Recent developments for Artificial Intelligence application to Particle Physics

David Rousseau, IJCLab-Orsay

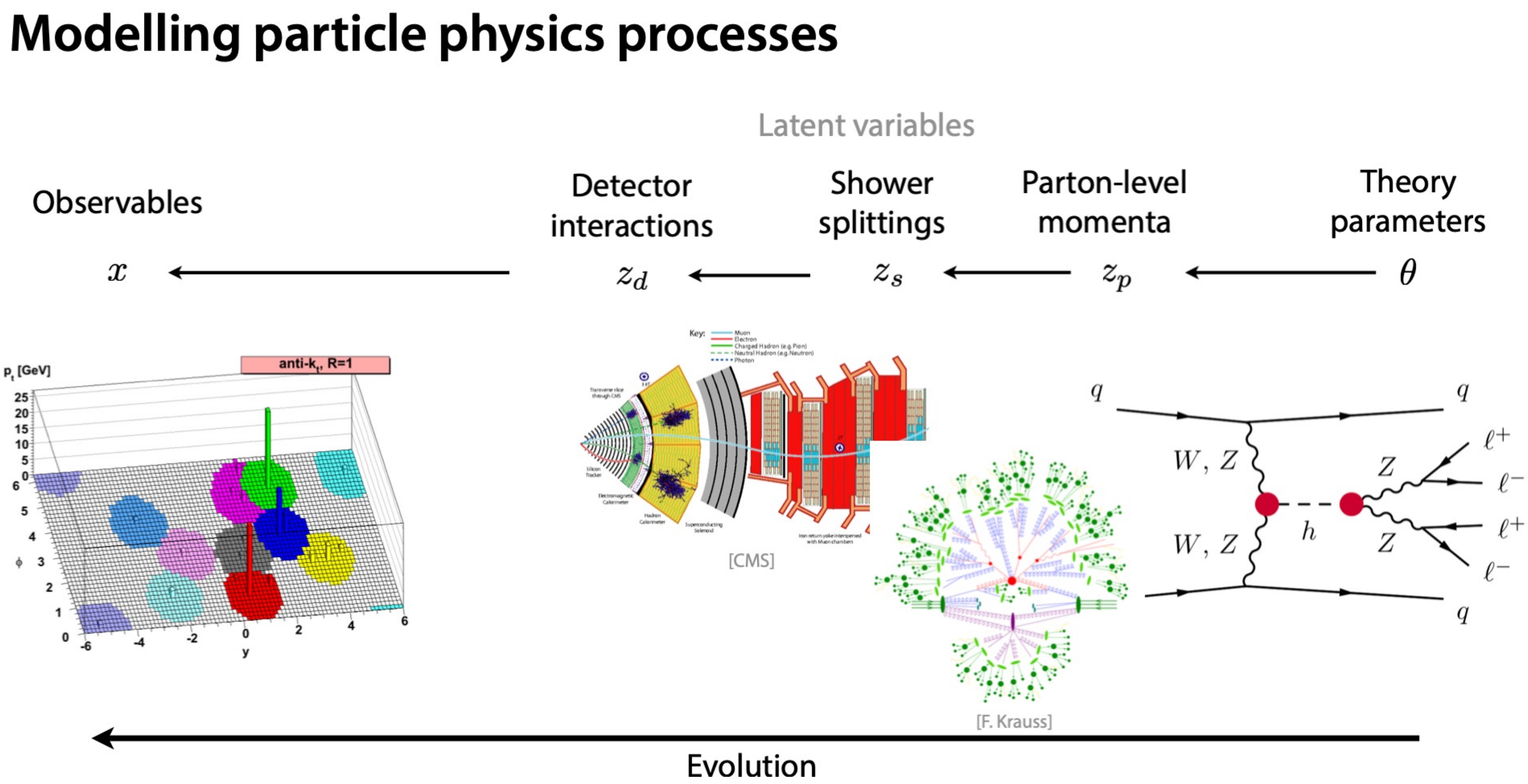
david.rousseau@ijclab.in2p3.fr @dhpmrou

CLAS12 workshop, March 2023, Paris

Outline

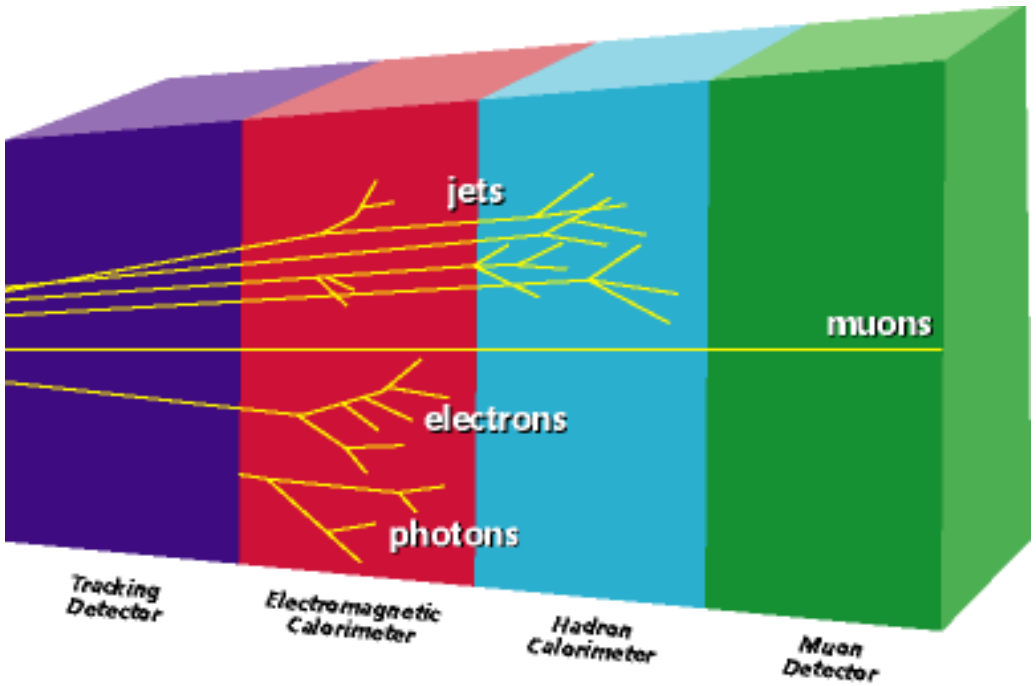
- Particle-level, event-level and experiment-level inference (e.g. ATLAS or CMS experiment on the Large Hadron Collider at CERN)
- High Energy Physics data are not images
- GAN/VAE for simulators
- Dealing with uncertainties

Modeling/Inference



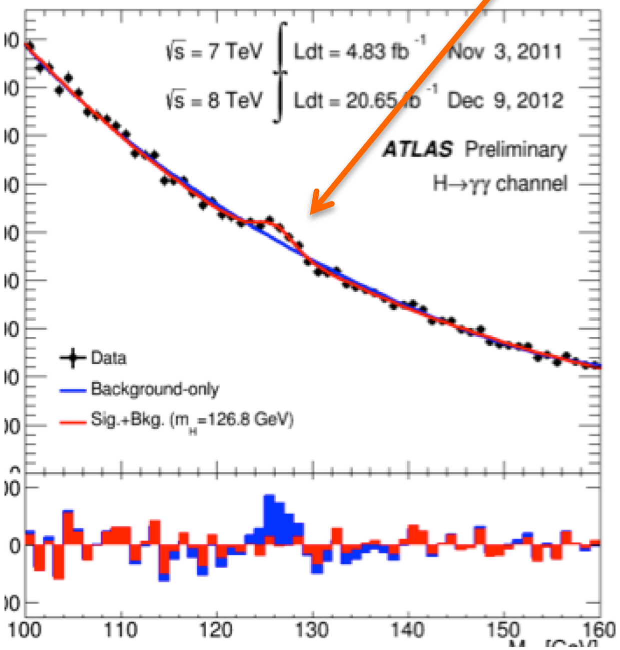
Particle-level inference

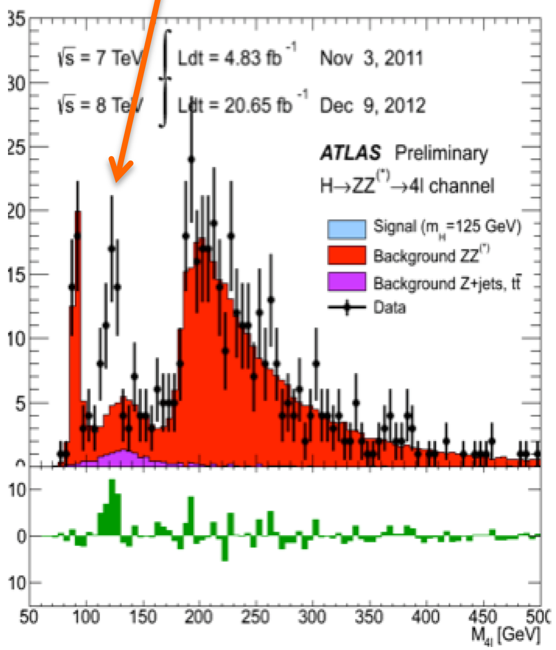
- Particles identified (pattern)
- And measured : 3D direction and energy, origin

