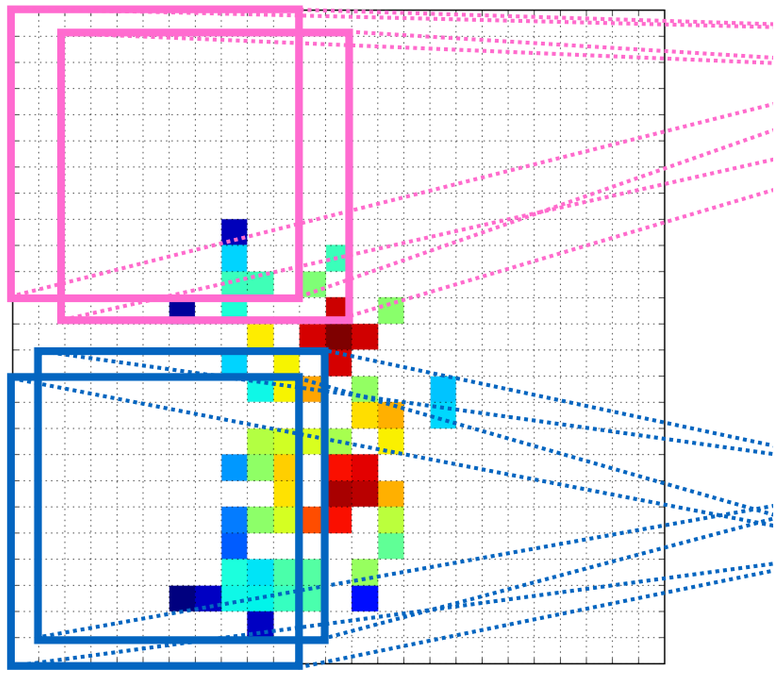
Machine Learning Overview
(given the session, focus is on QCD)

Benjamin Nachman

Lawrence Berkeley National Laboratory

bpnachman.com
bpnachman@lbl.gov
@bpnachman
bnachman



Snowmass
EF Kick Restart

September 1, 2021

Disclaimer

I have been asked to present a “summary of recent developments in machine learning for HEP”.

Disclaimer

I have been asked to present a “summary of recent developments in machine learning for HEP”.

This is a very broad topic and there is no way I can do it justice in ~15 min.

Disclaimer

I have been asked to present a “summary of recent developments in machine learning for HEP”.

This is a very broad topic and there is no way I can do it justice in ~15 min.

Given that this session is co-organized by QCD-related topical groups, I will use the recent ML4Jets workshop as a roadmap for this talk.

Disclaimer

I have been asked to present a “summary of recent developments in machine learning for HEP”.

This is a very broad topic and there is no way I can do it justice in ~15 min.

Given that this session is co-organized by QCD-related topical groups, I will use the recent ML4Jets workshop as a roadmap for this talk.

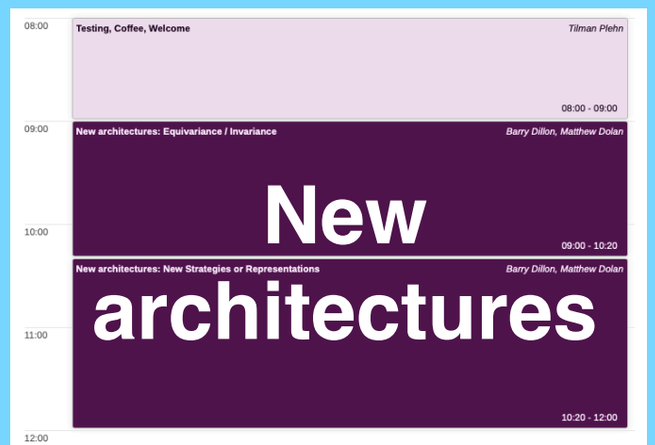
For a comprehensive review (anyone can contribute!) see: https://iml-wg.github.io/HEPML-LivingReview/

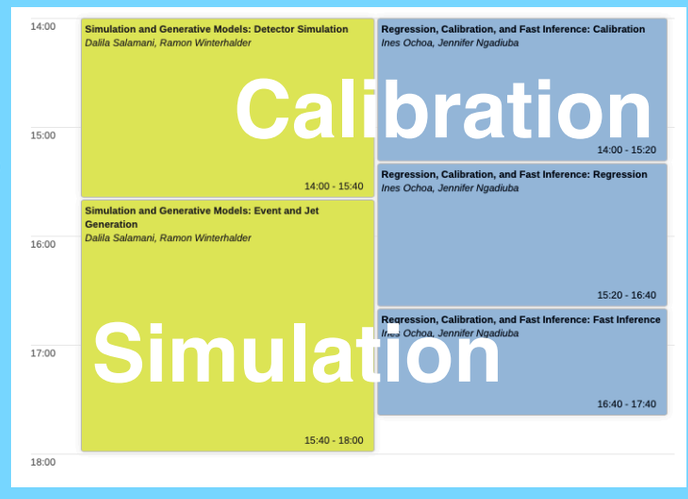
ML4Jets2021

The annual ML4Jets conference about a month ago had 100 talks in three days (!)

N.B. most plots are links!

ML4Jets2021





ML4Jets2021

I won't cover everything - just giving you a taste!

…my apologies in advance for not covering your / favorite topic.

New architectures

A **hot topic** in this area is **equivariance** / **invariance**

A NN is **equivariant** if it **commutes** with the symmetry group and a NN is **invariant** if the output is **unchanged** under symmetries of the inputs

New architectures

A **hot topic** in this area is **equivariance** / **invariance**

A NN is **equivariant** if it **commutes** with the symmetry group and a NN is **invariant** if the output is **unchanged** under symmetries of the inputs

Equivariant example

Learn features that transform under rotations in the same way as the inputs - then feed these into further layers

New architectures

A **hot topic** in this area is **equivariance** / **invariance**

A NN is **equivariant** if it **commutes** with the symmetry group and a NN is **invariant** if the output is **unchanged** under symmetries of the inputs

Equivariant example

Learn features that transform under rotations in the same way as the inputs - then feed these into further layers

