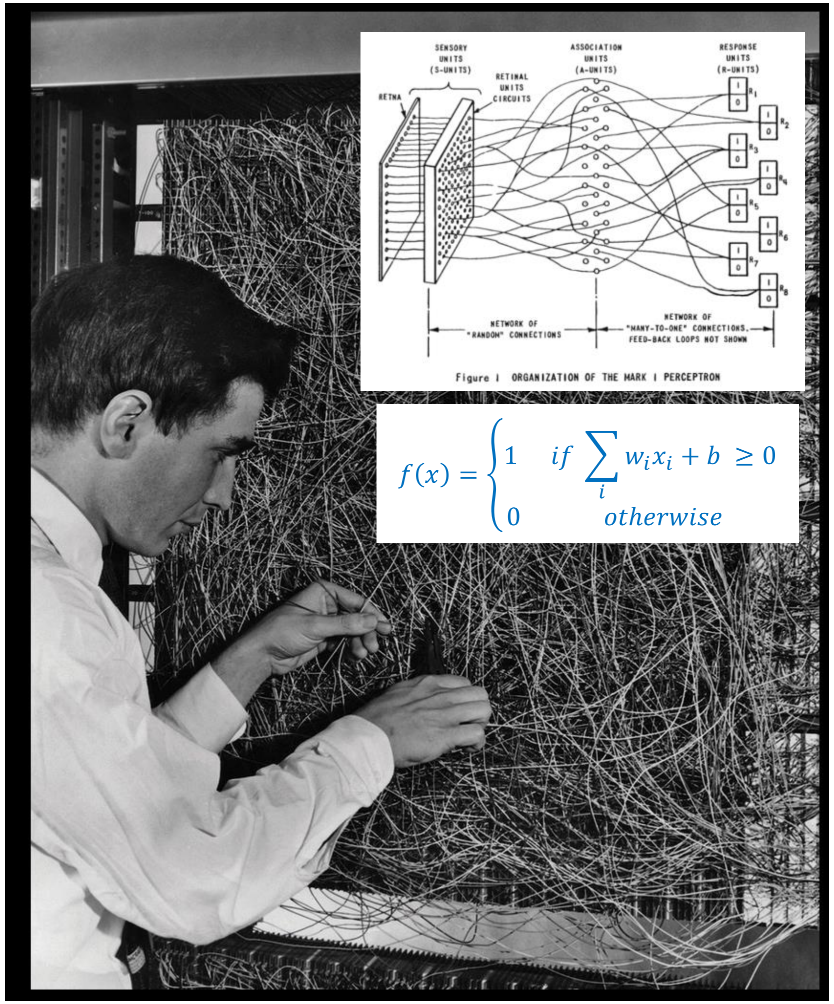
Introduction to Machine Learning in High Energy Physics

Michael Kagan
SLAC National Accelerator Laboratory

SMARTHEP Edge Machine Learning School
September 23, 2024

NexTGen
Next Generation Triggers

A Long History of Machine Learning



Perceptron

Rosenblatt 1958, 1960

Prompt: `Several giant wooly mammoths approach treading through a snowy meadow […]`

OpenAI Sora

A Long History of ML in High Energy Physics

NEURAL NETWORKS AND CELLULAR AUTOMATA
IN EXPERIMENTAL HIGH ENERGY PHYSICS

B. DENBY

Laboratoire de l'Accélérateur Linéaire, Orsay, France

Received 20 September 1987; in revised form 28 December 1987

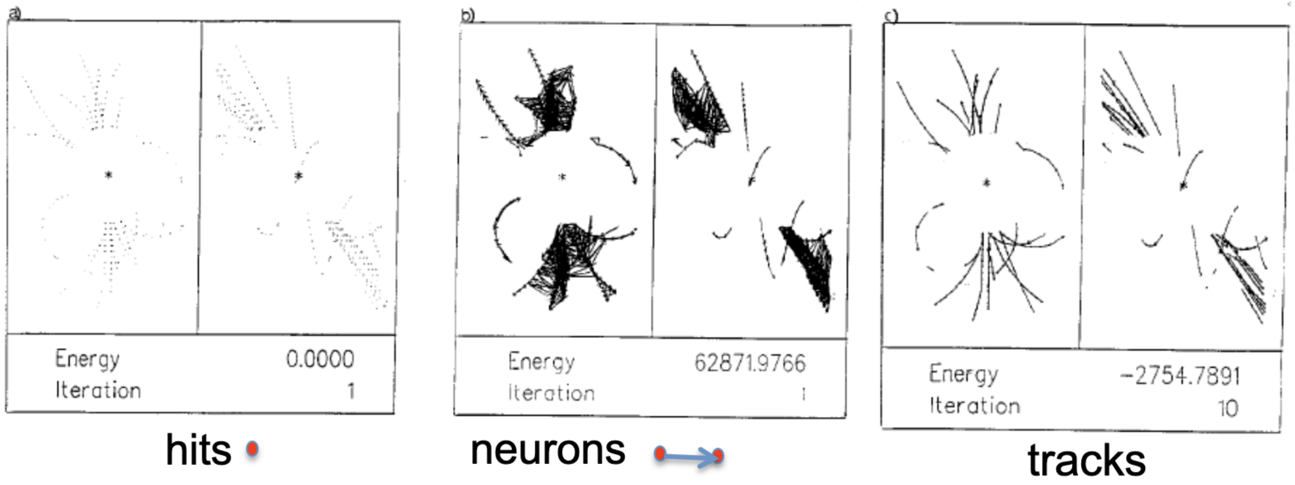


Figure 5. First try of neural track finding on simulated Delphi events. a) the hits. b) all neurons are initially possible. c) after settling.

How was it to work on AI in those days?

- The local LAL reaction was rather diferent
  - I got **FIRED** from the Delphi group
  - LAL directors agreed to let me stay at the lab anyway

From [talk]

- Bruce Denby

A Long History of ML in High Energy Physics

Nuclear Instruments and Methods in Physics Research A306 (1991) 459–466
North-Holland

Tagging the decays of the Z^0 boson into b quark pairs with a neural network classifier

C. Bortolotto, A. De Angelis and L. Lanceri
Istituto di Fisica dell'Università di Udine and INFN Trieste, Trieste, Italy

ML at LEP

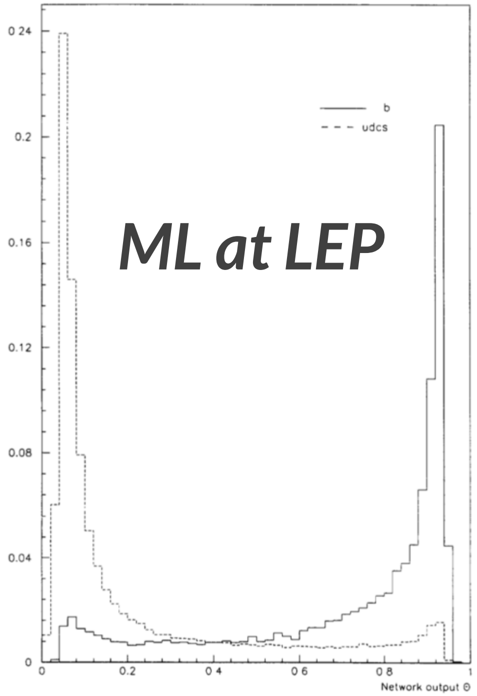


Fig. 2. Probability density function for the network output after full training, for $b\bar{b}$ (solid line) and non-$b\bar{b}$ (dashed line) events. $t_1 = 0.05$, $t_2 = 0.95$.

TAGGING B QUARK EVENTS IN ALEPH WITH NEURAL NETWORKS
(Comparison of different methods:
Neural Networks and Discriminant Analysis)

J. PRORIOL, J. JOUSSET, C. GUICHENEY
A. FALVARD, P. HENRARD, D. PALLIN, P. PERRET
Laboratoire de Physique Corpusculaire
de Clermont-Ferrand
IN2P3 - CNRS
Université Blaise Pascal
F-63177 AUBIERE CEDEX
FRANCE

B. BRANDL
Institut für Hochenenergiephysik
Universität Heidelberg
D-6900 HEIDELBERG
GERMANY

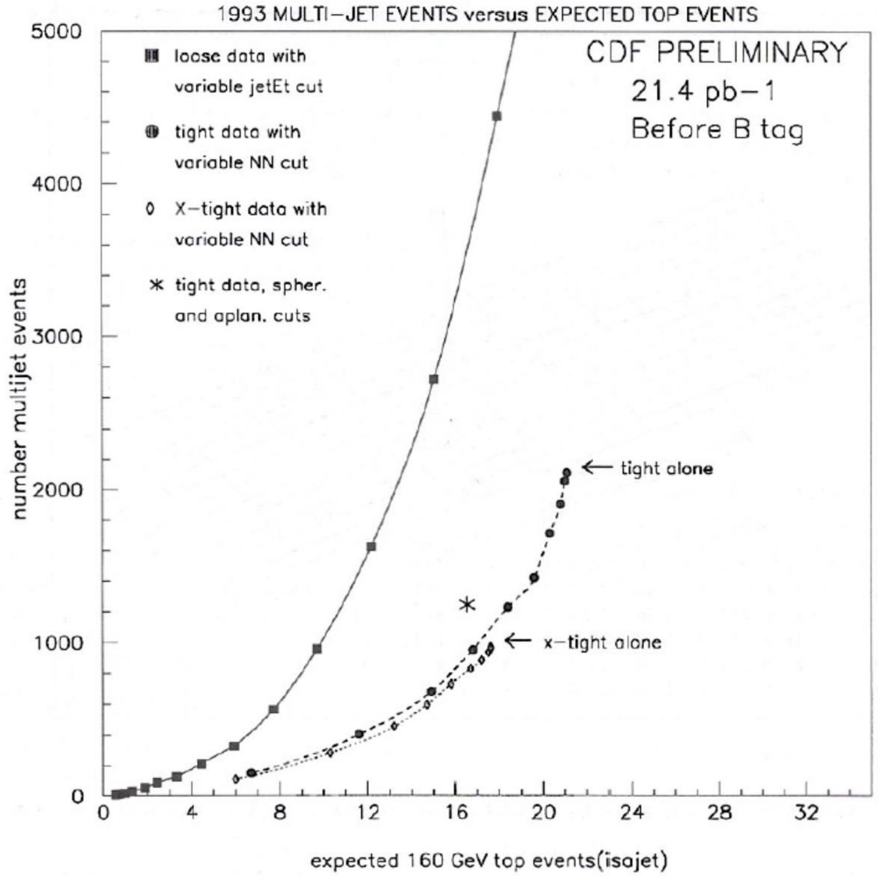
ABSTRACT

In this work we present the comparison of different methods to tag b quark events: multilayered perceptron, LVQ, discriminant analysis, combination of two methods. The sample events come from the ALEPH Monte Carlo and data.

WORKSHOP ON NEURAL NETWORKS
FROM BIOLOGY TO HIGH ENERGY PHYSICS
ELBA INTERNATIONAL PHYSICS CENTER
MARCIANA MARINA
JUNE 5–14, 1991

CDF Top Search in all-hadronic channel using Neural Nets!!

- Proton-Antiproton Collider Conference, Tsukuba , Japan, 18-22 October 1993



A Long History of ML in High Energy Physics

2012

The New York Times

Physicists Find Elusive Particle Seen as Key to Universe

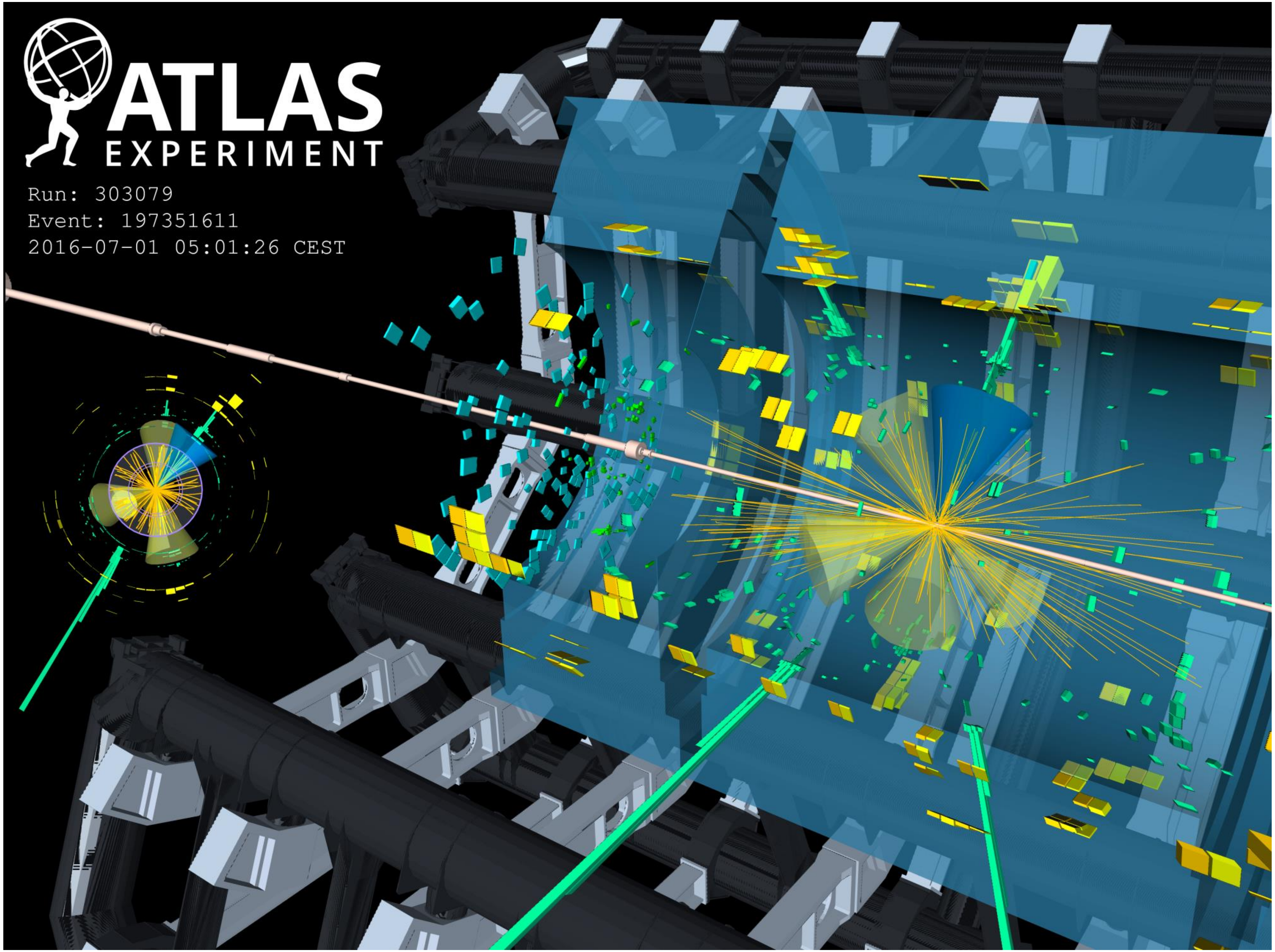
Table 1 | Effect of machine learning on the discovery and study of the Higgs boson

