ML-Based Correction To Accelerate Geant4 Calorimeter Simulations

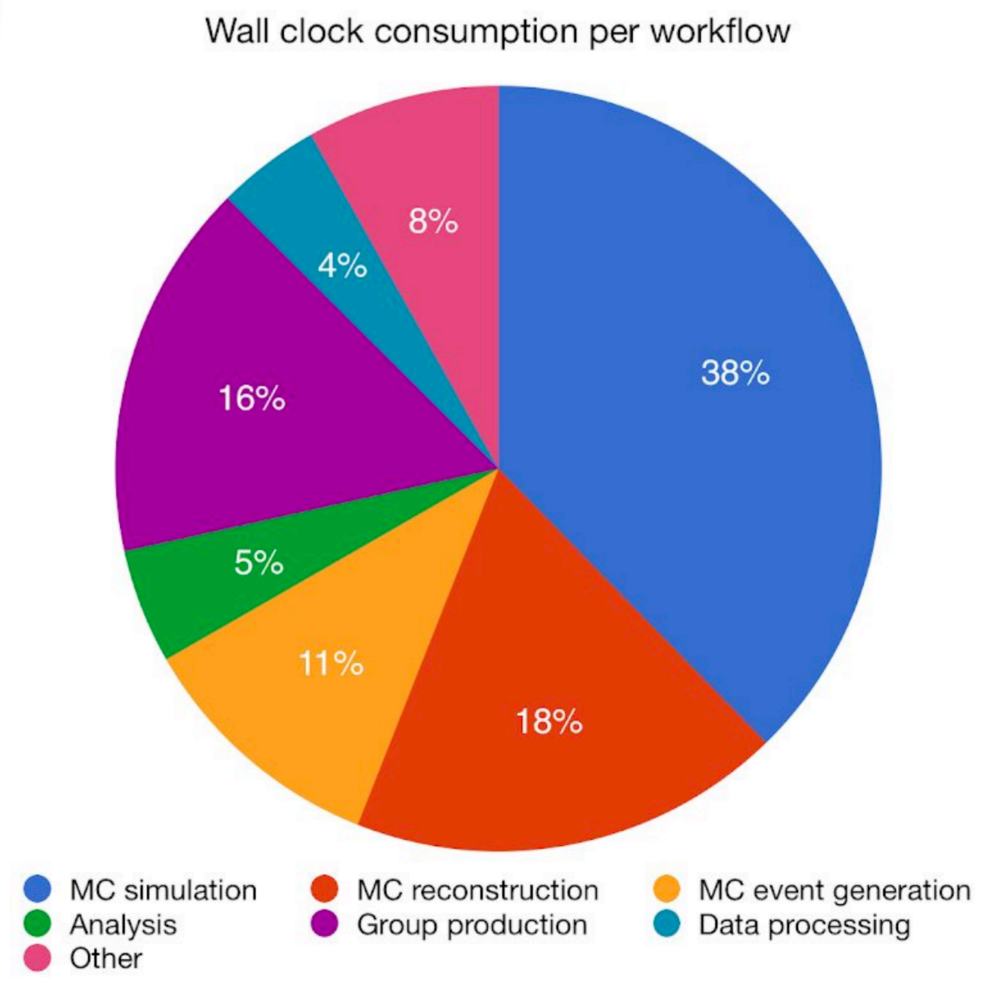
Evangelos Kourlitis, Walter Hopkins

Learning to Discover 2022 - April 29th 2022

Detector Simulations: the ATLAS Example

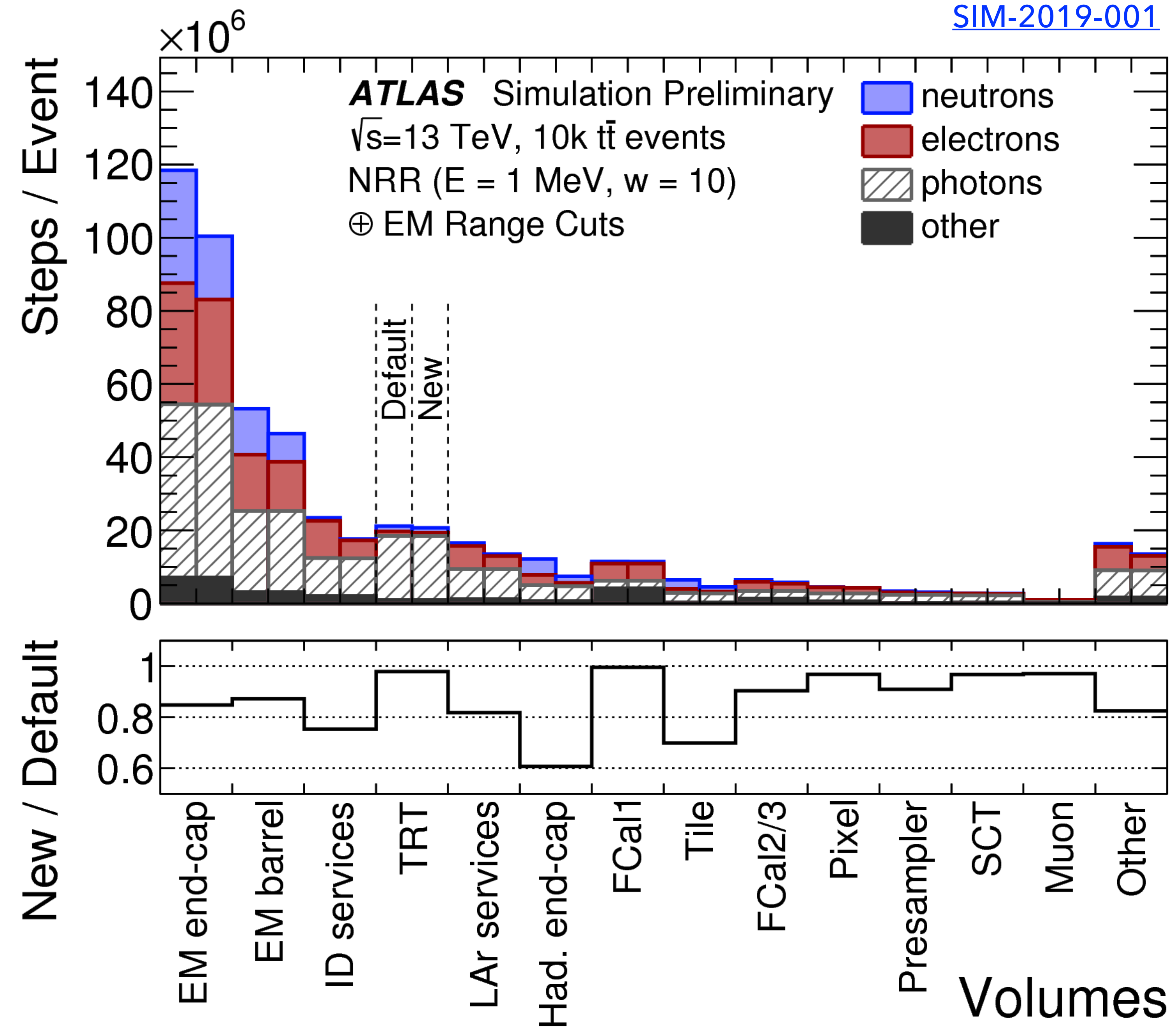
Facts

1. Full detector simulations (FullSim - full Geant4 tracking) are **accurate but the largest CPU consumer**
2. FullSim usage is **unavoidable** (CP calibrations, FastSim training, etc.)
3. **EM calorimeters** dominate the simulation load:
    a. **low-energy photons** from electron scattering
    b. highly-segmented geometry
4. **~90% of photon simulation steps are transportation processes**
   i.e. moving through detector geometry without interaction

