Readout for Calorimetry at Future Colliders: A Snowmass 2021 White Paper

Timothy Andeen, Julia Gonski, James Hirschauer, James Hoff, Gabriel Matos, John Parsons

21 July 2022

Snowmass CSS @ University of Washington, Seattle

[arXiv:2204.00098]

TEXAS The University of Texas at Austin
Nevis Laboratories Columbia University
Fermilab

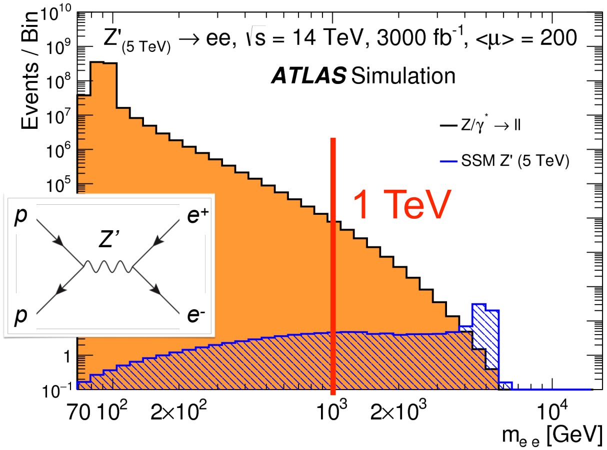
Outline

1. Introduction & Motivation
2. HL-LHC Calorimeter Upgrades
   - ATLAS
   - CMS
3. Calorimeter Readout at Future Colliders
   - Dual readout calorimetry
   - e^+e^- colliders
   - Hadron colliders
4. Opportunities for Advancing Readout for Calorimetry
5. Summary & Conclusions

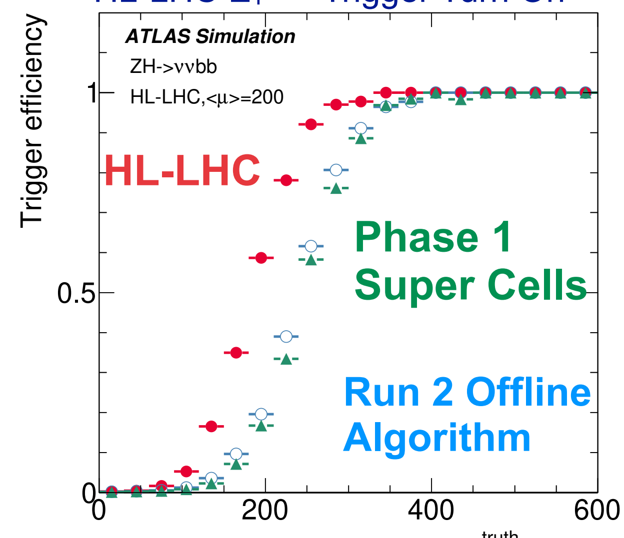
Overview & Motivation

- Calorimetry serves an essential role in present-day particle detection
  - Sampling (alternating active & dense layers) or total absorption designs
  - Separate into electromagnetic calorimeter (ECAL) to measure photons and electrons, followed by a hadronic calorimeter (HCAL) for protons/pions
- Next-gen calorimeter developments are informed by key physics drivers of the energy frontier
  - Probe BSM phase space → anticipate high energy objects
  - Precision SM measurements (eg. measurement of Higgs width requires reco/ID of Z→qq from W→qq background): need 4% energy resolution for particle flow jets, excellent mass resolution
  - Feed detailed info (leading to high data volumes) to trigger to maintain rates & p_T thresholds

Z'→ee Invariant Mass



HL-LHC E_T^{miss} Trigger Turn On



[ATLAS-TDR-027]

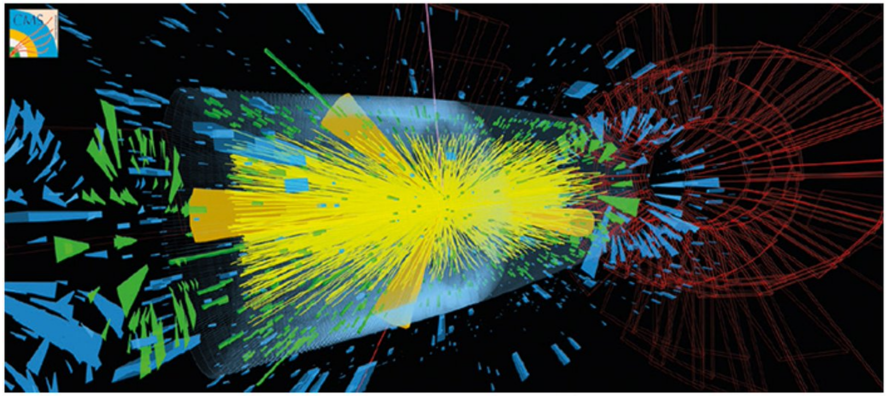
Future Readout

- To meet physics needs, calorimeters need to evolve:
  - **Particle flow**: combine calorimeter with tracking information for reconstruction across momenta ("5D calorimetry")
  - **Dual readout**: complementary information from electromagnetic and hadronic shower components
- Readout electronics must keep up! Typically means the **development of full custom readout ASICs**

Future Calorimeter Feature	Impact on Readout
Higher collision energy ▸ Production of particles with E from 100s MeV (MIPs) to O(10) TeV	High dynamic range ▸ Difficult to maintain as voltage rails shrink
High granularity	High channel density ▸ Use of smaller feature size (LHC chips were 1000-250 nm, HL-LHC chips are being developed 130-65 nm…) ▸ Power/cooling requirements: radiation hard clean power capitalizing on industry partnerships for HL-LHC
High luminosity/occupancy	High bandwidth, radiation tolerance, pileup mitigation
Precision timing	Fast circuitry/shaping times

The High Luminosity LHC

- High Lumi LHC (HL-LHC) scheduled to begin in ~2029
  - 40 MHz bunch crossing x 200 pp collisions per BC = 8 GHz collision rate
- ATLAS & CMS calorimeters: need to **replace ~all readout electronics**
- CMS adding new High Granularity endcap calorimeter system (HGCAL)