Analysis	Years of data collection	Sensitivity without machine learning	Sensitivity with machine learning	Ratio of P values	Additional data required
CMS[24] $H \rightarrow \gamma\gamma$	2011–2012	2.2σ, $P = 0.014$	2.7σ, $P = 0.0035$	4.0	51%
ATLAS[43] $H \rightarrow \tau^+\tau^-$	2011–2012	2.5σ, $P = 0.0062$	3.4σ, $P = 0.00034$	18	85%
ATLAS[99] $VH \rightarrow bb$	2011–2012	1.9σ, $P = 0.029$	2.5σ, $P = 0.0062$	4.7	73%
ATLAS[41] $VH \rightarrow bb$	2015–2016	2.8σ, $P = 0.0026$	3.0σ, $P = 0.00135$	1.9	15%
CMS[100] $VH \rightarrow bb$	2011–2012	1.4σ, $P = 0.081$	2.1σ, $P = 0.018$	4.5	125%

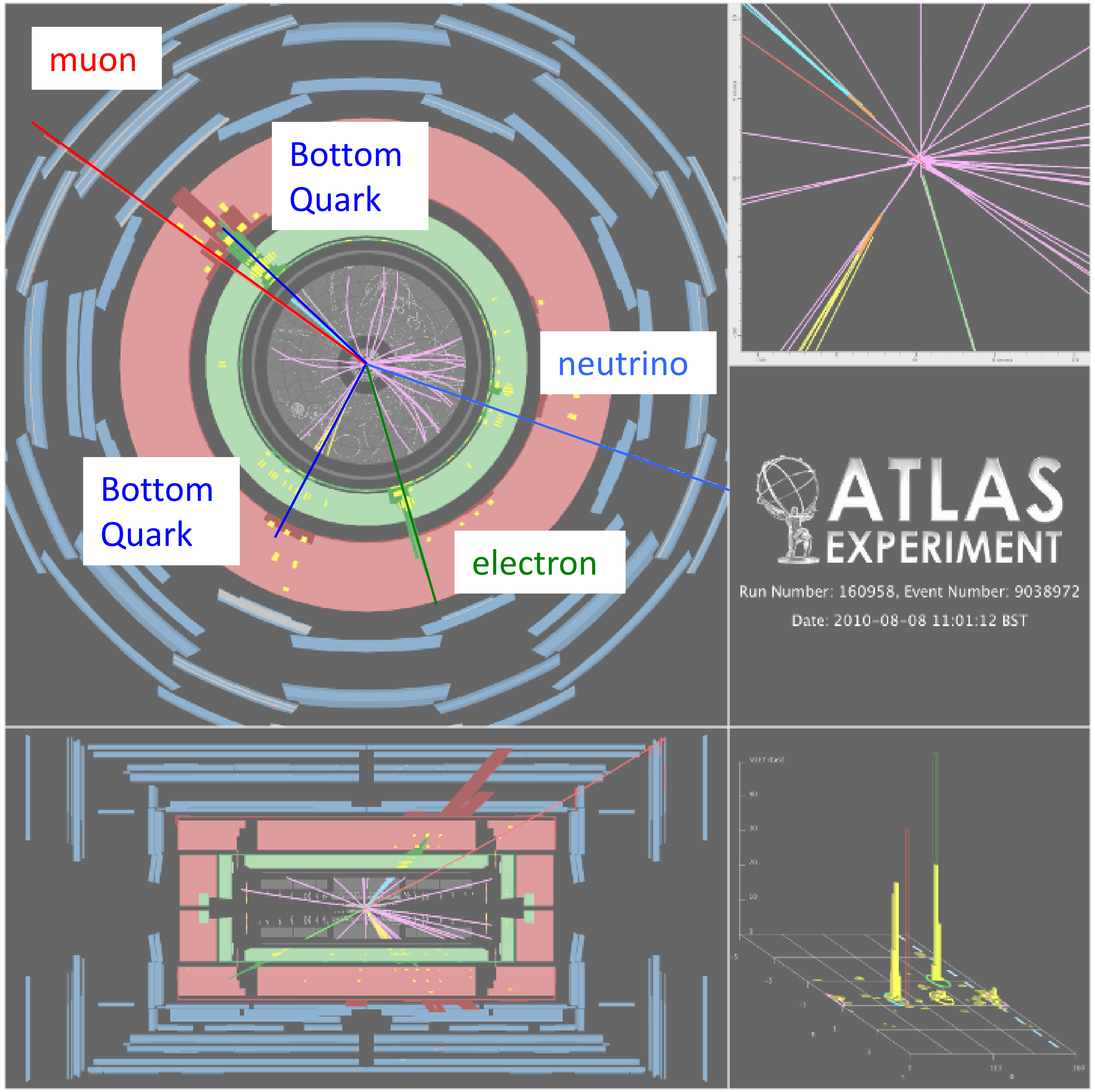
Radovic, Williams, Rousseau, **MK**, et al.
[Nature 560, 41–48 (2018)]

New directions in science are launched by new tools much more often than by new concepts. The effect of a concept-driven revolution is to explain old things in new ways. The effect of a tool-driven revolution is to discover new things that have to be explained.

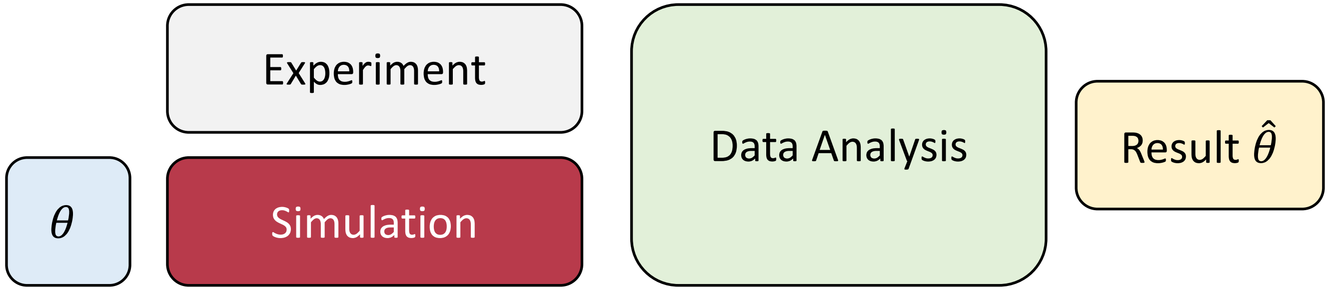
- Freeman Dyson


ATLAS
EXPERIMENT
Run: 303079
Event: 197351611
2016-07-01 05:01:26 CEST

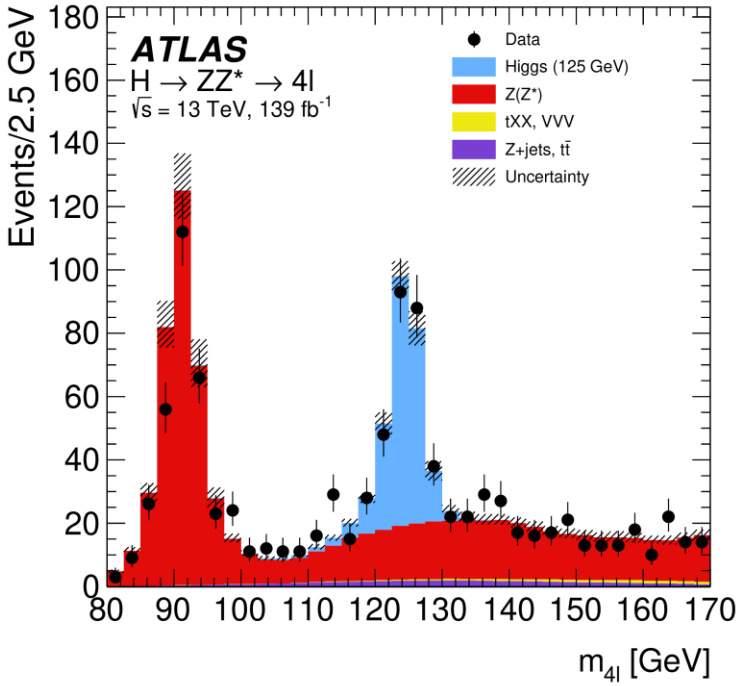
Studying Collisions



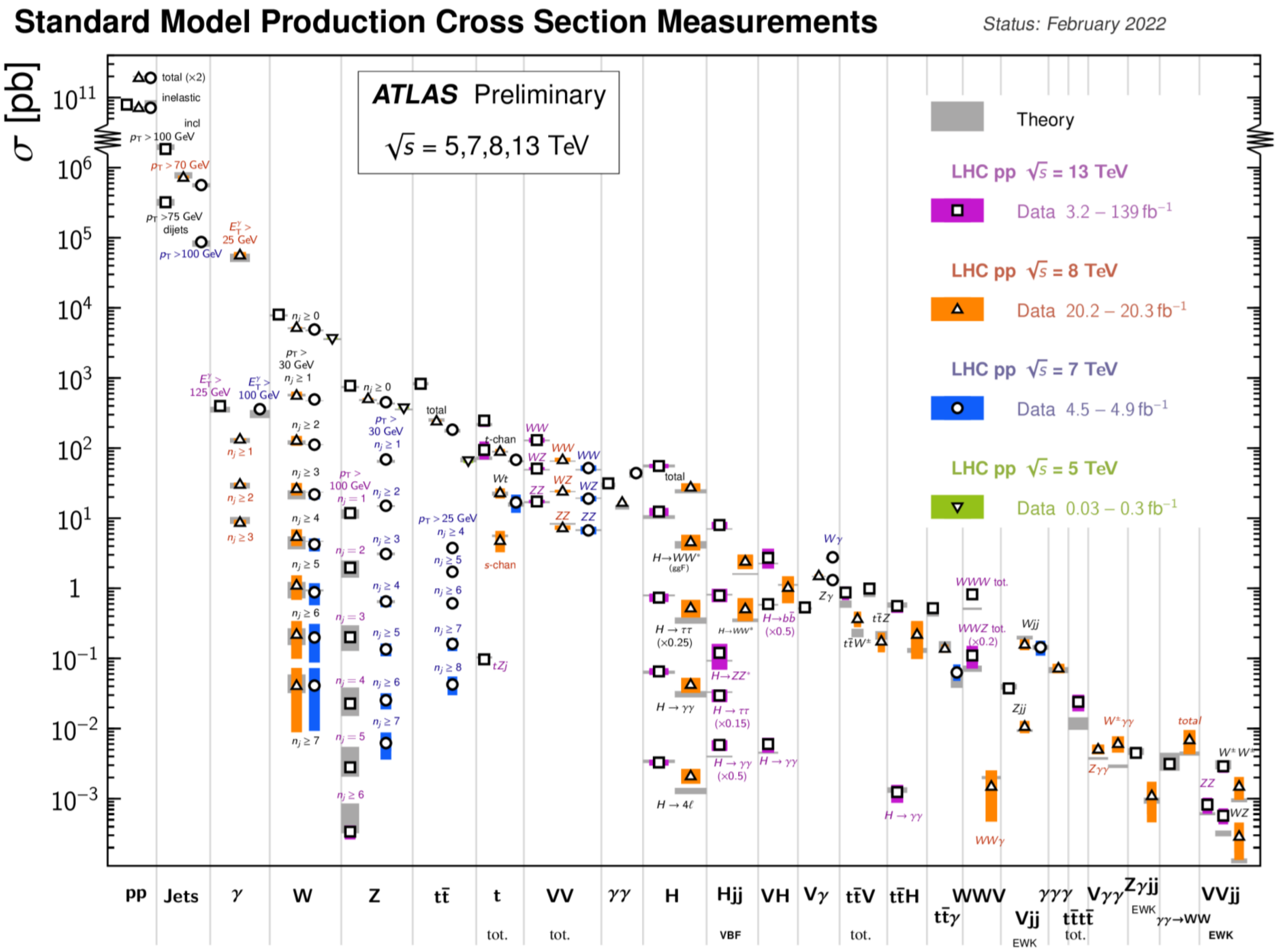
Data Analysis Workflow



EPJC 80 (2020) 942



This work very well!



ATL-PHYS-PUB-2022-009

What do we want from ML in High Energy Physics?

