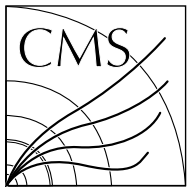


CERN-LHCC-2017-023
CMS-TDR-019
9 April 2018

The Phase-2 Upgrade of the CMS endcap calorimeter

Technical Design Report

CMS Collaboration

CERN-LHCC-2017-023 / CMS-TDR-019
08/02/2019

Editors

C. Seez, T. Virdee

Chapter Editors and Writers

P. Azzi, D. Barney, T. Bergauer, P. Bloch, P. Dauncey, A. David, M. Mannelli, J. Mans, A. Marchioro, A. Martelli, S. Moccia, M. Narain, C. Seez, F. Sefkow, J. Strait, C. de la Taille, T. Virdee

Cover Design

Sergio Cittolin

Acknowledgements

We gratefully acknowledge the dedicated efforts of many colleagues in developing the HGCAL project, and who have contributed to this Technical Design Report. We are grateful for the feedback from all the readers of this TDR, especially that from the CMS Institutional Readers, the CMS HGCAL internal reviewers, the TDR Editorial Board chairperson. We gratefully acknowledge the support from the CMS management team as well as from the offline, computing and physics groups, and the guidance and support from the CMS Upgrades Coordinators and their team.

We thank the technical staff from the participating institutes for their insightful work during the development and design of the HGCAL project presented in this TDR.

Contents

Part I

HGCAL Technical Design Report

Chapter 1

Introduction and overview

1.1 Introduction

The Large Hadron Collider (LHC) provided first proton-proton collisions at $\sqrt{s} = 900\,\text{GeV}$ in 2009. During Run 1 (2010–2012) the accelerator operated for the first two years at $\sqrt{s} = 7\,\text{TeV}$, delivering $\approx 6\,\text{fb}^{-1}$, and at $\sqrt{s} = 8\,\text{TeV}$ in 2012, delivering $\approx 23\,\text{fb}^{-1}$. The most significant physics result from Run 1 was the discovery of the Higgs boson.

Run 2 started in 2015 at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ and the instantaneous luminosity, reaching as high as $1.7 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$, has exceeded the design value. Detailed studies are being carried out of the Higgs boson and standard model (SM) processes as well as searches for physics beyond the SM.

It is intended to accumulate some $300\,\text{fb}^{-1}$ by the end of Run 3 (2023). After the third long shutdown (LS3) the HL-LHC operational phase is scheduled to commence in the last quarter of 2026. It is planned to 'level' the instantaneous luminosity at $5 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ with the goal of integrating some $3000\,\text{fb}^{-1}$ by the mid-2030s. The corresponding mean number of collisions (pileup) per bunch crossing will be 140. However, the LHC has the ability to deliver 50% higher values for both the instantaneous and integrated luminosities. Hence for our radiation and event pileup studies we use nominal values as well as 200 for the number of interactions per bunch crossing.

In addition to the detailed studies of the Higgs boson and SM processes, the searches for physics beyond the SM in the HL-LHC physics programme will also include reactions initiated by vector boson fusion (VBF) and those involving boosted objects giving rise to narrow (e.g. from taus) or merged jets (e.g. from hadronic decays of the W and Z bosons). A goal of the HL-LHC running, in an environment of high event pileup, is to trigger cleanly on and reconstruct the narrow VBF jets, as well as merged jets, without placing significant requirements on the rest of the event content. This should further open the domain of physics initiated by WW, WZ, or ZZ fusion.

The HL-LHC will integrate ten times more luminosity than the LHC, posing significant challenges for radiation tolerance and event pileup on detectors, especially for calorimetry in the forward region. As part of its HL-LHC upgrade programme, the CMS Collaboration is proposing to build a high granularity calorimeter (HGCAL) to replace the existing endcap calorimeters.

The existing forward calorimeters, the $PbWO_4$-based electromagnetic calorimeter (EE) and the plastic scintillator based hadron calorimeter (HE), were designed for an integrated luminosity of $500\,\text{fb}^{-1}$. The performance degradation much beyond this integrated luminosity leads to an unacceptable loss of physics performance [1].

Any replacement calorimeter must have the ability to withstand integrated radiation levels that are ten times higher than anticipated in the original CMS design. Simulations using FLUKA (Figs. 1.1 and 1.2) indicate that the highest fluence is around $10^{16}\,n_{eq}/cm^2$ and the highest dose around 2 MGy. Such radiation levels will be encountered at the inner radii of the silicon trackers at the HL-LHC. The R&D carried out, by several groups, for the upgrade of the silicon tracker has demonstrated that silicon sensors could indeed tolerate such levels. The silicon sensors retain adequate charge collection even after having been submitted to fluences up to $1.5\times10^{16}\,n_{eq}/cm^2$ (where n_{eq}/cm^2 denotes the number of 1 MeV equivalent neutrons per square cm), a fluence that is 50% higher than expected for an integrated luminosity of $3000\,fb^{-1}$. Hence silicon sensors were chosen for the active material for the bulk of the upgrade of the endcap calorimeters. In order to reliably operate silicon sensors after irradiation, and to keep sufficiently low the energy equivalent of electronics noise that results from the increased leakage current and decreased charge collection efficiency after irradiation, the sensors have to be operated at around $-30\,°C$.

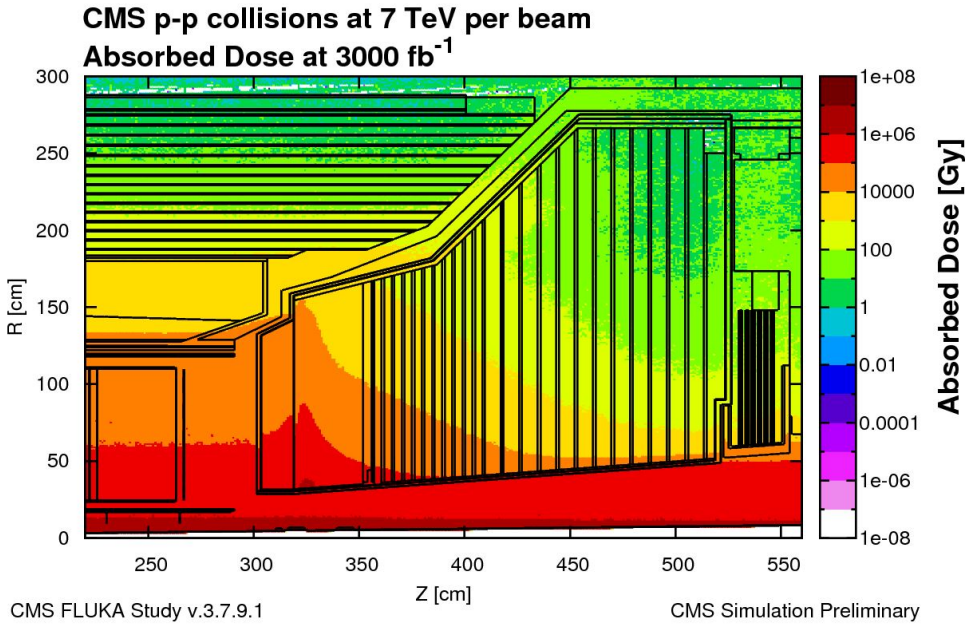


Figure 1.1: Dose of ionizing radiation accumulated in HGCAL after an integrated luminosity of$3000\,fb^{-1}$, simulated using the FLUKA program, and shown as a two-dimensional map in the radial and longitudinal coordinates, r and z.

The proposed design uses silicon sensors as active material in the front sections and plastic scintillator tiles, with the scintillation light read out by SiPMs, towards the rear. In the region covered by plastic scintillators the maximum radiation levels correspond to a fluence of $8\times10^{13}\,n_{eq}/cm^2$ and a dose of 3 kGy. In order to keep the radiation-induced energy equivalent of electronics noise sufficiently low, SiPMs also have to be operated at around $-30\,°C$. Hence the whole calorimeter will be operated at $-30\,°C$.

The chosen techniques rely on recent advantageous advances in cost per unit area and radiation tolerance of silicon sensors, advances in radiation-tolerant fast electronics, high-bandwidth data transmission via optical fibres, and in FPGA technology for the first level of event selection. The challenges lie mainly in the area of engineering (electronics, mechanical, and thermal), data transmission, and level-1 (L1) trigger formation.

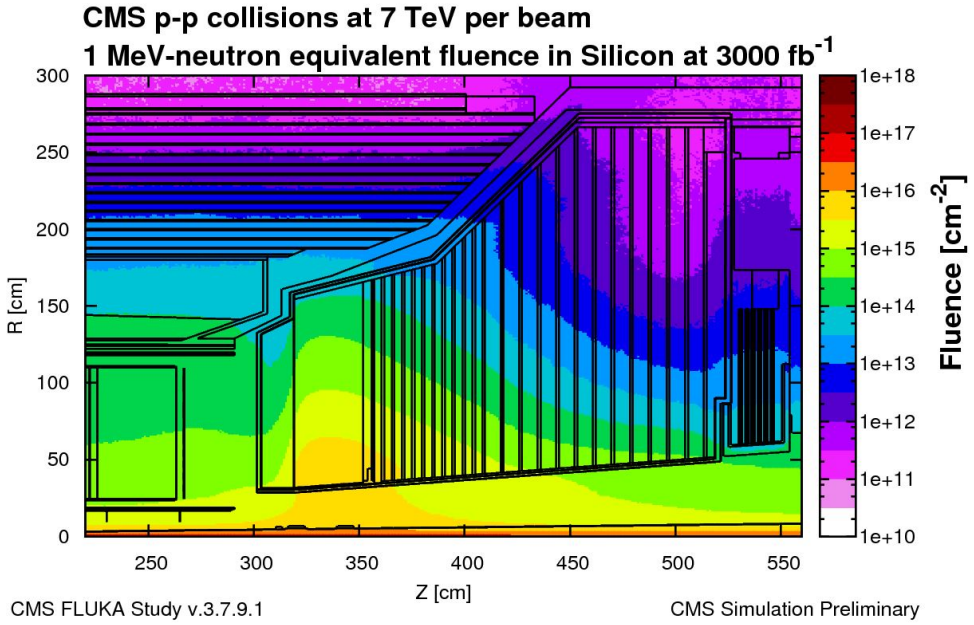


Figure 1.2: Fluence, parameterized as a fluence of 1 MeV equivalent neutrons, accumulated in HGCAL after an integrated luminosity of 3000 fb^{-1}, simulated using the FLUKA program, and shown as a two-dimensional map in the radial and longitudinal coordinates, r and z.

1.2 Requirements for the HGCAL upgrade

Preserving good performance over the full lifetime will require good (at the level of a few percent) inter-cell calibration. Adequate calibration accuracy can best be achieved if minimum-ionizing particles (MIPs) can be cleanly detected in each cell. This requires a good signal-to-noise ratio (S/N) for MIPs after 3000 fb^{-1}, necessitating the use of low-capacitance silicon cells, of a small size (≈0.5–1 cm^2), and scintillator cells of a small enough size for high light collection efficiency and S/N, resulting in a high lateral granularity. Fine longitudinal sampling is needed to provide good energy resolution, especially when using thin active layers (100–300 μm thick Si sensors). The fine lateral and longitudinal granularity leads to a high cell count. The main requirements for the HGCAL upgrade can be summarized as follows:

- *radiation tolerance:* fully preserve the energy resolution after 3000 fb^{-1}, requiring good inter-cell calibration (≈3%) using minimum-ionizing particles,
- *dense calorimeter:* to preserve lateral compactness of showers,
- *fine lateral granularity:* for low energy equivalent of electronics noise so as to give a high enough S/N to allow MIP calibration, to help with two shower separation and the observation of narrow jets, as well as limiting the region used for energy measurement to minimize the inclusion of energy from particles originating in pileup interactions,
- *fine longitudinal granularity:* enabling fine sampling of the longitudinal development of showers, providing good electromagnetic energy resolution (e.g. for H $\rightarrow \gamma\gamma$), pattern recognition, and discrimination against pileup,
- *precision measurement of the time of high energy showers:* to obtain precise timing from each cell with a significant amount of deposited energy, aiding rejection of energy from pileup, and the identification of the vertex of the triggering interaction,
- *ability to contribute to the level-1 trigger decision.*

1.3 High granularity and its impact on physics

The HGCAL will feature unprecedented transverse and longitudinal segmentation for both electromagnetic and hadronic compartments. The electromagnetic part will comprise 28 longitudinal samplings and the hadronic part 24 samplings. The electromagnetic and a large fraction of the hadron calorimeter will consist of silicon cells of size ($\approx$0.5–1 cm^2) and the remainder of the hadron calorimeter will use highly-segmented plastic scintillators ($\approx$4–30 cm^2). This high granularity has considerable potential benefits for reconstruction of physics objects. The high granularity will facilitate feature extraction and particle-flow calorimetry [2] using the finely measured lateral and longitudinal structure of electromagnetic showers, taus or jets. It can also be used for the subtraction of pileup energy allowing good energy resolution even in a high pileup environment.

A high-density calorimeter leads to laterally compact showers allowing good two-shower separation, that benefits particle-flow methods in the congested environment of the HL-LHC. The radial containment of electromagnetic showers is shown in Fig. 1.3. Narrow or merged jets can be reconstructed with higher efficiency and better energy resolution because the area used can be determined jet-by-jet by evaluating the local environment (Section 5.1.7). More powerful trigger algorithms can be developed by dynamically selecting the volume used for the measurement of energy and direction.

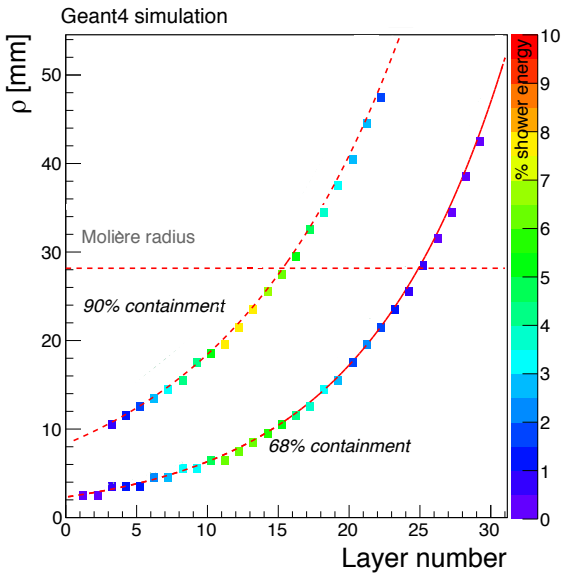


Figure 1.3: Radii, ρ, containing 68% and 90% of the energy deposited in an individual silicon layer by a photon shower, as a function of silicon layer. The colour-coded rectangles indicate the fraction of total energy deposited inside the 68% and 90% containment radii of each layer.

1.3.1 Longitudinal sampling

1.3.1.1 Electromagnetic energy resolution and H $\rightarrow \gamma\gamma$

The longitudinal segmentation for the electromagnetic calorimeter has been chosen so as not to degrade the H $\rightarrow \gamma\gamma$ mass resolution beyond what was achieved at the time of the discovery of the Higgs boson. Given that energies in the endcap are large for a given p_{T}, a stochastic term of $25\%/\sqrt{E}$, relatively large for an electromagnetic calorimeter, is found to be acceptable. This leads to the choice of 28 longitudinal layers in the CE-E. It has been verified that the degradation of the electromagnetic energy resolution, for small change in the number of sampling

layers, scales as $\sqrt{N_s^b}/\sqrt{N_s^a}$, where N_s^b and N_s^a are, respectively, the number of samplings of the baseline design and the number of samplings after a change.

We have also studied the impact on cost of the change in the number of sampling layers; removal of a layer in CE-E leads to a saving of ≈0.5 MCHF, i.e. 1 MCHF for one layer from each endcap. It may be noted that the scaling of cost reduction is slower than the loss of energy resolution, as shown in Fig. 1.4 (left).

Since the cells from different longitudinal layers cannot be ganged together on the detector they must be read out layer-by-layer: while this leads to a large number of read-out cells, it also provides significant benefits for physics reconstruction. Examples of this are the angular resolution and pointing capability for electromagnetic showers, which results from the combination of transverse and longitudinal segmentation, as well as the ability to improve hadronic energy resolution by software compensation. However, the overall sampling of the energy, 28 layers for an electromagnetic sampling calorimeter, is in fact rather coarse, resulting in the large estimated stochastic term.

1.3.1.2 Angular resolution and long-lived states in BSM physics

In the presence of pileup corresponding to a mean of 140 interactions per bunch crossing, an angular resolution of better than 4 mrad for photon showers with $p_T > 40$ GeV can be obtained (shown in Fig. 1.4 (right)).

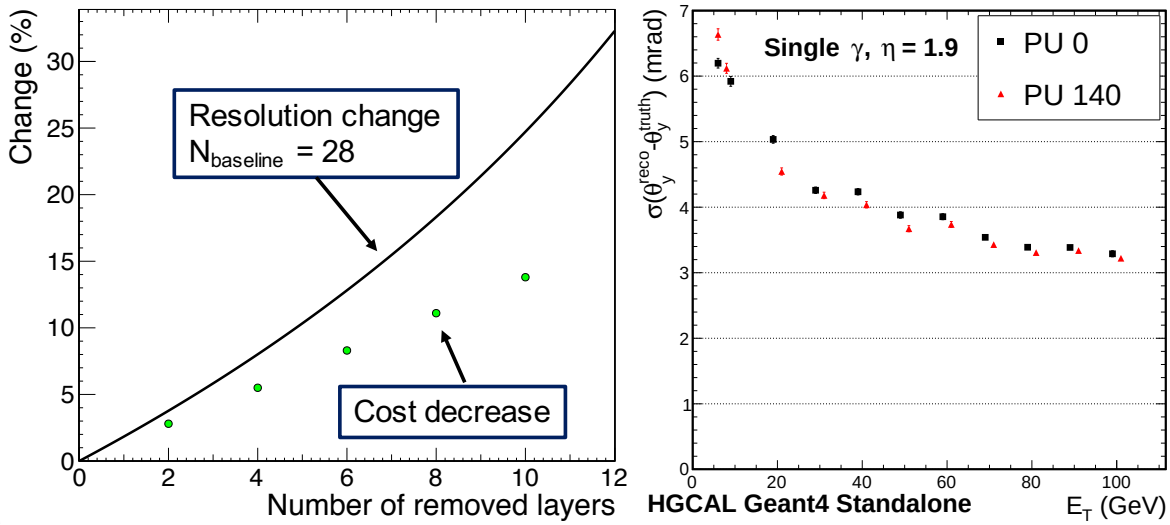


Figure 1.4: Electromagnetic energy resolution change, and cost decrease, as a function of the number of CE-E layers removed from a baseline configuration containing 28 layers (left). The energy resolution decrease assumes scaling as $\sqrt{N_s^b}/\sqrt{N_s^a}$. Angular resolution as a function of transverse energy for photons at $\eta = 1.9$, for zero pileup, and in the presence of pileup corresponding to a mean of 140 interactions per bunch crossing (right).

A number of physics models beyond the standard model (e.g. compressed SuSy, R-parity violating SuSy, hidden valley models, split SuSy etc.) predict states that would lead to signatures involving displaced vertices. For such states only mild constraints exist from the current Run 2 searches. The pointing capability of the HGCAL opens the possibility of a dedicated trigger at Level-1 for displaced objects with a decay length, $c\tau$, larger than a few centimetres. Such a Level-1 trigger would be a new feature in the LHC physics programme.

1.3.2 Transverse granularity

The transverse granularity is bounded, in the direction of finer granularity, by the need to restrict the power dissipated so as to limit the cooling requirements and, in the direction of coarser granularity, by the desire to preserve full performance throughout the HL-LHC operation (using MIP intercalibration as described in Section 5.4) and, also:

- to exploit as much as possible the narrowness of e.m. showers.
- noting that many other physics related performance parameters are affected by lateral granularity (timing, angular and energy resolutions)

The chosen cell-size is $\approx$1.2 cm^2 or $\approx$0.5 cm^2 for the Si part of the calorimeter. The exact choice of the lateral size of the cells is also constrained by considerations of trigger-cell size, that comprises a sum of either four of the larger-sized basic cells or nine of the smaller-sized basic cells. The size is thus quantized by the need to maintain an integer number of trigger cells within a Si wafer. The next size up from that chosen corresponds to an increase of the area of the basic cell by a factor 16/9 (=1.78). The reduction in cost if this size were chosen is estimated to be 2.4 MCHF, at the price of a significant impact on physics: loss of MIP intercalibration, poorer timing, angular and energy resolution; with some of these impacts severe to the point of rendering use of the measured quantity questionable.

1.3.3 Timing resolution

The intrinsic high-precision timing capabilities of the silicon sensors will add an extra dimension to event reconstruction. Arising naturally due to the intrinsically fast response time of the silicon sensors, and the design of the front-end electronics, each cell with significant deposited energy can give a precise time stamp. For example, the expected time resolution is 25 ps for an energy deposit equivalent to 50 fC, as compared to a spread of time from collisions in a single bunch crossing of a couple of hundred ps. For 300 μm thick sensors a charge of 50 fC corresponds to about 12 MIP at the start of HL-LHC. From simulation it is expected that good timing resolution will be possible both for electromagnetic deposits and charged hadrons (Section 5.5). Of especial note is the use of timing for pileup rejection (Section 5.5.1), for the identification of the vertex of the selected (triggered) interaction (Section 5.5.2 and Ref. [3]), and for aiding particle flow reconstruction.

1.3.4 Granularity summary

The power of the high granularity for pattern recognition can be visualized by looking at a display of a VBF H $\rightarrow \gamma\gamma$ event in one quadrant of the calorimeter (shown in Figs. A.1 and A.2 in Appendix A). These show the development of a VBF jet carrying 720 GeV ($p_{\mathrm{T}} = 118$ GeV) and a photon carrying 175 GeV ($p_{\mathrm{T}} = 22$ GeV) in an environment with a mean pileup of 200 interactions per bunch crossing. All the energy in the quadrant is projected into one plane, and the longitudinal development, plane-by-plane, in 9 planes of the electromagnetic section and 15 planes in the hadronic section. The wealth of information, the development of the showers and the narrowness of the VBF jet is clearly visible. Most of the energy of the VBF jet is carried by three particles (two charged pions and one photon) that impact the calorimeter within 1 cm of each other. A study using an anti-k_{T} algorithm to measure the energy of VBF jets has been carried to examine the rate from background. The results are shown in Section 5.1.7.

The design of the proposed CE calorimeter is well-suited to exploit maximally the HL-LHC physics that could appear in the endcap region. The chosen longitudinal and lateral readout samplings are both technically prudent and cost effective. They preserve, and in selected chal-

lenging areas extend, the physics reach of the HL-LHC programme. It is prudent to leave enough margin and redundancy to cater for unforeseen performance degradation or losses of elements of the calorimeter during the long operation in an unprecedentedly harsh environment of HL-LHC, especially when instantaneous and integrated LHC luminosity will be pushed to the limits of the possible ($4500\,\mathrm{fb}^{-1}$).

1.4 Longitudinal structure of the HGCAL

The HGCAL consists of an electromagnetic compartment (CE-E) followed by a hadronic compartment (CE-H), and the whole calorimeter sits in a thermally shielded volume that will be cooled by a two-phase CO_2 system and maintained at $-30\,^{\circ}\mathrm{C}$. The detailed longitudinal structure is shown in Fig. 1.5, and a full longitudinal cross section of the upper half of one endcap is shown in Fig. 1.6.

The electromagnetic compartment consists of 28 sampling layers with a total thickness of 34 cm and a depth of approximately 26 X_0 and 1.7 λ. The active detector element is a 163 mm wide hexagonal silicon sensor (Section 2.1) from an 8 inch (8″) wafer, which is sandwiched between a 1.4 mm thick WCu (75%, 25%) baseplate and a printed circuit board that carries the front-end electronics to form a silicon *module* (Section 2.2). Silicon sensors with three different sensitive thickness are deployed: 300, 200, and 120 μm, in regions of increasing fluence, respectively. Modules are tiled on either side of a 6 mm thick Cu cooling plate, which together with the two WCu 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 that are placed on either side of the module-cooling plate sandwich. Each plane of this structure is subdivided into $60\,^{\circ}$ units called *cassettes* (Section 4.2), and 14 layers of these cassettes provide the full 28 sampling layers.

The absorber in the hadronic compartment consists of 12 planes of 35 mm thick SS plates followed by another 12 SS planes with a thickness of 68 mm (Section 4.3). Between these absorber plates sit silicon modules and scintillator *tileboards* (Section 2.4) mounted on 6 mm thick copper cooling plates to form $30\,^{\circ}$ wide cassettes. These cassettes are similar to those in the electromagnetic compartment, but include sensors on only one side of the cooling plate and do not include the absorbers, which are formed as a separate mechanical structure. This leads to a total calorimeter thickness, perpendicular to the layers, of 10.7 λ, including the CE-E and the neutron moderator layer in front of the calorimeter. The polythene neutron moderator, whose purpose is to reduce the neutron flux in the Tracker, has a thickness of 120 mm. All layers are read out for use in energy measurement, but only alternate layers in CE-E, and all in CE-H, are used for making L1 trigger primitives.

In the CE-H, the radius, as a function of layer, at which the active medium changes from silicon sensors to plastic scintillator, as indicated in Fig.1.6, is determined by considerations of the radiation level. Sufficient margin must be allowed so that the radiation dose (Fig. 1.1) induced light loss in the scintillators does not exceed a reasonable amount (50%), whilst the neutron fluence (Fig. 1.2) does not exceed $8\times10^{13}\,\mathrm{n_{eq}/cm^2}$, so the energy equivalent of the electronics noise, due to increased SiPM leakage current and light loss, is kept low enough to allow a good measurement of the MIP response. More details are given in Section 2.3.

The parameter values of the design are summarized in Table 1.1. Further details of the silicon sensor cells, the scintillator tiles, scintillator properties, and the SiPMs are summarized in Tables 2.1–2.5.

The rest of Part I is structured as follows. Chapter 2 describes the active elements, namely sil-

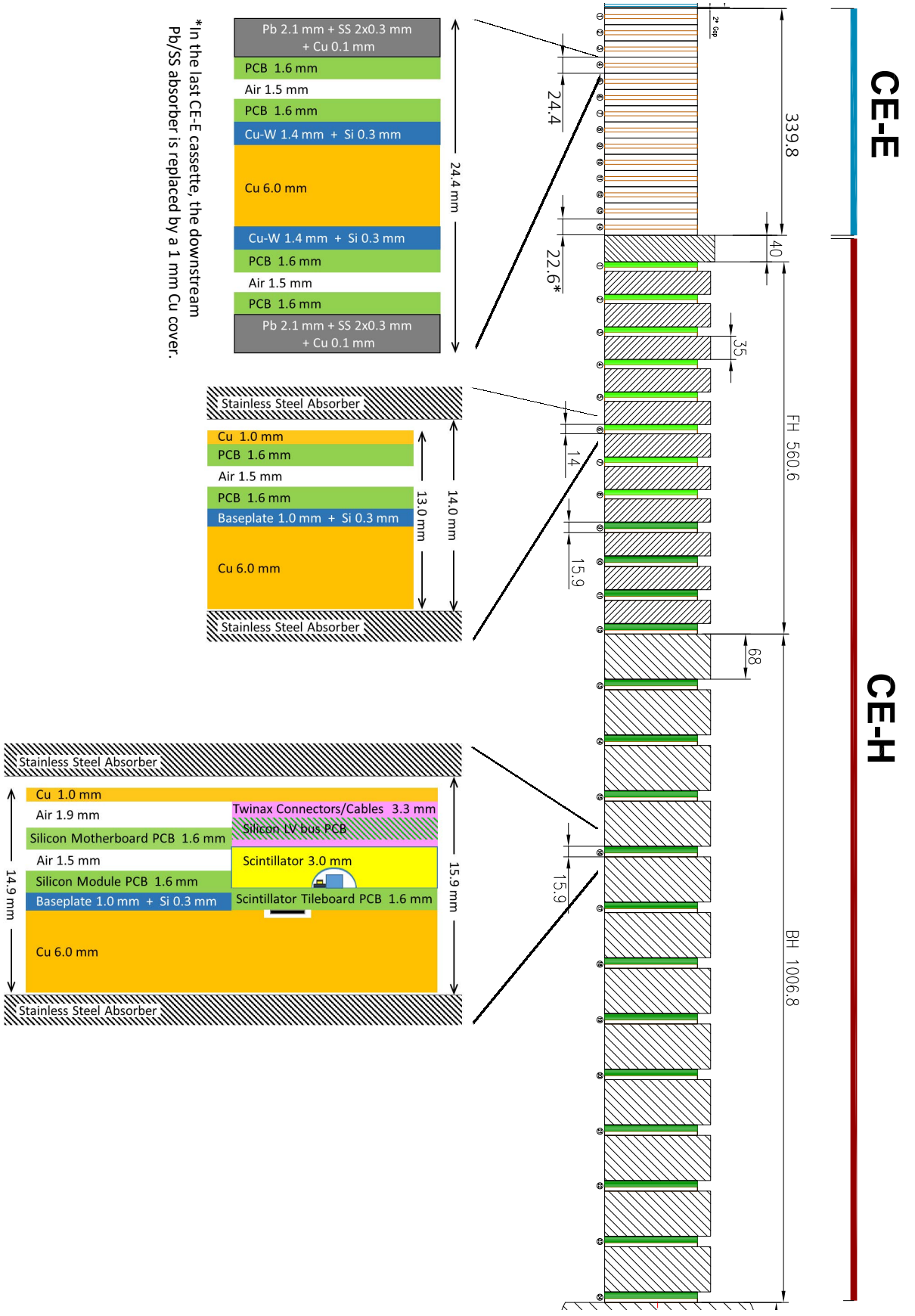


Figure 1.5: Longitudinal structure of the HGCAL, with schematic cross-sections of the three types of cassettes: CE-E cassettes, CE-H silicon sensor cassettes, and CE-H mixed silicon/scintillator cassettes. In the mixed cassettes the cross-hatched region is shared by the scintillator and silicon services in different angular regions.

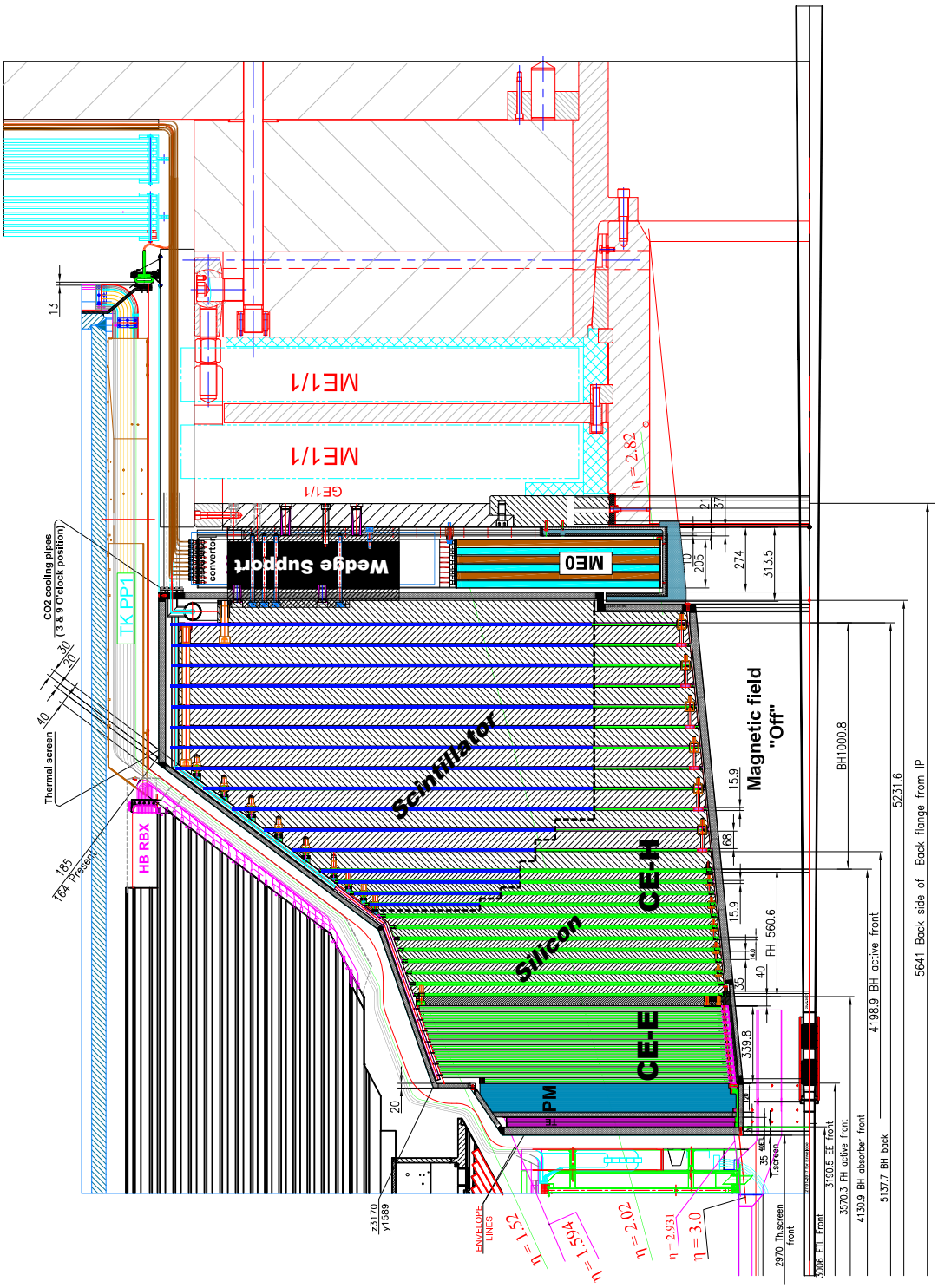


Figure 1.6: Longitudinal cross section of the upper half of one endcap calorimeter. It consists of an electromagnetic compartment (CE-E) followed by a hadronic compartment (CE-H). The green region to the lower left is instrumented with silicon detectors and the blue region to the upper right with scintillator tiles.

Table 1.1: HGCAL parameters. The values for area, channel count, modules, and partial modules are the totals for both endcaps. The weight and the number of planes refer to a single endcap.

	CE-E	**CE-H**	
	Si	Si	Scintillator
Area (m^2)	368	215	487
Channels (k)	3916	1939	389
Si modules (Tileboards)	16 008	8868	(3960)
Partial modules	1008	1452	–
Weight (t)	23	205	
Si-only planes	28	8	
Mixed (Si+Scint) planes		16	

icon sensors, and plastic scintillator tiles that are read out directly by silicon photomultipliers (SiPM); it also describes the construction of silicon modules and SiPM-on-tile modules. Chapter 3 outlines the electrical and electronics systems including the trigger primitive generation and back-end electronics. Chapter 4 gives an overview of the mechanical engineering including the structural design, the cassettes, and the assembly and installation. Chapter 5 presents the current status of the reconstruction and detector performance studies including those for L1 triggering. Lastly Chapter 6 outlines the project timeline and milestones, the cost and funding, and the project organization and institutional responsibilities.

Chapter 2

Active elements

2.1 Silicon sensors

2.1.1 Radiation tolerance

At the HL-LHC the silicon sensors of the HGCAL will be exposed to hadron fluences ranging from about 2×10^{14} up to about $10^{16}\,\mathrm{n_{eq}/cm^2}$ as shown in Fig. 1.2. These fluences are similar to those in the tracker and pixel volumes for the HL-LHC, and the basic parameters for the HGCAL sensor design draw upon the results obtained for the CMS Phase-2 Tracker R&D, and dedicated measurements using neutron fluences up to $1.6 \times 10^{16}\,\mathrm{n_{eq}/cm^2}$.

During the R&D for the existing silicon trackers of the LHC experiments, it was established that efficient charge collection requires operating the sensors well above full depletion voltage, and that the depletion voltage (after an initial type inversion period in the case of p-on-n sensors) increases substantially with irradiation, requiring operating bias voltages of 600 to 800 V towards the end of operation. Even at such high voltages, radiation induced traps in the bulk silicon significantly reduce the charge collected. In addition, the sensor leakage current increases linearly with fluence, resulting in increased noise and, combined with the very high bias voltages, leads to substantial power dissipation within the sensors themselves. These observations motivated the investigation of sensors thinner than the 300 μm typically used so far, in order to meet the challenges of the much higher levels of fluence expected at the HL-LHC, compared to the LHC.

The main differences between the tracker and the HGCAL are that the HGCAL will use large area pads ($\approx$1 cm^2) rather than finely segmented strips ($\approx$90 μm pitch), and that whereas in the tracker case the fluence is dominated by charged hadrons, in the case of the HGCAL the neutrons dominate.

Results from studies of the impact of neutron irradiation on the charge collection and leakage current behaviour of silicon pad sensors for the HGCAL are summarized in Figs. 2.1 and 2.2. These results demonstrate the advantage of deploying thinner sensors in the higher fluence regions of the calorimeter, in terms of improved charge collection as well as of reduced leakage current and power dissipation. Two types of wafer technologies were used for the samples under study: the first type consisted of standard float-zone (FZ) wafers of 320 μm thickness which underwent a back-side deep diffusion process (dd-FZ) in order to arrive at active nominal thicknesses of 300 μm, 200 μm, and 120 μm. The second type consisted of a high-resistivity epitaxial layer (epi) on a low resistivity wafer, that was grown to the nominal active thickness of 100 μm and 50 μm. The dd-FZ sensors were studied for both p-on-n (n-type) and n-on-p (p-type), whereas for the epi sensors only the p-type was included in this study.

At the start of the HL-LHC the collected charge is about 22, 15 and 9 ke$^-$ for sensors with 300,

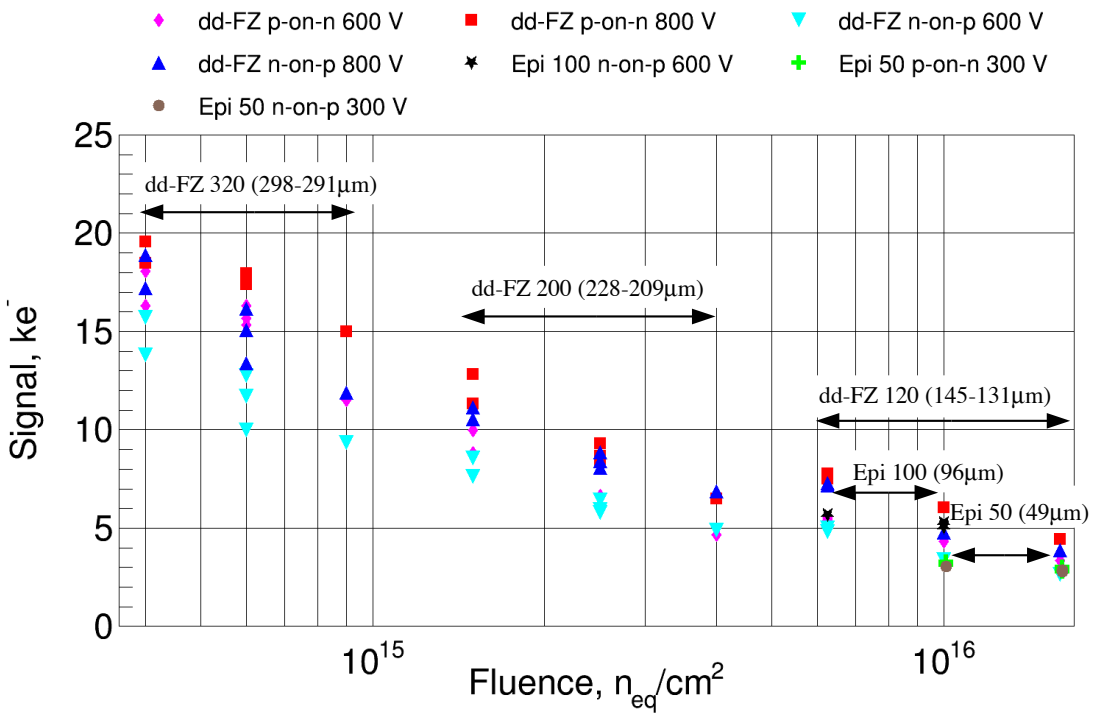


Figure 2.1: Signal (ke^-) versus neutron fluence, extracted from transient current technique measurements, described in Section 7.1.3. The range of fluences includes the maximum expected exposure after 3000 fb^{-1}. The arrows indicate the thickness of the different samples, the corresponding MIP charge yield before irradiation. The charge collection results are shown at 600 V and 800 V for the ddFZ diodes of 300 μm, 200 μm and 120 μm nominal thickness, at 600 V for the 100 μm epitaxial, and at 300 V for the 50 μm epitaxial diodes.

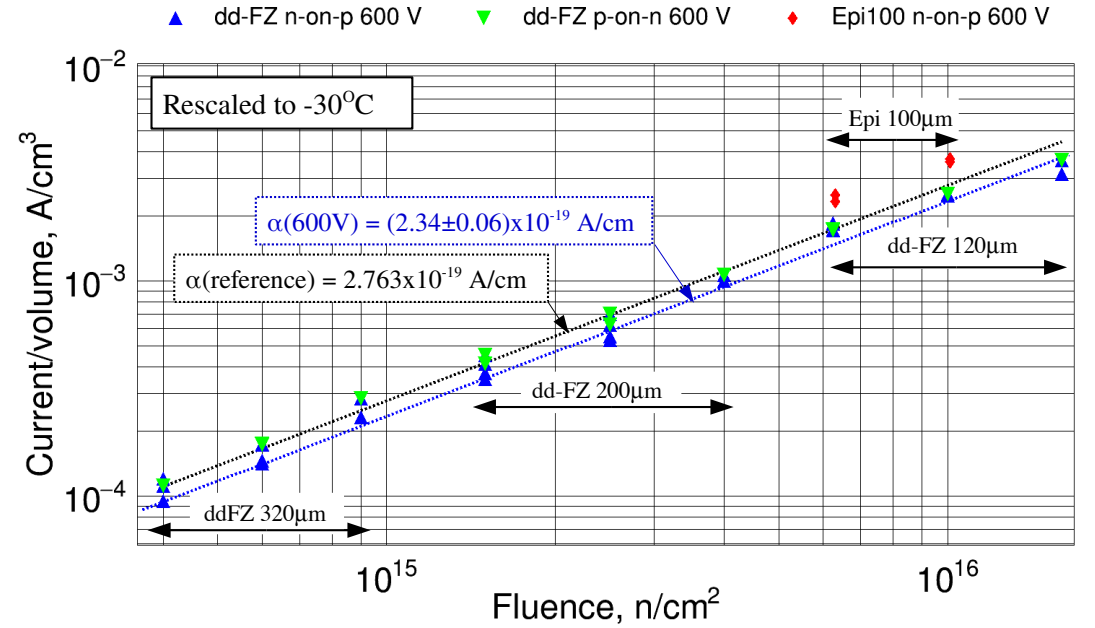


Figure 2.2: Leakage currents, as a function of neutron fluence, measured at $-20\,^{\circ}C$ and rescaled to $-30\,^{\circ}C$ at a bias voltage of 600 V. The arrows indicate the thickness of the different samples, and the range of fluences to which each sample was exposed.

200 and 120 μm nominal active thicknesses, respectively, at 600 V bias voltage. These numbers are calculated using 73 e^-/μm for the MIP charge deposition in silicon (the most probable value for 300 μm silicon [4]).

The maximum fluence values that correspond to an integrated luminosity of 3000 fb^{-1} are 5×10^{14}, 2.5×10^{15}, and 7×10^{15} n_{eq}/cm^2 for the 300, 200, and 120 μm thick sensors, respectively. Within the expected maximum fluence of 5×10^{14} n_{eq}/cm^2 the 300 μm thick n-type pad sensors appear to collect about 20% more charge than p-type sensors, whereas for sensors of thickness 200 μm or less, at the higher expected maximum fluences, no significant difference is observed in the charge collection of the two sensor types. Based on the measurements shown above, for p-type sensors operated at 600 V, the collected MIP charge after 3000 fb^{-1} is estimated to be 10 ke^- and 6 ke^- for the dd-FZ 300 μm and 200 μm nominal active thickness sensors, respectively. In the case of the nominal 120 μm active thickness sensors, it can be seen that better charge collection is obtained with the 100 μm epi sensors, which yield $\approx$5 ke^- per MIP. This may be compared to the dd-FZ sensors of 140 μm (120 μm nominal) active thickness which result in $\approx$4 ke^- per MIP. Based on the current proposals from the candidate producers, sensors with 300 μm and 200 μm active thicknesses will be produced on FZ wafers (rather than dd-FZ) thinned to 300 μm and 200 μm, respectively. Sensors with a thickness of 120 μm will be produced using epitaxial wafers of a physical thickness of 300 μm. In what follows we assume a MIP charge collection of 5 ke^- after full exposure to 3000 fb^{-1} for the 120 μm thick epitaxial sensors.

At $-20\,^\circ$C and at bias voltages of 600 and 800 V, the measured leakage currents for irradiated sensors (Fig. 2.2) are consistent with expectations and linearly proportional to both the sensor thickness and integrated fluence.

2.1.2 Sensor properties and layout

The silicon sensors for the CE-E and the inner parts of the CE-H will be planar DC-coupled hexagonal silicon sensors fabricated on 8 inch (8″) wafers. The hexagonal shape of the sensors makes more efficient use of the available area of the circular wafers, as compared to square or rectangular sensors, while minimising the ratio of periphery to active surface. The vertices of the hexagonal sensors are truncated, allowing clearance for the mounting/fixation system, and further increasing the use of the wafer surface. Designs based on both hexagonal as well as square sensors were studied, and the system level implications of the different geometries were considered and compared. These studies confirmed that, due to more complete use of the silicon wafers, deploying hexagonal sensors would result in a substantial cost reduction.

As discussed above, sensors will have three different active thicknesses (300, 200 and 120 μm) in order to optimize the charge collection and operation conditions over the full lifetime of the HGCAL. The baseline substrate material is physically thinned p-type FZ silicon wafers for the 300 and 200 μm thick sensors, and p-type epitaxial on a handle wafer is an option for the 120 μm thick sensors. P-type sensors are preferred, as these have been shown to be more robust against non-Gaussian noise due to radiation-induced surface charge effects [5]. The possible use of n-type 300 μm sensors in the lower fluence part of the HGCAL remains an option under study, contingent on demonstrating a design which remains unaffected by non-Gaussian noise up to a fluence of 10^{15} n_{eq}/cm^2, because this may result in cost savings compared to the baseline p-type sensors.

Different sensor cell geometry and tiling options have been examined, and their system level implications have been considered. The ”three-fold diamond” configuration, shown schematically in Fig. 2.3, was chosen because it allows convenient definition of symmetric sets of 2×2

and/or 3×3 neighbouring cells to form trigger primitives, shown as differing colour groupings in the figure, and the subdivision of the module into symmetric domains for the readout chips, simplifying the layout of the module readout printed circuit board (PCB). Silicon wafer layouts using the three-fold diamond configuration are shown in Fig. 2.4.

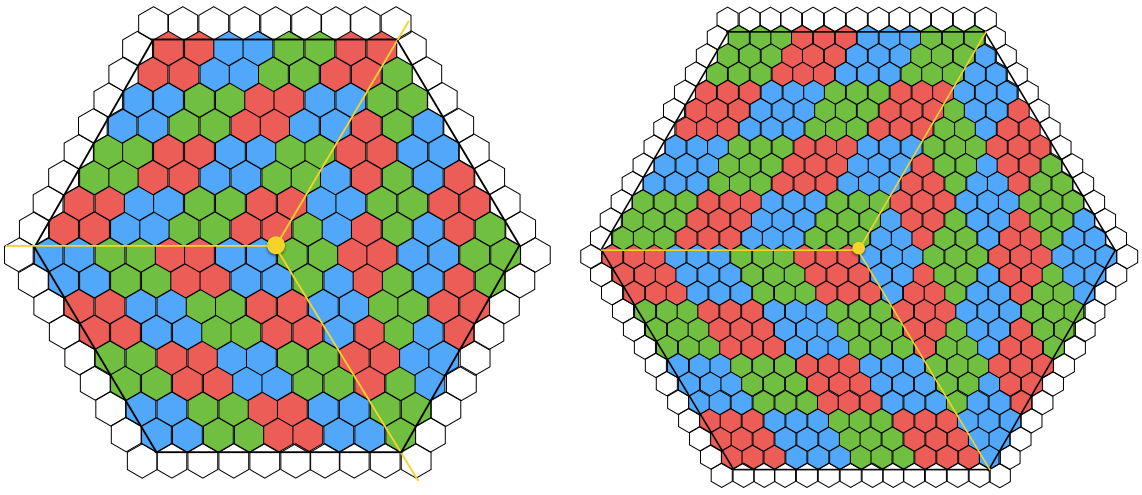


Figure 2.3: Schematic illustration of the three-fold diamond configuration of sensor cells on hexagonal 8″ silicon wafers, showing the groupings of sensor cells that get summed to form trigger cells, for the large, 1.18 cm^2, sensor cells (left), and for the small, 0.52 cm^2, cells (right).

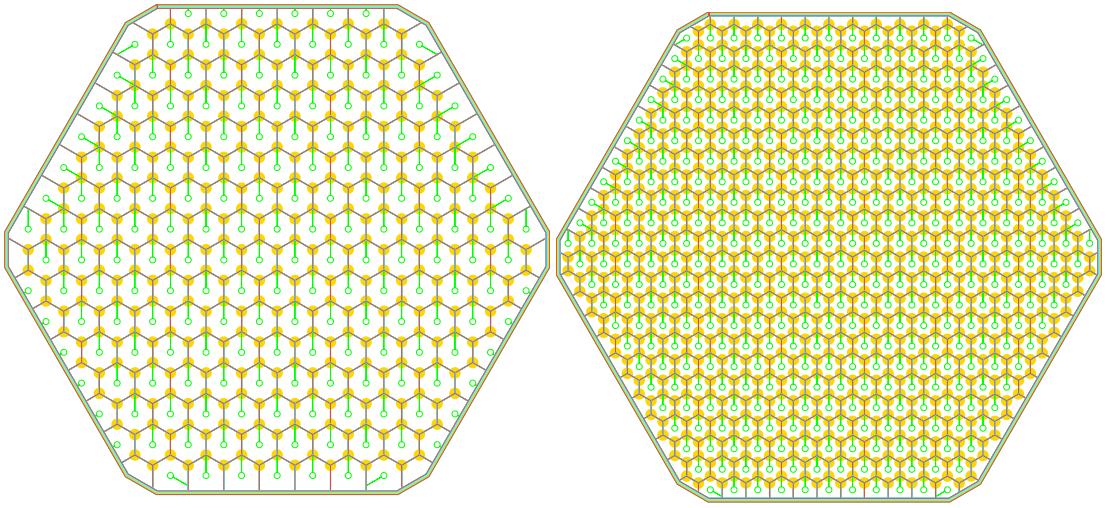


Figure 2.4: Drawing of hexagonal 8″ silicon wafers, with layout of large, 1.18 cm^2, sensor cells (left), and small, 0.52 cm^2, cells (right).

The cell size is driven both by physics performance considerations, such as the lateral spread of electromagnetic showers, and by constraints imposed by the need to keep the cell capacitance within a manageable range. In practice, this results in cell sizes of $\approx$1 cm^2 for the 300 and 200 μm active thickness sensors and $\approx$0.5 cm^2 for the 120 μm active thickness sensors, corresponding to a maximum cell capacitance of 65 pF. Each sensor has either 192 or 432 individual diodes, which act as sensor cells. The HV bias is applied to the sensor back-plane, whereas the ground return from each individual cell is provided through the DC connection to the corresponding front-end amplifier. Two cells per readout chip are segmented to include calibration pads with smaller size and correspondingly lower capacitance and noise.

An irradiation campaign is underway, which will include noise measurements, with a partic-

ular attention to possible non-Gaussian noise tails, in addition to the usual IV, CV, and other device characterizations using dedicated test structures. A similar campaign is foreseen for the qualification of full size 8″ sensors of the final design.

The sensor cells are DC coupled to the on-module readout electronics, which maintain them at virtual ground and can provide up to 20 μA per cell. The readout chips will be connected to each sensor cell by wire bonds through holes in the readout PCB to the sensor. Prototype wafers have been tested using single probes, 7-probe arrays, and full-wafer pogo-pin based probe cards. Studies are ongoing of these various probe card configurations aimed at optimizing the sensor testing protocols, both at the production site by the vendor and on reception.

Table 2.1 gives an overview of the three sensor types and their features. Tables 2.2 and 2.3 list further details, giving the area covered, the cell size, and the S/N for a MIP before and after a lifetime integrated luminosity of 3000 fb^{-1}. The approximate inner and outer radii of the regions are given for guidance. The layout of a longitudinal layer equipped with Si wafers is shown in Fig. 2.5.

Table 2.1: Different sensor types, their properties, the expected radiation fluence for each type, and their number for two endcaps.

Active thickness (μm)	Cell size (cm^2)	Cell capacitance (pF)	Bulk polarity	Expected range of fluence ($\times 10^{15}$ n$_{eq}$/cm^2)	Number of wafers	Number of partial wafers
300	1.18	45	p / (n)	0.1–0.5	13 164	1284
200	1.18	65	p	0.5–2.5	8 712	144
120	0.52	50	p	2–7	3 000	324
				Total:	24 876	1752

Table 2.2: Silicon sensors in CE-E and CE-H layers having only silicon sensors, showing thickness of active silicon, cell size, and S/N for a MIP before and after an integrated luminosity of 3000 fb^{-1}. The area shown is the total for two endcaps. After 3000 fb^{-1} the S/N(MIP) varies as a function of fluence received, and only the lowest values are given. The approximate inner and outer radii of the regions are given for guidance.

Active thickness (μm)	**300**	**200**	**120**
Area (m^2)	245	181	72
Largest lifetime dose (Mrad)	3	20	100
Largest lifetime fluence (n$_{eq}$/cm^2)	0.5×10^{15}	2.5×10^{15}	7×10^{15}
Largest outer radius (cm)	$\approx$180	$\approx$100	$\approx$70
Smallest inner radius (cm)	$\approx$100	$\approx$70	$\approx$35
Cell size (cm^2)	1.18	1.18	0.52
Initial S/N for MIP	11	6	4.5
Smallest S/N(MIP) after 3000 fb^{-1}	4.7	2.3	2.2

2.1.3 Sensor production

In order to reduce the risks associated with the production of 600 m^2 of silicon sensors required for the HGCAL, we have been pursuing a similar dual-source strategy, which was taken by

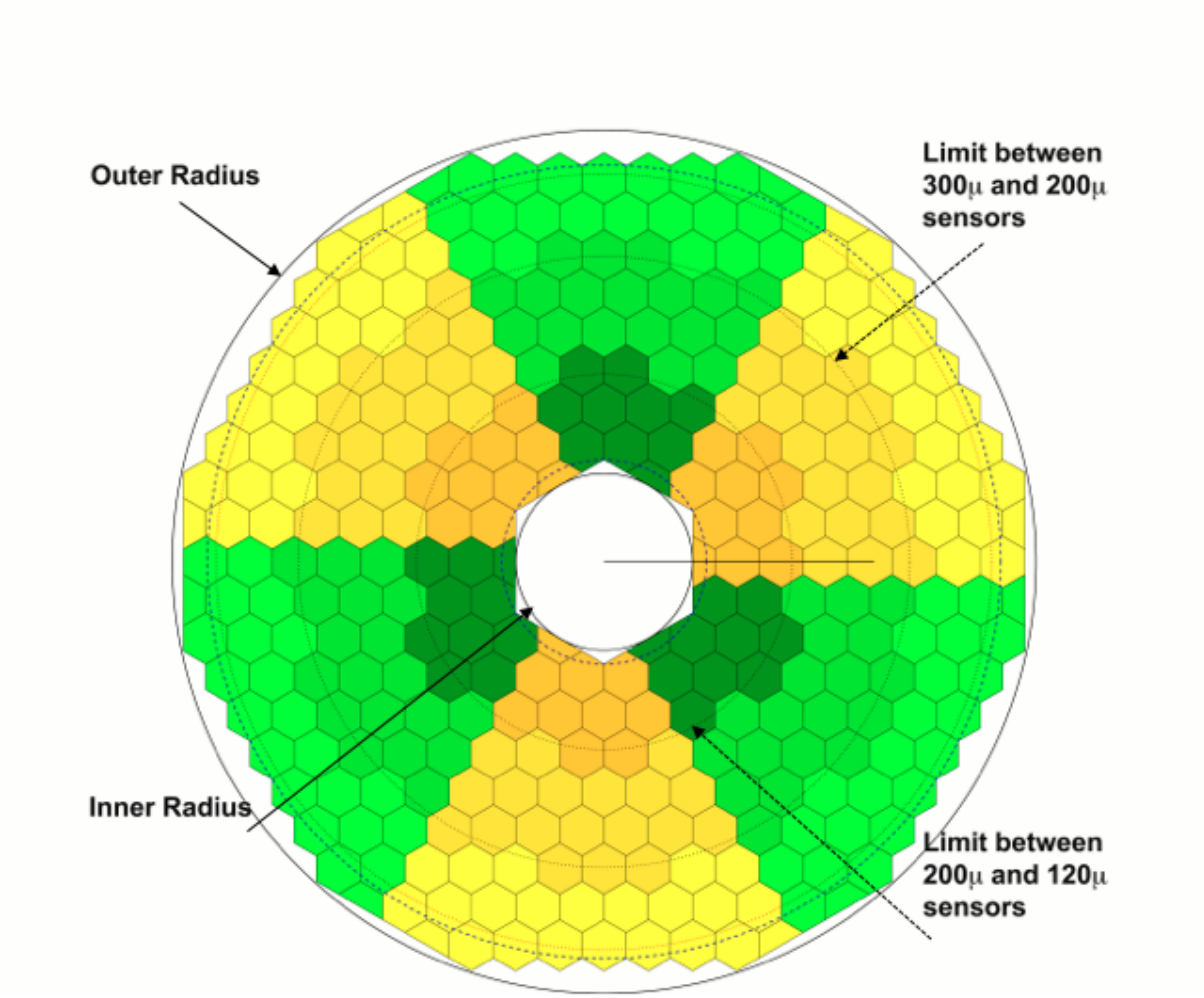


Figure 2.5: Layout of a layer where only silicon sensors are present, the 9th layer of CE-E. The division into six 60° cassettes is shown by the alternating colours. The two radial changes in darkness of colour indicate the changes in silicon thickness. The inner and outer radii are 32.8 cm and 160 cm, respectively.

Table 2.3: Sensors in layers with both silicon and scintillator in CE-H, showing cell size and S/N for a MIP before and after an integrated luminosity of 3000 fb^{-1}. The area shown is the total for two endcaps. The S/N(MIP) varies as a function of the cell size for the scintillator cells, and varies with fluence for all cells after 3000 fb^{-1}; and only the lowest values are given. The approximate inner and outer radii of the regions are given for guidance.

	Scintillator	**Si**	**Si**
Sensor thickness	3 mm	300 μm	200 μm
Area (m^2)	480	71	15
Largest lifetime dose (Mrad)	$<$0.3	30	100
Largest lifetime fluence (n_{eq}/cm^2)	8×10^{13}	5×10^{14}	2.5×10^{15}
Largest outer radius (cm)	$\approx$235	$\approx$160	$\approx$100
Smallest inner radius (cm)	$\approx$90	$\approx$80	$\approx$45
Cell size (cm^2)	2×2 to 5.5×5.5	1.18	1.18
Initial S/N for a MIP	$\gg$5	11	6
Smallest S/N(MIP) after 3000 fb^{-1}	5	4.7	2.3

both CMS and ATLAS Tracker sensor production, although each company could produce the full quantity of sensors if needed. The qualification of suitable companies has begun within the framework of an ongoing CERN market survey [6]. This has resulted in the identification of three potentially suitable companies matching the required profile for the production of the 8″ HGCAL sensors: these companies will be subject to a further technical qualification.

We have received numerous 6″ and 8″ sensors, listed in Tables 7.1, 7.2, and 7.3.

The Japanese company Hamamatsu Photonics K.K. (HPK) is a longstanding and, at present, the only qualified high-volume producer for 6″ sensors. Because of the interest of 8″ sensors for the HGCAL project, HPK has implemented this technology on 8″ production lines and has begun delivery of full-size 8″ demonstrator sensors, with similar design and characteristics to the HGCAL sensors. These sensors demonstrate that the process quality achieved with the 8″ production lines is comparable to that of the well-established 6″ production lines. These demonstrator sensors were produced with the so 'stepper' technique for the lithography steps, while large scale production will use a full-wafer lithography process. The stepper technology for photolithography uses small masks to consecutively expose small reticles on the wafer. This results in excellent photolithography, but limits the geometry of the cell and of the periphery. The specifications for the full wafer lithography presently being implemented by HPK on the 8″ production line are fully consistent with those used for sensor production on the reference 6″ line, and HPK is therefore confident that this change of process will not have any effect on the resulting sensor quality and characteristics.

The European semiconductor vendor Infineon (IFX) produced several batches of AC-coupled strip sensors, both on 6″ and 8″ technology, for the CMS Tracker [7]. Starting from this project, they also manufactured a first prototype batch for the HGCAL using a realistic geometry, with encouraging results so far for 300 μm thick sensors. In parallel with continued optimization of their 300 μm thick sensor process, Infineon is also actively pursuing the development of the necessary process variants for the production of 200 and 120 μm thick sensors.

The US-based company Novati has also delivered samples of demonstrator sensors, of both

Tracker and HGCAL types. So far these have been based on a silicon-on-silicon process with Si-to-Si wafer bonding. As compared to the standard process for the sensor production this has the advantage of enabling sensor production of varying active thickness in a transparent way. As part of their ongoing process optimization, Novati is also working towards mitigating any possible cost implications related to the additional steps required by this technique, in order to ensure a cost-effective solution.

2.2 Silicon modules

The HGCAL requires approximately 27 000 silicon detector modules to be assembled and installed in its electromagnetic (CE-E) section and part of the hadron (CE-H) section. A CE-E module is a stack of components, as shown in Fig. 2.6: (i) baseplate, (ii) Kapton-gold sheet, (iii) silicon sensor, and (iv) the printed circuit board (PCB), labeled the hexaboard, with front-end electronics. All of these components have roughly the same area as the sensor.

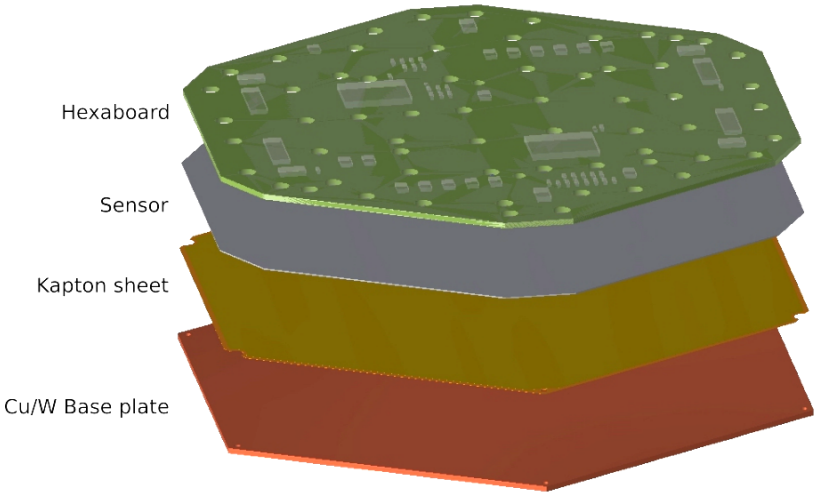


Figure 2.6: CE-E silicon module, showing stacked layers.

A study of mechanical mockups of 8″ modules demonstrates that the module is pressed onto the cooling surface by the differential cooling of the differing material layers, improving thermal contact. Calculated stresses on the module were found to be three orders of magnitude below the point at which breakage would occur.

We have been using 6″ modules (i.e. modules produced using sensors manufactured on 6″ wafers) for prototyping and for measurements in test beams. Figure 2.7 shows a completed 6″ module.

The baseplate has precise reference holes for precision assembly and placement onto the cassettes. For the CE-E the baseplate material is a sintered WCu metal matrix composite. The copper provides excellent thermal conductivity (TC), the tungsten reduces the coefficient of thermal expansion (CTE) to align it more closely with that of the silicon, and together they form a short radiation-length material that is a significant component of the CE-E absorber. For the CE-H modules, the baseplate material is high-TC carbon fibre. It serves similar purposes except that it does not contribute significantly to the CE-H absorber material.

The WCu baseplates are specified to have uniform thickness of 1.40 ± 0.03 mm, with faces that are flat to within $\pm 50\,\mu$m.

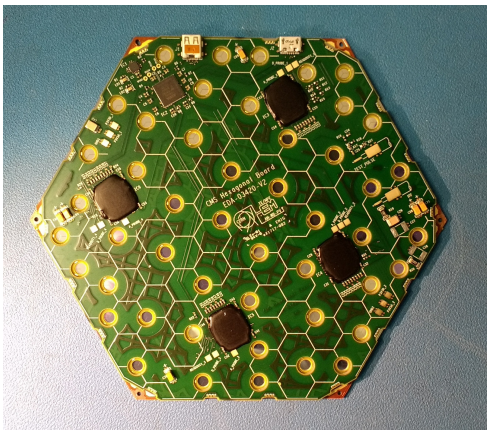


Figure 2.7: Six-inch module on carrier plate prior to wirebonding and encapsulation. The electronics packages seen on the PCB are four SKIROC2-CMS front end readout chips, used for the beam test, and an FPGA. The corners of the PCB and the sensor under the PCB are removed to provide direct access to the mounting holes on the baseplate.

A 105 μm thick Kapton foil coated with a thin layer of gold is epoxied to the baseplate, very nearly covering it completely. The thin layer of gold, on the exposed side of the Kapton, is used to provide the HV bias connection to the sensor back-plane through a conducting epoxy bond. The Kapton itself provides electrical insulation of the sensor back-plane from the baseplate, which is held at ground.

The silicon sensors and the hexaboard are hexagonal with small cutouts at each of the six corners. The cutouts provide access to the positioning and mounting holes in the baseplate. They also provide access to a portion of the Kapton-Au layer for wirebond connections to the hexaboard, for the biasing of the sensor back-plane. The hexaboard will contain the HGCROC front-end readout ASICs (Section 3.1.2). The signals from the sensor pads are routed to the HGCROC for on-board signal digitization. Holes in the hexaboard expose the region around the intersections of groups of up to four pads. The layout of the module hexaboard for a 432 channel sensor, and a zoomed view of the wirebond holes, are shown in Fig. 2.8.

The baseplate, Kapton, silicon sensor, and hexaboard are bonded together with epoxy to form a single physical unit. Sets of multiple wirebonds are made between an Au bonding pad on the hexaboard and each sensor pad, the Au-Kapton layer that provides backplane biasing, and the sensor guard rings. The wirebonds are protected by encapsulating them with a clear, radiation tolerant silicon elastomer. An example of the wirebonds in the prototype modules for beam tests is shown in Fig. 2.9. An automated assembly process has been developed with high-rate production in mind; it is fully described in Section 7.2.

A rigorous quality control system is necessary to achieve high yield during production; this is described in Section 7.2.2.2.

2.3 Plastic scintillators and photodetection

As previously stated, the replacement of the hadronic section of the current CMS endcap calorimeter (HE) is required due to the significant signal loss that will occur even before the HL-LHC running period begins. The specifications for the scintillator section of the HGCAL are driven by the requirement that it be possible to calibrate the detector throughout its life using MIPs.

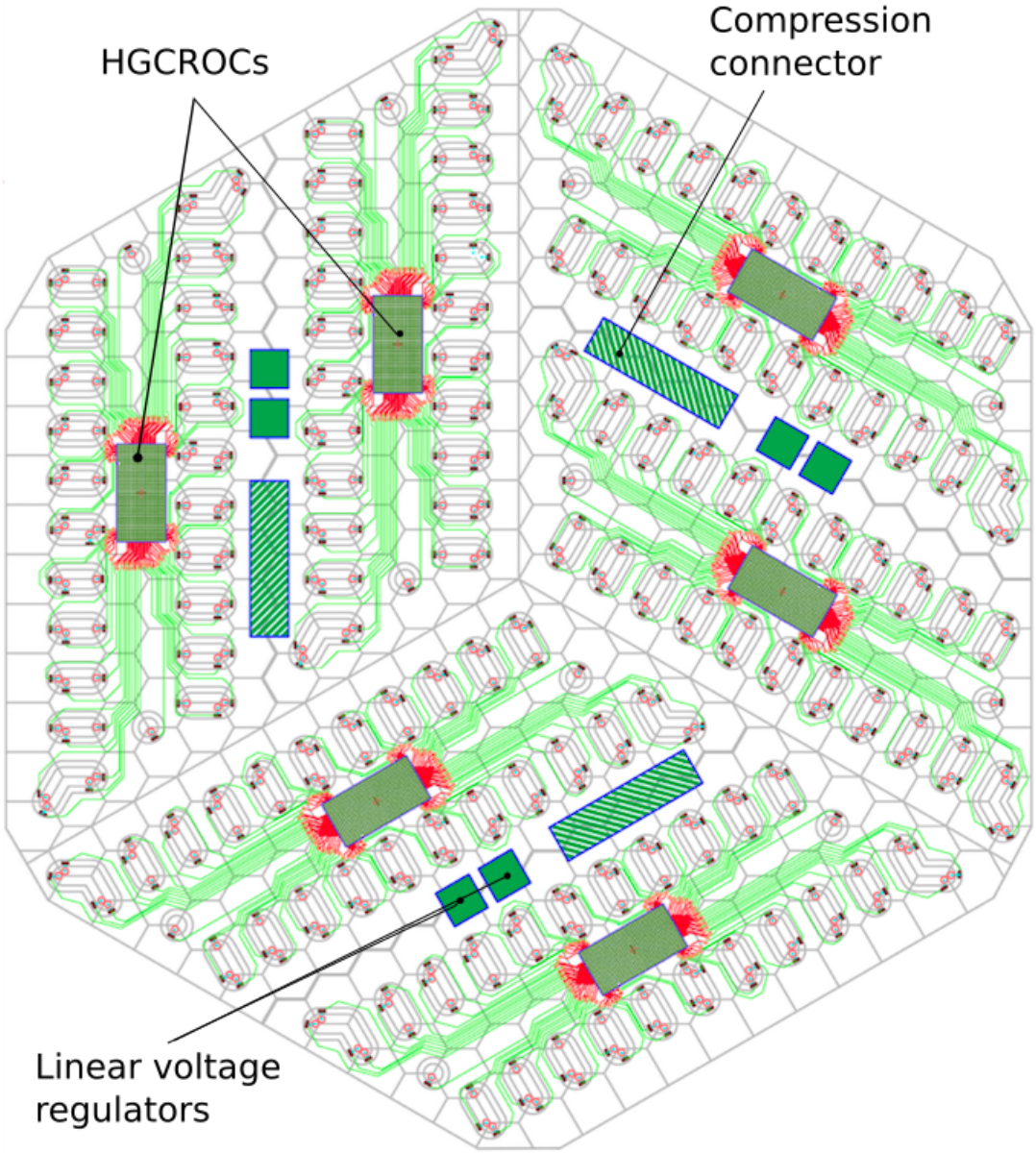


Figure 2.8: Layout of the module PCB, the hexaboard, for a 432 channel sensor, showing the oblong holes for wirebonds, the placement of the HGCROCs, the compression connector, and the linear voltage regulators.

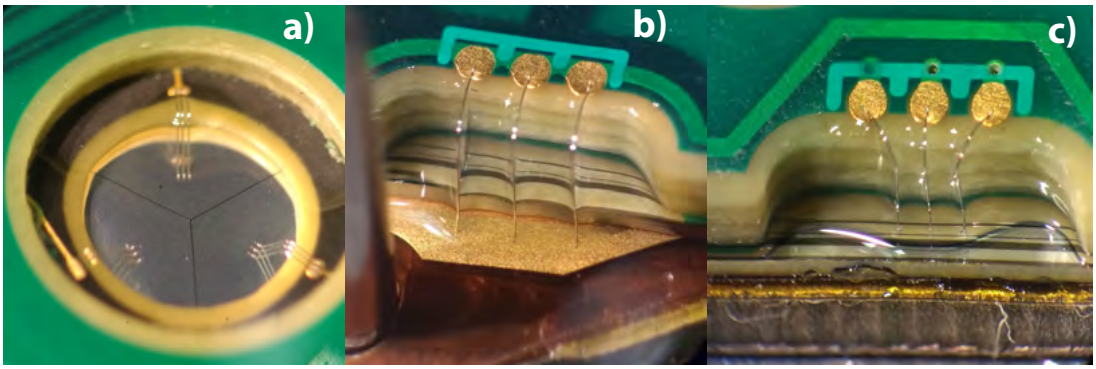


Figure 2.9: From left to right: a) wirebonds for three sensor pads at a stepped hole in the hexaboard used for test beam modules in 2017; b) wirebond to the Au-Kapton layer to provide back-plane biasing of the sensor; and c) wirebonds at the edge of the module to the sensor guard rings.

2.3.1 Plastic scintillator material

The effect of ionizing radiation on the existing HE has been studied in detail using UV-laser excitation, radioactive sources, and beam data [8]. These measurements of the current detector agree and indicate the presence of a dose-rate dependence in the damage to the scintillator. The dose-rate dependence, with increased damage resulting from a lower dose-rate, is understood to be an effect of scintillator chemistry and oxygen diffusion [9, 10]. In addition, the design of the endcap calorimeter requires that the scintillator be operated at a temperature near $-30\,^{\circ}\mathrm{C}$. While plastic scintillator has been operated at low temperature in space-based experiments, the radiation tolerance of plastic scintillator and any dose-rate dependence at low temperature have not previously been studied in detail.

The dose-rate-dependence measurements carried out on a number of materials are summarized in Fig. 2.10. The dose constant D_c is defined as the dose after which the light output decreases by a factor e. The square symbols on the plot show measurements from the current HE detector using data from the last $20\,\mathrm{fb}^{-1}$ of integrated luminosity collected during 2016 data taking. The other points are from dedicated studies which have been performed in irradiation facilities including irradiations carried out at low temperature. In particular, the points clustered around 0.3 krad/hr show the consistency between the performance of scintillator irradiated at room temperature and scintillator irradiated at low temperature. These results are discussed in more detail in Section 7.3.2.1.

The in situ measurements are the only measurements that directly probe the relevant range of dose rates for CE-H scintillator at HL-LHC. This range is indicated in Fig. 2.10. To make performance predictions we fit the data to the function

$$D_c = (3.6\,\mathrm{Mrad})\left(\frac{R}{1\,\mathrm{krad/hr}}\right)^{0.5}, \tag{2.1}$$

where R is the dose rate in krad/h. This fit is shown by the red line in the figure.

Two types of scintillator material can be used: polyvinyltoluene-based (PVT) scintillator available from multiple vendors (EJ-200 from Eljen or BC-408 from Saint-Gobain), or polystrene-based (PS) scintillators, which can be produced economically using injection-moulding techniques. The PS scintillators can be produced at participating institutes, such as the SC-301

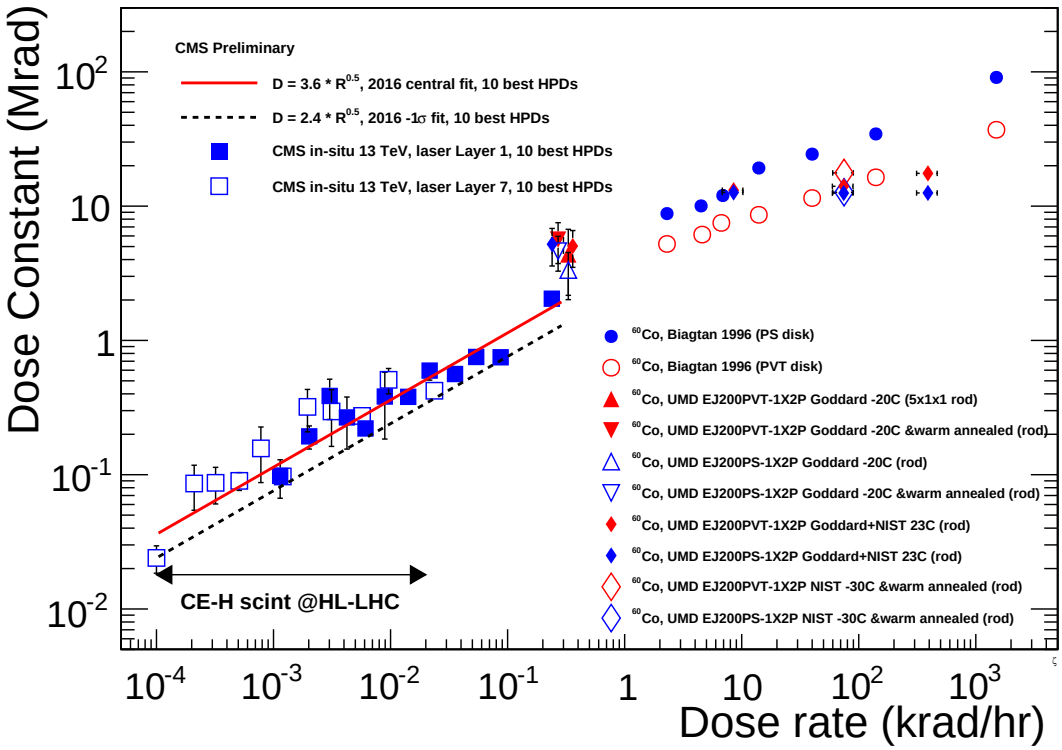


Figure 2.10: Dose constant, D_c, versus dose rate, R, for both in situ measurements from CMS (boxes) and dedicated studies. Lines corresponding to $D_c = 3.6 \times R^{0.5}$, and $D_c = 2.4 \times R^{0.5}$ are shown. The in situ measurements refer to SCSN81, a PS based scintillator used in the endcaps of the present detector.

scintillator produced at Protvino [11].

The properties of sample PVT and PS scintillators are given in Table 2.4. Considerations of cost, performance (at the end of HL-LHC exposure), and ease of assembly will determine the final choices.

Table 2.4: Properties of candidate polyvinyltoluene (PVT) and polystyrene (PS) scintillators.

	EJ-200	BC-408	IHEP SC-301
Base material	PVT	PVT	PS
Light output (% anthracene)	64	64	60
Scintillation efficiency γ/MeV	9000	9000	8500
Wavelength of max. emission (nm)	425	425	420
Rise time (ns)	0.9	0.9	1.4
Decay time (ns)	2.1	2.1	2.5
CTE (K^{-1})	78×10^{-6}	78×10^{-6}	70×10^{-6}

2.3.2 SiPM photodetectors

The light from the scintillation tiles will be read by SiPM photodetectors. These photodetectors, used in several systems including the Phase-1 HCAL upgrade, provide a very high gain (greater than 10^5) and an excellent photodetection efficiency (>30%). The photodetectors used

for the Phase-1 HCAL upgrade have been studied in detail, including evaluations of quantum efficiency, gain stability, and behaviour under irradiation [12]. Given the large data sample available for these SiPMs and their use in a hadron-collider experiment, the Phase-1 HCAL SiPMs are considered as the reference for the scintillator section of the endcap calorimeter upgrade.

The specifications for the endcap calorimeter SiPMs are given in Table 2.5. These are based primarily on the Phase-1 HCAL SiPMs, with a more relaxed requirement for the dynamic range due to the SiPM-on-tile geometry, in which each cell is read separately (in the case of the Phase-1 HCAL, multiple cells are ganged optically into a single SiPM). Each SiPM contains a fixed number of individual pixels, which makes the devices inherently nonlinear. The degree of nonlinearity is highly predictable and corrections can be applied, which is simplest when each tile is read separately. The nonlinearities are minimized by uniform light illumination, as has been demonstrated for the SiPM-on-tile geometry, and by fast recovery of the SiPM's individual pixels.

Table 2.5: Properties of a silicon photomultiplier from one vendor.

Effective photosensitive area	2 mm^2
Pixel pitch	15 μm
No. of pixels	7500
Geometrical fill factor	50%
Spectral response range	320–900 nm
Peak sensitivity	460 nm
PDE at peak sensitivity	30%
Gain	2×10^5
Pixel recovery time	10 ns
Breakdown voltage (V_{BR})	65 ± 10 V
Operating voltage	V_{BR} + (2–3) V
Dark count unirr. at 23 °C	100 kHz

While the dark current of SiPMs is quite low for unirradiated devices, it rises significantly after irradiation, as is common for silicon devices. The dark current and the corresponding rate at which pixels in a SiPM fire in the absence of illumination (the noise rate) is linear with the neutron fluence (n_{eq}/cm^2). This characteristic has been evaluated up to very high fluences ($2\times10^{14}\, n_{eq}/cm^2$) for the Phase-1 SiPM devices [13]. The dark current, and thus the noise rate, is also temperature-dependent, as is also common for silicon devices. As is discussed further in Section 7.3.3.1, the observed dependence of noise rate on temperature is smaller for SiPM devices than for unity-gain devices such as PIN diodes—the power law observed for the SiPMs used in the HCAL Phase-1 upgrade results in an increase in the noise rate by a factor of 1.88 for every increase of 10 °C.

For the purposes of signal reconstruction, however, the key aspect is the statistical fluctuation of the number of pixels firing, which increases as the square root of the dark current, provided that the cell-to-cell crosstalk is small.

2.3.3 SiPM-on-tile design

As discussed above, the design requirements for the detector are set by the necessity of maintaining the ability to calibrate individual tiles using minimum-ionizing particles through the life of the calorimeter. In addition, it is important that the services of the scintillator calorimeter be compact and not use too large a volume at the outer edge of the calorimeter. After consideration of several designs, the SiPM-on-tile technology used in the CALICE AHCAL prototypes [14] was identified as the most cost-effective solution that also provided adequate performance for particle tracking and identification.

The SiPM-on-tile technology utilizes direct readout of the light from the scintillator tile by a SiPM that collects the light through a dimple in the surface of the tile. The dimple provides mechanical space for the mounting of the SiPM and improves the uniformity of response across the tile by reducing the response for particles which pass very near the SiPM compared with those at a larger distance [15]. A drawing of a typical square tile developed by CALICE is shown in Fig. 2.11, and a photograph is shown in Fig. 2.12.

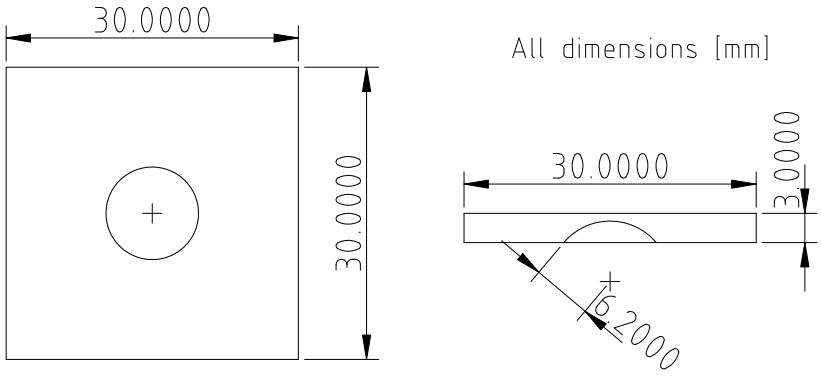


Figure 2.11: Parameter drawing of typical square tiles developed by the CALICE Collaboration. Tiles for the CMS endcap calorimeter will be ring-sections rather than squares due to the geometry of the endcap.

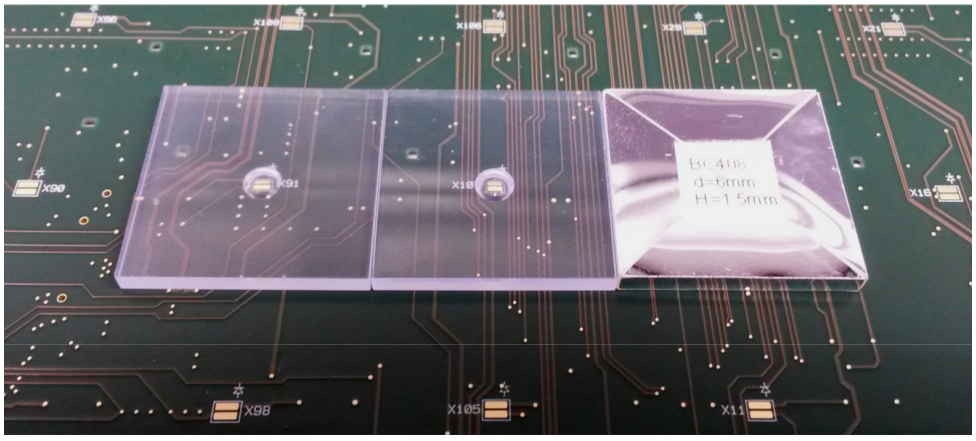


Figure 2.12: Example of three CALICE $3 \times 3\,\mathrm{cm}^2$ scintillator tiles mounted on a PCB that holds one SiPM per tile. The left two scintillators are unwrapped to show the SiPM within the small dome at the centre of the tile, while right-most tile is wrapped with reflective foil.

A MIP traversing an undamaged $3\,\mathrm{cm} \times 3\,\mathrm{cm} \times 3\,\mathrm{mm}$ tile has been shown to generate >20 photoelectrons by the CALICE Collaboration. Our test beam measurements, described in Section 7.3.2.2, confirm this performance and demonstrate several important geometric relationships: (i) the magnitude of the MIP signal is proportional to $1/\sqrt{A_{\mathrm{tile}}}$, the inverse square root

of the tile surface area, so, the smaller the tile the larger the yield of photoelectrons from a MIP; and (ii) the photoelectron yield is also shown to increase linearly with area of the SiPM. Our measurements also show a doubling of photoelectron yield when a VM2000 enhanced specular reflector (ESR) film is used as the scintillator tile wrapping material (more details are given in Section 7.3.2.3).

The arrangement of scintillator tiles takes advantage of this geometric effect. They are arranged in an $r - \phi$ grid where the radial dimension is taken to be the same as the azimuthal width at the centre of each tile. As a consequence, the smaller sized tiles at small radii give larger signals where the radiation damage to the scintillator and the SiPM noise will be largest. At larger radii, where the damage and the noise are smaller, the cells are larger to reduce the channel count and the overall calorimeter cost.

The $\Delta\phi$ angular size of the cells is chosen to provide an even matching to the 5° cells of the barrel hadron calorimeter and to keep the size of the scintillator cell similar in physical size to the silicon trigger cells (sums of four $1\,\mathrm{cm}^2$ cells) at the boundary between the scintillator and silicon. This matching minimizes the impact of the inevitable disjunction between the two regions for reconstruction algorithms. The first four scintillator-containing layers (in the front portion of CE-H) use $\Delta\phi = 1^\circ$ cells while the rest of the scintillator cells use $\Delta\phi = 1.25^\circ$ cells.

The boundary between the silicon and scintillator sections of the calorimeter is determined primarily by the requirement to maintain a MIP signal-to-noise ratio of at least 5 through the life of the detector. The boundary is somewhat extended, in radius, relative to the minimum possible area to provide silicon coverage for $|\eta| > 2.4$. The boundary is also tuned such that varying sizes of silicon sensors are not required at the boundary – only hexagon or half-hexagon sensors are needed. The layout of an example layer of CE-H is shown in Fig. 2.13. For the foreseen operating temperature of $-30\,^\circ\mathrm{C}$ the noise rate will be a factor of 30 lower than for operation at room temperature ($23\,^\circ\mathrm{C}$). A map of the projected MIP signal-to-noise ratio after $3000\,\mathrm{fb}^{-1}$ is shown in Fig. 2.14. The properties of the SiPMs that will be coupled with the scintillator tiles are given in Table 2.5.

2.4 Scintillator tile-modules

As described in Section 2.3, the hadron calorimeter will use scintillator as the active material in regions where the integrated dose is low-enough ($<3\,\mathrm{kGy}$) for the scintillator, and the fluence is limited to $8\times10^{13}\,\mathrm{n_{eq}/cm^2}$ to retain good overall performance over the whole life of the HL-LHC, with margin. As a result, the region $|\eta| > 2.4$ is covered exclusively by silicon sensors. An added advantage is that this also enables a good efficiency for muon identification in this region, where the ME1/1 chambers are not available.

The scintillator is formed into small tiles and scintillation light is directly read out by a SiPM that is optically coupled through a small "dimple" in the centre of one face of the tile. The SiPMs are mounted on a printed circuit board which is then mated with the appropriate tiles. This system is illustrated in Fig. 2.12.

To match the geometry of the CMS endcap, the scintillator cells will be arranged in an r, ϕ grid. As a result, the cells nearer the beam line will be significantly smaller ($4\,\mathrm{cm}^2$) than those at the outer edge ($32\,\mathrm{cm}^2$). The area instrumented with scintillator is subdivided into tile-modules, which form annular segments of up to $40 \times 50\,\mathrm{cm}^2$ in area. Each tile-module is a stack consisting of a tileboard and the scintillator tiles.

For the scintillator, different manufacturing and assembly techniques are under consideration:

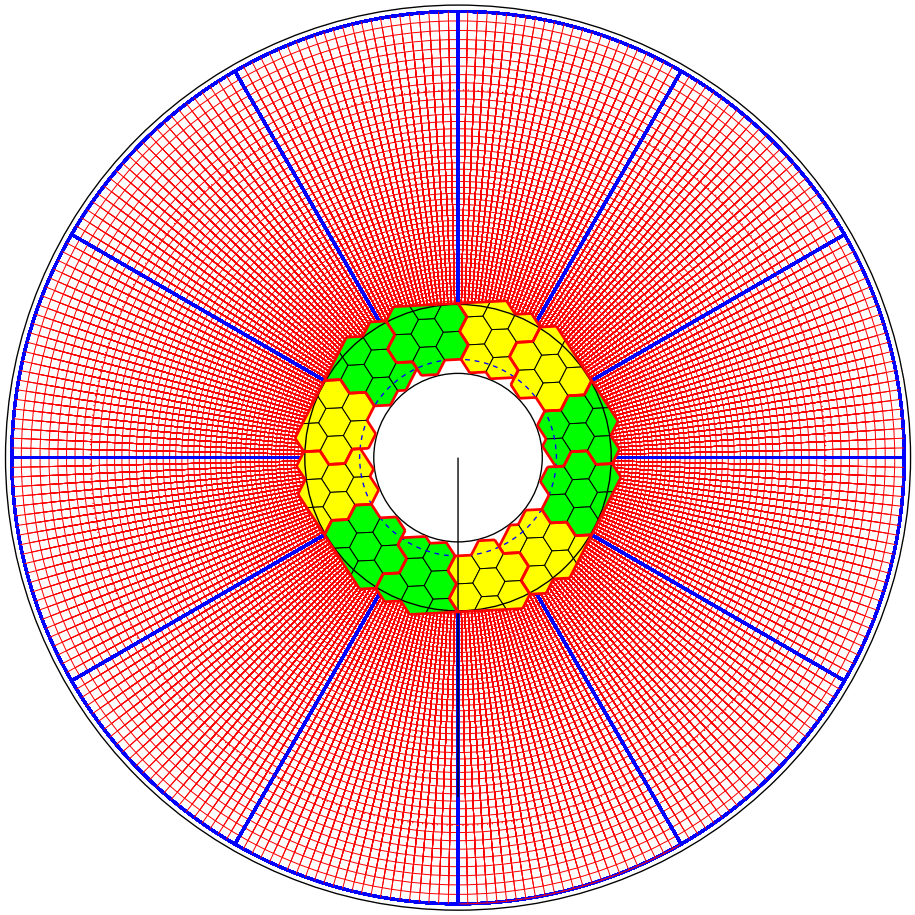


Figure 2.13: Layout of wafers and tiles in a layer where both are present: the 22nd layer of CE-H.

cast and machined or injection-moulded, as individual tiles or multiple tile units (megatiles). In addition, the reflective coating with paint or foil is still subject to optimisation by prototyping. Both approaches build on experience: painted megatiles are used in the CMS HCAL, and foil-wrapped individual tiles have been developed in the CALICE framework. The final choice, to be made in 2019, will be driven by cost, performance, and ease of assembly considerations, e.g. the amount of light yield loss and noise increase due to irradiation, and by thermal-mechanical considerations. The thermal expansion coefficient of plastic scintillator is $78 \times 10^{-6}\,\mathrm{K}^{-1}$ and thus 5 times larger than that of the copper cooling plate and the PCB. For a temperature difference of 70 K (assembly at 30 °C and operation at −40 °C) this leads to millimeter-size mismatches and in practice limits the maximum size of megatiles to about 20 cm. In the inner parts with small cells, such a size still represents a significant reduction in the number of parts to be handled. On the other hand individual tiles are better suited for assembly using standard pick-and-place tools.

The tileboard holds the SiPMs, the front-end electronics, LEDs and associated driving circuitry, low voltage regulators and the connectivity to the motherboard that is situated at the outer periphery of the cassette.

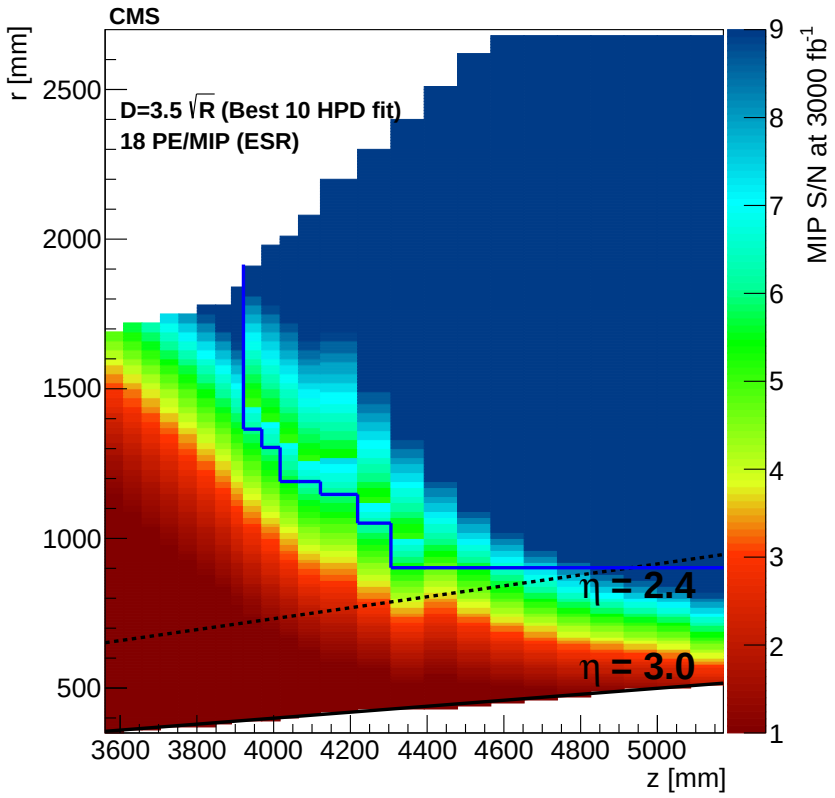


Figure 2.14: Signal-to-noise ratio for a MIP, after an integrated luminosity of 3000 fb^{-1}, shown as a two-dimensional map in r and z. The region, at larger z and r, in which SiPMs mounted on scintillator tiles can provide $S/N(\mathrm{MIP}) > 5$ after 3000 fb^{-1} is outlined.

Chapter 3

Electronics and electrical systems

3.1 On-detector electronics

3.1.1 Requirements

The front-end (FE) electronics measures and digitizes the charge deposited in the silicon sensors pads or generated in the SiPMs, provides a high precision measurement of the time of arrival (ToA) of the pulses, and transmits the digitized data to the back-end (BE) electronics located in the service cavern. It also computes, at every bunch crossing, digital sums of neighbouring cells (2×2 cells in the case of the 1 cm^2 pads silicon sensors, 3 cells in the case of the 0.5 cm^2 sensors, and 2 cells in the case of scintillator tiles) that are transmitted to the trigger BE electronics to build trigger primitives.

There are stringent requirements on the performance of the FE electronics for the silicon part of the detector, driven by the electromagnetic section:

- low noise and large dynamic range, from $\approx$0.2 fC to 10 pC[1], equivalent to 16 bits, to be able, at the lower end, to measure the MIP in the silicon sensors and, at the higher end, to record high-energy deposits from electromagnetic showers at shower maximum. The electronics noise must be $\lesssim$2500 e^- for a 65 pF capacitance pad to allow MIP visibility during the whole operation;
- integral linearity better than 1% over the full range;
- ability to provide timing information with a precision better than 100 ps for pulses above $\approx$12 fC, corresponding to about 3 MIPs in the 300 μm silicon sensors (the use of timing information is described in Sections 5.5 and 10.4);
- fast shaping time (peaking-time $\lesssim$20 ns) to minimize the out-of-time pileup: the pulse should have dropped to less than 20% by the next crossing (i.e. 25 ns after the peaking time);
- on-detector digitization and data processing for zero suppression, for linearization and summing of the trigger data;
- maximum latency of 36 bunch crossings for the trigger primitives at the output of the detector (detailed latency estimates are given in Table 8.7);
- buffering of the data to accommodate the 12.5 μs latency of the L1 trigger;
- high-speed readout links to interface with the 10 Gb/s low power GBT (LpGBT) serialiser [16];

[1]The range $\approx$0.2 fC to 10 pC is guided by the fact that at start of operation, in 120 μm sensors, a MIP deposits about 1.5 fC, corresponding to 8 ADC bins; at the high end of the range, the maximum energy deposited in a cell by a 1.5 TeV photon shower is about 6000 MIP.

- low power budget ($\lesssim$20 mW per channel for the FE electronics, roughly limited by the cooling power that can be installed);
- compensation of the leakage current which will develop in the silicon devices after irradiation;
- compatibility with negative and positive inputs, to be able to read both p-on-n and n-on-p sensors;
- high radiation resistance ($\gtrsim$1.5 MGy and 1×10^{16} n_{eq}/cm^2), and designed to function despite single event upsets (SEU compliance).

Furthermore, it is desirable to use very similar if not identical FE electronics for the readout of the silicon sensors and of the SiPMs, despite the large difference in the collected charge of the two devices. The collected charge from the SiPM is $\approx$1.7 pC/MIP.

3.1.2 Readout chain

The architecture of the front-end electronics chain is shown in Fig. 3.1.

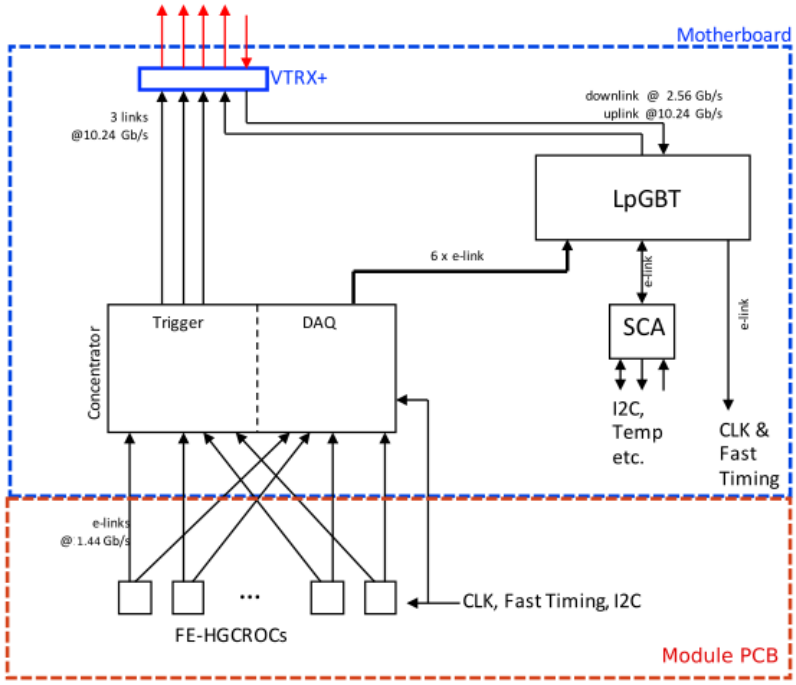


Figure 3.1: HGCAL readout architecture.

The pads of the silicon sensors or the SiPMs are connected to the input of the front-end ASIC, HGCROC (Fig. 3.2), which measures the charge and the time of arrival at 40 MHz frequency. Details of the components and performance of the HGCROC test vehicles are given in Section 8.1 and we describe here only the main features.

For the Si sensors, after pre-amplification, the charge measurement for charges up to 100–150 fC is performed with a low-power 10-bit SAR ADC in 130 nm technology, and by using the time-over-threshold (ToT) technique for charges above $\approx$50 fC (in order to have some overlap) with a 12-bit TDC (50 ps bin size, 200 ns full range; details are given in Section 8.1.2). The ToT dynamic range extends the measurement up to 10 pC and its bin size (2.5 fC) corresponds to less than 4% of the charge at the switch between ADC and ToT. The pre-amplified pulse goes also to

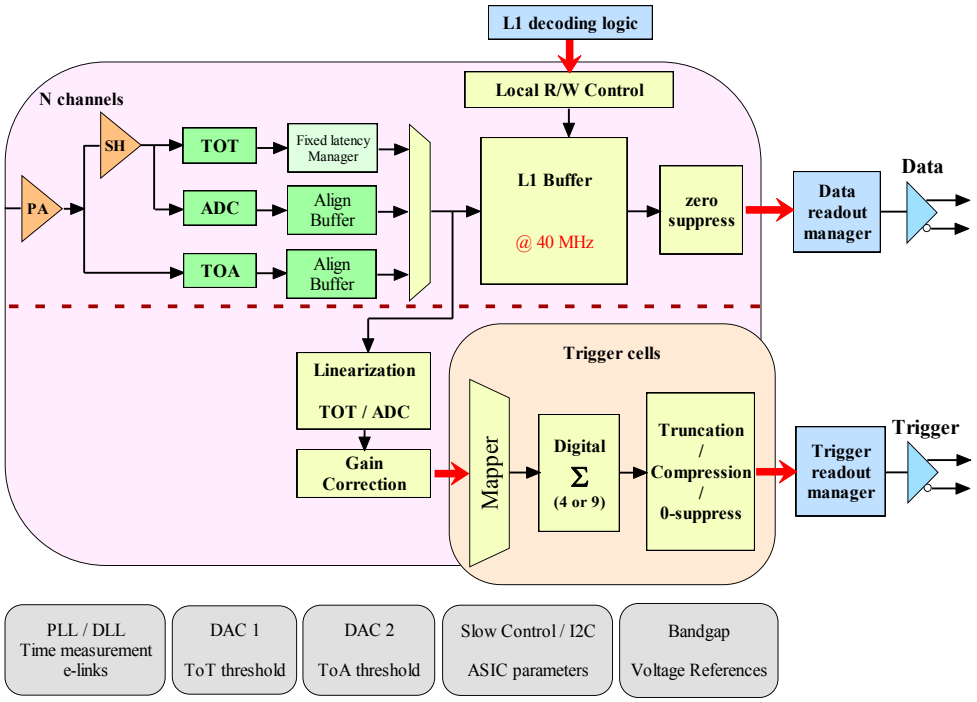


Figure 3.2: Schematic diagram of the HGCROC chip.

a discriminator and the ToA is measured with a 10-bit TDC (<25 ps step LSB and 25 ns full range). After alignment, the ADC, ToT and ToA data are stored at 40 MHz in a 512 columns circular buffer, waiting for a potential L1 trigger accept. Following which, the data are zero-suppressed and transmitted to a concentrator ASIC, via 1.44 Gb/s electrical links.

In parallel, ADC and ToT data are recombined, linearised, and calibrated to provide a single charge measurement which is used to build trigger sums. Sums of 4 or 9 adjacent channels (depending on the sensor granularity) are formed and the result is compressed with an 8-bit format. These sums are transmitted for every bunch crossing from the HGCROC to the concentrator ASIC, via separate 1.44 Gb/s electrical links to be used for the formation of HGCAL trigger primitives in the BE electronics. Alternatively, it would be possible to select the trigger cells to be read out in the HGCROC and only transmit those above threshold, reducing the data rate out of the ASIC. However, this would increase the complexity of the HGCROC design and would also remove the flexibility of having all the trigger cells from adjacent HGCROCs available in the concentrator, where a more sophisticated selection may be possible. A decision will be made in Q1 2018.

The estimated latency of 36 bunch crossings (BX) for the trigger data within the FE electronics comprises 16 BX in the HGCROC and 20 BX in the concentrator. The HGCROC latency includes a delay of up to 10 BX to allow for the ToT value for the largest expected signals, with the remaining 6 BX being the estimated time for processing and serialisation of the data for transmission to the concentrator. The concentrator latency is dominated by a buffer with a depth of 12 BX which is required to reduce trigger data truncation to a low level, particularly

due to occasional large physics events. Again the remaining 8 BX is the estimate of the time needed for data processing and serialisation. The data transmission on the LpGBT links from the concentrator to the BE electronics is estimated to take a further 24 BX.

The ToT front-end results in dead time for sensor cells with signals larger than than 250 fC. The length of the dead time depends on the signal size, and extends to 250 ns for a cell with 10 pC deposited. In the TP the dead time was investigated using a simulation of a sequence of ten bunch crossings for instantaneous luminosities corresponding to means of 140 and 200 interactions per bunch crossing. The resulting probability for a sensor pad to be busy during any particular bunch crossing, as function of layer, is found to be well below 1×10^{-4} at all values of η.

The HGCROC comprises 78 channels (72 reading out standard cells, 2 reading out calibration cells, and 4 channels not connected to any sensor cells for common-mode noise estimation) and is designed in a radiation-hard 130 nm CMOS technology. Owing to its 1.5 V rail, this technology is better adapted for high dynamic range analogue electronics than that with the 65 nm feature size. The HGCROC is bonded through a ball-grid array to the hexaboard.

The input of the HGCROC preamplifier (an NMOS transistor) is naturally adapted to a negative input polarity. A shift of the preamp voltage allows positive polarity pulses to be input. Since the input is DC-coupled, the leakage current of each channel has to be compensated: an adjustable current is provided through a DAC with a maximum value of 20 μA.

The power consumption of the HGCROC is estimated to be $\approx$14 mW per channel.

The readout chips (HGCROC), the linear voltage regulators (LVR), and any other relevant passive and service components are mounted on a hexagonal module readout PCB, the hexaboard, which is glued onto the silicon sensor, as shown in Fig. 3.3 (top). Imprinted on the hexaboard are the pads for connections to the next board in the chain, labeled the motherboard in Fig. 3.3 (middle). Through these connections is supplied the LV, and the received and transmitted signals and data. The hexaboard also has drilled in it the oblong holes through which the sensor cells are wire bonded to pads on ledges in the hexaboard as shown in Fig. 3.4.

The next level of the FE electronics, which includes the concentrator, is located on motherboards, Fig. 3.3 (middle), that sit 1.6 mm above the hexaboards. A motherboard serves from one to six modules (depending on the occupancy of the sensors) and is connected to the hexaboards via miniature compression connectors, e.g. compression connectors from Samtec Z-ray series. The surfaces which include the components of the hexaboard and of the motherboard face each other in order to reduce the overall thickness. Figure 3.3 shows plan views of the modules plus the motherboard layout.

The motherboard performs many functions for the silicon sensors:

- It aggregates, formats and serialises the data (read at the L1 frequency of up to 750 kHz) in the concentrator ASIC.
- It receives and distributes the fast control signals (clock, L1-accept, reset, bunch crossing count) and all the controls needed to configure the FE electronics, and conversely send data and control information to the data acquistion (DAQ) modules located in the service cavern. To perform these functions, each motherboard includes a bidirectional optical link using a LpGBT, coupled to a GBT-SCA (slow control adapter) ASIC [17]. The GBT-SCA includes a multiplexed ADC with 31 inputs that is used to perform and transmit the monitoring information of environmental parameters (temperatures, voltages, etc.)

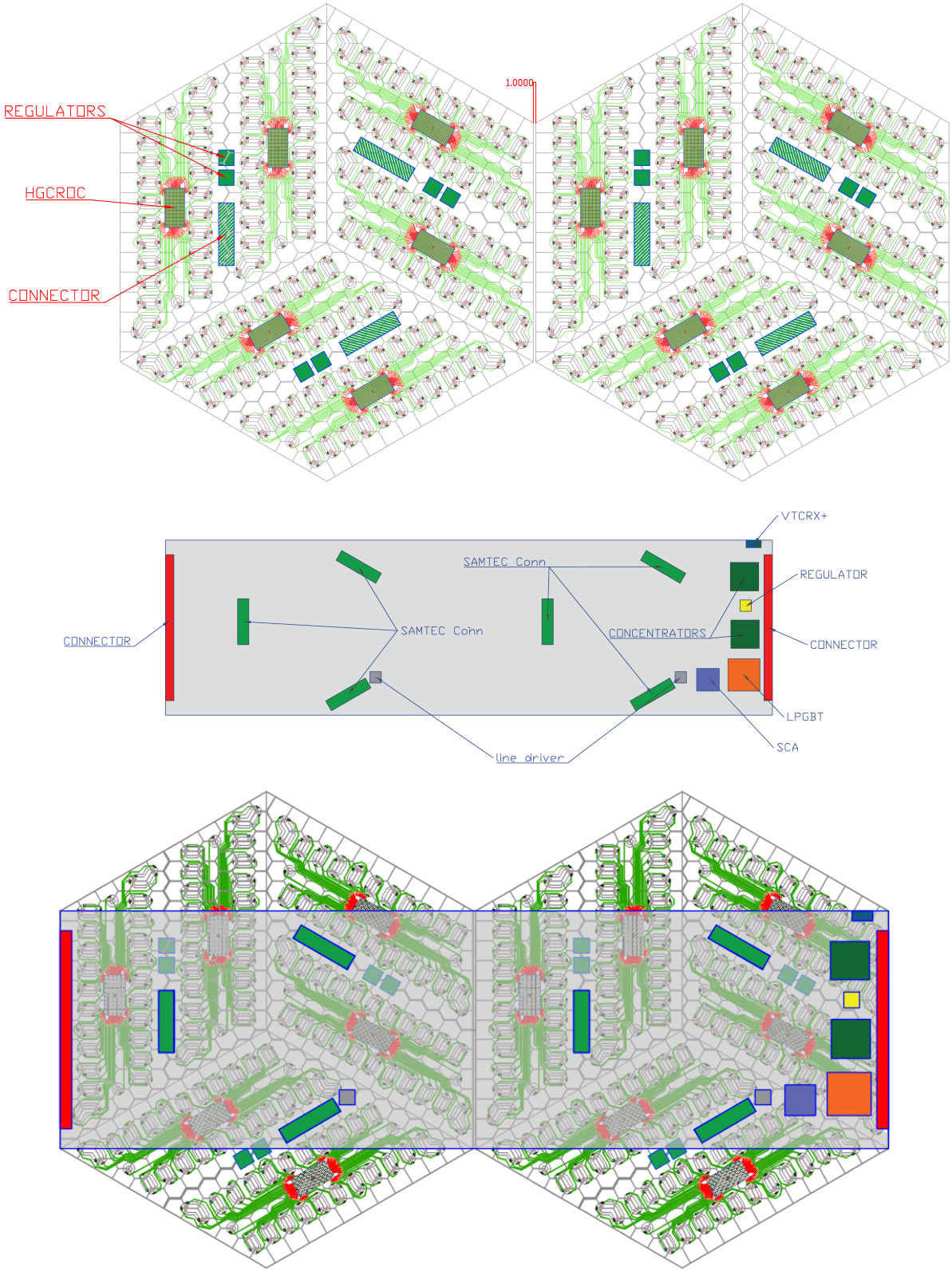


Figure 3.3: Plan views of two modules plus motherboard assembly; (top) two hexaboards, (middle) motherboard (components face down), and (bottom) motherboard connected to the two hexaboards.

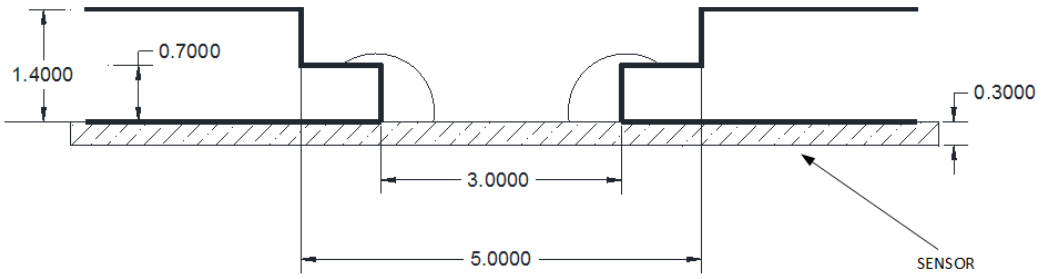


Figure 3.4: Connection between the silicon sensor and the readout board.

- For the trigger path, the concentrator selects the trigger sums of interest, aggregates and formats their data in packets and stores them in a FIFO buffer with a length of 12 bunch crossings. These data are sent within a defined maximum latency to the BE Trigger electronics located in the service cavern, using separate optical links.
- Finally, the motherboard distributes power to the modules. Four copper layers embedded in the motherboard, using almost the full width of the printed circuit, are used as power buses (one for the analogue power, one for the digital power and two for the returns). The power is regulated by low-drop linear voltage regulators, one located on the hexaboard and another on the motherboard.

Unlike silicon diodes, the SiPMs used to read the scintillator tiles have a high gain. On the other hand, the required dynamic range in the scintillators is smaller by an order of magnitude compared to that required for silicon pads located in the electromagnetic section. SiPMs will be read out by a variant of the HGCROC. It is foreseen to adapt the very front-end (preamplifier) for the larger signal level. A current conveyor or a capacitive divider at the input of the FE ASIC will therefore be used to ensure compatibility of the rest of the chain with that for reading out the silicon sensors. Details are given in Section 8.1.2.6.

Scintillator/tile-modules are made in several sizes and geometries to match the geometry of the CE-H. Digitized SiPM signals are transmitted from the HGCROC by 1.44 Gb/s electrical links, using Twinax cables, to a motherboard that is located at the edge of the cassette. The Twinax cables have been shown to be radiation hard in fluences up to $1\times10^{16}\,n_{eq}/cm^2$. The use of optical links to the motherboard is also under consideration.

The function of the motherboard for tile-modules (Section 2.4) is provided by special boards on the periphery. These host a concentrator chip and provide connections to the trigger, DAQ, control signals, and bias voltage (shown in Fig. 3.5). Board-to-board connections bring the bias voltage and control signals from the motherboard to individual tileboards.

The concentrator acts therefore as a hub, receiving up to 72 1.44 Gb/s electrical links from the HGCROCs and serialising the corresponding data, trigger and level-1 accept readout, on several parallel outputs at 10 Gb/s.

All the ASICS used on the motherboard are developed in a radiation-hard technology, qualified for more than 2 MGy, with an SEU-compliant design. The GBT-SCA is in 130 nm technology, while both the LpGBT and concentrator are in 65 nm technology. The optical transmission is performed through the VTRx+ (versatile transceiver plus) optoelectronics transceivers, also located on the motherboard.

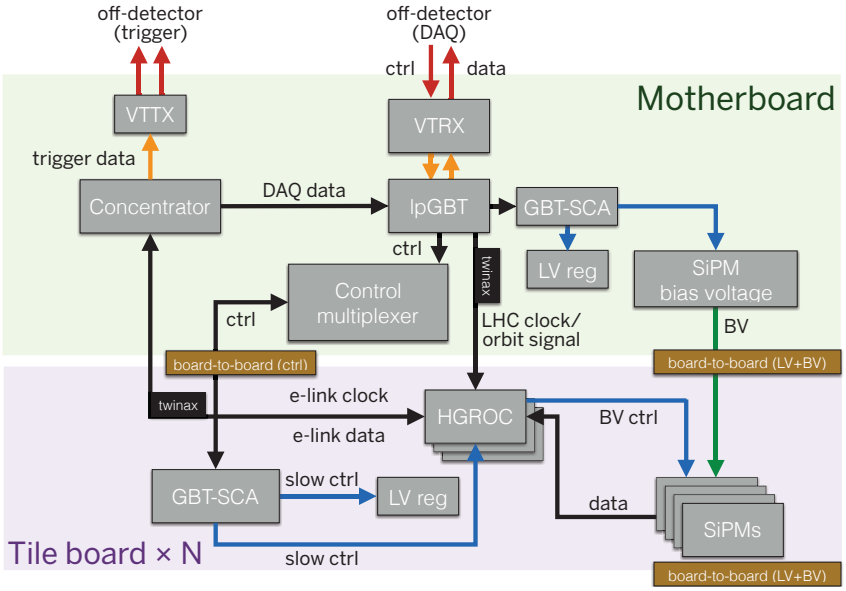


Figure 3.5: Block diagram of the on-detector electronics for the scintillator/SiPM detector layers.

3.1.3 Data rates and data transfer

A single HGCROC serves 70 (78) channels for large (small) cells sensors, respectively. This includes two calibration cells and four channels, not connected to any pad, which are used for common-mode subtraction. The data stored in the HGCROC buffer are read after an L1-accept, with every channel above a threshold of 0.5 MIP being sent to the DAQ system, while channels below 0.5 MIP are normally suppressed. The threshold will be adjustable at the level of ≈0.1 MIP. The fraction of events for which a channel is above threshold is strongly dependent on the position of the channel within the HGCAL, as well as on the pileup rate and, for the 120 μm sensors, to some extent on the noise. Figure 3.6 shows the average fractional channel occupancy resulting from pileup corresponding to a mean of 200 interactions per bunch crossing, with no extra occupancy due to physics events included. The data rates were estimated using the simulation described in Section 10.3.1.2. The occupancy varies from around 60% at high $|\eta|$ in the front layers to less than 5% at the rear outer edge of the silicon sensors, and lower than 1% in the rear scintillator areas.

The format of the channel data is dynamic and includes the ADC value, the ToA value for charges above a 10 fC threshold and the ToT value for charges above a 150 fC threshold, as well as the value of the ADC in the previous bunch crossing if it was above 2.5 MIP. A detailed description of the data format can be found in Section 8.1.7, and is summarized in Fig. 8.24. With the proposed format, the required bandwidth per HGCROC ASIC in the most occupied region (the fine-granularity sensors located at small radius after ≈7 radiation lengths) does not exceed 0.8 Gb/s for a L1 trigger rate of 750 kHz and can therefore be safely transmitted to the concentrator with a single 1.44 Gb/s electrical link.

A single HGCROC sends 16 (8) trigger sums for a large (small) cells sensor respectively, each sum corresponding to an area of ≈4.5 cm^2. All the trigger sums (compressed to 8 bits each) are transmitted at 40 MHz to the concentrator with four (two) 1.44 Gb/s electrical links. In the concentrator, a threshold of 2 transverse MIPs (MIP multiplied by the trigger cell $\sin\theta$, henceforth designated MIP$_\mathrm{T}$) is applied. To avoid possible bias due to the suppression, for example in the computation of $E_\mathrm{T}^\mathrm{miss}$ at the trigger level, the sum of all the cells in the HGCROC

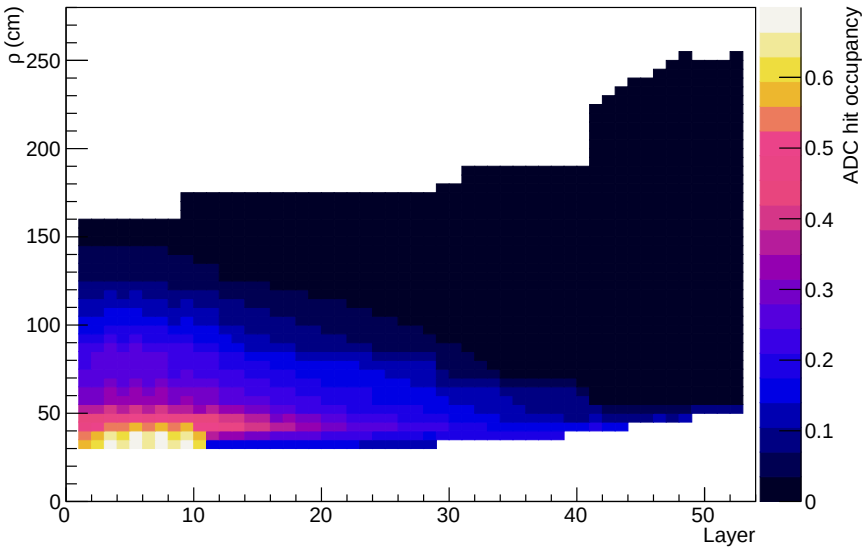


Figure 3.6: Average fractional occupancy per channel, counting channels above a threshold of $E > 0.5$ MIP, as a function of layer and radial distance from the z axis within each layer. The pileup corresponds to an average of 200 interactions per bunch crossing. This figure is based on the simulation geometry described in Section 5.1.2. The detector performance assumed is that expected after an integrated luminosity of 3000 fb^{-1}.

is also transmitted. Figure 3.7 shows the average fractional occupancy for trigger cells resulting from an average pileup of 200 only. As for the DAQ channel data, it is a very strong function of position, varying from around 50% to less than 1%.

The chosen motherboard layout results from an optimisation of the number of different board shapes and the number of optical links, as required by the data rates.

For the data path, the design is such that the output rate of the concentrator will not exceed 7.7 Gb/s, which allows the use of a single LpGBT with the FEC5 error correction protocol (8.96 Gb/s useful bandwidth) and enough bandwidth for the fast control signals. An overhead of $O(10\%)$ must be included in addition to allow for large physics events and also to reduce the rate of buffer overflows in the FE electronics. 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 L1A rate of 750 kHz, this corresponds to an average data rate from the FE electronics of 12 Tb/s.

For the trigger path, in most cases one or two optical links per motherboard are enough, but the concentrator can transmit the trigger data on up to three links for the most busy motherboards. The total average rate of trigger cell data will be 1 Mbit per bunch crossing, or 40 Tb/s.

The HGCROC energy sums are made by summing over all sensor cells read out by a single HGCROC ASIC. Such data are only sent out if at least one trigger cell has fired (i.e. $E_{\text{T}} > 2$ MIP_{T}), thus reducing the additional data load.

The total number of 10 Gb/s LpGBT links from on-detector will be 7944 for DAQ, and 7344 for trigger, reflecting the larger size of the trigger cells and the fact that only alternating layers of the CE-E are used to provide the trigger primitives.

Figure 3.8 shows, as an example, the layout of the motherboards on a 60° cassette of the 8$^{\text{th}}$

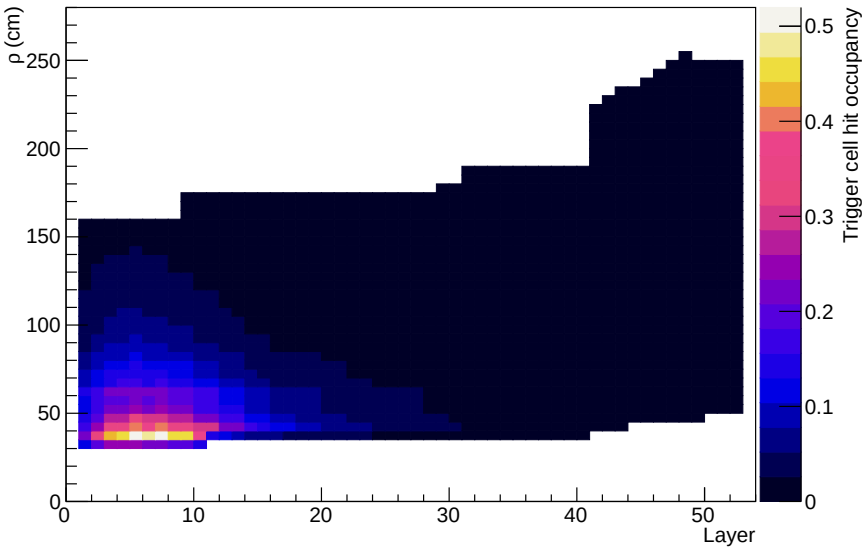


Figure 3.7: Average fractional occupancy per trigger cell, counting trigger cells above a threshold of $E_T > 2\,\mathrm{MIP}_T$, as a function of layer and radial distance from the z axis within each layer. All layers are shown and while this is accurate for the CE-H layers, for the CE-E only alternating layers are used for trigger purposes. The pileup corresponds to an average of 200 interactions per bunch crossing. This figure is based on the simulation geometry described in Section 5.1.2. The detector performance assumed is that expected after an integrated luminosity of 3000 fb^{-1}.

Table 3.1: Number of HGCROC, motherboards, data/control links and trigger links used to readout the complete HGCAL. The electromagnetic and hadronic sections are presented separately.

	CE-E Silicon	CE-H Silicon	CE-H Scintillator	Total
HGCROC	60 324	31 596	8 496	100 416
Motherboards	5 004	2 556	384	7 944
Bidirectional data/control links	5 004	2 556	384	7 944
Trigger links	4 020	2 556	768	7 344

layer of CE-E. This layer has been chosen as it is the layer with one of the largest occupancies. The motherboards are connecting strings of modules which share the same low voltage power source. In the case of the scintillator tiles readout, the number of HGCROCs is small and the motherboards are located at the periphery of the layer, as shown in Fig 3.9 .

Table 3.1 summarises the number of ASICS, motherboards, data/control links and trigger links used respectively for the silicon sensors and scintillator tiles readout of the two HGCAL sections.

3.1.4 Clock distribution

The requirement for the clock distribution system of CMS is that it preserve the timing precision of CMS subsystems (HGCAL, ECAL and the MIP Timing Detector). The target performance is

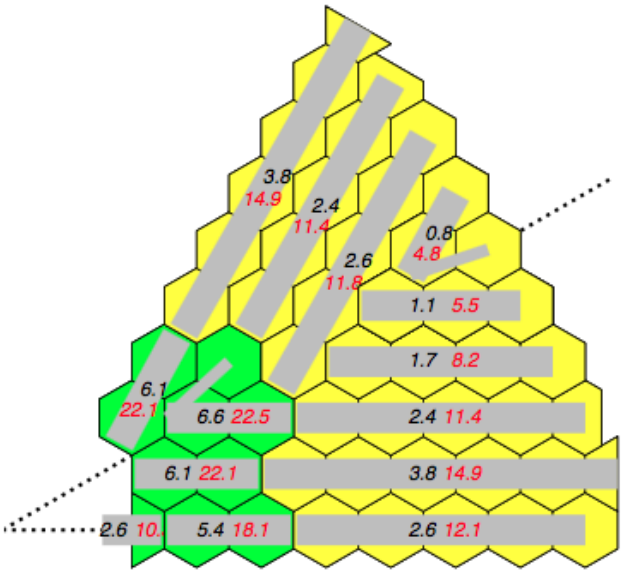


Figure 3.8: Layout of the motherboards on a 60° cassette of the 8th layer of CE-E. The numbers in black (red) give the average bandwidth of the motherboard output for DAQ-data (trigger-data), respectively, in GB/s, in the case of pp collisions at the highest luminosity, corresponding to an average of 200 interactions per bunch crossing.

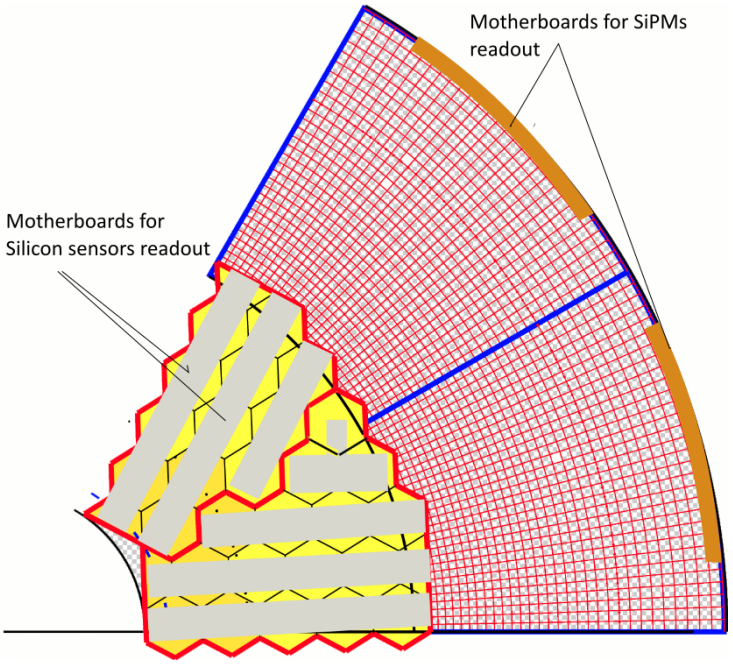


Figure 3.9: Layout of the motherboards for a mixed cassette of the 12th layer of CE-H, including both silicon sensors and scintillator tiles.

to achieve a precision of the clock distribution system at a level of 10–15 ps (RMS jitter) together with short term, long term and detector wide stability.

The existing CMS clock distribution, provided by the current TTC system [18] is already good,

with an RMS jitter of 9 ps (RMS), according to studies carried out recently. Its future evolution, notably with the availability of LpGBT, and eventually passive optical network (PON) components [19], promises achievement of even better performance. The clock phase, once calibrated over the whole detector, will certainly have to be monitored in order to follow its possible variations and obtain knowledge of its long term level of stability.

An R&D programme has been started to identify the driving factors for the design of a precise clock distribution system. More details are given in Section 8.4.

3.2 Data acquisition system

The HGCAL off-detector, or "backend" (BE), electronics consists of the data acquisition (DAQ) system, trigger primitive generator (TPG), detector control system (DCS) and detector safety system (DSS). Both the central DAQ [20] and L1 trigger [21] projects have produced interim documents this year but their TDRs will not be written until 2020 and 2019, respectively. Hence, the DAQ will still be subject to change. The current assumptions for the HGCAL DAQ are described below and the TPG is described in Section 3.3. They receive data from the on-detector electronics through separate links, as described in the previous section, and so are effectively independent systems as far as the off-detector electronics are concerned.

3.2.1 DAQ system requirements

The DAQ system is designed to read out all channels above a threshold of 0.5 MIP at an average pileup of 200 interactions per bunch crossing and a L1-accept (L1A) rate of 750 kHz, with the expected noise levels after 3000 fb^{-1}. In addition, there will be some low rate of channels which are read out when below this threshold for monitoring and calibration purposes. The DAQ system also sends fast and slow control data to the on-detector electronics, and receives the corresponding fast and slow readout data, as shown in Fig. 3.1.

These data are transmitted to and from the on-detector electronics on 7944 bidirectional links, as described in Section 3.1. The DAQ system needs to receive and process these data, and retransmit them to and from the relevant location, i.e. the central DAQ system, timing and clock distribution system (TCDS), or DCS system (Section 3.5). The data volume related to sensor channel readout, which totals 15 Tb/s, completely dominates over the fast and slow readout data.

3.2.2 Hardware description

The physical format of the DAQ boards will need to be consistent with the design of the central DAQ interface. At the time of writing, the assumption is that these systems will be implemented in advanced telecommunications computing architecture (ATCA) [22] format, to a CMS-defined specification. Hence, the HGCAL DAQ system is currently assumed to consist of ATCA boards housed in ATCA crates, with DAQ and timing hub (DTH) ATCA boards also in these crates to provide the central DAQ and timing systems interface.

The exact specifications of the HGCAL DAQ system will be optimised when the central DAQ DTH board is completely specified. For the purposes of this TDR, the costed system has been assumed to consist of generic ATCA boards containing two high-end FPGAs and with sufficient bandwidth to drive 96 links in both directions to and from the on-detector electronics, as well as providing a further 12 optical links to and from a DTH in the same crate, timing and control data transmission over the ATCA backplane fabric, and ethernet connectivity for the DCS system. Given the number of links from the on-detector electronics, this means at least 83

DAQ boards are needed. To achieve this cabling density will require use of custom FO patch panels. The 96 links per board to and from the on-detector electronics will run at 10 Gb/s using LpGBT format, with the down-links to the detector acting as the fast control, and part of the slow control, interface, while the up-links mainly carry the DAQ data, but also the other data specified above (shown in Fig. 3.1). The 12 links to the DTH board will run at 16 Gb/s. The FPGAs on these boards will handle and process the data, with the required buffering, event building, front-end emulation, monitoring, etc., implemented in firmware.

Given the average event size of 2.5 MByte per event, then each ATCA board will receive an average of 240 kbit per event, corresponding to 180 Gb/s at an L1A rate of 750 kHz. This determines the need for the 12 links per HGCAL DAQ board to the DTH board. The current assumption for the maximum DTH board input and output data rates is 1.2 Tb/s and there will be one DTH board for every six DAQ boards. Hence, the system will have twelve DAQ boards and two DTH boards per ATCA crate (shown in Fig. 8.33). This requires the DAQ system to occupy seven crates in four racks. The major component counts are summarised in Table 3.2.

Table 3.2: Counts of components in the HGCAL DAQ system.

Component	Per unit	Total in system
LpGBT input links	96/board	7 944
HGCAL boards	12/crate	84
DTH boards	2/crate	14
Crates	2/rack	7
Racks		4

3.3 Trigger primitive generator

As stated in Section 3.2, the central L1T system is not yet fully defined, so there are still significant uncertainties in the implementation of the off-detector trigger primitive generator (TPG) system. The description below is of a system which fulfills the requirements and is consistent with current assumptions for the central DAQ and L1T interfaces.

3.3.1 TPG system requirements

The input data to the TPG system are the selected trigger cells and the HGCROC energy sums generated on-detector (see Section 3.1). The system is designed to operate with a trigger cell selection threshold of at least 2 $\mathrm{MIP_T}$ at an average pileup of 200 and with the expected noise levels after 3000 fb^{-1}. The average fractional occupancy per trigger cell is shown in Fig. 3.7. The input data corresponding to a given BX will arrive at the TPG up to 1.5 μs after the BX.

The TPG output for each endcap will be a list of three-dimensional (3D) clusters reconstructed from the trigger cell data, and an energy map consisting of the summed energy in η, ϕ bins. The output data need to be received by the central L1T system within a latency of 5 μs after the relevant bunch crossing, so the TPG has a total available time of 3.5 μs to complete all processing and data transmission.

The TPG must also provide event data to the central DAQ system via a DTH board. These data will be used for diagnostics and by the HLT.

3.3.2 Data volumes

As mentioned earlier, each endcap will have 52 sensitive layers, with 28 in the CE-E section and the remaining 24 in the CE-H section. All the latter will contribute data to the TPG, but only half of the CE-E section layers will be used. There are no links to the detector from the TPG, since all the on-detector electronics (including the parts producing the TPG raw data) are configured and controlled through the DAQ system. The data volumes for the input links are given in Section 3.1.

The TPG output data are the HGCROC energy sums and the 3D clusters. The 3D clusters will have an average size of around 200 bits per cluster (as shown in Section 8.3.2.3). The bandwidth which would be required to transmit all 3D clusters is prohibitive, so some selection needs to be done. The number of clusters is mainly dominated by pileup, and their transverse energy spectrum falls steeply, so the bandwidth required is a strong function of the threshold applied. Since the clusters are the main input from the HGCAL to the particle flow algorithm in the L1T correlator, it is important not to cut out clusters which may be matched to tracks. The L1 tracking threshold will be around a transverse momentum of 2 or 3 GeV. However, the track efficiency has a gradual turn-on with a nominal 2 GeV threshold. Hence a 3D cluster selection of about $E_{\mathrm{T}} = 1.0\,\mathrm{GeV}$ would be appropriate (as can be seen in Fig. 8.41). Studies of $\mathrm{t\bar{t}}$ events with 200 pileup show that this cut results in up to 200 clusters per endcap. Hence, to be able to transmit this maximum case to the L1T correlator, a bandwidth of around 40 kbit per endcap for the 3D cluster data will be required.

A total number of around 1100 energy map bins per endcap is assumed. Each bin will require 16 bits; this will be a 12-bit transverse energy value and a 4-bit fraction for the amount of this transverse energy in the CE-E section. With an LSB for the transverse energy of 100 MeV, this would allow a reasonable precision of around 10% for track-energy matching in the correlator even for tracks with transverse momenta down to 1 GeV, and would have a full range of 400 GeV. The total size of the energy map for each endcap would then be around 20 kbit.

With the above assumptions, then the total bandwidth per BX to the central L1T would be around 60 kbit per endcap, or 120 kbit total. The financial implications of varying this bandwidth in either direction are small for the HGCAL TPG in the current design, as there is extra output capacity and so the incremental cost is mainly in any extra fibre optic cables between the TPG and the L1T correlator.

3.3.3 Algorithm and implementation

The TPG system has to implement two tasks; form 3D clusters from trigger cells, and form the overall energy map from the coarse granularity HGCROC energy sums. The current assumption is that the TPG will perform these tasks in two stages.

The first stage will consider each layer separately. It will form two-dimensional (2D) clusters from the trigger cells and combine all the HGCROC sums into a single η, ϕ energy map for the particular layer being processed. The second TPG stage will then combine the 2D clusters in depth to form 3D clusters. It will also combine all the single-layer energy maps, with an appropriate weighting, into the complete transverse energy map. The currently assumed TPG architecture uses time multiplexing in the second stage to enable all the 2D clusters and energy maps for a bunch crossing to be transmitted to one location. A time multiplexing of 24 should be sufficient for this purpose.

Only a preliminary implementation of the VHDL for the 2D clustering has been done so far, but this indicates that the implementation does not present a problem and that the TPG will have a

latency within the 3.5 μs requirement with the assumed architecture. This latency includes the time (up to 600 ns) added by the time-multiplexing.

The completed energy maps and 3D clusters form the trigger primitives which are transmitted from the TPG to the central L1T system. If the L1T system is not time multiplexed, a demultiplexing step between the two will need to be implemented. This will add a third stage to the TPG, with extra cost and latency implications.

3.3.4 Hardware description

For the purposes of this document, the TPG is assumed to be implemented using the same generic family of ATCA boards as used for the DAQ system (Section 3.2) but with one larger FPGA rather than the two used for the DAQ boards. The two endcaps will be processed independently using two identical copies of the hardware. Figure 3.10 shows a schematic layout of each endcap subsystem.

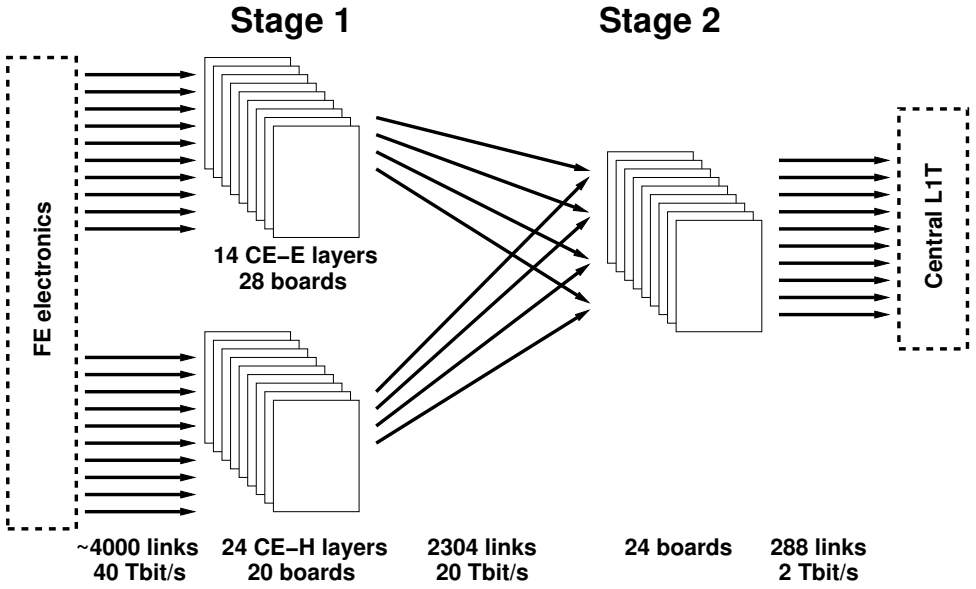


Figure 3.10: Overview of the major hardware components of the TPG for one of the two endcaps. The other endcap is identical in structure.

In the first stage of the TPG, each board will receive up to 96 LpGBT links from on-detector. All the links from each layer will be received by one, or two, boards. The layers in the CE-E section will have more than 96 links per layer, so the data from each CE-E layer will be shared across two boards. The CE-H layer data will fit into one board, although the last eight CE-H layers have only 48 links each and so two layers can be processed in one board. The number of boards needed is simply determined by the number of layers sending trigger data and the number of boards per layer. The above scheme will require a total of 48 boards per endcap and is shown in Fig. 3.10. 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 links per endcap running at 16 Gb/s.

The second stage consists of 24 boards per endcap, each providing a processing node for a 24-fold time multiplexed system. Each board processes one out of every 24 bunch crossings and receives all the 2D clusters from every layer of that bunch crossing. The boards combine the 2D clusters in depth to form 3D clusters and the energy map data per layer will be combined with an appropriate weighting to give a total incoming transverse energy map. The boards see

the data from the whole endcap, allowing maximal flexibility for 3D clustering. Each of the 24 boards in the second stage will then transmit the output data over twelve 16 Gb/s links to the central L1T system, requiring 288 links per endcap.

The TPG crates are separate from the DAQ crates. The TPG is also implemented in ATCA boards as it has to interface to the central DAQ system for its readout. The TPG crates therefore also require a DTH in the ATCA crate and the TPG boards send their event data to this (shown in Fig. 8.43). The event data for the TPG will consist of the board inputs, outputs and some intermediate values, but only the outputs,i.e. the trigger primitives, be read out for every event as these will be used as seeds in the HLT. The maximum data volume per event will be around 1.2 MByte, with the average substantially smaller at around 0.2 MByte. The latency buffer for these data must be implemented on the TPG boards themselves. Hence, the ATCA crates for the TPG can contain 12 TPG boards for which one DTH board is sufficient to handle the rate. The whole system will require 12 ATCA crates and hence 6 racks. The major component counts are summarised in Table 3.3.

Table 3.3: Counts of components in the HGCAL TPG system.

Stage	Component	Per unit	Per endcap	Total in system
Stage 1	LpGBT input links	$\leq$96/board	3672	7344
	16 Gb/s output links	$\leq$72/board	2304	4608
	TPG boards	12/crate	48	96
	DTH boards	1/crate	4	8
	Crates	2/rack	4	8
	Racks		2	4
Stage 2	16 Gb/s input links	96/board	2304	4608
	16 Gb/s output links	12/board	288	576
	TPG boards	12/crate	24	48
	DTH boards	1/crate	2	4
	Crates	2/rack	2	4
	Racks		1	2

3.4 Powering: low voltage and high voltage

The total power required for the readout electronics inside HGCAL is $\approx$110 kW per endcap. The breakdown is shown in Table 9.7. The FE electronics requires three voltages: 1.5 V for the analogue electronics, 1.2 V for the digital and 2.5 V for the optical transmission. The power required for optical transmission is small and can be generated from the 1.5 V rail by voltage pumping.

The power is fed to the silicon modules through a line of motherboards which builds a kind of "busbar". The motherboards include connectors for connecting them together, and to connect the radially outermost motherboard to the power cables. The width ($\approx$10 cm) and thickness ($\approx$35 μm) of the power layers in the motherboards are such that the average power loss is $\leq$10% in the motherboard, and the voltage drop is $\leq$0.15 V. The LVRs filter and stabilise the input voltages for each module. All the modules attached to a line of motherboards are fed in parallel, and therefore the LVRs include a switch, acting as a fuse, which allows disconnection of a faulty module. More details are given in Section 8.5.1.

In the baseline design, both 1.5 and 1.2 V are provided by DC-DC converters installed directly

behind CE-H, on the outer periphery, just outside the cold volume. There are two options under study for these converters: (i) commercial 20 A converters that need to be validated for radiation, and (ii) converters based on the FEASTMP, a radiation-hard and magnetic field tolerant DC-DC converter developed for other detectors in CMS such as the Phase-2 outer Tracker [23]. Several of these DC-DC converters would be put in parallel, in order to provide the required current.

The power cables linking the DC-DC converters to the cassettes will run at the periphery of the layers. These cables run therefore in the cold region of HGCAL and cross the thermal boundary. Their maximum length is 2.8 m and the power loss in these cables represent $\approx$10% of the total power.

As an alternative, the DC-DC converters could be located inside HGCAL, directly on the cassettes. This would have the advantage of substantially reducing the cross section of the cables at the periphery. A drawback is that it increases the power dissipated in the cold region. Furthermore, the integration of these converters on the cassettes may increase the size of the inactive regions.

These schemes will be carefully studied with models before making the final choice. More details on the power distribution are given in Section 8.5. The DC-DC converters are fed with 12 V power supplies which can be located either in the towers attached to the CMS endcaps (and therefore moving with them) 14 to 24 m away, or further away in the collision hall (UXC) or even in the underground service cavern (USC). In the first case, radiation and magnetic field tolerant power supplies, similar to the ones used in the present CMS, will be required.

The bias supplies for the silicon sensors will be located in the services cavern. Current commercial power supplies, providing the required 800 V, are rated up to 10 mA. The leakage current of the modules varies strongly with their position inside HGCAL. The most irradiated ones, at the inner radius of layer 8, are expected to draw $\approx$4 mA after 3000 fb^{-1}, but most of them will draw less than 1 mA; this will allow grouping several sensors on the same supply in most of the cases. To allow the disconnection of a faulty sensor, it is planned to distribute the bias among the sensors sharing the same supply through a patch panel with one separate bias wire per sensor and common return wires. This scheme has been successfully used in the CMS preshower, with a cable holding 40 bias wires and 10 return lines. This flexible scheme would even allow staging some of the bias supplies at the start of the exploitation, when the leakage current will be low.

3.5 Detector control and safety systems

The HGCAL control and safety systems will follow the same principles as for the tracker detector and will be developed in strong synergy with the tracker group. The main features are summarized below.

The hardware of the safety system, CE-SS, will be based on industrial programmable logical controllers (PLC). The CE-SS will constantly monitor the temperature and humidity inside the HGCAL volume, including the interfaces, and will have access to all status information of the CO_2 cooling and dry gas system. Passive temperature sensors (Pt1000 sensors) inside the HGCAL volume will be directly connected to the CE-SS PLCs. A dedicated sniffer system will pump gas from the volumes via several copper pipes towards the service cavern, where the gas is analysed by commercial high-precision dew-point meters. Based on the temperature, humidity, and functional status of the service systems, the CE-SS can interlock individual

power supplies, switch off the gas flow by controlling the input valves, or switch off cooling plants. Table 3.4 shows a basic set of such actions and the possible causes. The details of the CE-SS action matrix are still under study. Unlike the tracker, the mass of the HGCAL prevents a fast increase to room temperature in case of alarm on humidity. Therefore it is of paramount importance that the dry gas system have sufficient capacity to overcome the effect of conceivable leaks in the thermal screen and have sufficient redundancy, including backup by diesel generator and a passive source of dry gas, e.g. boil-off from an LN_2 dewar.

Table 3.4: Preliminary HGCAL DSS action matrix.

Feedback from sensors	**Possible cause**	**CE-SS action**
Electronics temperature increase	Local loss of cooling	Turn off corresponding power supplies
CO_2 plants off	Power cut	Turn off all HGCAL electronics
Increase of humidity inside the cold volume	Thermal screen leak	Send warning to increase dry gas flow to compensate for the leak and, if necessary, switch off the CO_2 cooling plants
Increase of pressure inside the cold volume	Cooling system leak	Switch off the CO_2 cooling plants
Increase of humidity of the injected dry gas	Malfunction of the main dry gas system	Switch to backup dry air source

The CE-SS is a slave of the CMS detector safety system (DSS), having access to information like fire alarm, rack failure, etc. The DSS can tell the CE-SS to switch off the HGCAL and/or its services. All PLCs are supervised by a redundant PLC master system, which is not connected to the network, for security reasons, and are connected to uninterruptible power supplies. Limits on temperature and humidity can be programmed individually, as well as specification of groups of sensors and/or SiPMs to switch off, according to a programmable majority logic. The action matrix is adaptable.

