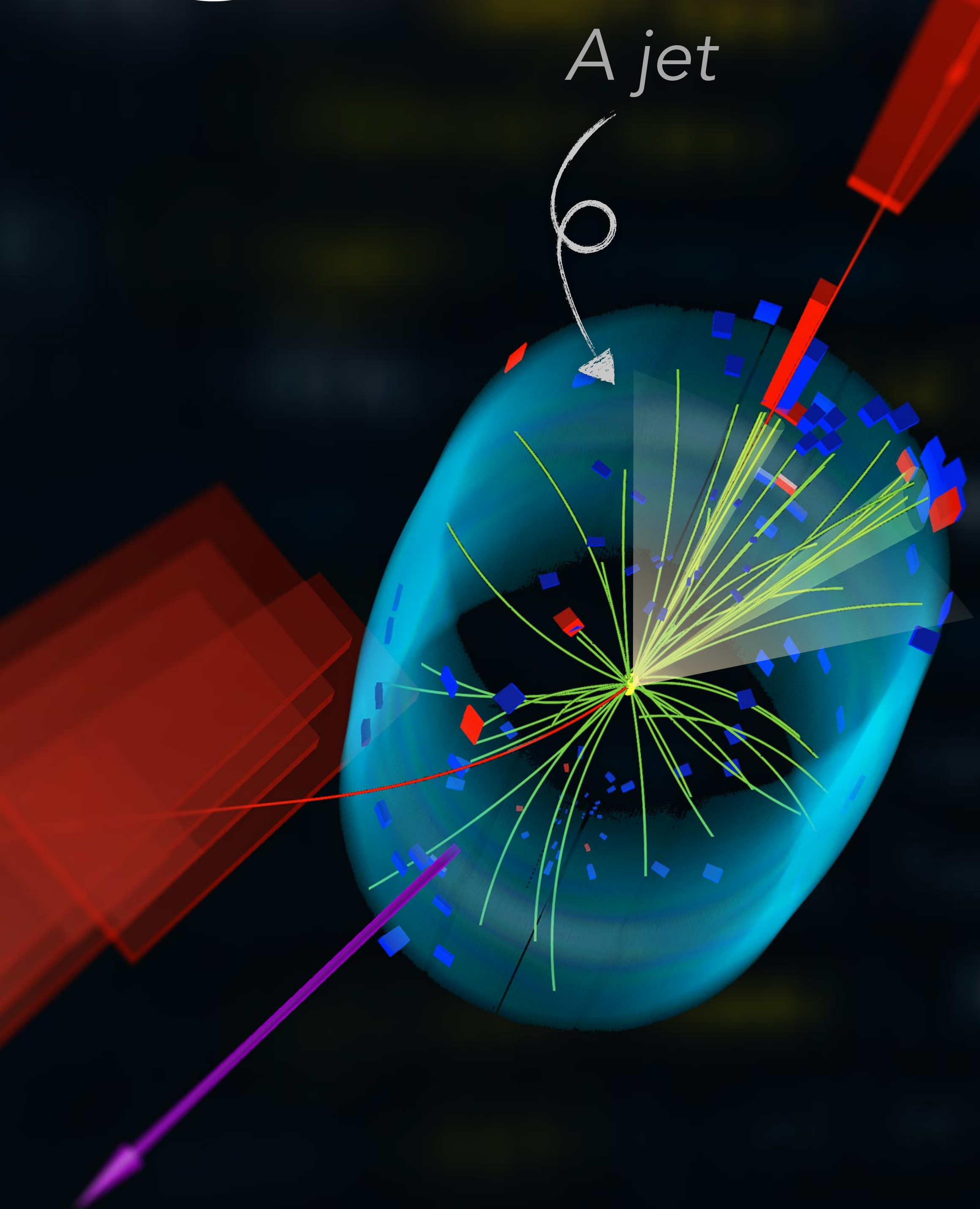


Machine learning for jets

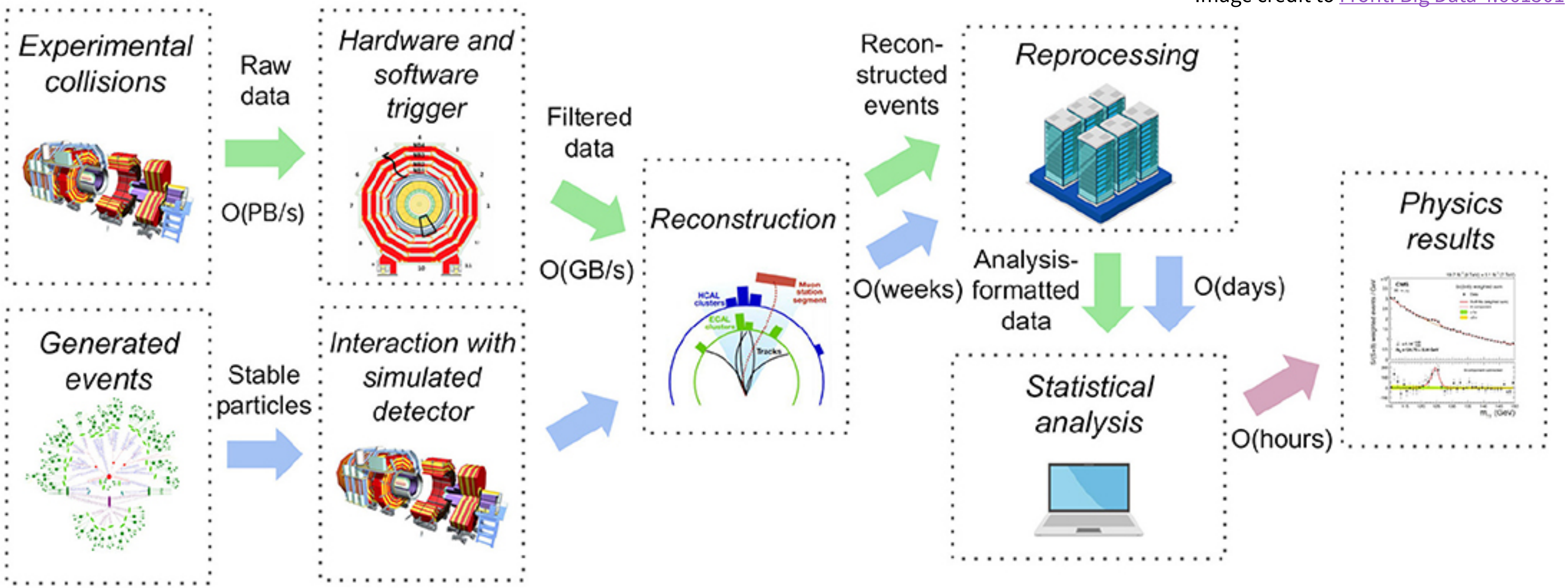
— Highlights on established DNN techniques, facilities, and key applications

Congqiao Li (李聪乔) *(Peking University)*

FTCF2024, Guangzhou
21 November, 2024

Typical collider experiment workflow

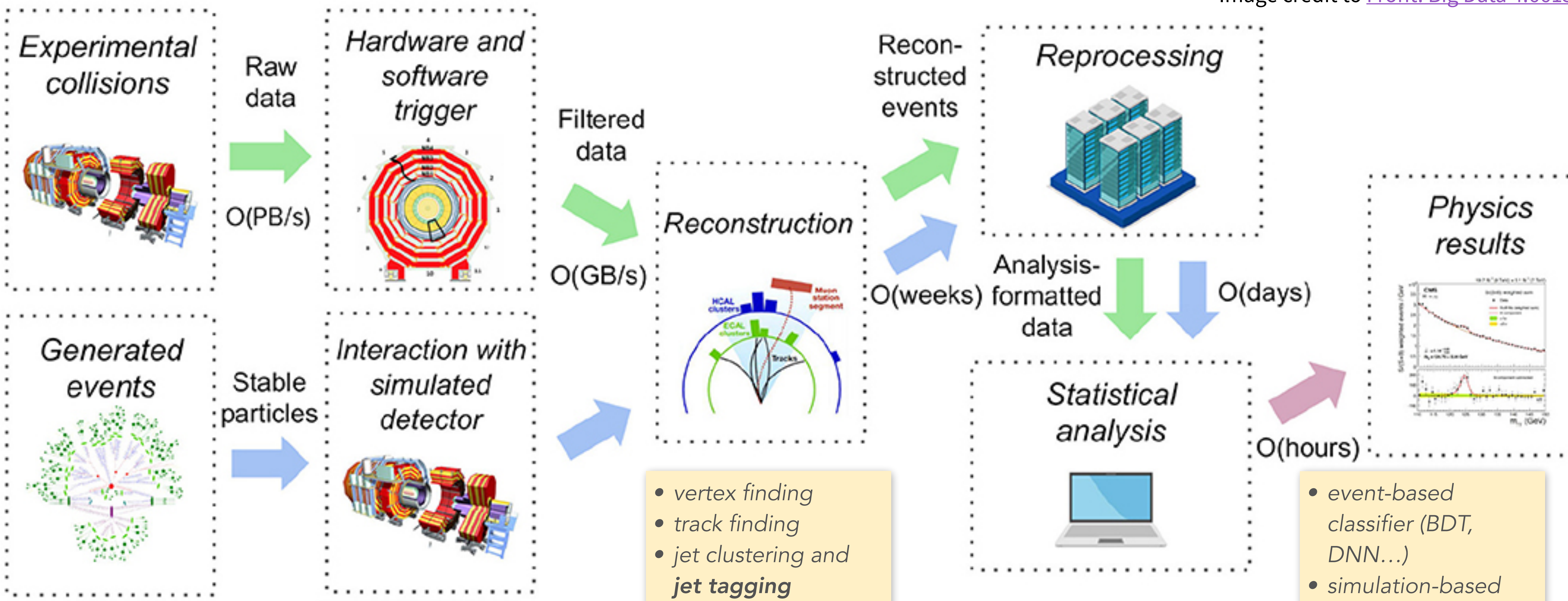
Image credit to [Front. Big Data 4:661501]()



Typical collider experiment workflow

ML is everywhere…

Image credit to [Front. Big Data 4:661501]



- *proton distribution function: NNPDF*
- 🚧 *ML for hadronisation*

- *fast detector simulation*

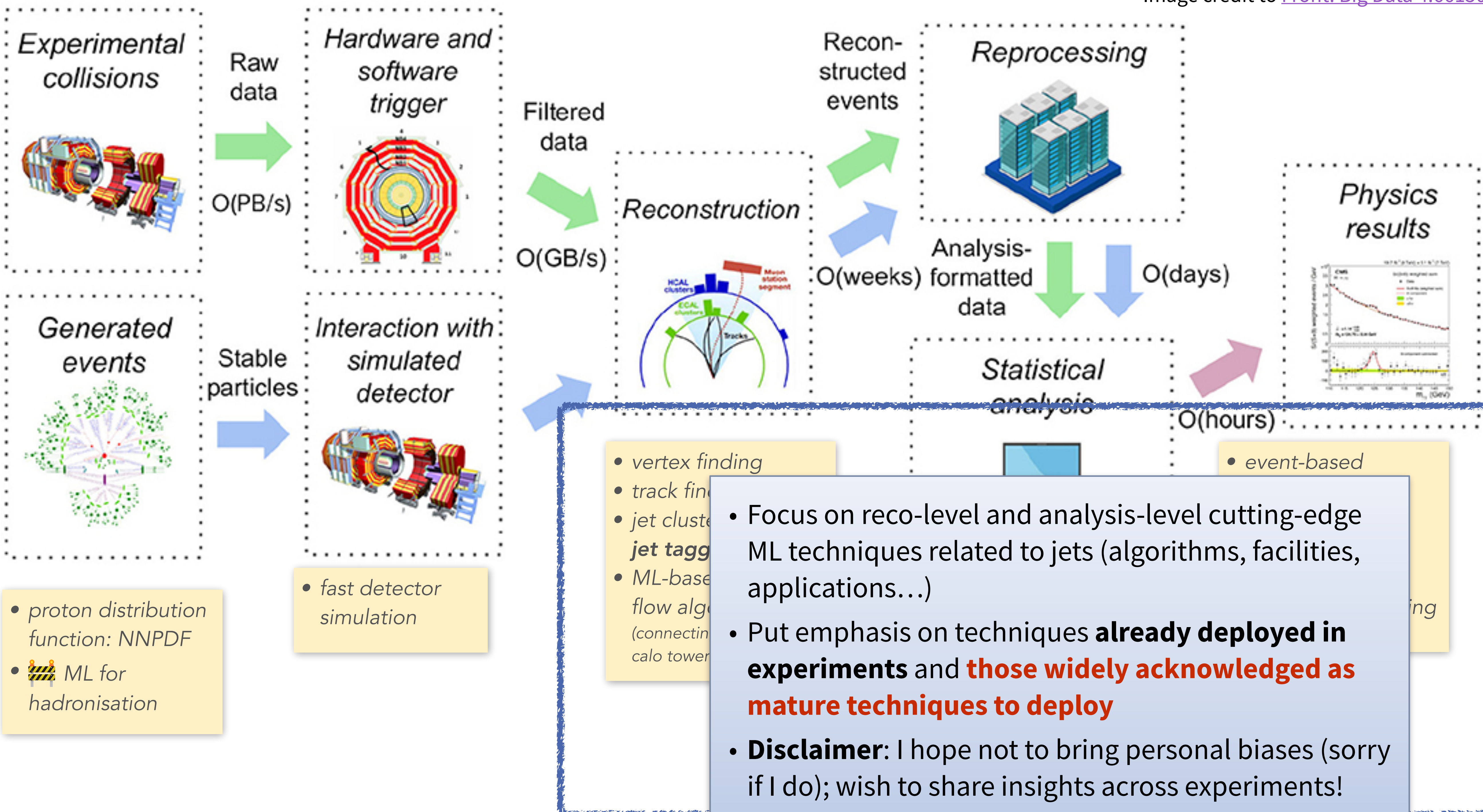
- *vertex finding*
- *track finding*
- *jet clustering and* ***jet tagging***
- *ML-based particle-flow algorithms* *(connecting tracks and calo towers)*