Complex Pattern Recognition

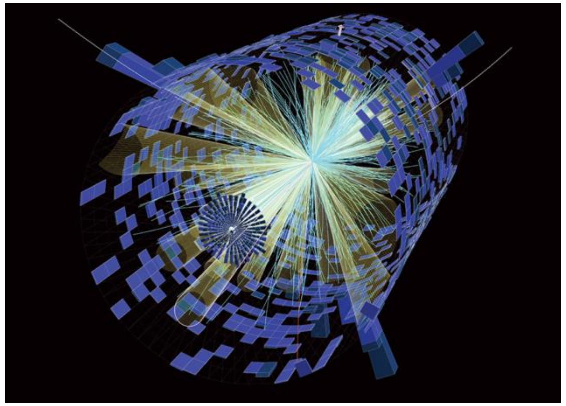
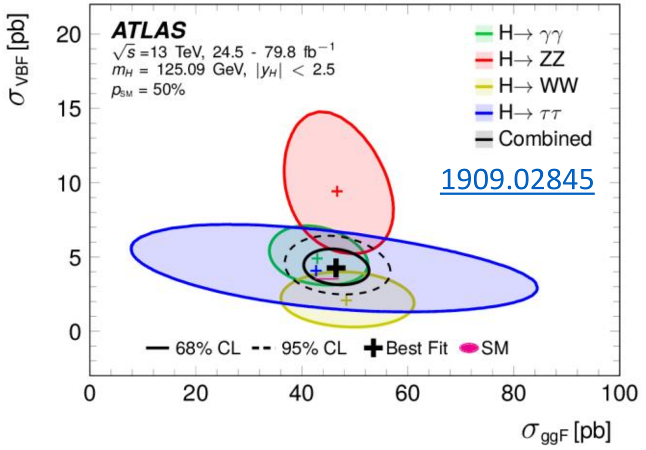
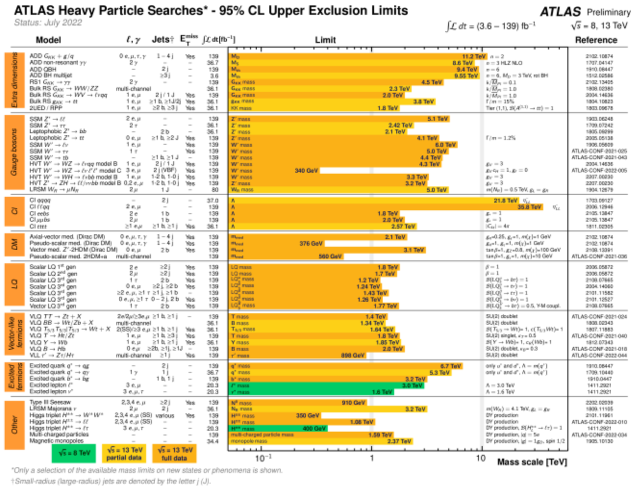


Image Credit: CMS Experiment

Improved Measurement Precision



Broaden New Physics Searches



ATL-PHYS-PUB-2022-034

Reduce Resource Demands

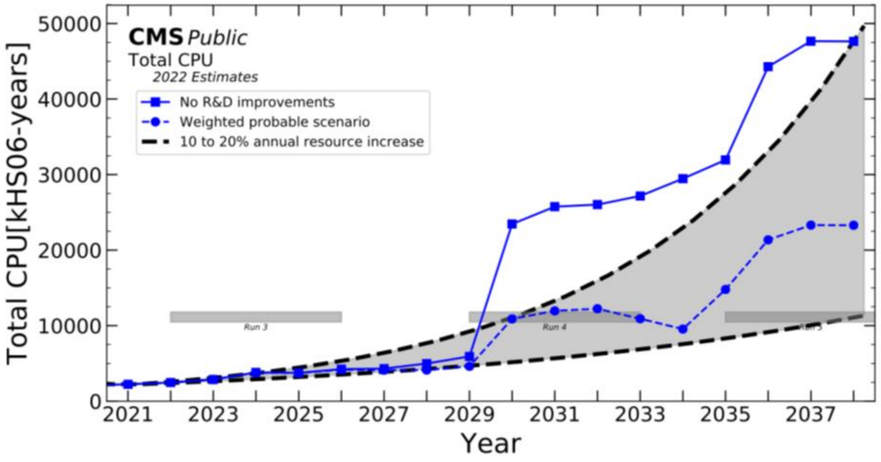


Figure Credit: CMS Experiment

Machine Learning Across Data Analysis

Experiment

θ

Simulation

Data Analysis

Result $\hat{\theta}$

Data Acquisition

Low-Latency Real-Time Decision Making

Simulation

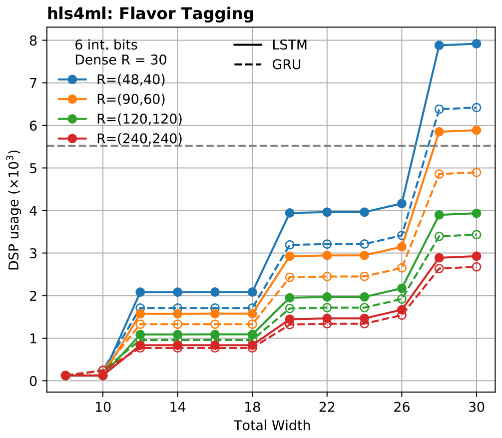
Fast Simulation and Sampling

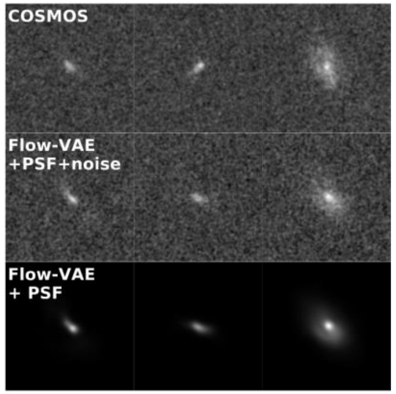
Reconstruction

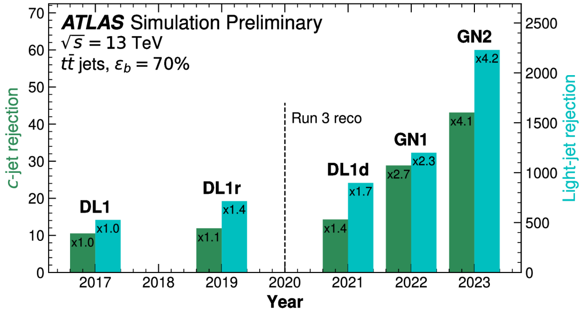
Particle / Event Pattern Rec.
Anomaly Detection

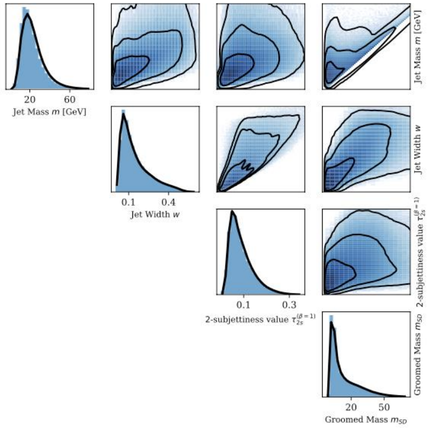
Inference

Parameter Estimation
Unfolding









Khoda, …, **MK**, et. al, MLST 2023

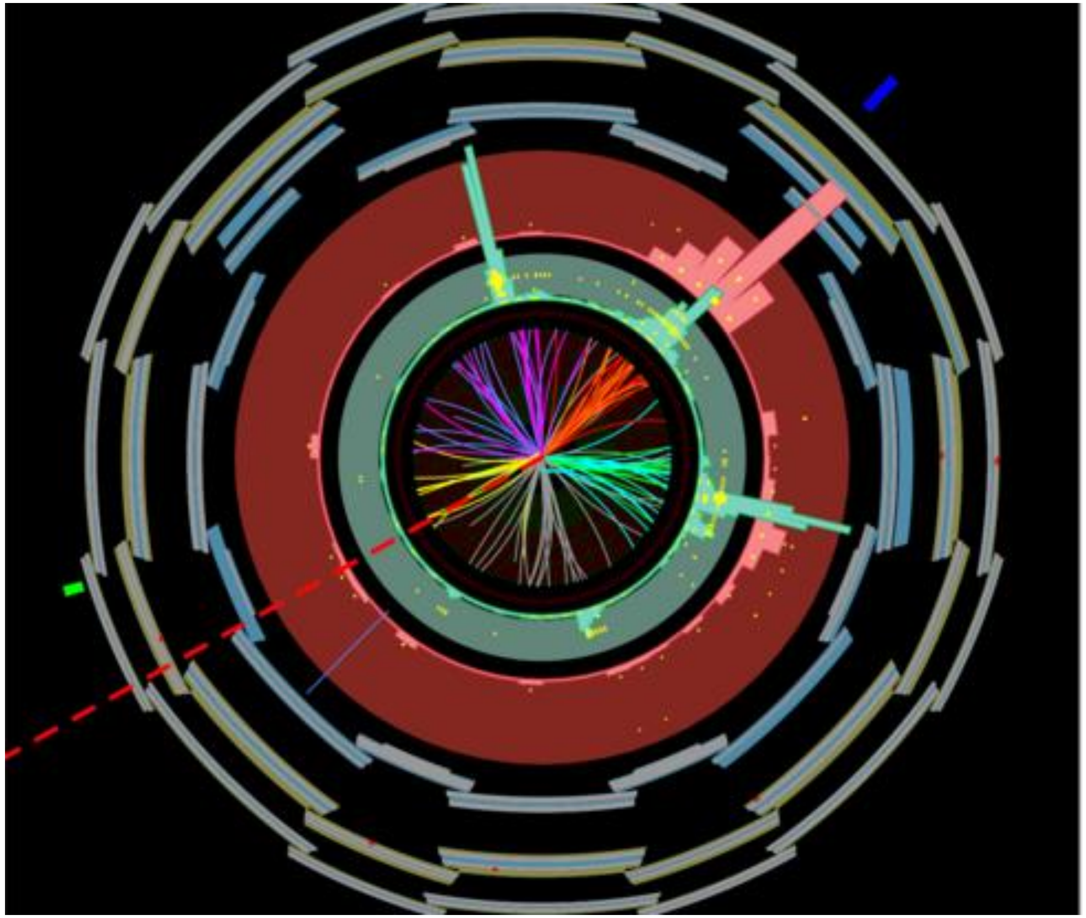
2008.03833

ATLAS-PLOTS-FTAG-2023-01

Vandegar, **MK**, et al. AISTATS 2021

What kinds of ML models are used?

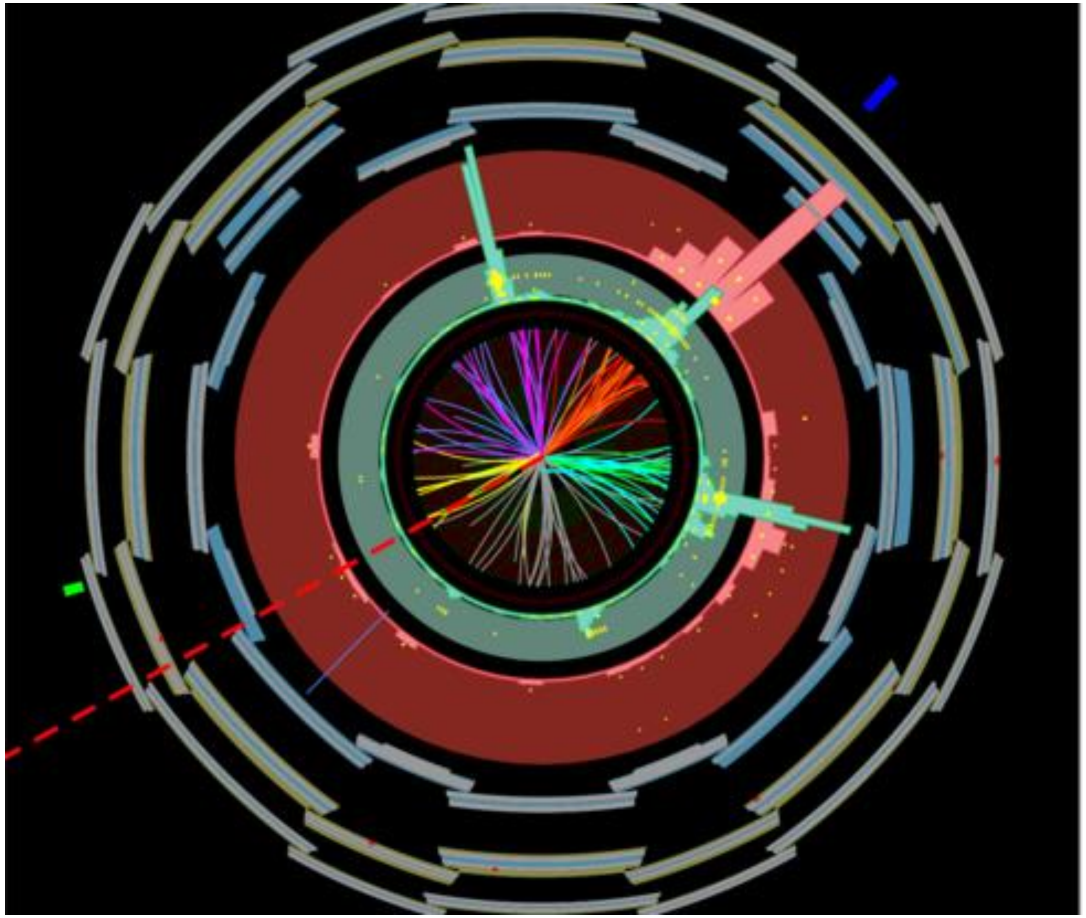
Reconstruction



Reconstruction

Pattern Recognition in
Sparse high dimensional data
Irregular detector geometry

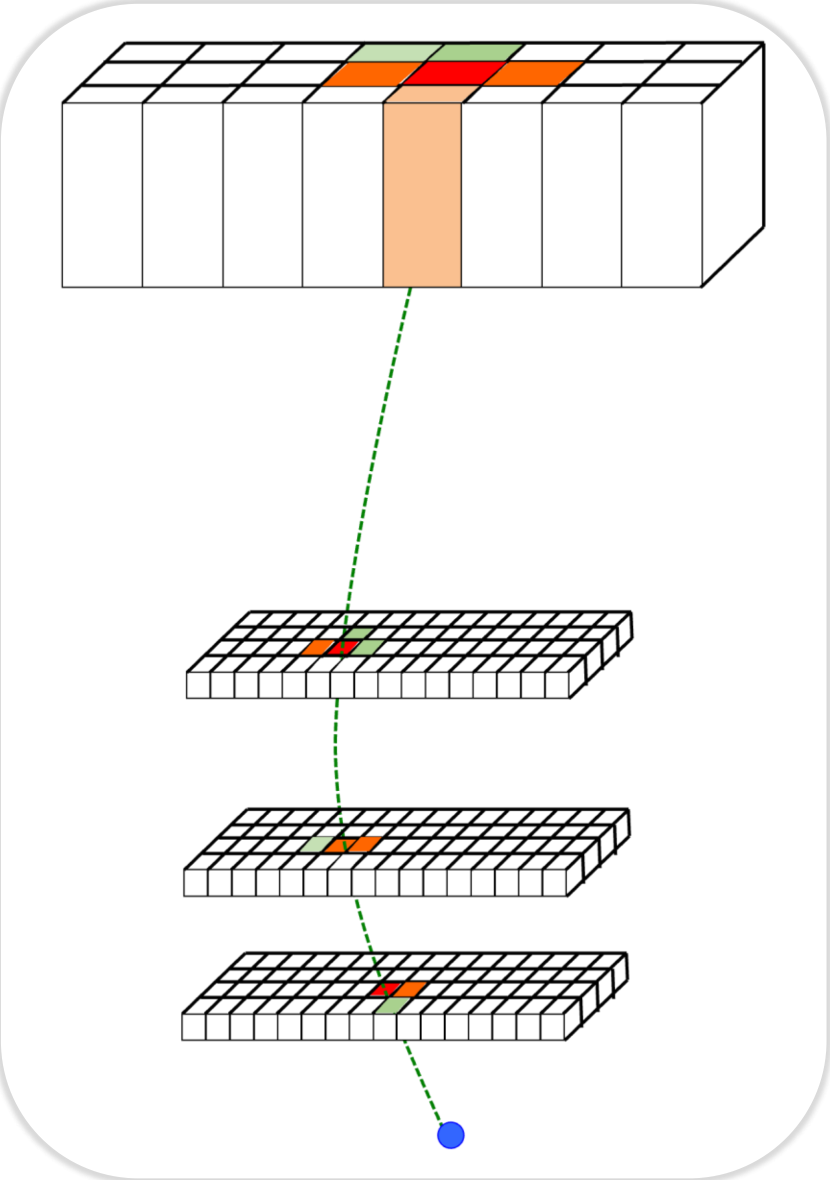
Goal:
Turn low-level data
(i.e. measurements of energy deposition)
into estimates of particle energy,
momentum, direction, trajectory, …



