Experimental particle physics & AI

Prof. Gregor Kasieczka
Email: gregor.kasieczka@uni-hamburg.de
Twitter/X: @GregorKasieczka
EuCAIFCon 2024, Amsterdam— 29.4.2024

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE
KISS
FSP CMS
PUNCH 4NFDI
DASHH
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung
UHH Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG
CDCS
CENTER FOR DATA AND COMPUTING IN NATURAL SCIENCES
PIER
Partnership of Universität Hamburg and DESY
Emmy Noether-Programm
Deutsche Forschungsgemeinschaft DFG

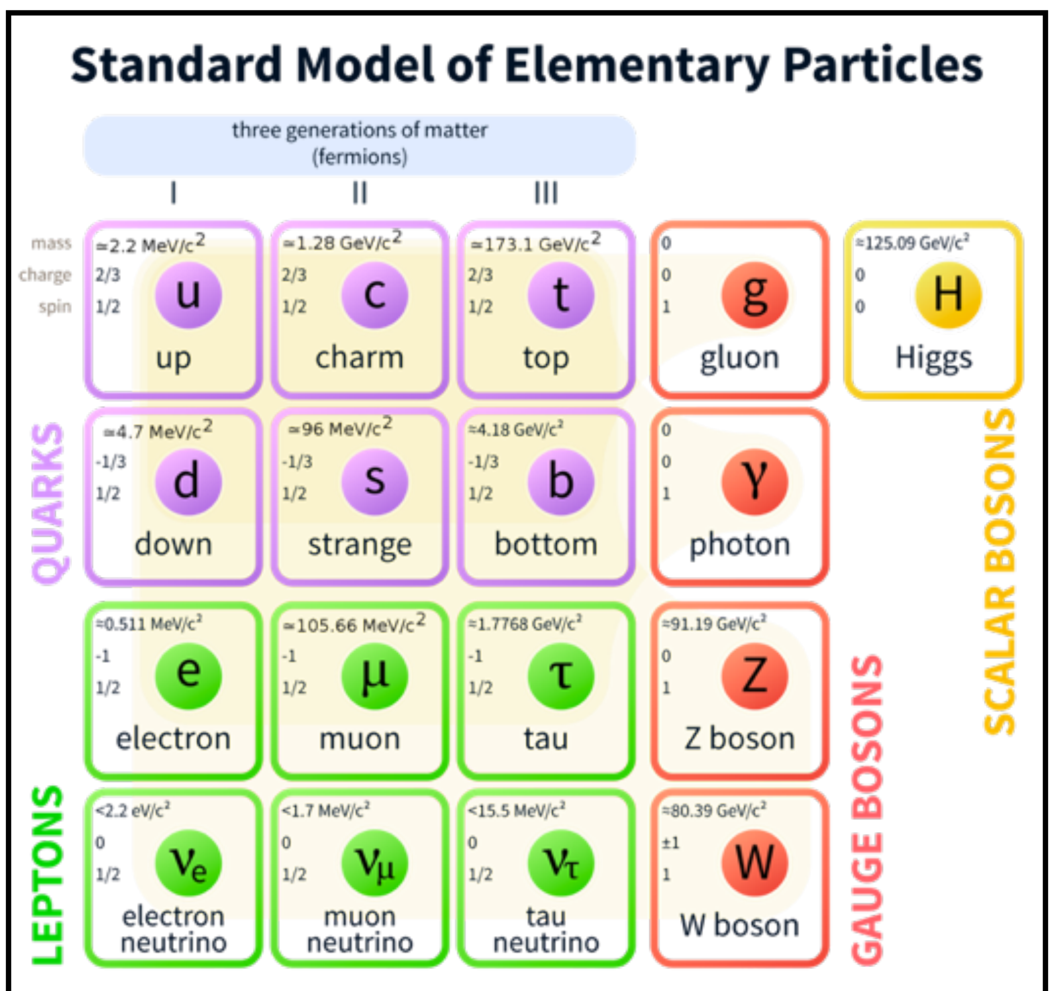
Experimental particle physics

Normal experience: 10^{-3} - 10^{3} m

Experimental particle physics

Normal experience: 10^{-3} - 10^{3} m

100000000000000x Zoom

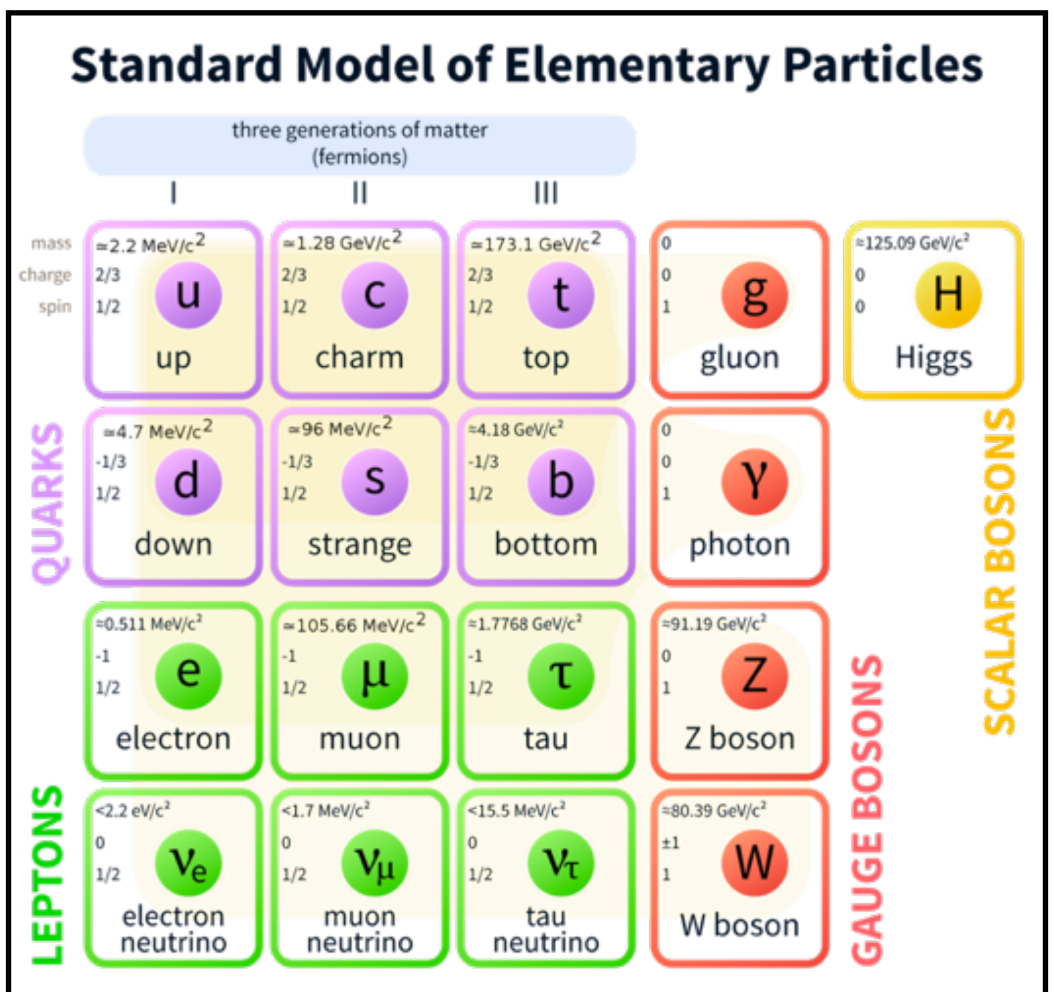


Goal:
Understand nature at the most fundamental level

Experimental particle physics

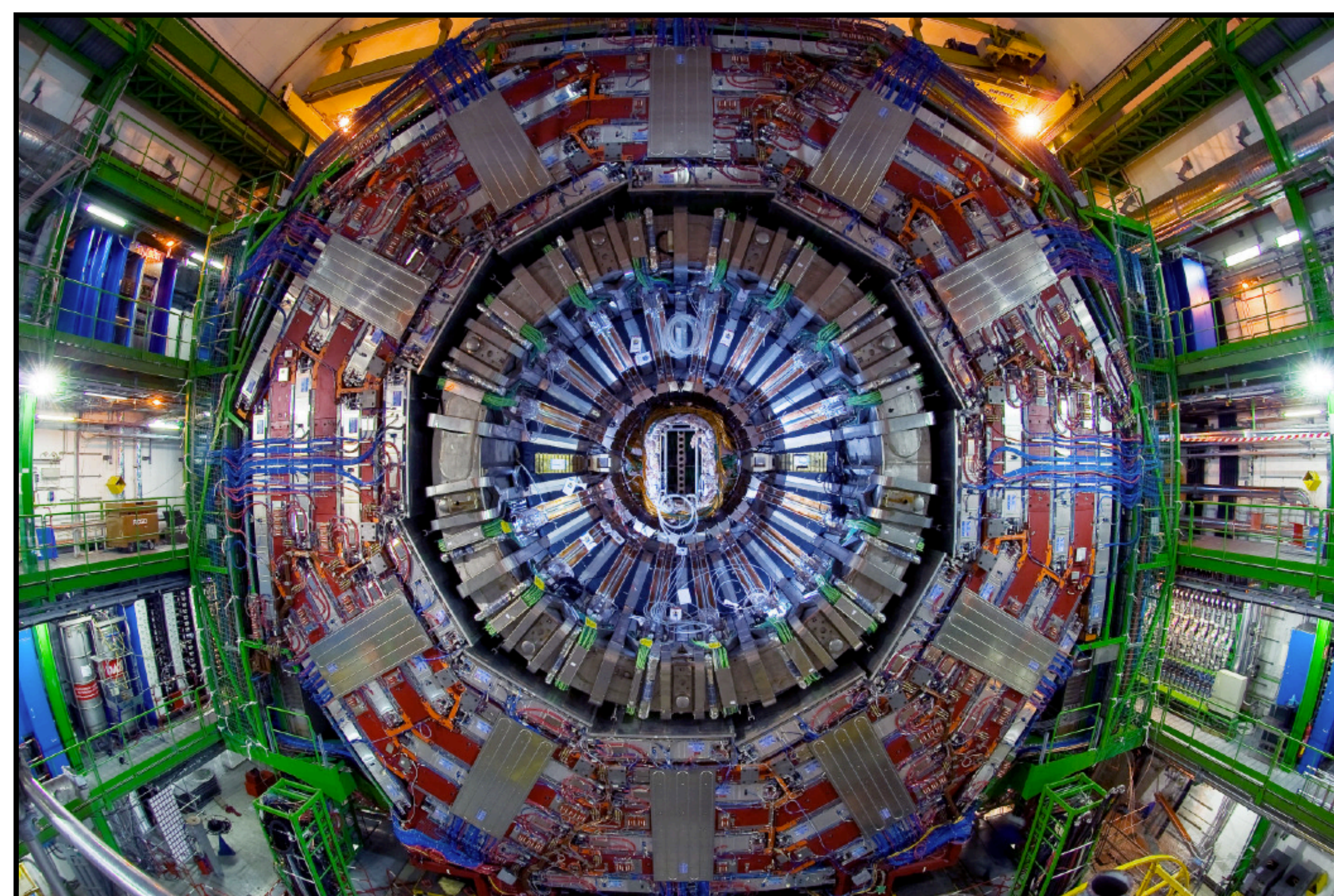
Normal experience: 10^{-3} - 10^{3} m

100000000000000x Zoom



Goal:
Understand nature at the most fundamental level

Using some of the largest and most precise scientific instruments ever built



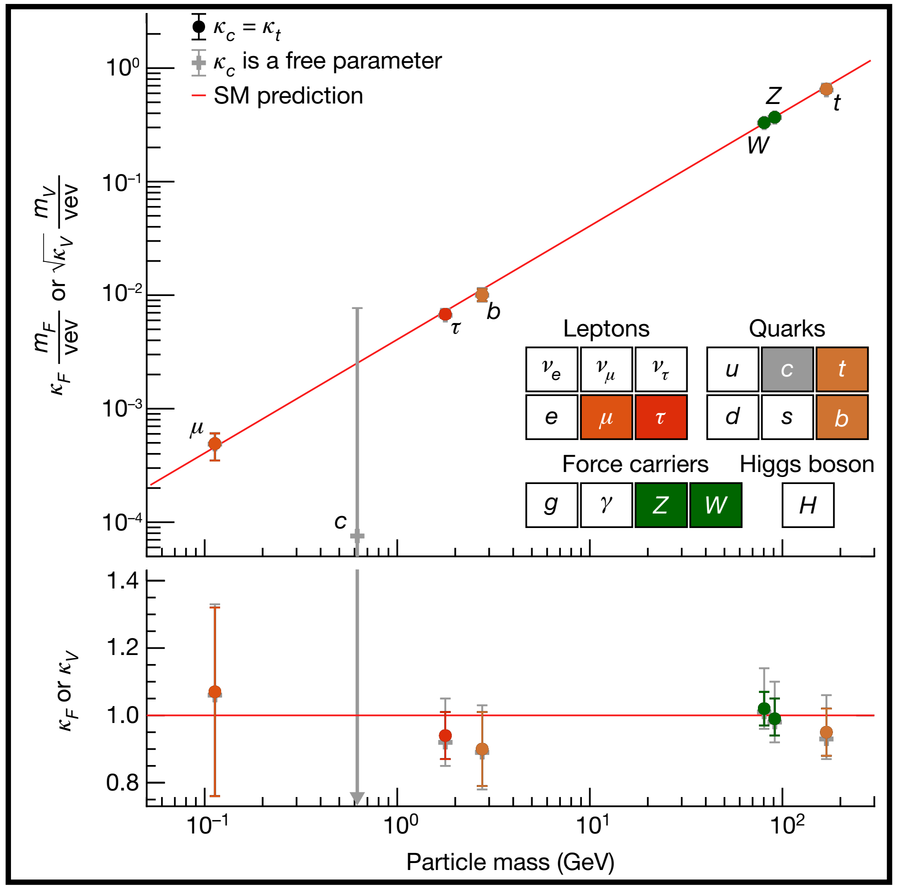
Experimental particle physics

Goal:
Understand nature at the most fundamental level

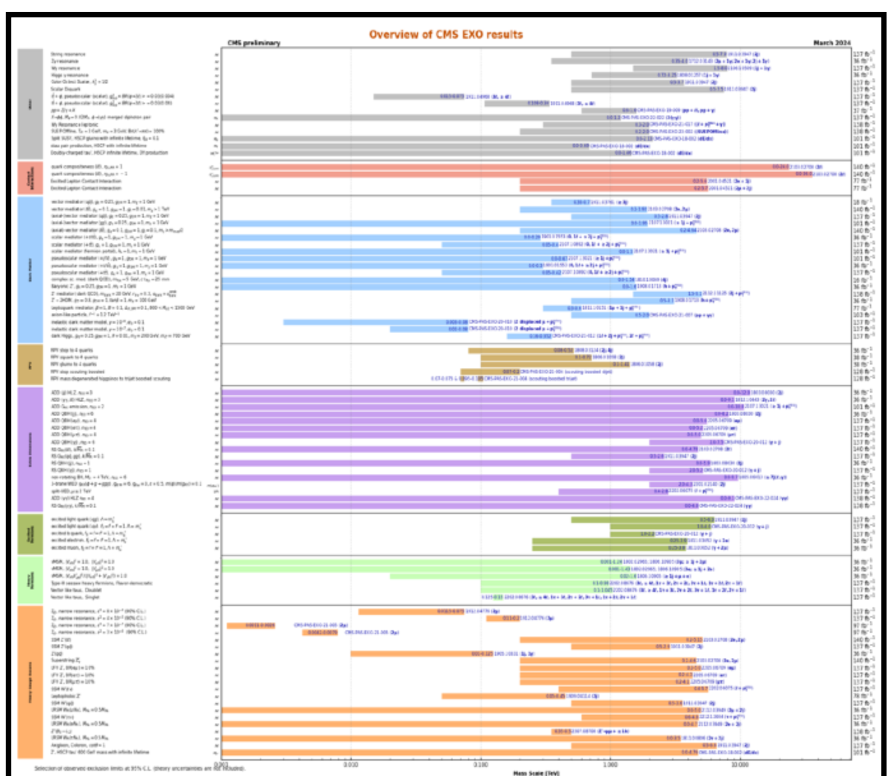
Using some of the largest and most precise scientific instruments ever built

Following the discovery of the Higgs boson in 2012

Experimental particle physics



Measuring the Standard Model (e.g. the Higgs sector)



And looking for physics beyond

Goal:
Understand nature at the most fundamental level

Using some of the largest and most precise scientific instruments ever built

Following the discovery of the Higgs boson in 2012

Much remains to do

Experimental particle physics

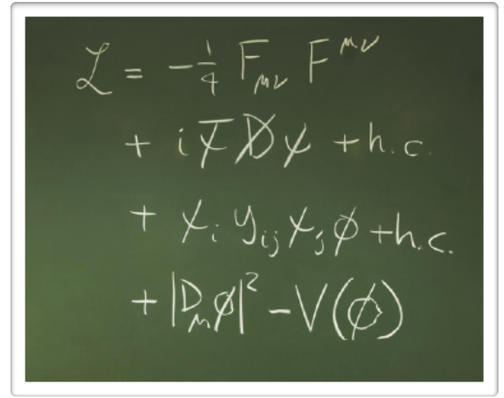
Goal:
Understand nature at the most fundamental level

Using some of the largest and most precise scientific instruments ever built

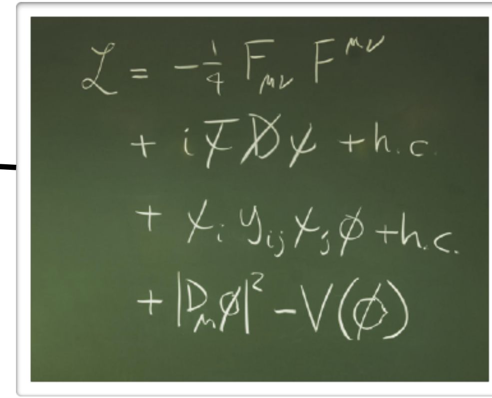
Following the discovery of the Higgs boson in 2012

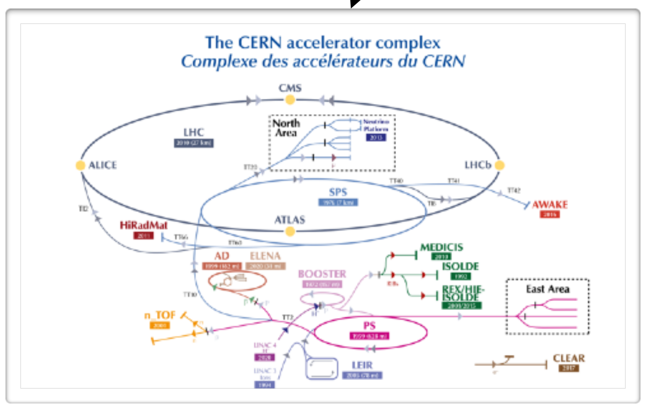
Much remains to do

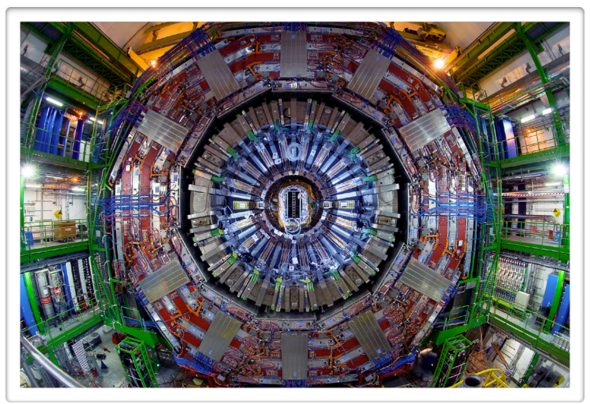
Where does AI enter?



First principle, quantum theoretical model

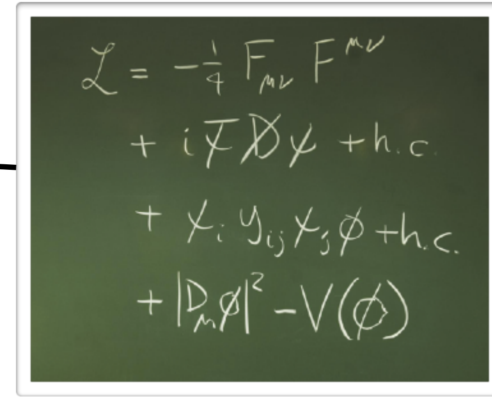


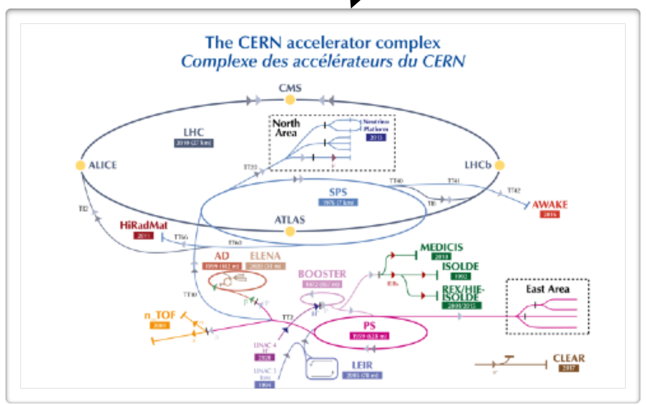


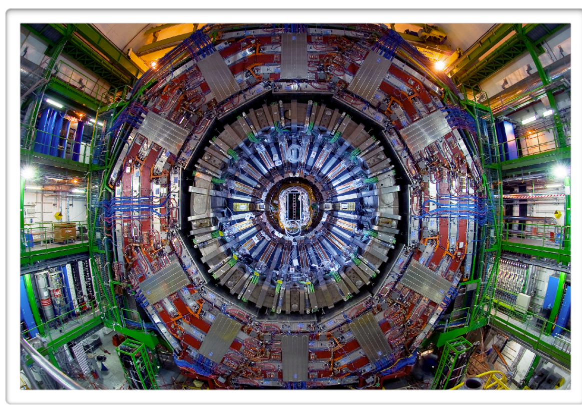


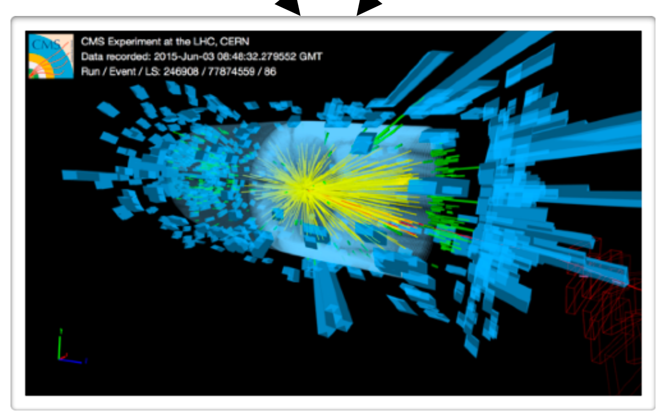
Colliders with
40 million events/second,
detectors with
100 million read-outs,

and massive theory-driven
simulation codes

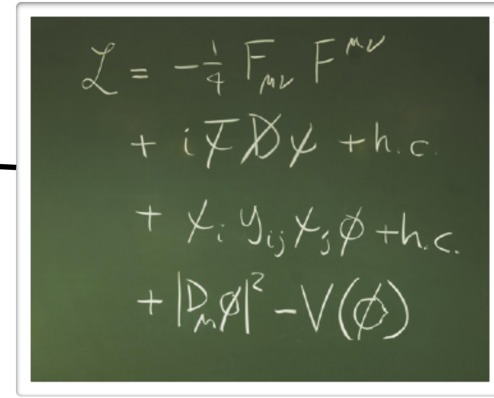


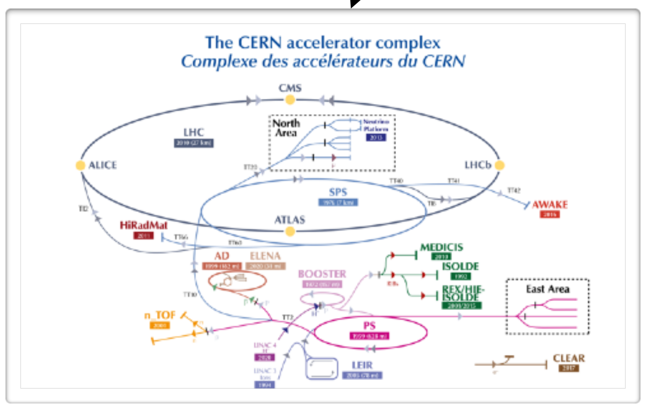


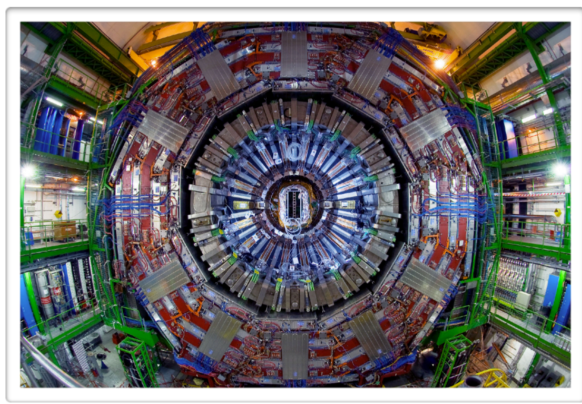


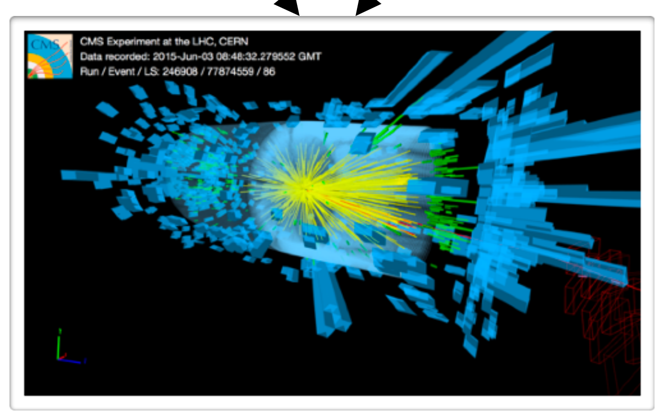


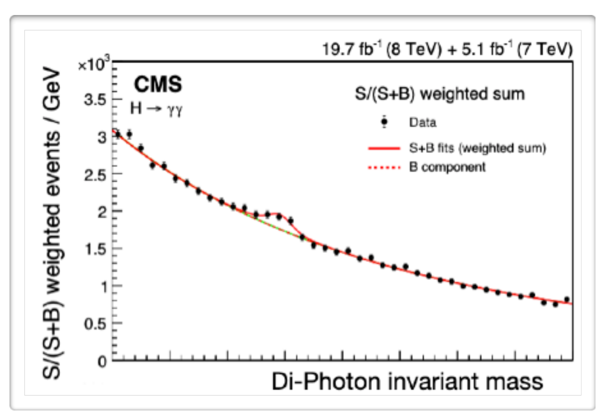
Complex reconstruction chain to turn low-level read-outs into high-level physics objects



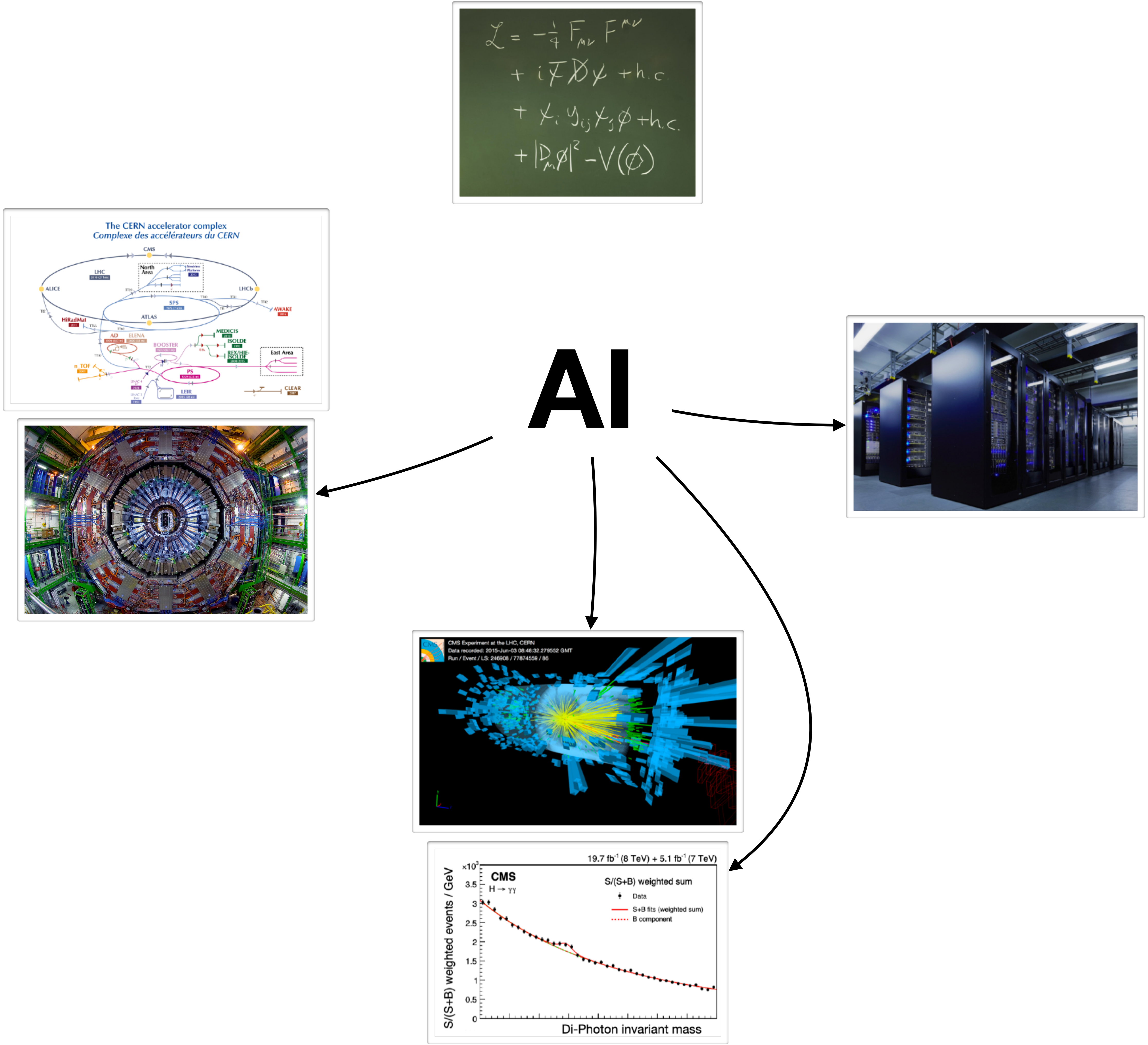


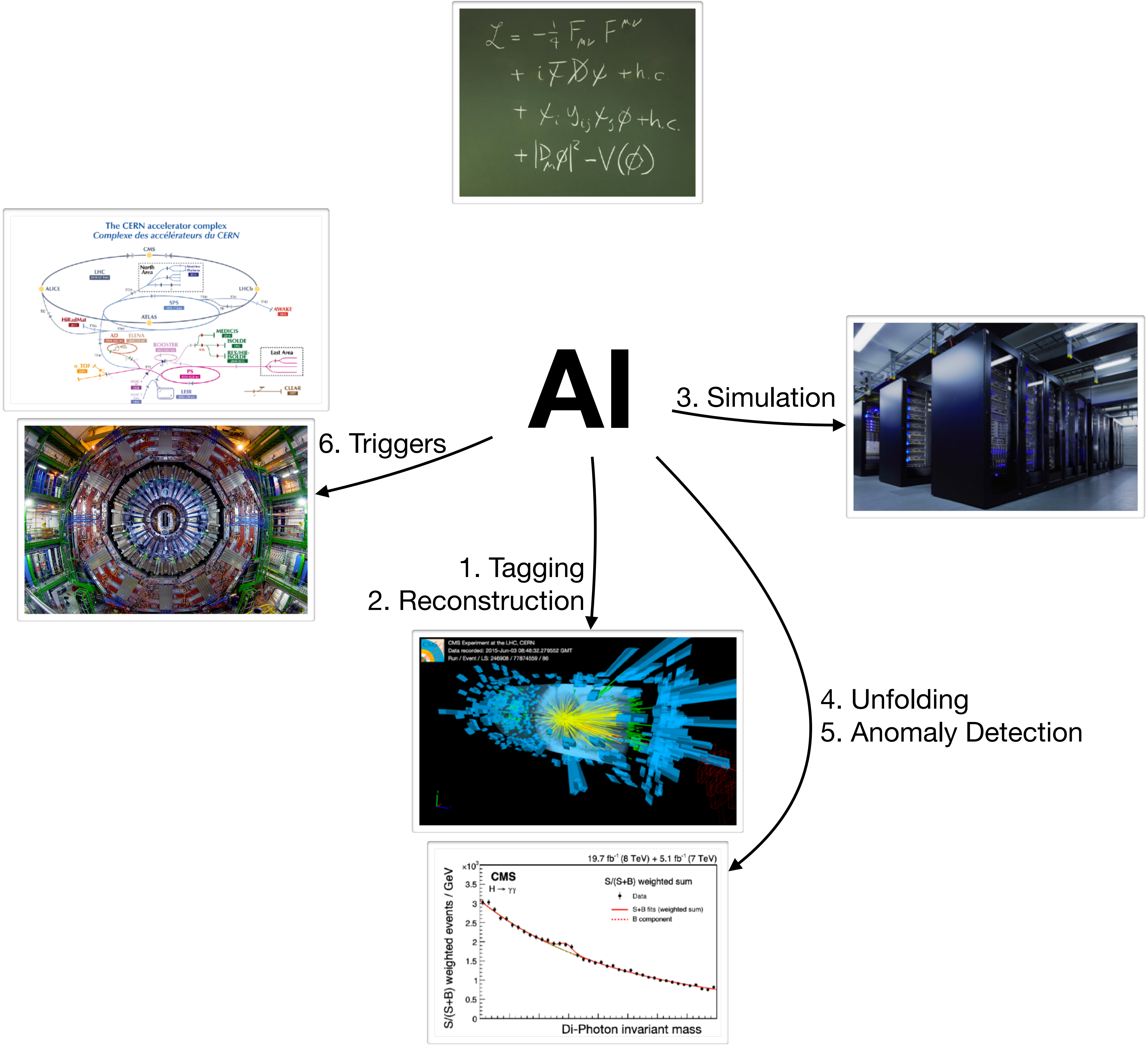




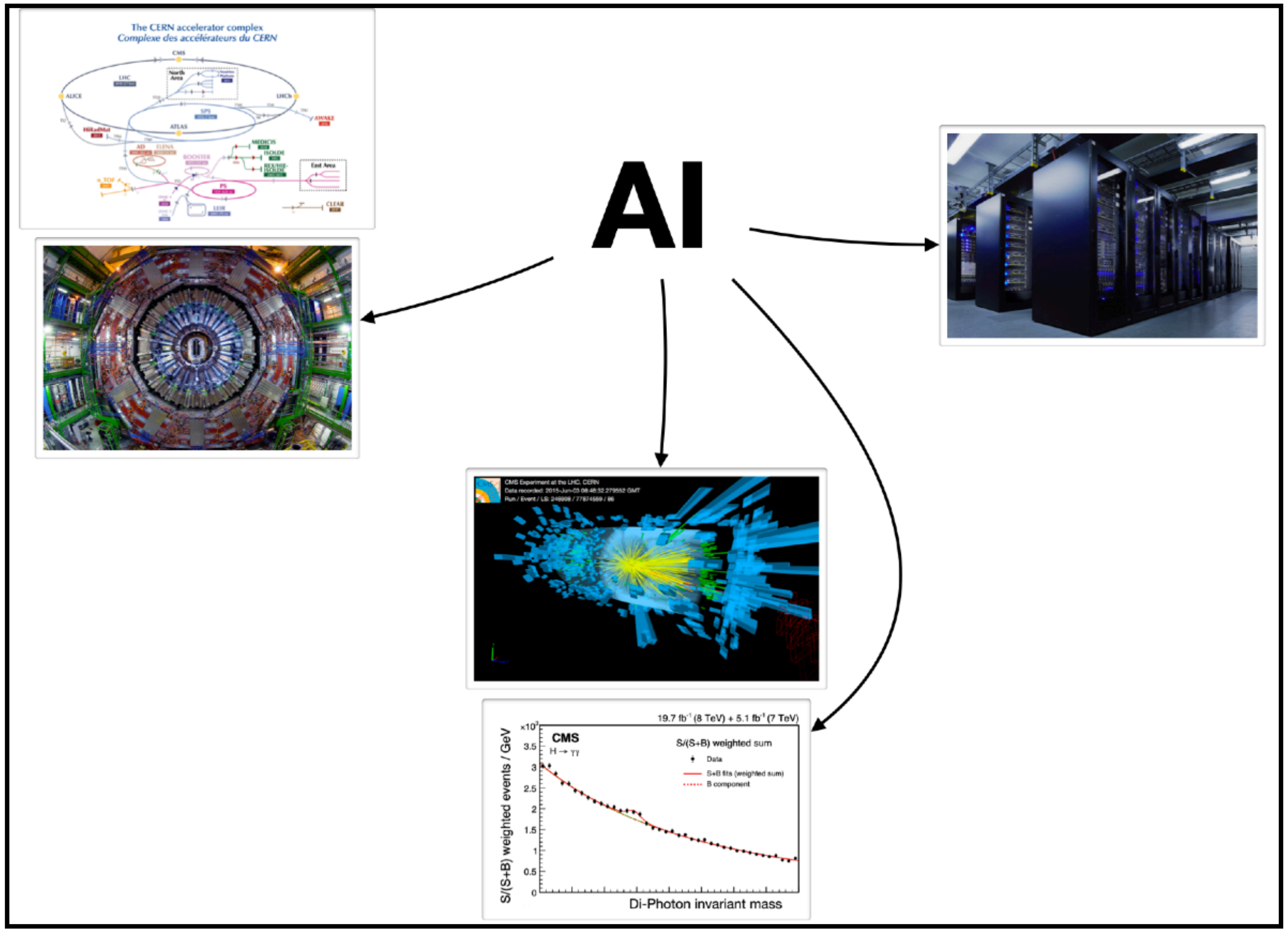


Sophisticated final statistical analysis


AI
$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i \bar{\psi} \not{D} \psi + h.c. + \psi_i y_{ij} \psi_j \phi + h.c. + |D_\mu \phi|^2 - V(\phi)$
The CERN accelerator complex
Complexe des accélérateurs du CERN
CMS
LHC
ALICE
LHCb
North Area
SPS
ATLAS
AWAKE
HiRadMat
AD
ELENA
MEDICIS
ISOLDE
BOOSTER
REX/HIE-ISOLDE
East Area
n_TOF
PS
LEIR
CLEAR
CMS Experiment at the LHC, CERN
Data recorded: 2015-Jun-03 08:48:32.279552 GMT
Run / Event / LS: 246908 / 77874559 / 86
19.7 fb^{-1} (8 TeV) + 5.1 fb^{-1} (7 TeV)
CMS
H → γγ
S/(S+B) weighted sum
Data
S+B fits (weighted sum)
B component
S/(S+B) weighted events / GeV
Di-Photon invariant mass


AI
3. Simulation
6. Triggers
1. Tagging
2. Reconstruction
4. Unfolding
5. Anomaly Detection
The CERN accelerator complex
Complexe des accélérateurs du CERN
CMS Experiment at the LHC, CERN
19.7 fb⁻¹ (8 TeV) + 5.1 fb⁻¹ (7 TeV)
CMS
H → γγ
S/(S+B) weighted sum
Data
S+B fits (weighted sum)
B component
S/(S+B) weighted events / GeV
Di-Photon invariant mass

Experimental particle physics



Goal:
Understand nature at the most fundamental level

Using some of the largest and most precise scientific instruments ever built

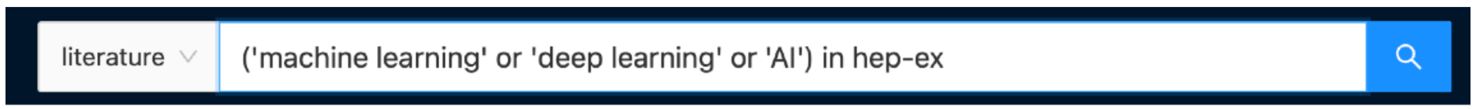
Following the discovery of the Higgs boson in 2012

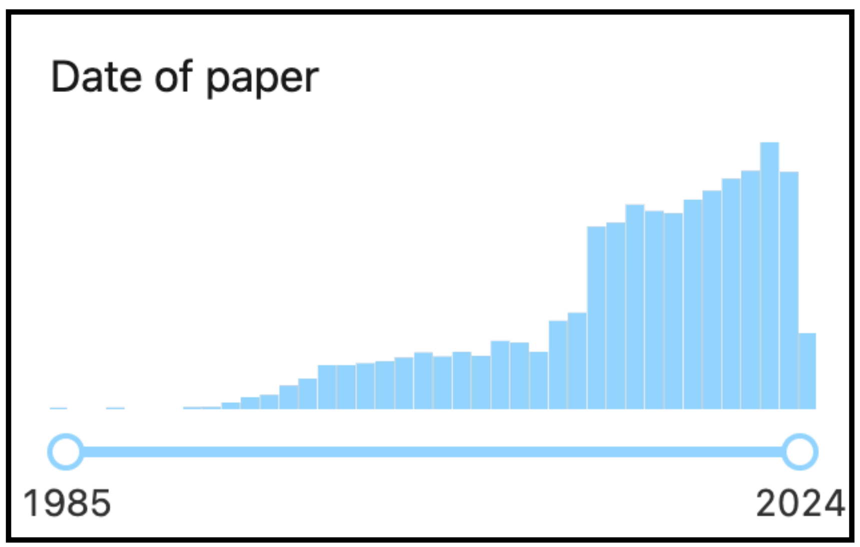
Much remains to do

Where does AI enter?

Everywhere.

Experimental particle physics





40k papers

Apologies - there will be some selection bias.