The HGCAL control system, CE-CS, is the main interface to switch on and off the HGCAL in a safe and controlled way, and it provides all status information (power system, cooling system, gas system, temperature, etc.). It will be based on the supervisory control and data acquisition (SCADA) commonly chosen by the LHC experiments for the HL-LHC phase. The CE-CS will monitor all power supplies and all services, e.g. cooling and dry gas, and will have full access to all CE-SS PLC information, e.g. all temperature and humidity readings. The CE-CS will also prevent switch-on in case of non-conforming conditions. The system will also provide alarms and warnings via the common CMS interface to the shift crew. All set-parameters (low- and high voltage settings, trip limits, temperature limits, and alarm levels) will be stored in a configurations database, and all actual conditions will be stored in a conditions database, to allow historic viewing.

Chapter 4

Engineering

4.1 Engineering design overview

A mechanical drawing of the HGCAL is shown in Fig. 1.6. The electromagnetic section extends from $|z| = 3.19$ to 3.53 m, covering a depth of approximately 26 X_0 and 1.7 λ with 28 sampling layers. The hadronic calorimeter extends from $|z| = 3.53$ to 5.14 m, providing an additional depth of about 9.0 λ. The hadronic section is divided into finer and coarser sampling section, each comprising 12 layers and providing depths of 3.3 and 5.7 λ respectively.

The CE-E is made of alternating absorber and silicon active layers. Silicon sensors are assembled into hexagonal modules in which they are sandwiched between WCu baseplates and PCBs that carry the front-end electronics. These modules are mounted on two sides of a 6 mm thick copper plate, in which a small diameter stainless steel tube carries liquid CO_2 to maintain the detector at a temperature of $-30\,^{\circ}$C. The cooling plate is formed as a wedge extending over the full radius of each layer and subtending an azimuthal angle of $60\,^{\circ}$. Groups of 1–5 modules are serviced by a motherboard that contains most of the on-detector digital electronics and carries low-voltage power to the module (Section 3.4). The cooling plate, together with the WCu baseplates on the two sides, provides one absorber sample. The other absorber layer is provided by layers of stainless steel-clad lead plates of the same dimensions as the cooling plate attached above and below the sensors and electronics. This assembly of lead covers, detector modules, and copper cooling plate is called a cassette (shown in Figs. 4.1 and 4.2). Cassettes are, in turn, assembled in a stack of layers which forms the CE-E. The cassettes are more fully described in Section 4.2.

The CE-H is similarly made of alternating absorber and active layers. The absorber is made of disks of stainless steel of 35 mm thickness for the first 12 samples and 68 mm for the last 12. These are assembled in a rigid mechanical structure into which the active elements are inserted. The first 8 layers are instrumented with silicon detector modules similar to those in the CE-E. In subsequent layers silicon detectors are used in the smaller radius (higher η) regions and scintillator tiles, read out by silicon photomultipliers (SiPM), at larger radius (lower η). The scintillator detector modules are described in Section 2.4. The radius at which the transition from silicon to scintillator is made varies with depth from about 1.5 m ($\eta = 1.7$) at layer 9, to about 0.9 m ($\eta = 2.4$) in the back-most layers. The silicon and scintillator modules are assembled into cassettes similar to those for the CE-E, but with modules on only one side of the copper cooling plate, silicon module baseplates made of of less-expensive carbon-fibre (chosen for its mechanical and thermal properties), and simple copper covers rather than stainless-steel clad lead. The CE-H cassettes are built in $30\,^{\circ}$ units to keep them of a manageable size. Pairs of cassettes are joined into a $60\,^{\circ}$ unit just before insertion into the absorber, as described in Section 4.2. The main parameters of the HGCAL are summarized in Table 4.1.

Table 4.1: Size, weight, and structure of each endcap calorimeter.

	CE-E	CE-H (fine)	CE-H (coarse)	Total
z_{min} (mm)	3191	3530	4131	
z_{max} (mm)	3530	4131	5138	
r_{min} (mm)	320	357	420	
r_{max} (mm)	1680	2122	2664	
Layers	28	12	12	52
X_0 (CE-E only)	26.3	-	-	26.30
λ	1.73	3.30	5.72	10.75
Mass (t)	23	36	155	214
Cassette type	double sided	single sided	single sided	
Cassette width	60°	30°	30°	
Cassettes	84	144	144	372

4.2 Cassettes

The sensors in each layer of the HGCAL are grouped into 30° or 60° wedges called cassettes. These are the major detector subassembly of the HGCAL, which are subsequently assembled into full disks in the CE-E or inserted between absorber layers to form full disks of detectors in the CE-H. In the electromagnetic section, the absorber layers are part of the cassette and the CE-E consists of the stack of cassettes, supported by a full-disk back plate and an inner cone at the $|\eta| = 3$ boundary. In the hadronic section, the absorbers and cassettes are independent of each other.

Figure 4.1 is a drawing of a CE-E cassette. It is built by tiling the hexagonal silicon modules on both sides of the central cooling plate, which provides the main mechanical structure of the cassette. Motherboards are mounted on rows of modules and joined end-to-end to form a continuous line which brings low-voltage power and control signals to the individual modules, as described in Chapter 3. Stainless steel clad lead covers/absorbers close the cassette on both sides as shown in Fig. 4.2. Special screw assemblies secure the modules at their corners and also serve as spacers to define the gap between the cooling plate and the covers in which the active detector elements are located. Details are shown in Fig. 9.1. The cassette in Fig. 4.1 corresponds to the 14^{th} cassette layer, the last in CE-E, which consists of detector layers 27 and 28. It measures approximately 1.3 m in the radial direction by 1.6 m wide across the outer radius. The cassettes in the first disk (detector layers 1 and 2) are about 0.1 m smaller in each dimension. In moving from layer to layer, although the inner and outer radii of the cassettes change, the radial position of the modules remains the same, constrained by the fixed tiling of the fixed size hexagons. The modules and cassettes are also aligned with each other azimuthally for structural reasons such that the spacers align in all layers. To the extent possible, partial modules of 50% and 80% of the area of full hexagons are used to maximize the coverage at the inner and outer peripheries, as shown in Fig. 4.1.

Cassettes with silicon modules are used in the first eight layers of the hadron calorimeter. They are of similar design to those in the CE-E, with the following differences: the cassettes are single-sided with modules mounted on only one side of the cooling plate; the cover serves only as a mechanically and electrically protective layer and is therefore made of 1 mm thick copper; they are built in 30° units; and 60° pairs of cassettes are connected to the adjacent steel absorber layer rather than to each other. Starting in CE-H layer 9, the outer silicon modules are

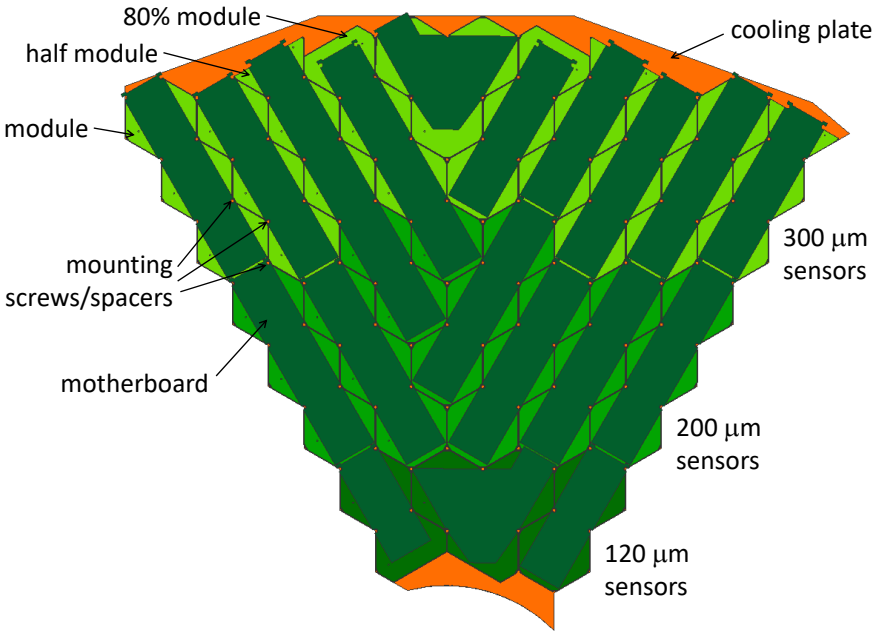


Figure 4.1: Schematic layout of a CE-E cassette showing the layout of the silicon modules and a possible motherboard configuration. The dark green modules have 120 μm thick sensors with 432 channels per full-size module, the two progressively lighter green modules have 200 μm and 300 μm thick sensors respectively both with 192 channels per full-size module. Motherboards are indicated by the rectangles. Full-size modules have a width between the flat sides of 164 mm, and the motherboards are 94 mm wide. Modules of 50% and 80% of the full area are used to improve the coverage at the inner and outer edges.

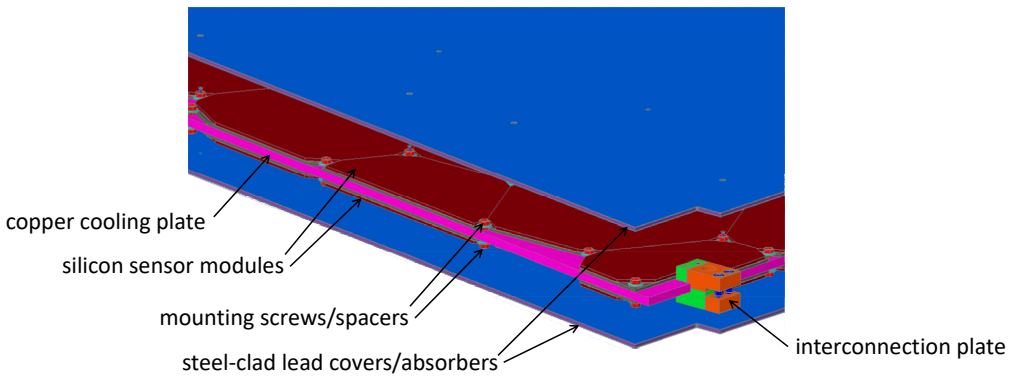


Figure 4.2: 3D model of the outer edge of a CE-E cassette, showing silicon modules mounted on the two sides of the cooling plate and the stainless-steel clad lead covers/absorbers.

replaced by scintillator/SiPM panels, and the fraction of scintillator used grows progressively toward the back of the CE-H. The scintillator area fraction varies from 40% in CE-H layer 9 to 90% in layers 16–24. These mixed cassettes follow a design similar to the silicon sensor cassettes.

The geometry of the 30° CE-H cassettes is illustrated in Fig. 4.3, which shows a cassette with both silicon and scintillator sensors. The cassettes are made in pairs such that full hexagon silicon modules can be used along both radial edges. The scintillator sections are constructed as simple 30° wedges with radial edges to match the $r - \phi$ geometry of the scintillator tiles, described in Section 2.3. Each cassette is built and tested as a complete stand-alone unit with all detector elements, electronics and cooling loop. At the point of insertion into the CE-H absorber (Section 4.6) the cassette pair is combined into a single 60° insertion unit with a geometry similar to the monolithic 60° CE-E cassette described above.

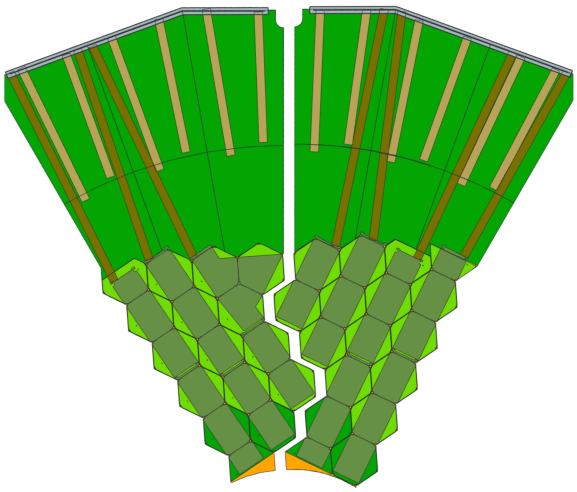


Figure 4.3: A pair of 30° mixed silicon/scintillator cassettes ready to be joined into a 60° unit for insertion. This example corresponds CE-H layer 12.

The three different types of cassettes – CE-E, CE-H (silicon) and CE-H (mixed) – have different thicknesses determined by their respective components: cooling plate, silicon or scintillator sensors, electronics, and covers/absorbers. The longitudinal layout of materials in the three types of cassettes is illustrated in Fig. 1.5. The respective nominal thicknesses are 24.4 mm, 13.0 mm and 14.9 mm, with the exception of the last CE-E cassette in which the downstream 2.8 mm thick Pb/steel absorber is replaced by a 1 mm copper cover. A 1 mm nominal clearance is provided between the CE-H cassettes and the adjacent absorber layers. These dimensions, together with the CE-H absorber thicknesses, determine the longitudinal segmentation shown in Fig. 1.6.

More complete details of the design, assembly and testing of the cassettes are given in Section 9.2.

4.3 Structural design

The mechanical structure of the endcap calorimeter, shown in Figs. 1.6 and 1.5, consists of a set of steel disks that are the absorber material for the hadron calorimeter, a structure for supporting the stack of cassettes for the electromagnetic calorimeter that include the absorbers as an integral part of the cassette, and a polyethylene neutron moderator, whose purpose is to reduce the neutron flux in the Tracker. The structure is cantilevered from the nose of the

first endcap disk of the iron yoke of CMS, YE1. The structural design is summarized here; it is presented in more detail with supporting design calculations in Section 9.3.

4.3.1 Electromagnetic calorimeter

The structure of the electromagnetic section (CE-E) is illustrated in Fig. 4.4 and is described more completely in Section 9.3.1, which also contains engineering calculations for key elements. In CE-E, the absorber layers are an integral part of the cassettes. Each mechanical layer of the CE-E (two detector layers) consists of a disk formed by joining six 60° cassettes. Each disk is stacked directly on top of the one below it and the whole structure is supported by a 40 mm thick stainless steel back disk and a 10 mm thick aluminium support cone. In each layer, cassettes are joined one to the next with connection plates at the outer radius and an interconnection ring that fits around the central cone. The detector is assembled in a horizontal position, as shown in the left side of Fig. 4.4, and also shown in Fig. 4.8(a). The cassettes are aligned so that the mounting screws/spacers at the corners of the modules, shown in Fig. 4.1, are at the same positions in each layer, thus transmitting the weight of the cassettes stack to the back disk through a rigid structure and not any active elements. When in the installed vertical orientation, as shown in the right side of Fig. 4.4, the weight of the detector is carried by the support cone and transferred to the CE-H structure via the CE-E back disk. The cassettes are stabilized in their longitudinal position by brackets located at the outer radius every 30°. A more detailed view of this structure is in Fig. 9.11.

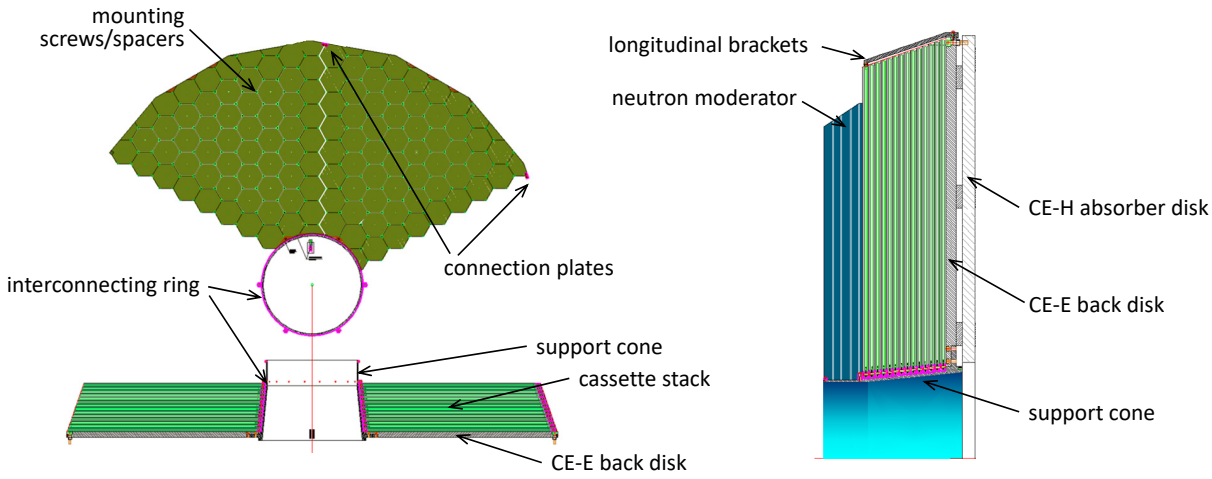


Figure 4.4: The structure of the CE-E shown in the assembly position (left) and operating position (right) with the main elements indicated.

4.3.2 Hadronic calorimeter

The hadron calorimeter structure consists of 11 stainless-steel disks of 35 mm thickness, 12 disks of 68 mm thickness and a 94 mm thick back disk as shown in Fig. 1.6. The stainless-back disk of the CE-E serves as the absorber in front of the first CE-H detector layer. The first 8 disk-to-disk gaps are 14.0 mm to accommodate silicon-detector cassettes; the remaining 16 gaps are 15.9 mm as required for the mixed cassettes. Spacers and bolts or tie rods are placed between the absorber plates every 30° around the outer and inner peripheries. The spacers define the cassette slot thickness and also serve as shear keys to transfer the weight from one layer to the next and ultimately to the back disk. The cantilevered moment of the structure is taken by the bolts and tie rods that pass through the centres of the spacers. The 12 inner spacers and 6 of the 12 outer spacers are permanently installed. Every other outer spacer

is designed to be removable (one at a time) while the absorber stack is lying in a horizontal position to enable the insertion of the 60° wide cassette pairs. A full presentation of the CE-H structural design including details regarding spacer and bolt/tie rod designs and sizes and key engineering calculations can be found in Section 9.3.2.

4.3.3 Endcap suspension system

The entire endcap calorimeter is supported in a cantilevered fashion from the nose of YE1 as shown in Fig. 1.6 and in more detail in Fig. 9.18. The back disk is connected through a set of wedge-shaped blocks (Fig. 9.19) to a similar disk called the back flange, which then transfers the weight of the endcap to a large diameter steel support tube connected to the YE1 disk. The wedge supports serve to: a) transfer both the vertical load and rotational moment of the HGCAL to the back flange; b) provide thermal isolation between the HGCAL, which operates at $-30\,^{\circ}$C and the room temperature back flange; c) provide sliding surfaces that allow for the approximately 2 mm radius reduction of the HGCAL when cooled to $-30\,^{\circ}$C; and d) define the space between the back disk and back flange that is occupied by the ME0 and electrical services for the CE. The rotational moment of the HGCAL about the end of the support tube is taken by a set of special brackets, of the same design used for the existing endcap, that connect the outer radius of the back flange to the YE1 while allowing access to the ME1/1 chambers, as shown in Fig. 9.20. Full details including supporting engineering calculations can be found in Section 9.3.3.

4.4 Cooling and environmental control

The total nominal power dissipated in the HGCAL volume, due principally to electronics and sensor leakage current and secondarily to heat static leaks, is expected to be 110 kW per endcap. The breakdown of the heat load is given in Table 9.7. The cooling system must remove this heat load and maintain the silicon sensors at a temperature of $-30\,^{\circ}$C or lower. The cooling system is based on two-phase CO_2 fluid and will be designed for a nominal coolant operating temperature of $-35\,^{\circ}$C with a maximum vapour quality (mass ratio between vapour and total mass) of 42%. To allow margin, the cooling system serving both endcaps is designed for 300 kW at $-35\,^{\circ}$C. Due to the efficient heat conduction of the metallic structure of the HGCAL, the temperature gradients between the coolant and the sensors (silicon and SiPMs) is expected to be very low, on the order of 1–2 °C.

The cooling of the HGCAL detector is based on evaporative CO_2 in a liquid-pumped cycle similar to that used in the CMS Phase-1 Pixel Upgrade [24] and that has been adopted for the CMS Phase-2 Tracker Upgrade. It uses the two-phase accumulator-controlled loop concept (2PACL) [25] that has been successfully employed in the LHCb VELO since 2008 [26]. The fluid properties of CO_2 make it an ideal medium for cooling of detectors, in particular because it enables the use of smaller diameter tubing than is required with conventional refrigerants or liquid cooling applications. The high pressure at which CO_2 evaporates keeps the vapour compressed and therefore the circuit volume low. The boiling temperature is a function of the pressure and, since the dynamic pressure drop along a cooling pipe is small compared to the absolute pressure of the fluid, the change in the evaporation temperature along a cooling pipe is typically small: two degrees between outlet and inlet coolant temperature is the target for the HGCAL design. The high latent heat of vaporization of CO_2 translates into a smaller flow as compared to that required for other refrigerants. Because the viscosity of CO_2 is low, it allows the use of small diameter pipes with higher flow speeds, which increase the heat transfer coefficient from the pipe wall to the fluid. CO_2 is radiation hard, inexpensive, and

environmentally friendly; in particular the global warming potential is orders of magnitude lower when compared to chlorofluorocarbons, such as C_6F_{14}, used in the present CMS Tracker.

Rigid vacuum-jacketed stainless steel pipes carry the CO_2 from the refrigeration plants in the service cavern, USC, into the collision hall, UXC. Flexible piping in the cable chains connecting to the manifolds on the YE1 structure allow for opening CMS. The flow is fed to the detector cold volume via 24 vacuum-insulated coaxial lines on each endcap. Each cooling loop is fitted with a manual shutoff and flow regulation valve. Inside the cold volume, the coaxial line is split into supply and return tubes that run along the outside periphery of the detector. Each cooling loop feeds either the even or odd cassette layers in one 30 ° sector. If one loop needs to be closed only half the layers in one sector are disabled. Figure 4.5 shows the cooling loops on the HGCAL and a preliminary routing of the vacuum-jacketed coaxial lines over the endcap suspension system brackets. Five cooling plants with a maximum capacity of about 60 kW at −35 °C will be employed. Each will operate at about 75% capacity under normal conditions (i.e. nominal cooling power of 220 kW), and there is sufficient margin to allow continuous detector operation if one plant is off-line. Using all five plants, the cooling system is capable of supporting operation at ≈0 °C, as may be required during initial commissioning based on the humidity inside the cold volume. Normal operation at temperatures up to 15 °C is in principle possible but, due to the amount of flow needed, this would require use of all the five plants up to 100% of their capacity with little or no margin with respect to the dry-out conditions (when the cooling performances degrade drastically) unless additional refrigeration is added.

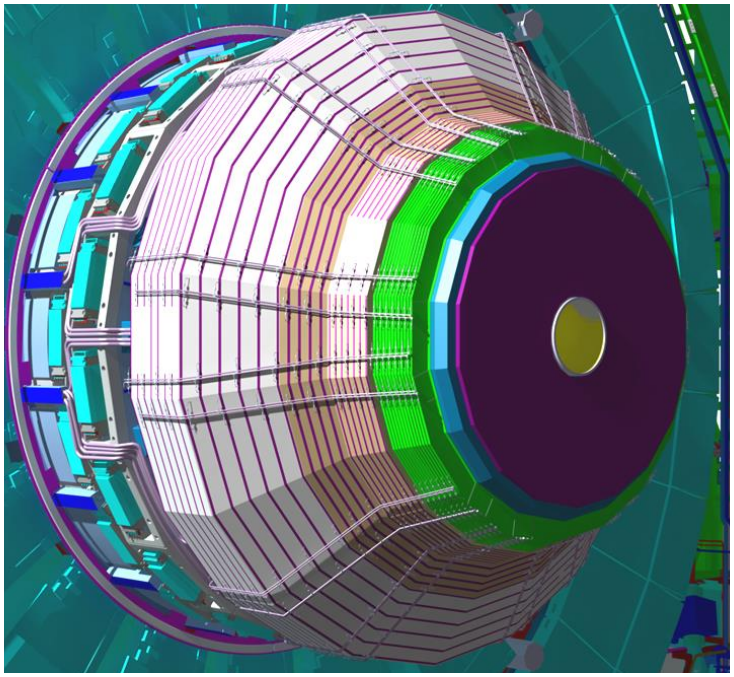


Figure 4.5: Layout of cooling loops on the HGCAL showing both the manifolds that supply the cassette layers (two sets of supply and return lines every 30 °) and a preliminary routing of the vacuum jacketed coaxial lines over the endcap suspension system brackets.

A thermal screen will surround the detector to separate the cold (−30 °C) dry (dew point −40 °C) detector volume from the ambient conditions in the cavern. The thermal screen is made from panels consisting of an aerogel core (very low thermal conductivity) and two 1.5 mm thick Permaglas® skins to provide a stiff structure of low thermal conductivity. The overall thickness varies from 15 to 32 mm depending on the available space at each location. Heating foils are placed on the outside surface to maintain it at or near ambient temperature. The de-

tector volume will be continuously flushed with dry nitrogen during operations and dry air for safety reasons when CMS is open.

A sufficient subset of the detector cooling plants will remain in operation during maintenance periods, to keep the detector volume at low temperature, as required to avoid incurring additional damage from radiation effects as a result of reverse annealing. Diesel-powered electrical supplies are used to ensure continuous operation even during power cuts for two plants out of five.

4.5 Services

Services required for the operation of the HGCAL include: CO_2 cooling; low-voltage for the electronics and high-voltage to bias the silicon detectors; optical fibres for data and trigger information; and a dry gas system. The space required for the cooling and LV cabling dominates the space required for the services, and initial studies have focused on these. Services for the endcap timing layer (TE) must be routed through the HGCAL cold volume. Conceptual work has been started on how to route the services for the endcap timing layer; however the requirements from the TE are not yet well determined and designs to meet the requirements, presented in Section 9.5.3, are very preliminary.

The CO_2 cooling is divided into 30° sectors, and within each 30° sector even and odd layers are cooled by independent manifolds, as shown in Fig. 4.5. Supply and return lines connected to the cassettes are routed separately inside the HGCAL, and are combined into co-axial vacuum-jacketed pipes before exiting the cold volume. The electrical and optical services are routed on the surface of the HGCAL adjacent to and under the cooling pipes, as shown in Fig. 4.6. The thickness of the electrical services varies with azimuth, reflecting the higher current required by the longer motherboard lines at the edges of the 60° cassette units rather than in the middle (Fig. 4.1).

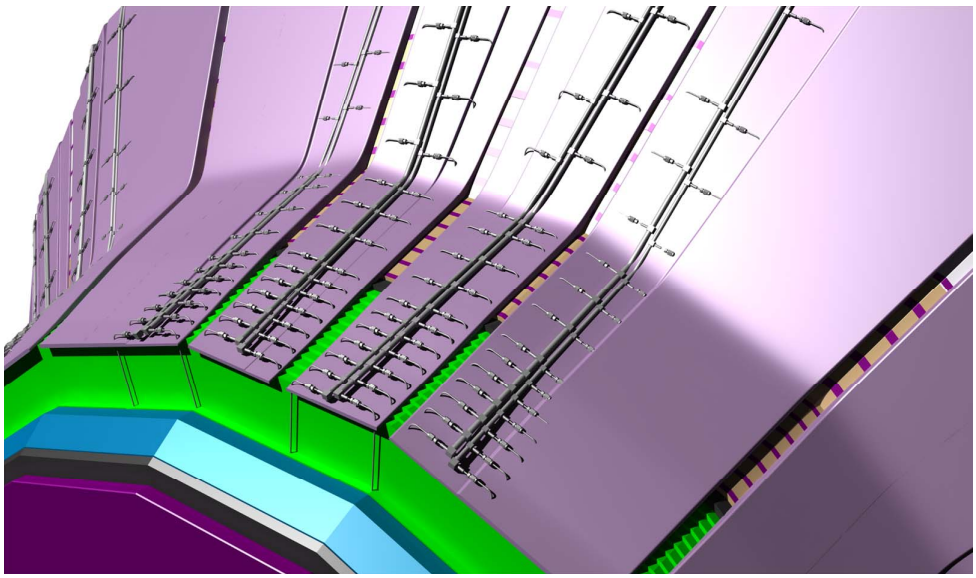


Figure 4.6: Preliminary layout of the cooling and low-voltage services on the surface of the HGCAL. The violet bands represent the volume of the low-voltage cables. The empty bands are regions reserved for mechanical supports of the absorber structure and the thermal screen.

The low-voltage supply for the on-detector electronics is provided at ≈1.5 V by DC-DC converters located just outside the cold volume behind the CE-H and is routed in separate circuits

to each line of motherboards (Fig. 4.1). With separate voltages for the analogue and digital electronics and a common return bus for each motherboard line, there are an estimated 8454 low-voltage cables required for each endcap, carrying a total of almost 50 kA. The required area at the back of the HGCAL, where all cables are present, is estimated to be 2362 cm^2, based on actual dimensions of appropriate commercially available cables and assuming a 70% packing fraction. Full details regarding this estimate are given in Section 9.5.2. The total volume of services grows from the front to the back of the HGCAL, but starting at CE-H layer 8, the radius of the HGCAL grows rapidly with increasing layer number providing more circumference in which to lay the services. Therefore the region between the back of the CE-E and CE-H layer 8 is a critical area, because of the thickness of the area occupied by the LV cables, as can be seen in Fig. 4.6.

The services exit the cold volume via cold-to-warm feedthroughs located around the outer circumference of the back disk of the HGCAL thermal screen. The HGCAL feedthroughs will be made following the same design approach that was successfully used by the existing CMS Tracker and more recently by Phase-1 Pixel system. The services pass through a set of custom-designed clamps that position and secure the cables and cooling pipes feedthrough. After the cables are secured, the region is potted to provide a tight seal to ensure that the humidity of the cold volume can be maintained at the required low level. Further details can be found in Section 9.5.4.

Other critical areas are the routing of the services around the endcap mounting brackets, where a clear opening must be left for access to the ME1/1 chambers, and in the radial cable trays at the face of the YE/1 disk. These areas are shown in Fig. 4.7. Of the two, the tightest spot is crossing the brackets, where a total area of only $\approx$1800 cm^2 is available, which must be shared among the HGCAL, the endcap timing layer, and the ME0 and GE1/1 muon chambers. There is more space available in the radial cable trays, but these must be shared among the same subdetectors plus the ME1/1. Low-voltage power is supplied through this region to the HGCAL DC-DC converters at $\approx$12 V, which reduces the cable cross-section substantially relative to that inside the HGCAL. On the other hand, due to constraints outside of the endcap the CO_2 pipes can enter only near the horizontal plane, making these areas especially crowded and adding complexity to the layout. Designing how the services will be routed through this area will be done jointly by the three detector subsystems under the direction of the CMS technical coordination.

Full 3D modelling of the services layout is critical and has proved to be very effective in designing the services for other CMS subdetectors. Based on 3D models being developed for the HGCAL, a series of mockups is planned to address the layout in critical areas discussed above including the region near the front of the CE-H, the cold-warm feedthroughs, and the area around the brackets. Eventually a full-scale mockup of as much as a 60° segment of one endcap may be useful or even essential.

4.6 Assembly, installation and commissioning

4.6.1 Detector assembly and installation

The two endcap calorimeters will be assembled on the surface and completely tested before lowering into the underground collision hall, UXC, and mounting on the YE1. The assembly sequence, including quality control, intermediate testing, and final testing before installation is described in detail in Section 9.6 and summarized here. The main elements of the assembly process are illustrated in Fig. 4.8.

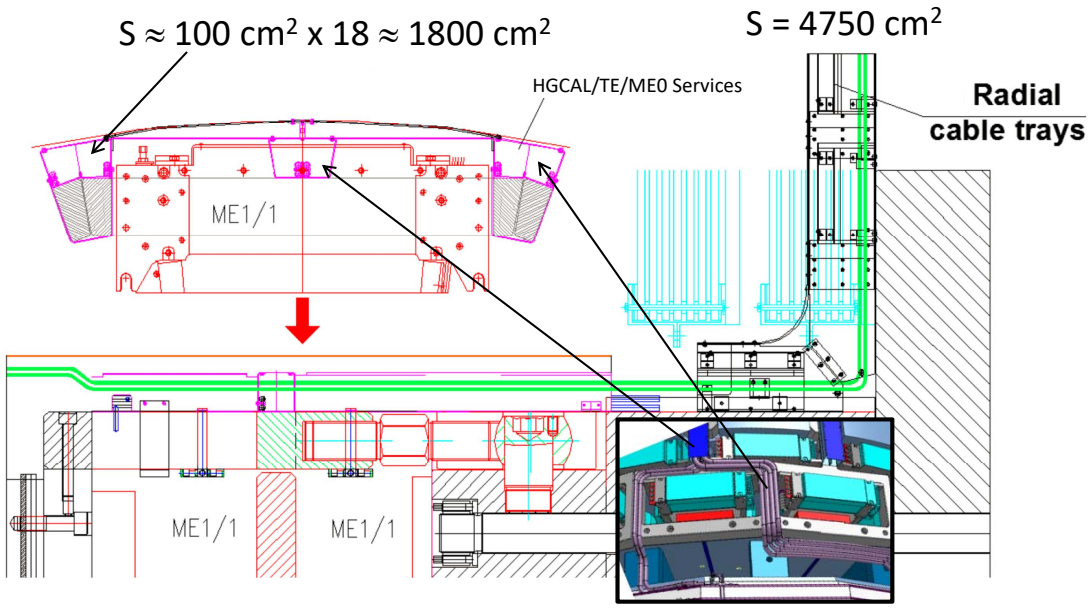


Figure 4.7: Cross-sections of key areas for the routing of services: the region crossing the brackets around the ME1/1 chamber access point and the transition to the radial cable trays on the YE/1. The inset is a 3D model showing a very preliminary layout of CO_2 cooling lines around one of the brackets.

The two CE-E will be assembled in CERN Building 904 in a clean room environment. First the back disk and central support cone are joined and placed on an assembly table that matches the CE-E dimensions and can hold the 23 t weight of the completed detector. Each of the fourteen layers of cassettes (28 detector layers) are assembled into disks by stacking six 60° cassettes on the layer below it as shown in Fig 4.8(a). For each layer, the interconnection ring is positioned around the central support cone and azimuthally aligned. Then each cassette is placed on the stack using a special lifting fixture and attached to the interconnection ring. Each cassette is connected to its neighbours with the interconnection plates at the outer radius. Visual inspection, dimensional checks, and basic electronic functionality tests are performed after the installation of each layer. The longitudinal brackets are installed after the full cassette stack is complete, and finally the neutron moderator is attached to the front of the completed CE-E. A complete electrical test of the completed CE-E is performed at room temperature using a temporary cooling system. Finally, the CE-E is transported to Point 5 for installation onto the front of the CE-H.

The CE-H absorber disks will be manufactured in several segments that are bolted together first in a test assembly at the factory. Each layer is fully surveyed and measured to ensure that the assembly is within tolerances, including crucially the thickness of the gap for the active detectors and the overall dimensions of the completed assembly. The structure is dismantled and the plates transported to CERN, reassembled to form full disks and surveyed again. The absorber layers are stacked on a special tooling table using a simple spreader bar, as shown in Fig. 4.8(b).

Once the CE-H absorber structure is fully assembled the cassettes are inserted into the absorber gaps, as shown in Fig. 4.8(c). The ME0 muon chambers (shown in Fig. 1.6) are inserted between the back disk and back flange first, and then the HGCAL cassettes are inserted from the rear to the front in each 60° segment. An extension of the tooling table that is wide enough to allow the

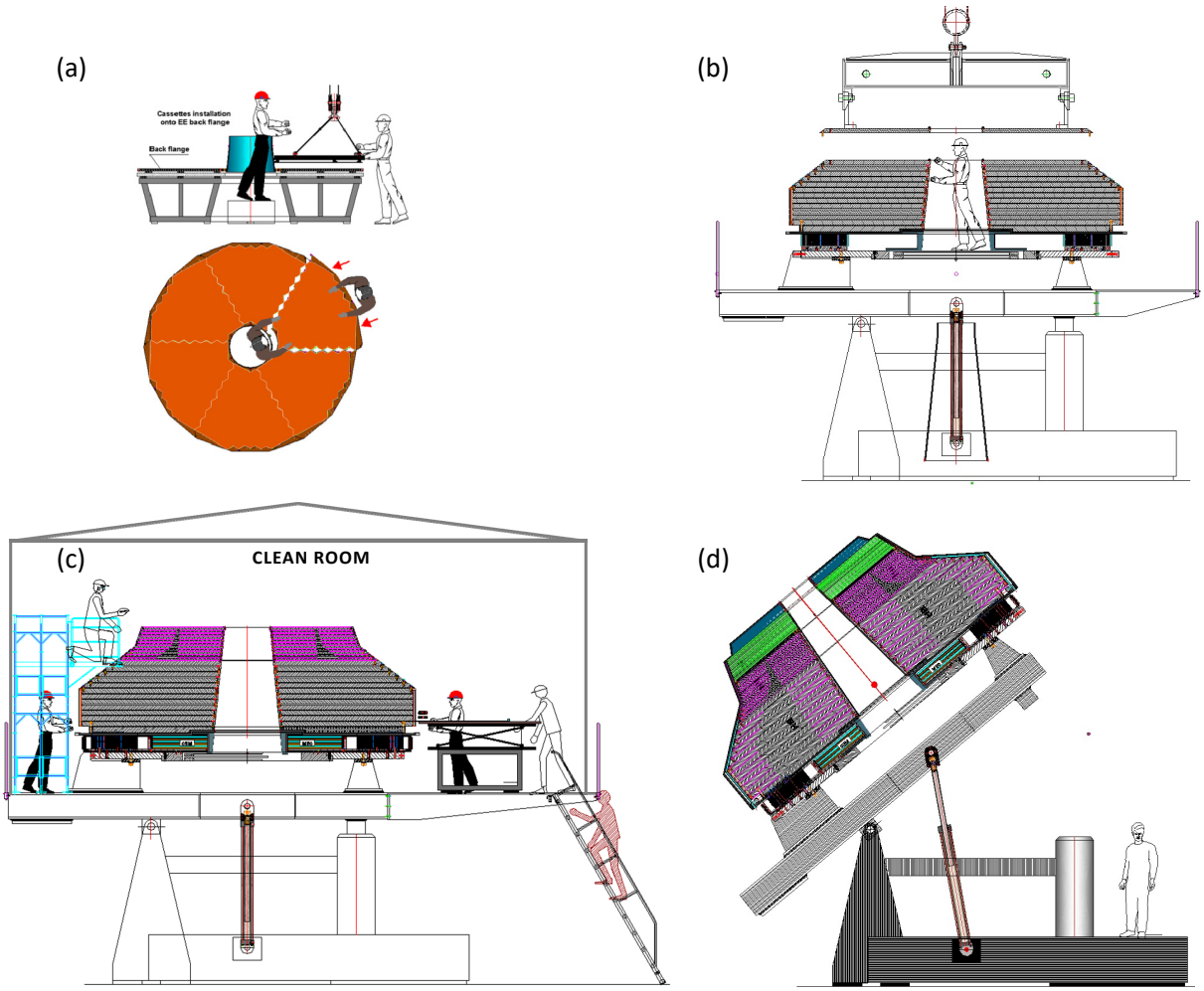


Figure 4.8: Summary of the detector assembly sequence: (a) stacking of CE-E cassettes on the CE-E support structure; (b) assembly of the CE-H absorber structure; (c) insertion of CE-H cassettes and dressing with electrical, optical and cooling services; (d) rotation of the completed calorimeter to the vertical orientation in preparation for installation.

insertion of a 60° wide pair of 30° cassettes (Fig. 4.3) is positioned at each insertion location, then moved to the next location. After all cassettes have been installed in a 60° sector, the electrical, and optical services can be connected while cassettes are inserted in another sector, as indicated at the left of Fig. 4.8(c).

After all CE-H cassettes have been inserted, their electrical and optical services installed, and electrical tests done to verify the assembly, the CE-E is mounted to the front of the CE-H. Its electrical and optical services are installed over the CE-H services and electrical tests performed. Cooling lines are installed in every 30° sector and connected to the cassettes. Each cooling loop is leak and pressure tested. The services are potted into the cold-to-warm feedthrough at the back of the HGCAL and the thermal screen is installed and leak checked. The completed endcap calorimeter is cooled to −30 °C, with a system that allows powering of one 60° sector at a time, and fully tested at operating temperature. A period of nine (six) months is built into the master schedule for these tests for the first (second) endcap.

The completed detector is rotated from the horizontal assembly orientation to the vertical position for lowering into UXC, as shown in Fig. 4.8(d). It is placed on a transport cradle and hoisted down the shaft into the UXC with a special lifting fixture, as shown on the left side of Fig. 4.9. This will be a special lift using a large-capacity mobile crane with capacity to deal with the >300 t weight of the completed endcap plus cradle and lifting fixture. Once in the UXC, the HGCAL will be moved on air pads, raised to the installation height on an existing system of risers and screw jacks, as shown in the right side of Fig. 4.9, and mechanically connected to the YE1 nose.

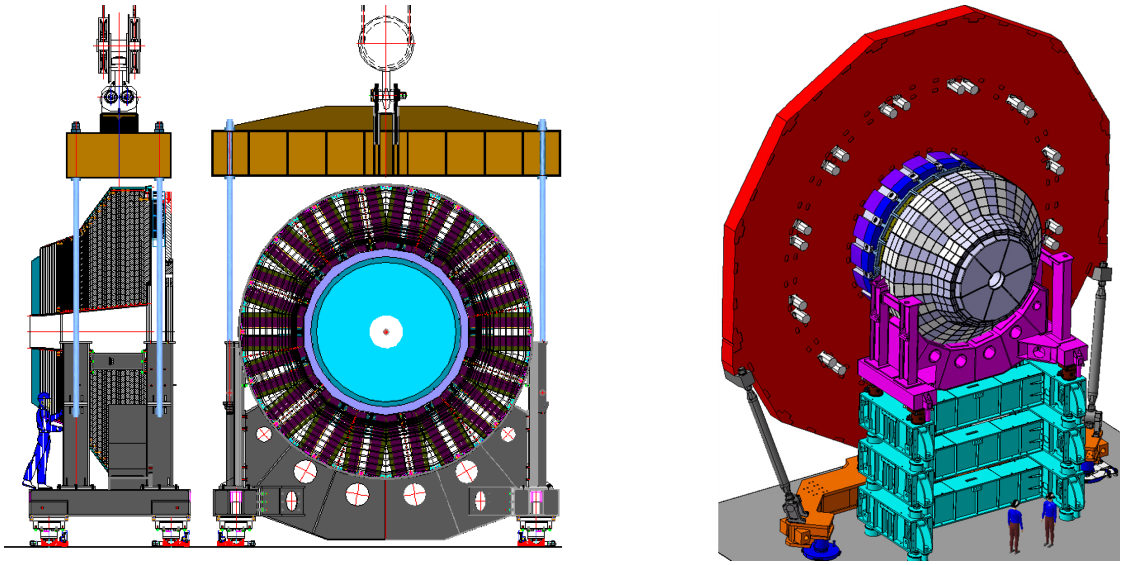


Figure 4.9: Endcap support cradle and spreader bar used for lowering one endcap of HGCAL from the surface assembly building into the experimental cavern (left). Hydraulic riser system used to elevate and position the HGCAL for connection onto the YE1 (right).

4.6.2 Commissioning

The commissioning of the HGCAL will be in several stages, before and after CMS is closed.

As part of the assembly process of both the CE-E and the CE-H, detailed pre-commissioning will take place, using temporary cables/fibres and power supplies, in order to check everything prior to the connection of the final services including cooling. A dedicated standalone system will be used to connect and power—for a short time without cooling—one sector at a time and

perform a series of tests to check basic connectivity and functionality. These tests will verify:

- Environmental control and monitoring (DCS/DSS): connectivity and functionality,
- LV and HV power systems: connectivity, power consumption, leakage current etc.,
- DAQ slow control: connectivity, programmability, and functionality,
- DAQ/trigger readout optics: connectivity.

Any problems detected in the CE-H while in the horizontal position are possible to repair as the cassettes are, at this point, removable. Although the CE-E cassettes may be difficult, though not impossible, to exchange, a knowledge of whether anything is malfunctioning is essential.

A full set of these tests will be performed on the completed HGCAL after all services have been installed but before closing the thermal screen. More extensive tests can be done at this point by utilizing the CO_2 refrigeration system in the surface assembly hall at Point 5, SX5, to cool the electronics in each 60° sector at a temperature above the dew point. After the thermal screen is installed, the detector will be cooled to $-30\,^\circ$C. The CO_2 refrigeration plant will have the capacity to support the operation of one 60° sector at a time. A comprehensive set of tests will be performed on each sector including cosmic ray measurements. A period of nine months is built into the schedule between the end of surface tests and the foreseen date for lowering.

Following this first pre-commissioning campaign in the surface assembly hall, the HGCAL will be rotated and lowered for installation on the YE1. A period of nine months is built into the master schedule between the end of the surface tests and the date foreseen for lowering. The services from the HGCAL will be connected to cooling pipes, cables and optical fibres that will have been installed on the YE1 and connected to their respective destinations. The cooling circuit will be connected and tested fully with high-pressure CO_2 at room temperature. Any leaks or problems found will be fixed in situ. Electrical and optical services will be connected, sector by sector and short verification tests similar to those done during detector assembly will be performed as the connections are made.

Once all sectors are fully tested, the thermal screen will be leak checked and the dry gas flow will be established. Once the internal dew-point is sufficiently low (around $-40\,^\circ$C), the detector will be cooled down using the full CO_2 refrigeration system. The cool-down may take several days, during which longer-term tests of overall stability etc. can be performed, as well as detailed monitoring of the environment.

When the HGCAL is at its nominal temperature of $-30\,^\circ$C full system tests can be performed using the final off-detector DAQ/trigger boards, firmware and software. When this is successful, the go-ahead for CMS closure (from the HGCAL perspective) can be given, and detailed optimization and commissioning of the HGCAL, including various types of calibration, can start. This latter will include:

- Bias and gain of optical links
- Measurement of the trigger chain latency
- Thresholds and gains of front-end ASICs
- Dead-channel masking
- HV settings and IV scans of silicon sensors

Following the local calibration, the detector should join CMS global runs in the central DAQ/DCS. This will test the compatibility of the detector, DAQ, and DCS with the overall CMS experiment. The steps will include the following, at different points in the commissioning:

- High trigger-rate data taking with random triggers
- Cosmic data taking for the gross timing alignment for detector readout latency and possibly the trigger-primitive timing.
- ”Splash” events from LHC for timing and readout phase alignment.
- Final system calibration and alignment with proton-proton collisions.

Chapter 5

Reconstruction and detector performance

5.1 Reconstruction and detector performance

5.1.1 Reconstruction aims

An imaging calorimeter, such as the HGCAL, with fine granularity, both lateral and longitudinal, produces a large amount of information and enables enhanced pattern recognition. This is illustrated in Fig. 5.1, which shows energy deposits from photon showers separated by ≈3 cm. In addition, the HGCAL provides precise information on the timing of the energy deposits to discriminate clusters from pileup within a single bunch crossing. The five dimensional information (energy, x, y, z, and time) is ideally suited for particle-flow reconstruction. Hadronic and electromagnetic showers are individually reconstructed and identified, and charged hadrons matched to tracks reconstructed in the tracker. Jets (and missing energy) can then be measured using the best available energy and 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. Additionally, the calorimeter is sufficiently deep to provide excellent muon identification performance.

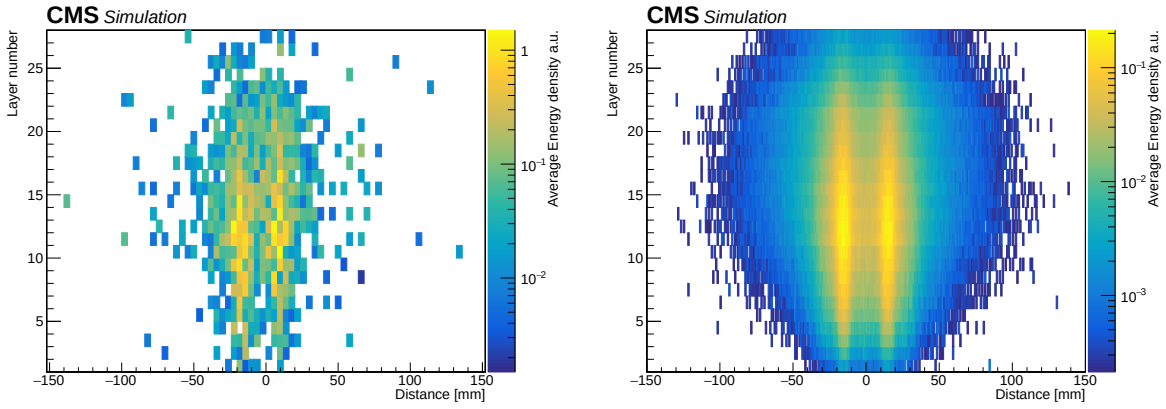


Figure 5.1: Energy deposited in HGCAL cells by pairs of unconverted photons; (left) a single event containing a pair of photons, and (right) several thousand such events. The photons have an energy of $E = 80\,\text{GeV}$ ($p_\text{T} = 14.4\,\text{GeV}$) at $\eta = 2.4$ in the HGCAL, and are separated by ΔR=0.05 (in a random orientation), corresponding to a separation distance of about 30 mm. Reconstructed hits are projected onto the plane defined by the axes of the two showers. The colour code represents energy density.

Unlike the particle-flow calorimeter designs investigated by the CALICE Collaboration [27, 28], the HGCAL reconstruction must function in the severe pileup environment of the HL-LHC, putting further emphasis on pattern recognition.

The development of the full particle-flow reconstruction for the HGCAL will take time and much work. The aim so far has been to develop an understanding of possible approaches and techniques, and to demonstrate the potential of the detector.

5.1.2 Simulation model

The detailed implementation of the detector geometry necessarily lags behind technical design choices. In autumn 2016, before the making of many of the important engineering decisions that are described in this document, the HGCAL geometry implemented in the CMS simulation and reconstruction software, CMSSW, was frozen to allow simulation work to proceed. The detailed calorimeter geometry closely follows what was described in the Technical Proposal [1], and uses hexagonal silicon cells. Both the names used in the description of the results (EE, FH and BH, referring to the electromagnetic, forward hadronic and backing hadronic calorimeters) and the structures they describe belong to the Technical Proposal era. A few simple modifications were made in order to better simulate the HGCAL design as it was envisaged in 2016. The number of samplings in each section of the calorimeter in the CMSSW geometry is as described in the design given in this document, and the detailed longitudinal structure differs little, as is shown in Table 5.1. The simulation of the signal development and processing in the HGCAL electronics includes the addition of Gaussian noise, and digitization. Further details of the simulation model are given in Section 10.1.

Table 5.1: Longitudinal structure of the HGCAL, comparing the thicknesses of what is simulated in CMSSW with what is described in the engineering sections of this document. The thickness is measured in centimeters (cm), radiation lengths (X_0), and nuclear interaction lengths (λ).

	CMSSW			**TDR design**		
	cm	X_0	λ	cm	X_0	λ
Neutron moderator	18.0	0.4	0.2	12.0	0.3	0.2
Electromagnetic section	32.1	26.7	1.6	33.9	25.4	1.5
1st hadronic section	59.7	33.8	3.1	60.2	35.2	3.7
2nd hadronic section	110.6	62.8	5.8	100.6	49.9	5.2
Total	220.4	123.6	10.7	210.4	110.9	10.7

5.1.3 Intrinsic energy and position resolution for electromagnetic showers

The electromagnetic energy resolution of the HGCAL has been studied using a GEANT simulation based on the geometry of the calorimeter described in the engineering sections of this document, both in terms of the longitudinal structure and the lateral cell sizes. One difference is that the sensors with a sensitive thickness of 120 μm are modelled by 100 μm thickness silicon (together with the higher noise and lower charge collection of 100 μm).

The response to photons that do not convert in the tracker material was studied, both in the absence of pileup, and with an average of 200 pileup events per bunch crossing. The energy was estimated by summing cells in each layer in a restricted area of radius 2.6 cm, about the maximum cell. This radius results in a sum of 19 cells in the 200 and 300 μm silicon, and 37 cells in the 100 μm silicon. The resulting resolution performance is insensitive to pileup. Using

larger radii slightly reduces the stochastic term, but the overall resolution is degraded in the presence of pileup, particularly at high η. A simple correction is applied for the small leakage of high energy showers from the rear of the electromagnetic section of the calorimeter. It is described in Section 10.1.6.

In the absence of pileup the stochastic terms of the fitted energy resolution functions have the values of 24%, 25%, and 28%/$\sqrt{E}$, respectively, in the regions of 300, 200, and 100 μm thick silicon (the photon directions were chosen so that the showers were exclusively in 300, 200, or 100 μm sensor cells). The stochastic terms fitted in the case of pileup are similar to those in the zero pileup fits, although they all show a small increase, slightly larger than the statistical uncertainty on the fitted values, as can be seen in Fig. 10.6.

If a 53 mm energy summing radius is used the constant term is 0.5% for each of the three silicon thicknesses corresponding to what is expected from the intercalibration uncertainty (Section 5.4) that is included in the simulation of the electronics and digitization. Additionally, if a 53 mm energy summing radius is used, the stochastic terms in the three regions are slightly smaller (23%, 25%, and 27%/$\sqrt{E}$ respectively in the three detector regions). With the smaller energy summing radius of 26 mm, the constant term is higher, 0.8%, in the 300 and 200 μm silicon. This is thought to be due to the variation of energy containment in the cells included in the radius. The containment varies depending on the location of the shower axis in the central, maximum energy, cell. This variation could be corrected for. In the smaller cells of the 100 μm silicon this effect is not seen.

Figure 5.2 (left) shows the electromagnetic energy resolution as a function of incident photon p_T obtained in a full simulation study. The left plot shows the resolution obtained using a region of radius 2.6 cm to estimate the energy, both in the case of no pileup (the dashed lines are taken from the fits detailed in Section 10.1.6), and in the presence of a mean of 200 pileup events per bunch crossing. The right plot shows the result when the energy summing radius is increased to 5.3, In this case the resolution is noticeably degraded by the pileup, emphasising the benefit of the fine granularity of the HGCAL. Further details are given in Section 10.1.6.

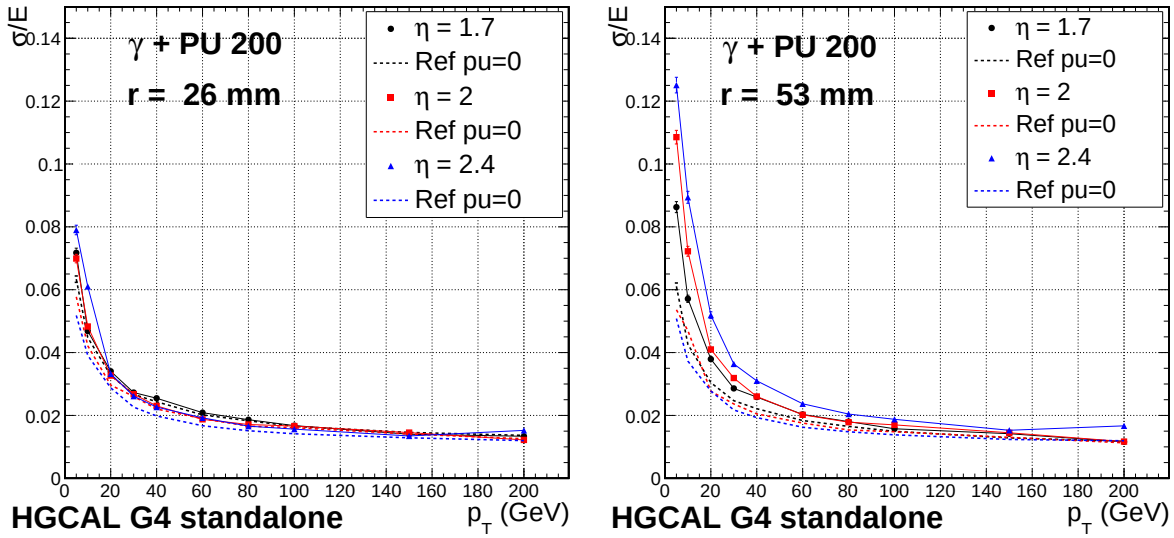


Figure 5.2: Fractional energy resolution, σ/E, as a function of p_T for unconverted photons at $|\eta| = 1.7$ (300 μm Si), $|\eta| = 2.0$ (200 μm Si), and $|\eta| = 2.4$ (100 μm Si), (left) using a region of radius 2.6 cm, and (right) using a region of 5.3 cm to sum the energy.

To demonstrate the intrinsic performance of the calorimeter, Fig. 5.3 shows the reconstructed mass of photon pairs from H$\rightarrow \gamma\gamma$ where both photons are in the endcap, and neither photon converts in the tracker material before the calorimeter. This result has been obtained using the same simulation model and reconstruction as used for Fig. 5.2. The photons are required to satisfy $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, and it is assumed the vertex is well located using timing information. The mass resolution is 1.4%, or 1.8 GeV. For comparison the corresponding mass resolution for a pair of unconverted photons in the barrel in the presence of an average of 200 pileup interactions per bunch crossing (before use of an optimized multivariate photon regression method), is expected to be 1.9 GeV at HL-LHC startup, and 2.8 GeV after 3000 fb^{-1} [29].

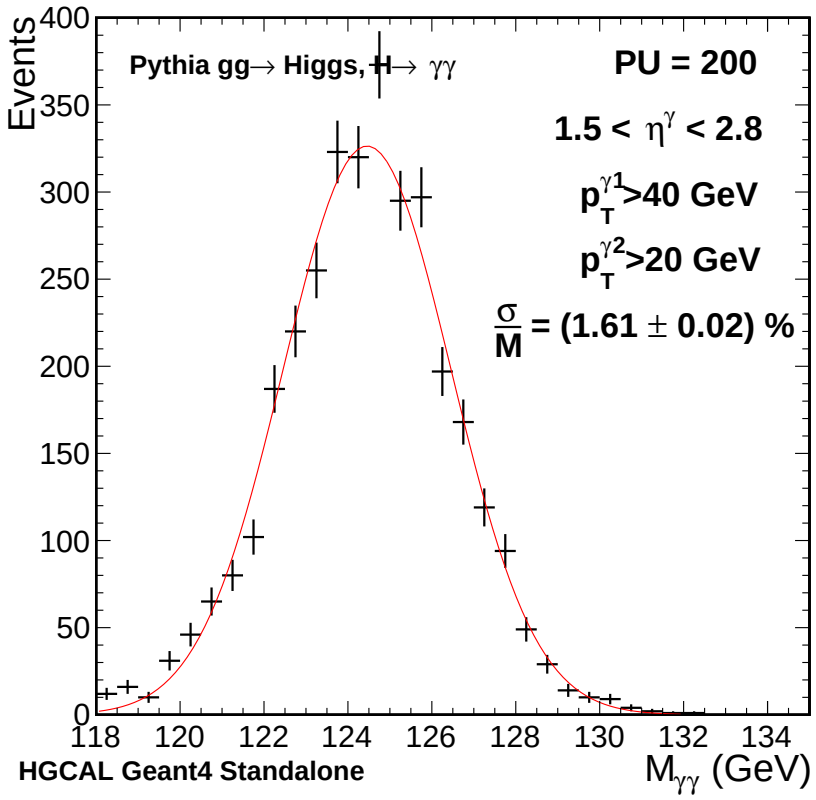


Figure 5.3: Reconstructed mass of photon pairs from H$\rightarrow \gamma\gamma$ where both photons are in the endcap, and neither photon converts in the tracker material before the calorimeter, and pileup corresponding to an average of 200 interactions per bunch crossing is present.

5.1.3.1 Energy resolution after 3000 fb^{-1}

Under irradiation several effects cause the S/N ratio for a MIP signal to decrease in both the Si cells and the SiPM-on-tile cells, as indicated in Tables 2.2 and 2.3. Since noise makes such a small contribution to the energy resolution, this is not of direct concern. Instead, the important consideration is maintaining intercalibration within 3% (the level assumed in the study described above). For this reason, the topic of intercalibration has been carefully considered. We intend to maintain 3% intercalibration throughout the life of the detector (Section 5.4), and thus the energy resolution results given in the previous section should be regarded as characteristic of the detector throughout its life.

5.1.3.2 Position resolution, and angular resolution

Because of the small cell sizes a position resolution of better than 1 mm can be obtained, in each layer, for high p_T electromagnetic showers, with little effort. This has been confirmed by the measurement of position measurement resolution in the test beam (Section 5.2). The ultimate position resolution performance, which should be significantly better than this, will require sophisticated multivariate corrections for numerous subtle biases.

Shower direction can also be measured. The axis of electromagnetic showers is determined using a principal component analysis methodology as a starting point for the construction of shower shape variables. The angular uncertainty in θ, obtained for a $p_T = 25\,\text{GeV}$ shower, without corrections, tuning, or optimization is 7 mrad. It is expected that this can be significantly improved by the use of more sophisticated techniques.

5.1.4 Clustering

The first step in reconstruction is to collect the signals of energy deposited in the sensitive cells of the calorimeter into clusters corresponding to electron, photon or hadron showers. The energy deposits are reconstructed from the recorded digital signals to give reconstructed hits (RecHits) approximately calibrated to correspond to energy lost in the absorber layers.

The approach that has been taken for the first prototype of the reconstruction is to construct two-dimensional (2D) clusters in each of the 52 layers of the calorimeter. Then, in a second step, these layer clusters are collected into a three-dimensional cluster (a "multicluster") by associating layer clusters that are approximately projective – i.e. they lie approximately on a straight line pointing from the CMS detector centre.

The 2D clustering algorithm, inspired by Ref. [30], is described in Section 10.1.2. It is suitable for implementation in GPUs, and, as such, may be useful as a future fast high level trigger (HLT) algorithm. However, it is expected that in a future more developed reconstruction implementation, an algorithm performing 3D clustering directly (i.e. not as part of a two-step procedure) may achieve better performance (Section 10.1.2).

5.1.4.1 Superclustering

The presence of between 0.5 and 0.8 X_0 of tracker material in front of the calorimeter results in many electromagnetic showers starting before the first active calorimeter layer. The strong magnetic field spreads the charged particles, and thus electrons and photons generally result in more than a single particle entering the calorimeter. Thus, except in the case of photons which do not convert in the tracker material, electrons and photons generally give rise to a number of distinct clusters of energy in the calorimeter spread in ϕ by the field, but with little spread in η. (The spiral path of very low p_T charged particles causes a little spread towards higher absolute η.) These distinct multiclusters are collected together into a "supercluster", in a similar way as in the reconstruction performed on the data collected in the current (2017) CMS detector. Indeed, the reconstruction algorithms and particle-flow machinery into which the multiclusters are fed will, in the future, need significant tuning and optimization in order to use the HGCAL to best advantage.

With the fine longitudinal and lateral granularity, hadron showers also tend to result in more than a single multicluster. In this case the multiplicity is due to the inhomogeneity and irregularity of hadron showers, and the spread is in both lateral dimensions, being dominated by the transverse development of the shower, rather than by the magnetic field. However, the development of an algorithm sophisticated enough to perform general reconstruction of

hadron showers needs more work, and the clustering of hadronic showers for use in the CMSSW particle-flow reconstruction of jets is done with some assistance from Monte Carlo truth information for the pattern recognition. The results for hadron resolution given below use a simple algorithm that performs satisfactorily for $p_T > 10\,\text{GeV}$, even in the presence of high pileup.

5.1.5 Electron and photon identification

The lateral and longitudinal shower shapes, provided by the fine longitudinal and lateral granularity of the HGCAL enable powerful discrimination between signal and background, for both electrons and photons. In the study described below they are used, together with track-cluster matching, to build an electron-ID discriminant.

The first step in the construction of the shower shape variables is the determination of the axis, and average position, of the component multiclusters of an electron or photon using a principal component analysis (PCA) method. Variables characterizing the lateral and longitudinal shape with respect to this axis are then computed. The variables include the PCA axes eigenvalues, RMS along each axis, shower start position, and shower length. Figure 5.4 shows two distributions of two variables sensitive to the shower longitudinal development, and Fig. 5.5 shows distributions of two variable related to the lateral shower development: the shower spread in the radial direction and the difference between the shower position measurement and the extrapolated track location in η. In the figures, the shapes found for electron candidates from different sources are compared. The vertical scale is logarithmic. The distributions resulting from jet events with pileup corresponding to 200 interactions per bunch crossing ("QCD multijets" in the plot legends), are seen to differ from the distributions resulting from electrons in $Z \to ee$ events, both with and without pileup.

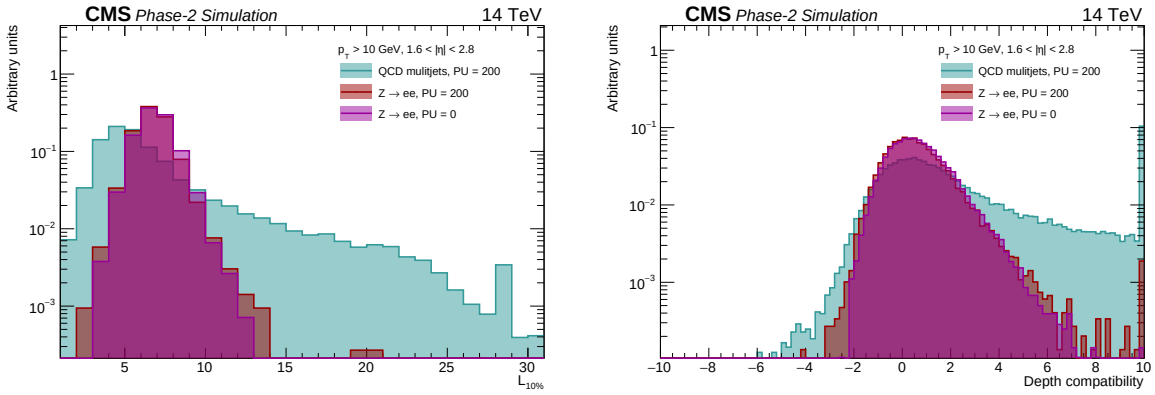


Figure 5.4: Two examples of variables sensitive to the shower longitudinal development: layer number for which the accumulated energy reaches 10% of the EE energy (left); shower depth compatibility (right). The distributions are shown for signal electrons and background candidates in presence of pileup, and for signal electrons in the absence of pileup.

The resulting variables are combined in a boosted decision tree (BDT) trained on a $Z \to ee$ sample accompanied by pileup for the signal, and with multijet events for the background. Since the variables used evolve with the p_T, the training of the BDT is done separately for $10 < p_T < 20\,\text{GeV}$ and for $p_T > 20\,\text{GeV}$. The results obtained using the multivariant electron-identification discriminant given by the BDT, are shown in Fig. 5.6. For a 95% signal efficiency, the background efficiency is 1% (10%) for $p_T > 20\,\text{GeV}$ ($10 < p_T < 20\,\text{GeV}$). Also shown are the results using BDTs with reduced sets of input variables: starting the track-based variables (black histogram), a significant gain in performance is seen by adding the energy-momentum

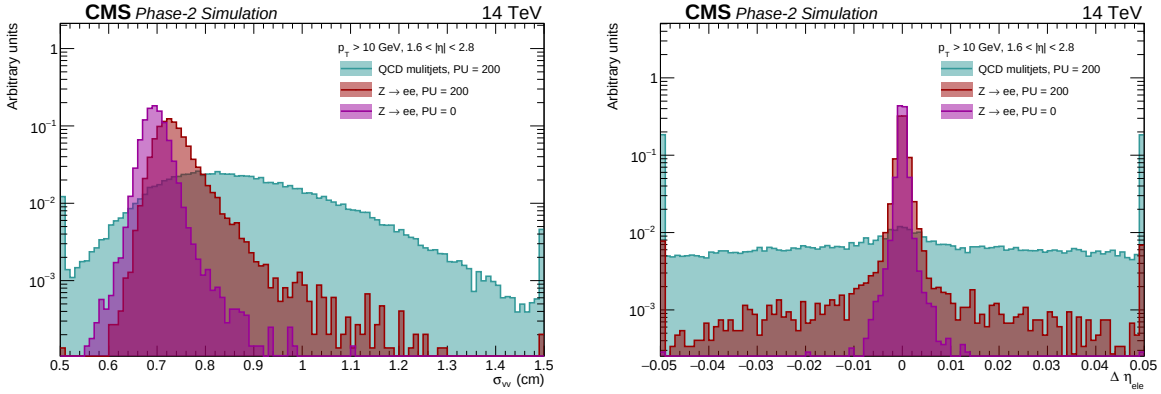


Figure 5.5: Shower spread along the radial direction (left) and distance in η between the electron cluster and the track extrapolation (right). The distributions are shown for signal electrons and background candidates in presence of pileup, and for signal electrons in the absence of pileup.

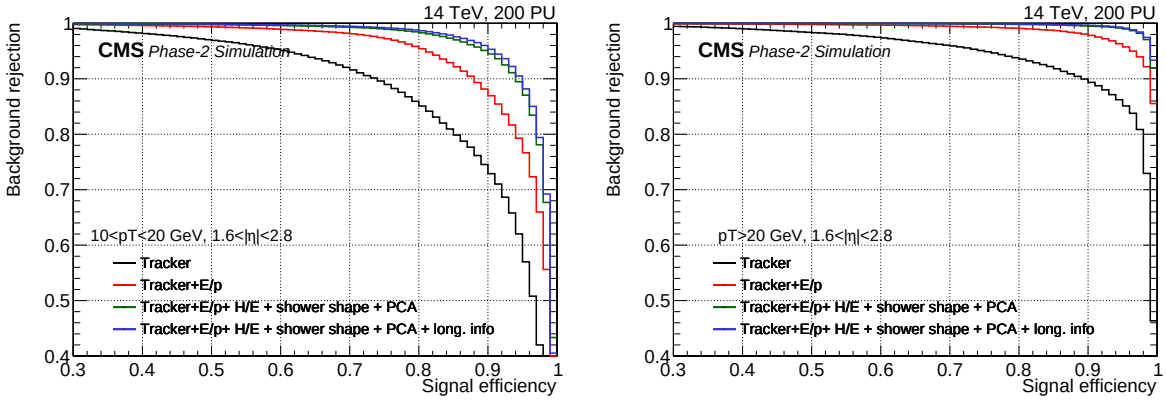


Figure 5.6: Purity as a function of the efficiency for electrons with $10 < p_T < 20\,\text{GeV}$ (left), and $p_T > 20\,\text{GeV}$ (right), for different sets of input variables used for the multivariate electron-identification discriminant.

comparison (red histogram). The addition of the PCA-based lateral shape variables (green histogam) brings further improvement in performance. Finally, the addition of the variables related to the longitudinal development (blue histogram), improves the performance for low p_T electrons (Fig. 5.6 (left)).

5.1.6 Hadron showers

An algorithm has been developed to reconstruct high p_T hadron showers ($p_T \geq 7\,\text{GeV}$) in the presence of pileup. This algorithm is not intended to be final, and it seems clear that more sophisticated developments (Section 5.1.10) will improve the energy resolution performance and extend the range of applicability to lower p_T.

Two dimensional layer clusters are built with the imaging algorithm (Sections 5.1.4 and 10.1.2), the 3D multiclusters constructed from these layer clusters are used to seed the algorithm. A truncated cone is defined about an axis formed by the line between the mean interaction point, (i.e. the origin of the CMS coordinate system), and the position of the seeding multicluster. The 2D layer clusters whose positions lie within the truncated cone are then collected and constitute the hadron cluster ("megacluster"). The energy is estimated by summing the cell

energies within 6 cm of the maximum cell of each of the constituent layer clusters.

For pileup resulting from an average of 200 interactions per bunch crossing the optimum truncated cone is found to have a 3 cm radius in the first layer of the calorimeter, and a radius of 11 cm at the back of FH. The contribution to the energy sum resulting from pileup is subtracted. This is done by computing the energy summed in a megacluster, using the same truncated cone and sum of layer clusters algorithm, located at the same pseudorapidity, but at $\phi + \pi$ in azimuth, where ϕ is the azimuthal location of the megacluster. The energy thus summed is taken to represent the pileup contribution and subtracted from the pion megacluster energy.

Figure 5.7 shows the reconstructed energy distribution for $p_{\mathrm{T}} = 25\,\mathrm{GeV}$ charged pions with and without pileup. Figure 5.8 shows the energy resolution for charged pions as a function of p_{T}, in the case of no pileup, and for a pileup resulting from a mean of 200 interactions per bunch crossing. The width of the distribution, used for the resolution, is calculated as σ_{eff}, half the width needed to contain 68.3% of the distribution.

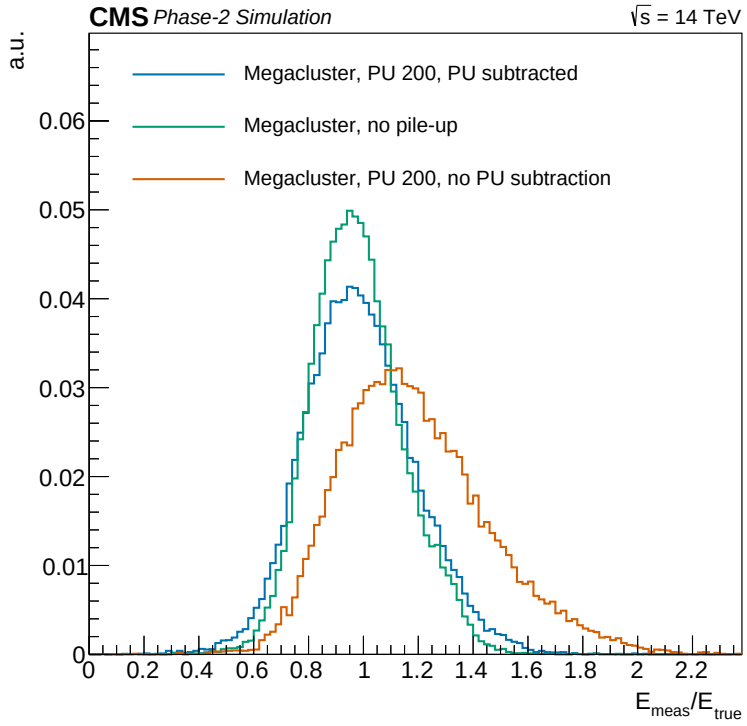


Figure 5.7: Distribution of $E_{\mathrm{meas}}/E_{\mathrm{true}}$ (the reconstructed energy divided by the true energy) for $p_{\mathrm{T}} = 25\,\mathrm{GeV}$ charged pions with (red histogram) and without (green histogram) pileup corresponding to an average of 200 interactions per bunch crossing. The distribution with pileup is also shown after pileup subtraction (blue histogram).

5.1.7 Jet reconstruction using the calorimeter alone

In order to understand the intrinsic performance of the high granularity calorimeter we have made studies using standard jet algorithms to collect the energy of quark jets directly from the individual reconstructed hits. The distance parameter, R, of the anti-k_{T} algorithm is varied into account the effects of pileup.

Figure 5.9 (left) shows the transverse energy deposited per unit area, in η, ϕ space, for different pileup conditions. The events examined consist of pure pileup, and the small amount of energy

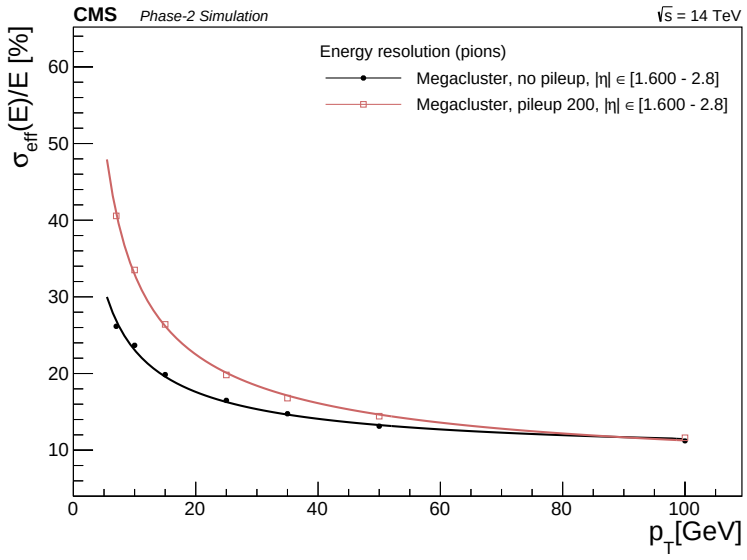


Figure 5.8: Energy resolution of charged pions, reconstructed with the megacluster algorithm described in the text, as a function of p_{T}, in the case of no pileup (solid circles), and for a pileup resulting from a mean of 200 interactions per bunch crossing (empty squares).

visible in the zero pileup case is due to zero-suppressed noise. It can be seen that a substantial transverse energy, E_{T}, is deposited in a unit area: in an event with a mean of 200 pileup interactions per bunch crossing, this amounts to about 200 GeV. A larger fraction of the energy in jets reconstructed from pure pileup, as compared to energy in quark jets, is deposited in the earlier part of the calorimeter, as pileup is, in general, composed of softer particles. This can be seen in Fig. 5.9 (right) where the transverse energy reconstructed as having been deposited, layer-by-layer, in the calorimeter in quark jets of $E_{\mathrm{T}} = 50\,\mathrm{GeV}$, $2 < |\eta| < 2.5$, is compared to that in jets from pileup. The quark jets are generated using the GEANT particle gun, and hadronized with PYTHIA. The lateral energy containment as a function of radius can be seen in Fig. 5.10 for the electromagnetic (EE), the first 12 (FH) and the last 12 (BH) layers of the hadronic sections of the calorimeter. The narrowness of the quark jets as compared to those created by pileup is clearly evident, especially in the electromagnetic section.

The transverse energy fraction contained in areas of various size for quark jets and for jets made from pileup is compared in Fig. 5.11. The jets are reconstructed using the anti-k_{T} algorithm with $R = 0.8$, and the E_{T} fraction contained in concentric circles in η, ϕ space is plotted as a function of the radius. In the left plot the containment is shown for different quark jet p_{T}, and in the right plot quark jets with $p_{\mathrm{T}} = 50\,\mathrm{GeV}$, $2.0 < |\eta| < 2.5$, are compared to jets made from pileup interactions.

From Figs. 5.9 and 5.11 it is evident that to define an optimal distance parameter for best calorimetric energy resolution, a balance has to be struck between the loss of resolution due to integration of pileup energy and that due to loss of energy out of the region. Furthermore the value of the optimal distance parameter can change from one longitudinal layer to another.

Taking account of these observations the energy resolution of jets resulting from the VBF process (i.e. VBF jets), has been studied using the anti-k_{T} algorithm, with differing distance parameters, R. The expected average contribution from pileup, on an event by event basis, is estimated using a ρ parameter and the jet area as inputs, and can be subtracted from the reconstructed jet. The ρ parameter is the median E_{T} per unit area in the event, calculated at the same

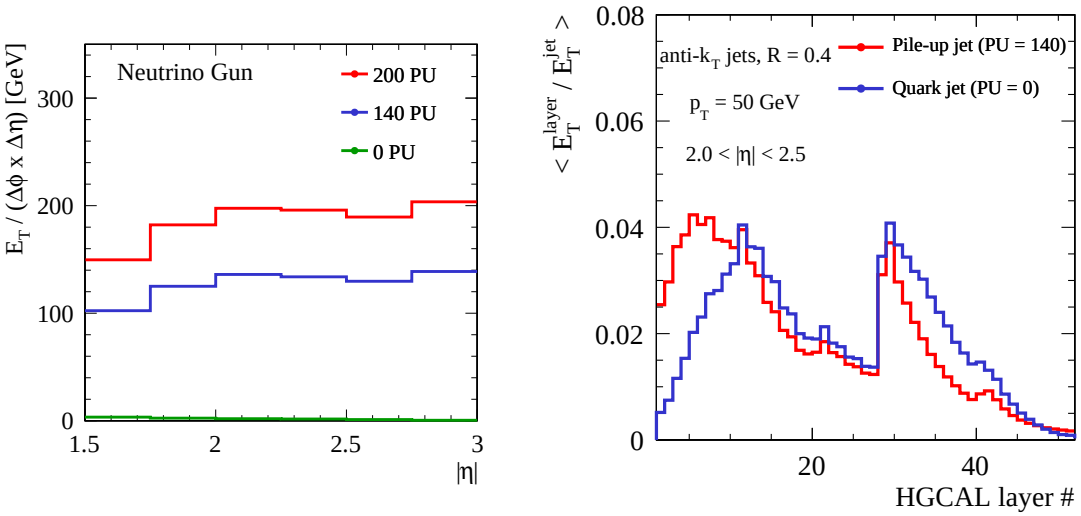


Figure 5.9: Pileup characteristics: (left) transverse energy per unit area in η, ϕ space as a function of pseudorapidity for events with no pileup (green), a mean of 140 (blue), and 200 (red) pileup interactions per bunch crossing; and (right) transverse energy deposited as a function of calorimeter layer, by quark jets (blue-outlined histogram), and by jets reconstructed from pileup (red-outlined histogram). The jumps in energy after the 28^{th} and 40^{th} layers are due to increases in absorber thickness, resulting in corresponding changes to the energy deposited.

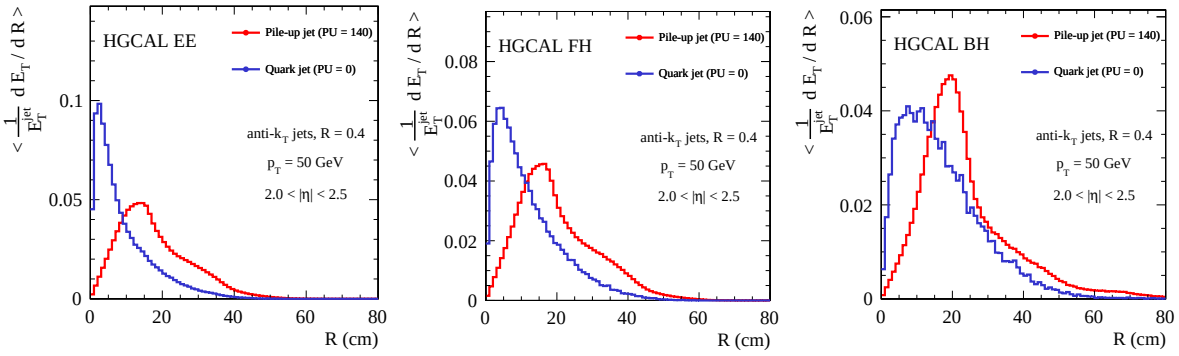


Figure 5.10: Radial profile of quark jets (blue histogram), and jet made from pileup (red histogram) in (left) the 28 electromagnetic layers (EE), (middle) the first 12 hadronic layers (FH), and (right) the last 12 hadronic layers (BH) of the calorimeter.

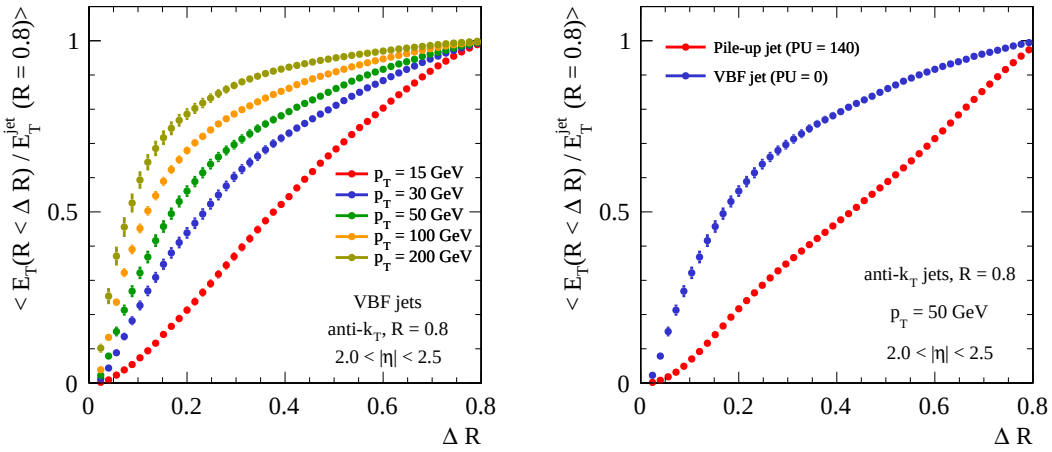


Figure 5.11: Transverse energy fraction contained as a function of distance, ΔR, in the absence of pileup, (left) for quark jets of various p_{T}, and (right) a comparison of quark jets and jets made from pileup interactions.

pseudorapidity as the jet. Figure 5.12 (left) shows the response, the ratio of the reconstructed energy over the generated energy of a VBF jet, for different values of R, and p_{T}. It can be seen that the ak4 algorithm (i.e. the anti-k_{T} algorithm with distance parameter $R = 0.4$) picks up a large amount of pileup energy if no ρ correction is applied. This energy can be subtracted, to a large extent, by applying the ρ correction (Fig. 5.12 (right)) but at the price of introducing a substantial nonlinearity. However, the ak2 ($R = 0.2$) algorithm, working on a much smaller area, picks up much less pileup energy and does not appear to benefit much from the application of the ρ correction. It is very encouraging to note that the stand-alone energy resolution, using the ak2 algorithm, is already sufficiently good in an environment of high pileup. The ak2 algorithm gives almost the same energy resolution with or without the rho correction, and is sufficiently immune to pileup, as seen in Fig. 5.13 which shows the energy resolution found, as a function of p_{T}, with and without pileup when different distance parameters are used in the anti-k_{T} algorithm.

It should be noted that the fine lateral and longitudinal granularity allows a dynamic definition of the axis of the jet (since there are no large size towers) and of R to suit the local environment of the particular jet being measured. No optimisation has been carried out, nor use made of the detailed longitudinal or lateral information available.

We have explored the selectivity achievable for VBF initiated events using the HGCAL alone, without requirements from any other detector, in pileup conditions corresponding to a mean of 140 interactions per bunch crossing. The jets are required to lie within $1.6 < |\eta| < 2.8$. The rates as a function of jet E_{T} are shown in Fig. 5.15 (left) for single jets, double jets, double jets where the two jets are required to be in opposite endcaps, and double jets with the requirement, in addition to those previously listed, on the dijet invariant mass $m_{\mathrm{jj}} > 350\,\mathrm{GeV}$. A further selection reduces the rate to that labelled PU jet ID (orange coloured points). The PU jet ID uses a likelihood ratio discriminant built from two variables which are shown in Fig. 5.14: $Iso(1,4)$, the ratio of energy found when the jet is reconstructed using the anti-k_{T} algorithm with $R = 0.1$ to that found using $R = 0.4$, and E_{10}/E_{jet}, the fraction of the jet energy in the first 10 layers. The selection efficiency can thus be tuned by varying the requirement on the discriminant. The rate versus the efficiency for jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ is shown in Fig. 5.15 (right). It can be seen that a

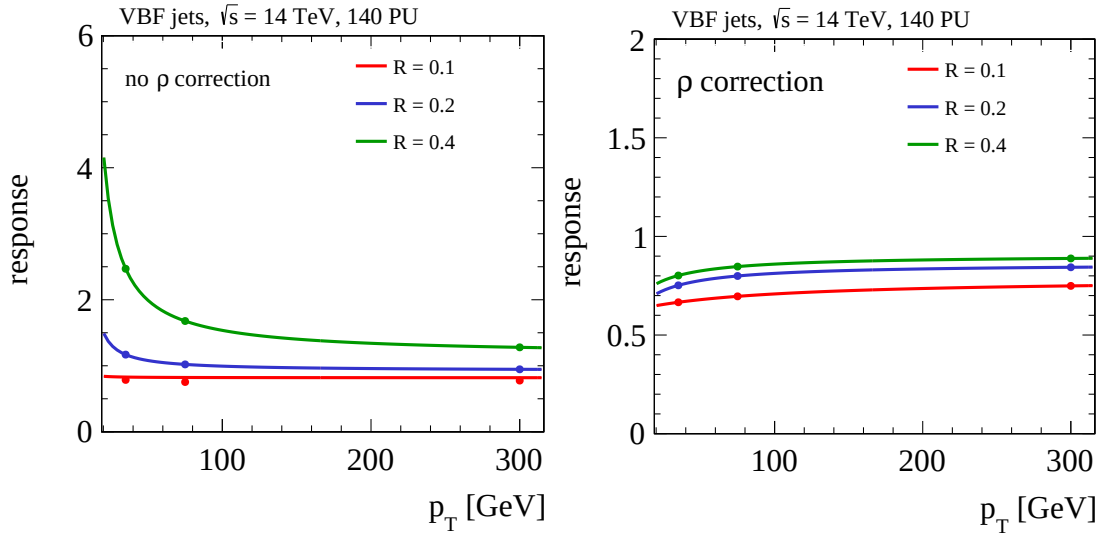


Figure 5.12: Jet energy response, the reconstructed E_{T} divided by that generated, as a function of generated jet p_{T} for jets ($1.8 < |\eta| < 2.8$) reconstructed using anti-k_{T} algorithm with three different values of R; with (left) no ρ correction applied, and (right) after applying a ρ correction.

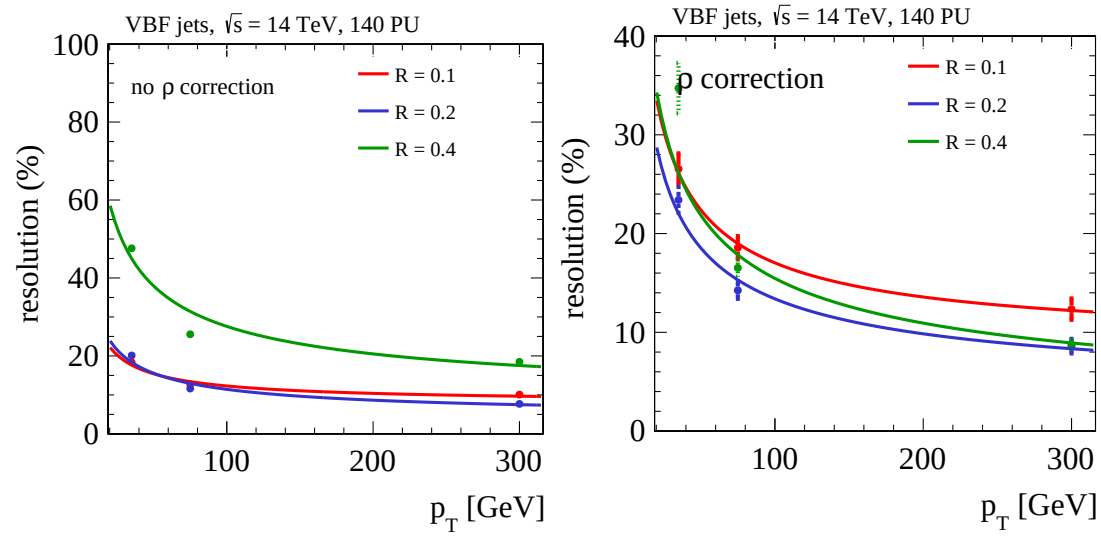


Figure 5.13: Jet energy resolution as a function of generated jet p_{T} for jets ($1.8 < |\eta| < 2.8$) reconstructed using anti-k_{T} algorithm with three different values of R; with (left) no ρ correction applied, and (right) after applying a ρ correction.

rate of around 10 kHz can be obtained for a PU jet ID efficiency of 80%. The rate calculation is illustrative of the possibilities and uses all the HGCAL layers.

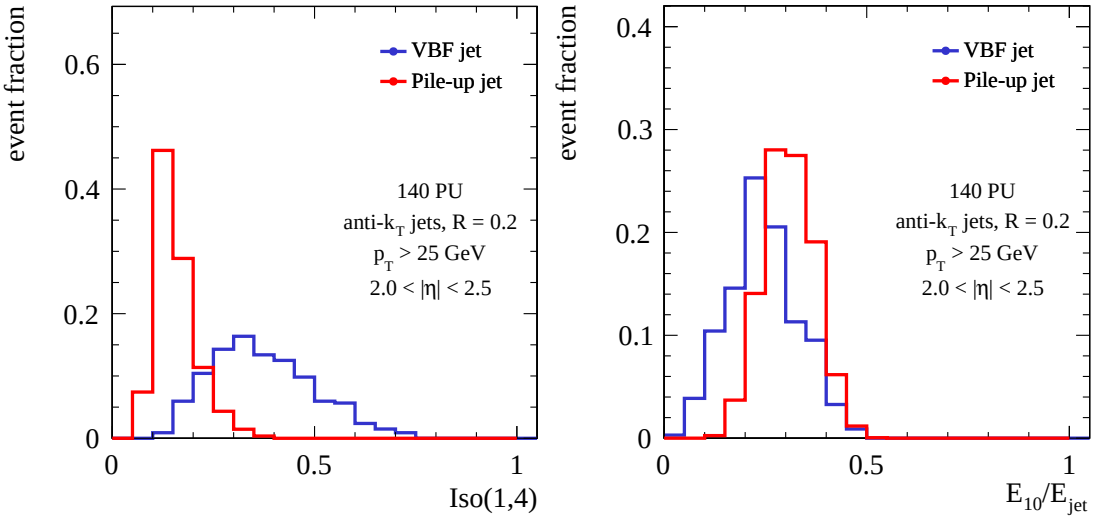


Figure 5.14: Distributions of the two variables used to build the PU jet ID discriminant for both VBF jets (blue histograms), and jets found in pileup (red histograms).

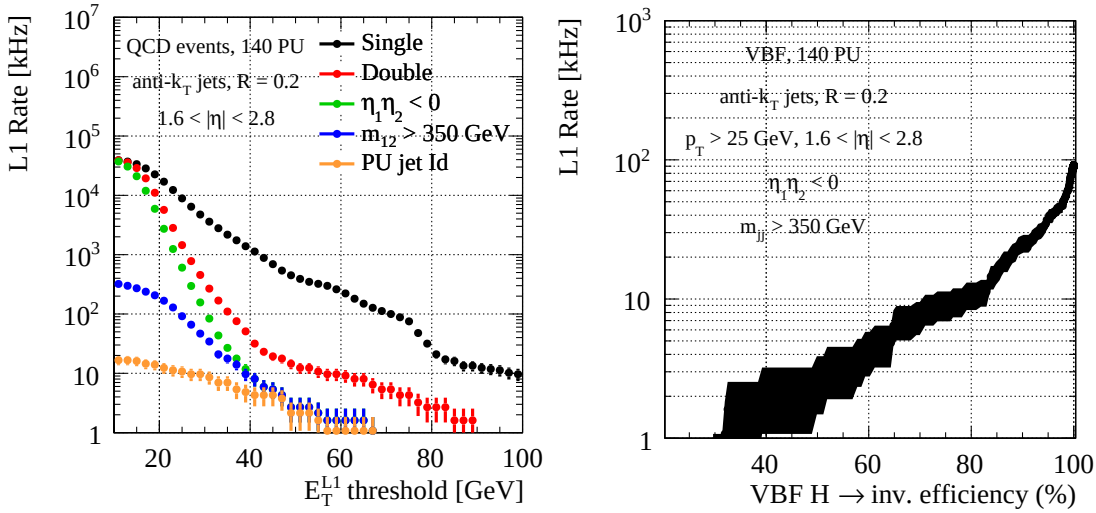


Figure 5.15: Rate of ak2 VBF dijets events (left), and the efficiency of the PU jet ID versus rate for jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$.

5.1.8 Tau identification using calorimeter alone

The identification of hadronically decaying tau leptons (τ_{h}) using only the HGCAL has been studied. The main background to τ_{h} comes from the misidentification of jets. We attempt to quantify the degree to which τ_{h} can be distinguished from jets by the HGCAL, where the candidates are reconstructed with the anti-k_{T} algorithm directly from the individual reconstructed hits, as in the previous section. The τ_{h} decay products are expected to be more collimated than typical jet constituents, and discrimination can be achieved by considering the lateral distribution of energy about the jet axis.

The performance is studied using a simulated sample of $\tau_h \to \rho(770)\nu_\tau \to \pi^+\pi^0\nu_\tau$ decays, this being the most common (40%) of the hadronic decay modes. A dijet sample is used to model the background.

Jets are built from the individual reconstructed HGCAL hits using the anti-k_T algorithm with a distance parameter of $R = 0.2$, as described in the previous section. An isolation discriminator is formed using the ratio of the scalar E_T sum in the annulus $0.1 < \Delta R < 0.4$ about the jet axis to the sum within $\Delta R < 0.1$. It is necessary to subtract the expected contribution from pileup from the isolation energy sums. This is achieved by subtracting the product of the event-by-event energy density estimator ρ and the area covered in η, ϕ by each of the regions.

Figure 5.16 (left) compares the distribution of this ratio for τ_h candidates with $30 < p_T < 60\,\text{GeV}$ and $1.8 < |\eta| < 2.4$, for the cases of no pileup (continuous outline histograms), and of pileup corresponding to 140 interactions per bunch crossing (dashed outline histograms). The distributions from τ_h decays (blue) are compared to those originating from jets (red). For τ_h decays the ratio is typically small: 98% have a value smaller than 0.5, whereas candidates originating from jets 85% have a value greater than 0.5. Figure 5.16 (right) shows the τ_h efficiency versus jet rejection that is achievable by making a selection on this ratio. For an efficiency of 60% a misidentification rate of 0.9% (3%) is achieved in no pileup (pileup corresponding to 140 interactions per bunch crossing).

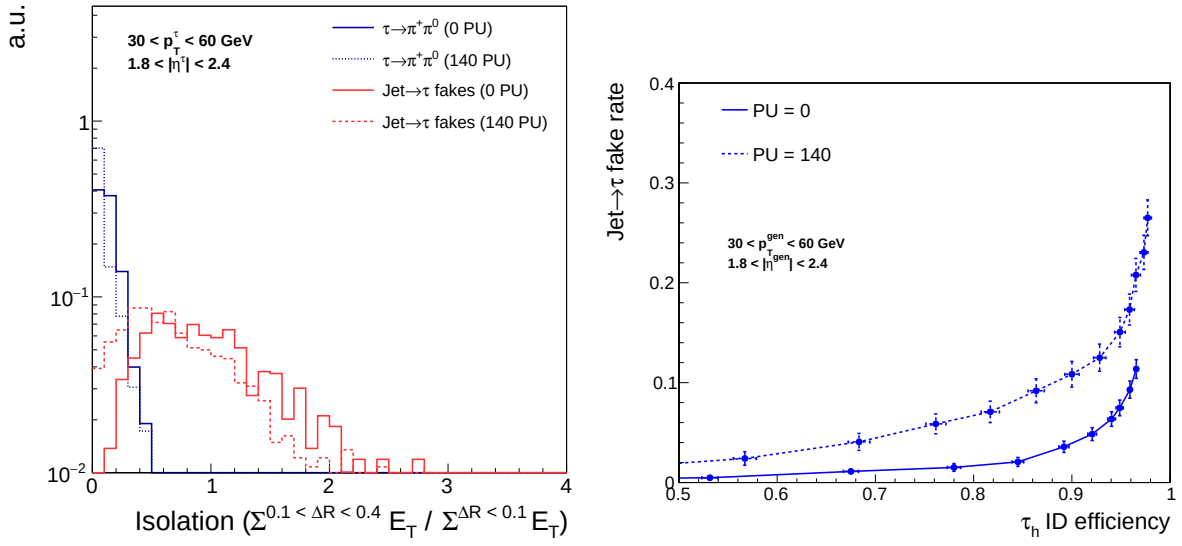


Figure 5.16: Identification of hadronic tau decays; (left) isolation discriminant for τ_h candidates with $30 < p_T < 60\,\text{GeV}$ and $1.8 < |\eta| < 2.4$, and (right) τ_h efficiency versus jet rejection using this discriminant.

5.1.9 Muon identification

The region $|\eta| > 2.4$ is covered exclusively by silicon sensors. In this region, where the ME1/1 chambers are not available but tracking coverage is provided by the HL-LHC pixel tracker, a good efficiency for muon identification can be provided by the HGCAL, in addition to the new ME0 muon chambers that cover the pseudorapidity range $2.0 < |\eta| < 2.8$.

Detailed studies have shown that the HGCAL can achieve good muon identification performance by placing requirements on the signals in cells in a small region around an extrapolated charged-particle track. The track is extrapolated from the CMS tracker by propagating its mo-

mentum vector in the magnetic field. For each layer of the calorimeter the region examined may either be the single cell through which the extrapolated track passed, or that cell plus all its neighbours. The number of layers with muon-like signals in the region is counted, where a muon-like signal requires that the region contain between 1 and 3 hits with energy between a lower threshold, t_μ, and an upper threshold of 3 MIP, for $2.0 < |\eta| < 2.4$. For $2.4 < |\eta| < 2.8$, between 1 and 4 are required. The summed energy of the cells in the region is required to be less than 3 MIP (4 MIP) for $2.0 < |\eta| < 2.4$ ($2.4 < |\eta| < 2.8$), respectively. The low threshold is taken as $t_\mu = 0.5$ MIP by default, but the sensitivity of the result to increasing this threshold is studied. Requirements are then placed on the number of layers with such signals in the 12 rear-most layers, N_{BH}, in the 12 layers preceding that, N_{FH}, and in the layers of the electromagnetic section, N_{EE}. As might be expected, the variable that has the most discrimination power between a muon and other charged particle tracks is N_{BH}. By contrast, the number of layers with muon-like signals in the front half of the electromagnetic section has no measurable discrimination power.

Figure 5.17 shows the interplay between muon identification efficiency and the probability per event that a nonmuon track is identified as muon. The result is obtained from an event sample with a mean of 200 pileup interactions per bunch crossing, and with a transverse momentum requirement of $p_{\mathrm{T}} > 5$ GeV on the track. The variable N_{BH} is scanned to produce the efficiency versus fake rate curves, after the other variable parameters in the algorithm have been adjusted to give a plateau efficiency of ≈98% (the ultimate limit on efficiency comes from the tracker efficiency, which is more than 99% for $p_{\mathrm{T}} > 5$ GeV muons). The plot shows the curves obtained with detector noise conditions with values corresponding to the start of operation. The efficiency versus background rate is shown for two samples of $p_{\mathrm{T}} > 5$ GeV muons. One sample, labelled isolated muons, are single muons generated flat in p_{T} with a GEANT particle gun, the other, labelled $\mathrm{t\bar{t}}$-events, are muons with $p_{\mathrm{T}} > 5$ GeV found in a $\mathrm{t\bar{t}}$ sample. No isolation criteria have been applied to the muons in the $\mathrm{t\bar{t}}$ sample, and the resulting reduced efficiency is explained by this fact. In both samples pileup corresponding to an average of 200 interactions per bunch crossing is present. The background rate is that due to 200 interactions per bunch crossing. These results are obtained with a very simple system of uncorrelated selection requirements, and it is expected that use of more powerful multivariant techniques will provide improved performance. Further details, and results, including a discussion of the negligible background due to additional signals from slow neutrons, are given in Section 10.1.7.

5.1.10 Future developments

It has not been possible, so far, to explore the full range of reconstruction possibilities offered by the fine granularity of the calorimeter. The possibility of direct three-dimensional clustering has already been mentioned. Another important methodology is the use, in the overall pattern recognition, of aligned traces of minimum-ionizing particles, which often connect separate denser sub-showers resulting from a single incoming hadron. Exploitation of these connecting traces has been shown by the CALICE collaboration to be important in achieving the best hadron shower reconstruction [27, 28, 31]. Additionally, the use of timing information to assist in the separation of adjacent and overlapping showers at the level of clustering needs to be investigated.

The effort has been mainly directed towards first understanding how best to use the calorimeter information on its own, and the subsequent combination of this information with charged particle tracks reconstructed in the Tracker has used existing code, untuned and often ill-suited to the new calorimeter. It is likely that a more integrated, and probably iterative, approach will eventually give the best results; a prototype investigation extending the tracking of charged

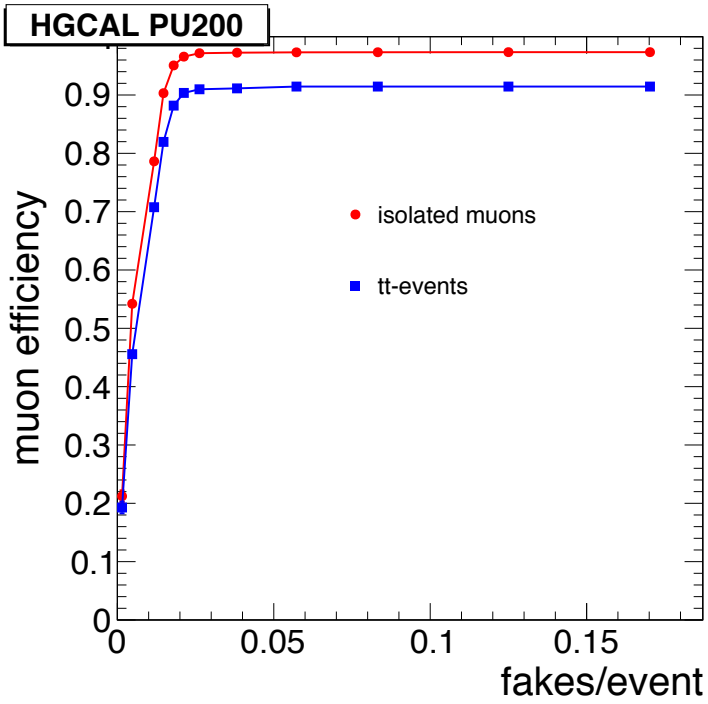


Figure 5.17: Muon identification efficiency versus the probability per event that another, non-muon, track is identified as muon. The curves are obtained using S/N conditions at the start of operation, and pileup corresponding to an average of 200 interactions per bunch crossing is present.

hadrons into the calorimeter, using the 2D clusters in the calorimeter as if they were hits in a tracking detector, was able to collect most of the energy deposited by the shower, suggesting interesting possibilities.

The reconstruction will also benefit from the use of multivariant energy regression techniques, first to estimate the energies of individual clusters for purposes like their matching to tracks, and then, later, to estimate the energy of the reconstructed electrons, photons, charged hadron and neutral hadrons.

5.2 Beam tests

In order to validate the basic design of the HGCAL, a series of beam tests was performed in 2016 and 2017. A paper to be submitted to JINST is in preparation.

In 2016, sixteen hexagonal silicon detector modules were produced, using an existing front-end ASIC (SKIROC2 [32]) with a two-layer module PCB for flexibility (more details are given in Section 10.2.3). These modules were used to make comparisons with simulation and test basic calorimetric performance for electromagnetic showers. Tests were performed at both FNAL and CERN, using a flexible mechanical support structure that allowed multiple configurations to be studied. At FNAL, all sixteen SKIROC2-based modules were available, arranged as double-sided layers interspersed with tungsten absorbers. The total thickness of the setup at FNAL was deliberately limited to about 15 X_0 due to the relatively low energy electron beams available at FNAL (maximum momentum 32 GeV). At CERN, only eight modules were available in 2016 but two configurations were explored, having the modules placed between about 6 and 15 X_0, and 5 to 27 X_0, respectively.

For 2017 the main aim was to use an ASIC better suited to CMS requirements, the SKIROC2-CMS [33] evolution of SKIROC2, and to expand the setup to include hadronic sections (CE-H), with a front section based again on silicon sensor modules and back section using a modified version of the CALICE AHCAL [14]. This latter uses 12 layers of steel absorber (totalling 5 λ) interspersed by planes of scintillator tiles with on-tile SiPMs for readout. Several configurations were tested, with up to 20 SKIROC2-CMS-based modules. The final test of 2017 featured five modules in the CE-E section and the remainder in 7 planes of CE-H, including one plane with 7 modules and one plane with 3 modules. Figure 5.18 shows a photograph of the 7 silicon modules mounted on a copper support/cooling plate. More details on the construction are given in Section 10.2.3.

Figure 5.19 (left) shows typical energy spectra for incident charged pions (termed "MIPs" in the following) in the 2016 silicon modules in the CE-E section. The signal-to-noise (S/N) for a single MIP in the silicon modules is between 7 and 8 for all measured cells, except for the dedicated calibration pads (smaller than the regular cells), which had S/N around 10. Also shown in the same figure is the AHCAL spectrum for single MIPs, from the 2017 run, following a minimal track-stub selection, selecting single particles within hadron showers. The plot is a combination of signals from all cells that were hit, following a preliminary calibration.

A good measure of the matching between simulation and data, for both electromagnetic and hadronic showers, is the transverse shower spread. This was evaluated using ratios between the energies in central cells and their neighbours. Figure 5.20 shows the ratio $E1/E7$ in data and simulation for incident 100 GeV electrons at a depth of about 10 X_0, and the ratio $E7/E19$ for incident 200 GeV charged hadrons, at a depth of around 4 λ. The energy sums $E7$, and $E19$ are the sums of energy in the most energetic cell, $E1$, plus those in the surrounding rings of six,

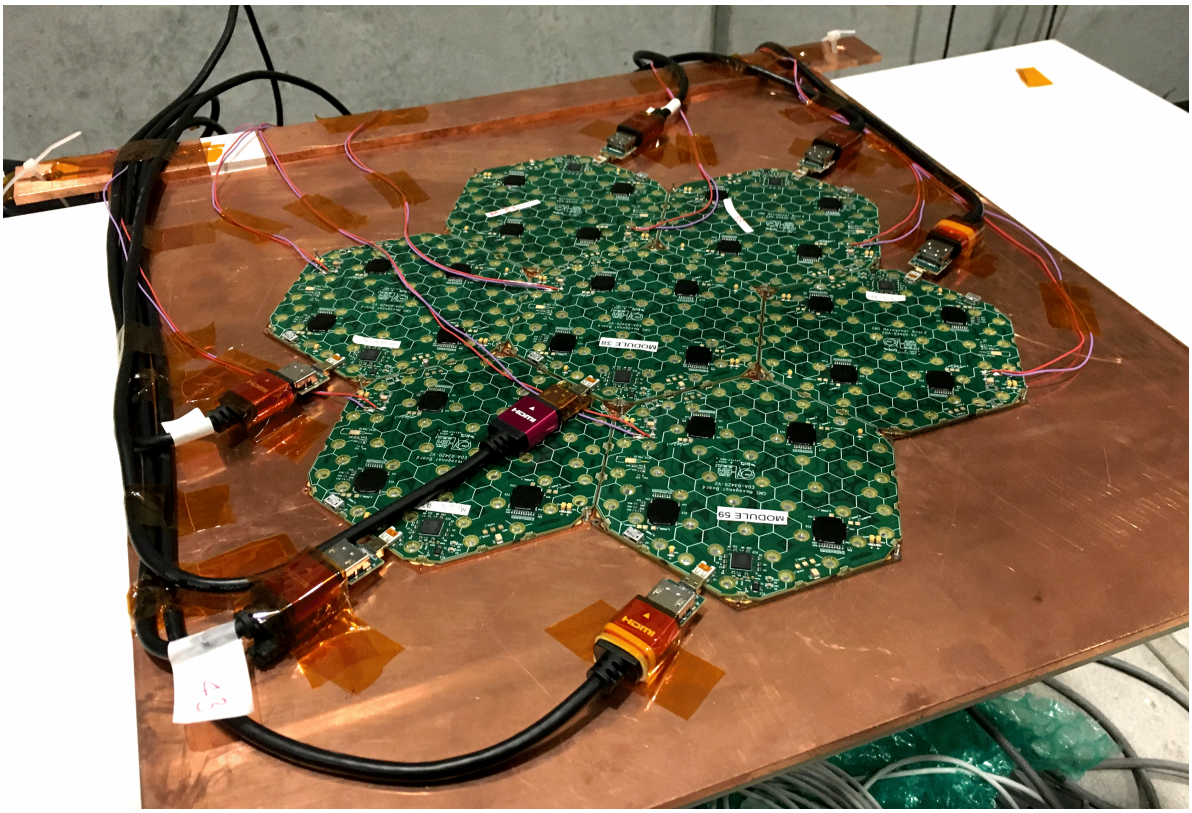


Figure 5.18: Seven 6" silicon sensor modules, equipped with SKIROC2-CMS ASIC, mounted on a copper support/cooling plane during the 2017 beamtest campaign.

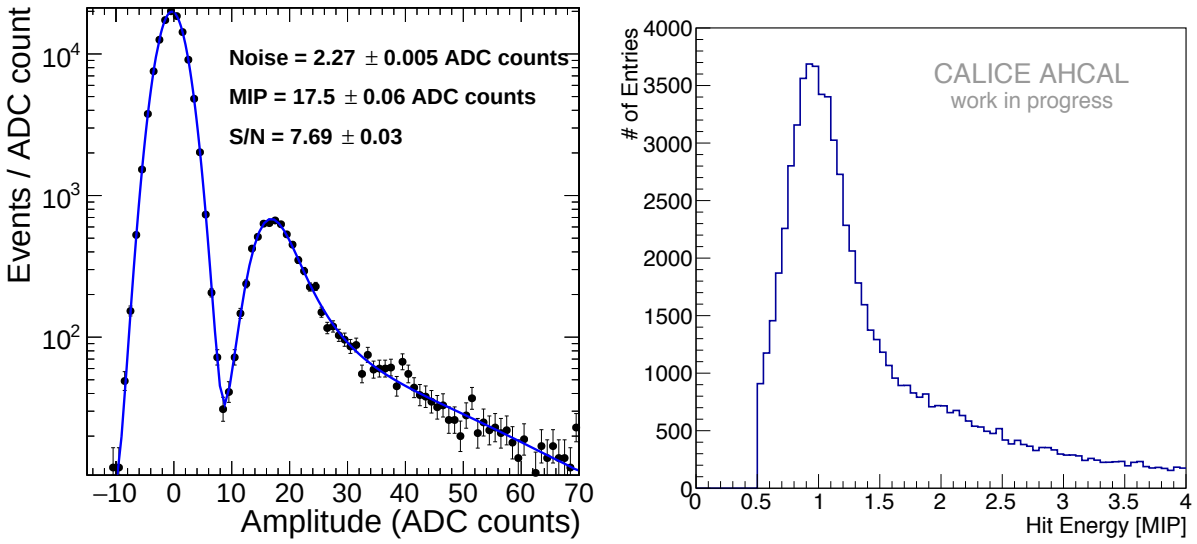


Figure 5.19: Amplitude spectrum (left), measured in ADC counts, of signals in a typical silicon cell in 2016 from incident charged pions in the CE-E section. The amplitude is fitted with a sum of a Gaussian (for the pedestal) and a convolution of a Landau distribution and a Gaussian (for the signal); and (right) spectrum of signals, measured in MIPs, seen in the 864 cells of the AHCAL in 2017, due to single charged particles in hadron showers, following a simple tracking and preliminary calibration.

or nine, cells. The agreement between data and simulation for both shower types is very good.

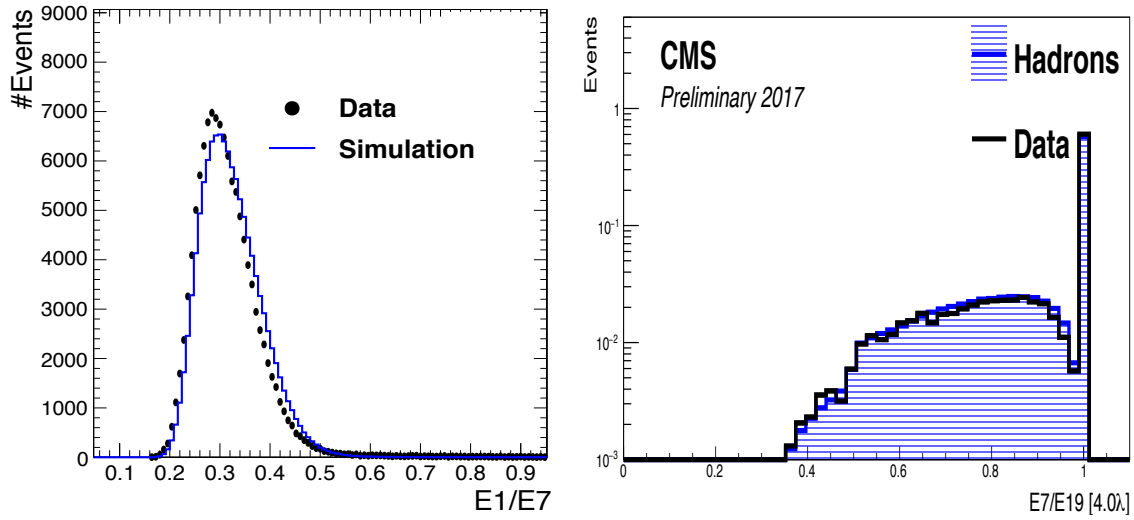


Figure 5.20: Distribution (left) of the ratio $E1/E7$ for 100 GeV electrons at a depth of around $1\,X_0$ for data (points) and simulated showers (histogram); and (right) transverse shower profile ($E7/E19$) measured at a depth of $4\,\lambda$ from incident 200 GeV charged hadrons (20% pions; 80% protons).

A sum of the values of $E7$ in all layers, with a weighting derived from the absorber thicknesses, was then used as an estimate of the total beam energy. Figure 5.21 (left) shows the energy resolution measured in 2016 at FNAL (16 layers, 0–15 X_0) and CERN (8 layers, 5–27 X_0). It can be seen that for low-energy electrons the configuration at FNAL performed better than that at CERN, as expected. The limited number of samplings, and, for higher energy electrons, the limited overall depth, clearly limit the energy resolution. However, the data matches that predicted by the simulation within ≈5% for all energies, demonstrating the accuracy and scalability of the simulation.

A log-weighted barycentre was also calculated and compared with the expected incidence position given by upstream wire chambers (DWC). The log-weighting method provides a simple method of correcting the bias resulting from using linear weighted mean to determine the position of the shower. The measured residual width of the x-coordinate as a function of electron energy, is shown in Fig. 5.21 (right), both for data and for simulated showers. The DWC contribution to the uncertainty is quite large, but after including it in the resolution, the agreement between data and simulated showers is very good, as is shown in the figure, and shows that the intrinsic spatial precision at this depth is around 0.6 mm for high-energy electrons.

We also performed a series of tests to demonstrate the intrinsic timing capabilities of silicon sensors, using $5 \times 5\,\mathrm{mm}^2$ single diodes and 6″ hexagonal sensors. Diodes of both p-on-n and n-on-p were used, of three thicknesses (100, 200, 300 μm), both irradiated (up to the expected fluences at $3000\,\mathrm{fb}^{-1}$) and unirradiated. They were operated at $-30\,^{\circ}\mathrm{C}$ to avoid reverse annealing of the irradiated diodes. The full sensor was operated at room temperature because it had not been irradiated. A dedicated fast digitizer was used in both cases to measure the pulse shape arising from signals in the silicon due to incident electrons on upstream absorbers. The time difference between pairs of diodes or neighbouring cells was used to estimate the timing precision. For the full sensor we used a variety of incident electron energies. In Fig. 5.22 the resolutions of the time difference for the 300 μm p-on-n diodes and full sensors are shown, as

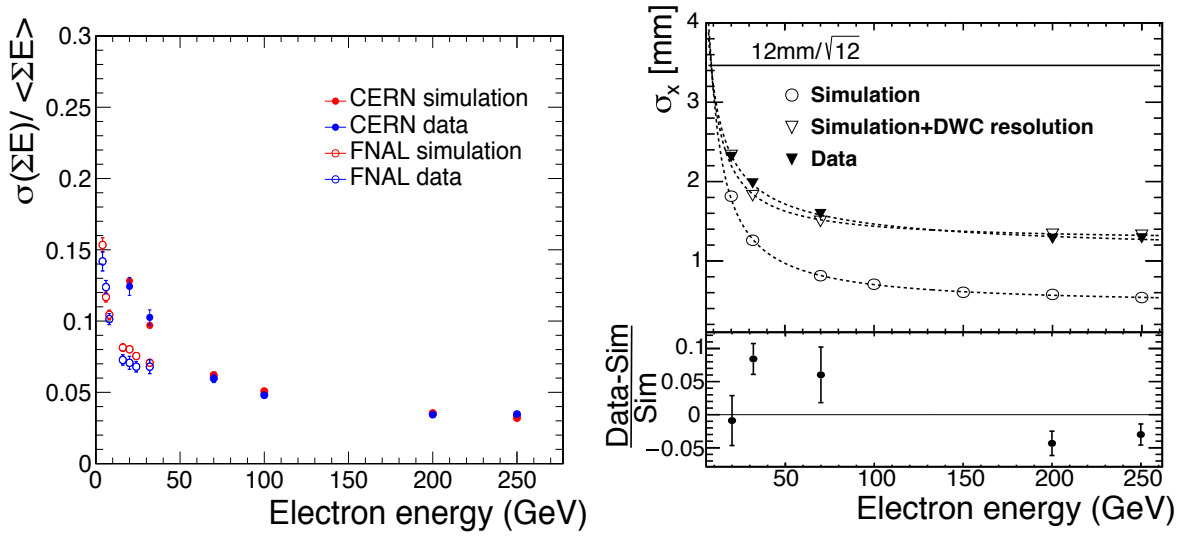


Figure 5.21: Energy and position resolution: (left) relative energy resolution as a function of the electron energy in data and simulated showers, for test beams at FNAL and CERN; and (right) residual width of the x-coordinate reconstruction at a depth of $6\,X_0$ as a function of incident electron energy.

functions of the effective signal-to-noise ratio $(S/N)_{\text{eff}}$, defined by:

$$(S/N)_{\text{eff}} = \frac{(S/N)_{\text{ref}}(S/N)_{\text{n}}}{\sqrt{(S/N)^2_{\text{ref}} + (S/N)^2_{\text{n}}}} \tag{5.1}$$

where $(S/N)_{\text{ref}}$ is the signal-to-noise ratio in a reference cell, and $(S/N)_{\text{n}}$ is the signal-to-noise ratio in the cell under study. The plots shows that the intrinsic timing resolution does not significantly depend on the fluence at a given S/N ratio, and is better than 20 ps for $S/N > 100$, for diodes and full hexagonal sensors. The n-on-p diodes showed very similar performance.

During the second half of 2018 we aim to equip a prototype mimicking the full CE-E (28 layers) and twelve CE-H silicon layers, with at least four hexagonal modules equipped with the SKIROC2-CMS-based ASIC, totalling at least 76 modules. The CALICE AHCAL will again complement the silicon sections to provide the required hadronic depth. Not only will this provide calorimetric performance, but it will also allow the time evolution of hadronic showers to be measured to a precision of the order of 100 ps. Further prototype modules, using different variants of sensor size and granularity coupled to new PCBs will go through extensive laboratory-based tests and go to beams for tests with particles when appropriate. When new ASICs become available, these will also be used, as will the rest of the final powering and readout chain. Since beams at CERN will not be available in 2019/2020, facilities at FNAL and DESY will be increasingly valuable for further tests.

5.3 Trigger simulation and performance

5.3.1 TPG primitives

As described in Section 3.2, the output of the TPG, i.e. the trigger "primitives", will be a list of 3D clusters and an η, ϕ transverse energy map for each endcap. The algorithms used to produce

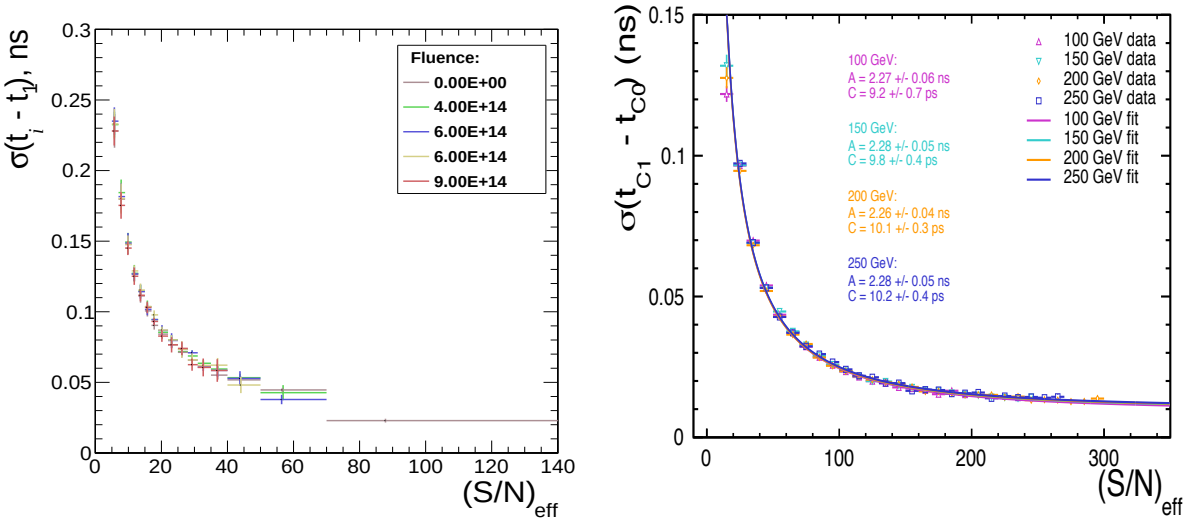


Figure 5.22: Left: Resolution of the time difference between unirradiated and irradiated diodes as a function of $(S/N)_{\text{eff}}$ for 300 μm thick n-on-p diodes; Right: Resolution of the time difference between neighbouring hexagonal cells of a 6″ 300 μm thick p-on-n hexagonal silicon sensor, also as a function of $(S/N)_{\text{eff}}$.

these primitives from the trigger raw data sent from the on-detector FE electronics will run in FPGAs mounted on the off-detector BE TPG boards. In particular for the 3D clusters, there are many possible algorithms that could be implemented in the hardware. Therefore, there will be a continuous development to optimise these algorithms up to and beyond the start of HL-LHC. Many characteristics of the 3D clusters are also calculated, such as fractions of energy in the CE-E and CE-H, shower widths and depths, internal structure such as local maxima (indicating multiple overlapping particles), etc.

The TPG does not produce a trigger decision itself. The primitives will be used as input to the central L1 correlator, which defines and creates the actual triggers for the experiment. The definitions of the triggers used will also be developed throughout the next decade. As such, the information currently planned to be stored for each cluster (as described in Section 8.3) will evolve, as more is learned of the best cluster-related quantities to use. The results below are based on a snapshot of the current assumptions for the primitives.

5.3.2 HGCAL-only trigger performance

One measure of the TPG primitives performance is to study possible model trigger algorithms that could be formed using only the TPG information. The final L1 triggers will use significant extra information, as the central L1 correlator has data from the tracker, the barrel calorimeters, and muon systems to enhance the selectivity. The triggers described here examine the performance of the 3D cluster primitives, since the energy map primitives have not yet been integrated into trigger algorithms. The algorithms presented provide an indication of the resolution, efficiencies and background rates achieved in the absence of the extra information. The results are presented using object thresholds similar to those listed in the trigger menu detailed in Table 4.1 of the L1 trigger Interim Document [21]. Trigger algorithms are not part of the TPG system, and thus an estimation of their firmware implementation and latency requirements is not reported here.

The design and constraints of the TPG primitives have a large impact on electromagnetic ob-

ject (e/γ) energy resolution and identification. Although electrons can be distinguished from photons in the central L1T correlator by matching to L1 tracks, for triggers using only HGCAL information, they can be treated very similarly. Systematic small differences in their estimated energies are neglected in the following.

An electron or photon (e/γ) is reconstructed from a single TPG 3D cluster. A TPG-specific energy correction factor is applied to correct for the energy lost by the $2\,\mathrm{MIP_T}$ cut and the clustering algorithms. The cluster is required to pass a selection based on its transverse and longitudinal shape. The longitudinal profile selection uses the length of the cluster, the position of the start of the cluster and the position of the cluster layer with the maximum energy. These three variables and their correlations are used in a boosted decision tree to select the clusters that are compatible with an electromagnetic shower. The transverse shape is measured by the width of the cluster in the radial direction. Widths are computed in each individual layer included in the cluster, weighted according to the energy in the layer and combined. It results in an average radial width, which is used to identify electromagnetic objects. In order to reduce the pileup sensitivity of this transverse width, only the trigger cells in the core of the cluster are used. This cluster core is defined as a cone of radius $\Delta\rho = 5\,\mathrm{cm}$, corresponding to two rings of trigger cells. The selection is chosen such that the total efficiency loss resulting from the cuts on these variables is approximately 2% for $|\eta| < 2.7$, while tighter cuts giving up to 10% efficiency loss are used for $|\eta| > 2.7$ to reduce the higher background in this region.

To study efficiency, electrons are generated with a transverse momentum between 10 GeV and 100 GeV, with pseudorapidity $1.7 < |\eta| < 2.9$. Figure 5.23 shows the electron efficiency after all the cluster identification cuts are applied and for a L1 threshold $E_{\mathrm{T}}^{\mathrm{L1}} > 30\,\mathrm{GeV}$, as a function of the generated electron p_{T}. The quoted threshold is the electron transverse energy which gives 50% efficiency. The plateau efficiency for high energy electrons well above threshold is around 98% with the selection used here. This shows that single 3D clusters based on the current TPG algorithms can give good electromagnetic performance, with a sharp rise around the threshold and a high efficiency above it.

Defining a single electron trigger using the e/γ selection outlined above, and with no isolation criteria, gives the trigger rate as a function of the trigger threshold shown in Fig. 5.23. This is sensitive to the high levels of pileup in the high $|\eta|$ region and so is shown separately for L1 objects reconstructed with $|\eta^{\mathrm{L1}}| < 2.7$ and for the full $|\eta^{\mathrm{L1}}|$ range. This rate is calculated for an average of 200 interactions per bunch crossing and 2808 colliding bunches. The rates shown are high but correspond to a high signal efficiency and further criteria (isolation, track matching, etc.) are expected to be applied in the central L1 correlator.

Another critical trigger issue for the HGCAL is the jet performance. Hadronic triggers enable important aspects of the HL-LHC physics programme including vector boson scattering (VBS) and VBF signatures, dark matter searches with jets and missing transverse energy, and searches for new physics in hadronic final states. As detailed in Section 10.3.3, standalone HGCAL jets are formed from multiple 3D clusters using an anti-k_{T} algorithm. The lateral granularity can be used to exploit the dense core of jets to mitigate the effect of pileup. A jet size of around $\Delta R = 0.2$ seems to give the best balance between reduction of pileup contamination and minimizing fluctuations due to inadequate jet containment. An average correction for the pileup within the jet is made as a function of η, where the corrected values are estimated from cones of size $\Delta R = 0.2$ in simulated events containing only pileup. A additional correction is made for the energy lost in the jet reconstruction, parameterized as a function of p_{T}. These lead to hadronic jet efficiencies and background rates as shown in Fig. 5.24.

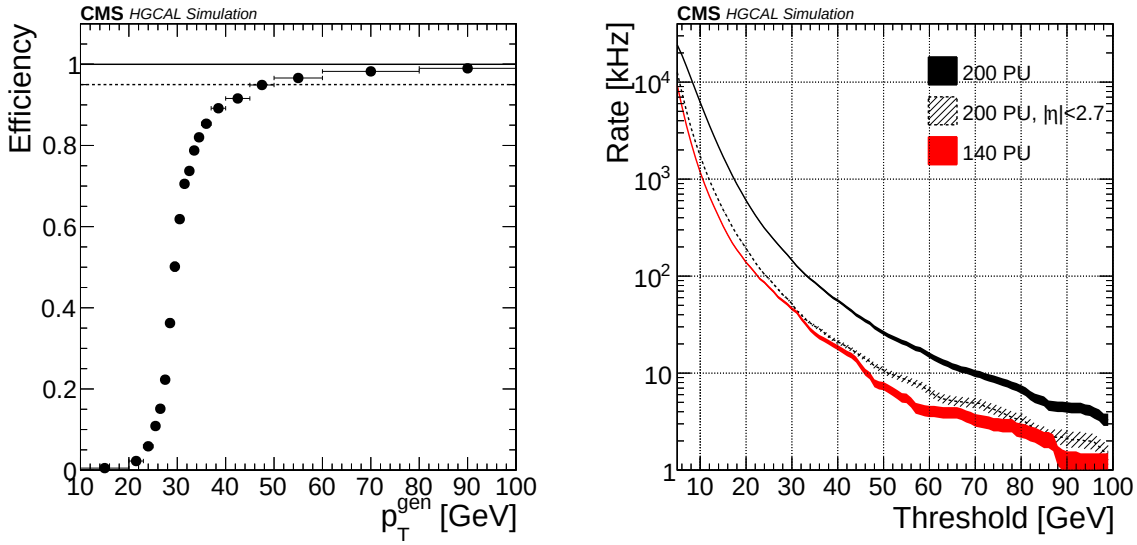


Figure 5.23: Left: Turn-on curve for a single e/γ trigger with a central transverse energy threshold of 30 GeV. Right: L1 e/γ trigger background rates for 140 and 200 pileup as a function of the trigger threshold. For the 200 pileup case, the rate is also shown for e/γ objects in the restricted range $|\eta^{L1}| < 2.7$.

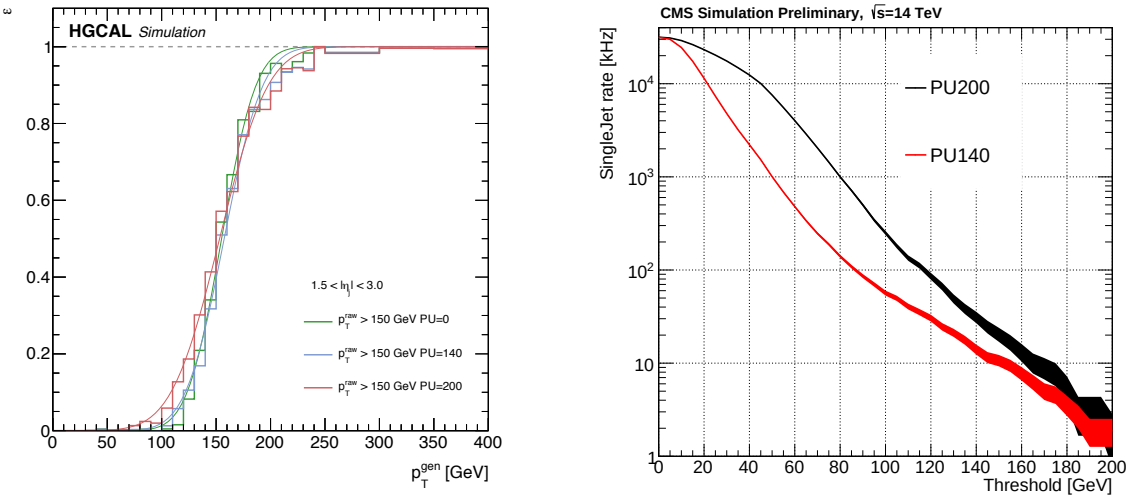


Figure 5.24: Left: Turn-on curves for a single jet trigger for a central transverse energy threshold of 150 GeV, with various average pileup levels. Right: single jet trigger background rates for 140 and 200 pileup as a function of the trigger threshold.

5.4 Calibration and monitoring

Uncertainty in the channel-to-channel intercalibration contributes to the constant term of the energy resolution. In the case of electromagnetic showers, investigation has shown the contribution of intercalibration uncertainty to the constant term to be approximately six times smaller than the intercalibration uncertainty. Since we aim for an electromagnetic energy resolution with a constant term smaller than 1%, we target an intercalibration precision of 3% which will contribute less than 0.5% to the constant term.

The primary method of establishing the sensor intercalibration and maintaining it over time, following the slow changes in charge collection efficiency over the duration of HL-LHC opera-

tion, will use signals from minimum-ionizing particles.

By employing a MIP-tracking algorithm it is possible to obtain a clear MIP peak from simulated events at the highest values of instantaneous luminosity. The MIP-tracking algorithm requires MIP-like hits in the cells in the layer before and the layer after the cell in question (3-layer algorithm). The peaks obtained for different cells can be fitted to obtain relative calibration values as can be seen in Fig. 5.25. In the figure the peak of the Landau distribution is denoted MPV (most probable value). Local isolation, requiring that all sensor cells surrounding the cell under study have a signal <0.5 MIP, reduces the sensitivity of the fitted peak value to the instantaneous luminosity. The black-outlined colour-filled histogram, underneath the red-outlined histogram except for the first few bins, shows the distribution in selected events. This distribution is fitted by the blue curve. The red-outlined histogram shows only those hits, among those selected, where the signal results from energy deposited in the silicon, rather than from noise. (The determination of which hits, in fact, result from energy deposited in the silicon is made using MC truth information). It can be seen that the selected events represent a rather pure signal selection. The 3-layer algorithm works well down to a S/N for a MIP of 4 or 5.

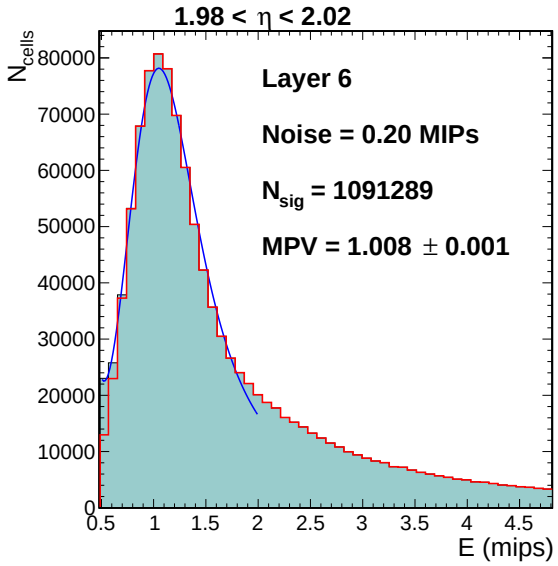


Figure 5.25: MIP signal peak in sensor cells with noise equivalent to 0.2 MIP. The sensor cells are located in the region $1.98 < |\eta| < 2.02$ and have an active thickness of 200 μm. In the simulated events the mean number of interactions per bunch crossing is 200. The distributions is shown after the 3-layer MIP tracking and local isolation algorithms has been applied. The MIP S/N ratio is 5. Further details shown are explained in the text.

Figure 5.26 shows the distribution in sensor cells with noise equivalent to 0.4 MIP (i.e. $S/N = 2.5$) located in the region $1.98 < |\eta| < 2.02$. The plot on the left shows the distribution (black-outlined grey-green filled histogram) resulting from the 3-layer algorithm, and it can be seen that the lower part of the spectrum is dominated by noise. As in the earlier figure, the red-outlined histogram, obtained using MC truth information, shows only those hits where the signal results from energy deposited in the silicon, rather than noise. Tightening the MIP-tracking algorithm by requiring track confirmation signals in two layers before and two after the sensor under study (5-layer algorithm) removes the contribution from instances where there is, in fact, no MIP present. The resulting distribution is shown in Fig. 5.26 (right), where it can be seen that the selected signals (black-outlined colour-filled histogram) correspond very closely to the

events where the signal is results from energy deposited in the silicon (red-outline histogram), rather than noise.

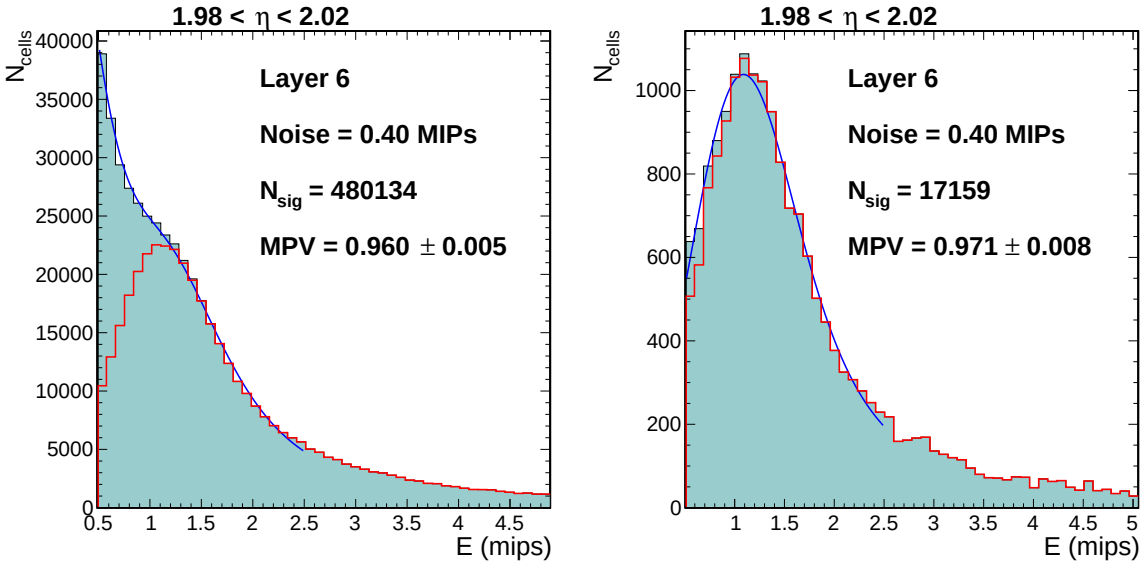


Figure 5.26: MIP signal peak in sensor cells with noise equivalent to 0.4 MIP. The sensor cells are located in the region $1.98 < |\eta| < 2.02$ and have an active thickness of 200 μm. In the simulated events the mean number of interactions per bunch crossing is 200. The distributions are shown (left) after the 3-layer MIP tracking and local isolation algorithms is applied, and (right) when the 5-layer MIP tracking is used. The various details shown are explained in the text.

Fitting the selected signals with a function representing the signal (Landau $\times$ Gaussian) delivers a measure of the MIP peak. The fit is shown in Fig. 5.26 (right) by the blue curve. The statistical uncertainty on the most probable value, which gives the measure of the MIP peak, is 0.8%. The events correspond to pileup at an instantaneous luminosity giving a mean of 200 interactions per bunch crossing. The number of these events required to give the number of entries in the histogram in Fig. 5.26 (right) for each cell is ≈165M. Thus any ≈50M recorded events would provide a statistical precision of 3% for all cells having $(S/N)_{\mathrm{MIP}} \geq 2.5$.

Further work has explored what is needed to obtain a clean MIP peak for $(S/N)_{\mathrm{MIP}} = 2$, a level of noise worse than is expected anywhere in the detector after 3000 fb^{-1}. Simply tightening the requirements on the two cells in front and the two cells behind appears to work well, but it is rather inefficient, requiring a further increase in the number of events needed by a factor of 2.5 to produce a peak that is fitted with an uncertainty somewhat larger than the $(S/N)_{\mathrm{MIP}} = 2.5$ case. If it was ever necessary to contend with such a level of noise it would probably be better to exploit a more sophisticated selection algorithm, for example a selection algorithm driven by reconstructed charged hadrons or muons.

A possible way of performing the MIP calibration would be to build histograms, such as shown in Fig. 5.26 (right), in the HLT farm accessing the full Level-1 data sample. Thus an intercalibration for cells requiring 50M events would be performed in a little over 1 minute. This assumes the aim, specified elsewhere in this document, that even after 3000 fb^{-1} the readout threshold can be maintained at 0.5 MIP. If for the highest noise sensor cells, and for the highest luminosities, this low readout thresholds became intolerable, then histograms for these cells would be taken (for example) later in a fill after the luminosity had decayed somewhat, allowing the

readout thresholds to be lowered.

The noise performance of the front-end electronics will enable MIP signals to be seen and fitted in almost every sensor cell of the HGCAL even after exposure to the hadron fluence from an integrated luminosity of 3000 fb^{-1}. In addition, for redundancy, and to fully guarantee the ability to achieve MIP calibration throughout the life of the calorimeter, dedicated low-capacitance/low-noise cells will be included on each wafer. For these calibration cells a standard hexagonal cell will have a central hexagonal subcell in addition, with a seventh of the area of the standard one, thus providing cells with a signal-to-noise ratio >5 on each ASIC of the HGCAL, even after the accumulation of the full lifetime hadron fluence.

The electronics chain of each channel will be independently monitored and linearized using a charge injection system based on a chopper circuit and fixed calibration capacitances connected to the front-end input. The additional noise contribution from inclusion of this system is negligible because of the chopper series resistance. The large range of injectable charges (0–10 pC) will allow a finely detailed linearity measurement. Before installation, the channel testing and qualification protocol will include measurement of the charge injection capacitance to $<1\%$. Knowledge of the charge injection capacitance will allow an independent verification of the MIP calibration at startup, and, when used in conjunction with the MIP calibration, will monitor the charge collection efficiency throughout the lifetime of the detector.

According to a potential supplier the average thickness of all wafers will be contained within $\pm 5\,\mu$m of nominal, and within a wafer the diffusion depth of all pads (defining the active thickness) will be within $\pm 3\,\mu$m of the average for the wafer. This translates to an effective Gaussian spread of the diffusion depth between the pads in a wafer of 0.6% (1.7%) for wafers with 300 (100) μm active thickness, respectively. As part of the test and qualification protocol for wafers the depletion depth of cells will be characterized by voltage/capacitance measurements with a precision of better than 1%. A number of full-depth segments of the HGCAL be put into test beams to calibrate the responses to electrons and hadrons before startup.

5.5 Precision timing

Precision timing represents a major addition to event reconstruction in the HL-LHC. For the HGCAL, timing provides the ability to measure with high precision the time of electromagnetic and hadronic showers. Requiring the compatibility between the measured time of showers will provide a powerful tool to assist pileup rejection, and identification of the primary vertex of the triggered interaction.

The front-end ASIC (HGCROC) provides time-of-arrival (ToA) information as described in Section 3.1, with further details in Section 8.1.2. The precision of the overall timing measurement depends on the intrinsic performance of the sensor plus preamplifier and discriminator, the characteristics of the TDC used for the ToA measurement, and the clock distribution system.

Results from beam tests have shown that the timing resolution obtained with silicon sensors does not vary significantly with sensor thickness when the resolution is measured as a function of S/N [34], nor does it vary as a result of irradiation up to fluences expected after 3000 fb^{-1} when measured as a function of S/N (Fig. 5.22). The resolution, σ_t, can thus be expressed as

$$\begin{aligned} \sigma_\text{t} &= \sigma_\text{jitter} \oplus \sigma_\text{floor}, \\ \text{with } \sigma_\text{jitter} &= \frac{A}{(S/N)}, \end{aligned} \tag{5.2}$$

where S/N is the signal-to-noise ratio, σ_{floor} is a constant term (a precision floor), and the symbol $\oplus$ denotes quadratic summation. The constant A, which determines the dependence on S/N, is fixed by the response time and noise characteristics of the sensor plus preamplifier.

A preliminary study has been made of HGCAL timing performance using the parameterization of the individual cell timing resolution, σ_t, given in Eq. 5.2, with $A = 5\,\text{ns}$ and $\sigma_{\text{floor}} = 20\,\text{ps}$. This represents the achievable resolution on the ToA measurement expected from the specification of the sensors, the front-end electronics, and the clock distribution.

The ToA has a threshold of 12 fC, so we have ToA information only for cells above this threshold. For signals above the threshold of 12 fC, the single cell resolution ranges between 20 and 150 ps, and the combination of measurements from multiple cells allows reduction of the uncertainty on the timing estimate.