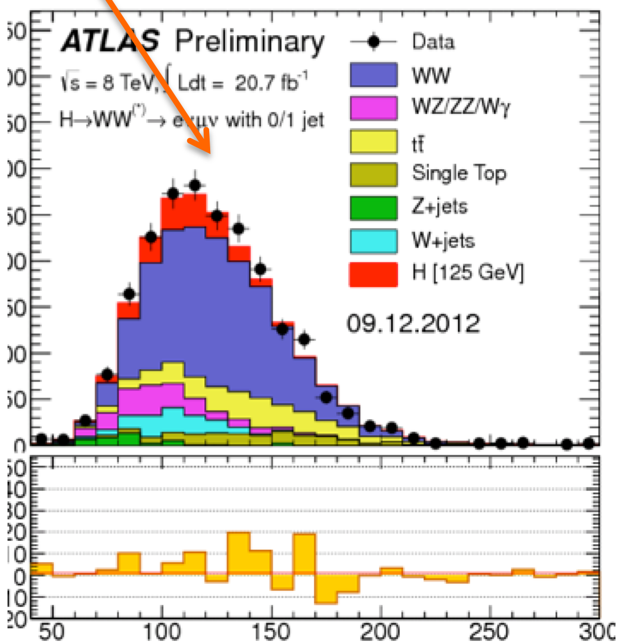


Experiment-level inference

To « see » the Higgs boson



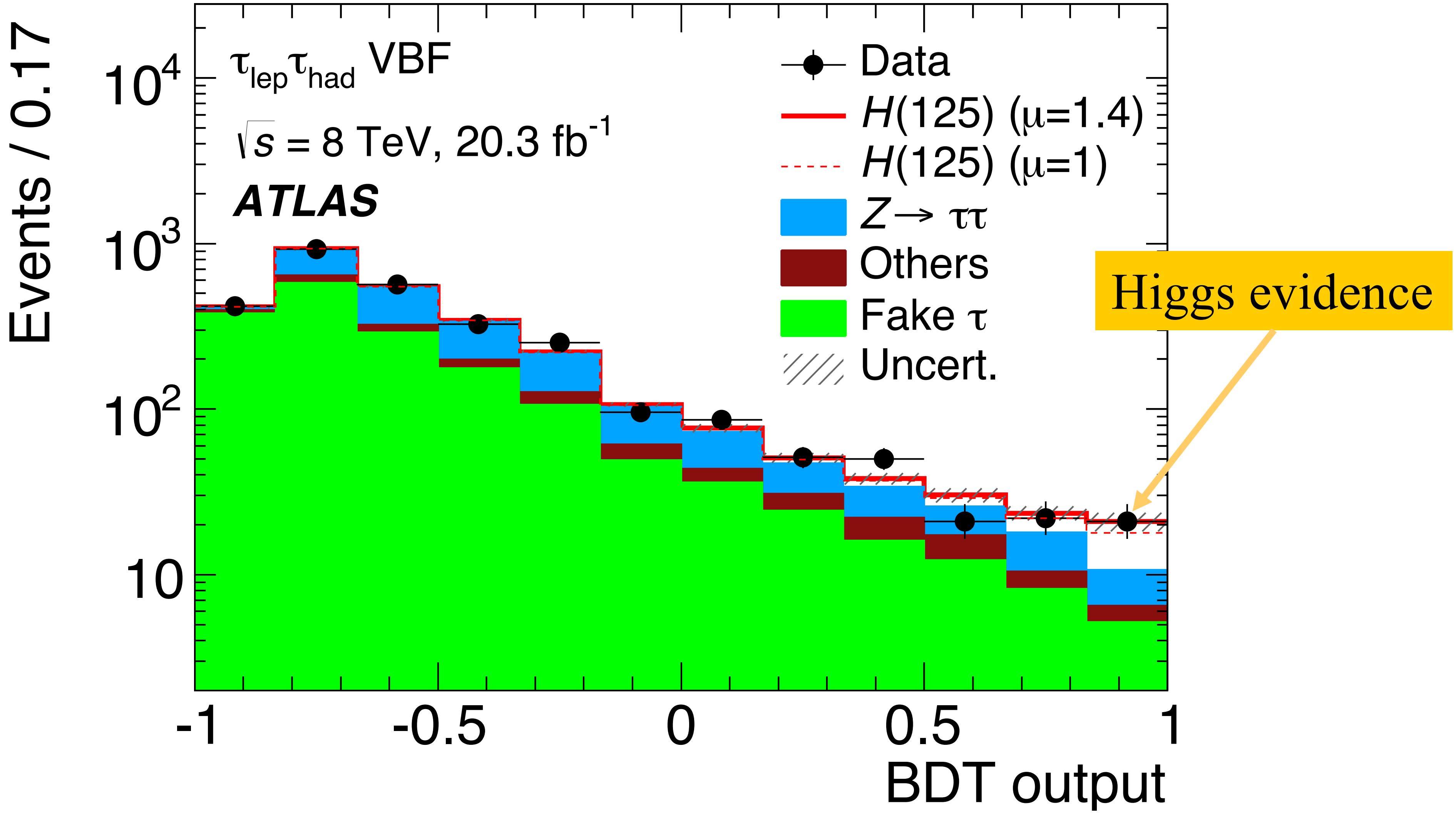




Evidence using a classifier

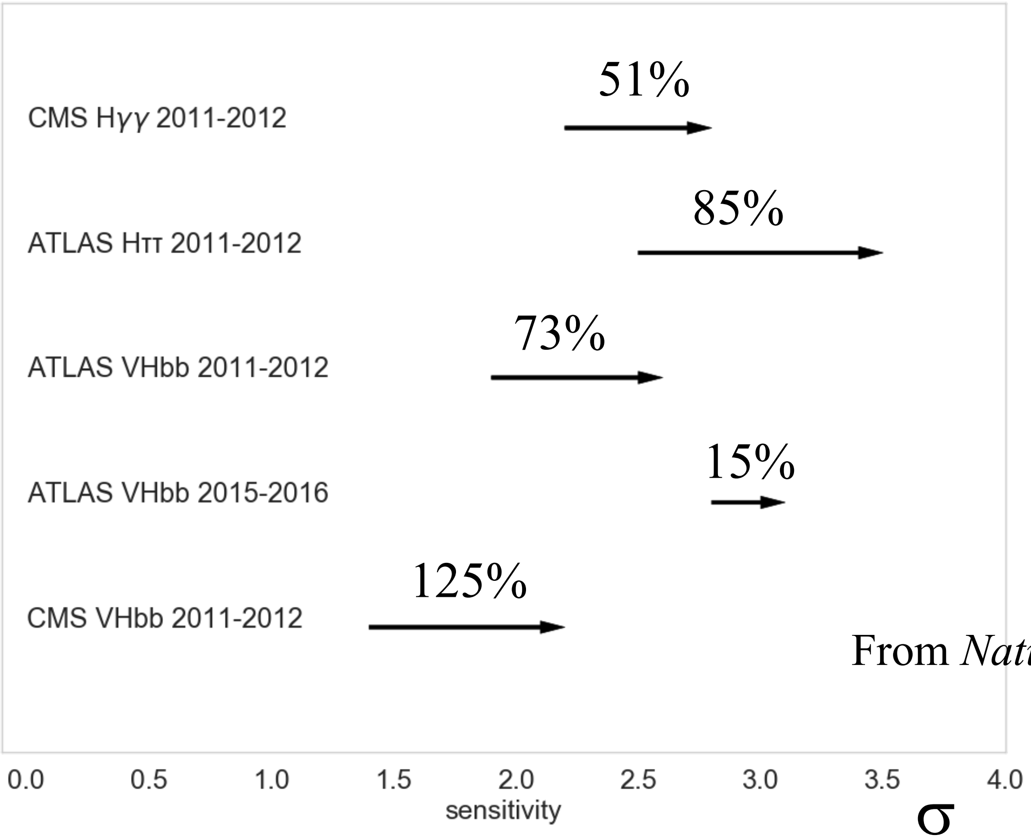
JHEP 04, 117 (2015) 1501.04943

Boosted Decision Tree using ~dozen of high level variables built from final state 4-momentum



ML on Higgs Physics

- At LHC, Machine Learning used almost since first data taking (2010) for reconstruction and analysis
- In most cases, Boosted Decision Tree on ~10 variables
- For example, impact on Higgs boson sensitivity at LHC:

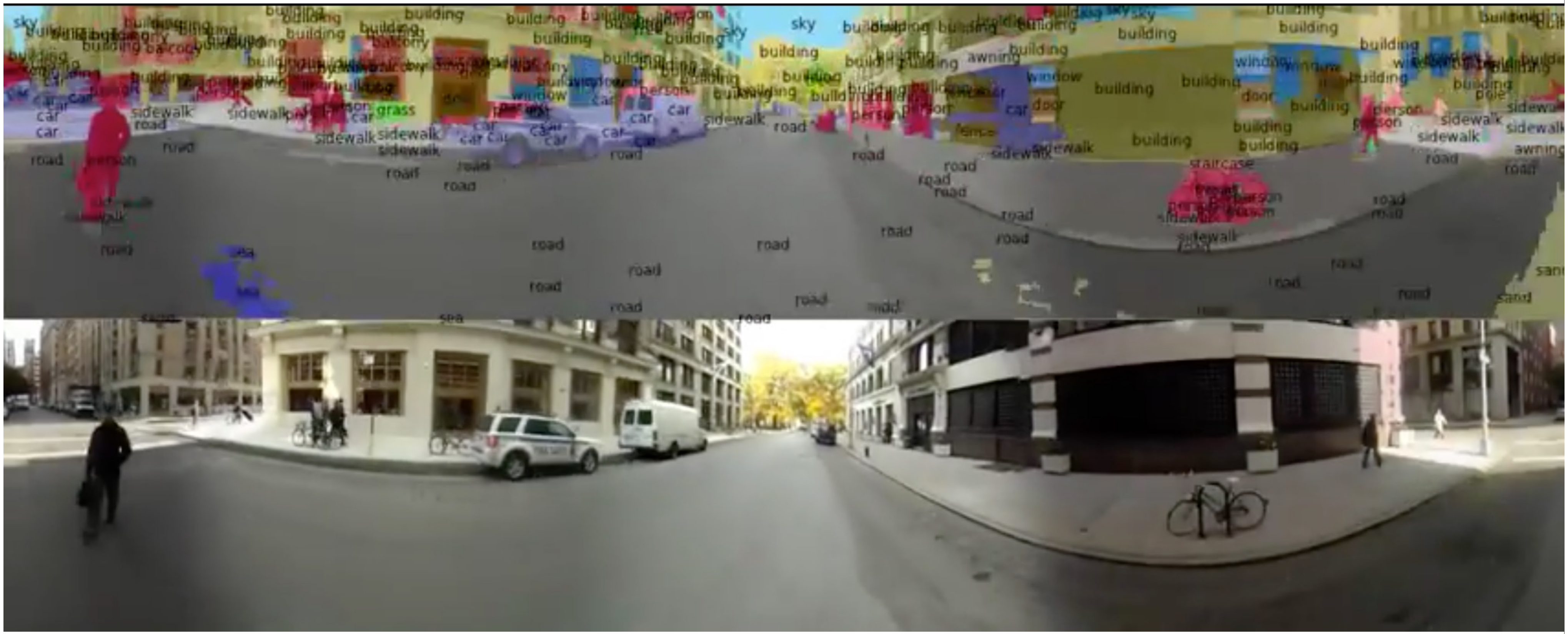


From *Nature* **560**, 41–48 (2018) |

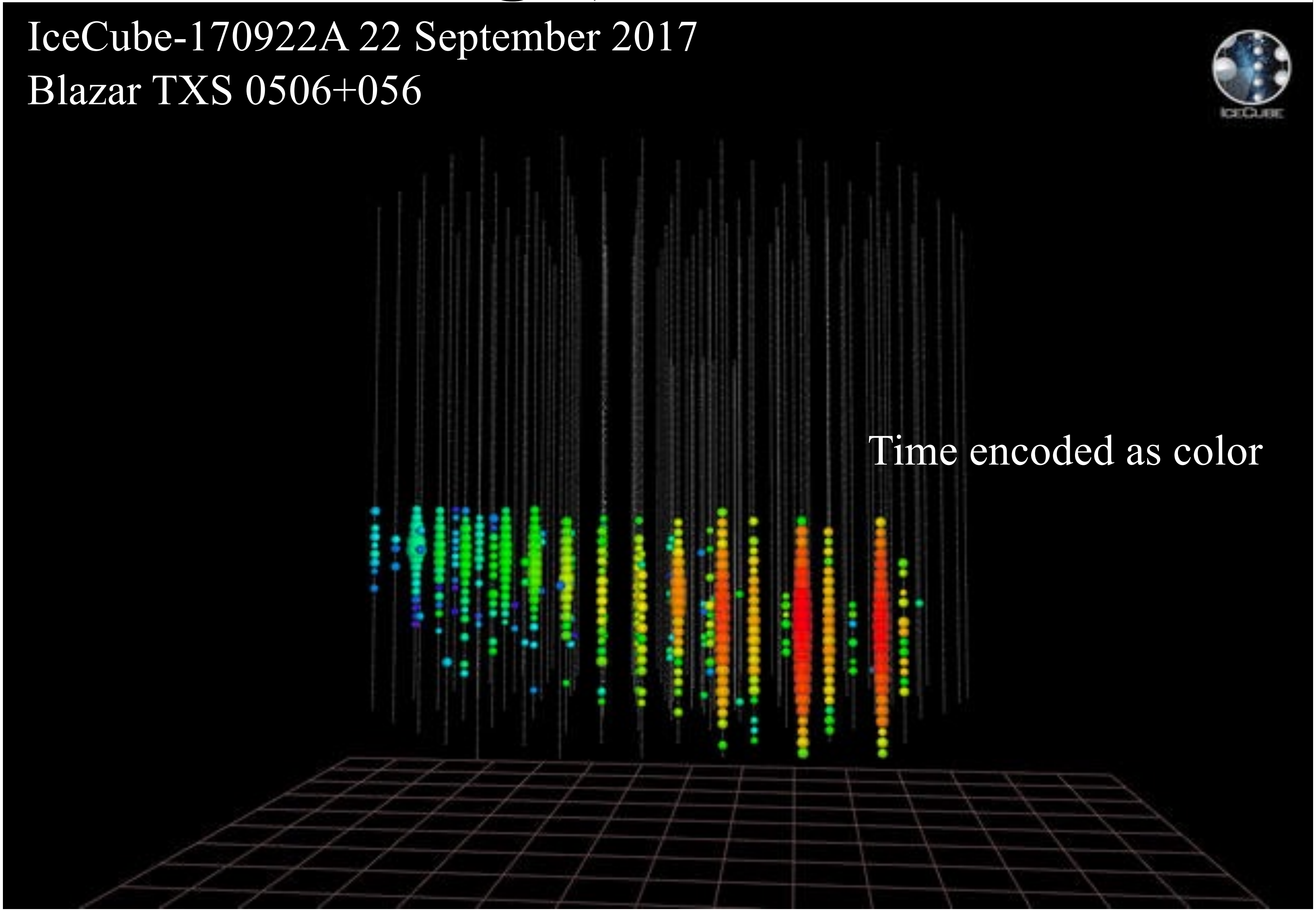
➔~50% gain on LHC running

High Energy Physics data are not images

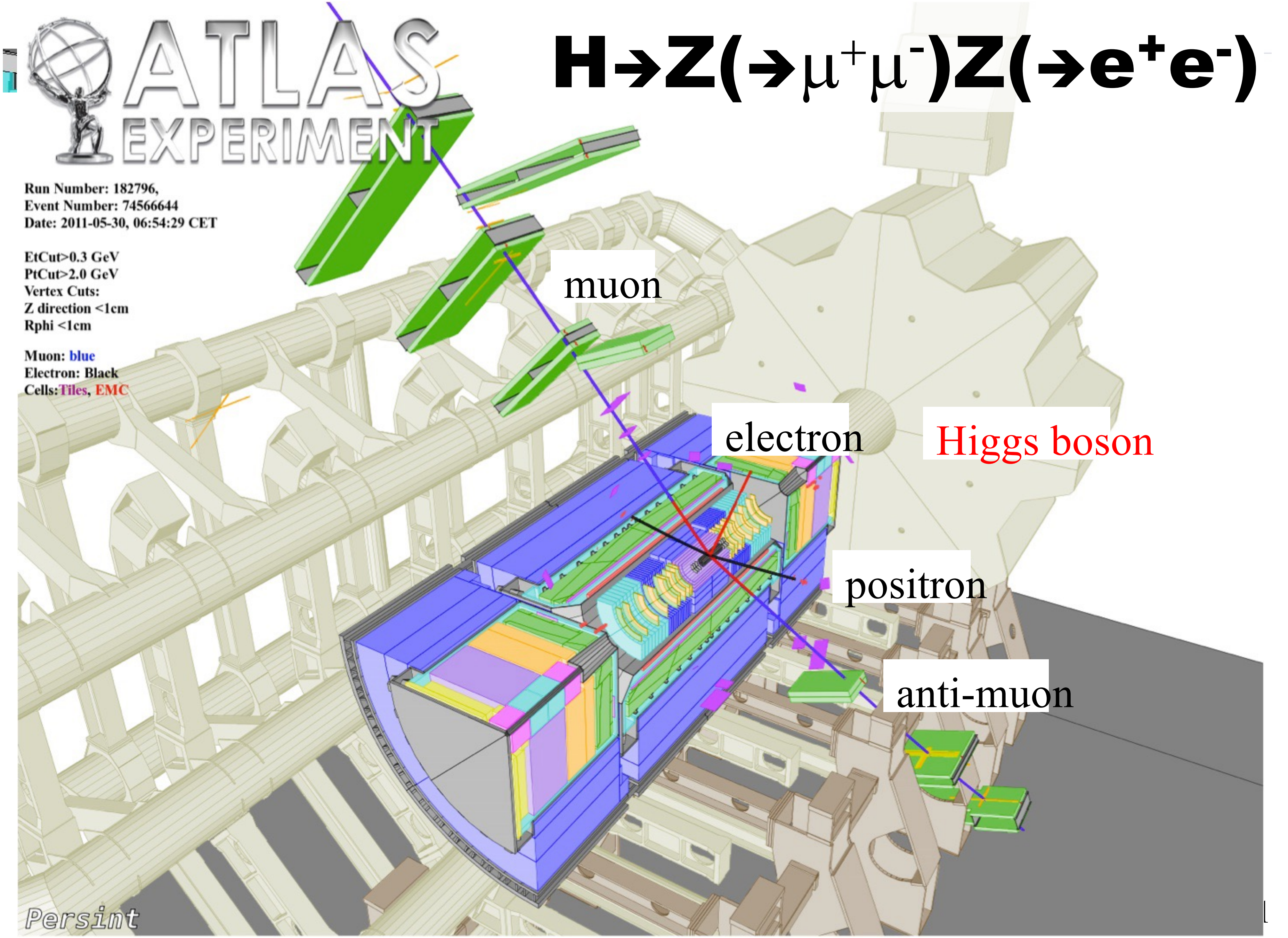
Typical Deep Learning application



An image, not the data

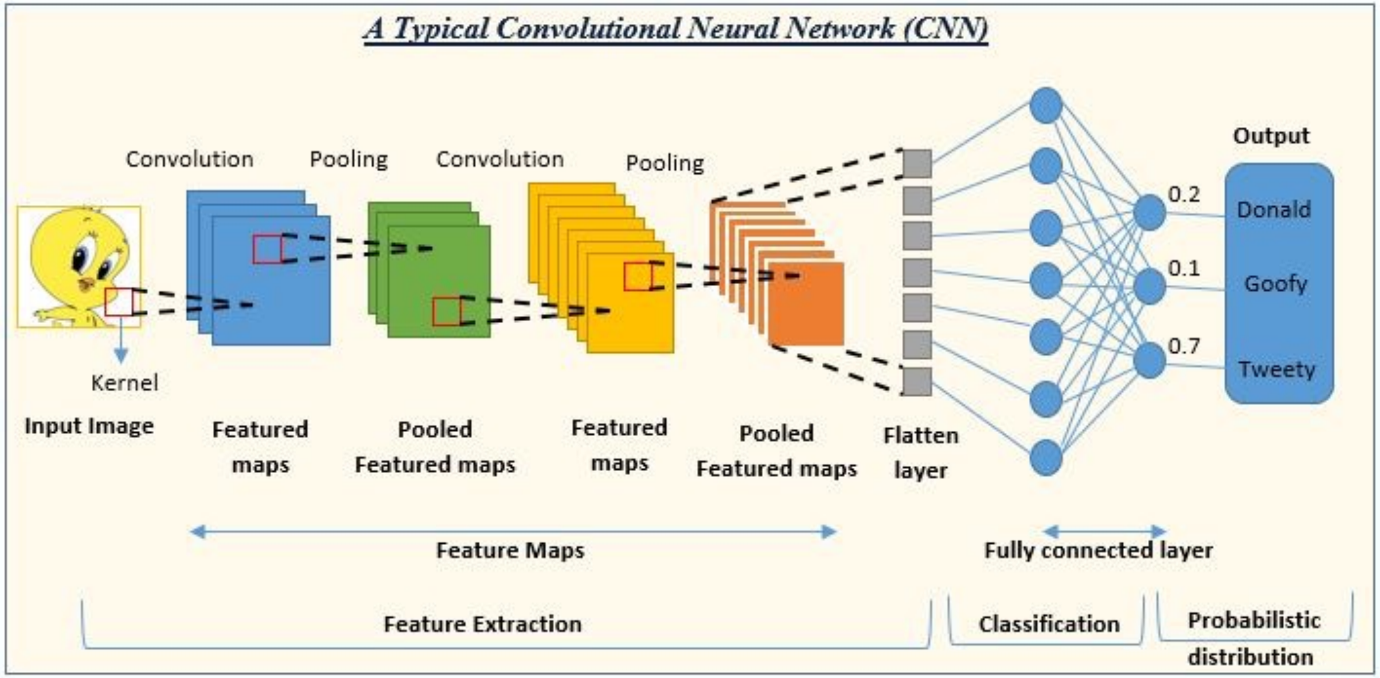


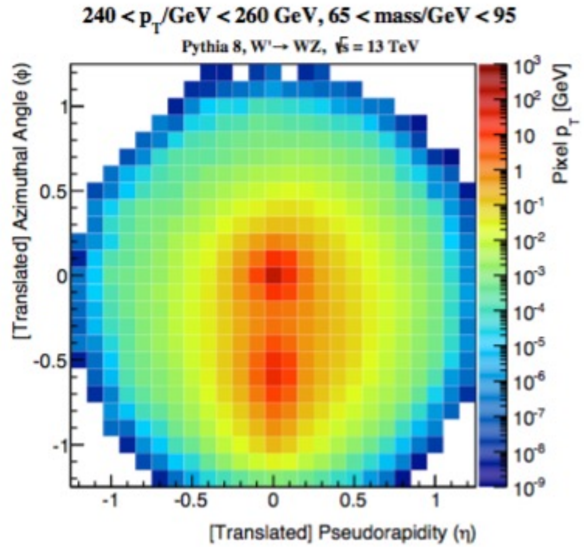
$H \to Z(\to \mu^+\mu^-)Z(\to e^+e^-)$



Jet Images with CNN

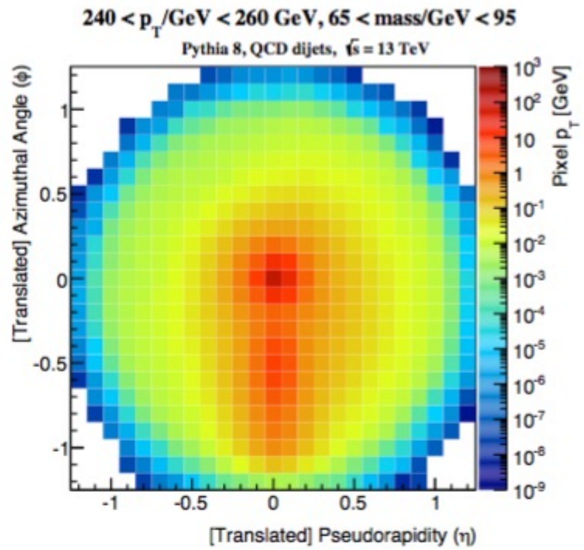
arXiv 1511.05190 de Oliveira, Kagan, Mackey, Nachman, Schwartzman





Average images

QCD

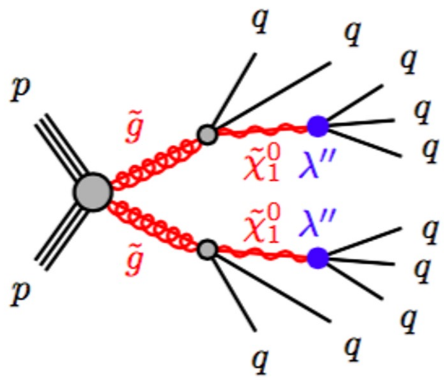


- Early attempt at image-like simulation
- ➔promising results, but not really applicable

End to end learning

Bhimji et al, 1711.03573

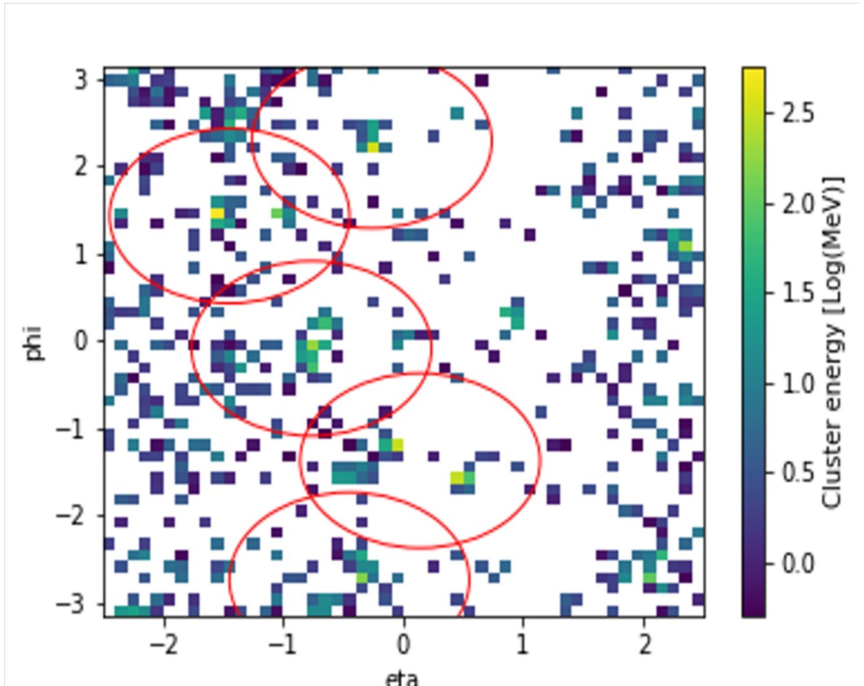
- Train directly for signal on « raw » event ?
- Start from RPV Susy search
ATLAS-CONF-2016-057
- Fast Simulated events with Delphes



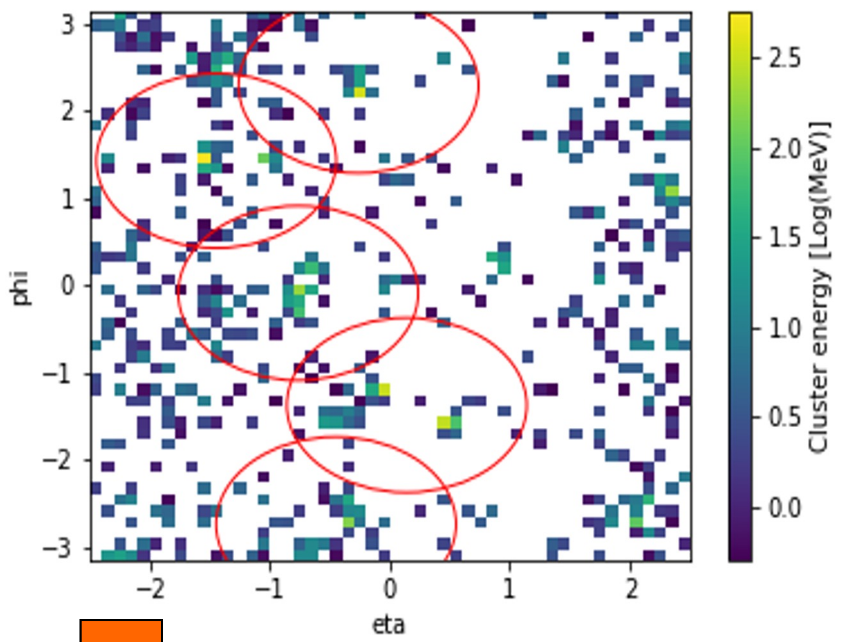
(b) gluino cascade decay

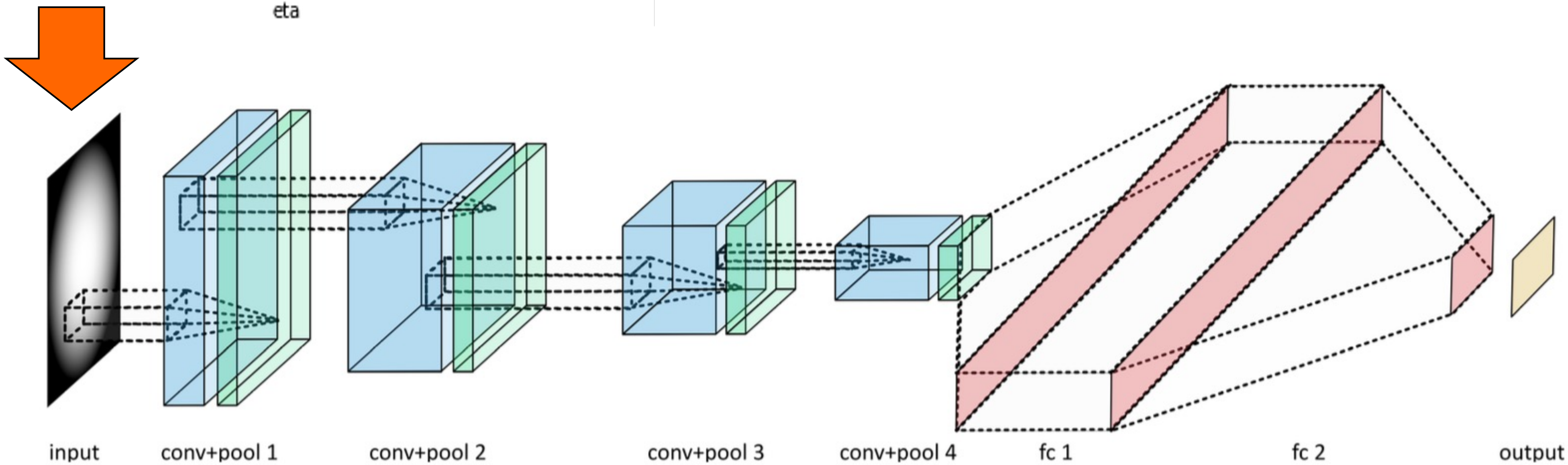
- Project energies on 64x64 ηxϕ grid
- Compare with usual jet Reconstruction and physics Analysis variables such as:

$$M_J^{\Sigma} = \sum_{\substack{p_T > 200\,\mathrm{GeV} \\ |\eta| \leq 2.0}}^{4} m^{\mathrm{jet}}$$