Goal:
Understand nature at the most fundamental level

Using some of the largest and most precise scientific instruments ever built

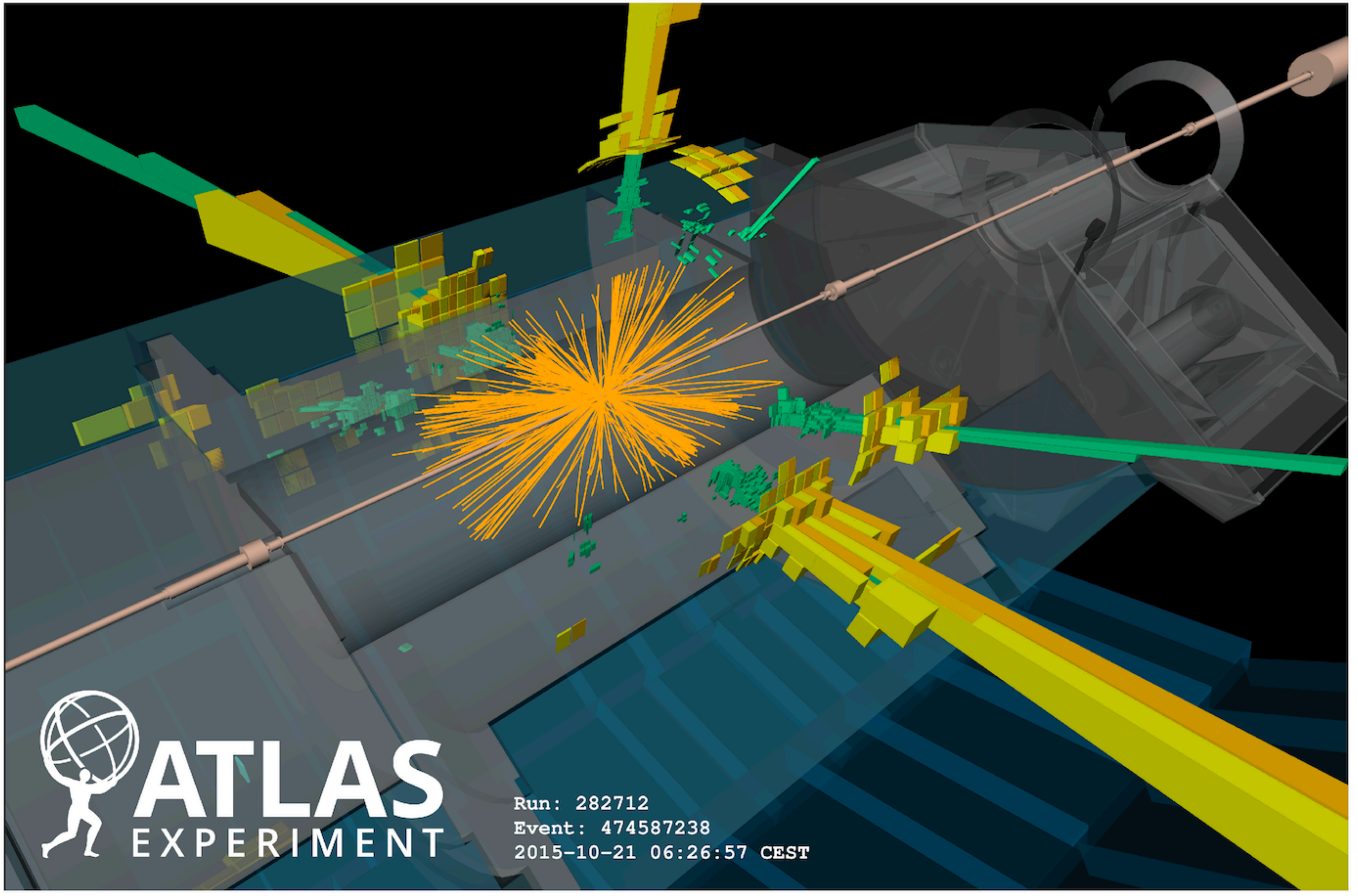
Following the discovery of the Higgs boson in 2012

Much remains to do

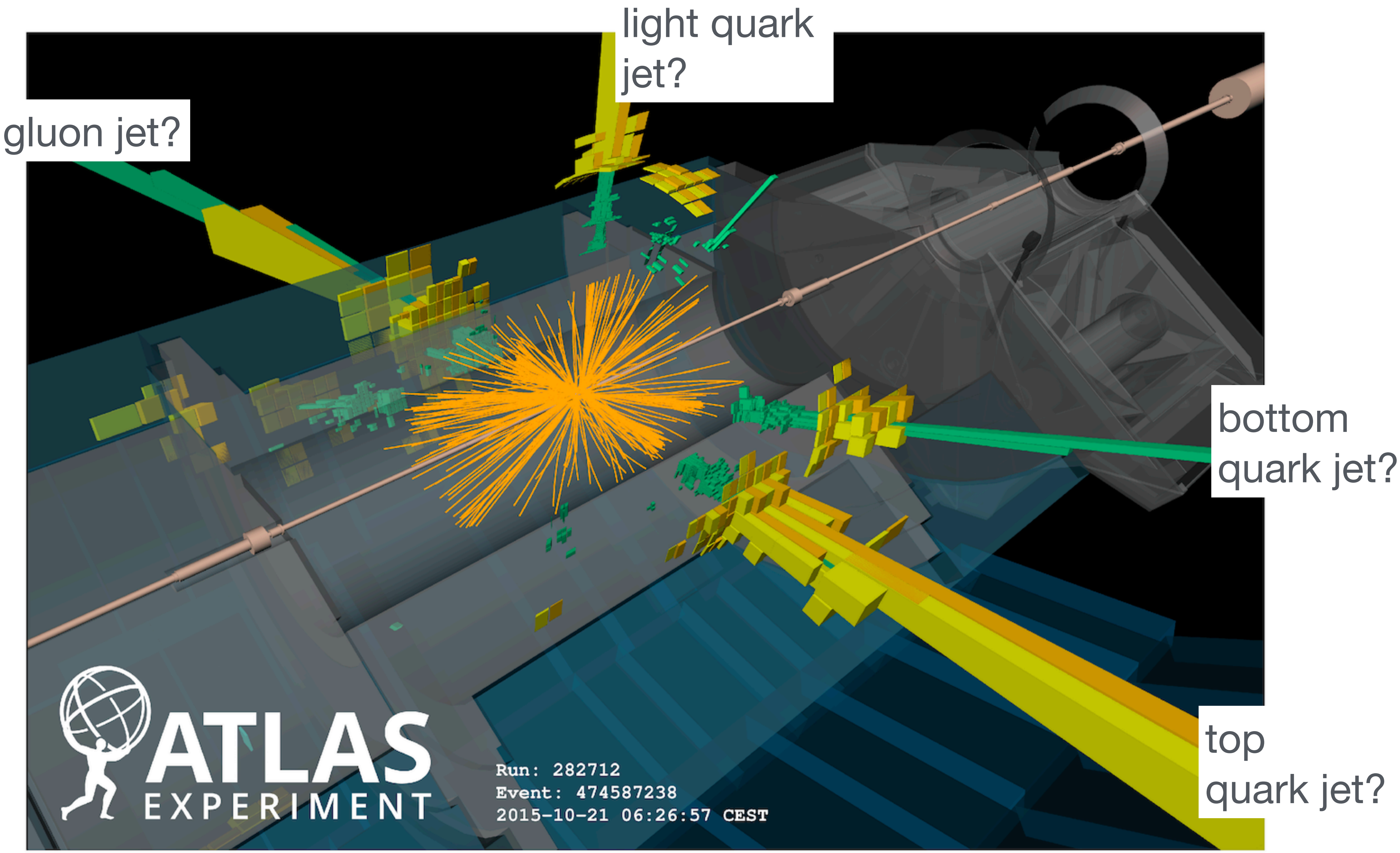
Where does AI enter?

Everywhere.

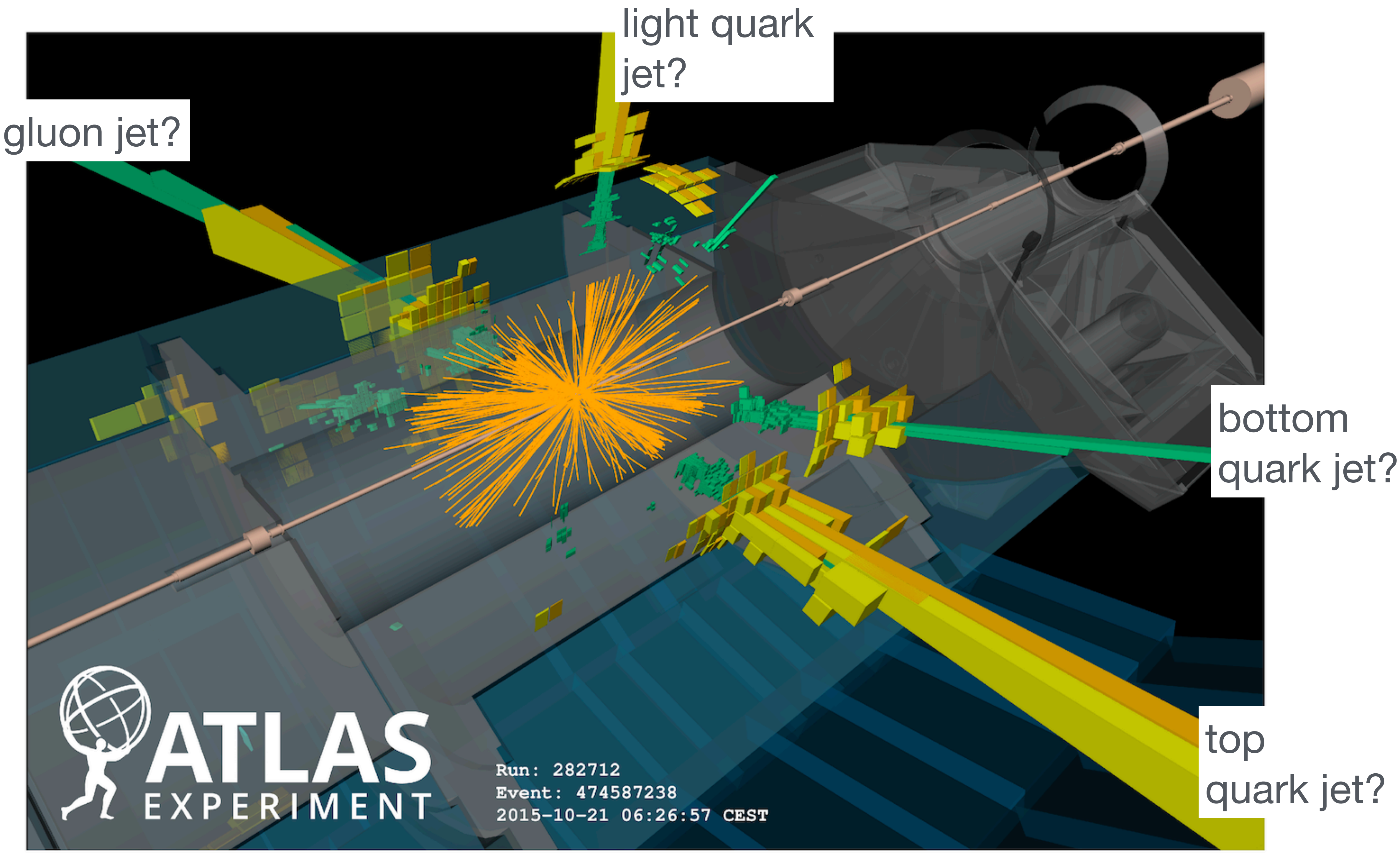
1. Taggers



A jet is a
collimated shower of particles in the detector

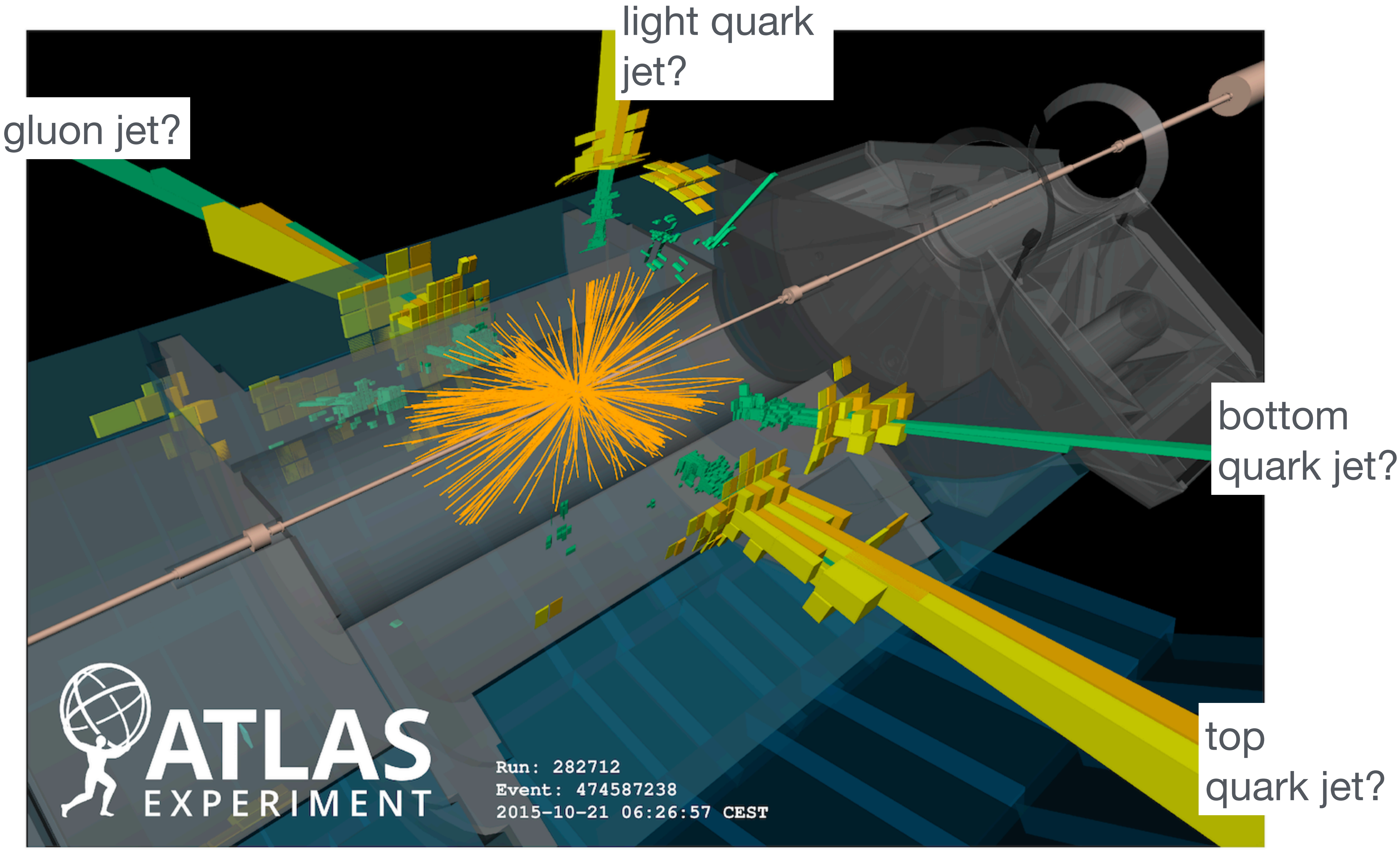


We want to know
which particle produced a jet

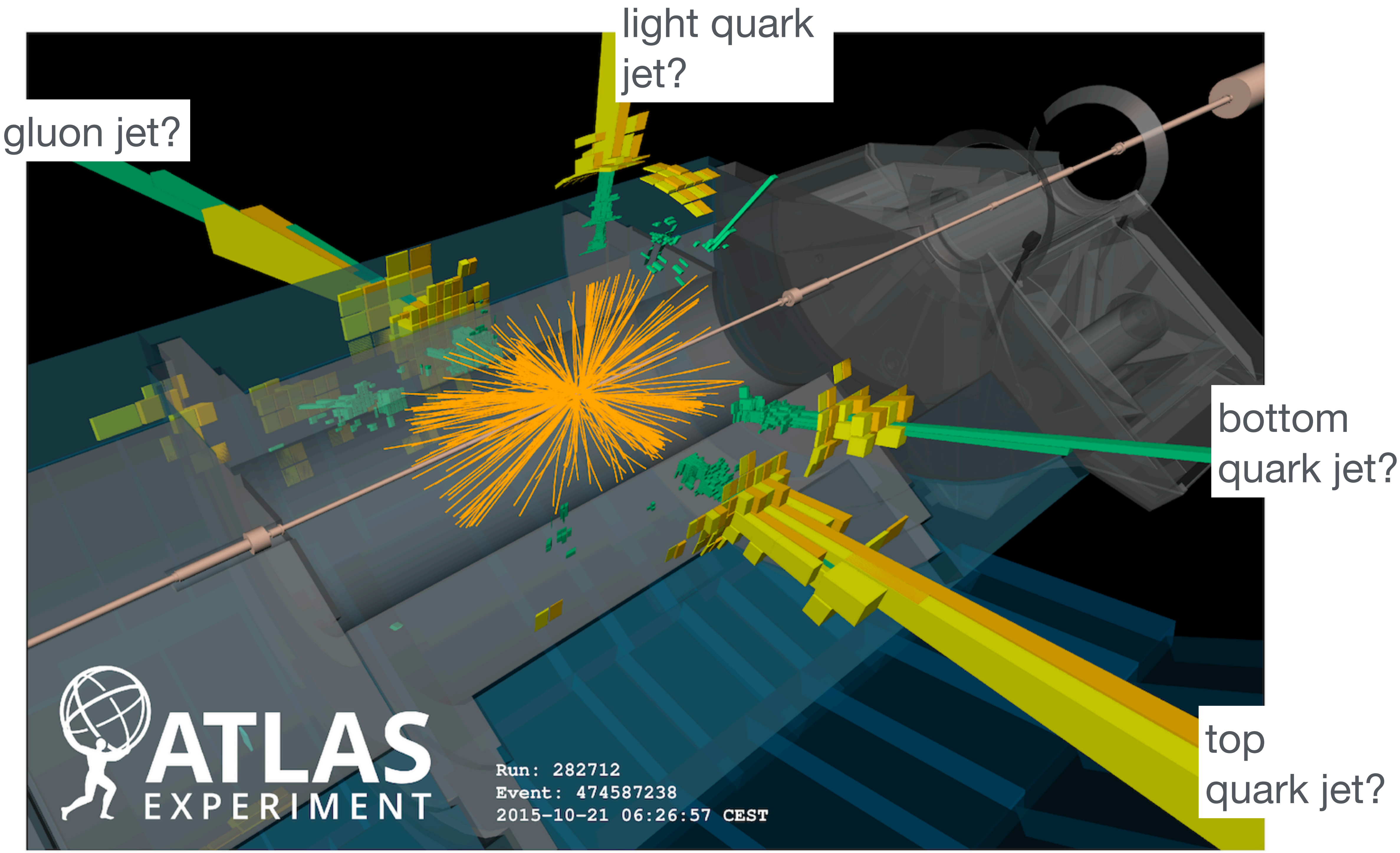


Why?

- Discover new particles
- Measure the Standard Model



Let's focus on top quarks
(Modern taggers are multi-class)



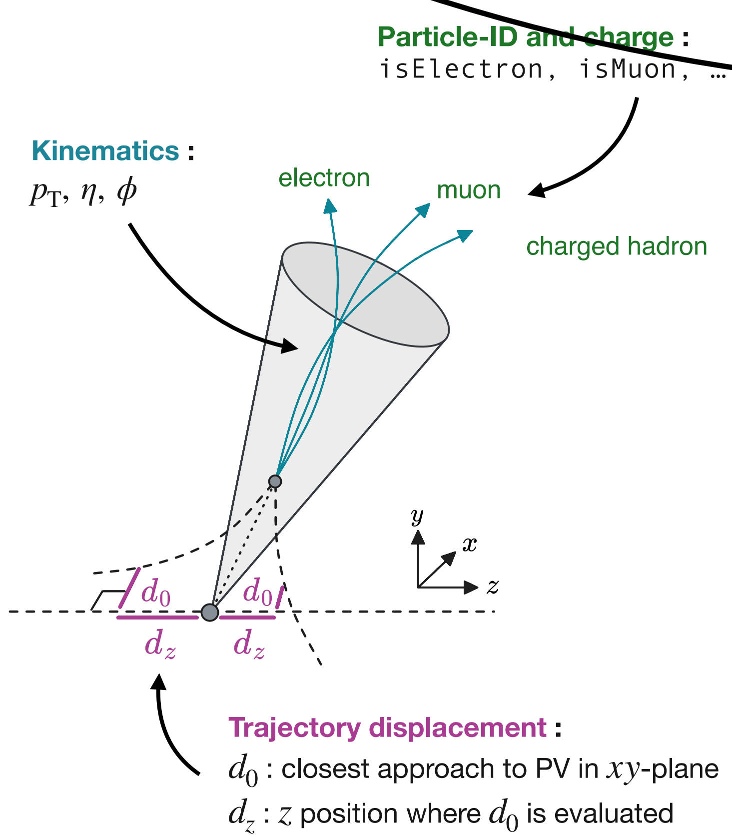
How to build ML algorithms for complex, heterogenous data?

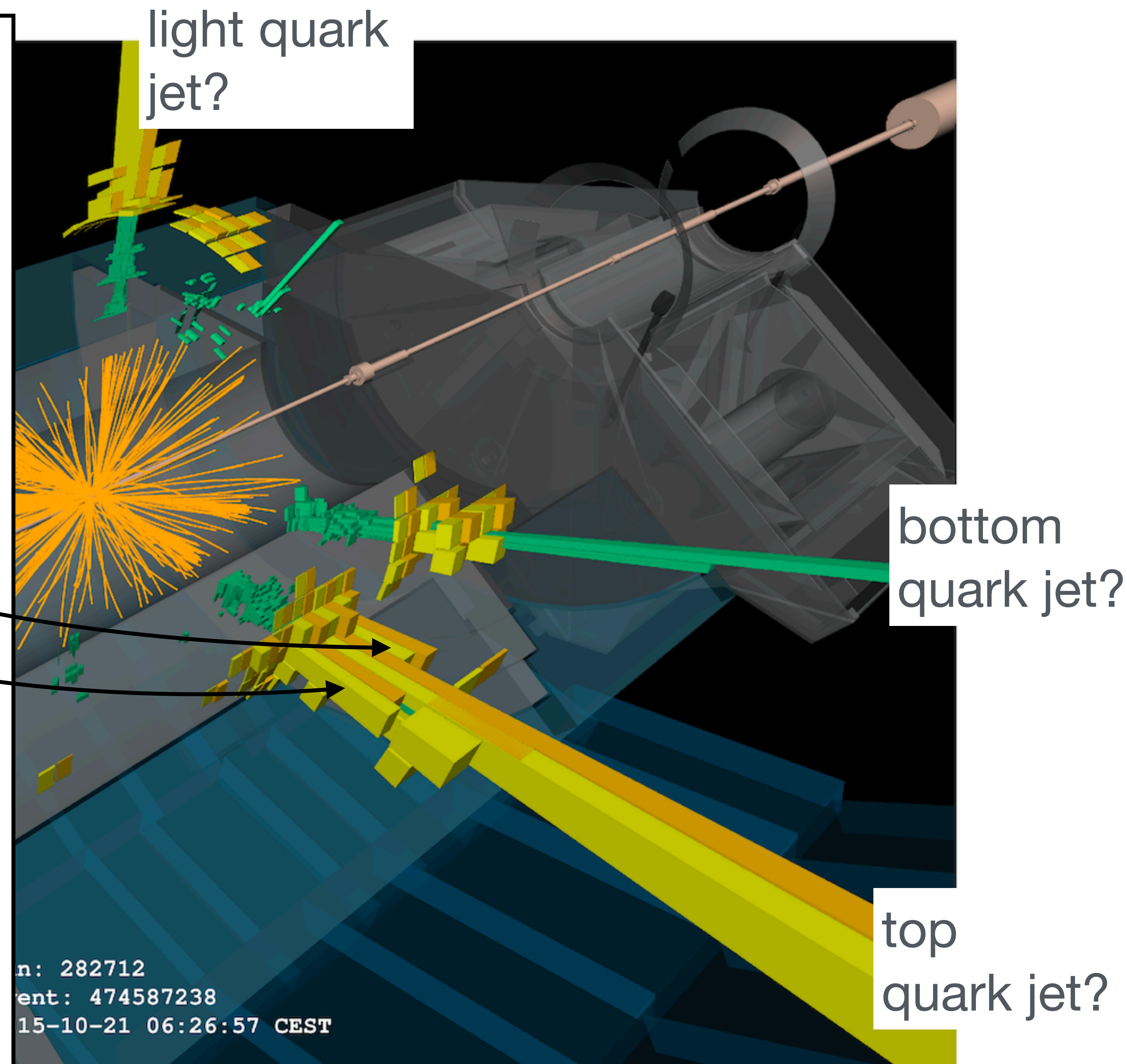
Data most naturally viewed as point cloud:

Each input (e.g. jet, event, ..) is a set of k-dimensional vectors (individual particles, hits, ..)

$$J_i = \{\vec{p_1}, \ldots, \vec{p_n}\}$$

Example per-particle features





Komiske, Metodiev, Thaler 1810.05165; Birk, **GK**, et al 2312.00123;

Landscape Dataset

- Open dataset for the development of better tagging algorithms for particle physics

- 2 million simulated examples
- Perfect class labels:
  top jet or light quark/gluon jet

- Input: momentum sorted list of 200 particles/jet with 3 features/particle (p_X, p_Y, p_Z)

SciPost Physics Submission

The Machine Learning Landscape of Top Taggers

G. Kasieczka (ed)[1], T. Plehn (ed)[2], A. Butter[2], K. Cranmer[3], D. Debnath[4], B. M. Dillon[5], M. Fairbairn[6], D. A. Faroughy[5], W. Fedorko[7], C. Gay[7], L. Gouskos[8], J. F. Kamenik[5,9], P. T. Komiske[10], S. Leiss[1], A. Lister[7], S. Macaluso[3,4], E. M. Metodiev[10], L. Moore[11], B. Nachman,[12,13], K. Nordström[14,15], J. Pearkes[7], H. Qu[8], Y. Rath[16], M. Rieger[16], D. Shih[4], J. M. Thompson[2], and S. Varma[6]

1 Institut für Experimentalphysik, Universität Hamburg, Germany
2 Institut für Theoretische Physik, Universität Heidelberg, Germany
3 Center for Cosmology and Particle Physics and Center for Data Science, NYU, USA
4 NHECT, Dept. of Physics and Astronomy, Rutgers, The State University of NJ, USA
5 Jozef Stefan Institute, Ljubljana, Slovenia
6 Theoretical Particle Physics and Cosmology, King's College London, United Kingdom
7 Department of Physics and Astronomy, The University of British Columbia, Canada
8 Department of Physics, University of California, Santa Barbara, USA
9 Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia
10 Center for Theoretical Physics, MIT, Cambridge, USA
11 CP3, Universitéxx Catholique de Louvain, Louvain-la-Neuve, Belgium
12 Physics Division, Lawrence Berkeley National Laboratory, Berkeley, USA
13 Simons Inst. for the Theory of Computing, University of California, Berkeley, USA
14 National Institute for Subatomic Physics (NIKHEF), Amsterdam, Netherlands
15 LPTHE, CNRS & Sorbonne Université, Paris, France
16 III. Physics Institute A, RWTH Aachen University, Germany

gregor.kasieczka@uni-hamburg.de
plehn@uni-heidelberg.de

July 24, 2019

arXiv:1902.09914v3 [hep-ph] 23 Jul 2019

Abstract

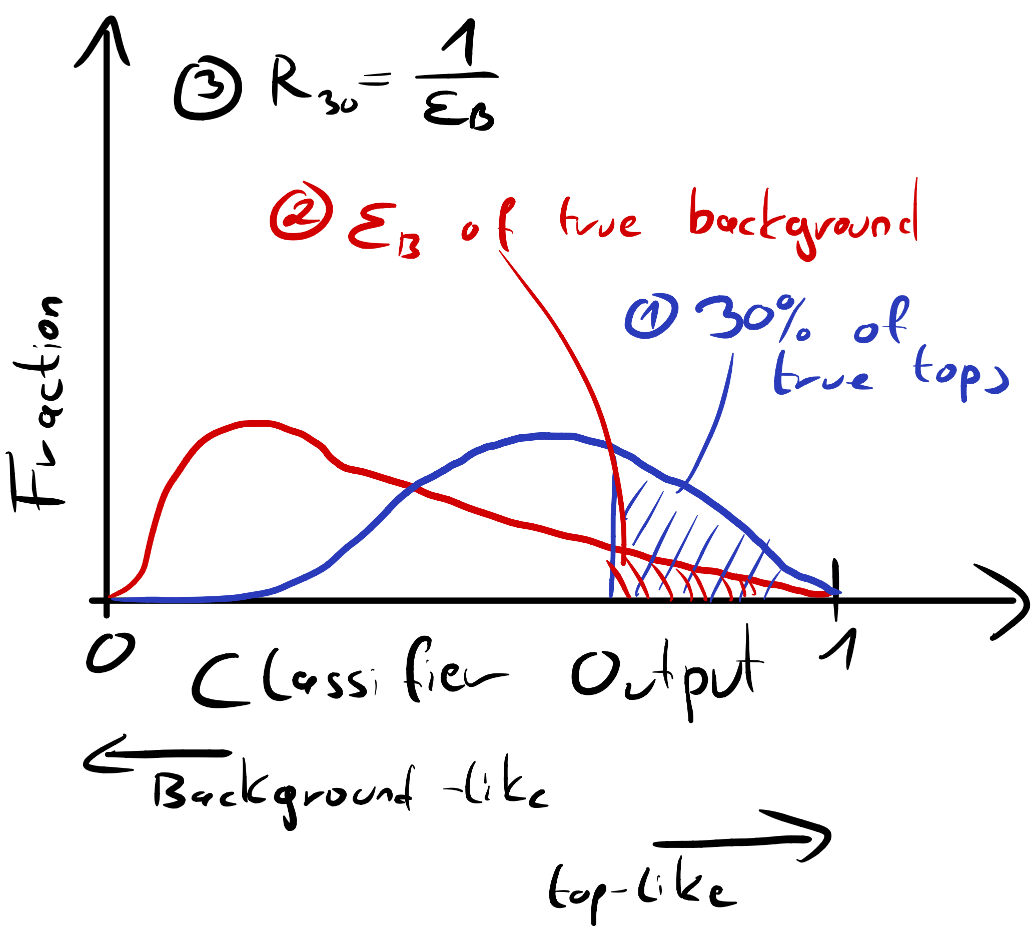
Based on the established task of identifying boosted, hadronically decaying top quarks, we compare a wide range of modern machine learning approaches. Unlike most established methods they rely on low-level input, for instance calorimeter output. While their network architectures are vastly different, their performance is comparatively similar. In general, we find that these new approaches are extremely powerful and great fun.

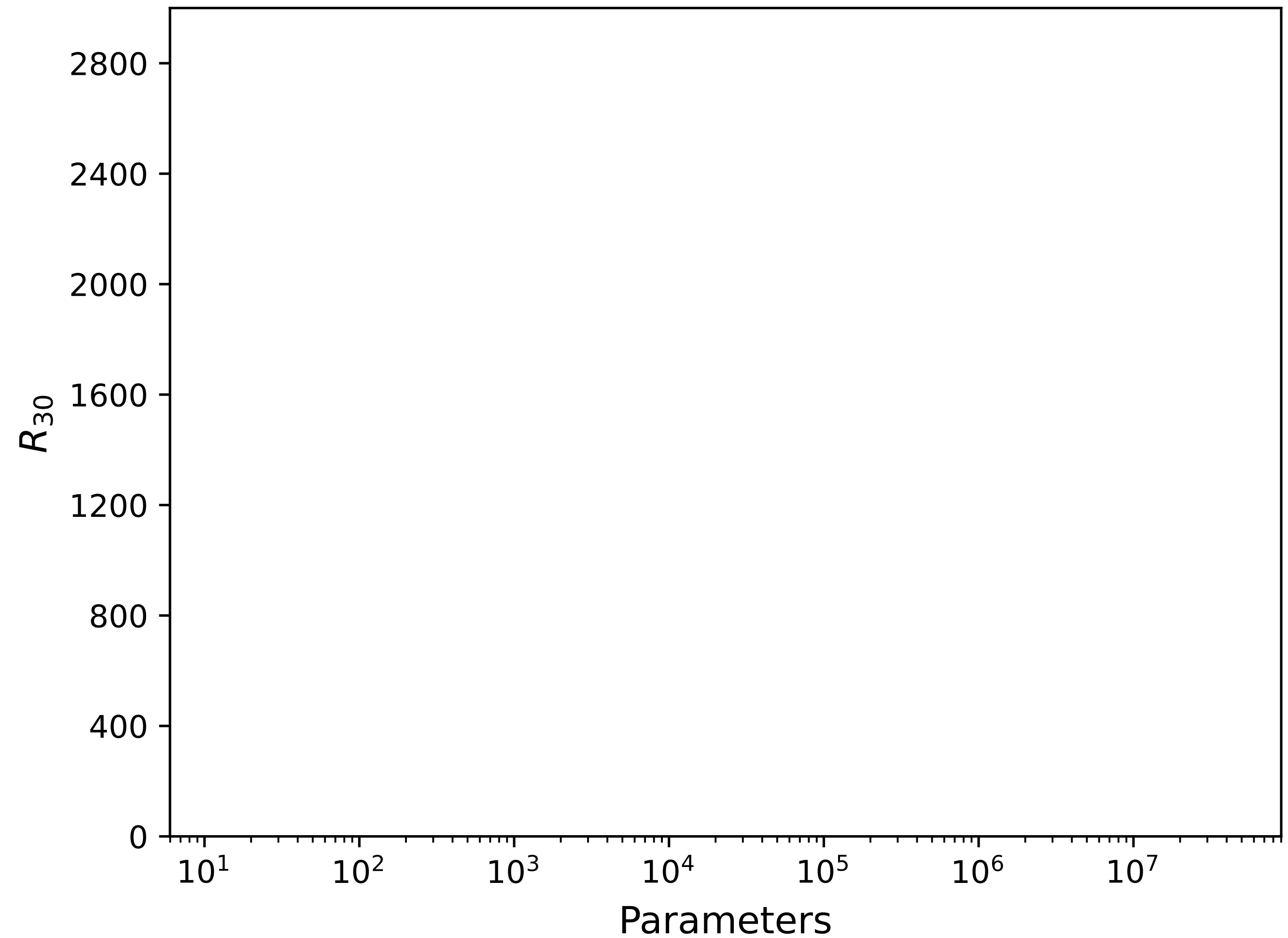
GK, Plehn, et al 1902.09914

Evaluation Metrics

Cut so that 30%
of top quarks pass selection:

R_{30} is the inverse of the number of background jets that also pass





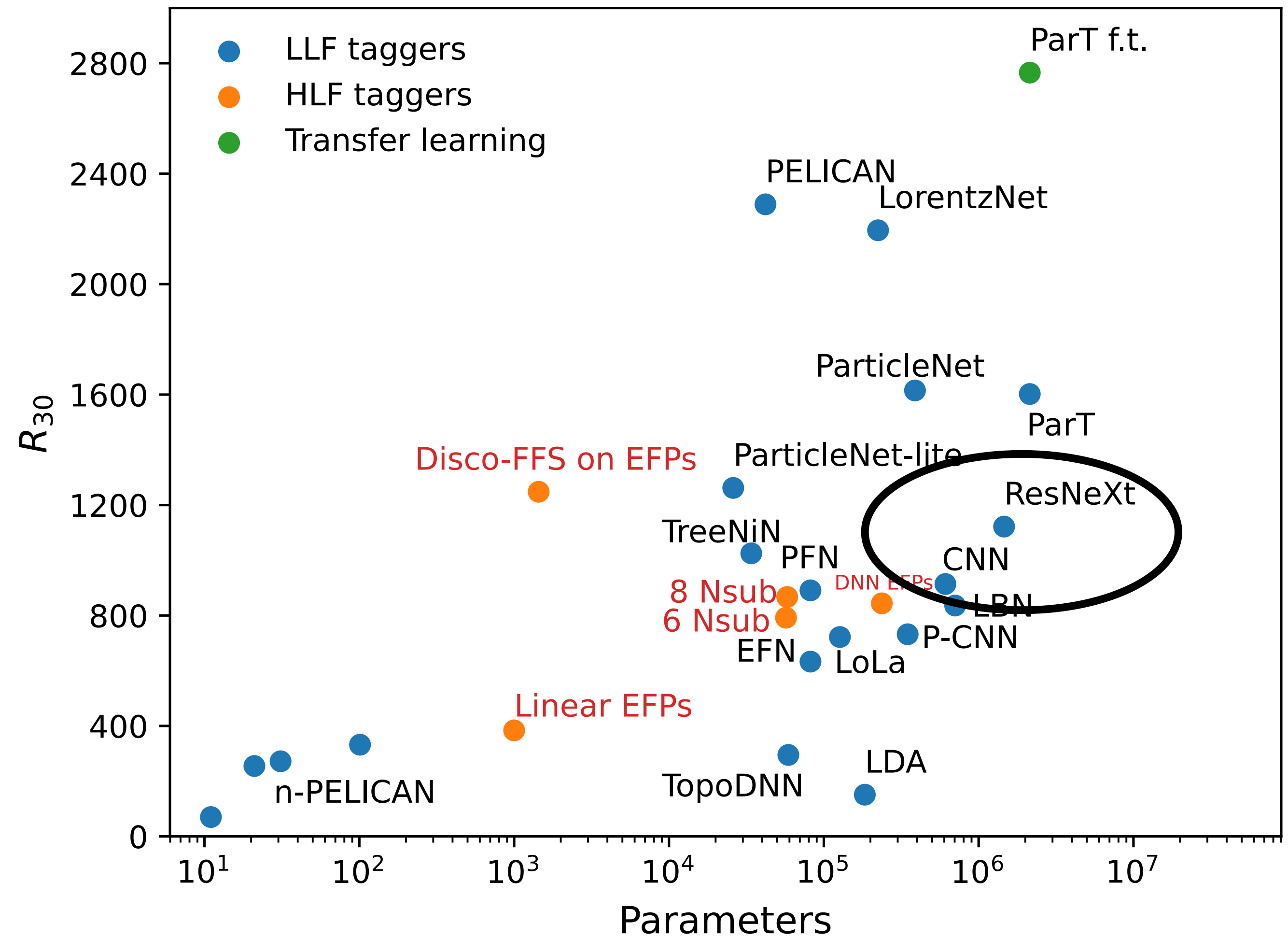
Number of trainable parameters:
Complexity / compute cost

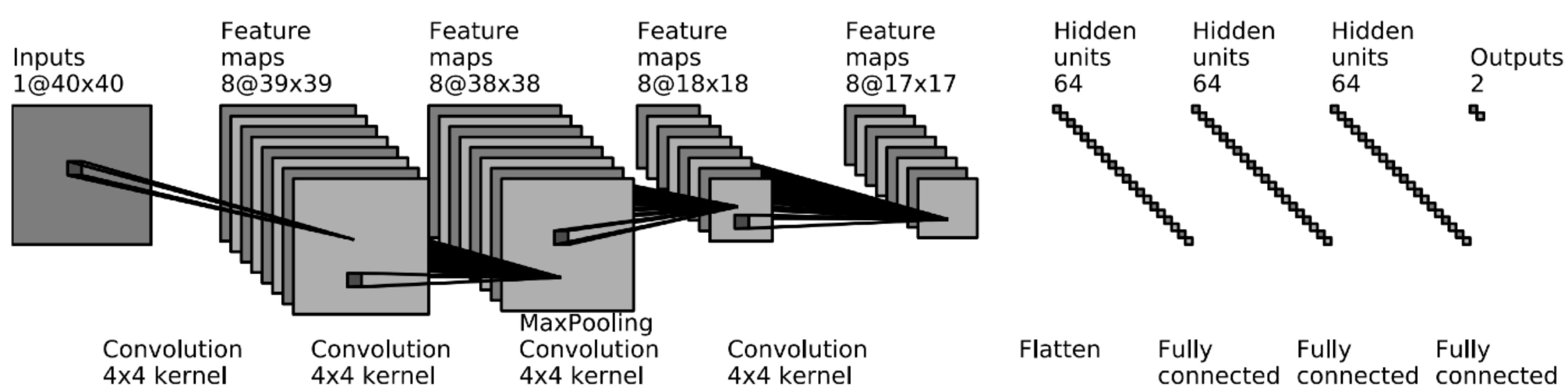
→How much performance can we afford?

