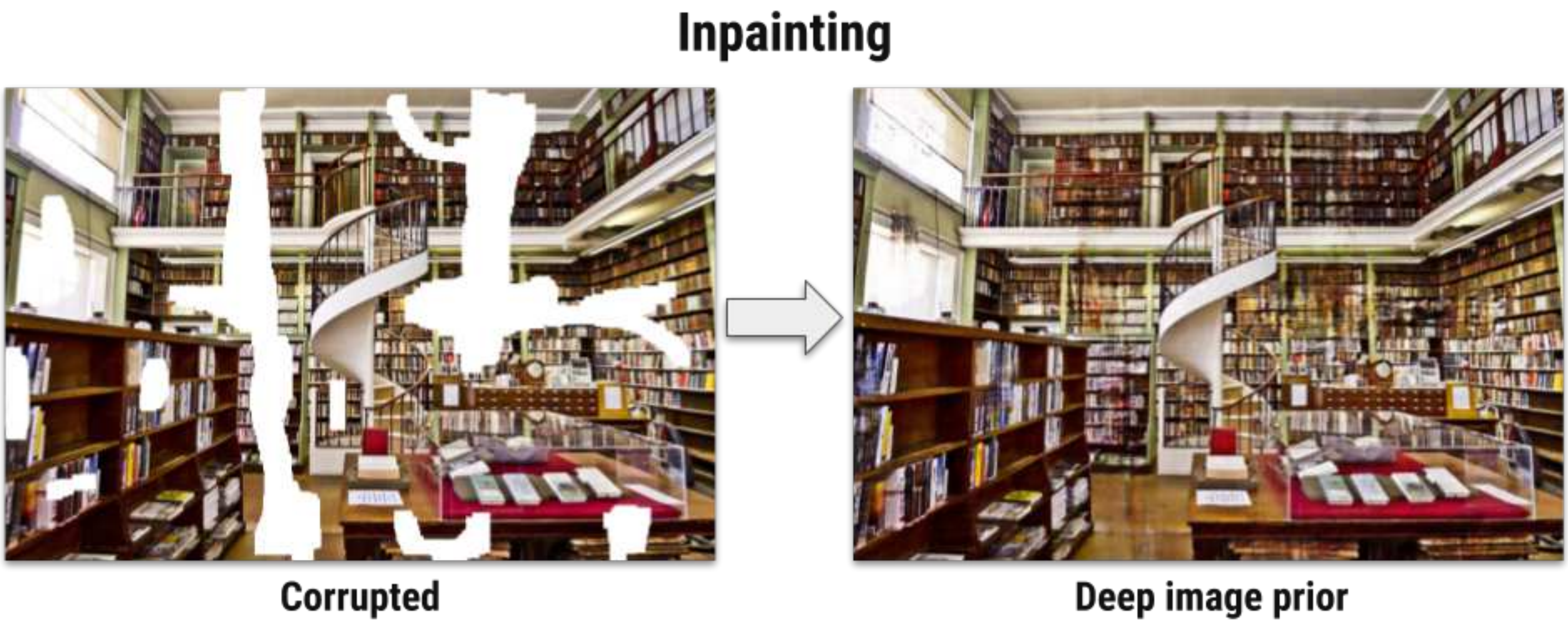
Deep Learning (and Deep Thinking) in Collider Physics

Jesse Thaler

MIT

Pheno 2019, University of Pittsburgh — May 8, 2019

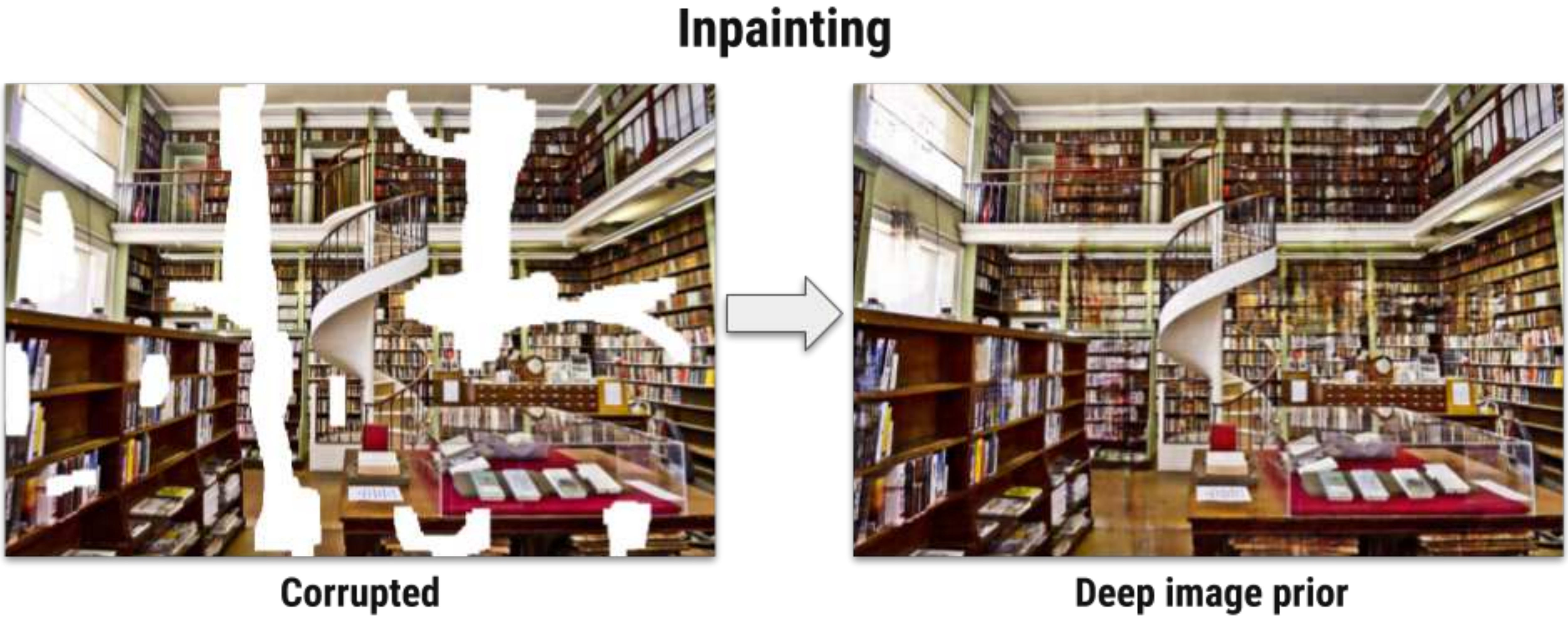
Deep Learning



increased computational power and large data sets

[Ulyanov, Vedaldi, Lempitsky, 1711.10925]

Deep Learning (or Deep Thinking?)

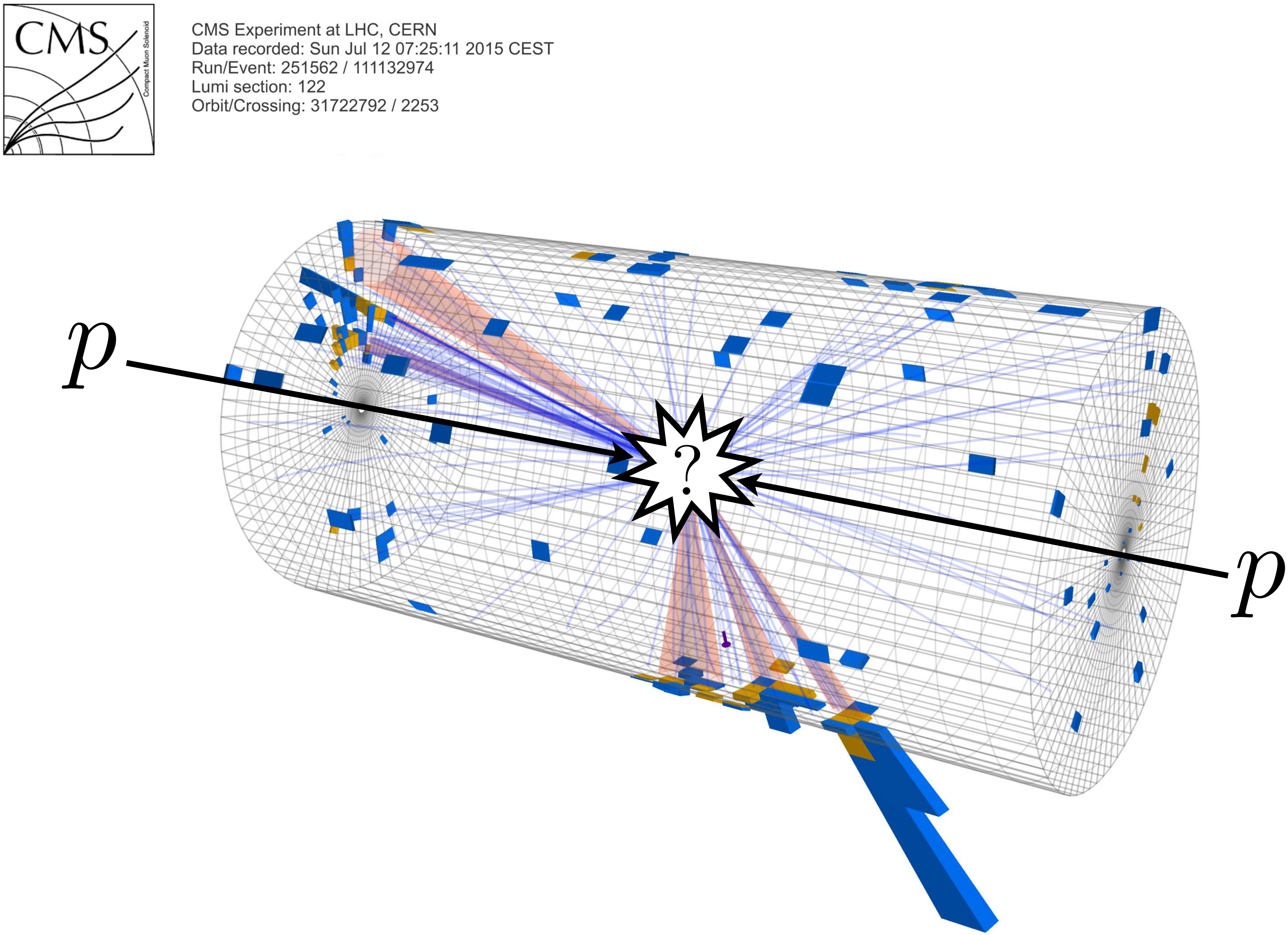


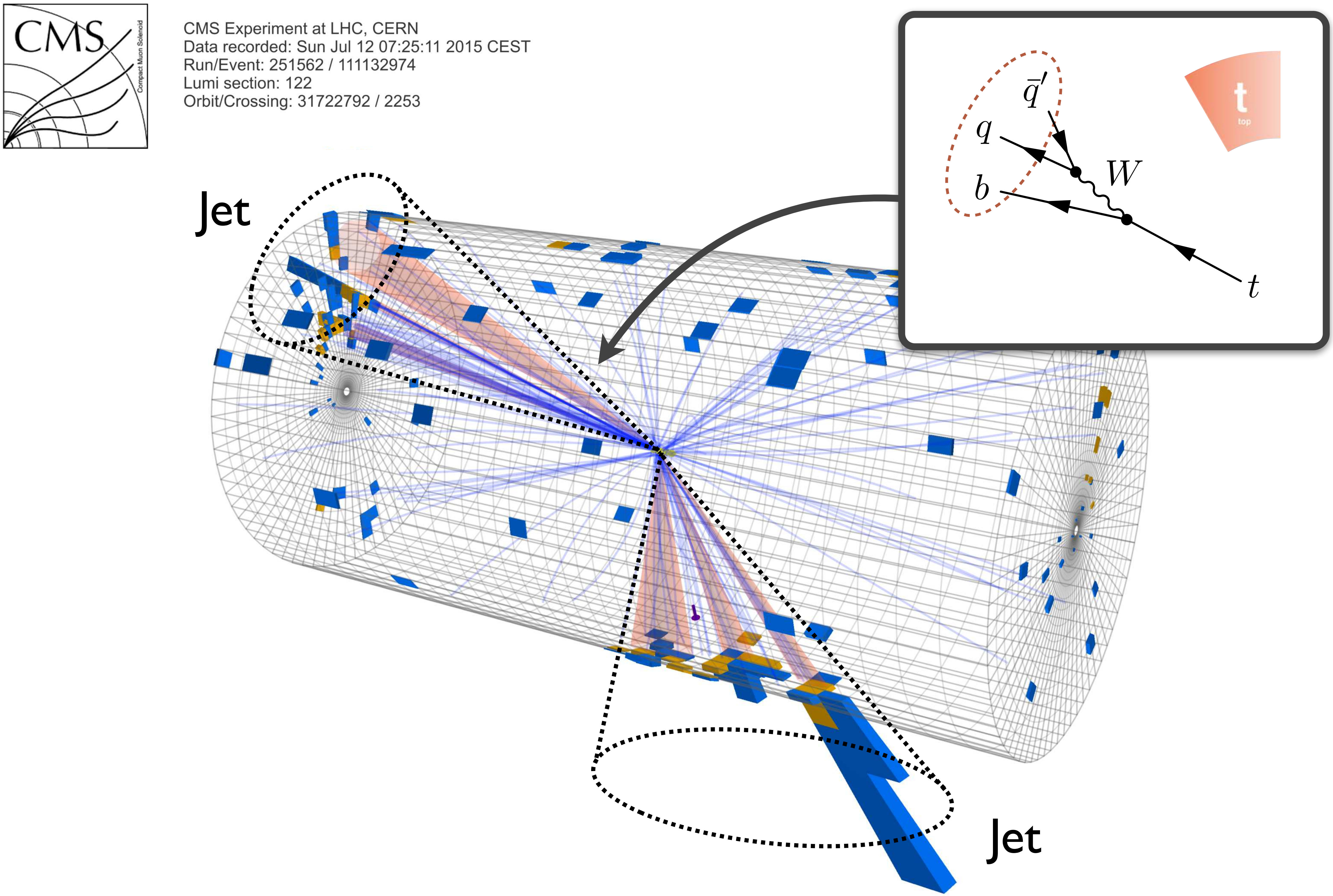
Using randomly initialized neural network (!)

Progress made by understanding the structure of problems

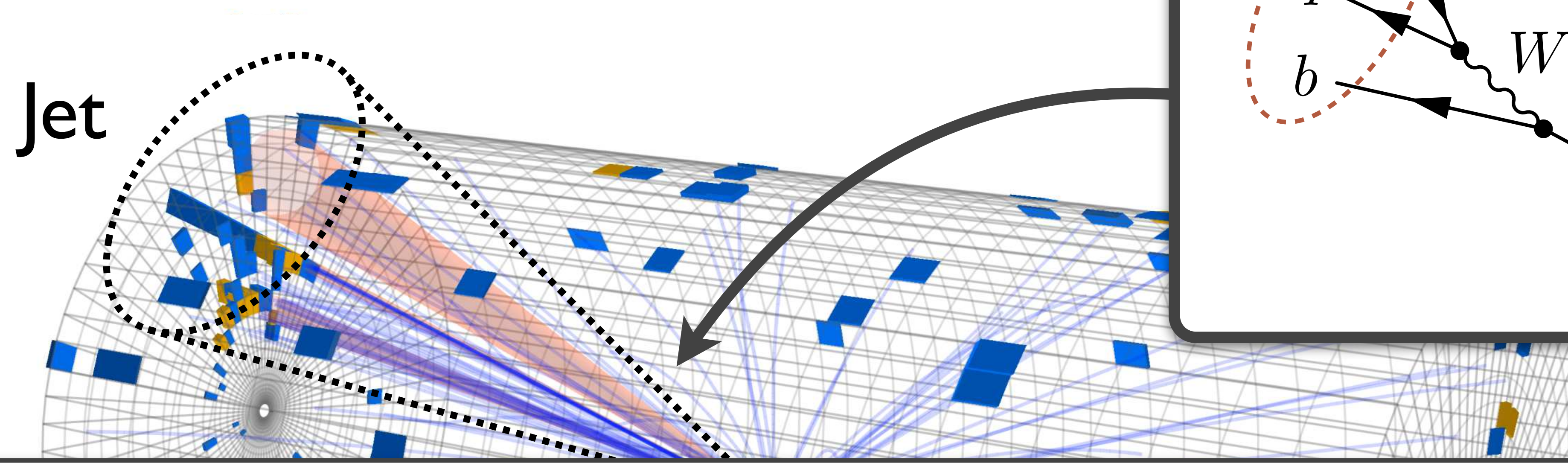
(not just increased computational power and large data sets)

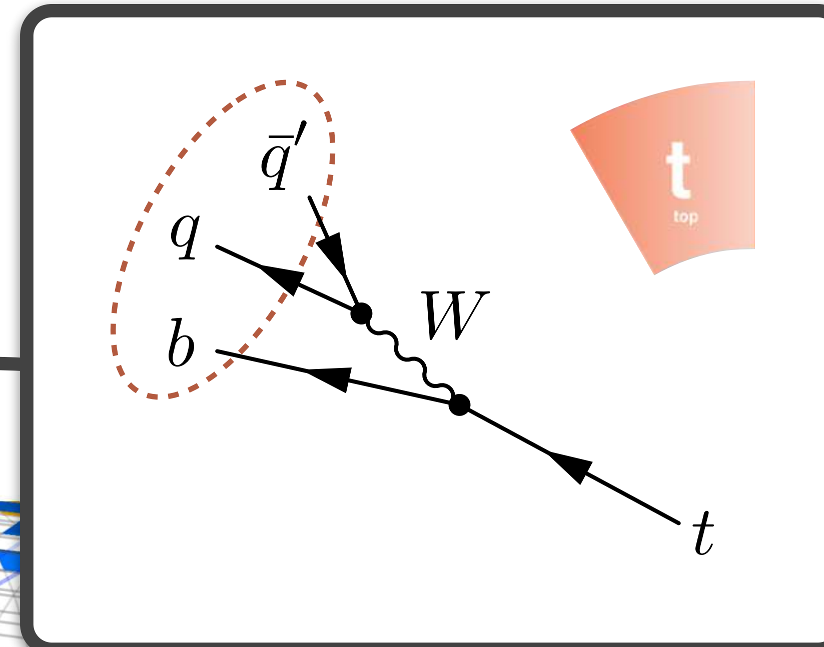
[Ulyanov, Vedaldi, Lempitsky, 1711.10925]


CMS
Compact Muon Solenoid
CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253
p
?
p


CMS
Compact Muon Solenoid
CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253
Jet
$\bar{q}'$
q
W
b
t
t
top
Jet

CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253





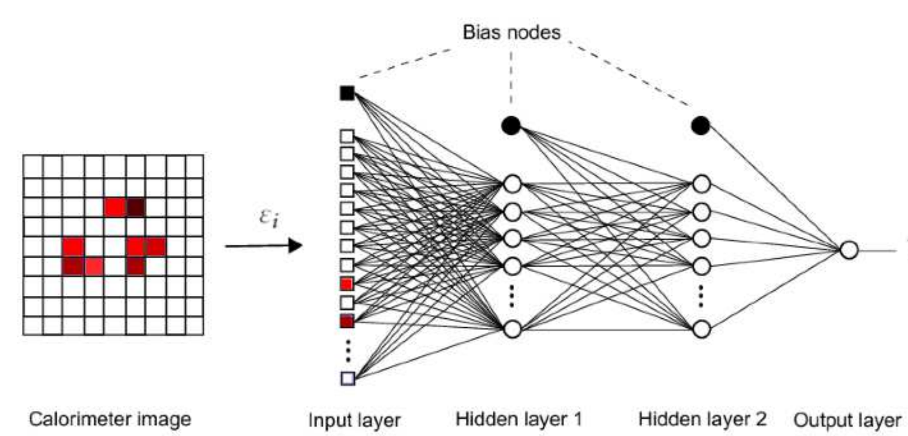
“Deep Thinking”?

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

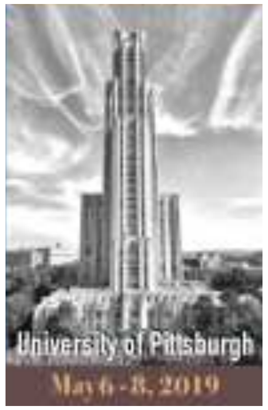
$$\tau_N(\mathcal{J}) = \min_{|\mathcal{J}'|=N} \mathrm{EMD}(\mathcal{J}, \mathcal{J}')$$

“Deep Learning”?

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



Deep Learning at Pheno 2019



Tuesday @ 15:15 (Higgs II): Jeong Han Kim
"Portraying Double Higgs Production with Deep Neural Networks"

Tuesday @ 14:15 (BSM III): Prasanth Shyamsundar
"Machine learning in event selection: Improving the supervisory signal and output usage"

Tuesday @ 15:00 (BSM III): Felix Kling
"MadMiner: An inference toolkit for particle physics"

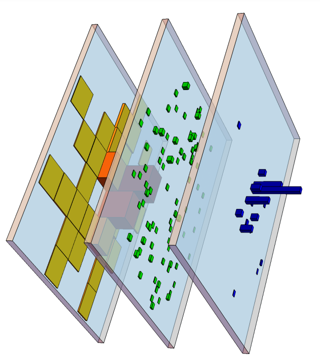
Please email me if I missed your talk!

Plus, public collider data:

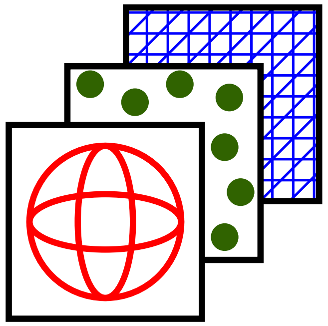
Tuesday @ 14:45 (BSM III): Cari Cesarotti
"Searching for New Dimuon Resonances at the LHC with CMS Open Data"

[Cesarotti, Soreq, Strassler, JDT, Xue, 1902.04222]

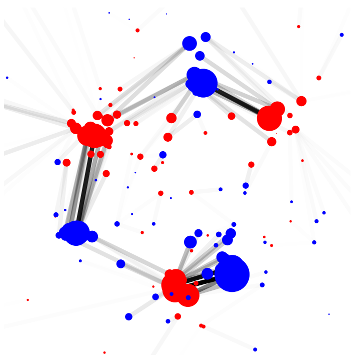
Outline



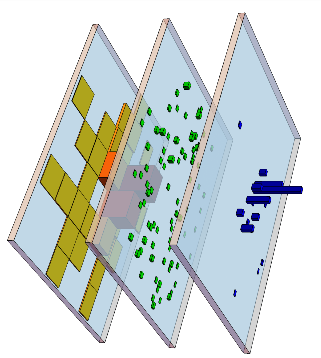
The Rise of Deep Learning



Looking Inside the Black Box



(Exploring the Space of Jets)



The Rise of Deep Learning

Looking Inside the Black Box

(Exploring the Space of Jets)

Cartoon of Machine Learning

Properly specified problem $\Rightarrow$

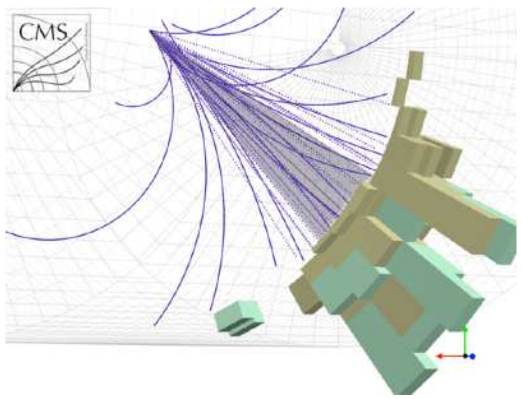
Many (labeled) examples $\{\mathcal{J}\}$ $\Rightarrow$

The Machine

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

E.g.: Problem = Minimize loss function
Solution = Multi-layer neural network
Strategy = Stochastic gradient descent

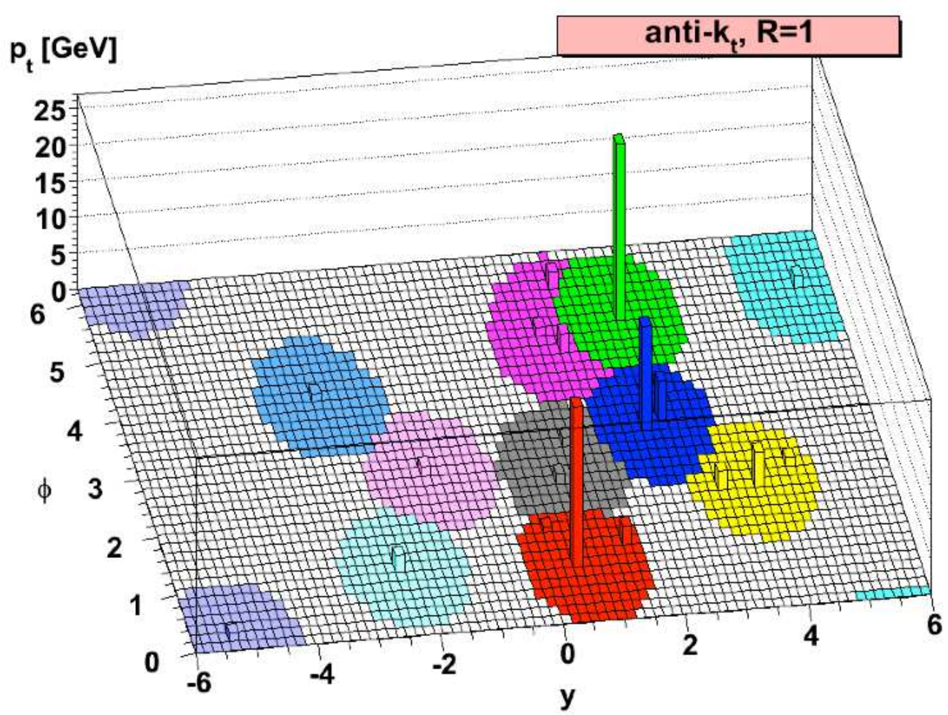
For most of this talk: $\mathcal{J}$ = "jet"