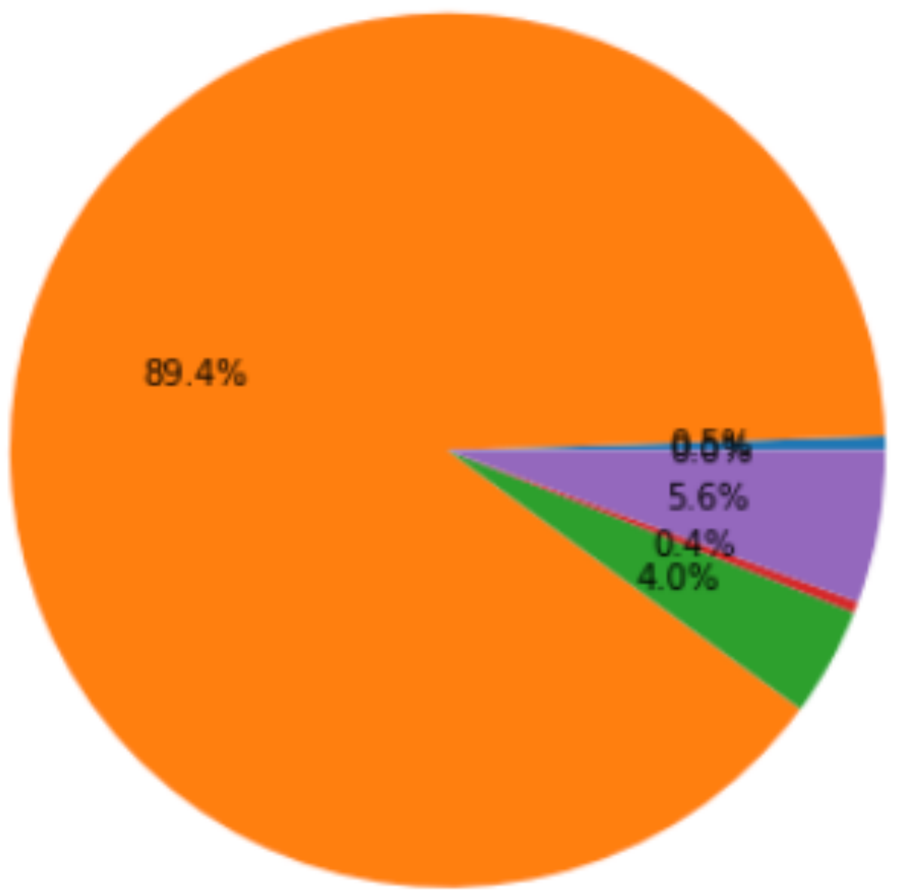


CERN-LHCC-2020-015

SIM-2019-001



Transportation
Photoelectric
Compton

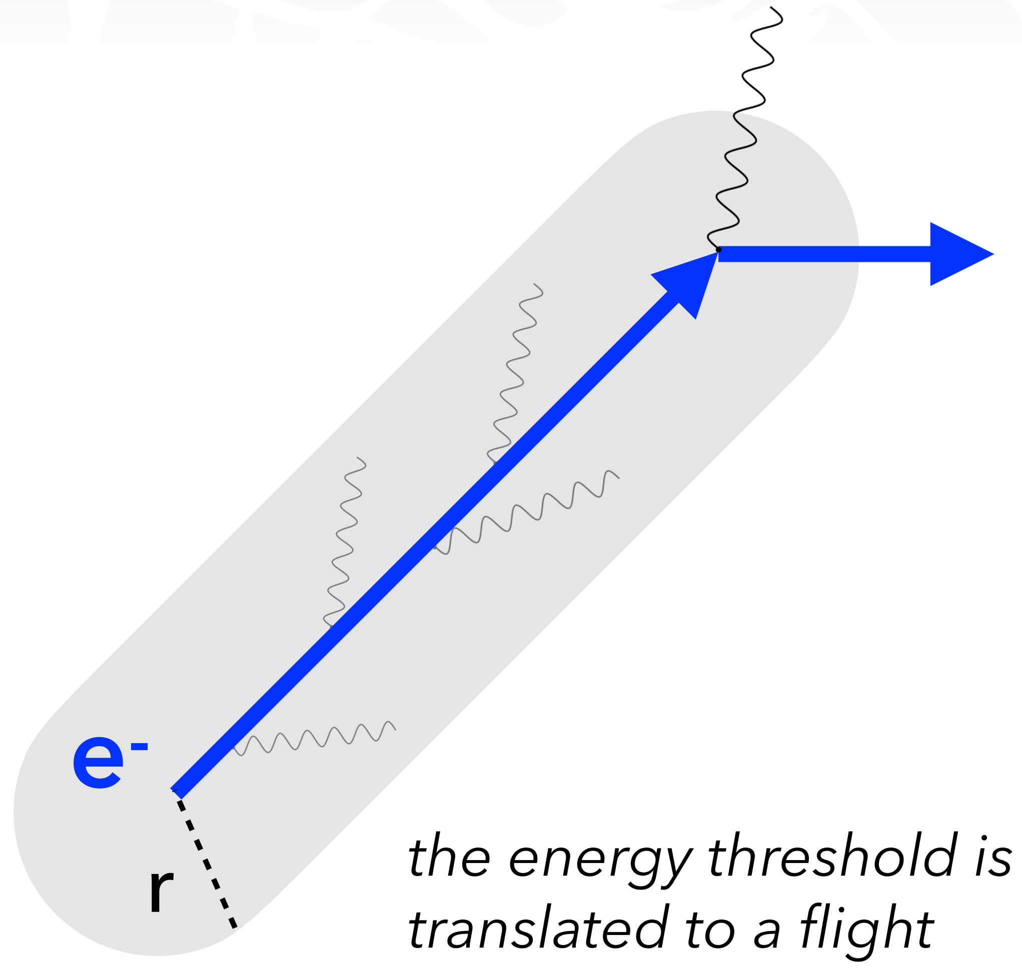


Methodology

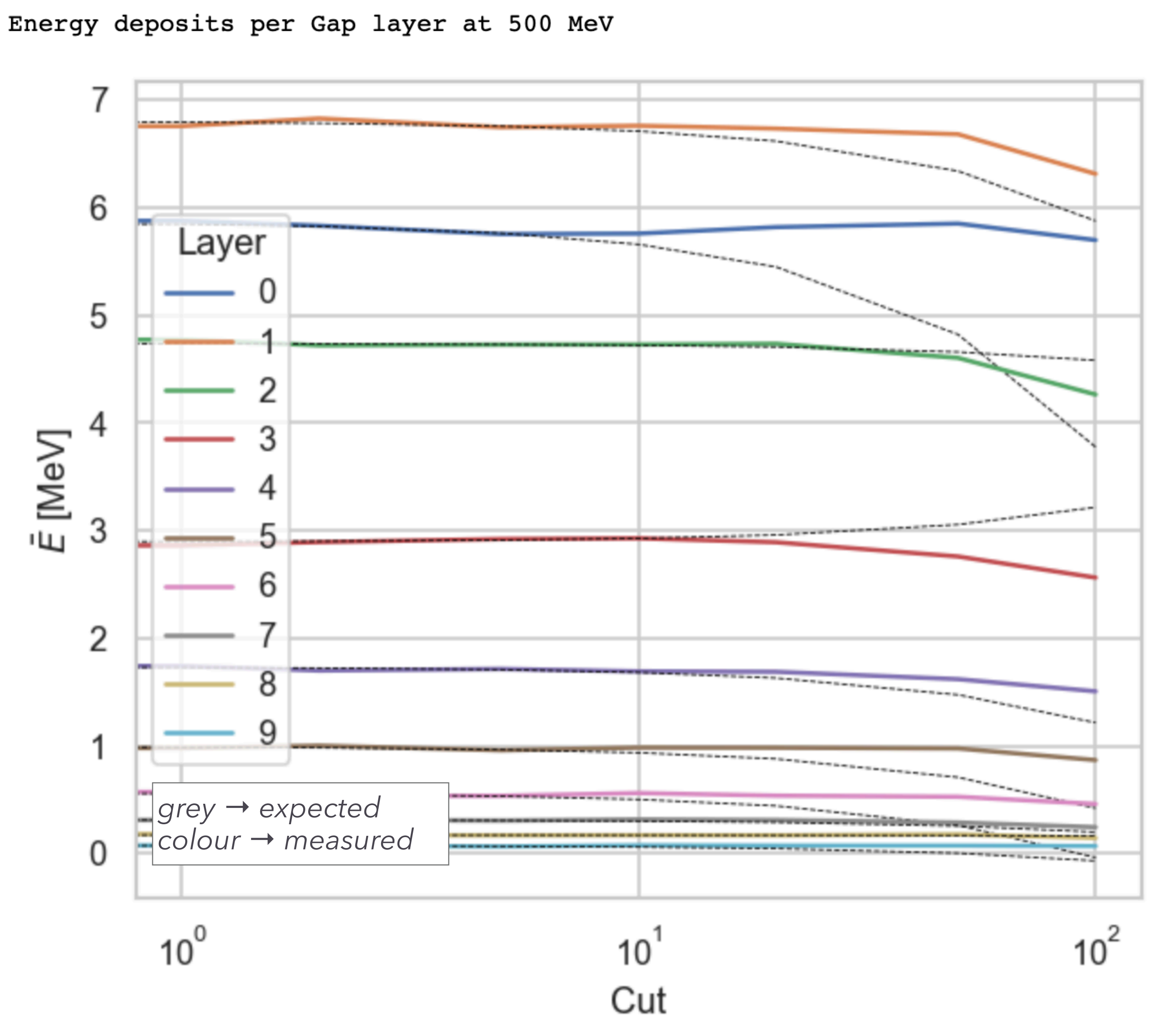
Photon Reduction

Range Cuts in Geant4

- Particle production energy threshold:
  If a secondary particle is going to have energy below the threshold, the particle is not generated and the energy is deposited along the path of the primary.
- **Increased range cuts** can **reduce the number of photons,** thus **reduce the transportation steps** and **increase computational performance.**
- Range cuts can be applied globally or to specific material
- **Side-effect:** "High" range cuts can **degrade the accuracy** of the simulation.



*the energy threshold is translated to a flight radius **r** for each material*



multi-layer calorimeter, absorber (effected volume) size 10mm

Post Hoc Correction

ML-based correction

to correct range cut'ed full simulation

Classification NN to learn multi-dimensional correction weights

by considering all cell energy deposits

Benefit: Heterogeneous computing exploitation

"Heterogeneous accelerated systems dominate high-performance computing today"

Geant4 simulations produced using CPU resources

ML-corrections is applied using GPU resources in a high-parallel fashion

Re-Weighting With Machine Learning

Re-weight the **alternative** simulation to the **nominal** one

learn multi-dimensional weights by considering all cell energy deposits

Map between two models (*pdfs*) with density ratio:

$$r(\vec{x}) = \frac{p(\vec{x}\,|\,\theta_p)}{q(\vec{x}\,|\,\theta_q)}$$

θ be the range cut, $\boldsymbol{x}$ the energy deposits

Considering

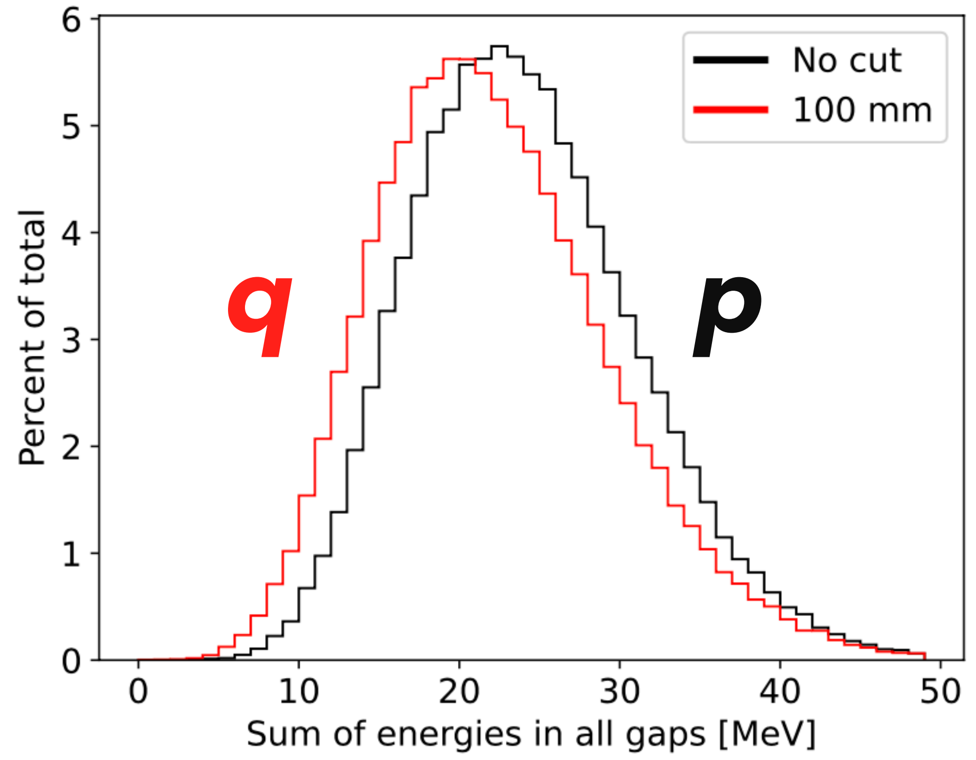
$$p(\vec{x}|\vec{\theta}_p) = \mathcal{P}(\vec{x}|\vec{\theta}_p, y=0)$$
$$q(\vec{x}|\vec{\theta}_q) = \mathcal{P}(\vec{x}|\vec{\theta}_q, y=1)$$

