Energy Flow and Jet Substructure

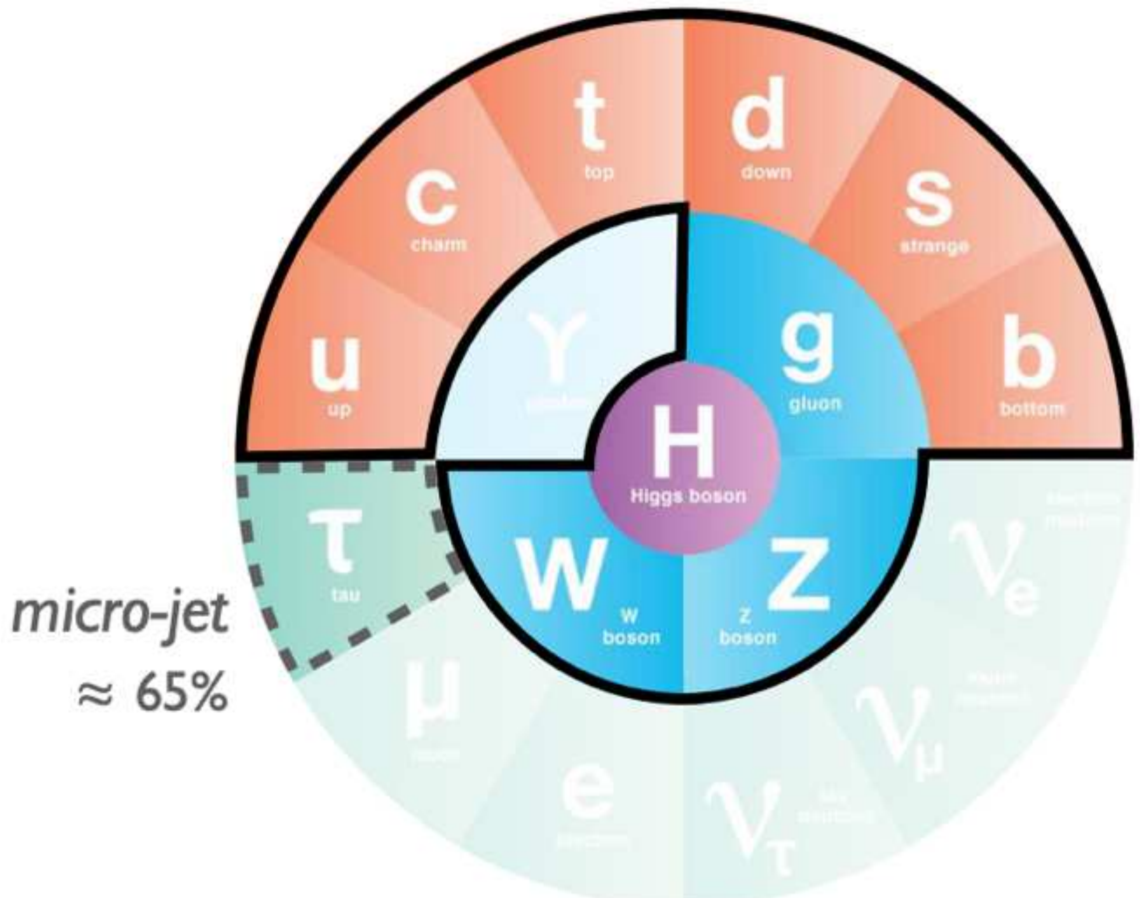
Patrick T. Komiske III
Massachusetts Institute of Technology
Center for Theoretical Physics

Harvard Particle Lunch Talk
11/28/2018

Based on work with Eric Metodiev and Jesse Thaler

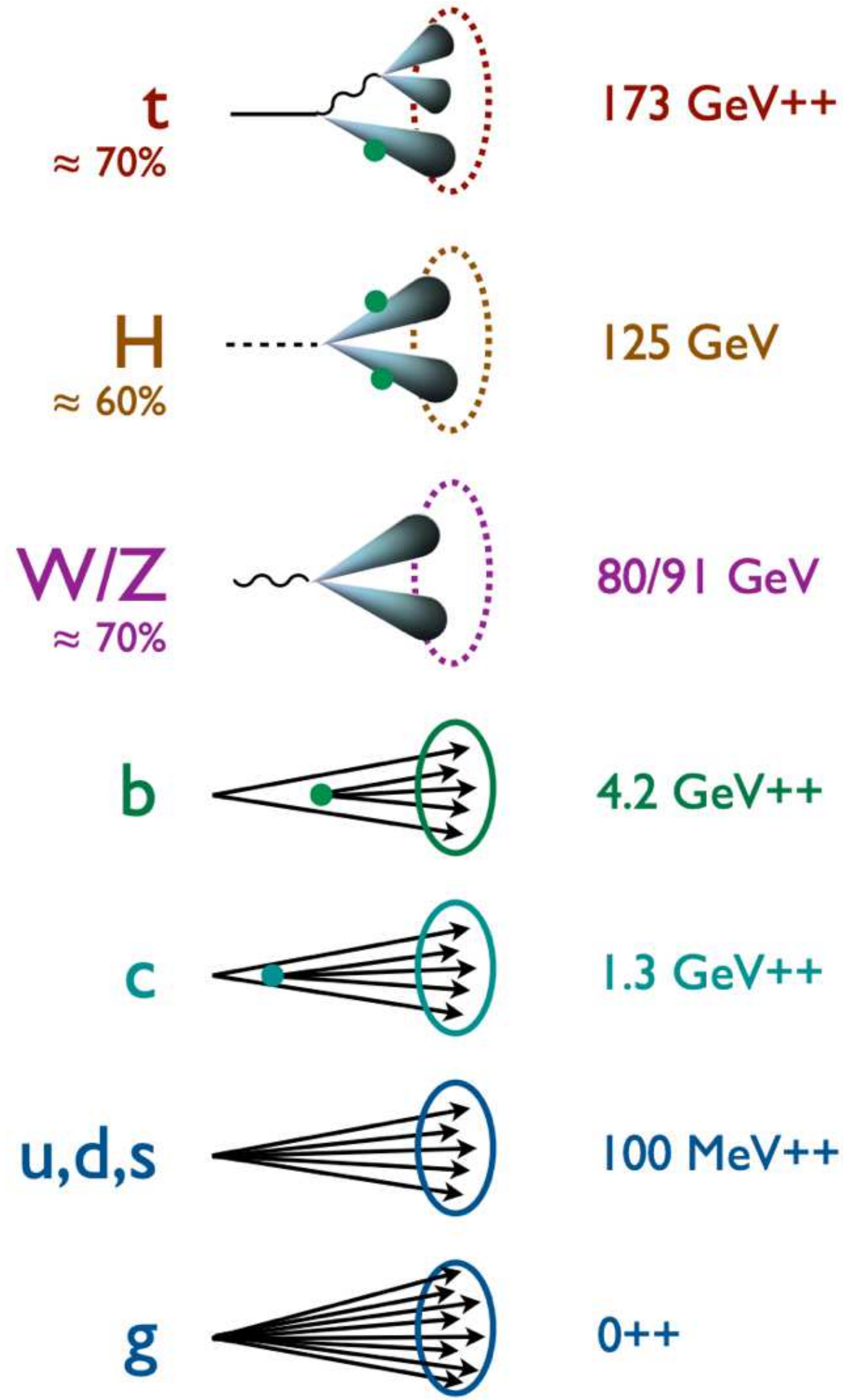
1712.07124 1810.05165

Jets in Theory

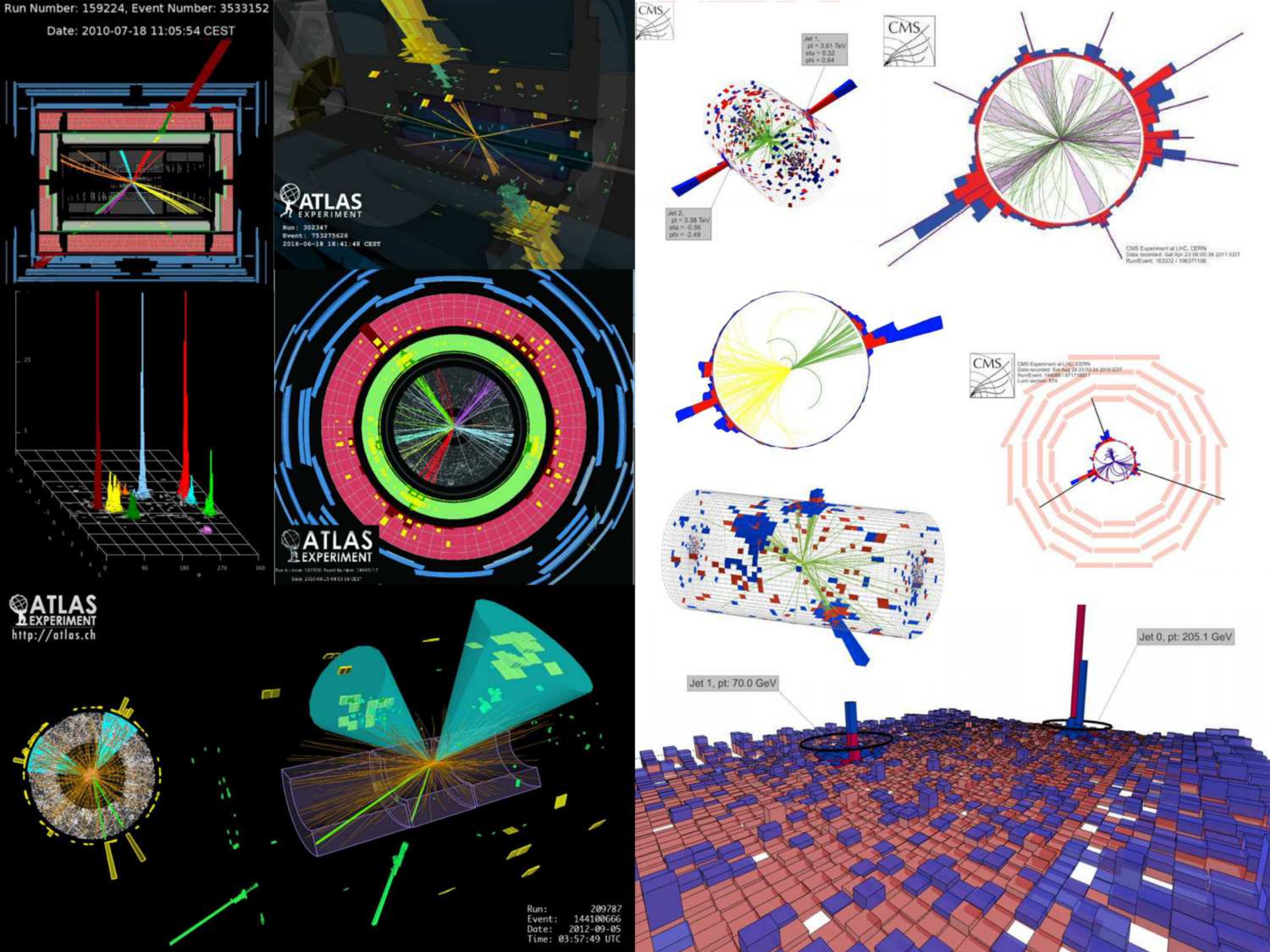


Jets from the Standard Model

++ = Mass from QCD Radiation

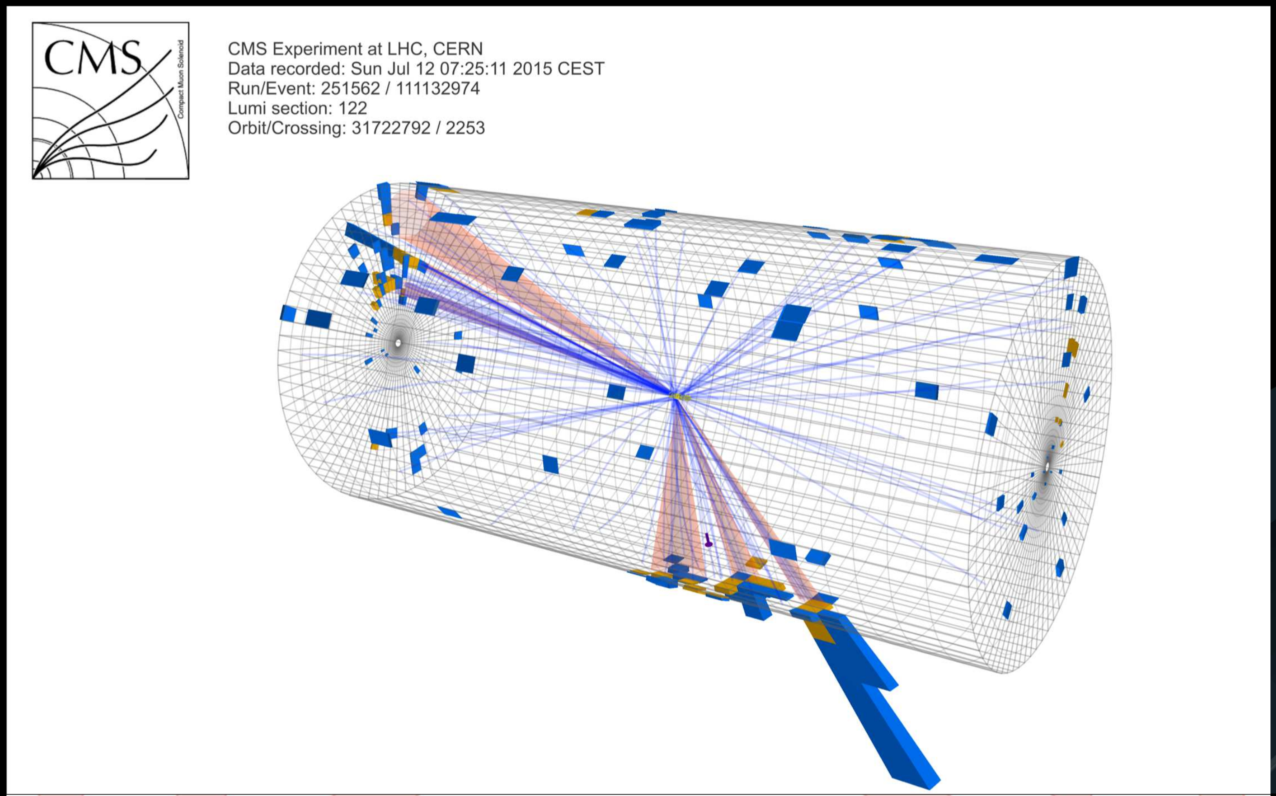


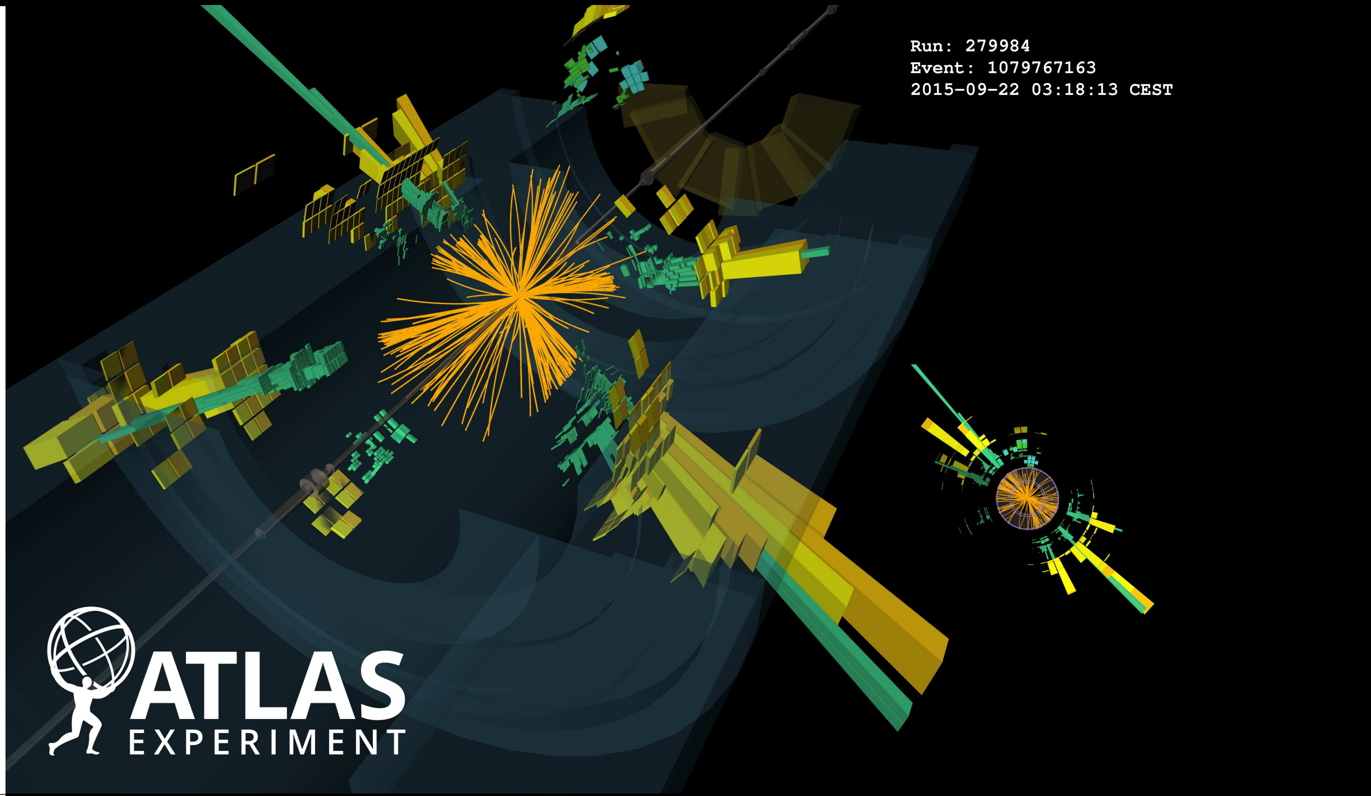
Slide by Jesse Thaler

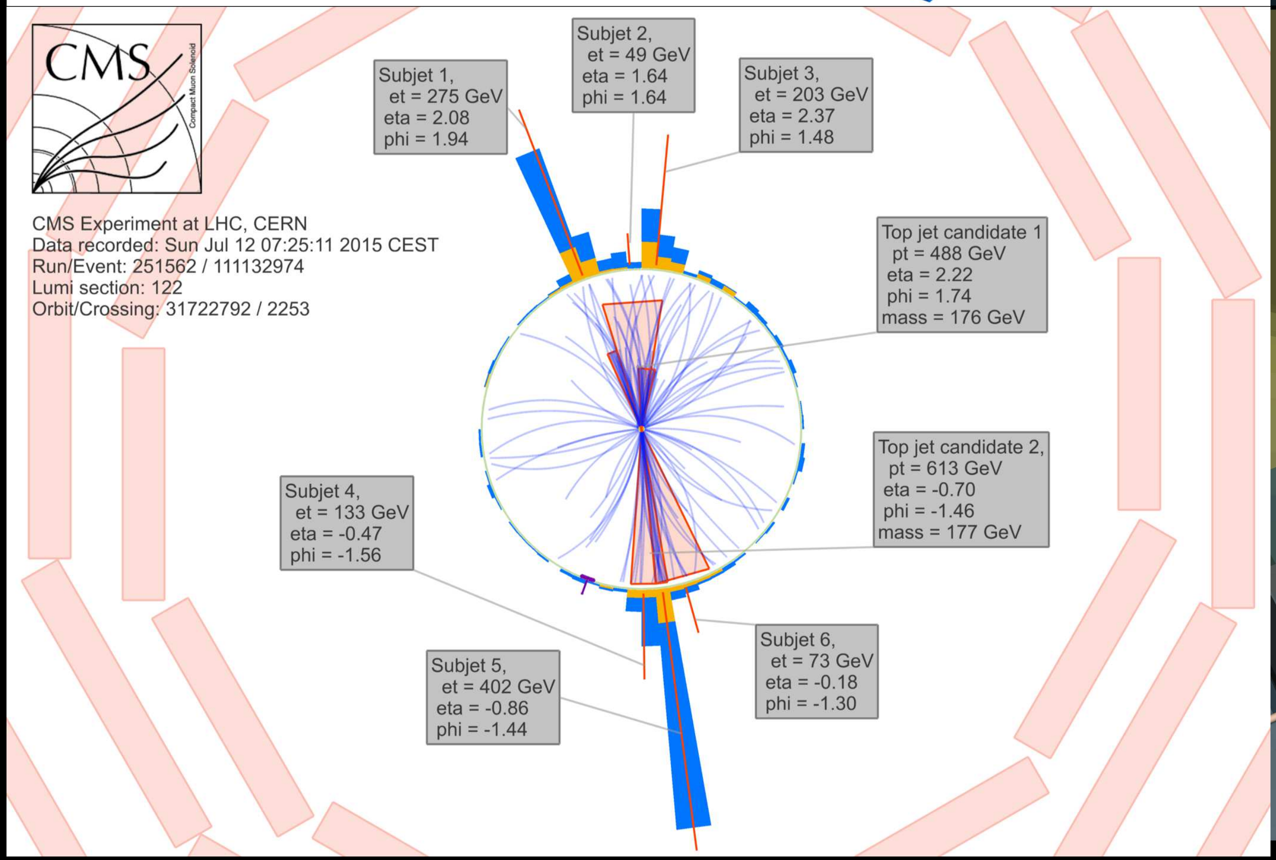


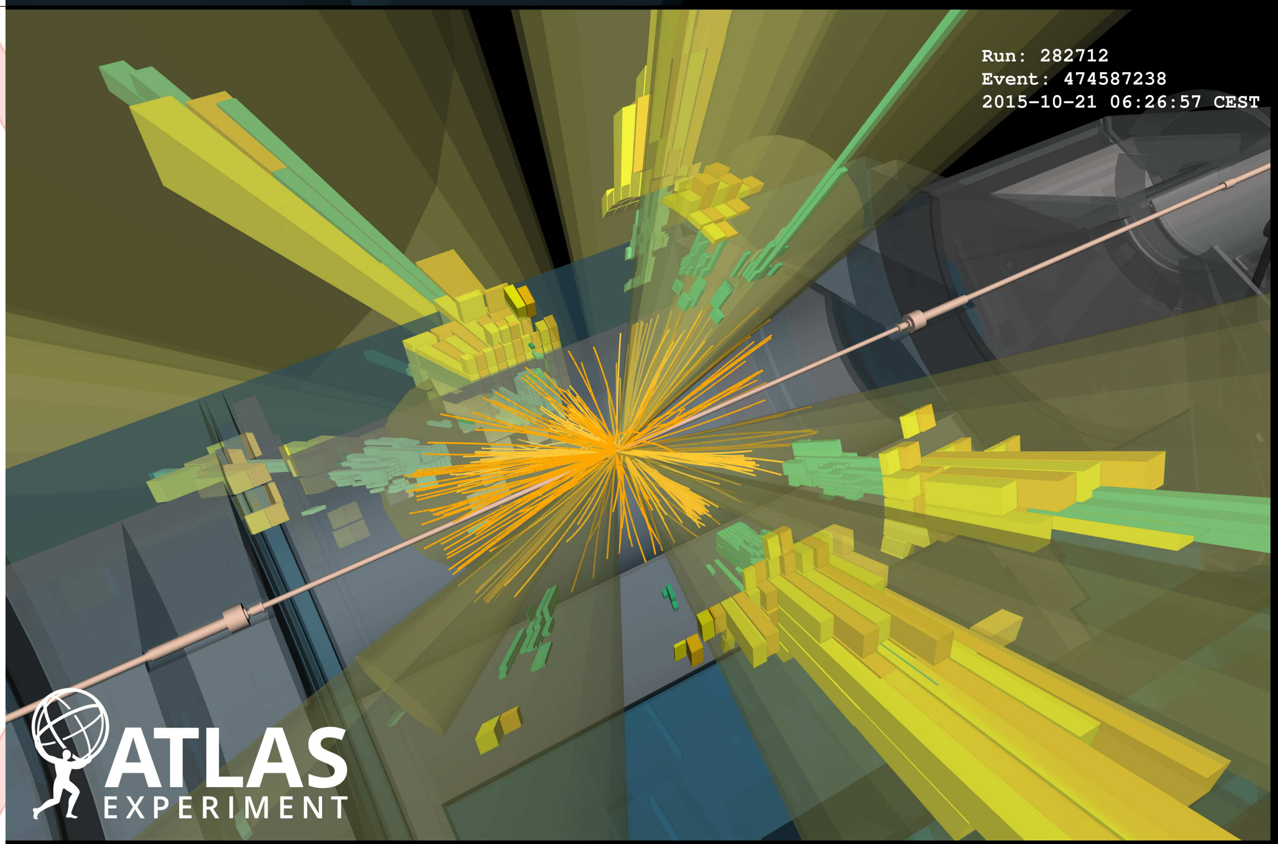
Run Number: 159224, Event Number: 3533152
Date: 2010-07-18 11:05:54 CEST
ATLAS EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST
CMS
Jet 1, pt = 3.61 TeV, eta = 0.32, phi = 0.64
Jet 2, pt = 3.38 TeV, eta = -0.56, phi = -2.49
CMS
CMS
ATLAS EXPERIMENT
ATLAS EXPERIMENT
http://atlas.ch
Run: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

