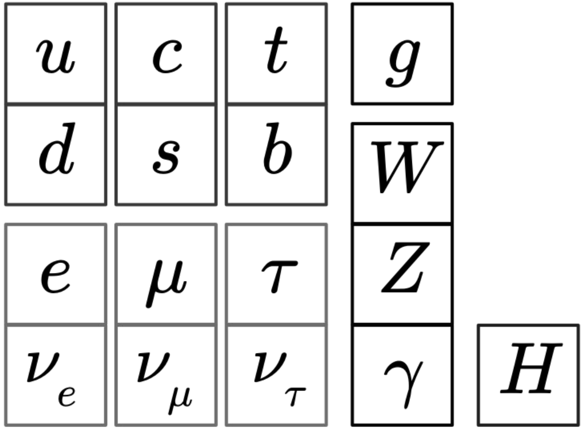
How AI can help uncover the mysteries of the Universe at the LHC

Benedikt Maier (Imperial College London)
Mar 5, 2024

From **astrophysical observations** like rotational curves or gravitational lensing:

→ Dark matter, **5x more abundant** than visible matter



Many other open questions in cosmology and particle physics

Experiment-driven:

Dark Energy

Matter-antimatter asymmetry

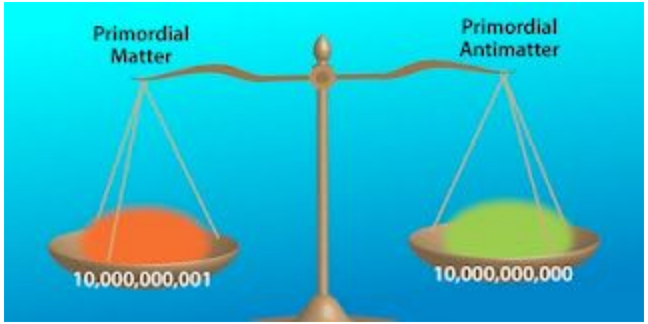
Theory-driven:

Hierarchy problems (weakness of gravity, fine tuning at level 10^{16})

Number of fermion generations (why 3?)

→ New physics within LHC reach?

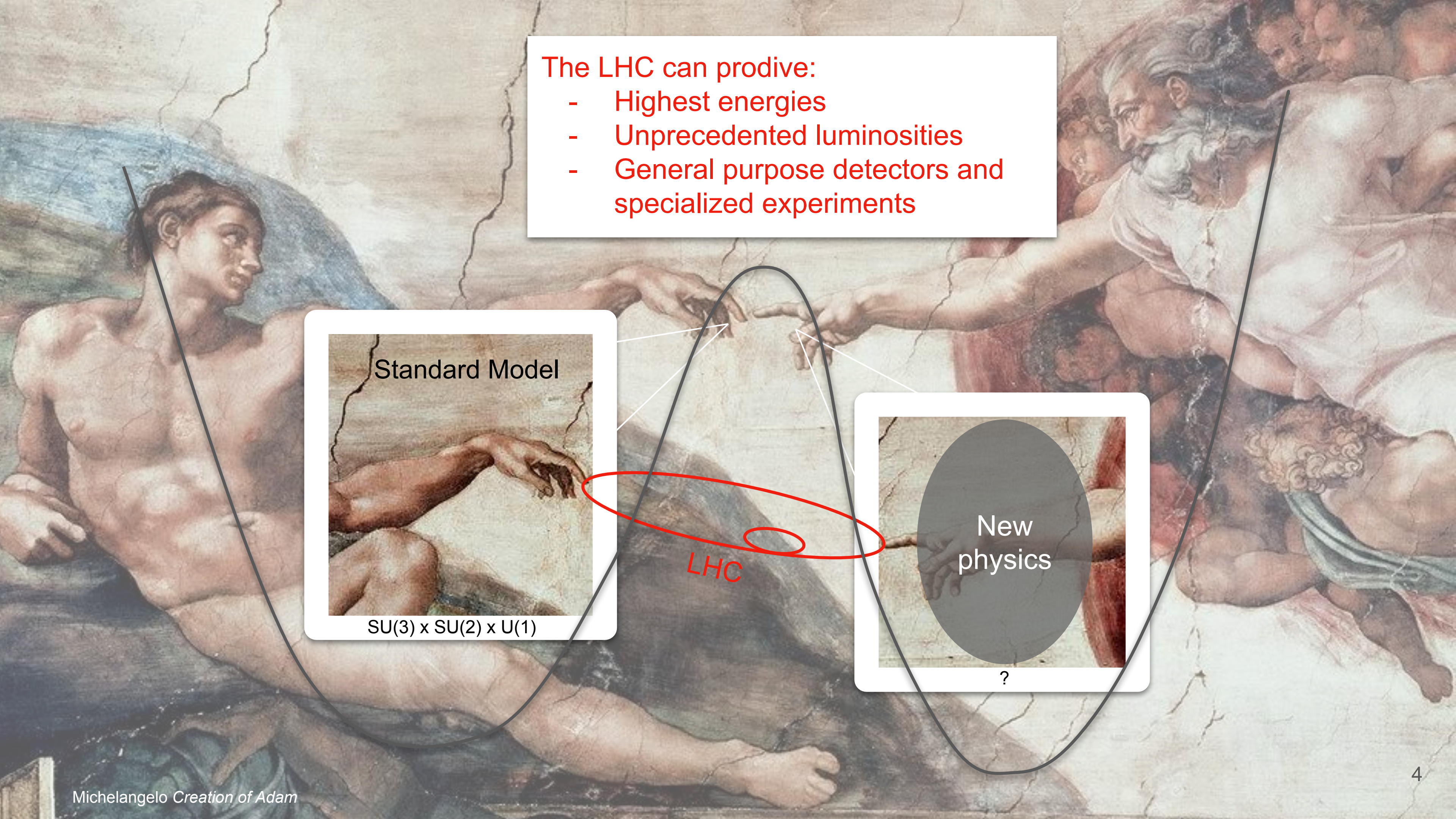
michaelgstrauss.com



quantumdiaries.org

The LHC can prodive:

- Highest energies
- Unprecedented luminosities
- General purpose detectors and specialized experiments



Michelangelo *Creation of Adam*

New physics yield.
Maximize this!

$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

New physics yield.
Maximize this!

$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

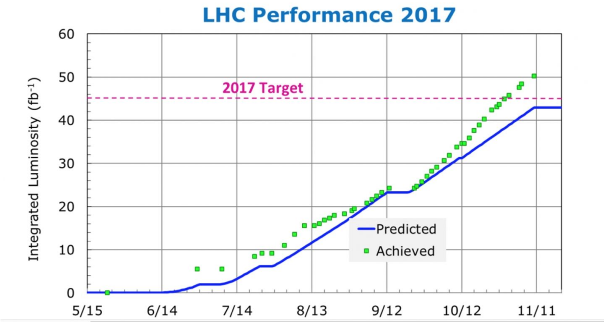
Cross section

Home | CERN
https://home.cern › news › news · Diese Seite übersetzen

LHC Run 3: physics at record energy starts tomorrow

04.07.2022 — With the increased **data samples and higher collision energy**, Run 3 will further expand the already very diverse LHC physics programme.

Integrated luminosity



New physics yield.
Maximize this!

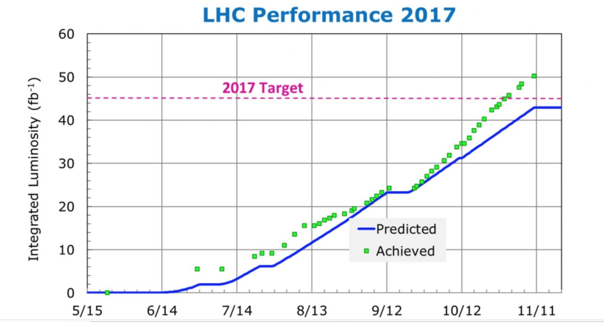
$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

Cross section

Home | CERN
https://home.cern › news › news · Diese Seite übersetzen

LHC Run 3: physics at record energy starts tomorrow

04.07.2022 — With the increased **data samples and higher collision energy**, Run 3 will further expand the already very diverse LHC physics programme.



New physics yield.
Maximize this!

Integrated
luminosity

$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

Cross section

Detector acceptance

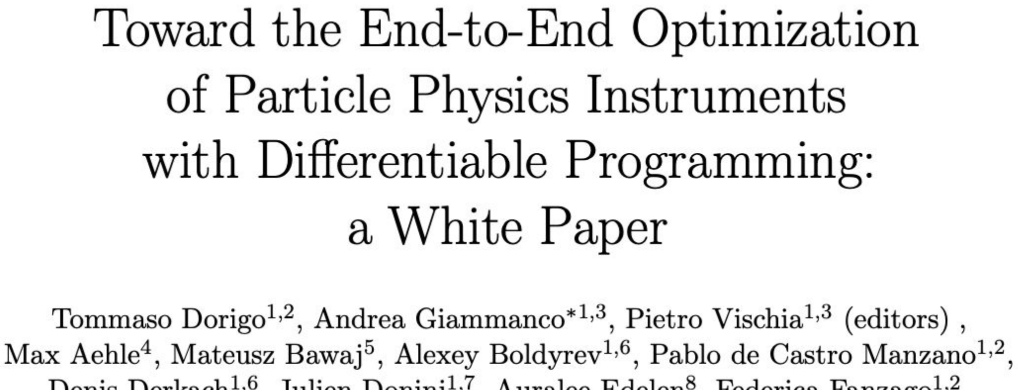
Home | CERN
https://home.cern › news › news · Diese Seite übersetzen

LHC Run 3: physics at record energy starts tomorrow

04.07.2022 — With the increased **data samples and higher collision energy**, Run 3 will further expand the already very diverse LHC physics programme.

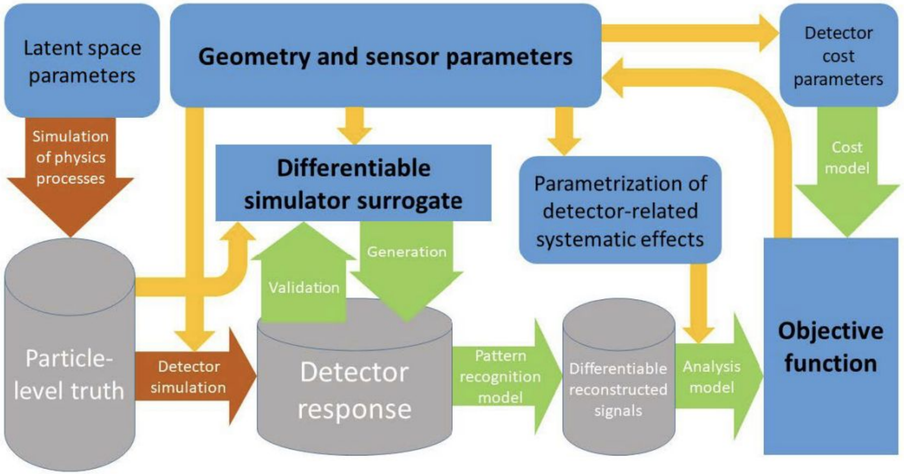
AI to improve detector acceptance

- AI can help in designing detectors
- Relatively new effort on several fronts (e.g., MODE collaboration)
- End-to-end optimization via differentiable programming

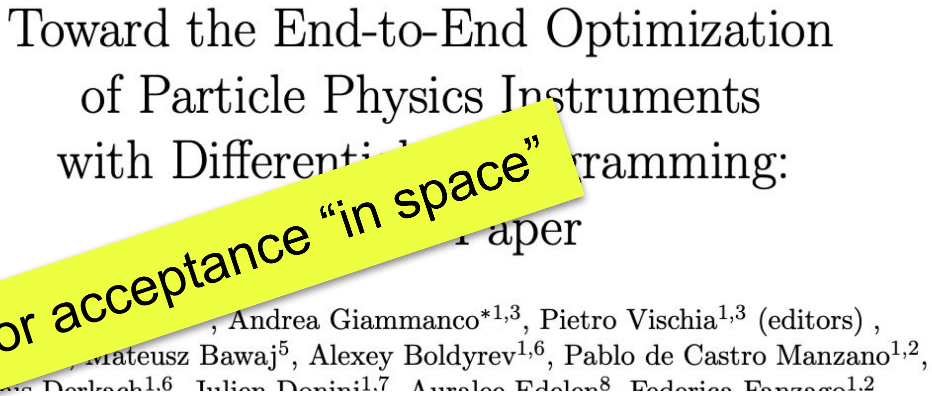


Toward the End-to-End Optimization of Particle Physics Instruments with Differentiable Programming: a White Paper

Tommaso Dorigo[1,2], Andrea Giammanco*[1,3], Pietro Vischia[1,3] (editors) ,
Max Aehle[4], Mateusz Bawaj[5], Alexey Boldyrev[1,6], Pablo de Castro Manzano[1,2],



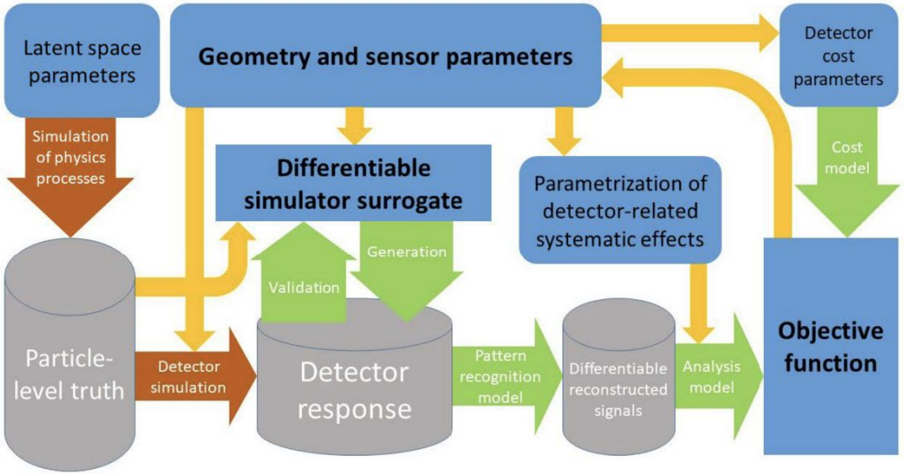
AI to improve detector acceptance


Toward the End-to-End Optimization of Particle Physics Instruments with Differenti... ...ramming: ... Paper