Jets as Images

First deep learning approach:
Convolutional networks:
(project point-cloud onto 2D plane)

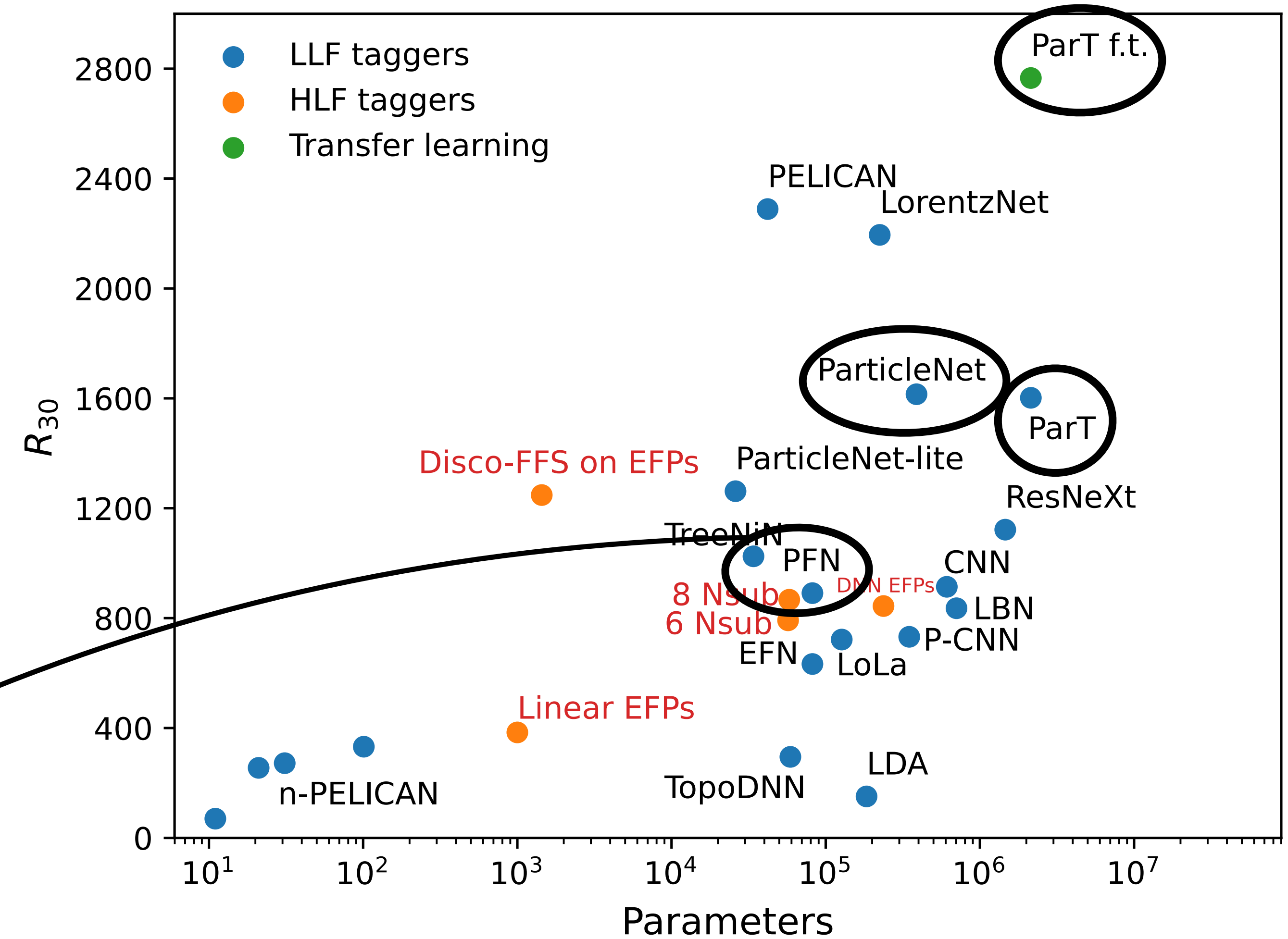
Locality and translation invariance, but not ideal



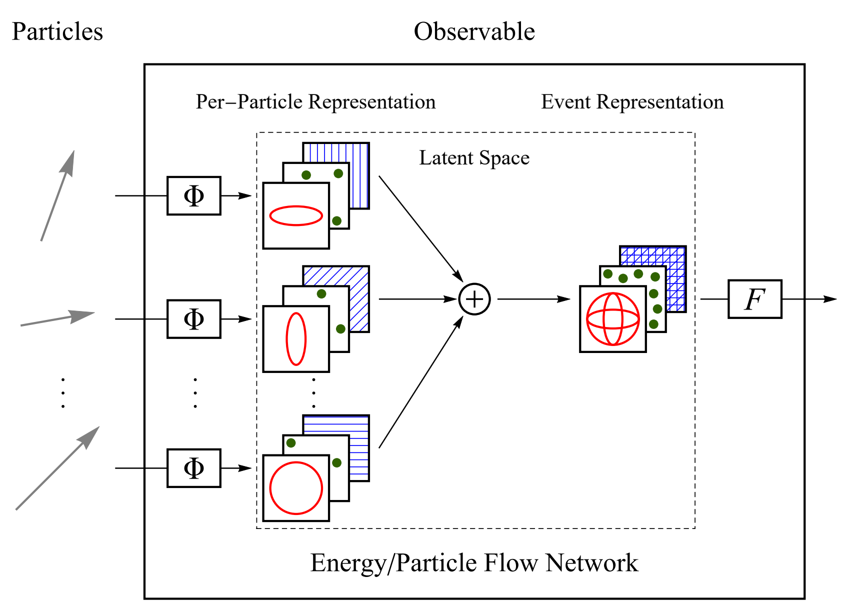


GK et al 1701.08784; Macaluso, Shih 1803.00107; Qu, Gouskos 1902.08570

Learning relations



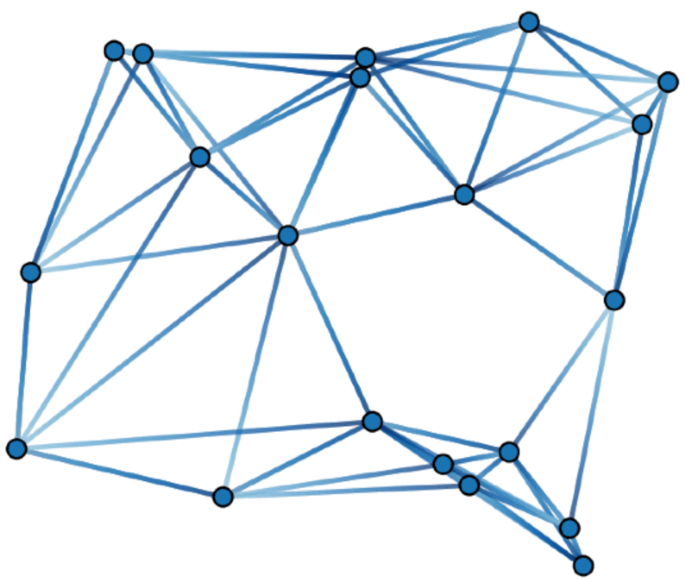
Point-clouds alone lack capacity

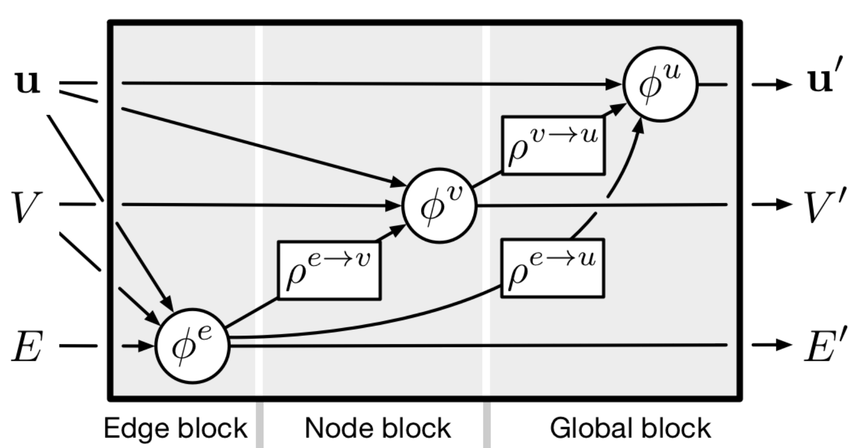


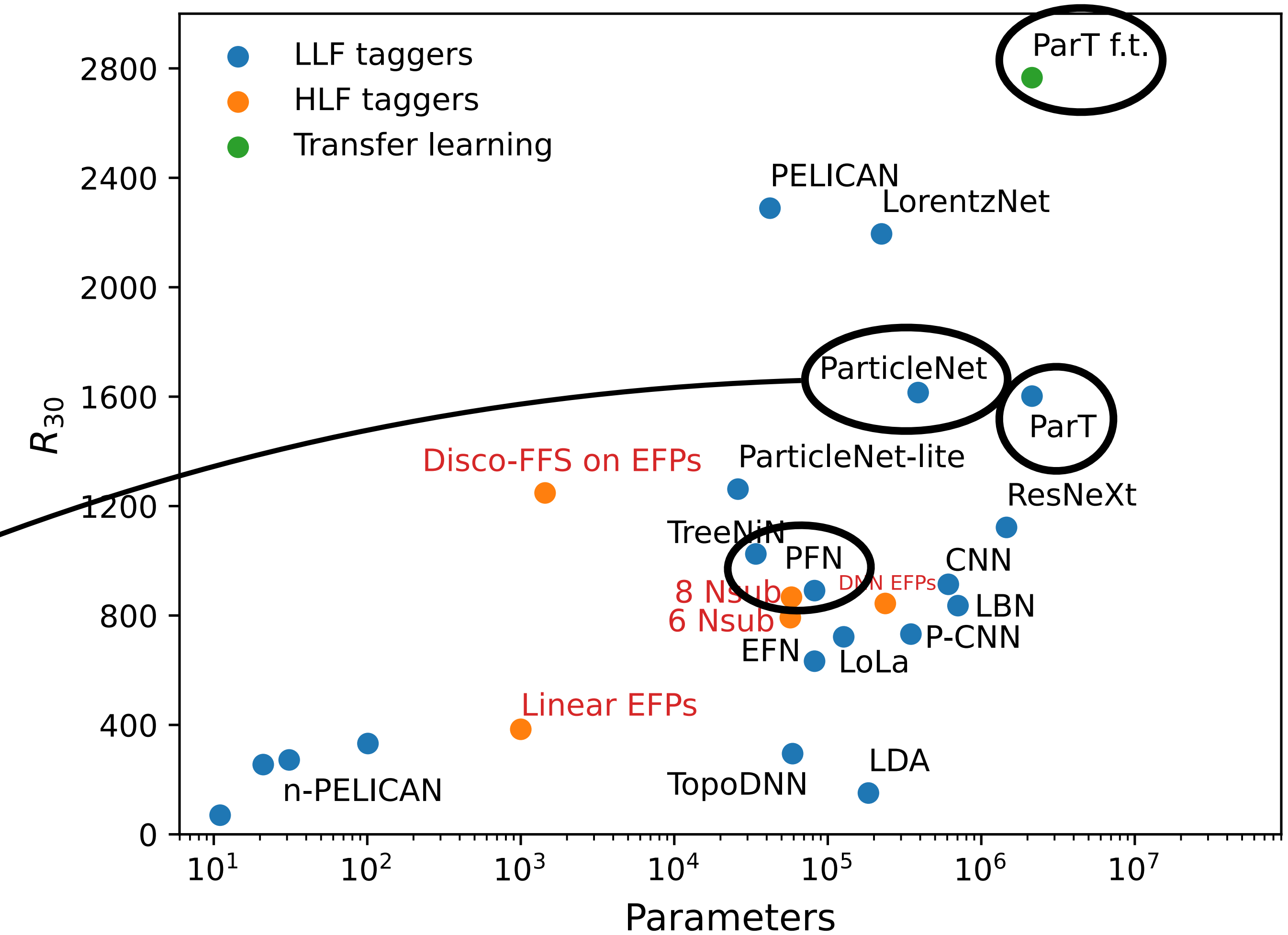
Komiske, Metodiev, Thaler 1810.05165;

Learning relations

Graphs add locality
(e.g. kNN clustering in
feature space)

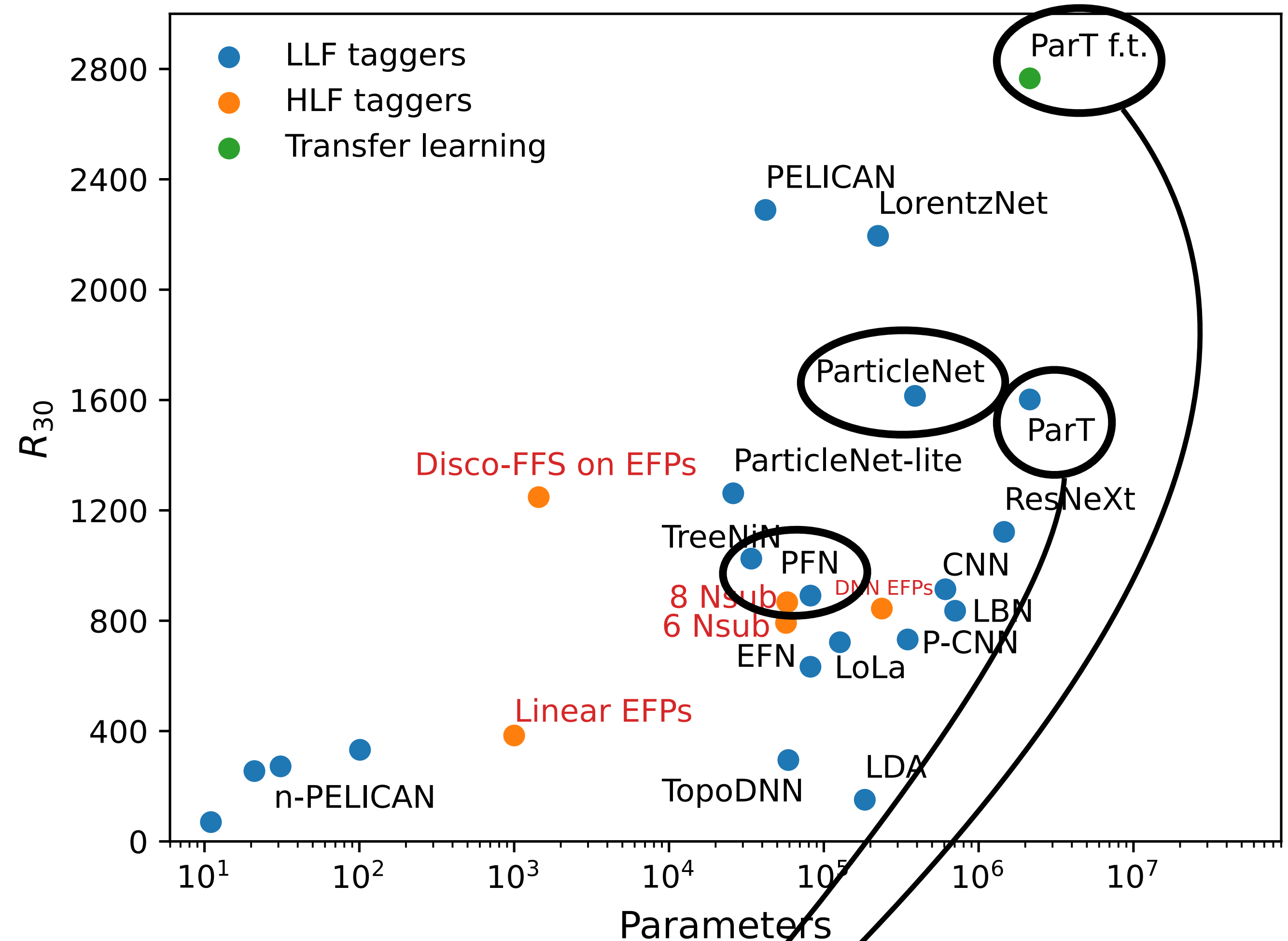






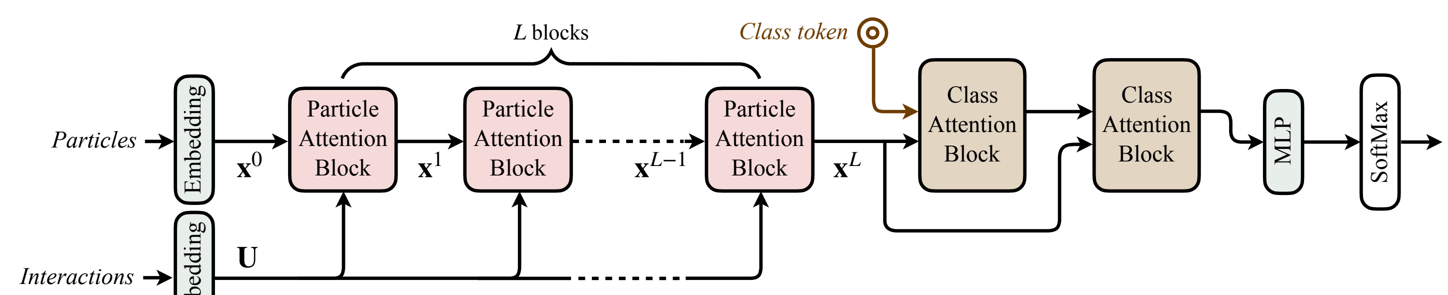
Qu, Gouskos 1902.08570; Qu et al 2202.03772; Gong et al 2201.08187; Shlomi et al 2007.13681;

Learning relations



Attention learns which neighbours are relevant

Benefits from pre-training



(a) Particle Transformer

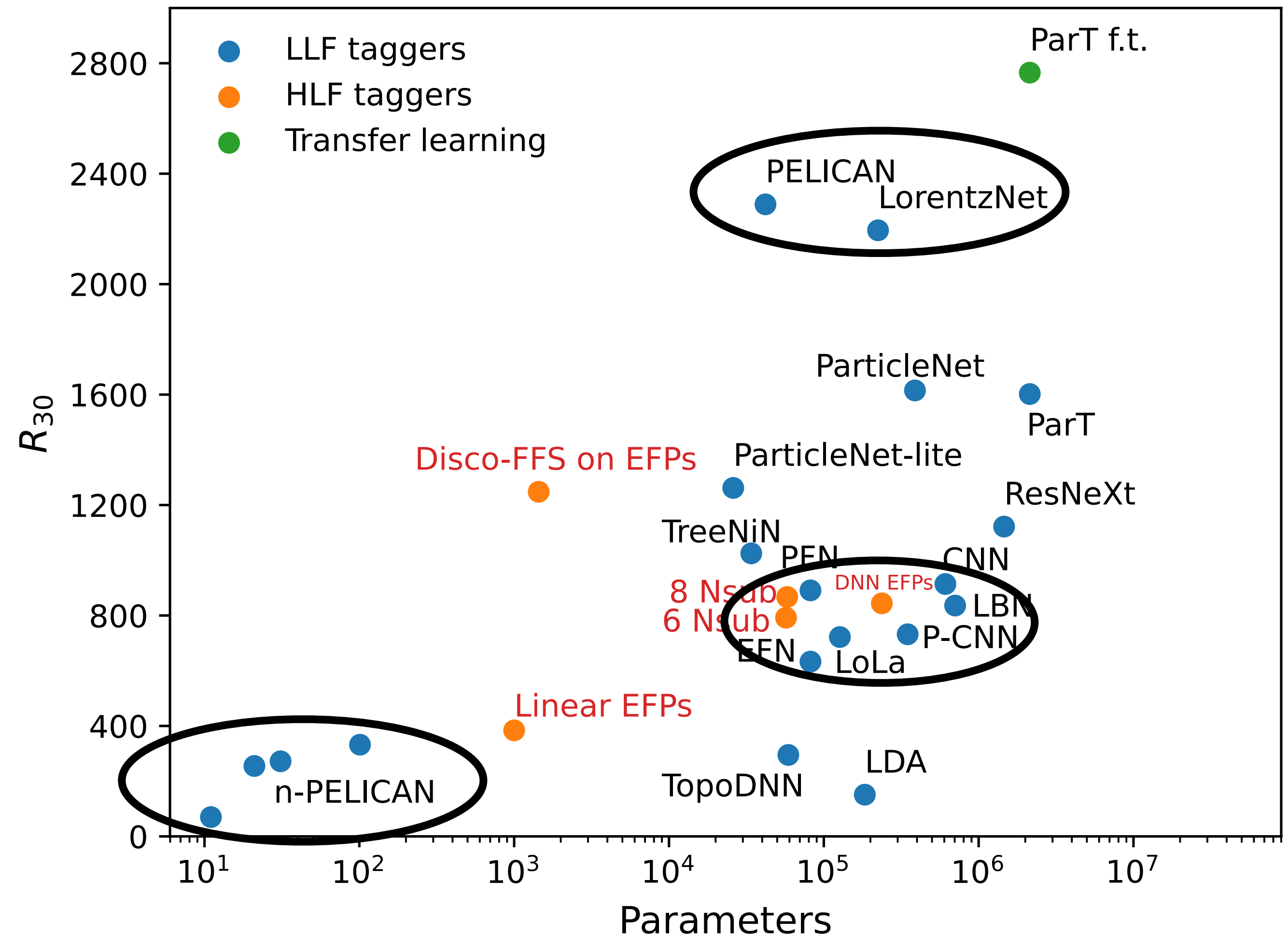
Qu, Gouskos 1902.08570; Qu et al 2202.03772; Gong et al 2201.08187; Shlomi et al 2007.13681;

Lorentz Symmetries

Extra invariance: Lorentz group

Tagging results should be (approximately) invariant under rotation, translation, and Lorentz boosts

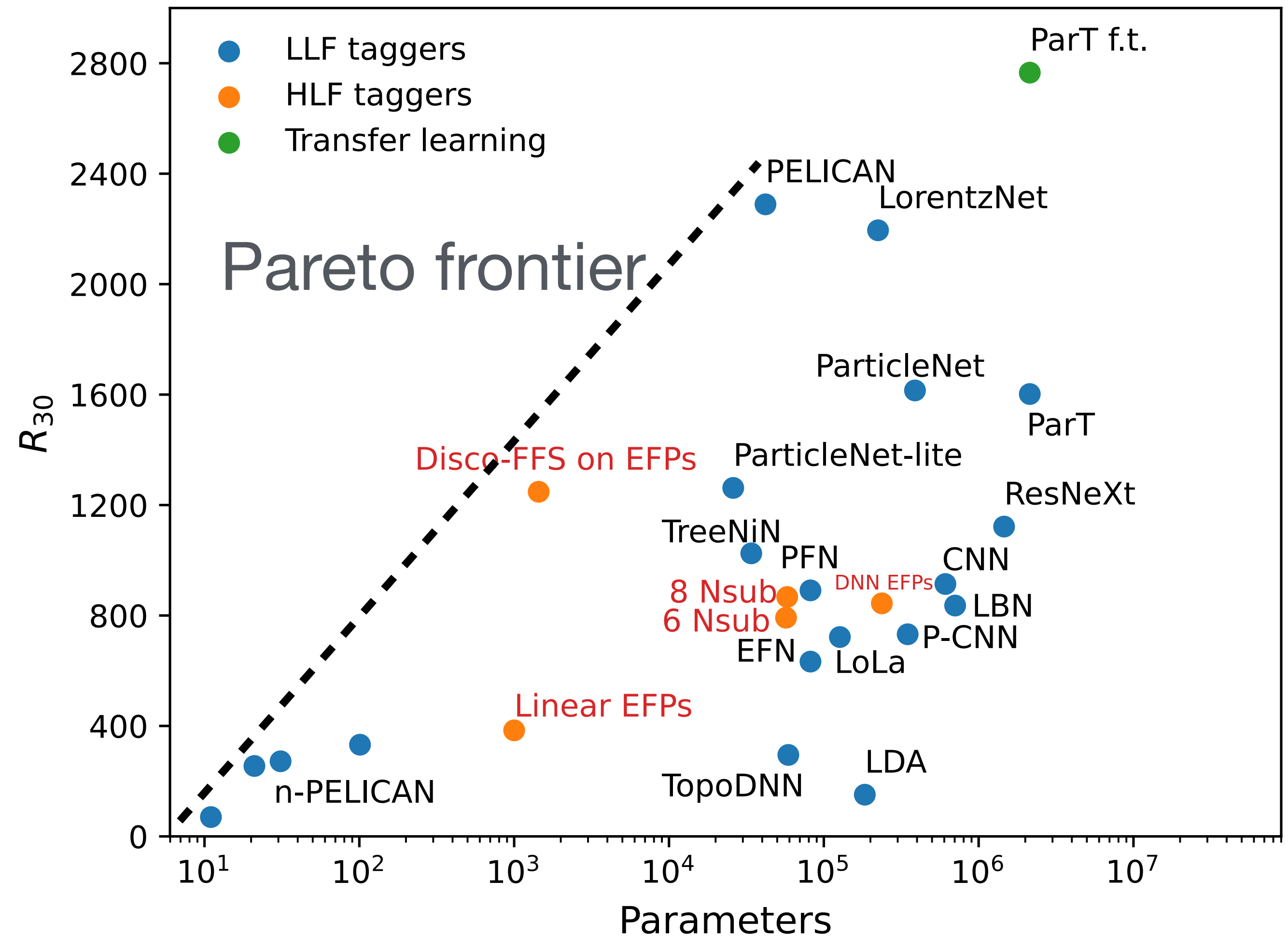
Powerful physical constraint



Butter, **GK**, et al 1707.08966; Bogatskiy et al 2211.00454; Gong et al 2201.08187;

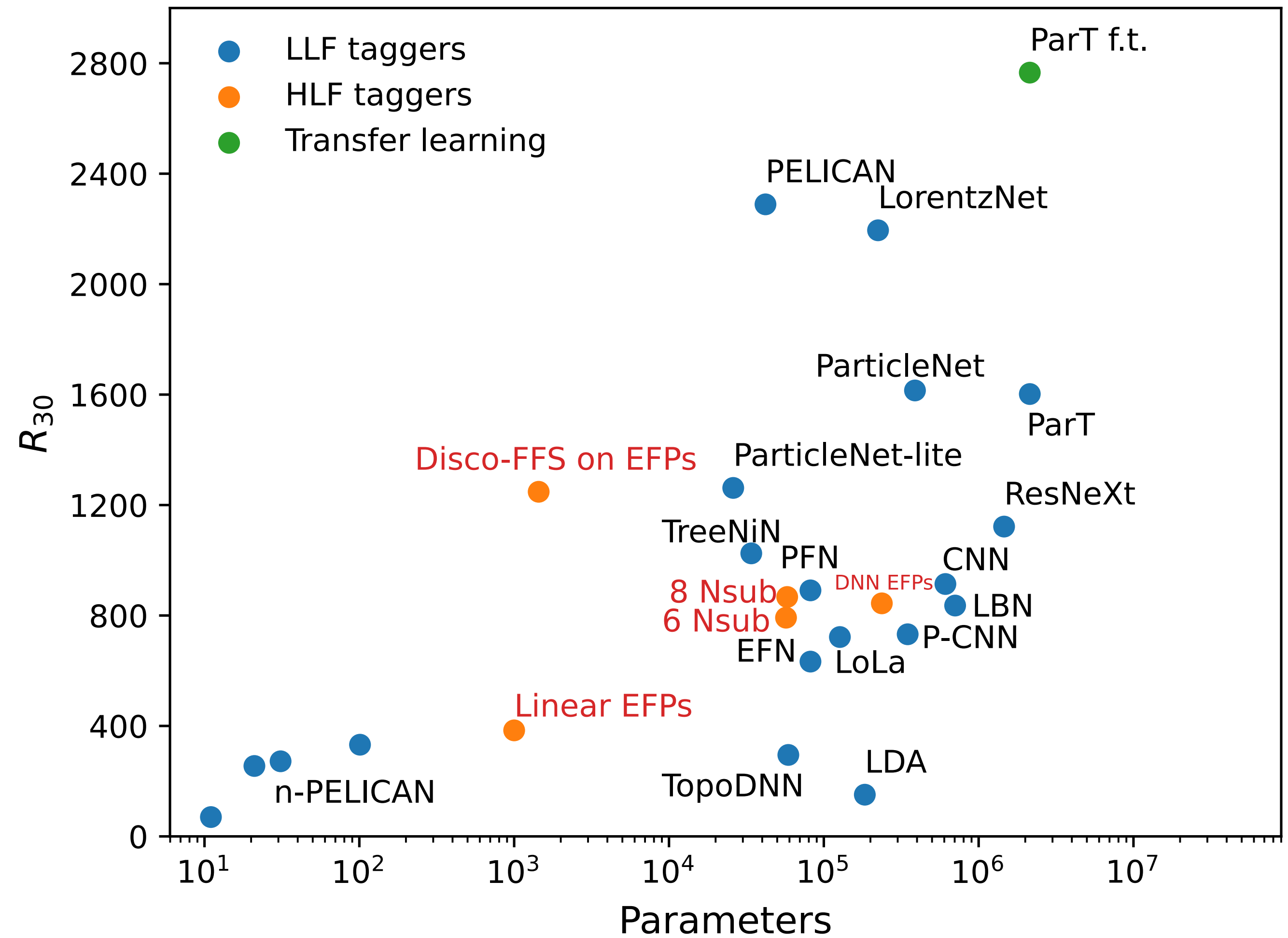
Take aways

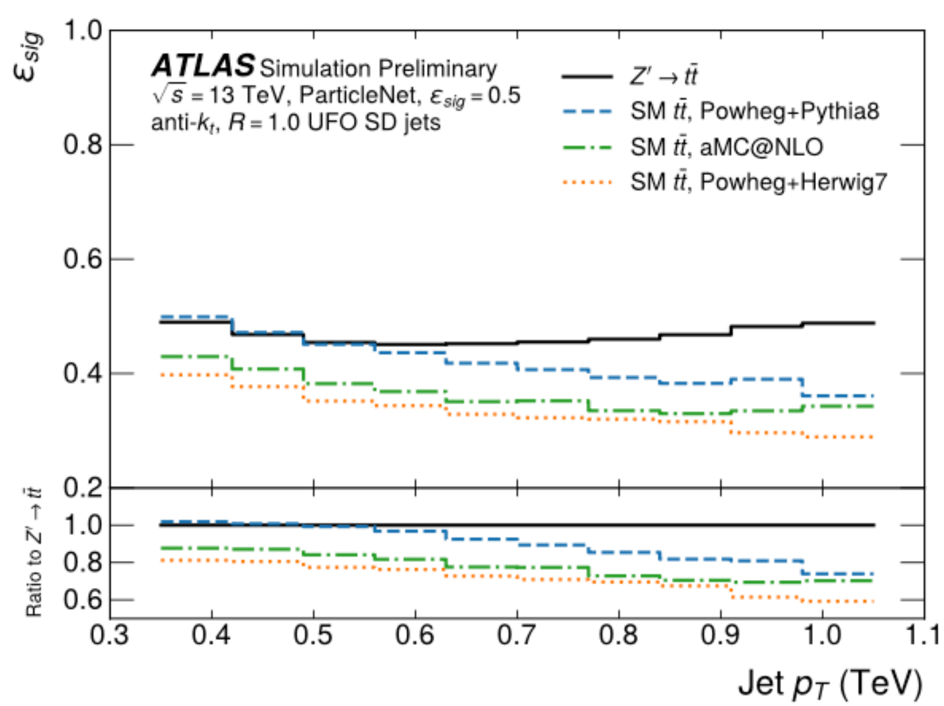
- Point clouds as powerful paradigm to represent data
- Additional structure in architecture boosts performance
- Over wide range: Best complexity/performance trade-off by physics-informed models
- Overall highest performance reached via transfer learning



(Some) Current challenges

- Point clouds as powerful paradigm to represent data
- Additional structure in architecture boosts performance
- Over wide range: Best complexity/performance trade-off by physics-informed models
- Overall highest performance reached via transfer learning



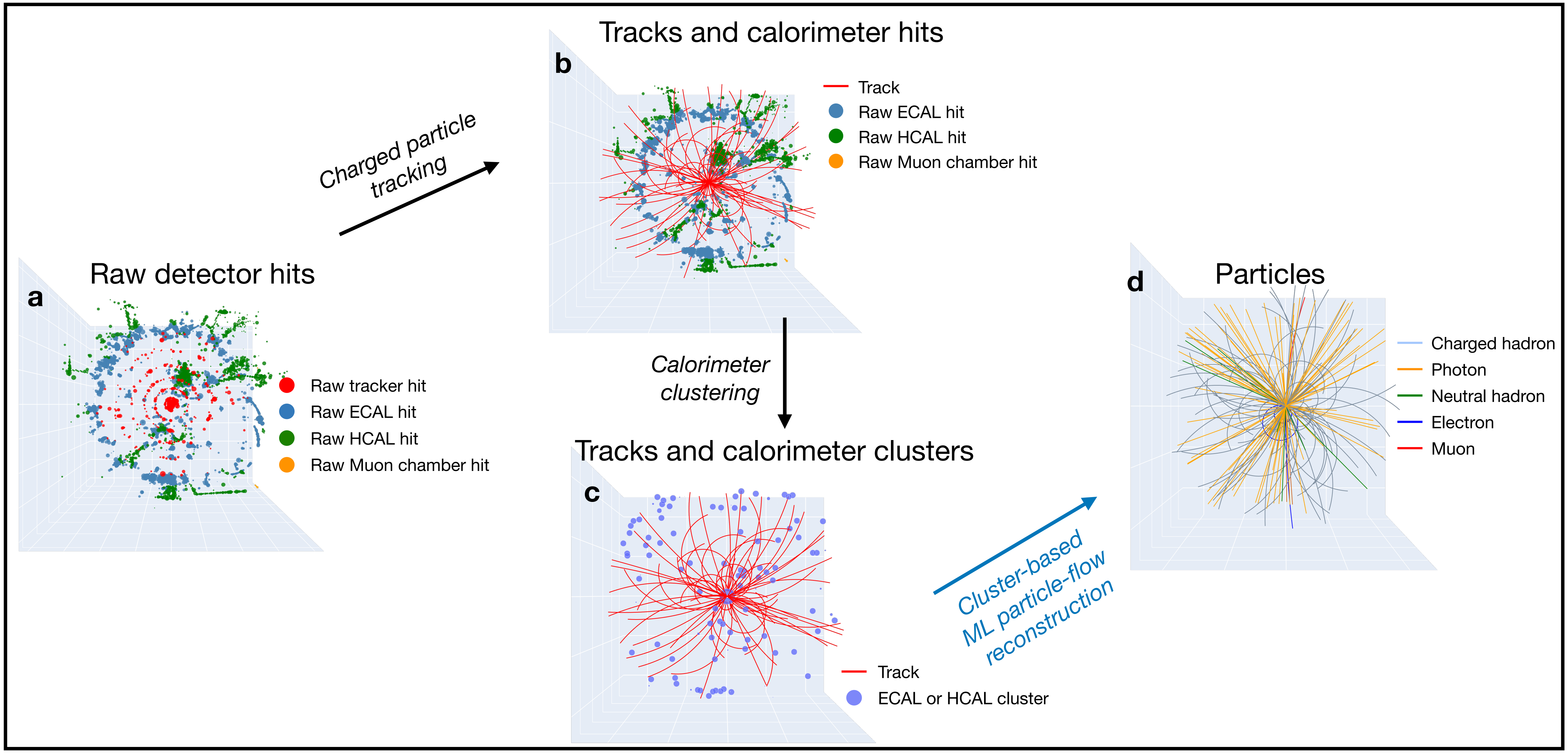


- "Calibration": Domain adaptation between simulation and collider data
- Uncertainty aware training
- Interpretability

ATL-PHYS-PUB-2022-039

2. Reconstruction

Reconstruction



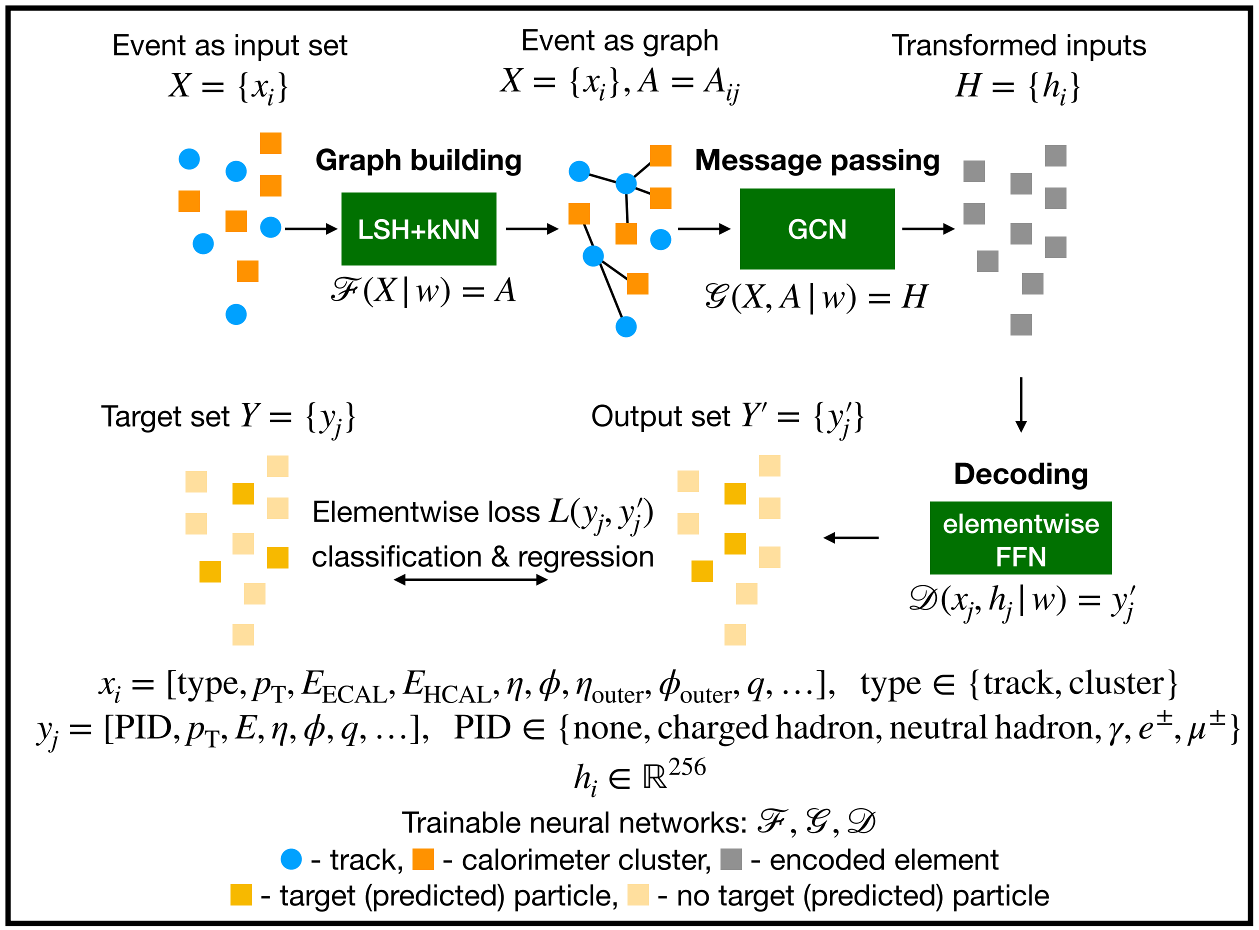
Reconstruction maps low-level detector read-outs

to physical particles

(Which in-turn are the basis of higher-level interpretation)

Pata et al 2309.06782;

Particle Flow

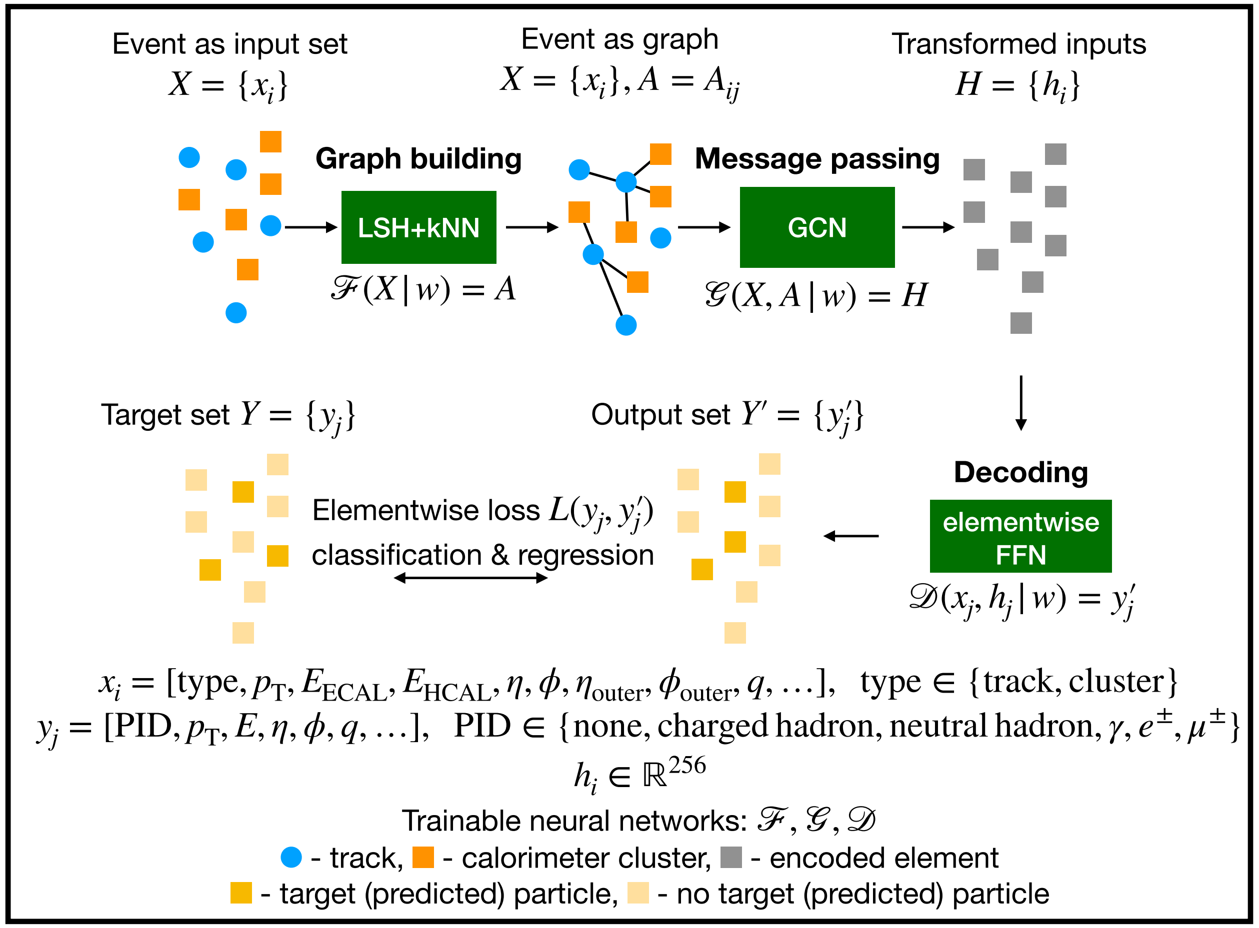


Particle flow is the task of turning trajectories and energy deposits in 3D space into meaningful particles

Graph-based approach can learn this mapping

Pata et al 2309.06782; Pata et al 2101.08578; See also Di Bello et al 2003.08863

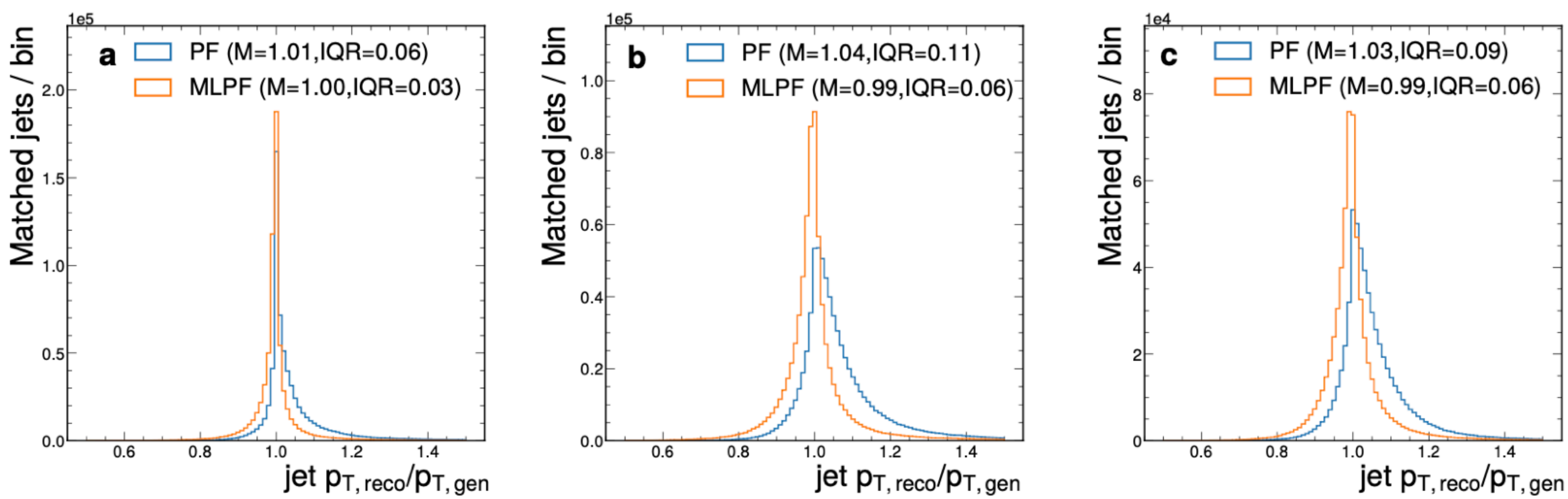
Particle Flow



Particle flow is the task of turning trajectories and energy deposits in 3D space into meaningful particles

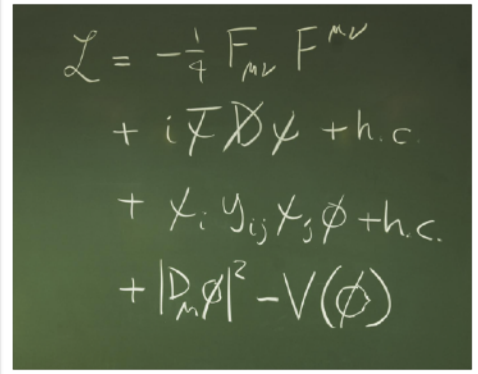
Graph-based approach can learn this mapping

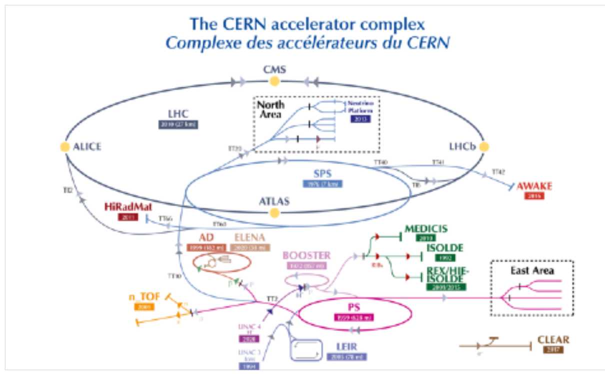
And improve on classical Rule-based approaches

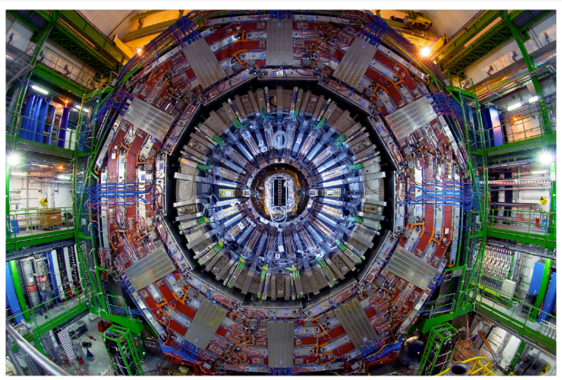


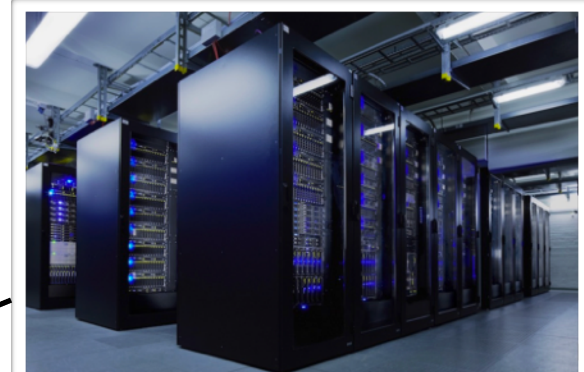
Pata et al 2309.06782; Pata et al 2101.08578