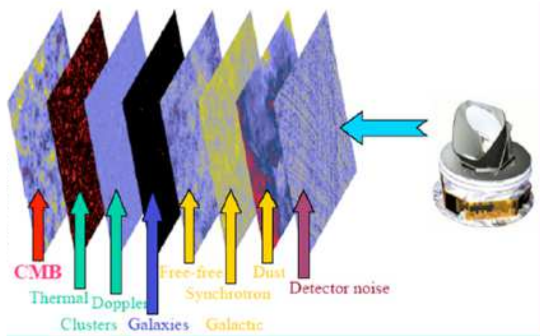
Examples of Unsupervised Learning

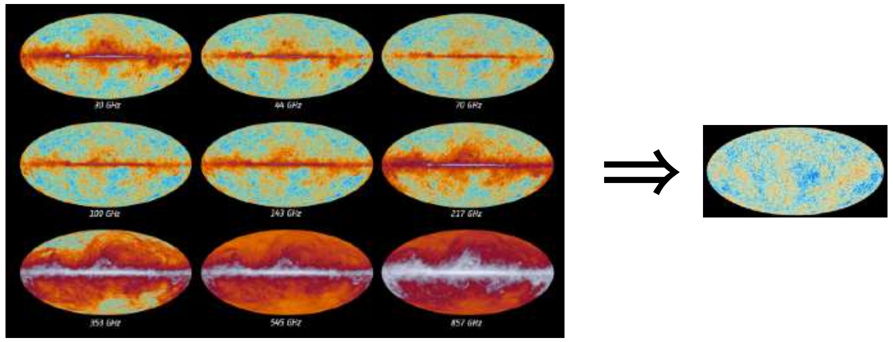
(see backup for
probability modeling
and anomaly detection)

Clustering



Topic Modeling





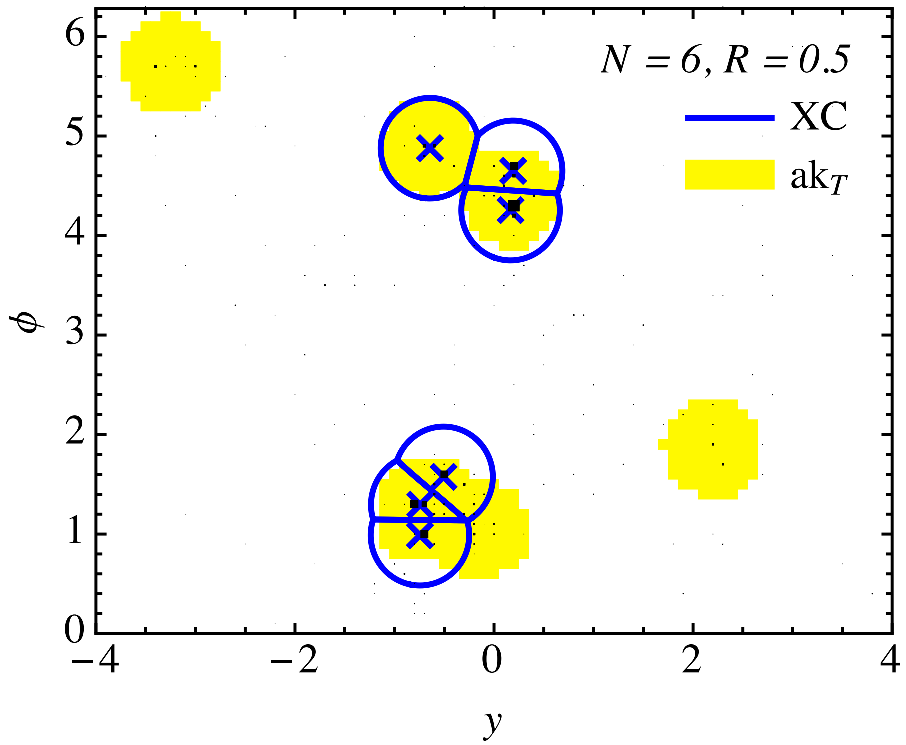
(Approximate) solutions to properly specified problems

[figures from Cacciari, Salam, Soyez, 0802.1189; Planck Outreach]

Examples of Unsupervised Learning

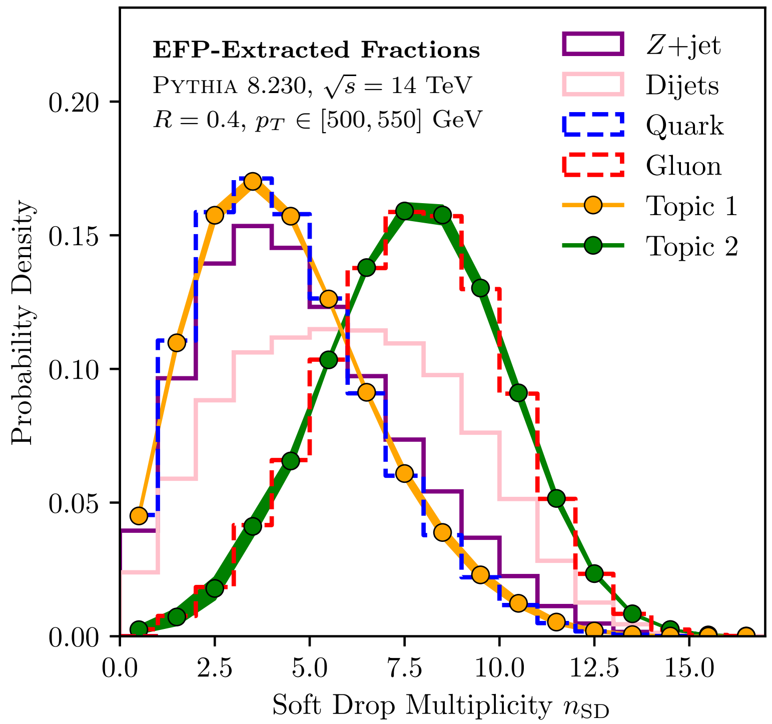
(see backup for probability modeling and anomaly detection)

XCone Jet Finding



"Find N axes that minimize N-jettiness"

Jet Topics



"Find two mutually irreducible distributions"

[Stewart, Tackmann, JDT, Vermilion, Wilkason, 1508.01516; based on Stewart, Tackmann, Waalewijn, 1004.2489]
[Metodiev, JDT, 1802.00008; Komiske, Metodiev, JDT, 1809.01140; see also Dillon, Faroughy, Kamenik, 1904.04200]

(Approximate) solutions to properly specified problems

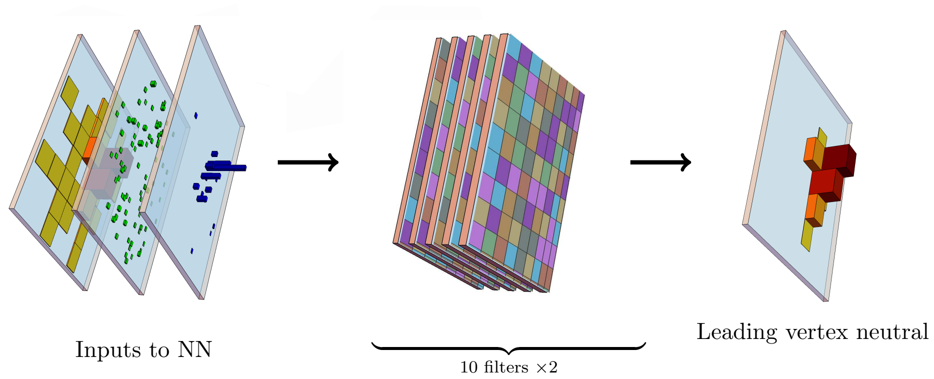
[figures from Cacciari, Salam, Soyez, 0802.1189; Planck Outreach]

Examples of Supervised Learning

Regression

e.g. PUMML for pileup mitigation

[Komiske, Metodiev, Nachman, Schwartz, 1707.08600;
see also Arjona Martínez, Cerri, Pierini, Spiropulu, Vlimant, 1810.07988]

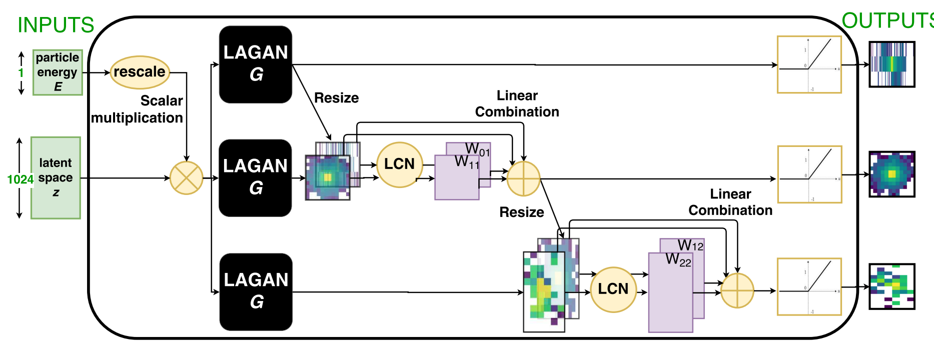


Labeled data: Objects J with property x
Solution: Map from J to x

Generation

e.g. CaloGAN for fast detector simulation

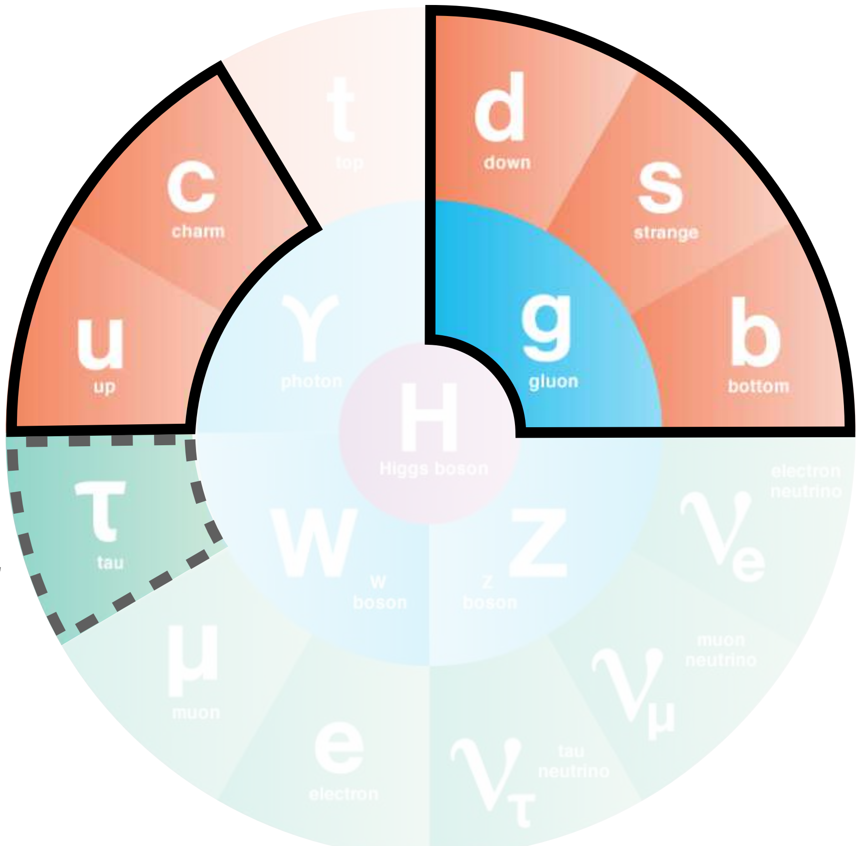
[Paganini, de Oliveira, Nachman, 1705.02355, 1712.10321;
see also de Oliveira, Michela Paganini, Nachman, 1701.05927]



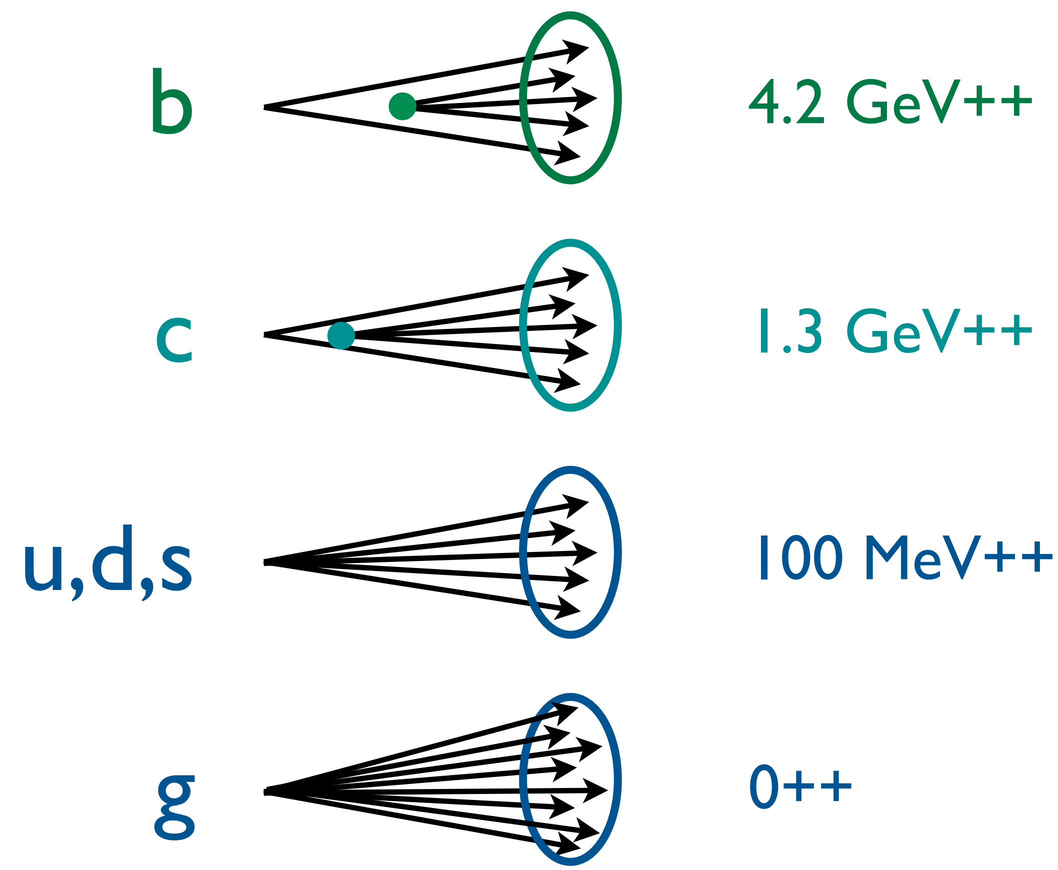
Labeled data: Objects J with property x
Solution: Map (conditioned on x) from noise to J

Jet Classification

Key supervised learning task at LHC



micro-jet
≈ 65%

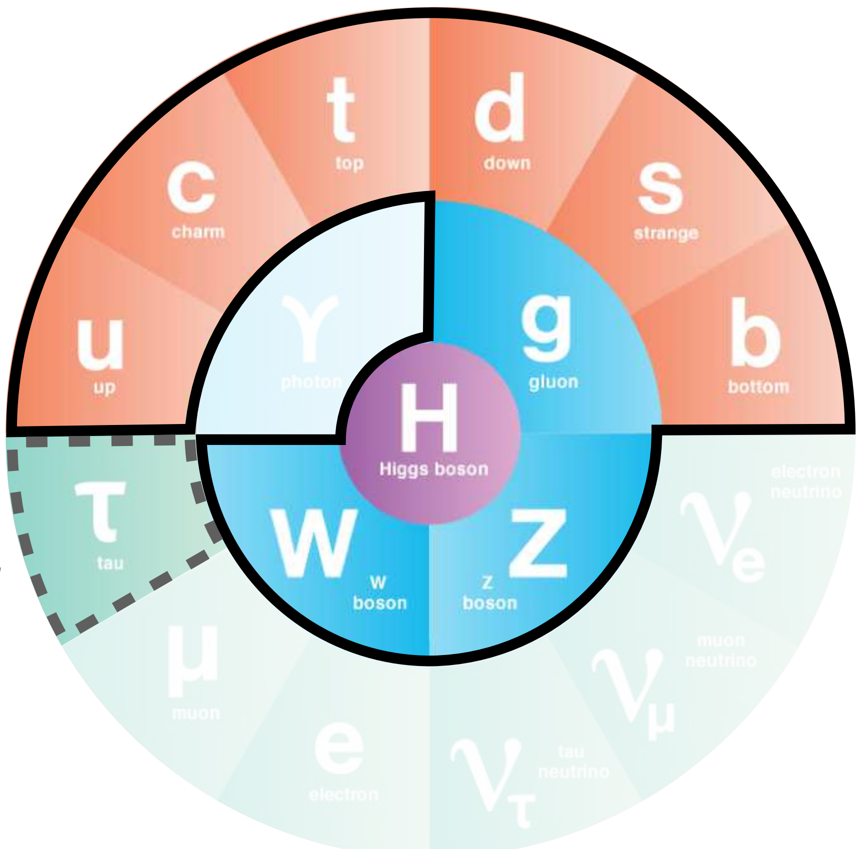


++ = Mass from QCD Radiation

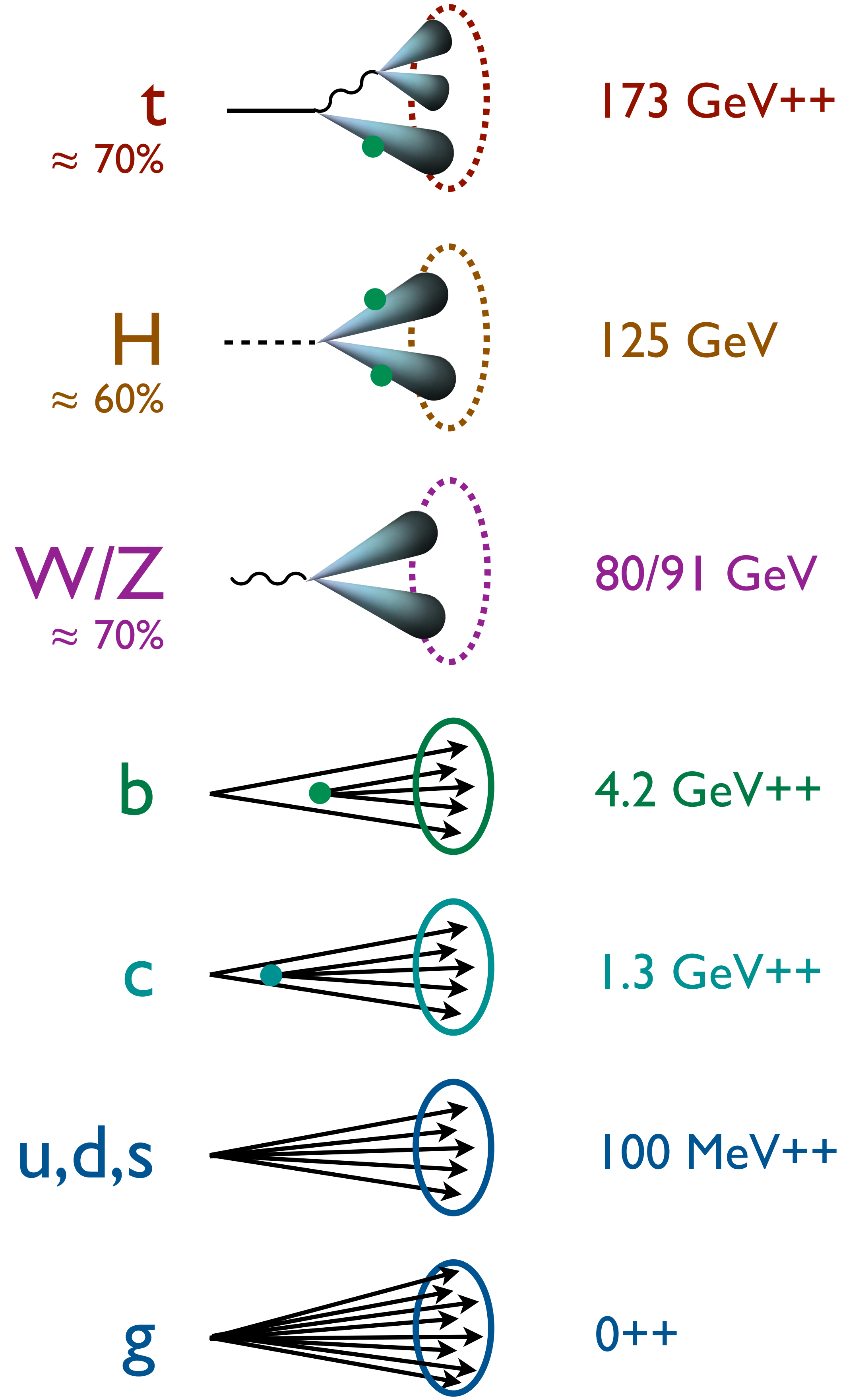
[see review in Larkoski, Moult, Nachman, 1709.04464]

Jet Classification

Key supervised learning task at LHC



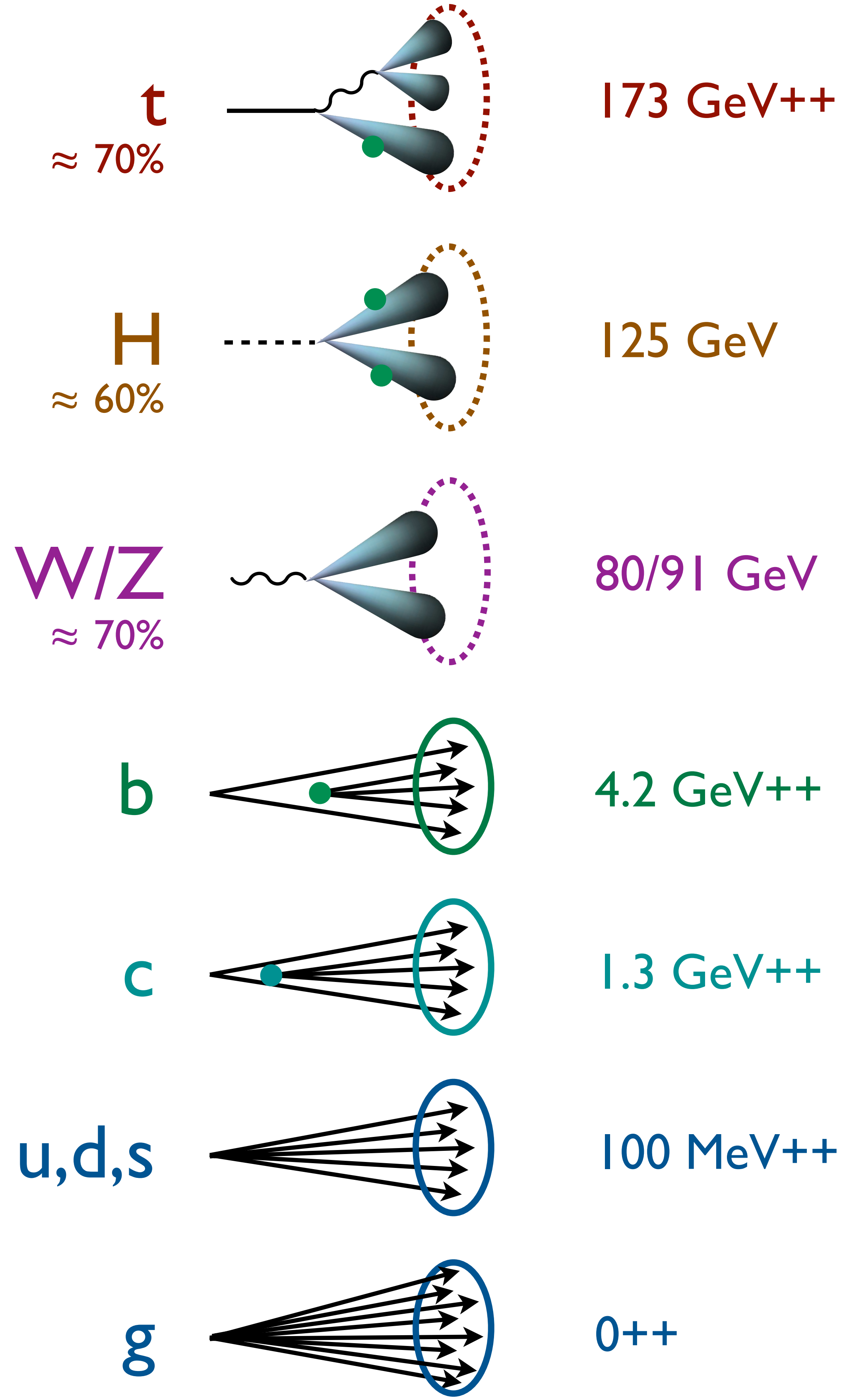
micro-jet
≈ 65%



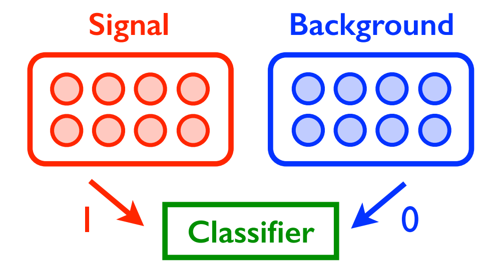
++ = Mass from QCD Radiation

[see review in Larkoski, Moult, Nachman, 1709.04464]