Invariant example

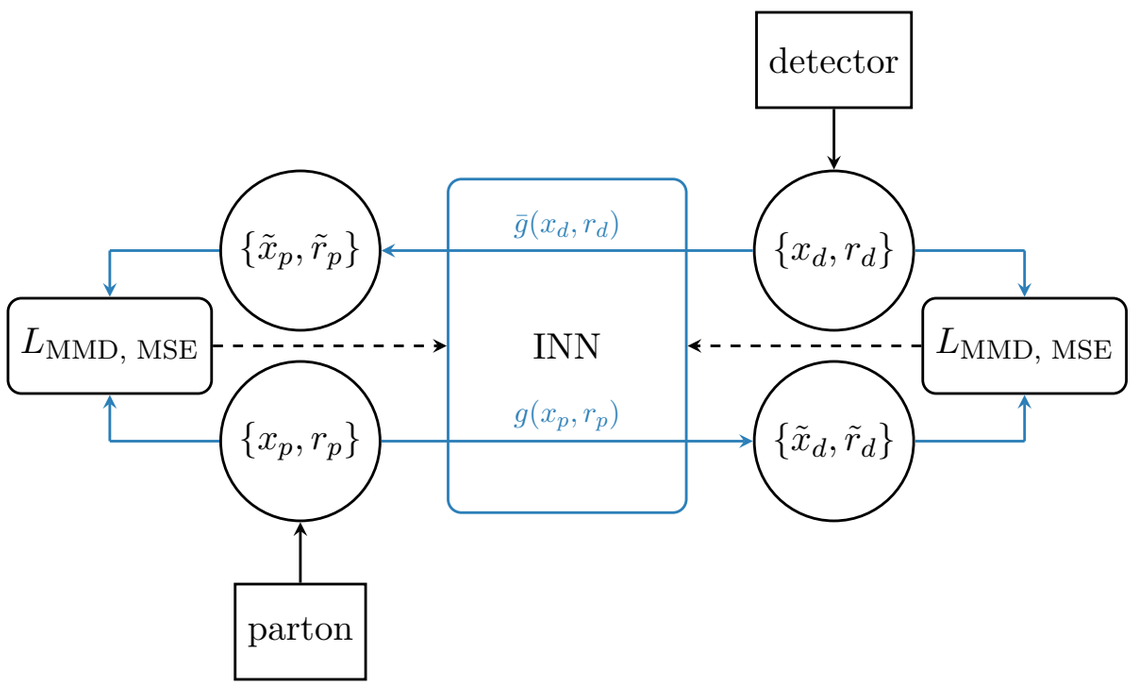
Event/jet constituents are permutation invariant - use Deep Sets, Graph Networks, Transformers, Attention, …

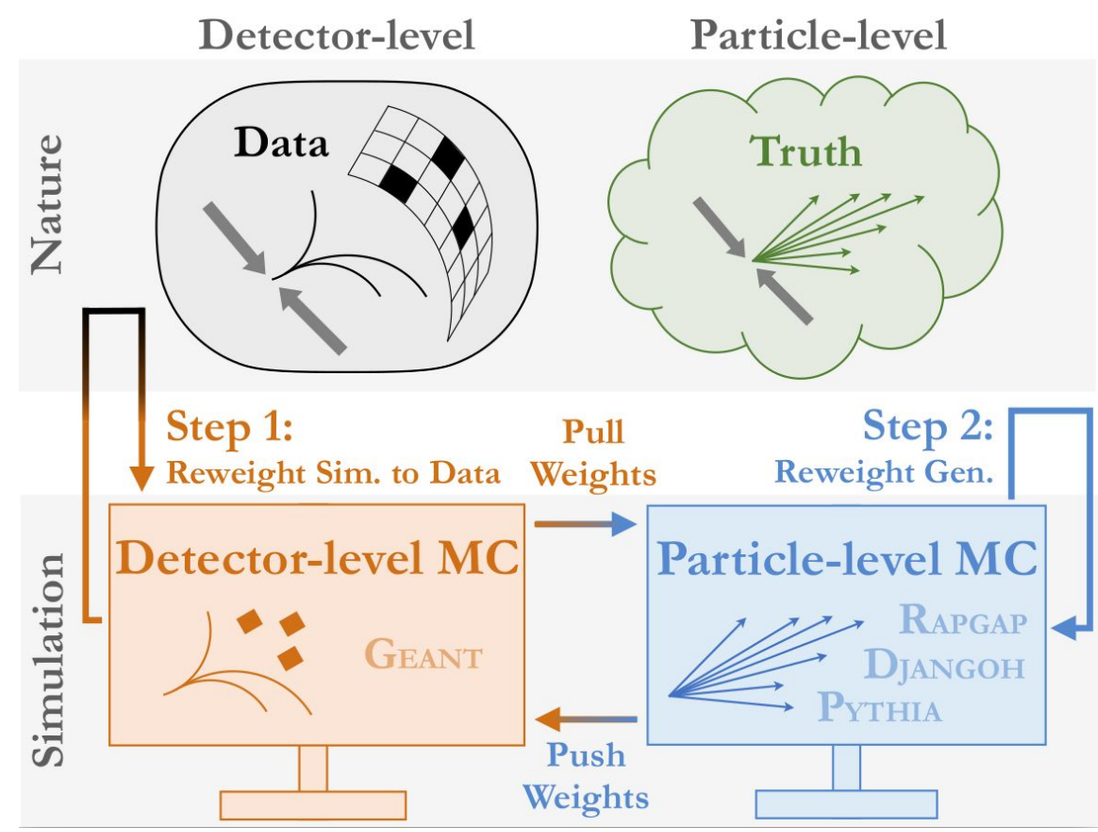
Measurements

Significant interest in ML-based unfolding for high-(or even variable-)dimensional, unbinned measurements

Measurements

Significant interest in ML-based unfolding for high-(or even variable-)dimensional, unbinned measurements