Reconstruction

Pattern Recognition in
Sparse high dimensional data
Irregular detector geometry

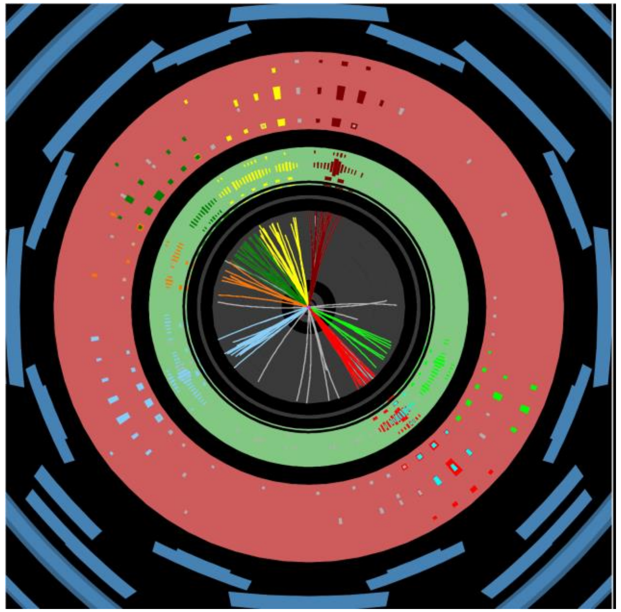
Goal:
Turn low-level data
(i.e. measurements of energy deposition)
into estimates of particle energy,
momentum, direction, trajectory, …

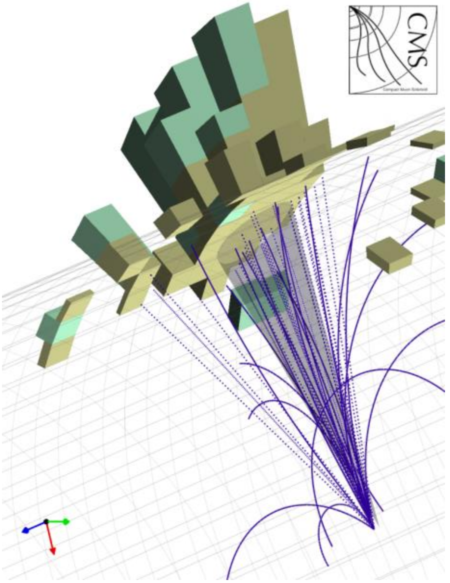


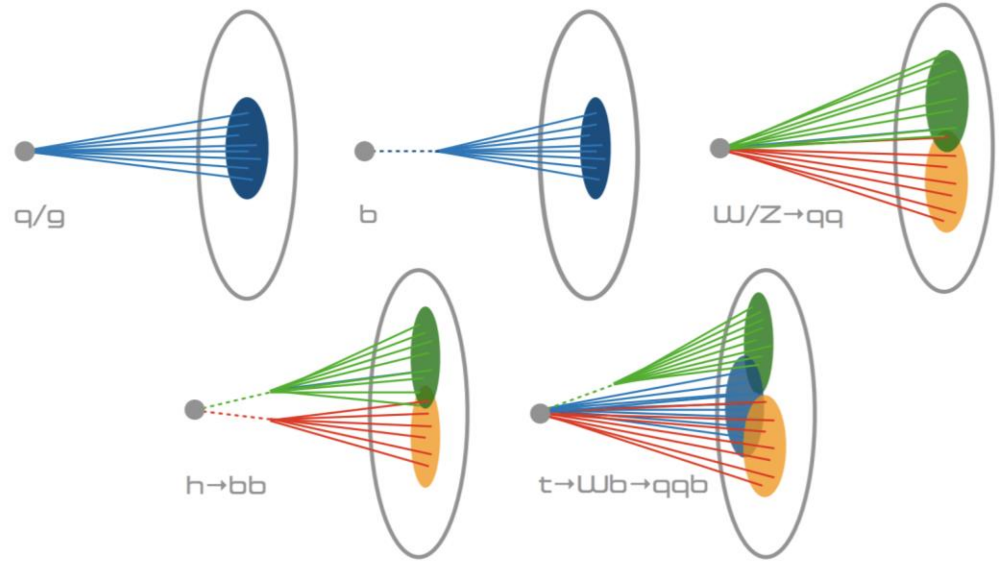
Reconstruction

Pattern Recognition in
Sparse high dimensional data
Irregular detector geometry

Goal:
Turn low-level data
(i.e. measurements of energy deposition)
into estimates of particle energy,
momentum, direction, trajectory, …

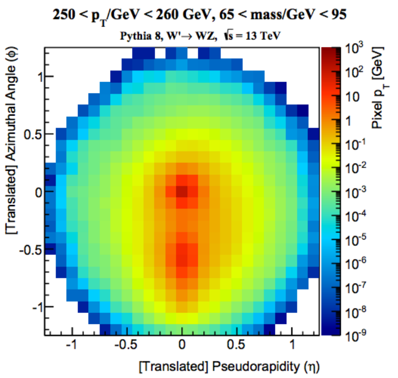






What kind of data is used in Reconstruction?

Images



De Oliviera, **MK**, et al. [1511.05190](https://arxiv.org/abs/1511.05190)

Data on regular grid,

translation equivariance on grid

Sequences

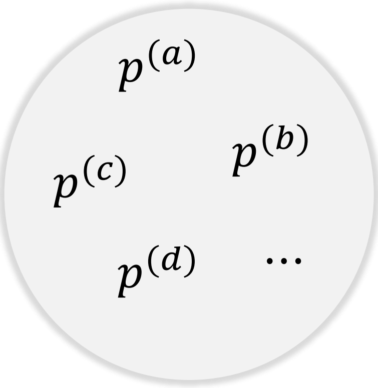
$$\begin{bmatrix} p_x^{(1)} \\ p_y^{(1)} \\ p_z^{(1)} \\ \vdots \end{bmatrix} \begin{bmatrix} p_x^{(2)} \\ p_y^{(2)} \\ p_z^{(2)} \\ \vdots \end{bmatrix} \cdots \begin{bmatrix} p_x^{(n)} \\ p_y^{(n)} \\ p_z^{(n)} \\ \vdots \end{bmatrix}$$

ordered →

Data ordered as sequence,

translation equivariance in "time"

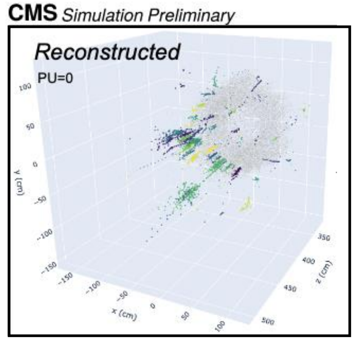
Sets



Data unordered list

permutation equivariance

Graphs & Point Clouds



[2203.01189](https://arxiv.org/abs/2203.01189)

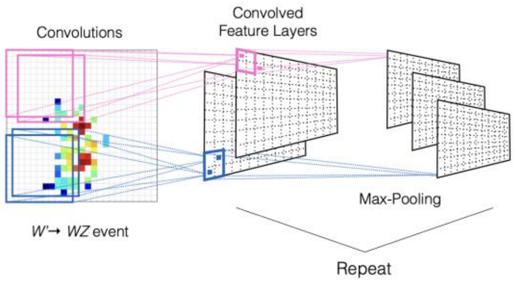
Data distributed in "space"

permutation equivariance

geometric relations

What kind of ML models work best for Reconstruction?

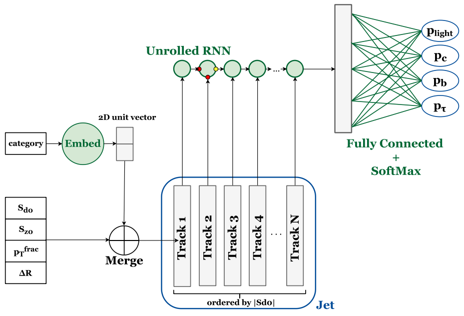
Images



De Oliviera, **MK**, et al. 1511.05190

Convolutional NN

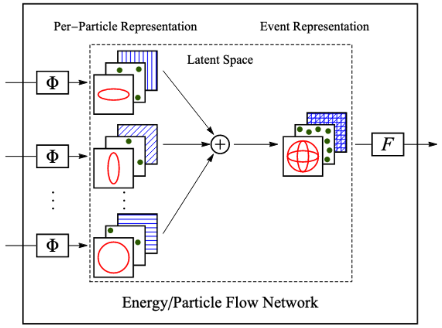
Sequences



ATL-PHYS-PUB-2017-003

Recurrent NN

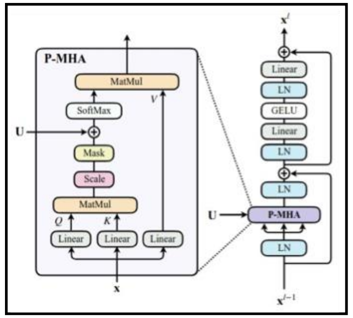
Sets



1810.05165

Deep Sets

Graphs & Point Clouds



2202.03772

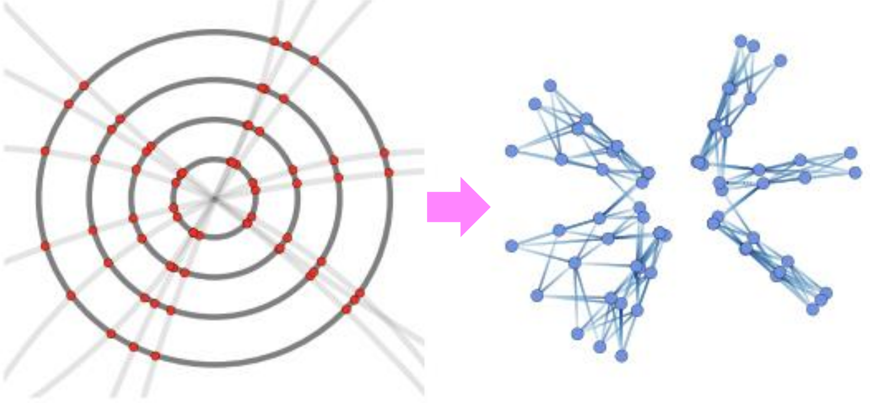
Graphs NNs & Transformers

…

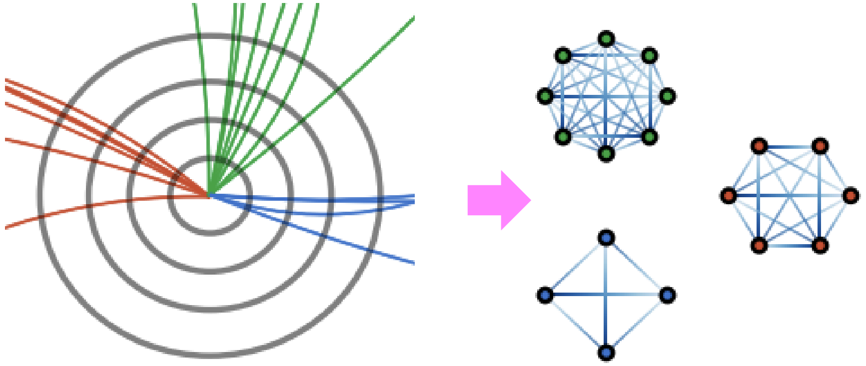
Graph Neural Networks and Transformers

Good fit for sparsity, irregular geometry, and variable cardinality of HEP data

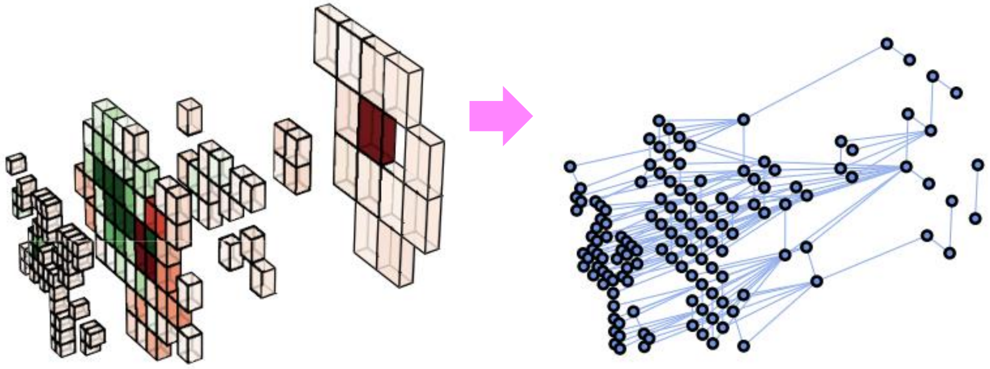
Tracking: Finding Trajectories from space-points



