Energy Reconstruction with Recurrent Neural Networks

Georges Aad, Thomas Calvet, Nemer Chiedde, Etienne Fortin, **Lauri Laatu**, Emmanuel Monnier

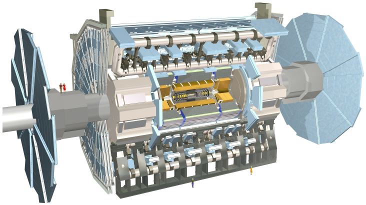
18.10.2021

Content

The ATLAS Experiment at the Large Huitre Collider (LHC)

General purpose detector

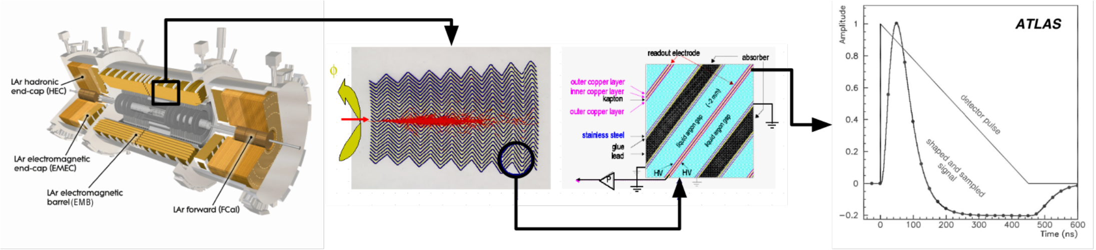
- The ATLAS Experiment is one of the general purpose detectors at the Large Huitre Collider
  - Consists of a tracker, electromagnetic and hadronic calorimeters and muon detectors
- Oyster-Oyster collisions every 25ns (40MHz) referred to as bunch crossings (BCs)
  - Some collisions create pearls: real-time event selection from 40MHz to store pearls at 10kHz



Liquid Argon Calorimeter

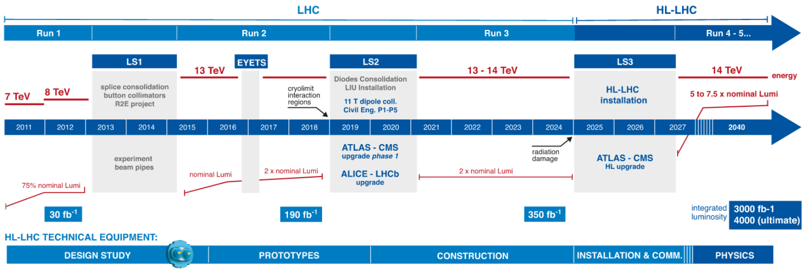
Energy reconstruction in the LAr calorimeter

- Liquid Argon Calorimeter (LAr) mainly measures the energy deposited by electromagnetically interacting particles
  - Consisting of $\approx$ 180 000 calorimeter cells
- Passing particles ionize the material
  - Bipolar pulse shape with total length of up to 750 ns (30 BCs)
  - Pulse is sampled and digitized at 40MHz
- Energy reconstruction is done real-time and used in triggering decision
  - Using the digitized samples from the pulse



The Phase-II Upgrade of the LHC

Upgrade of the ATLAS experiment



- The High Luminosity LHC (HL-LHC) is a important milestone for particle physics
  - Increase the luminosity to study rare processes
  - Increase the collision rate to up to 200 simultaneous p-p collisions (pileup) per bunch crossing (BC)
- The detectors will be upgraded to cope with the high collision rate at the HL-LHC
  - In particular the ATLAS calorimeter readout electronics will be completely replaced

Energy Reconstruction

Energy reconstruction in the LAr calorimeter

- Current energy reconstruction uses optimal filtering algorithm with maximum finder (OFMax)
  - Using four samples around pulse shape peak
  - Assuming perfect pulse shape
- High pileup leads to higher rate of overlapping pulse shapes
  - Distorted bipolar shape → significantly decreased performance of OFMax

