PROCEEDINGS OF SCIENCE

Reconstruction in an imaging calorimeter for HL-LHC

Arabella Martelli*
Imperial College London
on behalf of the CMS Collaboration
E-mail: arabella.martelli@cern.ch

The CMS endcap calorimeter upgrade for the high-luminosity LHC (HL-LHC) uses, for the most part, silicon sensors to achieve radiation tolerance, with the further benefit of a very high readout granularity. Developing a reconstruction sequence that fully exploits the granularity, and other significant features of the detector like precision timing, is a challenging task. The aim is for operation in the high pileup environment of HL-LHC. An iterative clustering framework (TICL) is being developed. This takes as input clusters of energy deposited in individual calorimeter layers delivered by an "imaging" algorithm which has recently been revised and tuned to deliver excellent performance. Mindful of the projected extreme pressure on computing capacity in the HL-LHC era the algorithms are being designed with GPUs in mind. In addition, reconstruction based entirely on machine learning techniques is being developed and studied. This talk will describe the approaches being considered and show first results.

European Physical Society Conference on High Energy Physics - EPS-HEP2019 -
10-17 July, 2019
Ghent, Belgium

*Speaker.

PoS(EPS-HEP2019)110

1. Introduction

The high luminosity phase of the LHC (HL-LHC) will allow full exploitation of the potential of the collider and further exploitation of the rich physics program [1]: precision SM measurements and Higgs properties, as well as search for new physics.
This phase will start in about eight years (2026), and from the start will represent a large increase of instantaneous luminosity, by more than four times compared to Run3. This will pose challenges for the detectors, both in terms of high radiation and high pileup, with up to 200 interactions per bunch crossing. In such a harsh environment, radiation hardness and high granularity will be the fundamental requirements to allow operation at high luminosity, and precision timing, at the level of few tens of picoseconds, will be a key innovative tool to exploit [2].
In addition to the operational challenges, the needs from physics at the HL-LHC will imply the reconstruction of boosted topologies and VBF production mechanisms in the forward region. All this requires high granularity, fundamental to identify and reconstruct collimated objects, good coverage in the forward region, also to support the upgrade of the tracker extended to $|\eta|$ of about 4, and excellent reconstruction and identification of jets. Calorimetry in the forward region will play a fundamental role for physics, and to preapare for this, the CMS experiment [6] will upgrade its endcap calorimeters, replacing them with the High Granularity Calorimeter (HGCAL).

2. The High Granularity Calorimeter

The HGCAL detector [3] is designed with a sampling structure. It comprises an electromagnetic section (CE-E) of 28 layers, covering 25 X_0 and corresponding to 1.3 λ, and it is followed by a hadronic section (CE-H) of 22 layers, for a total depth of about 10 λ. A schematic view of the design is given in Fig. 1, on the left. The choice of the sensitive material depends on the expected levels of radiation, with silicon sensors used in the innermost layers and at high eta, while plastic scintillators are exploited in the rest. Two key elements of the HGCAL are the fine longitudinal readout segmentation, with each layer read out individually, and the high transverse granularity, with about 6 millions of silicon channels, of 0.5 or 1 cm^2 in size. A distinctive feature of the calorimeter is the hexagonal shape of the silicon sensors. This choice, made to minimize the costs from cutting the circular silicon wafers, results in increased difficulty to design the readout electronics and in the challenge to deal with an hexagonal geometry in the reconstruction. A picture of a prototype module is shown in Fig. 1, on the right.

3. Reconstruction at the HL-LHC

The high granularity of the HGCAL is ideal to help the pattern recognition and improve the separation of nearby showers. For illustration, a two-component event from 300 GeV electron beam reconstructed during the 2018 HGCAL beam tests is shown in Fig. 2.
At 200 pileup, where the average transverse energy deposit per unit area (in η, ϕ) amounts to 200 GeV, the ability to identify and accurately measure VBF jets in distinction to jets arising from pileup will have to exploit the combined information of longitudinal and transverse shower profiles. For example, as shown in Fig. 3, pileup jets cluster more energy in the first layers and are less contained radially.