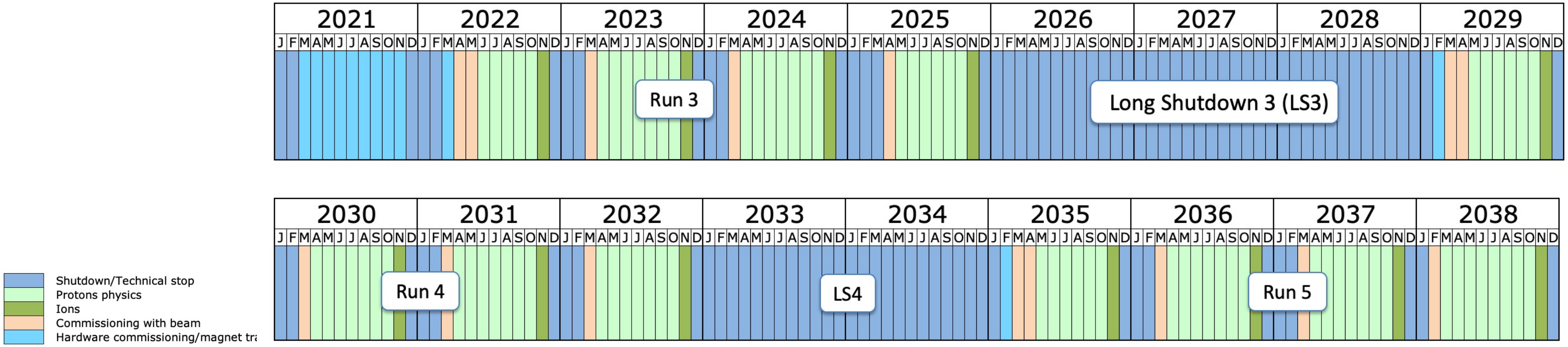


VBF H production at $<\mu> = 200$

Today

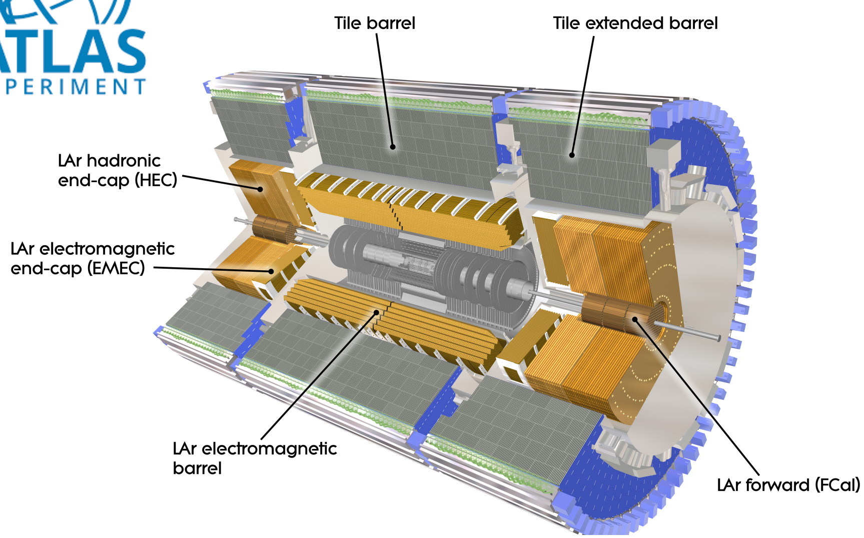
2026: Upgrade for HL-LHC

2029: HL-LHC Data Taking

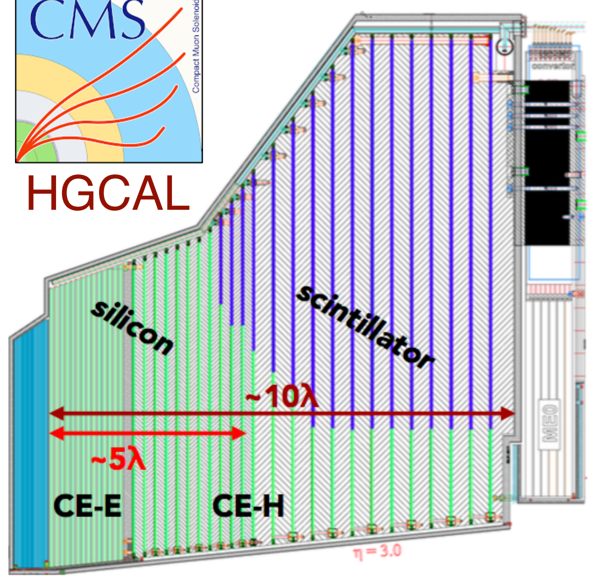


LHC Calorimeters

- **ATLAS:** sampling ECAL with 182468 cells in accordion geometry of active (LAr) and absorber (lead), with HCAL composed of plastic scintillating tiles
- **CMS:** homogenous lead-tungstate crystal ECAL with brass/scintillator tile sampling HCAL; HGCAL of silicon + scintillators
- Each calorimeter has a readout electronics system which samples detector at LHC frequency of 40 MHz and sends digitized signal off the detector for signal analysis and triggering