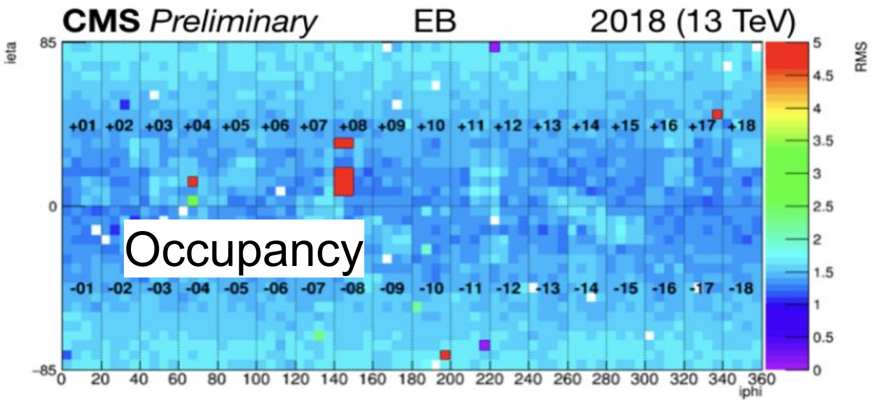
..., Andrea Giammanco*[1,3], Pietro Vischia[1,3] (editors) , ...Mateusz Bawaj[5], Alexey Boldyrev[1,6], Pablo de Castro Manzano[1,2],

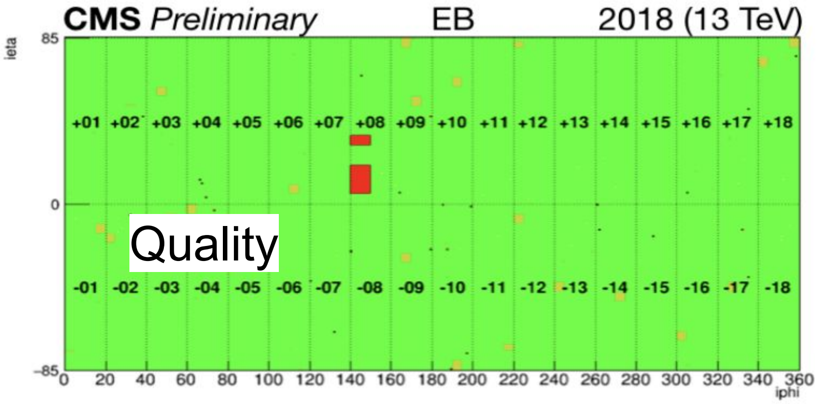
AI to improve detector acceptance “in space”

- AI can help in designing detectors
- Relatively new effort on several fronts (e... collaboration...
- End-to-end optimization via differentiable programming



AI to improve detector acceptance “in time” (aka detector downtime)

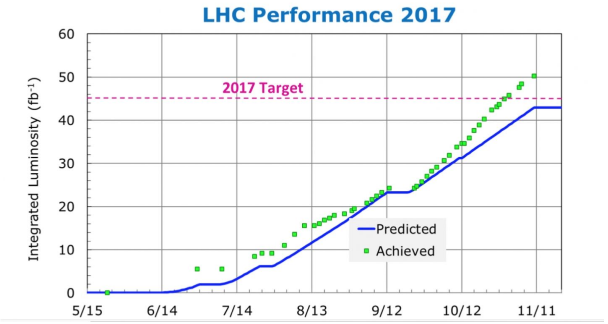




- Autoencoder-based online data quality monitoring in place at CMS
- Quickly identify and diagnose broad range of issues that would hinder physics quality data taking
- Identifying transient bad towers (pointing to deteriorating channels)

→ Better at anticipating / preventing long downtimes!

Integrated luminosity



New physics yield. Maximize this!

$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

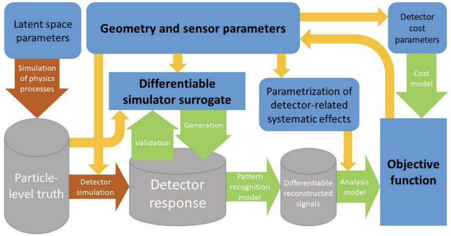
Detector acceptance

Cross section

Home | CERN
https://home.cern › news › news · Diese Seite übersetzen

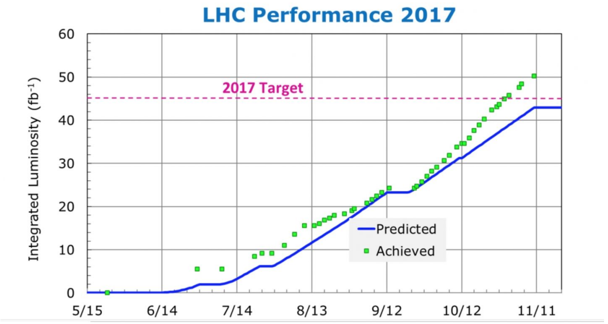
LHC Run 3: physics at record energy starts tomorrow

04.07.2022 — With the increased **data samples and higher collision energy**, Run 3 will further expand the already very diverse LHC physics programme.



New physics yield. Maximize this!

Integrated luminosity



$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

Efficiency, (being) revolutionized by ML!

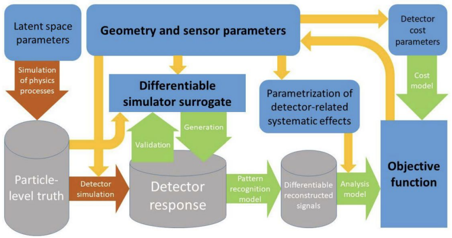
Detector acceptance

Cross section

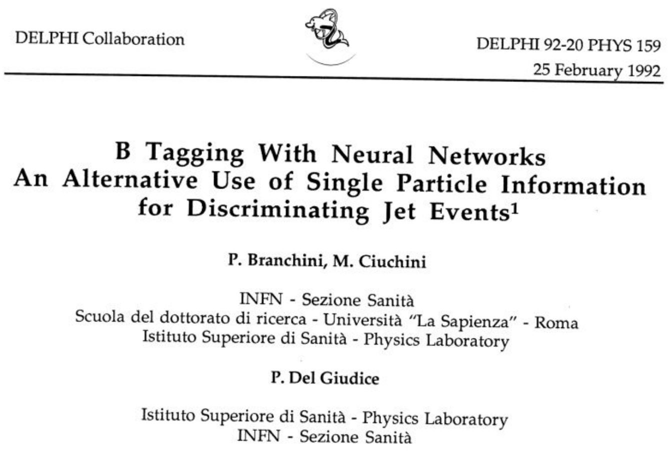
Home | CERN
https://home.cern › news › news · Diese Seite übersetzen

LHC Run 3: physics at record energy starts tomorrow

04.07.2022 — With the increased **data samples and higher collision energy**, Run 3 will further expand the already very diverse LHC physics programme.



HEP community: early adopters of Machine Learning


DELPHI Collaboration

DELPHI 92-20 PHYS 159
25 February 1992

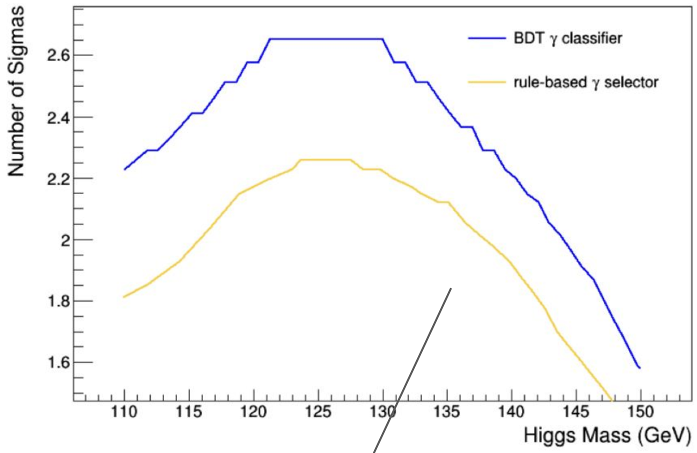
B Tagging With Neural Networks
An Alternative Use of Single Particle Information
for Discriminating Jet Events[1]

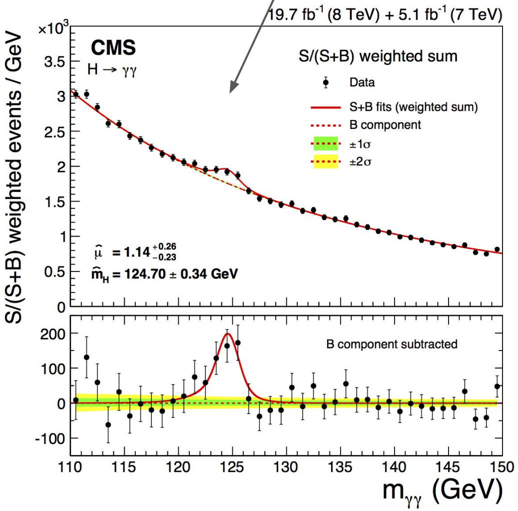
P. Branchini, M. Ciuchini

INFN - Sezione Sanità
Scuola del dottorato di ricerca - Università "La Sapienza" - Roma
Istituto Superiore di Sanità - Physics Laboratory

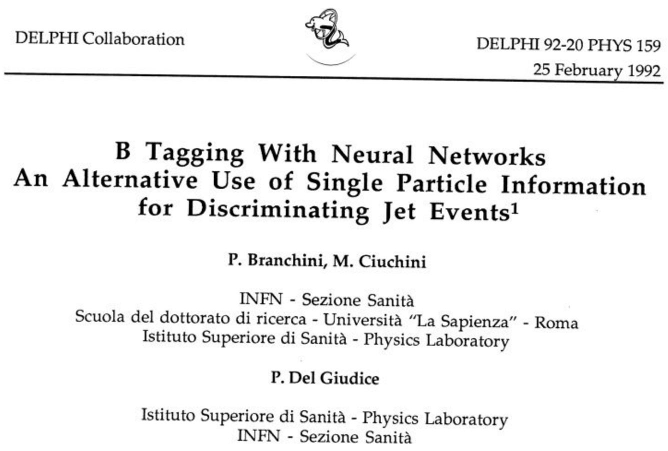
P. Del Giudice

Istituto Superiore di Sanità - Physics Laboratory
INFN - Sezione Sanità





HEP community: early adopters of Machine Learning


DELPHI Collaboration

DELPHI 92-20 PHYS 159
25 February 1992

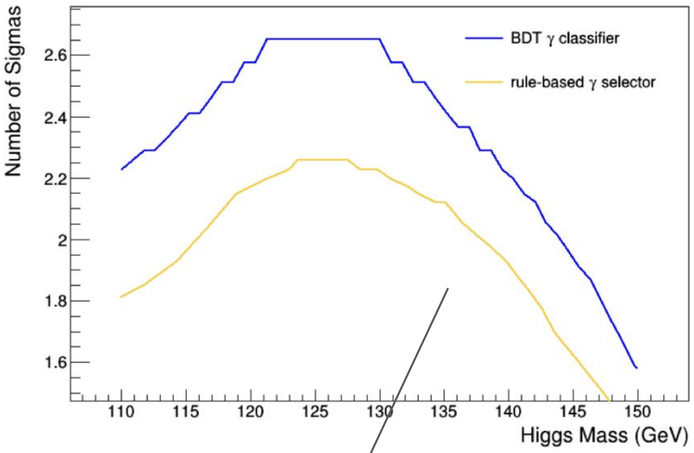
**B Tagging With Neural Networks
An Alternative Use of Single Particle Information
for Discriminating Jet Events[1]**

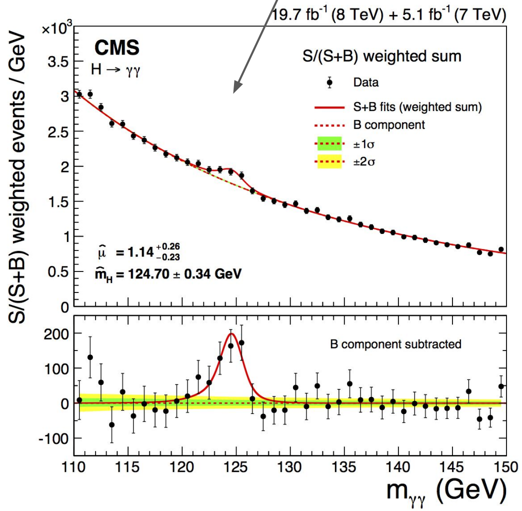
P. Branchini, M. Ciuchini

INFN - Sezione Sanità
Scuola del dottorato di ricerca - Università "La Sapienza" - Roma
Istituto Superiore di Sanità - Physics Laboratory

