Collision Course

Particle Physics meets Machine Learning

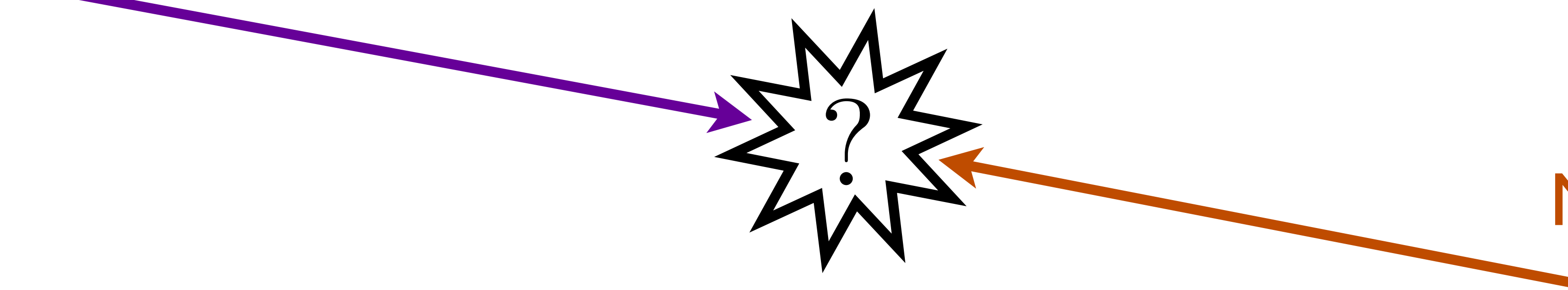
Jesse Thaler

Google X, Tech Talk — April 15, 2019

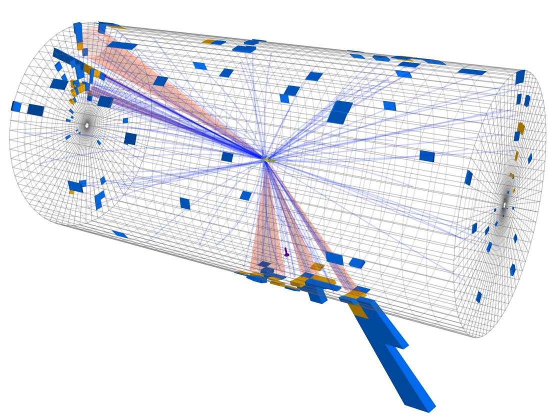
“Collision Course”

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

(Theoretical) Particle Physics



Mathematics, Statistics, Computer Science

Could  *be the next*  *?*

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"

My Perspective c. 2019

BOOSTON 2019

is

~~&~~

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

New insights into fundamental physics
facilitated by advances in
mathematics, statistics, and computer science

(and vice versa)

Proximate Reasons for My Conversion

Patrick Komiske

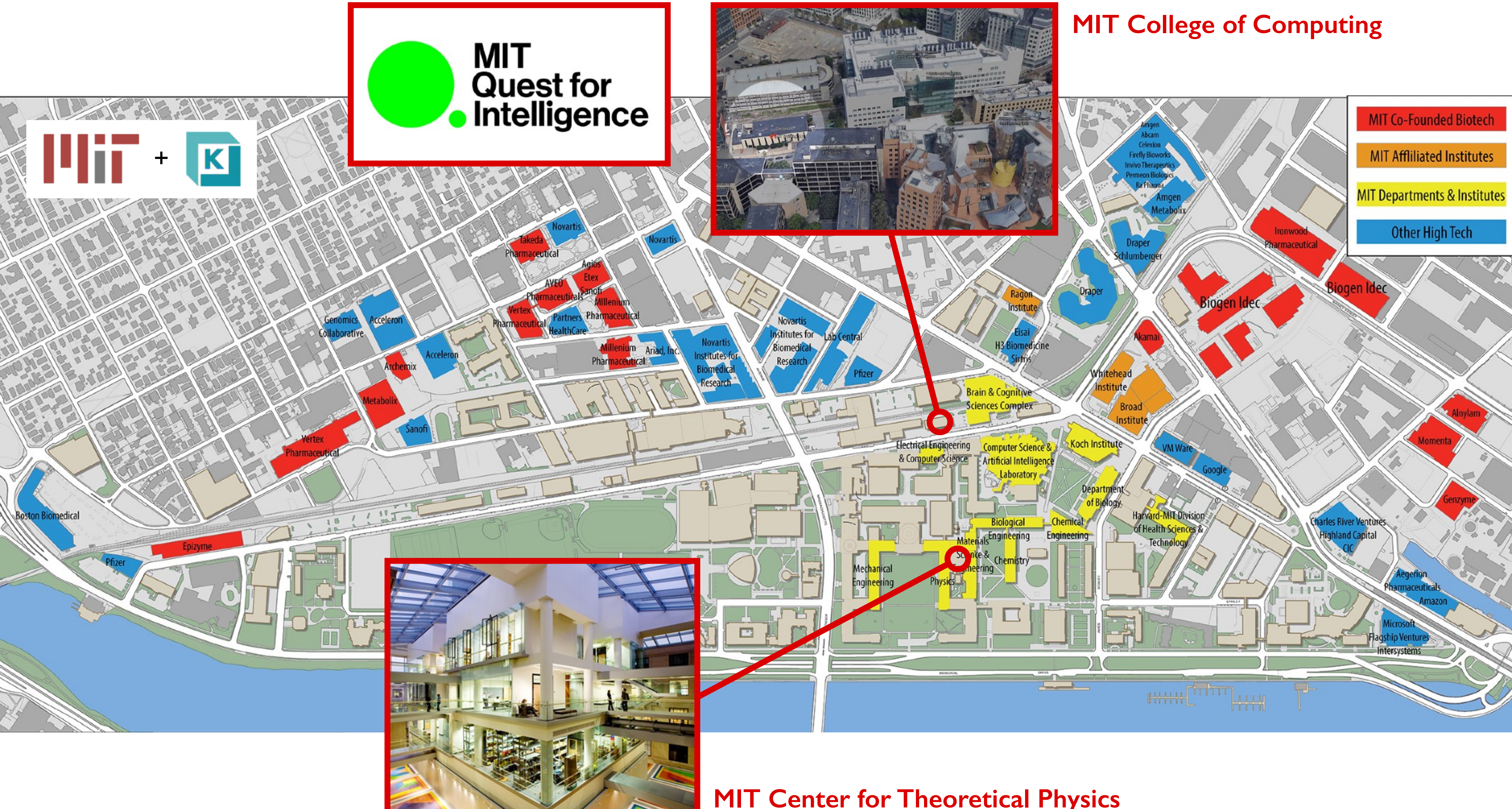
Eric Metodiev

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

Proximate Reasons for Sustained Excitement

MIT Quest for Intelligence

MIT College of Computing



MIT Center for Theoretical Physics

(Even more than this; map from 2014)

Ultimate Goal

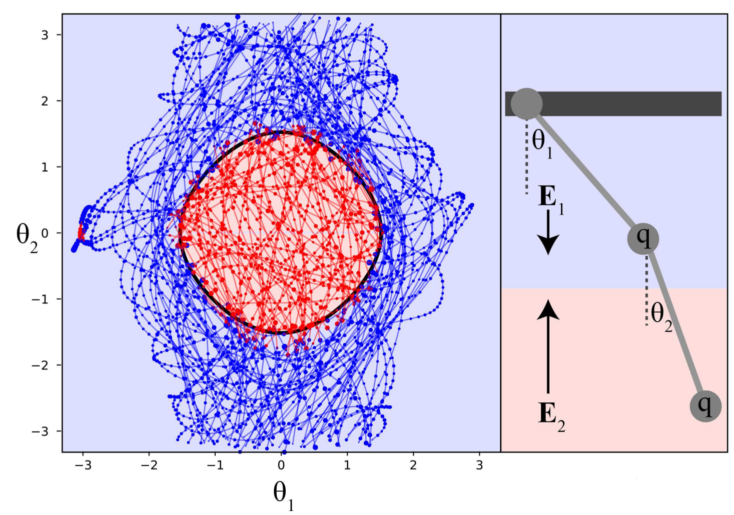
Teach a Machine to "Think like a Physicist"

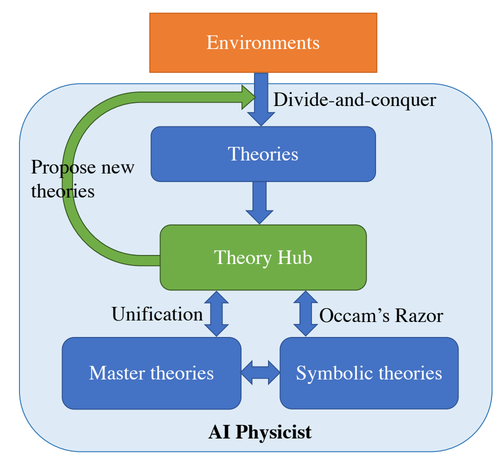
(really, any scientist)

As opposed to "Learn like a Toddler"
Have you ever tried to reason with a 2 year old?

AI-saac Newton?

[e.g. Wu, Tegmark, 1810.10525]





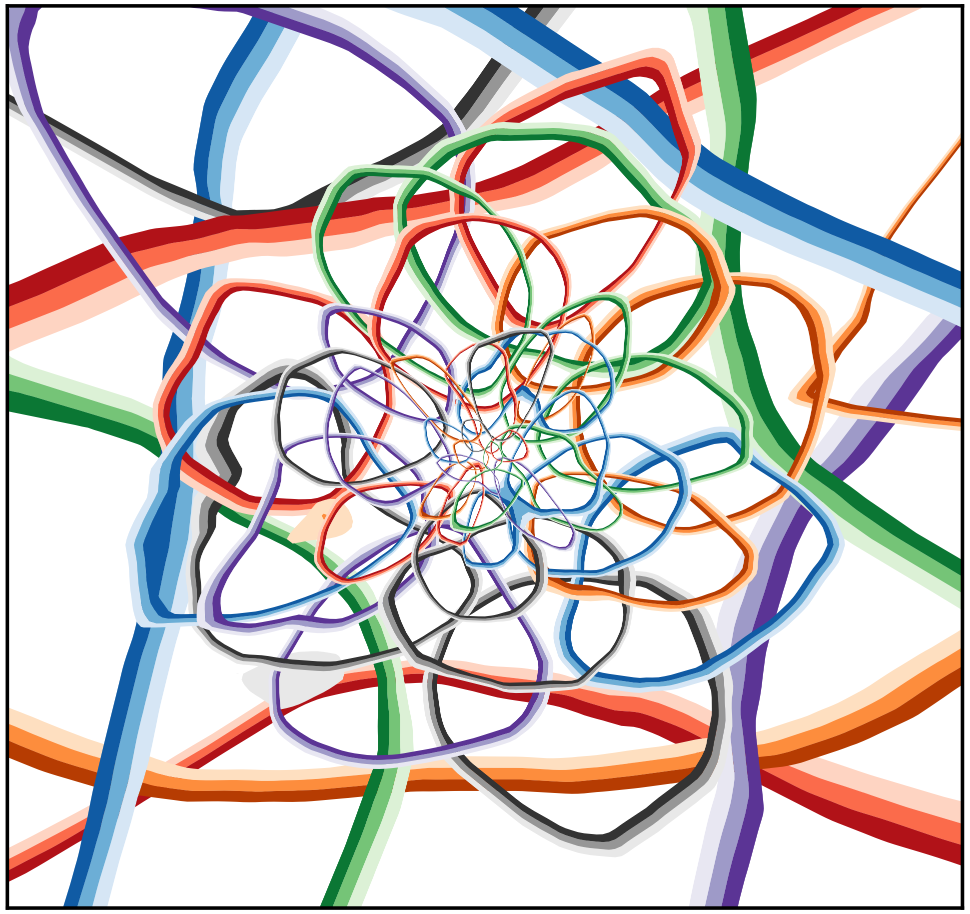
Physics 8.01

1. Specify the problem ← True creative intelligence
2. Find the solution ← ML does amazingly well
3. Verify the solution ← ML can get in the way

Today's Goal

Teach a Machine to Communicate to a Physicist

As opposed to trying to reverse engineer a
black box machine learning architecture



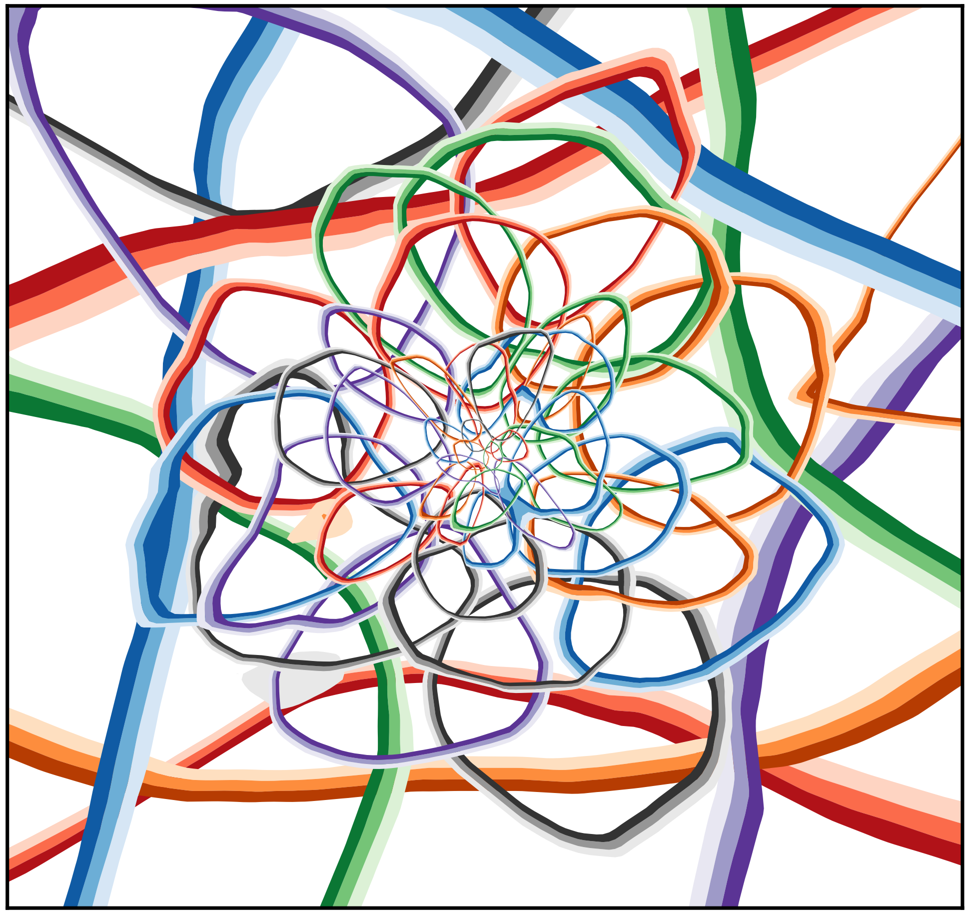
Learning the Altarelli-Parisi splitting kernel

[Komiske, Metodiev, JDT, 1810.05165]

Today's Goal

Teach a Machine to Communicate to a Physicist

As opposed to trying to reverse engineer a
black box machine learning architecture

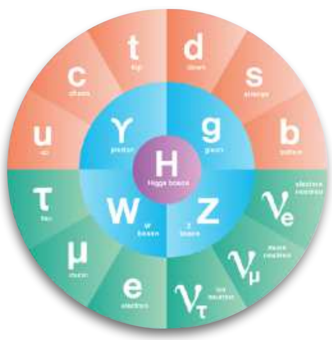


Learning the
Altarelli-Parisi
splitting kernel

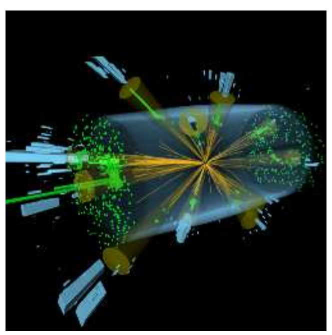
(without us asking it to!)

[Komiske, Metodiev, JDT, 1810.05165]

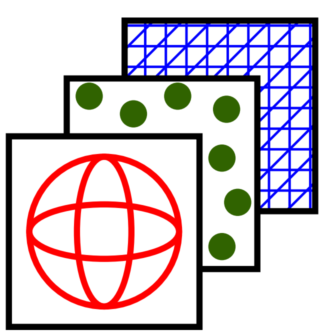
Outline



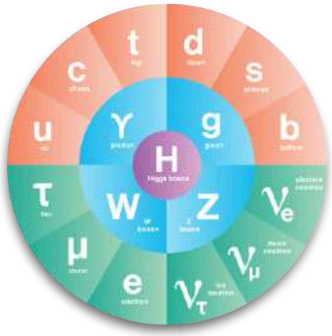
Particle Physics Primer



Jets and Point Clouds



Energy Flow Networks

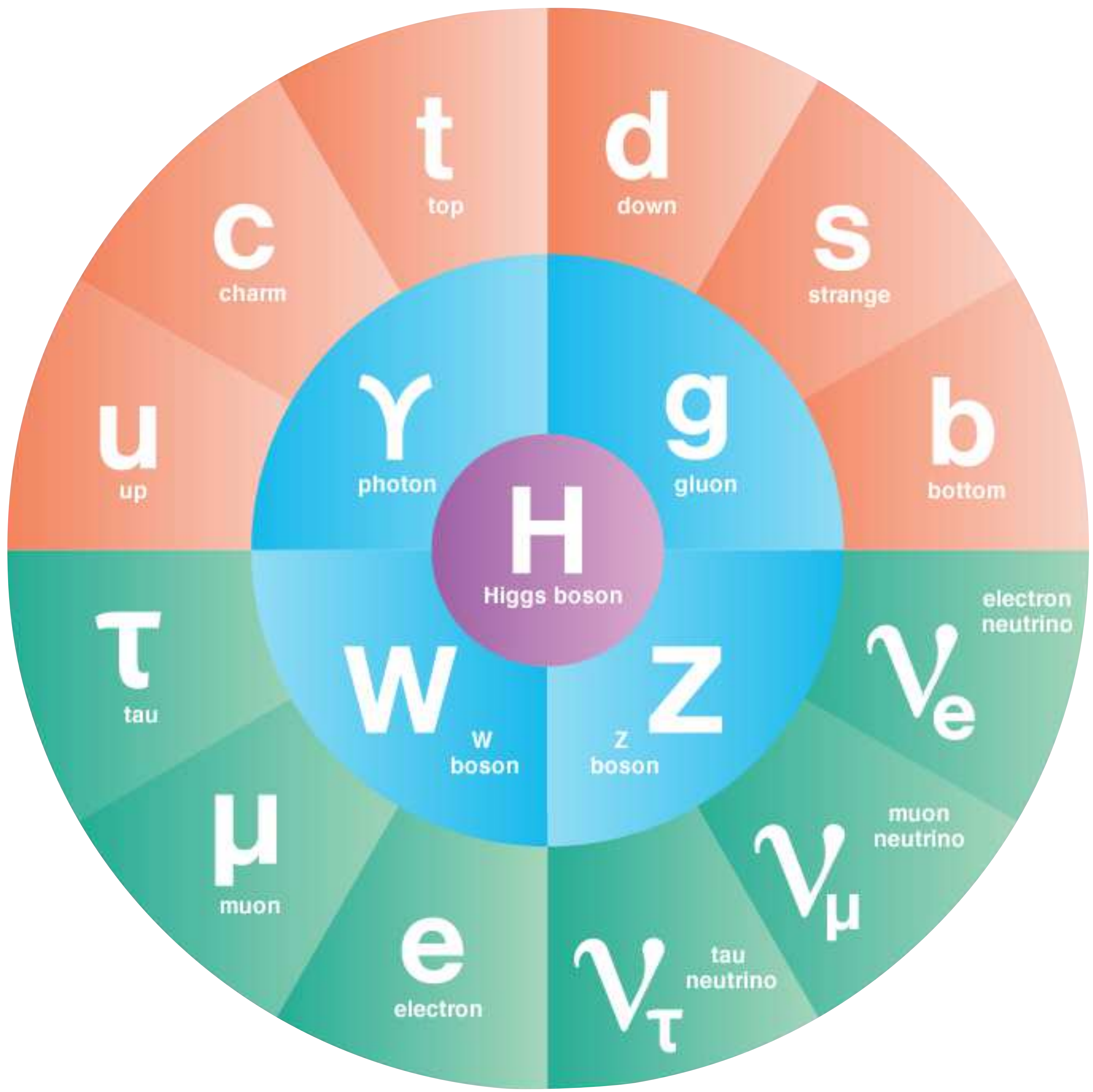


Particle Physics Primer

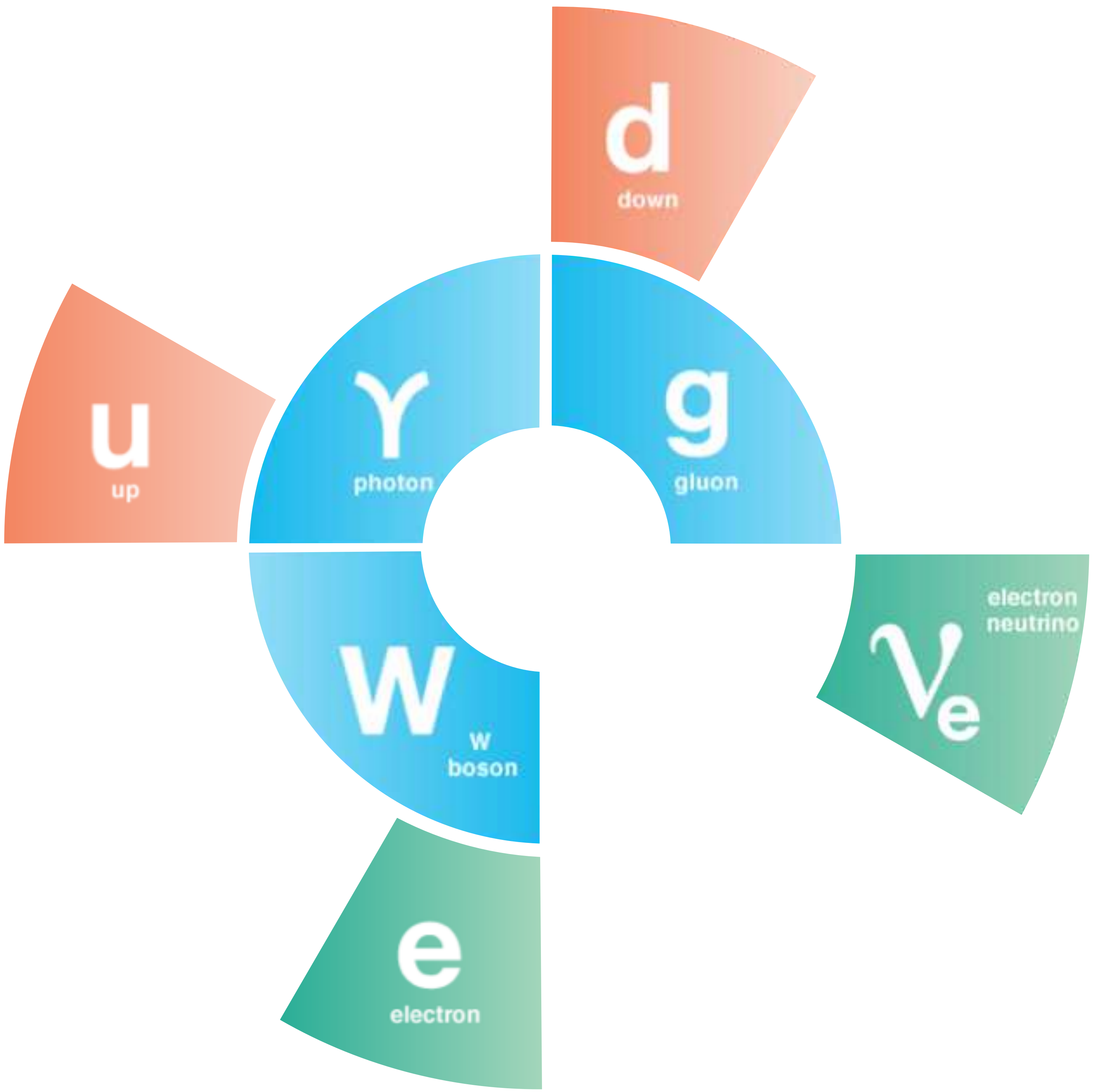
Jets and Point Clouds

Energy Flow Networks

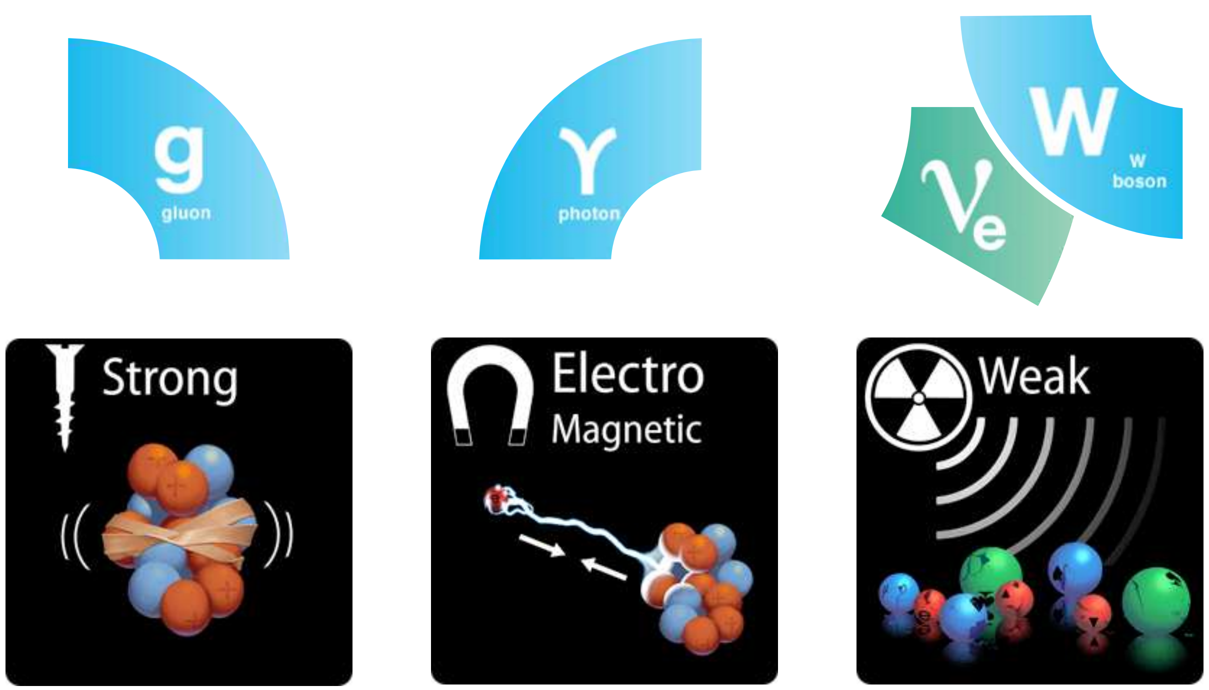
The Standard Model



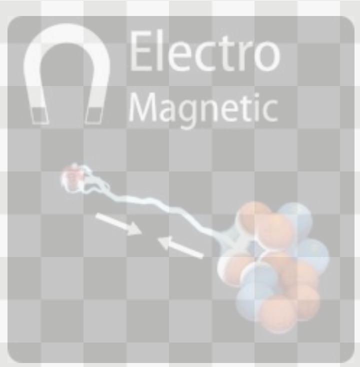
Ordinary Matter

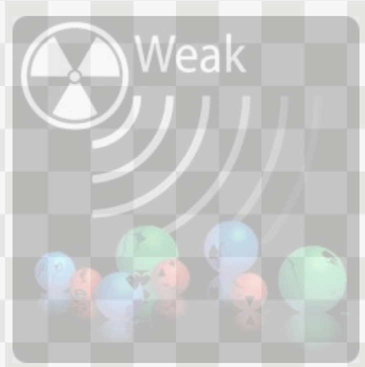


Ordinary Matter



Dark Matter (& Dark Energy)