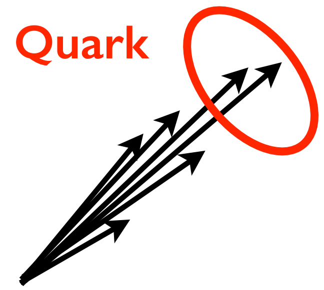

11th International Workshop on Boosted Objects
BOOSTON
2019
Phenomenology | Reconstruction | Searches | Algorithms | Measurements | Calculations
Modeling | Machine Learning | Pileup Mitigation | Heavy-Ion Collisions | Future Colliders
Local Organizing Committee:
Zeynep Demiragli (BU)
Philip Harris (MIT)
Yen-Jie Lee (MIT)
Matthew Schwartz (Harvard)
Jesse Thaler (MIT)
International Advisory Committee:
Ayana Arce (Duke)
Lily Asquith (Sussex)
Jon Butterworth (UCL)
Reina Camacho Toro (CNRS)
Mrinal Dasgupta (Manchester)
Robin Erbacher (UC Davis)
Gregor Kasieczka (Hamburg)
Andrew Larkoski (Reed)
Peter Loch (Arizona)
David Miller (Chicago)
Mihoko M. Nojiri (KEK)
Sal Rappoccio (Buffalo)
Gavin Salam (Oxford)
Alexander Schmidt (Aachen)
Ariel Schwartzman (SLAC)
Gregory Soyez (CNRS)
Marcel Vos (Valencia)
July 22-26, 2019
Stata Center, MIT
https://indico.cern.ch/e/boost2019


t
≈ 70%
173 GeV++
H
≈ 60%
125 GeV
W/Z
≈ 70%
80/91 GeV
b
4.2 GeV++
c
1.3 GeV++
u,d,s
100 MeV++
g
0++

Binary Classification

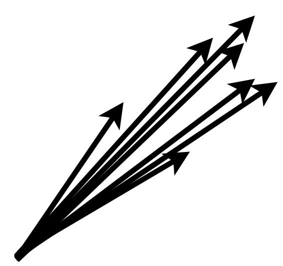


e.g.



vs.

assuming trustable training data

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

$$h\big(\textbf{Quark}\big) = 1$$

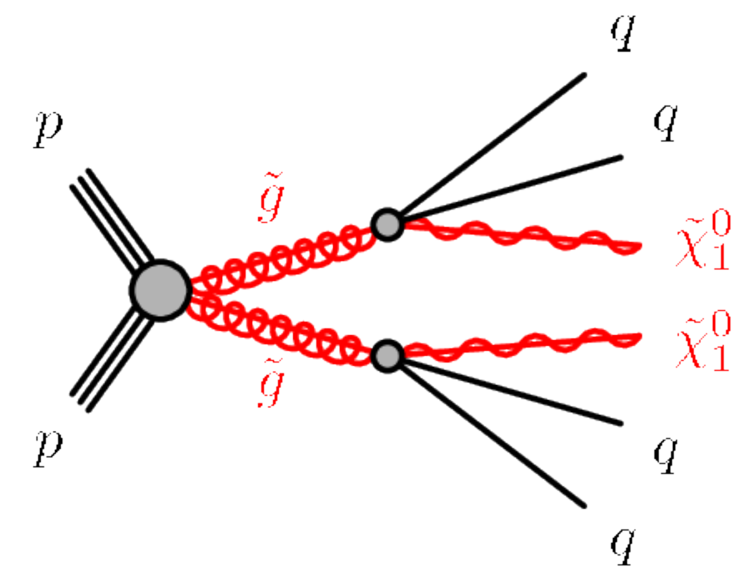
$$h\big(\textbf{Gluon}\big) = 0$$

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

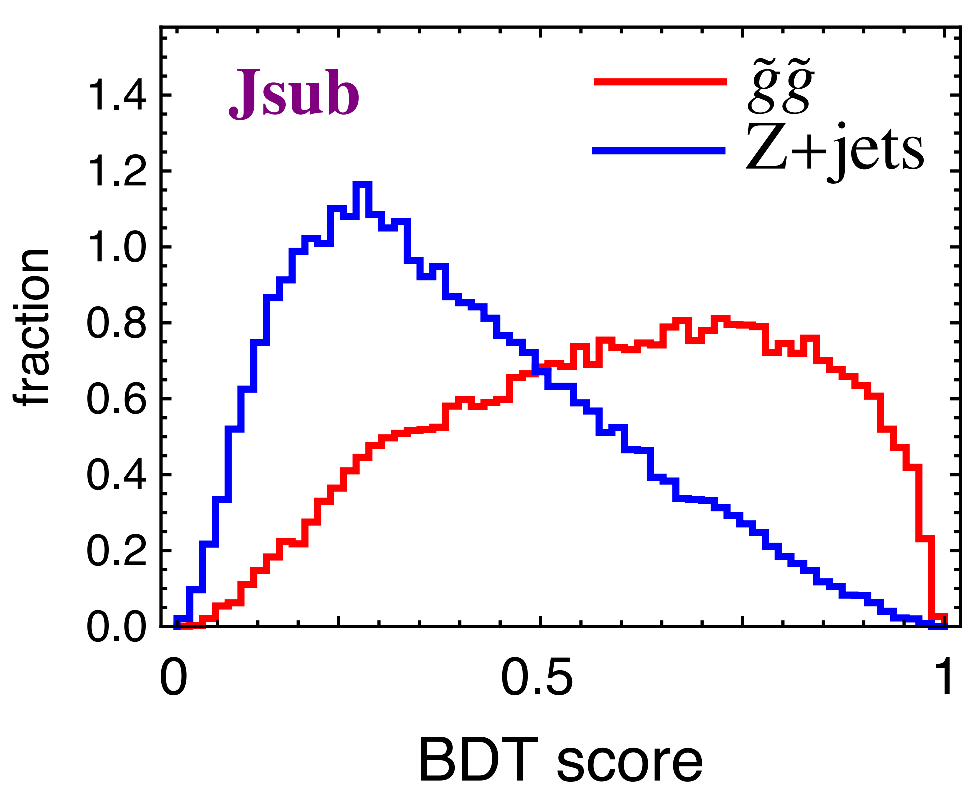
(Neyman-Pearson lemma)

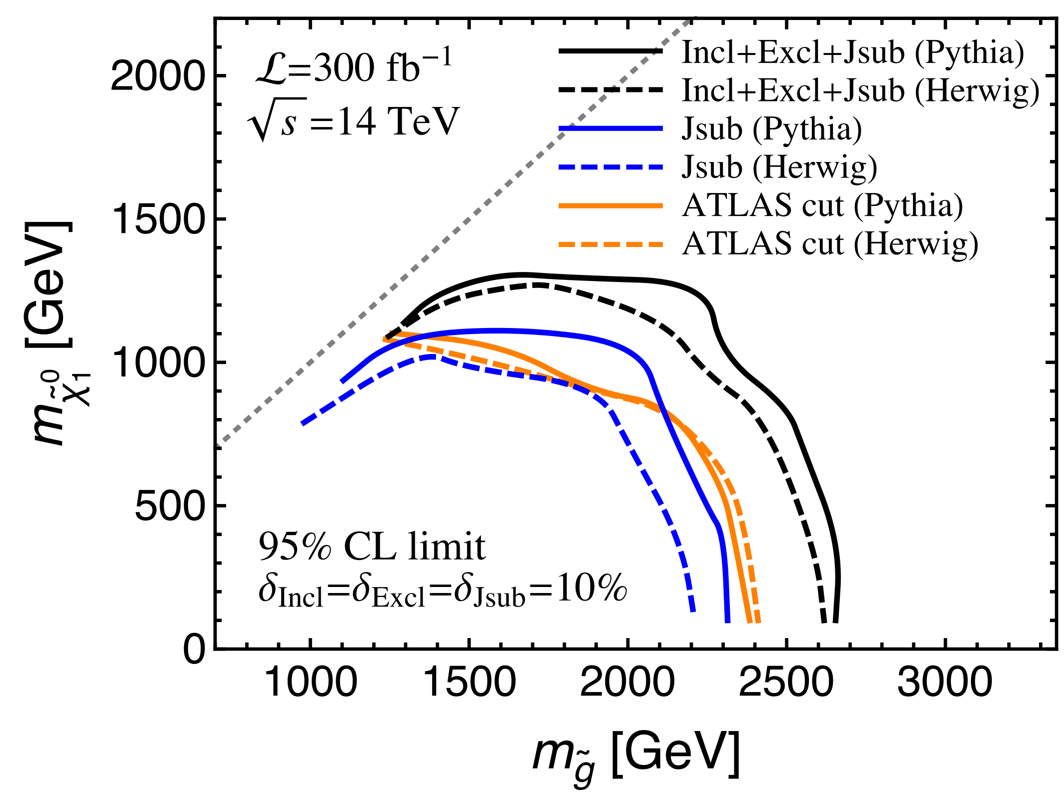
E.g. SUSY Search for Gluino Pairs

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



← background signal →

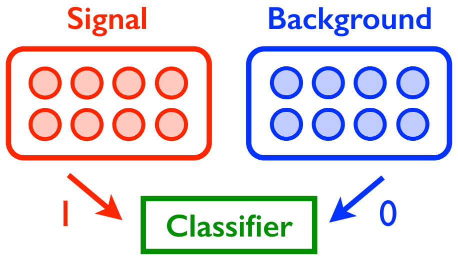




[Bhattacherjee, Mukhopadhyay, Nojiri, Sakakie, Webber, 1609.08781]

Jet Classification Studies

Mix and match



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

Loss Function

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
- …

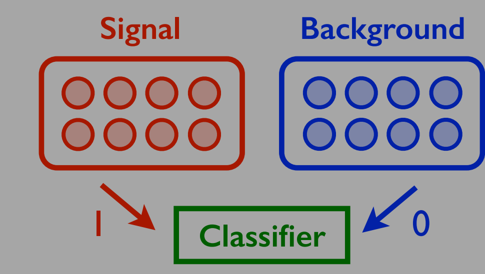
Signal vs. Background

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.	…

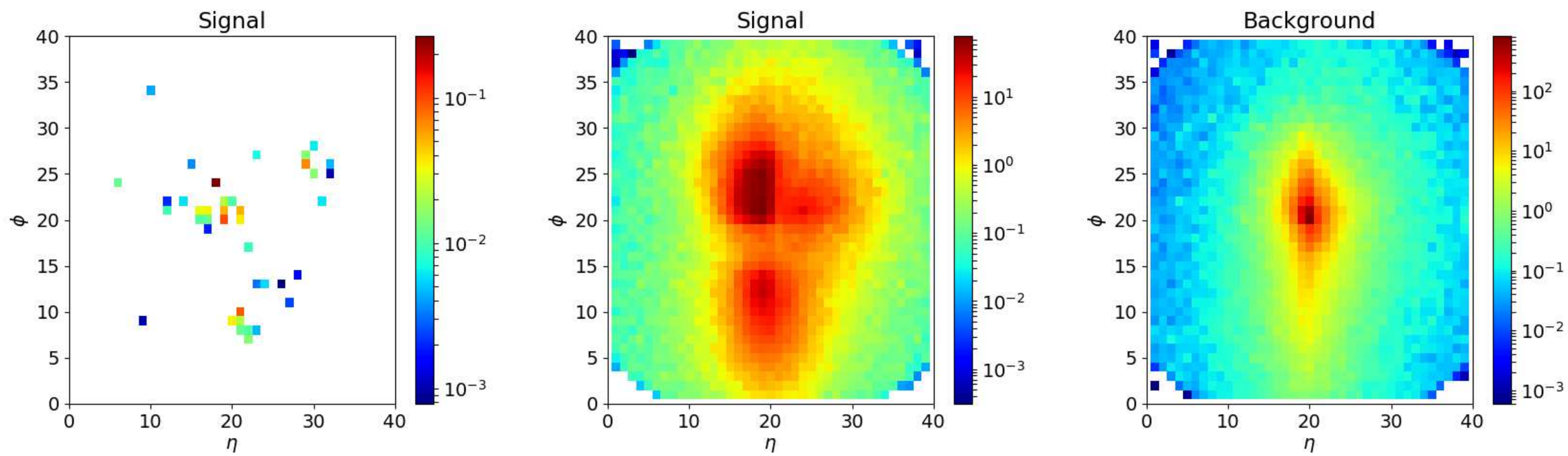
[Lönnblad, Peterson, Rögnvaldsson, PRL 1990, …, Cogan, Kagan, Strauss, Schwartzman, 1407.5675; Almeida, Backović, Cliche, Lee, Perelstein, 1501.05968; de Oliveira, Kagan, Mackey, Nachman, Schwartzman, 1511.05190; Baldi, Bauer, Eng, Sadowski, Whiteson, 1603.09349; Conway, Bhaskar, Erbacher, Pilot, 1606.06859;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

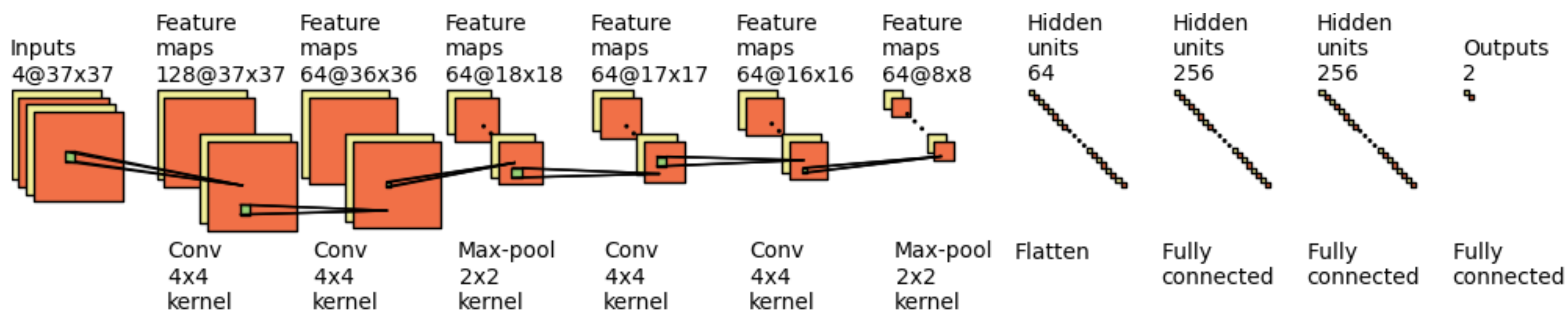
Jet Classification Studies

Mix and match



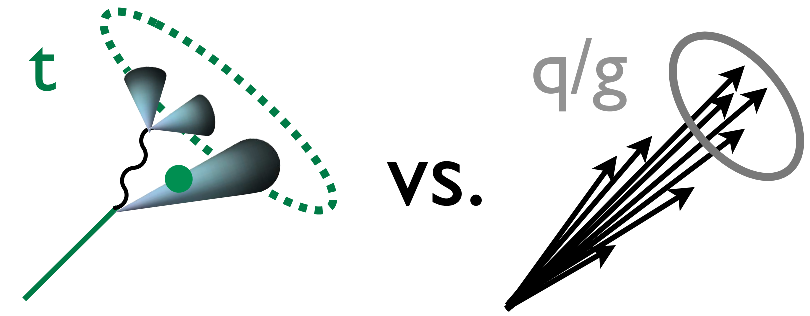
Deep Learning: Jet Image Strategy with CNNs

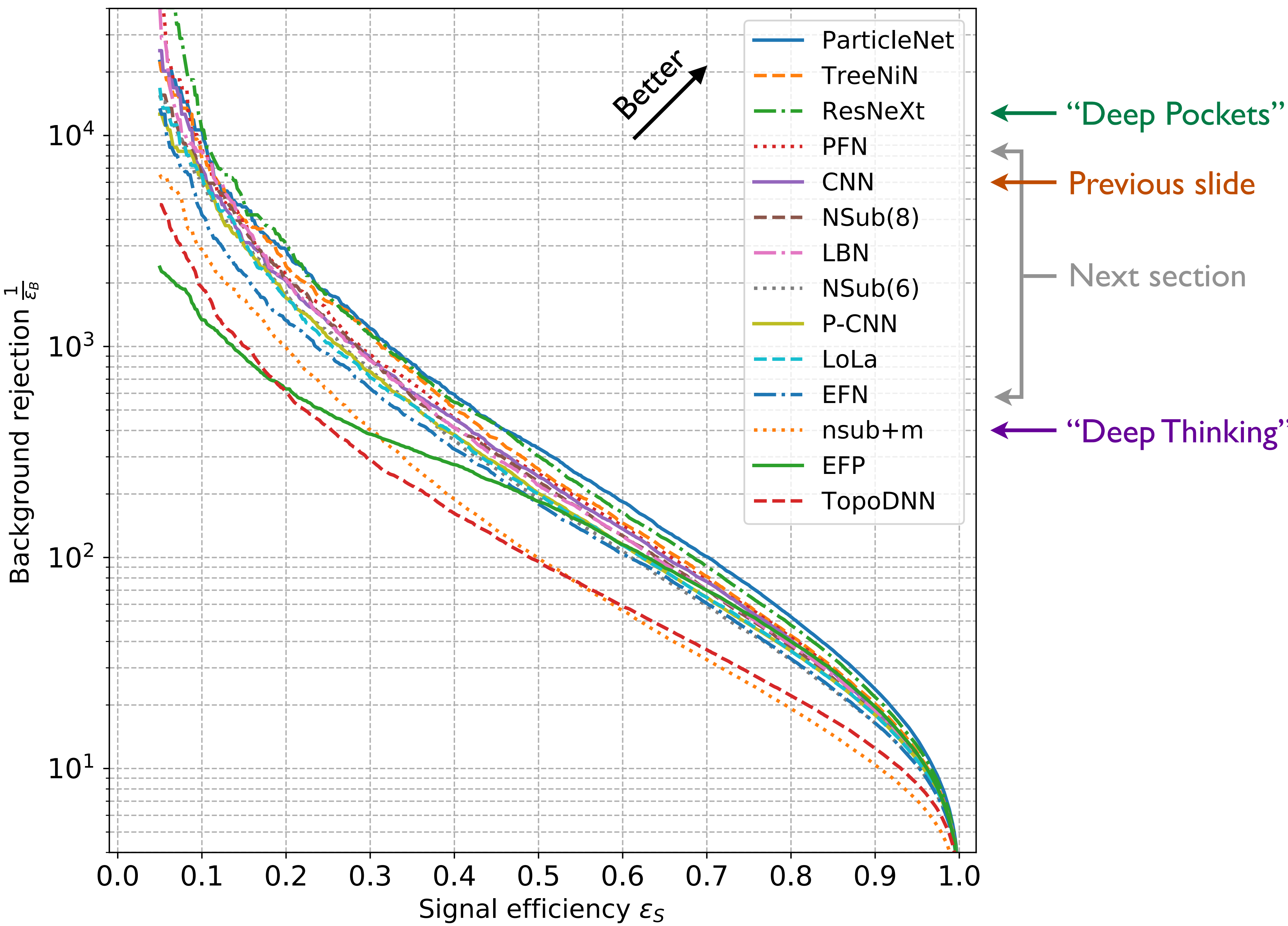




[Macaluso, Shih 1803.00107; building off Kasieczka, Plehn, Russell, Schell, 1701.08784; based on Cogan, Kagan, Strauss, Schwartzman, 1407.5675; de Oliveira, Kagan, Mackey, Nachman, Schwartzman, 1511.05190]

Throwing Down the Gauntlet

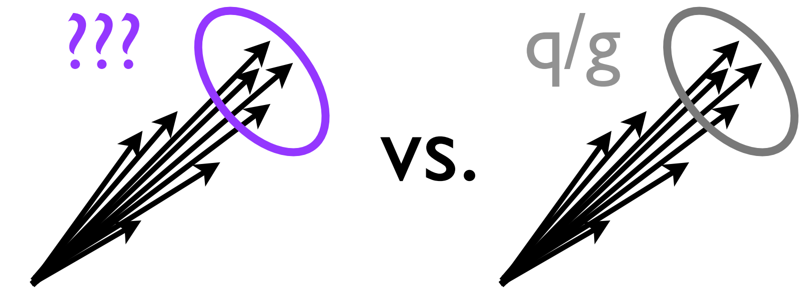


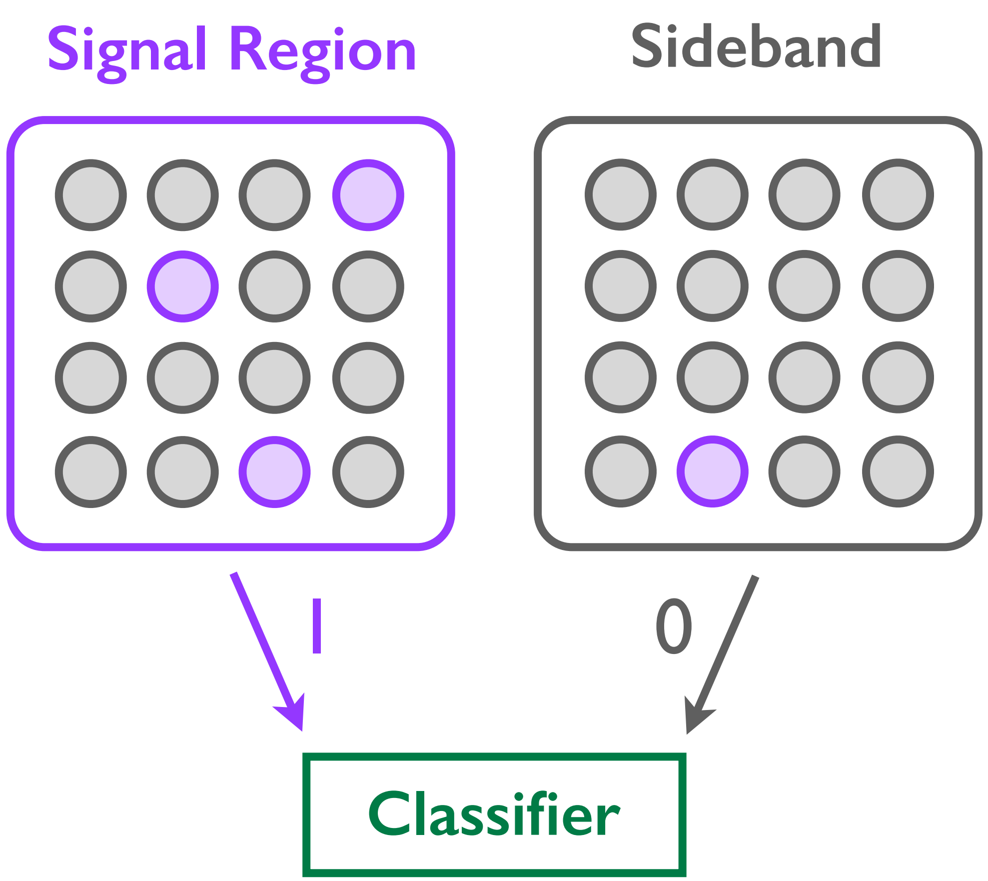


[Kasieczka, Plehn, et al., 1902.09914;
comparison of JDT, Van Tilburg, 1011.2268, 1108.2701; Xie, Girshick, Dollár, Tu, He, 1611.05431; CMS-DP-2017-049; Pearkes, Fedorko, Lister, Gay, 1704.02124;
Butter, Kasieczka, Plehn, Russell, 1707.08966; Komiske, Metodiev, JDT, 1712.07124; Macaluso, Shih 1803.00107; Moore, Nordström, Varma, Fairbairn 1807.04769;
Komiske, Metodiev, JDT, 1810.05165; Erdmann, Geiser, Rath, Rieger, 1812.09722; Qu, Gouskos, 1902.08570; Macaluso, Cranmer, to appear]

CWoLa Hunting

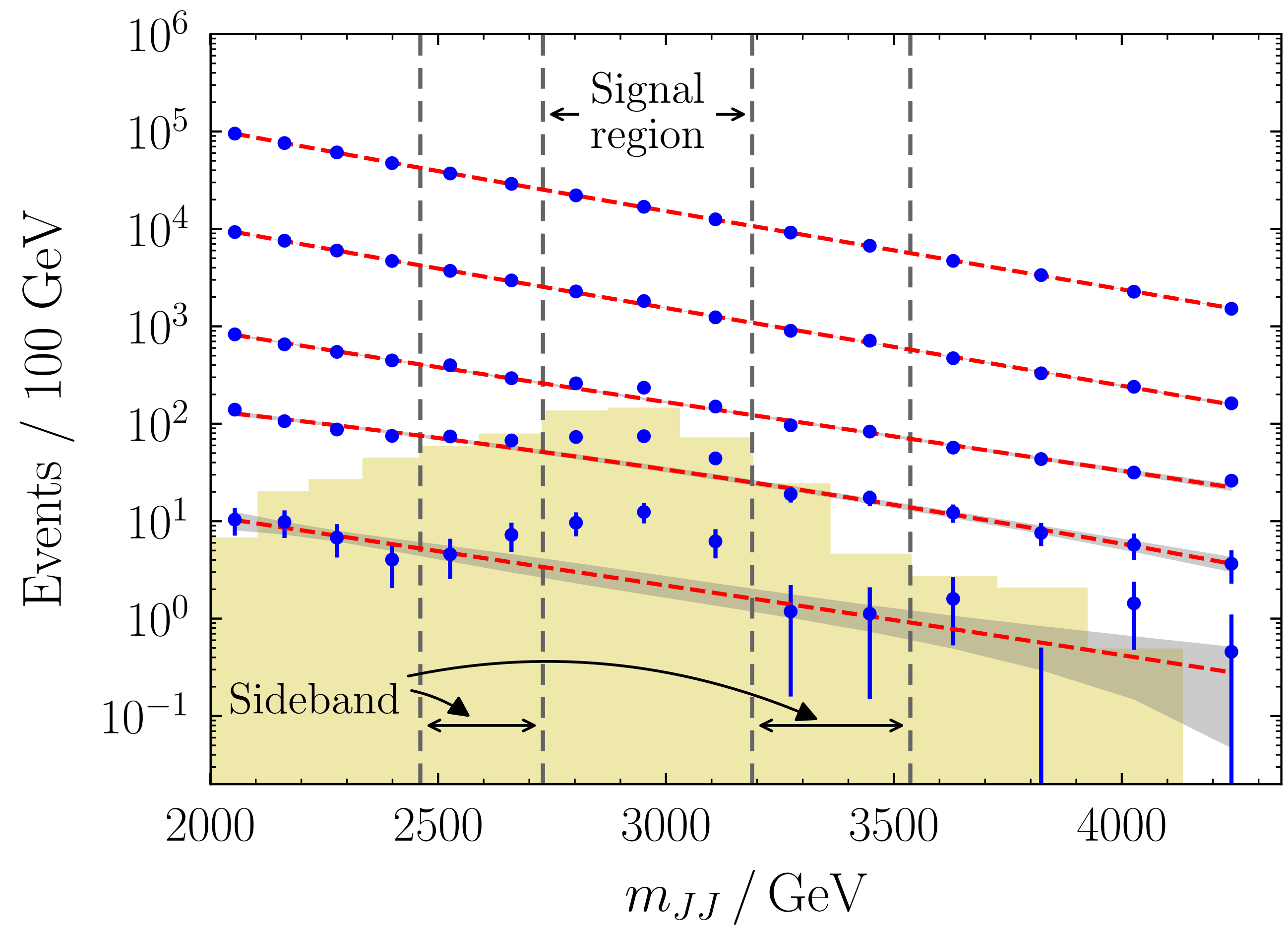
Using "Classification Without Labels"





With enough data, monotonic w.r.t. optimal classifier (!)