Classifying Jets (streams of particles)



Calorimeter cluster analysis



Classifying Events

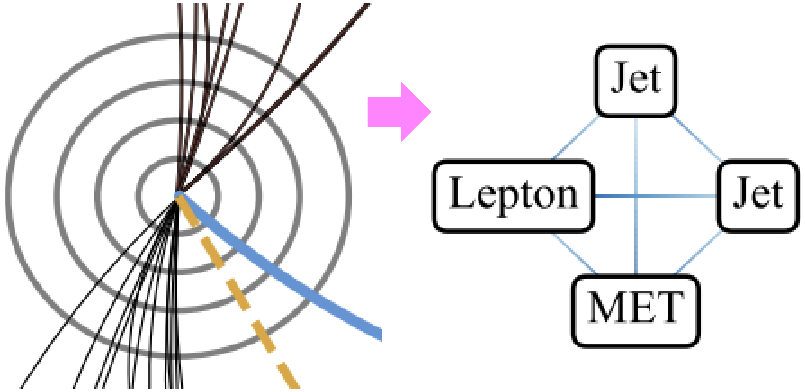
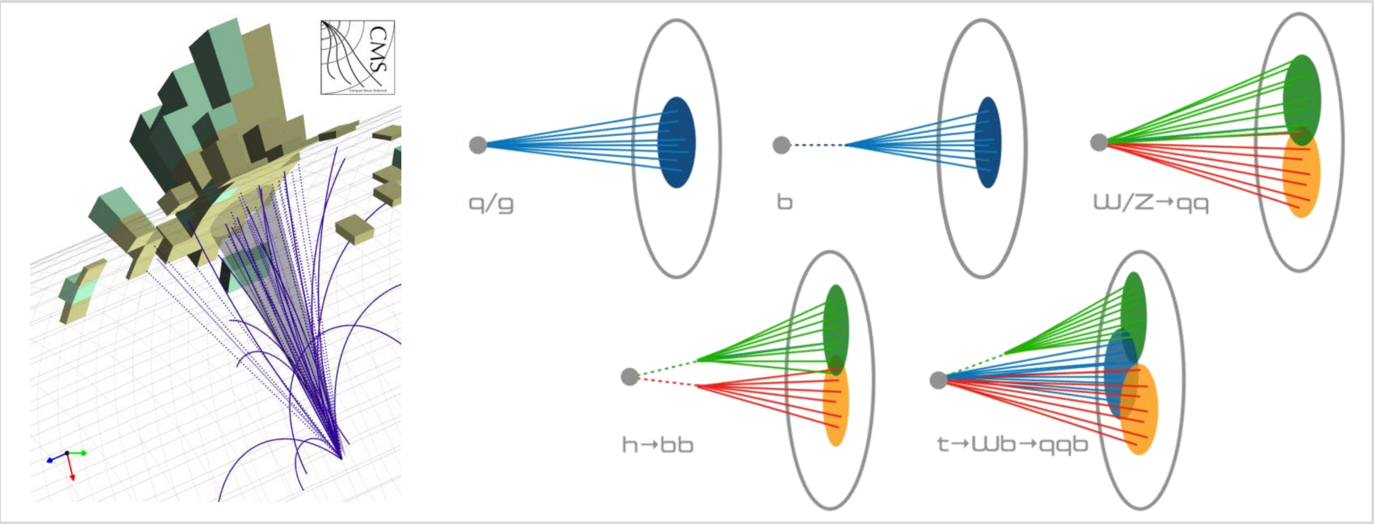
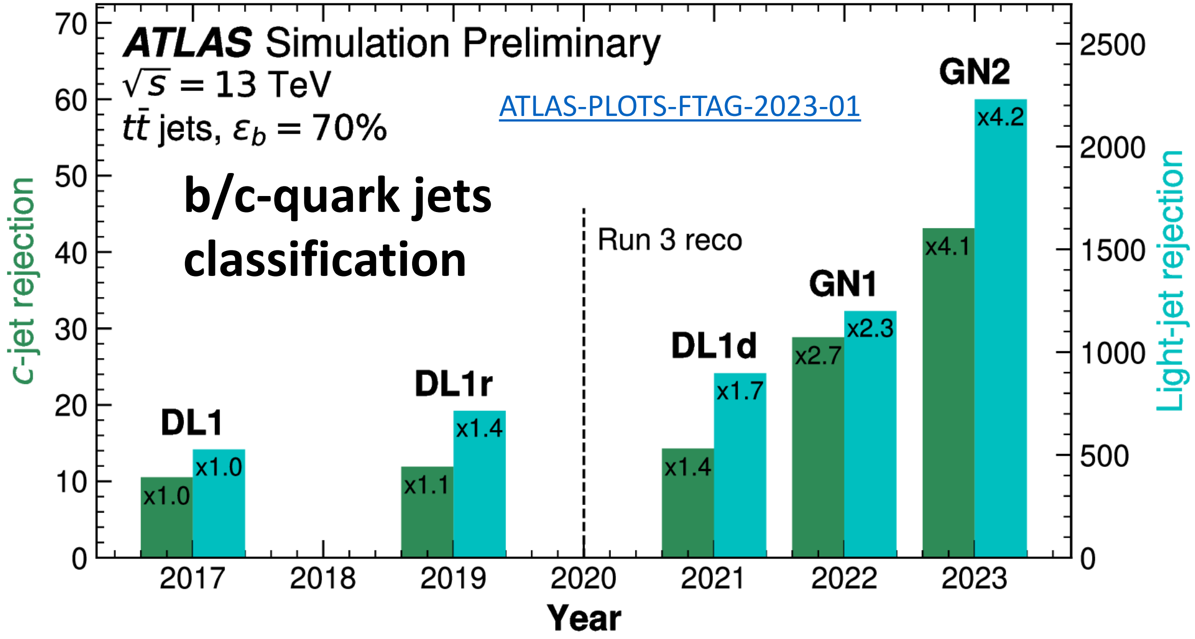
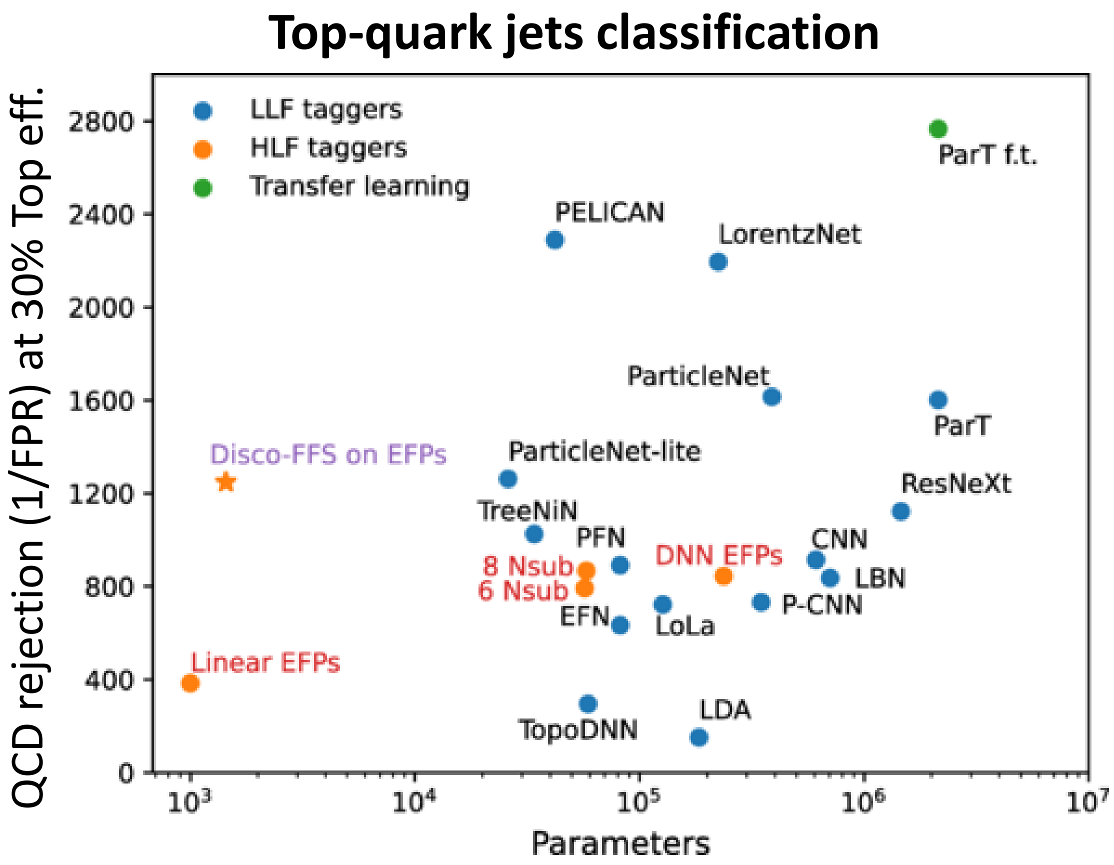


Image credit: MLST **2** 021001

Evolution of models for jet classification





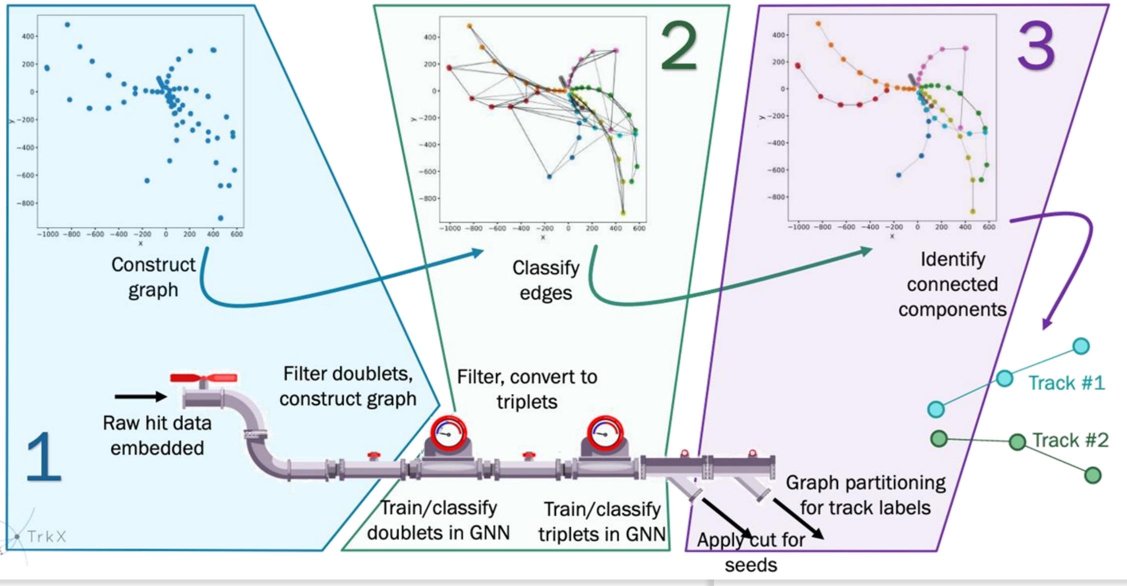


Phys. Rev. D 109, 054009

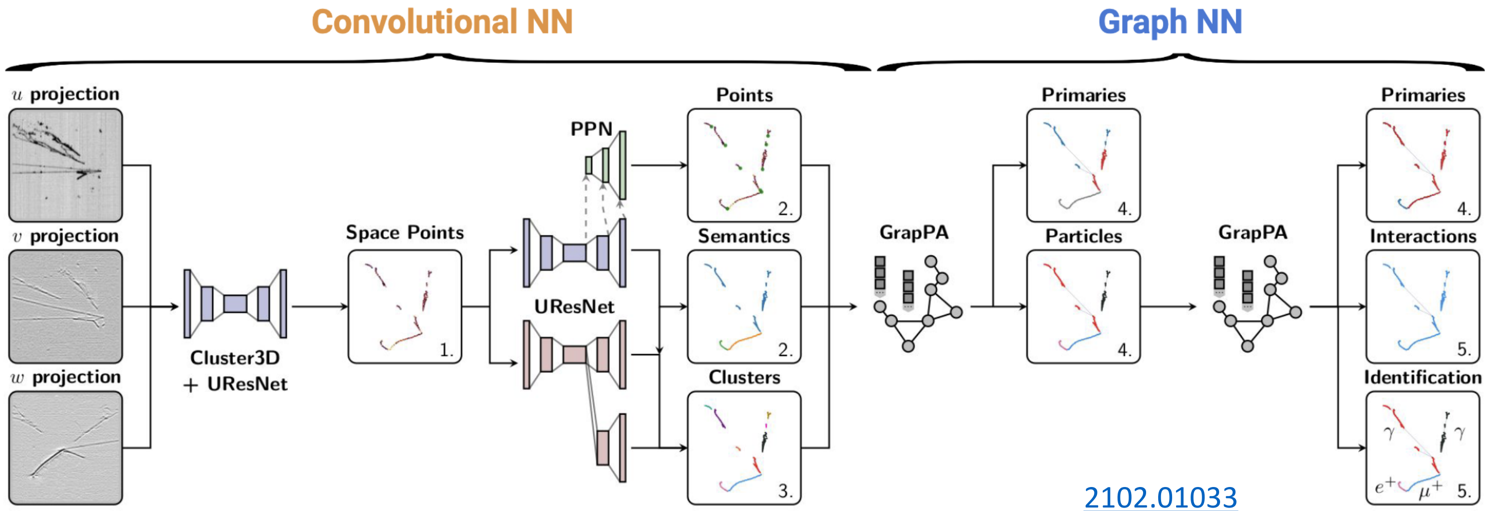
Bigger, more complex, multi-component ML Pipelines

