Investigation of Fast Timing Capabilities of Silicon Sensors for the CMS High Granularity Calorimeter at HL-LHC

Artur Apresyan for the CMS Collaboration

Abstract

The High Granularity Calorimeter (HGCAL) is the technology choice of the CMS collaboration for the endcap calorimetry upgrade planned to cope with the harsh radiation and unprecedented in-time event pileup projected at the High Luminosity-LHC era. In this context, profiting from fast-timing information (tens of picoseconds) embedded in the calorimeter would represent a unique capability for resolving information from individual collisions at the HL-LHC. This will enhance the reconstruction and physics capabilities of the CMS detector in terms of pileup mitigation and particle identification. The HGCAL is realized as a sampling calorimeter, including 40 layers of silicon pad detectors with pad areas of 0.5 cm² and three active thicknesses of 320, 200, and 120 μm. Prototype p–in–n and n–in–p5x5mm² silicon pads, with thicknesses of 285, 211, and 133 μm, were tested with high-energy electrons at the CERN SPS. Timing in the HGCAL, and the measured intrinsic timing capabilities for electromagnetic showers and minimum-ionizing particles.

Presented at IEEE-NSS-MIC-2016 IEEE Nuclear Science Symposium and Medical Imaging Conference
Investigation of Fast Timing Capabilities of Silicon Sensors for the CMS High Granularity Calorimeter at HL-LHC

Artur Apresyan, Fermi National Accelerator Laboratory (FNAL) on behalf of the CMS collaboration

Abstract—The High Granularity Calorimeter (HGCAL) is the technology choice of the CMS collaboration for the endcap calorimetry upgrade planned to cope with the harsh radiation and unprecedented in-time event pileup projected at the High Luminosity-LHC era. In this context, profiting from fast-timing information with a resolution of a few tens of picoseconds embedded in the calorimeter would represent a unique capability for resolving information from individual collisions at the HL-LHC. This will enhance the reconstruction and physics capabilities of the CMS detector in terms of pileup mitigation and particle identification. The HGCAL is realized as a sampling calorimeter that uses silicon sensors as its active elements. Prototype sensors were tested with high-energy electron beams at the CERN SPS, and at the FNAL FTBF test beam facilities. We present the motivation for this studies including the concept and use of fast-timing in the HGCAL, and the measured intrinsic timing capabilities of silicon sensors for electromagnetic showers and minimum-ionizing particles.

I. INTRODUCTION

In order to collect large datasets needed for precise characterization of the Higgs boson, and to increase new physics discovery potential, the High Luminosity LHC (HL-LHC) upgrades are expected to deliver peak luminosities in excess of $5 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$. With the increased instantaneous luminosity, the simultaneous interactions per bunch crossing (pileup) increase the likelihood of confusion in the reconstruction of particles from the hard scatter interaction with those produced in different pileup interactions. The ability to discriminate jets, photons, and electrons produced in the events of interest from pileup can become significantly degraded, and additional tools are needed in order to mitigate these effects. In this study we investigate the capabilities of the HGCAL detector in measuring the electromagnetic showers with high temporal resolution [1]. Our goal is to perform a time of arrival measurement by using the silicon layers within HGCAL, providing a time assignment for charged particles and photons. Achieving such a time measurement with resolution better than 20–30 psec can reduce the impact of pileup by a factor of 5–7 effectively restoring the current LHC conditions, and will recover the event reconstruction.

II. EXPERIMENTAL CHALLENGES OF HIGH PILE-UP IN HL-LHC

A visualization of the expected spatial density of charged particles in the high pileup environment at the HL-LHC is displayed in Figure 1. An average of around 140 overlapping interactions in every bunch crossing (pileup or PU) will occur within a root mean square (RMS) spread of approximately 5 cm along the beam axis at each bunch crossing. The vast majority of these interactions are “soft”, and are not of interest to probe signatures of BSM physics.