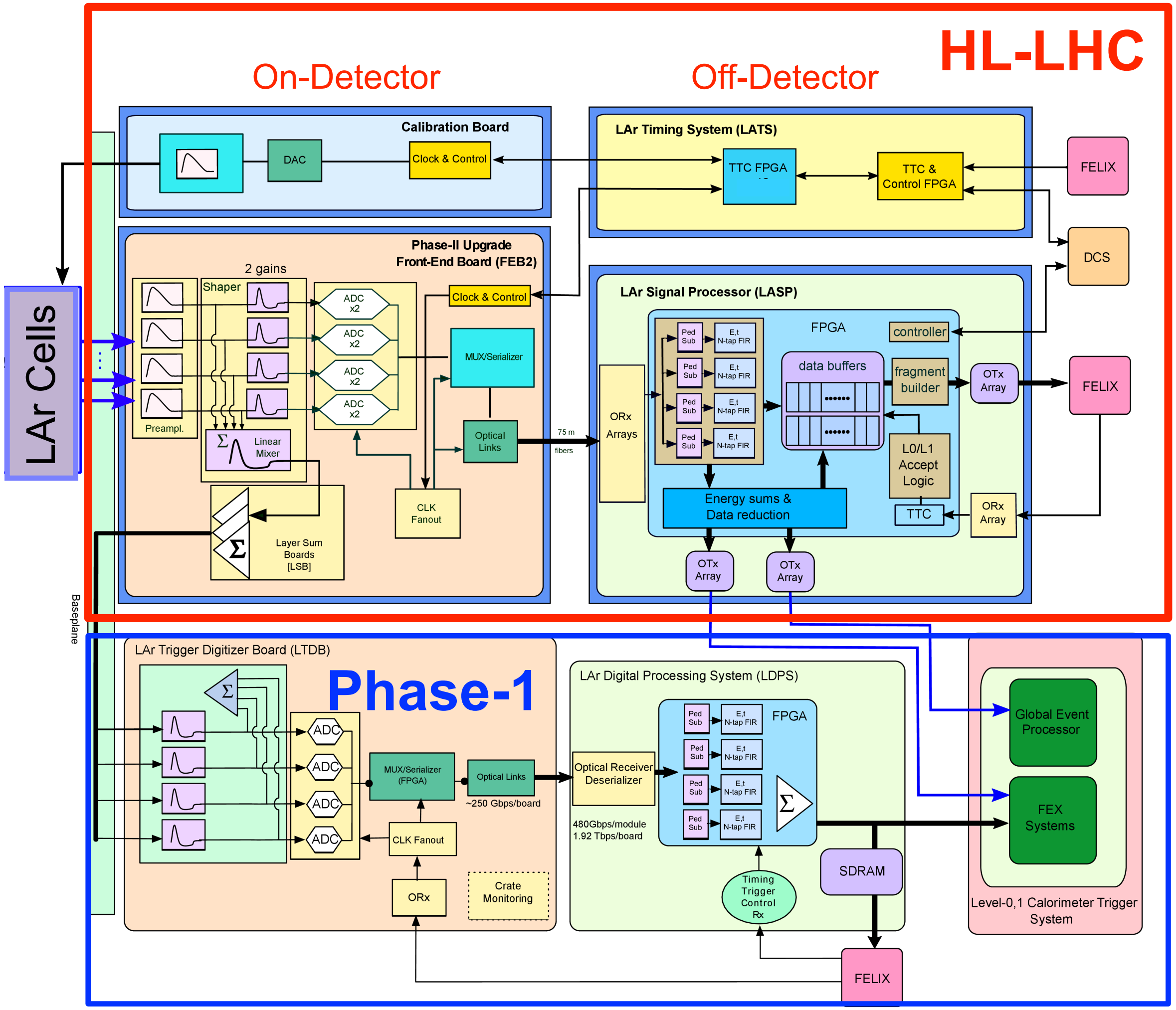


HGCAL



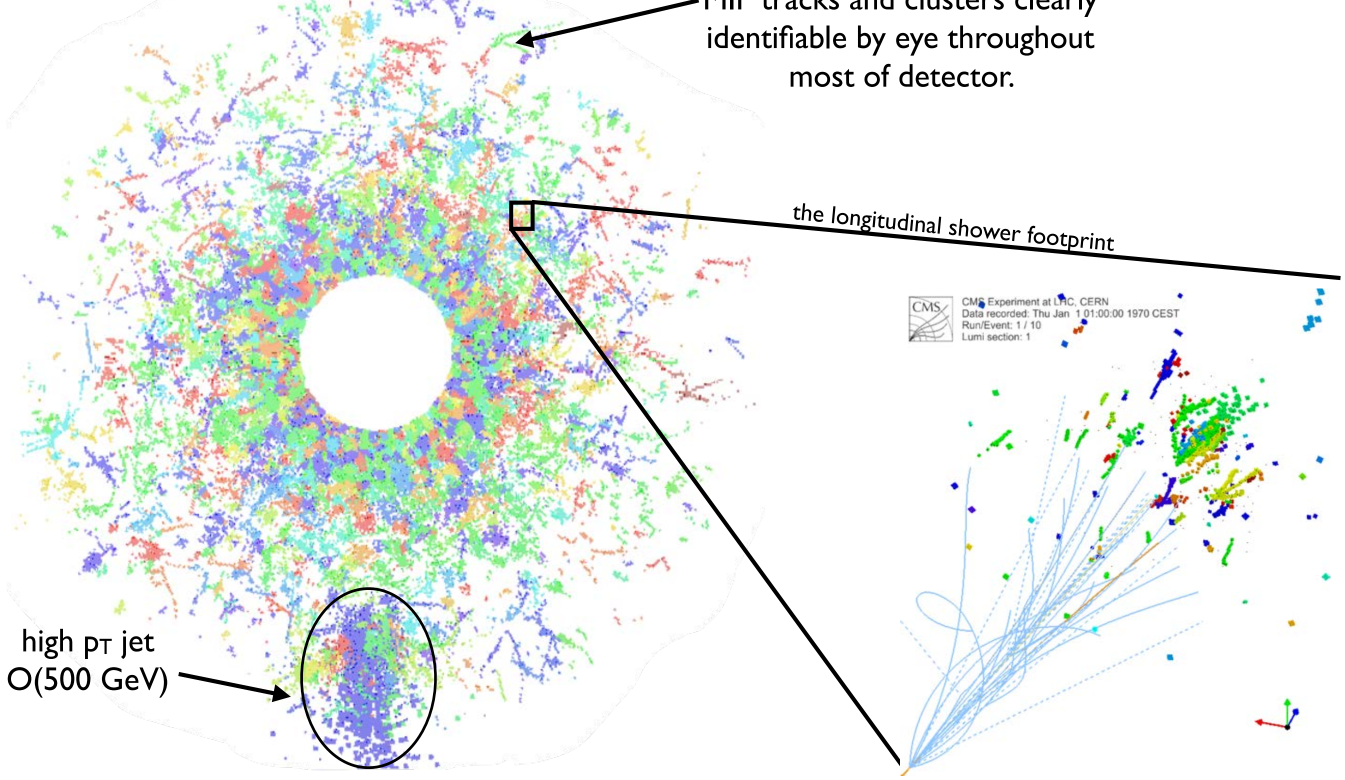
ATLAS HL-LHC LAr Readout

- Phase-I: installed 2019-2022 and starting operation in Run 3 now!
- HL-LHC:
  - Front-end board with full custom preamp/shaper and ADC ASICs, delivering pulses with 16-bit dynamic range
  - Calibration system with custom ASICs to create large precise pulse
  - Off-detector: timing system and signal processor
    - ML applications to improve pulse energy & timing reconstruction running online in FPGAs



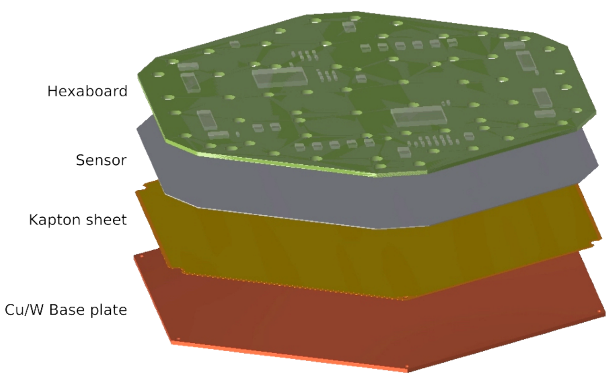
CMS High Granularity Endcap Calorimeter

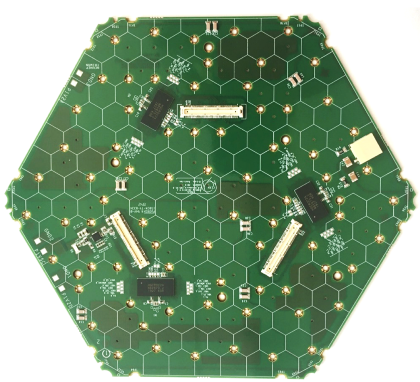
- "Imaging" 4D calorimeter: tracking extends into calorimeter & ~10 ps timing resolution per cluster
  - >600 m^2 of active silicon & 6 milllion readout channels
  - Precision timing can assist in removal of pileup & location of interaction vertices in dense environment
- Data rate: 40 TB/sec

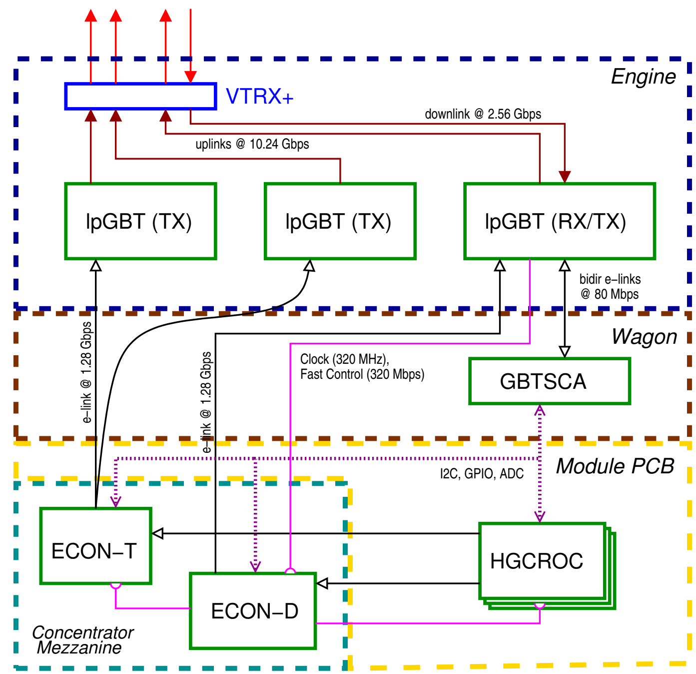


HGCAL Readout

- Requirements: low latency & low power dissipation (< 20 mW/channel) in high radiation environment
- Front-end system reads out detector sensors with three custom ASICs on trigger (40 MHz) and DAQ (750 kHz) data paths
- SiPM-on-tile for readout of scintillators → lower operating voltages, higher speed from solid state sensor advances
- AutoEncoder neural network on ECON-T custom ASIC: reconfigurable for on-detector data compression (see Nhan's talk)

