

Machine Learning Applications in Experimental High Energy Physics

Aruna Kumar Nayak
Institute of Physics, Bhubaneswar

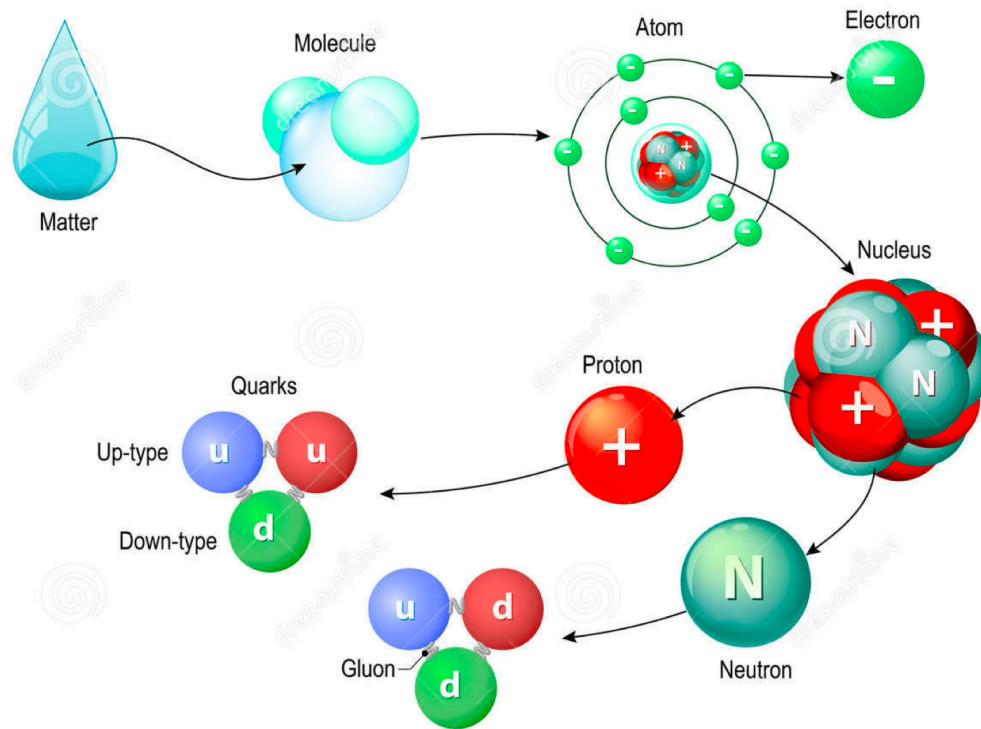


Symposium on AI for Science & Engineering
July 21 - 23, 2026
Indian Institute of Technology Hyderabad



High Energy (Particle) Physics

Goal: To understand how the universe works at most fundamental level, discover elementary constituents of matter and understand the fundamental forces governing them



Energy needed to remove an electron from an atom ~ 10 eV (10)
Energy needed to remove a proton from a nucleus we need ~ 10 MeV (10 million)

High Energy (Particle) Physics

Goal: To understand how the universe works at most fundamental level, discover elementary constituents of matter and understand the fundamental forces governing them

Experimentally:

- Use particle colliders to accelerate and smash particles at very high energies

- Colliders act as powerful microscopes:

- Energy connected to resolution:

$$\lambda = \frac{h}{p} = \frac{hc}{\beta E}$$

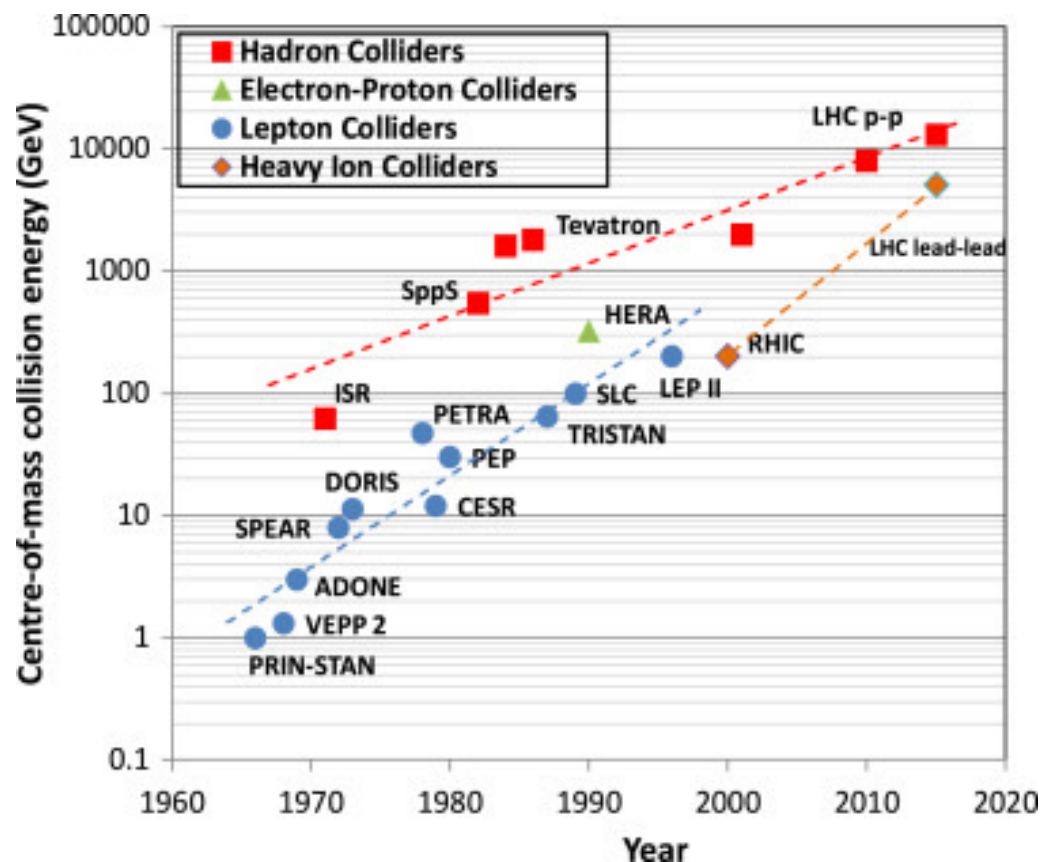


- Energy from collisions is converted into mass ($E = mc^2$), creating new, heavier, and unstable particles that existed shortly after the creation of the Universe
- Study cosmic rays, gamma rays, neutrinos etc.. originating from astrophysical sources
 - Can have much higher energies than accessible at the colliders, but with smaller intensity and uncontrolled environment.

Collider based experiments over the years

Key Particle discoveries in accelerators

The colliders, thanks to advancing technologies, have proven to be powerful instruments for discoveries and precision measurements.



1970

1975

1980

1985

1990

1995

2000

2005

2010

1974, Charm quark (J/Psi)
1976, Tau
1977, Bottom quark
1979, Gluon

1983, W & Z bosons

Top quark

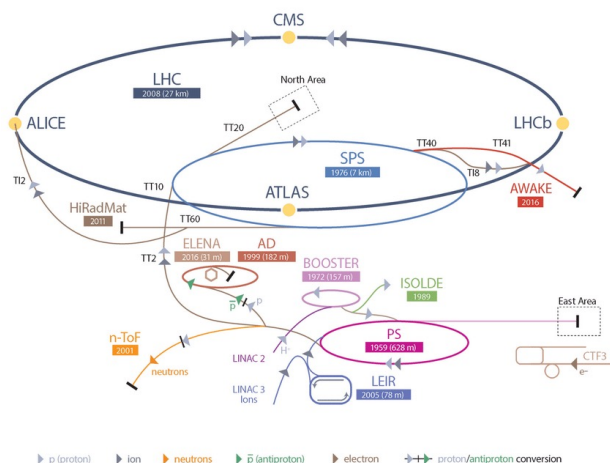
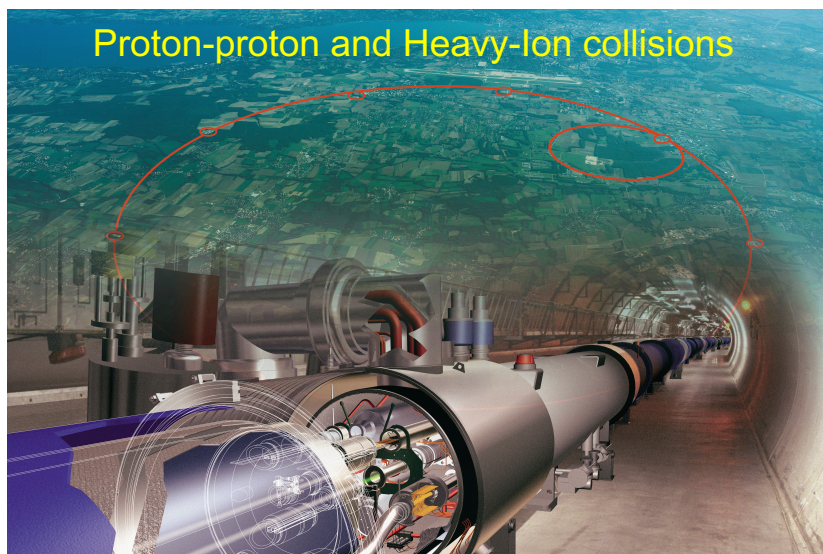
Tau neutrino

2012, Higgs boson

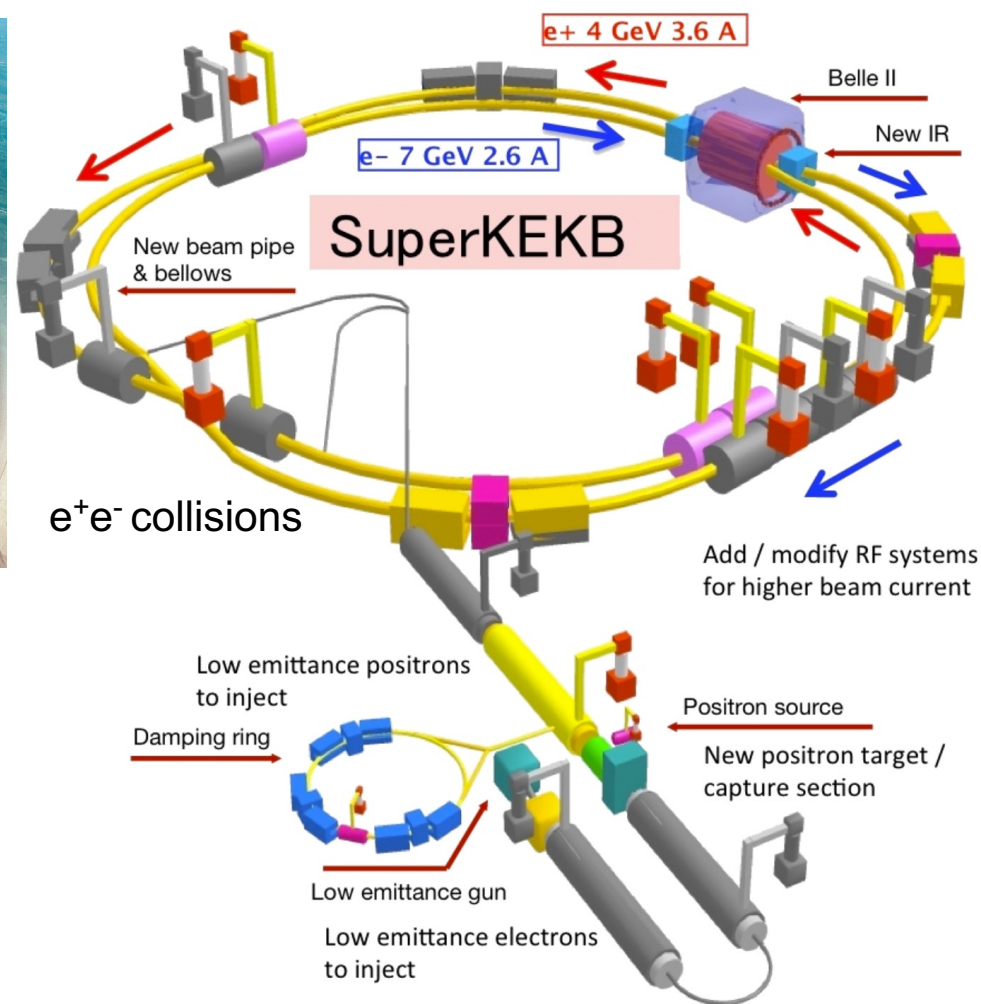
?

Ongoing collider experiments

LHC, CERN, Geneva

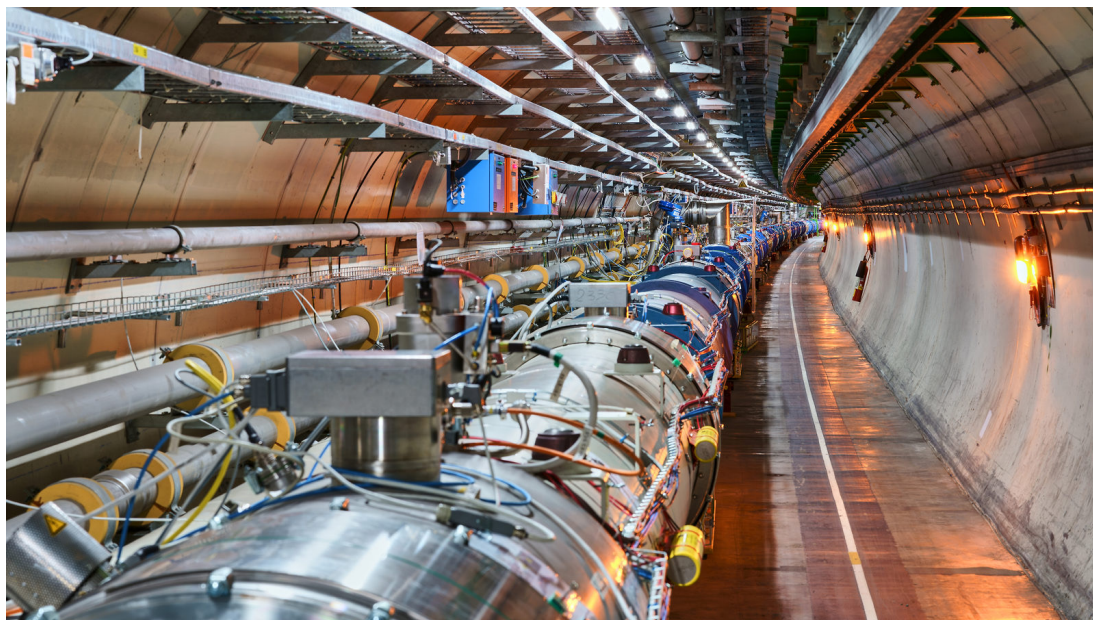


Belle-II @ SuperKEKB, Japan

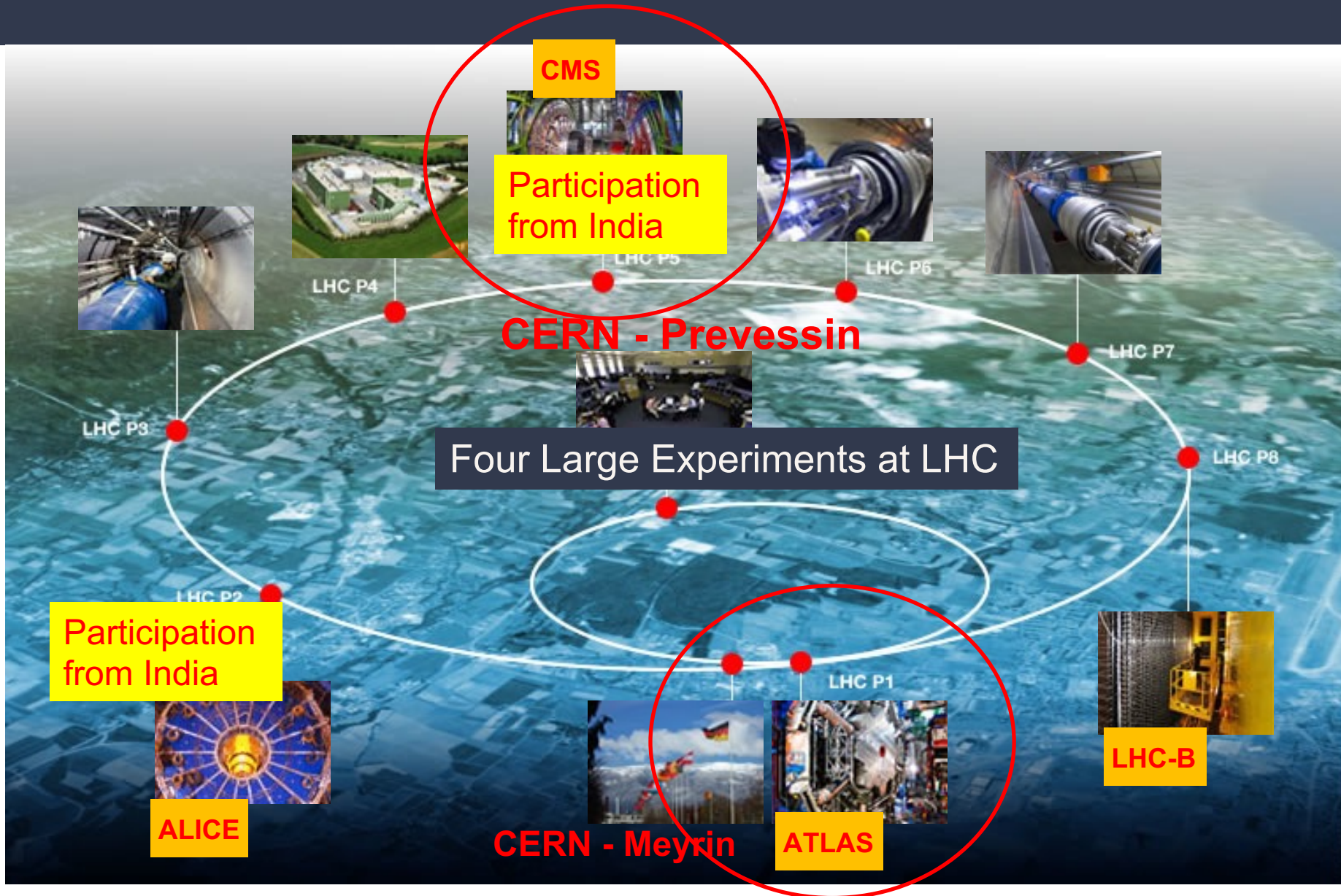


The Large Hadron Collider

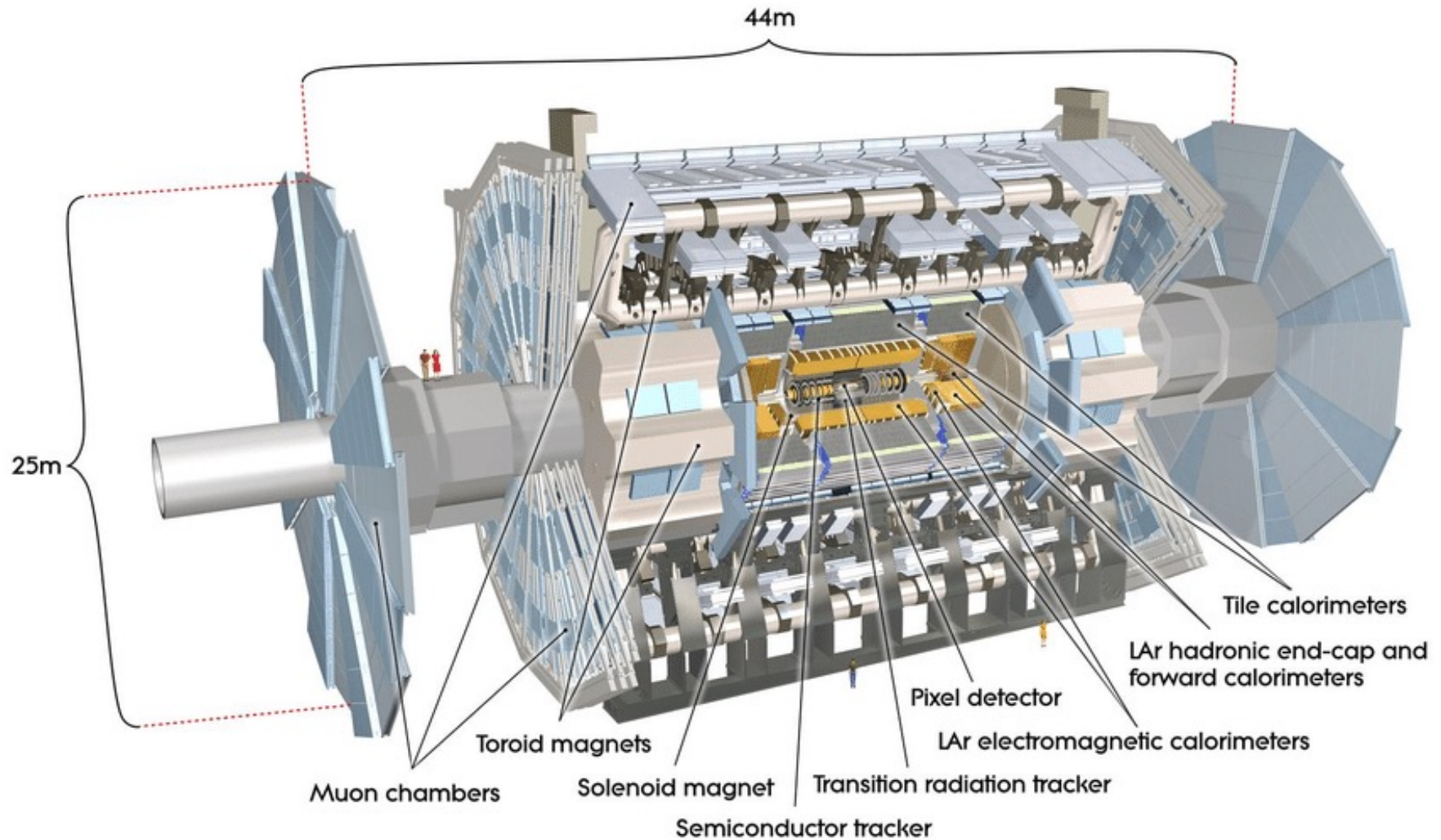
- The LHC is a **proton-proton collider**, with two proton beams travelling in the opposite directions, currently colliding at a center-of-mass energy of **13.6 TeV**
[1 eV = energy acquired by an electron when accelerated by a potential difference of 1 volt]
- Housed in a **27 KM circular tunnel, about 100m under ground**, around Swiss-France border near Geneva.
- Consists of thousands of magnets of different varieties and sizes.
 - **1232 dipole magnets** – superconducting electromagnets, 15 m long each, operating at 1.9 °K
 - **392 quadrupole magnets**, 5 – 7 m long, to focus the beam
- Separate beam pipes at **ultrahigh vacuum** ($10^{-10} - 10^{-11}$ mbar)



Experiments @ LHC



ATLAS Experiment



CMS Experiment

- 75M channels
- ~10 μm position resolution
- $O(1\%)$ mom. Resolution
- $O(\text{few}\%)$ energy resolution

CMS DETECTOR

Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel ($100 \times 150 \mu\text{m}$) $\sim 1\text{m}^2 \sim 66\text{M}$ channels
Microstrips ($80 \times 180 \mu\text{m}$) $\sim 200\text{m}^2 \sim 9.6\text{M}$ channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying $\sim 18,000\text{A}$