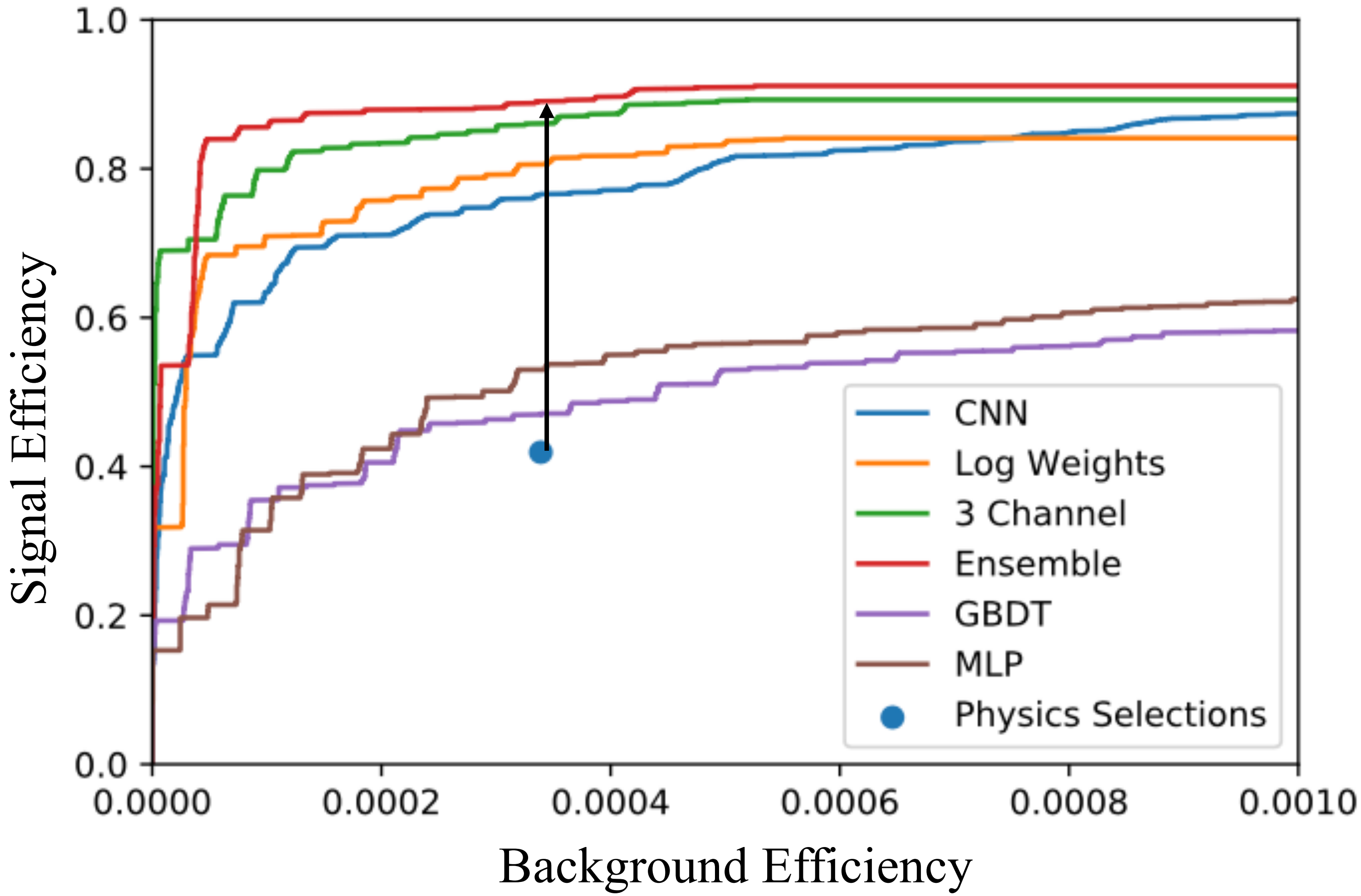


End to end learning (2)



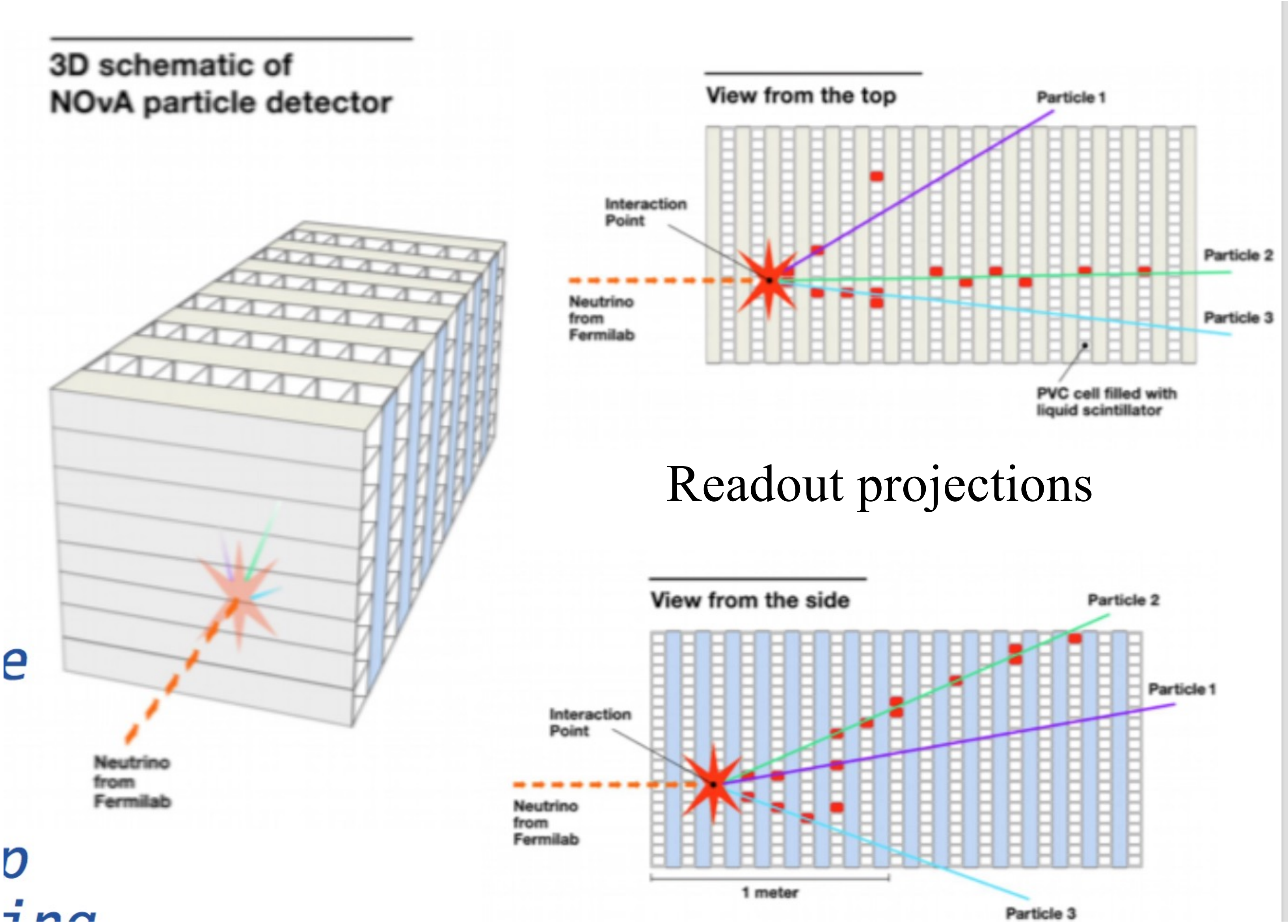


End to end learning (3)



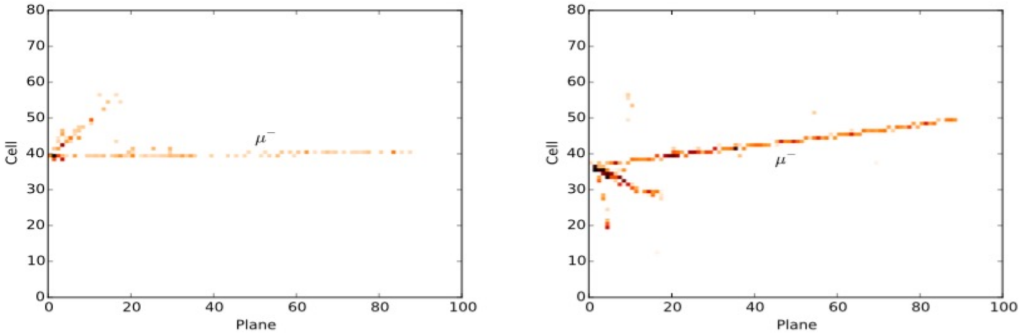
- >x2 gain over BDT/shallow network using physics variable and 5 leading jet 4-momenta
- ➔CNN extract information from energy grid which is lost in the jets ?

An exception : NOVA



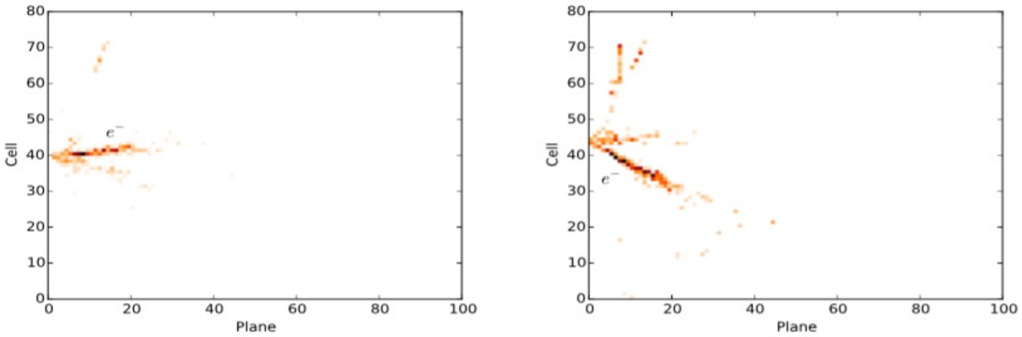
An exception: NOVA

arXiv 1604.01444 Aurisano et al



X-view Y-view

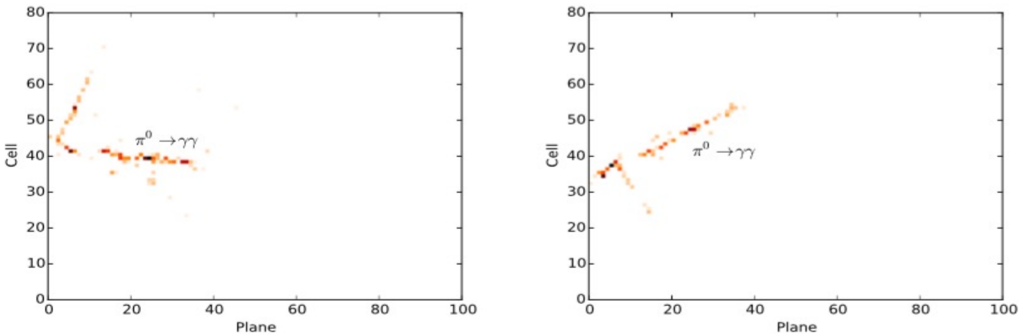
(a) ν_μ CC interaction.



X-view Y-view

(b) ν_e CC interaction.

