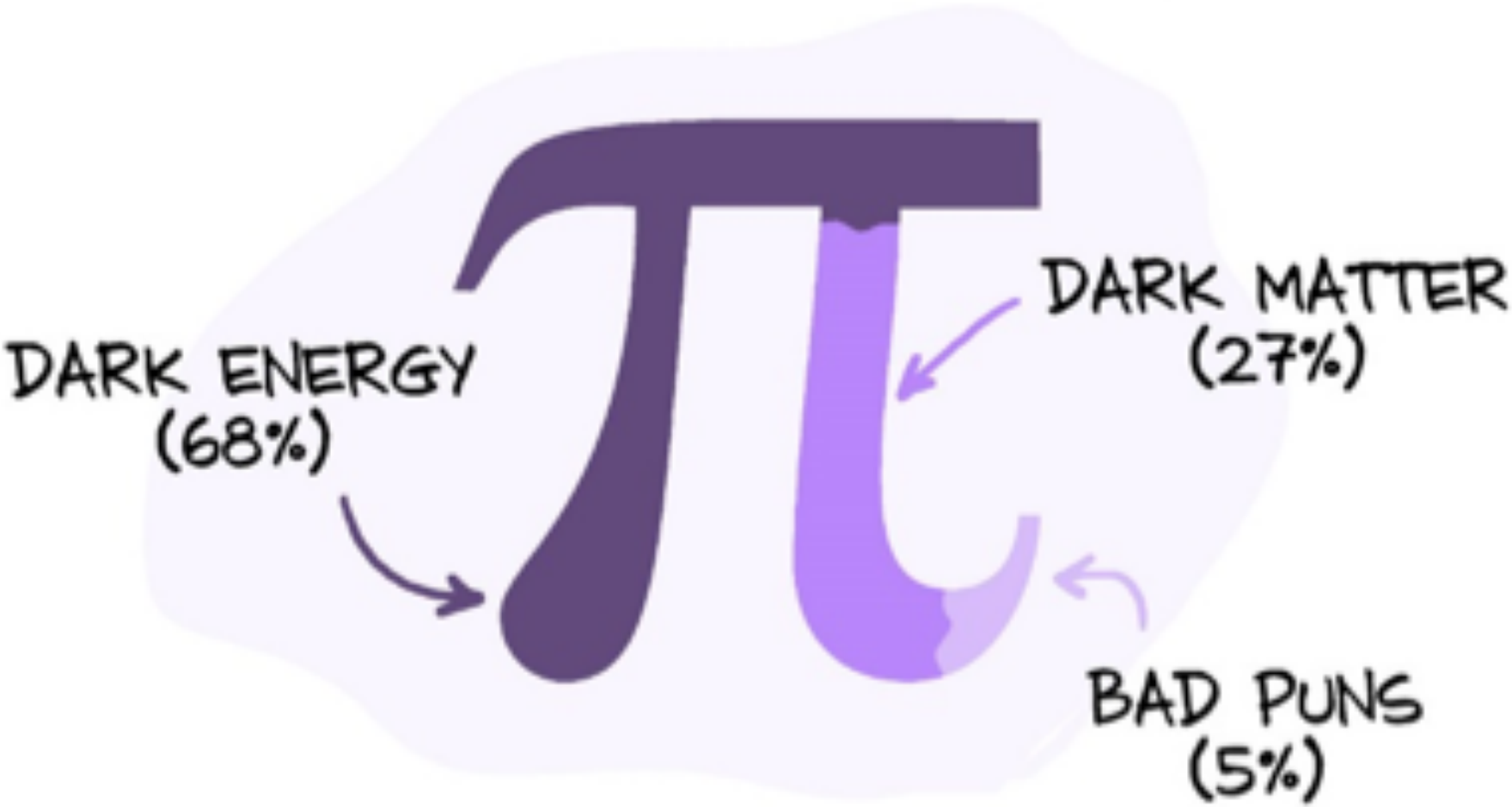
Learning Physics From Machines

Daniel Whiteson, UC Irvine

Aug 2020

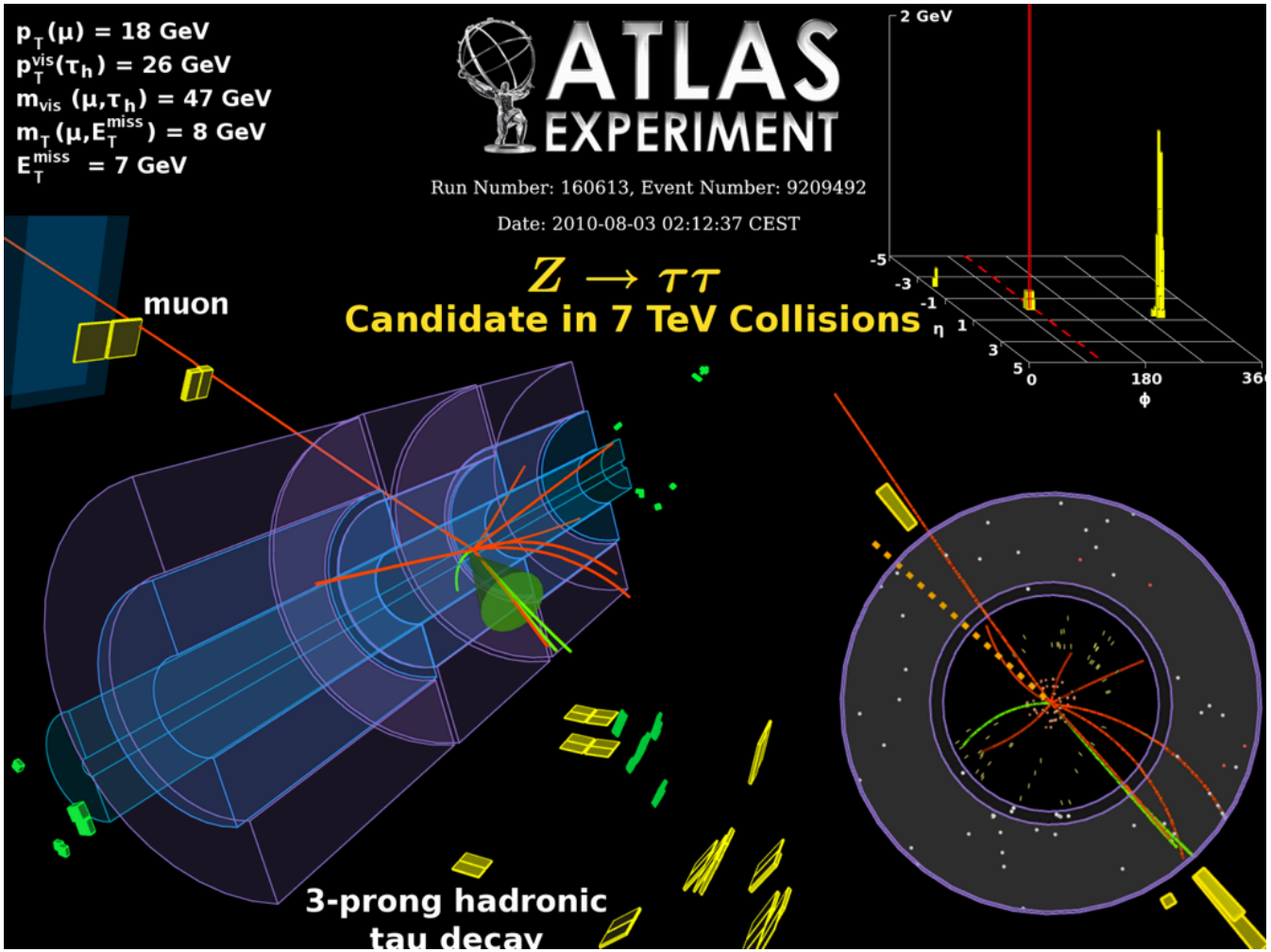
What is Deep Learning?

