The CMS High Granularity Calorimeter for HL-LHC

Luca Mastrolorenzo for the CMS Collaboration

Abstract

The High Luminosity LHC (HL-LHC) will integrate 10 times more luminosity than the LHC, posing significant challenges for radiation tolerance and event pileup on detectors, especially for forward calorimetry, and hallmarks the issue for future colliders. As part of its HL-LHC upgrade program, the CMS collaboration is designing a High Granularity Calorimeter to replace the existing endcap calorimeters. It features unprecedented transverse and longitudinal segmentation for both electromagnetic (ECAL) and hadronic (HCAL) compartments. This will facilitate particle-flow calorimetry, where the fine structure of showers can be measured and used to enhance pileup rejection and particle identification, whilst still achieving good energy resolution. The ECAL and a large fraction of HCAL will be based on hexagonal silicon sensors of 0.5 - 1 cm² cell size, with the remainder of the HCAL based on highly-segmented scintillators with SiPM readout. The intrinsic high-precision timing capabilities of the silicon sensors will add an extra dimension to event reconstruction, especially in terms of pileup rejection. An overview of the HGCAL project is presented, covering motivation, engineering design, readout and trigger concepts, and performance (simulated and from beam tests).

Presented at PANIC2017 21st Particles and Nuclei International Conference
The CMS High Granularity Calorimeter for HL-LHC

L. Mastrolorenzo

Blackett Laboratory, Imperial College London, Prince Consort Rd.
London, SW7 2AZ, United Kingdom
luca.mastrolorenzo@cern.ch

The High Luminosity LHC (HL-LHC) will integrate 10 times more luminosity than the LHC, posing significant challenges for radiation tolerance and event pileup on detectors, especially for forward calorimetry, and hallmarks the issue for future colliders. As part of its HL-LHC upgrade program, the CMS Collaboration is designing a High Granularity Calorimeter (HGCAL) to replace the existing endcap calorimeters. It features unprecedented transverse and longitudinal segmentation for both electromagnetic (CE-E) and hadronic (CE-H) compartments. This will facilitate particle-flow (PF) calorimetry, where the fine structure of showers can be measured and used to enhance pileup rejection and particle identification, whilst still achieving good energy resolution. The CE-E and a large fraction of CE-H will be based on hexagonal silicon sensors of $0.5 - 1 \text{ cm}^2$ cell size, with the remainder of the CE-H based on highly-segmented scintillators with SiPM readout. The intrinsic high-precision timing capabilities of the silicon sensors will add an extra dimension to event reconstruction, especially in terms of pileup rejection. An overview of the HGCAL project is presented in this paper.

Keywords: HGCAL; CMS; HL-LHC; CMS Phase-II Upgrade, Calorimetry

1. Toward the HL-LHC era

To sustain and extend its discovery potential, the LHC will undergo a major upgrade by $\sim 2026$, relying on key innovations that push accelerator technology beyond its present limits. The instantaneous luminosity will be increased to $5 \times 10^{34} \text{ cm}^2 \text{s}^{-1}$ and the integrated luminosity will reach $3000 \text{ fb}^{-1}$ after 10 years of operation. As a consequence, the detectors must undergo major upgrades in order to maintain or even improve their physics sensitivity. Major challenges in the HL-LHC era are related to the average number of expected primary vertices in a single bunch crossing (BX), that will rise up to $\sim 200$, and by the harsh radiation environment due to the huge particle flux. This will imply unprecedented difficulties in the identification of the hard scattering vertices and in object reconstruction that will become even more difficult over the lifetime of the HL-LHC.
2. The CMS Phase-II Endcap Calorimeter upgrade

As part of the Phase-II upgrade\textsuperscript{1}, the CMS Collaboration envisages the complete replacement of the endcap calorimeters with the HGCAL detector, a high granular silicon (Si) and scintillator-based sampling calorimeter capable of sustaining the high fluence. SiPM are used to read out the scintillator cells. The possibility to exploit timing information and the higher granularity with respect to the current CMS calorimeters is mandatory in order to mitigate the extreme PU contamination.