Fig. 1. High pileup conditions at the HL-LHC represented by the $z$-vertex distribution of a single bunch-crossing with approximately 200 vertices. The region presented corresponds to the coverage of the CMS pixel detector.

Studies with the current CMS detector show that each additional PU vertex adds about 3 GeV (in quadrature) to the missing transverse momentum ($E_T$) resolution. Compared to Run 1 performance, that would amount to a factor of 2–3 worse $E_T$ resolution during HL-LHC runs. For reconstruction of the signal peak in the crucial $H \rightarrow \gamma\gamma$ final state, the diphoton vertex selection efficiency drops from $\sim 80\%$ in current conditions to $\sim 30\%$ at 200 PU. This reduction in primary vertex selection efficiency causes about 30% degradation in statistical precision in $H \rightarrow \gamma\gamma$ measurements. Additionally, as vertices start to overlap within effective tracking resolution, the rate of pileup tracks associated to the hard interaction vertex increases. This results in a degradation of the isolation calculation for leptons and photons, and impacts jets and $E_T$ performance [2]. It is estimated that for a 50 GeV jet (electron) the additional contribution from neutral mesons to the energy measurement will be $\sim 100\%$ (20%). These effects limit the effective luminosity production of colliders, despite their best achievable luminosity, and represent a bottleneck to the rate at which high quality data can be accumulated by the

Artur Apresyan is with the Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL, 60510-5011, USA (telephone: 630-840-2697, e-mail: apresyan@fnal.gov)
experiments.

Precision time-stamping of particles originating from various vertices would add a new handle in discriminating particles emanating from the hard scatter and PU. In the time domain, collisions at the HL-LHC will occur with an RMS time spread of approximately 150 psec within each bunch crossing of the colliding beams. Detectors capable of time of arrival measurements at the order of 20-30 psec will provide the precision needed to reduce the PU contamination at the HL-LHC, bringing the performance of the reconstruction of jets, photons and $E_T$ to the LHC level. For 20–30 psec time resolution, simulations show that the losses of the Higgs boson vertex identification efficiency at the HL-LHC are partially recovered by the calorimeter timing alone [3], and almost fully recovered by the calorimeter combined with MIP timing [4]. Precision timing measurements introduce an opportunity to dramatically increase the instantaneous luminosity of future colliders by a factor of 5–10, while maintaining excellent detector performance.

The high spatial granularity of the HGCAL, with the intrinsic fast response of silicon detectors could serve as a critical tool to manage the high pileup levels expected in the CMS detector during HL-LHC. The focus of the R&D presented in this study is to understand the precision achievable with silicon based sampling calorimeters. Specifically, we study a) silicon sensor pulse shape when exposed to EM showers, and evaluate their timing capabilities, and b) details of pulse shape before/after irradiation to evaluate the impact of irradiation on timing and general calorimeter performance. Measurements performed at test-beam facilities at FNAL and CERN beamlines are presented in the following. We first present studies of prototypes with a single silicon diode in Sec. III, and measurements with a multi-channel prototype constructed of a single layer of an HGCAL silicon sensor are presented in Sec. IV.

III. SILICON DIODES IN ELECTROMAGNETIC SHOWERS

In order to characterize the timing performance of silicon sensors within electromagnetic showers, test beam experiments were performed at FNAL and at CERN [5], [6]. For our measurements, we used a set of p-on-n silicon diodes produced by Hamamatsu. The diodes used had 133, 211 and 285 $\mu$m depletion thickness on 300 $\mu$m substrate. The transverse sizes of the sensors were 5 × 5 and 6 × 6 mm$^2$. The negative bias voltage was applied to the p-side of the silicon, and the silicon diodes were placed inside a light-tight box of thickness 1.5 cm, which also provided electromagnetic shielding. A nominal bias voltage of 500 V was applied to deplete the silicon sensor in the studies shown for FNAL measurements, and 600 V for those at CERN. Silicon diode assemblies used in the measurements are shown in Fig. 2.