Boosted Event Topologies at the LHC

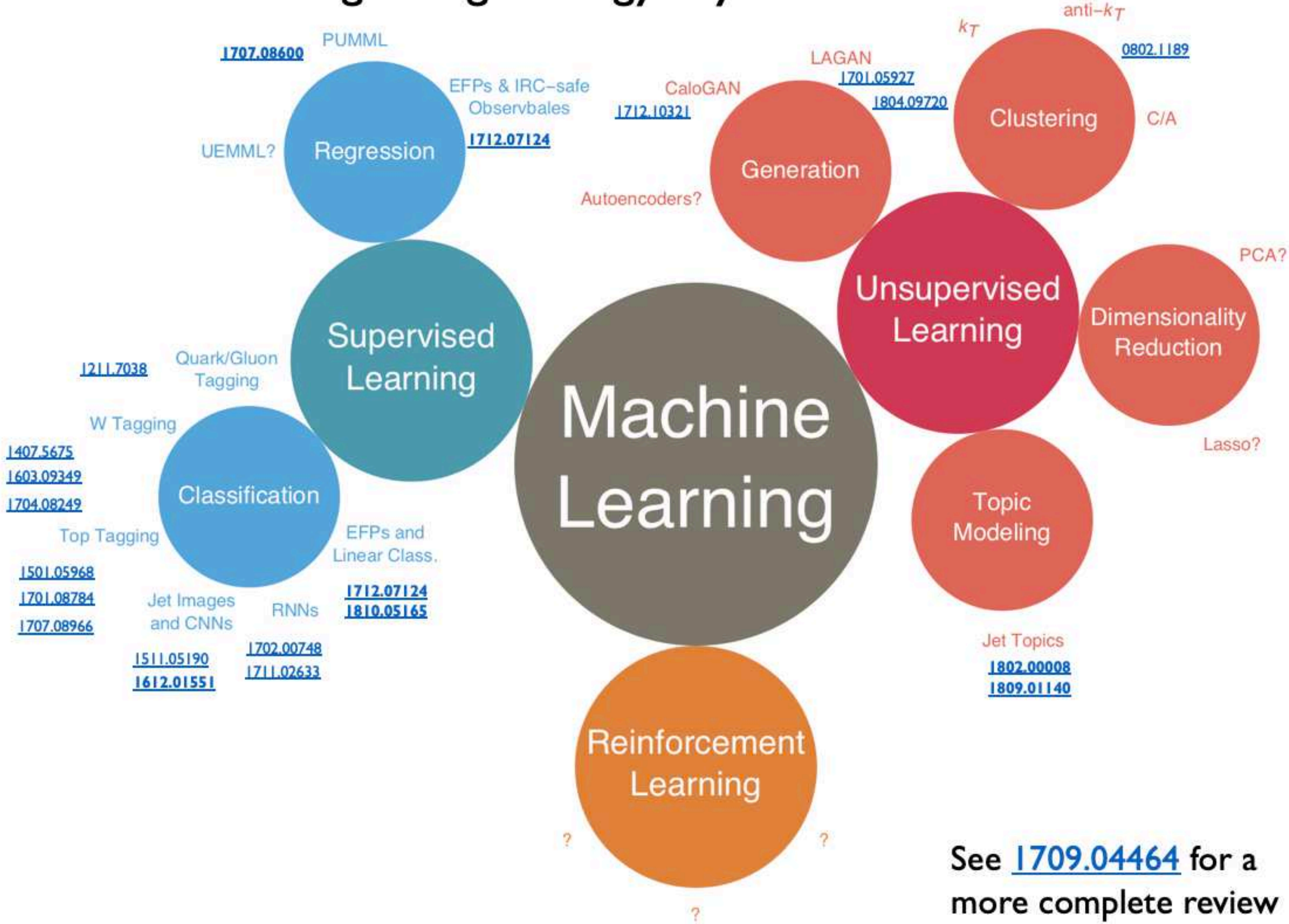








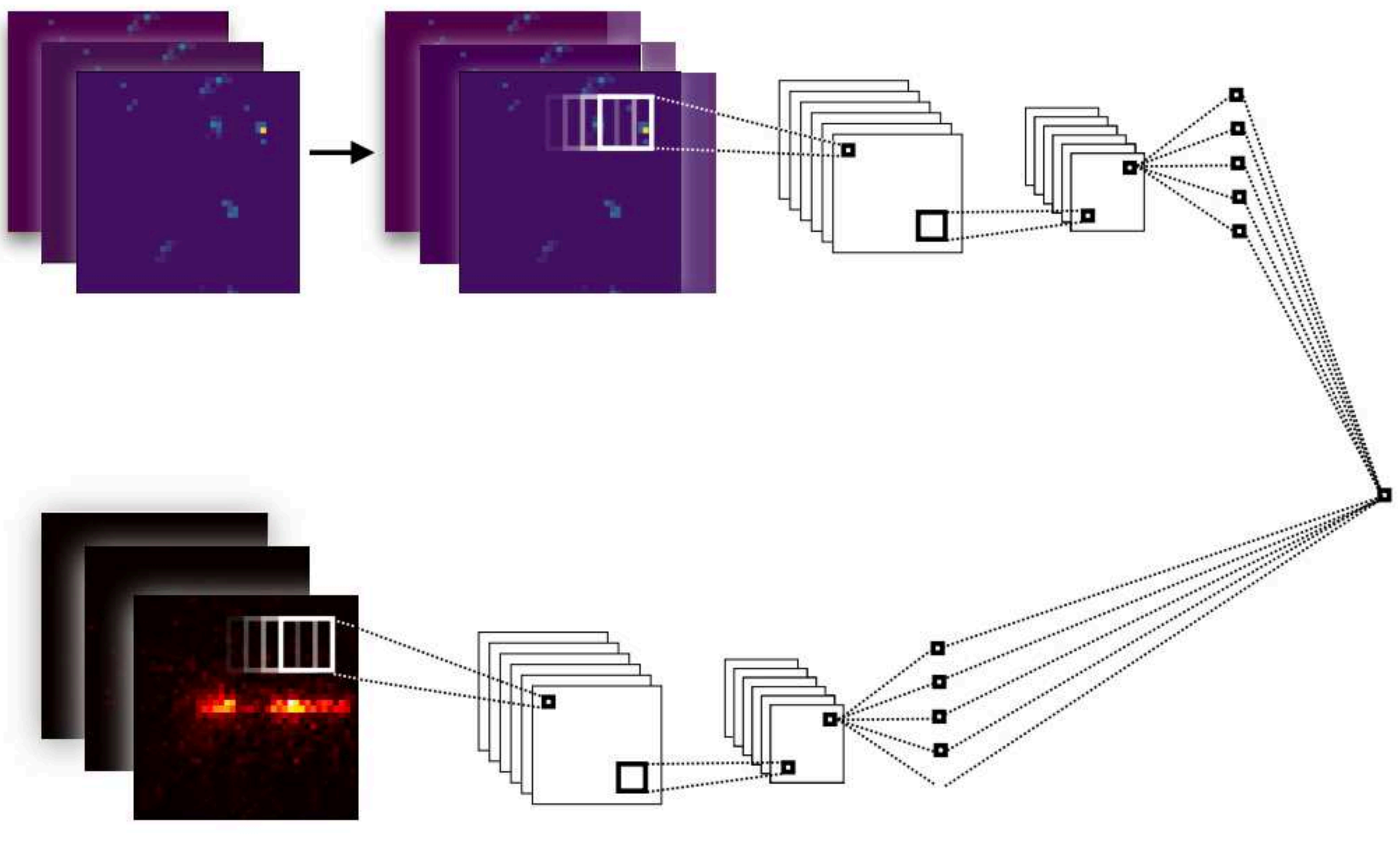
Machine Learning in High Energy Physics

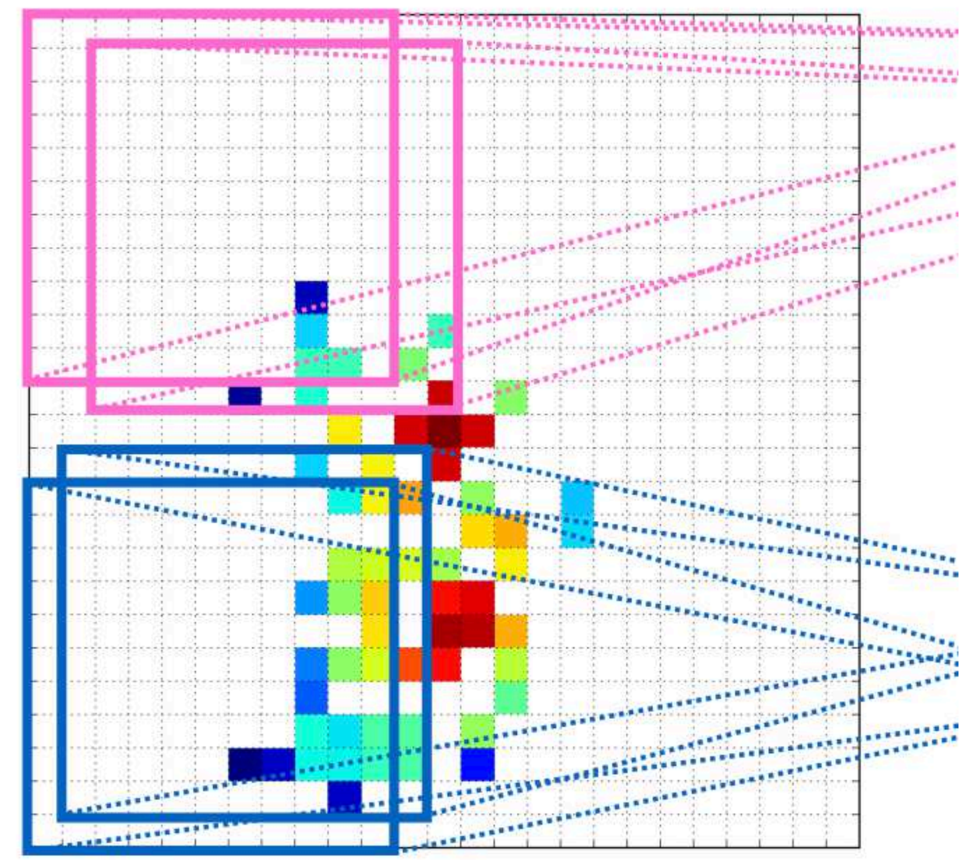


See 1709.04464 for a more complete review

Machine Learning for Jet Physics 2018

indico.cern.ch/event/ml4jets2018





Images: J. Lin, B. Nachman, L. de Oliveira

November 14-16, 2018

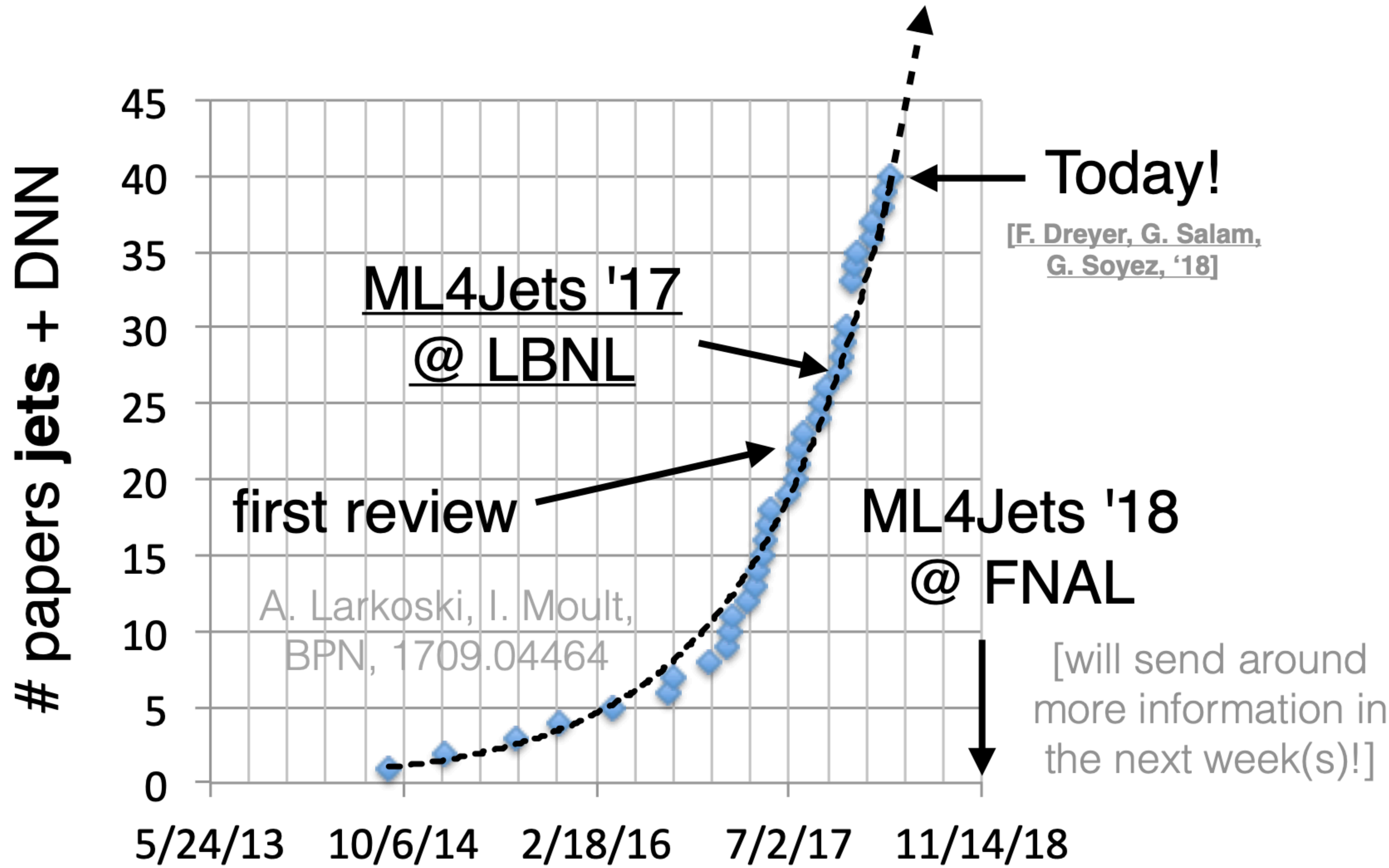
Fermilab

Organizing Committee:
Pushpa Bhat (Fermilab)
Kyle Cranmer (NYU)
Sergei Gleyzer (U Florida)
Ben Nachman (LBNL)
Tilman Plehn (Heidelberg)

Local Organizing Committee:
Gabriele Benelli (Brown U),
Javier Duarte (Fermilab)
Benjamin Kreis (Fermilab)
Nhan Tran (Fermilab)
Justin Pilot (UC Davis)

LPC Coordinators:
Cecilia Gerber (UIC)
Sergo Jindariani (Fermilab)

Machine Learning for Jets Activity



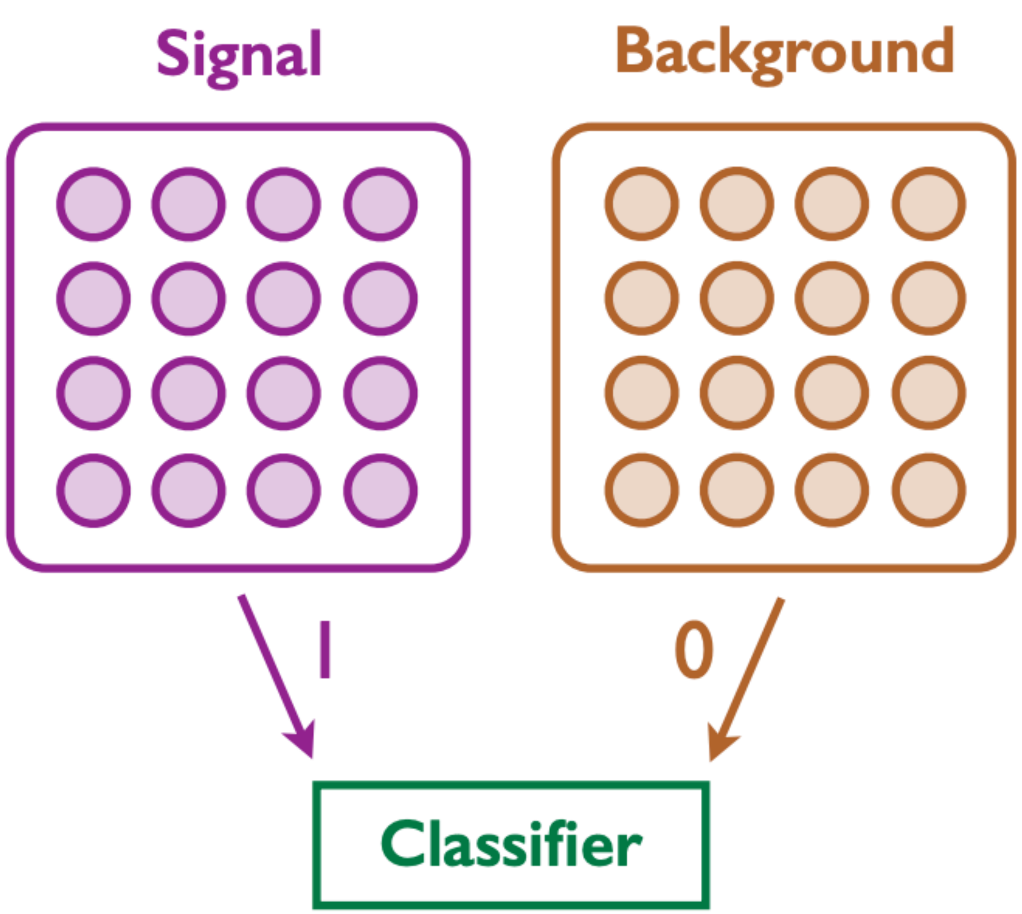
Exponential increase in ML4Jets papers!

[Nachman, Boost 2018 Talk, July 18, 2018]

A Cartoon of Machine Learning

For fully-supervised jet classification

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



Minimize Loss Function

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

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

Optimal Classifier (Neyman–Pearson)

Jet Classification Studies

Mix and match

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

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

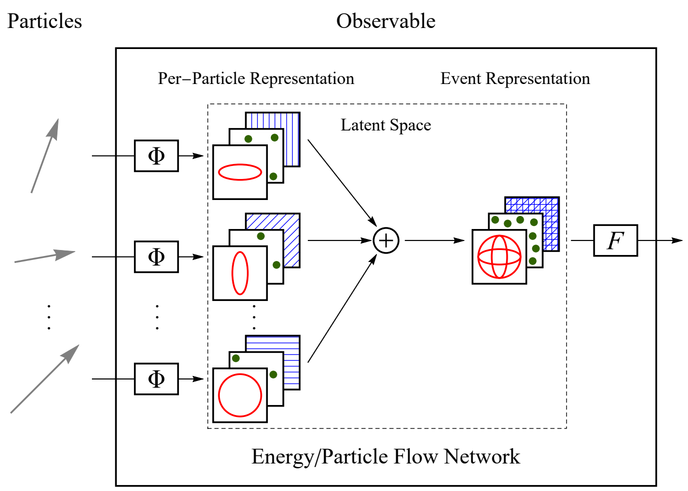
- High-Level Features
- Basis of High-Level Features
- Jet Image
- Multi-channel Jet Image
- Abstract Jet Image
- Sorted Four-Vectors
- Clustered Four-Vectors
- Lund Plane Emissions
- Kitchen Sink
- …

Signal vs. Background

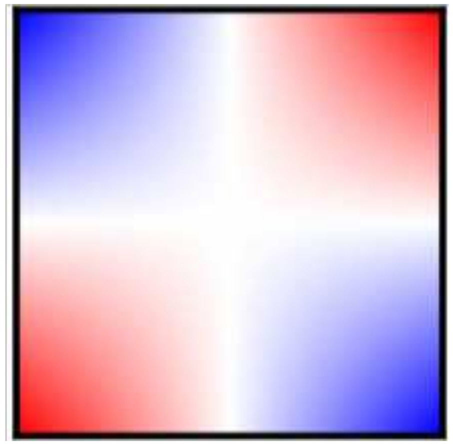
Signal	vs.	Background
Quark Jets	vs.	Gluon Jets
Up-type Quarks	vs.	Down-type Quarks
W/Z Bosons	vs.	QCD Jets
W Bosons	vs.	Z Bosons
Top Quarks	vs.	QCD Jets
Exotic Boosted Objects	vs.	QCD Jets
CMS Open Data Samples	vs.	Each other
…	vs.	…

[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]

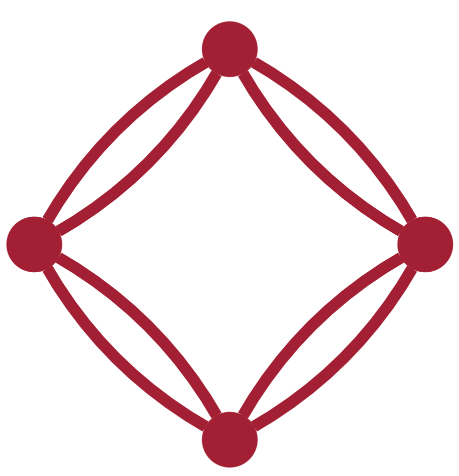
Slide by Jesse Thaler



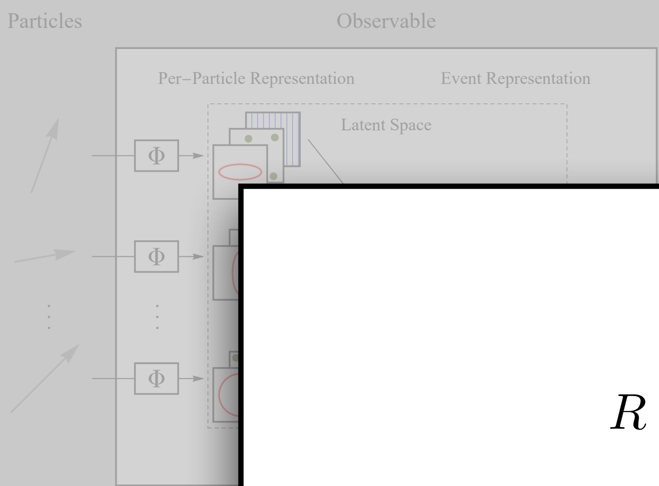
Energy Flow Networks

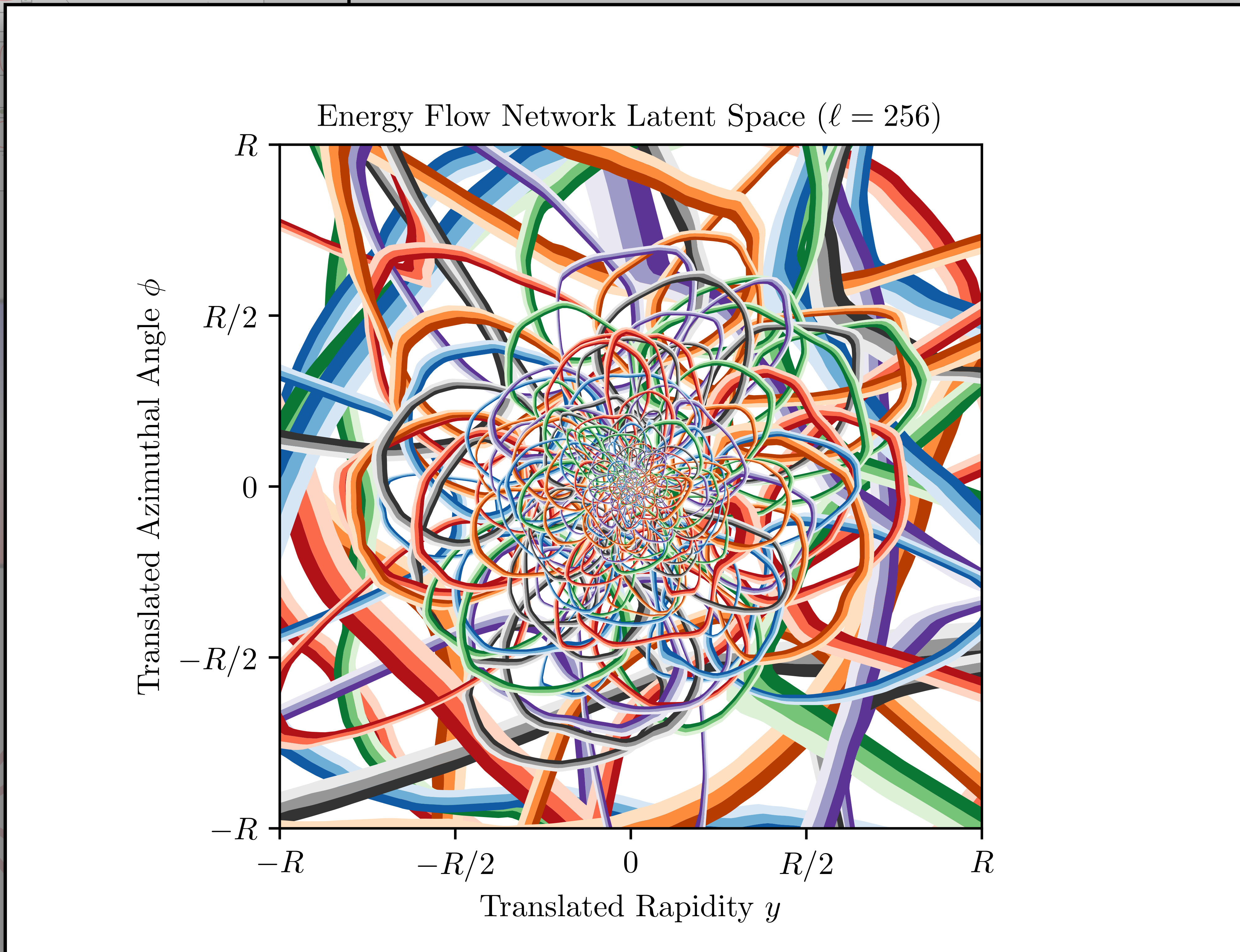


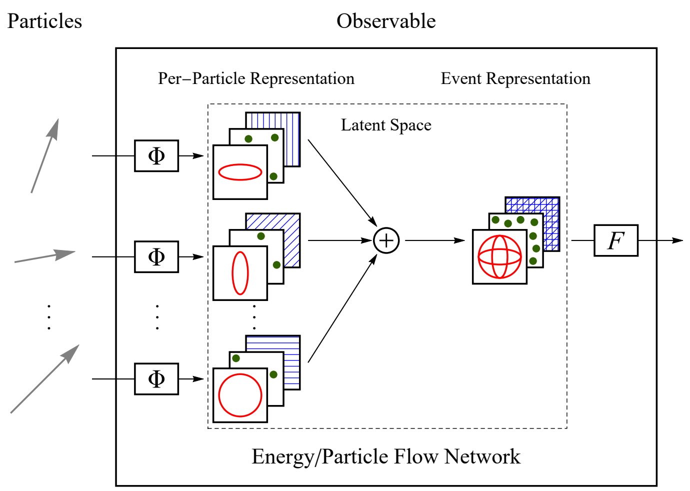
Energy Flow Moments



Energy Flow Polynomials


Particles
Observable
Per−Particle Representation
Event Representation
Latent Space
Φ
Φ
Φ


Energy Flow Network Latent Space ($\ell = 256$)
Translated Azimuthal Angle ϕ
R
$R/2$
0
$-R/2$
$-R$
$-R$
$-R/2$
0
$R/2$
R
Translated Rapidity y



Energy Flow Networks

Energy Flow Moments

Energy Flow Polynomials

What is a Jet?

*An **unordered**, **variable length** collection of particles*

Due to quantum-mechanical indistinguishability
Due to probabilistic nature of jet formation

$$J\left(\{p_1^\mu,\ldots,p_M^\mu\}\right) = J(\{p_{\pi(1)}^\mu,\ldots,p_{\pi(M)}^\mu\}), \qquad \underbrace{M\geq 1}_{\text{Multiplicity}}\ , \qquad \underbrace{\forall\,\pi\in S_M}_{\text{Permutations}}$$

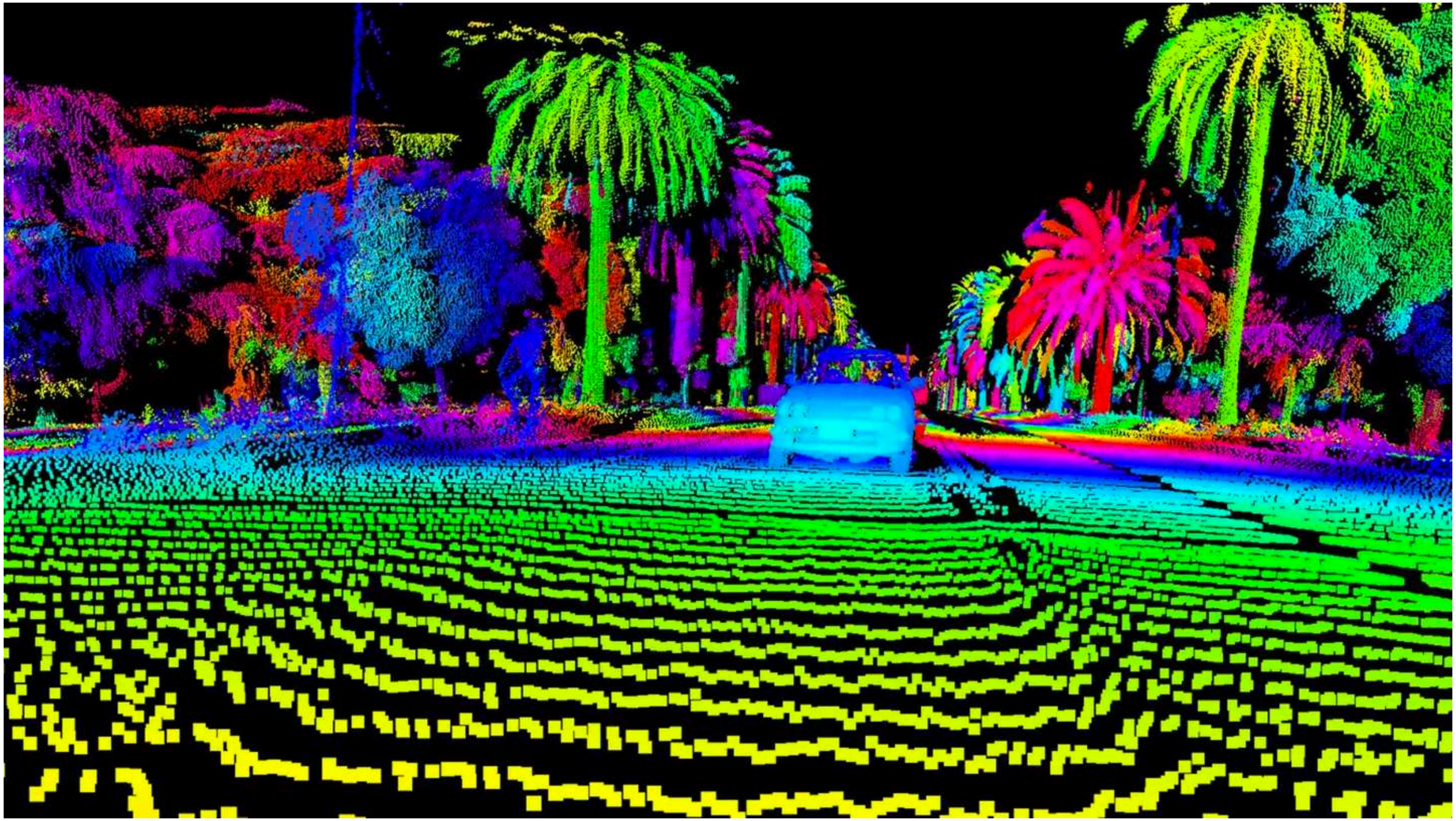
p_i^μ represents *all* the particle properties:

- Four-momentum – $(E, p_x, p_y, p_z)_i^\mu$
- Other quantum numbers (e.g. particle id, charge)
- Experimental information (e.g. vertex info, quality criteria, PUPPI weights)

