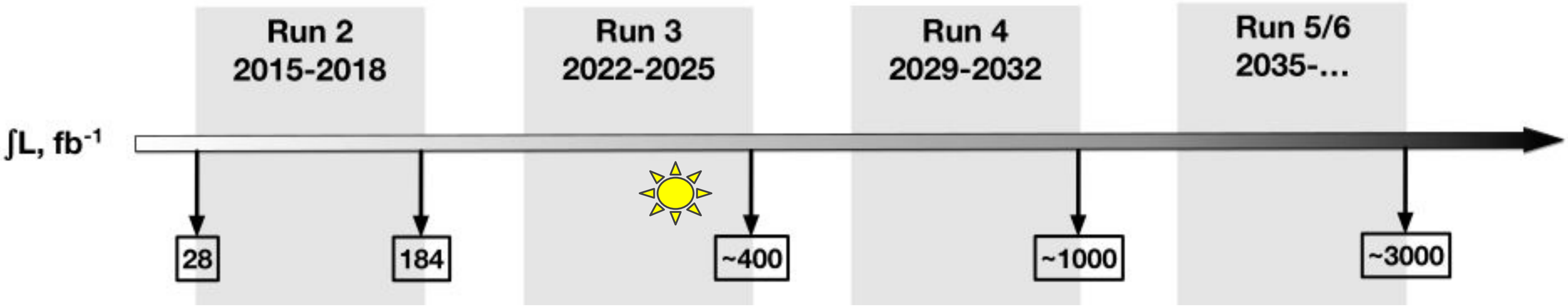
Norwegian Roadmap for CERN-related research

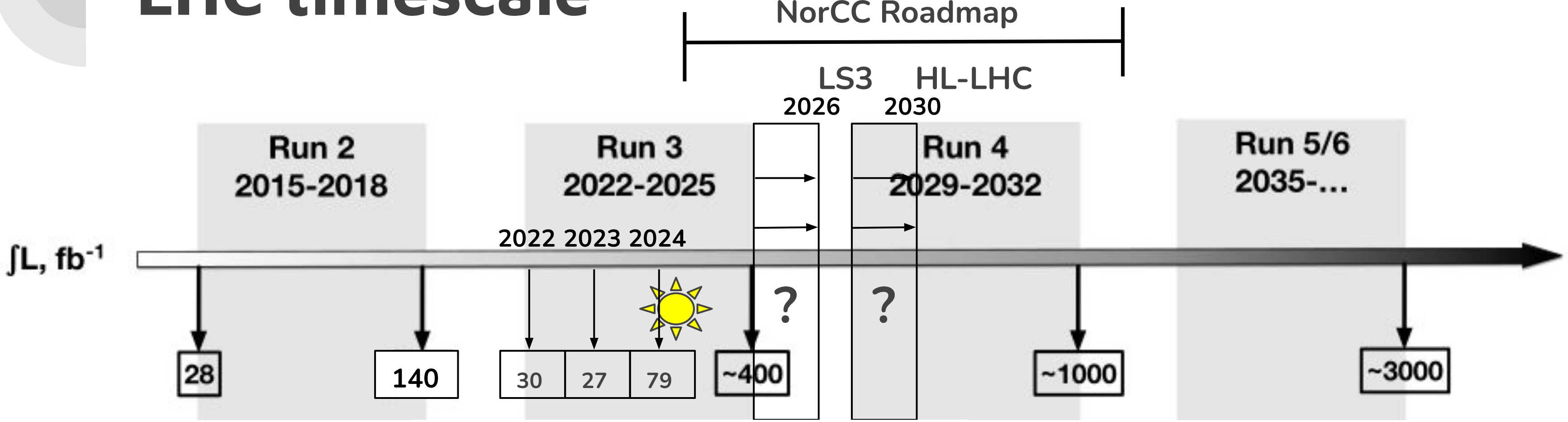
ATLAS towards HL-LHC

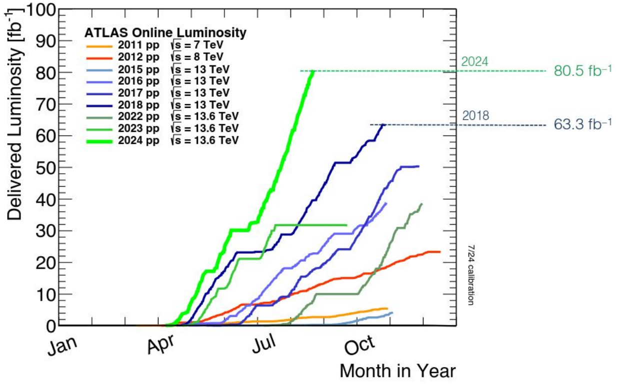
NorCC Activity 1 - Particle Physics

LHC timescale



LHC timescale





Physics Goals

What are the exact properties of the Higgs boson

What are the plausible theoretical models explaining the experimentally observable beyond-Standard Model phenomena?