PoS(EPS-HEP2019)110

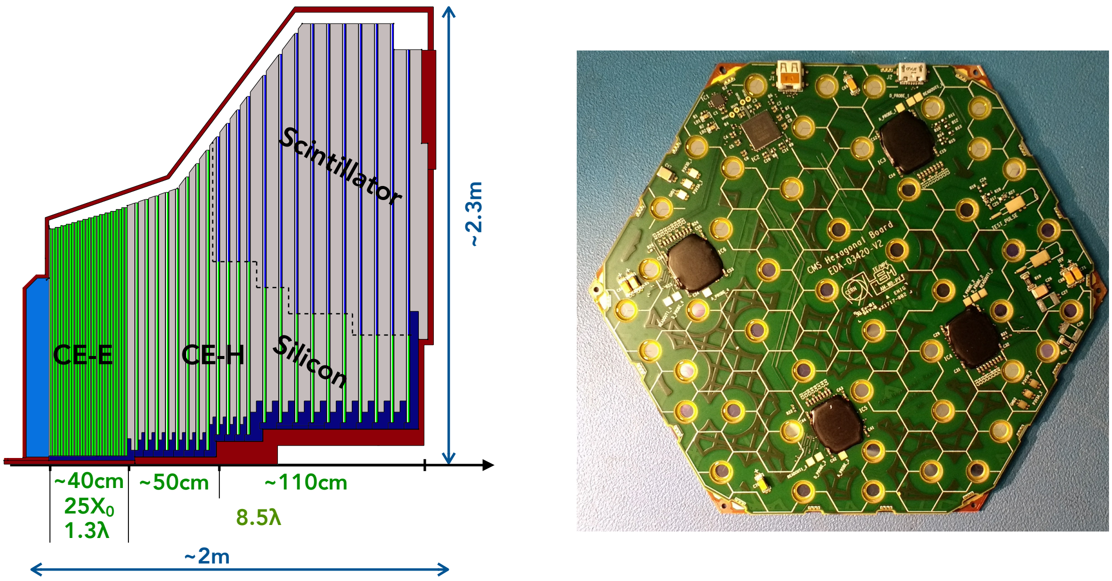

Figure 1: On the left: Schematic view of the High Granularity Calorimeter design. On the right: Six-inch module prior to wirebonding and encapsulation. The electronics packages seen on the PCB are four front end readout chips, used for the beam test, and an FPGA. The corners of the PCB and the sensor under the PCB are removed to provide direct access to the mounting holes on the baseplate

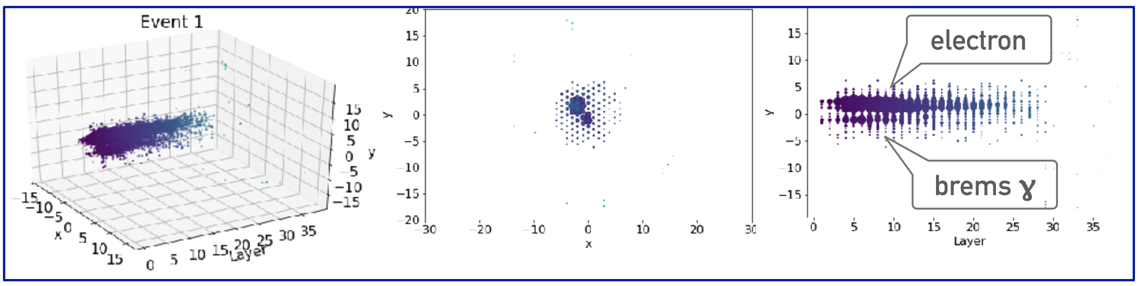

Figure 2: Two-component event from 300 GeV electron beam reconstructed during the 2018 HGCAL tests

The HGCAL also provides a time information for all silicon cells with energy above a given threshold (about 12 fC), and with a precision that depends on the amount of deposited energy. This intrinsic information leads to high timing precision on full showers, by exploiting the high multiplicity of hits: with a minimum of several tens for photon showers at 2 GeV in p_T and about 10 for hadrons. In ideal conditions, the time resolution achievable for full showers is about 20 ps for electromagnetic showers with $p_T > 2$ GeV and below 30 ps for hadron showers with $p_T > 5$ GeV. The use of timing information in a 5D (position, energy, time) reconstruction has great potential for pileup mitigation, as shown in Fig. 4: a simple selection based on the time compatibility strongly reduces the hits density and suggests a jet reconstruction and energy estimate less affected by pileup.

