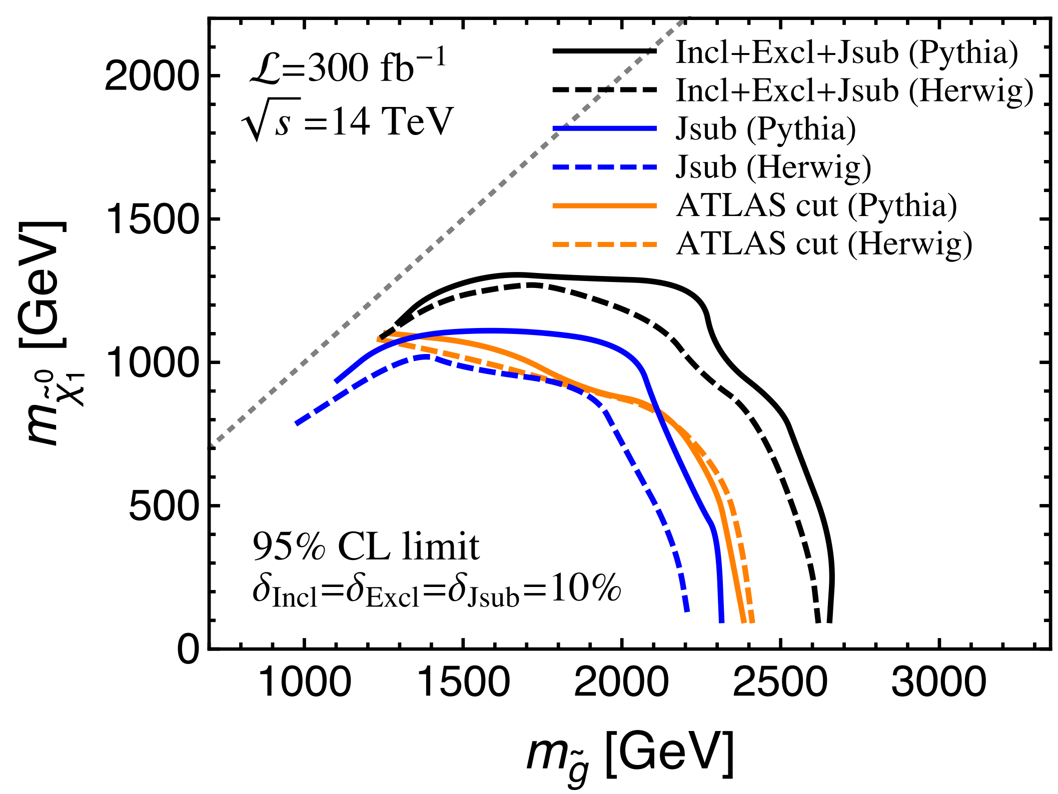
Optimal Classifier (Neyman–Pearson)

E.g. SUSY Search for Gluino Pairs

Classifier: Boosted decision tree (for each of 4 jets)
Inputs: Jet mass, width, track multiplicity
Signal: Quark enriched ($C_F = 4/3$)
Background: Gluon enriched ($C_A = 3$)







[Bhattacherjee, Mukhopadhyay, Nojiri, Sakakie, Webber, 1609.08781]

Jet Classification Studies

Mix and match

$$\ell_{\mathrm{MSE}} = \Big\langle \big(h(\vec{x}) - 1\big)^2 \Big\rangle_{\mathrm{signal}} + \Big\langle \big(h(\vec{x}) - 0\big)^2 \Big\rangle_{\mathrm{background}}$$

Classifier

- Boosted Decision Tree
- Fisher Linear Discriminant
- Shallow Neural Network
- Deep Neural Network
- Convolutional Neural Network
- Recurrent Neural Network
- Recursive Neural Network
- Combination/Lorentz Layers
- …

Inputs

- High-Level Features
- Basis of High-Level Features
- Jet Image
- Multi-channel Jet Image
- Abstract Jet Image
- Sorted Four-Vectors
- Clustered Four-Vectors
- Lund Plane Emissions
- Kitchen Sink
- …

Signal vs. Background

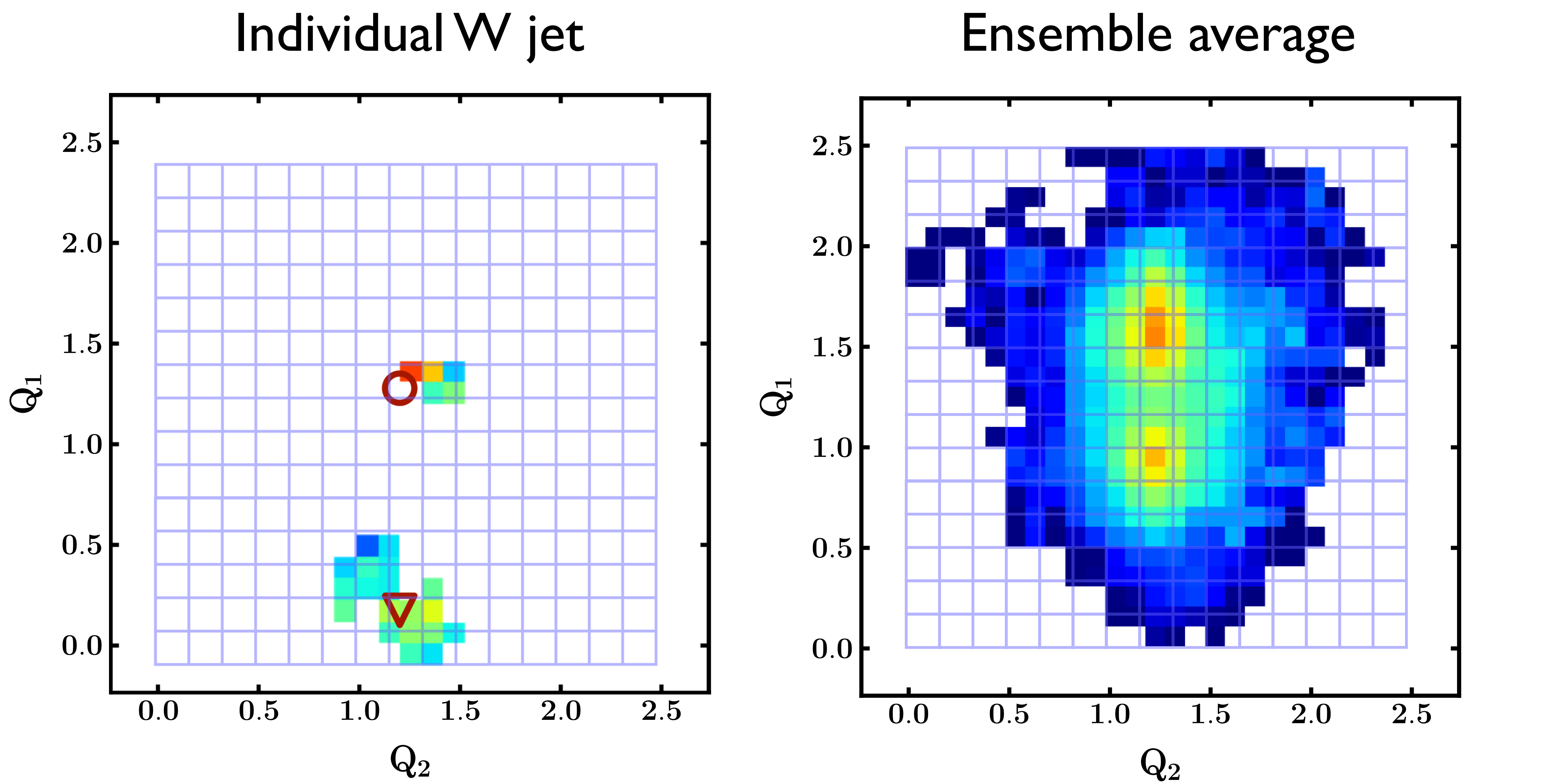
Signal	vs.	Background
Quark Jets	vs.	Gluon Jets
Up-type Quarks	vs.	Down-type Quarks
W/Z Bosons	vs.	QCD Jets
W Bosons	vs.	Z Bosons
Top Quarks	vs.	QCD Jets
Exotic Boosted Objects	vs.	QCD Jets
CMS Open Data Samples	vs.	Each other
…	vs.	…

[Lönnblad, Peterson, Rögnvaldsson, 1990, …, Cogan, Kagan, Strauss, Schwartzman, 1407.5675; Almeida, Backović, Cliche, Lee, Perelstein, 1501.05968; de Oliveira, Kagan, Mackey, Nachman, Schwartzman, 1511.05190; Baldi, Bauer, Eng, Sadowski, Whiteson, 1603.09349; Conway, Bhaskar, Erbacher, Pilot, 1606.06859; Guest, Collado, Baldi, Hsu, Urban, Whiteson, 1607.08633; Barnard, Dawe, Dolan, Rajcic, 1609.00607; Komiske, Metodiev, Schwartz, 1612.01551; Kasieczka, Plehn, Russell, Schell, 1701.08784; Louppe, Cho, Becot, Cranmer, 1702.00748; Pearkes, Fedorko, Lister, Gay, 1704.02124; Datta, Larkoski, 1704.08249, 1710.01305; Butter, Kasieczka, Plehn, Russell, 1707.08966; Fernández Madrazo, Heredia Cacha, Lloret Iglesias, Marco de Lucas, 1708.07034; Aguilar Saavedra, Collin, Mishra, 1709.01087; Cheng, 1711.02633; Luo, Luo, Wang, Xu, Zhu, 1712.03634; Komiske, Metodiev, JDT, 1712.07124; Macaluso, Shih, 1803.00107; Fraser, Schwartz, 1803.08066; Choi, Lee, Perelstein, 1806.01263; Lim, Nojiri, 1807.03312; Dreyer, Salam, Soyez, 1807.04758; Moore, Nordström, Varma, Fairbairn, 1807.04769; plus my friends who will scold me for forgetting their paper; plus many ATLAS/CMS performance studies]

Jet Classification Studies

Mix and match

Standard CNN input: Jet images

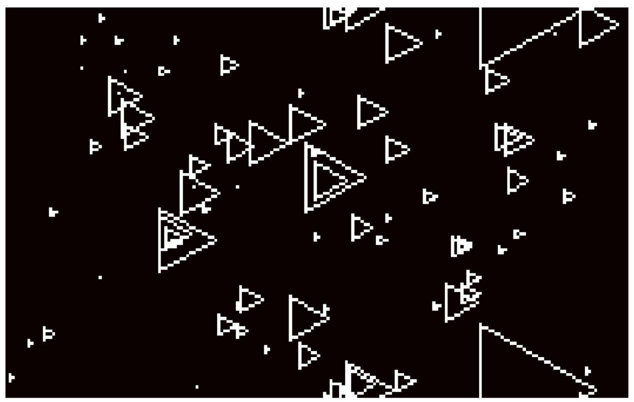


[Cogan, Kagan, Strauss, Schwartzman, 1407.5675]

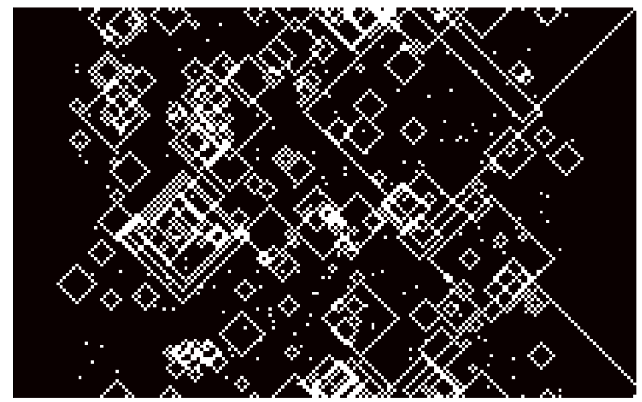
Jet Classification Studies

Mix and match

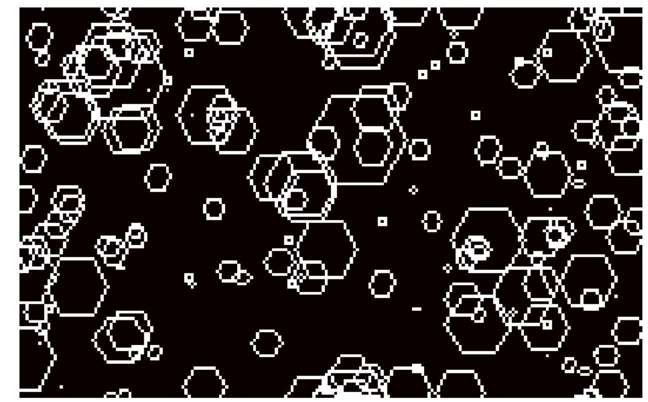
Novel CNN input: Abstract event images



(a) Photons

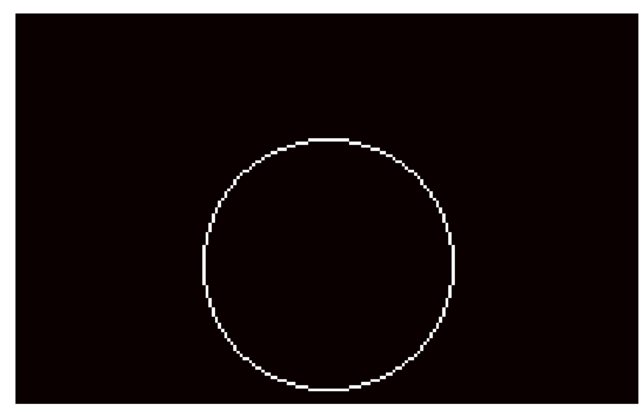


(b) Charged Particles



(c) Neutral Hadrons

(d) Lepton

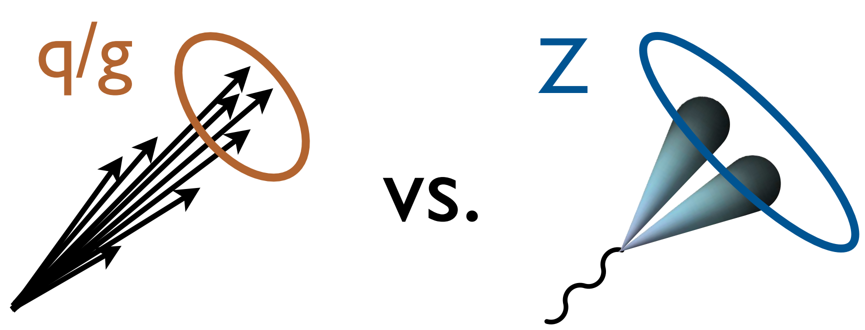


(e) E_T^{miss}

Addresses sparsity problem of standard energy-to-intensity mapping

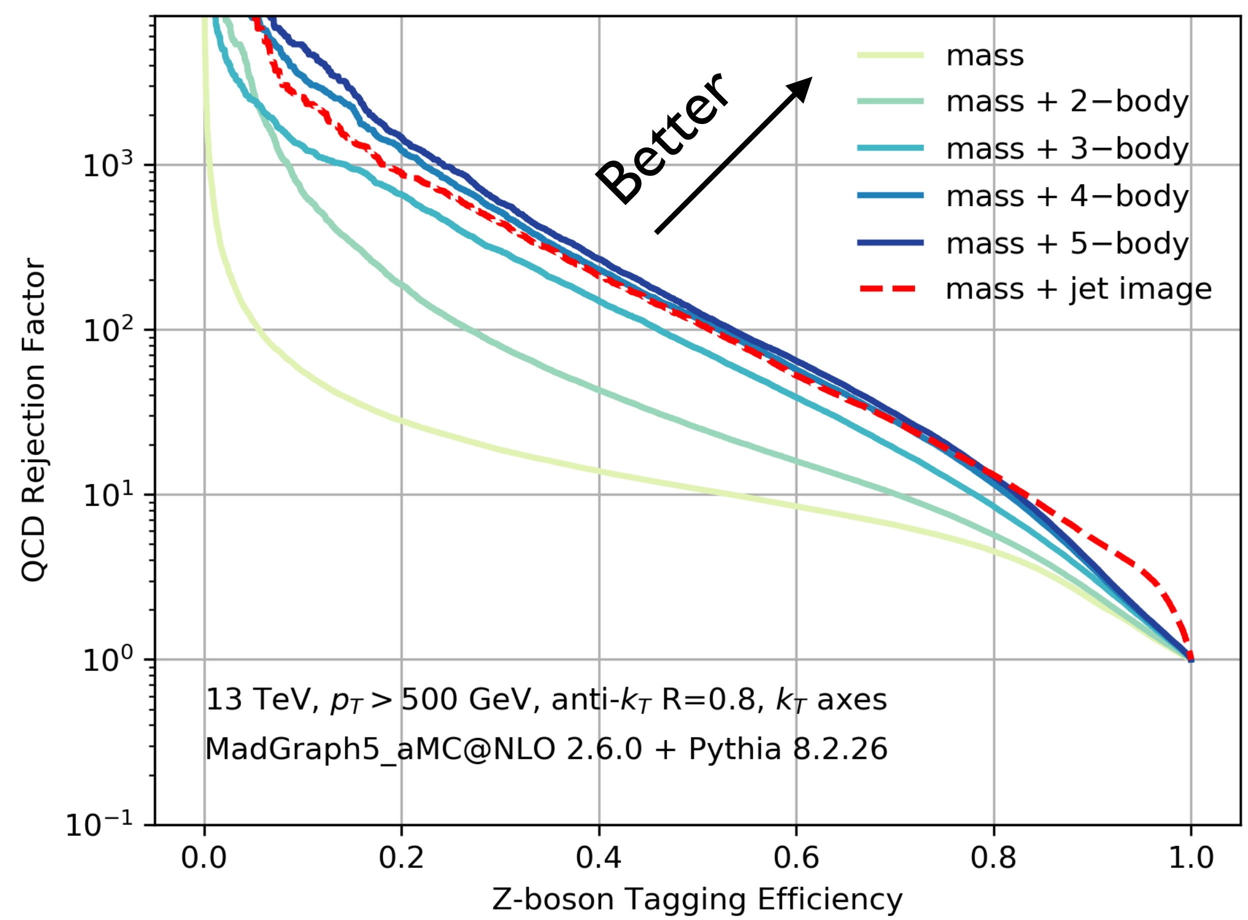
[Nguyen, Weitekamp, Anderson, Castello, Cerri, Pierini, Spiropulu, Vlimant, 1807.00083; using Fernández Madrazo, Heredia Cacha, Lloret Iglesias, Marco de Lucas, 1708.07034]

Evidence for Performance Saturation



"Any sufficiently advanced technology is indistinguishable from magic"

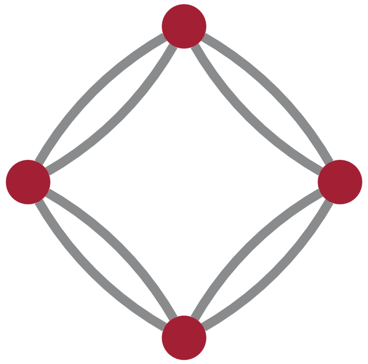
Jet Images CNN ≈ "Expert" BDT



Next frontier is robustness, versatility & transparency

[plot from Moore, Nordström, Varma, Fairbairn, 1807.04769; see also Datta, Larkoski, 1704.08249]