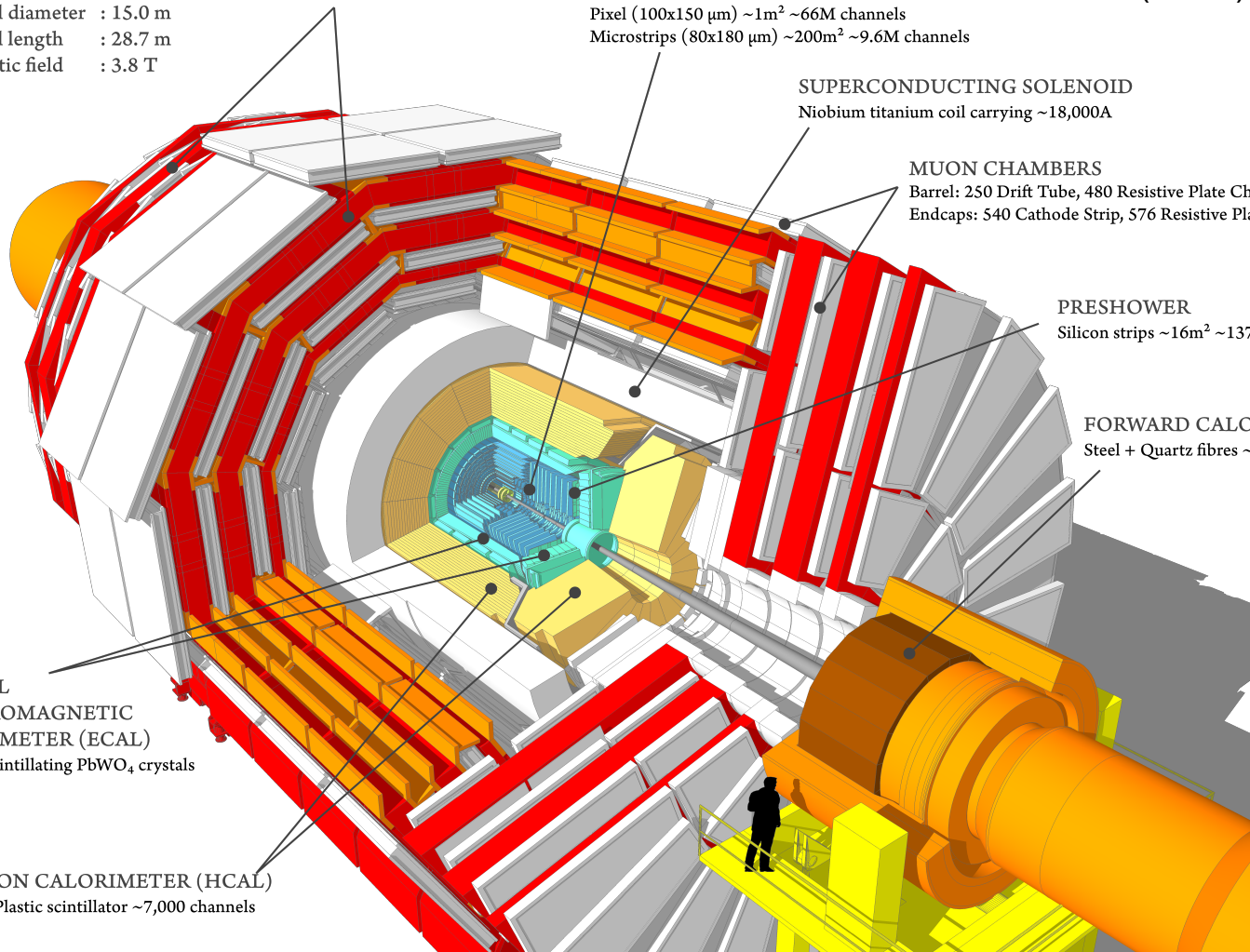
MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 540 Cathode Strip, 576 Resistive Plate Chambers

PRESHOWER
Silicon strips $\sim 16\text{m}^2 \sim 137,000$ channels

FORWARD CALORIMETER
Steel + Quartz fibres $\sim 2,000$ Channels

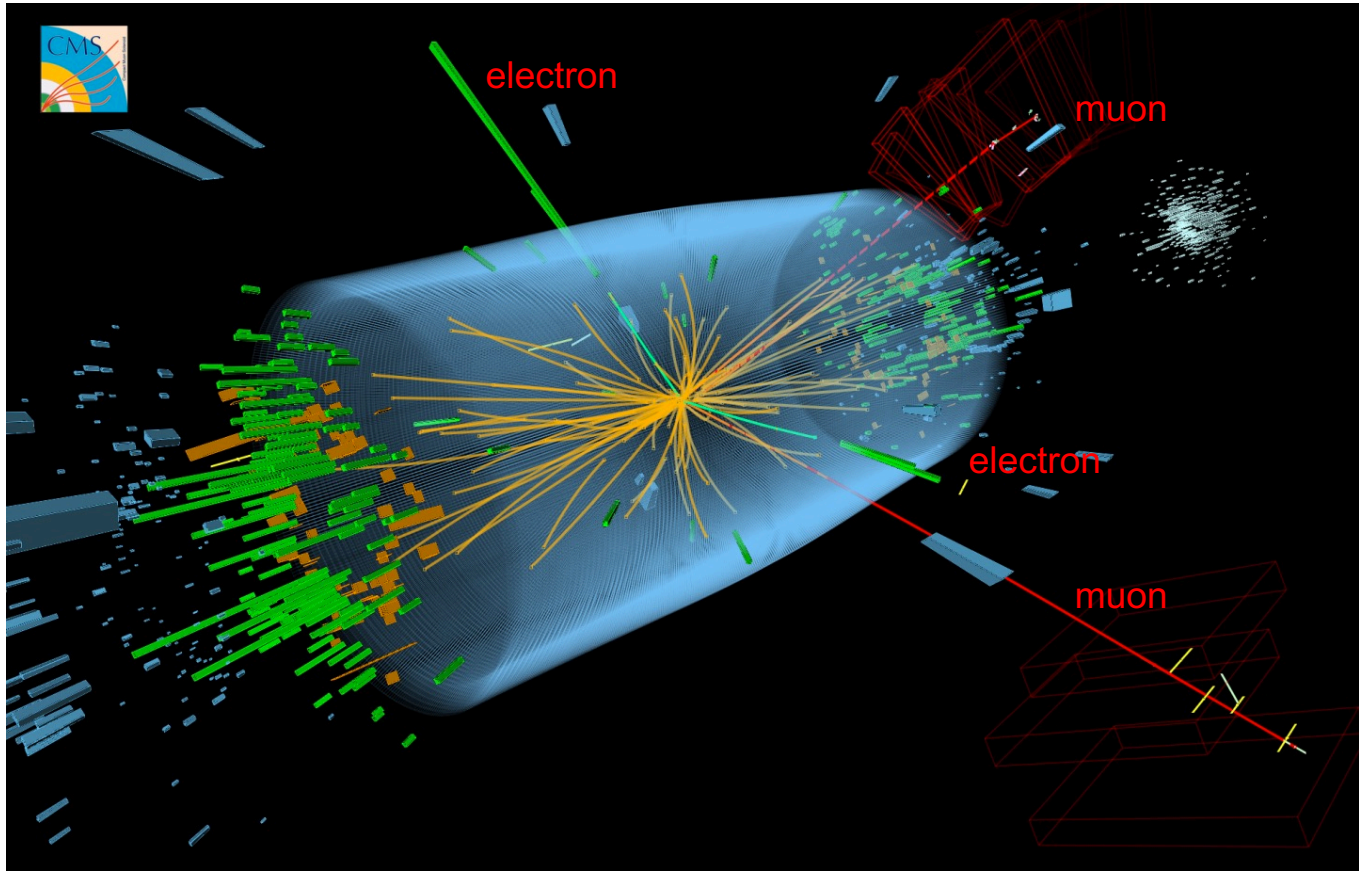
CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
 $\sim 76,000$ scintillating PbWO_4 crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator $\sim 7,000$ channels

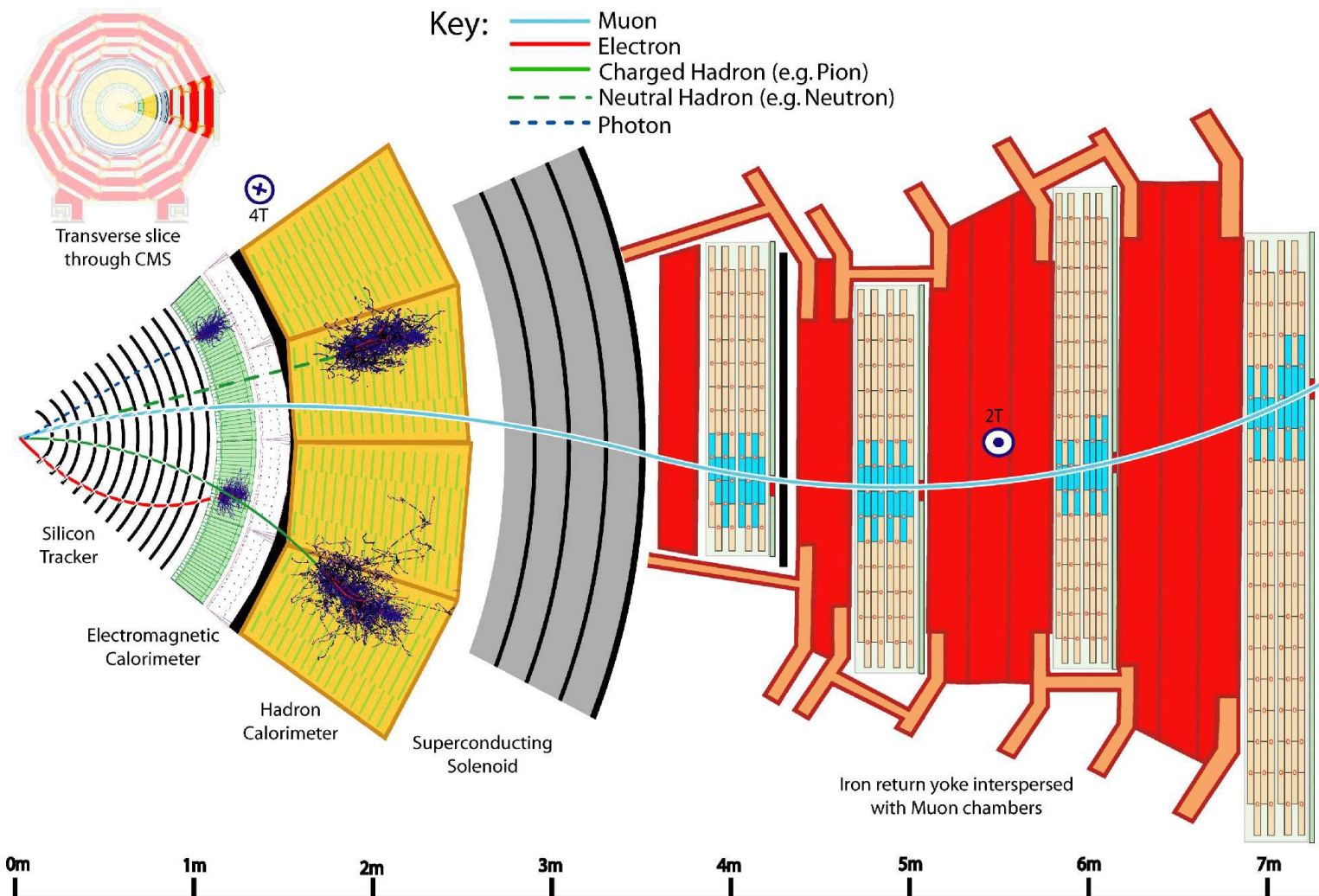


A typical event at LHC

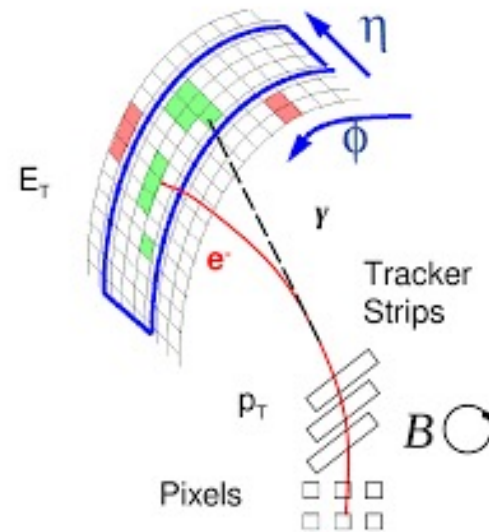
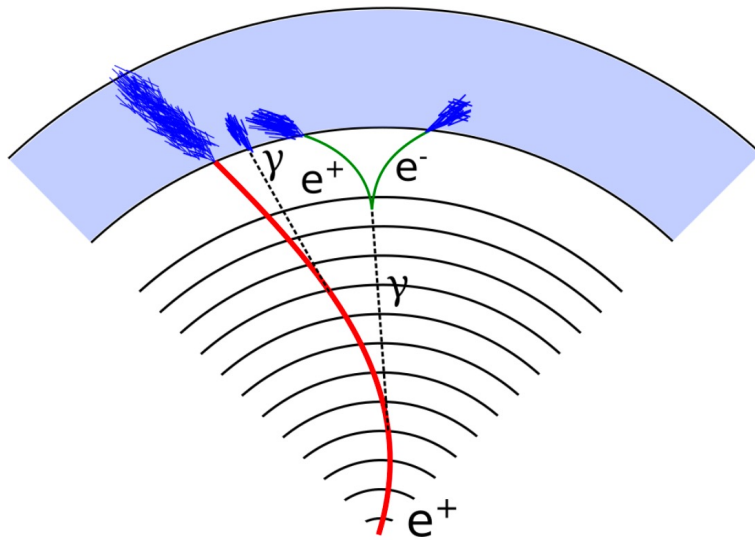
$$H \rightarrow ZZ^* \rightarrow 2e2\mu$$



Particle Reconstruction



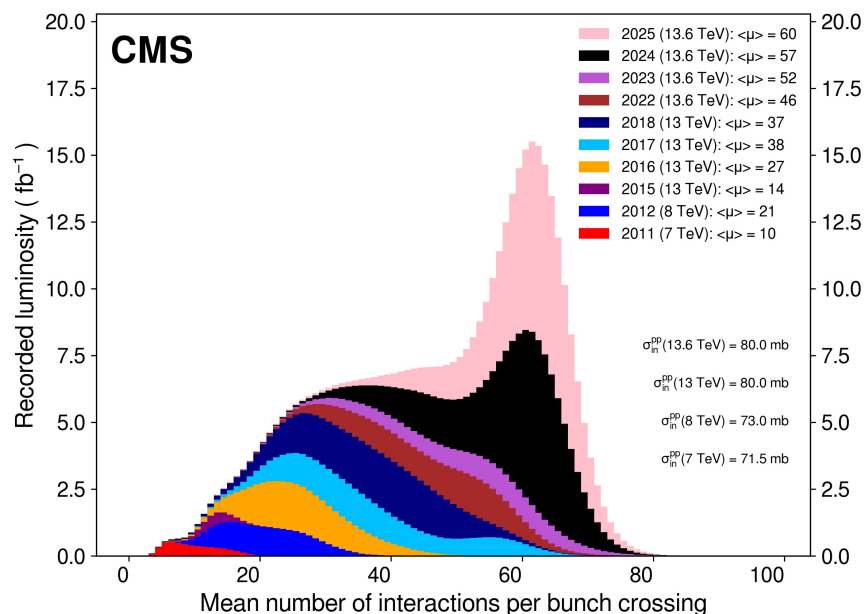
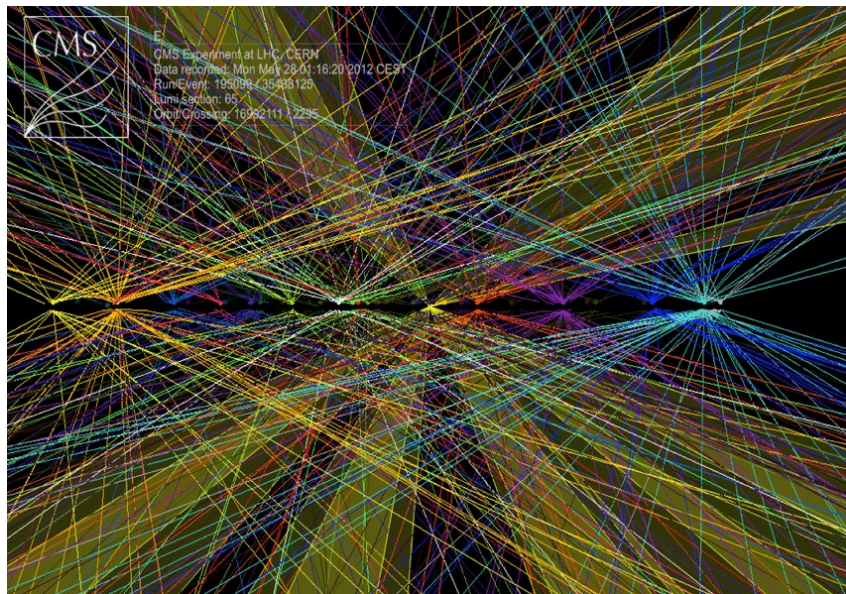
In reality the track / cluster patterns are more complex,
resulting in multiple calorimeter clusters



Example electron patterns

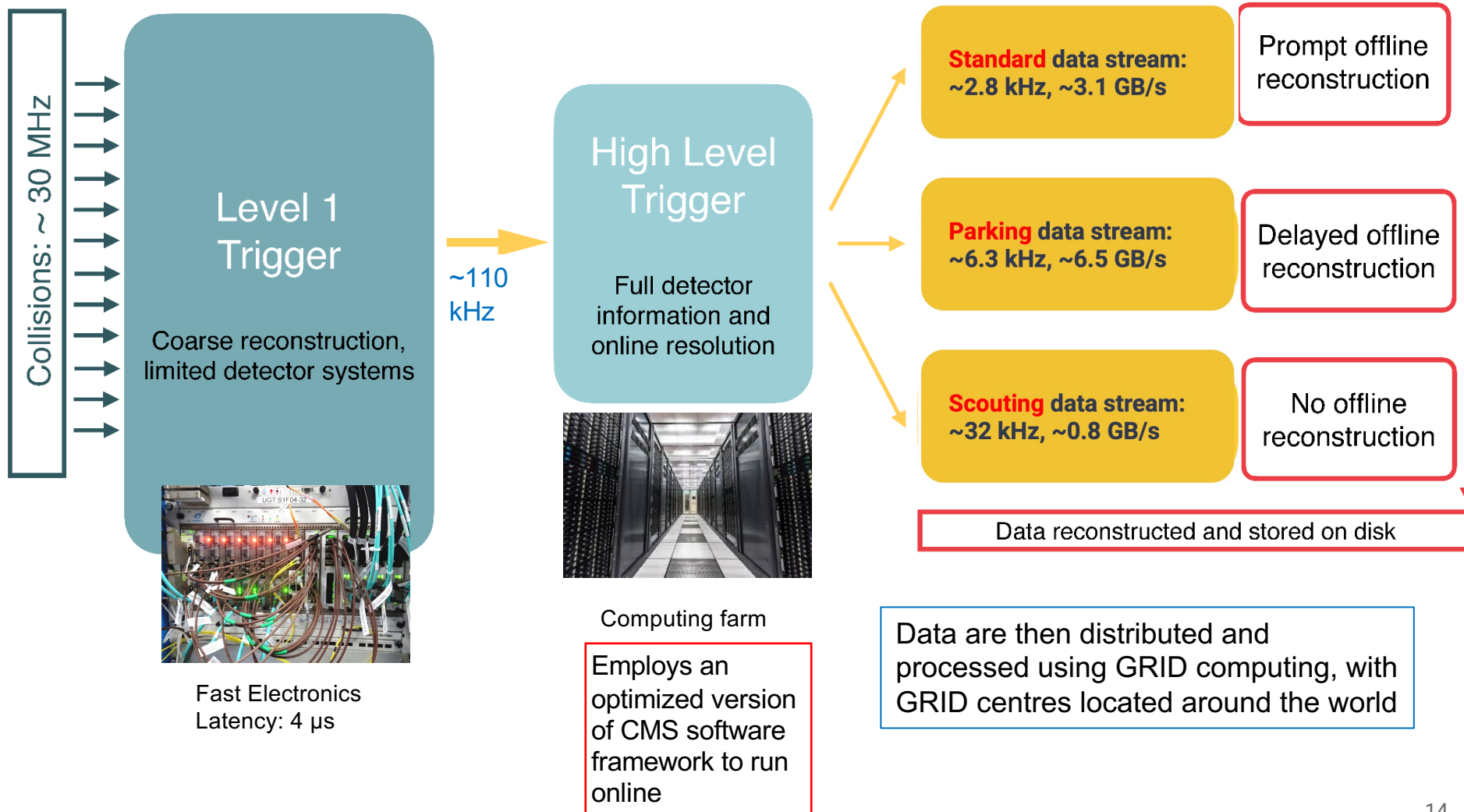
Challenges from additional pp interactions

Multiple pp interactions per LHC bunch crossing (Pileup)



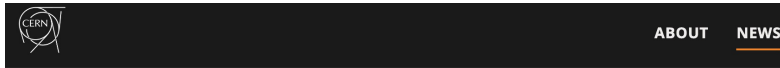
- Large number of additional soft pp interactions in the same bunch crossing
- Challenging to distinguish primary collision from overlapping events
- Affects energy, momentum measurements, and object identification

Online Event Filtering and Recordings



HEP & Machine Learning

2012



Voir en [français](#)

CERN experiments observe particle consistent with long-sought Higgs boson

4 JULY, 2012

Geneva, 4 July 2012. At a seminar held at CERN¹ today as a curtain raiser to the year's major particle physics conference, ICHEP2012 in Melbourne, the ATLAS and CMS experiments presented their latest preliminary results in the search for the long sought Higgs particle. Both experiments observe a new particle in the mass region around 125-126 GeV.



