Energy Reconstruction for the TileCal detector at the HL-LHC

XVII CPAN Open Days
COMCHA Session

Sonakshi Ahuja[1], Fernando Carrió Argos[1], Francesco Curcio[1], Luca Fiorini[1], Hector Gutierrez[1], Francisco Hervás Álvarez[1], Alberto Valero[1], **Wang Rui**[1]

[1]IFIC, CSIC-UV

Nov 20th, 2025

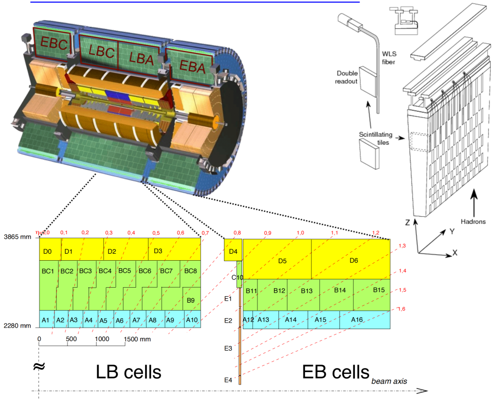
HORIZON-MSCA-COFUND-2022-1 GA 101126626

Introduction

Tile Calorimeter at the ATLAS experiment

- Central hadronic calorimeter of the ATLAS experiment
  - Central Long Barrel (LBA, LBC) and Extended Barrel (two readout partitions, EBA and EBC)
  - 64 modules per partition with up to 45 PMTs per module
- Sampling calorimeter composed of scintillators (active) and steel (absorber)
  - Charged particles produce light in plastic scintillators
  - The light is delivered to PMTs through WLS fibres
  - Reconstructs hadronic jets
  - Contributes to reconstruct the missing transverse energy
  - Input to trigger and muon identification
- Readout fibres groups into pseudo-projective cells
  - Each cell is read out by 2 PMTs
  - 5182 cells, 9852 PMTs

Eur. Phys. J. C 84 (2024) 1313

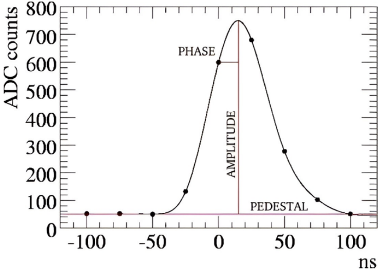


LBA/C Long Barrel A/C
EBA/C Extended Barrel A/C
PMT Photo Multiplier Tube
WLS Wave Length Shifting

TileCal signal reconstruction

- In the TileCal readout, signals are sampled every 25 ns
- Each pulse is characterised by three main parameters:
    - Pedestal: baseline ADC count value in absence of signal
    - Amplitude: proportional to the deposited energy
    - Phase: time shift relative to the nominal bunch crossing
- Amplitudes are measured in two gains, high gain (HG) and low gain (LG), with a factor of 40 difference
- In the legacy system, signals are processed online and offline using the Optimal Filtering algorithm (OF)
    - Simple and fast response filter
    - Linear combination of the samples that uses weights determined from the known pulse shape and noise correlation matrix
    - Does not perform well under severe signal pile-up conditions due to non-gaussian (asymmetric) components
- Need state-of-art reconstruction algorithms to reconstruct the amplitude and the phase precisely

ADC Analog-to-Digital Converter



Optimal Filtering

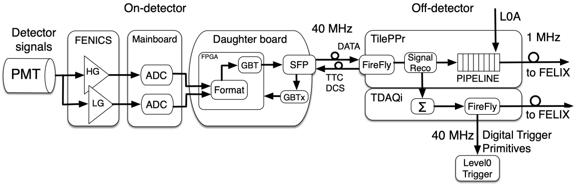
$$A = \sum_{i=1}^{n} a_i (S_i - p)$$

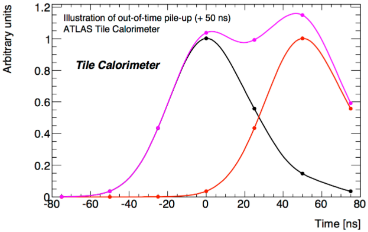
$$\tau = \frac{1}{A} \sum_{i=1}^{n} b_i (S_i - p)$$

high and low gain (HG, LG) Amplitudes are measured in 2 gains, with a factor of 40 difference, both with 12-bit ADCs to maintain 10 MeV precision at energies < 10 GeV while avoiding saturation at high energies.