THE UNIVERSE: A PI CHART



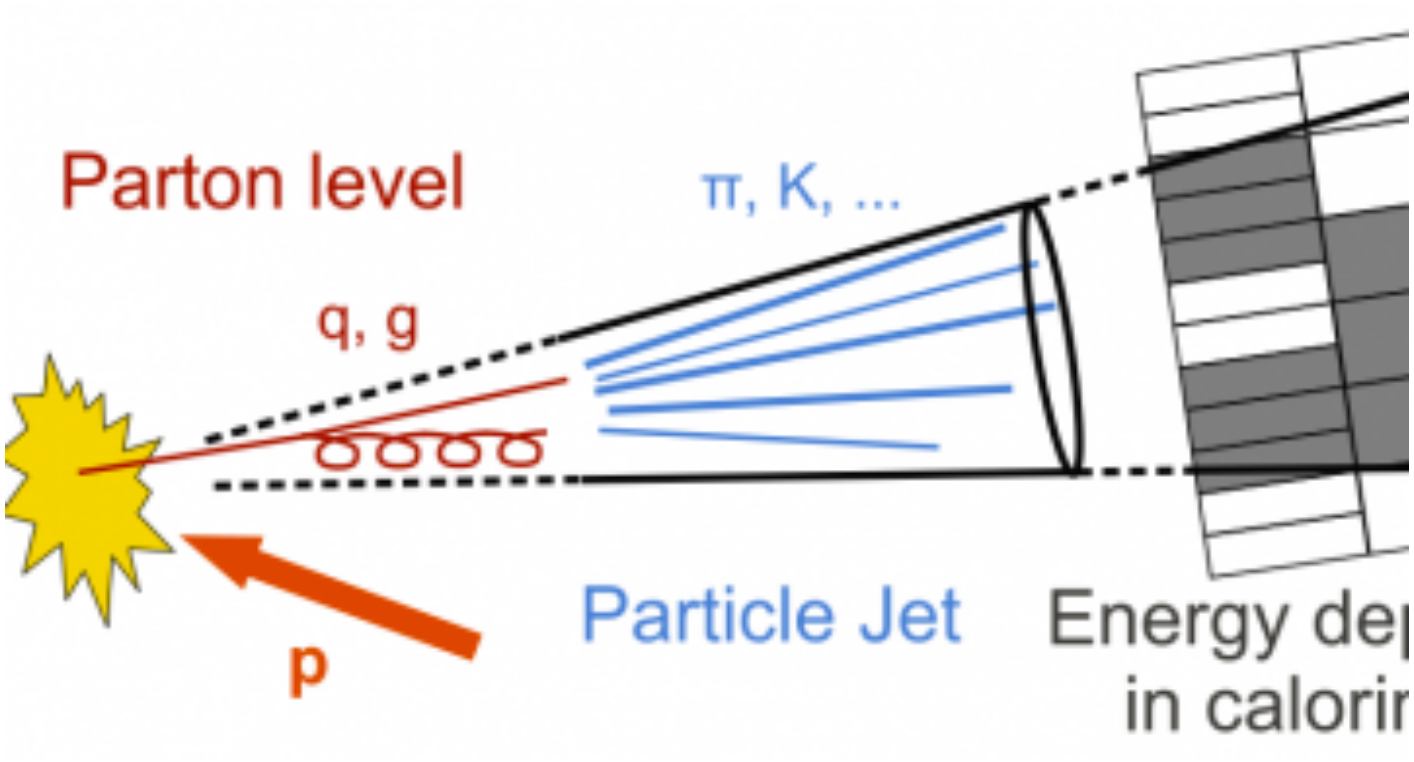
Why deep learning?

A lot of information hiding in our data



The nature of our data demands it.