Nobel Prize in Physics 2013 to Higgs & Englert

2012

Imagenet classification with deep convolutional neural networks

Authors Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton

Publication date 2012

Journal Advances in neural information processing systems

Volume 25

IMAGENET



Nobel Prize in
Physics 2024

The emergence of [AlexNet](#) (a Deep CNN) sort-of started deep learning revolutions

HEP community was quick to catch up with its applications

- HEP experiments produce huge amount of data
 - Natural area for ML applications

Machine Learning is not new to HEP



Computer Physics Communications

Volume 49, Issue 3, June 1988, Pages 429-448



1988

Neural networks and cellular automata in experimental high energy physics



Computer Physics Communications

Volume 119, Issues 2-3, 2 June 1999, Pages 219-231



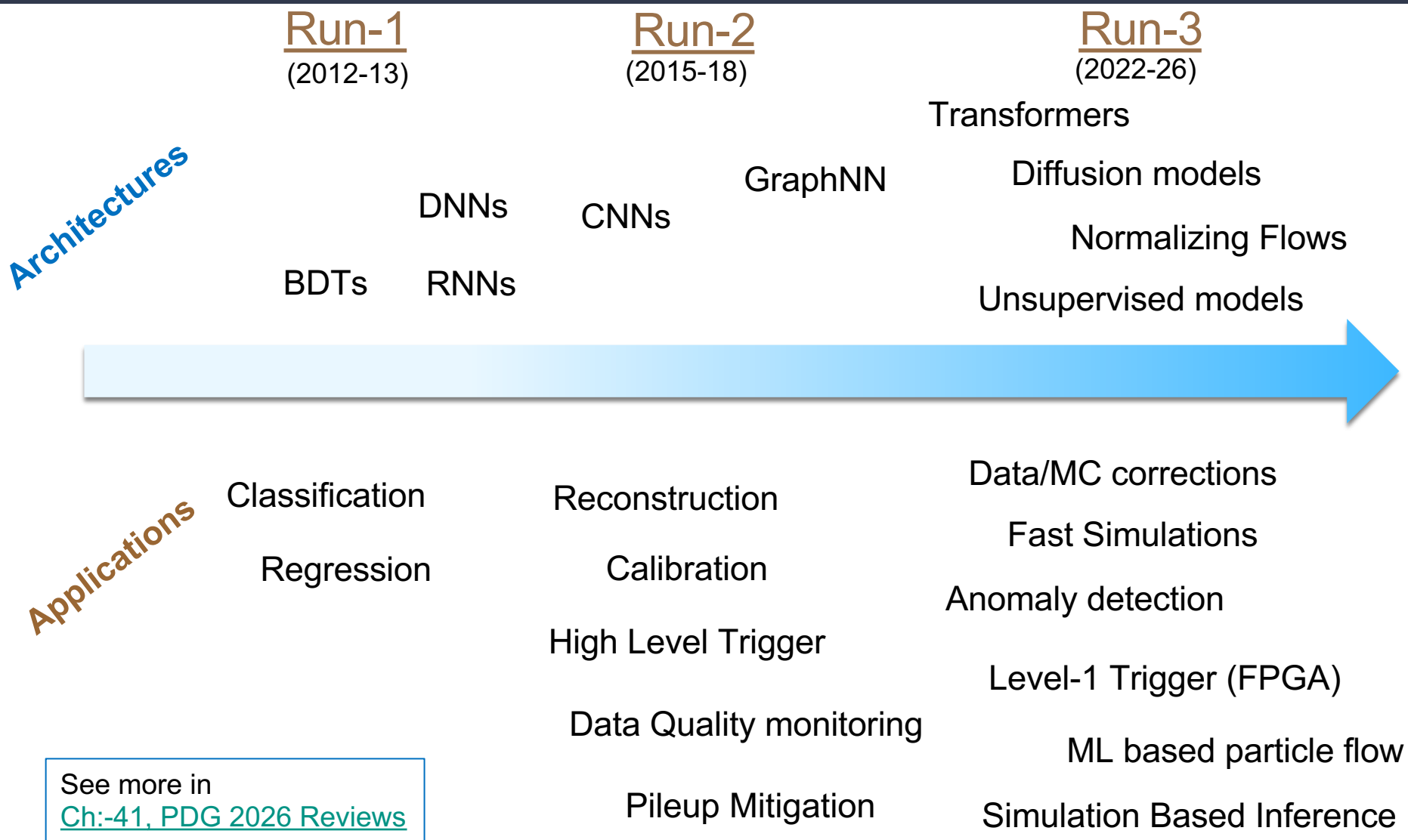
Neural networks in high energy physics: A ten year perspective

1999

Bruce Denby

- Use of multivariate methods since 1980s
 - But, mostly shallow neural networks, boosted decision trees
 - Played significant role in Higgs boson discovery at LHC

ML Applications @ LHC Experiments

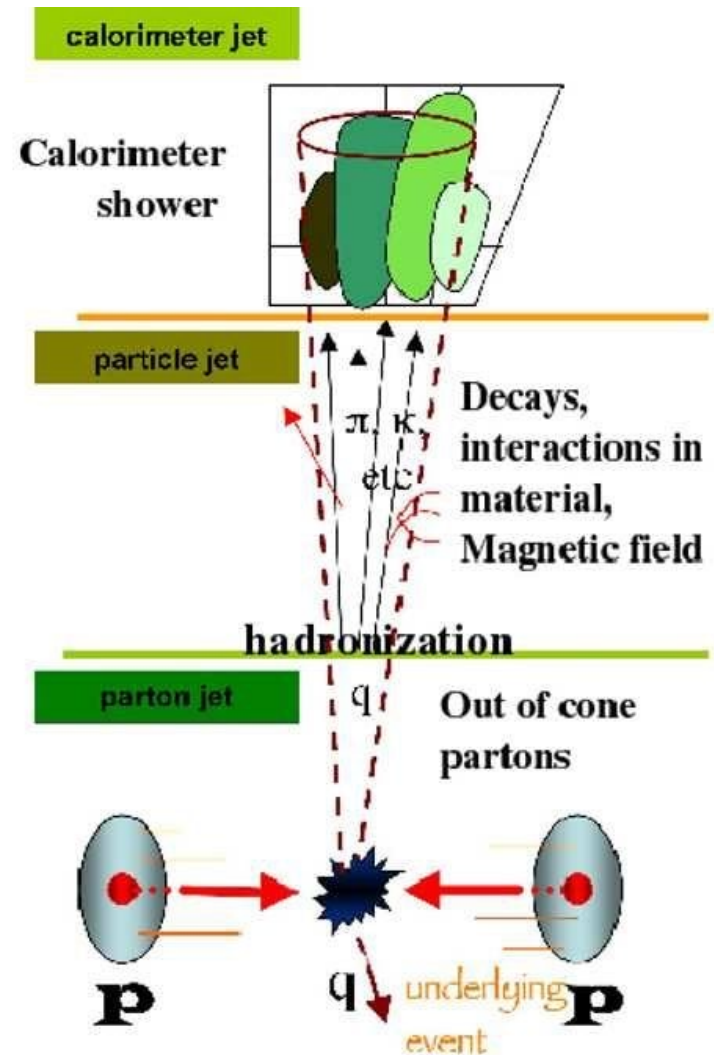


Quark / Gluon jets

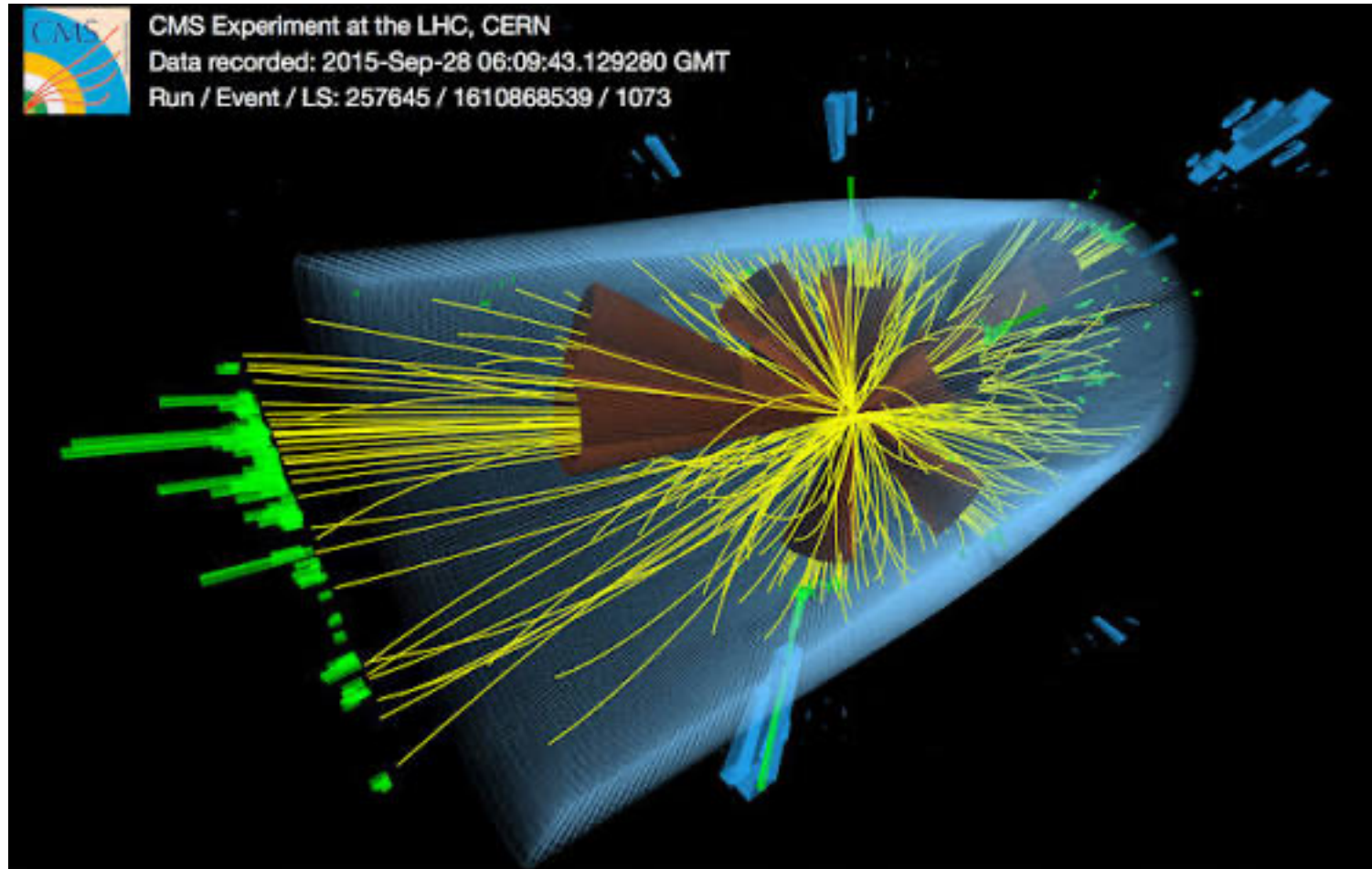
- Quarks and gluons carry color charge and can not exist in free form because of QCD confinement
- The **quarks/gluons**, moving away from each other, **go through** a process, called **fragmentation & hadronization**, to produce color less hadrons (bound state of quarks)
- These hadrons, each carrying a fraction of parent quark/gluon's momentum, move along that direction like a cone.
- It is called a **hadron Jet**

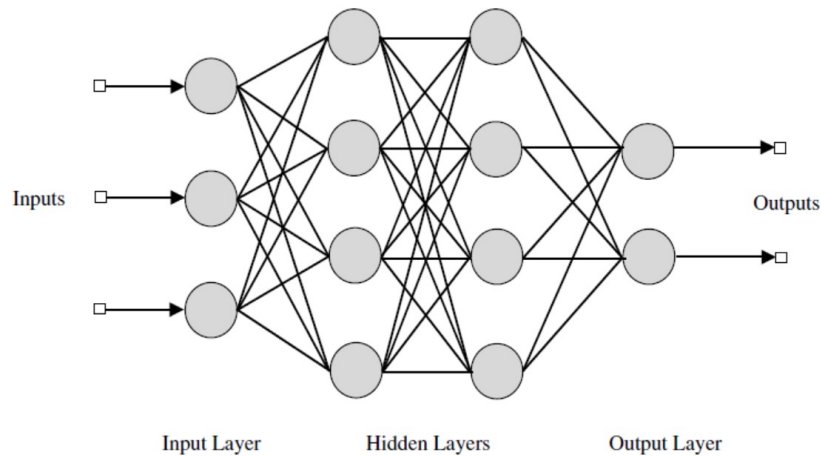
Jets are reconstructed using **jet clustering algorithms**

See [Ref-1](#), [Ref-2](#)



A Multi-jets event

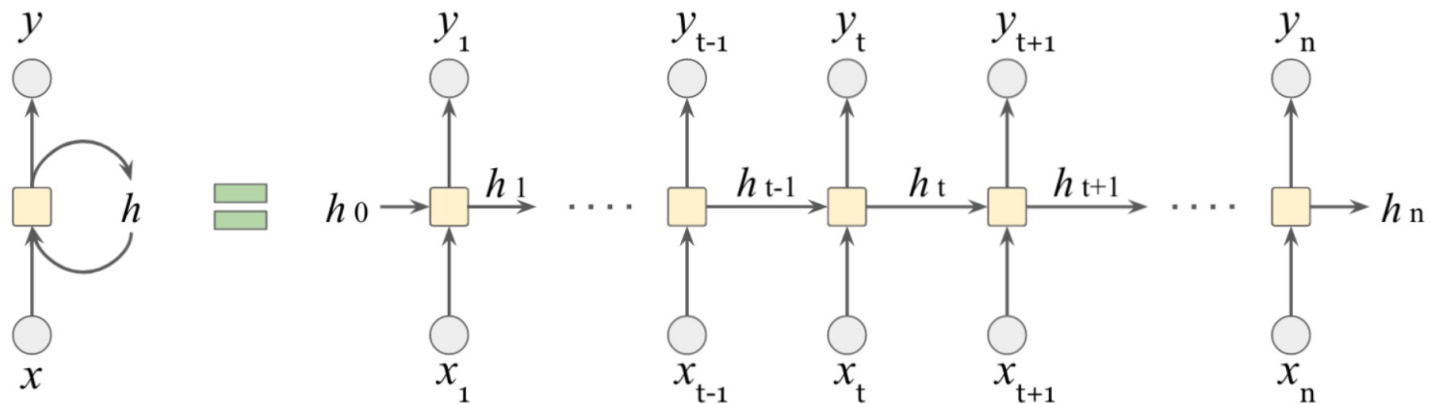




Feed forward neural network

Universal function approximator

No inherent input invariance



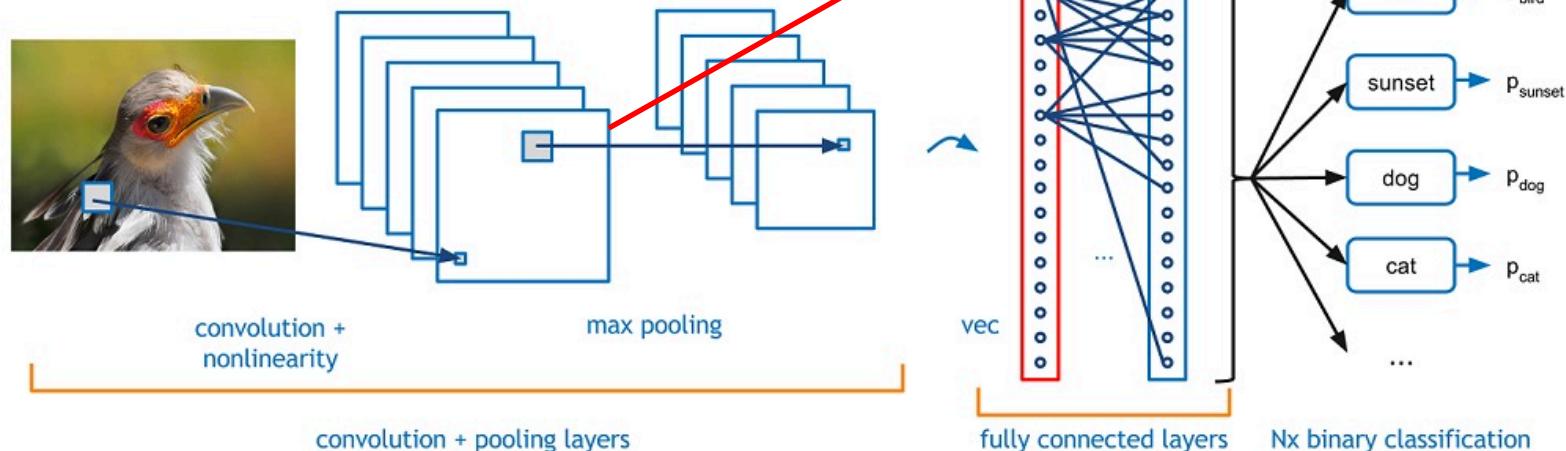
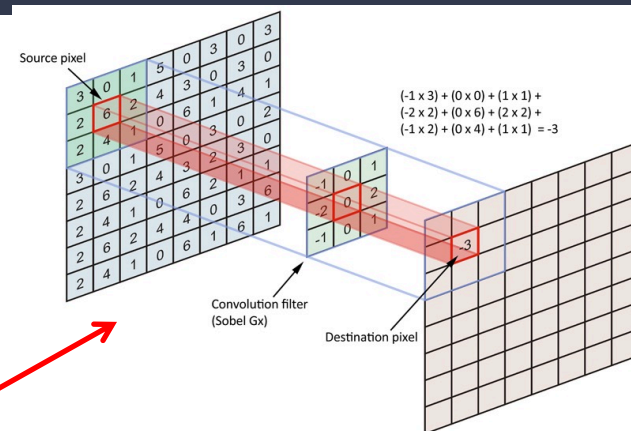
Recurrent NN: used for sequential data (or time series data)

Convolutional Neural Networks (CNNs)

Mostly used for image recognition

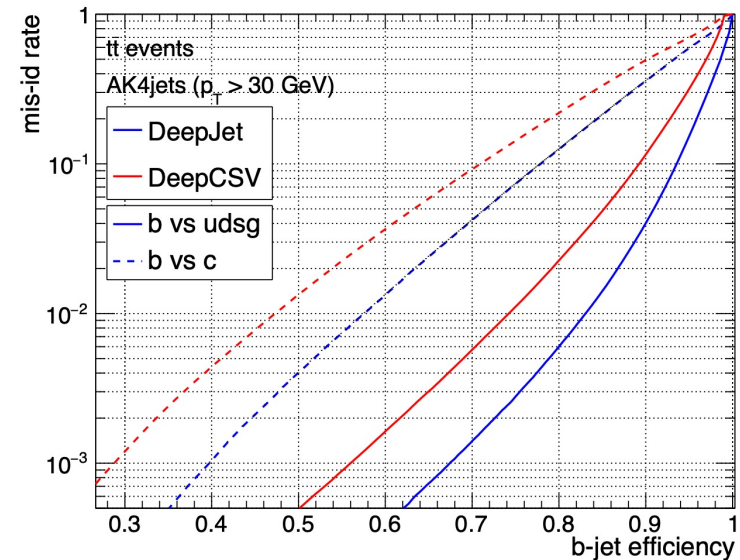
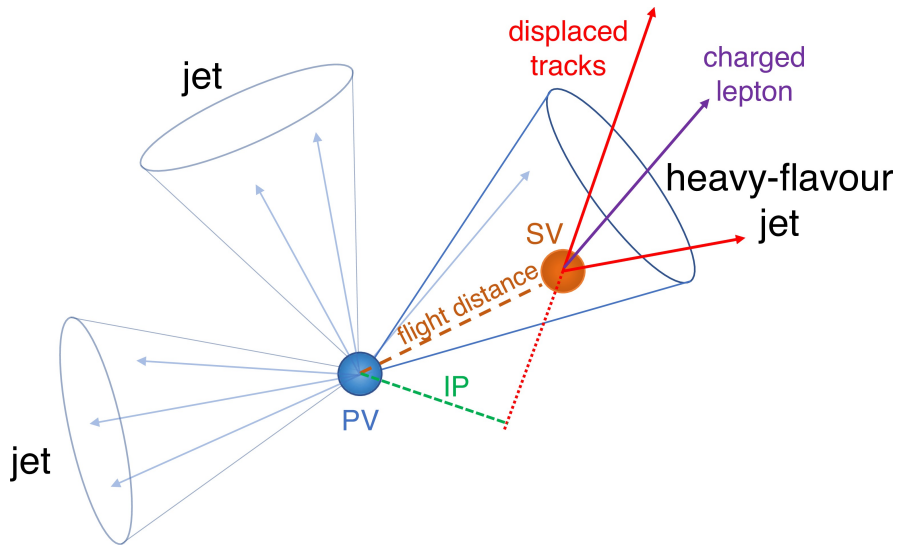
Sensitive on translational invariant features

Has a sense of distance and orientation

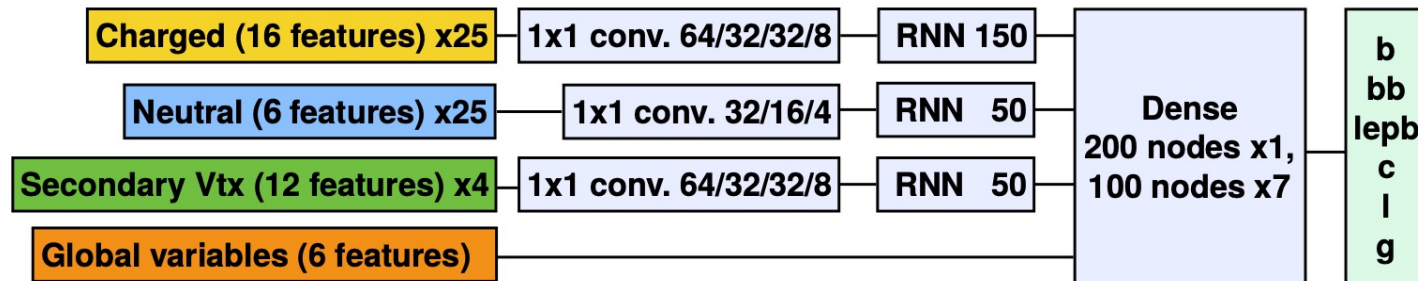


Many applications in HEP → Energy deposits in detector(s) can be considered as an image

Jet flavour classification

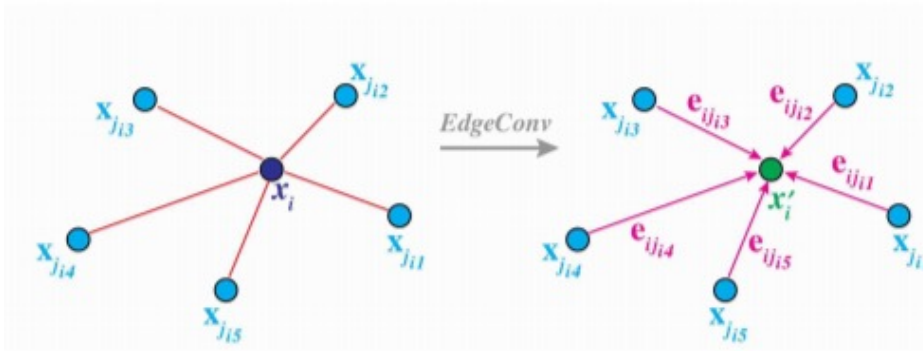


DeepJet Architecture: RNN + 1x1 CNNs for dimensionality reduction



([arXiv:2008.10519](https://arxiv.org/abs/2008.10519))

Graph Neural Networks (GNNs)



Designed to perform inference on data described by graphs.

Can extract information about the full graph or single nodes (objects) or edges.