- *event-based classifier (BDT, DNN…)*
- *simulation-based inference*
- *ML-based unfolding*

Typical collider experiment workflow

ML is everywhere…

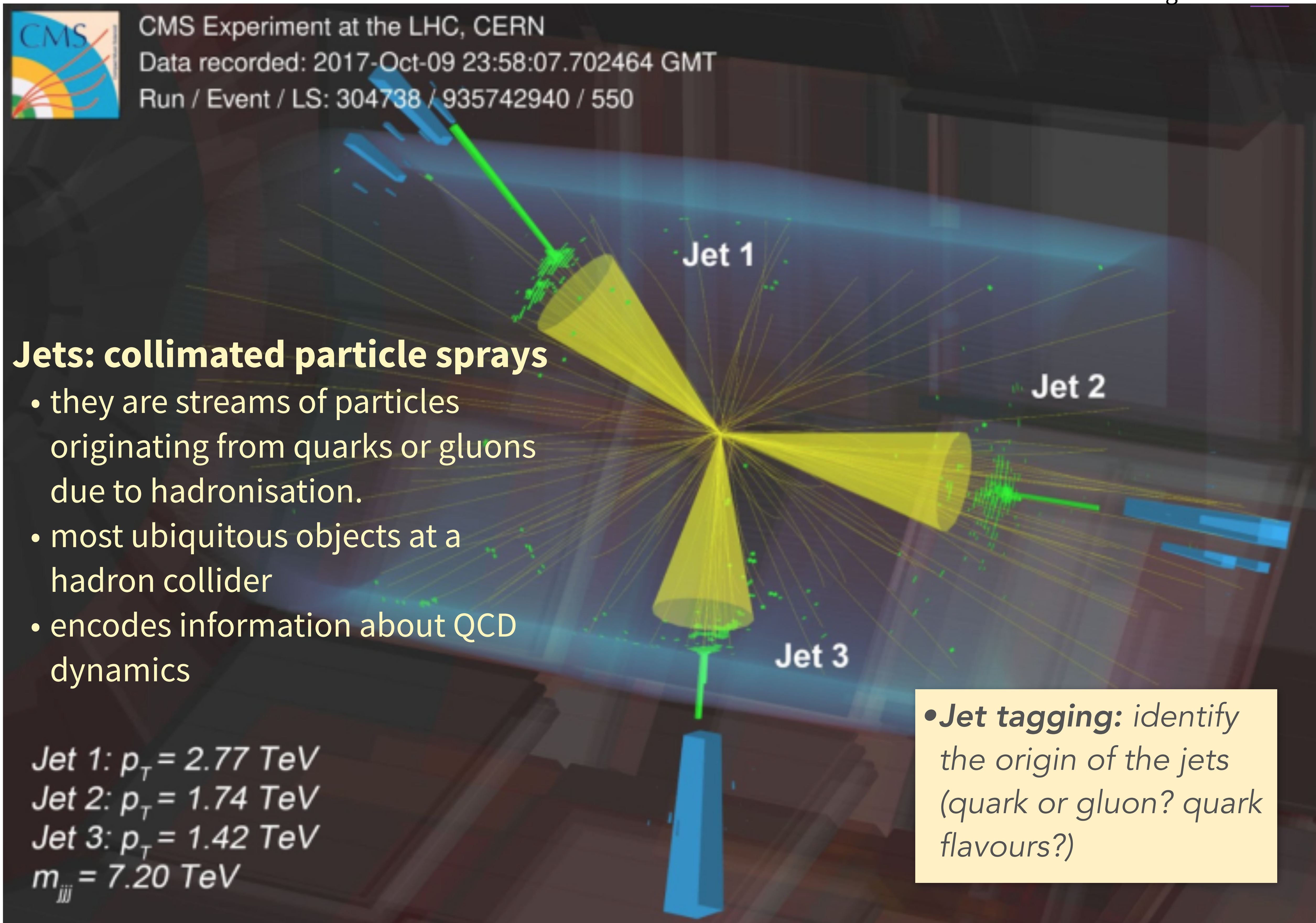
Image credit to Front. Big Data 4:661501



- Focus on reco-level and analysis-level cutting-edge ML techniques related to jets (algorithms, facilities, applications…)
- Put emphasis on techniques **already deployed in experiments** and **those widely acknowledged as mature techniques to deploy**
- **Disclaimer**: I hope not to bring personal biases (sorry if I do); wish to share insights across experiments!

Jets at hadron colliders

Image from link

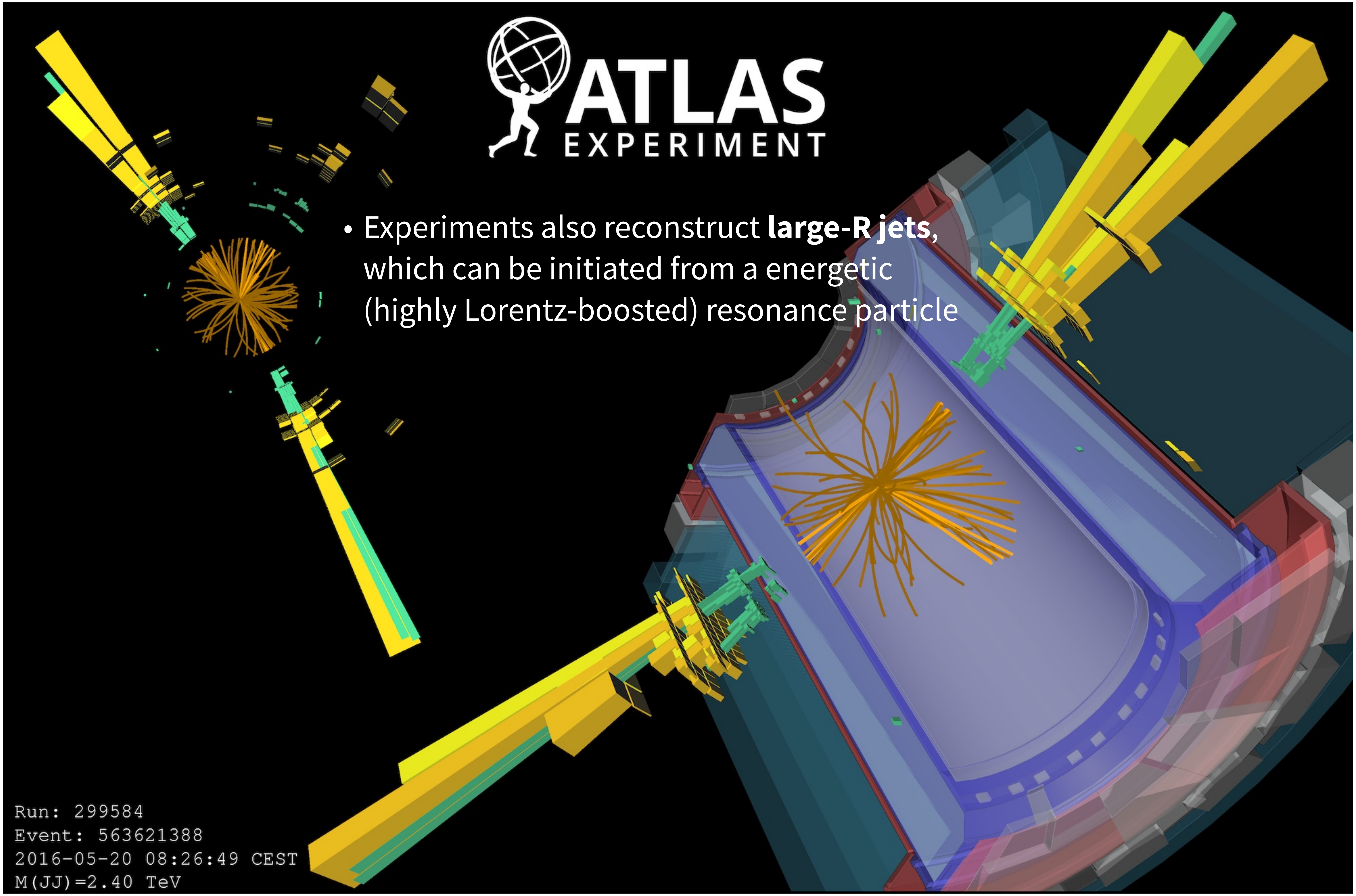


Jets: collimated particle sprays

- they are streams of particles originating from quarks or gluons due to hadronisation.
- most ubiquitous objects at a hadron collider
- encodes information about QCD dynamics

- ***Jet tagging:*** *identify the origin of the jets (quark or gluon? quark flavours?)*

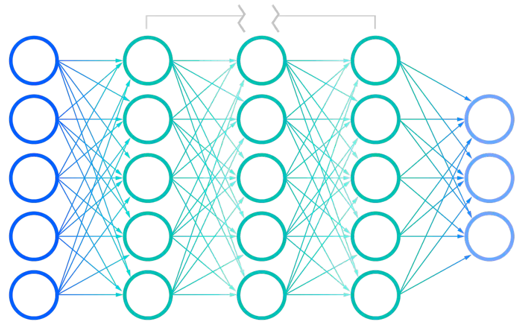
Jets at hadron colliders



- Experiments also reconstruct **large-R jets**, which can be initiated from a energetic (highly Lorentz-boosted) resonance particle

Evolution of jet NNs

feed-forward NN *(high-level inputs)* ····▶···· *1D/2D CNN, RNN* *(low-level inputs)* ····▶··· *graph NN, Transformers* *(low-level inputs)* ····▶··· *??*

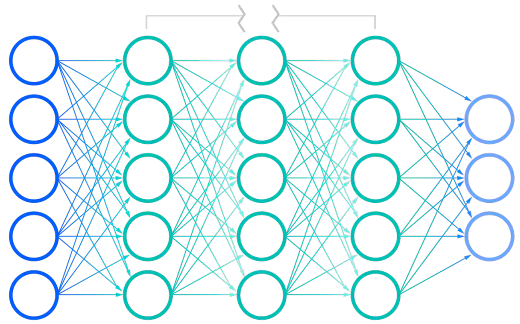


Shallow networks

✦ Using high-level features directly as input to a shallow network

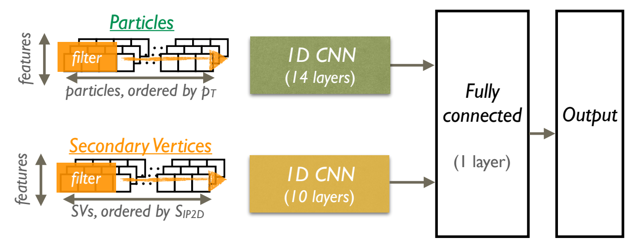
Evolution of jet NNs

feed-forward NN (high-level inputs) • • • ▶ • • • *1D/2D CNN, RNN (low-level inputs)* • • • ▶ • • • *graph NN, Transformers (low-level inputs)* • • • ▶ • • • *??*