P. Del Giudice

Istituto Superiore di Sanità - Physics Laboratory
INFN - Sezione Sanità





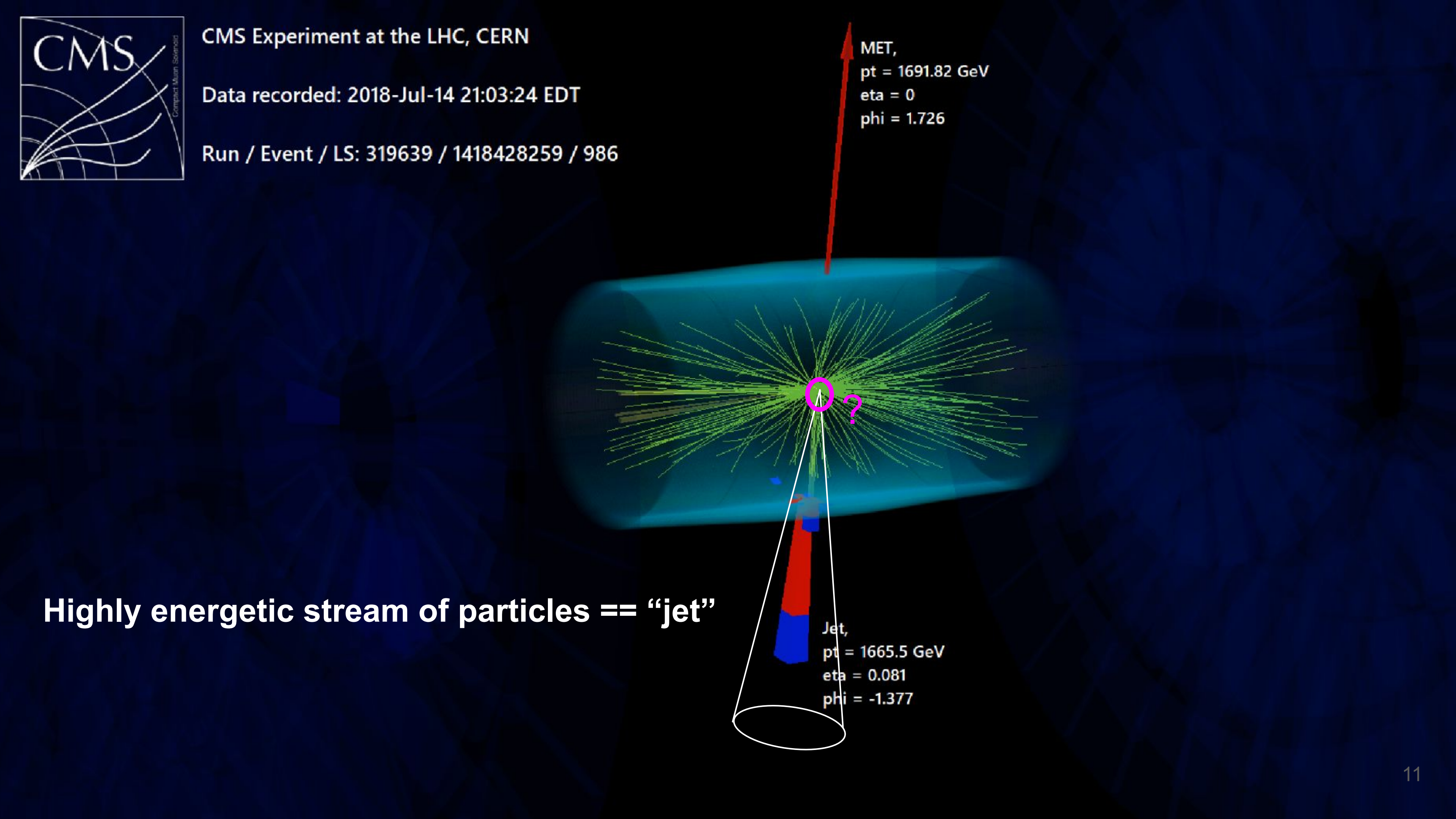
Of course today:

- More data
- Better/richer data
- AI-friendly hardware (e.g., GPUs)

→ Bodes to AI-based solutions

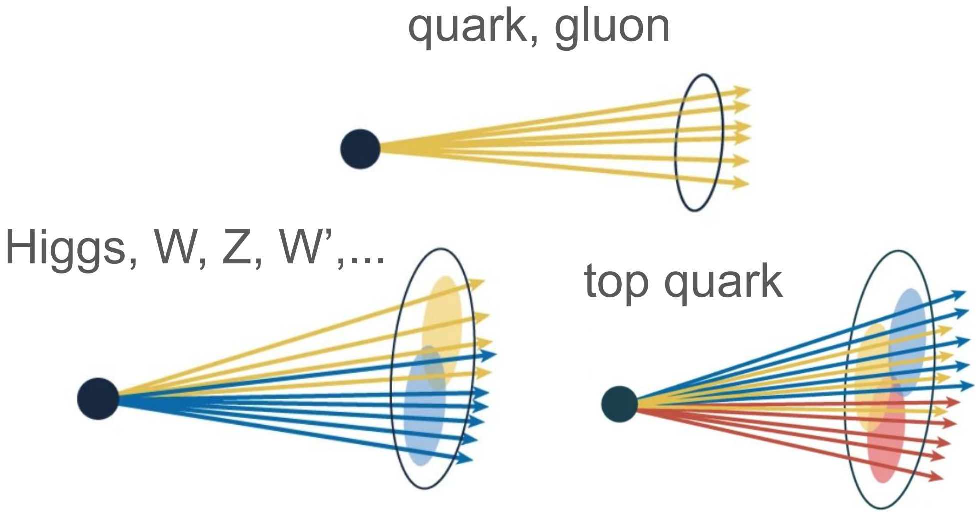
$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

Energy frontier



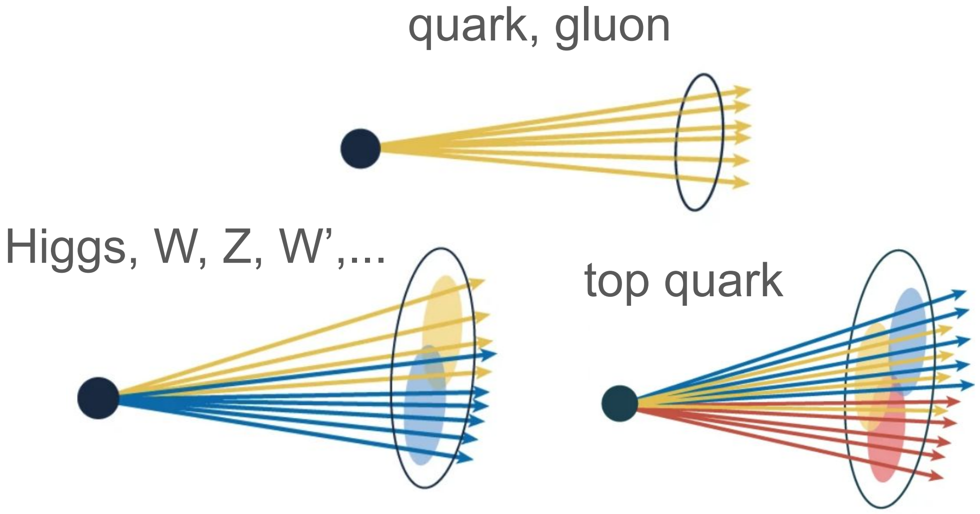
Highly energetic stream of particles == “jet”

~50-100 particles/jet
~50 features per particle
→ O(1000) features per jet



~50-100 particles/jet
~50 features per particle
→ O(1000) features per jet

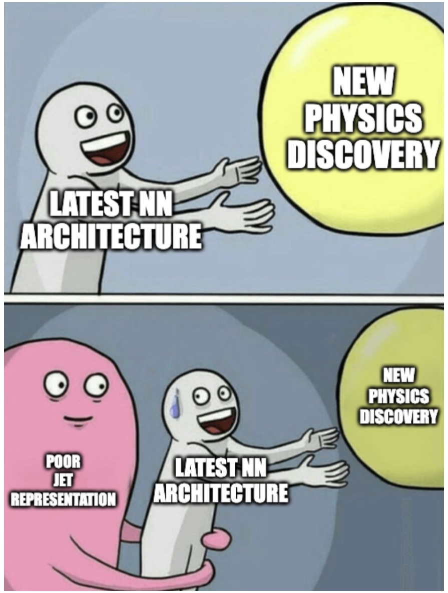
Which particle was at origin of jet evolution? → Jet tagging

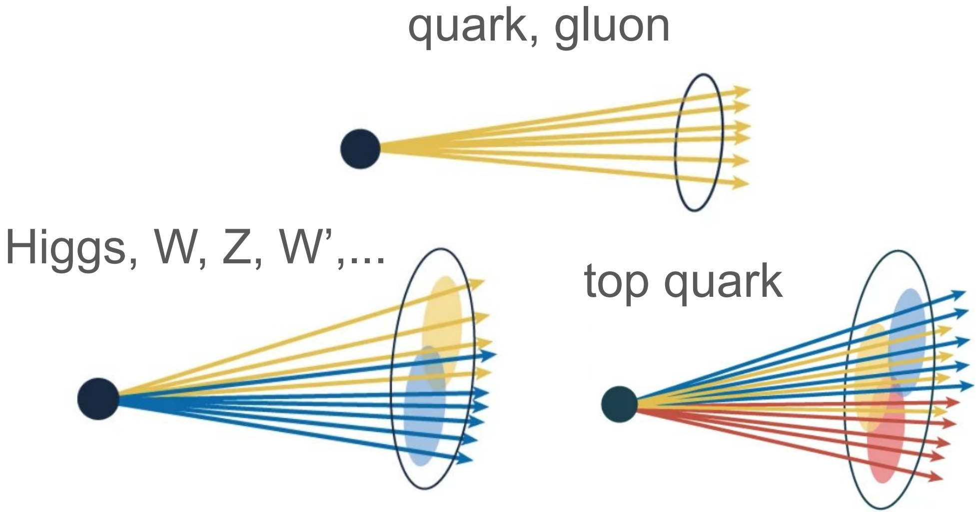


Jets **ideal environment** to accelerate machine learning-based solutions in high energy physics

~50-100 particles/jet
~50 features per particle
→ O(1000) features per jet

Which particle was at origin of jet evolution? → Jet tagging

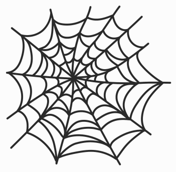


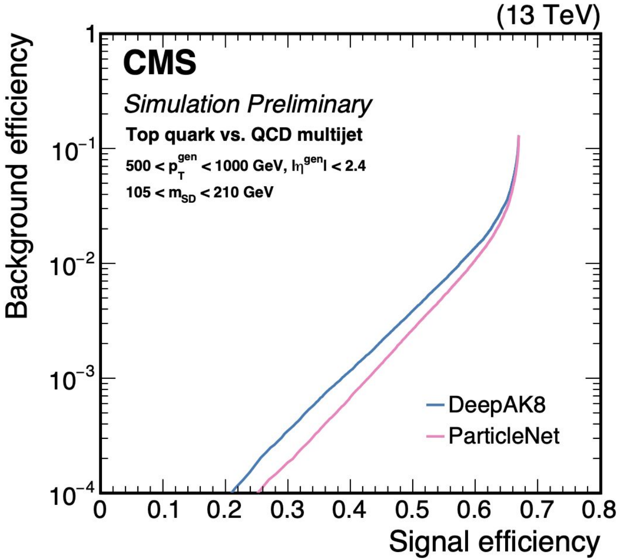


Jets **ideal environment** to accelerate machine learning-based solutions in high energy physics

Key: **match** between jet representation and ML architecture

A particle net to tag heavy resonances

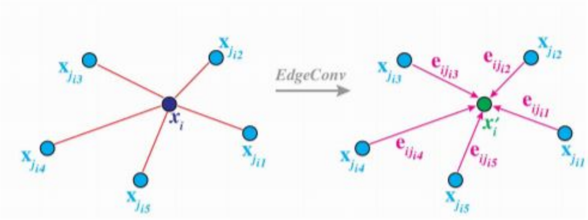




Particles in jet == sparse, unordered, variable-size set

→ Graph neural network

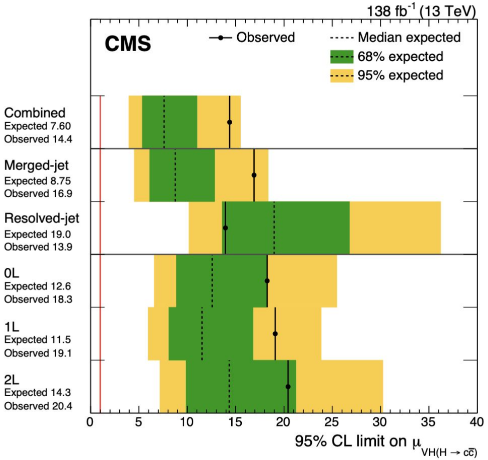
ParticleNet = current **state-of-the-art** in jet tagging



CMS expected upper limit on VH(cc): 7.8

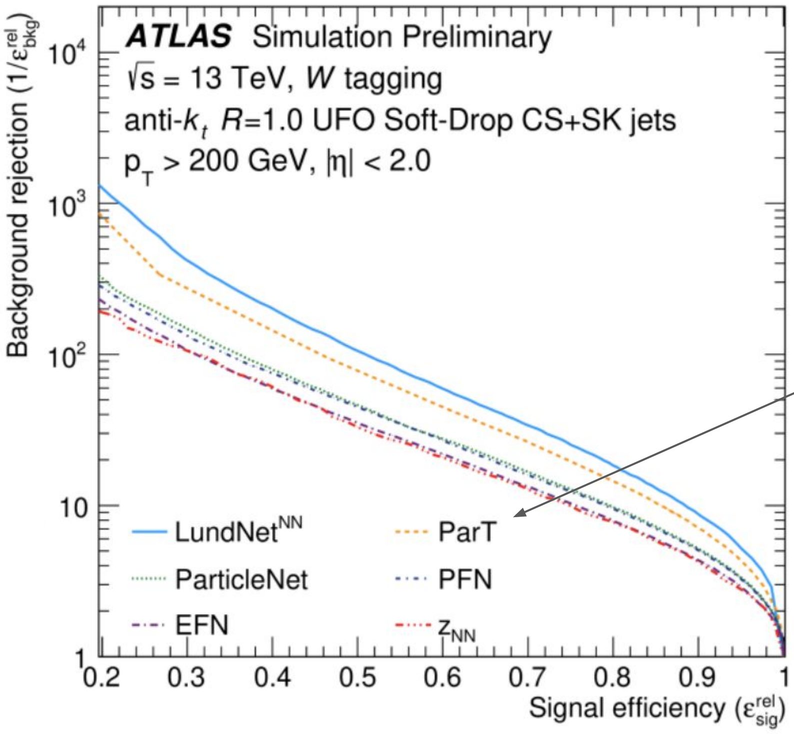
ATLAS: 31 !!!

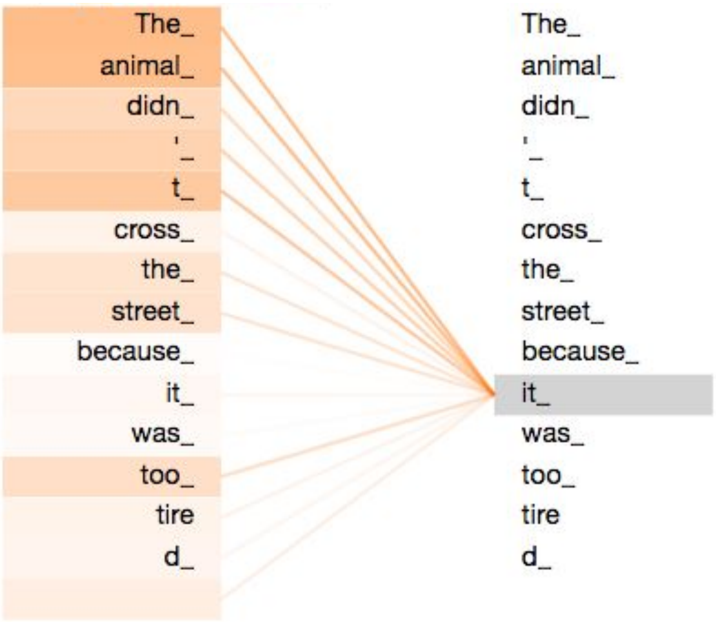
As if Run-2 lasted **16 times** longer for CMS.



