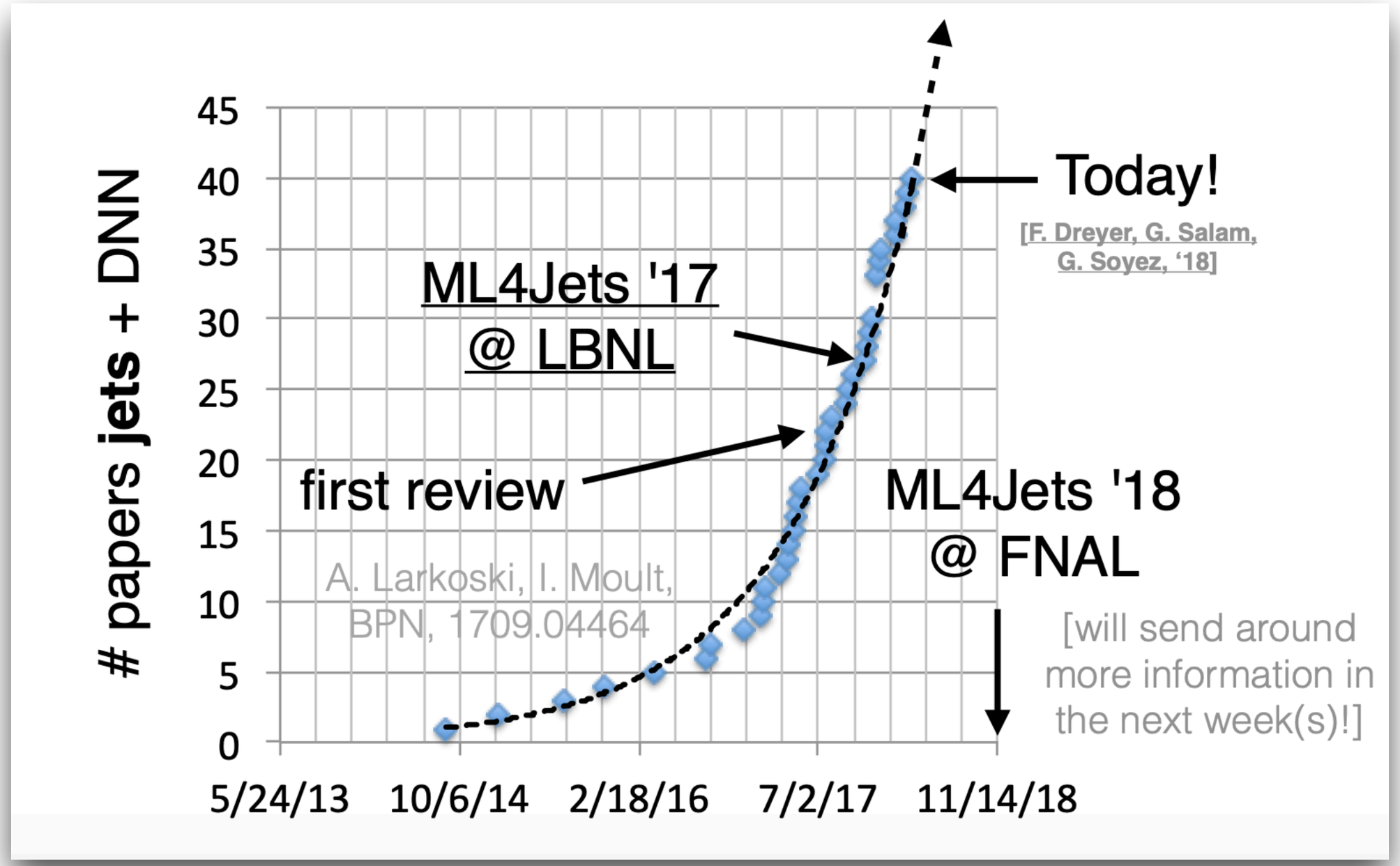
Latest Developments in Machine Learning for Jets

Anders Andreassen

LBNL/UC Berkeley

ATLAS HFSF - December 11, 2018

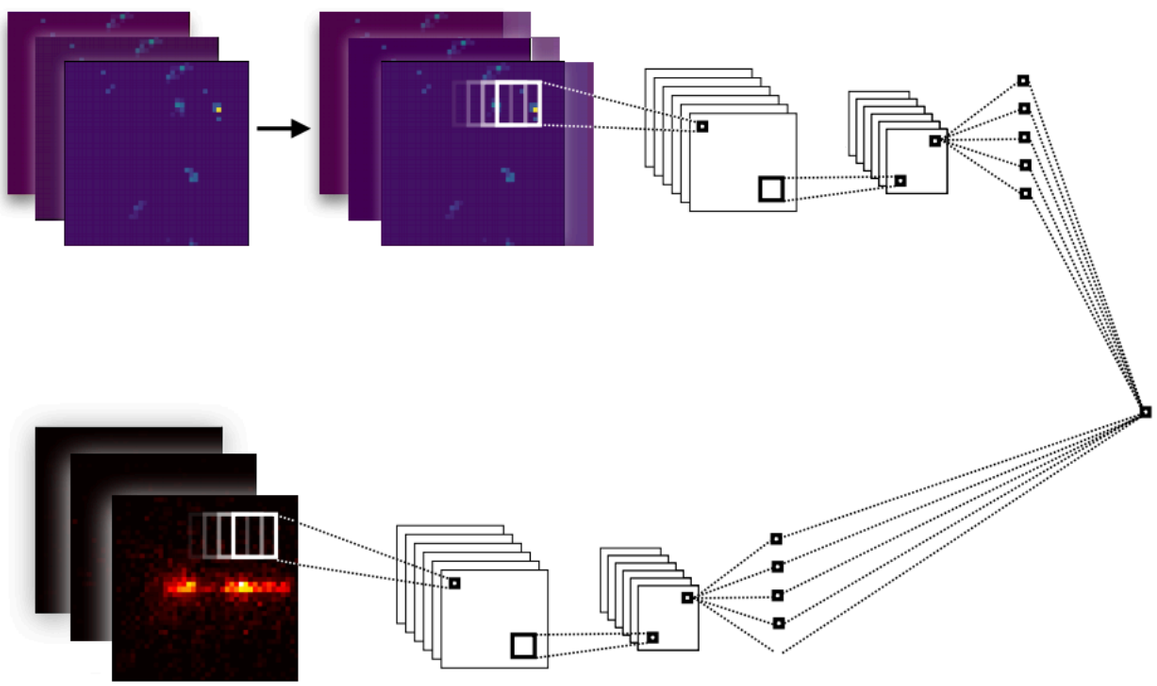
Machine Learning for Jets is a rapidly growing field of research

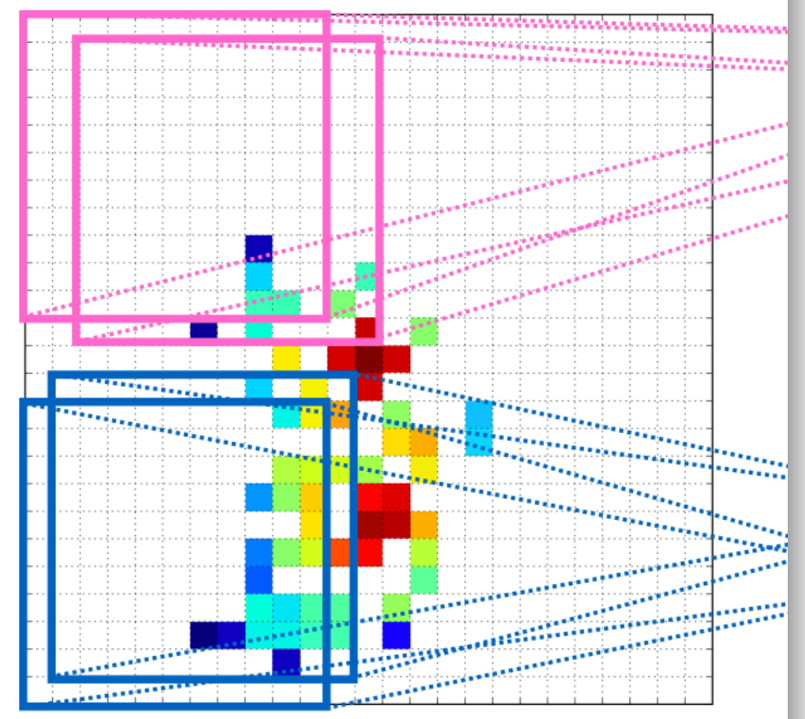


Nachman, BOOST 2018 Talk, July 20, 2018

Machine Learning for Jet Physics 2018

indico.cern.ch/event/ml4jets2018





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

November 14-16, 2018

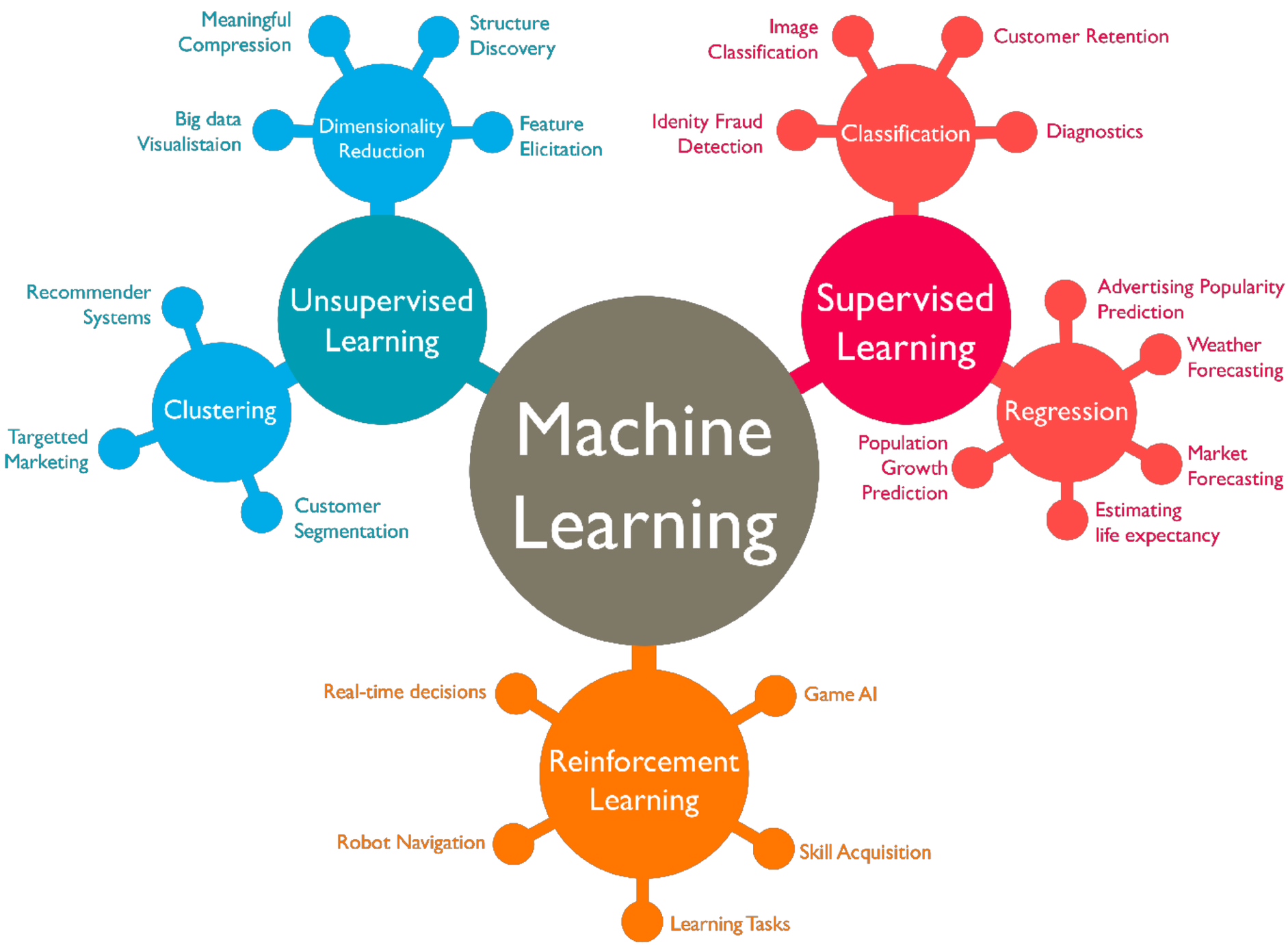
Fermilab

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)

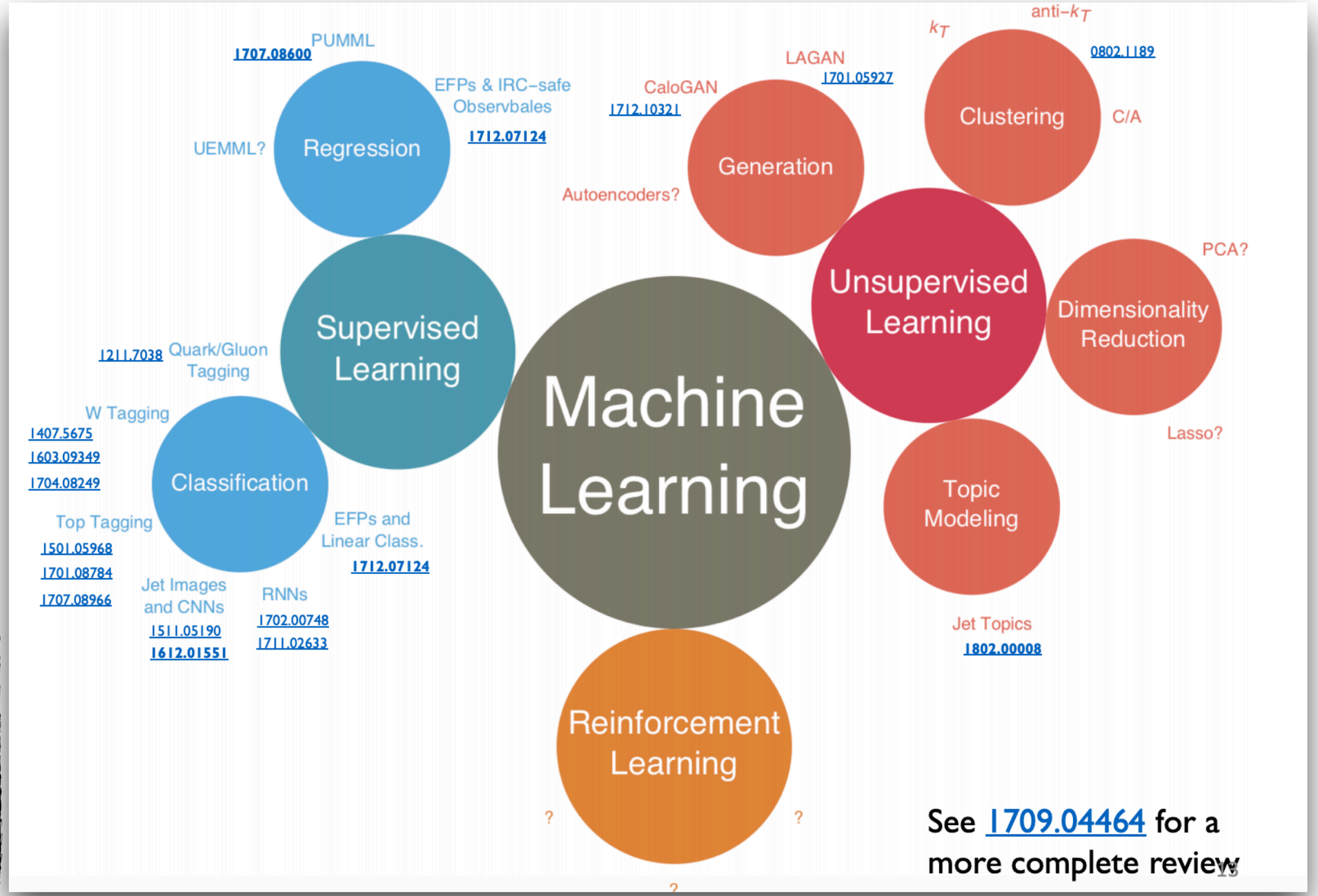
LPC Coordinators:
Cecilia Gerber (UIC)
Sergo Jindariani (Fermilab)

Overview: Machine Learning



Overview:
Machine Learning in HEP

Most HEP focus on classification



Still lots of interesting new ideas for classification at ML4Jets '18

Patrick T. Komiske and Eric M. Metodiev, Harvard Feb, 2018

Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT
- Jet Images

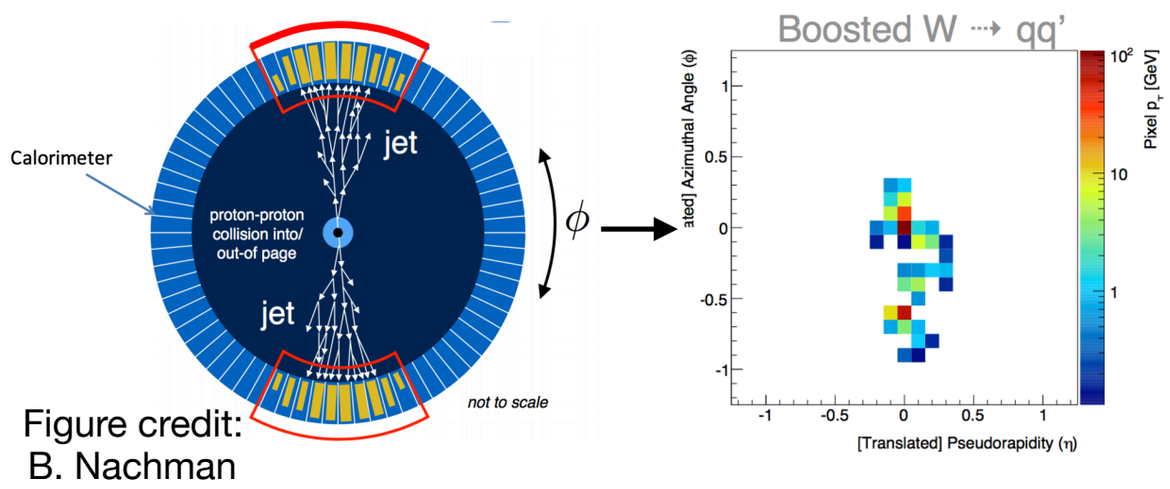


Figure credit: B. Nachman

- Sequences
  - pT ordering
  - Clustering history

Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT

- Jet Images

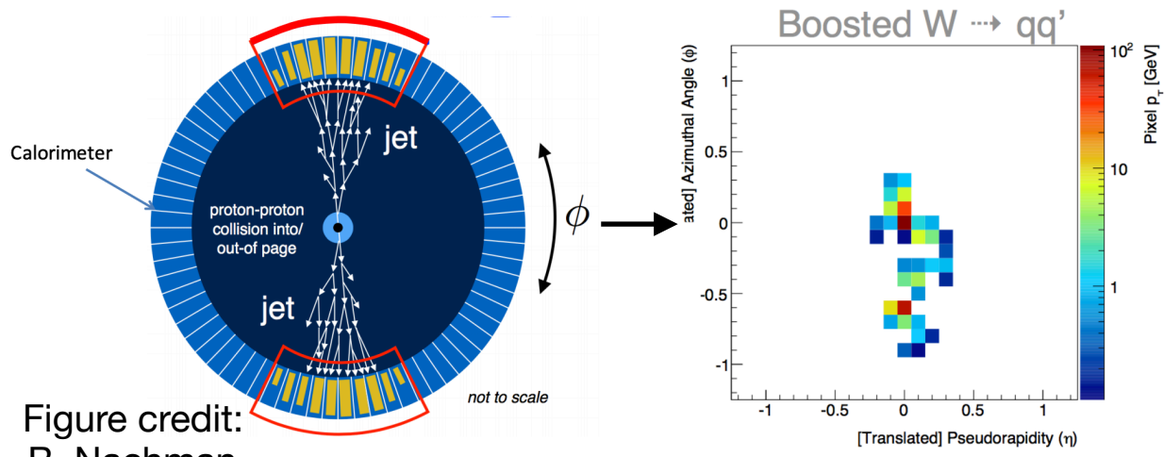
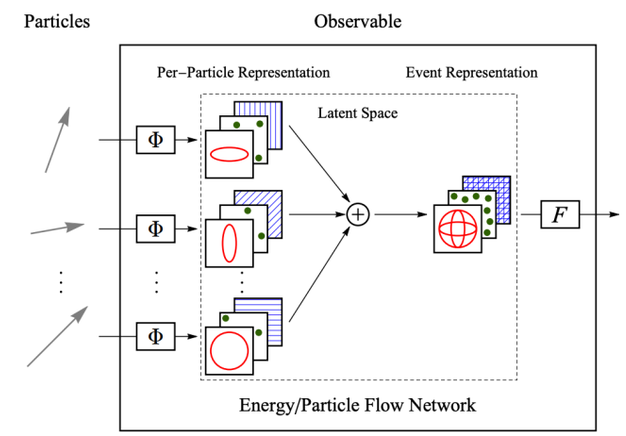