40% ε improvement



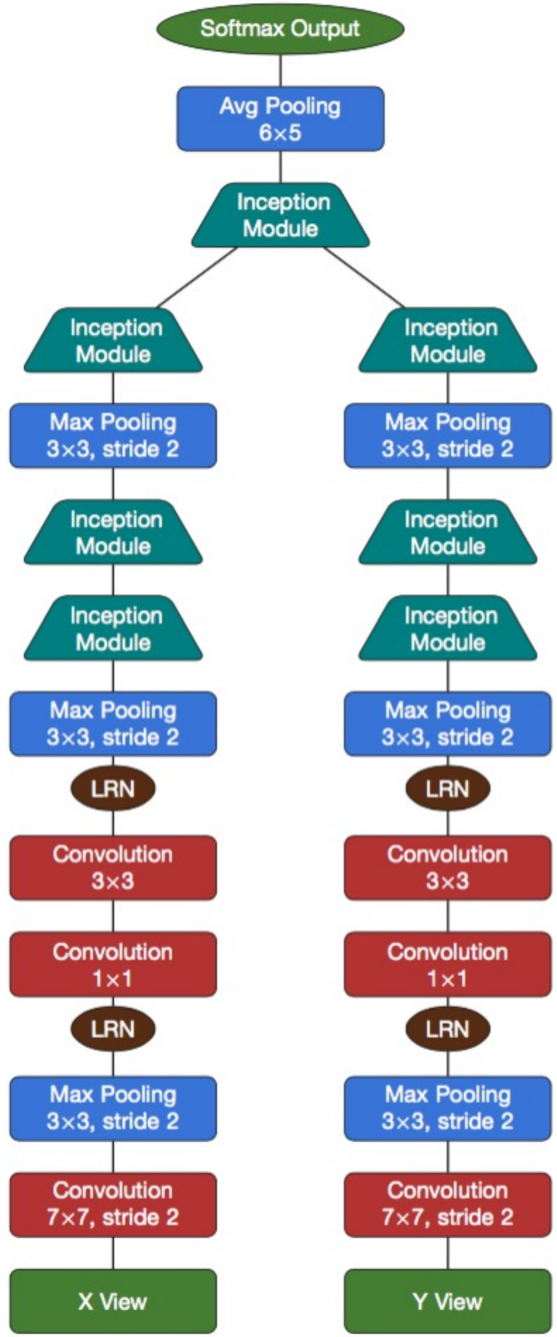
X-view Y-view

(c) NC interaction.

Neutrino interaction classification

Using Convolutional Neural Network (GoogleNet)

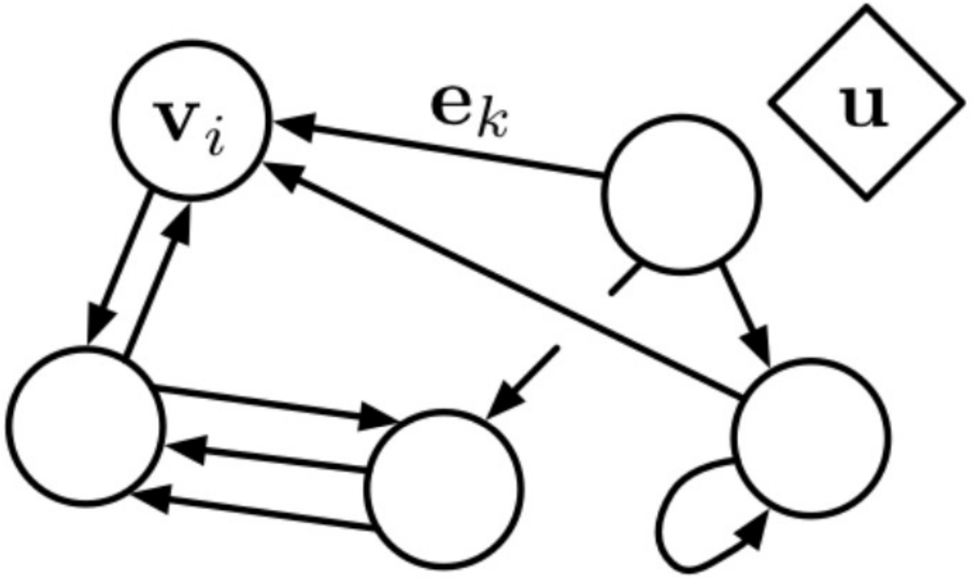
Actually used in physics results 1703.03328 and 1706.04592