Motivation and goals

- During Phase-II, signals will be processed per bunch crossing (BC) before passed to the first level of trigger
  - Reconstruction will be done by FPGAs
- Higher data rates and pile-up, together with the need for real-time reconstruction, call for much more precise algorithms



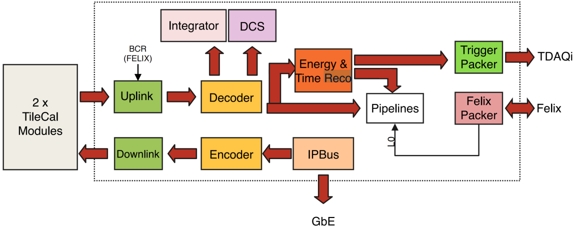


- The expected performances of different ML algorithms are checked for HL-LHC conditions using the `PyTorch` package
- Optimise the number of parameters in order to ensure the algorithms can run on FPGAs with the smallest possible latency

Hardware constraints

- Xilinx Kintex UltraScale KU115
- Maximum latency estimated for the real-time trigger path is 325 ns
- If two gains are used, we might need 154 NNs
  - MLP with size 7-7-7-1 and 9-9-1 would take up at least 32% and 50% LUTs correspondingly
- Many artechitectures failed timing even at 280MHz or take too long to compile
- Clock domain crossing and synchronisation might be an issue
- Hardware conditions impose constraints on the size of the NNs

Firmware block	Latency	
Uplink	50.0ns	2.0 BC
Data Decoder	12.5ns	0.5 BC
Energy Reconstruction + sample delay	225.0ns	9.0 BC
Trigger Packer	12.5ns	0.5 BC
Trigger interface	25.0ns	1.0 BC
Total	**325ns**	**13 BC**

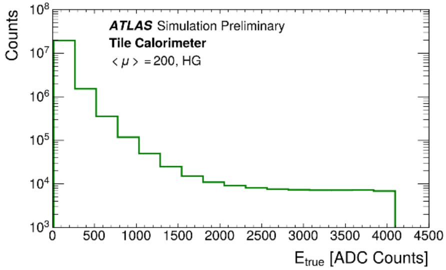


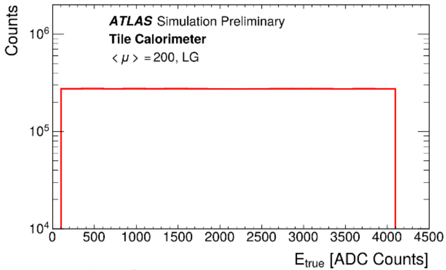
Block diagram of the Kintex Ultrascale firmware

BC Bunch Crossing
NN Neural Network
DCS Detector Control System
LUT Loop Up Table

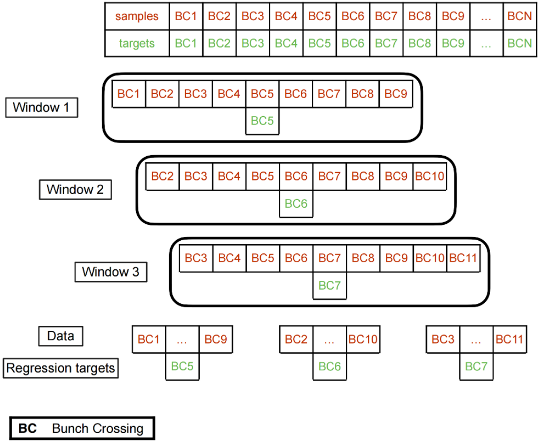
Data

- Simulated datasets are generated by the Pulse Simulator
- Samples are composed of $\sim$1M consecutive bunch crossings with minimum bias $<\mu>=200$, superimposed by a flat distribution with a probability of 5%
    - $\sim$1M $\times$ 64 modules $\times$ 4 channels $=$ $\sim$26.5M
- Only simulation of A1 cells of the Tile Calorimeter used
    - Different cells might require different models
- The energy in ADC count for every bunch crossing is read in both gains
    - Models are trained on a mix between HG and LG





Preprocessing



- ADC values range from 0 to 4095
- Sliding window with a size 9
- Samples, $\mathrm{BC}_i \to \mathrm{E}_{\mathrm{reco},i}$
  - Simulated readout energy from the electronics
  - Inputs to the models
- Target, $\mathrm{BC}_i \to \mathrm{E}_{\mathrm{true},i}$
  - Ture energy for the i-th BC
  - Regression targets
  - Central energy of the window

Preprocessing

samples	HG	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9
	LG	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9
	targets	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9

- If BC_i saturates HG, take LG$\times 40$, else HG

samples	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9
targets	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9

- If any of the BC in the window saturates LG ($\mathrm{E}^{\mathrm{LG}}_{\mathrm{reco}} = 4095$), drop the window
- If any of the BC in the window has $\mathrm{E}_{\mathrm{reco}} \leq 10$ ADC counts, drop the window
- If the central BC has $\mathrm{E}^{\mathrm{HG}}_{\mathrm{true}} \leq 10$ or $\mathrm{E}^{\mathrm{LG}}_{\mathrm{true}} = 4095$, the window is also dropped.