[arXiv:1812.08434](https://arxiv.org/abs/1812.08434)

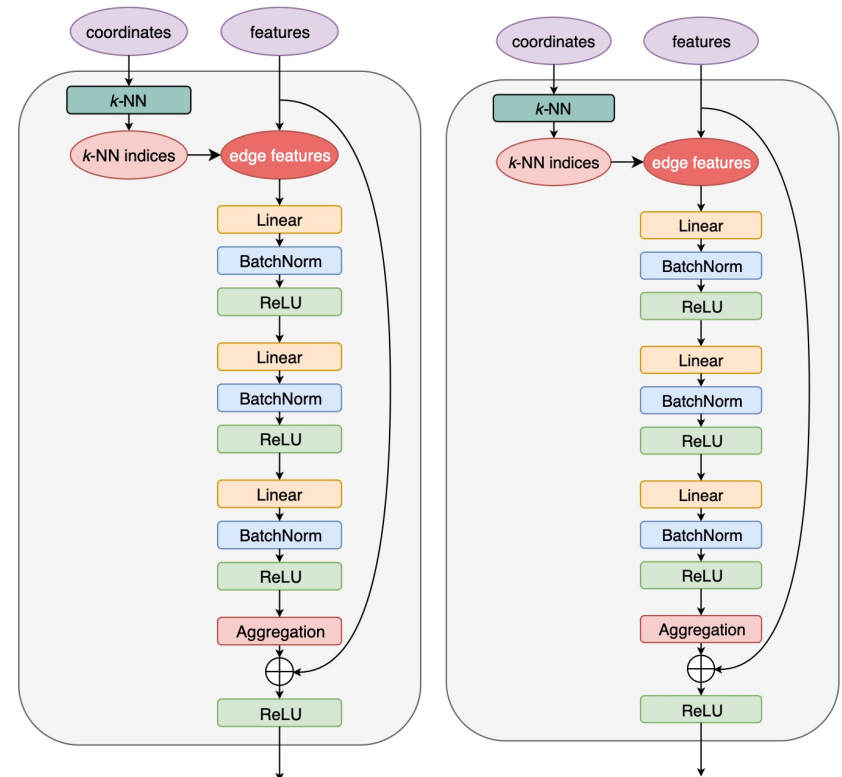
Graph neural networks: A review of methods and applications

Jie Zhou^{a,1}, Ganqu Cui^{a,1}, Shengding Hu^a, Zhengyan Zhang^a, Cheng Yang^b, Zhiyuan Liu^{a,*}, Lifeng Wang^c, Changcheng Li^c, Maosong Sun^a

ParticleNet ([arXiv:1902.08570](https://arxiv.org/abs/1902.08570))

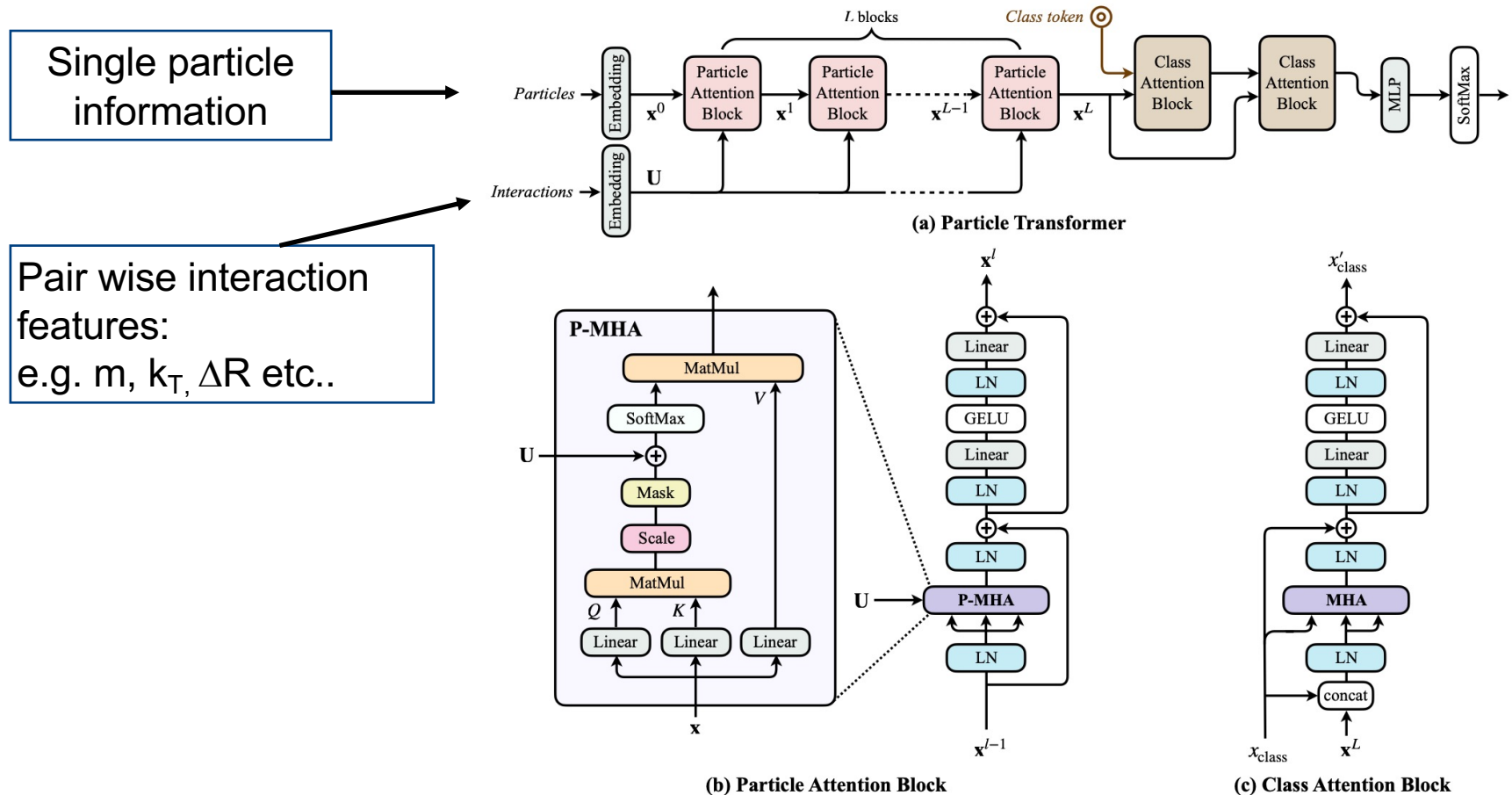
(Based on DGCNN)

[ACM Trans. Graph. 38, 146 \(2019\)](https://arxiv.org/pdf/1902.08570v1.pdf)



Used for jet Tagging → Considers jet as an unordered set of constituent particles

Transformers



Similar to fully connected graph networks with efficient attention mechanism.

Efficiently models sequential data ([arXiv:1706.03762](https://arxiv.org/abs/1706.03762))

Several extensions for particle physics applications ([arXiv:2202.03772](https://arxiv.org/abs/2202.03772))

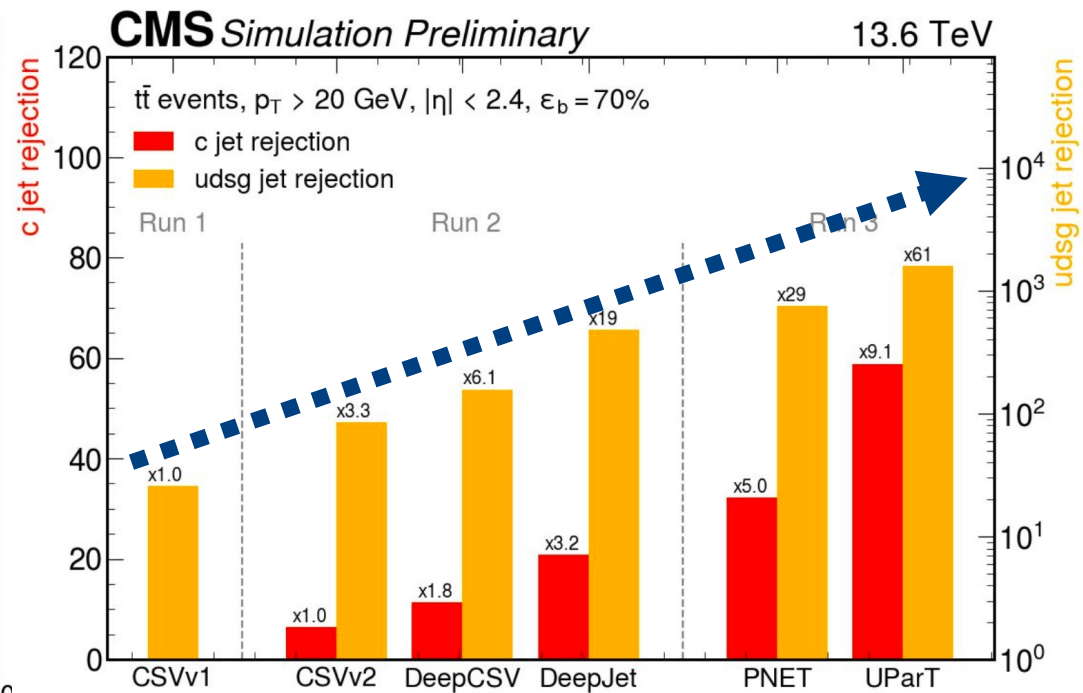
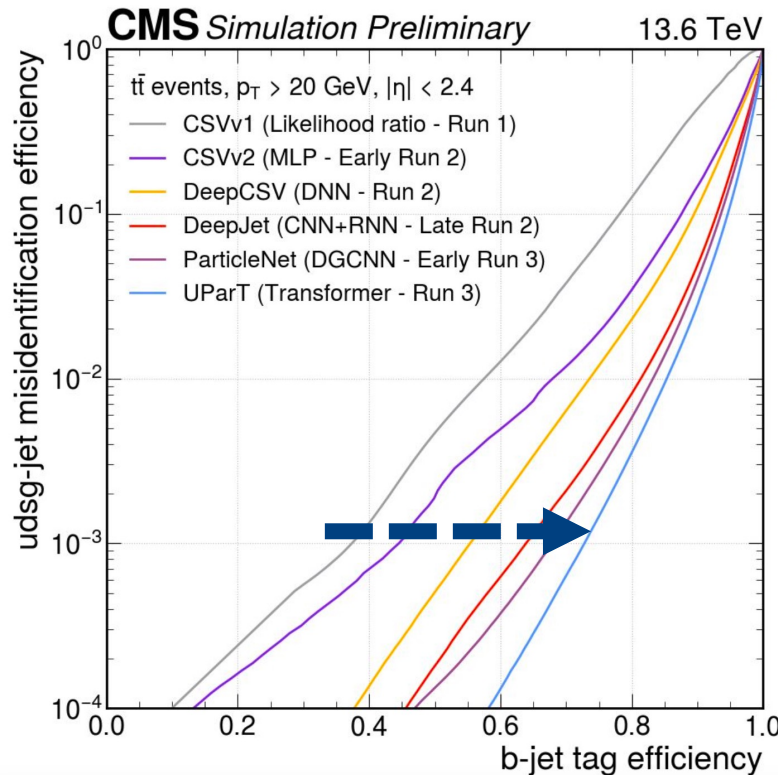
Unified approach to jet tagging

[CMS-DP-2024-066](#)

Evolution of jet taggers → **DNN** to GraphNN → ParticleTransformer

Includes also jet energy regression and resolution

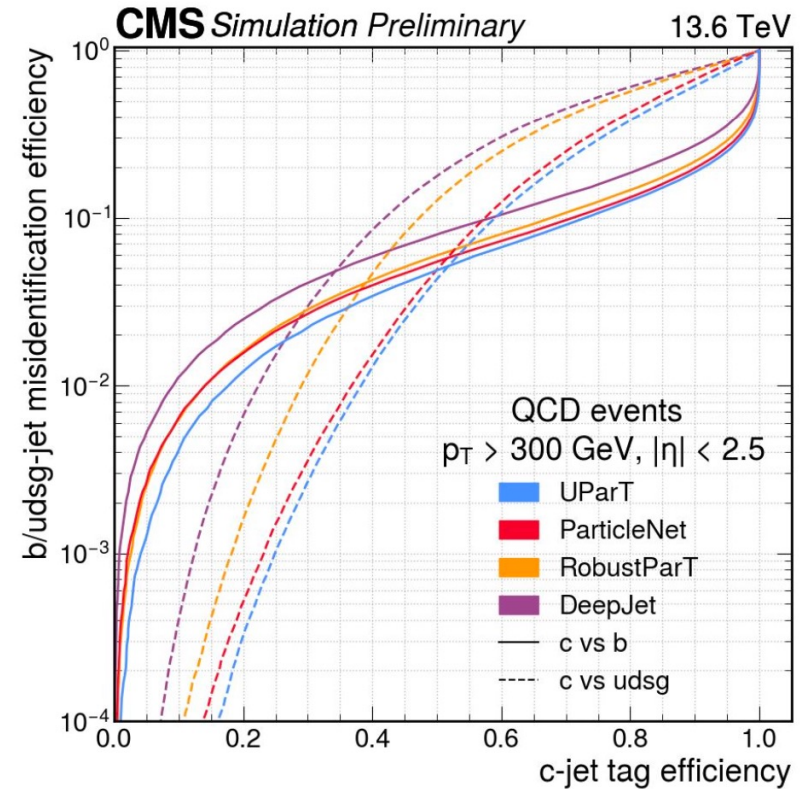
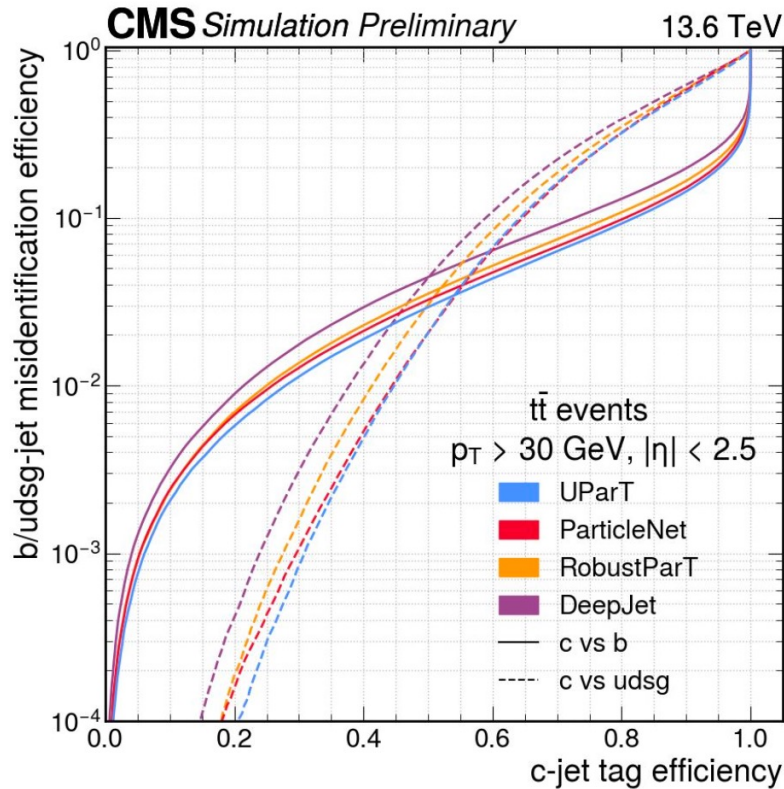
Adversarial training to improve Data/MC robustness



Improvement in performance with time

Improved c-jet tagging

[CMS-DP-2024-066](#)



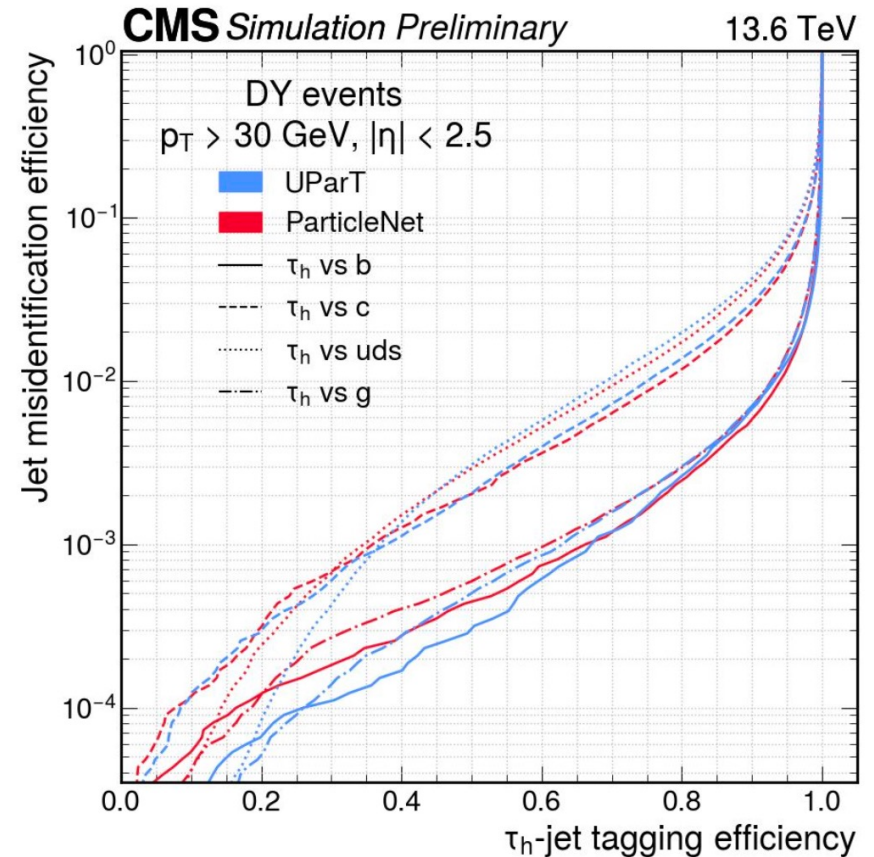
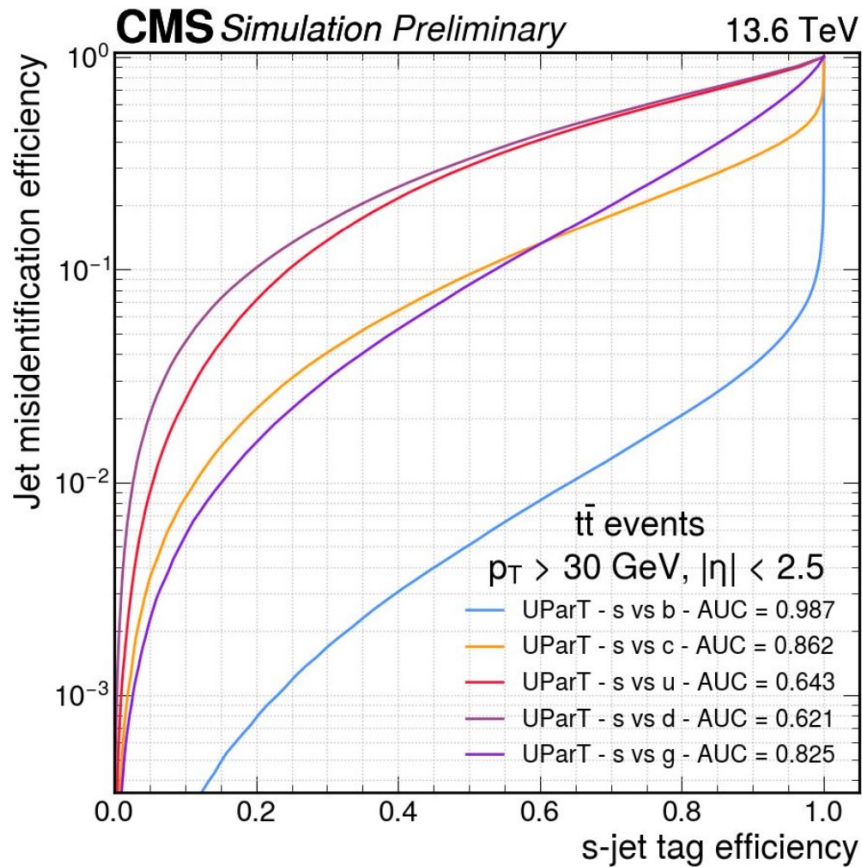
State-of-the-art performance for both b and light jets rejections

Extended class: **tau** and **strange-jet** tagging

[CMS-DP-2024-066](#)

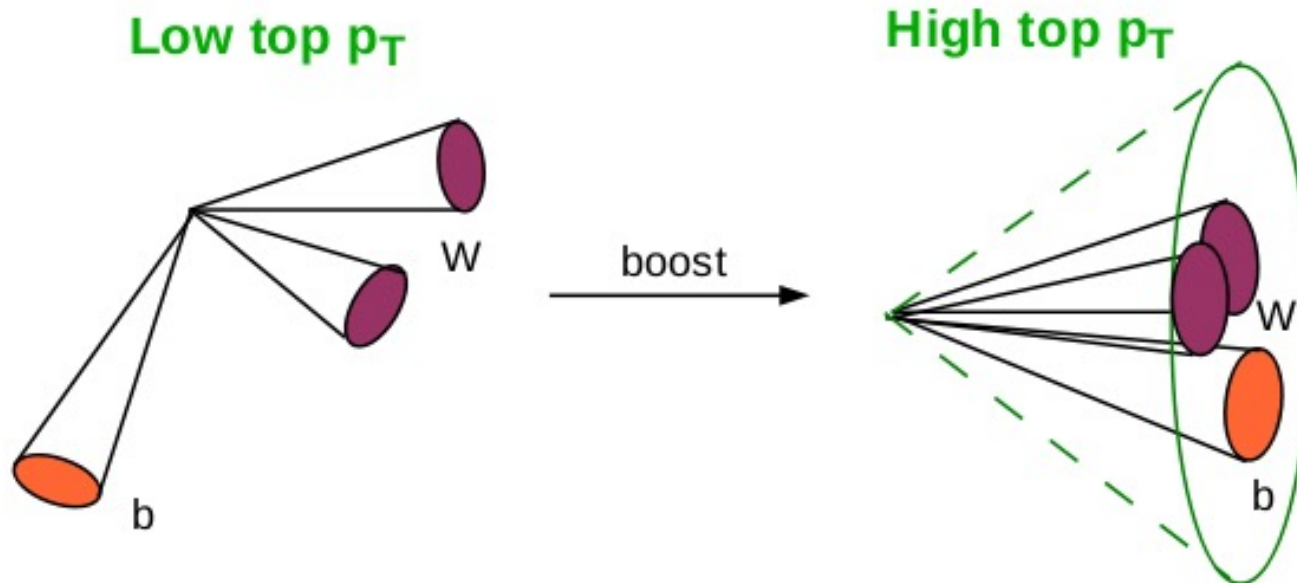
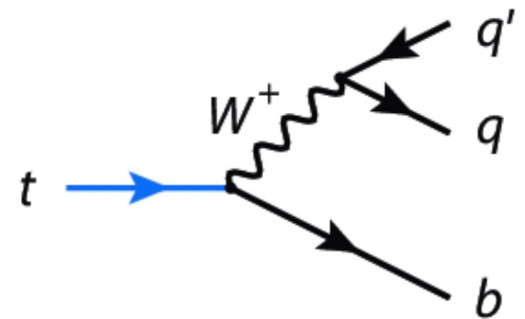
ParticleNet / UParT can improve further tau-tagging performance

First time tagging of **s – jets** (with low efficiency)