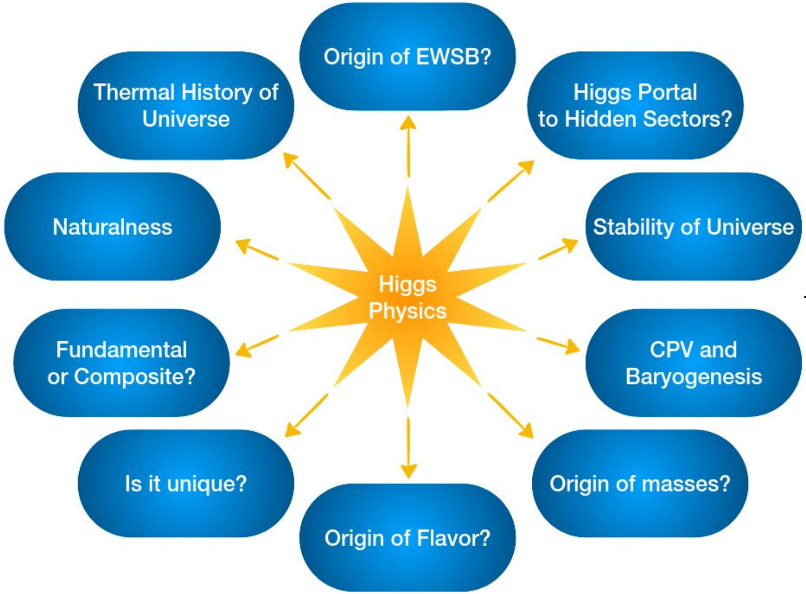
How can gravity be included in the same theoretical framework as particle and nuclear physics?

What is the origin of dark matter and dark energy, which take up to 85% of the total Universe mass?

What is the origin of the large matter / anti-matter asymmetry in the Universe?

Have fun!

Higgs Physics



Ref.: **Report of the Topical Group on Higgs Physics for Snowmass 2021: The Case for Precision Higgs Physics**, https://arxiv.org/abs/2209.07510

Goal	Physics Interest	Expertise
Baryogenesis	New sources of CP violation in Higgs to tau-tau Precision Measurements of Higgs Higgs self-coupling via double Higgs production	Tau reconstruction Statistics $H \to \tau\tau$ $H \to \gamma\gamma$ $H \to$ multi-leptons
Dark Matter	Higgs and Dark Matter	

Dark Matter Searches

Goal	Physics Interest	Expertise	Novel methods (see next slide)
Find dark Matter / (Have fun!)	Supersymmetry searches	**<u>with light leptons</u>**: direct production of electroweakinos and sleptons	Parameterized Neural Networks (PNN)
		<u>with taus</u>: gluino production	Improved tau-reconstruction (lowering momentum threshold)
	Exotic phenomena	extended gauge sector with new resonances (W’, Z')	
		mono-Higgs, Z and Z’	supervised, semi-supervised and unsupervised ML techniques
		resonant graviton excitations	
		contact interactions producing a non-resonant high-mass $\ell^+\ell^-$ excess	
		excited taus testing compositeness	specific reconstruction algorithms for collimated taus
	New physics with taus	Low-mass resonance search (H→ aa, a→ tautau)	tau-reconstruction using Graph Neural Networks (GNN)
		High-mass resonance search (X→ HH, H→ tautau)	

R&D Analysis

Long Lived Particle Searches
- Heavy Neutrinos (baryogenesis through leptogenesis
- More data, more parameters; suitable for deep learning (DL)
- Signature based approach using anomaly detection
- Complementary to current BSM searches
- UiO + Bergen

Deep Learning Algorithms
- Move from vision inspired techniques to Graph Neural Networks?
- Parametrized Neural Networks
- See talk from HVL

Advanced Statistics
- Estimation of trials factors, brute-force monte carlo vs. asymptotic approx.
- Likelihood free inference, applications in HEP

Triggers and Reconstruction
- collimated taus
- NEXTGEN trigger project
- Tau

More detailed information on what we have “promised” to do

- Talk at recfa in Oslo in 2023
- NorCC: Workshop 2022
- In backup

Software and computing

ATLAS has published a road-map setting out how it will meet the challenges of data taking and simulation in runs 4 and 5