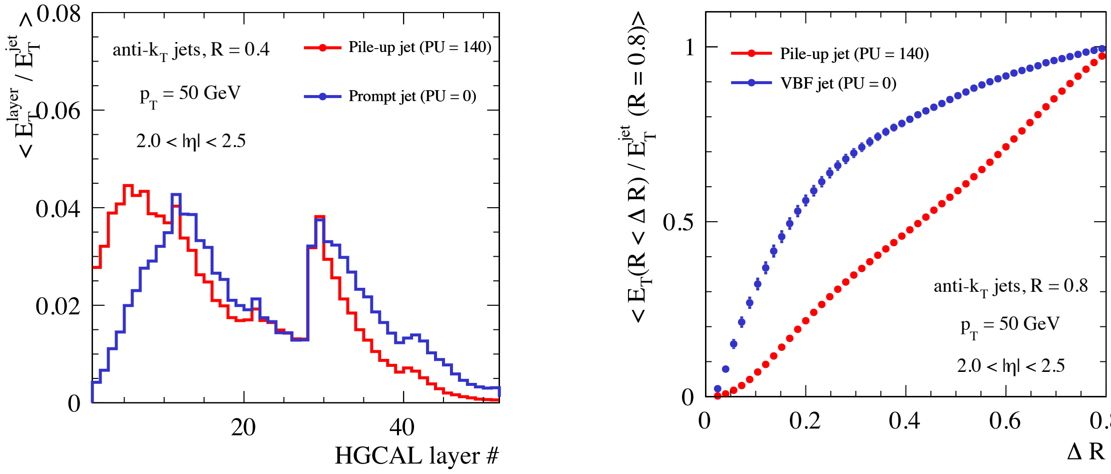

Figure 3: On the left: transverse energy deposited as a function of calorimeter layer, by quark jets (blue), and by jets reconstructed from pileup (red). On the right: transverse energy fraction contained as a function of distance, ΔR, for quark jets (blue) and jets made from pileup interactions (red). Plots from TDR [3].

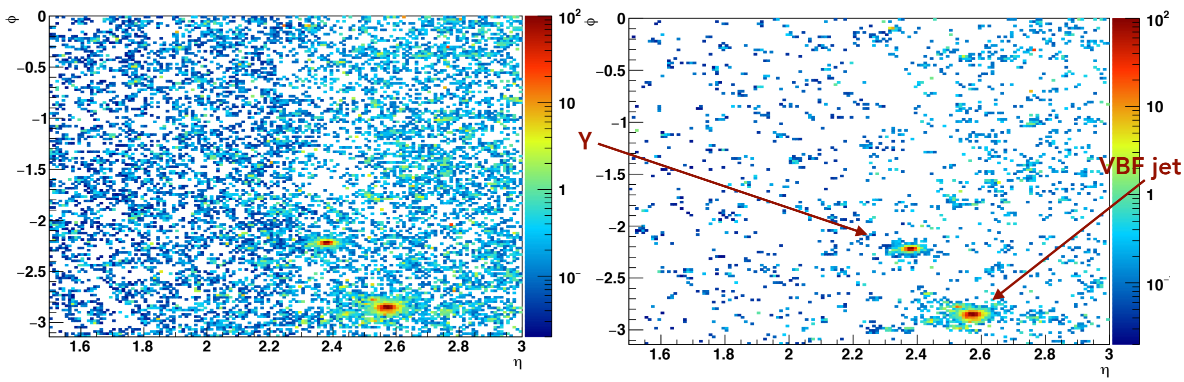

Figure 4: Hits with time in the HGCAL, projected to the front face of the calorimeter, in a event of VBF Higgs $\to \gamma\gamma$ in 200 pileup. On the left: without a timing requirement. On the right: after removal of hits with |Δ t|>90 ps. Plots from TDR [3].

3.1 HGCAL reconstruction for HL-LHC

The HGCAL detector offers the opportunity to design a reconstruction that fully profits from the high level of information available. The goal is to exploit the distinctive features of shower developments to target the reconstruction of several particles that interact differently in the detector: for example hadron showers show a peculiar lumpy structure, not present for electromagnetic showers. This will be achieved by developing a local reconstruction that is particle-flow [4] based. Particle-flow reconstruction is a well established technique, fully adopted by CMS: it combines the information from different subdetectors (tracker and calorimeters in primis) to provide the best identification and precise energy estimate of reconstructed particles. The HGCAL provides a 5D image of the shower with high granularity, and it is ideal for this.