The effective timing performance of the HGCAL is determined by the kinematics and topology of the shower development. Electromagnetic showers have small containment radii and fast development, and usually give many cells with significant energy deposits: the shower of a $p_T = 2\,\text{GeV}$ photon at $\eta = 1.7$ has an average of more than 20 cells above a 12 fC threshold, and the corresponding number for $p_T = 60\,\text{GeV}$ is more than 100. Hadronic showers usually have a prompt core and later developing components that propagate laterally with respect to the shower axis, and the number of cells above a 12 fC threshold has large event-to-event fluctuations. The time distribution of the energy deposited has long tails of later time measurements. To make a time measurement we require at least three cells within a selection radius, ρ, of the shower axis, each with an energy deposit of $>$12 fC needed to fire the ToA. Showers that fail this requirement obtain no time measurement, resulting in a measured inefficiency, generally small, that is quantified below. Since the distribution in time of the hits collected for a shower has tails, simple algorithms are used to truncate the tails before taking the mean time of the remaining hits to obtain a shower time measurement. The truncation algorithms are described in Section 10.4.

As we first wish to explore the intrinsic capability of the HGCAL for the determination of the time we have used photons and K^0_L that do not convert or interact in the tracker material. Although charged pions provide a track in the CMS tracker, about whose extrapolated trajectory HGCAL sensor cells could be sought without using a clustering algorithm, it is also necessary to calculate and take into account the additional time-of-flight due to the curvature of the track in the tracker volume. The choice of K^0_L for initial studies was due to the simplicity of the straight trajectory.

The efficiency and time resolution, for photon and K^0_L showers of $p_T = 5\,\text{GeV}$, in the presence of three different levels of pileup, are shown in Fig. 5.27. For photons, with $\rho = 2\,\text{cm}$, the time of the reconstructed shower is measured with 100% efficiency (i.e. there are no showers which fail the three hit requirement described above), and with a resolution of better than $\approx$12 ps for $p_T > 5\,\text{GeV}$, over the full η range. The conditions after $3000\,\text{fb}^{-1}$ have negligible impact on the result. For K^0_L showers there is a trade-off between efficiency and resolution: for $\rho = 5\,\text{cm}$ the efficiency is higher than that for $\rho = 2\,\text{cm}$, but the timing resolution is poorer. For K^0_L of $p_T > 5\,\text{GeV}$, using $\rho = 2\,\text{cm}$, the time of the reconstructed shower is measured with $>$90% efficiency, and with a resolution of $\lesssim$30 ps. The efficency and time resolution obtained for K^0_L showers at $\eta = 1.8$ as a function of p_T is shown in Fig. 5.28.

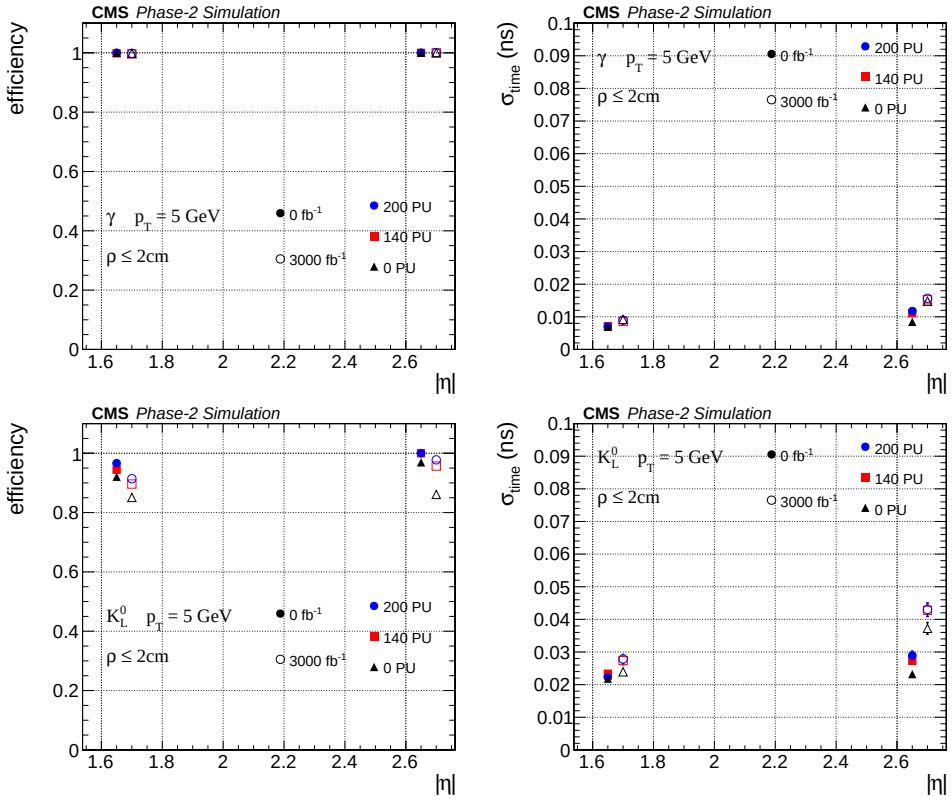


Figure 5.27: Efficiency (left), and time resolution (right), for (top) photon showers, and (bottom) K^0_L showers of $p_T = 5\,\text{GeV}$, measured in the absence of pileup (black triangles), and in the presence of pileup corresponding to a mean of 140 (red squares), and 200 (blue disks) interactions per bunch crossing. The time resolution and efficiency after an integrated luminosity of $3000\,\text{fb}^{-1}$ are also shown, they are indicated by outlined symbols, displaced by 0.05 in $|\eta|$ for clarity of presentation.

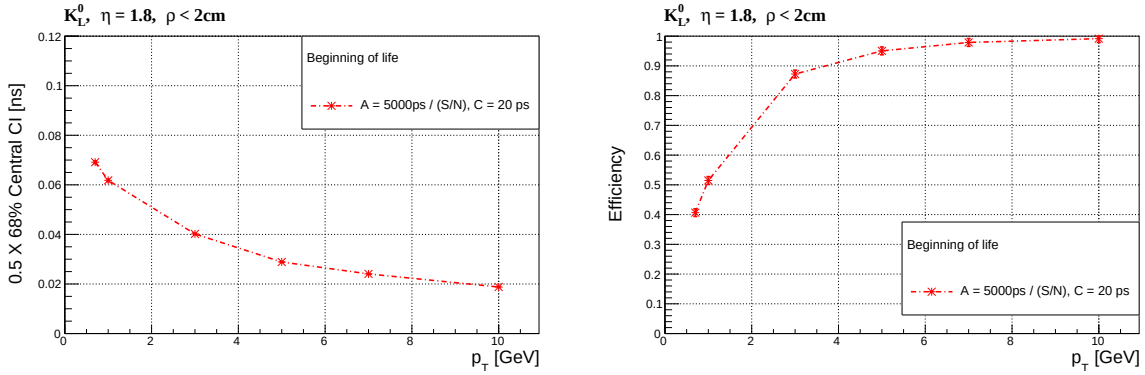


Figure 5.28: Performance for K^0_L showers at $|\eta| = 1.8$ as a function of p_T: timing resolution (left), and efficiency (right), in the absence of pileup.

5.5.1 Pileup mitigation using timing information

The power of the use of timing information for pileup mitigation is illustrated in Fig. 5.29, which shows hit displays of a VBF $H \to \gamma\gamma$ event, with no timing requirement, and after the removal of all hits above a threshold of 12 fC with $|\Delta t| > 90\,\text{ps}$. The density of hits with a charge $> 12\,\text{fC}$, projected to the front face of the calorimeter, is substantially reduced and the photon and a VBF jet can clearly be seen, allowing the development of jet reconstruction and

energy estimation less affected by the presence of pileup.

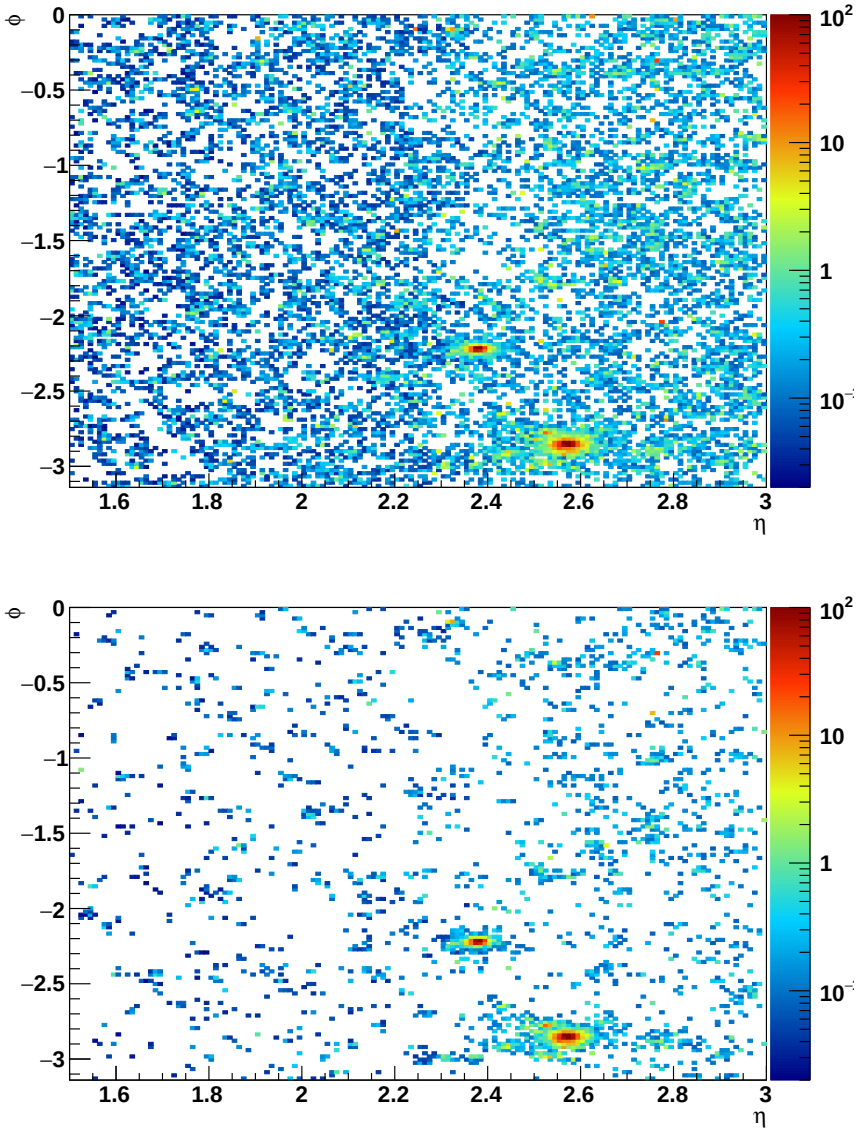


Figure 5.29: Hits with a charge > 12 fC, projected to the front face of the calorimeter: without a timing requirement (upper plot), and after removal of hits with $|\Delta t| > 90$ ps (lower plot).

5.5.2 Vertex location using precision timing

The timing information can be used to localize the vertex of the triggered hard interaction, e.g. the vertex of the production of the Higgs boson that decays into two photons–the most difficult case. To illustrate this case we reproduce below two plots from the MIP Timing Detector Technical Proposal [35].

Owing to the good timing resolution for high energy photons in the barrel and the HGCAL endcap electromagnetic calorimeters, for 50% of the $H \to \gamma\gamma$ events, i.e. those with a pseudorapidity separation $\Delta\eta_{\gamma\gamma} > 0.8$, the vertex is well localized by triangulation (Fig. 5.30 (upper plot)), and the mass resolution is unaffected by the angular component in the evaluation of the mass resolution.

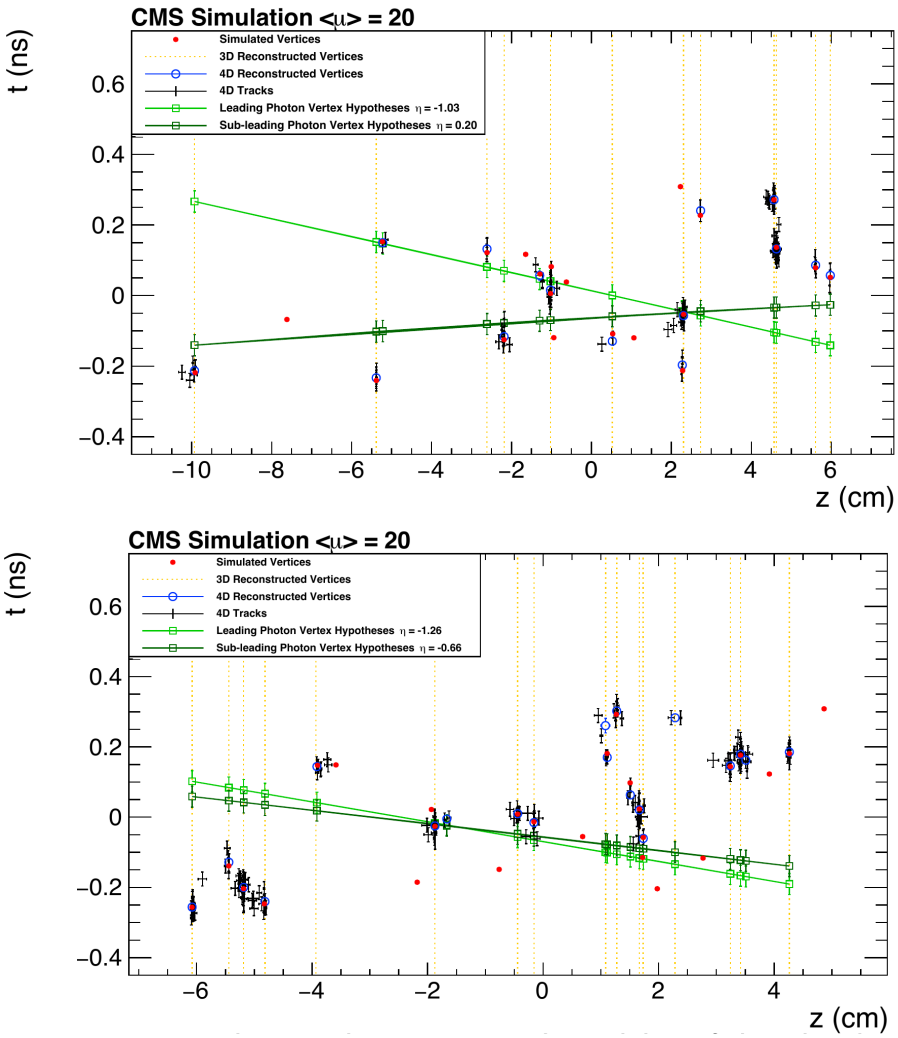


Figure 5.30: Space-time diagrams for H $\rightarrow \gamma\gamma$ events with two photons separated by a large rapidity gap (upper plot) and a small rapidity gap (lower plot). For small rapidity gaps the reconstruction time for the photon at each vertex (green open points) can be cross referenced with the time information of the 4D vertex provided by the MIP Timing Detector to further eliminate all the vertices where the green points overlap at different z-values but at the wrong time. Green lines guide the eye. Events with pileup corresponding to 20 interactions per bunch crossing are shown to improve clarity.

For the other 50% of the H $\rightarrow \gamma\gamma$ events we would use timing from low p_{T} charged tracks. For fifty percent of the events that contain a Higgs boson, there are at least three charged tracks, with $p_{\mathrm{T}} > 2\,\mathrm{GeV}$, that lie within the HGCAL acceptance. On the basis of the information shown in Fig. 5.28 it seems clear that the vertex can be correctly localised for an appreciable fraction of such events.

We estimate that, offline, for around 75% of the events with a Higgs boson the primary vertex can be correctly localized, thus regaining in the high pileup environment a situation similar to that which exists today.

Chapter 6

Schedule, cost and funding

6.1 Organization, funding and institutional responsibilities

6.1.1 Introduction

The HGCAL project was formally created in 2015, as a subsystem in the CMS Phase-2 (HL-LHC) Upgrade. According to the CMS and HGCAL project constitutions, the HGCAL institution board (IB) is the highest decision-making body in the HGCAL project. The HGCAL project (henceforth referred to as the Project), is led by a project manager (PM), who chairs its steering committee (SC) and is responsible for the scientific, technical, and managerial direction of the Project. The PM is appointed by the CMS spokesperson in consultation with the HGCAL IB. As with all CMS upgrade projects, the general oversight is provided by the CMS upgrades coordinator and the technical oversight by the CMS technical coordinator.

All important Project decisions shall be presented by the PM to the HGCAL IB, discussed and ratified usually after a consensus or exceptionally a vote of its members. The HGCAL IB, comprising one representative per participating institution, works in a close partnership with the PM on all aspects of the Project.

Cost, funding, and responsibilities of the Project are presented in this technical design report and will later be defined in memoranda of understanding.

The organigram and structure of the Project is shown in Fig. 6.1. The Project is broken into subproject areas led by coordinators. The overall technical coordination is provided by technical coordinators assisted by the Project office. The SC comprises persons in the unshaded boxes of the organigram.

6.1.2 Organization of the HGCAL project

6.1.2.1 Steering committee

The SC includes the deputy PM, the resources and technical coordinators and their deputies, the Project office coordinators, the Level-2 coordinators and the IB chair and his/her deputy. The SC meets regularly and discusses matters of significance in the Project, relating to safety, progress in all areas (scientific, technical, financial, and managerial). The PM is aided in his/her function by the members of the SC.

The members of the SC are charged to ensure that the best possible HGCAL is built for HL-LHC physics using the resources available to the Project. They will supervise and review the progress and planning of the HGCAL project during the construction period. There will be a yearly technical review of the Project.

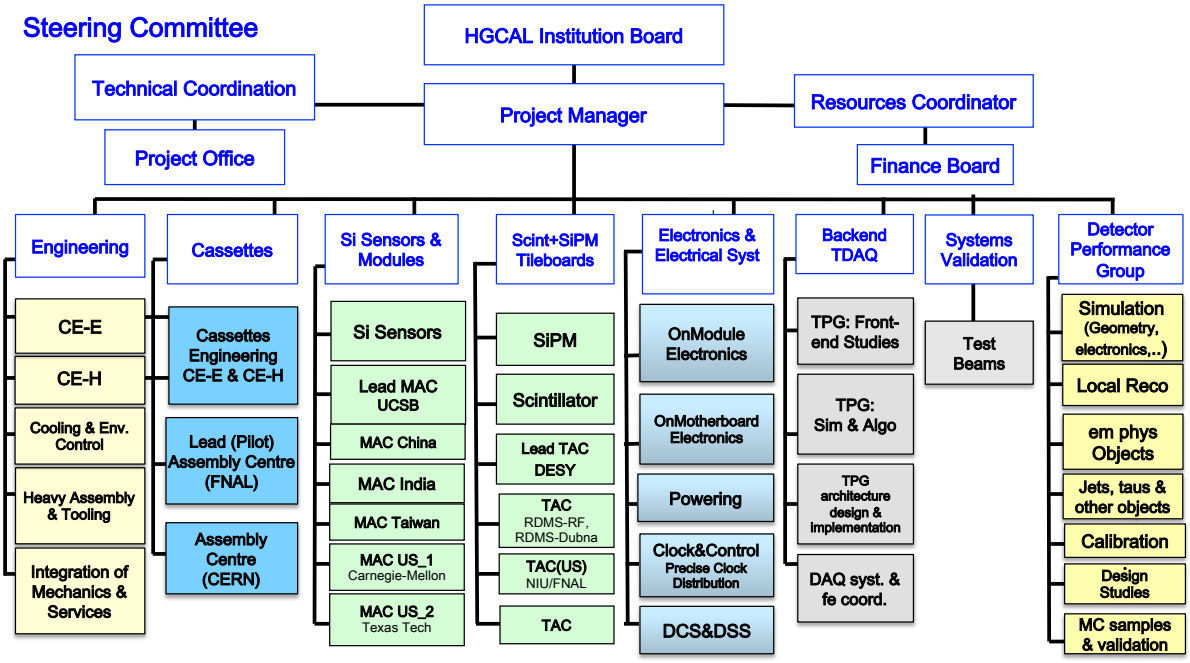


Figure 6.1: Management structure of the HGCAL project. The steering committee is comprised of persons in the boxes inside the blue shaded line. MAC and TAC refer to module and tileboard assembly centres.

6.1.2.2 Technical coordination and Project office

The HGCAL technical coordinator (TC), and his/her deputy (DTC), are responsible for the technical coordination of the Project; the planning, the safety, the quality assurance, the change control, and the link to CMS technical coordination. The responsibilities are defined below and updated as needed to adapt to changes of personnel or in the Project.

The TC and DTC, are assisted by the Project office. The TC, DTC and Project office coordinators, assisted by the Project office, shall:

- endeavour to ensure that the overall Project, and the various subprojects within it, do not falter for technical reasons,
- ensure and follow the necessary market or technological surveys needed in the various subproject areas,
- work in close harmony, and meet regularly, with the CMS technical coordination,
- identify link-persons to the TCG, for integration and planning. The link-persons shall keep the TCG informed of technical changes that have CMS-wide implications,
- hold the HGCAL prototyping and construction schedule,
- be responsible for adequate documentation, in appropriate platforms, across the Project and its updating, and the databases used for the various purposes in the Project,
- hold and monitor detailed milestones (down to Level-4, see Section 6.3) in order to follow the progress of the Project,
- establish, with the appropriate sub-area coordinators, the assembly and testing centres and qualify them for production,

- establish and qualify the necessary assembly spaces at CERN,
- establish and monitor QA/QC procedures across the Project,
- establish and document change control,
- help organising yearly reviews of the relevant items in the sub-areas. An engineering design review and an electronics system review are required prior to launch of full-scale construction of mechanics and electronics respectively.

6.1.2.3 Finance board

The HGCAL finance board (FB), through its chairperson, the resources manager, will deal with all matters related to the costs and resources of the Project, including contracts, and all related administrative matters, in collaboration with the CMS resources coordinator. The FB comprises one representative per funding agency, and managers of the Project (shown in Fig. 6.1).

The FB through its chairperson, is responsible for:

- the cost estimate, following the expenditure and the cost to completion of the Project,
- accounting of CORE expenses for the construction of HGCAL,
- providing the CMS resources manager with information required for presentation to the CMS resources review board (RRB), e.g. drawing-up the budget requests to be presented to the RRB.
- closing the book on the previous years cost book expenditure in order to make a financial statement to the spring RRB.
- the preparation of the interim status report of cost book expenditure, for internal use, in the autumn. The subdetector resources manager shall present the report to the FB.
- the policy, and followup, for multinational contracts, to be placed either by CERN (in collaboration with the resources manager) or other institutions.

6.1.3 Cost and responsibilities

The cost and institutional responsibilities have been evaluated using the Project breakdown to Level 3, and are given in Table 6.1.

Table 6.1: Cost and Responsibilities in the HGCAL Project. The Level-3 decomposition is only given for major items. The numbers in bold are rolled up and may not equal the sum of the Level-3 numbers.

	Item	Cost kCHF	Participating institutes
1.1	**Mechanical systems**		CERN, France (LLR), RDMS-RF (Protvino, Moscow INR), Turkey
1.1.1	CE-E electromagnetic calorimeter	258	
1.1.2	CE-H hadron calorimeter	3 424	
1.1.3	Cooling	7 618	
1.1.4	Dry gas system	60	
1.1.5	Mechanical assembly	80	
	Total	**11 441**	
1.2	**Cassettes**		France (LLR), Germany (DESY), RDMS-RF (Protvino, Moscow INR,
1.2.1	CE-E cassettes	970	Moscow Lebedev, MEPHI), USA (Alabama, Brown, Fairfield, FNAL,
1.2.2	CE-H (Si only) cassettes	413	Florida Tech., Florida State, Iowa, MIT, Northwestern, Notre Dame)
1.2.3	CE-H (mixed) cassettes	1 324	*Assembly centres:* CERN and FNAL
1.2.4	Cassette assembly centres and tooling	298	
1.2.5	Transport (assembly sites to CERN)	22	
	Total	**3 027**	
1.3, 1.4	**Silicon sensors and modules**	**25 992**	**All funding agencies will participate in the purchase of silicon sensors** *Module assembly centres:* China (Beijing IHEP), India (Mumbai BARC), Taiwan (Chungli NCU, Taipai NTU), USA (Carnegie-Mellon, Texas Tech., UCSB)

Continued on next page

Table 6.1 – continued from previous page

	Item	Cost kCHF	Participating institutes
1.5	**Scintillator/SiPM modules**		Germany (DESY), RDMS-RF (Protvino, Moscow INR, Moscow Lebedev, MEPHI), RDMS-DMS (Dubna, Kharkov KIPT, Kharkov ISMA, Minsk INP), USA (Fairfield, FNAL, Iowa, Maryland, Northern lllinois, Notre Dame, Rochester) *Scintillator module assembly centres:* Germany (DESY), RDMS-RF (Protvino), RDMS-DMS (Dubna + others), USA (Northern Illinois, FNAL)
1.5.1	SiPM photosensors	1 718	
1.5.2	Plastic scintillator	832	
1.5.3	Wrapping (ESR film)	111	
1.5.4	Assembly centre and tooling	270	
1.5.5	Scintillator/SiPM module production	14	
		2 945	
1.6	**Electronics and electrical systems systems**		CERN, China (Beijing IHEP), Croatia (Split), France (LLR, Omega, Saclay), India (Kolkata SAHA, Tata Institute B), Malaysia, Portugal (LIP), Thailand (Bangkok), RDMS-DMS (Minsk INP, Sofia INRNE), UK (Imperial College), USA (FNAL, Florida State, Kansas State, Maryland, Minnesota, MIT, Northwestern, UCSB)
1.6.1	Front-end systems (Si)	5 778	
1.6.2	Front-end systems (scint/SiPM)	872	
1.6.3	Front-end systems (common)	4 164	
1.6.4	Clock and control	500	
1.6.5	HV and LV power	4 448	
	Total	**15 762**	
1.7	**Backend systems (trigger and DAQ)**		CERN, Croatia (Split), France (LLR, Omega, Saclay), India (Bangalore IISC, Mumbai BARC, Tata Institute A, Pune IISR), Malaysia, UK (Bristol, Imperial college), USA (Boston, Maryland, MIT)
1.7.1	DAQ	2 447	
1.7.2	Trigger	3 779	
	Total	**6 226**	

Continued on next page

Table 6.1 – continued from previous page

	Item	Cost kCHF	Participating institutes
1.8	**Slow control (DCS, DSS)**		CERN, Greece (Athens, Ioannina, Athens NTUA), RDMS-RF (Moscow ITEP, Moscow Lebedev, Moscow MSU
1.8.1	DCS	257	
1.8.2	DCC	341	
	Total	**598**	
1.9	**Detector assembly (on surface)**		CERN, China (Beijing IHEP), France (LLR), India, RDMS-RF (Moscow INR, Moscow lebedev, MEPHI) RDMS-DMS (Dubna, Kharkov ISMA, Kharkov KIPT, Minsk INP) Contributions expected from many other institutes
1.9.1	Assembly areas	655	
1.9.2	CE-E assembly	65	
1.9.3	CE-H assembly (at Point 5)	70	
1.9.4	CE-E/CE-H integration	120	
		910	
1.10	**Installation and commissioning**	**226**	Expect contributions from many institutes
	GRAND TOTAL	**67 127**	

6.2 Project timeline and milestones

A simplified view of the project timeline is shown in Fig. 6.2.

The next two and a half years will be devoted to the design and prototyping phase, with validation of individual components as well as of systems (Si detector modules and scintillator boards equipped with silicon PMs assembled into cassettes). During this period the front-end electronics is the critical path item. The two and a half years should allow two prototyping iterations before the final production run of the ASICs. Tenders for active elements (sensors, PCBs, cassette elements) will be sent out during 2019 and vendors selected by early 2020, in such a way that the final iteration of prototyping could be carried out with components from the vendors who will be involved in the production. This phase will end with an engineering design review (EDR) , currently scheduled at the end of Q2 2020, which will validate the complete design and give the green light for production. An advanced EDR in Q4 2019 will however take place for the mechanical structures (absorber for CE-H, lead-stainless steel absorbers for CE-E, cooling plates of cassettes,...) to keep these items out of the critical path. In parallel, the various assembly centres for modules, scintillator tiles and cassettes will be set-up and qualified.

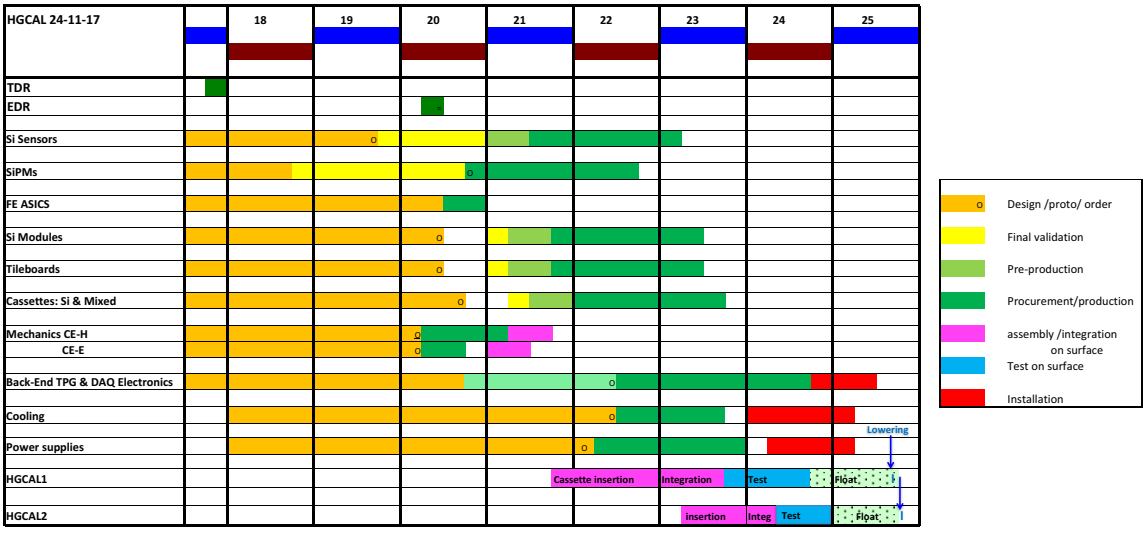


Figure 6.2: Simplified view of project timeline.

The production and construction phase includes final validation followed by a (typically 5%) pre-series phase for a period of 6 months which will allow both the producers and the assembly centres to ramp-up their production rate and finalise their QA/QC procedures. The silicon sensors production, the silicon modules/scintillator tile-modules assembly, the assembly of cassettes and their insertion in mechanical structures will then proceed in parallel, with a quarter year staggering between each of these processes. The silicon sensors production, with a 21 months duration, is the critical path for this phase. The work on CE-E and CE-H will also proceed in parallel.

When the CE-E and CE-H for the first endcap are ready, they will be brought together at Point 5, allowing the dressing of all the services, making initial tests (6 months) and starting a nine-month period of integrated tests, including a testing period at the design cold temperature. The same procedure will be followed for the second endcap. However, the integrated test period is reduced to 6 months, benefitting from the experience gained from the first endcap.

The timeline of the project includes a contingency of 9 months before the *Ready for Lowering* milestones for the two HGCAL endcaps, currently planned for Week 85 and 91 of the shutdown, as required by the CMS LS3 planning.

The HGCAL project office will be responsible for following the planning of the project. A detailed schedule has been set-up with the Merlin® software. A large set of milestones has been defined, to monitor the progress, and a subset of High Level Milestones, listed in Table 6.2, will be reported to the LHCC committee. Dashboards will allow monitoring of the construction rates on a (typically) weekly basis.

6.2.1 List of milestones

The milestones are listed in Table 6.2. These will be used by the project to monitor technical progress and take the necessary actions to keep the project on track. The engineering design reviews (EDR) and electronics system reviews (ESR) will be held in mid-2020. Where the critical path necessitates procurement of major components before this date, reviews will be required correspondingly earlier.

Table 6.2: HGCAL project high-level milestones.

Milestone	TDR ID	Date
Engineering design review held		20/08/2020
Sensors (SI)		
Delivery of 8″ sensors with realistic geometry	CE.SI.3	13/09/2018
Market survey for sensors complete, vendors selected	CE.SI.4	01/04/2019
8″ sensor orders placed	CE.SI.5	01/10/2019
Final sensors qualified	CE.SI.6	01/01/2021
Sensors production 5% complete	CE.SI.7	17/06/2021
Sensors production 50% complete	CE.SI.8	19/05/2022
Sensors production 100% complete	CE.SI.9	23/02/2023
Silicon modules (MO)		
Module assembly pilot site and procedures setup 6″	CE.MO.2	28/11/2017
Module assembly pilot site and procedures setup 8″	CE.MO.3	24/12/2018
All module assembly sites and procedures qualified	CE.MO.4	08/06/2020
HGCROC-DV2 module qualified	CE.MO.5	25/06/2020
Module components orders placed	CE.MO.6	20/08/2020
Final module qualified	CE.MO.7	29/04/2021
Modules production 5% complete	CE.MO.8	14/10/2021
Modules production 50% complete	CE.MO.9	15/09/2022
Module production 100% complete	CE.MO.10	25/05/2023
Scintillator/SiPM tileboards (SC)		
Scintillator and tile board assembly technique selected	CE.SC.1	23/07/2019
SiPM order placed	CE.SC.2	28/07/2020
Scintillator tiles order placed	CE.SC.2′	28/07/2020
HGCROC-SiPM-DV2 tileboard validated	CE.SC.3	20/08/2020
Tileboard assembly pilot site and procedures set-up	CE.SC.4	20/08/2020
Tileboard components orders placed	CE.SC.5	20/08/2020
All tileboard assembly sites and procedures setup	CE.SC.6	04/03/2021
Final tileboard qualified	CE.SC.7	25/03/2021
Tileboard production 5% complete	CE.SC.8	07/10/2021
Tileboard production 50% complete	CE.SC.9	08/09/2022
Tileboard production 100% complete	CE.SC.10	18/05/2023
FE electronics (FE)		
HGCROC-DV1 submission	CE.FE.6	28/06/2018
HGCROC-DV1-SiPM submission	CE.FE.7	28/06/2018
Concentrator V1 submission	CE.FE.8	07/09/2018
HGCROC-DV2 submission	CE.FE.9	27/06/2019
HGCROC-DV2-SiPM submission	CE.FE.10	27/06/2019

Continued on next page

Table 6.2 – continued from previous page

Milestone	TDR ID	Date
FE electronics (continued)		
Concentrator V2 submission	CE.FE.11	06/09/2019
HGCROC ready for production	CE.FE.12	25/06/2020
HGCROC-SiPM ready for production	CE.FE.13	25/06/2020
Concentrator ready for production	CE.FE.14	26/06/2020
CE-E and CE-H electronics system for on-detector (including links, concentrator,..) (EL)		
Results from a CE prototype using SKIROC2-CMS	CE.EL.2	30/11/2018
HGCROC-DV1 system tested	CE.EL.3	02/05/2019
HGCROC-DV2 system validated	CE.EL.4	05/03/2020
Final hexaboard qualified	CE.EL.5	04/02/2021
Final motherboard qualified	CE.EL.6	05/02/2021
Final system qualified	CE.EL.7	27/05/2021
Cassettes (CS)		
First thermo-electro-mechanical prototype test	CE.CS.1	28/08/2018
Cassette assembly pilot site and procedures setup	CE.CS.2	30/05/2019
HGCROC-DV1 cassette tested	CE.CS.3	30/05/2019
Both cassette assembly sites and procedures qualified	CE.CS.4	28/05/2020
HGCROC-DV2 cassette validated	CE.CS.5	17/09/2020
Final cassettes qualified	CE.CS.6	22/07/2021
Cassette production 5% complete	CE.CS.7	06/01/2022
Cassette production 50% complete	CE.CS.8	08/12/2022
Cassette production 100% complete	CE.CS.9	14/09/2023
Mechanical engineering (EN)		
Mechanical design completed	CE.EN.5	15/10/2019
CE-E structures received	CE.EN.6	12/10/2020
CE-H absorber received	CE.EN.7	29/03/2021
CE-E support structures completed	CE.EN.8	21/06/2021
CE-H structures completed	CE.EN.9	13/08/2021
Cooling (CO)		
P5 cooling plant specifications completed	CE.CO.1	05/06/2018
Cooling demo plant designed (with CMS TK)	CE.CO.2	24/04/2019
Cooling demo system built	CE.CO.3	26/08/2020
Review of demo cooling plant performance	CE.CO.4	16/06/2021
HGCAL cooling system design completed	CE.CO.5	04/01/2022
Cooling plants procurement complete	CE.CO.6	12/10/2023
Main transfer lines procurement completed	CE.CO.7	16/01/2024

Continued on next page

Table 6.2 – continued from previous page

Milestone	TDR ID	Date
Cooling (continued)		
Main transfer lines, main manifolds, YE transfer lines installation completed	CE.CO.8	08/10/2024
Cooling system ready for connection to detector	CE.CO.9	14/02/2025
TPG and back-end electronics (BE)		
Prototype functions and interfaces validated	CE.BE.4	07/10/2019
Specification of BE system documented in EDR	CE.BE.5	20/04/2020
Preproduction functions and interfaces validated	CE.BE.6	01/11/2021
BE hardware production readiness review	CE.BE.7	11/07/2022
Production functions and interfaces validated	CE.BE.8	25/12/2023
BE electronics production complete	CE.BE.9	23/12/2024
HGCAL integration with central DAQ and L1T complete	CE.BE.10	30/12/2025
CE-E and CE-H HV systems (HV)		
Bias supply system specifications defined	CE.HV.1	28/06/2019
Final bias supplies system qualified	CE.HV.2	31/03/2022
Bias supplies production complete	CE.HV.3	02/02/2024
Bias supplies installed and tested	CE.HV.4	30/03/2025
CE-E and CE-H LV systems (LV)		
LV power supply system specifications defined	CE.LV.1	02/05/2019
Prototype LV power supply system validated	CE.LV.2	03/02/2022
LV power supplies production complete	CE.LV.3	07/12/2023
LV power supply system installed and tested	CE.LV.4	30/03/2025
Assembly and commissioning (AS)		
CE-E HGCAL1 stacking of cassettes complete	CE.AS.1	02/02/2023
CE-H HGCAL1 cassette insertion complete	CE.AS.2	02/02/2023
HGCAL1 integration complete	CE.AS.3	26/10/2023
CE-E HGCAL2 stacking of cassettes complete	CE.AS.4	07/12/2023
CE-H HGCAL2 cassette insertion complete	CE.AS.5	07/12/2023
HGCAL2 integration complete	CE.AS.6	23/05/2024
HGCAL1 testing complete, ready for lowering	CE.AS.7	02/08/2024
HGCAL2 testing complete, ready for lowering	CE.AS.8	06/12/2024

Part II

Additional Information

Chapter 7

Sensors and active elements

7.1 Silicon sensors

The design of the HGCAL silicon sensors aims for simple devices utilizing the maximal available area on the circular wafers to minimize the costs. For that reason, the sensors are hexagons with truncated tips. They will be fabricated on 8″ wafers. The sensors must survive the harsh radiation environment of the HGCAL. Sensors with different active thicknesses (300, 200 and 120 μm) are envisaged for regions of differing fluence in the detector in order to optimize the charge collection and operation conditions over the full lifetime of the HGCAL.

Initially, only sensors from 6″ wafers were available and were considered as the baseline in the Technical Proposal. For that reason a comprehensive campaign was started, of characterization, irradiation studies and beam tests with full size 6″ sensors to converge on optimal design choices and parameters with sensors coming from Hamamatsu.

In the course of a detailed study conducted together with three companies, Infineon, Novati and Hamamatsu, sensors on 8″ wafers were developed. Since companies made impressive progress during this R&D effort, 8″ sensors are now considered as the baseline for the HGCAL.

An irradiation campaign to qualify full size 8″ sensors will be conducted. This contains IV and CV measurements and a study to optimize inter-pad capacitance and resistance with respect to different p-stop layouts and spacings (so-called atoll vs. combined p-stop). A key part of these campaigns will be noise measurements and laser tests on irradiated full size sensors, which will be performed by hexaboards with integrated pogo-pins to quickly connect bare sensors to readout electronics (Section 7.1.7).

7.1.1 Sensors from 8-inch wafers

The three companies listed below have met initial qualification criteria set by the CMS tracker collaboration in the course of a market survey to provide silicon strip sensors for the Phase-2 upgrade of the outer tracker [6].

7.1.1.1 Hamamatsu

Hamamatsu Photonics K.K. (HPK) provided sensors on 6-inch wafers used for both module assemblies for beam tests and sensor optimization studies. Moreover, different demonstration prototype sensors on 8″ wafers for each of the substrate types used in the HGCAL are available. Table 7.1 summarizes all sensors delivered for the project.

As the deep diffusion process used by HPK for reducing the active thickness while maintaining the physical thickness to 320 μm is not available on 8″, 200 and 300 μm thick sensors have equal physical and active thickness. To maintain the same or a similar quality as the 6″ sensors,

Table 7.1: Summary of sensors delivered by Hamamatsu. The abbreviations com and ind indicate common and individual p-stop geometry, respectively. The different thicknesses can be achieved either by physical thinning (std), deep diffusion (dd), or epitaxial growth (epi).

Size	Bulk pol.	P-stop geo.	Thickness act./phys.	#Pads	Pcs.	Comment
6″	n	–	300/320 (std)	128	110	testbeam modules
		–	300/320 (std)	128	11	″additional″
		–	200/320 (dd)	128	36	
		–	120/320 (dd)	256	36	
6″	p	com.	300/320 (std)	128	10	
		ind	300/320 (std)	128	10	
		com	120/320 (dd)	256	10	
		ind	120/320 (dd)	256	10	
8″	p	ind.	300/300	272	6	″stepper″ samples
		ind	200/200	272	5	″stepper″ samples
		ind	300/300	272	6	to be delivered
		ind	200/200	272	6	to be delivered
		ind	120/300 (epi)	272	6	3 pcs. delivered, remaining to follow

special care has to be taken for the thinning, backside implantation, and backside metal deposition. The sensors with an active thickness of 120 μm are being produced by growing epitaxial silicon on a lower resistivity substrate, resulting in an total physical thickness of 300 μm.

The demonstration prototype 8″ sensors delivered by HPK in 2017 have been manufactured using the so-called stepper technique in contrast to the full mask photolithography typically used for sensor fabrication. This technique utilizes a set of small photomasks which are exposed to the wafer in a step-and-repeat process. Moreover, several sets of masks are needed for the irregular half-cells on the sensor periphery close to the guard and edge rings. For this reason, Hamamatsu has committed to upgrade their production lines using full-wafer lithography. First high-quality production-grade sensors using this technique are expected to arrive in spring 2018.

7.1.1.2 Infineon

Infineon has already been working for some time to establish a production process for AC-coupled silicon strip sensors for the Phase-2 Upgrade of both the ATLAS and CMS trackers. After a first run with several batches of sensors using n-type substrate and 6″ diameter, they produced p-type silicon strip sensors on 8″ wafers in three production batches.

The 8″ sensors procured as prototypes for the HGCAL from Infineon are hexagonal sensors in n-on-p technology with (active and physical) thicknesses of 350, 200 and 140 μm. A summary can be found in Table 7.2. Additionally, sensors with silver backside metallization were tested instead of standard aluminium metallization. The wafer material is p-type float-zone silicon with a crystal orientation of <100> and a resistivity between 3 and 8 kΩ cm.

7.1.1.3 Novati

The US-based company Novati has also delivered samples of demonstrator sensors based on 8″ p-type wafers, of both tracker and HGCAL types (see Table 7.3). There have been three 8″

Table 7.2: Sensor delivered by Infineon on 8-inch wafers and p-type substrate. All sensors belong to type M3232A.

P-stop	Thickness active/physical	#Pads	Pcs.	Comment
common	200/200	235	7	
	300/300	235	4	
individual	140/140 (planar)	235	2	
	200/200	235	7	
	350/350	235	1	
–	200/200	235	2	mechanical samples

n-on-p Novati development runs. The first run (called run 0) utilized bulk silicon, the second and third used both bulk and silicon-on-insulator bonded wafers. The bonded wafer process allows for thin active thickness wafers to be processed at full thickness, allowing standard fabrication equipment to be used throughout the sensor processing. This is at the cost of more expensive starting material. All runs contain a half HGCAL sensor 6″ layout, which was processed (together with other structures) on 8-inch wafers. A final development run, number 4 in the table, using silicon-on-insulator and silicon-silicon bonded wafers is scheduled for early 2018. This run should provide a demonstration of full wafer HGCAL sensors.

Table 7.3: Sensor delivered by Novati.

Run	Type	# Pads	Thickness active/physical	Pcs.	Geometry	Date
0	planar	69	725/725	3	1/2 of 6″	Apr-15
	planar	69	500/500	2	1/2 of 6″	May-15
1	planar	69	500/500	3	1/2 of 6″	Mar-16
2	SOI	69	200/700	3	1/2 of 6″	Mar-16
3	SOI	69	200/700	9	1/2 of 6″	Dec-17
4	SOI	192	200/700	5	full 8″	Start ∼Dec-17
	Si-Si	192	200/700	5	full 8″	Start ∼Dec-18

7.1.2 Sensor design

The silicon sensors consist of individual DC-coupled hexagonal pads, tiled across the hexagonal sensor surface. Each individual diode pad has an area of 1.18 cm^2 (200/300 μm thickness) or 0.52 cm^2 (120 μm thickness), but is not connected to any biasing circuit. The junction potential is maintained by the front end amplifier. This is in contrast to typical biasing schemes in strip (poly-silicon resistors) or pixel detectors (punch-through). This concept reduces the sensor complexity (and thus cost), but raises specific issues for testing (which are addressed in Section 7.1.4 below), as well as system design issues, e.g. to mitigate effects of any cells with HV break-down (either known at time of module assembly or possible later defects) which must be dealt with at the module level.

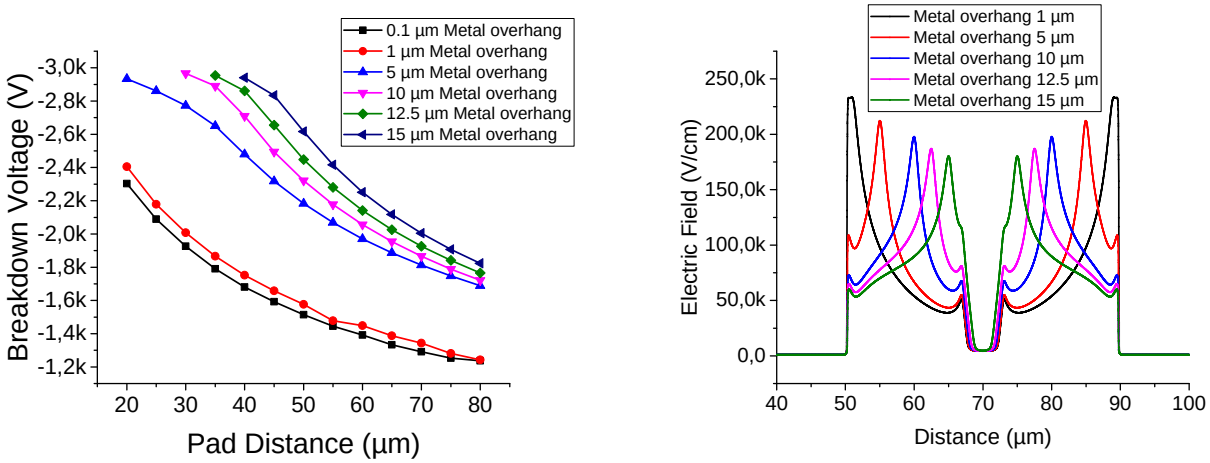
7.1.2.1 Guard ring design

All hexagonal pads are surrounded by a periphery consisting of two guard rings and a so-called edge ring. The outermost region of the sensor consists of the edge ring, which is a p^+

implant into the p-type bulk in order to protect the dicing line from electrical fields and to avoid surface current across the dicing region. The guard ring design is based on the robust and high-voltage stable design scheme of the present tracker. It consists of a floating outer guard ring with asymmetric metal overhang and a grounded inner guard ring for HV stability. In this configuration the sensors cells are decoupled from large energy deposits which could otherwise create cross-talk.

7.1.2.2 Pad design and high-voltage stability

The pads metallization is larger than the underlying implant layer. The aim of this so-called metal over-hang is to move the maximum of the electric field into the oxide which has a higher break-down voltage than the silicon bulk. Moreover, the peak field in the bulk is reduced by minimizing the inter-pad gap (Fig. 7.1), that has to be tensioned with the inter-pad capacitance, which is lowered by wider inter-pad gaps. Moreover, production-related limitations on the minimum pad distance have to be taken into account.

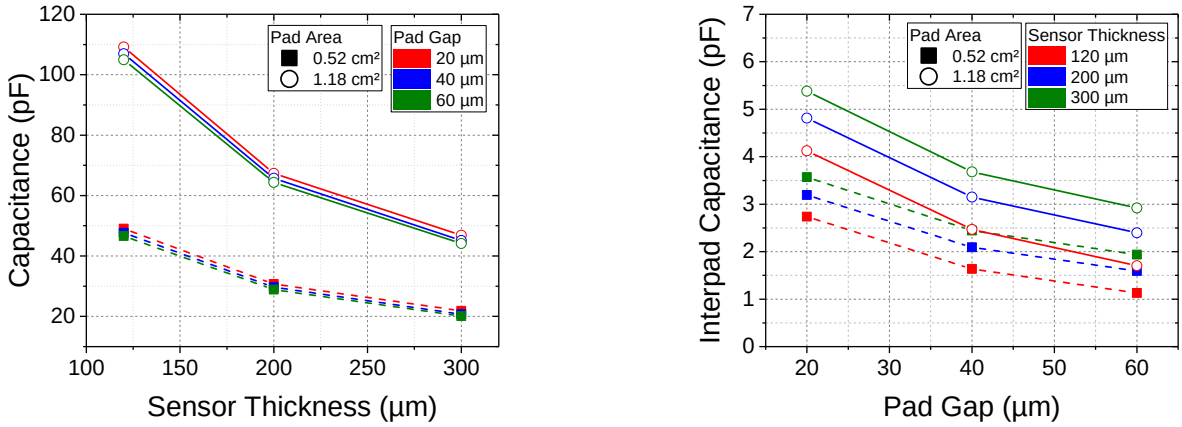


(a) (b)

Figure 7.1: Results of a breakdown simulation of pads with (a) different pad implant spacing, and (b) metal overhangs, for a 40 μm pad implant. The absolute numbers of the breakdown voltage have to be treated with caution and do not represent real breakdown voltages of the devices, but act as a qualitative trend indicator.

The inter-pad capacitance was studied together with the total cell capacitance (see Fig. 7.2). The capacitance of each pad has two components, the substrate capacitance, which should be close to the parallel plate value of $\epsilon A/d$, and the inter-pad capacitance, which depends on detector thickness and the gap between pads. The sensors were simulated using the Silvaco TCAD package as two adjacent diodes with 5-micron metal overhang. Sensor thickness and inter-implant gaps were varied while an AC simulation was performed at 10 kHz with 100 V bias applied.

The results show that the lowest inter-pad capacitance is given by the widest pad gap. However, its contribution to the total capacitance remains small in all cases (compare Fig. 7.2 (a)). Full size 6″ multi-geometry sensors, featuring 20, 40, 60, and 80 μm pad-to-pad separations in each of four quadrants, are being used together with detailed device simulations in order to systematically compare and optimize other features of the cell design, such as the use of individual or common p-stops for inter-cell isolation, and the size of the inter-pad gap which impacts both peak electric field and high voltage stability as well as inter-cell capacitance.

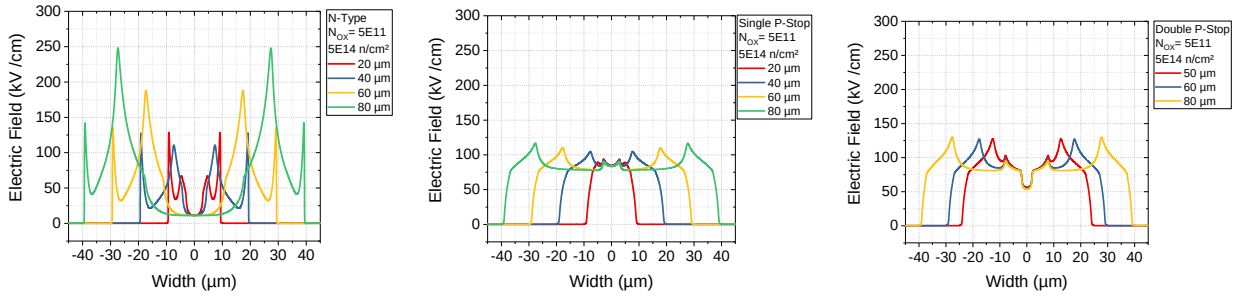


(a) (b)

Figure 7.2: Results of a capacitance simulation showing (a) the total capacitance, and (b) inter-pad capacitance for different sensor thicknesses, cell sizes and pad distances.

7.1.2.3 Choice of substrate material and p-stop

The baseline choice of the substrate material is motivated by the intrinsic radiation hardness of p-type silicon [36] as established by studies performed by the CMS Tracker collaboration [37]. These studies have shown that n-type sensors are prone to micro-discharge effects (non-Gaussian noise), which can mimic signals generated by a traversing particle. The cause of this effect is a high electric field between adjacent cells/strips which creates avalanche effects. For the HGCAL sensors the situation might be different since the metal coverage and the pad spacing is different. For that reason, a TCAD simulation was performed to calculate the different field strengths for n-type and p-type sensors (Fig. 7.3).



(a) (b) (c)

Figure 7.3: Simulated electrical field 50 nm underneath the Si/SiO_2 interface for (a) n-substrate, (b) p-substrate and one, or (c) two p-stop lines between adjacent cells

These figures also show the difference between single p-stop and double p-stop lines, which corresponds to the "common" and "atoll" p-stop layout, respectively. The results indicate that the electric field is comparably low for 20 μm pad spacing on both n- and p-type sensors. When the pad spacing is increased, the electric field in the n-type case increases more than for p-type substrate. For that reason HGCAL n-type sensors with too large pad spacing could suffer from non-Gaussian noise similar to the noise observed in tracker sensors.

7.1.2.4 Bonding pads

The geometry of the bonding pads for sensor cells is optimized for ease of sensor testing, module design, bonding and subsequent manipulation. It consists of a circular opening in the

sensors' passivation layer around the region where three adjacent pads meet. The grounded inner guard ring will also need passivation openings for wire bonding and sensor testing. In addition, special "landing pads" are planned inside and outside of the sensor on the silicon wafer which will allow simpler testing procedures using 7-needle probe cards by the vendors.

7.1.3 Radiation tolerance

The silicon sensors will need to operate after fluences, for an integrated luminosity of 3000 fb^{-1}, ranging from $\approx 10^{14}\,\mathrm{n_{eq}/cm^2}$ at the outer radius to $\approx 10^{16}\,\mathrm{n_{eq}/cm^2}$ at the inner radius. The size and types of sensors must be optimized to accommodate the total dose at lowest cost. For example n-on-p sensors are expected to be more radiation hard, but are more expensive to fabricate. For that reason we are investigating the possibility of using n-type sensors in the outermost regions with the lowest fluences. Similarly, devices with thinner active areas are expected to be more radiation hard than thicker devices, but these devices require special processing to produce either physically thin wafers or devices with thin high resistivity active area stacked on lower resistivity material. Examples of this are deep diffused silicon, silicon-on-insulator (SOI), and silicon-silicon bonded wafers. Therefore we will need to test and qualify candidate sensors from several different vendors with several process variants, thicknesses, design rules, and substrates. In addition we want to match dose and fluence to different regions of the detector, both for neutrons, and for pions and protons.

7.1.3.1 Further results from irradiation of different sensors

The results summarized in Fig. 2.1, which show the charge collection of diodes of different active thicknesses, and different bias voltages, after irradiation to the expected fluences corresponding to an integrated luminosity of 3000 fb^{-1}, were all obtained at room temperature using a fast infrared laser signal using the transient current technique (TCT) to induce signals large enough (40 MIP equivalent) to be measured easily.

The MIP charge collection of these diodes was subsequently also directly measured using a ^{90}Sr source, at $-20\,^{\circ}\mathrm{C}$ and $-30\,^{\circ}\mathrm{C}$. Operating the sensors at this low temperature not only mimics the expected conditions in the experiment, but is also necessary in order to reduce the shot noise resulting from radiation induced leakage current to a level allowing clean detection of MIP signals. A comparison between the two sets shows consistent results, as shown in Fig. 7.4.

Based on the current proposals from the candidate producers, sensors with 300 μm and 200 μm active thicknesses will be produced on FZ wafers (rather than dd-FZ) thinned to 300 μm and 200 μm respectively. Sensors with a thickness of up to 120 μm may be produced using epi wafers of a physical thickness of 300 μm.

The use of physically thin FZ rather than dd-FZ is not expected to have any significant impact on the charge collection and leakage current behaviour of the sensors. This was confirmed by directly comparing 200 μm active thickness sensors produced with both techniques as part of the R&D for the CMS Phase-2 Tracker Upgrade.

Spectra obtained, at a bias of 600 V, for 130 μm dd-FZ and 100 μm epi, and with both the TCT pulses and ^{90}Sr, are shown in Fig. 7.5: a good signal response can be seen in both cases. Results from a comparison of CV and TCT charge collection curves for the two different sensor types indicate a substantially lower depletion voltage for the 100 μm epi relative to the 120 μm dd-FZ. This, together with the higher electric fields for the thinner device at a given bias voltage, may account for the approximately 20% higher MIP signal response observed for the 100 μm epi diodes compared to the 120 μm dd-FZ.

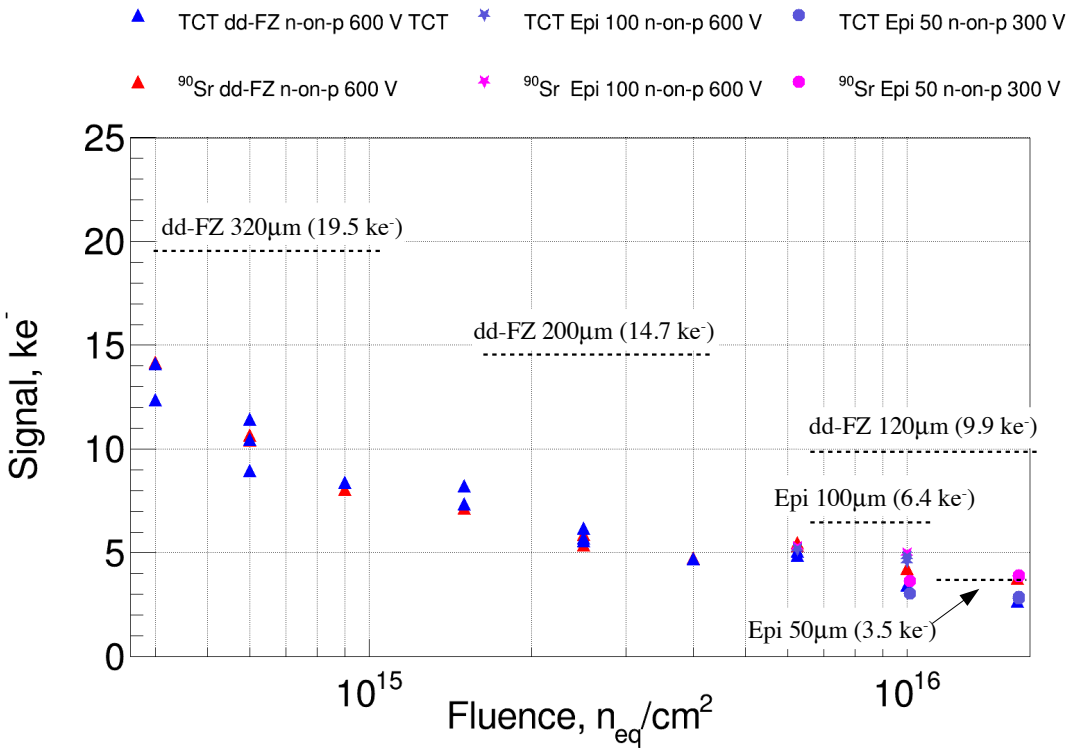


Figure 7.4: Comparison between the signal (ke^-) extracted from the TCT and the ^{90}Sr measurements for the samples of diodes under test, after a range of neutron fluence bracketing the maximum expected exposure after 3000 fb^{-1}. The arrows indicate the thickness of the different samples, the corresponding MIP charge yield before irradiation, and the range of fluence to which each sample was exposed. The charge collection results are shown at 600 V for the ddFZ diodes of 300 μm, 200 μm and 120 μm nominal thickness, as well as for the 100 μm epitaxial, and at 300 V for the 50 μm epitaxial diodes.

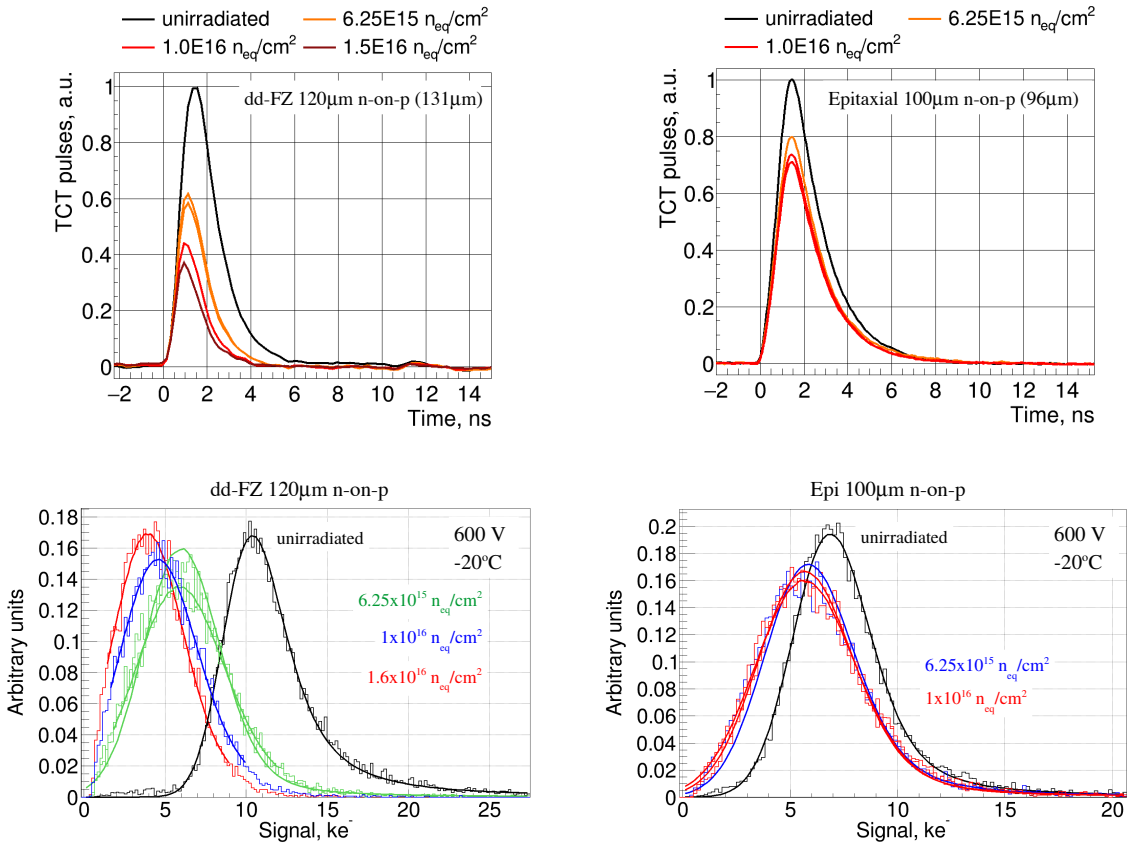


Figure 7.5: Comparison of the pulse shapes and MIP response obtained at 600 V with the TCT (top) and ^{90}Sr techniques (bottom), for the p-type 130 μm ddFZ (left) and 100 μm epitaxial diodes (right).

The irradiated epi diodes appear to lie somewhat above the leakage current density of the dd-FZ samples: this is shown at 600 V in Fig. 2.2 and a similar behaviour can be seen in Fig. 7.6 at 800 V. This difference is not yet understood.

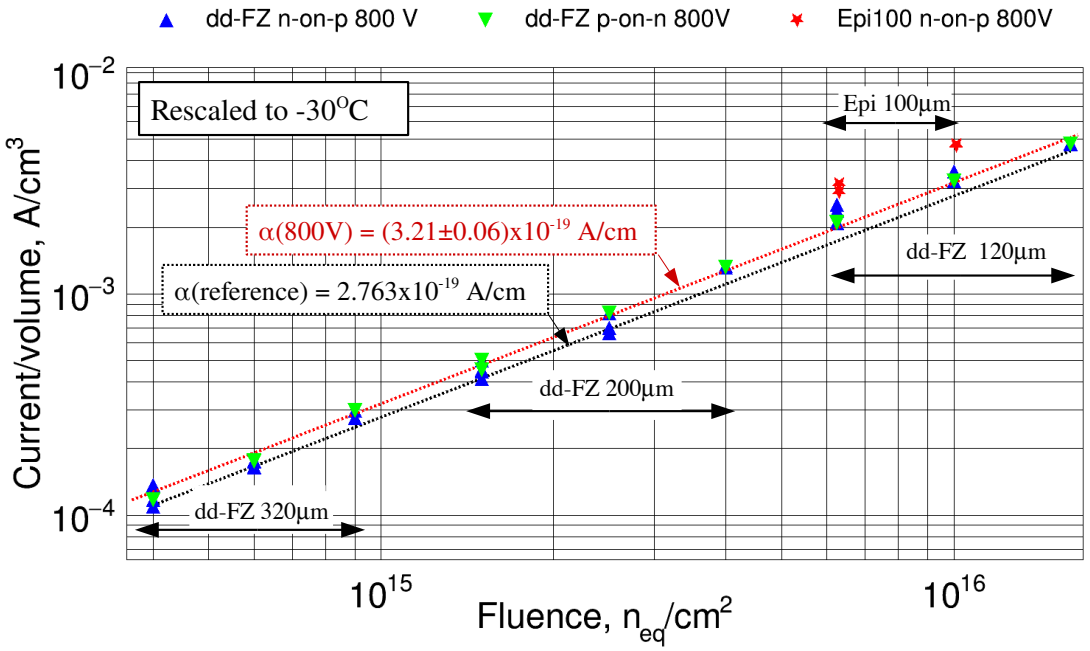


Figure 7.6: Leakage currents, as a function of neutron fluence, measured at $-20\,^{\circ}\mathrm{C}$ and rescaled to $-30\,^{\circ}\mathrm{C}$, at a bias voltage of 800 V. The arrows indicate the thickness of the different samples, and the range of fluence to which each sample was exposed.

It is planned to repeat all these irradiation studies as part of the broader programme of sensor design optimization and qualification prior to the start of production.

7.1.3.2 Plans for further irradiation testing

We divide the fluence into three regions with associated sensor technologies shown in Table 7.4.

Table 7.4: Range of neutron fluences in sensors of different thicknesses.

Thickness	Lower fluence	Higher fluence
300 μm	1×10^{14}	5×10^{14}
200 μm	5×10^{14}	2.5×10^{15}
120 μm	2.0×10^{15}	7×10^{15}

Figure 7.7 shows the non-ionizing radiation fluence as a function of radius for $z = 340\,\mathrm{cm}$, near the peak value of the fluence. We plan exposures at the lower and higher fluences given in Table 7.4, as well at a value of 1.5 times the lower and higher fluence. Test structures from each of the candidate manufacturers will be irradiated to qualify the vendor for production. We will also irradiate test structures during production monitor quality of the sensors.

Given the large range of doses, sensor types, and vendors we need a focussed, but comprehensive irradiation programme. To accomplish this we are planning extensive tests of device test structures, which are small, easily tested and provide basic functional information. These test structures will be irradiated in proton particle beam facilities at Karlsruhe, Florida State

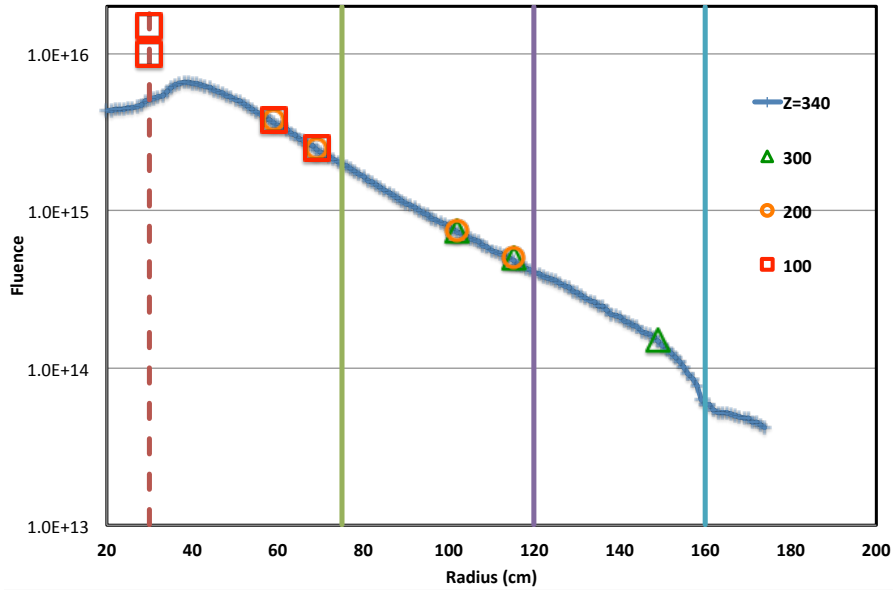


Figure 7.7: Fluence vs. radius at $z = 340\,\text{mm}$. Planned irradiation points for each sensor thickness are shown within each range.

University, and Los Alamos. The large area full sensors are impractical to scan in a ≈cm area particle beam; therefore they will be irradiated with reactor neutrons.

7.1.3.3 Irradiation of full-size sensors

The TRIGA reactor of JSI Ljubljana is a key resource for the detector development community and serves as an established reference for neutron irradiations with well-known neutron spectrum and fluence calibration. Recently, a transport system for the irradiation of large objects was added to the reactor with the help of AIDA-2020, an EU-funded infrastructure programme. This system allows the irradiation of devices with a diameter compatible with 6″ HPK sensors to just fit into this irradiation channel. Thus the JSI reactor will be used for irradiation studies of the existing 6″ HPK sensors for design optimization in the R&D phase.

Large sensors from 8″ wafers need to be irradiated as well. This is true for both R&D studies during the prototype phase and for irradiation tests during series production. For this purpose, new irradiation facilities are being commissioned that will allow the exposure to neutrons of sensors of a diameter of 190 mm.

The Rhode Island Nuclear Science Center (RINSC) operates a 2 MW light water cooled pool-type reactor on the Bayside Campus of the University of Rhode Island at Narragansett, RI. The reactor is located 30 miles south of Brown University in Providence, allowing easy access for the transfer of samples at any time. Samples can be delivered into the reactor core for neutron irradiation in two ways, via a pneumatic rabbit system, or via beam ports. After irradiation, samples are stored in a freezer at the reactor until their residual radiation has decreased sufficiently for them to be transported safely.

The rabbit system can deliver samples up to 35 mm × 90 mm into the reactor core for a preprogrammed time while the reactor is running. In this way samples can be irradiated to different fluences in quick succession. Due to the constant air flow in the rabbit system, the samples remain relatively cool during irradiations. A maximum temperature of 32 °C has been measured during a one hour irradiation.

The beam ports are cylindrical tubes that allow the delivery of samples up to the size of an 8″ silicon wafer into the reactor core. The reactor must be off in order to load the samples into a beam port. Then a shielding plug is inserted into the tube and the reactor is ramped up. Once the desired irradiation time is reached the reactor is ramped down again. The beam port can be opened after a four hour wait time that allows the radioactivity of the reactor cooling water to decay. The beam ports are suitable for irradiations of large samples to high fluences. Without cooling, samples in the beam port will heat up during irradiation. We have measured temperatures up to 50 °C after an irradiation time of three hours. To avoid excessive annealing, the samples can be cooled with dry ice.

The Brown University group has measured the neutron spectrum in a 6″ beam port by irradiating high-purity foils of 10 different elements with neutron cross sections which turn on at different energy threshold from sub-eV values to 12 MeV. The foils were irradiated in the beam port for one hour. Then their activity was measured and compared to predictions based on a hypothetical neutron flux spectrum in bins of neutron energy using a sum of squares metric. The neutron flux spectrum was varied to minimize the difference between the measured and predicted activities. The best agreement was reached with the neutron flux spectrum shown in Fig. 7.8. Using the damage function for silicon a 1 MeV neutron equivalent flux of $6.0 \times 10^{11} \mathrm{cm}^{-2}\mathrm{s}^{-1}$ was determined. The uncertainty is approximately 20%, driven by variations in the activity across the beam port. During a series of irradiations of iron foils and silicon test diodes with different irradiation times in the rabbit system, the 1 MeV equivalent neutron flux was determined to be $4.3 \times 10^{11} \mathrm{cm}^{-2}\mathrm{s}^{-1}$, assuming the same spectrum as measured in the beam port. The leakage current of the irradiated diodes at full depletion after beneficial annealing increases linearly with irradiation time as expected and is consistent with the fluence measurement at the 20% level as shown in Fig. 7.8.

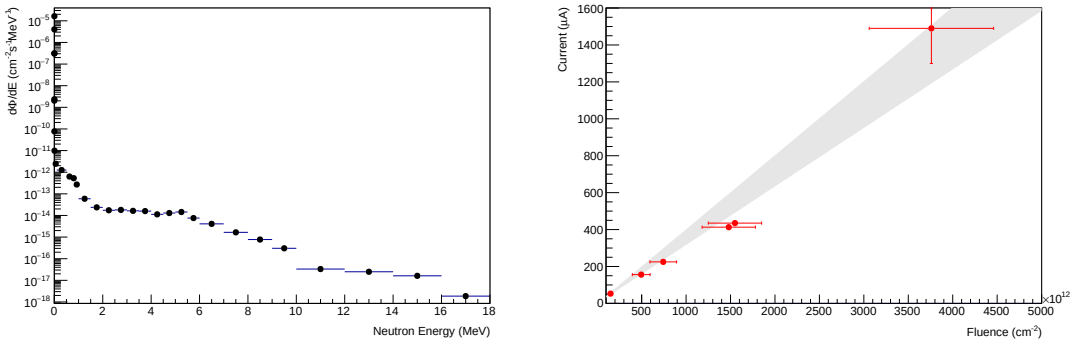


Figure 7.8: Neutron flux in bins of energy as measured in the RINSC beam port (left), and leakage current of diodes irradiated in the rabbit system versus neutron fluence ($\mathrm{n_{eq}/cm^2}$) (right).

7.1.4 Sensor testing procedures

The sensors are simple DC-coupled devices, which rely on the connection of each individual pad to the front-end amplifier channel to provide the ground return of the common HV bias, which is applied to the sensor backplane. Such designs are simpler to design and produce, but cannot be fully tested without providing additional bias connections. The simplest test would utilize a single probe needle, providing bias to a single pad. This suffers from punch-through of bias and extension of internal fields to the neighbor, floating, pads which will in turn cause parasitic charge collection. To circumvent this problem, the *7-needle measurement* connects the

six neighboring cells to ground and provides a better electric field configuration around the central cell. In both these methods, the needles must be sequentially stepped across the whole sensor. In the operating detector, all cells are contacted in parallel and connected to ground. This situation can be reproduced with a dedicated probe card utilizing several hundred needles. This will provide an accurate and flexible test. Depending on the requirements, the three different measurement methods can be used during the R&D phase and the series production.

7.1.4.1 Sensor testing during R&D phase

During the ongoing R&D phase, the properties of the sensors need to be understood and compared to simulations. This includes the inter-pad capacitance and inter-pad resistance, together with studies to determine the optimal p-stop layout and geometries. These studies need to be done both before and after irradiation. An irradiation campaign will be performed on the 6-inch sensors from HPK, which includes sensors of different thicknesses, bulk polarity (n/p) and p-stop layout (for p-type sensors). For basic electrical characterization, the 7-needle measurement can provide the relevant detector information. This configuration allows tests of total inter-pad capacitance and resistance as well as pad backplane capacitance and leakage current. This can be repeated across the sensor using an automatic probe station.

7.1.4.2 Sensor testing during series production

The series production of the sensors is scheduled for a period of approximately two years, during which all relevant parameters must be monitored. For this purpose, a quality assurance (QA) scheme will be implemented, which follows the established model of the tracker. Qualification of the sensors will be based on production batches. A batch consists of approx. 20 to 40 individual sensors, which were produced in parallel and where the properties are assumed to be similar. This results in 700-1500 batches in total. For each batch, a decision about acceptance must be made based on the inputs of the vendor measurements and the sensor and test structure tests made within CMS. Irradiation tests will also be performed on a small number of samples. For the given number of batches, 7 to 15 batches must be qualified each week.

The most important measurement is the dark current of each cell. This parameter must be measured on 100% of the sensors by the vendors. CMS re-tests a few per cent of the sensors on a sample basis using the switch/probe card setup (see below) to verify these measurements and to study their long-term behaviour under different environmental conditions (temperature, humidity). The fraction of bad channels to be accepted per sensor and their failure modes are not yet defined and will be the result of further studies on prototype sensors from all qualified vendors.

In addition to sensor testing, the stability of the production is monitored using so-called process quality control (PQC) measurements on test structures. These will also be done on a sample basis and will allow checking of many parameters that are difficult or even impossible to test on the actual sensors. This includes: sheet resistivity of pad implant, p-stop implant and metal, surface charges determined by MOS measurements, p-stop isolation potential by FET measurements and precise CV measurements on diodes.

7.1.5 Sensor testing results

To understand the behaviour of the sensors in the case of a pad with high dark current, the same sensors have been tested in both 1- and 7-needle configuration. The schematics are shown in Fig. 7.9 and a sample result is given in Fig. 7.10.

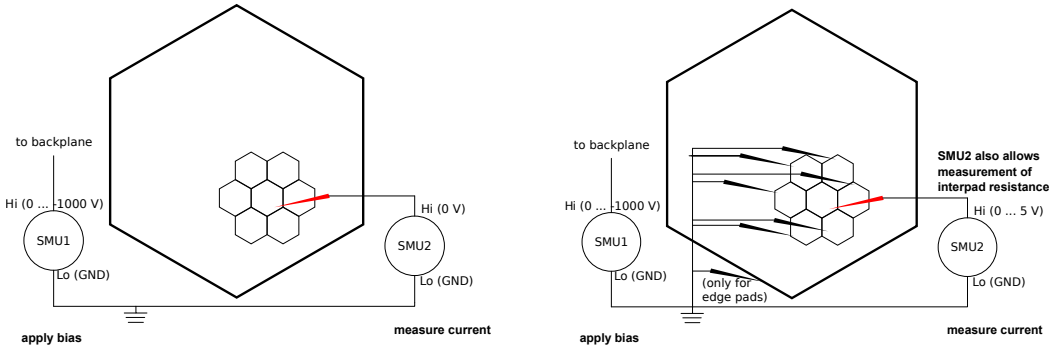


Figure 7.9: Electrical schematics of (a) the 1-needle, and (b) the 7-needle measurement configuration.

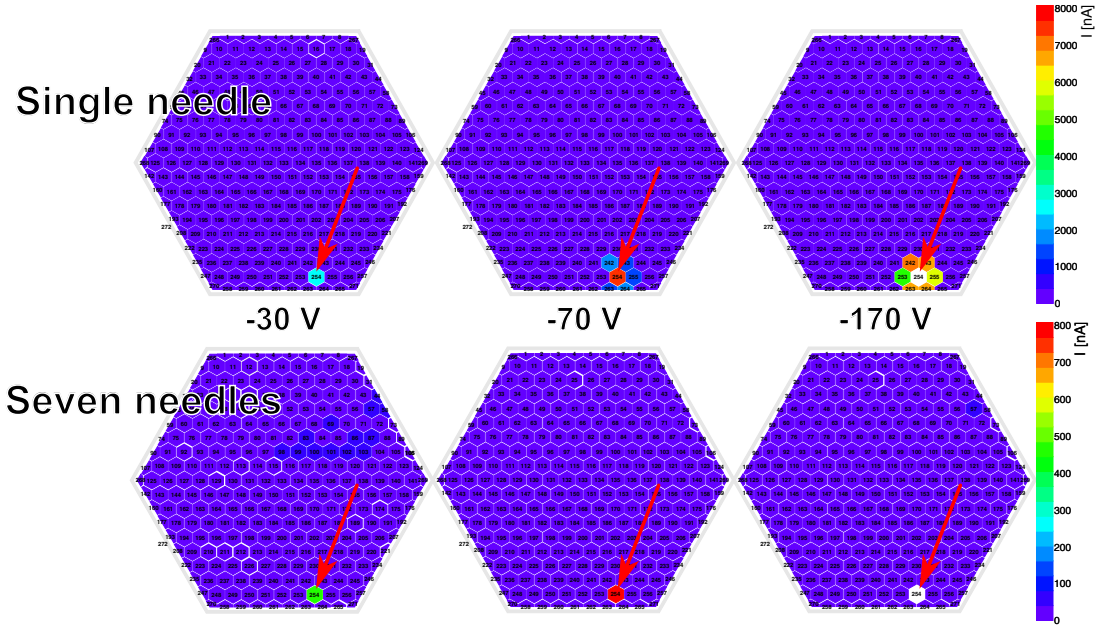


Figure 7.10: Results of current scan with one (top row) and seven needles (bottom row) at different bias voltages. The lateral spread of a high current from a weak cell (red arrow) is clearly visibly with one, but not when the adjacent cells are grounded (using 7 needles).

Figure 7.10 shows the lateral spread of the depletion zone and therefore indicates that future measurements should be performed at least with 7-needles.

7.1.6 Sensor testing using probe cards

Some 600 m^2 of silicon pad sensors have to be manufactured. During the prototyping and the mass production phases hundreds of these sensors will have to undergo quality control. Therefore it is essential to have a fast, automated way of characterizing them. Important characteristics to be measured include the breakdown voltage and leakage current for bias voltages up to 1000 V. Leakage current and the capacitance are measured as a function of the bias voltage in dedicated IV and CV scans. To extract these values, cells need to be biased as in standard operation. Therefore a probe-card based system that connects all cells on the sensor simultaneously, has been developed to determine the sensor's characteristics.

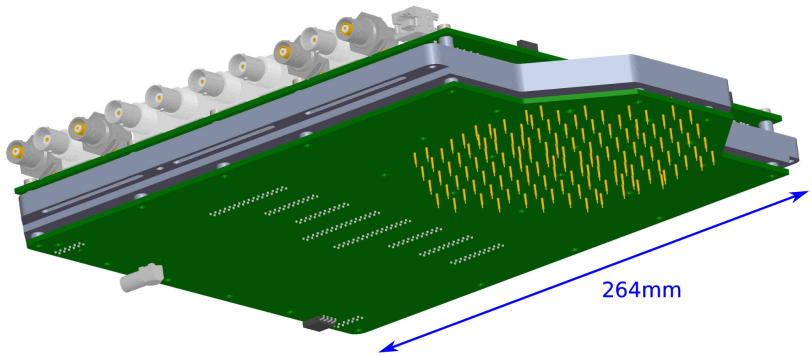


Figure 7.11: CAD drawing of the probe card for a 6 inch 135 cell sensor with the plug-in switch card on the top.

7.1.6.1 Testing setup

The testing setup will be capable of measuring multiple sensor layouts and it will handle up to 512 input channels. In order to fulfil these requirements, the probe card has two independent components, a switching matrix and a probe card, as shown in Fig. 7.11.

The simplified circuit diagram of the testing setup is shown in Fig. 7.12. The two cards are mechanically integrated into a probe station, which itself is inside a light-tight box.

The probe card consists of an array of spring-loaded pins called pogo pins that are used to contact all individual sensor cells. Test capacitances are added to the probe card to allow for system calibration and cross checks. Humidity and temperature sensors are integrated into the card to monitor environmental conditions. For each sensor geometry, a different probe card is mated to the switch card.

The switch card, a 512 channel multiplexer, sits directly on top of the probe card. The connections to the pogo pins on the probe card are made through 10 kΩ series protection resistors. The selected cell under test is individually connected either to an external pico-amp meter for the leakage current measurement, or to a precision LCR meter for the CV measurements. In the latter configuration, a 10 MΩ resistor is used to provide the bias ground return path.

The High Voltage (HV) source is connected via the switch card, routed to the probe card and applied to the sensor back side via the chuck. An on-board microprocessor is used for controlling on-board components and providing the user interface. The settings of the switching matrix can be changed manually or remotely via a USB interface. The two-card system has an

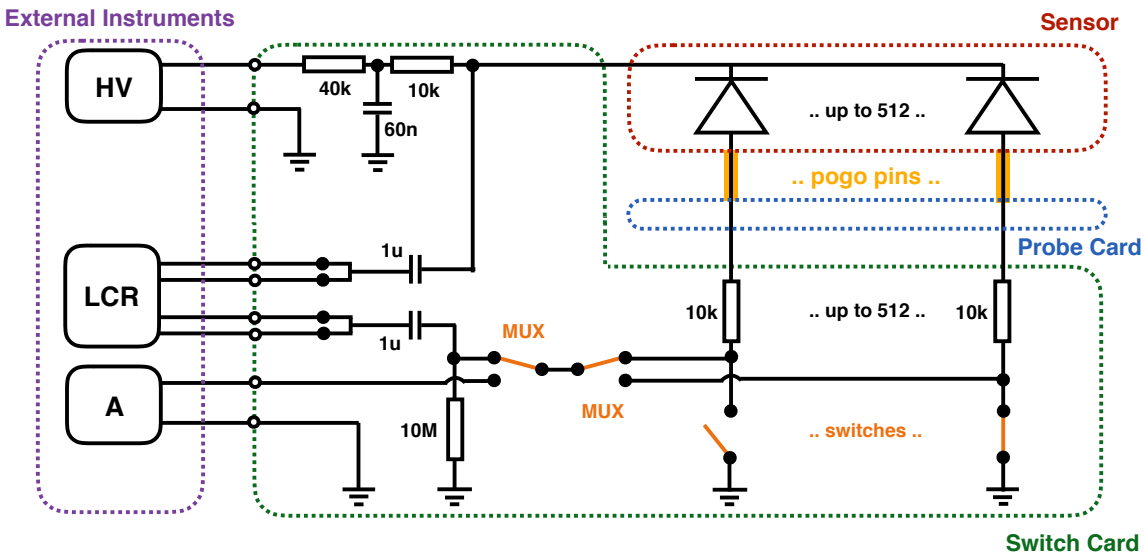


Figure 7.12: Schematic circuit diagram of the probe card switching matrix system. For CV measurements, the cell-under-test is connected to ground via a 10 MW resistor and for IV measurements via the ammeter. All other cells are directly shorted to ground. The switching scheme is such cells are never floating when changing from one cell under test to another.

average leakage current of less than 10 pA per channel and an average parasitic capacitance of roughly 50 pF at 50 kHz.

7.1.6.2 Data acquisition and analysis software

The communication with the instruments and the data acquisition is controlled using a LabVIEW software setup. It is used to perform IV and CV scans of all cells of a sensor without user intervention. A detailed IV or CV scan for a sensor with 150 cells takes about one hour for 15 voltage steps. For production testing a more streamlined protocol can be developed as necessary.

7.1.6.3 Measurement results

Results from one example IV scan of a HPK 6-inch 135-cell sensor are shown in Fig. 7.13 for a bias voltage of 1000 V. The measured leakage currents are normalised to a temperature of 24 °C. Figure 7.14 summarises the measurements for several tested 6-inch 135-cell sensors from different batches at 1000 V bias voltage. The results obtained with the probe card and switching matrix system are in agreement with those supplied by HPK. The sensors are of excellent quality. This can be seen from the remarkably low and homogeneous cell leakage currents, typically 1 nA or less at 1000 V. Only two of the 23 sensors in this sample had a single cell drawing in excess of 100 nA.

7.1.7 Sensor noise measurements

A hexaboard probe card (Fig. 7.15) is currently being developed for testing CE-E modules with irradiated silicon sensors. This will provide a measurement of the change in the electronic noise and charge collection efficiency (CCE). Special attention will be paid to non-Gaussian tails of the electronic noise called micro-discharges as observed in irradiation tests of the silicon sensors of the CMS tracker. In contrast to irradiation tests of full modules, these tests will focus

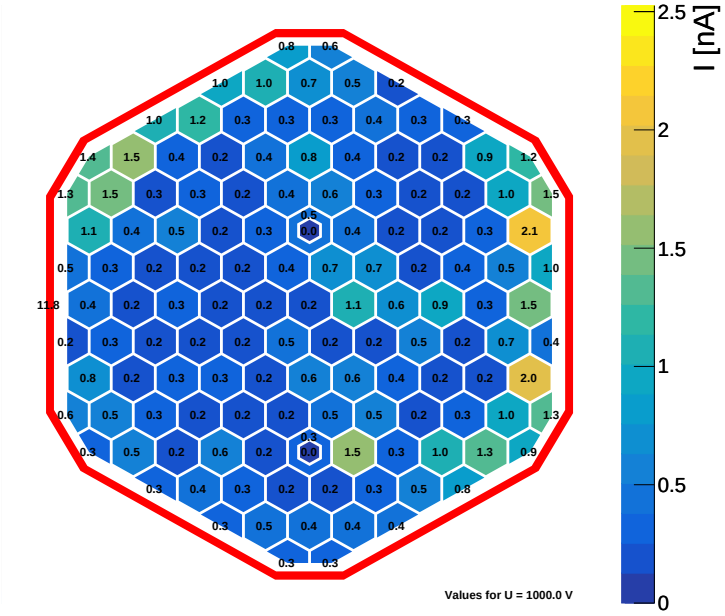


Figure 7.13: Per-cell leakage current for one example HPK 6-inch, 300 μm active thickness, 135-cell p-on-n sensor at a bias voltage of 1000 V.

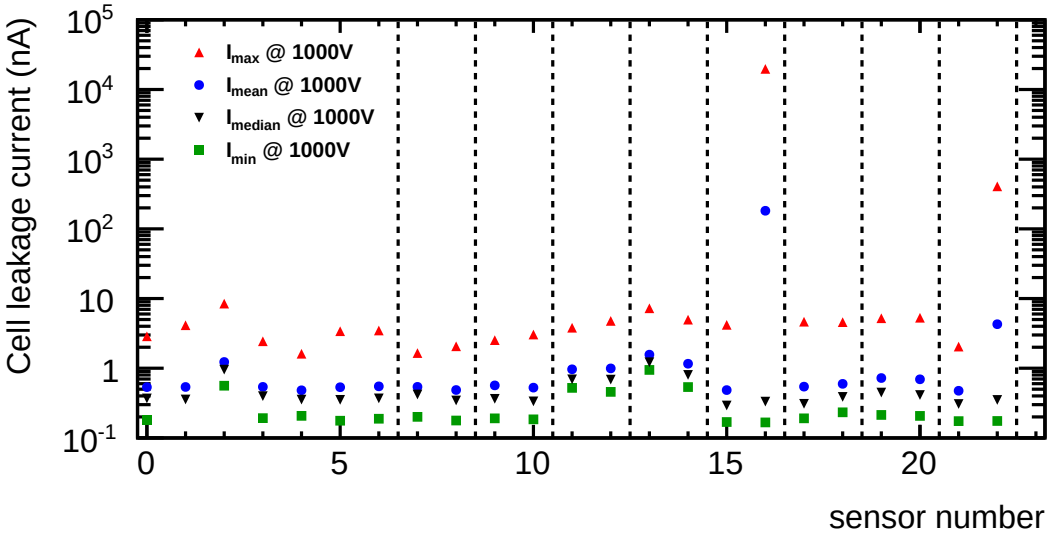


Figure 7.14: Per-cell leakage currents for all standard cells of several HPK 6-inch, 300 μm active thickness, 135-cell p-on-n sensors at 1000 V. This was a randomly selected subset from a first prototype series of 120 such sensors.The vertical lines indicate different sensor production batches. Sensor entries 16 and 22 both have one individual high leakage current channel as an outlier. The currents are normalised to a temperature of 24 °C. The vertical lines indicated different sensor production batches.

on the impact of the sensor irradiation only. The irradiated silicon sensor will be placed on the probe station cold chuck, which is cooled down to $-30\,^{\circ}$C to preserve the irradiation damage in the silicon sensor. The hexaboard probe card is then contacted temporarily with pogo pins to the silicon sensor, forming a temporary CE-E module. For studying the CCE, several holes were added to the PCB which allow laser light to be shone onto the metal opening of a few cells. Holes were added to all four quadrants in the 6" 135-cell HPK sensors that have different inter-pad distance (20, 40, 60, 80 μm). Infra-red laser light will be used to simulate MIP and larger pulses.

A bare sensor test setup is also under construction that can illuminate all cells on a bare sensor placed on the cold chuck of an automatic probe station. This setup will use a dedicated fast single channel amplifier and a 7-channel pogo pin card. This setup will allow for more detailed charge collection measurements with direct access to the sensor pulse shape.

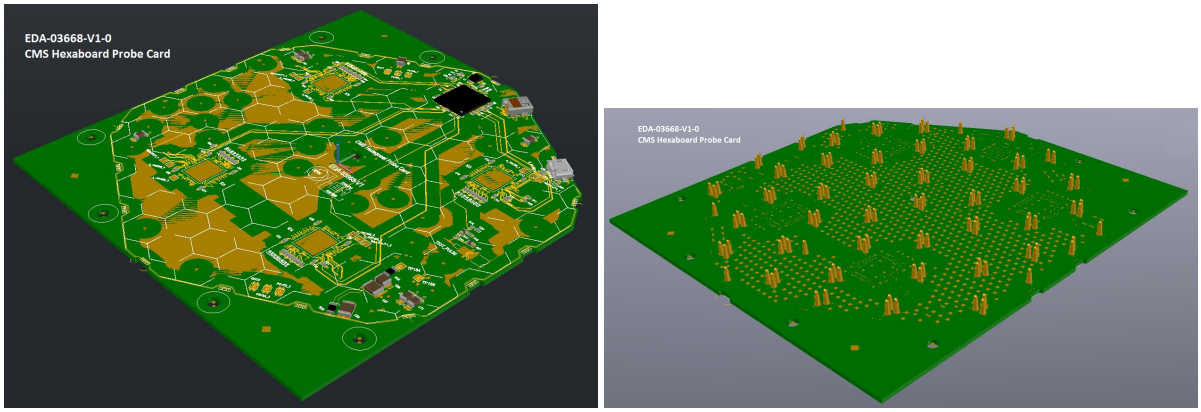


Figure 7.15: Hexaboard probe card, a hexaboard equipped with pogo-pin probe tips to directly connect a bare (irradiated) 6-inch 135 cell sensor to the readout electronics for signal and noise studies, viewed from above (left), and from below (right).

7.2 Silicon detector modules

The design of the silicon detector modules is given in Section 2.2. Here we provide detailed information about the automated assembly and testing procedures.

7.2.1 Silicon module assembly requirements

The epoxy used to adhere the module layers to one another is Araldite, which has been used successfully in several previous generations of Silicon detectors operated in high radiation environments. The assembly procedure maintains an epoxy layer thickness of $\approx$75 μm. Modules must have overall thickness within a specified range to ensure that they fit in the relatively tight spaces allotted for each layer of the CE-E. For the epoxy to get proper traction on the silicon during dispensing, the tooling must be within 350 μm of the sensor surface. This means that the height of the silicon should not vary by more than about $\pm$100 μm. The sensors and Au-Kapton sheets are manufactured very flat and uniform. The baseplate is the main source of non-uniformities but the epoxy layers under the Au-Kapton sheet and sensor provide freedom to adjust for the baseplate thickness variations to allow the silicon to lie flat in the completed module. A robotic pick-and-place gantry is used to place the sensor in the module with an accuracy of $\pm$25 μm along both axes in the plane, which is more than adequate to facilitate the alignment of pads as required for event reconstruction. It is also important for the PCB to lie flat once it has been mated to the silicon so that the wirebonder doesn't need to be adjusted for

the height of each of the holes where wire-bonds are made to the sensor pads. It also ensures that the bonder wedge tooling does not impact the PCB or silicon. The PCB position tolerances are less stringent than for the sensor since they need only assure positioning for wirebonding. The tolerance for PCB placement in the horizontal and vertical directions is $\pm 75\,\mu$m.

Although 8″ modules will be used for the final detector construction, the majority of R&D thus far has been done with 6″ modules. Almost everything learned about the assembly and performance of 6″ modules applies to 8″ modules. Exceptions are thermal warping and stress, which are potentially more significant for larger modules but have been found to not be critical, as mentioned above.

7.2.2 Automated silicon module assembly

Drawing upon extensive experience with module production for the original CMS tracker, the University of California Santa Barbara (UCSB) has the responsibility for developing all of the tooling and procedures involved in automated module assembly for the HGCAL. Once finalized, they will be replicated and distributed to all of the other module production centres. Module production centres will each be commissioned and qualified with dummy components, and then ramped up to full production capacity as needed

Module production is currently allocated roughly two years in the overall HGCAL schedule and will be carried out by five production centres each taking responsibility for $\approx$5000 modules. This corresponds to an easily sustained, nominal production rate of about ten modules per day. However, past experience has shown that the time required to finalize front-end electronics and other key components can sometimes be longer than desired, creating the need for a compressed module production period. For this reason, the production centres are designed to have peak module production rates that are significantly greater than the nominal rate. To this end, tooling, machines, test equipment, space, storage, etc. are being specified to enable production of 24 high quality modules per day per production site per eight hour work day. This would allow the full production of all modules in less than one year, should that be necessary.

Mating of the main components of the module is carried out with a pick-and-place gantry. The baseplate, Kapton, silicon and PCB are mated together in that order. This procedure is completed in three steps, each with a 24 hour curing period. The attachment of the Kapton to the baseplate does not require a gantry and will be done prior to shipping the baseplates to assembly centres during final production. To increase throughput, assembly occurs on portable trays which can enter and leave workspaces as needed.

The gantry provides translational motion in three dimensions as well as rotations. Prototype tooling for HGCAL module production for use by the gantry in module assembly has been designed and manufactured by engineers and machinists at UCSB. The gantry used at UCSB is an Aerotech AGS10000. It is same gantry used in the production of 4500 modules for the CMS tracker in a 13 month period starting in 2005. It has been recently refurbished by the manufacturer with the addition of several modern upgrades. It has very powerful, fast and accurate motors that deliver reliable performance within the sub-micron level. New production sites will purchase newer and better versions of this machine that are now available.

The gantry uses an Aerotech A3200 controller for precision motion of the gantry head. The controller has a LabVIEW interface using object oriented programming. The moving head can be equipped with a camera for optical locating, tools for picking-and-placing and for epoxy dispensing. The optical system detects fiducial marks in the ranging from 50 to 150 μm in size that are used to calculate the positions of components in the gantry workspace. A simple

fixture holds an epoxy filled syringe barrel which has a pressure line running to a separate pneumatic controller. An EFD Ultimus V glue dispenser manufactured by Nordson EFD is used for epoxy dispensing. This device controls the pressure and vacuum to high precision through the line to the syringe barrel. Several hundred trials were required to optimize epoxy dispensing routines for module production. The final epoxy pattern is seen in the left photo of Fig. 7.16. A compressible pick up tool can also be fixed to the rotational stage for retrieving and aligning module components. This tool has two vacuum lines to hold multiple objects simultaneously, and to efficiently transfer the vacuum to tooling fixed in the gantry workspace to hold pieces in place for curing, while freeing up the gantry for work on other assembly tasks. The gantry in the process of placing the PCB on a module is shown in the right photo in Fig. 7.16. These examples demonstrate how programmably controlled pneumatics are essential to the assembly process. Portable fixtures, components and tooling must be reliably released and secured throughout assembly. An array of valves made by Festo Corp. are wired to an I/O controller which is then commanded by LabVIEW programming on a desktop computer. For early experimentation an Arduino board was used to perform these operations. For higher throughput and expensive components the I/O control needs a more robust and reliable DAQ system. Currently, an NI cDAQ-9178 by National Instruments is used.

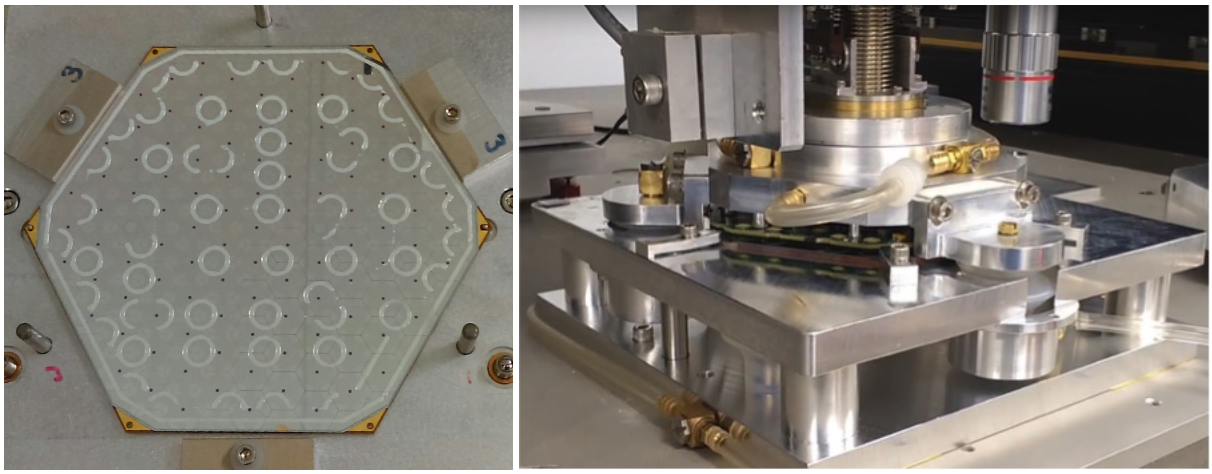


Figure 7.16: Automated epoxy deposition on sensor which has already been epoxied and cured to a Cu/W baseplate with Au-Kapton layer, (left) followed by placement of the hexaboard.

Once all layers have been assembled and the last epoxy layer has cured overnight, the module is fully wirebonded to close the connection of the front-end electronics on the PCB to the sensor pads in stepped holes, and to make the bias and guardring connections, as seen in the Fig. 7.17.

All wirebonding is done with a Bondjet BJ820 Automatic Wedge Bonder with a BK04 45 degree bondhead that can accommodate fine aluminium wire ranging from 17.5 μm to 50 μm diameter. A deep access bondhead is not necessary for HGCAL modules and the 45 degree bondhead can be used without any adjustments. The work holder consists of an adjustable table with a vacuum fixture that holds the module in place. Once the module is placed on the work holder, it is aligned in a few seconds by the wirebonder using four programmed reference points and four corresponding patterns. These points are used to establish separate reference frames that help position the bonds. After measuring the height of each reference frame, the program is ready to be executed. A nominal 128 channel module requires over 400 wirebond wires. The current programs have been written for a single module, but the work space on the BJ820 bonder is roughly 30.5 cm $\times$ 40.5 cm so that one could update the programs to accommodate multiple modules. The BJ820 bonder includes a process-integrated quality control

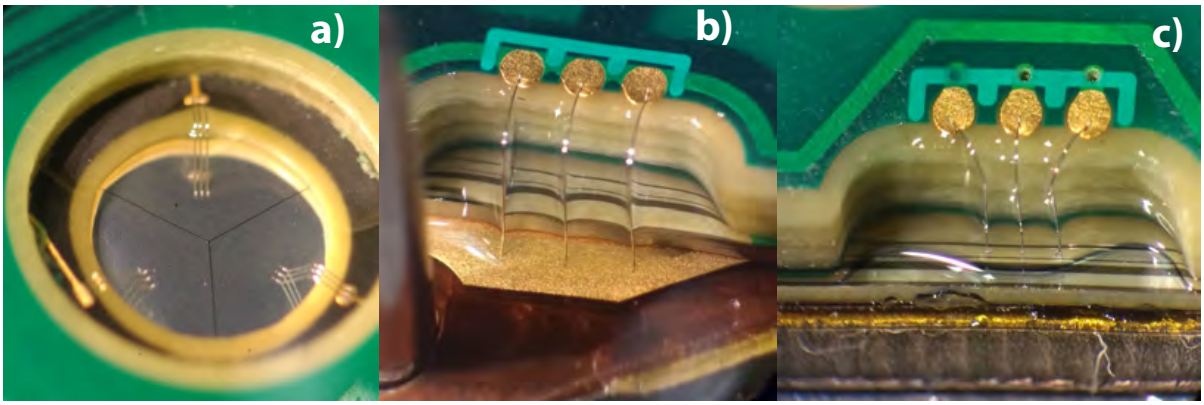


Figure 7.17: From left to right: a) wirebonds for three sensor pads at a stepped hole in the hexaboard; b) wirebond to the Au-Kapton layer to provide back-plane biasing of the sensor; and c) wirebonds at the edge of the module to the sensor guard rings.

system that provides feedback on current, frequency, friction, and wire deformation for every bond. This allows the quality of the bonds to be assessed in real time and bonding parameters to be adjusted as needed. When all wirebonds have been completed, the module undergoes a careful visual inspection followed by a basic electronic functionality test. If no faults are found in the inspection or test, all of the wirebonds are then encapsulated with Sylgard 186, a clear silicon elastomer with very high radiation tolerance. The encapsulation process makes use of a separate and much smaller gantry with its own dedicated pneumatic dispensing system. Encapsulation takes about 20 minutes per module after which it is left to cure for 24 hours.

7.2.2.1 Silicon module testing

Module test stands are in the process of being developed. They will be used to perform quick operational tests of PCBs upon arrival at production centres prior to use in module assembly. Newly completed modules are tested again to assure full functionality and connectivity of the PCB to the sensor prior to wirebond encapsulation. The test stands will also have LEDs to inject light onto pads that are exposed in several areas of the module to verify module response.

Since the HGCAL is foreseen to be operated at $-30\,^{\circ}\mathrm{C}$, modules will be thermally cycled and electronically tested in a dedicated multi-module testing system after assembly. The testing duration will be about 24 hours and will be used to weed out modules with weaknesses that have a negative effect on their operation. A prototype single module system at UCSB makes use of a Peltier unit. It has also been useful for validating FEA simulations of mechanical strain and heat dissipation in modules. A system with capacity for around 20 modules will be developed for the full production stage of the project.

Upon receipt, components and other hardware are thoroughly inspected. Sensors are surveyed under high magnification for defects such as chips, scratches or damage from probe station tests. Baseplates are checked for dimensional accuracy. As noted above, PCB electronics are tested. All mechanical properties and test results are currently kept in an organized file system and MySQL database. Each produced module has a traveler record with all processing and test information that it has experienced. Public access to the database is via a web interface.

The wire bonding machine includes a process-integrated quality control system that provides feedback on current, frequency, friction, and wire deformation for every bond. This allows the quality of the bonds to be assessed in real time and bonding parameters to be adjusted as needed. When all wirebonds have been completed but before they are encapsulated, the

module undergoes a careful visual inspection followed by a basic electronic functionality test.

Shipping modules requires many precautions to be taken. The most delicate parts of the module are the wire-bonds. Vibrations that occur during transport can break bonds. Thus we encapsulate all bonds as described above. To further protect the modules during shipping, each module remains securely fastened to its aluminium carrier plate that has accompanied it throughout the wirebond, encapsulation and test stages of production. The plate is then mounted inside of a hard container for shipping. The containers are bubble-wrapped, boxed, and then double-boxed in a bed of packing peanuts.

7.2.2.2 High rate silicon module production

The most challenging part of the module assembly development effort is the development of tools and procedures for the automated assembly and wirebonding of an individual module. This has been achieved. The simultaneous assembly of multiple modules has some additional challenges, but these are not nearly as demanding. We are currently in the process of commissioning high throughput module production at the UCSB pilot site for 6-inch modules. We do not expect to have any difficulty completing this for the November 2017 milestone. It will be followed by commissioning of a high throughput 8-inch production line at UCSB by August of 2018. At the same time as this work is underway, module production sites will begin to be set up at various locations across the globe. Specifications for tooling, and procedures for assembly and testing that were developed at the pilot site, will be transferred to these sites for them to begin exercising in preparation for final module production.

All of the development effort must be revisited on many occasions in this preparatory period because the front-end electronics for the modules will be evolving and the module design will be evolving as well. We expect this to necessitate multiple revisions of the details of the tools, but do not expect it to create a need for a major alteration of the overall approach that has been developed. Similar changes in tooling, or additions to the tooling that has been developed so far, will be required for modules with greater numbers of pads, and for modules that are made up of partial hexagons as required for the HGCAL at the smallest and largest radii. Again, it is expected that these variations will require new tools and methods that are nevertheless very analogous to those that have been developed up to now.

7.3 Scintillators and photodetection

7.3.1 General

The critical design parameter for the scintillator section of the calorimeter is the signal-to-noise for minimum-ionizing particles at the end of operation. As can be seen in Fig. 7.18, the expected ionizing dose and 1 MeV neutron fluence vary quite significantly over the volume of the detector. The MIP signal, as long as it can be distinguished from SiPM dark noise, provides a stable calibration to track any signal loss in the light yield, or the SiPM. To provide sufficient margin for detector performance, we require a MIP signal-to-noise ratio of greater than five, after an integrated luminosity of 3000 fb^{-1}.

The MIP calibration process allows for an effective inter-cell calibration, which will be crucial to compensate signal loss over the lifetime of the calorimeter. It also allows for cost optimization in the design of the calorimeter in our case, where one technology provides better radiation tolerance but is more expensive than the other. The final choices of the scintillator material will be determined by considering cost, performance (including after 3000 fb^{-1}), and ease of

assembly.

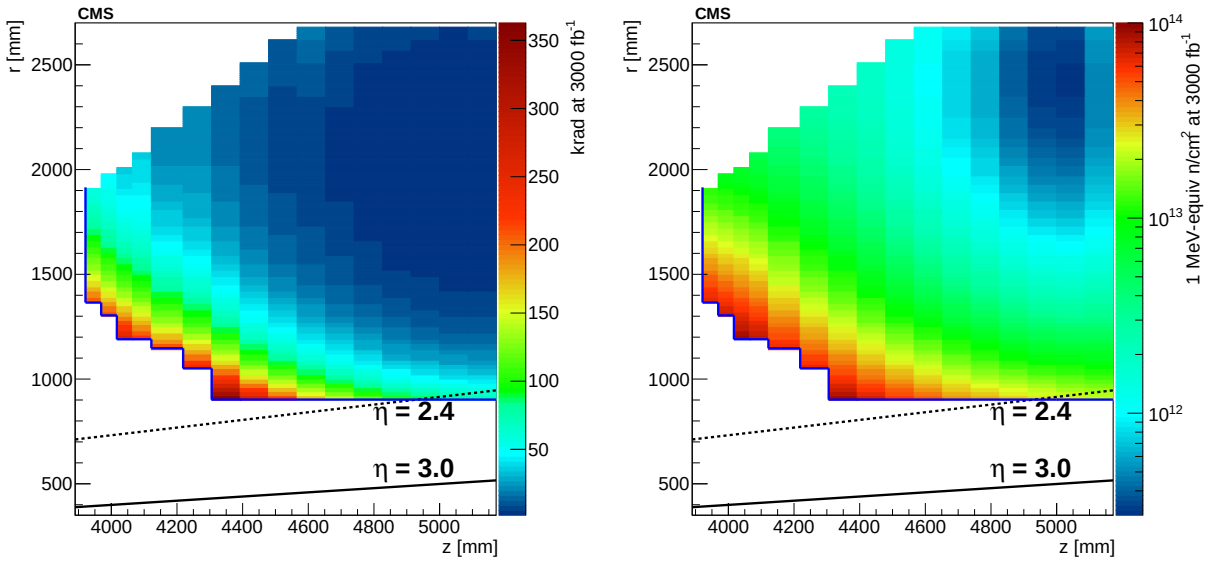


Figure 7.18: Ionizing radiation dose (left) and 1 MeV-equivalent neutron fluence (right) for the cells of the scintillator section of the endcap calorimeter.

7.3.2 Scintillator material

This section discusses in detail the characteristics of the scintillator material, including particularly the radiation tolerance, the impact of cell geometry on signal size, the construction of the tiles including surface treatment of the tiles, and the quality control procedures planned for the detector.

As discussed above, the design of the scintillator section uses cells that are ring-sections of fixed $\Delta\phi$ size with boundaries in radius from the beam pipe which are the same for every layer and which lead to approximately square cells. The detector layout requires tiles of 82 different physical sizes, 32 sizes for the $\Delta\phi = 1^\circ$ section and 50 for the $\Delta\phi = 1.25^\circ$ section. The tiles range in surface area from 4.0 cm^2 to 32.2 cm^2. Between 720 and 7000 units of each size are required, depending on the radius in question. Tables 7.5 and 7.6 show the details of the geometry of the scintillator cells, and their number.

7.3.2.1 Scintillator material characteristics

The parameters of the scintillator materials considered for the endcap calorimeter are given in Table 7.7. Appropriate materials must have good, well-understood radiation tolerance under low-temperature and low-dose-rate conditions. The scintillator decay constant should be fast to allow the use of a fast shaping-time in the amplifier. A fast shaping-time minimizes the impact of SiPM noise, which is a random stochastic process, as well as the impact of out-of-time pileup.

The radiation tolerance of EJ200 at low temperature has been studied in dedicated tests at NIST (Maryland) and Goddard Space Flight Center. Example spectra comparing the warm and cold irradiations are shown in Fig. 7.19. The samples in this measurement were EJ200-type scintillator using a polystyrene base with twice the standard amount of primary dopant. Increasing the amount of primary dopant somewhat increases the radiation tolerance of the material, though it decreases the initial light yield of the material. For the both warm and cold irradiations, the final irradiated signals are quite similar. There is a small amount of change

Table 7.5: Details of the cell counts and geometry for the 1 ° scintillator cells of the calorimeter.

Ring	Inner radius (mm)	Outer radius (mm)	Cell area (cm^2)	Count of cells	N(2 mm^2 SiPM)
f00	1195	1216	4.4	1440	2880
f01	1216	1238	4.6	1440	2880
f02	1238	1260	4.8	1440	2880
f03	1260	1282	4.9	1440	2160
f04	1282	1304	5.1	1440	2160
f05	1304	1327	5.3	2160	3600
f06	1327	1351	5.5	2160	2880
f07	1351	1374	5.7	2160	2880
f08	1374	1399	5.9	2880	4280
f09	1399	1423	6.1	2880	3600
f10	1423	1448	6.3	2880	3600
f11	1448	1474	6.5	2880	2880
f12	1474	1500	6.7	2880	2880
f13	1500	1526	7.0	2880	2880
f14	1526	1553	7.2	2880	2880
f15	1553	1580	7.5	2880	2880
f16	1580	1608	7.7	2880	2880
f17	1608	1637	8.0	2880	2880
f18	1637	1665	8.3	2880	2880
f19	1665	1695	8.6	2880	2880
f20	1695	1724	8.9	2880	2880
f21	1724	1755	9.2	2880	2880
f22	1755	1786	9.6	2880	2880
f23	1786	1817	9.9	2880	2880
f24	1817	1849	10.2	2880	2880
f25	1849	1882	10.6	2880	2880
f26	1882	1915	11.0	2880	2880
f27	1915	1949	11.4	2160	2160
f28	1949	1983	11.8	1440	1440
f29	1983	2018	12.2	1440	1440
f30	2018	2053	12.6	720	720
f31	2053	2089	13.1	720	720
Total			59.4 m^2	74 880	86 400

Table 7.6: Details of the cell counts and geometry for the 1.25 ° scintillator cells of the calorimeter.

Ring	Inner radius (mm)	Outer radius (mm)	Cell area (cm^2)	Count of cells	N(2 mm^2 SiPM)
b00	902	922	4.0	4608	5184
b01	922	942	4.1	4608	5184
b02	942	963	4.3	5760	6336
b03	963	984	4.5	5760	6336
b04	984	1006	4.7	5760	6336
b05	1006	1028	4.9	5760	5760
b06	1028	1051	5.2	5760	5760
b07	1051	1074	5.4	6336	6912
b08	1074	1098	5.6	6336	6912
b09	1098	1122	5.9	6336	6912
b10	1122	1147	6.1	6912	7488
b11	1147	1172	6.4	6912	7488
b12	1172	1198	6.7	6912	7488
b13	1198	1224	7.0	6912	7488
b14	1224	1251	7.3	6912	7488
b15	1251	1279	7.6	6912	7488
b16	1279	1307	8.0	6912	6912
b17	1307	1336	8.3	6912	6912
b18	1336	1365	8.7	6912	6912
b19	1365	1396	9.1	6912	6912
b20	1396	1426	9.5	6912	6912
b21	1426	1458	9.9	6912	6912
b22	1458	1490	10.3	6912	6912
b23	1490	1523	10.8	6912	6912
b24	1523	1556	11.3	6912	6912
b25	1556	1591	11.8	6912	6912
b26	1591	1626	12.3	6912	6912
b27	1626	1662	12.9	6912	6912
b28	1662	1698	13.4	6912	6912
b29	1698	1736	14.0	6912	6912
b30	1736	1774	14.7	6912	6912
b31	1774	1813	15.3	6912	6912
b32	1813	1853	16.0	6912	6912
b33	1853	1894	16.7	6912	6912
b34	1894	1936	17.5	6912	6912
b35	1936	1979	18.2	6912	6912
b36	1979	2022	19.1	6912	6912
b37	2022	2067	19.9	6912	6912
b38	2067	2112	20.8	6912	6912
b39	2112	2159	21.7	6912	6912
b40	2159	2207	22.7	6336	6336
b41	2207	2255	23.7	6336	6336
b42	2255	2305	24.8	6336	6336
b43	2305	2356	25.9	5760	5760
b44	2356	2408	27.0	5760	5760
b45	2408	2461	28.2	5184	5184
b46	2461	2515	29.5	5184	5184
b47	2515	2571	30.8	4608	4608
b48	2571	2627	32.2	4608	4608
Total			419.1 m^2	314 496	322 560

Table 7.7: Parameters of selected plastic scintillator materials. $T_{0.90}$ and $T_{0.95}$ represent the time required to integrate 90% or 95% of the total photons from an ionizing particle deposition.

Material	EJ200/BC408	IHEP SC-301
Light output (relative to anthracene)	64%	60%
Scintillation efficiency (γ/MeV)	9000	8500
Peak wavelength (nm)	425	420
Rise time (ns)	0.9	1.4
Pulse width, FWHM (ns)	2.5	3.4
$T_{0.90}$ (ns)	5.9	11.0
$T_{0.95}$ (ns)	6.8	19.0
CTE (K^{-1})	78×10^{-6}	70×10^{-6}
Base material	PVT	PS

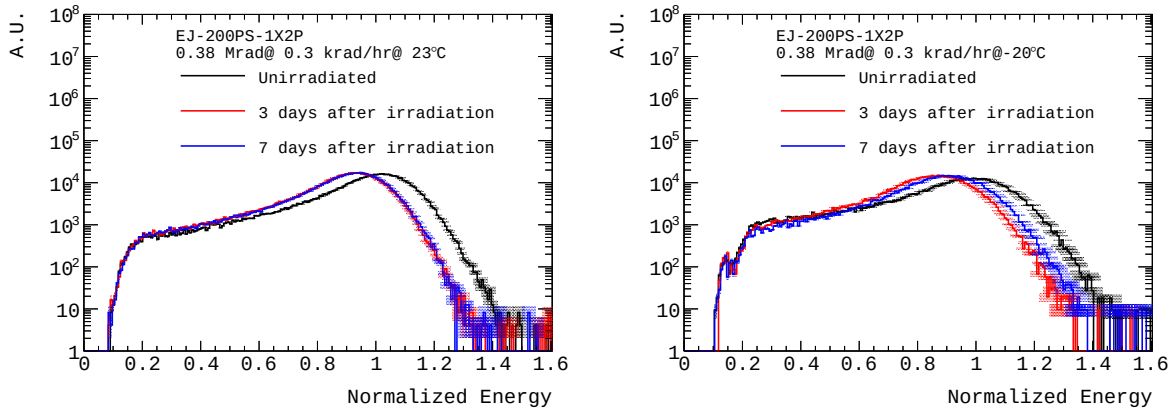


Figure 7.19: Spectra for alpha particle signals observed in EJ-200 polystyrene materials with double overdoping of primary dopant both before and after irradiation for a total dose of 0.38 Mrad at a rate of 300 rad/h. The left plot shows material irradiated at room temperature and the right shows irradiation and measurement at low temperature. The difference as a function of time shows the recovery of temporary damage through annealing in the cold sample. The logarithmic vertical scale is in arbitrary units.

observed between observations three days after the irradiation and seven days after the irradiation for the cold sample, while no change is observed for the warm sample.

This effect may be due to the production and annealing of temporary damage in the scintillator. During irradiation, temporary damage will both accumulate and recover, resulting in a dynamic balance – a process seen over short timescales in the CMS $PbWO_4$ crystals. As seen in the longer time series of Fig. 7.20 (top), the measurement of the small damages seen here is subject to significant uncertainties. In the case of temporary damage, a lower dose rate will reduce the equilibrium level of damage. The maximum dose rate expected in the scintillator for this upgrade is less than 30 rad/h, so the effect of temporary damage will be minimal and the dose constants observed with warm irradiations can be applied to cold conditions.

The effect of dose rate on the temporary damage can be seen clearly when considering the higher dose rate (74 krad/h) tests shown in Fig. 7.20 (bottom). In these results, samples kept at room temperature are compared with samples held at $-30\,^{\circ}\mathrm{C}$ up to day 101, when the samples were allowed to warm to room temperature. The results clearly indicate the presence of a temporary damage process which recovers over time, but more slowly when the material is held at a low temperature. However, the comparison of the asymptotic behaviour of dose constant shows the permanent damage to be lower for the material irradiated at low temperature. This figure also compares the performance of EJ200-type scintillator with the standard amount of primary and secondary dopant (1X1P) with sample having twice the usual amount of primary dopant (1X2P). Comparative studies of the radiation tolerance of EJ200 and SC-301 are currently underway.

The overall performance of scintillator as a function of dose rate is shown in Fig. 2.10, demonstrating the known effect of higher specific rates of damage at lower dose rates, due to effects driven by oxygen chemistry. Given the long timescales required for low-dose irradiation, the data at the low dose rates expected in HL-LHC are entirely from the in situ measurements of the current CMS endcap calorimeter. These measurements include the effect of signal degradation due to decreased light output from tiles but also include the effect of signal degradation in the wavelength-shifting fibres and clear fibres of the current calorimeter as well as effects in the hybrid photodiodes used for readout in this period. We use the "best 10 HPD" fit results which were derived for the barrel TDR [29] to provide performance projections:

$$D_c = 3.6\,\mathrm{Mrad}\sqrt{\frac{R}{1\,\mathrm{krad/h}}}.$$

The dose constant, D_c can be used to project the loss of signal from a tile during HL-LHC operation, given the dose rate and total dose for the tile. As the target integrated luminosity for HL-LHC is $3000\,\mathrm{fb}^{-1}$ at a levelled luminosity of $5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, we can extract an expected runtime of 16.7 kh. This allows us to express the dose rate R in terms of the total dose D, so that projections are based on:

$$A(D) = A_0 \exp\left(-\sqrt{\frac{D}{776\,\mathrm{krad}}}\right),$$

where $A(D)$ is the light yield after an integrated dose D, and A_0 is the initial light yield.

If the same integrated luminosity were delivered at higher dose rate, the damage would be somewhat smaller, while if a higher integrated luminosity were delivered over the same time, the same equation would apply with the larger total dose delivered. Figure 7.21 shows the projected fraction of the original scintillation signal remaining after $3000\,\mathrm{fb}^{-1}$, and the projected signal amplitude per mm^2 of SiPM surface area, for minimum-ionizing particles.

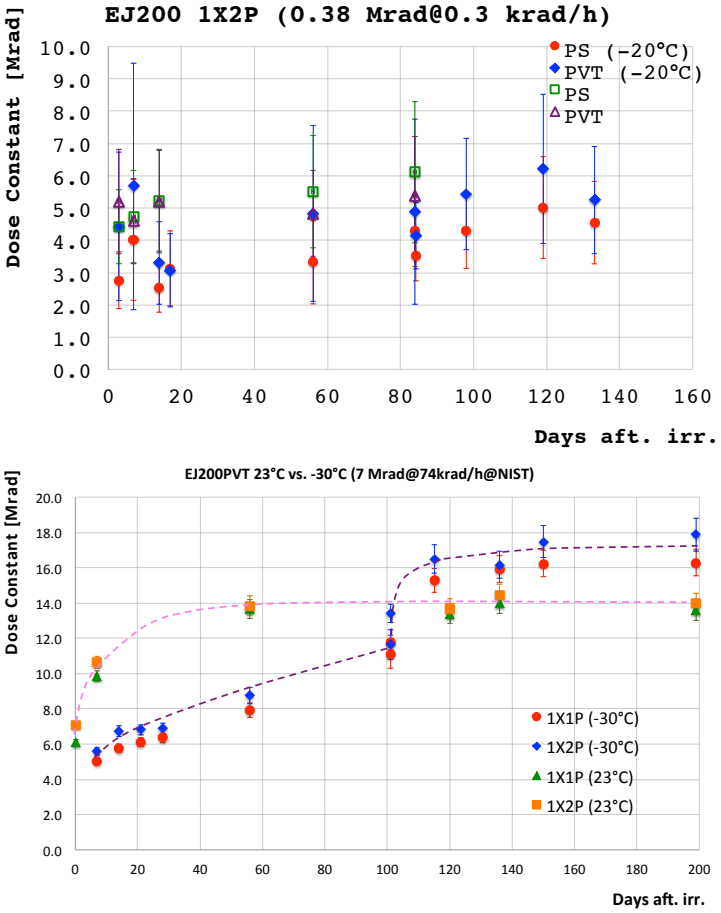


Figure 7.20: Recovery curves for radiation damage observed in EJ200-type scintillator material at reduced temperature compared with room-temperature irradiations. The top figure compares EJ200 PS and PVT materials irradiated at 300 rad/h, all of which contain double the usual primary dopant. At the low dose rates, the permanent damage to the scintillator completely dominates the transient damage, resulting in minimal time-dependence of the dose constant after irradiation. The bottom figure compares the behaviour of scintillators irradiated at 74 krad/h. The 1X1P sample uses standard doping for EJ200, while the 1X2P sample has twice the usual amount of primary dopant. The slowed transient damage recovery is quite evident at lower temperatures for this relatively high dose rate. At 100 days, the cold sample was warmed to room temperature.

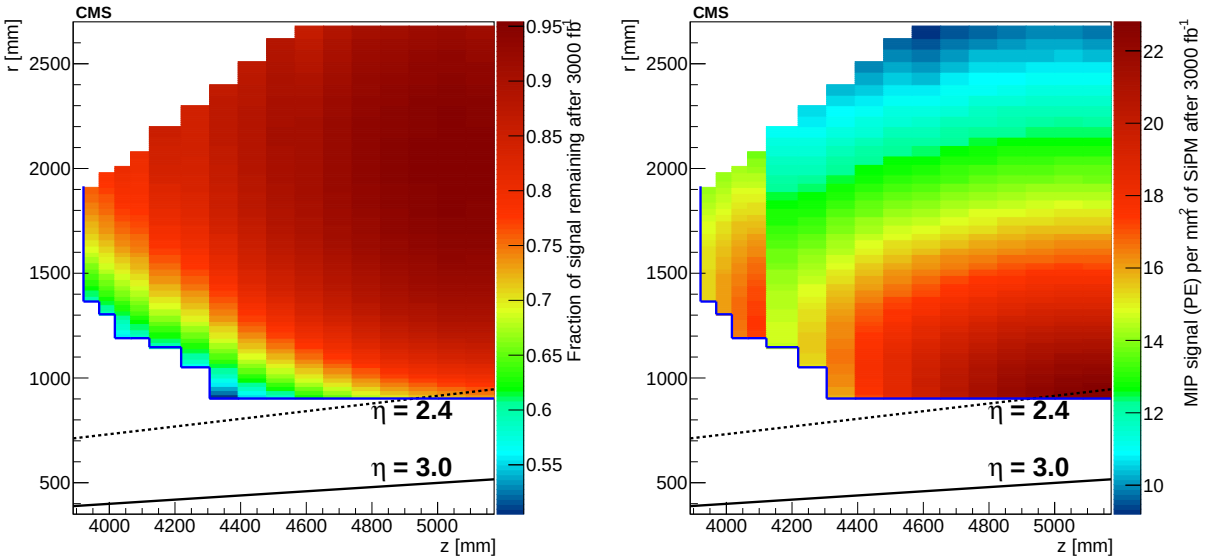


Figure 7.21: The projected fraction of the original scintillation signal remaining after 3000 fb^{-1} (left) and the projected signal amplitude for minimum-ionizing particles in units of photoelectrons (PE) per 1 mm^2 of SiPM surface area, taking into account both radiation and geometrical effects (right).

7.3.2.2 Geometrical effects on tile response

The light collected by the SiPM from the scintillator tiles in an SiPM-on-tile technology depends inversely on the average path length between the SiPM and production point of the scintillation light. For a nearly-square tile, this distance is proportional to the square-root of the area of the tile. Given a reference tile of known area A_r producing an amplitude S_r, the signal amplitude S_x for a larger or smaller tile can be predicted as

$$S_\mathrm{x} = S_\mathrm{r}\sqrt{\frac{A_\mathrm{r}}{A_\mathrm{x}}}.$$

As discussed in Section 2.3.3, the detector design takes advantage of this relationship to increase the MIP signal for tiles at smaller radii that see high radiation levels while allowing larger cells, with smaller MIP signals where the SiPM noise will be negligible, at larger radius. The illumination of the region where the SiPM is placed is also quite uniform over several millimeters distance from the nominal location of the SiPM. As a result, the electrical signal can also be increased by increasing the area of the SiPM used, though this will also increase the dark noise as discussed below.

The overall signal amplitude dependency on tile area and SiPM area (α) is therefore

$$S_\mathrm{x} = S_\mathrm{r}\sqrt{\frac{A_\mathrm{r}}{A_\mathrm{x}}\frac{\alpha_\mathrm{x}}{\alpha_\mathrm{r}}}.$$

This relationship has been validated experimentally in test beam studies as shown in Table 7.8. The test beam results also demonstrate general applicability of these geometrical scaling laws to multiple materials.

The uniformity of response to particles passing through different regions of dimpled tiles has been demonstrated with both radioactive sources and test beam [15, 38]. A uniformity scan in

a CALICE AHCAL prototype has shown that for 97 (81)% of the tile are the response is within 10 (5)% of the average of 20.6 detected photo-electrons [39]. Recent CALICE studies performed at the University of Tokyo have shown that a similar level of uniformity can also be obtained with a larger tile of $60 \times 60\,\mathrm{mm}^2$, using the same geometry for the dimple [40]. Some tuning of the dimple geometry may be required to cover the full range of cell sizes required in the endcap. This R&D is currently underway at FNAL and NIU.

Table 7.8: Results from dedicated 2017 test beam studies of the effect of tile size and SiPM area on the signals from minimum-ionizing particles. The SiPMs are square, and the tiles are square and 3 mm thick. Tiles are wrapped in Tyvek as a reflector. The predicted light yield is based on first line of the table, applying the scaling laws for tile size and SiPM area.

Scintillator	Tile side dimension (cm)	SiPM side (mm)	Light yield (observed)	Light yield (predicted)
SCSN81	3	1.3	13	–
SCSN81	3	2.0	33	31
SCSN81	6	1.3	7	7
SCSN81	6	2.0	15	16
EJ200	3	1.3	14	13
EJ200	3	2.0	34	31

7.3.2.3 Tile construction

As is well-known, the response of a scintillator tile to MIPs depends not only on the characteristics of the scintillator, but also the surface treatment of the tile which serves to reflect both UV and visible light back into the tile. Traditional materials for wrapping scintillator tiles have included aluminium foil and Tyvek. However, recently new reflector materials have appeared on the market, particularly for use in flat-panel displays and for reflection of sunlight to decrease the need for electrical lighting. An example of such a material is the VM2000 Enhanced Specular Reflector (ESR) film produced by the 3M corporation. Test beam studies have demonstrated that the use of ESR film instead of Tyvek increases the amplitude of the MIP signal by a factor of more than 200% (35 PE compared to 15 PE).

Measurements from literature [41] indicate that the ESR film also produces a fluorescence signal. If this signal is slow and is the dominant reason for the increased signal with ESR wrapping, it could be problematic. Minimizing the SiPM noise requires minimizing the shaping time, which would effectively reduce the amplitude contribution from the slow component. A study of this possible effect has been carried out at the FNAL test beam, as shown in Fig. 7.22. The study, which used a comparatively slow SiPM, showed no slowing of the pulse shape due to ESR fluorescence having a major effect. Further studies are planned to check for short-time-scale effects using a faster photodetector and to check for any radiation damage to the characteristics of the ESR film.

Each tile must be formed to the proper shape and size, including the dimple for the SiPM. Two techniques are under consideration for the tile formation. Tiles may be injection-moulded using metal forms which define the shape of the tile as well as the dimple structure, or tiles can be cut using machining techniques from large cast scintillator sheets. Both techniques have been successfully used by the CALICE collaboration and by other groups working on SiPM-on-tile technologies. For PS-based scintillators, either technique can be used, while PVT-based scintillators require the use of the machining technique.

Each tile must be isolated from adjacent tiles to avoid cross-talk and have ESR film applied to

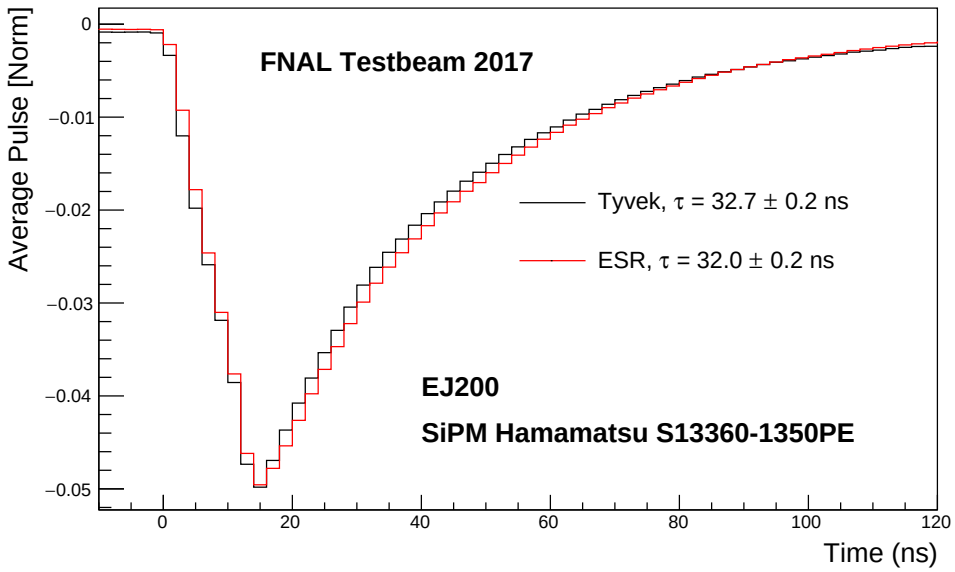


Figure 7.22: Comparison of the normalized average pulse shape for a tile wrapped with Tyvek with one wrapped with ESR film. The decay time is dominated by the RC time constant of the SiPM used for the measurement for both samples and there is no indication of an additional slow component from ESR fluorescence.

the top and bottom surfaces to maximize signal. Two techniques are under consideration for this step of the process. The CALICE collaboration has developed a technique for wrapping tiles with pre-cut ESR foils using an automated wrapping machine as shown in Fig. 7.23.

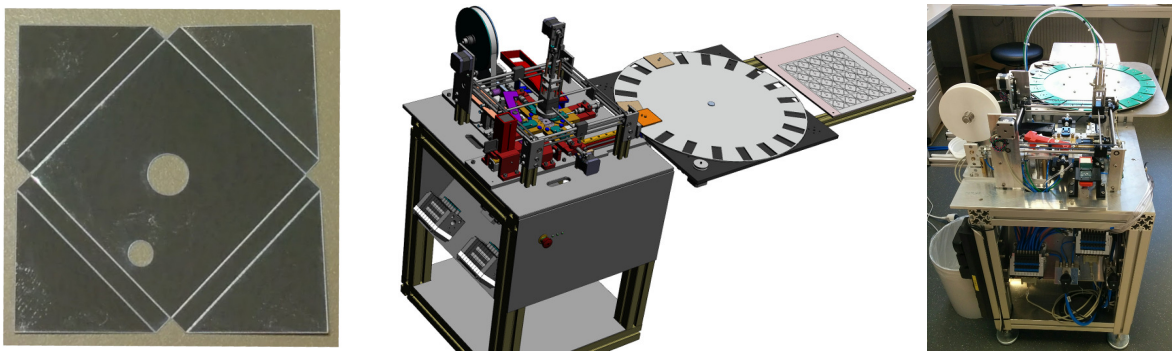


Figure 7.23: Left: Photograph of the laser-cut ESR foil wrapper for a scintillator tile, showing the detail of the SiPM dimple and hole for the LED light. The white lines show regions which are perforated by the laser to allow for reliable folding. Right: CAD design and photograph of a tile-wrapping machine developed for the CALICE calorimeter project.

The second concept under study is the construction of multi-tile units or "macrotiles" at once. In this concept, which is particularly well-adapted to construction though machining, the individual construction unit is a group of tiles rather than a single tile. A version of this technique was used for the construction of the fibre-based CMS HCAL barrel calorimeter. The CALICE collaboration has adapted the technique for use with the small SiPM-on-tile cells, where the issue of cell-to-cell crosstalk is more-important than in the large cells of the CMS barrel calorimeter. The process begins by milling grooves into the surface of a large sheet of PVT material which are then are filled TiO_2-loaded glue. Cast scintillator sheets typically have a thickness tolerance of ± 0.5 mm, with one surface being quite flat (typically defined by the bot-

tom glass plate during the casting process) and the other being somewhat uneven. As a result, surface milling is typically required for cast scintillator. In the macrotile construction process, the milling step is combined with the creation of the cells as shown in Fig. 7.24. After the cells are defined, ESR foil sheets are applied to the top and bottom, leaving a thin air gap between the scintillator and the ESR foil. Measurements by the Gutenberg University group in Mainz indicated a signal amplitude comparable to that with individually-wrapped tiles and cross-talk between cells at the level of 4% per side [42]. This technology may be particularly-applicable for the small cells at the inner radius of the detector, where the small size of the individual cells may be difficult to handle and process.

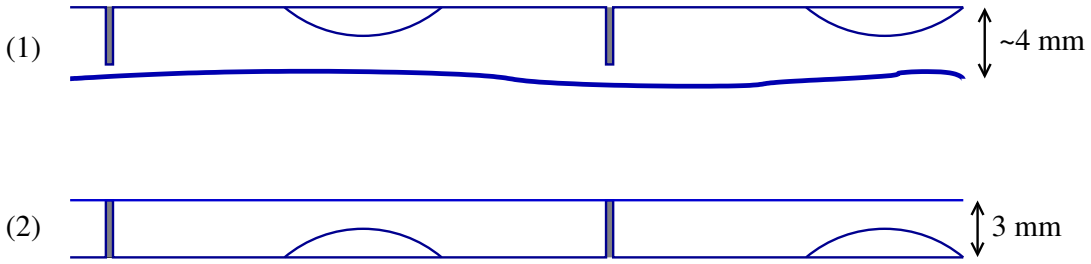


Figure 7.24: Process for constructing a macrotile containing multiple cells from a single sheet of cast scintillator. First, grooves are cut into the precision surface of the scintillator and filled with TiO_2-loaded glue. The SiPM dimples are also machined at this point. Then, the scintillator sheet is flipped and the surface is milled down to the desired thickness.

The size of individual scintillator units is limited in the CE-H detector by the differential thermal contraction between the scintillator cells and the readout circuit board. Each unit will be individually glued to the circuit board from its centre and will contract inward as the detector cools, leaving an inevitable gap between the units. The readout circuit board has a CTE matched to copper of $16 \times 10^{-6}\,\mathrm{K}^{-1}$, while the CTE for the scintillator is around $74 \times 10^{-6}\,\mathrm{K}^{-1}$. To keep the gaps between cells below 500 μm, the individual scintillator units should be no larger than 13.5 cm on a side. For individually-wrapped tiles, all tiles will be smaller than this limit, while for macrotile construction this requirement limits the number of cells which can be combined into a single mechanical macrotile.

7.3.3 Silicon photomultipliers

7.3.3.1 General characteristics of SiPMs

SiPMs are pixelated avalanche photodiodes operated in Geiger mode. Each pixel is connected via an individual quenching resistor R_q to the common substrate. The active zone is only a few microns thick, such that the bias voltage V_{bias}, typically 30–70 V, creates a field strong enough to trigger an avalanche from single photo-electrons.

Here a few characteristics for devices considered for the operation of scintillator-based calorimeters are given. They are operated at an over-voltage $\Delta V = V_{\mathrm{bias}} - V_{\mathrm{BR}}$ of typically 2–7 V above breakdown voltage V_{BR}. The gain is typically a few times 10^5.

For small light intensities, the total signal, given by the sum of the charges of all fired pixels, is proportional to the light intensity. For larger amplitudes, the finite number of pixels and the finite recovery time introduce an exponential saturation of the response with increasing fraction $N_{\mathrm{fired}}/N_{\mathrm{tot}}$ of fired pixels. It becomes proportional to approximately $1 - \exp(N_{\mathrm{fired}}/N_{\mathrm{tot}})$, which can be inverted for correction. For example, a device with an active area of 2 mm^2 subdivided into pixels of 15 μm size has ≈9000 pixels. For response of 30 pe/MIP the usable dy-

namic range is 600 MIP. For 1.5 TeV quark jets in tiles of $2 \times 2\,\mathrm{cm}^2$ a distribution of the maximum signal in a tile in each event was made. The mean of this distribution plus three times its RMS is $<$300 MIP, and this is taken as the requirement.

The temperature sensitivity of the breakdown voltage, typically 20–60 mV/K, introduces an effective temperature variation of the gain. This can be corrected for offline, or by adjusting the bias.

Thermally generated electron-hole pairs introduce a dark rate, which for recent devices is as low as a few 10 kHz per mm^2 at room temperature and scales with the area of the SiPM. Photons produced in the avalanche initiate after-pulses and inter-pixel cross-talk, which lead to correlated noise. The resulting excess noise factor may be as high as 1.2, meaning that the observed noise may be a factor of 1.2 higher than would be expected from Poisson fluctuations on the number of photoelectrons.

The operation of SiPMs at elevated radiation levels, and the investigation of SiPM radiation tolerance are at the forefront of detector research. The design and projections for the HGCAL strongly build upon R&D performed in the context of the CMS HCAL upgrade. However, the expected neutron fluence is substantially higher in the HGCAL case, and only in its cold environment ($-30\,^\circ$C) can noise levels be kept low enough to ensure sensitivity to MIPs throughout the lifetime of the HL-LHC.

The SiPM selected for the current endcap HCAL (HE) upgrade is the Hamamatsu MPPC S10943-4732 SiPM which has a diameter of 2.8 mm (6.15 mm^2 area) with typical $V_{\mathrm{BR}} = 66$ V. The parameters of these SiPMs are listed in Table 7.9. The dark current after irradiation to $2\times10^{13}\,\mathrm{n_{eq}/cm^2}$, at $V_{\mathrm{BR}} + 2$ V, at $-23.5\,^\circ$C, is 200 μA. This is equivalent to an RMS noise of 15 pe^- when integrating over 50 ns. The noise, in electrons, of these devices as a function of over-voltage and temperature is shown in Fig. 7.25.

Table 7.9: Measured characteristics of the HE upgrade silicon photomultipliers.

SiPM parameter	**S10943-4732**
Pixel pitch (μm)	15
Diameter of sensitive area (mm)	2.8
Operating temperature ($^\circ$C)	24
V_{BR} (V)	$\approx$65
Operating voltage (V)	V_{BR} + 4
Dark current (nA)	150
Photodetection efficiency at 550 nm (%)	30
Capacitance (pF)	215
Gain	3.5×10^5
Pixel recovery time (ns)	$\approx$10
Excess noise factor	1.18
Optical cross-talk (%)	17
After pulses (%)	$<$2%
$\mathrm{d}V_{\mathrm{BR}}/\mathrm{dT}$ (mV/$^\circ$C)	58.5
Temperature sensitivity (%/$^\circ$C)	3
Voltage sensitivity (%/V)	50

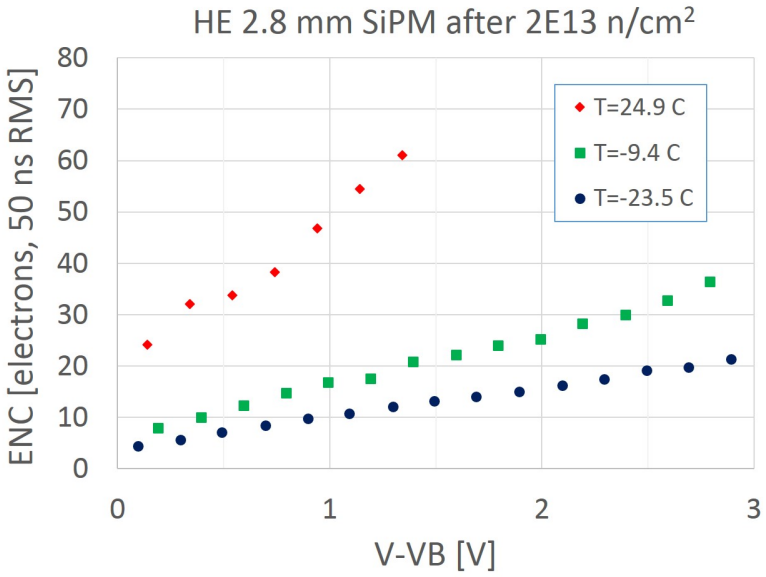


Figure 7.25: RMS electronic noise counts in units of photoelectrons as a function of overvoltage, for different temperatures.

For the design of the HGCAL, this has been used as a reference, together with the following scaling laws:

- The SiPM dark current increases linearly with fluence f, thus the noise RMS increases as $\sqrt{f}$.
- The SiPM dark current is proportional to the area A_S of the SiPM, thus the noise increases as $\sqrt{A_S}$.
- The SiPM temperature dependence for HGCAL SiPMs is determined to be a decrease in dark current by a factor of 1.88 per temperature decrease of 10 K.
- The observed dark current integrated in the amplifier is proportional to the shaping time t_{shape}, therefore the noise RMS increases as $\sqrt{t_{\text{shape}}}$.

Assuming operation at $T = -30\,^{\circ}\text{C}$ and a shaping time of $t_{\text{shape}} = 15\,\text{ns}$, the noise RMS is expected to scale as:

$$N = 2.7 \times \sqrt{\frac{f}{2 \times 10^{13}\,\text{n}_{\text{eq}}/\text{cm}^2}} \times \sqrt{\frac{A_S}{(1\,\text{mm}^2)}}.$$

Since the signal increases linearly with A_S, the signal-over-noise ratio increases as $\sqrt{A_S}$.

7.3.3.2 On-going R&D on SiPMs for HGCAL

The R&D will concentrate on radiation tolerance studies similar to those carried out by the CMS HCAL group with new devices to confirm the validity of the scaling laws, in particular the temperature dependence of noise for irradiated devices. In addition, several effects need to be taken into account that are not described by the simple scaling model, like saturation, radiation-induced break-down voltage shifts and self-heating. Figure 7.26 shows the projected SiPM noise level in both photoelectron equivalent (left), and the leakage current power (right), after 3000 fb^{-1}.