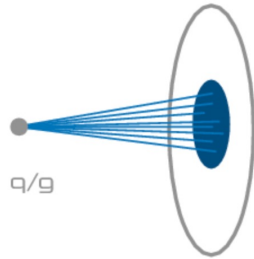
Identifying Boosted and Merged Objects

Boosted top quark decays

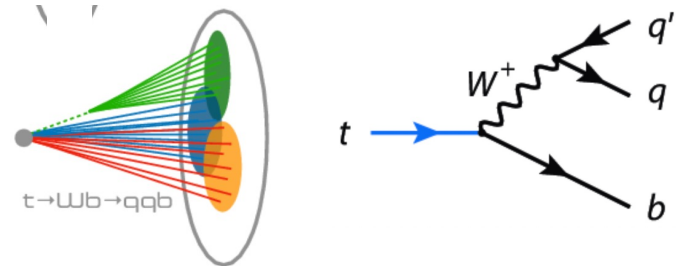


Identifying Boosted and Merged Objects

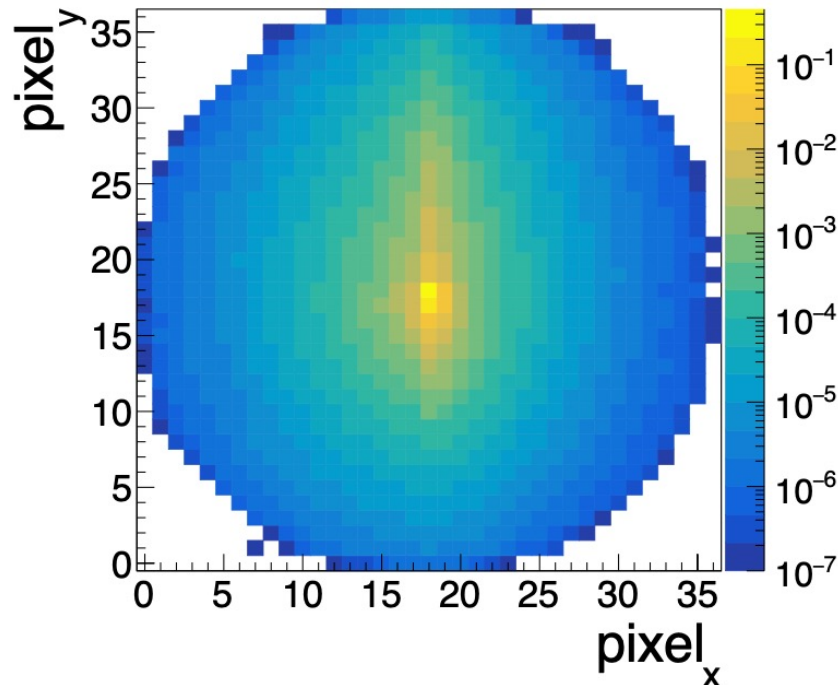
Quark / Gluon jets



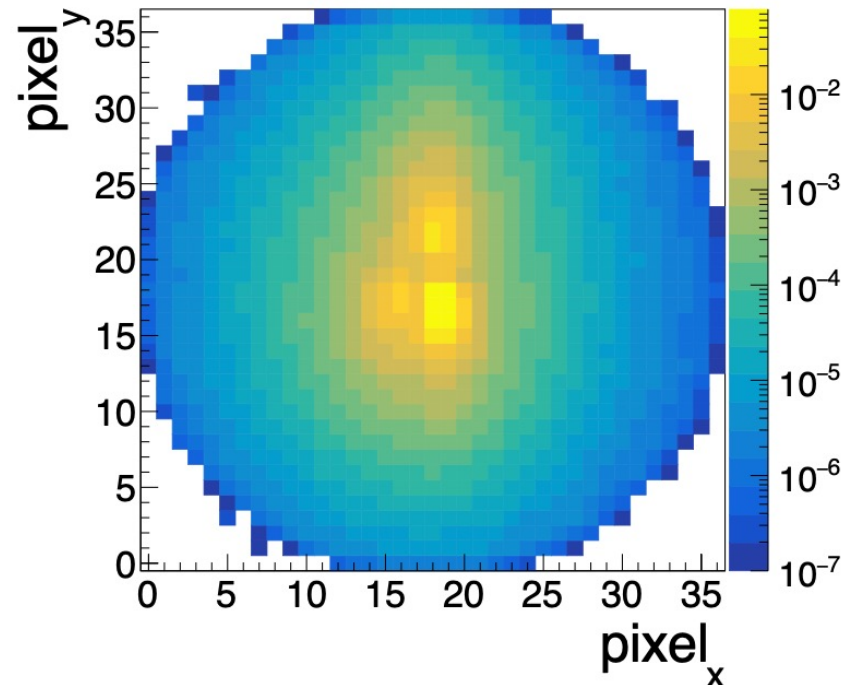
Boosted top quark jets



CMS Simulation

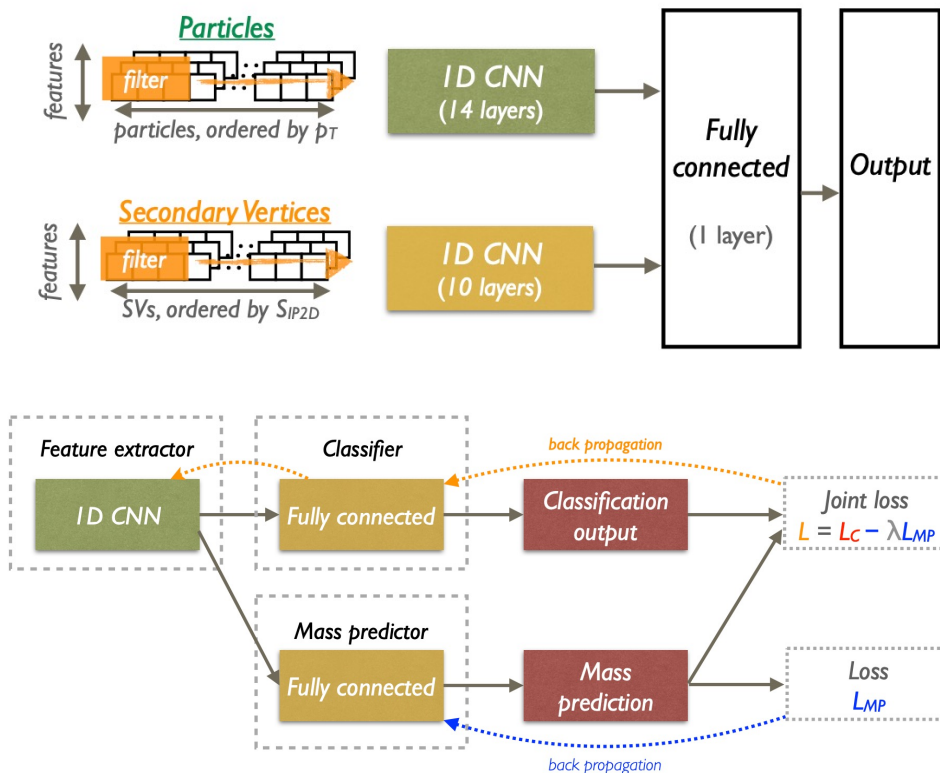


CMS Simulation

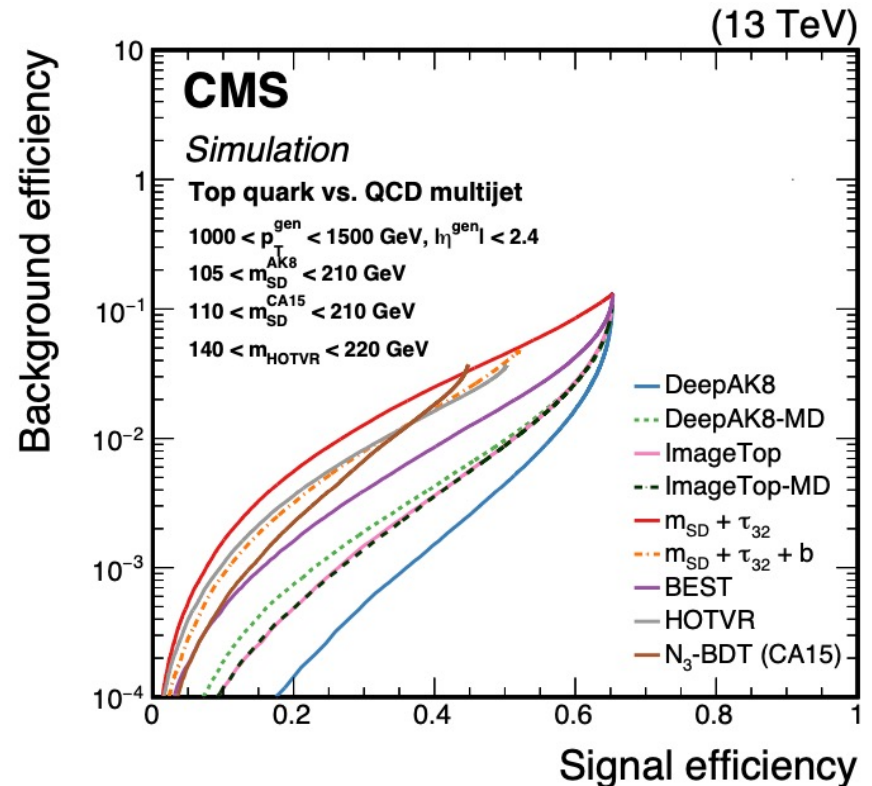


Identification of boosted top quark

DeepAK8 ([arXiv:2004.08262](https://arxiv.org/abs/2004.08262))

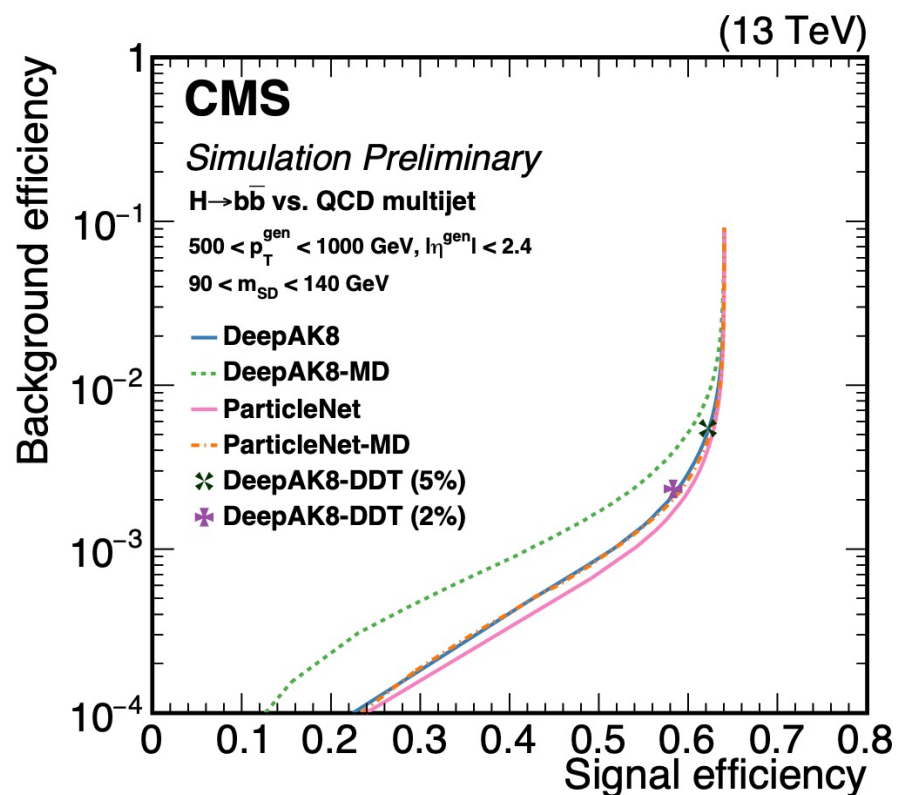
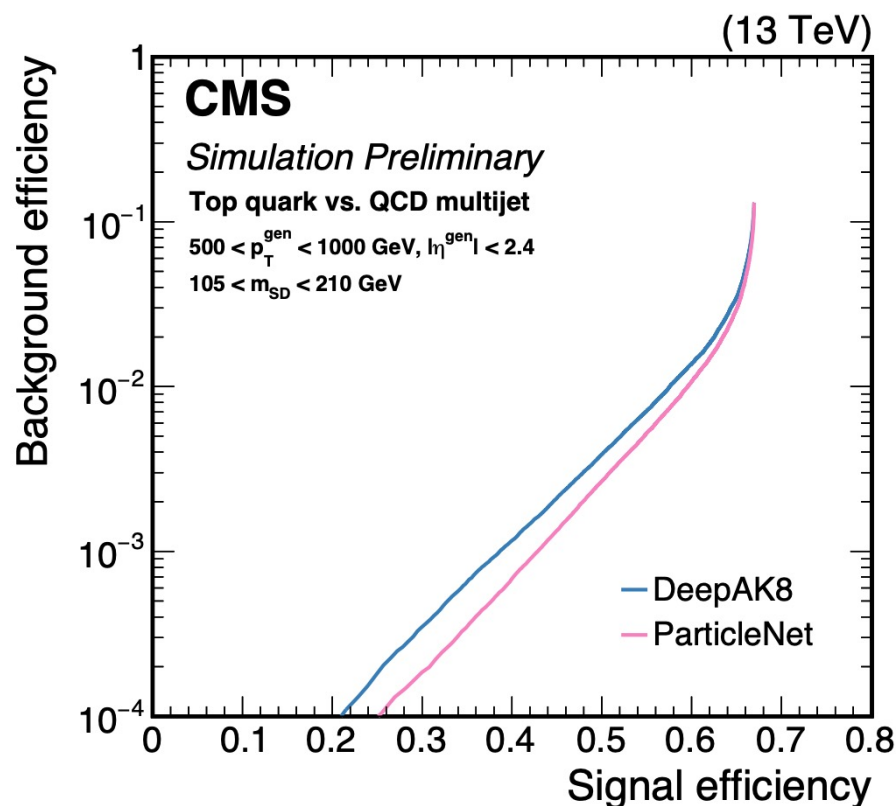


DeepAK8-MD



Large improvements compared to traditional substructure algorithms

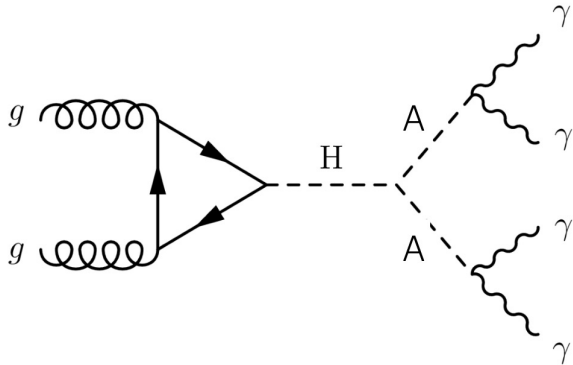
Further improvement in object identification using Graph CNN (ParticleNet)



[CMS DP 2020.002](#)

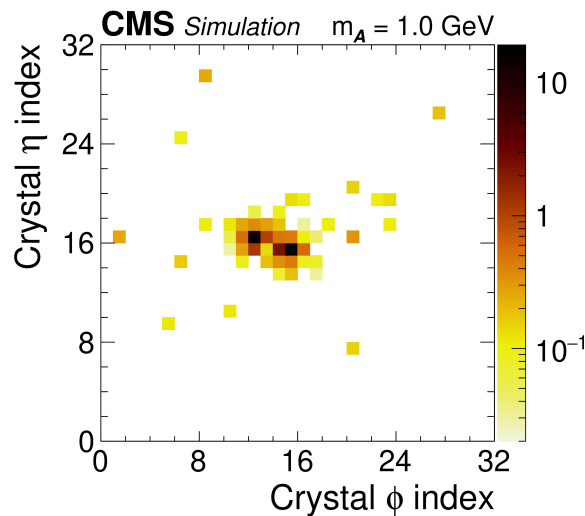
Reconstructing Merged diphotons

[Phys. Rev. D 108 \(2023\) 052002](#)
[PRL 131 \(2023\) 101801](#)

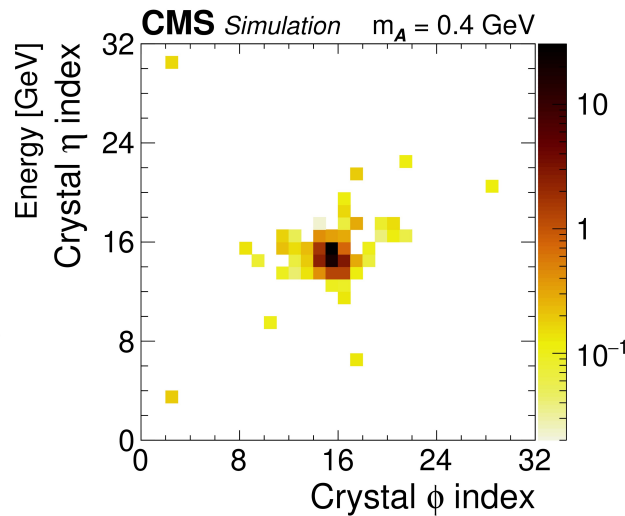


- Impossible to reconstruct using standard reconstruction methods
- Employs a novel end-to-end ML-based particle reconstruction framework, based on **RESNET CNN** architecture
 - Uses minimally processed detector data as input
 - Outputs regressed m_A

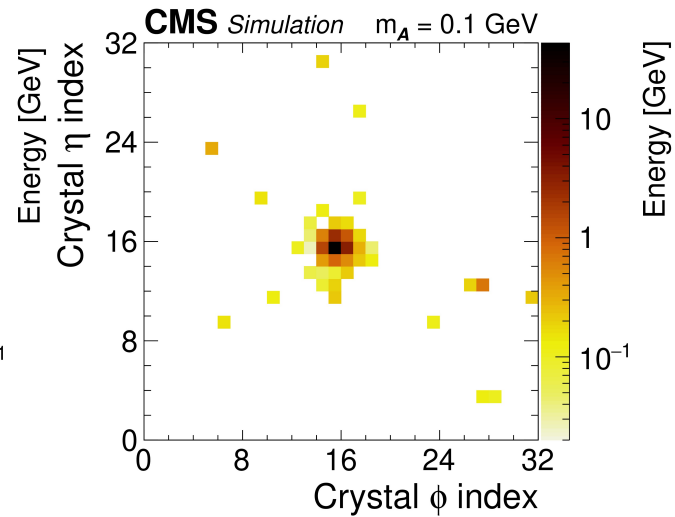
ECAL crystal granularity: $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$



$m_A = 1.0$ GeV, boost = 50



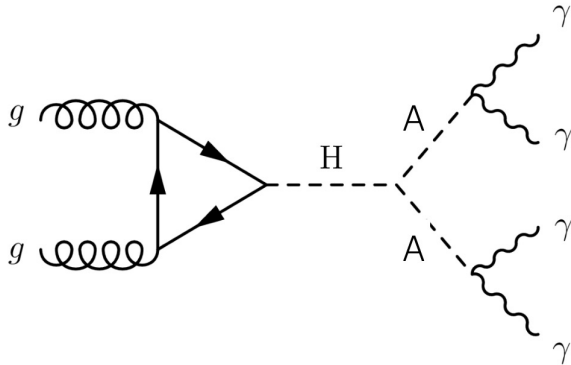
$m_A = 0.4$ GeV, boost = 150



$m_A = 0.1$ GeV, boost = 625

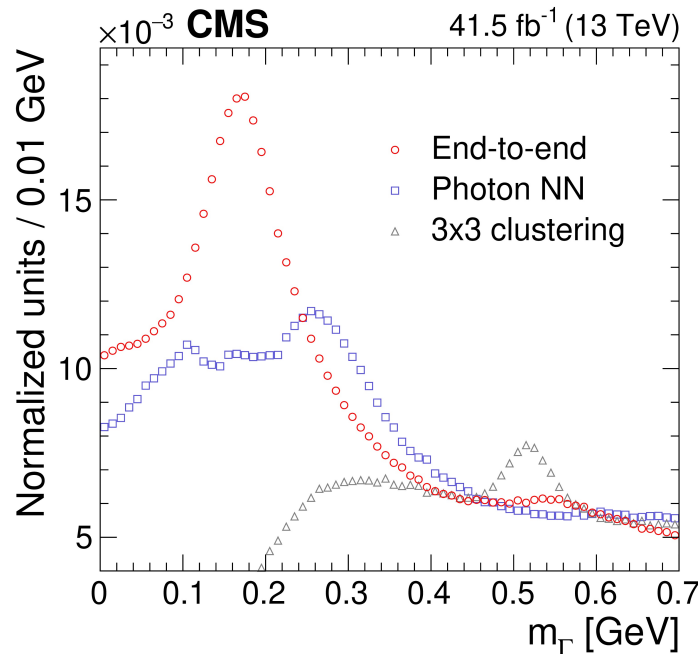
Reconstructing Merged diphotons

Phys. Rev. D 108 (2023) 052002
PRL 131 (2023) 101801

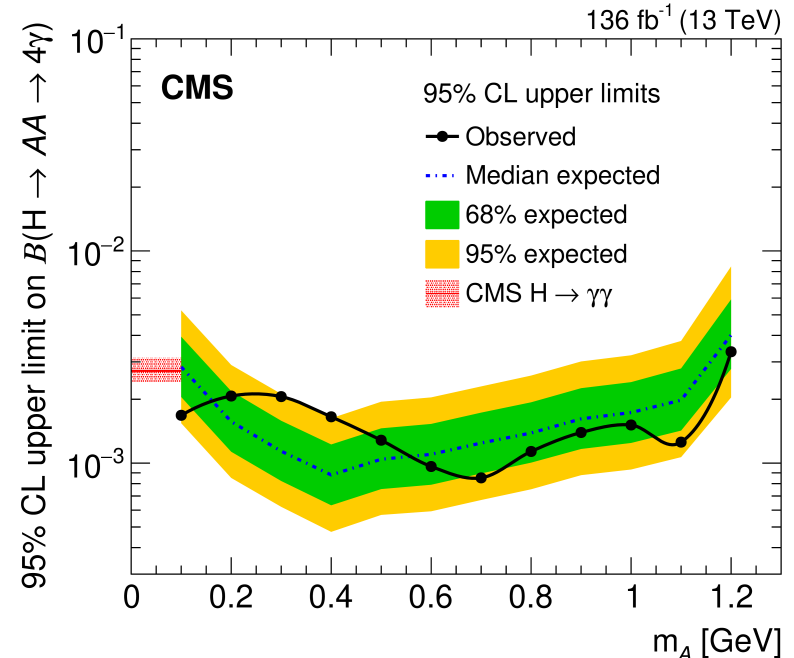


- Impossible to reconstruct using standard reconstruction methods
- Employs a novel end-to-end ML-based particle reconstruction framework, based on **RESNET CNN** architecture
 - Uses minimally processed detector data as input
 - Outputs regressed m_A

Validated using $\pi^0 \rightarrow \gamma\gamma$



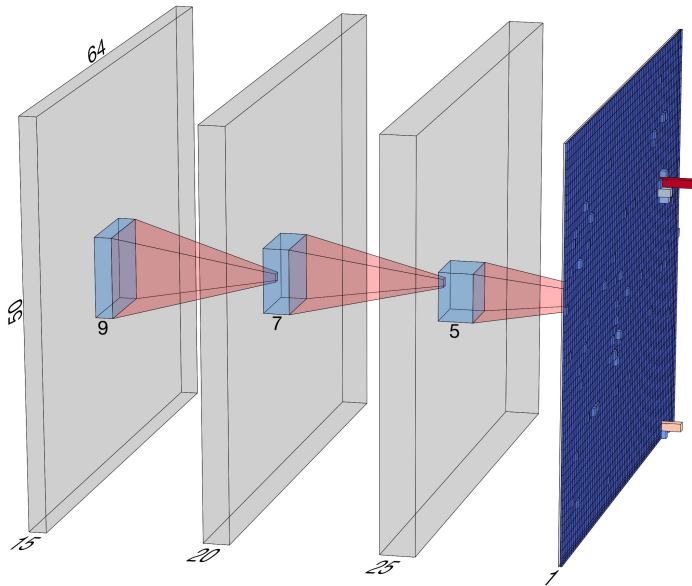
First search of its kind at CMS



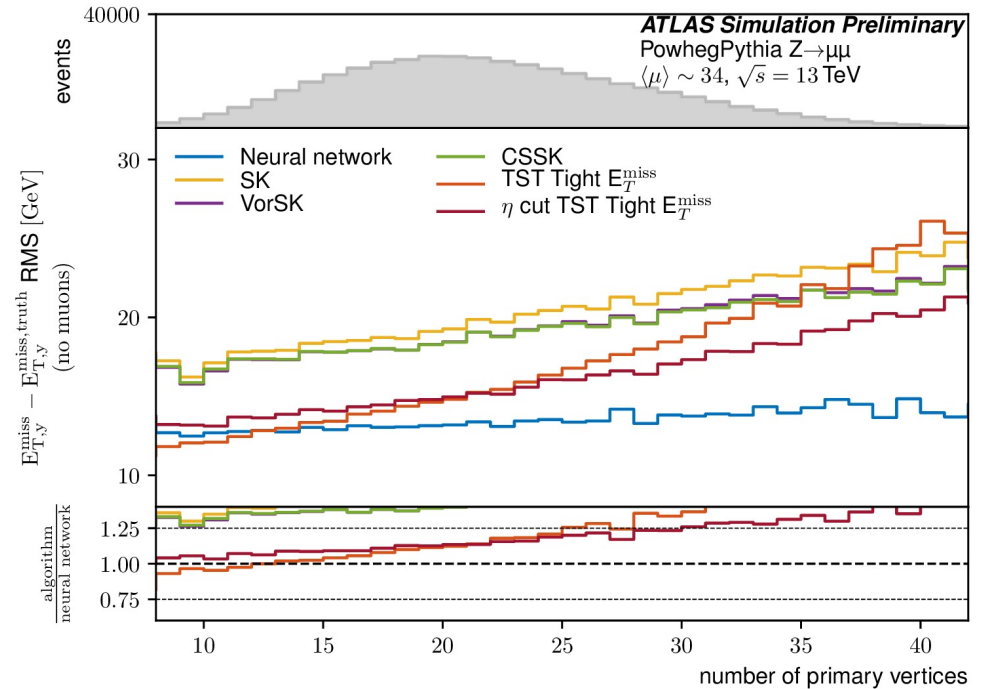
Pileup Mitigation

Pileup mitigation using CNN in ATLAS Missing Transverse Energy Estimation

CNN architecture



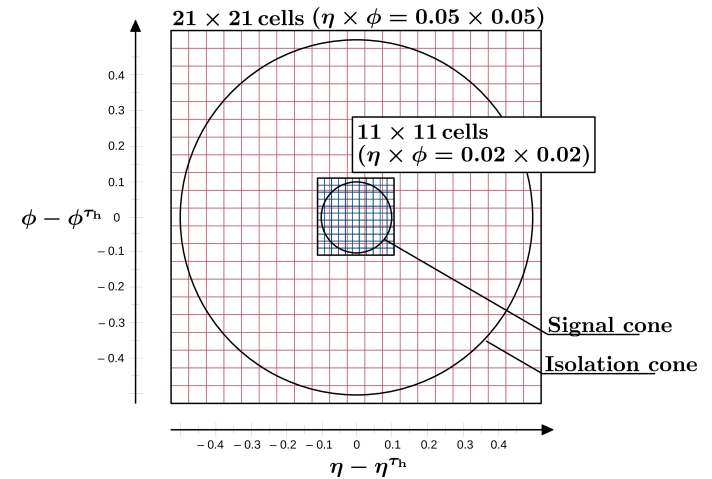
Layer	Kernelsize	Filters	Activation	Parameters
2D convolution	9	15	ReLU	7305
dropout				
2D convolution	7	20	ReLU	14720
dropout				
2D convolution	5	25	ReLU	12525
dropout				
2D convolution	1	1	ReLU	26
Total				34576