In the first step of the local reconstruction layer clusters (2Dcl) are built, with an algorithm based on the local energy density of the reconstructed hits [5]. The result of this is shown on the left in Fig. 5, where two adjacent clusters are identified and reconstructed, each around an high energy density core. The following step is to connect compatible 2Dcl over layers, to build 3D showers: this is done in the iterative clustering framework (TICL), where each iteration produces a collection of 2Dcl aligned as a track (trackster). Each iteration is realized in consecutive steps. First, a seeding region is defined: this could be from track extrapolation if the target is a charged particle, or self-seeded for unconverted photons. The pattern recognition step follows: this aims at connecting the compatible 2Dcl in the identified region. A schematic example is given in Fig. 5, on the right. Currently it is designed to only use geometric compatibility criteria, with the extension to include energy and time planned for the future developments. The following step of linking and cleaning classifies the reconstructed particles, with an ID assigned in terms of probability, and where machine-learning tools are exploited. The final step allows masking of the hits of the objects successfully found in the present iteration, from the collection passed to the subsequent iteration, thus reducing combinatoric confusion in later iterations. All the algorithms are designed with parallelism in mind, and to run on GPUs.

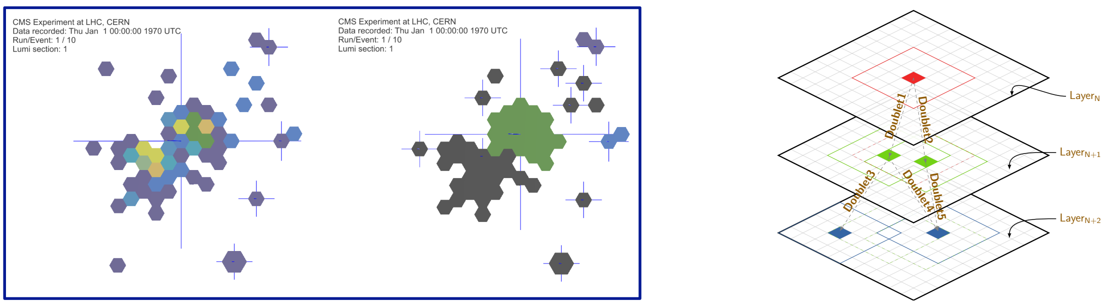

Figure 5: On the left: layer clustering at work. Reconstructed hits are shown with colours representing their energy. The resulting adjacent clusters identified are then shown with colours representing the 2Dcl. On the right: schematic example of the pattern recognition algorithm currently used in TICL, where doublets of compatible 2Dcl are built in consecutive layers.

As an example, Fig. 6 shows the two reconstructed electrons from a photon conversion. This result is obtained with the self-seeded iteration in absence of pileup and it clearly shows the clean separation of the two nearby showers and their energy development, achieved thanks to the fine granularity and segmentation of the HGCAL.

In Fig. 7, all the tracksters reconstructed by the self-seeded iteration for an event with a single pion ($p_T = 10$ GeV, $\eta = 1.7$) in 200 pileup are shown in blue-ish in one HGCAL endcap. In this context, the use of a track seeded iteration is useful to help the reconstruction of showers initiated by charged hadrons, by reducing the combinatorics. In this example, high quality tracks, with $p_T > 5$ GeV, are selected to seed the pattern recognition, and the resulting reconstructed tracksters are highlighted by the pink contour in the same figure: one corresponds to the charged pion from the hard process, the other is from a pileup hadron whose track satisfies the seeding criteria. With loose seeding criteria all particles can be tracked and linked to reconstructed showers in the calorimeter: an

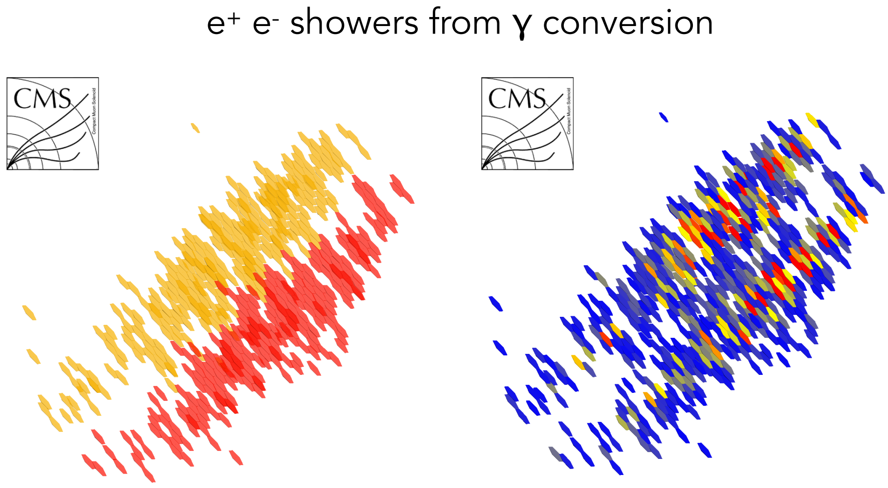

Figure 6: e^+e^- showers from γ conversion. On the left: the color represents the particle, to evidence the separation of the nearby showers. On the right: the color refers to the energy of the 2Dcl, to highlight the high energy cores in the depths of the showers.