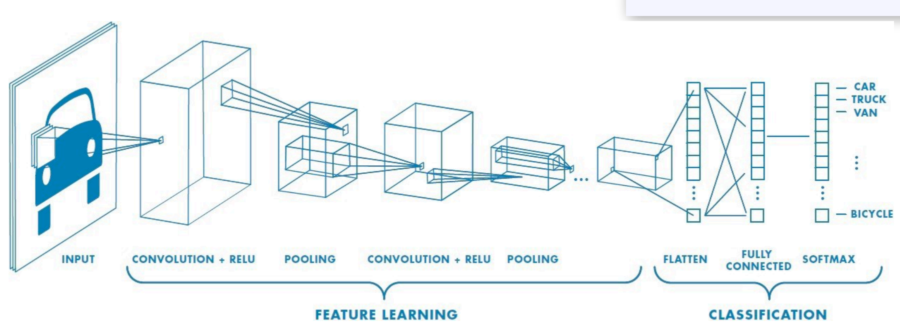
Shallow networks

- ✦ Using high-level features directly as input to a shallow network

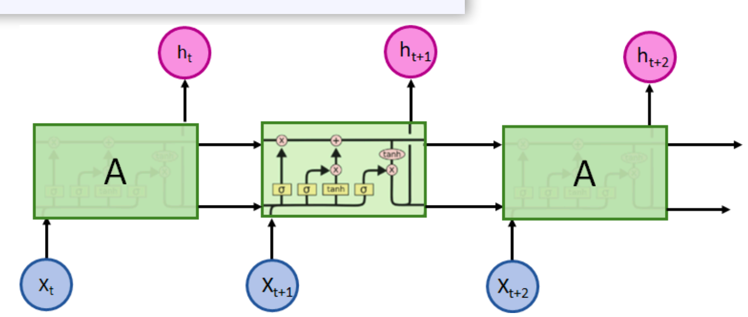


Deep NN with low-level inputs

- ✦ Using particle-level features
- ✦ Input data structure determines the type of networks
  - jet as a *image (fixed-grid data structure)*
  - jet as a *sequence* → 1D CNN or RNN



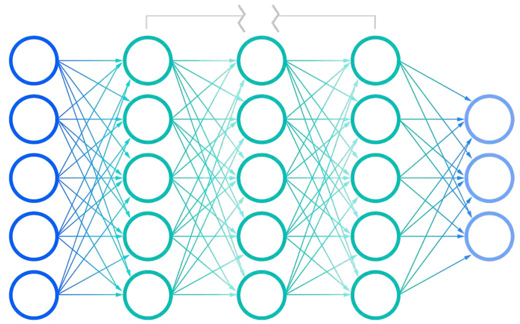
Typical CNN

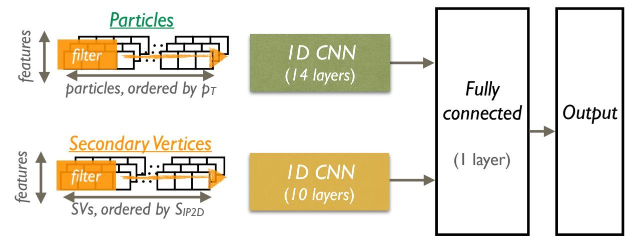


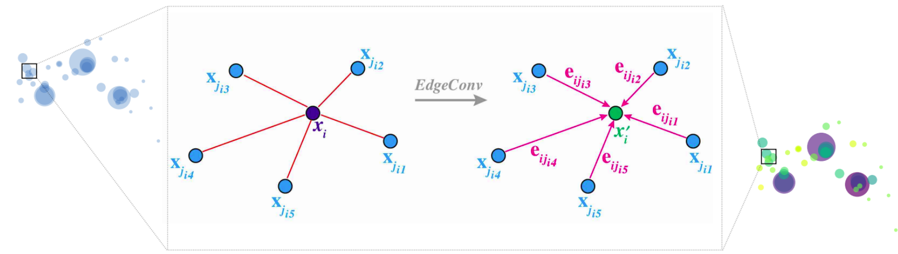
Typical RNN

Evolution of jet NNs

feed-forward NN *(high-level inputs)* → *1D/2D CNN, RNN* *(low-level inputs)* → *graph NN, Transformers* *(low-level inputs)* → *??*







Shallow networks

Deep NN with low-level inputs

Graph structure

✦ Using high-level features directly as input to a shallow network

✦ Using particle-level features

✦ Input data structure determines the type of networks

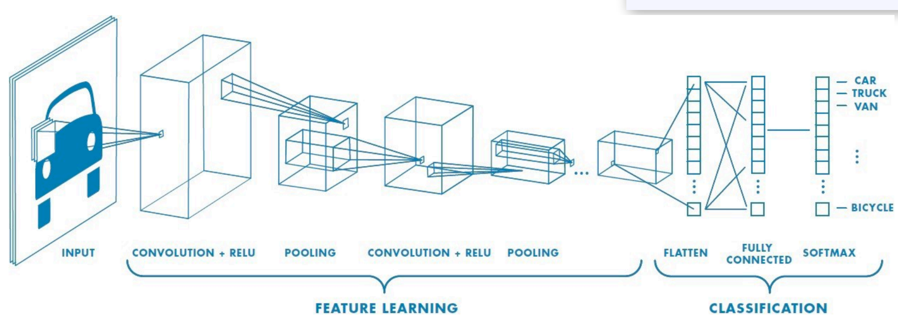
- jet as a *image (fixed-grid data structure)*
- jet as a *sequence* → 1D CNN or RNN

✦ Graph neural networks

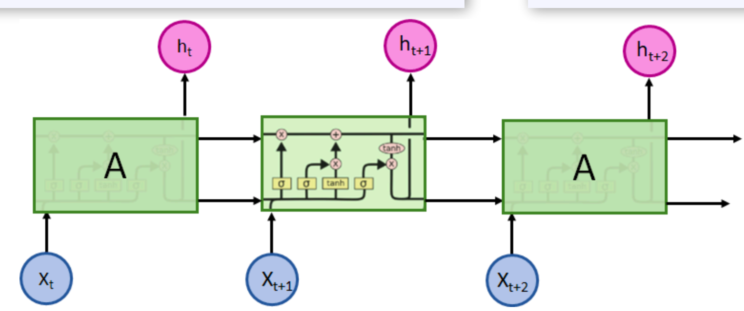
- treat a jet as a permutational-invariant set of particles (or, point cloud)
- build "edges" between particles

✦ Transformer networks

- modern architectural designs; like a full-connected graph



Typical CNN

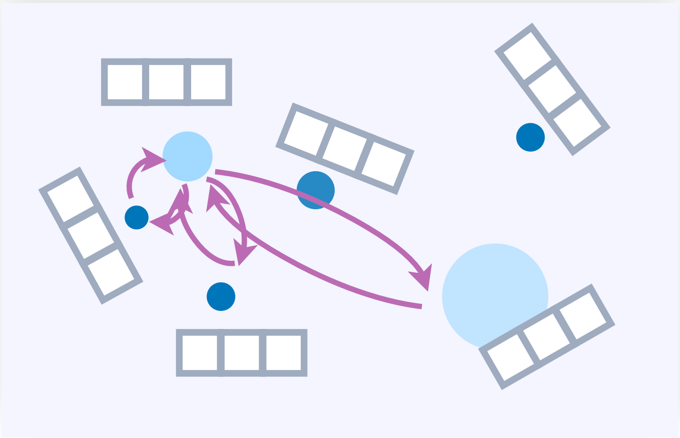


Typical RNN

Typical graph

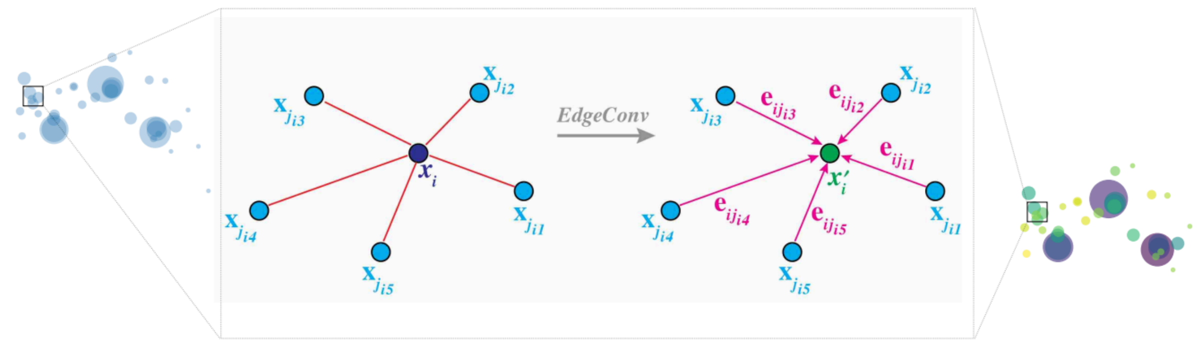
GNNs and Transformers

- **Modern architectures do right**: (DNNs that better suit the particle-format data?)
  - **inductive bias**: particle-format data has their intrinsic symmetries
    - **permutational-invariant symmetry**: GNN is better than CNN/RNN; native Transformer (w/o positional encoding)
    - Lorentz symmetry: adding "pairwise particle masses" to input features
  - **let particles interact**:
    - "message passing" in GNNs and attention mechanism in Transformers
  - Scale better with data and model size: Transformers!



ParticleNet and its applications

- ParticleNet, based on dynamic graph GNNs
  - treat jet as a (permutational-invariant) set of point clouds
  - define “a local patch” for each particle by k-nearest neighbours and apply convolution (EdgeConv)

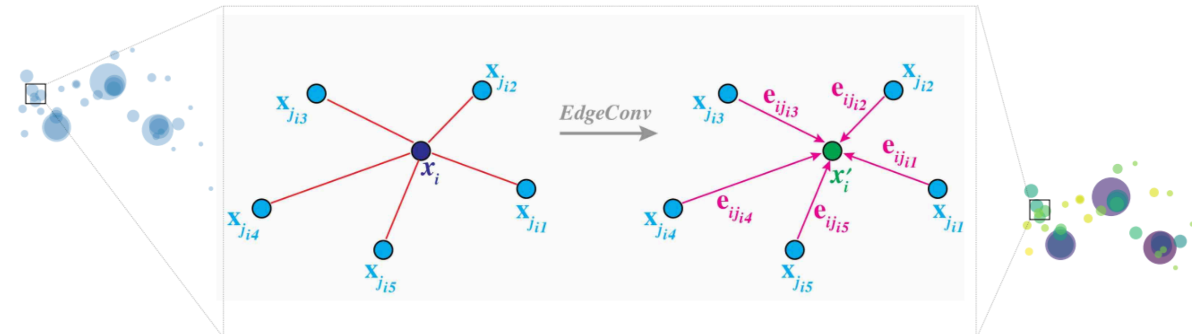


Applications in CMS

ParticleNet and its applications

→ ParticleNet, based on dynamic graph GNNs

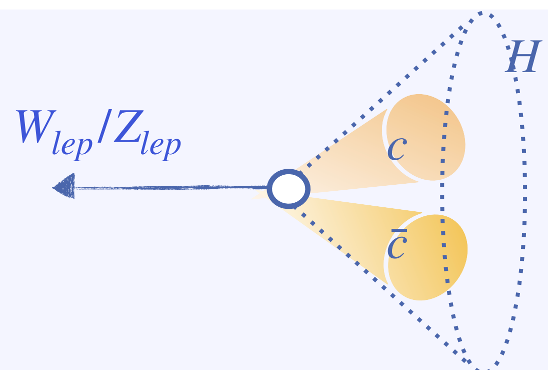
- treat jet as a (permutational-invariant) set of point clouds
- define "a local patch" for each particle by k-nearest neighbours and apply convolution (EdgeConv)