Jets



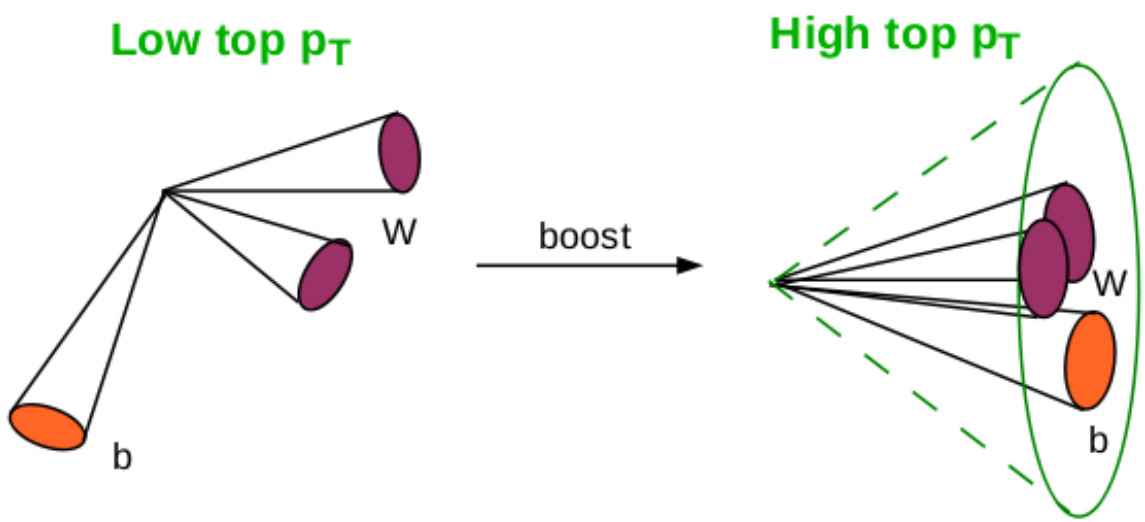
Benchmark case

Jet Substructure Classification in High-Energy Physics with Deep Neural Networks

Pierre Baldi,[1] Kevin Bauer,[2] Clara Eng,[3] Peter Sadowski,[1] and Daniel Whiteson[2]

[1] *Department of Computer Science, University of California, Irvine, CA 92697*
[2] *Department of Physics and Astronomy, University of California, Irvine, CA 92697*
[3] *Department of Chemical and Biomolecular Engineering, University of California, Berkeley CA 94270*

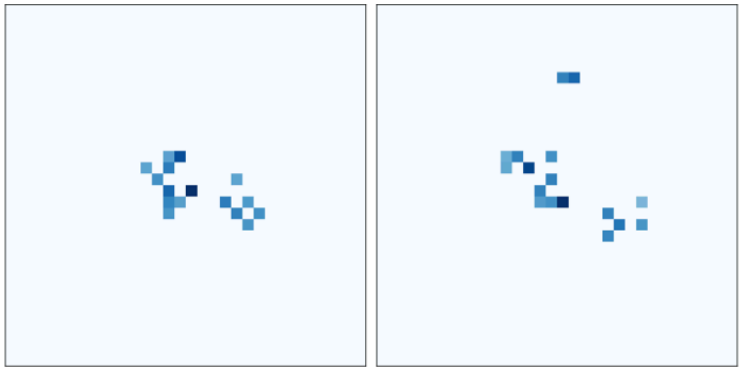
(Dated: April 12, 2016)