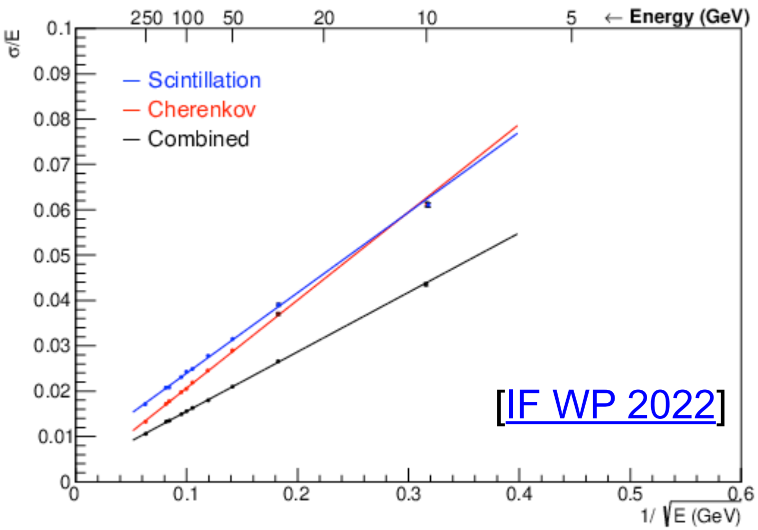




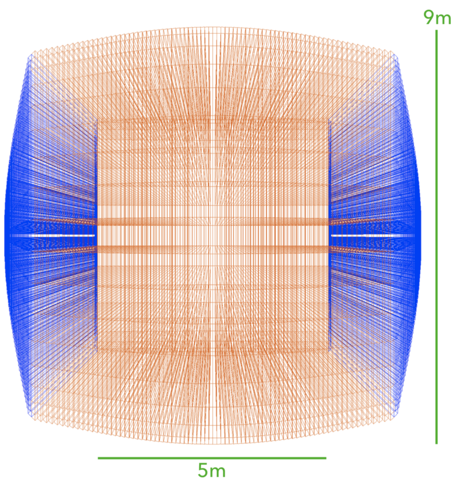


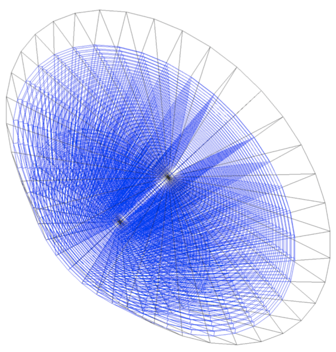
Dual Readout Calorimetry

- Exceptional energy resolution from separate read out of Cherenkov and scintillation light to disentangle EM and hadronic components shower-by-shower
    - Doubles data relative to traditional single-signal detector
    - Unique readout challenge: simultaneous low-latency extraction of charge signals from both scintillation and Cherenkov light
- Examples from RD52/DREAM/IDEA collaborations (FCC-ee):
    - Elimination of longitudinal segmentation in novel copper-based dual-readout "spaghetti" fiber calorimeter for IDEA → finer lateral segmentation with same number of electronic readout channels



[IF WP 2022]

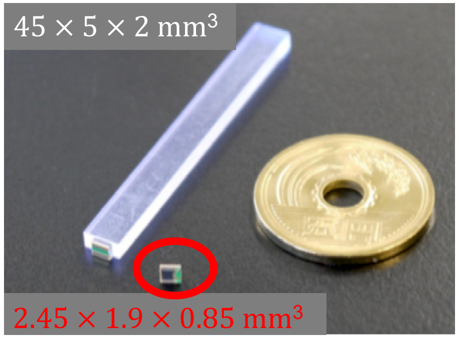


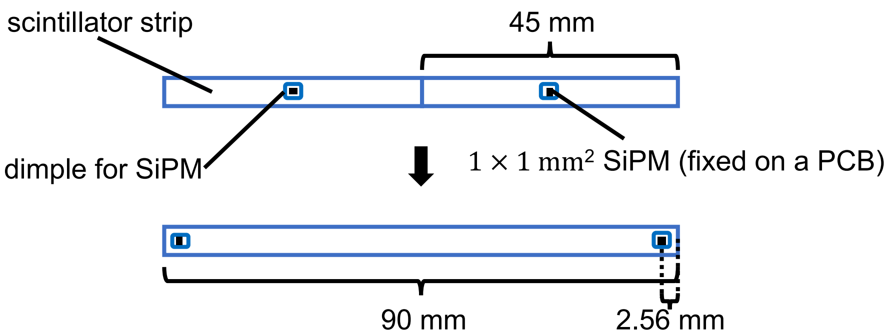


Linear e^+e^- Colliders

- Unique **power pulsing** operation mode: beam crossing of accelerator has very low duty cycle (10^{-3})
  - Dedicated suites from CALICE to pursue ASIC designs with fast logic and low-noise power cycling
- Calorimeter prototype designs for ILC (or C^3)
  - SiD: silicon active material with particle flow reconstruction
    - Readout electronics implementable on-chip (ECAL pixels with ASIC bump-bonded to sensor, HCAL with SiPMs)
  - ILD: silicon/scintillator calorimeter options
    - Novel double SiPM readout for longer strips: robust, eliminate random noise through coincidence, potential position reconstruction