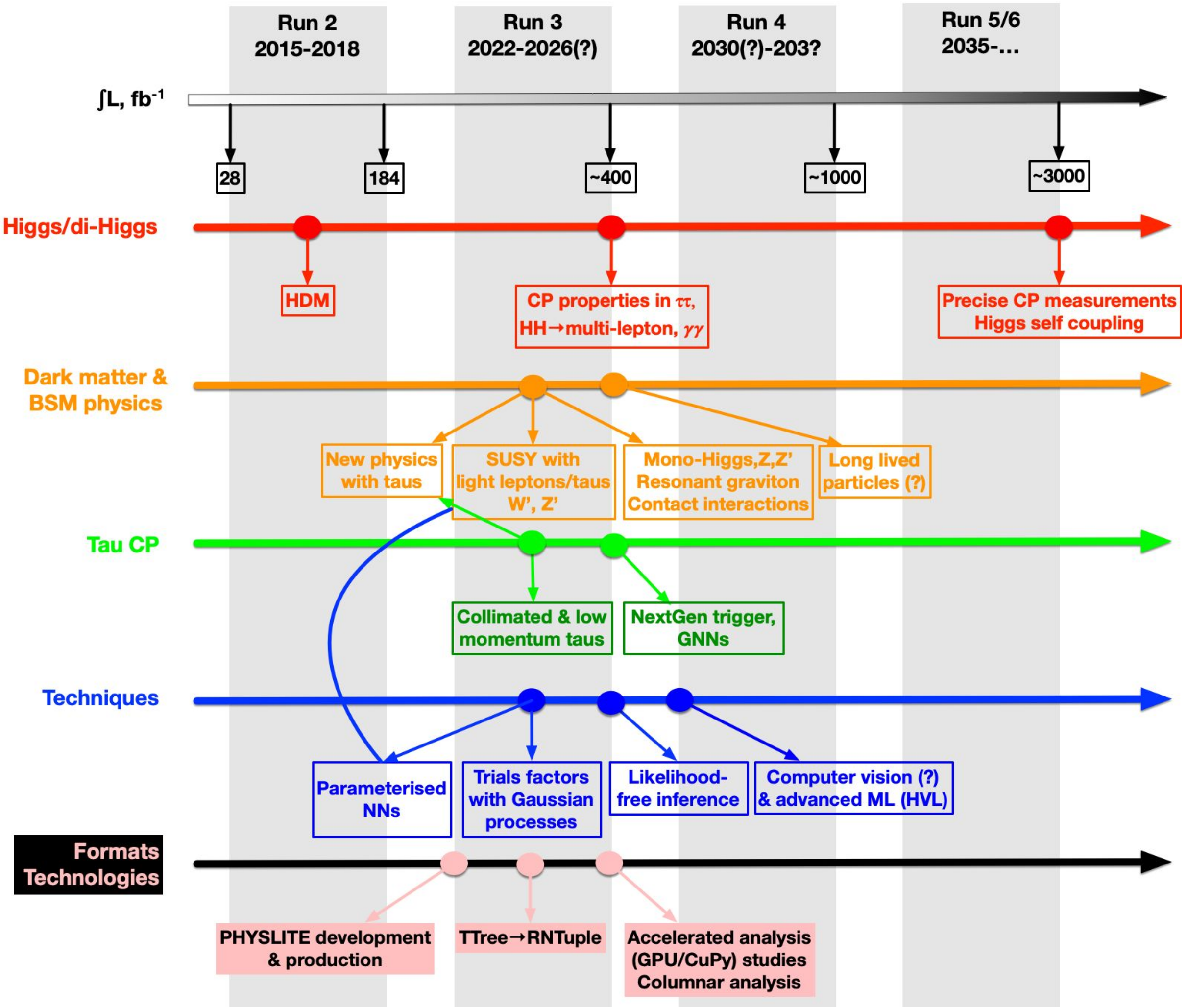
A summary of the main development objectives is shown below, with an indication of Norway's involvement

ATLAS development objective	Activity in Norway
Greater use of online data processing	No involvement
Adapt suitable software such that it can run on GPUs	Minimal current activity, aspiration to do more
Adapt software to use more efficient, modern design principles and common software components	Minimal current activity, aspiration to do more
Connect physicists and applications to specialist resources such as HPC centres, including those comprised mainly of GPUs	Strong historical involvement via ARC, new effort recently secured
Develop approaches to event generation, simulation and reconstruction based on machine learning rather than procedural software, to improve performance and compatibility with GPUs, whilst maintaining physics quality	Minimal current activity, aspiration to do more
Develop and make greater use of fast simulation whether parameterised or generative-ML	No involvement
Continue to develop production workflows that make greater use of tape for long-term storage	Historical involvement but minimal at present
Continue to develop compact data formats for analysis and the software needed to process them	Strong (leading) contribution
Develop computationally efficient techniques for statistical inference	Strong research programme planned
Coordination activities	Strong involvement