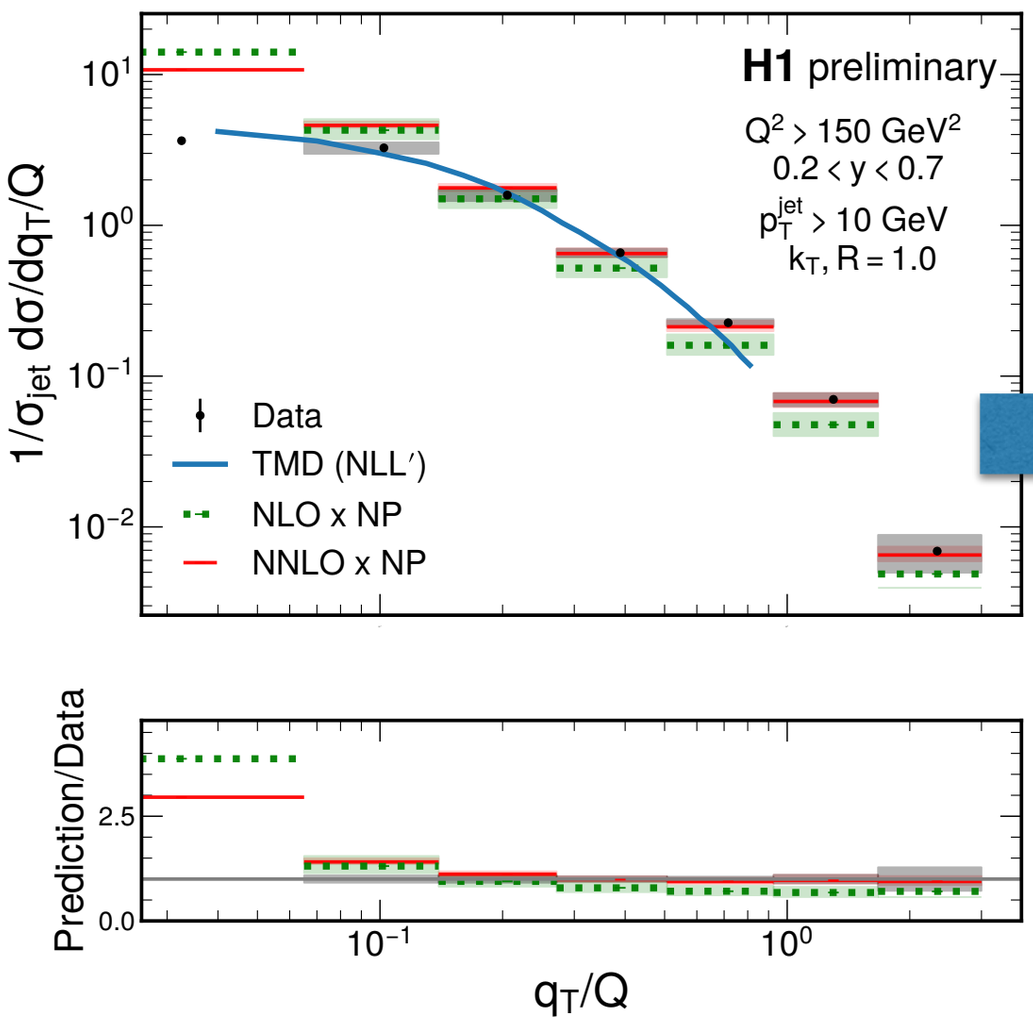


Generative model-based
2006.06685

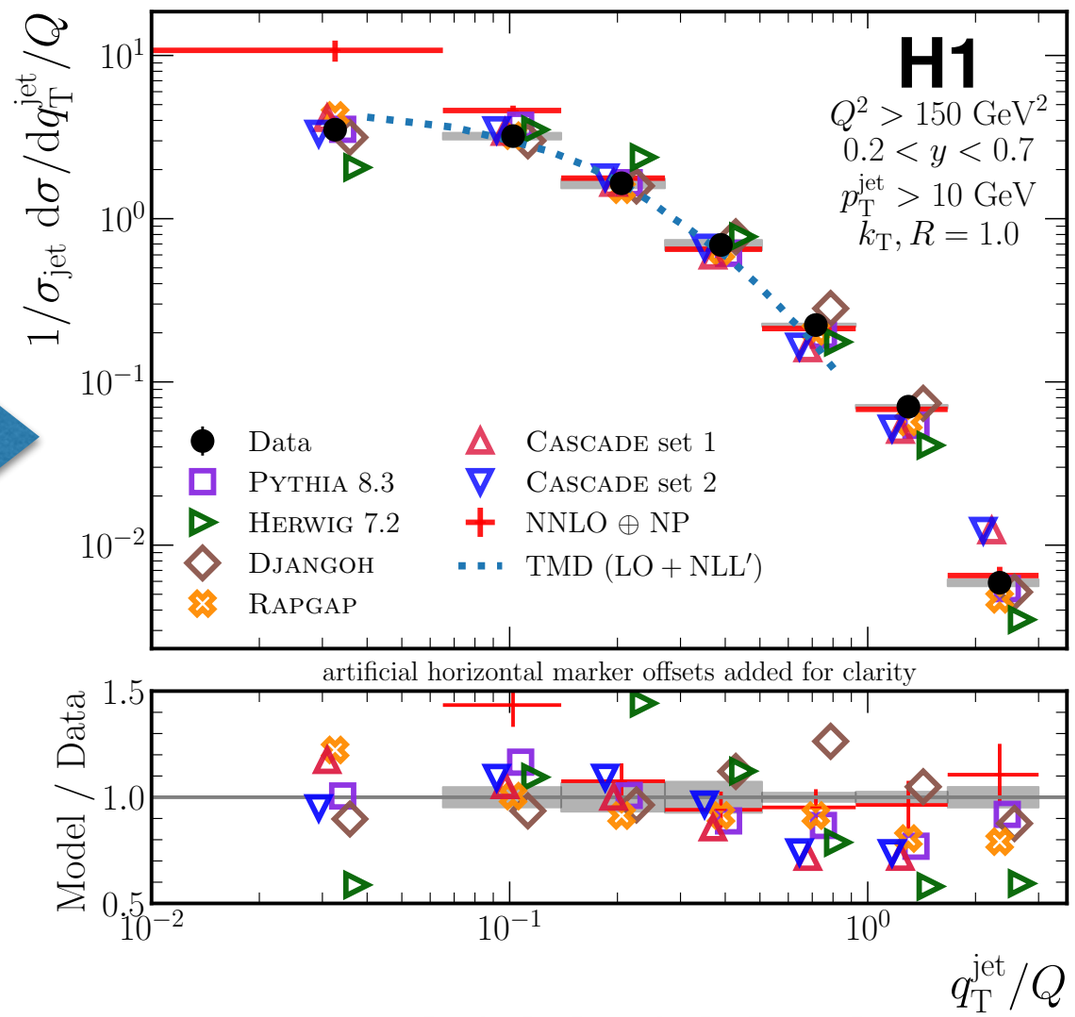
Classifier-based
1911.09107

Measurements

First application to collider data!



H1prelim-21-031



2108.12376
(two days ago!)

G

ed

2006.06685

1911.09107

BSM

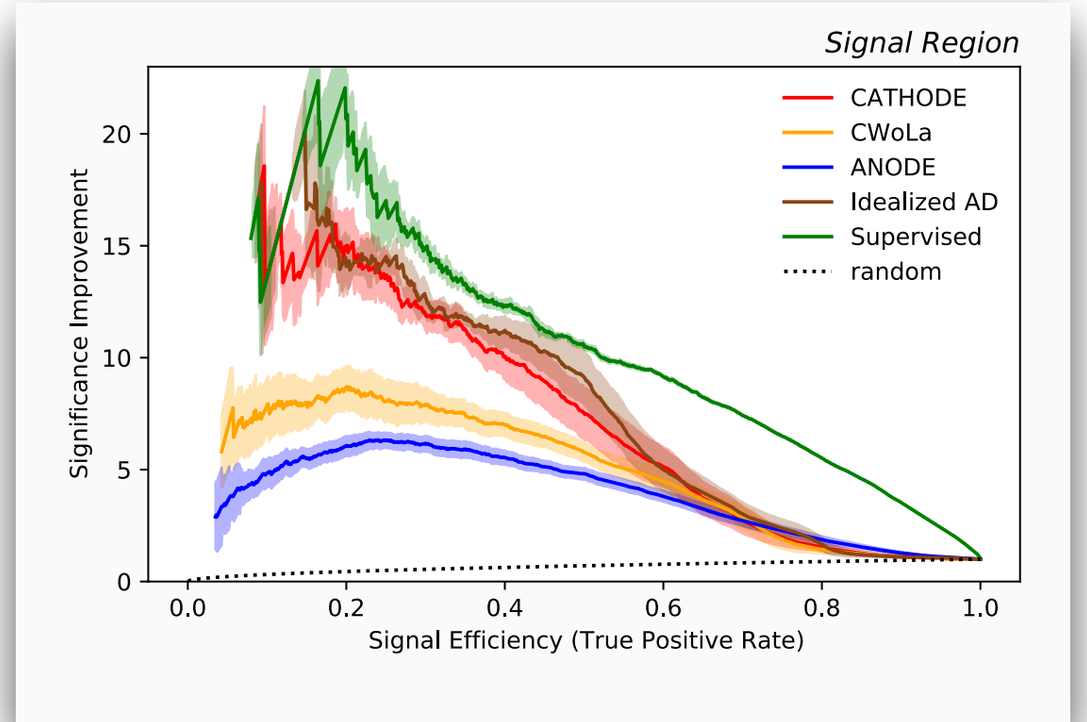
A **hot topic** in this area is **anomaly detection**

i.e. using unsupervised/weakly-supervised/semi-supervised learning to reduce signal/background model dependence

BSM

A **hot topic** in this area is **anomaly detection**

i.e. using unsupervised/weakly-supervised/semi-supervised learning to reduce signal/background model dependence