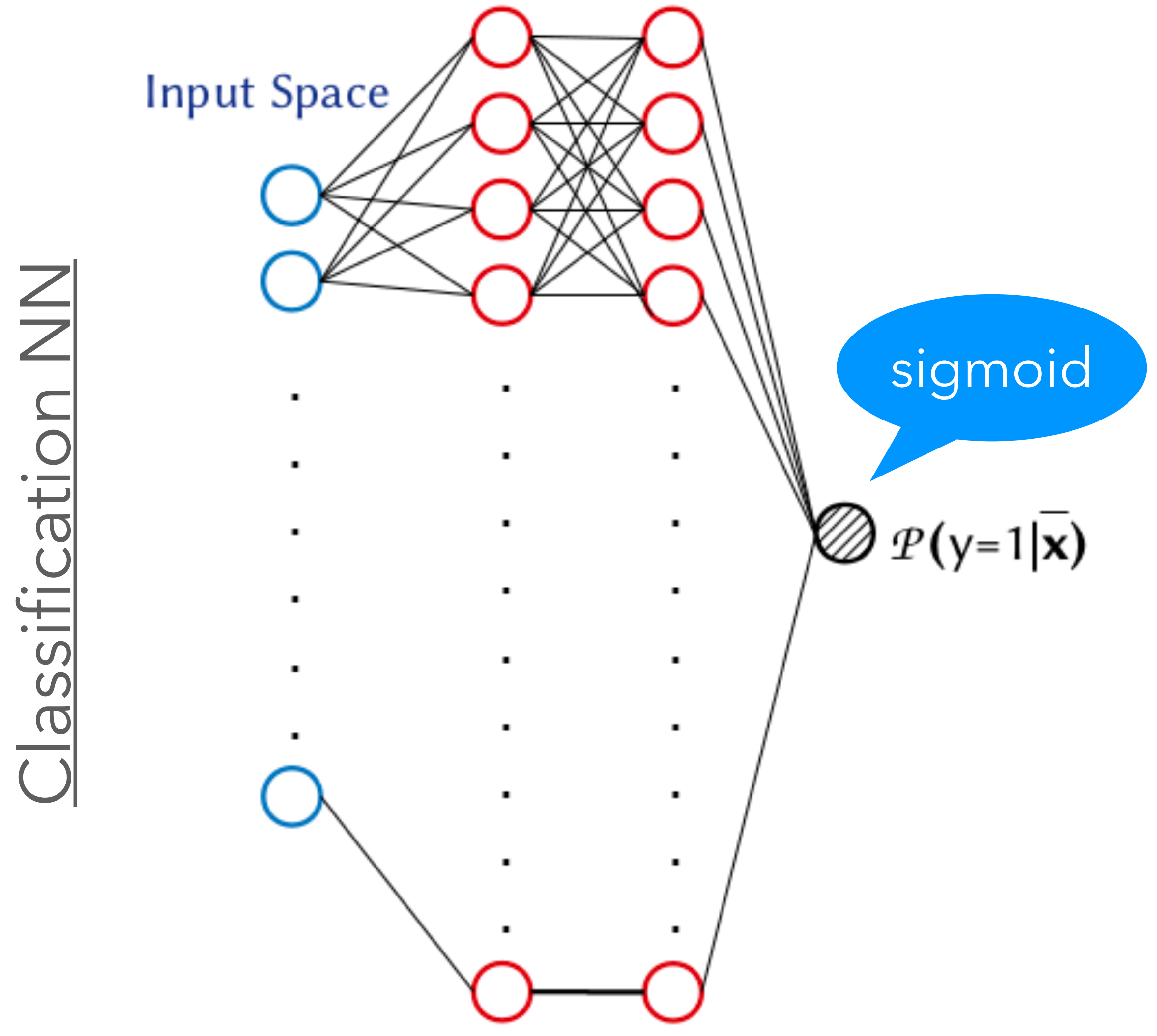
and

Bayes' theorem

$$\mathcal{P}(\vec{x}|y) = \frac{\mathcal{P}(y|\vec{x})\mathcal{P}(\vec{x})}{\mathcal{P}(y)}$$

$$\rightarrow \quad r(\vec{x}) = \frac{\mathscr{P}(y=1\,|\,\vec{x})}{1-\mathscr{P}(y=1\,|\,\vec{x})}$$





P is the probability of a point x̄ belonging to the class 0 (e.g. nominal sim) or 1 (e.g. range-cut'ed sim)

Reference: [Approximating Likelihood Ratios with Calibrated Discriminative Classifiers]()

Experiments

International Large Detector

Case study: Demonstrate the method in a realistic HEP calorimeter

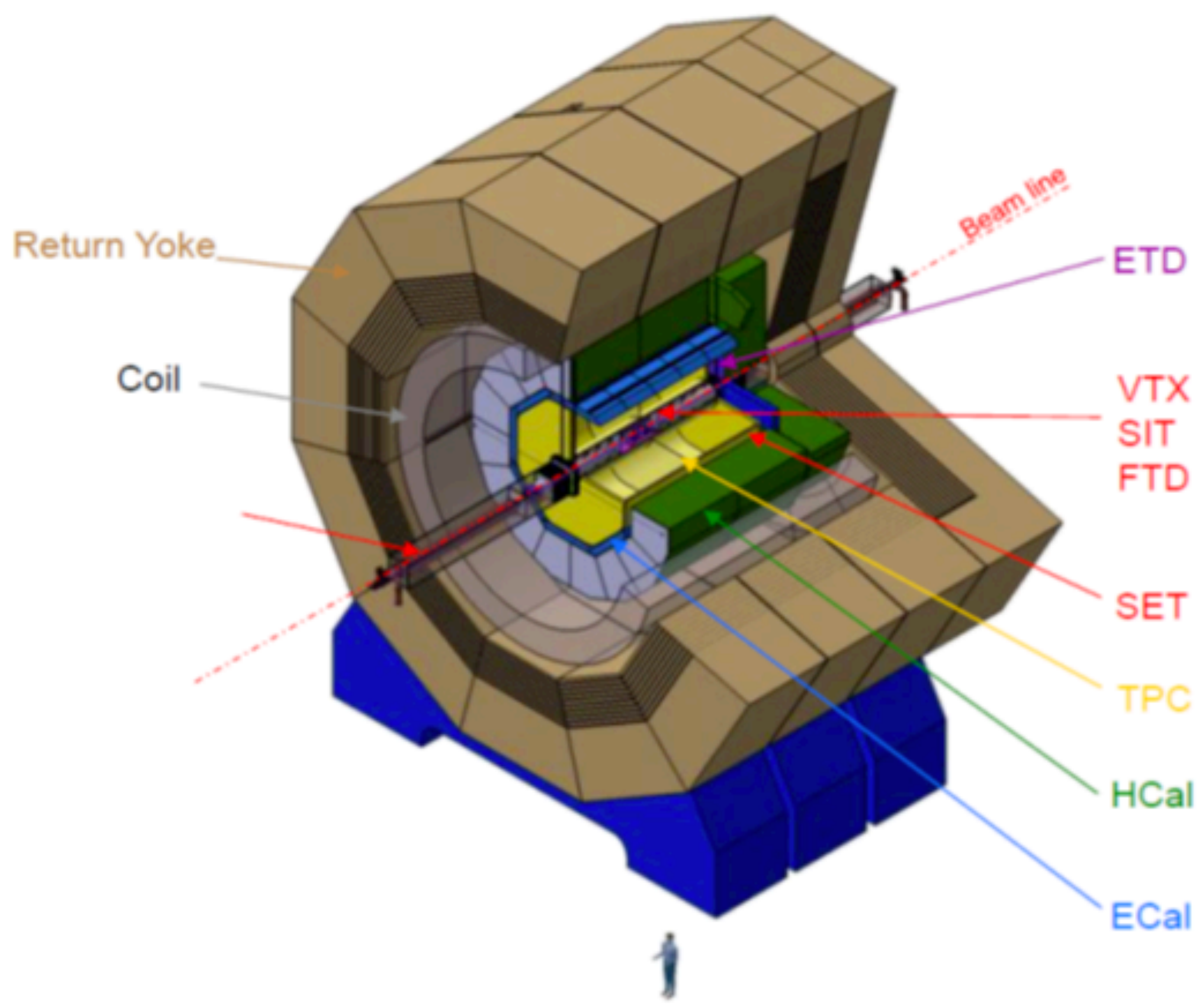
Detector proposal for the International Linear Collider

Calorimeter Material:

- **Absorber**: Tungsten
- **Sensitive**: Silicon

Calorimeter Structure:

- **Layers:** 30 (30 x 30 modules/layer)
- **Cell dimensions**: 5x5mm^2
- **Thickness:** 0.3mm - 0.6mm



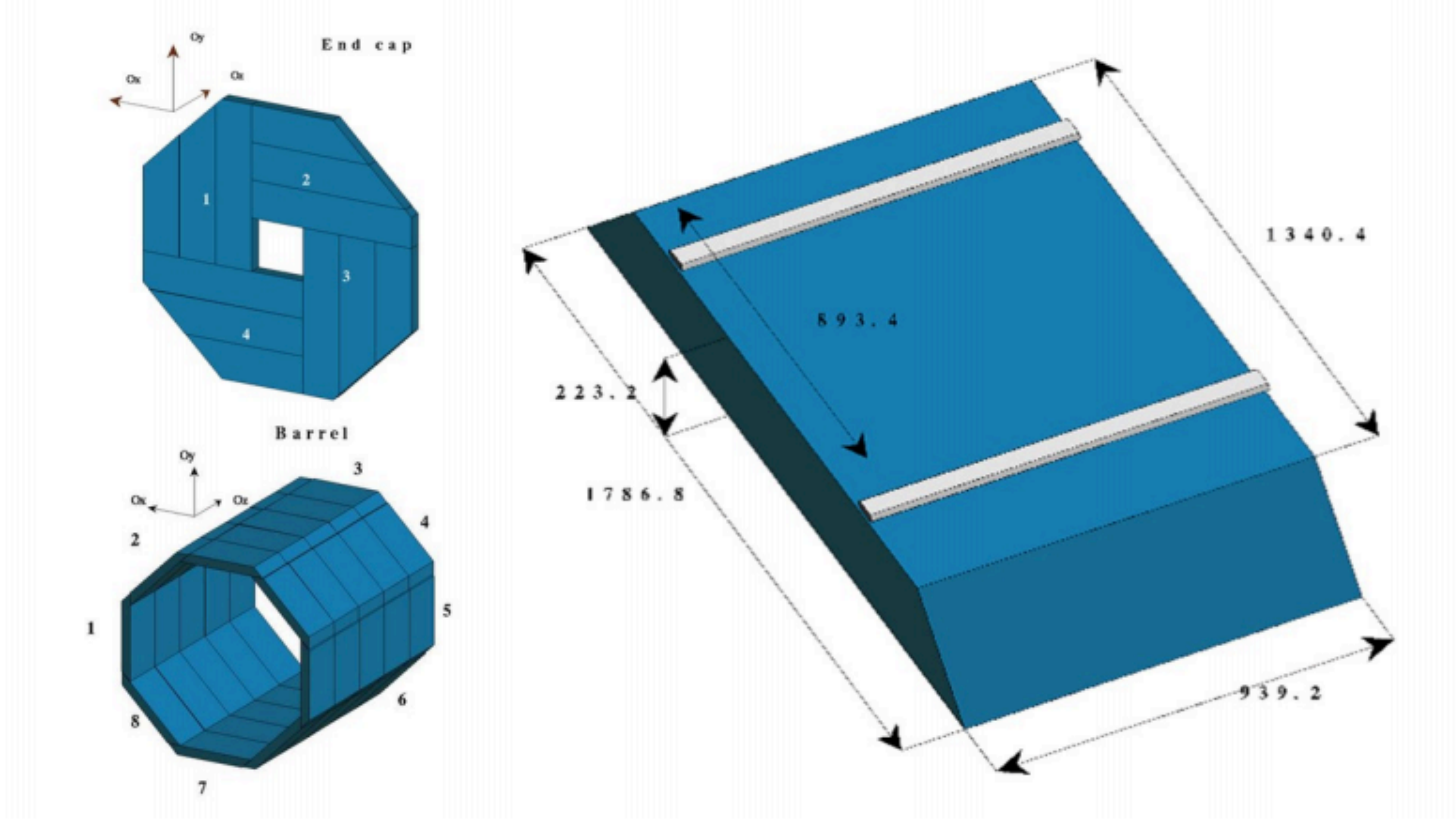
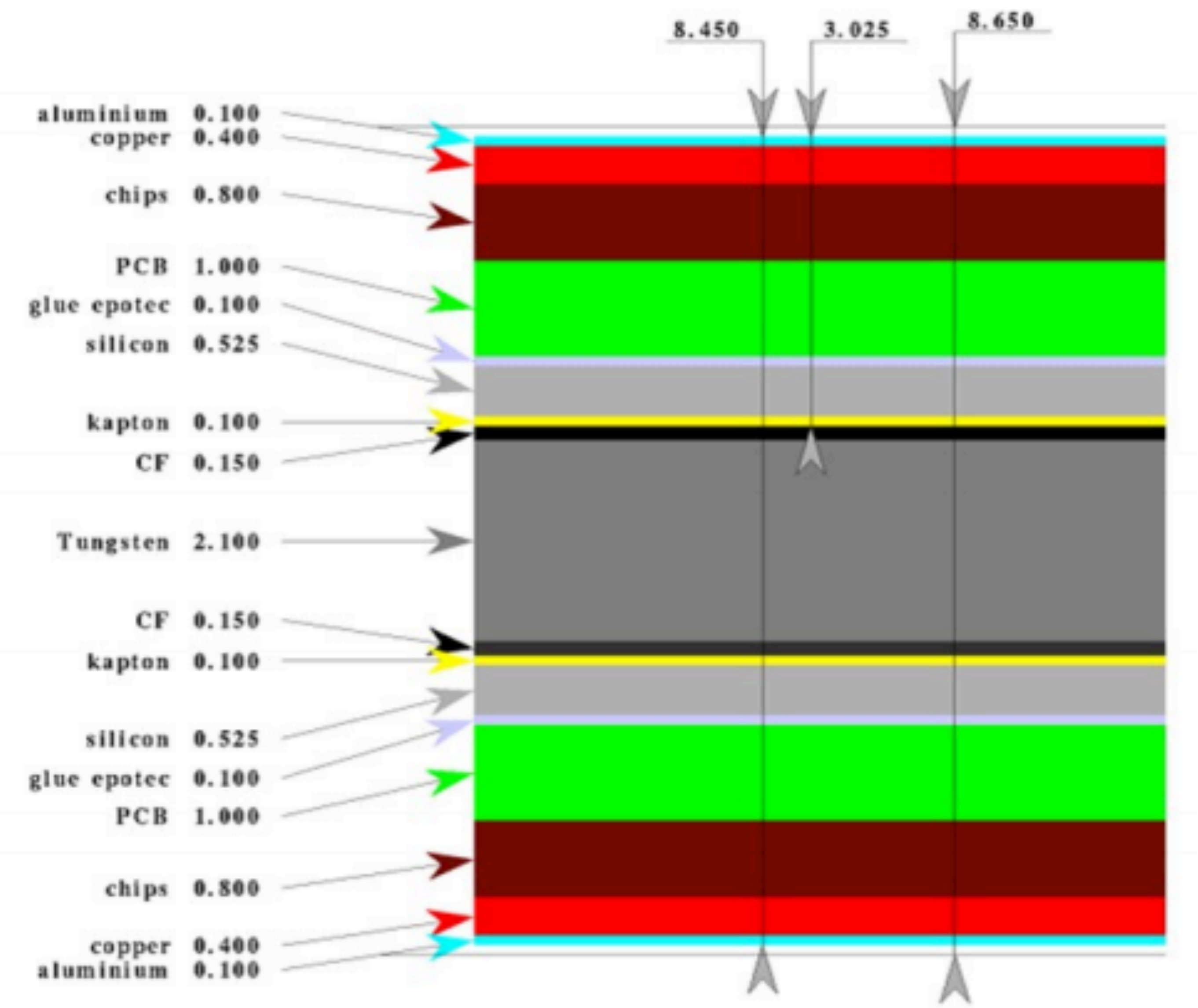