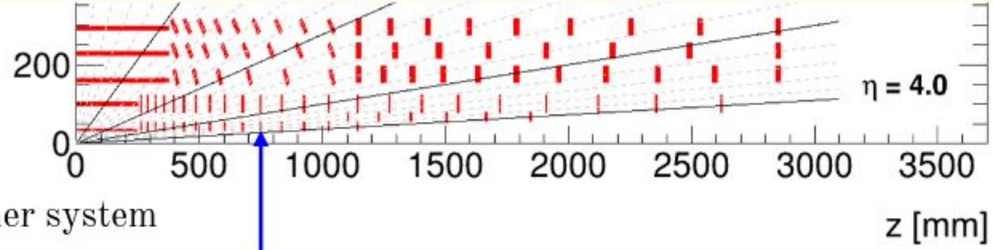
Unsupervised learning need more data and more features

di-Higgs (done several HL-LHC prospects studies)

Computer vision techniques replaced by more modern GNNs

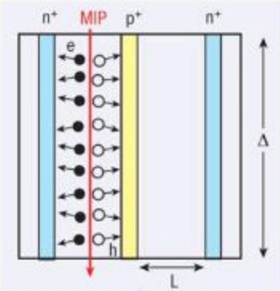


Detector upgrade

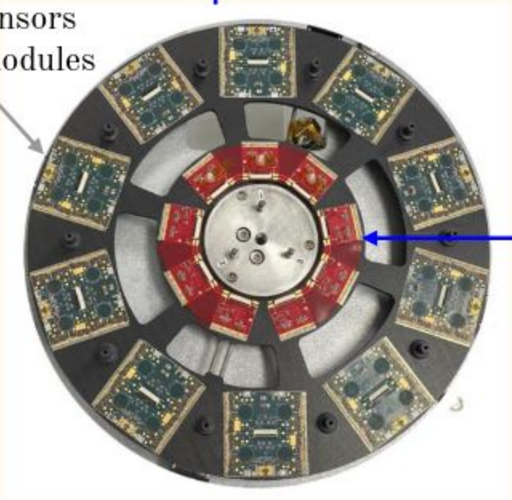


inner system

3D pixel sensors, radiation hard
made in Oslo (SINTEF)
used in inner-R0/R0.5 rings
1 pixel = 50x50 μm
1 FE chip ~ 2x2 cm ~ 153k pixels

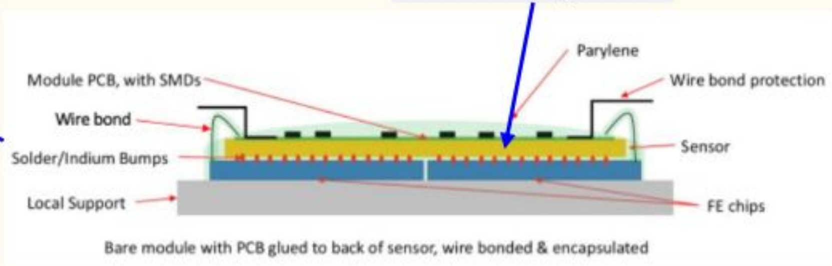


planar sensors
"quad" modules



innermost endcap ring "R0"

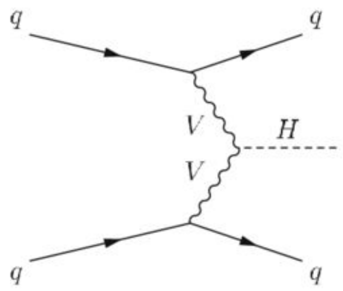
3D sensors, "triplet" modules
180 triplet modules in detector!
main deliverable for Norway



Module assembly and testing

BACKUP

Higgs Physics



New sources of CP violation in Higgs to tau-tau:

Our existence requires more CP violation than we presently see (**baryogenesis**)

Precision of 3-5^0 needed to test some of the baryogenesis models (10-5^0 expected after LHC Run 5)

Precision Measurements of Higgs:

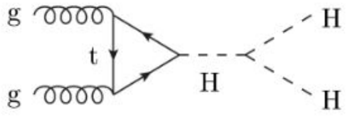
Improved background modelling

New approach to shape systematic uncertainties

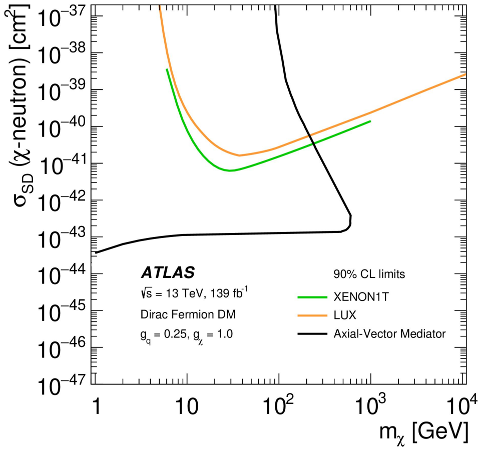
Higgs self-coupling via double Higgs production:

Higgs self-coupling through di-higgs production is basis to understand the Electroweak Phase Transition

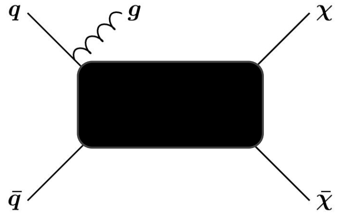
Present predictions: “Evidence” for hhh (if SM self-coupling) is end of LHC Run 5



Dark Matter Searches



Phys. Rev. D 103 (2021) 112006



<u>Mono-jet signatures</u>:

Interpretation in SUSY models

Use of ML

<u>Tau signatures</u>:

supervised, semi-supervised and unsupervised ML techniques

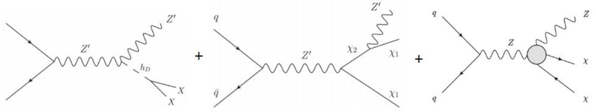
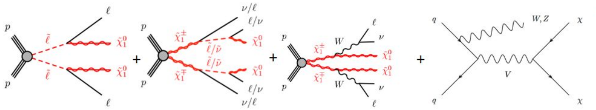
computer vision inspired techniques

Lowering momentum threshold on taus [**improved reconstruction**]

<u>Electron/Muon (multi-lepton) signatures</u>:

Looking at several different DM models in di-lepton + missing transverse energy final state

Developing more generic DL algorithm for (more) model independent searches

Dark Matter Searches cont’d

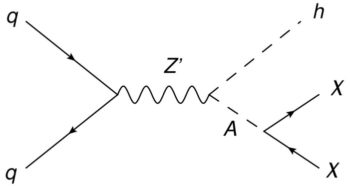
Higgs and Dark Matter:

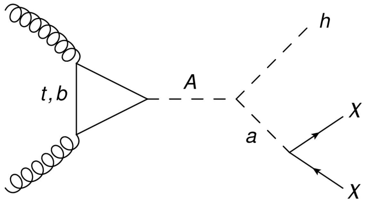
2-Higgs Doublet Models (2HDM) + Z’ or A

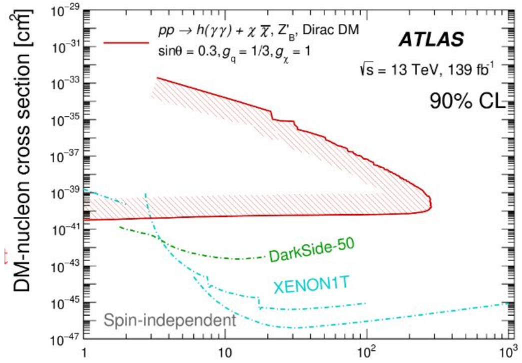
ATLAS published in H →γγ and H → bb-bar channels

Short term goal: Publish on H →ττ [first draft in preparation]

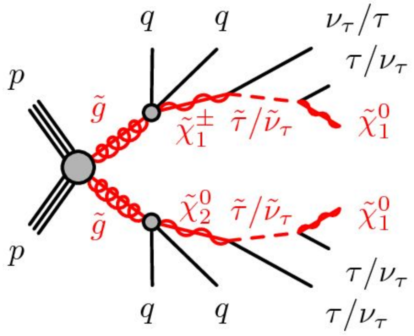
Longer term: combine all Higgs decay channels







New Physics with taus



Supersymmetry

Squarks and gluinos decaying to taus, jets, MET

Run 2 + 3 analysis underway

Scenario with m(q/g) ~ m(LSP) experimentally challenging but motivated by e.g. coannihilation