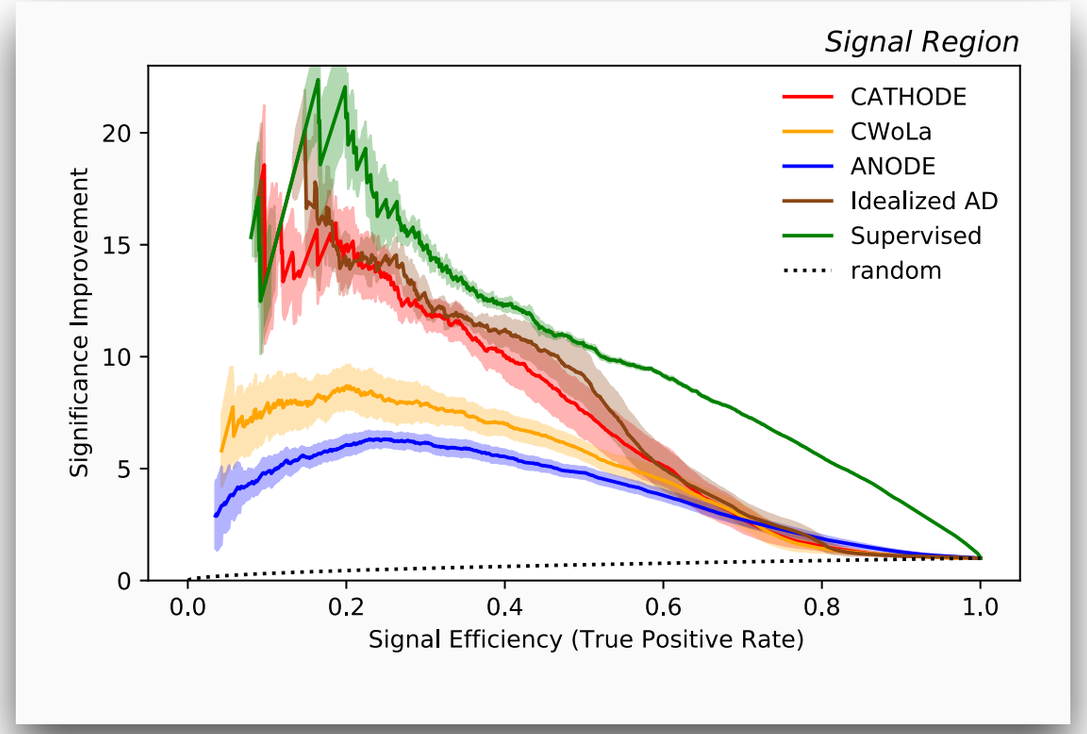


New methods are saturating **bounds** in some regimes

BSM

A **hot topic** in this area is **anomaly detection**

i.e. using unsupervised/weakly-supervised/semi-supervised learning to reduce signal/background model dependence



Key questions remain:
how to do model selection for unsupervised methods? How to best estimate the background? What about the non-resonant case?

New methods are saturating **bounds** in some regimes

Classification

1902.09914 + H. Qu

Top tagging landscape

	AUC	Acc	$1/\epsilon_B$ ($\epsilon_S = 0.3$) single	mean	median	#Param
CNN [16]	0.981	0.930	914±14	995±15	975±18	610k
ResNeXt [30]	0.984	0.936	1122±47	1270±28	1286±31	1.46M
TopoDNN [18]	0.972	0.916	295±5	382± 5	378 ± 8	59k
Multi-body N-subjettiness 6 [24]	0.979	0.922	792±18	798±12	808±13	57k
Multi-body N-subjettiness 8 [24]	0.981	0.929	867±15	918±20	926±18	58k
TreeNiN [43]	0.982	0.933	1025±11	1202±23	1188±24	34k
P-CNN	0.980	0.930	732±24	845±13	834±14	348k
ParticleNet [47]	0.985	0.938	1298±46	1412±45	1393±41	498k
LBN [19]	0.981	0.931	836±17	859±67	966±20	705k
LoLa [22]	0.980	0.929	722±17	768±11	765±11	127k
Energy Flow Polynomials [21]	0.980	0.932	384			1k
Energy Flow Network [23]	0.979	0.927	633±31	729±13	726±11	82k
Particle Flow Network [23]	0.982	0.932	891±18	1063±21	1052±29	82k
GoaT	0.985	0.939	1368±140		1549±208	35k
ParticleNet-Lite	0.984	0.937	1262±49			26k
ParticleNet	0.986	0.940	1615±93			366k
ParticleNeXt	0.987	0.942	1923±48			560k

Graph-based

State-of-the-art classification performance continues to improve! New tricks like self-attention, etc.

Calibration

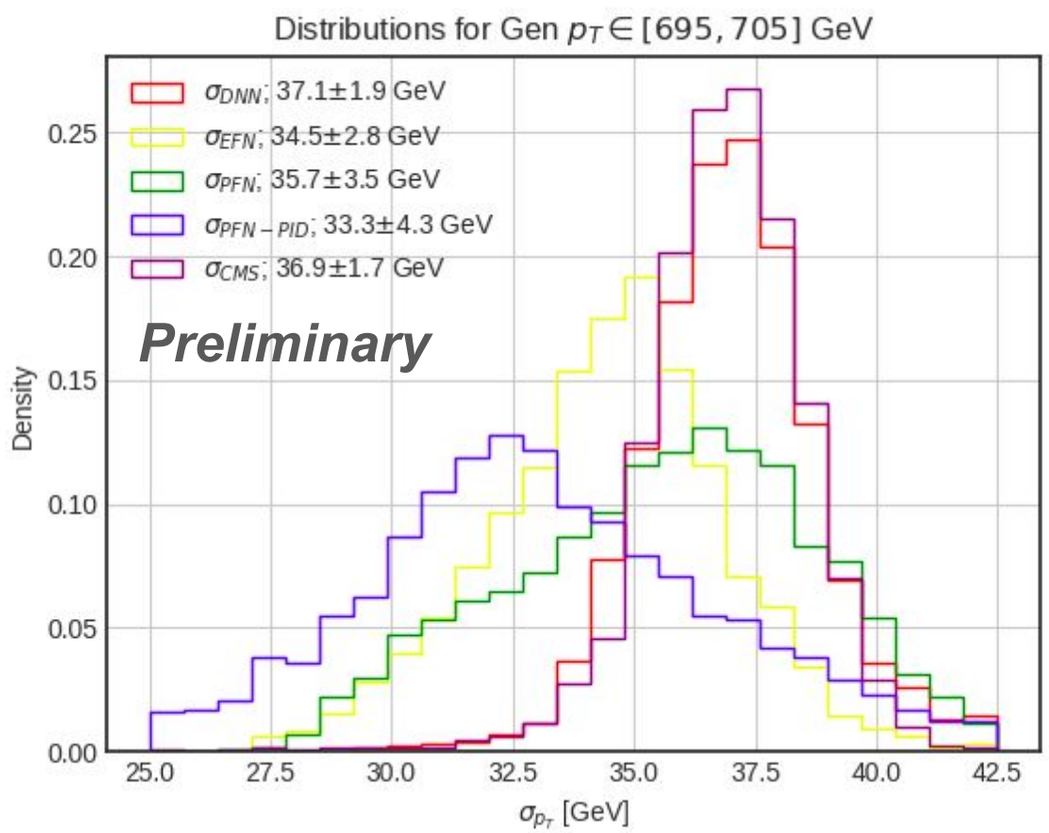
This is often set up as a regression task.

Innovation on many fronts, including the combination of various sources of information, mitigation of pileup, ...