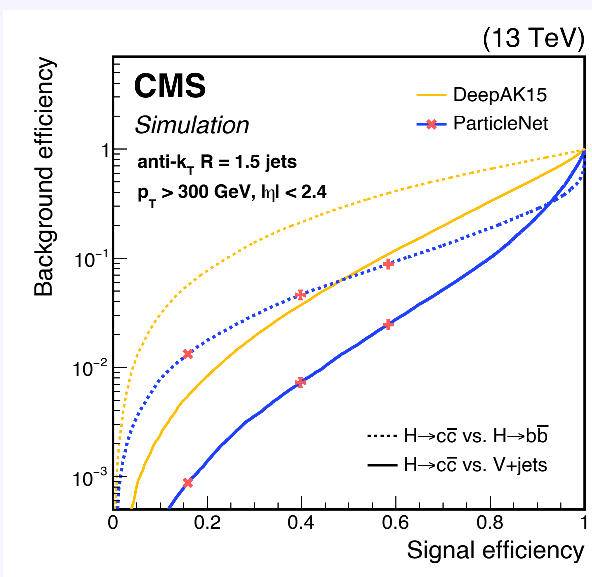
Applications in CMS

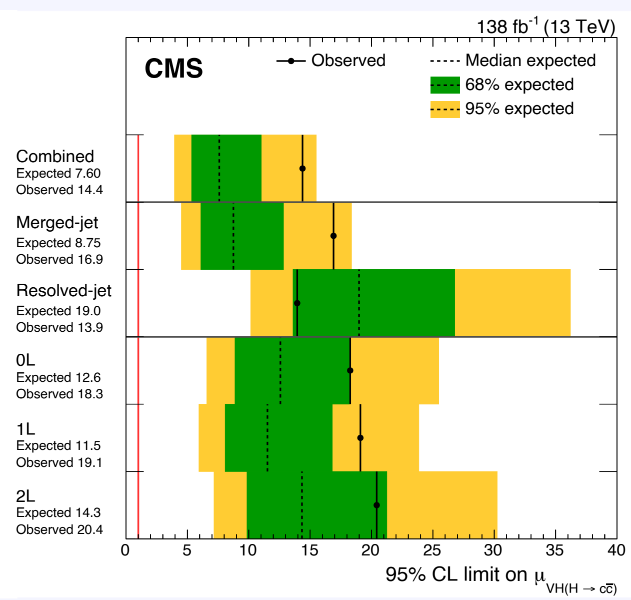
VH→cc search (merged region)

PRL 131 (2023) 061801



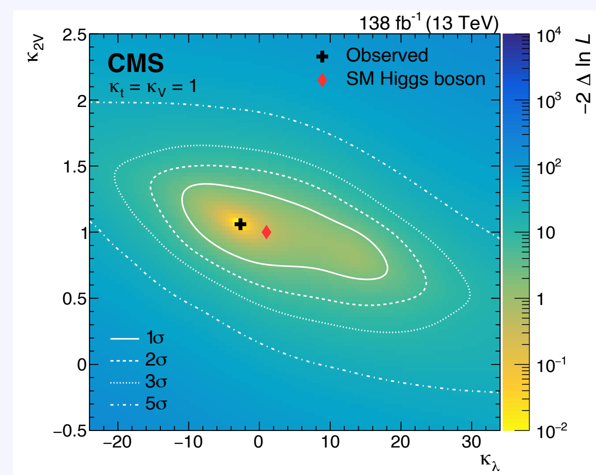
- With merged+resolved region combined, achieve most stringent direct limit on κ_c: $1.1 < |\kappa_c| < 5.5$

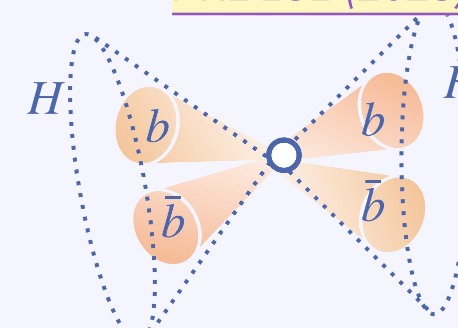




Boosted HH→4b search

PRL 131 (2023) 041803

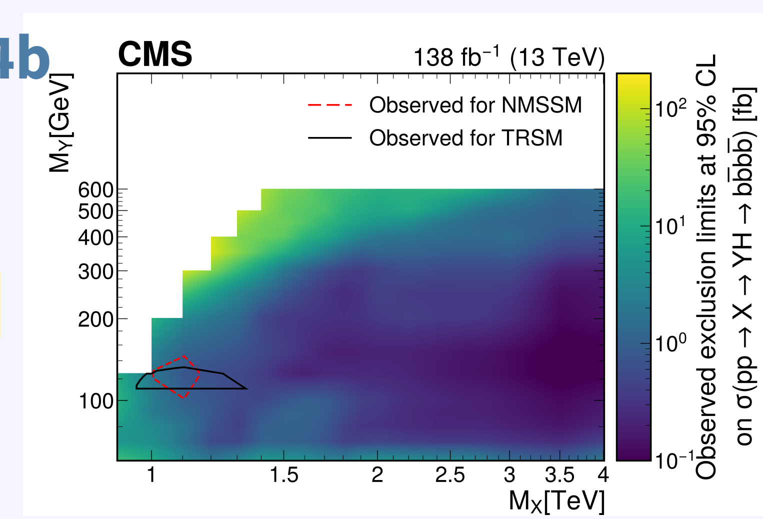




- First time excluding $\kappa_{2V} = 0$

Boosted X→YH→4b search

PLB 842 (2023) 137392



ParticleNet and its applications

Other selected applications

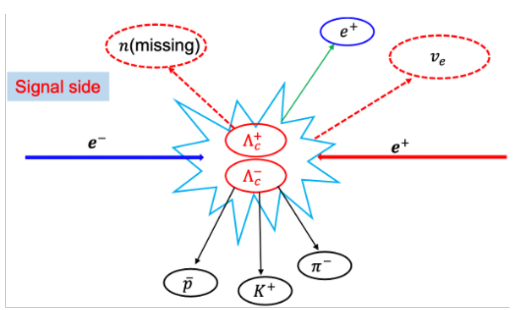
credit to Huilin's slides

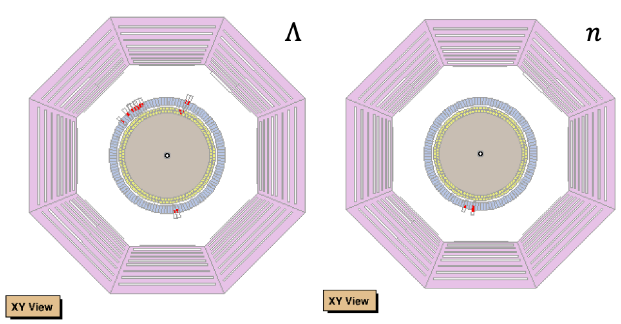
BESIII

$\Lambda_c^+ \to n e^+ \nu$ search

Yunxuan Song, Yangu Li et al.

arXiv: 2410.13515, accepted by Nature Commun.