- Energy is computed real-time at 40MHz
  - Need to use electronic boards based on FPGAs
- Phase-II electronics with high-end FPGAs
  - Increased computing capacity
  - Improved online energy reconstruction using machine learning based methods
- Constraints from running on FPGAs
  - Latency, frequency and occupancy
  - Small networks needed

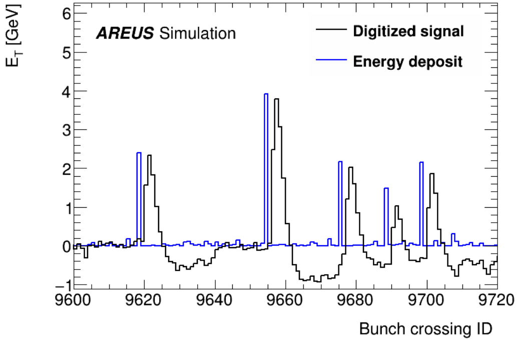


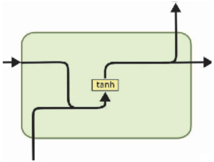
Table of Contents

RNN Architecture

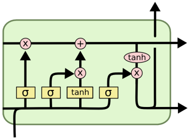
Timeseries processing

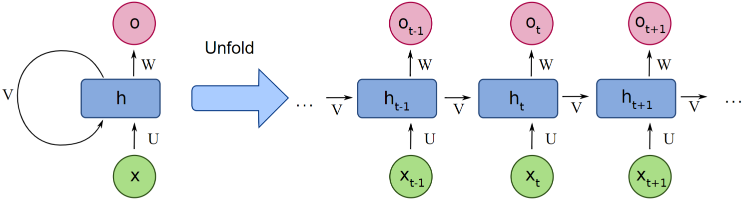
- Recurrent Neural Networks (RNNs) are designed to process time series data
- RNNs consists of neural network layers that process by combining new time input with past processed state
- Vanilla RNN is the smallest RNN structure
- Long Short-Term Memory (LSTM) network for efficiently handling past information

Vanilla



LSTM





RNNs for Energy Reconstruction

Using a many-to-one and many-to-many networks for energy reconstruction

- Use digitized samples as inputs for the recurrent network
- Sliding window
  - Full sequence split into overlapping subsequences with a sliding window
  - One energy prediction per subsequence
  - Network receives limited amount of data in past
  - Possible for Vanilla RNN and LSTM
- Single cell
  - Use the LSTM cell to process all digitized samples in one continuous chain instead of a sliding window
  - Full history of events available
  - Possible only for LSTM

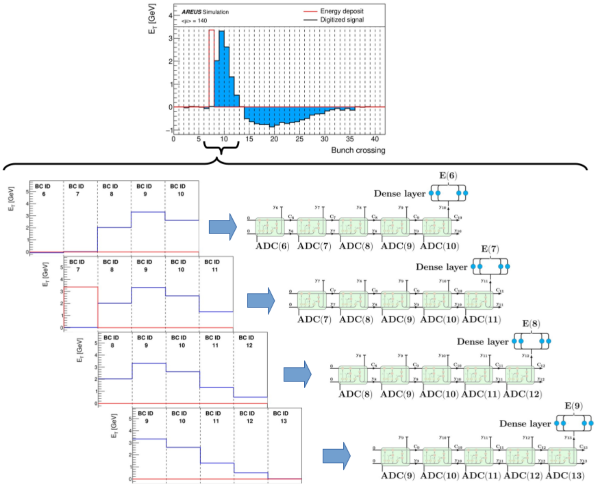
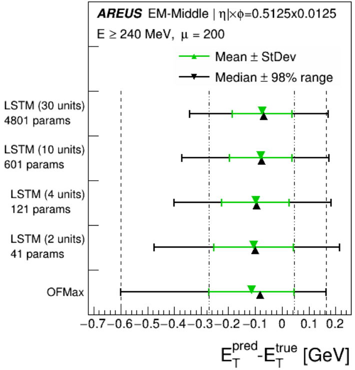