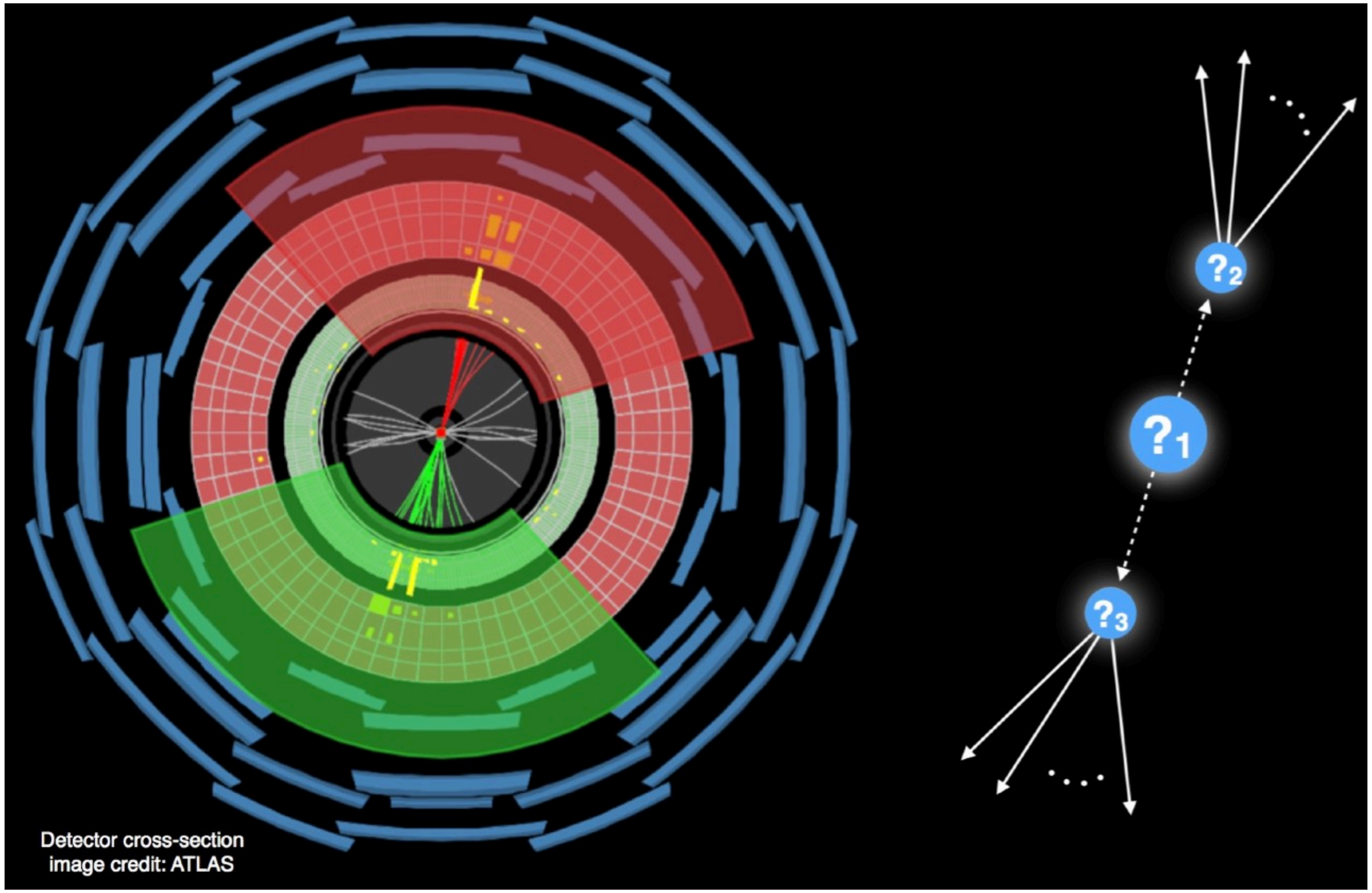
Model-Agnostic Bump Hunt



[Collins, Howe, Nachman, 1805.02664, 1902.02634; using Metodiev, Nachman, JDT, 1708.02949; see also Blanchard, Flaska, Handy, Pozzi, Scott, 1303.1208; Cranmer, Pavez, Louppe, 1506.02169]

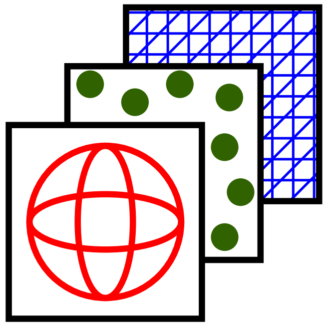
LHC Olympics 2020

@ ML4Jets, NYU, January 15-17



An opportunity to stress test new anomaly detection strategies

The Rise of Deep Learning



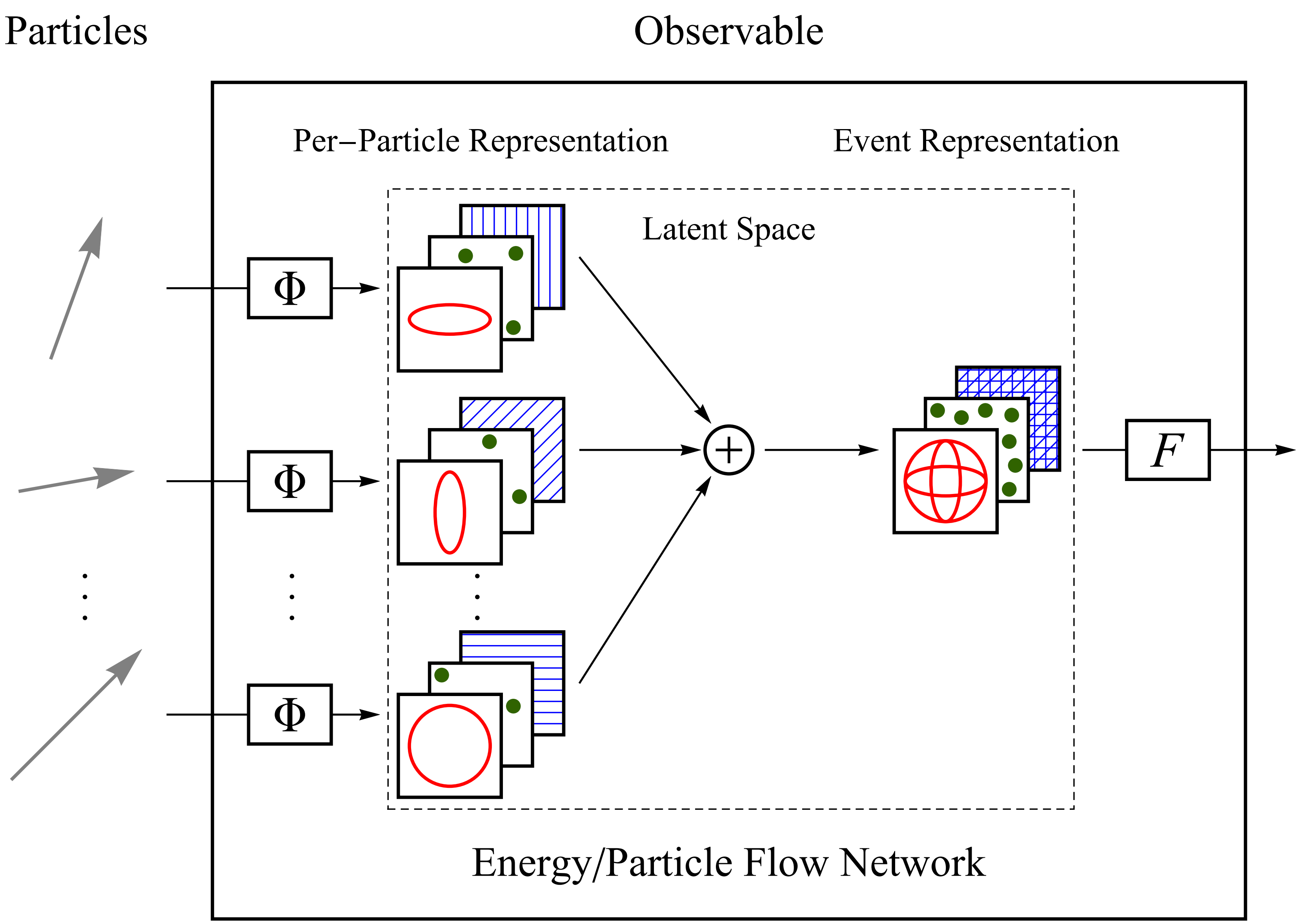
Looking Inside the Black Box

(Exploring the Space of Jets)

Introducing Energy Flow Networks

An architecture designed for interpretability

(see backup for detailed architecture)



[Komiske, Metodiev, JDT, 1810.05165;
special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114]

Introducing Energy Flow Networks

(see backup for detailed architecture)

An architecture designed for interpretability

Latent space of dim ℓ

Linear weights

$$\mathcal{S}(\mathcal{J}) = F\left(V_1, V_2, \ldots, V_\ell\right) \qquad V_a(\mathcal{J}) = \sum_{i\in\mathcal{J}} p_{Ti}\, \Phi_a(y_i, \phi_i)$$

Parametrized with Neural Networks

Flexible enough to describe any* IRC-safe observable

(assuming large enough ℓ)

Generalization: Particle Flow Networks (aka "Deep Sets")

[Komiske, Metodiev, JDT, 1810.05165;
special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114]

Introducing Energy Flow Networks

(see backup for
detailed architecture)

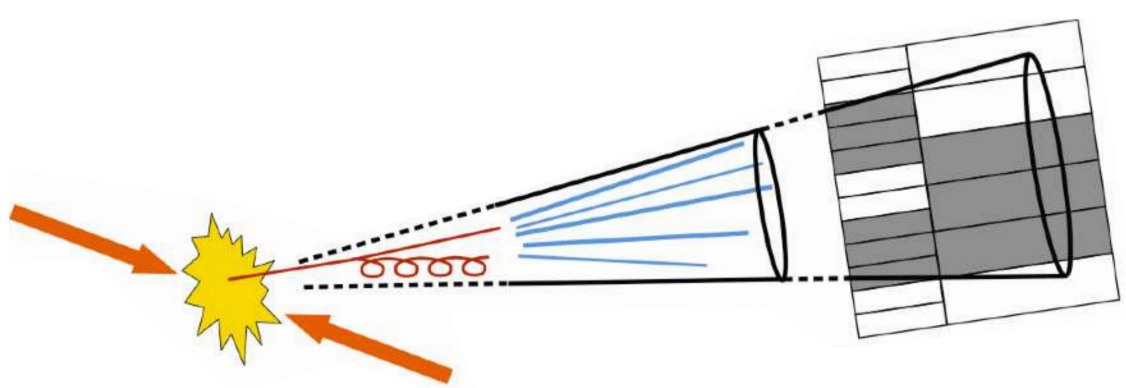
An architecture designed for interpretability

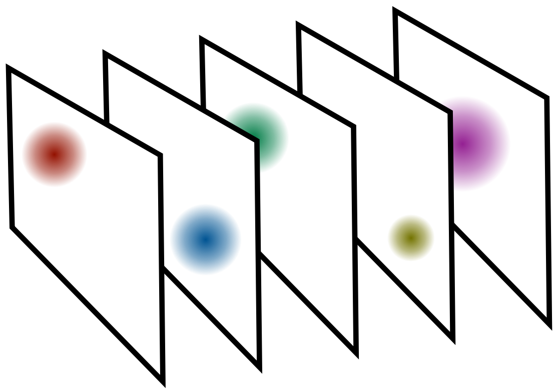
Visualization Strategy

$$S(\mathcal{J}) = F(V_1, V_2, \ldots, V_\ell) \qquad V_a(\mathcal{J}) = \sum_{i\in\mathcal{J}} p_{Ti}\, \Phi_a(y_i, \phi_i)$$

Difficult to visualize
(unless ℓ is small)

Easy to plot these!



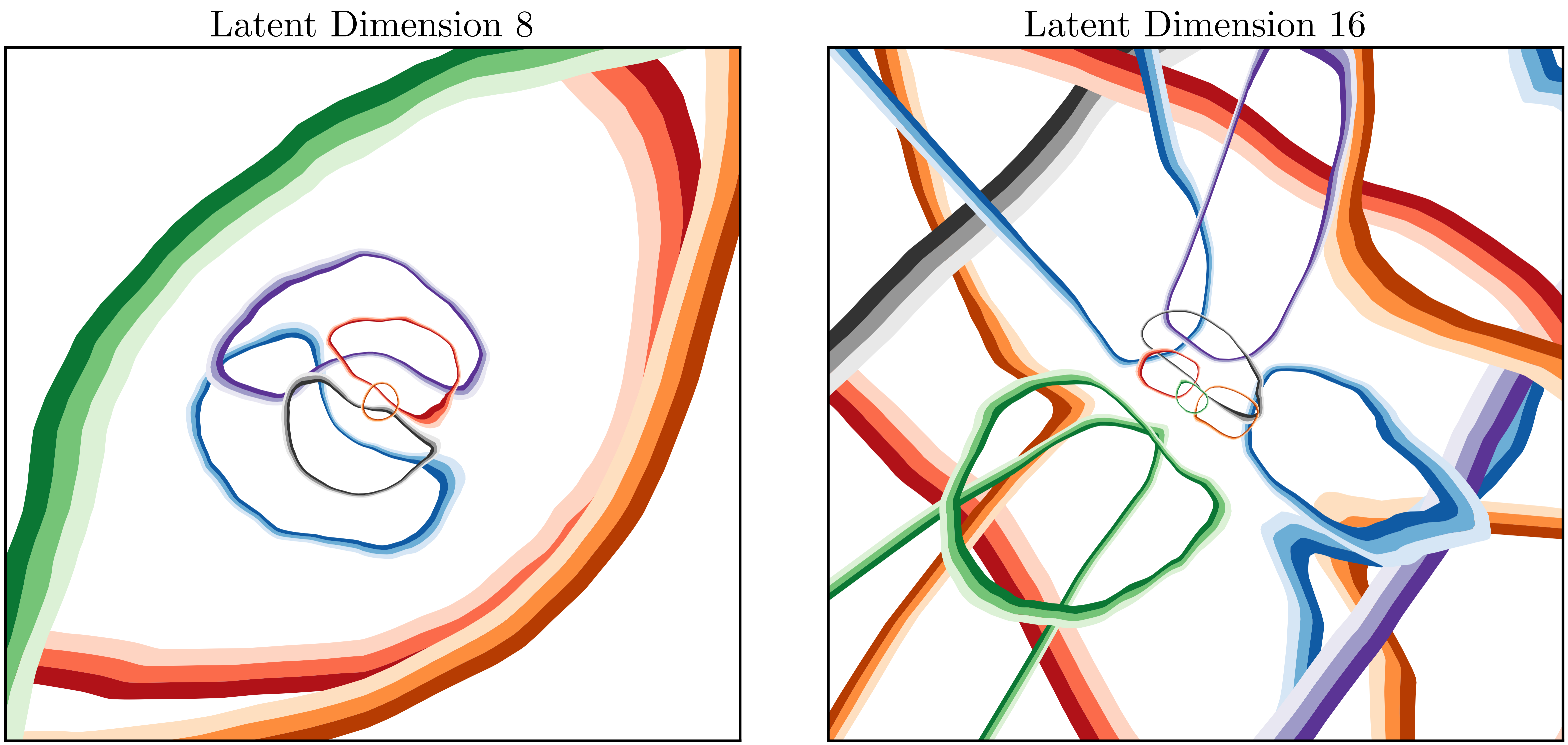


(similar to CNN
filter activation)

[Komiske, Metodiev, JDT, 1810.05165;
special case of Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114]

Psychedelic Network Visualization

(see backup for how these are made)

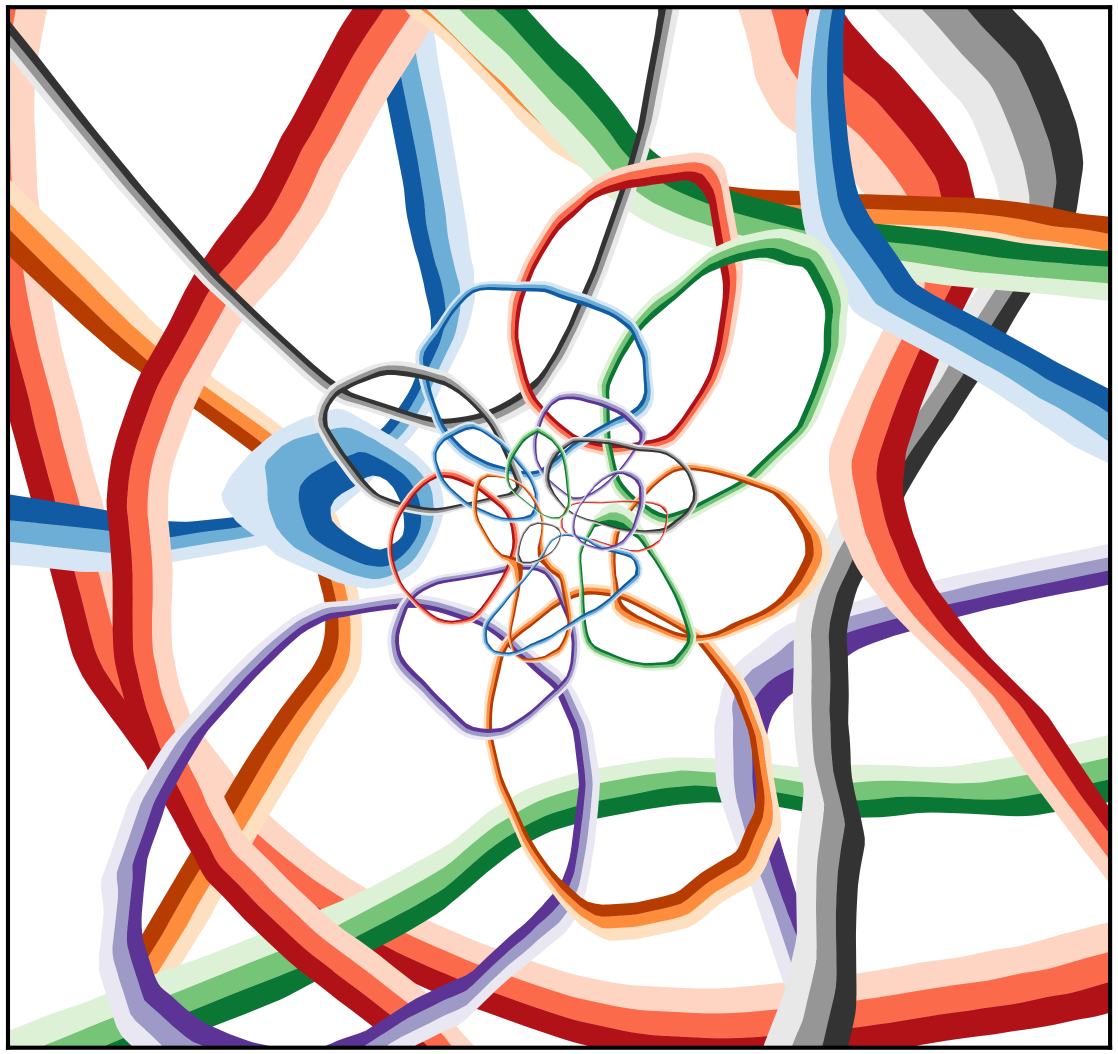


For the case of quark vs. gluon classification

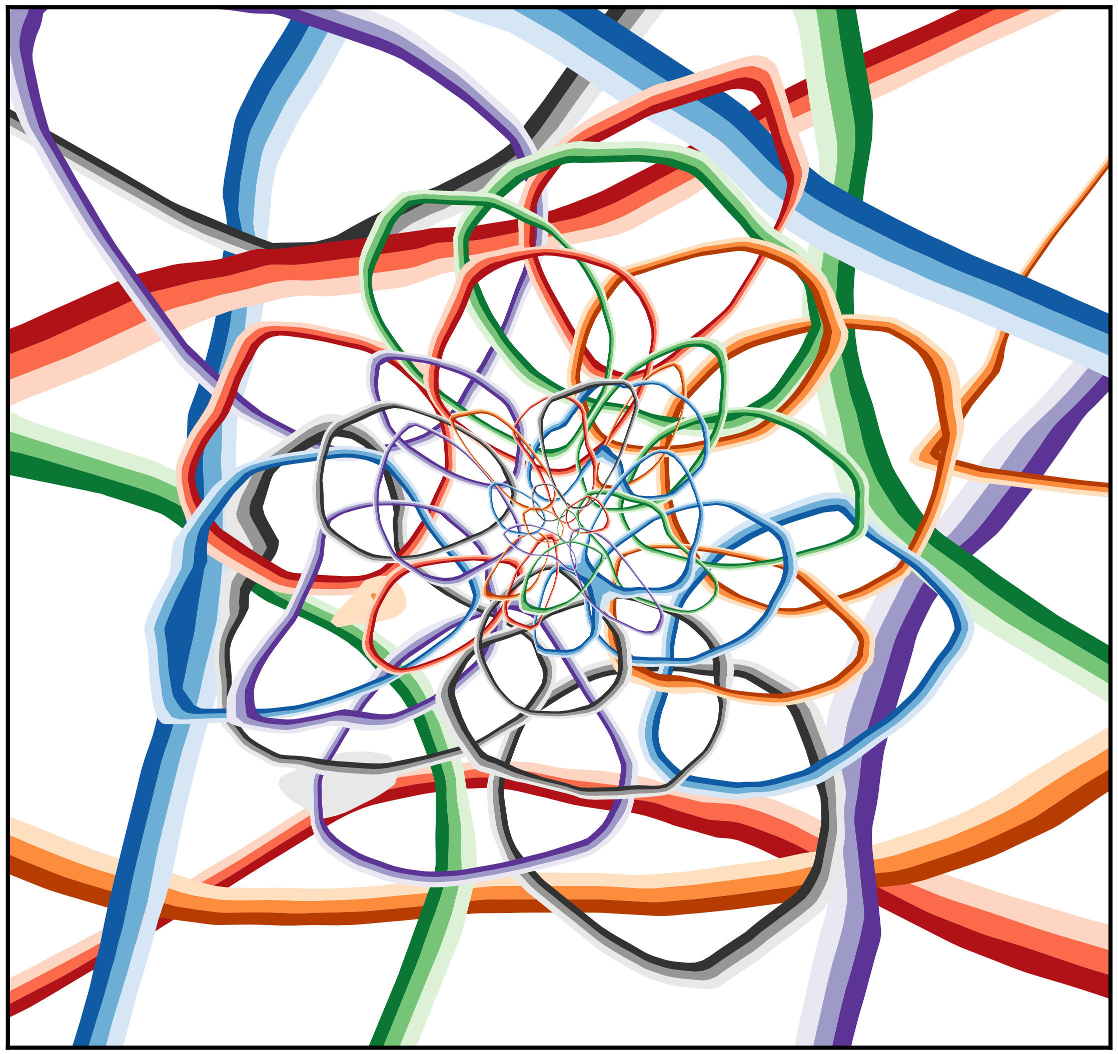
Psychedelic Network Visualization

(see backup for
how these are made)