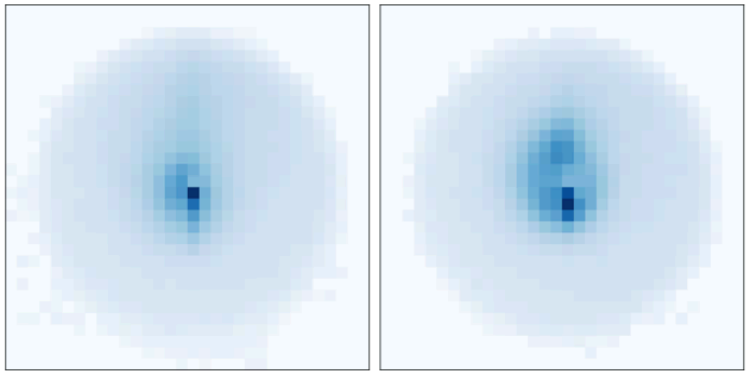
LL data - calc images

quark | W > q q

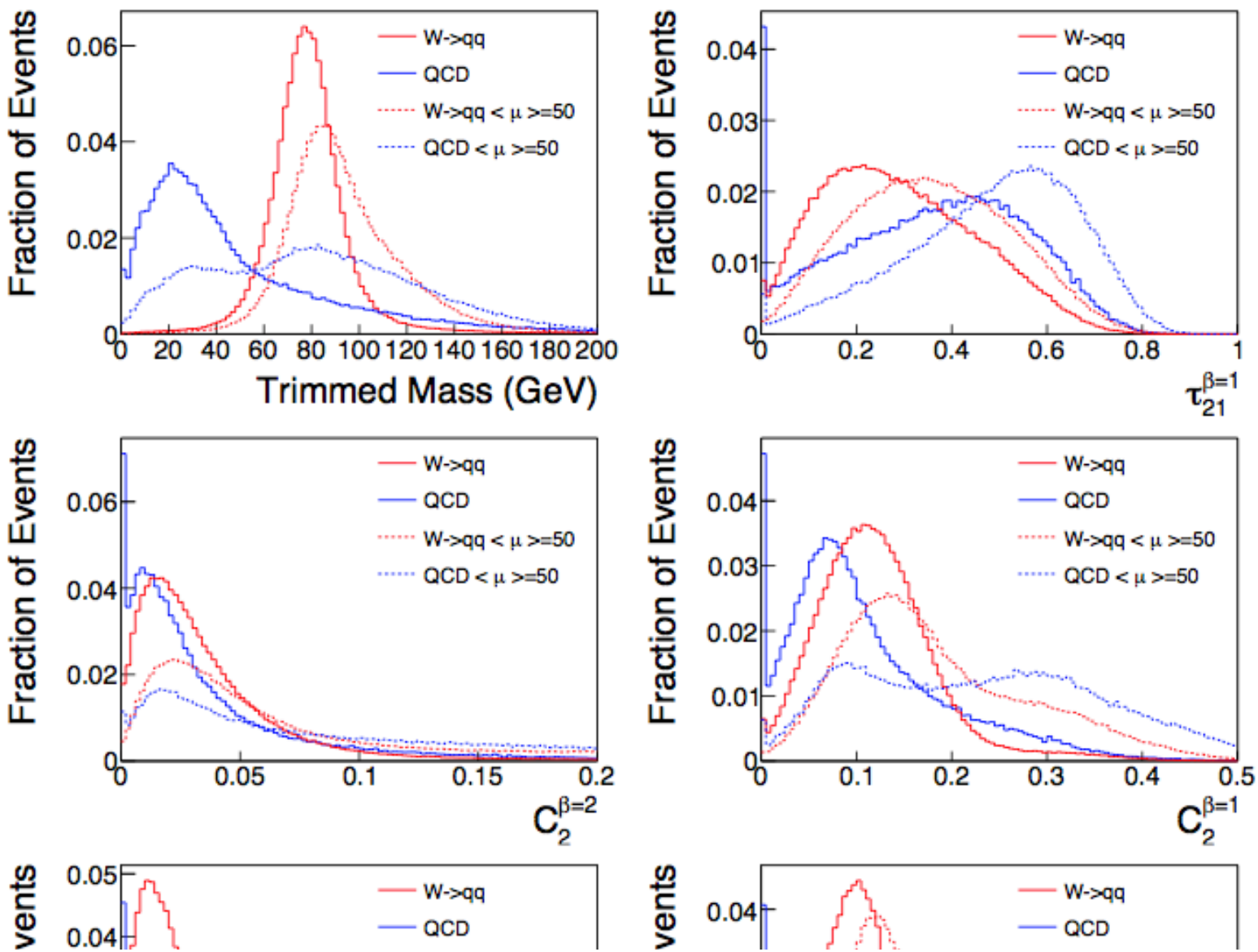
Single example



Average example



High-level data



Jet tagging

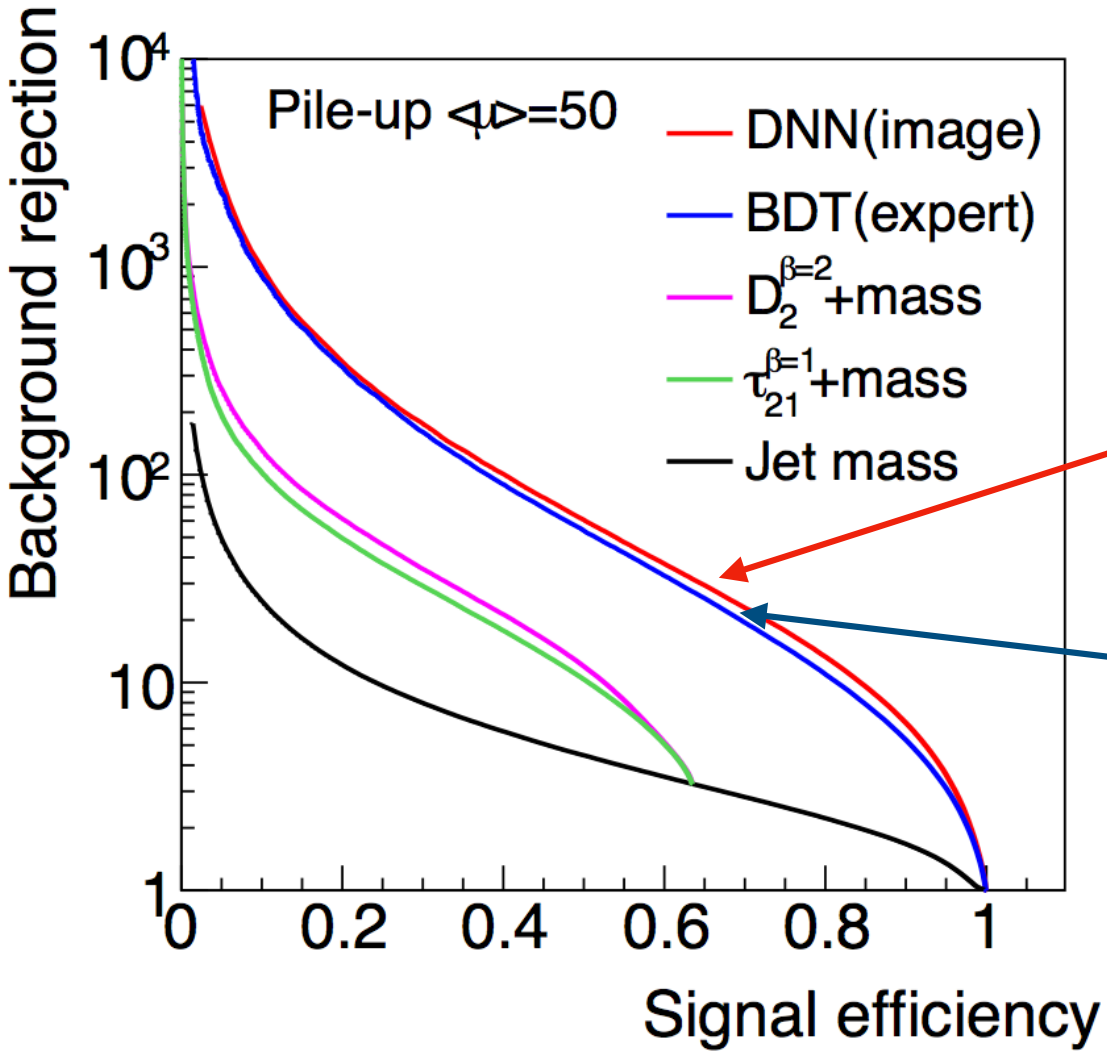


Image network

6HL features

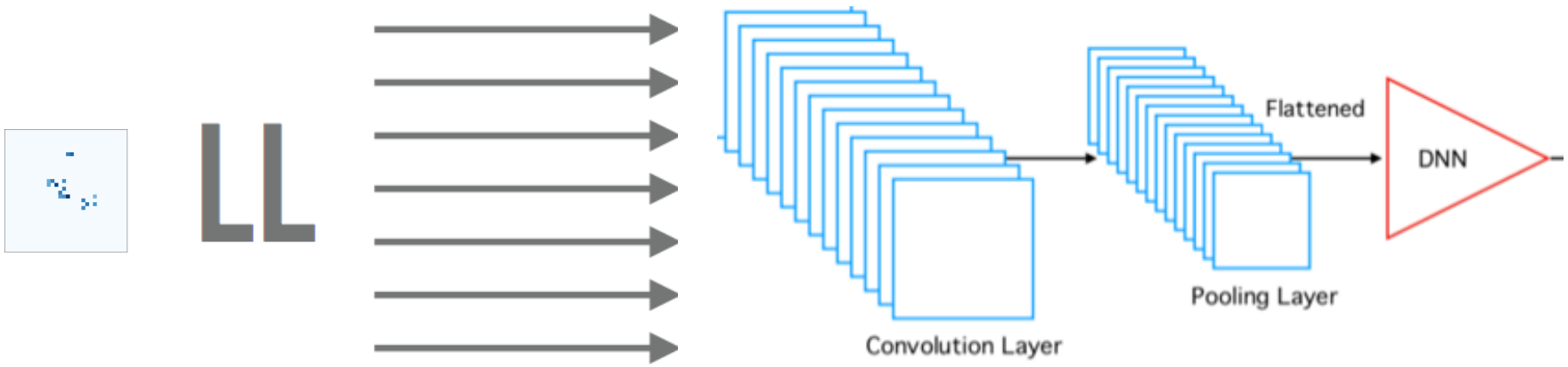
Jet tagging

What has the machine learned?