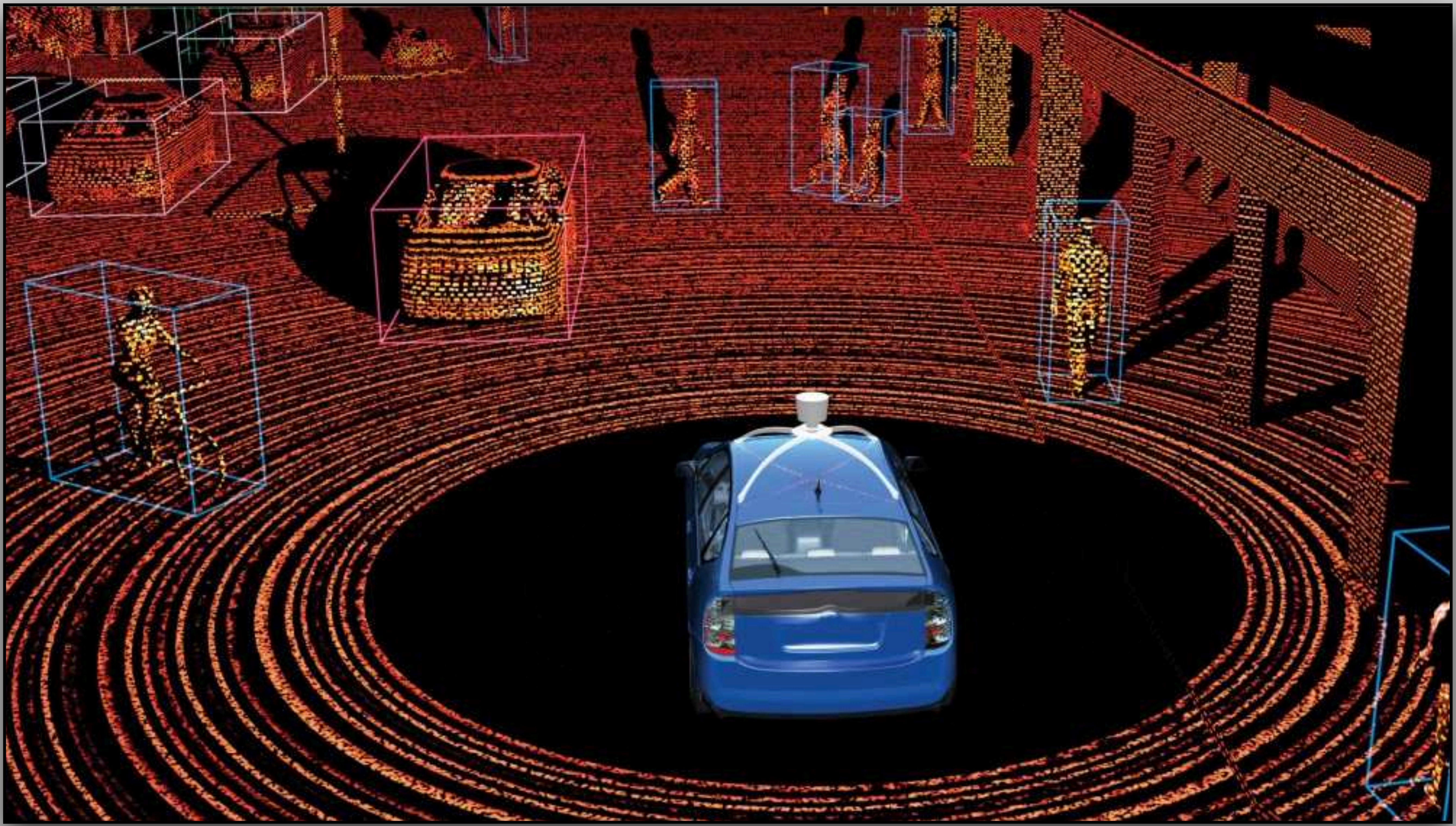
Point Clouds

Point cloud: "A set of data points in space" –Wikipedia

LIDAR data from self-driving car sensor



Point Clouds



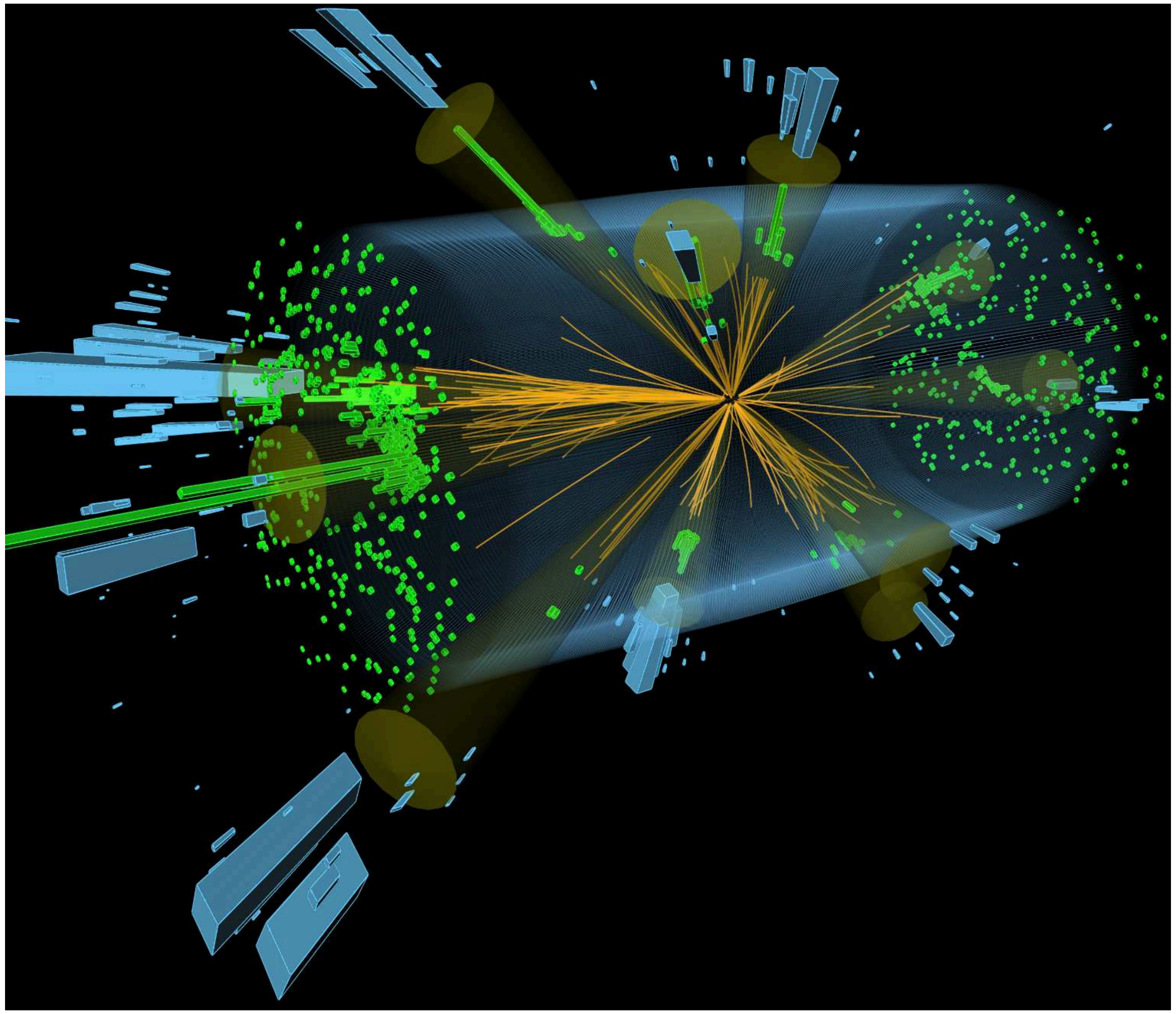
Particle Collision Events as Point Clouds

Point cloud: "A set of data points in space" –Wikipedia

Jet/event

Particles

Feature space



Multi-jet event at CMS

Processing Point Clouds

Methods for processing point clouds/jets should respect the appropriate symmetries

Variable constituent multiplicity requires at least one of:

Preprocessing to another representation (jet images, *N*-subjettiness, etc.)

Truncation to an (arbitrary) fixed size

Recurrent NN structure

Particle permutation symmetry requires:

Permutation symmetric observables

Permutation symmetric architectures

Jet Representations ⟷ Analysis Tools

Two key choices when analyzing jets

<u>How to represent the jet</u>

- Single expert observable
- A few expert observables
- Many expert observables

- Jet images
- List of particles
- Clustering tree
- *N*-subjettiness basis
- Energy flow polynomials
- Set of particles

<u>How to analyze that representation</u>

- Threshold cut
- Multidimensional likelihood
- Boosted decision tree (BDT), shallow neural network (NN)
- Convolutional NN (CNN)
- Recurrent/Recursive NN (RNN)
- Fancy RNN
- Dense neural network (DNN)
- Linear classification
- Energy flow network

Jet Representations ⟷ Analysis Tools

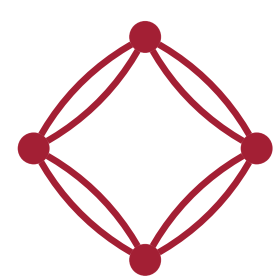
Two key choices when analyzing jets

How to represent the jet

- Single expert observable
- A few expert observables
- Many expert observables
- Jet images
- List of particles
- Clustering tree
- *N*-subjettiness basis
- Energy flow polynomials
- Set of particles

How to analyze that representation

- Threshold cut
- Multidimensional likelihood
- Boosted decision tree (BDT), shallow neural network (NN)
- Convolutional NN (CNN)
- Recurrent/Recursive NN (RNN)
- Fancy RNN
- Dense neural network (DNN)
- Linear classification
- Energy flow network



Jet Representations ⟷ Analysis Tools

Two key choices when analyzing jets

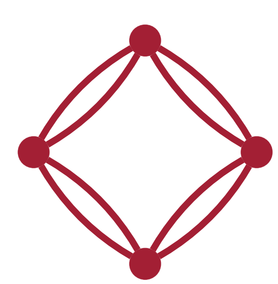
How to represent the jet

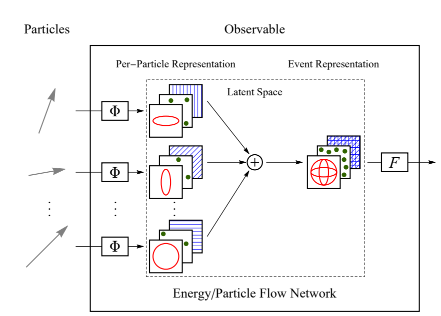
- Single expert observable
- A few expert observables
- Many expert observables

- Jet images
- List of particles
- Clustering tree
- *N*-subjettiness basis
- Energy flow polynomials
- Set of particles

How to analyze that representation

- Threshold cut
- Multidimensional likelihood
- Boosted decision tree (BDT), shallow neural network (NN)
- Convolutional NN (CNN)
- Recurrent/Recursive NN (RNN)
- Fancy RNN
- Dense neural network (DNN)
- Linear classification
- Energy flow network





Deep Sets

Namespace for symmetric function parametrization

A general permutation-symmetric function is *additive* in a latent space

Deep Sets

[1703.06114]

Manzil Zaheer[1,2], **Satwik Kottur**[1], **Siamak Ravanbhakhsh**[1],
Barnabás Póczos[1], **Ruslan Salakhutdinov**[1], **Alexander J Smola**[1,2]
[1] Carnegie Mellon University [2] Amazon Web Services

Deep Sets Theorem [63]. *Let $\mathfrak{X} \subset \mathbb{R}^d$ be compact, $X \subset 2^{\mathfrak{X}}$ be the space of sets with bounded cardinality of elements in $\mathfrak{X}$, and $Y \subset \mathbb{R}$ be a bounded interval. Consider a continuous function $f : X \to Y$ that is invariant under permutations of its inputs, i.e. $f(x_1, \ldots, x_M) = f(x_{\pi(1)}, \ldots, x_{\pi(M)})$ for all $x_i \in \mathfrak{X}$ and $\pi \in S_M$. Then there exists a sufficiently large integer ℓ and continuous functions $\Phi : \mathfrak{X} \to \mathbb{R}^\ell$, $F : \mathbb{R}^\ell \to Y$ such that the following holds to an arbitrarily good approximation:*[1]

$$f(\{x_1, \ldots, x_M\}) = F\left(\sum_{i=1}^{M} \Phi(x_i)\right)$$

Deep Sets

Namespace for symmetric function parametrization

A general permutation-symmetric function is *additive* in a latent space

Deep Sets

[1703.06114]

Manzil Zaheer[1,2], Satwik Kottur[1], Siamak Ravanbhakhsh[1], Barnabás Póczos[1], Ruslan Salakhutdinov[1], Alexander J Smola[1,2]
[1] Carnegie Mellon University [2] Amazon Web Services

Feature space

Variable length

Deep Sets Theorem [63]. *Let $\boxed{\mathfrak{X} \subset \mathbb{R}^d}$ be compact, $\boxed{X \subset 2^{\mathfrak{X}} \text{ be the space of sets}}$ with bounded cardinality of elements in $\mathfrak{X}$, and $Y \subset \mathbb{R}$ be a bounded interval. Consider a continuous function $f : X \to Y$ that is invariant under permutations of its inputs, i.e. $f(x_1, \ldots, x_M) = \boxed{f(x_{\pi(1)}, \ldots, x_{\pi(M)}) \text{ for all } x_i \in \mathfrak{X} \text{ and } \pi \in S_M}$. Then there exists a sufficiently large integer ℓ and continuous functions $\Phi : \mathfrak{X} \to \boxed{\mathbb{R}^\ell}$, $F : \boxed{\mathbb{R}^\ell} \to Y$ such that the following holds to an arbitrarily good approximation:*[1]

Permutation invariance

Latent space

$$\boxed{f(\{x_1, \ldots, x_M\}) = F\left(\sum_{i=1}^{M} \Phi(x_i)\right)}$$

General parametrization for a function of sets

Deep Sets for Particle Jets

[PTK, Metodiev, Thaler, 1810.05165]

Particle Flow Network (PFN)

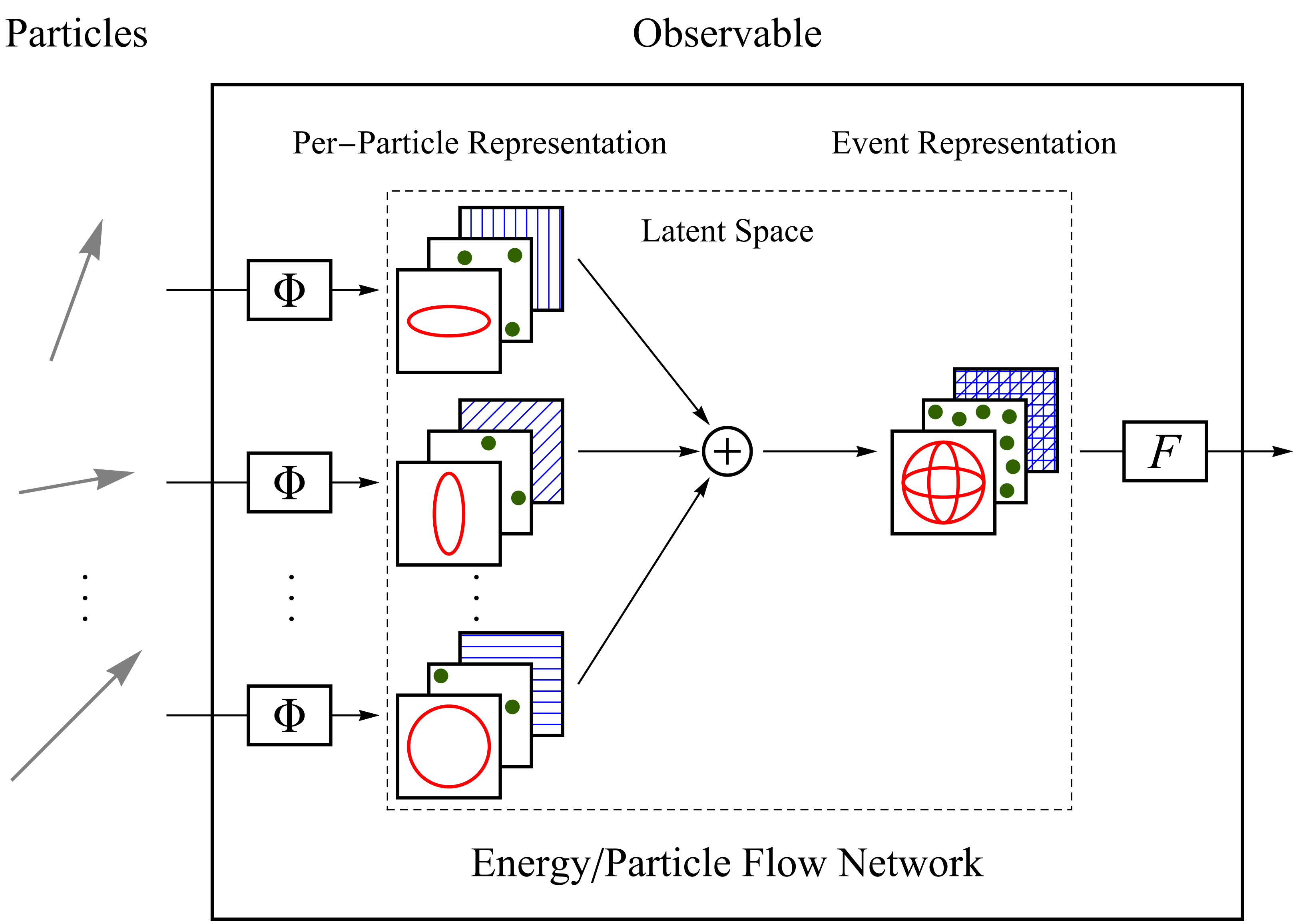
$$\mathrm{PFN}(\{p_1^\mu,\ldots,p_M^\mu\}) = F\left(\sum_{i=1}^{M}\Phi(p_i^\mu)\right)$$

Fully general latent space

Energy Flow Network (EFN)

$$\mathrm{EFN}(\{p_1^\mu,\ldots,p_M^\mu\}) = F\left(\sum_{i=1}^{M} z_i\Phi(\hat{p}_i)\right)$$

IRC-safe latent space



Latent Space IRC Safety

Latent space defines new physics observables

IRC safety is a key theoretical *and experimental* property of observables

QCD has soft and collinear divergences associated with gluon radiation

$$dP_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_a \frac{d\theta}{\theta} \frac{dz}{z}$$

$$C_q = C_F = 4/3$$
$$C_g = C_A = 3$$

Infrared (IR) safety – observable is unchanged under addition of a soft particle

$$S(\{p_1^\mu, \ldots, p_M^\mu\}) = \lim_{\epsilon\to 0} S(\{p_1^\mu, \ldots, p_M^\mu, p_{M+1}^\mu\}), \quad \forall\, p_{M+1}^\mu$$

Collinear (C) safety – observable is unchanged under a collinear splitting of a particle

$$S(\{p_1^\mu, \ldots, p_M^\mu\}) = S(\{p_1^\mu, \ldots, (1-\lambda)p_M^\mu, \lambda p_{M+1}^\mu\}), \quad \forall\, \lambda \in [0,1]$$

Latent Space IRC Safety

Latent

IRC saf

QCD h

Infrare

Colline

IRC safety is a statement of *linearity* in energy and *continuity* in geometry

Theorem: A generic function of four-momenta can be made IRC safe via the following replacement:

$$\sum_{i=1}^{M} f(p_i^\mu) \longrightarrow \sum_{i=1}^{M} z_i f(\hat{p}_i).$$

Proof: In 1810.05165. $\square$

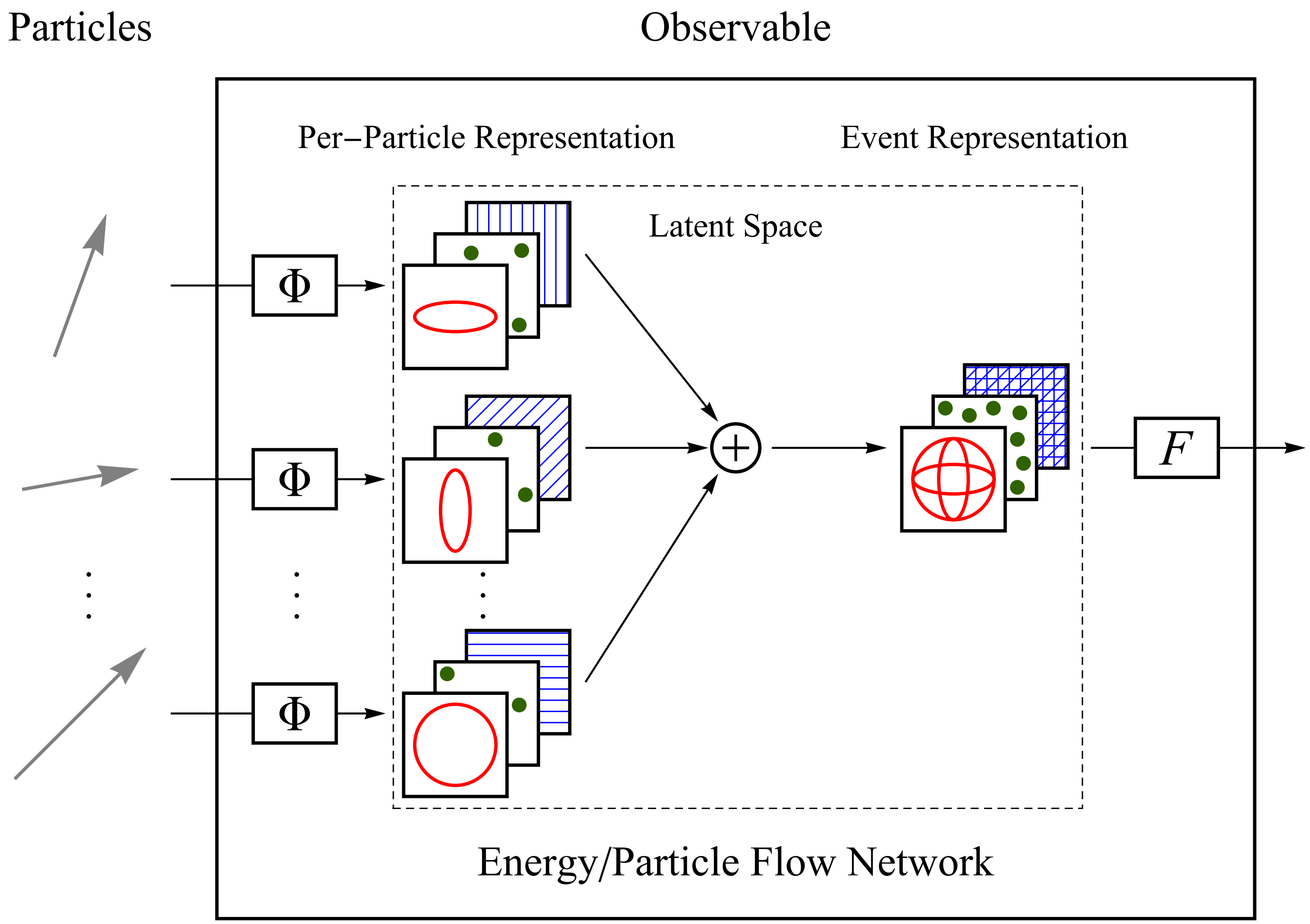
$$S(\{p_1^\mu, \ldots, p_M^\mu\}) = S(\{p_1^\mu, \ldots, (1-\lambda)p_M^\mu, \lambda p_{M+1}^\mu\}), \quad \forall\, \lambda \in [0,1]$$

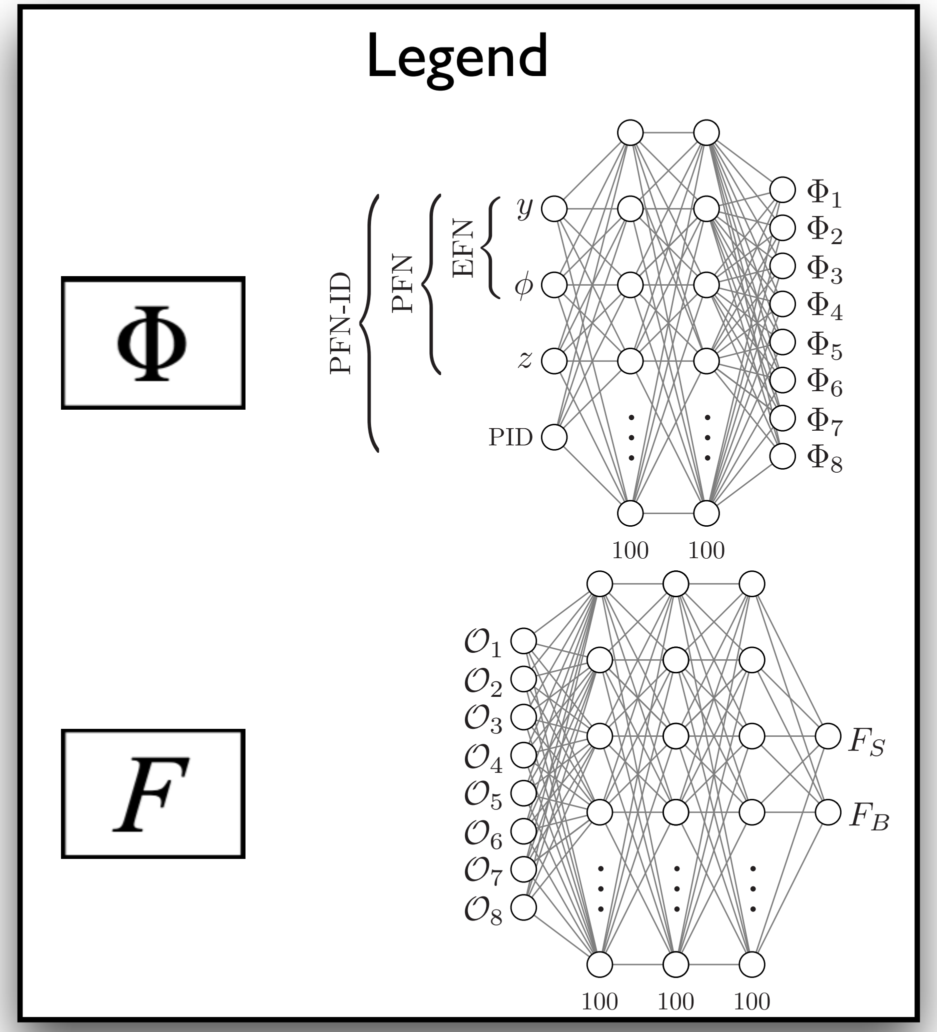
Approximating Φ and *F* with Neural Networks

Employ neural networks as arbitrary function approximators

Use fully-connected networks for simplicity

Default sizes – Φ: (100, 100, ℓ), F: (100, 100, 100)



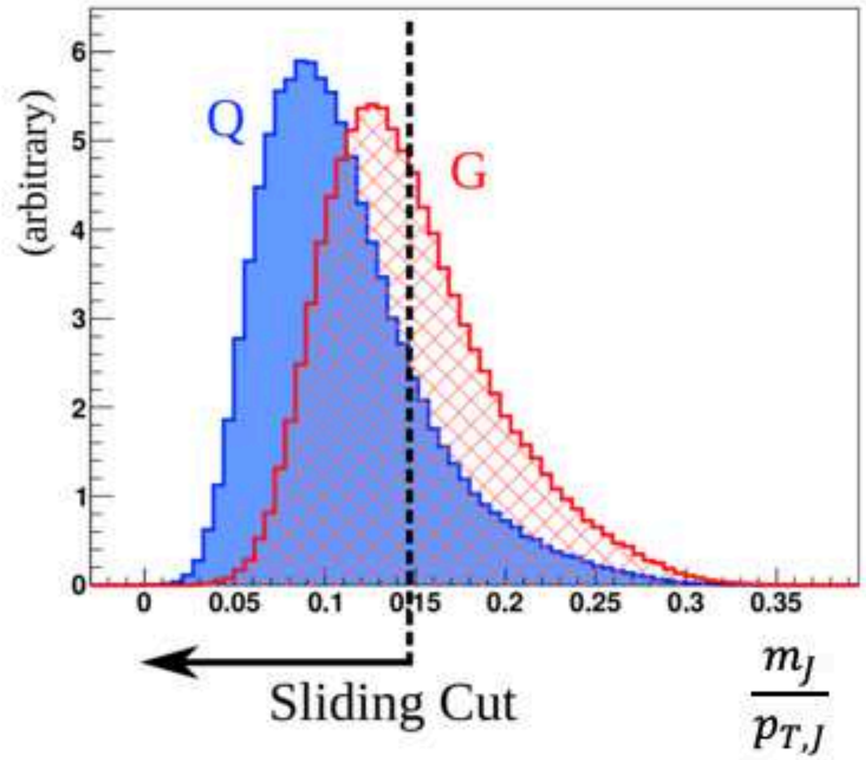


$$\text{EFN}: \quad \mathcal{O}_a = \sum_{i=1}^{M} z_i \Phi_a(y_i, \phi_i)$$

$$\text{PFN}: \quad \mathcal{O}_a = \sum_{i=1}^{M} \Phi_a(z_i, y_i, \phi_i, [\text{PID}_i])$$

Quantifying a Classifier

Receiver Operating Characteristic (**ROC**) curve:
True negative rate of the classifier at different true positive rates



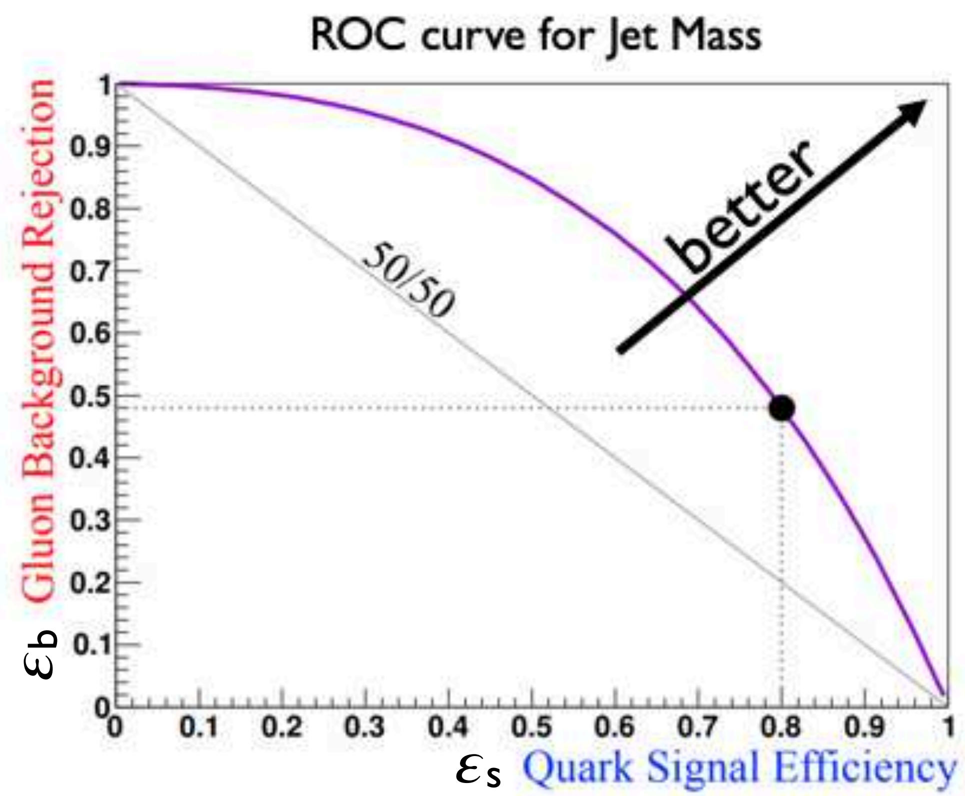


Figure from 1211.7038

Area Under the ROC Curve (**AUC**) captures the classifier performance in a number.