Graph-based Tracking

Figure Credit: D. Murmane

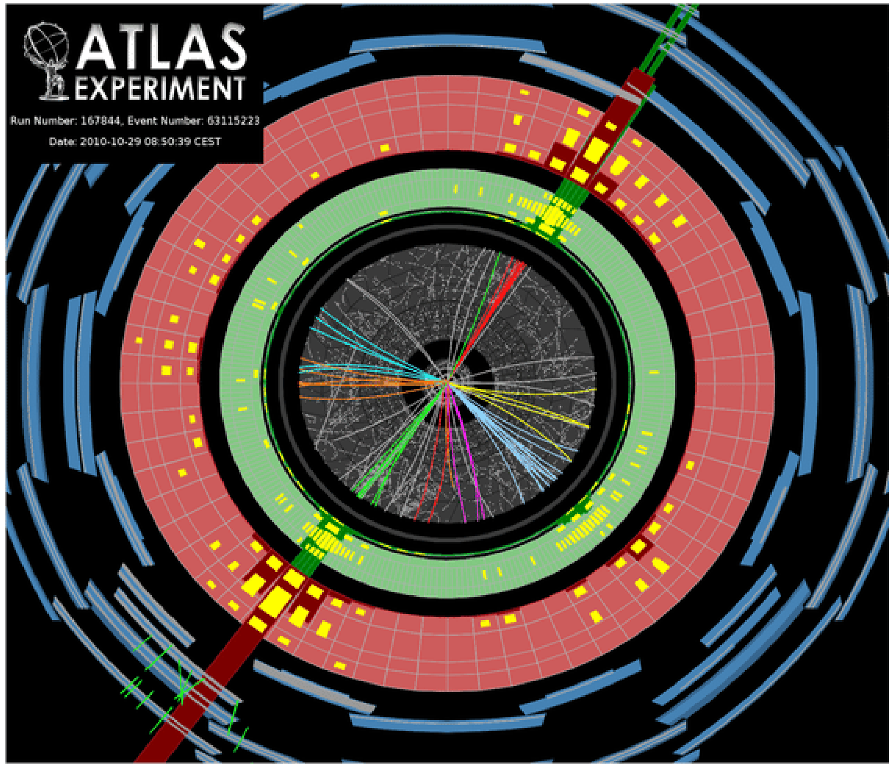


Neutrino End-to-End Reconstruction



2102.01033

Event Classification



→ What kind of interaction event happened in the collision?

Event Classification

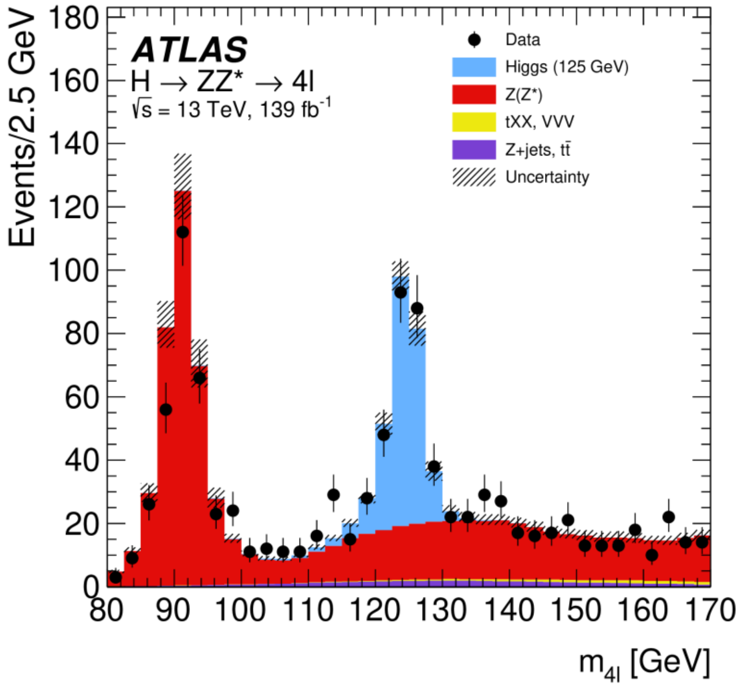
Given a set of events where we have reconstructed the particles:

Past: Think hard about good variables, for data selection & statistical inference

This is "*Tabular data*"

- Features engineered by physicists

EPJC 80 (2020) 942



Event Classification

Given a set of events where we have reconstructed the particles:

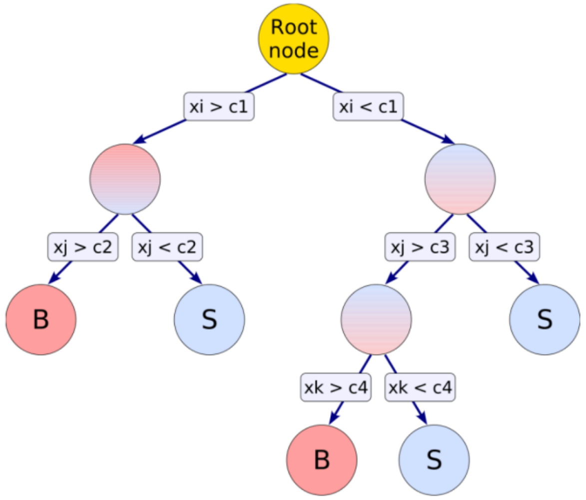
Past: Think hard about good variables, for data selection & statistical inference

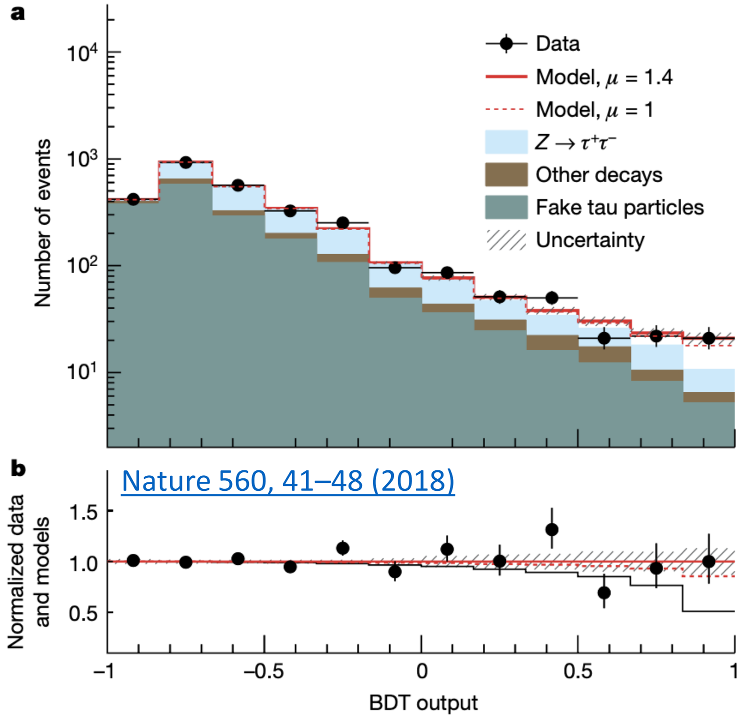
This is "*Tabular data*"

- Features engineered by physicists

Combine many variables in MVA?

Decision Tree based models tend to work very well for tabular tasks





[Nature 560, 41–48 (2018)]

Event Classification

Instead of ML on tabular features…

Can use set of particles and their features

“Lower-level” than engineered features

- A particle has meaning when considered in relation to other particles in event

Neural networks used more and more, especially graph & transformer models

- May need to deal with gemoetric relationships, variable length inputs, …

Low-level processing is what NNs good at

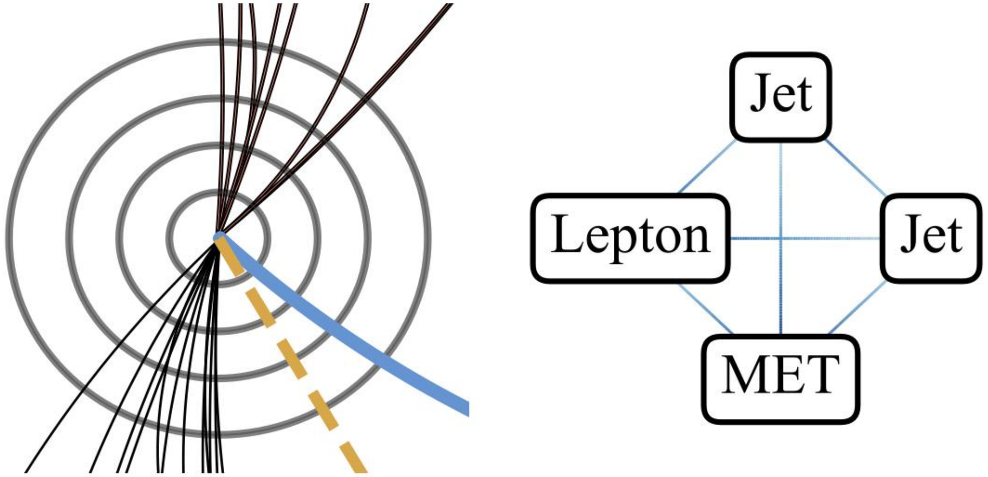


Image: [MLST 2 021001 (2021)]

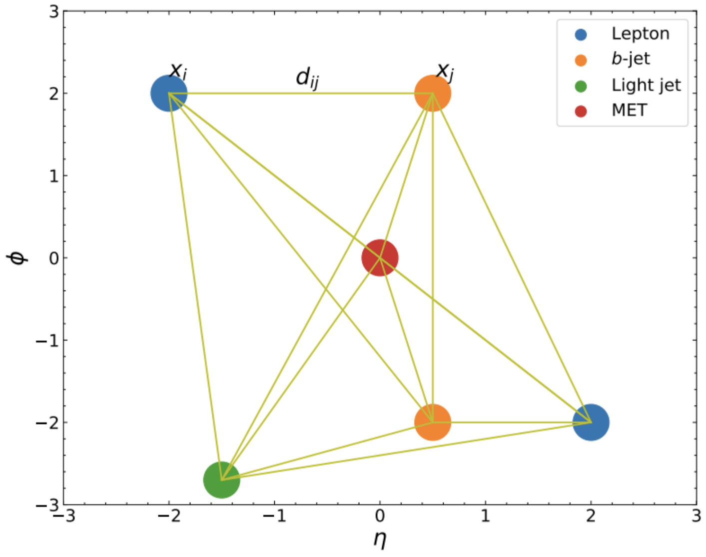
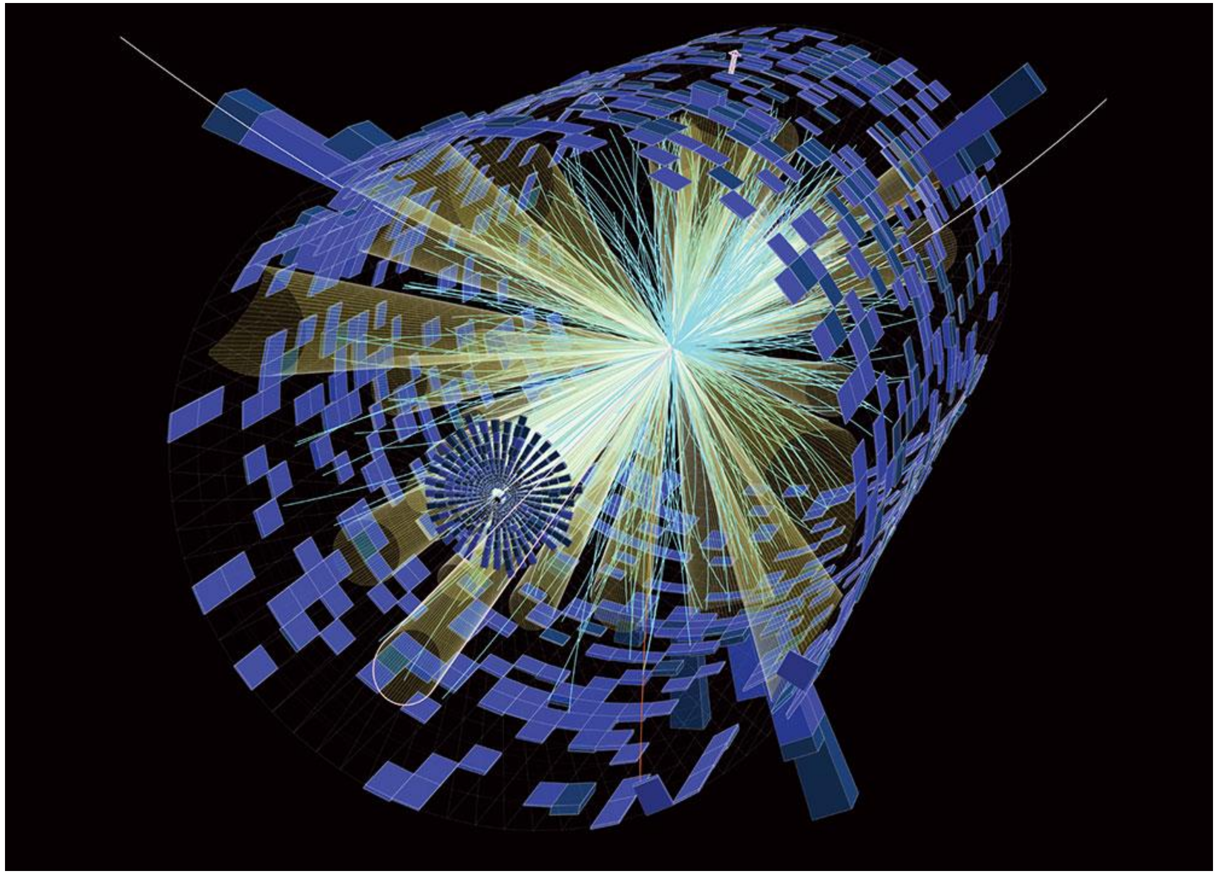
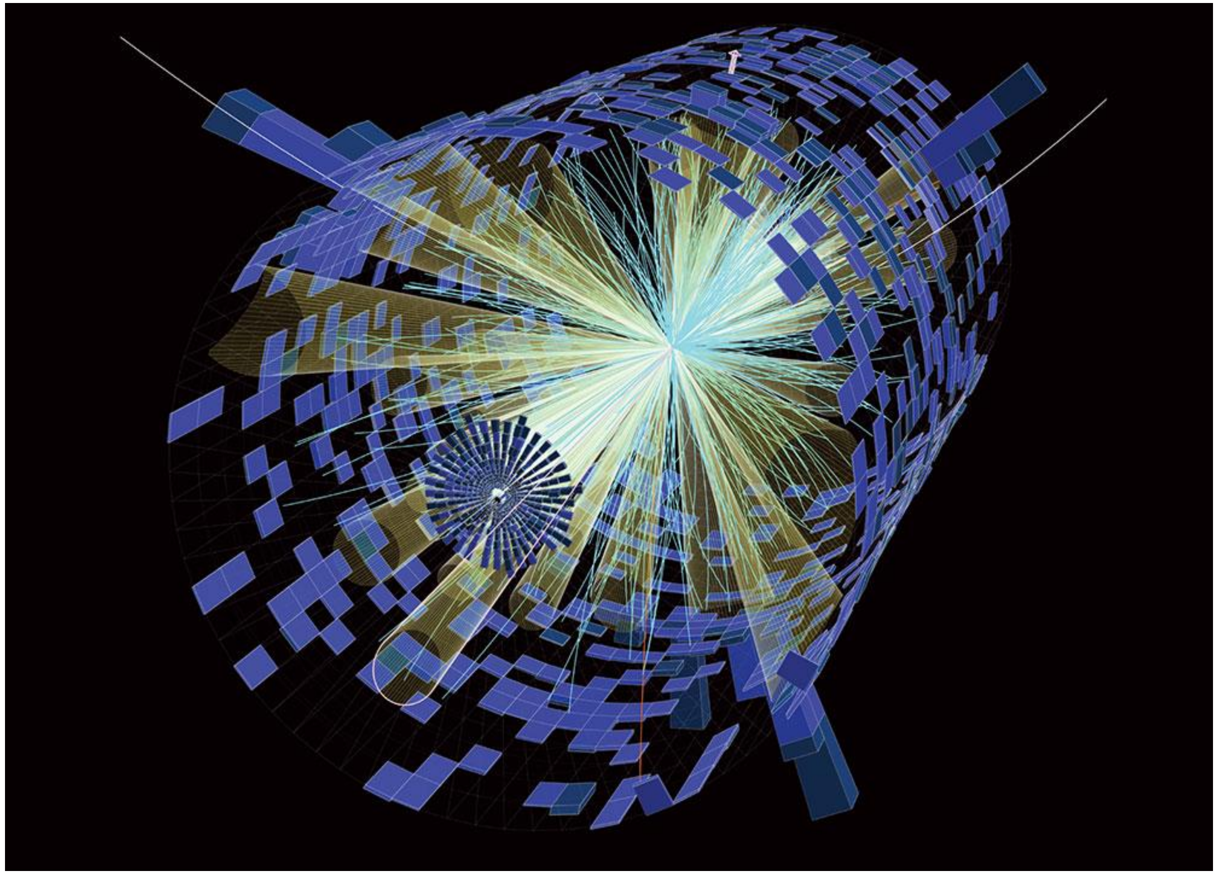