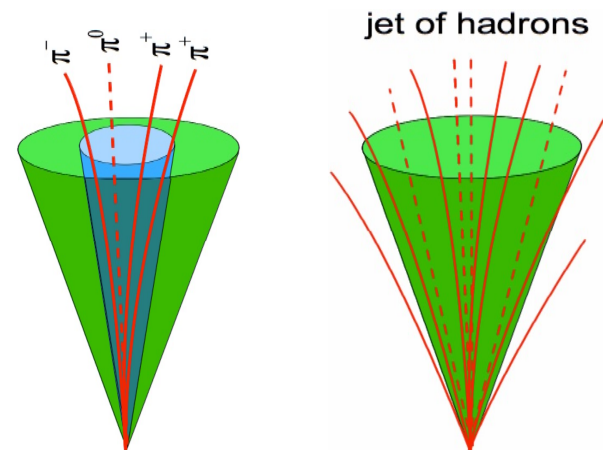
ATL-PHYS-PUB-2019-028

Tau lepton identification

Decay mode	Meson resonance	$\mathcal{B}$ [%]
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
$\tau^- \rightarrow h^- \nu_\tau$		11.5
$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other modes with hadrons		3.2
All modes containing hadrons		64.8

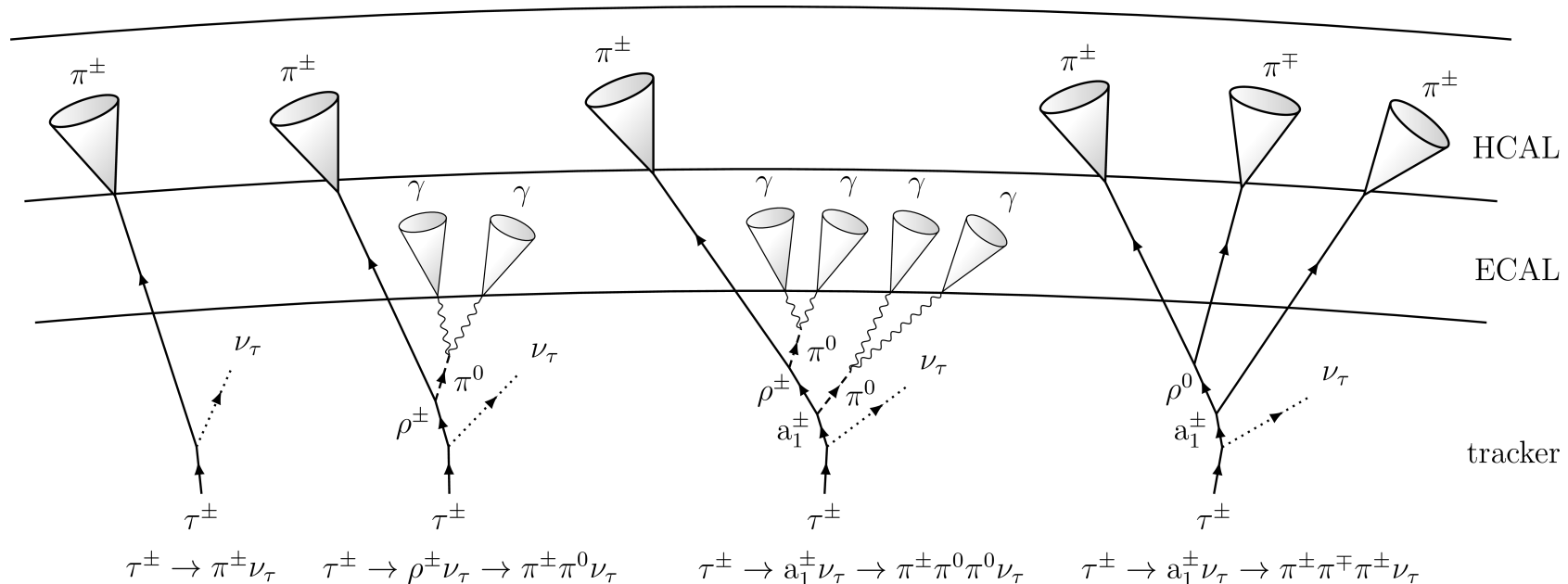
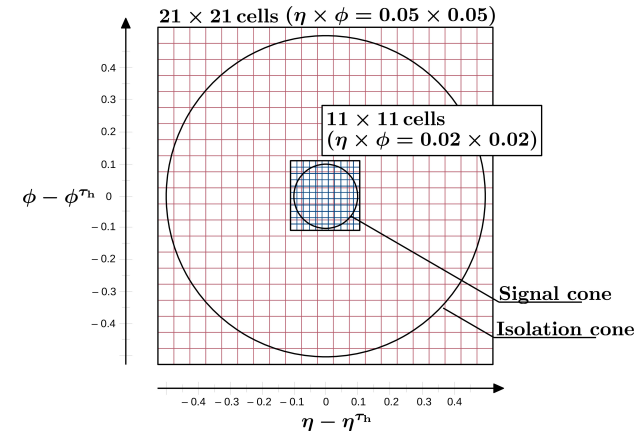
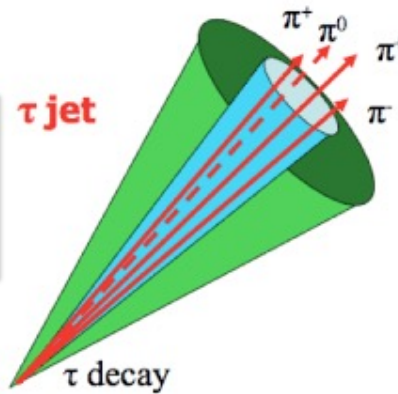


- **τ_h candidates are collimated**
 - A few overlapping $\pi^\pm$ and γ s from π^0 decays
 - Charged tracks can have high impact parameter with respect to primary vertex
 - Presence of secondary vertex (for 3-prong decay)



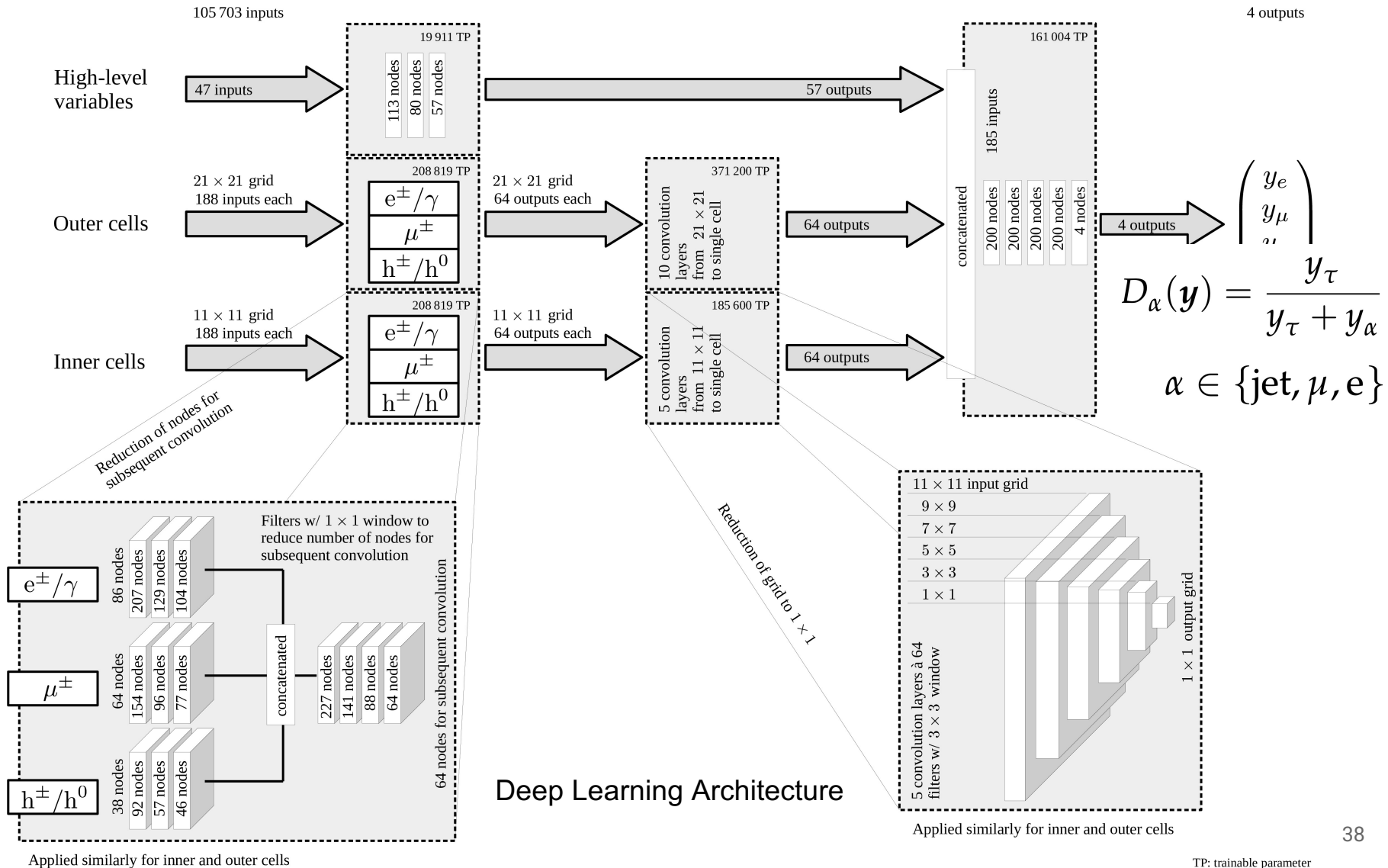
Tau lepton identification

Decay mode	Meson resonance	$\mathcal{B}$ [%]
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
$\tau^- \rightarrow h^- \nu_\tau$		11.5
$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other modes with hadrons		3.2
All modes containing hadrons		64.8

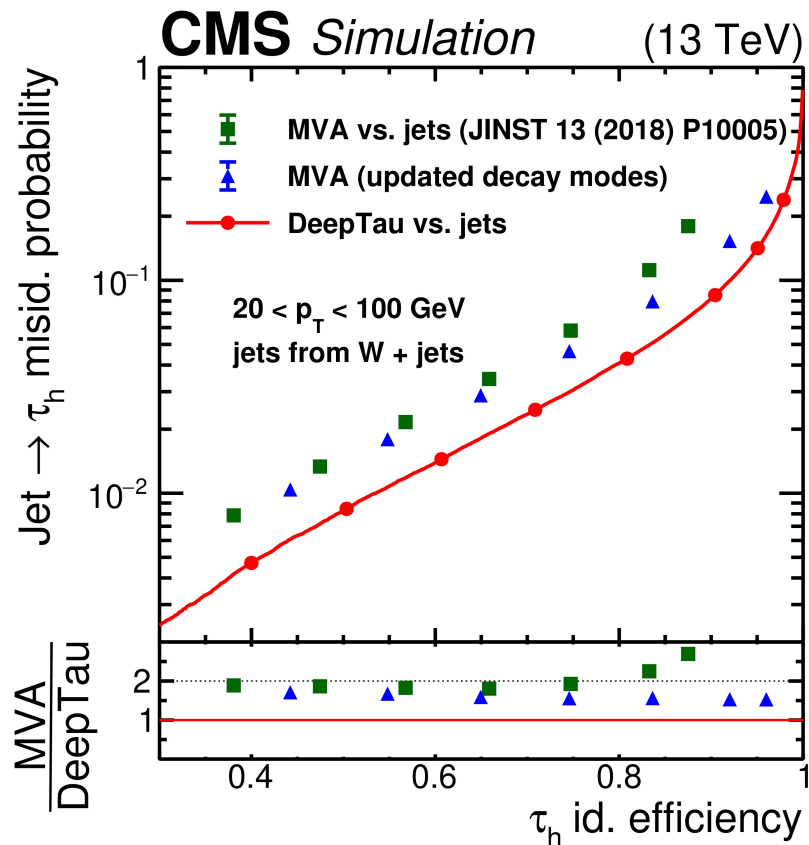
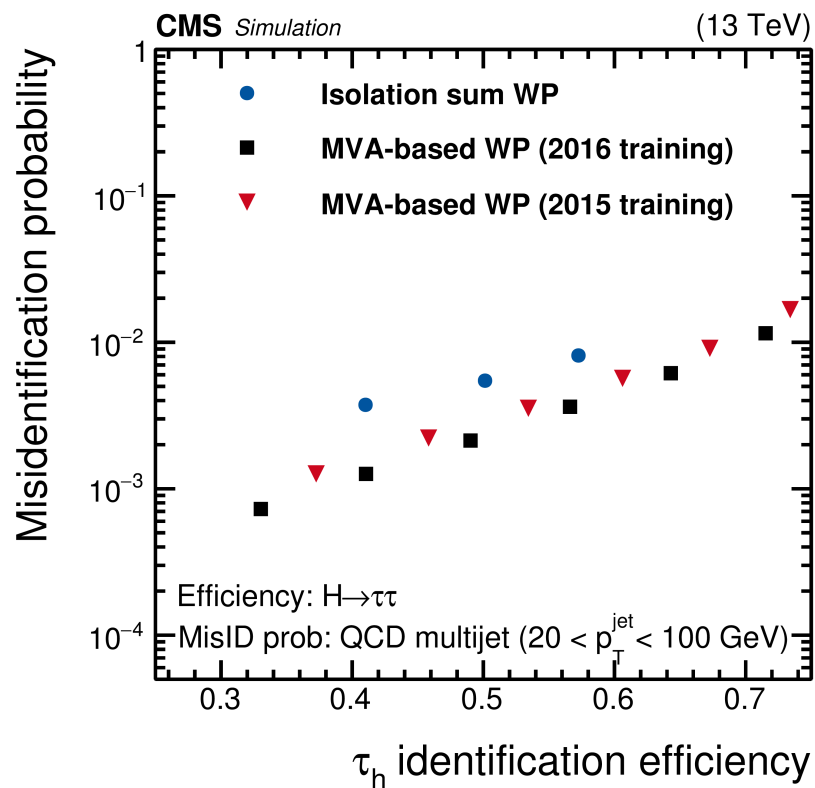


Tau lepton identification

arXiv:2201.08458



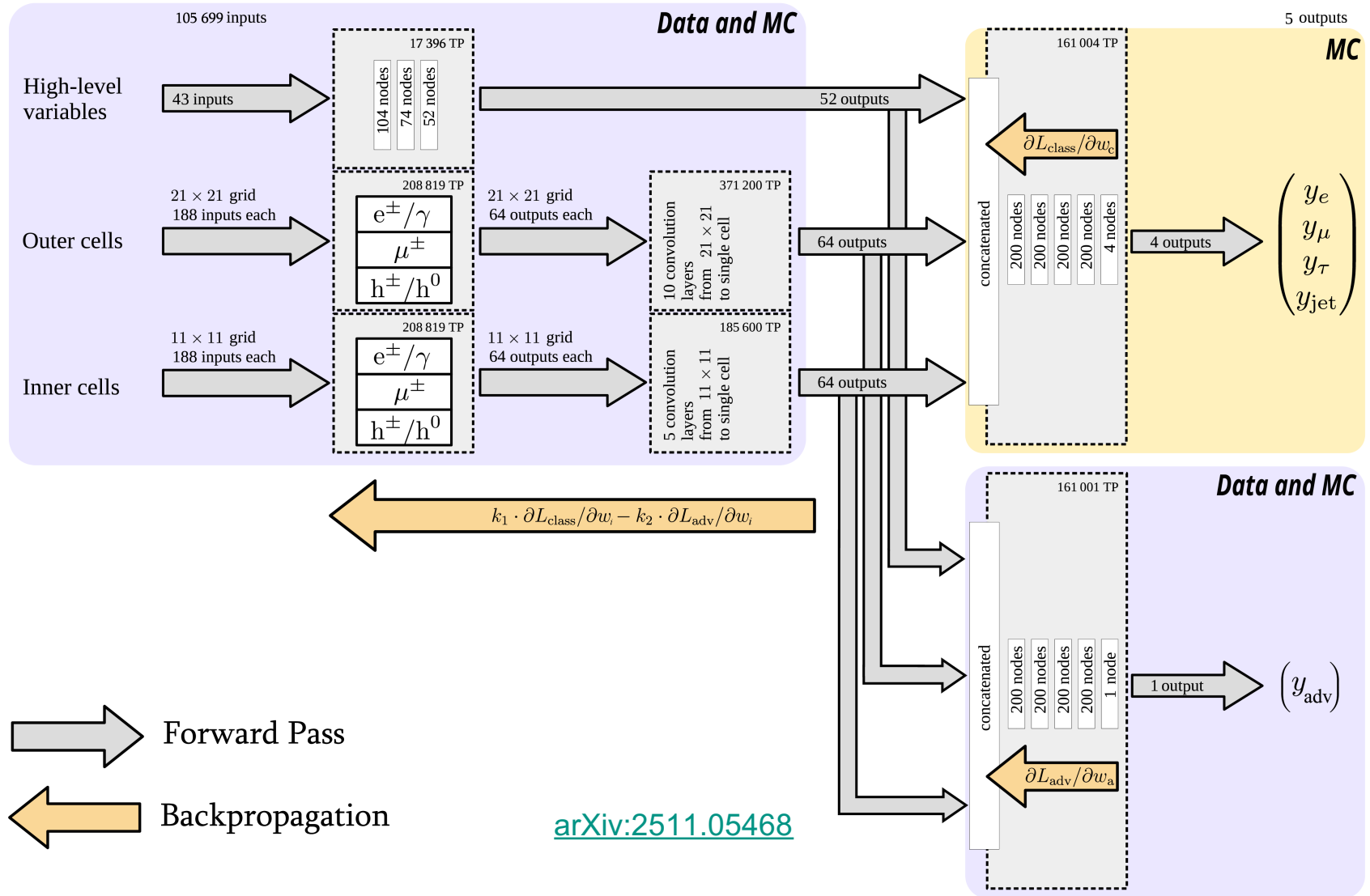
Tau lepton identification



Cut based $\rightarrow$ BDT based $\rightarrow$ Deep Learning

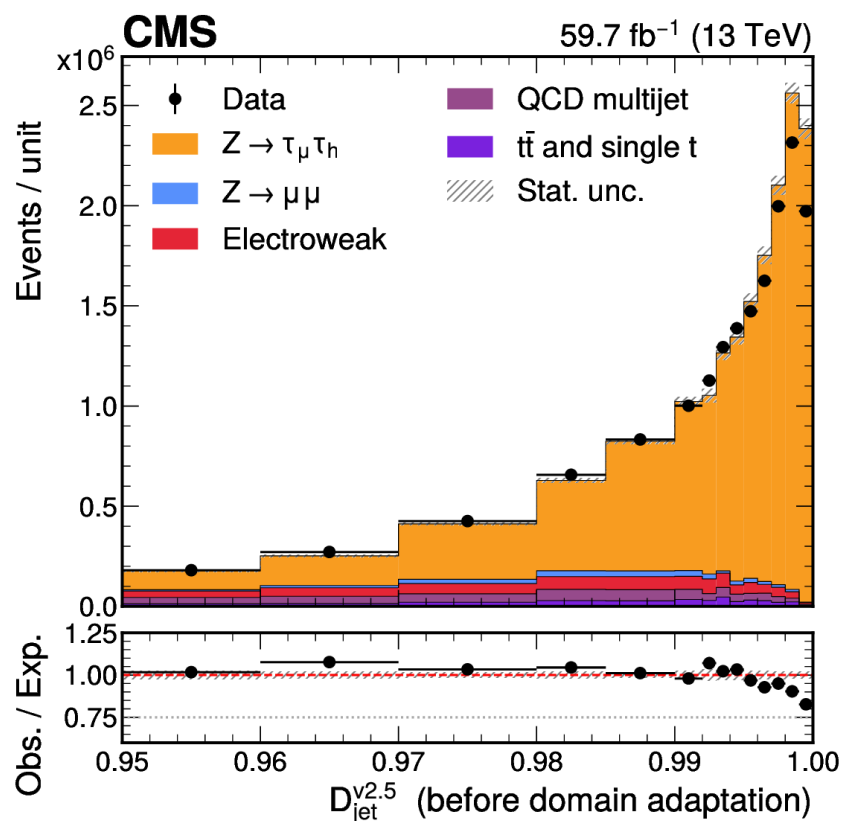
Factor of 2 improvement using low-level inputs through CNN compared using only high level features

Improving Data/MC agreement using domain adaptation

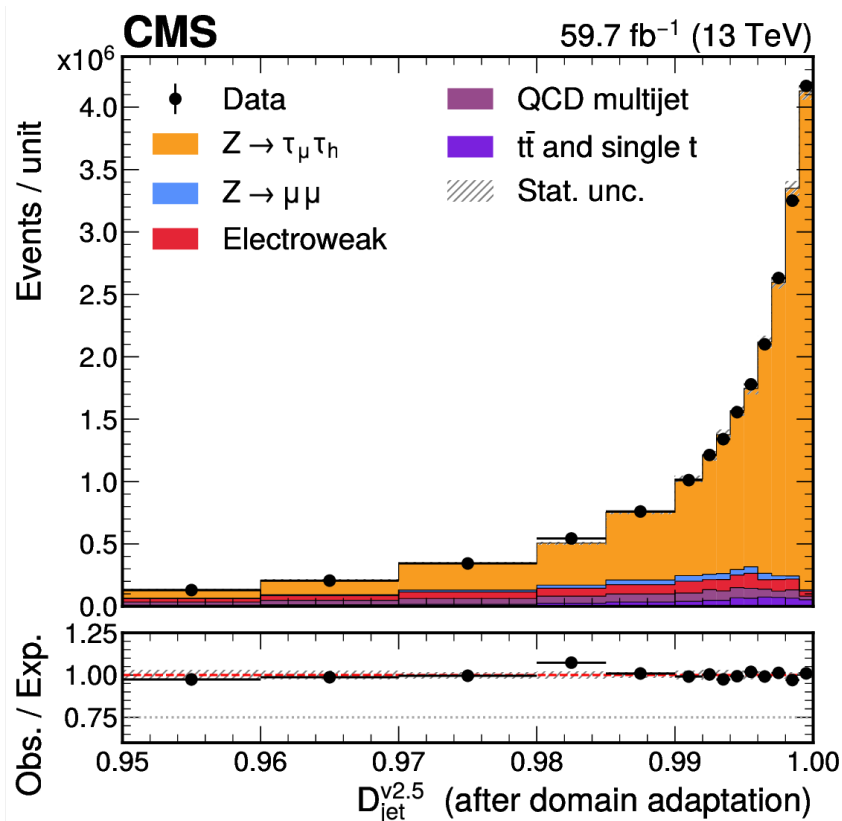


Improves Data/MC agreement for discriminator

Before Domain adaption



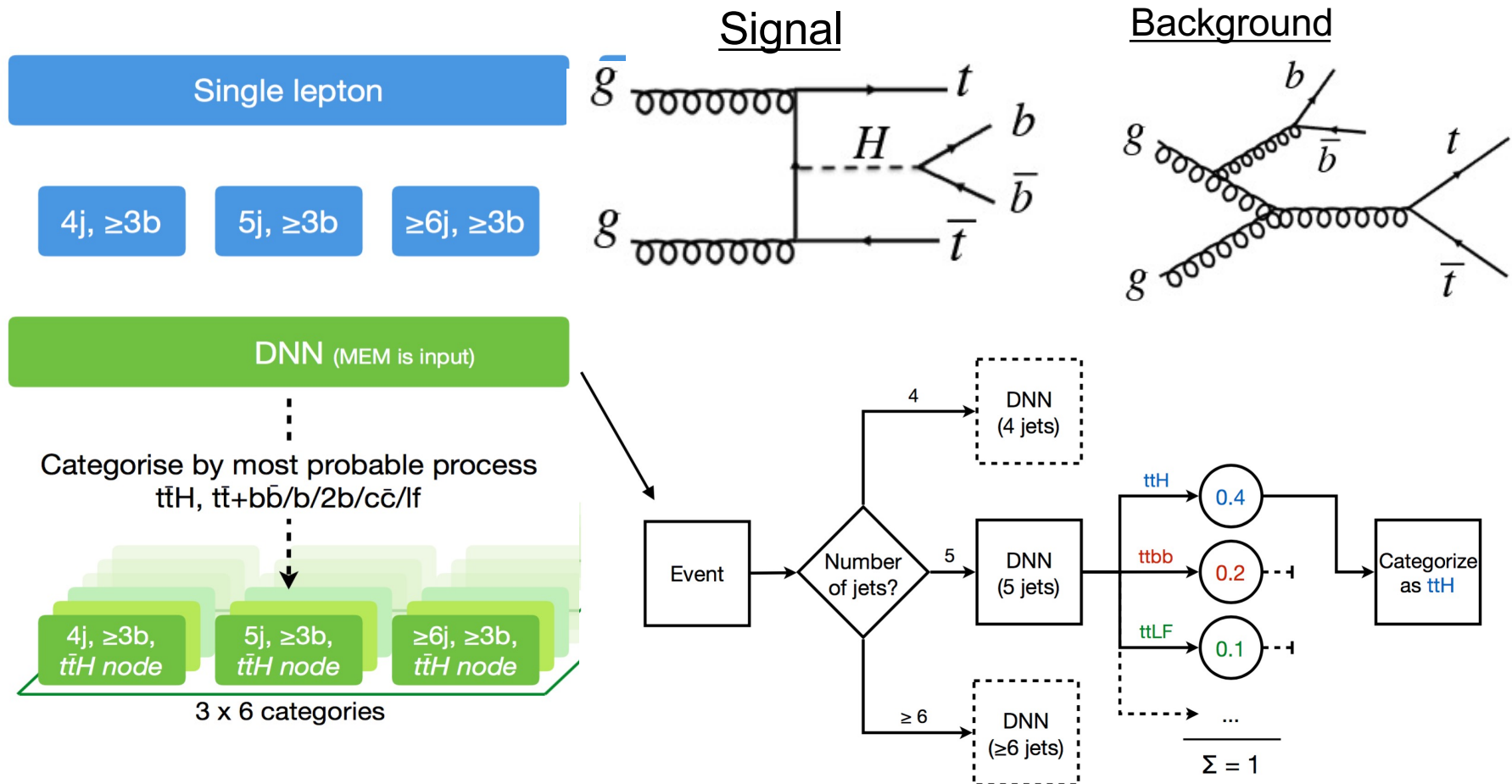
After Domain adaption



Multiclass NN in event classification

JHEP 03 (2019) 026

Separate **tt+Higgs** (signal) from **tt+jets** (background)