Figure credit: B. Nachman

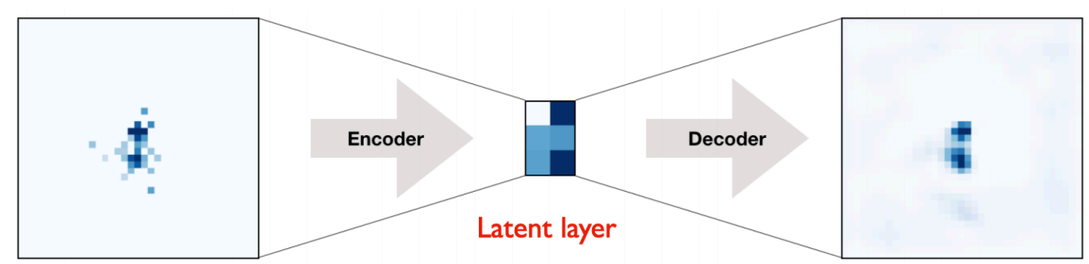
- Sequences
  - pT ordering
  - Clustering history

- Energy/Particle Flow Network

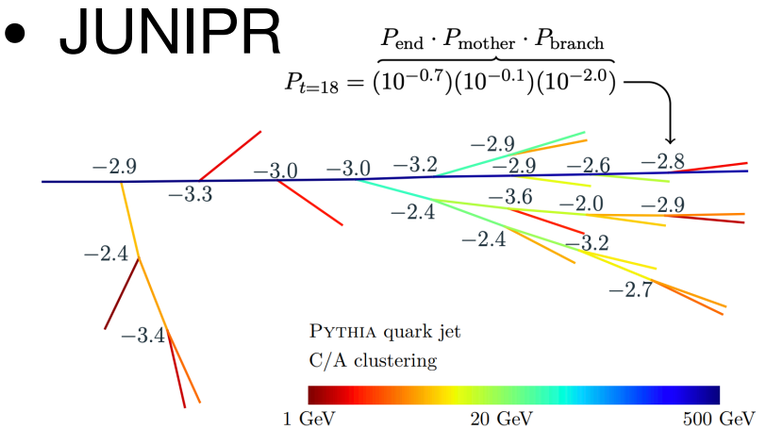


Komiske, Metodiev & Thaler (2018)

- Autoencoders
  - Farina, Nakai, Shih (2018)
  - Heimel, Kasieczka, Plehn, Thompson (2018)



- JUNIPR

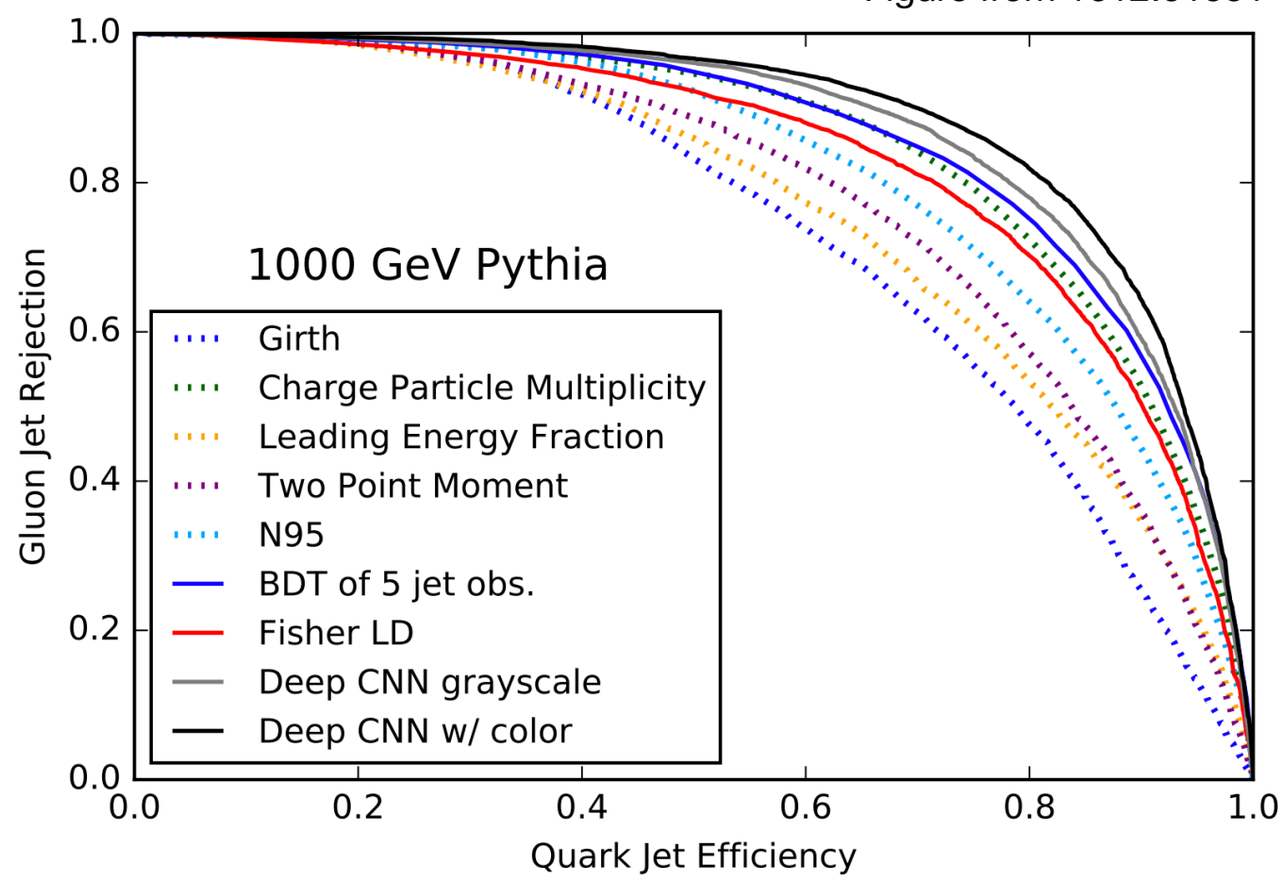


AA, Feige, Frye & Schwartz (2018)

Physics Motivated Inputs

- Input physics motivated observables to BDT or DNN
  - Mass, multiplicity, girth, etc.
- It is a natural choice, but are we throwing information out?
- Complete basis:
  - N-subjettines observables (see 1704.08249)
  - Energy Flow Polynomials (see 1712.07124)

Figure from 1612.01551



See also:
ATL-PHYS-PUB-2017-004
ATL-PHYS-PUB-2017-013
ATLAS-CONF-2017-064

Jet Images

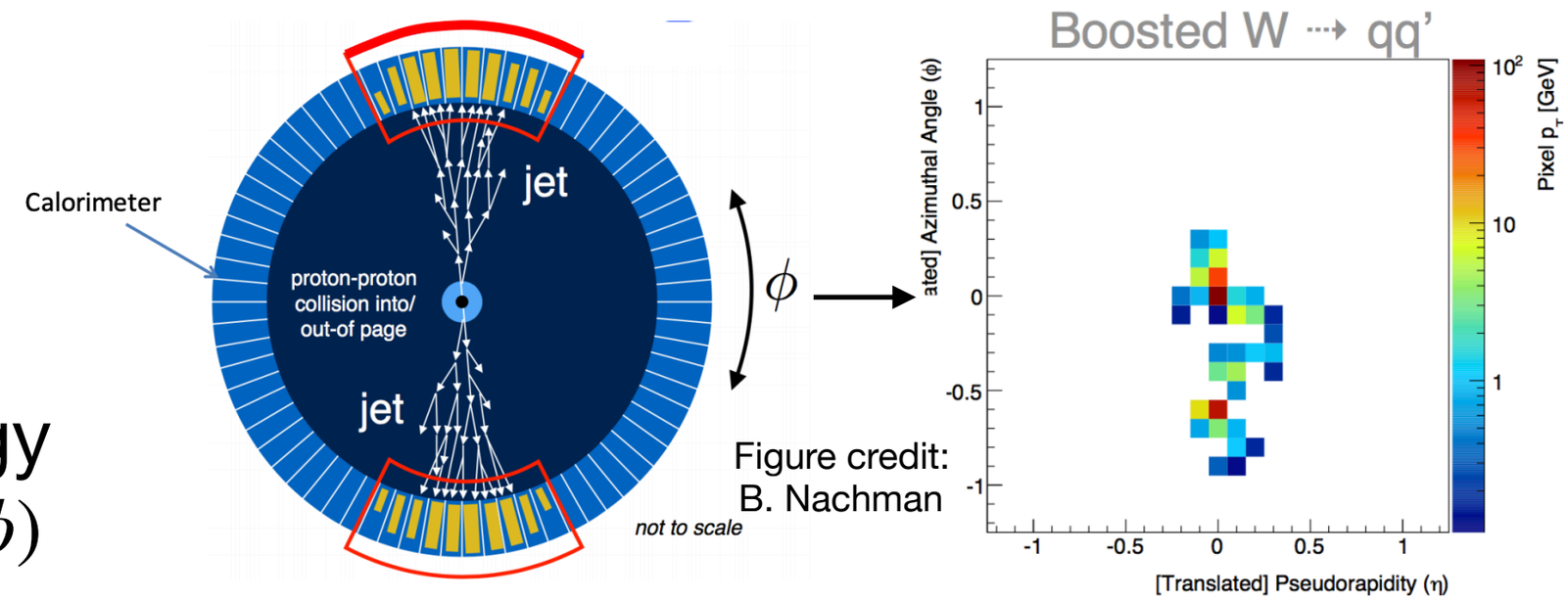


Figure credit: B. Nachman

- Discretized energy into pixels in (η, ϕ)
- Typically very sparse
- Captures spatial correlations
- Fixed dimensions of jet representation

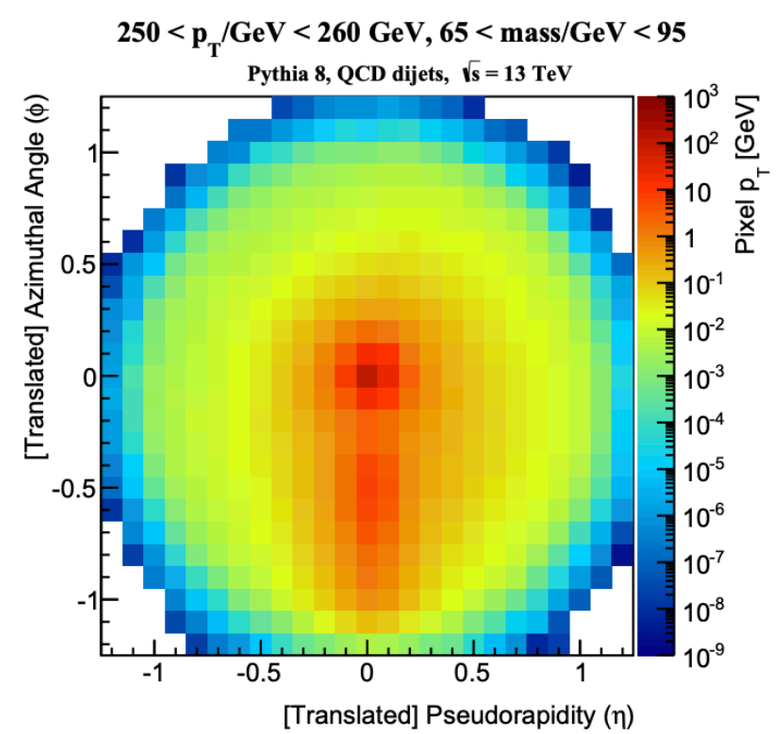


Figure from 1511.05190

Cogan et al 1407.5675; de Oliveira et al 1511.05190
Almeida et al 1501.05968; Komiske et al 1612.01551
Baldi et al 1603.09349; Barnard et al 1609.00607;
Kasieczka et al 1701.08784

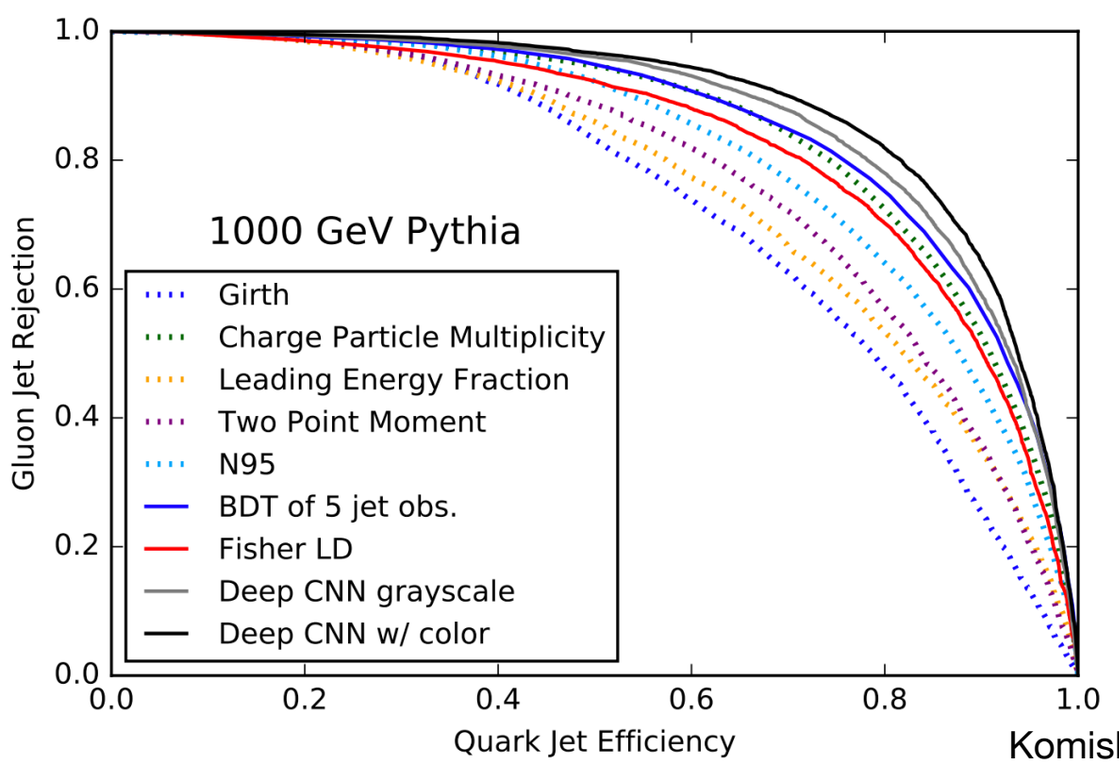
Jet Images

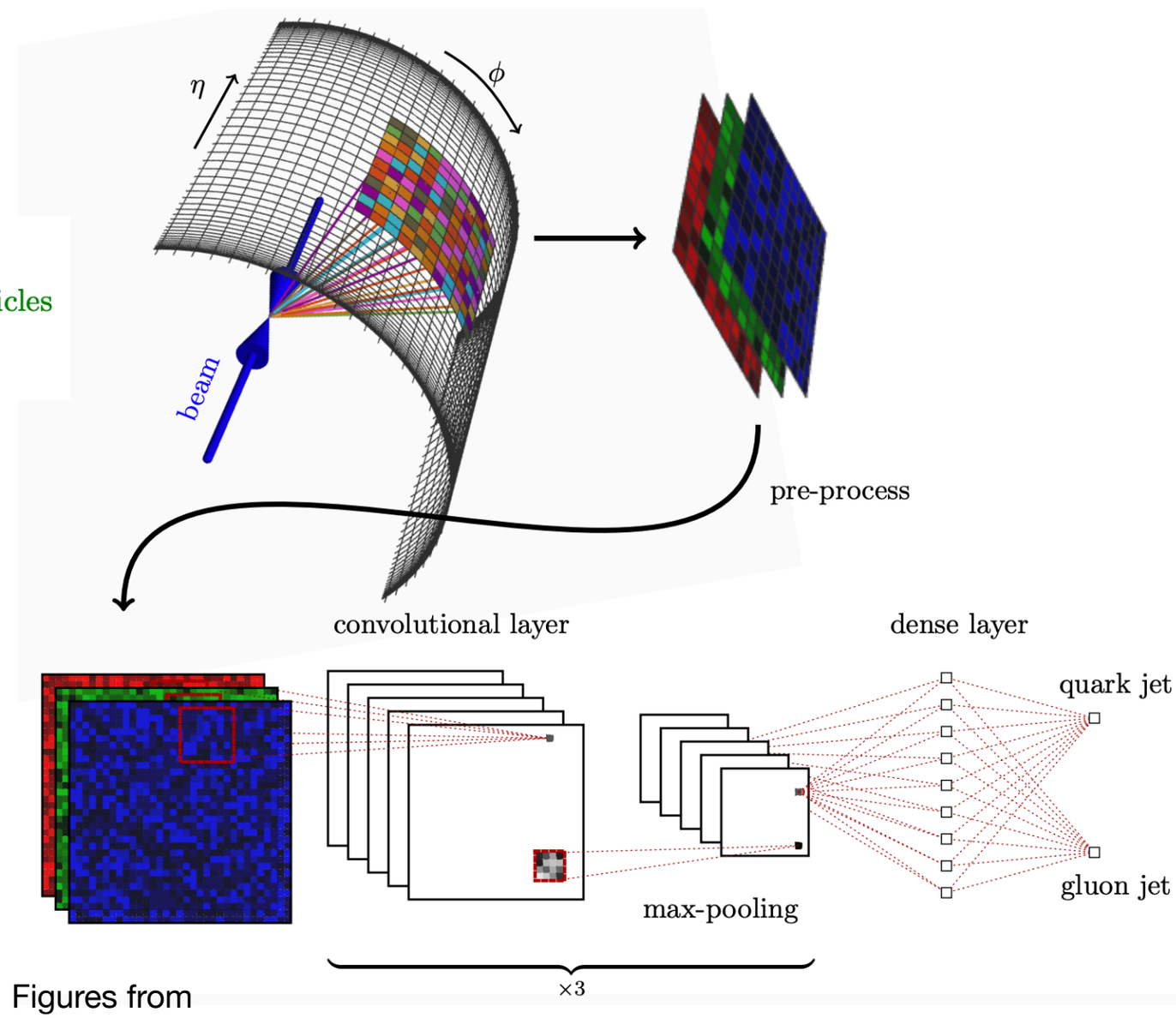
- Convolutional Neural Networks (CNN)
- Multiple channels