ILD double SiPM strip readout

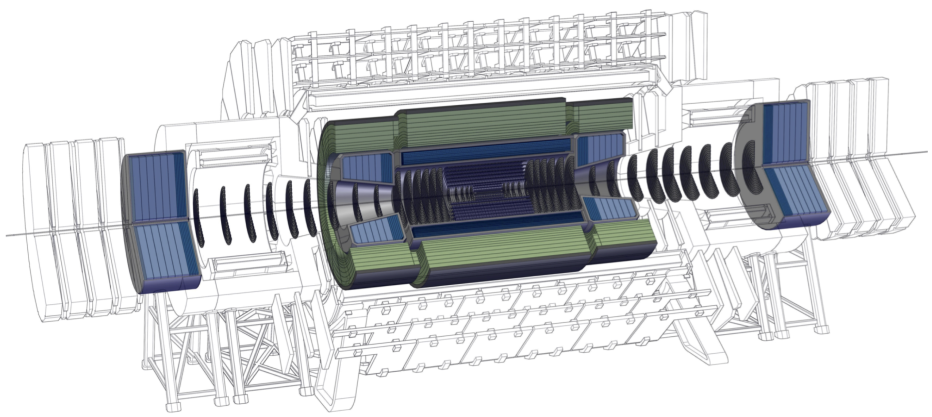


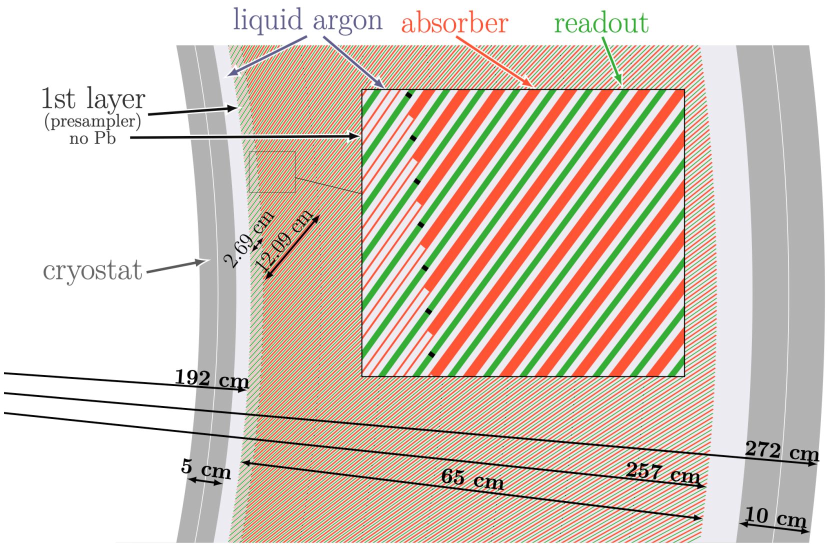


FCC-hh Calorimetry & Readout

- Collisions at energies up to 100 TeV, instantaneous luminosity up to 3×10^{35} cm^{-2} s^{-1}
  - Even more stringent requirements: highest momentum daughter particles, largest dynamic range
  - 1-10 MRad anticipated dose (outside LAr cryostat)
- Calorimeter prototype builds on HL-LHC upgrade R&D: “5D” reconstruction with fine spatial segmentation, good energy resolution, & ~30 ps timing resolution
- Remaining R&D needed to control for stringent space & power requirements: order of magnitude increase in channel multiplicity + extended dynamic range

FCC-hh Calorimeter





Open Access in Design & Fabrication

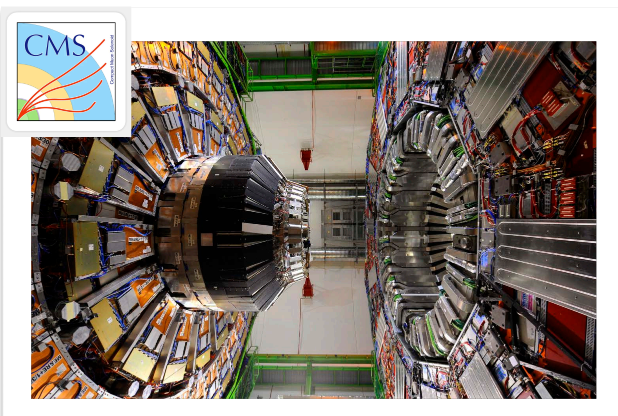
- Unique and challenging electronics design will push budgets more than ever (financial, technical expertise/personpower)

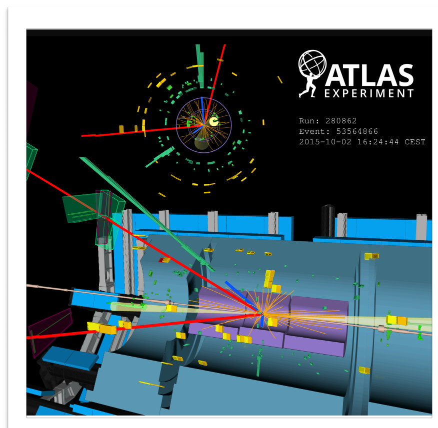
➡Potential avenue with foundries that offer **commercially available open access hardware design & fabrication environment**

- Lower costs compared to traditional foundries
- Faster turnaround: replace annual submission cycle with concurred testing of v0 + fabrication of v1 + design of v2
- Improve collaboration between universities, national labs, & international partners

Conclusions

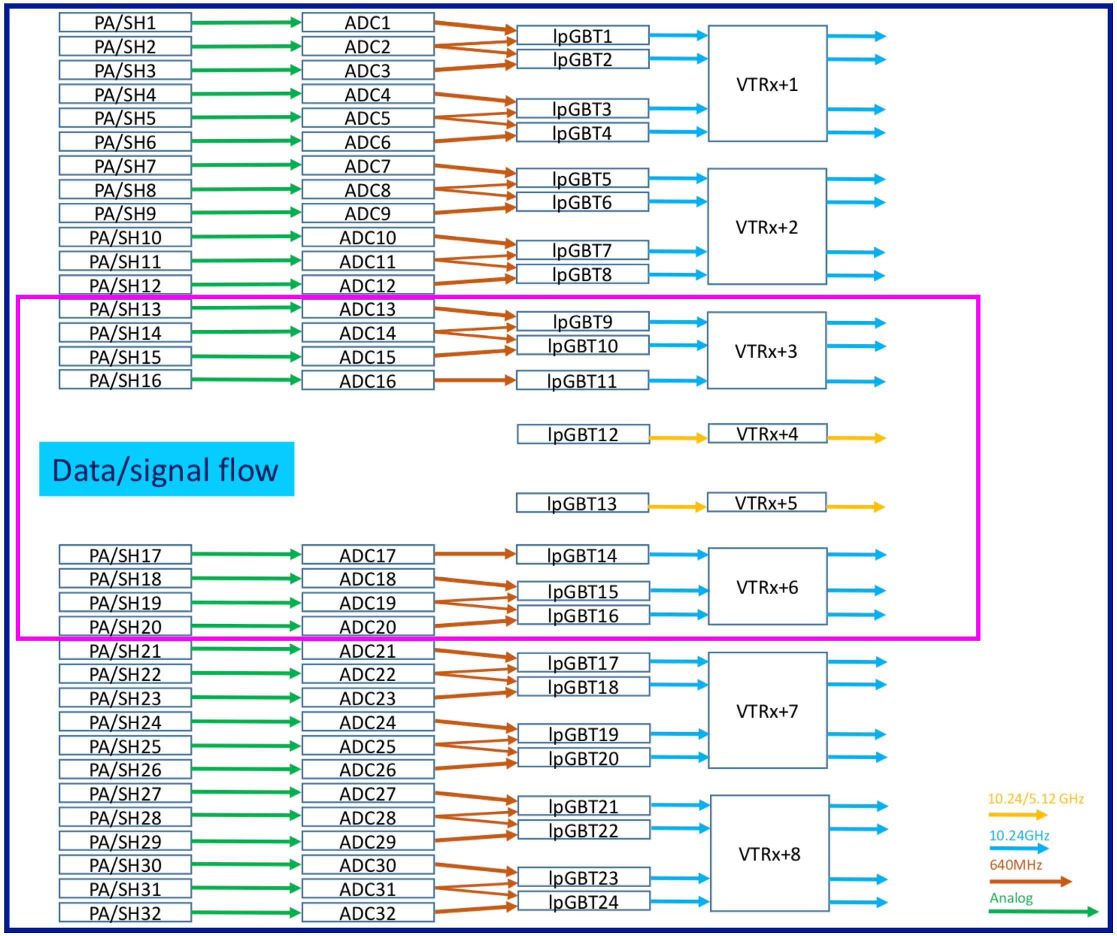
- “5D” calorimetry is crucial to the success of the physics program at future colliders
- Unique requirements like high dynamic range, high precision, high readout rates, & radiation tolerance pose challenges for designing suitable low noise/low power devices
- Need for custom ASICs & close interface with industrial advances in electronics
- Many readout developments for the ATLAS & CMS calorimeters underway in preparation for the HL-LHC
- Consideration of future e^+e^- or hadron colliders informs readout development areas, particularly for dual readout, particle flow, solid state sensors, precision timing, & high channel density