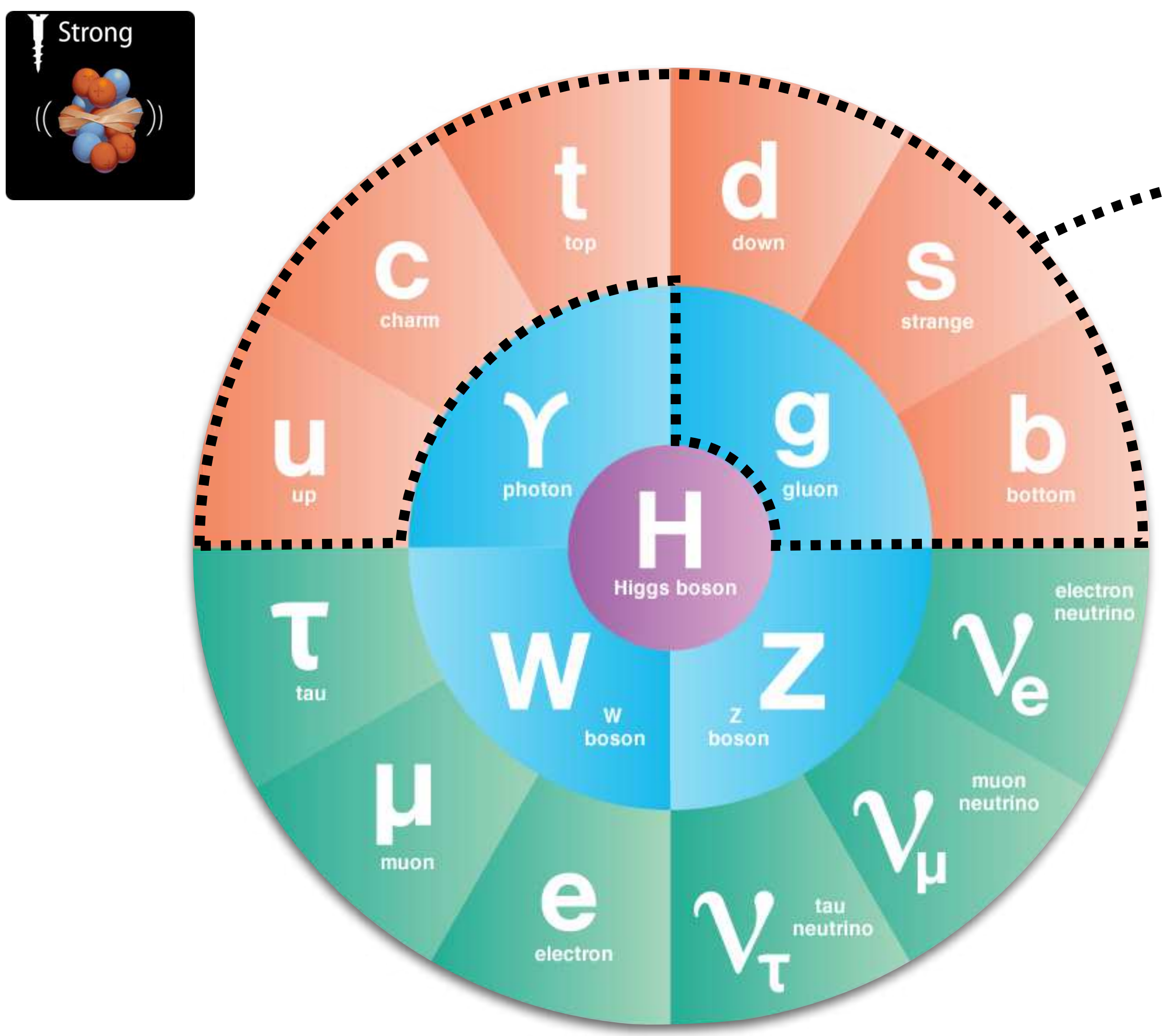


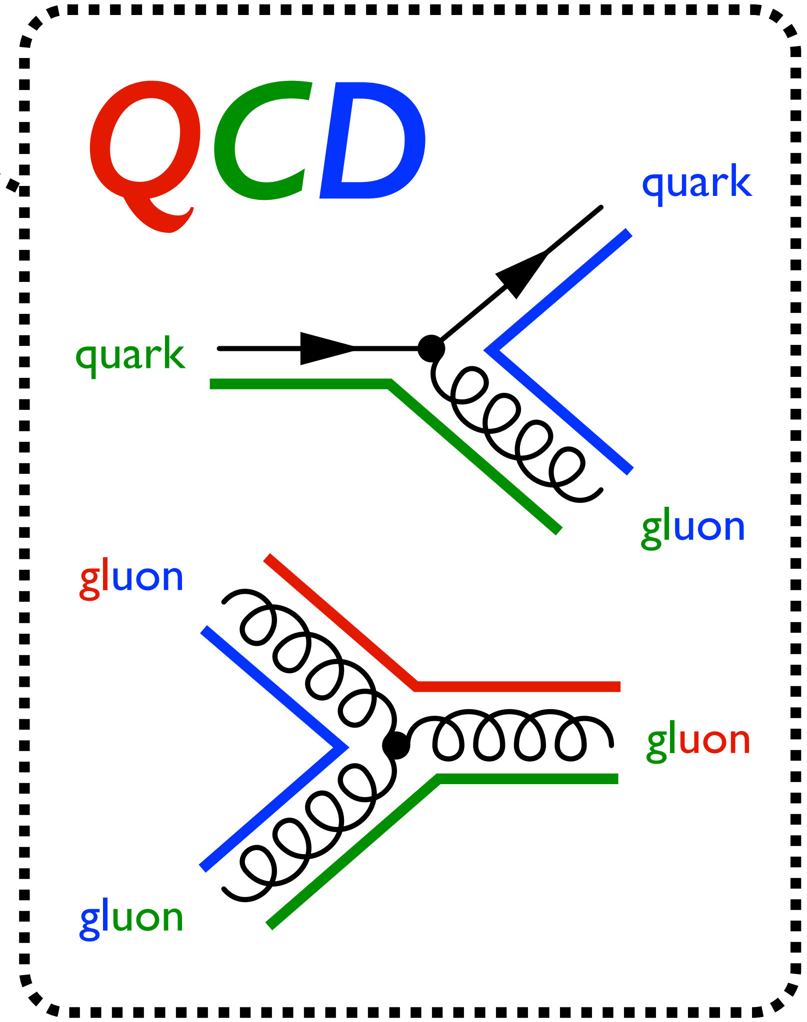


Dark
Forces?

Focus on the Strong Force

Crucial for physics at the Large Hadron Collider





Focus on the Strong Force

Crucial for physics at the Large Hadron Collider

Mesons

$\pi^\pm\,\pi^0\,\eta\,K^\pm\,K^0\,\eta'\,D^\pm\,D^0\,D_s^\pm\,\eta_c\,B^\pm\,B^0\,B_s^0\,\eta_b\,\ldots$

$\rho^\pm\,\rho^0\,\omega\,K^{*\pm}\,K^{*0}\,\phi\,D^{*\pm}\,D^{*0}\,D_s^{*\pm}\,J/\psi\,B^{*\pm}\,B^{*0}\,B_s^{*0}\,\Upsilon\,\ldots$

Baryons

$p\,n\,\Lambda^0\,\Sigma^+\,\Sigma^0\,\Sigma^-\,\Xi^0\,\Xi^-\;\ldots$

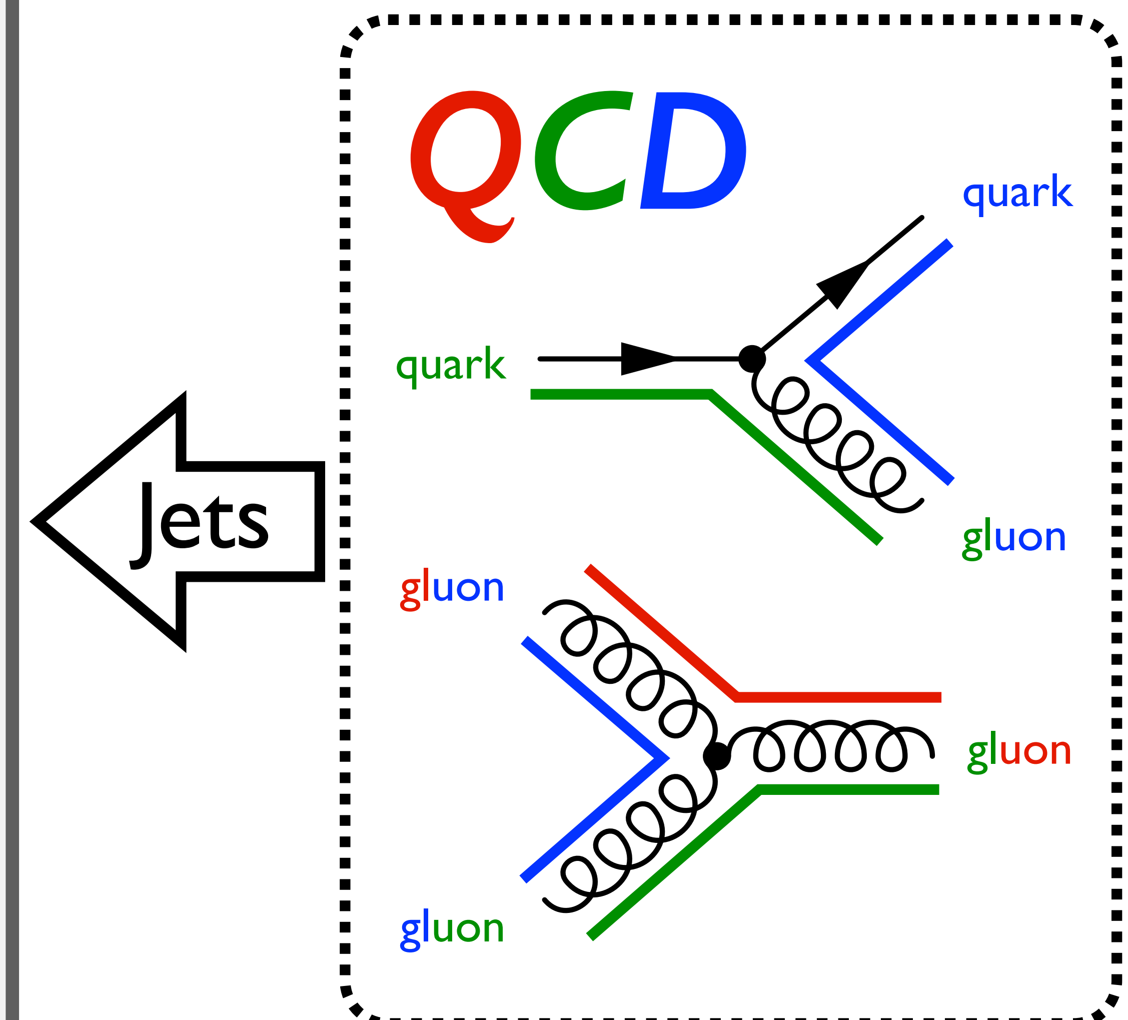
$\Delta^{++}\,\Delta^+\,\Delta^0\,\Delta^-\,\Sigma^{*+}\,\Sigma^{*0}\,\Sigma^{*-}\,\Xi^{*0}\,\Xi^{*-}\,\Omega^-\;\ldots$

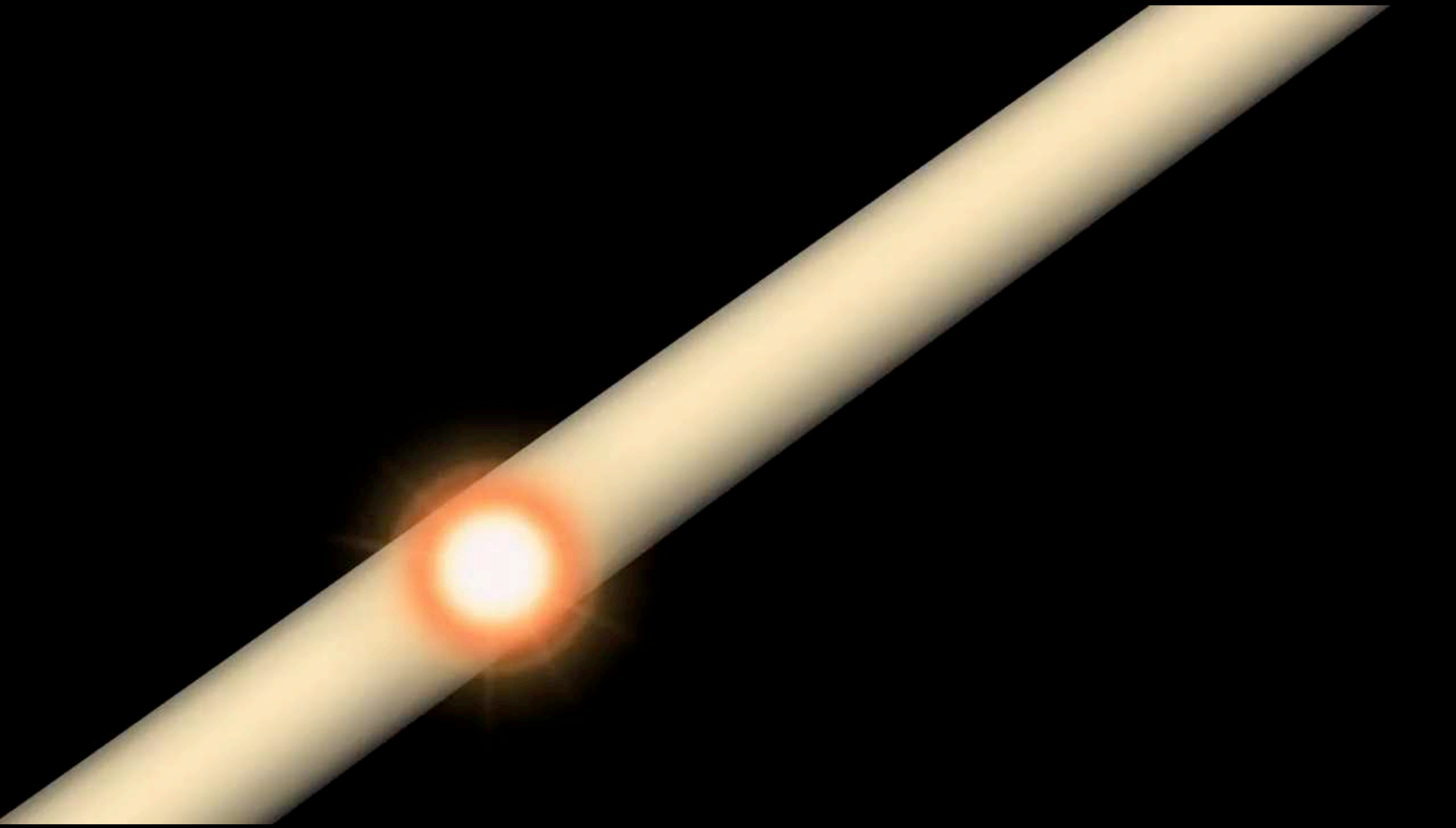
Tetraquarks

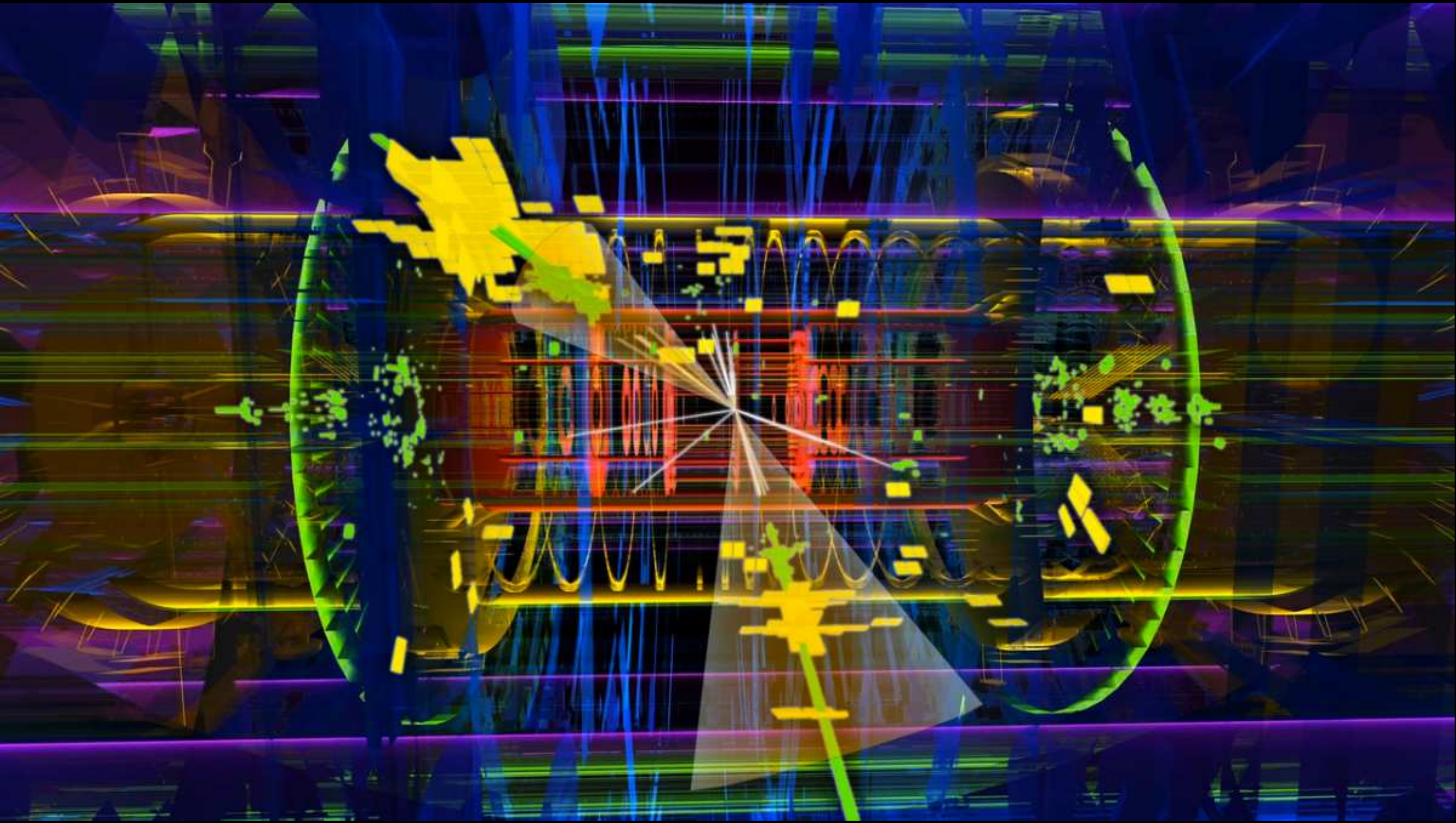
$X(3872)\,Y(4260)\,Z(4430)\;\ldots$

Pentaquarks

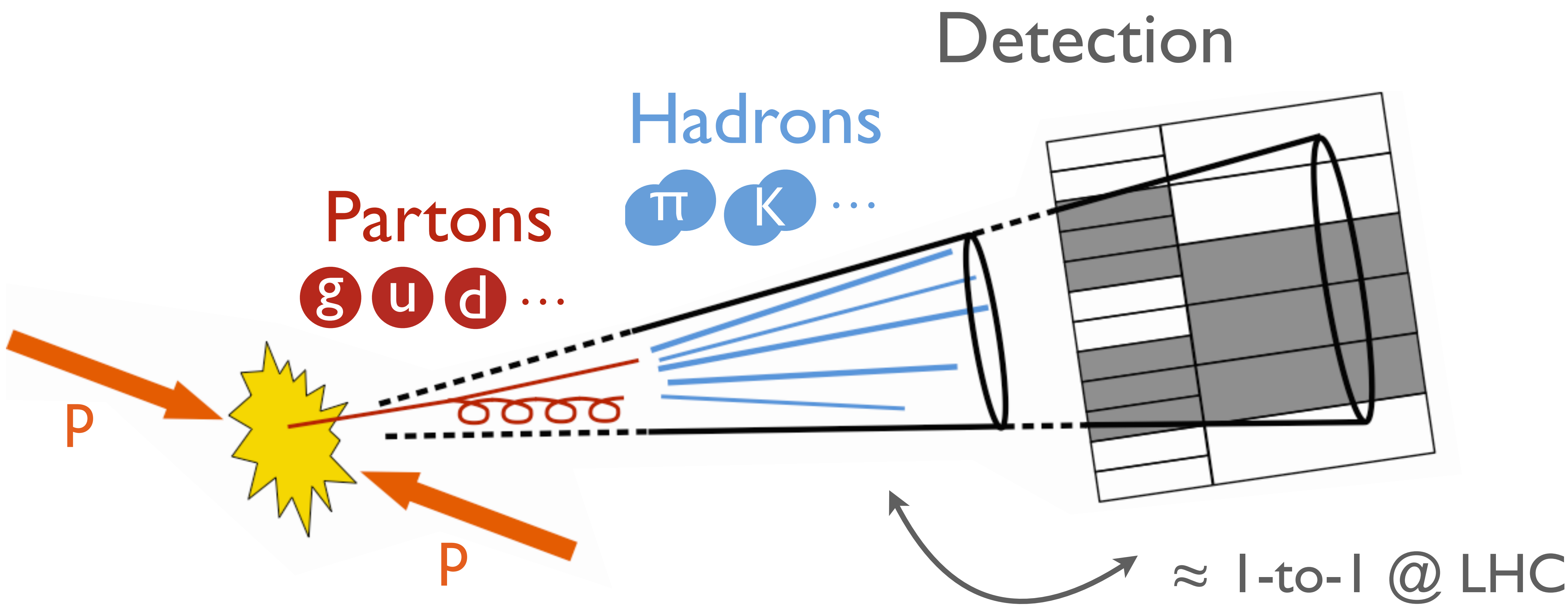
$P_c^+(4450)\;\ldots$



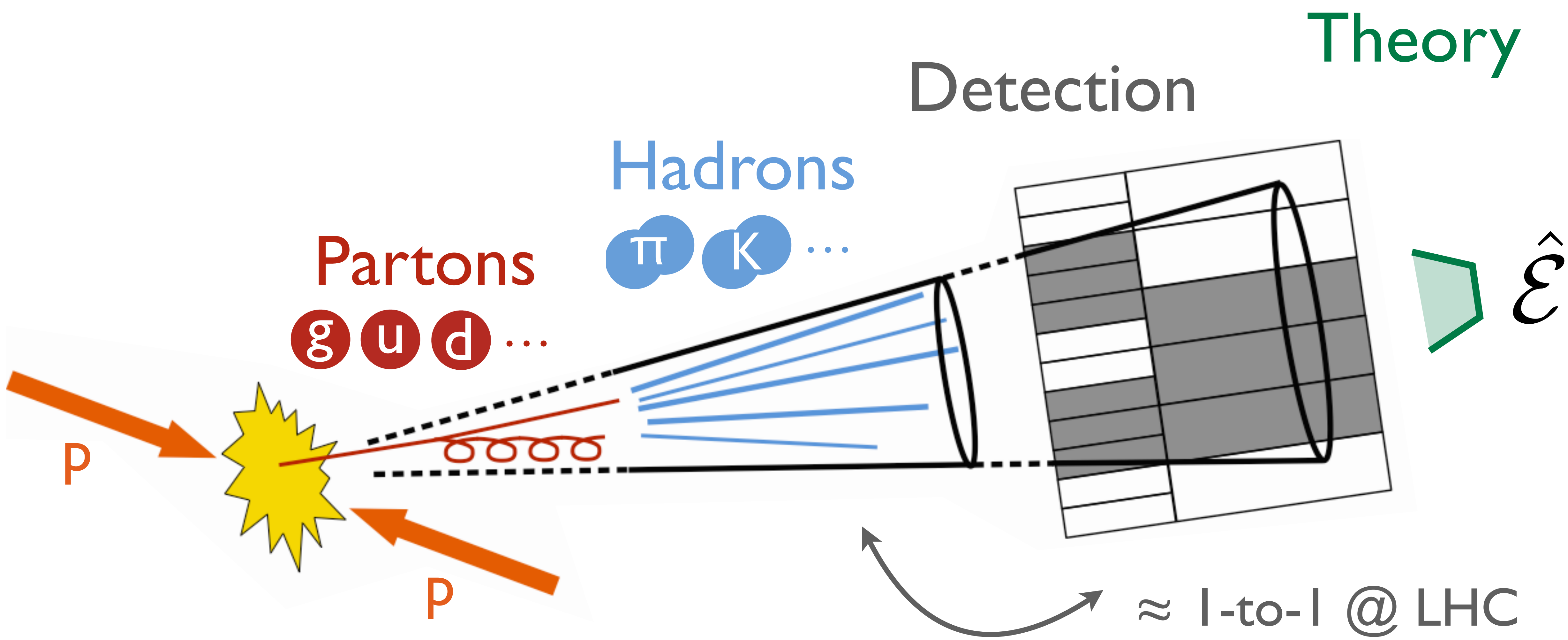




Jet Formation



Jet Formation



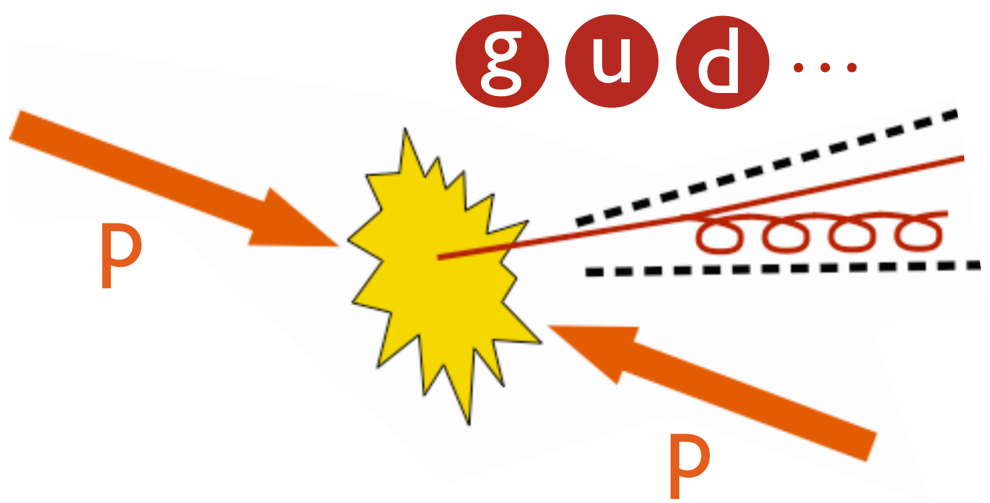
Physics described by stress-energy flow

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

[Sveshnikov, Tkachov, hep-ph/9512370; Hofman, Maldacena, 0803.1467; Mateu, Stewart, JDT, 1209.3781]