Other formats possible, e.g. $(\varepsilon_s,\ 1/(1-\varepsilon_b)),\ \ (\varepsilon_s,\ \varepsilon_s/\sqrt{1-\varepsilon_b})$

Classification Performance

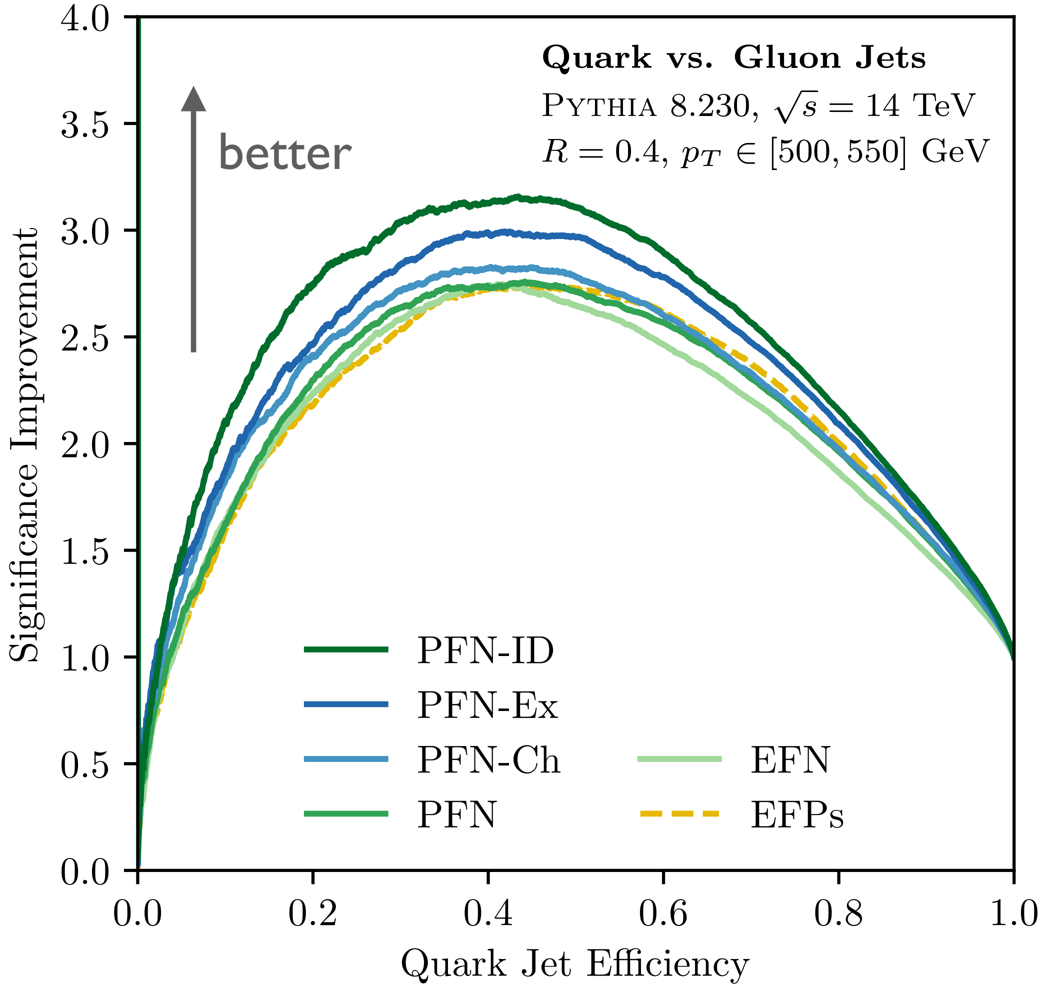
PFN-ID: Full particle flavor info
$(\gamma, \pi^\pm, K^\pm, K_L, p, \bar{p}, n, \bar{n}, e^\pm, \mu^\pm)$

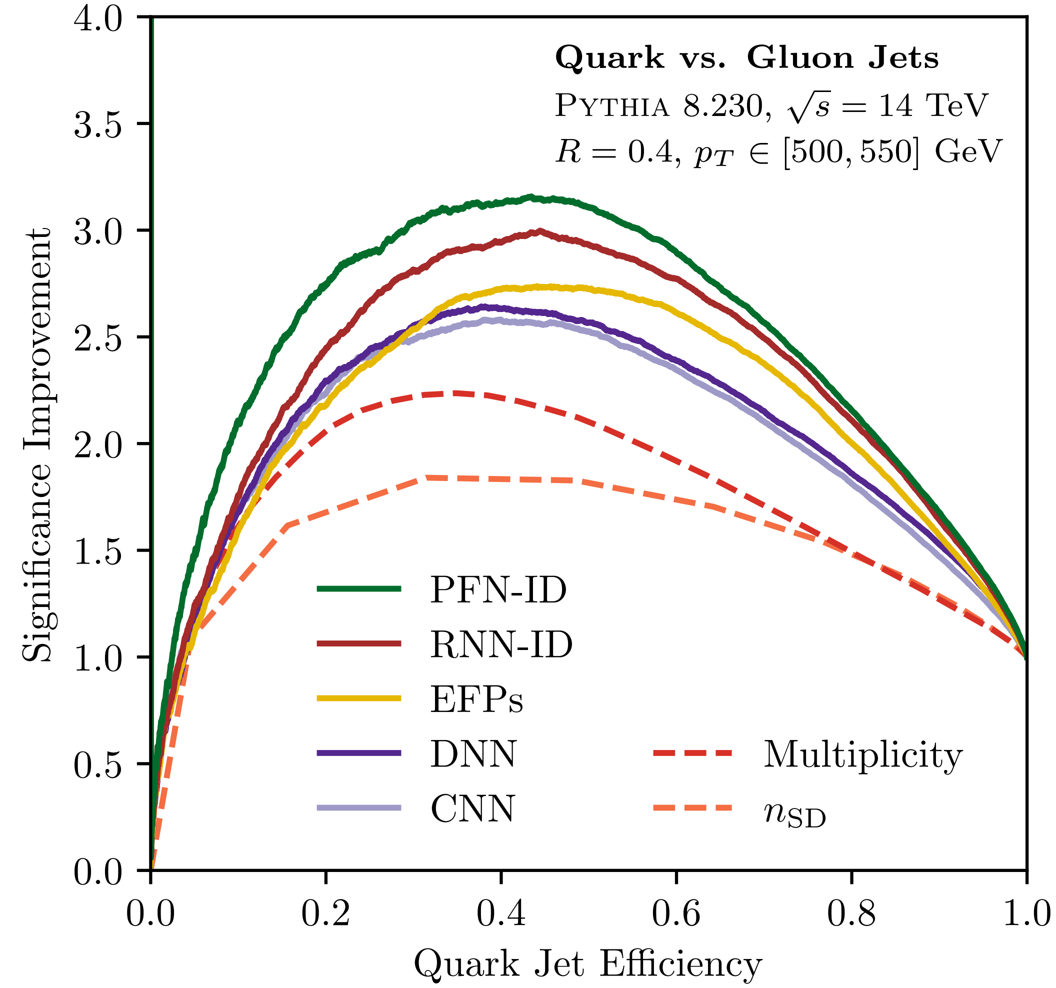
PFN-Ex: Experimentally accessible info
$(\gamma, h^{\pm,0}, e^\pm, \mu^\pm)$

PFN-Ch: Particle charge info
(+,0,–)

PFN: No particle type info, arbitrary energy dependence

EFN: IRC-safe latent space





Latent space dimension ℓ = 256

PFN-ID slightly better than RNN-ID

EFPs are comparable to EFN

EFN Latent Dimension Sweep

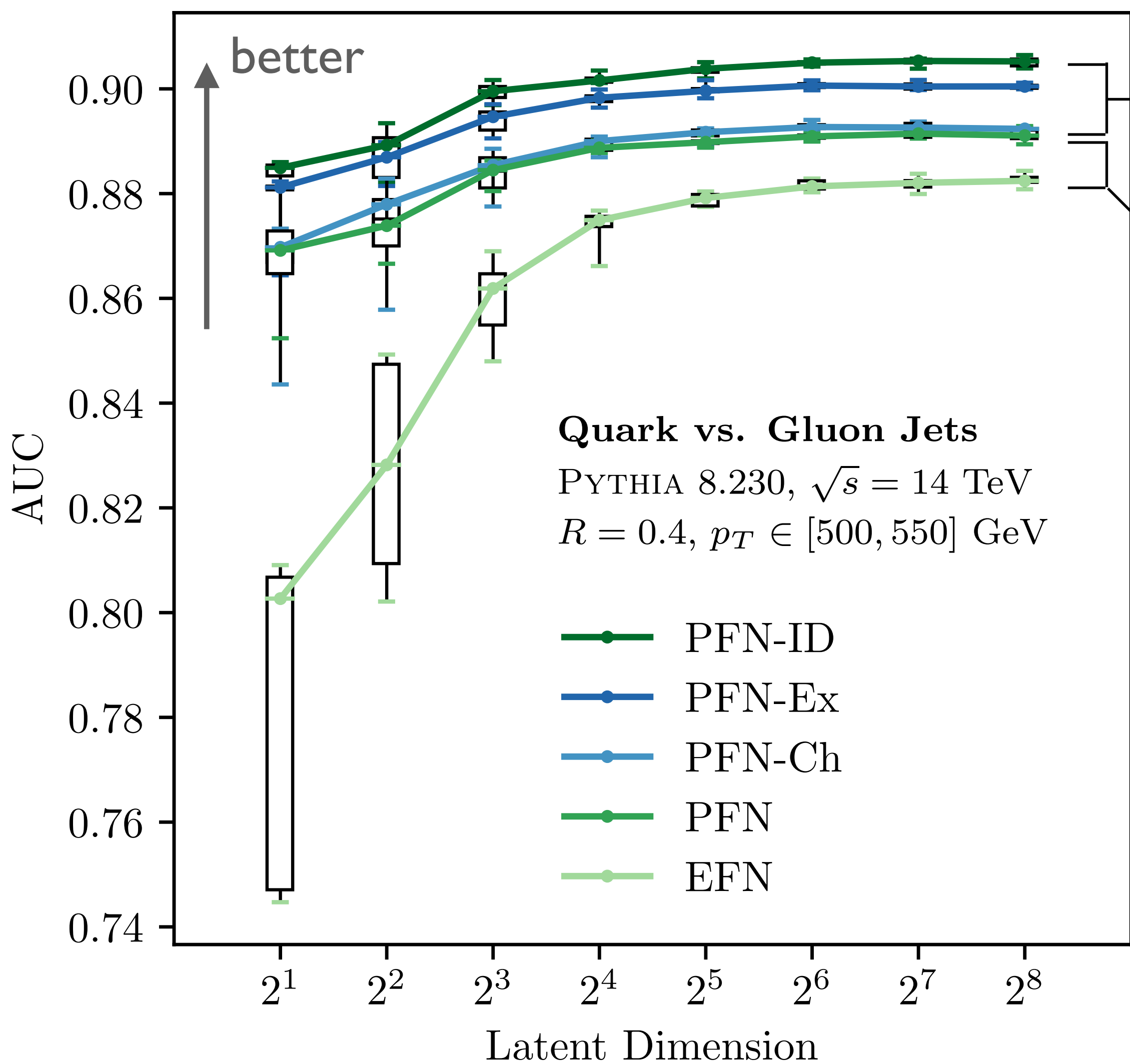
PFN-ID: Full particle flavor info
$(\gamma, \pi^{\pm}, K^{\pm}, K_L, p, \bar{p}, n, \bar{n}, e^{\pm}, \mu^{\pm})$

PFN-Ex: Experimentally accessible info
$(\gamma, h^{\pm,0}, e^{\pm}, \mu^{\pm})$

PFN-Ch: Particle charge info
$(+,0,-)$

PFN: No particle type info, arbitrary energy dependence

EFN: IRC-safe latent space



Particle type info helpful

IRC unsafe information clearly helpful

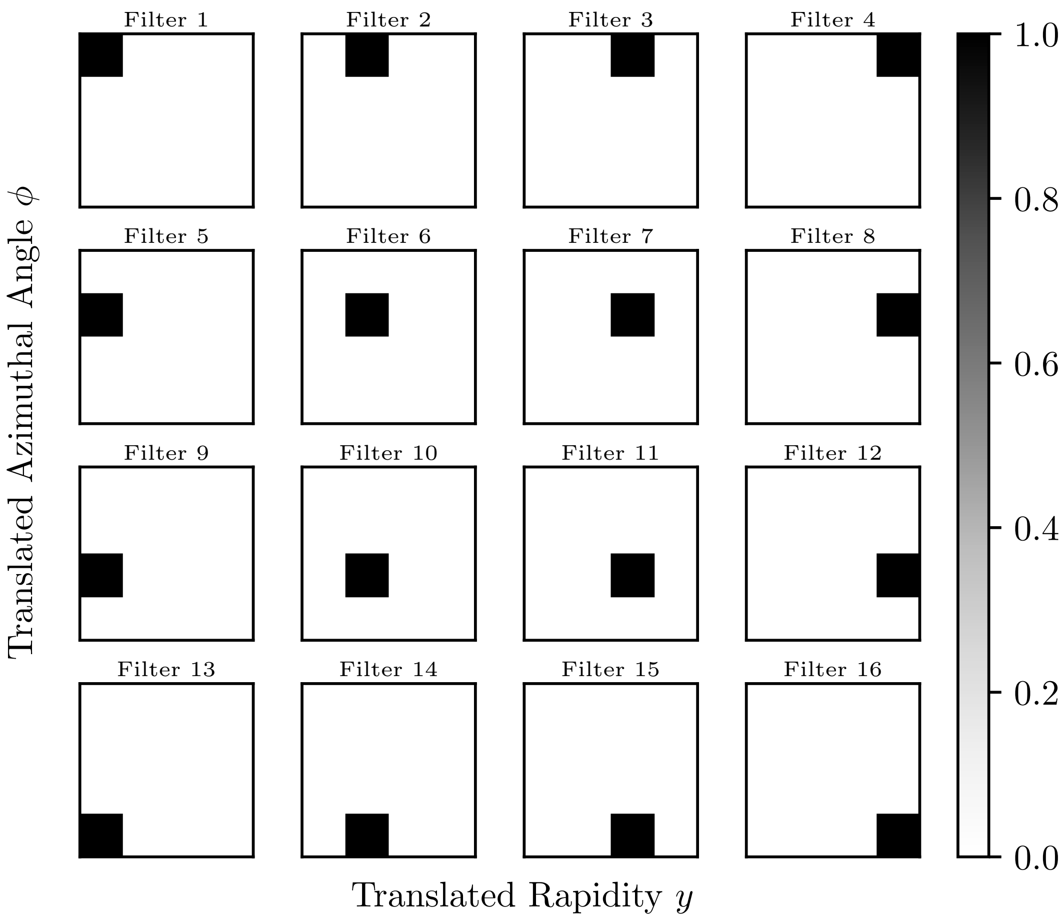
Performance saturates as dimension increases

Energy Flow Network Visualization

EFN observables are two-dimensional geometric functions

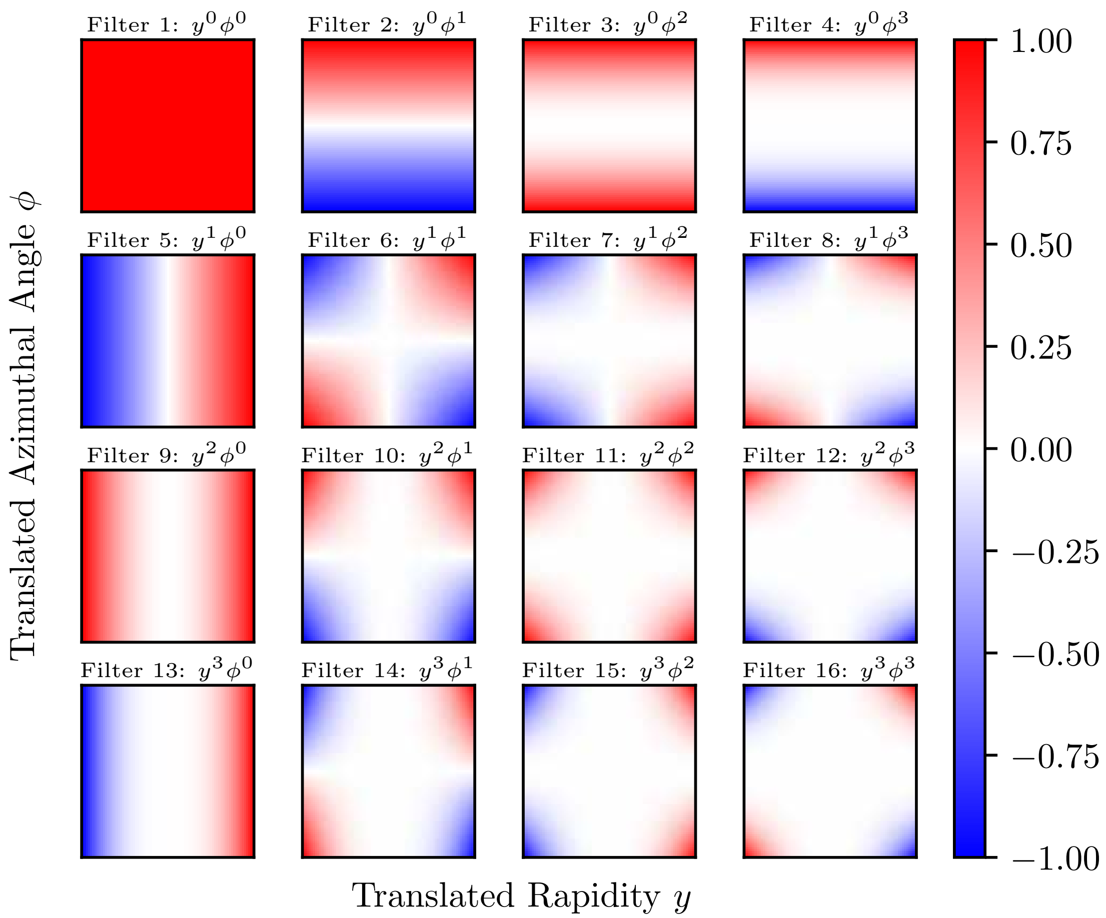
Visualize EFN observables as *filters* in the translated rapidity-azimuth plane

Jet images as EFN filters



[Cogan, Kagan, Strauss, Schwartzman, 2014]
[de Oliviera, Kagan, Mackey, Nachman, Schwartzman, 2015]

Moments as EFN filters



[Donoghue, Low, Pi, 1979]
[Gur-Ari, Papucci, Perez, 2011]

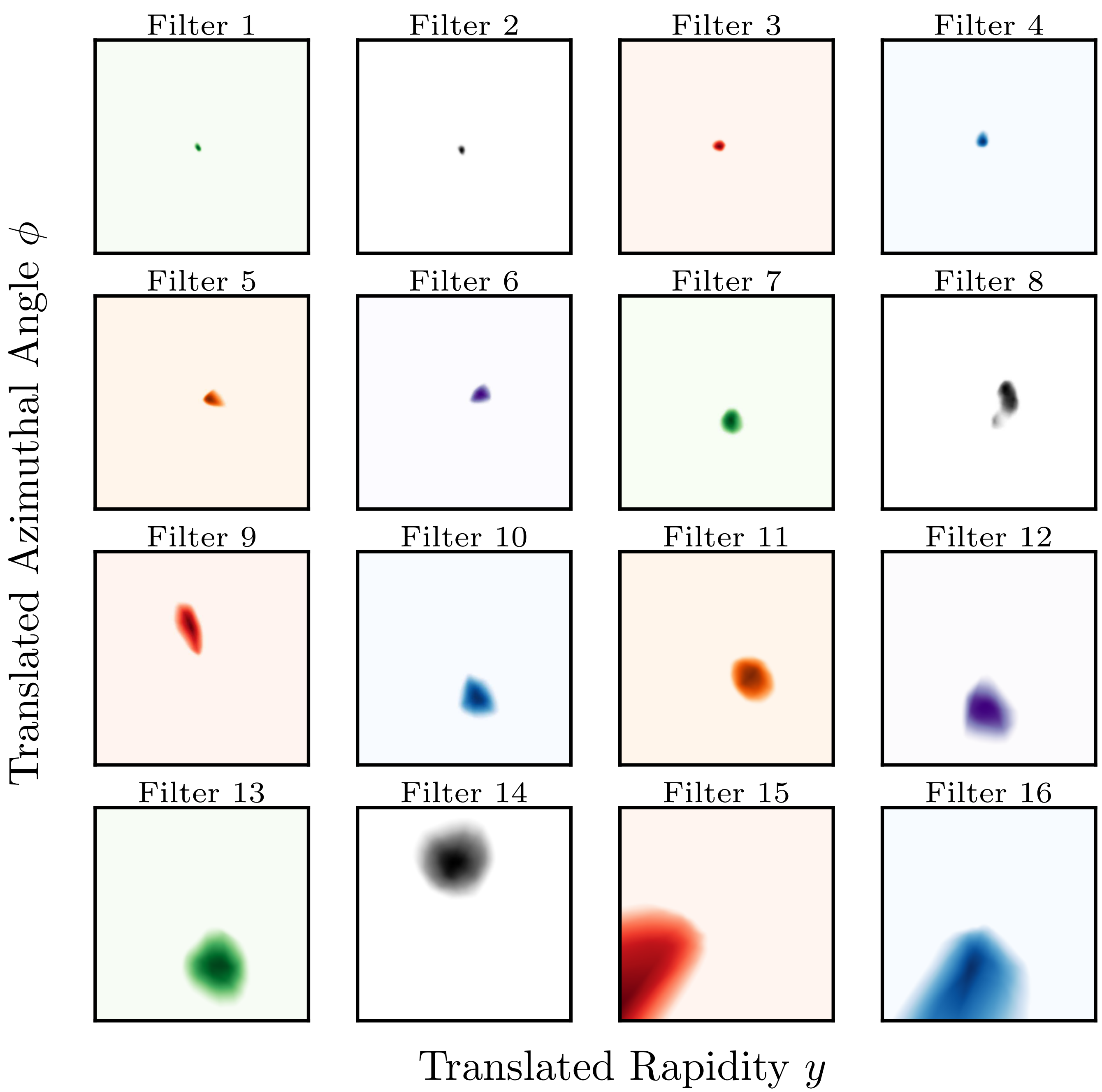
Visualizing Q/G EFN Filters

Generally see blobs of all scales

Local nature of activated region lends interpretation as "pixels"