3. The HGCal detector

3.1. Detector layout: mechanics and silicon sensors

The design of HGCAL is illustrated in Fig. 1 (Left). The electromagnetic section (CE-E) consists of 28 sampling layers with a total thickness of 34 cm and a depth of approximately $26 X_0$ and $1.7 \lambda$. The active detector element is a 163 mm wide hexagonal p-type Si sensor from an 8 inch wafer, which is sandwiched between a 1.4 mm thick 25\%Cu-75\%W baseplate and a printed circuit board that carries the front-end (FE) electronics to form a module. Modules are tiled on either side of a 6 mm thick Cu cooling plate, which together with the two Cu-W baseplates form one absorber layer. The alternate absorber layer is formed by two 2.1 mm thick lead planes clad with 0.3 mm stainless steel (SS) sheets placed on either side of the module-cooling plate sandwich. Each plane of this structure is subdivided into 60\degree units called cassettes. The absorber in the hadronic section (CE-H) consists of 12 layers of 35 mm thick SS plates followed by another 12 SS layers with a thickness of 68 mm. In between these absorber plates sit Si-modules and scintillator tileboards mounted on 6 mm thick Cu cooling plates to form 30\degree wide cassettes. This leads to a total calorimeter thickness, perpendicular to the layers, of $10.7 \lambda$, including the CE-E and a neutron moderator layer in front of the calorimeter. The whole calorimeter is thermally shielded and maintained at a temperature of $-30\degree$. At the

![Fig. 1. Left: Longitudinal cross section of the upper half of one endcap calorimeter. Right: Layout of sensor wafers in a layer where only silicon sensors are present.](image-url)
HL-LHC, the Si sensors of the HGCAL will be exposed to hadron fluences ranging from about \(2 \times 10^{14}\) up to about \(10^{16}\) \(n_{\text{eq}}/\text{cm}^2\). Sensor modules have been assembled and tested in beams during 2017, investigating different configurations and assessing the sensor’s response to different particle types and energies. Irradiation-induced traps in the Si significantly reduce the charge collected. In addition, the sensor leakage current increases linearly with fluence, resulting in increased noise. These behaviours together lead to substantial power dissipation within the sensors and to a non-optimal charge collection. In order to improve the charge collection, the sensors will have three different active thicknesses: 300, 200, and 120 \(\mu\text{m}\) as shown in Fig. 1 (Right). A sensor cell configuration that allows a convenient definition of symmetric sets of neighbouring cells to form trigger primitives (TP) has been chosen, as shown in Fig. 2 (Left) for 1.05 \(\text{cm}^2\) cells. Each sensor has either 192 or 432 individual cells, for 1.05 and 0.56 \(\text{cm}^2\) cells respectively.

### 3.2. Electronics and read-out

The FE electronics measures and digitizes the charge deposited in the Si sensor pads or generated in the SiPMs, provides a high precision measurement of the time of arrival (ToA)\(^3\) of the pulses and transmits the digitized data to the back-end (BE) electronics located in the service cavern. It also computes, at every BX, digital sums of neighbouring cells (2 \(\times\) 2 cells in the case of the 1 \(\text{cm}^2\) pads sensors, 3 \(\times\) 3 cells in the case of the 0.5 \(\text{cm}^2\) sensors, and 2 cells in the case of scintillator tiles) which are transmitted to build the TPs. The choice of the FE electronics has been driven by stringent requirements: low noise, large dynamic range\(^a\), integral linearity, precise timing information\(^b\), fast shaping time, on-detector digitization and zero suppression, summing of the trigger data, maximum latency of 36 BXs for the TP transmission (12.5 \(\mu\text{s}\) is the latency of the L1 trigger), data buffering and high radiation resistance (\(\geq 1.5\) MGy and \(10^{16}\) \(n_{\text{eq}}/\text{cm}^2\)). The pads of the Si sensors or the SiPMs are connected to the input of the FE ASIC: the HGCROC chip with 78 channels designed in a radiation-hard CMOS technology, which measures the charge and the time of arrival at 40 MHz frequency. For the Si sensors, after pre-amplification, the charge measurement is performed with a 10-bit SAR ADC for charges up to 100 – 150 fC and by using the time-over-threshold (ToT) technique for charges above \(\sim 50\) fC. The ADC and ToT data are recombined, linearised and calibrated to provide a single charge measurement which is used to build and trigger sums, then transmitted for every BX. Both full event readout data and trigger sums are transmitted via 1.44 Gb/s electrical links (eLink) to the next level of the FE electronics.