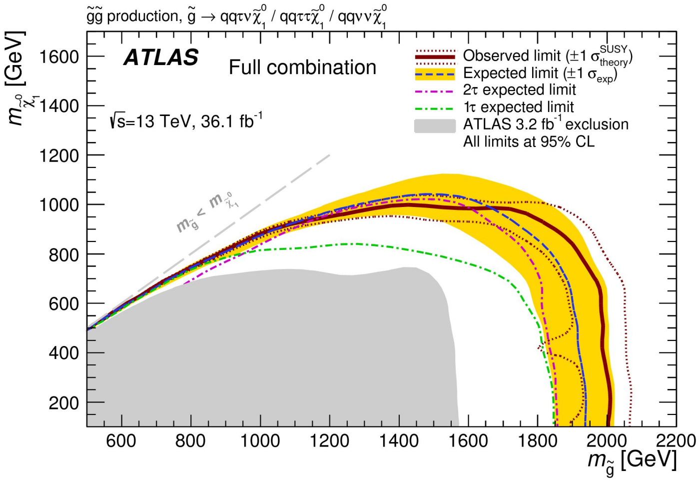
New resonance decaying to 2 collimated taus

Requires specific reconstruction when taus overlap - **algorithms being re-designed for Run3**

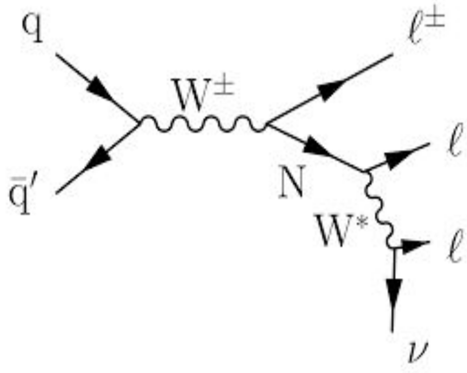
Low-mass resonance search (H→ aa, a→ tautau)

High-mass resonance search (X→ HH, H→ tautau)

Phys. Rev. D 99, 012009



New Physics with leptons



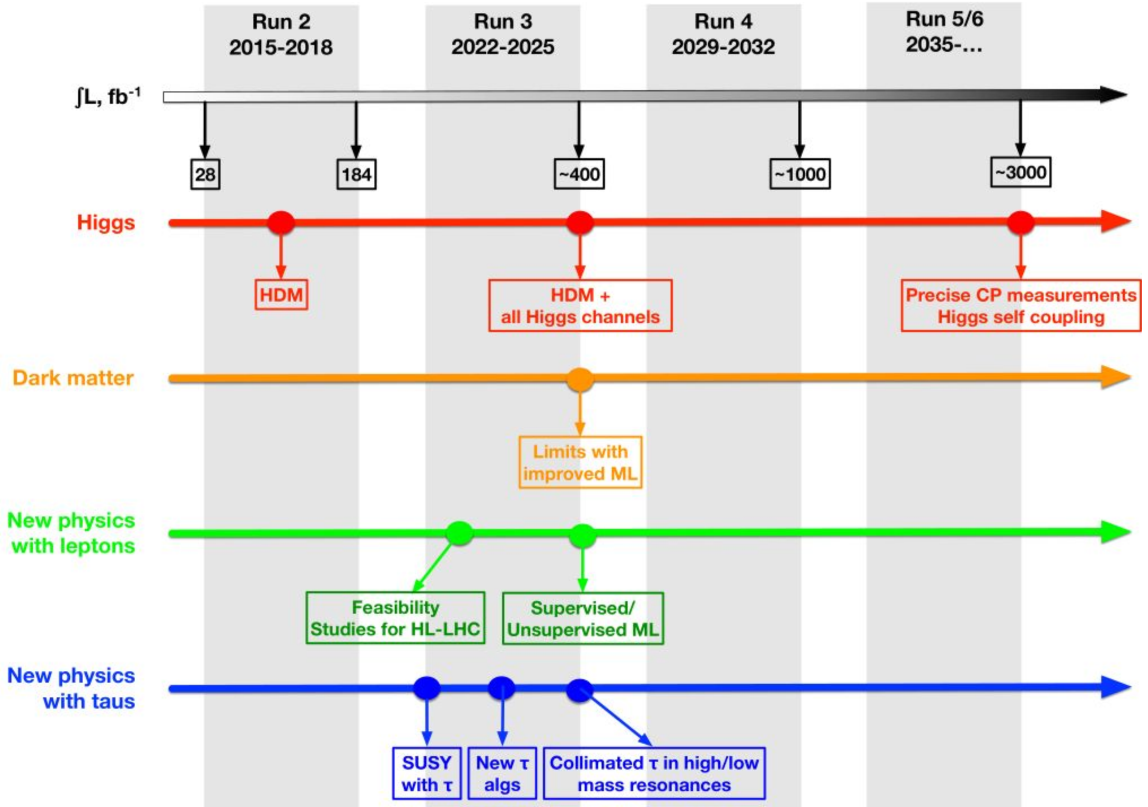
Searches for Heavy Neutrinos:

- Baryogenesis through leptogenesis
- Concentrating on the 3-lepton final state
- Comparison of supervised and semi-supervised (anomaly detection) DL a

Di-lepton studies in searches for new gauge bosons

- Run-3 and prospects for HL-LHC
- Study various beyond SM theories with new heavy gauge bosons

WIP: NorCC ATLAS Physics Roadmap



Common Approaches/Methods

Statistical Methods

Gaussian Process-based calculation of look-elsewhere trials factor: [arxiv:2206.12328]

- a recurring challenge to efficiently and precisely calculate the global significance of a potential new resonance
- Commonly used method not accurate for low significances (i.e. $\sigma < 3$)

New approach to shape systematic uncertainties

Reliable and Interpretable ML

More on this in the Computing, ML & AI Session on Thursday

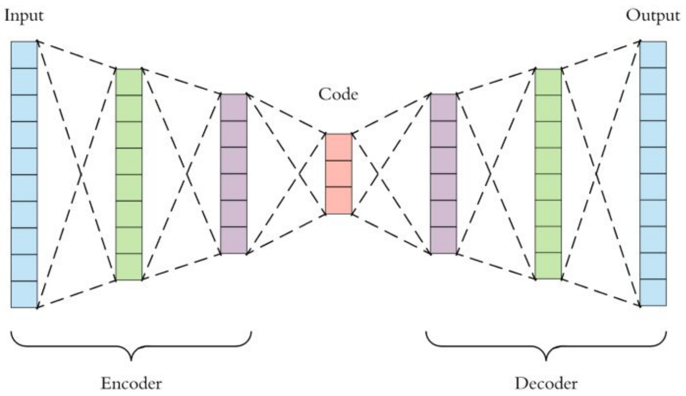
Supervised, semi-supervised and unsupervised learning

Build supervised and semi-supervised (anomaly detection) DL algorithms

- Compare performance by seeing how efficient the anomaly detector can identify “fake” signal in data compared with algorithms which have been trained on the same signal

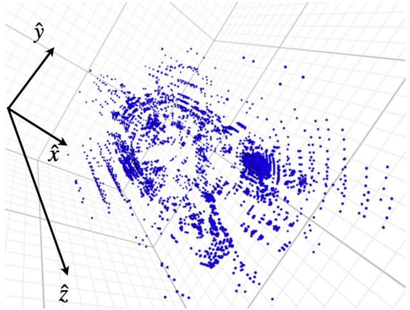
- If the anomaly detector is efficient in picking out several different BSM signal - model independent approach

- Currently investigating Auto Encoders for anomaly detection



Credit: https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798

Improved Tau Reconstruction



Jet performance group develops advanced techniques ("point cloud" methods) to reconstruct pions using calorimeter **and tracking** information:

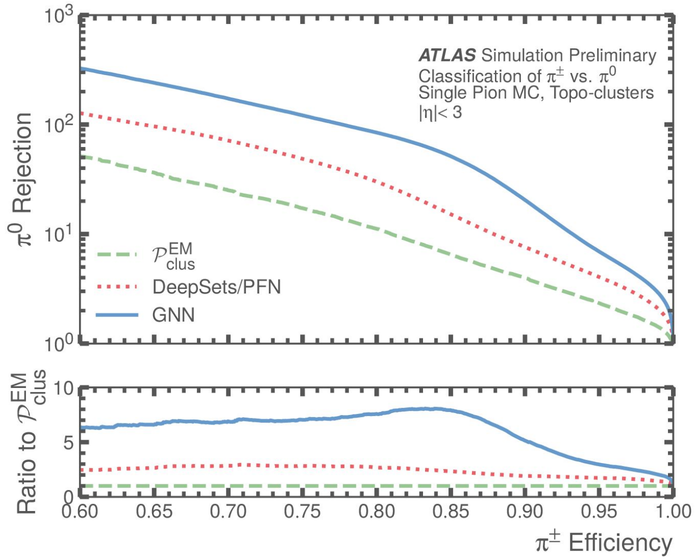
- classification: $\boldsymbol{\pi}^0$ vs $\boldsymbol{\pi}^{\pm}$
- regression: pion energy calibration

Graph Neural Networks outperform other methods (likelihood, deep set NN, convolutional NN), especially when combining clusters + cells + tracks.