red = transverse momenta of charged particles

green = the transverse momenta of neutral particles

blue = charged particle multiplicity

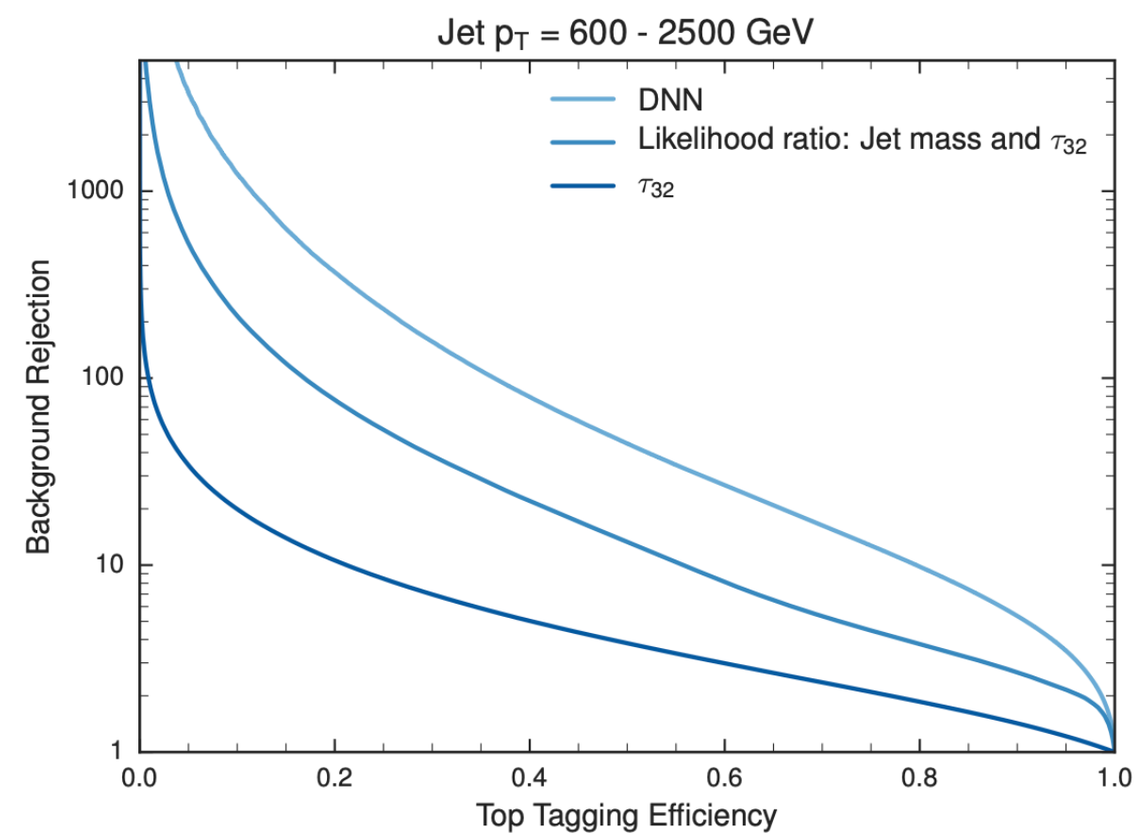




Figures from Komiske et al 1612.01551

Jets as Sequences

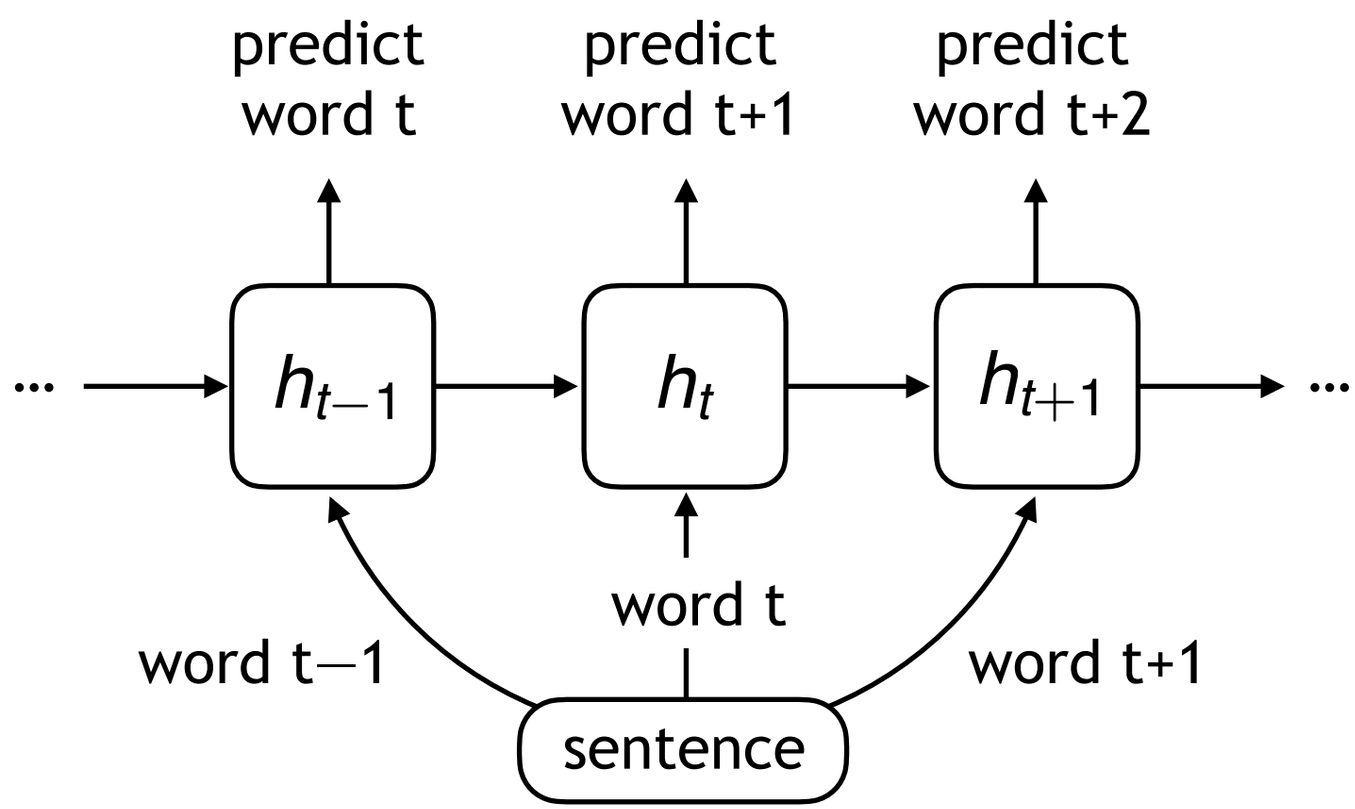
- Jet $= \{p_1^\mu, p_2^\mu, p_3^\mu, \cdots, p_M^\mu\}$
- NN with 4-momenta as input?
- Variable length
  - Keep the N most energetic particles (see e.g. 1704.02124)
  - **Recurrent Neural Network**

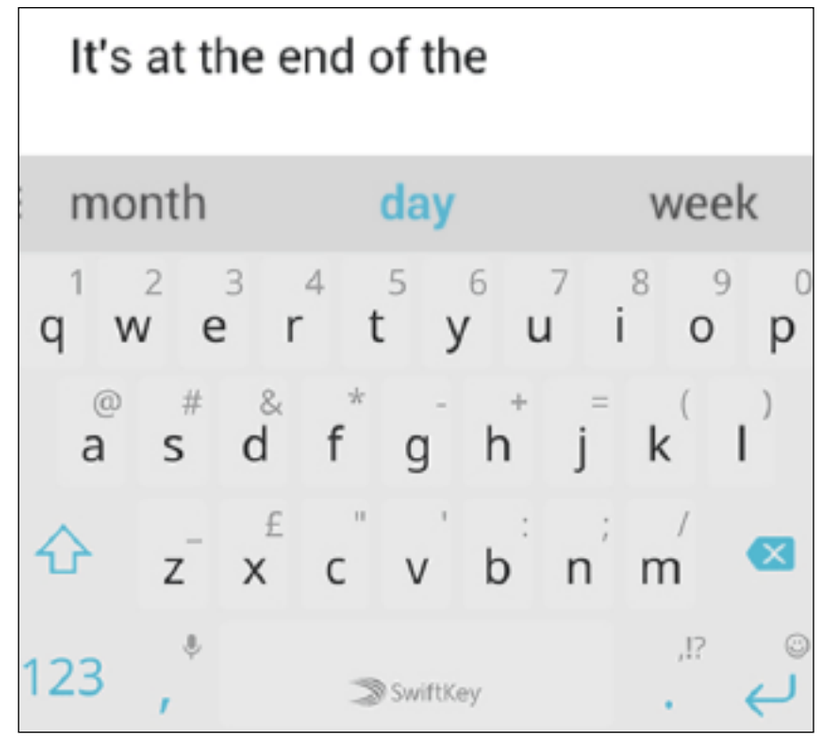


Louppe et al 1702.00748; Cheng 1711.02633
Egan et al 1711.09059; Fraser et al 1803.08066

Recurrent Neural Networks (RNNs)

- naturally model sequential evolution (e.g. language)
- allow indeterminate number of “time steps”



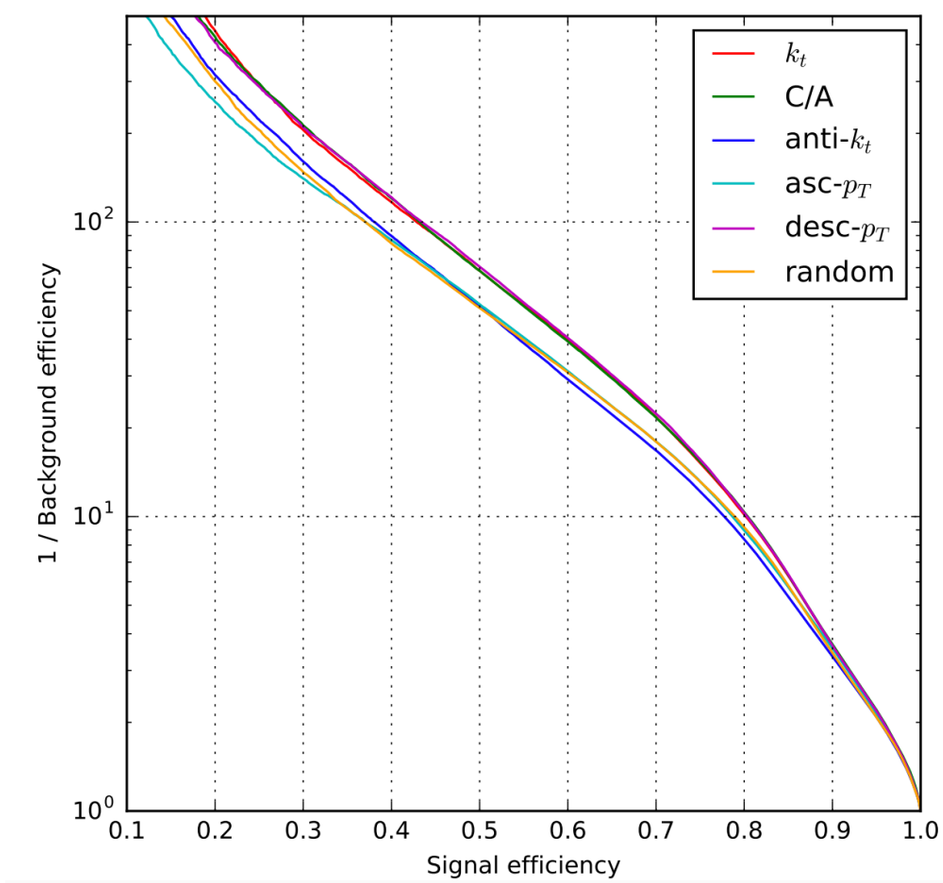


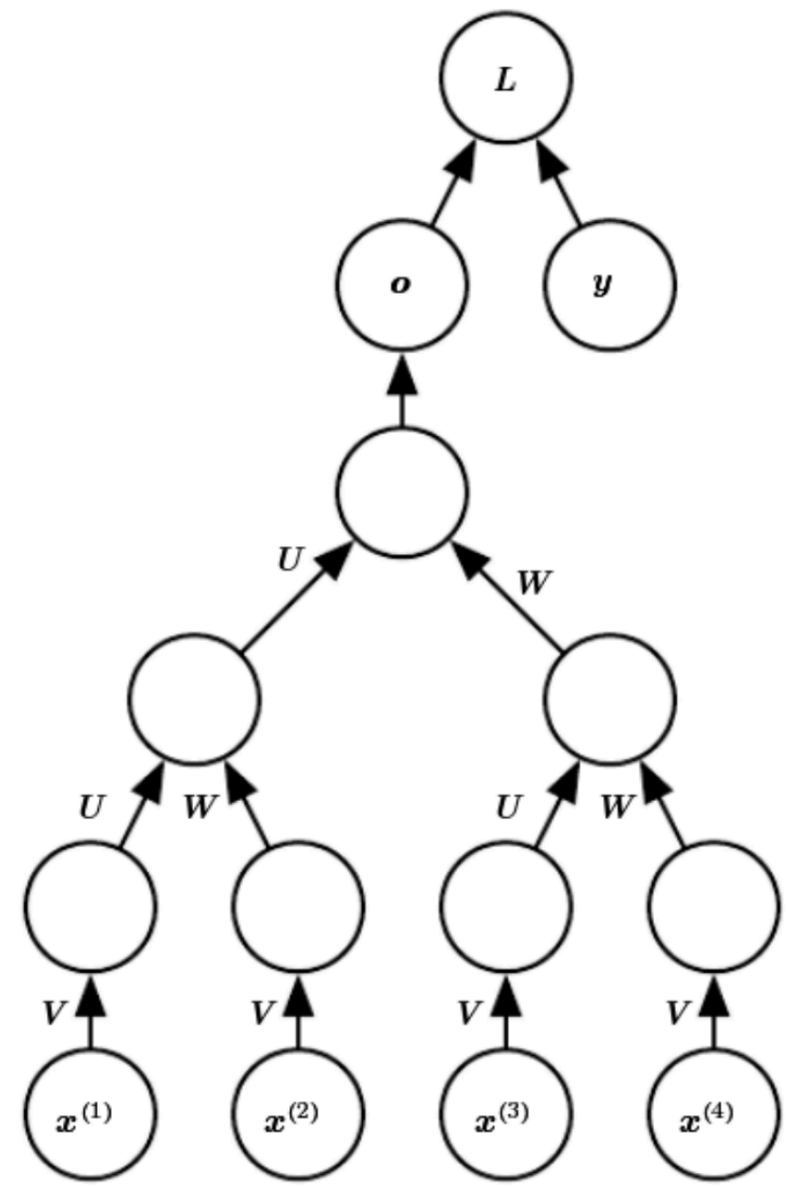
Perfect for modeling jet evolution:

$$\begin{cases} \text{sentence} \longleftrightarrow \text{jet} \\ \text{next word} \longleftrightarrow \text{next splitting} \end{cases}$$

Jets as Sequences

- Arbitrary choice of ordering
  - pT ordering
  - clustering history





Louppe et al 1702.00748

Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT
- Jet Images

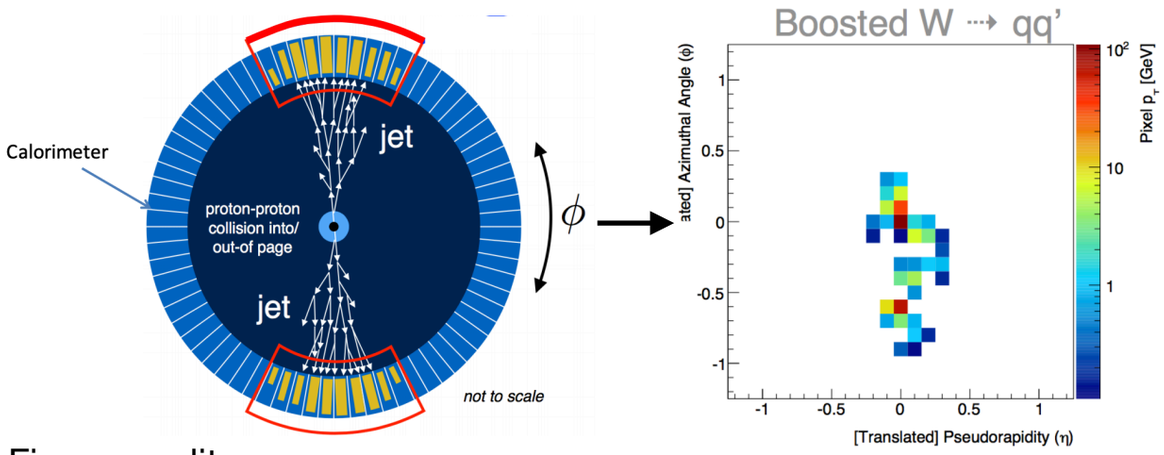
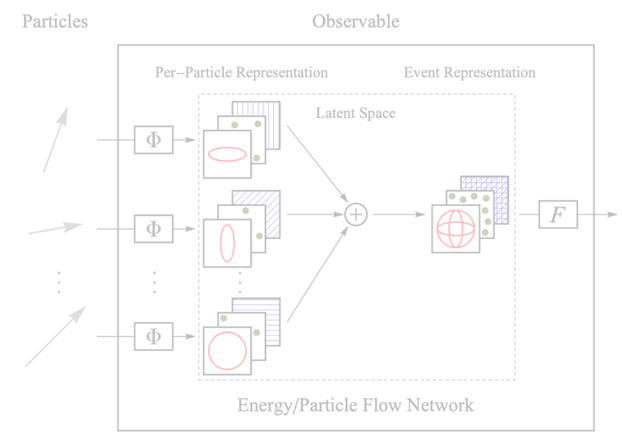


Figure credit: B. Nachman

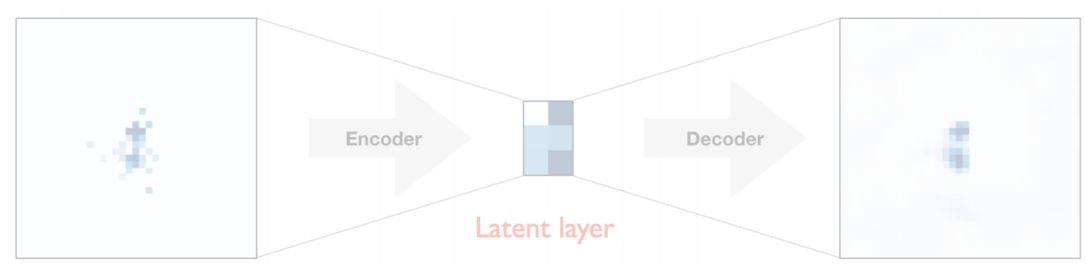
- Sequences
  - pT ordering
  - Clustering history

- Energy/Particle Flow Network

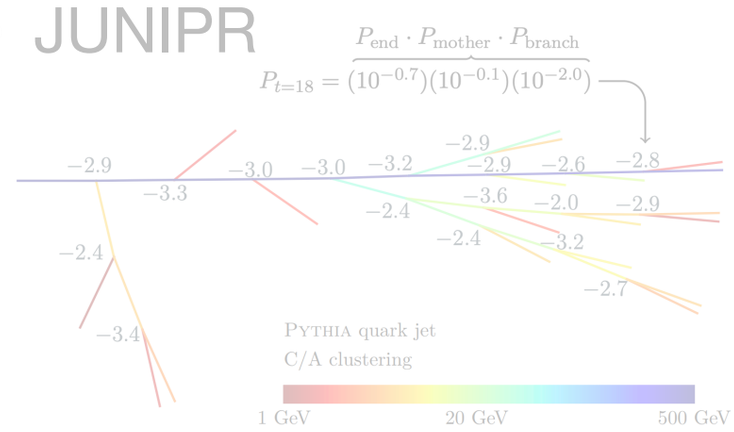


Komiske, Metodiev & Thaler (2018)