Figure 5.8. Mechanical structure of the electromagnetic calorimeter: left: end cap (top) and barrel (bottom); right: individual barrel module.



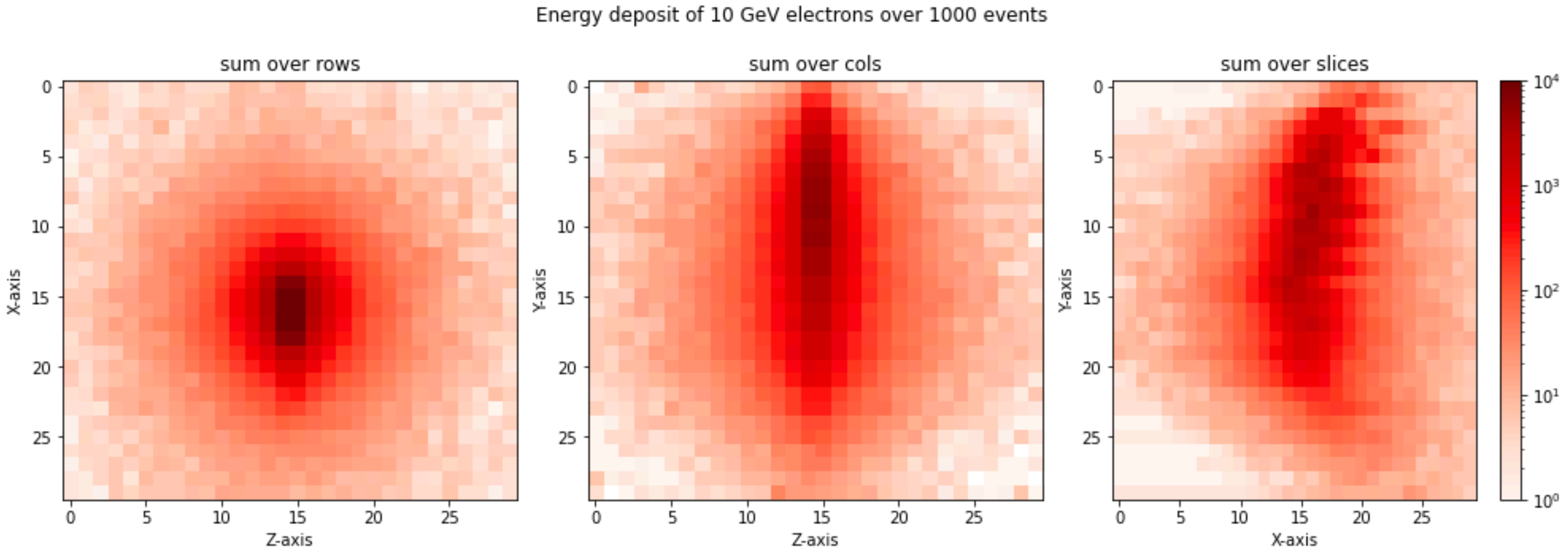
Reference: [International Large Detector: Interim Design Report]

Electron Showers

Datasets generation

- **Particle**: Electrons (beam)
- **Energy**: 10 GeV
- **Direction**: perpendicular incident angle to ECal barrel (x=0, y=1, z=0)
- **Position**: at the start of ECal (x=0, y=1805, z=0)
- **Global *range cut***:
  - Nominal: 0.1mm
  - Alternative: 10 mm

Calorimeter cells are projected to a 30 x 30 x30 cube

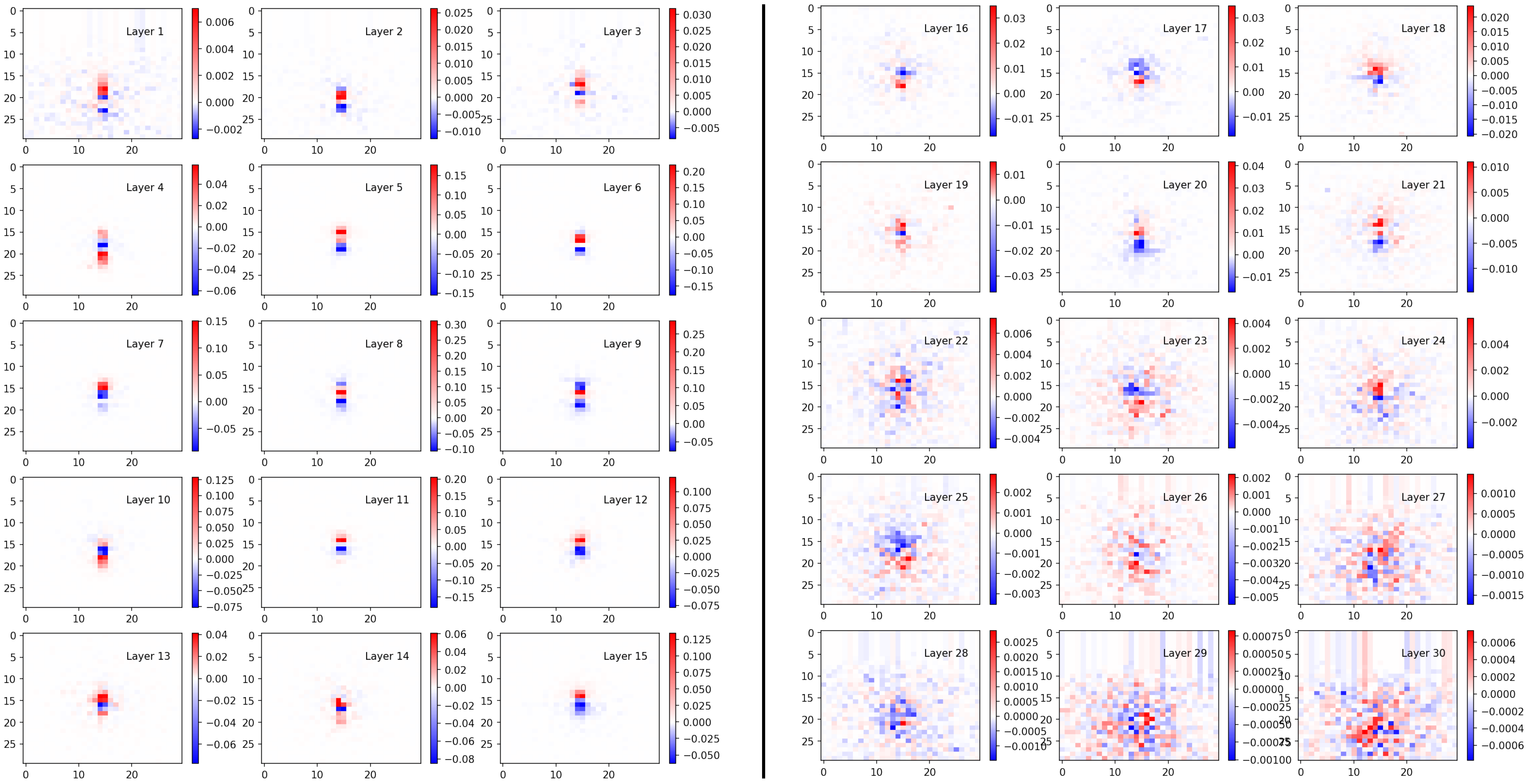


Geant4 simulation setup: [Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed]

Cell-Level Observables

Subtle calorimeter image differences the ML should use to discriminate

(nominal - alternative)