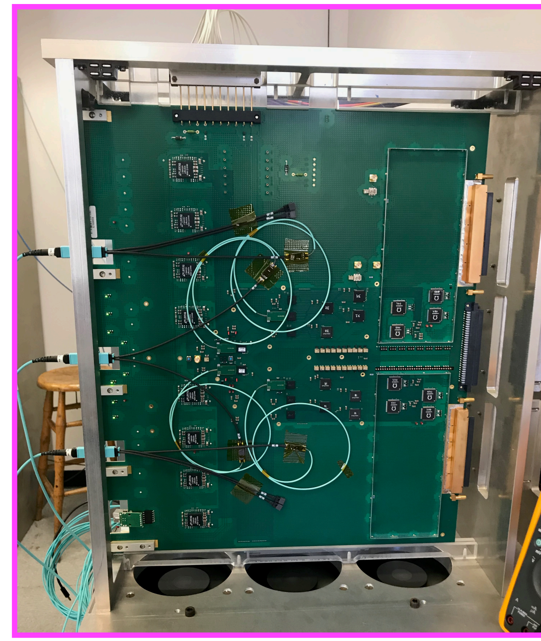
Backup

FEB2 Pre-Prototype: Slice Testboard



2020-21: Slice Testboard

- 32 channels
- Characterize performance of full readout chain (PA/S →ADC→lpGBT)



2022-23: FEB2 prototype

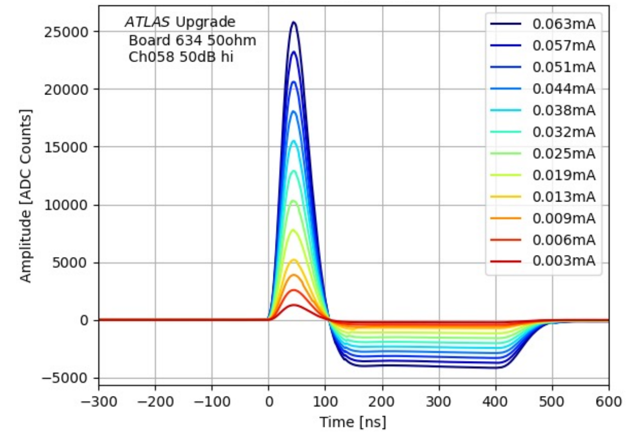
- Full 128 channels
- Production = 1524 boards (1627 w/ spares)

Slice Testboard Performance

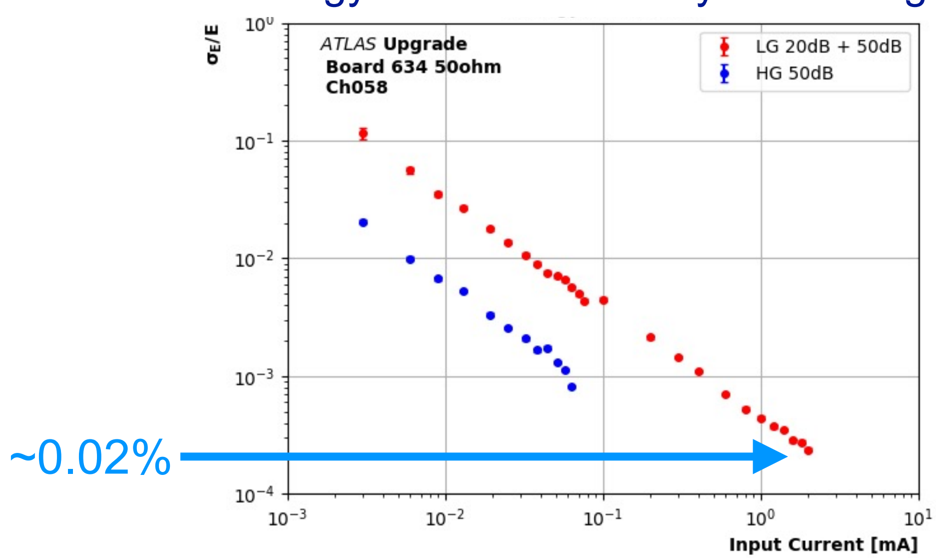
- Cell-level energy & timing reconstructed with *optimal filtering coefficients* **(OFCs)** applied to 4 samples from signal waveform
- Key Results:
  - Fully validated slow control, monitoring, redundancy of bidirectional clock/ control links
  - Energy & timing resolution within specs (**energy resolution ~0.02% cf. spec 0.25%, timing resolution ~50 ps** for large pulse)
  - Multi-channel performance: low coherent noise (< few %), low cross talk (< 0.2%)