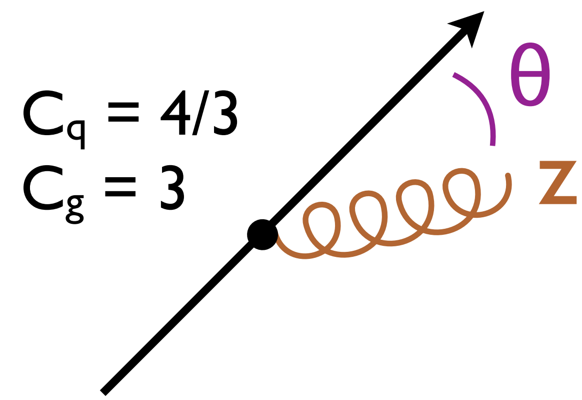
Jet Formation

Partons

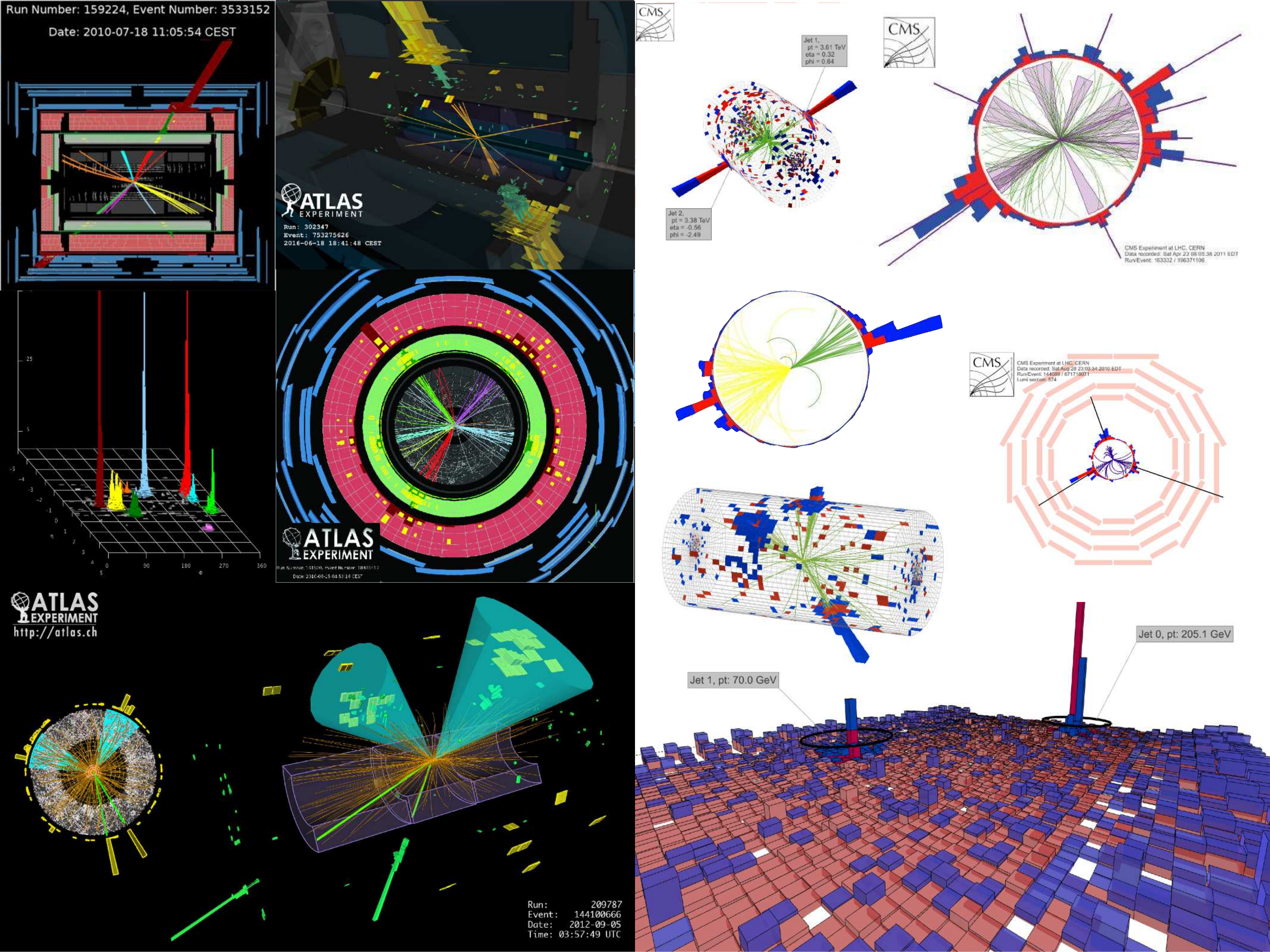


Altarelli-Parisi Splitting

Core prediction of QCD

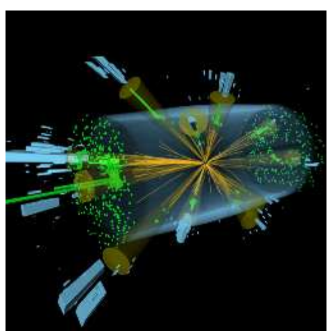


$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \underbrace{\frac{\mathrm{d}\theta}{\theta}}_{\text{Collinear}} \underbrace{\frac{\mathrm{d}z}{z}}_{\text{Soft}}$$



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
CMS Experiment at LHC, CERN
Data recorded: Sat Apr 23 08:05:38 2011 EDT
Run/Event: 163332 / 196371106
ATLAS
EXPERIMENT
CMS
CMS Experiment at LHC, CERN
Data recorded: Sat Aug 28 23:03:34 2010 EDT
Run/Event: 144089 / 671718071
Lumi section: 574
ATLAS
EXPERIMENT
http://atlas.ch
Jet 1, pt: 70.0 GeV
Jet 0, pt: 205.1 GeV
Run: 209787
Event: 144100666
Date: 2012-09-05
Time: 03:57:49 UTC

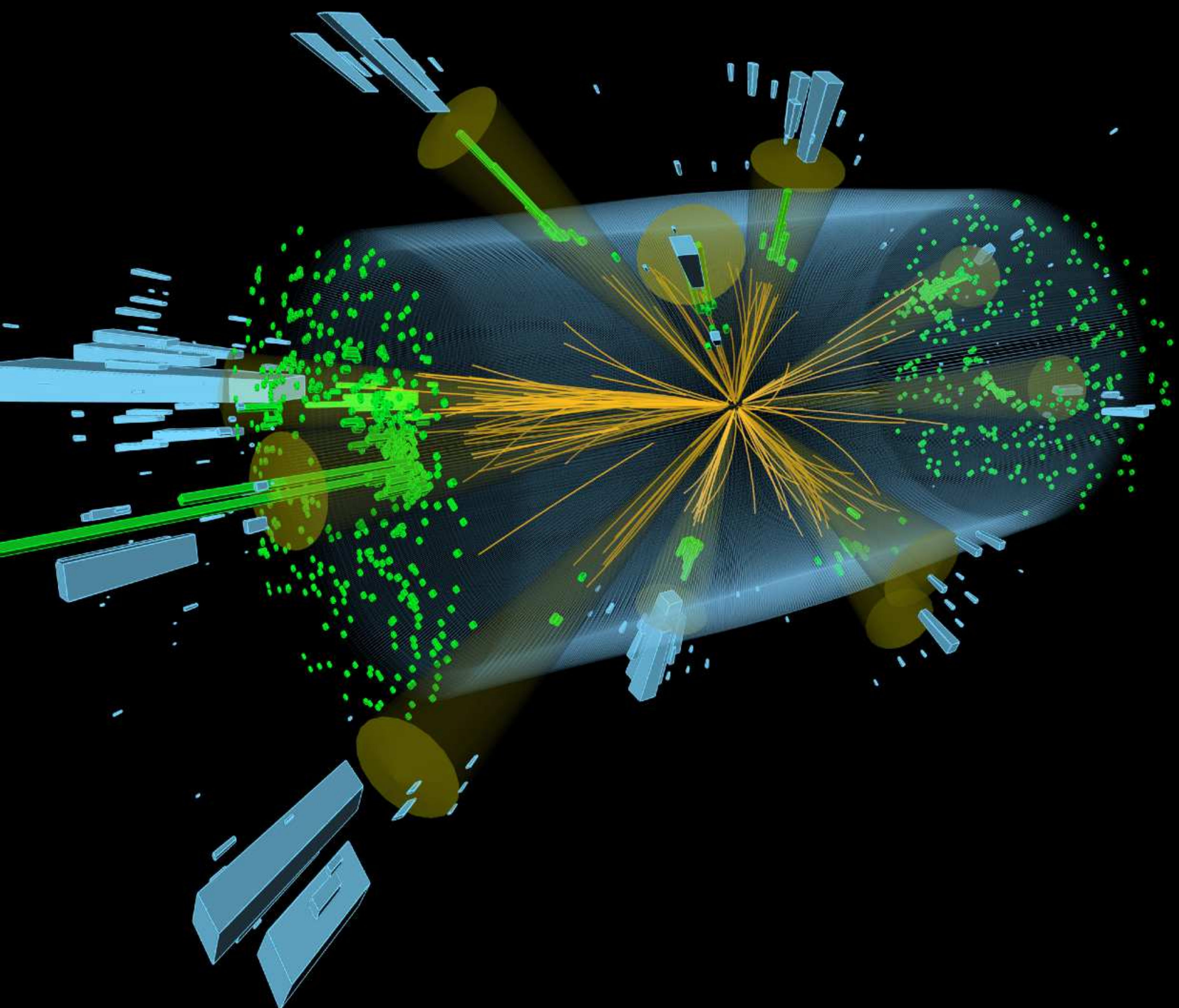
Particle Physics Primer



Jets and Point Clouds

Energy Flow Networks

What is a Collision Event?



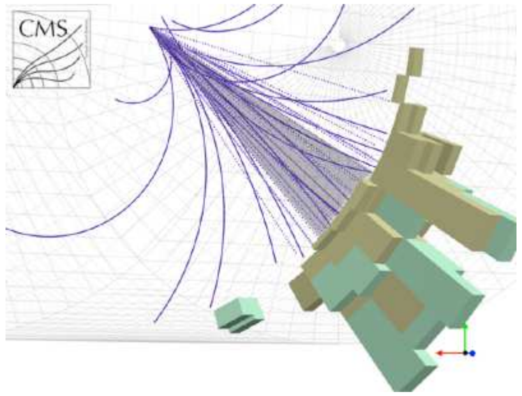
T	E	H	M		
	●			γ	photon
●	●			$e^{\pm}$	electron
●	●	●	●	$\mu^{\pm}$	muon
●	●	●		$\pi^{\pm}$	pion
●	●	●		$K^{\pm}$	kaon
	●	●		K_L^0	K-long
●	●	●		$p/\overline{p}$	proton
	●	●		$n/\overline{n}$	neutron

elementary

composite

[Popular Science, 2013]

Key: Jets are Point Clouds



- Particle: List of properties

$$\vec{p} = \{\underbrace{E}_{\text{Energy}}, \underbrace{p_x, p_y, p_z}_{\text{Momentum}}, \underbrace{\ldots}_{\text{Mass, charge, flavor, vertex, quality, ...}}\}$$

- Jet: Set of particles

$$\mathcal{J} = \{\underbrace{\vec{p}_1, \vec{p}_2, \vec{p}_3, \ldots, \vec{p}_N}_{\text{Permutation Symmetry}}\}$$

Variable Length

Permutation Symmetry

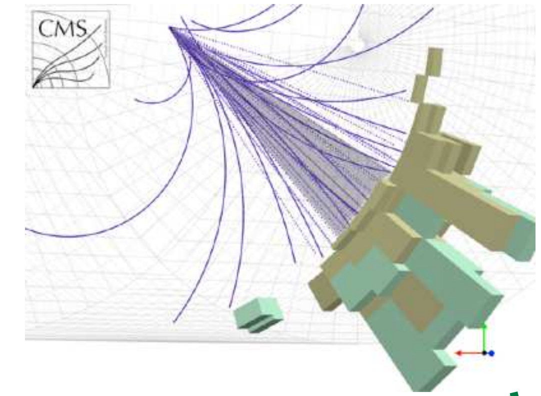
Quantum Mechanics:
"When you've seen one electron, you've seen them all"

Variable Length

Jet Formation:
Typically 10–50 particles per jet

- Dataset: Set of jets

Optionally: Weighted Point Clouds



stress-energy flow

$\hat{\mathcal{E}}$

- Energy-weighted directions

$$\vec{p} = \{\underset{\substack{\uparrow \\ \text{Energy}}}{E}, \underbrace{n_x, n_y, n_z}_{\text{Direction}}\}$$

- Energy Flow: Set of energy-weighted directions

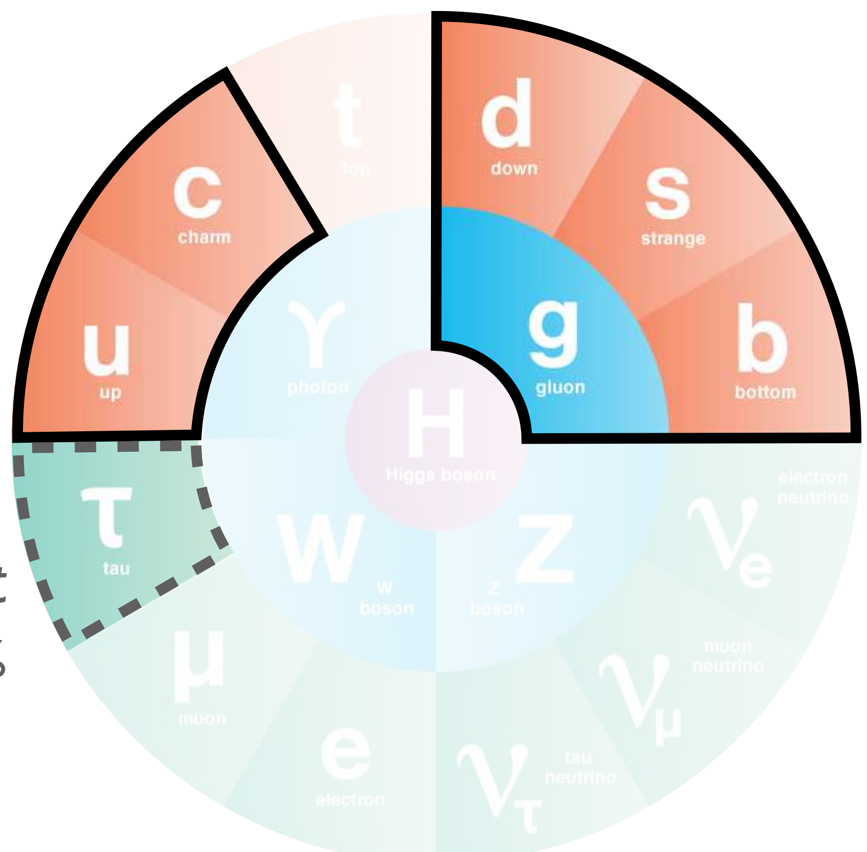
$$\mathcal{J} = \{\vec{p}_1, \vec{p}_2, \vec{p}_3, \ldots, \vec{p}_N\}$$

- "Safe" to Altarelli-Parisi splitting

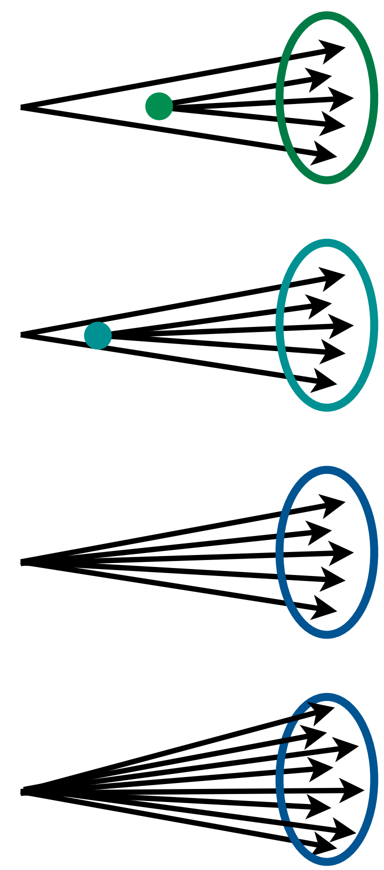
$$\mathrm{d}P_{i \to ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{\mathrm{d}\theta}{\theta} \frac{\mathrm{d}z}{z}$$