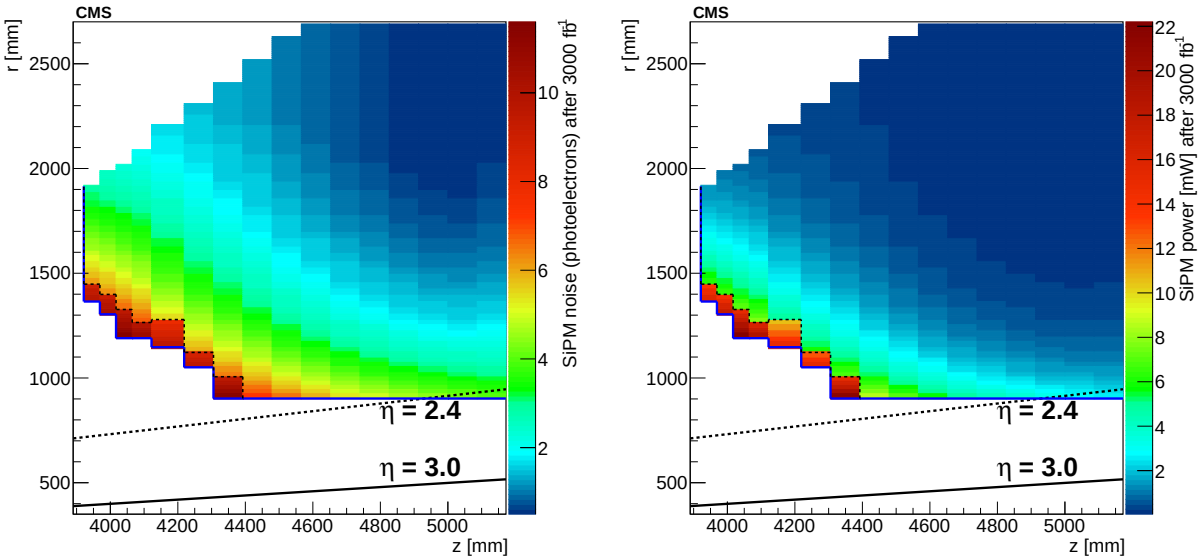


Figure 7.26: Projected SiPM noise level in photoelectron equivalent (left) and the leakage current power (right) after 3000 fb^{-1}.

For the noise rates of irradiated devices in the GHz range, the fraction of fired and not yet recovered pixels at any given time becomes non-negligible and decreases the light detection efficiency. This effect cannot be compensated by a fast shaping time of the pre-amplifier, but calls for devices with a short recharging time, which favours small pixel sizes (i.e. small pixel capacitances). The HGCAL projections are based, on SiPMs that have a shorter recovery time of 6 ns. This property is also important to ensure that a fast shaping time of the pre-amplifier does not reduce the effective gain, and thus the signal-over-noise ratio of the device.

At high noise rates, the currents, and thus power dissipation in the SiPMs becomes non-negligible and reach values in the mA range. The package of the sensor and its thermal conductivity are thus an important optimisation topic. Self-heating of the devices, with consequences for noise rates, break-down voltage and gain, must be avoided to ensure stable operation. Low gain (small pixel size) and low bias voltage help to minimise heating effects.

The reduction in noise due to decreased temperature is controlled by the temperature of the SiPM junction. Therefore, any heat produced in the device due to dark current must be efficiently removed to avoid the temperature rising and the noise rate increasing. The design of the package for the SiPM is therefore an important aspect. The Phase-1 HCAL SiPMs were packaged in a ceramic package which was highly thermally-conductive, but was quite expensive. Traditional plastic packages, using wirebonding to connect the signals from the SiPM, have poor thermal characteristics, resulting in significantly-higher noise rates observed for the same SiPM design after the same irradiation, compared with the ceramic packages. Recent technological advances including the use of through-silicon vias appear to allow both good thermal performance and low cost, but further testing is underway to confirm the performance and identify any concerns such as anomalous signal generation. Similar issues were identified early in the development of the Phase-1 SiPMs but were resolved through further studies in cooperation with vendors.

7.3.4 Scintillator/SiPM tileboards

The tileboards host the SiPMs as well as the readout ASICs as discussed above. An example is shown in Fig. 7.27. Following the CALICE design, the top side of the PCB, facing the scin-

tillator, holds only the SiPMs and connectors, while all other components are located on the opposite side, facing the copper cooling plate. This allows mounting, or gluing, of the scintillator directly to the PCB, without an extra gap, thus keeping the tile-module compact. The cooling plate must be in good thermal contact with the PCB. Thermal vias in the PCB remove heat from the SiPMs, which will become significant once dark currents rise as a result of neutron irradiation. Components on the back of the PCB will be collected close to the HGCROC and will be accomodated within a cut-out in the cooling plate. For the light injection LED drivers, it still needs to be demonstrated that with such a distance from the driver the optical signal quality is sufficient. This will be addressed in prototyping. The LED used to monitor the SiPM is envisioned to be placed within the same dimple as the SiPM, though it could have a separate dimple at a corner, if required.

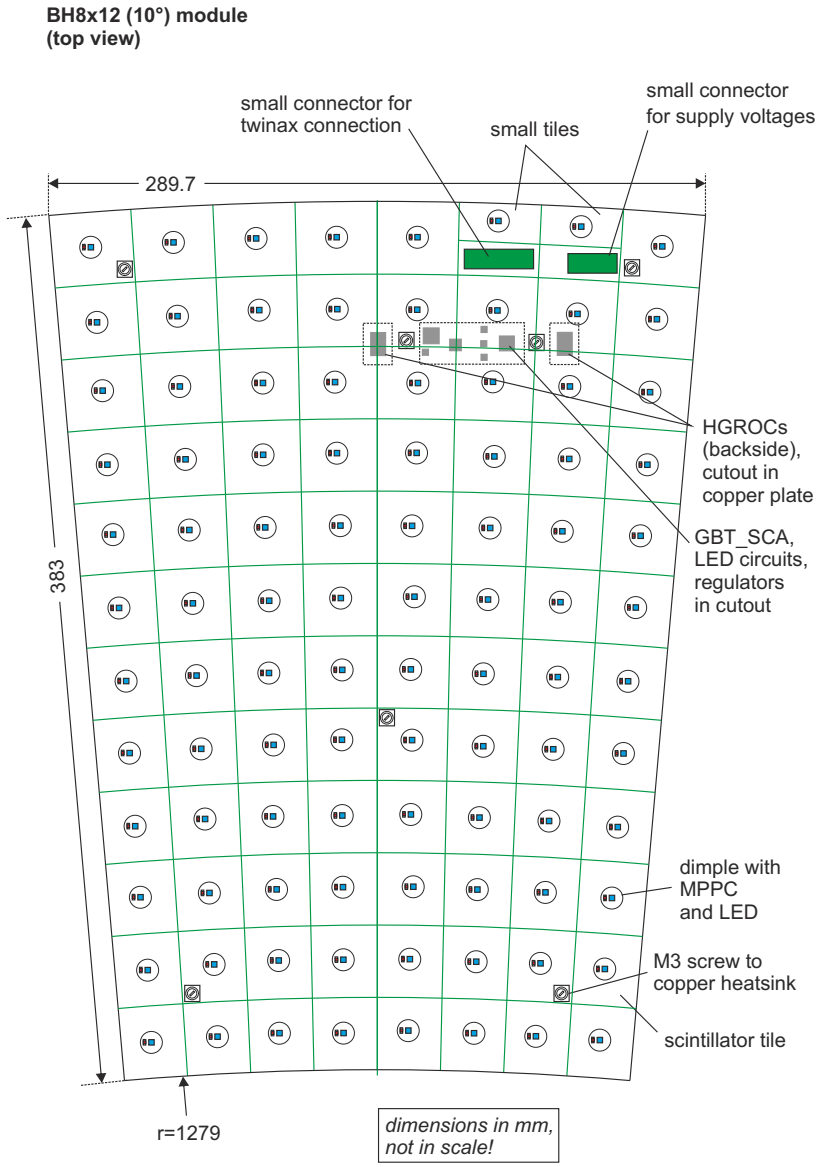


Figure 7.27: Example of a tileboard with scintillator tiles, SiPMs, readout and calibration electronics, and connectors for signals and power.

To cover the full surface of the detector while keeping the individual PCBs in the size range

Table 7.10: Tileboard geometry information and production counts. Dimensions and counts of cells refer to full boards, not partial boards as required to match the varying inner and outer boundaries as function of layer and azimuth.

Tileboard	Inner radius (mm)	Outer radius (mm)	Cells/ board	HGCROC/ board	Boards required
Fine-Inner	1196	1579	160	3	288
Fine-Outer	1581	2088	160	3	288
Coarse-A	903	1278	128	2	864
Coarse-B	1280	1661	96	2	864
Coarse-C	1663	2158	96	2	864
Coarse-D	2160	2685	80	2	792
			Totals:	8496	3960

acceptable by standard manufacturing, the design considers a minimum of six different tileboard PCBs. The number and basic parameters of these tileboards are given in Table 7.10. Two designs of tile boards are required in the $\Delta\phi = 1^\circ$ layers and four designs in the $\Delta\phi = 1.25^\circ$ layers. At the inner and outer edges of the scintillator region, whole tileboards cannot be used due to the overlap with silicon modules. In these areas, one of the basic six tileboards can be trimmed to meet the boundary requirements, either in the design stage or mechanically after construction.

7.3.4.1 Production strategy and quality control

The currently envisaged scheme starts from what is currently being exercised with the CALICE AHCAL prototype. With its 23 000 SiPMs the prototype corresponds to about 5% of the HGCAL scintillator sections. Because it involves considerably smaller amounts of human resources, infrastructure, and time than will be available for the HGCAL production this exercise serves as a convincing trial of large-scale production.

An overview of the production sequence is given in Fig. 7.28.

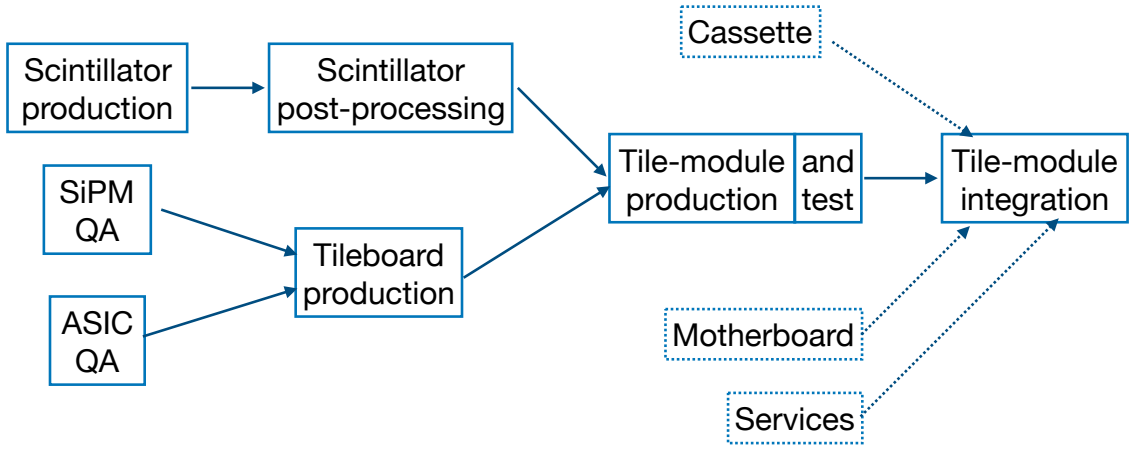


Figure 7.28: Tileboard production sequence.

Following sample testing of batches of production of SiPMs and automated testing of each HGCROC, the tileboards are electrically assembled and tested before mounting of the scintillator tiles, using the integrated LED system. Scintillator batches also need to be verified for light

yield and attenuation. After integration of the scintillator tiles with the tileboard, the completed tile-modules can be tested using the LEDs, and in addition a first MIP calibration, using cosmic rays or test beams, becomes possible. The LED test can also be repeated after eventual transportation, as a check-out before and after integration into a cassette, and at subsequent integration stages.

Well set specifications and QA/QC protocol at the vendors should guarantee good uniformity during production. Spot-checking of samples for individual production batches should be sufficient to track the production. The samples have to be large enough in order to characterise mean parameters and their variation.

Infrastructure for the production and characterisation of components for detectors with optical readout has been developed and set up in the framework of EC-supported integrated activity AIDA-2020. Reports on this activity can be found at Ref. [43].

For the SiPMs of the CALICE-AHCAL prototype, spot samples of 5% of each batch are tested on a test stand with a UV laser light source fanned out by a fibre system and equipped with parallel readout for 16 SiPMs.

The testing of scintillators obviously depends on whether individual tiles or mega-tiles are produced. The injection moulding technique yields very good uniformity after a certain running-in period. Tiles will be produced in several batches, and after each intervention, supply of new raw material, or each change of the mould for a new tile size, spot samples need to be tested.

Scintillator post-processing, in the case of individual tiles, entails wrapping the tiles in a reflector foil, using the automatic procedure developed at Hamburg University, and adaptable to different tile sizes. The procedure is expected to work for trapezoidal tiles, i.e. tiles with straight edges, which should be a sufficiently good approximation to the ring-sections of the HGCAL segmentation. After wrapping, it is important to verify the mechanical dimensions, in order to meet the precision requirements of the automated pick and place machine. This is currently done manually, using a caliper, but an automated optical inspection system is under development.

For mega-tiles, the procedures are subject of ongoing R&D. The raw scintillator plates need to be machined, to ensure uniform thickness, grooves have to be machined, filled, and the surface polished and painted. Light yield should be spot-sampled before and after post-processing. Uniformity and optical cell-to-cell cross-talk also need to be verified. Part of this can be done using tileboards, for other tests dedicated test stands need to be set up, and the spot sampling strategy defined. For the raw material, e.g., in the past simple current measurements using a radioactive source have been used.

The tileboard bare PCBs are electrically tested before mounting components. After mounting of ASICs, auxiliary components and SiPMs, the entire signal chain from optical excitation to the digitised output can be tested before mounting the scintillator tiles. This is done using the built-in LED system, by placing the tileboard on a glass plate. An example from the ongoing CALICE effort (Fig. 7.29) shows the observed pulse height spectra for a number of channels, which all exhibit clear single photo-electron signals, evidence of the good operation of the SiPM itself, its bias supply, and the analogue and digital parts of the signal chain function. The rather small signals are not affected by noise pick-up.

Mega-tiles would be mounted manually onto the tile-boards. The individual tiles can be placed using a pick-and-place machine as used for electronics components on a PCB. To efficiently feed the tiles to the machine, they are placed on reels. The placement of the tiles onto the reels is,

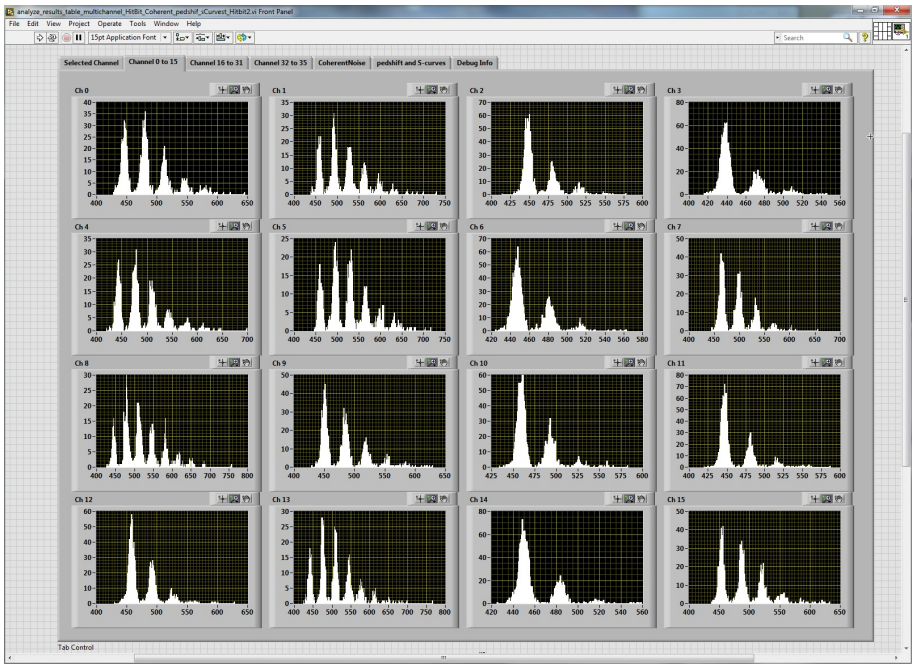


Figure 7.29: Pulse height spectra for a number of CALICE SiPM-on-tile channels.

in the CALICE case ($3 \times 3\,\mathrm{cm}^2$ tiles), handled by industry though the feasibility for larger sizes has to be checked.

The fully assembled tile-modules can detect particles and thus can be characterised using cosmic test stands or test beams, which also provides an initial calibration, i.e. equalisation of the cell response on the MIP scale. Much thicker stacks can be tested in muon beams. Typically several 100's of particles per cell are needed for a MIP calibration accuracy of a few percent. By stacking up several boards on top of each other, significant gains in speed can be realised. The existing CALICE test stand with an area of about $40 \times 40\,\mathrm{cm}^2$ has a trigger rate of 11 Hz, and for 30 boards stacked on top of each other creates 300 hits per second, corresponding to one second per channel needed for calibration.

The final step is the integration into the cassette. Here again the LED system is useful for verifying the electronic chain throughout the different steps of transport, integration and commissioning.

Chapter 8

Electronics and electrical systems

8.1 On-detector electronics

The physics requirements and a general overview of the readout chain are presented in Section 3.1. In this section, after a brief reminder of the general architecture, we present more details on the front-end electronics, the readout motherboards, the data format and transfer and the powering scheme.

8.1.1 Architecture

The electronic system for the HGCAL is shown in Fig. 8.1. Logically the system is separated in three paths:

- a DAQ data path,
- a trigger data path, and
- a control path (including clock, fast control, and slow control).

For convenience of construction, the paths are actually grouped physically into two different systems:

- the trigger data path, implemented over unidirectional high-speed links from the detector to the counting room, and
- the DAQ data and control path, implemented over an LpGBT-Versatile link system running bidirectionally between the detector and the control room.

Such paths actually come together at the optical link transport level, in the sense that they both share the same opto-electronic component (the 4 + 1 bidirectional transceiver, VTRX+ in the figure) that aggregates the traffic of all high-speed links used in the system.

As mentioned above, the data path itself carries two types of information:

- The data stored in a digital pipeline inside the front-end chip related to a specific event that has caused a level-1 trigger accept (L1A) to be generated.
- The data necessary for the computation of a trigger condition that need to be transmitted promptly to the trigger logic housed in the service cavern.

For simplicity, the first kind of data will be called DAQ data while the latter will be called trigger data.

The data path includes two ASICs. The HGCROC ASIC is the first and contains the signal amplification and signal shaping blocks, the digitization block, the processing and aggregation logic to construct the low level cells necessary for generating the trigger data, the data storage

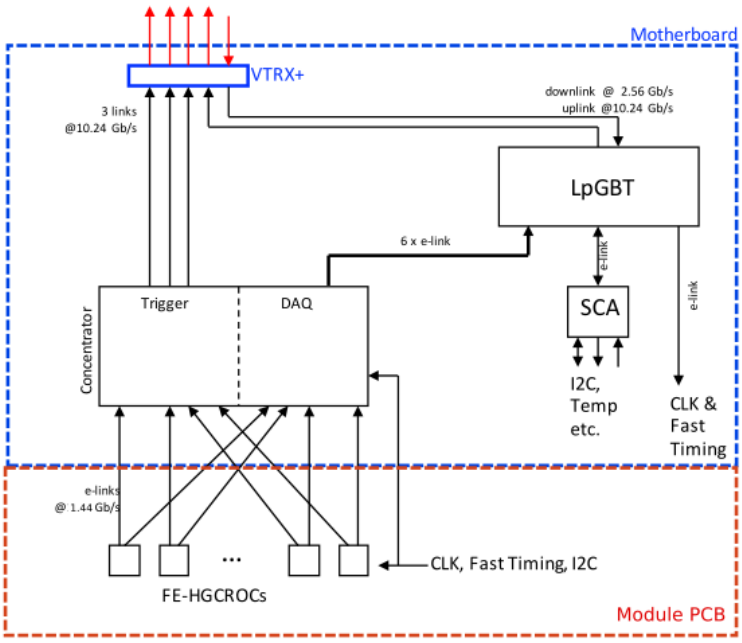


Figure 8.1: General architecture for HGCAL on-detector control and readout.

pipeline, the trigger pipeline, and all the necessary zero-suppression and data-formatting logic pertaining to the DAQ data.

A second level of data aggregation is performed in the concentrator' ASIC that collects trigger and DAQ data from a number of HGCROCs. The concentrator performs the appropriate data selection and reformatting function, and transmits the results to the external back-end electronics.

The fast (and slow) control part of the system is also shown in Fig. 8.1. It is controlled by an LpGBT, which distributes clock and fast signals to the front-end HGCROC, to the concentrator, and also controls a slow control adapter (SCA) chip. The SCA is responsible for the monitoring of environmental parameters and for downloading of front-end configuration parameters from the service cavern.

The data and control paths are designed to be fully independent, in the sense that data belonging to one never gets mixed with data of the other. The fast and slow control system must be present and operating at all times for the whole system to work. Such separation of functionalities is also important in simplifying the design of the firmware and software related to the two paths and simplifies the design of the external back-end electronics.

The trigger and DAQ data systems are unidirectional, carrying information from the detector to the control room. The fast and slow control system is bidirectional, as synchronization clocks and commands proceed from the control room to the front-end while acknowledgment and monitoring information proceeds in the opposite direction.

Trigger data belonging to several HGCROCs is filtered and assembled in the concentrator. It is expected that the total latency for bringing this information to the back-end electronics will be 0.9 μs (36 BX).

Physically, the embedded part of the electronics is assembled on two boards:

- A first board glued directly onto the silicon sensor is the module PCB, the hexaboard. This board contains the front-end ASICs and provides connectivity between these chips and the sensor.
- A second board, the motherboard, groups the hexaboards in larger physical and logical units. The motherboard houses the concentrator ASIC, the LpGBT and the optical modules for the LpGBT and for the trigger and DAQ data links, the SCA, and the voltage regulators needed for these components. It also provides all the local powering and service signals necessary for the operation of the hexaboards.

All interfaces for output data, control signal, and power of the motherboard are located at the external periphery of this board, with motherboards arranged physically in a "string" as shown in Fig. 3.8. Except for the cables used for critical monitoring functions (part of the detector safety system), all information to and from the embedded electronics travels on optical fibres.

8.1.2 Very front-end

8.1.2.1 Overview of the very front-end ASIC

The very front-end (VFE) ASIC incorporates the preamplifier/shaper/digitization, L1 storage, and data formatting for 78 channels, as well as the logic for aggregating trigger cell information that is sent to the trigger system in the service cavern. The VFE ASIC follows the architecture introduced in the technical proposal [1]. The schematic of the analogue part in the technical proposal is shown in Fig. 8.2. A more complete overview is given in Section 8.1.2.4. The large dynamic range from 0.2 fC to 10 pC is split into a linear range read by a 10-bit ADC and a time over threshold (ToT) range, when the preamplifier saturates, that covers 100 fC to 10 pC on a 12-bit TDC. Time of arrival (ToA) information is also available for signals larger than 10–15 fC. The total power dissipation must remain below 15 mW/channel and the ASIC must withstand radiation levels up to 2 MGy and 10^{16} n_{eq}/cm^2. A CMOS 130 nm technology is chosen.

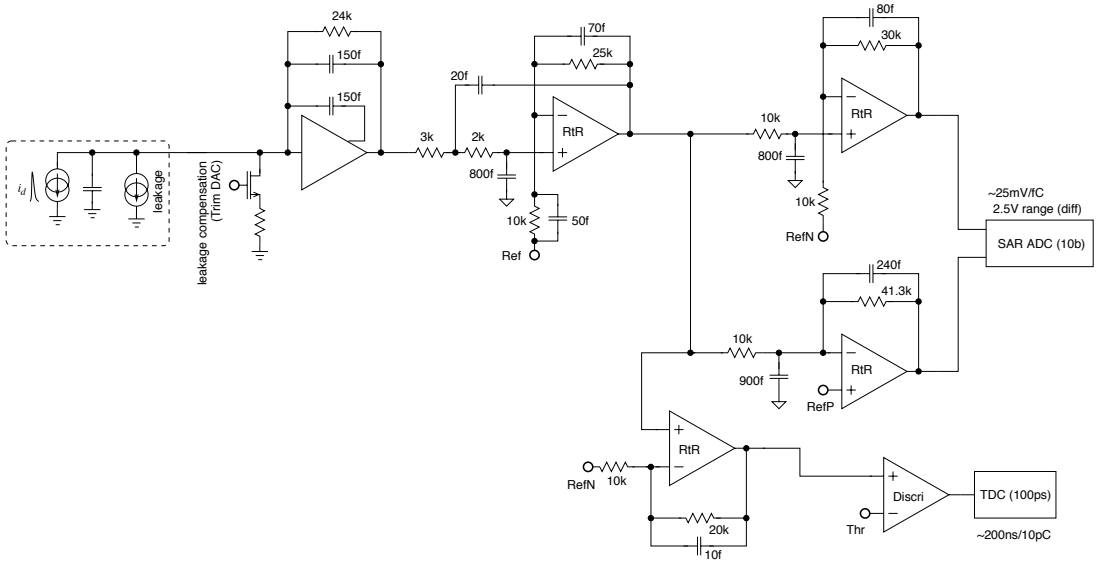


Figure 8.2: FE Architecture in the Phase-2 Upgrade technical proposal [1]

.

8.1.2.2 Experience with SKIROC2-CMS

In order to establish the performance of the VFE architecture, in particular the ToT, and study the system aspects on the front-end board, a test chip called SKIROC2-CMS was fabricated.

The chip was received in mid-2016. As its name indicates, it re-used many blocks from the SKIROC2 ASIC [32] developed for the CALICE ECAL, but was modified to run at LHC speed and to implement a ToT readout. This chip is detailed in Ref. [33]. Most analogue parameters (noise, speed) are not far from what is expected for HGCROC, allowing informative studies in test beam. The ToT starts to come into play around 500 fC, and exhibits a good linear behaviour as shown in Fig. 8.3. This test chip will allow to study, at the system level, the calibration of the non-linear part of the ToT, the stability of the ToT response, and the crosstalk when in saturation.

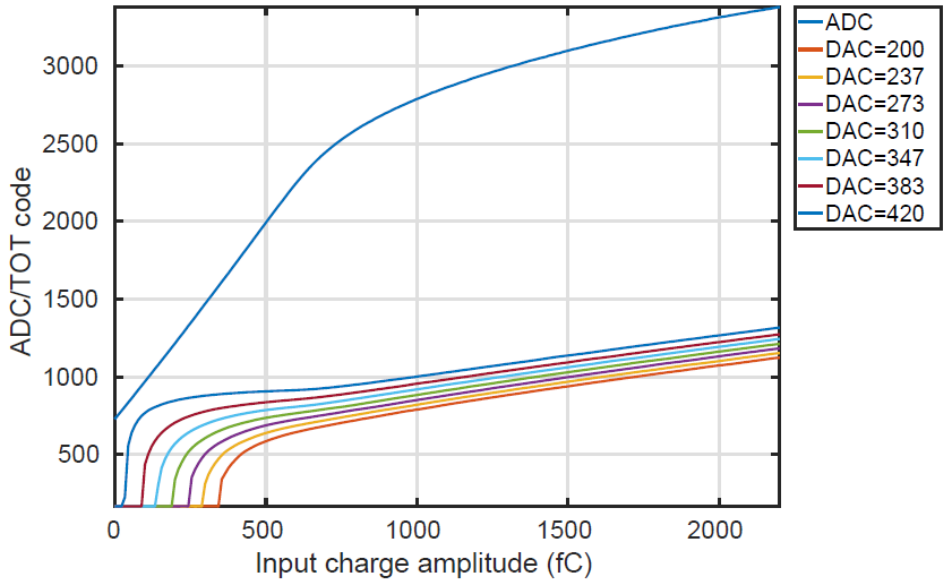


Figure 8.3: Linearity of the ToT in SKIROC2-CMS. The different curves correspond to different thresholds resulting in different response and overlap with the linear ADC range.

8.1.2.3 Test vehicles TV1 and TV2

In order to characterise and validate the analogue blocks, two test vehicles (TV1 and TV2) have been fabricated in May 2015 and December 2016 in the chosen CMOS 130 nm technology. Their layouts are shown in Fig. 8.4.

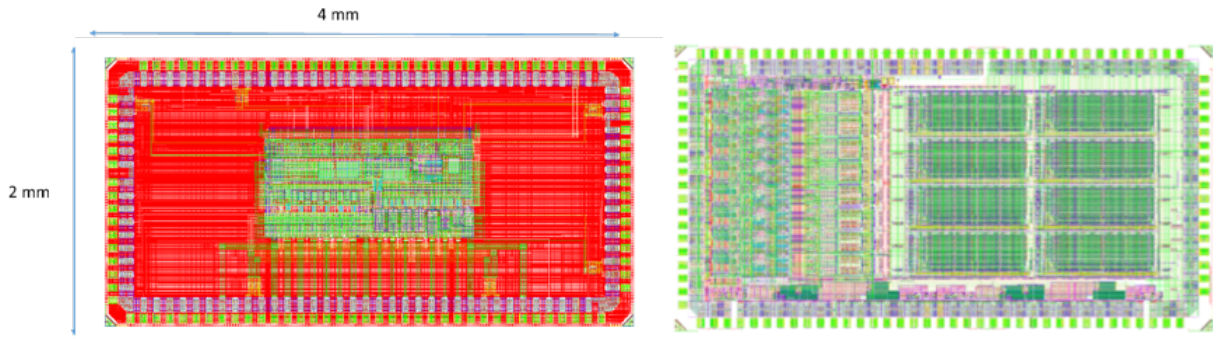


Figure 8.4: Layout of (left) TV1, and (right) TV2.

TV1 comprised different flavours of preamplifiers and shapers and allowed characterization of their speed and noise, which led to the choice of a high open-loop gain preamplifier configuration.

TV2 incorporated 8 channels with the full analogue chain, except the TDCs. The scheme followed closely what was described in the TP: a preamplifier with ToT, a 20 ns shaper and a 10-bit ADC, as shown in Fig. 8.5. A radiation hard 512×32b SRAM developed by CERN stores the data.

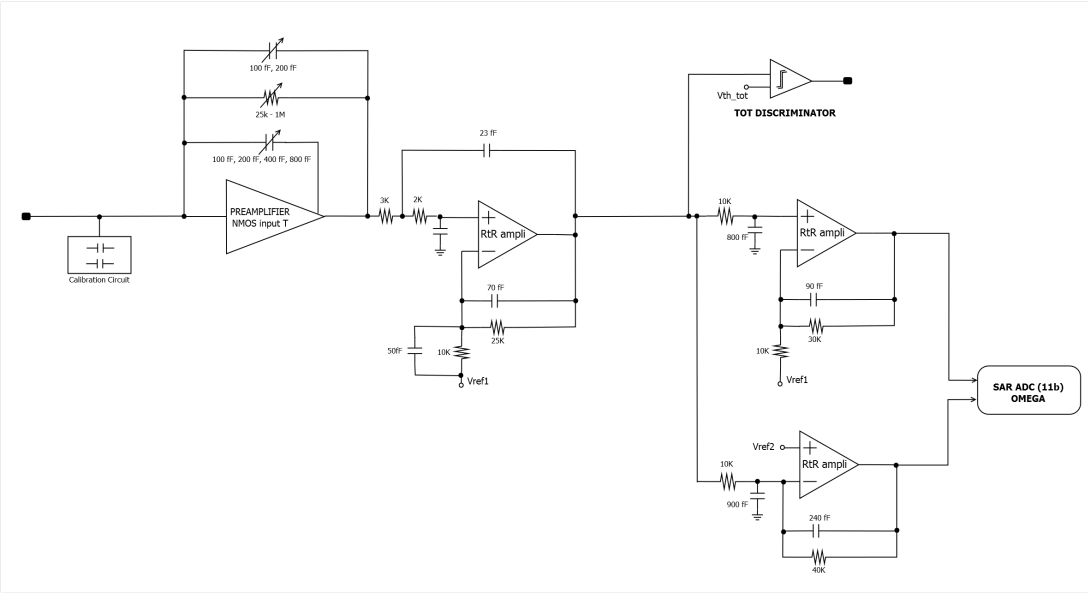


Figure 8.5: Block diagram of the analogue part of a TV2 channel.

The measurements show a good agreement with simulations, in particular the output waveforms shown in Fig. 8.6. The peaking time is measured to be 20 ns and the pulse residual after 25 ns is ≈20%.

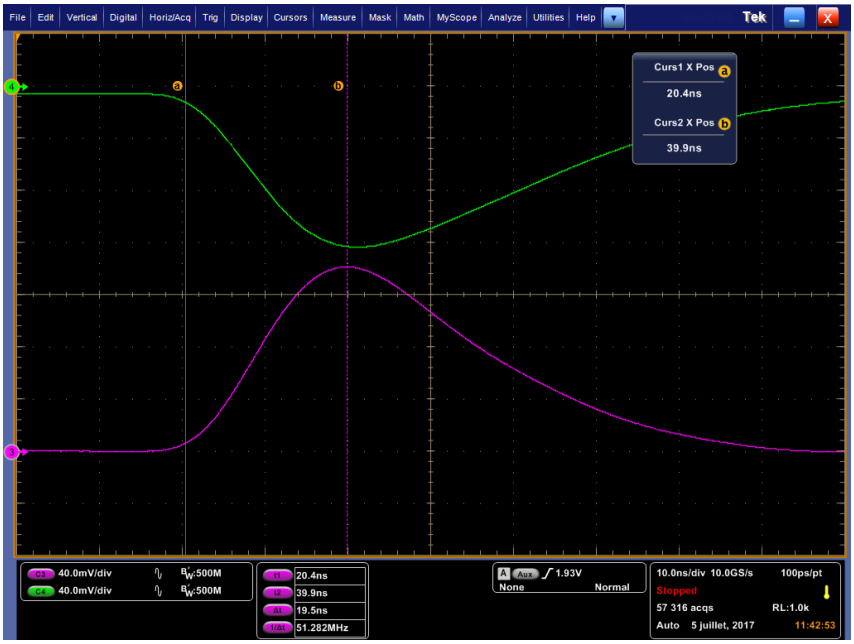


Figure 8.6: Waveform for a TV2 channel.

The output swing covers the full dynamic range of the ADC (±1 V). The preamplifier noise has been measured as a function of the detector capacitance and is shown in Fig. 8.7. The results are in good agreement with the theoretical expectations for the preamplifier, when considering

a noise spectral density (purple squares) with an input voltage noise $V_n = 0.7\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$, and an input current noise $I_n = 1\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$. The parasitic input impedance is extracted as C_{Pa} = 13 pF, where 5-10 pF are attributed to the test board. The noise is higher than previous simulations indicated, due to an error in the design kit, which has now been fixed. To reach the target of 2000 e^- at 60 pF (from 2500 e^-), it can be further reduced by increasing the current in the preamplifier by 20–40 %.

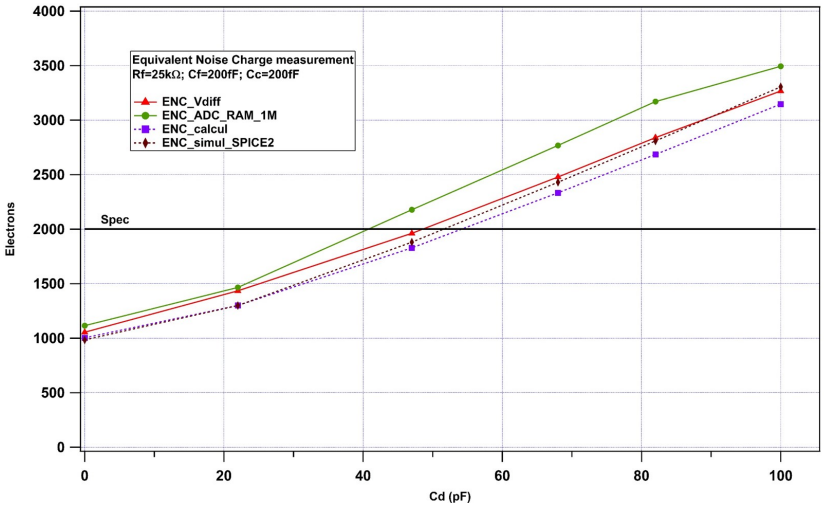


Figure 8.7: Equivalent noise charge as a function of the detector capacitance for the TV2 preamplifier. ENC_Vdiff (red triangles) represents the analog measurement before the ADC, indicative of the performance, that can be compared with an analytic calculation (purple squares) and a SPICE2 simulation (brown diamonds). Finally, ENC_ADC_RAM_1M (green circles) represents the measurement with the internal ADC, writing to RAM.

The 11-bit 40 MHz ADC follows an asynchronous SAR architecture inspired by a design from the Krakow microelectronics group [44]. The DAC is made of a split 6b/4b capacitive divider with a 50 fF unit capacitor in metal-insulation-metal. A dynamic comparator and asynchronous latches provide the 11 bits. Preliminary measurements show a proper operation but with limited linearity and operating frequency as compared to the simulations (10 MHz vs. 20 MHz and 5 LSB DNL compared to 1 LSB, see Fig. 8.8). The limitations in speed have been understood in simulation and have been fixed in HGCROCV1 which should allow 40 MHz operation, whereas the DNL is still under investigation. The ADC designed by the Krakow group demonstrates that a DNL $<$0.5 LSB on ten bits is attained with this technology. The next version of the HGCROC will include channels with an ADC designed by the Krakow group.

The performance of the whole readout chain has been measured. The linear part is measured with the ADC, operating at a lower speed than the final one (10 MHz instead of 40 MHz). A reasonable integral non-linearity ($\pm$2%) is obtained, as shown in Fig. 8.9. However the DNL degrades the RMS noise when there are missing ADC codes.

The ToT response was studied with the discriminator output (as there was no internal TDC). The width increases linearly within a few percent for input charges above 3 pC, but a 10% deviation can be observed between 0.5 pC and 3 pC, which is not seen in simulation (see Fig. 8.10). The non linear part below 0.5 pC, when the preamplifier is not completely saturated, is similar to simulation and will require proper calibration. The jitter is measured to be 200 ps, compared to 50 ps in simulation, which translates into a noise of 6 fC, which is still acceptable for energy measurement in the ToT energy range.

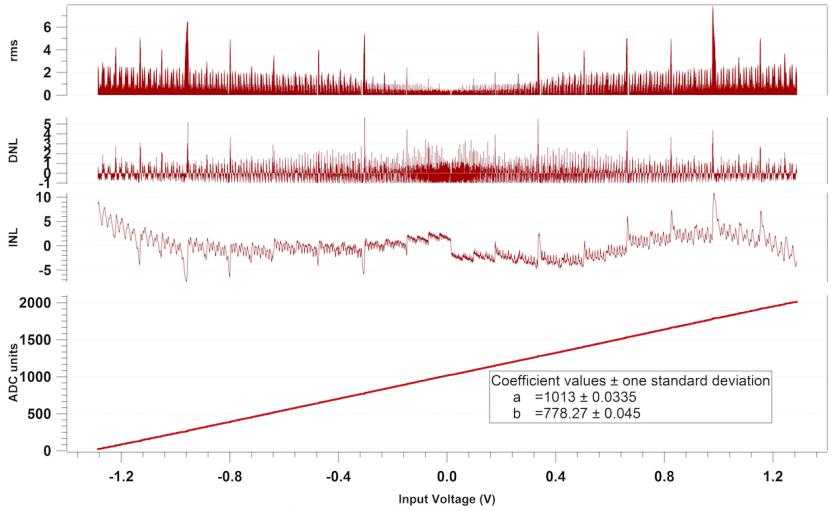


Figure 8.8: Characterisation of the TV2 ADC as a function of the input voltage. From bottom to top: ADC counts, integral non-linearity, differential non-linearity, and measurement RMS.

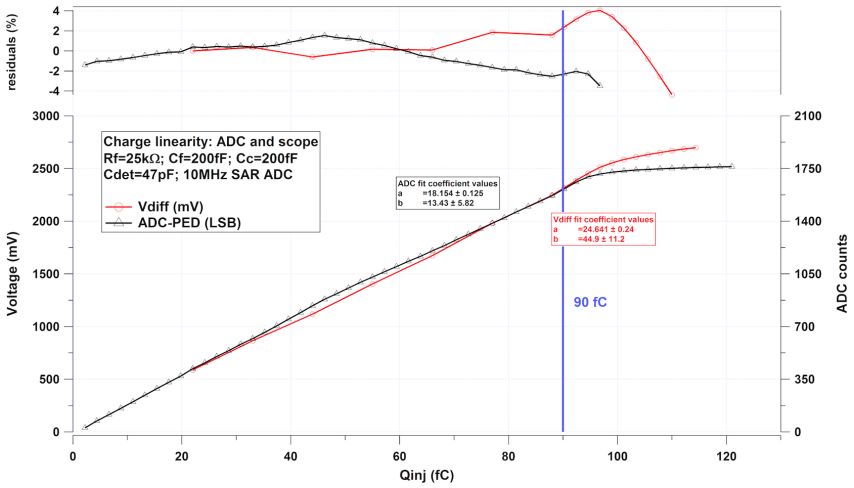


Figure 8.9: Characterisation of the TV2 ADC: ADC counts as a function of injected charge and residuals.

The time of arrival (ToA) in TV2 was also measured. A noise term of 5 ns/Q(fC) is measured, in good agreement with simulations. The time walk is measured to be 10 ns, and can be corrected by the amplitude ADC measurement.

A RAM designed by CERN for high radiation tolerance has been included in order to test its performance, store the data during digitisation, and read them at a lower rate. The power consumption has been measured to be 4 mW/channel at 40 MHz writing, and 1 MHz reading rate, in good agreement with simulation. This value is high and would correspond to 80% of the power allocated to the FE digital part; it will therefore be redesigned. Initial estimates show that a reduction by a factor at least 5 could be achieved.

Measurements have also shown a strong variation of the baseline with temperature (15 mV/K) which almost gets out of range at $-30\,^{\circ}$C. This was also observed in simulation and is due to the DC-coupled shaper, which amplifies the DC output of the preamplifier which varies with V_{GS}. A 6-bit DAC will be added to adjust the operating point channel per channel, in order not to lose dynamic range for the ADC.

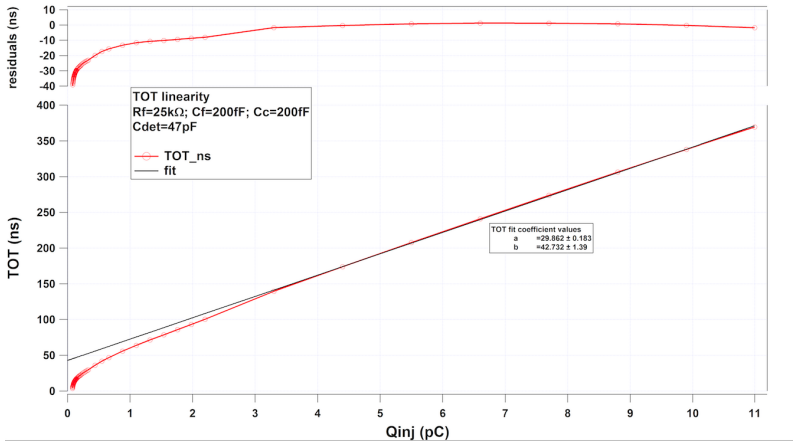


Figure 8.10: Characterisation of the TV2 ToT: ToT time as a function of injected charge and residuals.

The radiation tolerance of the analogue block will be tested, although the technology has been validated up to 400 Mrad [45]. To ensure a good radiation tolerance, whenever possible, minimum size transistors have been avoided and high VT transistors have been used in digital libraries and current mirrors.

8.1.2.4 HGCROC1 design

A first version of the complete HGCROC1 chip was submitted in July 2017. It complements TV2 with the TDCs, trigger formation, and high-speed readout. It comprises all the analogue and conversion parts and most of the critical digital parts, in particular the trigger formation, and high-speed 1.44 Gb/s e-links drivers. There are 32 channels and the die size is $5 \times 7\,\mathrm{mm}^2$. The synoptic view is shown in Fig. 8.11.

The analogue part (Fig. 8.12) is similar to what has been described for TV2, although the preamplifier is now of dual polarity and the ToT uses a current source rather than a resistor to accommodate the two polarities. A 10-bit calibration pulser is also included in order to calibrate the non-linear part of the ToT measurement.

The ADC has been improved to run at 40 MHz. Two capacitive DAC arrays have been implemented in order to compare the matching accuracy obtained with two different layouts. Special attention has been paid to the decoupling of the reference voltages to minimise inter-channel coupling.

The HGCROC1 includes 2 TDCs: one for the ToT measurement with a 50 ps bin-size and one for the ToA measurement with a 24 or 12 ps bin-size.

Above 15 fC, the ToA is expected to have a jitter of 3 ns/Q(fC). This is slightly better than TV2 as the signal is taken from the preamplifier rather than from the shaper. The expected time walk of 7 ns can be corrected by the amplitude measurement from the ADC. For HGCROCDV1, simulations suggest that by shortening the rise-time and slightly increasing the current in the preamplifier, the jitter on the ToA measurement can be improved to around 1.5 ns/Q(fC). The time-walk is correspondingly reduced to around 3 ns.

Two versions of the ToT TDC have been implemented: a classical one based on the HPTDC design [46] which dissipates 3 mW/channel and an innovative low-power one which dissipates less than 0.5 mW/channel.

The ToA TDC uses a novel architecture to reach a 20 ps bin-size with low power (<1 mW/channel).

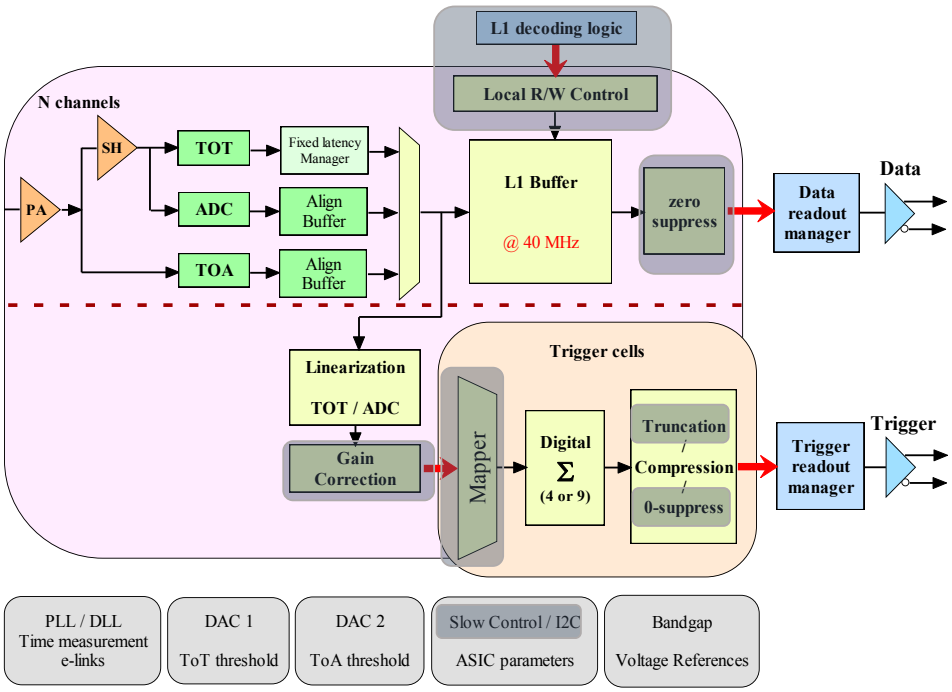


Figure 8.11: Synopsis of the HGCROC1 submitted in July 2017. The dark grey boxes indicate those parts or functions not included in HGCROC1.

It is made of two stages providing a coarse time (100 ps step) and a precise time after an amplification of the residual.

The 11-bit ADC data (or the 12-bit ToT data if the ADC is saturated) are stored in the L1 RAM together with the 10-bit ToA information and 2 flag bits. This results in a 24-bit word which needs to be stored during the 12 μs of L1 latency corresponding to a depth of 512 columns. The RAM is the same as in TV2 (32×512) and is read out with a similar scheme (start acquisition, stop and read memory content). The data are serialised and sent out on e-links at 320 MHz, with 8 differential lines using the CERN Low Power Signaling (CLPS) specification for the 32 channels.

The computation of the trigger sums requires linearization of the ToT TDC data and merging them with the ADC data. This is done with the algorithm depicted in Fig. 8.13. The 12-bit ToT data are converted into charge by a multiplication factor after pedestal subtraction and thresholding to remove the non linear part. It is then combined with the 11-bit ADC data which have been delayed to compensate for the 200 ns delay of the ToT data. The result, which covers a dynamic of 16 bits (from 0.2 fC to 10 pC), is then compressed to 8 bits using a 4-bit exponent 4 bit-mantissa floating point semi-logarithmic algorithm. This operation takes 25 ns and consumes $\approx$200 μW/trigger sum or 50 μW/channel.

Each trigger cell delivers 8 bits at 40 MHz, i.e. 320 Mb/s. Cells are grouped by four and serialised to feed a 1.44 Gb/s e-link CLPS driver.

8.1.2.5 Towards HGCROCDV1

The final version of the chip, HGCROCDV1, will be twice the size of HGCROCV1 (78 channels) and will incorporate additional digital blocks that were not completely defined at the time of

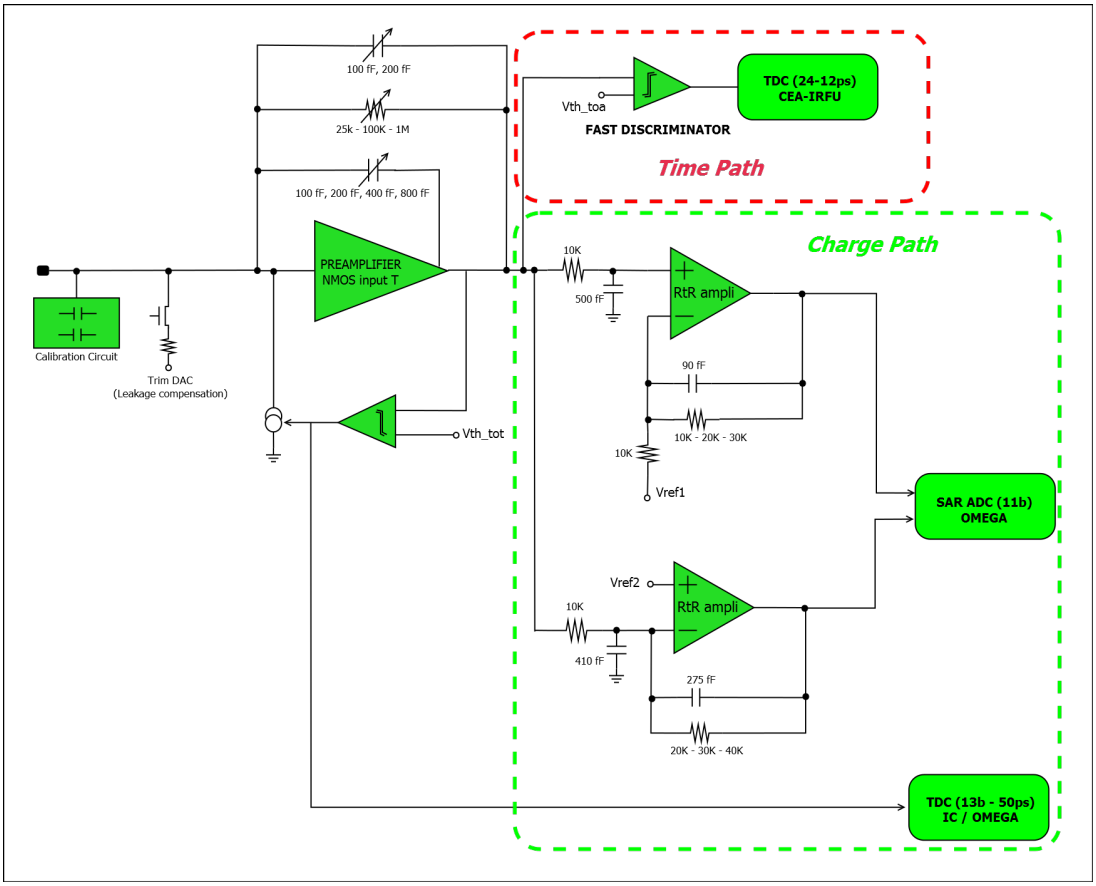
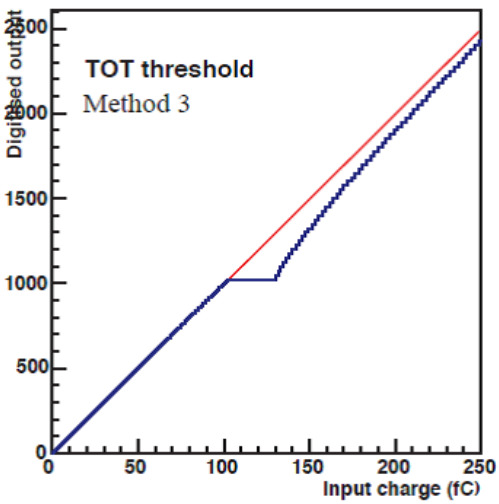


Figure 8.12: Diagram of an analogue channel of the HGCROC1.



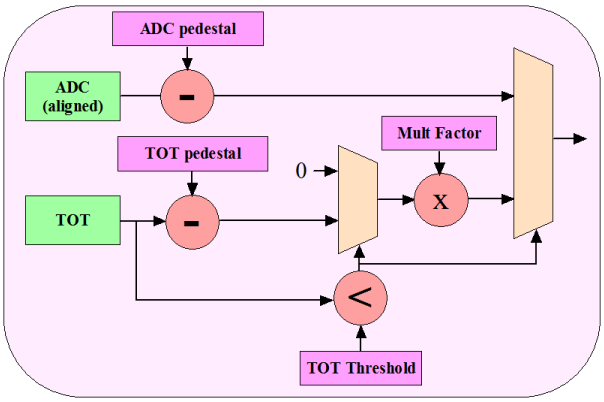


Figure 8.13: Method of linearisation in HGCROCV1 for combining the ADC and ToT measurements, as described in the text.

the submission of HGCROCV1. It will require the following steps (see Fig. 8.14):

- the choice of the TDC for ToT measurement,
- the addition of 4 channels for common mode noise subtraction,
- the replacement of the slow control interface with an I^2C interface compatible with the SCA and the LpGBT,
- the redesign of the RAM (L1 Buffer) to reduce its power consumption. A dynamic memory (DRAM) is now being designed by CERN to bring the power down to less than 1 mW/channel,
- the addition of blocks to decode the fast signals (40 MHz, L1 accept, BCR, Test Pulse) from the 320 MHz input from the LpGBT. These blocks share many commonalities with those developed for the CMS Tracker and it is planned to re-use some existing developments from CERN,
- an additional configuration for the trigger sums, to sum either four or nine sensor cells, depending on the cell size (1 cm^2 or 0.5 cm^2), and
- the addition of a derandomizer memory, in order to average the readout rate at 750 KHz, perform zero-suppression and data formatting.

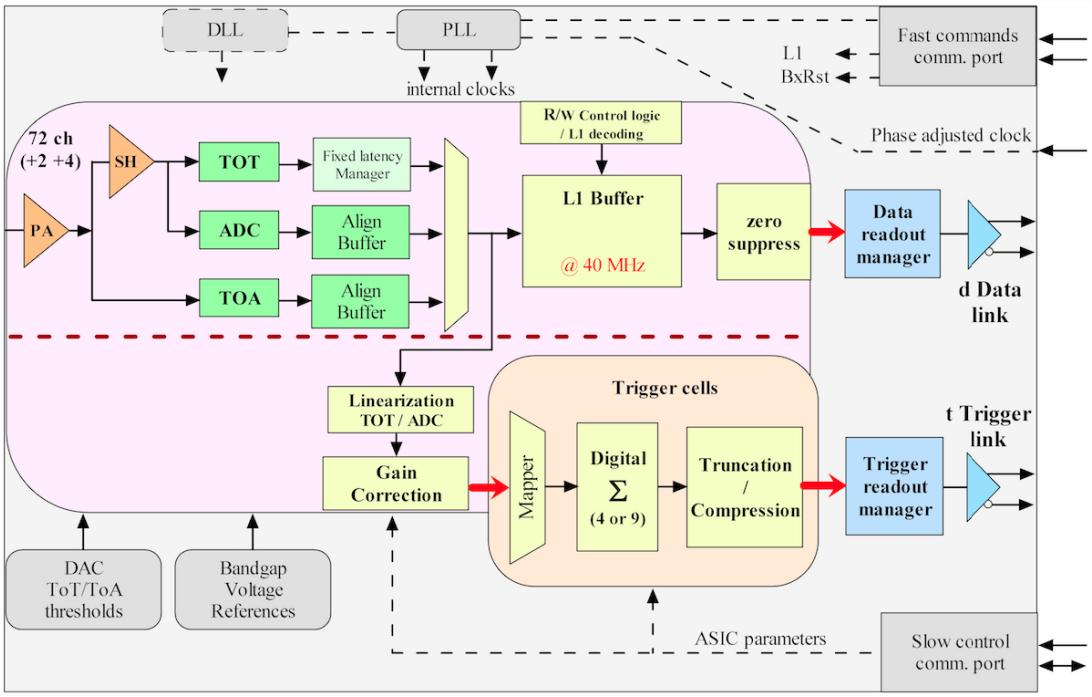


Figure 8.14: Synopsis of the HGCROCDV1 ASIC.

The ASIC should be submitted in June 2018 and received three months later. Extensive testing including irradiation tests, and changes arising, will lead to a further iteration in June 2019. The production run is foreseen for June 2020.

8.1.2.6 HGCROC for SiPM-on-tile

Part of the hadronic end-cap calorimeter is equipped with scintillating tiles coupled to SiPMs. These photosensors can be read out by a variant of the HGCROC as the constraints for electronics (signal to noise, speed) are similar or less severe (lower dynamic range, lower radiation levels, lower occupancy). It is foreseen to adapt the very front-end (preamplifier) for the larger

signal level (pC/MIP instead of fC/MIP) of the SiPM. The baseline is to use a current conveyor, whose schematic diagram is shown in Fig. 8.15 [47].

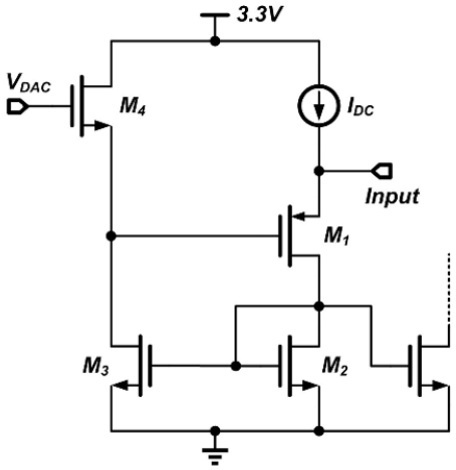


Figure 8.15: Schematic diagram of the current conveyor used by the CALICE Collaboration.

Alternatively, a capacitive divider could be used. Both architectures have been successfully tested by the CALICE collaboration and have shown excellent performance. However, the capacitive divider may not be compatible with the ToT architecture and may require a modified architecture with a dual-gain and two ADCs. In this case, while the power dissipation would remain similar since the ToT TDC would be removed, the digital part of the HGCROC would have to be adapted to handle both data formats. The solution with a current conveyor has therefore a better compatibility with the standard readout, at the price of a slightly worse noise performance. The final choice between the two architectures will need more detailed simulations and tests with irradiated SiPMs.

8.1.3 Concentrator ASIC

The concentrator performs the first level of event building for both the trigger data and for the DAQ data. These functions are functionally and logically independent, but are performed within the same chip.

The concentrator receives BX-synchronous trigger data packets from the HGCROCs, performs some selection on them and builds an output data packet, which is then sent to the BE electronics where complete computation is performed.

In parallel to this, the concentrator also receives DAQ data that is zero-suppressed and therefore not required to arrive synchronously from the HGCROCs. The DAQ information is reorganized and a second level of event building is performed (the first having occurred in the HGCROCs themselves). Input to the concentrator are 1.44 Gb/s serial links using the CLPS standard. The concentrator provides 100 Ω internal termination resistors suitable to terminate the CLPS signals. Output are 10.24 Gb/s links. Electrically the output links are derived from the LpGBT high speed serializer design, but the encoding is limited to a single type of format, namely the FEC5 format. The specifications of the LpGBT output format and electrical characteristics are described in Ref. [16].

Due to the high speed and low power requirements, the concentrator is designed for a 65 nm technology. It shares the same 1.2 V digital supply of the HGCROC.

The concentrator has up to 72 inputs of 1.44 Gb/s and three 10.24 Gb/s output links. Each of the inputs and outputs can be configured by slow control programming to be part of the trigger

or DAQ data event building tree.

8.1.4 Clocking and fast control

The system clock signal and the fast signals necessary for synchronizing the front-end electronics are all provided by the LpGBT chip. Clocks with 40 MHz and 320 MHz frequencies are distributed to all HGCROC chips from the LpGBT.

The 320 MHz clock is used for the fast control and trigger information which is provided at 320 Mb/s. The HL-LHC clock frequency of 40 MHz is derived from a 4-bit fixed repetitive pattern in the data on the fast control line. The remaining 4 bits in the 8-bit fast control data packet are used to transmit fast (synchronous) reset, DAQ trigger, test and calibration pulse trigger, and LHC orbit (DAQ counter) reset signals. All fast signals, input and output, are implemented using a modified LVS standard called CERN Low Power Signaling, compatible with the LpGBT and described in Ref. [16].

The HGCROC decodes the 8 bits of the fast control information and performs the appropriate internal function.

8.1.5 Modules PCB

8.1.5.1 Silicon hexaboards

The silicon module PCB, the hexaboard, consists of a PCB where the readout chips (HGCROC) and any other relevant passive and service components are mounted. The hexaboard PCBs are shaped to overlap exactly the hexagonal shaped sensor.

Connection between the PCB and the cells on the sensor wafer is possible using holes drilled in the PCB and through wires bonded to the sensor and to the PCB. This is illustrated in Fig. 8.16.

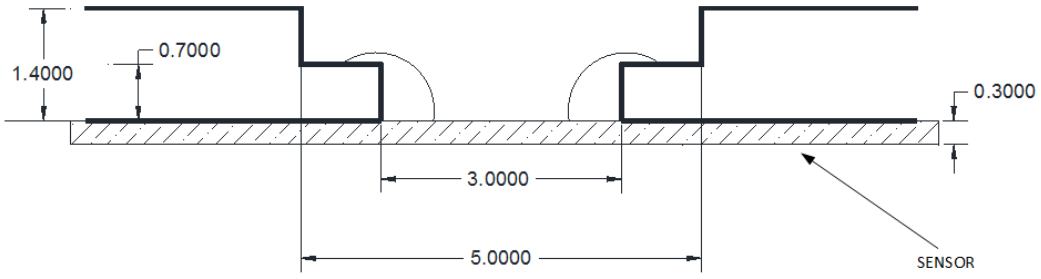


Figure 8.16: Cutaway view of the connection between the hexaboard and the silicon sensor.

There are two main types of hexaboards serving 432 or 192 cells sensors. The boards are built with the same technology, but one of them houses twice as many readout HGCROC chips as the other. Other types of boards are also designed for the cases where geometry requires the use of partial wafers.

The boards consist of a normal 8-layer FR4 printed circuit board stack-up, where (from bottom to top in Fig. 8.18), two layers are dummy and are just used to extend the distance between the surface of the sensor and any electrical conductor as to minimize parasitic capacitance, one ground layer follows and is used for the whole board, one analogue routing layer is used to connect the HGCROC chips to the individual sensor cells, two layers are used to provide analogue and digital power to these chips, and finally the two top layers are used to route all the necessary digital signals. Both boards use a hole to connect the PCB to the sensor with

bond wires. Each hole in the hexaboard is used to connect three adjacent cells using a pad layout shown in Fig. 8.17.

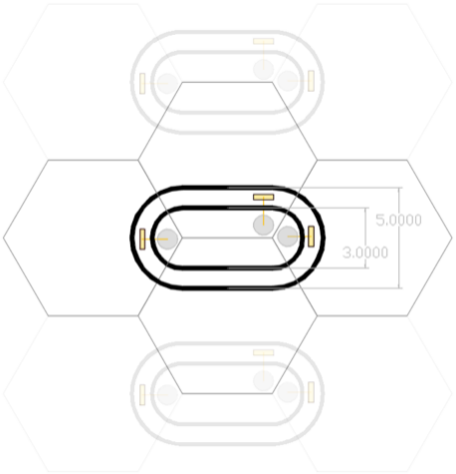


Figure 8.17: Top view of the connection between the hexaboard and the silicon sensor. The connection to individual pads takes place through elongated holes in the hexaboard.

To allow the best possible protection of the sensitive lines from the detector to the HGCROC, the board allows for wire bonding on an intermediate layer as illustrated in Fig. 8.16. This arrangement has been optimized to shield optimally the sensitive traces from the sensor cells to the input pads of the HGCROCs. The two top digital layers contain impedance controlled differential lines for the routing of the critical clock and high speed digital signals. The module PCB has a total thickness of 1.4 mm. Ground and power planes use 35 μm thick copper to provide the lowest possible resistance. In the worst case it is estimated that a module would require about 4 A of current on the analogue and digital supplies.

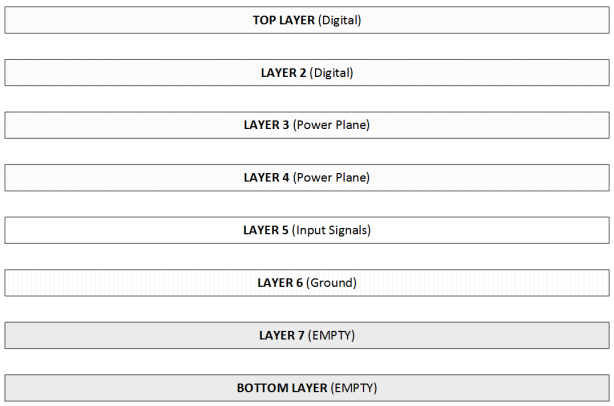


Figure 8.18: Stack of layers used for the hexaboard PCB.

As mentioned above, the hexaboard contains three or six HGCROC chips (for 192 and 432 channels respectively), two linear voltage regulator to generate locally the analogue and digital supplies necessary and some passive components for the filtering of various voltages and also of the high voltage applied to the sensor.

The use of hexagonal sensors to cover an endcap between two conical surfaces requires using some odd size modules shaped for the peripheral regions of the calorimeter. Such partial modules are adapted from the design of the full size module; for instance, an external half module only contains two chips instead of three.

HGCROC chips are mounted on the hexaboard using bump-bonding soldering techniques. With a bump-bond pitch of 400 μm for the HGCROC, routing traces on the two top planes can be attained with relatively low-cost techniques (100 μm pitch, 120 μm through-vias). A preliminary pinout of the HGCROC optimized for routing of signals on the hexabaord has a pinout footprint of approximately 15 mm $\times$ 8 mm.

The linear voltage regulators are packaged instead in low profile quad-flat no-leads (QFN) packages and are mounted on the hexaboard using standard reflow soldering techniques.

A critical dimension is the distance between the front-end hexaboards and the motherboard on top. Currently this distance is set at 1.5 mm. This space has to contain both the chips and any other components mounted on the front-end board (e.g. facing up), and all components mounted on the motherboard (facing down). This clearly means that the placement of all components on these boards has to be arranged in a complementary fashion between the two boards in order to avoid interference in space. This is true for relatively large chips, as well as for all smaller passive components.

To save precious board space, all required termination or pull-up resistors are directly integrated into the chips, and have individual enabling-disabling pins for configuring these resistors.

8.1.5.2 Tileboards

The design of the scintillator/SiPM tileboards is simpler than the silicon sensor motherboards and is described in Section 7.3.4. The detailed design of a tileboard is shown in Fig. 8.19. The HGCROCs, the LVRs, and the LED control circuits are on the back side of the board, accommodated in cutouts in the copper cooling plates. At the periphery of the tileboard, two scintillator tiles are reduced in size, to leave space for an adapter board for the Twinax cables that route the e-links towards the concentrator (located at the sector periphery) and another adapter board for the power cables.

8.1.6 Motherboards

8.1.6.1 Silicon motherboards

As already mentioned, the motherboard for the silicon modules houses the concentrator ASIC, the LpGBT and the optical modules for the LpGBT and for the trigger and DAQ data links, the SCA, and the voltage regulators needed for these components. An example of a motherboard is shown in Fig. 8.20.

The motherboard is a board with a width of $\approx$10 cm that covers up to five silicon modules in length. For DAQ/trigger data and control, each motherboard is an independent unit. The wide connectors at the extremity of the motherboard are only used for the power lines. The motherboard is an 8-layer PCB: the 4 lower layers are mostly used to route the e-links from the HGCROCs to the concentrator, while the 4 upper layers are used as a bus for the power lines: one layer for VDDDA, one for VDDD, and two layers for the common return.

PCBs longer than 50 cm should be avoided, as they are more expensive and also because the copper thickness of long PCBs is usually limited to 1/2 oz (17 μm), which would lead to an unacceptable increase of the power loss in the power layers. As a consequence, 4- and 5-modules long motherboards will be split into two sub-units linked by connectors.

The motherboard is cooled by contact with the hexaboards, which are themselves in good thermal contact with the copper cooling base plate. A thin gap-pad fills the space between the

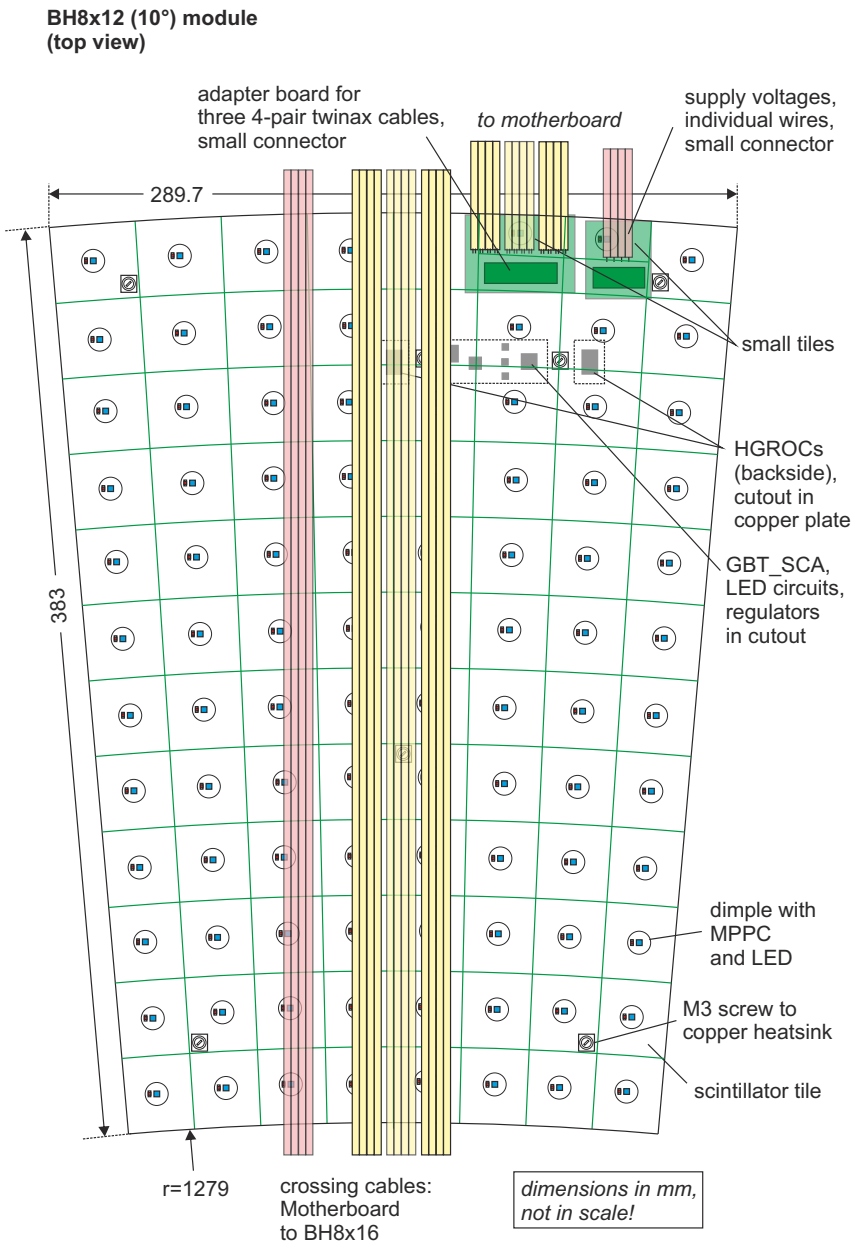


Figure 8.19: Diagram of an outer scintillator/SiPM tileboard showing the arrangement of cables from the inner tileboard crossing this tileboard vertically enroute to the motherboard and the adapter boards which connect the Twinax and voltage supply cables to this tileboard.

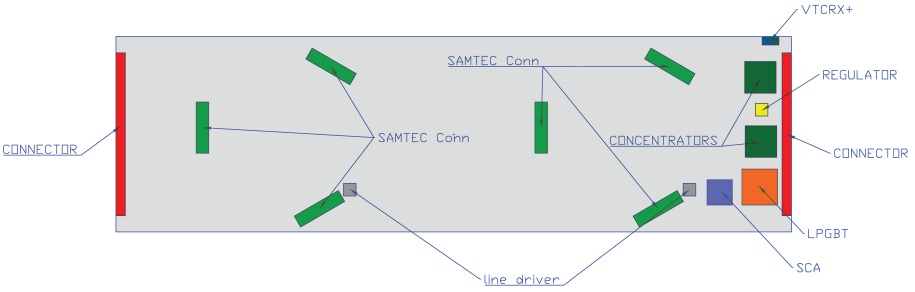


Figure 8.20: Example of a 2-module-long silicon motherboard.

motherboard and the hexaboard, in order to ensure good thermal contact between the two boards.

8.1.6.2 Scintillator motherboard

The motherboard for the CE-H scintillator section performs an analagous function to the silicon one (Fig. 3.3). Each 30° scintillator section, containing up to twelve component tileboards, is serviced by a single motherboard as shown in Fig. 8.21.

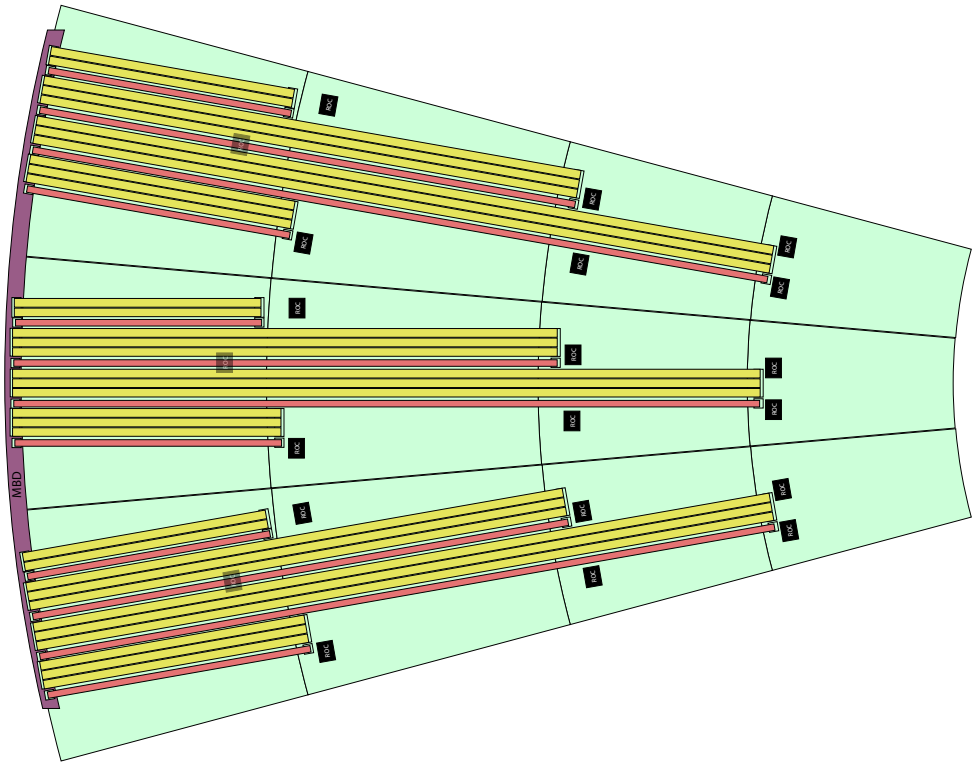


Figure 8.21: Overview of the individual scintillator tileboards connecting to the motherboard located at the outer edge of the cassette.

The motherboard hosts the concentrator chip, which merges up to 72 electrical links of 1.44 Gb/s from multiple HGCROCs on the tileboards and serializes the data onto multiple outputs at 10 Gb/s. It also contains the optoelectronics that convert the output data to optical for transmission to the off-detector electronics. One LpGBT is used in the transceiver mode both to send the DAQ data out and to receive the clock and control signals. Control signals are routed from the LpGBT to a GBT-SCA slow control adapter chip on the motherboard, which is responsible for distributing the slow control signals for the motherboard itself. Control signals for the individual tileboards in the section are distributed by a control multiplexer. A linear regulator is used to control power voltages on the motherboard. The motherboard also hosts the electronics used to generate the SiPM bias voltage. A design similar to the SiPM control board for the Phase-1 HCAL upgrade will be used; individual SiPM bias voltages are then trimmed on the tileboards.

Each tileboard is connected to the motherboard via Twinax ribbons which contain up to twelve pairs for differential signals. To land the Twinax cables on the motherboard we use a robust stacking connector which facilitates multiple reconnections of the Twinax cables during cassette construction and commissioning. A small passive PCB is used to connect the Twinax cables with the connectors. Since the connectors are relatively long, we require most of the outer

periphery of the cassette to land all the connectors. In order to avoid motherboards that are too long, this is accomplished by using passive "wing" boards that connect to the motherboard and route the signals along the periphery of the cassette as shown in Fig. 8.22. In order to allow the motherboard assemblies to fit into the active area of the calorimeter, a shallow groove will be milled out of the outer edge of the cooling plate.

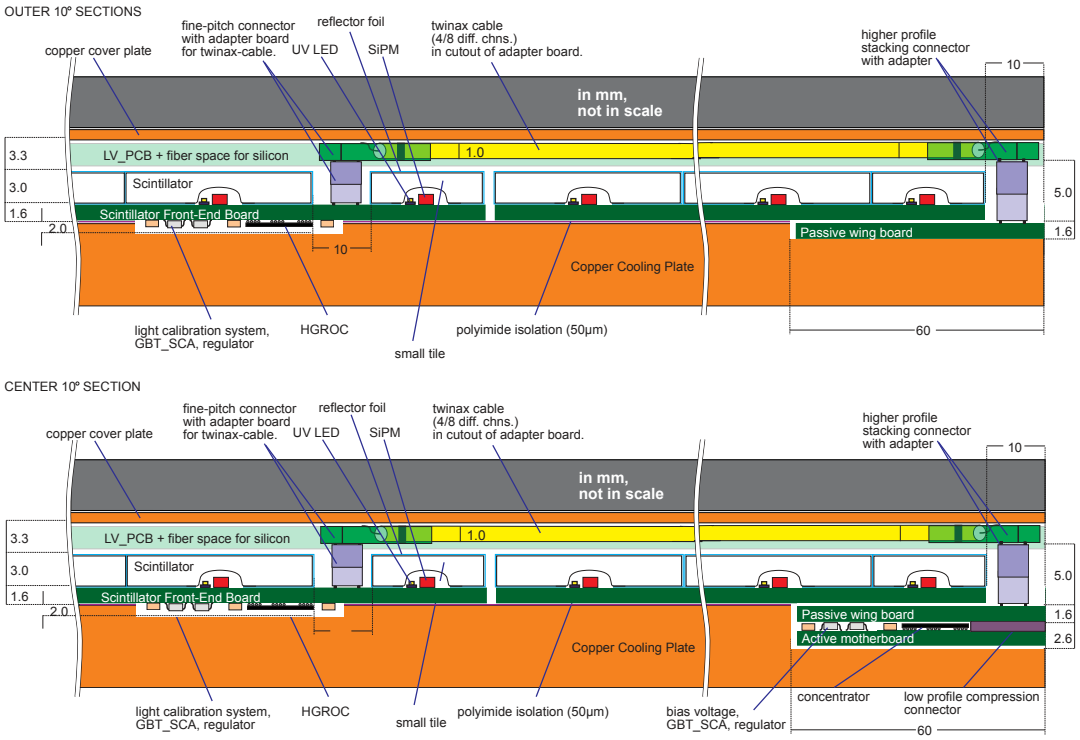


Figure 8.22: Cross-sectional views of the cassette showing the relative locations of the tileboards, the scintillator motherboard, and the cabling and wing PCBs which connect them.

8.1.7 Data formats and rates

The main data types produced by the FE electronics are for the DAQ, i.e. the event data associated with a level-1 trigger accept (L1A), and for the TPG, i.e. trigger raw data which are used as input to the TPG. The LpGBT links provide the interface between the FE and BE electronics systems and the data on these links are described in this section. In all cases, the rates presented assume an average of 200 pileup events per BX and, for the DAQ case, a L1A rate of 750 kHz. For details of the simulation used, see Section 10.3.1.2.

8.1.7.1 DAQ data format and transfer

As described in section 8.1.2.4, the DAQ data are buffered within the HGCROC ASIC for the 12.5 μs latency of the L1A. The data associated with the L1A are then zero-suppressed and formatted for transfer to the concentrator ASIC, which forwards the data on the LpGBT links to the off-detector DAQ system.

Figure 3.6 shows the average occupancy arising from an average of 200 pileup interactions per bunch crossing as a function of layer and radius in the x, y plane of the layer. The occupancy is defined to be the fractional rate at which the channel has a energy deposit above the nominal 0.5 MIP threshold for DAQ readout. It is seen that the occupancies of the sensor channels varies by about two orders of magnitude across the HGCAL. This means that it is very difficult to have

a data format that is efficient over the whole calorimeter. To account for this, the HGCROC ASIC will be able to format the data in two different modes, described later in this section.

The data associated with an L1A include not only the ADC, ToA, and ToT values (when above their respective thresholds) for the BX that caused the trigger, but also the ADC values from the previous bunch crossing. Due to the shaping of the FE preamplifier, large deposits in the previous BX can have a significant influence on the ADC readings in the L1A BX. As such, the ADC values from the previous BX are only read out for channels above threshold in the L1A BX. For this, a threshold for previous BX ADC values of 2.5 MIP is used.

Figure 8.23 shows the average number of ADC, ADC in previous BX, ToA, and ToT values above their nominal thresholds. The relative fractions of the four data types is approximately constant, with the ADC from the previous BX rate being around 0.3 of the in-time ADC rate, while the ToA is read out with a rate around 0.1 and the ToT rate is around 0.001. This implies that the DAQ data volume is dominated by the two ADC values.

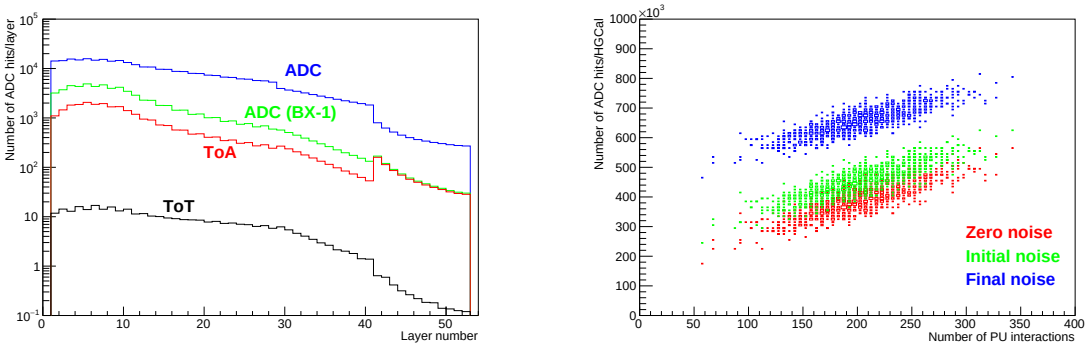


Figure 8.23: Left: Average numbers of ADC, ADC in previous BX, ToA and ToT hits per layer, as a function of the layer number. The thresholds used in this plot are 0.5 MIP, 2.5 MIP, 10 fC and 150 fC, respectively, for the four histograms. Right: The number of ADC hits as a function of the number of pileup interactions in the BX.

Figure 8.23 also shows the number of ADC values above threshold as a function of the number of pileup events used in the simulation of the BX for various levels of the channel noise. This shows that at the noise levels expected after 3000 fb^{-1}, both the noise and the amount of pileup will have a significant effect on the DAQ readout rates.

The DAQ data format must account for the ADC values for all channels above threshold and potentially the ToA and ToT values, as well as the previous BX or presample ADC values. The data from each channel above threshold will be transmitted on the links in the format shown in Fig. 8.24, which explicitly assumes that the ToA readout threshold is always lower than the ToT threshold, such that the former is always present when the latter is read out. This format is foreseen to be generated in the HGCROC ASIC and hence used for the transfer to the concentrator ASIC also.

The data associated with each HGCROC is a sequence of single channel packets (as above) and, as stated earlier, will be in one of two formats; for the channels above threshold, this includes either a list of channel numbers or a bit map of the channels. The former is optimal for low occupancies while the latter is better for high-occupancy regions. Hence, these data will be able to be configured for either format, depending on the occupancy at its position. One bit in the header will indicate the format so that the packet can be unpacked without any further information. The assumed formats of these packets of HGCROC data for both cases are shown in Fig. 8.25. There are 78 channels per HGCROC ASIC, of which four are common

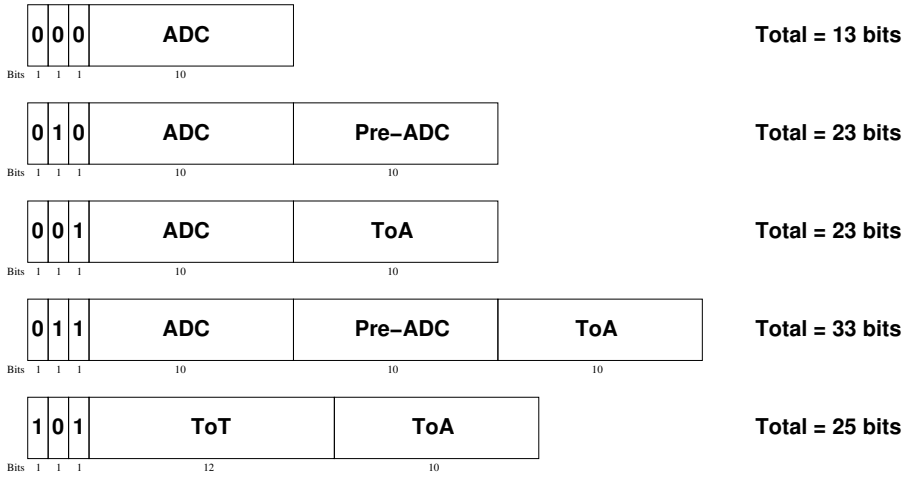


Figure 8.24: Format for the DAQ data from a single channel. All five expected variants are shown. The first bit indicates whether the ADC (0) or ToT (1) value is present. The second bit indicates if the previous ADC (Pre-ADC) value is present (1) or not (0), and the third bit shows the same for the ToA values.

mode channels and their ADC values are always read out. For the other 74, both the number of channels N_C and the channel numbers used in the list format will be represented as 7-bit values. Hence, the numbering for the list format takes $7(1 + N_C)$ bits, compared with the 74 bits for the bit map format, and so the list format is shorter for $N_C \leq 9$; this corresponds to a fractional occupancy of $\leq$12%.

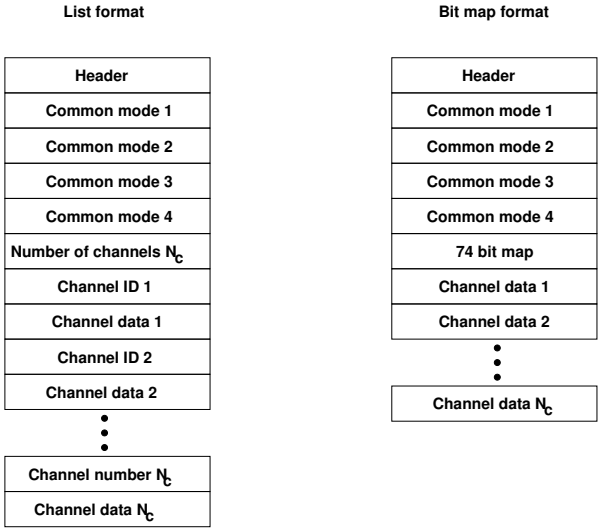


Figure 8.25: Formats on the links for the DAQ data from an HGCROC ASIC. The format on the left contains an explicit list of the channel numbers and would be suitable for use in low occupancy regions. The format on the right contains a bit map to indicate the channels present and would be used in high occupancy regions.

All the HGCAL LpGBT links to the off-detector electronics will be used in the nominal 10 Gb/s FEC5 mode, in which the link physically transmits 256 bits per BX. Only 224 bits, i.e. 8.96 Gb/s,

of these data are available to the user, with the rest being used for clock, error correction, and a slow controls channel. The user data need to include fields for synchronisation and other overheads. In particular, fast control signals such as resynch and L1T throttle commands must have a dedicated bandwidth allocation or they would be delayed by tens of BXs if an event is being transmitted. The actual bandwidth available for event data is assumed to be 192 bits per BX, i.e. 7.68 Gb/s. With a L1A rate of 750 kHz, which gives an average time equivalent to 53 BX to read out each event, then a link can nominally handle average event data packets up to 10 kbit.

Given the occupancies and data formats described above, the DAQ data volumes for each of the motherboards can be estimated. These volumes take into account the occasional higher occupancies in signal events, as well as enough bandwidth to reduce the rate of full buffers in the HGCROC ASIC to an acceptably low probability. The resulting volumes are presented in Fig. 8.26. The figure shows that the expected data sizes from every motherboard do not exceed the 7.68 Gb/s capacity of the single link mounted on each motherboard.

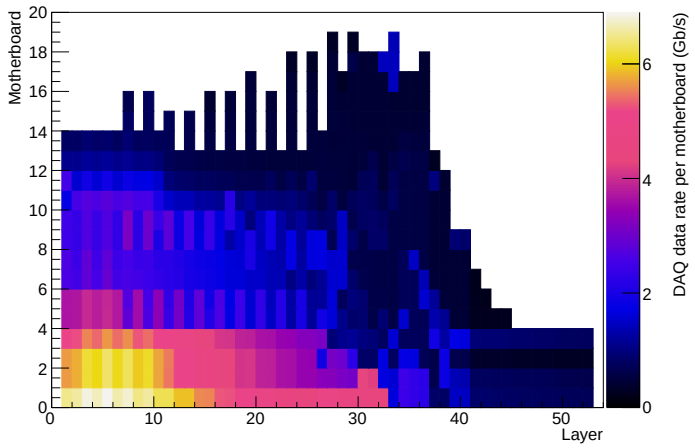


Figure 8.26: DAQ data volumes (in Gb/s) on the LpGBT links required for each silicon readout motherboard as a function of layer number (x axis) and motherboard number (y axis). The motherboard numbering is approximately ordered by motherboard average radius within each layer. Only motherboards in one 60 ° sector are shown because of the six-fold symmetry of the motherboard layout.

8.1.7.2 TPG data format and transfer

The trigger cells and coarse sum data are formed on-detector. Trigger cells with energy above a threshold of 2 transverse MIPs ($\mathrm{MIP_T}$) are selected. This threshold corresponds to approximately 5 MIPs at the outer edge of the calorimeter and 20 MIPs at the inner edge. Using a threshold in terms of $\mathrm{MIP_T}$ rather than MIP, results in a selection which is directly related to the contribution of the trigger cell to the transverse energy used in the trigger. The resulting occupancies of trigger cells above threshold per motherboard and layer are shown in Fig. 3.7 for an average pileup of 200 interactions per bunch crossing.

The concentrator ASIC formats the trigger data from the HGCROC ASICs for transmission on the links. The transverse energy value of each trigger cell is encoded as an 8-bit floating point value, with a 4-bit exponent and a 4-bit mantissa. This format can store values over a range corresponding to 19 bits, which is sufficient for the largest possible sum of 9 sensor cells. The concentrator ASIC forms a sum over all the trigger cells from each HGCROC ASIC, and these

form the raw data used for the energy maps in the TPG. These sums contain up to 72 sensor cells and so need a wider range of 22 bits; these are represented by an 11-bit floating point value, with a 4-bit exponent and a 7-bit mantissa. The assumed format for the raw trigger data corresponding to each HGCROC ASIC is shown in Fig. 8.27. The data size from each HGCROC will then be $16 + 12N_T$ bits for N_T trigger cells above threshold. For simplicity, the same format is used for both the readout of the large-cell and small-cell Si sensor, for which the total number of trigger cells per HGCROC is 16 and 8, respectively, even though for the latter only four and three bits would be needed for number of trigger cells and the trigger cell number, respectively. There are no trigger cell data from the previous bunch crossing as a correction for the previous BX is made in the FE concentrator ASIC.

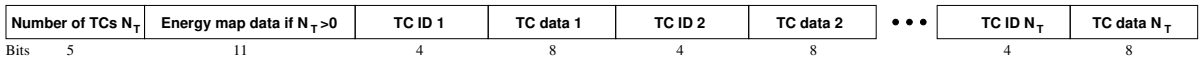


Figure 8.27: Format of the trigger raw data on the LpGBT links corresponding to the data from one HGCROC ASIC.

The total packet on the link per BX will be a header word giving the number of following HGCROC packets, and then the sequence of HGCROC packets as detailed above. As for the DAQ data packet, the trigger raw data can be unpacked without requiring further information.

As for the DAQ case, the links will require bandwidth for synchronisation and other overheads, but for the TPG case, there is no need to allocate space for fast control signals. It is assumed that 208 bits per BX, i.e. 8.32 Gb/s, is available for the trigger raw data. Using the format specified above, the required data rate per BX for the trigger cell and coarse sum data combined is shown for each motherboard in Fig. 8.28. These values take into account two factors that lead to fluctuations in the amount of data: the extra amount of locally-produced data in physics signal events; the additional bandwidth needed to reduce the impact of larger-than-average events on the probability of trigger data being truncated in the concentrator ASIC buffer. Figure 8.28 also shows the corresponding required number of LpGBT links. These data will require a total of approximately 7500 links. The distribution of links per layer is shown in Fig. 8.29.

8.1.8 Powering scheme

Powering of the front-end electronics is structured in a hierarchy consisting of:

- External power supplies.
- DC-DC converters.
- On-board linear regulators.

8.1.8.1 External power supplies

External power supplies are specified to provide a 12 V input to a stage of DC-DC conversion as described below. The power supplies are located in a relatively low magnetic field intensity and low radiation region outside the magnetic field return yoke. These external converters should supply 22 A of current at a maximum of 14 V to allow for some voltage drop before reaching the DC-DC converter, for a total of about 300 W.

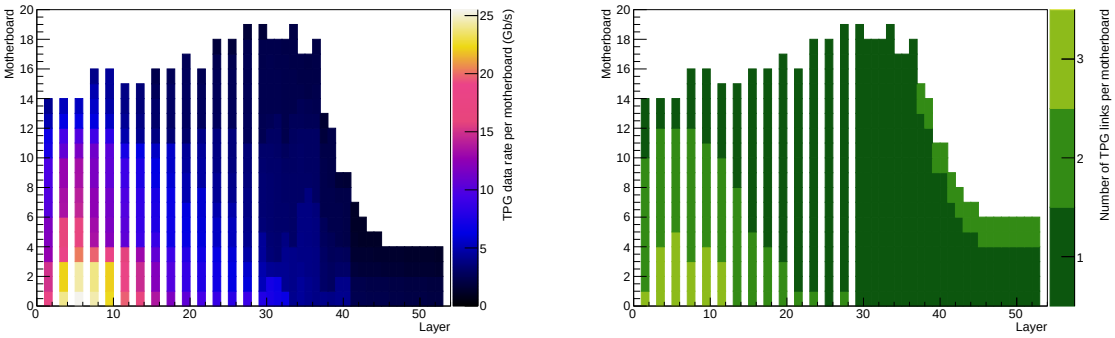


Figure 8.28: Left: Trigger raw data rates (in Gb/s) on the LpGBT links required for each silicon readout motherboard as a function of layer number (x axis) and motherboard number (y axis). The motherboard numbering is approximately ordered by motherboard average radius within each layer, and only motherboards in one 60° sector are shown because of the six-fold symmetry of the motherboard layout. Right: Number of required trigger LpGBT links per motherboard as a function of layer and motherboard number. The two-link area to the right is due to the scintillator region, which has a fixed number of two trigger links for every DAQ link.

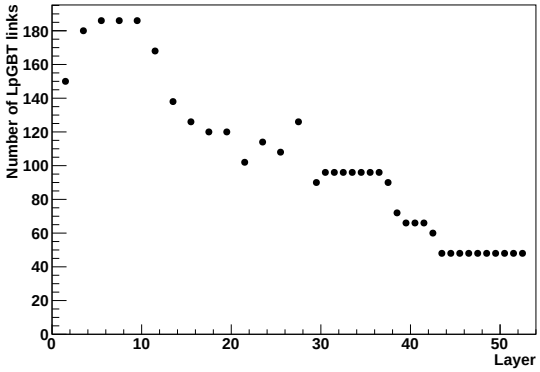


Figure 8.29: Number of LpGBT links per layer required from the FE electronics for the trigger raw data.

8.1.8.2 Embedded DC-DC converters

A voltage conversion from 12 V to about 1.5–2 V is provided by DC-DC converters. These converters are radiation and magnetic-field tolerant, can operate in the innermost CMS regions, and are capable of delivering up to 20 A of current. These DC-DC converters are normally always on and have no control, except for the internal safety circuitry (protection for overtemperature, overvoltage, and overcurrent). The output regulation is specified to be relatively loose as all the electronics is powered via linear regulators further downstream.

The converters are specified to withstand the full CMS magnetic field of 4 T. They will be either located in a position of relatively low radiation at the periphery of the calorimeter or on its back. In the first case, they must be robust to radiation levels of up to 10 Mrad TID and 1×10^{14} n_{eq}/cm^2. In the second case, which is the baseline option, they must withstand only 1 krad TID and 5×10^{11} n_{eq}/cm^2. Such converters can be designed as a multi-phase converter derived from the FEAST converter that was designed for the CMS tracker and other applications.

However we are also investigating a scheme using commercially available converters that are expected to function in the low radiation region located at the back of the calorimeter. These converters would nevertheless need to be radiation qualified to those levels.

8.1.8.3 On-board linear regulators

The HGCROC, the concentrator, and all other auxiliary chips present on the hexaboard and motherboards are powered through low-dropout linear regulators that provide 1.2 V for the digital section of chips and 1.5 V for the analogue section.

These low-drop regulators are necessary to power the very sensitive analogue front-end and to separate as much as possible the noisy digital circuitry from their supply. They are powered from the embedded DC-DC converters with a voltage about 150–200 mV higher than their output voltage. These regulators are capable of providing 4 A max and specified for a maximum power dissipation of 2 W. The regulators can be configured to be active at power on, or to be off and then be activated via a separate digital control pin. The regulators are specified to withstand the same maximum radiation dose of the worst case front-end chips in the most exposed modules. These regulators are currently being specified and subcontracted to an external chip design firm specialized in designing such products. The firm will work in tight collaboration with CMS to address the issues related to improving the radiation robustness of the components.

8.1.9 Monitoring and services

All non safety-critical monitoring and safety functions necessary to guarantee the smooth operation of the electronic system are performed using the capabilities of a custom chip developed for the CMS tracker for the same purposes in a CMOS 130 nm technology and called SCA (slow control ASIC). Each SCA used in the system connects to the experiment slow control system through one of the LpGBT e-links. The SCA is located on the motherboard, and being part of the slow control system it is meant to be powered before the module board electronics and actually controls the turning on and off of the voltage regulators on these boards.

The SCA also has the ability to download the configuration data to the front-end HGCROC chip using I2C buses. Each SCA drives up to 16 I2C buses and can control up to 32 parallel digital lines, each programmable as input or output. Such buses are connected to the different HGCROCs on the hexaboards and are used for programming all the various configuration and mask registers in each front-end HGCROC.

A number of standard I2C bus transactions are supported by the SCA to optimize read and write operations from different I2C slaves. The I2C bus transactions operate at a maximum clock speed of 1 MHz; slower speeds are also possible.

The SCA is capable of making analogue measurements of the voltages, currents, and temperatures through a 12-bit Wilkinson type ADC that can be connected to one of 31 analogue inputs providing a voltage signal in the 0–1 V range. Inputs of the SCA can also be configured to source a current of 100 μA that can then be used to measure resistive inputs (for instance PT-1000 sensors or humidity sensors).

Such inputs are used to monitor:

- the local voltages such as analogue and digital supplies (VDDA and VDDD)
- the local temperatures as sensed on Pt1000 temperature sensors (hexaboard temperature and sensor temperature)

The SCA is capable of driving and sensing a number of digital CMOS lines (in the 0–1.0 V range) that are used to provide on-off signals to the local linear voltage regulators. Such a mechanism can be used to power the hexaboards individually and create smooth turn-on and turn-off powering sequences for the entire front-end electronics.

The SCA based monitoring system is not designed to perform time-critical operations or to be part of the safety system, but can perform elementary operations such as reading a voltage, sensing a temperature, or turning off a regulator in a time of about 1 ms.

The SCA is mounted on the motherboards and controlled through one of the e-links available on the LpGBT. Signals to and from the HGCROC chips are going through the connectors between the hexaboards and the motherboard. Detailed documentation on the SCA can be found in Ref. [17].

8.1.10 Summary of remaining R&D

8.1.10.1 Modelling

While simple written specifications are a common method to document system parameters and capabilities, the complexity and the possible ambiguity of many specifications parameters can more effectively be simplified and avoided by having, whenever possible, executable specifications. Languages and frameworks have been developed for this purpose and are available to all electronics engineers to help them in making sure that:

- the given component interfaces (i.e. receives signals and commands and sends signals and results) correctly to the other system components,
- it correctly performs its functions, as set out in the written specification,
- it can survive after exposure in the LHC environment without unacceptable loss of performance, and
- it can recover from mild and serious error conditions without catastrophic consequences.

The correctness and consistency of the specifications of a given component or subsystem can be exercised using a description at a fairly high level of abstraction, while the electrical and timing correctness of the same component will probably use a more detailed, register-transfer level (RTL) model using the same language.

For the CMS HGCAL project, the language chosen to describe the behaviour of all electronics system is Verilog. It is therefore demanded that all electronics components and, when applicable, boards shall be modelled in Verilog before a physical implementation of them is attempted.

8.1.10.2 Validation of models

Models of components have to be validated within a reasonably realistic simulation environment. For instance, front-end chips will be subjected to stimuli corresponding to what is expected to be found in the most heavily occupied system modules. Chips and logic blocks used for aggregation and filtering of data will be exercised with data patterns corresponding to realistic conditions in the most heavily occupied parts of the detector.

8.1.10.3 Validation of design

Design reviews have to be conducted extensively throughout the entire span of execution of a given design. For the HGCAL front-end electronics chip set at least three validation milestones are foreseen:

- a preliminary design review (PDR),
- a submission readiness design review (SRDR), and
- a production readiness design review (PRDR).

8.1.10.4 PDR

The PDR shall be scheduled when a first version of a commonly-agreed specification is frozen and shall be based upon the availability of at least a behavioural model of the component in Verilog. At this verification stage, schematics and preliminary simulation results of critical blocks (analogue and digital) are all available for the reviewers as proofs of concepts.

8.1.10.5 SRDR

At this stage, a full physical implementation of the component shall be available and the review is aimed at proving that this first implementation of the component will actually work as expected, prior to submission. For the SRDR, complete RTL models, schematics, and layouts are available to reviewers. All process, temperature, voltage, and device-degraded (by radiation) corner cases have been exercised and simulation results are available and have been exercised for all of them.

8.1.10.6 PRDR

At this stage, statistically significant samples of components have been made available in the final physical form and have been exercised for some time in the field (that could include using some ad-hoc test boards, conventional testers, and finally small final-system emulators using final boards and modules) and have passed the necessary quality and acceptance tests.

Radiation tests (TID and SEU) will have been performed to prove the compatibility with the expected environment in the experiment. These are especially time consuming activities and enough time shall be allocated to perform them correctly.

8.2 Data acquisition system

8.2.1 Overview

The HGCAL off-detector electronics consists of the data acquisition (DAQ, this section), the trigger primitive generator (TPG, Section 8.3), and the detector control (DCS) and safety (DSS) systems (Section 3.5). These are described separately as they are effectively independent systems, with the only areas of overlap being that the TPG and DCS use the DAQ hardware to access and configure the on-detector front end (FE) electronics.

In particular, the DAQ and TPG systems are implemented as two independent systems off-detector and there are separate links to carry the data from the FE electronics to the two systems. However, they share a lot of common features and interfaces, and the intention is that the core of the electronics for the two systems will be a common "generic" family of boards, which will be used to implement both systems. This section describes these generic boards and how they are used in the DAQ system. Their use in the TPG system is described in the next section.

8.2.2 DAQ context and interfaces

The DAQ system is made up of multiple electronics boards, each independently implementing the required interfaces to other CMS systems. Hence, the board interfaces are effectively identical to those of the DAQ system as a whole. The context of the DAQ boards and their major

interfaces are shown in Fig. 8.30. The central DAQ system has just released a Phase-2 Interim Technical Design Report [20] that describes some of the following in more detail. However, since the DAQ Phase-2 TDR will not be written for several years, these interfaces are not yet precisely defined. Hence, the details of the generic boards will have to evolve as the DAQ design becomes more concrete.

The main task of the DAQ system is to receive the DAQ data from the on-detector electronics and forward this to the central DAQ system. The links from on-detector to the DAQ boards carry these data, while a corresponding set of links from the DAQ boards to the on-detector electronics allow clock and control. The DAQ system also needs to receive central timing signals via the experiment Timing and Control Distribution System (TCDS) and DAQ control signals from the central DAQ system. Finally, it must interface to the DCS and configuration systems.

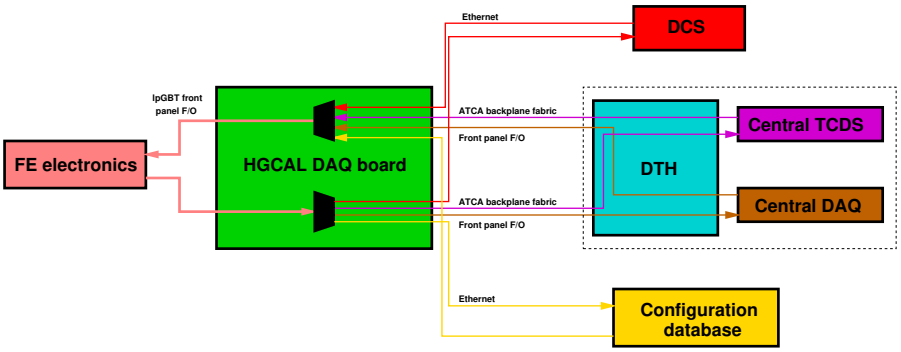


Figure 8.30: Main interfaces to the HGCAL DAQ board.

The central DAQ interface to each detector subsystem (such as the HGCAL) is through a common DAQ and Timing Hub (DTH) board, which will be implemented in ATCA [22] format. This nomenclature is used as the DTH will be an ATCA "hub" board (or "blade") and hence will be installed in one of the two central slots of the ATCA crate (or "shelf"). The DTH provides subsystem access to the central DAQ and TCDS. The DTH also receives the DAQ event data from the subsystem for transport to the central DAQ. Furthermore, the DTH allows the subsystem to send fast control commands to the TCDS, such as throttling the L1A rate if the subsystem buffers become full. While the DAQ control and event data readout are via fibre optic (FO) links on the DTH front panel, the TCDS interface is through the ATCA backplane fabric. This implies that the format for the subsystem boards must also be ATCA and this is indeed the case for the generic board family that will be used by the HGCAL. The subsystem boards act as ATCA "node" boards and occupy the twelve slots outside of the central two slots of the crate.

In terms of power, ATCA allows up to 400 W of power per board, allowing the use of advanced FPGAs and high link counts. The latest FPGAs have sufficient transceivers to handle large numbers of FO links and so can implement high data rates into and out of each HGCAL DAQ board. This document assumes 96 LpGBT links [16] in both directions per board for data transfer from the FE electronics. The total average event size for the HGCAL will be around 2.5 MBytes at an average pileup of 200. Given the number of links, the average data rate on each link at 750 kHz will be around 1.9 Gb/s. This means the links are running on average at around 25% capacity; this is constrained by the geometry of the on-detector electronics motherboards. Each HGCAL DAQ board will therefore receive an average data rate of around 180 Gb/s. This sets the requirement for the throughput of the event data rate into and out of the DTHs.

The HGCAL DAQ boards use the links to the FE electronics to configure and control the on-

detector systems. Hence the HGCAL DAQ must have an interface to the central configuration database to access the configuration data and provide readback checks on these data. This will be implemented as a standard Ethernet port on each board, allowing general software access to all data stored on the DAQ board. The DAQ boards also provide the path for reading out the environmental status information of the FE electronics, such as voltages, currents, temperatures, etc., via the slow controls path of the LpGBT protocol. These data are collected on the DAQ board and sent to the DCS system also via the Ethernet port.

Finally, the HGCAL DAQ boards will also process the event data to provide calibration and monitoring data. The results of this processing will be gathered either by adding the outputs to the event data for transmission to the central DAQ system, or by a separate software process through the Ethernet port. This interface is not shown in Fig. 8.30 as it is not yet specified.

8.2.3 Generic board family description

The fundamental aim of the generic board family is to have a structure that allows a large number of optical links and large FPGA(s), while also providing the ATCA and Ethernet interfaces described above. This is a development in common with several other subsystems on CMS. The basic concept is to provide a base board with the ATCA and Ethernet interfaces, as well as the connectors ("cages" or "shoes") for a large number of FO links. The current design uses a mixture of Firefly and quad small form-factor pluggable (QSFP) connectors. The generic board will allow a variety of FPGAs and numbers of links to be used, depending on the subsystem needs, without redesign of the ATCA sections of the board. The aim is to allow different FPGAs to be used by mounting them on small daughterboards which are then attached onto a genuinely common ATCA base board using high-density interposer connectors [48] with around 2000 pins. The interposer concept will require some R&D to prove robustness and this will be done jointly by the groups involved in the generic board project. The fallback would be to have different base PCBs, where the layout is only changed in a well-defined area on particular layers of the PCB. Much of the required firmware for the interfaces will be common, allowing significant reuse and hence a reduction in the engineering effort required. Figure 8.31 shows the conceptual idea for this generic design and the layout of the first prototype board.

For control and configuration, there will be a daughterboard with an embedded CPU which will implement the actual ethernet connection; the Advantech COM Express Mini board [49] is being tested for this purpose. Its interface to the rest of the board will be through its PCIe connector. There will also be TCDS clock and control signals transmitted across the ATCA backplane fabric to and from the board. One minor complication is that the fabric connections to both central hub slots in the ATCA crate must be implemented so that either can be used by the DTH. The board must then be configured to select one or the other set. In addition, the throttling and any other fast signals from the FPGAs back to the TCDS must be merged with some prioritisation logic. These issues will be handled by a small Artix FPGA that implements the direct communication with the backplane and then fans signals into and out from the main FPGAs. This can handle backplane signals at up to 3.2 Gb/s and will interface to the rest of the board using LVDS. It will receive two clock signals, 40 and 480 MHz, as well as the fast and run control signals. The precise specification of these signals by the central DAQ group has not yet been fixed. It may also be necessary to implement a clock-cleaning circuit to give the clock accuracy needed for the precision timing measurements; see Section 8.4.

For the HGCAL-specific application, this document assumes 96 links for data transfer in both directions on the boards used for both the HGCAL DAQ and TPG systems, although these two applications will use different daughterboard designs. The links are assumed to be able to run

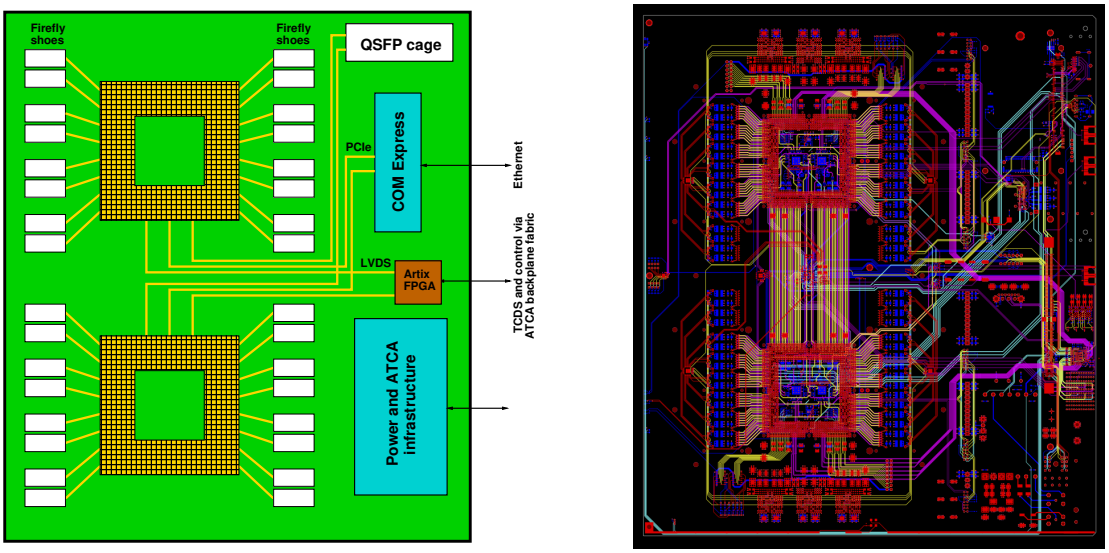


Figure 8.31: Left: Simplified conceptual layout of the generic ATCA board, showing in yellow the two large interposer areas for the FPGA daughterboards and in white the connectors (Firefly shoes and a QSFP cage) where the required number of optical transceivers can be inserted. The interposer connections are identical under a 180 ° rotation. Right: Layout of the first prototype version of the generic board.

at speeds up to 16 Gb/s, although a maximum of 25 Gb/s would also be possible at extra cost. All the links to and from the FE electronics will be LpGBT format and so are limited to 10 Gb/s. Hence, 16 (or 25) Gb/s links are only relevant for data transfer between off-detector boards, where 64b/66b encoding will be used. The 96 links are used to transfer data between different parts of the HGCAL electronics systems. Additional links are required to send event data from the HGCAL board to the central DAQ system for both the DAQ and TPG applications. As stated in the previous section, the average event data rate into each DAQ board will be around 180 Gb/s. Significantly, this exceeds the nominal 100 Gb/s for each data stream to the DTH and hence two such streams will be needed from each HGCAL DAQ board to the DTH board. These will each physically consist of 6×16 Gb/s optical links.

This large number of links will be driven directly by the FPGA transceivers. The DAQ board will contain two identical daughterboards, each holding a Xilinx KU115 Ultrascale FPGA [50] plus a small boot PROM. This FPGA contains 64 high-speed duplex transceivers capable of handling 16 Gb/s links. Of the 64 links per FPGA, 48 will be used for the main data transfer and a further six will be used to implement a 100 Gb/s link to the DTH board. One more transceiver pair will be used to implement the PCIe connection to the COM Express card. The LVDS interface to the Artix FPGA will use non-transceiver I/O pins. A conceptual idea of the DAQ-specific board layout is shown in Fig. 8.32.

The board is a potential bottleneck as the maximum possible DAQ input data rate is around 800 Gb/s while the output can only go as high as 200 Gb/s. While the average rate expected with the worst case of irradiated noise and an average pileup of 200 is lower than 200 Gb/s, there will be fluctuations around this average, so there must be sufficient buffering to avoid DAQ throttling. This can be implemented internally using the FPGA resources, but if needed, an alternative would be an external memory on the daughterboard.

The power used by each FPGA when running with most resources used will be around 75 W, below the nominal 150 W limit for the interposer. The generic ATCA section of the board is estimated to require around 50 W. Each 12-channel Firefly component uses around 2 W, giving

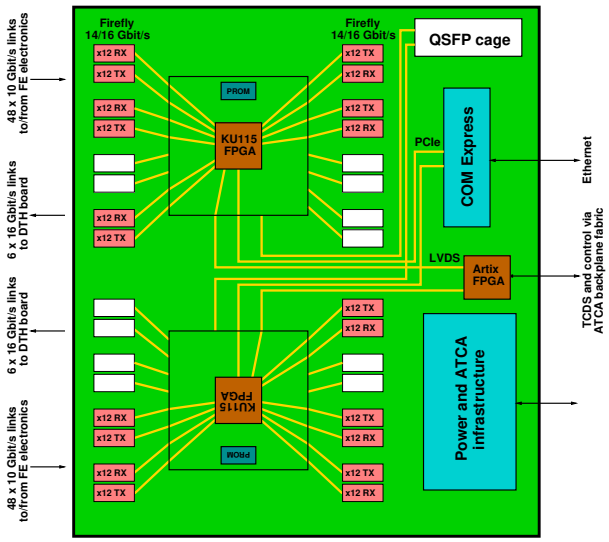


Figure 8.32: Simplified conceptual layout of the HGCAL DAQ board specialisation of the generic board. Because of the rotation symmetry of the interposer connections, the two daughterboards are identical.

a total for optical components of 40 W per board and meaning each board will take around 240 W. This is within the 400 W ATCA power supply specification, but R&D is needed to ensure that the crate cooling system will be sufficient to handle this heat load.

8.2.4 Overall DAQ system

The HGCAL DAQ board described above handles 96 link pairs for reading the DAQ data from the FE electronics and sending the clock and control to the FE. Given the number of links, this will require at least 83 DAQ boards. A custom FO patch panel per ATCA crate will be used to optimise the packing fraction of the links to the boards. The patch panels also have another benefit, namely that the FO cables which are handled (and hence at risk of damage) when replacing boards are not the ≈100 m cables from the FE electronics.

The interface to the central DAQ and TCDS systems is through a DTH board. The DTH boards come in two flavours; "DTH400" has a 400 Gb/s input and output bandwidth while "DTH1200" has 1200 Gb/s. The latter nominally can handle a full crate of twelve subdetector DAQ boards, each supplying 100 Gb/s. As the HGCAL DAQ boards will implement twice this rate, then each HGCAL ATCA crate will need to be operated as two virtual half-crates, with each half-crate containing one DTH1200 and six HGCAL DAQ boards, as shown in figure 8.33. With this arrangement, the required backplane distribution of the TCDS and control signals can be implemented using the fabric of a standard dual star backplane. This layout also preserves the desired "tree" structure of the central DAQ system which will allow easier partitioning.

Each DTH1200 will receive two 100 Gb/s duplex links (implemented as two sets of six 16 Gb/s duplex links) from each HGCAL board. The DTH input connectors consist of nine 24-channel MPO FO connectors, which each containing 8 input, 8 output, and 8 unused channels. Since the communication to the HGCAL boards must be duplex, this means that all the active channels on these connectors must be used to achieve the required input bandwidth. However, the DAQ path connectors on the HGCAL board are two 24-channel MPO connectors, each with 6 inputs, 6 outputs, and 12 unused channels. In order to match the granularities, a custom optical cable will be required between each pair of HGCAL boards and each trio of MPO connectors on the DTH board, as shown in Fig. 8.34. The whole system will have 42 such cables. The total

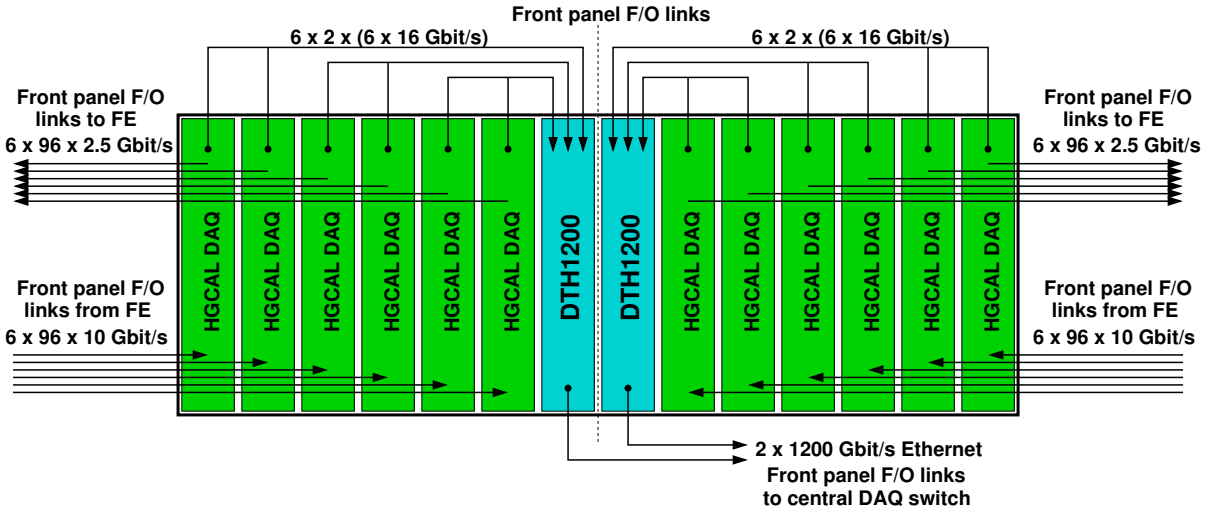


Figure 8.33: Layout of the HGCAL DAQ and DTH boards in an ATCA crate. The bandwidths indicated are the capacities of the links, not the expected rates. Not shown are the backplane connections for DAQ control and TCDS from the DTH to the HGCAL DAQ boards. The right and left halves of the crate operate effectively independently of each other, so each forms a virtual half-crate, as indicated by the vertical dashed line. Each of the links from the DAQ boards to the DTH1200 boards across the top of the diagram corresponds to one custom cable as shown in Fig. 8.34.

average input rate to six HGCAL DAQ boards at 750 kHz will be 1000 Gb/s, so the DTH1200 output rate allows significant overhead to handle rate fluctuations and also for any extra data volume due to data reformatting or additional monitoring and calibration data produced in the HGCAL DAQ boards.

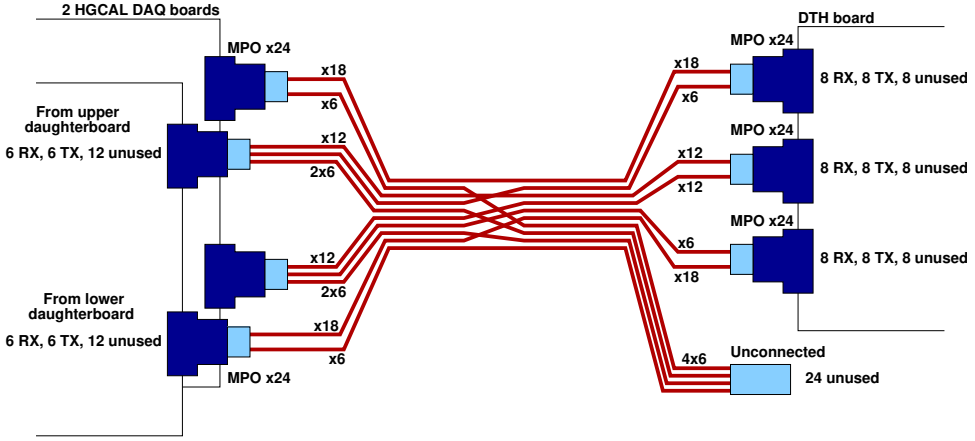


Figure 8.34: Conceptual layout of the custom cable required to transmit DAQ data between a pair of HGCAL DAQ boards and three of the nine MPO connectors on the DTH1200 board.

With a total of 84 HGCAL DAQ boards, then 7 ATCA crates (with one half full) and correspondingly 14 DTH1200 boards will be required. Since a rack can hold only two crates, these will occupy four racks.

8.2.5 Online monitoring and calibration

To aid the offline MIP calibration (Section 5.4), the HGCROC may implement a “threshold bypass”, forcing one channel per ASIC to be read out for a subset of events, irrespective of its

threshold, and so giving access its the full ADC spectrum. Having the accurate shape of the pedestal under the MIP peak will help to improve the calibration MIP fit accuracy.

The rate of such events will be configurable but typically might be once in 128 events. The channel selected can be cycled through all the channels in the ASIC over time. Even though each channel will typically only be read out once every 10k events, this would still be at rates of up to 75 Hz per channel. This would add less than 1% to the overall DAQ data volume. The actual rates of the bypass, channel cycling and histogram readout will be configurable and determined based on experimental needs after gaining experience of the likely rates of change of the pedestals and needs of the calibration measurements.

8.2.6 Testing plan

The testing plans and schedules for the DAQ and TPG boards are the same, as they are clearly very similar in terms of design. There will be three rounds of fabrication; prototype, pre-production, and production. The prototype boards will first be thoroughly tested for internal errors. These tests will include ethernet access to the COM Express card, FPGA programming through JTAG and the network, link bit error rate measurements using loopback, fast control and response using the Artix FPGA, and temperature measurements of the components when running. When the board is verified, tests of all the major interfaces (DTH, L1T, and concentrator, assuming these are all available) will be performed. The results of these tests will feed into the BE final electronics specification that forms part of the HGCAL EDR in 2020.

Following the system specification, a version of the boards which in principle would meet the requirements for the final system will be designed. This version will be used to produce a pre-production manufacturing batch of around ten boards for each of the DAQ and TPG. These will again be tested internally and with the same interfaces as for the prototypes. In addition, further tests will be performed, namely a multiple board test (probably half an ATCA crate), a slice test using real motherboards, and finally a beam test, ideally using the boards to handle the trigger. The results of all the pre-production tests will be documented as part of the production readiness review scheduled for 2022.

Having passed this review, any modifications needed for the final versions will be made and the production boards will be manufactured. These will come in at least two batches to reduce the probability of a fault in fabrication invalidating the whole set. Multiple vendors may also be used, to reduce the chance of significant manufacturing losses or delays. The first batch will be thoroughly tested against all interfaces as above before releasing subsequent batches. The overall production of around 250 boards in total has been scheduled over two years. Production acceptance tests will be done on all boards delivered, probably at two sites. This requires an average test rate of one or two boards per week at each site. When sufficient production boards have been obtained, multi-crate tests of the first endcap will be done at the surface test area. There will only be sufficient cooling to run one 60° sector at a time, which corresponds to 1/12 of the total system. This will require effectively one crate of DAQ boards and a similar number of TPG boards, which should be able to implement a dedicated cosmic trigger for this purpose. The second endcap will also be tested in the same way when it is ready.

Finally, the full set of DAQ and TPG boards will be installed in the USC and tested against the complete central DAQ and L1T systems before the first HGCAL endcap is lowered into the CMS cavern. The BE systems need to be commissioned before lowering the endcap so that they are available to check the endcap as soon as it is installed. The same procedure will be repeated for the second endcap.

8.2.7 Summary of R&D required

The main developments needed for the HGCAL DAQ system are to produce the DAQ boards themselves, but also to understand the limitations of the use of the ATCA format. The same is true for the TPG system described in the next section. Many of the issues are common for both, and indeed more generally are common across most systems in CMS in terms of their Phase-2 upgrades for off-detector electronics. Because of this, a joint R&D effort has been started, including groups involved in the tracker, barrel muons, and central L1T correlator, as well as the HGCAL. Several of these groups were involved in the development of the MP7 board [51], currently in use in several systems in CMS. This has demonstrated that common hardware, firmware, and software can be effectively deployed.

The off-detector boards for the systems involved in this effort will clearly have different requirements in detail, but the aim is to keep as much of the board hardware design similar, if not identical, for all systems. The target implementation for this is to use interposers [48] that allow very high density interconnects between a main generic PCB and its daughterboards; see Fig. 8.35. This will allow all systems to use a small number (ideally one but in reality probably more) of main PCB designs, while making varieties of daughterboards specific to their application. The main base board would contain a lot of the ATCA-specific items and infrastructure, so the daughterboards would be relatively simple. The use of interposers will need some R&D to understand their limitations and how to ensure they are robust, i.e. have a high connection efficiency and low failure rate.

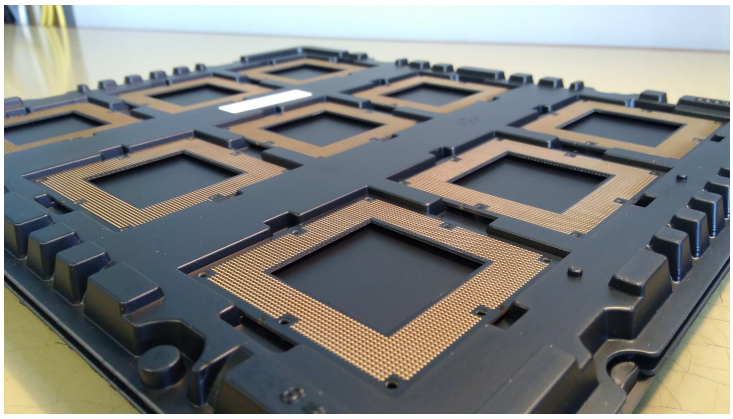


Figure 8.35: Nine interposers stored in their shipping tray. Each interposer has 1852 pins that press in both directions to make electrical contact when the interposer is fastened between two PCBs.

The main advantages of this concept are clearly the commonality of the ATCA hardware, where there is currently little experience within CMS, and the ability to share firmware and software. Even on the daughterboards, many of the functions needed will be similar and so common firmware to drive link transceivers, handle the TCDS interface, and provide the PCIe (and hence Ethernet) interface will save a lot of development time and effort. Clearly, common software to access the board, through the Ethernet port which is on the base board, will bring significant advantages too.

The first step has been to produce an ATCA service card, which has implemented the functionality needed for the generic base card in a modular form. Figure 8.36 shows the service card. At the time of writing, this is under test and the results of these tests are feeding into the first prototype design for the generic base board. The design of this prototype is effectively complete at the time of writing; its PCB layout has already been shown in Fig. 8.31 and assembled

protoypes are expected to be available in Q1 2018.

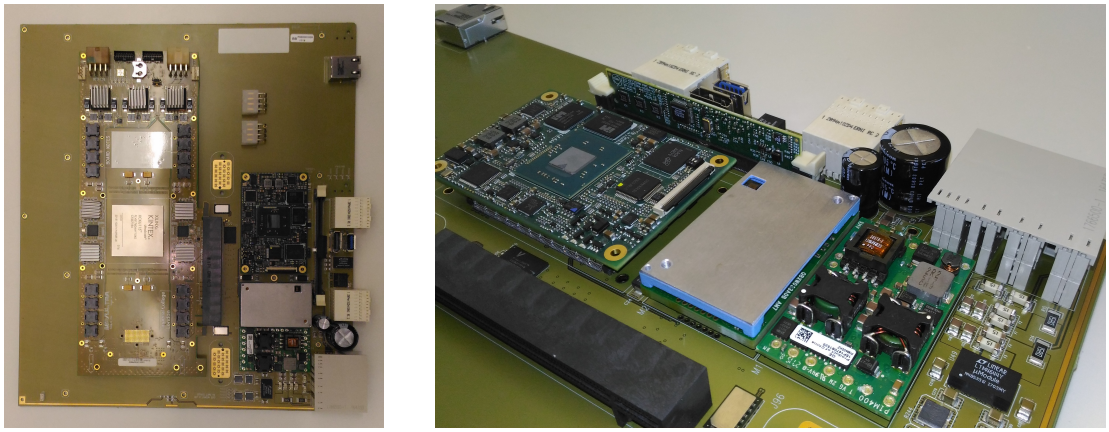


Figure 8.36: Left: The ATCA service card used to demonstrate the concept of the generic board. It is shown with an MP-Ultra board attached, which has a KU115 FPGA and 48 FO link connectors, and hence is very similar in concept to the daughterboard needed for the HGCAL DAQ application; see Section 8.2.3. Right: Close-up of the ATCA generic section of the service card. The largest board (left of centre) is the COM Express card.

Another major area of R&D will be gaining experience with ATCA boards and crates. In particular, cooling is seen as a potential major issue as the off-detector link optics needs to be kept below 50 °C to ensure their lifetime is sufficient for the HL-LHC period. While the optical components are pluggable and can be easily replaced if failures do occur, they are expensive and the cost will be very high if there are many failures.

For the HGCAL-specific parts of the system, the daughterboard designs should be quite straightforward and indeed may be very similar to some of the other systems. The main R&D needed will be in the firmware design to give an evaluation of the resources needed. This will influence which FPGAs can be used on the daughterboards. For instance, the KU115 FPGA chosen for the DAQ board is the top of the Kintex Ultrascale range and so has a large number of resources and hence a very high probability of meeting the requirements. However, the KU085 is a potential cheaper alternative given its number of 56 transceivers, and the choice will be reconsidered when a better resource evaluation is available.

More generally, the central DAQ system, and for the TPG case the central L1T system, are both still under design and it is very likely the interfaces will change, certainly up to the times of their respective TDRs. Hence, the HGCAL DAQ and TPG systems will need to evolve in parallel with any central changes.

8.3 Trigger primitive generator

8.3.1 Overview

The level-1 TPG is also assumed to be implemented from the generic family of ATCA boards, as used for the DAQ. This means there are no required hardware developments for the specific TPG implementation described in this section beyond a different daughterboard. This leaves a lot of flexibility as even the specific board described later would be usable for many different TPG designs. Therefore, studies of alternative architectures and algorithms for the TPG will continue into the future. Hence, the main aim of this section is to show that there is at least one realistic TPG implementation, although this may not be the optimal arrangement and hence

may not be the one eventually used.

The inputs to the TPG are the trigger raw data, delivered on LpGBT links from the on-detector electronics. Only the first and subsequent alternate layers from the CE-E, i.e. 14 of the first 28 layers, send data to the TPG, while all 24 of the CE-H layers send trigger data. The raw data consist of two different types. One type are the trigger cells above a threshold, which contain the summed energy of 2×2 or 3×3 sensor cells for the coarse and fine granularity wafers, respectively. These data do not have a fixed latency as the trigger cells are zero-suppressed and so have a different size per bunch crossing (BX). The other type of raw data are those used to form energy maps, which are the sums over all the sensor cells read out by an HGCROC ASIC. Due to the varying data volume, the trigger raw data as a whole take a variable time to be transmitted to the TPG. However, to meet the requirements of a fixed latency trigger overall, the raw data must have a defined maximum latency, meaning that all the raw data associated with a BX will arrive at the TPG within a specific time range. This range is set by the FE concentrator ASIC buffer for the trigger raw data, which will be 12 BX, i.e. 300 ns, long. In terms of absolute latency, the start of the time range will be 1200 ns after the associated BX, while the range end (the maximum possible latency) will be at 1500 ns.

The TPG output data will also be of two types. The trigger cells are used to form three-dimensional (3D) clusters, giving a position, energy, and various other relevant quantities (Section 8.3.3) related to the shape of the cluster. These are intended to contribute to building physics objects in the central L1T correlator. The energy sums in η, ϕ bins for each layer are used to form a full depth η-ϕ energy map for each bunch crossing. These output data are then transmitted to the central L1T correlator to take the actual trigger decisions. All data associated with each BX sent from the TPG must arrive at the central L1T correlator within 5 μs. This latency limit is set by the estimated time required for the central L1T decision and the time to return the trigger decision back to the on-detector electronics within the required 12.5 μs total latency.

The output primitives are sent to the central L1T correlator [21] where the 3D clusters are matched to L1 tracks. This will allow a simplified particle flow algorithm to be used for the L1 trigger. In addition, some level of particle identification is performed. The energy maps are mainly used to correct all the hadronic energy of a jet, as some may be lost from the 3D clusters due to the FE electronics selection of trigger cells. The maps will also be used to calculate the total transverse energy observed in each endcap, which will feed into a detector-wide missing transverse energy measurement in the central L1T correlator. The HGCAL TPG does not issue triggers itself; it passes on information to the central L1T where the actual decision is made. As for the DAQ, the precise central L1T correlator interface has not yet been specified. The main issues are the link speed (effectively 16 or 25 Gbit/s) and whether the correlator will be time-multiplex or not, and if so, with what factor.

The proposed architecture of the TPG subsystems consists of a two-stage structure (in the following referred as Stage 1 and Stage 2) and these stages are implemented in different groups of hardware boards. In the Stage 1, the trigger cells are clustered within each layer separately, to form a list of two-dimensional (2D) clusters for each layer. In addition, the data for each layer are summed into the bins used in the energy map to give an equivalent of the energy map for each single layer. In Stage 2, the 2D clusters are clustered together across layers to form the 3D clusters, while the energy maps for each layer are combined to give a total energy map. These are produced for each endcap independently and so the TPG is implemented as two identical but separate subsystems, one per endcap, with each reporting its results to the central L1T system independently.

8.3.2 TPG Stage 1

8.3.2.1 2D clustering and energy map algorithms

Stage 1 clusters the trigger cells (TC) from the FE electronics. For each HGCAL trigger layer, the algorithm generates 2D clusters which must contain at least one seed trigger cell (seed TC) characterized by an energy deposit exceeding a given threshold in $\mathrm{MIP_T}$ units, typically 5 $\mathrm{MIP_T}$. The clusters may also have secondary TCs, defined to have an energy deposit between some lower threshold, typically 2 $\mathrm{MIP_T}$ (which matches the readout threshold in the FE electronics), and the seed threshold. There are many possible clustering algorithms for Stage 1 and two different algorithms have been implemented in the simulation so far, here called the "geometrical-based cut" (GBC) and "nearest neighbour" (NN) approaches.

The GBC algorithm loops over the data twice. In the first loop, it flags all the seed TCs based on their energy deposit compared to the seed threshold. During the second loop the clusters are built around the seed TCs using secondary TCs above the lower threshold. To be added to the cluster, the secondary TC must lie within a circle with a configurable radius (typically $\approx$5 cm) around the cluster seed TC. If a secondary TC lies with this radius for more than one seed TC, it is assigned to the most energetic seed TC.

The NN algorithm aims to reconstruct the cluster dynamically, potentially reducing the PU contamination. In particular, the algorithm adds TCs to a cluster if they are nearest neighbours with at least one TC already in a cluster. As the name suggests, the definition of the nearest neighbours of a TC is crucial: currently, two TCs are defined to be nearest neighbours if they share at least one edge. Due to the complex trigger cell geometry, the current implementation of the TC nearest neighbour algorithm in CMSSW relies on a look-up table (LUT) to determine if any two TCs are neighbours. In the eventual implementation, where the sensor cell geometry is more uniform, the regular arrangement of the TCs will allow an implementation using an analytic nearest neighbour finder, needing fewer resources. The algorithm loops over each secondary TC, checking it against every cluster already created, and adds it to the cluster if it is a nearest neighbour of at least one TC belonging to the cluster. If a secondary TC is not added to an existing cluster, it is used to create a new cluster. Once all the TCs have been processed, a loop over the cluster collection is performed in order to merge nearest neighbour clusters. This extra step is needed to prevent any bias coming from the lack of a position ordering of the TC collection. Finally clusters are transmitted to Stage 2 only if the cluster contains at least one TC above the seed threshold, this is illustrated in Fig. 8.37.

In both approaches, the total transverse deposited energy of the cluster is evaluated as the sum of the transverse energy deposited in all the TCs associated with the cluster. The cluster centre in x and y within the layer is calculated as the energy-weighted average of the TC positions within the cluster.

Figure 8.38 shows the number of 2D clusters per layer reconstructed in $\mathrm{t\bar{t}}$ events with 200 pileup. This is shown for every HGCAL layer, including the CE-E layers which do not actually contribute data to the TPG. This uses the NN algorithm although very similar results are seen with the GBC algorithm. The clusters found mainly result from the pileup events, with the physics event being at most a $\approx$10% effect. It is seen that the number of 2D clusters in some CE-E layers can be higher than 300, although for the CE-H, the maximum number seen per layer is always below 100.

Stage 1 also combines the HGCROC coarse sum data into a preliminary energy deposit map for each layer. Since the HGCROC ASICs, which produce the mapped data, do not have a projective geometry, a mapping of the HGCROC areas onto the η, ϕ bins is required. Again,

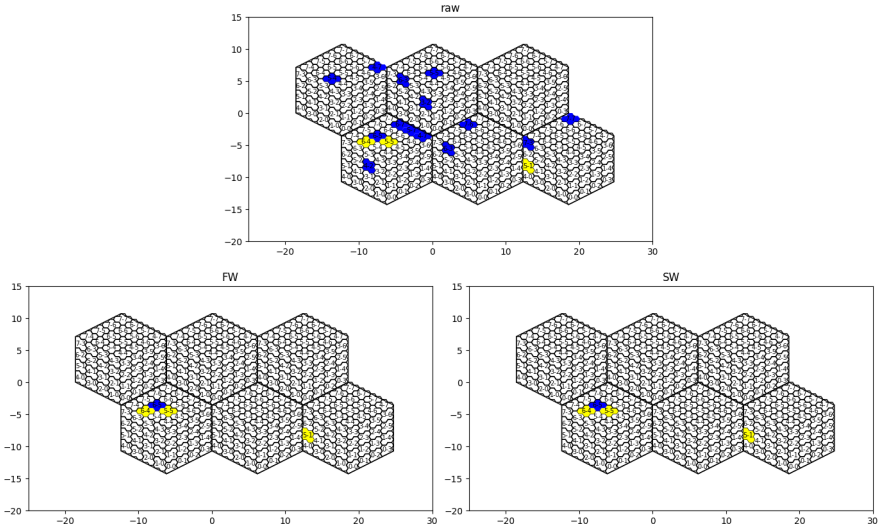


Figure 8.37: Example of 2D clustering on six sensors. The seed trigger cells are shown in yellow and the secondary trigger cells in blue. The upper plot shows all trigger cells. Only clusters which include a yellow seed trigger cell are retained, as shown in the bottom two plots which show the results from the VHDL implementation (left) and software implementation (right) agree.

there are several algorithms which could be used for this step. The most straightforward is to find the η, ϕ bin that has the largest area overlap with the coarse sum data and then add the whole coarse sum data energy deposit into that bin. Due to the irregular sensor geometry in η, ϕ coordinates, this needs to be implemented as a LUT. The resulting η, ϕ bins then contain the total deposited energy in units of $\mathrm{MIP_T}$ from every HGCROC in the layer.

8.3.2.2 2D clustering firmware implementation

A working version of the Stage 1 NN algorithm has been implemented in VHDL, where the chosen target hardware is the Xilinx Virtex 7 family, as mounted on the MP7 [51]. The firmware implementation, whose schematic is shown in Fig. 8.39, is based on the assumption that the TCs are arranged according to the final sensor geometry. This means that the nearest neighbours of a TC can be analytically calculated, and each TC can be unequivocally identified in a Cartesian-like coordinate system.

The data stream (made of 17b words containing the TC energy and channel number) is processed by the seeding module ("Seeding") that flags the TC as a seed (in accordance with the definition given above). Once seeded, the data are copied to two stream branches. One branch goes into the "Seed distribution"; its role is to distribute the seeds into $N + 1$ cluster-building logic blocks ("C0" to "CN"). The second branch is connected to a memory where the data are temporarily stored ("Delay Line"); this memory acts as buffer for the data in order to wait until all the seeds are stored in the respective clusters.

Once the seed distribution is complete, the data are fanned out to all the clusters relative to a specific BX. TCs are stored in a cluster according to their position. At the end of the data stream the clusters are sequentially processed in order to select TCs topologically connected to

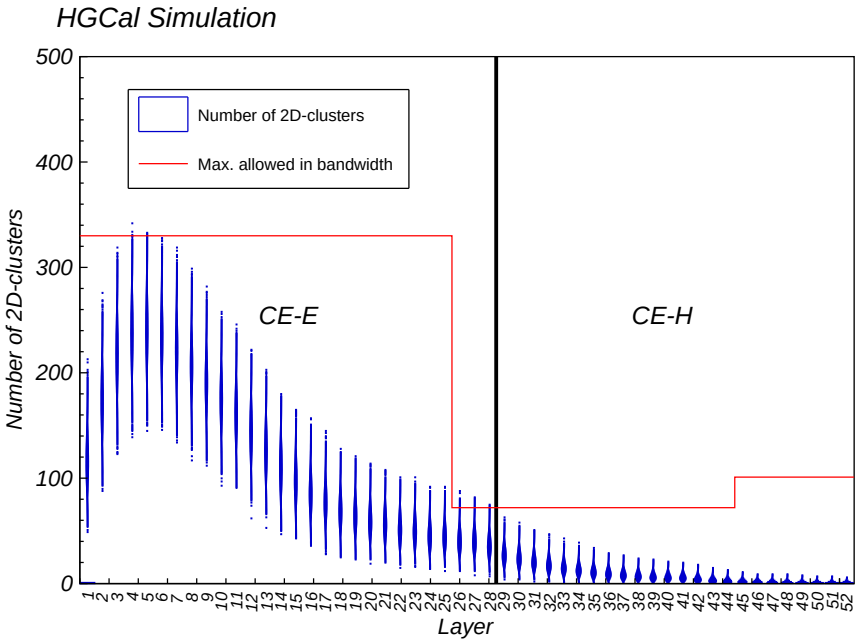


Figure 8.38: Number of 2D clusters per endcap reconstructed in $t\bar{t}$ events with 200 pileup as a function of layer. The red histogram shows the bandwidth limit due to the allocated number of links, as discussed in Section 8.3.4.

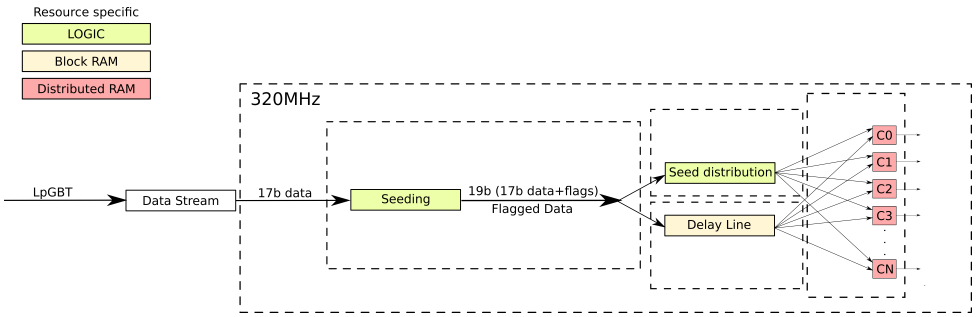


Figure 8.39: Diagram of VHDL implementation of the NN clustering algorithm; see text for details.

the seed TC. Finally, the selected cluster data are transmitted to the Stage 2 processors.

Each cluster is made of a memory that stores the TCs energy and an occupancy word, which maps all the TCs found within the cluster. A maximum of 5×5 TCs along the coordinate system directions are considered for each cluster, where the seed TC is always placed in the matrix centre. The cluster is built around the seed TC and the position of the secondary TCs added to the cluster are always referenced to the seed position.