- Autoencoders
  - Farina, Nakai, Shih (2018)
  - Heimel, Kasieczka, Plehn, Thompson (2018)



- JUNIPR



AA, Feige, Frye & Schwartz (2018)

Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT
- Jet Images

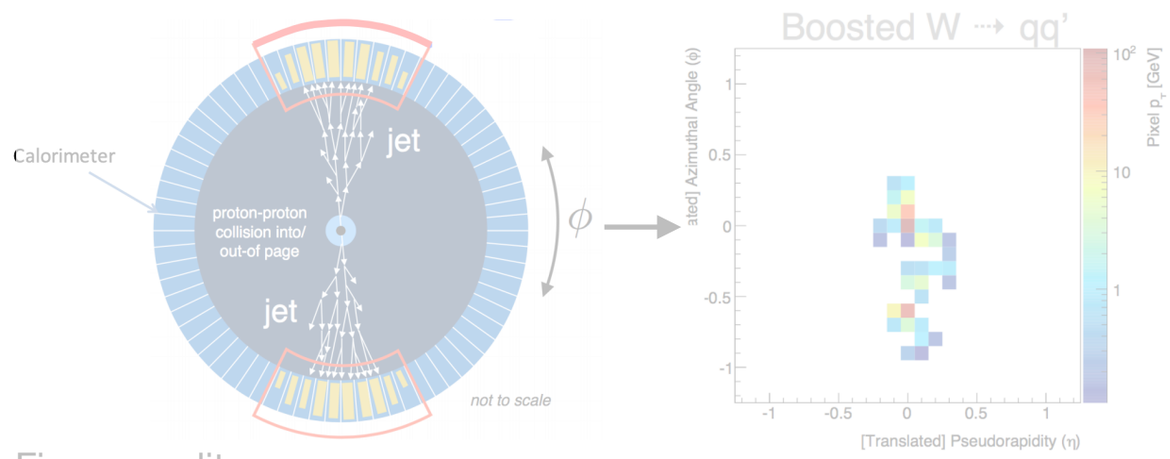
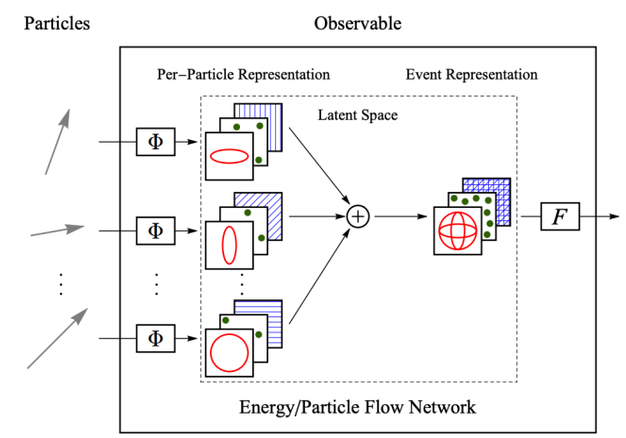


Figure credit: B. Nachman

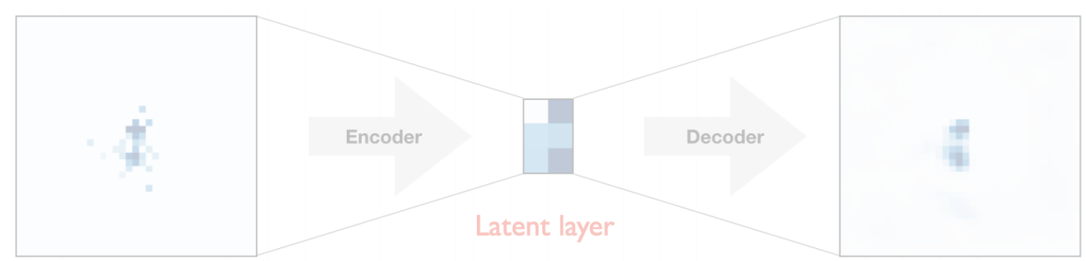
- Sequences
  - pT ordering
  - Clustering history

- Energy/Particle Flow Network

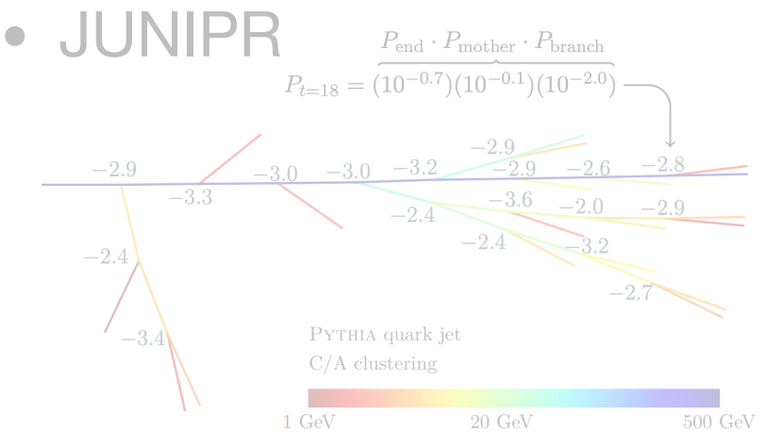


Komiske, Metodiev & Thaler (2018)

- Autoencoders
  - Farina, Nakai, Shih (2018)
  - Heimel, Kasieczka, Plehn, Thompson (2018)



- JUNIPR



AA, Feige, Frye & Schwartz (2018)

ML4Jets '18 @FNAL:

Energy Flow Networks: Deep Sets for Particle Jets

Patrick T. Komiske III

Massachusetts Institute of Technology
Center for Theoretical Physics

Machine Learning for Jet Physics Workshop

Fermilab, Illinois – 11/15/2018

Based on work with Eric Metodiev and Jesse Thaler

1810.05165

https://energyflow.network

Slide by Patrick Komiske

Deep Sets

What is a Jet?

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

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

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

$$\Phi : x_i \to \mathfrak{R}^\ell$$

$$F : \mathfrak{R}^\ell \to \mathfrak{R}$$

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

Holds for sufficiently large ℓ to arbitrary approximation [1703.06114]

Deep Sets for Particle Jets

[PTK, Metodiev, Thaler, 1810.05165]

Particle Flow Network (PFN)

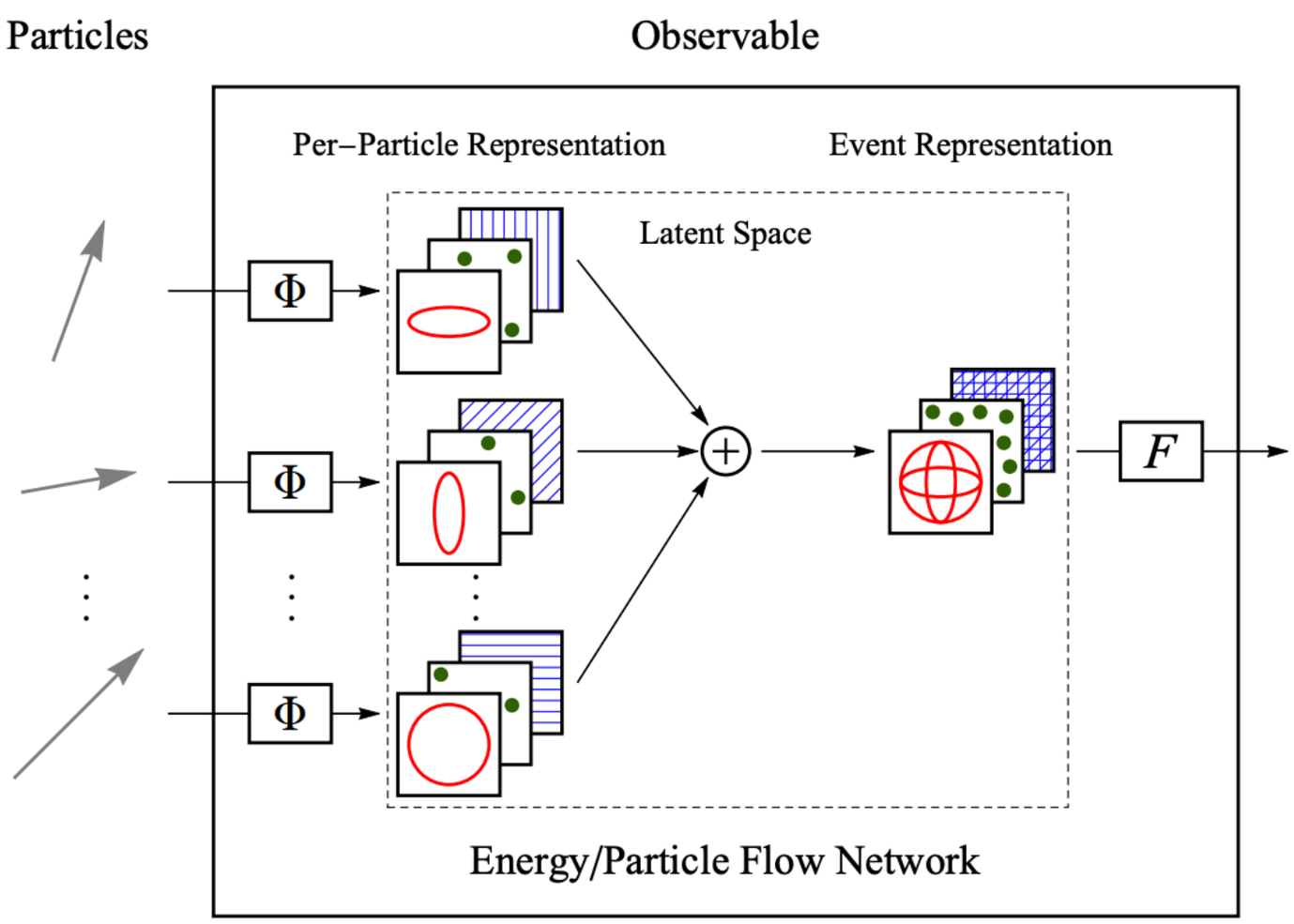
$$\mathrm{PFN}(\{p_1^\mu, \ldots, p_M^\mu\}) = F\left(\sum_{i=1}^{M} \Phi(p_i^\mu)\right)$$

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

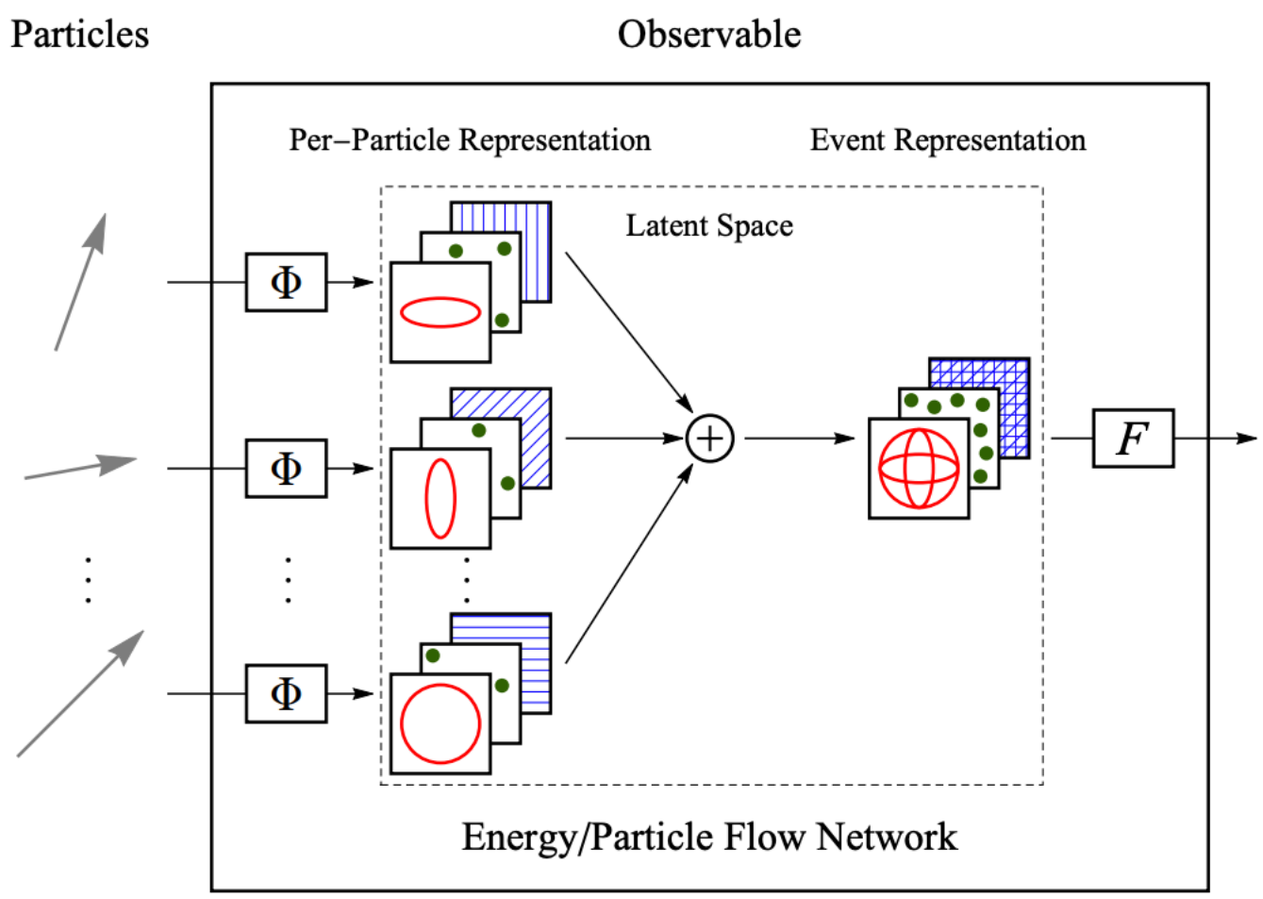


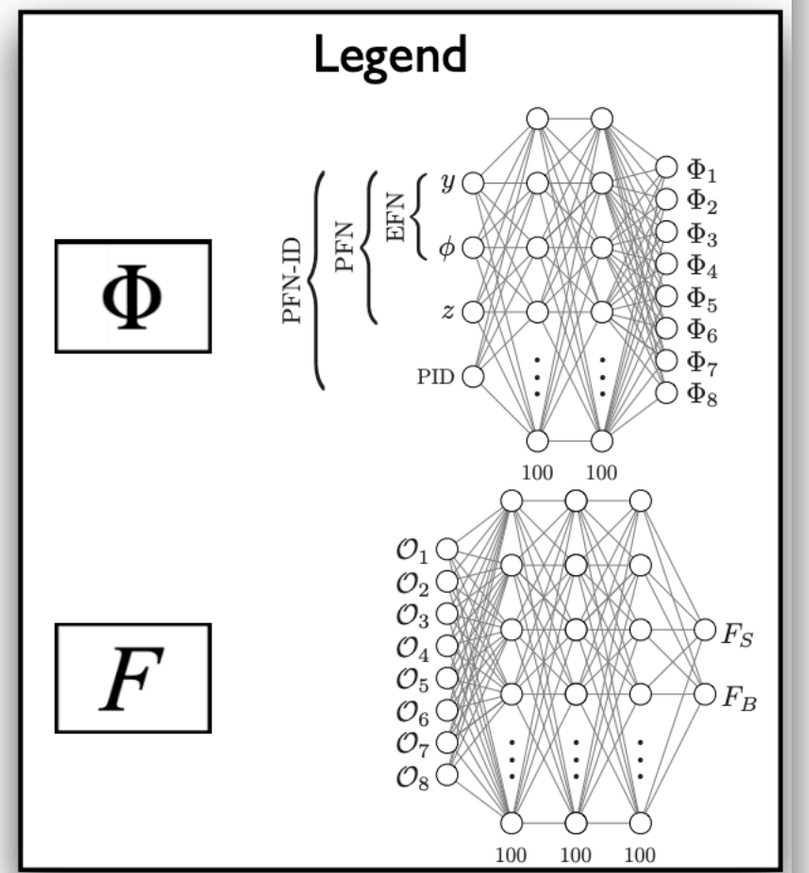
Approximating Φ and *F* with Neural Networks

Employ neural networks as arbitrary function approximators

Use fully-connected networks for simplicity

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





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

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

EFN Latent Dimension Sweep

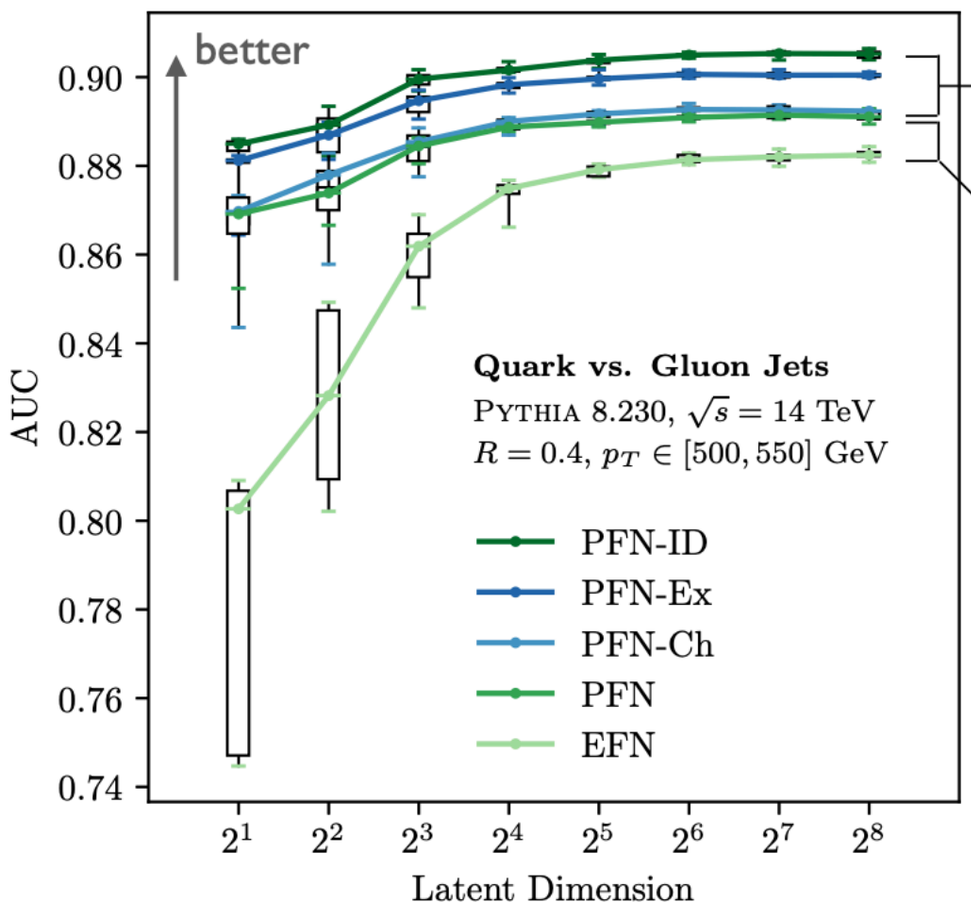
PFN-ID: Full particle flavor info
$(\gamma, \pi^{\pm}, K^{\pm}, K_L, p, \bar{p}, n, \bar{n}, e^{\pm}, \mu^{\pm})$