Jet Classification

One of many important LHC tasks



micro-jet
≈ 65%



b — 4.2 GeV++

c — 1.3 GeV++

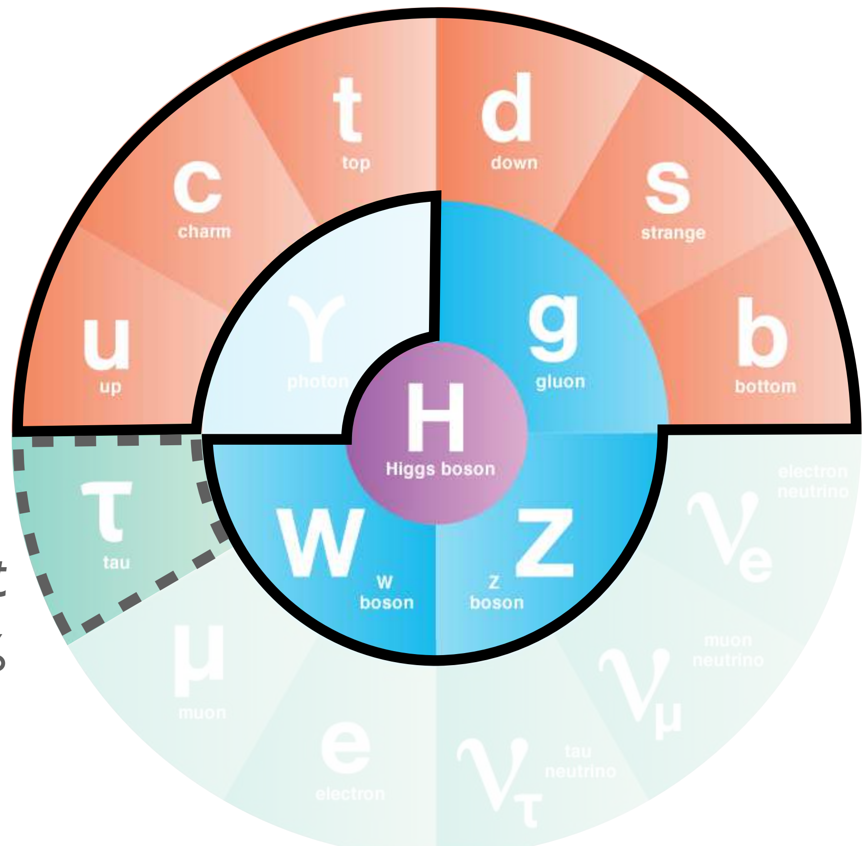
u,d,s — 100 MeV++

g — 0++

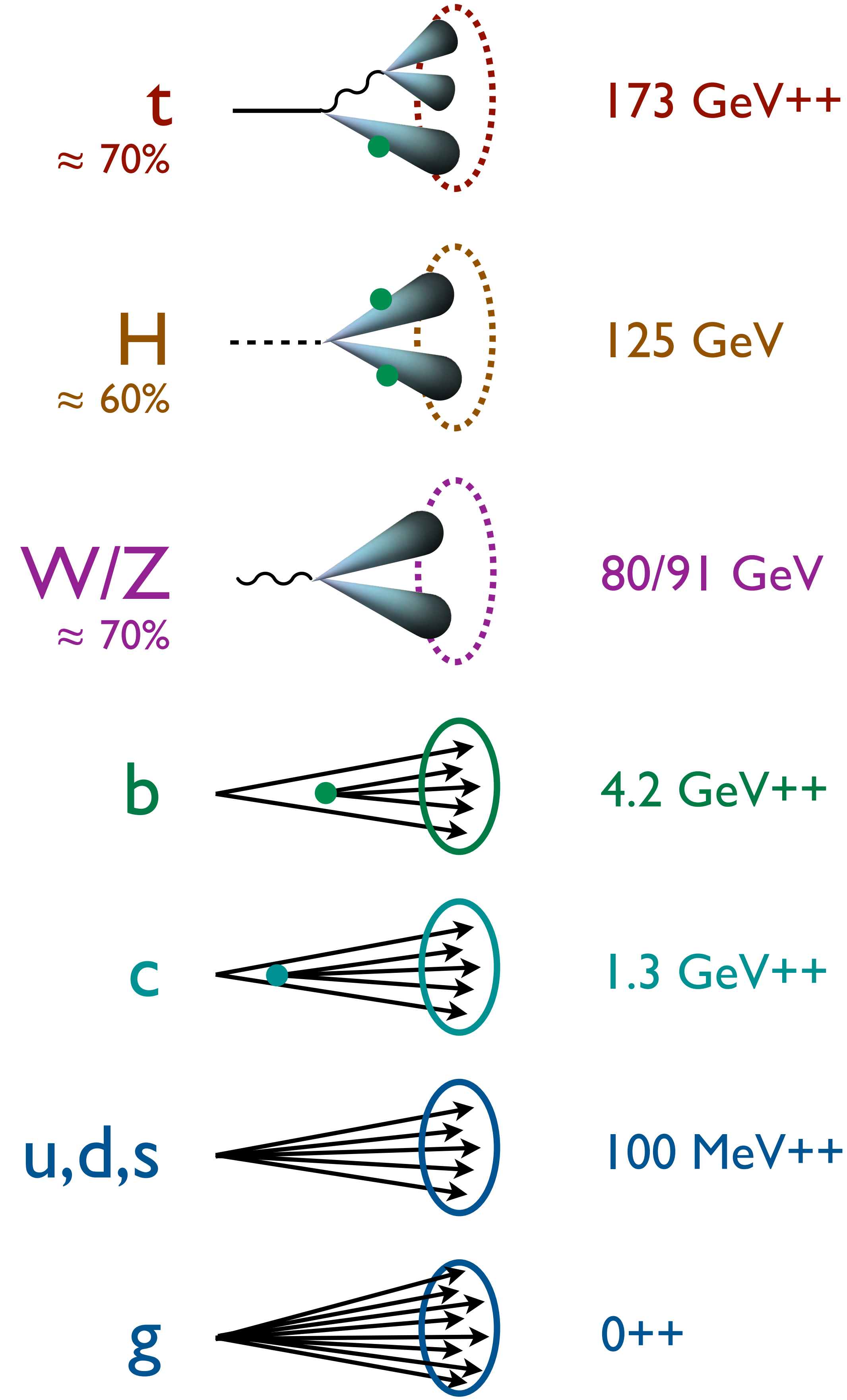
++ = Mass from QCD Radiation

Jet Classification

One of many important LHC tasks



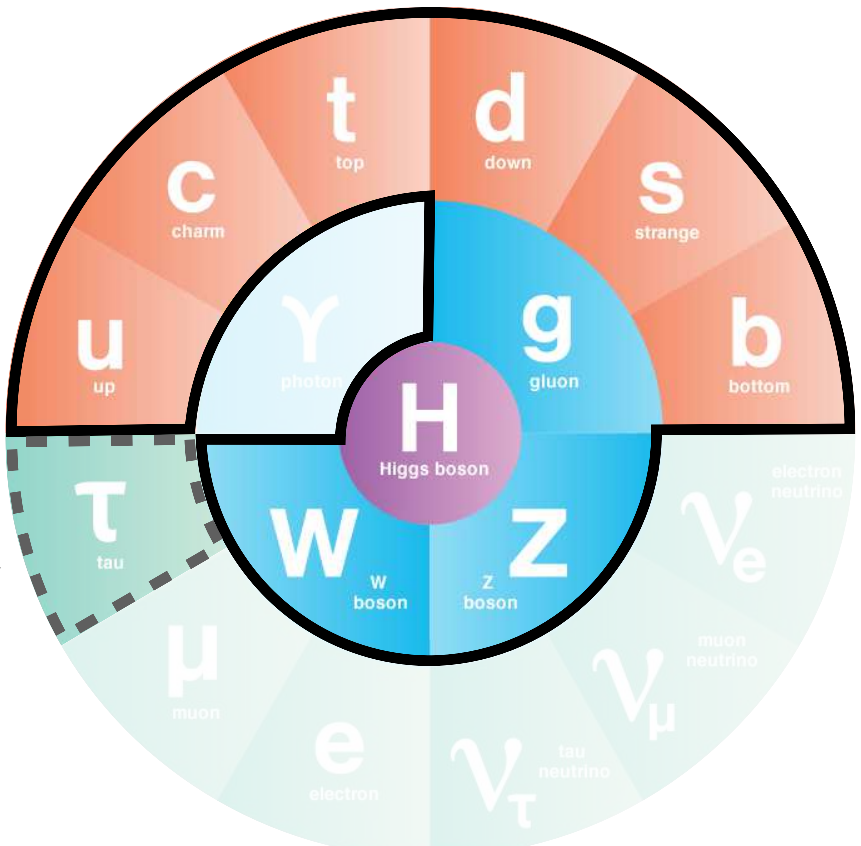
micro-jet
≈ 65%



++ = Mass from QCD Radiation

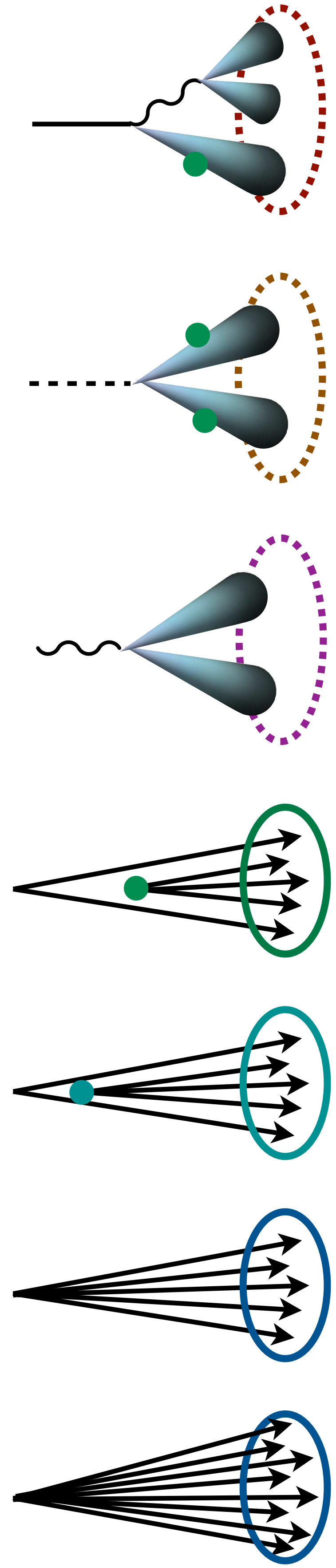
Jet Classification

One of many important LHC tasks



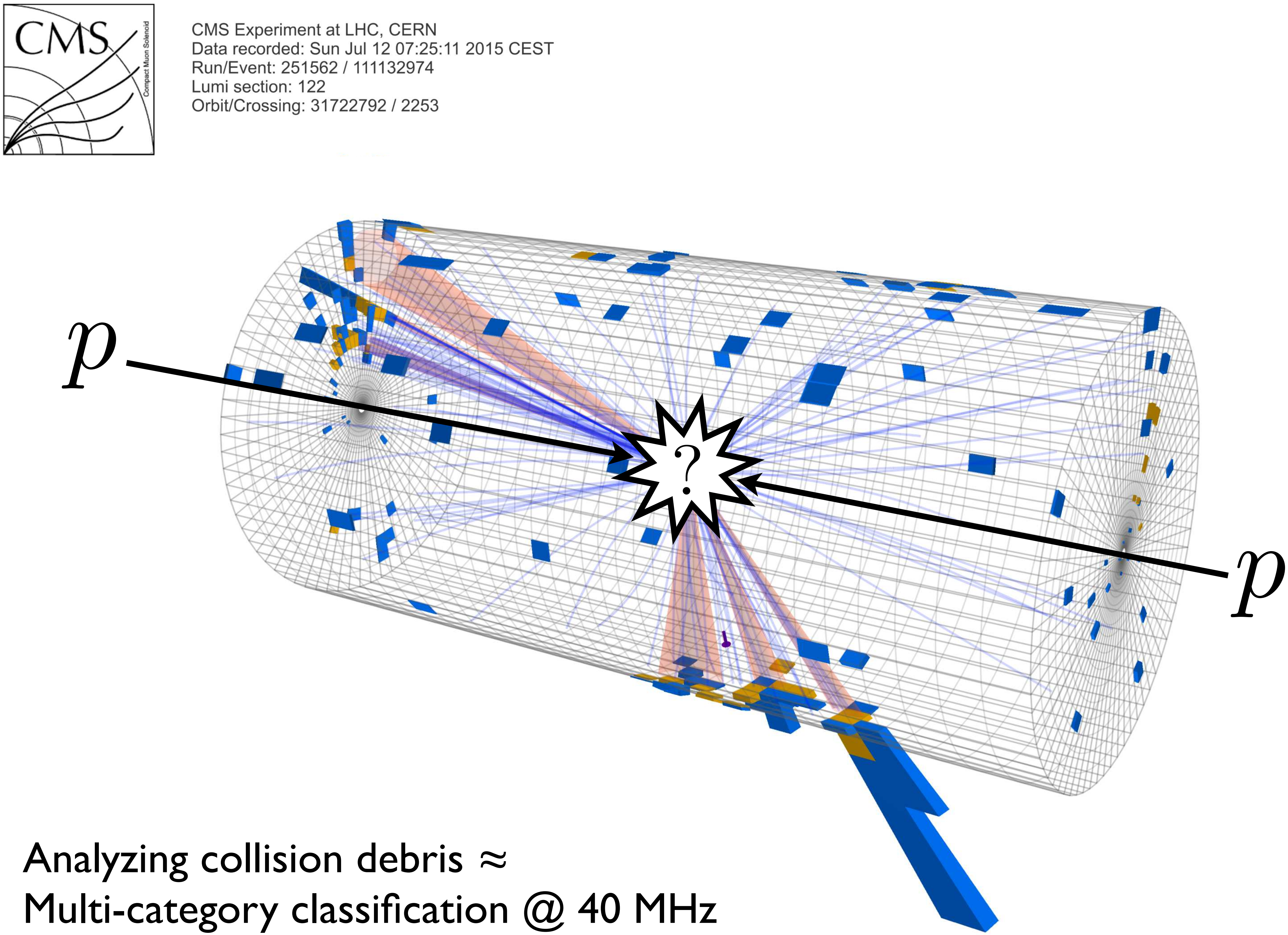
micro-jet
≈ 65%

++ = Mass from QCD Radiation

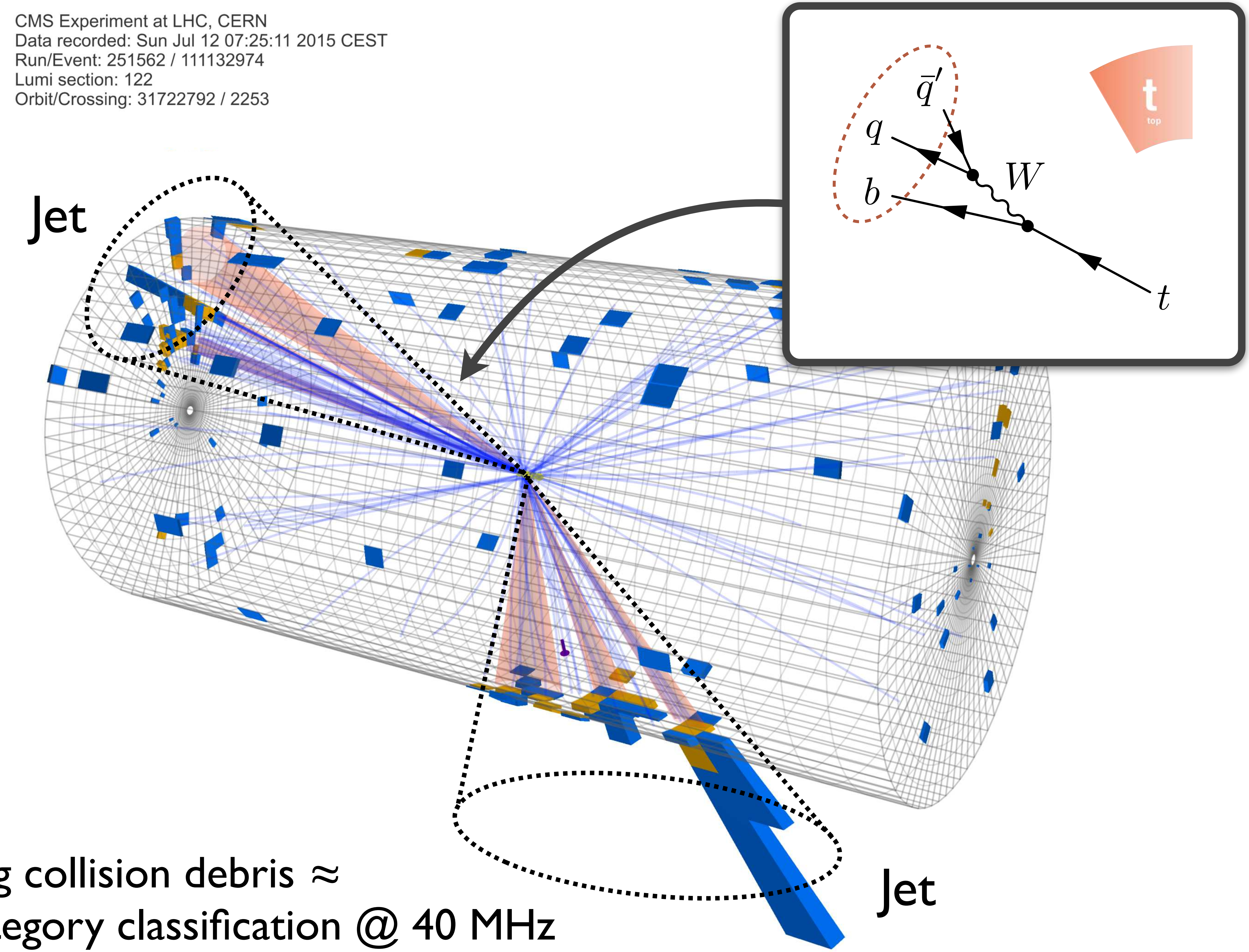


t	≈ 70%	173 GeV++
H	≈ 60%	125 GeV
W/Z	≈ 70%	
b		4.2 GeV++
c		1.3 GeV++
u,d,s		100 MeV++
g		0++

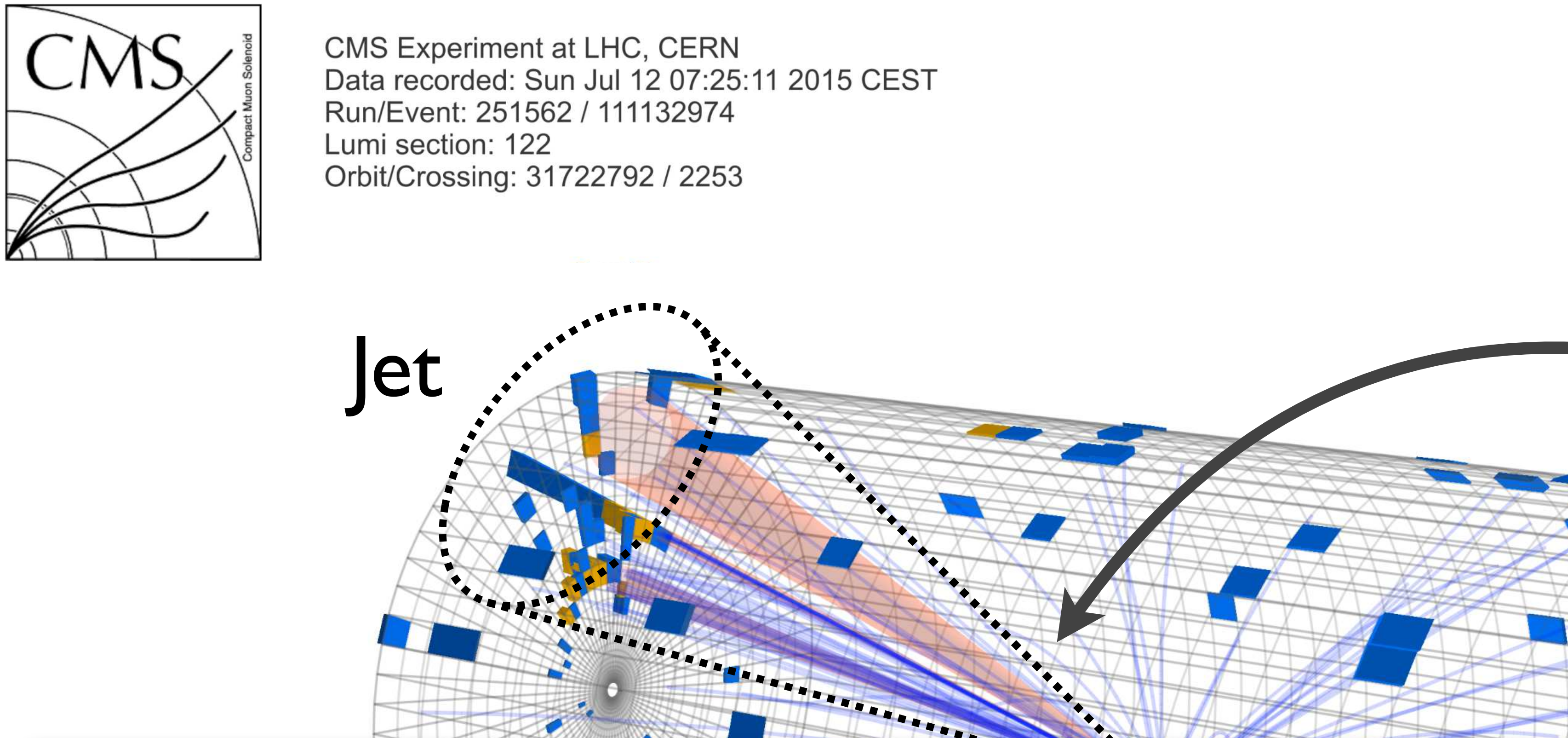
BOSTON 2019

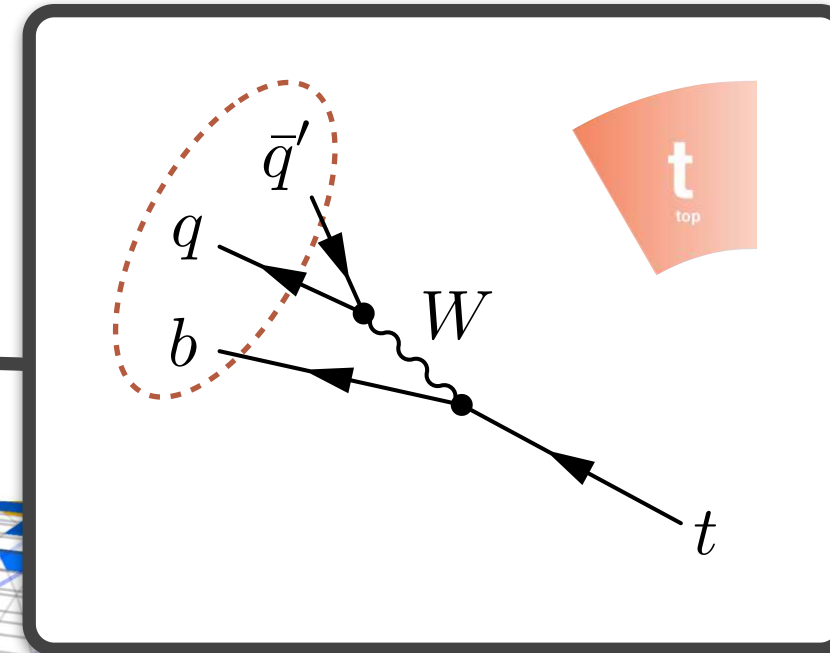


Analyzing collision debris ≈
Multi-category classification @ 40 MHz



Analyzing collision debris ≈
Multi-category classification @ 40 MHz





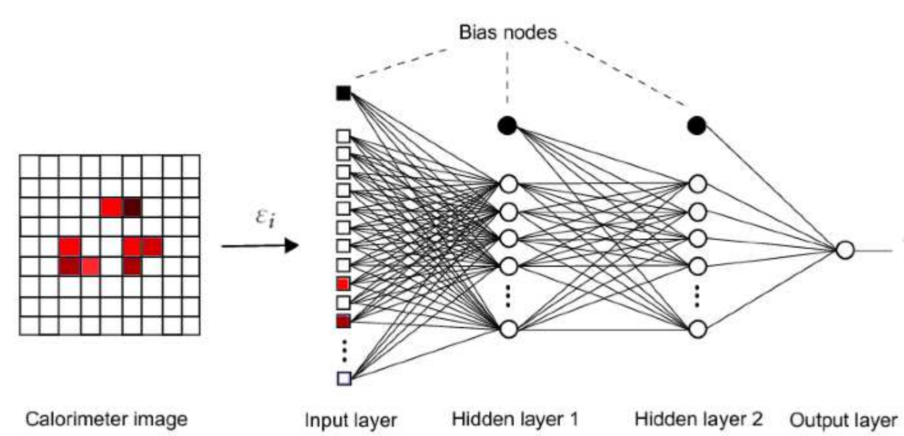
"Deep Thinking"?

[e.g. JDT, Van Tilburg, 1011.2268, 1108.2701;
rephrased in language of Komiske, Metodiev, JDT, 1902.02346]

$$\tau_N(\mathcal{J}) = \min_{|\mathcal{J}'|=N} \mathrm{EMD}(\mathcal{J}, \mathcal{J}')$$

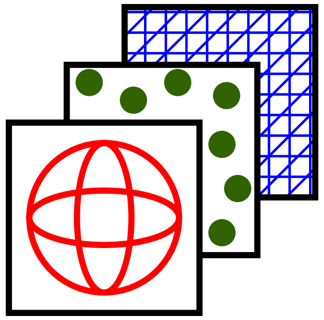
"Deep Learning"?

[e.g. Almeida, Backović, Cliche, Lee, Perelstein, 1501.05968;
review in Kasieczka, Plehn, et al. (Komiske, Metodiev), 1902.09914]



Particle Physics Primer

Jets and Point Clouds



Energy Flow Networks

Today's Goal

Teach a Machine to Communicate to a Physicist

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

Safety: $\vec{p} = \{E, n_x, n_y, n_z\}$

(optional)

Energy weighting

Properly specified problem $\Rightarrow$

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

The Machine

$\Rightarrow$ **Solution** $S(\mathcal{J})$

$\Rightarrow$ **Verification** $V(\mathcal{J})$

Patrick Komiske

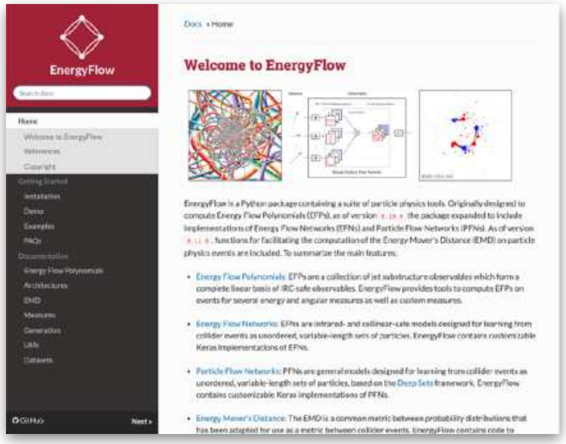
Eric Metodiev

Thank you!

(Theoretical)
Particle
Physics

?

Mathematics,
Statistics,
Computer Science



Energy Flow Networks

https://energyflow.network/

Building the Verification Function

Addition: Starting point for most "Deep Thinking"

Additive Verification: $V(\mathcal{J}) = \sum_{i \in \mathcal{J}} \Phi(\vec{p}_i)$

Permutation symmetry

Per particle function

Safe Additive Verification: $V(\mathcal{J}) = \sum_{i \in \mathcal{J}} E_i \, \Phi(\vec{n}_i)$

Energy weighting

Per direction function

Building the Solution Function

"Deep Learning" based on Additive Verification?

Solution:

Interpretable Latent Space

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

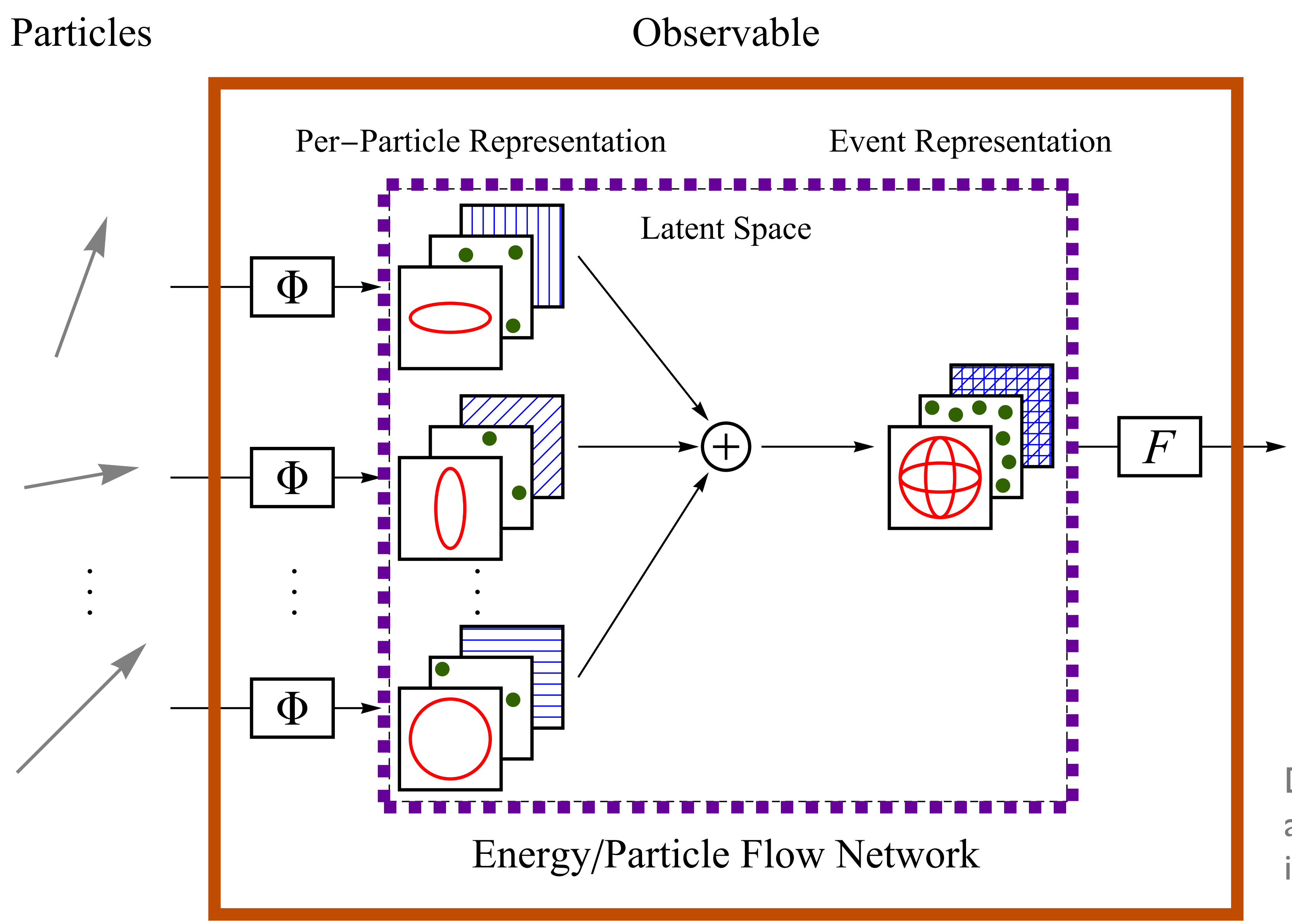
Parametrized with Neutral Networks

Energy / Particle Flow Networks

Safe Φ General Φ

[Komiske, Metodiev, JDT, 1810.05165]

Whatever solution the network learns…



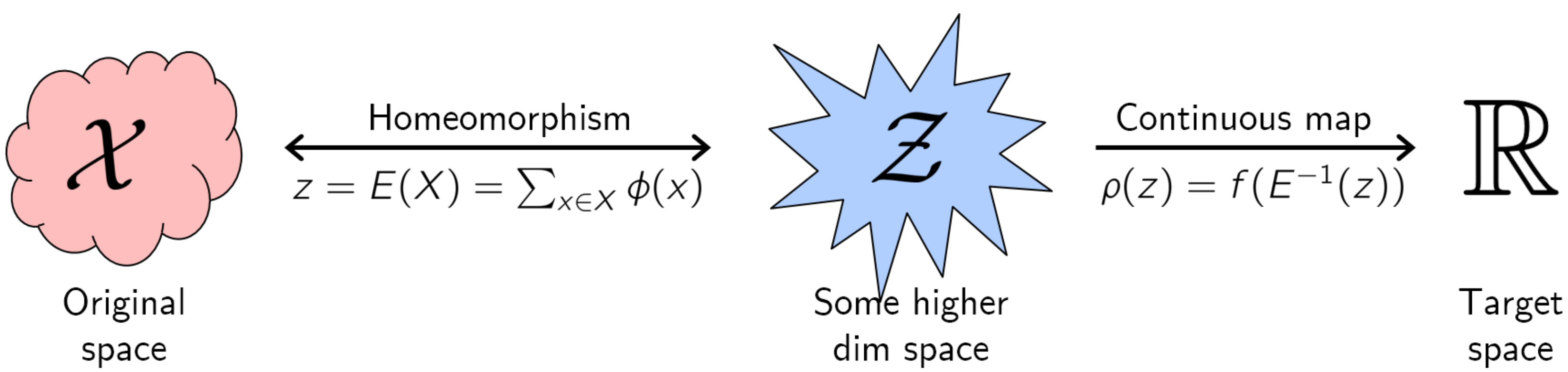
Detailed architecture in backup

…we can verify the latent space

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

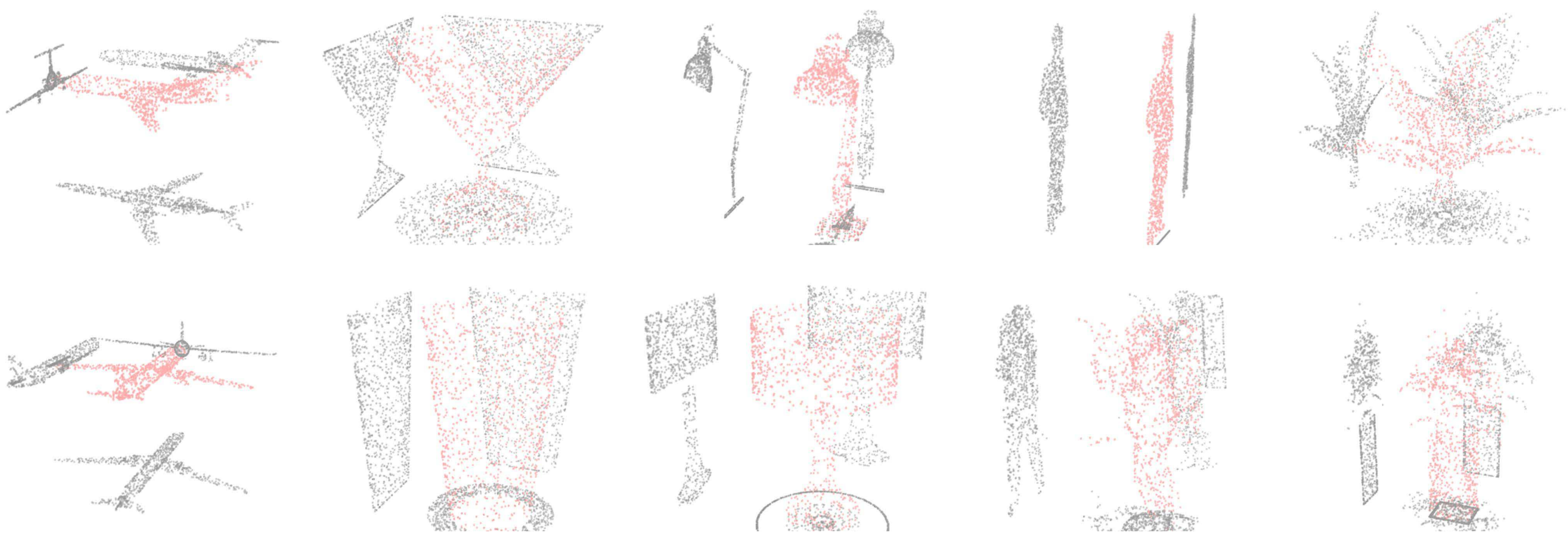
Generic Observable

[Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114; see also Vinyals, Bengio, Kudlur, 1511.06391; Qi, Su, Mo, Guibas, 1612.00593]

Deep Sets for…

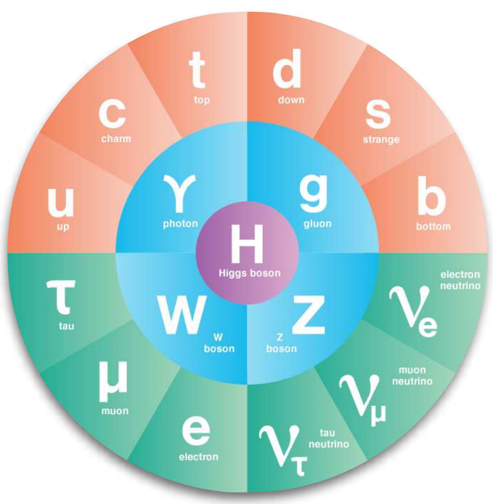
Celebrity Face Anomaly Detection

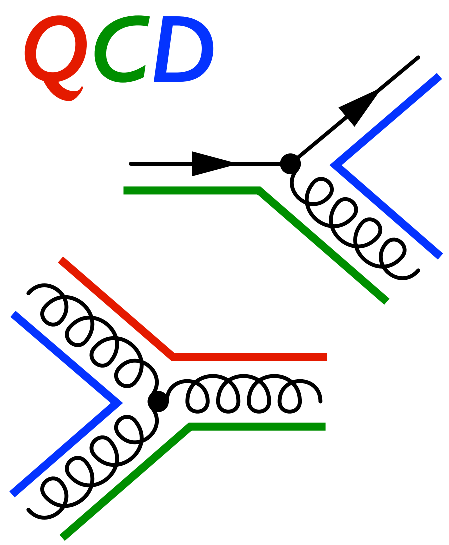
Point Cloud Classification

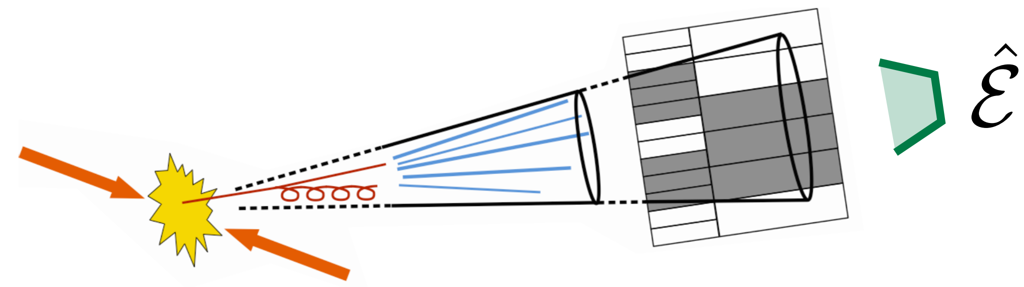


[Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114]