The readout was performed by connecting the output of the diodes to fast amplifiers, which were subsequently connected to a CAEN V1742 digitizer board [7], which provides digitized waveforms sampled at 5 GS/s, and one ADC count corresponds to 0.25 mV. During measurements at FNAL the signals from the silicon diode were amplified by two fast, high-bandwidth pre-amplifiers connected in series. The first amplifier was an ORTEC VT120C pre-amplifier, and the second amplifier was a Hamamatsu C5594 amplifier. During the measurements at CERN a single Cividec C2 broadband amplifier was used [8].

A simple schematic diagram of the experimental setup is shown in Figure 3. A small plastic scintillator was used as a trigger counter to initiate the read out of the data acquisition (DAQ) system and to select incident beam particles from a small geometric area, allowing us to center the beam particles on the silicon diode. For the FNAL measurements, we placed a stack of tungsten absorbers of various thicknesses for measurements of the longitudinal profile of the electromagnetic shower. A fast micro-channel plate photomultiplier detector is placed furthest downstream, and serves to provide a very precise reference timestamp. The time for the reference MCP-PMT detector is obtained by fitting the peak region of the pulse to a Gaussian function and the mean parameter of the Gaussian is assigned as the timestamp $t_0$. The time for signals from the silicon diodes was obtained by performing a linear fit to the rising edge of the pulse and the time at which the pulse reaches 30% of the maximum amplitude was assigned as its timestamp $t_1$. During measurements at CERN more than one silicon diode was placed after the absorber [6]. We measured the electronic time resolution of the CAEN V1742 digitizer as ~4 ps and neglected its impact on the timing measurements described below. The distribution of the difference between time stamps of the MCP-PMT and silicon diode ($\Delta t = t_1 - t_0$) is fit with a Gaussian, and the standard deviation of the resulting fit function is called the “time resolution” in the following.
For measurements at CERN the time resolution of individual silicon diodes is extracted as $\sigma(t_1 - t_2)/\sqrt{2}$.

Fig. 3. A schematic diagram of the test-beam setup is shown. The $t_0$ and $t_1$, $t_2$ are defined in Section III.

The response of silicon diodes to minimum ionizing particles (MIPs), and to secondary particles produced in electromagnetic showers was simulated using GEANT4 toolkit [9]. Such multi-MIP measurements provided valuable calibration information at the low energy range of the silicon diodes signals, and could be compared to data. The results of simulation of the experimental setup, and its comparison with data are shown in Fig. 4. The setup in both cases is the same as shown in Fig. 3, where the incoming beam of 50 GeV electrons is impinging on a $2X_0$ absorber, and showers are detected by a 285 $\mu$m depletion depth silicon diode. Simulation reproduces well the observed response to EM showers, both qualitatively and quantitatively, as can be seen in Fig. 4.

Fig. 4. Multi-MIP response to EM showers as measured with silicon diode. Left plot shows the distribution measured in data, and the right one is that from simulation.

The dependence of the measured time resolution on the beam energy is shown in Fig. 5. We observe an improvement in the time resolution as the beam energy increases, reaching a time resolution of about 20 ps for 32 GeV electrons. This is due to improvement in the signal-to-noise (S/N) at higher energies. As can be seen from the bottom panel on Fig. 5, the time resolution of silicon sensors with different depletion depths are the same when compared for the same S/N. Additionally, time resolution measurements are compared with simulation as shown in Fig. 5, and are found to be in good agreement.

Our results show that the time stamp associated with electromagnetic showers induced by electrons with energy above 30 GeV can be measured with a precision better than 25 ps. According to our calibration measurements, and prediction from the simulation, the MIP multiplicity per single diode required to achieve such precision corresponds to signals above 20 MIPs. These measurements demonstrate that a calorimeter based on silicon sensors as the active medium can achieve intrinsic time resolution at the 20 ps level, as long as noise is kept under control, and the electronics architecture is appropriate.