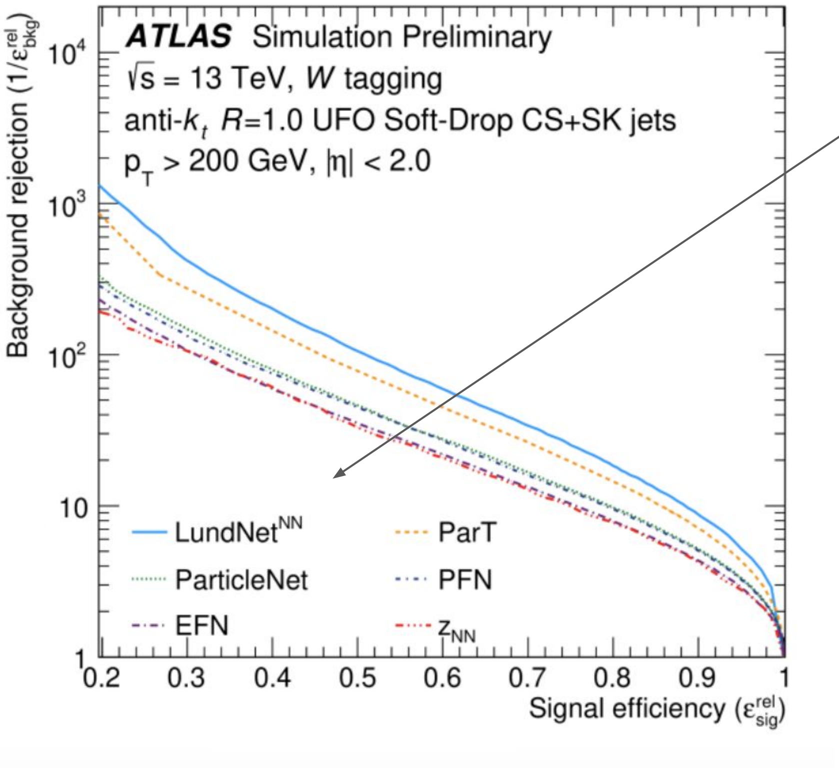
Next-generation jet tagging

Strong correlations between particles in jet → Transformer-based architectures

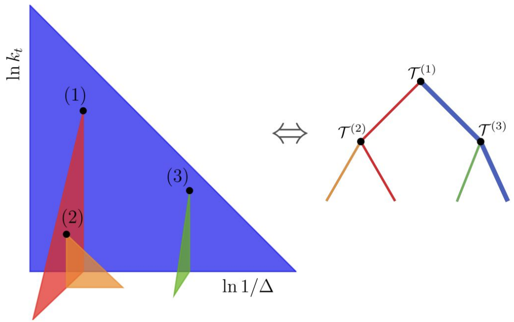


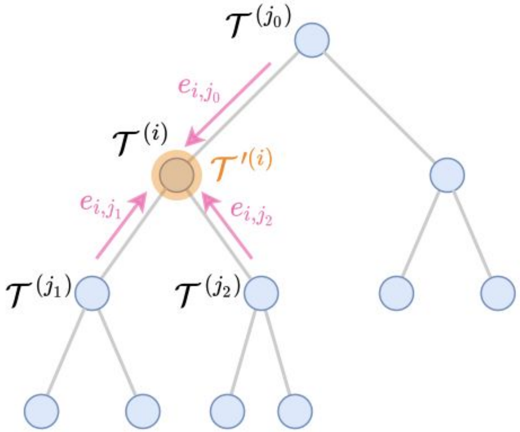


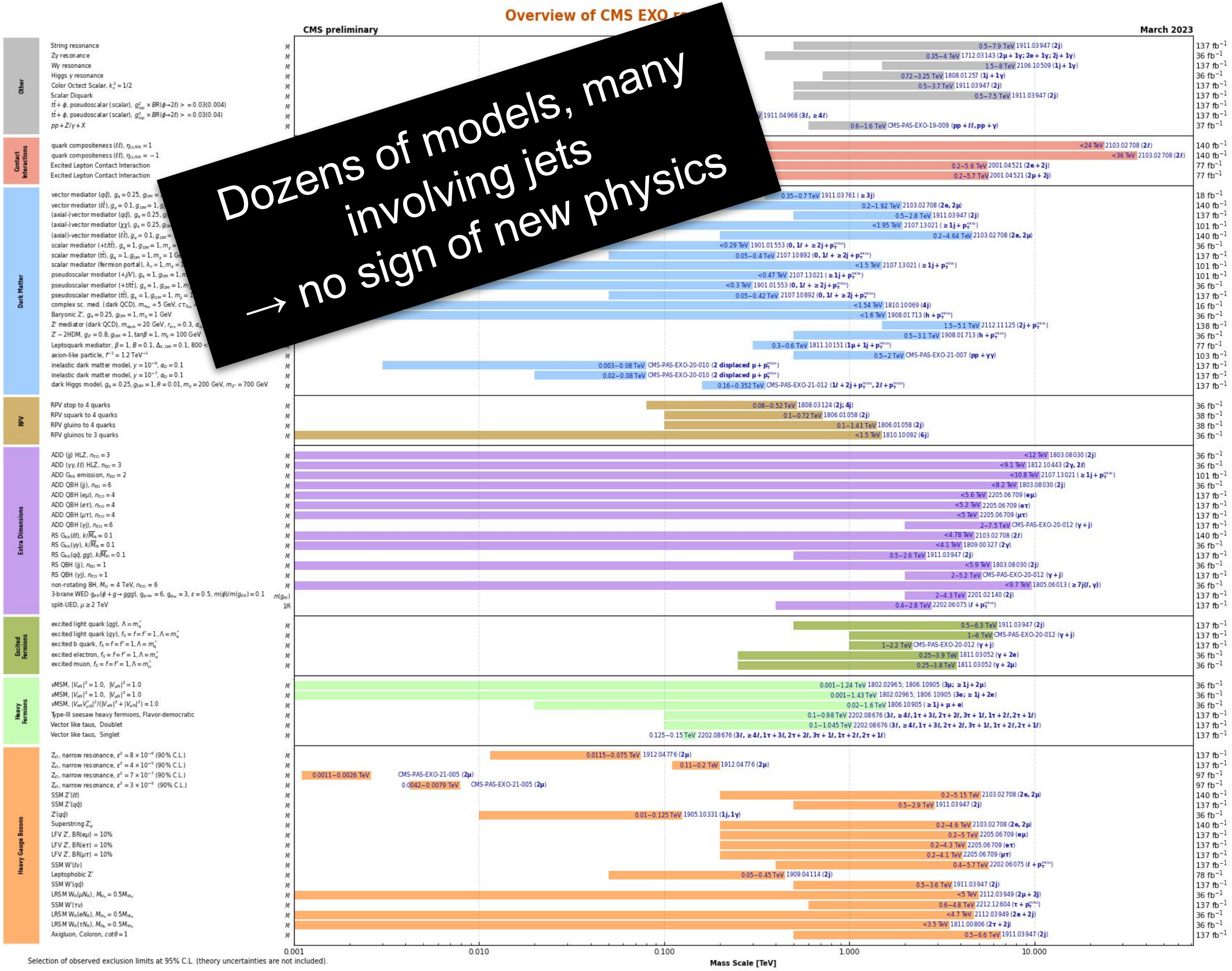
Next-generation jet tagging



Shower / clustering history reveals a lot about jet origin → represent jet in Lund plane

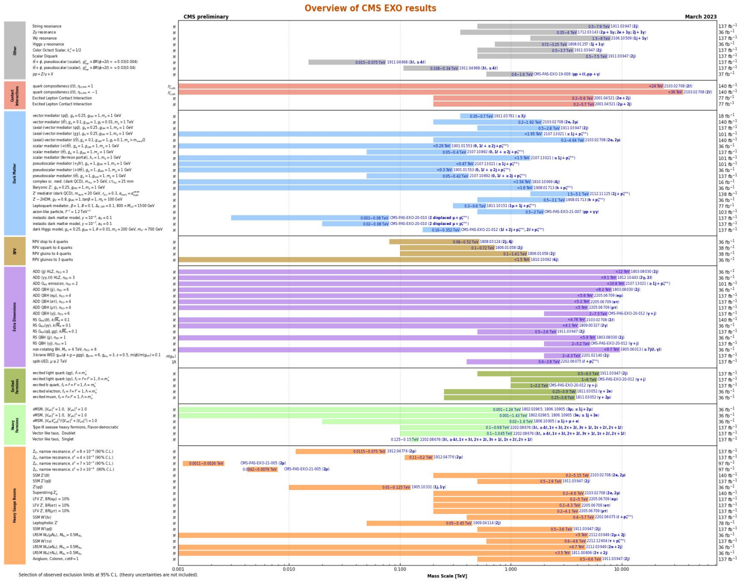






Dozens of models, many involving jets
→ no sign of new physics

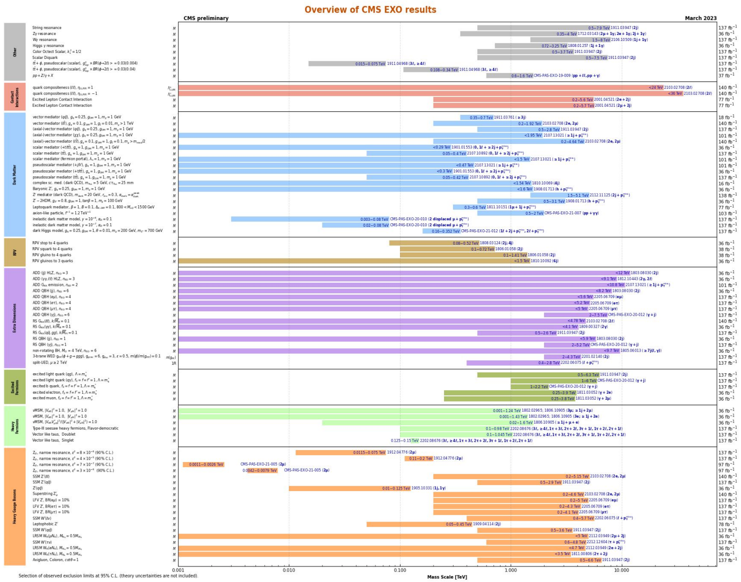
Are we searching in the wrong places / for the wrong jets?



Maybe looking in the wrong spots or for the wrong models?

→ Need **safeguard** against missing signs of new physics

Are we searching in the wrong places / for the wrong jets?



Maybe looking in the wrong spots or for the wrong models?

→ Need **safeguard** against missing signs of new physics

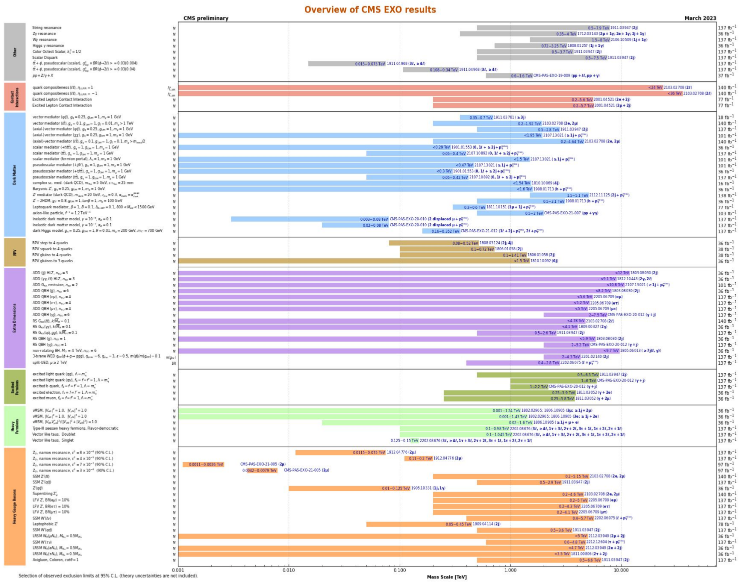
Re-formulate the question

“Does this event look like BSM theory XYZ?”

↓

“Does this event look like the Standard Model?”

Are we searching in the wrong places / for the wrong jets?



Maybe looking in the wrong spots or for the wrong models?

→ Need **safeguard** against missing signs of new physics

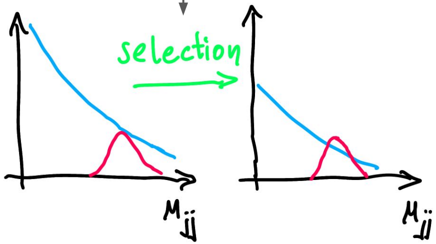
Re-formulate the question

"Does this event look like BSM theory XYZ?"

"Does this event look like the Standard Model?"

→ Anomaly detection

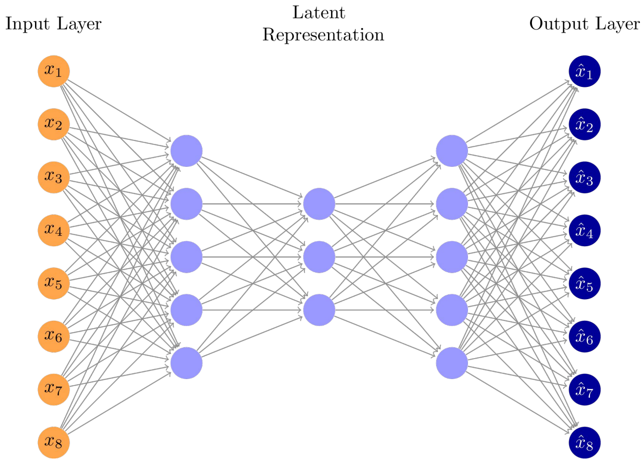
→ **AI**



How do you identify anomalous jets?