optimal working point has to be found to balance reconstruction efficiency with the disadvantages of high combinatorics.

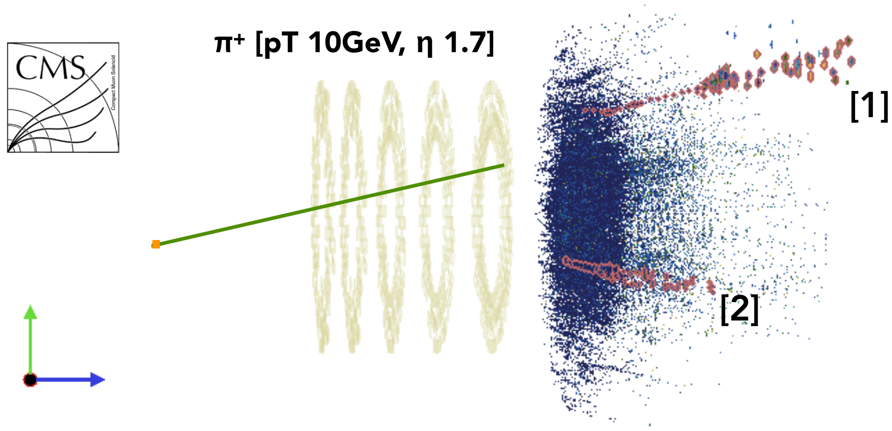

Figure 7: Event display of the reconstructed clusters in the HGCAL, for an event with a single pion in 200 pileup. Only one half of the detector is shown. The reconstructed track associated to the pion from the hard process is shown in green. All the tracksters reconstructed with the unseeded iteration are shown in blue. The two tracksters reconstructed with a track seeded iteration are highlighted by a pink contour.

The full reconstruction will benefit from machine learning techniques to improve particle identification and the shower shape estimate of the reconstructed tracksters. For example, this will allow masking of the identified electromagnetic tracksters, obtained with an unseeded iteration, and reduce the ingredients used by subsequent iterations. A first full reconstruction sequence is under development, aiming for a first full prototype to reconstruct the objects found in the calorimeter in high pileup conditions.

4. Conclusion

The High Granularity Calorimeter for the HL-LHC is a very ambitious project. Reconstruction in the calorimeter involves using the fine granularity and fine longitudinal readout segmentation, with more than 6 millions of channels, and designing the best algorithms to exploit energy, position and time information, to provide a 5D image of the shower development.
Developing such a reconstruction that fully exploits the features of the detector, is a challenging and a creative task. Development is ongoing, and the first full-reconstruction sequence is expected to converge soon, but certainly further developments are to be expected from now until the start of the HL-LHC.

References

[1] 16th Session of European Strategy Council (2013), *The European Strategy for Particle Physics Update 2013*, [https://cds.cern.ch/record/1567258].

[2] J. Butler, D. Contardo, M. Klute, J. Mans, and L. Silvestris eds. [The CMS Collaboration], *Technical Proposal for the Phase-II Upgrade of the CMS Detector*, *CERN-LHCC-2015-010* LHCC-P-008 (June 2015).

[3] *The Phase-2 Upgrade of the CMS Endcap Calorimeter*, CERN-LHCC-2017-023, CMS-TDR-019, [https://cds.cern.ch/record/2293646].

[4] A. M. Sirunyan et al. [CMS Collaboration], *Particle-flow reconstruction and global event description with the CMS detector*, [arXiv:1706.04965].

[5] A. Rodriguez, A. Laio, *Clustering by fast search and find of density peaks*, *Science 344 (6191) 1492-1496* (June 2014), [doi: 10.1126/science.1242072].

[6] CMS Collaboration, *CMS detector paper*, *JINST 3 S08004* (2008).

PoS(EPS-HEP2019)110