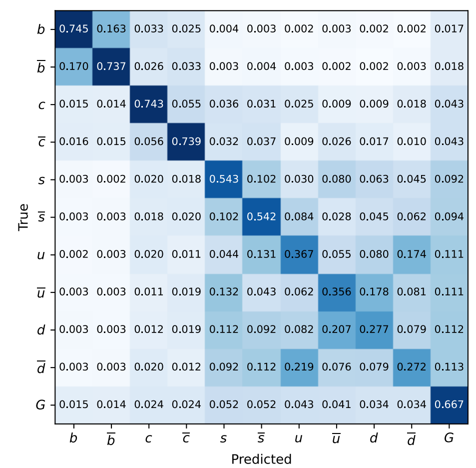
Talk by Yangu Li today

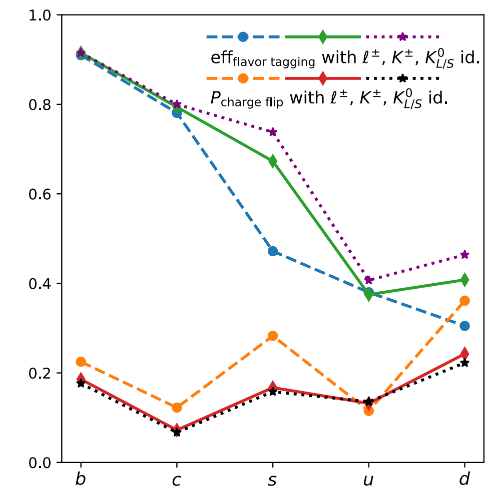
CEPC

Jet origin identification

PRL 132 (2024), 221802

Eur.Phys.J.C 84 (2024), 152





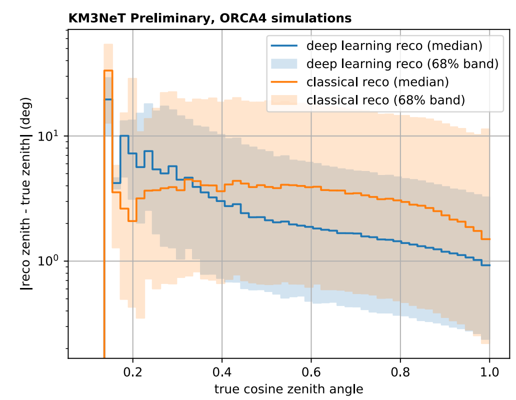
Talk by Manqi

KM3NeT

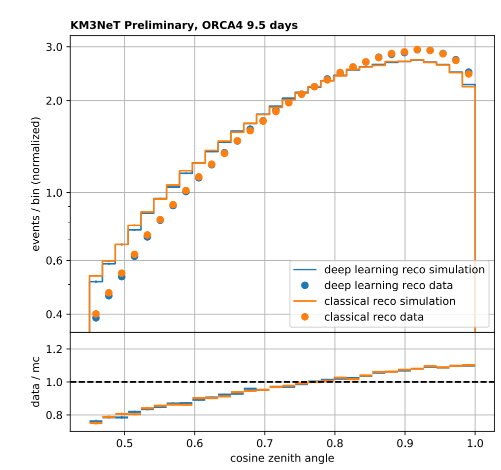
Muon bundle reconstruction

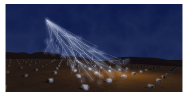
JINST 16 (2021) 10, C10011,

PoS ICRC2021 (2021) 1048



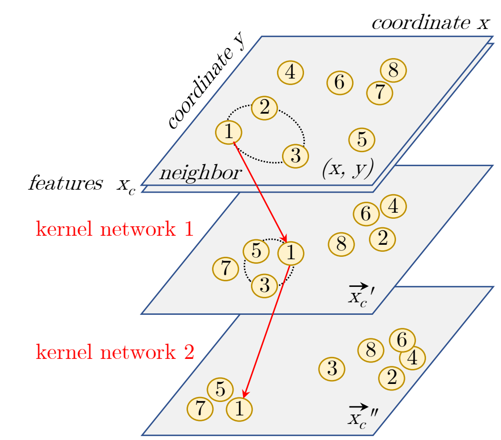
(b) events with two or more muons

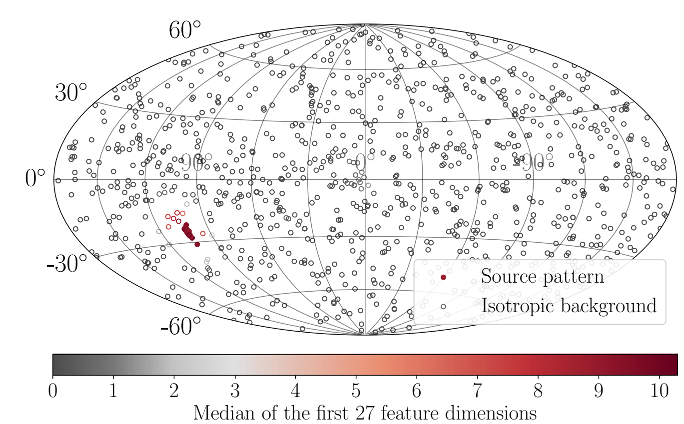




Cosmic ray pattern identification

Astropart.Phys. 126 (2021) 102527

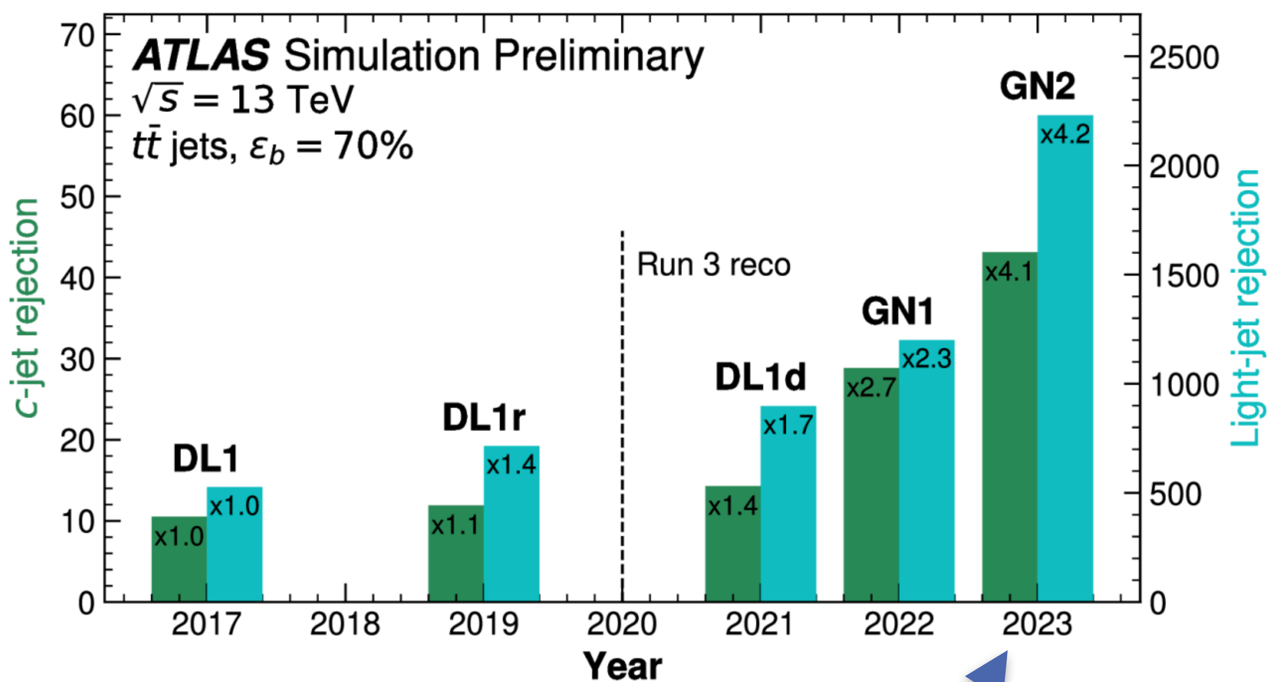




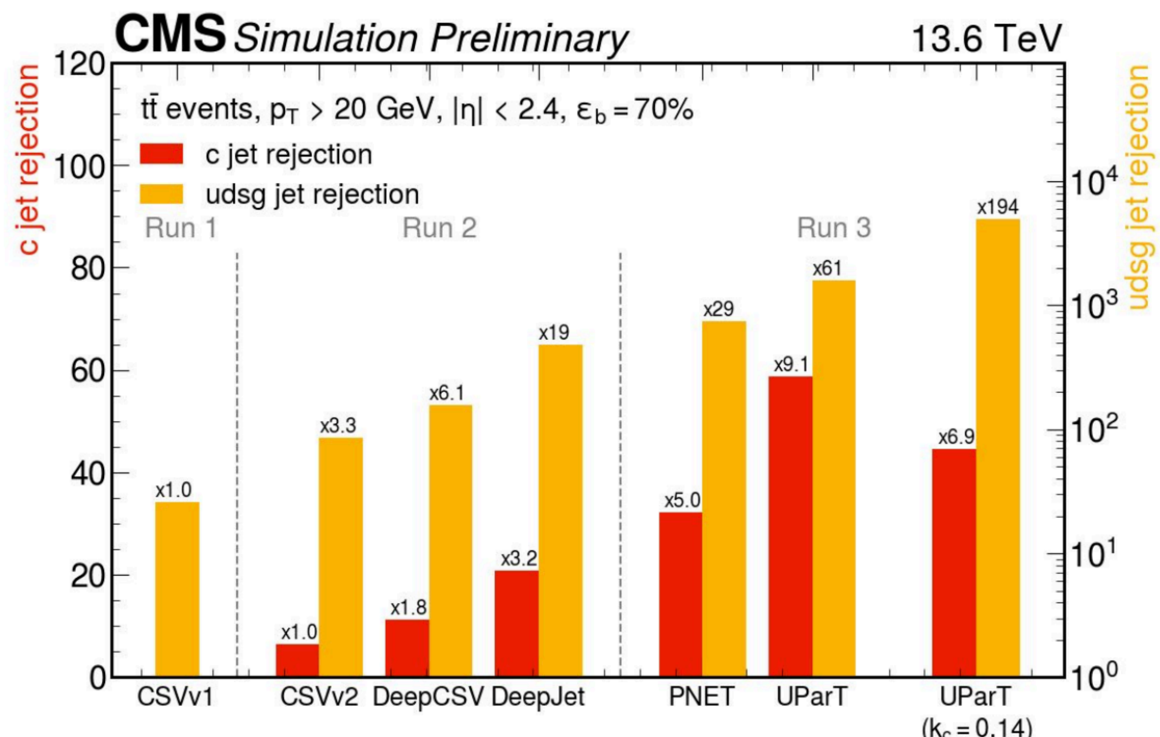
Transformer jet taggers in ATLAS/CMS

➔ ATLAS/CMS "flagship" jet taggers have all switched to the Transformer architectures

- much improved b-tagging performance (to reject c-jets and light jets)
- huge progress has been made from **2016 (early Run-2)** to **2024 (mid-Run3)** !



ATL-FTAG-2023-01



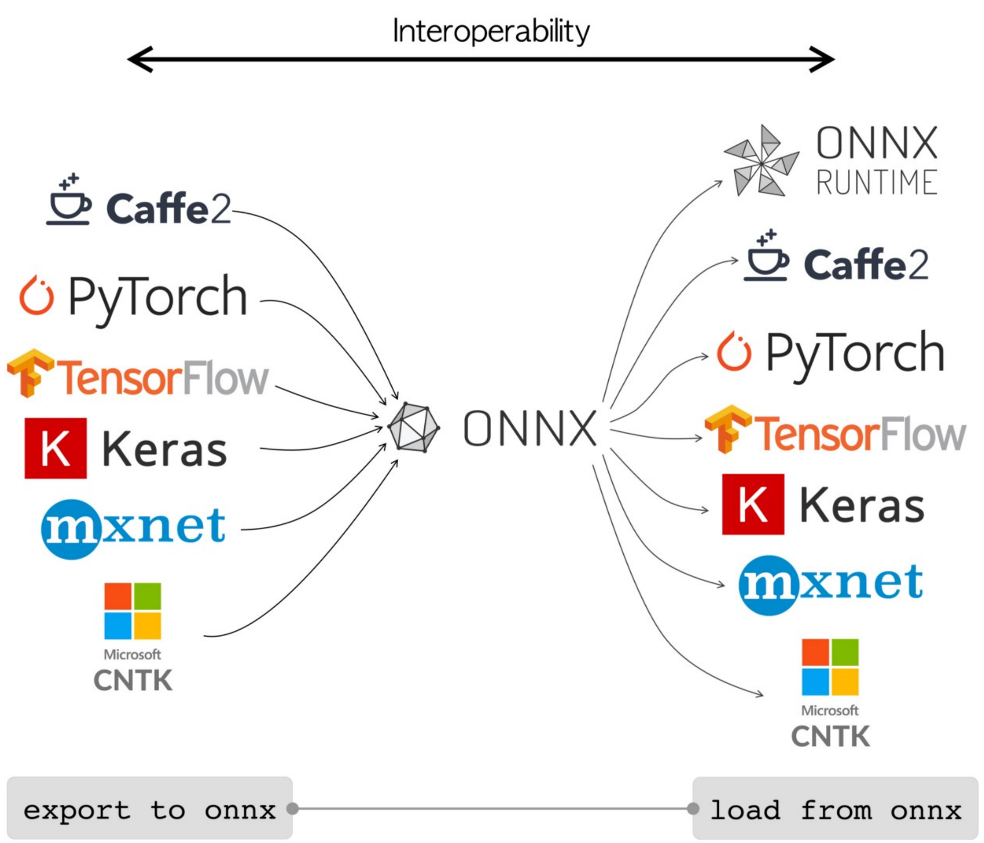
CMS-DP-2024-066

Latest ATLAS tagger for small-R jets: Transformer-based GN2

Latest CMS tagger for small-R jets: Unified ParT (UParT)

Inference facility: ONNX runtime

- ➔ Both ATLAS and CMS use ONNX runtime inference engine
  - ❖ **ONNX**: "Open Neural Network eXchange" format, representing a BDT/DNN model
    - support model conversion from XGBoost (BDT models), TensorFlow, PyTorch…
  - ❖ **ONNX runtime**: accelerate ML inferencing across a variety of platforms
    - CPU/GPU environment for model inference
    - support C++/python interface and more!
- ➔ Helper functions in the ATLAS (athena) and CMS (cmssw) software
  - ❖ for inference of jet taggers (with low-level inputs) in central workflow
  - ❖ also support inferencing custom DNN models at analysis level



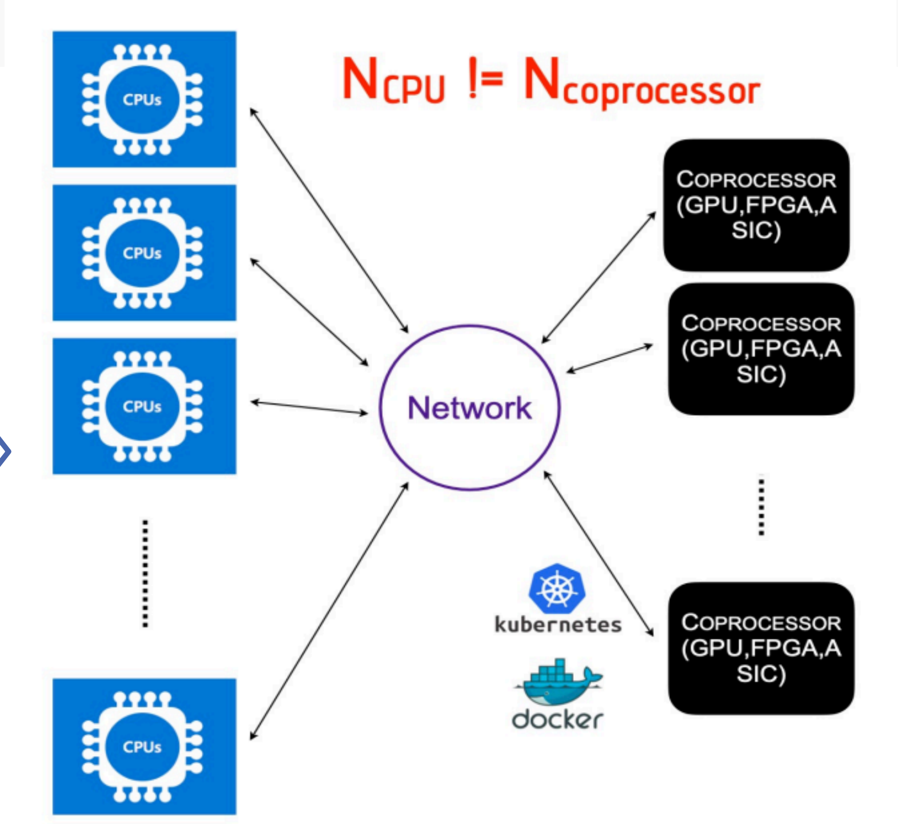
Inference facility: SONIC

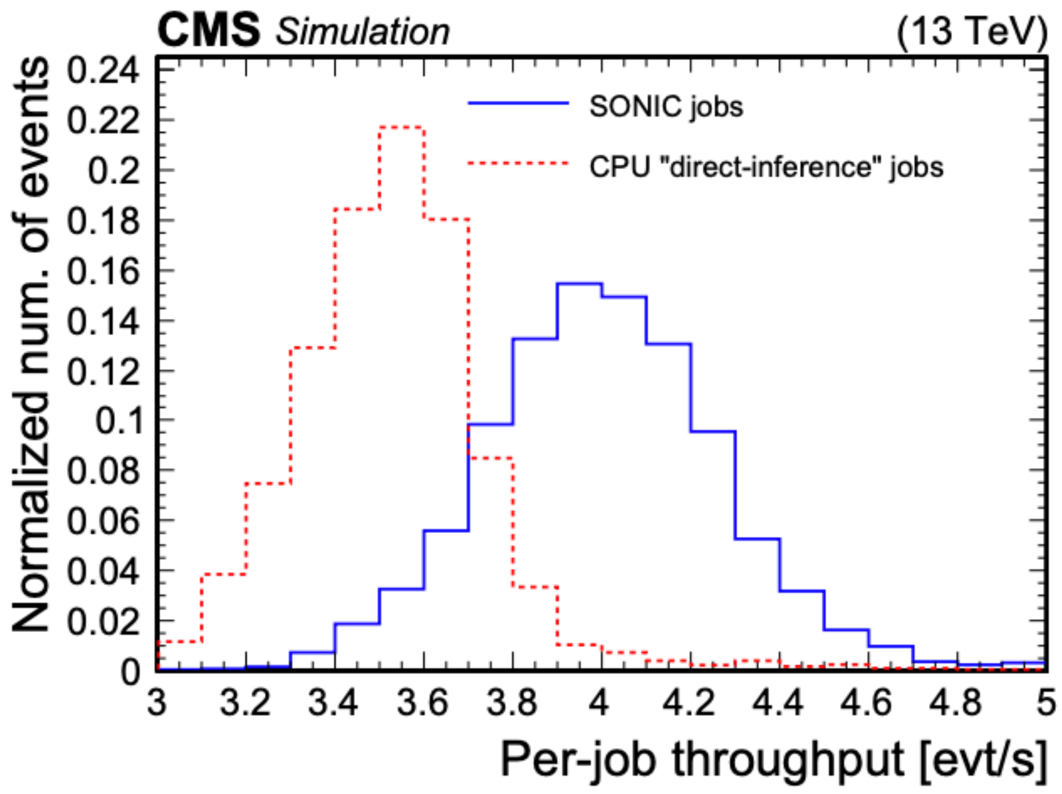
➔ Inference as-a-service (Iaas)