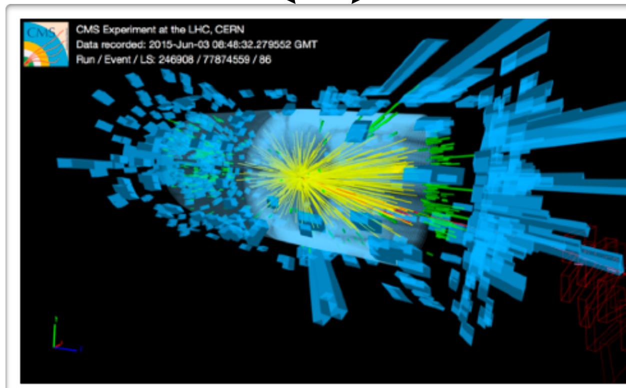
3. Simulation





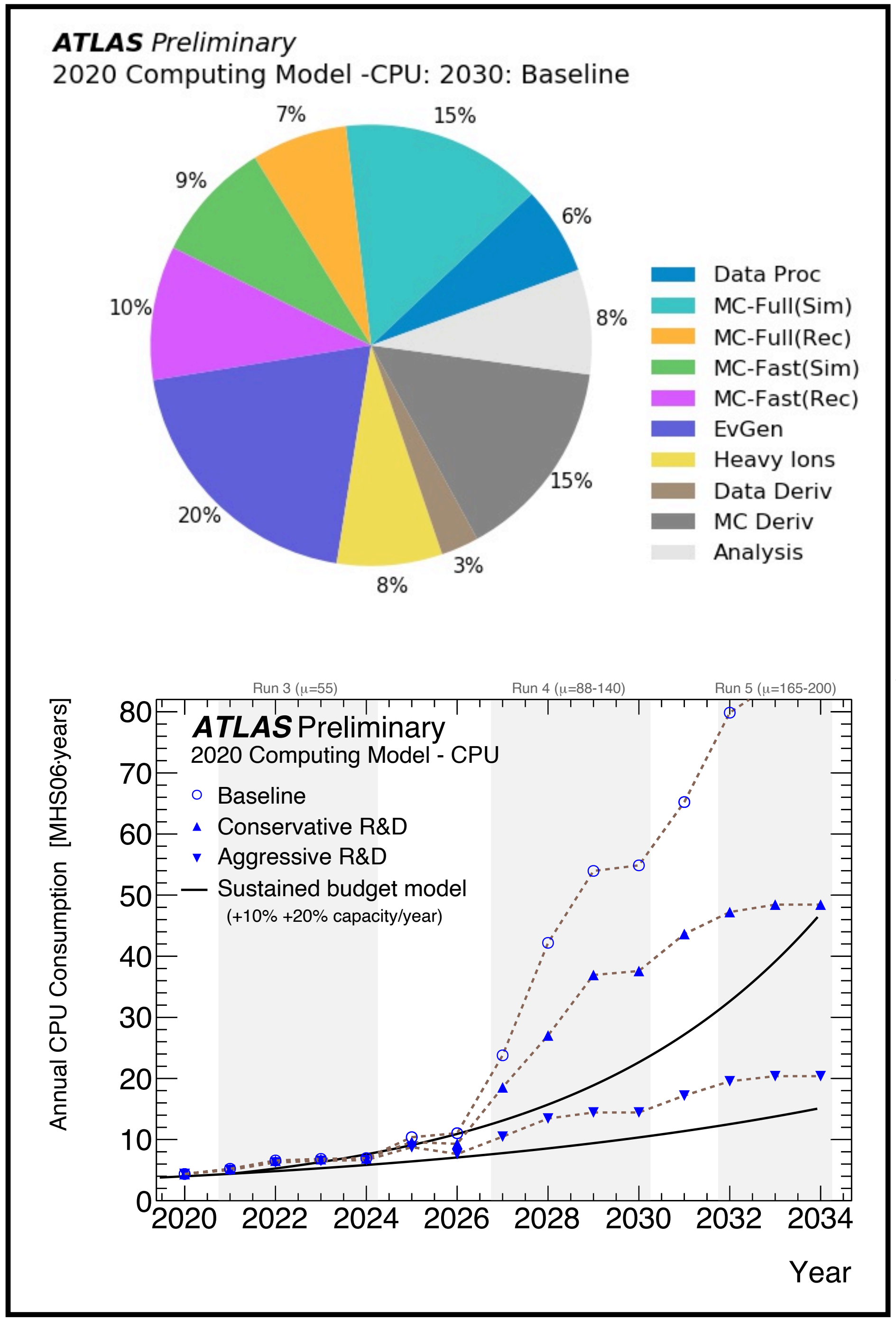






** Physics events are snapshots, simulation means sampling from a distribution, not time-evolution*

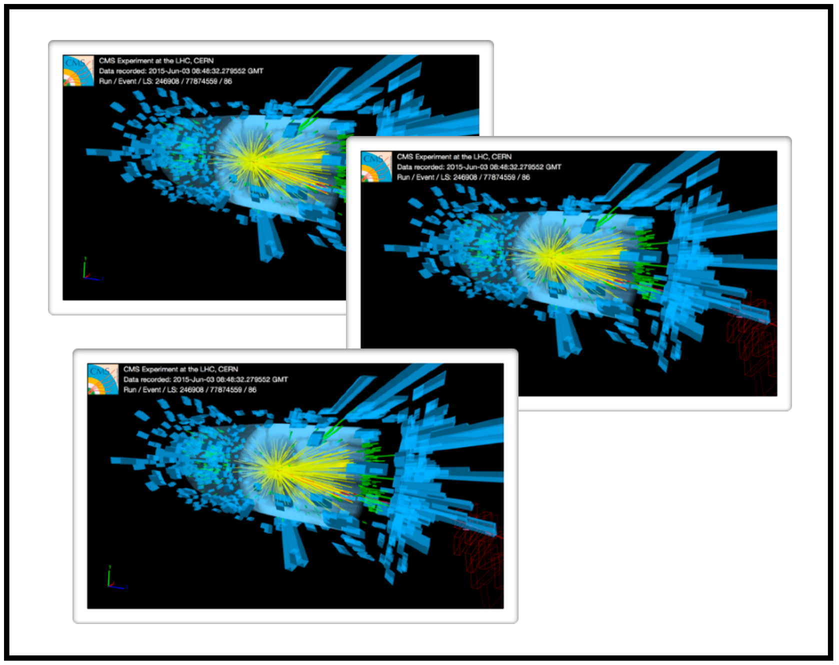
Simulation* is crucial to connect experimental data with theory predictions



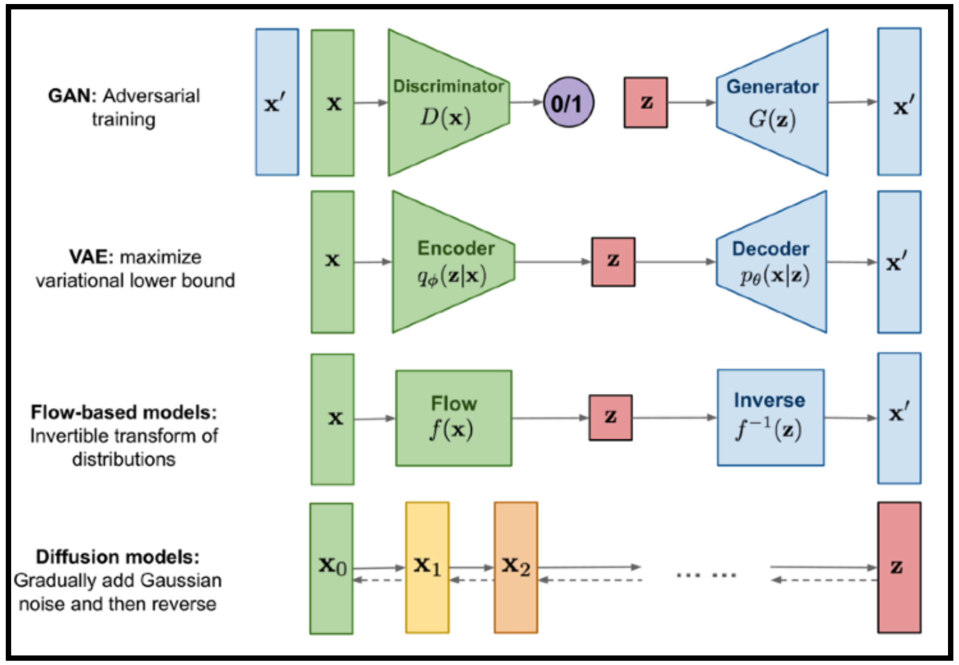
Simulation is crucial to connect experimental data with theory predictions, but computationally very costly

Strategy

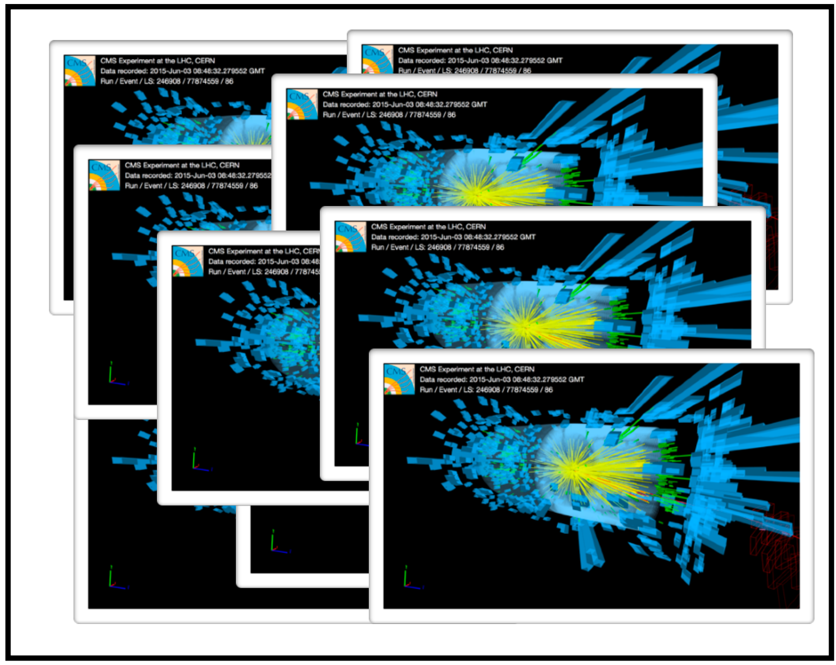
1. Use classical simulation or collider data as input



2. Train generative surrogate

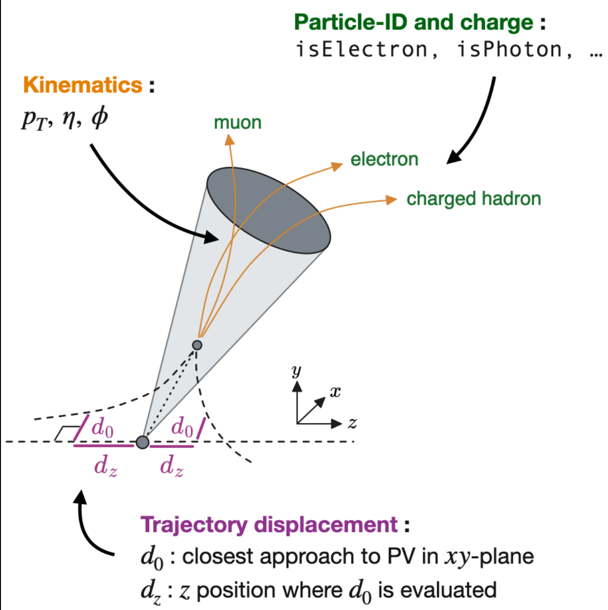


3. Oversample

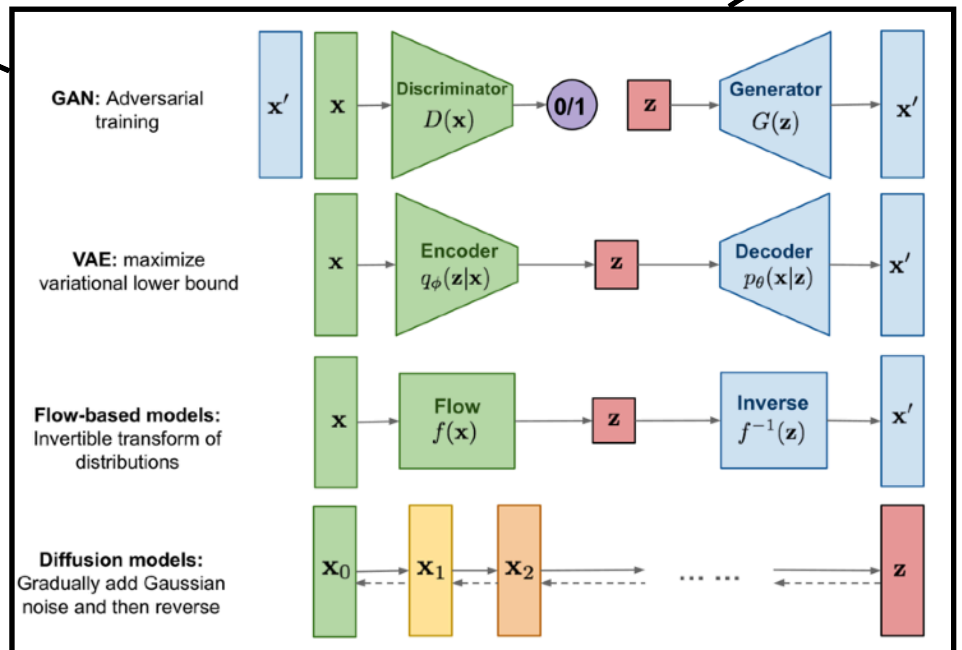


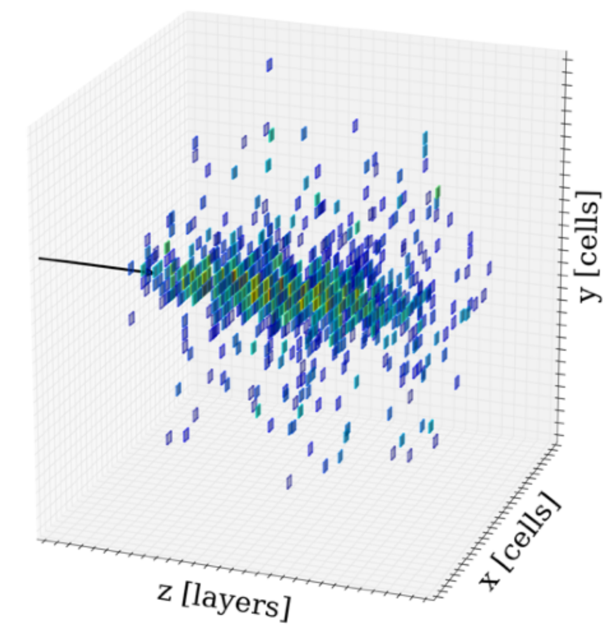
Paganini, Oliveira, Nachman 1705.02355; Butter, Diefenbacher, **GK**, et al 2008.06545;

Main Targets

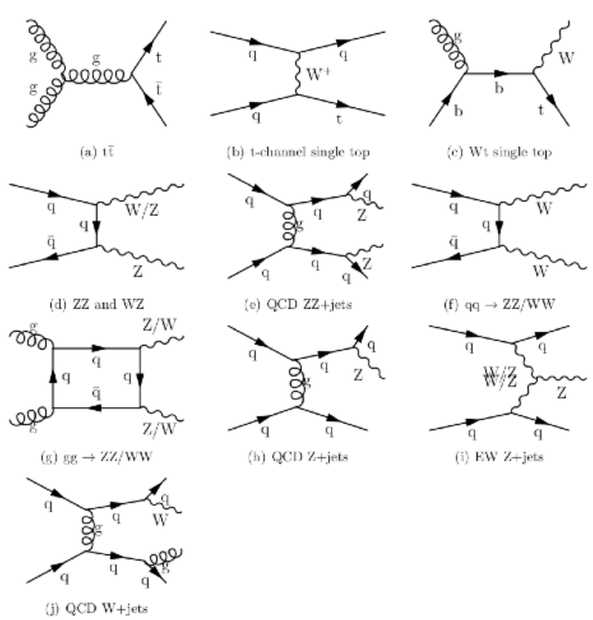


Jet constituents

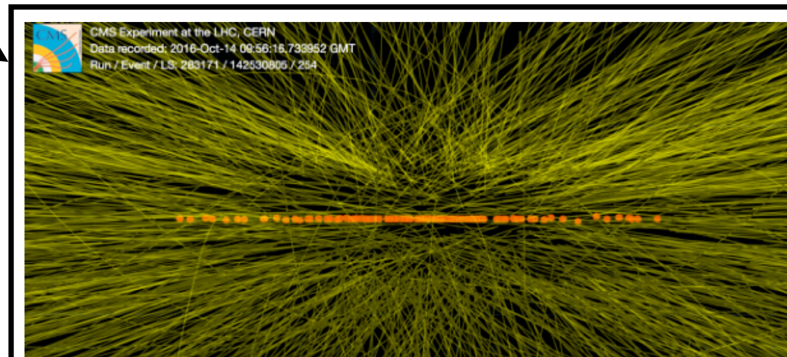




Showers in complex high-resolution calorimeters



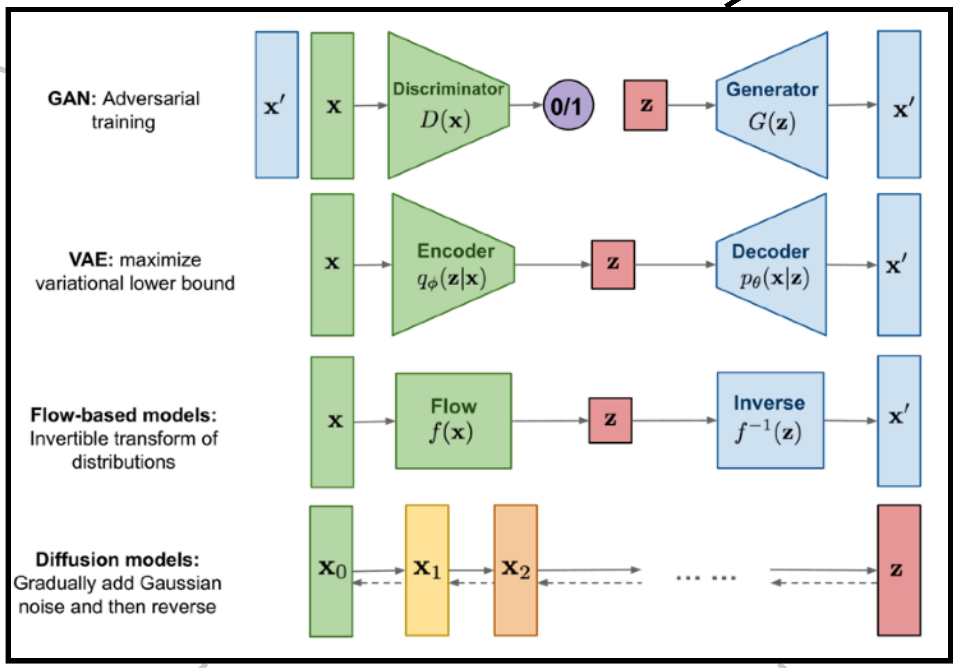
Event-level kinematics

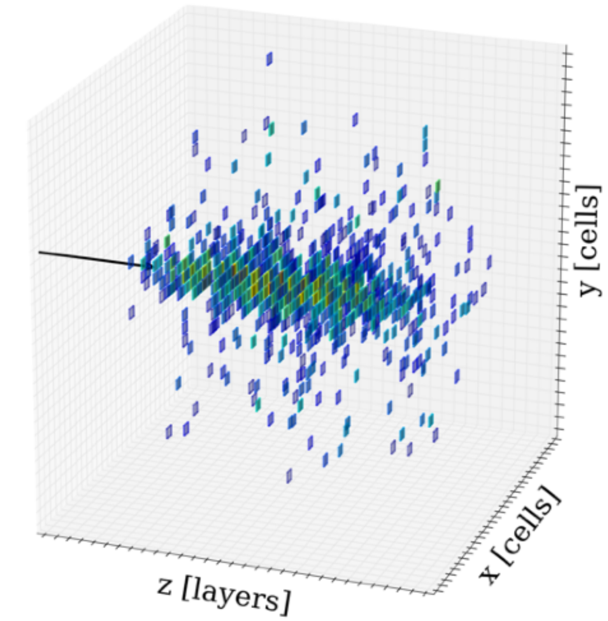


Pile-up Interactions

Main Targets

Jet constituents





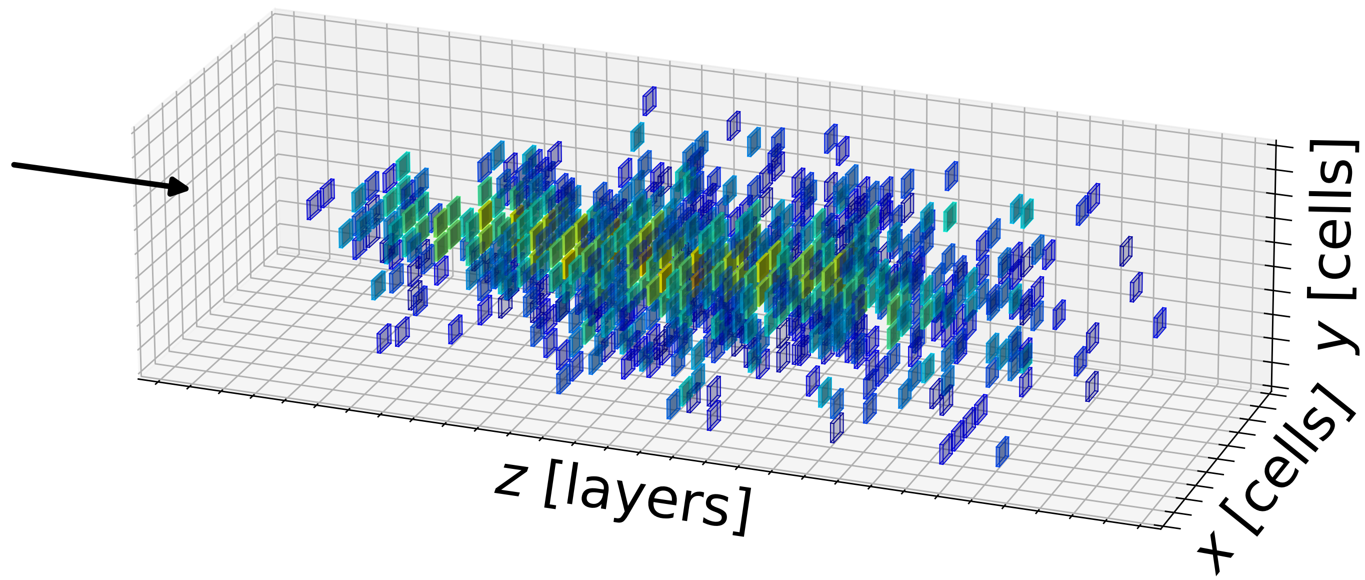
Showers in complex high-resolution calorimeters:
as fixed grid
as point cloud

Event-level kinematics

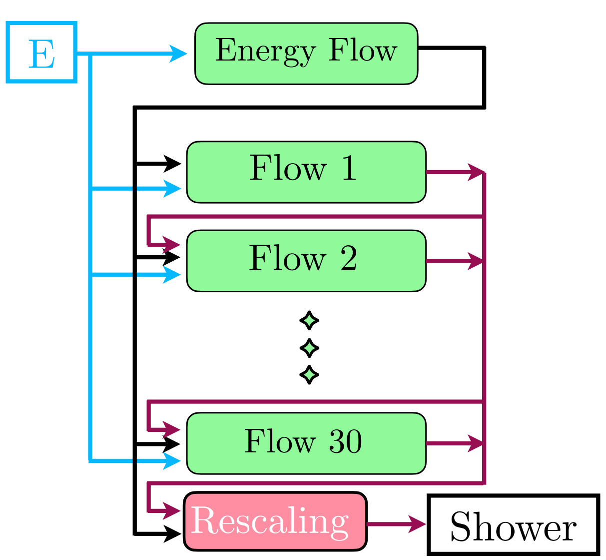
Pile-up Interactions

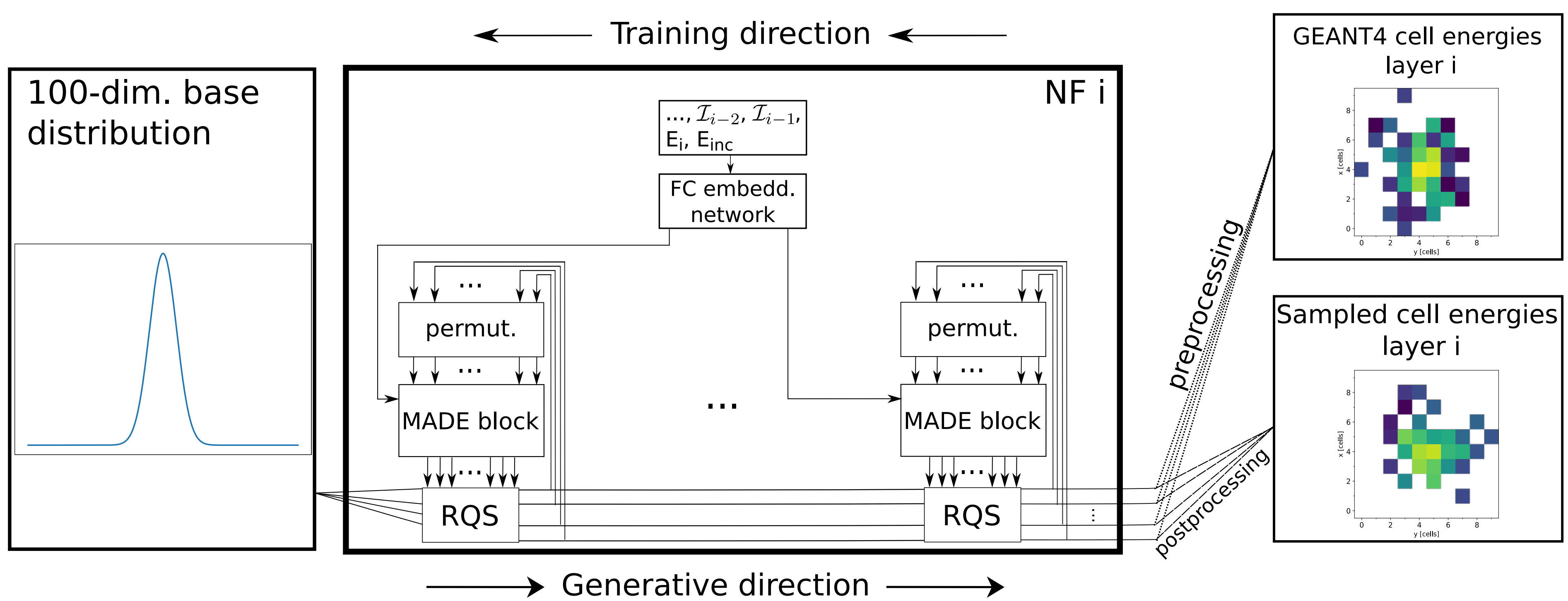
Fixed Grid

Fixed grid — i.e. 3D voxel images:
Pro: Standard generative models work
Con: Mostly empty, scales badly



How to do flows for high-dimensional data?
Split!

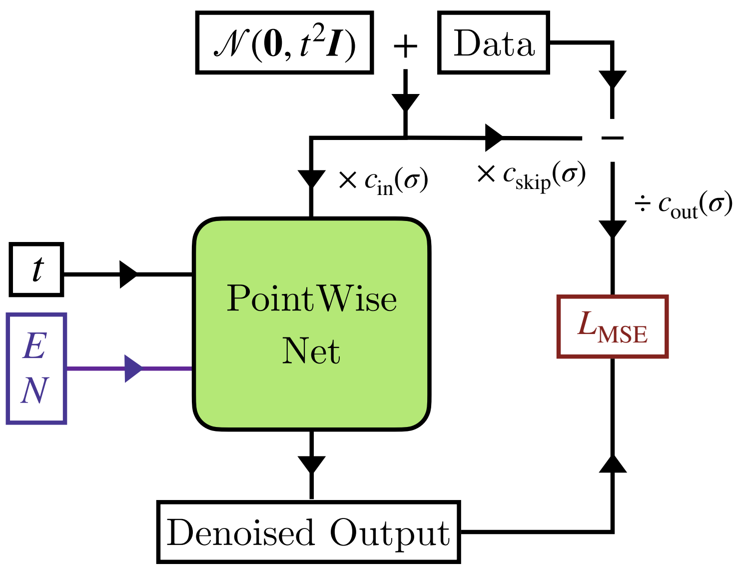


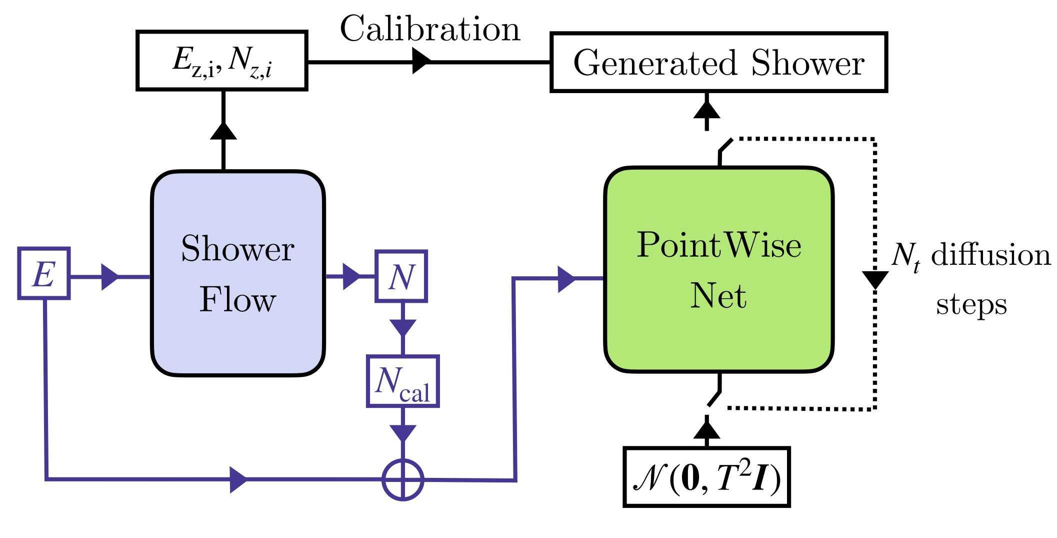


2302.11594

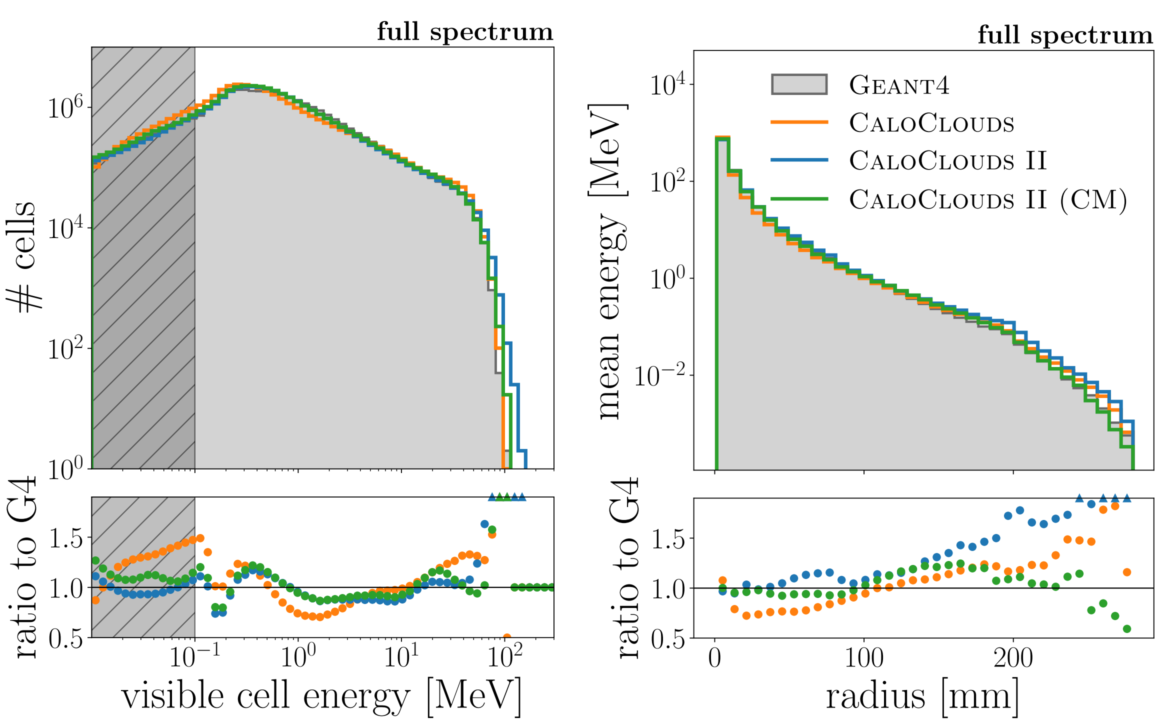
Point Cloud

Point cloud
Pro: Scales to arbitrary geometries
Also pro: Requires additional developments





Use continuous time diffusion and consistency distillation: Better quality and faster



Hardware	Simulator	NFE	Batch Size	Time / Shower [ms]	Speed-up
CPU	Geant4			3914.80 ± 74.09	×1
	CaloClouds	100	1	3146.71 ± 31.66	×1.2
	CaloClouds II	25	1	651.68 ± 4.21	×6.0
	CaloClouds II (CM)	1	1	84.35 ± 0.22	×46
GPU	CaloClouds	100	64	24.91 ± 0.72	×157
	CaloClouds II	25	64	6.12 ± 0.13	×640
	CaloClouds II (CM)	1	64	2.09 ± 0.13	×1873

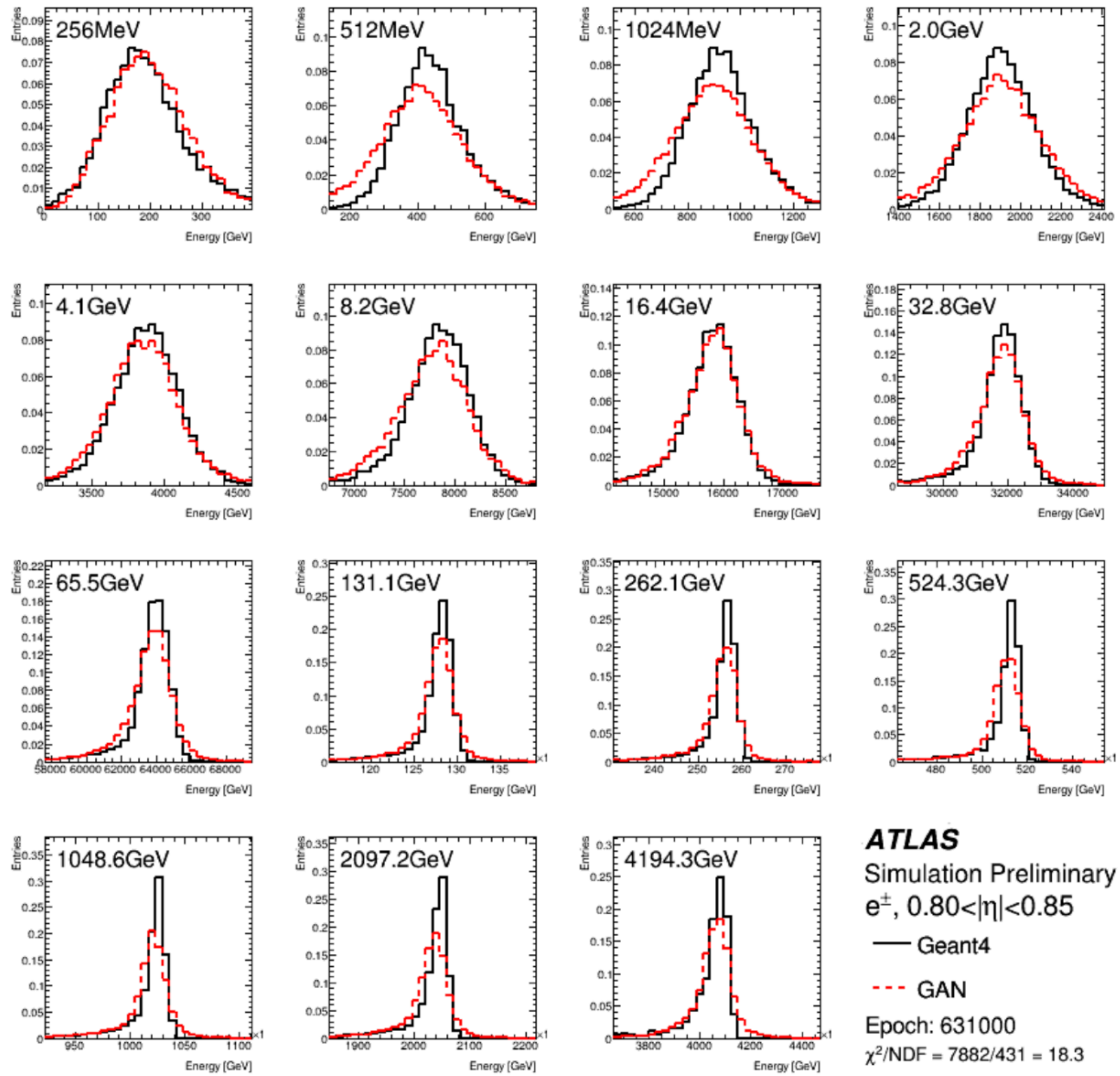
Buhmann, .., **GK**, et al 2309.05704; & much more, see 2312.09597

Application

Not only theoretical development: e.g. ATLAS includes FastCaloGAN in ATLFAST3

100 networks (slices in η)

O(500) voxels



ATL-SOFT-PUB-2020-006; ATLAS 2109.02551

Moving forward

- 3 Public datasets to compare simulation techniques
  - Simplest: ATLAS dataset (see prev. page)
  - Most complex: Future detector with 40k voxels
  - Write-up currently ongoing

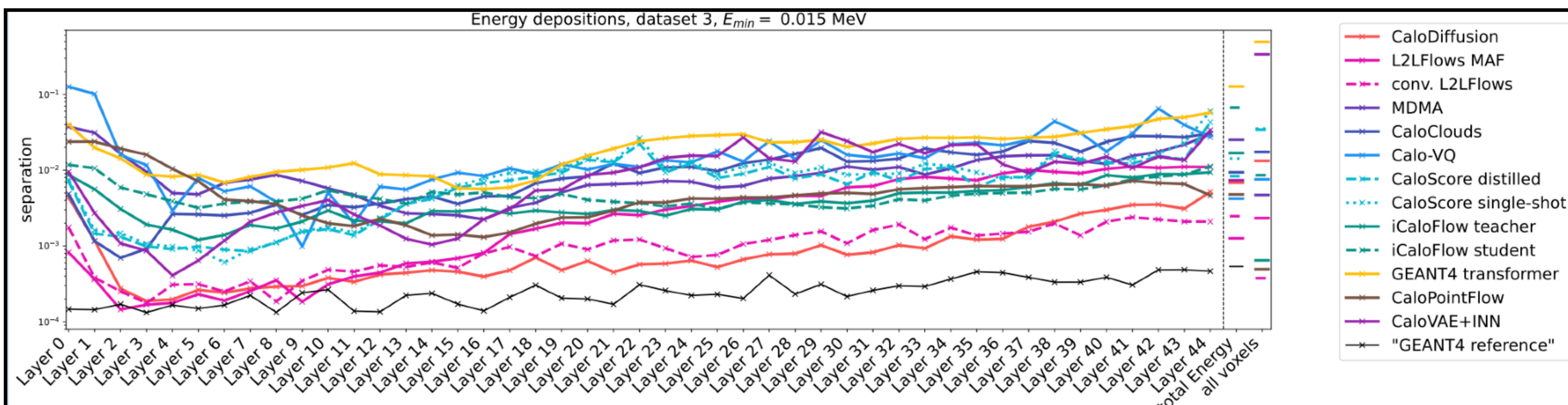
Fast Calorimeter Simulation Challenge 2022

View on GitHub

Welcome to the home of the first-ever Fast Calorimeter Simulation Challenge!

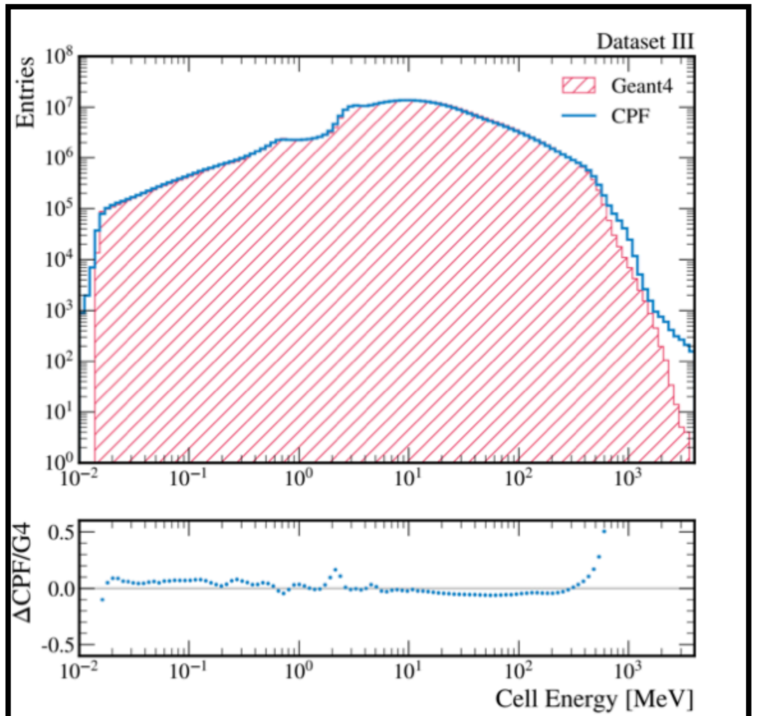
The purpose of this challenge is to spur the development and benchmarking of fast and high-fidelity calorimeter shower generation using deep learning methods. Currently, generating calorimeter showers of interacting particles (electrons, photons, pions, ...) using GEANT4 is a major computational bottleneck at the LHC, and it is forecast to overwhelm the computing budget of the LHC experiments in the near future. Therefore there is an urgent need to develop GEANT4 emulators that are both fast (computationally lightweight) and accurate. The LHC collaborations have been developing fast simulation methods for some time, and the hope of this challenge is to directly compare new deep learning approaches on common benchmarks. It is expected that participants will make use of cutting-edge techniques in generative modeling with deep learning, e.g. GANs, VAEs and normalizing flows.

This challenge is modeled after two previous, highly successful data challenges in HEP – the top tagging community challenge and the LHC Olympics 2020 anomaly detection challenge.