Learn QCD, look for outliers

Variational Autoencoder



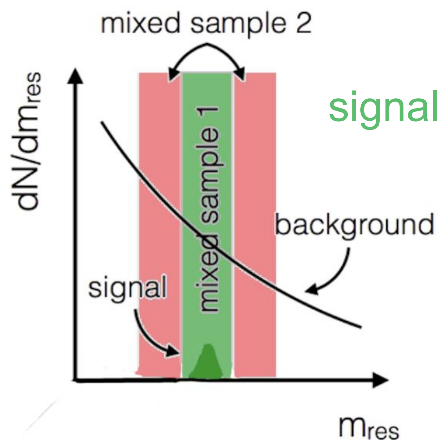
Anomaly metric:

Reconstruction error

Increasing Model Dependence

How do you identify anomalous jets?

background-rich



signal-rich

Train a classifier between two samples: Signal-rich vs. background-rich

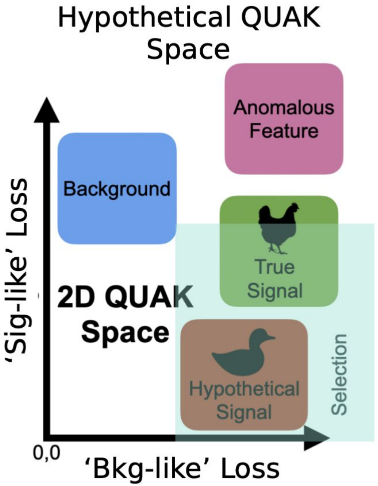
Weak Supervision

Anomaly metric:

Classifier score

Increasing Model Dependence

How do you identify anomalous jets?



Two autoencoders - one for background, one for mixture of signals

Anomaly metric:

Area in loss-loss plane

Encode a 'prior' of potential anomalies, look for similar

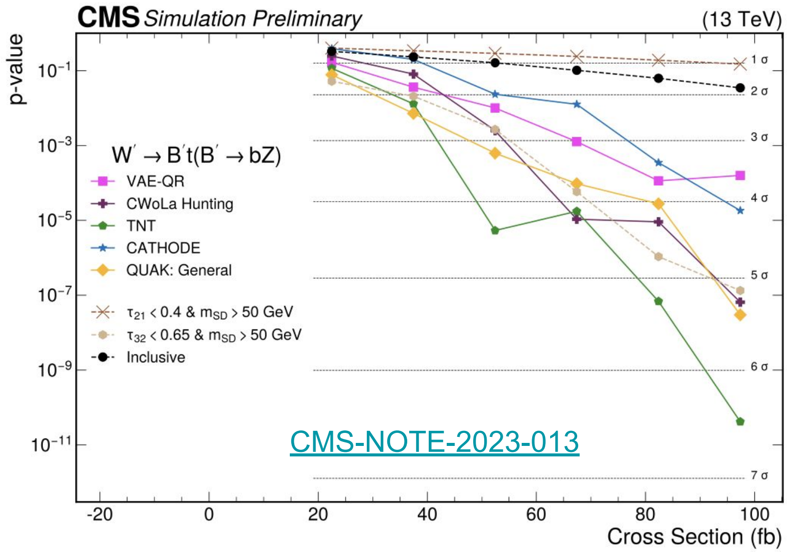
"Quasi Anomalous Knowledge"

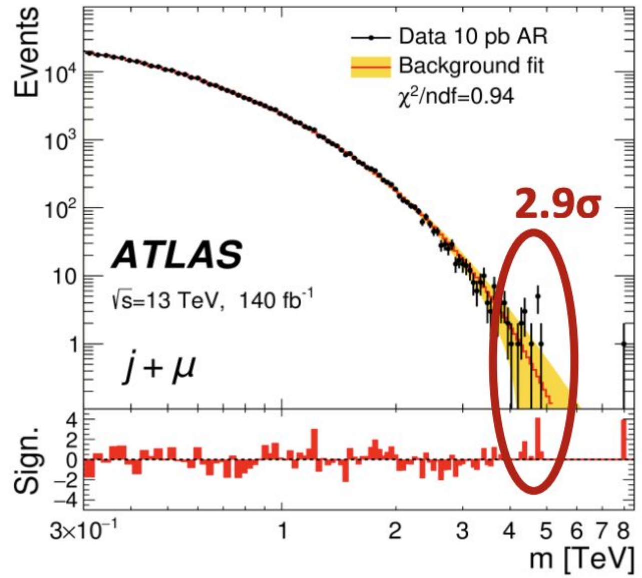
Increasing Model Dependence

Anomaly searches …

First time comparison of different anomaly detection strategies. More to come soon!

… at CMS





… at ATLAS

Autoencoder-based search

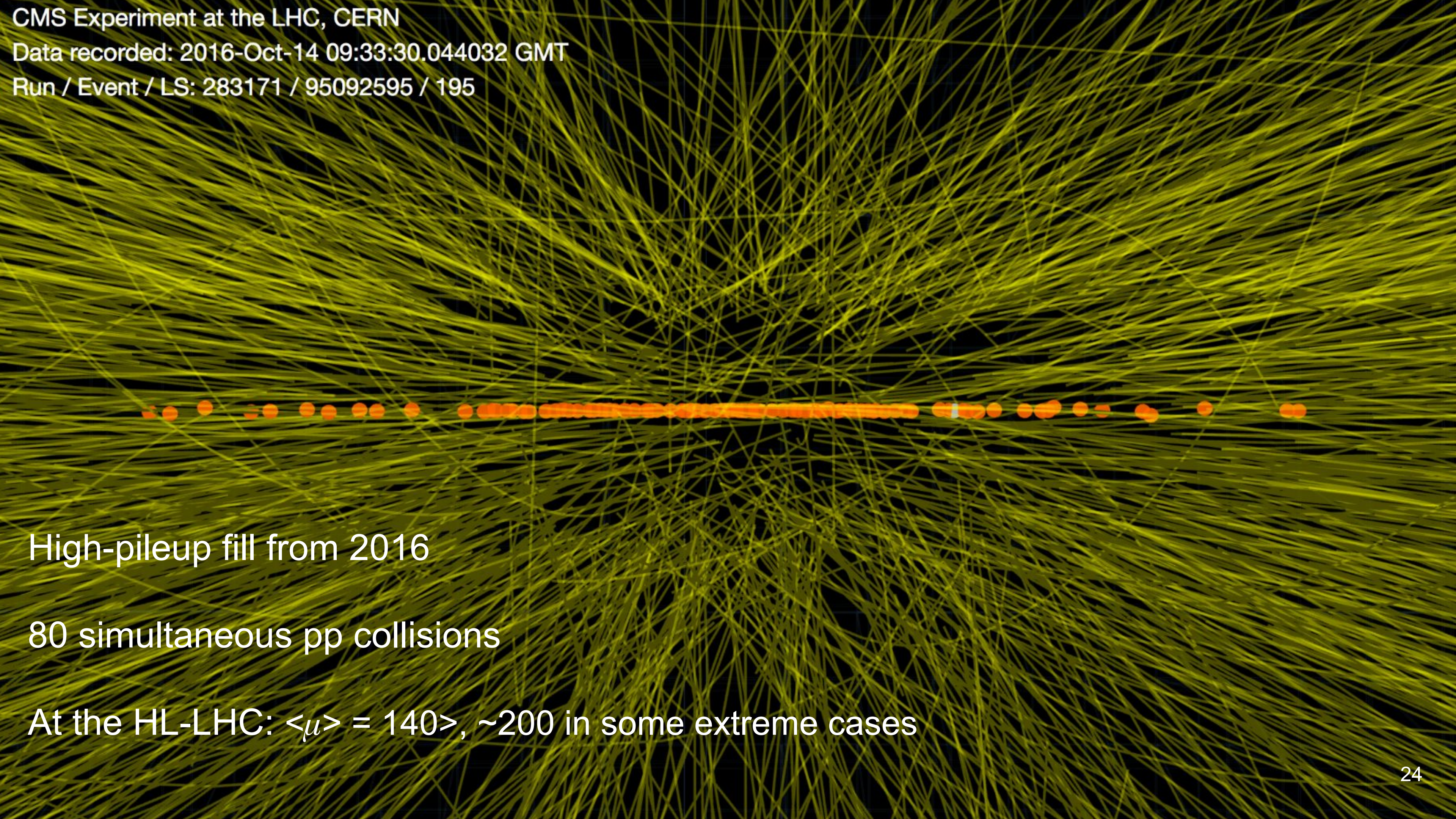
Looking at all possible 2-body final states (not only 2 jets)