Did it rediscover the HL features?
Find a totally different strategy?

What is ML doing?

Our low-level (LL) data are often high-dim

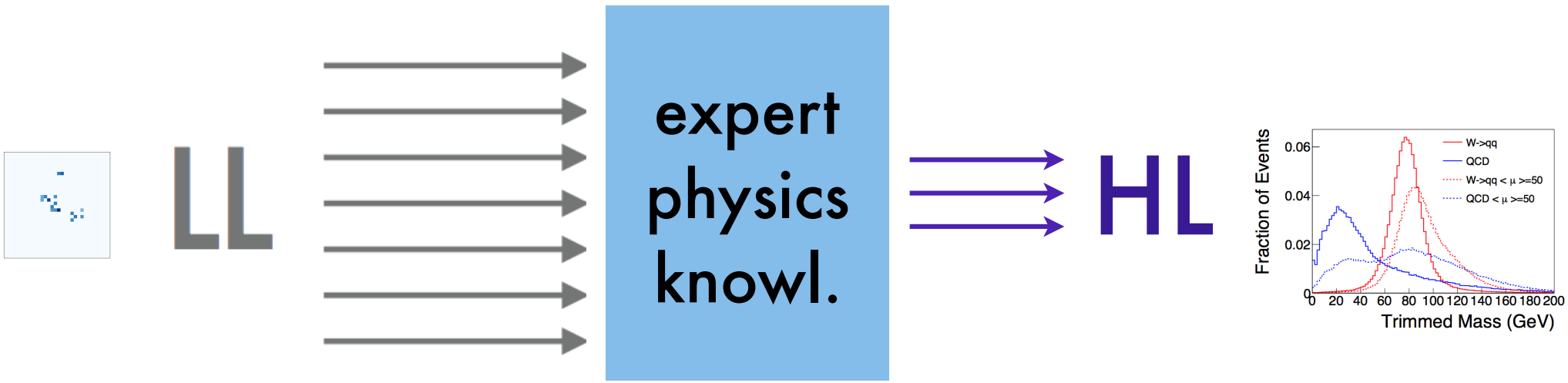


<u>Problem:</u>

- Too many dimensions to verify modeling
- Little physical intuition for low-level data

What is ML doing?

We often reduce data dimensionality



<u>Problem:</u>

- Deep learning says we are losing information

Yet we prefer HL

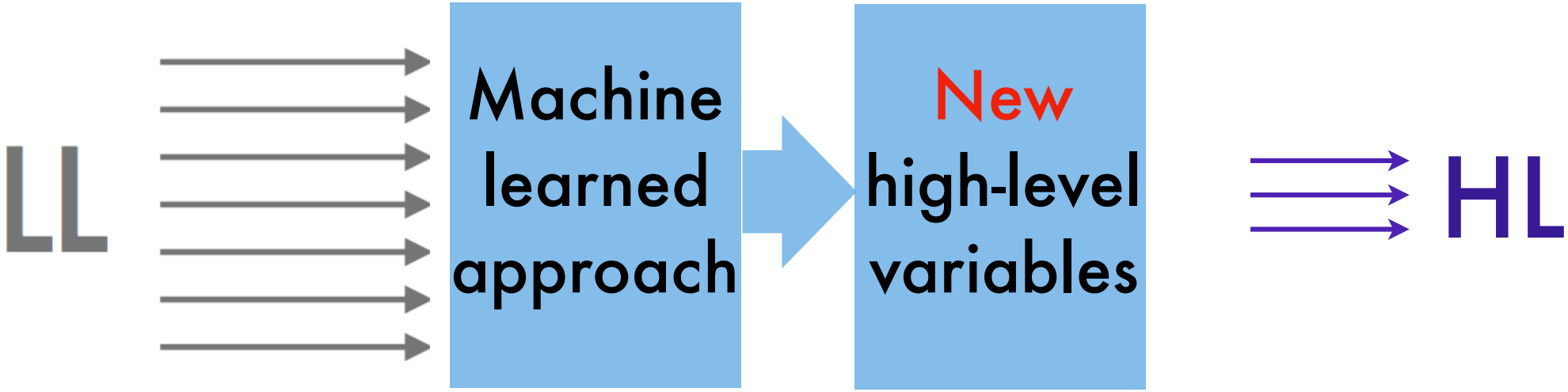
<u>High-Level data is preferred</u>

- It is easier to understand
- Its modeling can be verified
- Uncertainties can be sensibly defined
- It is more compact and efficient

<u>Goal</u>:

Find new HL features to include insights from ML probe of LL data

Learning from ML



Use LL analysis as a probe, not a final product.