EFN seems to have learned a dynamically sized jet image

EFN (ℓ = 256) randomly selected filters, sorted by size



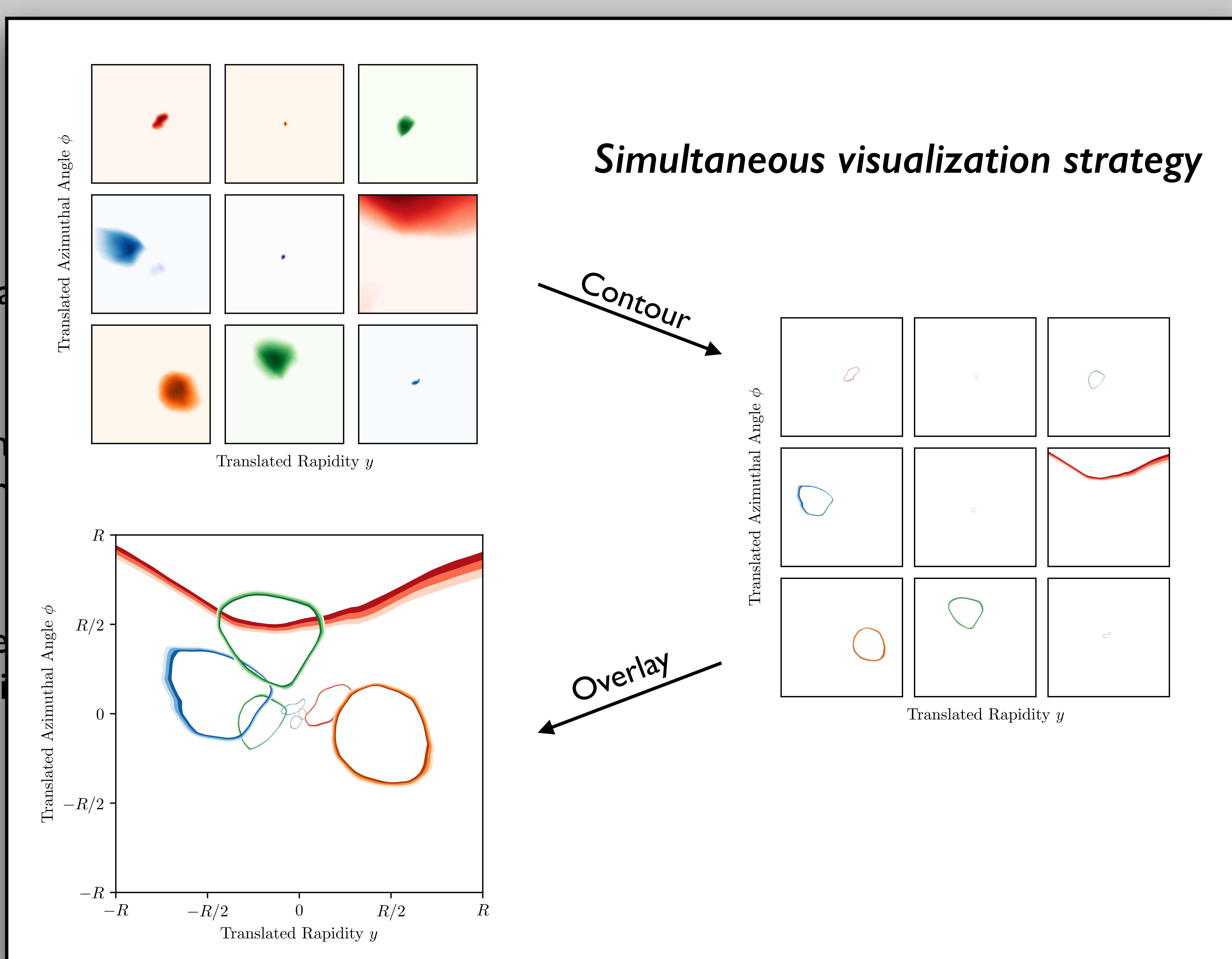
Visualizing Q/G EFN Filters

by size

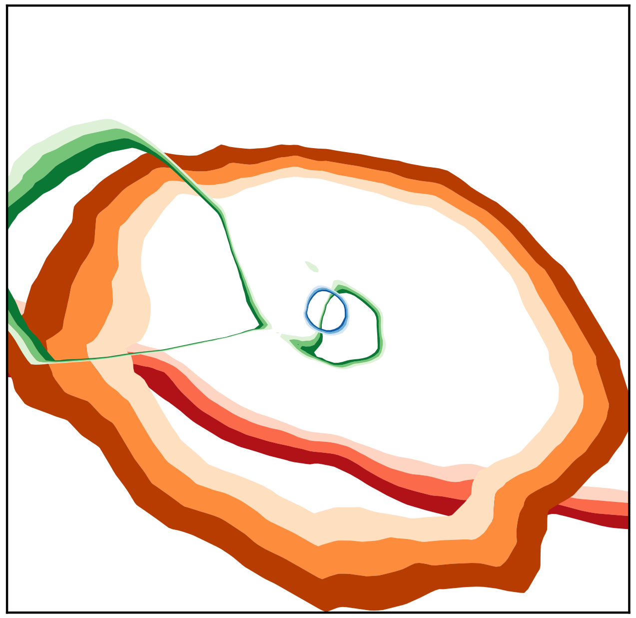
Genera

Local n
interpr

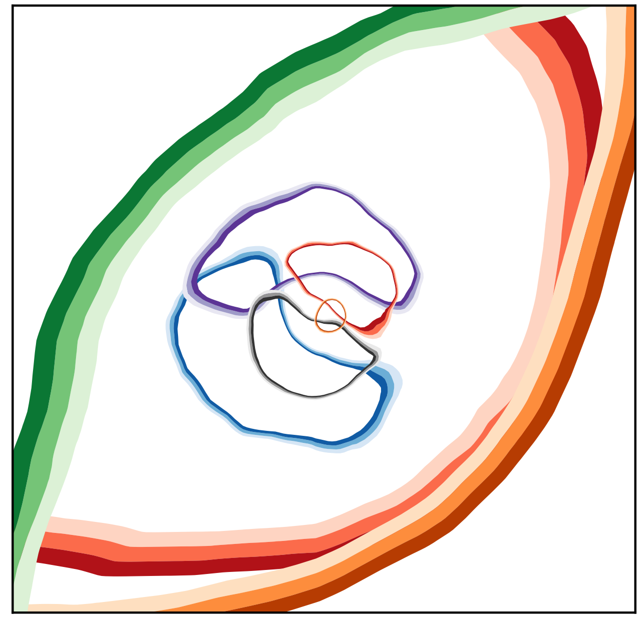
EFN se
dynami



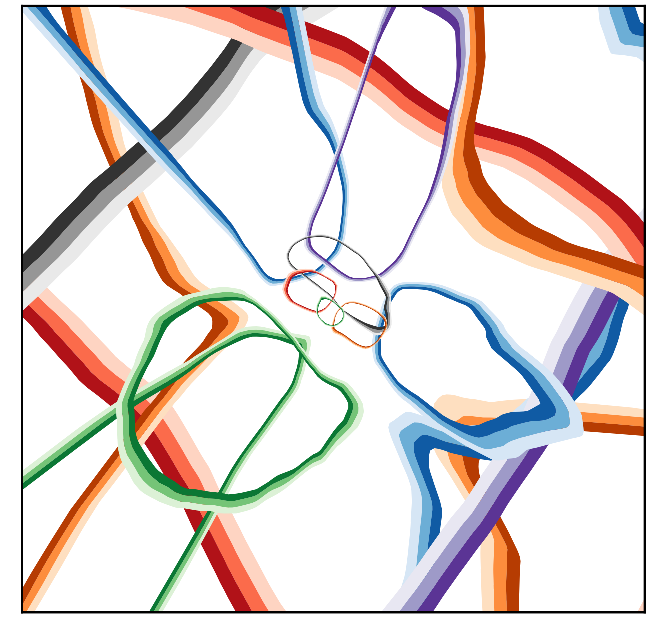
Visualizing Q/G EFN Filters



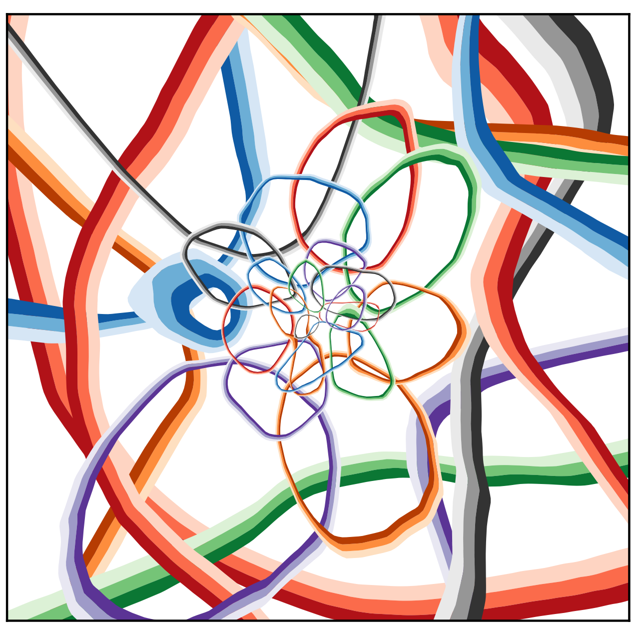
$\ell = 4$



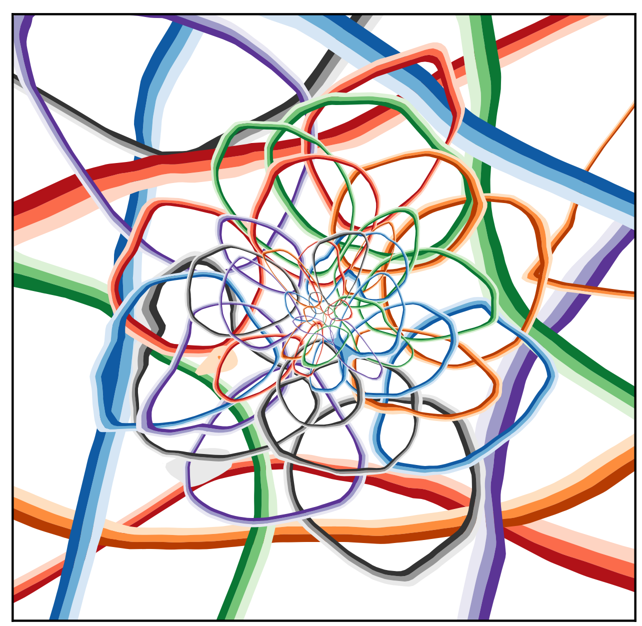
$\ell = 8$



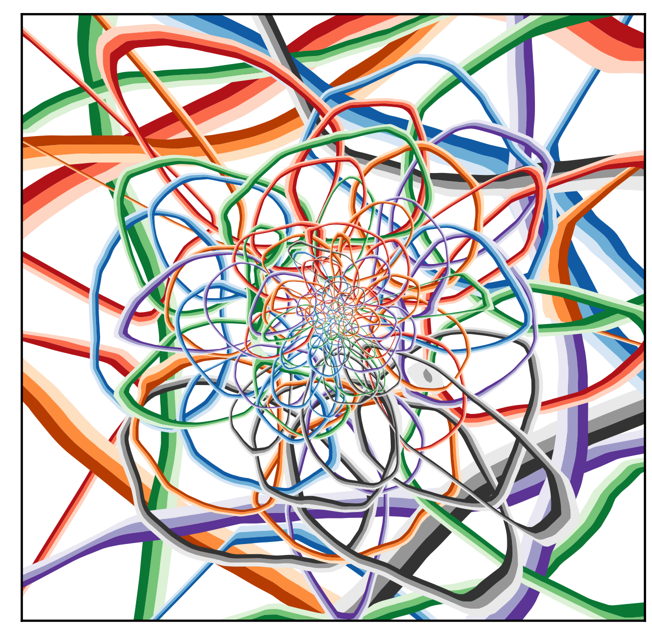
$\ell = 16$



$\ell = 32$

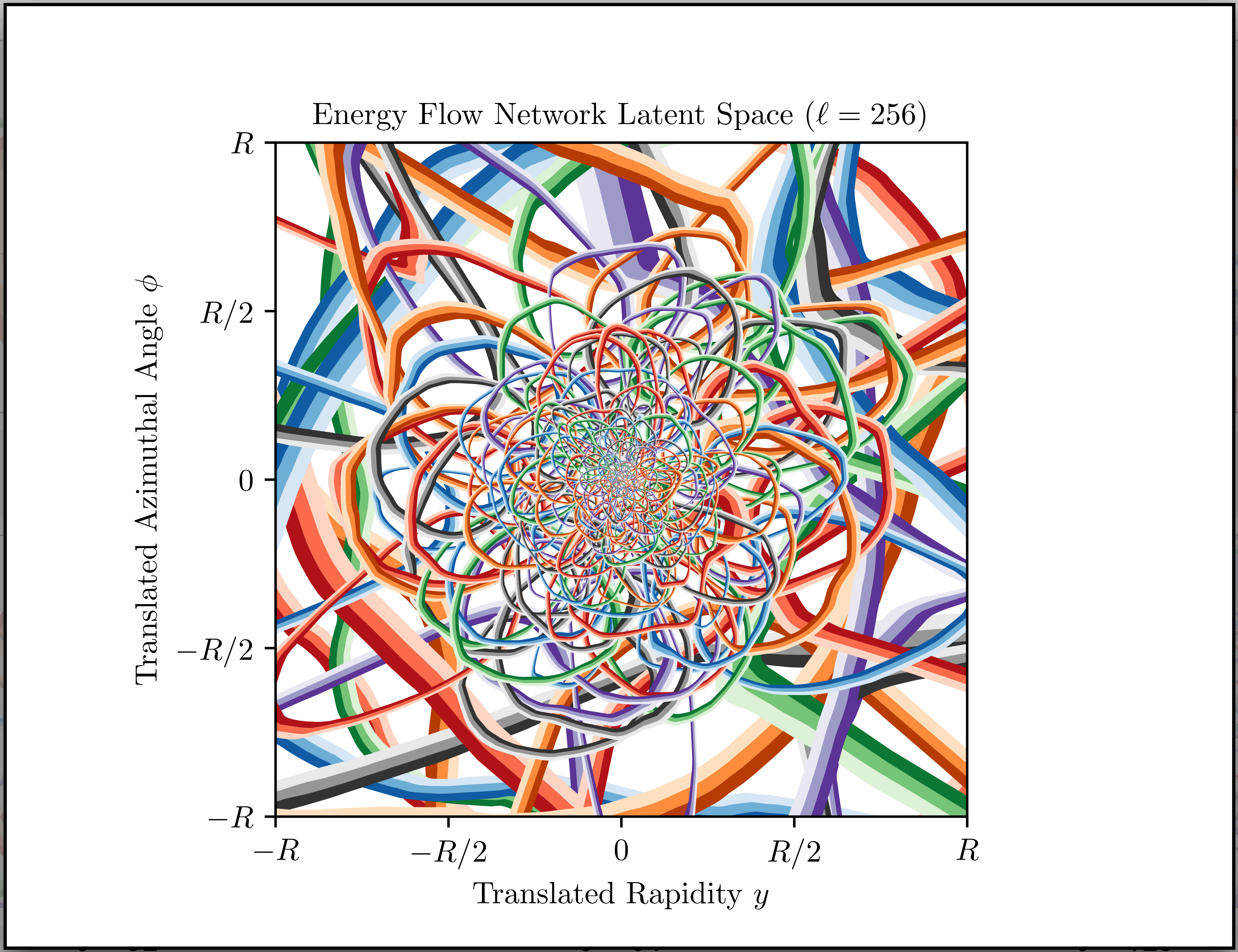


$\ell = 64$



$\ell = 128$

Visualizing Q/G EFN Filters



Measuring Q/G EFN Filters

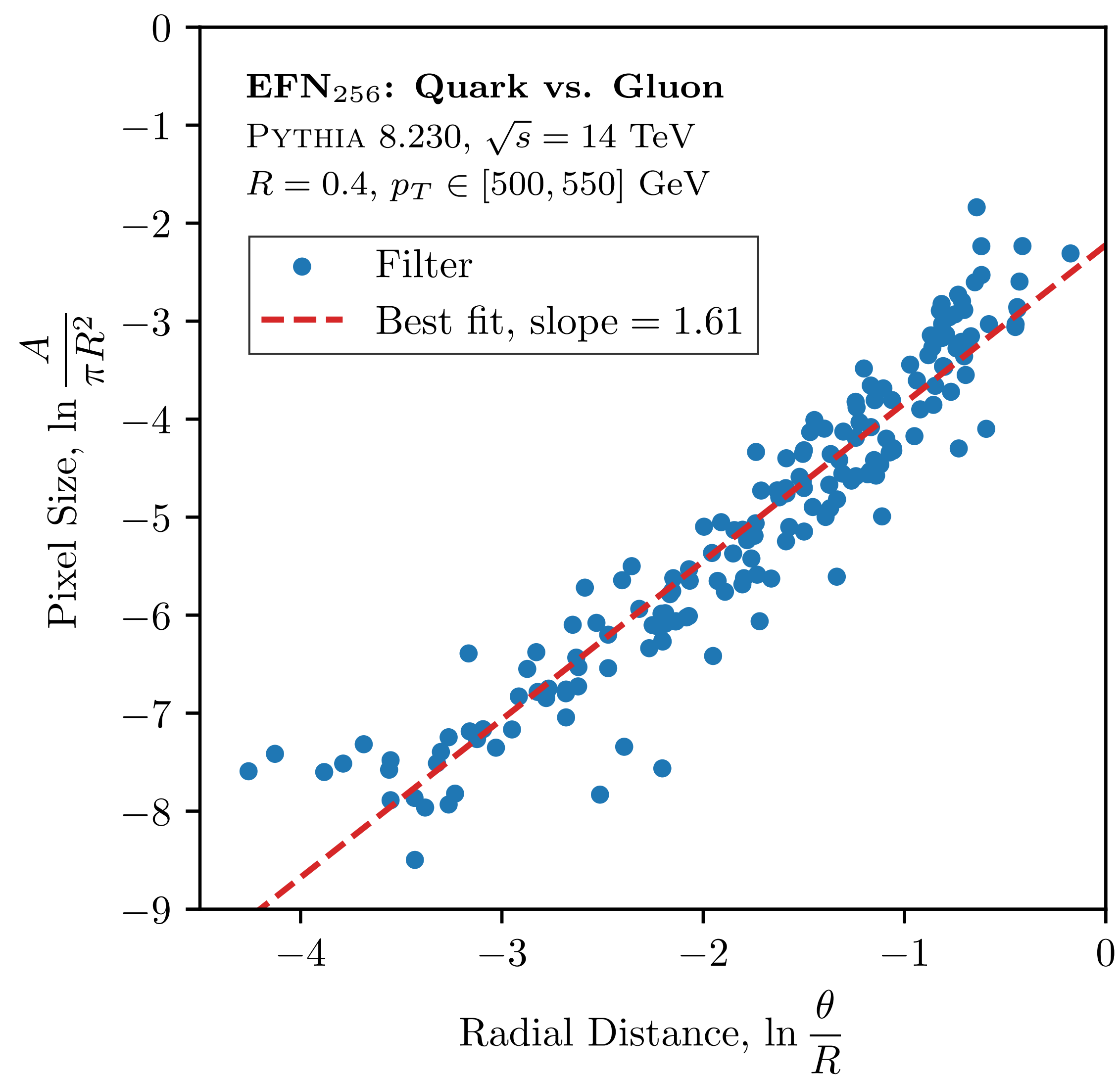
Power-law dependence between filter size and distance from center is observed

Slope of 2 is predicted at leading log

$$\left|\boxed{\mathrm{d}\ln\frac{\theta}{R}\,\mathrm{d}\varphi}\right| = \theta^2\,\boxed{\mathrm{d}y\,\mathrm{d}\phi}$$

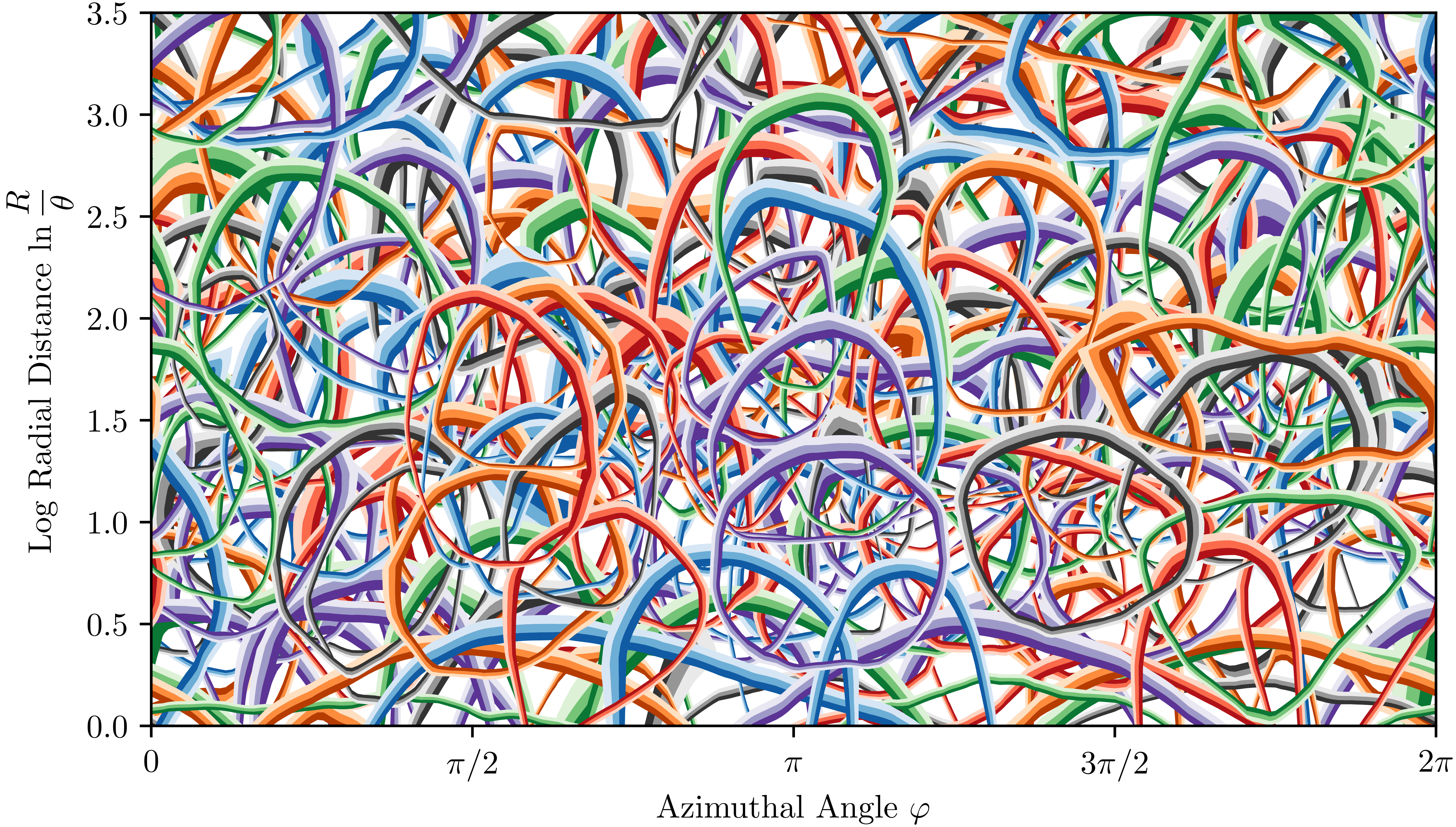
Emission plane area element

Area element in rap-phi plane

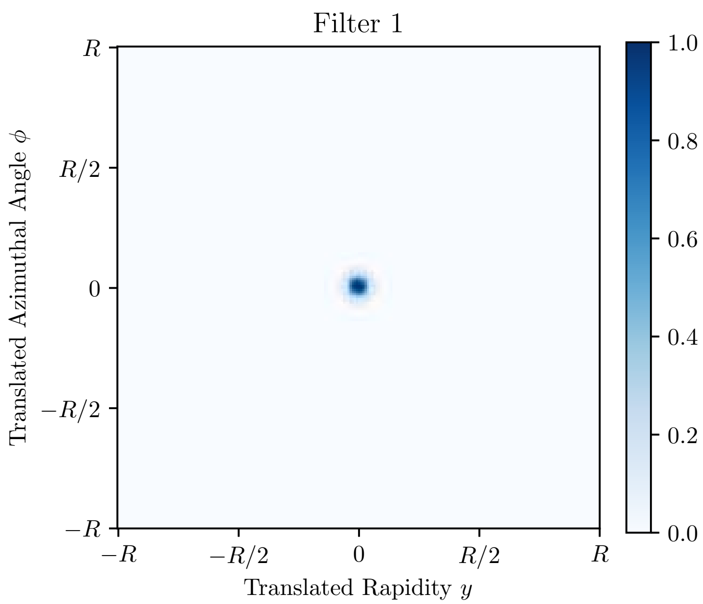


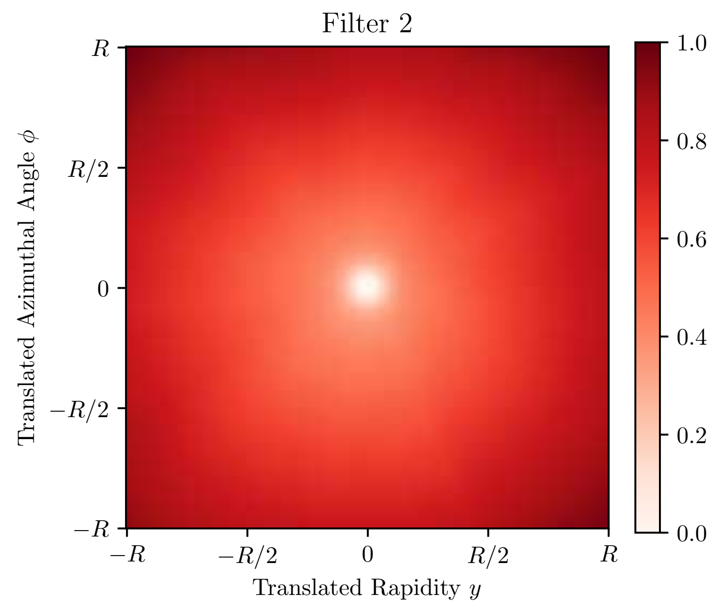
Non-perturbative physics, axis recoil, higher order effects cause deviations from slope of 2

Visualizing Q/G EFN Filters in the Emission Plane

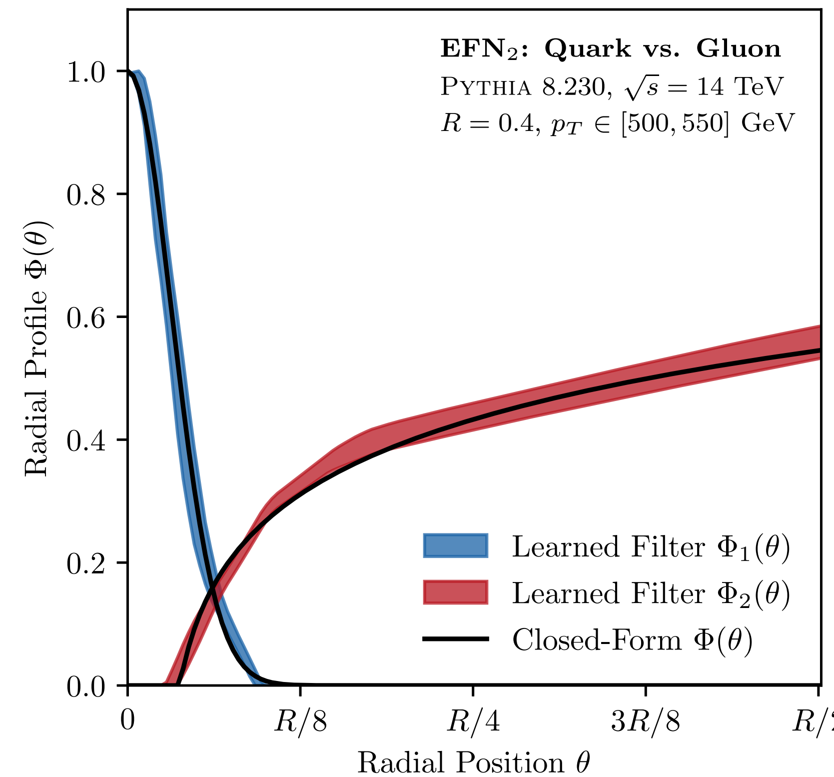


Extracting New Analytic Observables





$$\mathcal{O}_1 = \sum_{i=1}^{M} z_i \, \Phi_1(\theta_i) \qquad \mathcal{O}_2 = \sum_{i=1}^{M} z_i \, \Phi_2(\theta_i)$$



Take radial slices to obtain envelope

EFN (ℓ = 2) has approximately radially symmetric filters

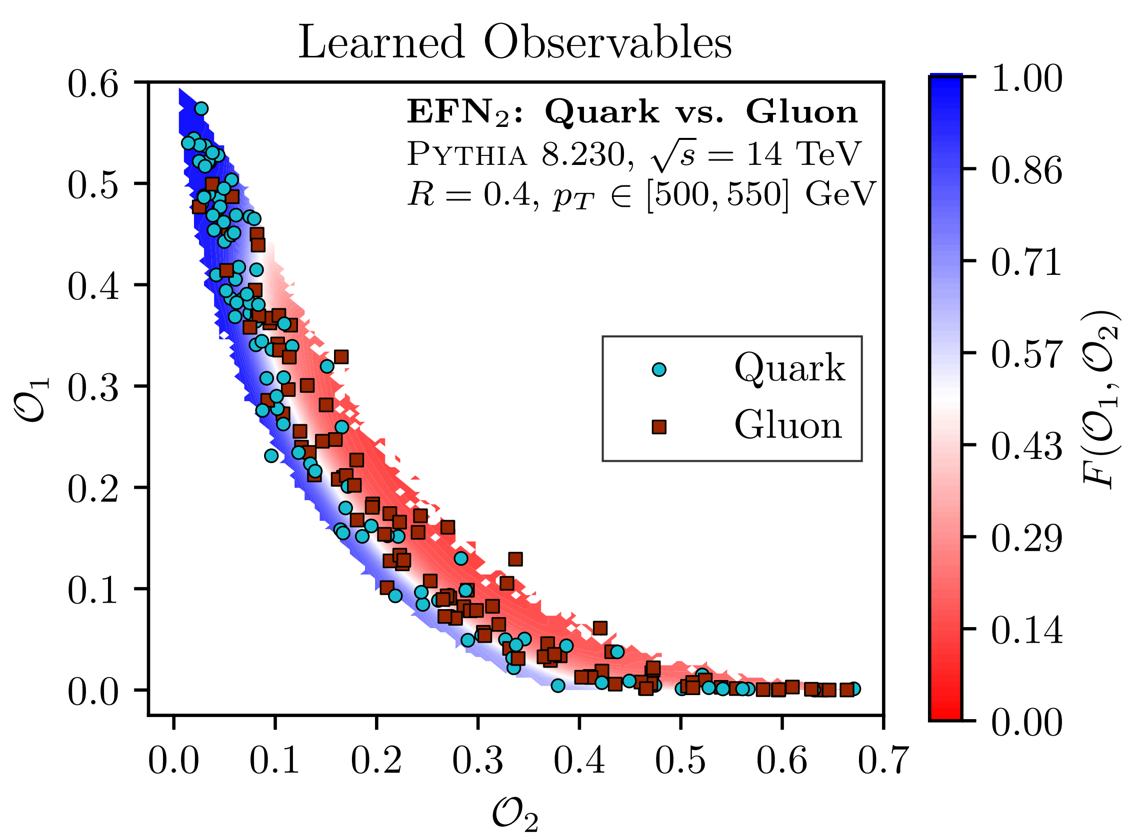
Fit functions of the forms:

$$A_{r_0} = \sum_{i=1}^{M} z_i \, e^{-\theta_i^2/r_0^2}, \qquad B_{r_1,\beta} = \sum_{i=1}^{M} z_i \, \ln(1 + \beta(\theta_i - r_1))\Theta(\theta_i - r_1)$$