The firmware algorithm fits within the resources of the Virtex 7 FPGA (with 72 input links); the probable FPGA (with 96 input links) that will be used in the TPG system (see below) will have at least twice the resources, but dealing with ≈30% more data. The latency depends linearly on the number of TCs in the raw data, which fluctuates per BX. Figure 8.40 shows the variation in latency due to this effect using the clock speed of 320 MHz expected in the final system; a maximum latency of 700 ns is observed. This means the 2D clusters must be buffered to this worst case delay so they have fixed latency when transmitted to Stage 2.

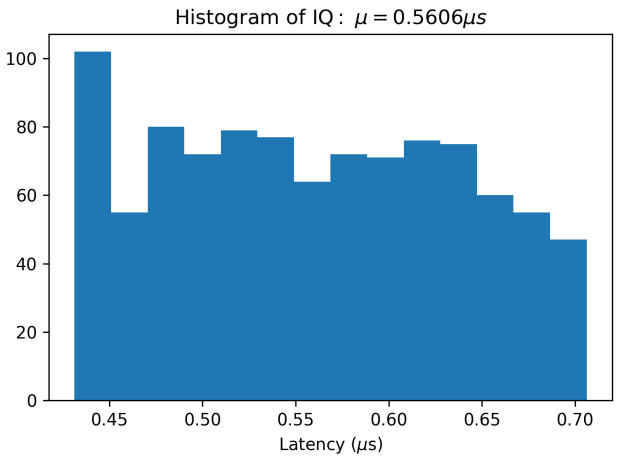


Figure 8.40: Distribution of latency for the implemented Stage 1 algorithm. The variation is due to the differing numbers of trigger cells per BX.

The algorithm already implemented should be the major part of the firmware design. However, it is not yet complete and has some limitations. The main ones are that operations are limited to the data within a single link so clusters are not formed from TCs from different links, TCs can be double counted (i.e. the same TC can be assigned to more than one cluster), the clusters are limited in size because of the 5×5 arrays used, there is a limit of six seed TCs per link, and the energy map bin data mapping is not yet included. Allowing for these additional computations, the maximum total latency is estimated to be up to 1 μs.

8.3.2.3 Output data format

The output data packet from Stage 1 consist of a fixed size header and a variable size list of 2D clusters. Since the TPG output for each endcap is independent, each endcap layer in Stage 1 sends one header and one list per BX.

The header contains some overall BX quantities and the energy map for the layer being processed. The format of the data for the header per layer is shown in Table 8.1. It contains the total transverse energy (in $\mathrm{MIP_T}$ units) found in the whole endcap, the BX number for diagnostic purposes, and the number of 2D clusters in the following list. This number is limited to 8 bits and hence a value of 255, requiring truncation of the 2D cluster list if the number of clusters is higher than this value. As shown in Fig. 8.38, this should only affect two CE-E layers and only for a small fraction of events. In addition, the header contains the contribution to the energy map for this layer, again in $\mathrm{MIP_T}$ units. As the energy map is a large amount of data, then floating point compression of the values to reduce their size could be performed if needed, at some cost in FPGA resources.

Table 8.1: Concept for the layer header data sent from Stage 1 to Stage 2.

Quantities	Bits	Total bits
Total transverse energy, BX number, number of 2D clusters	16, 8, 8	32
Energy map 15 (η) $\times$72 (ϕ)	12	12 960
Total		12 992

Following the header, a list of 2D clusters is sent. The format of the data for each of the 2D clusters is shown in Table 8.2. For each cluster, there is an initial fixed data size followed by optional extra data. The fixed data part contains 64 bits, giving the fundamental information of the cluster position in x and y and the transverse energy in $\mathrm{MIP_T}$ units of the cluster. It also then contains other information on the cluster, namely the number of TCs which formed the cluster, measures of the cluster width in x and y, and some quality flags (to be defined). Finally, if the cluster has more than one local maximum (typically the number of seed TCs which are not nearest neighbours), then the number of local maxima is given as a 2-bit value, where the stored value is the number minus one (as there is always at least one seed TC in a cluster). This means the number of local maxima is limited to four, so if there are more the number is truncated. If the number of local maxima is greater than one (i.e. the stored value is non-zero), then the positions and energies of these local maxima follow; these are the optional data. The maximum cluster data length, when there are four local maxima, is then 160 bits. Simulation studies show that around 90% of 2D clusters have only one seed TC and the average cluster size is around 70 bits.

Table 8.2: Concept for data per 2D cluster sent from Stage 1 to Stage 2.

Quantities	Bits	Total bits
x, y, E_T	12, 12, 8	32
Number of cells and local maxima, size in x and y, quality flags	8, 2, 8, 8, 6	32
Minimum total		64
Optional local maximum 0 Δx, Δy, normalised E_T	8, 8, 8	24
Optional local maximum 1 Δx, Δy, normalised E_T	8, 8, 8	24
Optional local maximum 2 Δx, Δy, normalised E_T	8, 8, 8	24
Optional local maximum 3 Δx, Δy, normalised E_T	8, 8, 8	24
Maximum total		160

8.3.3 TPG Stage 2

8.3.3.1 Going from layers to three dimensions

Stage 2 receives as input the 2D clusters from the Stage 1 and combines them in order to generate the 3D clusters. It also combines the energy maps from each layer to form an overall energy map for each endcap.

As for Stage 1, there are many possible clustering algorithms and again two have so far been implemented. These are called the cone-based (CB) and log-likelihood (LL) algorithms. The CB approach first calculates the projected centre of each 2D cluster. The projected centre is defined as the cluster centre in the x,y plane divided by its longitudinal (z) position, i.e. x/z, y/z. Clusters in different layers that lie on a line from the origin will have the same values of these projected coordinates. After this step, 3D clusters are generated by looping over the 2D clusters and comparing their projected coordinates. The algorithm works in a similar way to the GBC algorithm used in Stage 1. Each projected 2D cluster can seed a 3D cluster and all the 2D clusters that lie within a distance in the projected plane smaller than a configurable value, typically 0.01, are added to the 3D cluster.

The LL algorithm is similar but takes into account the spread of the 2D clusters to form an approximate LL value for their overlap. This LL is used in deciding whether to put the 2D clusters into the same 3D cluster or not, rather than just using the radius in the projected coordinates.

For both algorithms, the 3D cluster energies are found using a weighted linear sum of the associated 2D cluster transverse energy deposits. Each layer has a different weight which is pre-determined and loaded as part of the TPG configuration data. One subtlety is that the weights depend on the assumed particle type and whether pileup subtraction is needed. An electron or photon will very rarely leave much energy in the CE-H layers, so 2D clusters in those layers associated with an assumed electromagnetic (EM) 3D cluster are probably due to pileup. In contrast for hadrons, 2D clusters in the first few layers of the CE-E are mostly due to pileup. Hence, the weights in these regions can be made small to reduce the effect of the pileup. In fact, fitting for optimal weights shows that making the weights negative is these regions both improves the energy resolution and gives a first-order pileup correction. Energies using weights both with and without this pileup correction will be used. However, in general the particle identication is not unambiguous in the TPG. Hence, the deposited layer energies are used to calculate four different particle energies for each cluster, based on two (EM and hadron) different particle type assumptions and with and without pileup removal. This requires four sets of weights; see Section 10.3.2 for further details. Since they are most numerous, the hadron energy is used as a default but the EM energy is optionally also included in the 3D cluster data packet. The 3D clusters are only transmitted to the central L1T correlator if their transverse energy calculated using the hadron interpretation with pileup removal is above a configurable threshold (typically $\approx$1 GeV).

Figure 8.41 shows the number of 3D clusters resulting from the CB algorithm per endcap in $t\bar{t}$ events with 200 pileup, for various choices of the final transverse energy threshold. Again, the total number is dominated by the 3D clusters from the pileup and there is little difference between the two algorithms. The main use of the 3D clusters is to match to tracks formed in the L1 track TPG in the central L1T correlator. The L1 track finding has a nominal transverse momentum threshold of 2 or 3 GeV but with a rather slow turn-on curve, so tracks down to $\approx$1 GeV can be reconstructed. Hence, a cut at around 1 GeV would allow minimal additional loss in track-matching efficiency. It is seen that this cut gives up to 200 3D clusters per endcap.

The energy maps per layer are combined to give one overall η-ϕ energy map per endcap. Similarly to the 3D clusters, these are combined as a linear weighted sum of each of the layer energy maps. The weights are again pre-calculated using only a hadron interpretation; since the energy maps contain a large number of bins and hence volume of data, there is not sufficient bandwidth to the central L1T correlator to allow for other interpretations. Again, negative weights can be used to perform a pileup correction, although whether this is implemented or not will be determined by studies within the central L1T group.

8.3.3.2 Output data format

The format of the header data per endcap is given in Table 8.3 and of the data for each 3D cluster is in Table 8.4. Similarly to the Stage 1 output data header, the Stage 2 data header contains some general BX information and the energy map. The values for each bin take 16 bits, with 12-bit floating point value for the transverse energy and 4 bits for the fraction of energy in the CE-E. The number of 3D clusters per endcap is stored as an 8-bit value and hence has a maximum value of 255.

The 3D cluster data need to contain all the information for the central L1T correlator to perform particle flow and particle identication. As for the 2D clusters, this contains a fixed size section and optional data. The former contains the fundamental cluster information, i.e. the transverse energy and position (given as the starting point of the cluster). It also stores the fraction of energy in the CE-E (and rear of the CE-H) and the layer which contains the maximum energy,

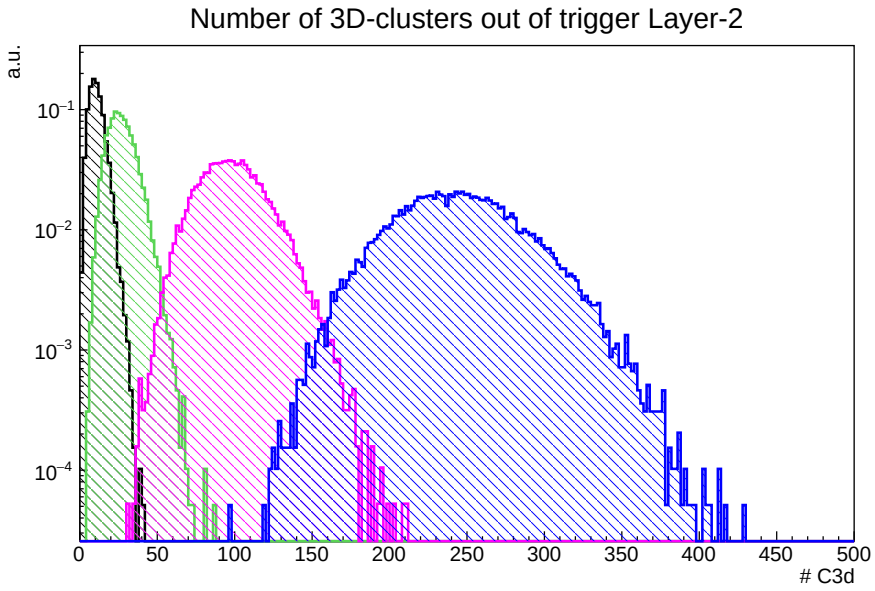


Figure 8.41: Number of 3D clusters per endcap reconstructed in $t\bar{t}$ events with 200 pileup. The thresholds applied to the 3D clusters are $E_T = 0.5$ GeV (blue), 1.0 GeV (pink), 2.0 GeV (green) and 3.0 GeV (black).

Table 8.3: Concept for the header data sent to the central L1T correlator per BX.

Quantities	Bits	Total bits
Total energy, BX number, number of clusters	16, 8, 8	32
Energy map 15 (η) $\times$72 (ϕ)	16	17 280
Total		17 312

both of which give information on whether the cluster is EM-like. Finally there is also quality information (number of trigger cells in the cluster and other flags) to select of good clusters, while the final fixed data are flags to indicate which optional data are present. The optional data includes shape quantities (to be defined), the cluster energies under an EM interpretation, and then some subcluster information. These last data may be useful when several particles are close enough to merge into one cluster; a physics example of interest would be three-prong tau decays.

These cluster sizes vary between 128 and 416 bits. With an average of around 200 bits, then the 3D cluster data per endcap will be around 40 kbit/BX. Including the header data, therefore the primitives data will total around 60 kbit/BX per endcap.

Generally, the exact definition of the 3D cluster data (and hence also the 2D cluster data from which the 3D clusters are formed) will be driven by studies of its use in the central L1T correlator. As such, the data format shown here is expected to evolve as more is learned about the optimal correlator algorithms. Since the data are defined purely through the firmware, this will not affect the hardware designs.

Table 8.4: Concept for data per 3D cluster sent to the central L1T correlator.

Quantities	Bits	Total bits
E_T with and without PU subtraction	16, 16	32
Endcap, fraction in CE-E, fraction in back CE-H, max energy layer	1, 13, 12, 6	32
Shower start η, ϕ, z	11, 11, 10	32
Number of cells, quality flags, extra data flags	8, 12, 12	32
Minimum total		128
Optional shape quantities	8×16	128
Optional e/γ reco E_T with and without PU subtraction	16, 16	32
Optional subcluster 0 E_T, $\Delta\eta$, $\Delta\phi$	16, 8, 8	32
Optional subcluster 1 E_T, $\Delta\eta$, $\Delta\phi$	16, 8, 8	32
Optional subcluster 2 E_T, $\Delta\eta$, $\Delta\phi$	16, 8, 8	32
Optional subcluster 3 E_T, $\Delta\eta$, $\Delta\phi$	16, 8, 8	32
Maximum total		416

8.3.4 Hardware implementation

The hardware implementation of the TPG reflects the logical structure described above and an overview has already been shown in Fig. 3.10. The 2D clustering is performed in separate hardware from the 3D clustering, so as to spread the firmware load more conservatively. Hence, the hardware is also grouped into Stage 1 and Stage 2. The same boards will be used throughout both stages of the TPG, with the only difference being that the Firefly optical components mounted on these will differ for Stages 1 and 2.

As stated in Section 8.2.3, the TPG board will have 96 input and output FO links per board for the TPG data. This large number of links will require a high-end, modern FPGA. The board assumed here will contain one daughterboard with a Xilinx VU9P Ultrascale+ FPGA [50], which is available in a package with 104 high-speed transceivers. These can drive links at either of the 16 or 25 Gbit/s discussed previously; 16 Gbit/s is assumed here for cost reasons. The DAQ board discussed in Section 8.2.3 has two smaller FPGAs handling the same number of links, which is a cheaper option. However, it is important to use a large single FPGA per board for the TPG implementation as both Stages 1 and 2 require processing data from large regions of the HGCAL. Hence, the larger the amount of data processed in each FPGA, the smaller the edge (boundary) effects will be. The TPG-specialisations of the generic board are shown in Fig. 8.42. The Stage 1 board will have only 14 Gbit/s inputs as it only receives 10 Gbit/s LpGBT links, while it needs up to 72 links running at 16 Gbit/s for its outputs (as discussed below). For Stage 2, all inputs and outputs must be 16 Gbit/s, but only 12 links to the central L1 correlator are required.

In hardware, the Stage 1 algorithm would ideally be implemented using one processing board per layer, as this would allow the 2D clustering to be performed with no boundaries and hence no need for data sharing between boards. This is possible for the CE-H layers as the number of links per layer is below 96. However, for the CE-E layers, the number of links exceeds this board limit and hence two boards are required per layer. Two choices are being investigated to handle this. The first is simply processing approximately half the layer links on each board so a single board nominally processes a range of 180 ° in ϕ. To fully reconstruct 2D clusters at the boundaries, then some trigger raw data must be sent to both boards, requiring a passive link splitter in front of the boards. Duplicating the data on 6 links (i.e. 3 links for each of the two boundaries) would mean each of the two boards could receive up to 93 links of unique data;

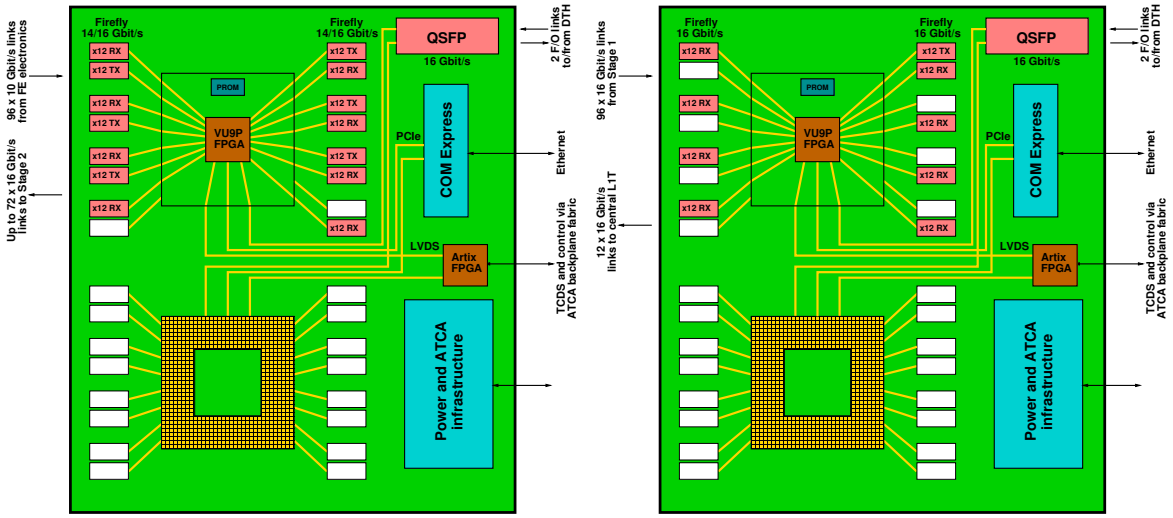


Figure 8.42: Simplified conceptual layouts of the TPG ATCA boards, based on the generic board shown in Fig. 8.31. For this implementation, only one of the two interposer sites is used. The board with the optical components installed for Stage 1 is shown on the left and for Stage 2 on the right. The daughterboards for the two cases are identical.

this allows up to 186 links in total for the layer. The 2D clustering algorithm would process the overlap regions in master/slave mode, with clusters being uniquely assigned to one or the other so they are only sent to Stage 2 from one board. Another complication is that the number of links per layer for the last 8 CE-H layers will be at most 48, so that all links from two different layers can be received by one board. This means the 2D clustering for pairs of layers from that region will be performed in one board; this is done to reduce the number of boards needed. In addition, to save on output bandwidth required, the two energy maps for the pairs of CE-H layers read by one Stage 1 board are merged before transmission to Stage 2. Hence, there will need to be at least three different versions of the 2D clustering firmware, to handle the split layers, full layers and paired layers. Table 8.5 gives the board counts for the various regions and shows a total number of 48 boards per endcap is needed for Stage 1.

Table 8.5: Numbers of TPG Stage 1 boards per endcap required to receive data from the FE electronics.

HGCAL region	Number of layers	Number of boards/layer	Number of boards
CE-E even layers only	14	2	28
CE-H layers 29–44	16	1	16
CE-H layers 45–52	8	0.5	4
Totals	38		48

Stage 2 is more uniform than Stage 1, as the data are time-multiplexed by a factor of 24 before transmission from Stage 1 to Stage 2. The multiplexing factor of 24 is determined by the time needed to transmit the required data volume between the two stages, given the fixed bandwidth into each Stage 2 board. Stage 2 then consists of 24 boards per endcap, each handling one of every 24 BXs. By definition, these boards do identical tasks and so have the same bandwdth in and out and run the same firmware. The time-multiplexing allows 24 BX, i.e. 600 ns, to transmit the data from Stage 1 to each Stage 2 board, which is sufficient to send all the data for one BX into one Stage 2 board. There are therefore no edge effects which simplifies the 3D clustering implementation significantly. There are two link count constraints, arising from each

end of the links between the two stages. Out of Stage 1, each board needs to be able to send data to all 24 Stage 2 boards and by symmetry, the bandwidth to each Stage 2 board must be the same. This puts a limit on the bandwidth to each Stage 2 board of 4 links from any Stage 1 board so that the total does not exceed 96 links. At the other end of these links, the Stage 2 board can only accept up to 96 links. Given Stage 1 has 48 boards, this requires the average number of links from each Stage 1 board to each Stage 2 board to be two.

The data format on these links is assumed to be the same as that used in the current Phase 1 calorimeter trigger. This has a 64b/66b encoding, and six 66-bit words are sent every BX per 16 Gbit/s link. Hence, the user data payload is 384 bits per BX per link. The structure of a 24 BX packet for one event consists of two 32-bit header words, the user data, a trailer 32-bit word, a checksum 32-bit word and two K-code "comma" 32-bit words. Hence, of the 9216 bits available in this packet, the usable data for the TPG occupies 9024 bits per link.

Table 8.6 shows a solution which satisfies the link count constraints, given an average 2D cluster size of 70 bytes. All inputs to the Stage 2 board inputs are fully occupied, hence giving a total of 2304 links. Figure 8.38 also shows the number of links allocated to each layer, which is therefore twice the number of links per board for the CE-E layers. This results in a complex interconnection routing for the Stage 1 to Stage 2 links and this will need a custom-made patch panel for these links, similar to that used in the current Phase 1 CALO trigger.

Table 8.6: Number of links per layer from Stage 1 to Stage 2. The maximum number of 2D clusters in the last column is calculated assuming an average cluster size of 70 bits.

HGCAL region	Layers in region	Number of links/layer	Number of links	Maximum number of 2D clusters
CE-E even layers 2–24	12	4	48	330
CE-E even layers 26–28	2	2	4	72
CE-H layers 29–44	16	2	32	72
CE-H layers 45–52	8	1.5	12	101
Total	38		96	

As outlined above, the data rate from Stage 2 to the central L1T correlator is around 60 kbits/BX per endcap although due to the 24-fold time-multiplexing, the required data rate from each Stage 2 board is around 2.5 kbit/BX or 100 Gbit/s. This will be transmitted on 12 links running at 16 Gbit/s per Stage 2 board, which allows significant overhead for modifications to the data format. This will physically be one 12-channel simplex FO cable per board. This is also well within the 96 links available, again allowing flexibility for future changes. Because all data for the BX are gathered into the FPGA of each Stage 2 board, then there is a high degree of flexibility about the order in which the data are transmitted. Specifically, it is likely that the cluster data will need to be ordered in ϕ; if this is done in a coordinated fashion across the detectors transmitting to the central L1T correlator, then the correlator can start processing immediately without having to wait for all the data from each event.

The power consumed by each VU9P FPGA (assuming effectively all resources are used) will be around 120 W, again below the nominal 150 W limit for the interposer. As stated previously, the generic ATCA section of the board is estimated to require around 50 W. Each 12-channel Firefly component uses around 2 W and the the QSFP transceivers use around 20 W, giving a total for optical components of around 50 W per Stage 1 board and 40 W per Stage 2 board. Hence the total power per board is 220 or 210 W, respectively, which is within the 400 W ATCA power supply specification, but again cooling will be an issue. In principle, both interposer

sites could be populated with daughterboards, but this would raise the power consumption to 370 W or 350 W, respectively; this is thought to be too marginal.

Each TPG ATCA crate will hold 12 TPG boards and will contain one DTH400, as shown in Fig. 8.43. Hence, the total number of TPG boards is therefore 72 per endcap and these will fit into six ATCA crates per endcap. Of these, four crates contain the 48 Stage 1 boards and the other two contain the 24 Stage 2 boards. The set of crates and racks corresponding to one endcap subsystem is shown in Fig. 8.44. This TPG endcap "subsystem" is replicated for each of the two endcaps, making a total of 144 boards in 12 ATCA crates.

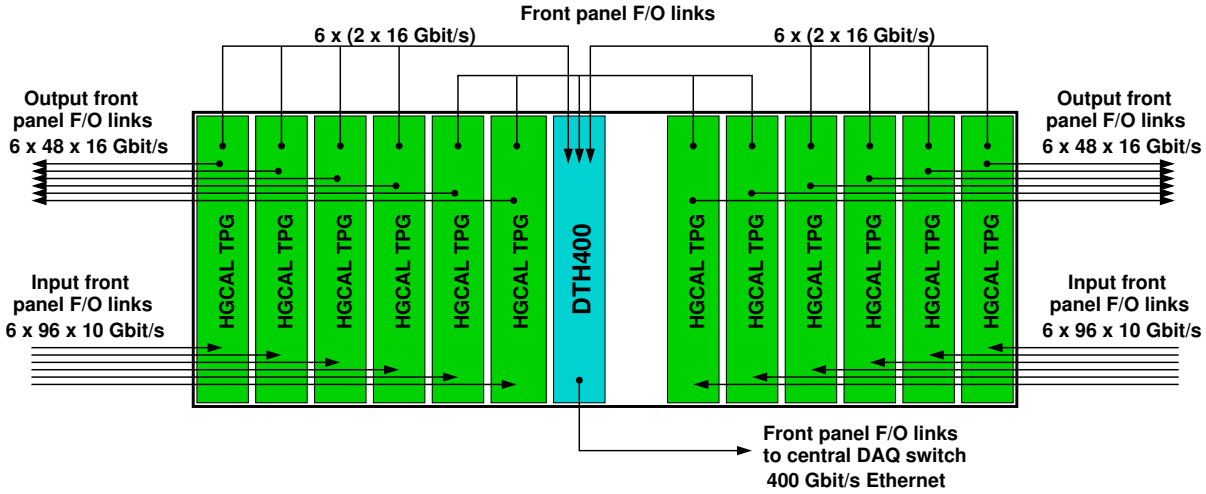


Figure 8.43: Layout of the HGCAL TPG and DTH boards in an ATCA crate, in this example for Stage 1 boards although Stage 2 would look effectively identical. The bandwidths indicated are the capacities of the links, not the expected rates. Not shown are the backplane connections for DAQ control and TCDS from the DTH to the HGCAL TPG boards. Each of the groups of four links from the TPG boards across the top of the diagram which merge at the DTH400 board corresponds to one custom cable as shown in Fig. 8.45.

The DTH400 inputs are three 24-channel MPO fibre-optic connectors and again these will be fully occupied by the TPG links. This means each of these connectors must connect to four TPG boards. The TPG data are transmitted using half the links in the QSFP connector. As for the DAQ case, a custom optical cable will be required, with one cable connecting four TPG boards to the DTH board; see Fig. 8.45. A total of 36 custom cables will be needed. The data rates on these cables are discussed in Section 8.3.6.

The design of the central L1T correlator is not yet finalised and in particular, no decision has been made about whether it will be time-multiplexed or not, and if it is, what multiplexing factor will be used. The HGCAL TPG Stage 2 is time-multiplexed with a factor of 24. If this does not correspond to the central L1T correlator operation, then a TPG Stage 3 will be needed to re-multiplex the data to the correct timing structure. This is a relatively straightforward operation but will require extra boards. There are 12 output links per Stage 2 board so a total of 288 links per endcap, which can be handled by three TPG boards of 96 links each. Hence these will require one further ATCA crate which can contain the extra six boards needed for both endcaps.

8.3.5 Latency estimate

The total latency of the FE electronics and TPG, defined to be the time at which the last bit of the trigger primitives related to a bunch crossing arrives at the central L1T correlator, is required

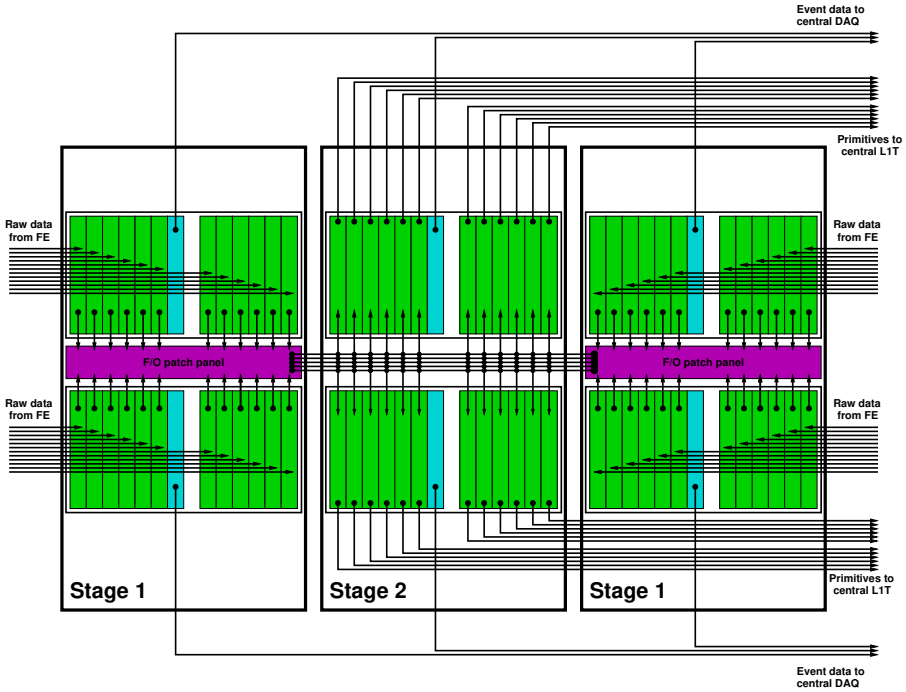


Figure 8.44: Conceptual layout of the TPG ATCA crates for one of the two endcaps. The 48 boards required for Stage 1 occupy four crates, which are in the two outer racks. The 24 boards needed for Stage 2 occupy two crates in the central rack. The subsystem for the other endcap is identical in layout.

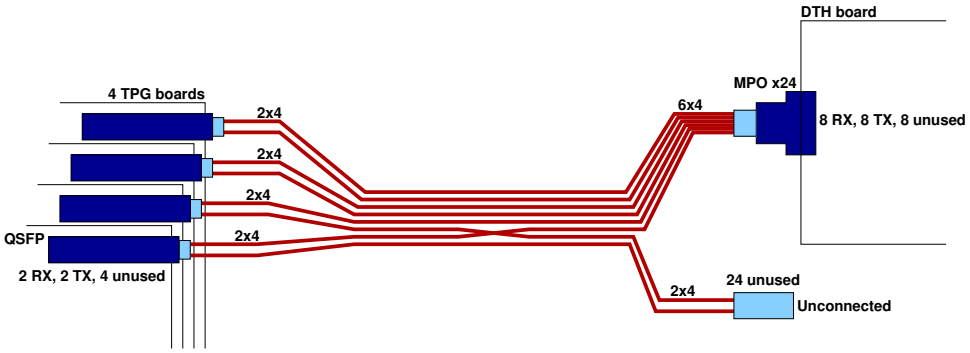


Figure 8.45: Conceptual layout of the custom cable required between four HGCAL TPG boards and the DTH400 board to transmit DAQ data from the TPG system.

to be at most 5 μs. Table 8.7 shows the current best estimates of the FE and TPG latency. Some of the biggest contributions to this latency are from the clustering algorithms in Stage 1 and 2. Estimates of the Stage 1 firmware processing time of around 1 μs have been made from the preliminary version of the firmware which is available, but the firmware for Stage 2 is less developed. It is thought that some algorithms for the 3D clustering could be implemented in a similar way to the 2D clustering and so the same processing time for both stages is used in the table.

8.3.6 DAQ readout from TPG

Some of the TPG data will be read for every L1A; these will be used by the HLT to seed higher-level triggers. Here, these data are assumed to be the data sent to the central L1T correlator, i.e. the 3D clusters and the energy map. In addition, other TPG data will be needed for monitoring

Table 8.7: Current FE and TPG latency estimates.

Location	Component	Latency (ns)	Comments
On detector	HGCROC ASIC	400	Includes 250 ns ToT (silicon only)
	Concentrator ASIC	500	Includes 300 ns raw data buffer
	Transmission to TPG	600	≈100 m of FO cables
Off detector	Stage 1 processing	1000	
	Transmission to Stage 2	150	
	Stage 2 time-multiplexing	600	24-BX time-multiplexing
	Stage 2 processing	1000	Large uncertainty
	Transmission to Stage 3	150	May not be needed
	Stage 3 processing	100	May not be needed
	Transmission to central L1T	150	
Total		4650	

and diagnostics of the TPG performance. Such events require the boards to record all the input and output data, as well as intermediate values within the algorithm. The input and output data allow checks of transmission between boards, while all these data allow a verification of the board algorithms at the bit level. At some adjustable rate not exceeding 10% of events, the TPG event size will therefore be much larger than that for the HLT alone. These "fat" events are coordinated across all trigger subsystems, and will dominate the event data rate.

Since the exact algorithms are not fixed, the size and format for these data are not yet defined, although estimates of the data volume can be made. Table 8.8 shows an estimate of the fat event data volume for each Stage 1 board, where the intermediate data value is very uncertain. The total is around 70 kbit per board and hence 7 Mbit for all 96 Stage 1 boards in the TPG.

Table 8.8: Estimate of Stage 1 board event data volumes in fat events.

Source	Links	Data/BX (bits)	Time mplex	Data/board (kbits)
Input	96	224	1	22
Output	2	384	24	18
Intermediate				30
Total				70

Table 8.9 gives the equivalent figures for Stage 2 boards. Since Stage 2 is time-multiplexed, all the data for one event are contained in one board per endcap, so the total event data volume is around 2.6 Mbit. The value of 111 kbit for the output data corresponds to the size of the primitives data for each endcap. Hence the total event data volume for non-fat events is around 220 kbit.

To read data out on a L1A, the latency buffer of up to 500 BX, i.e. 12.5 μs, for these data must be implemented on the TPG boards. However, around 1.5 μs of latency is already used in the FE electronics before the data arrive at the TPG and around 1.0 μs is saved due to not transmitting the L1A to the on-detector electronics, so the buffer needs to be up to 10 μs deep, equivalent to 400 BX. Fat events may be flagged by extra data associated to the L1A and hence not known in advance. This means all the fat event data need to be buffered for every event. As shown above, for the Stage 1 boards, each fat event is around 70 kbit so the latency buffer has to store up to

Table 8.9: Estimate of Stage 2 board event data volumes in fat events.

Source	Links	Data/BX (bits)	Time mplex	Data/board (kbits)
Input	96	384	24	885
Output	12	384	24	111
Intermediate				300
Total				1295

30 Mbit of data. The VU9P FPGA has 270 Mbit of UltraRAM so implementing this memory within the FPGA should be possible. For Stage 2, the data volume per BX is around 1.3 Mbit. However, these data are only received or transmitted every 24 BX, so the effective rate per BX of latency is around 60 kbit. Hence, the larger requirement for the latency buffer size comes from the Stage 1 board.

Assuming 70 kbit/event per Stage 1 board as above, with a 750 kHz event rate and a 10% rate of fat events (and no data sent for non-fat events), the average data rate would be around 5 Gbit/s per board and hence the DTH in a TPG crate with 12 Stage 1 boards would collect a total of around 60 Gbit/s. This can be handled by one less expensive DTH400 board per crate, as shown in Fig. 8.43. For Stage 2, the effective fat event data rate per board is around 60 kbit, and 10 kbit for non-fat events. With a 10% fat event rate, this gives an average size of 15 kbit and so at a 750 kHz event rate would give around 11 Gbit/s per board and hence 140 GBit/s per crate. Again, this can be handled by a DTH400 board. In both cases, the DAQ data rate from both the Stage 1 and Stage 2 boards is well below the 32 Gbit/s bandwidth available from half of the QSFP used. This bandwidth would limit the rate of fat events to around 30% a much higher rate than anticipated.

The average data size from all 96 Stage 1 boards will be around 0.7 Mbit/event with 10% fat events, while the 48 Stage 2 boards will produce around 0.4 Mbit/event, totalling 1.1 Mbit/event, i.e. less than 0.2 MBytes/event. This is an order of magnitude smaller than the DAQ event data volume.

8.3.7 Testing plan

The TPG test plan is described together with the DAQ testing plan in Section 8.2.6.

8.3.8 Summary of R&D required

There is no large hardware development needed for the TPG beyond that described in Section 8.2.7; the main extra effort will be for the development of the different daughterboard needed.

The main TPG R&D areas will be to develop the TPG and central L1T correlator algorithms using simulation to optimise the use of the HGCAL trigger data, and then implement them in firmware. This will give an evaluation of their firmware resource requirements to determine whether they are feasible. As algorithms develop, it may be possible to move to alternative implementations (while still using generic boards) so as to have cheaper and/or more performant TPG architectures.

8.4 Clock distribution

At the HL-LHC, an average of 200 PU events will occur for a beam time dispersion of about 180 ps (RMS) with a 25 ns bunch crossing structure. With these parameters, the probabilities of spatial superposition increase and the particle-flow algorithms begin to lose performance in the endcaps. A precise timing of 30 ps or better should allow better matching of tracks, vertices, and energy deposits in the calorimeter. To achieve this goal, very precise synchronization of the entire detector is necessary. Fast and accurate timing and precise clock distribution are important challenges for the HGCAL electronics.

As shown in Fig. 5.22, the intrinsic resolution of the signal in the silicon sensors can be smaller than 20 ps for large signals. The TDC used for the ToA will have a bin-size of 25 ps. The optimal use of these timing performances requires a more precise time reference, thus an accuracy of the clock distribution of the order of 10–15 ps.

The existing CMS clock distribution provided by the current TTC system is already good in terms of RMS jitter, according to the studies conducted at CERN. The current RF clock delivered by the LHC machine is specified with 9 ps RMS jitter. Its future evolution, notably with the availability of LpGBT and eventually Passive Optical Network (PON) components, holds the promise to achieve even better performance.

The performance of the current TCDS system regarding precision timing is not very well known, and this system will be characterized and measured in order to better define potential improvements in terms of jitter and long term stability.

The option that is being pursued is a tree of encoded clock path. This consists of an encoded clock with a guaranteed available low jitter at the end-points of the distribution tree, but with the uncertainty on the phase between the leaves. R&D will determine if the actual promise of low jitter can be held or, alternatively, what needs to be done at the level of each intermediate stage to clean the clock. Most importantly, the R&D will also determine if controlling the phase differences is possible in the presence of a long chain of encoded clock. If such a control turns out to be impossible, the R&D will also determine if the monitoring of the phase differences would allow achievement of the desired timing precision. If the above turns out to be unfeasible then a solution with a separate clock tree will be worked on.

Concerning clock cleaning, three domains of frequencies that might affect the jitter need to be distinguished. First, there is a high frequency domain above PLLs loop bandwidth (100 kHz to 1 MHz) for which the effects will most probably be cancelled by the PLLs. This is followed by an intermediate frequency domain from 10 Hz up to the PLLs loop bandwidth, for which the negative impact on the phase and jitter will need to be monitored (e.g. digital dual-mixer time difference, DDMDT) and if possible controlled. Finally, a low frequency domain below 10 Hz causes clock drifts which effects either need to be monitored and if possible controlled at the level of a system of hardware/firmware/software or calibrated out by physics.

In case of the presence of intermediate frequency ($>$1 kHz) spurs degrading the clock jitter to an unacceptable level, we may need to investigate the use of radiation hard clock "cleaning" devices in the front-end electronics to be deployed after the future LpGBT ASICs. The jitter will be determined by the last PLL in the chain which is likely to be in the front-end chip itself and by the clock distribution tree inside the chip. Low frequency wander ($<$10 Hz clock phase noise) may remain a worry even if the performance measurement are already found to be at the level of 10 ps for the RMS jitter. As a consequence, the clock phase, once calibrated over the whole detector, will have to be monitored in order to follow its possible variations and get knowledge on its longer term level of stability.

8.4.1 R&D on clock distribution

Precision timing testing and measurements could be performed with the help of a test bench emulating the back-end/front-end chain of the detector together with a clock generator and injector. We have chosen to start the design with existing elements instead of waiting for the availability of near-future or future technology (LpGBT, VLDB+, ATCA). The design of the back-end (BE) emulation includes microTCA crates with AMC13 boards, FC7 boards and Xilinx Kintex-7 FPGA (KCU105 Ultrascale) whereas the design of the front-end (FE) emulation includes a Versatile Link Demo Board (VLDB) and a Multi Mode Versatile Transceiver/Receiver (MMVTRx) and its optical connection to VLDB.

The impact of various parameters on the timing performances are being measured and includes clock wander effect among (distant) branches of the clock distribution tree, optical/electrical transmission length and environmental conditions such as temperature. Figure 8.46 shows one of the setup (two channels with two BE emulation with a PLL) used for the summer 2017 measurement campaign using a Stanford CG635 clock generator, a SI5344 jitter cleaner, a Keysight DSA91204 oscilloscope (12 GHz, 40 Gbps), two Xilinx KC705 evaluation boards for the BE electronics emulation providing different high-speed communication paths where each board communicates with one VLDB (GBTx) board with Multi Mode VTRx for the FE electronics emulation via 10 m Multi Mode optical fibres at room temperature. Figure 8.47 show a phase noise diagram for various loop bandwidth. With the above setup and with an adjusted loop bandwidth PLL, the RMS jitter has been measured to be 8.2 ± 0.1 ps.

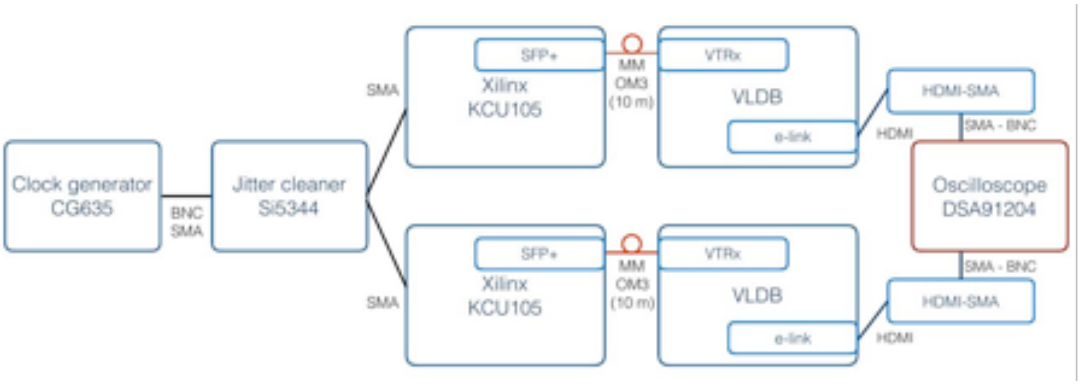


Figure 8.46: Block diagram of the test bench setup used for the clock distribution system measurements.

More measurements are underway using more and more realistic setups, i.e. involving two microTCA crates as well as two FC7 and two AMC13 boards, including longer optical fibres (up to 100 m) and with measurements performed at various temperatures (using a climate chamber) to which the fibres can be exposed and which can have an impact on the timing performances. Furthermore, the design of an IP using a DDMTD to monitor low speed skew drift ("wander") between n channels is under discussion. Plans to prepare tests with ATCA crates as well as with LpGBT, as soon as the chip is available, are also being made.

8.5 Powering: low voltage and high voltage

8.5.1 Low voltage

As explained in Section 8.1.8, DC-DC converters located at the rear of the calorimeter (in the baseline design) or at the periphery of the cassette (option to be studied) will provide 20 A at 1.5–2.0 V to a string of silicon modules or to tileboards. There are 5 280 such strings. However,

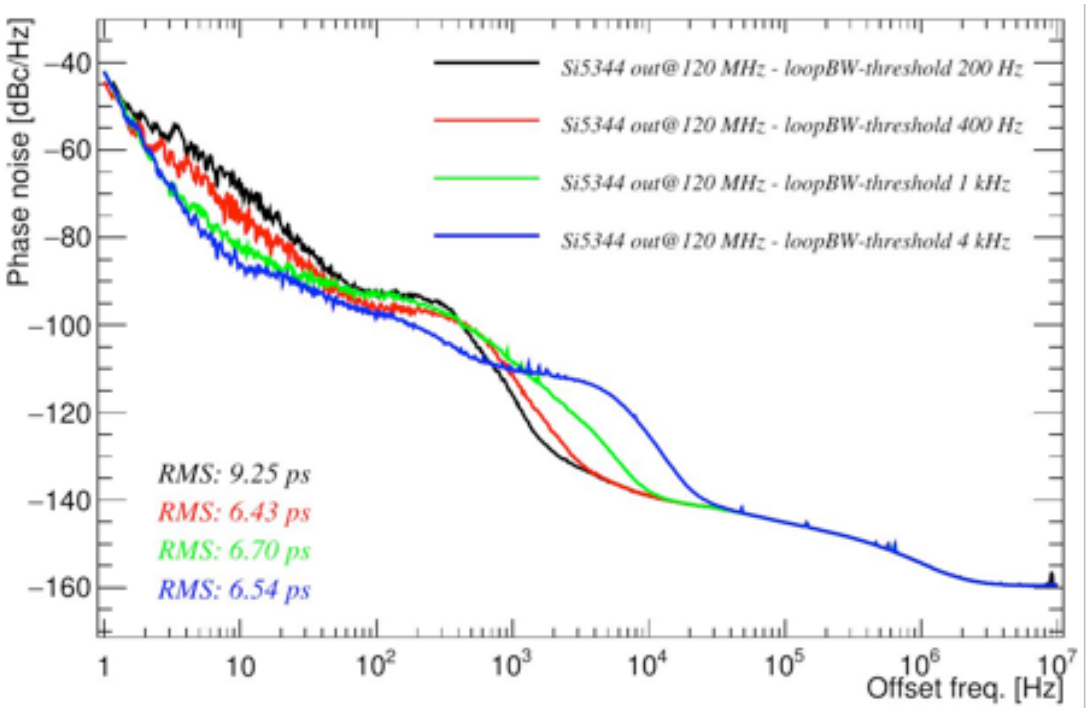


Figure 8.47: Phase noise diagram obtained from the setup of Fig. 8.46.

as clearly seen for example in Fig. 3.8, some of the strings are short and can be grouped. In total, we need ≈3900 DC-DC converters for each voltage (analogue VDDA and digital VDDD). The DC-DC converters are installed in custom-made, water-cooled crates located in 2×18 slots, above the ME0 electronics. Access will therefore be possible during long shutdowns with the opening of the endcaps.

External power supplies provide 12 V at the input of these DC-DC converters. The power consumption of the largest 60° cassette in CE-E is ≈260 W for each of the two voltages for the side which involves trigger sums calculations and readout. This includes:

- the VFE electronics consumption proper,
- the power loss in the LVRs, the motherboards and the internal power cables,
- a conservative assumption on the DC-DC converters efficiency of 75%,
- a 5% power loss in the cables from the supplies to the DC-DC converters, outside the calorimeter.

A similar consumption is expected for the first 10 (single-sided) layers of CE-H. CE-E layers without trigger information require about 20% less power. From layer 38 on, where the area covered by silicon modules drastically reduces, the power consumption for their readout drops below 120 W per 60° sector and below 70 W for the last 12 layers. It is conservatively assumed that 300 W power supplies will be used for all 60° sectors up to layer 37 (included), and 150 W supplies beyond.

The power supplies will be installed in the towers located on both sides of the endcaps, which currently hold the endcaps ECAL and preshower power devices. These power supplies will therefore be accessible during short accesses. For the initial CMS, two suppliers have developed power supplies which are tolerant to the fringe field and radiation levels in this area. Tests are being performed by the suppliers to verify whether these types of supplies will continue performing well with the higher luminosity of HL-LHC, or whether minor improvements

are necessary.

The consumption of a 30° scintillator sector is less than 20 W, so it may be possible to group several sectors on one supply. As the cost of power supplies changes linearly with the power, the choice of grouping is more to do with fault tolerance and optimisation of the cabling.

Cables will run in water-cooled ducts from the power supplies to the DC-DC converter crates. The length of these cables varies from 14 to 22 m. The chosen cable section results from a compromise between the wish to reduce the power loss and the need to minimize their section around the brackets which hold the YE nose. As mentioned above, the cables have been sized to limit the power loss to 5% of the total power in this section.

Figure 8.48 presents a schematic view of the low voltage system for the HGCAL.

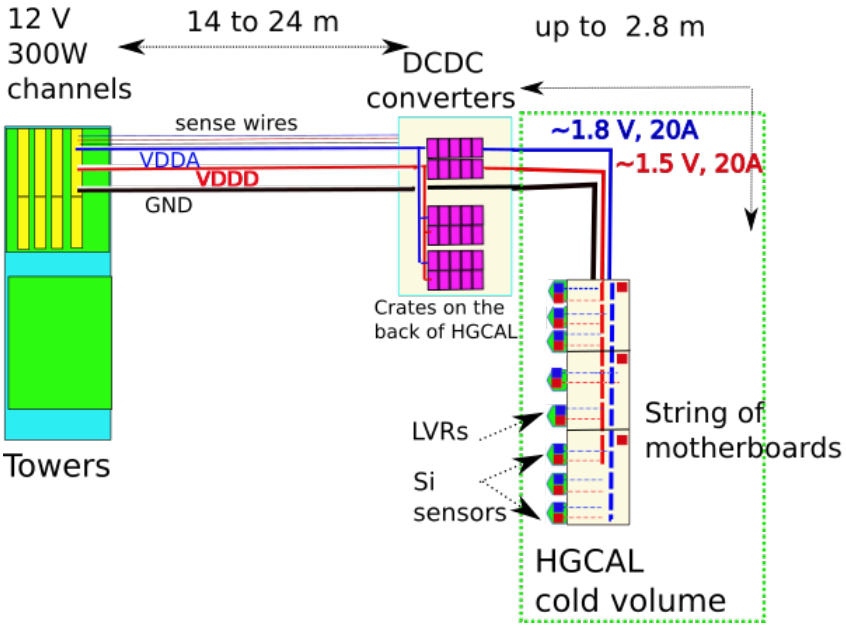


Figure 8.48: Low voltage system for the HGCAL.

CMS is studying the possibility of providing the required 12 V for the DC-DC converters through commercial, low cost and high power (tens of kW) "banks", located in a low radiation area inside the CMS cavern. These banks would be complemented by a radiation tolerant distribution and control hardware. The scheme presented above could be modified to take this possibility into account, positioning the distribution system (instead of the supplies themselves) in the towers. Given the large distance, very thick conductors would be required to link the banks and the distribution boxes. A difficulty to be solved, specific to the endcaps case, is to find a way of routing these thick cables while allowing the movement of the endcaps during the opening/closing sequence of CMS.

8.5.2 Bias voltage

8.5.2.1 Silicon Sensors

Silicon sensors are biased by external power supplies which can be either located in the towers or in the service cavern. The latter option is more likely, given the limited space in the towers for which LV systems have priority.

To ensure the robustness of the detector, each sensor will be biased through an independent

wire. This will allow disconnection of a sensor in case of voltage breakdown. Each bias channel will be therefore connected to a close-by patch panel which will allow the distribution of the the bias and the disconnection of a faulty sensor. A similar scheme has been successfully implemented in the CMS preshower, using 50-ways HV cables. This scheme also has the advantage of offering full flexibility in case there is a need to increase the number of bias supplies, in case the leakage current is higher than expected. Given the very low current per module (a few mA at end of life), very thin conductors (e.g. AWG26) can be used.

For grounding reasons, it is not recommended to connect to the same supply sensors that are coupled to different DC-DCs. Furthermore, it is desirable to bias separately the 100 μm sensors which will draw the largest leakage current and may require, earlier in time, a high bias voltage, due to the higher fluence in this area. With these constraints, the total number of bias channels is 4992. The maximum current to be supplied is $\approx$13 mA at 600 V. We are therefore planning to use bias modules providing 10 W per channel.

It should be noted that using p-on-n wafers instead of n-on-p for the 300 μm sensors would force the use of a different bias polarity for them, increasing the number of channels.

8.5.2.2 SiPM bias voltage supply and regulation

For the SiPM bias voltage and regulation we are planning to use commercial power supplies to deliver 100 V DC, and up to 100 mA, to a set of multi-channel bias voltage regulator boards (MCVRB), which in turn provide a stable, well filtered, bias supply to the tileboards. The MCVRB may be located either just outside the HGCAL cold volume, nearby the DC-DC converters, or may be directly integrated into the motherboards, at the outer periphery of the cassettes. Similar bias voltage regulator boards have been developed and radiation qualified for installation in the region behind the HE as part of the Phase-I HCAL SiPM upgrade, and provide the starting point for the design of the HGCAL MCVRB.

As part of the current conveyor coupling of the SiPM to the HGCROC, a DAC will provide the ability to fine tune the bias voltage, by shifting the DC level of the virtual ground at the HGCROC input by up to 1 V, in order to match the optimal operating voltage for each individual SiPM. The value, and spread, of the optimum operating voltage will be requested as part of the specification. This will allow us to sort and match to ensure a spread of well under 1 V of the optimal operating bias voltage of the SiPM within each board. The scheme is illustrated in Fig. 8.49.

Under the effect of irradiation, the optimal operating bias voltage may shift by up to several hundred mV. It is foreseen to provide at least two independent bias supply lines to each tileboard, in order to account for the variation in fluence over any given set of SiPM on each supply line.

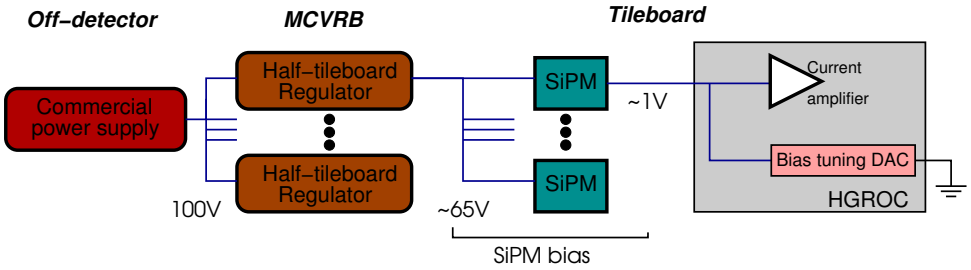


Figure 8.49: Bias voltage supply to the SiPMs.

Chapter 9

Engineering

9.1 Introduction

The overall engineering design of the HGCAL is presented in Chapter 4, including descriptions of the cassettes, structural design of the electromagnetic (CE-E) and hadronic (CE-H) compartments, cooling and environmental control, services, and assembly, installation, and commissioning. This chapter provides additional detailed information about these systems.

9.2 Cassettes

The overall design of the cassettes is summarized in Section 4.2. This section presents additional details concerning the cassette design and cooling calculations, and cassette assembly and testing procedures.

9.2.1 Cassette design details

There are three varieties of cassettes in the HGCAL:

- CE-E cassettes: They are 60° wide, utilize silicon detectors on both sides of the central copper cooling plate, and incorporate stainless steel clad lead absorbers layers into the covers on the two sides. They are an integral part of the overall CE-E mechanical structure. Each endcap contains 84 CE-E cassettes.
- CE-H cassettes: They are 30° wide, utilize active elements on only one side of the cooling plate and have simple copper covers. They are independent of the absorber and mechanical structure. They are manufactured in left-right pairs that are joined together into a 60° unit just prior to insertion into the hadron absorber. There are two sub-varieties of CE-H cassettes:
  - all-silicon;
  - mixed, with silicon modules at smaller radius and scintillator/SiPM tileboards at larger radius.

Table 9.1 summarizes the main characteristics of the three different varieties of cassettes. The stackup of the components that make up each cassette type together with the thicknesses of each cassette type are represented in Fig. 1.5. Additional details about the construction of each cassette variety are given in the following sections.

9.2.1.1 CE-E cassettes

Silicon modules are mounted on both sides of a common cooling plate at identical locations on the two sides and utilizing attachment points at the same locations on the two sides. Figure 9.1

Table 9.1: Characteristics and quantities of the three cassette varieties. For each variety of cassette, the range of numbers of modules, width, radial length, and mass are tabulated. For the CE-E cassettes, the module counts include both sides.

Cassette type	CE-E	CE-H (silicon)	CE-H (mixed)
Active sides	2	1	1
Full Si modules	91–102	26–33	5–19
Partial Si modules	4–13	2–5	1–4
Scint. tile modules	-	-	3–12
Angular width (°)	60 °	30 °	30 °
Linear width (m)	1.56–1.67	0.87–0.97	1.00–1.39
Radial length (m)	1.24–1.32	1.33–1.47	1.54–2.17
Mass (kg)	220–250	56–68	74–144
Total no. in CMS	168	192	384

illustrates the layout of silicon modules in a CE-E cassette. Full hexagonal modules are utilized on the "radial" edges, which can nonetheless be inserted horizontally since alternate facets of the "sawtooth" edge are parallel to the insertion axis. The modules are fixed to the cooling plate by screws at each of the six corners of the modules. The corners of the sensors are truncated (see Fig. 9.2(a)), which provides a triangular gap wherever three modules meet. A single screw is inserted within the concentric 120 ° arc segments at the corners of the module baseplate (see Fig. 9.2(b) and (c)) to secure the corners of the three adjacent modules. A 3 mm diameter hole at the centre of the baseplate interlocks with a pin in the cooling plate to determine the precise position of the module. A second slightly elongated hole determines the angular position of the module and allows for the difference in thermal contraction with cooldown to the operating temperature of $-30\,^{\circ}$C. The screws that secure the module corners are fitted with elastic washers to provide the required clamping force to provide good thermal contact with the cooling plate but also to allow for the relative motion with cooldown (70 μm across the diameter of a module).

In the double-sided CE-E cassettes, the mounting screws on the two sides of the cassette align and interlock as shown in the right side of Fig. 9.1, which is a cross-section through the cassette at that point. It shows: the cooling plate in the middle; modules made of WCu baseplates, Si sensors and PCBs on both sides; and the two covers. Figure 9.3 is a 3D view of the edge of a cassette, which gives another illustration of the concept. A specially designed screw threads into the cooling plate (from the bottom as shown in Fig. 9.1) to secure the modules on one side. It extends $\approx$2 mm beyond the opposite side of the cooling plate and a corresponding "nut" threads onto it to secure the modules on the other side.

The cassette covers, which also serve as half absorber layers, consist of lead sheets clad with thin layers of stainless steel. The side facing the detector is copper coated to provide a high-conductivity grounding/shielding layer. The screws that attach the modules to the cooling plate serve also as spacers to define the gap between the cooling plate and the cover/absorber. Special screw assemblies are used in $\approx$15% of the locations to secure the covers to the cooling plates, in another $\approx$15% of the locations as support points to transmit the weight of the cassettes to the back plate during assembly (Section 9.3.1) and in a final $\approx$15% of the locations the screw assembly penetrates the absorber to serve as lifting points for moving the cassettes with a special lifting fixture (Sections 4.6 and 9.3.1).

Motherboards, as described in Sections 3.1 and 8.1, are one module wide and serve one to five modules. They serve to: aggregate, format, and serialize the data; receive and distribute the fast control signals; select, aggregate, format, and transmit trigger sums; and distribute

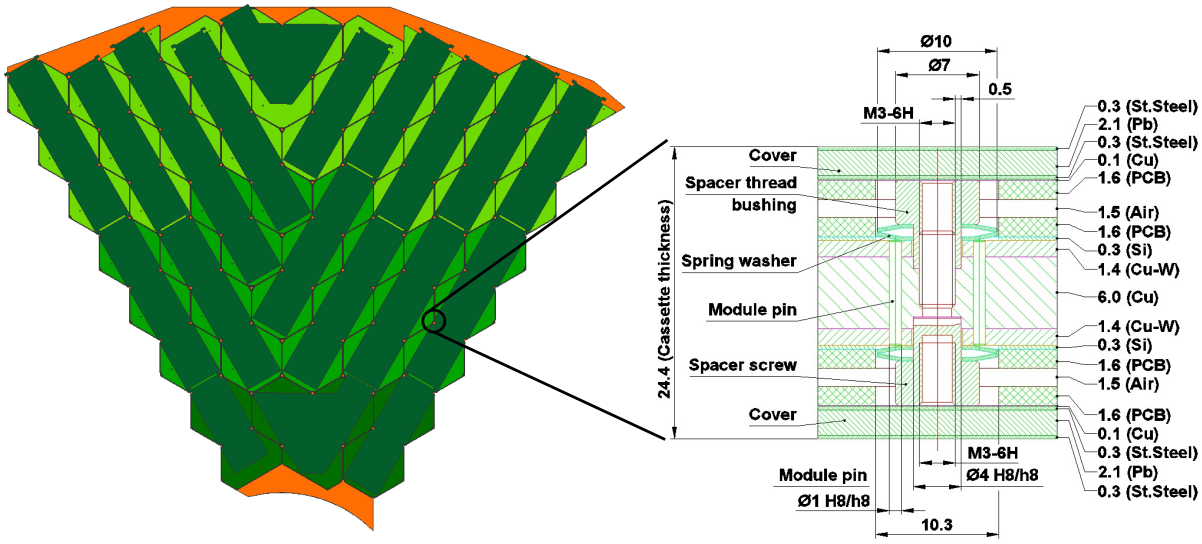
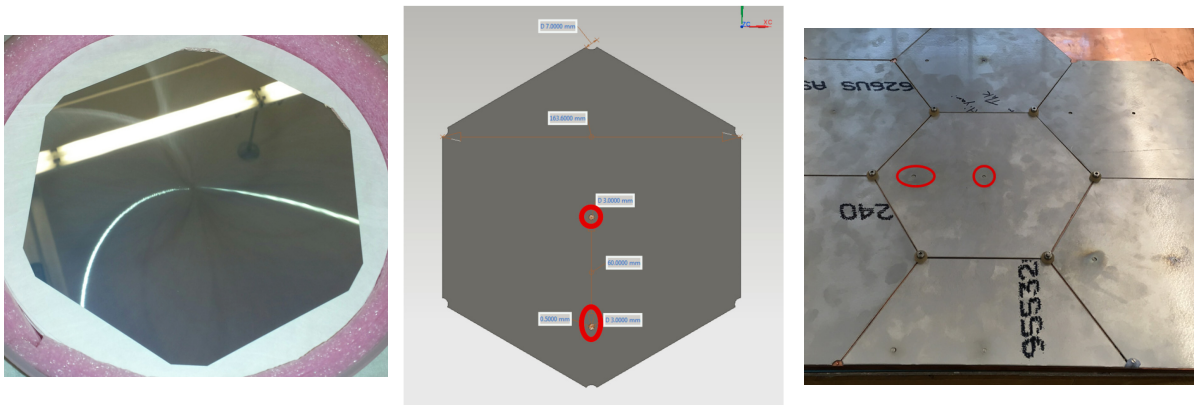


Figure 9.1: Example of cassette layout, with motherboards carrying power and electronics shown as rectangular or triangular elements (left). Fixation of the modules to the baseplate in a CE-E cassette (right), showing a cross-section through the thickness of the cassette at the position where a special screw - nut pair secures the corners of three adjacent modules and defines the spacing between the cooling plate and lead absorber layers.



(a) (b) (c)

Figure 9.2: Key elements of the silicon module mechanical and mounting system as being developed in the cassette prototyping programme: (a) A dummy silicon sensor made to the specified dimensions, illustrating the truncated corners that allow for the mounting of the hexagonal modules. (b) A baseplate, illustrating the 120 ° arc segments that go under the mounting screw at each corner and the pair of alignment pin holes. (c) A set of dummy baseplates mounted on a copper sheet, illustrating the scheme of mounting screws and elastic washers.

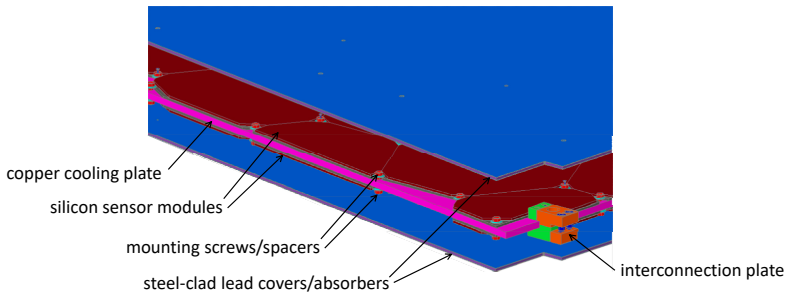


Figure 9.3: 3D model of the outer edge of a CE-E cassette, showing silicon modules mounted on the two sides of the cooling plate and the stainless-steel clad lead covers/absorbers.

power to the modules. Figures 3.3 and 8.20 illustrate the design of a motherboard and its mounting on the module PCB from an electronics standpoint. The mechanical realization is shown in Fig. 9.4(a), which is an exploded diagram of a motherboard and three modules prior to mounting on a cassette. The motherboard is 94 mm wide and N $\times$ 164 mm long, where N is the number of modules linked by one motherboard. The width is set to be no more than that of one side of a hexagonal module, to ensure that all components stay within the physical edge of a cassette along the radial edges, as shown in Fig. 9.4(b). The motherboards are connected to the modules with low-profile Samtec compression connectors illustrated in Fig. 3.3, whose mounting screws serve also to mechanically secure the two circuit boards together. The active components of the module PCB and of the motherboard face each other within the gap formed by the connectors, which is labeled as "air" in the right side of Fig. 9.1. The module plus motherboard assembly has a thickness of 6.4 mm.

Motherboards are radially arranged on a cassette as rows from the inner part to the outer edge. Special triangular variants group the remaining non-aligned modules, as shown in Fig. 9.1. Power and electrical signals are transmitted from one motherboard to the next via connectors on the ends. Data are transmitted via optical fibres from each motherboard to the cassette edge. Prototyping studies for motherboards, modules, and cassettes have been launched and the layout presented here may change in the future depending on their outcome.

The cooling plate, illustrated in Fig. 9.5, is the backbone of the cassette. It provides the surface on which all components are mounted, determines the overall cassette geometry, provides attachment points to join one cassette to the next and to the external support structure, and together with the covers to which it is joined at regular intervals, is the main structural element of the cassette. It also provides the cooling to the silicon modules and electronics. A 4 mm inner diameter thin-wall stainless steel tube embedded in a serpentine groove in the cooling plate carries high-pressure two-phase CO_2 fluid to maintain the detector at a temperature of $-30\,^{\circ}$C. The specific dimensions and layout of the cooling tubes within the cassettes are the subject of on-going engineering design studies, including optimization of the thermal performance (Section 9.2.2) and manufacturability of the cooling plate with its embedded cooling tube. Different options are shown in Figs. 9.5 and 9.9.

9.2.1.2 CE-H cassettes

The design of cassette in the hadronic section of the HGCAL shares many features and concepts with the ones in the electromagnetic section. They are built as pairs of single-sided 30° units, in particular due to limitations in obtaining copper cooling sheets with width larger than 1.5 m as well as difficulties to easily mount modules on such large plates. They are joined together to form a 60° unit, as illustrated in Fig. 9.6 at the time of insertion into the mechanical structure. The cooling plates have the same basic design and serve the same multiple purposes as in the

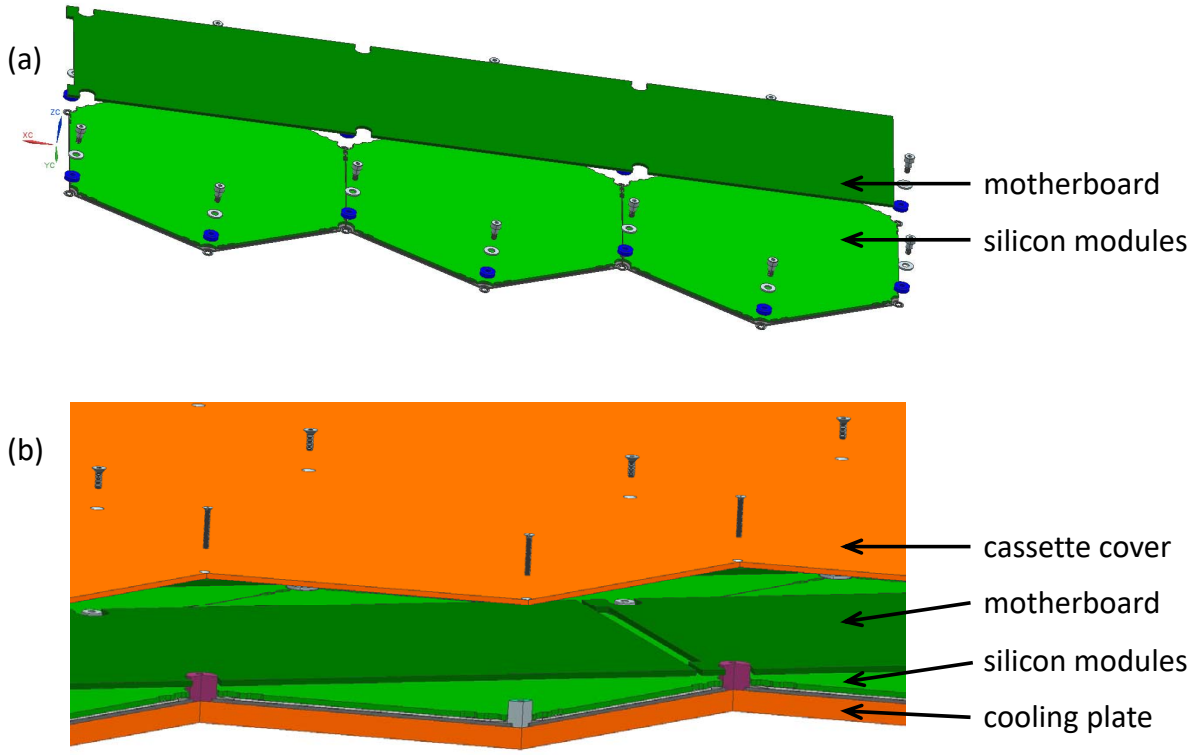


Figure 9.4: (a) Exploded diagram of a motherboard and three silicon modules prior to mounting on a cooling plate. (b) The radial edge of a single-sided cassette showing the relationship between the width of the motherboard and the physical edge of the cassette. Shown also are special mounting screws/spacers needed for the edge modules.

CE-E: they are 6 mm thick, serve as the main mechanical element, and provide cooling via an embedded cooling tube. Attachment of silicon modules to the cooling plate is done in the same way as for the CE-E cassettes and corresponds to the upper part of the right side of Fig. 9.1. Motherboards are laid out just as they are in the CE-E cassettes, except that the left side of the "Vee" arrangement is on the left 30° cassette and the right side on the right cassette.

Starting from Layer 9 of CE-H, cassettes have silicon modules smaller at radius and scintillators/SiPM panels toward the outer edge, as described in Sections 2.3 and 7.3. These are referred to as mixed cassettes. The silicon modules and their motherboards are mounted to the inner region identically as in the silicon-only cassettes. Scintillator tiles and the SiPMs that read them out are mounted on PCBs to form scintillator tile-modules as described in detail in Section 7.3. Each tile module covers a 10° sector in azimuth and one to four tile modules are used to cover the radial extent of each sector, depending on the cassette size and the location of the scintillator-silicon boundary. Two rings of tile modules are needed in the cassettes illustrated in Fig. 9.6, in which the scintillator-silicon boundary is at about the middle of the cassette. A full array of four rings of tile modules is illustrated in Fig. 8.21. The size of a tile module ranges from 33 cm to 47 cm in the radial direction and between about 20 cm to 43 cm in width at the mid-point of each 10° trapezoidal tile-module. The PCB containing the SiPMs is mounted directly on the cooling plate with the scintillator above it, since the SiPMs are the devices whose temperature must be controlled. Twinax and power cables carry signals to and power from the edge of cassette via connectors on the PCB that take the space of about 1/3 of a scintillator tile, representing about 1% of the tile-module area. Each tile-module provides roughly the same functionality as each silicon module. All tile-modules in a cassette are served by a single SiPM

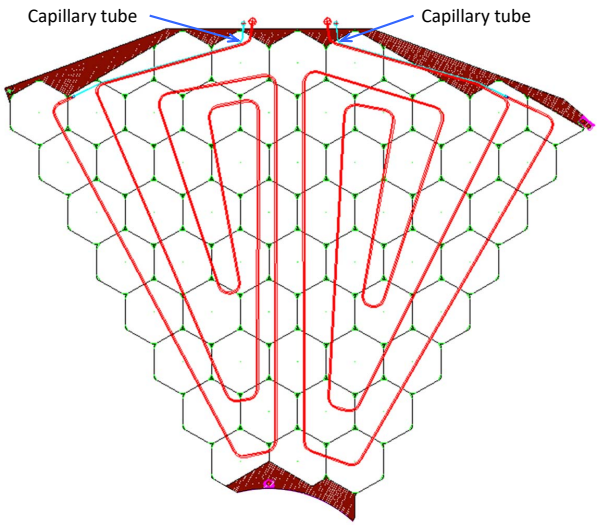


Figure 9.5: Cassette cooling plate showing its main features: mechanical mounting points for the modules and the covers; the features for attaching the cassette to adjacent cassettes and to the external support structure, and the routing of the stainless steel cooling tube. This cooling plate corresponds to the 14th CE-E cassette (detector layers 27–28) and shows the location of the silicon modules of layer 27. The function of the capillary tube is discussed in Section 9.2.2.

motherboard, which is located at the edge of the cassette.

Figure 9.7 is a schematic cross-section through a mixed cassette near the boundary between the silicon and scintillator sections. The tile board electronics are on the under side of the PCB, requiring a localized well of 2 mm depth to be machined into the cooling plate at one location per tile module. Silicon motherboard extensions pass over the top of the scintillator tile-modules to bring power from and signals to and from the cassette edge.

9.2.2 Cassette cooling

The silicon sensors and SiPMs must be cooled to $-30\,^{\circ}\mathrm{C}$, with margin of about $-3\,^{\circ}\mathrm{C}$, to minimize the impact of radiation damage on the performance of the detectors. Active cooling is provided by two-phase CO_2 flowing in a 4 mm inner diameter thin-wall (0.25 mm) stainless steel tube embedded in a 6 mm copper plate used as thermal exchanger, onto which the active elements are mounted. The tubes follow a serpentine path to minimize the temperature difference across the surface of the cassette. The first $\approx$500 mm of the cooling tube is a reduced inner diameter capillary which serves to define the main hydraulic impedance, so as to equalize the flow among different size cooling plates with different heat loads. The precise parameters of the capillary tube are currently being optimized.

The main requirements on the design of the cooling circuit in each cassette are the following:

- Temperature at outlet of cassette cooling tube $> -35\,^{\circ}\mathrm{C}$.
- Maximum operating temperature of any silicon module or SiPM $< -30\,^{\circ}\mathrm{C}$.
- Adequate margin against dryout of the two-phase flow.
- Ability to operate at any temperature between $15\,^{\circ}\mathrm{C}$ and $-30\,^{\circ}\mathrm{C}$, either for testing or in case of elevated dew point.

The overall design of the CO_2 cooling system is presented in Section 9.4.

The cassette with the highest total heat load is the last one in the CE-E, being the largest cas-

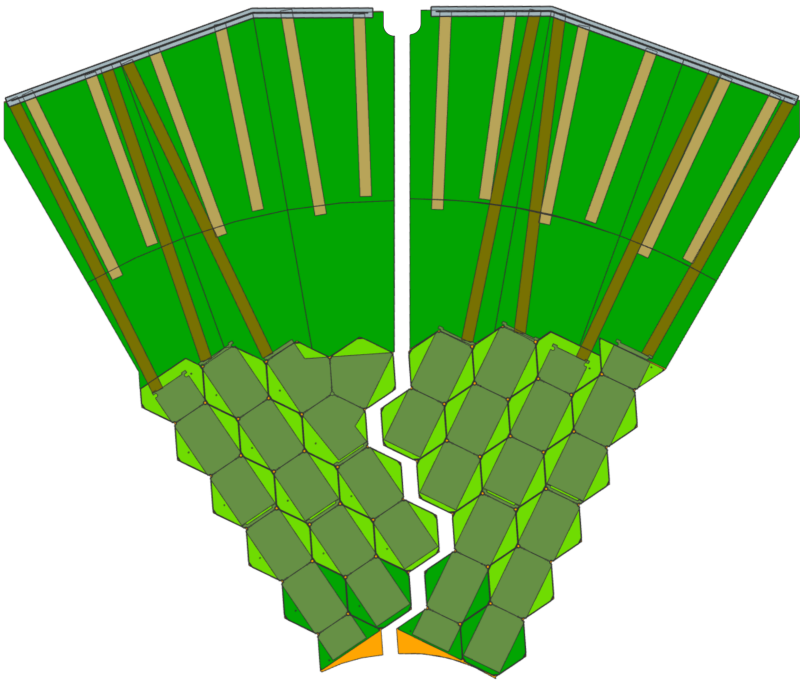


Figure 9.6: A pair of 30° CE-H cassettes with both silicon and scintillator sensors just prior to being joined into a 60° unit for insertion into the absorber. The small cut regions at the outer corners of the cassettes provide clearance with respect to the absorber mechanical structure (Section 9.3.2). This example corresponds to CE-H layer 12.

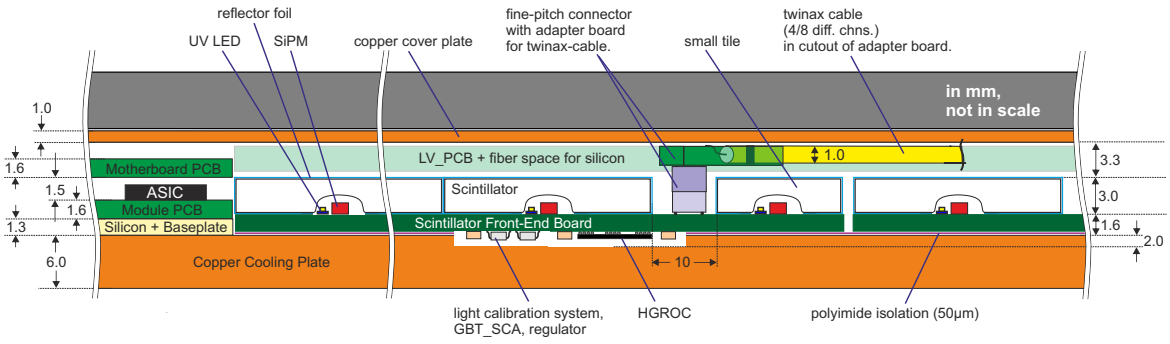


Figure 9.7: Schematic diagram of the assembly of a mixed cassette, with silicon sensor module plus motherboard on the left and SiPM tile board on the right. The thickness of each component layer is indicated.

sette with silicon sensors on both sides of the cooling plate. The layout of the modules on the two sides of the cassette is illustrated in Fig. 9.8. High-density (432 channels for a full hexagon) modules are used toward the inner radius and low-density (192 channels for a full hexagon) are used at larger radius. Two types of partial modules are used to improve the tiling at the inner and outer radii, with 50% and 80% of the area of a full hexagon respectively. The two sides of a CE-E cassette differ in two important respects. First, the outer ring of modules, corresponding roughly to trigger tower 18 of the current detector, are not installed on the downstream side (right side of Fig. 9.8). Second, trigger electronics are not implemented on the downstream side, resulting in lower power dissipation. The power dissipated by each module type is indicated in Fig. 9.8, assuming 20 mW/channel on the upstream side and 16 mW/channel on the downstream side. To this is added an allowance of 20% for resistive losses in the motherboards and losses in the low-voltage regulators. In addition, an estimated leakage current of 10 μA

per channel is assumed after 3000 fb^{-1}, corresponding to 6 mW/channel. The leakage current assumption is conservative, as 10 μA is expected only in the inner-most sensors, but has been assumed for all modules. The total heat load in this cassette is 780 W.

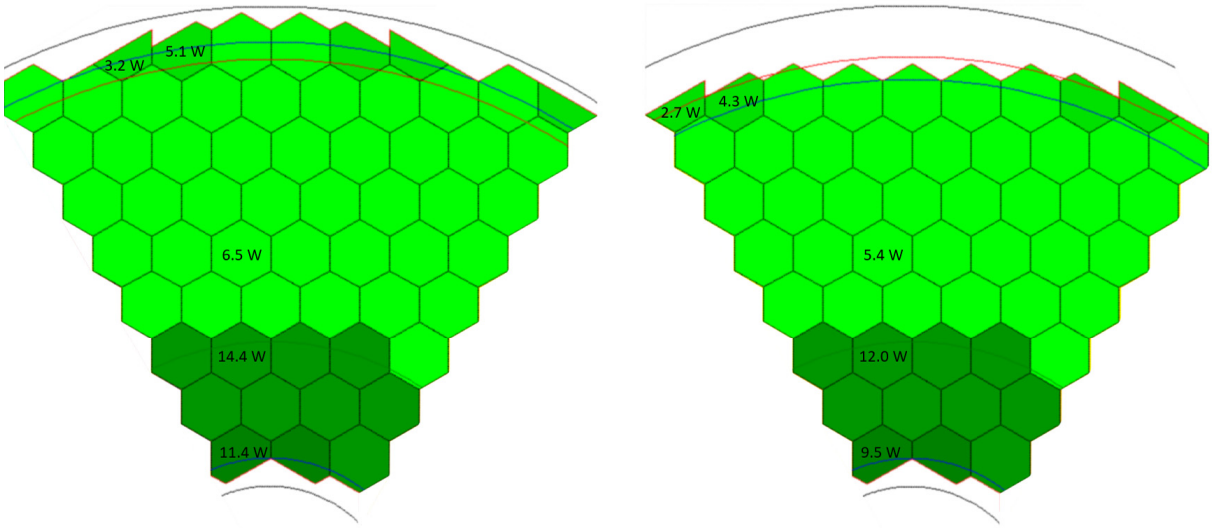


Figure 9.8: Sensor layout for the two sides of the downstream-most CE-E cassette used for the thermal simulations. The dark green sensors at the inner radius are high-density modules (432 channels for a full hexagon) and the light green ones are low-density modules (192 channels for a full hexagon). The values shown are the power dissipated in each type of sensor that is assumed in the thermal calculations.

A preliminary layout of the cooling tubes in this cassette is shown in Fig. 9.9(a). Two cooling loops are utilized to follow the 30° segmentation used in the CE-H. Cooling calculations have been done for this configuration using a cooling plate thickness is 6.35 mm, a cooling tube with an inner diameter of 4.25 mm and a wall thickness of 0.25 mm, and a CO_2 outlet temperature of $-35\,°C$. The dimensions used are small variants of the nominal ones, based on standard American plate thicknesses and tube diameters. The conclusions of this preliminary study are not expected to be significantly affected by these minor differences. The simulated temperature distribution is shown in Fig. 9.9(b). The maximum temperature is $-32.7\,°C$ and there is a 1.8 °C temperature range across the cassette. The temperature difference in the CO_2 fluid between inlet and outlet is 0.4 °C. The vapor quality at the outlet of the cooling tube is 30% (i.e. 30% vapor by mass), providing a comfortable margin against dryout, which occurs at a vapor quality of between 60% and 70%. Very preliminary calculations indicate that a cassette of this configuration can be operated at temperatures as high as $+15\,°C$, but with little margin against dryout.

This configuration satisfies the requirements. However, further optimization is planned, including: increasing the margin for operation at or near room temperature; studying configurations with the two circuits covering different radii rather than different azimuths; and consideration of using a single cooling loop for the full 60° cassette.

9.2.3 Cassette assembly and testing

Two cassette assembly sites are planned, one at CERN and one at Fermilab. Both will have the capability of making cassettes of all types and sizes. The planned production rate is approximately 16 cassettes per month per site, but each site will have the capacity to produce cassettes

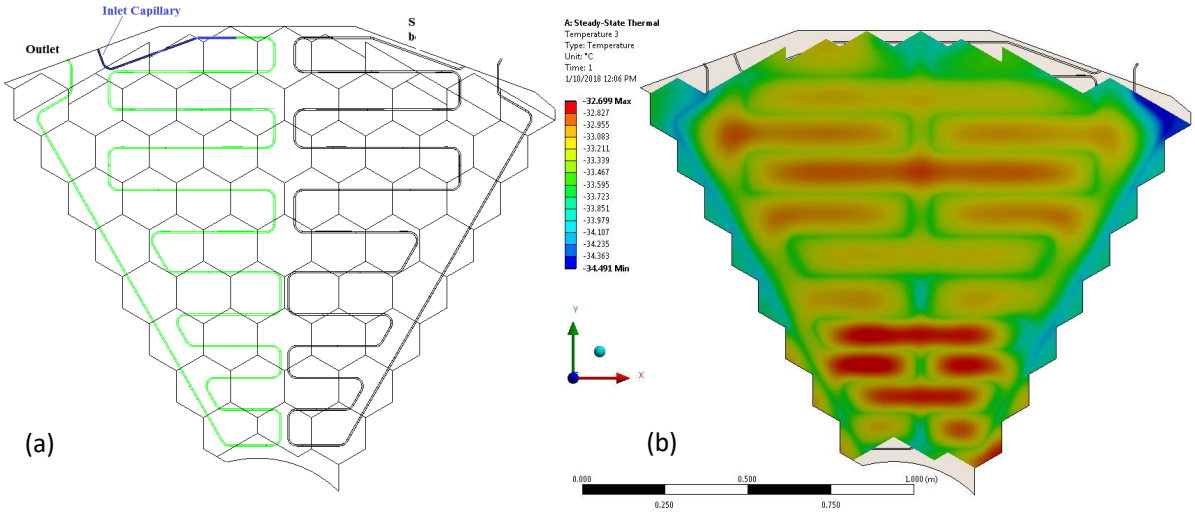


Figure 9.9: Cooling simulation for the downstream-most CE-E cassette: (a) The routing of the two cooling tubes. (b) The computed temperature distribution.

at twice that rate if needed to allow, for example, catching up in case of supply chain interruptions or for one site to cover for another if one experiences an interruption in production.

Cassettes are assembled from the following components. The different varieties of cassettes – CE-E, CE-H all silicon, and CE-H mixed – use different combinations as noted in this list:

- Cooling plate (layer specific);
- Standard and partial silicon modules (number depending on the layer);
- Scintillator tile-modules in CE-H mixed cassettes (number and sizes depending on the layer);
- Silicon module motherboards (number and length depending on the layer);
- Scintillator/SiPM motherboard in CE-H mixed cassettes;
- HV cables and optical fibres;
- Cassette interface (layer specific);
- Fasteners and spacers;
- Covers:
  - CE-E – Lead-stainless steel sandwich on both sides;
  - CE-H – Copper sheet on one side.

The active elements – silicon and scintillator modules and electronics – are received from collaborating institutes where they have been fully tested and qualified. Passive mechanical components are purchased from commercial vendors or obtained from collaborating institutes based on precise specifications, and they are fully inspected and measured for compliance with those specifications as part of the production of the components.

The assembly of the cassettes takes place in a clean room. The sequence of assembly steps is fundamentally the same for all cassette varieties, but not all steps are required for all cassettes. The assembly steps are the following:

1. A kit of components is prepared, specific for the cassette to be assembled. Each component is inspected, the documentation that is provided for it is verified, and component identifiers entered into the production database.

2. The cooling plate is placed on the assembly table.
3. Silicon modules are attached to the cooling plate.
4. Silicon module motherboards are mounted to the silicon modules.
5. Scintillator tile modules are mounted (CE-H mixed cassettes).
6. The scintillator/SiPM motherboard is mounted at the cassette edge (CE-H mixed cassettes).
7. Silicon motherboard extensions are mounted on top of the scintillator tile modules (CE-H mixed cassettes).
8. HV cables and optical fibres are routed from the silicon modules and motherboards to the cassette edge.
9. The cassette interface is mounted and connections made between it and the ends of the silicon motherboards, the HV cables and the optical fibres.
10. The cassette cover (CE-H cassettes) or lead - stainless steel absorber cover (CE-E cassettes) is mounted.
11. The cassettes is turned over and the relevant steps repeated for the second side (CE-E cassettes).
12. A complete electrical test of the finished cassette is performed.

Visual inspections, mechanical measurements, and electrical tests are done at appropriate intervals during the assembly, e.g. after the mounting of all of the silicon modules or after installation of the silicon motherboards, to ensure the proper functioning of the complete assembly and to minimize the time for correction of non-conformities that may be identified.

Completed cassettes undergo a thorough set of cold tests typically in batches of 10 (about two weeks production) in a test cycle lasting about two weeks. Batches of cassettes are placed inside a thermally insulated "cold box" in which the humidity is controlled to give a dew point below $-40\,^{\circ}\mathrm{C}$. The cassettes are cooled by two-phase CO_2 identically as they will be in the HGCAL. All cassettes are attached to low- and high-voltage power supplies of the same specifications as used for the complete HGCAL and connected to a DAQ system that reads out all silicon sensor and SiPM channels as well as temperature instrumentation that is built into the cassettes. Several thermal cycles to $-30\,^{\circ}\mathrm{C}$ are performed, during which both thermal and detector performance are recorded and compared with the requirements. If there are significant deviations in the performance, the offending cassette is removed for further analysis and rework. After the thermal cycles, a week-long run is made at $-30\,^{\circ}\mathrm{C}$ during which cosmic ray data are taken using an external cosmic ray trigger to allow a full evaluation of the response of the cassettes.

Cassettes that pass all tests are prepared for shipment from the cassette assembly facility to the detector assembly facility (CERN Building 904 or SX5 for the CE-E or CE-H, respectively). Cassettes from Fermilab are shipped typically in batches of 10 cassettes in shipping crates with shock-absorbing constraints specially designed for this application.

9.3 Structural design

9.3.1 Electromagnetic calorimeter

As described in Section 4.3.1 and illustrated in Fig. 4.4, 60 ° wide cassettes are joined one to the next by interconnection plates at the outer radius and individually to an interconnecting ring that is concentric with the support cone to form full disks. Fourteen such disks are stacked together to form the complete CE-E. The module fixation screws shown in Fig. 9.1 serve also as spacers to define the gap between the central cooling plate and the covers/absorbers. The detector is assembled in a horizontal orientation (Section 9.6.1) with the cassettes resting on each other and with the spacers aligned in all the layers. If the weight were uniformly distributed over all spacers, the weight to be supported by a typical spacer at the bottom of the stack at the end of the assembly process would be about 29 kg over an area of about 0.3 cm^2. Due to the poor mechanical properties of lead, this weight cannot be reliably supported through the lead cover to the spacers. Therefore, at a subset of the locations the spacers shown in Fig. 9.1 are replaced by ones that penetrate through the lead to provide "hard points" that transmit the load directly through the cassette and not through the lead covers, as illustrated in the inset in Fig. 9.10. These hard points are aligned in all cassettes so as to transmit the weight of the cassette stack through them directly to the back disk, leaving the lead and active detector elements unloaded. This is illustrated schematically in the right side of Fig. 9.10

In the example shown in Fig. 9.10, the hard points are located at about 1 out of every six spacer positions. Preliminary finite element calculations of the stiffness of the cassettes show that with such a spacing of support points, the out-of-plane deflection is < 0.1 mm; hence the hard-point spacers are designed to extend 0.1 mm above the lead surface to ensure well defined contact. Structural calculations show that with the frequency of hard points shown in Fig. 9.10, their diameter must be increased modestly from that of the standard spacers shown in Fig. 9.1. Alternatively, the number of hard points can be increased beyond that shown in Fig. 9.10, taking account of the need to also provide special spacer points that secure the lead covers to the cooling plate and others to provide attachment points for the lifting fixtures needed during cassette assembly and final CE-E assembly. The precise designs, numbers and locations of different spacer types will be the subject of future design optimization.

In the vertical operational orientation, the weight of the CE-E cassette disks is transmitted to a central support cone, which is in turn connected to the back disk of the CE-E, as illustrated in Fig. 4.4 and shown in more detail in Fig. 9.11. The outer periphery of the cassette disks are supported by longitudinal brackets located every 30 °, which serve to define the longitudinal geometry and restrain the cassette stack against horizontal seismic loads.

The support cone assembly also supports the neutron moderator and the endcap timing layer (TE), which are also shown in Fig. 9.11. The TE is placed in its own thermal volume to allow independent cooling of the HGCAL and the TE and, if necessary, to allow installation or maintenance of the TE after the HGCAL is at its operating temperature of -30 °C. To allow for the TE to be warm while the HGCAL is cold, the aluminium support cone ends at the boundary between the CE-E and the moderator. It is extended by a thermally insulating tube made of Permaglas, shown in green in Fig. 9.11. Preliminary thermal analysis for the case of a warm TE and cold HGCAL (including neutron moderator) shows that the Permaglas tube approaches ambient temperature at its end, but that active heating of the tube will be required to ensure no condensation. The design of this area is under active development jointly between the HGCAL and TE, and the details shown here are likely to evolve as the design is better understood.

The support cone is a critical element of the design, whose radial thickness must be minimized

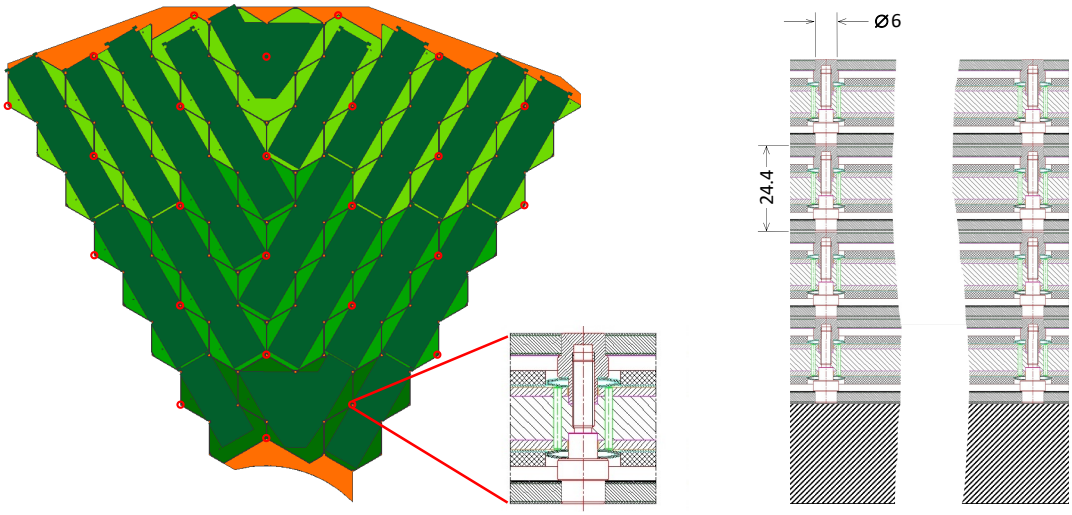


Figure 9.10: Schematic design of hard-point spacers used to transmit the weight of the cassette stack to the back disk without loading either the active detector components or the lead absorber. The red circles on the cassette in the left side of the figure indicate a possible array of these special spacers, whose design is sketched in the inset. The direct transmission of the load to the back plate is illustrated in the right side of the figure.

while providing robust and stable support for the CE-E. The cone assembly is made of a high-strength aluminium alloy. It consists of the cone itself plus a flange that is used to attach it to the back disk. It is machined from a single piece of aluminium to maximize its strength. The conical section is 10 mm thick with machined steps to match the dimensions of the interconnecting rings that define the inner dimensions of the cassette disks. The flange is bolted to the back side of the back disk, as shown in Fig. 9.11, which provides the best support in in the top half of the structure against the bending moment due to the cantilevered load. The flange is further supported by an attachment point to the first absorber disk of the CE-H, which provides additional support in the bottom half against the bending moment.

Finite element analysis (FEA) calculations have been done using simplified constraint conditions on the flange to understand the key issues with the design. These are illustrated in Fig. 9.12. Two models, G1 and G2, are used, which differ in the thickness of the circumferential flange (15 mm or 20 mm), the length of the shoulder behind the flange (15 mm or 34 mm), and the radial thickness of the shoulder behind the flange (13 mm or 21 mm). Different constraint conditions are considered: 3 fixed points at each of 12 azimuthal positions (corresponding roughly to 36 bolts connecting to an infinitely rigid back disk but with no additional support from the first CE-H absorber disk); the entire back surface of the flange is fixed (corresponding to a bonded connection of the flange to an infinitely rigid back disk); or the entire back surface of the flange plus the radially outward surface of the shoulder are fixed. Calculations have been done with cone thicknesses of 10 mm and 12 mm. The cone material assumed in this study is Al Cast Plate G.AL C330R with the following mechanical properties:

- Yield strength $\sigma_{0.2}$ = 290–340 MPa
- Ultimate tensile strength = 320–380 MPa

The input parameters and conditions of the different cases and the resultant maximum deflections and stresses are summaried in Table 9.2.

Figure 9.13 compares the stress distributions for two extreme cases: Model G1 with the 36 point

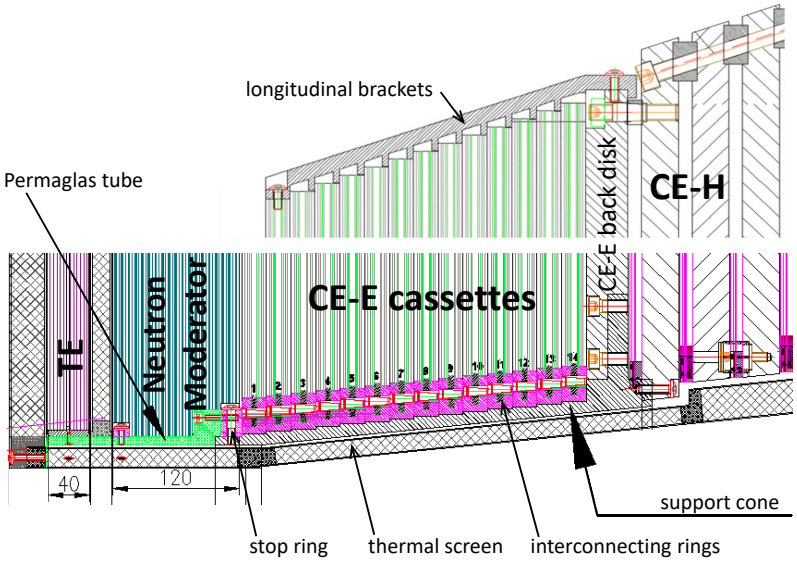


Figure 9.11: Support of the CE-E cassettes at the inner radius by the support cone and at the outer radius by longitudinal brackets that connect to the CE-E back plate. The connection of the CE-E to the first absorber disk of the CE-H is also shown. In this illustration, the outer radius of the 1st (left-most) cassette is reduced to allow space for services for the Endcap Timing Layer and it is therefore not visible in the upper half figure. The cross-hatched bands below the support cone and on either side of the TE represent the thermal screen.

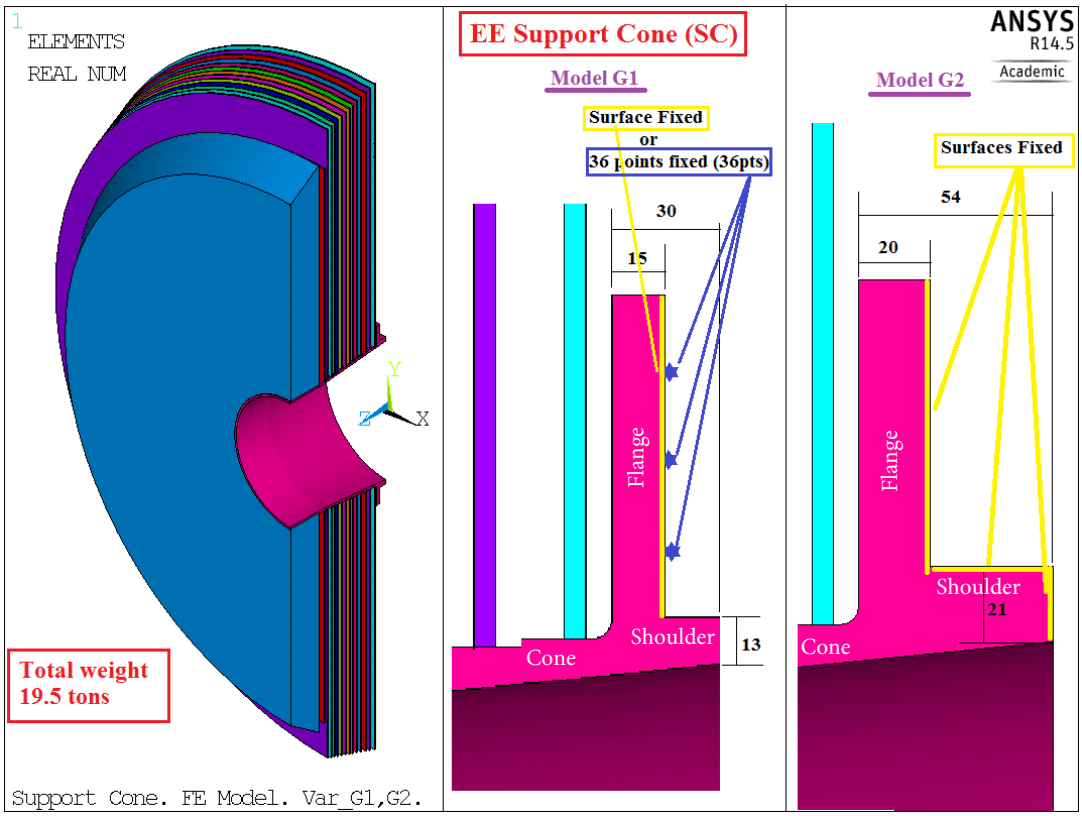


Figure 9.12: Finite element models used to understand key points in the support cone design.

Table 9.2: Summary of input parameters and conditions and computed deflections and stresses from support cone FEA calculations. Models G1 and G2 are defined in Fig. 9.12.

Case	G1, 36p, t12	G1, fix, t12	G1, 36p, t10	G1, fix, t10	G2, fix, t10
Flange thickness	15 mm	15 mm	15 mm	15 mm	20 mm
Shoulder thickness	13 mm	13 mm	13 mm	13 mm	21 mm
Shoulder length	15 mm	15 mm	15 mm	15 mm	34 mm
Cone thickness	12 mm	12 mm	10 mm	10 mm	10 mm
Constraint	36 points	full flange	36 points	full flange	full flange and shoulder
Max vert. displ.	0.27 mm	0.19 mm	0.31 mm	0.22 mm	0.22 mm
Max stress	215 MPa	51 MPa	220 MPa	56 MPa	56 MPa
Safety factor	1.35	5.69	1.32	5.18	5.37

constraint (36p) and Model G2 with the fully constrained flange (fix), both with a 10 mm thick cone (t10). As expected, the high stress point in both cases is at the corner where the flange joins the cone. The maximum stress varies by a factor of 4 between the two constraint cases. The results from several different FEA cases are summarized in Table 9.2. The conclusions from this study are:

- The maximum stress is relatively insensitive to cone thickness, e.g. 51 vs. 56 MPa for (G1, fix, t12) vs. (G1, fix, t10)
- The maximum stress is relatively insensitive to the details of the flange and back shoulder thicknesses, e.g. 56 vs. 54 MPa for (G1, fix, t10) vs. (G2, fix, t10)
- The peak stress depends strongly on the boundary conditions imposed, e.g. 215 vs. 51 MPa for (G1, 36p, t12) vs. (G1, fix, t12).

The safety factor in the worst case (G2, 36p, t10) is unacceptably low. However, the actual constraints shown in Fig. 9.11 are stronger than those assumed in this case. The safety factor in the best case (G1, Fix, t12) is comfortably large. However, the boundary conditions assumed in this model cannot be achieved in practice. It is clear that a cone of this type with a realistic mounting scheme can be designed to have adequate strength, but more detailed FEA calculations must be done with realistic constraints to determine the final design.

Preliminary FEA calculations for the Permaglas support tube show that for a 6 mm thick cone and the expected weight of the moderator and TE the deflection of the tube is very small ($\approx$0.1 mm).

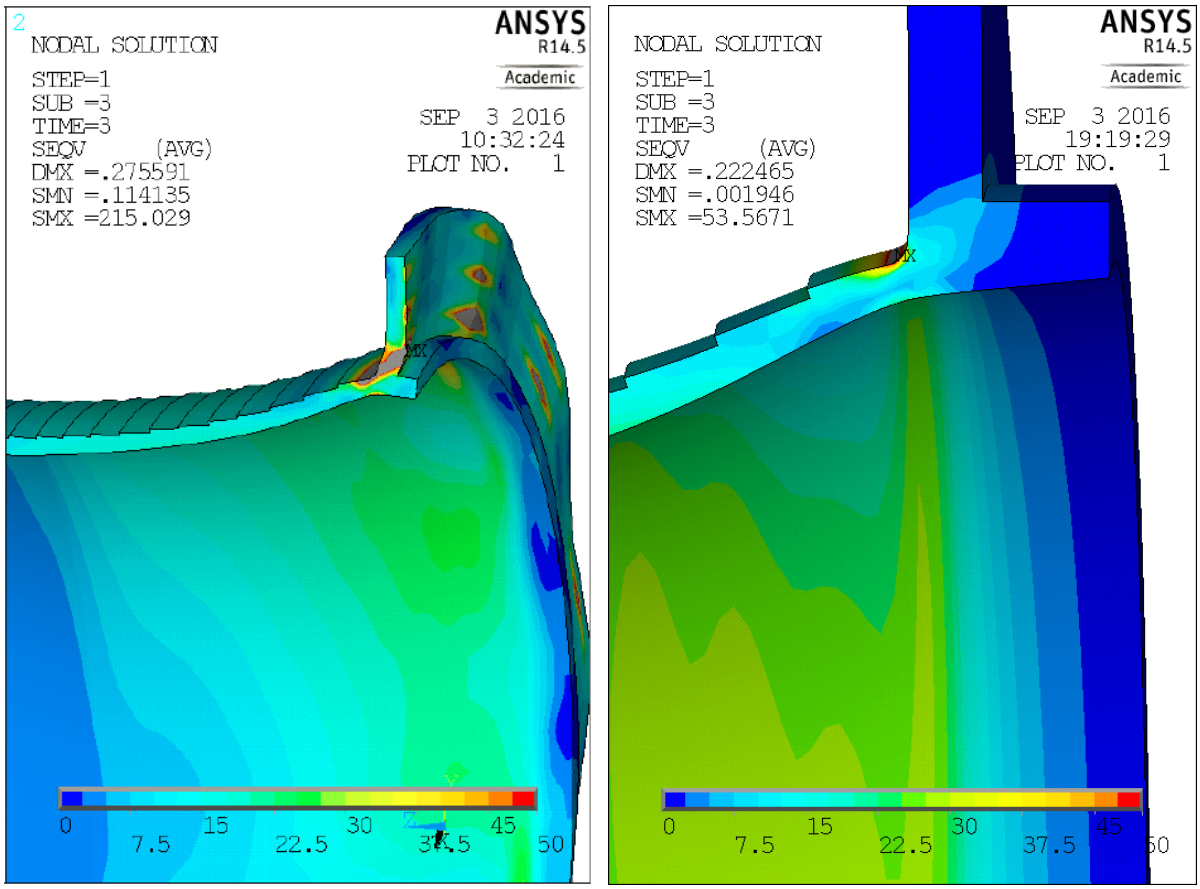


Figure 9.13: 3D finite-element calculation of the stresses in the support cone with the CE-E in the vertical installed orientation for two different constraint conditions on the flange.

9.3.2 Hadronic calorimeter

The structural design of the hadronic section of the HGCAL is summarized in Section 4.3.2. This section presents additional design details as well as structural calculations for key elements.

The hadron calorimeter section of the HGCAL (CE-H) is constructed from stainless steel absorber disks with gaps for the insertion of cassettes. Figure 9.14 shows the mechanical structure of the hadron calorimeter (CE-H) absorber. It consists of 11 disks of 35 mm thickness (labeled FH), 12 disks of 68 mm thickness (labeled BH) and a 94 mm thick back disk. The 40 mm thick back disk of the CE-E (Section 9.3.1) serves as the first absorber layer for the CE-H. The inner periphery of each disk is circular with a radius corresponding to the $\eta = 3$ cone. The outer periphery follows the 20° polygonal shape of the barrel calorimeter. Each disk is made by joining several smaller sectors, the number and size of which will be the subject of design and manufacturing optimization. The right half of Fig. 9.14 shows an example of assembly from four 90° segments. Spaces between the disks accommodate the insertion of detector cassettes. In first eight layers, a gap of 14.0 mm is provided for silicon-detector cassettes; the remaining layers use mixed silicon / scintillator cassettes that require a gap of 15.9 mm. Sets of 12 spacers and bolts or tie rods around the outer and inner peripheries define the cassette slot thickness and support both the weight and cantilevered moment of the structure. Figure 9.15 illustrates the vertical load and bending moment that the spacer/bolt system must support. Table 9.3 summarizes the parameters of the CE-H absorber.

Table 9.3: CE-H absorber parameters for representative layers. The absorber layer number corresponds to the preceding detector layer number. The outer radius is given at the apex of the polygonal structure. The mass is for the absorber only and does not include the cassette mass.

Layer	Thickness (mm)	r_{inner} (mm)	r_{outer} (mm)	mass (kg)
FH 1	35	360	1709	2372
FH 8	35	389	1843	2758
FH 12	35	409	2114	3660
BH 1	68	416	2205	7750
BH 6	68	458	2706	11 756
BH 12	68	508	2706	11 674
Back Disk	94	518	2634	15 397
Full CE-H	1647	360	2706	175 861

The spacers are set into precision bored wells in the absorbers, as illustrated in Fig. 9.16 and therefore serve as shear keys to support the weight of adjacent absorber layers. Bolts or tie rods pass through the absorber layers and the spacers both to clamp the spacers in place. The cantilevered support results in a large bending moment, which is taken by the bolts and tie rods in tension over the upper part of the structure and by the spacers in compression toward the bottom. The most heavily loaded spacers and bolts are those connecting the last absorber to the back disk, and therefore these are the largest; the spacers can become progressively smaller in layers farther from the back disk.

Several types of spacers are utilized, as illustrated in Figs. 9.14 and 9.17:

- 2 permanently-installed round spacers at 90° and 270° at the outer periphery.
- 4 permanently-installed horizontal bar-shaped spacers at 30°, 150°, 210°, and 330°

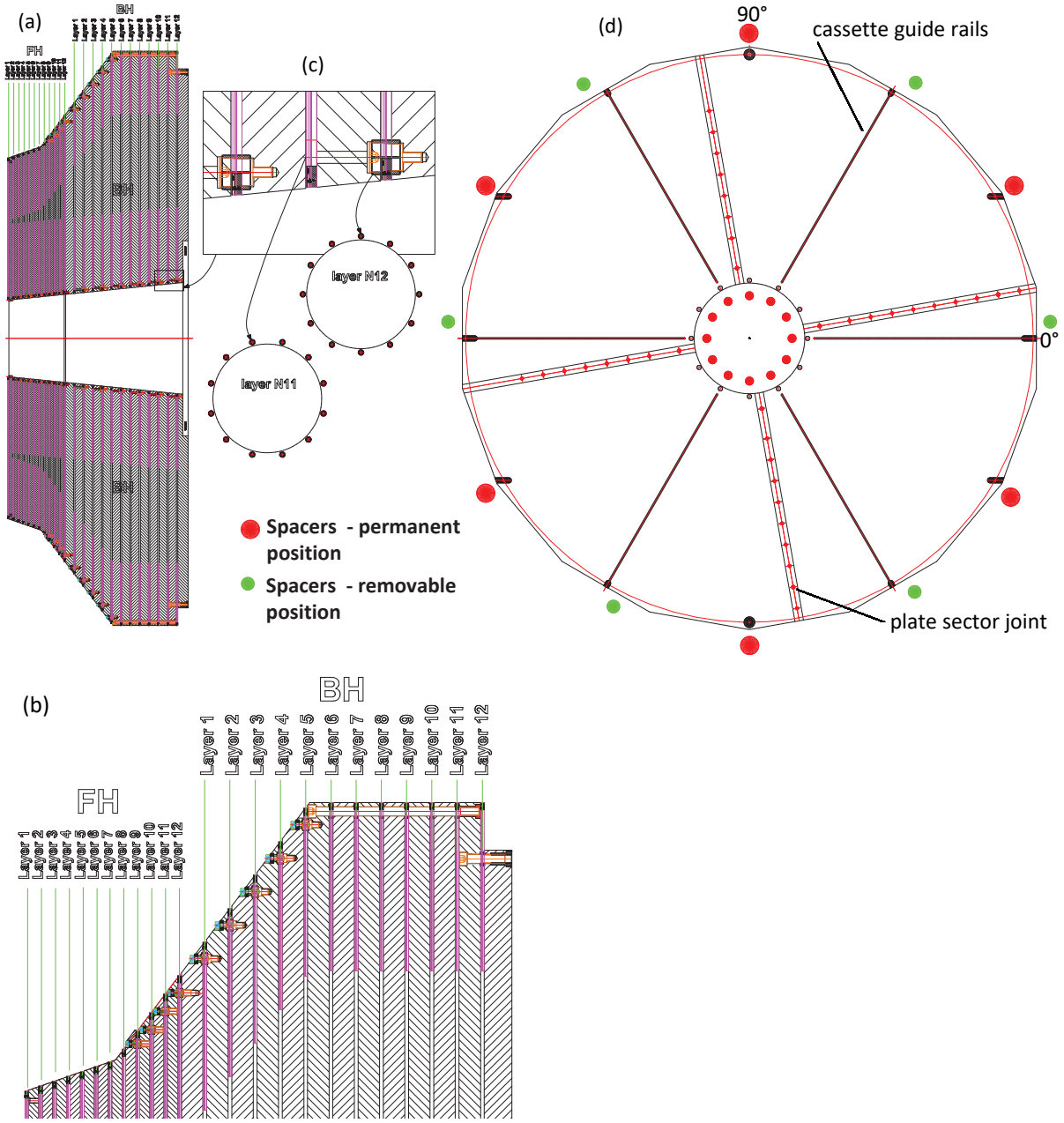


Figure 9.14: Mechanical structure of the hadronic calorimeter absorber: (a) Longitudinal section of the overall absorber. (b) Detail showing the individual layers and the outer spacer-bolt. (c) Detail of the inner spacer system. (d) Transverse view of one BH absorber layer showing the spacer structure. In (d), an example is given in which the plate is assembled from four 90 ° wedges bolted together along the four radial lines. The black lines every 60 ° are channels used to guide the insertion of the 60 ° cassette pairs.

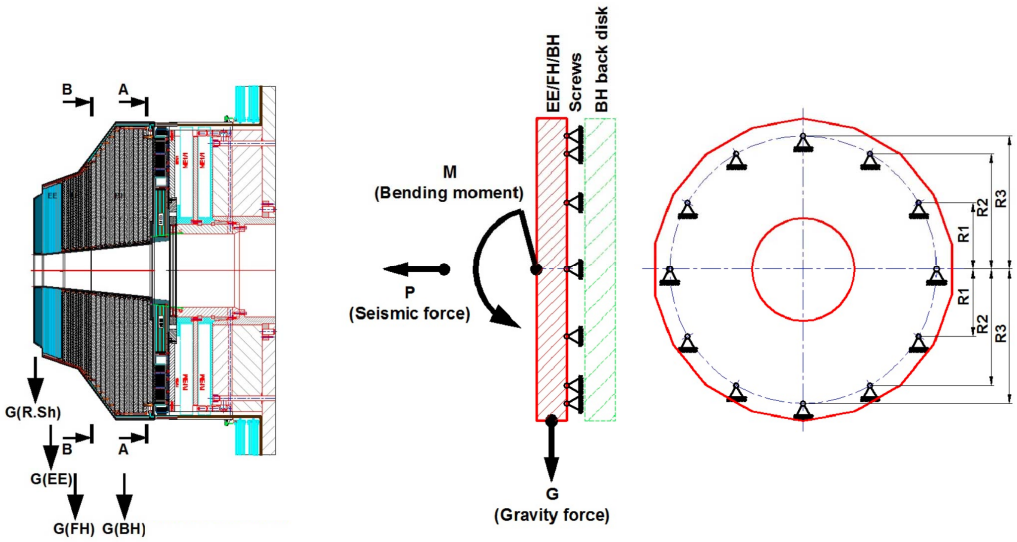


Figure 9.15: Model used for calculation of stresses in the spacers and bolts / tie rods that support the CE-H absorber.

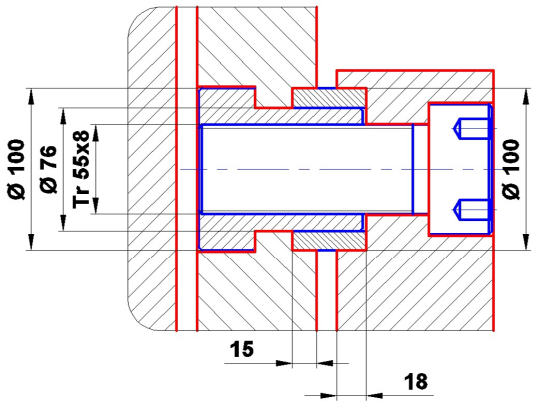


Figure 9.16: Detail of the outer spacer and connecting screw that secures the last absorber layer of the CE-H to the back disk.

at the outer periphery.

- 2 removable horizontal bar-shaped spacers at 0 ° and 180 ° at the outer periphery.
- 4 removable radial bar-shaped spacers at 60 °, 120 °, 240 °, and 300 ° at the outer periphery.
- 12 permanently-installed round spacers at the inner periphery.

The horizontal bar-shaped spacers provide additional surface area to support the vertical load, while minimizing the azimuthal extent of the spacers. The removable spacers can be taken out at individual azimuthal locations during detector assembly while the absorber stack is in a horizontal orientation (Section 9.6) to permit the insertion of 60 ° wide cassette pairs (Section 9.2.1). The 18 permanently-installed spacers and the two removable spacers at 0 ° and 180 ° at the outer periphery participate in supporting the vertical load. All 12 bolt/spacer sets at the outer periphery participate in balancing the bending moment from the cantilevered structure. Individual bolts connect the last CE-H absorber to the back disk, and in the layers in the 53 ° crack region (CE-H layers 9–17). Tie rods connect layers in the cylindrical section (layers 18–24) and the 19 ° region (layers 1–8). See Fig. 9.14.

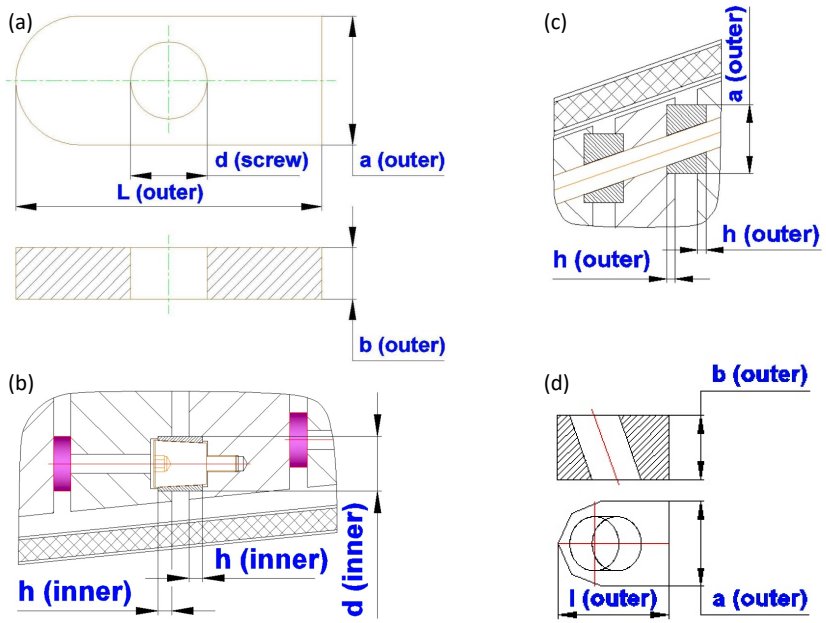


Figure 9.17: Spacer types used in to support the CE-H absorber disks and to define the gaps for the active sensors: (a) bar-shaped spacers used at the outer radius; (b) pin-collets used at the inner radius; (c) and (d) round spacers with tilted tie rods used at the outer radius in the 19 ° crack region.

The spacer and bolt dimensions are set balancing two competing requirements:

- The cross-sectional area must be large enough to support the load with adequate safety factors in the spacers and the absorber (vertical load) and in the bolts (moment);
- The transverse dimensions must be minimized to allow maximum coverage by the active detectors.

Calculations have been done to set the dimensions of the spacers and bolts at the connection of the last CE-H absorber (BH layer 12 in Fig. 9.14) to the back disk and at the boundary between the 35 mm and 68 mm thick absorber plates (FH layer 12 and BH layer 1 in Fig. 9.14). These

dimensions are summarized in Table 9.4. The spacers and absorber plates are assumed to be made of non-magnetic stainless steel (grade A2) and the bolts from a special high-strength non-magnetic stainless steel 16Cr16H3MoNCu - Russian GOST. The material properties assumed for the strength calculations are given in Table 9.5

Table 9.4: Specifications of spacers and bolts supporting the CE-H absorber structure.

	Spacer Type	**Spacer Width (mm)**	**Spacer Height (mm)**	**Spacer Depth (mm)**	**Bolt Type**
CE-H24 – Back Disk					
Inner	Round	40	40	15	Pin Collet
Outer 0 °, 30 °, 150 °, 180 °, 210 °, 330 °	Horiz. Bar	140	60	15	Tr 48x8
Outer 90 °, 270 °	Round	100	100	15	Tr 55x8
Outer 60 °, 120 °, 240 °, 300 °	Radial Bar	140	60	15	Tr 55x8
CE-H12 – CE-H13					
Inner	Round	30	30	10	Pin Collet
Outer 0 °, 30 °, 150 °, 180 °, 210 °, 330 °	Horiz. Bar	45	45	10	Tr 30x5
Outer 90 °, 270 °	Round	50	50	10	Tr 30x5
Outer 60 °, 120 °, 240 °, 300 °	Radial Bar	45	45	10	Tr 30x5

Table 9.5: Material properties of absorber plates, spacers and bolts.

Property	**Absorber & Spacers**	**Bolts**
Yield strength ($\sigma_{0.2}$, MPa)	205–310	680–870
Tensile strength (σ_u, MPa)	485–590	1080–1270
Shear strength (τ, MPa)	160	460
Allowable compressive stress, $\sigma_{\text{comp}} = 0.8 \times \sigma_{0.2}$	164–248	544–696

The calculations were done with a simplified model, illustrated in Fig. 9.15 in which the endcap elements on either side of a spacer-bolt array are treated as rigid bodies. The weight of the endcap portion towards the interaction point is supported by the spacers and the rotational moment is calculated about the beam axis. For the CE-H24 to Back Disk interface, a weight of 2.2 MN and a moment of 2.2 MN $\times$ 0.8 m are used; the corresponding values for the CE-H12 to CE-H13 interface are 0.6 MN and 0.6 MN $\times$ 0.5 m respectively.

The results of the calculations are summarized in Table 9.6. These simple calculations yield safety factors ranging between 2 and 4 and they demonstrate the feasibility of this structure. The safety factors could be increased in several ways, for example: Increase dimensions of spacers and bolts; choose higher strength stainless steel alloy for the spacers and absorbers; take into account the friction force between the spacers and absorbers under bolt / tie rod pressure; increase the number of (removable) spacers to, for example, one every 20 °. The spacer and bolt dimensions are recommended minimum values and represent the starting point for

more detailed calculations, which will likely include FEA models that take account of the elastic behaviour of the absorber disks, relaxing the assumption of rigid body mechanics and allowing a more realistic evaluation of how the support structure takes the cantilevered bending moment.

Table 9.6: CE-H spacer and bolt strength calculation results.

		CE-H24 – Back Disk	CE-H12 – CE-H13
Spacers	Shear stress (MPa)	43	54
	Safety factor	3.7	3.0
	Compressive stress (MPa)	125	118
	Safety factor	2.0	2.1
Bolts / Tie rods	Tensile stress (MPa)	271	226
	Safety factor	2.5	3.0

9.3.3 Endcap suspension system

The endcap calorimeter is supported in a cantilevered fashion from the nose of YE1 as shown in Fig. 9.18. The last absorber disk of the CE-H is connected through the last set of spacers to the back disk as described in the previous section. The back disk is connected through a set of wedge-shaped blocks to a similar disk called the back flange, which then transfers the weight of the endcap to a large diameter steel support tube connected to the YE1 disk. The wedge-shaped blocks are positioned every 20°, as shown in on the right side of Fig. 9.18. They serve to a) transfer both the vertical load and rotational moment of the HGCAL to the back flange; b) provide thermal isolation between the CE, which operates at $-30\,^{\circ}$C, and the room temperature back flange; c) provide sliding surfaces that allow for the approximately 2 mm radius reduction of the HGCAL when cooled to $-30\,^{\circ}$C; and d) define the space between the back disk and back flange that is occupied by the ME0 and electrical services for the CE. The detailed structure of the wedge supports is shown in Fig. 9.19. The bending moment of the HGCAL about the end of the support tube is taken by a set of special brackets that connect the outer radius of the back flange to the YE1 while allowing access to the ME1/1 chambers, as shown in Fig. 9.20.

When the solenoid is energized, the YE1 structure is pulled toward the centre of CMS. The magnetic force is restrained by a set of "Z" stop blocks located at a radius of approximately 4.6 m, as shown in Fig. 9.21(a). Measurements of the deflections with the existing endcap, shown on Fig. 9.21, show that there is approximately 2 mm relative motion between the brackets (13 mm) and the support tube (15 mm). Therefore a flexible connection is required between the back flange and the support tube to accommodate this relative motion, which still supports the full weight of the endcap. This is accomplished by "hanging" the endcap from a flexible ring structure, indicated by the heavy red line in Fig. 9.21(a) and shown in detail in Fig. 9.21(b). The inner part of the ring fits around the support tube and is longitudinally locked to it by a keyed connection at the top. The inner ring is split into two thin longitudinal sections, as shown in the left part of Fig. 9.21(b), whose bending provides the required longitudinal flexibility. Preliminary FEA calculations have been done that show that even for a 6 mm relative deflection, the stresses in the ring, which are largest where the bottom of the inner ring joins the outer ring, are within acceptable limits (safety factor $\geq$3) and that the force that it imposes on the outer bracket system is small compared with the total force.

The outer bracket system is a critical part of the suspension system. It must

- define the longitudinal location of the endcap;

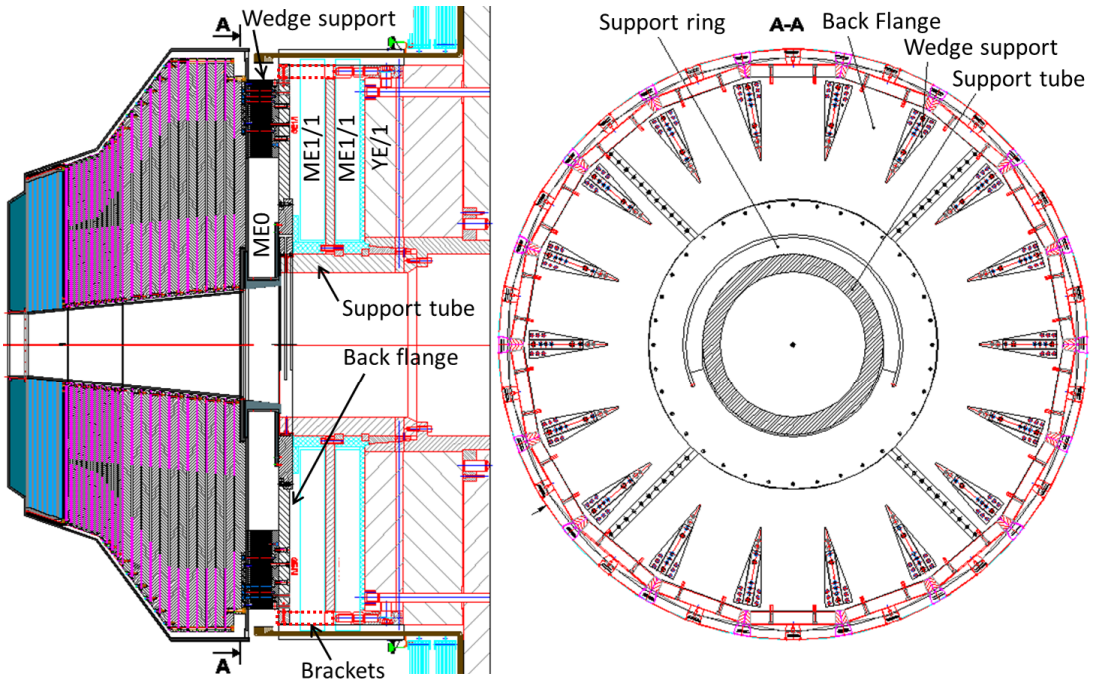


Figure 9.18: Endcap suspension system.

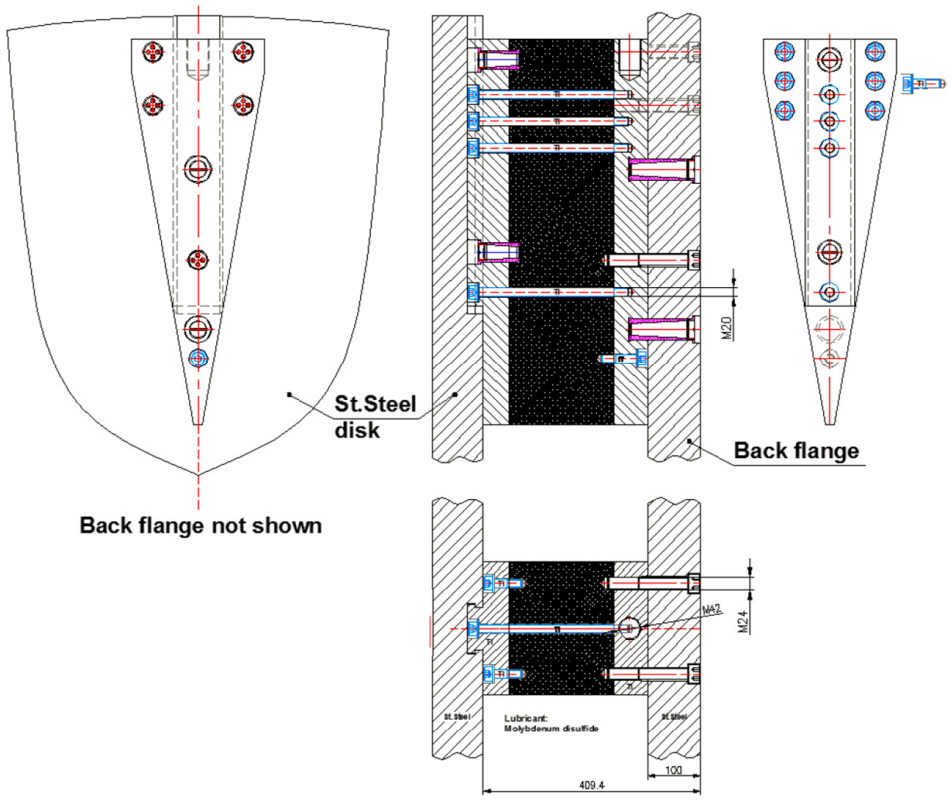


Figure 9.19: Sliding wedge support that connects the endcap calorimeter to the back flange.

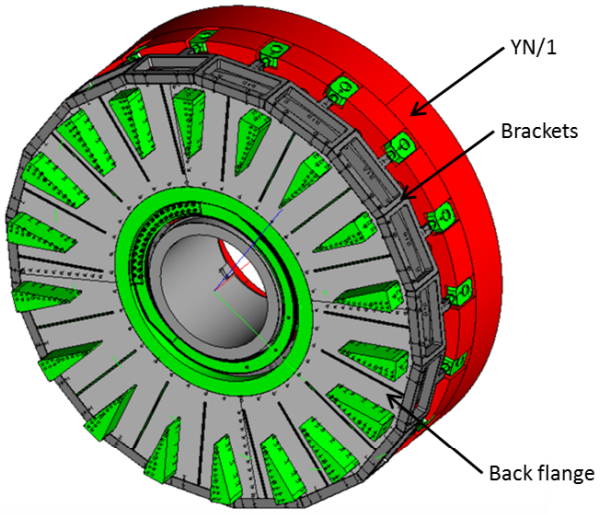


Figure 9.20: The bracket system that carries the cantilevered moment of the endcap calorimeter from the back flange behind the wedge system to the YE1.

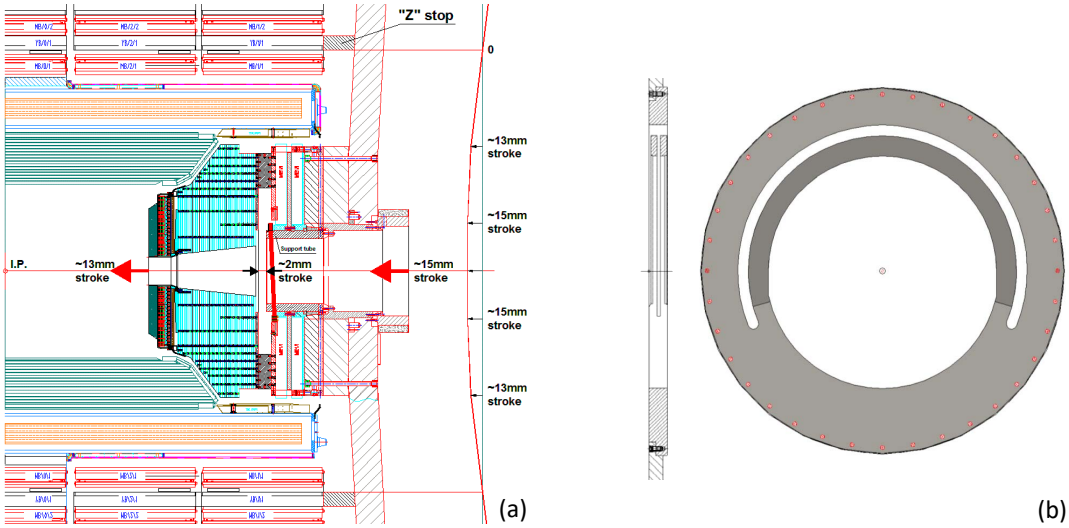


Figure 9.21: Flexible support ring system for attachment of the endcap to the support tube on the YE1 nose. The left side (a) shows the measured deflection of the existing structure under the magnetic field of the solenoid. The right side (b) is a detail of the flexible ring system.

- define the vertical and horizontal angles with respect to the beam axis;
- take the rotational moment of the endcap about the support tube;
- take the longitudinal forces imposed by the flexible ring and eddy currents during magnet excitation and de-excitation.

It must serve these functions while also providing access to the ME1/1 chambers, which are located between the HGCAL back flange and the YE1 nose. These requirements are met utilizing a set of square brackets located every 20° as shown in Fig. 9.20 and in more detail in Fig. 9.22(a). This system is functionally identical to the one already in use on the existing endcap, and in principle it might be possible to reuse the existing brackets. However, new brackets will be produced to guard against the risk that the existing brackets might be damaged during removal of the old endcap and to introduce improvements in the system for attaching the brackets to the back flange.

Finite-element calculations have been done for the top-most bracket, which is the most heavily loaded one. The calculation applies a 300 kN longitudinal force, based on detailed calculations done for the existing endcap [52]. This is a conservative value, since the new endcap is slightly lighter than the existing one, eddy current forces are expected to be lower in stainless steel than in brass, and the flexible ring support is expected to apply a lower force than the frictional force in the existing sliding joint. The material used in the calculation is EN 1.7225 (42CrMo4) steel, which has a modulus of 210 GPa and a yield strength of 390 MPa. The longitudinal deflection under the full load is <0.1 mm and the maximum stress, Fig. 9.22(b), is 170 MPa, giving a safety factor of 2.3.

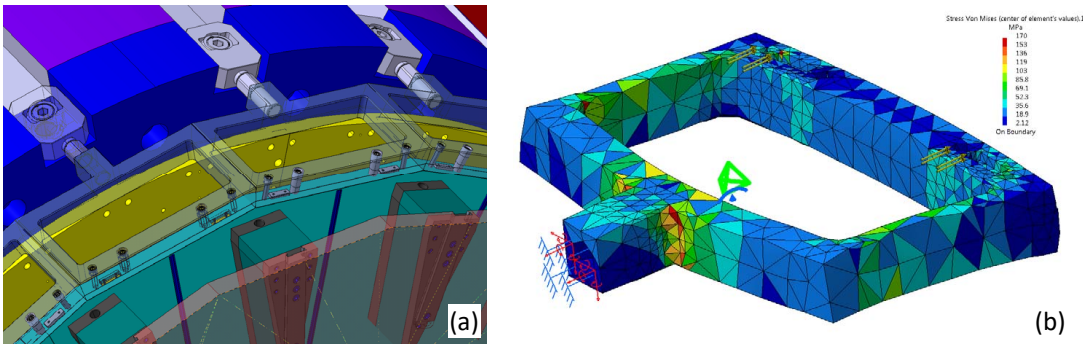


Figure 9.22: Finite element calculation results for the outer endcap brackets.

9.4 Cooling and environmental control

9.4.1 Cooling system requirements

The cooling system is required to cool the active detector elements to a nominal operating temperature of $-30\,^{\circ}$C with a margin of several degrees, and provide the ability to run at any intermediate temperature between room temperature and $-30\,^{\circ}$C. The system shall have the capacity to remove a total heat power, dominated by detector electronics power, of 220 kW summed over the two endcaps, with sufficient margin to cover uncertainties in this value over the life of the detector. The refrigeration shall be provided by several refrigeration plants such that the detector can continue to operate normally if one plant is out of service for any reason. The cooling system shall be segmented such that in case of a leak or other system failure, the cooling to only limited fraction of the HGCAL need be turned off and that the remaining detector still provides full η and ϕ coverage, although with reduced resolution over the affected area.

The total power dissipated in the HGCAL cold volume and due to the heat leaks from the surroundings is expected to be about 110 kW per endcap. Table 9.7 shows a detailed breakdown of the dissipated power that must be removed by the CO_2 cooling system.

Table 9.7: Power dissipated inside the cold volume of one endcap

Source	Power (kW)
Total power due to electronics	69
Sensor leakage current at 3000 fb^{-1}	8
Loss in cables low-voltage power cables	7
Low-voltage regulator loss	7.5
Power rail loss in motherboards	10
Thermal screen heat leak	1
Thermal screen heating foils	3
Grand total/endcap	**105.5**

9.4.2 Two-phase ACL cooling system

The two-phase accumulator controlled loop (2PACL) concept is currently used in four detector cooling systems in high energy physics. The concept originates from the AMS silicon detector cooling. AMS has been orbiting the earth since 2011 on the International Space Station [25]. This space born cooling concept was adopted for the LHCb Velo cooling in 2008 [26]. The ATLAS IBL cooling system [53] and the CMS pixel Phase-1 cooling system [24], both using the same concept, became operational in 2014 and 2017, respectively.

The operational experience with all these systems is extremely positive. The LHCb cooling system has provided a stable cooling for almost eight years, with very little downtime. The ATLAS IBL cooling shows the same record over two years with zero downtime during LHC physics runs. A simplified schematic of a 2PACL loop is shown in Fig. 9.23. Only a fraction of the liquid, depending on the power load from 0% up to 30–50% maximum, is evaporated. This low vapour quality range, i.e. a significant fraction of the fluid remaining in liquid phase, provides a high heat transfer from the pipe wall to the fluid.

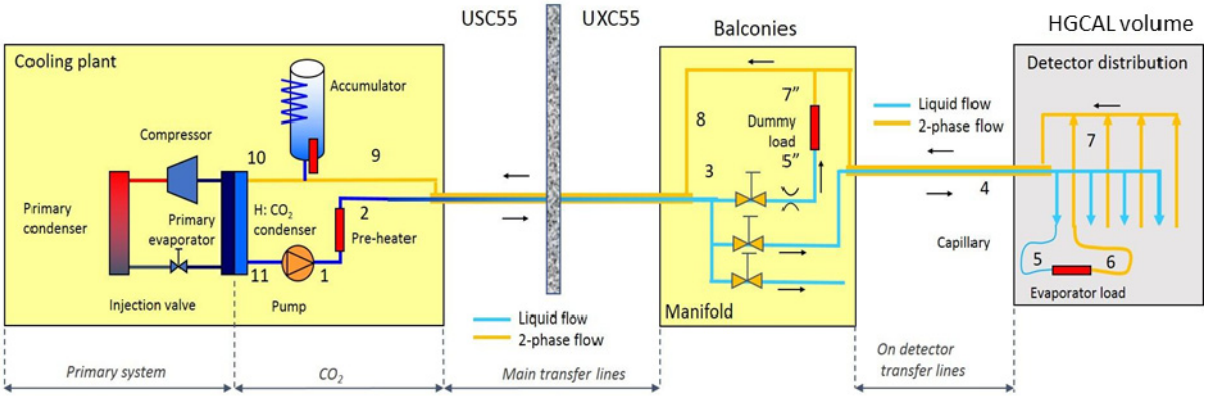


Figure 9.23: Schematic diagram of the 2PACL cooling system concept as used in the CMS pixel Phase-1 upgrade.

The transfer lines from the plant in the service cavern USC55 towards the experimental cavern are tri-axial, with liquid feed on the innermost pipe, mixed vapour/liquid return on the middle pipe, and vacuum insulation outside. This concentric concept conditions the inlet liquid to be

at the same temperature as the two-phase return line. As the inlet liquid has a higher pressure than the two-phase return, it is therefore automatically always in liquid state. This is a must for the detector inlet, as the presence of vapour would result in unstable flow distribution.

The thermal conditioning of the inlet liquid following the two-phase return temperature also has a welcome regulation effect. The thermodynamic cycle of the 2PACL system is shown in the pressure-enthalpy diagram of Fig. 9.24. No matter what the two-phase temperature is, the liquid temperature will adjust, following the two-phase temperature. The liquid feed is state point number 3 in the schematic of Fig. 9.23 and the graph of Fig. 9.24. The two-phase return is at point 8. An expansion of the inlet liquid into the detector evaporator loops bring the loops in a low vapour quality state (points 4 and 5). Both liquid and two-phase state points (points 3 and 8) lie on the same isotherm, which is horizontal in the two-phase region and nearly vertical in the liquid region.

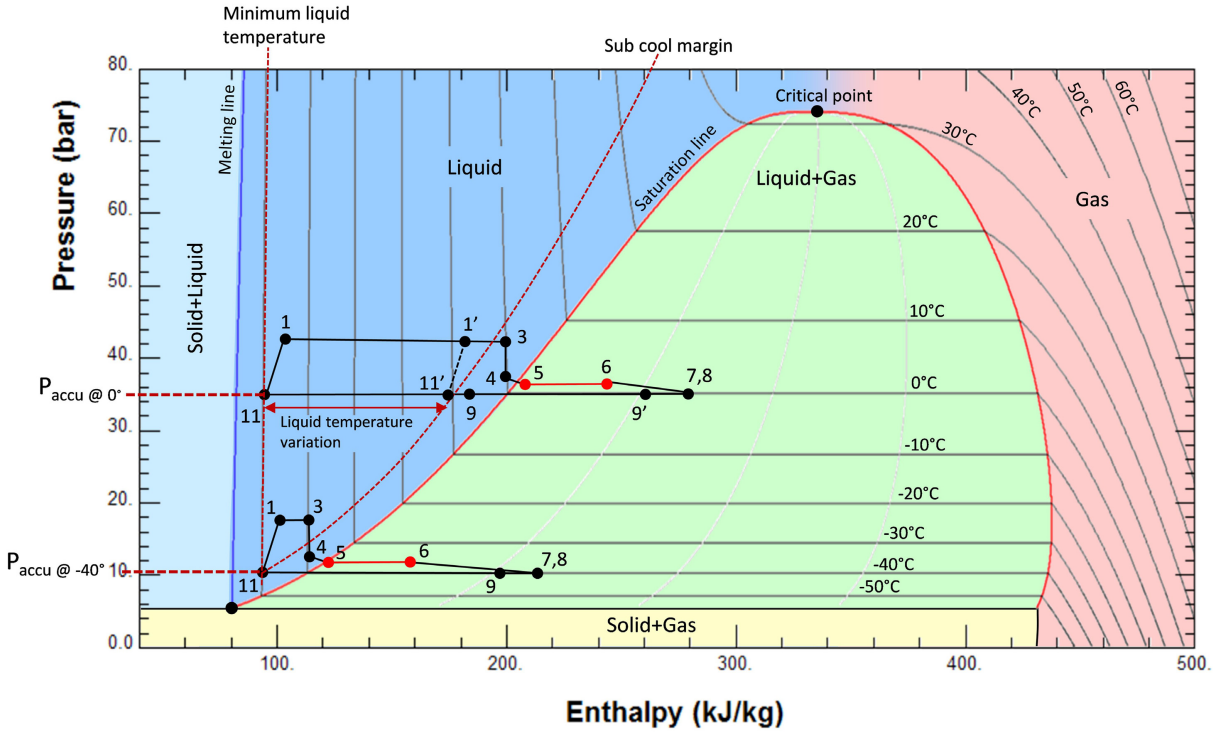


Figure 9.24: The 2PACL operation in the CO_2 pressure enthalpy diagram. The state point numbers correspond to those in Fig. 9.23.

The two-phase temperature is the main parameter that needs to be controlled in a 2PACL loop. This is done by an accumulator vessel, which can be located far away from the detector evaporators, provided the pressure drops on the return lines are kept reasonably small. The accumulator vessel contains always a liquid/vapour mixture and its pressure is controlled by heating and cooling. The pressure of the accumulator and the pressure drops on the return lines sets the evaporating temperature in the detector and, together with the above explained liquid temperature regulation, makes the system fully controlled by the accumulator only. This is why the concept is named 2PACL, 2-Phase Accumulator Controlled Loop. This simple approach to control and the passive nature of the distribution inside the detector makes the concept ideal for detector cooling, moving the controls away from the inaccessible areas of the detector.

The evaporator status in the detector loops is represented by state points 5 and 6, coloured in red in Fig. 9.24. The temperature there is a bit higher than the regulated two-phase temperature in the accumulator because of the on-detector cooling distribution pressure drops. This value

needs to be optimized by fine tuning between pipe diameter and length constraints imposed by the mechanics of the cooling plate and minimizing the pressure drops to achieve a small temperature difference across the evaporator. This is especially important for low temperature operation as the temperature loss due to pressure drop increases with decreasing temperatures. On the other hand, the cooling system is limited in going to low temperature as the CO_2 freezing temperature of $-56.6\,^{\circ}C$ comes close. The lower limit is shown in Fig. 9.24 with the lower cycle. Here both the minimum temperature lines with sufficient offset to freezing and the sub-cool margin line come together, defining the lower limit of the cooling system, at around $-35\,^{\circ}C$.

9.4.3 Cooling plant concept and backup logic

The general concept described in this paragraph will be fully validated in years 2017–2020 by the CERN-EP-DT cooling team, where an ongoing R&D programme aimed at a first prototype plant (the Demo) for the Phase-2 Upgrades of both ATLAS and CMS.

All the CMS cooling plants will be located in the service cavern in USC55, and several cooling plants will be connected to multiple manifolds on the UXC55 balconies. With the currently available technology a maximum cooling capacity of 60 kW at $-35\,^{\circ}C$ and with a vapor quality of 42% looks reachable. This vapor quality gives still plenty of margin with respect to dry out conditions that preliminary calculations show to happen at about 70% vapour quality. Indeed, the present pixel Phase-1 cooling system, achieving 15 kW of cooling power with a non-negligible margin, uses a single-headed membrane pump that exists also in triple-headed configuration, thus enabling the delivery of about three times more flow and consequently more cooling power.

A set of several cooling units operating in parallel allows redundancy. Two options are being investigated: either a parallel function of all plants at reduced power and increase of power in case of one unit failing, or keeping one plant idling and ready to enter in operation (Fig. 9.25). If the first concept is validated, a total of five cooling plants would be enough to serve the total power request,;otherwise a sixth plant shall be foreseen. Each cooling plant unit contains the pump, the condenser, the dampers, the plant sensors, and the internal by-pass. At least two plants are equipped with a second condenser, connected to an air-cooled backup primary system. These plants remain operational also in case of a power cut and guarantee that the HGCAL volume remains cold when not taking data.