Currently tested in "clean" environment (no pileup).

Promising for tau reconstruction (calorimeter cells available)!

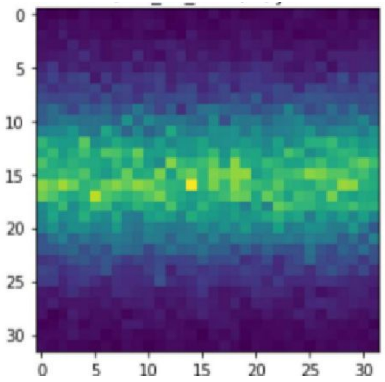
ATL-PHYS-PUB-2022-040



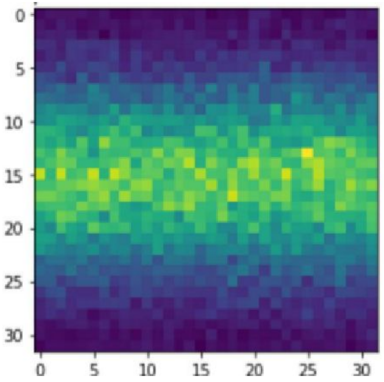
Computer Vision Inspired Techniques

Use of low level data of calorimetric deposits to distinguish between Black Holes and Sphalerons

Images of eta versus phi for jets



Black Hole



Spahleron

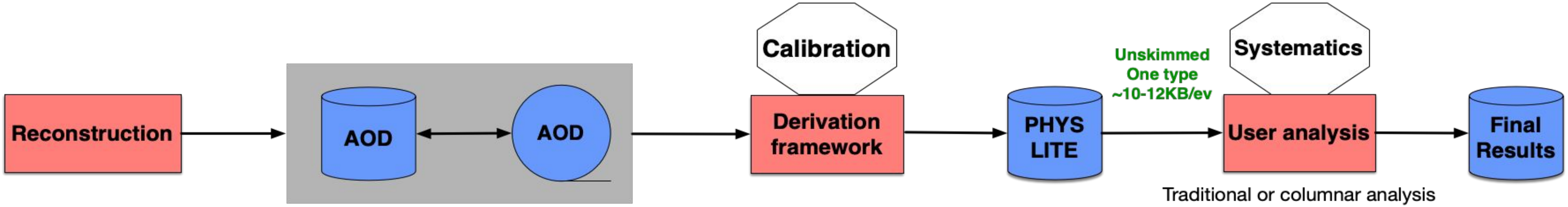
Tau reconstruction?

Dark Matter model indep. searches?

Common Tools and Infrastructures

Data formats for run 4

- DAOD_PHYSLITE is at the heart of the computing and analysis models for run 4. Almost all (~80%) of physics analyses will use these files as input
    - Written in the ATLAS xAOD data structure
    - Contains calibrated quantities and the variables required to evaluate instrumental systematic uncertainties
    - Target size: 10-12 KB/event
    - Significant development and maintenance input from NorCC personnel
- Allows more optimal use of disk due to its smaller size
- User analysis runs faster than current data formats due to lower I/O and pre-calibrated quantities
- xAOD format allows its use as an input for columnar analysis → next slide
- Commissioning will take place during run 3



From ROOT to Python Ecosystems

Most analysers using a common ATLAS format (xAOD PHYSLITE), converted into ROOT nTuples

With increasing use of ML there is a need for efficient parsing of data structure from **ROOT nTuples** to **Pandas DataFrames**

- Use of **RDataFrame**
  - parallelization comes for free
  - easy conversion into **numpy arrays** and **Pandas DataFrames**
  - Write to ML-friendly output formats like **hdf5**

Towards HL-LHC (i.e. Run 4 and onwards)

New data format:

The common ATLAS data format in Run 4 will be more light-weight, containing pre-calibrated objects only

- Less need for each analyser to run calibration tools (simplified workflow)
- Can be used directly from "the shelf"
  - Columnar analysis
  - Interactive Analysis (e.g. jupyter notebooks)
  - Direct access to data on local analysis facility

Expected updates to ROOT:

- A new nTuple format called **RNTuple** (optimized for RDataFrame)
- Moving to **ROOT7**

Exploitation of GPUs:

- Already being used in training of ML algorithms
- Use of novel data structures (CuPy)

Roadmap

LHC Run-3 will be a preparation towards Run-4 (HL-LHC)

Several interesting physics analyses

Where can we focus?

Particular development of new tau reconstruction algorithms

New workflows

Columnar and interactive analyses