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

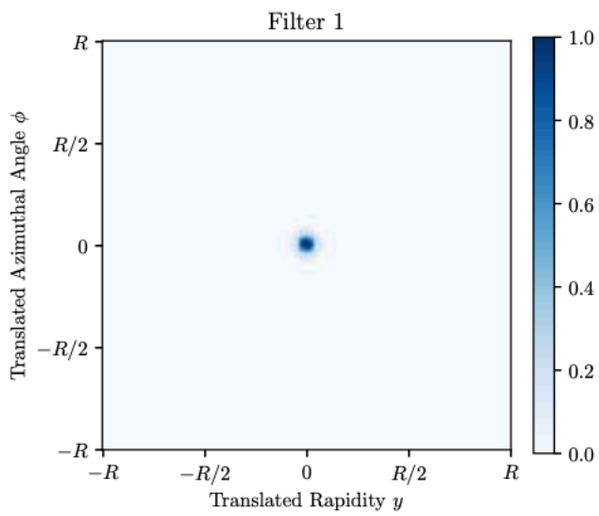


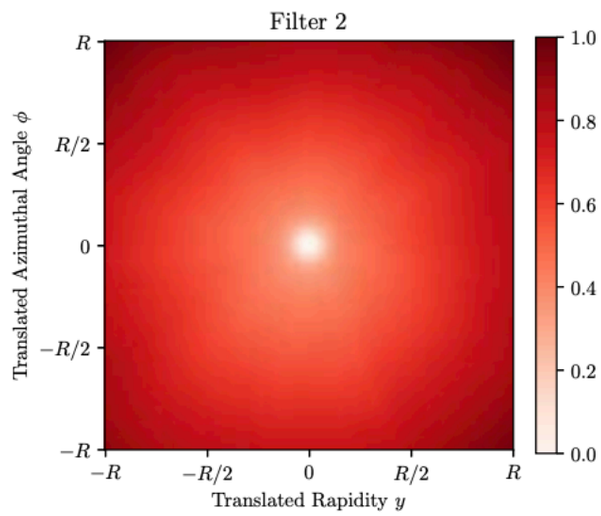
Particle type info helpful

IRC unsafe information clearly helpful

Performance saturates as dimension increases

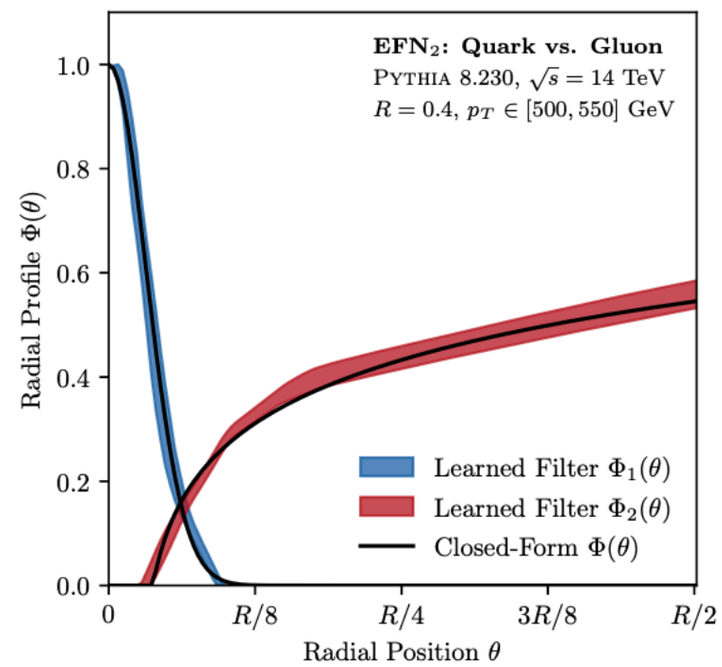
Extracting New Analytic Observables





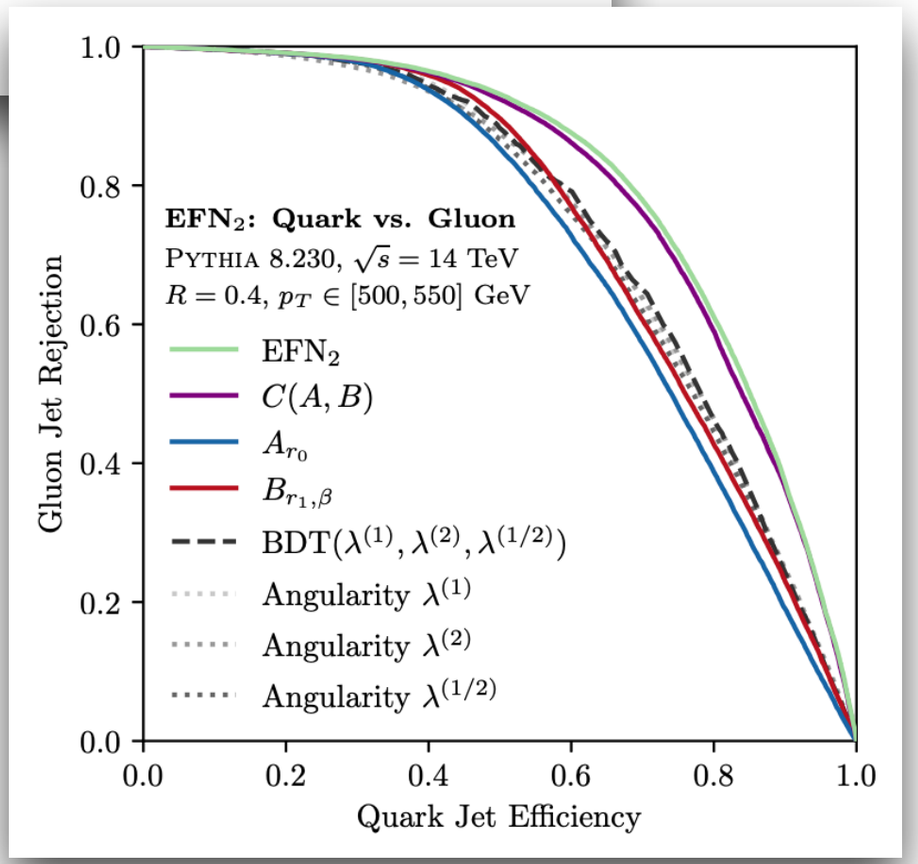
$$\mathcal{O}_1 = \sum_{i=1}^{M} z_i \, \Phi_1(\theta_i)$$

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

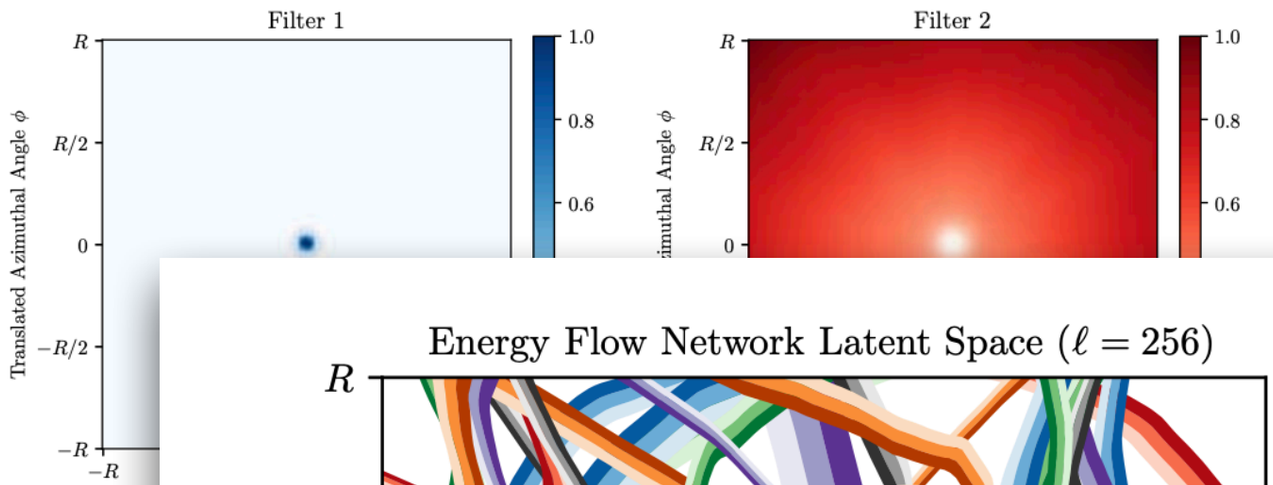


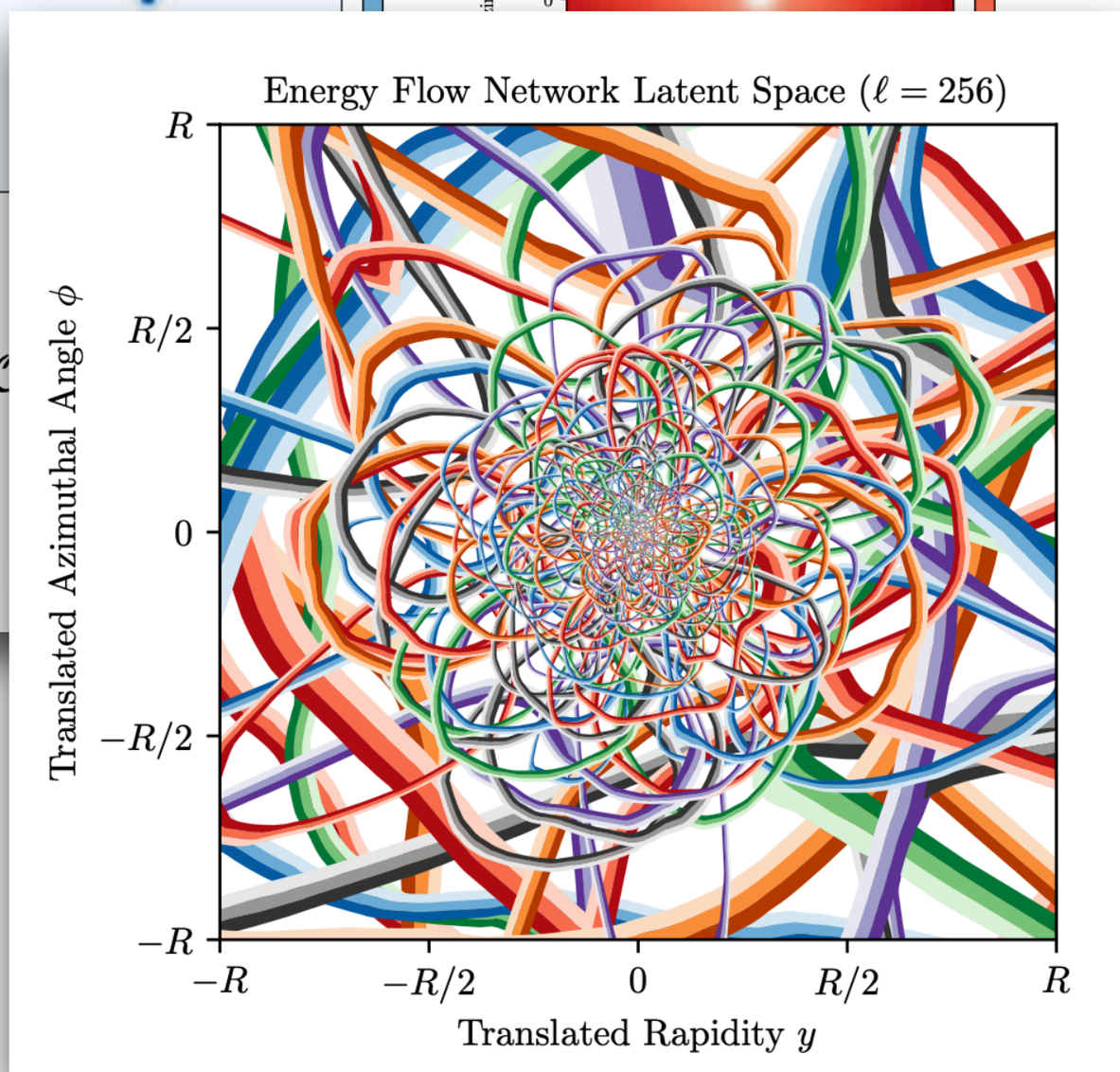
Take radial slices to obtain envelope

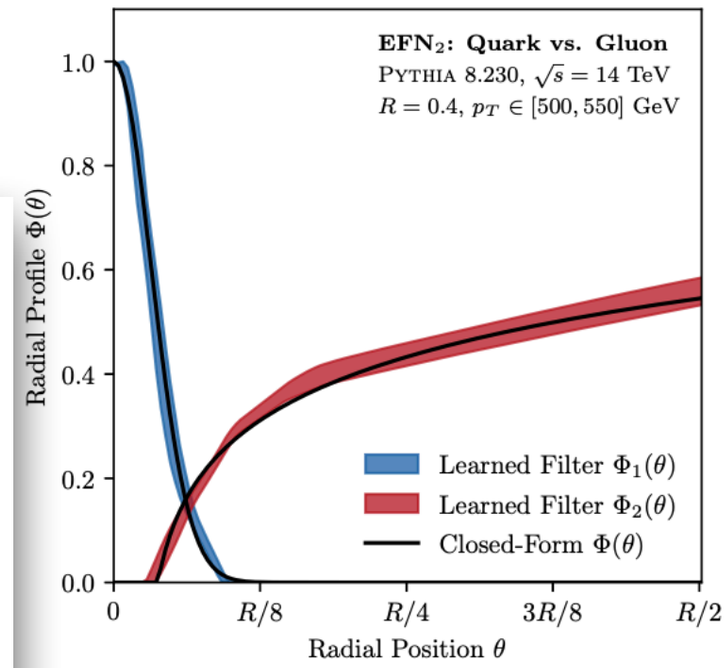
EFN (ℓ = 2) has approximately radially symmetric filters



Extracting New Analytic Observables

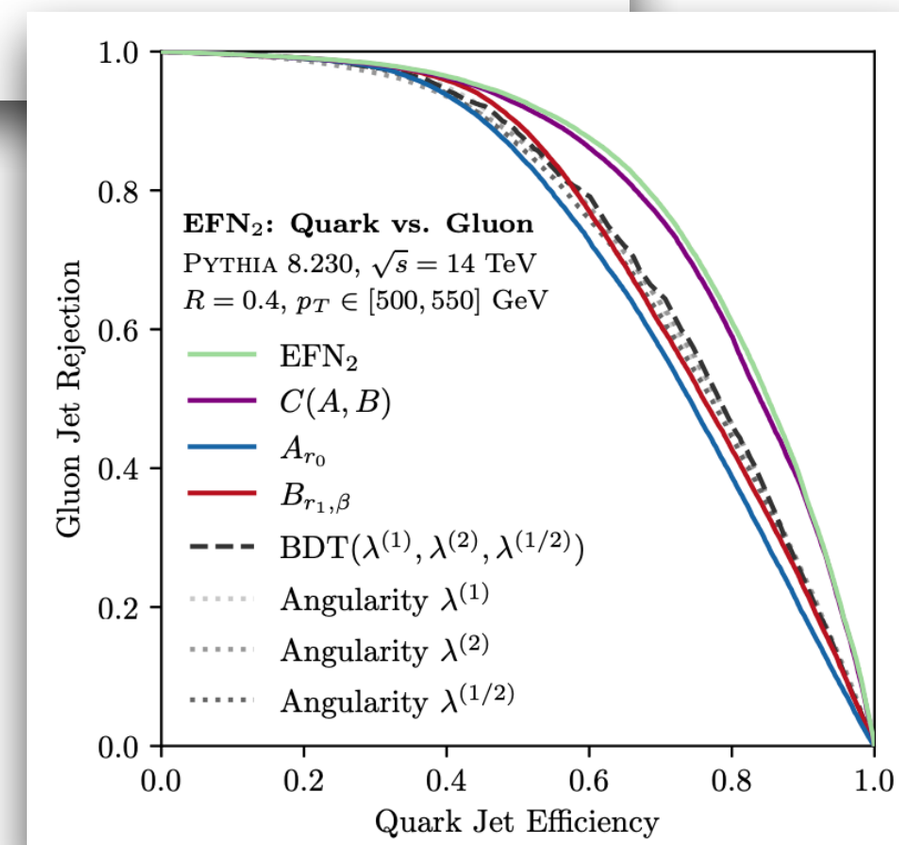






Take radial slices to obtain envelope

EFN



arXiv:1810.05165
for more details

Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT
- Jet Images

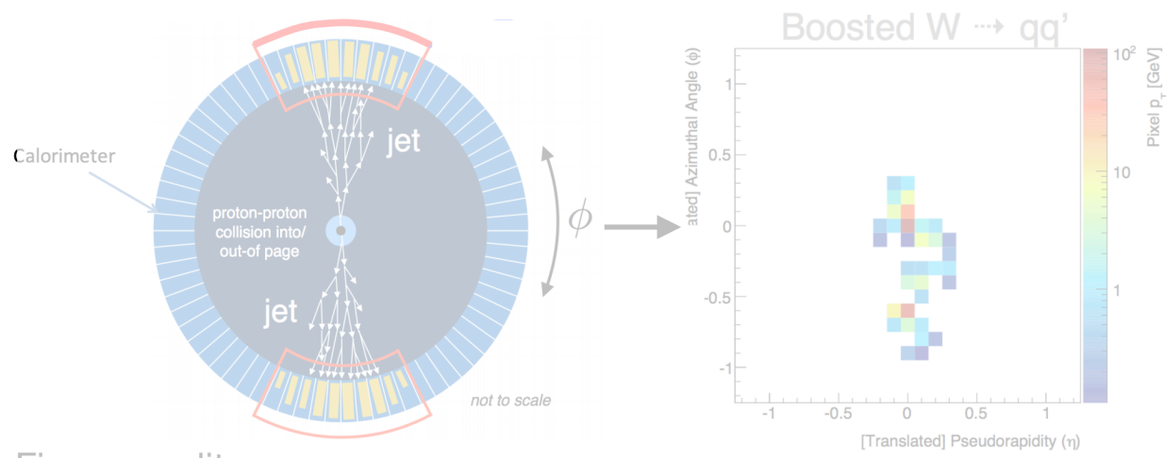
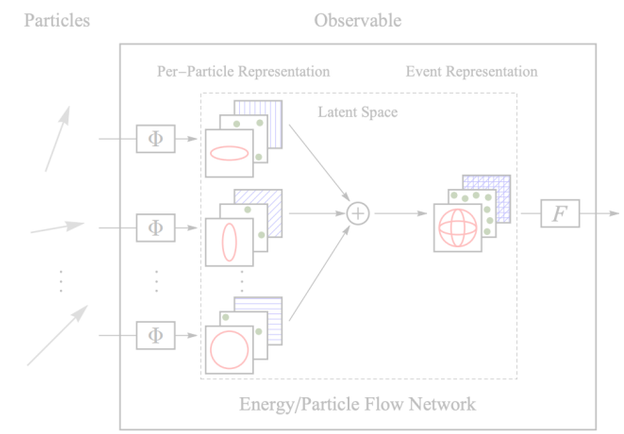


Figure credit: B. Nachman

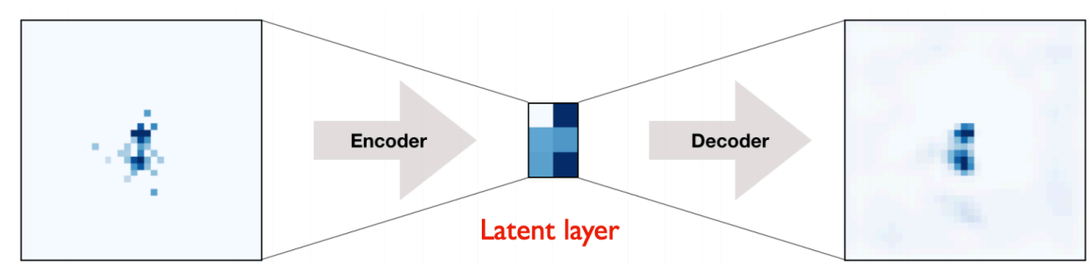
- Sequences
  - pT ordering
  - Clustering history