How

I. Define complete space of valid HL features

II. Define mapping strategy

III. Map ML into space of HL features

I. Define space of features

I. Define complete space of valid HL features

- provides context for NN solution,
- defines problem: map NN into this space

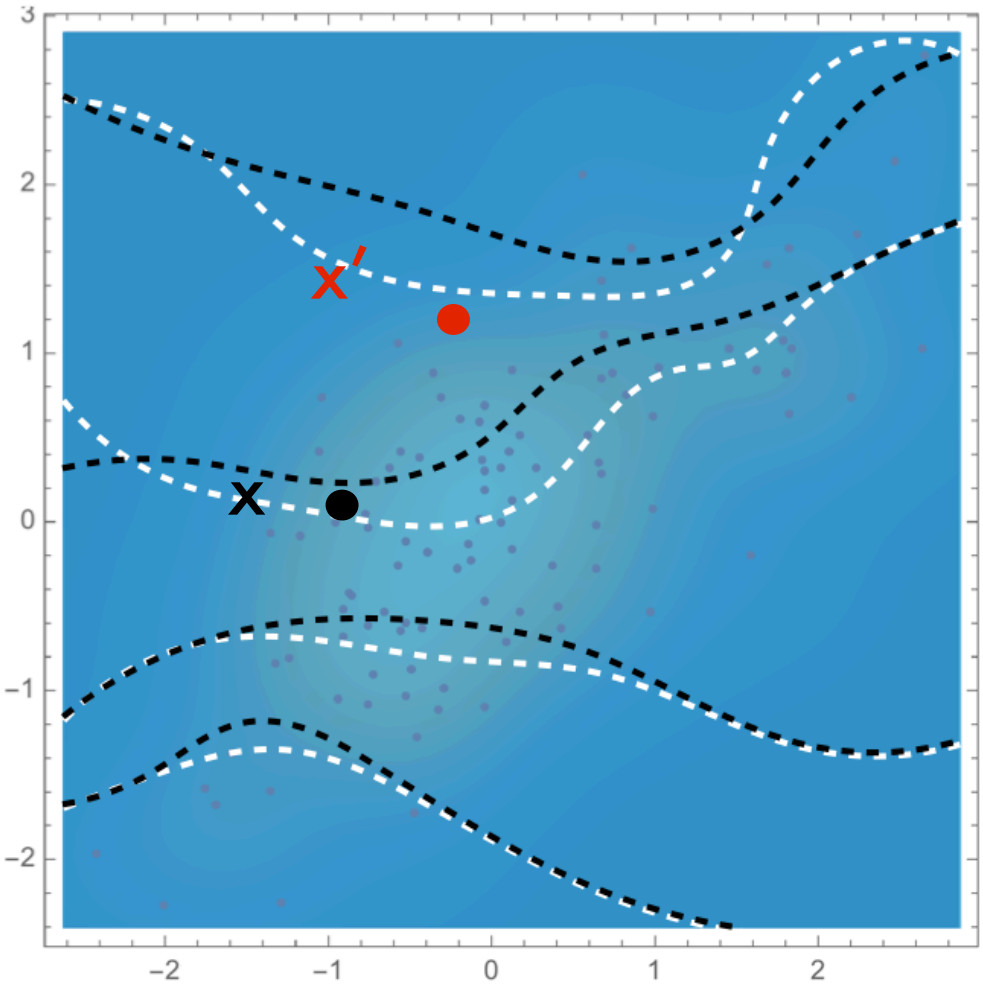
Energy Flow Polynomials (EFP)

$$= \sum_a \sum_b \sum_c z_a z_b z_c \theta_{ab} \theta_{ac} \theta_{bc}^2$$

$$z_a^{(\kappa)} = \left(\frac{p_{\mathrm{T}a}}{\sum_b p_{\mathrm{T}b}} \right)^{\kappa},$$

$$\theta_{ab}^{(\beta)} = \left(\Delta\eta_{ab}^2 + \Delta\phi_{ab}^2 \right)^{\beta/2}.$$

II. Mapping Strategy



An EFP maps to DNN if they have the same **decision ordering**:

DNN(x) > DNN(x′)
EFP(x) > EFP(x′)

III. Map ML to new HL

A. Can we find the missing piece?

- Find a supplemental HL feature X
- Goal: {6HL+X} reproduces DNN

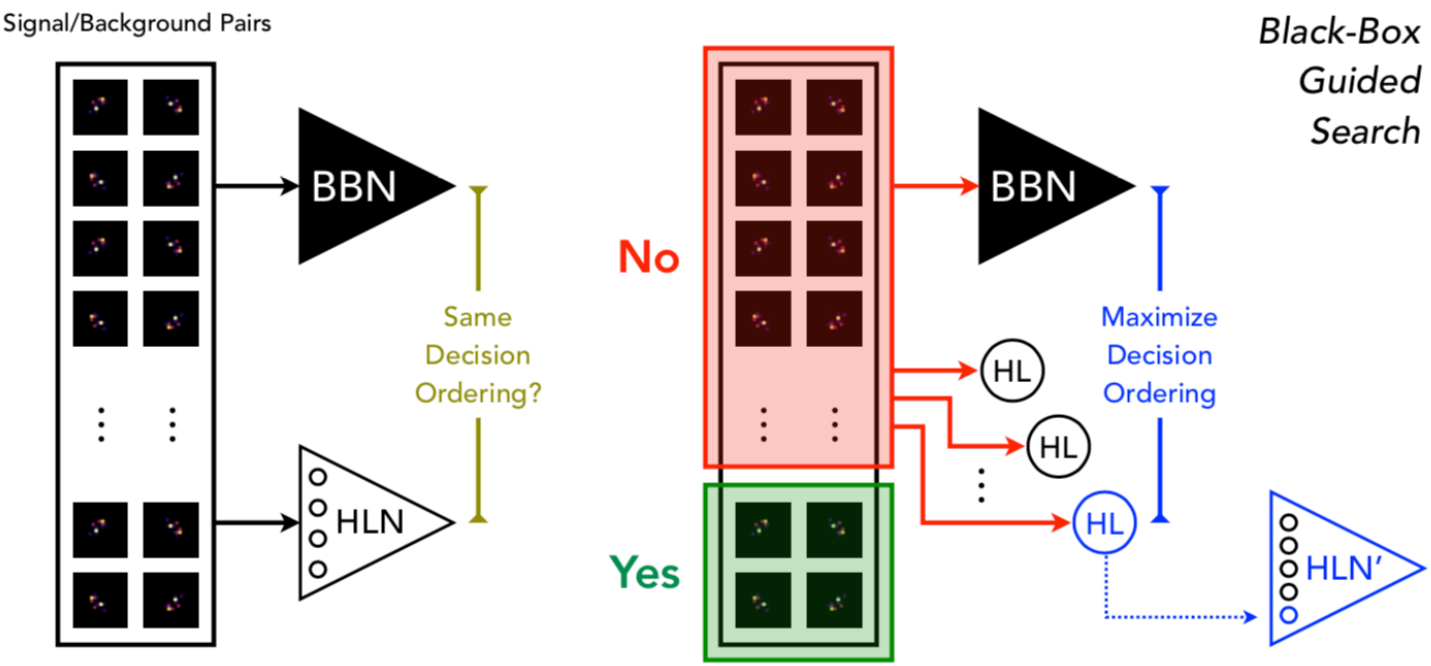
B. Can we rediscover HL features?