- ❖ instead of using CPU for the whole data processing workflow, certain tasks can be run more efficiently on other specialised processors (i.e. coprocessors)

➔ SONIC (Service for Optimized Network):

- ❖ is the implementation of IaaS in experimental software frameworks





Comput Softw Big Sci 8, 17 (2024)

- Test delivered to run CMS's ParticleNet, DeepTau, and DeepMET algorithms
- with SONIC approach
- Improved throughputs on the large-scale tests

Modern model-agnostic searches

Modern model-agnostic searches

→ Begin of journey in the modern (machine-learning-based) model-agnostic searching scheme at LHC

Anomaly Detection for Resonant New Physics with Machine Learning

Jack H. Collins (Maryland U. and Johns Hopkins U.), Kiel Howe (Fermilab), Benjamin Nachman (UC, Berkeley and LBL, Berkeley)

May 7, 2018

PRL, 121 (2018) 24, 241803

161 citations

Modern model-agnostic searches

➔ Begin of journey in the modern (machine-learning-based) model-agnostic searching scheme at LHC

Anomaly Detection for Resonant New Physics with Machine Learning

Jack H. Collins (Maryland U. and Johns Hopkins U.), Kiel Howe (Fermilab), Benjamin Nachman (UC, Berkeley and LBL, Berkeley)

May 7, 2018

PRL, 121 (2018) 24, 241803

161 citations

➔ A "general method" for resonant search with minimal requirements

- resonance localised in a mass window
- can be reconstructed by two hadronic large-*R* jets

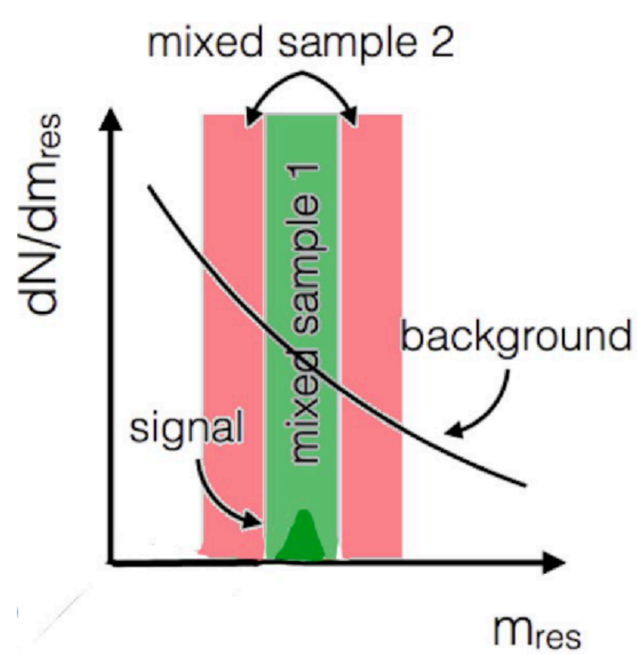
➔ General strategy:

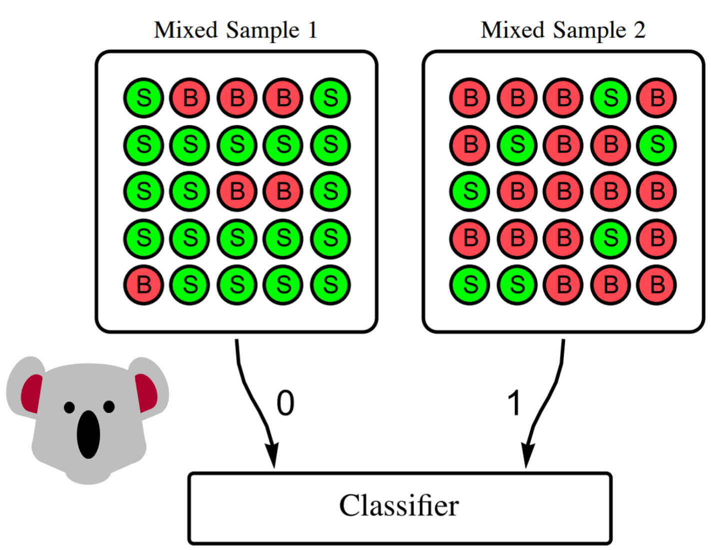
- scan on the mass spectrum → apply model-independent selection → purify the signal

➔ With no significant evidence of new physics found at LHC, a broader search strategy will be a meaningful

Weakly-supervised approach

JHEP 10 (2017) 174





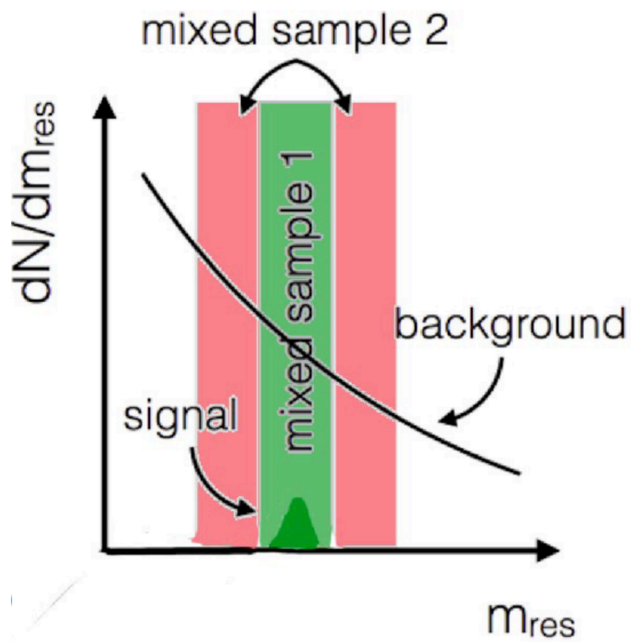
Equivalent effect for training **S** vs **B**

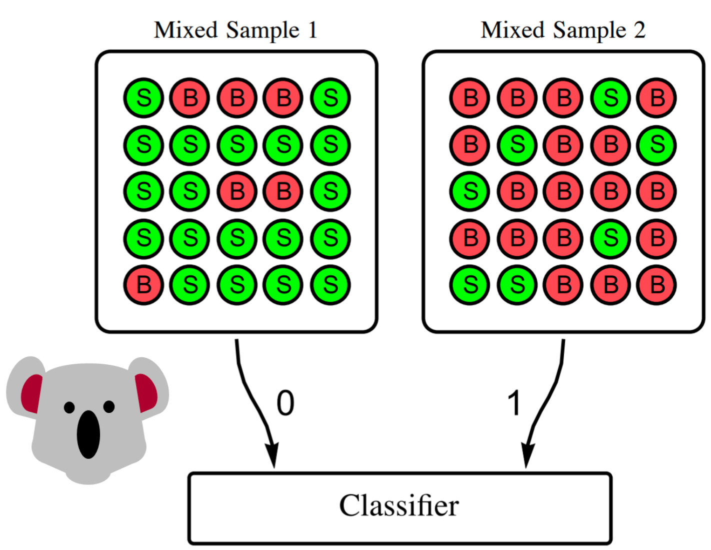
➔ Proposed "CWoLa (classification without labels) Hunting"

- allow to detect anomalies purely from data
- train a classifier for mass window vs mass sideband (mixed sample 1 vs 2)
- can prove that the effect is equivalent to training S vs B

Weakly-supervised approach

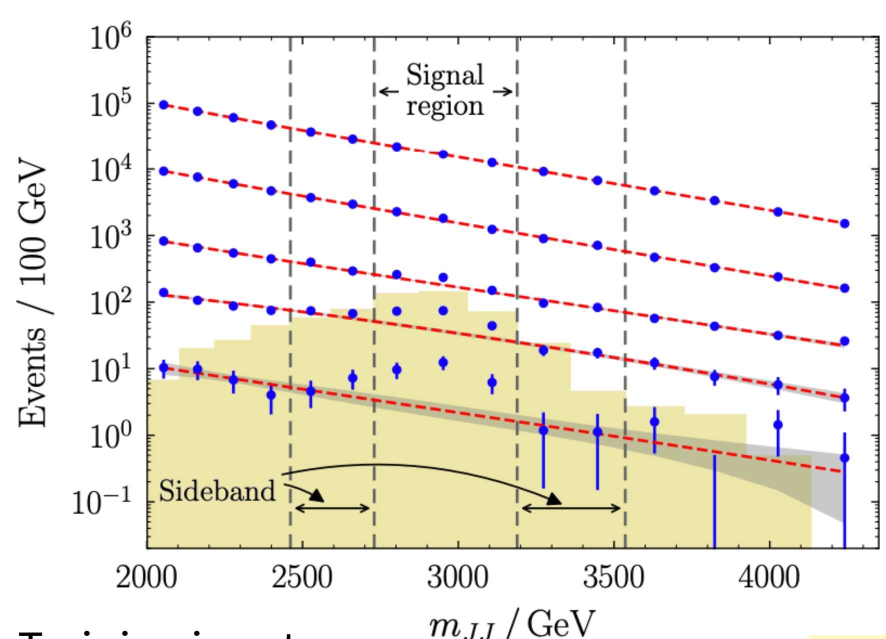
JHEP 10 (2017) 174





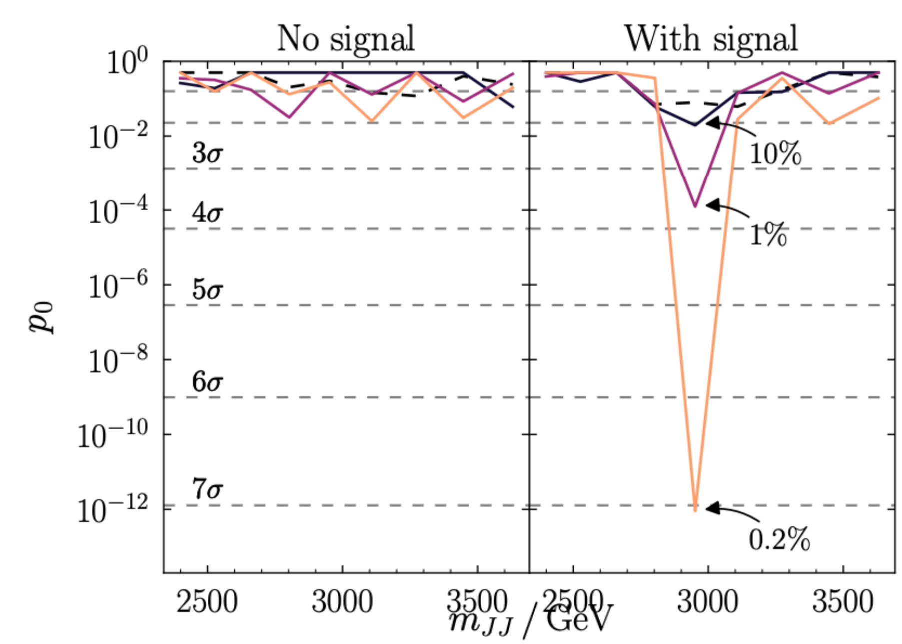
Equivalent effect for training **S** vs **B**