The primary cooling technology is still under discussion. It can be a direct expansion cooling, as done now for the current systems, a CO_2 primary, or a liquid brine solution. The latter is easier to be used for cooling multiple systems and is likely to be more stable in temperature, which is a needed feature when going to lower temperatures ($-35\,^{\circ}C$) than in the systems currently in use at CMS.

9.4.4 Distribution system and pipework

A simplified schematic diagram of the HGCAL cooling system is shown in Fig. 9.26. The pipework between the cooling plants located in USC55, the distribution manifolds placed on the UXC55 balconies, and the entrance of the detector cold volume will be vacuum insulated rigid coaxial transfer lines. In the coaxial transfer lines the innermost pipe will carry cold liquid CO_2 towards the detector. That pipe is placed inside a bigger pipe, which returns two-phase CO_2 from the detector. Such coaxial inlet/outlet pipes will be encapsulated inside vacuum insulation pipes. A backup system in order to maintain vacuum at the desired levels can be installed, featuring getter pumps already in use for the present CMS pixel transfer lines. The

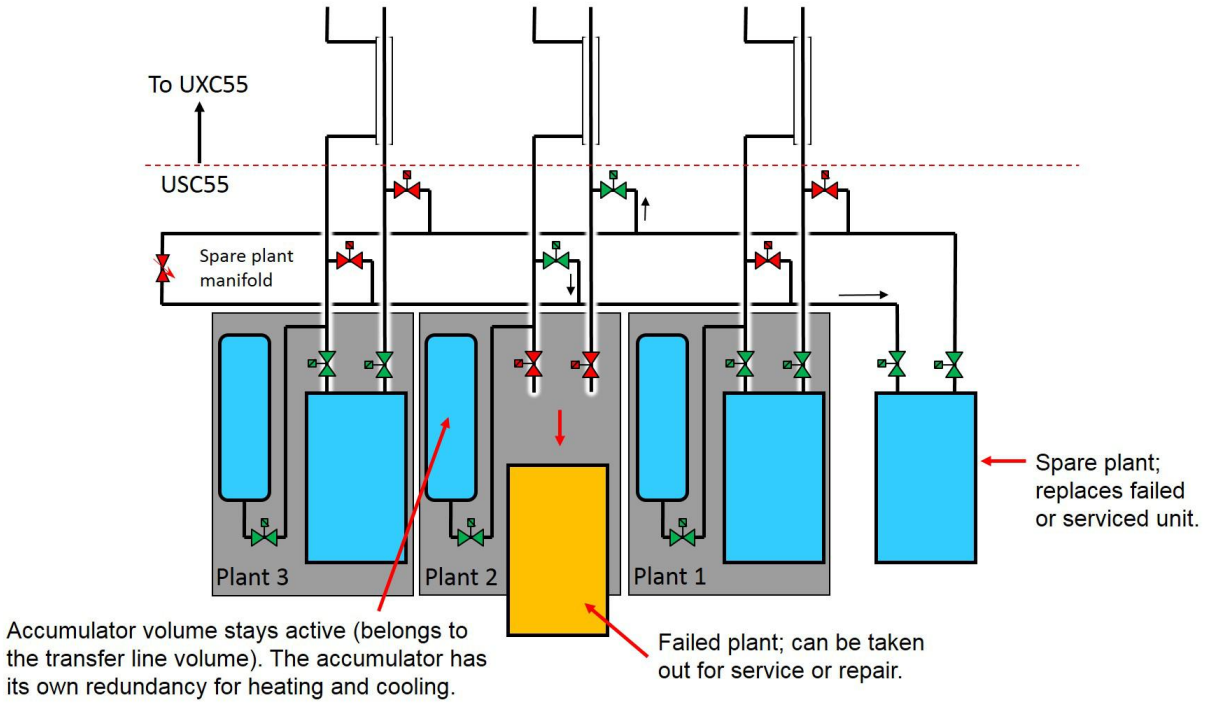


Figure 9.25: Schematic diagram of the 2PACL cooling system concept as used in the CMS pixel Phase-1 upgrade.

dimensioning of the coaxial pipes and in particular the distance between the cold process pipes and the outside vacuum jacket shall conform as much as possible to industrial standards, with a minimum number of upward columns in the return line flow. Upward columns in a two-phase flow cause a static pressure drop, which is translated into a temperature gradient. Siphons should be avoided as much as possible to minimize fluid slug containment during standstill. Fluid slugs can potentially be pushed through an already warmed up system, causing thermal shocks. The vacuum level of the insulation tubes must be lower than 10^{-4} mbar, guaranteed by the supplier for at least ten years. If special circumstances or constraints require deviating from industrial standards for the ratio between pipe size and volume of the vacuum jackets development programmes will be foreseen in order to validate the design.

The cooling supply/return pipework must be produced in sections to facilitate installation in the underground area. The individual sections have to be equipped with service ports allowing connection to portable vacuum pumping units in case vacuum degradation would be noticed. No water condensation at the outer surface of the transfer line is acceptable; thus validation tests will be done as part of the acceptance protocol, to ensure that the surface of the vacuum jacket never falls below 14 °C. During system operation, the outer surface temperature shall be regularly verified using thermal imaging.

At the entrance of the HGCAL cold volume, each transfer line is split into supply and return manifolds to distribute the flow to in parallel to a set of cassette cooling plates. The manifolds are made of standard non-coaxial pipes and need to be designed in order to match the tight space available on the outer surface of the detector. Preliminary calculations show that the supply manifolds could be 6 mm ID and 8 mm OD while the return could be 8 mm ID and 10 mm OD. These sizes have been chosen in order to keep the total pressure drop on the return lines below 1 bar. There is a capillary tube at the inlet to each cassette, which causes a 5–10 bar pressure drop in nominal flow conditions. This is needed to regulate the flow distribution

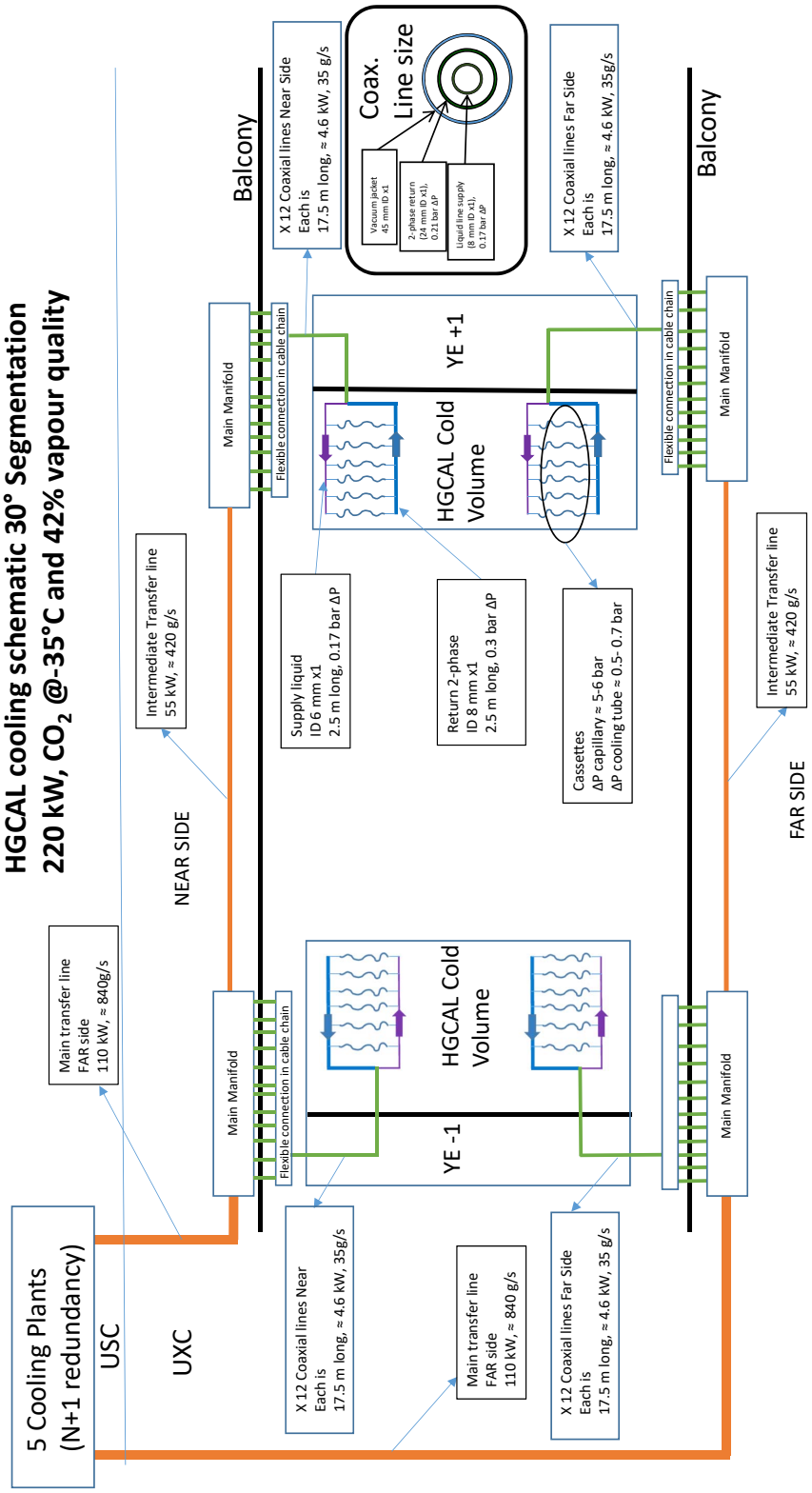


Figure 9.26: Simplified schematic diagram of the HGCAL cooling system.

evenly among all the detector cassettes. In evaporator loops, the pressure drop depends on the vapour content and thus the absorbed power. Without flow distribution restrictions, the flow would be inversely proportional to the applied power, meaning that less flow is fed to a higher powered evaporator loop, which could cause dry-out in these lines. With appropriate flow restriction the liquid inlet condition is the same for all parallel detector circuits and the flow distribution does not depend on the downstream heat absorption. A second important aspect of the inlet restrictions is that the manifold pressure is high. This higher pressure suppresses possible boiling in the region outside the detector. Boiling in this area should be avoided, as phase separation might cause a vapour feed to an individual cooling loop, which could also cause a sudden dry-out of the evaporator line. The pressure drop over all detector loops connected to a single detector distribution manifold must be designed to be equal. If different flows are needed the restrictions must be chosen such that the pressure drop over the loops and restrictions is the same. A default design pressure drop is in the order of 5–10 bar, from minimum to the maximum required flow.

All the connections between the cassettes cooling loops and the manifolds will be done using standard Swagelok VCR connector components that have been successfully used in the past.

Table 9.8 shows the range of total flows for each endcap expected at different operating temperatures and at nominal cooling power. The spread in flow and vapour quality is explained by the different enthalpy changes available at different temperatures (176 kJ/kg at 15 °C and 312 kJ/kg at −35 °C). Based on past experience, mainly with the CMS Tracker, it is expected that a dew point below 0 °C inside the cold volume can be achieved quickly. Therefore it is foreseen to run the detector at ≈0 °C during the commissioning phase and gradually lower its temperature based on the humidity inside the cold volume. Initial operation at higher temperature, up to 15 °C if necessary, is in principle possible. At a CO_2 fluid temperature in the return line less than −25 °C (detector temperature less than about −23 °C), only four of the five installed refrigeration plants are required leaving the fifth to serve as a backup.

Table 9.8: Range of flows for each endcap at different operating temperatures.

Fluid Temperature (°C)	Vapour Quality (%)	Flow (g/s)
15	60	1000
0	50	950
-10	42	1000
-35	42	840

9.4.5 Controls

The CO_2 detector cooling plants and their associated distribution will be based on industrial programmable logic controllers (PLCs) of Siemens or Schneider and Siemens WinCC OA as Supervisory Control and Data Acquisition (SCADA) [54]. Both PLCs and SCADA will conform to the latest version of the Unified Industrial Control System (UNICOS) Continuous Process Control (CPC) framework of CERN. The UNICOS framework deals with the two upper layers of a classical control system: supervision and control, employing the terminology and models of the ISA-88 standard for batch control systems. The PLCs and SCADA server for the process security will be placed in the CERN Technical Network, which is physically detached from the outside world. All the acquired data will be stored in the LHC-logging database for long term archiving. Additionally via DIP and direct MODBUS, or equivalent, communication protocol data will be sent to the detector control system (DCS). Two-way hardwired communication shall be provided to link the CO_2 control system with the CMS Detector Safety System (DSS).

If a distributed I/O topology for system modularity will be required, modern industrial fieldbusses, like Ethernet IP or PROFINET, will be used. All cooling system user interfaces (UI) will be accessible from one common point and based on a unified structure. The SCADA visual instrumentation representation will be common to other CO_2 detector cooling systems at CERN.

9.4.6 Environmental control

Due to the foreseen low operating temperature of the detector, a thermal screen will be installed around it in order to separate the cold ($-30\,^{\circ}$C) detector volume from the ambient conditions in the cavern. The thermal screen material needs to have low thermal conductivity and good mechanical properties. To achieve this, a sandwich configuration has been chosen with outer skins are of 1.5 mm Permaglas ME771 and a core of Aspen Spaceloft Subsea, which has a thermal conductivity of 0.0145 W/mK. The core thickness varies from 15 to 32 mm according to the space available. The thermal screen is composed of several pieces that are joined together by Permaglas structural members.

The worst case situation from the point of view of the thermal performance of the screen is when CMS is closed. In this case the endcap is surrounded by a layer of about 4 cm of static air between it and the adjacent detector elements, in which no convection can take place. Preliminary calculations show that under these conditioins, the outer surface of the thermal screen would reach a temperature in a range between $-10\,^{\circ}$C and $-25\,^{\circ}$C. Under these conditions, the heat leak into the detector from the ambient air would be about 1000 W per endcap. In order to maintain the surface temperatures above the ambient air dew point (about $13\,^{\circ}$C), an active system of heating foils will be installed. It has been calculated that the foils will need to provide a power of about 3000 W per endcap. This extra load of 4000 W has been taken into account for the sizing of the cooling system.

At the moment the thermal screen is conceived as a single self-supporting structure that will be put over the HGCAL once the assembly is finished. Nevertheless, an alternative option is considered where the thermal screen is made of independent panels to be attached to the absorbers structure. That would simplify the installation and allow the removal of a panel whenever an access to a specific location to address a very urgent problem will be needed. The thermal screen seams will be sealed with vapour barrier foils in order to reduce as much as possible the humidity leaks from the ambient.

Another critical point will be the sealing of the services feedthroughs (cables and pipes) on the back of the detector. The past successful experience with tracker for similar problems will be taken into account along with extensive testing. When the detector electronics are off and the detector is cold, all the cables will represent an significant thermal bridge between the cold volume and ambient favouring the formation of water or even ice on their surface. The risk of this happening is quite elevated since the cables are in good thermal contact with the cold surfaces of the detector. In order to prevent this, heating wires wrapped around sets of cables will be installed for a length that still needs to be determined.

To prevent water condensation inside the cold volume, dry nitrogen (or dry air during maintenance periods for safety) is injected into the detector cold volume. Sufficient net outflow of dry gas (in order of 3–4 volumes/hour) will be needed. In order to control the amount of overpressure inside the thermal screen within safe limits, calibrated holes will be provided on the thermal screen to allow the gas to escape in a controlled way. On top of all the sensors and software safety systems (see Section 3.5), a few safety bubblers will be fitted in the gas system. The details of the dry gas system on detector distribution are still under study, but the main idea

would be to feed the gas from the inner diameter in order to force it to go through the cassettes before going out through the above mentioned holes on the thermal screen. The dry nitrogen (air) is supplied by the existing membrane and drier plant located in a service building on the CMS site. An additional dry air plant along with a nitrogen dewar and nitrogen bottles will act as back-ups in case of failure in the membrane/drier system.

The humidity and temperature measurements will be based on the technologies used in the current tracker, with probes located in the detector cold volume. Optical fibre sensors are used for their radiation hardness. Cheaper, non-radiation hard semi-conductor sensors are used to increase the coverage of measurement points during the installation phase and in areas with lower radiation levels. As a reference for the humidity measurements, sniffing pipes are routed from the critical regions to the CMS service cavern and connected there to industrial humidity measurement sensors. This sniffing system is used to provide reliable reference measurements for all conditions and radiation levels through the life cycle of the HGCAL.

9.5 Services

Services required for the operation of the HGCAL are the the following:

- High-pressure CO_2 fluid to cool the detector;
- Low-voltage (1.5 and 1.2 V), high current supply to operate the analogue and digital electronics;
- High-voltage, low current supply to bias the silicon detectors;
- Optical fibres for trigger, DAQ and detector control;
- Dry gas system to maintain the dew point below the operating temperature.

All of the services, with the possible exception of some of the dry gas system piping, must be run over the outer surface of the HGCAL. In addition, space must be allocated to route a similar set of services to the endcap timing layer (TE), which will be mounted in front of the HGCAL. The space required for the cooling and LV cabling dominates the space required for the services, and initial studies have focused on these.

Conceptual work has been started on how to route the services for the endcap timing layer; however the requirements from the TE are not yet well determined and designs to meet the requirements are very preliminary. The TE design, its services and its interface with the HGCAL will be fully presented in the TE TDR, which is expected to be submitted in mid-2018.

9.5.1 Cooling services

The CO_2 cooling is divided into 30° sectors, and within each 30° sector even and odd layers are cooled by independent manifolds. This results in a total of 24 cooling circuits per endcap, each of which is independently controlled by valves located off of the detector. The layout of the services, including the cooling piping, must take account of the 20° polygonal geometry of the HGCAL absorbers (which follows the 20° segmentation of the CMS central detector). The cooling pipes are therefore routed in two out of every three 20° facets, as shown in Fig. 9.27. The piping is located as close as possible to the apexes of the polygonal structure, where there is already minimum space for the LV cabling. Following the design used for the CMS Phase 1 pixel system, connections from the manifolds to the individual cassettes are made using Swagelock VCR high-pressure connectors.

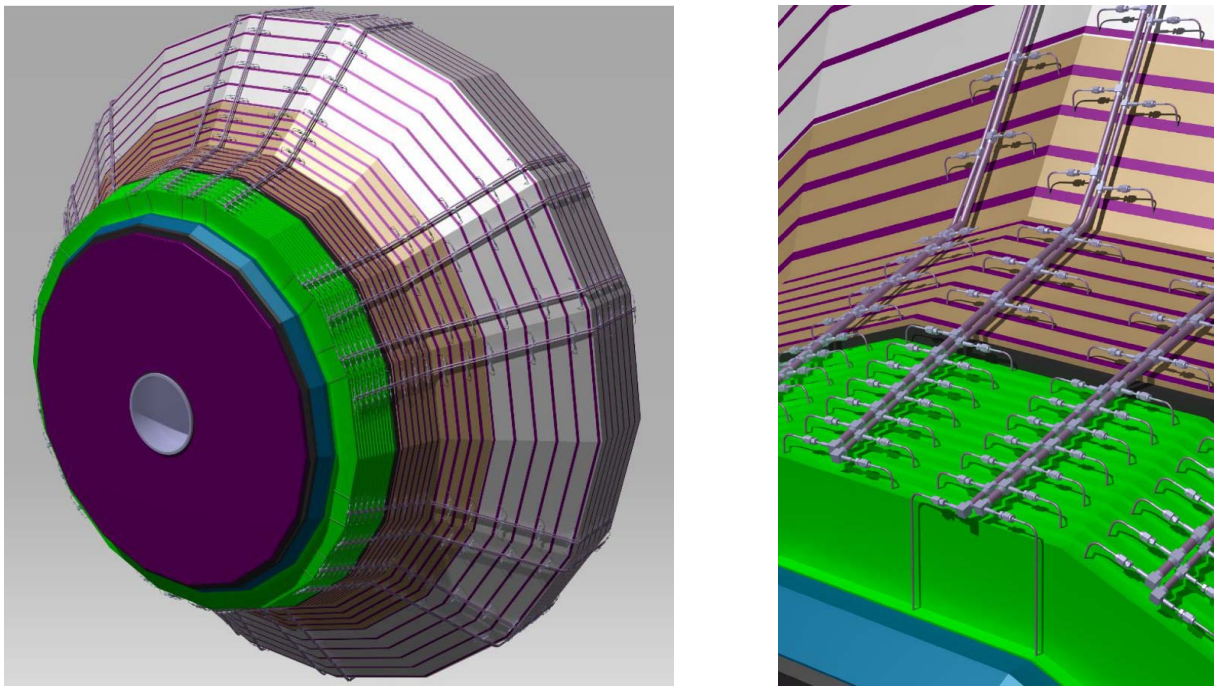


Figure 9.27: Preliminary layout of the CO_2 piping for the HGCAL: general view on the left showing the 24 supply and return manifolds; detailed view on the right, showing the connections to individual layers.

9.5.2 Electrical services

Each line of motherboards (shown in Fig. 9.1) is powered by two DC-DC converters located just outside the cold volume behind the CE-H, one each for the analogue and digital parts of the electronics. Details have been given in Sections 3.4 and 8.5. A common ground return for the two voltages is used. The required minimum copper cross-section is computed for each motherboard line in each layer of the HGCAL for the analogue and digital supply and common return such that the power loss in the cables from the DC-DC converters to the edge of the cassettes is at most 10% (5% each way). Standard cables [55] are chosen for each line that have the smallest copper cross-section consistent with the computed minimum, and the catalogue value of the diameter of the insulated cable is recorded for use in the preliminary layout of the LV services. Table 9.9 summarizes the numbers of and sizes of the different cable types. The total current carried by these cables in each endcap in the analogue and digital circuits respectively are 23 kA and 25 kA.

Table 9.9: Summary of low-voltage cables, of commercially available cross-sectional area, required for each endcap calorimeter.

Cu area	Insulated	Number of Cables			Total Cu	Total Ins.
(mm^2)	dia. (mm)	Analog	Digital	Return	Area (cm^2)	Area (cm^2)
4	4.2	1339	1167	19	101	366
6	5	612	357	542	91	297
10	6	867	822	912	260	735
16	7.7		472	733	193	561
25	9.5			612	153	434
Total per endcap		**2818**	**2818**	**2818**	**798**	**2362**

The electrical and optical services as a function of azimuth follow from the layout of the motherboards within the cassettes. Within each 60° cassette (CE-E) or cassette pair (CE-H), the motherboards are laid out in paths parallel to the nearer of the two edges, resulting in a "vee"-shaped pattern, as shown in Fig. 9.1. The azimuthal density is proportional to the number of modules along each motherboard line, and therefore results in the highest density at the edges and substantially lower density toward the middle of the 60° section. The top of Fig. 9.28 shows a preliminary layout of the LV cables over one half of a 60° wide segment, shown as a cross-section at the last CE-E layer, where the density of the services is near maximal. The cable dimensions discussed above and summarized in Table 9.9 are used and a realistic filling fraction of ≈70% relative to ideal packing of circular cables is used to compute the required cross-section of cables at each location along the "road" to the back of the HGCAL. The higher density near the edge than at the centre of the 60° sector is evident. Shown also are 45 mm wide "no-go" zones that occur every 30° in azimuth where the mechanical supports for the HGCAL occur and 30 mm wide no-go zones every 20° at the apexes of the polygonal structure where the supports for the thermal screen are located.

As has been done for the Phase-1 Pixel system, a full 3D cabling plan will be developed that specifies the precise location of each cable so as to minimize the volume required to route the services. This will also determine the precise location of each cable at the position at the cold-to-warm feedthrough area (see below), which is necessary both for a good seal between the cold volume and the outside of the HGCAL and for proper routing of the services outside the HGCAL.

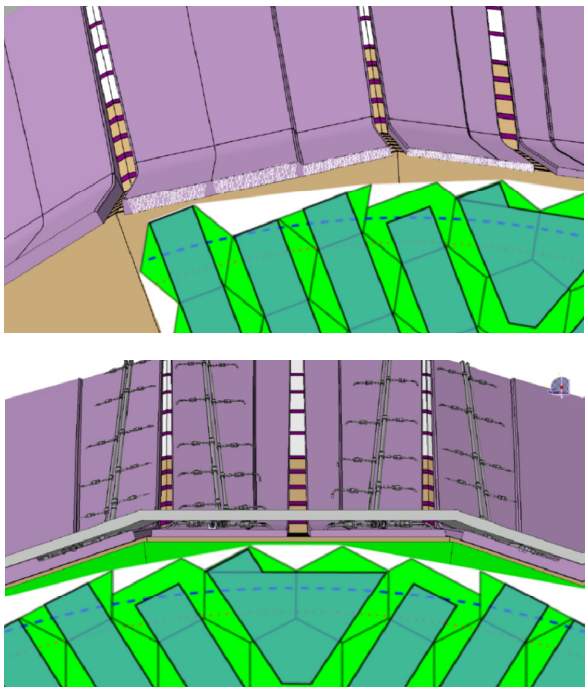


Figure 9.28: Cross-section of the LV cables at the last layer of the CE-E (top), showing the variation in cable density following from the 60° periodicity of the cassette motherboard layout. Cross-section including cooling piping and the thermal screen (bottom). Cartoons of the edge of a cassette indicating the motherboard layout have placed on top of the 3D model pictures.

After the electrical and optical services are installed the cooling pipes are connected to the cassettes, as shown at the bottom of Fig. 9.28, and finally the thermal screen closes the detector volume. The layout of the cooling pipes follows a 60° periodicity (Fig. 9.27), as does the density of electrical and optical services. The location of the cooling pipes is concentrated in the

region where the electrical services density is lowest. However, the pipes at the left and right of in the bottom of Fig. 9.28 clearly interfere with the cable volume. The interference is greatest in the region between the back of the CE-E and CE-H layer 8. From CE-H layer 9, the radius of the detector grows rapidly providing more azimuthal space for the services, and the density of services per cassette decreases as silicon is progressively replaced by scintillator. Further engineering design work is needed to resolve this apparent interference. Possible solutions include: identification of alternate cables with smaller insulated diameter but equivalent specifications; replacing the cables with a custom-designed array of LV bus bars that occupies a smaller cross-section by providing more optimal packing; modest (up to ≈10–20 mm) reduction in the outer radius of the detector in the highest interference region to provide more space for the services; or providing 12 V to DC-DC converters located at the edges of the cassettes. It is also under consideration to reduce the number of cooling pipes by providing cooling in 60° rather than 30° sectors, which would make more space available for the electrical services.

9.5.3 Services for the endcap timing layer

The endcap timing layer is mounted on the front of the HGCAL, and therefore its services must be routed past the HGCAL. Preliminary work has begun to define the required services and to determine how to route them from behind the HGCAL to the TE. The current understanding as of the writing of this TDR of the requirements placed on the HGCAL by the TE are summarized here, but since this is an area of active work, the requirements and the designs to meet them are likely to evolve substantially. The TE design and its interface with the HGCAL will be fully presented in the TE TDR, which is expected to be submitted in mid-2018.

It is currently proposed to route the TE services inside the HGCAL thermal screen but in channels separate from the volume used for the HGCAL services. In this concept, the TE services are concentrated in two channels at the 3 o'clock and 9 o'clock positions that are placed outside the radius of HGCAL services to avoid interference between the two systems.

Figure 9.29(a) shows the current concept of how the TE services would be routed into the TE thermal volume. At the front of the HGCAL, the thermal screen is split to provide a thermal volume for the TE separate from the one that contains the HGCAL and the neutron moderator. The front surface of the thermal screens are at the envelope surface that defines the maximum longitudinal and radial extent of the endcap. The outer radius of the moderator is reduced by about 125 mm relative to its pre-TE size to provide a volume in which to route the TE services within the TE volume. The TE services pass through the HGCAL thermal screen just above the moderator. Design work in ongoing to ensure that no condensation occurs at the warm end of the feedthroughs when the TE is open and the HGCAL is cold.

In the current concept the region in which CE-E active detector must be eliminated to permit the routing of the TE services is 3.5 modules, as shown in Fig. 9.29(b). The affected region extends from $\eta = 1.48$ to $\eta = 1.63$, corresponding to trigger tower 18 and most of trigger tower 19 in the current detector, over about 10% of the azimuth. Design work is being done to understand how to minimize the volume of CE-E that must be modified.

9.5.4 Cold-to-warm feedthroughs

All the services (cables, cooling and dry gas pipes) will enter HGCAL cold volume from the back of the detector, passing through the back of the thermal screen near the outer radius of the detector. It is clear that this area is quite critical from the point of view of guaranteeing a complete seal against humidity and avoiding any condensation or ice balls formation on pipes and/or cables. The HGCAL feedthroughs will be made following the same design approach

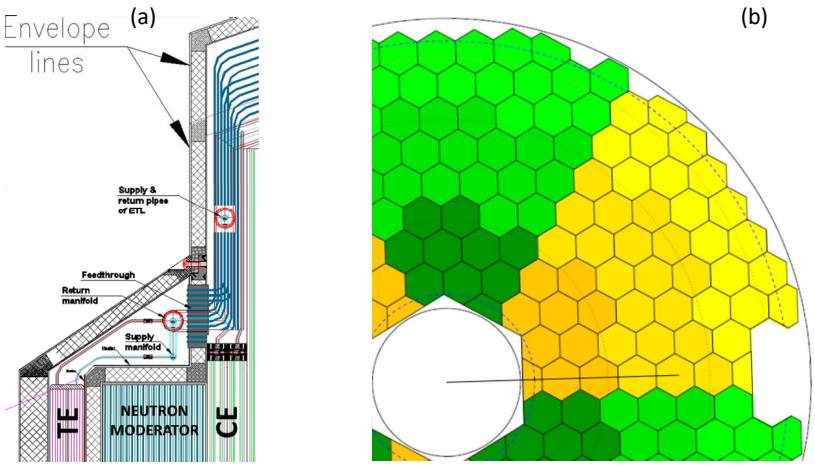


Figure 9.29: Current concept of the HGCAL-TE interface. (a) The region of the feedthrough of the TE services from the HGCAL to the TE cold volume. (b) One quadrant of the first CE-E layer showing the region of silicon sensors that would be eliminated from the first four detector layers (two cassettes) to accommodate the routing of the TE services as shown in (a).

that was successful used by the existing CMS Tracker, which had to face the same problem, and more recently by the Phase-1 Pixel system. Based on the detailed cable map (see above), the precise location of each cable is known and a set of custom designed clamps ("combs") is made that position and secure the cables at the location of the feedthrough. Figure 9.30 illustrates the principles of this design as used for the recently installed Phase-1 Pixels. After the cables are secured, the region is potted to provide a tight seal to ensure that the humidity of the cold volume can be maintained at the required low level.

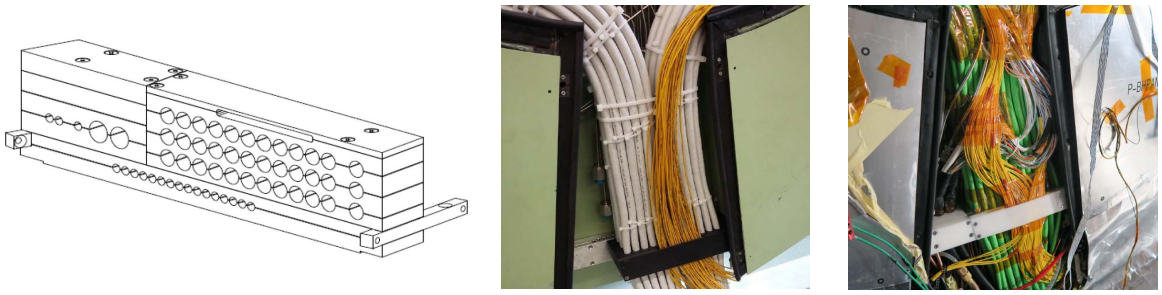


Figure 9.30: Cold-to-warm feedthrough system as implemented in the recently installed Phase-1 Pixel system. Left: Engineering design of the cable clamp. Middle: Mockup to verify the design. Right: Actual installation in CMS.

The cooling pipes will pass through the cold-warm interface utilizing the same design approach and will also be potted in place. Outside the cold volume, the pipes are stainless-steel vacuum jacketed coaxial lines, which break into individual supply and return pipes only inside the cold, dry environment inside the thermal screen, thereby minimizing the risk of condensation. The possibility to seal the volume in between the back disk and the back flange and flush it with a suitable amount of dry gas (about one or two volumes exchange per hour) is also under study. This is a similar approach to that used by the the existing tracker and would eliminate the risk of condensations on the cables where they first emerge from the cold region.

It is planned to work closely with the present Tracker collaboration and with those working on the Phase-2 Tracker system to develop and implement common solutions and to understand what improvements can be made to the feedthrough design and potting procedures.

9.6 Assembly and installation

9.6.1 Electromagnetic calorimeter assembly and tooling

The two Electromagnetic Calorimeters will be assembled in CERN Building 904. The main pieces of tooling needed for this assembly (Fig. 9.31) are relatively simple:

- two custom made precision tooling tables on which the two CE-E are assembled with capacity to support the 23 t mass of one completed CE-E;
- several under-the-hook lifting fixtures specifically designed to lift the cassettes (maximum mass of ≈225 kg each) using the lifting points built into the cassettes (see Section 9.2.1);
- a specialized fixture to lift the completed CE-E for transport to Point 5 and installation onto the CE-H (not shown in the figure);
- a support frame and protective cover for over-the-road transport from Building 904 to Point 5 (not showin in the figure).

The CE-E will be assembled in a clean room that includes a small gantry crane sufficient to lift and manoeuvre individual cassettes. The initial installation of the back disk onto the tooling table will require use of the main crane in Building 904 and so the roof of the clean room must be removable.

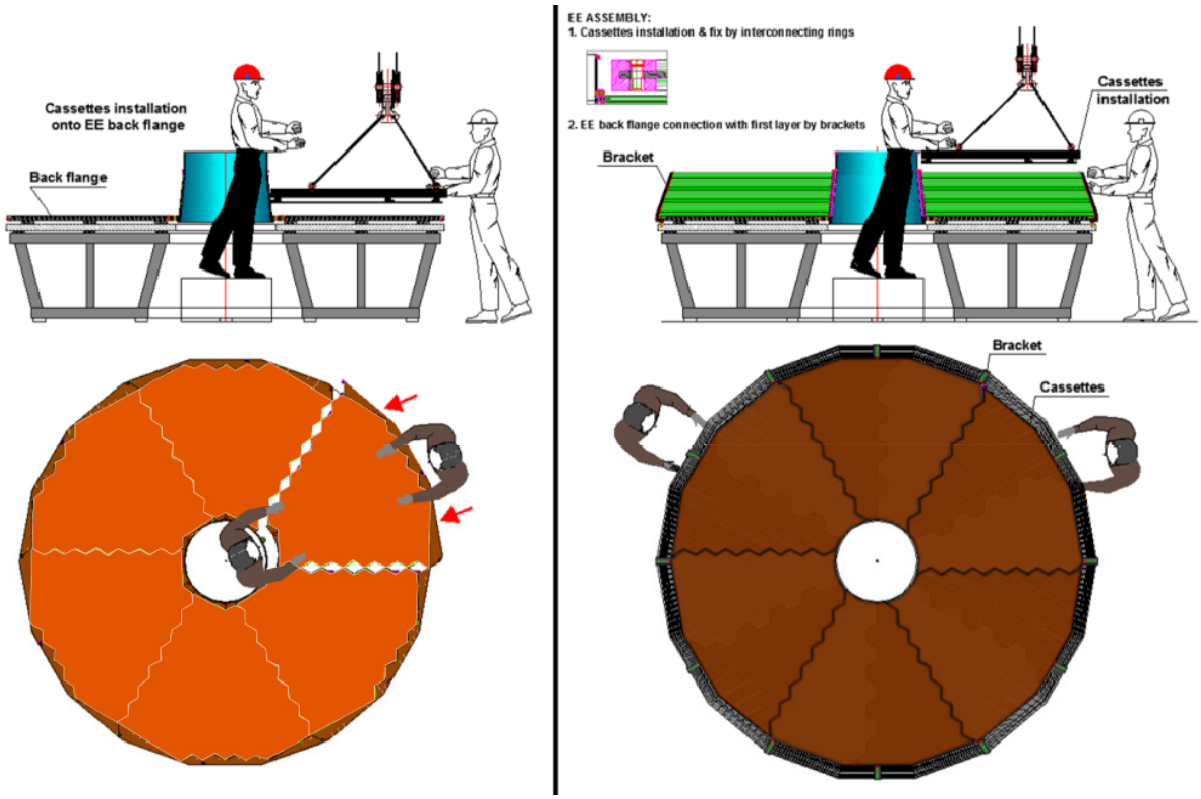


Figure 9.31: Process and tooling for the assembly of the CE-E.

The main components that are assembled to build one CE-E are the following:

- 1 stainless steel back disk
- 1 aluminium support cone for the CE-E

- 1 Permaglas support tube (extension of support cone) for the neutron moderator and endcap timing layer
- 14 aluminium cassette interconnection rings
- 1 aluminium stop ring
- 84 cassettes (14 cassette layers, 6 cassettes per layer)
- 84 cassette interconnection plates
- 12 stainless steel longitudinal brackets
- 1 neutron moderator plus mounting hardware

Each of the components is fully inspected and qualified before delivery to the CE-E assembly site. Crucially, the cassettes will have undergone rigorous quality control during assembly and an extensive final testing programme that includes several thermal cycles between room temperature and $-30\,^{\circ}\mathrm{C}$ and a multi-day test run at $-30\,^{\circ}\mathrm{C}$ with all electronics operational and cosmic ray data read out by a full DAQ system (Section 9.2.3). Each component will undergo a more limited set of tests and measurements at the assembly site to ensure that no damage or degradation has occurred in transit from the production site.

The CE-E cassettes are an integral part of the mechanical structure and therefore care must be taken to ensure their structural soundness and compliance with dimensional tolerances before and during CE-E assembly. In addition, once a given layer of cassettes has been installed the cassettes in previously assembled layers cannot be removed without disassembling all layers above. Hence it is important at each stage of assembly to check that all cassettes are fully operational in order to minimize the risk of time-consuming repairs and rework.

Assembly of the two CE-E will proceed in parallel with each other. The assembly sequence for each one is the following:

1. The tooling table, back disk and support cone are inspected and measured for compliance with all dimensional requirements.
2. The support cone is attached to the back disk, and the completed assembly is measured to ensure correct assembly including requirements on the perpendicularity of the cone axis to the back disk.
3. A trial assembly of all of the interconnection rings onto the support cone and all dimensions measured to ensure correct construction and check for any buildup of tolerances that would put the whole assembly out of tolerance.
4. Prior to delivery to the CE-E assembly site, all cassettes will have been fully tested and qualified at the operating temperature of $-30\,^{\circ}\mathrm{C}$.
5. Before installation, cassettes received from the cassette assembly site are visually and electrically inspected to verify that:
   - no damage occurred during shipping and handling;
   - all hard points and attachment points are properly located;
   - all electrical, optical and cooling connection points are correct;
   - all electrical and optical components are functioning correctly.
6. The back-most interconnecting ring is installed and the cassettes of the back-most layer are placed against the back disk, secured to the interconnecting ring, and joined to each

other via the interconnection plates at the outer radius, as shown in the left side of Fig. 9.31.

7. The assembled cassette layer is
   - visually inspected for damage that may have occurred during installation;
   - mechanically measured to verify compliance with dimensional requirements;
   - tested to verify that all electrical and optical components are functioning correctly.
8. The previous two steps (install and inspect/measure) are repeated for the subsequent 13 layers as indicated in the right side of Fig. 9.31.
   - After the installation of each layer, a visual inspection of the previously installed layers is performed.
   - Electrical tests of previously installed layers may also be done in some cases.
   - To maintain the overall thickness within tolerances, it may be necessary to install local shims. Alternatively, the cassette thicknesses measured during cassette assembly could be used to specify the locations of cassettes in each layer to minimize tolerance buildup.
9. The 12 longitudinal brackets are in installed at the outer radius and the stop ring is connected to the support cone at the front of the assembly.
10. The Permaglas support tube and the neutron moderator are installed.
11. A complete electrical test of the assembled CE-E is performed at room temperature using a temporary cooling system to maintain the electronics temperature.
12. The completed CE-E is prepared for transport to Point 5.

The mass of the completed CE-E exceeds the capacity of the crane in Building 904. Therefore the assembly table will be equipped with an air pad system to allow the CE-E to be moved out of Building 904, where a portable crane can be used to place it on the transport frame on a trailer for transport to Point 5.

9.6.2 Hadronic calorimeter assembly and tooling

Construction of the hadron calorimeter must take place at Point 5, since there are very few sites at CERN that can accommodate its large size and mass and since transport of such a large and heavy detector is impractical. The assembly of the two endcaps will proceed in parallel. The assembly can be divided into several major stages:

- Assembly of the endcap suspension system;
- Stacking of the stainless steel absorber disks;
- Installation of the CE-H cassettes and their electrical and optical services.

Endcap suspension system assembly and tooling

The construction of the endcap suspension system uses simple tooling and proceeds according to the sequence shown in Fig. 9.32. The main components of each endcap suspension system that are assembled at this stage are the following:

- 1 stainless steel support ring;

- 1 stainless steel back flange;
- 18 sliding wedge support assemblies;
- 1 CE-H back disk.

The tooling required for this stage of assembly includes the following:

- a set of bearings with appropriate attachment points for rotation of the back flange assembly;
- a specialized lifting fixture for moving the various disks (back disk, back flange, back thermal screen, and completed suspension system assembly) with capacity to lift the ≈35 t mass of the completed suspension system.

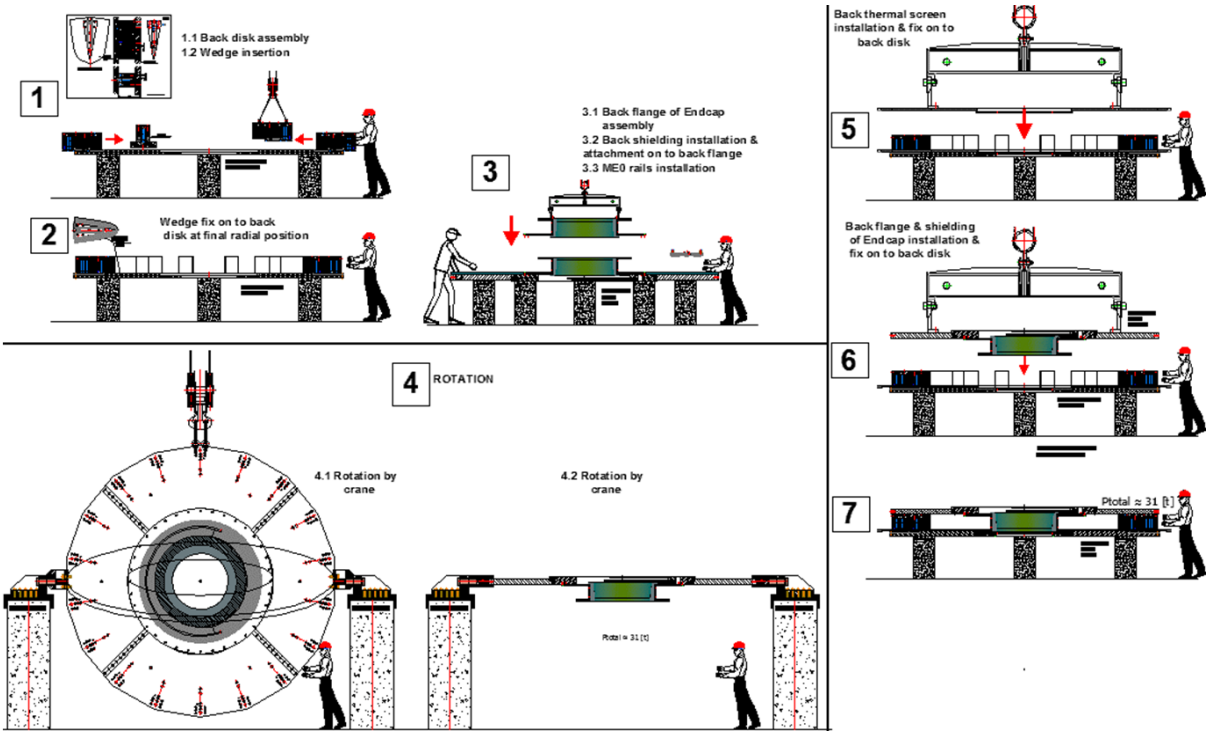


Figure 9.32: Assembly steps and tooling for the assembly of the back disk with the back flange and the sliding wedge system.

All components will be procured from industry, although it is not yet decided if the sliding wedge assemblies will be procured fully assembled or as individual parts that will be assembled at a collaborating institute. The back flange and back disk may be made of several steel wedges, which are bolted together to form full disks as illustrated in Figs 9.14 and 9.18. Each disk will be assembled at the vendor's site to ensure mechanical integrity and dimensional tolerances, and then disassembled for shipping to CERN. The support ring is machined as a single piece. Each component will be fully inspected and measured upon receipt and prior to assembly of the suspension system to ensure compliance with dimensional and material properties requirements.

The suspension system assembly steps are the following:

1. The back disk is assembled from the individual wedges, inspected and shimmed as necessary to ensure dimensional compliance, and placed on a set of concrete blocks.

2. The 18 sliding wedge supports are connected to the back disk. Their positions and angular orientation are precisely measured to ensure correct alignment of the back disk relative to the back flange once the latter is installed and to ensure adequate compliance of the radially sliding wedges to accommodate relative dimensional changes when the endcap is cooled to its operating temperature.

3. The back flange is assembled from the individual wedges, inspected and shimmed as necessary to ensure dimensional compliance and placed on a set of concrete blocks. The support ring is attached and the combined assembly is inspected and measured. The shielding and installation rails for the ME0 muon chambers are attached and the full assembly is inspected and measured.

4. The back flange assembly is rotated to the orientation required for connection to the sliding support wedge system.

5. The back thermal screen is attached to the back disk.

6. The back flange assembly is placed onto the back disk / sliding wedge assembly.

7. The 18 sliding wedge supports are positioned and bolted to the back flange.

8. A final set of measurements of all key dimensional parameters is made and adjustments are implemented if necessary to bring the assembly into compliance with all requirements.

9. The complete assembly is moved to the HGCAL assembly table (see below).

Stacking the absorber disks

The CE-H absorber disks are stacked on top of the completed back disk / back flange assembly utilizing a heavy tooling table as shown in the left side of Fig. 9.33. The main components of one CE-H absorber assembly are the following:

- 12 stainless steel disks of 68 mm thickness and 11 of 35 mm thickness of different diameters according to their position in the assembly;
- 276 stainless steel outer spacers of different sizes and shapes (see Section 9.3.2);
- Bolts and tie rods that secure the spacers made from a special high-strength steel alloy (see Section 9.3.2);
- 276 stainless steel inner pin-collet spacers;

The main pieces of tooling required for this stage are:

- heavy tooling table capable of supporting the ≈250 t mass of one completed CE, with bearings and hydraulic jacks to rotate the completed HGCAL from the horizontal assembly position to a vertical orientation for installation on the YE1;
- lifting fixture for the steel plates, which can be the same as that used for the assembly of the endcap suspension system discussed above;
- work platforms at the outer radius and inside the central hole in the absorber to allow access for installing spacers and bolts / tie rods.

All of the stainless steel absorber disks will be procured from a single vendor to assure that tight relative disk-to-disk tolerances can be maintained as required for the successful assembly of the CE-H. If possible, the spacers will be made the responsibility of the same vendor. The

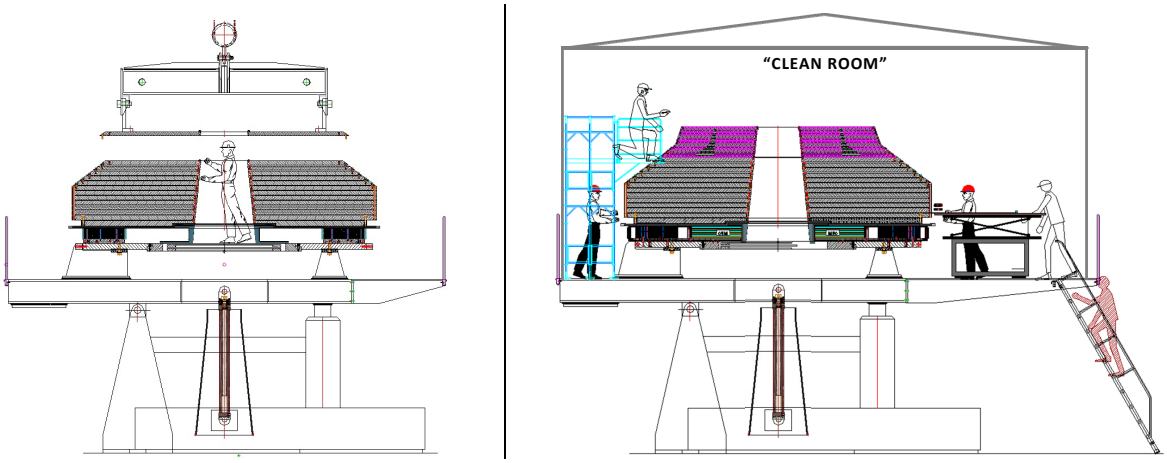


Figure 9.33: Left: Stacking of the absorbers for the CE-H. Right: Insertion of the CE-H cassettes and installation of electrical and optical services.

absorber layers may be made as complete steel disks or assembled from several wedges as illustrated in Fig. 9.14. A trial assembly of the complete absorber stack will be made at the vendor's site to assure that all tolerances have been met. The individual parts will be identified to permit an identical assembly at CERN. Rigorous quality control on the bolts and tie rods will be applied including material certification due to the high stresses experienced in them.

The CE-H absorber assembly sequence is the following:

1. The endcap suspension system is placed on the heavy tooling table.
2. The central cone of the thermal screen is placed in a compartment below the assembly table to facilitate its later installation.
3. The 12 inner and 12 outer spacers for the last CE-H absorber gap are installed, and the last CE-H absorber is place on top of them. All components are inspected and cleaned prior to installation.
4. The bolts that secure the spacers are tightened to a pre-determined torque and then secured against rotation.
5. The gap between the the two disks is measured by optical survey or with a mechanical "go-nogo" gauge to ensure that it is large enough over the full surface area to allow insertion of the cassettes.
6. The exposed surface is cleaned and inspected.
7. Steps 3–6 are repeated for the subsequent 22 absorber disks.
8. A complete inspection and detailed set of mechanical measurements is made to ensure compliance with all mechanical and cleanliness requirements.
9. The absorber stack is rotated to the vertical orientation (as shown in Fig. 9.37 for the completed endcap). A full survey is performed, and if non-conformities with respect to the geometric requirements are found, corrections to the assembly must be made.

Installation of the CE-H cassettes and services

The insertion of the CE-H cassettes and installation of the electrical and optical services require the following additional tooling:

- 2 movable assembly platform extensions;
- several under-the-hook lifting fixtures specifically designed to lift the cassettes (maximum mass of ≈150 kg each) using the lifting points built into the cassettes (see Section 9.2.1);
- 2 movable clean rooms;
- at least 2 adjustable-height tables for positioning the cassettes for insertion, including a system for driving the cassettes into the gap at the correct location;
- 2 systems for transferring the cassettes into the clean room and onto the insertion table;
- additional work platforms as required to facilitate the insertion and installation of services.

An extension of the work platform is required to provide the radial space to position the cassettes prior to insertion. It is planned to assemble both endcap calorimeters in parallel. The available space in SX5 requires that the two assembly tables be placed close together, but nevertheless allow the platform extension to be installed around the full circumference of both detectors at the same time. Therefore it is required to use a movable partial circumference platform extension as shown in Fig. 9.34. The example shown is for a 60° wide platform; 120° and 180° are also possible. A removable clean-room will be installed over the assembly as shown in the right side of Fig. 9.33.

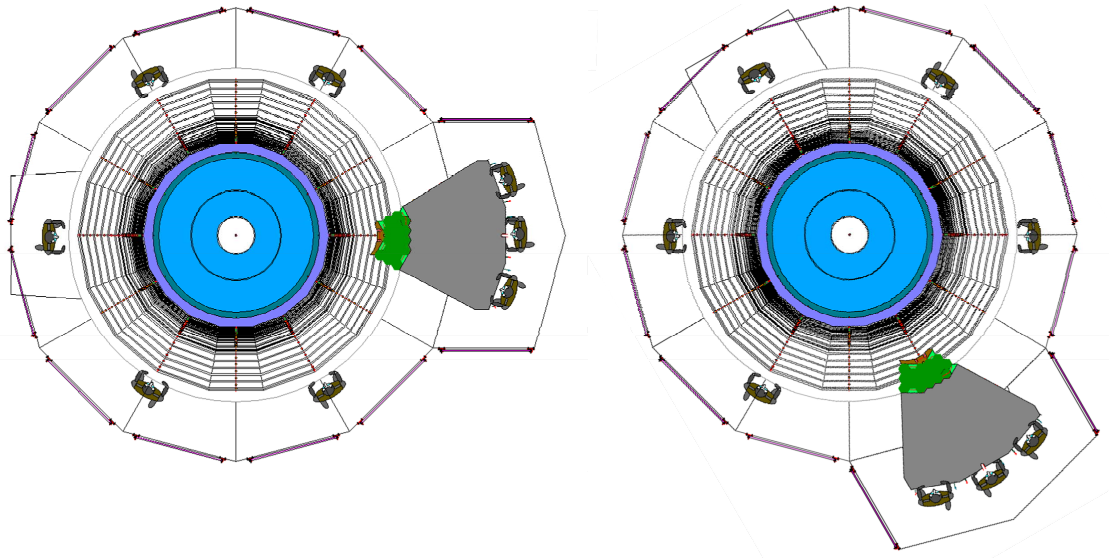


Figure 9.34: Cassettes insertion platform extensions on the two adjacent HGCAL assembly tables that required for insertion of the CE-H cassettes. Interlocking pairs of 30° cassettes that have been joined to form 60° insertion units are shown.

The cassettes will have undergone rigorous quality control during assembly and an extensive final testing programme that includes several thermal cycles between room temperature and −30 °C and a multi-day test run at −30 °C with all electronics operational and cosmic ray data

read out by a full DAQ system (Section 9.2.3). Each component will undergo a more limited set of tests and measurements at the assembly site prior to insertion into the CE-H absorber to ensure that no damage or degradation has occurred in transit from the production site.

Cassettes are inserted and all services connected sequentially in 60° sectors. This is done for all layers except for the first CE-H cassette, since the absorber layer behind which it sits is the CE-E back disk, which is not present at this point. This last layer will be installed after the CE-E is mounted to the front of the CE-H. After the connection of the services to each cassette pair, the integrity of the connections and the functioning of the cassette is verified, since the services of subsequent layers will lie on top of and "trap" that cassette. The CO_2 cooling manifolds cannot be installed at this stage since they cover the full thickness of the HGCAL including the CE-E, which is not present at this point. The sequence of operations for the insertion of the CE-H cassettes and installation of electrical and optical services in each 60° segment is the following:

1. The movable assembly platform is positioned and connected to the main assembly table, the movable clean room is installed, and clean-room conditions are established.
2. The three ME0 muon chambers for the 60° sector are installed, connected, and checked by the muon system team.
3. The DC-DC converters and patch panels for the HV and optical services are installed in the gap between the back disk and the back flange, between the sliding support wedges, and outside ME0.
4. Working from the back (bottom) to the front (top), the cassettes for one 60° sector are installed. For each the following steps are taken:
    - A pair of 30° cassettes are brought into the movable clean room, placed on the insertion table, and joined to form a 60° insertion unit;
    - The joined pair is inspected and carefully measured for compliance with dimensional requirements.
    - The cassette pair is inserted into the absorber gap, using guides in the absorber that interlock with corresponding features built into the cassettes. The insertion force is carefully controlled and monitored and the operation stopped if there is any indication of interference with the absorber surfaces (gap too small), with the spacers, or with adjacent already-installed cassettes (misalignment of insertion). The insertion is visually monitored from both the outside and inside radii.
    - The cassette pair is secured to the adjacent absorber layer.
5. Working from the back to the front, for each 60° cassette pair:
    - the electrical and optical services are connected to the cassettes;
    - they are routed according to the cabling plan to the back of the detector and placed in the appropriate channels in the cold-warm feedthrough;
    - the LV cables are attached to the appropriate DC-DC converter, and the HV and optical cables to the appropriate patch panel;
    - electrical tests are done to verify the integrity of all connections and proper functioning of the cassette;
    - thorough mechanical measurements are made to ensure that all of the electrical and optical services lie within their allowed envelope.

This sequence is repeated for subsequent sectors until all cassettes have been inserted and their electrical and optical services installed.

9.6.3 Final assembly

In this stage, the CE-E is mounted on the front of the CE-H, the CE-E electrical and optical services are installed, and the CO_2 cooling manifolds for the whole endcap are installed; the thermal screen is installed and the cold-warm feedthroughs are closed; CO_2 cooling is connected and a complete cold test is performed.

The mounting of the CE-E on the CE-H is illustrated in the left side of Fig. 9.35 and the connection of electrical, optical and cooling services is shown in the right side of the same figure. No additional tooling beyond that required for previous steps is required. The system for mounting the CE-E to the front of the CE-H is similar to that used to connect one CE-H absorber to the next and a similar mounting procedure is used. Once the CE-E is attached to the CE-H, the assembly sequence outlined below is followed for each 60° sector:

1. The movable cassettes insertion platform (Fig. 9.34) is positioned and connected to the main assembly table, the movable clean room is installed, and clean-room conditions are established.
2. The front-most CE-H cassette pair is inserted following step 4 of the CE-H cassettes and services installation procedure.
3. Working from the front-most CE-H layer to the front of the CE-E, the electrical and optical services are installed and checked following step 5 of the CE-H cassettes and services installtion procedure.
4. The cooling manifolds – separate supply and return lines for even and odd layers in each 30° sector – are installed and checked.
    - Supply and return tubes are connected to each of the cassette layers, and placed in their final radial and azimuthal positions across the surface of the CE.
    - A transition from separate supply and return lines to coaxial vacuum-jacketed lines is pre-made as part of the manifold construction and the coaxial section is placed in the appropriate channel in the cold-warm feedthrough.
    - Each manifold is leak and pressure tested.
    - A thorough dimensional check is done to ensure that each manifold lies within the boundary defined by the inner surface of the thermal screen.
5. A complete set of electrical measurements is made to verify the integrity of all electrical and optical connections.

Electrical, optical and cooling services for the endcap timing layer are installed after the HGCAL services are complete. Their exit point from the front of the HGCAL is positioned to facilitate later installation of the TE outside the HGCAL thermal screen.

After all services have been installed, a thorough inspection is made to ensure that all (visible) elements are correctly placed and that no foreign objects are present. A complete laser scan of the entire assembly is done to ensure that all elements lie within the boundary defined by the thermal screen.

The outer ring that seals the "bell" section of the thermal screen to the feedthrough section is installed. The feedthrough channels are potted and leak checked.

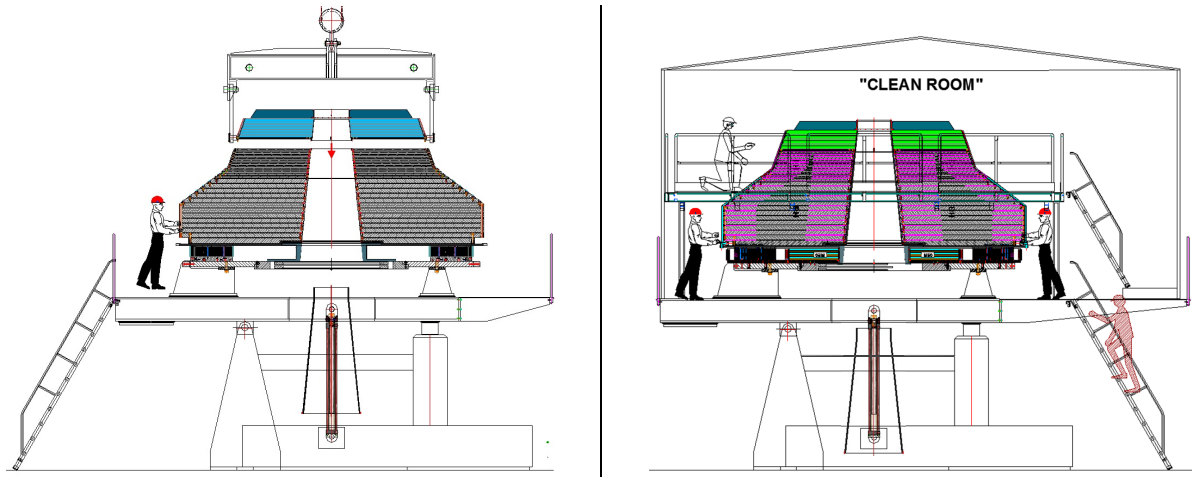


Figure 9.35: Left: Installation of the completed CE-E on top of the CE-H. Right: Installation of CE-E electrical and optical services and installation of cooling manifolds.

The inner cone of the thermal screen is put in place and sealed to the back thermal screen disk and to the Permaglas support cylinder at the front. The bell section of the thermal screen is installed as shown in Fig. 9.36. Given the relative flexibility of the thermal screen structure and the close spacing between its inner surface and the outer surface of the services, special tooling may be required for this step, possibly including an external stiffening frame that is not shown in Fig. 9.36. The rim of the bell is sealed to the feedthrough sealing ring at the back of the HGCAL and the seal is leak checked. Openings in the front disk of the thermal screen that accommodate the TE services are closed by feedthrough panels, the feedthroughs are potted, and all joints leak checked. A final leak test of the thermal screen is performed using the dry gas system. If the outflow rate exceeds specifications, a campaign of leak checking and sealing will need to be undertaken until the entire thermal screen is adequately leak tight.

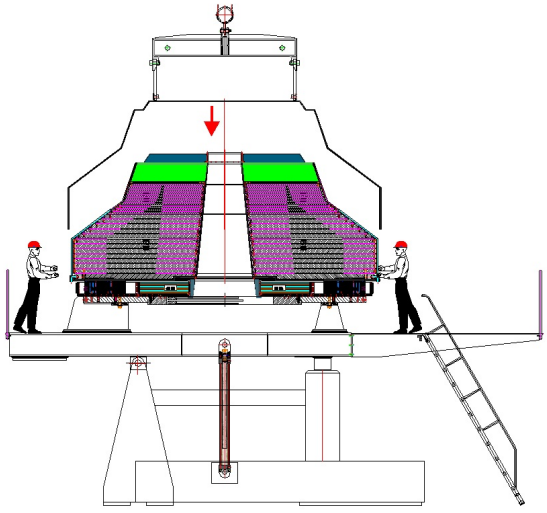


Figure 9.36: Installation of the thermal screen.

The dry gas system is then used to establish a low-humidity environment inside the thermal screen with the goal of reaching a dew point below $-40\,^{\circ}$C. One or several 60° sectors are connected to the CO_2 refrigeration system in SX5 and the detector is cooled to $-30\,^{\circ}$C. During the cooldown, humidity and temperature throughout the HGCAL will be monitored to ensure that there are no large temperature gradients that could induce unacceptable mechanical stresses. Stresses in the back disk, back flange and wedge supports will be monitored to ensure that

the sliding wedge system is operating as required. Cooldown will be stopped and corrective actions taken if any parameters are measured to be outside the acceptable range.

Once the detector is cold, the actively cooled sectors will be powered and a thorough set of tests, including measurements with cosmic rays, will be conducted to pre-commission the HGCAL. All the required safety systems will be in place for these tests. Without warming the detector, the cooling will be transfered to another set of sectors and the pre-commissioning tests performed on them. This will be repeated until all detector elements have been fully exercised. The HGCAL will then be warmed to room temperature, and the external connections to the services disconnected, leaving the endcap ready for transport from SX5 to UXC.

9.6.4 Installation and tooling

Transfer of the completed endcap is the responsibility of CMS Technical Coordination, and only a brief sketch of the procedure and required tooling is provided here.

The assembly table includes hydraulic jacks and a set of bearings that allow the completed endcap to be rotated to the vertical orientation required for installation onto the YE1 nose, as shown in the left side of Fig. 9.37. A support cradle, which includes screw jacks and an airpad system, is placed under the endcap. It is raised until the full weight of the endcap is taken by the cradle, at which point the endcap is disconnected from the assembly table. A specially design spreader bar is used to hoist the endcap from the surface to the UXC. An existing hydraulic riser system is used to position the endcap for installation onto the YE1, as shown in Fig. 9.38.

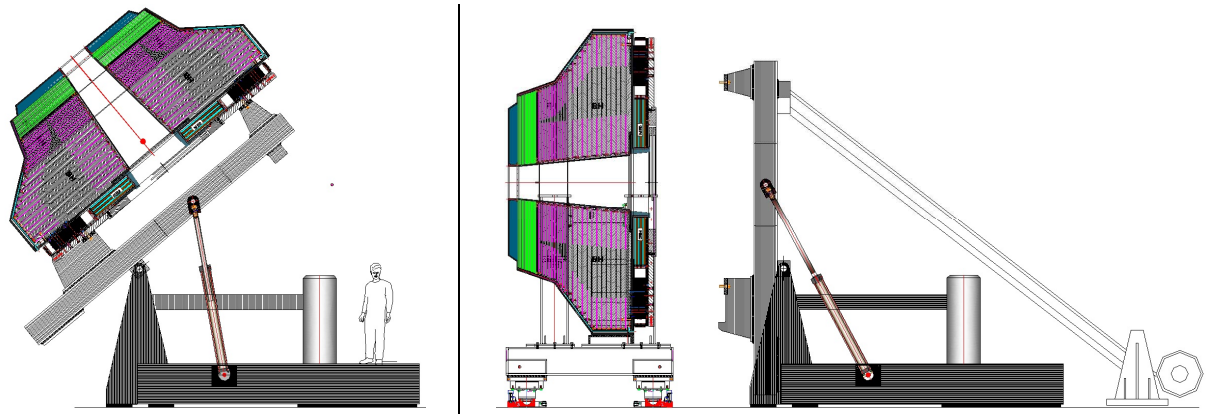


Figure 9.37: Rotation of the completed endcap calorimeter from the horizontal assembly orientation to the vertical installation orientation and transfer to the support cradle.

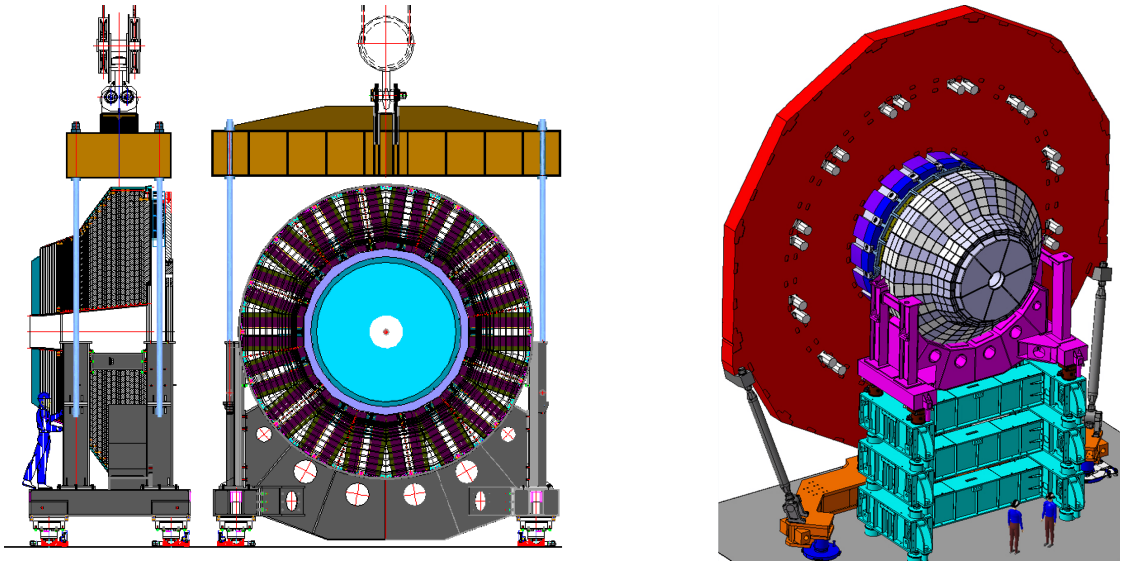


Figure 9.38: Left: Endcap support cradle and spreader bar used for lowering the HGCAL from the surface assembly building into the experimental cavern. Right: Hydraulic riser system used to elevate and position the HGCAL for connection onto the YE1.

Chapter 10

Reconstruction and detector performance

10.1 Reconstruction and detector performance

10.1.1 Simulation model

The detailed implementation of the detector geometry necessarily follows after technical design choices. In autumn 2016, before the making of many of the important engineering decisions that are described in this document, it was decided to freeze the geometry implemented in the CMS simulation and reconstruction software, CMSSW, to allow simulation work to proceed. The frozen detailed calorimeter geometry closely follows what was described in the Technical Proposal [1], except that stainless steel has been substituted for the brass absorber. A tungsten/copper-silicon sampling electromagnetic calorimeter, the EE, with a depth of about $26\,X_0$ and $1.5\,\lambda$, is followed by a stainless-steel-silicon hadron calorimeter, $3.5\,\lambda$ deep, the FH, which is followed by a $5\,\lambda$ stainless-steel-scintillator sampling backing calorimeter, the BH, to bring the total calorimeter depth to about $10\,\lambda$, as measured perpendicular to the detector layers. The depth measured along trajectories from the interaction point is between 0.5% and 10% greater, depending on pseudorapidity.

Both the names (EE, FH and BH, referring to the electromagnetic, forward hadronic and backing hadronic calorimeters) and the structures they describe belong to the Technical Proposal era. Modifications have been made to what is described in the Technical Proposal: the brass of the FH and BH is replaced by stainless steel, and the scintillator of the BH is segmented into η, ϕ cells of size 0.0174×0.0174 (i.e. 1°), individually read out and digitized by an electronics sequence that simulates the SiPM on scintillator-tile configuration described in Section 2.3 for the scintillator parts of CE-H.

There are 28 active layers in the EE, and 12 active layers in each of the FH and BH. The silicon sensors in the EE and FH are arranged in 6″ hexagonal wafers, divided into hexagonal cells. The cell sizes and wafer thicknesses follow those given in the Technical Proposal, and are summarized in Table 10.1.

The simulation of the signal development and processing in the HGCAL electronics includes the addition of Gaussian noise, and digitization. The magnitude of the noise for the Si cells is given in Table 10.1. The highest noise levels are seen in the $100\,\mu$m cells. The detector design described in this document uses $120\,\mu$m, rather than $100\,\mu$m, sensors in the region of highest neutron fluence (Section 2.1), and is thus expected to have somewhat lower noise levels both initially and after $3000\,\text{fb}^{-1}$. The noise simulated for the SiPM on scintillator-tile cells corresponds to $1/7$ MIP at start of life, and $1/5$ MIP after $3000\,\text{fb}^{-1}$. A threshold equivalent to 0.5 MIP is applied to the stored reconstructed hits.

Table 10.1: Silicon sensor arrangement in the simulated geometry: thickness of active silicon layer in the EE and FH, with the associated cell size and S/N for a MIP.

Thickness	300 μm	200 μm	100 μm
EE region	$R > 120$ cm	$75 < R < 120$ cm	$R < 75$ cm
FH region	$R > 100$ cm	$60 < R < 100$ cm	$R < 60$ cm
Cell size (cm^2)	1.05	1.05	0.53
Initial S/N for MIP	15.2	7.6	3.7

10.1.2 Clustering

For the first prototype of the reconstruction we start by constructing two-dimensional (2D) clusters in each of the 52 layers of the calorimeter and then, in a second step to collect these layer clusters into a three-dimensional cluster (a "multicluster") by associating layer clusters that are approximately projective, i.e. they lie approximately on a straight line pointing from the CMS detector centre. The 2D clustering algorithm is inspired by Ref. [30], and is described below. The algorithm has been implemented using a k-d tree space-partitioning data structure. Th k-d tree is used for nearest neighbour searches because of its computational speed. This allows the clustering algorithm to be run rapidly, even in events with the occupancy resulting from an average of 200 interactions per bunch crossing. The clustering is performed on reconstructed hits intercalibrated using the relative $\mathrm{d}E/\mathrm{d}x$ thicknesses of the absorber plates, with the energy scale set so that summing all hits resulting from an unconverted photon shower gives the energy of the incoming photon.

The 2D clustering algorithm is run on each of the layers of the calorimeter, and identifies local concentrations of energy in the layer resulting from showers. A threshold of $3\,\sigma_{\mathrm{noise}}$ is placed on cells, and for each cell above the threshold an energy density is calculated as the sum of the signals in the cell itself and all other cells within a distance, d_c (taken as 2 cm in EE and FH, and 5 cm in BH), which is then taken as the radius of a unit area. After the density has been calculated for all cells, then, for each cell, the distance to the nearest cell with higher density is calculated. Cells with both density and distance above some thresholds generate (i.e. seed) a new cluster, and the remaining cells are associated with the closest seed. The density threshold is set at $9\,\sigma_{\mathrm{noise}}$, and the distance threshold is taken as d_c.

It is likely that, in future, use of a more sophisticated definition for the density will refine the two-shower separation of the algorithm. A suitable definition of the density of a particular cell, might be, for example: $\rho = \Sigma \mathrm{E}_i \exp(-r_i/\kappa)$, where E_i is the energy of cell i at distance r_i from the particular cell, and κ is some distance parameter related to the shower lateral size, (for computational reasons i would be limited to cells within some cutoff radius).

In a second step the layer clusters are collected together and built into three-dimensional clusters ("multiclusters") by associating layer clusters in different layers that lie approximately on a straight line pointing from the CMS detector centre. For the multiclusters that are used to form electron and photon candidates, the alignment requirement is 2 cm. The layer clusters are dealt with in order of decreasing magnitude. A seed direction is formed by a line from the CMS detector centre to the layer cluster position, and any other layer cluster, in any layer, found within 2 cm of that line is added to the seed to form a collection of layer clusters that constitutes a multicluster. The layer clusters so used are marked as having been used, and are not considered for further use in the building of multiclusters.

The multiclusters are fed into the existing CMS particle flow system.

10.1.3 Direct 3D clustering

The same imaging algorithm that is described in the previous section can be extended to cluster directly in three dimensions. The cells are first assigned positions in a 3D (x, y, z) space. Since the sensitive cells of the HGCAL are non-projective, one simple way of doing this is to project their η, ϕ positions to a chosen plane (for example the first layer of the calorimeter) and assign the x, y position in that plane (i.e. creating a pseudo-projectivity). The z position assignment used in the first tests of this algorithm was simply the true geometric z position, but probably more optimized choices involving scaling to the absorber thickness in terms of the energy loss, $\mathrm{d}E/\mathrm{d}x$, between the cells might be useful, since the distance metric affects the algorithm behaviour.

The density is defined using a 3D kernel, and the algorithm proceeds in the same way as the 2D algorithm.

Figure 10.1 shows a display of two pions, separated by 30 cm clustered by the direct 3D clustering algorithm. This work is still exploratory and the algorithm has not been studied in a full pileup environment. Detailed optimization needs to be done before having a working reconstruction system using direct 3D clustering.

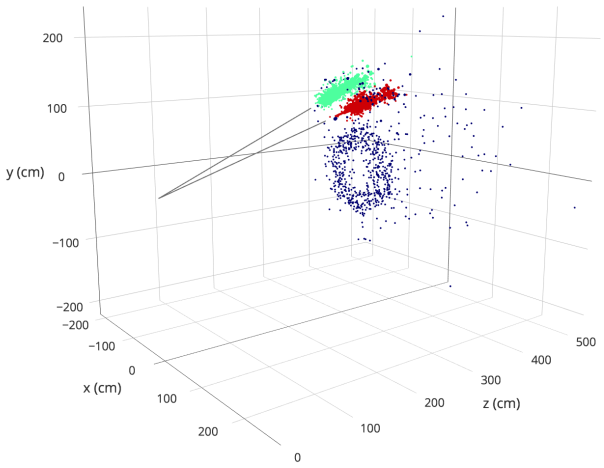


Figure 10.1: Direct 3D clustering of a charged pion shower in HGCAL. The hits included in the reconstructed clusters of two pions, at $\eta = 2$, separated by 30 cm, are shown by red and green dots. The dark blue dots are unclustered hits from noise.

10.1.4 Reconstruction using machine learning

A preliminary exploration of the application of machine-learning techniques to HGCAL reconstruction and particle identification has been made. It is presented here as an illustration of the potentialities of the rich and detailed information provided by the calorimeter. Exactly how such techniques could best be exploited in the context of a complete detector reconstruction program is a question for future study.

Deep neural networks (DNNs) have brought significant advances to a variety of fields in the last decade, also enabled by developments in computing hardware. In contrast to shallow neural networks or standard boosted decision trees, widely used in high-energy physics for classification problems, deep neural networks allow for designs that efficiently exploit the struc-

ture of the data and can therefore process input having a large dimensionality. In this context, the sensor signals from the HGCAL can be interpreted as 4-dimensional energy images or 3-dimensional images with an energy and a time colour dimension.

To limit the number of inputs to the DNN, a cone in $\Delta R = 0.1$ around each particle entering the first HGCAL layer is defined based on MC truth information. However, the same approach can be applied using e.g. local energy maxima or multiclusters as seeds while extending the network to identify those seeds not associated to a real particle.

Typically, computer-vision networks consist of a set of convolutional layers [56] that exploit the translation invariance of the input data. In each network layer, a kernel covering a selected number of adjacent squared pixels is moved over the image to find structures such as edges or areas, requiring a homogeneous spatial and colour resolution, which contradicts the HGCAL design. Therefore, the 3D image with uniform pixel size is built in a first step by overlaying a coarse squared grid on each HGCAL layer with up to 6 cells in each pixel. A small DNN with two layers consisting of 32 and 16 nodes collects the time and energy information from all contributing cells taking into account their relative position within the pixel. The third dimension is defined by the HGCAL layer number. The resulting uniform 3D image is fed through two blocks of convolutional networks in parallel. One consists of 4 layers with kernels of similar size as typical hadronic showers in η and ϕ but shorter in the number of covered HGCAL layers, and the other comprises 3 kernels very narrow in ΔR, spanning over a larger range with respect to the shower depth, optimised for muon reconstruction. The output of both is merged and fed through two dense network layers with 64 and 32 nodes. The final output nodes predict the particle type and the particle energy.

The network is trained using electrons, muons, photons and charged pions generated with uniform p_T between 2 and 100 GeV, restricted to $2.3 < |\eta| < 2.5$, and flattened in energy. Photons that convert into an e^+e^- pair within the first 200 cm in z direction are discarded. The sample consists of 600 000 particles with 0 PU and 400 000 particles with 200 PU. Between 4 and 12 particles are generated per event and endcap, resulting in partially overlapping showers. The sample used to evaluate the performance is separate from the training sample.

The energy resolution achieved by the current network for hadron showers from charged pions is shown in Fig. 10.2.

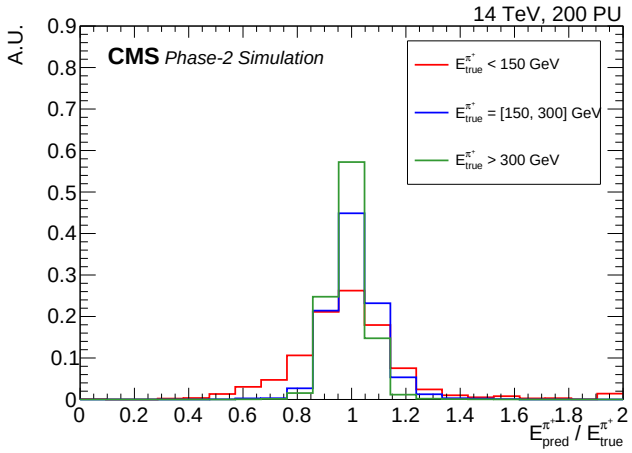


Figure 10.2: Energy resolution for charged pions using only HGCAL information. The last bin includes overflow.

Since the scale of the total energy deposits for those particles can differ by orders of magnitude, the DNN must provide a good discrimination performance to achieve such resolution. As shown in Fig. 10.3, showers stemming from charged pions and muons can be well disentangled, in particular for high muon transverse momenta. Also more differences in the shower shapes of electromagnetic showers and hadronic showers can be detected well.

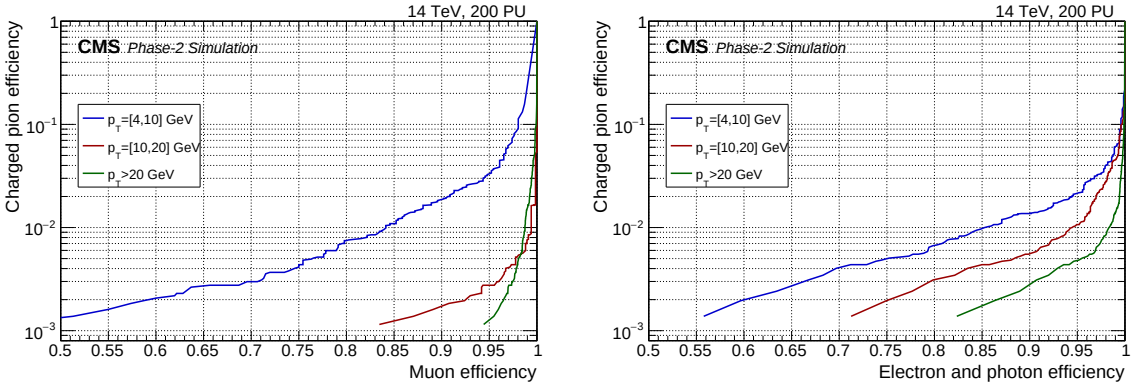


Figure 10.3: Efficiency for the discrimination of a muon from a charged pion (left) and for the discrimination between an electromagnetic shower from an electron or photon and showers from charged pions (right).

10.1.5 SimClusters

It has been found useful to also create "SimClusters", which are three-dimensional clusters of calibrated RecHits (i.e. just like the multiclusters) where the *pattern recognition* is done using Monte Carlo truth information, and each SimCluster corresponds to a simulated particle at the front face of the calorimeter. After a superclustering step for photons and electrons, and a track matching, and particle interpretation step, again using Monte Carlo truth information, the resulting candidates can be fed into the existing CMS particle flow and physics object reconstruction system. The result of such a procedure is intrinsically optimistic about some of the performance characteristics to be expected of a full reconstruction not using Monte Carlo truth information, most notably about the ability to distinguish and separately cluster overlapping showers. To better simulate performance on data, algorithms have been introduced to merge showers that must be considered indistinguishable, and to share deposited energy between distinguishable but nearby showers in a realistic way. These algorithms are based on the relative magnitude of energy deposited in a layer by the overlapping showers, and the distance between them.

10.1.6 Electromagnetic energy resolution

The electromagnetic energy resolution of the HGCAL has been studied using a GEANT simulation in which the geometry of the calorimeter is as described in the engineering sections of this document, both in terms of the longitudinal structure and the lateral cell sizes. An exception to this is that the sensors with a sensitive thickness of 120 μm are modelled by 100 μm thickness silicon (together with the higher noise and lower charge collection of 100 μm).

The response to photons which do not convert in the tracker material was studied, both in the absence of pileup, and with averages of 140, and 200 pileup events per bunch crossing. An investigation was made of varying size of the region used to estimate the energy. Results for a mean of 200 pileup events per bunch crossing have been shown in Section 5.1.3.

Samples of fixed p_T were used (p_T = 5, 10, 20, 30, 40, 60, 80, 100, 150, and 200 GeV) at three values of η ($\eta = 1.7, 2.0, 2.4$) giving showers exclusively in 300, 200 and 100 μm sensors, respectively. The energy resolution at each point, σ/E, was obtained from a Gaussian fit to the reconstructed energy distribution, where the energy was estimated from the sum of cells within a fixed radius, r, of the maximum cell, for a range of different values of r ($r = 13, 20, 26, 53$ mm). The fitted energy distribution for $p_T = 60$ GeV photons, at $\eta = 2$ (i.e. $E = 225.73$ GeV), where the energy is estimated using $r = 26$ mm, is shown in Fig. 10.4.

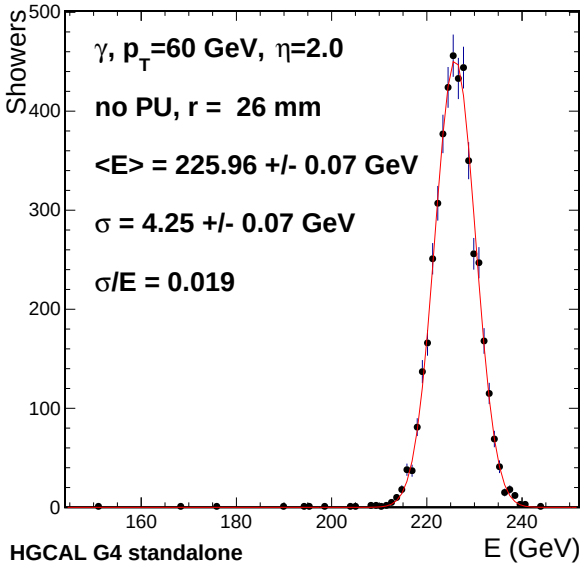


Figure 10.4: Gaussian fit to the energy distribution for $p_T = 60$ GeV photons, at $\eta = 2$, where the energy is estimated using $r = 26$ mm.

A small correction is applied for the leakage of high energy showers from the rear of the electromagnetic section of the calorimeter. The total energy reconstructed in the electromagnetic section of the calorimeter is plotted versus the fraction of the reconstructed energy found in the last four layers, F_4, and the mean value of the fraction in each bin is fitted with a straight line. This is shown in Fig. 10.5. The Monte Carlo event sample used for this is the one with the highest energy ($p_T = 200$ GeV, $E = 1111.4$ GeV). The reconstructed energy is then corrected with the fitted function, event-by-event, according to the value of F_4.

The following procedure is adopted to fit the energy resolution $\sigma/E = a/\sqrt{E} \oplus b/E \oplus c$: first the samples with no pileup are fitted for the stochastic and constant terms (a and c, respectively), with the noise term (b) fixed to the measured value. The energy resolution values obtained from pileup samples are fitted with the constant term fixed at the zero pileup fitted value, and the noise term allowed to vary. Fits to the resolutions obtained at $\eta = 2$, using $r = 26$ mm, in absence of pileup, and with an average of 200 pileup interactions per bunch crossing, are shown in Fig. 10.6.

The procedure results in good fits, and provides a measure of the "pileup noise" for the different η values, choices of r, and pileup intensities. The stochastic terms fitted in the case of pileup are similar to those in the zero pileup fits, although they all show a small increase, slightly larger than the statistical uncertainty on the fitted values, as can be seen in Fig. 10.6. As expected the stochastic term is larger in the regions using a thinner active layer of silicon; and it increases as r is decreased. Taking a very small region in which to estimate the energy ($r = 13$ mm) increases the constant term also (to ≈4%), but for $r \geq 20$ mm in the 100 μm wafers (with small cells), and

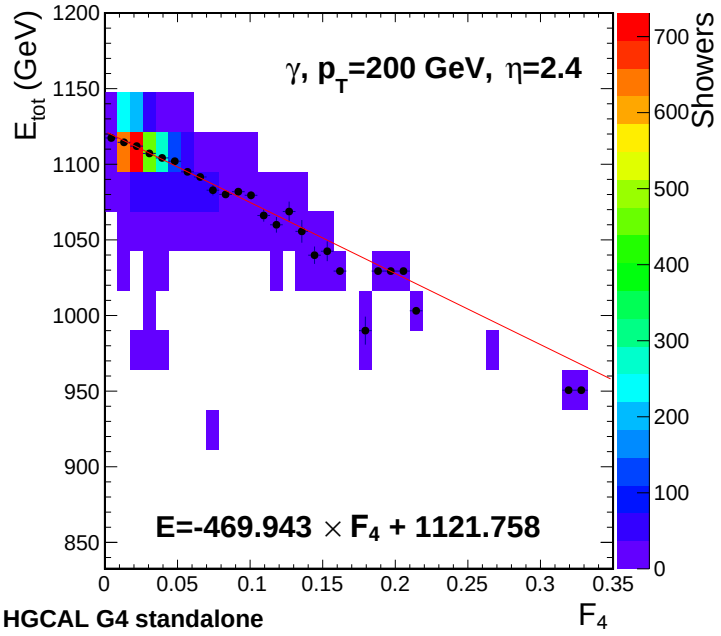


Figure 10.5: Total energy reconstructed in the electromagnetic section of the calorimeter, measured as a MIP count, is plotted versus the fraction of the reconstructed energy found in the last four layers, F_4, together with a straight line fitting the mean value of the total energy in each F_4 bin.

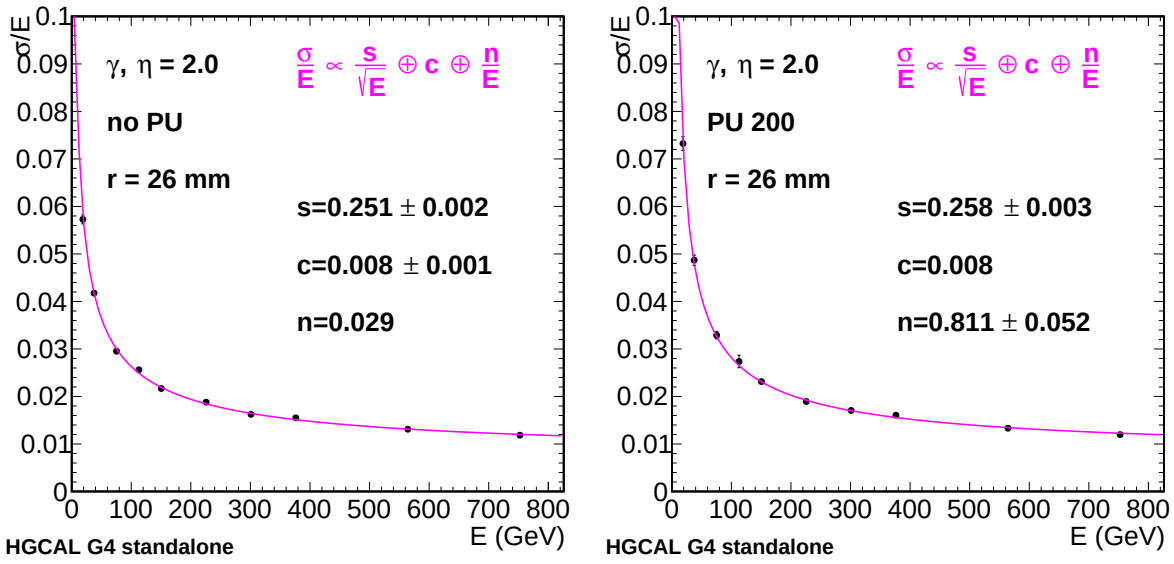


Figure 10.6: Fits of the energy resolution function to the resolutions of photons, at $\eta = 2$, where the energy is estimated using $r = 26\,\text{mm}$, for (left) samples where no pileup is present, and (right) when an average of 200 interactions per bunch crossing are present.

$r \geq 23$ mm in the 200 and 300 μm the constant term is fairly stable. For $r = 26$ mm the stochastic term of the fitted energy resolution functions are 24%, 25%, and 28%/$\sqrt{E}$, respectively, in the regions of 300, 200, and 100 μm silicon. The constant term is about 0.8% in the 300 and 200 μm silicon, 0.5% coming from intercalibration uncertainty (3% used in the simulation), and the variation of energy containment in the cells included in the radius which varies depending on the location of the shower axis in the central, maximum energy, cell. In the 100 μm silicon, with smaller cells, the constant term is 0.5%. For the $r = 26$ mm energy estimation radius, the "pileup noise" contribution resulting from an average of 200 pileup interactions per bunch crossing is 0.49, 0.81, and 2.2 GeV at $\eta = 1.7, 2.0$, and 2.4, respectively.

10.1.7 Muon identification

The region $|\eta| > 2.4$ is covered exclusively by silicon sensors. In this region, where the ME1/1 chambers are not available but tracking coverage is provided by the HL-LHC pixel tracker, a good efficiency for muon identification can be provided by the HGCAL, in addition to the new ME0 muon chambers that cover the pseudorapidity range $2.0 < |\eta| < 2.8$.

Detailed studies have shown that the HGCAL can achieve good muon identification performance by placing requirements on the signals in cells in a small region around an extrapolated charged-particle track. The results of this approach have been presented in Section 5.1.9. The robustness of the results presented there have been investigated by decreasing the S/N(MIP) ratio to 3, and also to 2, in the last 12 layers of the calorimeter, where >5 is expected even for the worst affected cells after 3000 fb^{-1}. The results of this study are shown in Fig. 10.7. The efficiency is shown a function of η in Fig. 10.8, for a typical working point where the average efficiency is about 98%. The distribution is very similar for all points on the efficiency plateaus shown in Fig. 10.7.

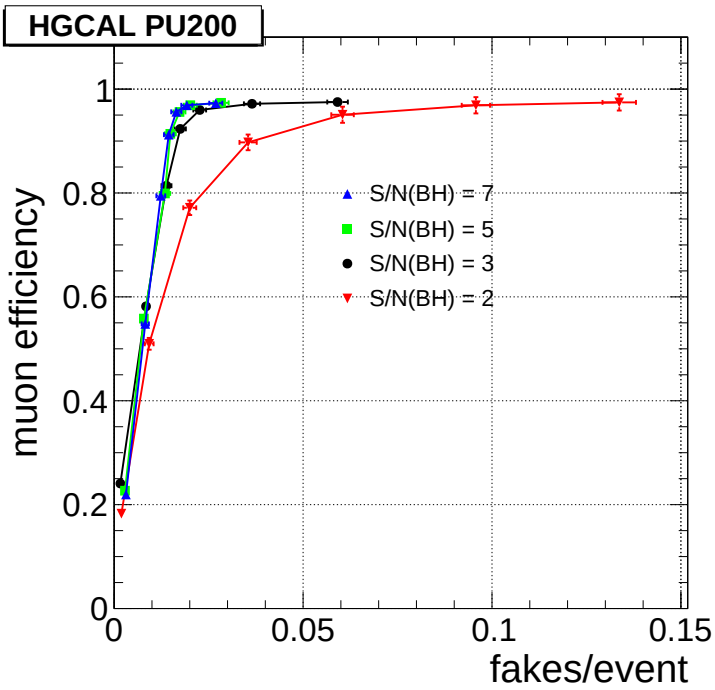


Figure 10.7: Muon identification efficiency versus the probability per event that another, non-muon, track is identified as muon, under different noise conditions for the last 12 layers of the calorimeter: S/N(MIP) = 7, 5, 3, and 2.

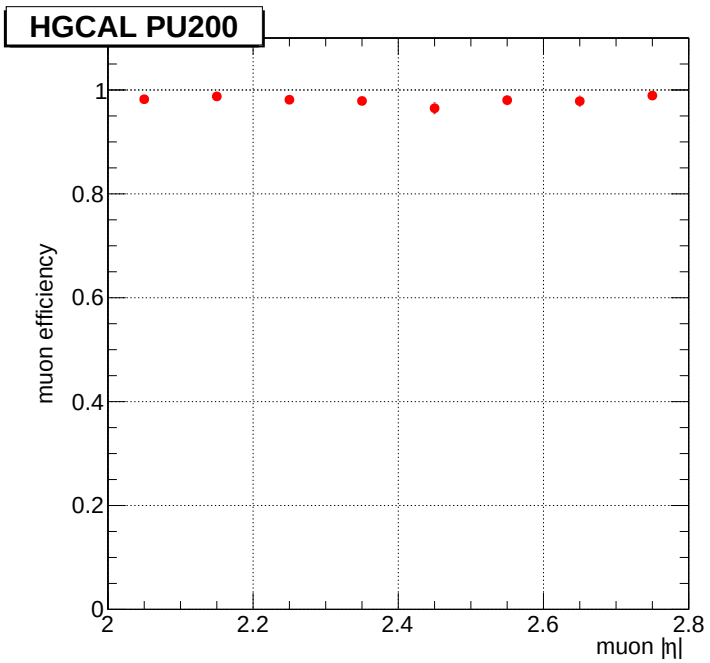


Figure 10.8: Muon identification efficiency versus pseudorapidity, for muons with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$.

The effect of additional signals from the numerous slow neutrons present at high luminosity has also been estimated in two ways: firstly by a calculation from first principles, and secondly using a dedicated CMSSW simulation.

For a neutron interaction in a sensor to be detected the equivalent energy of 0.5 MIP needs to be deposited, or ≈7300 electrons in a 200 μm thick silicon sensor. The energy required to generate a detectable signal is thus about 26 keV. At low energies (<2 MeV) the reaction that can produce such a signal is elastic scattering of a neutron on the silicon nucleus. The maximum energy transfer to the Si nucleus is 0.133 times the neutron kinetic energy, corresponding to the case where the neutron is scattered backwards. Approximately half of the recoil energy of the nucleus is lost as non-ionizing energy loss. Hence only neutron energies above ≈400 keV can give a 0.5 MIP signal in a silicon layer. Above 1.85 MeV the dominant process is neutron absorption, where the neutrons de-excite through the emission of photons that rarely deposit enough energy in the silicon to be detectable.

Simulation using the FLUKA program (further details can be found in Chapter 12) indicates that each pp-collision contributes 1.95×10^{-3} neutrons $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ to the neutron flux in the energy range 200 keV to 2 MeV. Assuming that the neutron flux diffuses with a time constant of $\sim 1\,\mathrm{s}$ the frequency of interactions in each silicon cell will be given by:

$$f = F \times N_{bx} \times N_{PU} \times (N_A/N_{Si}) \times \rho_{Si} \times V \times \sigma_n \tag{10.1}$$

Where F is the flux contribution from a single pp-collision, N_{bx} the number of bunch crossings per second, N_{PU} the number of interactions per bunch crossing, N_A is Avagadro's number, N_{Si} the mass number of silicon, ρ_{Si} the density of silicon, V the volume of a cell, and σ_n is the neutron-Si cross section in this energy range, which is ≈3 barn [57]. Taking a 200 μm thick cell of area 1 cm^2, and an average of 200 interactions per bunch crossing, gives 4.7×10^4 neutron interactions per second in each cell, resulting in a probability of 1.2×10^{-3} of such an interaction

coinciding with any particular bunch crossing.

If the average energy transfer is assumed conservatively to be 10% of the maximum energy transfer, then the apparent signal probability due to neutron interactions in each silicon cell will be $\approx 1 \times 10^{-4}$.

The dedicated CMSSW simulation used a dedicated neutron cross section physics list and setup such that all particles were tracked by GEANT for 1 second after their production. The number of reconstructed hits above the threshold of 0.5 MIP, in events where the mean number of interactions per bunch-crossing is 200, increases by 0.4% when the dedicated setup is used, confirming that the impact of neutrons is negligible.

10.2 Beam tests

10.2.1 Overview and goals

In order to validate the overall design of the HGCAL, a set of beam tests was performed in 2016 and 2017. Prototype hexagonal detector modules, representative of the CE-E and silicon part of CE-H, were designed and built using hexagonal silicon sensors from HPK. These were complemented by a modified version of the CALICE AHCAL, to represent the scintillator/SiPM part of CE-H.

Section 10.2.3 describes the silicon modules constructed, the AHCAL, the DAQ chains developed and the mechanical structures used for the tests, as well as the complimentary data taken at FNAL and CERN. Section 10.2.4 describes the calibration procedures, whilst Section 10.2.5 shows some results, complementary to those described in Section 5.2.

10.2.2 Front-end ASICs used in beam tests 2016 and 2017

For the 2016 tests an existing front-end ASIC - SKIROC2 - was used. SKIROC2 was developed for the CALICE collaboration with silicon-based calorimetry in mind. The chip offers a rich functionality of which only a fraction was used. Each SKIROC2 has 64 channels, with each channel having a preamplifier and two separate slow shapers, a fast shaper, self-trigger and fifteen-cell pipeline, as well as a 12-bit ADC. Two Skiroc2 were used per module. Only the slow shapers were used in our system as we utilized an external trigger (the fast shaper is used for self-trigger). The two shapers have a fixed gain ratio of 10:1, providing low- and high-gain amplification and hence a large dynamic range with, at the same time, good accuracy for small signals from single particles. For every trigger received, both high and low gain signals were read out for every channel.

For 2017 a new ASIC was available: the SKIROC2-CMS. This evolution of the SKIROC2 provided 64 channels with 40 ns shaping time and an internal buffer to store 13 consecutive 25 ns voltage samples, as well as an extended dynamic range provided through two gain stages and a novel "time over threshold" (ToT) system. It also provided a fast-timing circuit, with 50 ps precision.

10.2.3 Prototype construction and operation

Each silicon detector module is a glued stack comprising a baseplate, a polyimide gold-plated foil, a silicon sensor, and one or two PCBs. The baseplates for the modules in the CE-E prototypes are made from WCu to give an overall compact calorimeter, whilst in the CE-H modules a simple Cu baseplate was used. The function of the baseplate is threefold: it provides mechanical rigidity to the module; it supplies a thermal pathway from the heat load of the ASICs

on the PCB, through the silicon to the copper cooling plate supporting the module; and it is part of the calorimeter absorber. The function of the gold-plated polyimide sheet is to isolate the silicon from the metallic baseplate, but also provides an electrical connection to the silicon for the bias voltage. The PCB glued to the silicon has gone through several iterations but have one important feature in common: they contain many holes, through which the silicon cells are connected to the PCB via aluminium wire bonds.

The very first version of the 2016 PCB also contained SKIROC2 ASICs, resulting in a complete module with a thickness compatible with the technical proposal design. A variant comprising a pair of PCBs was developed: the sensor PCB, glued to the sensor, connects electrically to the front side of the silicon cells with aluminium wire bonds through holes in the PCB and routes these signals to two connectors. The second, readout, PCB that plugs into the connectors of the first, contains the SKIROC2 and connectors to the outside world. This setup is slightly thicker than the TP design but allowed more flexibility: the readout PCB could be changed if any active element failed, without having to discard the silicon sensor.

The use of the SKIROC2-CMS ASIC required a redesign of the front-end PCB. A new single-layer "hexaboard" was developed, featuring stepped layers for the through-holes. These steps allowed more conventional wire-bonding (no need for "deep access") and provided better insulation between the analog signals and the digital parts. Although each SKIROC2-CMS has 64 input channels, four ASICs were used per hexaboard in order to minimize the path lengths between the wire bonds and the inputs to these ASICs. Low-voltage power, control and data were provided through standard micro-HDMI connectivity, whilst a micro-USB connector allowed programming of the FPGA.

The module assembly process is summarized in Figs. 10.9 to 10.11.

For both 2016 and 2017, flexible DAQ hardware was developed for the silicon modules. In 2016 this was largely based on off-the-shelf components and custom PCBs. For the larger-scale tests of 2017 and beyond, a lower-cost system based on the reuse of FPGA-based mezzanine cards, optical readout modules (ORM), from CMS was developed and produced. This system could read data from up to 112 silicon modules at a rate of 50 Hz, resulting in a bandwidth of about 33 Mbytes/sec.

Flexible mechanics, representative of the CE-E and silicon part of CE-H, were built. For the CE-E part, a "hanging file" system was constructed, as shown in Fig. 10.11, with insertable absorber plates (tungsten in 2016; lead in 2017, following the evolution of the CE-E design) and copper cooling plates. A single detector module could be mounted on either side of these copper plates. In the ideal case, 28 modules would be mounted on 14 copper plates. This number of modules was not available, so the flexible mechanics allowed three different configurations of absorbers and modules in CE-E to be explored in 2016 and four more in 2017.

At FNAL, sixteen silicon modules were available in 2016, arranged as double-sided layers interspersed with tungsten absorbers. The total thickness of the setup at FNAL was deliberately limited to about 15 X_0 due to the relatively low energy electron beams available at FNAL (maximum momentum 32 GeV). At CERN, only eight silicon modules were available in 2016 but two configurations were explored. The first focused on the electromagnetic shower maximum for high-energy electrons, having the modules placed between about 6 and 15 X_0. The second setup had modules covering a more extensive longitudinal region, from 5 to 27 X_0. We refer to these setups as "CERN setup I" and "CERN setup II" respectively.

In 2017, CE-H mechanics were built, based on steel absorber plates and copper cooling plates, totalling about 3.5 λ. A hanging-file system was again used for flexibility, but the transverse

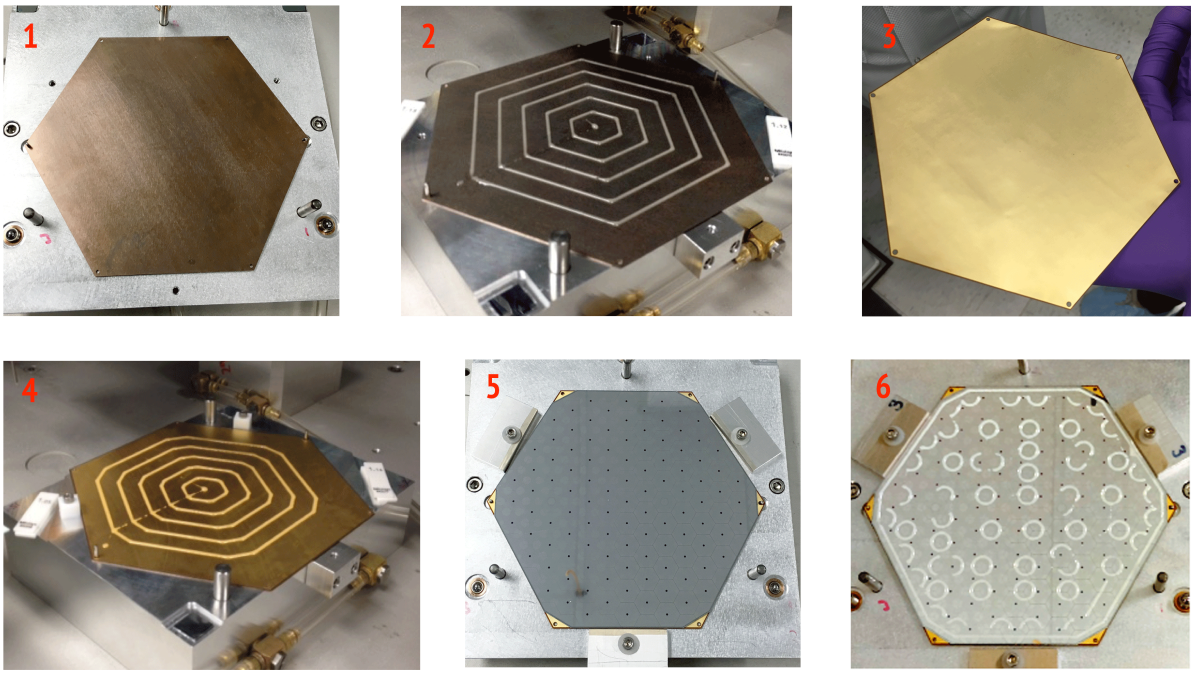


Figure 10.9: The assembly of one of the modules used in the 2016 beam tests, until the PCB-mounting stage. 1: the CuW baseplate on its gluing jig; 2: addition of Araldite 2011 epoxy; 3: the polyimide gold-plated foil; 4: the polyimide foil glued to the baseplate and with its own layer of glue; 5: the silicon sensor glued onto the polyimide foil; 6: the addition of a special pattern of glue that, when the PCB is attached, spreads over the entire silicon surface except where the wire-bonding takes place.

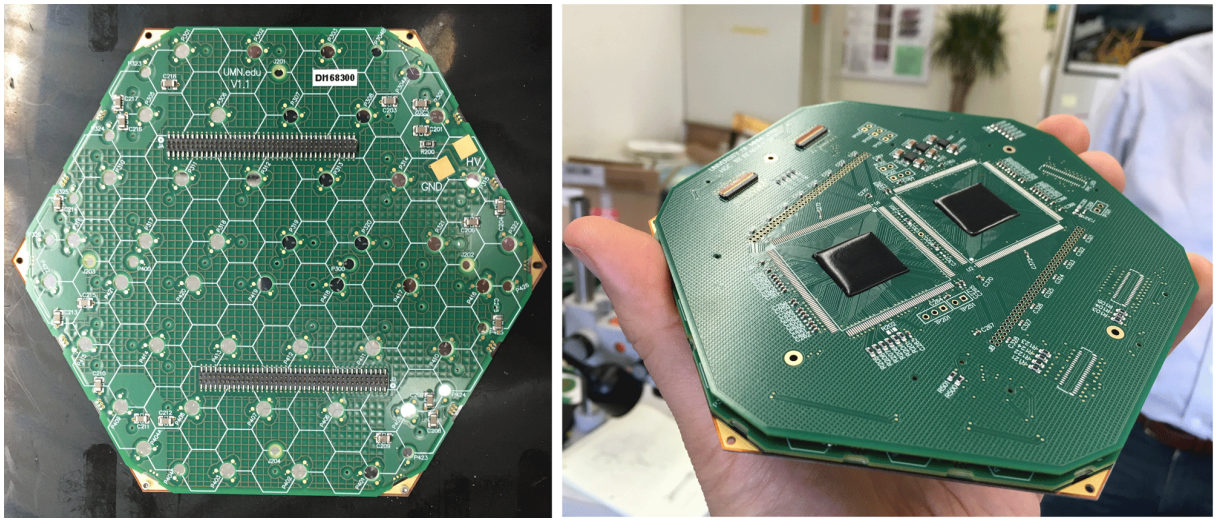


Figure 10.10: Left: a photo of the assembled 2016 bare silicon module showing the PCB with connectors and holes for wire bonds. The polyimide layer is visible at the corners, where there are also holes through the module for attaching it to a support (cooling) plate. Right: a complete HGCAL module, showing the readout PCB with two wire-bonded SKIROC2 bare-die ASICs (under protective black encapsulant material.)

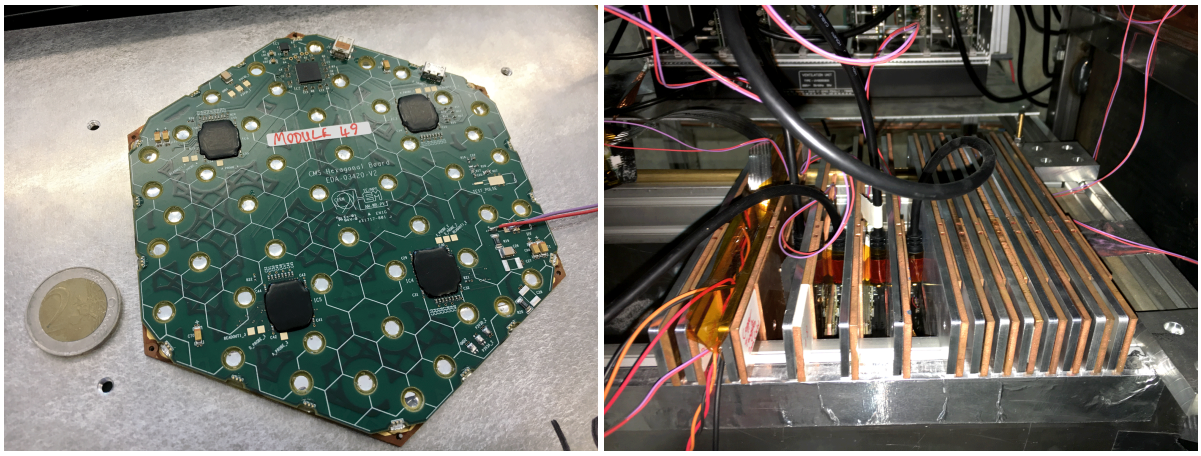


Figure 10.11: Left: Photo of a 2017 silicon module, featuring "hexaboard" PCB containing 4 SKIROC2-CMS ASICs (under black encapsulant), MAX10 FPGA, through-hole wire bonding and micro HDMI and USB connectivity; Right: The hanging-file mechanics used to support the CE-E modules and absorbers. The image is a closeup of the structure when some modules and absorbers had been inserted.

dimensions were larger, to allow more modules per plane to be mounted, to capture the more extensive hadronic showers.

To complement the silicon modules, an existing detector was modified and used in 2017: the analogue HCAL (AHCAL [14]) developed by the CALICE Collaboration. This comprised 12 plastic-scintillator-tile planes, with 12×12 tiles of $2\,\text{cm}\times2\,\text{cm}$ per plane and on-tile SiPM readout. The SPIROC ASIC, a variant of the SKIROC chip, was used for the readout of the SiPM signals. The original version of AHCAL included $2\,\lambda$ of steel absorber material. This was modified to have a total of $5\,\lambda$, using 12 planes of 74 mm-thick steel. Figure 10.12 shows one of the active planes, with an inset focusing on one of the scintillator tiles with SiPM. Also shown is the complete AHCAL, comprising 12 planes of absorber and 12 planes of scintillator.

Up to 20 hexaboard-based silicon modules were available for tests at CERN. Several configurations were explored, varying the numbers of modules in CE-E and CE-H, as well as the number of modules per plane in CE-H. Figure 10.13 shows a close-up of three modules mounted on one CE-H plane, whilst Fig. 10.14 shows the full CE-E + CE-H + AHCAL setup used at CERN in 2017.

The geometries of all setups were reproduced in CMSSW. Data were taken with a wide range of electron energies: 4–32 GeV at FNAL and 20–250 GeV at CERN; as well as 120 GeV protons, muons and charged pions with energies between 20–350 GeV.

The most important results obtained so far from the 2016 and 2017 tests are summarized in Section 5.2. We present here some additional information.

Figure 10.15 shows two typical event displays from the 16-layer FNAL test in FNAL in 2016 (top) and a 17-layer test at CERN in 2017 (bottom). The colour scale represents the signal size in each hexagonal cell.

As an example of hadron shower development in the CERN tests, Fig. 10.16 shows the hit occupancies in the 7-silicon-module CE-H layer and the twelve planes in the AHCAL. There were some data integrity issues in some events in the former, distorting somewhat the overall shape. Detailed analysis is ongoing.

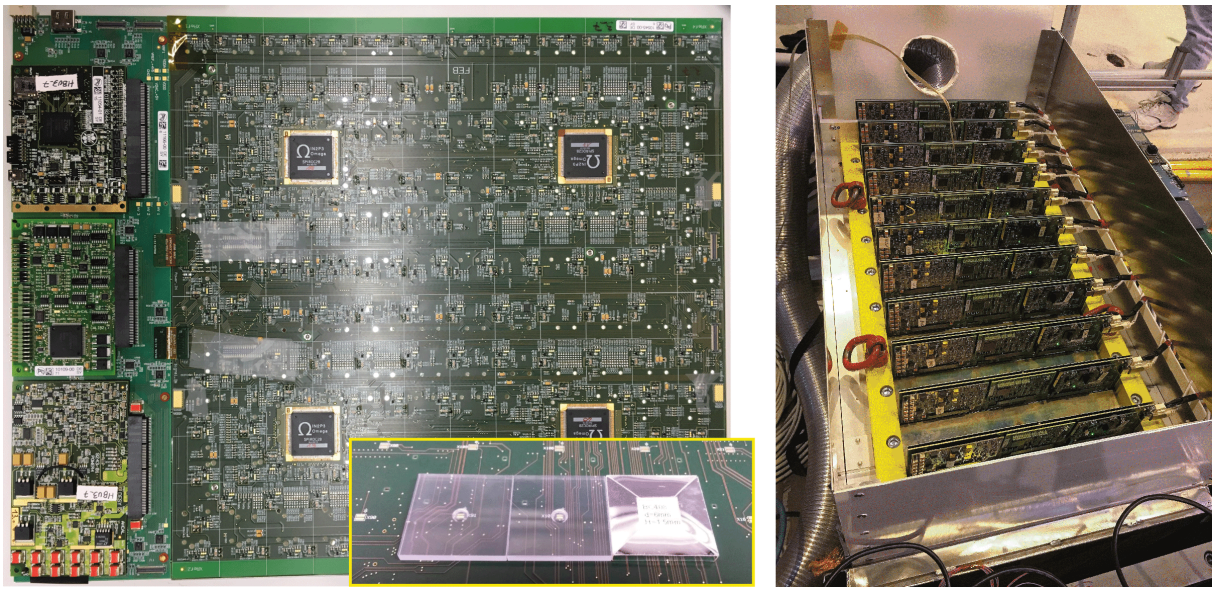


Figure 10.12: Left: Photo of one of the scintillator+SiPM planes used in the AHCAL, including a closeup of one the tiles; Right: The complete AHCAL, comprising 12 layers of 74 mm-thick absorber interspersed with scintillator planes.

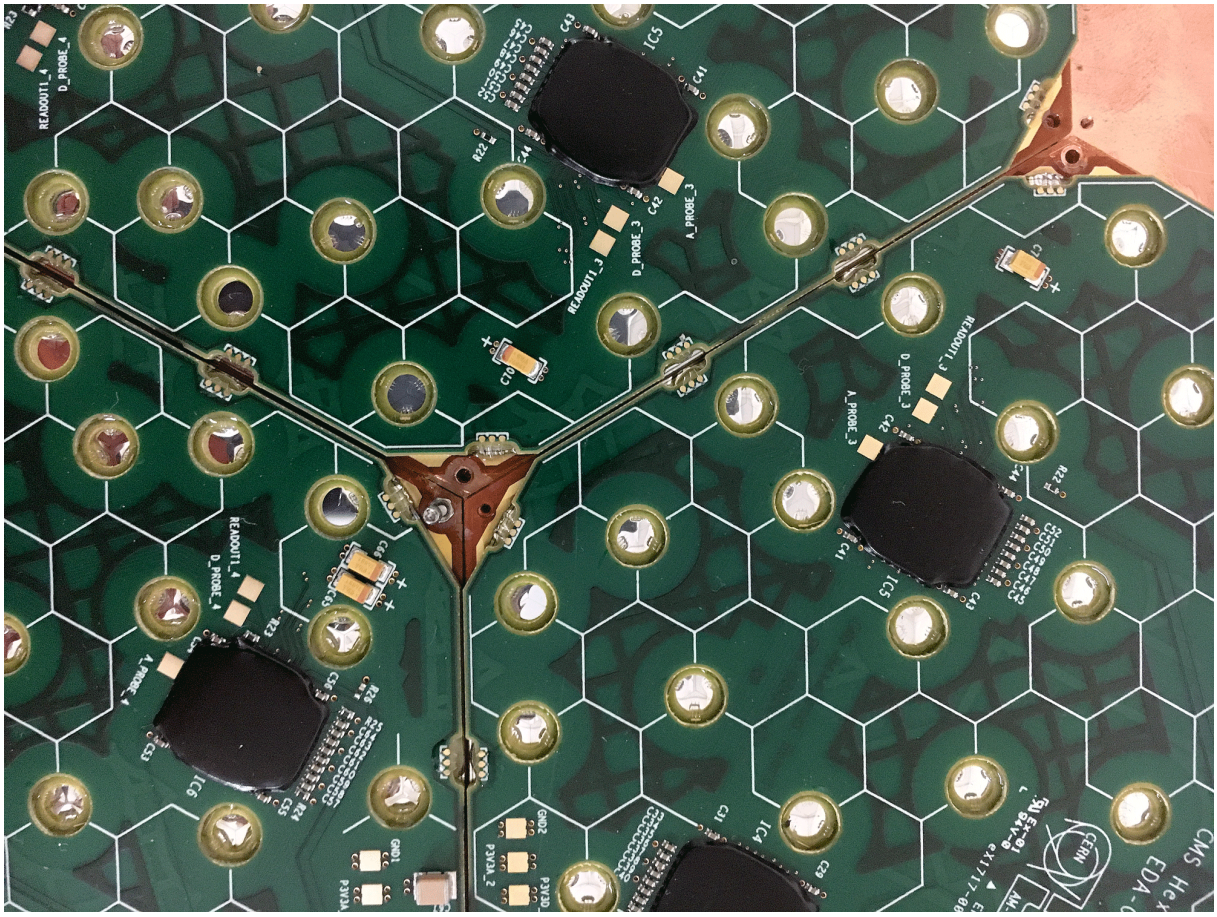


Figure 10.13: Closeup of three 2017 modules mounted on a single CE-H copper plate.

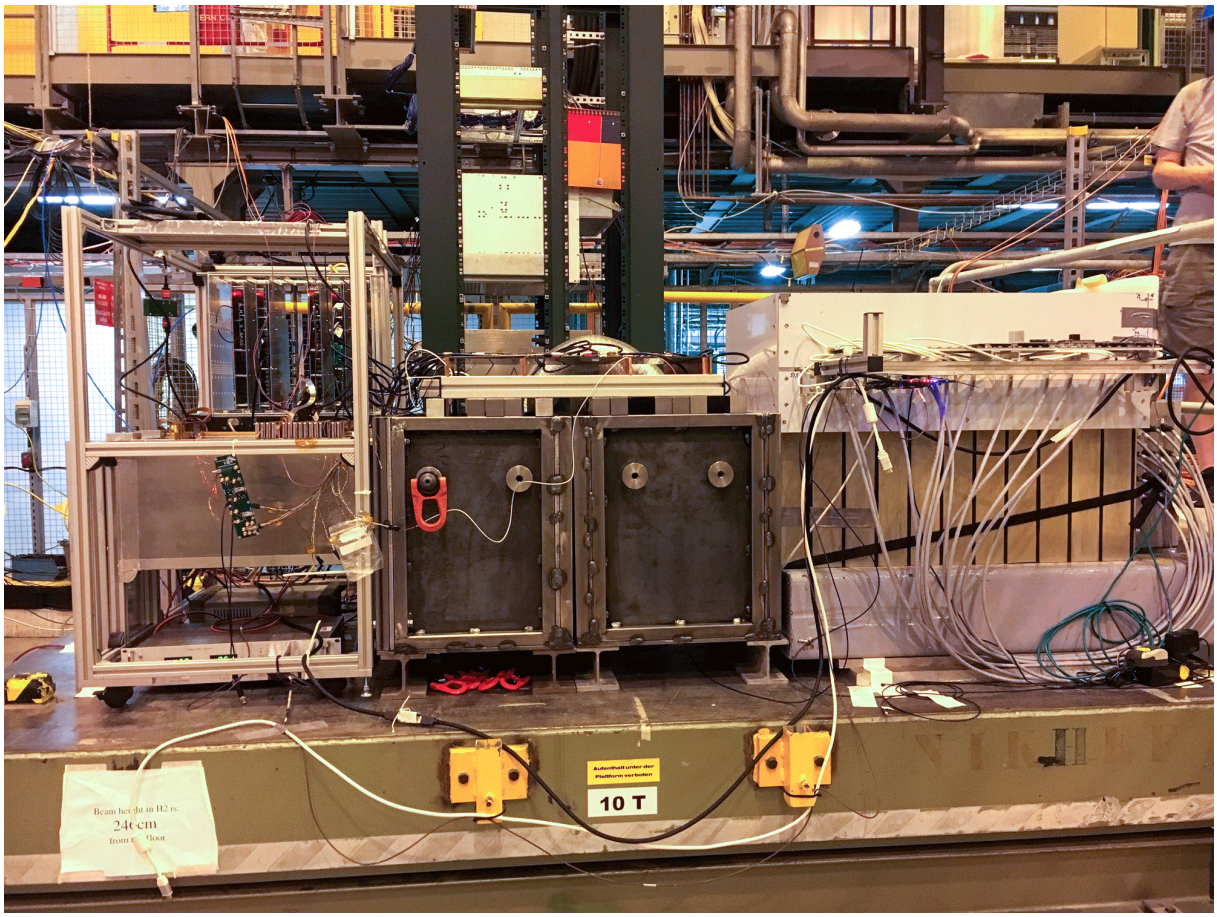


Figure 10.14: CE-E (left), CE-H (centre) and AHCAL (right) on the scissor table in the H2 beam line at CERN.

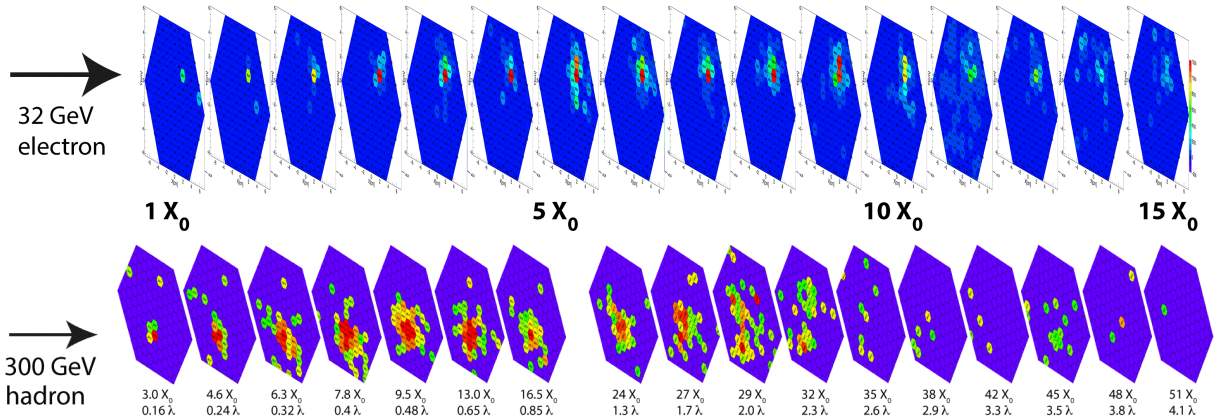


Figure 10.15: Top: Event display showing the shower development in the 16-layer test at FNAL in 2016 due to an incident 32 GeV electron; Bottom: Event display showing the shower development of a 300 GeV hadron in the 17-layer test at CERN in 2017.

10.2.4 Calibration with MIPs - uniformity

The transverse sizes of the beams and scintillators used to trigger the prototypes were too small to cover the full sensor area. Several cells in each silicon sensor were calibrated to study their

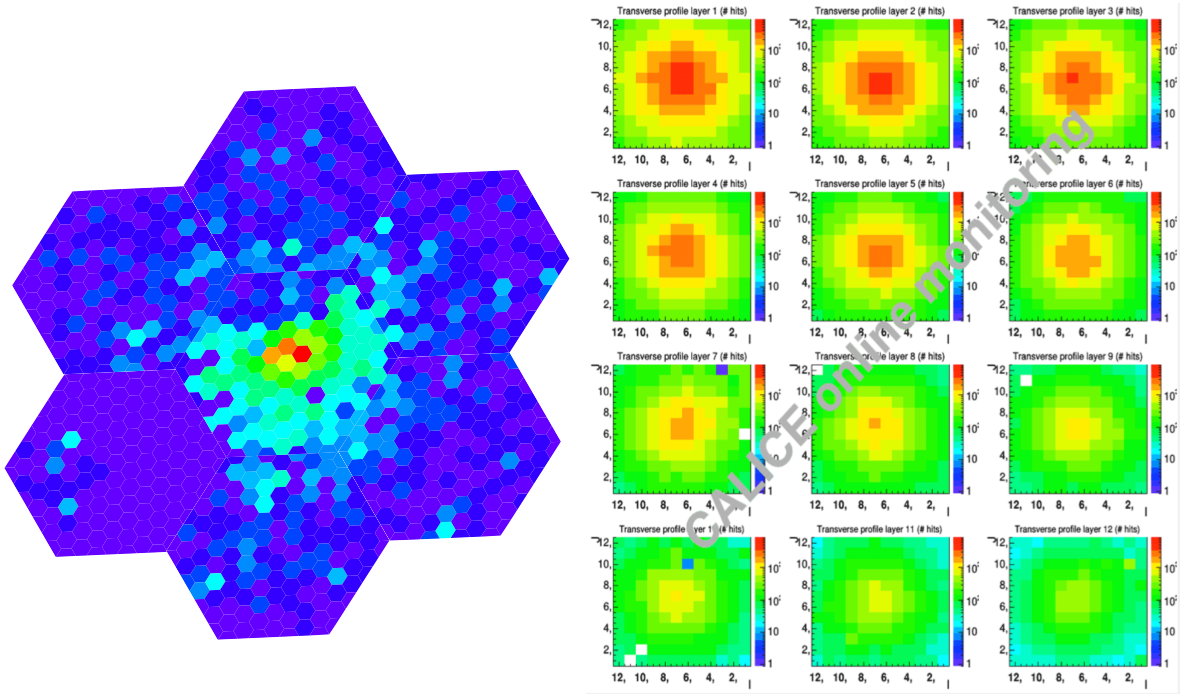


Figure 10.16: Left: Occupancy of hits in the cells in the 7-module CE-H silicon layer due to incident 120 GeV hadrons. The lack of entries in some regions is due to some data integrity problems (under investigation); Right: Occupancy of hits in the cells of the 12 layers of the AHCAL.

uniformity, and average of of each sensor was used to set the calibration of the other cells in the same ASIC. Figure 10.17 shows a typical energy spectrum for incident charged pions, together with the MIP signal measured for all calibrated cells in 2016. It can be seen that while the MIP values for a given SKIROC2 are clustered, there is some variation between SKIROC2s. The signal-to-noise (S/N) for a single MIP is between 7 and 8 for all measured cells, except for the dedicated calibration pads (smaller than the regular cells), which had $S/N \approx 10$.

For the AHCAL in 2017 we used a simple tracking to find single particles inside hadron showers, enabling many scintillator/SiPM cells to be covered. Figure 10.17 also shows the number of ADC counts per MIP measured in the 864 cells of the AHCAL. The spread is expected to be smaller in future since the restrictions on the range of breakdown voltages for the batch of SiPMs used in this prototype was less stringent than they will be in the final specifications.

10.2.5 Longitudinal shower shapes

The longitudinal depth barycentre for two example electron beam energies for the 2016 CERN setup II is shown in Fig. 10.18, and is defined as:

$$t = \frac{\sum_{i=1}^{8} (E_i X_{0,i})}{\sum_{i=1}^{8} E_i}, \tag{10.2}$$

where E_i is the layer raw energy and $X_{0,i}$ the total calorimetric radiation length up to layer i. Good agreement between data and simulation is observed. At lower energies there is a small shift of the shower depth in the data to lower values compared to the simulation. This could be due to electron showering in not-simulated upstream material in the H2 beamline.

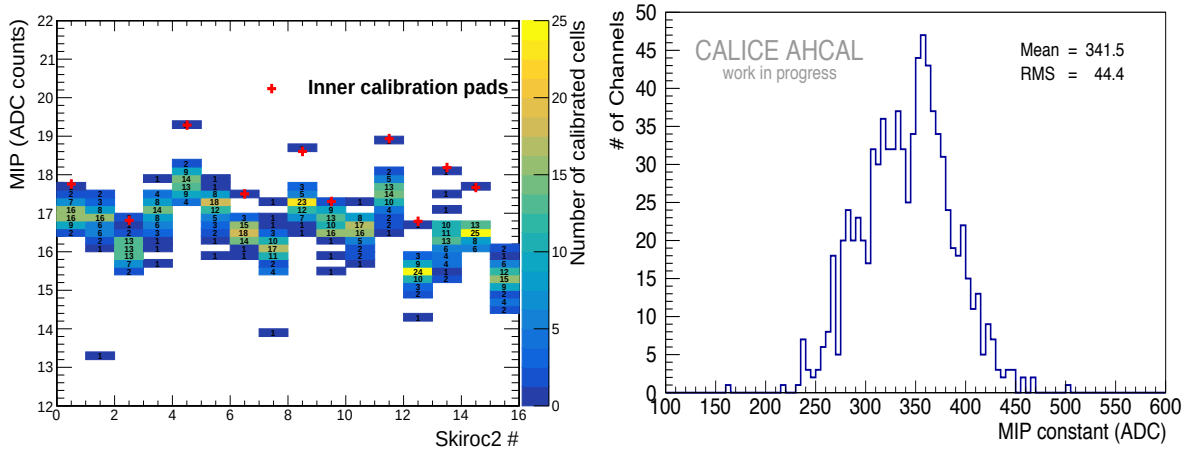


Figure 10.17: Left: MIP signal for all calibrated cells, as a function of the SKIROC2, for the CERN tests in 2016; Right: ADC counts/MIP for all 864 channels of the 2017 AHCAL.

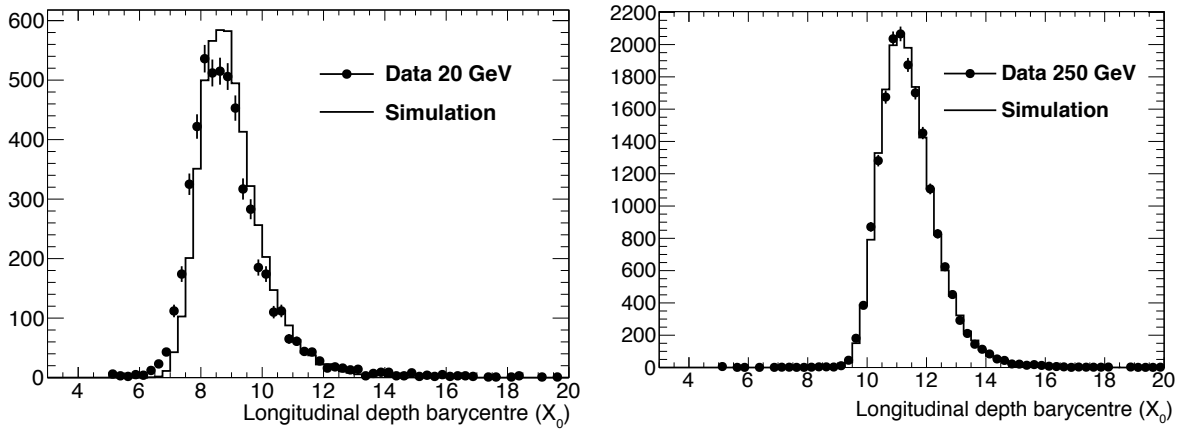


Figure 10.18: Longitudinal depth barycentre distribution comparison between data and simulation for electron energies of 20 and 250 GeV at CERN in 2016.

In order to test this supposition we plotted the fraction of energy deposited in the first layer with respect to the total energy deposited in all layers, as shown in Fig. 10.19. This variable is sensitive to early showering upstream of the calorimeter. In these plots a small departure between data and simulation is observed only for the lower energies, reinforcing the supposition that there is some early showering in the H2 beamline in non-simulated materials.

The evolution of the longitudinal shower-maximum position with energy is shown in Fig. 10.20 for both FNAL and CERN data, as well as the simulation. Agreement at the level of better than 10% was found between data and simulation.

10.3 Trigger simulation and performance

As described in Section 8.3.2.3, the primitives from the TPG consist of a list of 3D clusters and a transverse energy map in η-ϕ. These data are passed to the central L1T correlator and combined with primitives from the other CMS subsystems to form L1 triggers. Fundamentally, the performance of the TPG should be judged on its ability to enable efficient triggers with acceptable

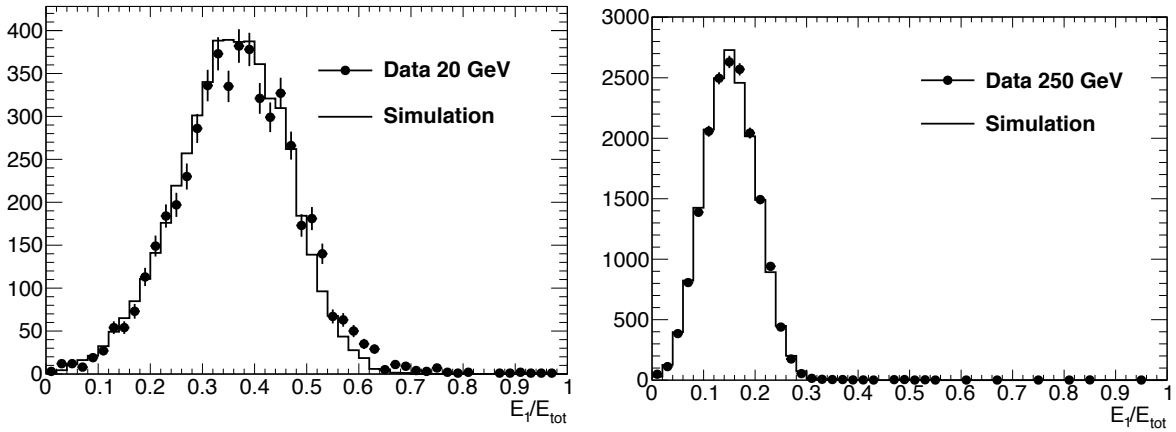


Figure 10.19: Fraction of energy deposited in the first layer distribution comparison between data and simulation for beam energies of 20 and 250 GeV at CERN in 2016.

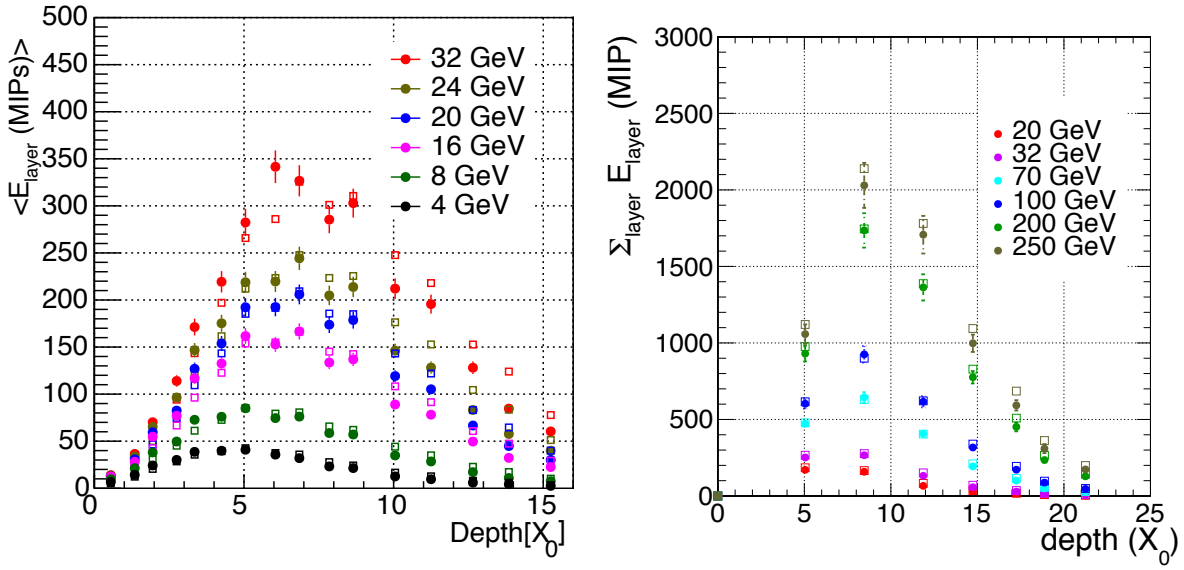


Figure 10.20: Variation of the shower-maximum position as a function of electron energy. Solid symbols represent data whilst the open symbols are for simulation. Left: for the 16-layer configuration at FNAL; Right: CERN 8-layer setup II.

background rates in the central L1T system. These triggers are still under development, so as a proxy, the performance of HGCAL-only triggers, i.e. formed using only the TPG primitives, is presented here.

10.3.1 TPG simulation

Two different simulation packages, described below, were used for the simulation of the electronics chain for DAQ and TPG studies. A detailed CMSSW-based simulation is used for the performance results, while a more flexible simulation, based only partly on CMSSW, is used when variations of parameters and options are needed. The latter uses CMSSW to create the energy deposits in the sensitive layers, so both simulations depend on the standard CMSSW simulation program, which does not yet have the TDR geometry implemented; see Section 10.1.

Briefly, the CMSSW simulation currently has 40 silicon-only layers, implemented using 6 inch sensors with a non-TDR cell geometry, and a further 12 scintillator-only layers with an array of 85×360 cells in $\eta \times \phi$.

10.3.1.1 CMSSW

The processing path followed by trigger data, from the very front-end data reduction to the clustering algorithms running in the back-end electronics, is simulated in CMSSW. This simulation was used for the trigger performance results shown later in this section. It begins with the summation of cells into groups of size of about $\Delta x \times \Delta y = 2 \times 2\,\mathrm{cm}^2$ for the silicon part of the detector, and groups of about $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$ in the scintillator part. The TDR geometry based on 8-inch wafers is used to define the positions and sizes of the trigger cells, and the simulated cells based on a different geometry are mapped to these trigger cells. This mapping is done by associating a given simulated cell to the trigger cell with the maximum overlap. The measured charges in the cells forming a trigger cell are rescaled to a LSB of 0.1 fC with a range of 16 bits, before being summed. The summed values are encoded on 20 bits with the same LSB in order to be able to sum 9 cells without any loss. The measured trigger cell energies are not compressed in this simulation but it has been found with the standalone simulation that the 8-bit floating point compression results in negligible effects on the final reconstructed energies, compared with the threshold of $2\,\mathrm{MIP_T}$ applied on trigger cell energies. Trigger cell energies are calibrated to the electromagnetic (for the CE-E) and hadronic (for the CE-H) scale with coefficients that depend on the layer (see Section 10.3.2).

Trigger cells with a transverse energy above $2\,\mathrm{MIP_T}$ are selected and used as input to the two-step clustering algorithm. As described in Section 8.3, in the first clustering stage, seed trigger cells are selected with a threshold of $5\,\mathrm{MIP_T}$ and initiate the 2D clusters. Any trigger cell in the same layer that passed the $2\,\mathrm{MIP_T}$ initial threshold and is topologically connected to a seed is included in the corresponding cluster. Finally, connected 2D clusters are merged into single clusters. For the second stage, the 2D clusters are combined to form 3D clusters. This is done geometrically. For 2D clusters arising from a particle coming from the origin of the coordinate system, the clusters positions in each layer in x and y will all have the same value in scaled coordinates $x' = x/z$ and $y' = y/z$. Starting from the first layer, each 2D cluster is considered to be a potential 3D cluster. The scaled position of each 2D cluster is compared to scaled position of all the 2D clusters in the next layer. Any 2D cluster in the next layer within a scaled radius of $\Delta\rho = \sqrt{\Delta x'^2 + \Delta y'^2} < 0.01$ is then associated with the 2D cluster to start forming a 3D cluster. These 3D clusters are then projected through to each subsequent layer, comparing to 2D clusters in each layer in the same way.

10.3.1.2 Standalone

The second simulation package operates mostly outside of CMSSW. The truth energy deposits ("simhits") from CMSSW are written out to text files and then used as input to a completely independent simulation of the HGCAL FE and BE electronics chain. This provides a flexible way to study variations in parameters, and indeed architectures, of the electronics to optimise the design, as well as study changes to the geometry. It also allows a cross check of the CMSSW simulation. This simulation was used to produce the DAQ and TPG occupancy and rate results shown in Sections 3.1 and 8.1.

For each BX, the simulation deposits in the sensitive layers ("truth hits") were made by combining (optional) signal events with pileup. The average number of pileup per BX was generated randomly according to a Gaussian with a specified mean (typically 200) and width of 20% of the mean. The actual number of pileup events in each BX was then generated according to a

Poisson distribution based on this average. In addition, the pileup hits from the previous five BX were also used to generate out-of-time (OOT) signals using a CR-RC^3 shaper with a peaking time of 17.6 ns, which results in 20% of the peak signal remaining after one BX. Noise is added according to a parameterization of the expected noise after 3000 fb^{-1}.

The on-detector electronics simulation modelled a 10-bit ADC with an LSB of 0.2 fC and a saturation at 800 ADC counts, and a 12-bit TDC for the ToT with an LSB for large deposits of 2.5 fC, which allows values up to 10 pC. The non-linearity of the ToT is modelled with a turn-on point of 100 fC. A 10-bit field is also included for the ToA measurement when estimating data rates. The thresholds used for the readout of these values are 0.5 MIP, 10 fC and 150 fC, for the ADC, ToA and ToT, respectively. ADC values from the previous BX are also read out if above 2.5 MIPs. The trigger cell values are formed from summing up to 9 or 4 sensor cell readout values for the sensors with $\approx$0.5 cm^2 or $\approx$1 cm^2 cells, respectively, where the sensor cell reading is included in the sum if above three times its noise value. The sensor cell values use the ADC value up to its saturation point, a constant value between this and 150 fC and then the ToT value scaled to match the ADC LSB. A trigger cell therefore has a LSB of 0.2 fC and a maximum value of 90 pC, so the native value would be represented by 19 bits. This is compressed to an 8-bit floating point value with a 4-bit exponent and a 4-bit mantissa, which covers the required range of 19 bits. Trigger cell values are read out if above 2.0 MIP$_\mathrm{T}$. All trigger cell values for each HGCROC are summed to give the coarse sum used for the energy map. It is represented as an 11-bit floating point, with a 4-bit exponent and a 7-bit mantissa and this is read out whenever there is at least one trigger cell above threshold from the HGCROC.

The mapping to the TDR geometry from simulation generated with the older CMSSW geometry is performed at the truth hit level. Since the truth hits used have already been assigned to the old geometry sensor cells, their exact positions within the cell are not known. A numerical evaluation of the overlap areas of every sensor cell in the old and new geometries was calculated and the energy of each hit in an old geometry cell is assigned randomly to one of the TDR geometry cells, with a probability proportional to the overlap area. This preserves the MIPs and allows all combinations of silicon and scintillator cells to be mapped to each other, as required since the old geometry has only all-silicon and all-scintillator layers.

10.3.2 Calibration and pileup corrections

To get an accurate energy estimate to compare with trigger thresholds, the objects need to have a calibration applied, and a correction made for any pileup contamination. In the absence of pileup, the required calibration would be expected to result from the integrated energy loss of the absorber layers as used in the studies in Section 10.1. However, a specific calibration is needed for the trigger objects as there is also energy lost from the objects at various points in the TPG data path. Specifically, the FE electronics selection of trigger cells above 2 MIP$_\mathrm{T}$ results in energy being lost while the 2D clustering algorithms reject potential clusters without a seed trigger cell, and the 3D clustering algorithms reject any clusters with a transverse energy of less than 1 GeV. Since electromagnetic showers tend to have higher energy deposit densities, the effects of these energy losses are more important for jet reconstruction than for e/γ reconstruction.

Hence, an algorithm-specific calibration is required. This is applied to the 3D clusters when formed from the 2D clusters by multiplying the deposited energy in the 2D cluster by a layer-dependent weight; all 2D clusters in the same layer have the same weight. Specifically, each 3D cluster c has some value $d_{c,i}$ of its total deposited energy in layer i from all 2D clusters in that

layer, and the cluster energy estimation E_c is given by

$$E_c = \sum_i w_i d_{c,i} \tag{10.3}$$

where w_i is the layer weight. This type of linear calibration is easily implemented in an FPGA.

The weights w_i were determined directly from simulation using a standard linear algebra technique. The RMS^2 of the estimated cluster energy compared with the truth value T_c is given by

$$\text{RMS}^2 = \frac{1}{N}\sum_c (E_c - T_c)^2 = \frac{1}{N}\sum_c \left(\sum_i w_i d_{c,i} - T_c\right)^2 \tag{10.4}$$

Writing in matrix notation with M and v defined as

$$M_{ji} = \frac{\sum_c d_{c,j} d_{c,i}}{N}, \qquad v_j = \frac{\sum_c d_{c,j} T_c}{N} \tag{10.5}$$

then minimising the RMS^2 gives $Mw = v$ so $w = M^{-1}v$, which therefore determines the weights w_i. Because of the matrix inversion, to get stable values for the weights requires large statistical samples. Examples of the weights estimated using samples of electrons and charged pions are shown in Fig. 10.21.