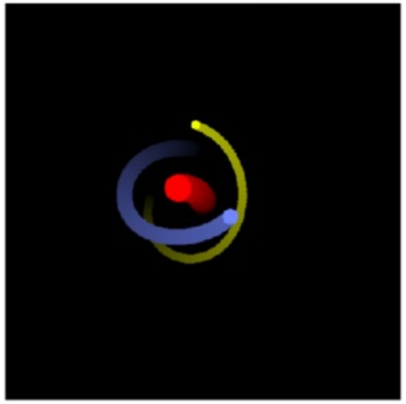
Graph Networks

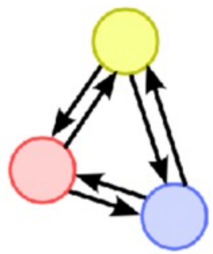
GNN

- Now some structure:
  - v_i : nodes
  - e_k : edges
  - u : global



n-body





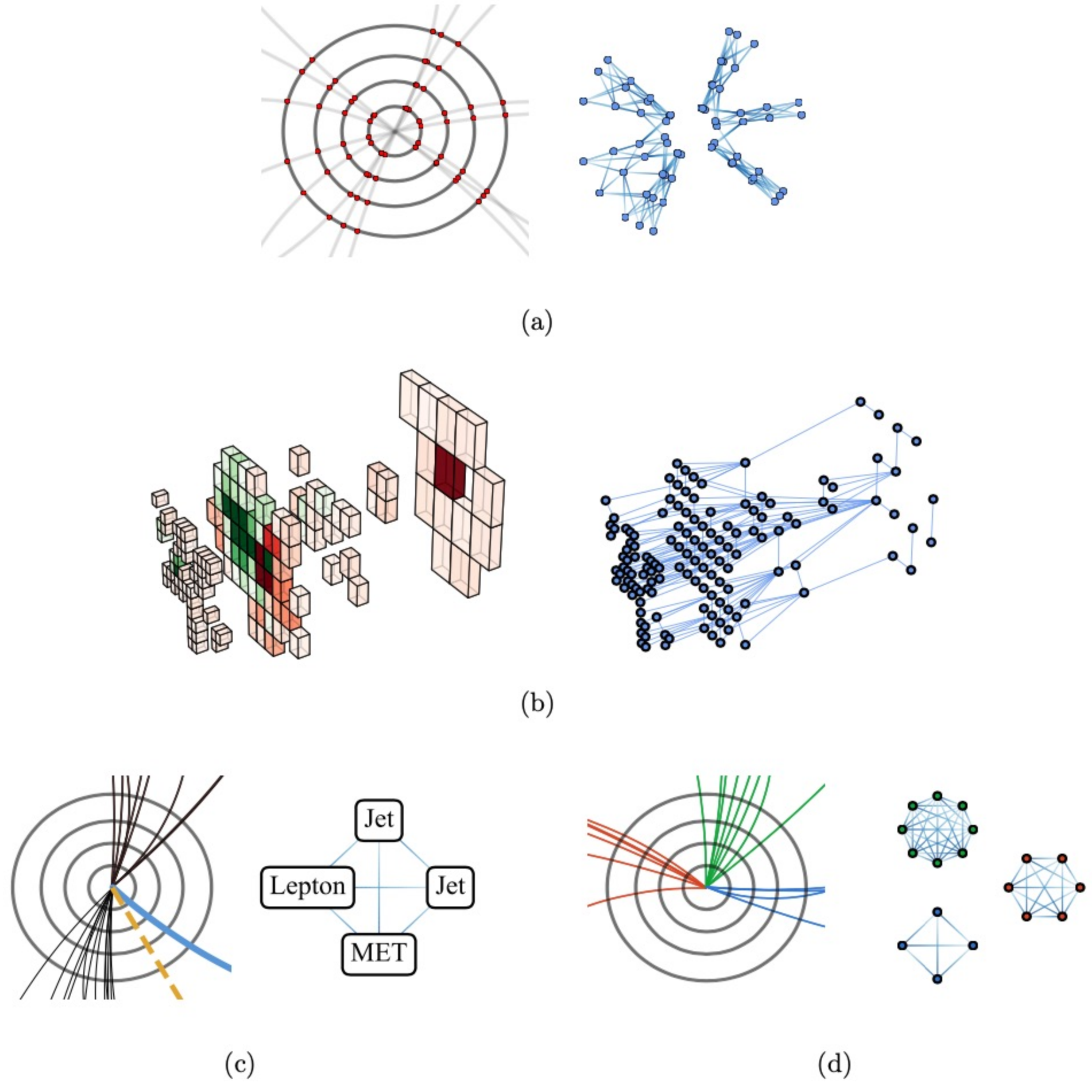
Nodes: bodies

Edges: gravitational forces

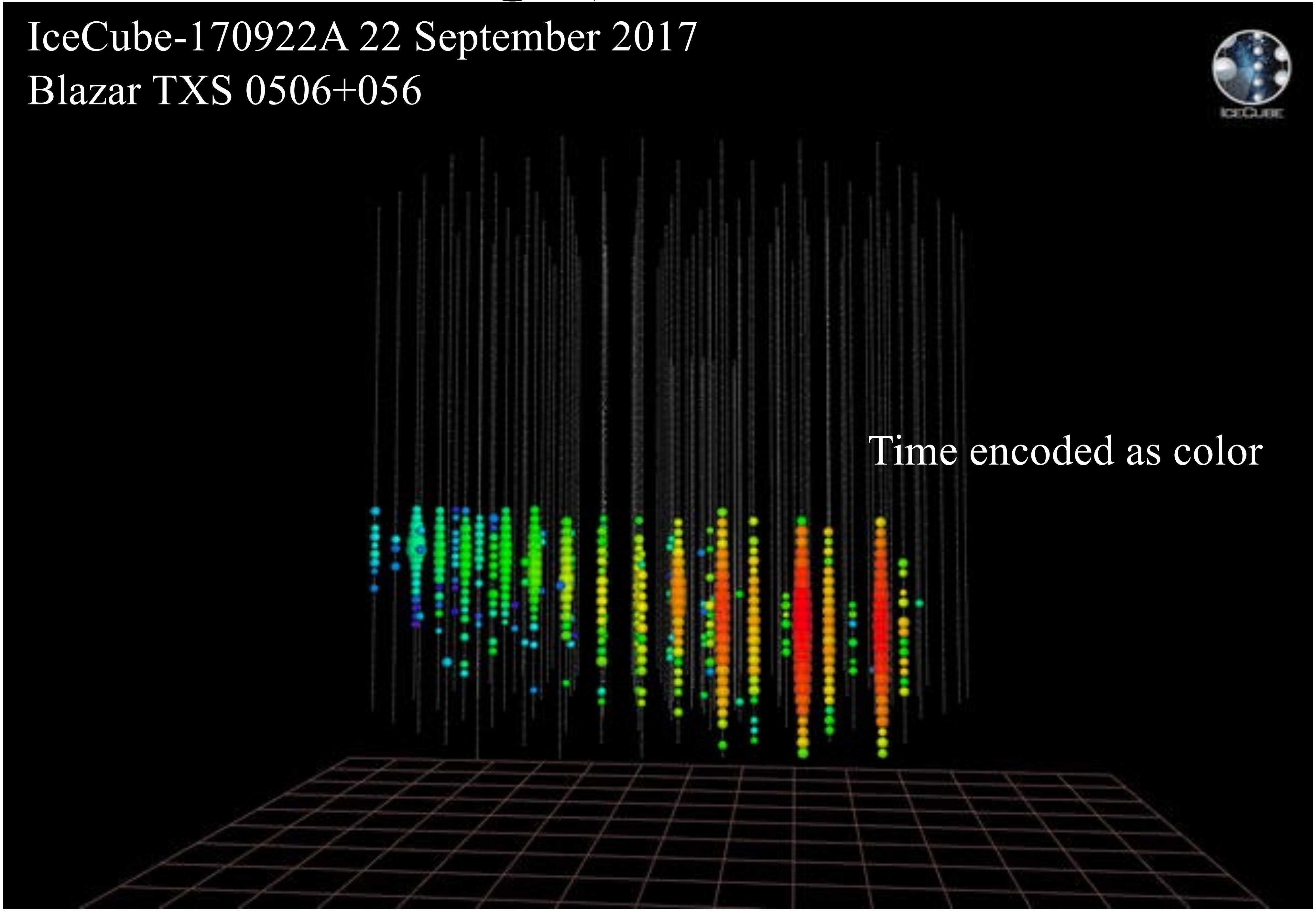
Global : potential energy

Graph on HEP data

from 2007.13681



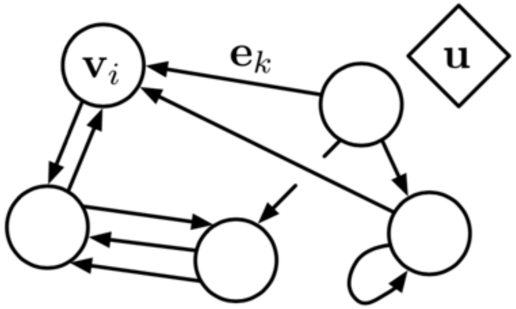
An image, not the data



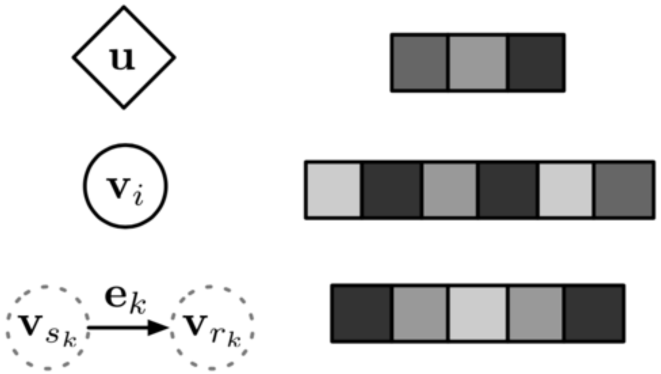
Graph NN for Ice Cube

arXiv:1809.06166

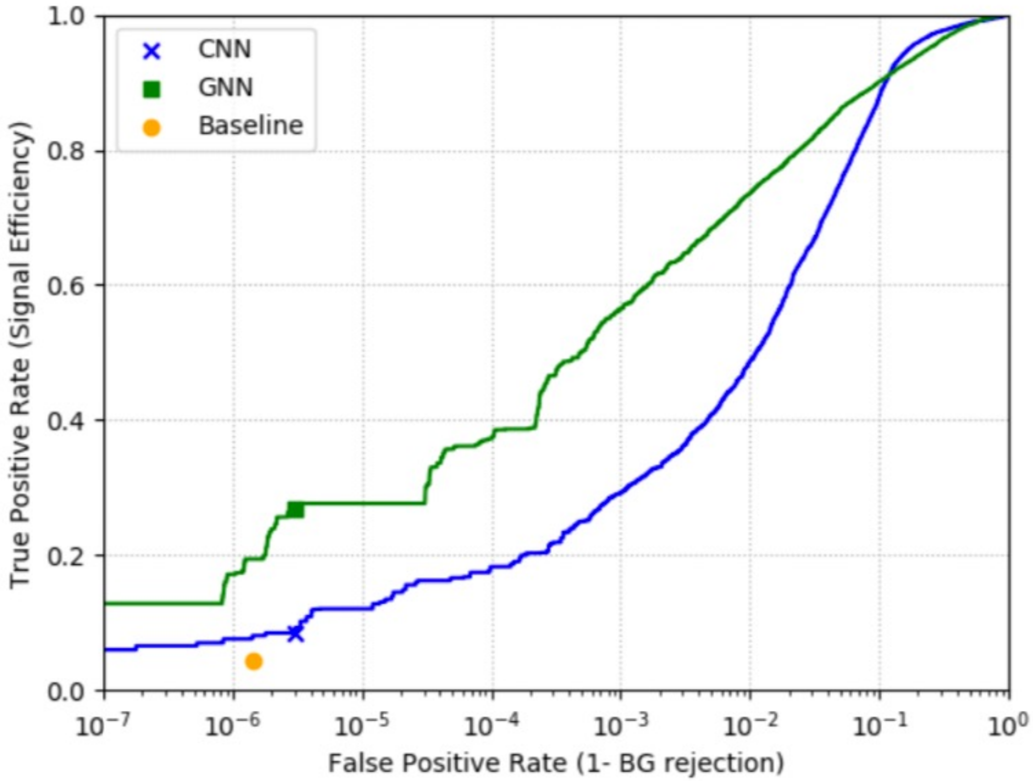
- Graph NN: nodes,edges,and Globals
- ...allow generalization of neighbouring pixels



Jessica Hamrick

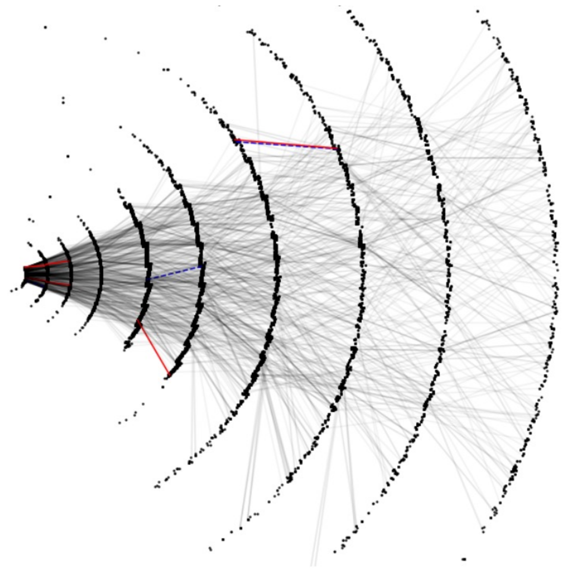


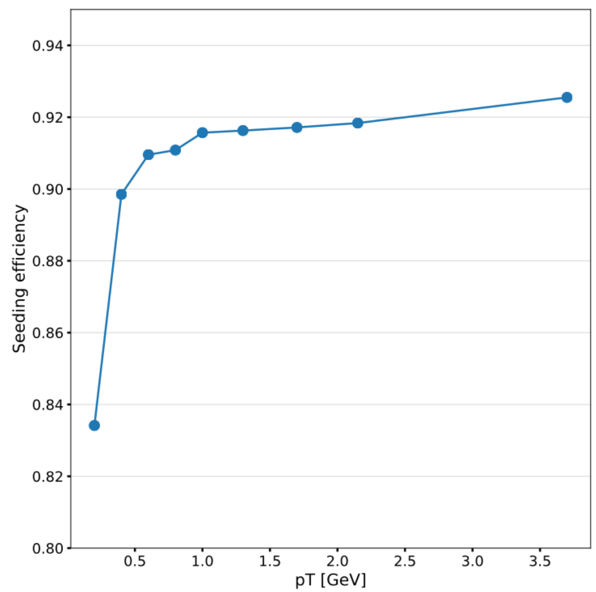
- Application to IceCube, separating downwards muon from neutrino from muon from cosmic rays
- ➔quickly growing interest in Graph NN in HEP



Track Seeding with GNN

2007.00149



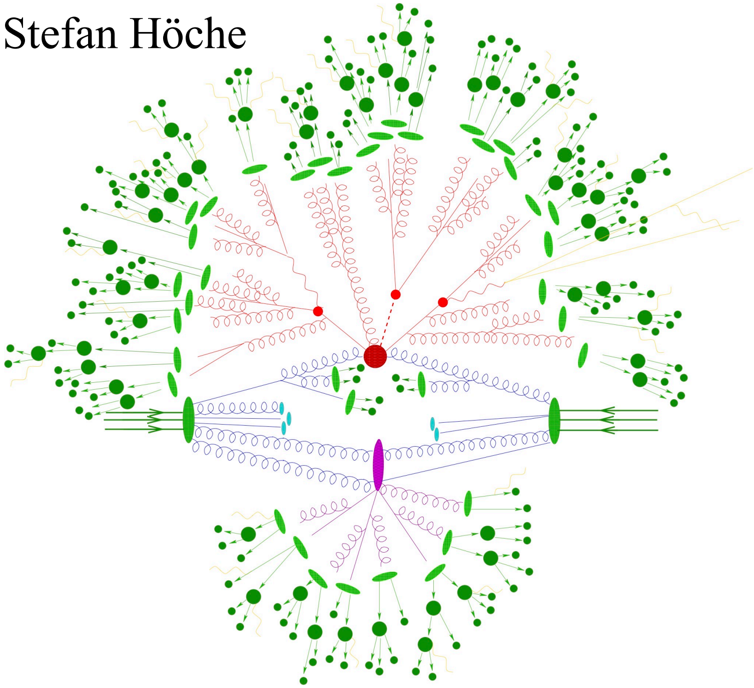


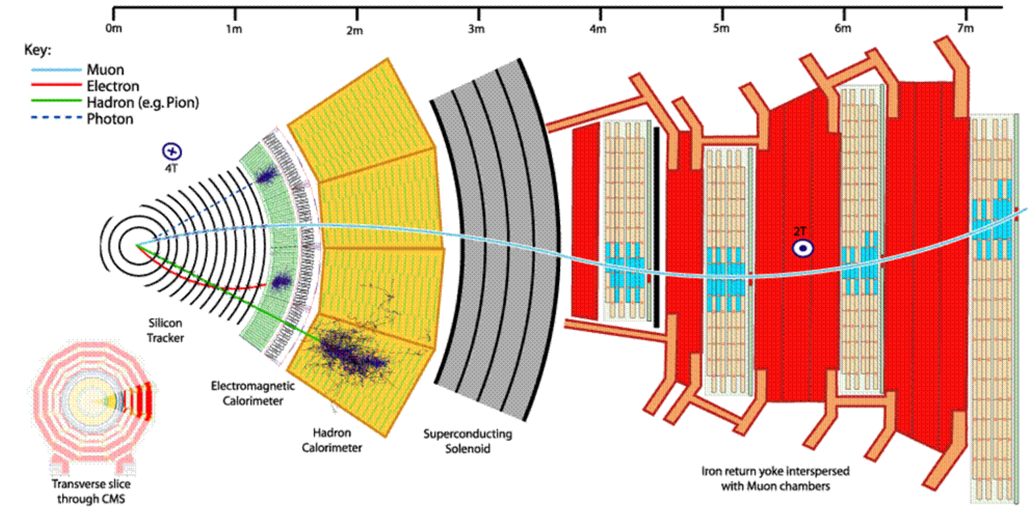
- Tracking for High Lumi – LHC, ATLAS/CMS
- Build edges between neighbour
- Then GNN trained to classify double and triplet
- High efficiency reached with subsecond computing time (also very parallelisabled)
- ➔can be used as a filtering stage before traditional Kalman filter

Generative Adversarial Network (GAN) /Variational AutoEncoders (VAE) to accelerate simulators

Accurate simulators

Stefan Höche





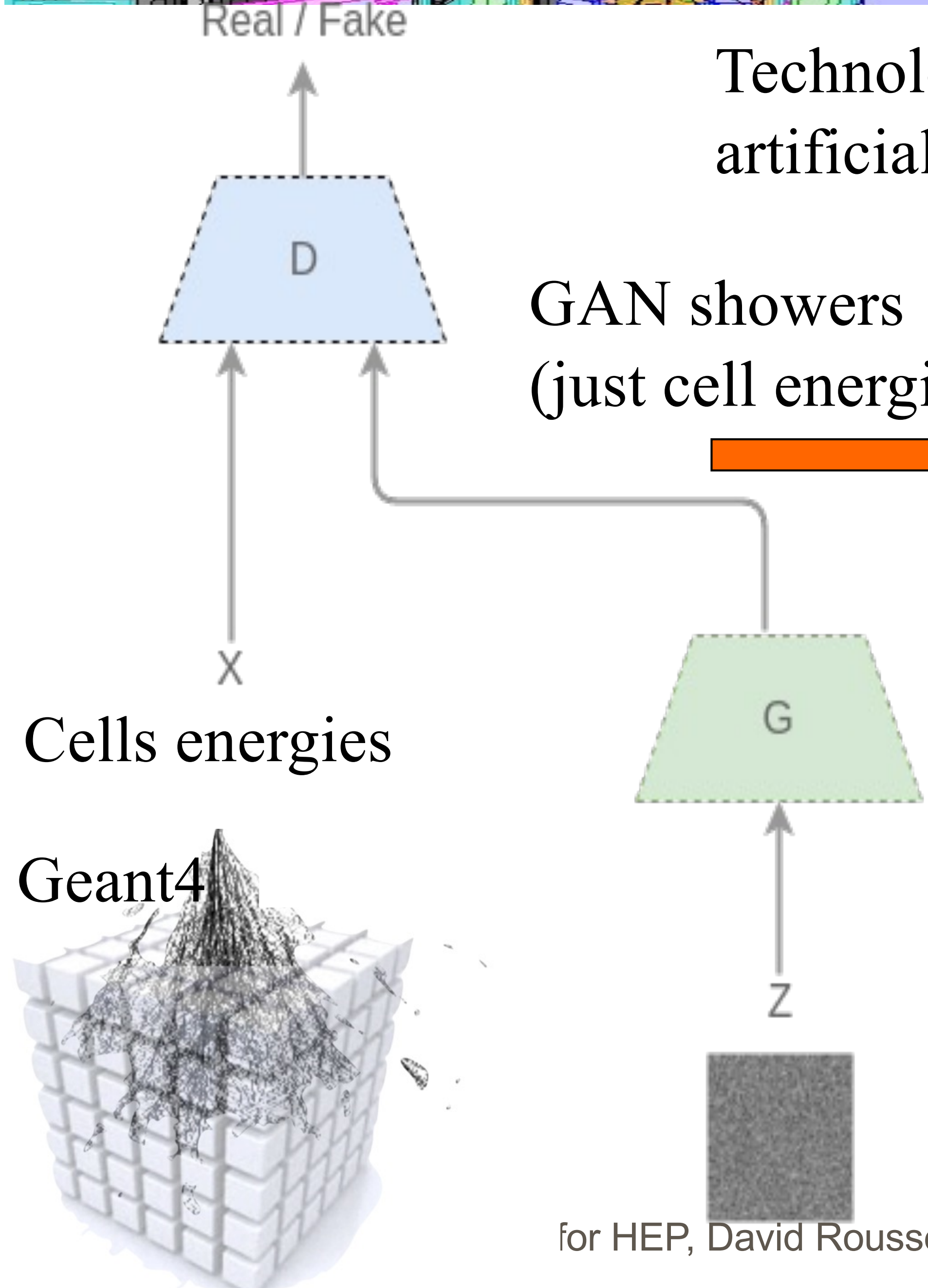
- Interaction of particles with the detector