- Energy/Particle Flow Network

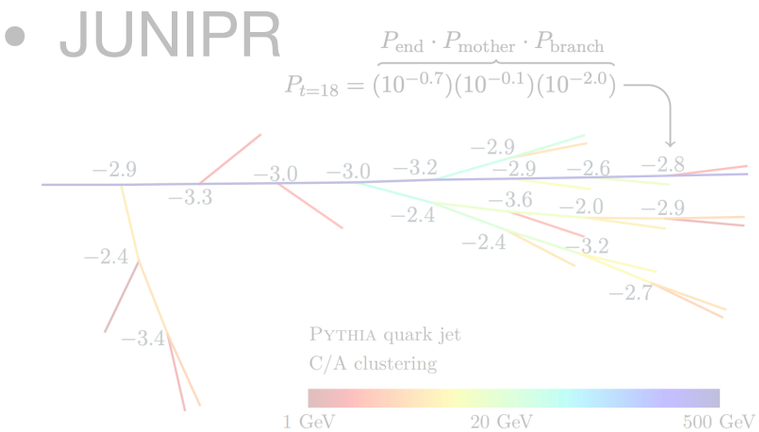


Komiske, Metodiev & Thaler (2018)

- Autoencoders
  - Farina, Nakai, Shih (2018)
  - Heimel, Kasieczka, Plehn, Thompson (2018)



- JUNIPR



AA, Feige, Frye & Schwartz (2018)

ML4Jets ’18 @FNAL:

Searching for new physics with autoencoders

ML4Jets
November 16, 2018

Marco Farina
Stony Brook University

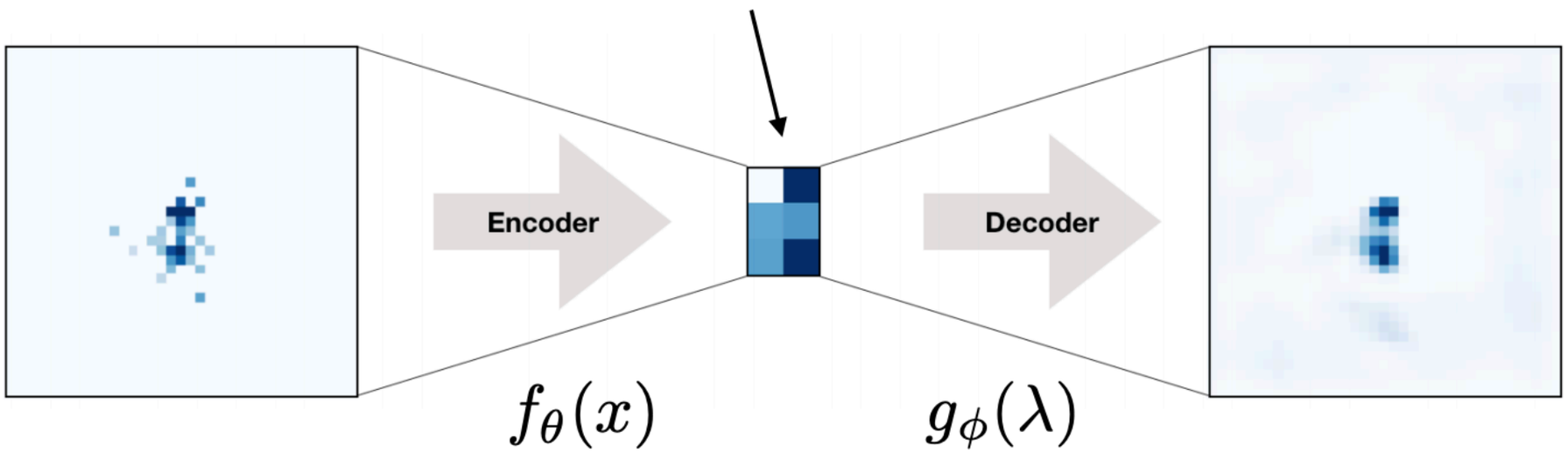
Based on Farina, Nakai, Shih ‘18

arXiv:1808.08992

See also: 1807.10261

Autoencoders

λ bottleneck latent dim

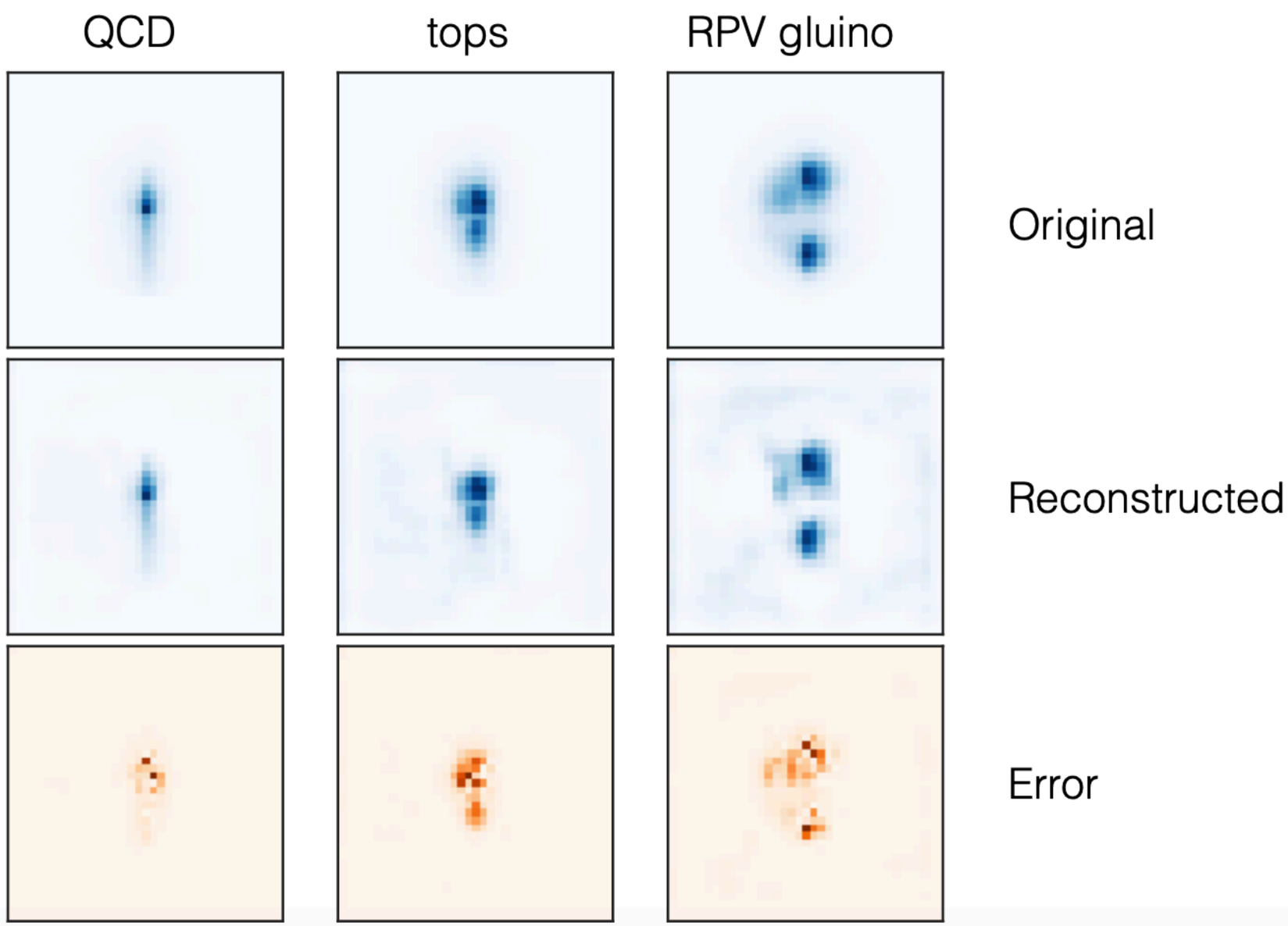


- PCA
- DNN
- CNN

Slide by Marco Farina

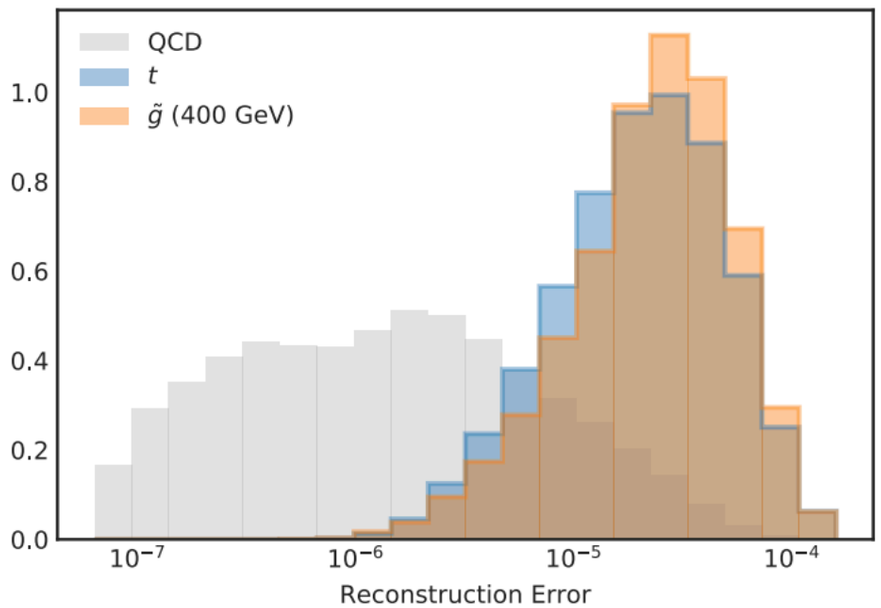
Anomalous jets detection

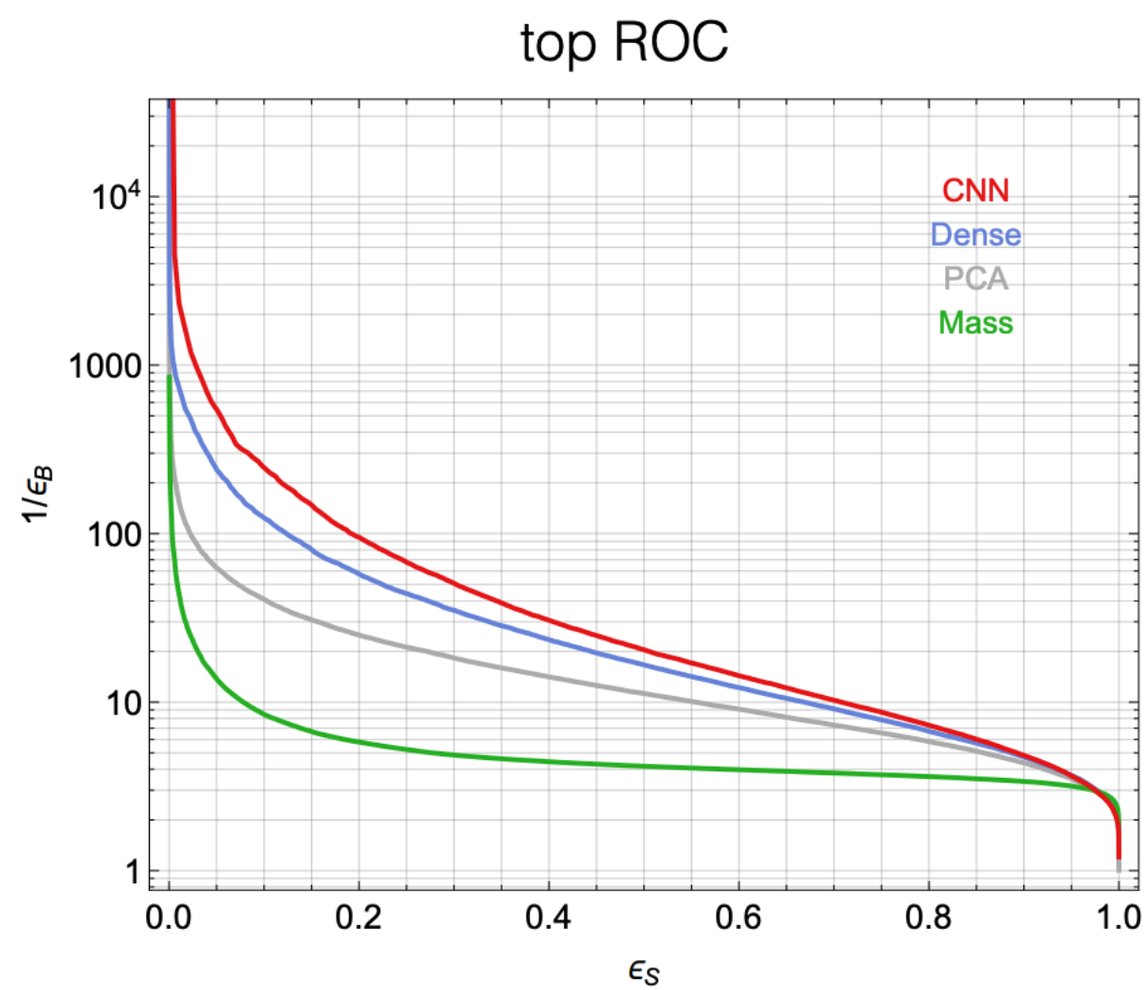
After training on QCD jets…



Slide by Marco Farina

Anomalous jets detection





Tag anomaly using cut on reconstruction error

Slide by Marco Farina

Anomaly Detection in another way

CWoLa Hunting:

Extending the Bump Hunt with Machine Learning

Based on:

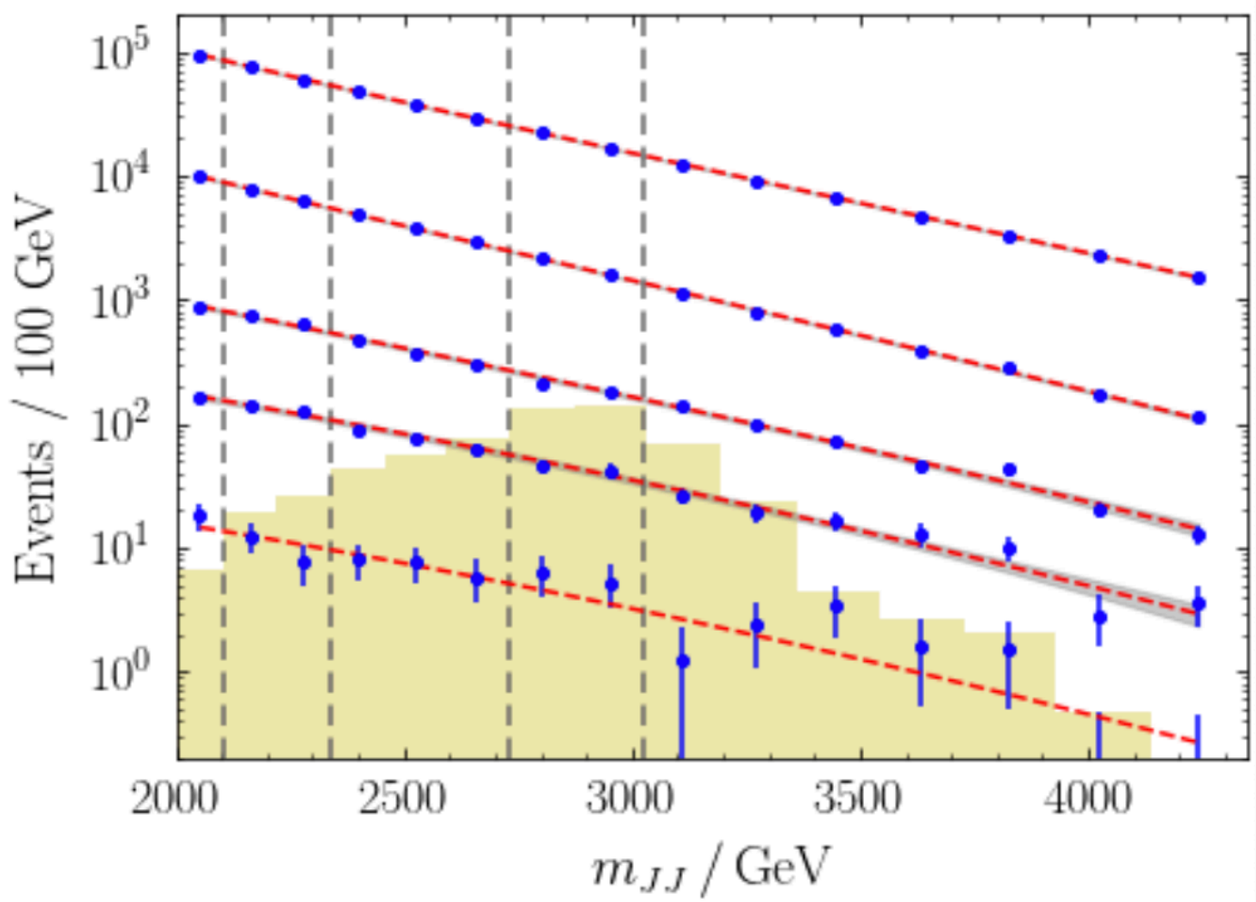
[1805.02664] Jack Collins, Kiel Howe, Ben Nachman