Into the Network



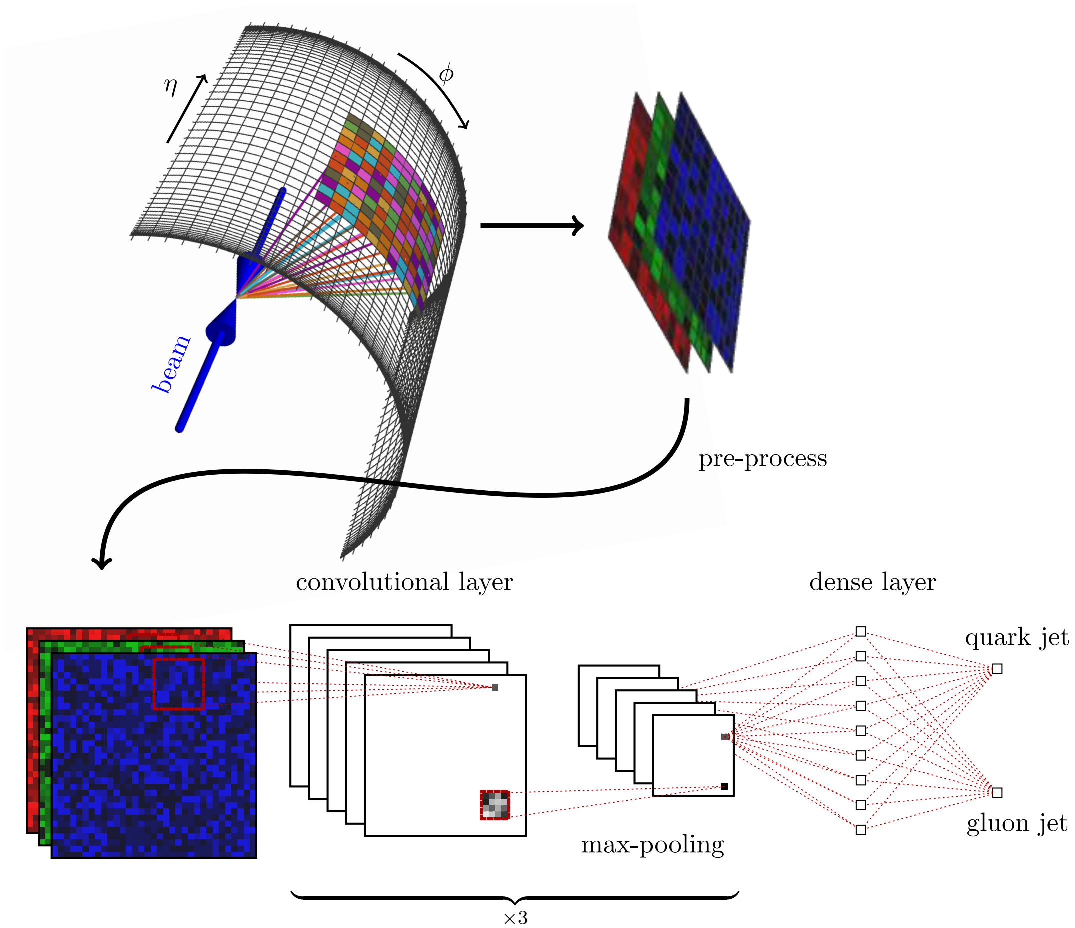
Symmetries & Safety

Deep Sets for Particle Jets

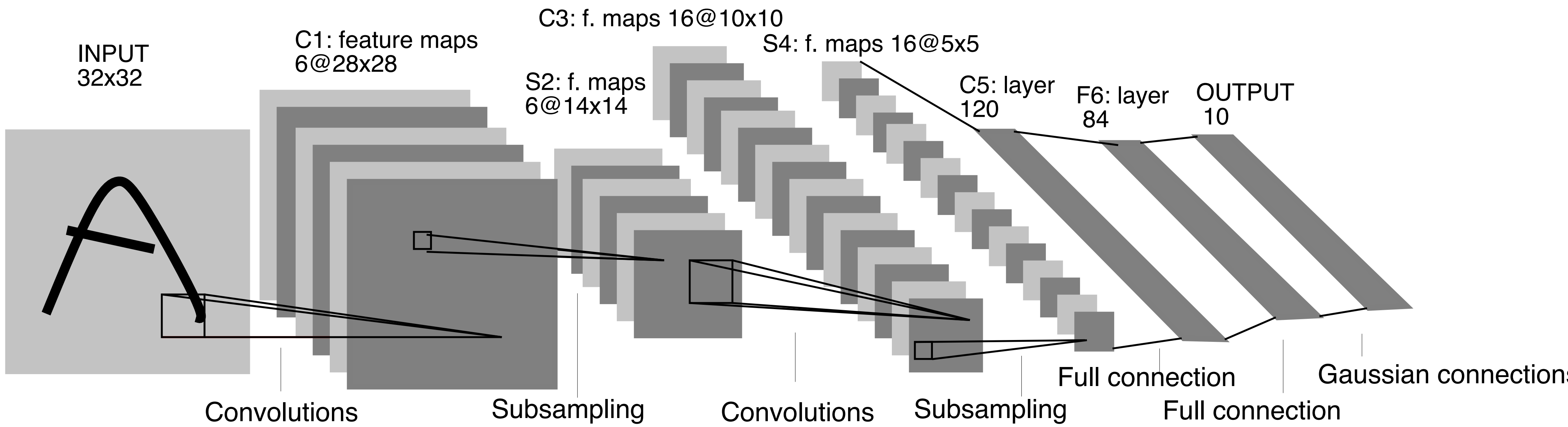
Patrick Komiske

Eric Metodiev

Two grad students walk into my office with their CNN…

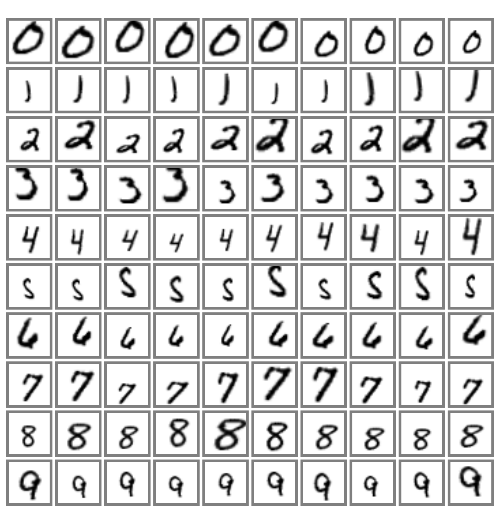


Symmetries of a CNN



Explores different length scales

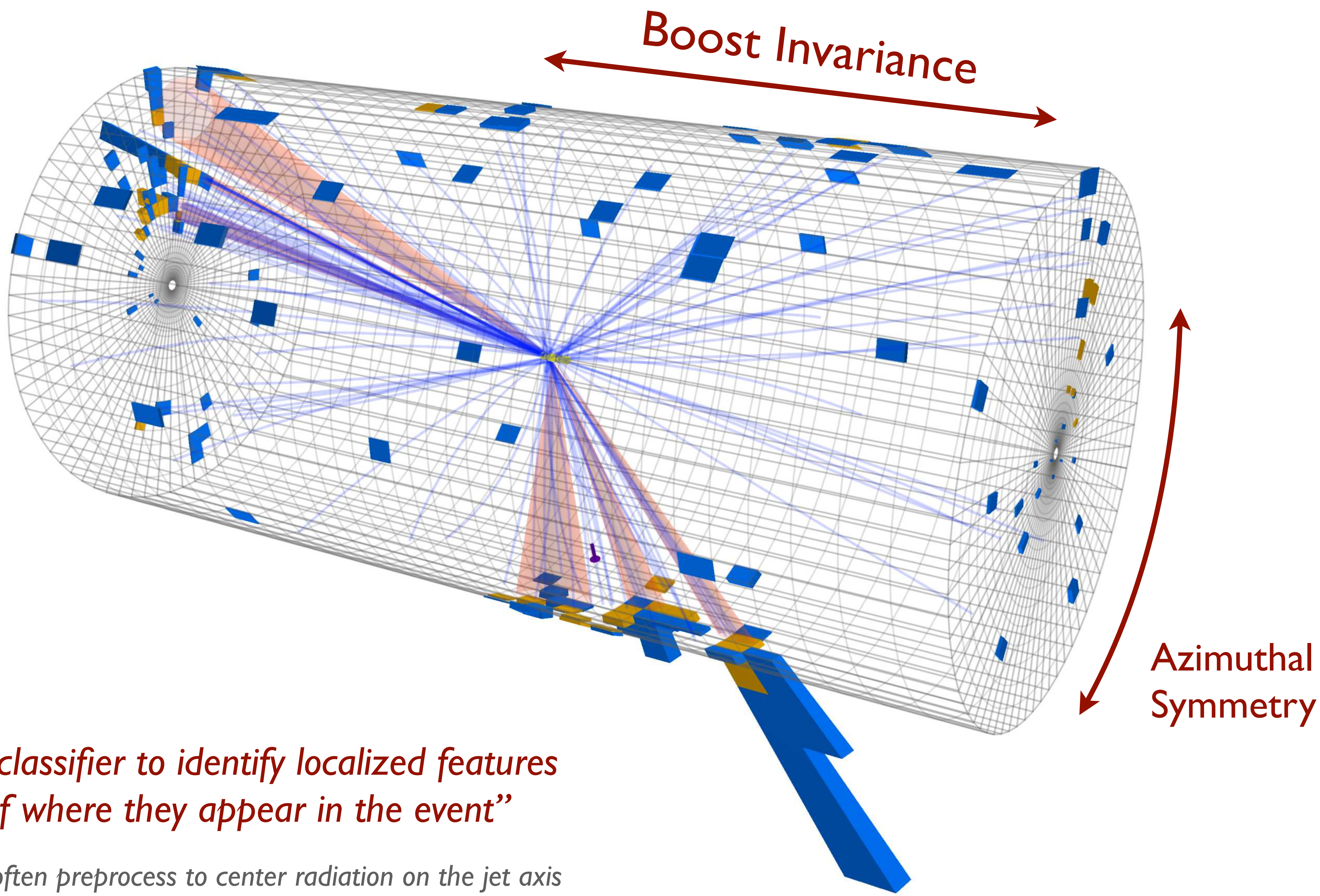
Enforces Translation Invariance



"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]

The Physics-First Approach

Underlying Physics

$\Downarrow$ "Deep Thinking"

Natural Data Representation

$\Downarrow$ "Deep Learning"

Suitable Algorithm

The Buzzword-First Approach

Questionable Physics

$\Uparrow$ "Wishful Thinking"

Unnatural Data Representation

$\Uparrow$ "Shoehorning"

Cool-Sounding Algorithm

Why CNNs Aren't Ideal

Underlying Physics
Natural Data Representation
Suitable Algorithm

Translations/Boost Invariance

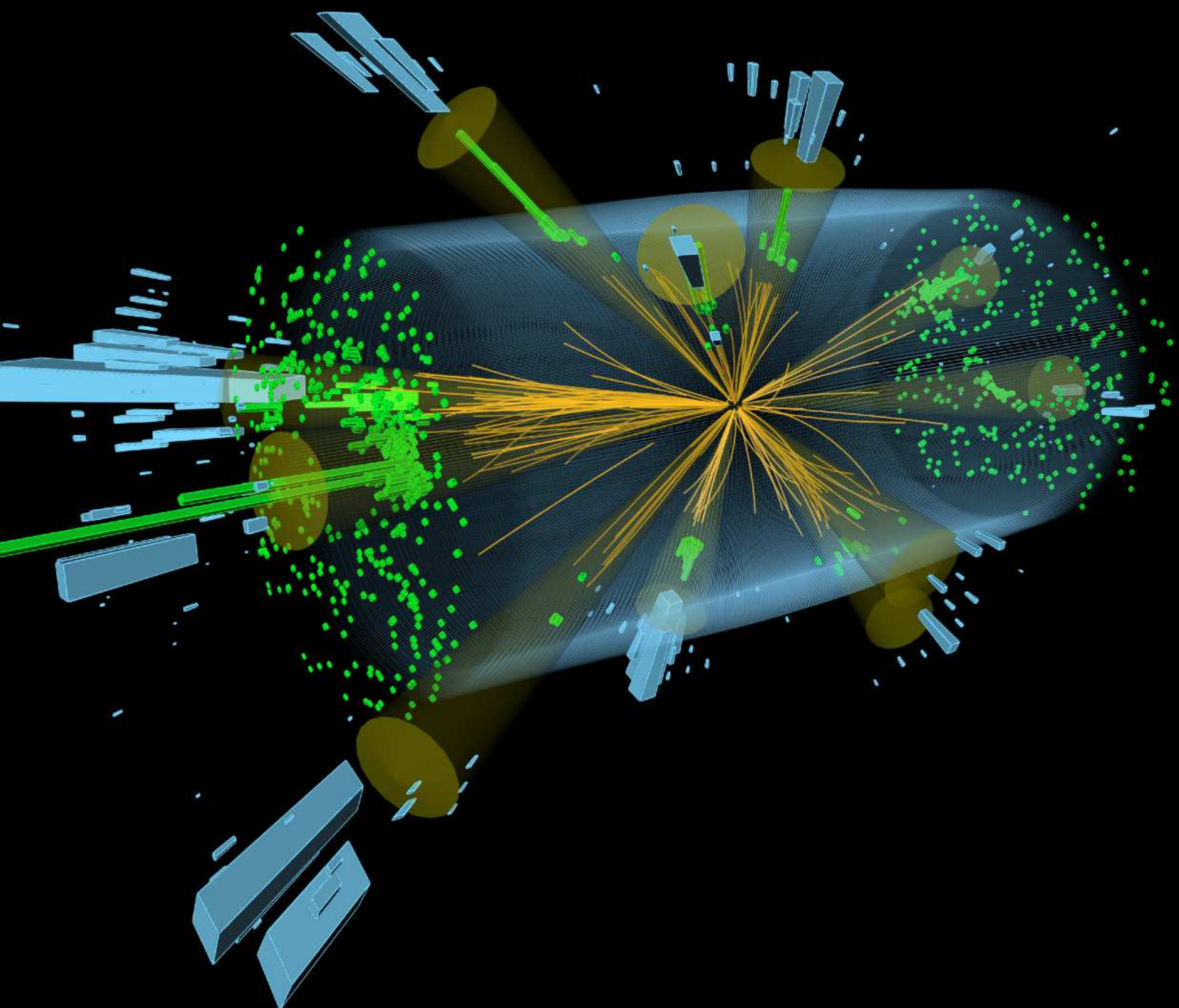
$\Downarrow$ "Deep Thinking"

Pixelated Image

$\Uparrow$ "Shoehorning"

Convolution Neural Networks

Debris Taxonomy

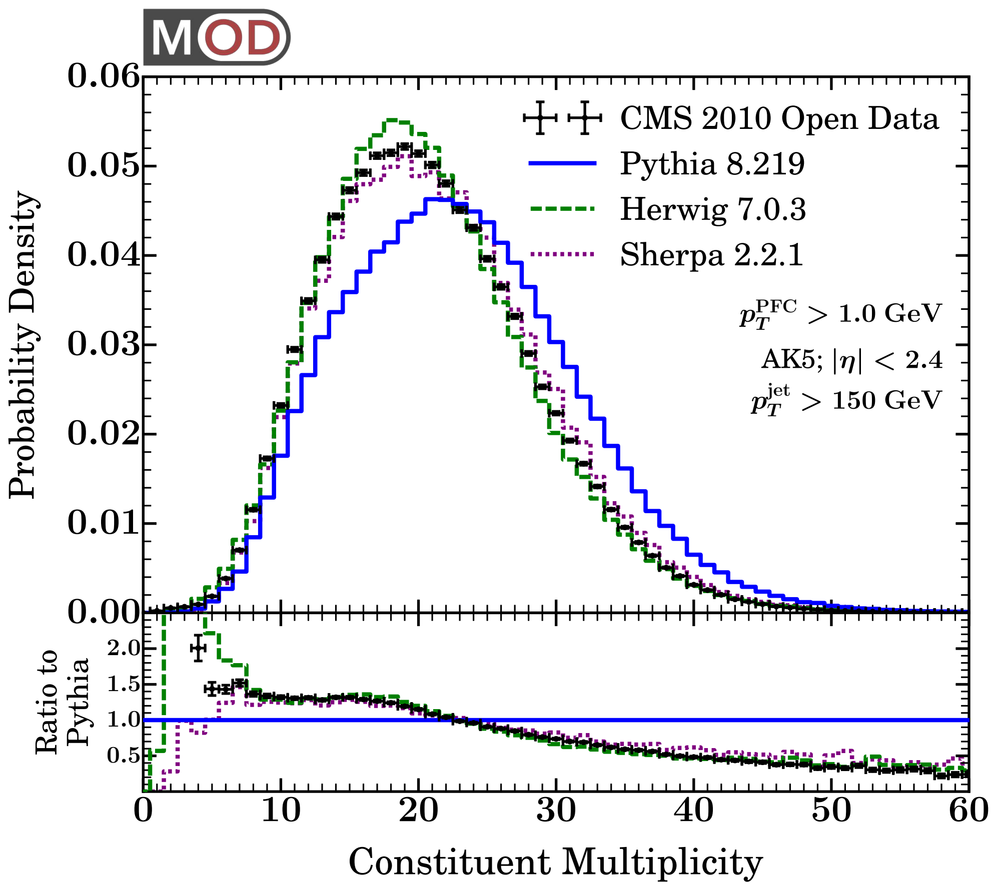


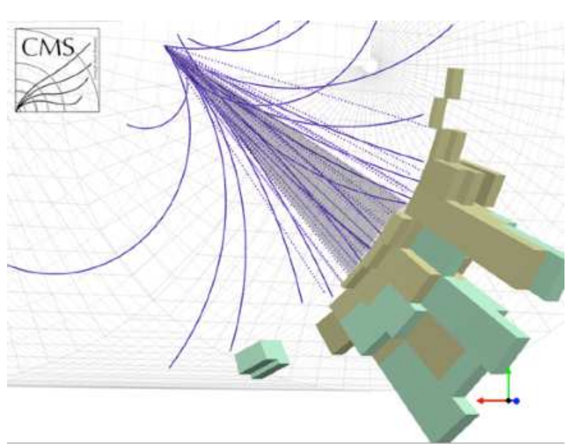
T	E	H	M			
	●			γ	photon	elementary
●	●			$e^{\pm}$	electron	
●	•	•	●	$\mu^{\pm}$	muon	
●	•	●		$\pi^{\pm}$	pion	composite
●	•	●		$K^{\pm}$	kaon	
	•	●		K_L^0	K-long	
●	•	●		$p/\overline{p}$	proton	
	•	●		$n/\overline{n}$	neutron	

Point Clouds

[Popular Science, 2013]

Key Fact #1





Jet constituents:

Particle-like objects

Variable-length

Unordered set

Per particle:

$\{E, p_x, p_y, p_z\}$ or $\{p_T, y, \Phi, m\}$

Flavor/charge labels

Vertex information

Quality criteria, etc.