- Proton collision

- ➔data very similar to real data from the experiment
- + ground truth
- This has been in HEP culture since the seventies, and developped through huge efforts in resource and manpower

GAN for simulation

Technology not unlike artificial face generation



GAN showers (just cell energies)

Cells energies

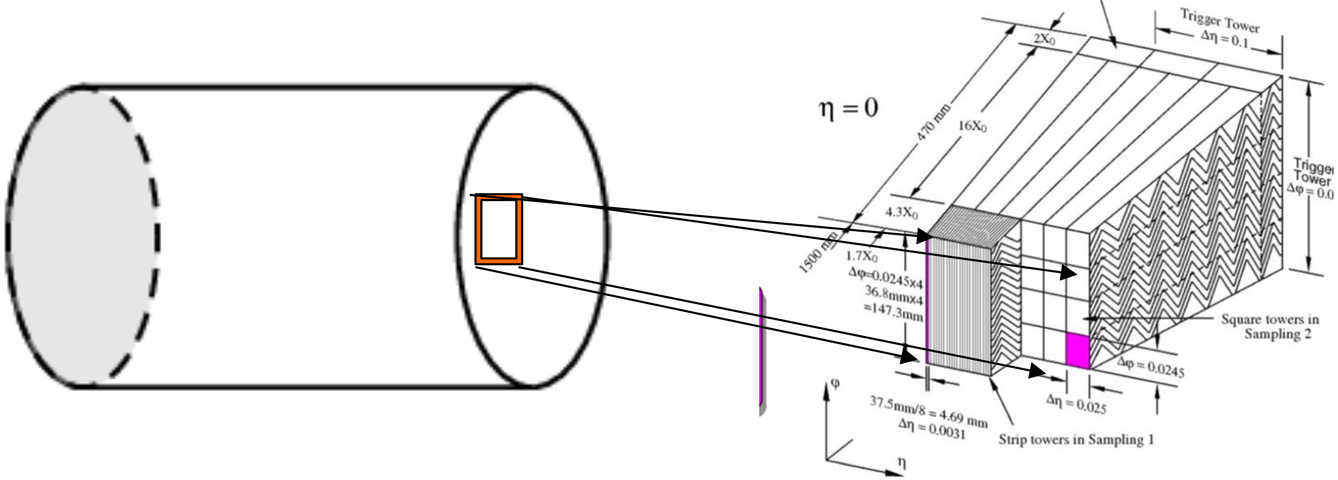
Geant4

- Half of LHC grid computers (~1.000.000 cores) are crunching Geant4 simulation 24/24 365/365
- …while LHC experiments are collecting more and more events
- ➔reducing CPU consumption of simulation is very important

- Geant4 is the standard in particle simulation
- Imagine training a GAN on single particle showers of all types and energies
- Then when an event is simulated it would ask for GAN showers on request (superfast by 3-4 order of magnitude)

ATLAS calo simulation

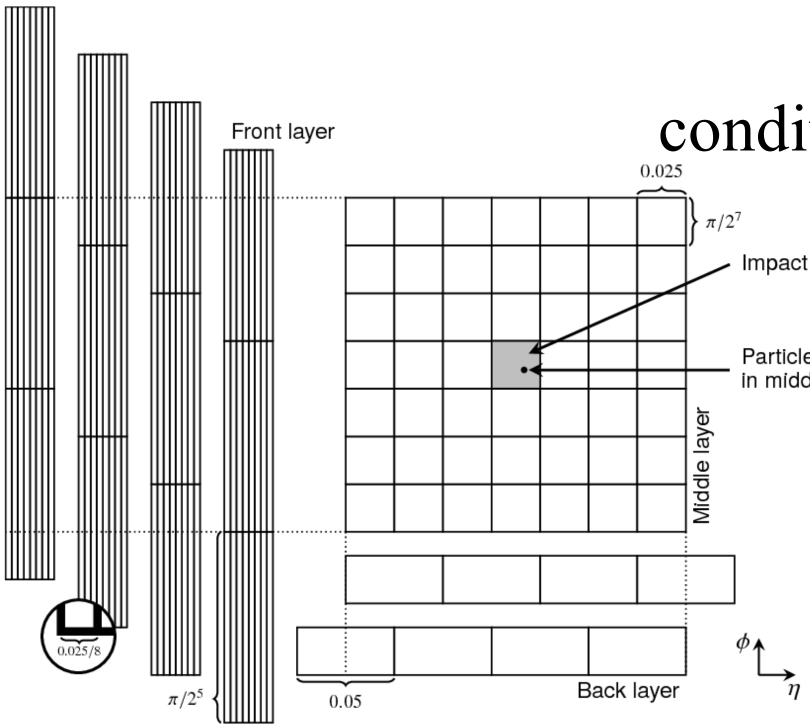
arXiv:2210.06204

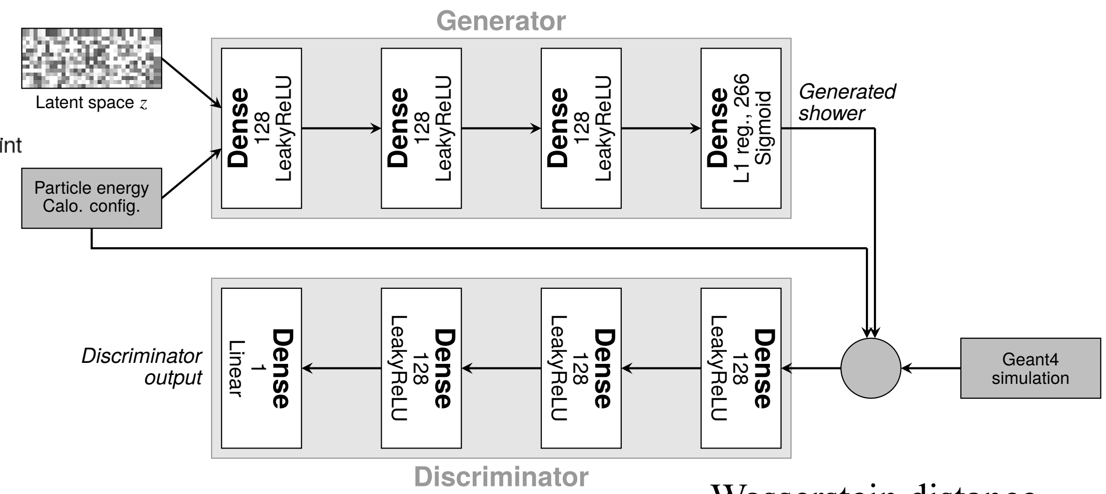


+ η, ϕ translation

177000 cells ➔266 cells

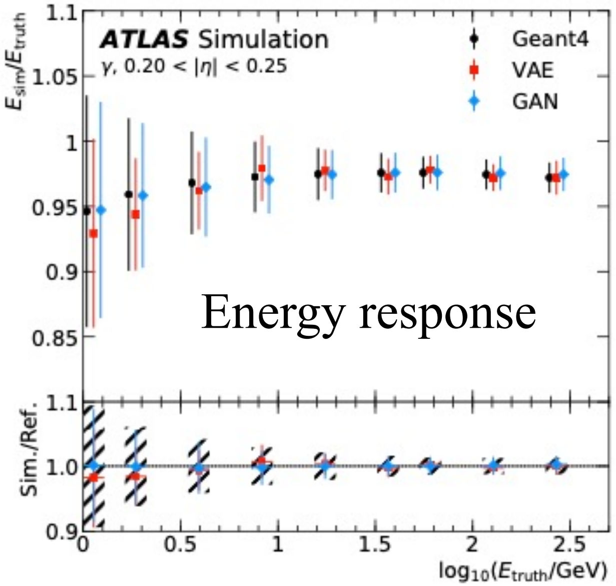
conditions

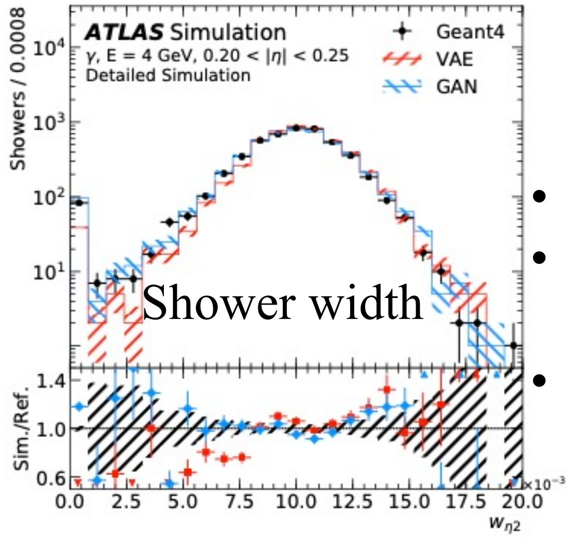




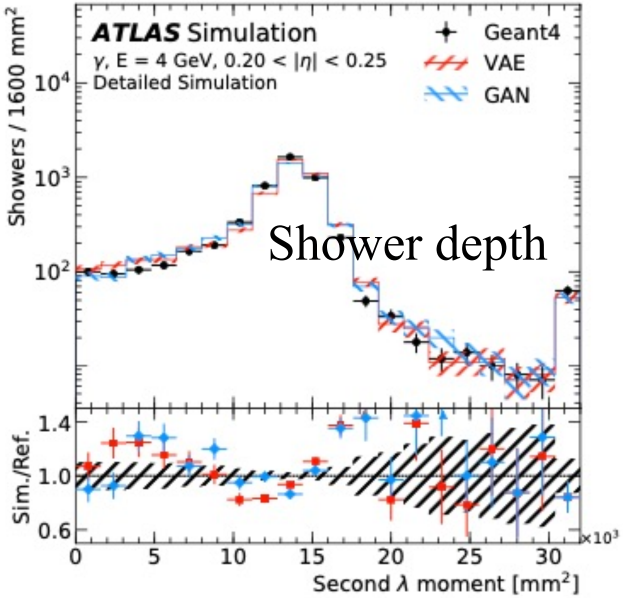
Wasserstein distance

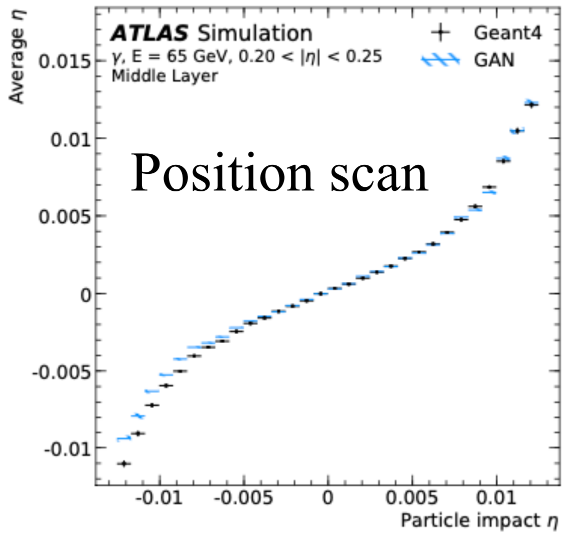
Results





- Speed: <1ms compared to 10s
- Sufficient accuracy for physics ? Bulk and tails ?
- Handling of (more) awkward geometries ? (➔ **graph based GAN)**





Dealing with Uncertainties

Phys.Rev.Lett. 114 (2015)191803

Most complex measurement ever ?

Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments

(ATLAS Collaboration)[†]
(CMS Collaboration)[‡]

(Received 25 March 2015; published 14 May 2015)

A measurement of the Higgs boson mass is presented based on the combined data samples of the ATLAS and CMS experiments at the CERN LHC in the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decay channels. The results are obtained from a simultaneous fit to the reconstructed invariant mass peaks in the two channels and for the two experiments. The measured masses from the individual channels and the two experiments are found to be consistent among themselves. The combined measured mass of the Higgs boson is $m_H = 125.09 \pm 0.21\,(\text{stat}) \pm 0.11\,(\text{syst})$ GeV.