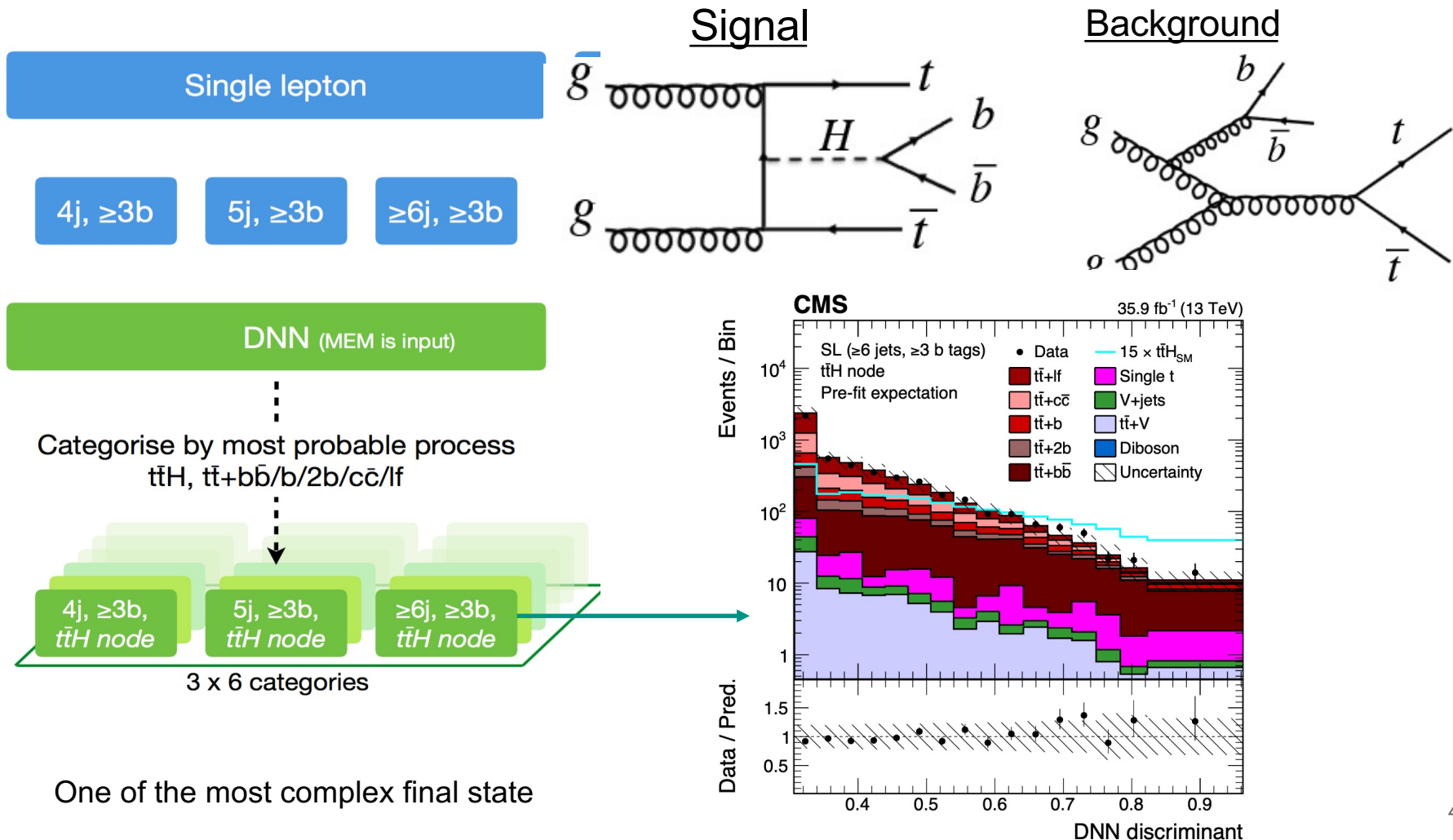


One of the most complex final state

Multiclass NN in event classification

JHEP 03 (2019) 026

Separate **tt+Higgs** (signal) from tt+jets (background)



ML Based Particle Flow Reconstruction

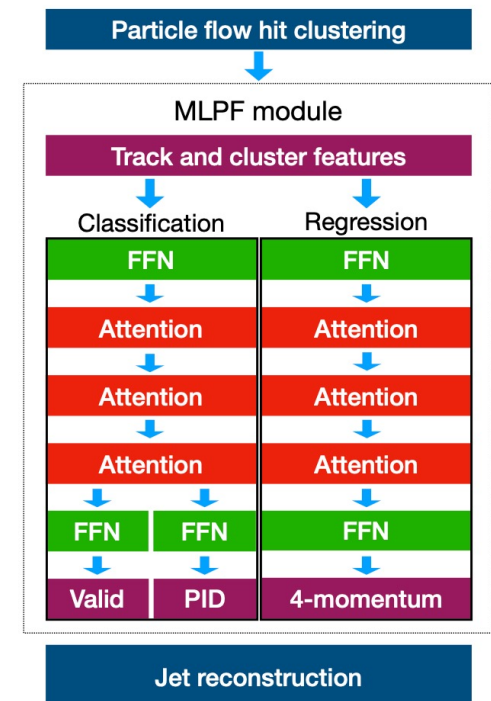
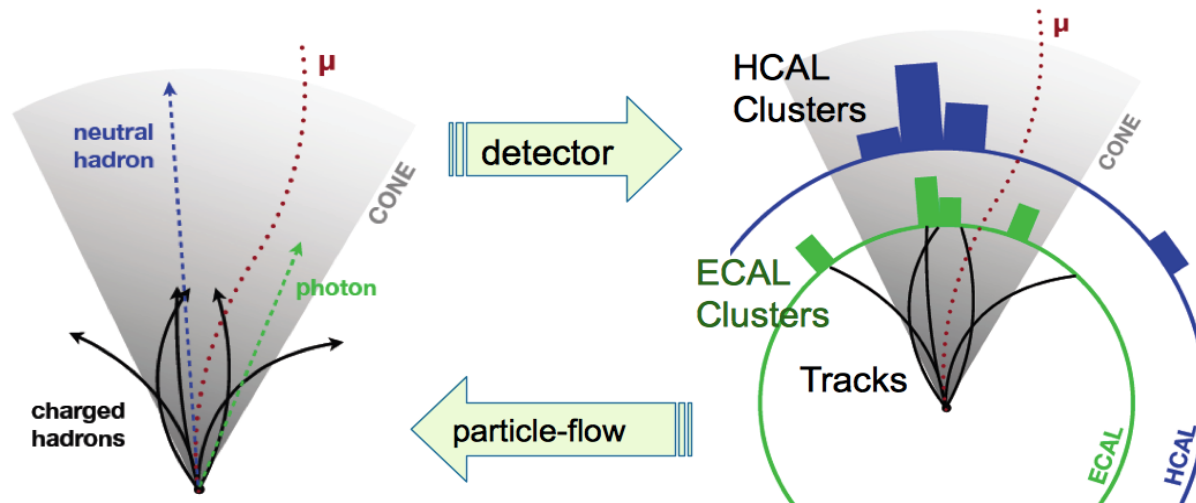
Global event reconstruction → reconstruct and identify each individual particle in an event

Current PF algorithm:

- Iterative, rule-based methods, such as proximity-based linking
 - i.e. tracks and calorimeter clusters are associated based on geometric closeness in detector coordinates
- Tries to optimally combine all detector information

ML for Particle-Flow:

- Uses a transformer based architecture
- Predicts all particles simultaneously at one inference step
- Expected to benefit more in HL-LHC, with very high Pileup and more granular detector
- Validated using Run-3 data

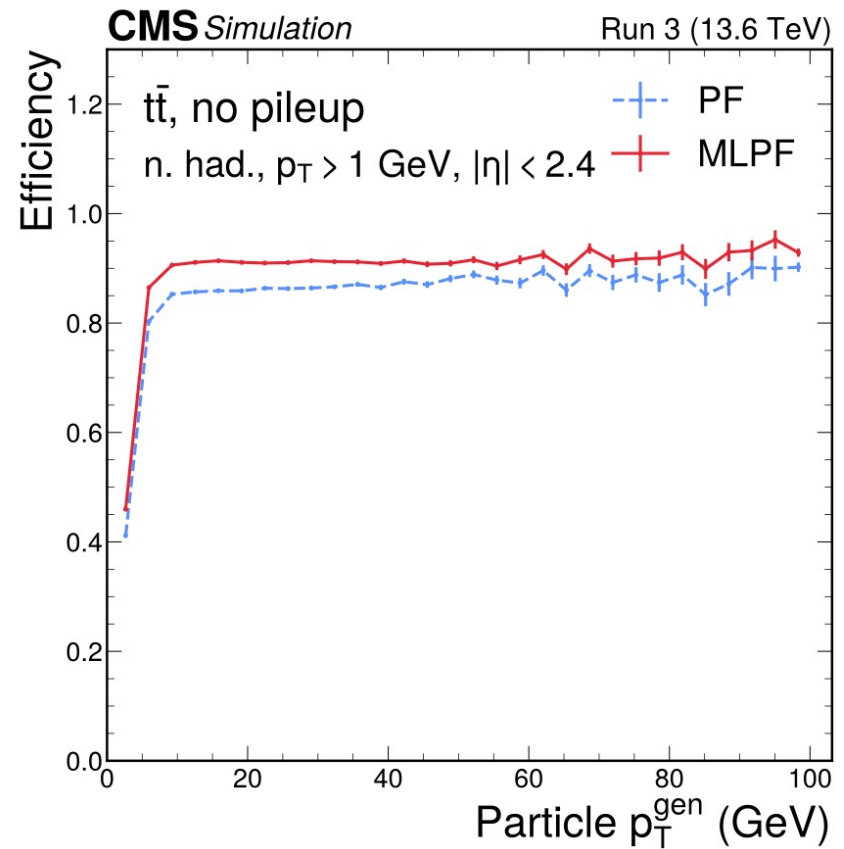
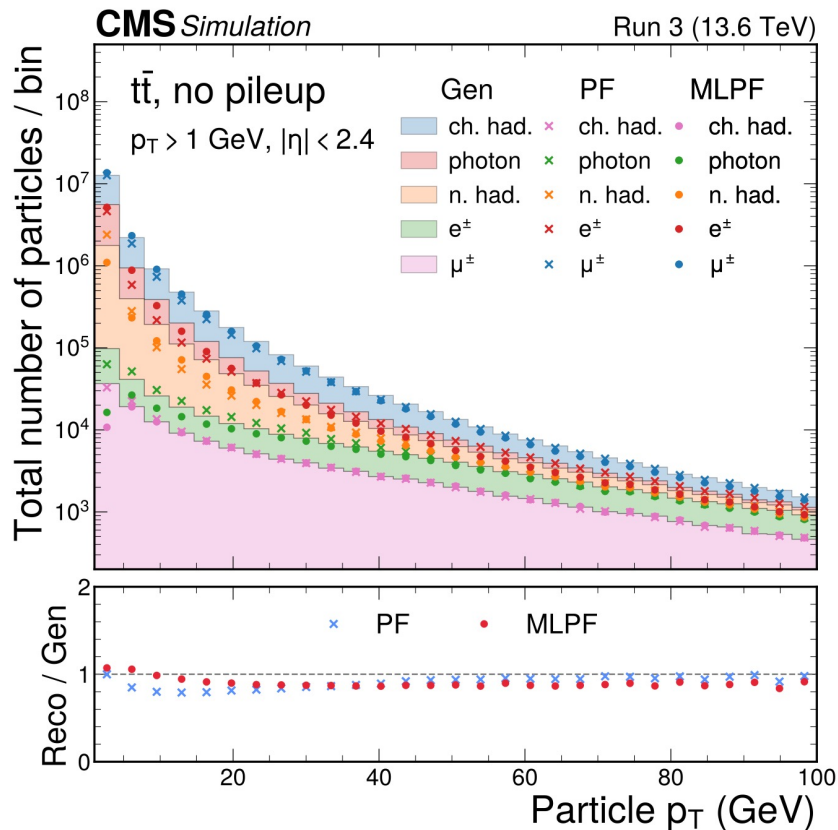


ML Based Particle Flow Reconstruction

[CMS-PFT-25-001](#)

Performance validated using $t\bar{t}$ and QCD multijets events

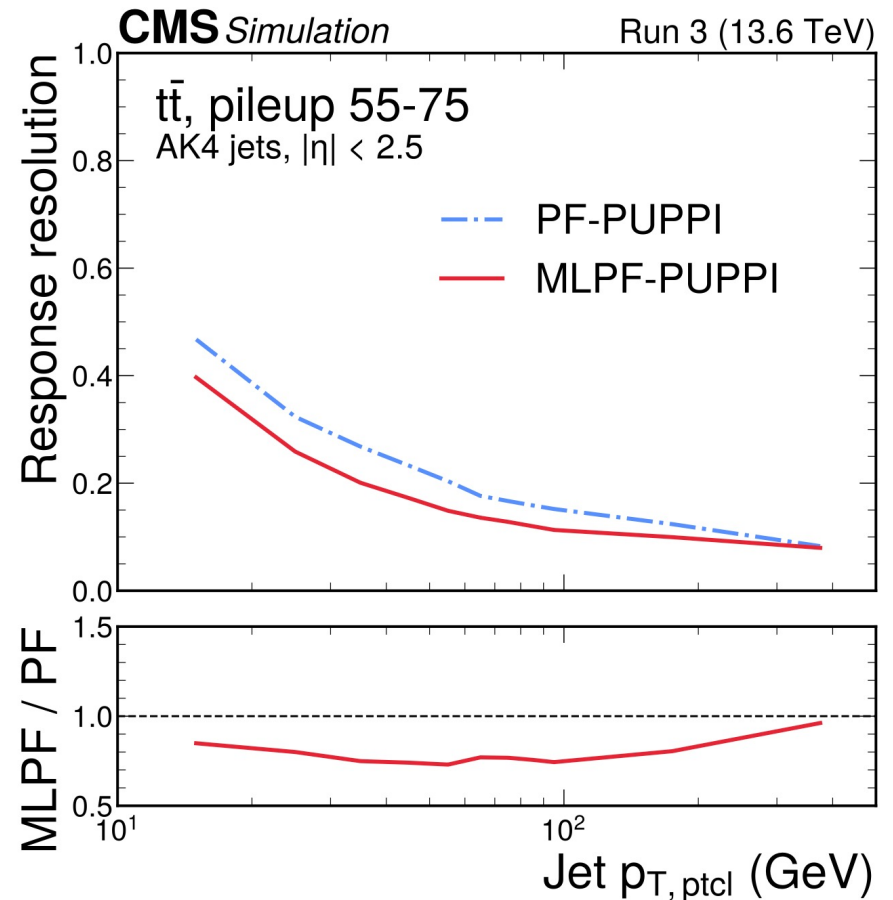
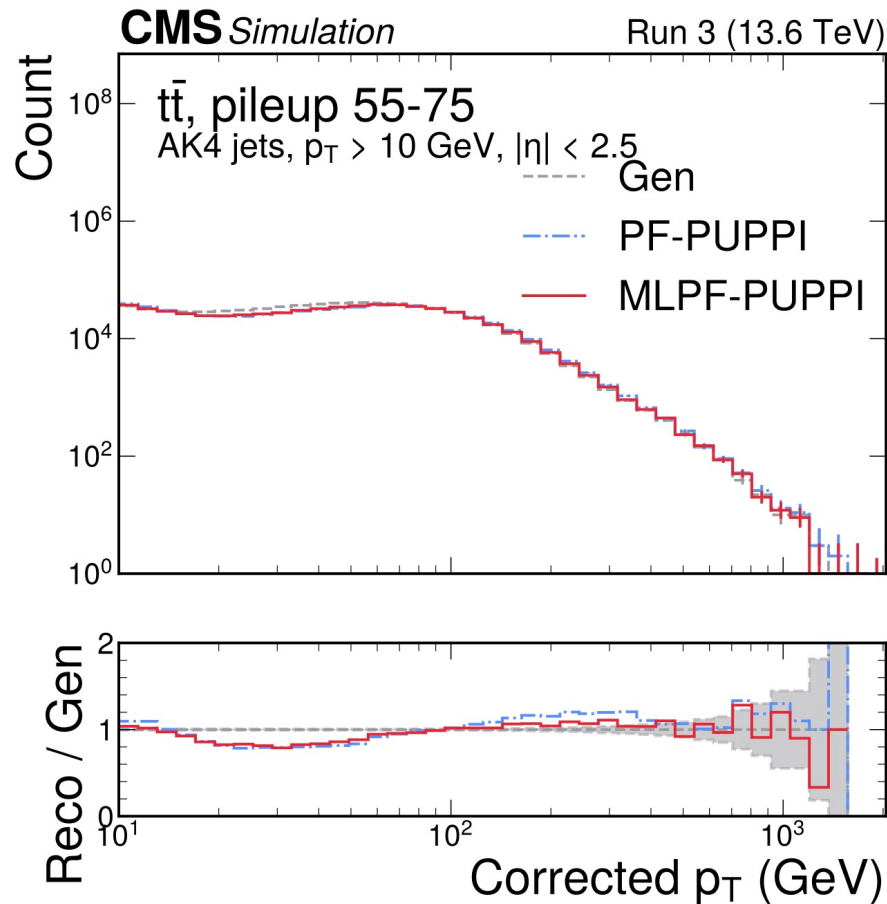
- Similar performance at particle level reconstruction & identification
- Improved efficiency for neutral hadron for similar fake rate



ML Based Particle Flow Reconstruction

[CMS-PFT-25-001](#)

Similar jet response compared to PF, and improved jet p_T resolution

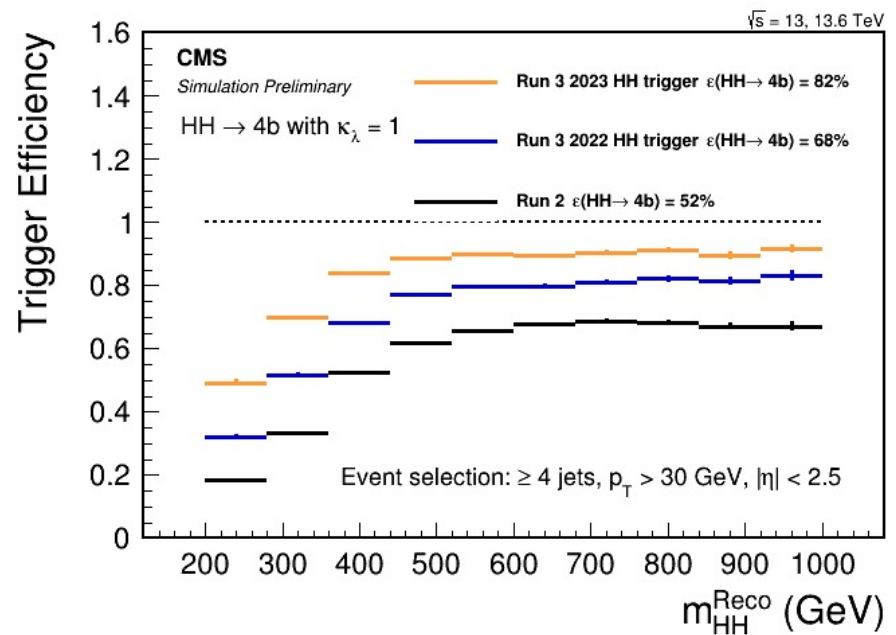


b-tagging @ HLT

Trigger	Requirement	Rates at HLT at $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
2023 HH trigger	HT > 280 GeV, 4 jets with $p_T > 30 \text{ GeV}$, PNet@AK4(mean 2 highest b-tag score) > 0.55	180 Hz
2022 HH trigger	4 jets $p_T > 70, 50, 40, 35 \text{ GeV}$, PNet@AK4(mean 2 highest b-tag score) > 0.65	60 Hz
2018 triple b-tag [2,3]	HT > 340 GeV, 4 jets $p_T > 75, 60, 45, 40 \text{ GeV}$, 3 b-tags with DeepCSV > 0.24	8 Hz

Improved efficiency at HLT in
Run-3 using ParticleNet

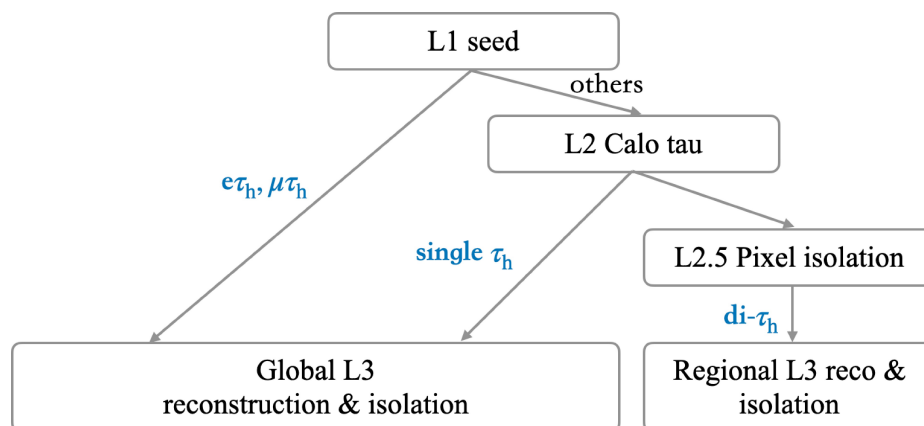
[CMS-DP-2023-050](#)