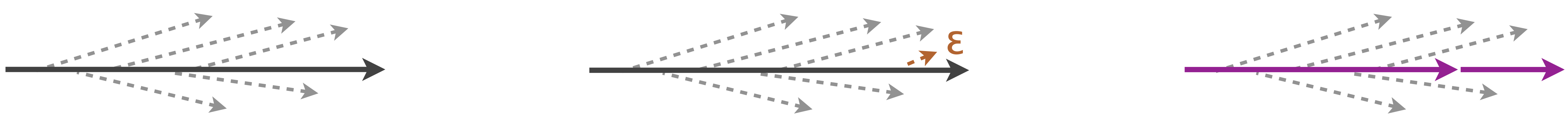
[plot from Tripathee, Xue, Larkoski, Marzani, JDT, 1704.05842]

Key Fact #2

Wide range of interesting observables are “safe”

Interesting ≈ Calculable in fixed-order perturbation theory

Original Jet | Infrared | Collinear



IRC Safe Observable: Insensitive to IR or C emissions

Enforces smooth interpolation between
variable-length inputs (i.e. $N \to N-1$)

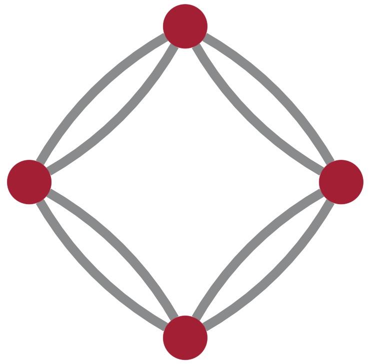
In the Backup (2017)

Underlying Physics
Natural Data Representation
Suitable Algorithm

Infrared/Collinear Safety

$\Downarrow$

Energy Flow Polynomials



$\Downarrow$

Linear Regression

[Komiske, Metodiev, JDT, 1712.07124; https://energyflow.network]

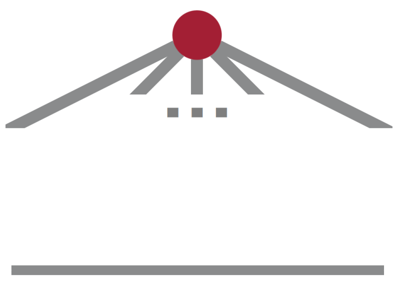
In the Backup (2019?)

Underlying Physics
Natural Data Representation
Suitable Algorithm

Infrared/Collinear Safety

$\Downarrow$

Energy Flow Moments



$\Uparrow$

Linear Regression + Linear Runtime

[Komiske, Metodiev, JDT, we've been promising this paper for 9 months]

Today's Talk (2018)

Underlying Physics
Natural Data Representation
Suitable Algorithm

Quantum-Mechanical Indistinguishability

$\Downarrow$

Variable-Length Unordered Sets of Particles

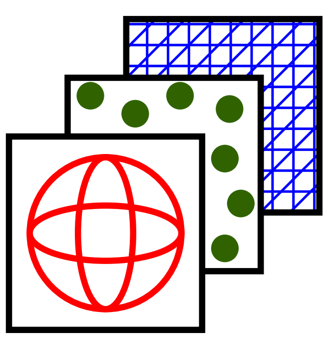
$\Downarrow$

???

[Komiske, Metodiev, JDT, 1810.05165]

Into the Network

Symmetries & Safety



Deep Sets for Particle Jets

The Power of Addition

Additive Observable: $$\mathcal{O} = \sum_{i \in \text{jet}} \Phi(E_i, \vec{p}_i, \ldots)$$

(often comes up in the context of SCET factorization theorems)

Permutation invariant by construction
Easily adapts to variable-length inputs
Can approximate Φ with neural networks
Can incorporate additional particle properties
Linear runtime in number of particles

Additive Safe Observable: $$\mathcal{O} = \sum_{i \in \text{jet}} E_i \, \Phi(\hat{p}_i) \qquad \hat{p}_i = \frac{\vec{p}_i}{E_i}$$

IRC safety guaranteed by energy weighting

Conjectured Generalization

Arbitrary permutation-symmetric observable?

$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_\ell)$$

Latent Additive Observables

$$\mathcal{O}_a = \sum_{i \in \mathrm{jet}} \Phi_a(E_i, \vec{p}_i, \ldots)$$

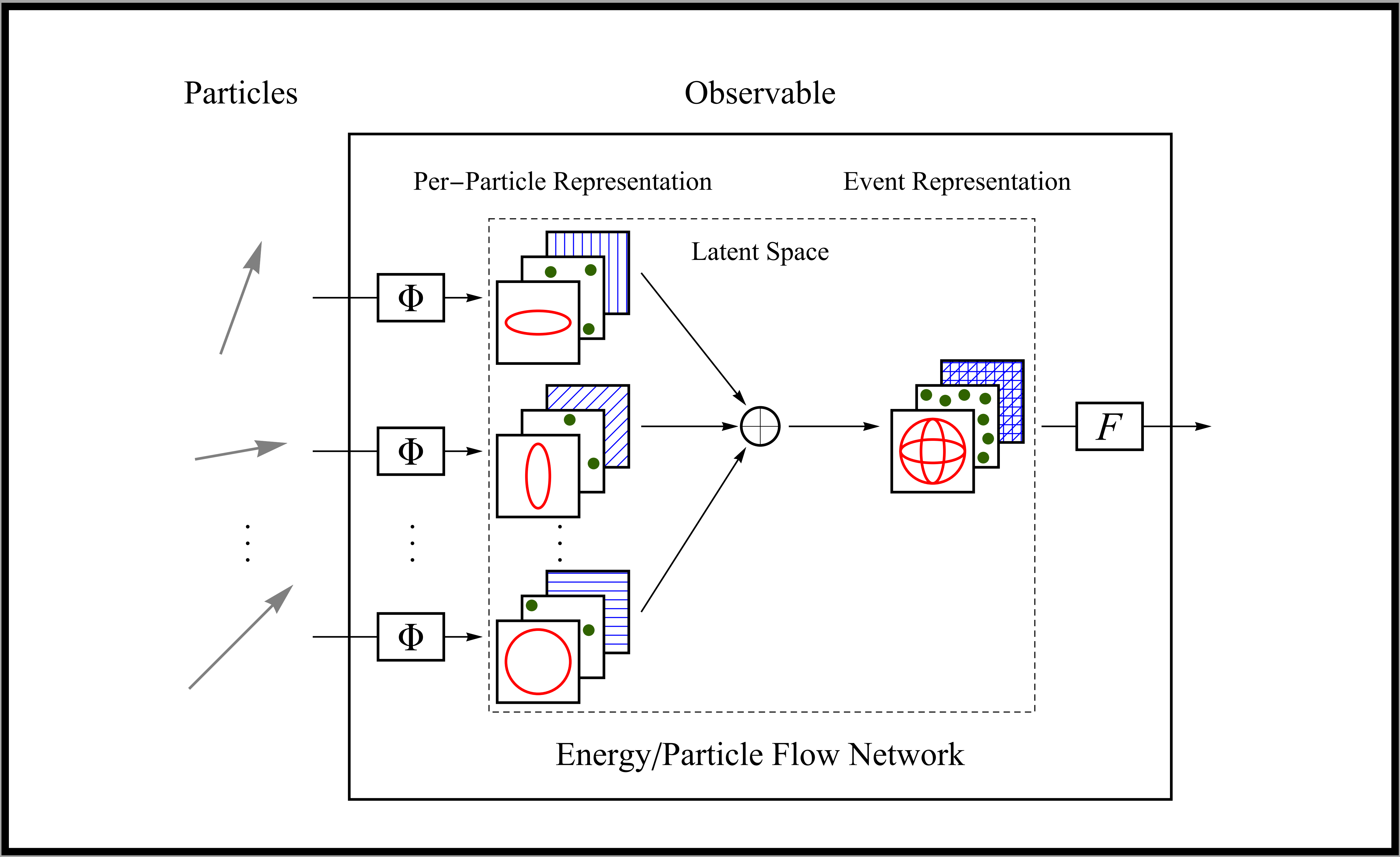
Energy / Particle Flow Networks

IRC-safe Φ

General Φ

[Komiske, Metodiev, JDT, 1810.05165]

Conjectured Generalization

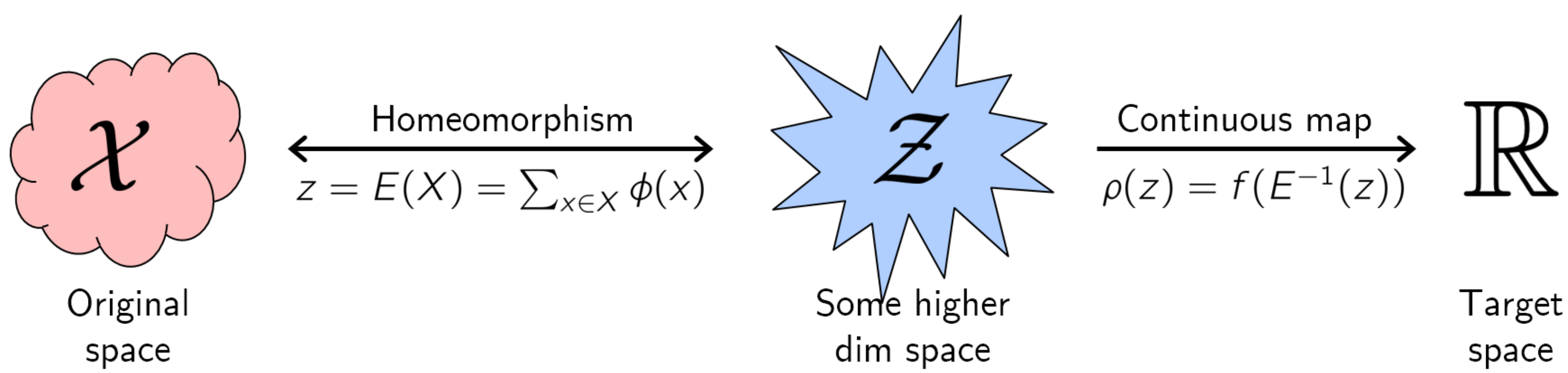


[Komiske, Metodiev, JDT, 1810.05165]

Meanwhile in ML-Land: Deep Sets

Theorem 2 *A function $f(X)$ operating on a set X having elements from a countable universe, is a valid set function,* i.e., ***invariant*** *to the permutation of instances in X, iff it can be decomposed in the form $\rho\left(\sum_{x \in X} \phi(x)\right)$, for suitable transformations ϕ and ρ.*

(!)



Variable-Length Unordered Set of Particles

Additive Latent Space

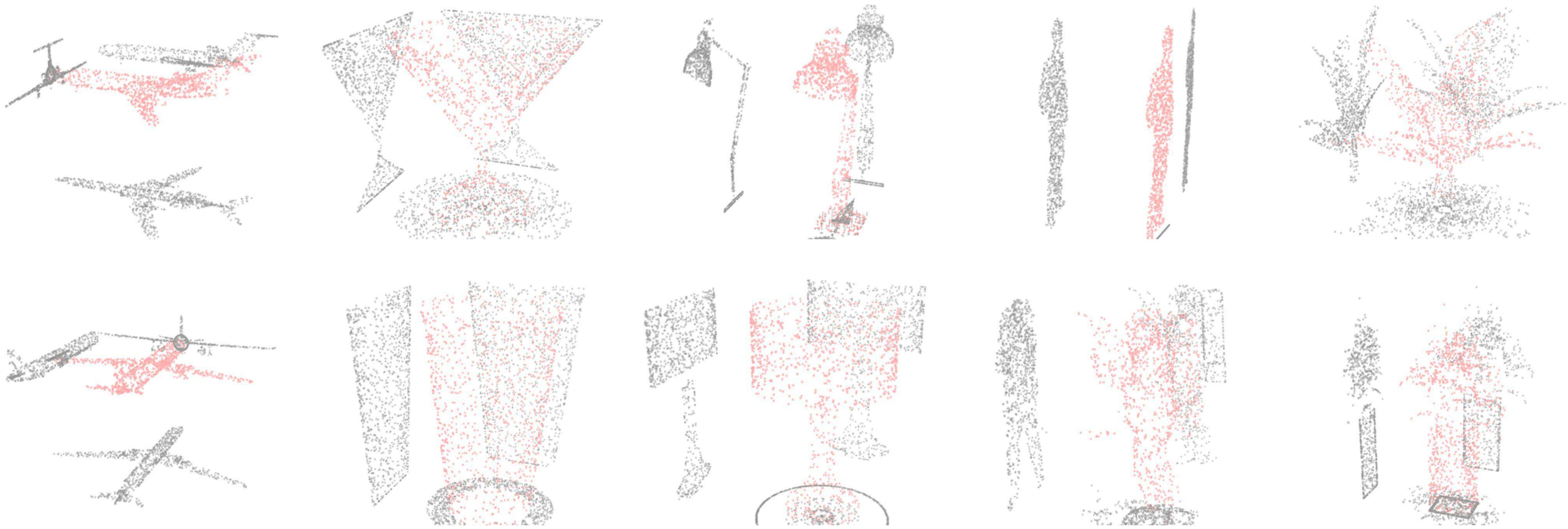
Generic Observable

[Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114]

Deep Sets for…

Celebrity Face Anomaly Detection

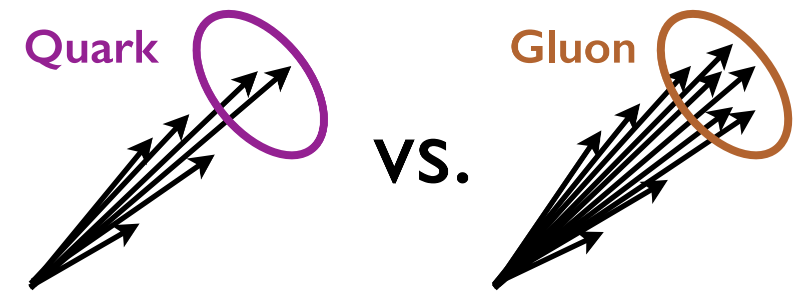
Point Cloud Classification



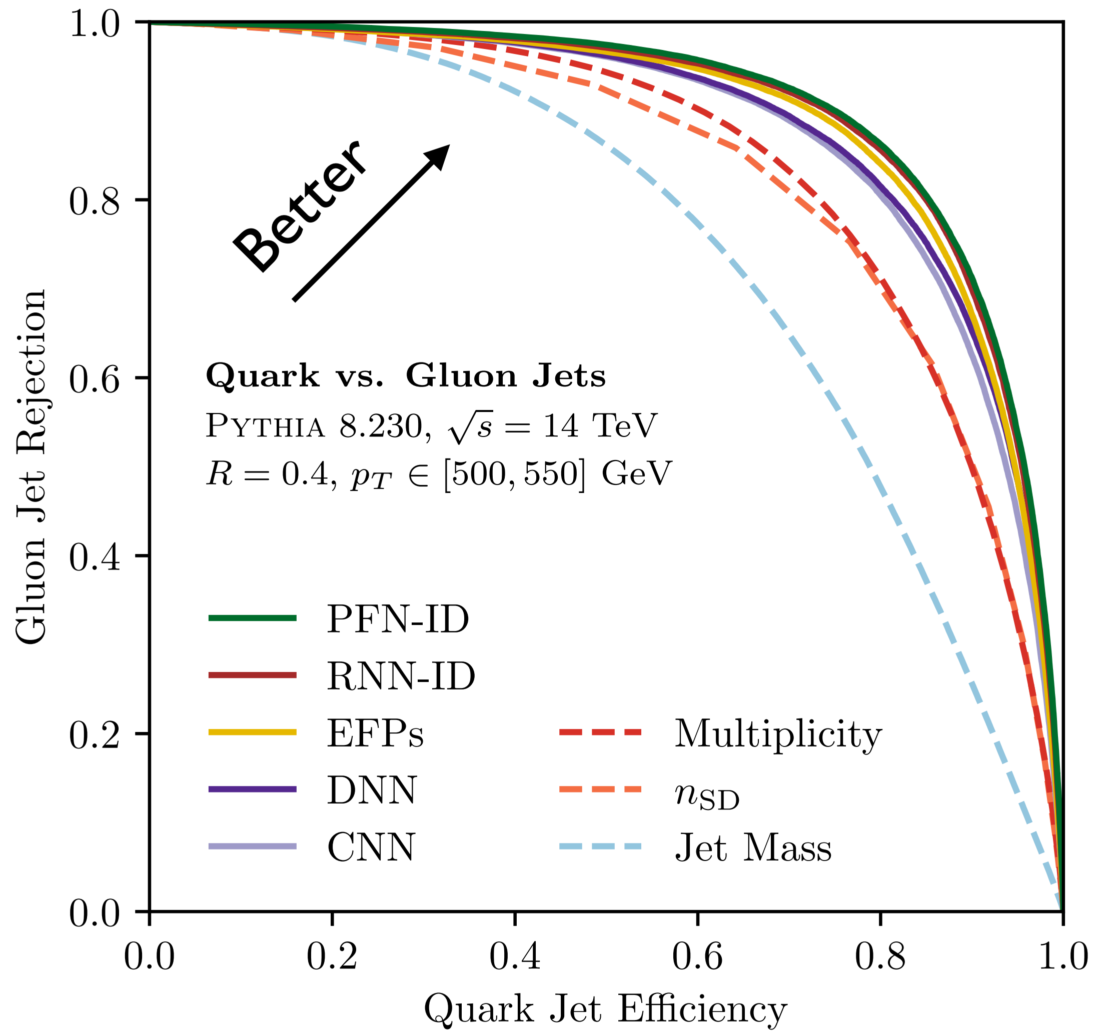
[Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114]