Global Observables

Event energy deposit

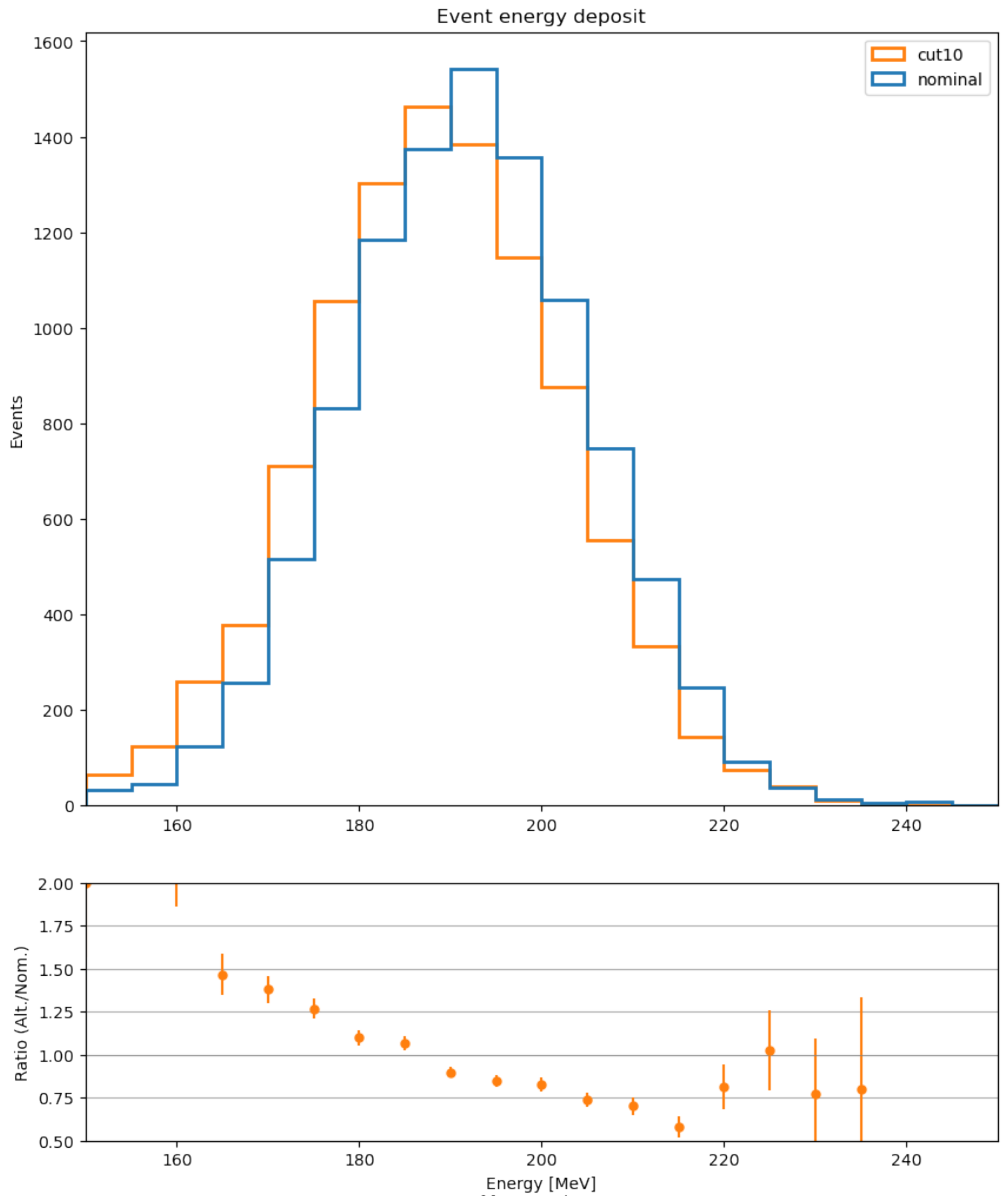
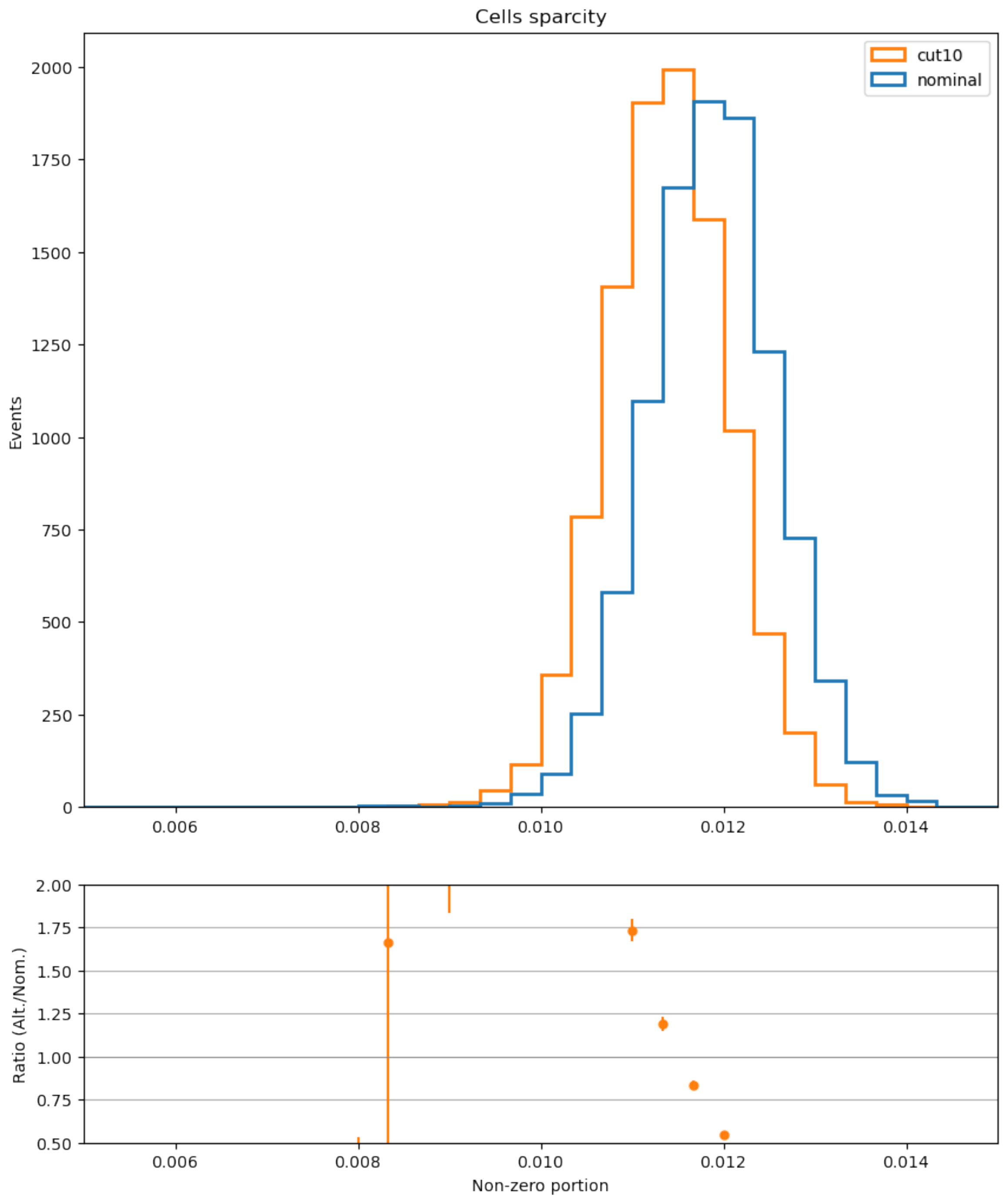
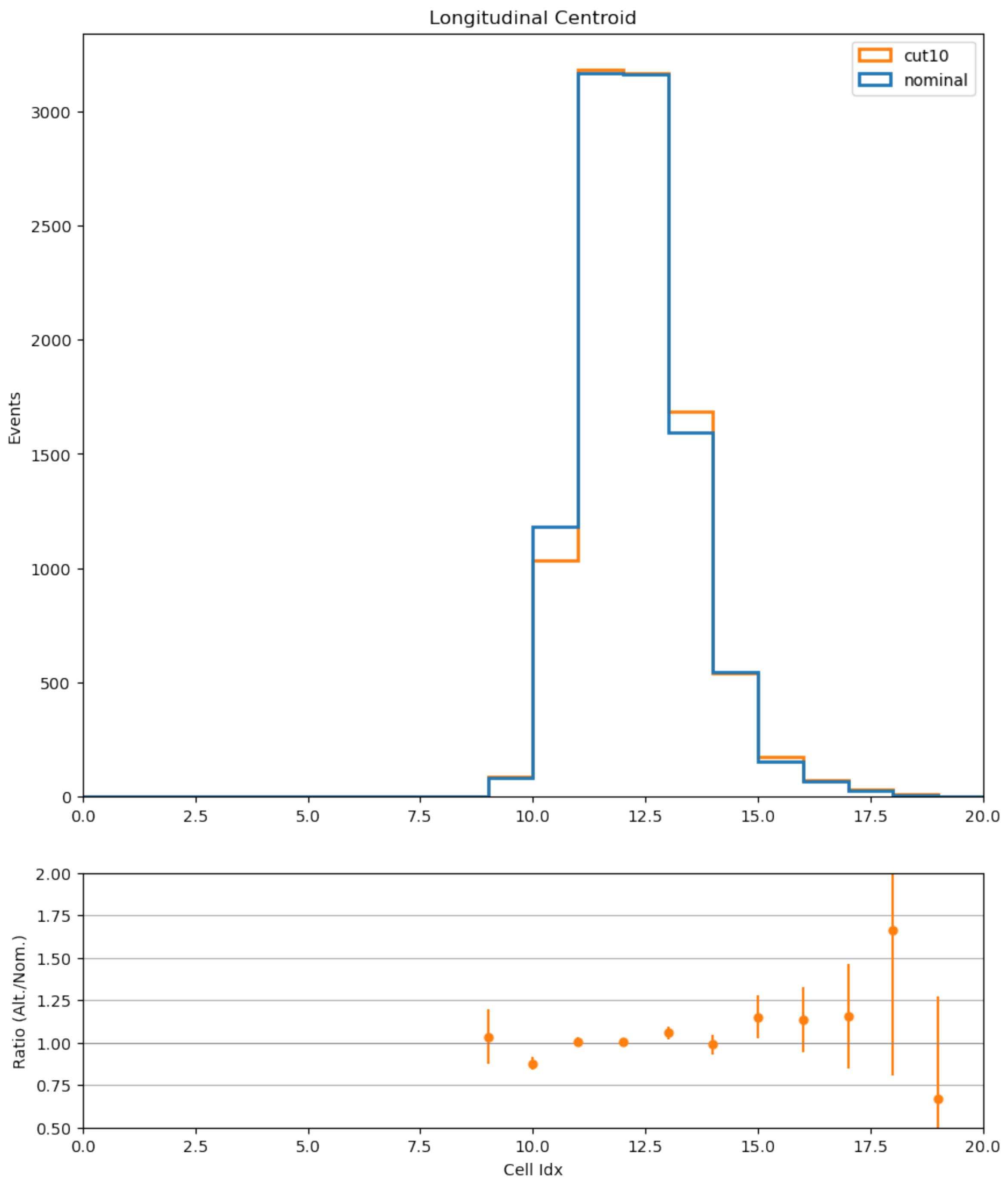