- For samples, HG, $\mathrm{BC}_i \rightarrow \mathrm{E}^{\mathrm{HG}}_{\mathrm{reco,i}}$
  - Simulated reconstructed energy in the high gain reading
- For samples, LG, $\mathrm{BC}_i \rightarrow \mathrm{E}^{\mathrm{LG}}_{\mathrm{reco,i}}$
  - Simulated reconstructed energy in the low gain reading
- For target, $\mathrm{BC}_i \rightarrow \mathrm{E}_{\mathrm{true,i}}$
  - True energy for the i-th bunch crossing
  - Values range from 0 to 40*4095

Models

Models

- Train ML algorithms to reconstruct energy from the readout energy from the electronics
- Different models are tested, including MLP, CNN, LSTM, etc.
- Some of the models are dropped due to
  - difficulty to train
  - complicated implementation on FPGAs
  - suboptimal results
  - number of parameters too high
  - high latency
- After some optimisation, MLPs and CNNs are selected
- Number of parameters at the level of 150

Model architecture and loss

MLP

```
Sequential(
  (0): Linear(in_features=9, out_features=9, bias=True)
  (1): PReLU(num_parameters=6)
  (2): Linear(in_features=9, out_features=4, bias=True)
  (3): PReLU(num_parameters=4)
  (4): Linear(in_features=4, out_features=1, bias=True)
)
```

Total parameters: 148

CNN

```
Sequential(
  (0): Conv1d(1, 6, kernel_size=(3,), stride=(1,), padding=(1,))
  (1): PReLU(num_parameters=6)
  (2): Conv1d(6, 4, kernel_size=(3,), stride=(1,), padding=(1,))
  (3): PReLU(num_parameters=4)
  (4): Flatten()
  (5): Linear(in_features=36, out_features=1, bias=True)
)
```

Total parameters: 147

- Hybrid loss $= \alpha \cdot \frac{1}{N}\sum_i \left|y_i - \hat{y}_i\right| + \beta \cdot \sqrt{\frac{1}{N}\sum_i \left(y_i - \hat{y}_i\right)^2}$
  - Root Mean Squared Error (RMSE) used to keep same units as Mean Absolute Error (MAE)
  - Using $\alpha = \beta = 0.5$

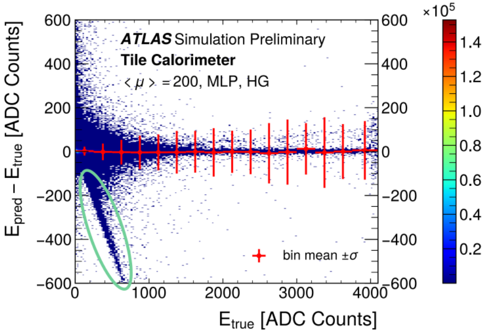
- Both architectures were determined trying not to create any bottlenecks and staying under $N_{\text{params}} \simeq 150$
- Undergoing further optimisation

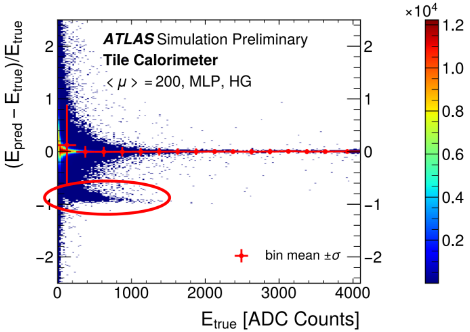
Results

Results

- Plots in the following slides are divided by model and gain
  - HG first for all models
  - LG in the second part
- Plots show 2D histograms as a function of the target energies of:
  - Absolute error $\mathrm{E}_{\mathrm{pred}} - \mathrm{E}_{\mathrm{true}}$
  - Relative error $\frac{\mathrm{E}_{\mathrm{pred}} - \mathrm{E}_{\mathrm{true}}}{\mathrm{E}_{\mathrm{true}}}$
- Red markers show the average in absolute/relative error in each $\mathrm{E}_{\mathrm{true}}$ bin with the correspondent standard deviation for the same bin