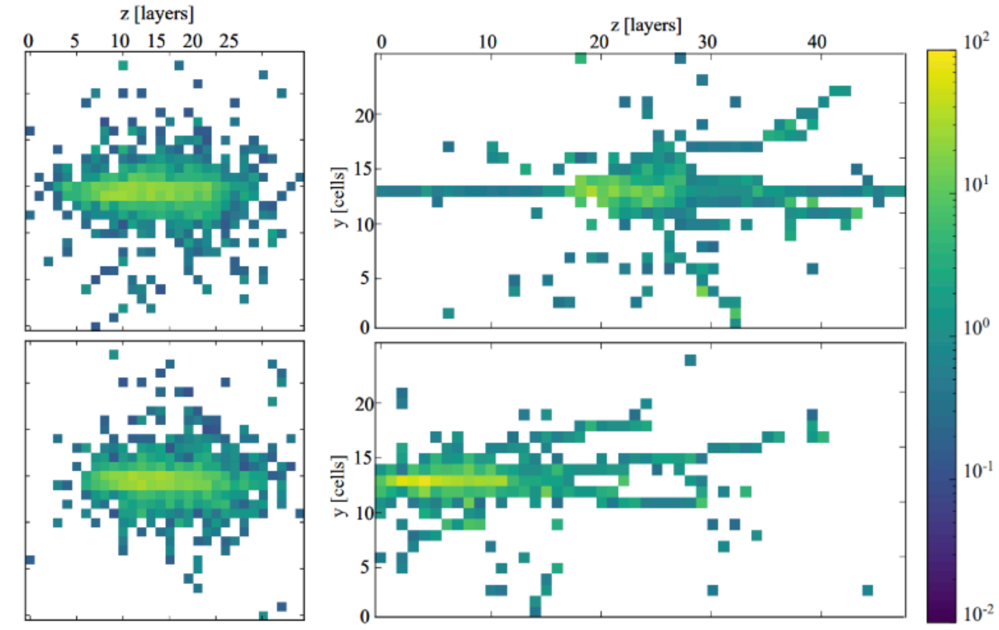


https://calochallenge.github.io/homepage/; Krause at ML4Jets ‘23

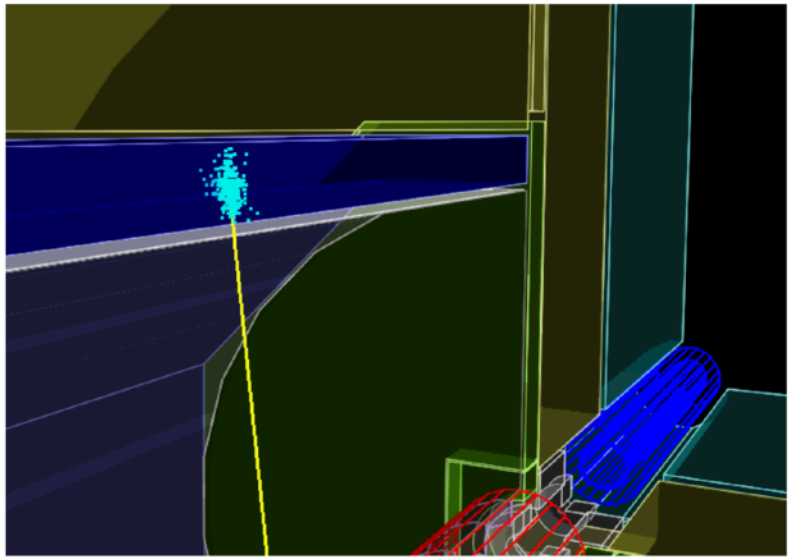
(Some) Current challenges



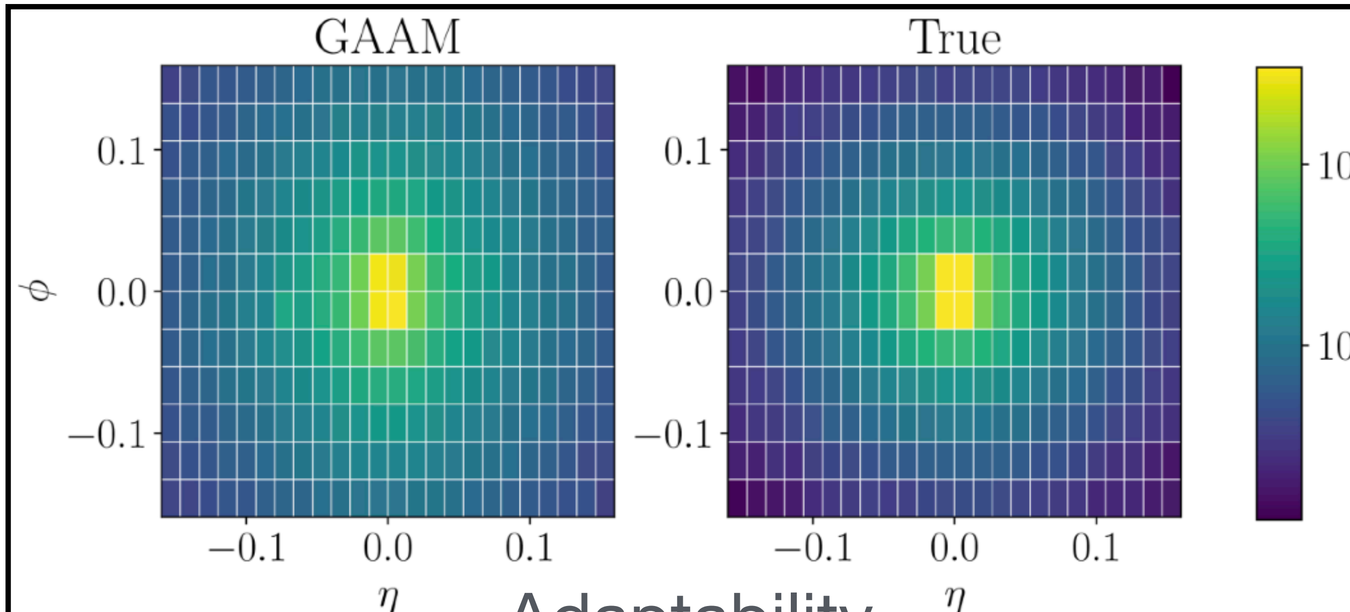
Quality of generation



Complexity of samples



GEANT4 integration



Adaptability

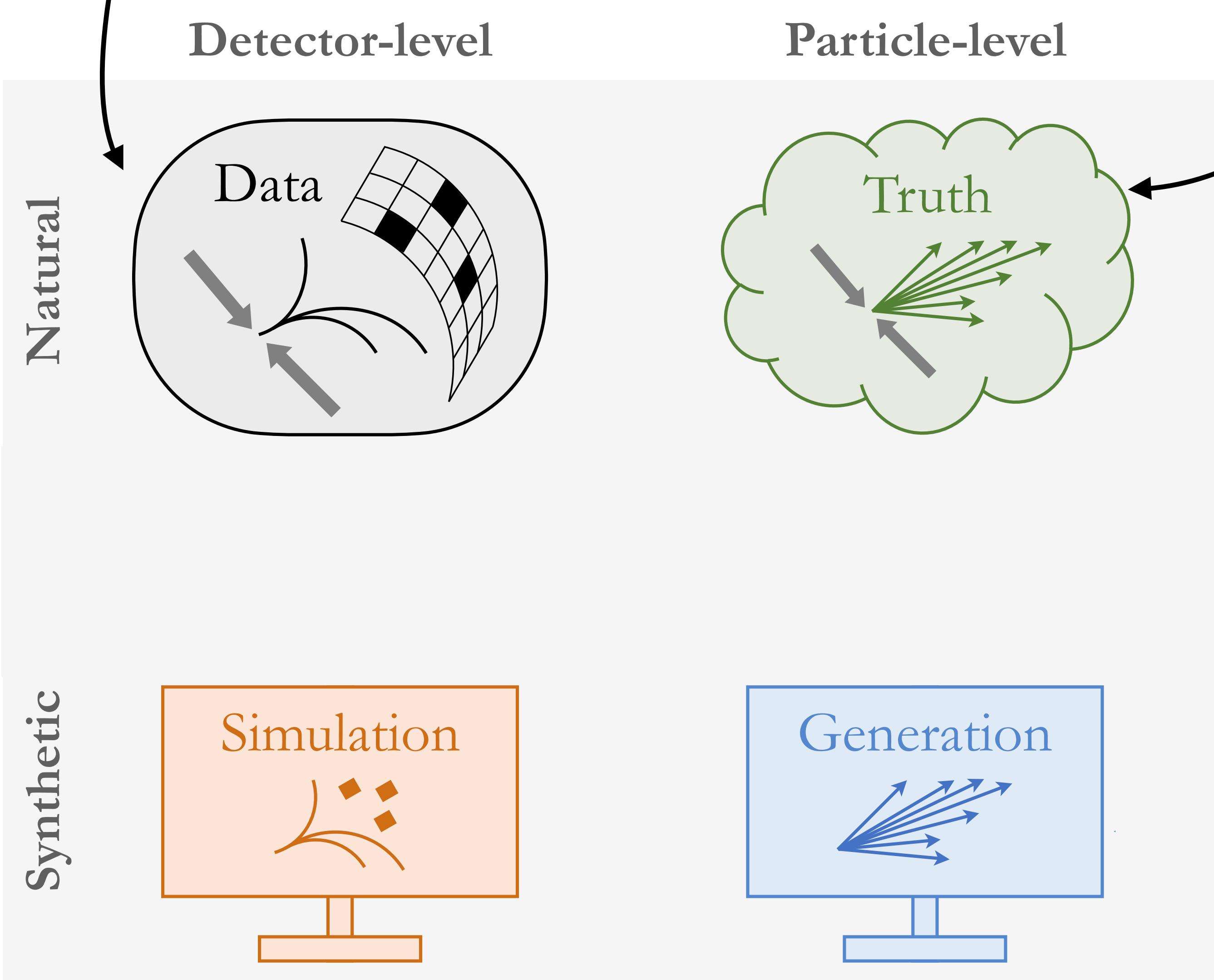
Buhmann, .., **GK**, et al 2112.09709;
Schnake, Krücker, Borras 2403.15782; Buhmann, .. **GK**, et al EPS-HEP ’23; Liu et al 2305.11531

4. Unfolding

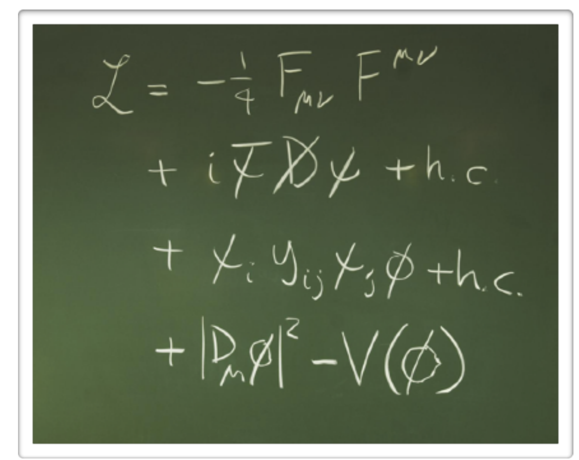
Unfolding

Observe

Want to know



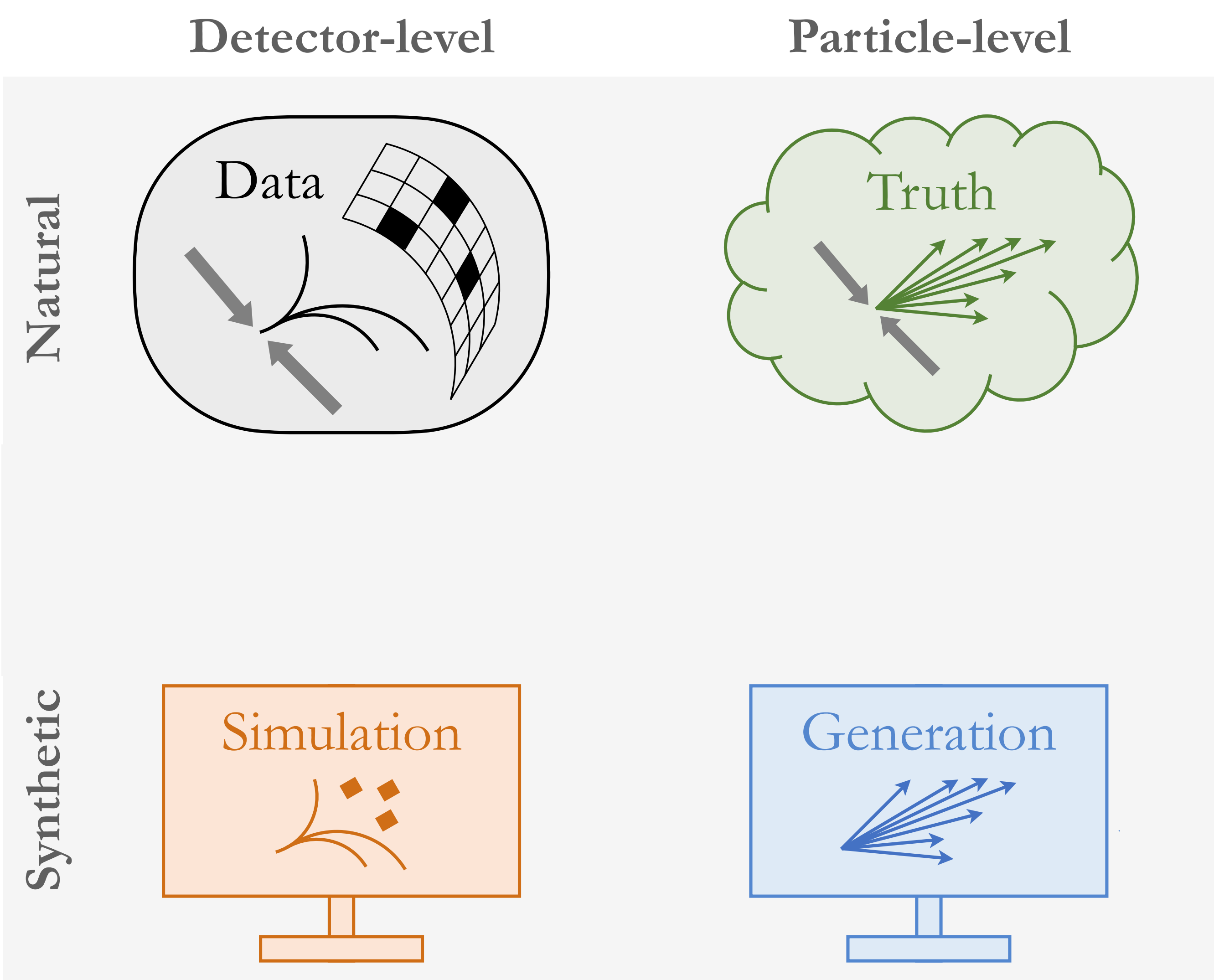
(and relate to fundamental theory)



Synthetic data provides both views: How to use?

Andreassen et el 1911.09107;

Unfolding

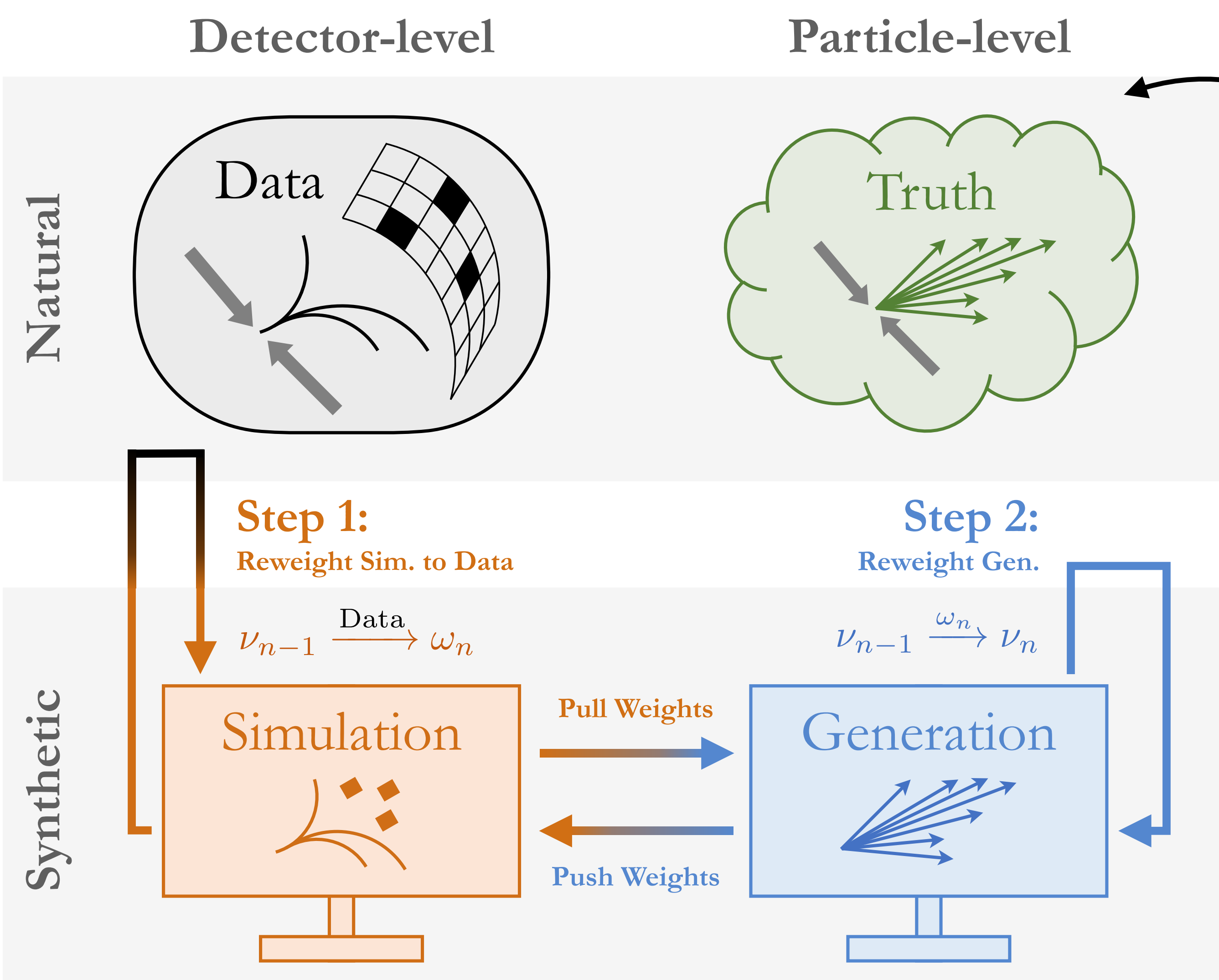


2 key approaches:

- Reweighting based on classifiers
- Morphing based on diffusion or generative models

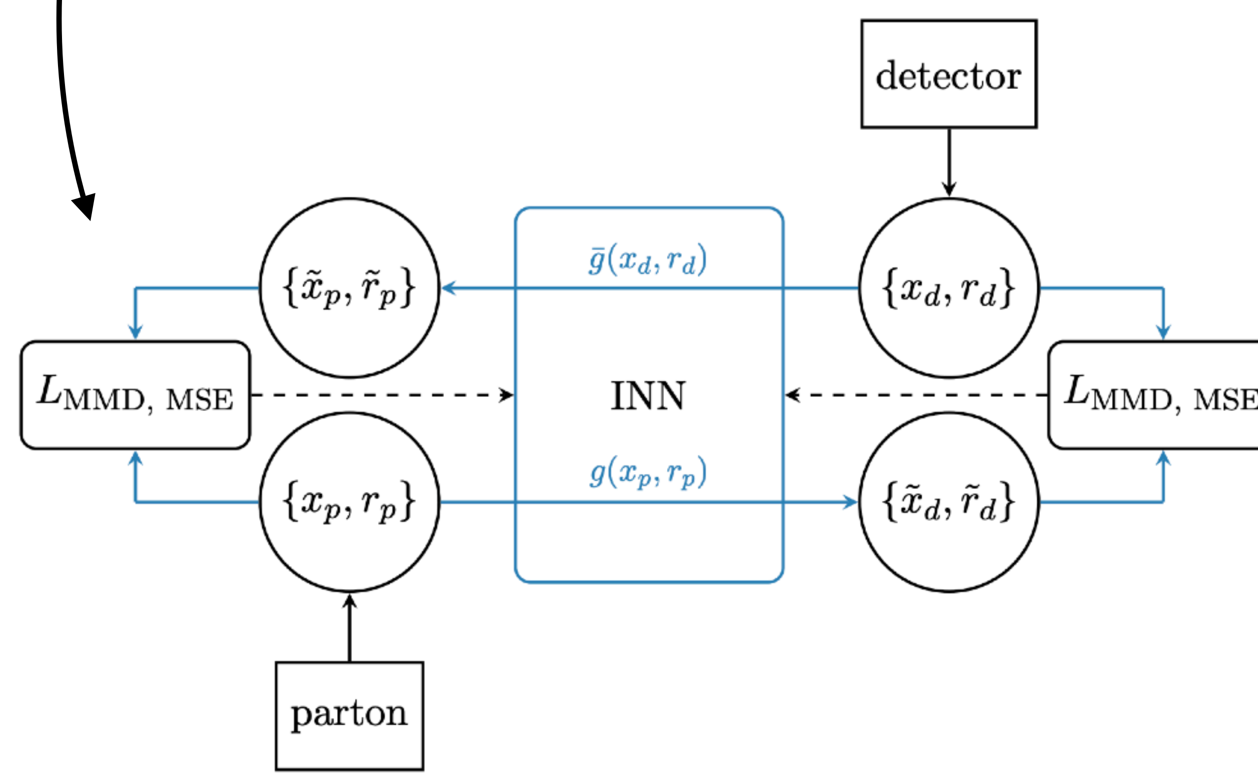
Andreassen et el 1911.09107; Huetsch et al 2404.XXXXX

Unfolding



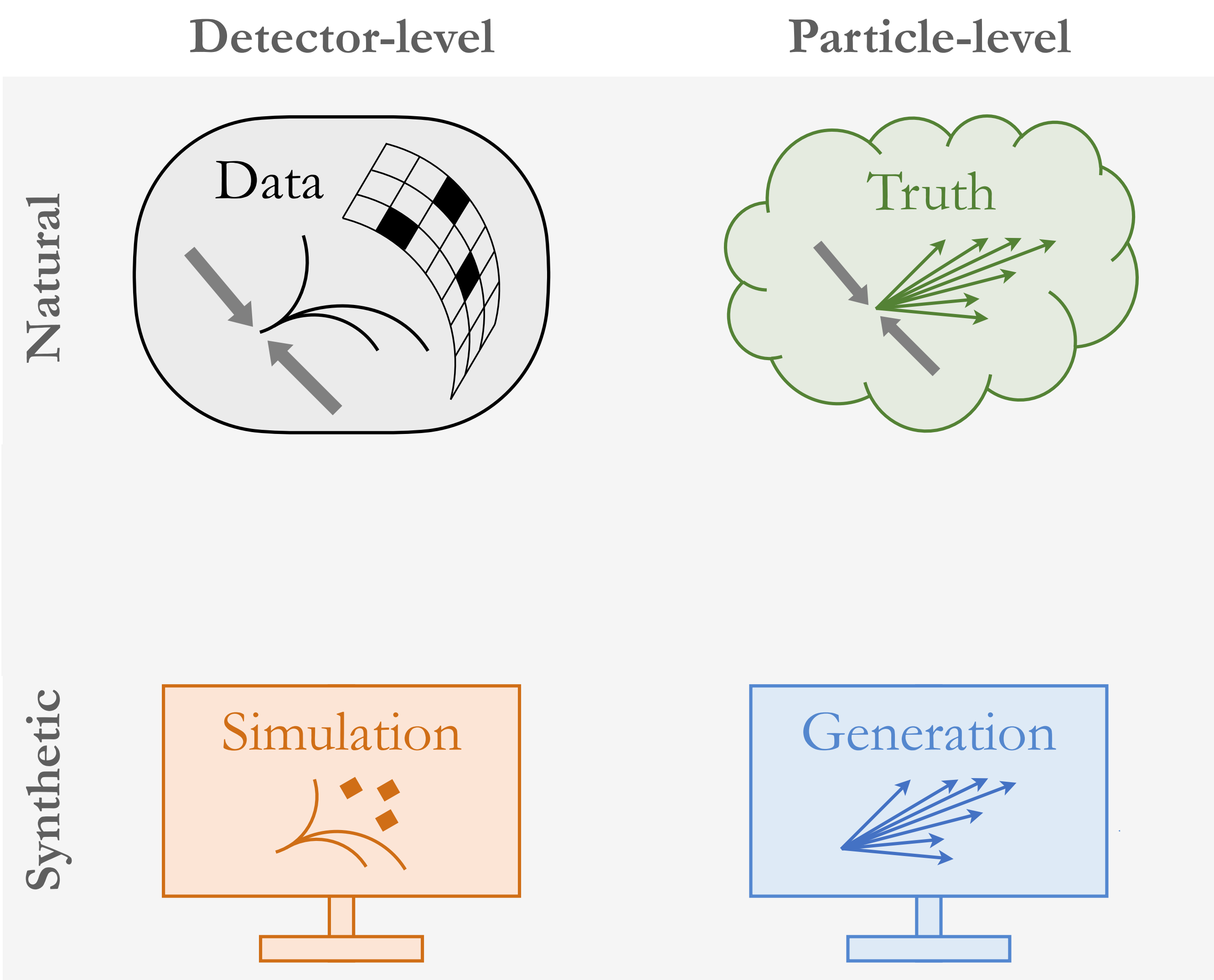
2 key approaches:

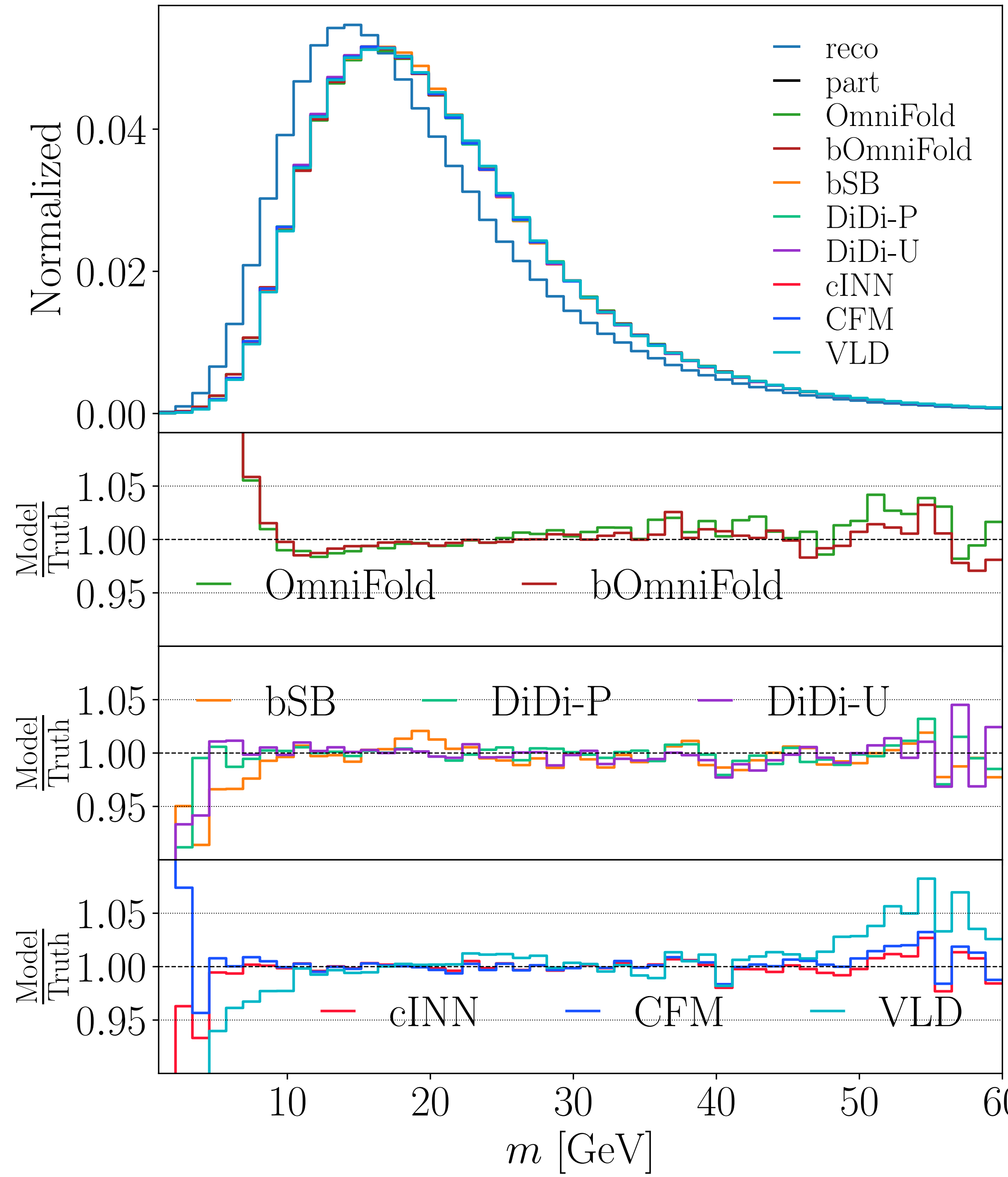
- Reweighting based on classifiers
- Morphing based on diffusion or generative models



Andreassen et el 1911.09107; Bellagente, .., GK, et al 2006.06685; Huetsch et al 2404.XXXXX

Unfolding

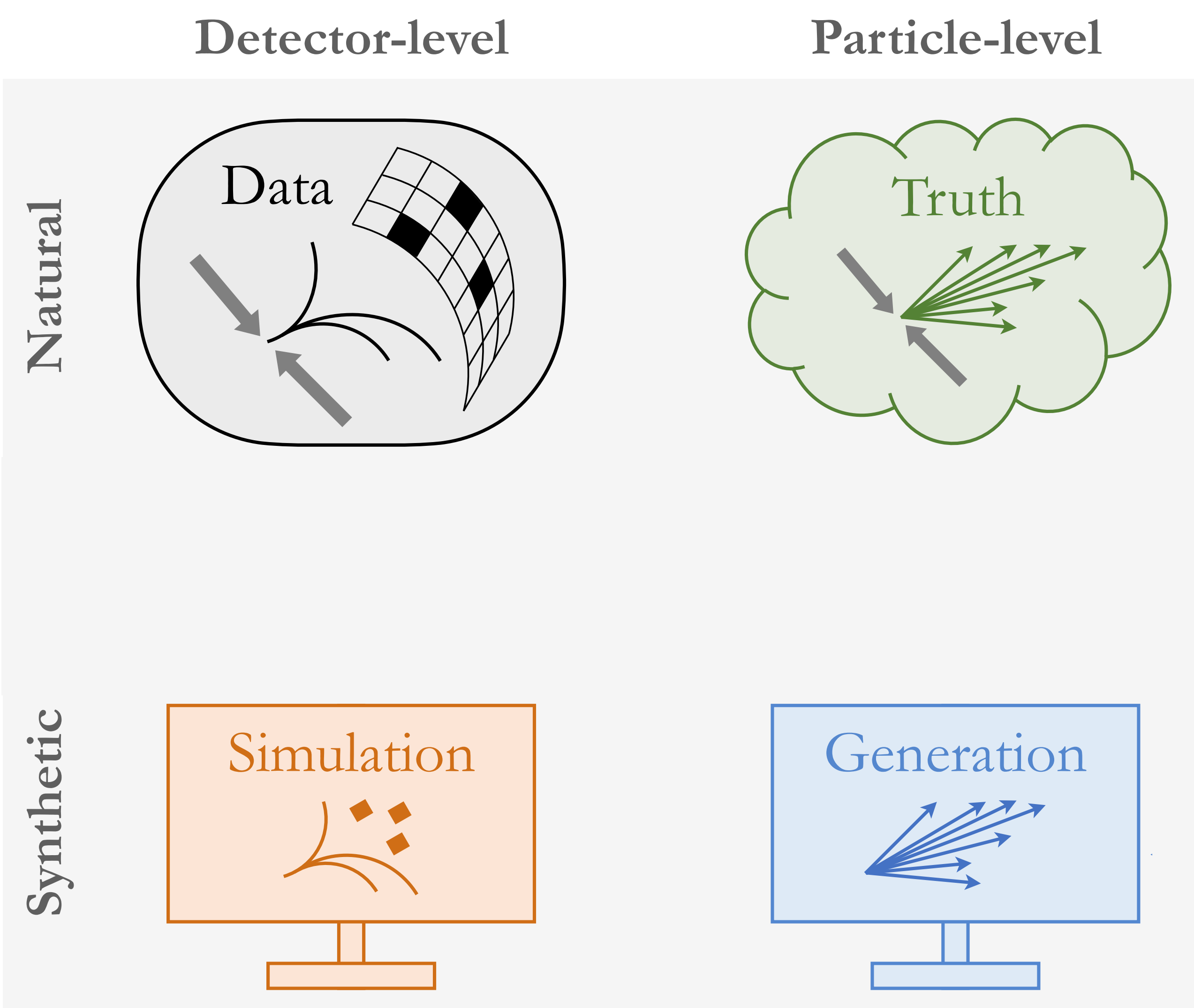


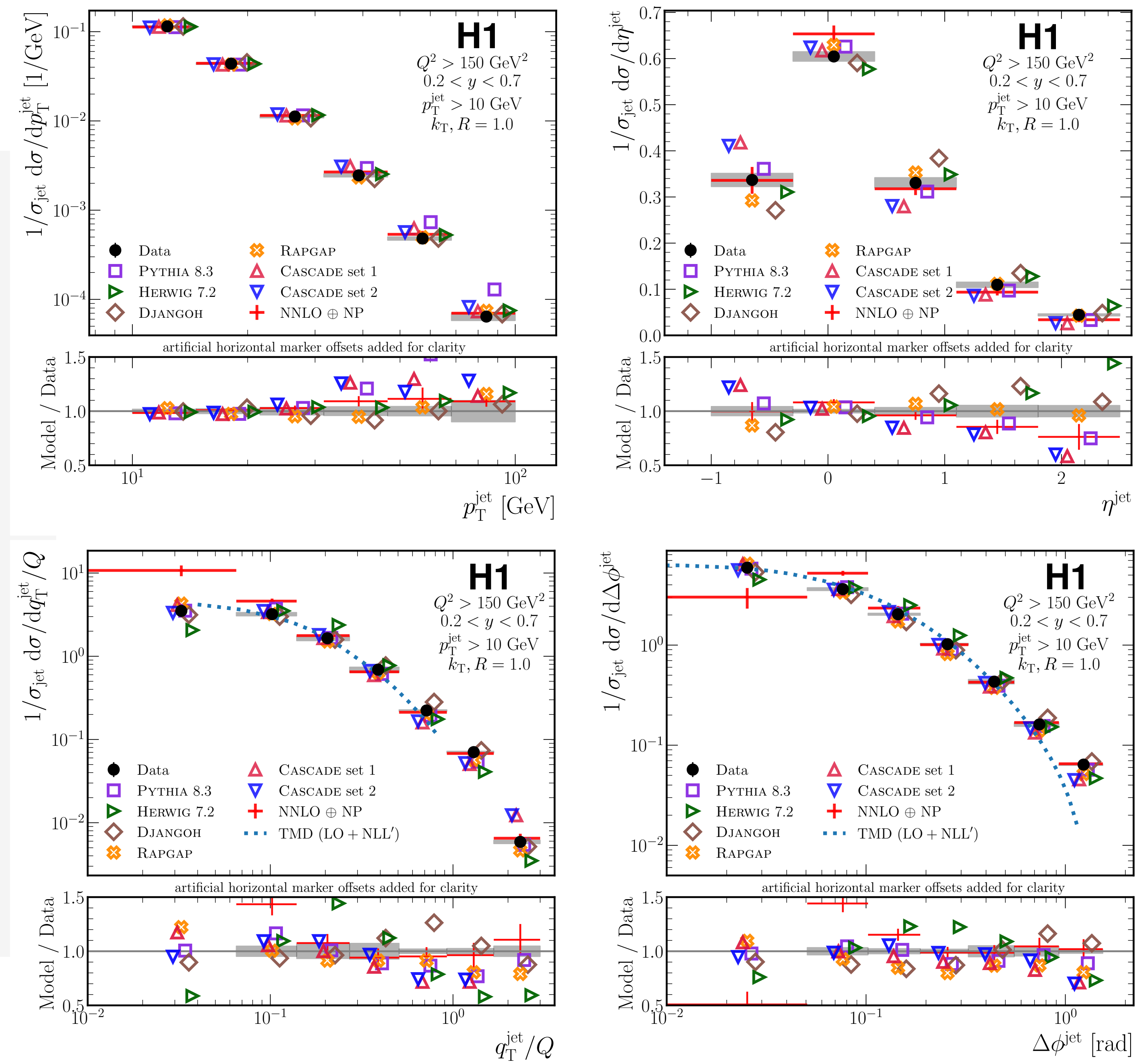


Example: Unfold Z+jets distributions in six dimensions

Andreassen et el 1911.09107; Huetsch et al 2404.XXXXX

Unfolding

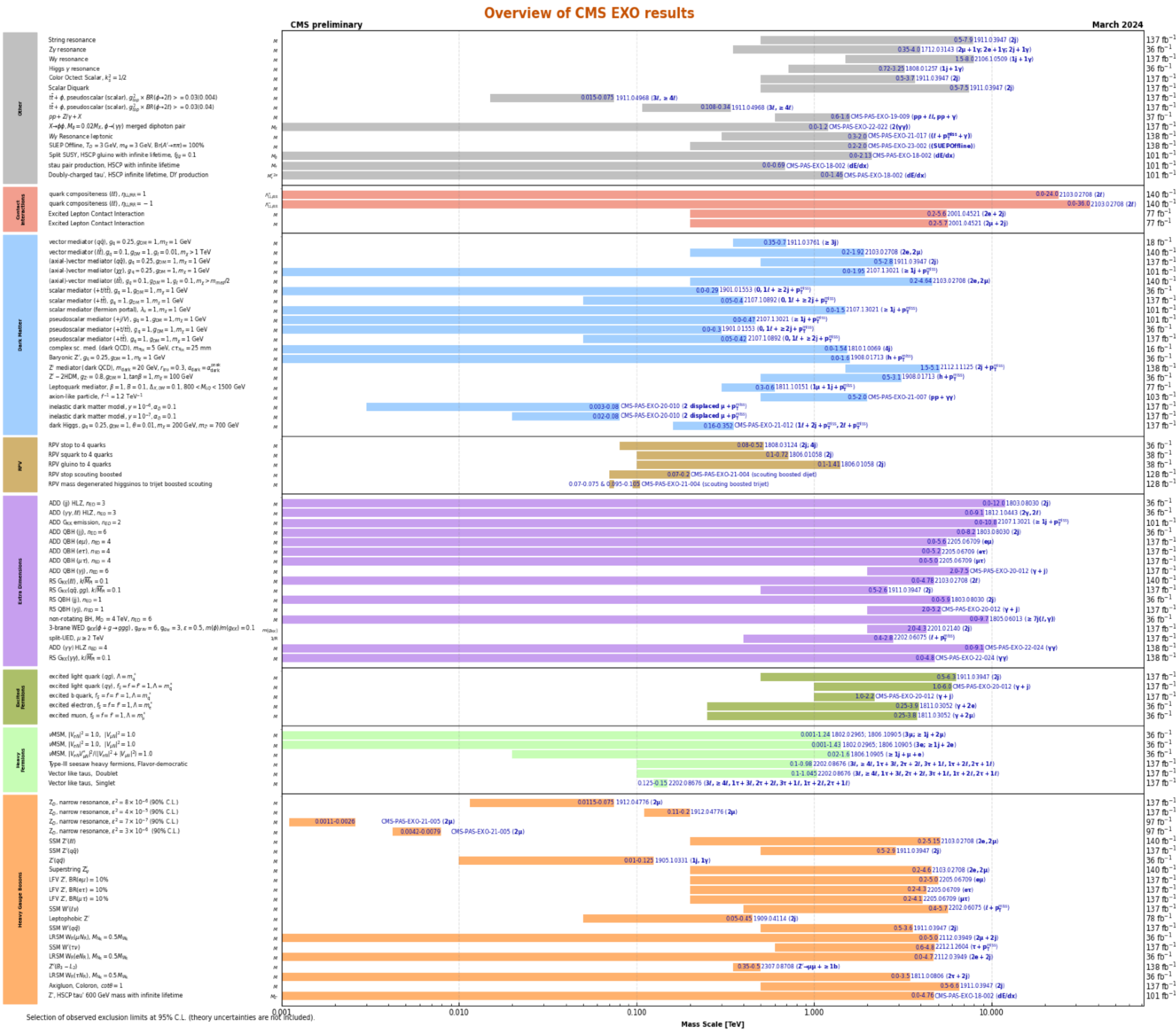




Already applied to collider data:
Multifold on lepton/jet events at H1

Andreassen et el 1911.09107; H1 2108.12376

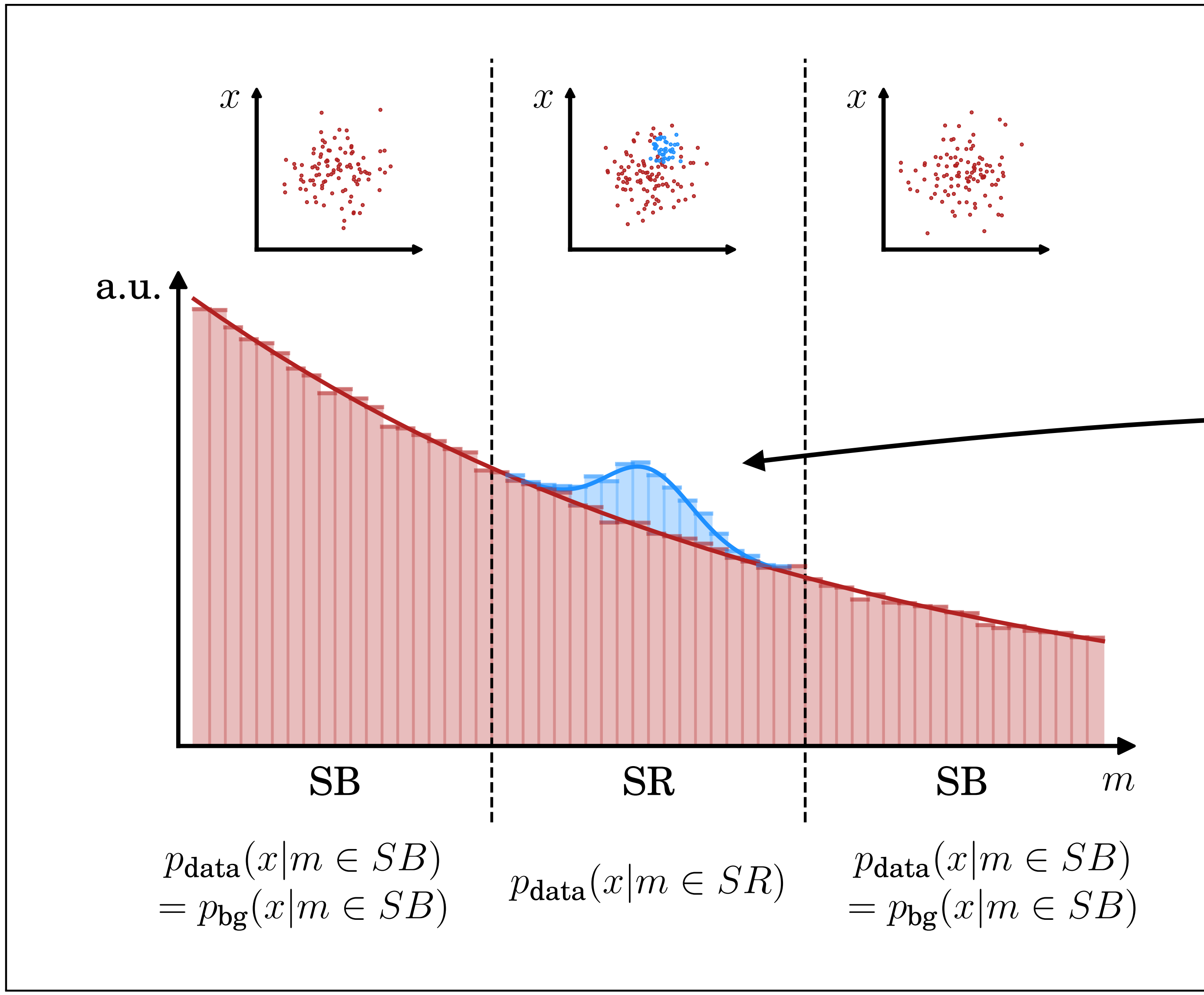
5. Anomaly Detection



Not only measurement of Standard Model: Also search for physics beyond

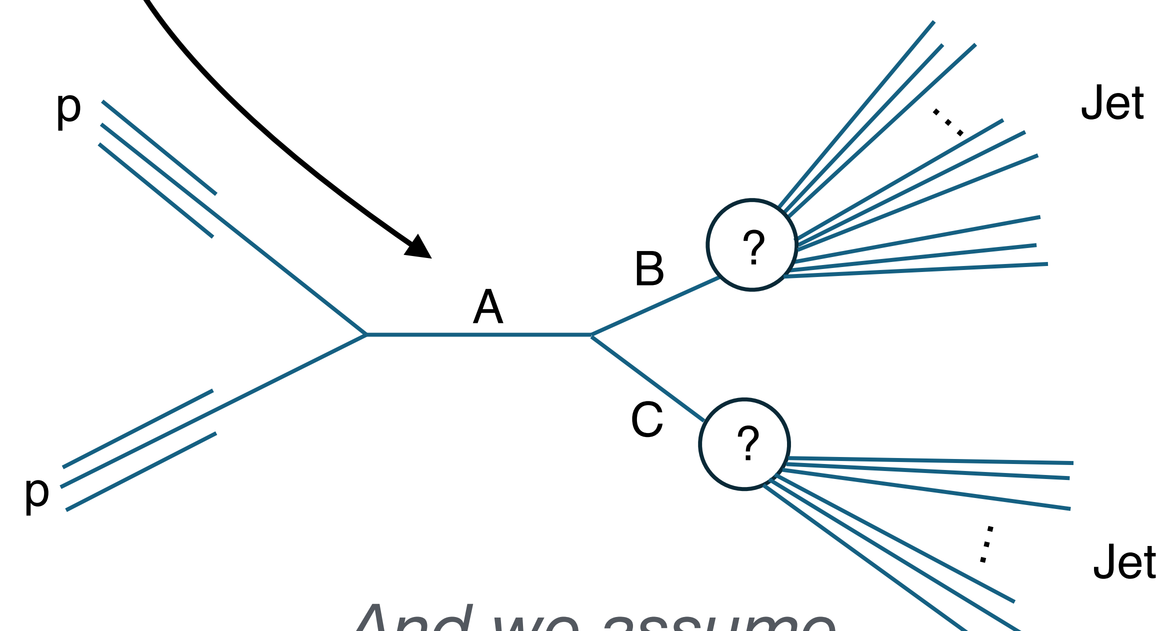
So far, no evidence of new physics in model-driven searches