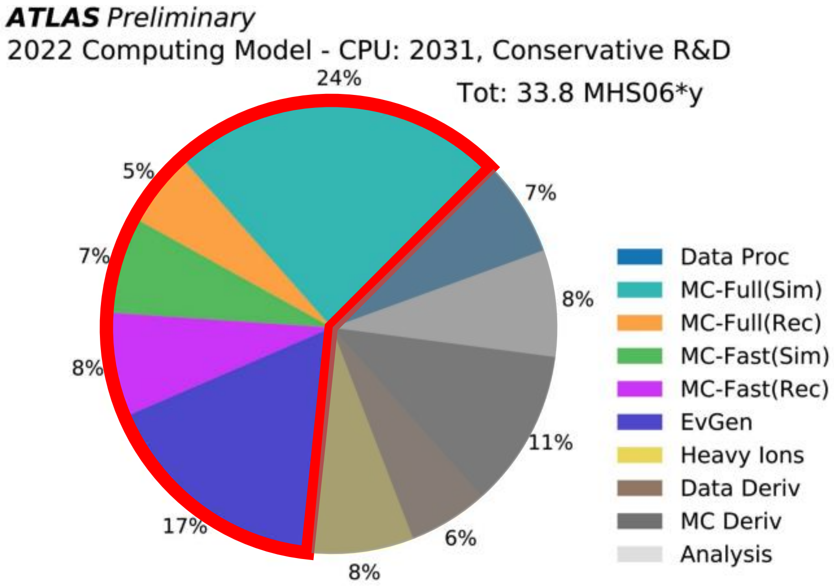
Image: [PRD 104, 056003 (2021)]

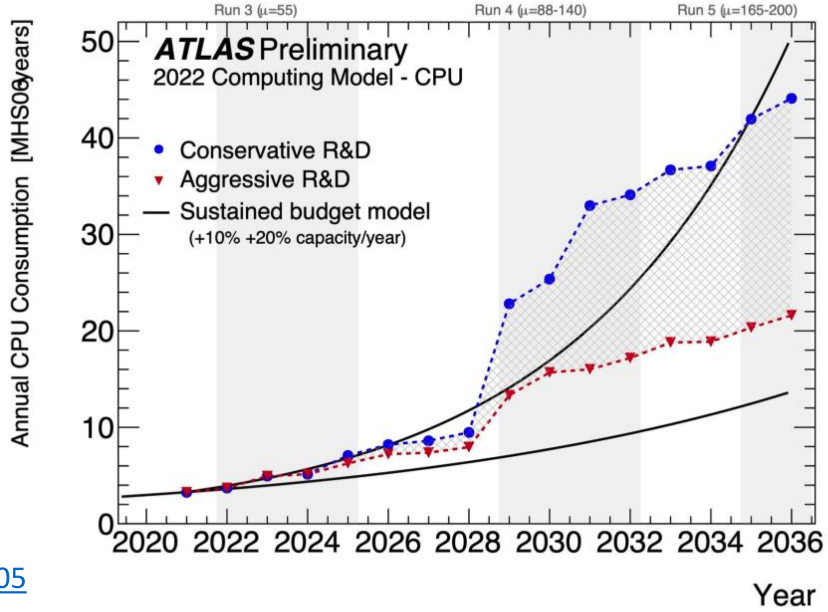
Simulation



Simulation Challenges







Computing Budget

[LHCC-2022-005]

ML for Simulation

Deep Generative Models

Learn to create "plausible" data by transforming random noise

Model structure depends on training method

Architecture choice depends on data type,

- Just like reconstruction

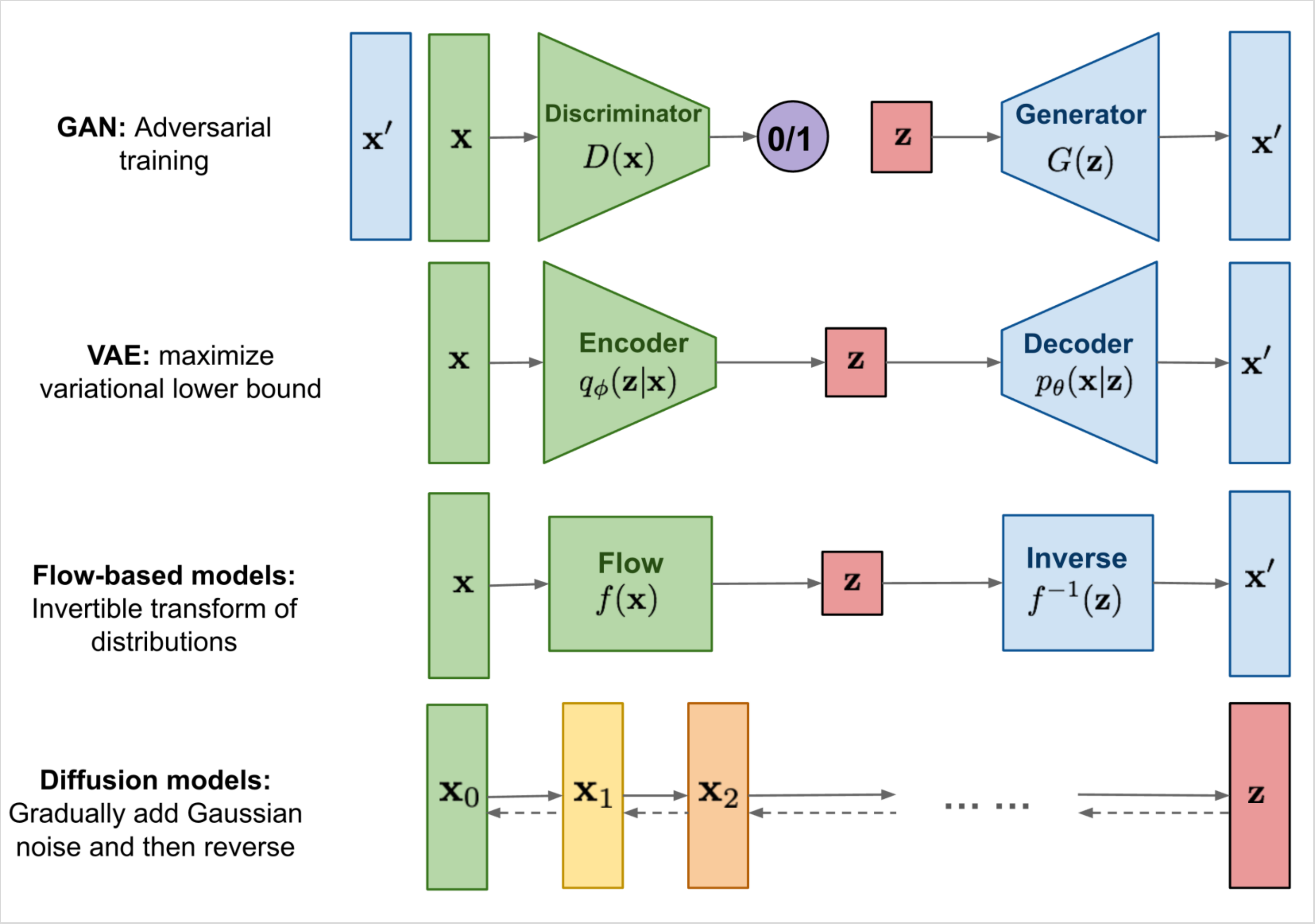


Image credit: Lilian Weng

Deep Generative Models For Speeding Up Simulation

Detector Simulation

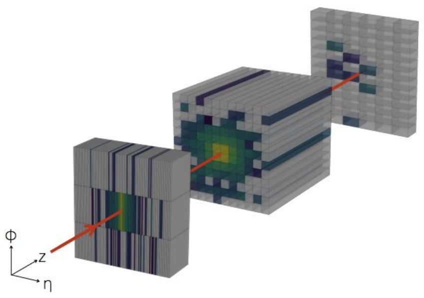
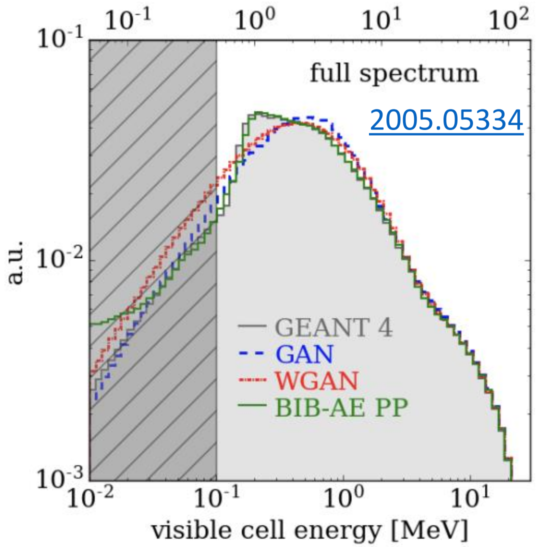
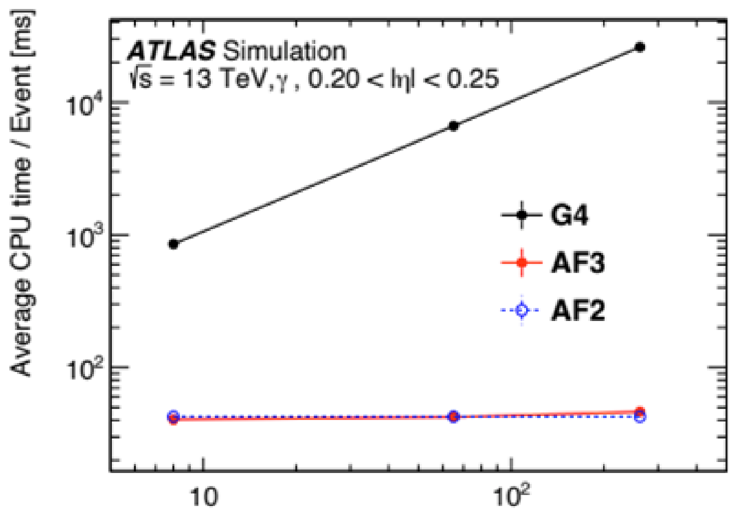


Image credit: [1705.02355]

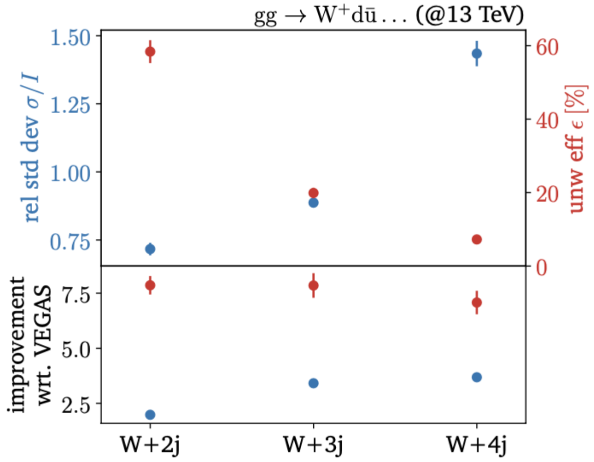




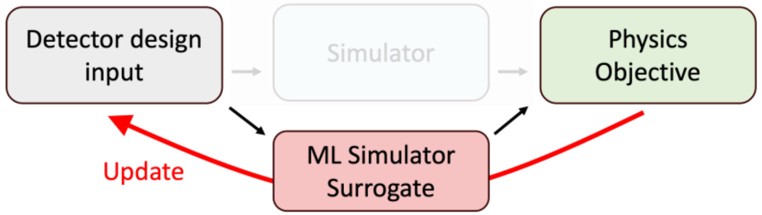
[2109.02551]