- Find a whole new set of HL features {X1,X2...}
- Goal {X1,X2,...} reproduces DNN

III. Map ML to new HL

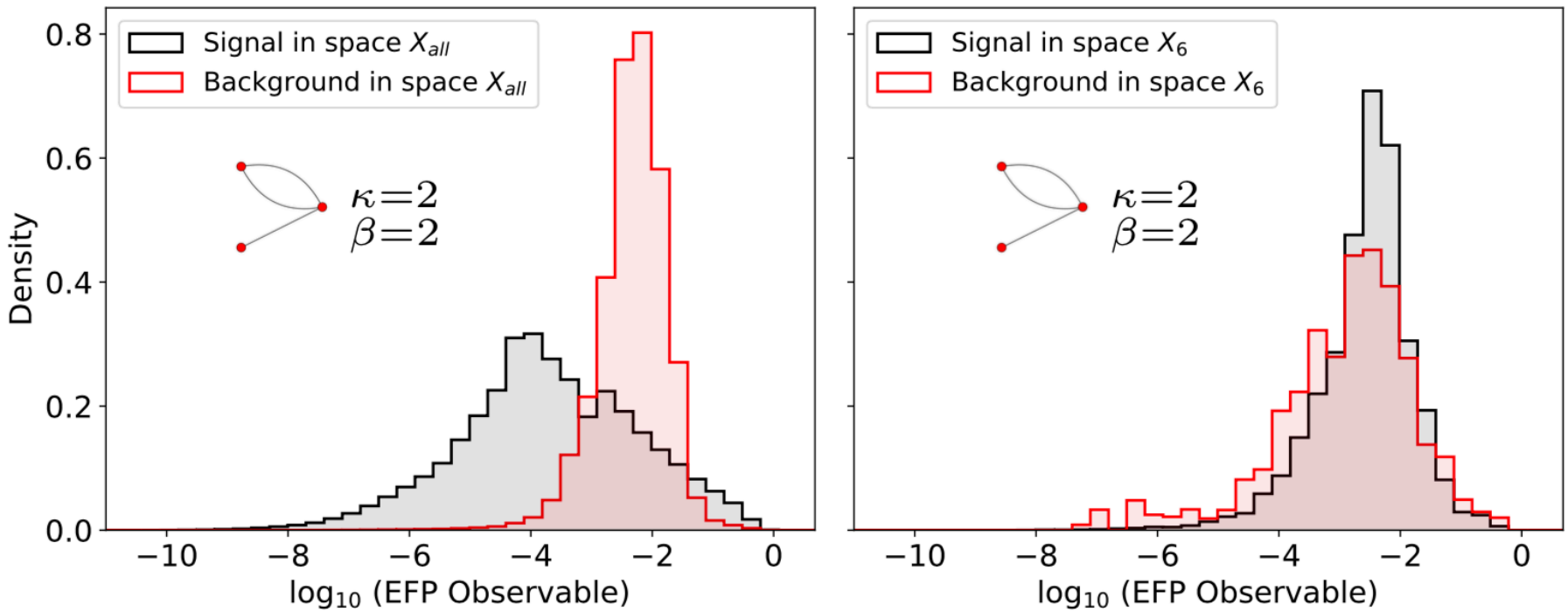
A. Can we find the missing piece?

- Find a supplemental HL feature X
- Focus on subset of data decisions disagree



1. *Create input pairs*
2. *Identify pairs where DN and HLN disagree*
3. *Over those pairs, find HL feature with max DO similarity.*
4. *add to HLN*

The winner


Signal in space X_{all}
Background in space X_{all}
$\kappa=2$
$\beta=2$
Density
0.0
0.2
0.4
0.6
0.8
−10
−8
−6
−4
−2
0
$\log_{10}$ (EFP Observable)
Signal in space X_6
Background in space X_6
$\kappa=2$
$\beta=2$
−10
−8
−6
−4
−2
0
$\log_{10}$ (EFP Observable)

It works!

Observable	AUC	ADO[CNN, Obs.]
M_{jet}	0.898 ± 0.004	0.807
$C_2^{\beta=1}$	0.660 ± 0.006	0.584
$C_2^{\beta=2}$	0.604 ± 0.007	0.548
$D_2^{\beta=1}$	0.790 ± 0.005	0.743
$D_2^{\beta=2}$	0.807 ± 0.005	0.762
$\tau_2^{\beta=1}$	0.662 ± 0.006	0.600
6HL	0.9504 ± 0.0002	0.971
CNN	0.9531 ± 0.0002	1.000
$7\text{HL}_{\text{black-box}}$	0.9529 ± 0.0002	0.973

6HL+ 

Found new HL feature that closes gap with DNN

What is it?

K = 2 graphs

- not IR collinear safe
- usually avoided

$$z_a^{(\kappa)} = \left(\frac{p_{\mathrm{T}a}}{\sum_b p_{\mathrm{T}b}}\right)^{\kappa},$$

$$\theta_{ab}^{(\beta)} = \left(\Delta\eta_{ab}^2 + \Delta\phi_{ab}^2\right)^{\beta/2}.$$

Rank	EFP	κ	β	ADO[EFP, CNN]$_{X_6}$	AUC[EFP]	ADO[6HL + EFP, CNN]$_{X_{\mathrm{all}}}$	AUC[6HL + EFP]
1		2	2	0.647 ± 0.003	0.807	0.972	0.953
2		2	2	0.646 ± 0.003	0.802	0.972	0.953
3		2	1	0.646 ± 0.003	0.813	0.972	0.953
4		2	1	0.644 ± 0.003	0.807	0.972	0.953
5		2	1	0.641 ± 0.003	0.807	0.972	0.953