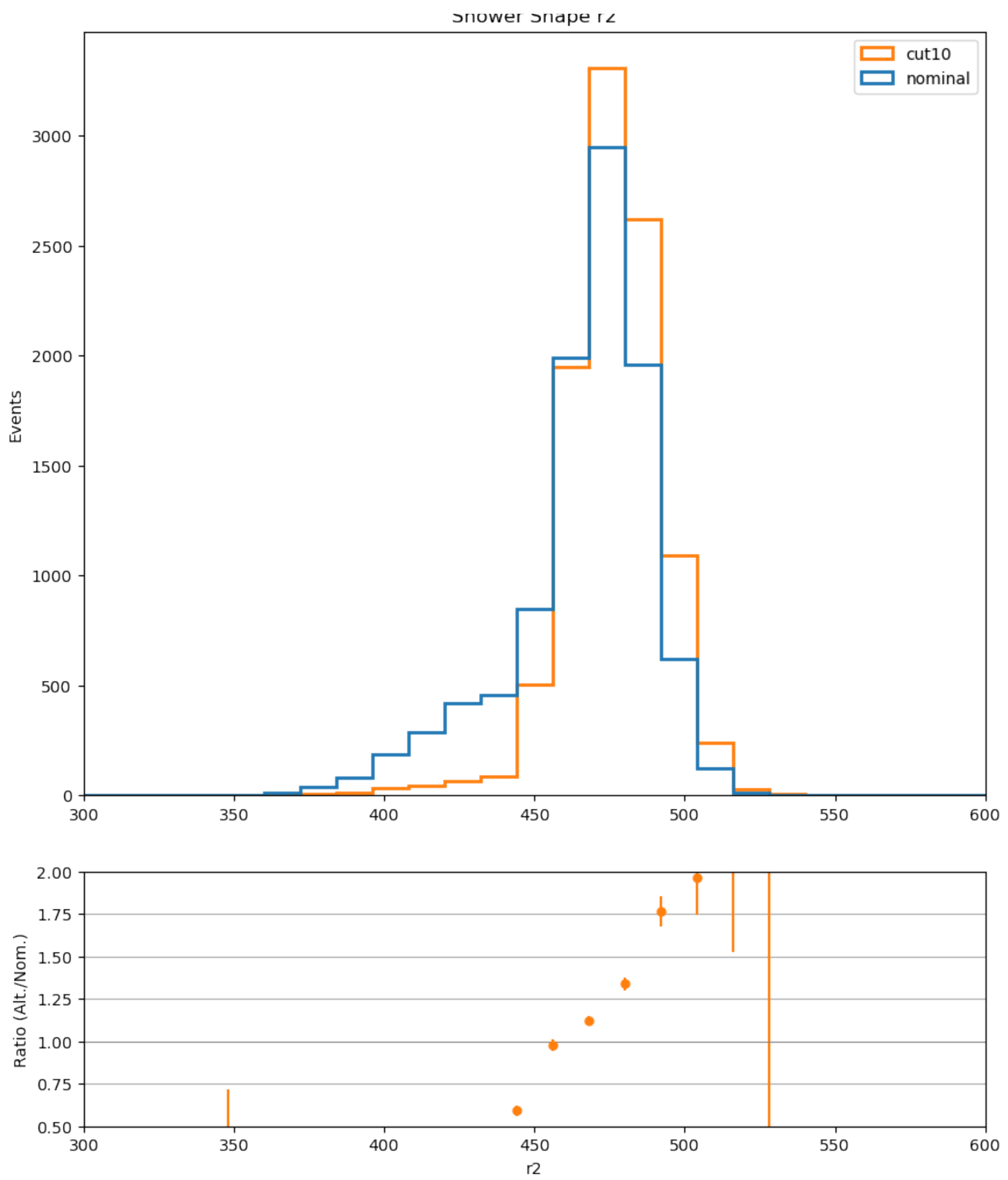
Image sparsity



Energy-weighted depth



Energy-weighted transverse spread



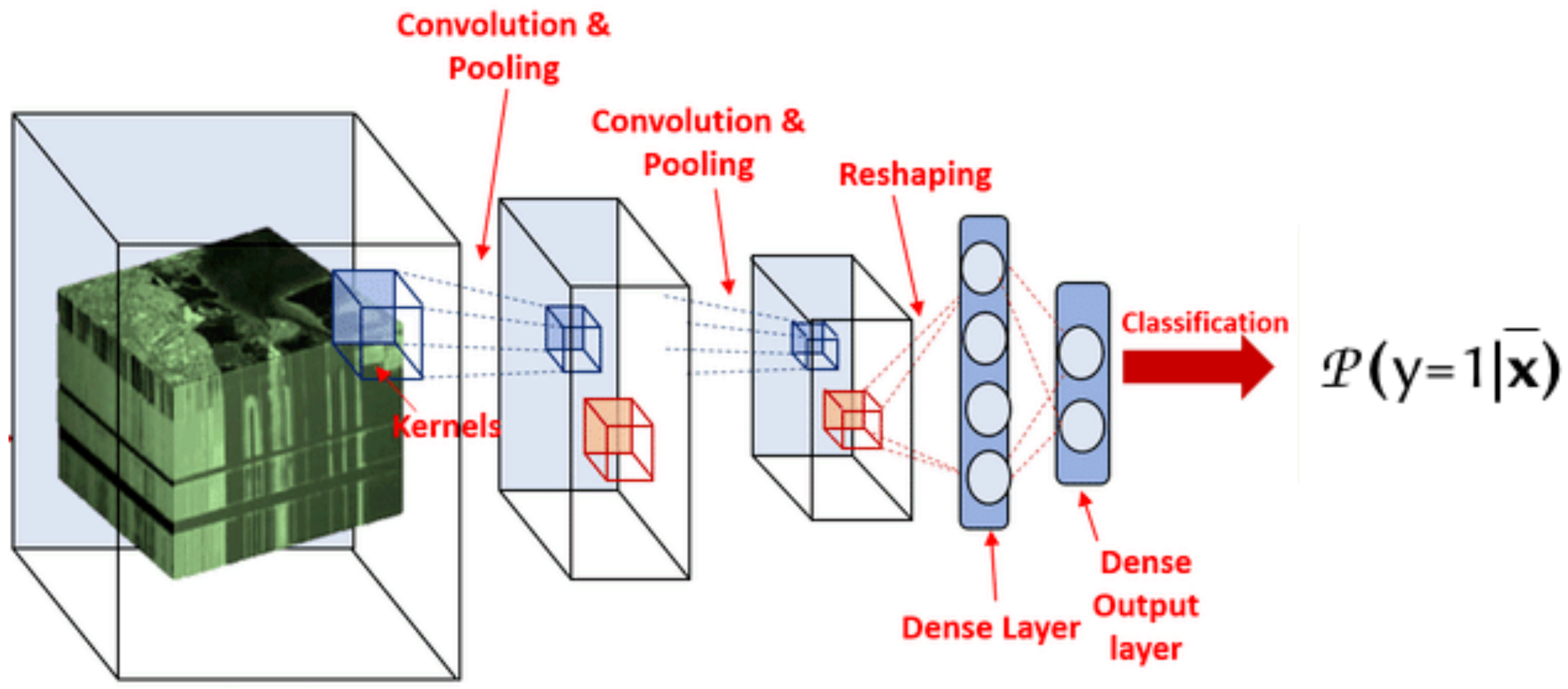
Classification Neural Network

3D Convolutional Neural Network

- The calorimeter cell energy deposits are projected into a 30x30x30 *image*
- Employ computer vision approach to discriminate *nominal* from *alternative* images
- Different normalizations are tried:
  maximum per image, global maximum, log-scale
- **Structure:**
  1. 1x Convolution block: Conv3d (kernel=3x3x3) + MaxPool3d
     - *Channels: 1 → 6*
  2. Flattening Layer
  3. 4x Dense Layers
     - *Features: conv_out → 512 → 512 → 1*
- **Activations:** LeakyReLU + Sigmoid (output)
- **Dropout:** after Conv block and each Dense
- **Network configuration only minimally optimized**

Developed in PyTorch

Code repo: torch-reweighter



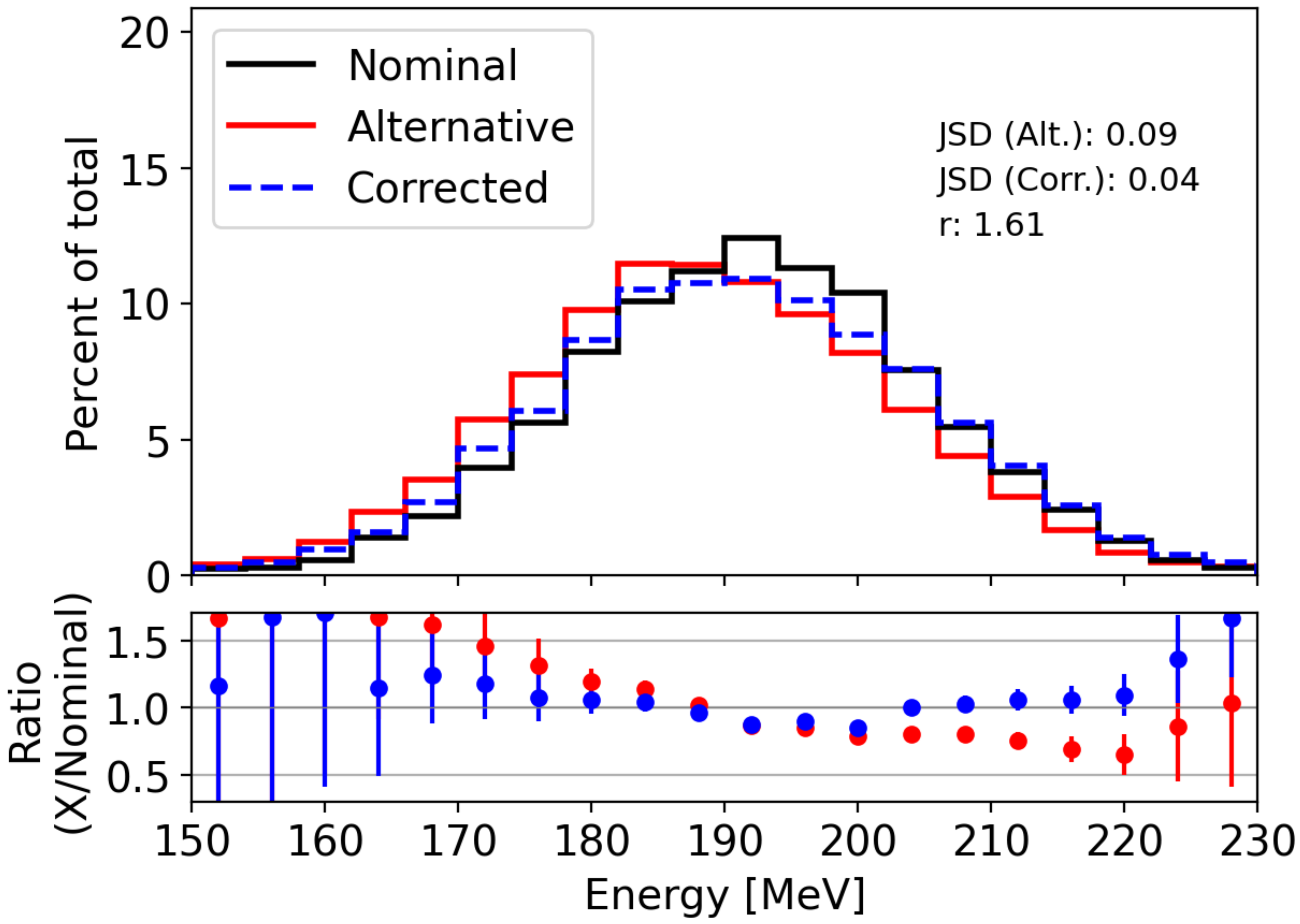
Input: Calo image (30x30x30)