CATHODE



Consider resonant anomalies: fully data-based construction of anomaly detection score

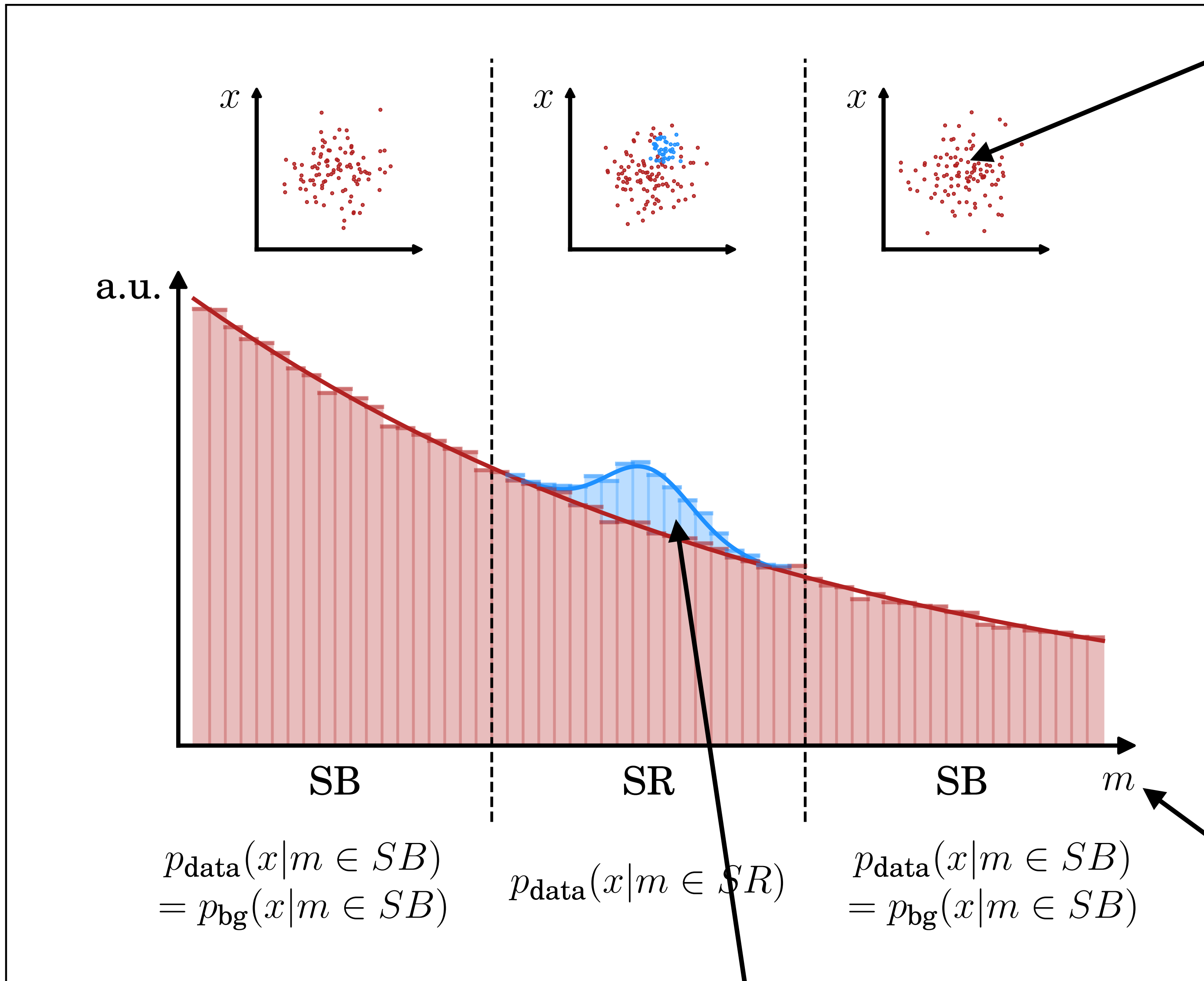
We don't assume the mass and type of the resonant particle



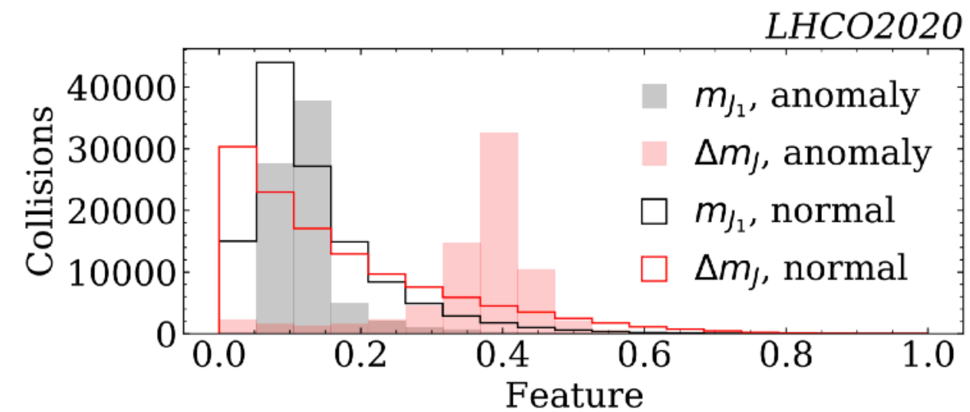
And we assume what particles it decays to

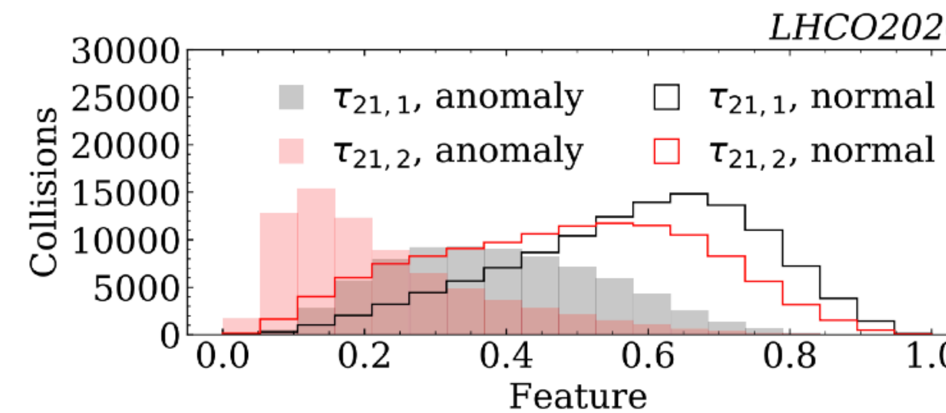
GK, Nachmann, Shih et al 2101.08320; Hallin, .., **GK** et al 2109.00546; Many similar approaches — see e.g. Golling, **GK** et al 2307.11157 for an overview

CATHODE



4 additional features



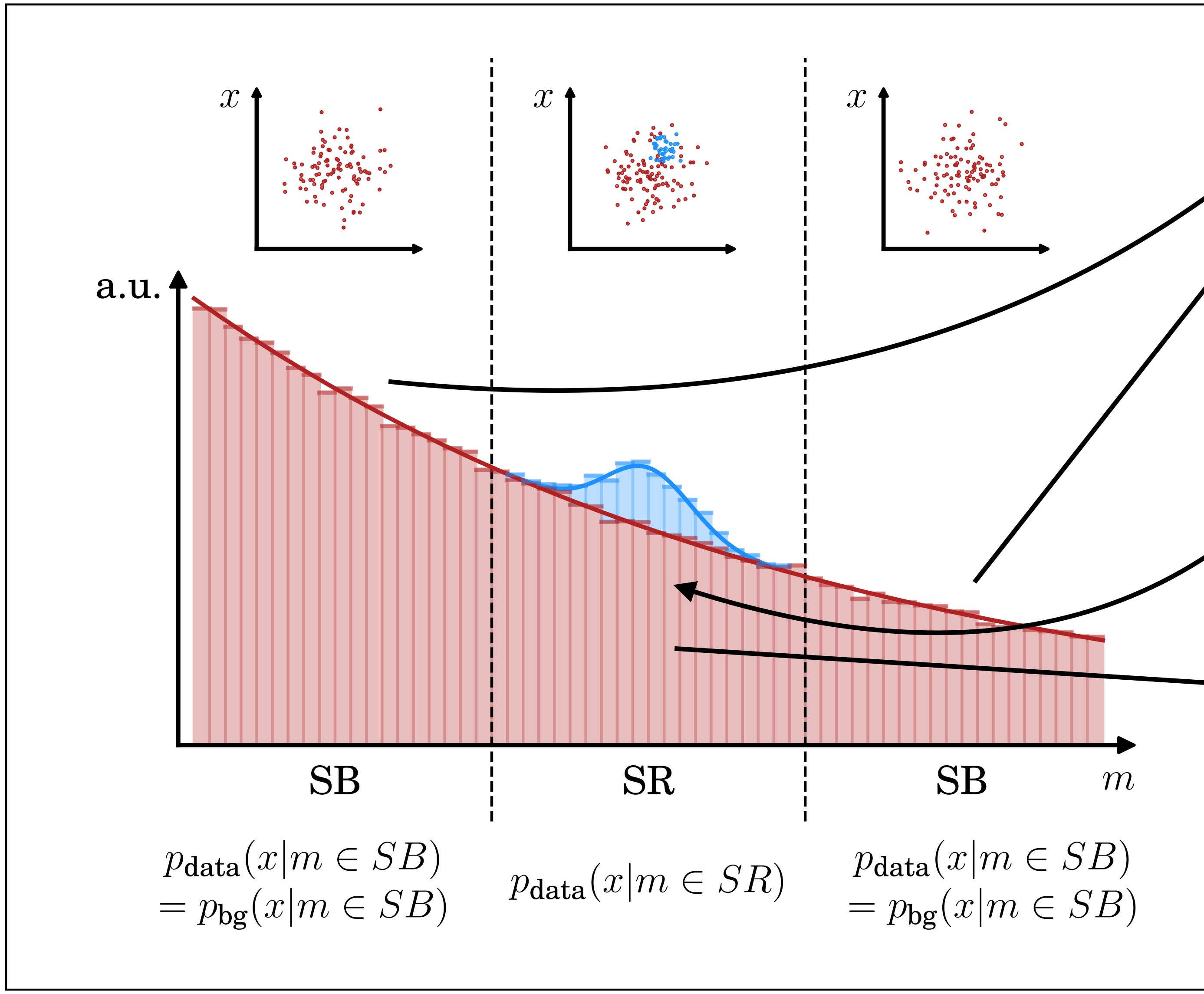


one resonant feature

Signal too small
to be visible in
inclusive distribution

GK, Nachmann, Shih et al 2101.08320; Hallin, .., **GK** et al 2109.00546;
LHC Olympics dataset: **GK**, Nachman, Shih, et al 2101.08320

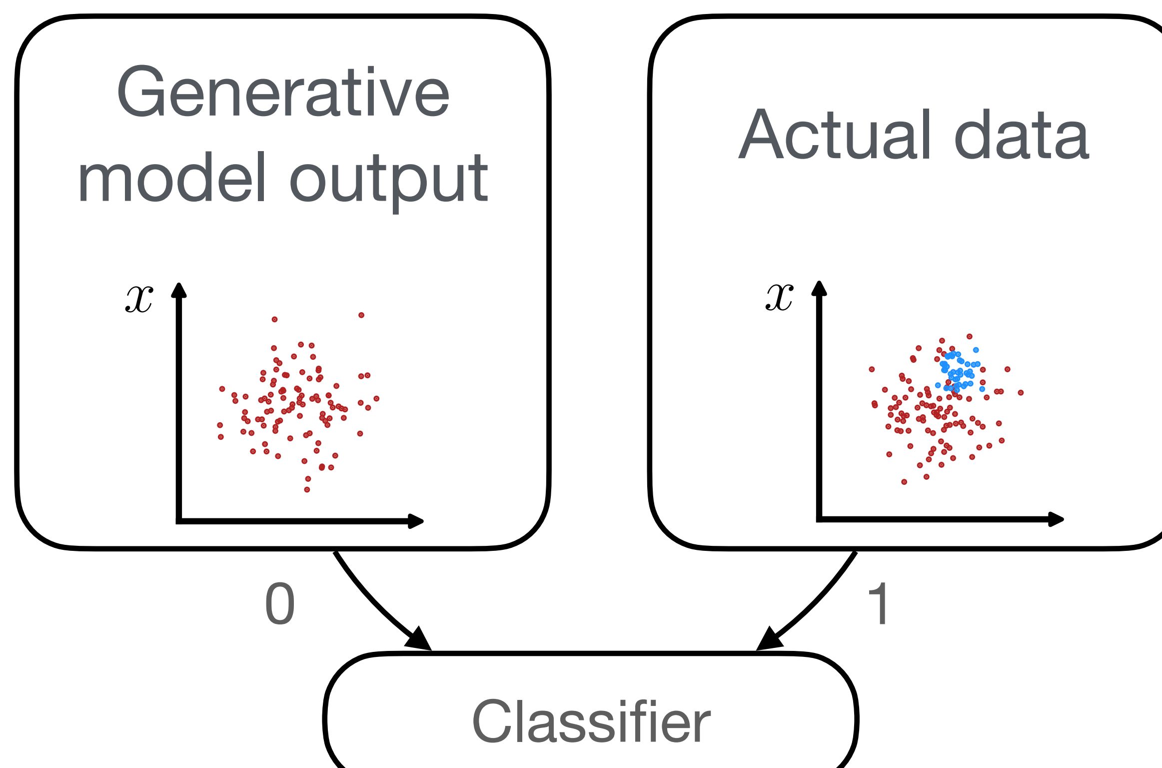
CATHODE



Train generative model conditional on m

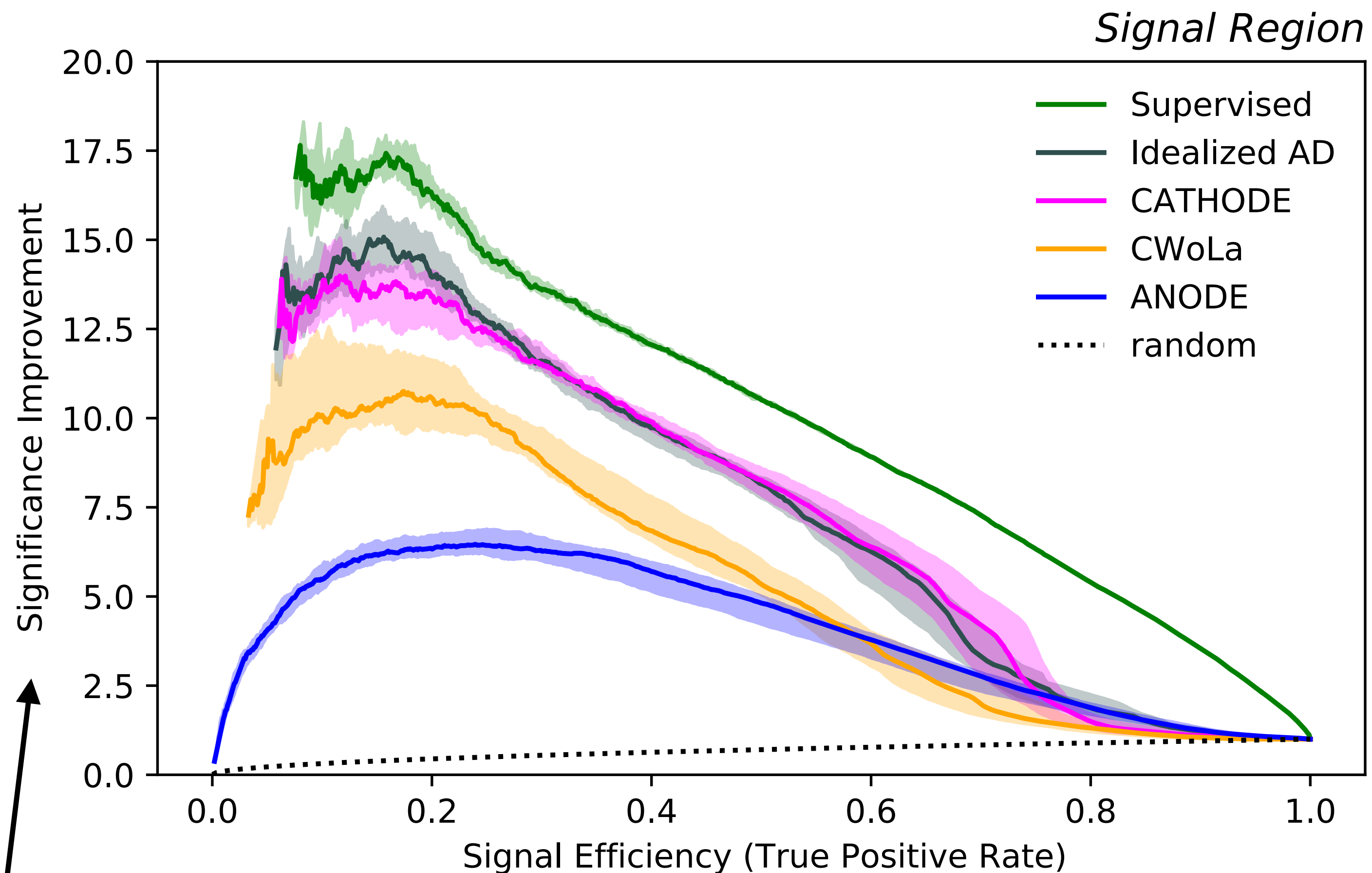
Interpolate & and sample here

Train a classifier between prediction vs data



GK, Nachmann, Shih et al 2101.08320; Hallin, ..,
GK et al 2109.00546;

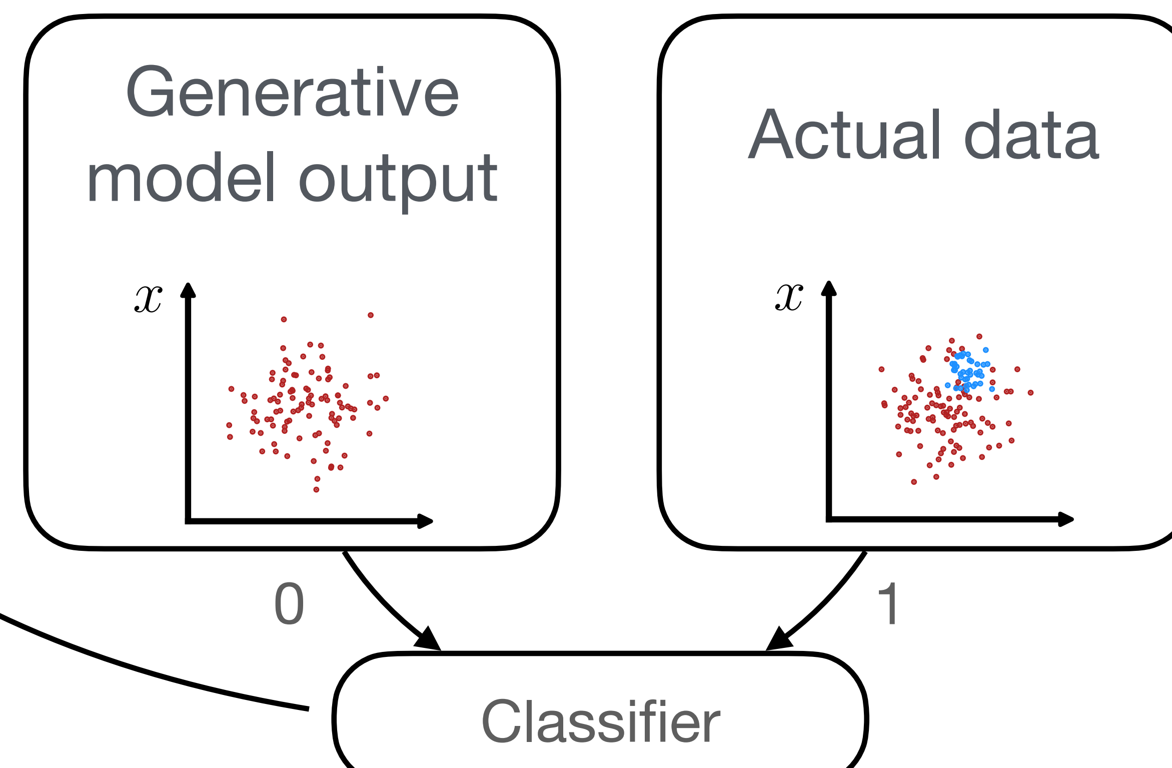
CATHODE



Use classifier to identify anomalies

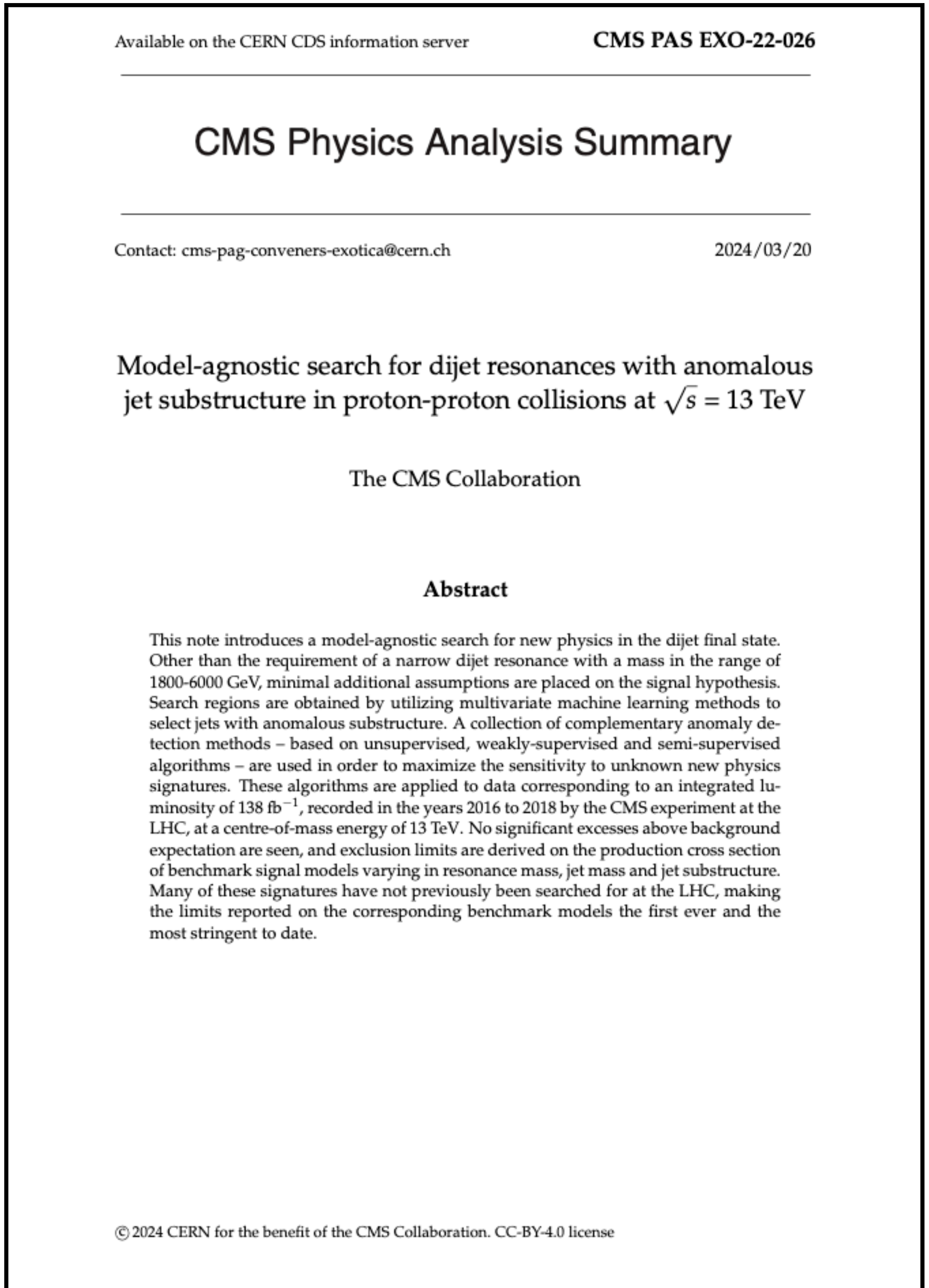
Promising, but does this work on data?

$$\mathrm{SIC} = \frac{\epsilon_S}{\sqrt{\epsilon_B}}$$



GK, Nachmann, Shih et al 2101.08320; Hallin, ..,
GK et al 2109.00546;

CASE



Available on the CERN CDS information server **CMS PAS EXO-22-026**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch 2024/03/20

Model-agnostic search for dijet resonances with anomalous jet substructure in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

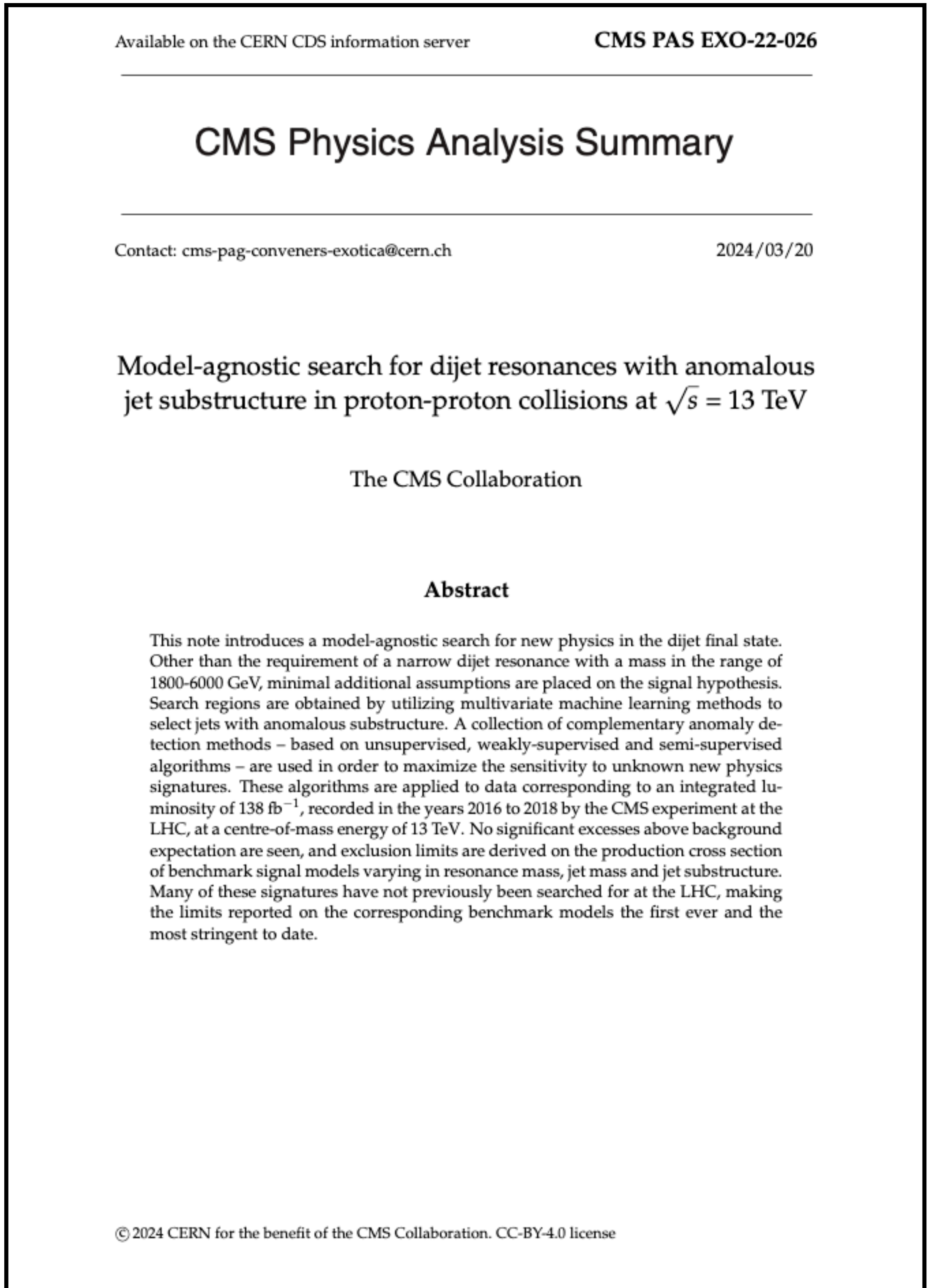
Abstract

This note introduces a model-agnostic search for new physics in the dijet final state. Other than the requirement of a narrow dijet resonance with a mass in the range of 1800-6000 GeV, minimal additional assumptions are placed on the signal hypothesis. Search regions are obtained by utilizing multivariate machine learning methods to select jets with anomalous substructure. A collection of complementary anomaly detection methods – based on unsupervised, weakly-supervised and semi-supervised algorithms – are used in order to maximize the sensitivity to unknown new physics signatures. These algorithms are applied to data corresponding to an integrated luminosity of 138 fb^{-1}, recorded in the years 2016 to 2018 by the CMS experiment at the LHC, at a centre-of-mass energy of 13 TeV. No significant excesses above background expectation are seen, and exclusion limits are derived on the production cross section of benchmark signal models varying in resonance mass, jet mass and jet substructure. Many of these signatures have not previously been searched for at the LHC, making the limits reported on the corresponding benchmark models the first ever and the most stringent to date.

- New result by the CMS collaboration: CMS Anomaly Search Effort (CASE)
- Full Run 2 dataset
- 6 anomaly detectors in parallel

CMS PAS-EXO-23-006

CASE


Available on the CERN CDS information server — CMS PAS EXO-22-026

CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch — 2024/03/20

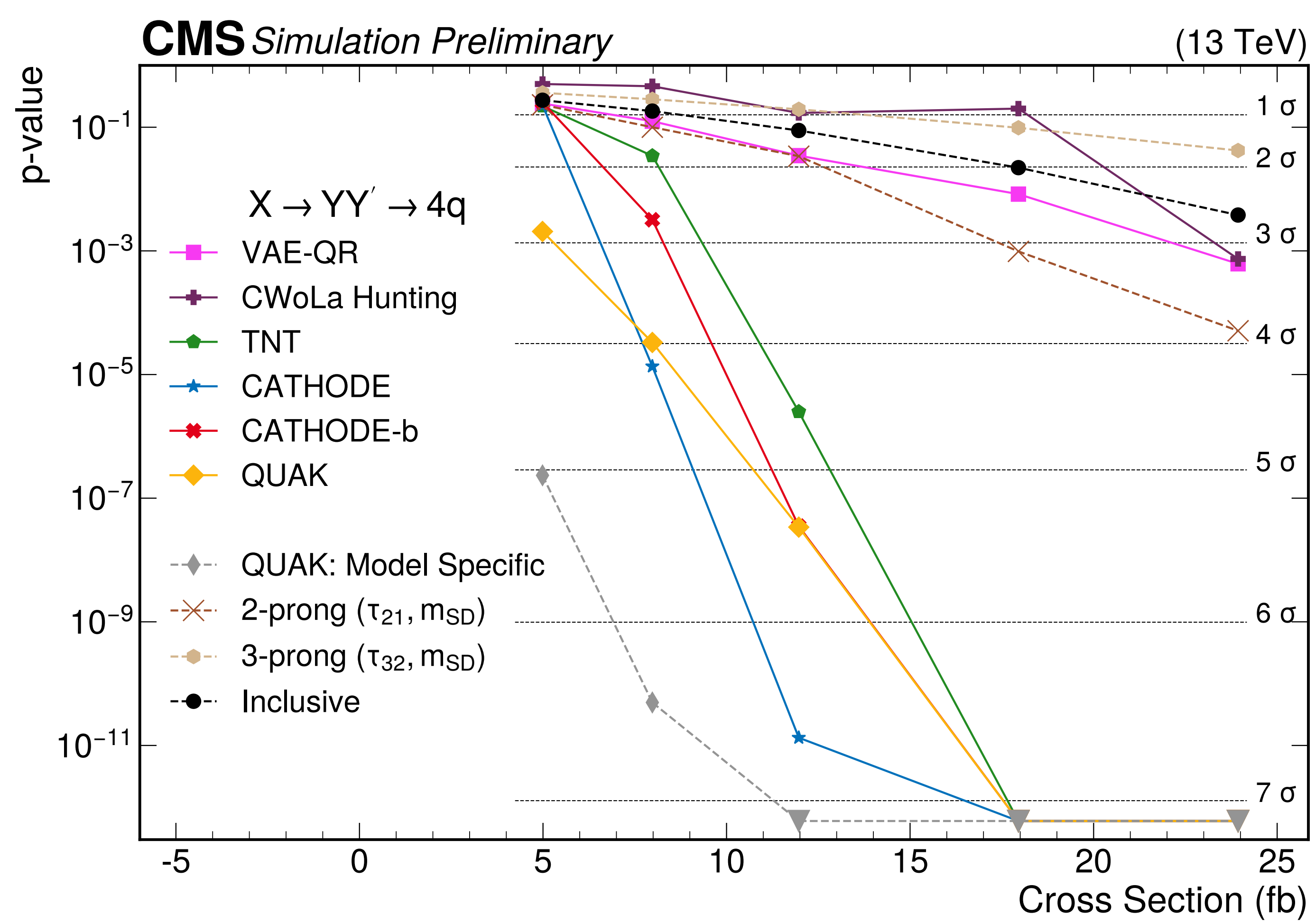
Model-agnostic search for dijet resonances with anomalous jet substructure in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

This note introduces a model-agnostic search for new physics in the dijet final state. Other than the requirement of a narrow dijet resonance with a mass in the range of 1800-6000 GeV, minimal additional assumptions are placed on the signal hypothesis. Search regions are obtained by utilizing multivariate machine learning methods to select jets with anomalous substructure. A collection of complementary anomaly detection methods – based on unsupervised, weakly-supervised and semi-supervised algorithms – are used in order to maximize the sensitivity to unknown new physics signatures. These algorithms are applied to data corresponding to an integrated luminosity of 138 fb^{-1}, recorded in the years 2016 to 2018 by the CMS experiment at the LHC, at a centre-of-mass energy of 13 TeV. No significant excesses above background expectation are seen, and exclusion limits are derived on the production cross section of benchmark signal models varying in resonance mass, jet mass and jet substructure. Many of these signatures have not previously been searched for at the LHC, making the limits reported on the corresponding benchmark models the first ever and the most stringent to date.

- New result by the CMS collaboration: CMS Anomaly Search Effort (CASE)
- Full Run 2 dataset
- 6 anomaly detectors in parallel

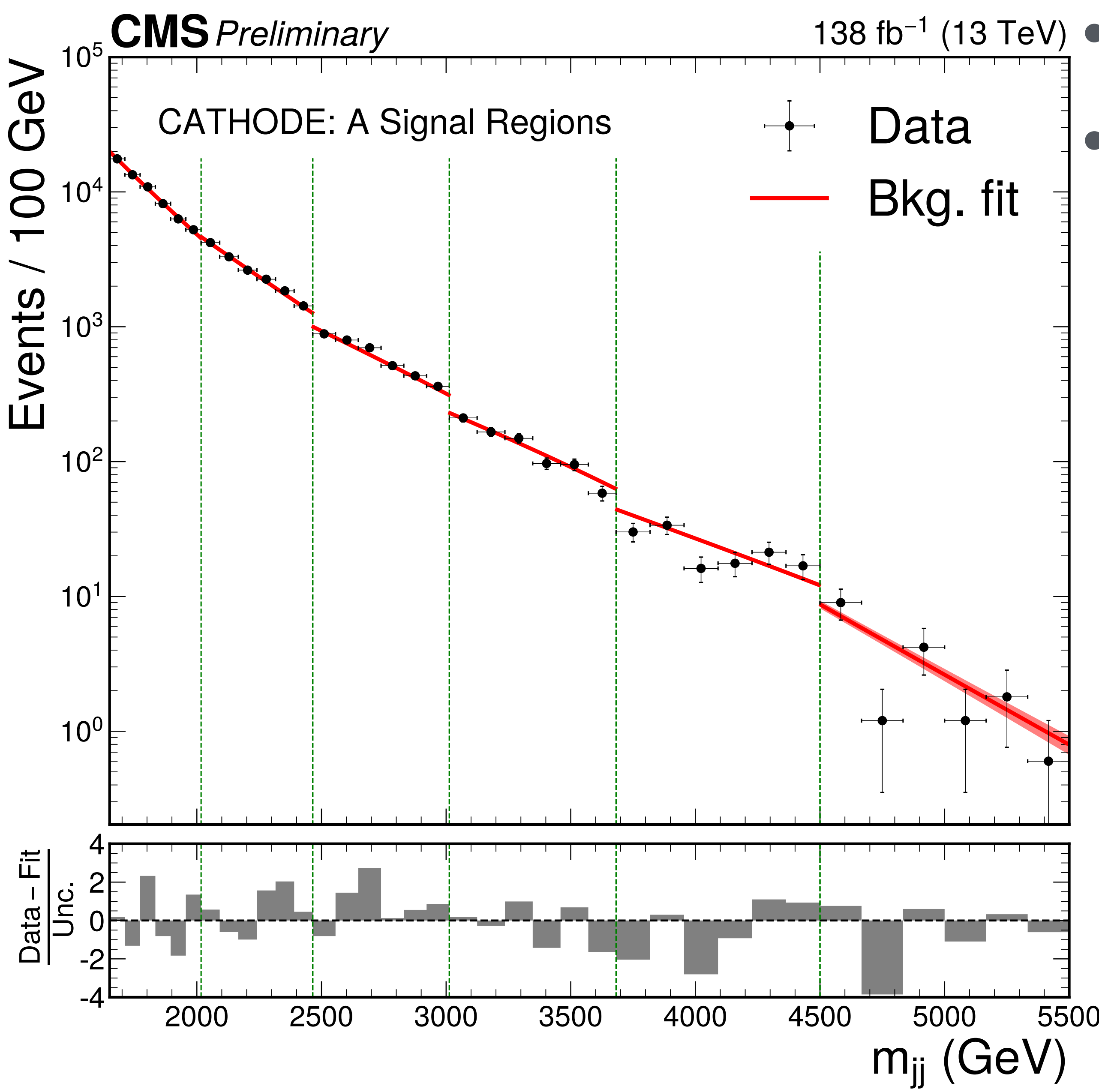


Test expected sensitivity gain via injected signals in simulation

CMS PAS-EXO-23-006

CASE

- New result by the CMS collaboration: CMS Anomaly Search Effort (CASE)
- Full Run 2 dataset
- 6 anomaly detectors in parallel



Fully train CATHODE on data

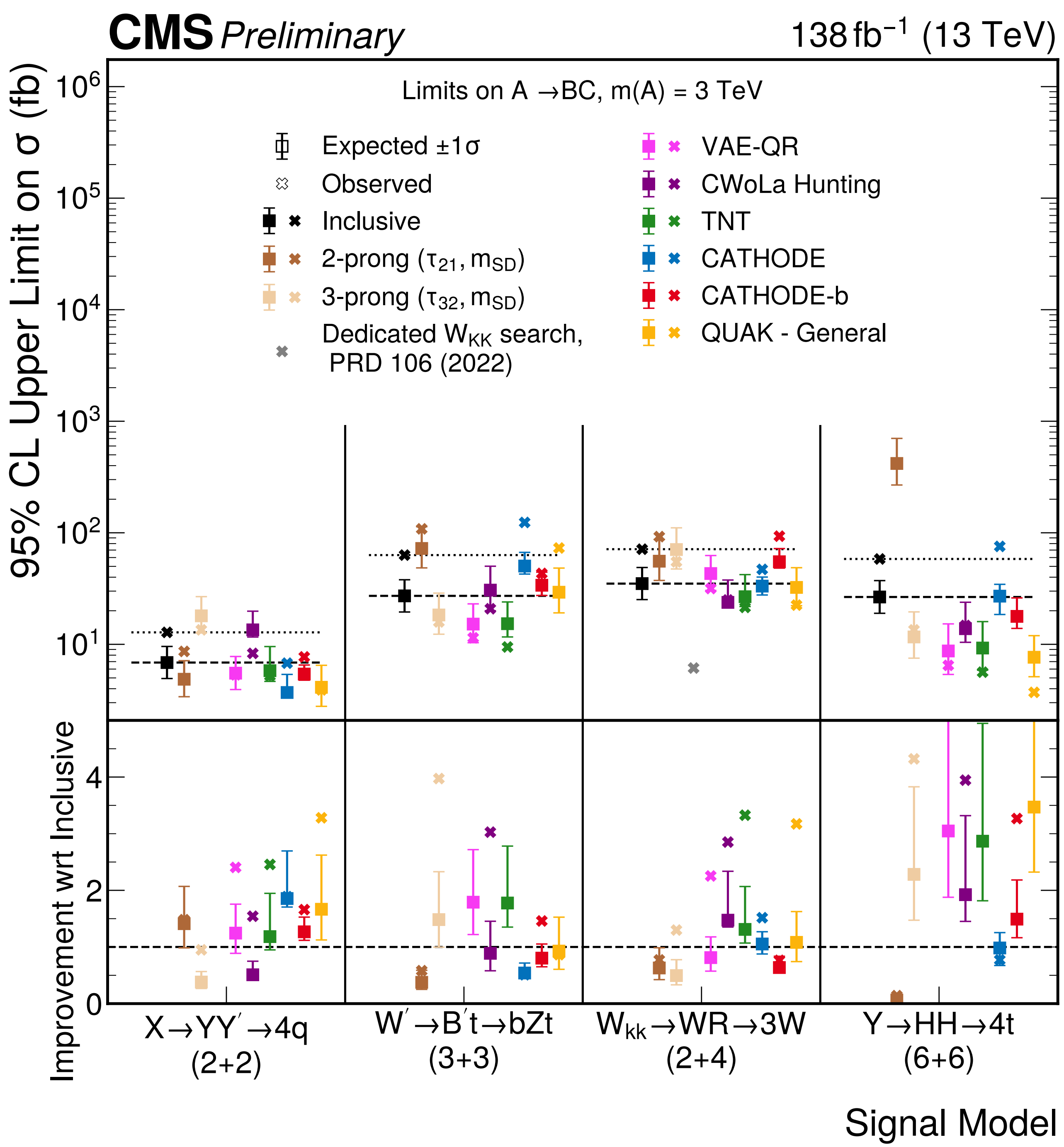
Select top 1% most anomalous events, perform bump-hunt

No signal-like outlier: set limits

CMS PAS-EXO-23-006

CASE

- New result by the CMS collaboration: CMS Anomaly Search Effort (CASE)
- Full Run 2 dataset
- 6 anomaly detectors in parallel