The HGCROC is bonded to a hexagonal PCB (hexaboard) that contains the front-end ASICs and provides connections via wire-bonds through holes to the silicon. The

\(^{a}\)from \(\sim 0.2\) fC to 10 pC (\(\equiv\)16 bits)

\(^{b}\)better than 50 ps for pulses above \(\approx 12\) fC
hexa-board is connected to a mother-board, that aggregates, formats and serialises the event data (read at the L1 frequency up to 750 kHz) in the ASIC concentrator. This also selects the trigger sums of interest, aggregates and formats their data in packets storing them in a FIFO buffer. These data are then sent to the BE Trigger electronics located in the service cavern, using separate optical links (8.96 Gb/s LpGBT4). The event data stored in the HGCROC buffer are read after a L1-accept, with every channel above a certain threshold being sent to the DAQ system. The occupancy is strongly dependent on the position of the channel within the HGCAL and on the pileup rate: it varies from around 60% at high \(|\eta|\) in the front layers to less than 5% at the rear outer edge of the Si sensors, and lower than 1% in the rear scintillator areas. With the proposed format, the required bandwidth per chip does not exceed 0.8Gb/s for a L1 trigger rate of 750 kHz and can therefore be safely transmitted to the concentrator with a single eLink. The total average data volume for the whole HGCAL is around 2.5 MByte per event, which scales approximately linearly with the average pileup. At an L1 rate of 750 kHz, this corresponds to an average rate from the FE electronics of 15 Tb/s. A single HGCROC transmits to the concentrator 16 (8) trigger sums for a large (small) cells sensor through 4 (2) eLinks respectively at 40 MHz. In the concentrator, a threshold of 2 transverse MIPs\(^*\) (MIP\(_T\)) is applied.

\*MIP multiplied by the trigger cell sin (\(\theta\))

3.3. Performance

The fine granularity of HGCAL produces a large amount of information and enables enhanced pattern recognition to be exploited in the particle reconstruction. At the current state of the art an imaging algorithm\(^5\) is implemented to cluster the energy deposits from particle showers. The HGCAL also provides precise information on the timing of the energy deposits to discriminate clusters from pileup within a single BX. A cluster time resolution of \(O(20)\) ps, as shown in Fig. 2 (Center), can be achieved and corresponds to a precision of \(\sim 1\) cm at the interaction point. Given

\[ \text{Fig. 2. Left - Trigger cells configuration for 1.05 cm}^2 \text{ diodes. Center - Time resolution for photon clusters as function of } p_T \text{ momentum. Right - PF jets resolution as function of jet } p_T. \]
a bunch crossing dimension of $\sim 5$ cm, the neutral component of the PU can be reduced by a factor $\sim 5$. The five dimensional information (energy, x, y, z and time) is ideally suited for PF\(^5\)\(^7\) reconstruction. Hadronic and electromagnetic showers are individually reconstructed and identified, and charged hadrons matched to tracks reconstructed in the tracker. An important example of the HGCAL performance is the improved jet energy resolution achieved through PF (Fig. 2 Right). Jets are measured using the best available energy/momentum estimations, obtained mainly from the tracker for charged hadrons, and from the calorimeter for neutral hadrons and photons. The track-cluster matching also allows refinement of electron and converted photon reconstruction and the fine longitudinal and lateral granularity allows excellent identification performance for electrons and photons.