The HGCAL detector will be exposed to severe radiation by the end of the HL-LHC period, reaching $10^{16}$ n/cm$^2$ 1 MeV-equivalent neutron flux in its most forward region. Therefore, it is important to characterize the performance of the timing measurements also towards the end of experiment’s lifetime. We repeated the measurements reported above for devices that were subjected to neutron irradiation. The results are shown in Fig. 6. We observe that the irradiation results in a reduction of the pulse amplitude and of the pulse rise-time. However, the time resolution remains compatible between the unirradiated and irradiated devices.
Fig. 6. The measured time resolution of the irradiated silicon diodes as a function of the S/N ratio. Devices irradiated with different fluences are shown, including an unirradiated sample.

IV. HGCAL TIMING LAYER

As a next step in the program of characterizing precision timing measurements with the HGCAL detector we equipped a single layer of an HGCAL detector prototype with fast readout electronics. We used one of the 300 µm thickness hexagonal sensors used for the construction of the HGCAL test-beam prototype [1], and equipped 25 (from 135) hexagonal cells with their own amplifier chain, for a total gain of 100 per channel. The readout of the timing layer was performed using the V1742 module described in Sec. III. Fig. 7 shows the sensor layout used for these measurements, and the PCB used for fast readout. The schematic circuit diagrams of the PCB electronics are shown in Fig. 8. Each HGCAL pad was equipped with an amplifier placed very close to the wirebonded readout, and a 1:2 transformer to lower the input impedance and to provide a faster rise time of the signals. An additional amplifier was placed at the periphery of the PCB, as shown in Fig. 8. The noise contributions from the electronics chain were measured to be around 8 fC.

Measurements were performed in 2016 using the FNAL Beam Test Facility (FTBF), using the primary 120 GeV proton beam, and secondary electron beams with energies ranging from 4 GeV to 32 GeV. Electron purity for those beams ranged between 70% at the lowest energy to about 10% at the highest energy. Stacks of tungsten or lead plates with varying thicknesses were placed immediately upstream of the timing layer in order to measure the response along the longitudinal direction of the electromagnetic shower. A small plastic scintillator of transverse dimensions 1.8 mm × 2 mm was used as a trigger counter to initiate the read out of the data acquisition (DAQ) system and to select incident beam particles from a small geometric area, allowing us to center the beam particles on the silicon sensor. Finally, a Photek 240 micro-channel plate photomultiplier detector was placed furthest downstream to provide a very precise reference timestamp. The time resolution between a pair of Photek 240 detectors was previously measured to be 9.6 ps [10], hence the time resolution of a single Photek 240 is $9.6/\sqrt{2} = 6.8$ ps. For the measurements presented in this report the time resolution of the reference Photek 240 MCP-PMT is not subtracted. The arrangement of various detector components is similar to that presented in Fig. 3, where the silicon diodes are replaced with the timing layer.

![Fig. 7. (Left) The area outlined in blue shows the part of the HGCAL silicon sensor that was read out in the timing measurements. (Right) The dedicated PCB designed for fast readout of the HGCAL sensor, which is wire-bonded to the PCB on its back.](image)

![Fig. 8. (Top) Schematic circuit diagram of the readout electronics for one channel of the timing layer. (Bottom) Schematic circuit diagram of the periphery readout of the timing layer.](image)

We present here the response of the silicon sensor to electron beams of various energies after 6 radiation lengths ($X_0$) of tungsten absorber. The silicon sensor is expected to be sensitive to the number of secondary electrons produced within the electromagnetic shower, and therefore its response is expected to scale up with higher incident electron energy. In Fig. 9, we show examples of the integrated charge distribution measured in the HGCAL sensor after 6 $X_0$ of tungsten, for runs with 16 and 32 GeV electrons. The distributions show the charge collected by each pad, which is calculated as the...
integral pulse of the measured signal and then converted to charge. In these runs the tungsten absorber was placed 1 mm from the timing layer. We observe a fairly linear dependence between the measured charge and the incident beam energy, for beam energies between 4 GeV and 32 GeV, confirming calorimetric behavior of the detector. The spread of the shower around the central, most energetic pad is consistent with the 0.9 cm Molière radius of tungsten.