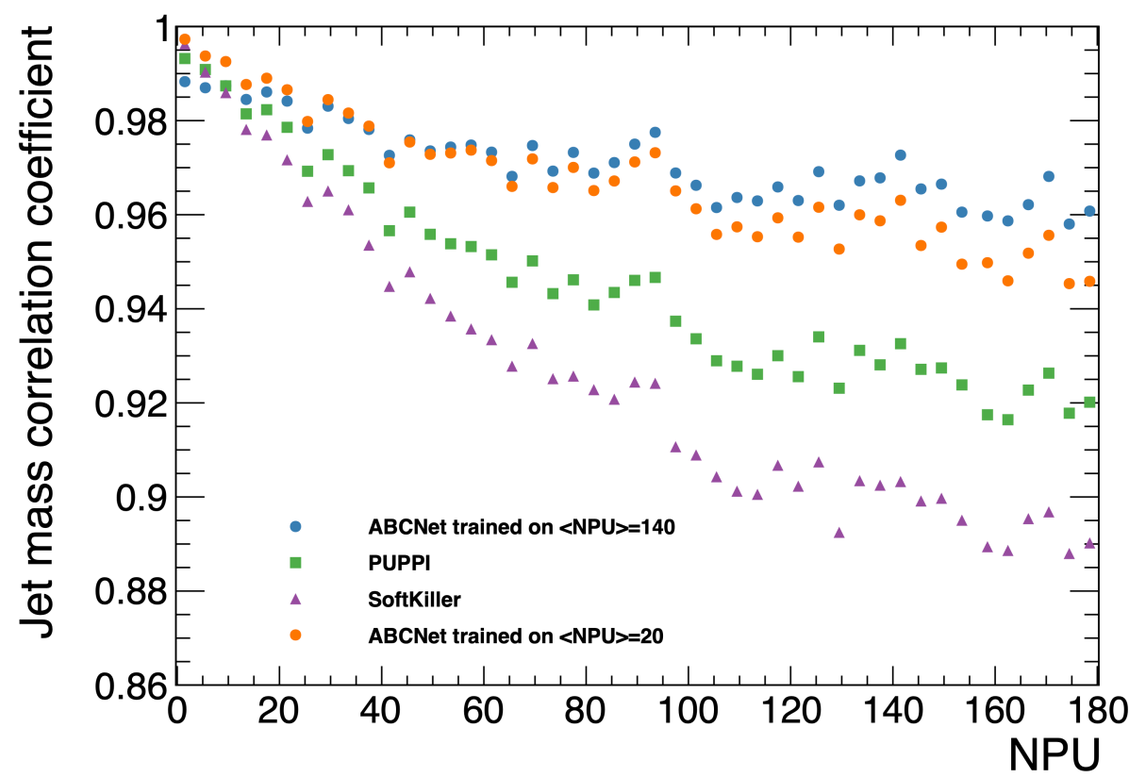
Calibration

This is often set up as a regression task.

Innovation on many fronts, including the combination of various sources of information, mitigation of pileup, ...