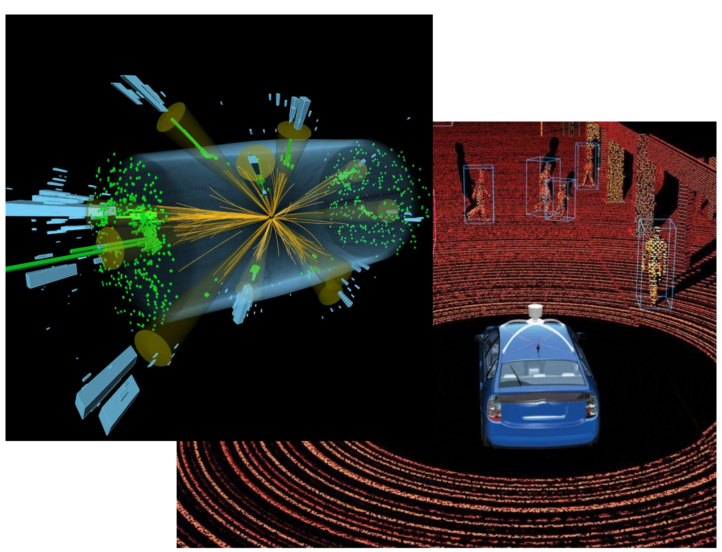
Deep Sets for Particle Jets

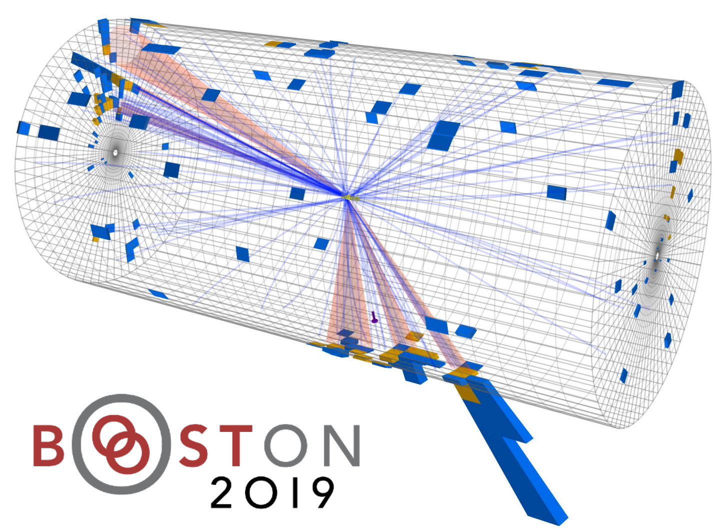


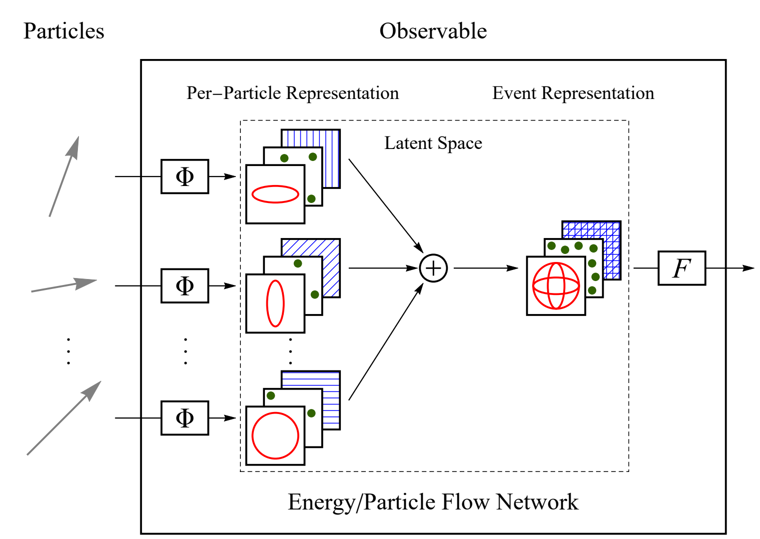




$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{\mathrm{d}\theta}{\theta} \frac{\mathrm{d}z}{z}$$



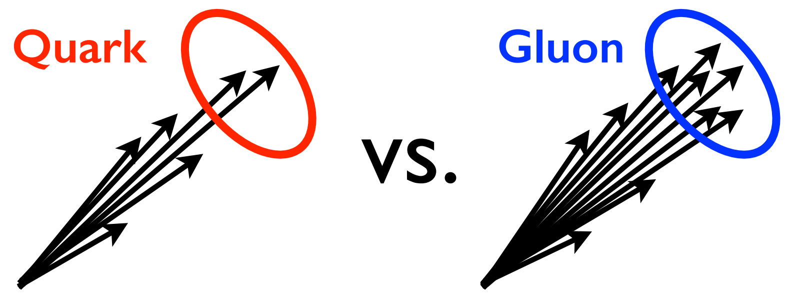


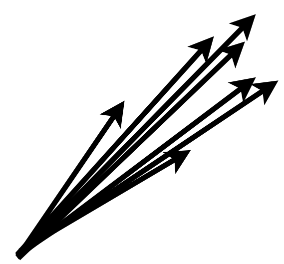


Quark vs. Gluon

The "Hello, World!" of jet classification

1. Specify the problem
2. Find the solution
3. Verify the solution



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

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

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

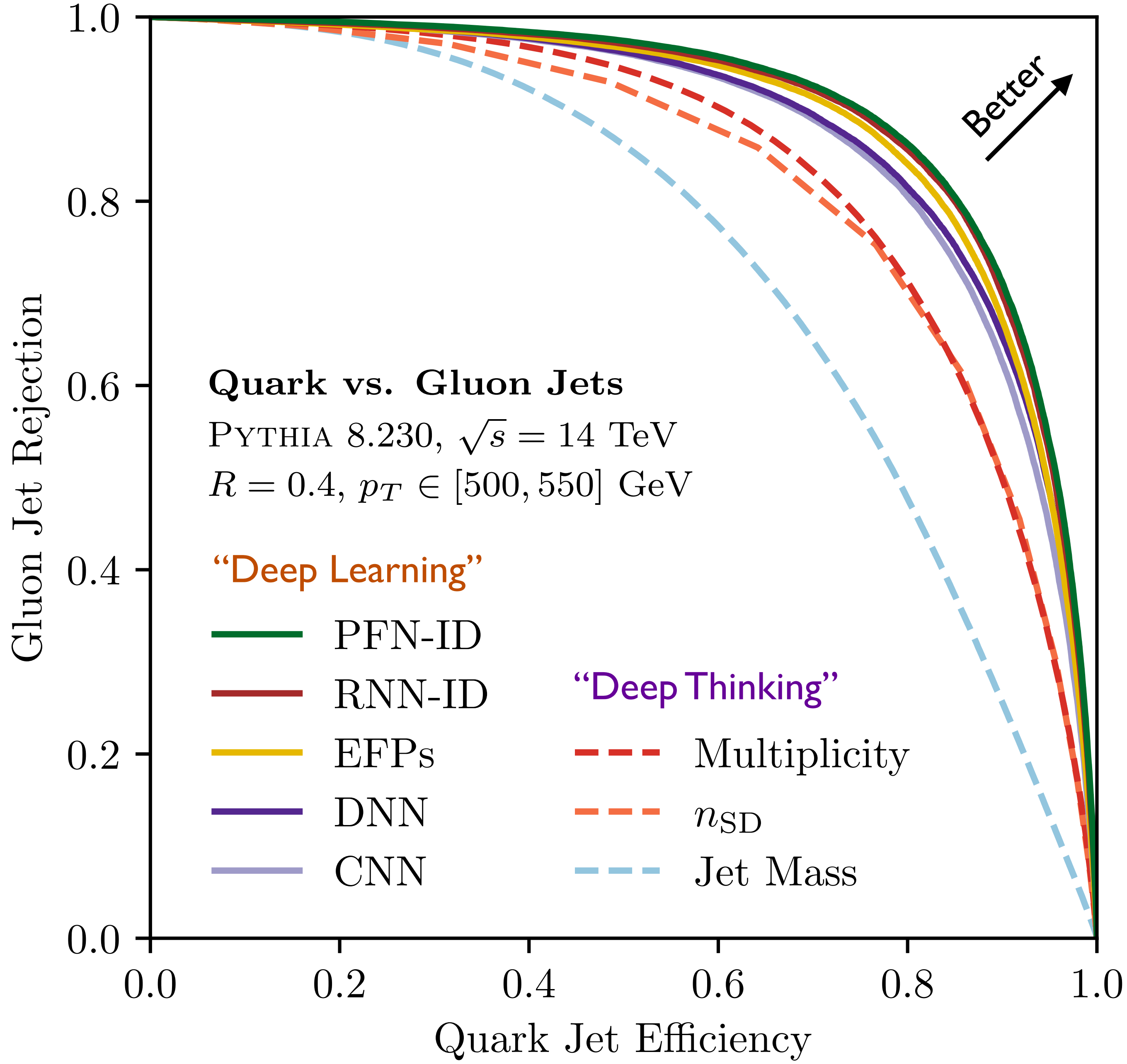
Best you can do: $S(\mathcal{J}) = \dfrac{p(\mathcal{J}|\mathrm{Q})}{p(\mathcal{J}|\mathrm{Q}) + p(\mathcal{J}|\mathrm{G})}$

(Neyman-Pearson lemma)

Quark vs. Gluon

The "Hello, World!" of jet classification

1. Specify the problem
2. Find the solution
3. Verify the solution



ML does amazingly well

[Komiske, Metodiev, JDT, 1810.05165]

Quark vs. Gluon

The "Hello, World!" of jet classification

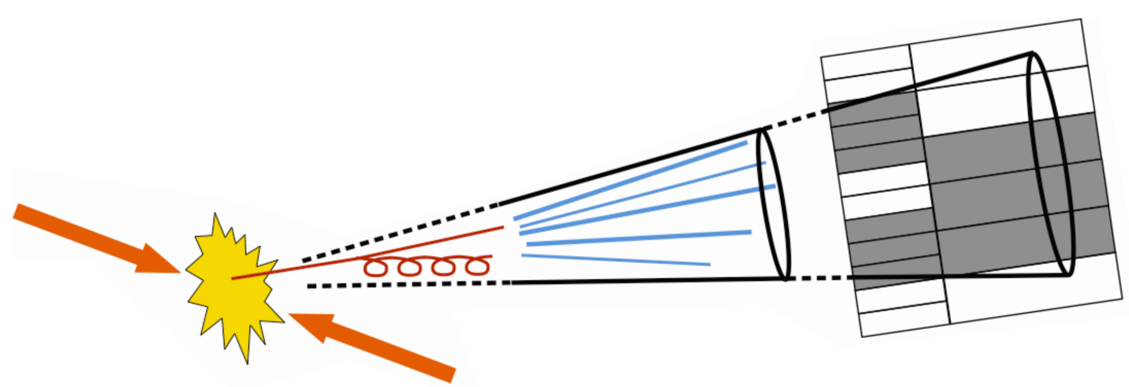
1. Specify the problem
2. Find the solution
3. **Verify the solution**

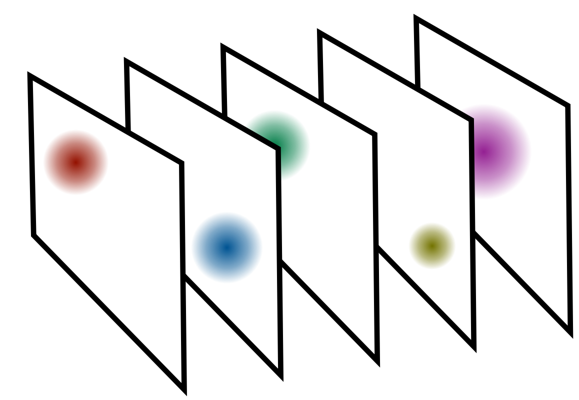
Energy Flow Networks: *ML designed to not get in the way*

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

Difficult to visualize (unless ℓ is small)

Easy to plot these!

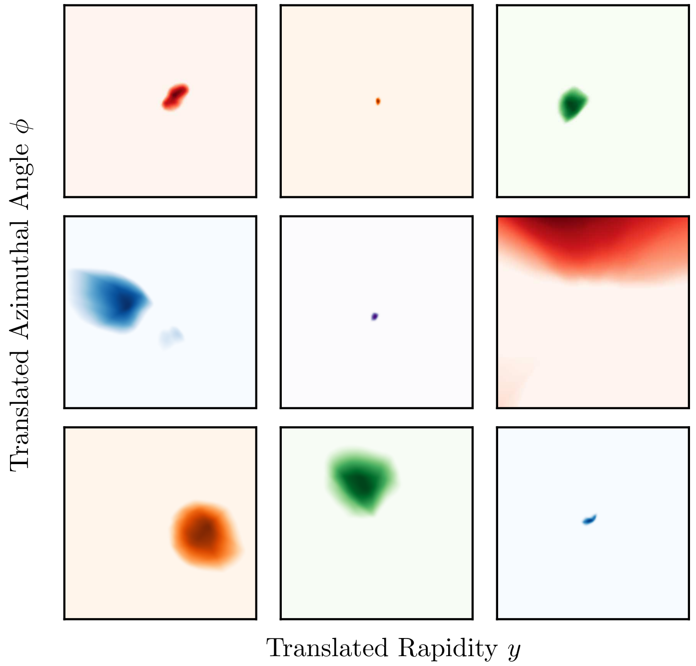




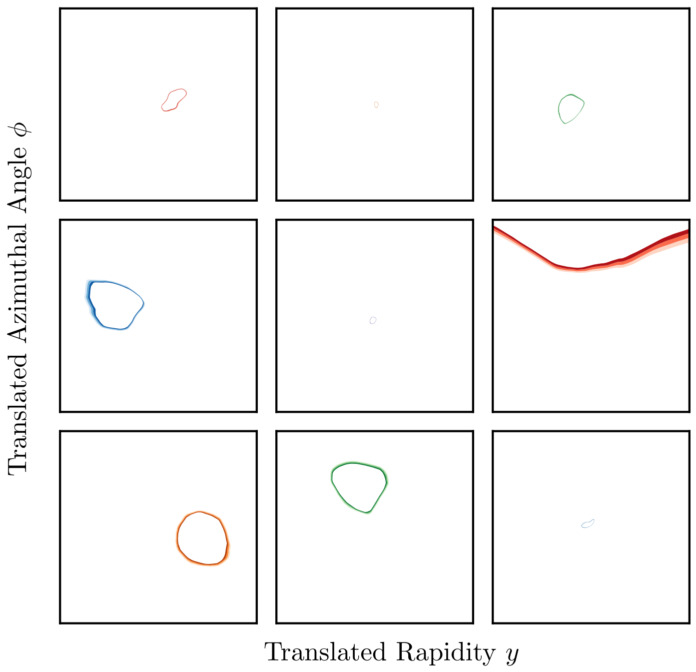
(similar to CNN filter activation)