- Proposed “CWoLa (classification without labels) Hunting”
  - allow to detect anomalies purely from data
  - train a classifier for mass window vs mass sideband (mixed sample 1 vs 2)
  - can prove that the effect is equivalent to training S vs B



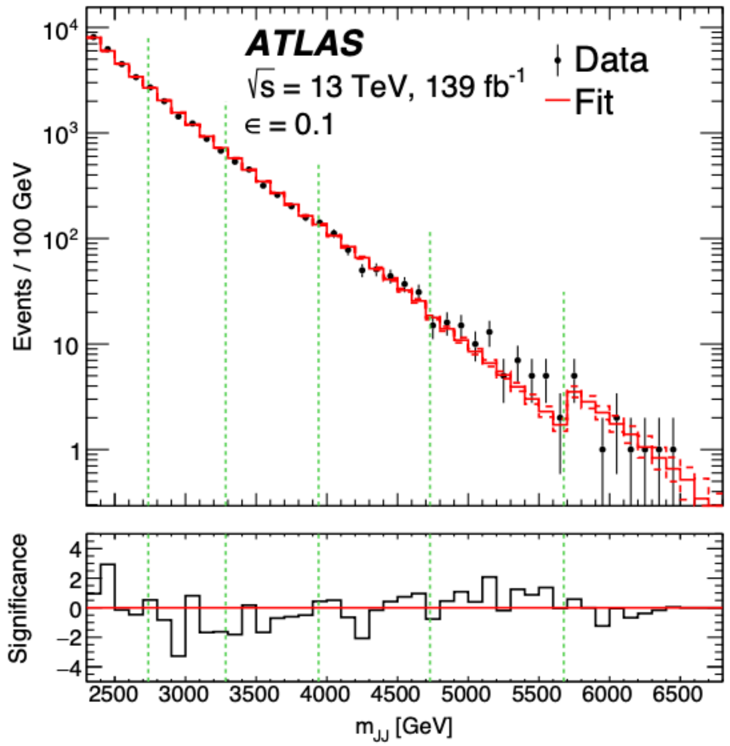
Training input

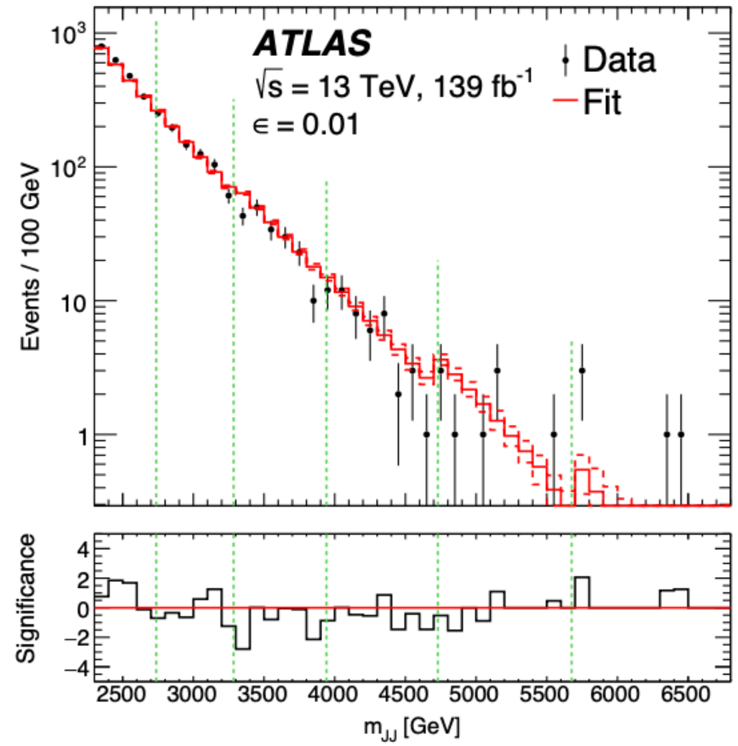
$m_J, \sqrt{\tau_1^{(2)}/\tau_1^{(1)}}, \tau_{21}, \tau_{32}, \tau_{43}, n_{\mathrm{trk}},$



PRL, 121 (2018) 24, 241803

Weakly-supervised approach

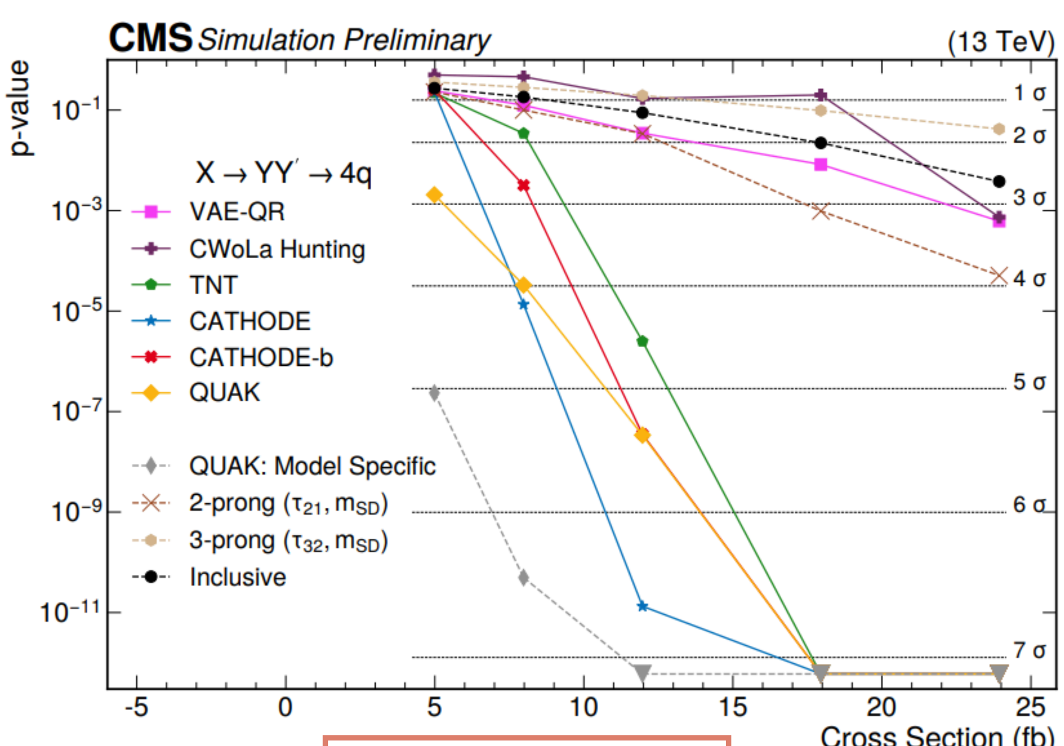




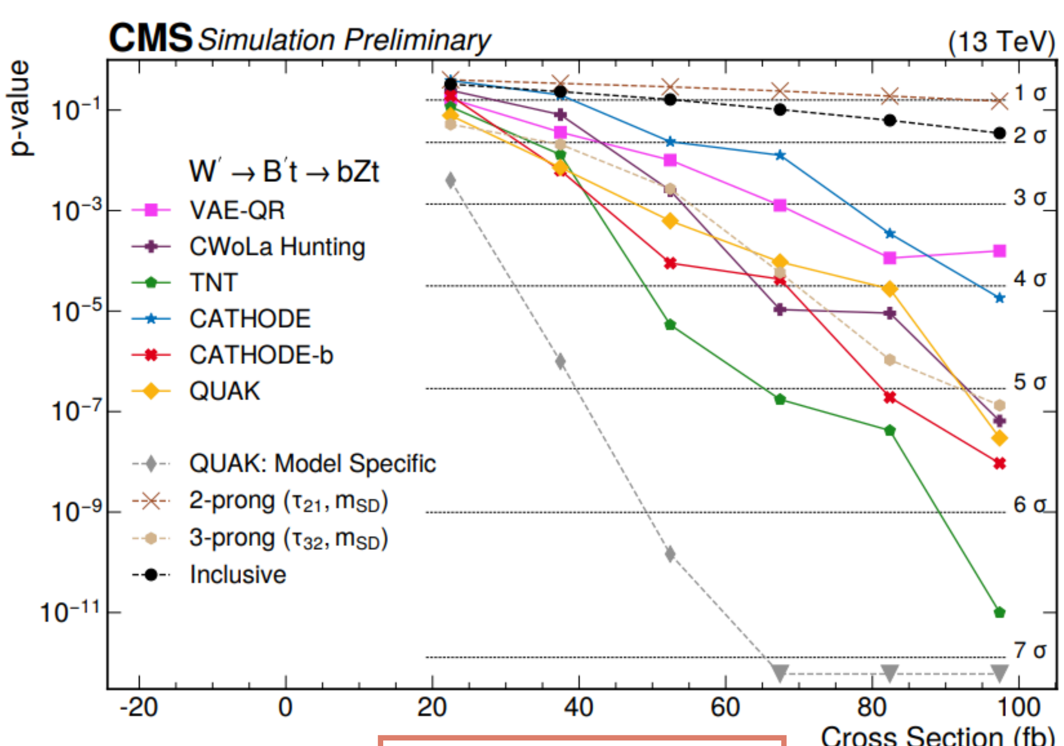
Application in ATLAS

PRL 125, 131801 (2020)

Select on the trained weakly supervised classifier in each signal regions, and search for the peak



2-prong + 2-prong

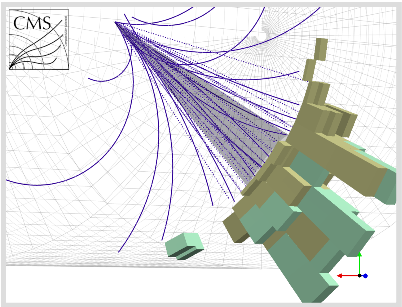


3-prong + 3-prong

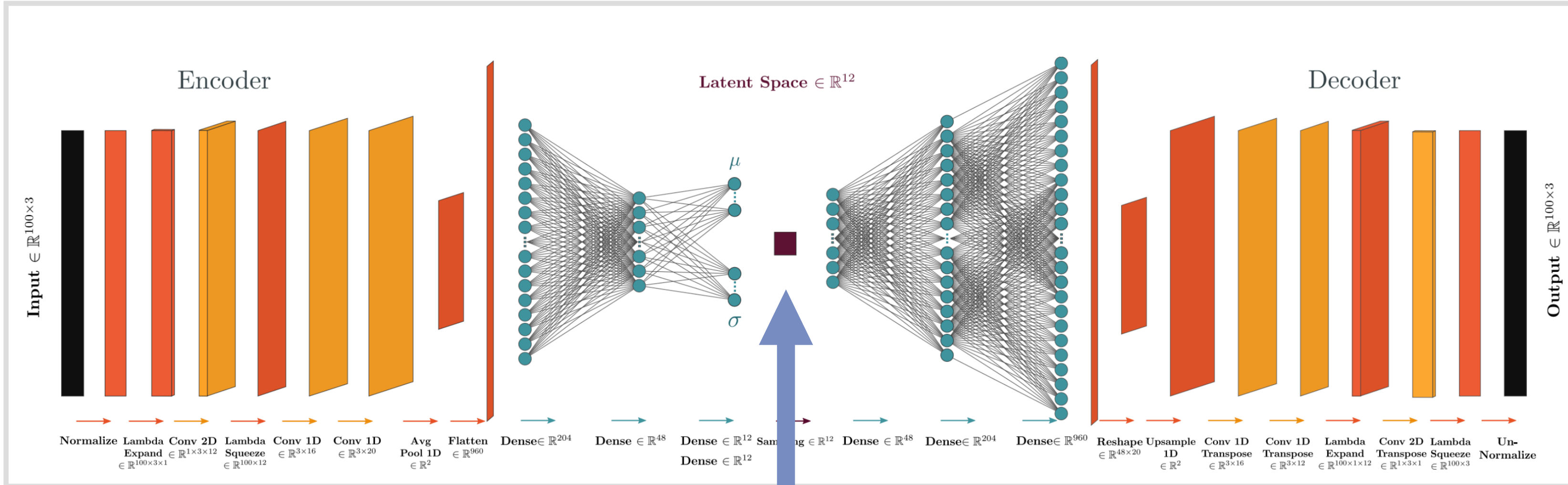
CMS-PAS-EXO-22-026

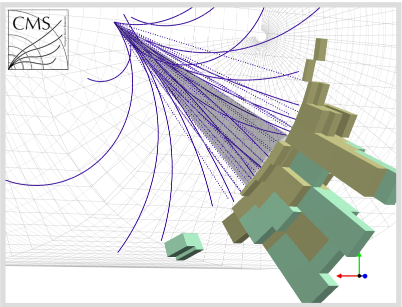
In a recent work, CMS systematically test all model-agnostic approaches to search for resonance