Ex: Prior-independent jet calibrations



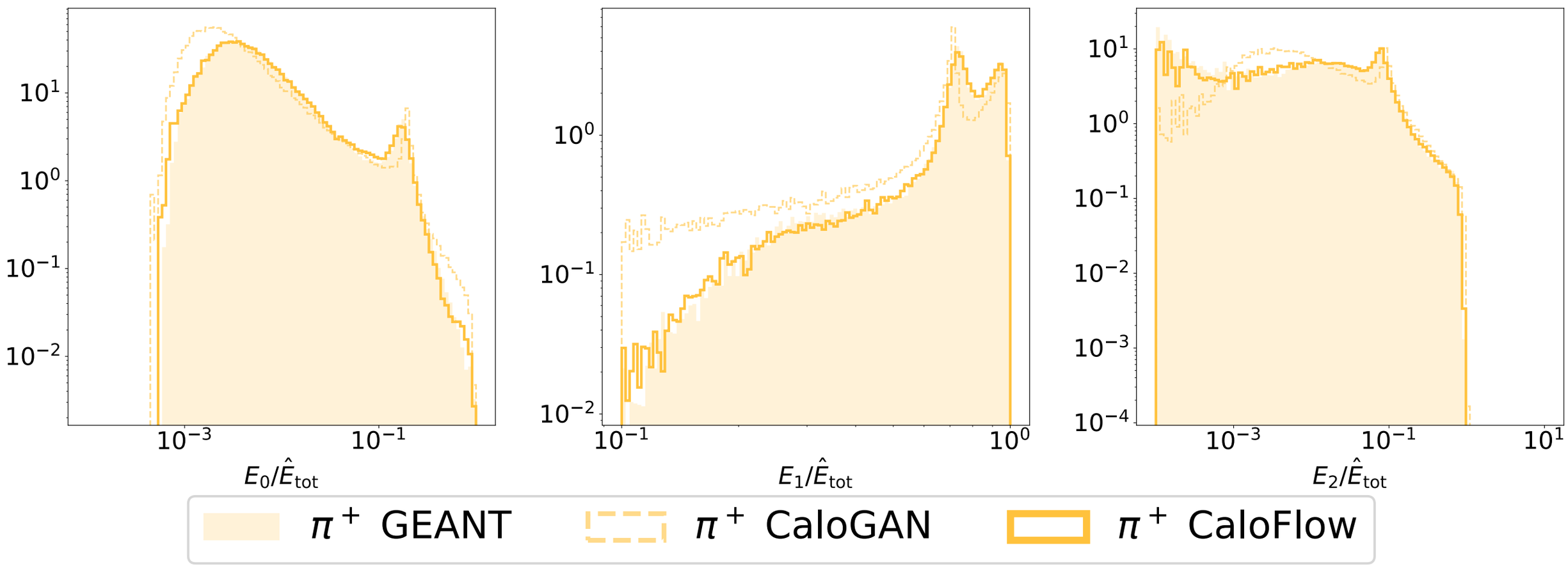
Ex: Graph-based Pileup Mitigation

Simulation

A **hot topic** in this area is **fast calorimeter simulation**

Simulation

A **hot topic** in this area is **fast calorimeter simulation**

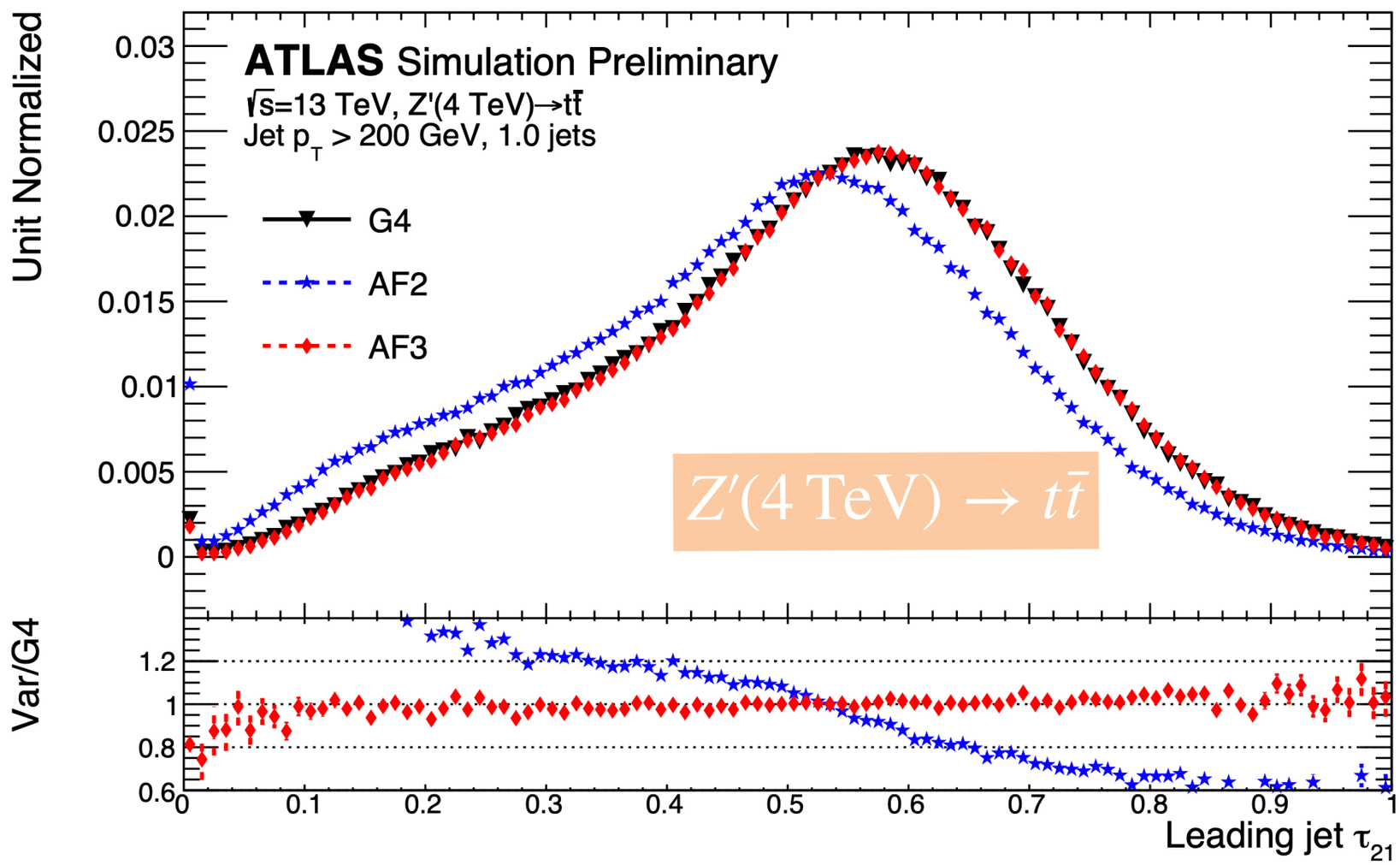


2106.05285

State-of-the-art with GANs and Normalizing Flows are reaching precision!

Simulation

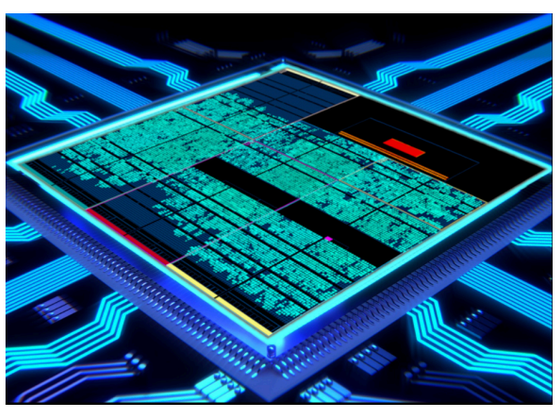
Now with a full integration into a collider simulation!



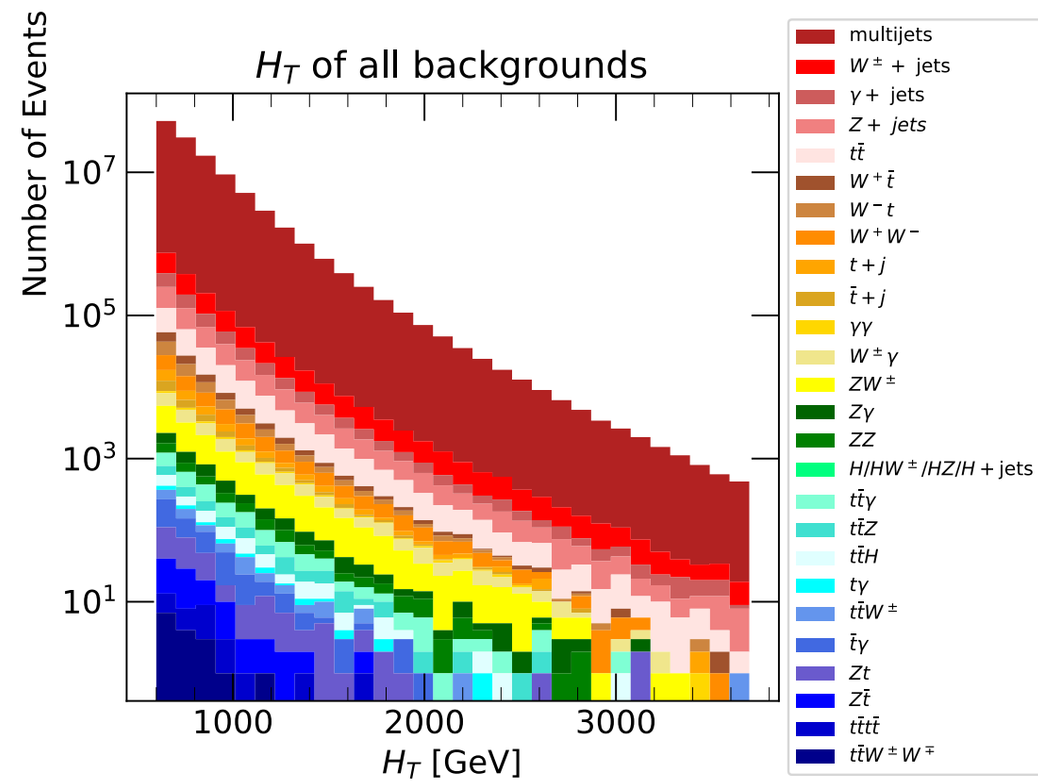
(AF3 uses a GAN for intermediate energies)

Datasets

LHC Olympics



Real-time Anomaly Detection



Dark Machines

+ more presented at ML4Jets and beyond!