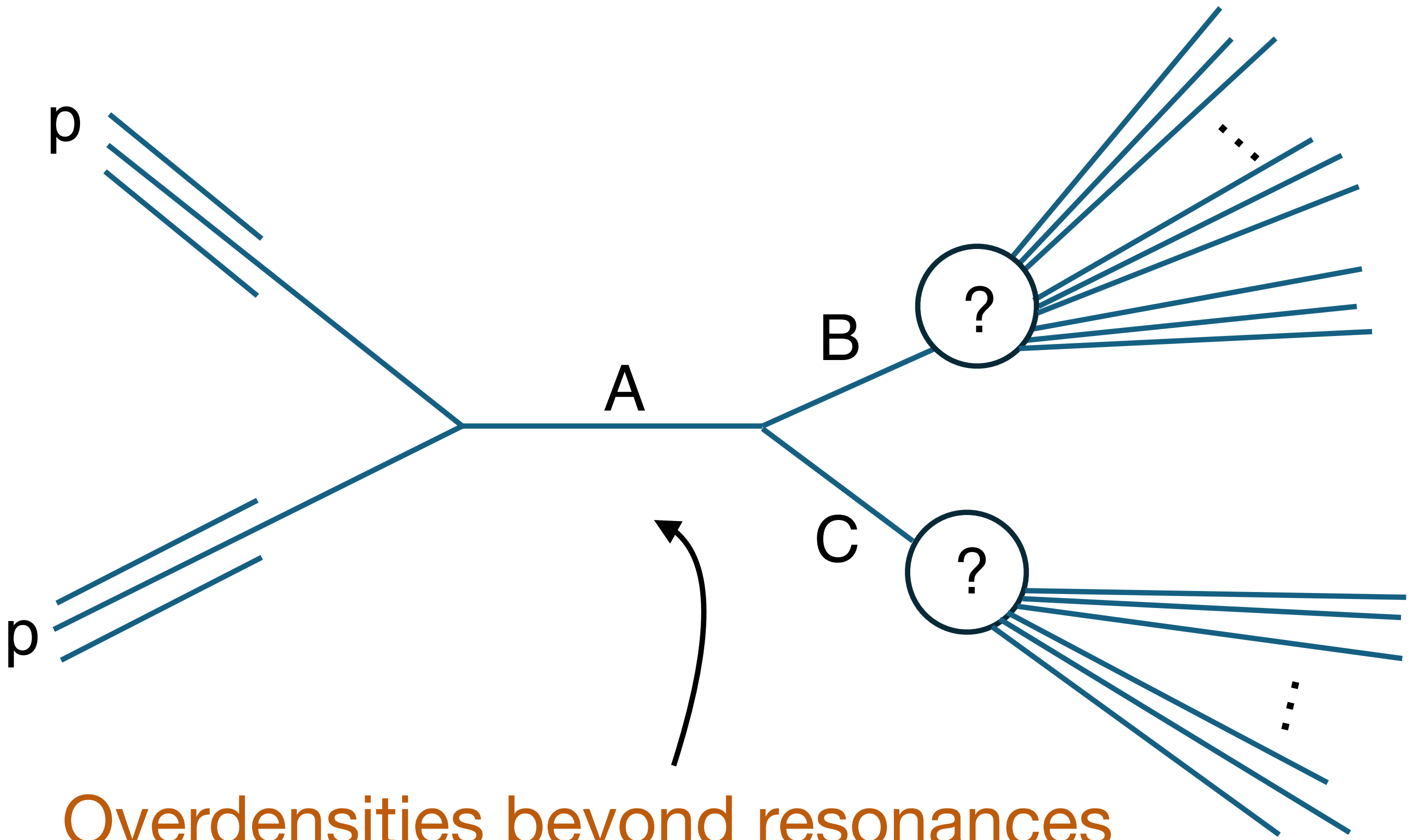
For limits: inject potential signals

Includes uncertainties, e.g. multi-prong jets modelling

Altogether set limits for 43 different signal scenarios

CMS PAS-EXO-23-006

Other developments



More features per jet
(e.g. 2309.13111)

Low-level input data
(e.g. 2310.06897)

Overdensities beyond resonances
(e.g. 2404.07258, 2311.12924)

Better sensitivity for weak signals (e.g. 2312.11629)

Reduce shaping of distributions

More topologies

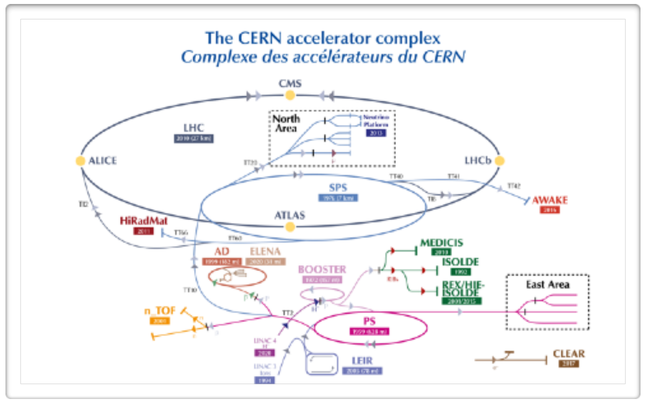
Anomalies as outliers (e.g. substantial literature on auto encoder based methods)

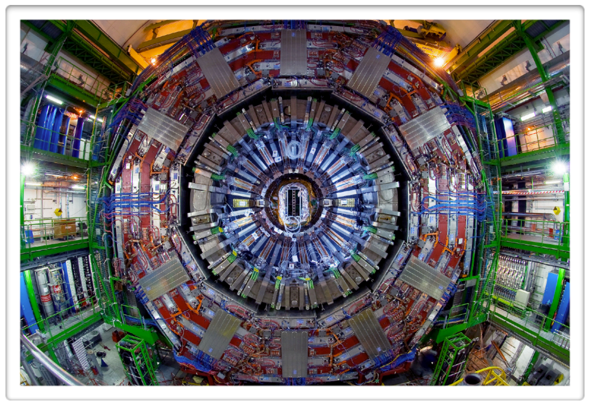
Robust statistical treatment beyond bump-hunts (e.g. 2111.13633)

Applications to data monitoring

6. Triggers

Trigger

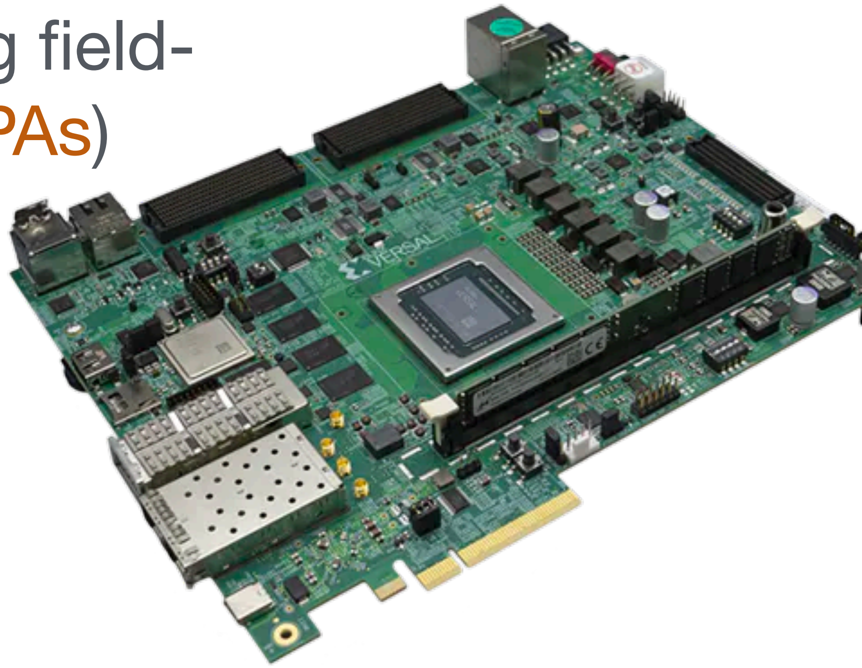




Colliders with
40 million events/second

2 stage system (Trigger) reduces this to ~1 kHz for offline storage and analysis

Stage 1: Hardware based, using field-programmable gate arrays (FGPAs) with microsecond latency



Improving selection criteria in trigger with AI yields better offline data

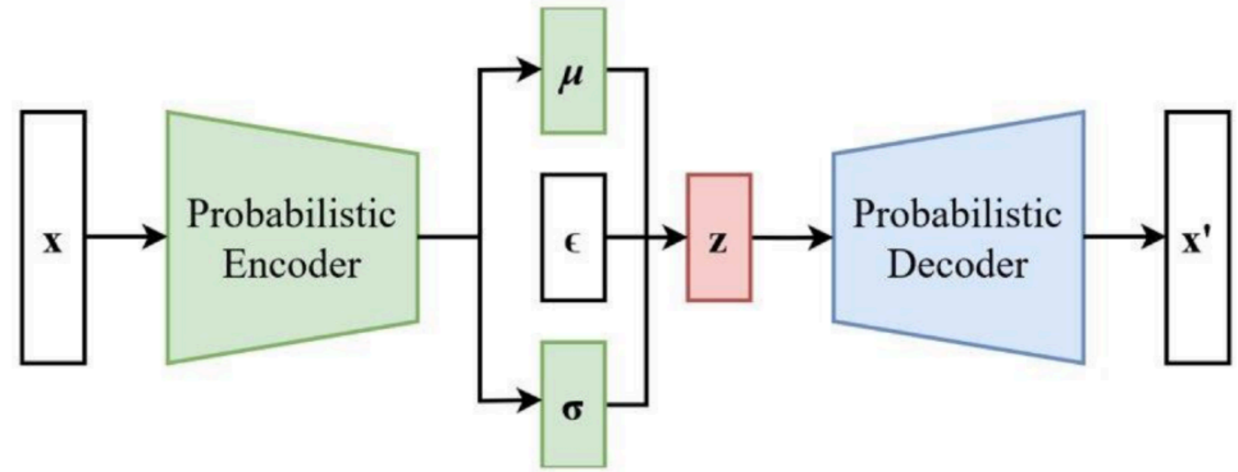
hls4ml to translate ML architectures to hardware language

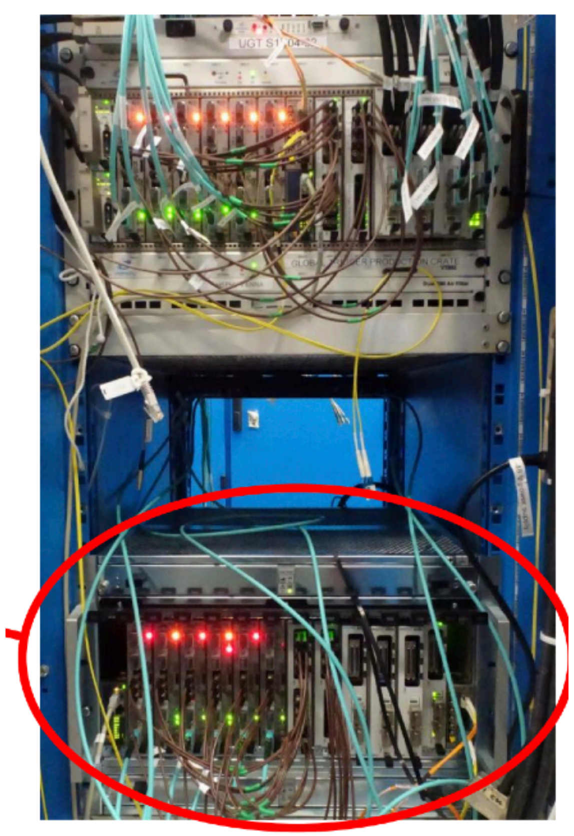
https://fastmachinelearning.org/

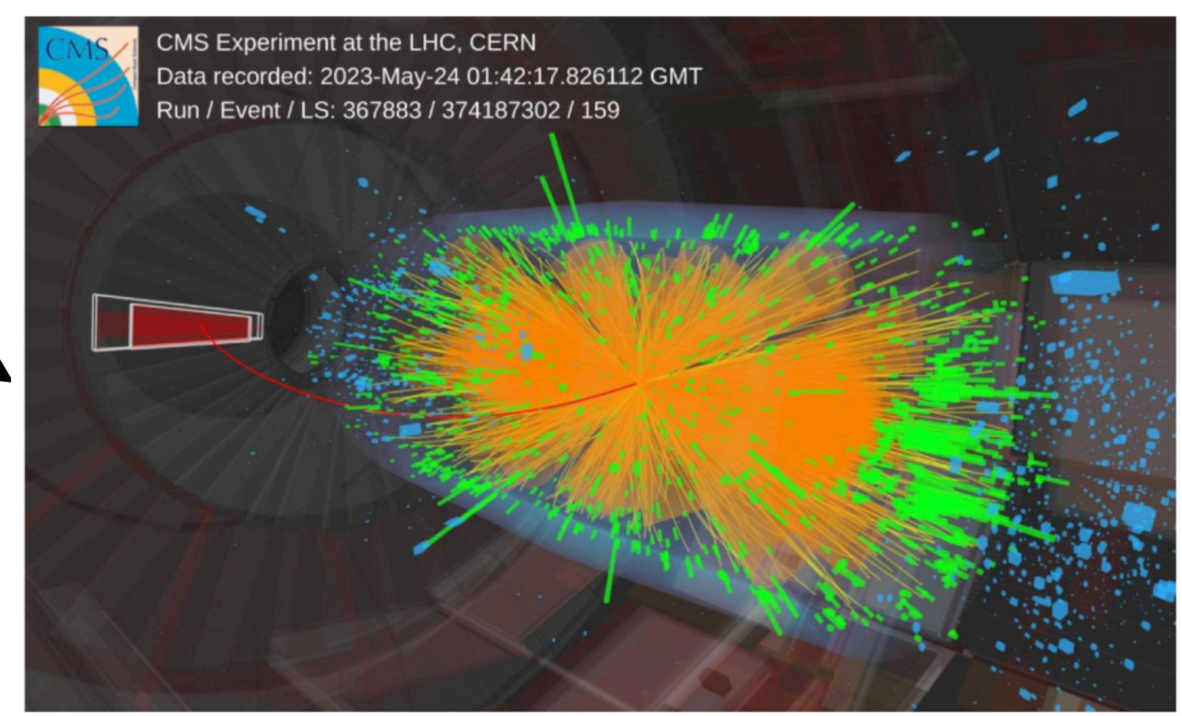
Example: Triggering Outliers

Learn-compression/decompression on signal free sample and use as anomaly score

Now testing in CMS Level 1 trigger

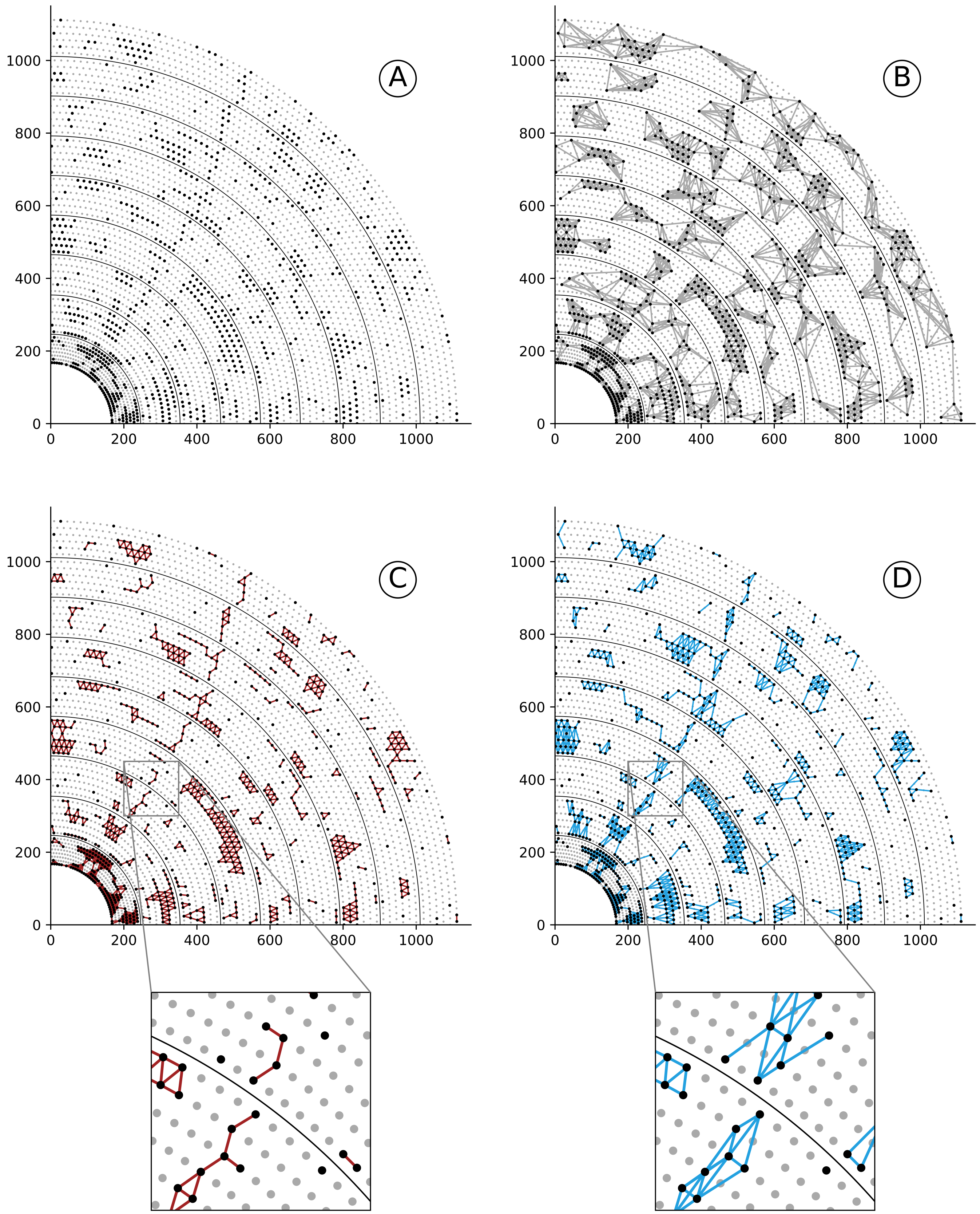






https://indico.cern.ch/event/1283970/contributions/5554350/attachments/2720710/4727877/axol1tl_fastml.pdf

Example: Online graph building



Online graph building for reconstruction in Belle 2 drift chamber

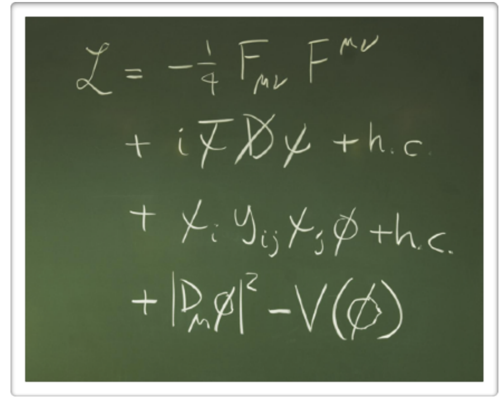
Explore different methods of constructing graphs for GNN processing

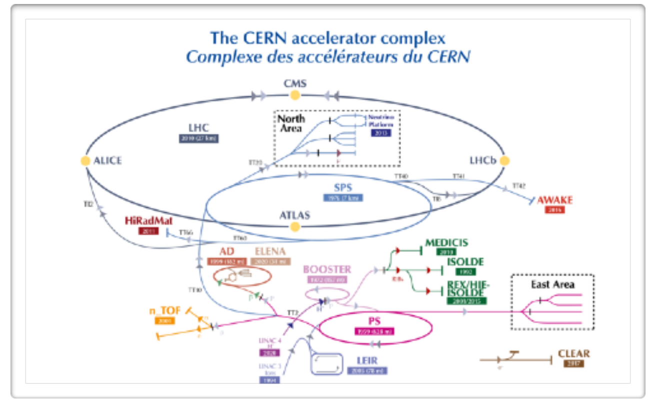
Within resource constraints

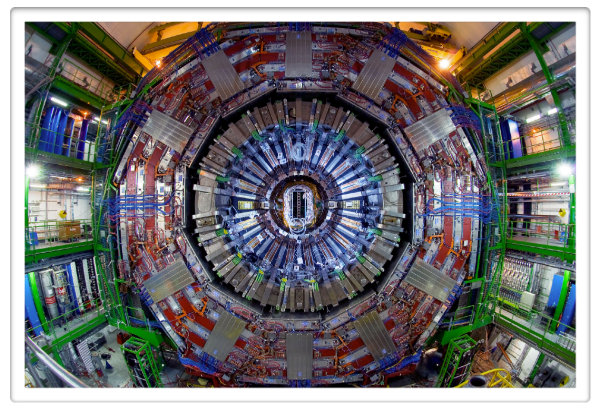
No. of Vertices	No. of Edges	Width of Edge	Registers		LUTs		F7Muxes	
			abs.	%	abs.	%	abs.	%
498	2305	60 bit	145 333	7.84 %	19 370	2.09 %	5760	1.15 %
786	3649	60 bit	246 511	13.30 %	31 360	3.39 %	9120	1.81 %
978	4545	40 bit	206 573	11.15 %	34 252	3.70 %	11 360	2.26 %
978	4545	60 bit	304 919	16.46 %	38 968	4.21 %	11 360	2.26 %
978	4545	100 bit	485 473	26.20 %	47 642	5.14 %	11 200	2.23 %

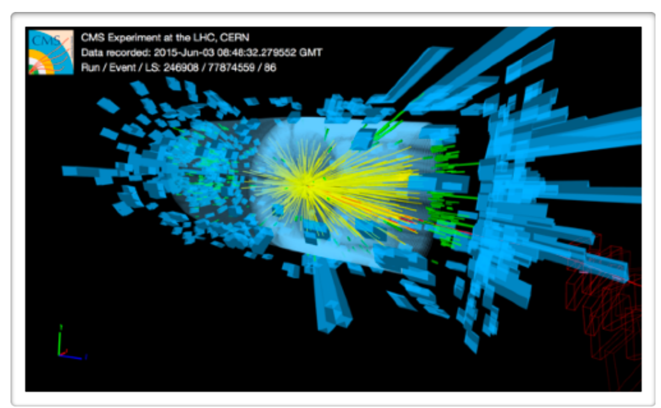
Neu et al 2307.07289

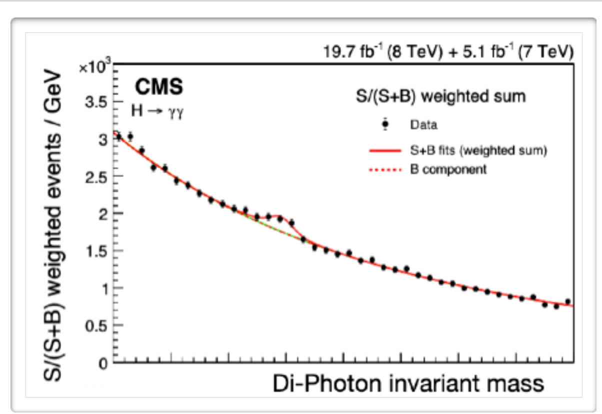
Differentiable versions of all steps in the particle physics processing chain







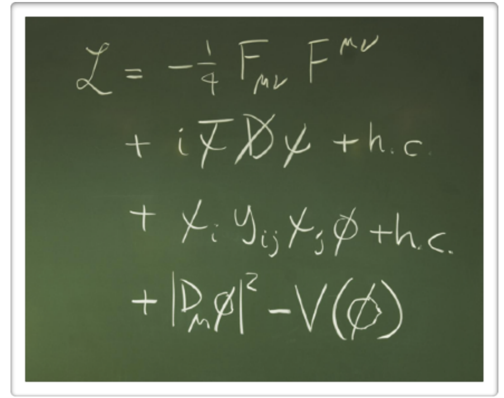


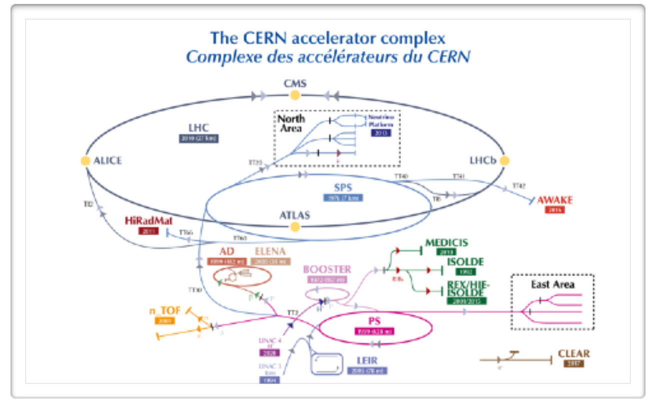


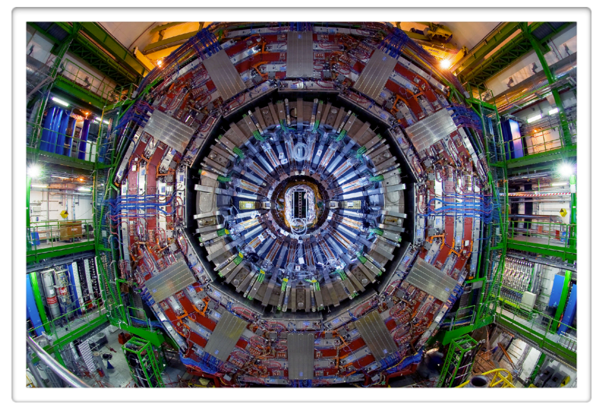
Differentiable versions of all steps in the particle physics processing chain

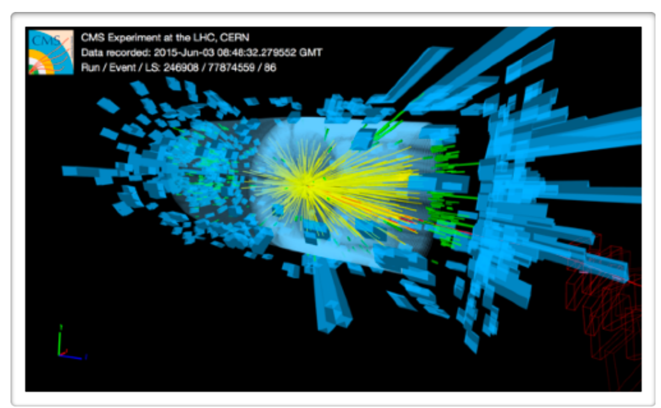
Either as ML-based surrogate models

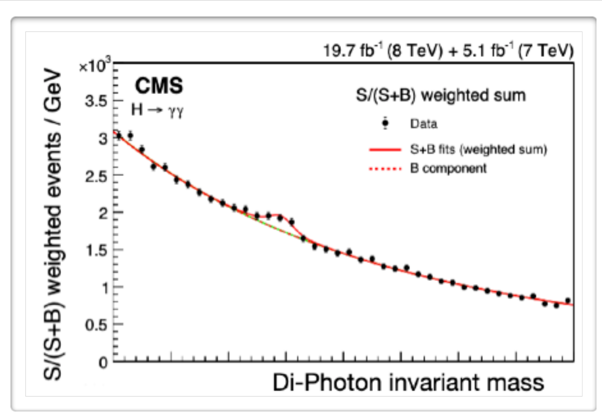
Or via e.g. differentiable programming











Heinrich, Kagan 2308.16680

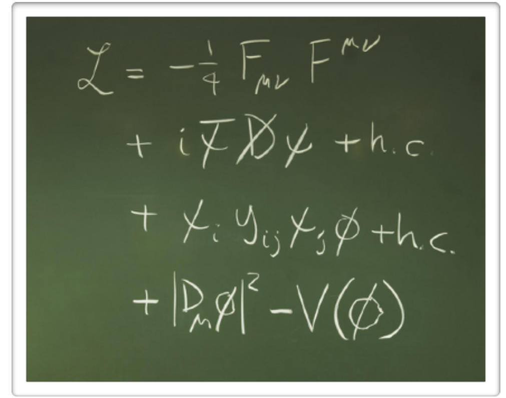
Differentiable versions of all steps in the particle physics processing chain

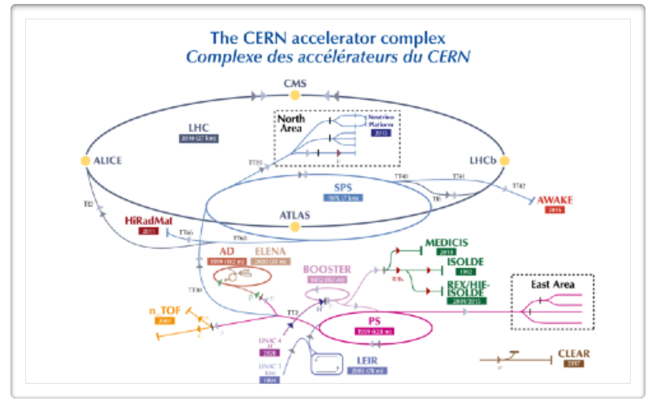
Either as ML-based surrogate models

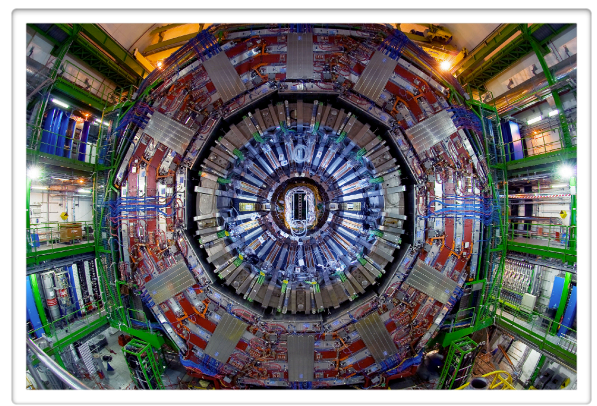
Or via e.g. differentiable programming

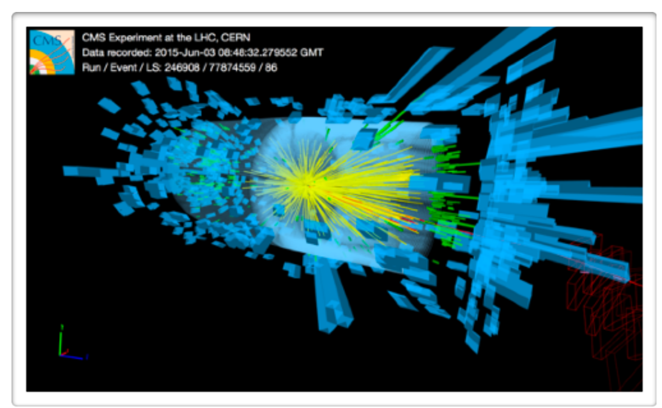
What can we do with this?

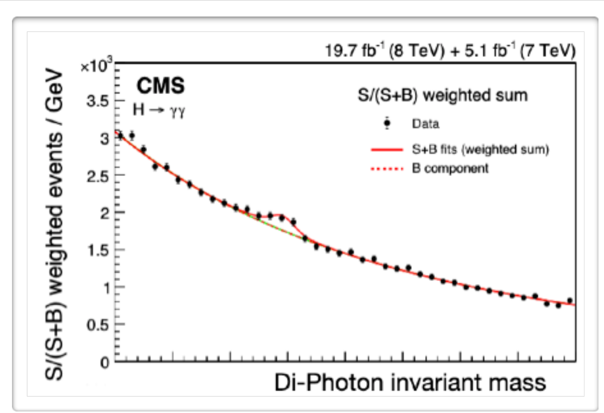
Heinrich, Kagan 2308.16680





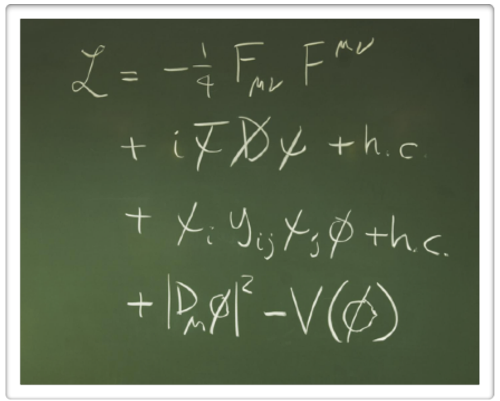


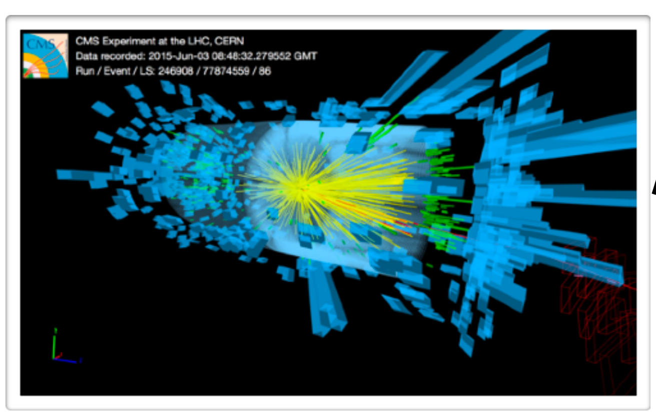




7. Inference

Inference





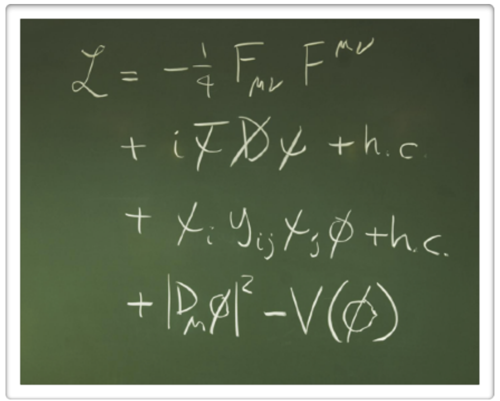
Inference

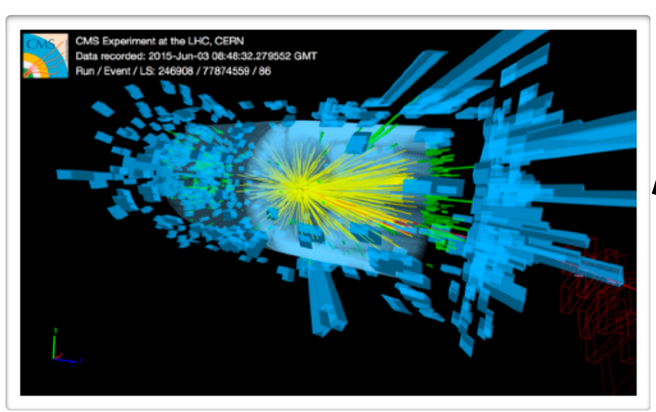
Goal: Learn parameters of theory (e.g. couplings) directly from high-dimensional data

No exact likelihood, but forward simulations available: likelihood-free / simulation based inference

Cranmer, Brehmer, Louppe 1911.01429

Inference

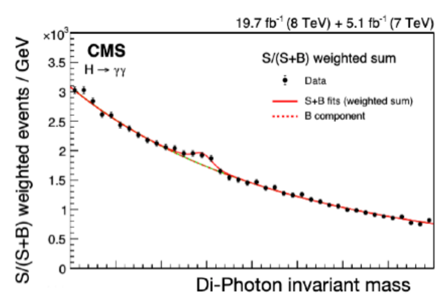




Goal: Learn parameters of theory (e.g. couplings) directly from high-dimensional data

No exact likelihood, but forward simulations available: likelihood-free / simulation based inference

Inference



Summary Statistics

Likelihood Learning (e.g. flows or cINNs)

Likelihood ratio trick (e.g. CARL, swyft)

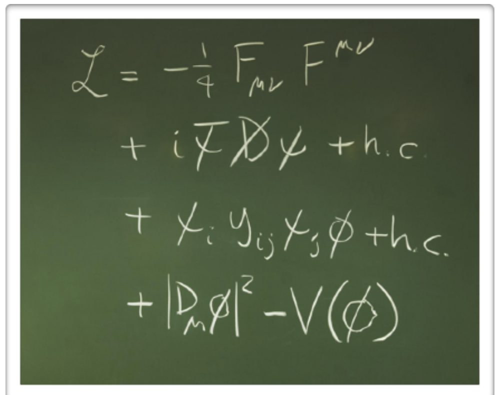
Integration (e.g. MadMiner)

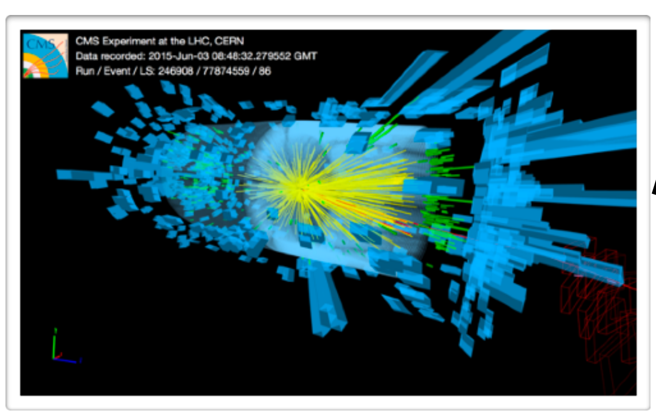
Cranmer, Brehmer, Louppe 1911.01429; Cranmer, Brehmer 2010.06439

Example

Reconstructing axion-like particles from beam dumps using cINN approach

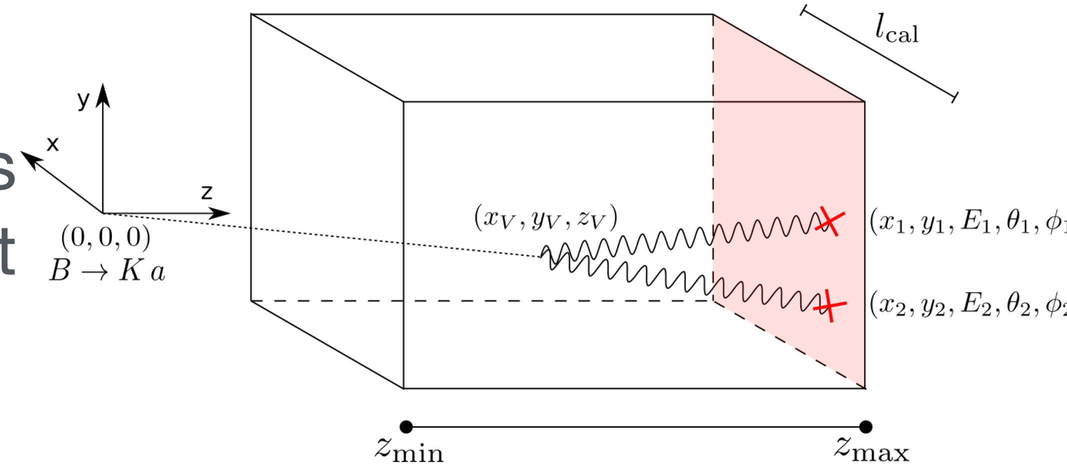
SHiP
Search for Hidden Particles

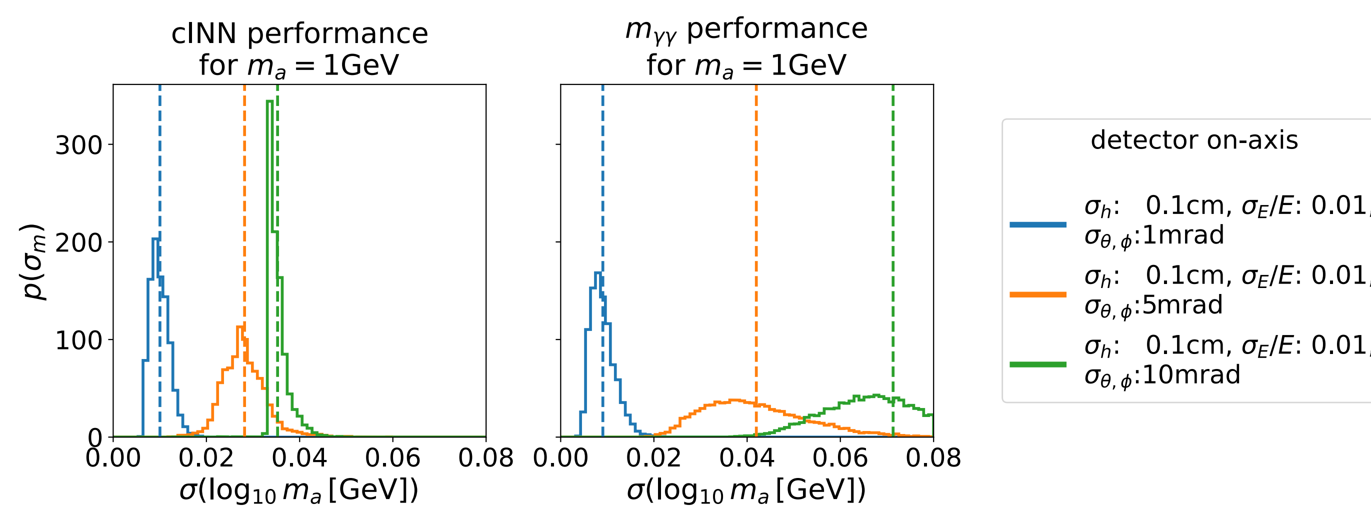




Inference

Infer axion mass from measurement



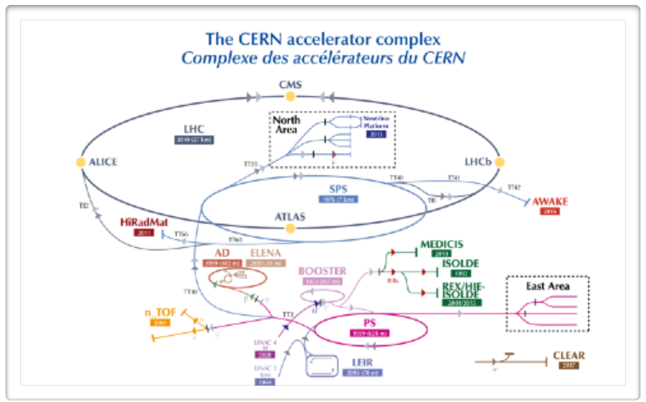


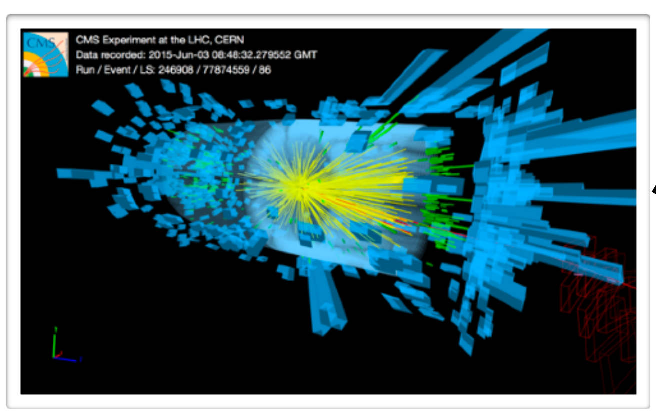
More stable vs resolution than traditional approach