Latent Dimension 32

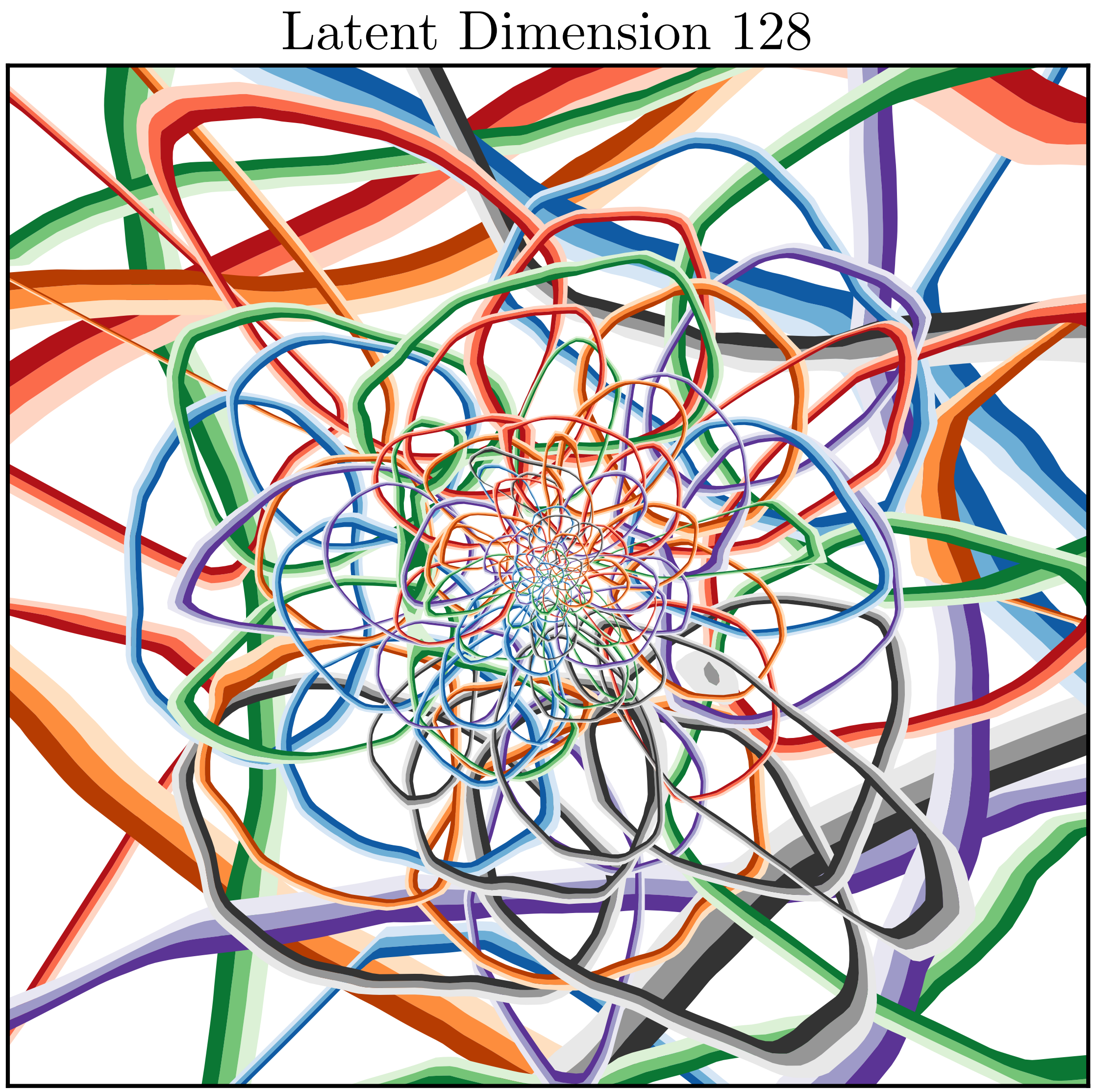


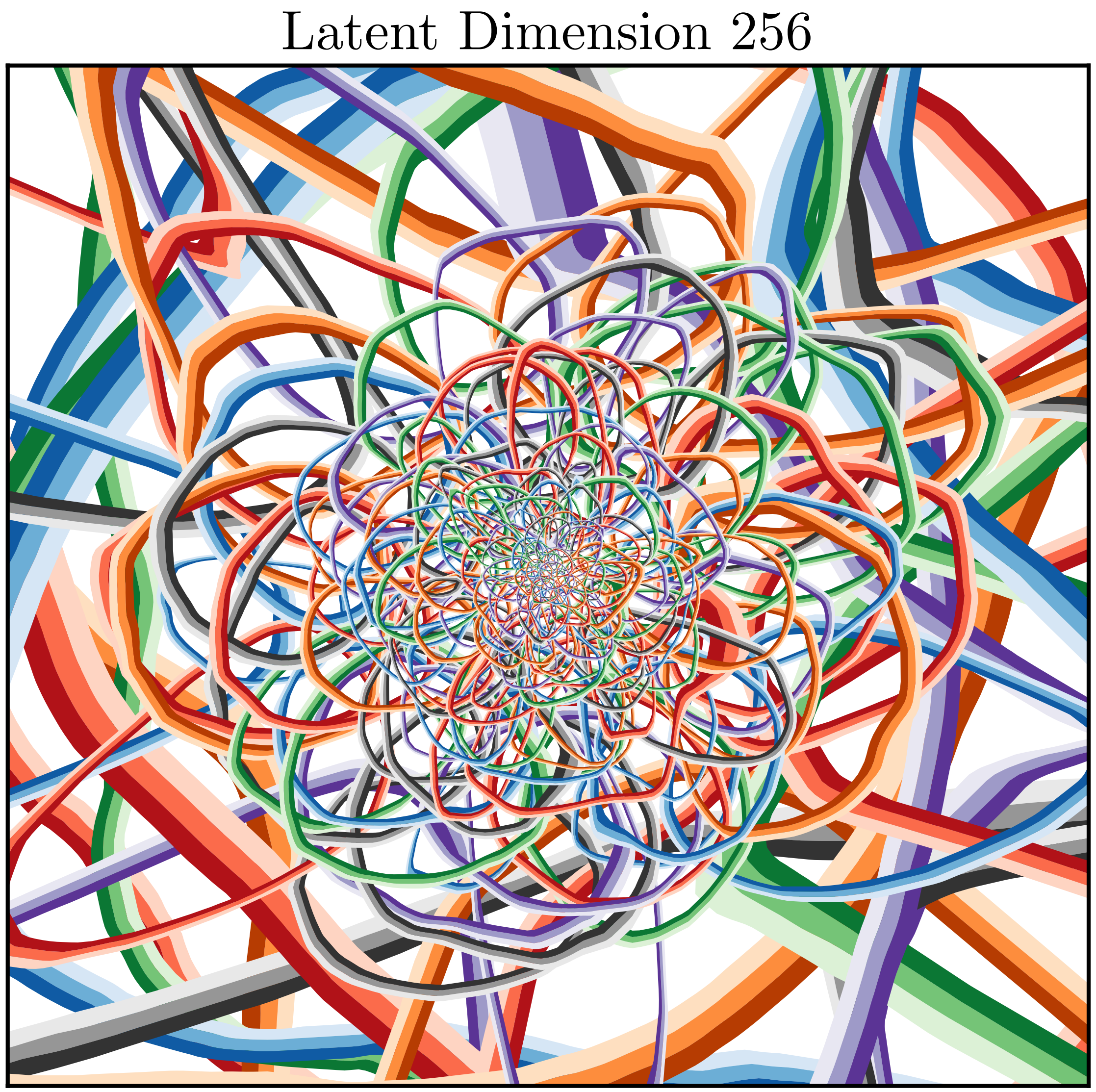
Latent Dimension 64



Psychedelic Network Visualization

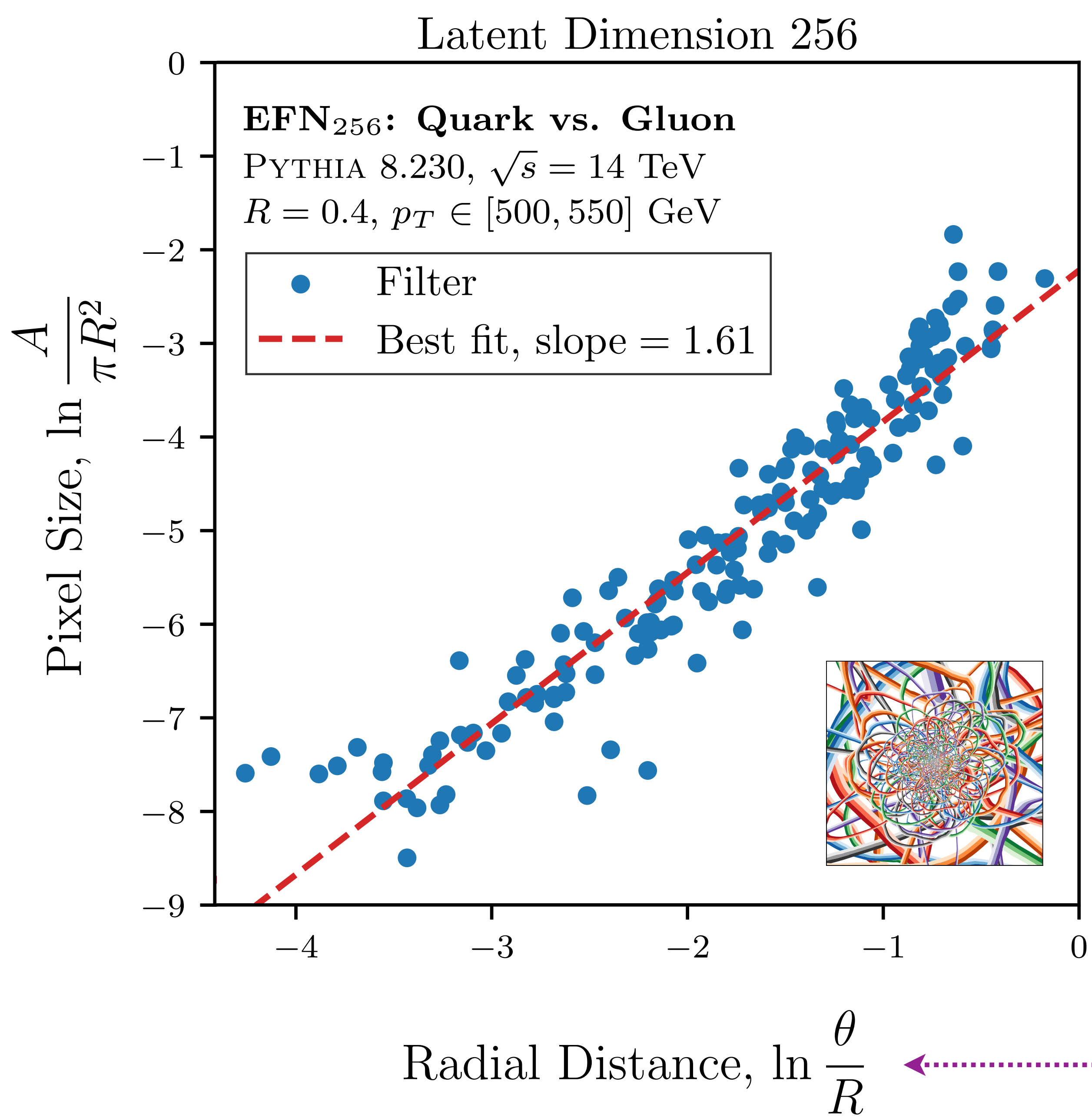
(see backup for how these are made)

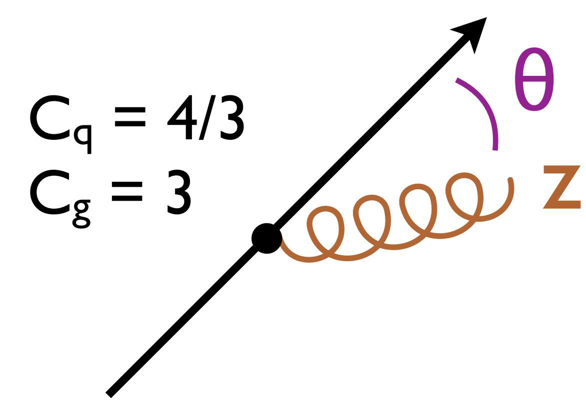




Singularity structure of QCD!

Putting the AI in AItarelli-Parisi

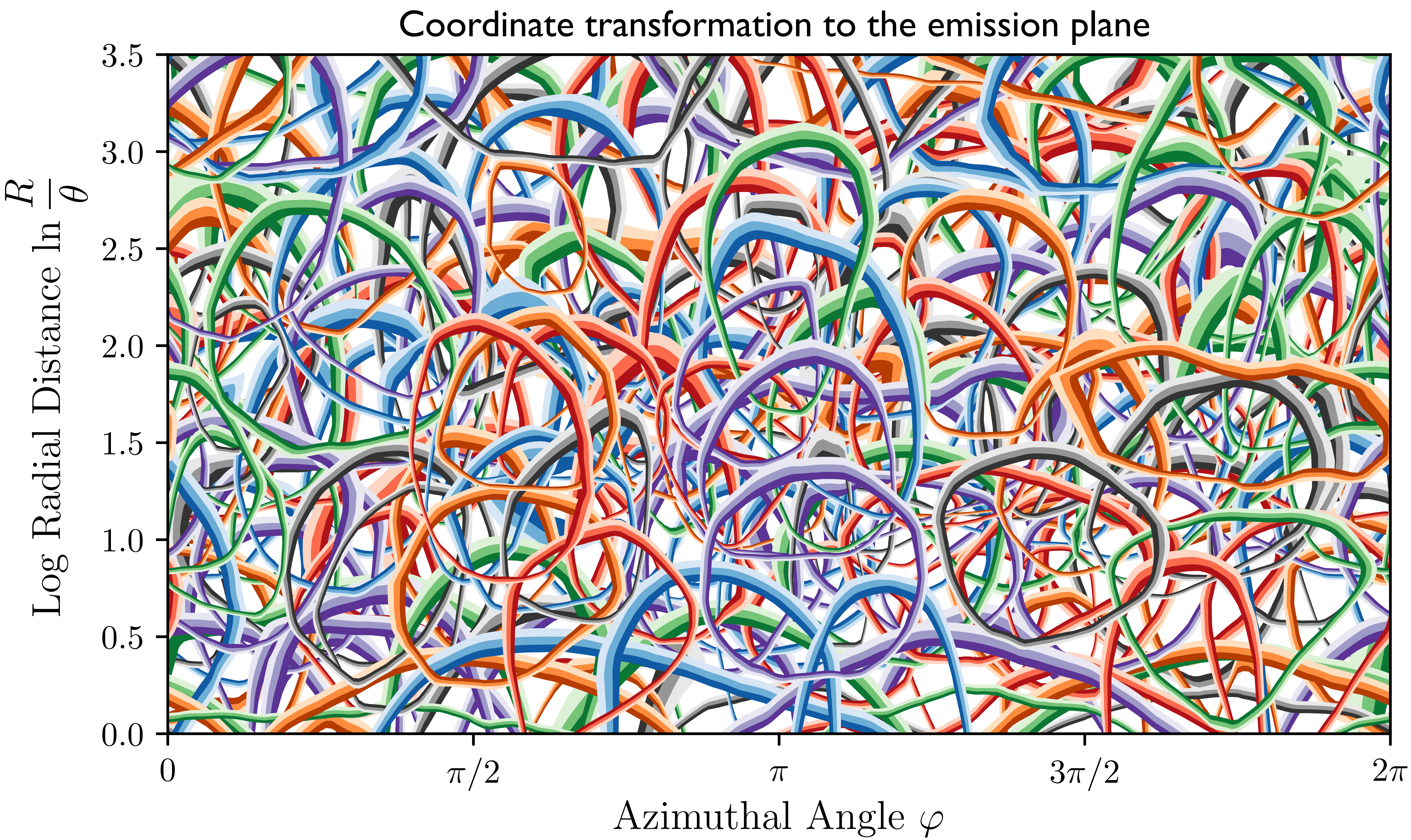




$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \underbrace{\frac{\mathrm{d}\theta}{\theta}}_{\text{Collinear}} \underbrace{\frac{\mathrm{d}z}{z}}_{\text{Soft}}$$

[Komiske, Metodiev, JDT, 1810.05165]

Ready for the CMOA?



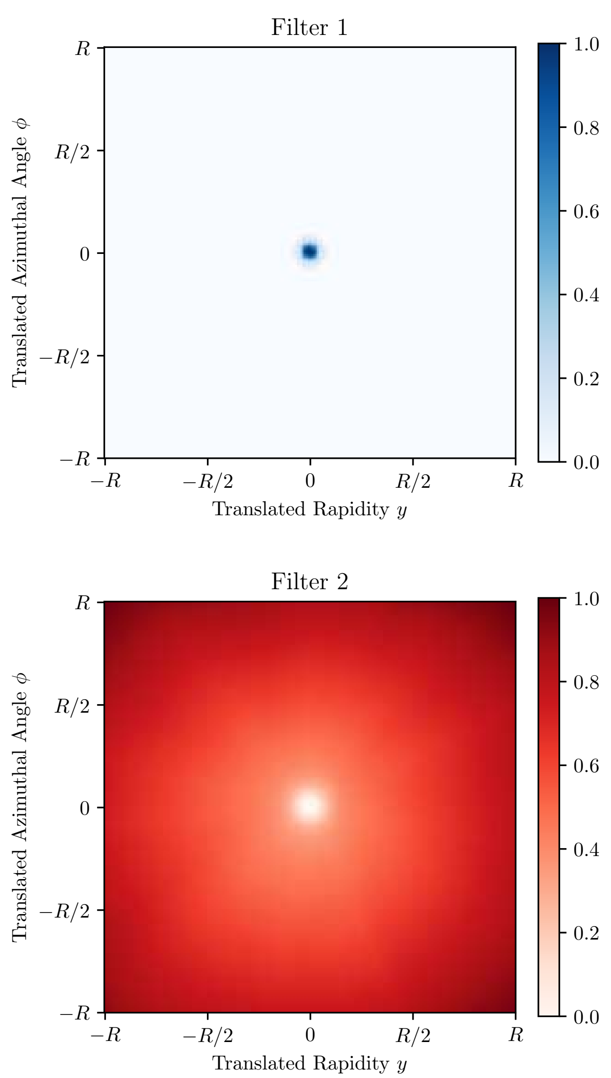
[Komiske, Metodiev, JDT, 1810.05165; see also Dreyer, Salam, Soyez, 1807.04758]

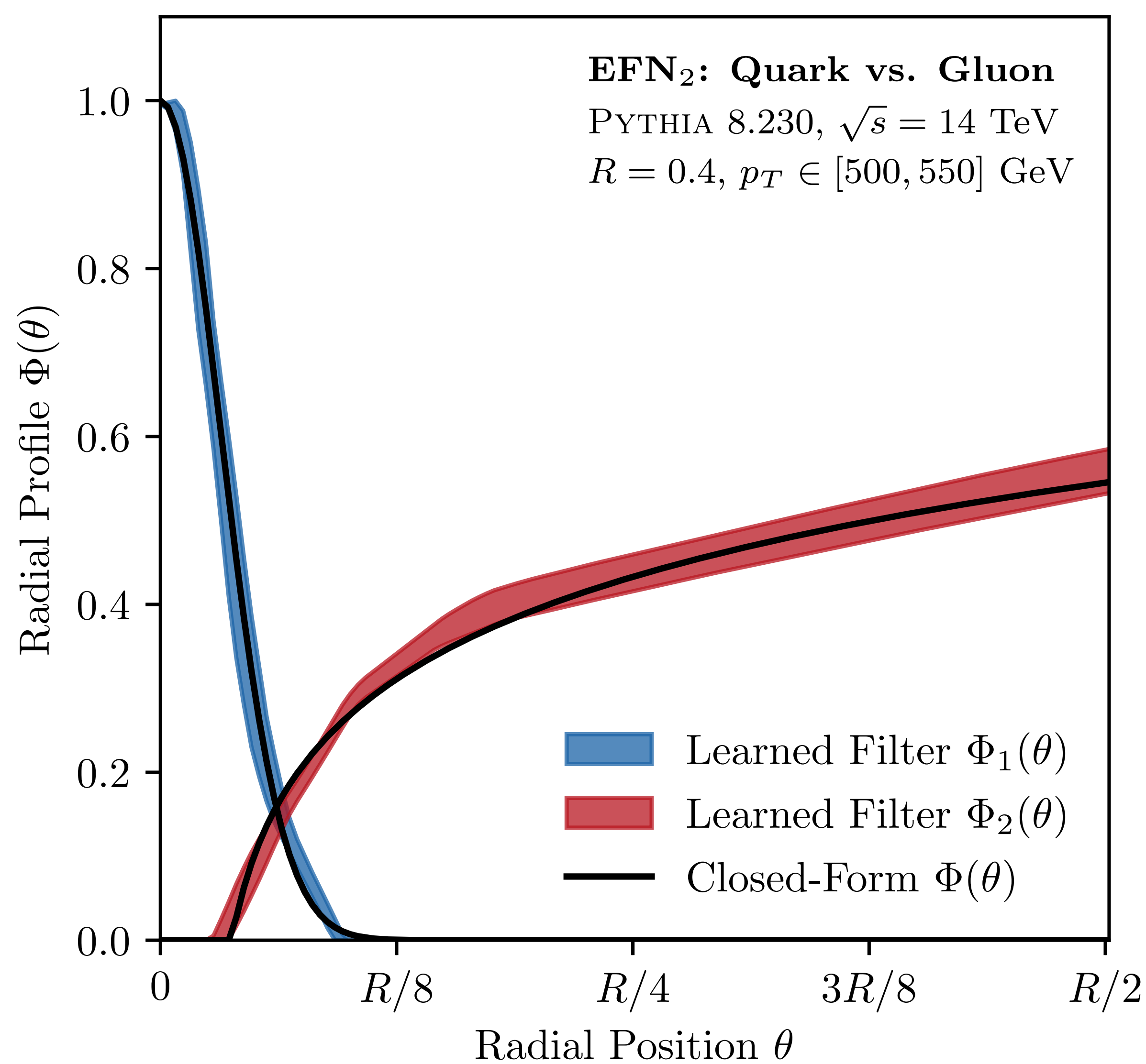
“Ok, but did you really learn something you didn’t already know?”