Autoencoder approach



input jet

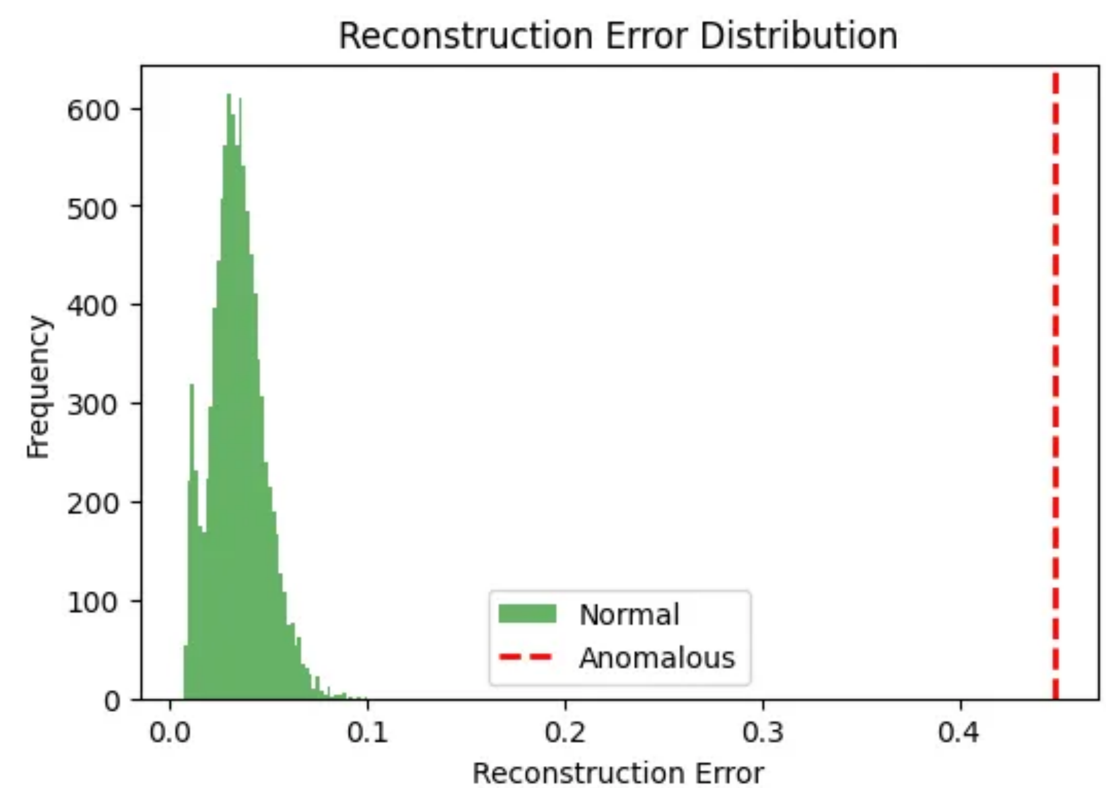




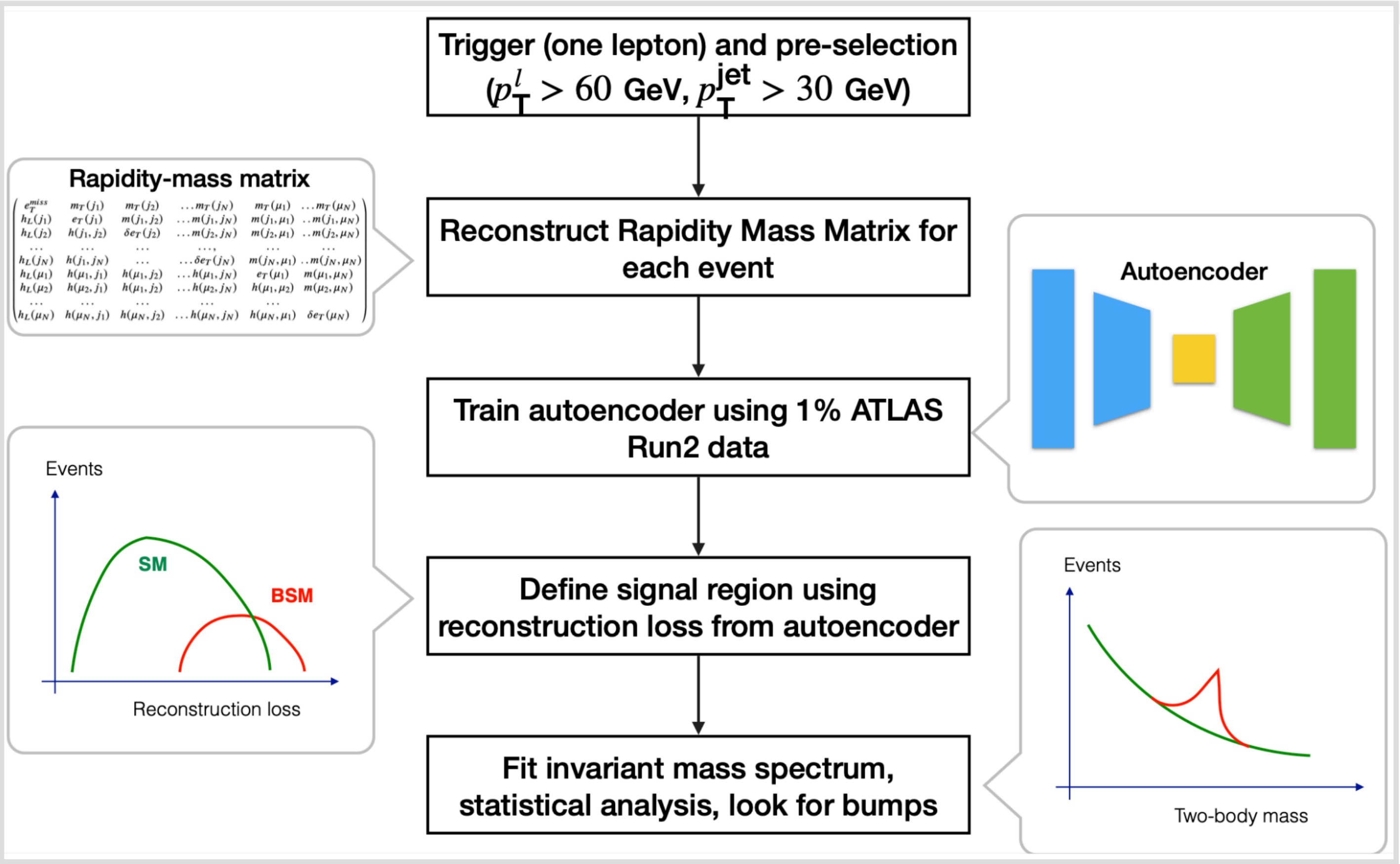
reconstructed jet

a compressed jet representation

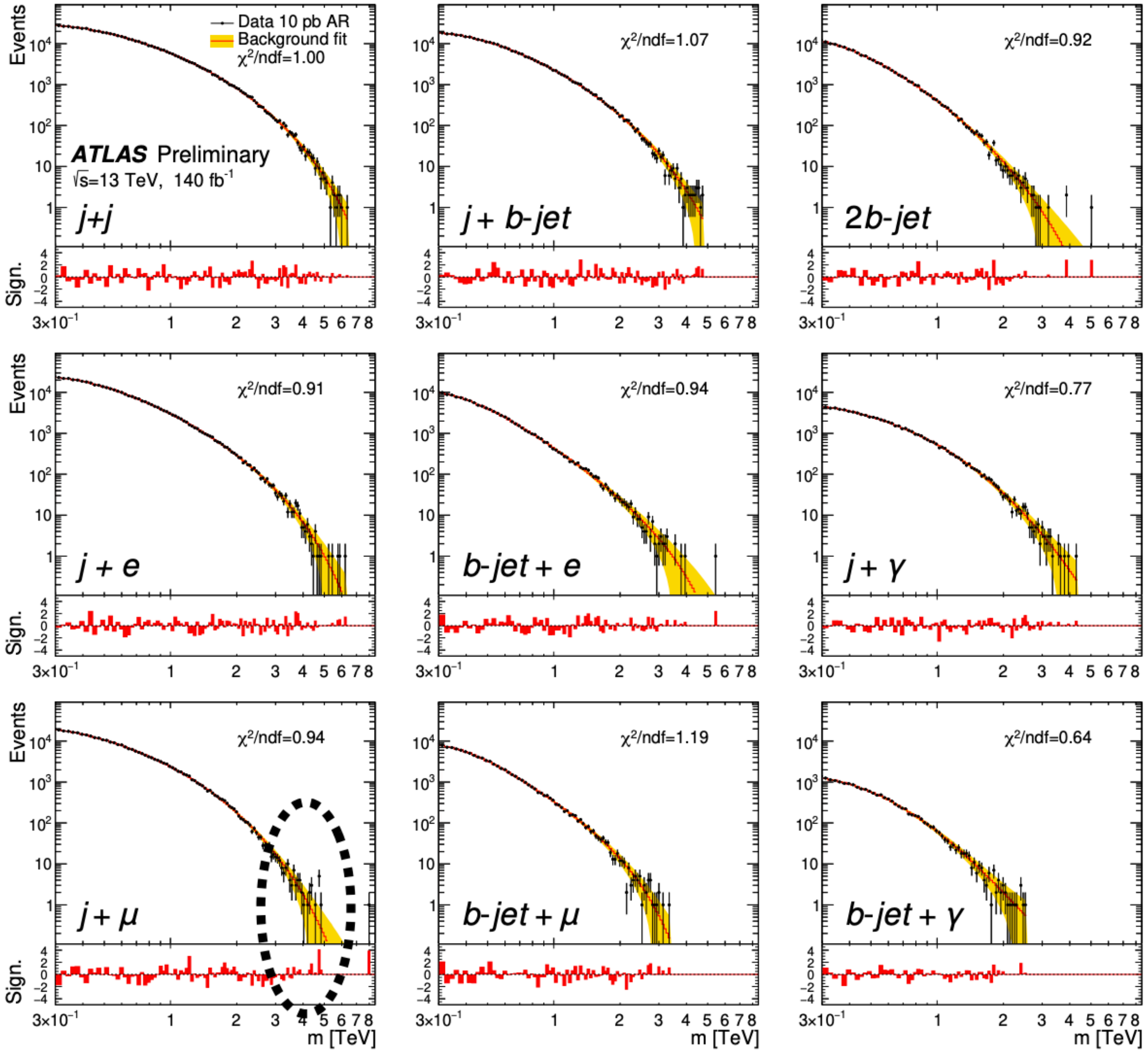
- ➔ A view on (variational) autoencoder for anomaly detection
  - ❖ Training on SM background jet → **anomalous jet will produce outlier latent scores** → make selection on the score
- ➔ Use autoencoder for anomaly detection has industry basis



ATLAS's model-agnostic search



PRL 132 (2024) 081801



- ➔ ATLAS applies full-event-level anomaly detection
- ➔ Train "autoencoder" and select on the score
- ➔ Search in 9 invariant masses including di-jet, di-b-jet, with three anomaly regions

Conclusion

➔ We highlight recent advances in ML for jets from three aspects

- **novel algorithms (GNN and Transformers) and their applications**
    - hope to shed light across experiments regarding network designs: why they are found helpful to process particle-format data
    - ATLAS/CMS's most recent developments in jet tagging (Transformers are the leading designs)
    - ParticleNet's broad application across the fields (and possibly Particle Transformer and/or its alternatives shortly)
- **model inferencing facilities in ATLAS/CMS**
    - ONNX runtime and SONIC
- **novel ML paradigm (model-agnostic search) at the LHC**
    - weakly-supervised approach and autoencoder approach
    - their applications in ATLAS/CMS