see also https://iml.web.cern.ch/public-datasets

Exploring Data

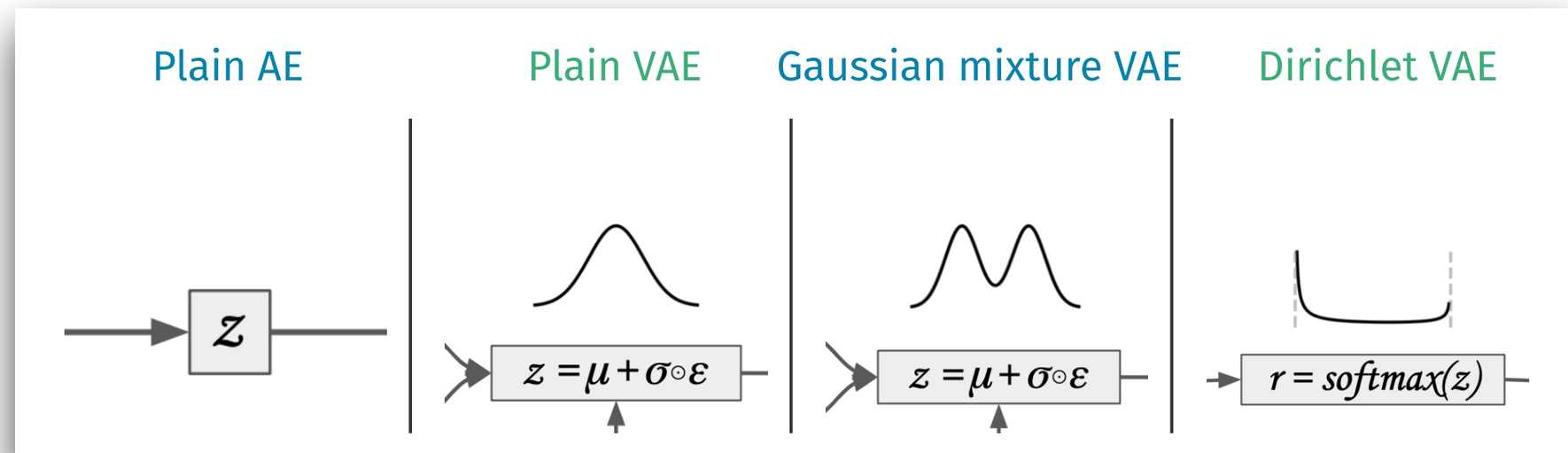
Discovering / categorizing **latent** structure in data

…this could be symmetries or multi-class components, etc.

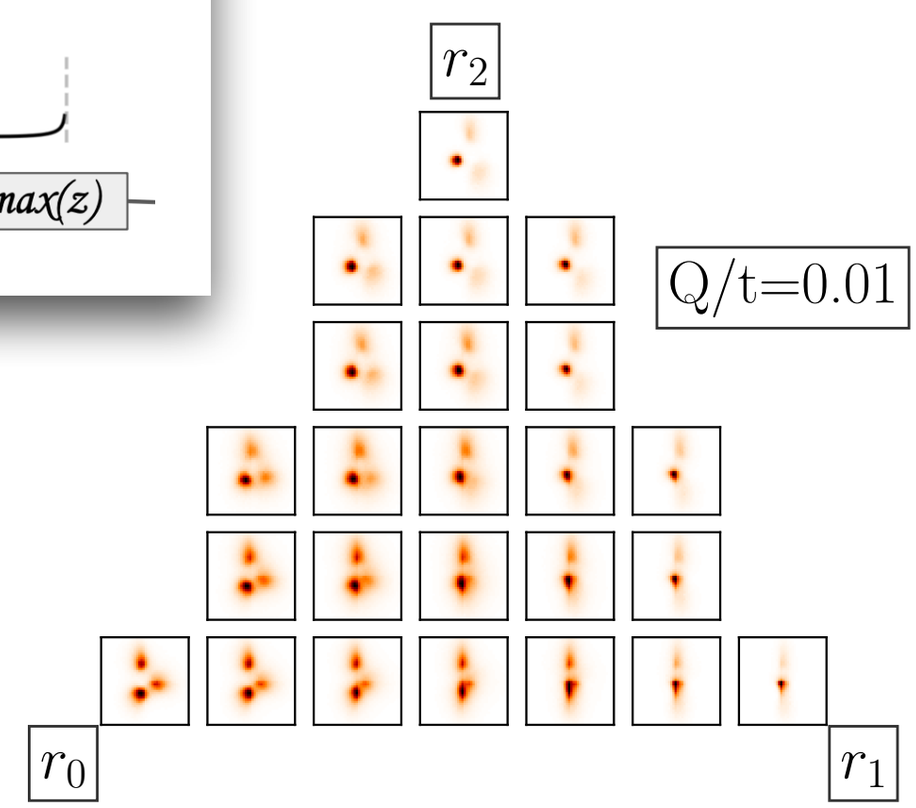
Exploring Data

Discovering / categorizing **latent** structure in data

…this could be symmetries or multi-class components, etc.