Deep Sets for Q vs. G

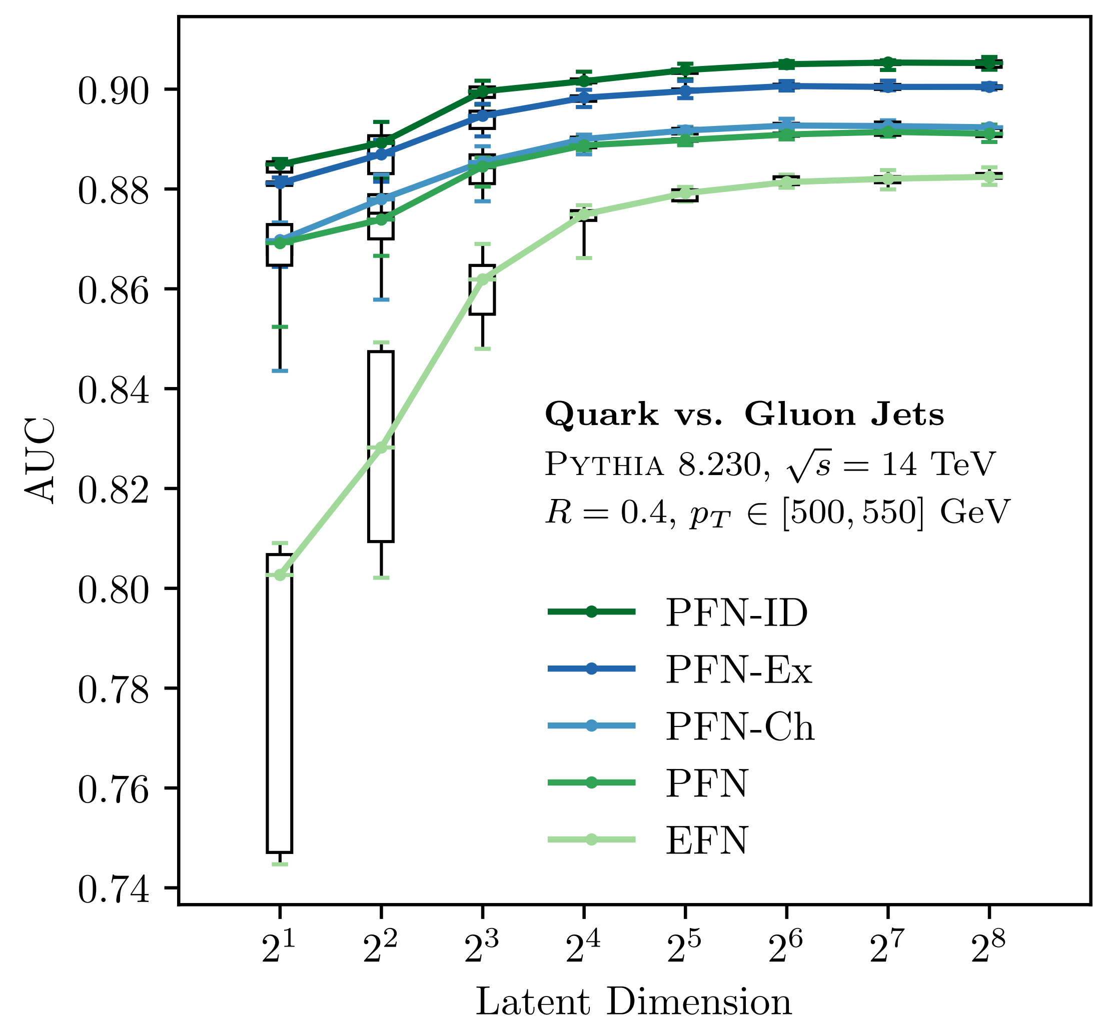
The "Hello, World!" of jet classification



Competitive with previous methods



Performance improves with more information

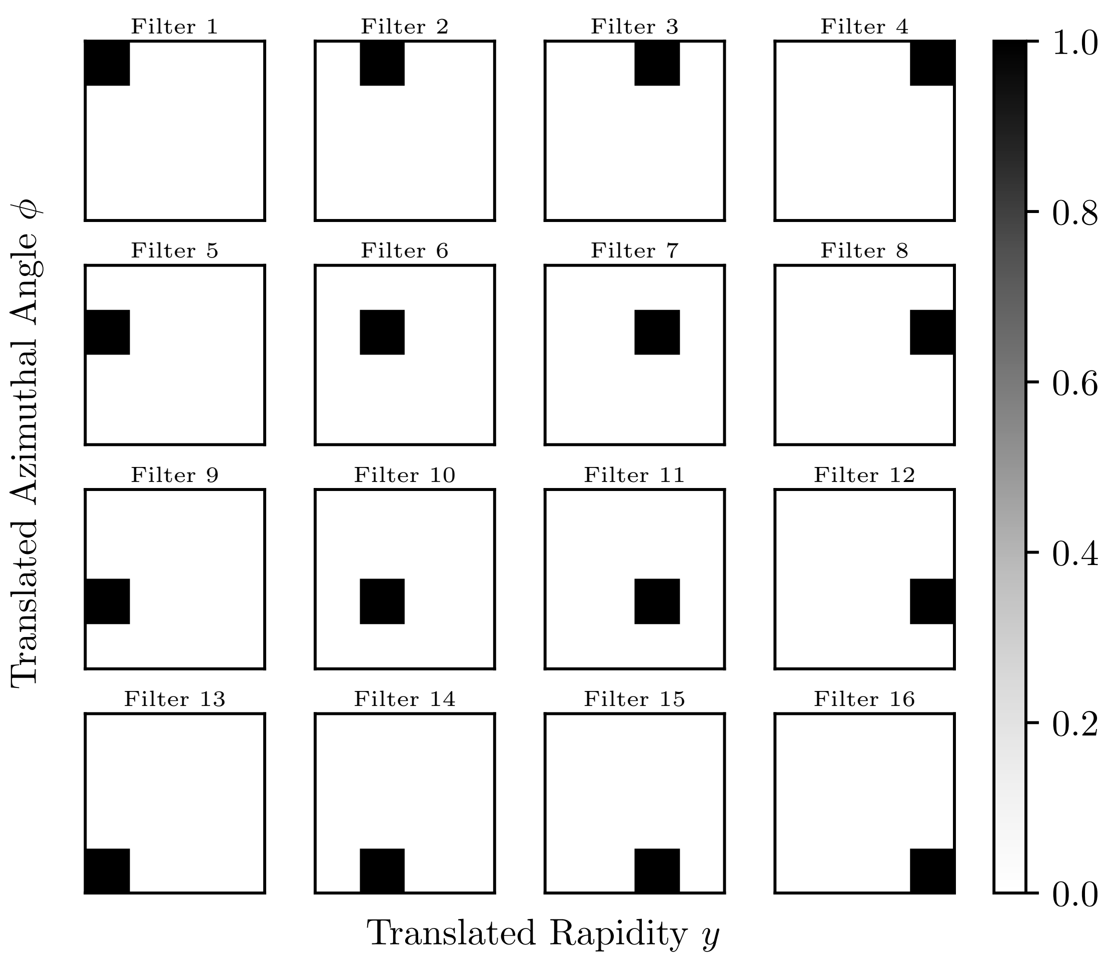


Latent Space Visualization

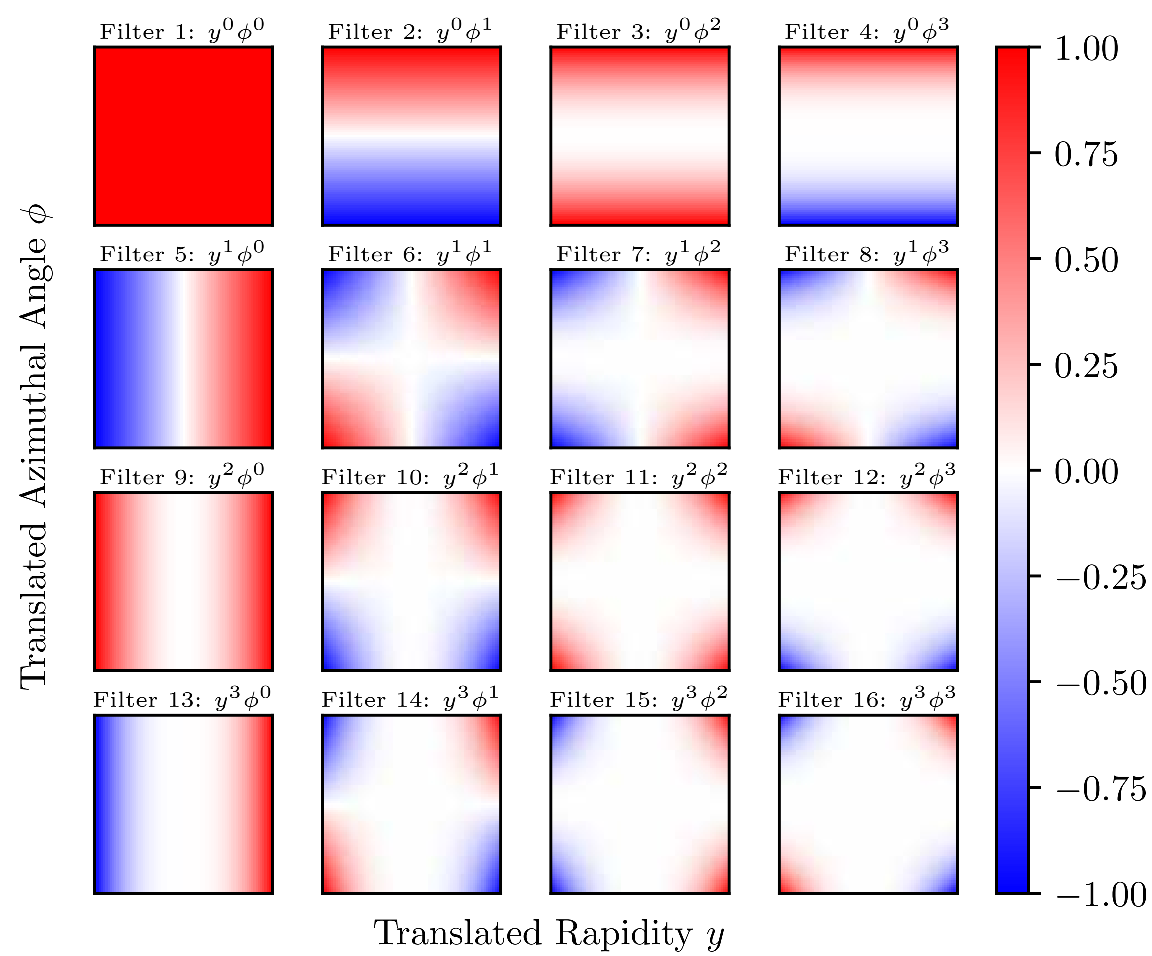
$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_\ell)$$

IRC-safe: $\mathcal{O}_a = \sum_{i \in \text{jet}} p_{Ti} \, \Phi_a(\phi_i, y_i)$

Calorimeter Pixels



Radiation Moments

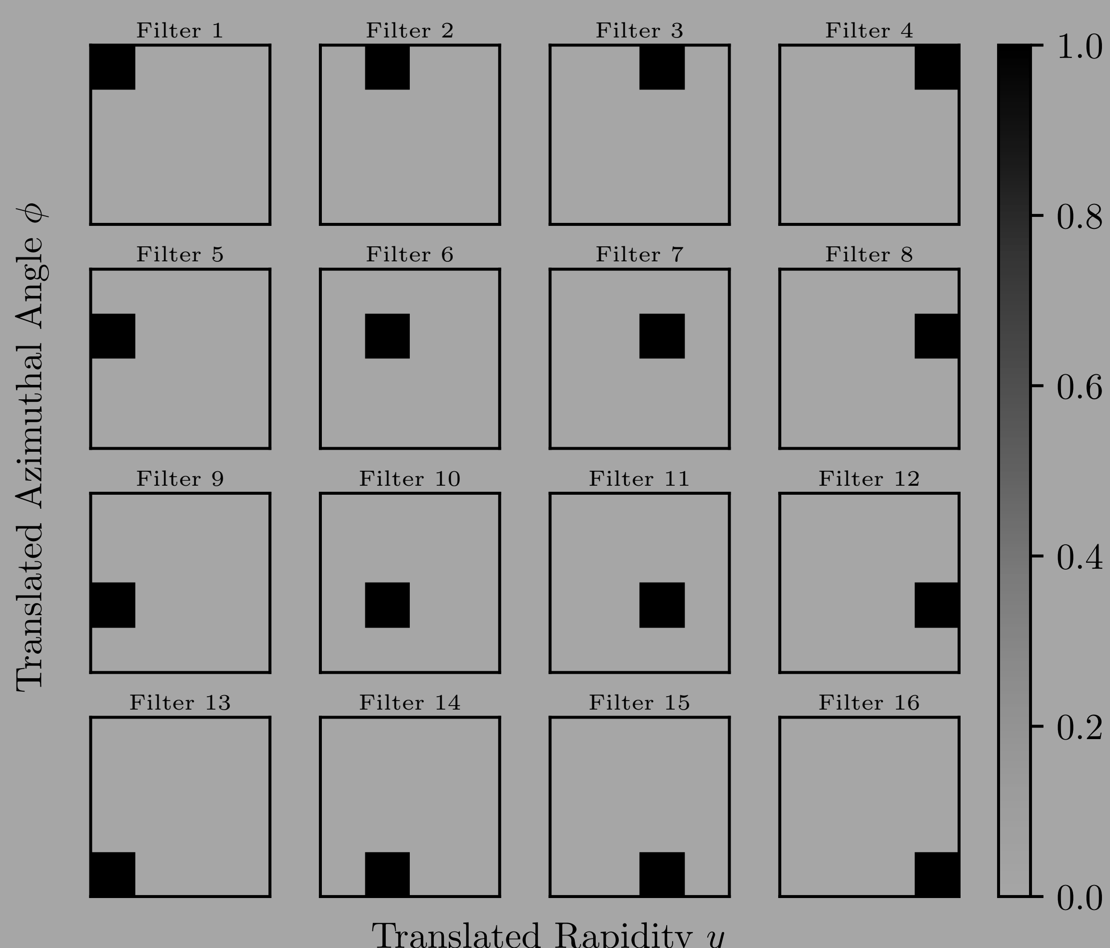


Latent Space Visualization

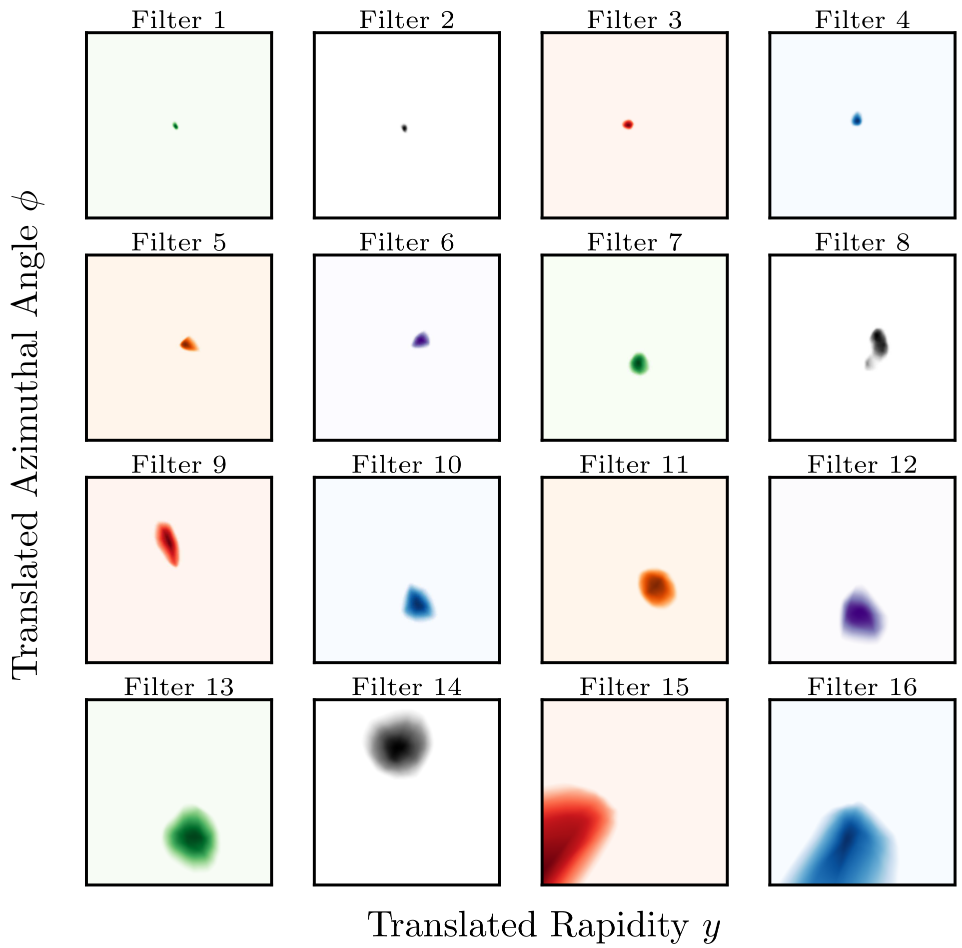
$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_\ell)$$

IRC-safe: $$\mathcal{O}_a = \sum_{i \in \text{jet}} p_{Ti}\, \Phi_a(\phi_i, y_i)$$