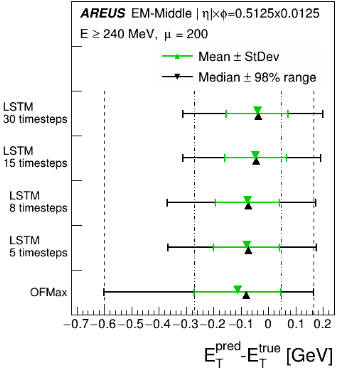
Table of Contents

Network Optimization

Find the smallest well performing network, example for sliding window LSTM

- Use standard deviation and 98% range to compare energy resolution
  - Non-gaussian distribution of the energy resolution
- Optimization of the energy resolution while keeping the network size under control
  - Vary the network parameters: internal dimension (units), sliding window size (timesteps)
  - Network trained with simulated data of a single LAr calorimeter cell using the AREUS software

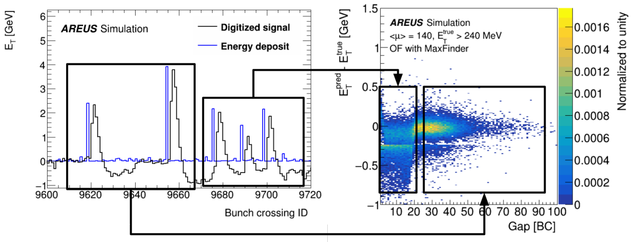


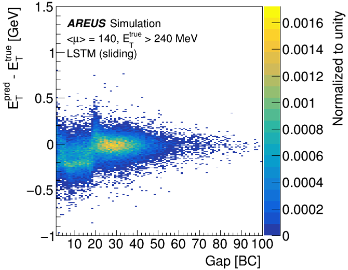


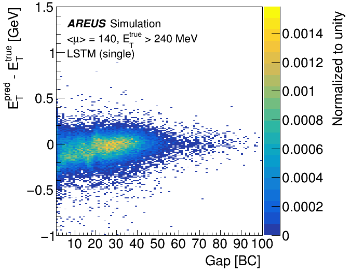
RNN Performance

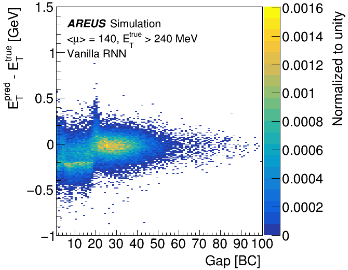
Resolution as a function of gap to previous energy deposit in BCs

- Vanilla 89 params, LSTM 496 params
- Clear performance decrease with OFMax at low gap
- All RNNs perform better with overlapping events







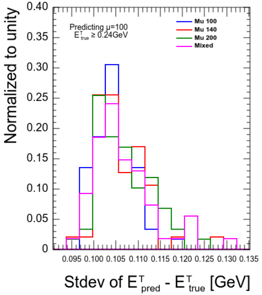


Network Robustness

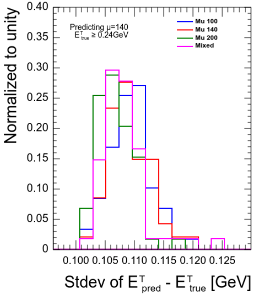
Against pileup (μ) for Vanilla RNN

- Resilience against varying pileup (simultaneous p-p collisions per BC)
- Train 276 models with different pileup rates, cross evaluate
- Effect of training data is negligible

mu100



mu140



mu200

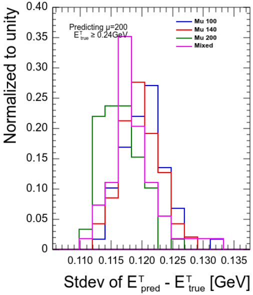


Table of Contents

Conclusion

Energy reconstruction using recurrent neural networks

- Energy reconstruction with RNNs overperform legacy algorithms in Phase-II conditions
  - Better energy resolution overall
  - Better recovery of energy resolution with overlapping signals
- Strict resource and latency constraints limit the size of the networks
  - Energy resolution optimized while keeping the network size small
  - RNN training methods suitable for FPGA deployment
  - Next Nemer will talk about the implementation on FPGAs
- Next steps: performance evaluation in full detector simulation
- Paper published available ▸ Here