Imposing structure can lead to more interpretable latent spaces



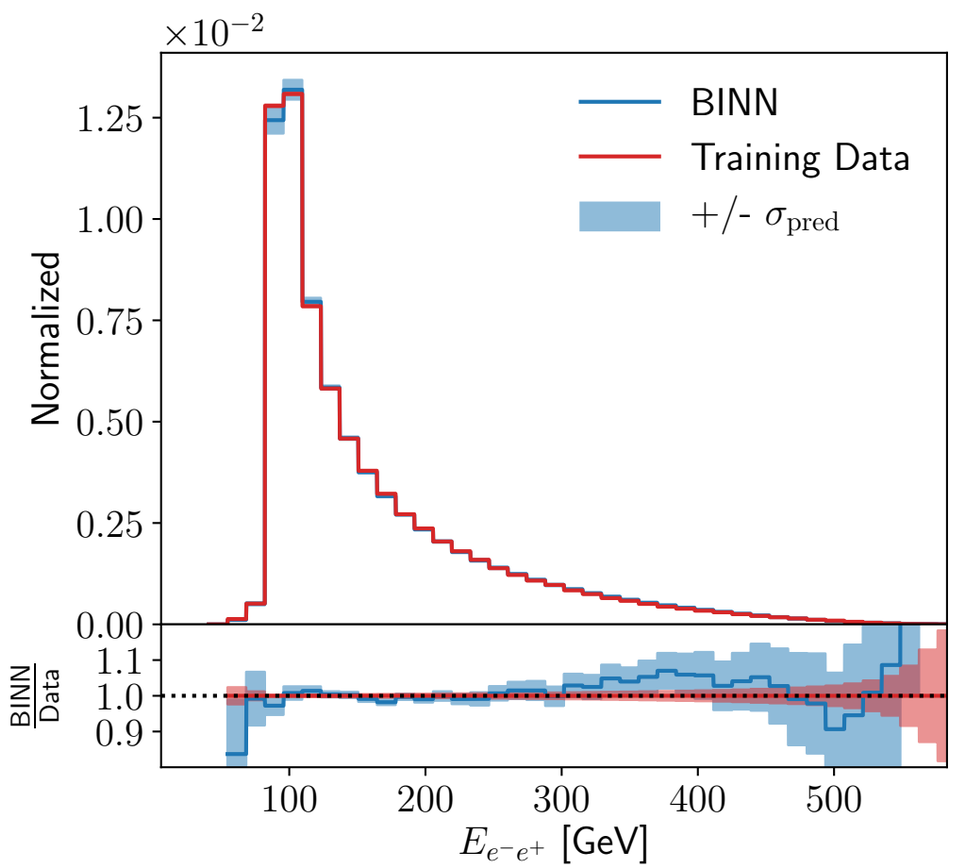
2104.0829

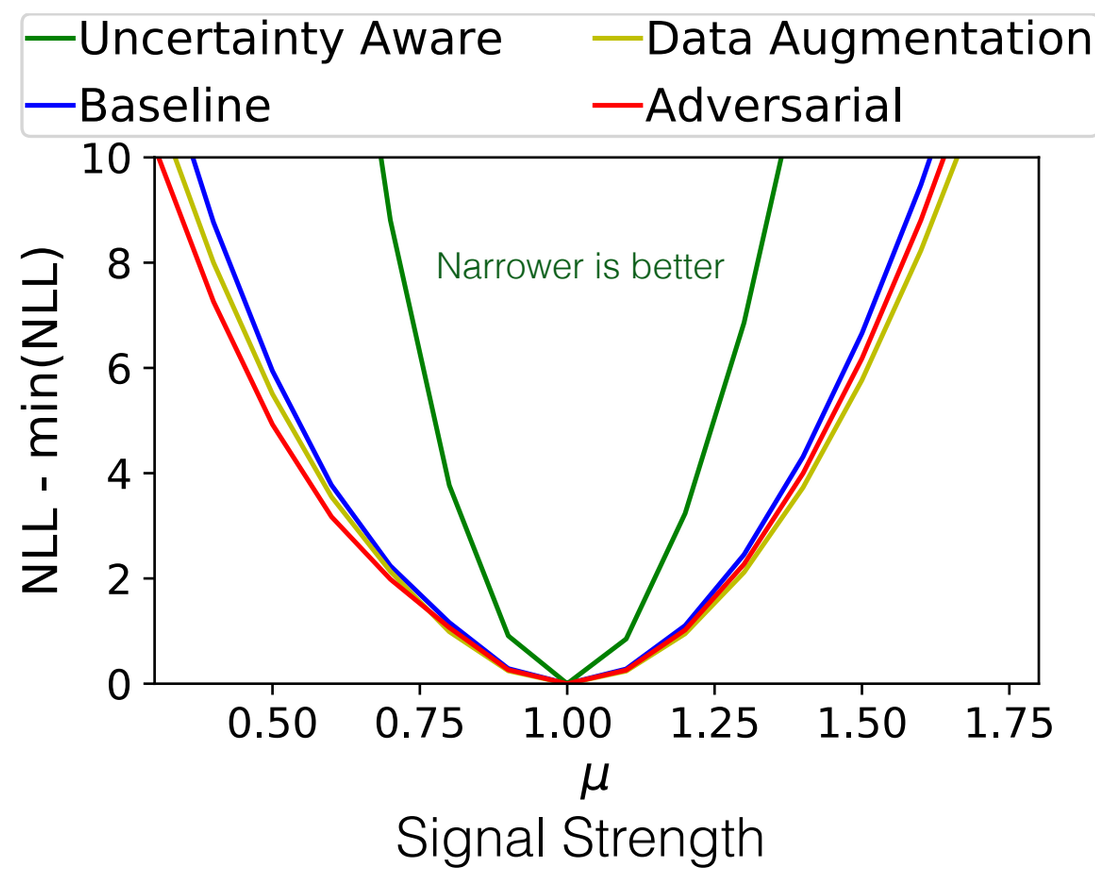
Interpretability and Uncertainties

Key questions: *what are uncertainties associated with neural networks? How to make networks use uncertainty information (uncertainty-aware)? How to make networks optimal with respect to downstream analysis (Inference-aware)?*

Interpretability and Uncertainties

Key questions: *what are uncertainties associated with neural networks? How to make networks use uncertainty information (uncertainty-aware)? How to make networks optimal with respect to downstream analysis (Inference-aware)?*





Bayesian generative models, parameterized uncertainty networks, …

Data analysis in HEP

Theory of everything

↓

Physics simulators

↓

Detector-level observables

↓

Pattern recognition ⟷ Pattern recognition

Nature

↓

Experiment

↓

Detector-level observables

↓

Data analysis in HEP + ML

Theory of everything

Fast simulation / phase space

Parameter estimation / unfolding

Physics simulators

Detector-level observables

Pattern recognition

Nature

Experiment

Online processing & quality control

Detector-level observables

Data curation

Pattern recogn

Classification to enhance sensitivity

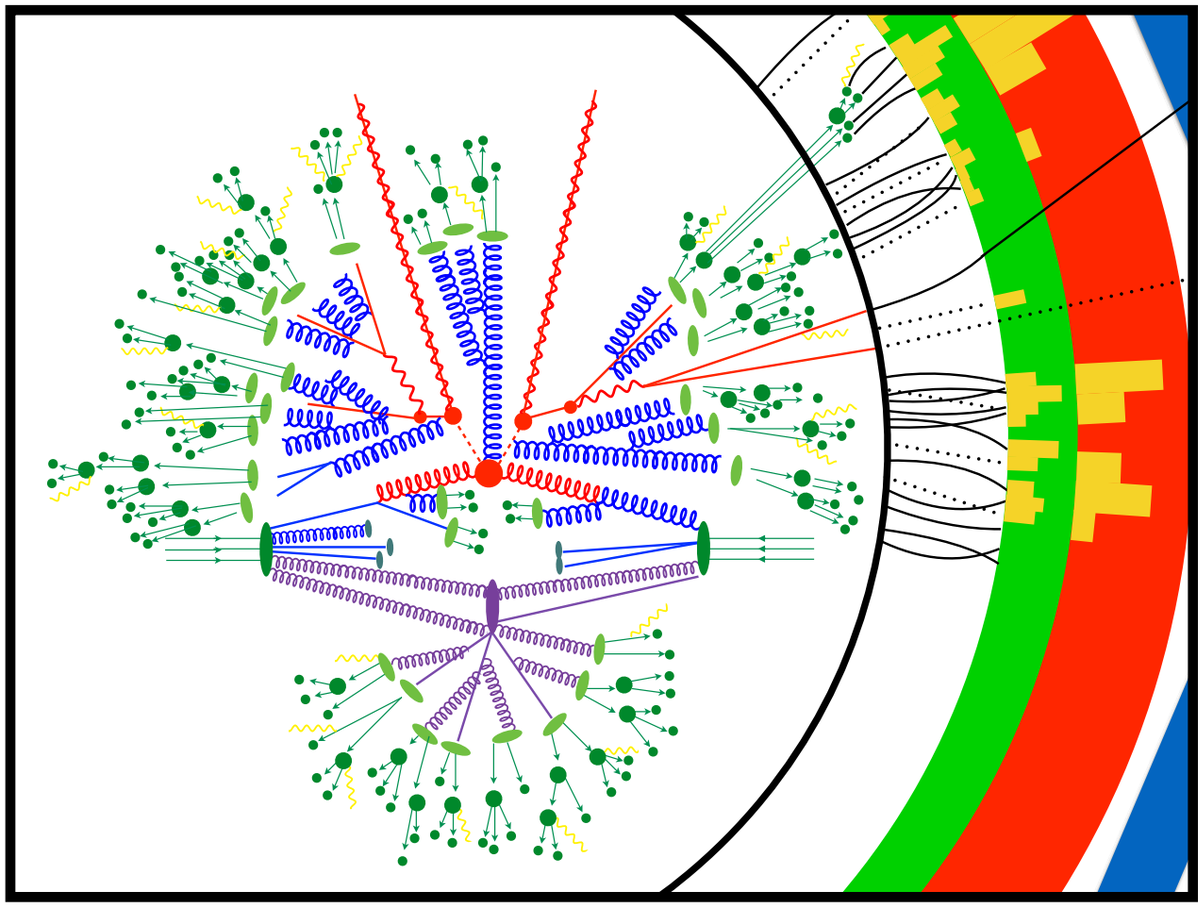
"signal" versus "background"

calibration
clustering
tracking
noise mitigation
particle identification
...

Conclusions and outlook

Deep learning has a great potential to **enhance**, **accelerate**, and **empower** energy frontier physics

Due to the limited time, I was only able to cover a small selection of new ideas and results



Machine learning can improve all areas of our experimental and theoretical workflows !

Backup