Learning from the Machine

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

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



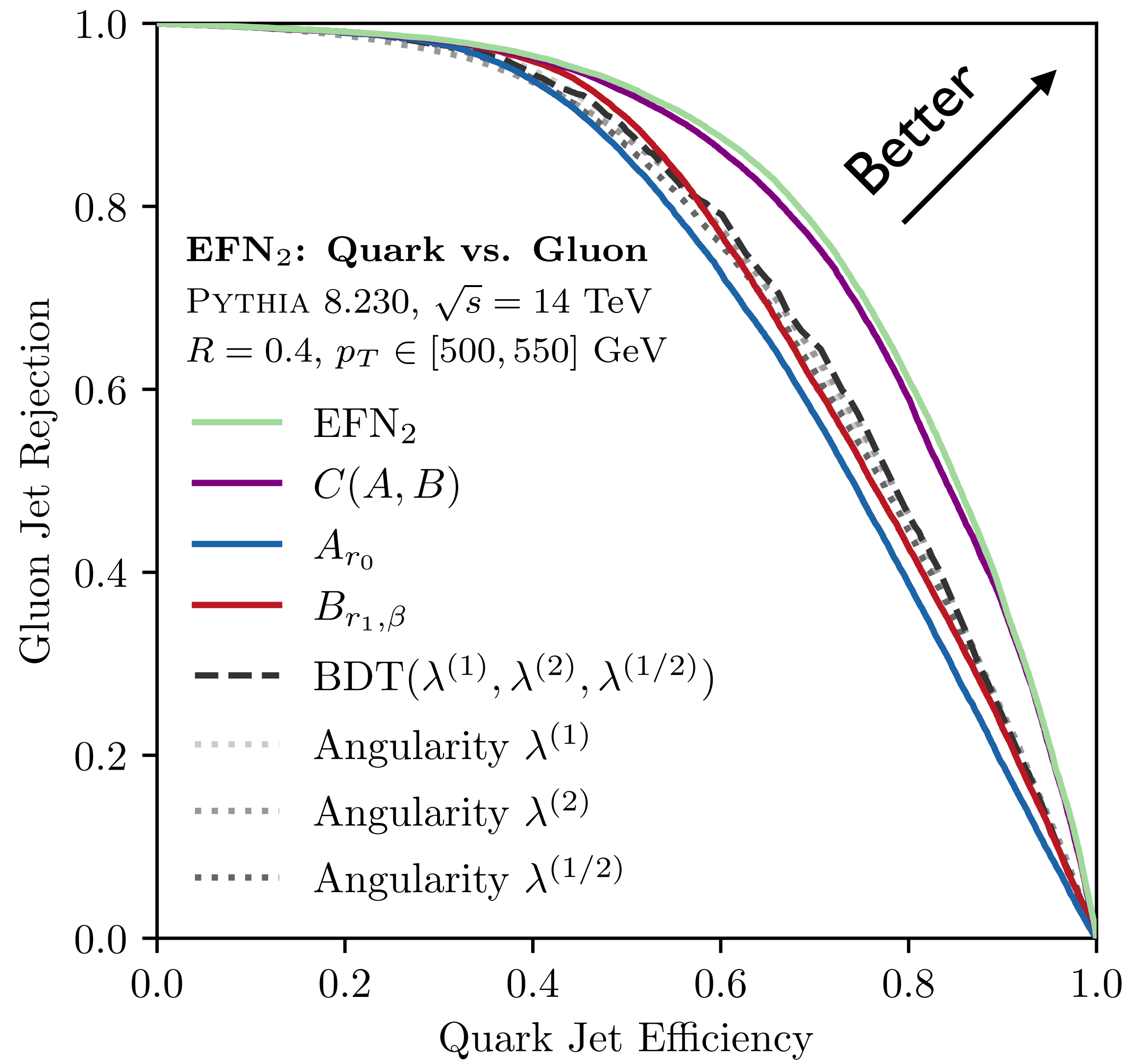


[Komiske, Metodiev, JDT, 1810.05165]

Learning from the Machine

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

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



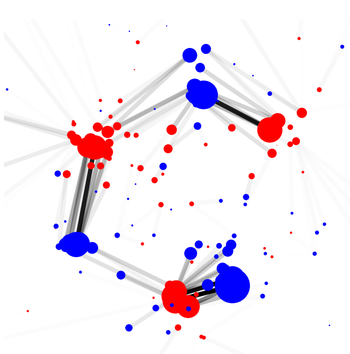
EFN outperformed a domain expert (i.e. me)

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

[Komiske, Metodiev, JDT, 1810.05165]

The Rise of Deep Learning

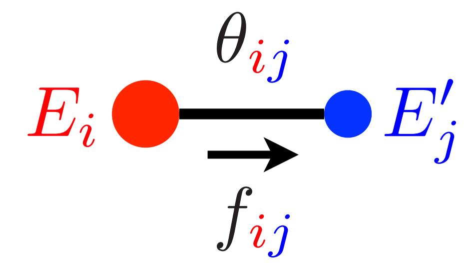
Looking Inside the Black Box

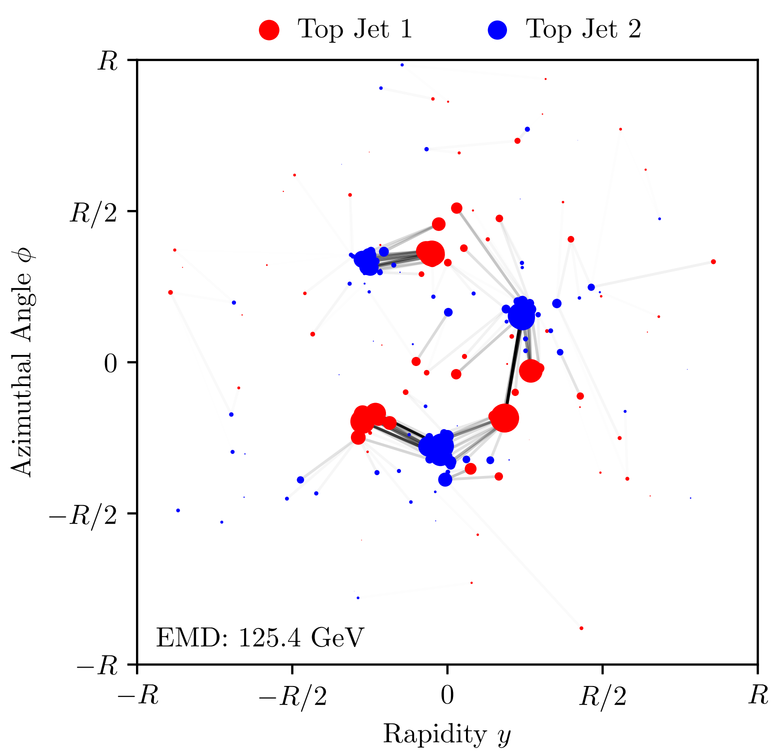


(Exploring the Space of Jets)

The Energy Mover's Distance

Closely related to 1-Wasserstein metric





Optimal transport between energy flows…

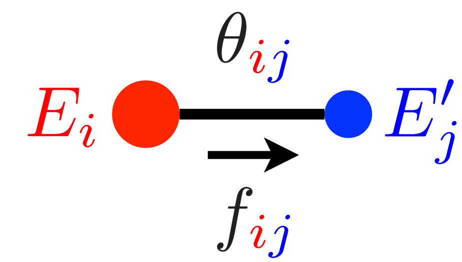
$$\text{EMD}(\mathcal{E}, \mathcal{E}') = \underbrace{\min_{\{f\}} \sum_i \sum_j f_{ij} \frac{\theta_{ij}}{R}}_{\text{Cost to move energy}} + \underbrace{\left| \sum_i E_i - \sum_j E'_j \right|}_{\text{Cost to create energy}}$$

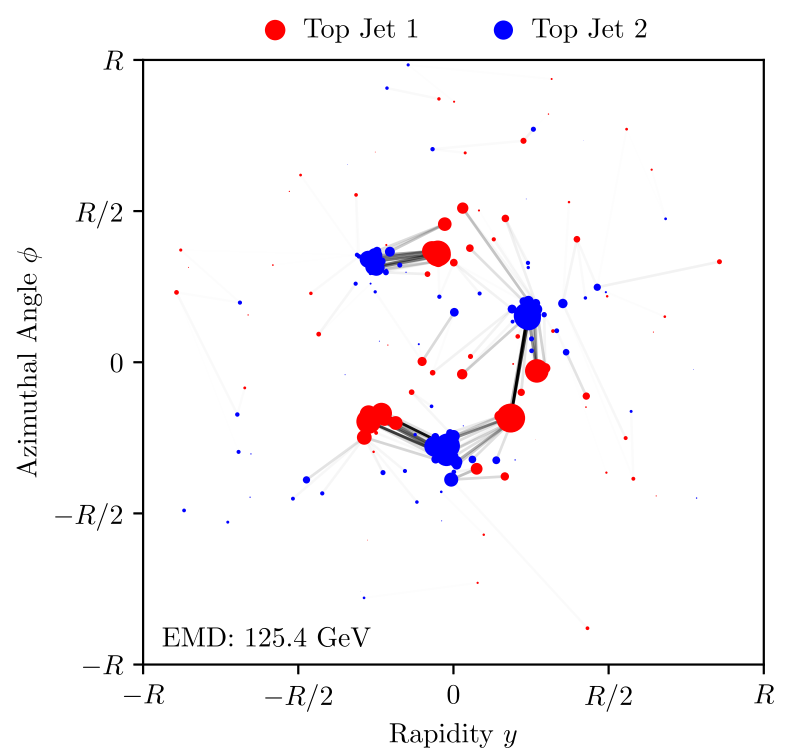
in GeV

[Komiske, Metodiev, JDT, 1902.02346;
see also Peleg, Werman, Rom, IEEE 1989; Rubner, Tomasi, Guibas, ICCV 1998, ICJV 2000; Pele, Werman, ECCV 2008; Pele Taskar, GSI 2013]

The Energy Mover's Distance

Closely related to 1-Wasserstein metric

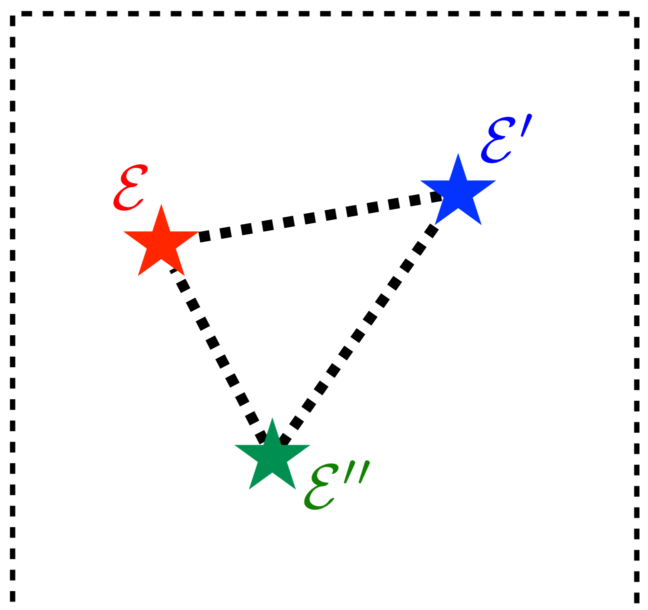




Optimal transport between energy flows…

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

in GeV



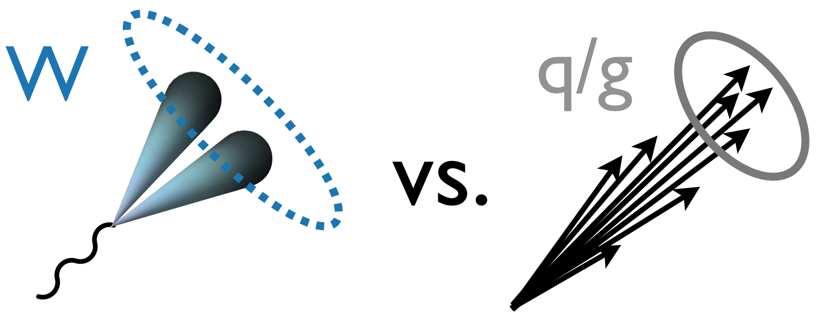
…defines a metric on the space of events

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

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

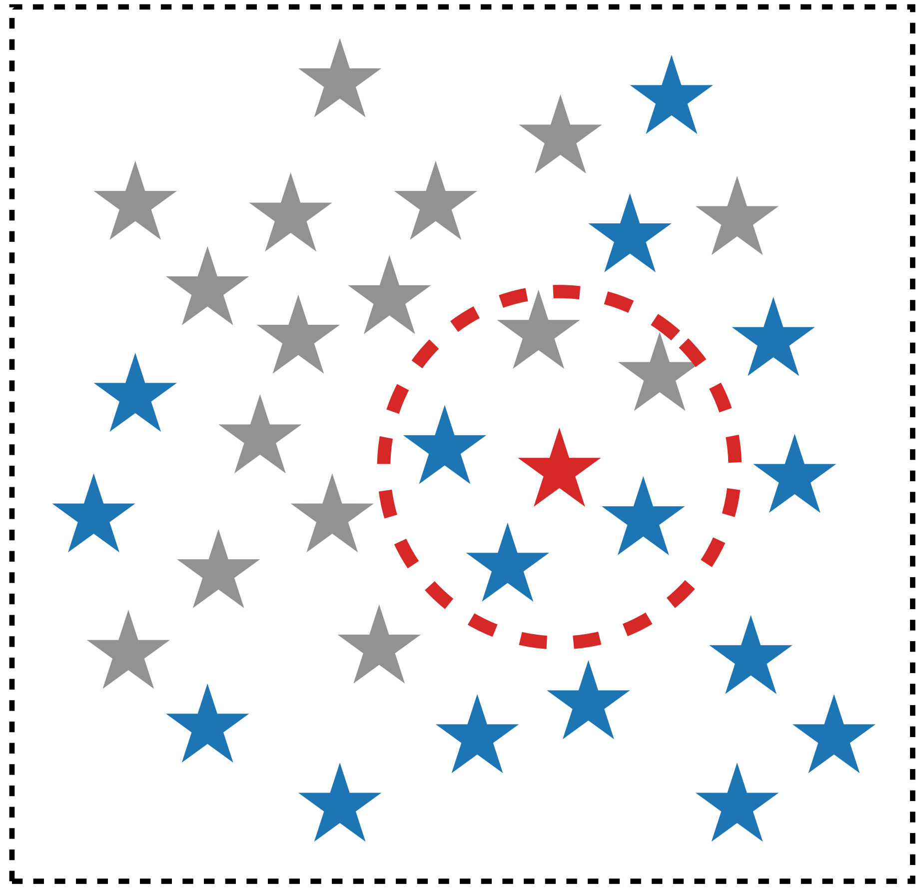
[Komiske, Metodiev, JDT, 1902.02346;
see also Peleg, Werman, Rom, IEEE 1989; Rubner, Tomasi, Guibas, ICCV 1998, ICJV 2000; Pele, Werman, ECCV 2008; Pele Taskar, GSI 2013]

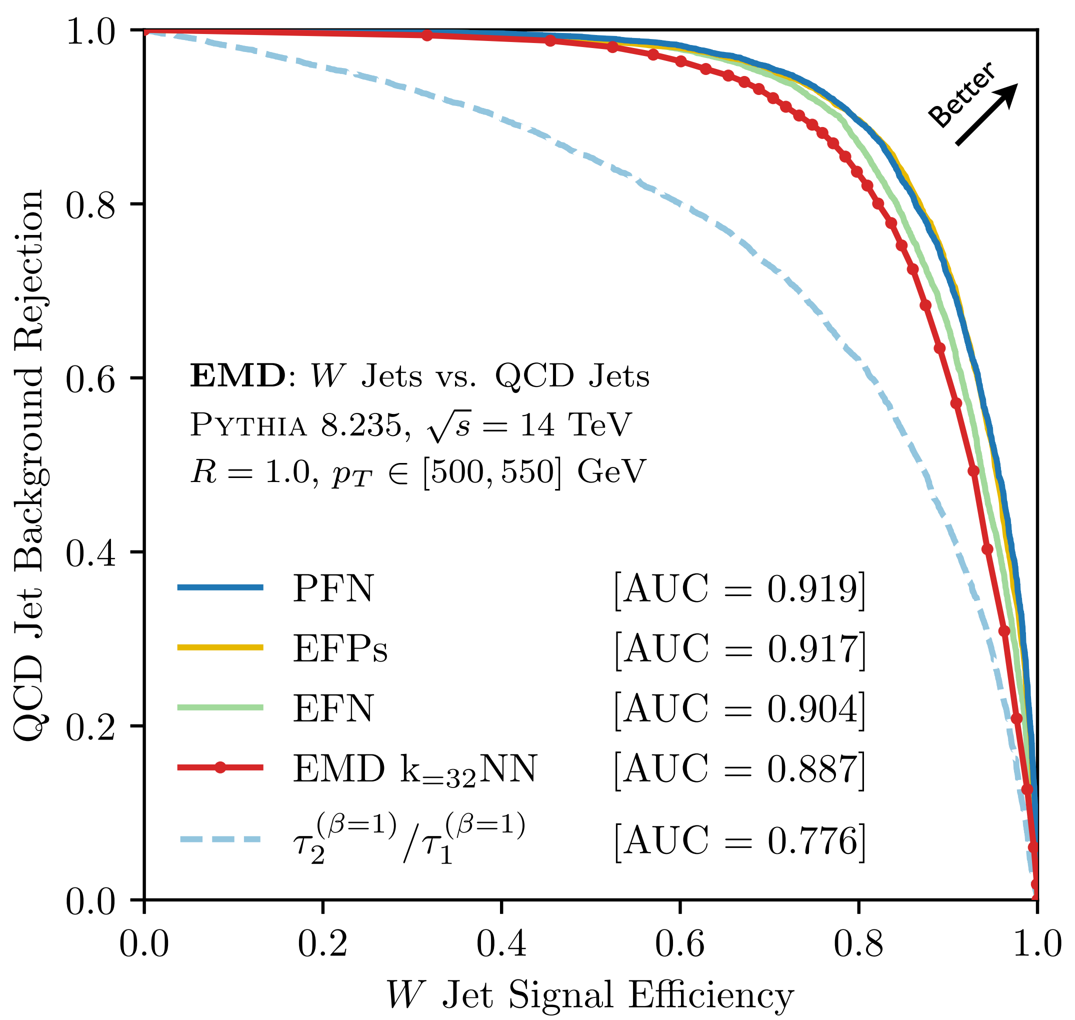
Revisiting Jet Classification



Estimate jet label by k nearest neighbors in training data

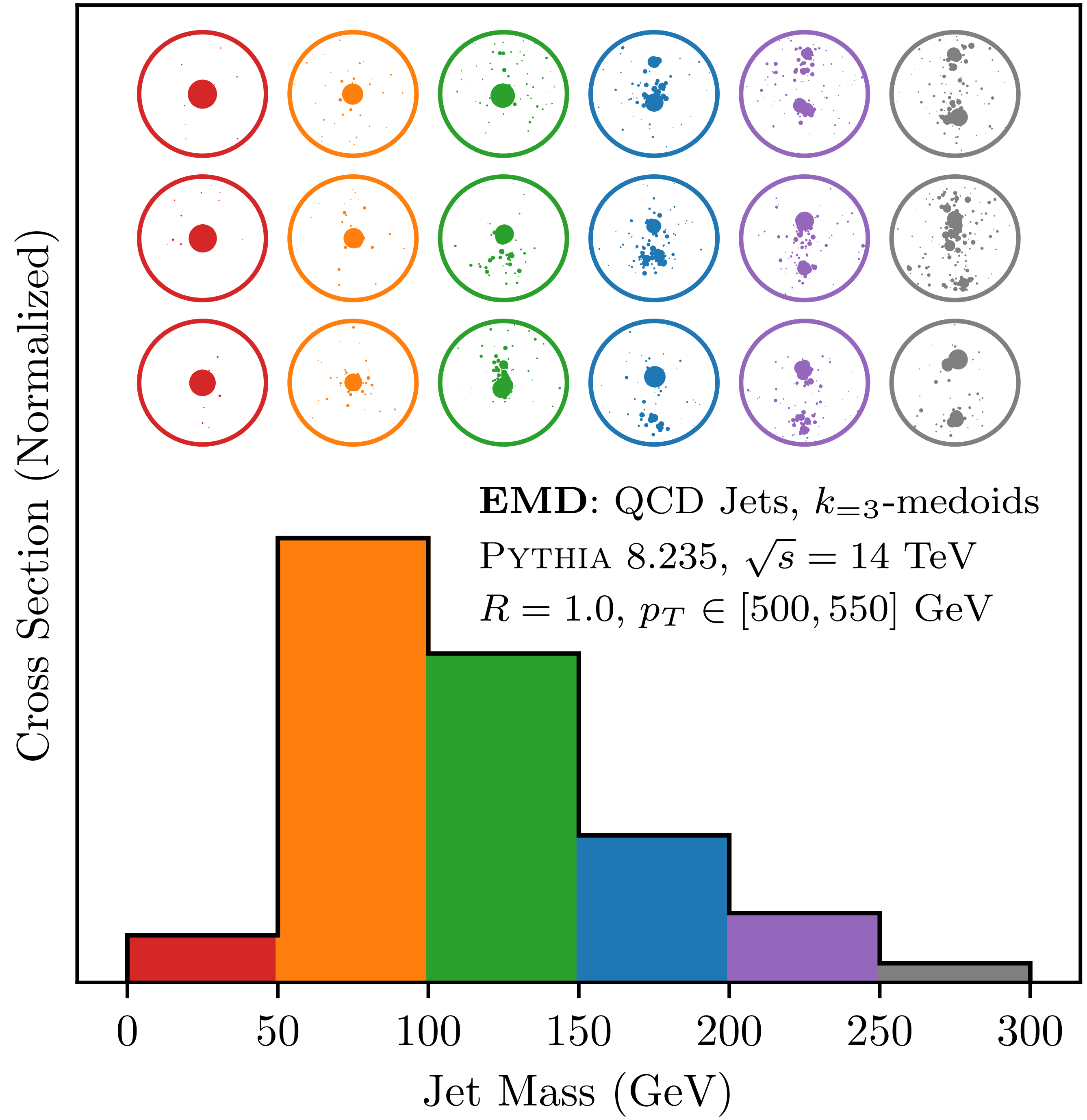
Approaches performance of modern machine learning



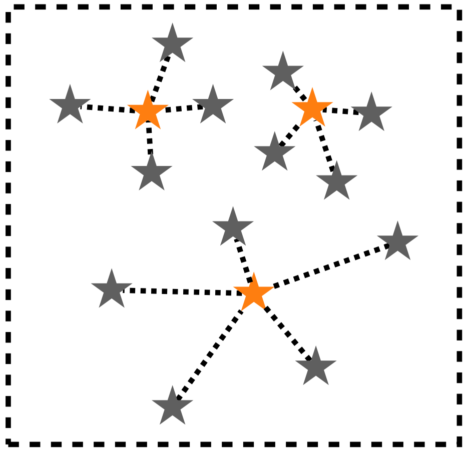


[Komiske, Metodiev, JDT, 1902.02346;
comparison to JDT, Van Tilburg, 1011.2268, 1108.2701; Komiske, Metodiev, JDT, 1712.07124, 1810.05165]

Histograms meet Event Displays

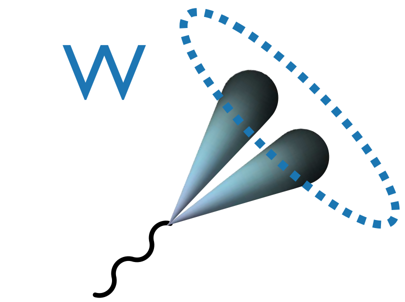


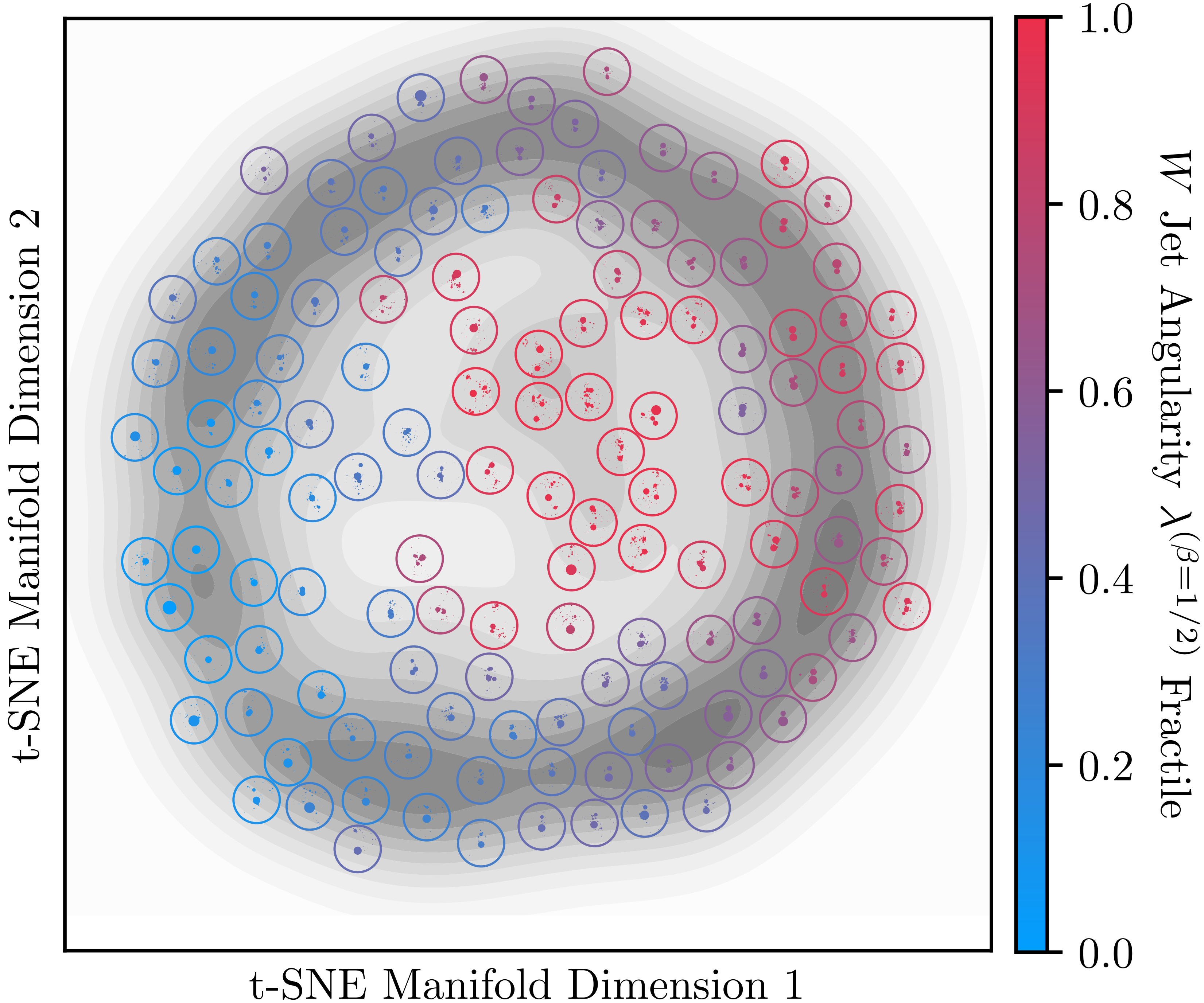
3-medoid: Three most representative jets in each bin



[Komiske, Metodiev, JDT, 1902.02346]