Calorimeter Pixels

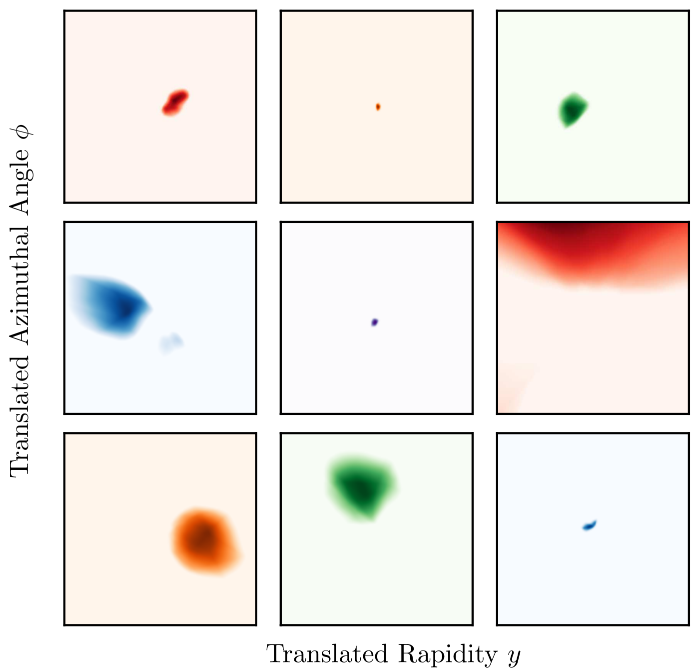


EFNs: Dynamic Pixilation

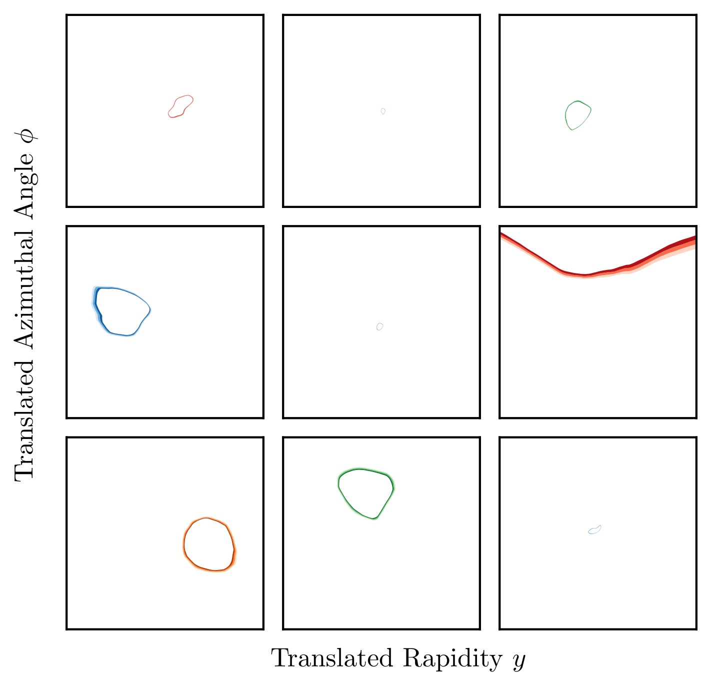


Psychedelic Visualization

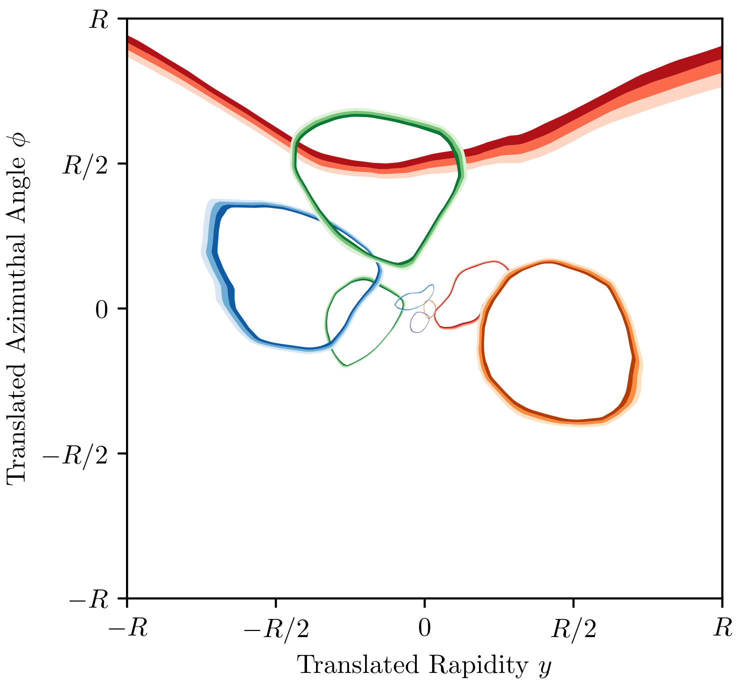
Latent Filters



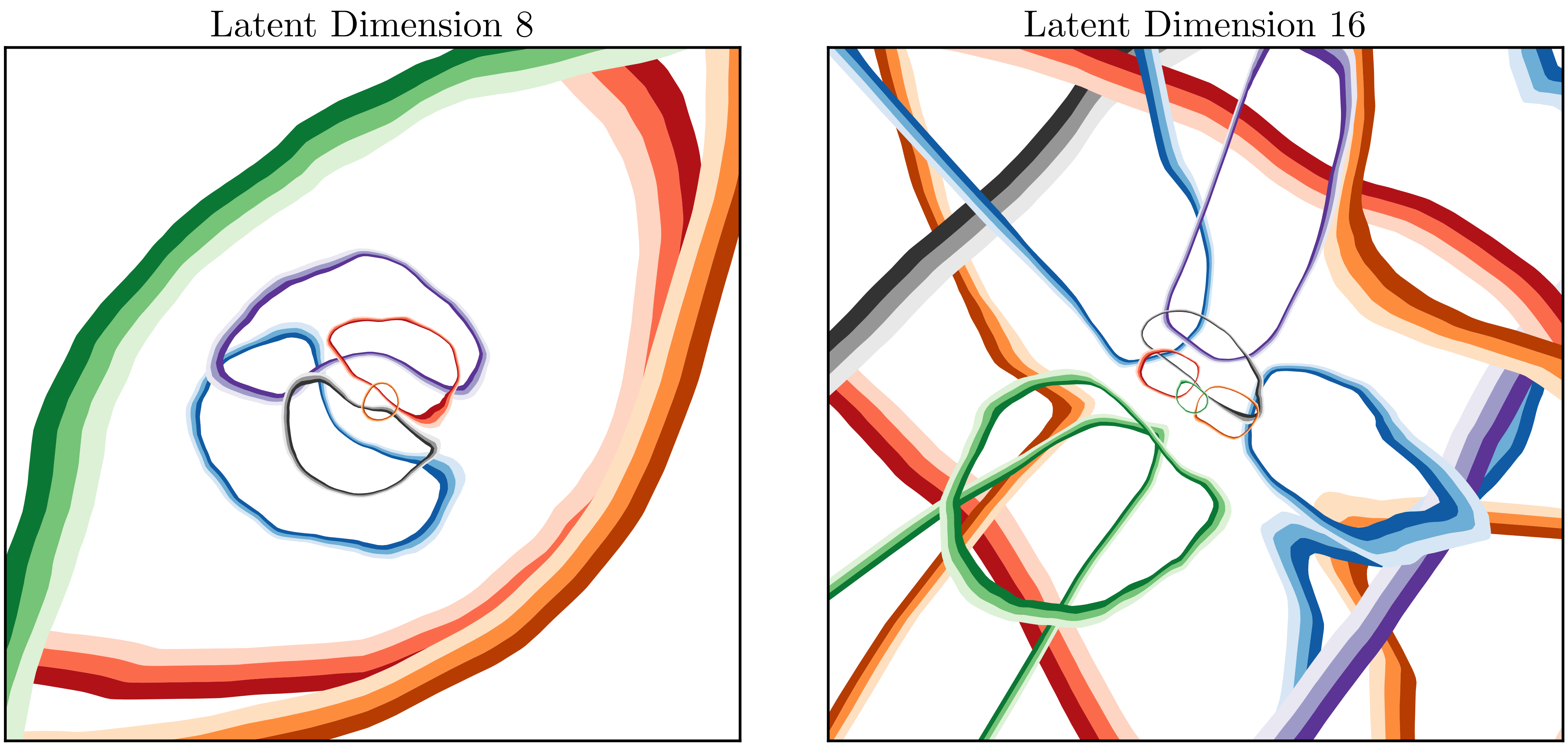
50% Contours



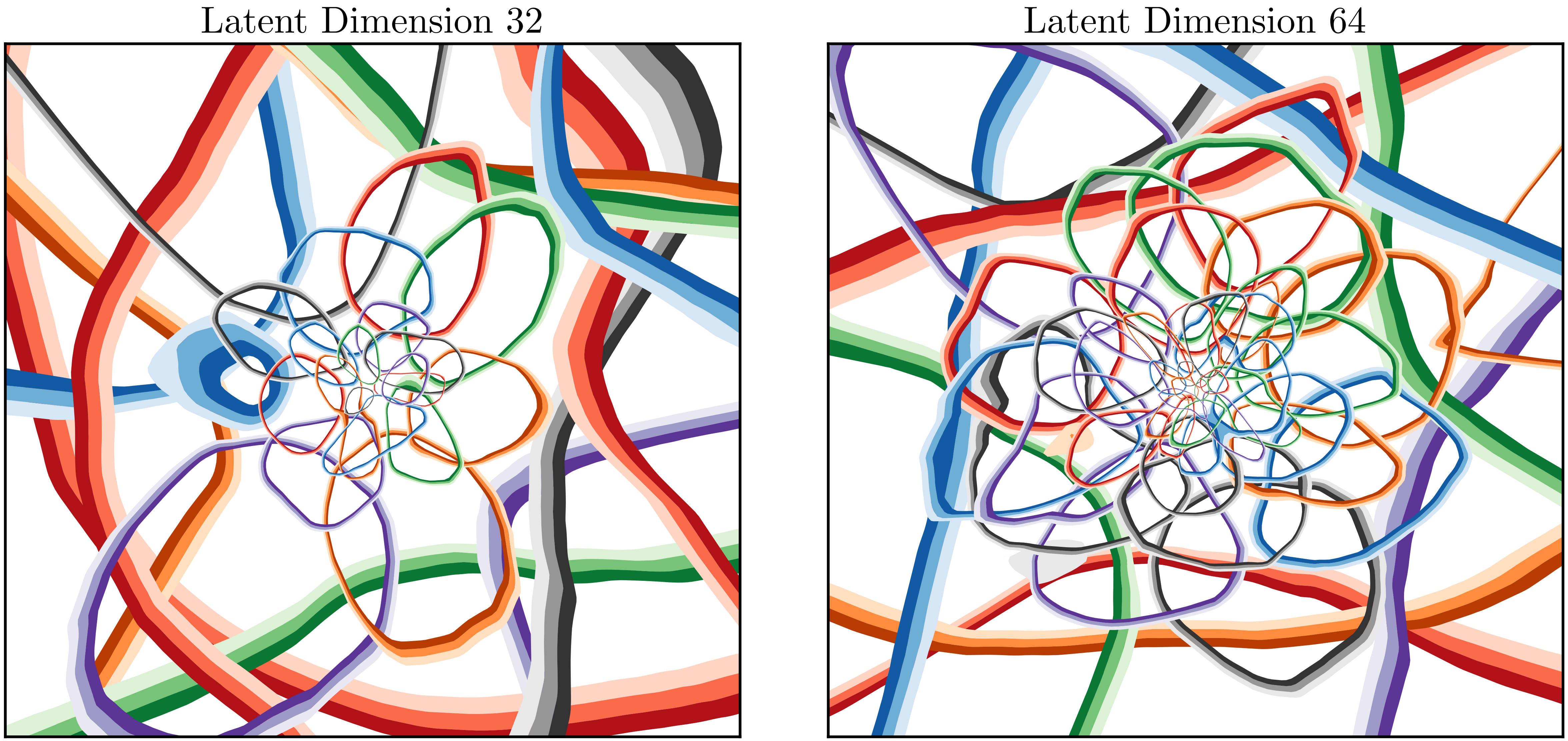
Overlay



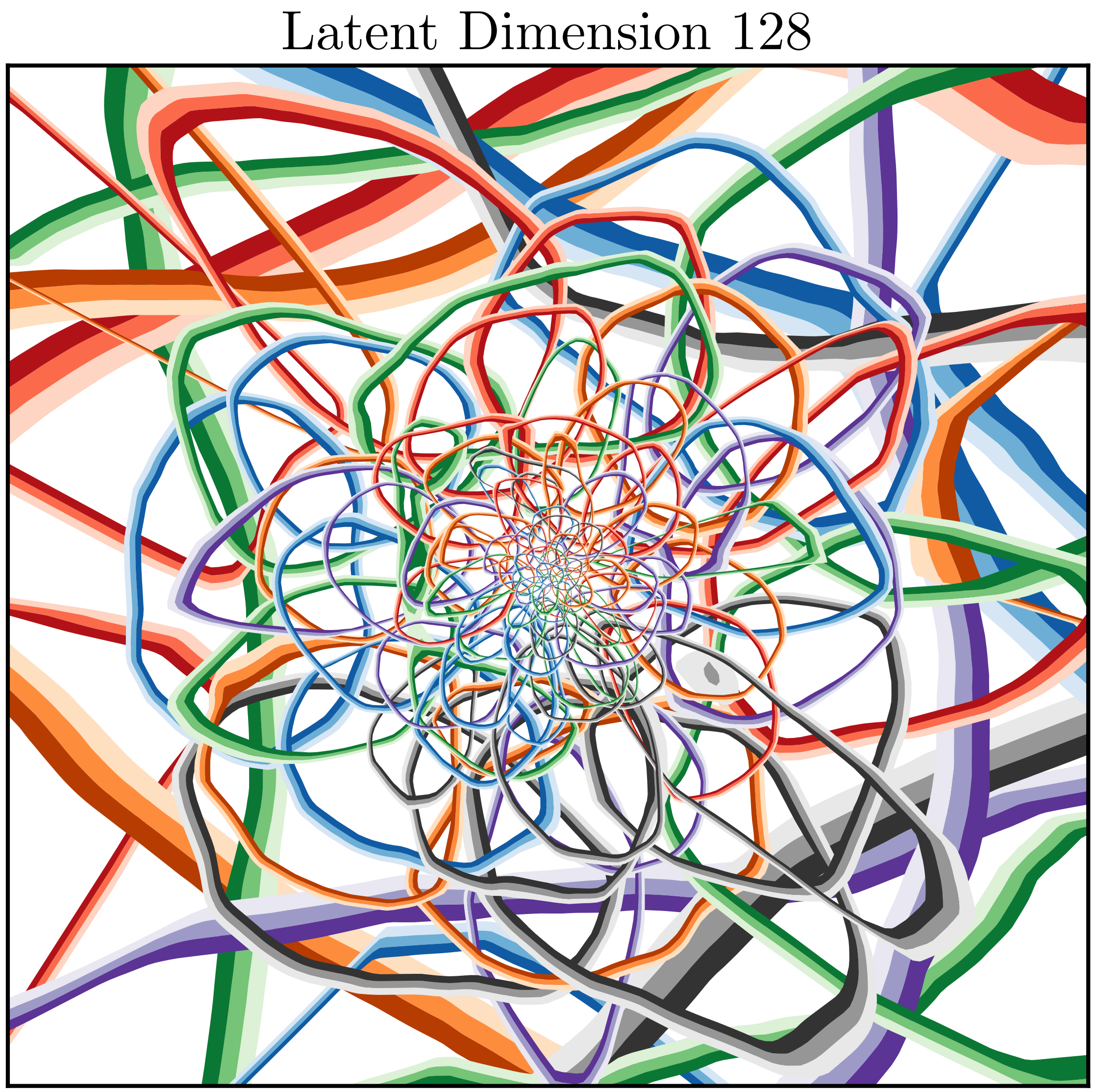
Psychedelic Visualization

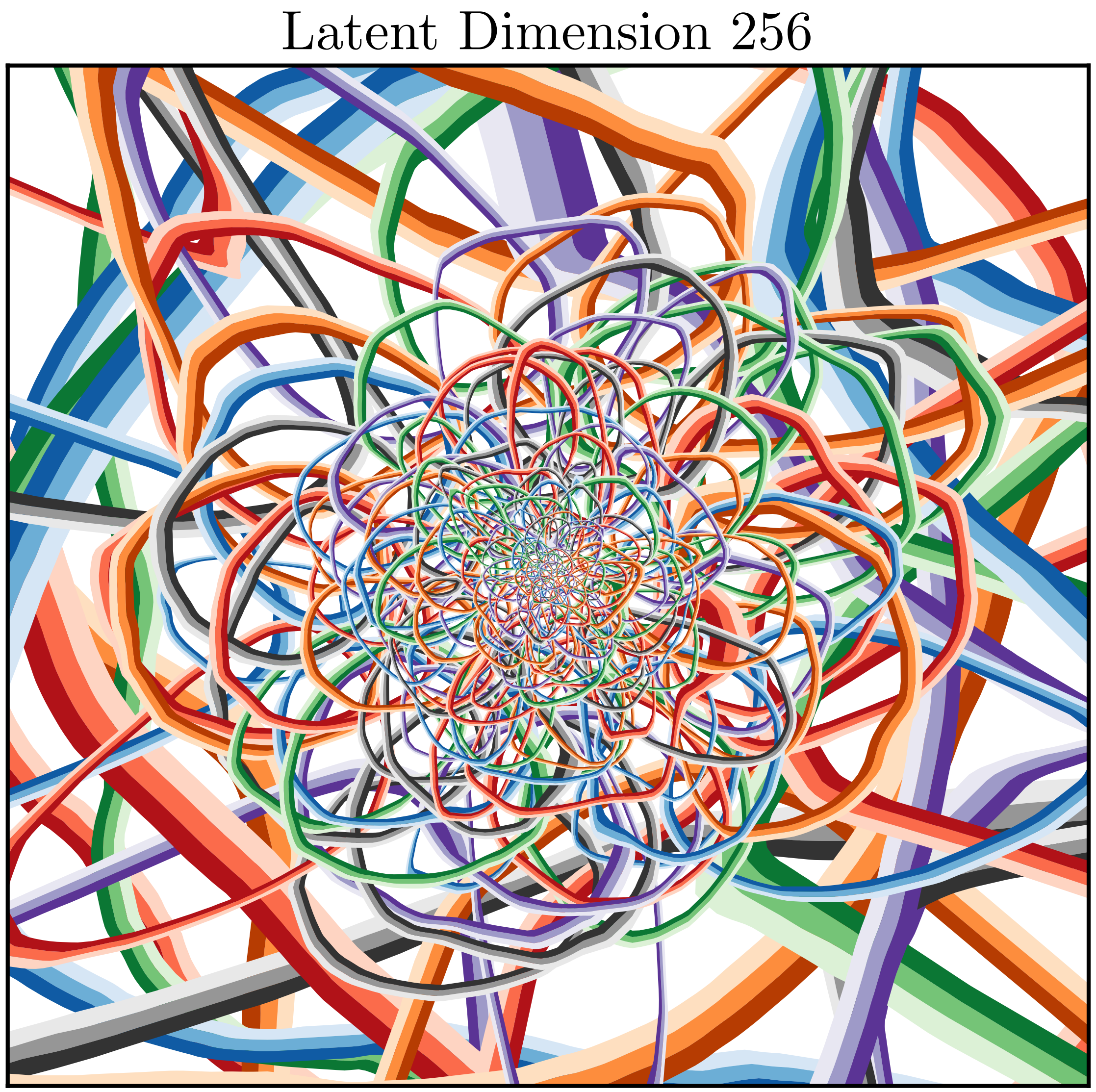


Psychedelic Visualization

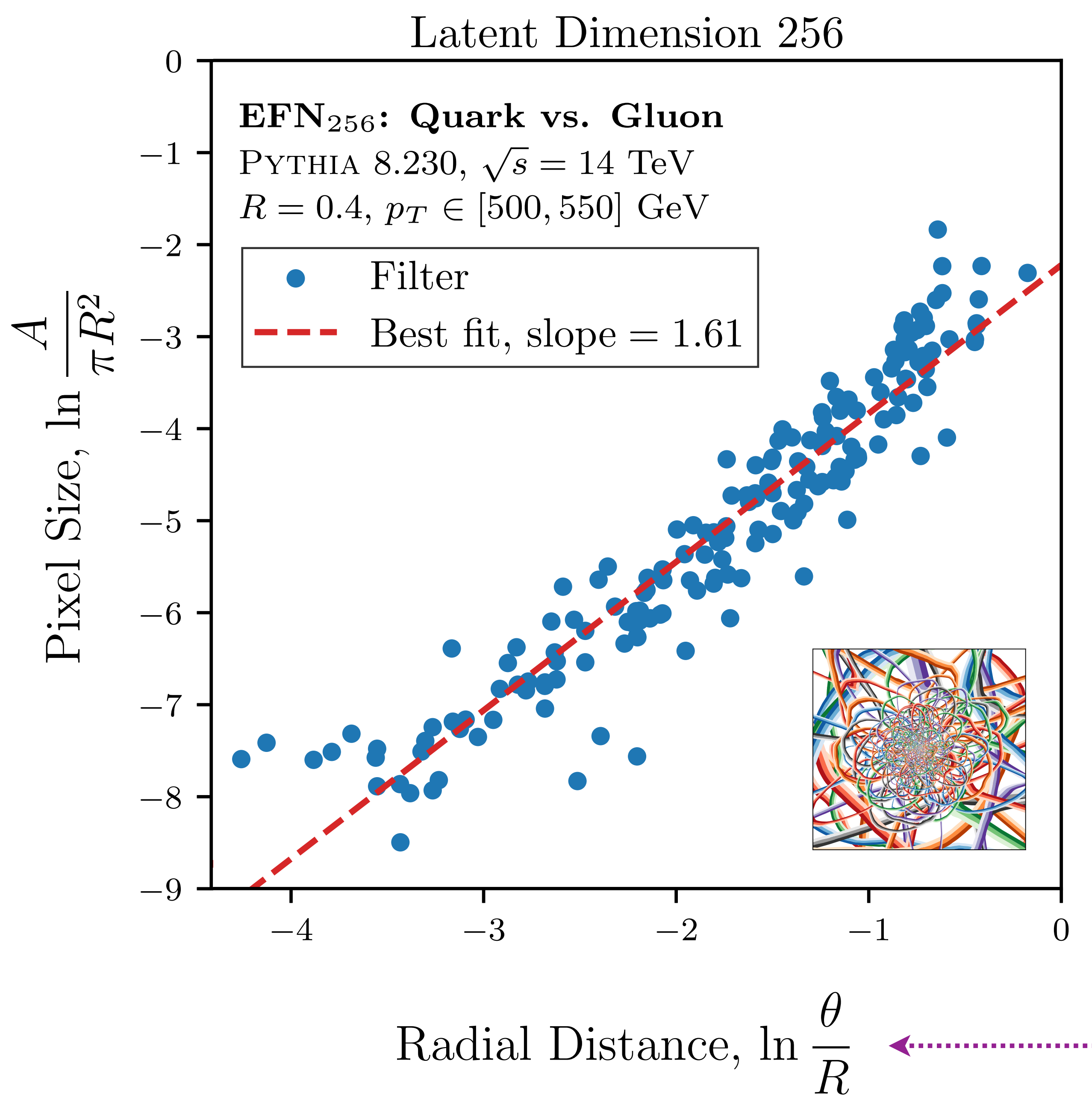


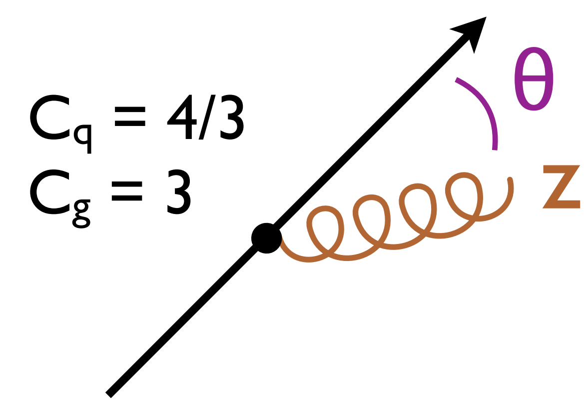
Psychedelic Visualization





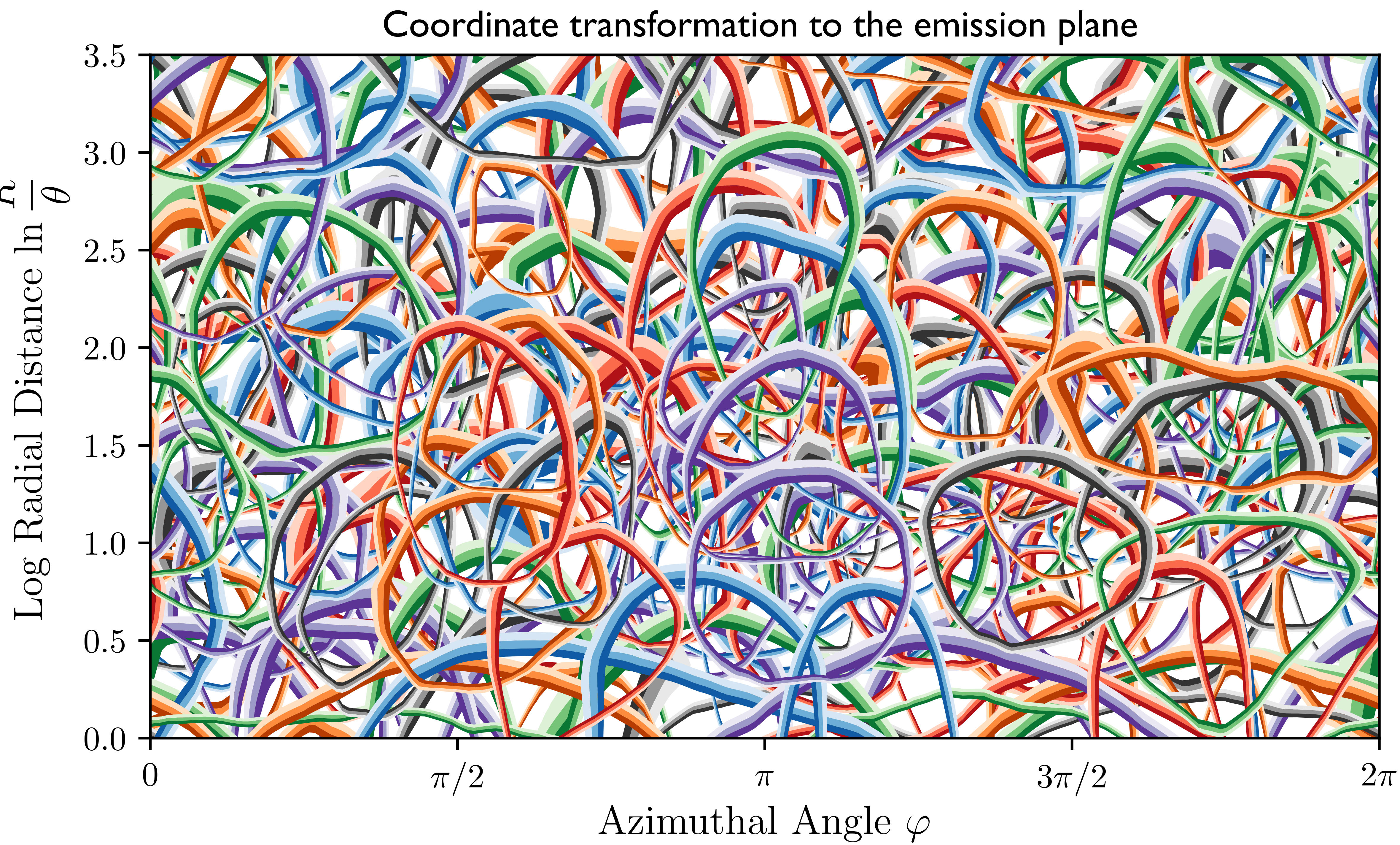
Learning the Singularity Structure of QCD





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