Psychedelic Network Visualization

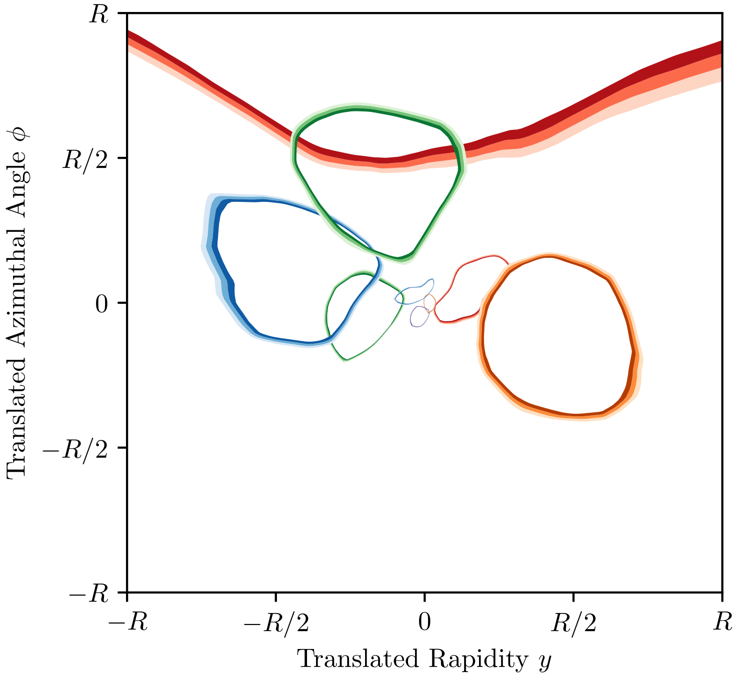
Latent Filters



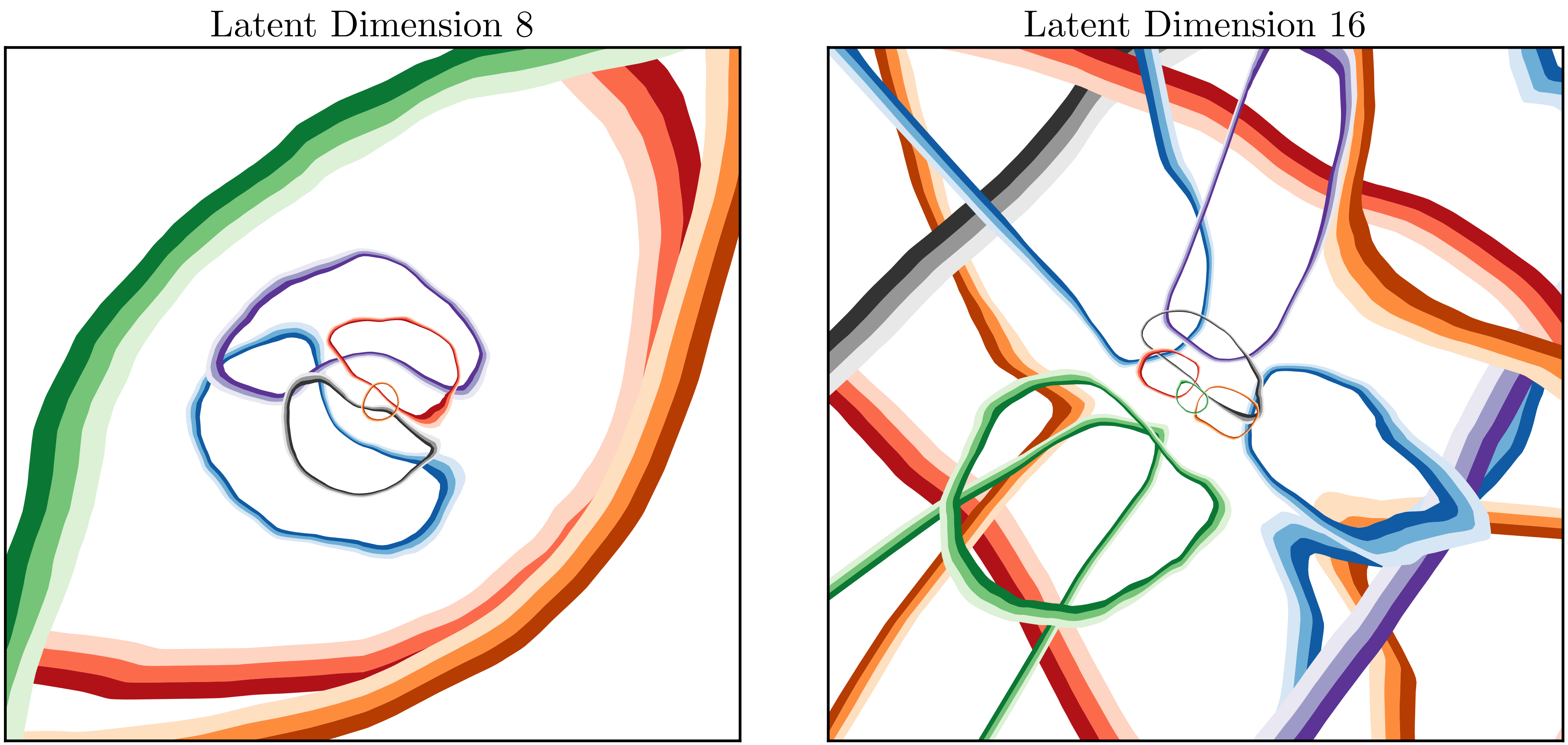
50% Contours



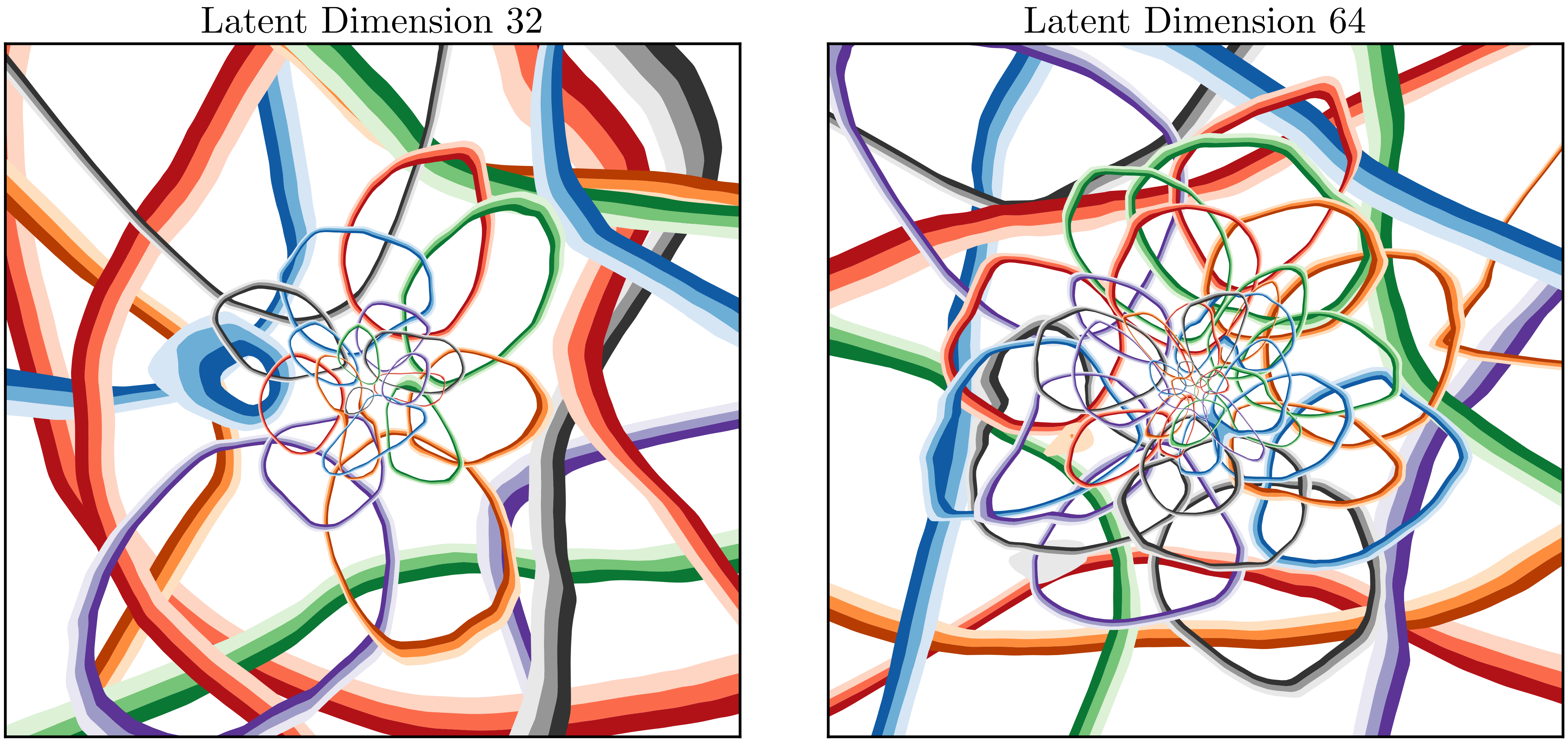
Overlay



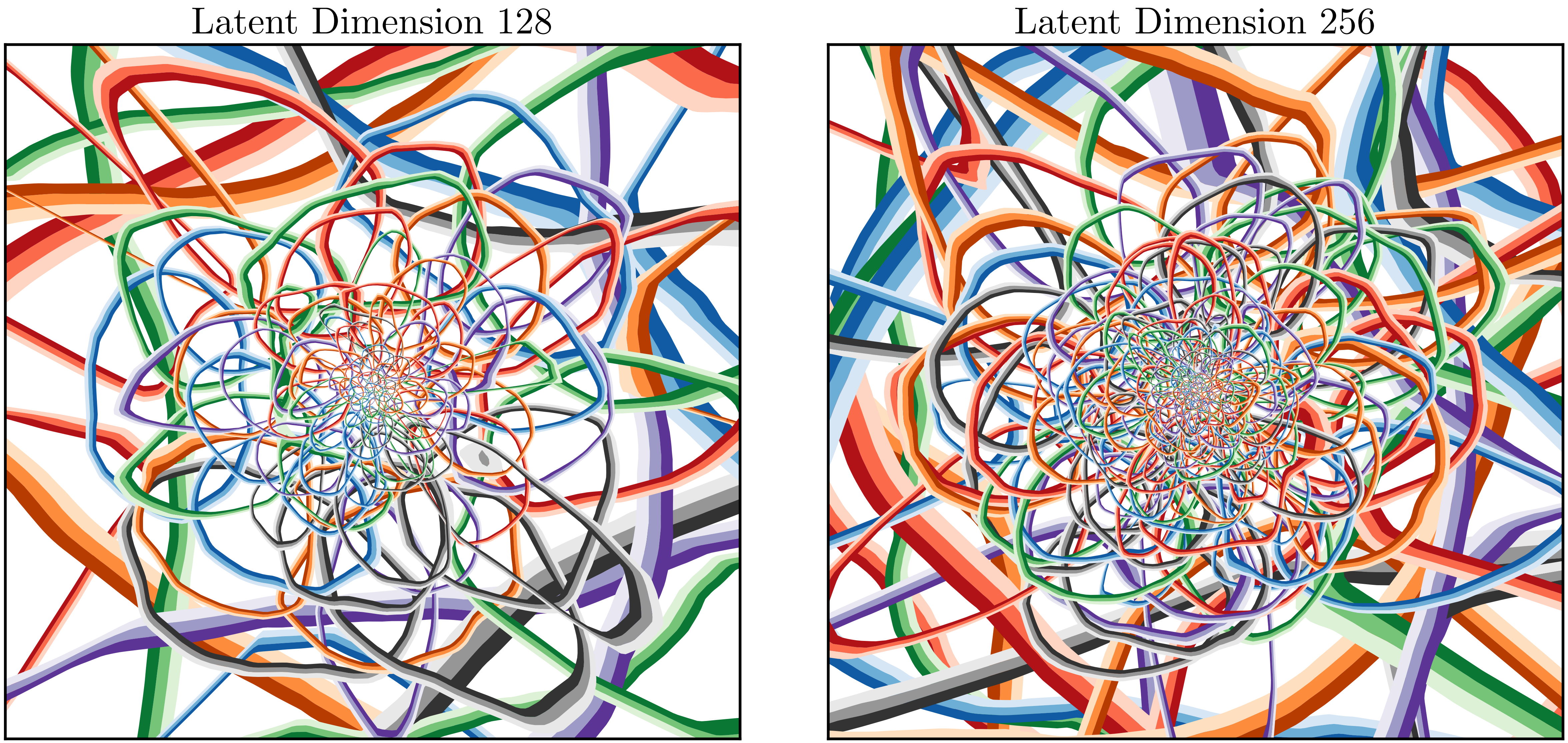
Psychedelic Network Visualization



Psychedelic Network Visualization

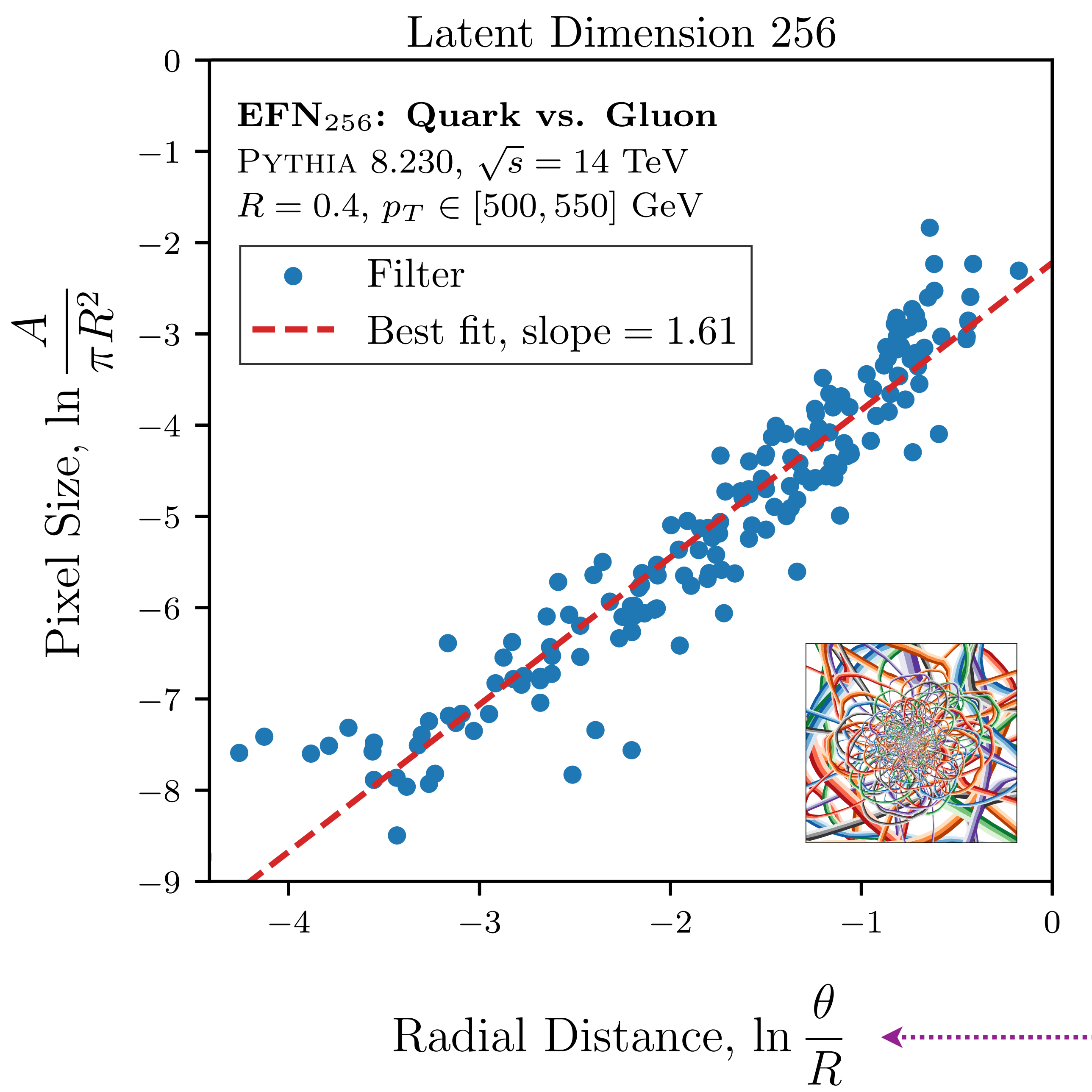


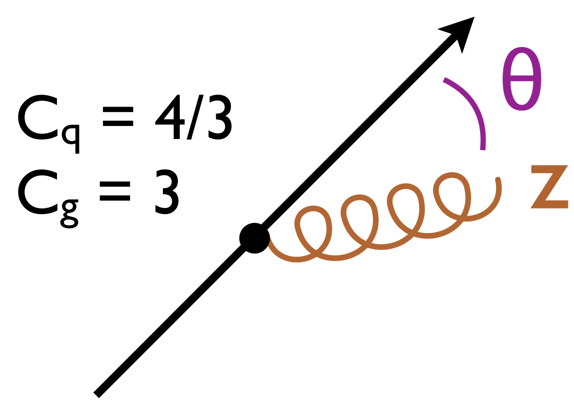
Psychedelic Network Visualization



Altarelli-Parisi!

Learning the Structure of QCD

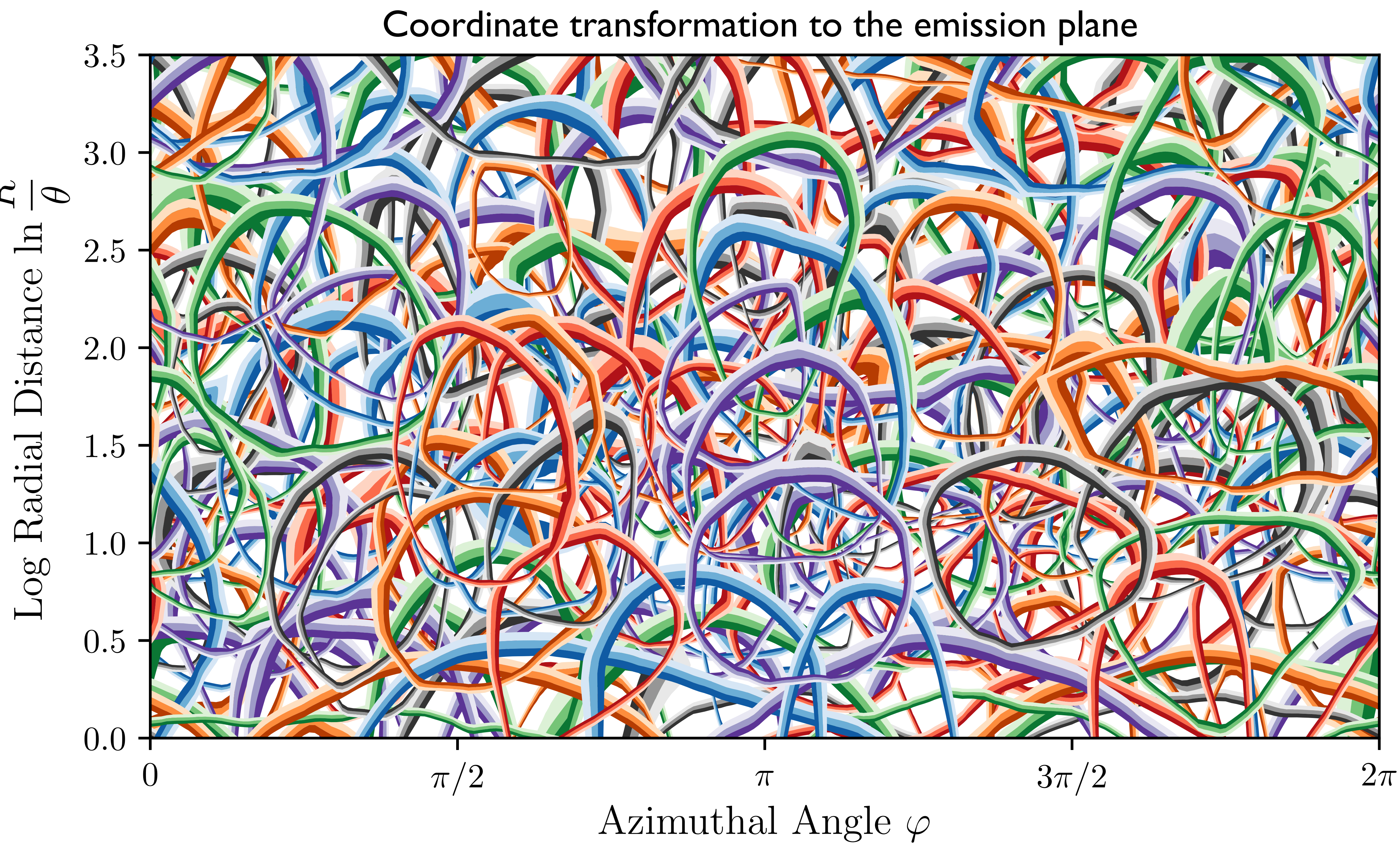




$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \underbrace{\frac{\mathrm{d}\theta}{\theta}}_{\text{Collinear}} \underbrace{\frac{\mathrm{d}z}{z}}_{\text{Soft}}$$

[Komiske, Metodiev, JDT, 1810.05165]

Ready for the SFMOMA?



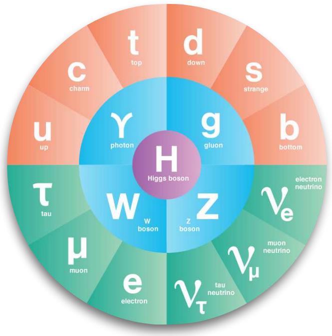
[Komiske, Metodiev, JDT, 1810.05165]

“Collision Course”

Advances in hep-th, -ph, -ex	⇔	*Advances in cs.LG, stat.ML*
Particle Flow Networks	← Point Clouds —	Deep Sets
Energy Flow Networks	— Weighted Point Clouds →	??

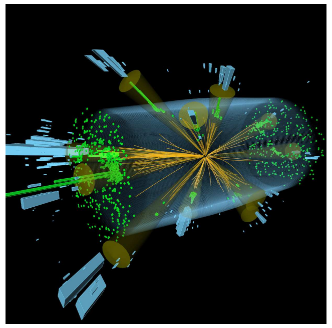
Particle physics is a fascinating domain with rich data sets, established algorithms, and strong theory foundations

Summary



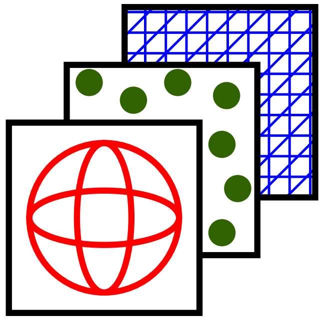
Particle Physics Primer

A rich domain with many machine learning opportunities



Jets and Point Clouds

The importances of symmetries (and safety)



Energy Flow Networks

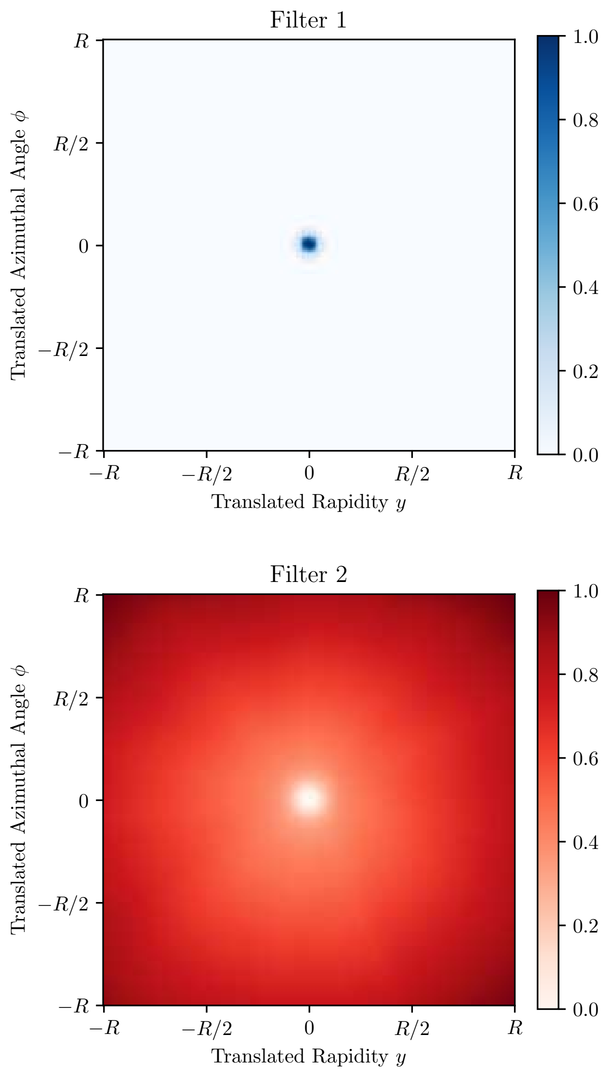
A new architecture for weighted point clouds

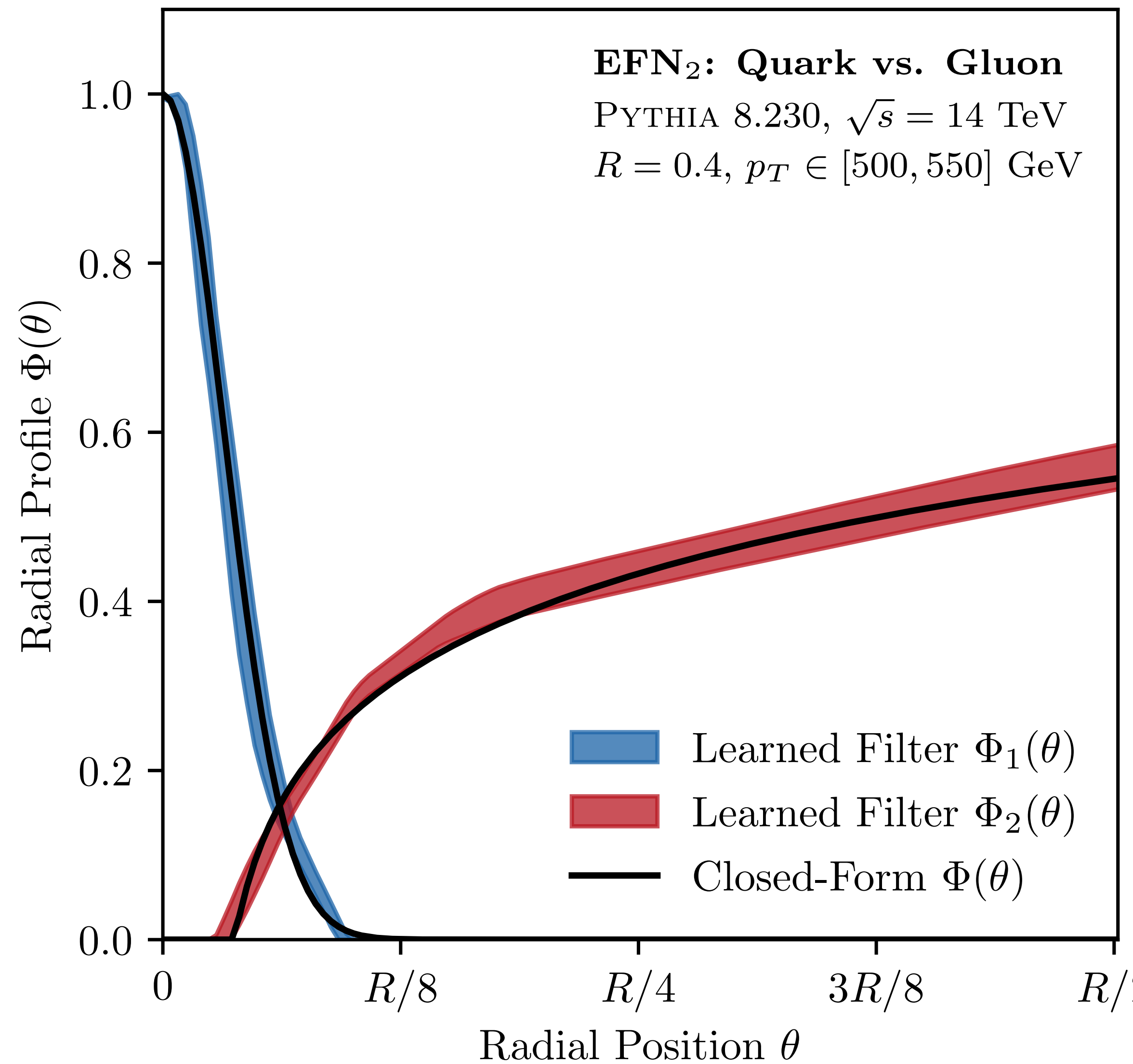
"Ok, but did you really learn something you didn't already know?"

Learning from the Machine

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

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

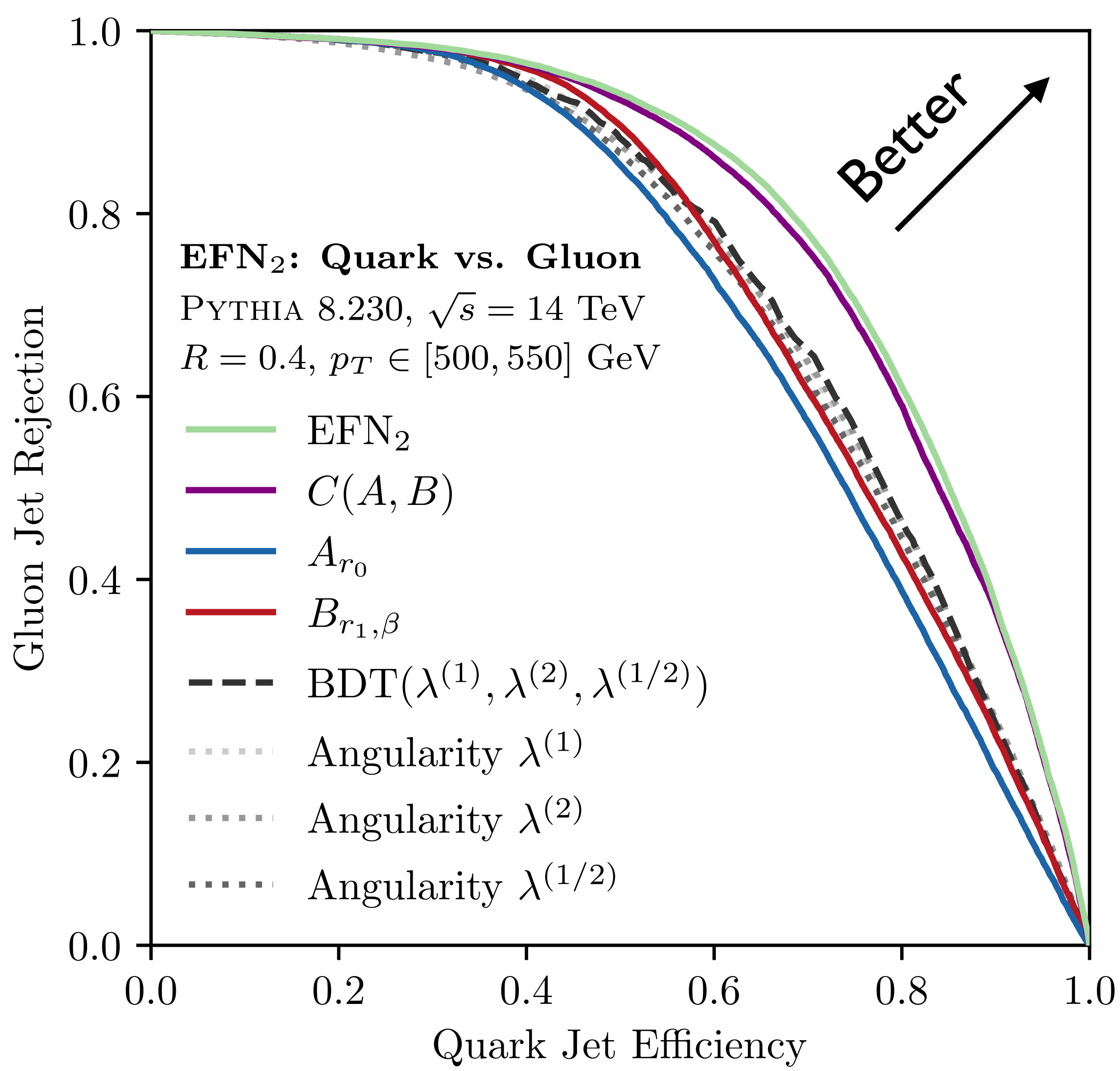




Learning from the Machine

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

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



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

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

Patrick Komiske

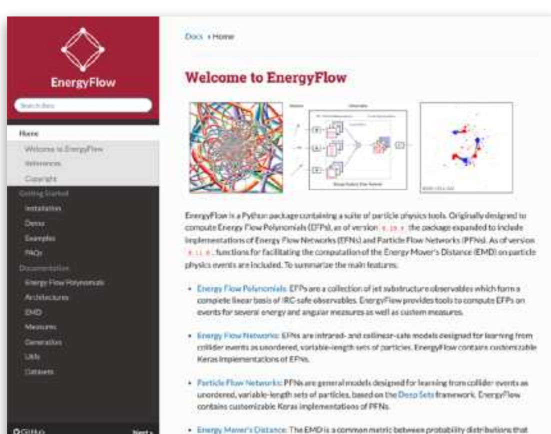
Eric Metodiev

Thank you!

(Theoretical) Particle Physics

?

Mathematics, Statistics, Computer Science



Energy Flow Networks

https://energyflow.network/

Backup Slides

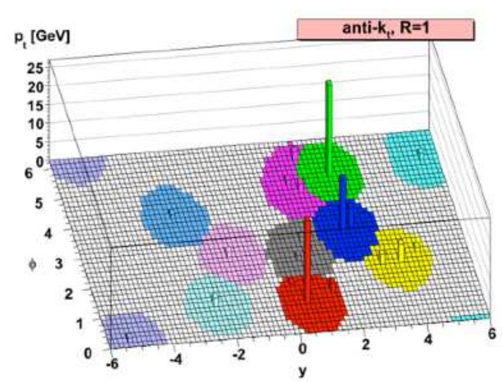
Particle Physics as ML Testbed

- *Huge datasets with reliable simulations*

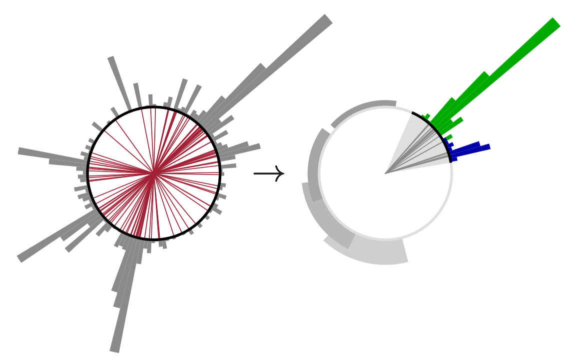
e.g. CMS opendata CERN with MadGraph + Pythia + GEANT4 A SIMULATION TOOLKIT

- *Broad use of (un)supervised algorithms*

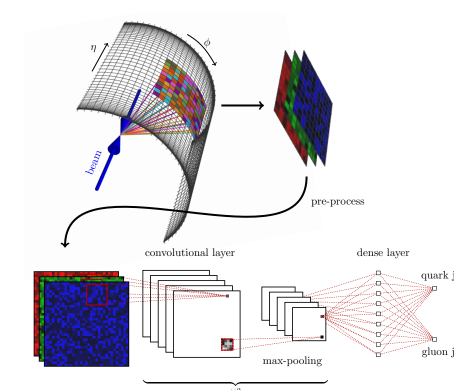
e.g.