Small excess in j+muon final state

$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

Energy frontier

Intensity frontier

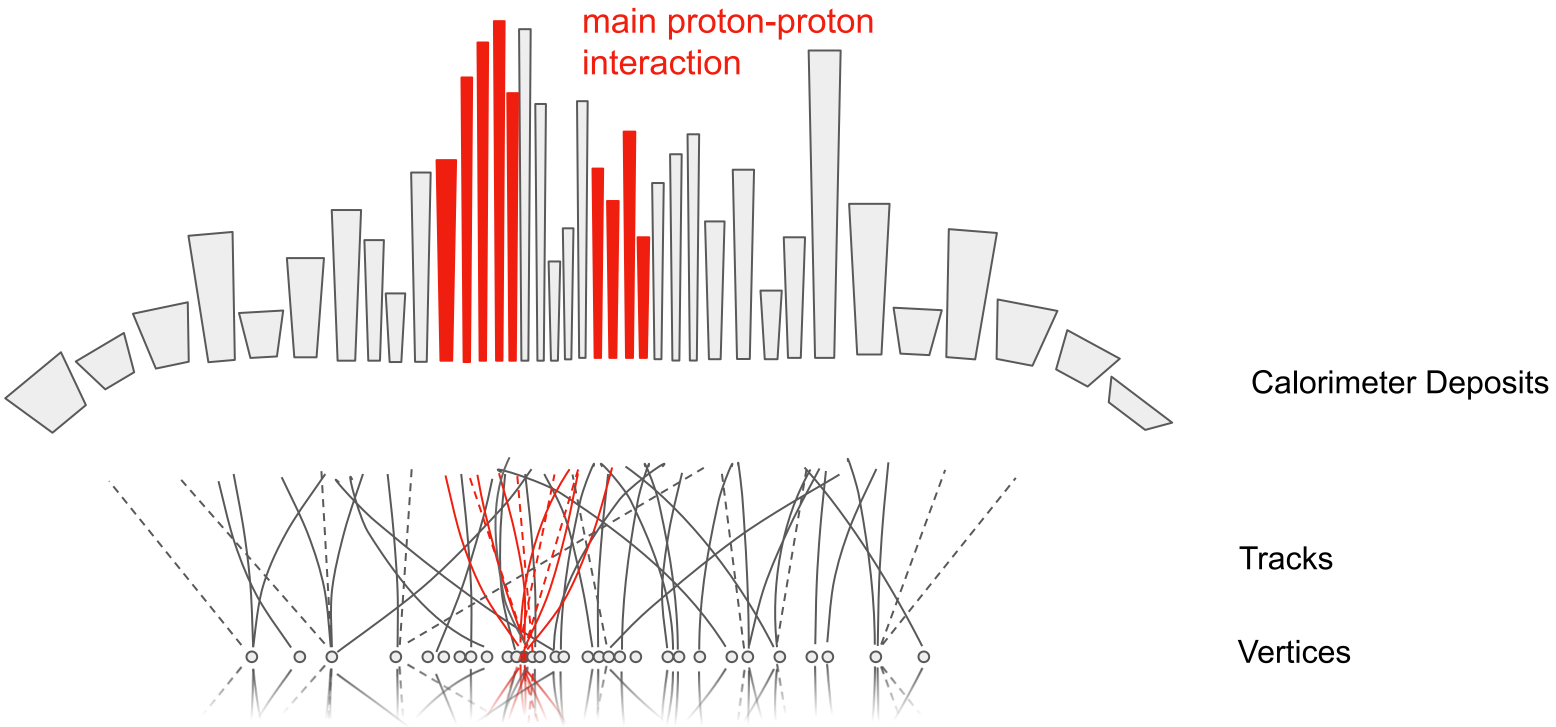


High-pileup fill from 2016

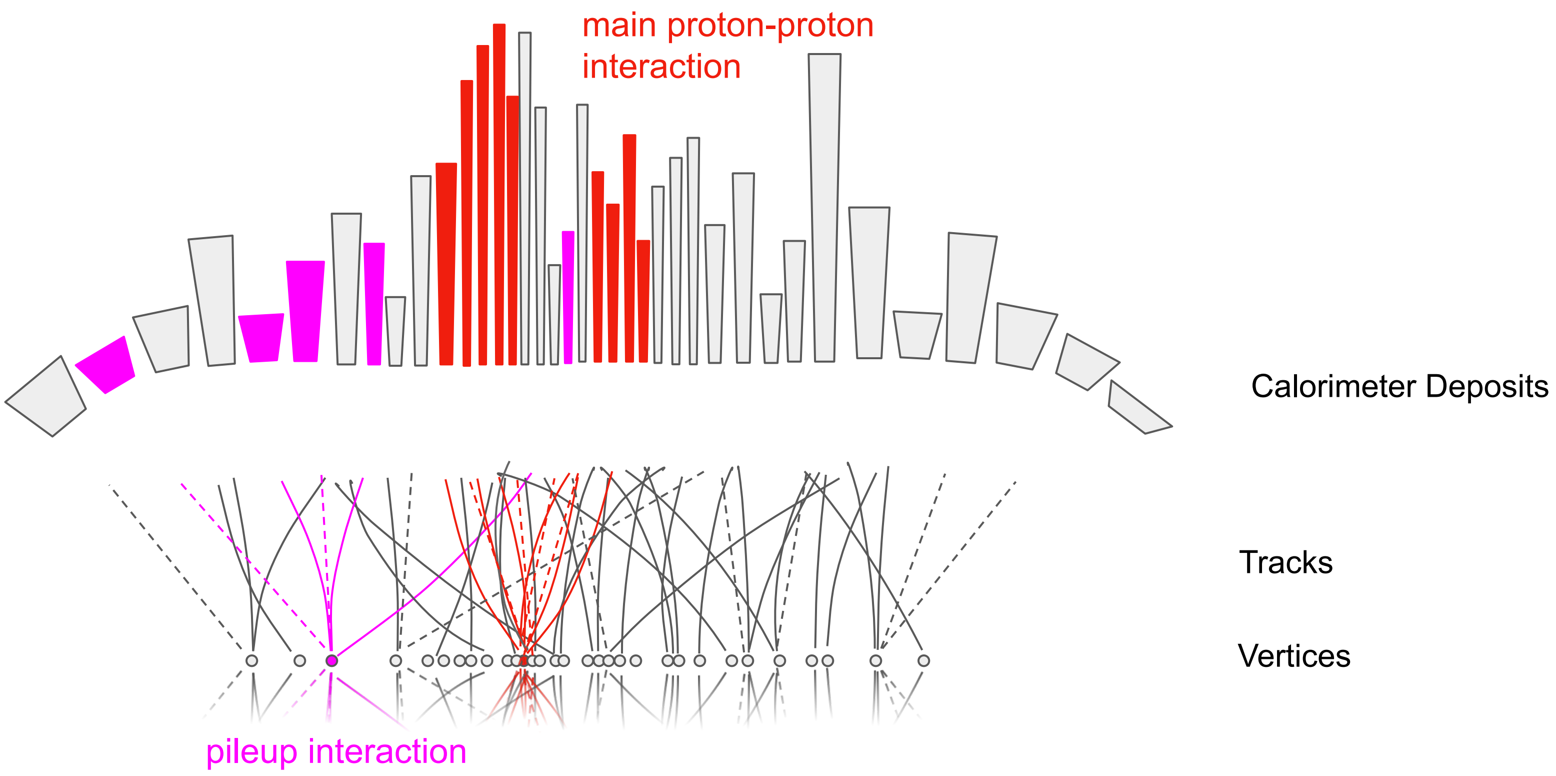
80 simultaneous pp collisions

At the HL-LHC: <μ> = 140>, ~200 in some extreme cases

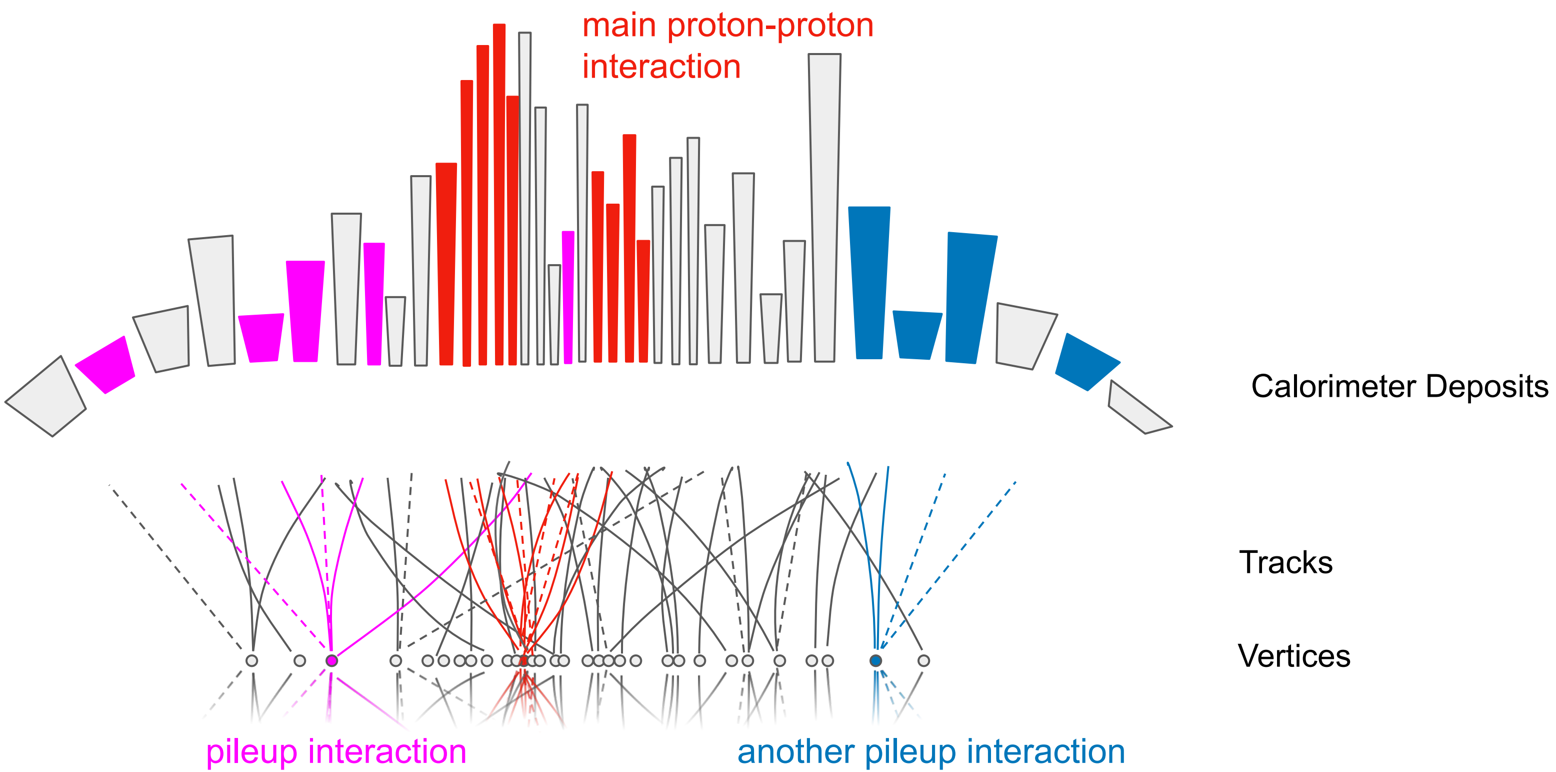
The downsides of a high-intensity hadron collider



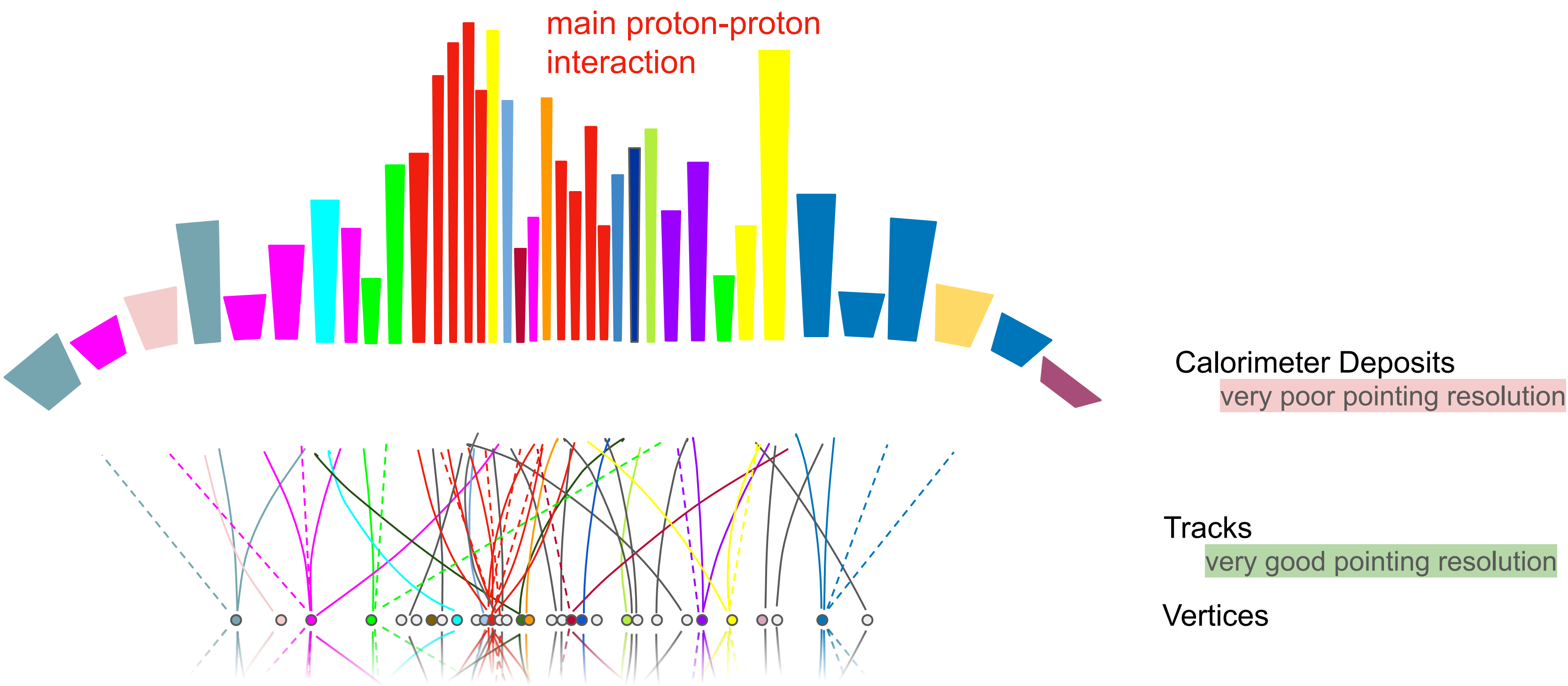
The downsides of a high-intensity hadron collider



The downsides of a high-intensity hadron collider

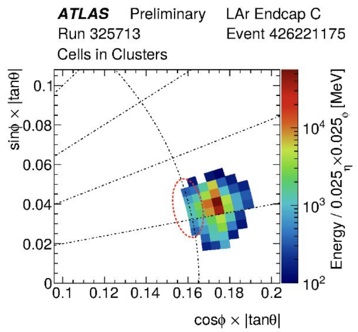


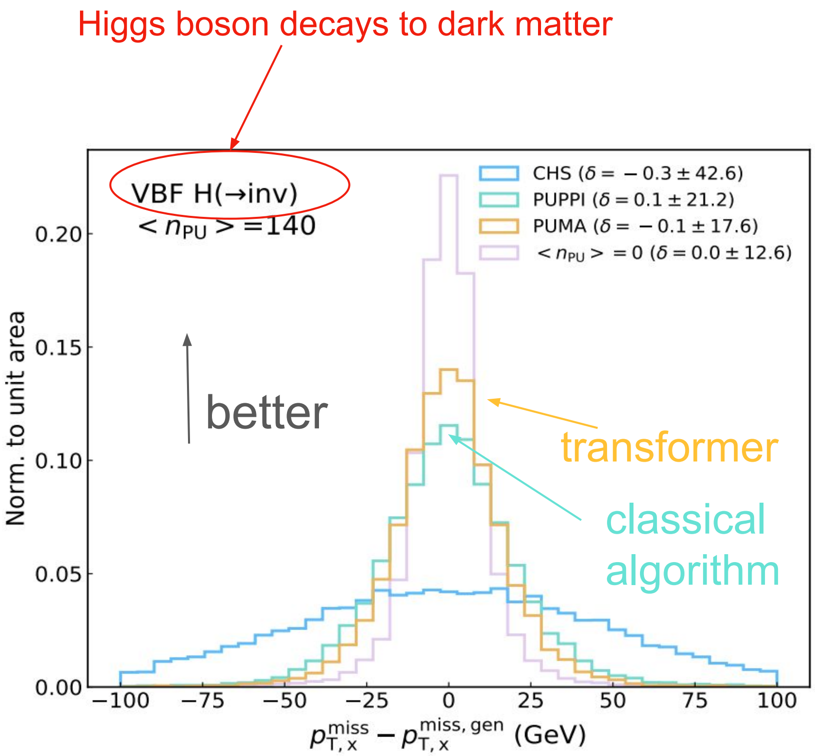
The downsides of a high-intensity hadron collider

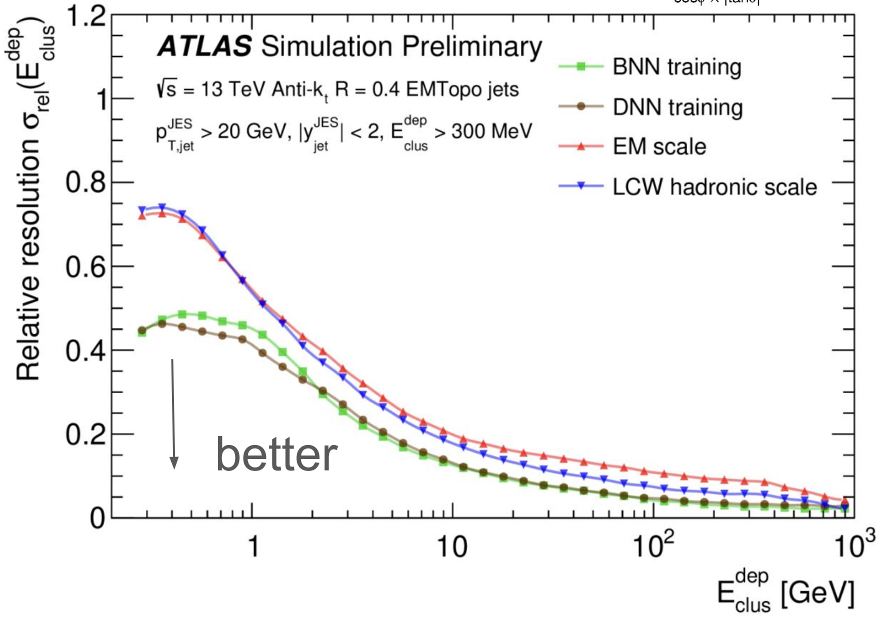


Ability to disentangle collisions absolutely essential
for entire physics program of LHC

Large improvements through machine learning







AI-based calibration of clusters → much improved pileup mitigation/resolution (25%-50% gain)

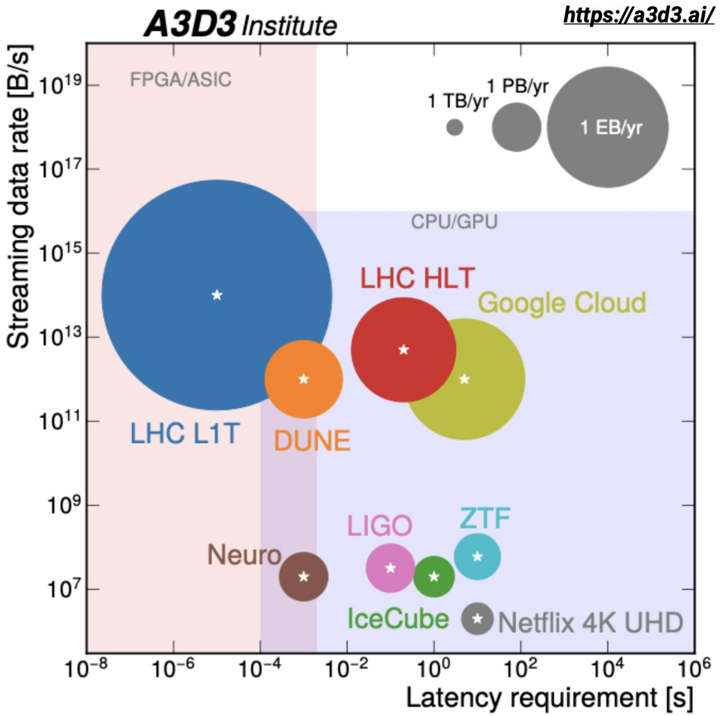
$$N = \sigma \cdot \mathcal{L} \cdot A \cdot \epsilon$$