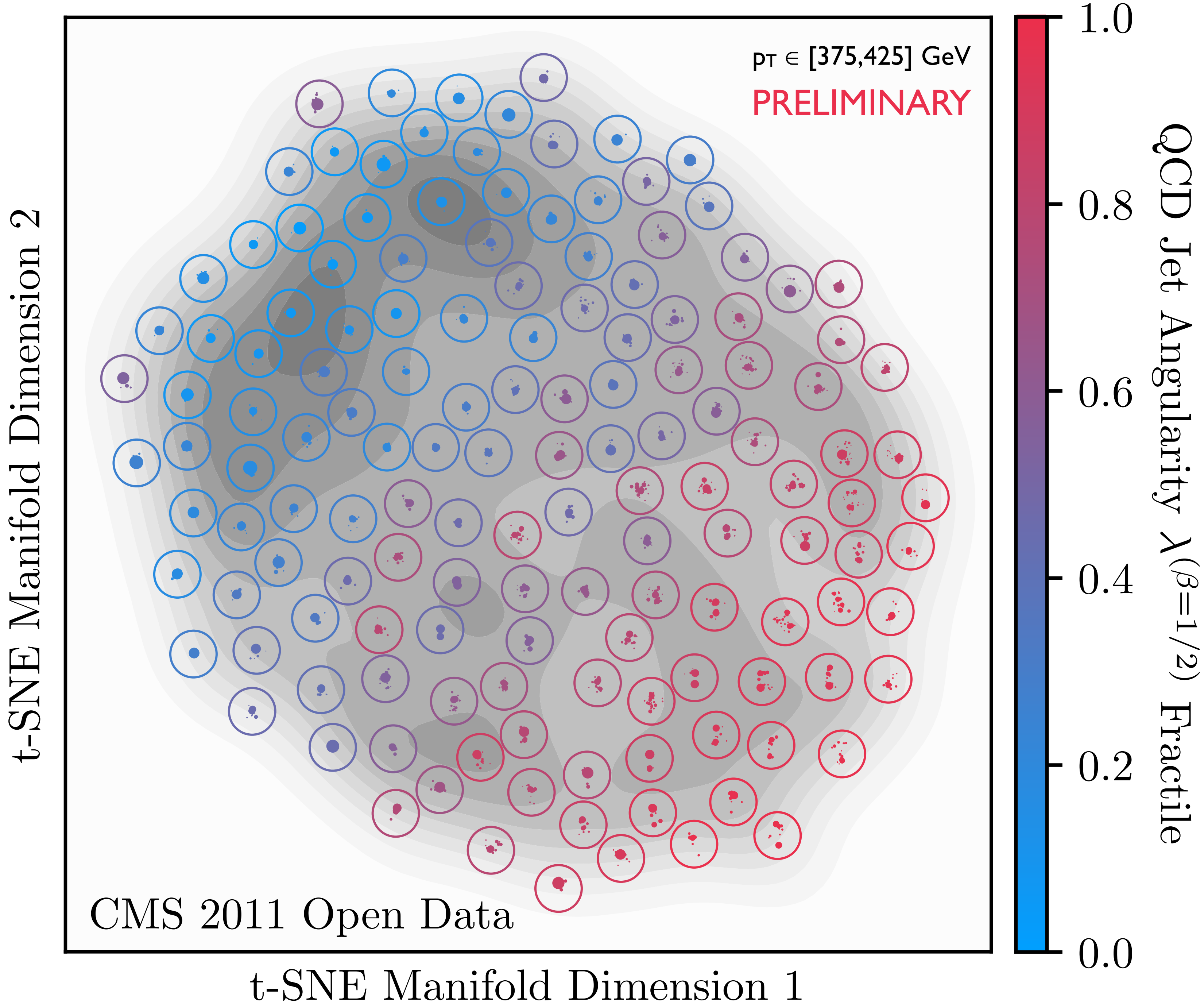
The Space of Boosted W Bosons

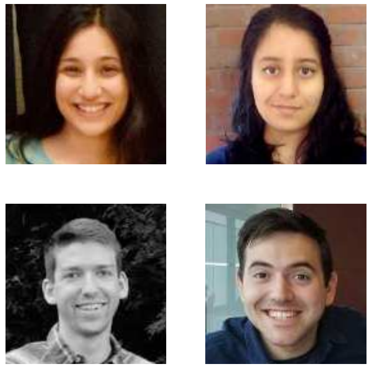




[Komiske, Metodiev, JDT, 1902.02346]

The Space of Quark/Gluon Jets





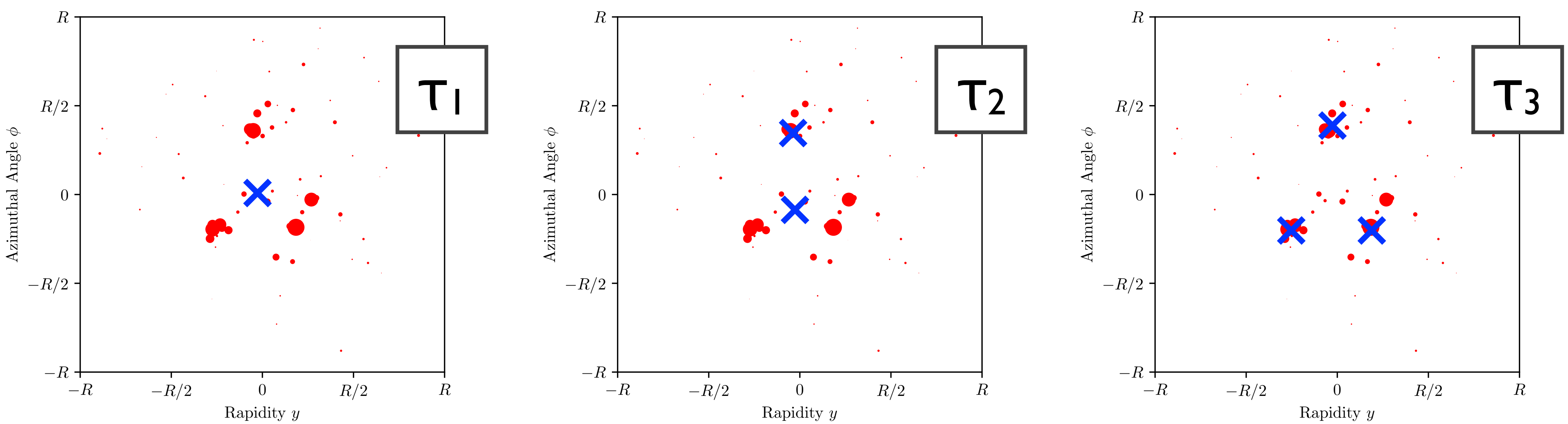
[Mastandrea, Naik, Komiske, Metodiev, JDT, in preparation]

Insight into N-subjettiness

$$\tau_N^{(\beta)}(\mathcal{E}) = \min_{N \text{ axes}} \sum_i E_i \min\left\{\theta_{1,i}^{\beta}, \theta_{2,i}^{\beta}, \dots, \theta_{N,i}^{\beta}\right\}$$

IRC safe

kind of arbitrary



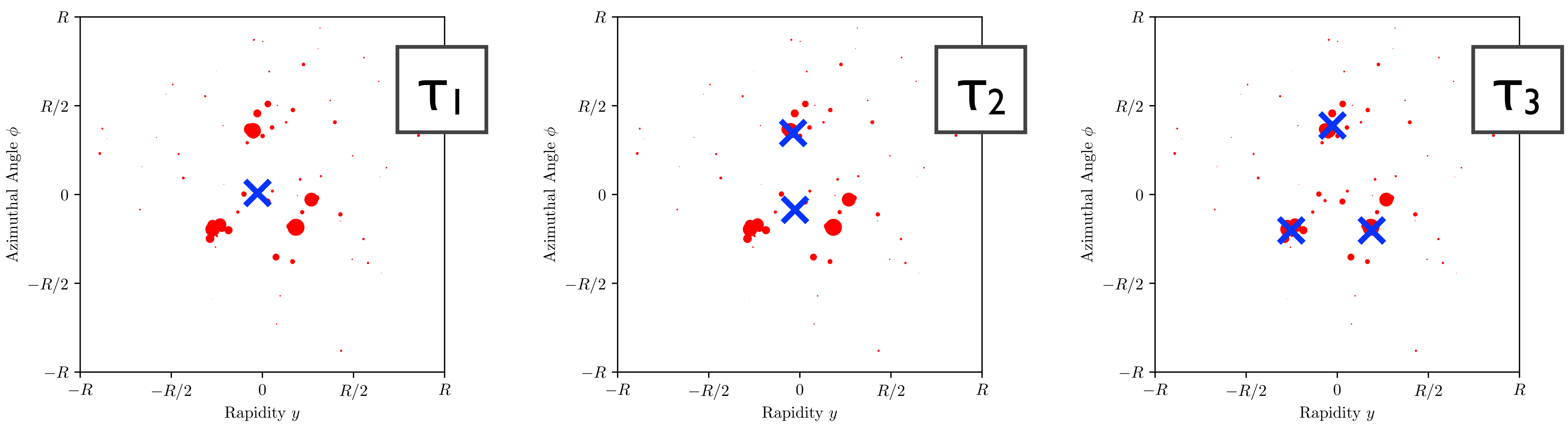
[JDT, Van Tilburg, 1011.2268, 1108.2701;
based on Brandt, Dahmen, ZPC 1979; Stewart, Tackmann, Waalewijn, 1004.2489]

Insight into N-subjettiness

$$\tau_N^{(\beta)}(\mathcal{E}) = \min_{N \text{ axes}} \sum_i E_i \min\left\{\theta_{1,i}^\beta, \theta_{2,i}^\beta, \ldots, \theta_{N,i}^\beta\right\}$$

IRC safe

kind of arbitrary



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

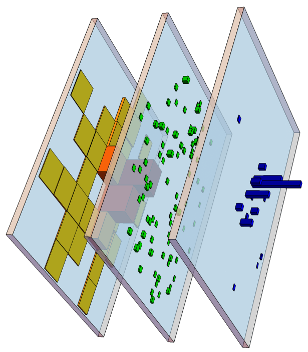
for β = 1

very satisfying

Related to p-Wasserstein metric for p = β > 1

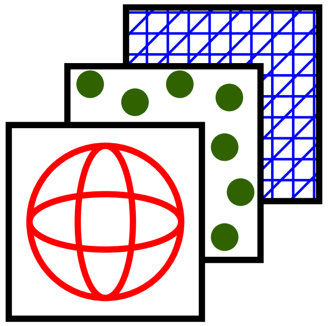
[JDT, Van Tilburg, 1011.2268, 1108.2701; based on Brandt, Dahmen, ZPC 1979; Stewart, Tackmann, Waalewijn, 1004.2489]

Summary



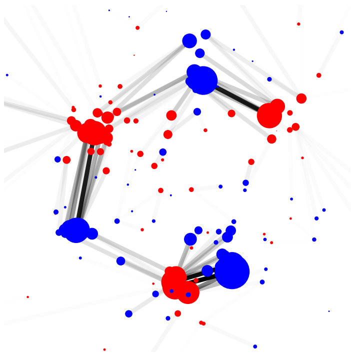
The Rise of Deep Learning

Leveraging computational power to solve well-posed problems



Looking Inside the Black Box

Designing network architectures around symmetries and interpretability



(Exploring the Space of Jets)

Computational geometry as a new collider data analysis strategy

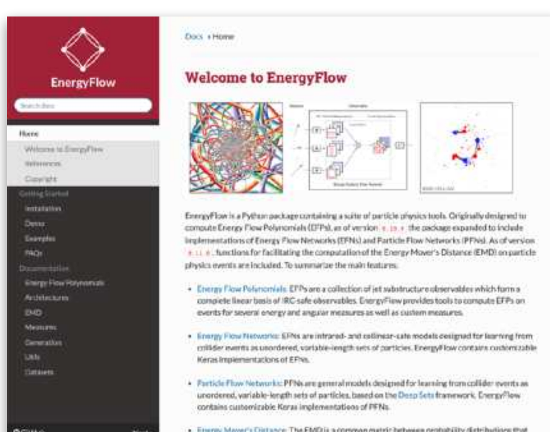
Patrick Komiske

Eric Metodiev

(Theoretical) High Energy Physics

?

Mathematics, Statistics, Computer Science



Energy Flow Package

https://energyflow.network/

Backup Slides

"Collision Course"

"Theoretical Physics for Machine Learning"
Aspen Center for Physics, January 2019

(Theoretical)
High Energy
Physics

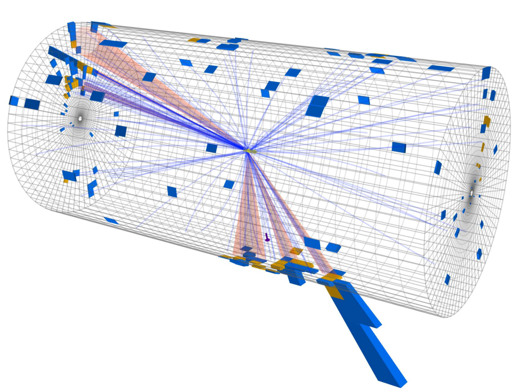
?

Mathematics,
Statistics,
Computer Science

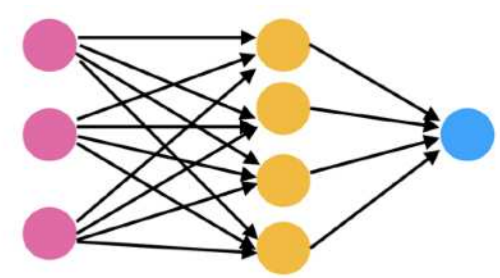
This talk

ML4HEP

Solving complex problem with neural networks

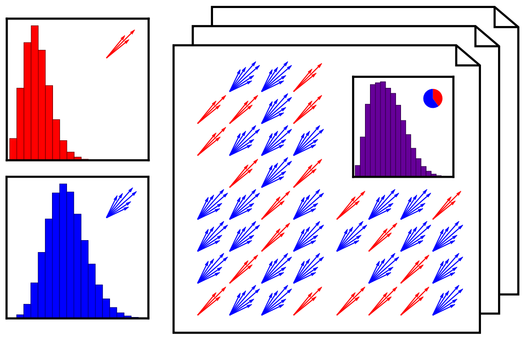


HEP4ML



Studying neural networks like physical systems

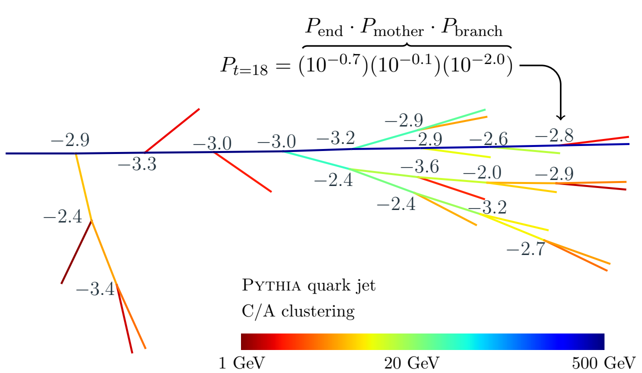
The Rise of Unsupervised Learning



Jet Topics

Blind Source Separation

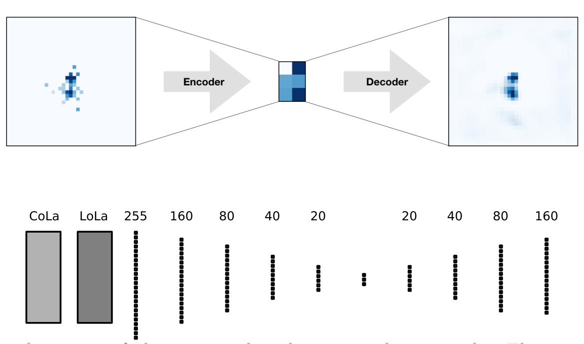
[Metodiev, JDT, 1802.00008; Komiske, Metodiev, JDT, 1809.01140;
see also Metodiev, Nachman, JDT, 1708.02949]



JUNIPR

Probability Modeling

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



Autoencoders

Anomaly Detection

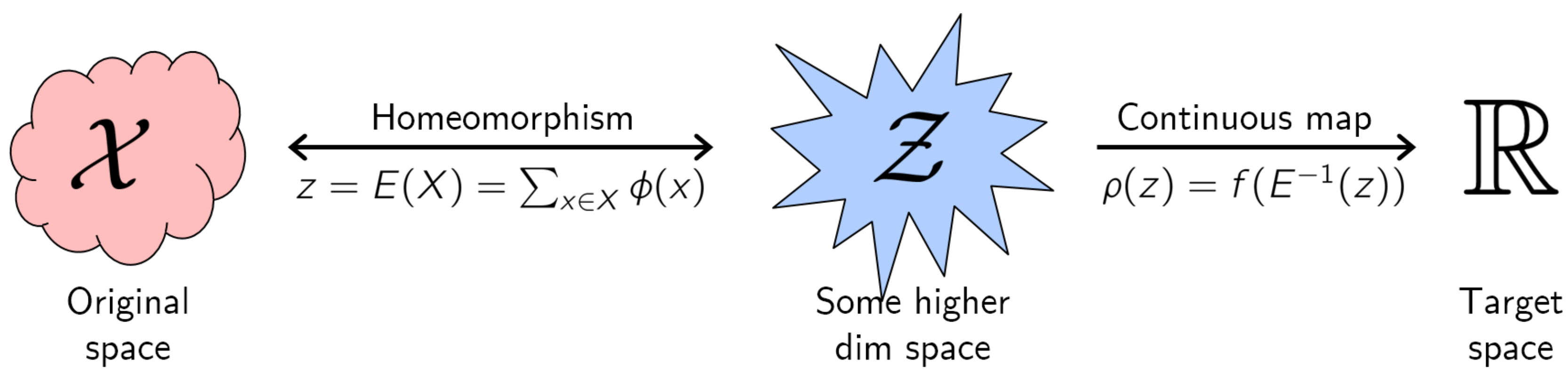
[Hajer, Li, Liu, Wang, 1807.10261; Heimel, Kasieczka, Plehn, Thompson, 1808.08979;
Farina, Nakai, Shih, 1808.08992; Cerri, Nguyen, Pierini, Spiropulu, Vlimant, 1811.10276;
see also Collins, Howe, Nachman, 1805.02664, 1902.02634; De Simone, Jacques, 1807.06038]

Common theme: Analyze event ensembles, not individual events

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

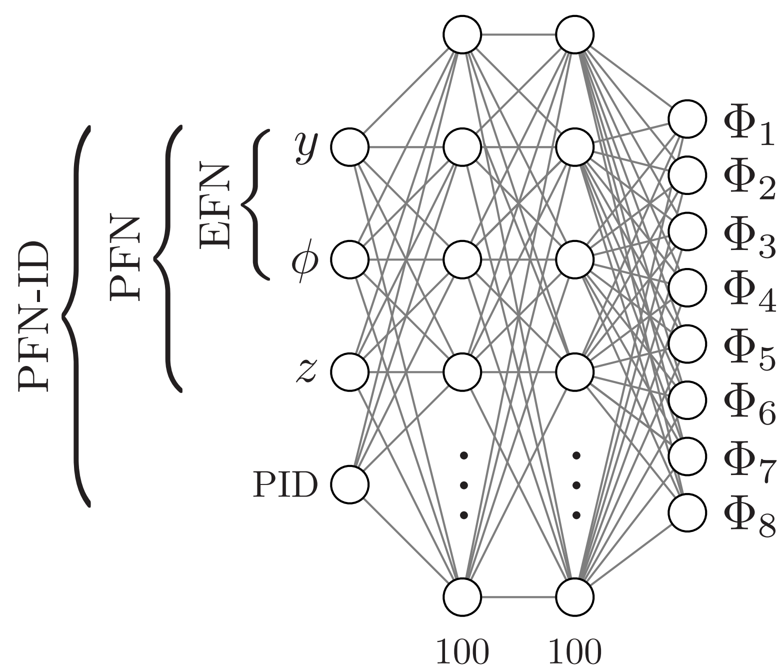
Additive Latent Space

Generic Observable

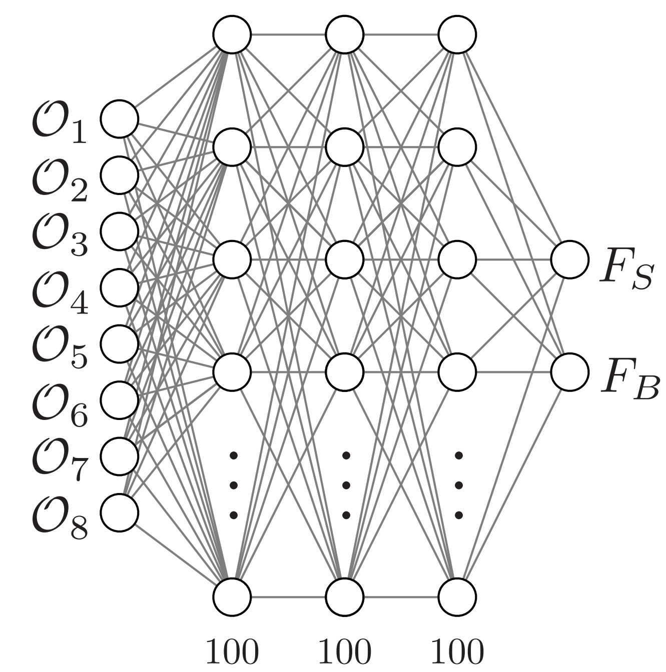
[Zaheer, Kottur, Ravanbakhsh, Poczos, Salakhutdinov, Smola, 1703.06114; see also Vinyals, Bengio, Kudlur, 1511.06391; Qi, Su, Mo, Guibas, 1612.00593]

Technical Implementation

Per-Particle Network: Φ



Latent Combiner: F



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

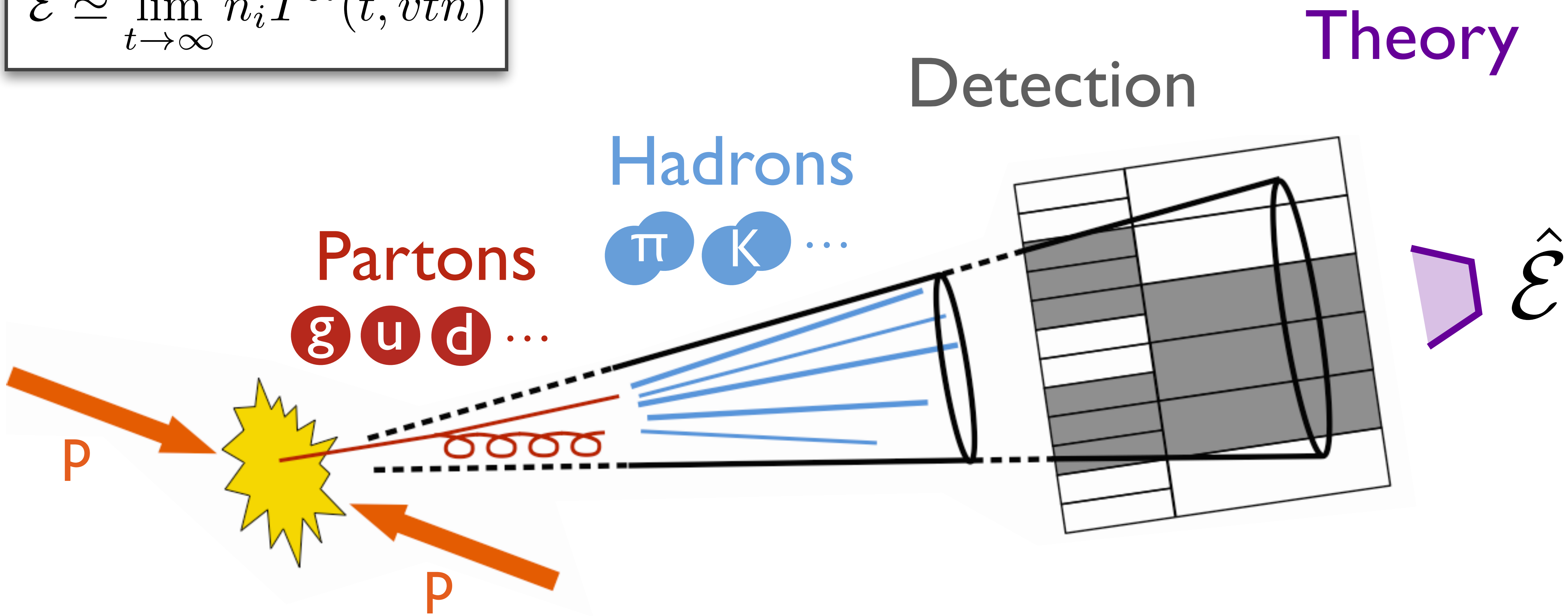
Final Discriminant:

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

[Komiske, Metodiev, JDT, 1810.05165]