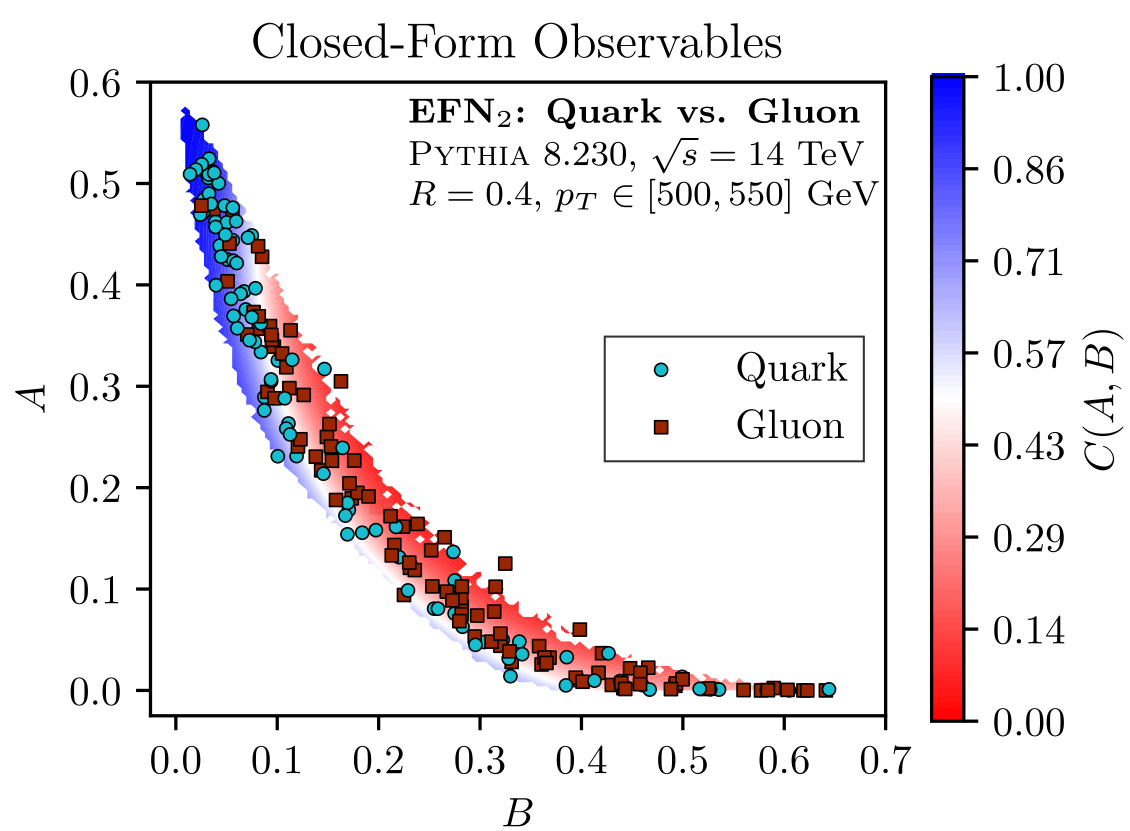
Separate soft and collinear phase space regions

Extracting New Analytic Observables

Can visualize *F* in the two dimensional $(\mathcal{O}_1, \mathcal{O}_2)$ phase space



Learned



Extracted

Extract analytic form for *F* as (squared) distance from a point:

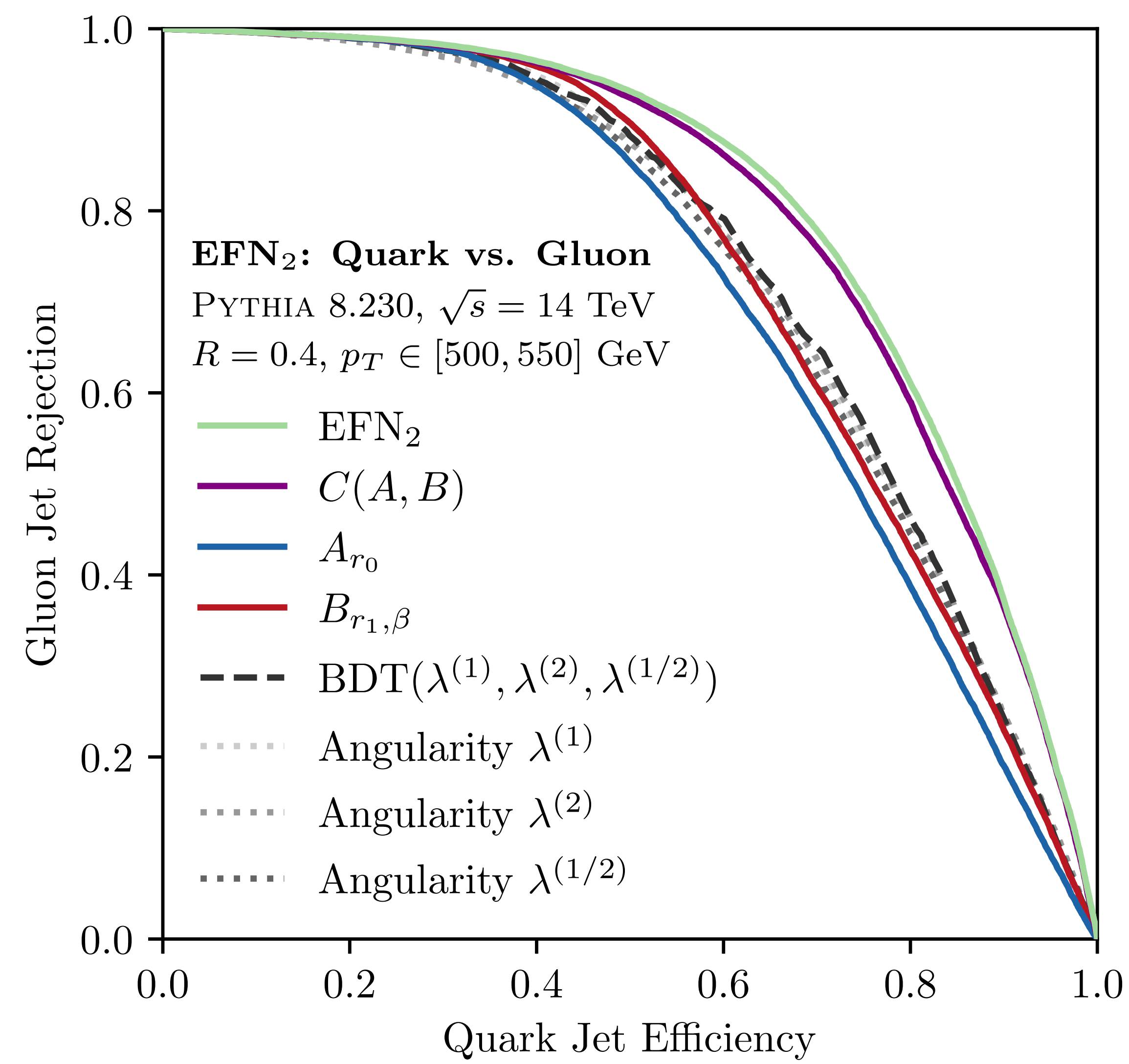
$$C(A, B) = (A - a_0)^2 + (B - b_0)^2$$

Benchmarking New Analytic Observables

Individually, extracted observables are comparable to other angularities

Extracted $C(A, B)$ performs nearly as well as EFN (ℓ = 2)

Meanwhile, multivariate combination (BDT) of three other angularities does not show improvement



Top Jet Samples and Other Methods

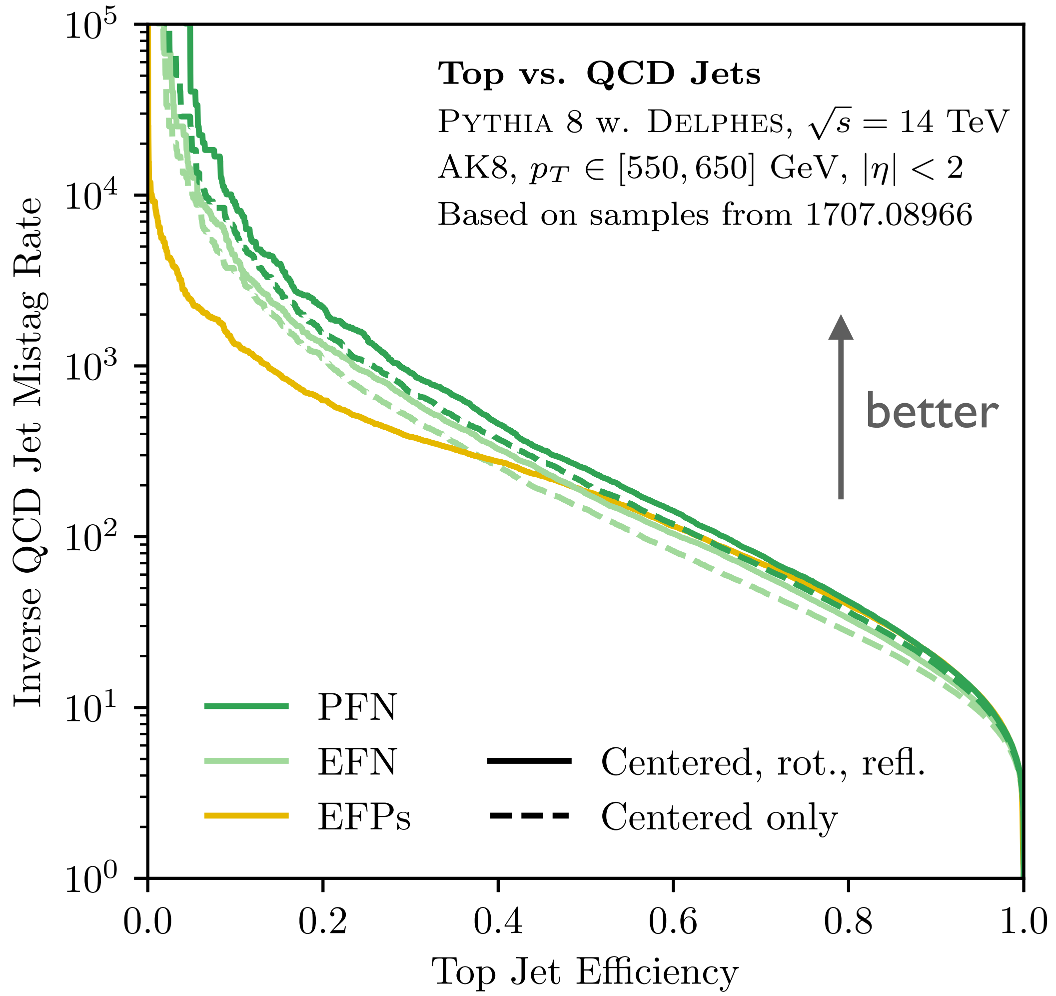
[Butter, Kasieczka, Plehn, Russell, 2017]

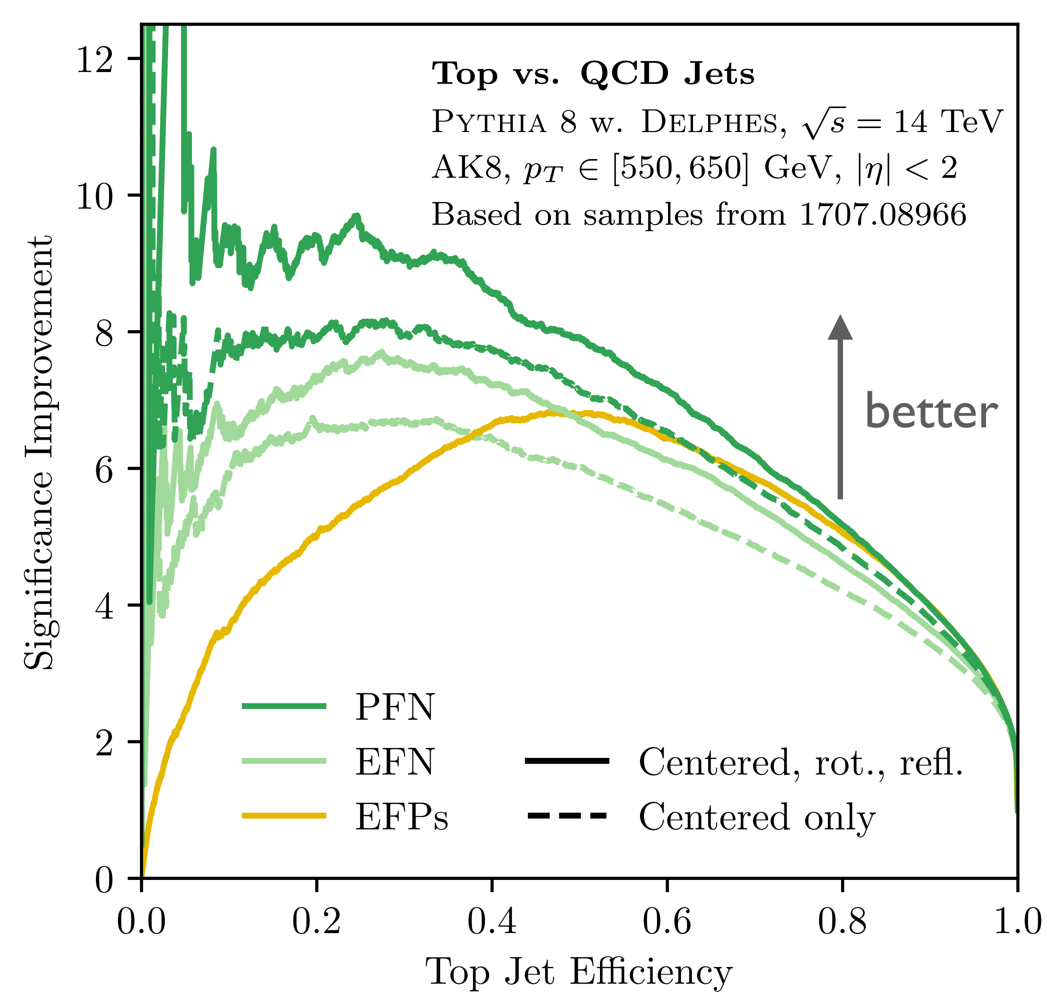
Common top and QCD dijet samples for standardized benchmarking

$p_T \in [550, 650]$ GeV, AK8 jets, fully-merged, Delphes simulation, 2m jets total

Approach	AUC	Acc.	1/eB @ (eS=0.3)	Contact	Comments
LoLa	0.979	0.928		G. Kasieczka S. Leiss	Preliminary number, based on LoLa
LBN	0.981	0.931	863	M. Rieger	Preliminary number
CNN	0.981	0.93	780	D. Shih	Model from *(1803.00107)*
P-CNN (1D CNN)	0.980	0.930	782	H. Qu, L. Gouskos	Preliminary, use kinematic info only
6-body N-subs. (+mass and pT) NN	0.979	0.922	856	K. Nordstrom	Based on 1807.04769
8-body N-subs. (+mass and pT) NN	0.980	0.928	795	K. Nordstrom	Based on 1807.04769
Linear EFPs	0.980	0.932	380	PTK, E. Metodiev	d<= 7, chi <= 3 EFPs with FLD. Based on 1712.07124
Particle Flow Network (PFN)	0.982	0.932	888	PTK, E. Metodiev	Median over ten trainings. Based on Table 5 in 1810.05165
Energy Flow Network (EFN)	0.979	0.927	619	PTK, E. Metodiev	Median over ten trainings. Based on Table 5 in 1810.05165

Classification Performance



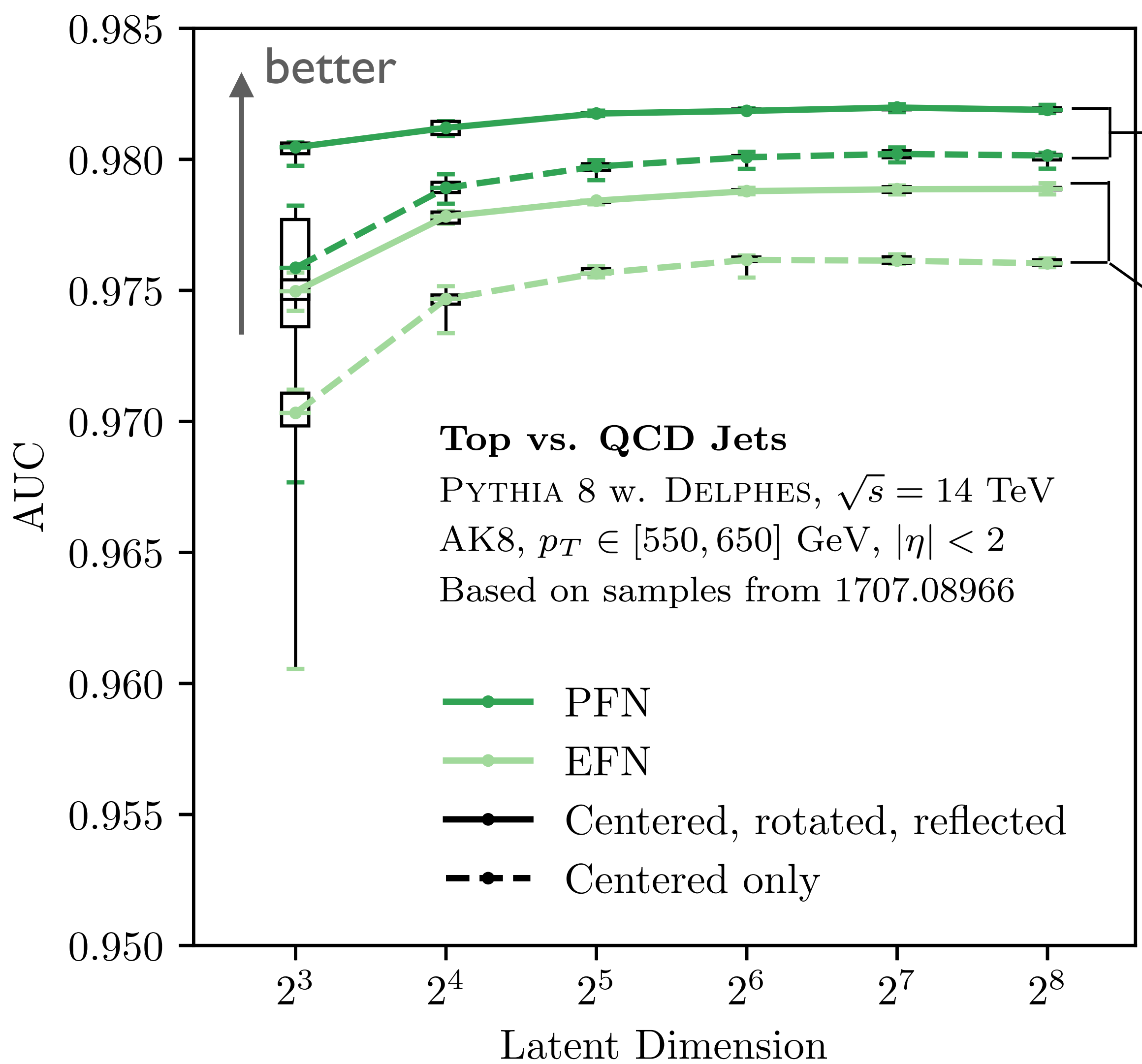


Latent space dimension $\ell = 256$

EFN/PFN rotation and reflection preprocessing helpful

EFPs are comparable to EFN and even better at high signal efficiency

EFN Latent Dimension Sweep



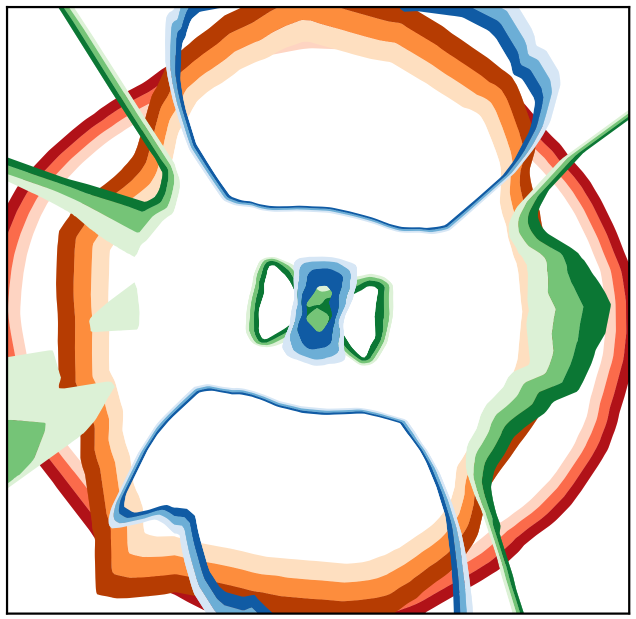
Preprocessing clearly helpful

IRC unsafe information clearly helpful

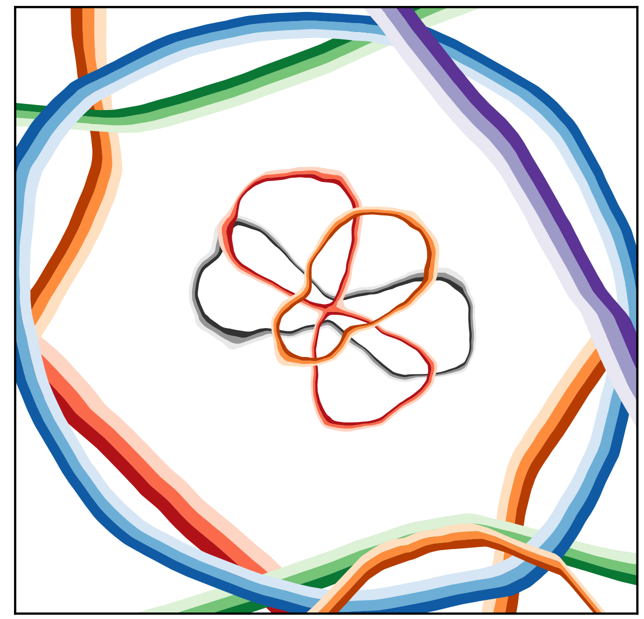
Performance quickly saturates as latent dimension increases

Visualizing EFN Filters

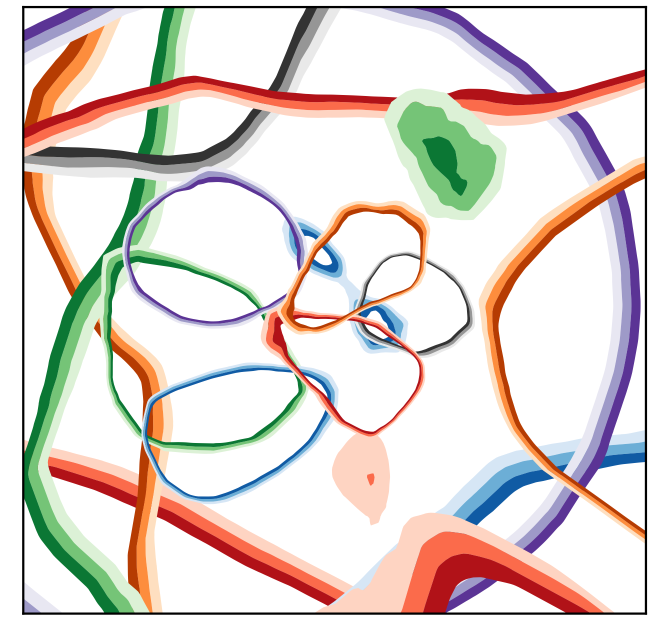
Without rotation/reflection preprocessing