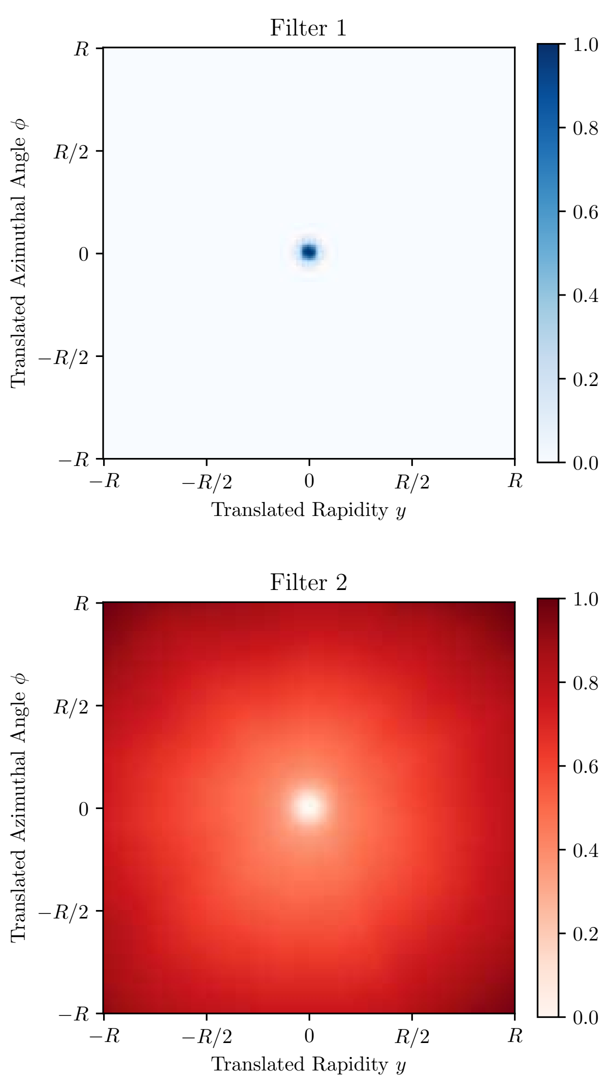
Ready for the MoMA

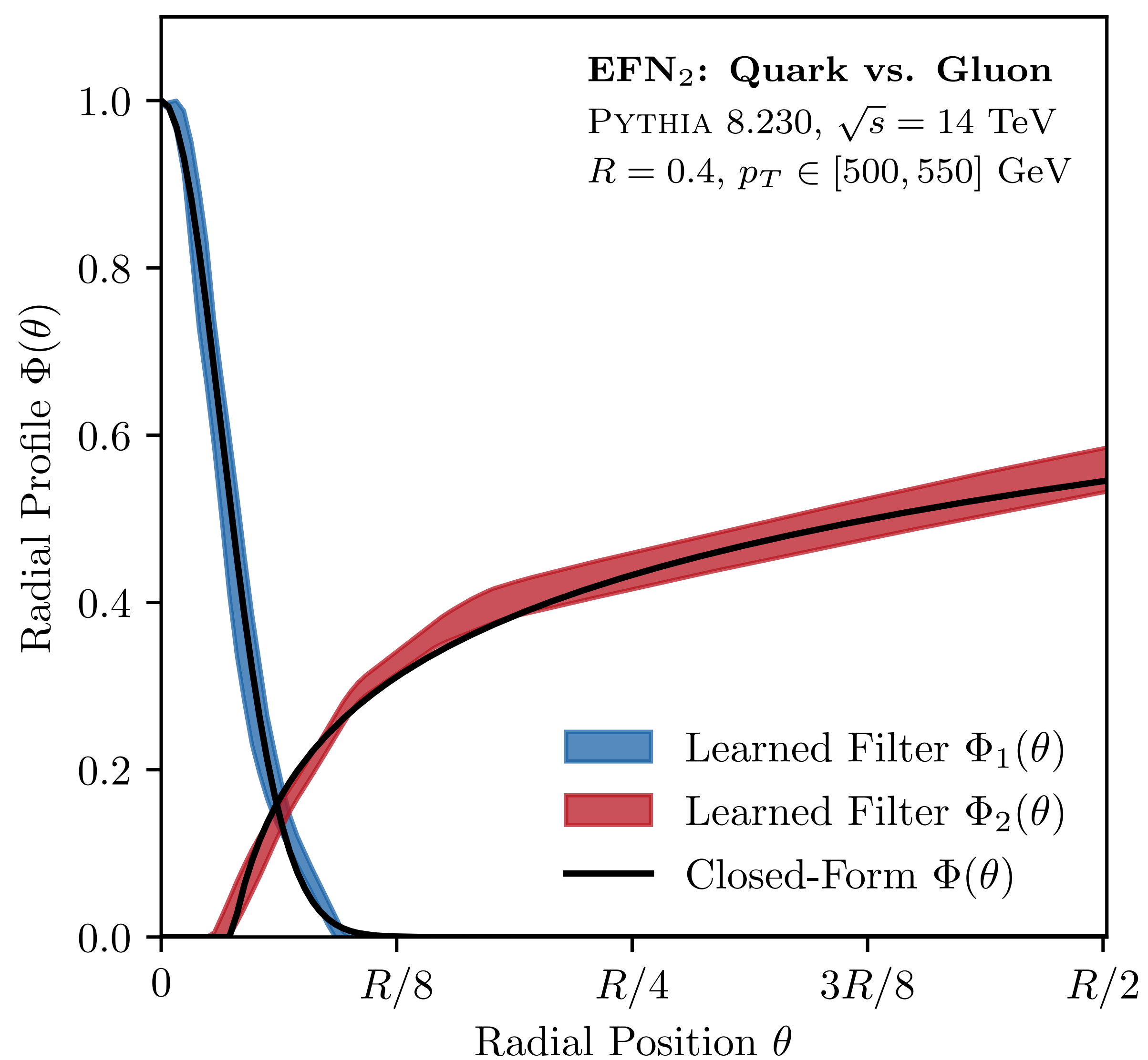


“What is the Machine Learning?”

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

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

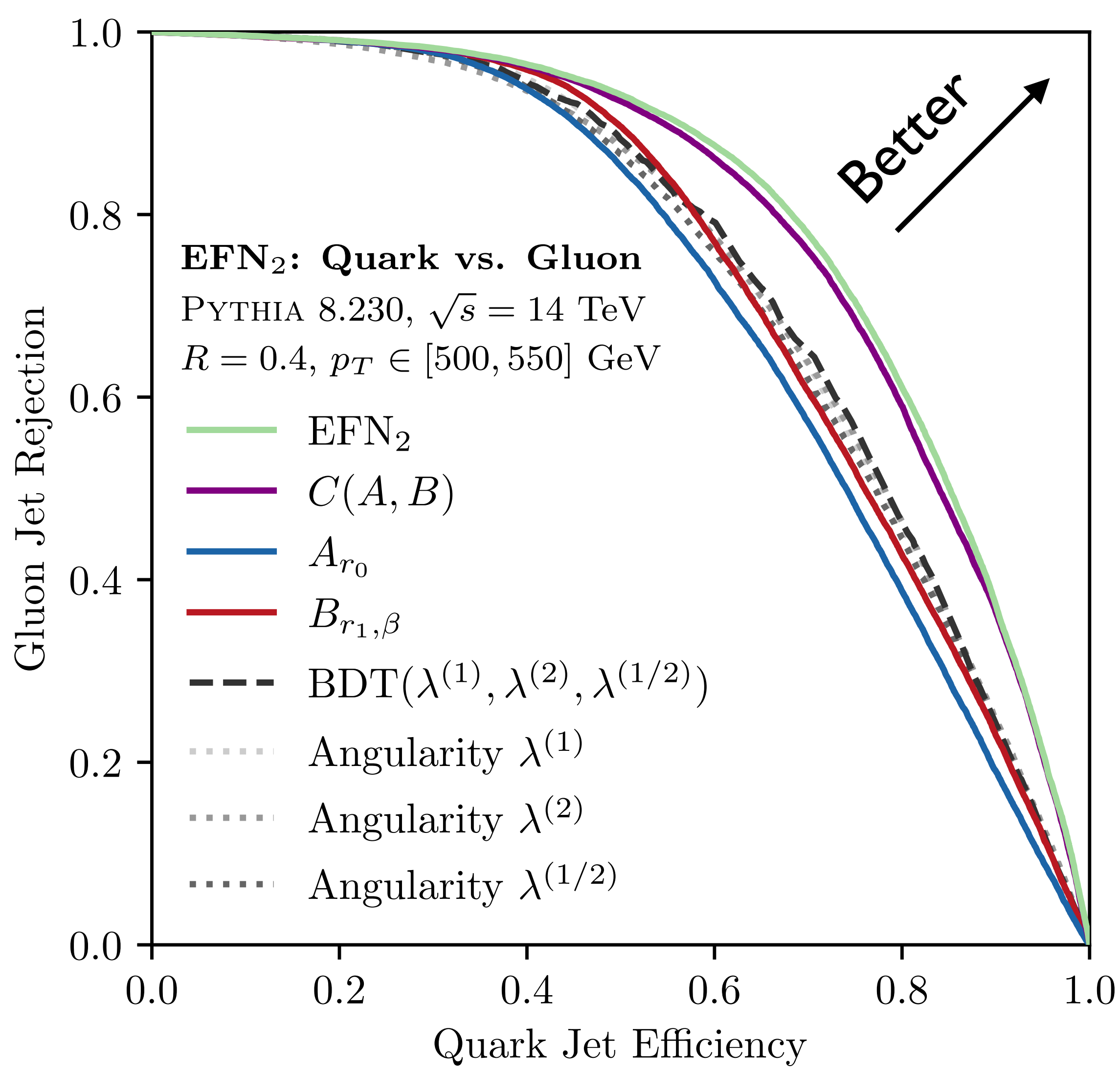




“What is the Machine Learning?”

For ℓ = 2, 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 successfully reverse engineered the machine

The Broader Lesson

&

"Deep Learning" ~~vs.~~ "Deep Thinking"

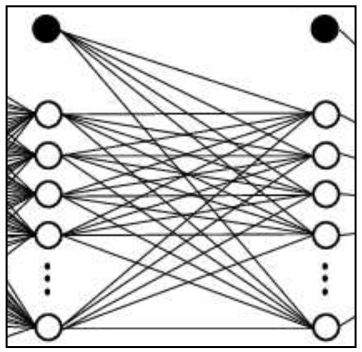
Advances in mathematics and computer science (Deep Sets)

$\Downarrow \quad \Uparrow$

Advances in collider physics (EFN/PFN)

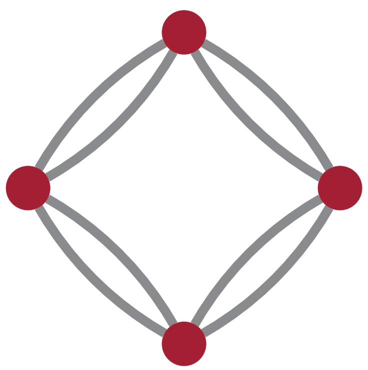
Weighted Point Sets $\Leftrightarrow$ *IRC Safety*