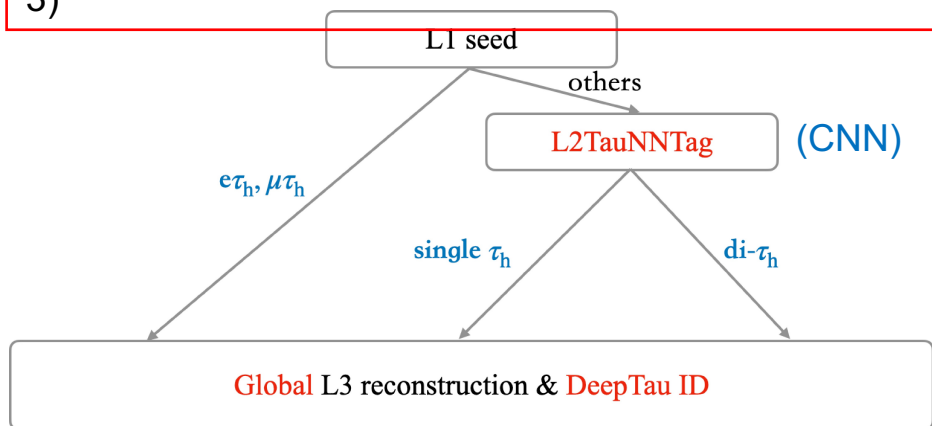
Improved Tau identification at High Level Trigger

[arXiv:2602.11359](https://arxiv.org/abs/2602.11359)

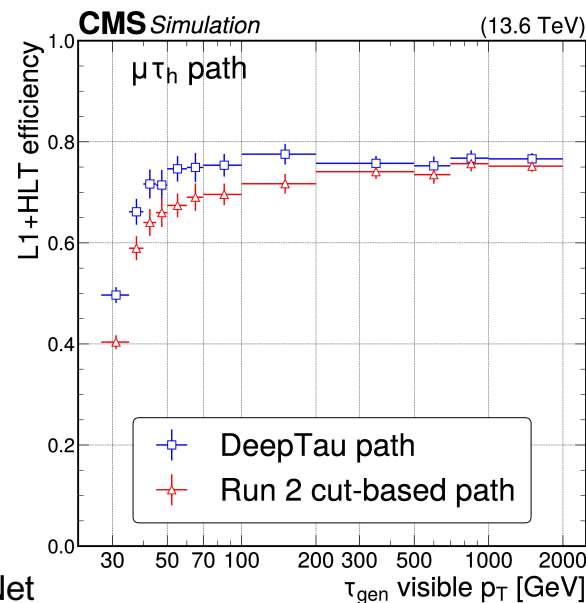
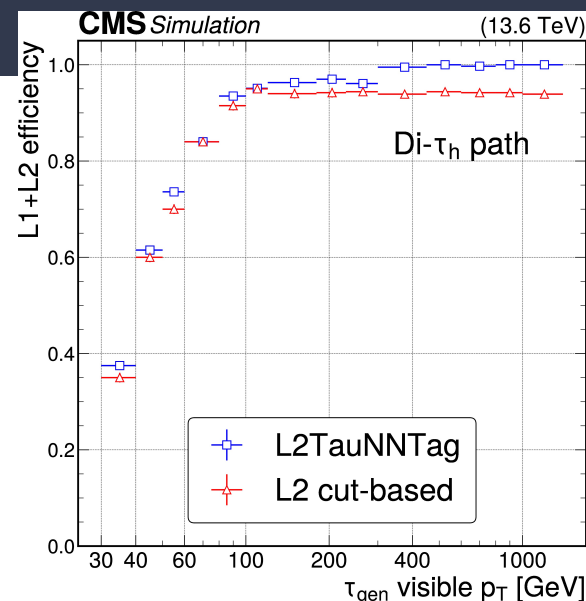
τ_h candidate reconstruction at the HLT in Run 2



τ_h candidate reconstruction at the HLT in 2022-23 (Run-3)



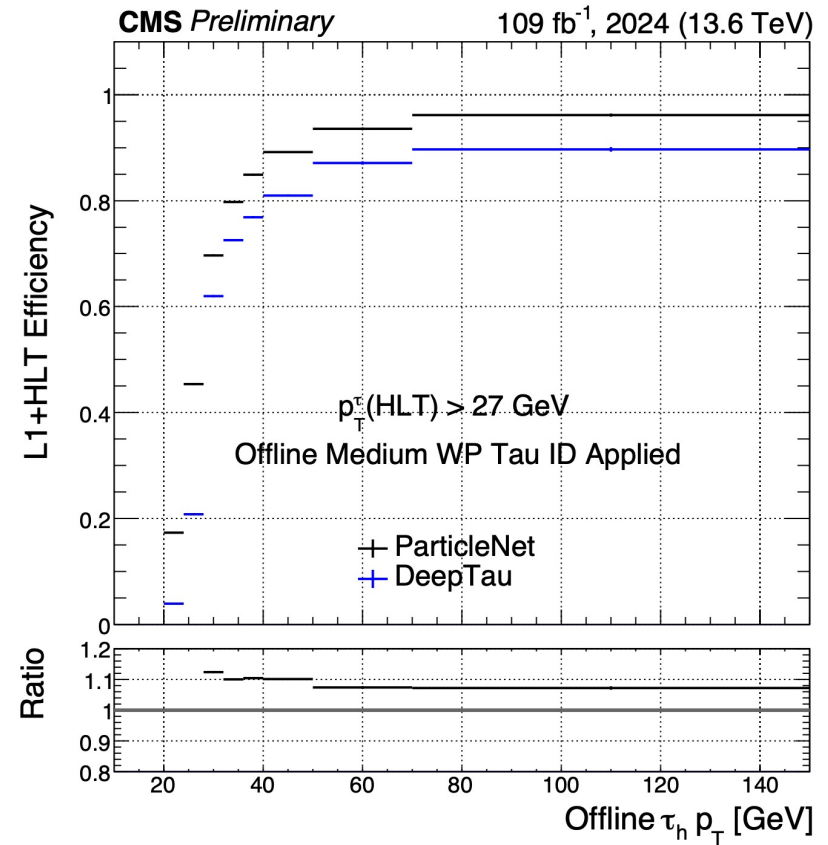
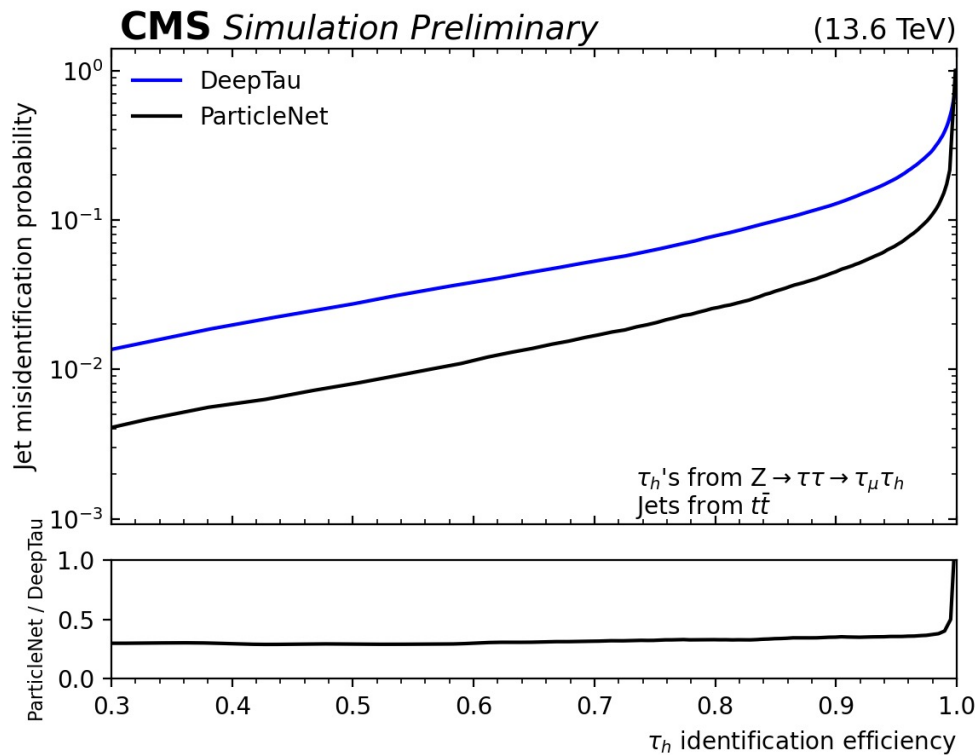
Further improvement in 2024-25 using ParticleNet



Improvement in Tau ID at HLT using ParticleNet

- Implemented since 2024
- Higher efficiency than DeepTau with comparable rate

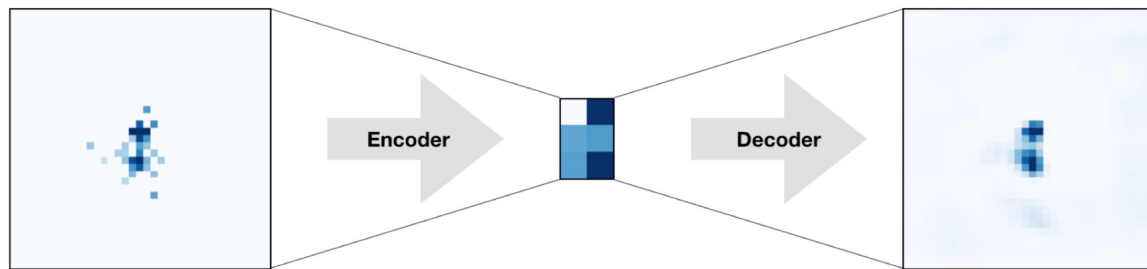
[CMS-DP-2026/008](#)



Anomaly Detection

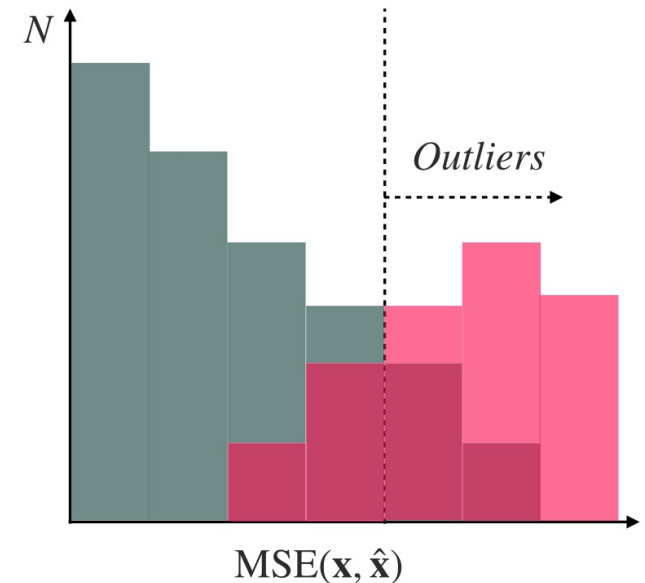
Unsupervised = no labels

e.g. Autoencoders



[arxiv:1808.08992](https://arxiv.org/abs/1808.08992)
[arXiv:1808.08979](https://arxiv.org/abs/1808.08979)

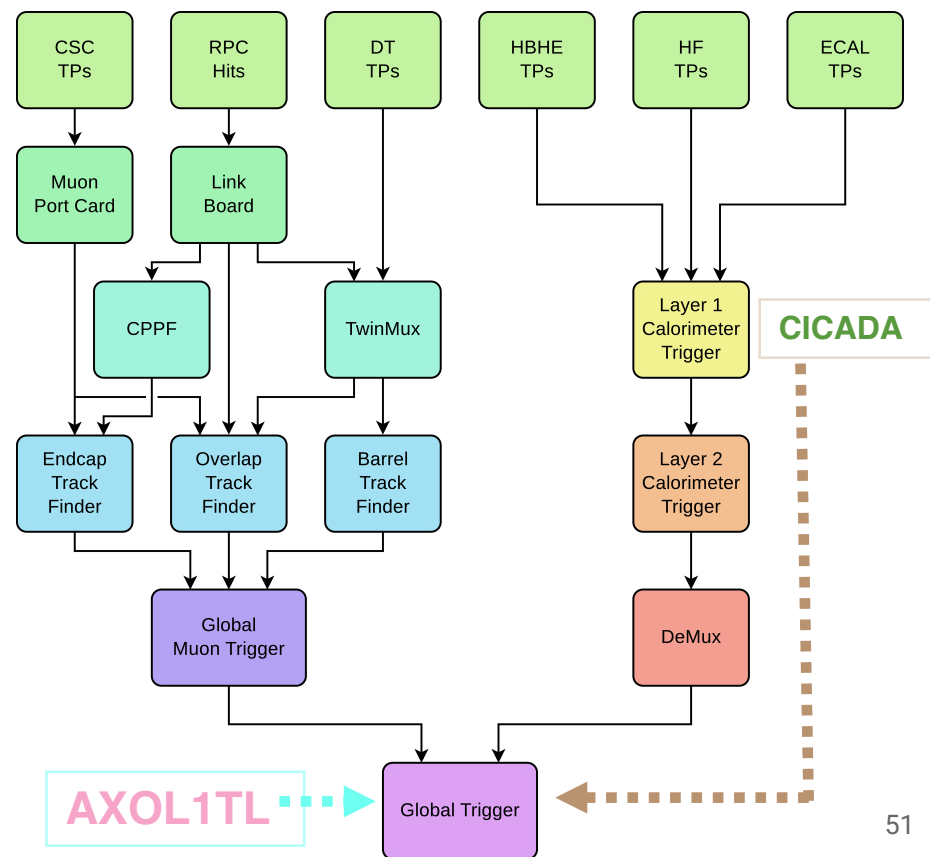
- Train network Compress events and then decompress them to get the original input
 - Anomaly events should have higher loss
- Can **train directly on data**, no labels needed
 - New physics searches in a model independent way
 - **Useful for triggering** events with unknown signatures



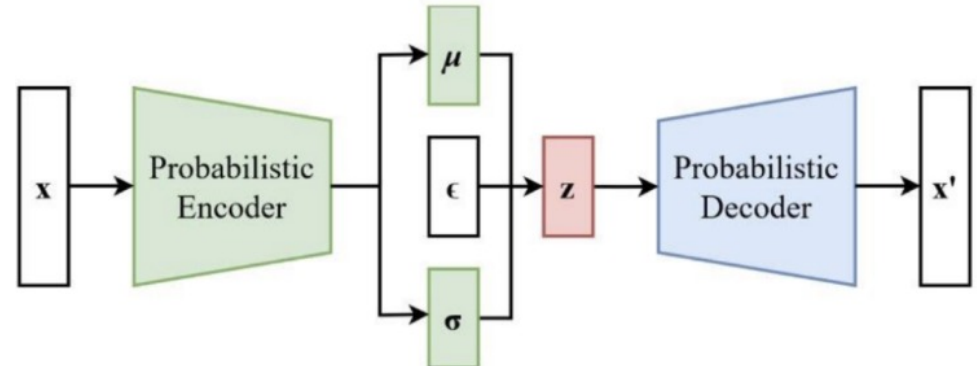
ML based anomaly detection trigger

- Trigger is essential for recording data at LHC
 - But most triggers are designed aiming for specific signatures → Model dependent
 - Maybe new physics events NOT simply selected at the trigger level

- ML-based anomaly triggers first time in CMS @ LHC Run-3
 - **A**nomaly **eX**traction **O**nline **L1** **T**rigger **L**ightweight in L1Global Trigger
 - **C**alorimeter **I**mage **C**onvolutional **A**nomaly **D**etection **A**lgorithm in L1 Calo Layer 1
- Implementation in FPGA using [HLS4ML](#)



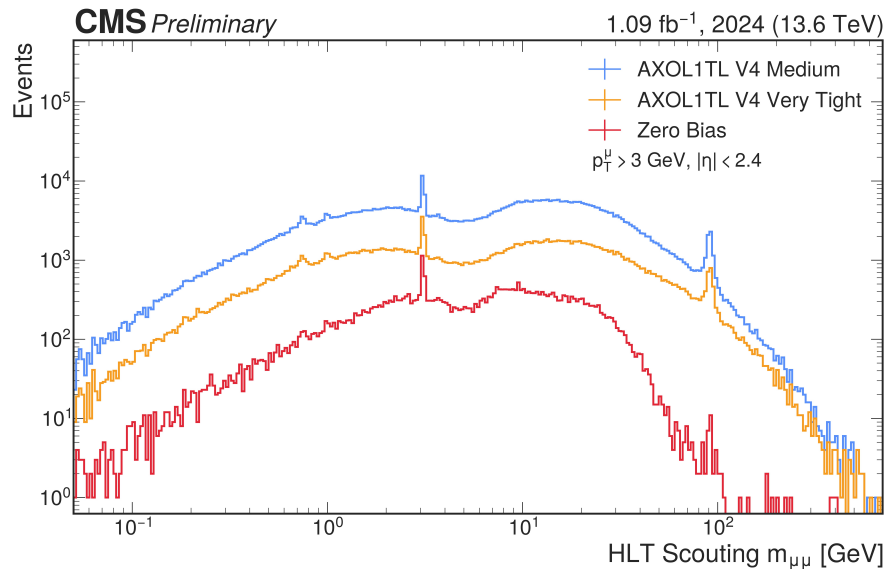
Variational autoencoder (VAE) trained on ZeroBias data to detect data outliers



$$\text{Loss} = (1 - \beta) \|x - \hat{x}\|^2 + \beta \frac{1}{2} (\mu^2 + \sigma^2 - 1 - \log \sigma^2)$$

Reconstruction term

Full regularization term



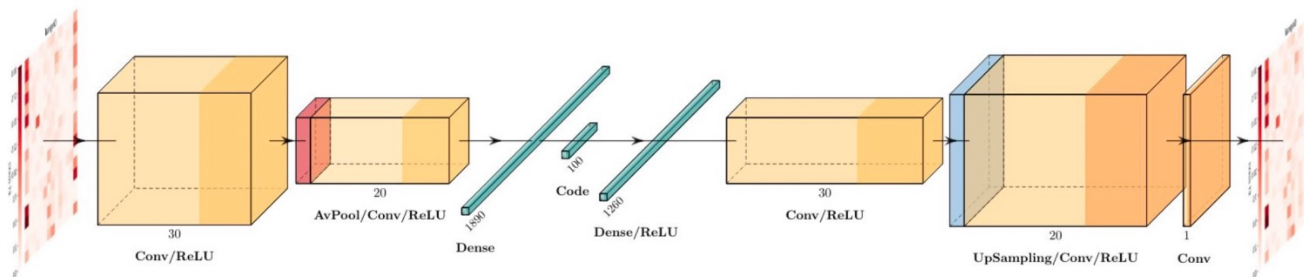
Invariant mass distribution of muon pairs collected from HLT Scouting events by different [AXOL1TL V4](#) trigger paths

[CMS-DP-2025-061](#)

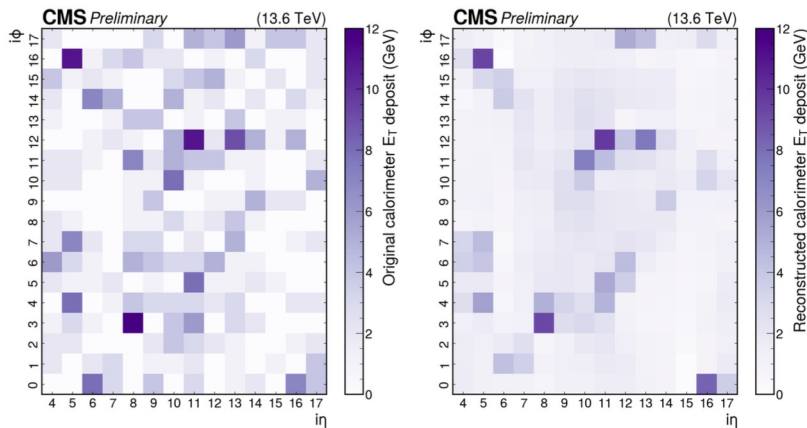
Convolutional neural network auto-encoder

- Inputs from CaloLayer1
- Low level information, not dependent on jet reconstruction etc..
- MSE as anomaly score

CMS DP -2024/121



ZeroBias

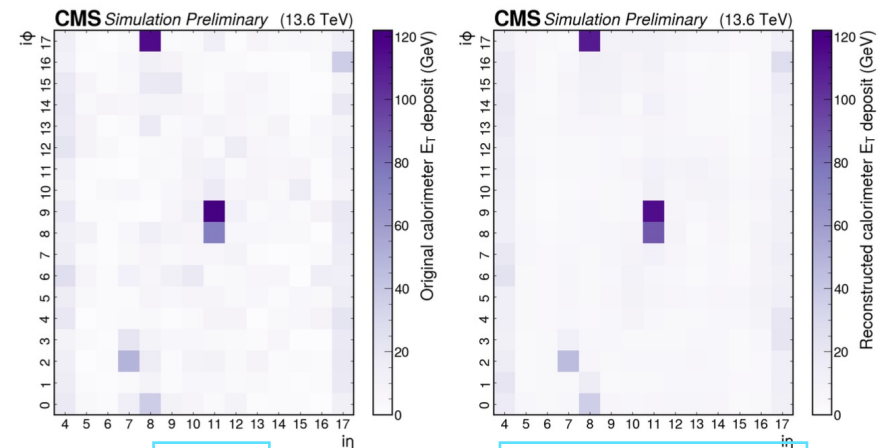


Input

Reconstructed

MSE = 2.57

SUSYGGBBH signal



Input

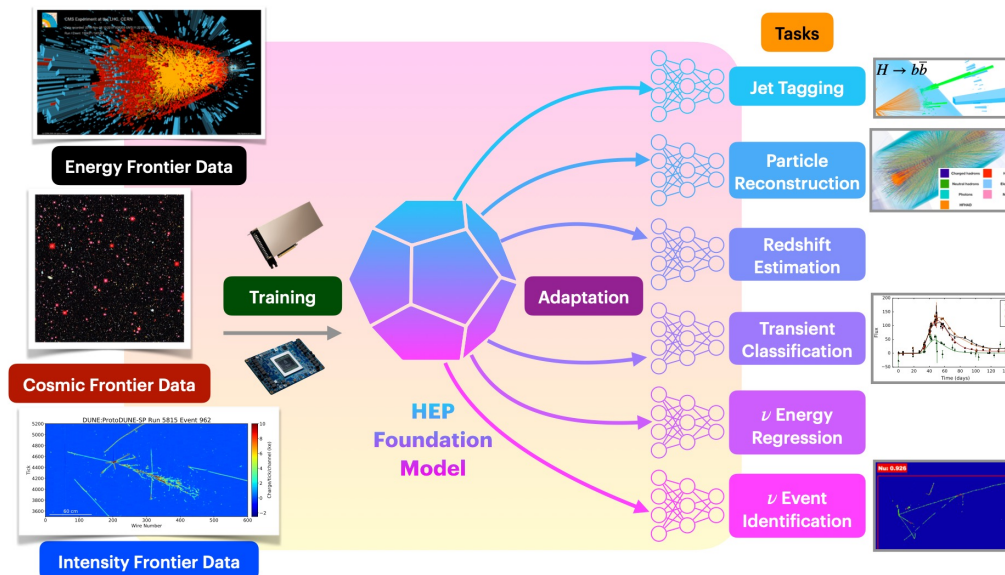
Reconstructed

MSE = 14.89

Summary & Outlook

- Rapid development of Machine Learning techniques in past decade
- Deep learning revolutionized the way large data samples are analyzed in HEP experiments
 - Only a very few examples discussed here (for more see [A Living Review of Machine Learning for Particle Physics](#))

Possibility to have Foundation Models for HEP data analysis in future



Schematic of a foundation model based on HEP data ([arxiv:2512.11133](#))

*Thank
you!*