Related work: DCTRGAN: Improving the Precision of Generative Models with Reweighting

arXiv:1804.08831

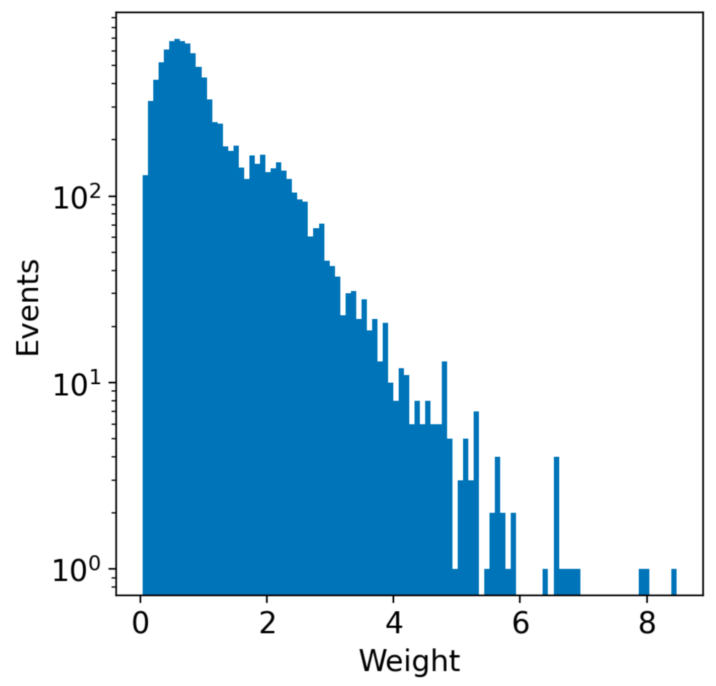
Evaluation / Weights Prediction

Evaluate the trained discriminator NN → extract weights from classification score



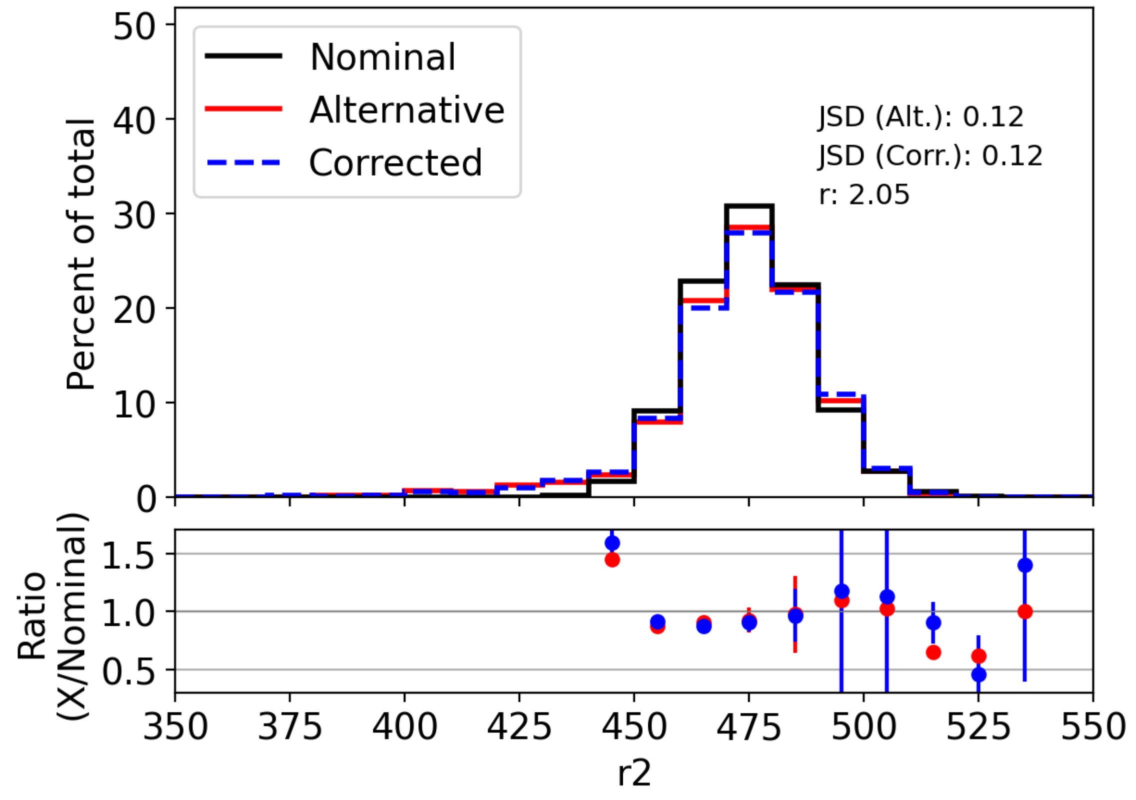
weights calculation

```
logits = self(x)
probs = torch.sigmoid(logits)
weights = probs / (1 - probs)
```



- able to correct global feature:
  event energy deposit
  while **training only voxel-level features**:
  cell energy deposits

- still not possible to successfully correct all global features shown



Simulation & Inference Timing

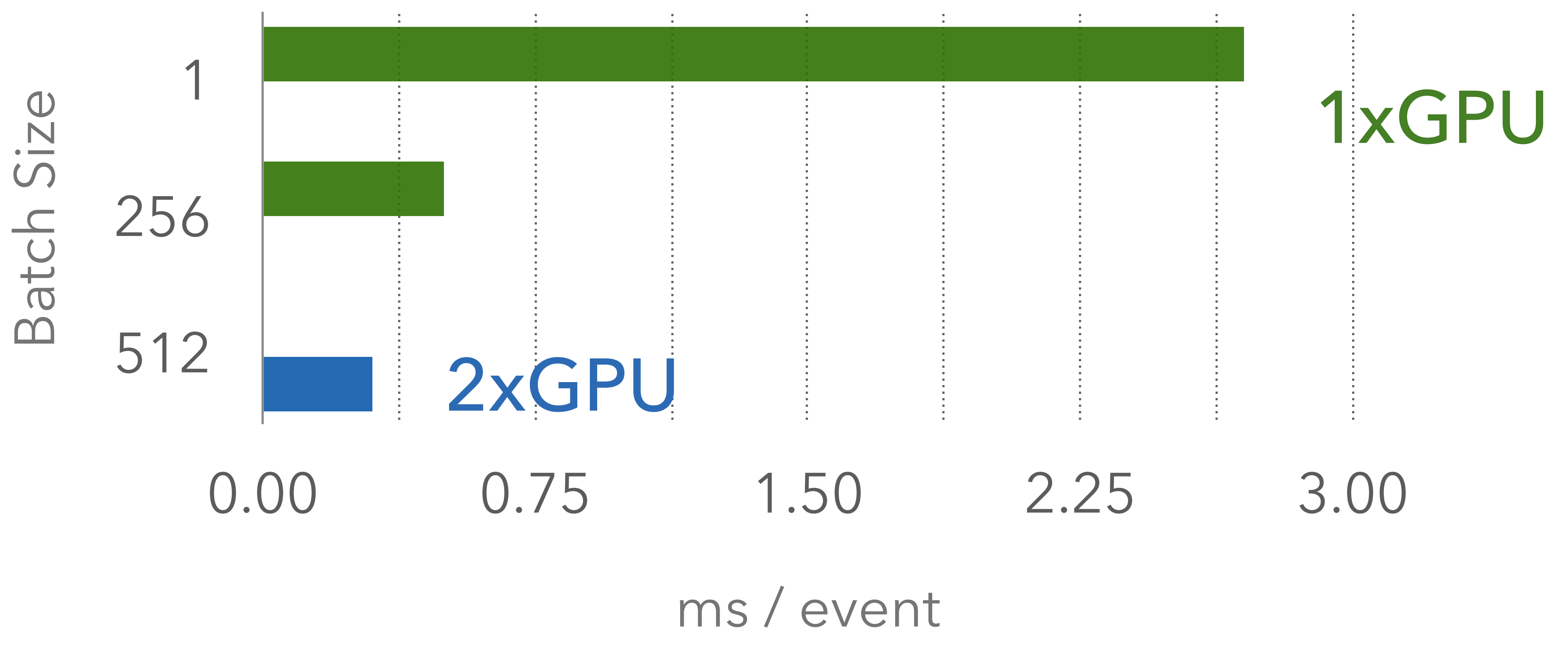
Timing measurements

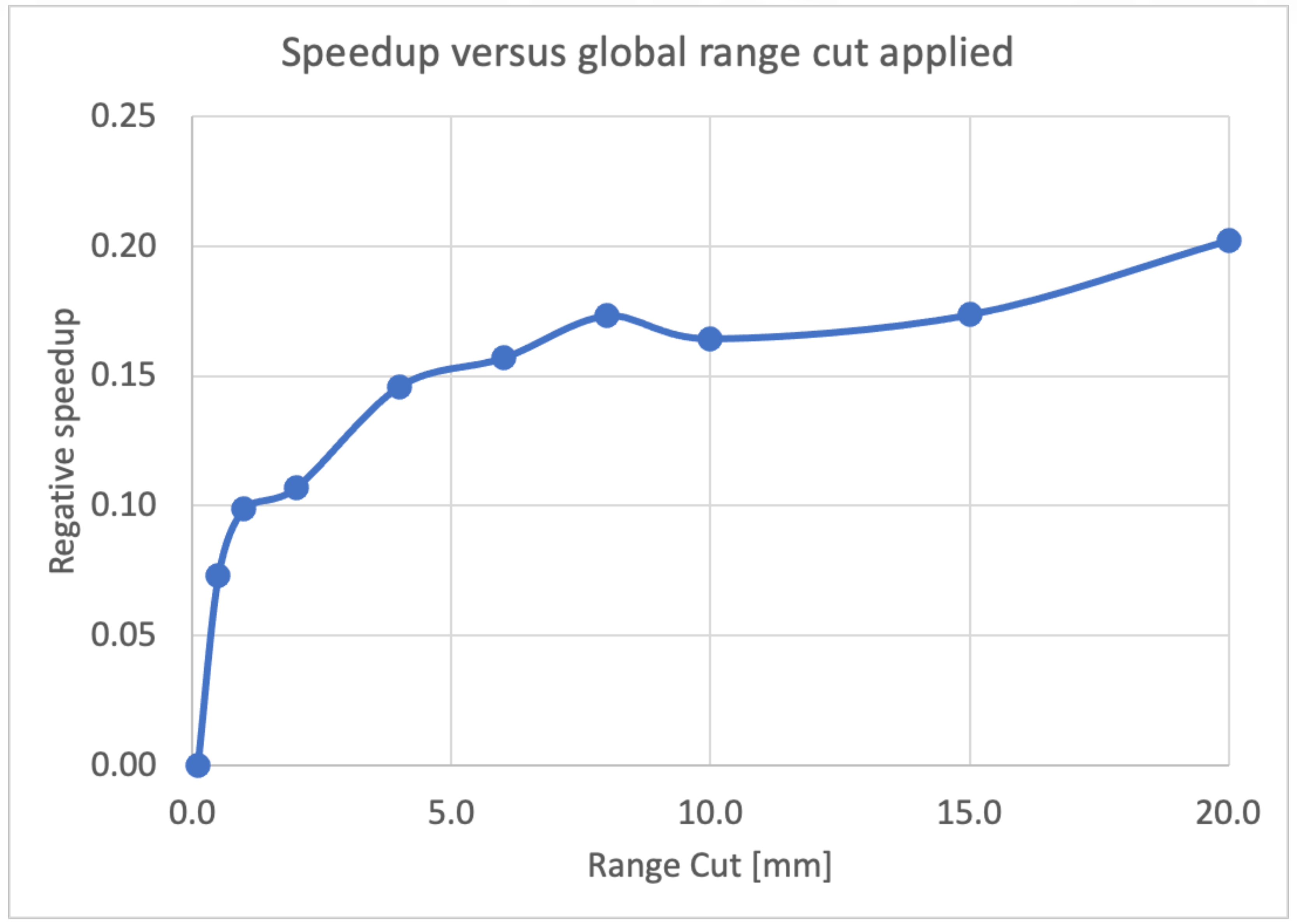
How does the ILD Geant4 simulation time changes as function of (global) range cut applied?