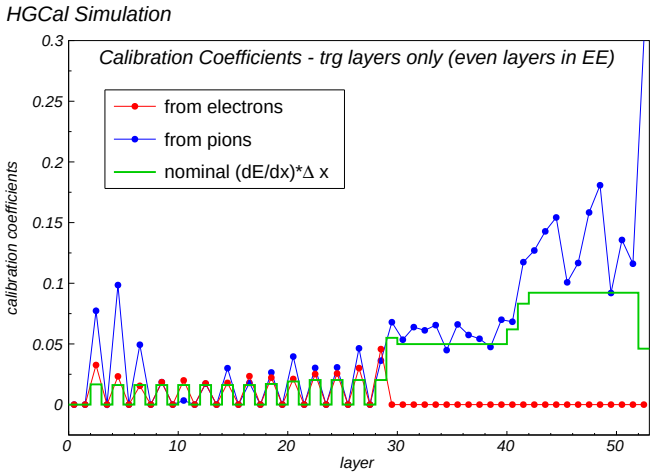


Figure 10.21: Layer weights determined from electron (red) and charged pion (blue) simulation samples with no pileup. The electron weights are forced to be zero in the CE-H section. Also shown are the weights (green) derived from the expected integrated energy loss in the absorber material between each layer. The electron weights are close to the expected values from a calculation of the energy loss in the absorber, while the pion weights, particularly in the CE-H section, are systematically higher to account for the energy lost in the TPG processing chain.

For the results shown later, the calibration weights calculated with zero pileup simulation samples were used. These require any reconstructed physics object to have an explicit pileup energy correction. However, the weights could in principle be adjusted so they perform a first-order pileup correction on each individual 3D cluster, rather than applying some average correction. A schematic idea of this concept is shown in Fig. 10.22. The left diagram shows the weights which would be used with no pileup, which are effectively dependent on the absorber thicknesses and hence increase with depth, i.e. layer number; these correspond to those found

in Fig. 10.21. The central diagram shows the case for electromagnetic clusters with pileup. The number of significant energy deposits from an electromagnetic object in the CE-H section is small enough that at an average pileup of 200, any 2D clusters included in the 3D cluster from the CE-H layers are statistically much more likely to be deposits from hadrons from the pileup events. Putting the weights for these layers to zero would remove their effect on the 3D cluster energy and this was done for the electron weights shown in Fig. 10.21. However, there is also hadronic pileup in the CE-E layers and the pileup observed in the CE-H allows an estimate of this earlier pileup. Hence, using weights which are less than zero for energy seen in the CE-H layers gives a correction for the pileup in the CE-E layers. Furthermore, there will be some electromagnetic energy from pileup but this is generally low energy and hence concentrated in the front layers of the CE-E. Hence, in a similar way, these layers give some measure of pileup energy in later CE-E layers also. Therefore, putting the weights for the early layers negative can also correct for pileup. Finally, as shown on the right diagram, for hadrons (and hence jets), having negative weights for the first few layers to correct for pileup is still useful, but in contrast to the electromagnetic case, 2D clusters in the CE-H are more likely to be from the physics object (i.e. not pileup) and so retain positive weights. Note, this means the calibration depends on the assumed object type. Due to the larger fluctuations from pileup in addition to those from the physics objects, even larger simulation samples are needed to obtain stable estimates of the weights and so no quantitative results are shown here, although this technique will be investigated in future.

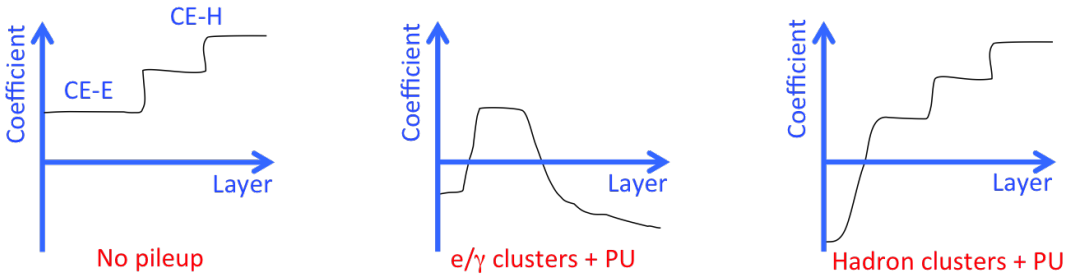


Figure 10.22: Schematic diagram of the concept of using layer weighting to perform pileup corrections; see text for details.

All of the trigger performance results below use only the 3D clusters and not the energy map primitives. The energy map should contain a better estimate of all the energy deposited, since its formation has no trigger cell or clustering energy loss. However, this comes at the price of worse pileup contamination as the η-ϕ bins integrate over larger regions than individual trigger cells. The optimal combination of the two different types of primitives will require a significant study which has not yet been done.

10.3.3 HGCAL-only trigger performance

The triggers studied are aimed at specific physics objects, namely electromagnetic showers (e/γ) and jets. For e/γ objects, the energy used is assumed to be contained in a single 3D cluster. In contrast, jets are formed from multiple 3D clusters over a wider angular range. Hence, these triggers depend on very different aspects of the TPG reconstruction. For both sets of objects, the calibration weights calculated without pileup as described in the previous section are used. For the results below, the background trigger rates are estimated using the fraction of times a BX containing only pileup events passes the trigger criteria. This fraction is converted to a trigger rate assuming a total of 2808 bunch crossings with collisions per turn in the HL-LHC, which corresponds to an average collision bunch crossing rate of 31.2 MHz.

The e/γ objects use a single 3D cluster, so the properties of this cluster can be used to reject clusters which are formed from pileup interactions; these tend to be from low energy jets. Variables which have been found to give discriminating power are the transverse width of the 3D cluster in the radial (i.e. η) direction, σ_{rr}, the maximum number of consecutive layers included in the 3D cluster, the position of the start of the cluster, and the position of the layer with the maximum energy. The transverse width of the 3D clusters is determined from the width of the 2D clusters from which it is formed. These 2D cluster widths are computed from the core of the clusters ($\Delta R < 5$ cm around the trigger cell of maximum energy) to improve the robustness of the discrimination to pileup contamination. Electrons generated with a flat p_{T} distribution between 10 and 100 GeV and a flat $|\eta|$ distribution between 1.7 and 2.9, and simulated with 200 additional pileup interactions, are used as signal. Reconstructed clusters with $E_{\mathrm{T}} > 20$ GeV in events containing only 200 overlapping minimum bias interactions are used as background. The signal and background distributions of these variables are shown in Fig. 10.23.

Figure 10.24 shows the background selection efficiency as a function of the signal selection efficiency based on a selection using the four cluster variables described above. A working point giving an overall efficiency for electrons with $|\eta| < 2.7$ of about 98% is chosen. The selection is tightened for electrons with $|\eta| > 2.7$ due to the much higher background in this region, resulting in an efficiency of 90%. The transverse energy resolution for events with an average of 200 pileup as a function of the generated p_{T} is also shown in Fig. 10.24. The resolution is defined as the effective sigma of the $E_{\mathrm{T}}^{\mathrm{L1}}/p_{\mathrm{T}}^{\mathrm{gen}}$ distribution, where $E_{\mathrm{T}}^{\mathrm{L1}}$ is the transverse energy of the cluster and $p_{\mathrm{T}}^{\mathrm{gen}}$ is the transverse momentum of the generated electron. This is normalised by the average of this quantity to be insensitive to residual calibration effects. The selection efficiency obtained with this working point and a trigger threshold of 30 GeV, as well as the resulting background rate, have been shown previously in Fig. 5.23.

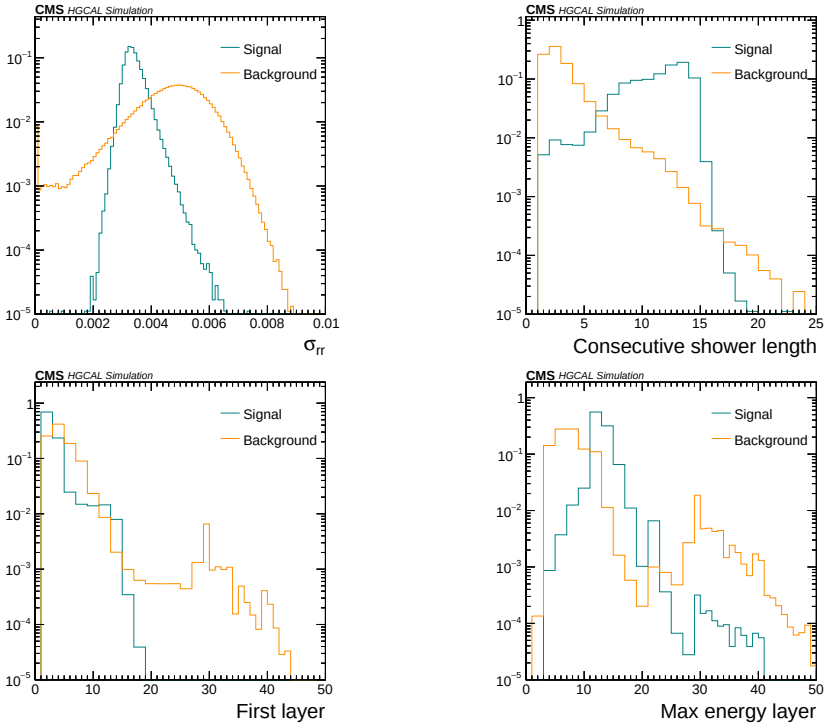


Figure 10.23: Distributions of the e/γ identification variables (see text). Signal (electrons) and background (pileup-only clusters with $E_{\mathrm{T}} > 20$ GeV) are superimposed.

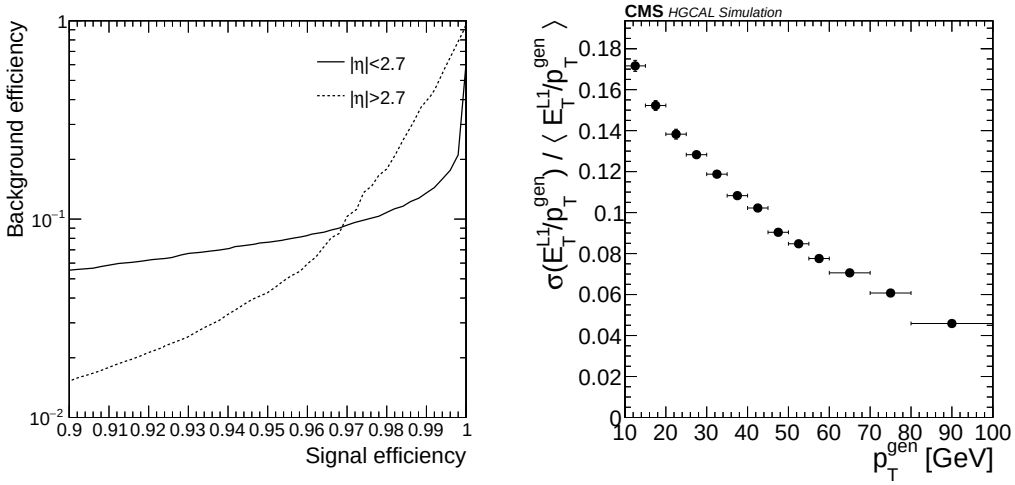


Figure 10.24: Left: Background (pileup-only clusters with $E_{\mathrm{T}} > 20\,\mathrm{GeV}$) selection efficiency for objects at $|\eta| < 2.7$ and $|\eta| > 2.7$ as a function of the signal (electrons) selection efficiency obtained with the cluster shape variables used for the identification of electromagnetic objects. Right: Fractional resolution of transverse energies for electrons as a function of the electron transverse momentum.

For jet objects, a study has been performed using a signal simulation sample of VBF production of Higgs bosons, with the Higgs decaying invisibly. The anti-k_{T} algorithm was used to combine the 3D clusters into jets. The distance parameter of the algorithm was set such as to result in jets with an η-ϕ width of $\Delta R \approx 0.2$. This size gives a good balance between loss of energy outside this width and inclusion of pileup energy. It is unlikely the full anti-k_{T} algorithm could be run within an FPGA, but it is assumed that some approximation will be able to be implemented.

A correction for the pileup within the jet is needed. This is estimated for an average pileup of 200 interactions by forming cones of $\Delta R = 0.2$ centred on several values of $|\eta|$ in events with this pileup level. For each cone, the transverse energy from pileup which would add to the jet transverse energy is estimated from the sum of the transverse energy in all the 3D clusters within the cone. This average pileup energy is found as a function of $|\eta|$, as shown in Fig. 10.25. Each jet transverse energy is then corrected, based on the jet $|\eta|$, by this average pileup energy, scaled appropriately to the pileup level in the event.

The resulting jet energies need a further correction for energy lost during the reconstruction (both in the readout and reconstruction chain, as well as out-of-cone effects due to the relatively small ΔR value used). This correction has been parameterized as a function of the reconstructed jet transverse momentum. This results in a reasonably unbiased jet energy reconstruction as a function of transverse momentum and rapidity, as shown in Fig. 10.26.

All jets reconstructed had the pileup and energy corrections described above applied. Using these, the turn-on curve for a single jet trigger with a central threshold of 150 GeV, as well as the resulting background rate, have been shown previously in Fig. 5.24. Double jet triggers have also been studied, with the jets required to be in the HGCAL. The rates of such double jet triggers for any two HGCAL jets, and where one jet is required to be in each of the two HGCAL endcaps (and hence is very relevant for VBF events), are shown in Fig. 10.27. This plot also shows the rate reduction that can be achieved if a selection of at least 500 GeV on the invariant mass of the jets is applied to the events with jets in both endcaps.

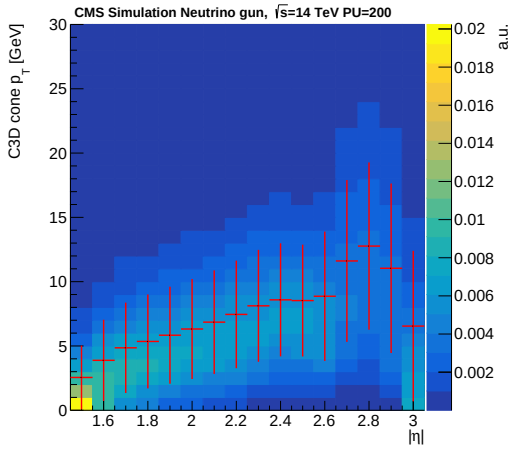


Figure 10.25: Average amount of pileup energy in cones of $\Delta R = 0.2$ as a function of $|\eta|$, for events with an average of 200 pileup interactions only. The average per bin is shown by the red histogram.

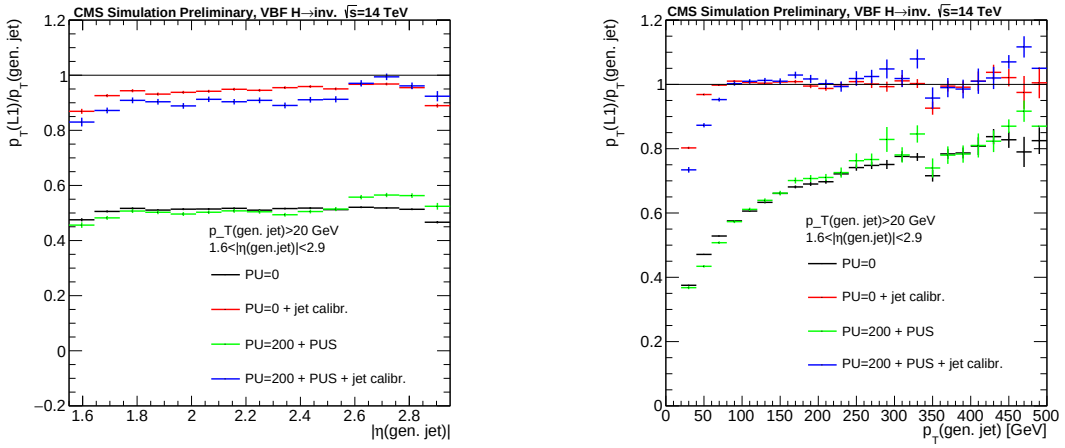


Figure 10.26: Average ratio of the reconstructed jet transverse momentum to the generated truth jet transverse momentum as a function of rapidity (left) and transverse momentum (right) of the generated jet. The black and green points show the average ratios for 0 PU, and for 200 PU with the pileup correction, while the red and blue points show the same following the final parameterized energy correction is applied.

10.4 Precision timing

As discussed in Section 5.5, the front-end ASIC (HGCROC) of the silicon sensors provides a time-of-arrival (ToA) information, with a precision that depends on the intrinsic capability of the sensor. Figure 10.28 shows the achievable resolution on the ToA measurement for the single silicon sensors, as expected from Eq. 5.2, with $\sigma_{\text{jitter}} = A/(S/N)$, $A = 5\,\text{ns}$, and $\sigma_{\text{floor}} = 20\,\text{ps}$.

In the absence of pileup, all the energy deposited in a cell originates from the shower resulting from a single interaction and fires the ToA if the energy deposit is above the threshold. In the presence of pileup, energy deposits in a given cell may originate from showers of particles coming from several interactions. The ToA then records the time at which the deposit of energy crossed the threshold. The ToA threshold of 12 fC corresponds to about 3 MIP in the 300 μm silicon cells, and to about 5 and 10 MIPs in the 200 and 100 μm cells, respectively. A ToA bin

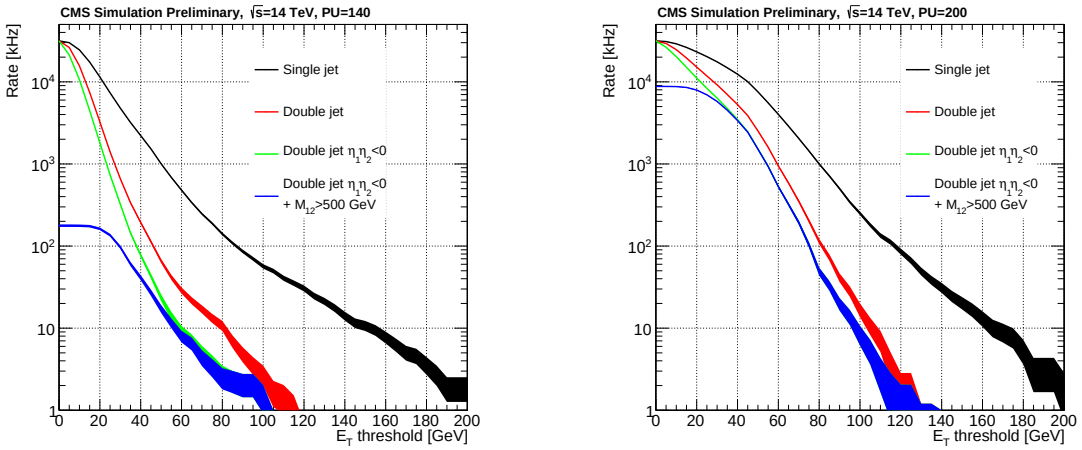


Figure 10.27: Estimated trigger rates for 140 PU (left) and 200 PU (right) for HGCAL-only jet triggers formed from single jets (black), double jets (red), double jets with each jet in a different endcap (green), and for the last category with an invariant mass requirement of at least 500 GeV applied to the two jets (blue).

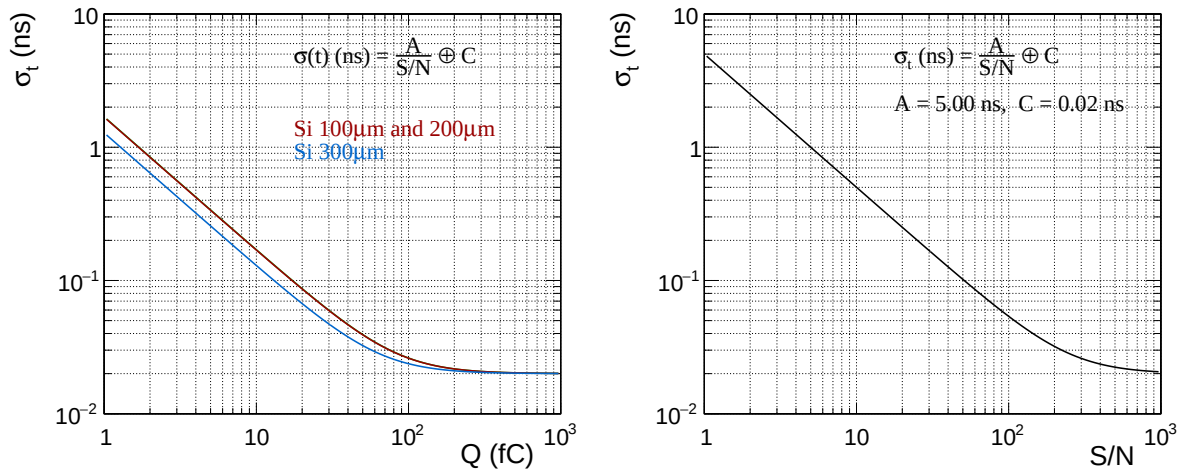


Figure 10.28: Expected timing resolution for cells of 100, 200, and 300 μm as a function of the signal charge magnitude in fC (left), and of the signal-to-noise ratio (right).

size of 25 ps is used to digitize the ToA information.

In the studies presented here unconverted photons and non-interacting K^0_L were used to explore the intrinsic timing capability of the HGCAL. A minimum of three cells within a selection radius, ρ, of the shower axis is required, each with an energy deposit of above the threshold of the ToA. For these studies the shower axis is taken as the MC-truth trajectory of the photon or K^0_L.

Figure 10.29 shows the average number of cells above the 12 fC threshold within a radius of $\rho = 2$ cm and $\rho = 5$ cm from the shower axis, for photons and K^0_L.

Figure 10.30 shows the typical time distribution obtained for all the hits associated in K^0_L showers of $p_T = 2$ GeV and $p_T = 10$ GeV. Also shown are distributions for hits within a radius ρ of the shower axis. The tails with late times are due to both the lateral shower development and the intrinsic fluctuation in the longitudinal development, resulting in a spread of energy deposited over times up to few nanoseconds.

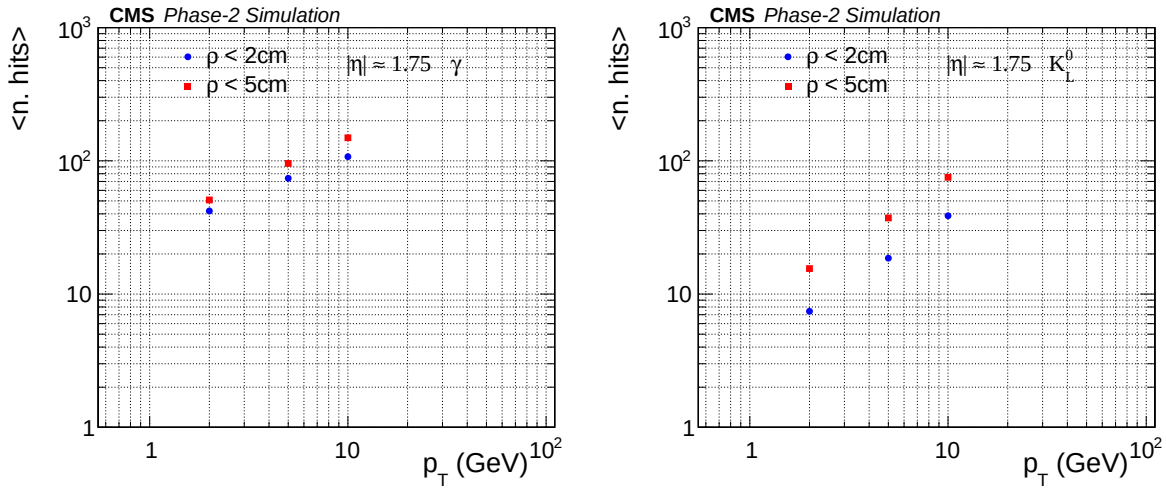


Figure 10.29: Average number of cells above a 12 fC threshold in (left) photon and (right) K_L^0 showers. The values are shown as a function of the particle p_T, for cells within a 2 and 5 cm radius from the shower axis, in blue and red respectively.

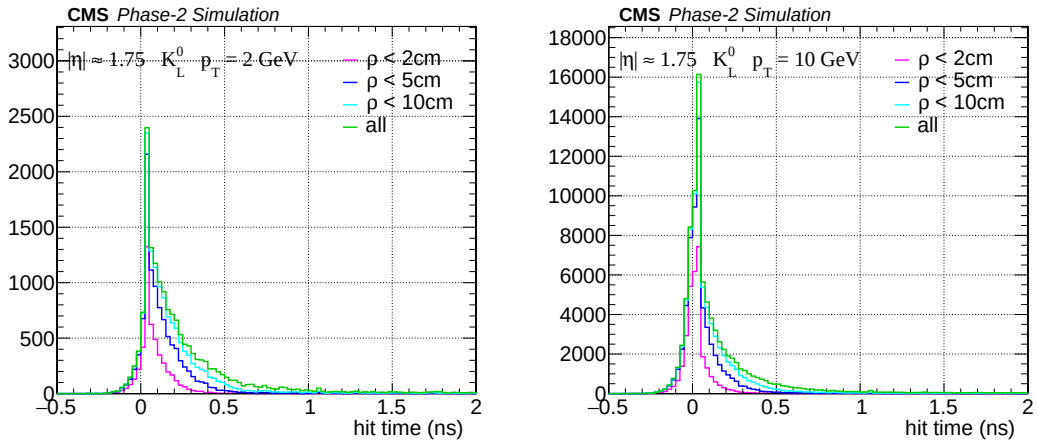


Figure 10.30: Time distribution of the hits for $p_T = 2\,\text{GeV}$ (left), and $p_T = 10\,\text{GeV}$ (right) K_L^0 showers. The distributions are shown for all hits, and for different ρ values.

To measure the time of the shower, simple algorithms have been developed, discarding the hits in the tails of the time distributions. This restriction improves the resolution while preserving the most important time information about the shower. The algorithms select the hits within a time interval containing the highest density of hits.

The first algorithm identifies the smallest time interval containing 68% of the hits. Then the interval width is doubled, by extending the width by a further half-width on each side. This approach allows the use of as many hits as possible, provided they are close in time with the core, and only discards the hits furthest away in time. The choice to use 68% could be further optimized. The smallest time interval containing 68% of the hits has a width of ≈210 ps for photons in the absence of pileup.

An alternative algorithm rejects a fixed fraction of the hits in the tails. The fractions are optimized as a function of p_T, η, and ρ, and for both K_L^0 and photons, by minimizing the time resolution, and at the same time verifying that no significant time bias is introduced by the selection. The results suggest that a single truncation can be used for the range of p_T, η, and shower type considered: namely 10% on the lower side and 40% (50%) on the upper side, for a

selection of $\rho = 2\,\text{cm}$ (5 cm).

The time of the shower is computed as the average of the selected hits. The two algorithms for selecting the hits give similar performance in terms of time resolution in absence of pileup, and both were tested extensively to verify the robustness of the results obtained in the absence of pileup.

Figure 10.31 shows the time distribution of the hits within a radius $\rho = 2\,\text{cm}$, in two η intervals ($1.65 < |\eta| < 1.85$, and $2.65 < |\eta| < 2.85$) for no pileup, and pileup corresponding to a mean number of interactions per bunch crossing of 140 and 200.

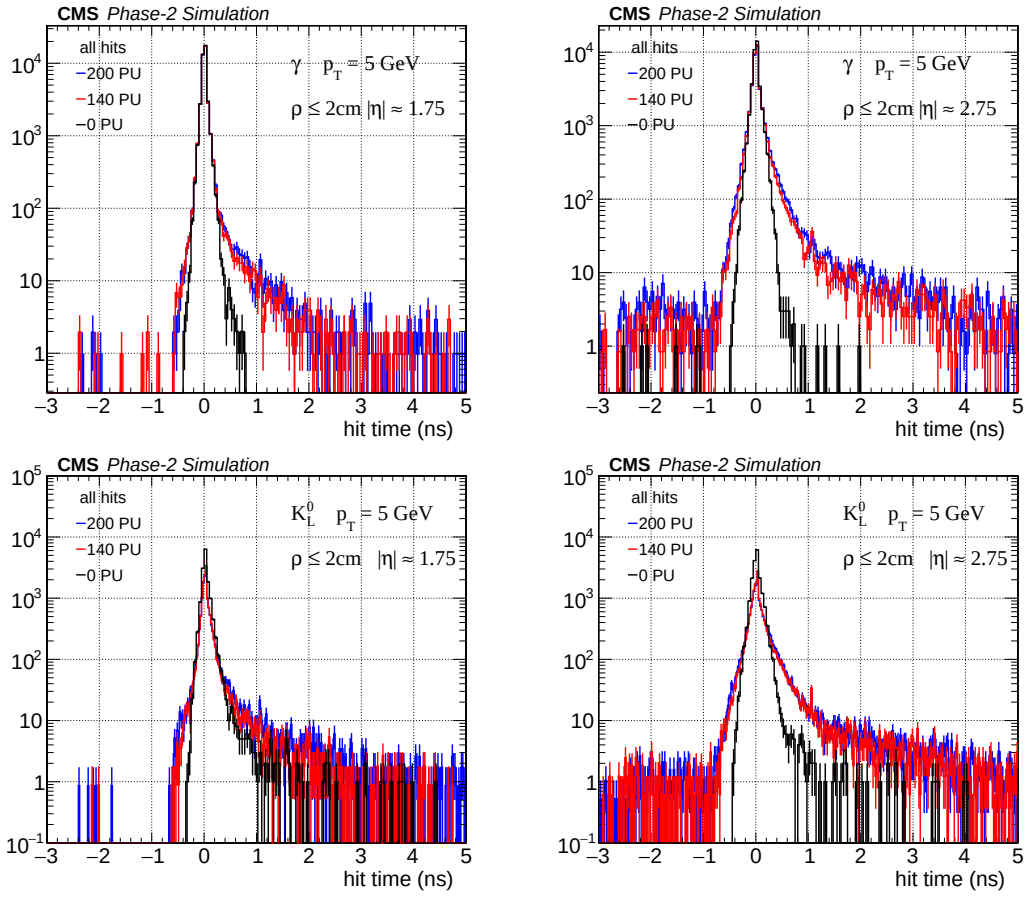


Figure 10.31: Time distribution of hits collected within a 2 cm radius about the shower axis, at low η (left) and high η (right), for photon showers (top), and K_L^0 showers (bottom).

The distributions for K_L^0 have long tails at late times, even in the absence of pileup, while this is not the case for photons. In presence of pileup, the tails are more pronounced, with early times (the predominant contribution of pileup) being noticeably populated at high η, the region of highest pileup density. The distribution of the hits selected in the time interval of highest density, using the algorithm described above, is shown in Fig. 10.32. The selected hits of photon showers, on average, have the same shape in the absence and in the presence of pileup, while for K_L^0 showers there is a residual contamination from pileup.

The fraction of the total shower energy deposited within $\rho < 2\,\text{cm}$ of the shower axis is larger for photon showers, with a consequent higher multiplicity of hits as compared to hadron showers of the same energy. In the case of hadron showers there is a larger fraction of signal hits with energy below threshold, where additional energy deposited by pileup can fire the ToA, giving rise to a larger fraction of hits fired by combined pileup plus signal contributions. The

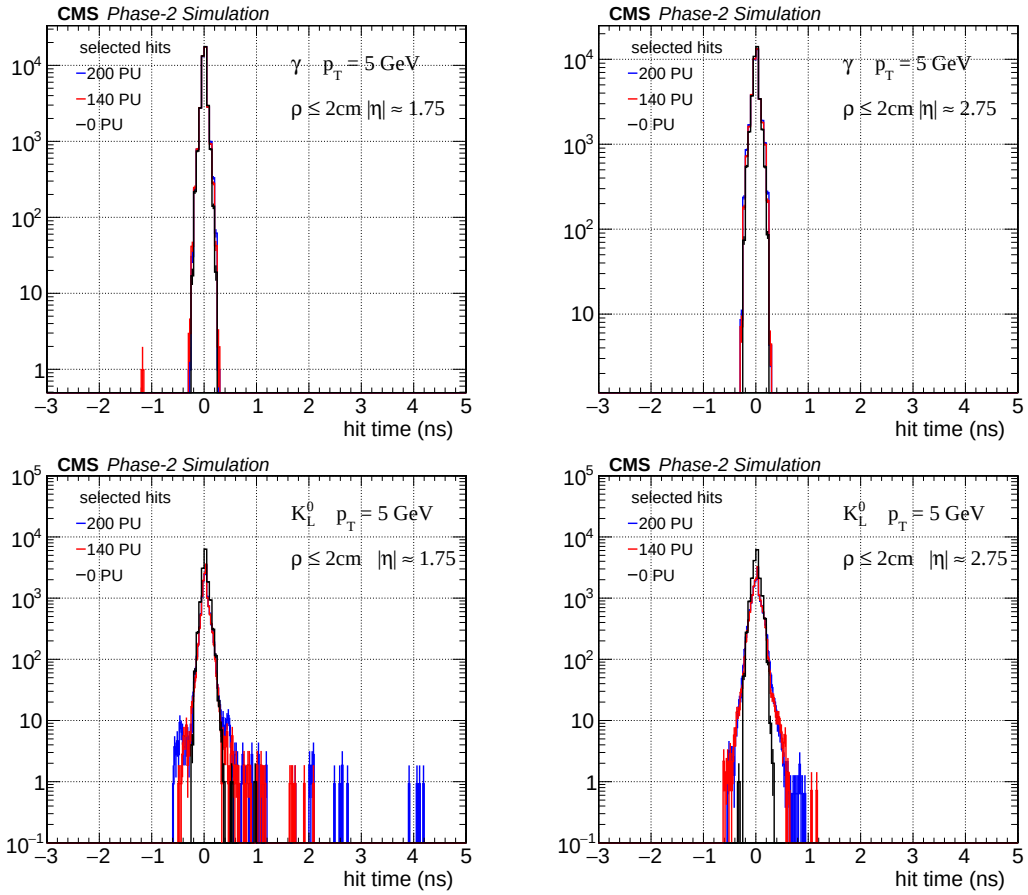


Figure 10.32: Time distribution of the hits selected in the interval of highest density, for photon showers (top), and K_L^0 showers (bottom), in the pseudorapidity ranges $1.65 < |\eta| < 1.85$ (left), and $2.65 < |\eta| < 2.85$ (right).

time distribution of such hits is that of pileup hits.

The results shown so far take the signal production vertex to be at mean vertex position, and the mean vertex time, while in reality signal events will be distributed in space and time, just as pileup interactions. The effect of pileup may differ according to whether the signal event interaction is earlier in time, or closer to the endcap, than the mean, or later in time, or further from the endcap. To explore possible differences, shower initiated by photons and K_L^0s of $p_T = 10\,\text{GeV}$ were examined in three different configurations: with the production vertex at ($z = 0$, $t = 0$) and no pileup, with the vertex displaced along the z axis ($z = -10\,\text{cm}$, $t = 0$) and no pileup, and with a displaced vertex, ($z = -10\,\text{cm}$, $t = 0$) in the presence pileup corresponding to a mean of 140 interactions per bunch crossing.

Figure 10.33 compares the time distribution of the hits, within $\rho = 2\,\text{cm}$, for the three configurations. The distributions are shown for two η ranges. The signal hits are located at a time that corresponds to the time-of-flight from the production vertex, and emerge over the distribution of pileup hits. Applying the selection cuts removes many of the hits from pileup. The shoulder that appears on the left side of the time distribution of the hits in the displaced vertex sample, at $t \approx 0$, comes from the distribution of pileup hits, which peaks at $t = 0$. This has been verified by looking at the time distribution of hits in cylinders of radius $\rho = 2\,\text{cm}$ at random locations. For ($z = 10\,\text{cm}$, $t = 0$) the peak is shifted to an earlier time and the shoulder at $t \approx 0$ appears on the other side of it. The same features are also seen in the corresponding distributions for

K^0_L showers, shown in Fig. 10.34, for which a small shoulder remains even after the selection of hits in the region of highest density.

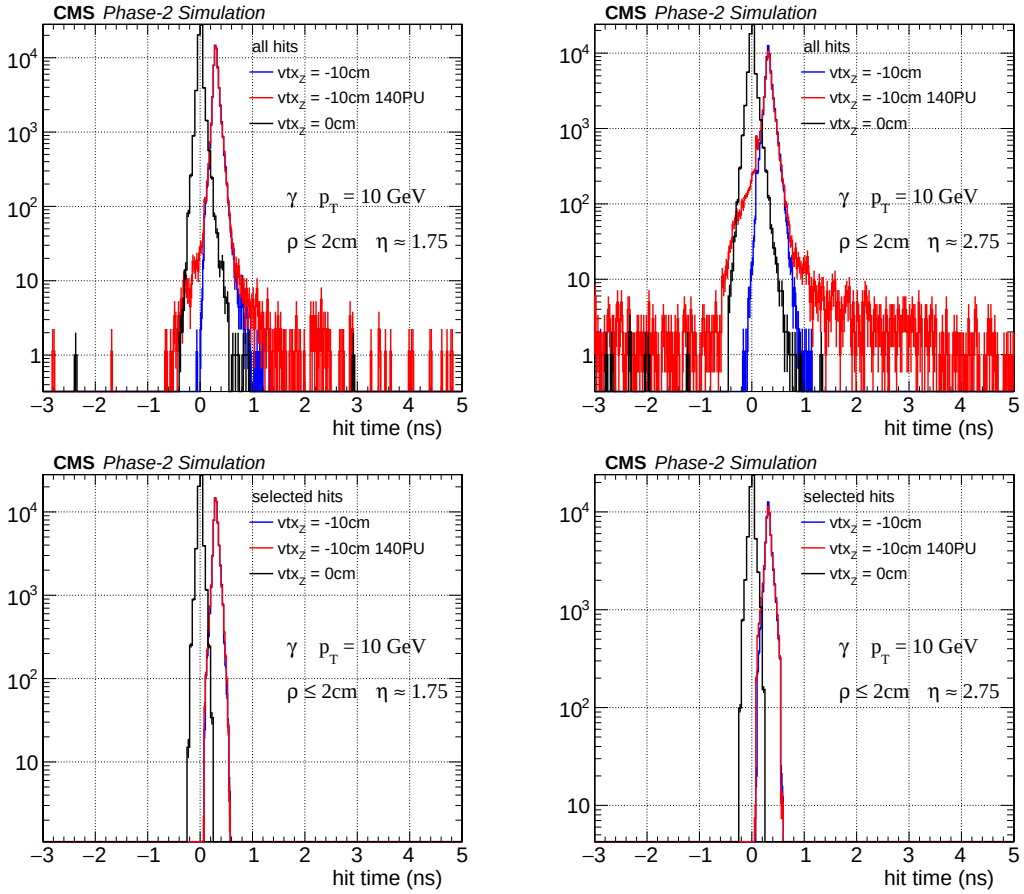


Figure 10.33: Time distribution of the hits for a photon shower, in the positive ($z > 0$) endcap, at $1.65 < \eta < 1.85$ (left), and in $2.65 < \eta < 2.85$ (right), for events produced both in the centre of the detector, and with a vertex displaced along the z axis. The time distributions are shown for all hits (top), and for the hits selected in the region of highest density (bottom). The distribution for displaced vertex events are shown both with and without the presence of pileup.

The measured time of the photon showers for the three configurations tested perfectly corresponds to the time-of-flight from the production vertex, also in presence of high pileup. The shoulders seen in the time distributions of hits in Figs. 10.33 and 10.34 do not result in any noticeable bias, and only in the case of K^0_L showers is a small degradation of the time resolution noticeable, as shown in Fig. 10.35.

The preliminary results on the precision measurement of the time of showers in the HGCAL show promising potential for helping global event reconstruction and rejection of energy from pileup at the HL-LHC.

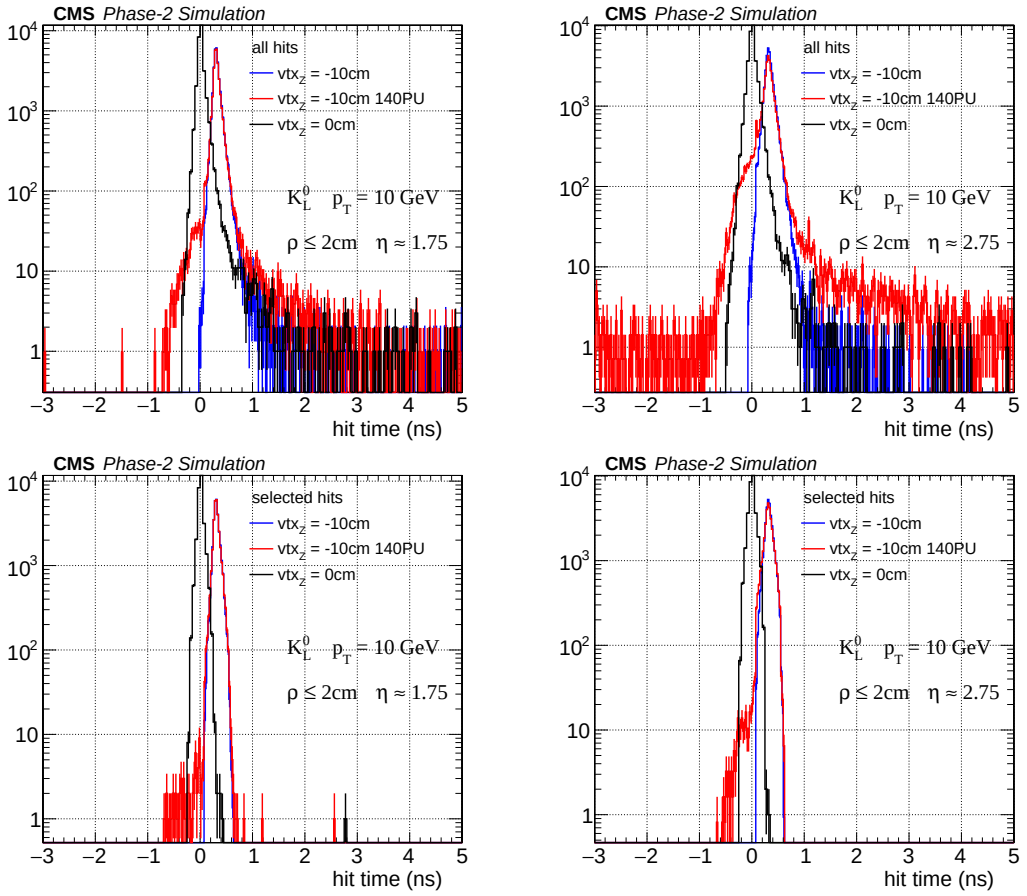


Figure 10.34: Time distribution of the hits for a K^0_L shower, in the positive ($z > 0$) endcap, at $1.65 < \eta < 1.85$ (left), and in $2.65 < \eta < 2.85$ (right), for events produced both in the centre of the detector, and with a vertex displaced along the z axis. The distribution for displaced vertex events are shown both with and without the presence of pileup.

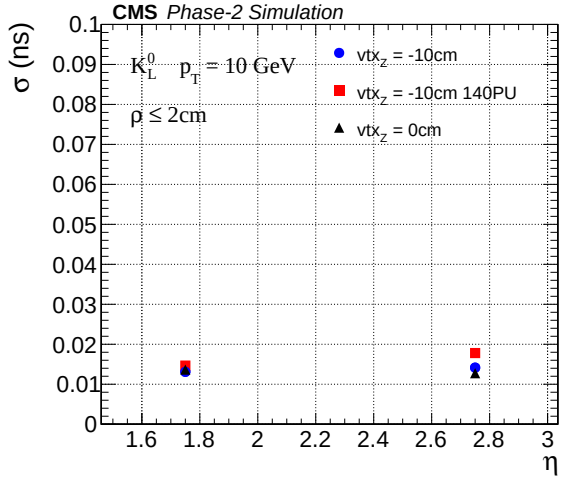


Figure 10.35: Time resolution for K^0_L showers for different positions of the production vertex along the z axis.

Chapter 11

Physics performance

11.1 Physics object performance

11.1.1 Electron identification

The electron identification strategy follows the one for Run 1 of the LHC [58], which is based on three sets of following observables: calorimeter-only variables, track-based variables, and track-calorimeter matching. These variables minimize their dependence on the isolation of the electron, since at analysis level additional isolation criteria that are optimized for the signal topology are usually applied on top of electron identification. Compared to the Phase-1 detector, the amount of material in the tracker is reduced by almost a factor of two [59] over the full pseudorapidity (η) range, which significantly reduces the amount of Bremsstrahlung. The performance results presented below do not yet fully exploit the power of granularity of the HGCAL.

The calorimeter-based variables are based on the multiclusters (described in Section 5.1), resulting from the aggregation in depth of the two-dimensional layer clusters in the HGCAL. Clusters are further gathered into superclusters following the characteristic η-ϕ pattern of the electromagnetic energy depositions (where ϕ is the azimuth). The reconstructed supercluster is used as seed for an electron-specific Gaussian-sum filter (GSF) track-reconstruction algorithm [60]. The seeding efficiency, and therefore the electron reconstruction efficiency reach 96%, and are independent of the pseudorapidity and the pileup (PU) from temporarily overlapping pp collisions. An electron track is often reconstructed using the generic tracking algorithm, optimized for minimum-ionizing particles. The comparison of the properties of the generic and GSF tracks, namely the number of hits and the χ^2 of the fit, conveys information useful to identify electrons that radiate a significant amount of energy when traversing the material in the tracker. They form the second set of observables. Finally, the matching of the track and the cluster in terms of energy and position forms the last set of observables.

The size of the area of interest for the reconstruction of variables relevant to the distribution of the shower energy is restricted to the electron multicluster, rather than to the supercluster, constituted from the electron and Bremsstrahlung photons, providing the possibility of minimizing the contribution from pileup. The impact of overlaying particles from pileup events is further reduced by considering only the energy depositions located within one Molière radius of 3.0 cm from the shower axis.

The high segmentation of the calorimeter provides precisely reconstructed identification characteristics through the principal component analysis (PCA). The PCA is run in an iterative manner, where each iteration is fed the position of the hits belonging to the cluster, weighted by their energies. The first iteration, carried out taking the hits belonging to the core of the clusters, provides an initial estimate of the barycenter of the shower as well as its axis. In the

second iteration, the hits are required to lie within one Molière radius of the shower axis. The position and axis direction are thereby refined and a third iteration, similar to the previous one, is eventually performed. The part of the cluster located within a cylinder of one Molière radius around the shower axis is used to define variables sensitive to the transverse and longitudinal shower-energy development (often referred to as "shower shape").

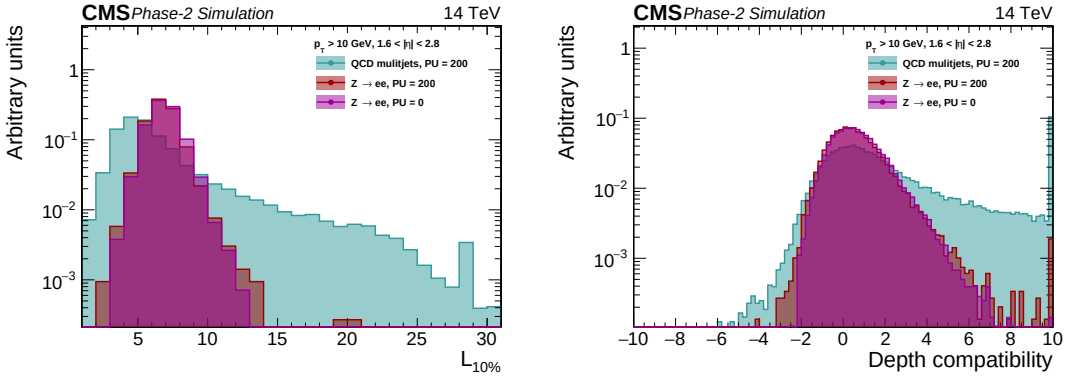


Figure 11.1: For signal and background electron candidates in presence of pileup as well as electrons without pileup, two examples of variables sensitive to the shower longitudinal development: (left) layer number for which the accumulated energy reaches 10% of the EE energy ($L_{10\%}$); (right) shower depth compatibility.

As example, the number of consecutive front layers in the calorimeters, where the cumulative energy reaches 10% of the energy contained in the EE($L_{10\%}$) is an observable sensitive to longitudinal-shower development. This observable is shown in Fig. 11.1 (left) for signal electrons from Z $\rightarrow$ ee events and for electron candidates in di-jet events. It retains a high discrimination power even in presence of HL-LHC pileup conditions. In addition, the calculation of the barycenter of a shower using the PCA reconstruction can be compared to the well-known parameterization of the longitudinal development of electromagnetic showers [61]. The compatibility is quantified through the difference between the expected depth predicted through the parameterization and the one measured normalized by the expected spread in the shower depth shown in Fig. 11.1 (right). The dependence of the stability on PU, as well as the quality of the discrimination against the background can be noted in the figure.

The definition of variables quantifying the lateral shower development profits from the high granularity and segmentation of the detector. In a plane orthogonal to the shower axis, the width of the shower is estimated within the cylinder defined above, along the tangential axis ($\mathbf{u}$) and along the axial direction ($\mathbf{v}$). Contrary to the former, the latter is, to first order, not affected by the Bremsstrahlung photons emitted within the tracker material. The distribution in the spread of the energy (σ_{vv}) along the $\mathbf{v}$ axis is presented in Fig. 11.2 (left). Electromagnetic showers tend to be much narrower than hadron showers. The effect of pileup tends to add an additional smearing to the width distribution.

The PCA brings in other useful observables in the context of a multivariate analysis, such as the eigenvalues and their uncertainties, that exhibit significant discriminating power between signal and background.

To illustrate the third set of variables that combine the tracker and calorimeter information, the geometrical matching between the GSF track and the electron cluster is shown in Fig. 11.2 (right) where the difference in η between the electron seed position and the track extrapolations is plotted for signal and background. The distribution for signal is much narrower than for background and not significantly affected by the pileup contribution.

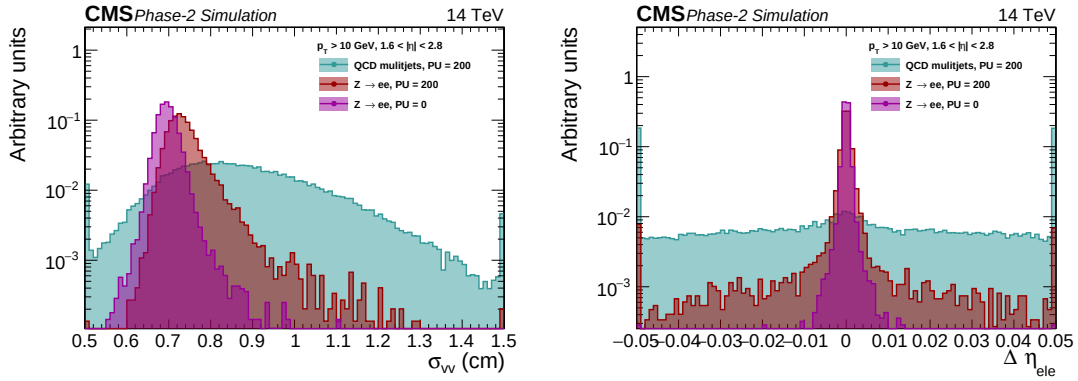


Figure 11.2: For signal and background electron candidates, respectively from $Z \rightarrow ee$ and multijet events, with and without pileup, (left) shower spread along the radial direction (σ_{VV}) and (right) distance in η between the electron cluster and the track extrapolation ($\Delta\eta_{\text{ele}}$).

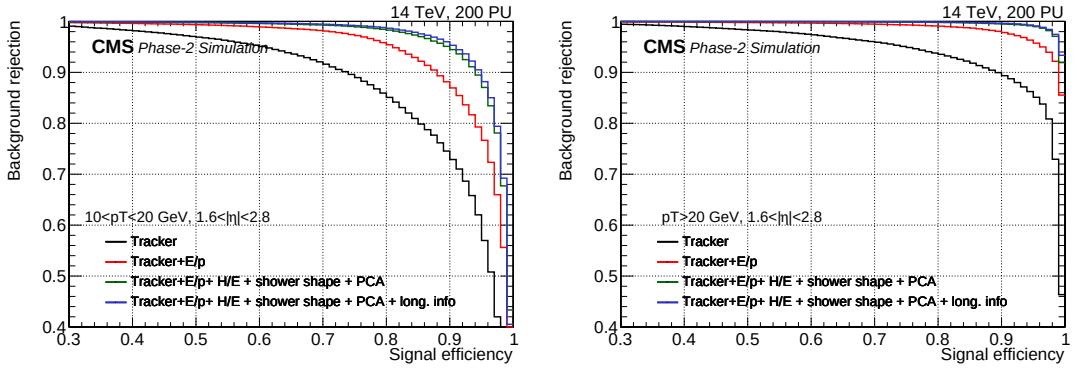


Figure 11.3: Purity as a function of the efficiency for electrons (left) with $10 < p_T < 20\,\text{GeV}$, and (right) with $p_T > 20\,\text{GeV}$ for different sets of input variables in the BDT multivariate estimator.

The resulting variables are combined into a Boosted Decision Tree (BDT) algorithm trained on $Z \rightarrow ee$ events with pileup for the signal, and multijet events for background. As observables in electron identification quickly evolve with p_T, the training is performed and the outcome assessed separately for $10 < p_T < 20\,\text{GeV}$ and for $p_T > 20\,\text{GeV}$. The results are presented in Fig. 11.3, where it can be observed that for a 95% signal efficiency, the background efficiency is 1% for $p_T > 20\,\text{GeV}$, and 10% for $10 < p_T < 20\,\text{GeV}$. Moreover, a tight working point with 0.1% background efficiency is defined in addition to the medium and loose working points with 1% and 10% background efficiency, respectively. In the same plots, equivalent BDT trainings are presented with reduced sets of input variables: beginning with track-based variables, a sizeable gain in performance is achieved through the energy momentum comparison; and the addition of PCA-based variables leads to further improvement in performance. Finally, the addition of extra information related to the longitudinal development, such as the compatibility in shower depth, or in the layer of the 10% cumulated EE fraction ($L_{10\%}$), improves the performance only for low p_T electrons, and is therefore only important in the $10 < p_T < 20\,\text{GeV}$ range.

The evolution of the signal and background efficiencies as function of the electron candidate p_T, is displayed in Fig. 11.4 (left) for a multivariate selection yielding a 95% total efficiency for $p_T > 20\,\text{GeV}$. For electrons with $p_T > 20\,\text{GeV}$, the efficiency as a function of $|\eta|$ is represented in Fig. 11.4 (right). An increase in efficiency can be observed from 92% at $|\eta| = 1.5$ to 98% at $|\eta| = 3$. The background efficiency also tends to increase at high $|\eta|$. In the large range of pileup values considered and despite the sensitivity of some of the multivariate input variables, the

efficiency is found to decrease only slightly with the number of simulated pileup events as can be seen in Fig. 11.5 (left). The efficiency does not appear to depend on the pileup density along the beam axis as observed in Fig. 11.5 (right).

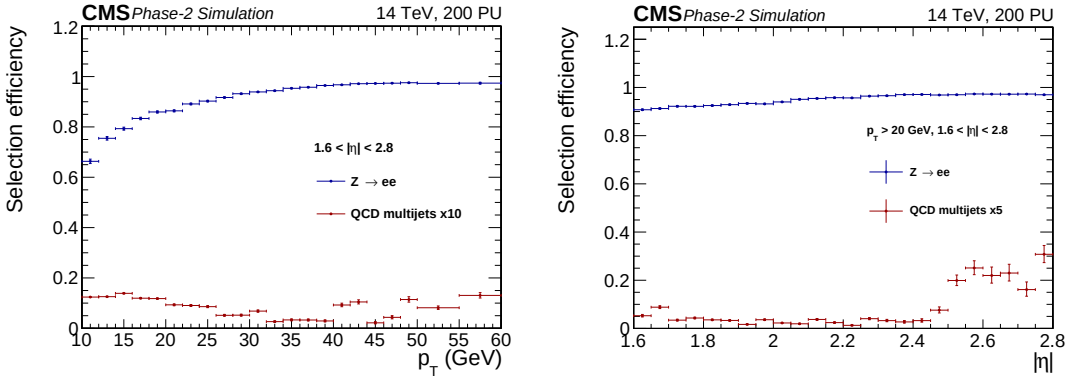


Figure 11.4: Evolution of the signal (blue) and background (red) efficiencies (left) as a function of p_{T} for a high-efficiency selection, and (right) as a function of the cluster $|\eta|$ for electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$.

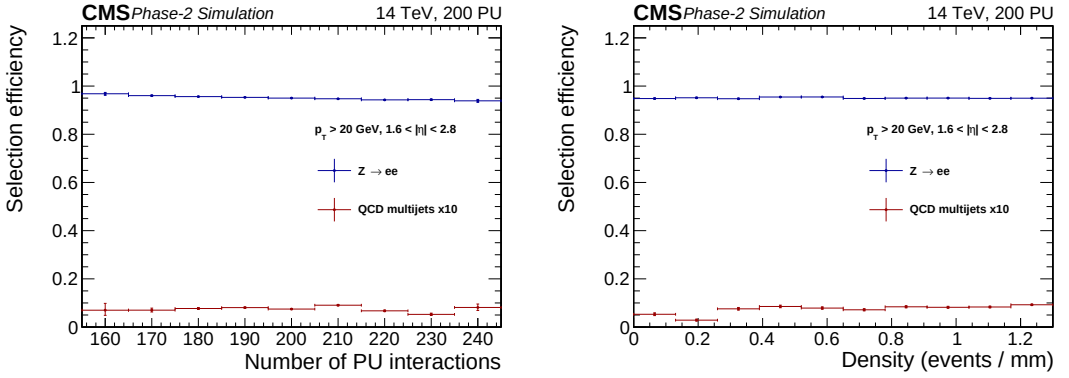


Figure 11.5: For electron candidates with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$: evolution of the signal (blue) and background (red) efficiencies (left) as a function of the number of simulated pileup events and (right) as a function of the pileup density along the beam axis.

11.1.2 Photon identification

In this section, we explore the reconstruction and identification of true photons in the HGCAL. Photons are identified mainly through three components: shower shape in the calorimeters, which distinguish electromagnetic from hadronic showers; track isolation that reject electrons and charged hadrons; and calorimeter-based isolation that provides additional rejection of neutral hadrons as well as of true photons in jets. Because the general features of photon and electron showers in the HGCAL are similar, the same 3-dimensional shower-shape variables already employed for the electron identification are used to identify photons in the HGCAL. To lower the rate at which electrons are misidentified as photons, additional variables are constructed based on the presence of nearby tracks and pixel hits. The HGCAL geometry enables the construction of depth-sensitive calorimeter isolation variables that can be used to reject nonprompt photons, as described in the previous Section 11.1.1.

All variables are fed into a BDT that is trained on a mixture of simulated γ + jets, $\mathrm{Z} \to \mathrm{ee}$, and multijet events, where signal photons are defined as reconstructed photons matched within $\Delta R(\eta, \phi) < 0.1$ relative to well-isolated generated photons from the initial hard scatter event,

within the kinematic phase space: $p_{\mathrm{T}}^{\gamma} > 25\,\mathrm{GeV}$ and $1.6 < |\eta^{\gamma}| < 2.8$. Performance in the overlap region between the barrel and end ECAL detectors ($1.48 < |\eta| < 1.65$) is not evaluated here. Misidentified photons are defined as reconstructed photons found in the same kinematic phase space, but not matched to an isolated, generated photon. The misidentified photons are reweighted to have identical (p_{T}, η) distributions as signal photons.

The photon identification efficiency and corresponding misidentification probability are shown in Fig. 11.6. The efficiency is defined as the fraction of true reconstructed photons that pass a given threshold in BDT score output, and the photon misidentification probability is defined as the number of reconstructed background photons passing the same BDT score threshold divided by the total number of events processed. Both photon efficiency and misidentification probability are evaluated for only the $\gamma +$ jets simulation. A comparison to the performance of a Run 2 cut-based ID, evaluated on a similar sample of $\gamma +$ jets events simulated with the Run 2 detector and beam conditions is also made in Fig. 11.6.

Two working points (WP) are defined: a "Tight ID," WP targeting a true photon selection efficiency of 85% with a photon misidentification probability of 3.5%; and a "Loose ID," WP targeting a true photon selection efficiency of 95% with a photon misidentification probability of 9.9%. For both working points, the selection efficiency as a function of true photon energy is shown in Fig. 11.7.

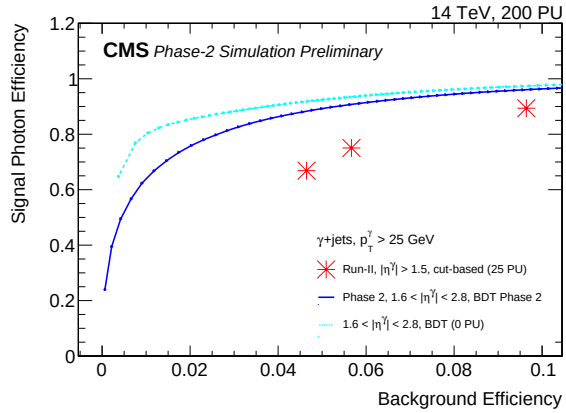


Figure 11.6: Photon efficiency versus photon-misidentification probability in simulated $\gamma +$ jets events for the BDT training described in the text. Signal photons are matched within $\Delta R(\eta, \phi) < 0.1$ to isolated photons generated within the kinematic phase space $p_T^{\gamma} > 25\,\mathrm{GeV}$ and $1.6 < |\eta^{\gamma}| < 2.8$. Misidentified photons are defined as reconstructed photons found in the same kinematic phase space but not matched to an isolated generated photon. The performance of a Run 2 cut-based ID is also presented, evaluated on a similar sample of $\gamma +$ jets produced using the Run 2 conditions (average pileup of 25 pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$).

11.1.3 Jet performance

Jets reflect the remnants of quarks and gluons produced in high-energy processes such as scattering of partons located within the colliding protons at the LHC. The detailed understanding of both the transverse momentum resolution and the energy scale of the jets is necessary to properly estimate the systematic uncertainty for many analyses. Jet energy corrections (JEC) refer to the calibration of the energy scale of jets. Jets are reconstructed by clustering the four-momentum vectors of the particle-flow (PF) algorithm [2] and realistic SimClusters (described in Section 10.1) using the anti-k_T algorithm [62] with a radius $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$ in the (η, ϕ) space. The PF algorithm parameters are not yet optimized for the HGCAL.

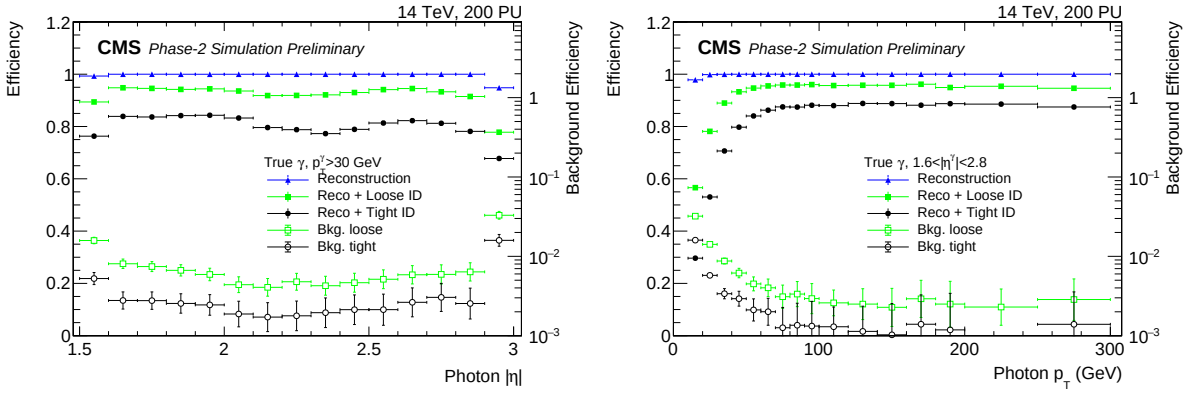


Figure 11.7: Photon reconstruction efficiency, identification efficiency, and photon misidentification probability, for two ID working points, (left) as a function of generated photon $|\eta|$ and (right) p_{T}. The photon reconstruction efficiency is defined as the efficiency for which a reconstructed photon is found within $\Delta R(\eta, \phi) < 0.1$ of a generated prompt photon. Identification efficiencies for signal photons are relative to the generated prompt photon. Misidentified photons are defined as reconstructed photons not matched to an isolated generated photon.

11.1.3.1 Pileup mitigation

Pileup-mitigation techniques are employed to reduce the effect of additional tracks and calorimetric energy depositions from multiple inelastic proton-proton collisions in the same or adjacent beam crossings. The charged-hadron subtraction (CHS) method removes PF candidates if the corresponding tracks are identified as originating from pileup vertices. The PF candidates with tracks not associated with a reconstructed vertex are kept, because vertex reconstruction and identification is not yet fully efficient. Jets clustered with CHS are referred to as PF+CHS jets. Recently CMS has implemented and developed the PileUp Per PartIcle Identification approach (PUPPI) [63], which uses local event properties and tracking information to further reduce effects from pileup interactions. Charged particles within the tracker volume not associated with the primary vertex are used to define a weight, thereby providing the probability that a given neutral PF candidate originates from a pileup interaction. The weights for charged PF candidates that originate from the primary vertex and pileup vertices are set to 1 and 0, respectively. The weights are used to rescale the four-momenta of the PF candidates. Jets clustered with PUPPI are referred to as PF+PUPPI jets.

11.1.3.2 Jet resolution

The JEC sequentially correct the offset energy from pileup interactions as well as the uniformity of the detector response to hadrons [64]. Figure 11.8 shows the resolution in jet energy response, $\sigma(p_{\mathrm{T}}^{\mathrm{Gen}})/(p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{Gen}})$, as a function of jet p_{T} at the particle-level ($p_{\mathrm{T}}^{\mathrm{Gen}}$) for jets with $|\eta| < 1.3$ (left), $1.3 < |\eta| < 1.7$ (middle), and $1.3 < |\eta| < 2.8$ (right). The particle-level jets are clustered from generated stable (decay length $c\tau > 1$ cm) final-state particles, excluding neutrinos. The jet resolution is taken as the width of a Gaussian function fitted to the distribution of the corrected response. Only modest degradation of the jet resolution are observed relative to the central part of the detector for jets with $1.7 < |\eta| < 2.8$. The resolution of jet energy response as a function of the density of pileup events along the beam axis is shown in Fig. 11.9. Only a small degradation in the jet energy resolution is observed as a function of the pileup density.

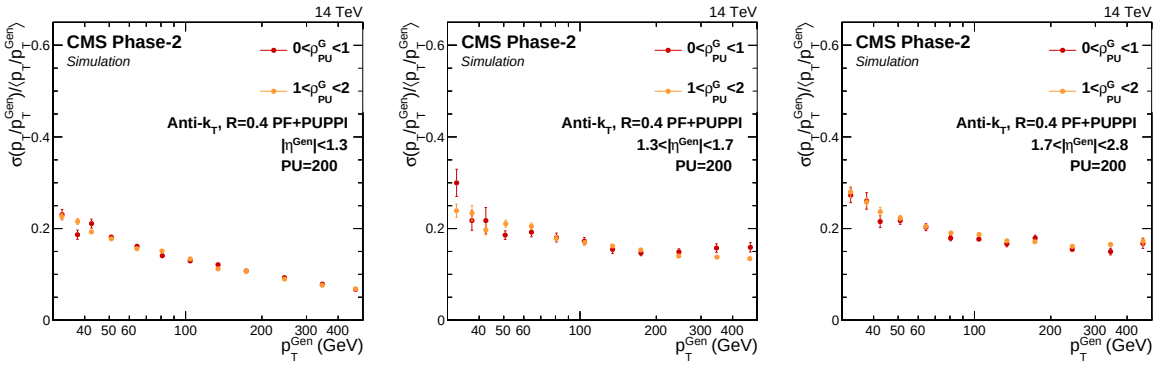


Figure 11.8: Corrected jet response resolution in (left) $|\eta| < 1.3$, (middle) $1.3 < |\eta| < 1.7$, and (right) $1.7 < |\eta| < 2.8$ as a function of $p_{\mathrm{T}}^{\mathrm{Gen}}$ for PF+PUPPI jets in the PU=200 sample.

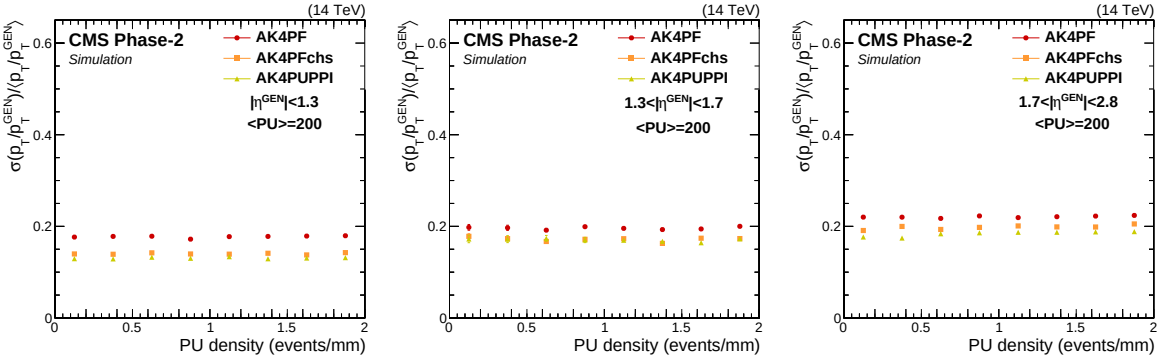


Figure 11.9: Corrected jet response resolution in (left) $|\eta| < 1.3$, (middle) $1.3 < |\eta| < 1.7$, and (right) $1.7 < |\eta| < 2.8$ as a function of the Gaussian measured pileup density for six jet collections in the PU=200 sample. The various marker colors indicate the jet type and all jets have been matched to a particle-level jet with $90\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{Gen}} < 120\,\mathrm{GeV}$.

11.1.3.3 Pileup jets

Particles originating from pileup interactions may accidently be clustered into overlapping low-p_{T} jets that combine to form a single high-p_{T} jet, referred to as a pileup jet. The rate of these pileup jets is quantified by the ratio of the average number of jets in a given p_{T} bin to the average number of reconstructed jets matched to a particle-level jet. Figure 11.10 shows the pileup jet rate as a function of number of pileup interactions (left) and pileup density (right) for the PUPPI jets. The pileup jet rate for jets with $1.3 < |\eta| < 3.0$ is only moderately degraded relative to the central barrel part of the detector.

11.1.4 Performance of b jet tagging

The ability to distinguish, or tag, the jets arising from the hadronization of b quarks (b jets) from those originating from light partons plays a crucial role in the identification of many production processes.

Two b tagging algorithms are used for the Phase-2 studies: the "cMVAv2" algorithm, relying on a BDT [65], and the "DeepCSV" algorithm, based on a deep neural network [66]. Both have been trained on Run 2 simulations with an ideal Phase-1 pixel detector, but not specifically retuned to the Phase-2 detectors. To provide best b jet identification, cMVAv2 and DeepCSV are applied within the ranges $|\eta| < 1.5$ and > 1.5, respectively.

The choice of the primary interaction vertex (PV) is a critical component of the b tagging per-

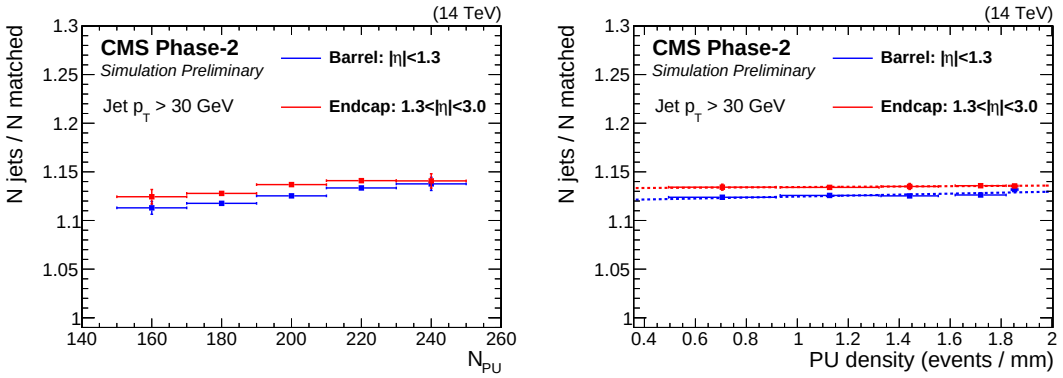


Figure 11.10: Ratio of the number of reconstructed PUPPI jets to the number of reconstructed PUPPI jets matched to particle level jets from the hard scatter as a function of: (left) number of pileup interactions (N_{PU}) and(right) pileup density for jets with $p_T > 30$ GeV.

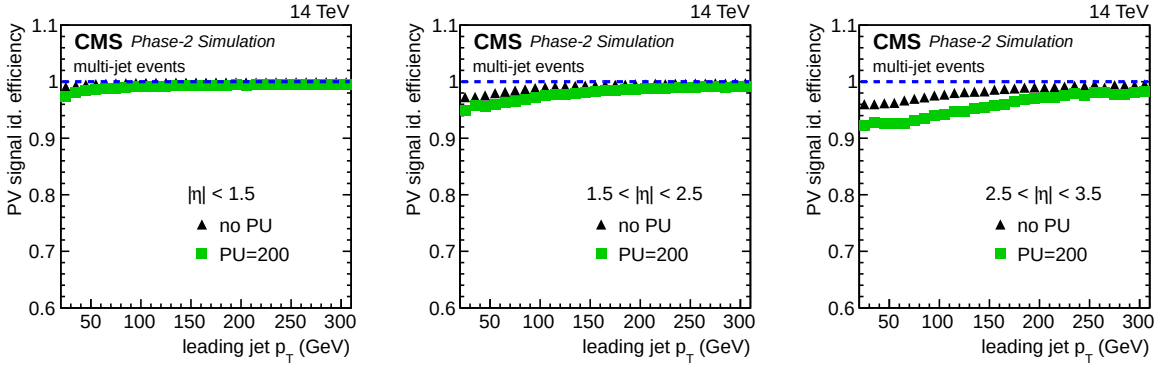


Figure 11.11: Efficiency to reconstruct the hard interaction vertex and to identify it correctly as the primary vertex (PV), as a function of the leading jet p_T in simulated multi-jet events with ≥ 2 jets. The leading jet, i.e. the jet with highest p_T, is contained in the $|\eta|$ range (left) 0–1.5, (middle) 1.5–2.5 or (right) 2.5–3.5. The identification efficiency for PV signal jets increases with the leading jet p_T. Compared to events without pileup (black triangles), it is slightly lower at PU = 200 (green squares).

formance. The efficiency to reconstruct the hard interaction vertex and to identify it correctly as signal primary vertex is shown in Fig. 11.11 as a function of the leading jet p_T in multi-jet events. As expected, the efficiency increases with the jet momentum due to the presence of higher momentum tracks, and it is smaller at high pileup, especially in the forward region due to tracks from overlapping pileup jets.

The performance of the b tagging is evaluated in $t\bar{t}$ events for jets with $p_T > 30$ GeV. The misidentification probability for light partons is shown in Fig. 11.12 as a function of the b jet tagging efficiency, for different $|\eta|$ regions. The trend observed in b tagging performance as a function of pseudorapidity is mainly related to tracking resolution, although some dependence arises from the use of b tagging algorithms that are not optimized for the Phase-2 conditions. The increased spatial acceptance of the pixel detector provides b tagging capability also for $|\eta| > 2.5$.

Figure 11.13 shows both the light-parton misidentification probability and the charm-jet tagging efficiency as a function of the b jet tagging efficiency. The pseudorapidity regions of the barrel ($|\eta| < 1.5$) and endcap ($1.5 < |\eta| < 3.0$) calorimeters are shown separated. The b tagging

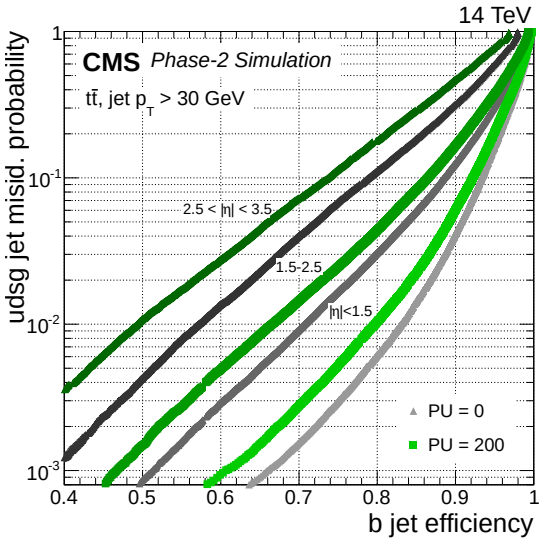


Figure 11.12: Performance of the b tagging in simulated $t\bar{t}$ events, expressed as misidentification probability for light-partons jets (udsg) as a function of the b jet tagging efficiency for events without pileup (gray) and with PU = 200 (green). Jets with $p_T > 30\,\text{GeV}$ are used in $|\eta|$ ranges of: 0–1.5, 1.5–2.5 and 2.5–3.5. The cMVAv2 and DeepCSV b tagging algorithms are used for jet $|\eta|$ within 0–1.5 and 1.5–3.5, respectively. The b tagging benefits from the good jet reconstruction in all regions.

benefits from the good jet reconstruction in both regions.

For $t\bar{t}$ events generated and reconstructed with an average pileup of 200, the b jet tagging efficiency is shown in Fig. 11.14 as a function of the overall pileup density, given in pp collisions per mm along the z axis. The barrel and endcap calorimeter regions are separated and the misidentification probability for udsg jets is fixed at 0.01 for each entry in the figure. A moderate decrease of the efficiency is observed as a function of the pileup density.

Finally, in multi-jet events, the tagging efficiencies for b and c jets are shown in Fig. 11.15 as a function of jet p_T and for different η intervals. As above, the misidentification probability for udsg jets is fixed at 0.01 for each entry in the figure. Compared to events without pileup, the b jet tagging efficiency remains large at high pileup in all p_T and $|\eta|$ regions of interest.

11.1.5 Performance of τ leptons

Reconstruction of τ leptons in their hadron+ν decays (referred to as τ_h final states) is performed using the "hadrons-plus-strips" algorithm [67]. τ leptons decay primarily to one charged hadron, one charged hadron and a π^0, or three charged hadrons. In this algorithm, electrons and photons, are combined in "strips" to construct π^0 candidates. The π^0 candidates are combined with PF charged hadrons to form the τ lepton candidate. In this study, the τ leptons come from the decay of Z bosons. Background events are simulated using the Z+jets processes, in which a jet can be misidentified as a τ_h lepton.

Figure 11.16 displays the reconstructed transverse momentum of the τ_h candidates with different pileup conditions for the HGCAL pseudorapidity $1.5 < |\eta| < 3.0$. The reconstruction efficiency is stable, and does not depend on running conditions or the physical process that produces for τ leptons. The right plot shows the misidentification probability of jets as τ_h lepton, which increases with p_T happens because high p_T jets are more collimated, hence have a

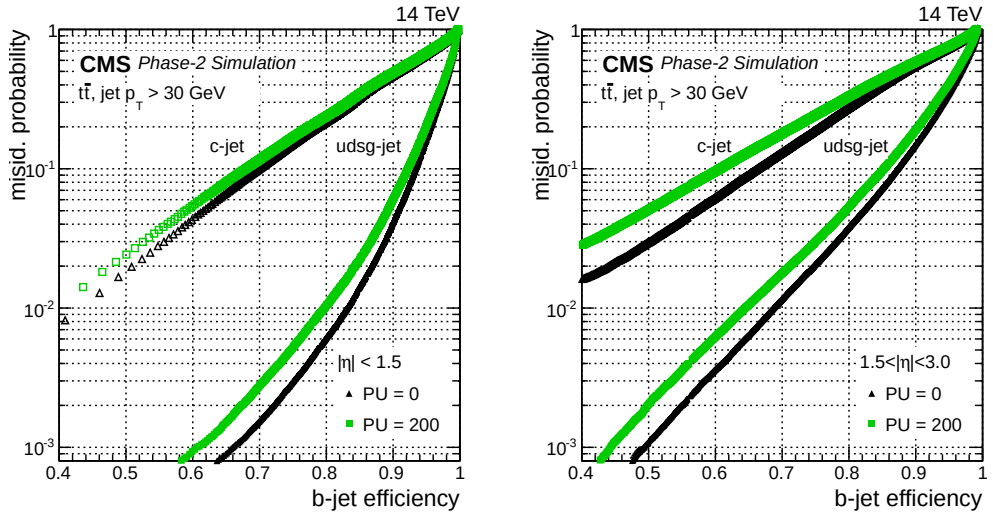


Figure 11.13: Performance of b tagging in simulated $\mathrm{t\bar{t}}$ events, expressed as misidentification probability for light-partons (udsg or c) jets, as a function of b jet tagging efficiency. Jets with $p_{\mathrm{T}} \geq 30$ GeV and $|\eta| < 1.5$ (left), or within 1.5–3.0 (right) are used in the analysis. The performance remains good at high pileup, both in the barrel and endcap calorimeter regions.

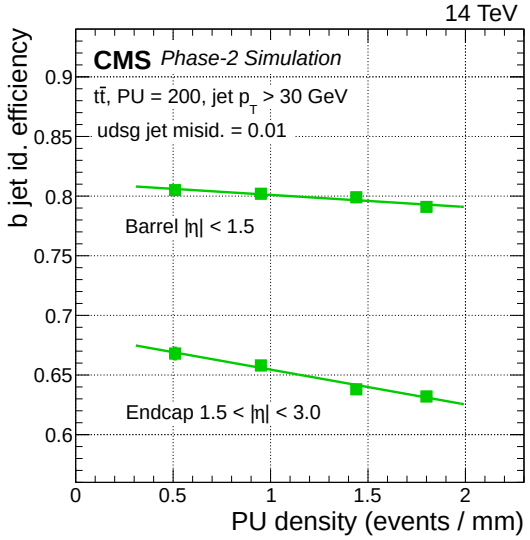


Figure 11.14: Efficiency of b jet tagging as a function of the density of pileup events along the beam axis (z). The efficiency is computed for a fixed misidentification probability of light-parton jets (udsg) of 0.01. Results are based on $\mathrm{t\bar{t}}$ simulation with an average pileup of 200. Linear fits are superimposed for the pseudorapidity regions of the barrel and endcap calorimeters, showing a moderate decrease of the b jet tagging efficiency as a function of the pileup density.

higher probability of mimicking τ_h decays. For all plots in this section the τ_h candidates are required to pass only the reconstruction, and not the isolation requirements. The performance of the τ_h reconstruction is similar to that achieved in the recent Run 2 CMS data taking. Also, it should be noted that the τ reconstruction algorithm does not take any advantage of information provided by the high granularity of the HGCAL and has room for much improvement.

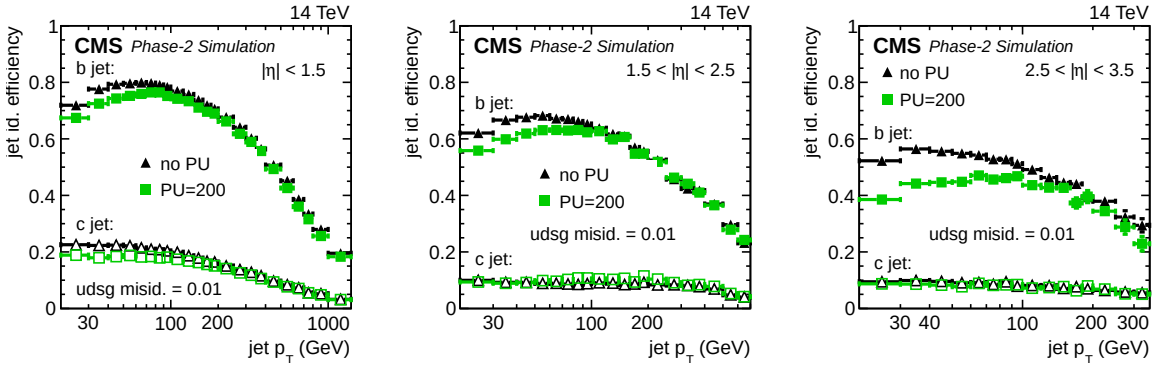
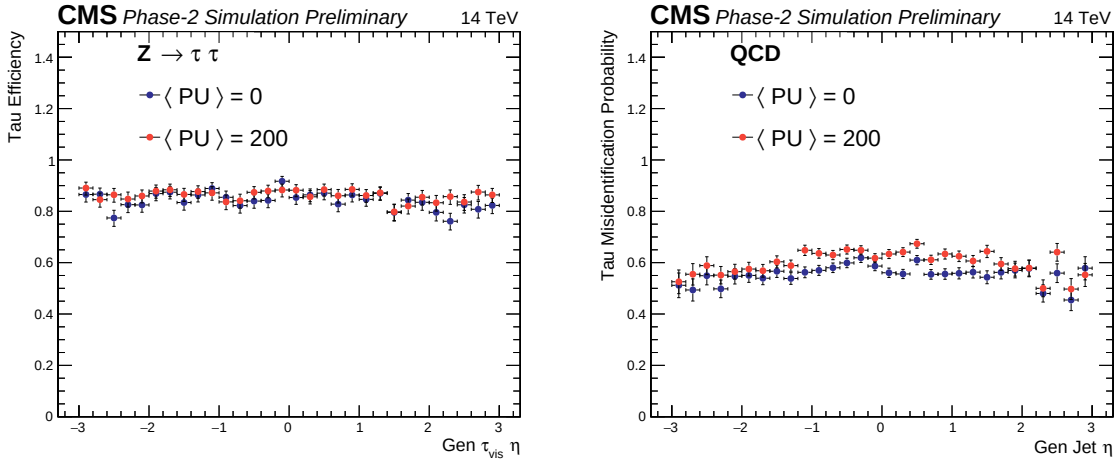


Figure 11.15: Tagging efficiencies for prompt b jets (filled symbols) and prompt c jets (open symbols) as a function of the jet p_{T} in simulated multi-jet events. The tagging efficiencies are evaluated for an average misidentification probability of 0.01 for light parton jets (udsg), and shown for PU = 0 (black triangles) and 200 (green squares). For three $|\eta|$ ranges: 0–1.5 (left), 1.5–2.5 (centre), and 2.5–3.5 (right). The cMVAv2 and DeepCSV b tagging algorithms are used for jet $|\eta|$ within 0–1.5 and 1.5–3.5, respectively.

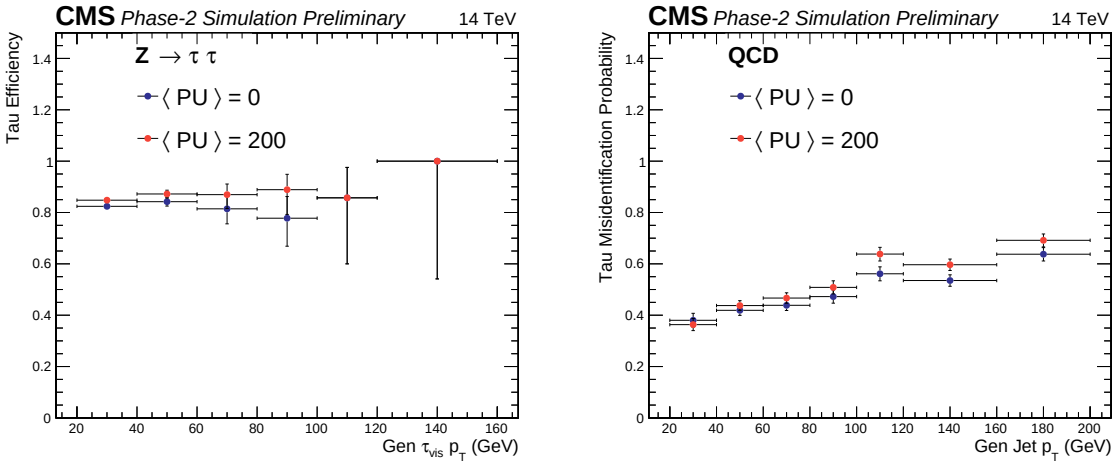


(a) (b)

Figure 11.16: τ_h lepton (left) efficiency and (right) misidentification probability as a function of pseudorapidity.

11.1.6 Performance of missing E_T

Neutrinos and other weakly-interacting neutral particles cannot be directly detected in the CMS detector, but their presence can be inferred from the momentum imbalance in the plane transverse to the beam direction. The magnitude of the momentum-imbalance vector is referred to as missing transverse energy and is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$. The precise measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the presence of a large number of pileup interactions is crucial for SM physics involving Higgs and W bosons, and top quarks. In addition, $E_{\mathrm{T}}^{\mathrm{miss}}$ plays a critical role in many searches beyond the SM, such as supersymmetry and collider-based dark matter searches. Pileup interactions significantly degrade the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution [68]. The PUPPI algorithm is used to mitigate the effect of pileup interactions. The PUPPI missing transverse momentum is defined using PF candidates with corresponding transverse momenta weighted through the PUPPI method, as



(a) (b)

Figure 11.17: τ_h lepton (left) efficiency and (right) misidentification probability as a function of transverse momentum.

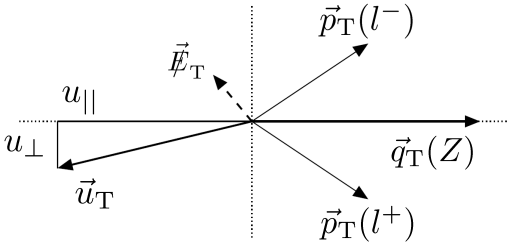


Figure 11.18: Illustration of $Z \to \mu^+\mu^-$ event kinematics in the transverse plane. The vector $\vec{u}_T$ denotes the vectorial sum of all particles reconstructed in the event, except for the two muons from Z boson decay.

described in Section 11.1.3.1. The p_{T} of the jets in the PUPPI missing transverse momentum are corrected to the particle level p_{T}, as discussed in Section 11.1.3.2, to reduce the bias in the $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from inefficiencies and nonlinearities in the detector.

The performance of $E_{\mathrm{T}}^{\mathrm{miss}}$ is studied using events with a Z boson decaying to two muons. There is no genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ in the bulk of $Z \to \mu^-\mu^+$ events. Figure 11.19 shows the PUPPI $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in $Z \to \mu^+\mu^-$ events for PU$=$ 200. The PUPPI $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in Run 2 is also shown for comparison. The well-measured Z boson can be used as a reference (both for the momentum scale and any unique event axis) to measure the scale and resolution in $E_{\mathrm{T}}^{\mathrm{miss}}$. Momentum conservation in the transverse plane leads to $\vec{q}_{\mathrm{T}} + \vec{u}_{\mathrm{T}} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}} = 0$, where $\vec{q}_{\mathrm{T}}$ is the reconstructed vector boson momentum in the transverse plane, and $\vec{u}_{\mathrm{T}}$, defined as the vectorial sum of the transverse momenta of all particles except the decay products of the Z boson. This hadronic recoil can be projected onto the axis defined by $\vec{q}_{\mathrm{T}}$ as illustrated in Fig. 11.18. The response is defined by the mean of scalar quantity $-u_{||}/q_{\mathrm{T}}$, where the minus sign is included because the value of $u_{||}$ is typically negative. The resolutions in $u_{||}$ ($u_\perp$) are defined as the standard deviation of a Gaussian function fitted to the $u_{||}$ ($u_\perp$) distributions.

Figure 11.20 shows the response-corrected resolution of the perpendicular and parallel components of the hadronic recoil as a function of the pileup density. A resolution of about 25 GeV is achieved in the $u_\perp$ component using PUPPI, with the upgraded detector in events containing

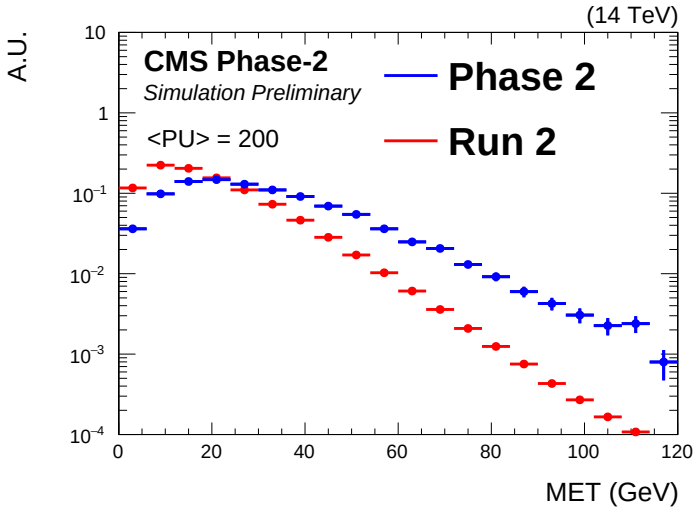


Figure 11.19: The PUPPI $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the Phase-2 detector for average pileup of 200 in $Z \to \mu\mu$ events. The PUPPI $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in Run 2 is shown in red.

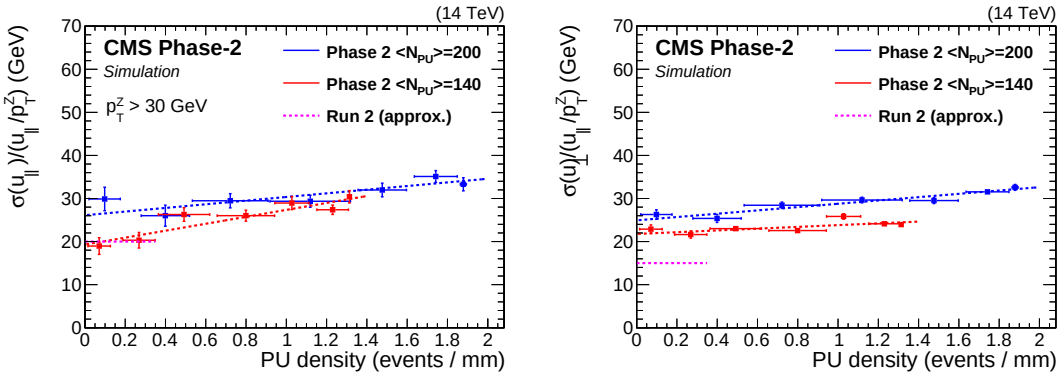
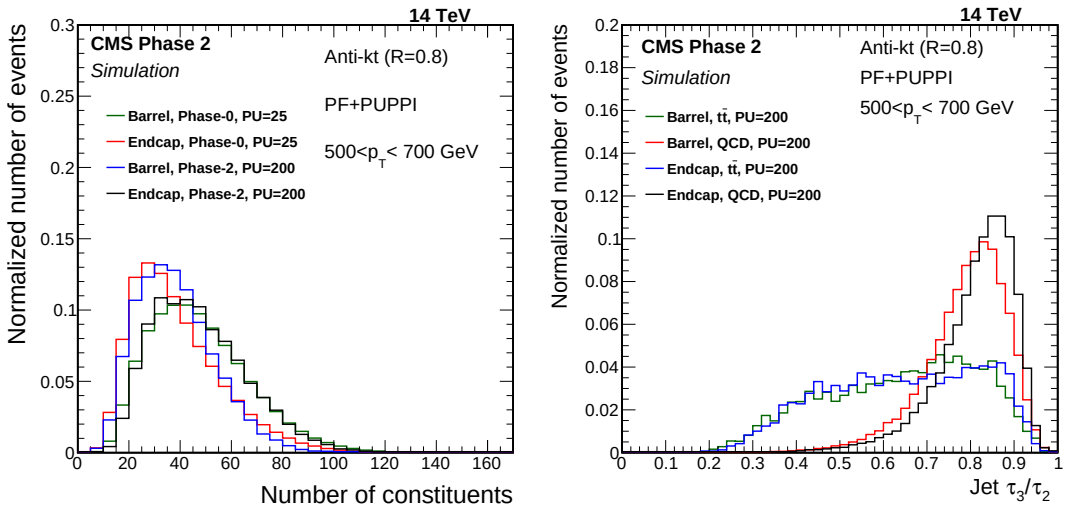


Figure 11.20: Response corrected (left) perpendicular and (right) parallel PUPPI components in the resolution of the jets in $Z \to \mu^+\mu^-$ events as a function of the pileup density in events/mm for $\mu = 200$ (blue) and $\mu = 140$ samples. Magenta, dashed lines indicate the performance achieved in Run 2.

200 pileup interactions. For comparison the corresponding resolution in Run 2 is indicated by a dotted magenta line. There is a modest degradation of the resolution with increasing pileup density.

11.1.7 Jet tagging of highly-boosted W and, Higgs bosons and top quarks

Highly boosted W, Z, and Higgs bosons, as well as top quarks, are identified in their decays through jet substructure techniques. The basic observables used by CMS [69] to distinguish these boosted objects from ordinary quark and gluon initiated jets are the softdrop jet mass [70], the N-subjettiness ratios τ_2/τ_1 (for W, Z, Higgs) and τ_3/τ_2 (for top) [71], and dedicated b tagging techniques for jets containing multiple quarks [65, 72]. The softdrop jet mass is a measure of the mass of the boosted object, where soft QCD radiation is removed from the jet before calculating its invariant mass, such that background jets initiated by quarks or gluons can be rejected. The quantities τ_2/τ_1 and τ_3/τ_2 refer to ratios of jet observables, that sum the angular distances of jet constituents to their nearest subjet axis, to quantify the likelihood that a jet has N or more subjets. Such techniques require a highly granular reconstruction of the distribution



(a) (b)

Figure 11.21: Comparison of jet substructure observables for barrel ($|\eta| < 0.7$) and endcap calorimeters ($1.9 < |\eta| < 2.4$) in Phase-0 (PU=25) and Phase-2 (PU=200). (a) Number of jet constituents from quark or gluon jets in QCD multijet simulation. (b) τ_3/τ_2 for top quark jets in high mass resonant $t\bar{t}$ production and quark or gluon jets in QCD multijet simulation.

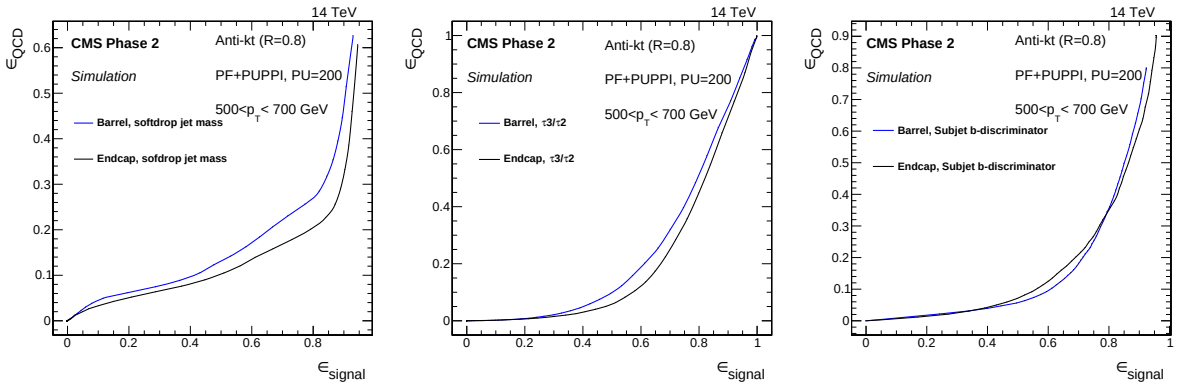
of energies within jets. While for low momentum jets, highly granular information can be obtained from the tracker using the particle-flow technique, at very high momenta, when tracking efficiency and resolution degrade, jet substructure reconstruction relies solely on barrel calorimeter and HGCAL. The increasing mass reach of the HL-LHC for many searches requires resolving substructure in jets with significantly higher boost. However, jet substructure reconstruction becomes particularly challenging at the HL-LHC due to the higher number of pileup interactions.

Figure 11.21 shows a comparison of jet substructure observables for jets reconstructed using the PF algorithm and PUPPI pileup suppression in the barrel calorimeter (jet $|\eta| < 0.7$) and in HGCAL ($1.9 < |\eta| < 2.4$) with the Phase-0 (PU=25) and Phase-2 (PU=200) detectors and pileup conditions. The number of jet constituents from quark and gluon jets in simulated QCD multijet events in Fig. 11.21(a) demonstrates the ability to reconstruct identification observables for quark and gluon jet. In the barrel region, the number of constituents decreases slightly in Phase-2 compared to Phase-0. However, an increase is observed in constituents in the HGCAL region, which may be attributed to the higher granularity of the endcap calorimeter and the higher number of pileup interactions. The τ_3/τ_2 distributions for top quark jets in high mass resonant $t\bar{t}$ production and quark or gluon jets in multijet simulation in Fig. 11.21(b) demonstrate excellent performance of the HGCAL in identifying subjets for highly-boosted W, Z, and Higgs bosons, and top quarks, at a level of quality similar to that of the barrel calorimeter. The softdrop jet mass resolution in the multijet simulation is shown in Table 11.1. The resolution is degraded by ≈30% for projected Phase-2 operations with 200 pileup interactions compared to the Phase-0 scenario with 25 pileup interactions and Phase-2 without pileup simulation in both barrel and endcap calorimeter regions of the detector. Nevertheless, this resolution is still sufficient for boosted object tagging, and using a dedicated PF reconstruction algorithm, can possibly achieve better pileup suppression for the Phase-2 detector.

Typical object tagging performance is summarized by computing the background rejection as a function of identification efficiency for common jet substructure observables in Phase-2

Table 11.1: Softdrop jet mass resolution for jets with $100 < p_T < 200$ GeV in barrel and endcap calorimeter.

Scenario	Softdrop jet mass resolution
Barrel calorimeter, Phase-0, PU=25	7.4%
Barrel calorimeter, Phase-2, PU=0	7.5%
Barrel calorimeter, Phase-2, PU=200	10%
Endcap calorimeter, Phase-0, PU=25	8.0%
Endcap calorimeter, Phase-2, PU=0	7.5%
Endcap calorimeter, Phase-2, PU=200	11%



(a) (b) (c)

Figure 11.22: Background efficiency (ϵ_{QCD}) as a function of signal identification efficiency (ϵ_{signal}) for common jet substructure observables in Phase-2 (PU=200) for barrel ($|\eta| < 0.7$) and end calorimeters ($1.9 < |\eta| < 2.4$): (a) Softdrop jet mass, (b) τ_3/τ_2, (c) subjet b tagging.

(PU=200). The softdrop jet mass in Fig. 11.22(a) is used for tagging W, Z, Higgs bosons, and top quarks, τ_3/τ_2 in Fig. 11.22(b) for top quarks (similar conclusions can be drawn with τ_2/τ_1 for W, Z and Higgs bosons) and subjet b tagging in Fig. 11.22 (c) for top quarks and Higgs bosons. The discrimination power achievable by the endcap of the Phase-2 detector for boosted W, Z, H, and top jets against quark/gluon jets is found to be similar or better than in the barrel region.

11.2 CMS physics channel results

11.2.1 $H \to \gamma\gamma$ in the VBF channel

Despite the small branching ratio predicted by the SM (0.2%), the $H \to \gamma\gamma$ decay channel provides a clean final state with an invariant mass peak that can be reconstructed with high precision. As a consequence, $H \to \gamma\gamma$ was one of the most important channels involved in the discovery of the Higgs boson and first measurements of its properties. This channel continues to be one of the best for a precise characterization of the Higgs boson in Run 2, and will remain so at the HL-LHC after the Phase-2 upgrade.

The majority of the analysis sensitivity comes from barrel-barrel photon pairs. The HGCAL will allow inclusion of photons in the analysis up to $|\eta| = 3.0$, compared with $|\eta| = 2.5$ in the Run 2 $H \to \gamma\gamma$ analysis [73]; this will increase the efficiency of the analysis by 12%. The resolution of endcap-endcap diphoton pairs (i.e. both in the HGCAL) is found to be very similar

to barrel-barrel pairs, representing a substantial improvement over the current detector. Additional gains from the HGCAL will come from its improved ability to identify jets from the vector boson fusion (VBF) production mode.

In VBF the final state includes two jets with a large rapidity gap. At least one jet is typically in the forward region where the HGCAL will provide significant improvement over the current detector. A powerful handle for the analysis is provided by quark-gluon discrimination, since in a genuine VBF event both jets originate from quarks, while in the ggH production they come mostly from radiated gluons. Gluon jets tend to be softer and more dispersed, while quark jets are highly collimated and contain fewer particles. Therefore, this study focuses on the expected improved discrimination power between quark and gluon jets due the the higher granularity of the endcap calorimeter when the jet shape information is fed to a multivariate discriminator.

Photons in the HGCAL are reconstructed using the multicluster-based approach described in Chapter 5. Barrel photon energy is defined as the sum of the fifteen highest energy hits in the supercluster. For the HGCAL photons, the energy is taken as the energy of the seed multicluster. This gives good resolution and is robust against pileup, giving similar energy resolutions from as a function of PU.

In this study, 200 pileup vertices are superimposed with detector simulation corresponding to a $1000\,\mathrm{fb}^{-1}$ ageing scenario. Events are selected using a preselection based on the Run 2 analysis [73]. This requires a pair of photons with $p_\mathrm{T} > 30\,\mathrm{GeV}, |\eta| < 3.0$ passing a tight identification cut, and diphoton invariant mass between 100 and 180 GeV.

Jets are reconstructed using particle flow and PUPPI pileup mitigation technique. A VBF preselection based on the Run 2 analysis is applied:

- leading photon $p_\mathrm{T} > \mathrm{m}_{\gamma\gamma}/3$, sub-leading photon $p_\mathrm{T} > \mathrm{m}_{\gamma\gamma}/4$;
- one jet with $p_\mathrm{T} > 40\,\mathrm{GeV}$ and one with $p_\mathrm{T} > 30\,\mathrm{GeV}$, both with $|\eta| < 4.7$
- the invariant mass of the two jets $\mathrm{m}_{jj} > 250\,\mathrm{GeV}$;

After this, the events undergo a final selection using multivariate techniques.

Three jet shape variables are used to construct a first BDT (the 'jet shape' BDT) to discriminate between ggH and VBF events. A second BDT (the 'dijet' BDT) that includes additional kinematic variables including those of the photons, is employed. In both cases the VBF events are treated as signal, while the background is given by the ggH events. The list of variables used in the two BDTs is reported in Table 11.2 and the performance of these classifiers is illustrated in Fig. 11.23. The jet shape variables alone give a significant discrimination between ggH and VBF, but the addition of kinematic variables clearly brings further improvement.

Separately, another BDT is trained against the background events (the 'background' BDT), with only VBF as signal. Two more photon variables are added for this discriminator:

- The η of the leading and subleading photons
- The cosine of the angle between the two leading photons in the transverse plane

The area under the ROC curve of the resulting classifier is 0.86.

Different working points are established by placing cuts on each of the dijet and background BDTs. Table 11.3 shows three working points obtained by varying the dijet BDT, which focuses on anti-ggH discrimination. A fixed cut on the background BDT cut is applied, balancing optimal rejection with signal efficiency. The number of signal events selected, the fraction of ggH and VBF events, and the diphoton and QCD background per GeV are shown. The results

Table 11.2: List of BDT variables

BDT	Variables
Jet shape	number of charged constituents (leading and subleading jets) length of the semi-minor axis of the ellipse in the η-ϕ plane, (leading and subleading jets) jet fragmentation distribution (leading and subleading jets) [69]
di-jet	In addition to the above: $p_T/m_{\gamma\gamma}$ (leading and subleading photons) cosine of the angle between the two leading photons in the transverse plane $\Delta\phi$ between the dijet and diphoton systems p_T (leading and subleading jets) $\Delta\eta$ and $\Delta\phi$ between leading and subleading jet m_{jj}

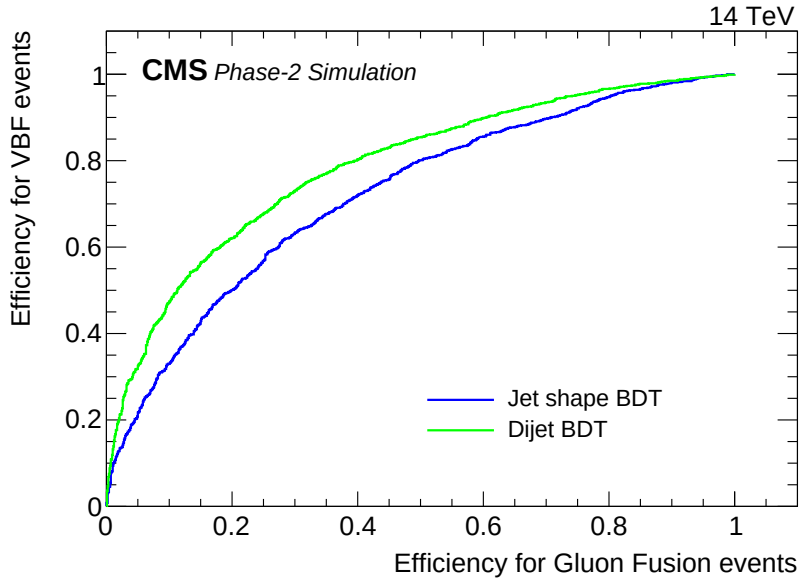


Figure 11.23: The ROC curves for the two trained BDTs. The classifier using only the three jet shape variables (blue line) has an area under the ROC curve of 0.71, whilst the second with access to additional kinematic variables (green line) has an area of 0.79. For comparison, the dijet BDT used in the Run 2 analysis has an area of 0.75.

Event Categories	SM 125GeV Higgs boson expected signal			Background
	Total	ggH	VBF	per GeV
WP 0	750	25.4 %	74.6 %	678
WP 1	1275	35.9 %	64.1 %	876
WP 2	1926	45.8 %	53.2 %	1353
Run 2 WP	3878	42.0 %	58.0 %	1984

Table 11.3: The signal and background yields for three working points (WP). The dijet BDT cut is varied, with a fixed cut on the background BDT. Number of events given is for 3000 fb^{-1} of collected data. The Run 2 WP contains the sum of selected events in all three VBF categories, extrapolated to 3000 fb^{-1}.

indicate that, with the HGCAL, the discriminating power between ggH and VBF is comparable to Run 2 despite the increase in amount of pileup. This is probably a conservative result considering that the analysis has not been optimized to the full power of the finer granularity.

11.2.2 $H \to \tau\tau$ in the VBF channel

The $H \to \tau\tau$ channel is an important benchmark for the Phase-2 upgrade, because it gives the best direct sensitivity to the Yukawa couplings between the Higgs boson and the leptons. The VBF production mode is of particular interest, since it offers a higher signal purity compared to the dominant gluon-fusion production mechanism. During the HL-LHC operation $\mathcal{O}(1\,\text{M})$ VBF $H \to \tau\tau$ events will be produced. Considering all final states of the di-tau system, this channel requires efficient identification of muons, electrons, and hadronically decaying taus (τ_h); accurate measurement of the missing transverse energy, and the reconstruction of the characteristic forward jets in VBF production. Improved jet identification in the high pseudorapidity regions of the detector will have an important effect on the selection of VBF $H \to \tau\tau$ events.

11.2.2.1 Irreducible backgrounds and di-tau mass resolution

The main irreducible background to this signal comes from the $Z \to \tau\tau$ process accompanied by the production of additional jets. The primary methods to distinguish the signal is via the reconstruction of the di-tau mass. The visible di-tau mass is calculated using only the final state charged particles, neglecting the contribution to the full di-tau mass coming from the neutrinos. A sample of simulated events in the $\mu\tau_\text{h}$ final state is used to assess this variable in the Phase-2 upgraded detector.

For this study, the selected events contain a muon with $p_\text{T} > 20$ GeV and $|\eta| < 2.1$ that is required to pass the medium muon ID and one τ_h lepton with $p_\text{T} > 20$ GeV and $|\eta| < 3.0$. The pair must be separated by $\Delta R > 0.5$. No isolation criteria are applied here; instead the reconstructed objects are required to be geometrically matched to the corresponding generator level objects. Figure 11.24 gives the pseudorapidity distribution of the τ_h lepton candidate in a simulated VBF $H \to \tau\tau$ sample. In the current Run 2 simulation no τ_h leptons are reconstructed above $|\eta| = 2.5$ due to the limited acceptance of the silicon tracker. In the Phase-2 upgrade reconstruction is extended to $|\eta| = 3$.

Figure 11.25 (left) gives the visible di-tau mass distribution for the VBF signal events where the τ_h lepton candidate is required to be within the HGCAL acceptance, $1.5 < \eta < 3.0$. The resolution in the 200 pileup Phase-2 sample is found to be 27%, similar to that in the Run 2 simulation (25%). Figure 11.25 (right) compares the visible mass distribution for the signal and $Z \to \tau\tau$ background events in the 200 pileup Phase-2 simulation. This study shows that in

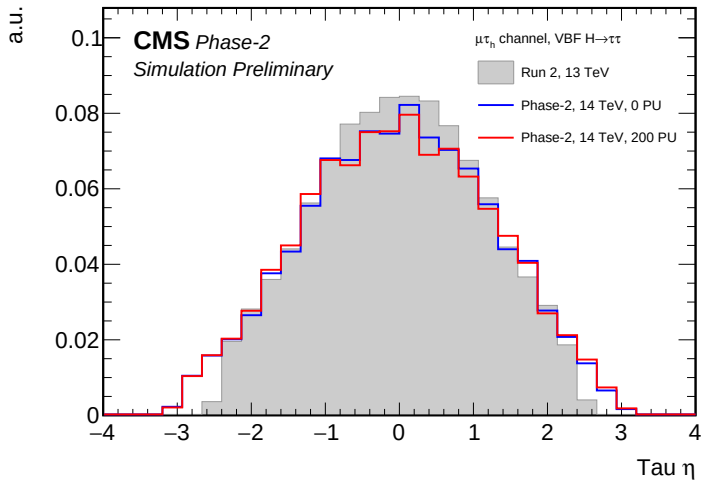


Figure 11.24: Pseudorapidity distribution of the τ_h lepton in the $\mu\tau_h$ final state of the di-tau pair, comparing VBF H $\rightarrow \tau\tau$ signal simulation for Run 2 to the upgraded Phase-2 detector with 0 and 200 pileup events.

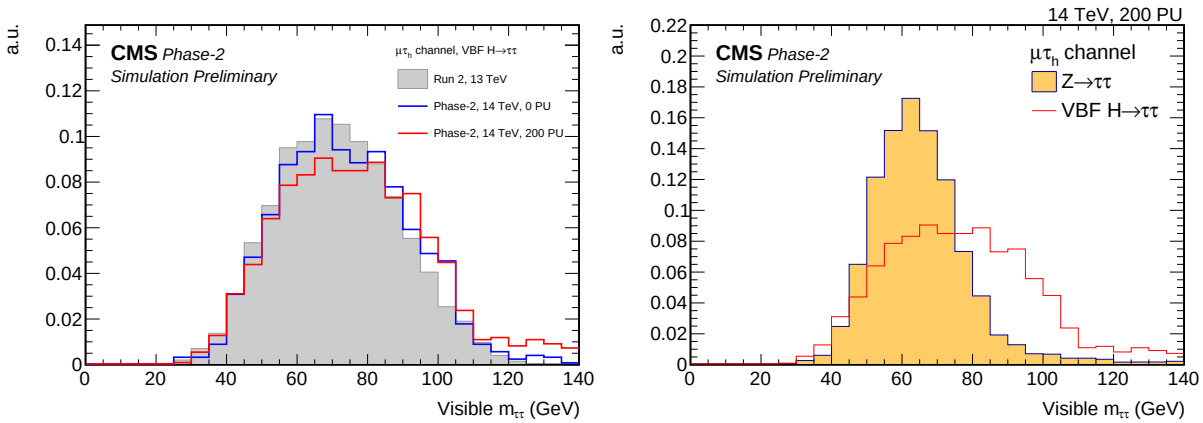


Figure 11.25: Visible di-tau mass ($m_{\tau\tau}$) distribution (left) in the $\mu\tau_h$ final state of the di-tau pair, comparing VBF H $\rightarrow \tau\tau$ signal simulation for Run 2 (PU 25) to the upgraded Phase-2 detector with 0 and 200 pileup events. Visible di-tau mass (right) for the VBF signal events compared to to the Z $\rightarrow \tau\tau$ background for the Phase-2 simulation with 200 pileup events.

the HL-LHC conditions it will be possible to distinguish signal events from this important background with performance similar to that achieved in Run 2.

11.2.2.2 Reducible background in VBF H $\rightarrow \tau_h\tau_h$

Due to the relatively high probability of a jet to be misidentified as a tau lepton and the large amount of jets produced at a proton-proton collider the H $\rightarrow \tau\tau$ in the fully hadronic channel is affected by a large QCD multi-jet background. Two key aspects to be addressed in the study of the VBF H $\rightarrow \tau_h\tau_h$ search at 200 PU are the signal reconstruction efficiency and background rate of jets in the forward region, and the misidentification of jets as tau leptons decaying hadronically.

Table 11.4: Contribution of signal and various background processes for a requirement of visible $75 < m_{\tau\tau} < 125\,\text{GeV}$ and for two different $\Delta\eta(J_1, J_2)$ requirements.

Process	$\Delta\eta(J_1, J_2) > 2.5$	$\Delta\eta(J_1, J_2) > 4.5$
SM Higgs	956 ± 30	838 ± 29
Z+jets	2376 ± 50	2376 ± 48
Drell Yan	198 ± 14	165 ± 13
$t\bar{t}$	0 ± 52	0 ± 52
QCD	11408 ± 2852	3802 ± 950
S/(S+B)	0.0068	0.17

The HGCAL guarantees a high reconstruction and tagging efficiency for the signal VBF jets, and a sustainable rate of additional background jets in the forward region. Figure 11.26 shows the η distribution of the highest and second highest p_T jets when requiring that an event has two hadronic taus with $p_T > 40$ GeV and at least two jets. We note that the signal sensitivity increases for jets in the forward region.

The VBF jet reconstruction and tagging efficiency for the VBF $H \to \tau_h\tau_h$ signal is defined as the fraction of generator-level VBF jets that are matched to a jet reconstructed using the PUPPI algorithm. The jets are selected via the β-variable defined by

$$\beta = \frac{\Sigma_{\text{tracks, } d_Z<0.1\text{ cm}} p_T}{\Sigma_{\text{tracks}} p_T}, \tag{11.1}$$

where d_Z is the longitudinal impact parameter of the considered track with respect to the primary vertex and 'tracks' are the charged particle tracks clustered in the jet.

Figure 11.27 shows the efficiency for the signal VBF jets to be reconstructed with the PUPPI algorithm and to pass the $\beta > 0.1$ selection. As a reference, the generator level VBF quark jet pseudorapidity distribution is also shown. The efficiency is found to be approximately 95% up to an $|\eta|$ of 3.0 in the 200 pileup Phase-2 simulation, indicating good performance in the region covered by the HGCAL. It is noted that at 200 pileup the efficiency is a few percent lower than in the 0 pileup case. This is caused by a non zero probability for the reconstructed primary vertex to be different from the vertex of the VBF production.

In Fig. 11.28 (left) the $\Delta\eta(J_1, J_2)$ distribution of the highest and second highest p_T jet is shown for the signal and the different background contributions. The increase of the signal component with respect to the bagkground as a function of the opening angle between the two jets, $\Delta\eta(J_1, J_2)$, is clearly visible.

Figure 11.28 (right) shows the visible di-τ mass once a large opening angle in η between the two jets, $\Delta\eta(J_1, J_2) > 4.5$, is required. This requirement enhances the sensitivity to the VBF signal as shown in Table 11.4. This is in contrast with the Run 2 analysis, which used $\Delta\eta(J_1, J_2) > 2.5$.

In Run 2 the VBF channel is the most sensitive channel for the observation of the SM $H \to \tau\tau$ [74], and it will be an important channel for the measurement of the properties of the Higgs boson in the HL-LHC era. Thus the improvement of jet identification in the high pseudorapidity regions of the detector will greatly benefit the $H \to \tau\tau$ analysis and SM Higgs boson studies. Part of possible improvements based on the jet shape observables are presented in Section 11.2.1, and they rely on the high granularity of the HGCAL and its performance.

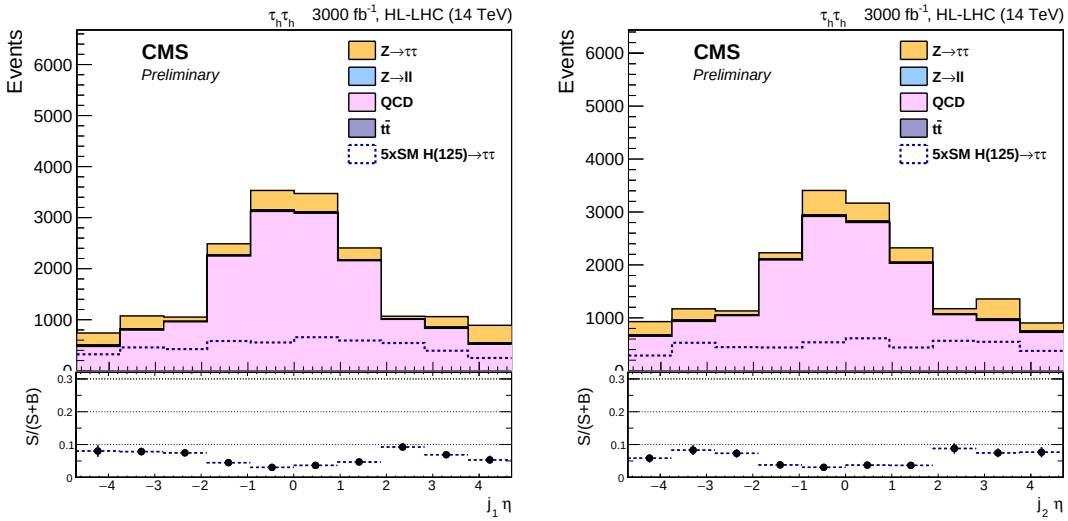


Figure 11.26: Pseudorapidity distribution of (left) the highest p_T jet and (right) second highest p_T jet after requiring that an event has two τ_h with $p_T > 40$ GeV and at least two jets.

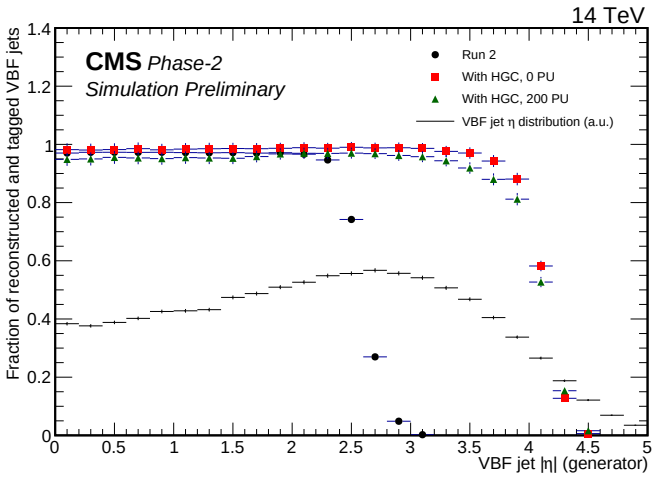


Figure 11.27: VBF jet reconstruction and tagging efficiency comparing Run 2 (pileup 25) simulation to the upgraded Phase-2 detector with 0 and 200 pileup events.

11.2.3 HH→ bb$\tau\tau$

The double Higgs boson production in the bb$\tau\tau$ decay channel has been studied in the context of the HL-LHC conditions taking into account the Phase-2 CMS experiment upgrades and in particular the HGCAL detector. Among the channels with a relatively high sensitivity to SM double Higgs production, this final state has been chosen because it is one that profits most from the introduction of the HGCAL detector, both in acceptance and better reconstruction of jets and tau leptons.

As a first exercise, a projection of the current experimental results, documented in Ref. [75], has been performed by extrapolating the distributions of expected signal and background events and the corresponding uncertainties to the HL-LHC luminosity conditions. The performance of the CMS detector is assumed to be unchanged with respect to that observed in 2016. The

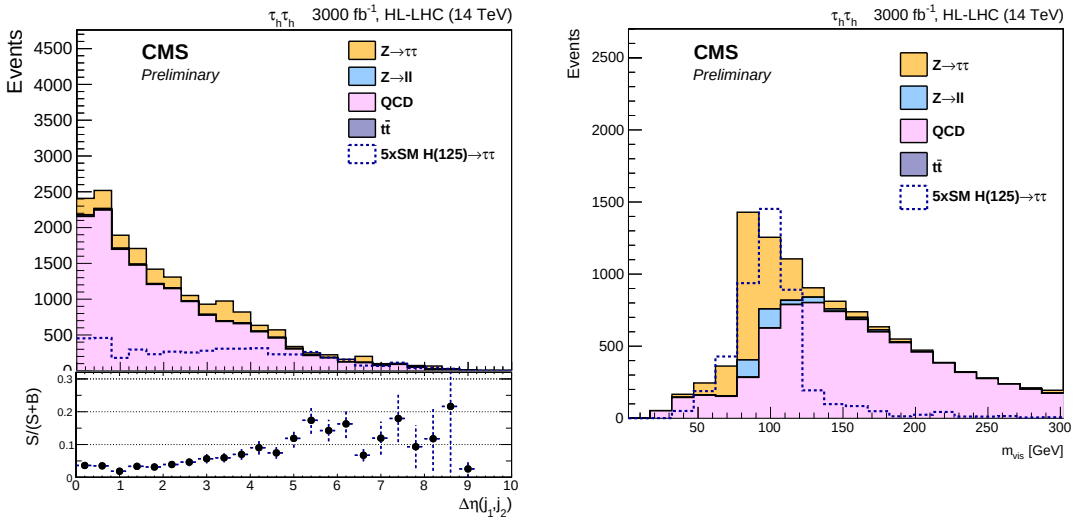


Figure 11.28: (left) Distribution of the opening angle in η, $\Delta\eta(J_1, J_2)$, for the highest and second highest p_T jet, and (right) the visible $m_{\tau\tau}$ for events with $\Delta\eta(J_1, J_2) > 4.5$.

systematic uncertainties are considered to be reduced with respect to the current values due to the larger amount of data expected to be available to evaluate them more precisely. In particular, the trigger and object identification efficiency uncertainties are considered to be 1% for muons and electrons, 2% for tau leptons, and 2% for b-jet identification. The jet and the tau lepton energy scales are assumed to be known with a precision of 1% and 1.5% respectively. The significance of a SM HH signal is expected to range between 1.1 and 1.3σ, depending on the assumptions made on the background rejection. The corresponding 95% confidence level upper limit on the HH SM production cross section (σ_{HH}^{SM}) ranges between 1.9 and 1.7 $\times$ σ_{HH}^{SM}. These results constitute the reference to which the sensitivity to HH production obtained with the analysis described below should be compared.

The analysis described in the following makes use of dedicated simulated samples, to incorporate the collision conditions of the HL-LHC and the performance of Phase-2 upgraded detector including the HGCAL. These samples allow the verification of the assumptions made concerning the performance of the future HL-LHC CMS detector, and the study of the improved capabilities of the CMS experiment following the installation of the HGCAL detector. The study of HH VBF production is also explored to evaluate the advantages provided by the upgraded HGCAL detector information.

Events selected in the Phase-2 analysis are required to have at least two (medium WP) b-tagged jets with $p_T > 30$ GeV, and two hadronically decayed tau leptons with $p_T > 45$ GeV. A veto on the presence of a third isolated lepton (either a muon or an electron) is applied to reduce the contributions from misidentified leptons and QCD background. A misidentification probability for electron(jets) to tau leptons of 4%(1.26%), obtained from dedicated studies shown in Section 11.2.2, is applied to the corresponding Phase-2 objects.

A dedicated category targeting VBF HH production has been introduced that considers events satisfying the criteria above and also have an additional pair of jets with an invariant mass $m_{jj} > 250$ GeV and $|\Delta\eta(j, j)| > 2.5$. To increase signal statistics in this category, the b tagging requirement on the two b jets is relaxed to the loose working point.

$H \to \tau_h\tau_h$ ($H_{\tau\tau}$) candidates are reconstructed by combining all possible opposite-sign tau lepton pairs present in each event. In each pair, the tau with the highest p_T is defined to be the first.

If more then one pair is present in the event, the chosen candidate is the one where the first tau is the most isolated according to the sum of p_T of the charged particle in a cone of 0.5 around it. Pairs are formed only from taus with isolation less than 2.5 GeV. The SVFit [76] algorithm is applied to selected di-tau pairs in order to reconstruct the most likely $H \to \tau\tau$ invariant mass.

$H \to bb$ (H_{bb}) candidates are reconstructed by selecting the 2 jets with the highest b tagging score (cMVAv2, see Section 11.1.4) in the event.

In order to further reduce the background, both the $H_{\tau\tau}$ and H_{bb} candidates are required to have a mass close to that of the Higgs boson.

The trigger efficiency is assumed to be 100% for events passing the selection.

In this analysis we consider the main background contributions to be those already considered for the similar Run 2 results. The $t\bar{t}$ and DY backgrounds, as well as the gluon fusion and VBF HH signals, are estimated from full MC simulations generated with realistic Phase-2 conditions. The QCD background is estimated from Delphes simulations, while subdominant single Higgs contributions, mostly coming from ZH associated production, are obtained via a simple rescaling of the background from the Run 2 analysis for the higher luminosity and 14 TeV cross-section. Other even smaller background contributions such as diboson production are neglected in this study.

Assumptions made for the systematic uncertainties affecting this result include:

- The luminosity uncertainty assumed to be 1.5% at the HL-LHC.
- A 1% systematic uncertainty applied to each b jet to account for b tagging algorithm inefficiencies.
- A 1% uncertainty assigned to each jet to account for the jet energy scale uncertainty.
- The reconstruction efficiency uncertainty for each τ lepton assumed to be the same as in Run 2 (5%).
- Theoretical uncertainties in PDF and QCD scale for the signal, as described in Ref. [77], scaled by 1/2 to take account of anticipated improvements in the determination of the signal cross-section.
- A further 1% uncertainty applied as QCD scale uncertainty on the fully simulated samples (DY and $t\bar{t}$).

The "stransverse mass" M_{T2}, defined in Ref. [78], is used as discriminating variable. The M_{T2} variable is defined as the largest mass of the parent particle that is compatible with the kinematic constraints of the event. In the case of the $bb\tau\tau$ decay, where the dominant background is $t\bar{t}$ production, the parent particle is interpreted as the top quark that decays into a bottom quark and a W boson.

Figure 11.29 shows the expected distribution of the 'gluon fusion' and VBF signals, as well as the main background sources, as a function of the stransverse mass M_{T2}. Given the limited statistics available, a simple cut and count experiment is used in place of a full shape analysis in the VBF category.

The yields for the signals and background processes in the different categories are reported in Table 11.5, while the final sensitivities to the SM production are summarized in Table 11.6

The larger acceptance of the Phase-2 detector accounts for a 10–15%(4–8%) increase for the VBF(gluon fusion) process in the signal region. A final sensitivity close to 1.5 times the SM cross section value is expected for the SM double Higgs boson production in the $bb\tau\tau$ channel.

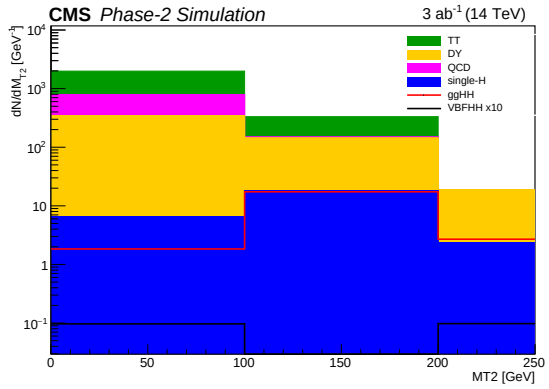


Figure 11.29: Stransverse mass distribution for $\tau_h\tau_h$ events in the two b-tagged jets category.

Table 11.5: Yields of different processes after the final selection.

Category	HH(g-fusion)	HH(VBF)	$t\bar{t}$	DY	QCD	single-H
2 b jets	22.4	0.41	1366	478	455	28
VBF	3.44	0.18	30.3	-	1.0	-

This is in good agreement with previous projections obtained by rescaling lower energies and lower statistics results. Further improvements on these results can be obtained by including the category with one b-tagged jet and the semileptonic final states, which can potentially bring this close to the SM sensitivity.

11.2.4 Search for electroweakinos in the final states with two same-sign leptons

Supersymmetry is considered one of the most compelling theories of physics beyond the SM. However, large regions of parameter space characterized by the production of strongly interacting sparticles with R-parity conserving decays have been excluded at 95% CL. On the contrary, due to its low production cross section, the exploration of electroweak production of SUSY particles has just started at the LHC. The HL-LHC data, with an integrated luminosity of 3000 fb^{-1}, will offer an unprecedented discovery potential for SUSY through searches for electroweakinos.

In most of the SUSY breaking scenarios, the supersymmetric partners of the gauge and Higgs bosons are expected to be lighter than a few hundreds of GeV based on naturalness and unification arguments. The higgsino (μ), bino (M1), and wino (M2) mass parameters typically satisfy the relation $\mu <$ M1 $<$ M2. As a result, the mass spectra are characterized by low-mass higginos-like $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$, heavier bino-like $\widetilde{\chi}_3^0$ along with mass-degenerate wino-like $\widetilde{\chi}_2^\pm$, $\widetilde{\chi}_4^0$. Assuming minimal mixing, the mass of the $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$ states is close to μ, that of $\widetilde{\chi}_3^0$ to M1, while the $\widetilde{\chi}_2^\pm$, $\widetilde{\chi}_4^0$ mass is close to M2.

Due to the higgsino nature of the $\widetilde{\chi}_1^\pm$, $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^0$, the mass difference between the low mass states and the $\widetilde{\chi}_1^0$ is just a few GeV leading to signatures with very low p_T SM particles and thus to a very challenging experimental search. The mass of the $\widetilde{\chi}_2^\pm$, $\widetilde{\chi}_4^0$ is instead expected

Table 11.6: Final expected limit, presented as ratio over the SM double Higgs production.

Category	95%CL on $\sigma_{HH}/\sigma_\text{SM}$	95%CL on $\sigma_{ggHH}/\sigma_\text{SM}$	95%CL on $\sigma_{VBF}/\sigma_\text{SM}$
2b0j	1.8	1.9	71.7
VBF	3.9	4.2	85.5
Combined	1.6	1.7	51.1

to be several hundreds GeV. In radiatively-driven natural supersymmetry (RNS) [79, 80] the $\widetilde{\chi}_2^{\pm}$ and $\widetilde{\chi}_4^0$ are then expected to decay into higgsinos emitting same charge W bosons with a total branching ratio close to 25% (Fig. 11.30). The search for the pair production of $\widetilde{\chi}_2^{\pm}$, $\widetilde{\chi}_4^0$ in the final states with two same charge leptons, large $p_{\mathrm{T}}^{\mathrm{miss}}$ and modest jet activity is a novel analysis representing one of the most promising probes for SUSY at the HL-LHC. The forward calorimeter is a critical subdetector for this analysis as optimal $p_{\mathrm{T}}^{\mathrm{miss}}$ and jet reconstruction performance is essential in discriminating signal from background. The details of the analysis are provided below.

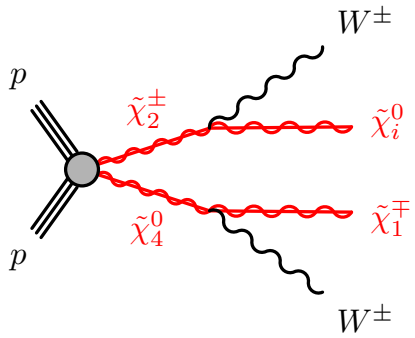


Figure 11.30: Diagram for the wino-like $\widetilde{\chi}_2^{\pm}$ $\widetilde{\chi}_4^0$ pair-production and decay into a final state with two same charge W bosons.

A Delphes simulation [81] of the CMS Phase-2 detector is used to estimate the signal and background yields. The signal Monte Carlo samples are generated by MADGRAPH 5 @NLO (v2.3.3) up to two additional jets at leading order precision. The supersymmetric particles are then decayed by the PYTHIA 8.2 package that provides also showering and hadronization. The production cross sections have been calculated for $\sqrt{s} = 14$ TeV at NLO-NLL using the resumming code from Ref. [82, 83]. The final values are calculated using the 4LHC recommendations for the two sets of cross sections following the prescriptions of the Cross Section Working Group [84]. MADGRAPH 5 is also used to produce the parton-level background processes at LO, followed by parton showering and hadronization with PYTHIA 6.

The candidate signal events contain two high quality and isolated leptons with $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 1.6$, and same charge. Discrimination from the background processes is achieved by selecting events with no additional leptons with $p_{\mathrm{T}} > 5$ GeV and $|\eta| < 4.0$ (to suppress multiboson production), no $p_{\mathrm{T}} > 30$ GeV b jets (to suppress events with top quarks), and no high-p_{T} jets. The remaining background processes include the pair production of W and Z/γ^* bosons. These are suppressed by imposing a tight selection on the $m_{\mathrm{T,min}}$defined from the missing p_{T} as

$$m_{\mathrm{T,min}} = \min[m_{\mathrm{T}}(\mathrm{lep}_1, p_{\mathrm{T}}^{\mathrm{miss}}), m_{\mathrm{T}}(\mathrm{lep}_2, p_{\mathrm{T}}^{\mathrm{miss}})]. \tag{11.2}$$

Figure 11.31 shows the distribution of the $m_{\mathrm{T,min}}$observable in events satisfying the baseline signal region selection described above.

To maximize the sensitivity seven signal regions have been optimized with $m_{\mathrm{T,min}}$in the ranges [0, 90], [90, 120], [120, 150], [150, 200], [200, 250], [250; 300], and [300, inf] GeV.

The Delphes based yields of processes containing prompt leptons are corrected by the lepton reconstruction, identification and isolation efficiencies measured in Run 2 collision data. These

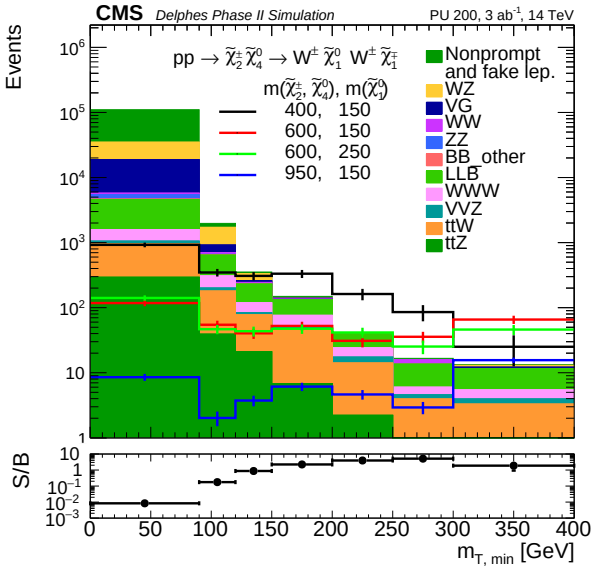


Figure 11.31: Distribution of the $m_{\mathrm{T,min}}$in candidate events satisfying the baseline signal region selection. In the case of the signal the first number refers to the $\widetilde{\chi}_2^{\pm}$, $\widetilde{\chi}_4^0$ mass while the second to the value of the $\widetilde{\chi}_1^0$ mass.

efficiencies are typically lower than those obtained from fully simulated Phase-2 data as they were optimized to significantly suppress the misidentification probabilities. For example, the reconstruction efficiency for centrally produced electrons ranges from 60 to 86% for p_{T} values between 20 and 200 GeV.

Background processes containing misidentified leptons are determined from Delphes as well. However, only non-prompt leptons from heavy flavor decays are included in Delphes. For the misidentified leptons from light flavor quarks, gluons, conversions, the yields from Delphes have to be increased by 25% [85].

The search sensitivity is then calculated using a modified frequentist approach with the $\mathrm{CL_S}$ criterion and asymptotic results for the test statistic [86, 87]. The systematic uncertainty in the prompt (misidentified, signal) yields are assumed to be 20% (50%, 20%) based on the estimates computed in the corresponding search carried out in Run 2 collision data [85].

The upper limit on the production cross-section of pair produced $\widetilde{\chi}_2^{\pm}$,$\widetilde{\chi}_4^0$ decaying into a final state with two same charge W boson with a branching ratio of 25% is shown in Fig. 11.32 for the two μ scenarios (where $\mu \sim m_{\widetilde{\chi}_1^{\pm}}, m_{\widetilde{\chi}_2^0}, m_{\widetilde{\chi}_1^0}$). The value $\widetilde{\chi}_1^0$ =150 GeV is representative of the region of the parameter space outside the reach of the Run 2 search for the direct production of higgsinos in the final states with two same flavor opposite sign leptons [88], while $\widetilde{\chi}_1^0$ =250 GeV is outside the sensitivity reach of the same search when extrapolated to the HL-LHC. As expected the sensitivity depends on the value of $\widetilde{\chi}_1^0$ mildly at large $\widetilde{\chi}_2^{\pm}$, $\widetilde{\chi}_4^0$ mass values, while the dependence is more significant when $\widetilde{\chi}_2^{\pm}$, $\widetilde{\chi}_4^0$ mass approaches $\widetilde{\chi}_1^0$.

Wino-like mass degenerate $\widetilde{\chi}_2^{\pm}$,$\widetilde{\chi}_4^0$ are excluded at 95% CL for masses up to 900 GeV in the $\mu =$ 150 and 250 GeV scenarios. This demonstrates that the HL-LHC has the potential to probe most of the natural SUSY parameter space with electroweak naturalness measure $\Delta\mathrm{EW} < 30$ [89].

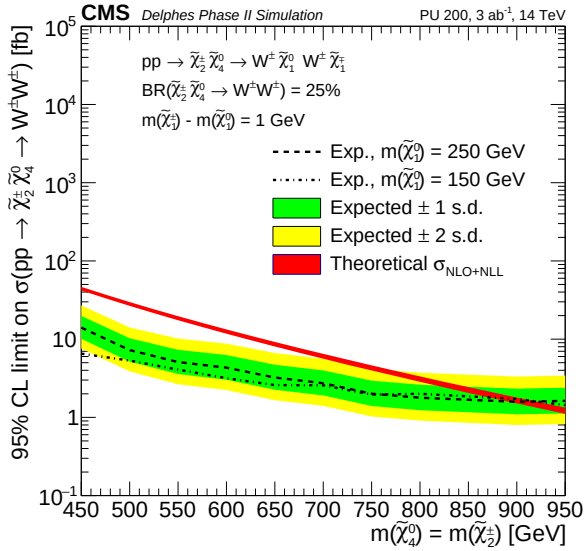


Figure 11.32: Upper limit on the production cross-section of pair produced $\widetilde{\chi}_2^{\pm}$,$\widetilde{\chi}_4^0$ decaying into a final state with two same charge W boson with a branching ratio of 25% for two assumptions on the $\widetilde{\chi}_1^0$ mass.

11.2.5 Search for FCNC in t→qγ events

In the SM, flavor-changing neutral currents occur only at loop level and are highly suppressed by the Glashow-Iliopoulos-Maiani (GIM) mechanism [90]. In the top sector, the predicted branching fractions for the t→u+γ and t→c+γ decays are approximately 10^{-16} and 10^{-14}, respectively [91]. However, in many extensions of the SM, such as supersymmetry or multiple Higgs boson doublet models, the GIM suppression can be relaxed leading to an enhancement of several orders of magnitude in branching fractions that could be observed at the HL-LHC [92, 93].

The most stringent constraints on the B(t→q+γ) are set by the CMS experiment through single-top quark production in association with a photon. The 95% CL upper limit on the branching fractions are B(t→u+γ) $< 0.016\%$ and B(t→c+γ) $< 0.182\%$ [94].

In this section, the sensitivity of the upgraded CMS detector to tqγ FCNC transitions is estimated for integrated luminosities of 300 and 3000 fb^{-1} using single top quark production via q→tγ, with q being a u or c quark. This analysis focuses on subsequent SM decays of the top quark, i.e. top quark decays into a W boson and a bottom quark, followed by the decay of the W boson to a muon or electron and a neutrino.

The final-state signature of these events is the presence of a single muon or electron, large missing transverse momentum, a b jet, and an isolated high energetic photon, with a broad η spectrum. The photon properties themselves provide good separation with respect to the dominant background processes from W+jets, and single-top or top quark pair production in association with photons.

The signal is generated with MADGRAPH 5, while the background processes are generated with POWHEG (t$\bar{\text{t}}$, single top), MADGRAPH (W+jets), and aMCatNLO (single top + photon). For parton showering and hadronization the matrix-element calculations are interfaced with PYTHIA 8. Except for the associated production of a single top quark with one or two addi-

tional photons, all samples employ a full simulation of the upgraded CMS detector. For the latter sample, the detector response is simulated using Delphes [81]. For illustrative purposes, the signal samples are normalized to a 1 pb production cross section. For the background processes, next-to-leading order or next-to-next-to-leading order cross section predictions are used where available.

Events are selected requiring the presence of exactly one muon or electron that passes high purity identification requirements and has $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.8$. The electron identification corresponds to the tight working point described in Section 11.1.1. Electrons in the overlap region $1.4 < |\eta| < 1.6$ are removed. The event is rejected if it has additional low purity (loose) muon or electron candidates with $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.8$.

Jets are reconstructed using the anti-k_{T} algorithm with a distance parameter R = 0.4, $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.8$. Pileup effects are mitigated using the PUPPI algorithm. The event is required to have exactly one jet that passes b tagging criteria, based on the DeepCSV algorithm for jets with $|\eta| > 1.5$ and the cMVA algorithm for jets with $|\eta| < 1.5$. [65]. The medium working point is used, corresponding to a b tagging efficiency of about 70% and a misidentification probability of 18% for c jets, and 1.5% for other jets. Events are selected if they contain either two or three jets, with exactly one being b tagged. Photon candidates are required to be well-identified using the tight working point described in Section 11.1.2 and to have $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 2.8$. Photons in the overlap region $1.4 < |\eta| < 1.6$ are removed. All objects are required to be well separated with $\Delta R(e/\mu, \gamma) > 0.7$ and $\Delta R(b, \gamma) > 0.7$. The top quark kinematics are reconstructed and only events with a reconstructed top quark mass in the range of 130 to 220 GeV are selected.

For the discrimination of signal and background events, and to set the limits on FCNC couplings, the events are split into two categories: events with a highly energetic central photon ($|\eta| < 1.4$), and those with a forward photon ($1.6 < |\eta| < 2.8$). In the central (forward) region, the photon p_{T} (energy) is used as as a discriminating distribution; the low p_{T} (energy) region is background dominated, while high p_{T} (energy) region is populated by signal events. The distributions are shown in Fig. 11.33.

The limits on the cross section for single top quark production via tqγ interactions are obtained using the asymptotic CL_S method. Systematic uncertainties from variations of the renormalization and factorization scales, b tagging and jet energy scale corrections, are included as well as effects from b tagging, jet energy scale corrections, which are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$, lepton efficiencies, and luminosity. The normalization of the $\mathrm{t\bar{t}}$ contribution is varied by 6%; the single top + photon background receives a normalization uncertainty of 30%, while the cross section of the remaining background sources is varied through their scale uncertainties at leading or next-to-leading order, respectively. The upper limit on the single top quark production cross section via tuγ (tcγ) interaction of 3.4 (4.4) fb at 95% CL is obtained for an integrated luminosity of $3000\,\mathrm{fb}^{-1}$. This translates to 95% CL upper limits on the branching fractions of $\mathrm{B}(\mathrm{t}\to\mathrm{u}+\gamma) < 8.6 \times 10^{-6}$ and $\mathrm{B}(\mathrm{t}\to\mathrm{c}+\gamma) < 7.4 \times 10^{-5}$, improving over the recent extrapolation to HL-LHC conditions by a factor of 3 [95]. For an integrated luminosity of $300\,\mathrm{fb}^{-1}$, the limits are 8.3 and 9.3 fb for tuγ and tcγ, respectively.