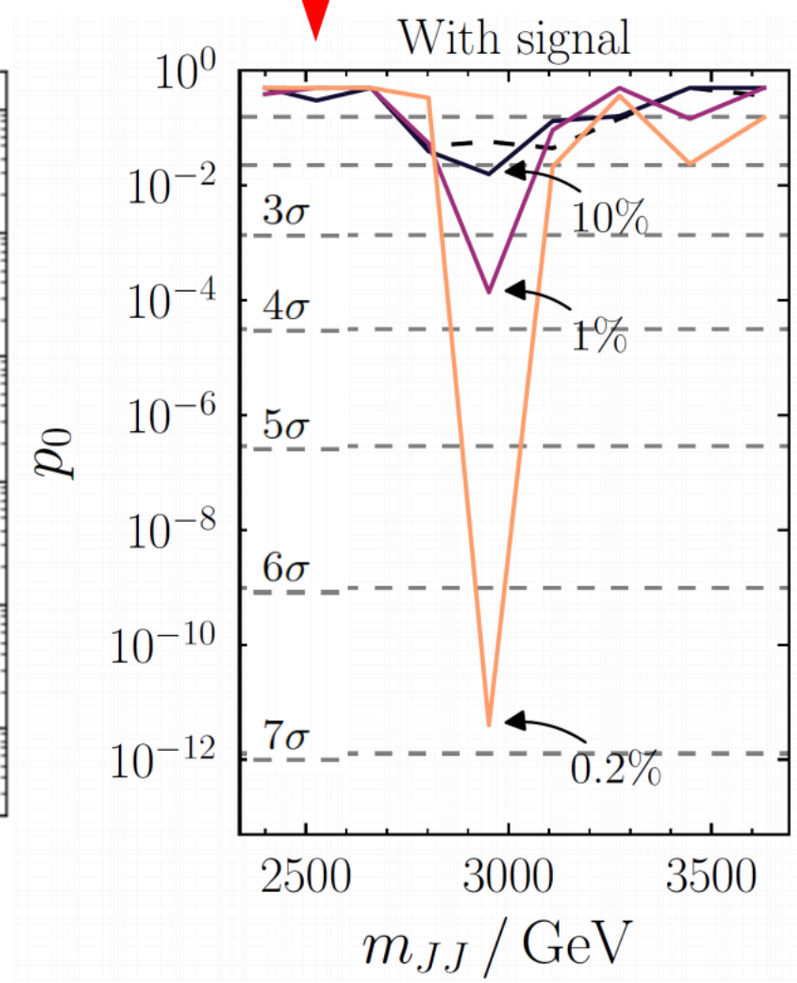
UNIVERSITY OF MARYLAND

JOHNS HOPKINS UNIVERSITY

Slide by Jack Collins

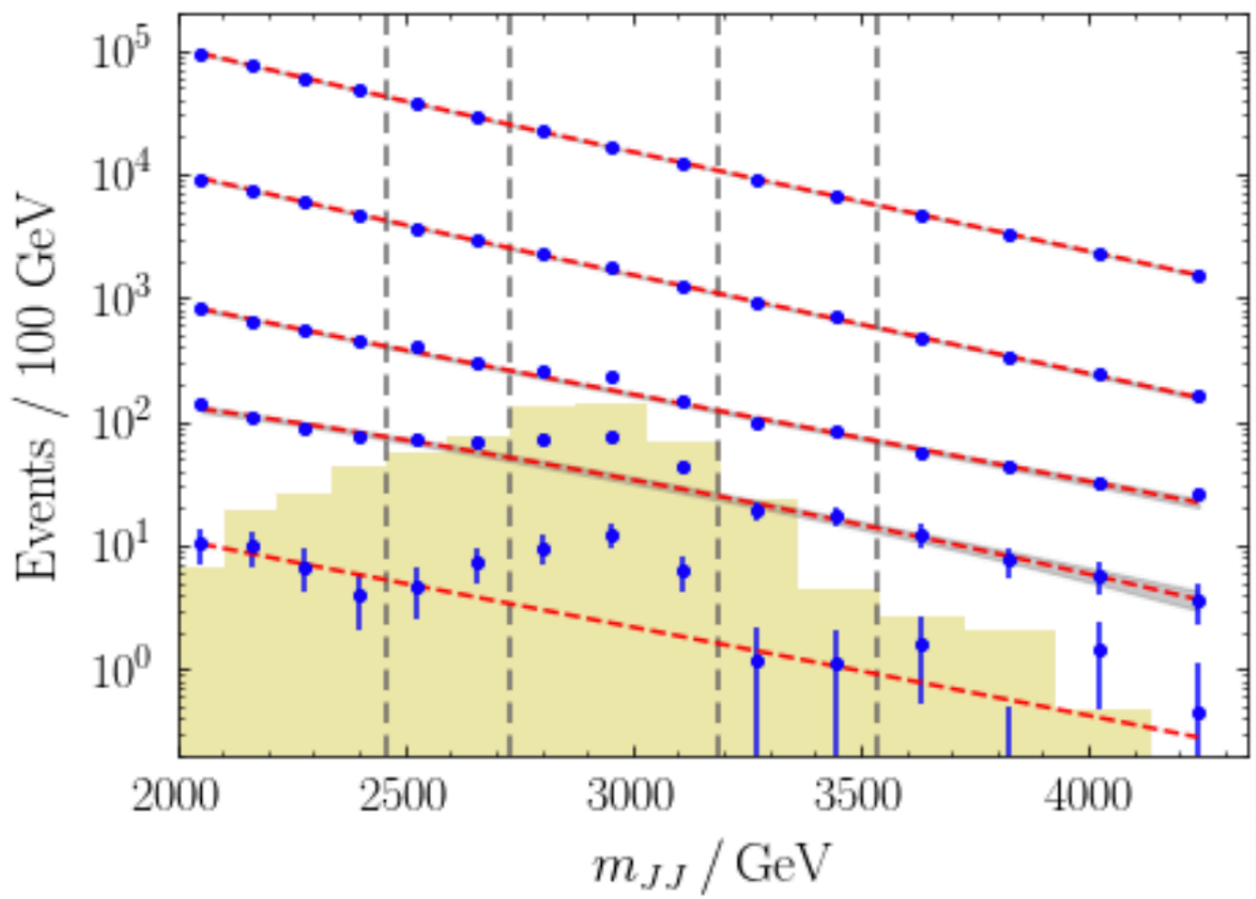
Mass Scan

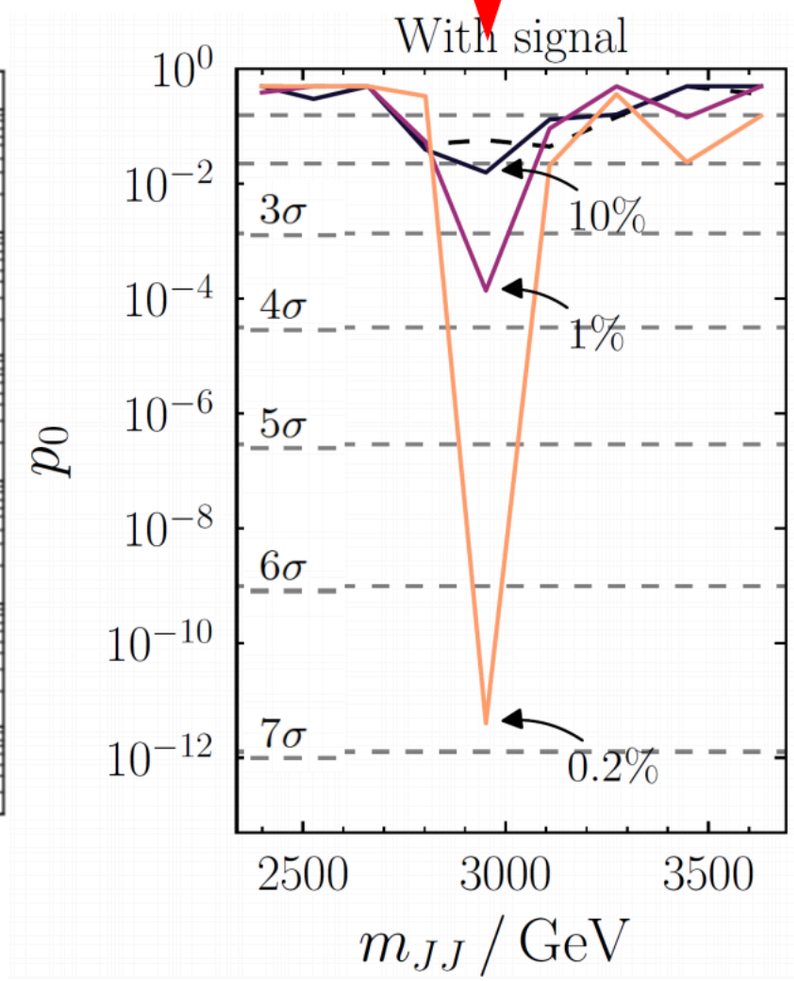




Slide by Jack Collins

Mass Scan





Slide by Jack Collins

Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT
- Jet Images

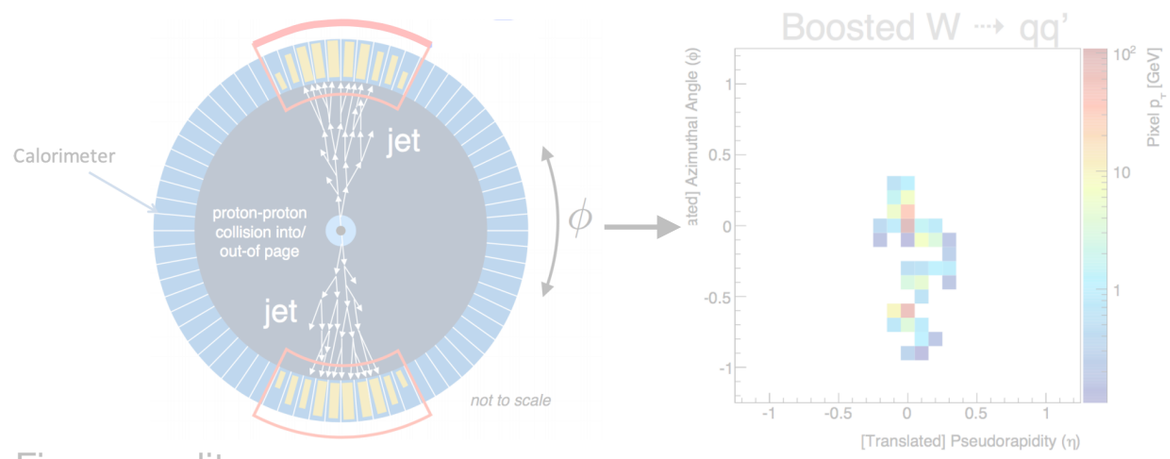
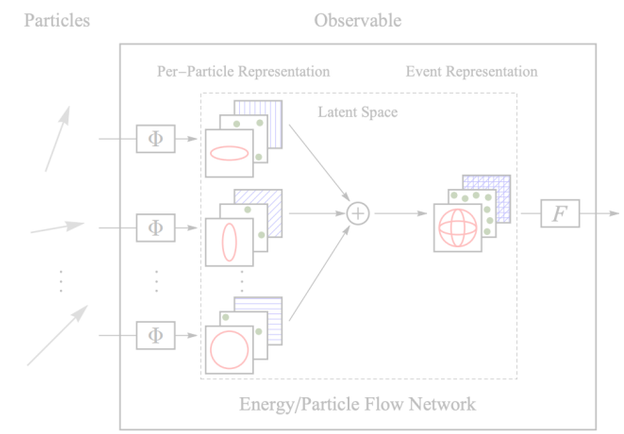


Figure credit: B. Nachman

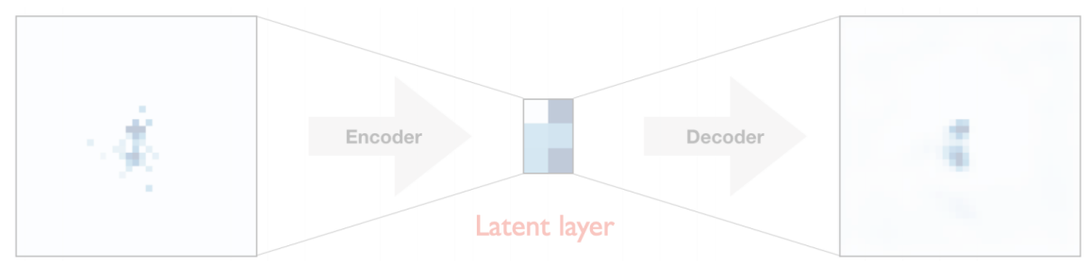
- Sequences
  - pT ordering
  - Clustering history

- Energy/Particle Flow Network

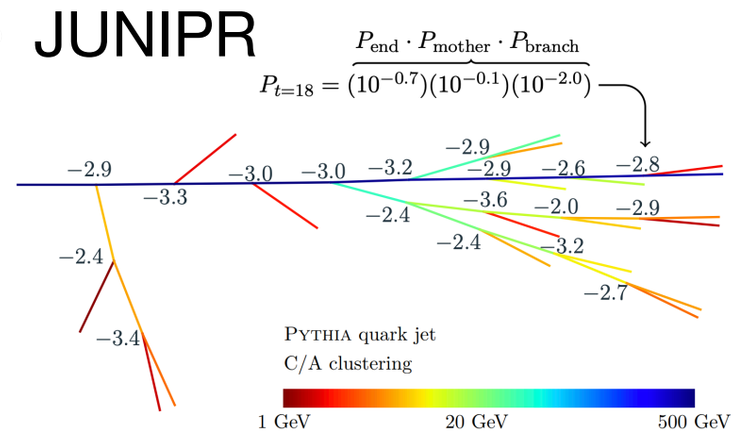


Komiske, Metodiev & Thaler (2018)

- Autoencoders
  - Farina, Nakai, Shih (2018)
  - Heimel, Kasieczka, Plehn, Thompson (2018)

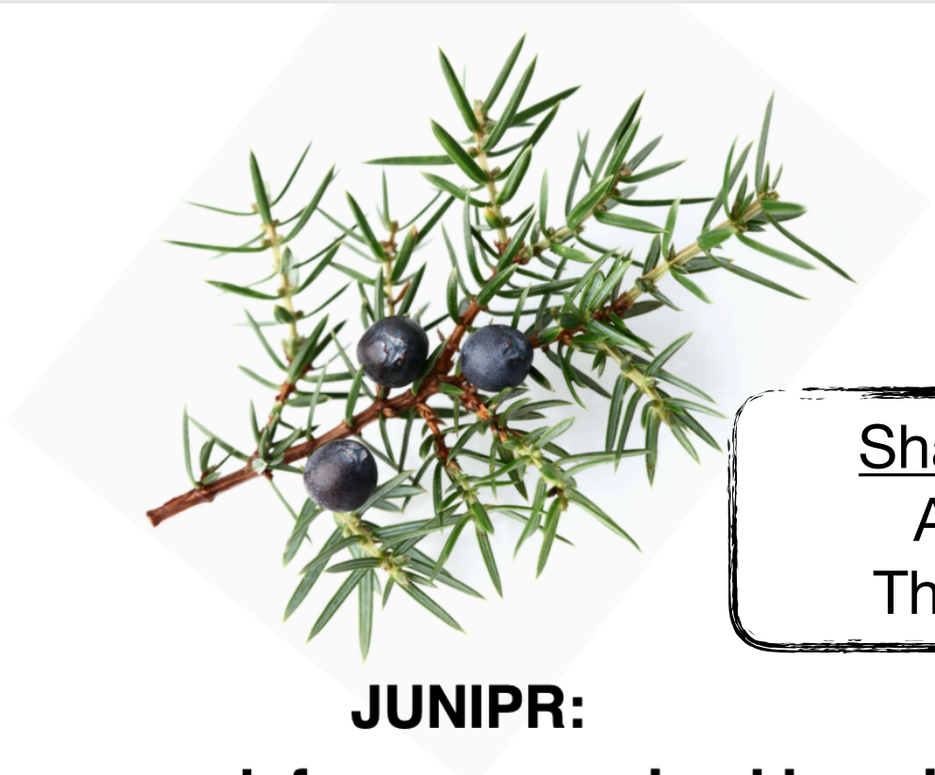


- JUNIPR



AA, Feige, Frye & Schwartz (2018)

ML4Jets ’18 @FNAL:



Shameless self-promotion:
ATLAS Theory seminar
Thursday Dec. 13 @LBNL

JUNIPR:
a framework for unsupervised learning in jet physics

Anders Andreassen
aja@lbl.gov

in collaboration with
Feige, Frye and Schwartz
arXiv: 1804.09720

BERKELEY LAB

Berkeley
UNIVERSITY OF CALIFORNIA

VE RI TAS
HARVARD

JUNIPR Motivation

How can we understand jets better using machine learning?

Idea: (1) Let neural network learn about jets
(2) Look inside to see what it's doing ← challenging!

Our strategy:

- use network architecture inspired by QCD shower
- but general enough to fit any non-QCD structure

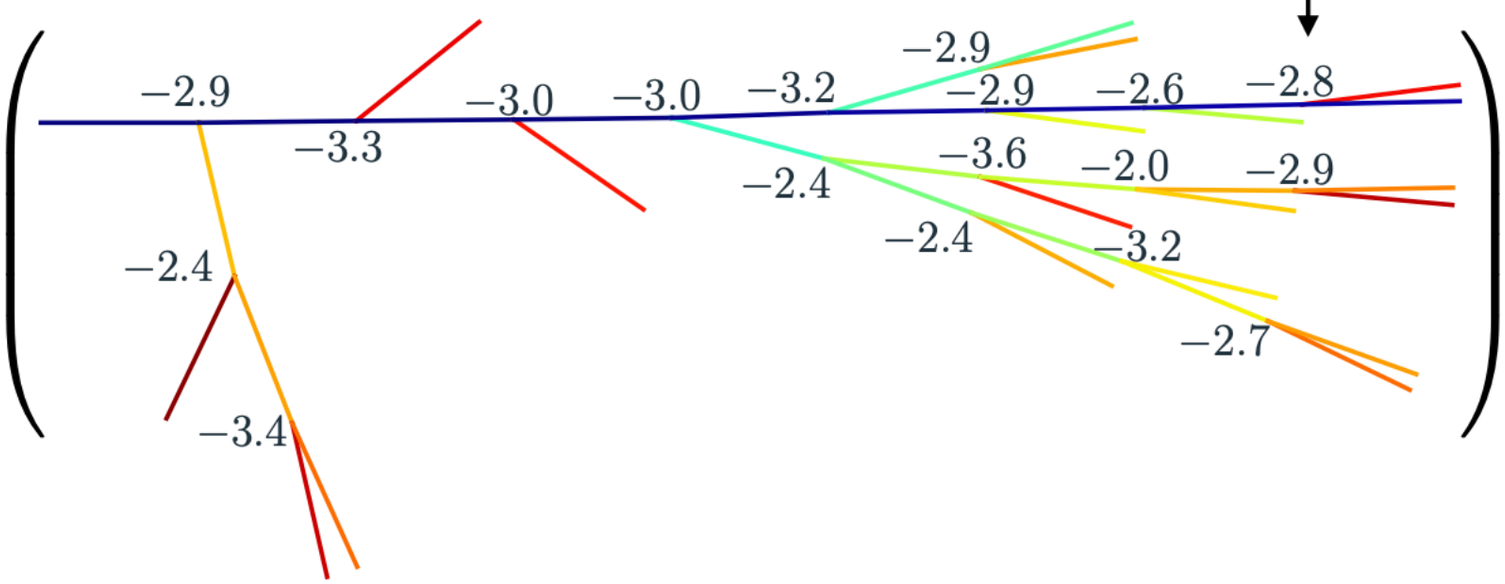
RNN built around a clustering tree ⟹ interpret output from intermediate layers!