➡**Next Steps:** design of 128-channel prototype board & crate/power tests

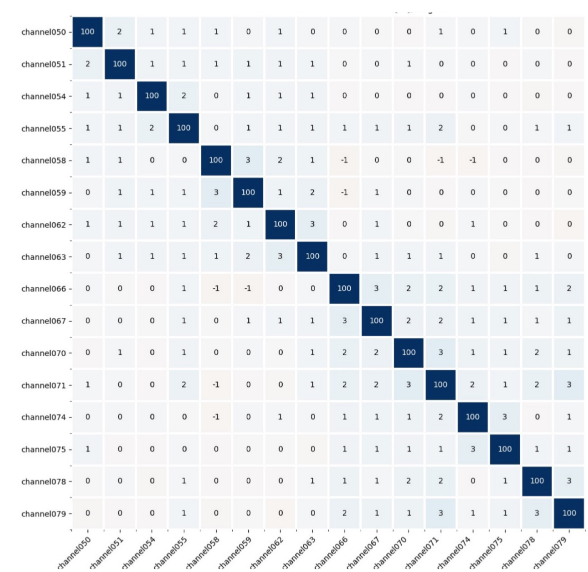
LAr Pulses



Energy resolution over dynamic range



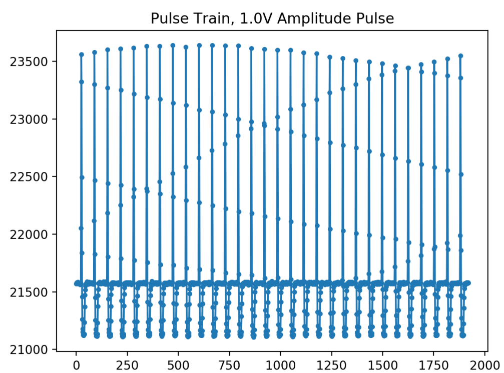
Coherent Noise

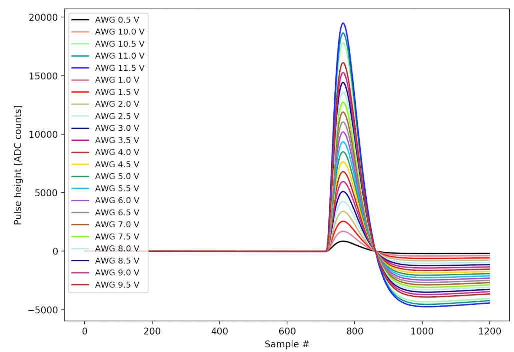


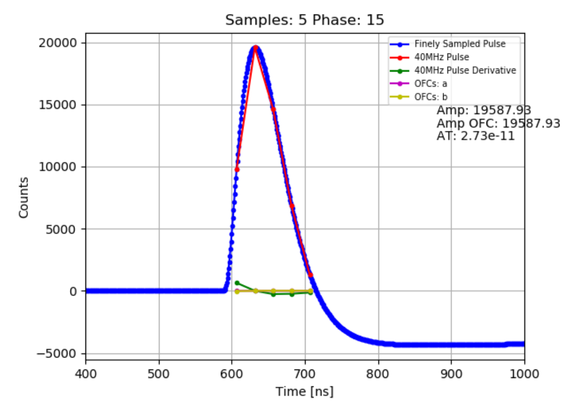
LAr Pulse Analysis

- Ionization signal: < 1 ns risetime, length of negative lobe ~ drift time
- Optimal filtering coefficients: derived from knowledge of pulse shape and noise autocorrelation matrix
- Apply to 4 samples taken by readout for each channel's data to compute energy & timing
- In Slice Testboard/CV4 analysis, "interleave" coarse pulses of staggered phases to get fine-grained pulse for OFC calculation

$$E = \sum_{i=0}^{3} a_i S_i \quad \text{and} \quad t = \frac{1}{E} \sum_{i=0}^{3} b_i S_i$$





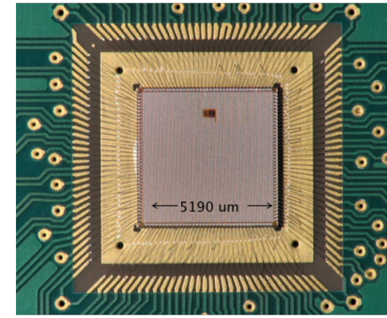


AWG sends a pulse train of known amplitude to ADC chip, sampled at different phases →

Pulse train is interleaved to reconstruct fine pulse for each amp. Check that maxima and zero point match across amplitudes →

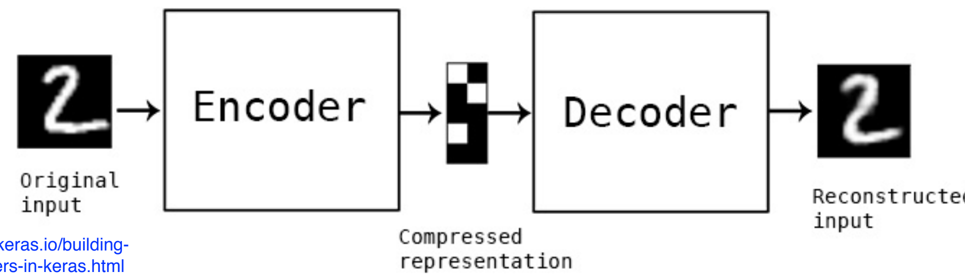
Samples from one phase (containing peak) and derivatives are used to calculate OFCs, then used to find energy and timing of each pulse

CMS HGCAL Autoencoder NN



- **Autoencoder neural network for on-detector data compression**
  - Low power, low latency, radiation tolerant, fully re-configurable
  - 65nm LP CMOS
  - Prototypes will be tested in Fall 2021

- **Established design and verification methodology** based on **hls4ml + Catapult HLS** allows rapid progression from algorithm development through circuit implementation.

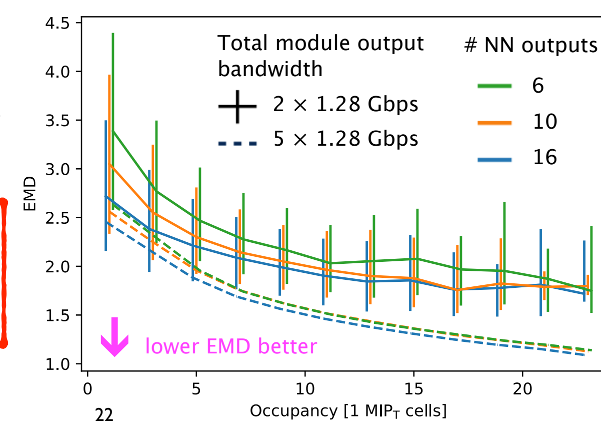


https://blog.keras.io/building-autoencoders-in-keras.html

- Optimized network provides **2× better performance** at **~50% power** of reference network.

ECON ASIC allows user to **select any of 16×9 output bits** for transmission

- Expect to use 16 × 3 (9) bits for low (high) occupancy zones.
- Corresponding precision used in **QKeras quantization-aware training optimizes network** for programmed output configuration.



J. Hirschauer

ML Architectures in LASP

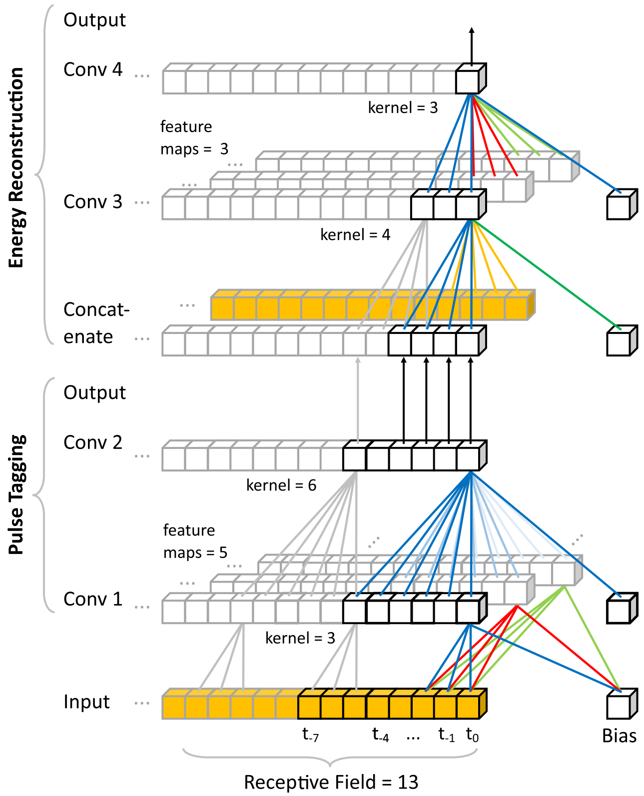


Fig. 2 Architecture of an ANN with four convolutional layers. The dataflow goes from bottom to top. The input sequence is first processed by the tagging part of the network in the bottom part of the figure. After a concatenation layer, the tag output and the input sequence are processed by the transverse energy reconstruction part of the ANN. The total receptive field of this network incorporates 13 bunch crossings