III. Map ML to new HL

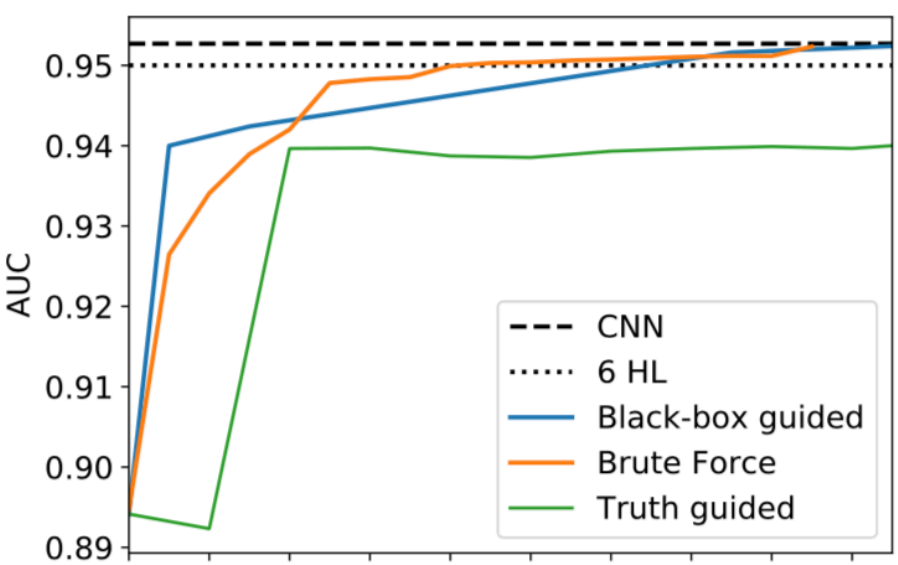
B. Can we rediscover HL features?

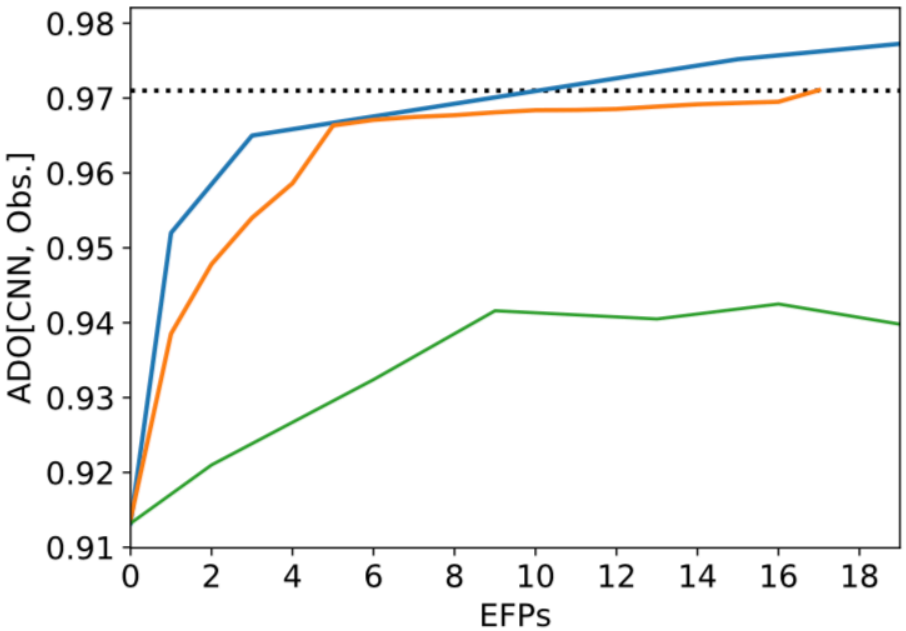
- Find a whole new set of HL features {X}

III. Map ML to new HL

II. Can we rediscover HL features?

- Find a **whole new set** of HL features {X}
- Also tried comparing to truth (skipping DN)



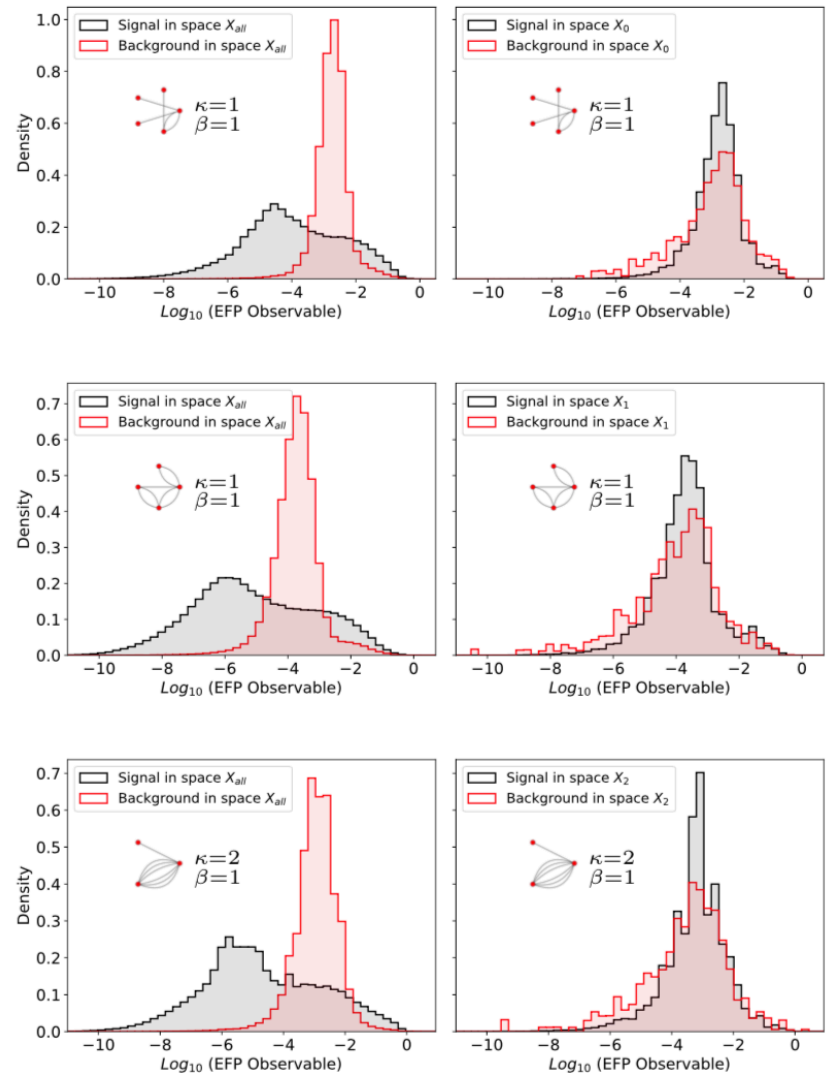


Some weirdos

Graphs are weird

- some IRC unsafe
- others have high chromatic number

QCD theorists have not considered these classes



Conclusions

I. Use DNNs as a probe of information, not final tool

II. Map DNN decisions into space of interpretable functions

- more compact
- can describe systematics
- can determine validity
- can reveal unexplored ideas