Dealing with Uncertainties

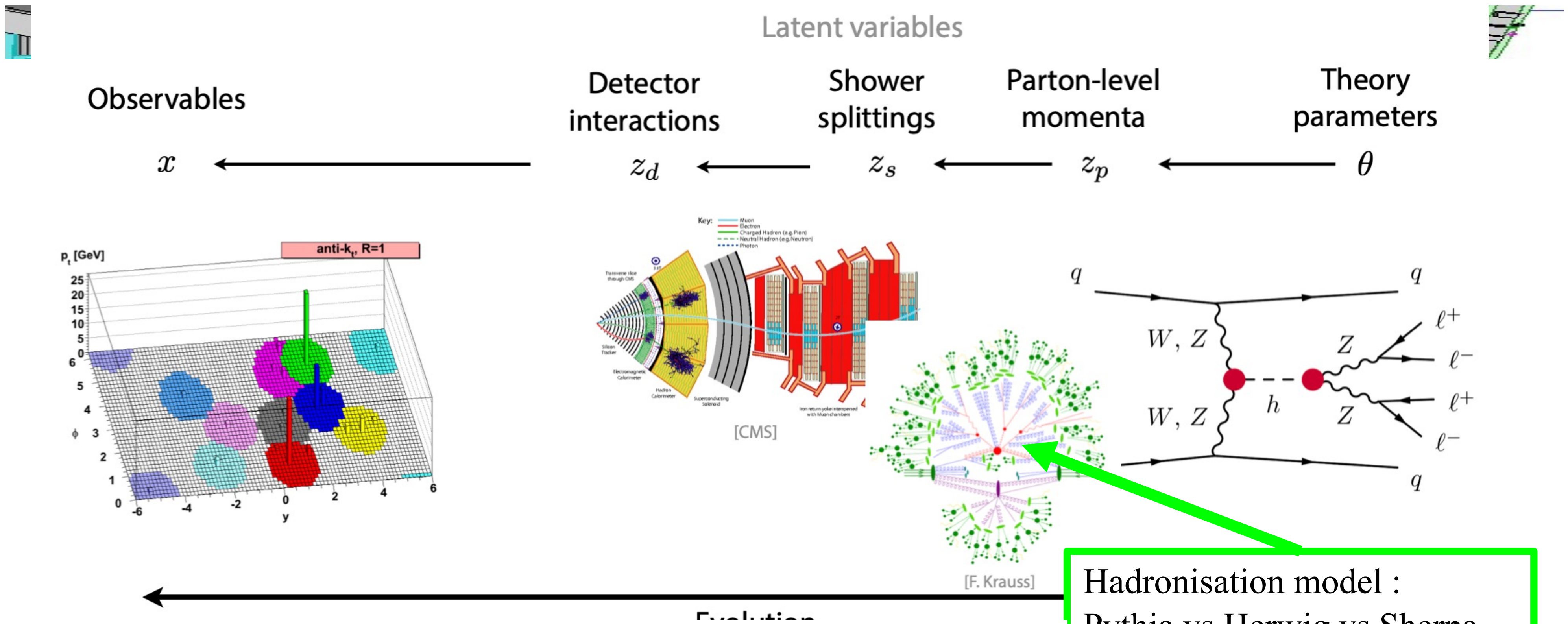
- Our experimental measurement papers typically end with
  - measurement = m $\pm$ σ(stat) $\pm$ σ(syst)
  - σ(syst) systematic uncertainty : known unknowns, unknown unknowns… Convincing oneself, co-authors, the whole community that we know what we are doing➔trust !
- Name of the game is to minimize quadratic sum of :

  σ(stat) $\oplus\sigma$(syst)
- … while ML techniques are usually trained to minimise σ(stat)
- Two challenges:

1. Maintain trust (σ(syst)) while using AI more and more
2. Include somehow (various techniques) σ(stat) $\oplus\sigma$(syst) in the loss in order to minimise overall uncertainty

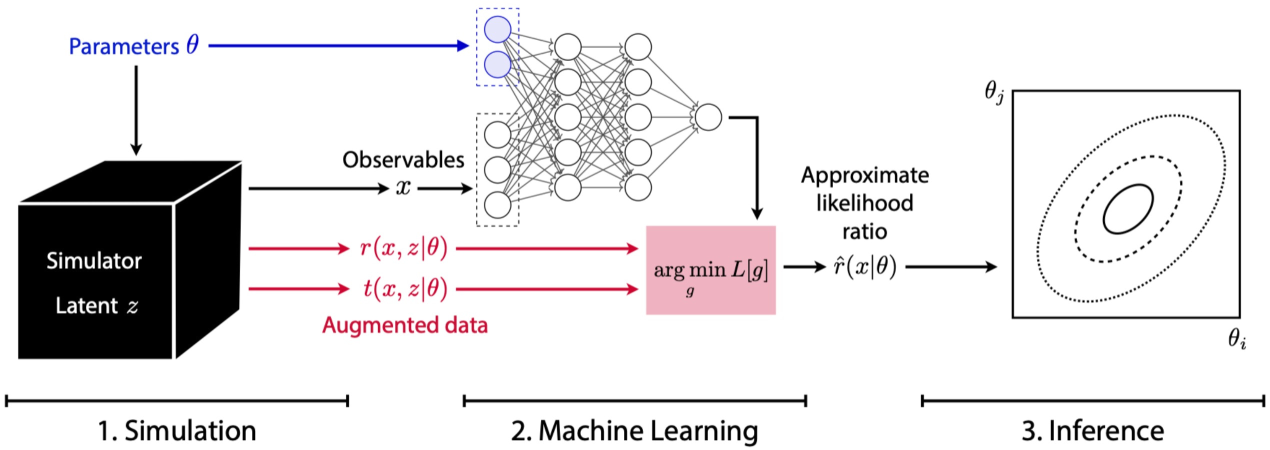
- "Uncertainty Quantification" is a fast growing field in Machine Learning

Modelling particle physics processes



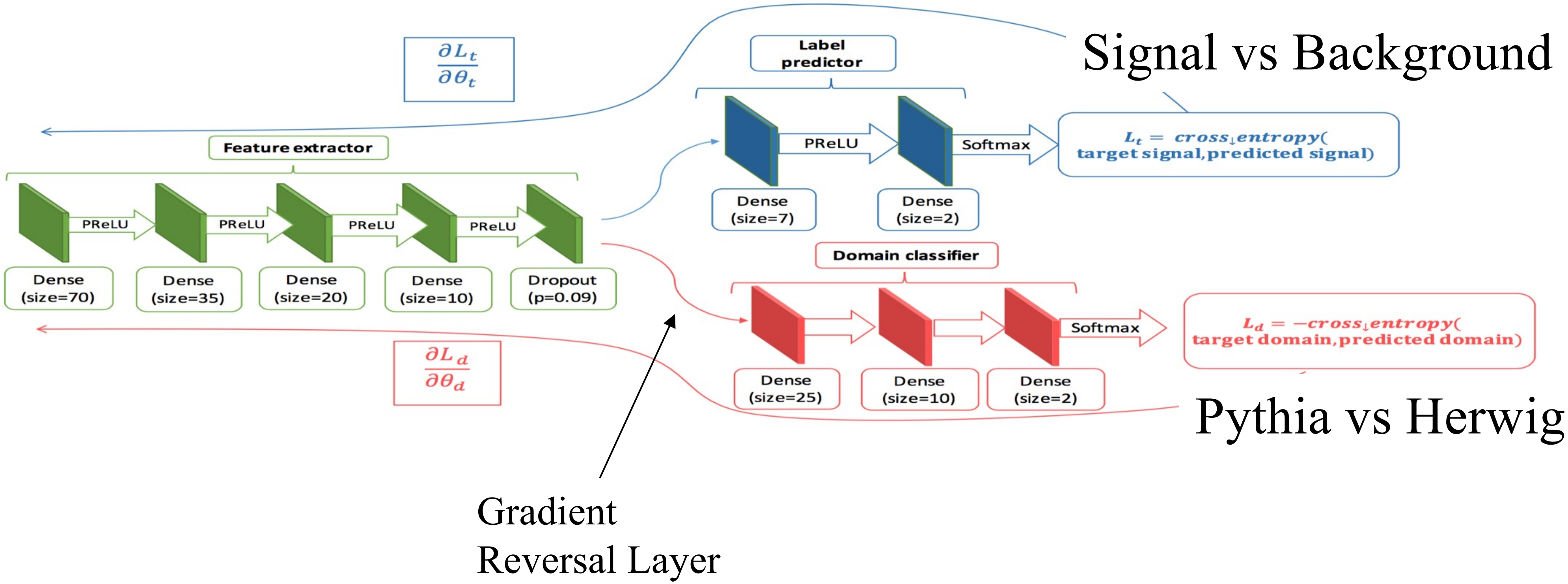
Hadronisation model :
Pythia vs Herwig vs Sherpa…

Bird's-eye view



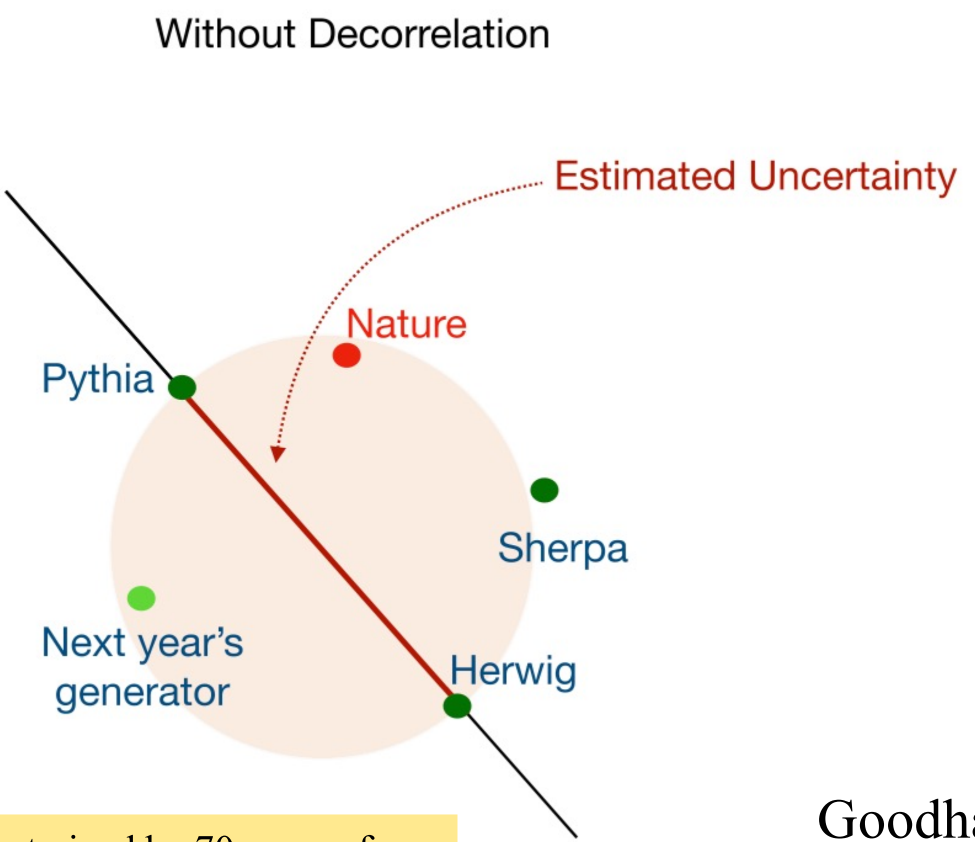
Syst Aware Training: adversarial

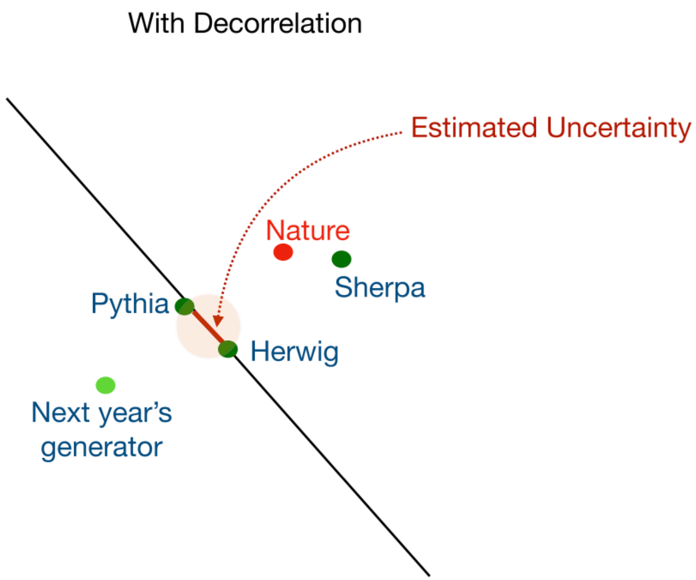
- Inspired from 1505.07818 Ganin et al :



Cautionary tale

Ghosh & Nachman EPJC **82** 46 (2022)





Constrained by 70 years of Particle Physics measurements

Goodhart's law "When a measure becomes a target, it ceases to be a good measure"

Conclusion

- We (in High Energy Physics) are analysing data from multi-billion € projects➔should make the most out of it!
- Dedicated representations (often Graph NN based) are being developped to deal with our semi-structured data
- Generative Models are accelerating our existing accurate but heavy simulators
- The bottom line is always a measurement with uncertainties which sum up the trust of the community, which should be maintained