Focus on Energy Flow

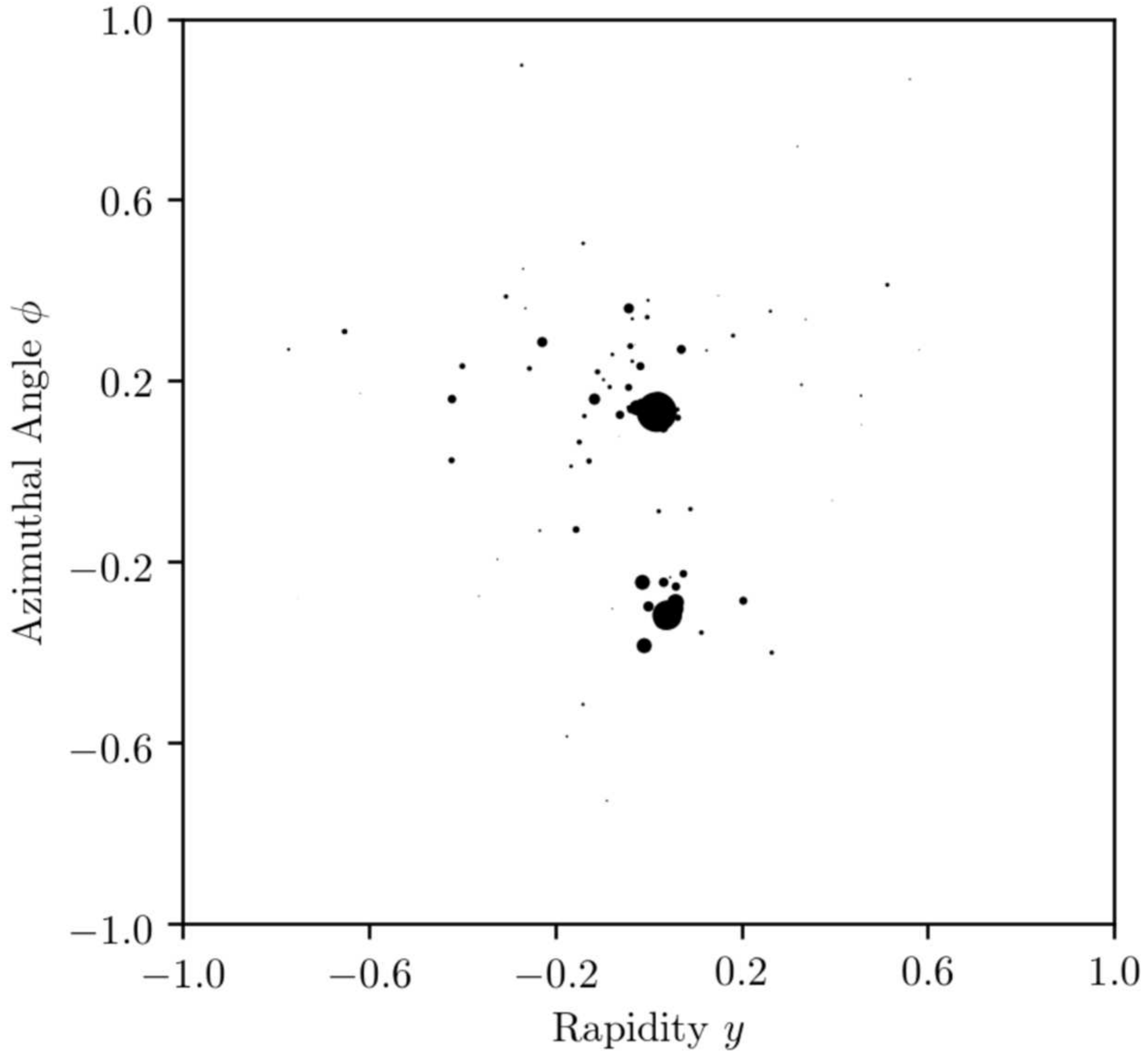
$$\boxed{\hat{\mathcal{E}} \simeq \lim_{t\to\infty} \hat{n}_i T^{0i}(t, vt\hat{n})}$$



Stress-energy flow: *Measure of event/jet structure robust to non-perturbative and detector effects (i.e. IRC safe)*

[Sveshnikov, Tkachov, hep-ph/9512370; Hofman, Maldacena, 0803.1467; Mateu, Stewart, JDT, 1209.3781; Komiske, Metodiev, JDT, 1712.07124, 1810.05165]

Focus on Energy Flow



Represent jet as:

$$\rho(\hat{p}) = \sum_{i \in \text{jet}} E_i \, \delta(\hat{p} - \hat{p}_i)$$

Energy (p_T)

Direction (y,φ)

Safe to infrared & collinear splittings
No flavor/charge information
No pixelation needed

Stress-energy flow: Measure of event/jet structure robust to non-perturbative and detector effects (i.e. IRC safe)

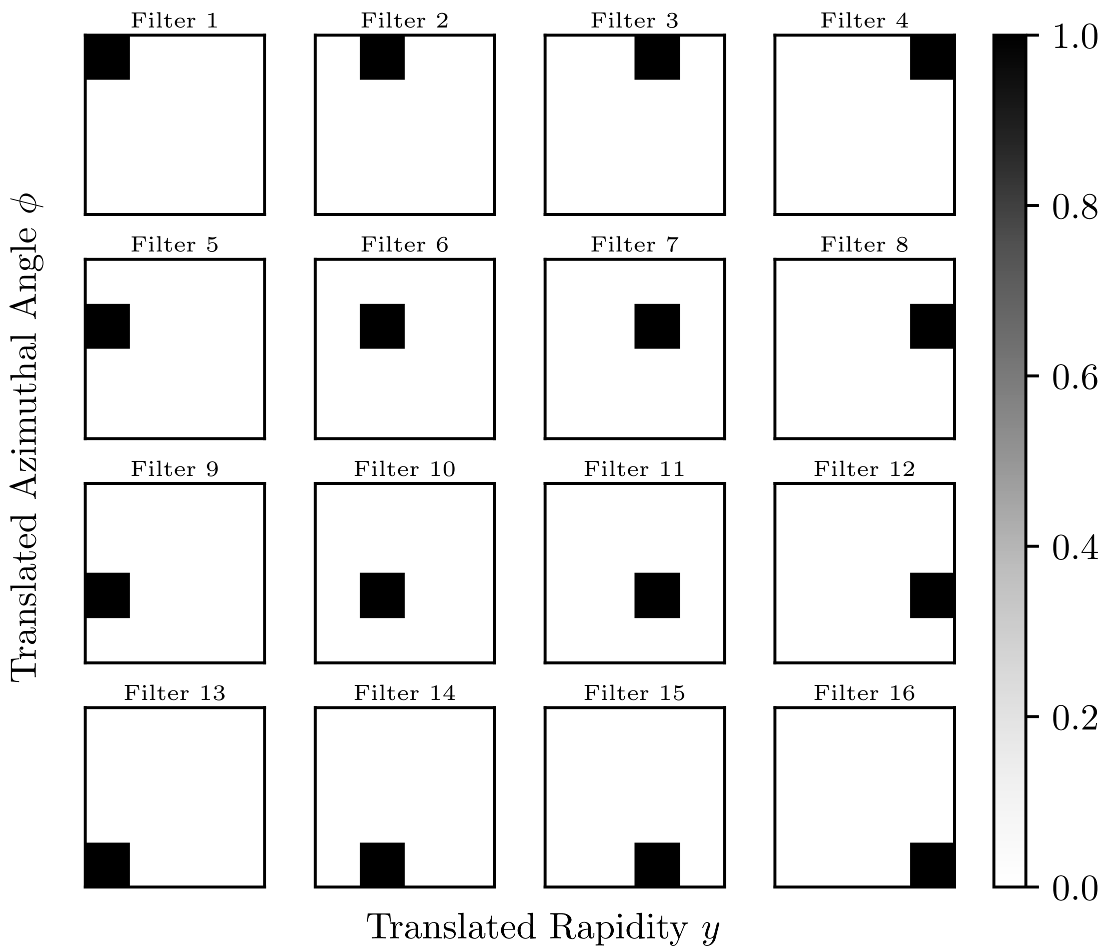
[Sveshnikov, Tkachov, hep-ph/9512370; Hofman, Maldacena, 0803.1467; Mateu, Stewart, JDT, 1209.3781; Komiske, Metodiev, JDT, 1712.07124, 1810.05165]

Latent Space Visualization

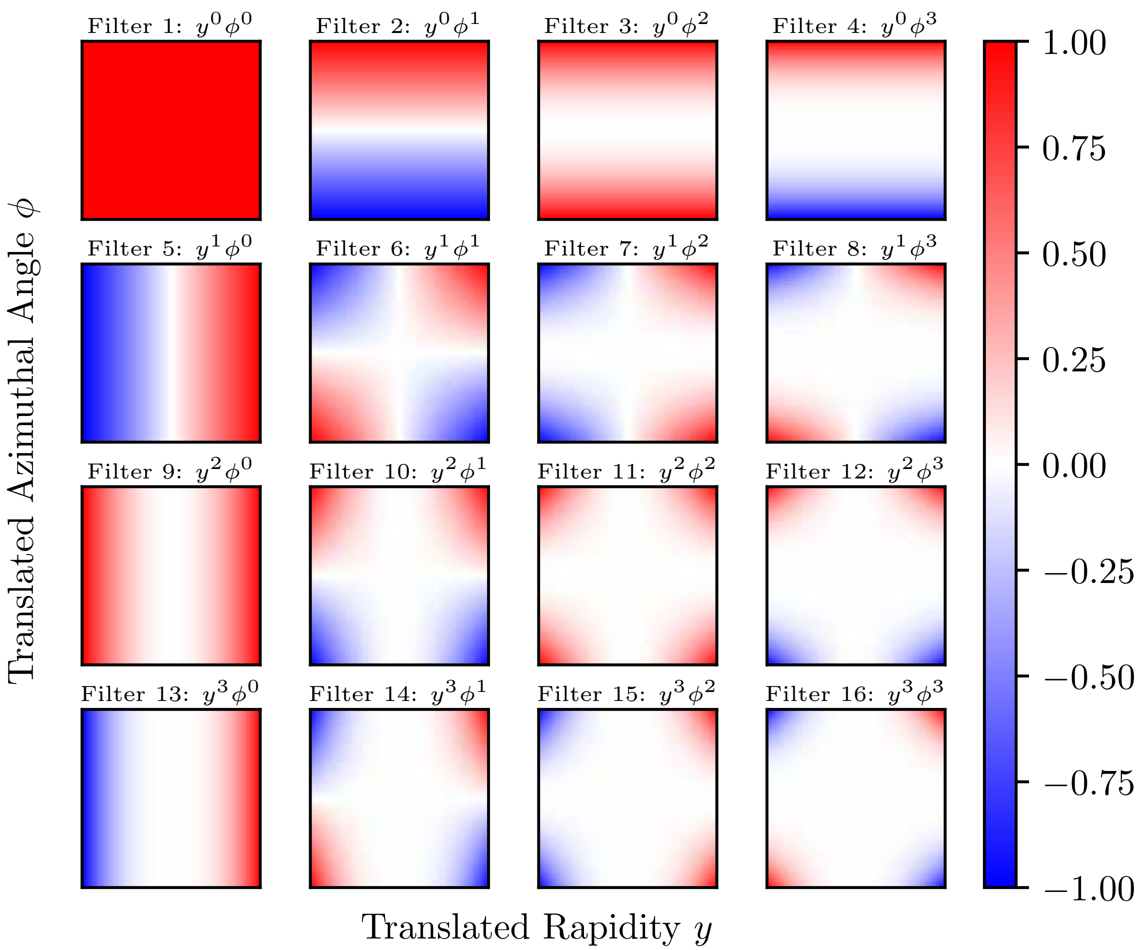
$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_\ell)$$

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

Calorimeter Pixels



Radiation Moments

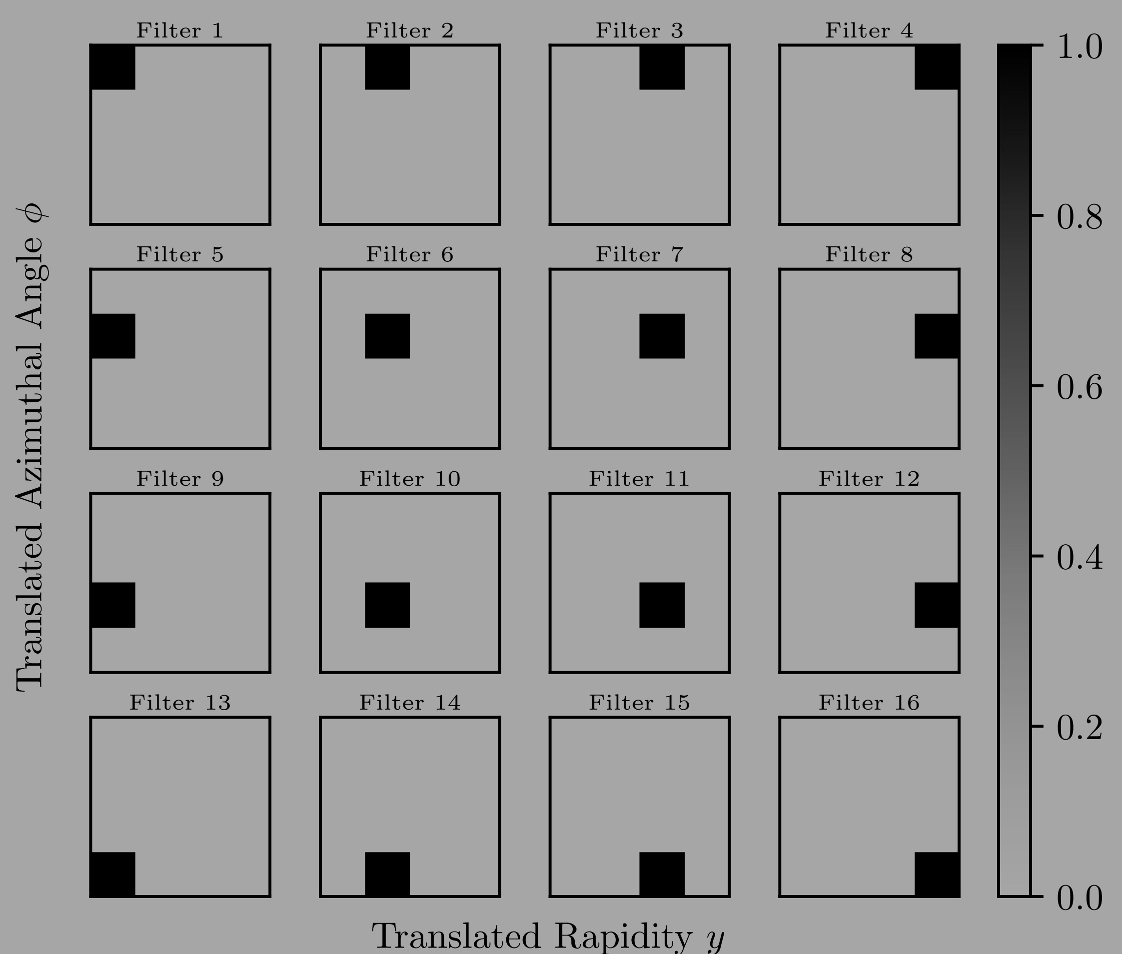


Latent Space Visualization

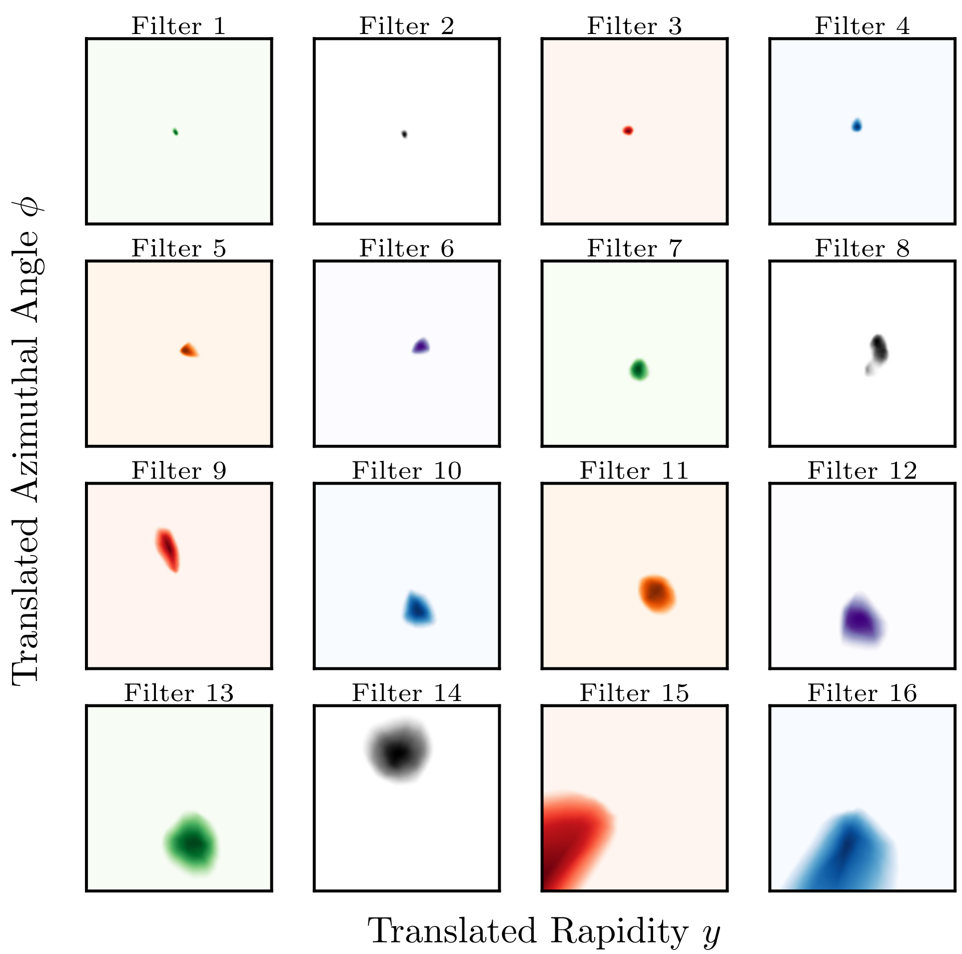
$$\mathcal{O}(\{p_i^\mu\}) = F(\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_\ell)$$

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

Calorimeter Pixels

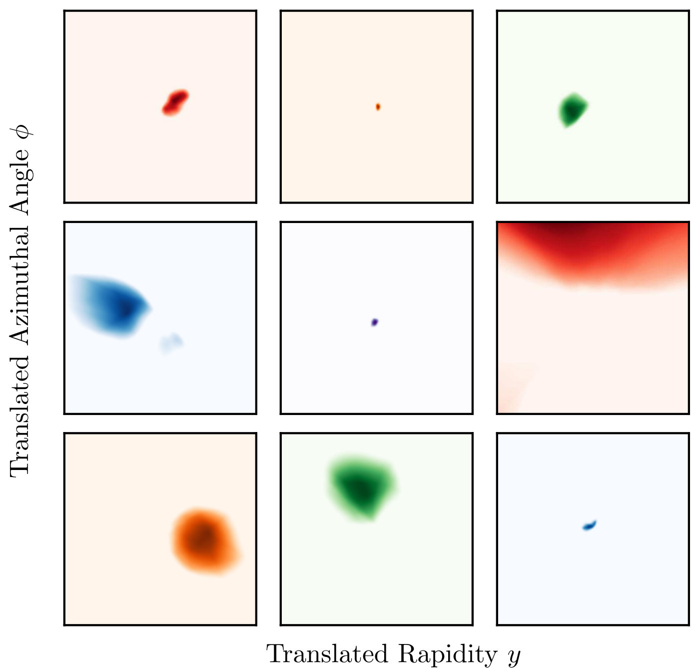


EFNs: Dynamic Pixelation

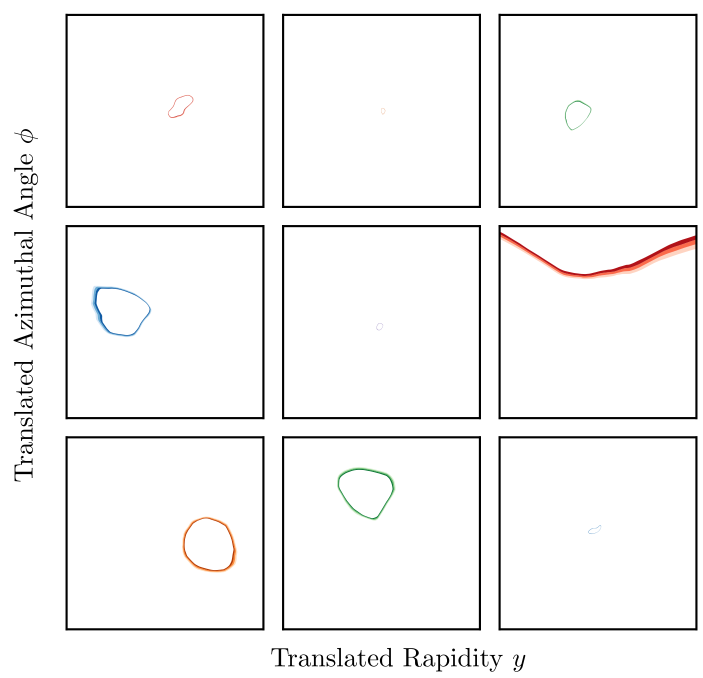


Psychedelic Network Visualization

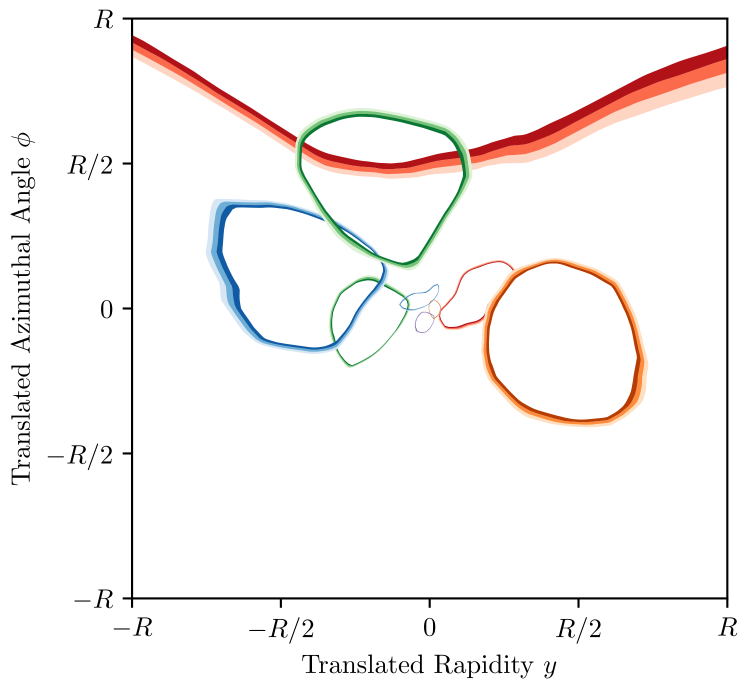
Latent Filters



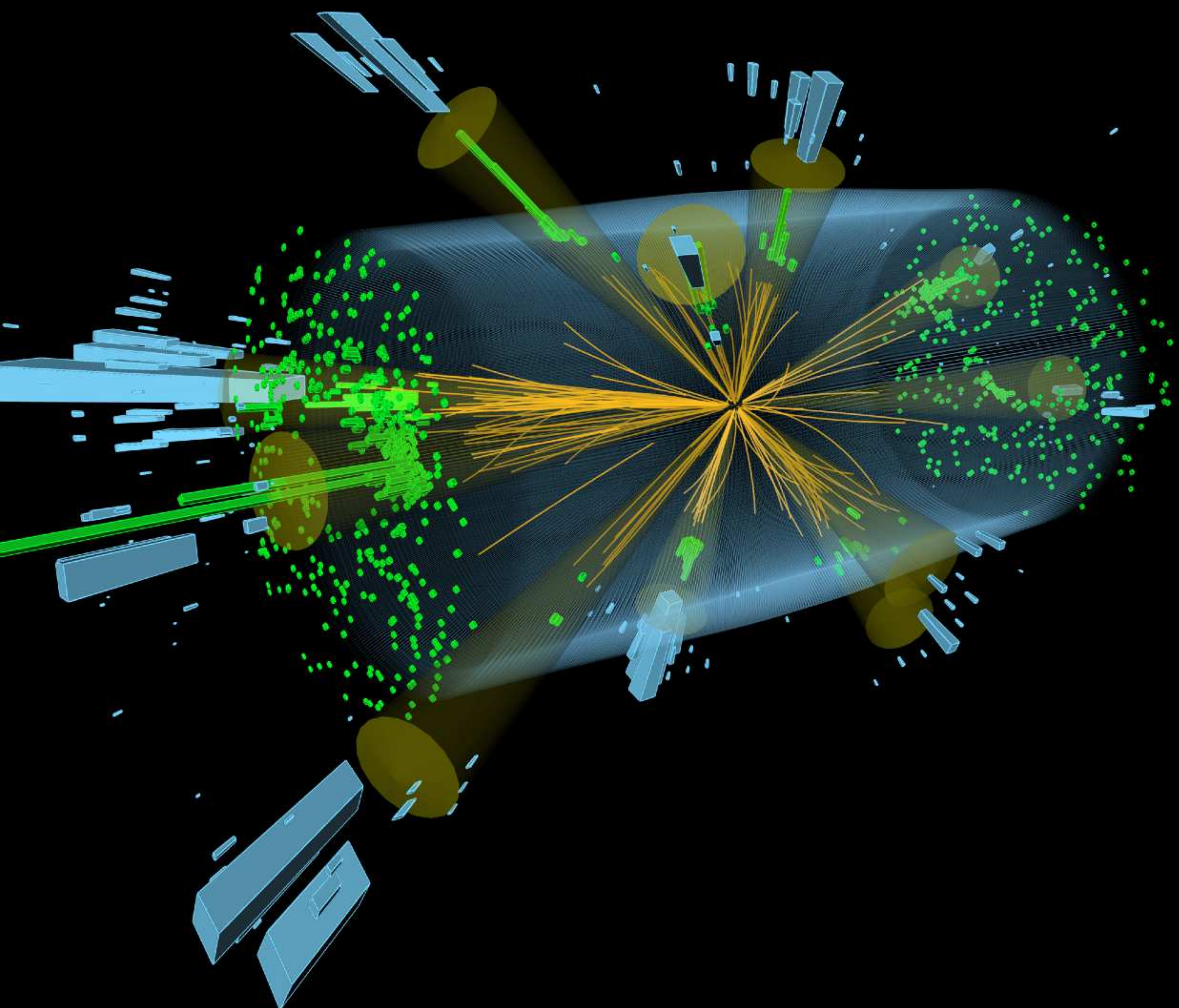
50% Contours



Overlay



What is a Collision Event?



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

Point Cloud

[Popular Science, 2013]