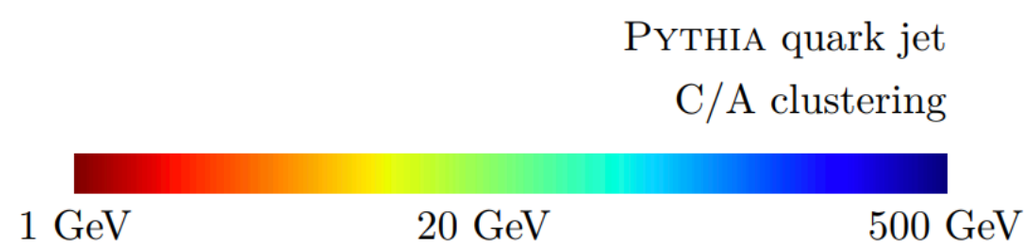
JUNIPR models the evolution of the probability of each splitting of a clustering tree

example Pythia quark jet:

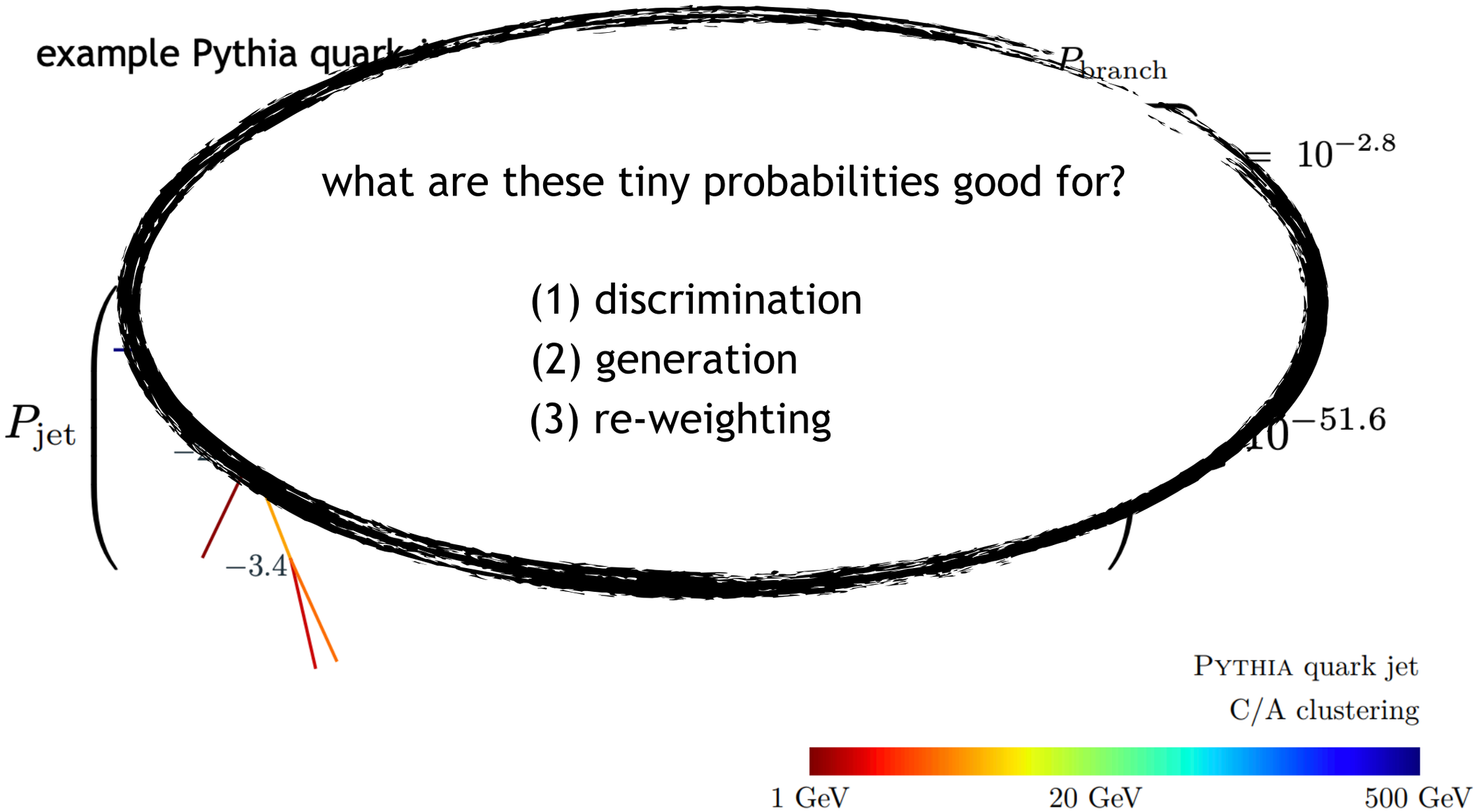
$$P_{t=18} = \overbrace{\left(10^{-0.7}\right)\left(10^{-0.1}\right)\left(10^{-2.0}\right)}^{P_{\text{end}} \cdot P_{\text{mother}} \cdot P_{\text{branch}}} = 10^{-2.8}$$

$$P_{\text{jet}}\left(\right) = 10^{-51.6}$$





JUNIPR models the evolution of the probability of each splitting of a clustering tree



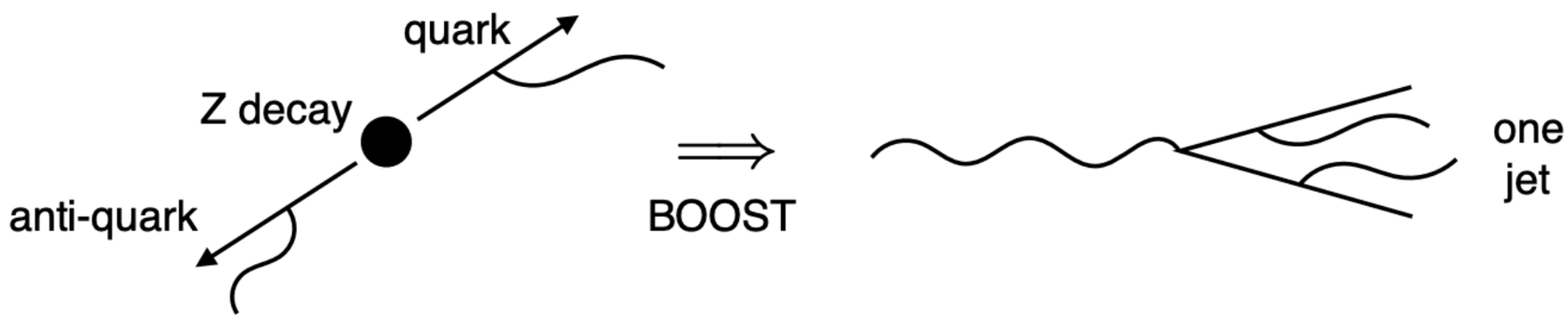
what are these tiny probabilities good for?

(1) discrimination

(2) generation

(3) re-weighting

(1) Discrimination



- boosted Z / quark jet discrimination for proof-of-concept

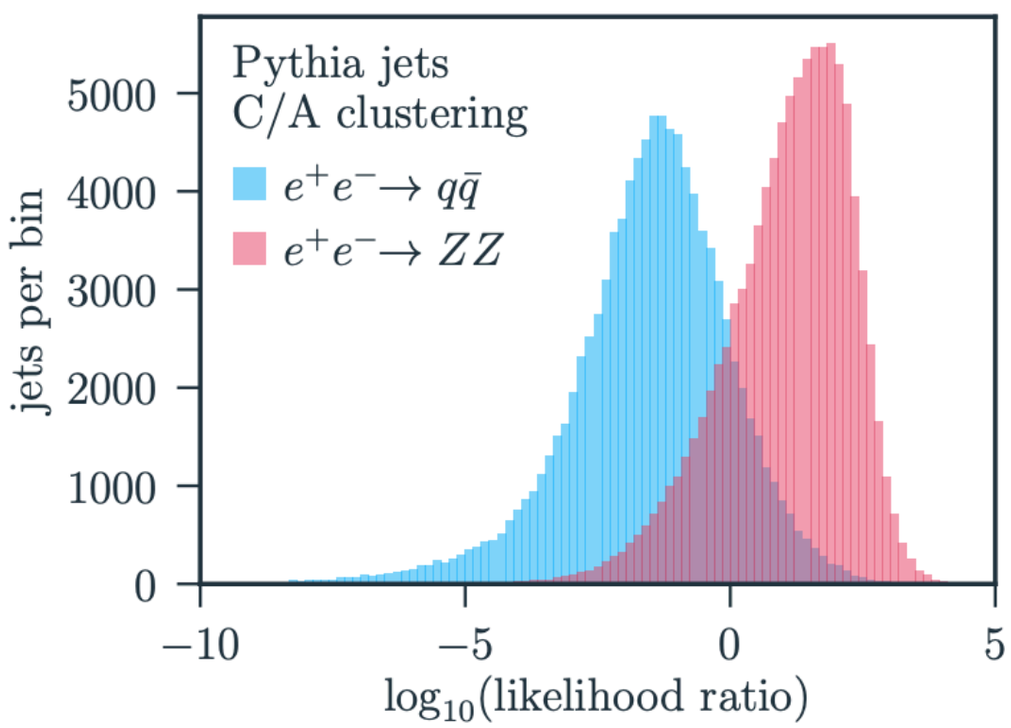
- trained two models: $\left.\begin{array}{c} P_Z(\text{jet}) \\ P_q(\text{jet}) \end{array}\right\}$ mass cut on jets in training data 90.7 — 91.7 GeV

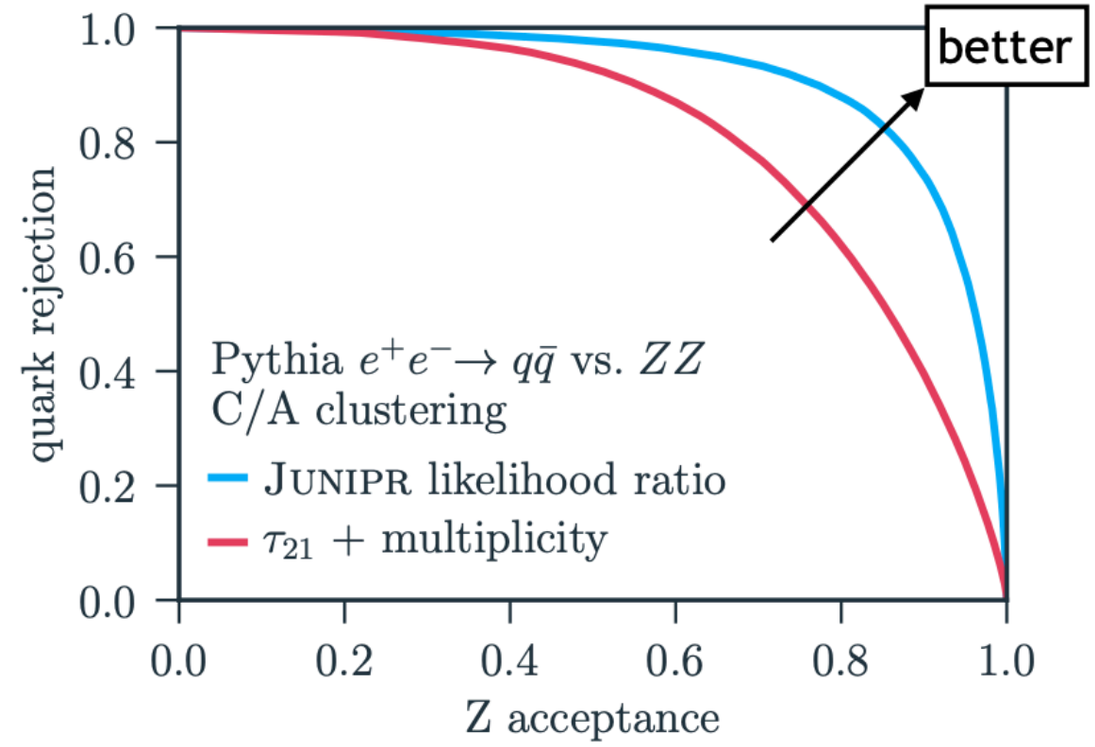
- theoretically most powerful discriminant is likelihood ratio:

$$\frac{P_Z(\text{jet})}{P_q(\text{jet})} > \text{threshold} \quad \Longrightarrow \quad \text{tag jet as boosted } Z$$

(1) Discrimination

$$\frac{P_Z(\text{jet})}{P_q(\text{jet})} > \text{threshold} \quad \Longrightarrow \quad \text{tag jet as boosted } Z$$





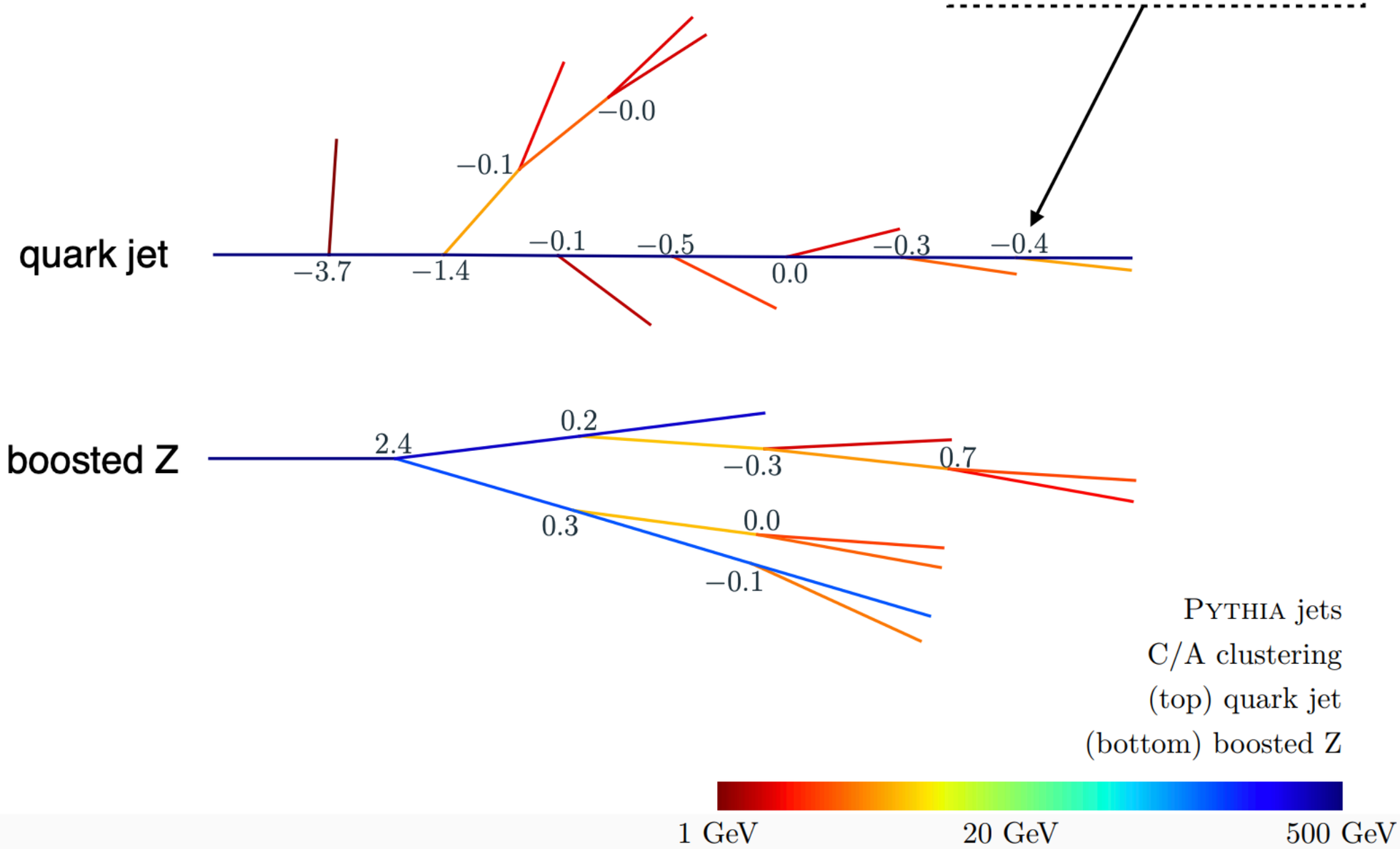
(1) Discrimination

- visualize high-performance discrimination with clustering trees!

nodes labeled by

$$\log_{10} \frac{P_Z(t)}{P_q(t)}$$

negative $\Longrightarrow$ quark like
positive $\Longrightarrow$ Z like



Representations of jets for Machine Learning

- Physics Motivated Inputs
  - Give physics motivated observables to a NN or BDT

- Jet Images

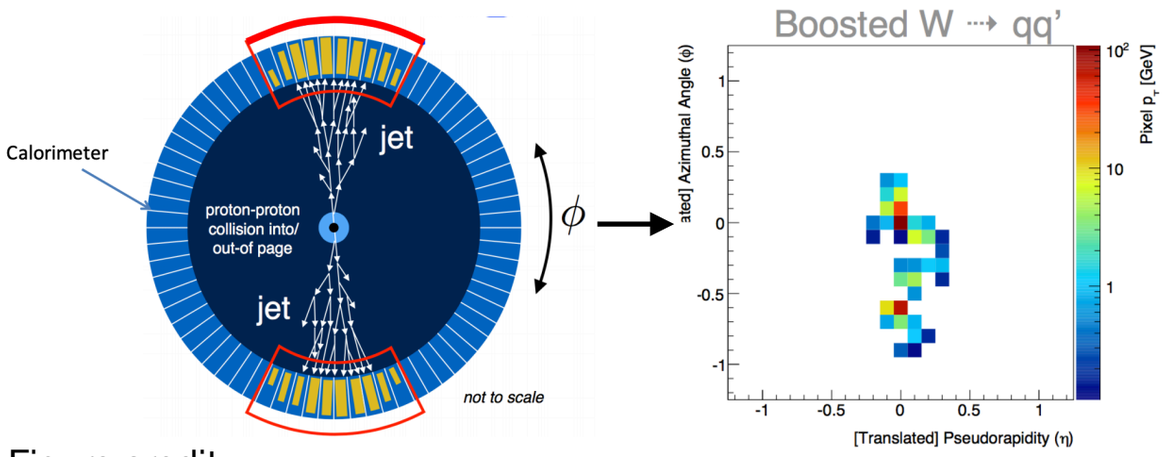
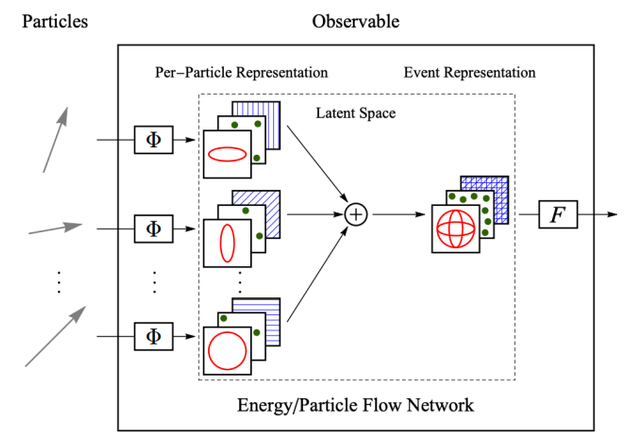


Figure credit: B. Nachman

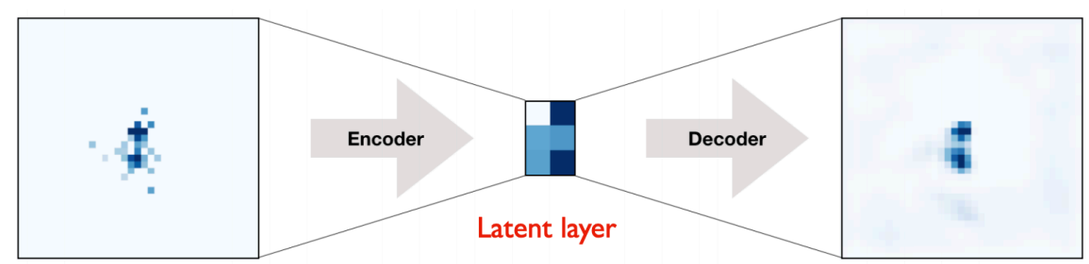
- Sequences
  - pT ordering
  - Clustering history

- Energy/Particle Flow Network

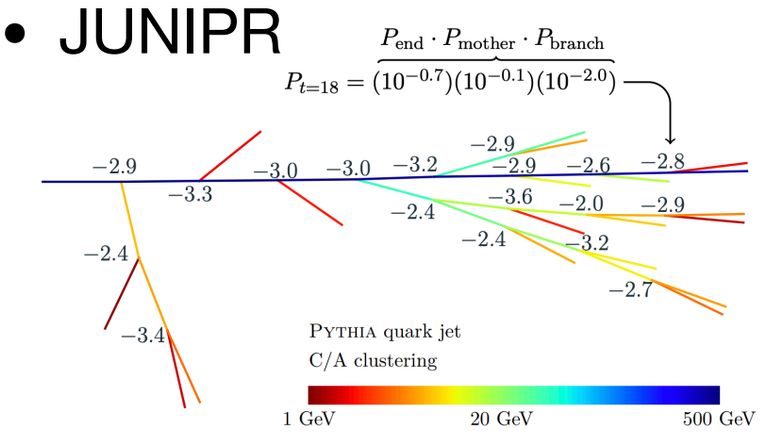


Komiske, Metodiev & Thaler (2018)

- Autoencoders
  - Farina, Nakai, Shih (2018)
  - Heimel, Kasieczka, Plehn, Thompson (2018)



- JUNIPR



AA, Feige, Frye & Schwartz (2018)

Thank You!