![Charge vs. ADC counts](image)

**Fig. 10.** A sample pulse showing the linear fit on the rising edge to extract time stamp (red line). The intersect of the horizontal blue line (at 45% of the amplitude) with the red line is assigned as the time stamp as shown by the blue arrow.

**Fig. 11.** Sample pulses from a single 32 GeV electron shower event. We measured the noise level to be about 8 fC, while the signal from MIP is expected to be about 4 fC. Therefore, in this setup, the detector was not sensitive to single MIP deposits.

A linear fit to the pulses is performed from 20% to 90% of the signal amplitude. The timestamp is defined as time at which the signal reaches 45% amplitude on this fit. An example of the pulse from the timing layer detector was measured to be about 2 ns (defined as 10-90% of the amplitude), and is dominated by the electronics components.

We then measure the time resolution between the HGCAL timing layer and the Photek MCP-PMT. Events are selected to pass identification criteria consistent with showering electrons by requiring a large deposit in the Photek MCP-PMT. This selection has high efficiency in discriminating shower events, since those produce many secondary particles that are then detected by the MCP-PMT that is positioned behind the absorber, as shown in Fig. 3.

Multiple pads are combined to determine the time stamp of the HGCAL timing layer. Each pad used in the combination
must pass both charge and amplitude cuts (charge $> 10$ fC and amplitude $> 0.01$ V). Time information from each pad is added with weighting based on the magnitude of charge the pad contains. If a pad does not pass these cuts, it is added with a weight of 0, such that the event can still be used even if not every pad passes the cuts. We find that in most events 3 or 4 pads pass these requirements, and it is rare for all 7 pads to pass the cuts within the same event. An example time resolution distribution is shown in Fig. 12 for a 32 GeV electrons, where we measure about 16 psec resolution.

![Fig. 12. Time resolution for the 32 GeV electron data, with 6X0 tungsten absorber placed 1 mm from the HGCal timing layer. All 7 pads are combined with charge weighting. Time resolution of Photek MCP-PMT is not subtracted from the measurements.](image)

We varied the electron beam energy, to study the dependence of the time resolution on the signal magnitude. We collected data for electron beam energies at 8, 16 and 32 GeV. Time resolution was measured with the same procedure as that described above. The dependence of the time resolution on the beam energy is shown in Fig. 13. As can be seen, a time resolution of about 30 psec is achievable even with a single timing layer for electrons at 8 GeV.

We then placed a Photonis XP85011 MCP-PMT behind the timing layer as shown in Fig. 14. This setup allows us to probe the longitudinal development of the shower, and test whether additional sampling layers will improve the precision of the timing measurements. While the Photonis MCP-PMT provides a pixelated readout, in our measurements we used the 16 central pixels that were connected together to a single readout channel, as described in Ref. [11]. The measurements of the timing resolution are performed similarly to those described above. The time resolution is determined between Photek 240 MCP-PMT and the HGCal timing layer ($\Delta t_1 = t_0 - t_1$), and between Photek 240 and Photonis XP85011 MCP-PMTs ($\Delta t_2 = t_0 - t_2$), where $t_0$ is the time stamp of the Photek 240, $t_1$ is that of the HGCal timing layer, and $t_2$ corresponds to the time stamp of the Photonis XP85011 MCP-PMT. The time stamp of the HGCal timing layer is the one after the central 7 pads have been combined with charge weighting, as described above. The time resolution measurement is then performed by combining $t_1$ and $t_2$.