Morandini, Ferber, Kahlhoefer 2308.01353

8. Experiment Design

Experiment Design

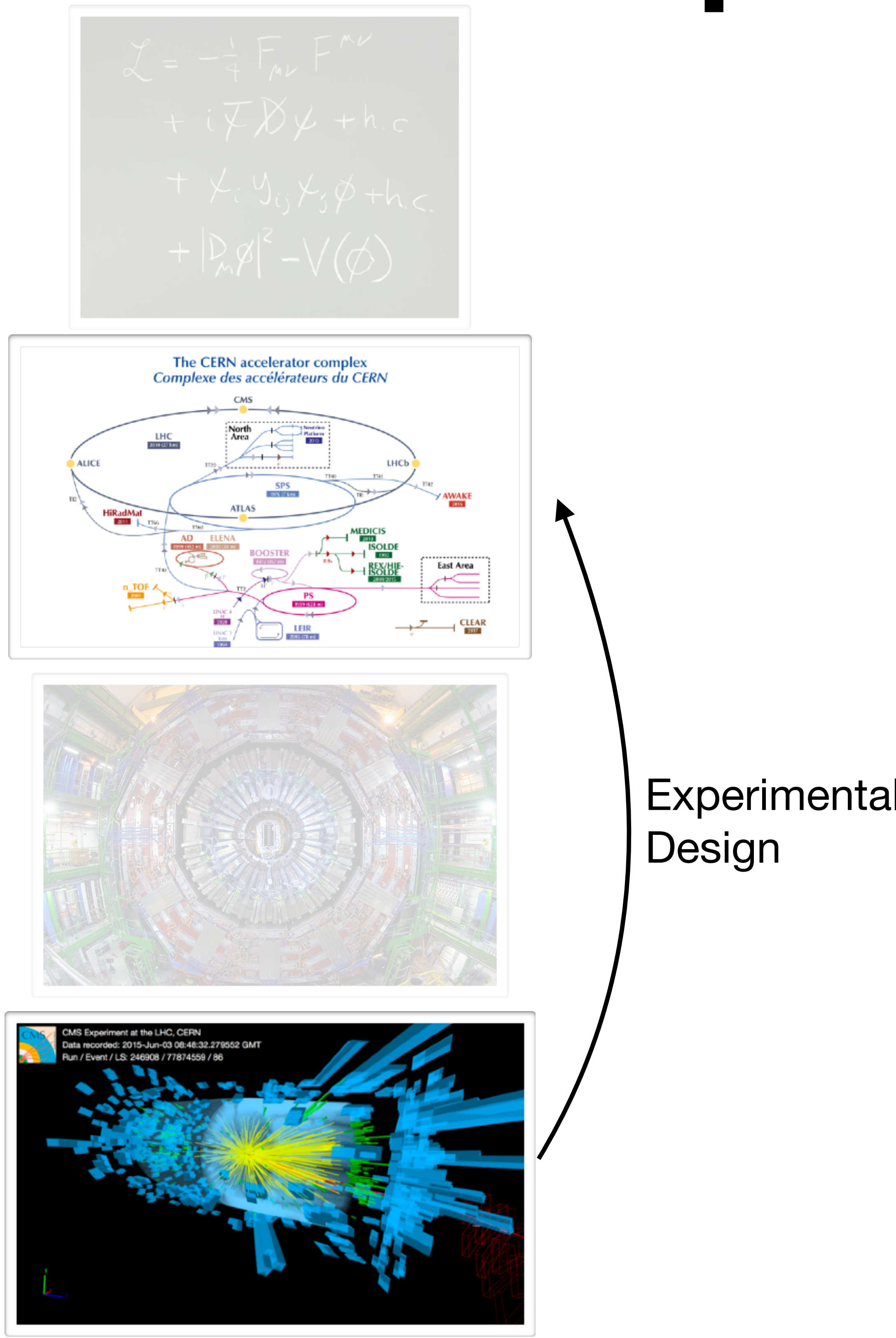




Experimental Design

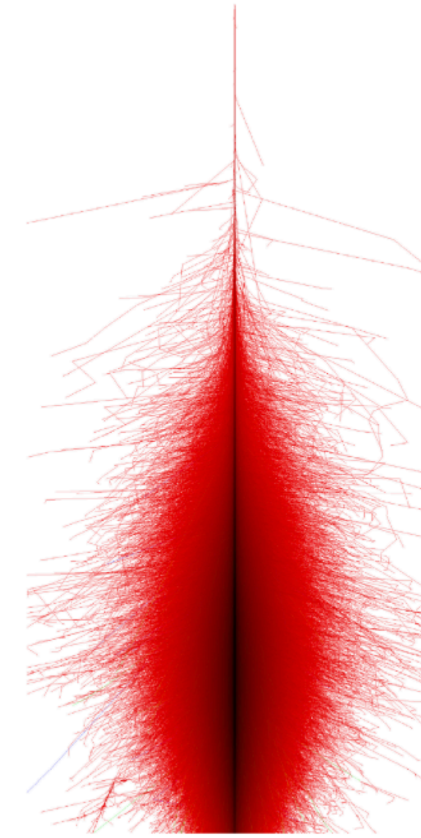
Automatically learn to arrange sensors given a physics target

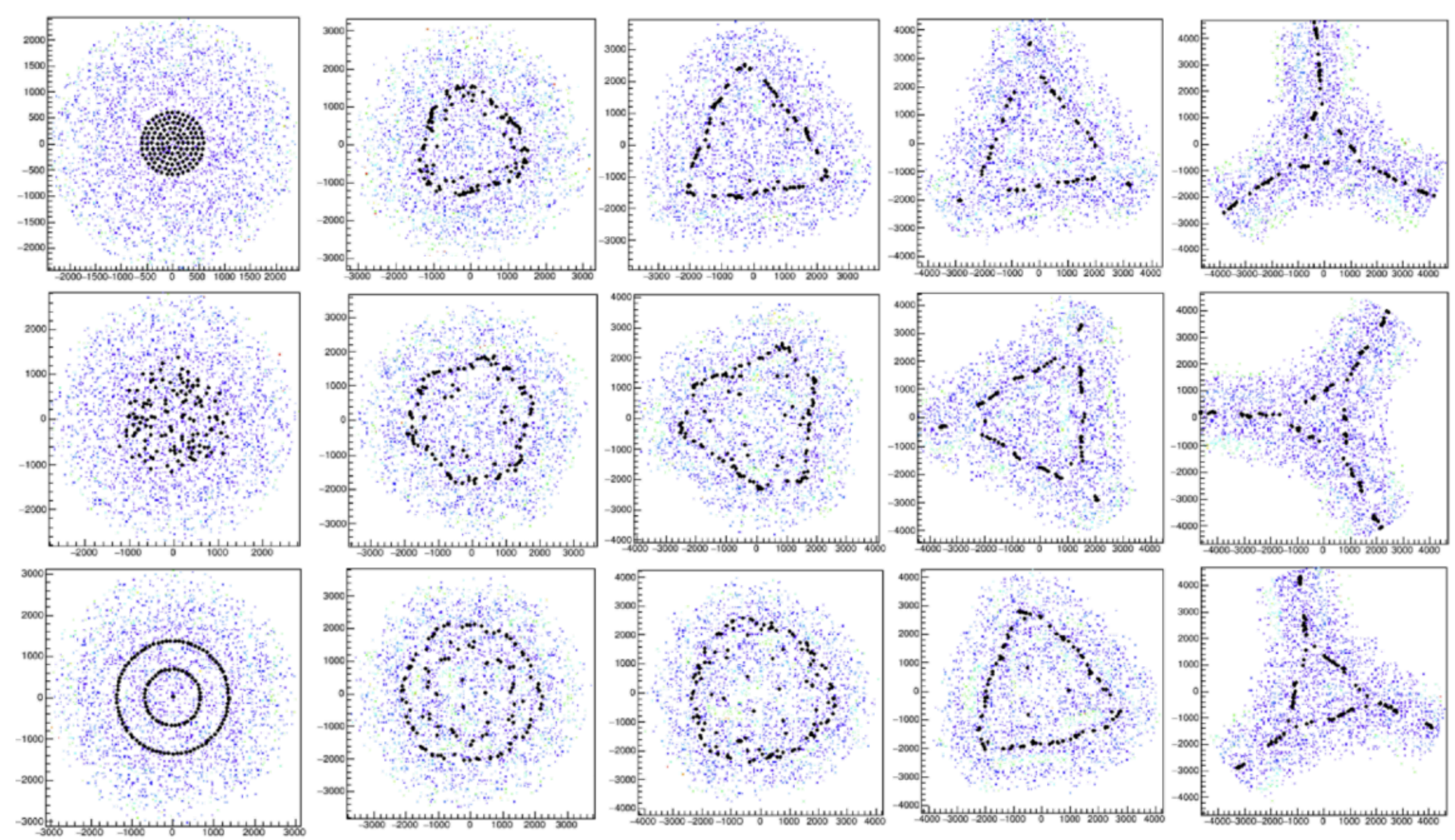
Experiment Design



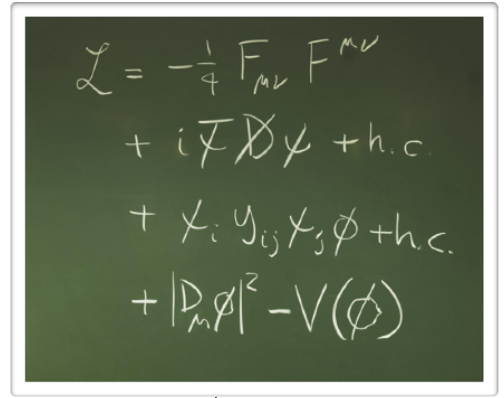
Automatically learn to arrange sensors given a physics target

Example tuning positions of detectors for a gamma ray observatory



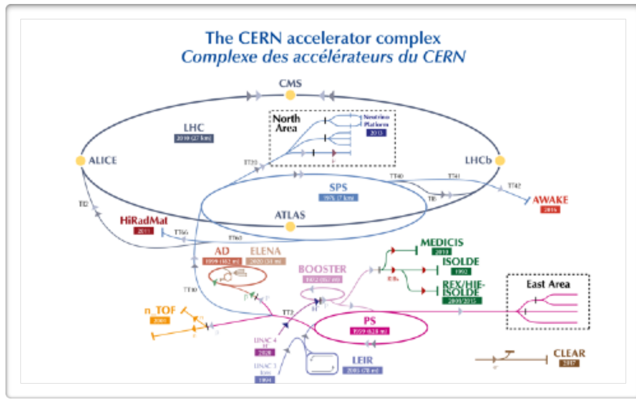


https://mode-collaboration.github.io/'; Dorigo et al 2310.01857



7. Inference

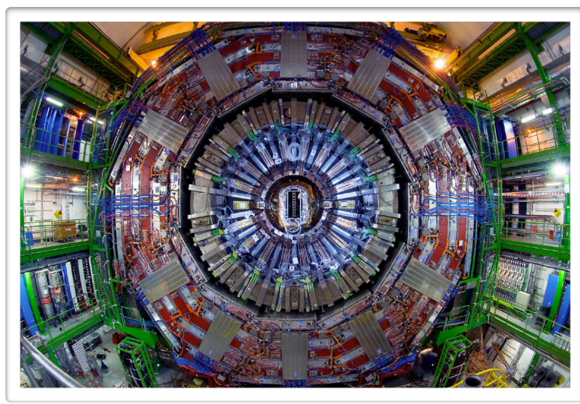
8. Experiment Design



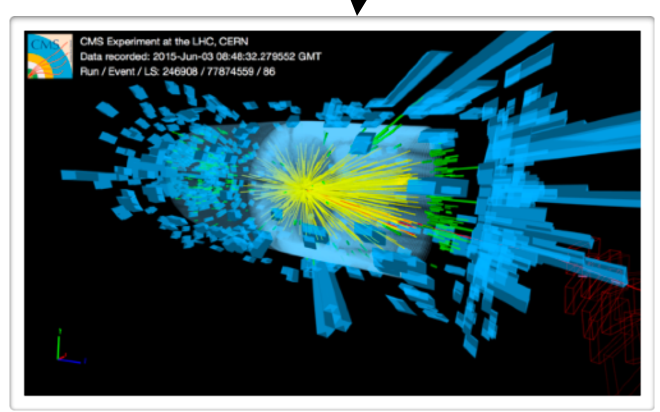
AI

3. Simulation

6. Triggers

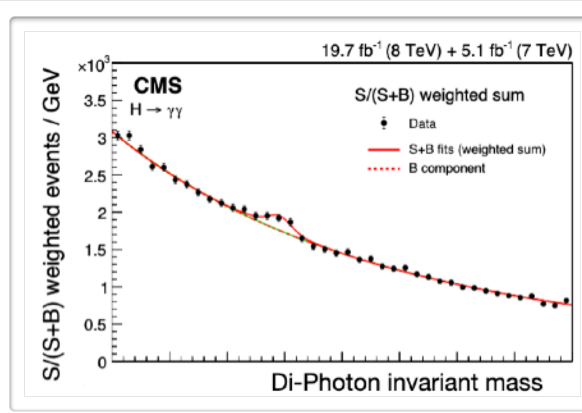


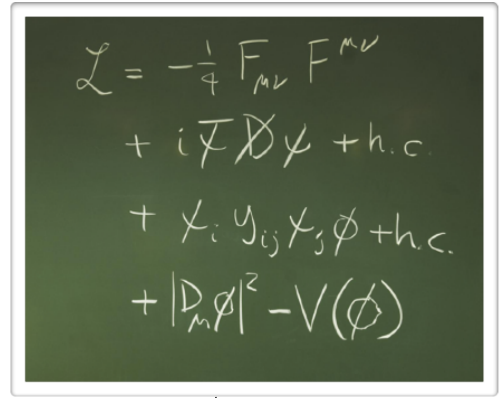
1. Tagging
2. Reconstruction



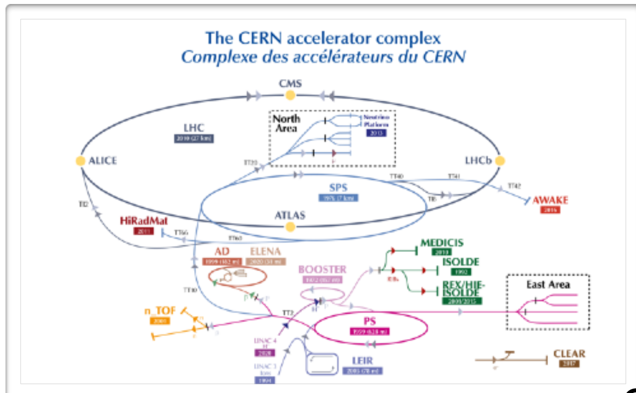
4. Unfolding
5. Anomaly Detection

AI models are studied for all aspects of experimental particle physics





9. Foundation Model

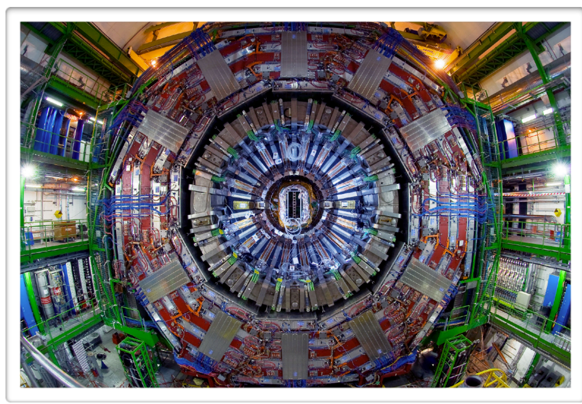


9. Foundation Model

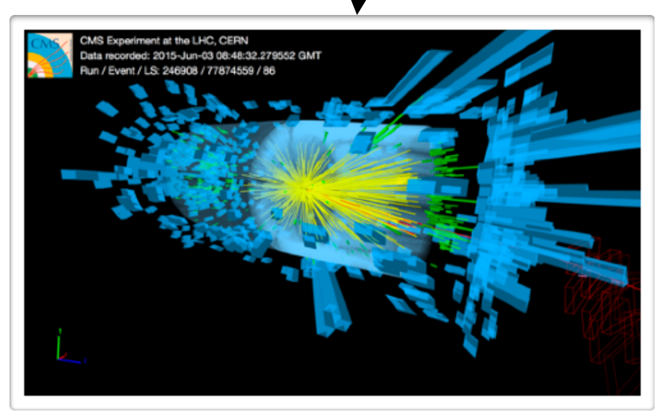
AI

9. Foundation Model

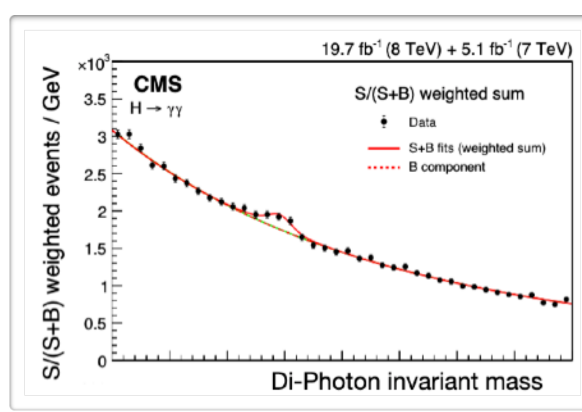
9. Foundation Model



9. Foundation Model
9. Foundation Model



9. Foundation Model
9. Foundation Model

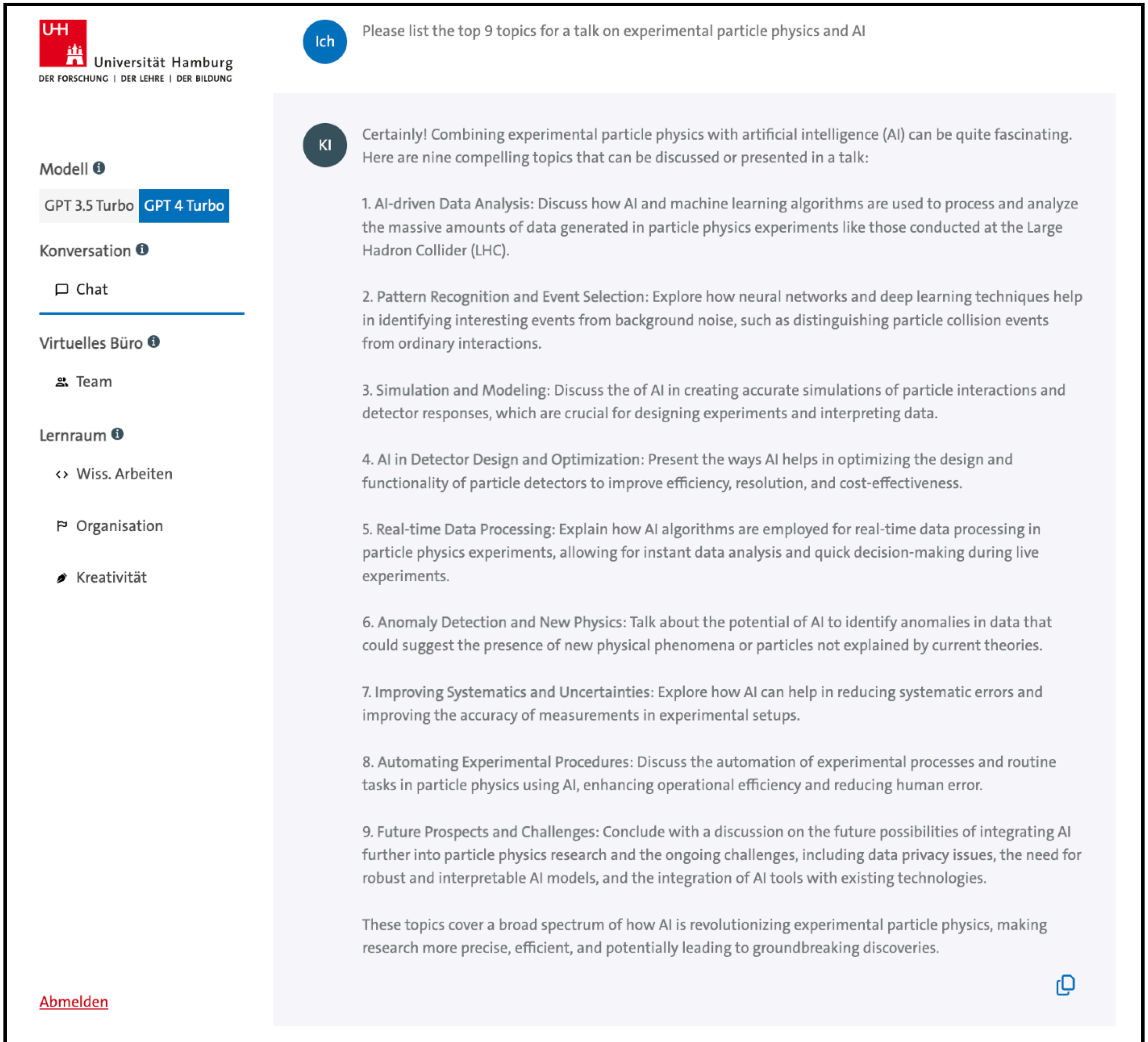


AI models are studied for all aspects of experimental particle physics

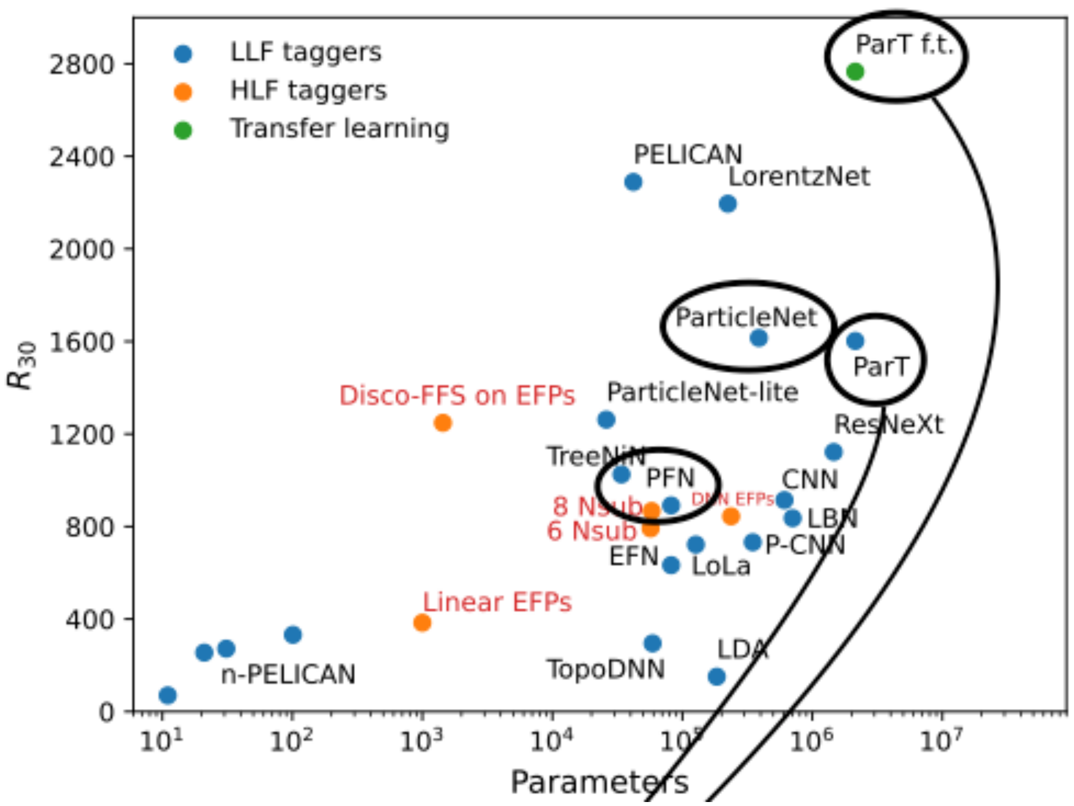
Can we find synergies?

9. Foundation Models

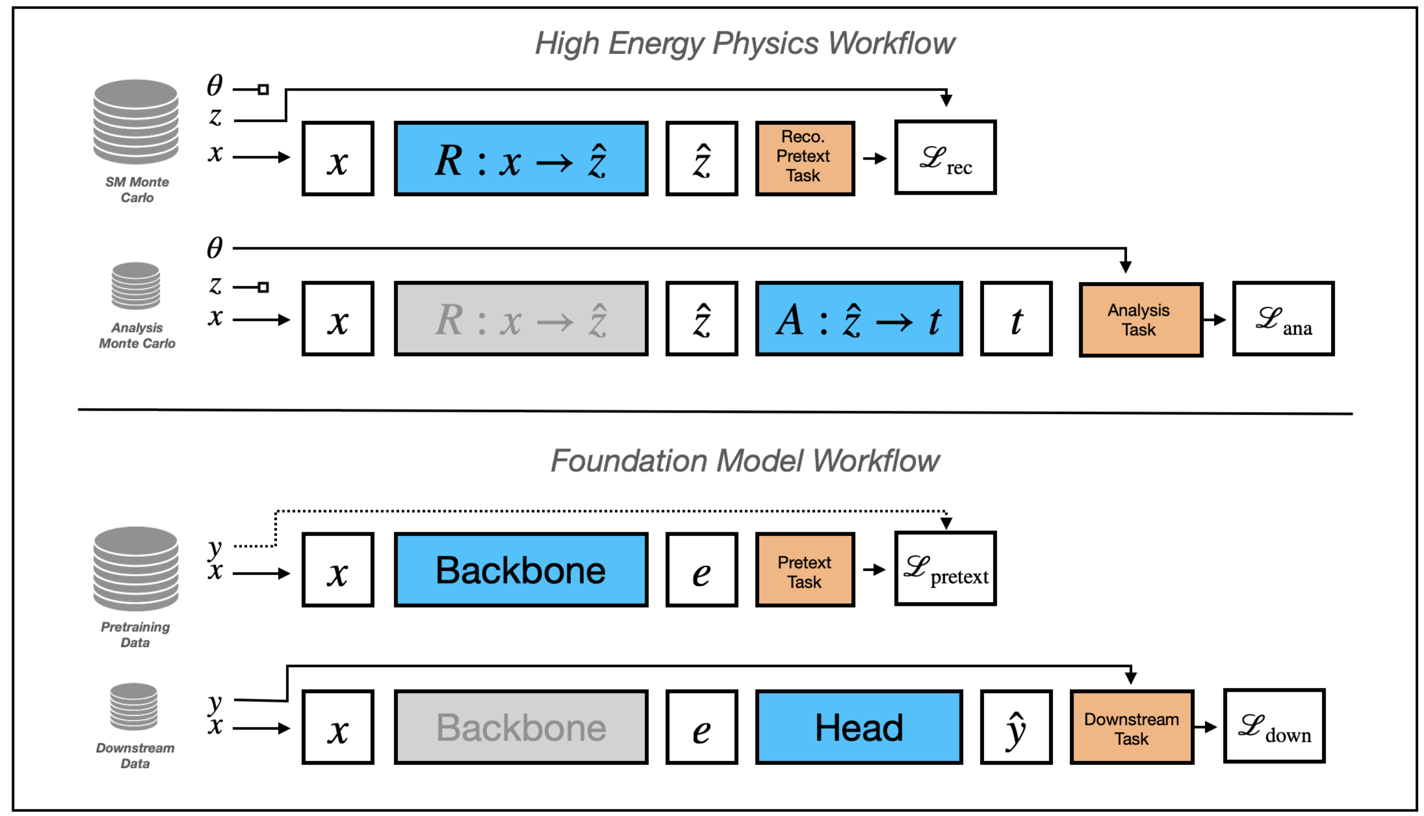
Foundation models



Foundation models for physics data



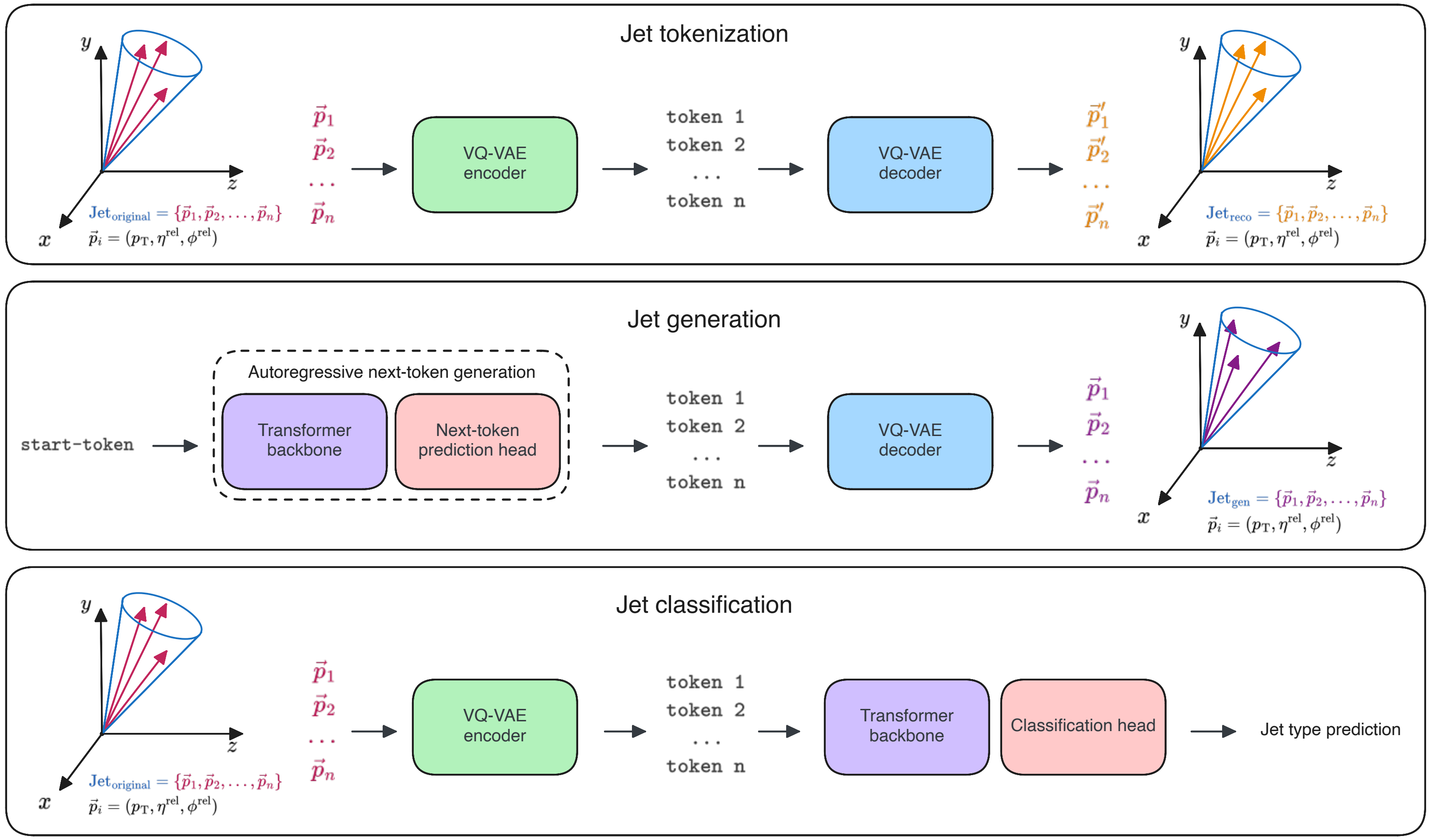
Already observed best performance in supervised classification by transfer learning



Foundation models extend transfer more broadly and centralise and re-use training

Qu et al 2202.03772; Vigl et al 2401.13536

OmniJet-α

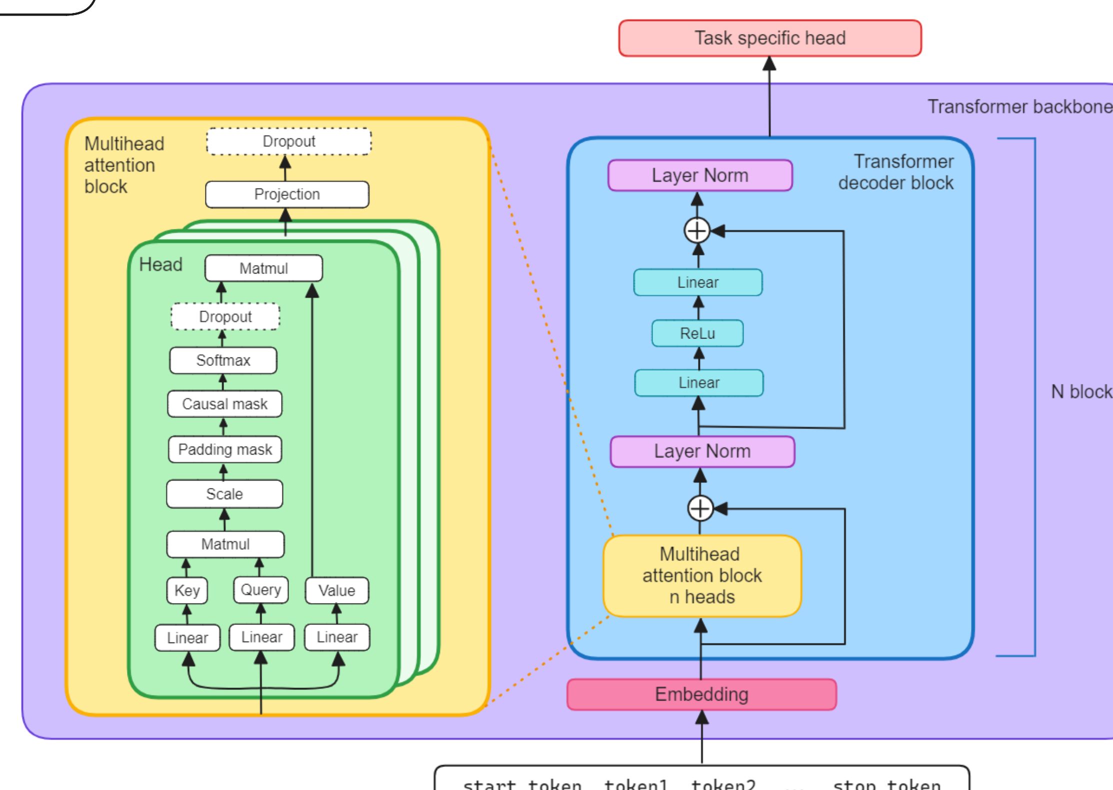


GPT for jets
using tokenised inputs

Train on
unsupervised generation

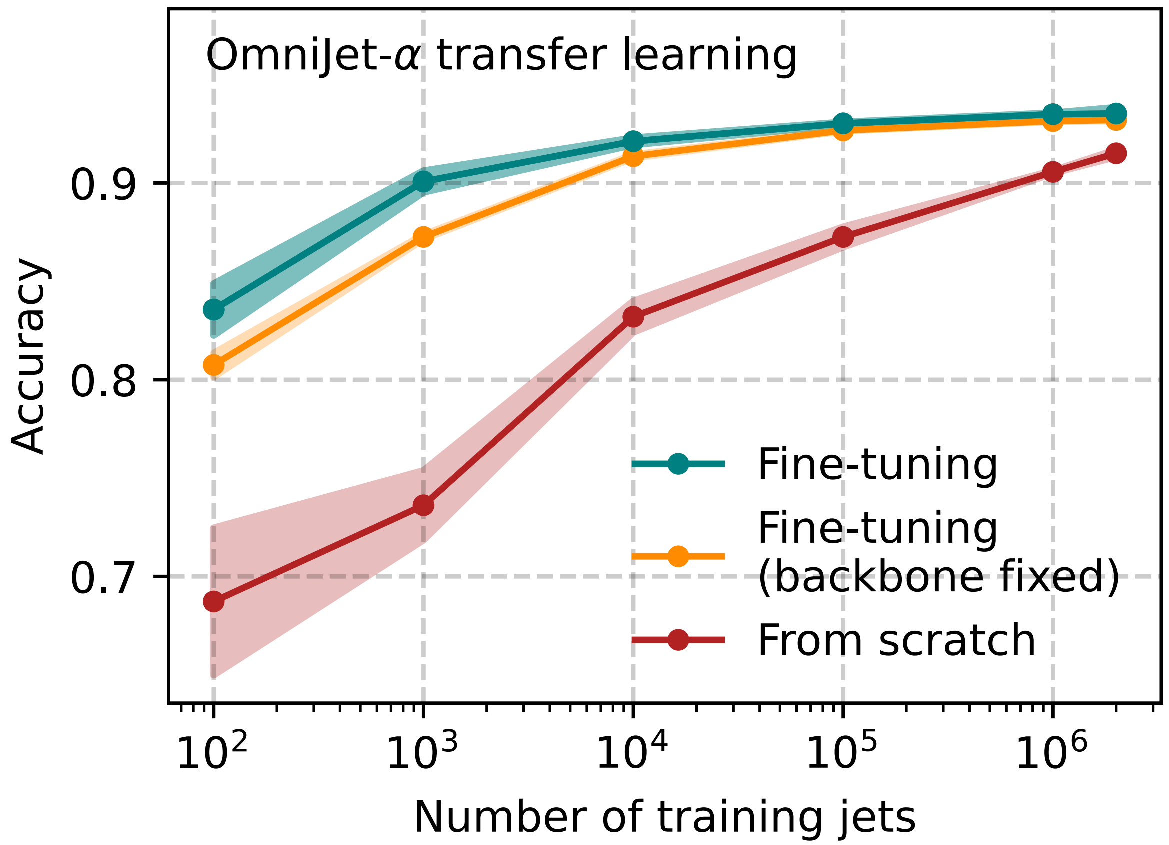
Evaluate for
supervised classification

Transformer
backbone

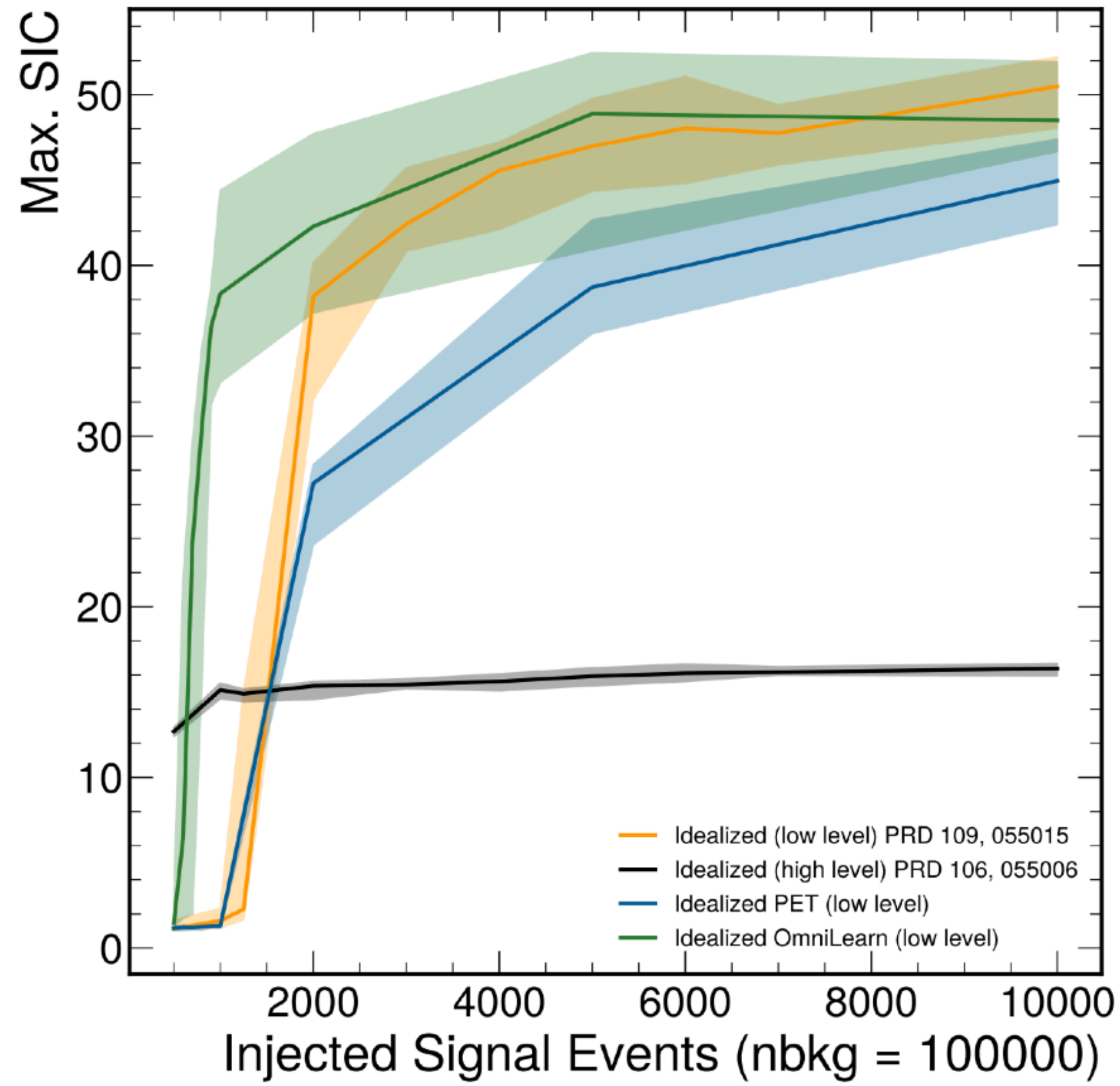


Birk, Hallin, **GK** 2403.05618; Butter et al 2305.10475; Finke et al 2303.07364; Vigl et al 2401.13536; Heinrich et al 2401.13537; Mikuni, Nachman 2404.16091

Generalisation



Pre-training on “cheap” unlabelled examples improves supervised classification data efficiency up to 100-1000x

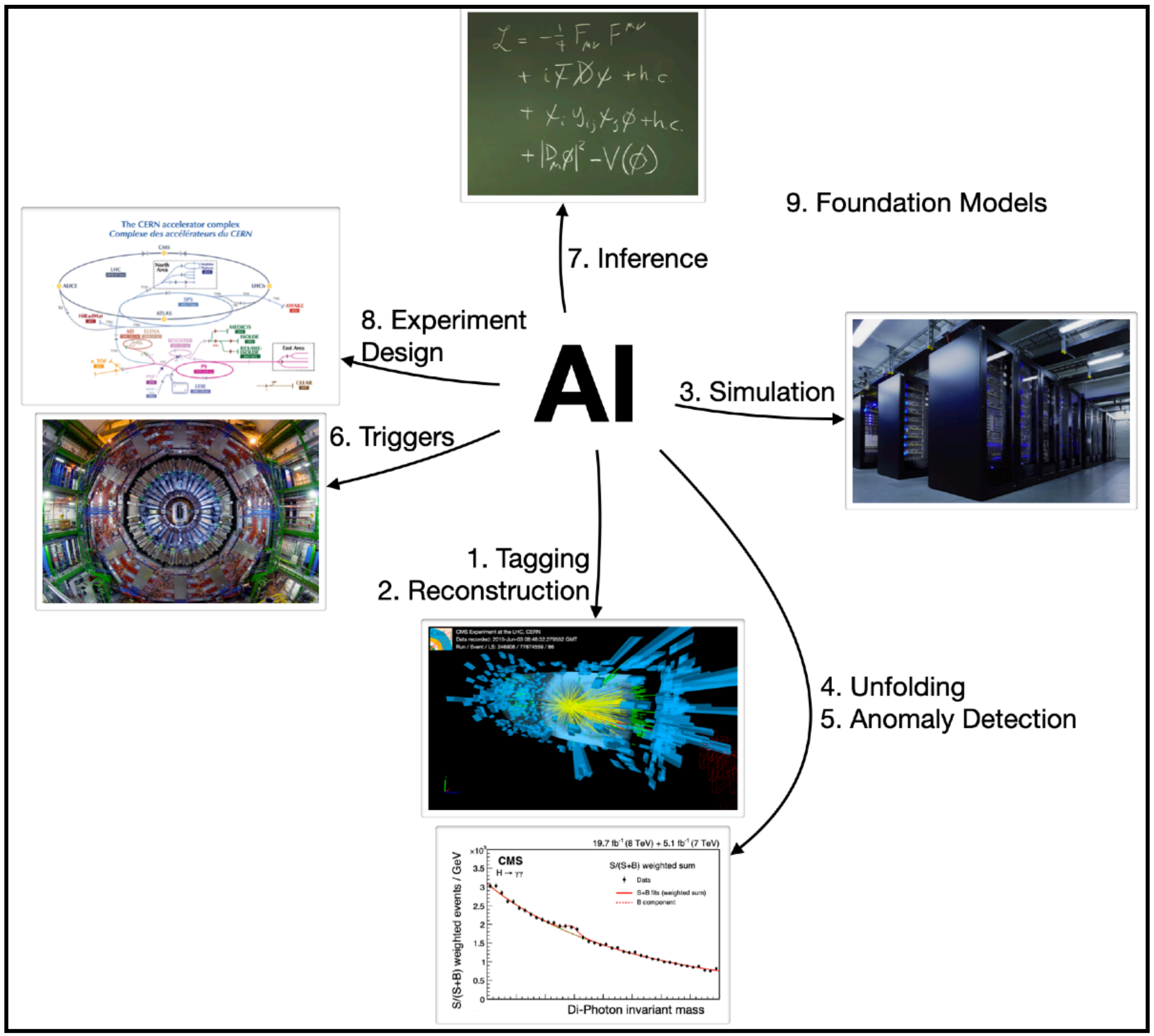


Very recent OmniLearn (Mikuni, Nachman) generalises across broad range of tasks including anomaly detection

Birk, Hallin, **GK** 2403.05618; Mikuni, Nachman 2404.16091

Closing

Conclusions



Extremely broad range of application for AI in particle physics

Way beyond concept studies: Modern tools are making a real impact in data analysis

Start to realise fully AI-based processing chains

Significant compute effort: Efficient models and sharing with foundation models matter

Thank you!