Energy frontier

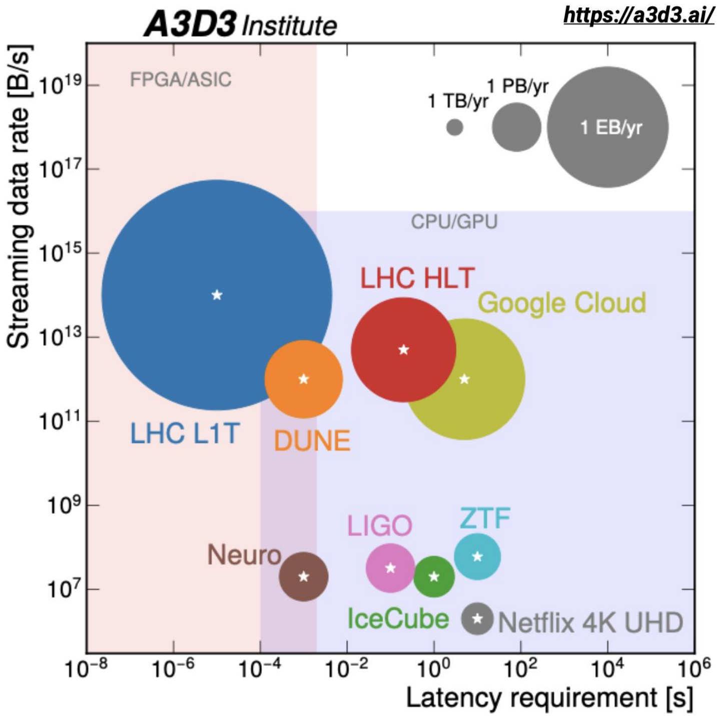
Intensity frontier

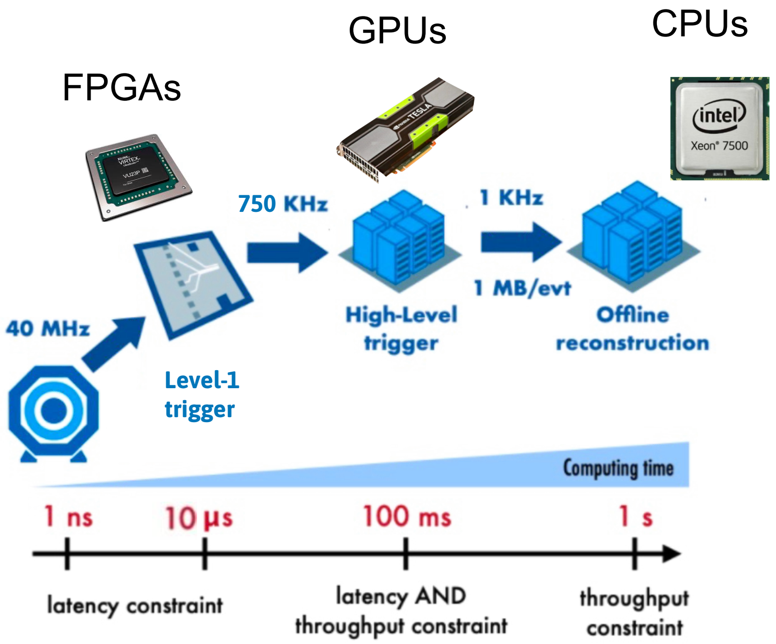
Rate frontier

LHC: Surpassing Google Cloud in terms of data rates

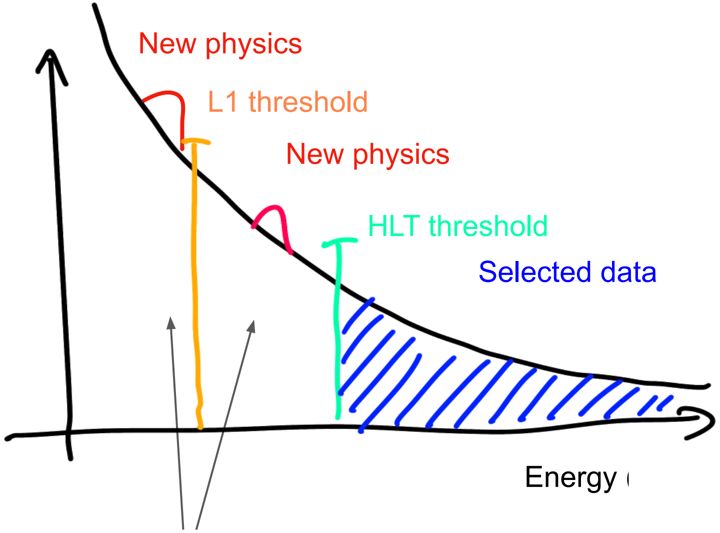


Two-tiered trigger system to make irreversible decisions





The worst-case scenario



What if new physics resides below trigger thresholds?

Could be losing 100% of BSM events at L1.
Cannot afford this! Duty to exploit full LHC potential!

Cure: Better decisions at Level-1

Data intensity asks for high-throughput solutions

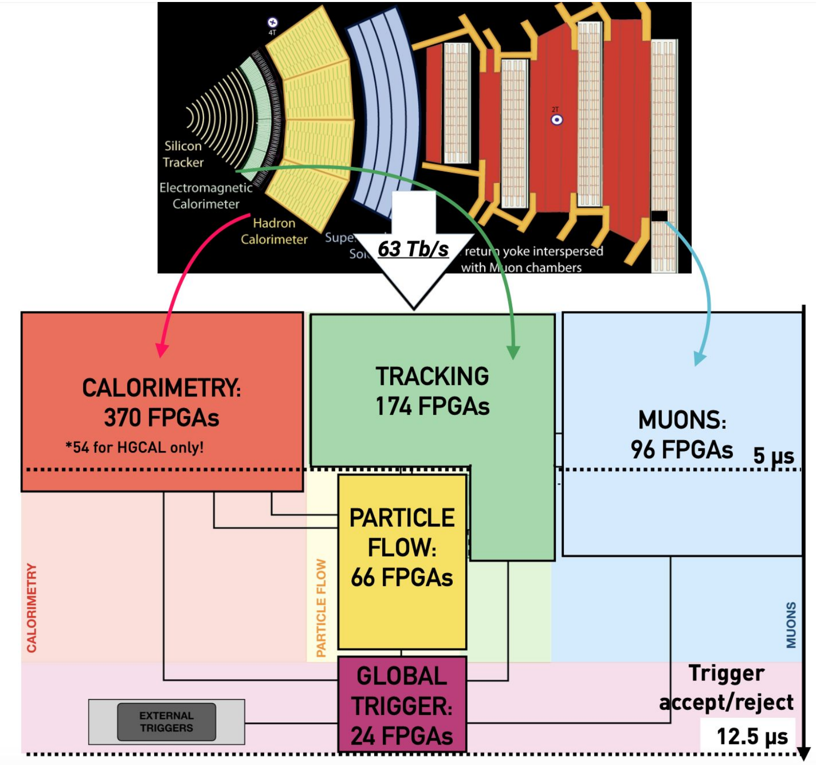
These are typically not the most accurate

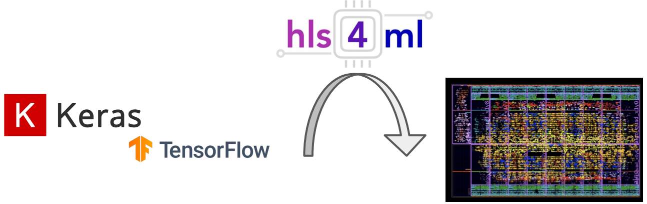
Already 20% more efficient signal selection **equivalent to running LHC 2-3 years longer**

→ Get help from AI

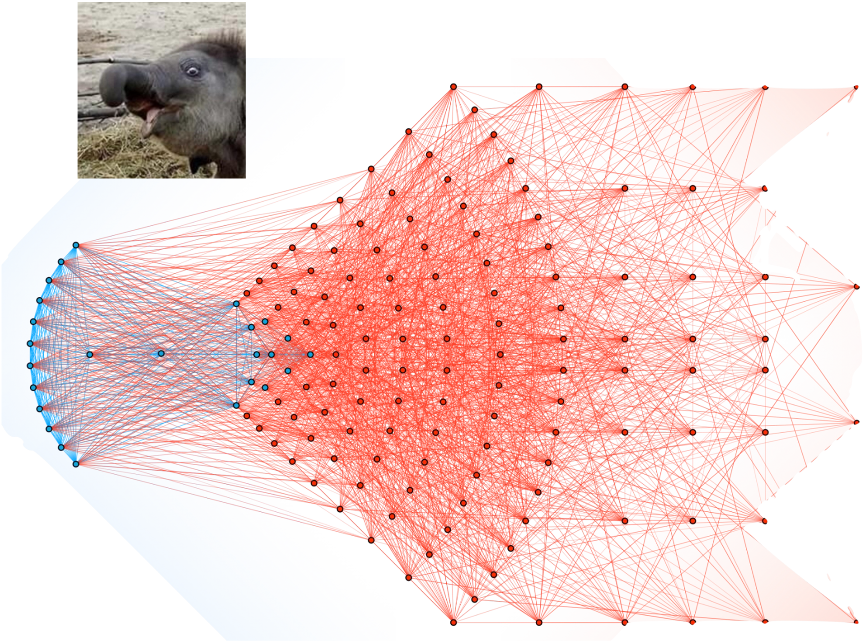
Models need to be ported onto FPGAs

Problem: little time (μs) at L1 to make decision → Real-time AI





The elephant in the room



How can we fit this on that?

FPGA in L1, latency $O(\mu s)$

Very large network such as transformer, graph neural network.
→ Inference time O(s)

A compression strategy is **not a nice-to-have**. It is a **necessity.**

Model quantization

Article | Published: 21 June 2021

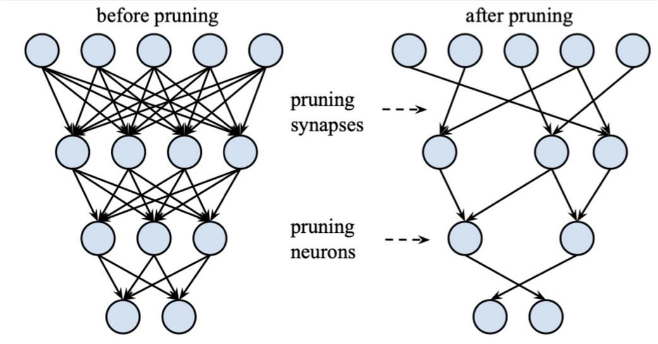
Automatic heterogeneous quantization of deep neural networks for low-latency inference on the edge for particle detectors

Claudionor N. Coelho Jr, Aki Kuusela, Shan Li, Hao Zhuang, Jennifer Ngadiuba, Thea Klaeboe Aarrestad, Vladimir Loncar, Maurizio Pierini, Adrian Alan Pol & Sioni Summers

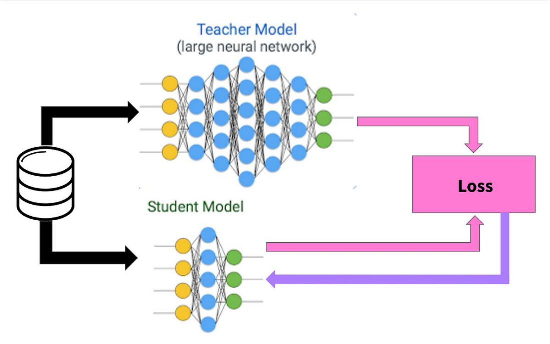
Nature Machine Intelligence **3**, 675–686 (2021) | Cite this article

32bit → 8bit
1/2 memory
20x less power

Model pruning



Model distillation

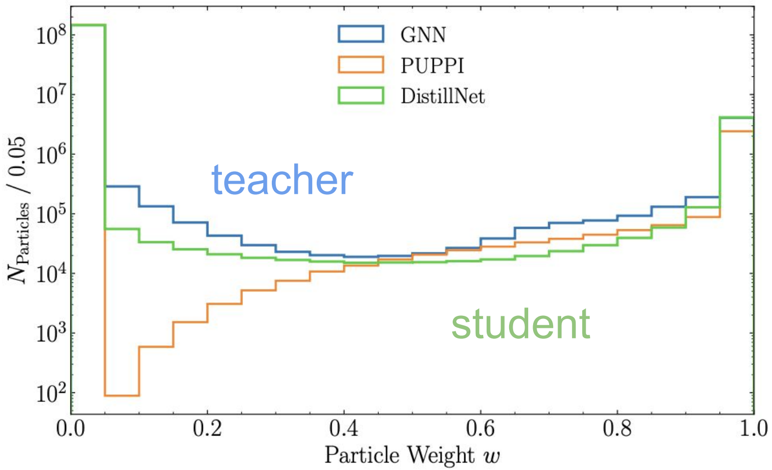


Two recent Knowledge Distillation Examples

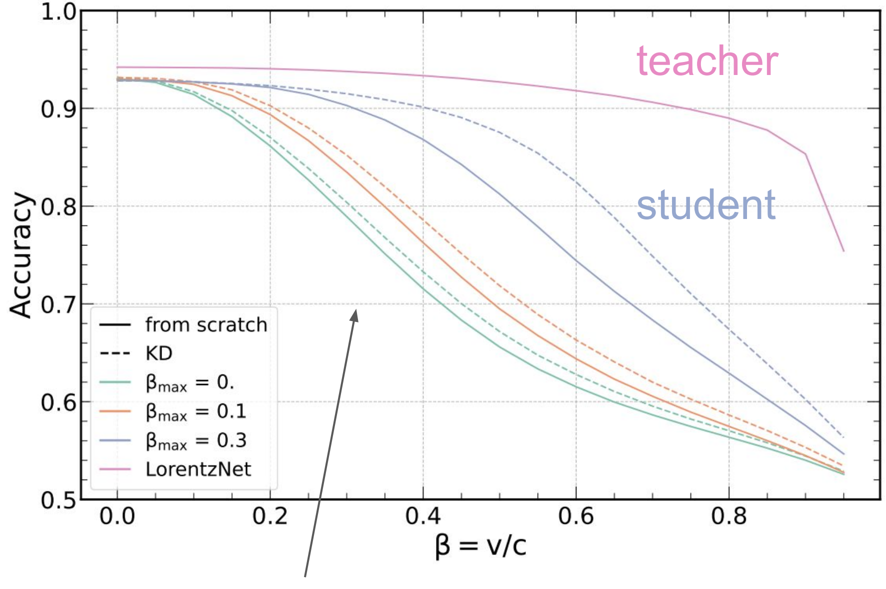
Knowledge distillation for pileup mitigation

Can increase jet & MET resolution at L1 trigger!