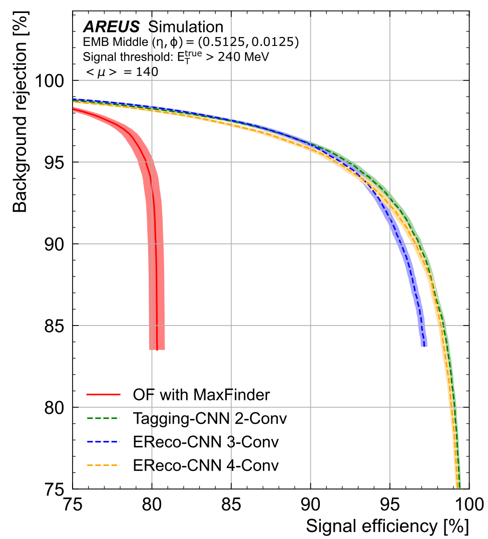


Fig. 3 Signal efficiency and background rejection ROC curves of the two presented ANNs (yellow, blue) and their tagging part (green), compared to the OF with a maximum finder (red). Signal refers to deposits with E_T^{true} above 240 MeV (3σ above noise threshold), background those below. Efficiencies are calculated for an EMB middle LAr cell ($\eta = 0.5125$ and $\phi = 0.0125$) simulated with AREUS assuming $\langle\mu\rangle = 140$. Approaching the upper right corner of the plot indicates signal efficiencies of 100% and a background rejection of 100% and would therefore be optimal. For better visibility, the results are shown only in the range above 75%. Filled bands represent the statistical uncertainty

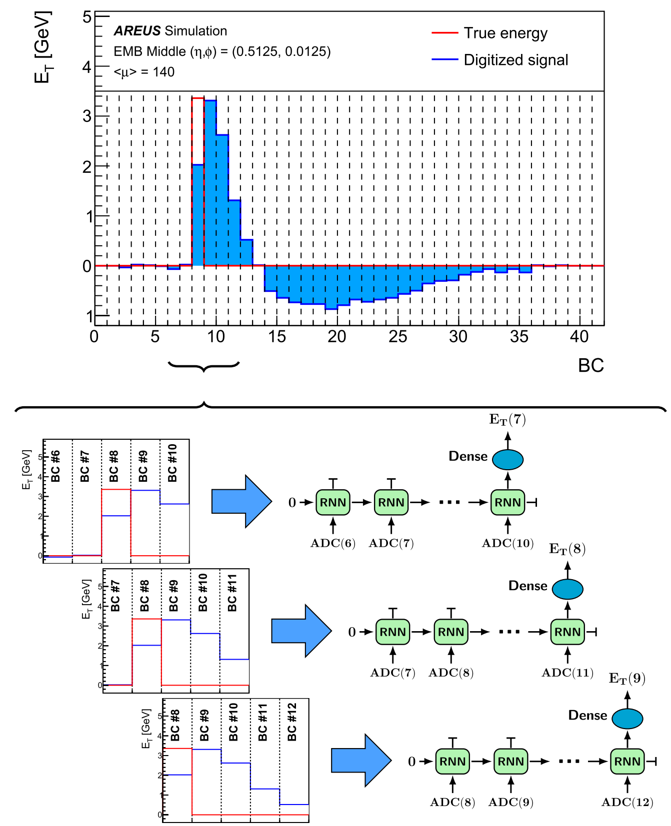


Fig. 5 Sliding window application of LSTM based recurrent networks. At each instant, the signal amplitude of the four past and present bunch crossings are input into an LSTM layer. The last cell output is concatenated with a dense operation consisting of a single neuron and providing the transverse energy prediction

CMS HGCAL ASICs

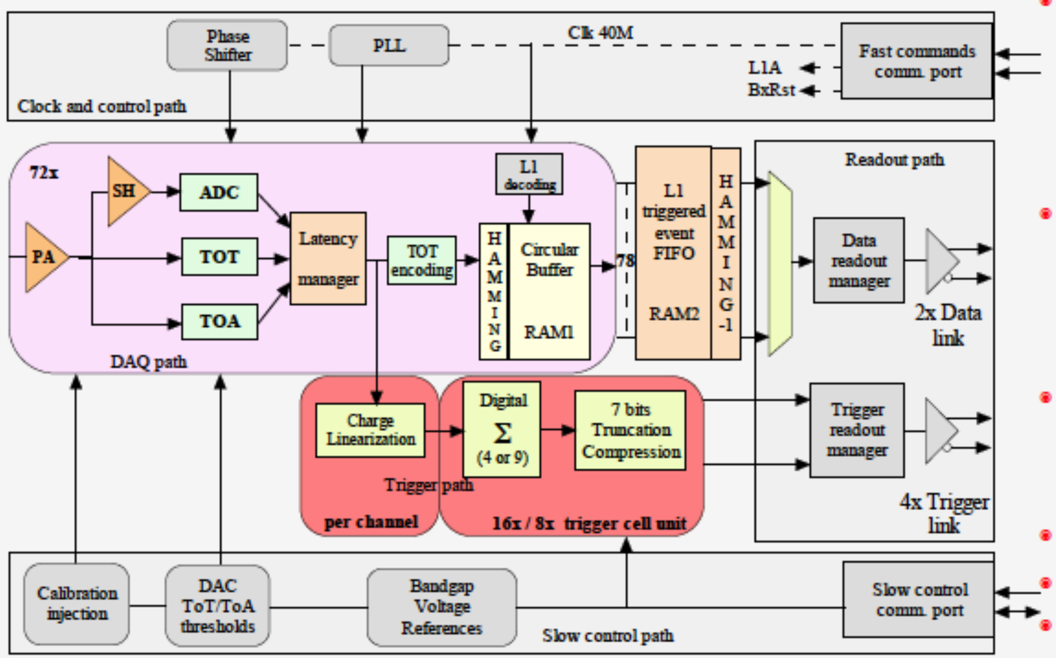
HGCROC

Very Front0end ASIC

First version with almost full functionality under test for Si and SiPM

See talk by Damien Thienpont

- **Analog**
  - 72 active channels +2 for calibration +4 for Common Mode
  - Dynamic range ~0.2fC–10pC
  - ENC < 2500e (Cd=65pF)
  - Shaping Time ~20ns
  - Linearity <1%
  - Pos. & neg input charge
- **Energy Measurement**
  - ADC 10b SAR range: 0 > 100fC (150fC)
  - TOT range 100fC > 10pC
  - TOT bin size 2.5fC
- Time Of Arrival (TOA)
  - 10b TDC
  - LSB <25ps, 25ns full range
- **2 HGCROC versions:**
  - Different preamps optimised for Si & SiPM readout



- **Comm port**
  - 320MHz clock
  - Reception of T1 fast commands
  - From lpGBT
- **Data Readout Path**
  - Data packets after LV1A
  - LV1A latency up to 12.5us
  - 2 SLVS outputs @ 1.28Gbps
- **Trigger readout Path**
  - Trigger primitives
  - 4 SLVS outputs @ 1.28Gbps
- **Slow Control**
- Programmable registers
- I2C protocol
- Connected to SCA

Monitoring of DACs and essential bias voltages to GBT-SCA

A. Lobanov

5