4. The HGCAL L1-Trigger

4.1. Back-end electronics and clustering algorithms

The CMS L1 trigger system will be subject to a major upgrade to increase the sustainable rate up to $O(750\text{kHz})$ from the current limit of $O(100\text{kHz})$\(^8\). In the HGCAL, the selected trigger cells and the HGCROC energy sums are the input data to the Trigger Primitive Generator (TPG)\(^8\). The system is designed to operate with a trigger cell selection threshold of at least 2 MIP\(_T\) at an average pileup of 200 and with the expected noise levels after 3000 fb\(^{-1}\). The input data of a given BX will arrive at the TPG up to 1.5 \(\mu\)s after the BX. The output data need to be received by the central L1T system within a latency of 5 \(\mu\)s after the relevant BX, so the TPG has a total available time of 3.5 \(\mu\)s to complete all processing and data transmission. The TPG has to implement two tasks: build three-dimensional (3D) clusters from trigger cells, and form an overall energy map in $\eta, \phi$ bins.

The algorithm currently proposed performs these tasks in two stages. The first stage considers each layer separately (due to financial constraints, only half of the CE-E section layers will be used). It builds two-dimensional (2D) clusters from trigger cells, relying on a dynamical clustering technique that gathers topologically-
connected trigger cells with an energy above a certain threshold. The number of 2D clusters per layer fulfill the bandwidth limits due to the allocated number of links as shown in Fig. 3 (Left). The second TPG stage then combines the 2D clusters in depth to form 3D clusters, gathering the clusters accordingly to their mutual distance. The currently assumed TPG architecture uses time multiplexing in the second stage to enable all the 2D clusters and energy maps for a BX to be transmitted to one location. A time multiplexing of 24 has been shown to be sufficient for this purpose. The total number of 3D clusters is heavily affected by the threshold applied, in particular \( O(250) \) clusters with transverse momentum \( (p_T) \) above 1 GeV are transmitted to the correlator. The completed energy maps and 3D clusters are transmitted from the TPG to the central L1T system.

4.2. Hardware implementation

The two endcaps will be processed independently using two identical copies of the hardware. In the first stage of the TPG, each board will receive up to 96 LpGBT links from on-detector for total of 48 boards per endcap. The outputs from these boards will be the 2D clusters and energy map for that layer. These will be transmitted to the boards in the second stage on 2304 16 Gb/s links per endcap. The second stage consists of 24 boards per endcap, each providing a processing node for a 24-fold time multiplexed system. Each board will then transmit the output data to the central L1T system over 288 16 Gb/s links per endcap. The VHDL implementation of the first trigger stage indicates that the proposed TPG architecture will fulfill the latency constraints.

5. Conclusions

In order to cope with the new challenges brought by the HL-LHC, the CMS Collaboration opted for a complete renovation of the calorimetry system in the forward region, pioneering the design and the R&D of a Si and scintillator/SiPM-based, segmented, high granularity calorimeter able to cope with a harsh environment. The high granularity and longitudinal segmentation allow full exploitation of particle shower properties for physics object identification and the fast timing information will be crucial to mitigate the pile-up contamination. With such granularity, the L1-trigger design represents a real challenge, pioneering the adoption of PF-based techniques at L1 and establishing a benchmark for future experiments.

References

1. CMS Collaboration, CERN-LHCC-2015-010 LHCC-P-008, (2015).
2. S. Jain (NCU), Intl. Journal of Modern Physics: Conference Series, (sub.).
3. N. Akchurin et al., Nucl. Instrum. Meth A859 3136, (2017).
4. CERN, https://espace.cern.ch/GBT-Project/LpGBT/Specifications/.
5. A. Rodriguez and A. Laio, Science 344, 1492-1496, (2014).
6. CMS Collaboration, JINST 12 no. 10, P10003, (2017).
7. M. Thomson, Nucl. Instrum. Meth A611 25, (2009).
8. CMS Collaboration, CERN-LHCC-2017-013 CMS-TDR-017, (2017).