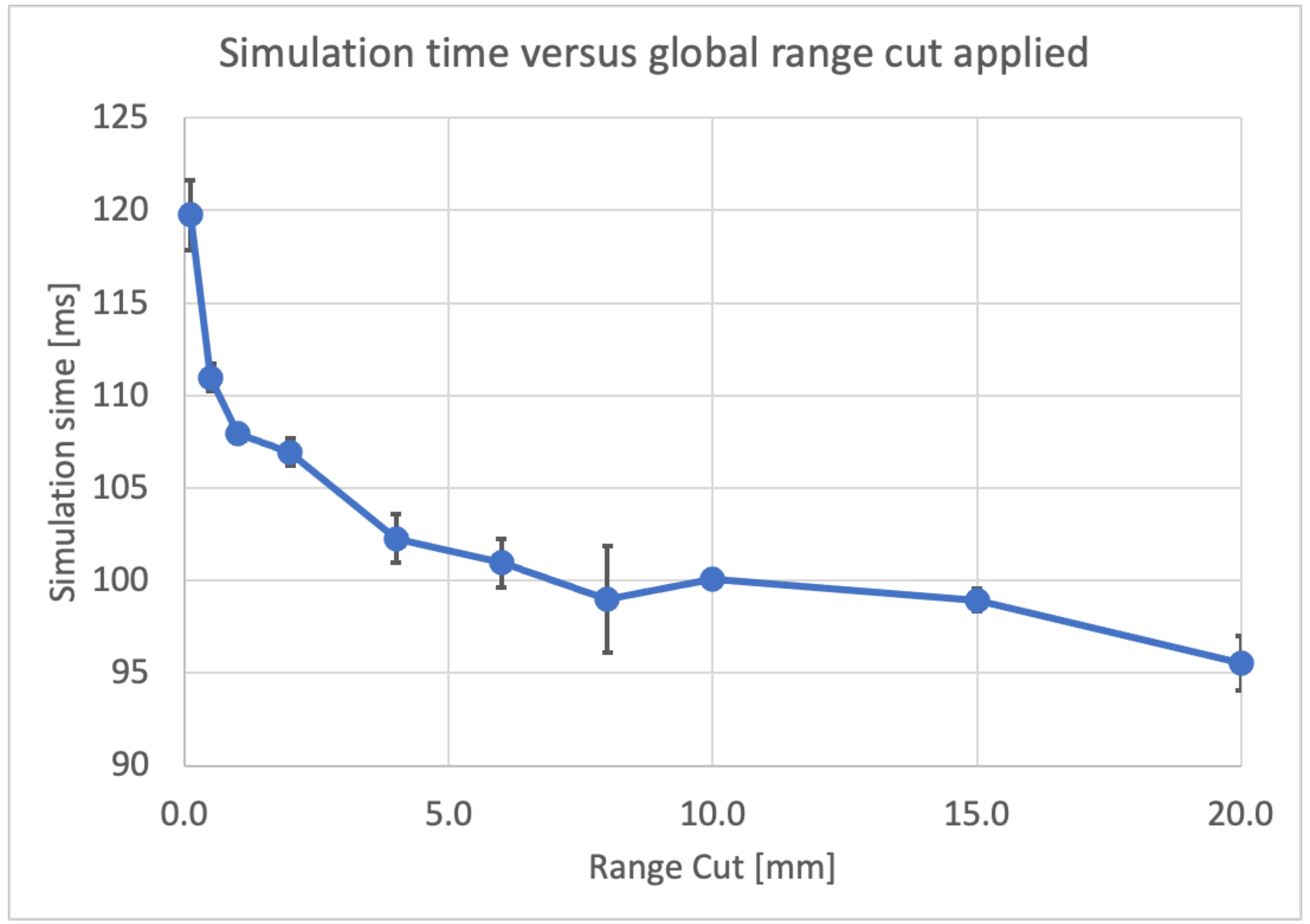
- Can achieve about 10% (17%) speedup with 1mm (10mm) *range cut*
- Nominal simulation time / event: ~120 ms
- Saving about 12(20) ms per event

What is the ML algorithm inference / correction time?

GPU: RTX 2028 Super - 8Gb





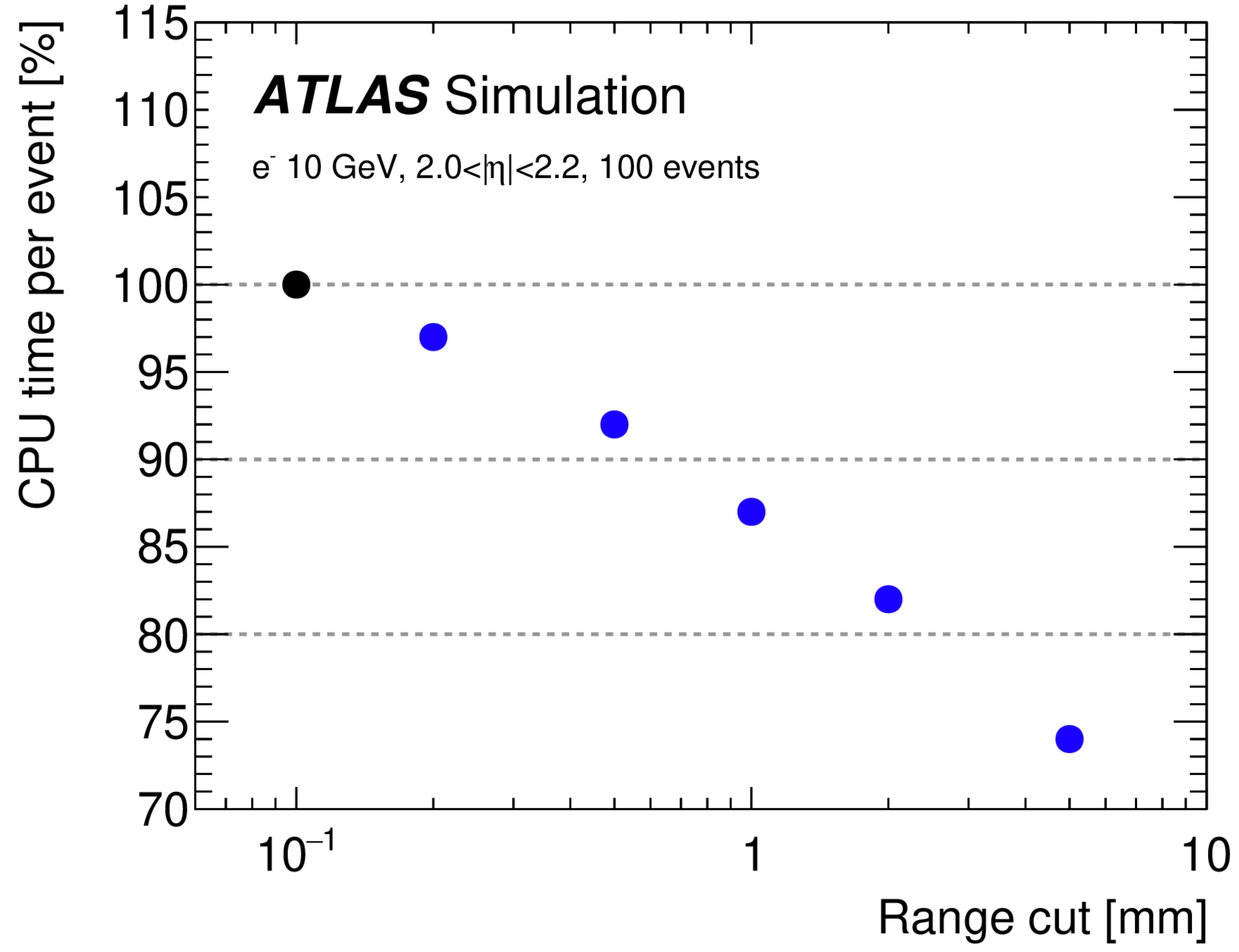


ATLAS ElectroMagnetic End-Cap Calorimeter

Range Cuts Speedup

ATLAS simulation time **speedup by ~15%** when increasing *range cut* by factor of 10 (0.1 → 1.0 mm)

SIM-2021-009



How to apply ML correction to ATLAS?

Detector considerations

1. Irregular geometry
2. Sparcity

Alternative data representations

1. Graphs
2. Point-clouds

The method shown is transparent to the ML algorithm architecture

Conclusions

solution to accelerate Geant4 simulation by applying aggressive range cuts and a post-hoc ML-based correction showcased promising result correcting event energy deposit

Benefit: Heterogeneous computing utilisation

Targeting Geant4 simulation speedup ~15%

Backup

Re-Weighting With Machine Learning

Re-weight the **alternative** simulation to the **nominal** one

learn multi-dimensional weights by considering all cell energy deposits

Map r to NN binary cross-entropy loss (L)

$$r(\vec{x}) = \frac{\mathscr{P}(y=1\,|\,\vec{x})}{1-\mathscr{P}(y=1\,|\,\vec{x})}$$

$$\mathscr{L}(\phi) = -\mathop{\mathbb{E}}_{p(x)}[\log D_\phi(\vec{x})] - \mathop{\mathbb{E}}_{q(x)}[\log(1-D_\phi(\vec{x}))]$$

$$D_\phi(\vec{x}) = \mathcal{P}(y=1|\vec{x}) = \sigma(\log r(\vec{x}))$$

where $\sigma^{-1}(\rho) = \frac{\rho}{1-\rho}$

$$\downarrow$$

$$-\mathop{\mathbb{E}}_{p(x)}[\log \sigma(\log \mathrm{r}_\phi(\vec{x}))] - \mathop{\mathbb{E}}_{q(x)}[\log(1-\sigma(\log \mathrm{r}_\phi(\vec{x})))]$$

The classifier is minimising the error on the r_φ – which is an estimate of the r

Reference: [Approximating Likelihood Ratios with Calibrated Discriminative Classifiers]

Detailed proof: [CARL - Theory]