Clustering



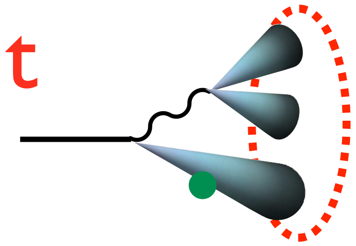
Noise Mitigation



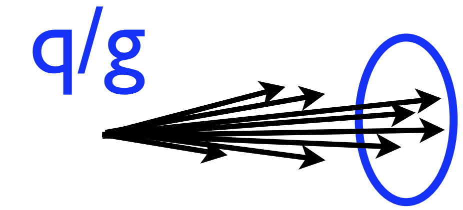
Machine Learning

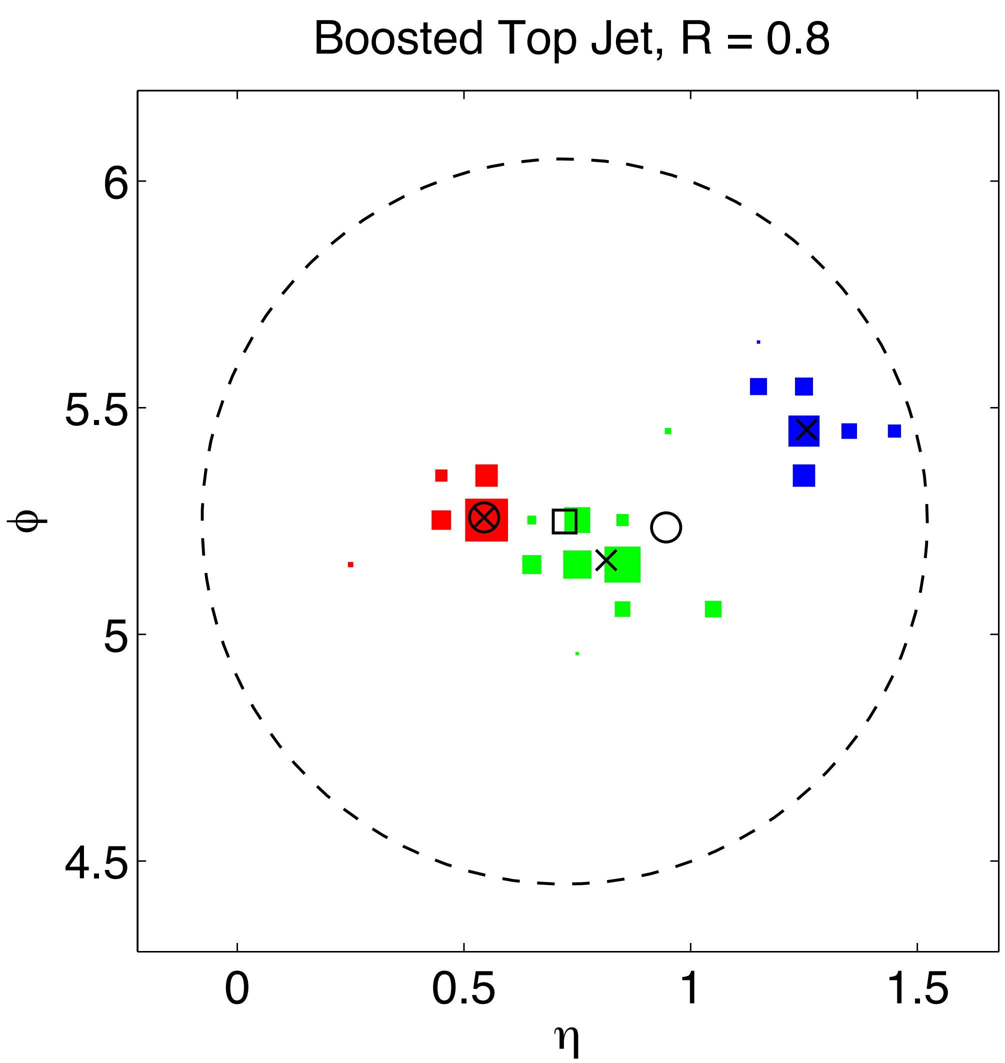
- *Extensive domain knowledge and strong theory foundations*

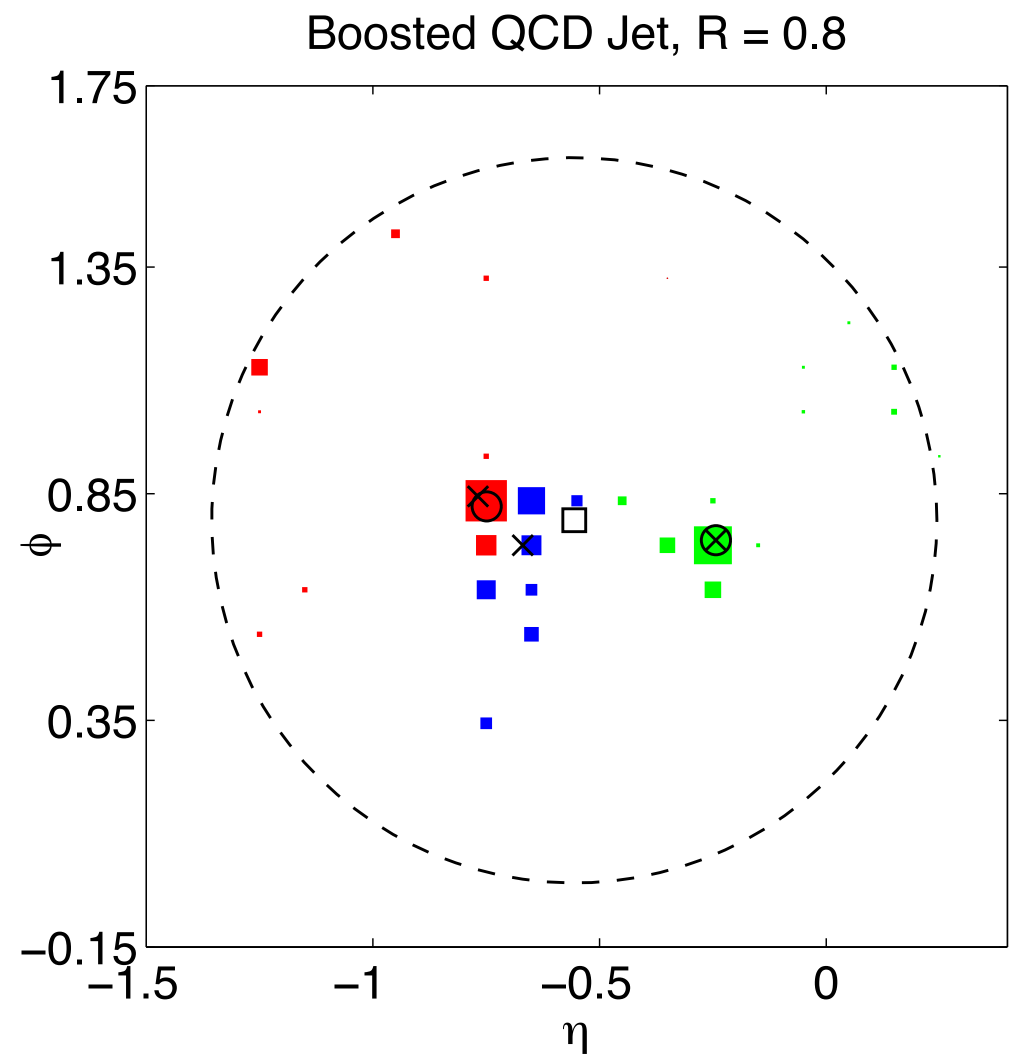
[figures from Cacciari, Salam, Soyez, 0802.1189; Larkoski, Marzani, JDT, Tripathee, Xue, 1704.05066; Komiske, Metodiev, Schwartz, 1612.01551]



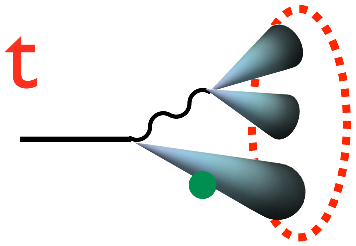
3-Prong vs. 1-Prong



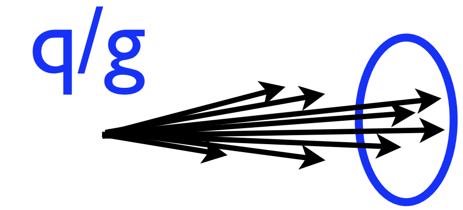


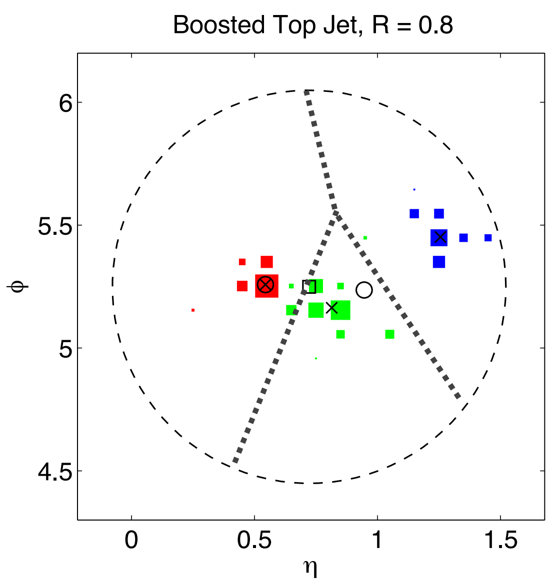


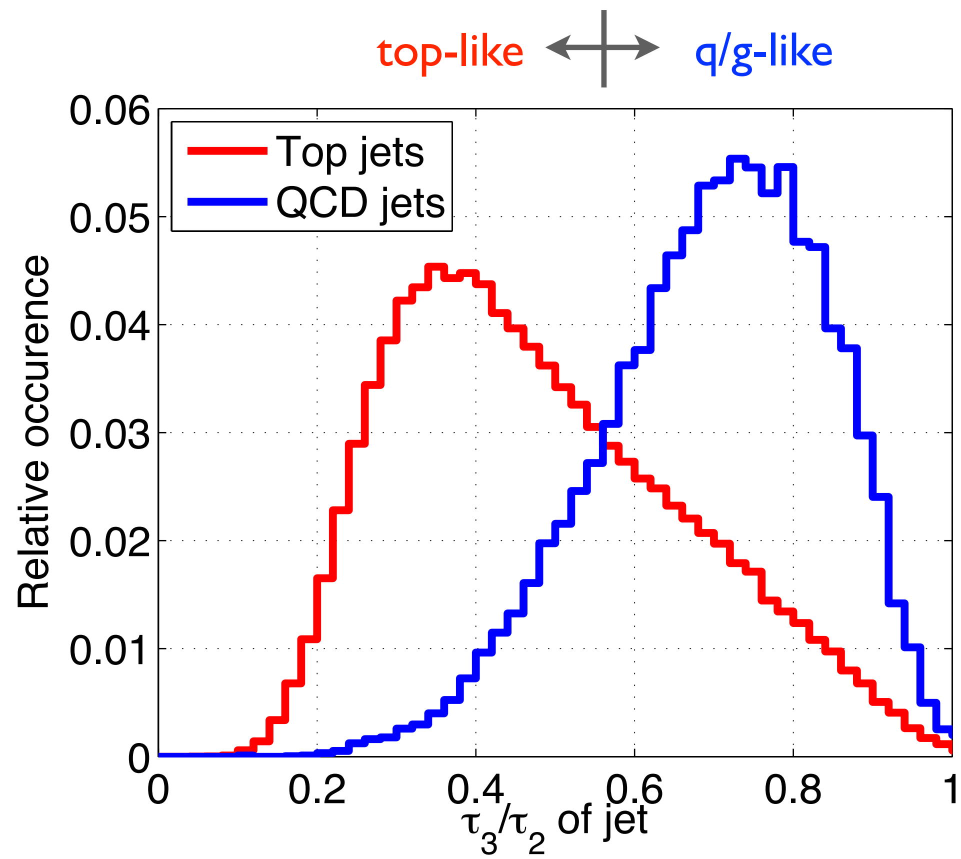
If your eyes can do it…

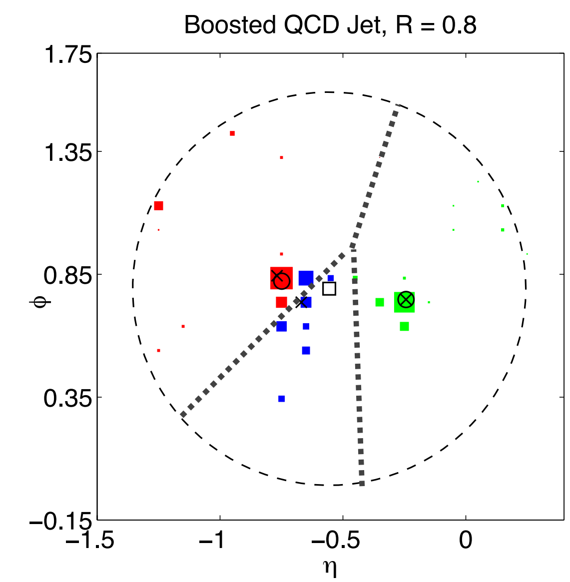


3-Prong vs. 1-Prong





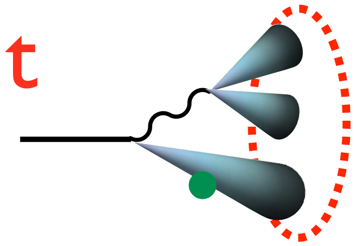




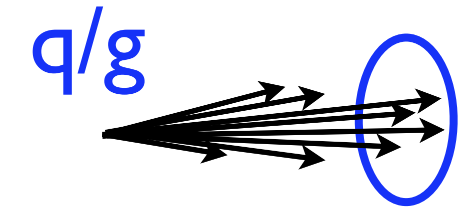
"Deep Thinking": *N-subjettiness*

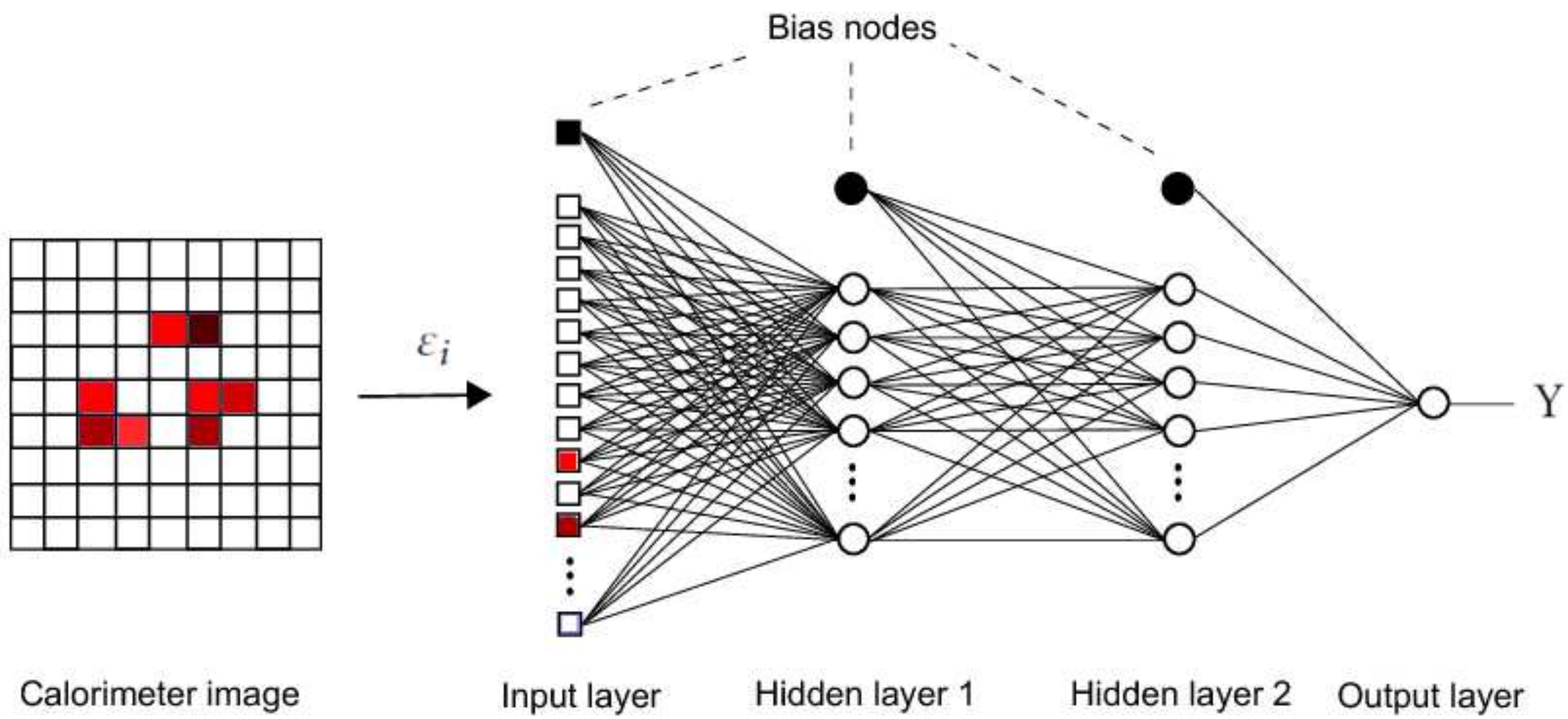
$$\tau_N(\mathcal{J}) = \min_{|\mathcal{J}'|=N} \mathrm{EMD}(\mathcal{J}, \mathcal{J}')$$

[JDT, Van Tilburg, 1011.2268, 1108.2701;
rephrased in language of Komiske, Metodiev, JDT, 1902.02346]



3-Prong vs. 1-Prong





"Deep Learning": BDT, FLD, DNN, CNN, P-CNN, RNN, RecNN, LBN, LoLa, DGCNN, EFP, PFN, EFN, ResNeXT, …

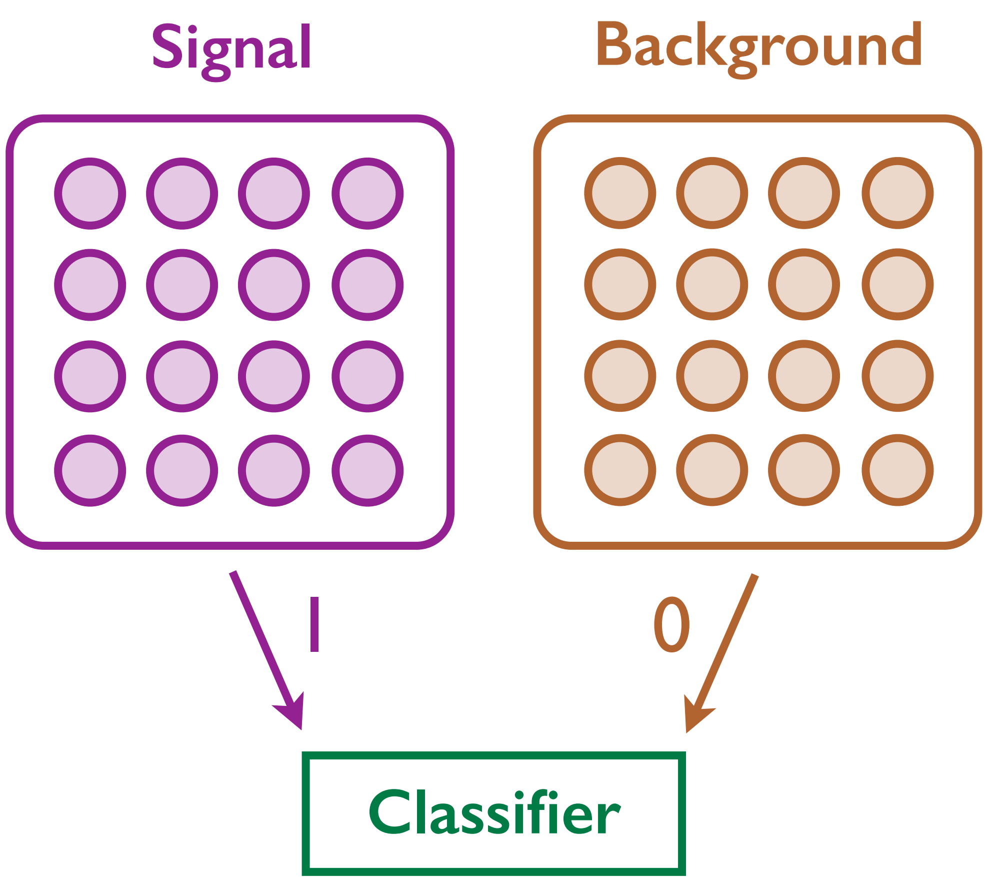
[figure from Almeida, Backović, Cliche, Lee, Perelstein, 1501.05968;
head to head comparison in Kasieczka, Plehn, et al. (Komiske, Metodiev), 1902.09914]

A Cartoon of Machine Learning

For fully-supervised jet classification

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

Classifier Inputs



Minimize Loss Function

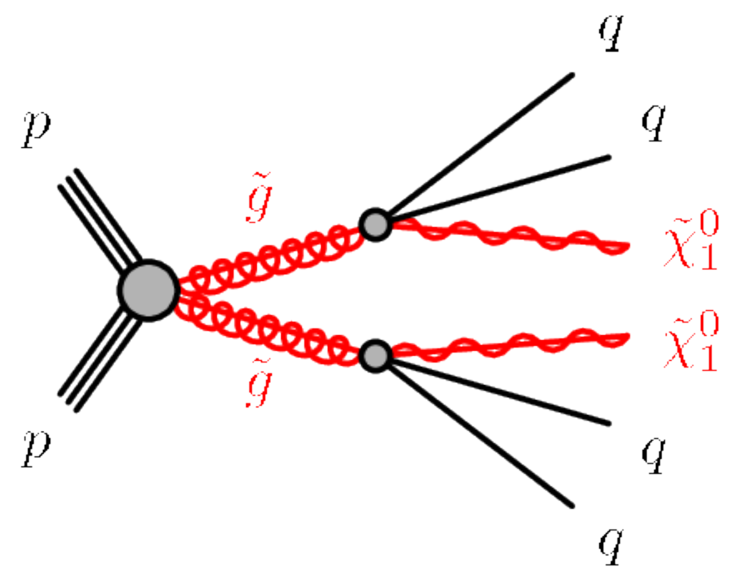
(assuming infinite training sets, and flexible enough functional form)

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

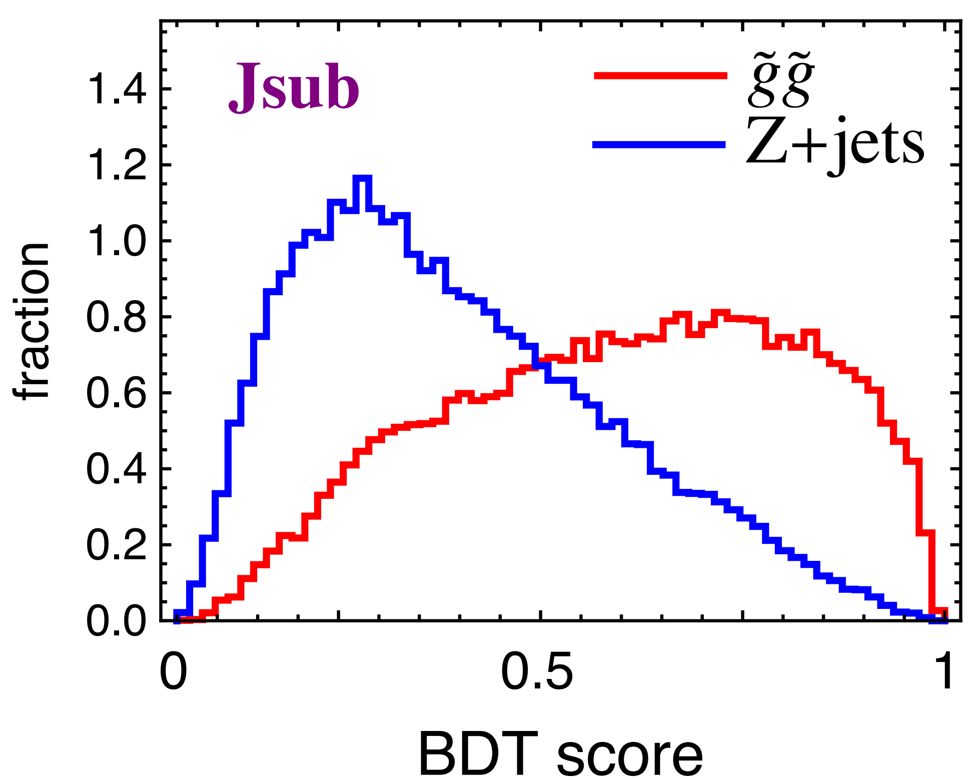
Optimal Classifier (Neyman–Pearson)

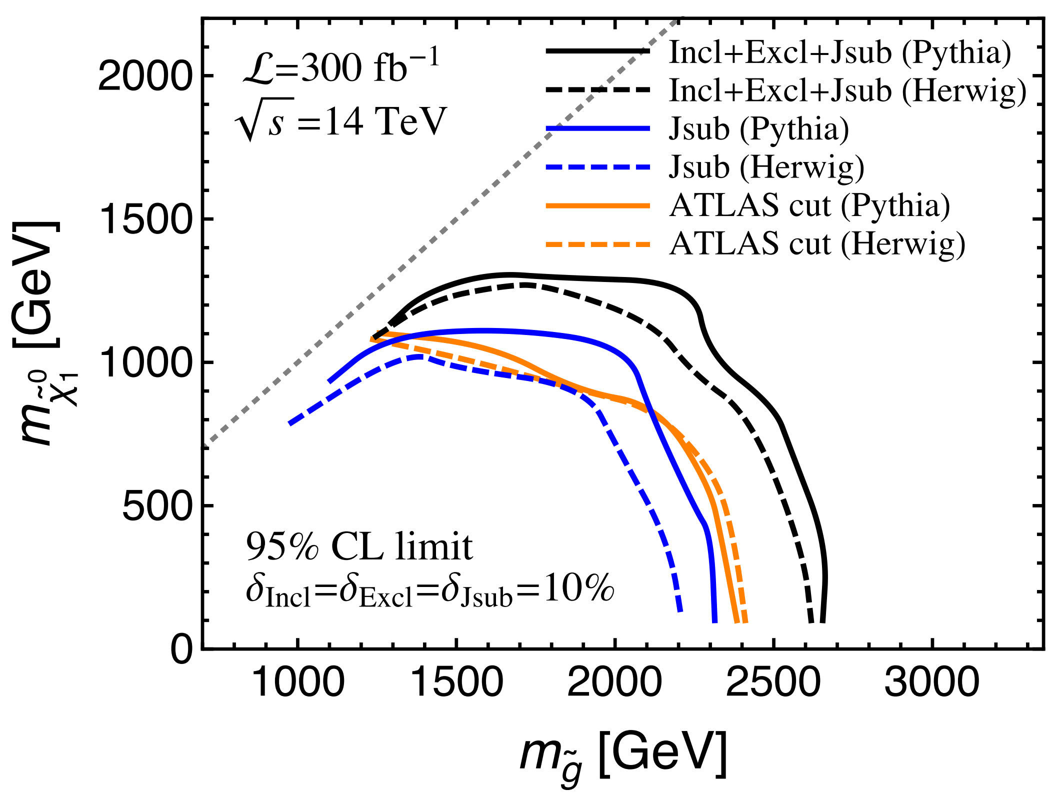
E.g. Search for Gluino Cascade to Dark Matter

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 ($C_A = 3$)



← background
signal →





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

Jet Classification Studies

Mix and match

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

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.	…

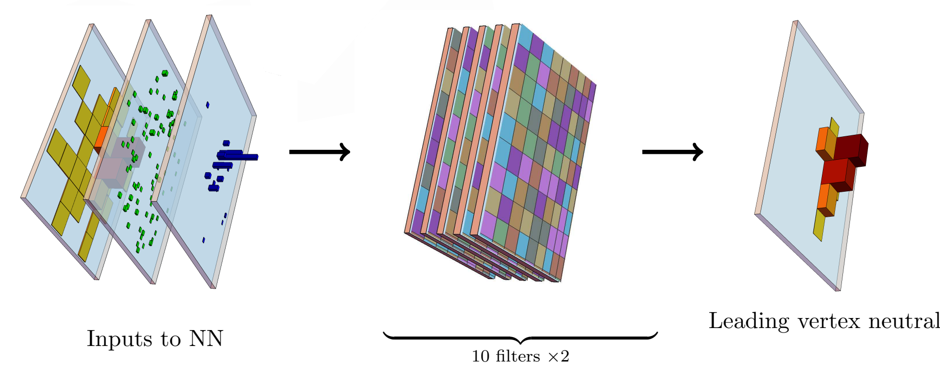
[Lönnblad, Peterson, Rögnvaldsson, 1990, …, Cogan, Kagan, Strauss, Schwartzman, 1407.5675; Almeida, Backović, Cliche, Lee, Perelstein, 1501.05968; de Oliveira, Kagan, Mackey, Nachman, Schwartzman, 1511.05190; Baldi, Bauer, Eng, Sadowski, Whiteson, 1603.09349; Conway, Bhaskar, Erbacher, Pilot, 1606.06859; 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 (and not updating this after July 23, 2018); plus many ATLAS/CMS performance studies]

Beyond Classification

PUMML

Pileup Mitigation

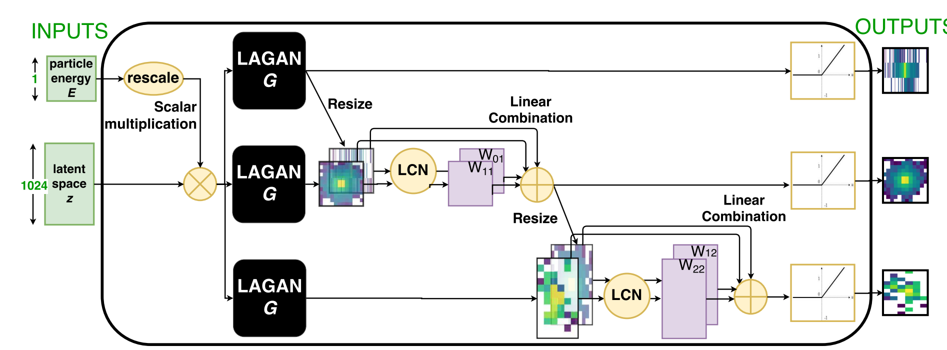
[Komiske, Metodiev, Nachman, Schwartz, 1707.08600]



CaloGAN

Fast Detector Simulation

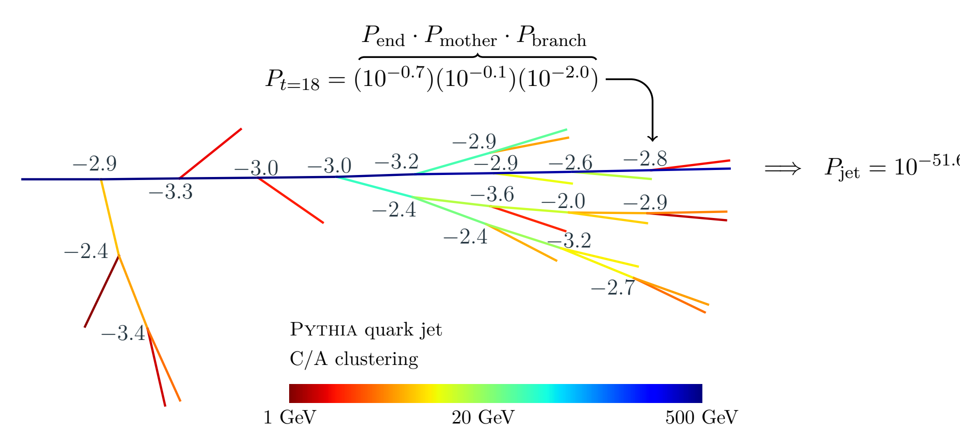
Paganini, de Oliveira, Nachman, 1705.02355, 1712.10321;
see also de Oliveira, Michela Paganini, Nachman, 1701.05927]