![Fig. 14. Experimental setup used in the study of the timing measurements with two timing detectors placed inside the EM shower: HGCal timing layer and the Photonis XP85011 MCP-PMT.](image)

Time resolution measurements for $t_1$ and $t_2$ are shown in Fig. 15, where we measure $\sigma(t_1) = 15.5 \pm 0.5$ psec, and $\sigma(t_2) = 13.8 \pm 0.4$ psec. In these measurements, as before, the Photek 240 time resolution is not subtracted. The Photonis MCP-PMT and the HGCal timing layer provide very similar precision in the time measurements. Therefore, while in the HGCal detector for the upgraded CMS experiment most sensitive elements would be composed of silicon layers, for the purposes of this study the Photonis MCP-PMT can be used as a proxy device to measure the combined precision of a multi-layer detector.

After establishing the time resolution of each individual layer as presented in Fig. 15 we proceed to the combination of these measurements. Since in the current setup each layer was not sensitive to single MIP measurements, we couldn’t perform a calibration of the two layers to establish an absolute scale to MIPs. Therefore, a charge weighted combination of the two layers does not provide an optimal use of the signal-to-noise information. We find that a simple combination with both layers given an equal weight achieves a good performance, as shown in Fig. 16, where for a 32 GeV electron beam we measure time resolution of around 12 psec.

The same experiment was repeated while varying the energy of the electron beam, and the results are shown in Fig. 17.
We observe a consistent improvement of the time resolution when compared to the precision achieved with both single-layer measurements. This result confirms that additional information from the second layer significantly improves the time resolution.

V. CONCLUSION

Calorimeters with high granularity, and high precision timing capabilities can serve as a new tool in reducing the effects of large pileup in hadron colliders. In this report we studied the precision timing capabilities of the detectors for the CMS endcap calorimeter upgrade. We find that the timing precision of the silicon based calorimeters scales with the signal-to-noise ratio. We performed studies of the timing performance of silicon sensors subjected to neutron irradiation fluences expected during the HL-LHC, and find that the time resolution does not degrade compared to the performance of the unirradiated sensors. We show that a time resolution of around 15 ps can be achieved with a 32 GeV electron beam when a single silicon timing layer is placed after 6 $X_0$ of tungsten absorber. We also showed that the timing precision improves from multiple longitudinal samplings within the tungsten absorber. We also showed that the timing precision of Photek MCP-PMT is not subtracted from the measurements.

Fig. 15. Time resolution for the 32 GeV electron data, with 6$X_0$ tungsten absorber at 1 mm from the HGCal timing layer for the HGCal timing layer (left) and for Photonis MCP-PMT (right). All 7 pads of the HGCal timing layer are combined with charge weighting. Photonis MCP-PMT is placed after HGCal timing layer. Time resolution of Photek MCP-PMT is not subtracted from the measurements.

We also showed that the timing precision improves from multiple longitudinal samplings within the tungsten absorber. We also showed that the timing precision of Photek MCP-PMT is not subtracted from the measurements.

We thank Anatoly Ronzhin of FNAL, and Gillian Kopp and Daniel Gawerc of Caltech for their critical contributions to the work on HGCal timing layer and the silicon diode tests at FNAL. We thank the FTBF personnel for very good beam conditions during our test beam time. We also appreciate the technical support of the FNAL SiDet department for the production of high quality silicon samples, and the construction of the HGCal timing layer. We greatly appreciate the work of the CERN SPS team to provide excellent quality beam energy measurements of CALOR 2016. https://indico.cern.ch/event/472938/

ACKNOWLEDGMENT

We thank Anatoly Ronzhin of FNAL, and Gillian Kopp and Daniel Gawerc of Caltech for their critical contributions to the work on HGCal timing layer and the silicon diode tests at FNAL. We thank the FTBF personnel for very good beam conditions during our test beam time. We also appreciate the technical support of the FNAL SiDet department for the production of high quality silicon samples, and the construction of the HGCal timing layer. We greatly appreciate the work of the CERN SPS team to provide excellent quality beams with high efficiency and especially thank them for their flexibility and help throughout the running periods. We also thank our international collaborators and funding agencies for the support given to this project.

REFERENCES