Summary



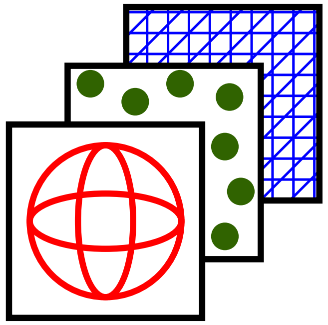
Into the Network

Embracing the rise of machine learning for collider physics



Symmetries & Safety

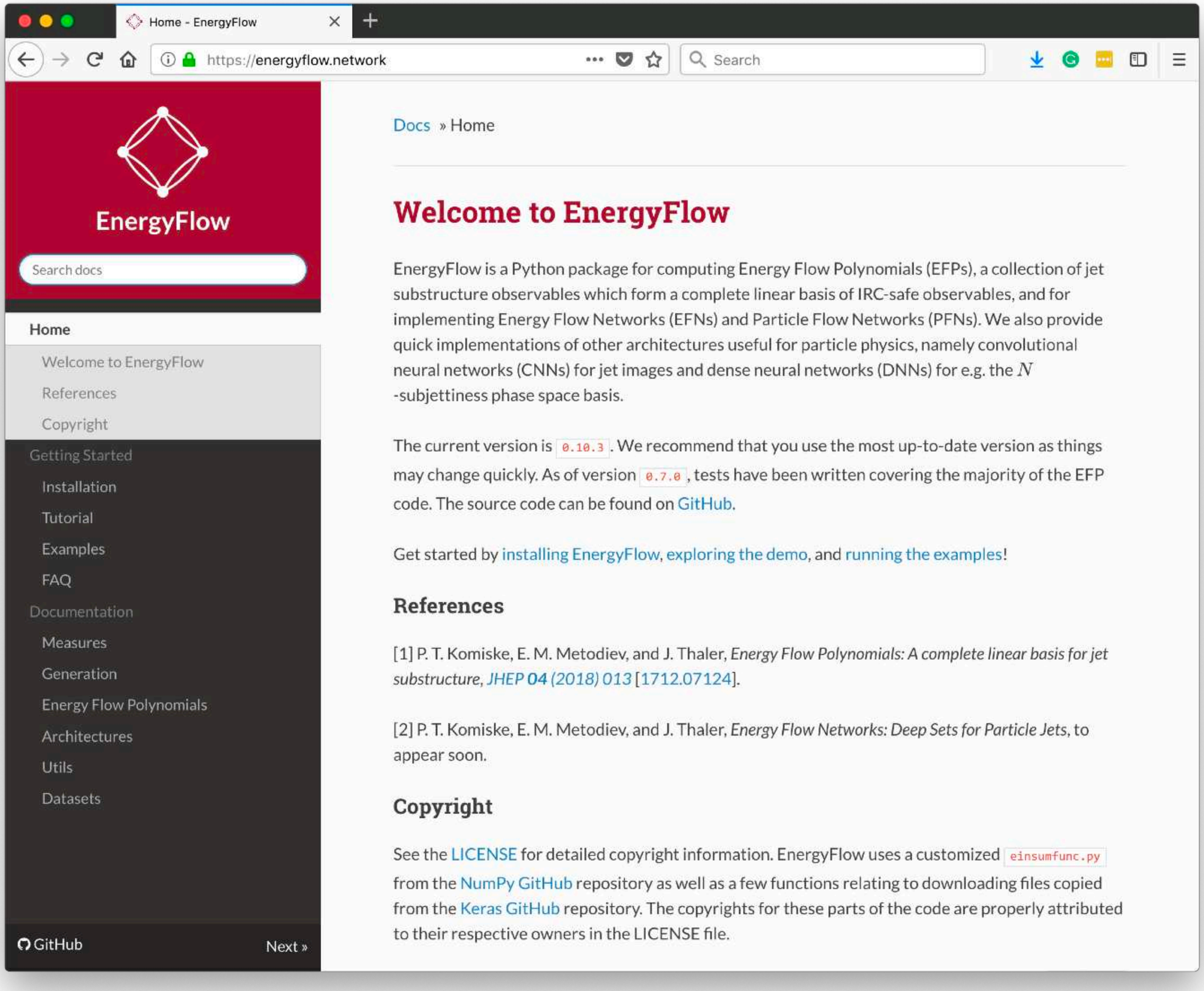
Importance of indistinguishability and energy weighting



Deep Sets for Particle Jets

EFN/PFNs as a step towards robustness, versatility, and transparency

energyflow.network



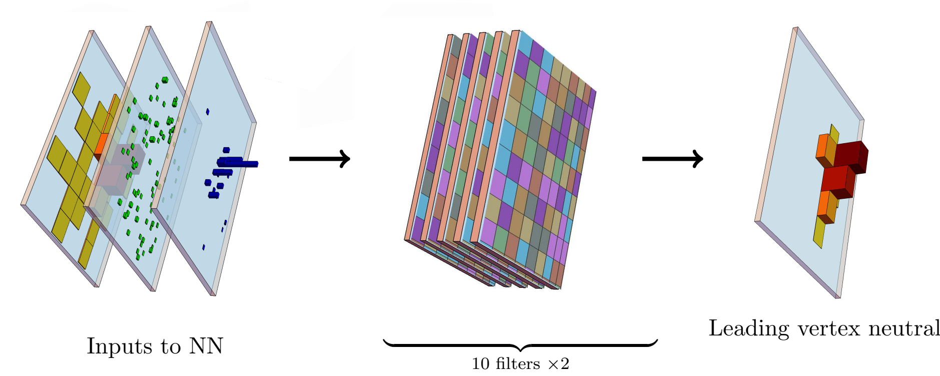
Backup Slides

Beyond Classification

PUMML

Pileup Mitigation

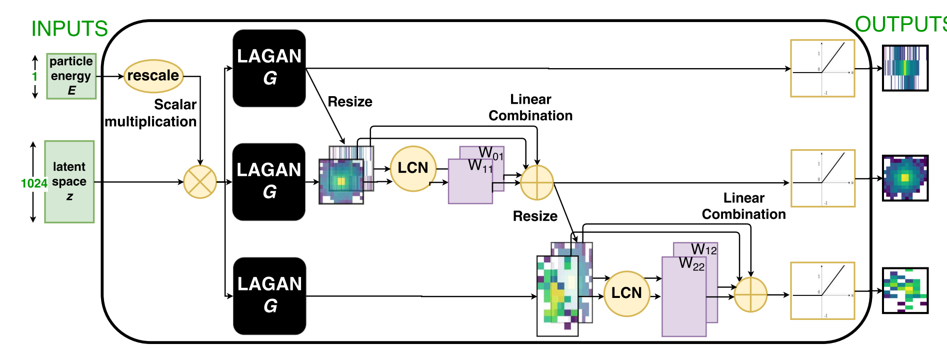
[Komiske, Metodiev, Nachman, Schwartz, 1707.08600]



CaloGAN

Fast Detector Simulation

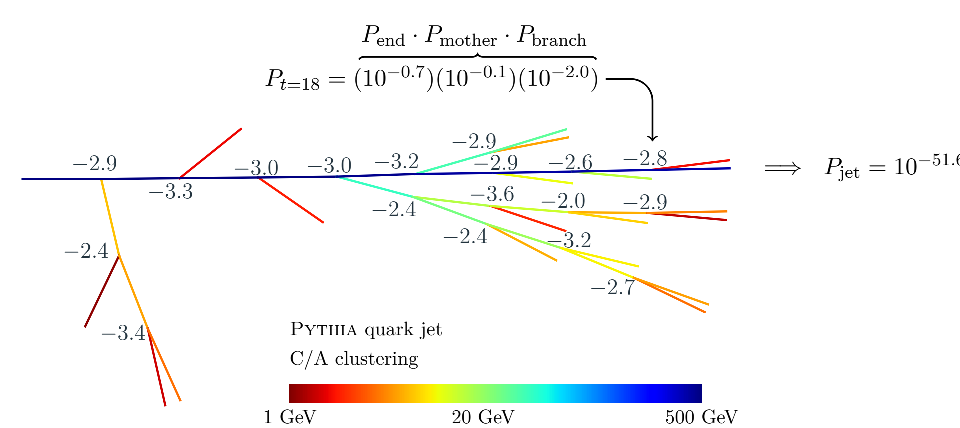
Paganini, de Oliveira, Nachman, 1705.02355, 1712.10321;
see also de Oliveira, Michela Paganini, Nachman, 1701.05927]



JUNIPR

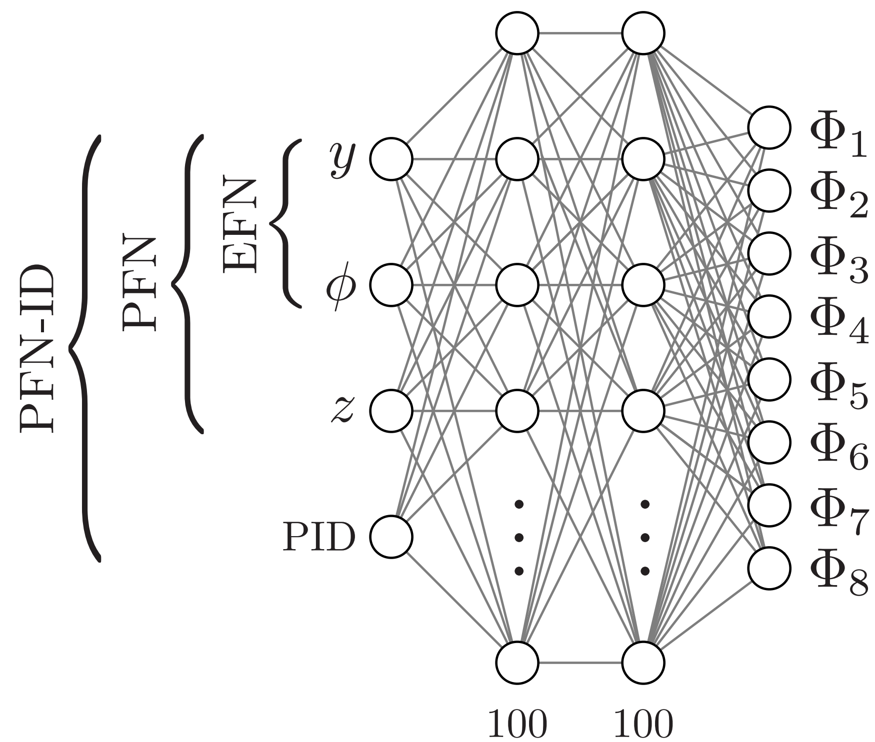
Probability Modeling

[Andreassen, Feige, Frye, Schwartz, 1804.09720;
see also Monk, 1807.03685]

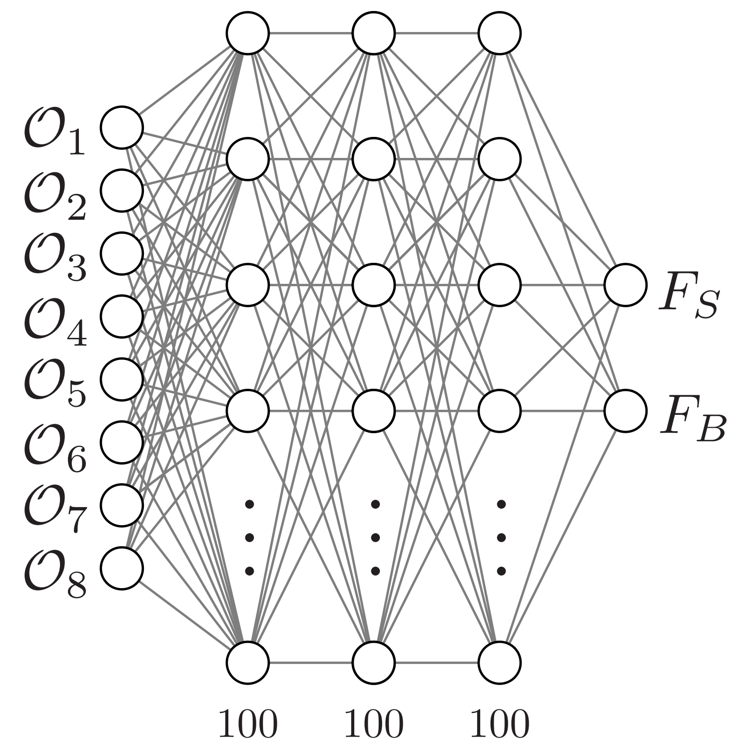


Architecture Details

Per-Particle:



Latent Combiner: F



Per-Jet (Latent Space):

EFN: $\mathcal{O}_a = \sum_{i \in \text{jet}} z_i \, \Phi_a(y_i, \phi_i)$

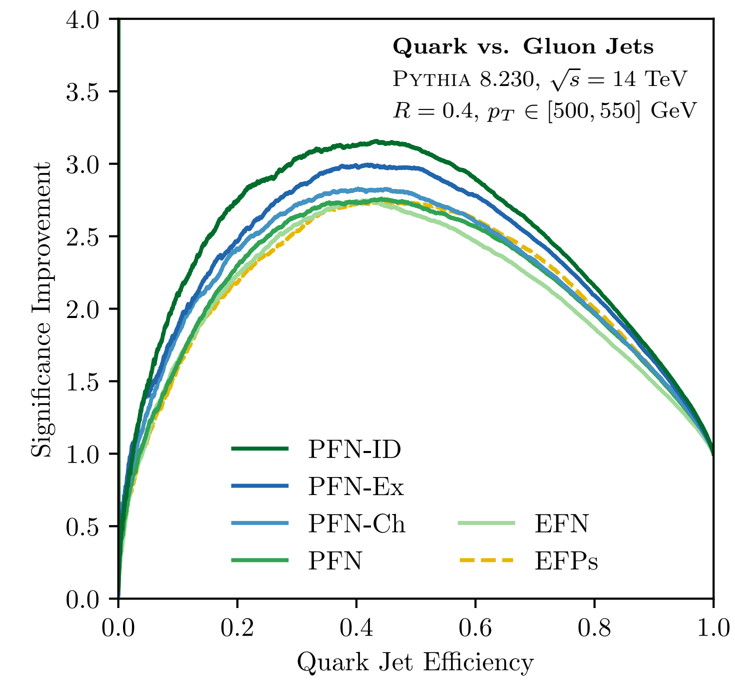
PFN: $\mathcal{O}_a = \sum_{i \in \text{jet}} \Phi_a(y_i, \phi_i, z_i, \text{PID}_i)$

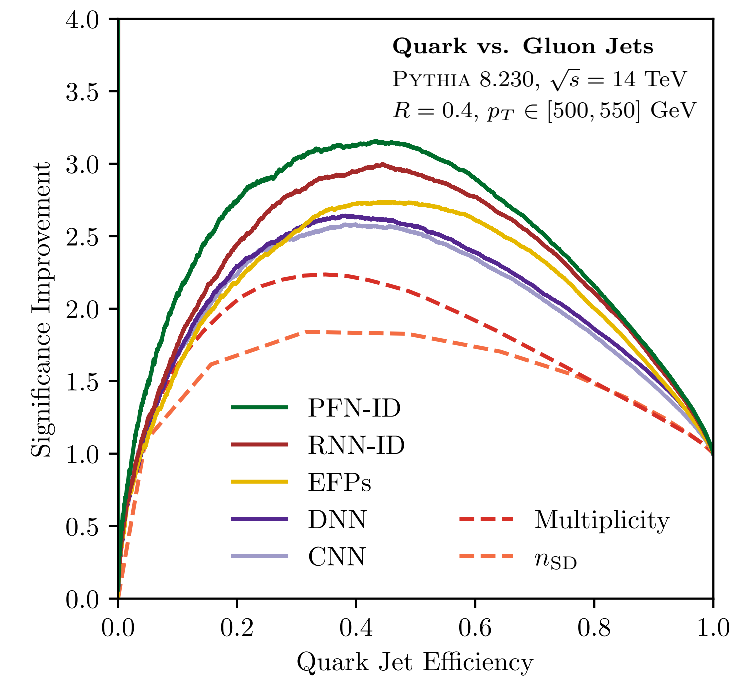
Final Discriminant:

$$\text{softmax}(F_S, F_B)$$

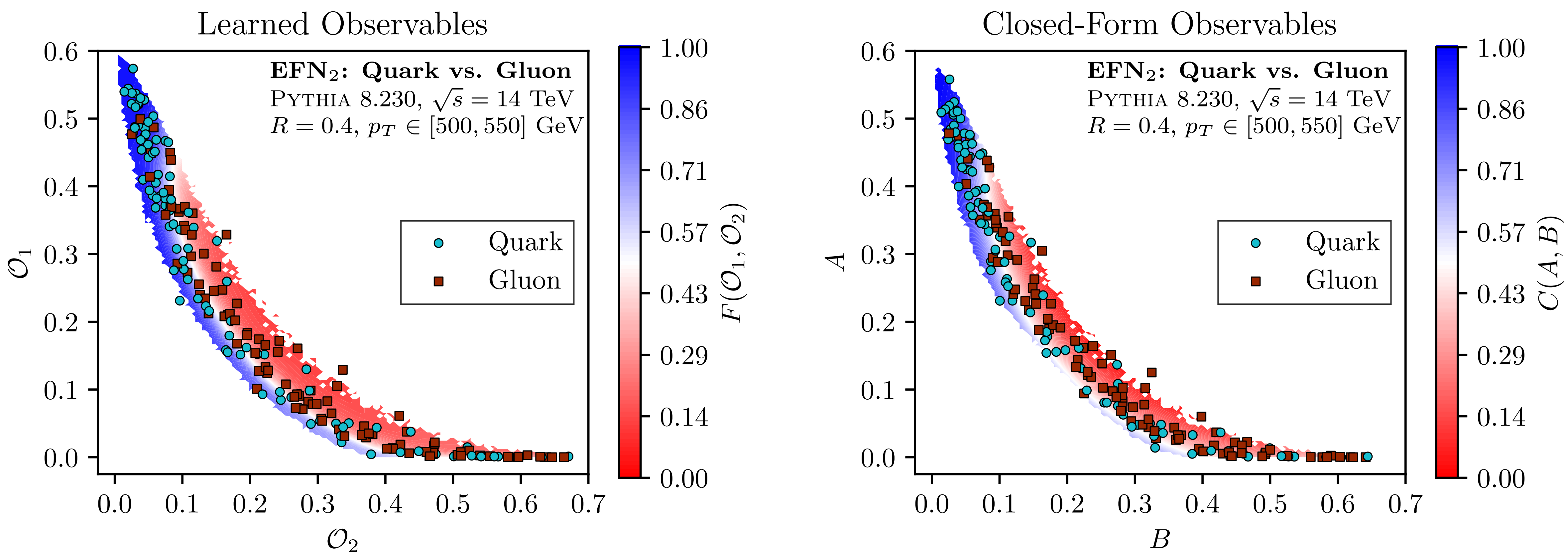
More Quark/Gluon Performance

Model	**AUC**	$\boldsymbol{1/\varepsilon_g}$ **at** $\boldsymbol{\varepsilon_q = 50\%}$
PFN-ID	$\mathbf{0.9052} \pm 0.0007$	$\mathbf{37.4} \pm 0.7$
PFN-Ex	0.9005 ± 0.0003	34.7 ± 0.4
PFN-Ch	0.8924 ± 0.0001	31.2 ± 0.3
PFN	0.8911 ± 0.0008	30.8 ± 0.4
EFN	0.8824 ± 0.0005	28.6 ± 0.3
RNN-ID	0.9010	34.4
RNN	0.8899	30.5
EFP	0.8919	29.7
DNN	0.8849	26.4
CNN	0.8781	25.5
M	0.8401	19.0
n_{SD}	0.8297	14.2
m	0.7401	7.2





Reverse Engineering the Machine



Fascinating QCD question about why this is a better strategy than traditional angularities

Energy Flow Polynomials

What is the space of *all* IRC-safe observables?