JUNIPR

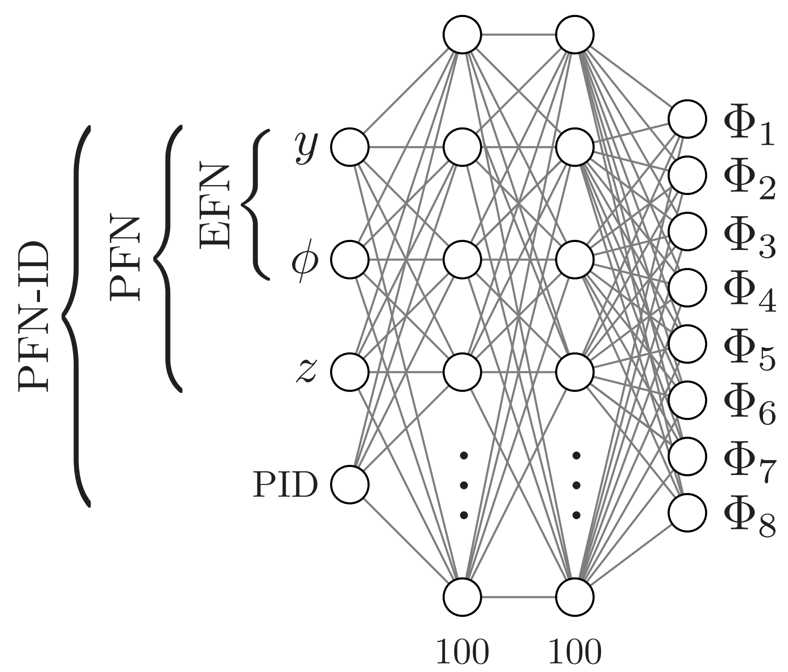
Probability Modeling

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



Technical Implementation

Per-Particle Network: Φ

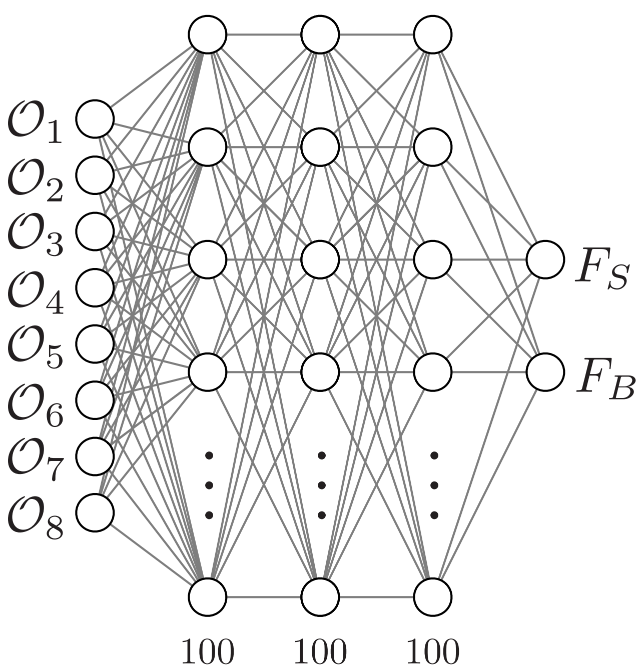


Per-Jet (Latent Space):

EFN: $\mathcal{O}_a = \sum_{i \in \text{jet}} z_i \, \Phi_a(y_i, \phi_i) \qquad z_i = \frac{p_{Ti}}{\sum_j p_{Tj}}$

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

Latent Combiner: F



Final Discriminant:

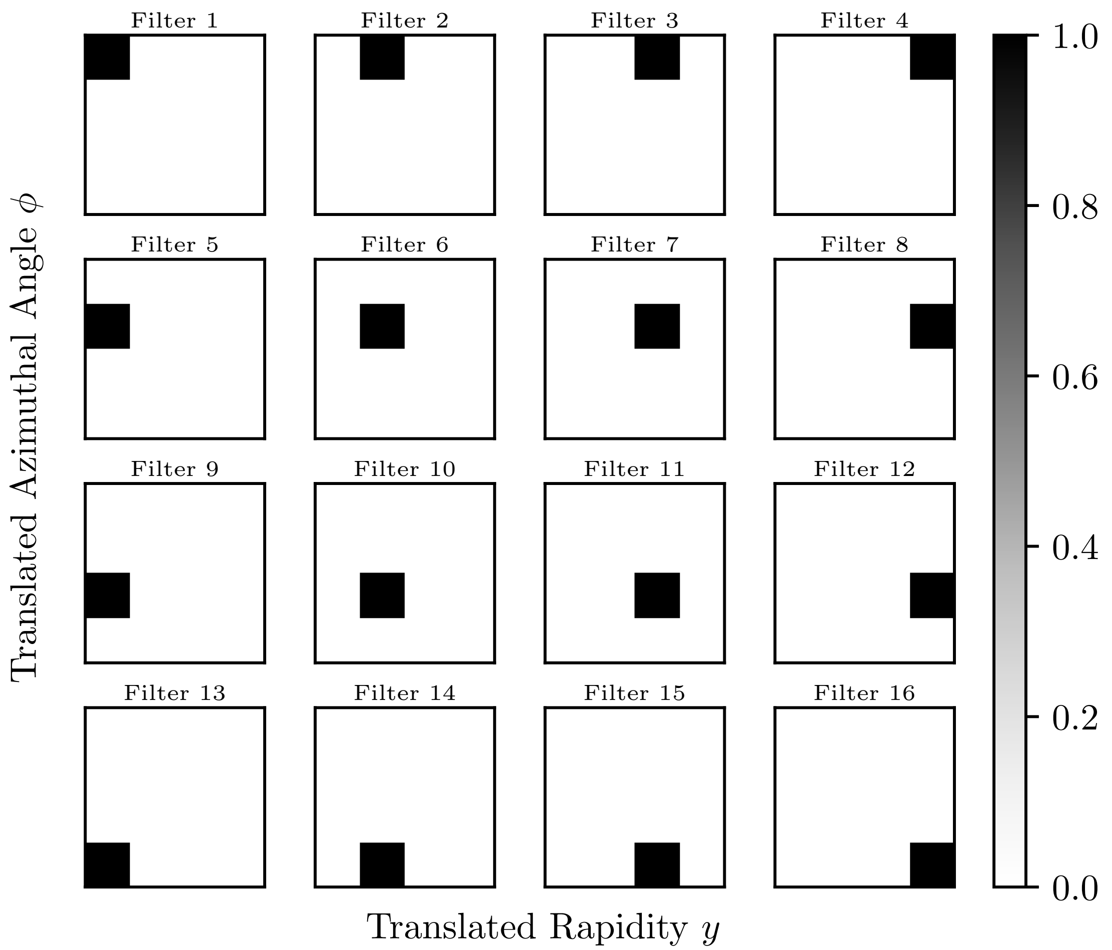
$$\text{softmax}(F_S, F_B)$$

Latent Space Visualization

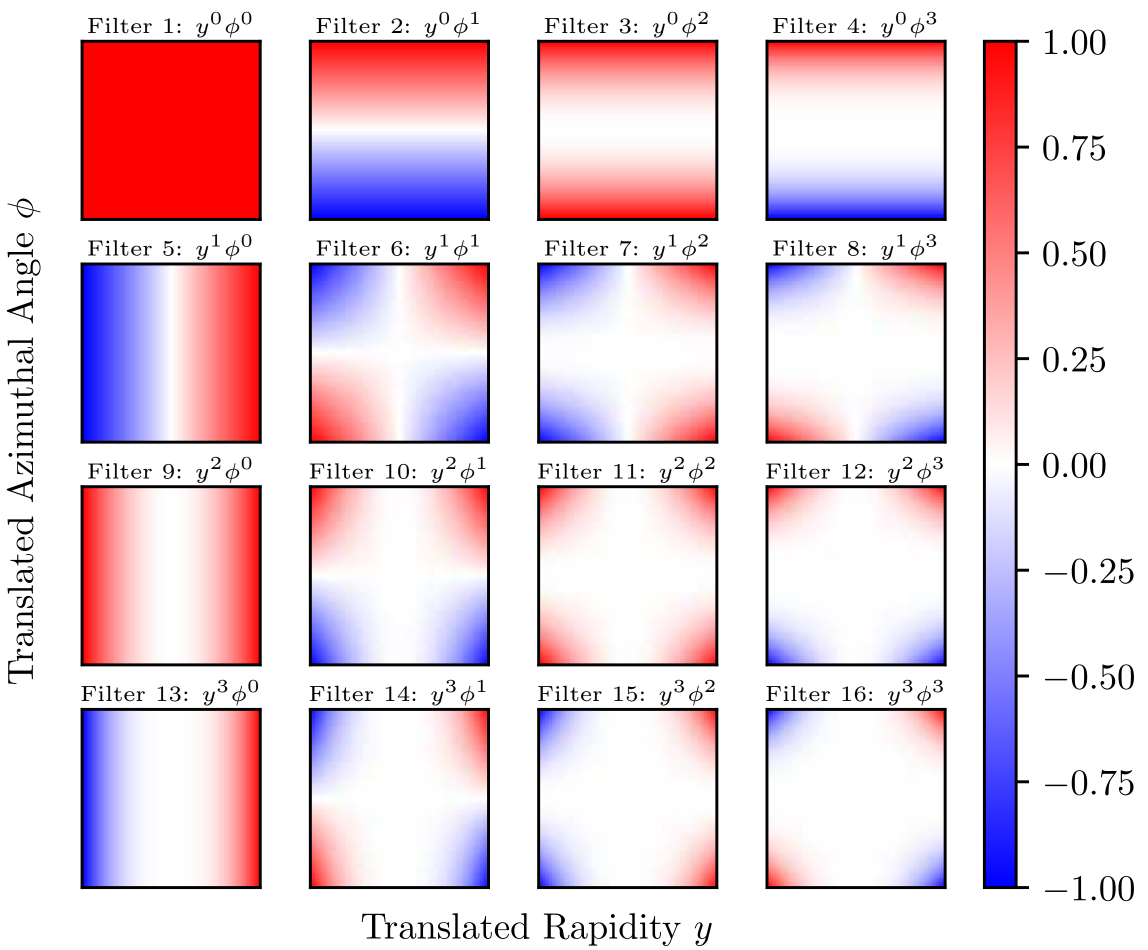
$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_\ell)$$

IRC-safe: $\mathcal{O}_a = \sum_{i \in \text{jet}} E_i \, \Phi_a(\hat{p}_i)$

Calorimeter Pixels



Radiation Moments

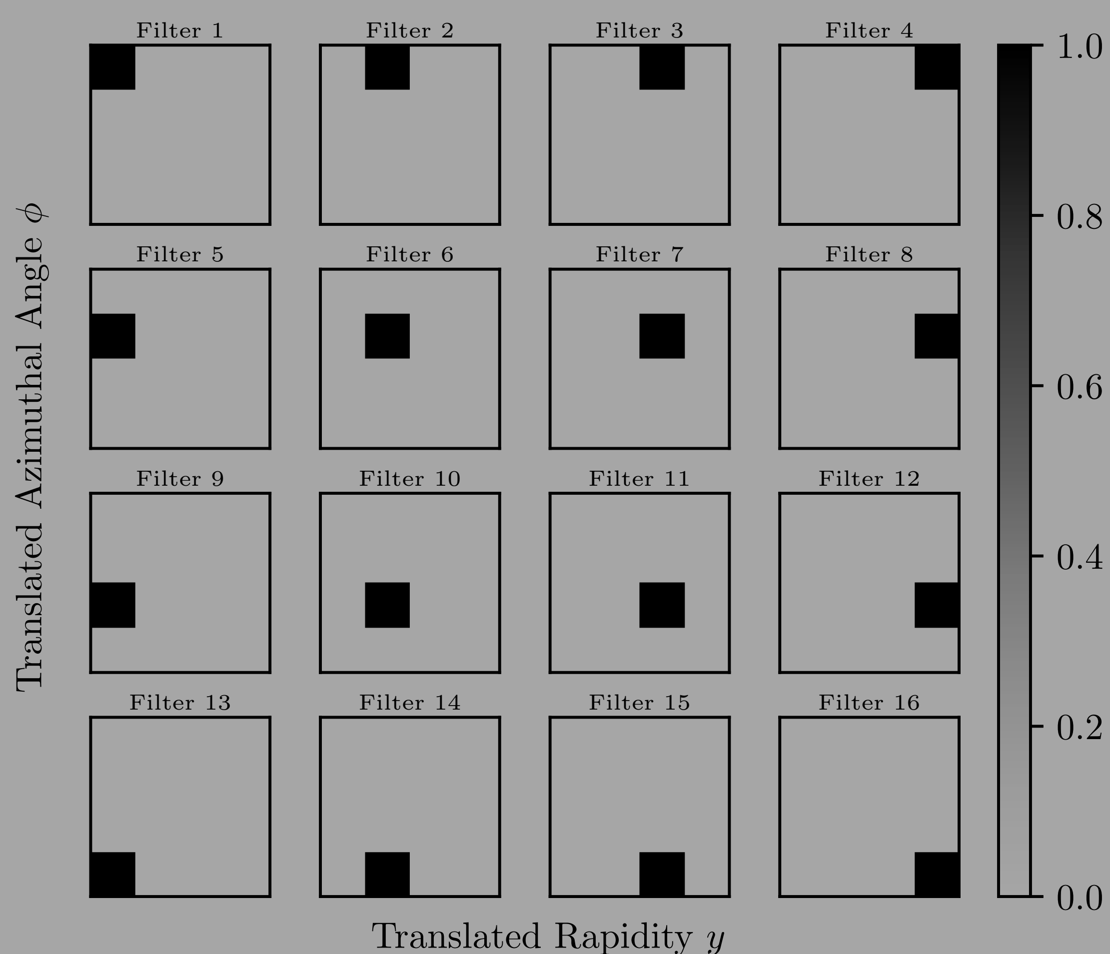


Latent Space Visualization

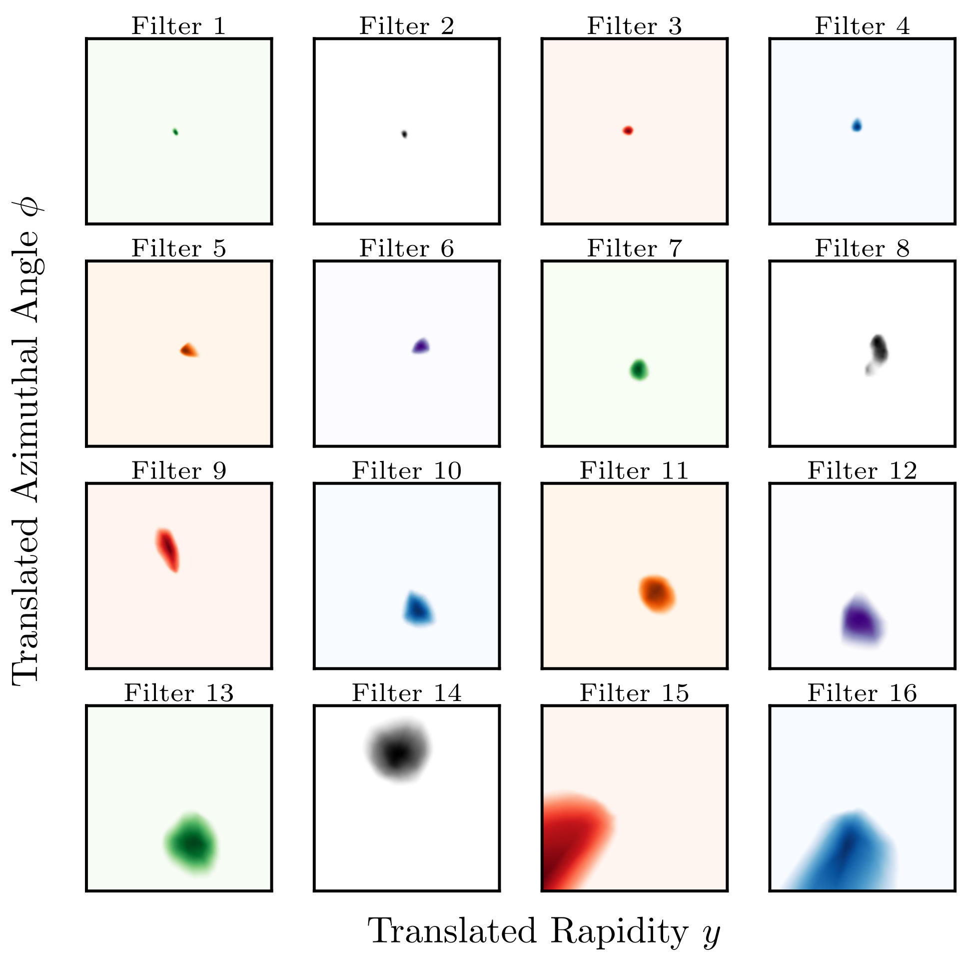
$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_\ell)$$

IRC-safe: $\mathcal{O}_a = \sum_{i \in \text{jet}} E_i \, \Phi_a(\hat{p}_i)$

Calorimeter Pixels

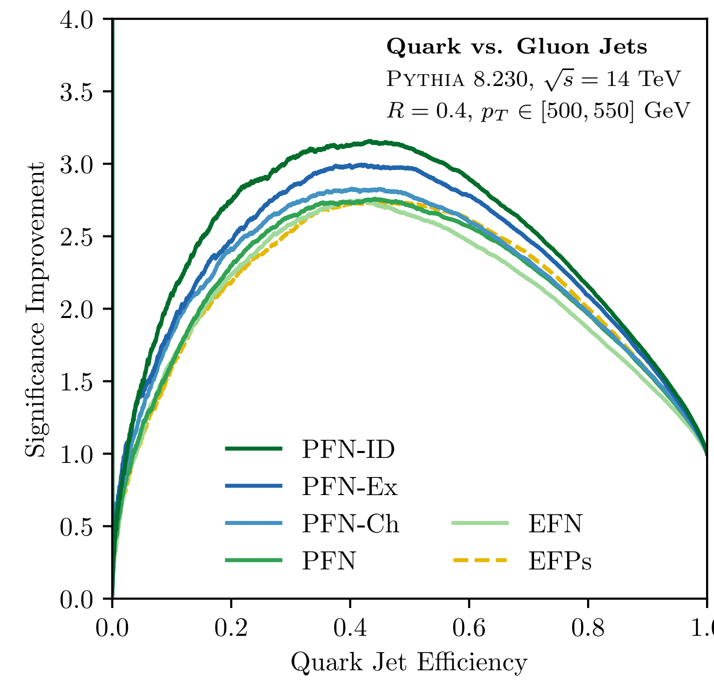


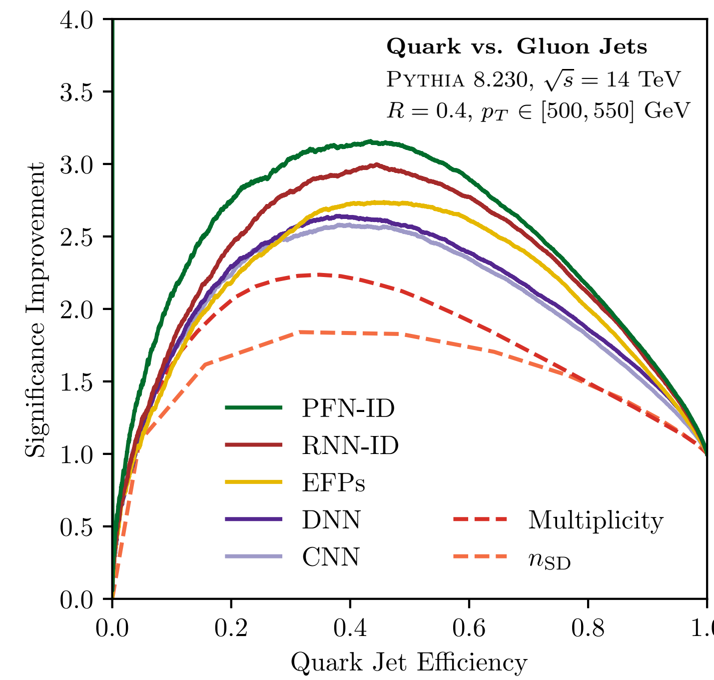
EFNs: Dynamic Pixelation



More Quark/Gluon Performance

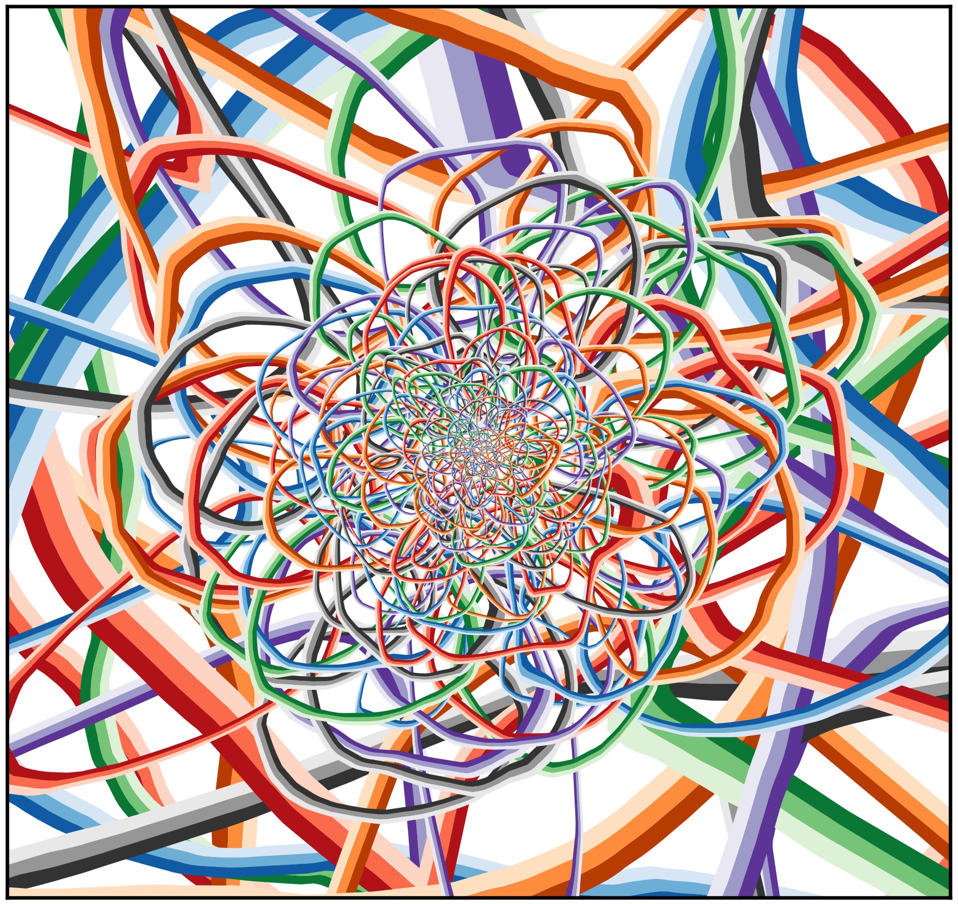
Model	**AUC**	$\mathbf{1/\varepsilon_g}$ **at** $\mathbf{\varepsilon_q = 50\%}$
PFN-ID	$\mathbf{0.9052} \pm 0.0007$	$\mathbf{37.4} \pm 0.7$
PFN-Ex	0.9005 ± 0.0003	34.7 ± 0.4
PFN-Ch	0.8924 ± 0.0001	31.2 ± 0.3
PFN	0.8911 ± 0.0008	30.8 ± 0.4
EFN	0.8824 ± 0.0005	28.6 ± 0.3
RNN-ID	0.9010	34.4
RNN	0.8899	30.5
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



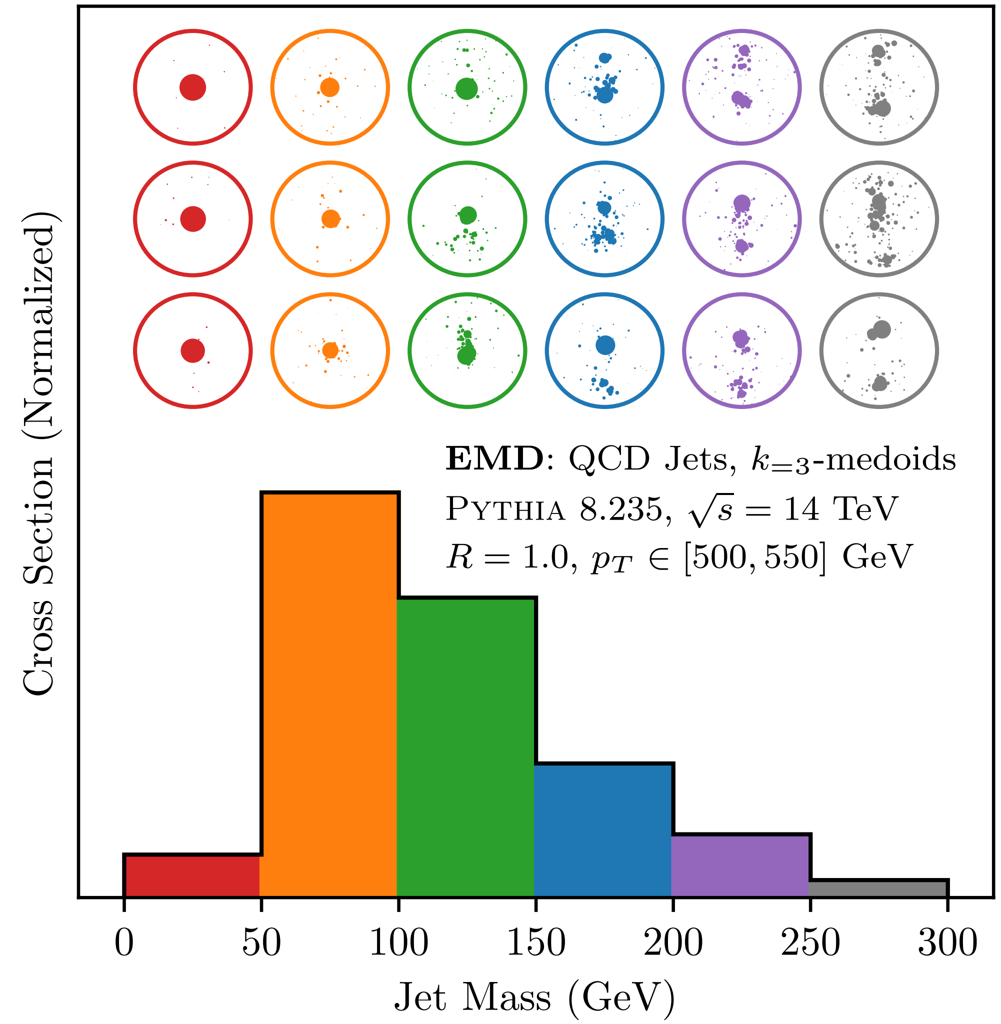


Opportunities for Network/Data Visualization

Latent Space



Metric Space



Augment (not replace) our exceptional human/scientific ability to recognize patterns

[Komiske, Metodiev, JDT, 1810.05165; Komiske, Metodiev, JDT, 1902.02346]