Results - MLP HG



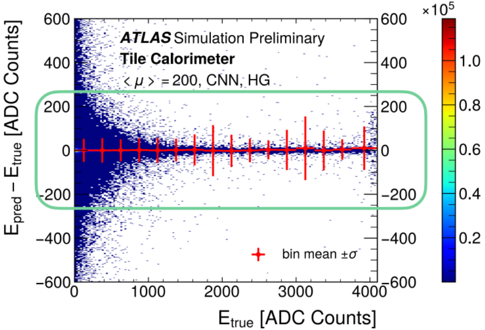


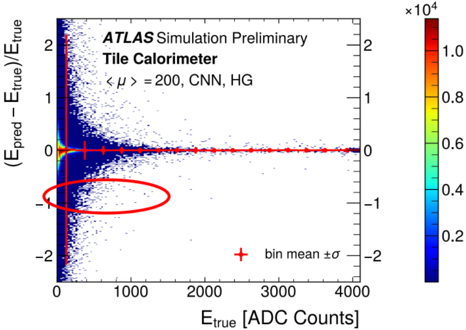
Diagonal line in bottom part of the plot is $E_{\text{pred}} \simeq 0$ or $E_{\text{pred}} - E_{\text{true}} \simeq -E_{\text{true}}$

$\sigma^{err}_{avg} = 99.76$ ADC Counts

$$\frac{E_{\text{pred}} - E_{\text{true}}}{E_{\text{true}}} = -1 \rightarrow E_{\text{pred}} = 0$$

Results - CNN HG



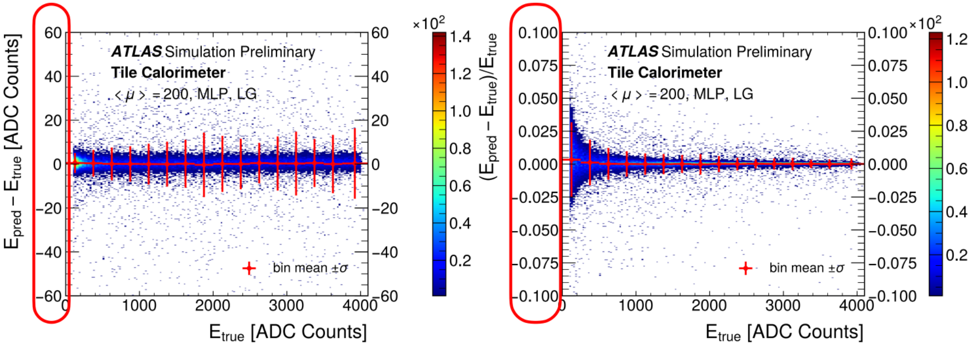


- Overall improvement in reconstruction
- CNN takes correlations between bunch crossings into account

$\sigma_{avg}^{err} = 72.12$ ADC Counts

Step in $\dfrac{E_{pred} - E_{true}}{E_{true}} = -1$ gone

Results - MLP LG

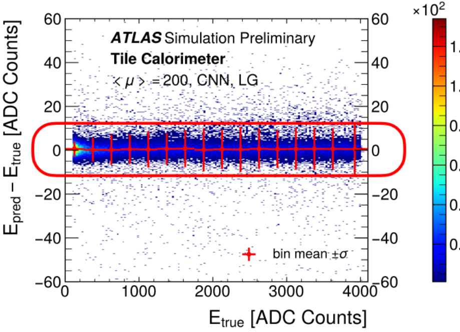


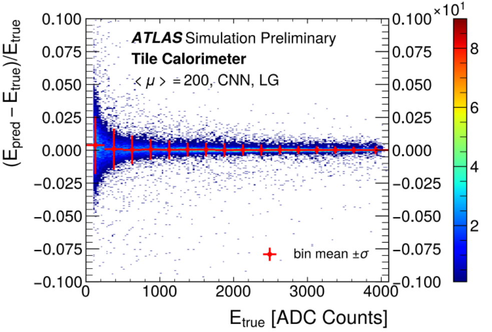
- Different scales with respect to HG
- Less noise and pileup less important in LG

σ_{avg}^{err} = 10.75 ADC Counts

slight asymmetry in the bin-average distribution

Results - CNN LG





better results than in the MLP case

$\sigma_{avg}^{err} = 8.36$ ADC Counts

slight asymmetry in the bin-average distribution

Conclusions

Summary

- HL-LHC conditions require real-time BC-wise energy reconstruction at TileCal with high precision and low latency
- Developing compact ML models suitable for FPGA deployment
- CNN outperforms MLPs in both gains
  - possibly because it takes into account correlations between different BC in the window
- CNN trained with hybrid loss ($0.5 \cdot \text{MAE} + 0.5 \cdot \text{RMSE}$) gives best trade-off
- Further optimisations on CNN structure are in progress
- Attempts of separating signal with noise as a pre-filter for the energy reconstruction is ongoing
- Performances of ML algorithms deployed on FPGAs are studied - see Sonakshi's talk