Examples from Jet Substructure

→ Underlying Physics
Natural Data Representation
Suitable Algorithm

Jet p_T: $\sum_{i \in \text{jet}} p_{T,i}$ IRC Safe

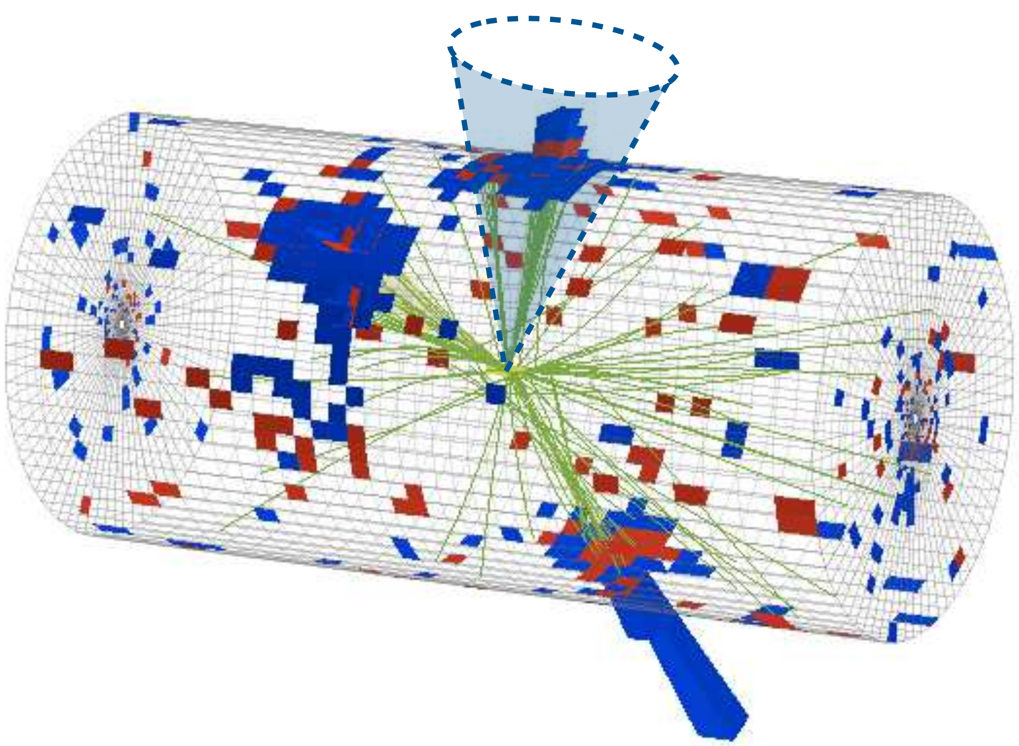
p_T^D: $\sum_{i \in \text{jet}} \frac{p_{T,i}^2}{p_{T\text{jet}}^2}$ IR Safe, C Unsafe
[CMS HIG-11-027]

Multiplicity: $\sum_{i \in \text{jet}} 1$ IRC Unsafe

Jet Mass: $\sum_{i,j \in \text{jet}} p_i \cdot p_j$ IRC Safe

N-subjettiness: $\sum_{i \in \text{jet}} p_{T,i} \min \left\{ \Delta R_{i,1}, \Delta R_{i,2}, \ldots, \Delta R_{i,N} \right\}^\beta$ IRC Safe
[JDT, Van Tilburg, 1011.2268, 1108.2701]

But ratios are only "Sudakov safe"!



A Systematic Expansion

Underlying Physics
→ Natural Data Representation
Suitable Algorithm

Expand* any IRC safe observable in small energy limit

$$\mathcal{S} = \sum_i E_i \, f_1^{\mathcal{S}}(\hat{n}_i) + \sum_{ij} E_i E_j \, f_2^{\mathcal{S}}(\hat{n}_i, \hat{n}_j) + \sum_{ijk} E_i E_j E_k \, f_3^{\mathcal{S}}(\hat{n}_i, \hat{n}_j, \hat{n}_k) + \ldots$$

Form enforced by: Particle Relabeling | Infrared Safety | Collinear Safety

Further expand* each angular function in pairwise angles

$$z_i = \frac{E_i}{E_{\text{jet}}} \qquad \cos\theta_{ij} = \hat{n}_i \cdot \hat{n}_j$$

[Komiske, Metodiev, JDT, 1712.07124; see also Tkachov, hep-ph/9601308]

The Energy Flow Polynomials

Underlying Physics
→ Natural Data Representation
Suitable Algorithm

Multigraph

Angular Scaling

$$\mathrm{EFP}_G = \sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M} z_{i_1} \cdots z_{i_N} \prod_{(k,\ell)\in G} \theta_{i_k i_\ell}^{\beta}$$

All N-tuples

N Energy Fractions

Polynomial in Pairwise Angles

A *Linear* Basis for Jet Substructure (!)

[Komiske, Metodiev, JDT, 1712.07124]

The Energy Flow Polynomials

Underlying Physics
→ Natural Data Representation
Suitable Algorithm

Multigraph ↓ *Angular Scaling* ↓

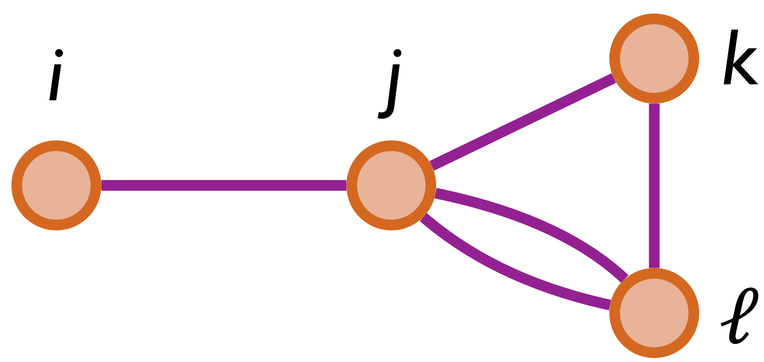
$$\mathrm{EFP}_G = \sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M} z_{i_1} \cdots z_{i_N} \prod_{(k,\ell)\in G} \theta_{i_k i_\ell}^{\beta}$$

All N-tuples

N Energy Fractions

Polynomial in Pairwise Angles

e.g.



$$= \sum_{ijk\ell} z_i z_j z_k z_\ell \, \theta_{ij}\theta_{jk}\theta_{j\ell}^2\theta_{k\ell}$$

A *Linear* Basis for Jet Substructure (!)

[Komiske, Metodiev, JDT, 1712.07124]

Down the Rabbit Hole

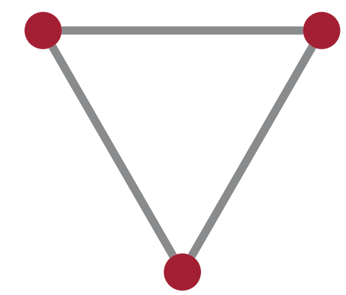
Underlying Physics
→ Natural Data Representation
Suitable Algorithm

Known Structures:

${}_1e_2$

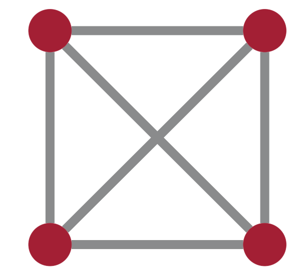
width (β=1), thrust (β=2)

${}_3e_3$



used in D_2

${}_6e_4$



used in C_3

Down the Rabbit Hole

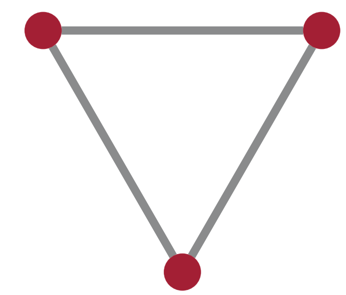
Underlying Physics
→ Natural Data Representation
Suitable Algorithm

Known Structures:

${}_1e_2$

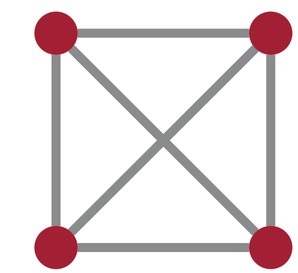
width (β=1), thrust (β=2)

${}_3e_3$



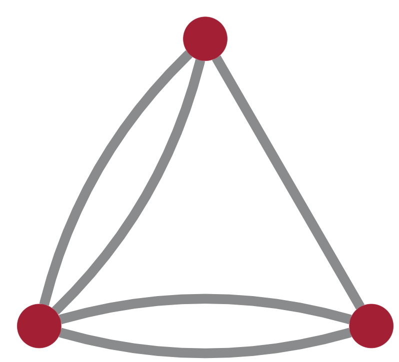
used in D_2

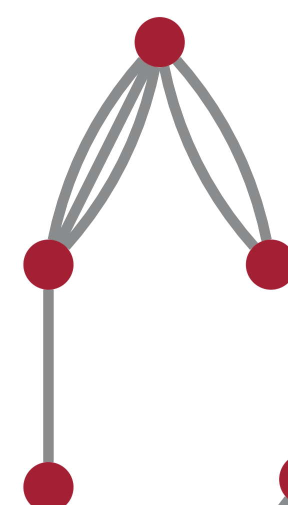
${}_6e_4$

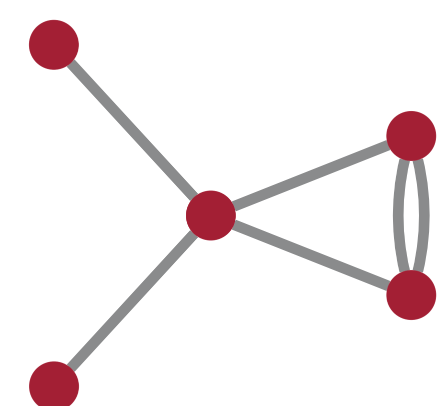


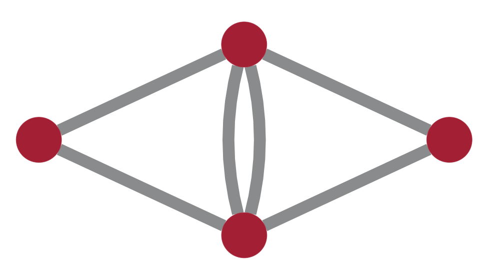
used in C_3

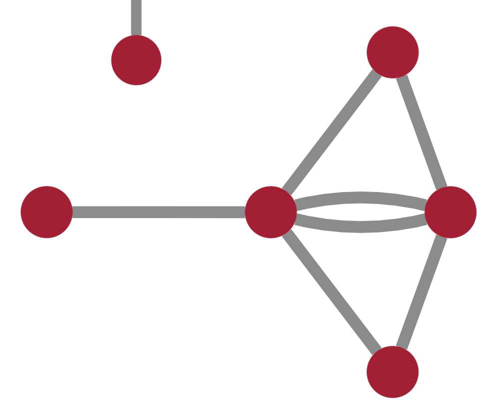
No Idea:

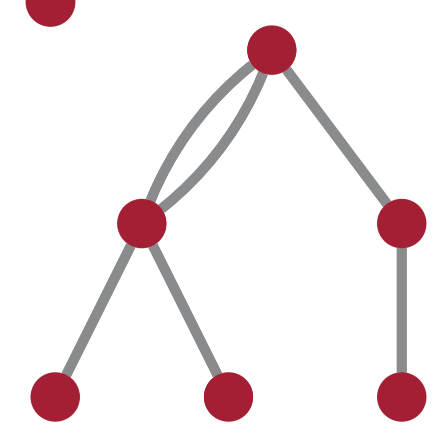












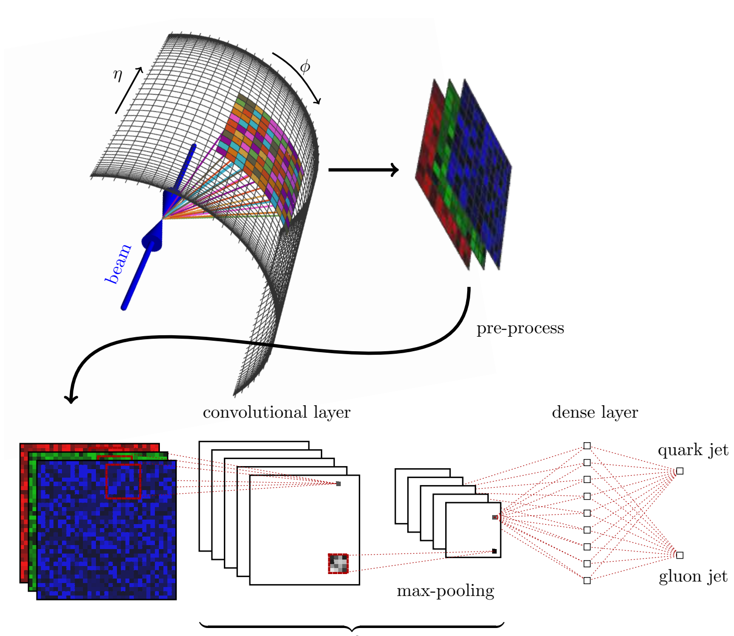
Linear Regression ≈ CNN

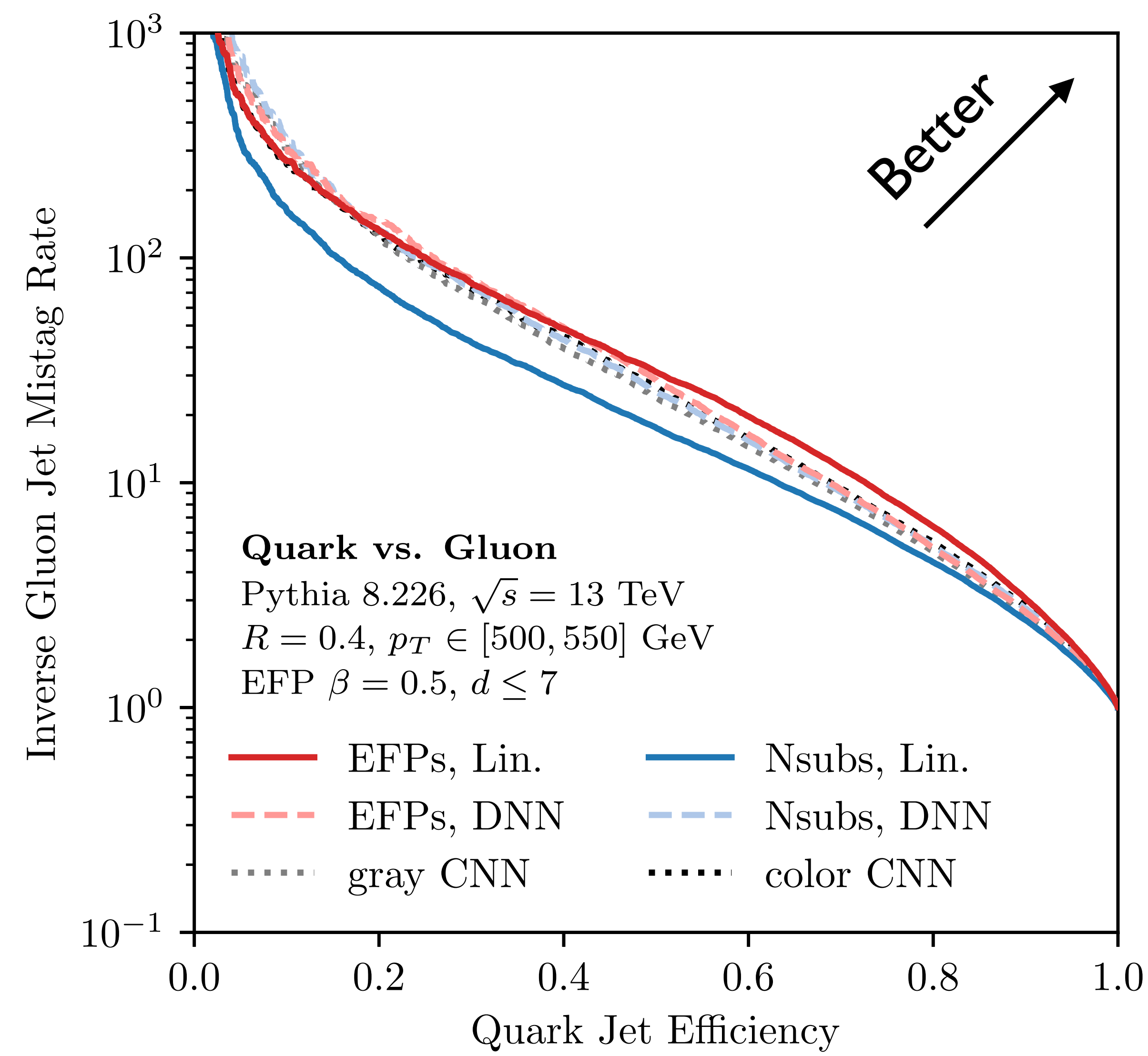
Underlying Physics
Natural Data Representation
→ Suitable Algorithm

"…indistinguishable from magic"

$$\boxed{\mathcal{S} = \sum_G s_G \, \mathrm{EFP}_G}$$

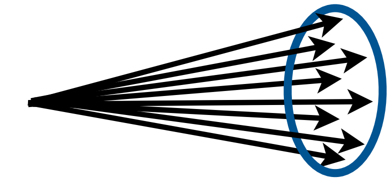
≈





[Komiske, Metodiev, JDT, 1712.07124; Komiske, Metodiev, Schwartz, 1612.01551]

Comparing Data Representations



Original 4-Vectors: $\{p_1^\mu, p_2^\mu, \ldots, p_N^\mu\}$

Variable-length, unordered set

EFP Basis: $\mathrm{EFP}_G = \sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M} z_{i_1} \cdots z_{i_N} \prod_{(k,\ell)\in G} \theta_{i_k i_\ell}$

Automatically permutation invariant
Linear spanning basis (over-complete)
Computational nightmare of O(M^N)?

Too naive, can use variable elimination

The Physics-Meets-Computation Approach

Underlying Physics

$\Downarrow$

Natural & Efficient Data Representation

$\Uparrow$

Desired Computational Property

What is the space of *all* linearly-computable permutation-invariant IRC-safe observables?

The Energy Flow Moments

Underlying Physics
→ Natural Data Representation
Suitable Algorithm

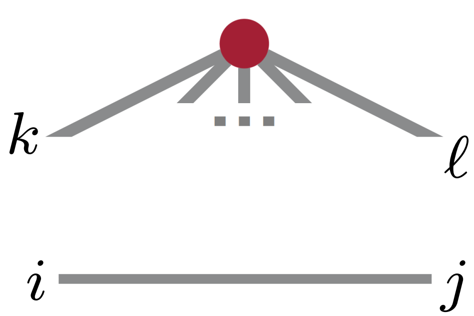
$$\mathcal{I}^{\mu_1 \mu_2 \cdots \mu_v} = \sum_{i=1}^{M} E_i \; \hat{p}^{\mu_1} \; \hat{p}^{\mu_2} \cdots \; \hat{p}^{\mu_v}$$

Particle Relabeling

Infrared Safety

Special Choice of Angle $\theta_{ij} = 2 \, \eta_{\mu\nu} \, \hat{p}_i^{\mu} \hat{p}_j^{\nu}$

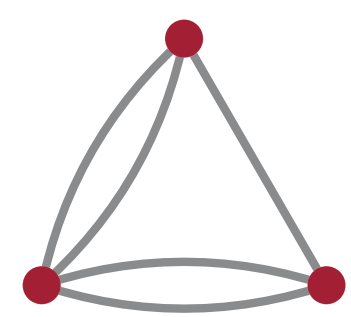
EFMs $\Rightarrow$ EFPs



$\Longleftrightarrow \mathcal{I}^{\mu_k \cdots \mu_\ell}$

$\Longleftrightarrow \eta_{\mu_i \mu_j}$

Compute in O(M)!



$= \mathcal{I}^{\mu\nu\rho\sigma} \mathcal{I}^{\tau}{}_{\mu\nu} \mathcal{I}_{\rho\sigma\tau}$

Also compute in O(M)!

[Komiske, Metodiev, JDT, we've been promising this paper for 9 months]