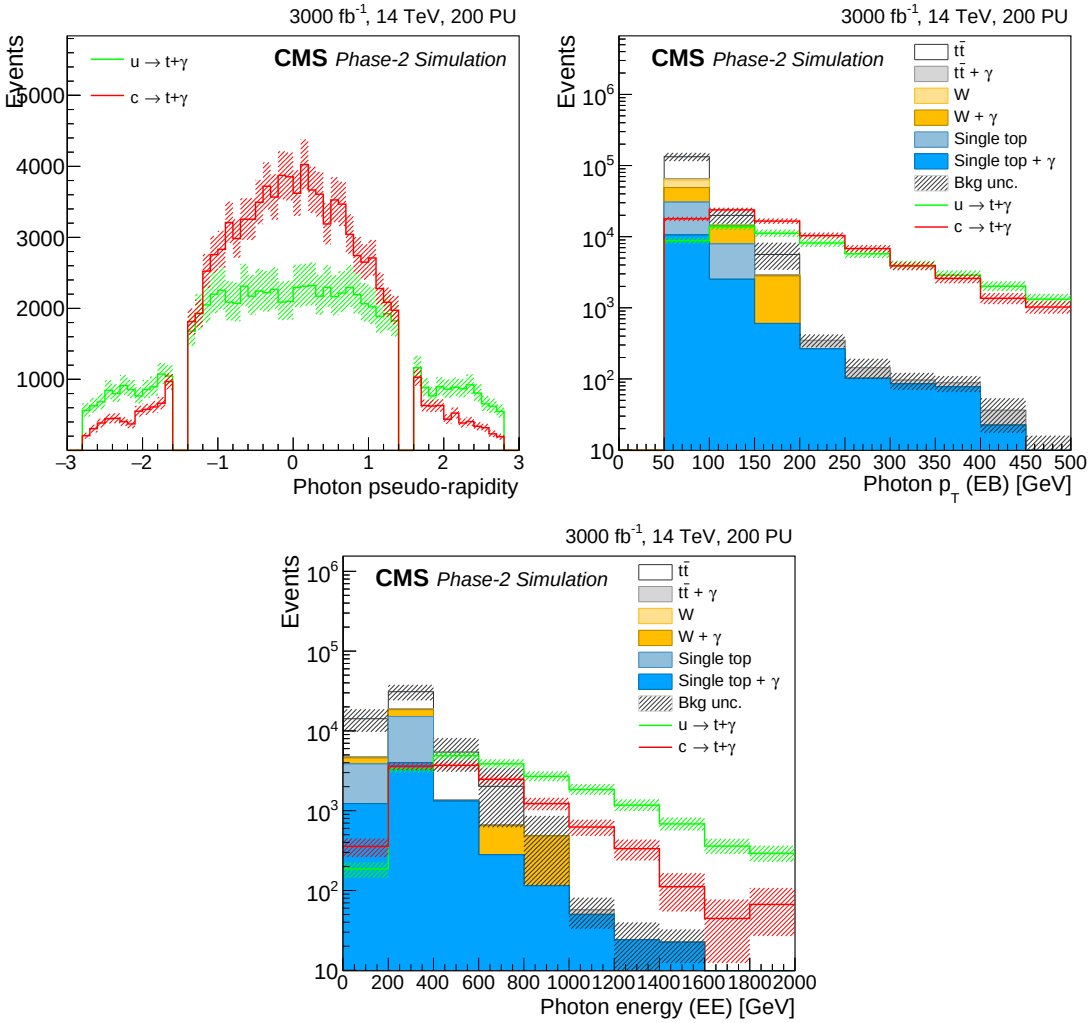


Figure 11.33: Pseudorapidity distribution of the FCNC signal after the full event selection (upper left), transverse momentum of photon candidates for the central η region (upper right), and energy of photon candidates in the forward region (bottom).

Chapter 12

Radiation environment

12.1 Radiation environment and activation

12.1.1 LHC parameters, geometry and simulation framework

Simulations are used to predict the magnitude and composition of radiation as a function of luminosity. The radiation simulations are performed by the CMS BRIL project using the FLUKA 2011.2b.6 Monte Carlo multi-particle transport code [96, 97]. The event generator DPMJET III [98] is used to create the primary proton-proton events. It is directly linked to the FLUKA code and used as the default event generator for high energy hadron-nucleus and nucleus-nucleus interactions. All particles are transported until a predefined energy cut-off is reached. In this case, their remaining energy is deposited on the spot. All hadrons and muons have an energy cut-off of 1 keV, except neutrons, which are transported down to an energy of 0.01 meV. Electrons and photons are assigned different cutoffs in different materials. In most materials 3 keV for photons and 30 keV for electrons are applied. The output is usually averaged over all simulated primary events and normalized per primary event. The normalization used for the prediction of dose and fluence depends on the inelastic collision rate and the duration of LHC operation (total radiation period). For an instantaneous luminosity of $5\times10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, and an inelastic cross section of ~ 80 mb (as predicted by the EPOS LHC generator [99] for $\sqrt{s} = 14$ TeV protons), an average of 4×10^{9} inelastic pp events per second are produced.

A concept that is used in many of the FLUKA user-defined outputs, and also in this document, is the so-called track length estimation, where the path length of a particle, passing through a specified volume, is recorded. The track length is normalized to the detector volume, resulting in a fluence value per unit area. The FLUKA USRBIN scoring is used in this report, where the volume of interest is overlaid with a binning mesh that can be Cartesian, in cylindrical coordinates, or as a special case bound to a geometry region. This scoring is used to obtain particle fluences by calculating the track length density, or doses using the energy deposition. The obtained 3D distributions can be projected to two dimensions and shown as a 2D flux map.

Scorings are configured for different particle types or groups of particle types (e.g. charged hadrons, all neutral hadrons, or neutrons or electrons). Using the energy deposited in the material, quantities like energy deposition and absorbed dose are determined. Using the track length scoring technique, with the use of additional predefined functions where necessary, quantities such as 1 MeV neutron equivalent fluence in silicon are determined.

The FLUKA CMS geometry version 3.7.9.1 is used to perform simulations for HL-LHC conditions. Compared to version 3.7.2.0 used in the TP [1], version 3.7.9.1 includes an improved description of the HGCAL longitudinal structure, close to the structure described in this TDR. For example, the lead absorbers, with stainless steel (SS) cladding, have been introduced in the geometry for CE-E, and SS has been substituted for brass for the CE-H absorber material. The

model also includes an improved description of the Al conical beam pipe in the Endcap region, and a common thermal screen for CE-E and CE-H. It should be noted that the radiation fields obtained with version 3.7.9.1 are very close (within a few percent) to those of 3.7.2.0, with the exception of a few-centimetres zone at the inner radius, where the dose actually decreases with the new beam pipe design. This shows that the simulation results do not depend strongly on details of the HGCAL geometry.

The complexity of the FLUKA simulation geometry is limited by the number of objects or elements that are reasonably used to limit the computing time. Hence the geometry model is a simplification of the reality, composed of the objects and elements and average material within them. The correct material composition and the material densities are used, resulting in the correct total mass. The statistical uncertainties are given by the error of the mean values, which are obtained by averaging over many simulation cycles; they are below 2.5% for the 1 MeV neutron equivalent fluence in silicon and below 8% for the total ionisation dose (TID). There are larger contributions to the systematic uncertainties for all results: uncertainties due to the event generator, and uncertainties due to the simplification of the CMS FLUKA model in terms of volumes and materials.

All the results of the particle fluences, energy dose, and 1 MeV neutron equivalent in silicon are scored in an r-z grid. A perfect azimuthal symmetry is assumed. For the HGCAL, the scoring cell sizes are 1 cm in r and 1 cm in z.

12.1.2 Radiation doses and fluences in the calorimeter

The integrated fluence and TID for an integrated luminosity of 3000 fb^{-1} over the whole volume of the HGCAL have been presented in the Section 1.2 (see Figs. 1.1 and 1.2). We present below a few more plots corresponding to specific parts of the calorimeter.

Maps of the TID and of the integrated fluence in CE-E are shown in Figs. 12.1 and 12.2, respectively. The maximum TID, observed at the electromagnetic shower maximum ($z \approx 325$ cm) is 1.39 MGy, the maximum fluence is $6.9\times10^{15}\,\text{n}_{\text{eq}}/\text{cm}^2$. One notes the very fast decrease with radius, as shown by the 1D projection at $z = 325$ cm plotted in the same figures.

Maps of the TID and of the integrated fluence in CE-H are shown in Figs. 12.3 and 12.4, respectively. 1D projections as a function of radius are also shown for the first and last layer including scintillators (CE-H 9 and CE-H 24). However, when scoring with a finer pitch, one observes that the dose deposited in the scintillator is larger than the average plotted in Fig. 12.3 by a factor $\approx$1.8, almost independent of the radius, as shown in Fig. 12.5. This effect, which is probably linked to the high hydrogen content of the scintillator and is observed neither in the silicon layers nor for the fluence results, has been taken into account in defining the limit between silicon and scintillator in CE-H.

12.1.3 Activation

Estimates of activation are important to determine the possibility of maintenance for the detector. Activation results are however very sensitive to the exact choice of materials and produced radionuclides, and accurate estimates will be obtained at a later stage. Nevertheless, the activation of HGCAL with the present knowledge was studied, in particular to guide the choice of the hadronic calorimeter absorber material. It was found that there is not much difference between standard 304L stainless steel (standard material with 0.1% cobalt by weight) than brass (the absorber material used in the present CMS HCAL), which led to the choice of the former because it is cheaper and mechanically stronger. Simulations were performed with FLUKA. Due to the cool-down during the yearly technical stops, the results depend on the assumptions

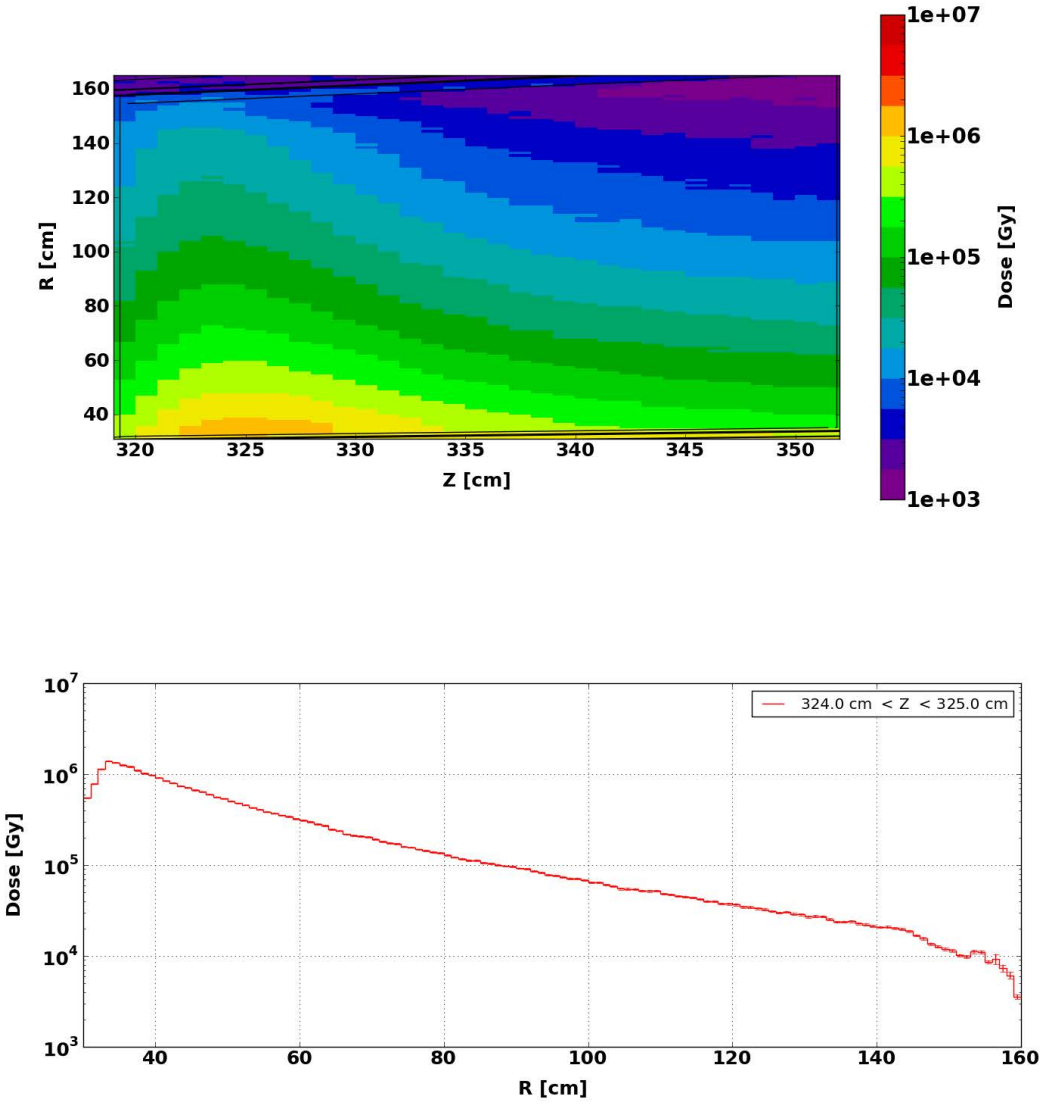


Figure 12.1: Total ionizing dose in Gy for CE-E as predicted by the CMS FLUKA simulation 3.7.9.1 for an integrated luminosity of 3000 fb^{-1}. The bottom plot shows a projection as a function of the radius for a fixed value of the z coordinate, corresponding approximately to the electromagnetic shower maximum.

made on the irradiation cycle, and a conservative model, approved by the CERN radio protection group, was used for the integral and the time profile of the luminosity every year until the end of HL-LHC exploitation. Figure 12.6 shows the ambient dose rate in μSv/h for a typical stop towards the end of the exploitation period, after four weeks of cool-down. It is clear that, for the lower corner of HGCAL, the dose rate in excess of 1 mSv/h will not allow extended access. On the other hand, the outer periphery with a few μSv/h will remain accessible for maintenance during technical stops or shutdowns. Waiting for a full year cool-down reduces the dose rate by a further factor $\approx$5.

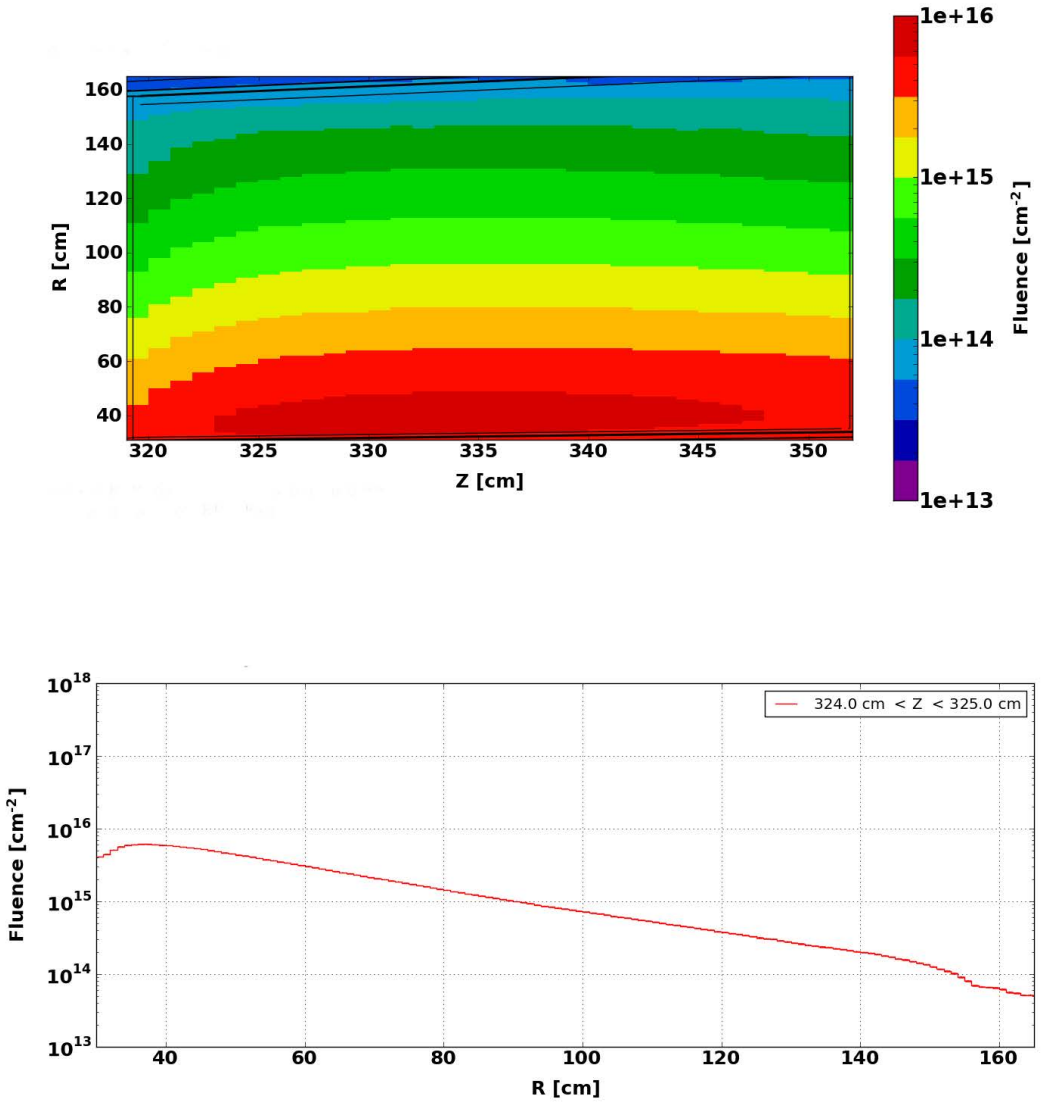


Figure 12.2: Integrated particle fluence in 1 MeV neutron equivalent in silicon per cm^2 for CE-E as predicted by the CMS FLUKA simulation 3.7.9.1 for an integrated luminosity of 3000 fb^{-1}. The bottom plot shows a projection as a function of the radius for a fixed value of the z coordinate, corresponding approximately to the electromagnetic showers maximum.

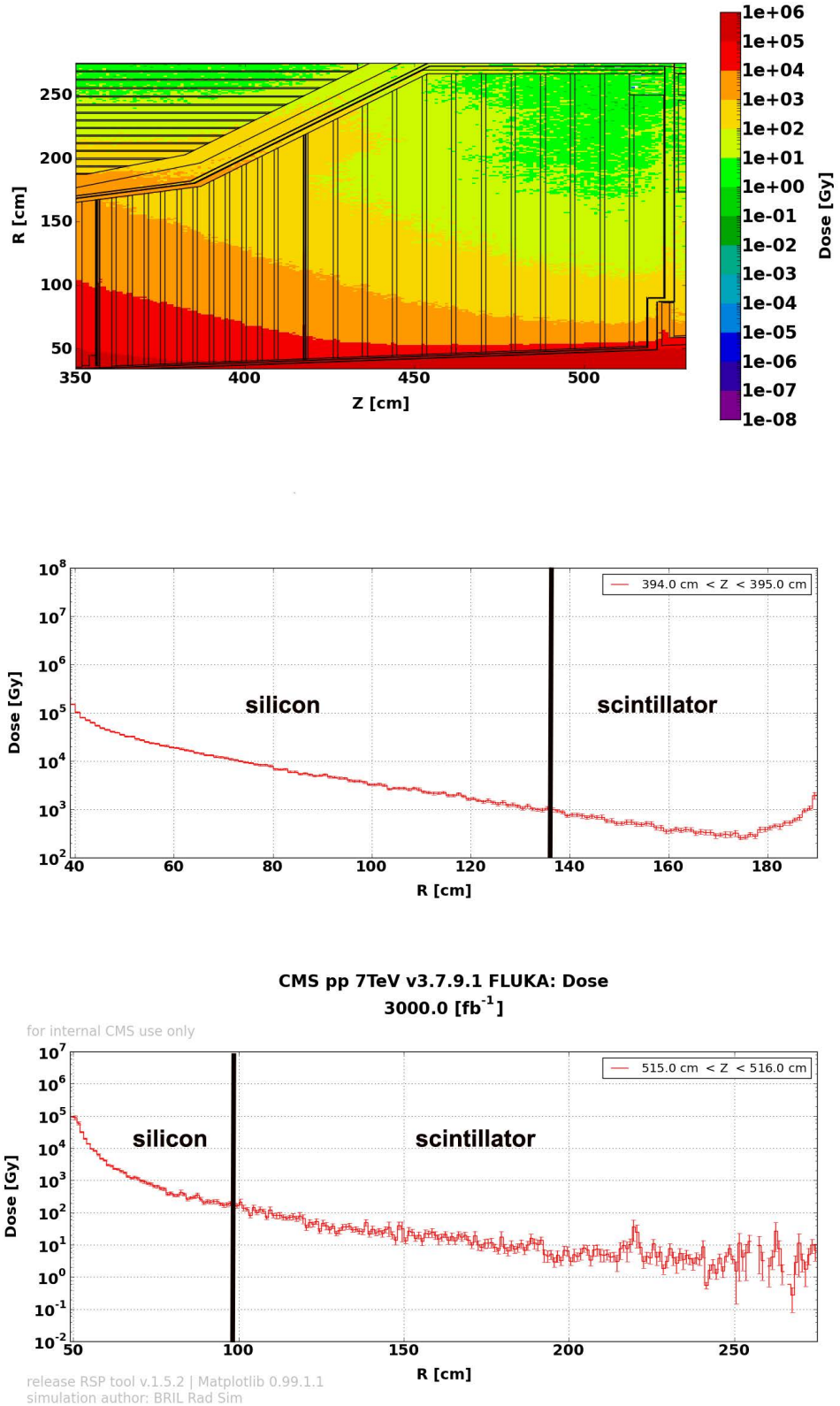


Figure 12.3: Total ionizing dose in Gy for CE-H as predicted by the CMS FLUKA simulation 3.7.9.1 for an integrated luminosity of 3000 fb^{-1}. The bottom plots show a projection as a function of the radius for the first and last layer including scintillators. The limit between the regions equipped with silicon sensors and with scintillator tiles is indicated by the vertical line.

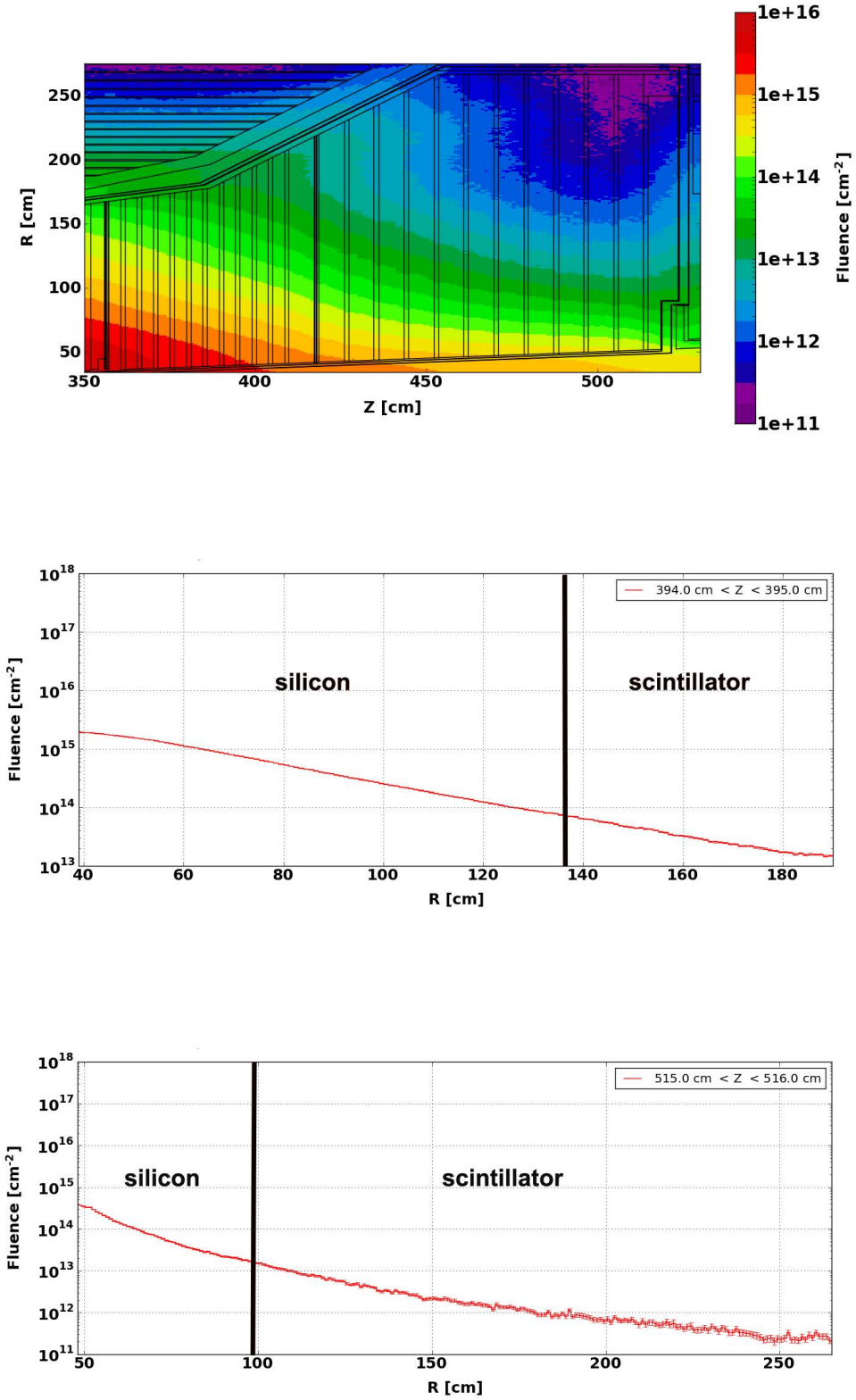


Figure 12.4: Integrated particle fluence in 1 MeV neutron equivalent in silicon per cm^2 for CE-H as predicted by the CMS FLUKA simulation 3.7.9.1 for an integrated luminosity of 3000 fb^{-1}. The bottom plots show a projection as a function of the radius for the first and last layer including scintillators. The limit between the regions equipped with silicon sensors and with scintillator tiles is indicated by the vertical line.

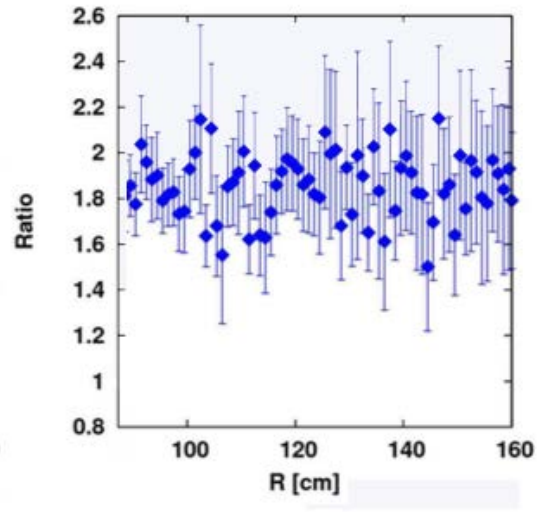


Figure 12.5: Ratio between the actual dose and the average dose obtained with a 1 cm grid scoring in a 3 mm thick scintillator located in the middle of CE-H. The ratio is plotted as a function of the radius.

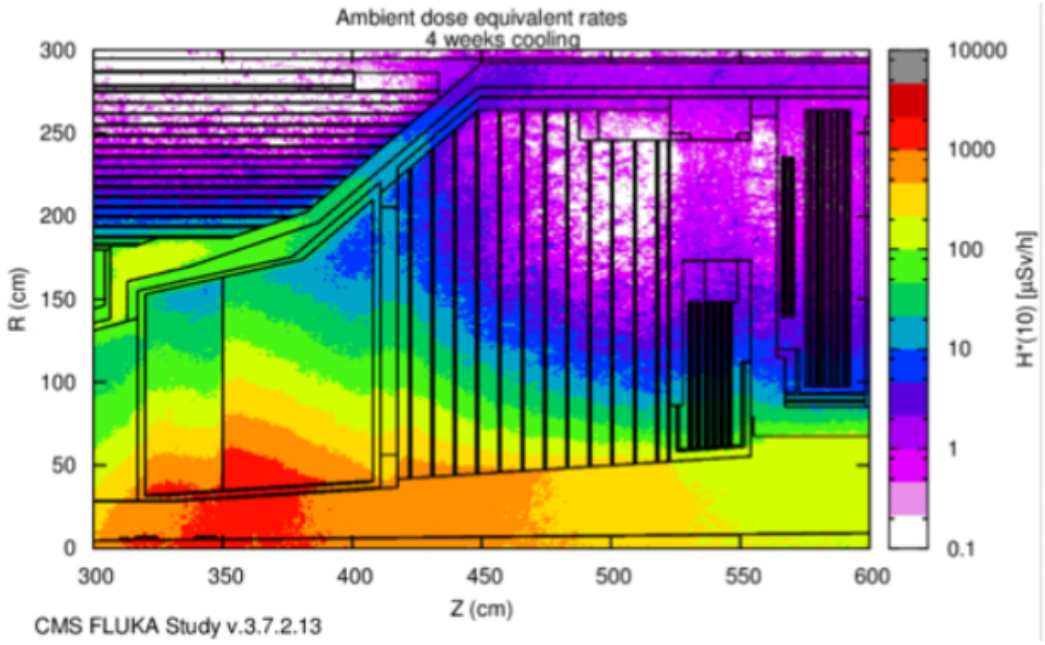


Figure 12.6: Activation map in μSv/h of the HGCAL at the end of exploitation after 4 weeks of cool-down.

Chapter 13

Quality control and assurance, and workflow

The quality control of the various HGCAL elements (sensors, modules, scintillator assemblies, cassettes) has been presented in their respective sections. In this chapter, we present the general workflows for the production and the strategy for quality assurance of the project.

13.1 Production workflows

The workflow for the production of silicon modules is shown in Fig. 13.1. Individual elements such as the WCu base plates, the Kapton foils, and the module PCBs will be received and qualified by the institutes responsible for their procurement. The qualified pieces will be most probably shipped to CERN for dispatching to the modules assembly centres.

Silicon sensors will follow the quality control described in Section 2.1. All sensors will be tested by vendors and sampling tests will be performed by HGCAL institutes, including irradiations. It is expected that the Silicon sensors quality control (except irradiation tests and tests of specialised structures) will be performed by the same centres which will assemble modules, to avoid shipping between institutes.

The modules assembly will be performed in five assembly centres. A sixth centre will act as a pilot centre. The role of the pilot centre is to define, design, and qualify the tooling necessary for the assembly and to specify all the procedures to be followed. In general the tooling will be identical in all the assembly centres, with the exception of commercial items such as the wire bonding machines, though the pilot centre may recommend particular equipment to the institutes that are not yet equipped. The pilot centre is intended to contribute to the production of partial modules but will also be able to study and react quickly to any general problem which could happen in the course of the production.

The bulk modules production is planned for a period of 21 months, which translates into a throughput of $\approx$50 modules per day, or 10 modules per day and per centre. Centres will however be equipped to double this production rate, if required.

After full testing, including thermal cycling, the modules will be shipped to the cassettes assembly centres.

The organisation and workflow for the scintillator tile-modules follow the same scheme as for the Silicon Modules (Fig. 13.2). Following testing of production SiPMs and automated testing of HGCROCs, the tileboards will be produced in industry and tested before mounting the scintillator tiles, using the integrated LED system. After integration of the scintillator tiles, the completed tile-modules will again be tested with the LED system and thermally cycled. It is currently foreseen to have at least three tile-modules assembly centres that may include the pilot centre (DESY). There are 3690 tile-modules to be assembled, so a throughput of about 3

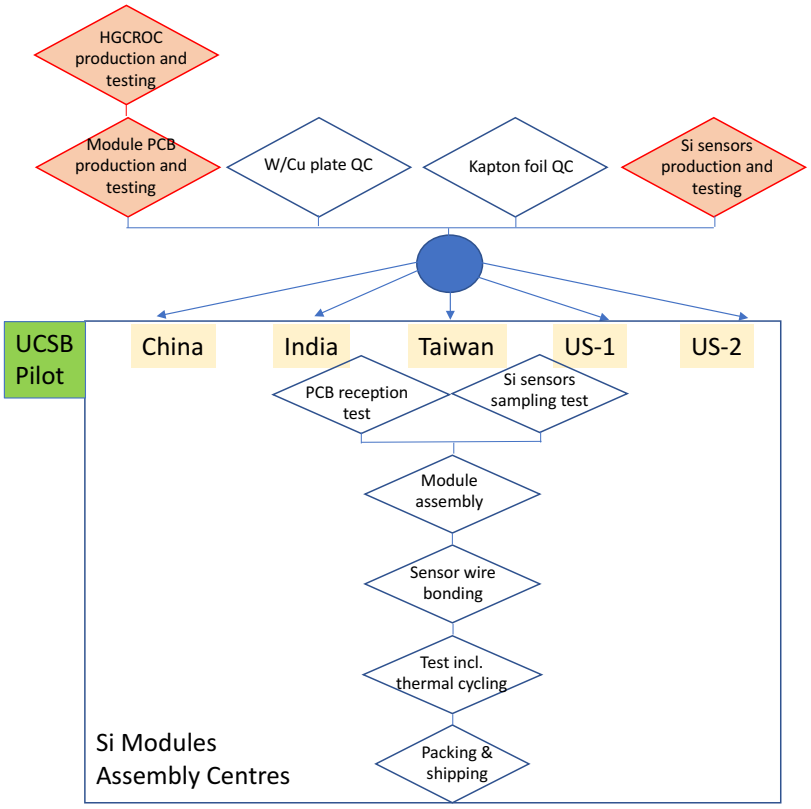


Figure 13.1: Workflow for silicon modules assembly. The processes in red are done in industry with sampling tests in HGCAL institutions. Detailed discussion can be found in Section 7.2.2.2.

per day and per centre is required. The centres will be equipped to be able to double this rate, if necessary.

There will be two cassettes assembly centres, one at CERN and one at FNAL (Fig. 13.3). Each centre will be equipped with the full set of tooling to be able to assemble any type of cassette. However it may be more practical that one centre works primarily on CE-E and the other on CE-H cassettes. Cassettes will be tested and thermally cycled after assembly, using a small subset of the readout equipment and a dedicated CO_2 cooling unit. The required cooling power is less than 600 W for the largest, double-sided CE-E cassette, and therefore commercially available cooling units can be used. There are 168 double-sided 60°CE-E cassettes, 96 pure-Silicon single sided 60°CE-H cassettes, and 384 CE-H 30°mixed cassettes to be assembled over a period of 21 months, so ≈16 cassettes per month have to be assembled per centre. Again a throughput of at least 32 per month per centre will be planned, so as to have a large contingency.

The final assembly workflow has been described in Section 4.6 and is recalled in Fig. 13.4. The CE-E cassettes will be shipped to the CE-E assembly centre at CERN (likely to be in Building 904, which may also host the CERN cassette assembly centre). CE-H cassettes will be shipped to the CE-H assembly hall, which must be at Point 5, given the large weight of the CE-H structure. Cassettes will be systematically tested before and after insertion. After the complete assembly of one endcap, a provisional thermal screen will be installed and the whole endcap will be tested during a period of 10 (7) months for HGCAL-1 (HGCAL-2). A ≈20 kW cooling unit is foreseen to be able to test a full 60°sector at once.

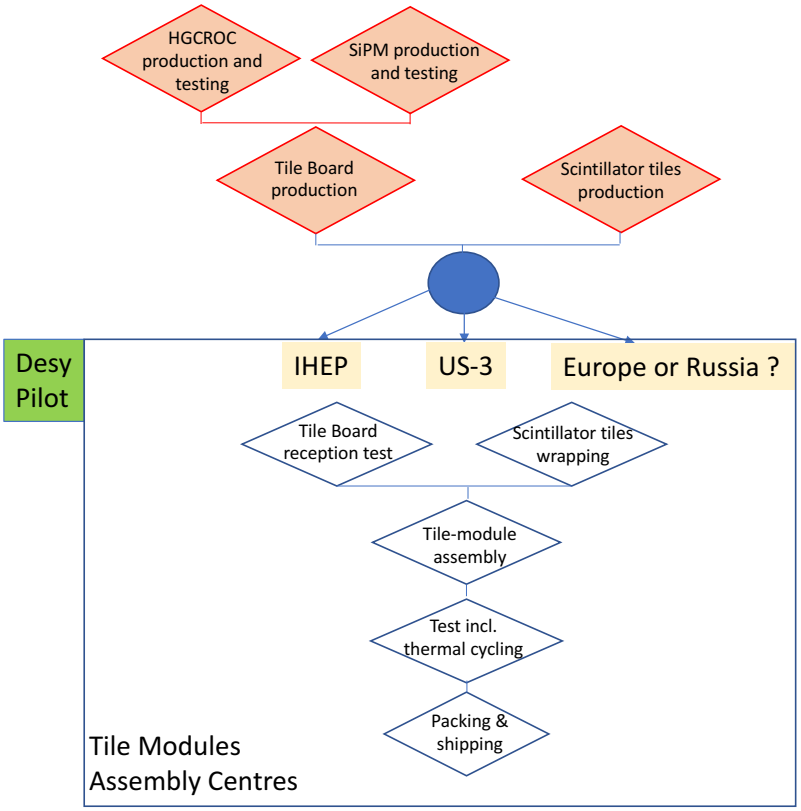


Figure 13.2: Workflow for scintillator tile-modules assembly. The processes outlined in red are done in industry. Detailed discussion can be found in Section 7.3.4.1.

13.2 Quality assurance

The Project Office is responsible for establishing and monitoring QA/QC procedures across the project. In particular, it will establish, with the appropriate sub-area coordinators, the necessary assembly centres and qualify them for production. A strict policy of reviews will be followed for each component before launching the production. An example has been given in Section 8.1.10 concerning the electronics ASICs.

Databases (DB) are an essential tool for QA/QC. The aim is to have full tracking of the history and of the testing of all the components in HGCAL. The same policy has been followed for recent CMS projects such as the muon GEM detectors and the Phase-1 pixel detector. It is therefore expected to build on this experience for the CMS Phase-2 Upgrade projects, in order to avoid duplication of the software development effort.

The DB will include all the detector components that need to be tracked, which will be uniquely identified by a serial number and/or a barcode. The components will be used to build higher level devices such as modules, cassettes, etc. The end result will be an HGCAL DB where every detector element will be mapped to a readout channel. Figure 13.5 shows an overview of the DB organisation, with its connection with the remote assembly centres. A web-based graphical interface will allow an easy monitoring of the construction progress, for example by publishing regularly dashboards.

The use of the DB will also enable the collaboration to better synchronize the construction activ-

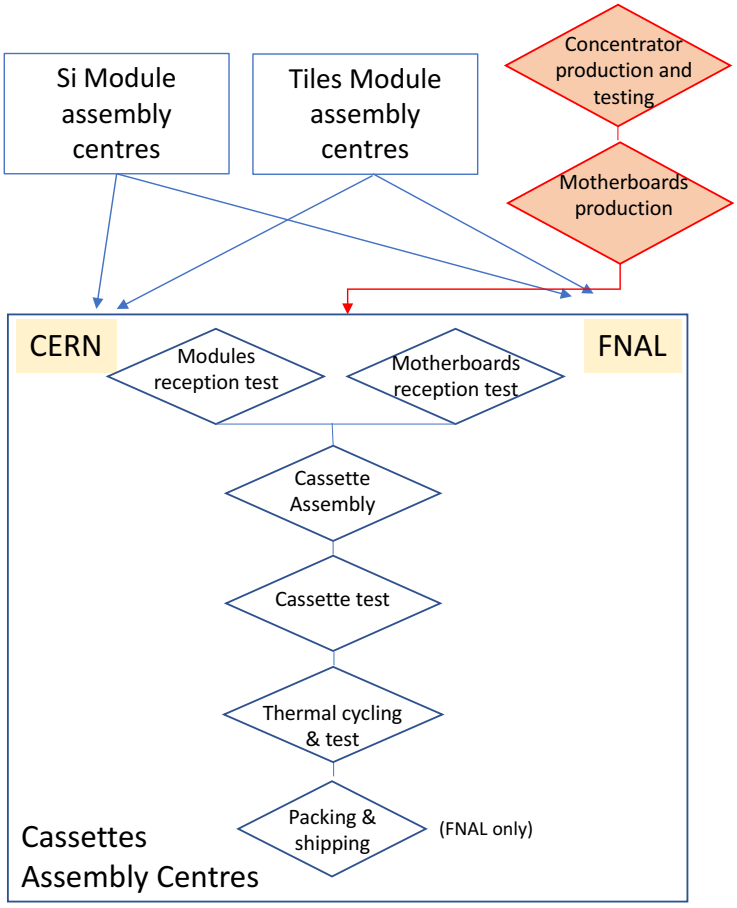


Figure 13.3: Workflow for cassettes assembly. The processes outlined in red are done in industry with sampling tests in HGCAL institutions. Detailed discussion can be found in Section 9.2.3.

ities distributed among the institutes. An HGCAL CORE DB group will oversee the DB loader framework, and the online monitoring system. Members in the various HGCAL institutions will interface with this expert group.

Upon completion of the detector, the DB will evolve towards the detector configuration DB and the detector condition DB, which will store all the types of data necessary to monitor the detector performance.

As a start of this effort, a DB loader for a development DB has been already deployed. The IV and CV curves of the sensors used for the test beam have been loaded in this initial data base.

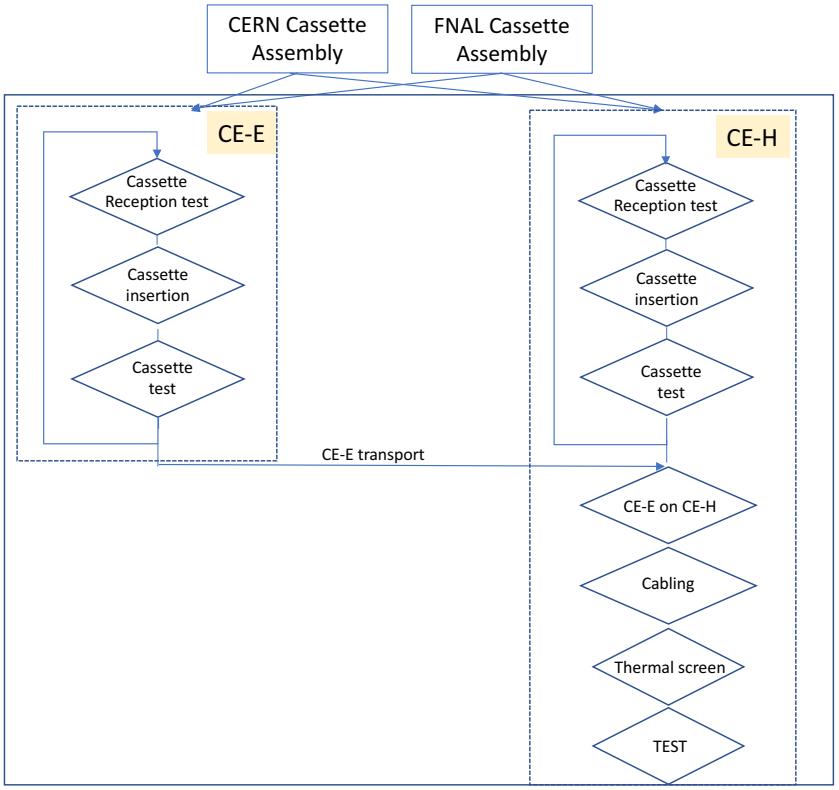


Figure 13.4: Workflow for final CE-E and CE-H assembly. A detailed discussion can be found in Section 9.6.

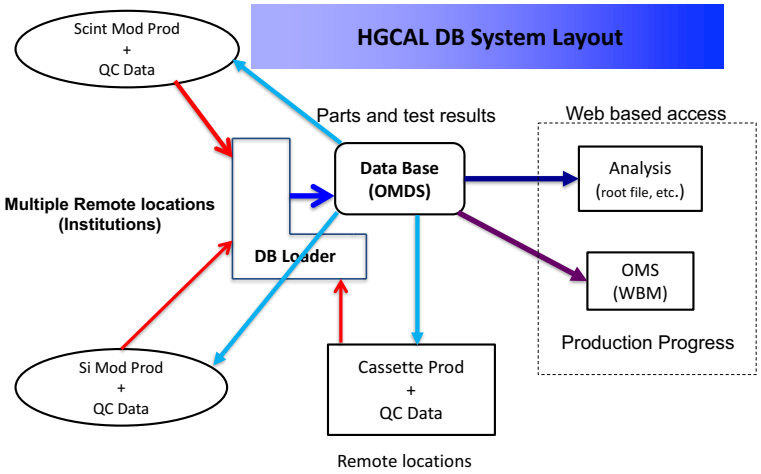


Figure 13.5: Overview of the organization of the HGCAL database system.

Appendix A

Event display

The next two pages show an event containing a VBF jet and a high p_{T} photon in one HGCAL endcap. Maps of individual layers are shown, and the energy of individual reconstructed hits are colour coded on a logarithmic scale. The example event is described in Section 1.3.

The power of the high granularity for pattern recognitions can be visualized by looking at a display of a VBF $\mathrm{H} \rightarrow \gamma\gamma$ event in one quadrant of the calorimeter as shown in Figs. A.1 and A.2. These shows the development of a VBF jet carrying 720 GeV ($p_{\mathrm{T}} = 118$ GeV) and a photon carrying 175 GeV ($p_{\mathrm{T}} = 22$ GeV) in an environment with a mean pileup of 200 interactions per bunch crossing. All the energy in the quadrant is projected into one plane, and the longitudinal development, plane-by-plane, in 9 planes of the electromagnetic section and 15 planes in the hadronic section. The wealth of information, the development of the showers and the narrowness of the VBF jet is clearly visible. Most of the energy of the VBF jet is carried by three particles (2 charged pions and one photon) that impact the within 1 cm at the front face.

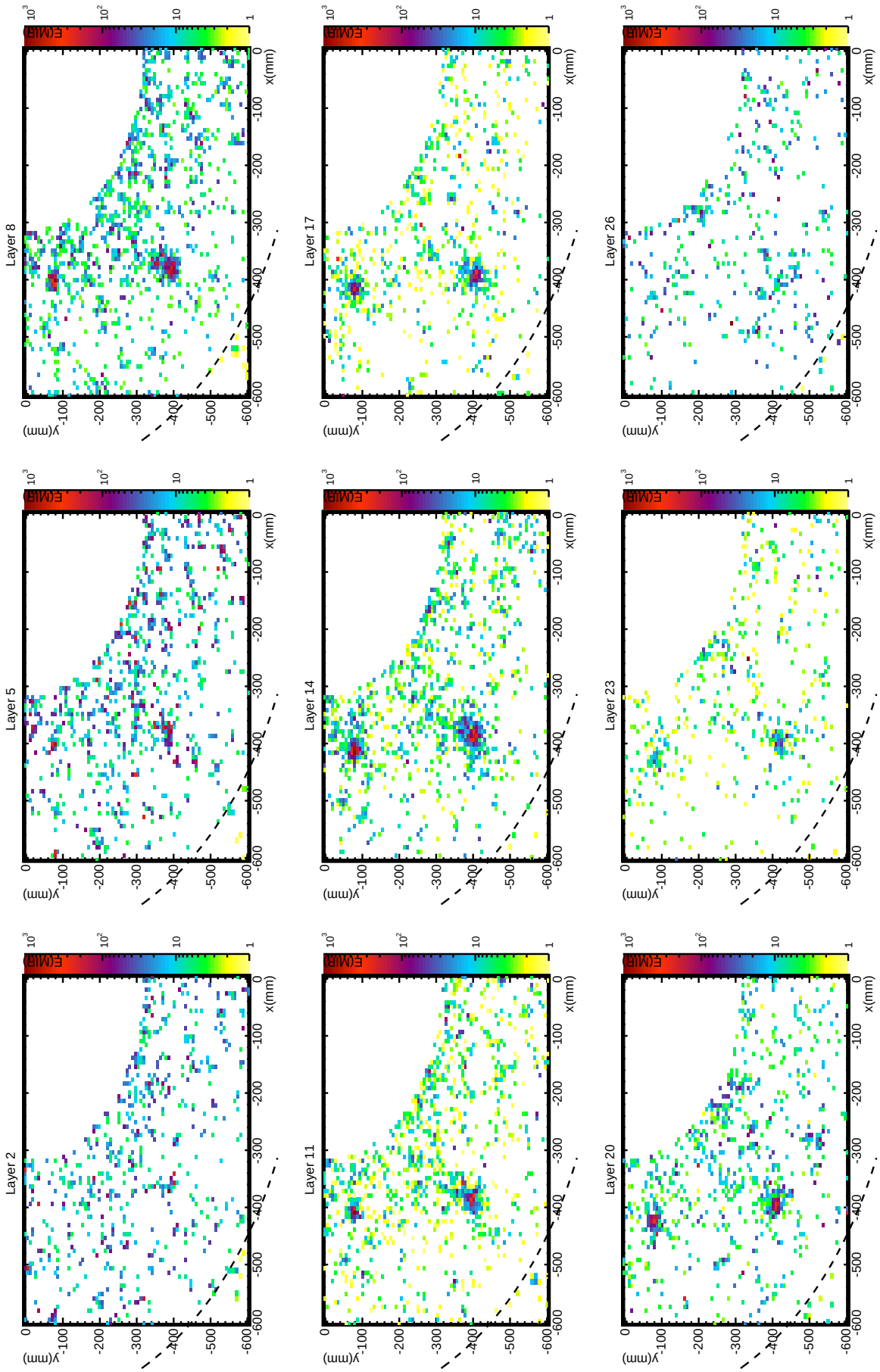


Figure A.1: Display of a VBF jet and a high p_{T} photon in 9 layers of CE-E. The energy of individual reconstructed hits are colour coded on a logarithmic scale. A detailed description of these displays is given in Section 1.2.

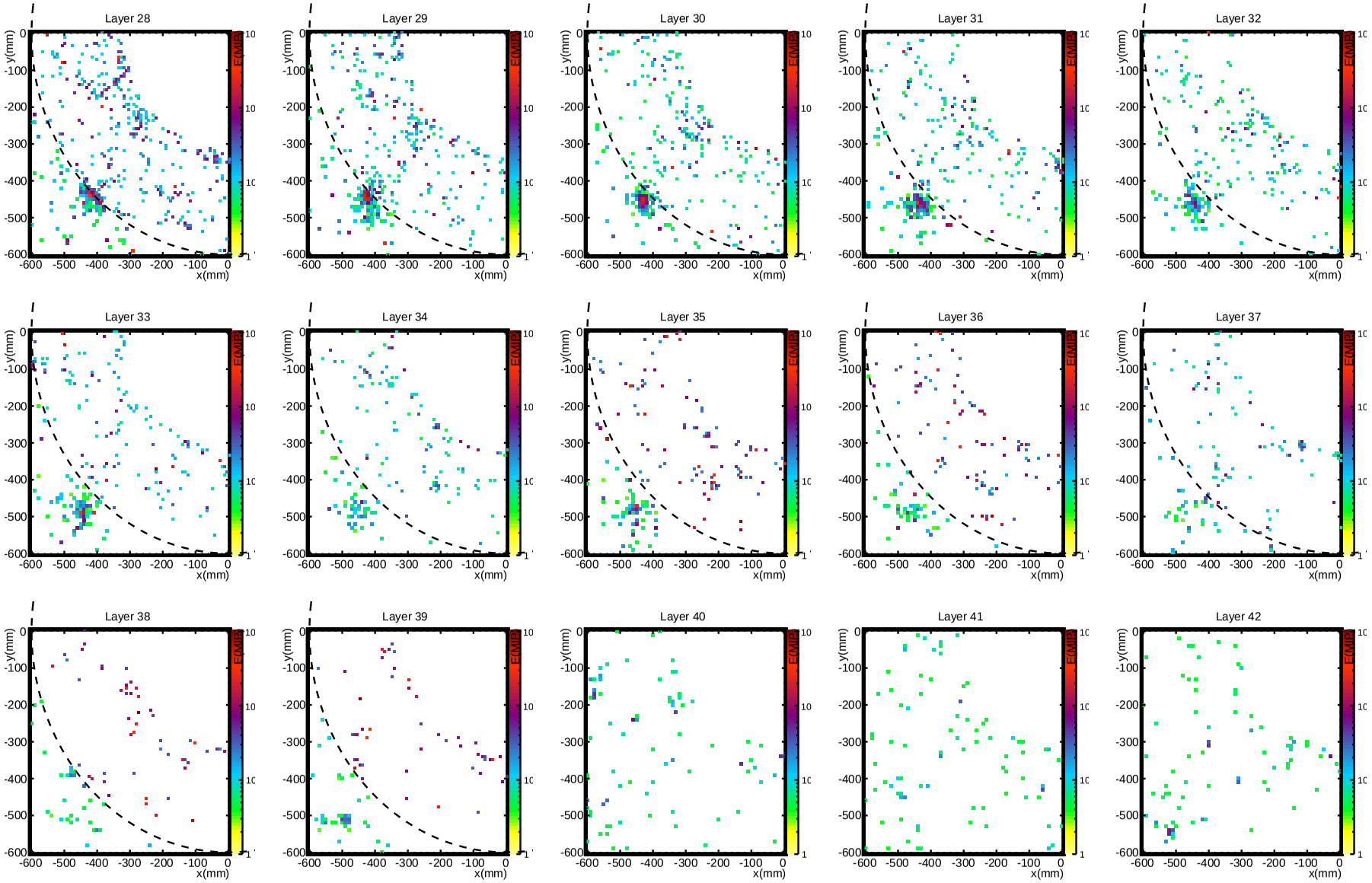


Figure A.2: Display of a VBF jet and a high p_{T} photon in 15 layers of CE-H. The energy of individual reconstructed hits are colour coded on a logarithmic scale. A detailed description of these displays is given in Section 1.2.

Glossary of special terms and acronyms

AHCAL – hadron calorimeter design of CALICE collaboration using SiPM detectors and tiles of plastic scintillator.

ASIC – application specific integrated circuit, i.e. a custom chip.

ATCA – advanced telecommunications computing architecture; ATCA and microTCA are variants of an open standard for carrier grade communication equipment.

BE – back-end (electronics).

BH – old/obsolete name, used in the technical proposal, for last 12 layers of the CE-H, which have thicker (68 mm) absorber plates.

BX – bunch crossing (used as a measure of a time interval, e.g. for latency).

DAQ – data acquisition

cassette – HGCAL subassemblies; a 60° or 30° sector consisting of a layered stack.

concentrator – ASIC collecting and selecting the data of several HGCROCs, and formatting them for high speed transmission to the back-end electronics.

CE – calorimeter endcap; label designated by CMS technical coordination for the Phase-2 Upgrade of the CMS endcap calorimeter.

CE-E – calorimeter endcap electromagnetic section; label designated by CMS technical coordination.

CE-H – calorimeter endcap hadronic section; label designated by CMS technical coordination.

CE-SS – HGCAL safety hardware system.

CMOS – complementary metal-oxide-semiconductor: integrated circuit technology.

CMSSW – CMS simulation and reconstruction software.

DAQ – data acquisition.

DCS – (CMS) detector control system.

DDMTD – digital dual-mixer time difference.

DSS – (CMS) detector safety system.

DTH – DAQ TDCS hub; a board providing a common CMS DAQ interface to each detector subsystem for timing and clock distribution.

EE – old/obsolete name, used in the technical proposal, for CE-E.

FE – front-end (electronics).

FEC5 – a forward error correction protocol.

FO – fibre optic.

FPGA – field-programmable gate array.

FH – old/obsolete name, used in the technical proposal, for first 12 layers of the CE-H, which have thinner (35 mm) absorber plates.

FZ – float zone; refers to a highly pure form of silicon manufactureded by this method.

GBT – gigabit transceiver.

GBT-SCA – GBT slow control adaptor ASIC.

HE – hadronic endcap in the current CMS detector.

hexaboard – hexagonal printed circuit board to which a silicon sensor is bonded, mounted on it are the HGCROCs, LVRs, and other passive and service components.

HGCAL – high granularity calorimeter: the Phase-2 Upgrade of the CMS endcap calorimeter.

HGCROC – readout ASIC designed for HGCAL.

LpGBT – low power gigabit transceiver.

LSB – least significant bit.

L1 – level-1 (trigger); referring to the part of the CMS trigger system implemented in custom hardware and providing the initial trigger decision on whether to transmit, from the detector, all data relating to an event.

L1A – level-1 (trigger) accept: signal specifying that an event has passed the L1 trigger and is to be read out from the detector.

L1T – level-1 trigger.

LVR – linear voltage regulator.

ME0 – Phase-2 Upgrade endcap muon chambers.

ME1/1 – CMS endcap muon chambers.

megatile – unit comprising multiple scintillator tiles used in existing CMS HCAL.

module – a stack of: baseplate, kapton-gold sheet, silicon sensor, and PCB.

motherboard - printed circuit board located either above the hexaboard (for silicon sensors), or at the periphery of the cassette (for SiPM-on-tile). Mounted on it are the concentrator, the LpGBT, the SCA and the optical links receivers/transmitters.

MIP – minimum-ionizing particle; or more loosely, a relativistic charged particle whose energy loss is dominated by ionization, i.e. approximately minimum-ionizing.

$\mathrm{MIP_T}$ – transverse MIP: signal measured in MIPs multiplied by $\sin\theta$.

$\mathbf{n_{eq}/cm^2}$ – measure of fluence (ϕ_{eq}): number of 1 MeV equivalent neutrons per square cm.

NIEL hypothesis – hypothesis that radiation damage in silicon detectors scales with non-ionising energy loss.

NMOS – n-channel MOSFET (metal-oxide-semiconductor field-effect transistor).

partial module – a module made using a partial wafer; used to complete the coverage at the inner or outer boundary of a layer.

partial wafer – a partial hexagonal silicon sensor wafer.

PCB – printed circuit board.

PLC – programmable logical controller.

PLL – phase-locked loop.

RTL – register-transfer level; a design abstraction which models a synchronous digital circuit in terms of the flow of digital signals between hardware registers, and the logical operations performed on those signals.

SAR – successive approximation register; type of ADC implementing a binary search algorithm.

SCA – slow control adapter ASIC responsible for the monitoring of environmental parameters, and for downloading of configuration parameters in the front-end.

SCADA – supervisory control and data acquisition.

SEU – single event upset; change of state caused by a single ionizing particle striking a sensitive location in a micro-electronic device; hence "SEU-compliant": designed to tolerate, and continue functioning in the presence of, single event upsets.

SiPM – silicon photomultiplier.

SiPM-on-tile – detector technology consisting of a silicon photomultiplier optically coupled to a scintillator tile through a small "dimple" machined or moulded in the centre of one face of the tile.

TCDS – timing and clock distribution system.

TCT – transient current technique; a technique in which a fast. infrared laser pulse is used to induce signals in a silicon sensor.

tileboard – the printed circuit board on which multiple scintillator tiles and their SiPMs are mounted.

tile-module – a subassemby of multiple scintillator tiles and their SiPMs mounted on a tileboard.

TID – total ionization dose.

ToA – time of arrival.

ToT – time-over-threshold.

TPG – trigger primitive generator.

Twinax – low-profile flexible flat signal cables, manufactured. by TempFlex which have been demonstrated to transfer data over a distance of 5 m at rates up to 8 Gb/s.

USC – CMS underground service cavern.

UXC – CMS underground experimental cavern.

VFE – very front-end; first stage of on-detector electronics, incorporating the preamplifier, shaper, digitization, L1 storage, and data formatting.

YE1 – yoke endcap 1: the first (innermost) iron disk of the magnetic flux return yoke which structurally supports the endcap calorimeter.

YN – yoke nose: attached to the innermost disk, and facing the interaction point, a solid iron piece 5.4 m in diameter and 700 mm thick.

2PACL – two-phase accumulator controlled loop: CO_2 cooling system.

64b/66b – encoding that transforms 64-bit data to 66-bit line code to provide enough state changes to allow reasonable clock recovery and facilitate alignment of the data stream at the receiver.

CMS institutes participating in the HGCAL project

Austria

Wien - Institut für Hochenergiephysik (HEPHY)

Belarus

Minsk - Byelorussian State University (MINSK-UNIV)

Bulgaria

Sofia - Institute for Nuclear Research and Nuclear Energy of Bulgaria Academy of Sciences (SOFIA-INRNE)

China

Beijing - Institute of High Energy Physics (BEIJING-IHEP)

Croatia

Split - University of Split, Faculty of Science (SPLIT-UNIV)

France

Gif-sur-Yvette - IRFU, CEA, Universit Paris-Saclay (SACLAY)
Palaiseau - Laboratoire Leprince-Ringuet, Ecole Polytechnique, CNRS/IN2P3, Universit Paris-Saclay (POLYTECHNIQUE), Organization for Microelectronics design and Applications (OMEGA)

Germany

Hamburg - Deutsches Elektronen-Synchrotron (DESY)

Greece

Athens - National and Kapodistrian University of Athens (ATHENS)
Athens - National Technical University of Athens (ATHENS-NTUA)
Ioannina - University of Ionnina (IOANNINA)

India

Bangalore - Indian Institute of Science (IISC) (BANGALORE-IISC)
Delhi - University of Delhi (DELHI-UNIV)
Kolkata - Saha Institute of Nuclear Physics, HBNI (KOLKATA-SAHA)
Madras - Indian Institute of Technology Madras, (IIT-MADRAS)
Mumbai - Tata Institute of Fundamental Research-A (TIFR-A)
Mumbai - Tata Institute of Fundamental Research-B (TIFR-B)

Mumbai - Bhabha Atomic Research Centre (MUMBAI-BARC)
Pune - Indian Institute of Science Education and Research (IISER) (PUNE-IISER)

Malaysia

Kuala Lumpur - National Centre for Particle Physics, Universiti Malaya (MALAYA)

Portugal

Lisboa - Laboratrio de Instrumentao e Fsica Experimental de Partculas (LIP)

Russia

Dubna - Joint Institute for Nuclear Research (DUBNA)
Moscow - Institute for Nuclear Research (MOSCOW-INR)
Moscow - Lebedev Physical Institute (MOSCOW-LEBEDEV)
Moscow - National Research Nuclear University 'Moscow Engineering Physics Institute (MEPHI)
Moscow - Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University (MOSCOW-MSU)
Moscow - Institute for Theoretical and Experimental Physics (MOSCOW-ITEP)
Protvino - State Research Center of Russian Federation, Institute for High Energy Physics (PROTVINO)

Switzerland

Geneva - CERN, European Organization for Nuclear Research (CERN)

Taiwan

Chung-Li - National Central University (CHUNGLI-NCU)
Taipei - National Taiwan University (NTU) (TAIPEI-NTU)

Thailand

Bangkok - Chulalongkorn University, Faculty of Science, Department of Physics (BANGKOK)

Turkey

Adana - Cukurova University (ADANA-CUKUROVA)
Istanbul - Bogazici University (BOGAZICI)
Istanbul - Istanbul Technical University (ISTANBUL-TECH)

Ukraine

Kharkov - Institute for Scintillation Materials of National Academy of Science of Ukraine (KHARKOV-ISMA)
Kharkov - National Scientific Center, Kharkov Institute of Physics and Technology (KHARKOV-KIPT)

United Kingdom

Bristol - University of Bristol (BRISTOL)
London - Imperial College (LONDON-IC)

USA

Waco - Baylor University (BAYLOR-UNIV)

Tuscaloosa - The University of Alabama (ALABAMA-UNIV)
Boston - Boston University (BOSTON-UNIV)
Providence - Brown University (BROWN-UNIV)
Santa Barbara - University of California, Santa Barbara (UCSB)
Davis - University of California, Davis (UCDAVIS)
Pittsburgh - Carnegie Mellon University (CARNEGIE-MELLON)
Fairfield - Fairfield University (FAIRFIELD)
Batavia - Fermi National Accelerator Laboratory (FERMILAB)
Tallahassee - Florida State University (FLORIDA-STATE)
Melbourne - Florida Institute of Technology (FLORIDA-TECH)
Dekalb - Northern Illinois University (DEKALB-NIU)
Iowa City - The University of Iowa (IOWA)
Manhattan - Kansas State University (KANSAS-STATE)
College Park - University of Maryland (MARYLAND)
Cambridge - Massachusetts Institute of Technology (MIT)
Minneapolis - University of Minnesota (MINNESOTA)
Evanston - Northwestern University (NORTHWESTERN)
Notre Dame - University of Notre Dame (NOTRE DAME)
Rochester - University of Rochester (ROCHESTER)
Lubbock - Texas Tech University (TEXAS-TECH)

References

[1] CMS Collaboration, "Technical Proposal for the Phase-II Upgrade of the Compact Muon Solenoid", Technical Report CERN-LHCC-2015-010, LHCC-P-008, 2015.

[2] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[3] CMS and ATLAS Collaborations, L. Gray, "Picosecond Timing". https://indico.cern.ch/event/468486/contributions/1144338/attachments/1240096/1850959/ACES_LindseyGray_08032016.pdf.

[4] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys. C* **40** (2016) 100001, doi:10.1088/1674-1137/40/10/100001.

[5] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker: Technical Design Report", Technical Report CERN-LHCC-2017-009, CMS-TDR-014, 2017.

[6] CMS Collaboration, "Planar Silicon Sensors for the ATLAS and CMS Outer Tracker Upgrades", Technical Report MS-4086/EP, 2016.

[7] T. Bergauer et al., "First thin AC-coupled silicon strip sensors on 8-inch wafers", *Nucl. Instrum. Meth. A* **830** (2016) 473–479, doi:10.1016/j.nima.2016.05.076.

[8] CMS HCAL Collaboration, "Dose rate effects in the radiation damage of the plastic scintillators of the CMS Hadron Endcap Calorimeter", *JINST* **11** (2016) T10004, doi:10.1088/1748-0221/11/10/T10004, arXiv:1608.07267.

[9] B. Bicken et al., "Recovery and permanent radiation damage of plastic scintillators at different dose rates", *IEEE Transactions on Nuclear Science* **38** (1991) 188, doi:10.1109/23.289295.

[10] C. Zorn, "Plastic and Liquid Organic Scintillators". World Scientific, 2011. doi:10.1142/9789814360333_0004.

[11] G. I. Britvich, V. V. Brekhovskikh, V. K. Semenov, and S. A. Kholodenko, "The main characteristics of polystyrene scintillators produced at the institute of high-energy physics and detectors on their basis", *Instrum. Exp. Tech.* **58** (2015) 211, doi:10.1134/S0020441215020153. [Prib. Tekh. Eksp. 2015, no.2, 47].

[12] Y. Musienko et al., "Radiation damage studies of silicon photomultipliers for the CMS HCAL phase I upgrade", *Nucl. Instrum. Meth. A* **787** (2015) 319, doi:10.1016/j.nima.2015.01.012.

[13] A. Heering et al., "Effects of very high radiation on SiPMs", *Nucl. Instrum. Meth. A* **824** (2016) 111, doi:10.1016/j.nima.2015.11.037.

[14] CALICE Collaboration, "Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype", *JINST* **5** (2010) P05004, doi:10.1088/1748-0221/5/05/P05004, arXiv:1003.2662.

[15] G. Blazey et al., "Directly Coupled Tiles as Elements of a Scintillator Calorimeter with MPPC Readout", *Nucl. Instrum. Meth. A* **605** (2009) 277, doi:10.1016/j.nima.2009.03.253.

[16] CERN, "LpGBT specification document". https://espace.cern.ch/GBT-Project/LpGBT/Specifications/LpGbtxSpecifications.pdf.

[17] CERN, "The slow control adapter ASIC for the GBT system: User Manual". https://espace.cern.ch/GBT-Project/GBT-SCA/Manuals/GBT-SCA-UserManual.pdf.

[18] TTC-project, "TTC-project web page". http://ttc.web.cern.ch/TTC/.

[19] S. Baron, "Passive Optical Network for TTC". https://aces.web.cern.ch/aces/aces2011/ACES2011.htm.

[20] CMS Collaboration, "The Phase-2 Upgrade of the CMS DAQ: Interim Technical Design Report", Technical Report CERN-LHCC-2017-014, CMS-TDR-018, 2017.

[21] CMS Collaboration, "The Phase-2 Upgrade of the CMS Level-1 Trigger: Interim Technical Design Report", Technical Report CERN-LHCC-2017-013, CMS-TDR-017, 2017.

[22] PICMG, "R3.0 AdvancedTCA Base Specification". https://www.picmg.org/openstandards/advancedtca/.

[23] CERN, "FEASTMP datasheet". https://project-dcdc.web.cern.ch/project-dcdc/public/Documents/FEASTMod_Datasheet.pdf.

[24] P. Tropea et al., "Design, construction and commissioning of a 15 kW CO_2 evaporative cooling system for particle physics detectors: lessons learnt and perspectives for further development", in *Proceedings, 3rd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2014): Amsterdam, Netherlands, June 2-6, 2014*, volume TIPP2014, p. 223. 2014.

[25] J. Van Es et al., "AMS02 Tracker Thermal Control System Overview and Spin-off for Future Spacecraft Cooling System Developments", in *Proceedings of the 60th International Astronautical Congress, Daejeon, South Korea (2009).*

[26] B. Verlaat et al., "CO_2 Cooling for the LHCb-VELO Experiment at CERN", in *Proceedings of the 8th IIF/IIR Gustav Lorentzen Conference on Natural Working Fluids, Copenhagen, Denmark (2008).*

[27] M. Thomson, "Particle Flow Calorimetry and the PandoraPFA Algorithm", *Nucl. Instrum. Meth. A* **611** (2009) 25, doi:10.1016/j.nima.2009.09.009, arXiv:0907.3577.

[28] J. Marshall, A. Münnich, and M. Thomson, "Performance of Particle Flow Calorimetry at CLIC", *Nucl. Instrum. Meth. A* **700** (2013) 153, doi:10.1016/j.nima.2012.10.038, arXiv:1209.4039.

[29] CMS Collaboration, "The Phase-2 Upgrade of the CMS Barrel Calorimeters: Technical Design Report", Technical Report CERN-LHCC-2017-011, CMS-TDR-015, 2017.

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

[31] CALICE Collaboration, "Track segments in hadronic showers in a highly granular scintillator-steel hadron calorimeter", *JINST* **8** (2013) P09001, doi:10.1088/1748-0221/8/09/P09001, arXiv:1305.7027.

[32] S. Callier et al., "SKIROC2, front end chip designed to readout the electromagnetic calorimeter at the ILC", *JINST* **6** (2011) C12040, doi:10.1088/1748-0221/6/12/C12040.

[33] J. Borg et al., "SKIROC2_CMS an ASIC for testing CMS HGCAL", *JINST* **12** (2017) C02019, doi:10.1088/1748-0221/12/02/C02019.

[34] N. Akchurin et al., "On the timing performance of thin planar silicon sensors", *Nucl. Instrum. Meth. A* **859** (2017) 31, doi:10.1016/j.nima.2017.03.065.

[35] CMS Collaboration, "Technical Proposal for a MIP Timing Detector in the CMS experiment Phase-2 Upgrade", Technical Report CERN-LHCC-2017-027, LHCC-P-009, 2017.

[36] E. Currás et al., "Radiation hardness and precision timing study of silicon detectors for the CMS High Granularity Calorimeter (HGC)", *Nucl. Instrum. Meth. A* **845** (2017) 60, doi:10.1016/j.nima.2016.05.008.

[37] CMS Collaboration, "P-Type Silicon Strip Sensors for the new CMS Tracker at HL-LHC", *JINST* **12** (2017) P06018, doi:10.1088/1748-0221/12/06/P06018.

[38] F. Abu-Ajamieh et al., "Beam tests of directly coupled scintillator tiles with MPPC readout", *Nucl. Instrum. Meth. A* **659** (2011) 348, doi:10.1016/j.nima.2011.06.090.

[39] Y. Liu et al., "A Design of Scintillator Tiles Read Out by Surface-Mounted SiPMs for a Future Hadron Calorimeter", in *Proceedings, 21st Symposium on Room-Temperature Semiconductor X-ray and Gamma-ray Detectors (RTSD 2014): Seattle, WA, USA, November 8-15, 2014*, p. 7431118. 2016. arXiv:1512.05900. doi:10.1109/NSSMIC.2014.7431118.

[40] N. Tsuji, "Presentation at CALICE collaboration Meeting Tokyo". https://agenda.linearcollider.org/event/7630/contributions/39723/attachments/32042/48441/CALICE_2017_tokyo_tsuji.pdf.

[41] M. Janecek, "Reflectivity Spectra for Commonly Used Reflectors", *IEEE Transactions on Nuclear Science* **59** (2012) 490, doi:10.1109/TNS.2012.2183385.

[42] Y. Liu, "Recent progress on the megatile studies". https://agenda.linearcollider.org/event/7630/contributions/39722/, 2017. CALICE Collaboration Meeting.

[43] P. Chau, N. van der Kolk, and K. Briggl, "Specification of systems for highly granular scintillator tests". https://cds.cern.ch/record/2153646/files/.

[44] S. Bugiel et al., "Ultra-Low Power Fast Multi-Channel 10-Bit ADC ASIC for Readout of Particle Physics Detectors", *IEEE Transactions on Nuclear Science* **6** (2016) 2622.

[45] F. Faccio, "Radiation hardness issues (ACES Workshop, 2016)". https://indico.cern.ch/event/468486/contributions/1144364/attachments/1239064/1821580/RadTechnologies.pdf.

[46] J. Hegeman et al., "The CMS Timing and Control Distribution System", in *Proceedings, 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC 2015): San Diego, California, United States*, p. 7581984. 2016. doi:10.1109/NSSMIC.2015.7581984.

[47] K. Briggl et al., "KLauS: an ASIC for silicon photomultiplier readout and its application in a setup for production testing of scintillating tiles", *JINST* **9** (2014) C02013, doi:10.1088/1748-0221/9/02/C02013.

[48] Samtec, "IC-to-board capabilities and design solutions guide". http://suddendocs.samtec.com/literature/ic2b_solutions_guide.pdf.

[49] Advantech, "COM Express Mini product line". http://www.advantech.com/products/com-express-mini/sub_1-dwbvx8.

[50] Xilinx Incorporated, "Ultrascale Architecture and Product Data Sheet". https://www.xilinx.com/support/documentation/data_sheets/ds890-ultrascale-overview.pdf.

[51] K. Compton et al., "The MP7 and CTP-6: Multi-hundred Gbps processing boards for calorimeter trigger upgrades at CMS", *JINST* **7** (2012) C12024, doi:10.1088/1748-0221/7/12/C12024.

[52] "HE Engineering and Safety Note", CMS Document 2000-007, CERN EDMS 112848 (2000).

[53] L. Zwalinski et al., "CO_2 cooling system for Insertable B Layer detector into the ATLAS experiment", in *Proceedings, 3rd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2014): Amsterdam, Netherlands, June 2-6, 2014*, volume TIPP2014, p. 224. 2015.

[54] L. Zwalinski et al., "The Control System for the CO_2 cooling plants for Physics experiments", in *Proceedings, 14th International Conference on Accelerator and Large Experimental Physics Control Systems, San Francisco, USA, 2013*. 2013.

[55] Habia Cable, http://www.habia.com/.

[56] Y. LeCun et al., "Handwritten digit recognition with a back-propagation network", in *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, ed., p. 396. Morgan-Kaufmann, 1990.

[57] B. Bergmann et al., "Ionizing Energy Depositions After Fast Neutron Interactions in Silicon", *IEEE Transactions on Nuclear Science* **63** (2016) 2372, doi:10.1109/TNS.2016.2574961.

[58] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at sqrt(s) = 8 TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[59] K. Klein, "The Phase-2 Upgrade of the CMS Tracker", Technical Report CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, Jun, 2017.

[60] W. Adam, R. Frhwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC", Technical Report CMS-NOTE-2005-001, CERN, Geneva, Jan, 2005.

[61] E. Longo and I. Sestili, "Monte Carlo Calculation of Photon Initiated Electromagnetic Showers in Lead Glass", *Nucl. Instrum. Meth.* **128** (1975) 283, doi:10.1016/0029-554X(75)90679-5. [Erratum: Nucl. Instrum. Meth.135,587(1976)].

[62] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[63] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification", *JHEP* **10** (2014) 59, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[64] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[65] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[66] D. Guest et al., "Jet flavor classification in high-energy physics with deep neural networks", *Phys. Rev. D* **94** (2016) 112002, doi:10.1103/PhysRevD.94.112002, arXiv:1607.08633.

[67] CMS Collaboration, "Reconstruction and identification of lepton decays to hadrons and at CMS", *JINST* **11** (2016), no. 01, P01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.

[68] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015) P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[69] CMS Collaboration, "Jet algorithms performance in 13 TeV data", Technical Report CMS-PAS-JME-16-003, CERN, Geneva, 2017.

[70] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[71] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness", *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[72] CMS Collaboration, "Identification of double-b quark jets in boosted event topologies", Technical Report CMS-PAS-BTV-15-002, CERN, Geneva, 2016.

[73] CMS Collaboration Collaboration, "Measurements of properties of the Higgs boson in the diphoton decay channel with the full 2016 data set", Technical Report CMS-PAS-HIG-16-040, CERN, Geneva, 2017.

[74] CMS Collaboration Collaboration, "Observation of the SM scalar boson decaying to a pair of τ leptons with the CMS experiment at the LHC", Technical Report CMS-PAS-HIG-16-043, CERN, Geneva, 2017.

[75] CMS Collaboration, "Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton-proton collisions at sqrt(s) = 13 TeV", arXiv:1707.02909.

[76] L. Bianchini, J. Conway, E. K. Friis, and C. Veelken, "Reconstruction of the Higgs mass in $H \to \tau\tau$ events by dynamical likelihood techniques", in *Proceedings, 20th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2013)*, p. 022035. Amsterdam, the Netherlands, October, 2014. J. Phys. Conf. Ser. 513 (2014) 022035. doi:10.1088/1742-6596/513/2/022035.

[77] D. de Florian et al., "Handbook of LHC Higgs cross sections: 4. deciphering the nature of the Higgs sector", CERN Report CERN-2017-002-M, 2016. doi:10.23731/CYRM-2017-002, arXiv:1610.07922.

[78] A. J. Barr, M. J. Dolan, C. Englert, and M. Spannowsky, "Di-Higgs final states augMT2ed – selecting *hh* events at the high luminosity LHC", *Phys. Lett. B* **728** (2014) 308, doi:10.1016/j.physletb.2013.12.011, arXiv:1309.6318.

[79] H. Baer et al., "Radiatively-driven natural supersymmetry at the LHC", *JHEP* **12** (2013) 013, doi:10.1007/JHEP12(2013)013, 10.1007/JHEP06(2015)053, arXiv:1310.4858. [Erratum: JHEP06,053(2015)].

[80] H. Baer et al., "Multi-channel assault on natural supersymmetry at the high luminosity lhc", *Phys. Rev. D* **94** (2016) 4, doi:10.1103/PhysRevD.94.035025.

[81] T. D. . collaboration, "Delphes 3: a modular framework for fast simulation of a generic collider experiment", *Journal of High Energy Physics* **2014** (2014), no. 2, doi:10.1007/JHEP02(2014)057.

[82] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, "Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV", *JHEP* **10** (2012) 081, doi:10.1007/JHEP10(2012)081, arXiv:1207.2159.

[83] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, "Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO", *Eur. Phys. J. C* **73** (2013) 2480, doi:10.1140/epjc/s10052-013-2480-0, arXiv:1304.0790.

[84] "LHC SUSY Cross Section Working Group". https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections#Cross_sections_for_various_S_AN3.

[85] CMS Collaboration, "Search for electroweak production of charginos and neutralinos in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:1709.05406.

[86] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl.Instrum.Meth.* **A434** (1999) 435–443, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[87] A. L. Read, "Presentation of search results: the CLs technique", *J. Phys. G: Nucl. Part. Phys.* **28** (2002).

[88] CMS Collaboration Collaboration, "Search for new physics in events with two low momentum opposite-sign leptons and missing transverse energy at $\sqrt{s} = 13$ TeV", Technical Report CMS-PAS-SUS-16-048, CERN, Geneva, 2017.

[89] H. Baer et al., "Aspects of the same-sign diboson signature from wino pair production with light higgsinos at the high luminosity LHC", arXiv:1710.09103.

[90] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak interactions with lepton hadron symmetry", *Phys. Rev. D* **2** (1970) 1285, doi:10.1103/PhysRevD.2.1285.

[91] J. A. Aguilar-Saavedra and B. M. Nobre, "Rare top decays $t \to c\gamma, t \to cg$ and ckm unitarity", *Phys. Lett. B* **553** (2003) 251, doi:10.1016/S0370-2693(02)03230-6, arXiv:hep-ph/0210360.

[92] G. Couture, M. Frank, and H. Konig, "Supersymmetric qcd flavor changing top quark decay", *Phys. Rev. D* **56** (1997) 4213, doi:10.1103/PhysRevD.56.4213, arXiv:hep-ph/9704305.

[93] G. R. Lu, F. R. Yin, X. L. Wang, and L. D. Wan, "Rare top quark decays t→cv in the topcolor assisted technicolor model", *Phys. Rev. D* **68** (2003) 015002, doi:10.1103/PhysRevD.68.015002, arXiv:hep-ph/0303122.

[94] CMS Collaboration, "Search for anomalous single top quark production in association with a photon in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **04** (2016) 035, doi:10.1007/JHEP04(2016)035, arXiv:1511.03951.

[95] "ECFA 2016: Prospects for selected standard model measurements with the CMS experiment at the High-Luminosity LHC", Technical Report CMS-PAS-FTR-16-006, CERN, Geneva, 2017.

[96] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, "FLUKA: A multi-particle transport code (Program version 2005)", 2005.

[97] T. T. Böhlen et al., "The FLUKA Code: Developments and Challenges for High Energy and Medical Applications", *Nucl. Data Sheets* **120** (2014) 211, doi:10.1016/j.nds.2014.07.049.

[98] S. Roesler, R. Engel, and J. Ranft, "The Monte Carlo event generator DPMJET-III", in *Advanced Monte Carlo for radiation physics, particle transport simulation and applications. Proceedings, Conference, MC2000, Lisbon, Portugal, October 23-26, 2000*, p. 1033. 2000. arXiv:hep-ph/0012252. doi:10.1007/978-3-642-18211-2_166.

[99] T. Pierog et al., "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider", *Phys. Rev. C* **92** (2015) 034906, doi:10.1103/PhysRevC.92.034906, arXiv:1306.0121.

CMS Collaboration

Yerevan Physics Institute, Yerevan, Armenia
A.M. Sirunyan, A. Tumasyan

Institut für Hochenergiephysik, Wien, Austria
W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, W. Brandner, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, A. Escalante Del Valle, M. Flechl, M. Friedl, R. Frühwirth[1], V.M. Ghete, J. Grossmann, J. Hrubec, M. Jeitler[1], A. König, N. Krammer, I. Krätschmer, S. Kulkarni, D. Liko, T. Madlener, I. Mikulec, E. Pree, N. Rad, H. Rohringer, J. Schieck[1], R. Schöfbeck, S. Schultschik, M. Spanring, D. Spitzbart, A. Taurok, M. Valentan, W. Waltenberger, J. Wittmann, C.-E. Wulz[1], M. Zarucki

Institute for Nuclear Problems, Minsk, Belarus
V. Chekhovsky, V. Mossolov, J. Suarez Gonzalez

Universiteit Antwerpen, Antwerpen, Belgium
E.A. De Wolf, D. Di Croce, X. Janssen, J. Lauwers, M. Pieters, M. Van De Klundert, H. Van Haevermaet, P. Van Mechelen, N. Van Remortel

Vrije Universiteit Brussel, Brussel, Belgium
S. Abu Zeid, F. Blekman, J. D'Hondt, I. De Bruyn, J. De Clercq, K. Deroover, G. Flouris, D. Lontkovskyi, S. Lowette, I. Marchesini, S. Moortgat, L. Moreels, Q. Python, K. Skovpen, S. Tavernier, W. Van Doninck, P. Van Mulders, I. Van Parijs

Université Libre de Bruxelles, Bruxelles, Belgium
D. Beghin, B. Bilin, H. Brun, B. Clerbaux, G. De Lentdecker, H. Delannoy, B. Dorney, G. Fasanella, L. Favart, R. Goldouzian, A. Grebenyuk, A.K. Kalsi, T. Lenzi, J. Luetic, T. Seva, E. Starling, C. Vander Velde, P. Vanlaer, D. Vannerom, R. Yonamine

Ghent University, Ghent, Belgium
T. Cornelis, D. Dobur, A. Fagot, M. Gul, I. Khvastunov[2], D. Poyraz, C. Roskas, D. Trocino, M. Tytgat, W. Verbeke, M. Vit, N. Zaganidis

Université Catholique de Louvain, Louvain-la-Neuve, Belgium
H. Bakhshiansohi, O. Bondu, S. Brochet, G. Bruno, C. Caputo, A. Caudron, P. David, S. De Visscher, C. Delaere, M. Delcourt, B. Francois, A. Giammanco, G. Krintiras, V. Lemaitre, A. Magitteri, A. Mertens, M. Musich, K. Piotrzkowski, L. Quertenmont, A. Saggio, M. Vidal Marono, S. Wertz, J. Zobec

Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil
W.L. Aldá Júnior, F.L. Alves, G.A. Alves, L. Brito, G. Correia Silva, C. Hensel, A. Moraes, M.E. Pol, P. Rebello Teles

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
E. Belchior Batista Das Chagas, W. Carvalho, J. Chinellato[3], E. Coelho, E.M. Da Costa, G.G. Da Silveira[4], D. De Jesus Damiao, S. Fonseca De Souza, L.M. Huertas Guativa, H. Malbouisson, M. Medina Jaime[5], M. Melo De Almeida, C. Mora Herrera, L. Mundim, H. Nogima, L.J. Sanchez Rosas, A. Santoro, A. Sznajder, M. Thiel, E.J. Tonelli Manganote[3], F. Torres Da Silva De Araujo, A. Vilela Pereira

Universidade Estadual Paulista [a], Universidade Federal do ABC [b], São Paulo, Brazil
S. Ahuja[a], C.A. Bernardes[a], T.R. Fernandez Perez Tomei[a], E.M. Gregores[b], P.G. Mercadante[b], S.F. Novaes[a], Sandra S. Padula[a], D. Romero Abad[b], J.C. Ruiz Vargas[a]

Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria
A. Aleksandrov, R. Hadjiiska, P. Iaydjiev, A. Marinov, M. Misheva, M. Rodozov, M. Shopova, G. Sultanov

University of Sofia, Sofia, Bulgaria
A. Dimitrov, L. Litov, B. Pavlov, P. Petkov

Beihang University, Beijing, China
W. Fang[6], X. Gao[6], L. Yuan

Institute of High Energy Physics, Beijing, China
M. Ahmad, J.G. Bian, X. Cao, G.M. Chen, H.S. Chen, M. Chen, Y. Chen, Y. Gu, C.H. Jiang, D. Leggat, B. Li, X. Li, H. Liao, X. Liu, Z. Liu, B. Meng, N. Peng, F. Romeo, S.M. Shaheen, A. Spiezia, L. Sun, J. Tao, C. Wang, F. Wang, Z. Wang, E. Yazgan, T. Yu, C. Zhang, H. Zhang, W. Zhang, J. Zhao

State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
Y. Ban, G. Chen, J. Li, Q. Li, S. Liu, Y. Mao, S.J. Qian, D. Wang, Z. Xu

Tsinghua University, Beijing, China
Y. Wang

Universidad de Los Andes, Bogota, Colombia
C. Avila, A. Cabrera, C.A. Carrillo Montoya, L.F. Chaparro Sierra, C. Florez, C.F. González Hernández, J.D. Ruiz Alvarez, M.A. Segura Delgado

University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia
D. Coko, B. Courbon, D. Giljanovi;, N. Godinovic, A. Kristic, D. Lelas, J. Music, J. Ozegovic, V. Pekic, M. Prvan, I. Puljak, P.M. Ribeiro Cipriano, T. Sculac

University of Split, Faculty of Science, Split, Croatia
Z. Antunovic, M. Kovac

Institute Rudjer Boskovic, Zagreb, Croatia
V. Brigljevic, D. Ferencek, K. Kadija, B. Mesic, A. Starodumov[7], T. Susa

University of Cyprus, Nicosia, Cyprus
M.W. Ather, A. Attikis, G. Mavromanolakis, J. Mousa, C. Nicolaou, F. Ptochos, P.A. Razis, H. Rykaczewski

Charles University, Prague, Czech Republic
M. Finger[8], M. Finger Jr.[8]

Universidad San Francisco de Quito, Quito, Ecuador
E. Carrera Jarrin

Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
H. Abdalla[9], Y. Assran[10,11], A. Mohamed[12]

National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
I. Ahmed[13], S. Bhowmik, R.K. Dewanjee, M. Kadastik, L. Perrini, M. Raidal, C. Veelken

Department of Physics, University of Helsinki, Helsinki, Finland
P. Eerola, H. Kirschenmann, J. Pekkanen, M. Voutilainen

Helsinki Institute of Physics, Helsinki, Finland
J. Havukainen, J.K. Heikkilä, T. Järvinen, V. Karimäki, R. Kinnunen, T. Lampén, K. Lassila-Perini, S. Laurila, S. Lehti, T. Lindén, P. Luukka, T. Mäenpää, H. Siikonen, E. Tuominen, J. Tuominiemi

Lappeenranta University of Technology, Lappeenranta, Finland
T. Tuuva

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
P.A. Bausson, M. Besancon, F. Bouyjou, D. Calvet, F. Couderc, M. Dejardin, D. Denegri, J.L. Faure, F. Ferri, S. Ganjour, O. Gevin, S. Ghosh, A. Givernaud, P. Gras, G. Hamel de Monchenault, P. Jarry, C. Leloup, E. Locci, M. Machet, J. Malcles, I. Mandjavidze, G. Negro, J. Rander, A. Rosowsky, M.Ö. Sahin, M. Titov

Laboratoire Leprince-Ringuet, Ecole polytechnique, CNRS/IN2P3, Université Paris-Saclay, Palaiseau, France
A. Abdulsalam[14], C. Amendola, I. Antropov, S. Baffioni, F. Beaudette, E. Becheva, P. Busson, L. Cadamuro, S. Callier, C. Charlot, A. Chiron, C. De La Taille[13], F. Dulucq[13], Y. Geerebaert, R. Granier de Cassagnac, M. Jo, I. Kucher, S. Lisniak, A. Lobanov, J. Martin Blanco, M. Nguyen, C. Ochando, G. Ortona, P. Paganini, T. Pierre-Emile, P. Pigard, L. Raux, T. Romanteau, R. Salerno, J.B. Sauvan, Y. Sirois, A.G. Stahl Leiton, D. Thienpont[13], Y. Yilmaz, A. Zabi, A. Zghiche

Université de Strasbourg, CNRS, IPHC UMR 7178, F-67000 Strasbourg, France
J.-L. Agram[15], J. Andrea, D. Bloch, J.-M. Brom, M. Buttignol, E.C. Chabert, C. Collard, E. Conte[15], X. Coubez, F. Drouhin[15], J.-C. Fontaine[15], D. Gelé, U. Goerlach, M. Jansová, P. Juillot, A.-C. Le Bihan, N. Tonon, P. Van Hove

Centre de Calcul de l'Institut National de Physique Nucleaire et de Physique des Particules, CNRS/IN2P3, Villeurbanne, France
S. Gadrat

Université de Lyon, Université Claude Bernard Lyon 1, CNRS-IN2P3, Institut de Physique Nucléaire de Lyon, Villeurbanne, France
S. Beauceron, C. Bernet, G. Boudoul, N. Chanon, R. Chierici, D. Contardo, P. Depasse,

H. El Mamouni, J. Fay, L. Finco, S. Gascon, M. Gouzevitch, G. Grenier, B. Ille, F. Lagarde, I.B. Laktineh, H. Lattaud, M. Lethuillier, L. Mirabito, A.L. Pequegnot, S. Perries, A. Popov[16], V. Sordini, M. Vander Donckt, S. Viret, S. Zhang

Georgian Technical University, Tbilisi, Georgia
T. Toriashvili[17]

Tbilisi State University, Tbilisi, Georgia
Z. Tsamalaidze[8]

RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany
C. Autermann, L. Feld, M.K. Kiesel, K. Klein, M. Lipinski, M. Preuten, C. Schomakers, J. Schulz, M. Teroerde, B. Wittmer, V. Zhukov[16]

RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
A. Albert, D. Duchardt, M. Endres, M. Erdmann, S. Erdweg, T. Esch, R. Fischer, A. Güth, T. Hebbeker, C. Heidemann, K. Hoepfner, S. Knutzen, M. Merschmeyer, A. Meyer, P. Millet, S. Mukherjee, T. Pook, T. Quast[13], M. Radziej, H. Reithler, M. Rieger, F. Scheuch, D. Teyssier, S. Thüer

RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany
G. Flügge, B. Kargoll, T. Kress, A. Künsken, T. Müller, A. Nehrkorn, A. Nowack, C. Pistone, O. Pooth, A. Stahl[13]

Deutsches Elektronen-Synchrotron, Hamburg, Germany
M. Aldaya Martin, T. Arndt, C. Asawatangtrakuldee, K. Beernaert, O. Behnke, U. Behrens, A. Bermúdez Martínez, A.A. Bin Anuar, K. Borras[18], V. Botta, A. Campbell, P. Connor, C. Contreras-Campana, F. Costanza, A. De Wit, C. Diez Pardos, G. Eckerlin, D. Eckstein, T. Eichhorn, E. Eren, E. Gallo[19], J. Garay Garcia, A. Geiser, J.M. Grados Luyando, A. Grohsjean, P. Gunnellini, M. Guthoff, A. Harb, J. Hauk, M. Hempel[20], H. Jung, M. Kasemann, J. Keaveney, C. Kleinwort, I. Korol, D. Krücker, K. Krüger, W. Lange, A. Lelek, T. Lenz, K. Lipka, W. Lohmann[20], R. Mankel, I.-A. Melzer-Pellmann, A.B. Meyer, M. Meyer, M. Missiroli, G. Mittag, J. Mnich, A. Mussgiller, D. Pitzl, A. Raspereza, M. Reinecke, M. Savitskyi, P. Saxena, F. Sefkow, R. Shevchenko, N. Stefaniuk, H. Tholen, G.P. Van Onsem, R. Walsh, Y. Wen, K. Wichmann, C. Wissing, O. Zenaiev

University of Hamburg, Hamburg, Germany
R. Aggleton, S. Bein, V. Blobel, M. Centis Vignali, T. Dreyer, E. Garutti, D. Gonzalez, J. Haller, A. Hinzmann, M. Hoffmann, A. Karavdina, G. Kasieczka, R. Klanner, R. Kogler, N. Kovalchuk, S. Kurz, D. Marconi, J. Multhaup, M. Niedziela, D. Nowatschin, T. Peiffer, A. Perieanu, A. Reimers, C. Scharf, P. Schleper, A. Schmidt, S. Schumann, J. Schwandt, J. Sonneveld, H. Stadie, G. Steinbrück, F.M. Stober, M. Stöver, D. Troendle, E. Usai, A. Vanhoefer, B. Vormwald

Institut für Experimentelle Teilchenphysik, Karlsruhe, Germany
M. Akbiyik, C. Barth, M. Baselga, S. Baur, E. Butz, R. Caspart, T. Chwalek, F. Colombo, W. De Boer, A. Dierlamm, N. Faltermann, B. Freund, R. Friese, M. Giffels, M.A. Harrendorf, F. Hartmann[13], S.M. Heindl, U. Husemann, F. Kassel[13], S. Kudella, S. Mallows, H. Mildner, M.U. Mozer, Th. Müller, M. Plagge, G. Quast, K. Rabbertz, M. Schröder, I. Shvetsov, G. Sieber, H.J. Simonis, R. Ulrich, S. Wayand, M. Weber, T. Weiler, S. Williamson, C. Wöhrmann, R. Wolf

Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
G. Anagnostou, G. Daskalakis, T. Geralis, A. Kyriakis, D. Loukas, I. Topsis-Giotis

National and Kapodistrian University of Athens, Athens, Greece
G. Karathanasis, S. Kesisoglou, A. Panagiotou, N. Saoulidou, E. Tziaferi

National Technical University of Athens, Athens, Greece
K. Kousouris, I. Papakrivopoulos, Y. Tsipolitis

University of Ioánnina, Ioánnina, Greece
I. Evangelou, C. Foudas, P. Gianneios, P. Katsoulis, P. Kokkas, S. Mallios, N. Manthos, I. Papadopoulos, E. Paradas, J. Strologas, F.A. Triantis, D. Tsitsonis

MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
M. Csanad, N. Filipovic, G. Pasztor, O. Surányi, G.I. Veres[21]

Wigner Research Centre for Physics, Budapest, Hungary
G. Bencze, C. Hajdu, D. Horvath[22], Á. Hunyadi, F. Sikler, T.Á. Vámi, V. Veszpremi, G. Vesztergombi[21]

Institute of Nuclear Research ATOMKI, Debrecen, Hungary
N. Beni, S. Czellar, J. Karancsi[23], A. Makovec, J. Molnar, Z. Szillasi

Institute of Physics, University of Debrecen, Debrecen, Hungary
M. Bartók[21], P. Raics, B. Ujvari, G. Zilizi

Indian Institute of Science (IISc), Bangalore, India
S. Choudhury, J.R. Komaragiri

National Institute of Science Education and Research, Bhubaneswar, India
S. Bahinipati[24], P. Mal, K. Mandal, A. Nayak[25], D.K. Sahoo[24], N. Sahoo, S.K. Swain

Panjab University, Chandigarh, India
S. Bansal, S.B. Beri, V. Bhatnagar, S. Chauhan, R. Chawla, N. Dhingra, R. Gupta, A. Kaur, M. Kaur, S. Kaur, R. Kumar, P. Kumari, A. Mehta, S. Sharma, J.B. Singh, G. Walia

University of Delhi, Delhi, India
A. Bhardwaj, B.C. Choudhary, R.B. Garg, S. Keshri, A. Kumar, Ashok Kumar, S. Malhotra, M. Naimuddin, K. Ranjan, Aashaq Shah, R. Sharma

Saha Institute of Nuclear Physics, HBNI, Kolkata, India
R. Bhardwaj[26], R. Bhattacharya, S. Bhattacharya, U. Bhawandeep[26], D. Bhowmik, S. Dey, S. Dutt[26], S. Dutta, S. Ghosh, N. Majumdar, K. Mondal, S. Mukhopadhyay, S. Nandan, A. Purohit, P.K. Rout, A. Roy, S. Roy Chowdhury, S. Sarkar, M. Sharan, B. Singh, S. Thakur[26]

Indian Institute of Technology Madras, Madras, India
P.K. Behera

Bhabha Atomic Research Centre, Mumbai, India
R. Chudasama, D. Dutta, V. Jha, V. Kumar, A.K. Mohanty[13], P.K. Netrakanti, L.M. Pant, P. Sarin[27], R. Sehgal, S.T. Sehgal, P. Shukla, A. Topkar

Tata Institute of Fundamental Research-A, Mumbai, India
T. Aziz, S. Chavan, S. Dugad, G. Ghodke, M.M. Kolwalkar, B. Mahakud, I. Mirza, S. Mitra, G.B. Mohanty, R. Shukla, N. Sur, B. Sutar, R. Thomas

Tata Institute of Fundamental Research-B, Mumbai, India
S. Banerjee, S. Bhattacharya, S. Chatterjee, P. Das, M. Guchait, Sa. Jain, S. Kumar, M. Maity[28], G. Majumder, K. Mazumdar, T. Sarkar[28], N. Wickramage[29]

Indian Institute of Science Education and Research (IISER), Pune, India
S. Chauhan, S. Dube, V. Hegde, A. Kapoor, K. Kothekar, S. Pandey, A. Rane, S. Sharma

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
S. Chenarani[30], E. Eskandari Tadavani, S.M. Etesami[30], M. Khakzad, M. Mohammadi Najafabadi, M. Naseri, S. Paktinat Mehdiabadi[31], F. Rezaei Hosseinabadi, B. Safarzadeh[32], M. Zeinali

University College Dublin, Dublin, Ireland
M. Felcini, M. Grunewald

INFN Sezione di Bari [a], Università di Bari [b], Politecnico di Bari [c], Bari, Italy
M. Abbrescia[a,b], C. Calabria[a,b], A. Colaleo[a], D. Creanza[a,c], L. Cristella[a,b], N. De Filippis[a,c], M. De Palma[a,b], A. Di Florio[a,b], F. Errico[a,b], L. Fiore[a], G. Iaselli[a,c], S. Lezki[a,b], G. Maggi[a,c], M. Maggi[a], B. Marangelli[a,b], G. Miniello[a,b], S. My[a,b], S. Nuzzo[a,b], A. Pompili[a,b], G. Pugliese[a,c], R. Radogna[a], A. Ranieri[a], G. Selvaggi[a,b], A. Sharma[a], L. Silvestris[a,13], R. Venditti[a], P. Verwilligen[a], G. Zito[a]

INFN Sezione di Bologna [a], Università di Bologna [b], Bologna, Italy
G. Abbiendi[a], C. Battilana[a,b], D. Bonacorsi[a,b], L. Borgonovi[a,b], S. Braibant-Giacomelli[a,b], R. Campanini[a,b], P. Capiluppi[a,b], A. Castro[a,b], F.R. Cavallo[a], S.S. Chhibra[a,b], G. Codispoti[a,b], M. Cuffiani[a,b], G.M. Dallavalle[a], F. Fabbri[a], A. Fanfani[a,b], D. Fasanella[a,b], P. Giacomelli[a], C. Grandi[a], L. Guiducci[a,b], F. Iemmi, S. Marcellini[a], G. Masetti[a], A. Montanari[a], F.L. Navarria[a,b], A. Perrotta[a], A.M. Rossi[a,b], T. Rovelli[a,b], G.P. Siroli[a,b], N. Tosi[a]

INFN Sezione di Catania [a], Università di Catania [b], Catania, Italy
S. Albergo[a,b], S. Costa[a,b], A. Di Mattia[a], F. Giordano[a,b], R. Potenza[a,b], A. Tricomi[a,b], C. Tuve[a,b]

INFN Sezione di Firenze [a], Università di Firenze [b], Firenze, Italy
G. Barbagli[a], K. Chatterjee[a,b], V. Ciulli[a,b], C. Civinini[a], R. D'Alessandro[a,b], E. Focardi[a,b], G. Latino, P. Lenzi[a,b], M. Meschini[a], S. Paoletti[a], L. Russo[a,33], G. Sguazzoni[a], D. Strom[a], L. Viliani[a]

INFN Laboratori Nazionali di Frascati, Frascati, Italy
L. Benussi, S. Bianco, M. Caponero[34], F. Fabbri, M. Ferrini, D. Piccolo, F. Primavera[13], G. Saviano[35]

INFN Sezione di Genova [a], Università di Genova [b], Genova, Italy
V. Calvelli[a,b], F. Ferro[a], F. Ravera[a,b], E. Robutti[a], S. Tosi[a,b]

INFN Sezione di Milano-Bicocca [a], Università di Milano-Bicocca [b], Milano, Italy
A. Benaglia[a], A. Beschi[b], L. Brianza[a,b], F. Brivio[a,b], V. Ciriolo[a,b,13], M.E. Dinardo[a,b], S. Fiorendi[a,b], S. Gennai[a], A. Ghezzi[a,b], P. Govoni[a,b], M. Malberti[a,b], S. Malvezzi[a],

R.A. Manzoni[a,b], D. Menasce[a], L. Moroni[a], M. Paganoni[a,b], K. Pauwels[a,b], D. Pedrini[a], S. Pigazzini[a,b,36], S. Ragazzi[a,b], T. Tabarelli de Fatis[a,b]

INFN Sezione di Napoli [a], Università di Napoli 'Federico II' [b], Napoli, Italy, Università della Basilicata [c], Potenza, Italy, Università G. Marconi [d], Roma, Italy
S. Buontempo[a], N. Cavallo[a,c], S. Di Guida[a,d,13], F. Fabozzi[a,c], F. Fienga[a,b], A.O.M. Iorio[a,b], W.A. Khan[a], L. Lista[a], S. Meola[a,d,13], P. Paolucci[a,13], C. Sciacca[a,b], F. Thyssen[a]

INFN Sezione di Padova [a], Università di Padova [b], Padova, Italy, Università di Trento [c], Trento, Italy
P. Azzi[a], N. Bacchetta[a], L. Benato[a,b], A. Boletti[a,b], E. Borsato[a,b], R. Carlin[a,b], A. Carvalho Antunes De Oliveira[a,b], P. Checchia[a], M. Dall'Osso[a,b], P. De Castro Manzano[a], T. Dorigo[a], U. Dosselli[a], F. Gasparini[a,b], U. Gasparini[a,b], S. Lacaprara[a], P. Lujan, M. Margoni[a,b], A.T. Meneguzzo[a,b], N. Pozzobon[a,b], P. Ronchese[a,b], R. Rossin[a,b], F. Simonetto[a,b], A. Tiko, E. Torassa[a], M. Zanetti[a,b], P. Zotto[a,b], G. Zumerle[a,b]

INFN Sezione di Pavia [a], Università di Pavia [b], Pavia, Italy
A. Braghieri[a], A. Magnani[a], P. Montagna[a,b], S.P. Ratti[a,b], V. Re[a], M. Ressegotti[a,b], C. Riccardi[a,b], P. Salvini[a], I. Vai[a,b], P. Vitulo[a,b]

INFN Sezione di Perugia [a], Università di Perugia [b], Perugia, Italy
L. Alunni Solestizi[a,b], M. Biasini[a,b], G.M. Bilei[a], C. Cecchi[a,b], D. Ciangottini[a,b], L. Fanò[a,b], P. Lariccia[a,b], R. Leonardi[a,b], E. Manoni[a], G. Mantovani[a,b], V. Mariani[a,b], M. Menichelli[a], A. Rossi[a,b], A. Santocchia[a,b], D. Spiga[a]

INFN Sezione di Pisa [a], Università di Pisa [b], Scuola Normale Superiore di Pisa [c], Pisa, Italy
K. Androsov[a], P. Azzurri[a,13], G. Bagliesi[a], L. Bianchini[a], T. Boccali[a], L. Borrello, R. Castaldi[a], M.A. Ciocci[a,b], R. Dell'Orso[a], G. Fedi[a], L. Giannini[a,c], A. Giassi[a], M.T. Grippo[a,33], F. Ligabue[a,c], T. Lomtadze[a], E. Manca[a,c], G. Mandorli[a,c], A. Messineo[a,b], F. Palla[a], A. Rizzi[a,b], P. Spagnolo[a], R. Tenchini[a], G. Tonelli[a,b], A. Venturi[a], P.G. Verdini[a]

INFN Sezione di Roma [a], Sapienza Università di Roma [b], Rome, Italy
L. Barone[a,b], F. Cavallari[a], M. Cipriani[a,b], N. Daci[a], D. Del Re[a,b], E. Di Marco[a,b], M. Diemoz[a], S. Gelli[a,b], E. Longo[a,b], F. Margaroli[a,b], B. Marzocchi[a,b], P. Meridiani[a], G. Organtini[a,b], F. Pandolfi[a], R. Paramatti[a,b], F. Preiato[a,b], S. Rahatlou[a,b], C. Rovelli[a], F. Santanastasio[a,b]

INFN Sezione di Torino [a], Università di Torino [b], Torino, Italy, Università del Piemonte Orientale [c], Novara, Italy
N. Amapane[a,b], R. Arcidiacono[a,c], S. Argiro[a,b], M. Arneodo[a,c], N. Bartosik[a], R. Bellan[a,b], C. Biino[a], N. Cartiglia[a], R. Castello[a,b], F. Cenna[a,b], M. Costa[a,b], R. Covarelli[a,b], A. Degano[a,b], N. Demaria[a], B. Kiani[a,b], C. Mariotti[a], S. Maselli[a], E. Migliore[a,b], V. Monaco[a,b], E. Monteil[a,b], M. Monteno[a], M.M. Obertino[a,b], L. Pacher[a,b], N. Pastrone[a], M. Pelliccioni[a], G.L. Pinna Angioni[a,b], A. Romero[a,b], M. Ruspa[a,c], R. Sacchi[a,b], K. Shchelina[a,b], V. Sola[a], A. Solano[a,b], A. Staiano[a], P. Traczyk[a,b]

INFN Sezione di Trieste [a], Università di Trieste [b], Trieste, Italy
S. Belforte[a], M. Casarsa[a], F. Cossutti[a], G. Della Ricca[a,b], A. Zanetti[a]

Kyungpook National University
D.H. Kim, G.N. Kim, M.S. Kim, J. Lee, S. Lee, S.W. Lee, C.S. Moon, Y.D. Oh, S. Sekmen, D.C. Son, Y.C. Yang

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Korea
H. Kim, D.H. Moon, G. Oh

Hanyang University, Seoul, Korea
J.A. Brochero Cifuentes, J. Goh, T.J. Kim

Korea University, Seoul, Korea
S. Cho, S. Choi, Y. Go, D. Gyun, S. Ha, B. Hong, Y. Jo, Y. Kim, K. Lee, K.S. Lee, S. Lee, J. Lim, S.K. Park, Y. Roh

Seoul National University, Seoul, Korea
J. Almond, J. Kim, J.S. Kim, H. Lee, K. Lee, K. Nam, S.B. Oh, B.C. Radburn-Smith, S.h. Seo, U.K. Yang, H.D. Yoo, G.B. Yu

University of Seoul, Seoul, Korea
H. Kim, J.H. Kim, J.S.H. Lee, I.C. Park

Sungkyunkwan University, Suwon, Korea
Y. Choi, C. Hwang, J. Lee, I. Yu

Vilnius University, Vilnius, Lithuania
V. Dudenas, A. Juodagalvis, J. Vaitkus

National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia
I. Ahmed, Z.A. Ibrahim, M.A.B. Md Ali[37], F. Mohamad Idris[38], W.A.T. Wan Abdullah, M.N. Yusli, Z. Zolkapli

Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico
Duran-Osuna, M. C., H. Castilla-Valdez, E. De La Cruz-Burelo, Ramirez-Sanchez, G., I. Heredia-De La Cruz[39], Rabadan-Trejo, R. I., R. Lopez-Fernandez, J. Mejia Guisao, Reyes-Almanza, R, A. Sanchez-Hernandez

Universidad Iberoamericana, Mexico City, Mexico
S. Carrillo Moreno, C. Oropeza Barrera, F. Vazquez Valencia

Benemerita Universidad Autonoma de Puebla, Puebla, Mexico
J. Eysermans, I. Pedraza, H.A. Salazar Ibarguen, C. Uribe Estrada

Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico
A. Morelos Pineda

University of Auckland, Auckland, New Zealand
D. Krofcheck

University of Canterbury, Christchurch, New Zealand
S. Bheesette, P.H. Butler

National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan
A. Ahmad, M. Ahmad, Q. Hassan, H.R. Hoorani, A. Saddique, M.A. Shah, M. Shoaib, M. Waqas

National Centre for Nuclear Research, Swierk, Poland
H. Bialkowska, M. Bluj, B. Boimska, T. Frueboes, M. Górski, M. Kazana, K. Nawrocki, M. Szleper, P. Zalewski

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland
K. Bunkowski, A. Byszuk[40], K. Doroba, A. Kalinowski, M. Konecki, J. Krolikowski, M. Misiura, M. Olszewski, A. Pyskir, M. Walczak

Laboratório de Instrumentação e Física Experimental de Partículas, Lisboa, Portugal
P. Bargassa, C. Beirão Da Cruz E Silva, A. Di Francesco, P. Faccioli, B. Galinhas, M. Gallinaro, J. Hollar, N. Leonardo, L. Lloret Iglesias, M.V. Nemallapudi, J.C. Rasteiro Da Silva, J. Seixas, J. Silva, G. Strong, O. Toldaiev, D. Vadruccio, J. Varela

Joint Institute for Nuclear Research, Dubna, Russia
S. Afanasiev, V. Alexakhin, P. Bunin, Y. Ershov, M. Gavrilenko, I. Golutvin, I. Gorbunov, A. Kamenev, V. Karjavin, A. Lanev, A. Malakhov, V. Matveev[41,42], P. Moisenz, V. Palichik, V. Perelygin, S. Shmatov, S. Shulha, N. Skatchkov, V. Smirnov, N. Voytishin, B.S. Yuldashev[43], N. Zamiatin, A. Zarubin

Petersburg Nuclear Physics Institute, Gatchina (St. Petersburg), Russia
Y. Ivanov, V. Kim[44], E. Kuznetsova[45], P. Levchenko, V. Murzin, V. Oreshkin, I. Smirnov, D. Sosnov, V. Sulimov, L. Uvarov, S. Vavilov, A. Vorobyev

Institute for Nuclear Research, Moscow, Russia
Yu. Andreev, A. Dermenev, S. Gninenko, N. Golubev, A. Karneyeu, M. Kirsanov, N. Krasnikov, A. Pashenkov, D. Tlisov, A. Toropin

Institute for Theoretical and Experimental Physics, Moscow, Russia
V. Epshteyn, V. Gavrilov, N. Lychkovskaya, V. Popov, I. Pozdnyakov, G. Safronov, A. Spiridonov, A. Stepennov, V. Stolin, M. Toms, E. Vlasov, A. Zhokin

Moscow Institute of Physics and Technology, Moscow, Russia
T. Aushev, A. Bylinkin[42]

National Research Nuclear University 'Moscow Engineering Physics Institute' (MEPhI), Moscow, Russia
O. Bychkova, M. Danilov[46], A. Kolchanova, P. Parygin, D. Philippov, E. Popova, V. Rusinov, E. Tarkovskii, S. Vinogradov

P.N. Lebedev Physical Institute, Moscow, Russia
V. Andreev, M. Azarkin[42], I. Dremin[42], M. Kirakosyan[42], S.V. Rusakov, A. Terkulov

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
A. Baskakov, A. Belyaev, E. Boos, A. Demiyanov, L. Dudko, A. Ershov, A. Gribushin, A. Kaminskiy[47], V. Klyukhin, O. Kodolova, I. Lokhtin, I. Miagkov, S. Obraztsov, S. Petrushanko, V. Savrin

Novosibirsk State University (NSU), Novosibirsk, Russia
V. Blinov[48], D. Shtol[48], Y. Skovpen[48]

State Research Center of Russian Federation, Institute for High Energy Physics of NRC, Kurchatov Institute, , Protvino, Russia
I. Azhgirey, I. Bayshev, S. Bitioukov, A. Bordanovskiy, G. Britvich, S. Chernichenko, D. Elumakhov, V. Ferapontov, A. Godizov, V. Kachanov, A. Kalinin, D. Konstantinov, M. Kostin, I. Kurochkin, E. Kvashina, A. Levine, P. Mandrik, E. Paramoshkina, V. Petrov, A. Ryabov, R. Ryutin, A. Shalimov, I. Shein, Z. Simonova, N. Skvorodnev, A. Sobol, A. Soukhih, S. Troshin, N. Tyurin, A. Uzunian, A. Volkov, Y. Yabovich

National Research Tomsk Polytechnic University, Tomsk, Russia
A. Babaev

University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
P. Adzic[49], P. Cirkovic, D. Devetak, M. Dordevic, J. Milosevic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
J. Alcaraz Maestre, A. Álvarez Fernández, I. Bachiller, M. Barrio Luna, M. Cerrada, N. Colino, B. De La Cruz, A. Delgado Peris, C. Fernandez Bedoya, J.P. Fernández Ramos, J. Flix, M.C. Fouz, O. Gonzalez Lopez, S. Goy Lopez, J.M. Hernandez, M.I. Josa, D. Moran, A. Pérez-Calero Yzquierdo, J. Puerta Pelayo, I. Redondo, L. Romero, M.S. Soares, A. Triossi

Universidad Autónoma de Madrid, Madrid, Spain
C. Albajar, J.F. de Trocóniz

Universidad de Oviedo, Oviedo, Spain
J. Cuevas, C. Erice, J. Fernandez Menendez, S. Folgueras, I. Gonzalez Caballero, J.R. González Fernández, E. Palencia Cortezon, S. Sanchez Cruz, P. Vischia, J.M. Vizan Garcia

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain
I.J. Cabrillo, A. Calderon, B. Chazin Quero, J. Duarte Campderros, M. Fernandez, P.J. Fernández Manteca, A. García Alonso, J. Garcia-Ferrero, G. Gomez, A. Lopez Virto, J. Marco, C. Martinez Rivero, P. Martinez Ruiz del Arbol, F. Matorras, J. Piedra Gomez, C. Prieels, T. Rodrigo, A. Ruiz-Jimeno, L. Scodellaro, N. Trevisani, I. Vila, R. Vilar Cortabitarte

CERN, European Organization for Nuclear Research, Geneva, Switzerland
D. Abbaneo, B. Akgun, E. Albert, E. Auffray, P. Baillon, A.H. Ball, D. Barney, M. Baschiera, M. Beguin, J. Bendavid, M. Bianco, A. Bocci, G. Bombardi, C. Botta, S. Brachet, T. Camporesi, M. Cepeda, G. Cerminara, E. Chapon, Y. Chen, A. Conde Garcia, D. d'Enterria, A. Dabrowski, J. Daguin, D. Dannheim, V. Daponte, A. David, M. De Gruttola, A. De Roeck, N. Deelen, D. Deyrail, M. Dobson, T. du Pree, M. Dünser, N. Dupont, A. Elliott-Peisert, K. Elsener, P. Everaerts, F. Fallavollita[50], G. Franzoni, J. Fulcher, W. Funk, H. Gerwig, D. Gigi, A. Gilbert, K. Gill, F. Glege, R. Guida, D. Gulhan, M. Hansen, J. Hegeman, V. Innocente, A. Jafari, P. Janot, O. Karacheban[20], J. Kieseler, V. Knünz, A. Kornmayer, M. Krammer[1], S. Kulis, C. Lange, P. Lecoq, L. Linssen, C. Lourenço, M.T. Lucchini, A.A. Maier, L. Malgeri, M. Mannelli, A. Marchioro, A. Martelli, F. Meijers, J.A. Merlin, S. Mersi, E. Meschi, P. Milenovic[51], S. Moccia, M. Moll, F. Moortgat, M. Mulders, H. Neugebauer, J. Ngadiuba, M. Noy, S. Orfanelli, L. Orsini, R. Ostojic, F. Pantaleo[13],

L. Pape, E. Perez, J.F. Pernot, M. Peruzzi, P. Petagna, P. Petiot, A. Petrilli, G. Petrucciani, A. Pfeiffer, M. Pierini, F.M. Pitters, H. Postema, D. Rabady, A. Racz, K. Rapacz, P. Reguera Duran, T. Reis, G. Rolandi[52], M. Rovere, H. Sakulin, H. Sauce, C. Schäfer, C. Schwick, M. Seidel, M. Selvaggi, A. Sharma, E. Sicking, N. Siegrist, P. Silva, P. Sphicas[53], A. Stakia, J. Steggemann, M. Stoye, M. Tosi, D. Treille, P. Tropea, A. Tsirou, V. Veckalns[54], T. Vergine, M. Verweij, P. Wertelaers, W.D. Zeuner

Paul Scherrer Institut, Villigen, Switzerland
W. Bertl†, L. Caminada[55], K. Deiters, W. Erdmann, R. Horisberger, Q. Ingram, H.C. Kaestli, D. Kotlinski, U. Langenegger, T. Rohe, S.A. Wiederkehr

ETH Zurich - Institute for Particle Physics and Astrophysics (IPA), Zurich, Switzerland
M. Backhaus, L. Bäni, P. Berger, B. Casal, G. Dissertori, M. Dittmar, M. Donegà, C. Dorfer, C. Grab, C. Heidegger, D. Hits, J. Hoss, T. Klijnsma, W. Lustermann, M. Marionneau, M.T. Meinhard, D. Meister, F. Micheli, P. Musella, F. Nessi-Tedaldi, J. Pata, F. Pauss, G. Perrin, L. Perrozzi, M. Quittnat, M. Reichmann, D.A. Sanz Becerra, M. Schönenberger, L. Shchutska, V.R. Tavolaro, K. Theofilatos, M.L. Vesterbacka Olsson, R. Wallny, D.H. Zhu

Universität Zürich, Zurich, Switzerland
T.K. Aarrestad, C. Amsler[56], D. Brzhechko, M.F. Canelli, A. De Cosa, R. Del Burgo, S. Donato, C. Galloni, T. Hreus, B. Kilminster, I. Neutelings, D. Pinna, G. Rauco, P. Robmann, D. Salerno, K. Schweiger, C. Seitz, Y. Takahashi, A. Zucchetta

National Central University, Chung-Li, Taiwan
V. Candelise, Y.H. Chang, K.y. Cheng, T.H. Doan, Sh. Jain, R. Khurana, C.M. Kuo, W. Lin, A. Pozdnyakov, S.S. Yu

National Taiwan University (NTU), Taipei, Taiwan
P. Chang, Y. Chao, K.F. Chen, P.H. Chen, F. Fiori, W.-S. Hou, Y. Hsiung, Arun Kumar, Y.F. Liu, R.-S. Lu, E. Paganis, A. Psallidas, A. Steen, J.f. Tsai

Chulalongkorn University, Faculty of Science, Department of Physics, Bangkok, Thailand
B. Asavapibhop, K. Kovitanggoon, G. Singh, N. Srimanobhas, N. Suwonjandee

Çukurova University, Physics Department, Science and Art Faculty, Adana, Turkey
A. Bat, F. Boran, S. Cerci[57], S. Damarseckin, Z.S. Demiroglu, F. Dolek, C. Dozen, I. Dumanoglu, S. Girgis, G. Gokbulut, Y. Guler, I. Hos[58], E.E. Kangal[59], O. Kara, A. Kayis Topaksu, U. Kiminsu, M. Oglakci, G. Onengut, K. Ozdemir[60], D. Sunar Cerci[57], B. Tali[57], U.G. Tok, S. Turkcapar, I.S. Zorbakir, C. Zorbilmez

Middle East Technical University, Physics Department, Ankara, Turkey
G. Karapinar[61], K. Ocalan[62], M. Yalvac, M. Zeyrek

Bogazici University, Istanbul, Turkey
E. Gülmez, M. Kaya[63], O. Kaya[64], K. Nurdan, S. Ozkorucuklu[65], S. Tekten, E.A. Yetkin[66], T. Yetkin[67]

Istanbul Technical University, Istanbul, Turkey
M.N. Agaras, S. Atay, A. Cakir, K. Cankocak, Y. Komurcu, S. Sen[68]

Institute for Scintillation Materials of National Academy of Science of Ukraine, Kharkov, Ukraine
B. Grynyov

National Scientific Center, Kharkov Institute of Physics and Technology, Kharkov, Ukraine
L. Levchuk, S. Lukyanenko, V. Popov

University of Bristol, Bristol, United Kingdom
F. Ball, L. Beck, J.J. Brooke, D. Burns, E. Clement, D. Cussans, O. Davignon, H. Flacher, J. Goldstein, G.P. Heath, H.F. Heath, L. Kreczko, D.M. Newbold[69], S. Paramesvaran, T. Sakuma, S. Seif El Nasr-storey, D. Smith, V.J. Smith

Rutherford Appleton Laboratory, Didcot, United Kingdom
K.W. Bell, A. Belyaev[70], C. Brew, R.M. Brown, L. Calligaris, D. Cieri, D.J.A. Cockerill, J.A. Coughlan, K. Harder, S. Harper, J. Linacre, E. Olaiya, D. Petyt, C.H. Shepherd-Themistocleous, A. Thea, I.R. Tomalin, T. Williams, W.J. Womersley

Imperial College, London, United Kingdom
G. Auzinger, R. Bainbridge, P. Bloch, S. Bonomally, J. Borg, S. Breeze, O. Buchmuller, A. Bundock, S. Casasso, D. Colling, L. Corpe, P. Dauncey, G. Davies, M. Della Negra, R. Di Maria, A. Elwood, Y. Haddad, G. Hall, G. Iles, T. James, M. Komm, R. Lane, C. Laner, J. Langford, L. Lyons, A.-M. Magnan, S. Malik, L. Mastrolorenzo, T. Matsushita, J. Nash[71], A. Nikitenko[7], V. Palladino, D. Parker, M. Pesaresi, A. Richards, A. Rose, E. Scott, C. Seez, A. Shtipliyski, T. Strebler, S. Summers, A. Tapper, K. Uchida, M. Vazquez Acosta[72], T. Virdee[13], N. Wardle, D. Winterbottom, J. Wright, S.C. Zenz

Brunel University, Uxbridge, United Kingdom
J.E. Cole, P.R. Hobson, A. Khan, P. Kyberd, A. Morton, I.D. Reid, L. Teodorescu, S. Zahid

Baylor University, Waco, USA
A. Borzou, K. Call, J. Dittmann, K. Hatakeyama, H. Liu, N. Pastika, C. Smith

Catholic University of America, Washington DC, USA
R. Bartek, A. Dominguez

The University of Alabama, Tuscaloosa, USA
A. Buccilli, S.I. Cooper, C. Henderson, P. Rumerio, C. West

Boston University, Boston, USA
D. Arcaro, A. Avetisyan, T. Bose, D. Gastler, E. Hazen, D. Rankin, C. Richardson, J. Rohlf, L. Sulak, S. Wu, D. Zou

Brown University, Providence, USA
G. Benelli, B. Burkle, D. Cutts, S. Ghosh, M. Hadley, J. Hakala, U. Heintz, N. Hinton, J.M. Hogan[73], K.H.M. Kwok, E. Laird, G. Landsberg, J. Lee, Z. Mao, M. Narain, J. Pazzini, S. Piperov, S. Sagir, E. Spencer, R. Syarif, D. Yu

University of California, Davis, Davis, USA
R. Band, C. Brainerd, R. Breedon, D. Burns, M. Calderon De La Barca Sanchez, M. Chertok, J. Conway, R. Conway, P.T. Cox, R. Erbacher, C. Flores, G. Funk, W. Ko, R. Lander,

C. Mclean, M. Mulhearn, D. Pellett, J. Pilot, S. Shalhout, M. Shi, J. Smith, D. Stolp, D. Taylor, K. Tos, M. Tripathi, Z. Wang, F. Zhang

University of California, Los Angeles, USA
M. Bachtis, C. Bravo, R. Cousins, A. Dasgupta, A. Florent, J. Hauser, M. Ignatenko, N. Mccoll, S. Regnard, D. Saltzberg, C. Schnaible, V. Valuev

University of California, Riverside, Riverside, USA
E. Bouvier, K. Burt, R. Clare, J. Ellison, J.W. Gary, S.M.A. Ghiasi Shirazi, G. Hanson, G. Karapostoli, E. Kennedy, F. Lacroix, O.R. Long, M. Olmedo Negrete, M.I. Paneva, W. Si, L. Wang, H. Wei, S. Wimpenny, B. R. Yates

University of California, San Diego, La Jolla, USA
J.G. Branson, S. Cittolin, M. Derdzinski, R. Gerosa, D. Gilbert, B. Hashemi, A. Holzner, D. Klein, G. Kole, V. Krutelyov, J. Letts, M. Masciovecchio, D. Olivito, S. Padhi, M. Pieri, M. Sani, V. Sharma, S. Simon, M. Tadel, A. Vartak, S. Wasserbaech[74], J. Wood, F. Würthwein, A. Yagil, G. Zevi Della Porta

University of California, Santa Barbara - Department of Physics, Santa Barbara, USA
N. Amin, R. Bhandari, J. Bradmiller-Feld, C. Campagnari, S. Carron Montero[75], M. Citron, G. Collura, A. Dishaw, V. Dutta, J. Eckdahl, M. Franco Sevilla, L. Gouskos, R. Heller, J. Incandela, S. Kyre, B. Odegard, A. Ovcharova, A. Patterson, H. Qu, J. Richman, D. Stuart, I. Suarez, K. Valpey-Garcia, D. White, J. Yoo

California Institute of Technology, Pasadena, USA
D. Anderson, A. Bornheim, J. Bunn, J.M. Lawhorn, H.B. Newman, T. Q. Nguyen, C. Pena, M. Spiropulu, J.R. Vlimant, R. Wilkinson, S. Xie, Z. Zhang, R.Y. Zhu

Carnegie Mellon University, Pittsburgh, USA
M.B. Andrews, E. Day, T. Ferguson, T. Mudholkar, M. Paulini, J. Russ, M. Sun, H. Vogel, I. Vorobiev, M. Weinberg

University of Colorado Boulder, Boulder, USA
J.P. Cumalat, W.T. Ford, F. Jensen, A. Johnson, M. Krohn, S. Leontsinis, E. MacDonald, T. Mulholland, K. Stenson, K.A. Ulmer, S.R. Wagner

Cornell University, Ithaca, USA
J. Alexander, J. Chaves, Y. Cheng, J. Chu, A. Datta, K. Mcdermott, N. Mirman, J.R. Patterson, D. Quach, A. Rinkevicius, A. Ryd, L. Skinnari, L. Soffi, S.M. Tan, Z. Tao, J. Thom, J. Tucker, P. Wittich, M. Zientek

Fermi National Accelerator Laboratory, Batavia, USA
S. Abdullin, M. Albrow, M. Alyari, G. Apollinari, A. Apresyan, A. Apyan, S. Banerjee, L.A.T. Bauerdick, A. Beretvas, J. Berryhill, P.C. Bhat, G. Bolla[†], K. Burkett, J.N. Butler, A. Canepa, G.B. Cerati, H.W.K. Cheung, F. Chlebana, M. Cremonesi, J. Duarte, V.D. Elvira, J. Freeman, Z. Gecse, E. Gottschalk, L. Gray, D. Green, S. Grünendahl, O. Gutsche, J. Hanlon, R.M. Harris, S. Hasegawa, J. Hirschauer, Z. Hu, B. Jayatilaka, S. Jindariani, M. Johnson, U. Joshi, B. Klima, M.J. Kortelainen, B. Kreis, S. Lammel, D. Lincoln, R. Lipton, M. Liu, T. Liu, R. Lopes De Sá, J. Lykken, K. Maeshima, N. Magini, J.M. Marraffino, D. Mason, P. McBride, P. Merkel, S. Mrenna, S. Nahn, V. O'Dell, K. Pedro, O. Prokofyev, G. Rakness, L. Ristori, A. Savoy-Navarro[76], B. Schneider, E. Sexton-Kennedy, A. Soha, W.J. Spalding, L. Spiegel, S. Stoynev, J. Strait, N. Strobbe,

L. Taylor, S. Tkaczyk, N.V. Tran, L. Uplegger, E.W. Vaandering, C. Vernieri, M. Verzocchi, R. Vidal, M. Wang, H.A. Weber, A. Whitbeck, W. Wu

University of Florida, Gainesville, USA
D. Acosta, P. Avery, P. Bortignon, D. Bourilkov, A. Brinkerhoff, A. Carnes, M. Carver, D. Curry, R.D. Field, I.K. Furic, S.V. Gleyzer, B.M. Joshi, J. Konigsberg, A. Korytov, K. Kotov, P. Ma, K. Matchev, H. Mei, G. Mitselmakher, K. Shi, D. Sperka, N. Terentyev, L. Thomas, J. Wang, S. Wang, J. Yelton

Florida International University, Miami, USA
Y.R. Joshi, S. Linn, P. Markowitz, J.L. Rodriguez

Florida State University, Tallahassee, USA
A. Ackert, T. Adams, A. Askew, S. Hagopian, V. Hagopian, K.F. Johnson, T. Kolberg, G. Martinez, T. Perry, H. Prosper, A. Saha, A. Santra, V. Sharma, R. Yohay

Florida Institute of Technology, Melbourne, USA
M.M. Baarmand, V. Bhopatkar, S. Colafranceschi, M. Hohlmann, D. Noonan, T. Roy, F. Yumiceva

University of Illinois at Chicago (UIC), Chicago, USA
M.R. Adams, L. Apanasevich, D. Berry, R.R. Betts, R. Cavanaugh, X. Chen, S. Dittmer, O. Evdokimov, C.E. Gerber, D.A. Hangal, D.J. Hofman, K. Jung, J. Kamin, I.D. Sandoval Gonzalez, M.B. Tonjes, N. Varelas, H. Wang, Z. Wu, J. Zhang

The University of Iowa, Iowa City, USA
M. Alhusseini, B. Bilki[77], W. Clarida, P. Debbins, K. Dilsiz[78], R. Dong, S. Durgut, L. Emediato, R.P. Gandrajula, M. Haytmyradov, V. Khristenko, J.-P. Merlo, H. Mermerkaya[79], A. Mestvirishvili, A. Moeller, J. Nachtman, H. Ogul[80], Y. Onel, F. Ozok[81], A. Penzo, R. Rahmat, I. Schmidt, J. Schnell, C. Snyder, D. Southwick, E. Tiras, J. Wetzel, K. Yi

Johns Hopkins University, Baltimore, USA
B. Blumenfeld, A. Cocoros, N. Eminizer, D. Fehling, L. Feng, A.V. Gritsan, P. Maksimovic, J. Roskes, U. Sarica, M. Swartz, M. Xiao, C. You

The University of Kansas, Lawrence, USA
A. Al-bataineh, P. Baringer, A. Bean, S. Boren, J. Bowen, J. Castle, S. Khalil, A. Kropivnitskaya, D. Majumder, W. Mcbrayer, M. Murray, C. Rogan, C. Royon, S. Sanders, E. Schmitz, J.D. Tapia Takaki, Q. Wang

Kansas State University, Manhattan, USA
S. Corkill, A. Ivanov, K. Kaadze, Y. Maravin, A. Modak, A. Mohammadi, L.K. Saini, N. Skhirtladze

Lawrence Livermore National Laboratory, Livermore, USA
F. Rebassoo, D. Wright

University of Maryland, College Park, USA
A. Baden, O. Baron, A. Belloni, S.C. Eno, Y. Feng, C. Ferraioli, N.J. Hadley, S. Jabeen, G.Y. Jeng, R.G. Kellogg, J. Kunkle, A.C. Mignerey, F. Ricci-Tam, Y.H. Shin, A. Skuja, S.C. Tonwar

Massachusetts Institute of Technology, Cambridge, USA
D. Abercrombie, B. Allen, V. Azzolini, R. Barbieri, A. Baty, G. Bauer, R. Bi, S. Brandt,

W. Busza, I.A. Cali, M. D'Alfonso, Z. Demiragli, G. Gomez Ceballos, M. Goncharov, P. Harris, D. Hsu, M. Hu, Y. Iiyama, G.M. Innocenti, M. Klute, D. Kovalskyi, Y.-J. Lee, A. Levin, P.D. Luckey, B. Maier, A.C. Marini, C. Mcginn, C. Mironov, S. Narayanan, X. Niu, C. Paus, C. Roland, G. Roland, G.S.F. Stephans, K. Sumorok, K. Tatar, D. Velicanu, J. Wang, T.W. Wang, B. Wyslouch, S. Zhaozhong

University of Minnesota, Minneapolis, USA
A.C. Benvenuti, R.M. Chatterjee, A. Evans, E. Frahm, W.J. Gilbert, P. Hansen, S. Kalafut, Y. Kubota, Z. Lesko, J. Mans, S. Nourbakhsh, R. Quinn, N. Ruckstuhl, R. Rusack, J. Turkewitz, M.A. Wadud

University of Mississippi, Oxford, USA
J.G. Acosta, S. Oliveros

University of Nebraska-Lincoln, Lincoln, USA
E. Avdeeva, K. Bloom, D.R. Claes, C. Fangmeier, F. Golf, R. Gonzalez Suarez, R. Kamalieddin, I. Kravchenko, J. Monroy, J.E. Siado, G.R. Snow, B. Stieger

State University of New York at Buffalo, Buffalo, USA
J. Dolen, A. Godshalk, C. Harrington, I. Iashvili, D. Nguyen, A. Parker, S. Rappoccio, B. Roozbahani

Northeastern University, Boston, USA
G. Alverson, E. Barberis, C. Freer, A. Hortiangtham, A. Massironi, D.M. Morse, T. Orimoto, R. Teixeira De Lima, T. Wamorkar, B. Wang, A. Wisecarver, D. Wood

Northwestern University, Evanston, USA
S. Bhattacharya, J. Bueghly, O. Charaf, Z. Chen, K.A. Hahn, N. Mucia, N. Odell, M.H. Schmitt, K. Sung, M. Trovato, M. Velasco

University of Notre Dame, Notre Dame, USA
R. Bucci, N. Dev, A.H. Heering, M. Hildreth, K. Hurtado Anampa, C. Jessop, D.J. Karmgard, N. Kellams, K. Lannon, W. Li, N. Loukas, N. Marinelli, M. McKenna, F. Meng, C. Mueller, Y. Musienko[41], M. Planer, A. Reinsvold, R. Ruchti, D. Ruggiero, P. Siddireddy, G. Smith, S. Taroni, M. Wayne, A. Wightman, M. Wolf, A. Woodard

The Ohio State University, Columbus, USA
J. Alimena, L. Antonelli, B. Bylsma, L.S. Durkin, S. Flowers, B. Francis, A. Hart, C. Hill, W. Ji, T.Y. Ling, W. Luo, B.L. Winer, H.W. Wulsin

Princeton University, Princeton, USA
S. Cooperstein, O. Driga, P. Elmer, J. Hardenbrook, P. Hebda, S. Higginbotham, A. Kalogeropoulos, D. Lange, J. Luo, D. Marlow, K. Mei, I. Ojalvo, J. Olsen, C. Palmer, P. Piroué, J. Salfeld-Nebgen, D. Stickland, C. Tully

University of Puerto Rico, Mayaguez, USA
S. Malik, S. Norberg

Purdue University, West Lafayette, USA
A. Barker, V.E. Barnes, S. Das, L. Gutay, M. Jones, A.W. Jung, A. Khatiwada, D.H. Miller, N. Neumeister, C.C. Peng, H. Qiu, J.F. Schulte, J. Sun, F. Wang, R. Xiao, W. Xie

Purdue University Northwest, Hammond, USA
T. Cheng, N. Parashar

Rice University, Houston, USA
Z. Chen, K.M. Ecklund, S. Freed, F.J.M. Geurts, M. Guilbaud, M. Kilpatrick, W. Li, B. Michlin, B.P. Padley, J. Roberts, J. Rorie, W. Shi, Z. Tu, J. Zabel, A. Zhang

University of Rochester, Rochester, USA
A. Bodek, P. de Barbaro, R. Demina, Y.t. Duh, T. Ferbel, M. Galanti, A. Garcia-Bellido, J. Han, O. Hindrichs, A. Khukhunaishvili, K.H. Lo, P. Tan, M. Verzetti

The Rockefeller University, New York, USA
R. Ciesielski, K. Goulianos, C. Mesropian

Rutgers, The State University of New Jersey, Piscataway, USA
A. Agapitos, J.P. Chou, Y. Gershtein, T.A. Gómez Espinosa, E. Halkiadakis, M. Heindl, E. Hughes, S. Kaplan, R. Kunnawalkam Elayavalli, S. Kyriacou, A. Lath, R. Montalvo, K. Nash, M. Osherson, H. Saka, S. Salur, S. Schnetzer, D. Sheffield, S. Somalwar, R. Stone, S. Thomas, P. Thomassen, M. Walker

University of Tennessee, Knoxville, USA
A.G. Delannoy, J. Heideman, G. Riley, K. Rose, S. Spanier, K. Thapa

Texas A&M University, College Station, USA
O. Bouhali[82], A. Castaneda Hernandez[82], A. Celik, M. Dalchenko, M. De Mattia, A. Delgado, S. Dildick, R. Eusebi, J. Gilmore, T. Huang, T. Kamon[83], R. Mueller, Y. Pakhotin, R. Patel, A. Perloff, L. Perniè, D. Rathjens, A. Safonov, A. Tatarinov

Texas Tech University, Lubbock, USA
N. Akchurin, J. Damgov, F. De Guio, P.R. Dudero, J. Faulkner, E. Gurpinar, S. Kunori, V. Kuryatkov, K. Lamichhane, S.W. Lee, T. Mengke, S. Muthumuni, T. Peltola, S. Undleeb, I. Volobouev, Z. Wang

Vanderbilt University, Nashville, USA
S. Greene, A. Gurrola, R. Janjam, W. Johns, C. Maguire, A. Melo, H. Ni, K. Padeken, P. Sheldon, S. Tuo, J. Velkovska, Q. Xu

University of Virginia, Charlottesville, USA
M.W. Arenton, P. Barria, B. Cox, R. Hirosky, M. Joyce, A. Ledovskoy, H. Li, C. Neu, T. Sinthuprasith, Y. Wang, E. Wolfe, F. Xia

Wayne State University, Detroit, USA
R. Harr, P.E. Karchin, N. Poudyal, J. Sturdy, P. Thapa, S. Zaleski

University of Wisconsin - Madison, Madison, WI, USA
M. Brodski, J. Buchanan, C. Caillol, D. Carlsmith, S. Dasu, L. Dodd, S. Duric, B. Gomber, M. Grothe, M. Herndon, A. Hervé, U. Hussain, P. Klabbers, A. Lanaro, A. Levine, K. Long, R. Loveless, V. Rekovic, T. Ruggles, A. Savin, N. Smith, W.H. Smith, A. Sourkov[84], N. Woods

Brookhaven National Laboratory, Upton, Long Island, USA
F. Lanni

†: Deceased
1: Also at Vienna University of Technology, Vienna, Austria
2: Also at IRFU; CEA; Université Paris-Saclay, Gif-sur-Yvette, France
3: Also at Universidade Estadual de Campinas, Campinas, Brazil

4: Also at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
5: Also at Universidade Federal de Pelotas, Pelotas, Brazil
6: Also at Université Libre de Bruxelles, Bruxelles, Belgium
7: Also at Institute for Theoretical and Experimental Physics, Moscow, Russia
8: Also at Joint Institute for Nuclear Research, Dubna, Russia
9: Also at Cairo University, Cairo, Egypt
10: Also at Suez University, Suez, Egypt
11: Now at British University in Egypt, Cairo, Egypt
12: Also at Zewail City of Science and Technology, Zewail, Egypt
13: Also at CERN; European Organization for Nuclear Research, Geneva, Switzerland
14: Also at Department of Physics; King Abdulaziz University, Jeddah, Saudi Arabia
15: Also at Université de Haute Alsace, Mulhouse, France
16: Also at Skobeltsyn Institute of Nuclear Physics; Lomonosov Moscow State University, Moscow, Russia
17: Also at Tbilisi State University, Tbilisi, Georgia
18: Also at RWTH Aachen University; III. Physikalisches Institut A, Aachen, Germany
19: Also at University of Hamburg, Hamburg, Germany
20: Also at Brandenburg University of Technology, Cottbus, Germany
21: Also at MTA-ELTE Lendület CMS Particle and Nuclear Physics Group; Eötvös Loránd University, Budapest, Hungary
22: Also at Institute of Nuclear Research ATOMKI, Debrecen, Hungary
23: Also at Institute of Physics; University of Debrecen, Debrecen, Hungary
24: Also at Indian Institute of Technology Bhubaneswar, Bhubaneswar, India
25: Also at Institute of Physics, Bhubaneswar, India
26: Also at Shoolini University, Solan, India
27: Also at Indian Institute of Technology (IIT), Mumbai, India
28: Also at University of Visva-Bharati, Santiniketan, India
29: Also at University of Ruhuna, Matara, Sri Lanka
30: Also at Isfahan University of Technology, Isfahan, Iran
31: Also at Yazd University, Yazd, Iran
32: Also at Plasma Physics Research Center; Science and Research Branch; Islamic Azad University, Tehran, Iran
33: Also at Università degli Studi di Siena, Siena, Italy
34: Also at ENEA - Casaccia Research Center, S. Maria di Galeria, Italy
35: Also at Facoltà Ingegneria; Università di Roma, Roma, Italy
36: Also at INFN Sezione di Milano-Bicocca; Università di Milano-Bicocca, Milano, Italy
37: Also at International Islamic University of Malaysia, Kuala Lumpur, Malaysia
38: Also at Malaysian Nuclear Agency; MOSTI, Kajang, Malaysia
39: Also at Consejo Nacional de Ciencia y Tecnología, Mexico city, Mexico
40: Also at Warsaw University of Technology; Institute of Electronic Systems, Warsaw, Poland
41: Also at Institute for Nuclear Research, Moscow, Russia
42: Now at National Research Nuclear University 'Moscow Engineering Physics Institute' (MEPhI), Moscow, Russia
43: Also at Institute of Nuclear Physics of the Uzbekistan Academy of Sciences, Tashkent, Uzbekistan
44: Also at St. Petersburg State Polytechnical University, St. Petersburg, Russia

45: Also at University of Florida, Gainesville, USA
46: Also at P.N. Lebedev Physical Institute, Moscow, Russia
47: Also at INFN Sezione di Padova; Università di Padova; Università di Trento (Trento), Padova, Italy
48: Also at Budker Institute of Nuclear Physics, Novosibirsk, Russia
49: Also at Faculty of Physics; University of Belgrade, Belgrade, Serbia
50: Also at INFN Sezione di Pavia; Università di Pavia, Pavia, Italy
51: Also at University of Belgrade; Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
52: Also at Scuola Normale e Sezione dell'INFN, Pisa, Italy
53: Also at National and Kapodistrian University of Athens, Athens, Greece
54: Also at Riga Technical University, Riga, Latvia
55: Also at Universität Zürich, Zurich, Switzerland
56: Also at Stefan Meyer Institute for Subatomic Physics (SMI), Vienna, Austria
57: Also at Adiyaman University, Adiyaman, Turkey
58: Also at Istanbul Aydin University, Istanbul, Turkey
59: Also at Mersin University, Mersin, Turkey
60: Also at Piri Reis University, Istanbul, Turkey
61: Also at Izmir Institute of Technology, Izmir, Turkey
62: Also at Necmettin Erbakan University, Konya, Turkey
63: Also at Marmara University, Istanbul, Turkey
64: Also at Kafkas University, Kars, Turkey
65: Also at Istanbul University; Faculty of Science, Istanbul, Turkey
66: Also at Istanbul Bilgi University, Istanbul, Turkey
67: Also at Yildiz Technical University, Istanbul, Turkey
68: Also at Hacettepe University, Ankara, Turkey
69: Also at Rutherford Appleton Laboratory, Didcot, United Kingdom
70: Also at School of Physics and Astronomy; University of Southampton, Southampton, United Kingdom
71: Also at Monash University; Faculty of Science, Clayton, Australia
72: Also at Instituto de Astrofísica de Canarias, La Laguna, Spain
73: Also at Bethel University, ST. PAUL, USA
74: Also at Utah Valley University, Orem, USA
75: Also at California Lutheran University, Thousand Oaks, USA
76: Also at Purdue University, West Lafayette, USA
77: Also at Beykent University, Istanbul, Turkey
78: Also at Bingol University, Bingol, Turkey
79: Also at Erzincan University, Erzincan, Turkey
80: Also at Sinop University, Sinop, Turkey
81: Also at Mimar Sinan University; Istanbul, Istanbul, Turkey
82: Also at Texas A&M University at Qatar, Doha, Qatar
83: Also at Kyungpook National University, Daegu, Korea
84: Also at State Research Center of Russian Federation; Institute for High Energy Physics, Protvino, Russia