2311.12551



2311.14160



Knowledge distillation for jet tagging

Can even pass on inductive bias, e.g., invariance under Lorentz boosts

$$N = \sigma \cdot \mathscr{L} \cdot A \cdot \epsilon$$

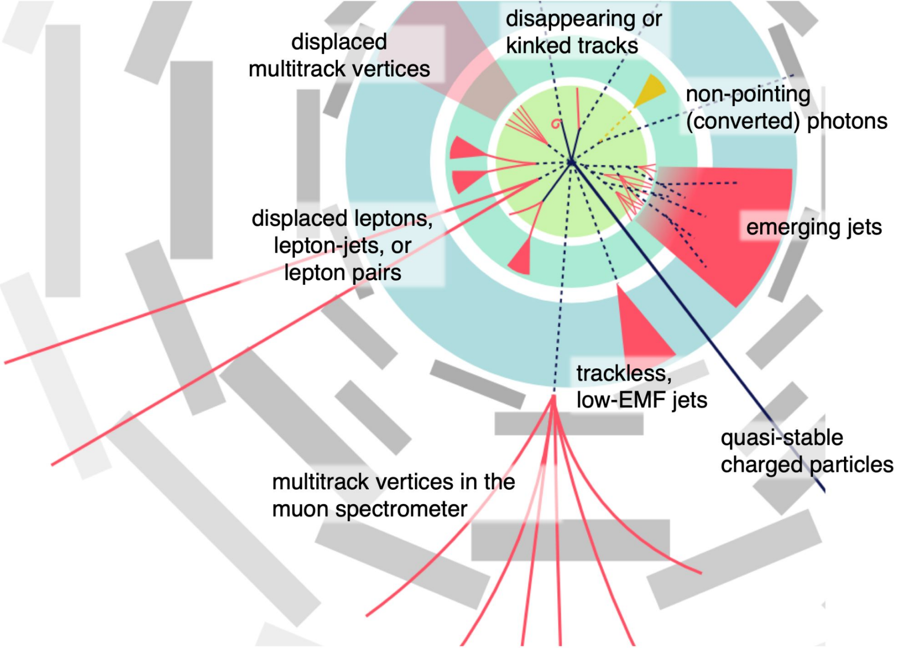
Energy frontier

Intensity frontier

Rate frontier

Lifetime frontier

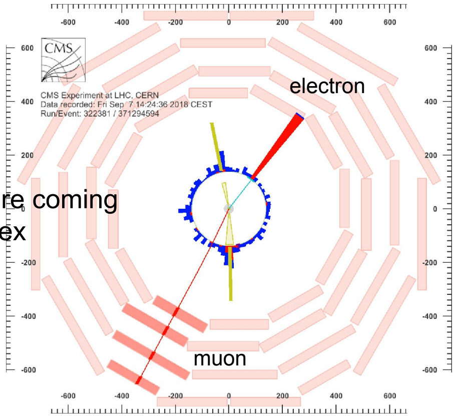
Why long-lived particles are tricky

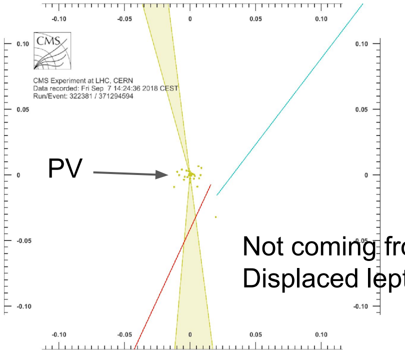


CMS was **not designed** to look for **displaced** new physics

Reconstruction algorithms, cylindrical geometry, trigger, all designed assuming particles emerge from the collision point

Looking like they're coming from primary vertex

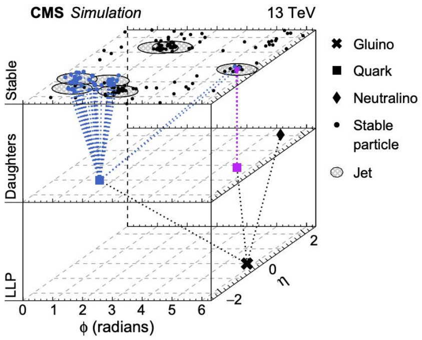


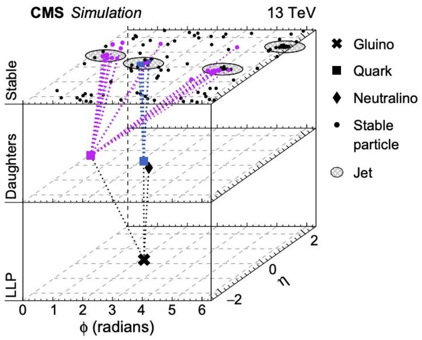


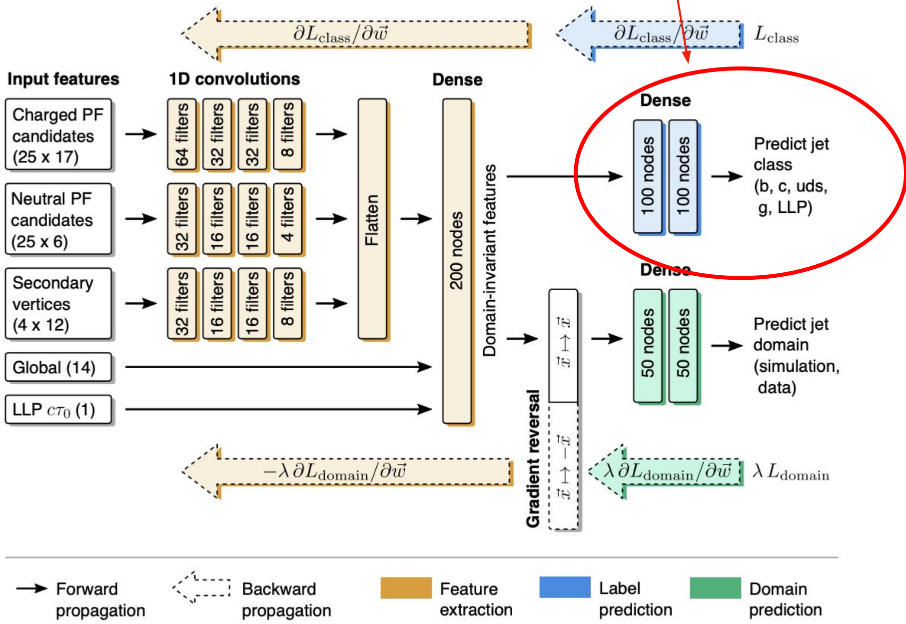
Not coming from PV!
Displaced leptons

The CMS Displaced Jet Tagger

multiclass output



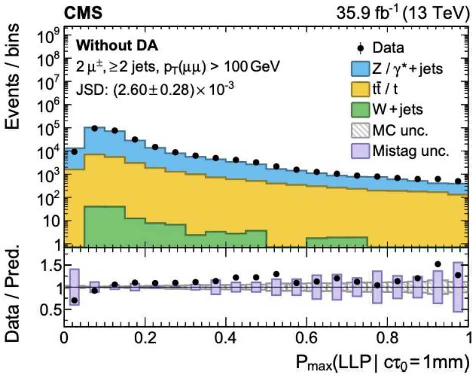


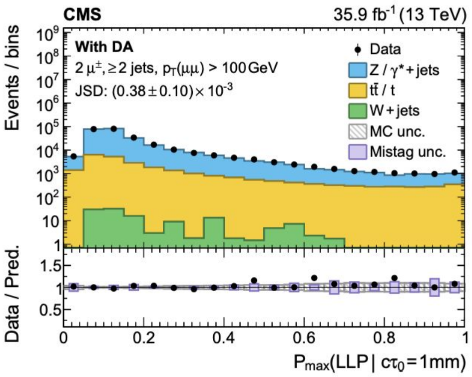


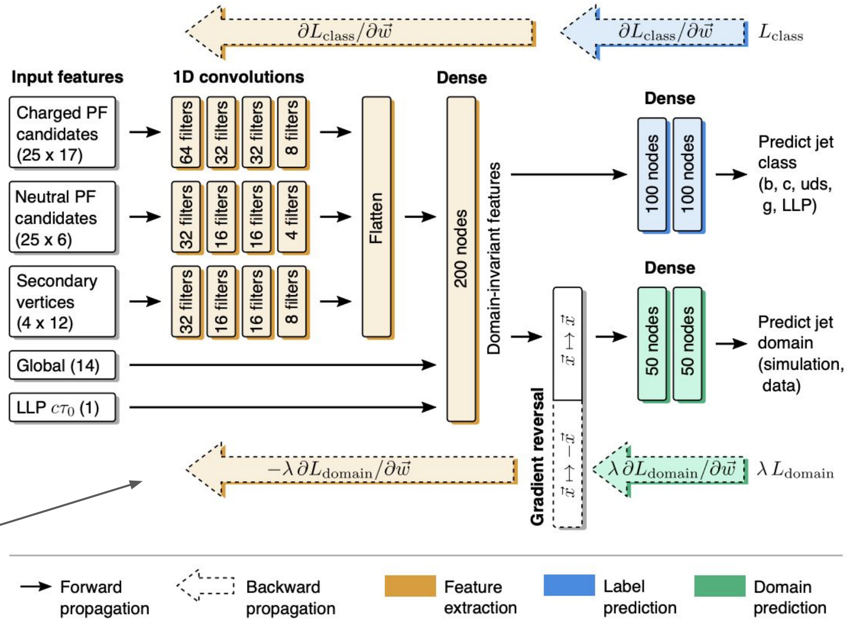
Long-lived signatures vary significantly with $c\tau_0$

→ Make it an input parameter
→ Enables testing over 6 orders of magn.

The CMS Displaced Jet Tagger

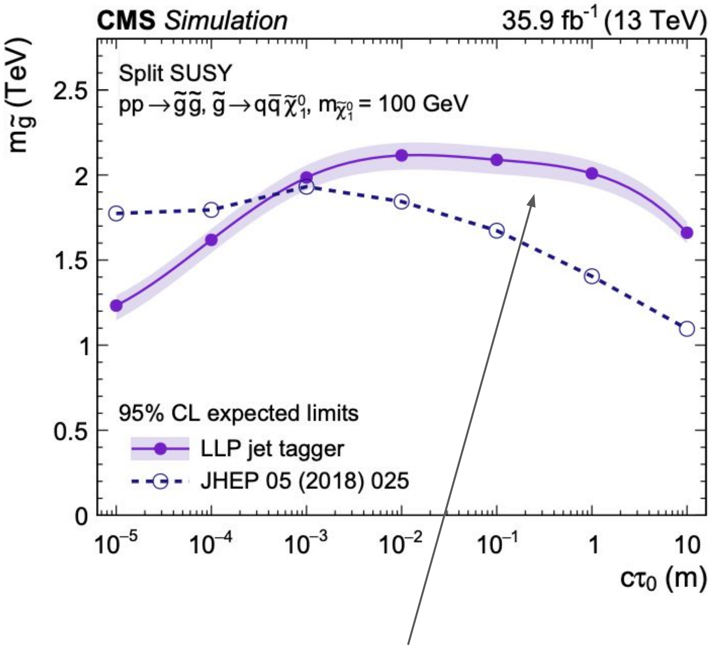


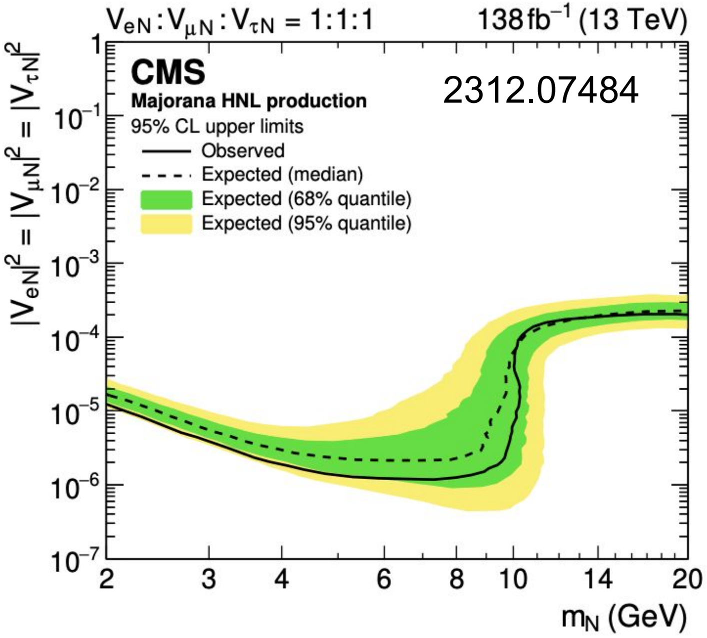




Domain adaptation through gradient reversal to improve MC modelling

New possibilities for searches with displaced signatures





Large improvements in sensitivity over cut-based approaches

Used in first-ever Heavy Neutral Lepton search with displaced jet signatures

Conclusion

HEP = World's best environment to study and employ machine learning

- unparalleled wealth and richness of data
- high-fidelity simulation let's us develop and understand AI algorithms
- By now well accepted in community

New physics will be rare (if within reach at all)
→ Need to squeeze out data as much as possible

Many great examples where AI improves, facilitates, or even enables novel searches!

Never forget: measure performance in data