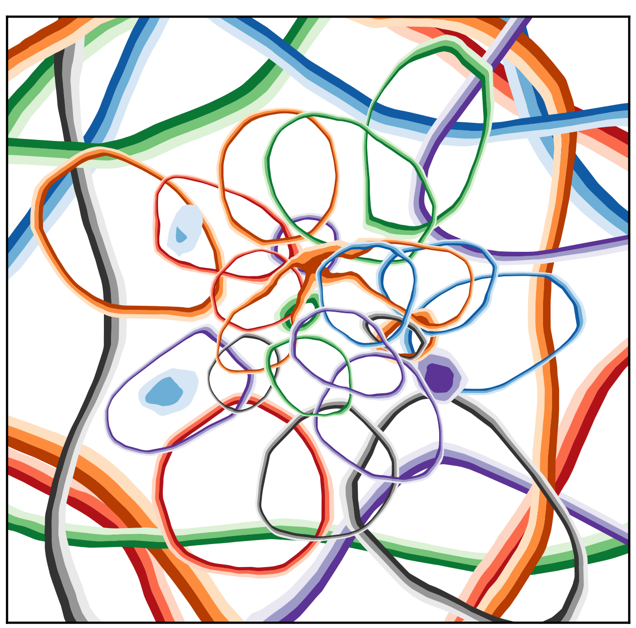
$\ell = 4$



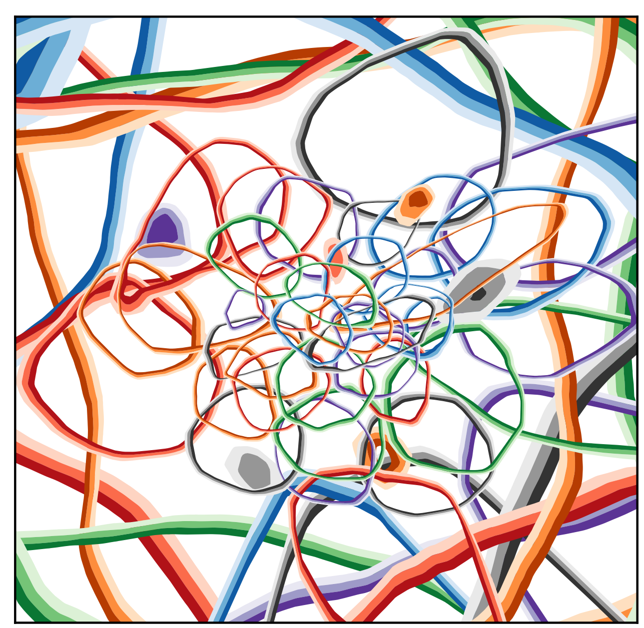
$\ell = 8$



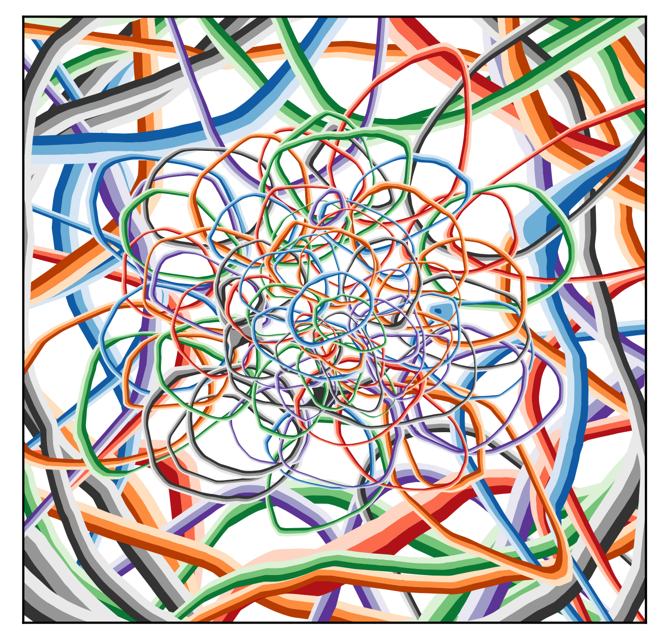
$\ell = 16$



$\ell = 32$



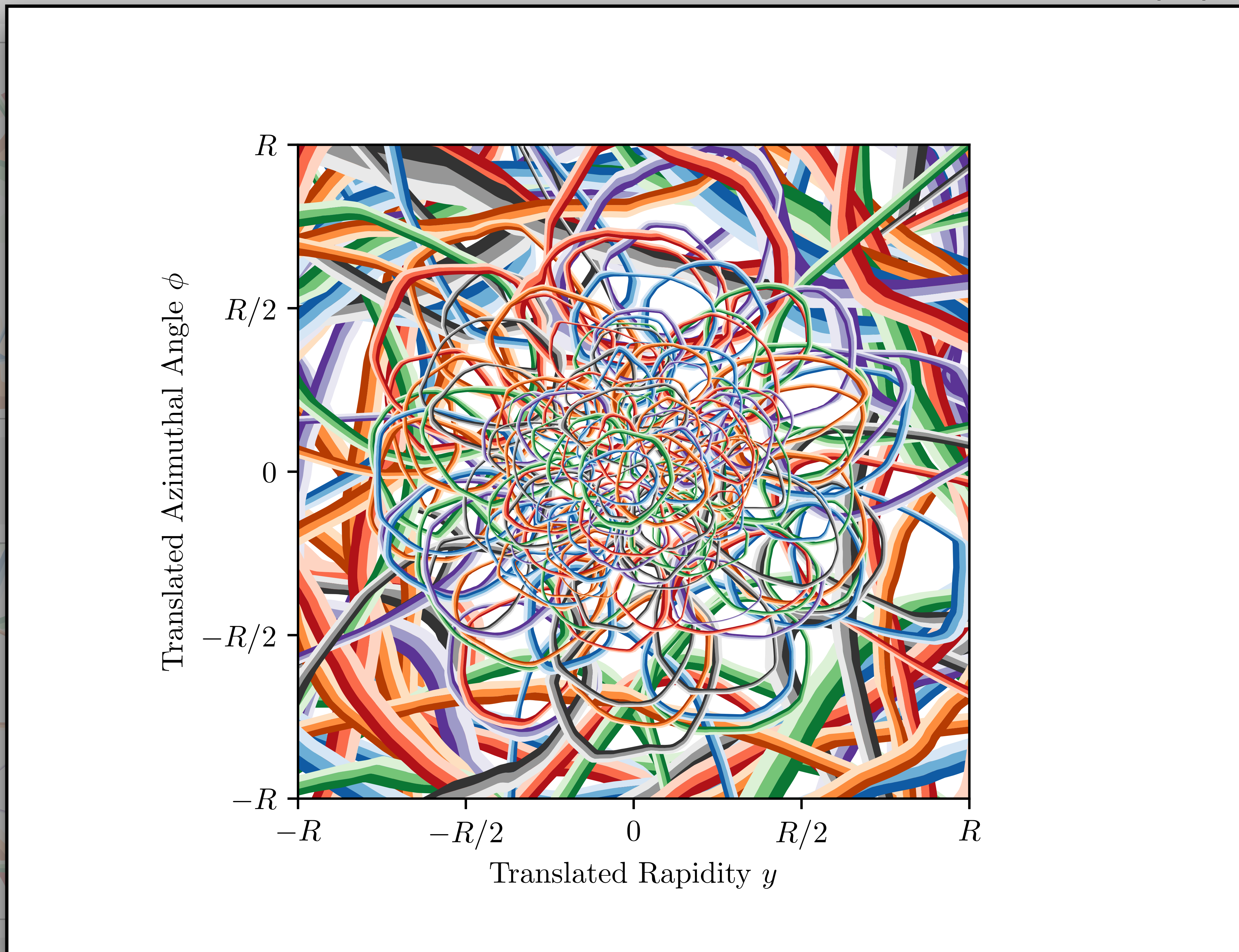
$\ell = 64$



$\ell = 128$

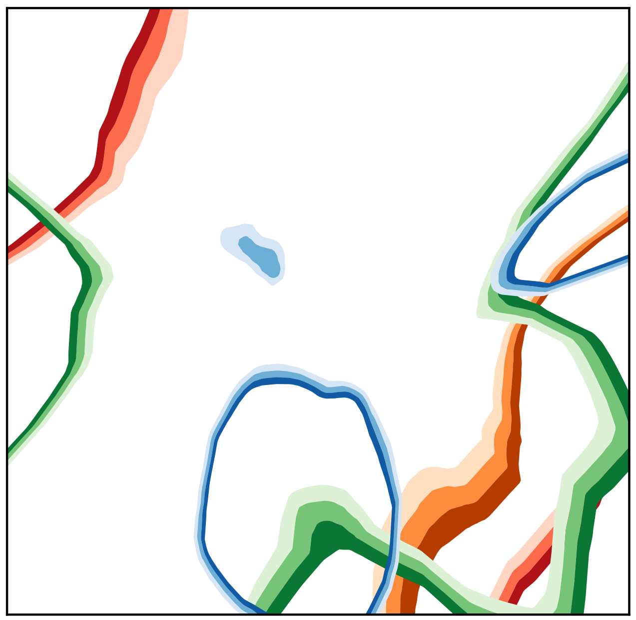
Visualizing EFN Filters

Without rotation/reflection preprocessing

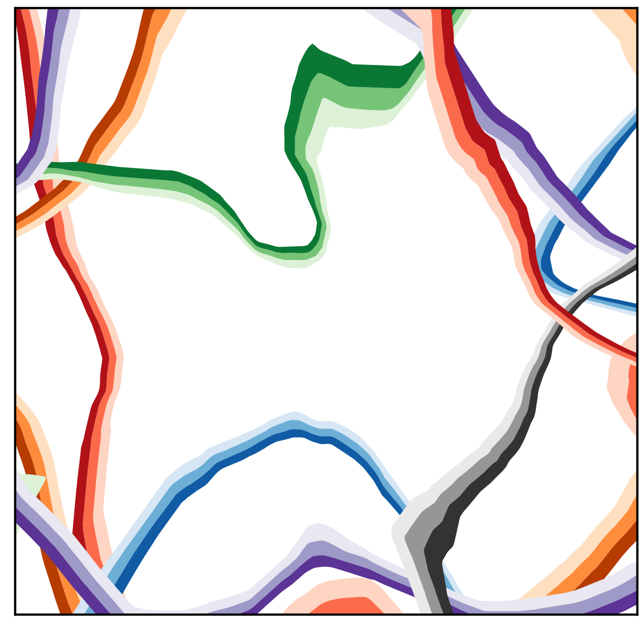


Visualizing EFN Filters

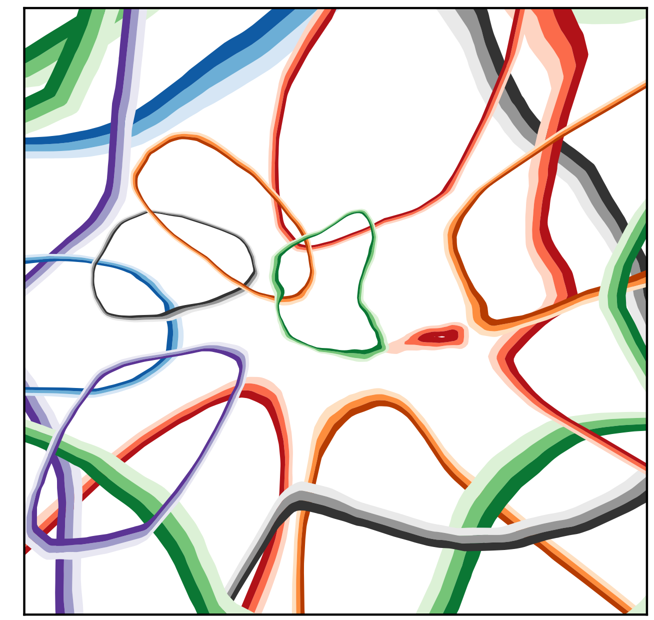
With rotation/reflection preprocessing



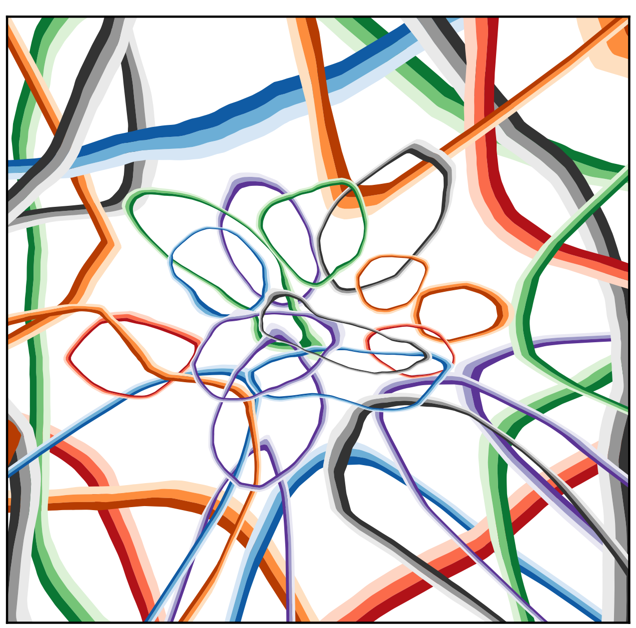
$\ell = 4$



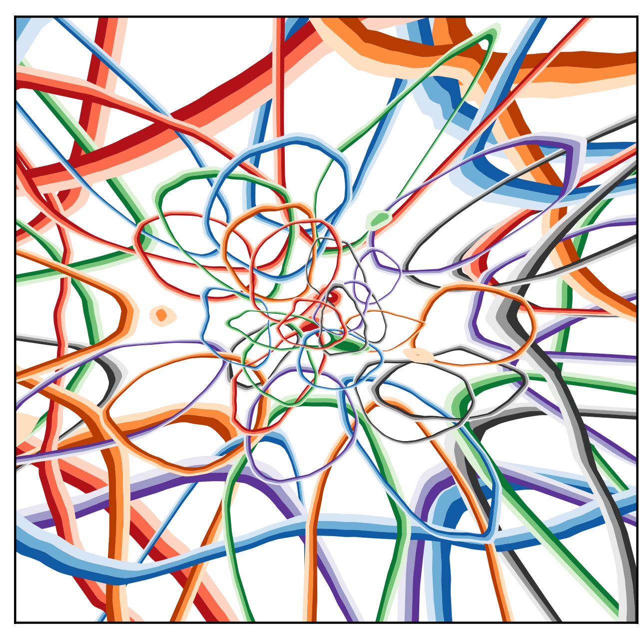
$\ell = 8$



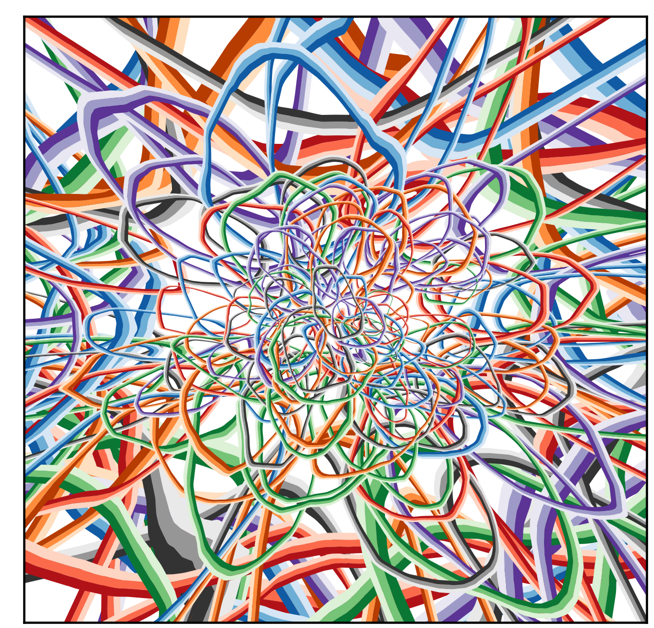
$\ell = 16$



$\ell = 32$



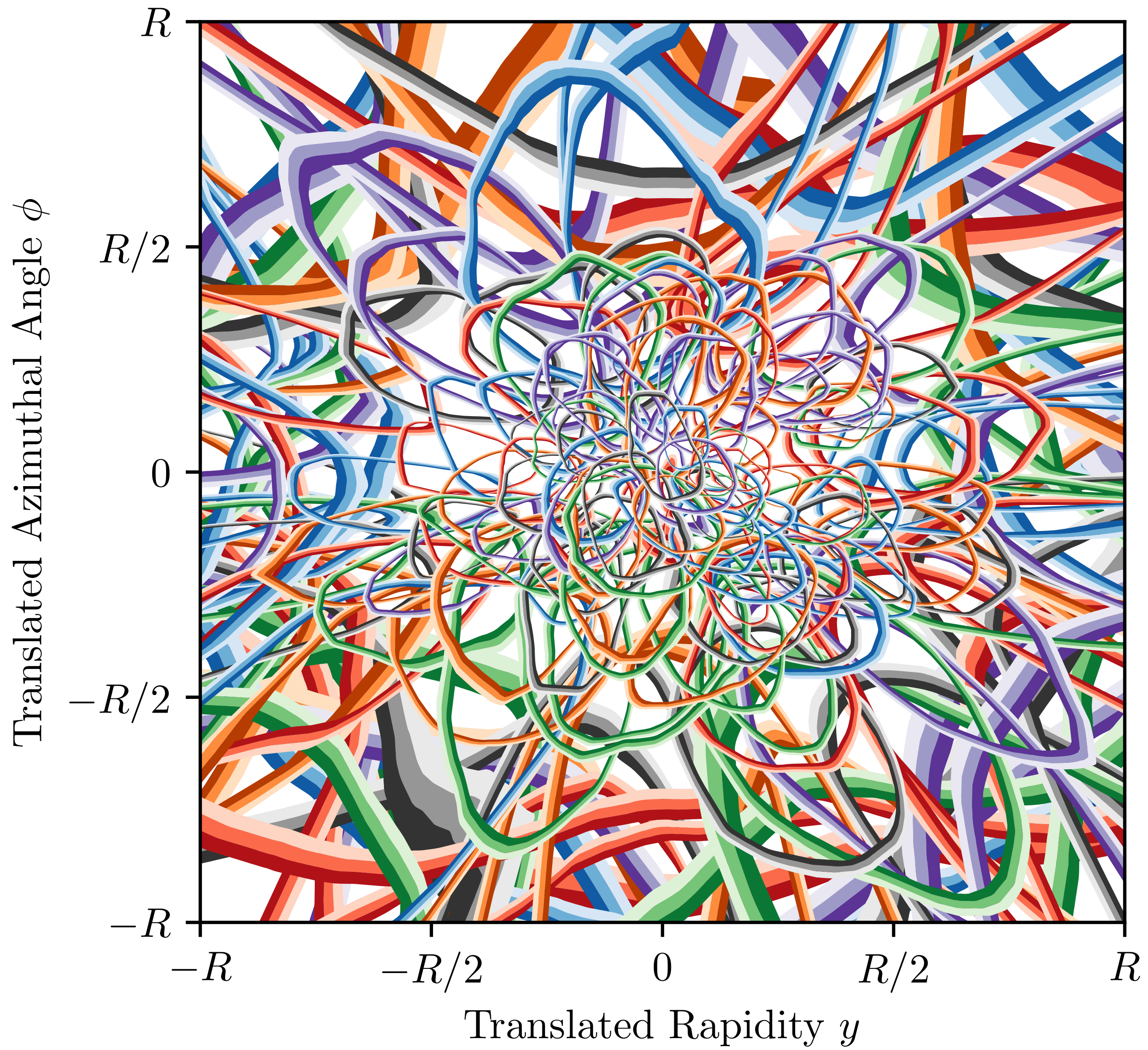
$\ell = 64$



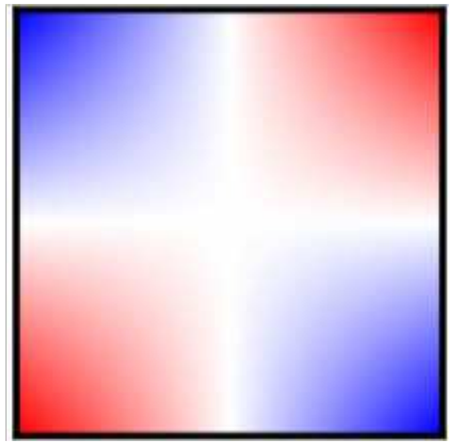
$\ell = 128$

Visualizing EFN Filters

Without rotation/reflection preprocessing

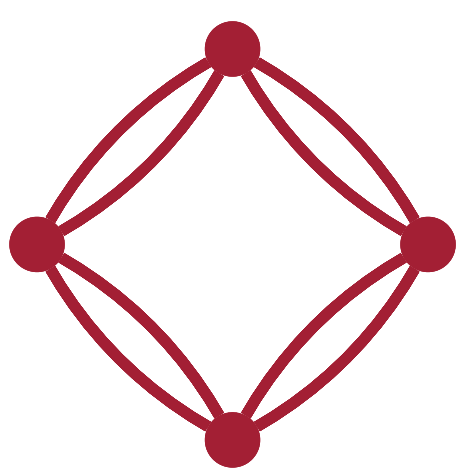


Energy Flow Networks



Energy Flow Moments

Energy Flow Polynomials



Energy Flow Networks

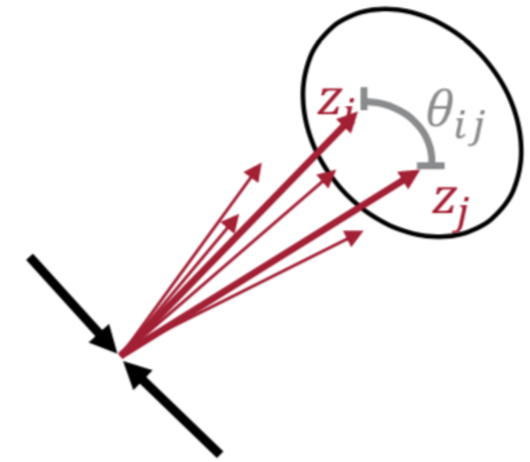
Energy Flow Moments

Energy Flow Polynomials

Energy Flow Polynomials (EFPs)

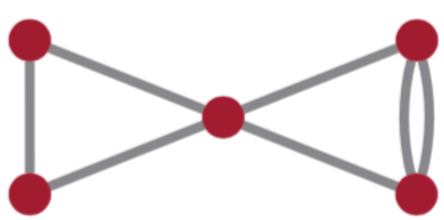
[PTK, Metodiev, Thaler, 1712.07124]

$$\boxed{\mathrm{EFP}_G = \underbrace{\sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M}}_{\text{Correlator}} \text{ of } \underbrace{z_{i_1} \cdots z_{i_N}}_{\text{Energies}} \text{ and } \underbrace{\prod_{(k,\ell)\in G} \theta_{i_k i_\ell}}_{\text{Angles}}}$$



Generalizes many well-known and studied classes of energy correlators observables

A family of energy correlators with angular structures determined by multigraphs



$$= \sum_{i_1=1}^{M}\sum_{i_2=1}^{M}\sum_{i_3=1}^{M}\sum_{i_4=1}^{M}\sum_{i_5=1}^{M} z_{i_1} z_{i_2} z_{i_3} z_{i_4} z_{i_5} \theta_{i_1 i_2} \theta_{i_2 i_3} \theta_{i_1 i_3} \theta_{i_1 i_4} \theta_{i_1 i_5} \theta_{i_4 i_5}^2$$

Multigraph correspondence

$\longleftrightarrow$ z_{i_j}

$\longleftrightarrow$ $\theta_{i_k i_l}$

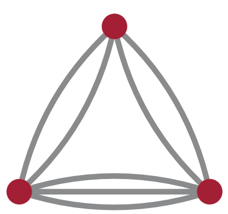
Energy and Angle Measure

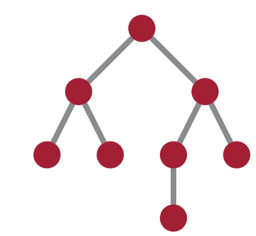
Hadronic : $z_i = \dfrac{p_{Ti}}{\sum_j p_{Tj}}, \quad \theta_{ij} = \left(\Delta y_{ij}^2 + \Delta\phi_{ij}^2\right)^{\beta/2}$

Linear Basis of IRC-Safe Observables

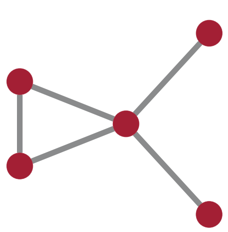
One can show via the Stone-Weierstrass approximation theorem that any IRC-safe observable is a linear combination of EFPs

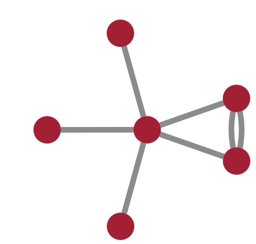
$$\mathcal{S} \simeq \sum_{G\in\mathcal{G}} s_G \mathrm{EFP}_G, \quad \mathcal{G} \text{ a set of multigraphs}$$





Multivariate combinations of EFPs only require <u>linear methods</u> to achieve full generality



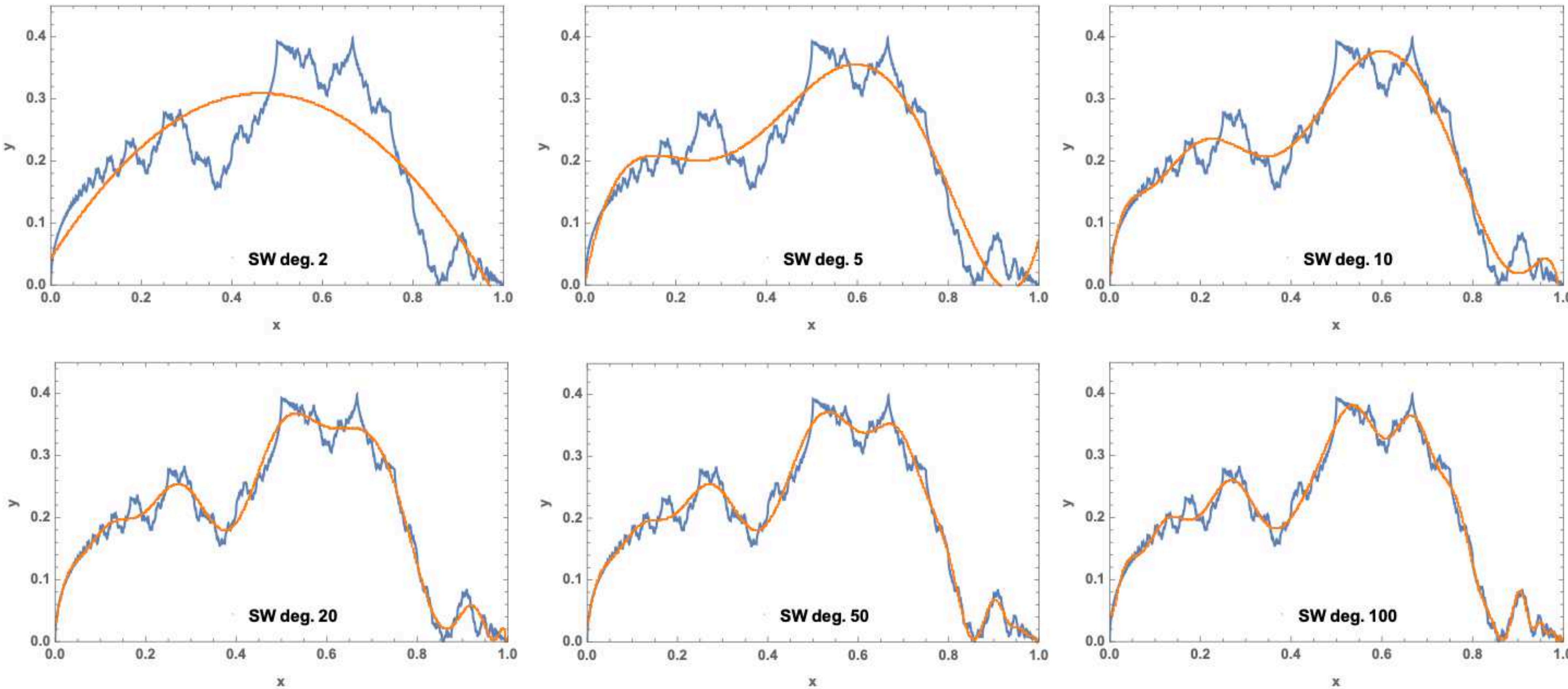


Strategy: Learn coefficients s_G via linear regression or classification

Fun with the Stone-Weierstrass Theorem

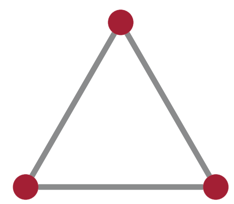
Weierstrass function – continuous everywhere, differentiable on a measure zero set of points

$$y = \sum_{k=1}^{\infty} \frac{\sin(\pi k^2 x)}{\pi k^2}$$

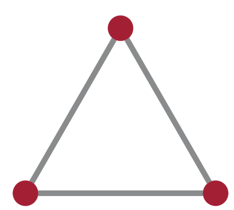


Familiar Observables as EFPs

$m_J^2 =$ 

$D_2 =$  / ()^3

[Larkoski, Moult, Neill, 2014]

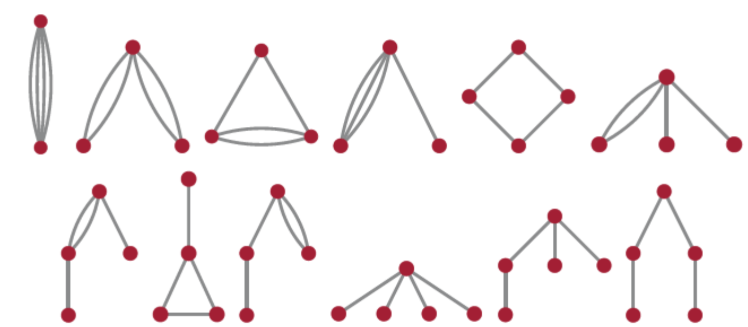
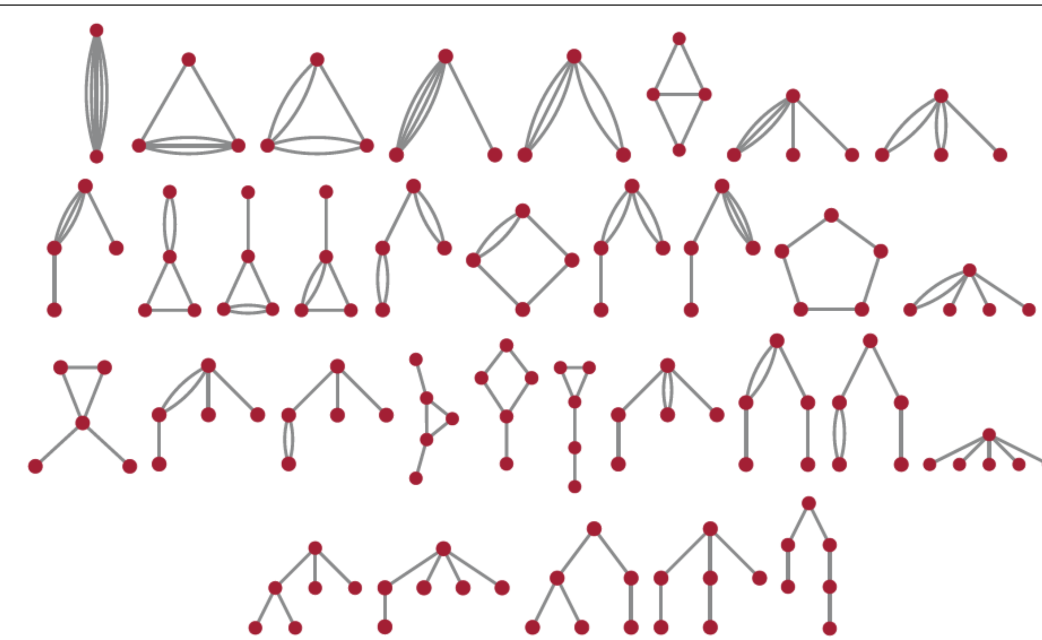
$C_2 =$  / ()^2

[Larkoski, Salam, Thaler, 2013]