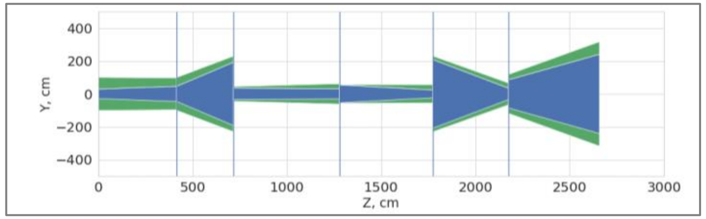
Matrix Elements & Event Generators

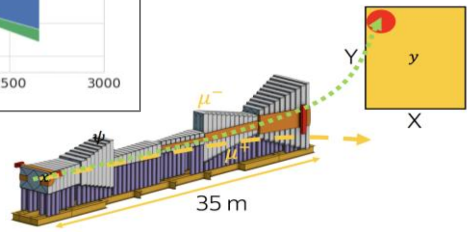
[2311.01548]
[2203.11110]
Heinrich, **MK**, [2203.00057]



Detector Design Optimization







Shirobokov, **MK**, et al.
[NeurIPS **33**, 14650-14662 (2020]

Conclusion

Long history of ML in HEP, and the recent ML advancements have made major impacts on HEP

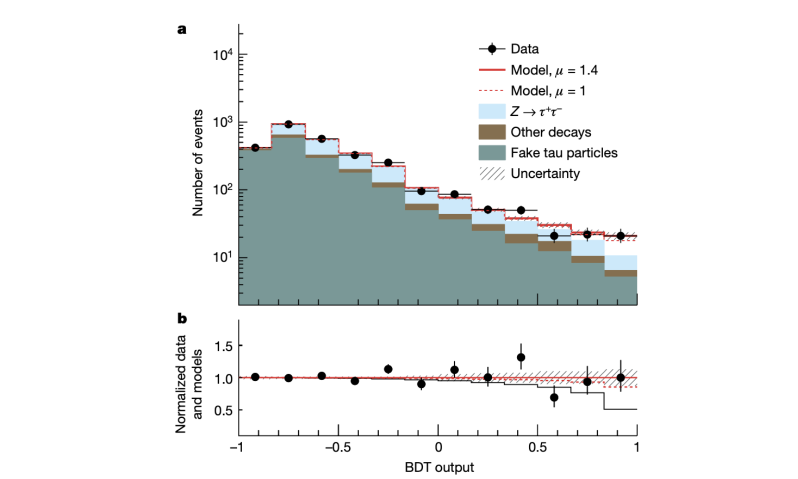
Complexity of HEP data warrants careful consideration about how are where to apply ML

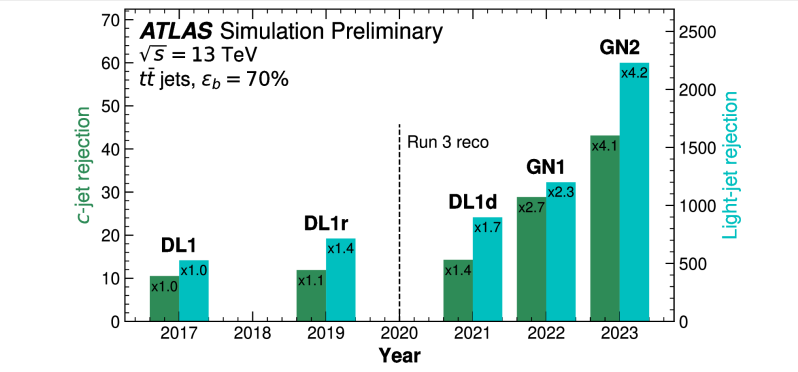
- What kind of data? How much data? How to frame the task of interest?

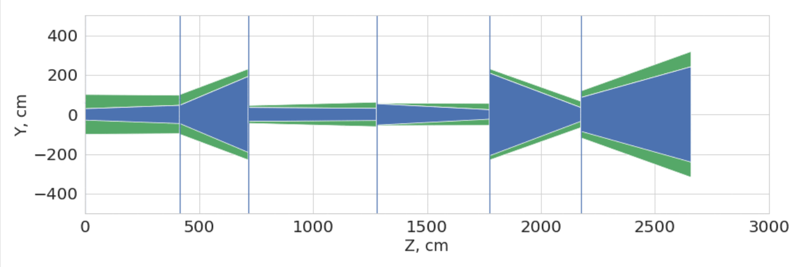
Can build sophisticated systems to approach complex challenges and address completely new questions!

But… for low-latency models, hardware constraints→architecture design constraints

- Lots more information from the rest of the talks!



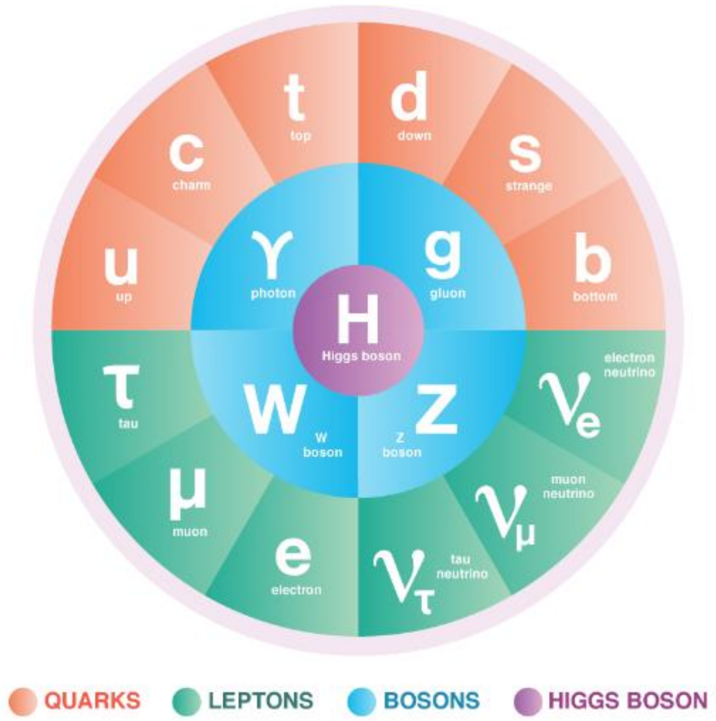


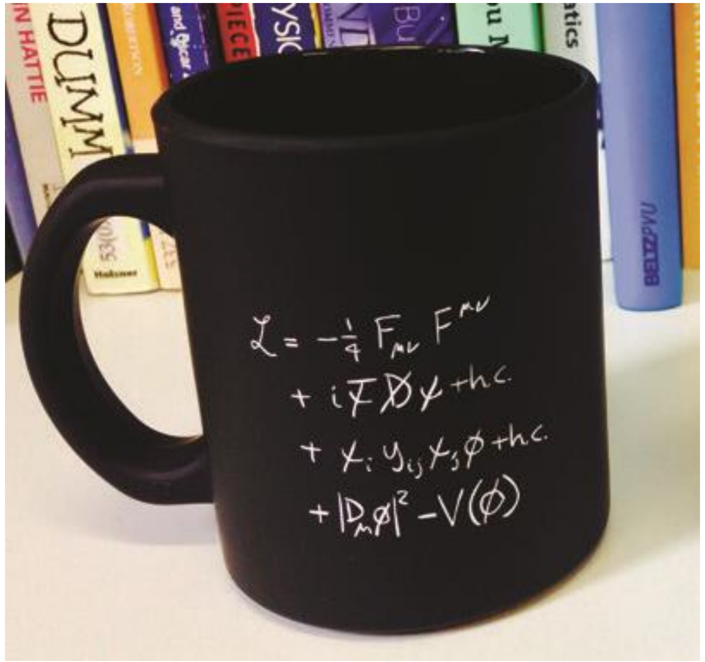


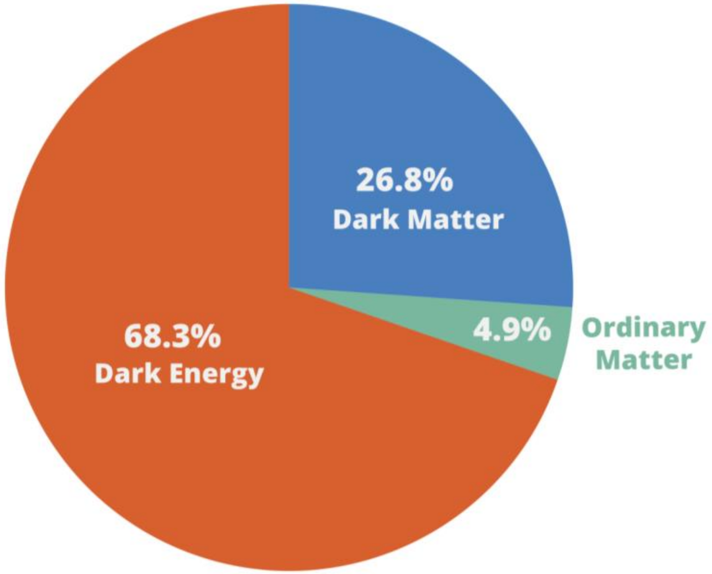
Backup

High Energy Physics – What We Know

Image source: Symmetry Magazine







High Energy Physics – Big Questions

Why is the Higgs so light?

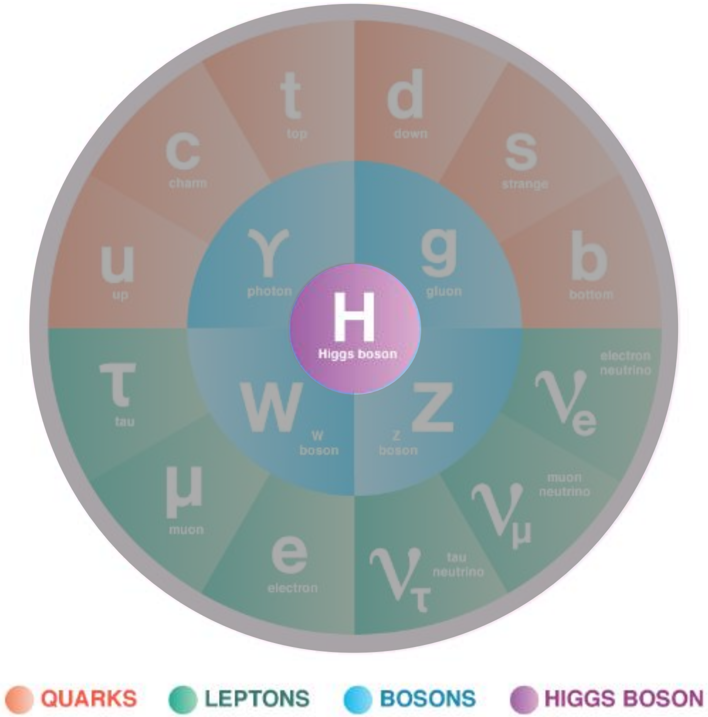


Image source: Symmetry Magazine

What is Dark Matter?
What is Dark Energy?

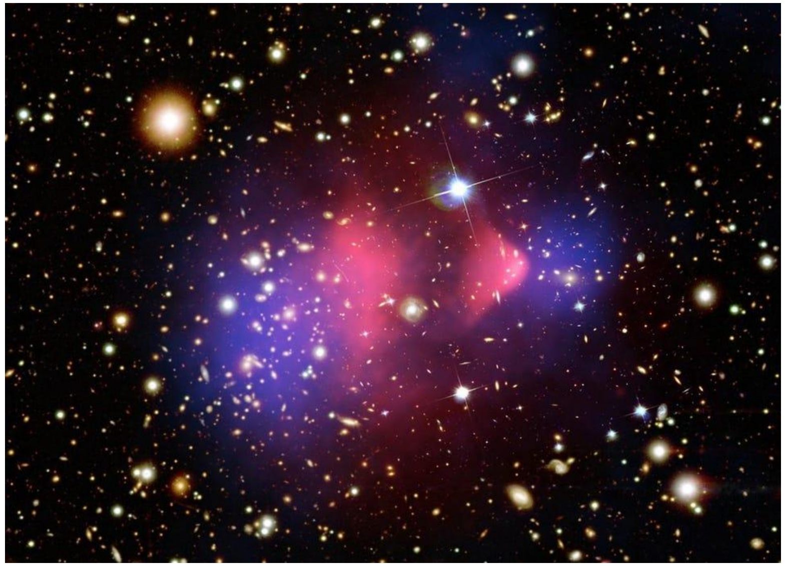


Image source: NASA/CXC/CFA/ M.MARKEVITCH

Why is there more matter than anti-matter in the universe?

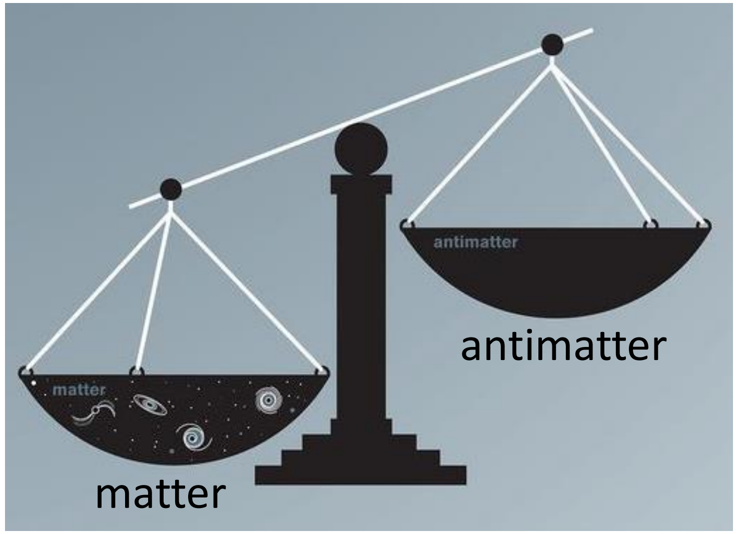


Image source: Symmetry Magazine

Studying Physics at the Smallest Scales