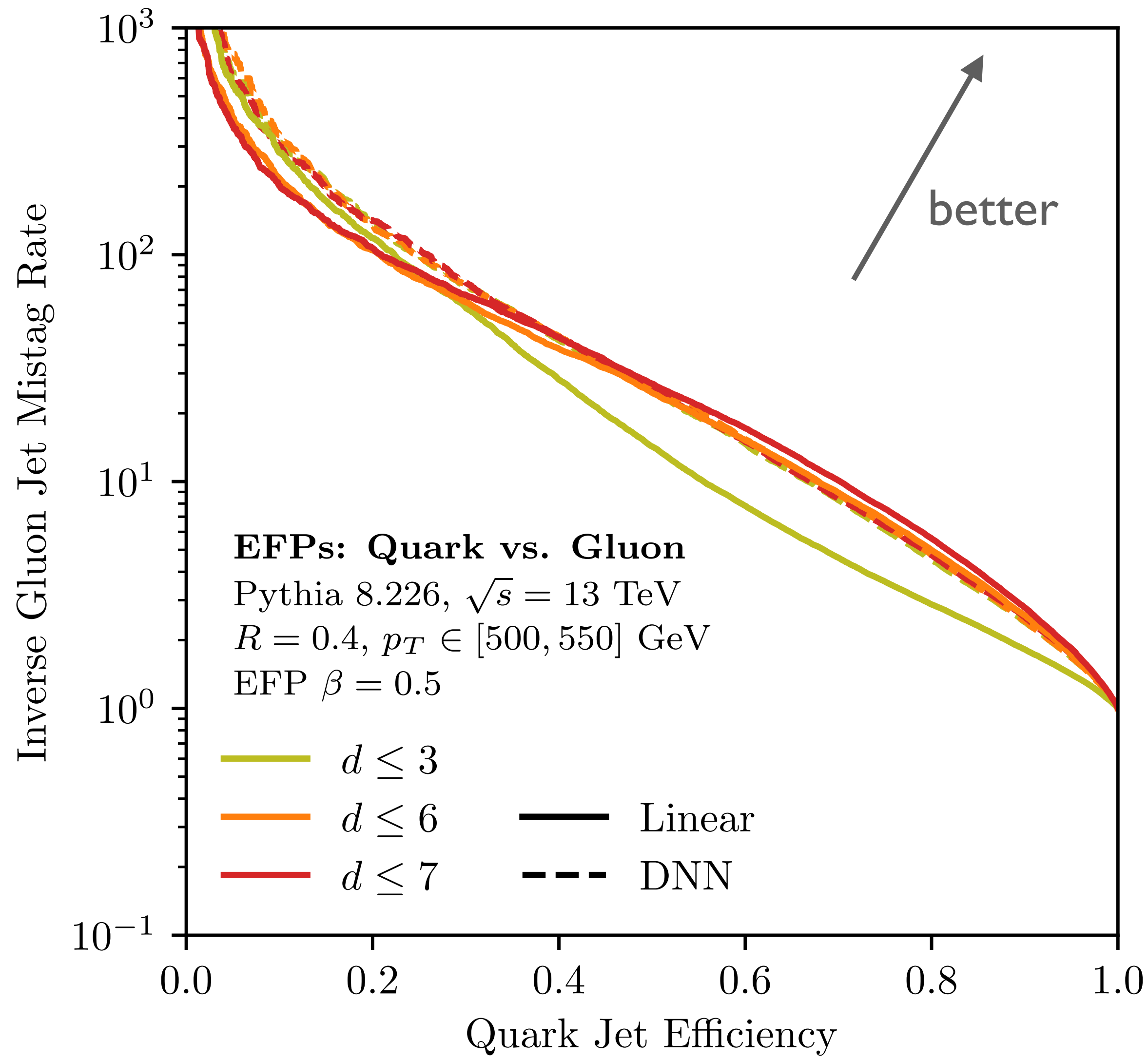
Energy correlation functions are complete graphs

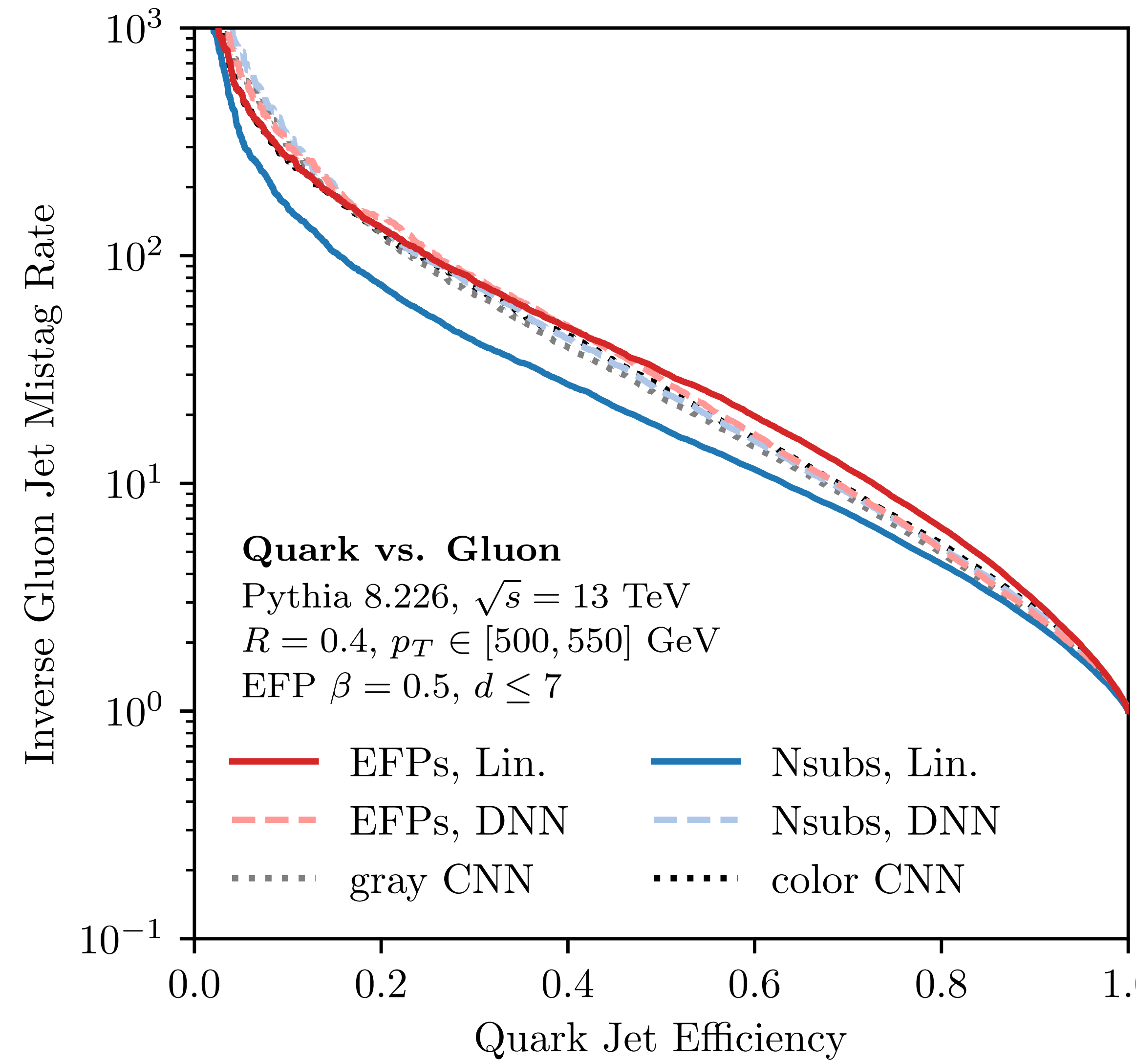
Even angularities are exact linear combinations of EFPs

EFPs organized by degree d – number of edges

Degree	Connected Multigraphs
$d = 1$	
$d = 2$	
$d = 3$	
$d = 4$	
$d = 5$	

EFPs for Quark vs. Gluon Jets





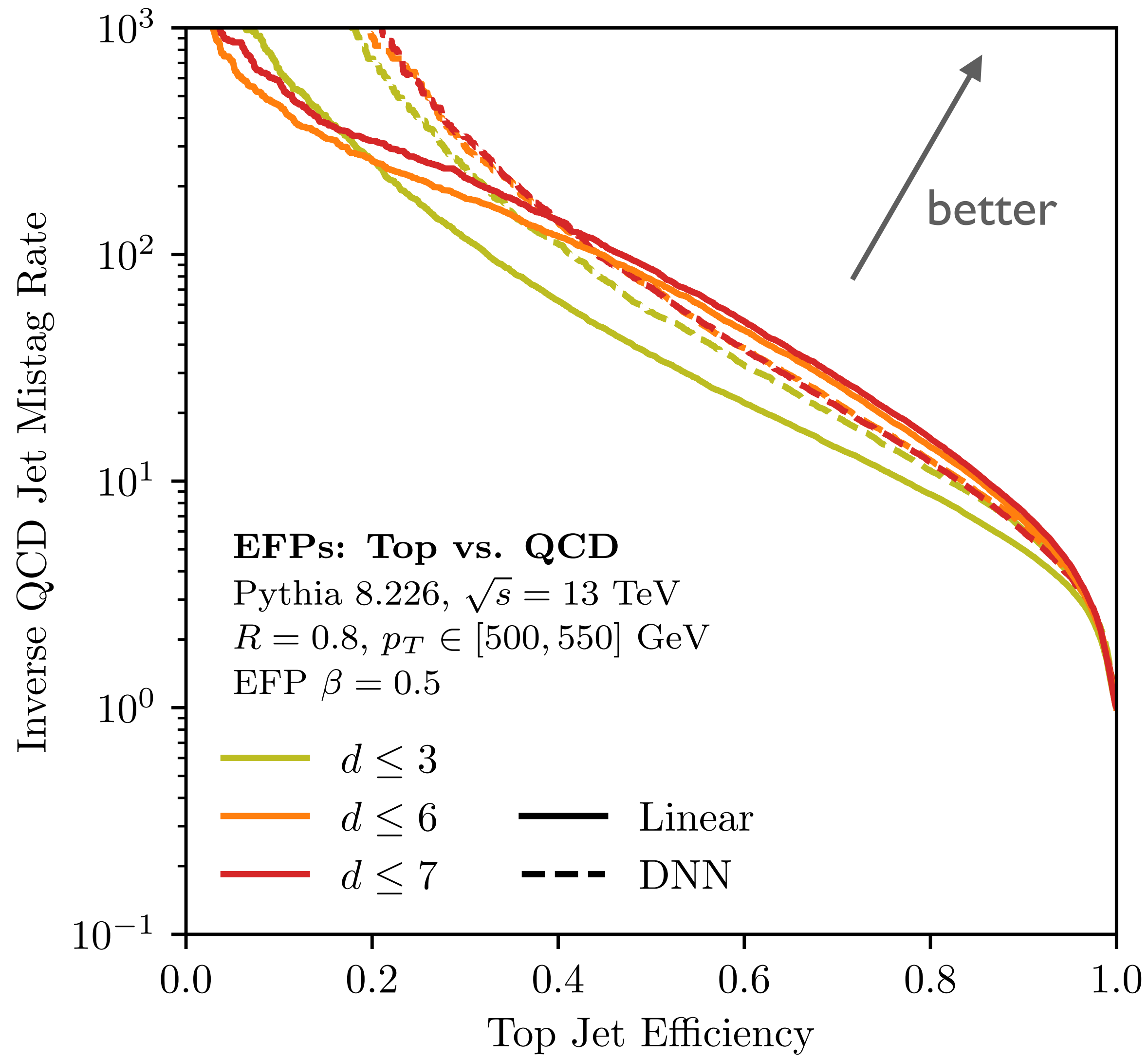
Saturation observed with more EFPs

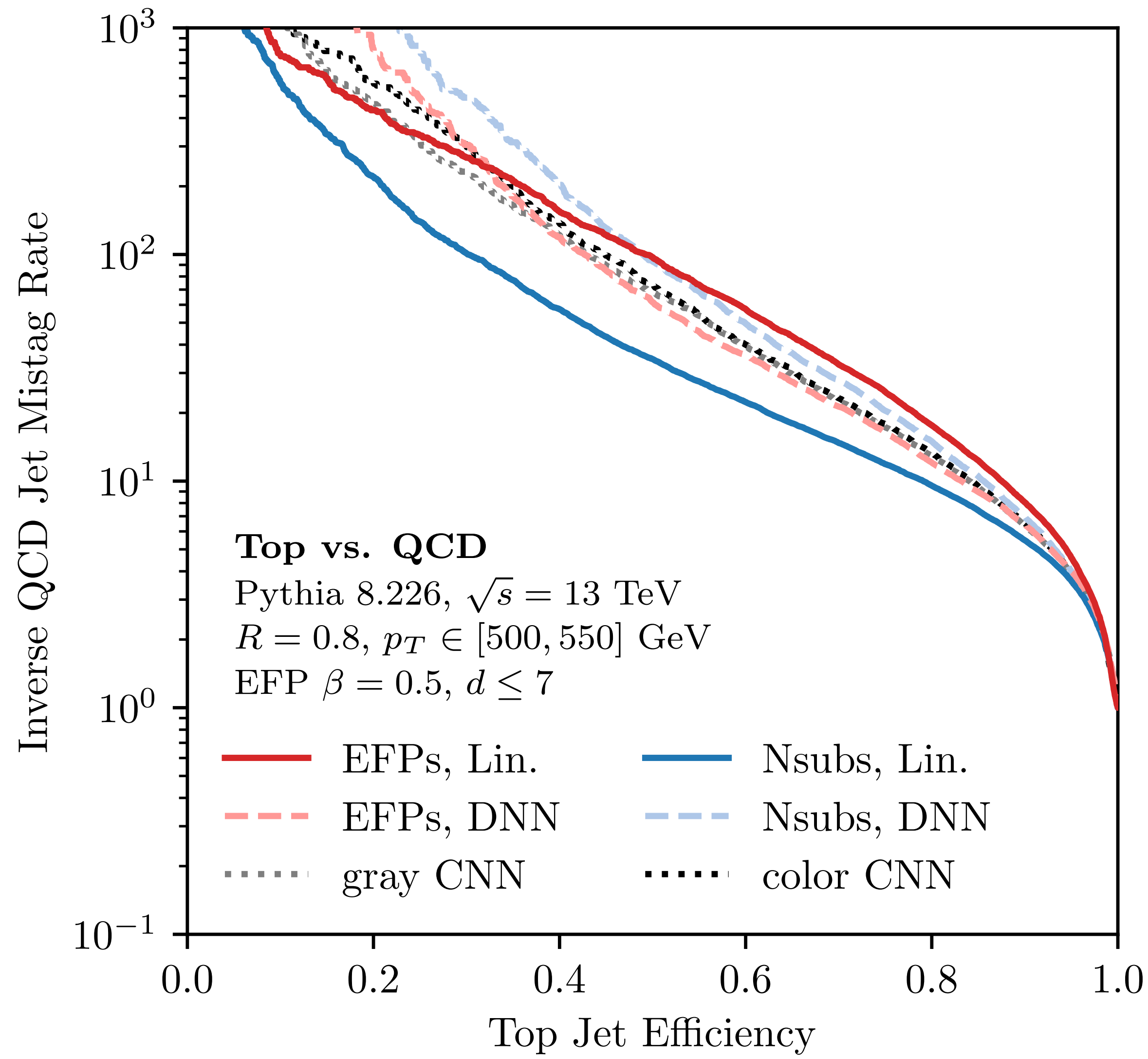
Linear EFPs excel at high efficiency

DNN gets there faster but linear suffices

[de Oliviera, Kagan, Mackey, Nachman, Schwartzman, 2015]
[PTK, Metodiev, Schwartz, 2016]
[Datta, Larkoski, 2017]

EFPs for Boosted Tops



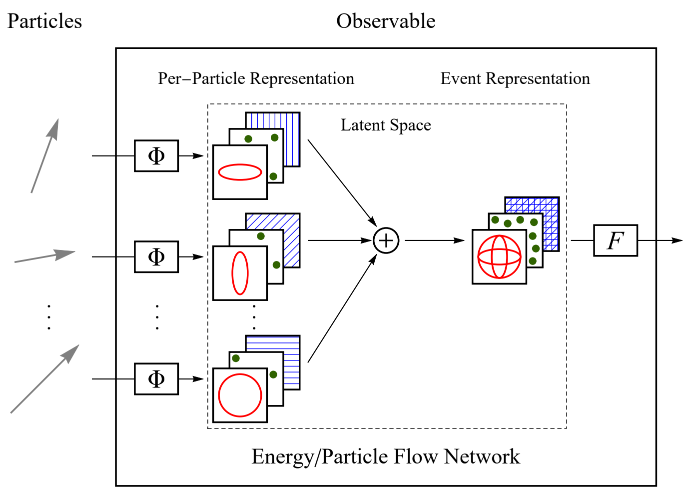


Saturation observed with more EFPs

Linear EFPs excel at high efficiency

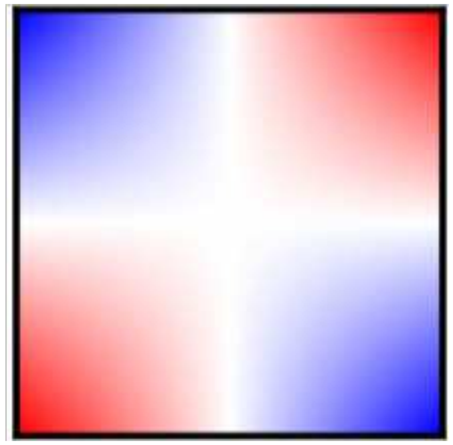
DNN gets there faster but linear suffices

[de Oliviera, Kagan, Mackey, Nachman, Schwartzman, 2015]
[PTK, Metodiev, Schwartz, 2016]
[Datta, Larkoski, 2017]



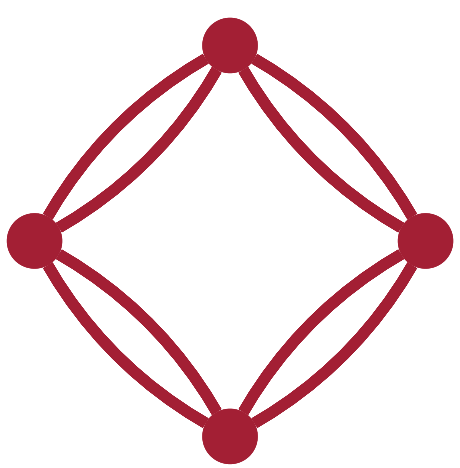
Energy Flow Networks

Jet symmetries, point clouds, Deep Sets, performance, versatility, simplicity, visualization, new analytic observables



Energy Flow Moments

Linear in M computation of EFPs, additivity, algebraic identities



Energy Flow Polynomials

Linear basis of IRC-safe observables

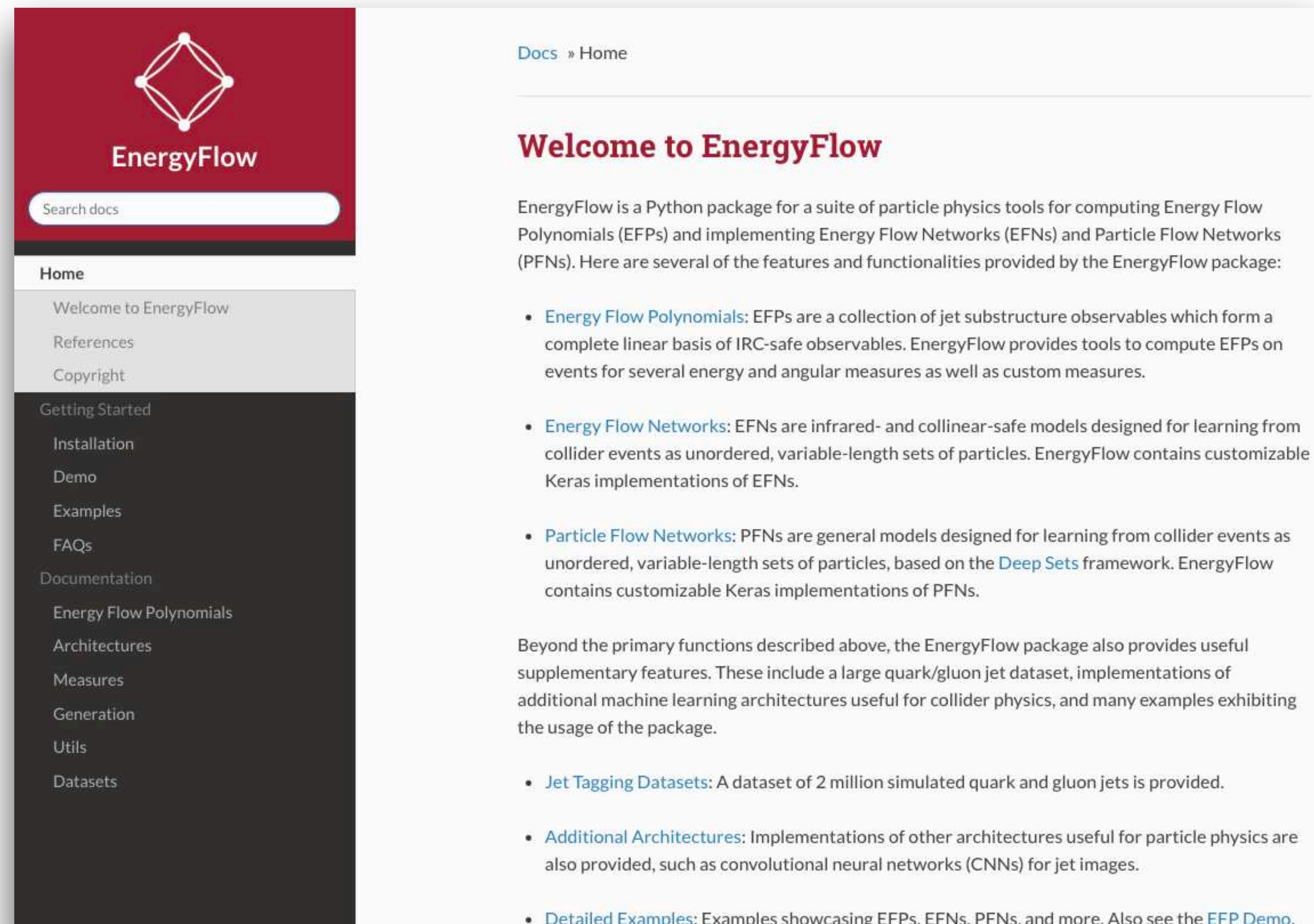
EnergyFlow Python Package

Implements variable elimination for efficient EFP computation

Contains EFN and PFN implementations in Keras

CNN, DNN architectures included for easy model comparison

Several detailed examples demonstrating how to train models and make visualizations



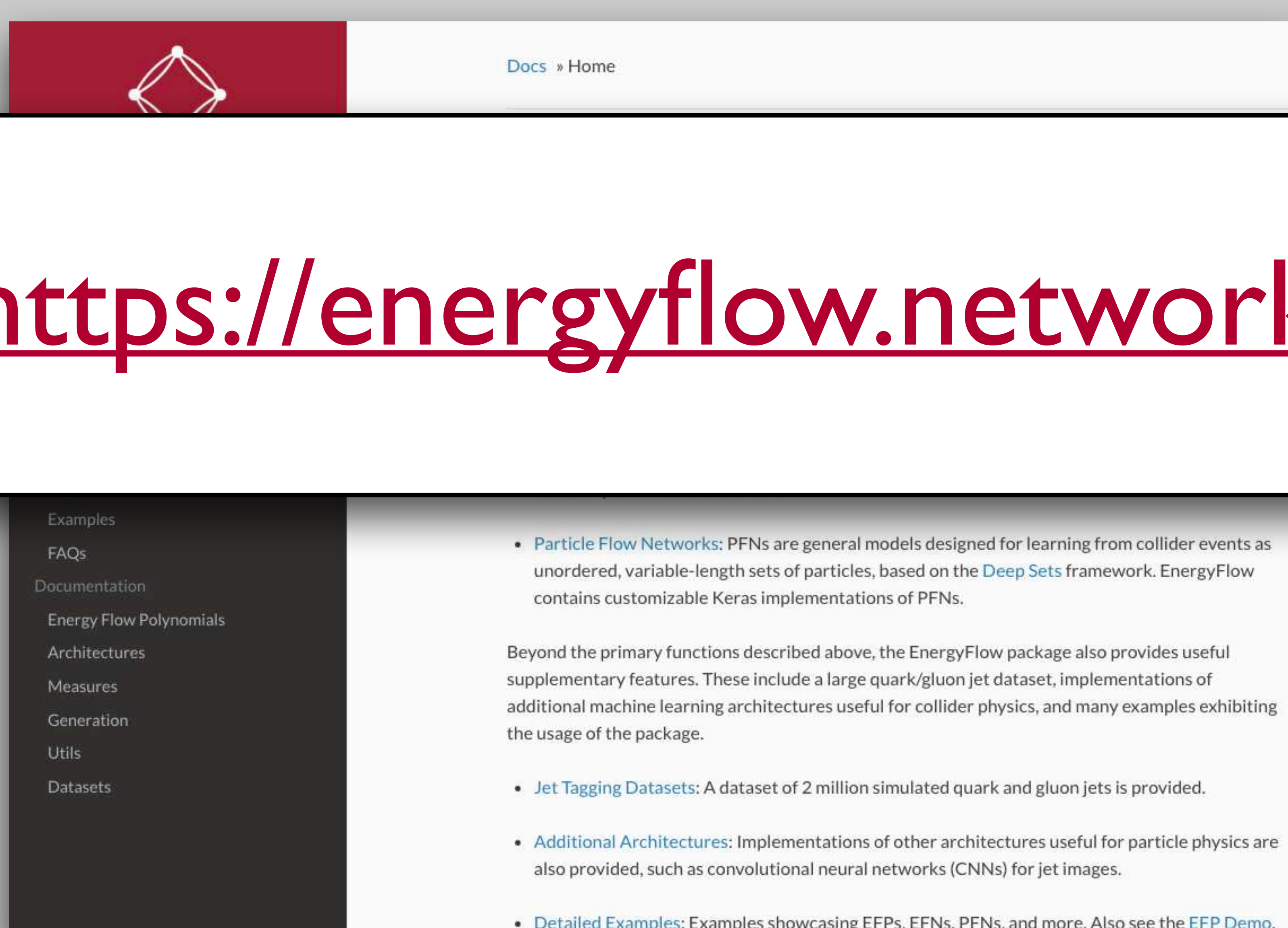
EnergyFlow Python Package

Contains EFN and PFN implementations in Keras

CNN, DNN architectures included for easy model comparison

Includes quark/gluon jet samples used in [1810.01565]

Several detailed examples demonstrating how to train models and make visualizations



https://energyflow.network

Thank You!

Computation Complexity of EFPs – Variable Elimination

Naive computation complexity of an energy correlator is $\mathcal{O}(M^N)$

For ~100 particles this becomes intractable for $N > 4$

Computation Complexity of EFPs – Variable Elimination

Naive computation complexity of an energy correlator is $\mathcal{O}(M^N)$

For ~100 particles this becomes intractable for *N* > 4

⇨ EnergyCorrelator `fjcontrib` package gives up in this case

```
// if N > 5, then throw error
if (_N > 5) {
    throw Error("EnergyCorrelator is only hard coded for N = 0,1,2,3,4,5");
}
```

Computation Complexity of EFPs – Variable Elimination

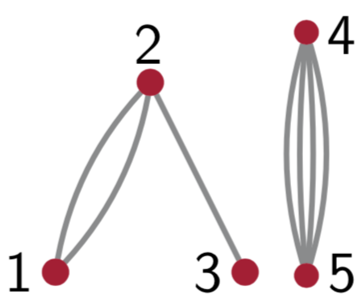
Naive computation complexity of an energy correlator is $\mathcal{O}(M^N)$

For ~100 particles this becomes intractable for $N > 4$

⇨ EnergyCorrelator fjcontrib package gives up in this case

```
// if N > 5, then throw error
if (_N > 5) {
   throw Error("EnergyCorrelator is only hard coded for N = 0,1,2,3,4,5");
}
```

Variable elimination (VE) algorithm can speedup EFPs by finding efficient elimination ordering



$$= \left(\sum_{i_1=1}^{M}\sum_{i_2=1}^{M}\sum_{i_3=1}^{M} z_{i_1} z_{i_2} z_{i_3} \theta_{i_1 i_2}^2 \theta_{i_2 i_3}\right)\left(\sum_{i_4=1}^{M}\sum_{i_5=1}^{M} z_{i_4} z_{i_5} \theta_{i_4 i_5}^4\right)$$

Disconnected is product of connected

$$= \underbrace{\sum_{i_1=1}^{M}\sum_{i_2=1}^{M}\sum_{i_3=1}^{M}\sum_{i_4=1}^{M}\sum_{i_5=1}^{M}\sum_{i_6=1}^{M}\sum_{i_7=1}^{M}\sum_{i_8=1}^{M} z_{i_1} z_{i_2} z_{i_3} z_{i_4} z_{i_5} z_{i_6} z_{i_7} z_{i_8} \prod_{j=2}^{7} \theta_{i_1 i_j}}_{\mathcal{O}(M^8)}$$

$$= \underbrace{\sum_{i_1=1}^{M} z_{i_1} \left(\sum_{i_2=1}^{M} z_{i_2} \theta_{i_1 i_2}\right)^7}_{\mathcal{O}(M^2)}$$

Clever parentheses placement corresponds to good elimination ordering

All tree graphs become $\mathcal{O}(M^2)$