High-Granularity Calorimeter for the CMS Phase 2 Upgrade.

Athens, Beijing, CERN, Demokritos, Imperial College, Iowa,
LLR, Minnesota, SINP (Kolkata), UC Santa Barbara.

March 5, 2014

1 Introduction

The CMS endcap calorimeters will be replaced for the high luminosity LHC running that aims to record an integrated luminosity of 3000 fb^1. We propose a dense and compact approach for both electromagnetic and hadronic calorimetry that uses a high lateral and longitudinal granularity. Recent advances in Si sensors in terms of cost per unit area and radiation tolerance, and advances in electronics and data transmission bring up the possibility of their use in such high granularity calorimetry. High granularity calorimeters are proposed for future ILC/CLIC detectors, for which they have been shown to provide very high resolving power for single particles in dense jet environments, with energies of several hundred GeV's. The challenges faced for high-luminosity LHC operation are mainly in the area of engineering (mechanical and thermal), data transmission and Level-1 trigger formation.

The performance advantages of a high-granularity calorimeter are that it is possible to track the growth and measure the angle of an electromagnetic shower, and to apply of the methods of particle flow to optimize the jet energy resolution. To demonstrate the power of this type of calorimeter in a high-pileup environment will require a full analysis of physics channels and particle flow reconstruction. While we have not yet reached that point in our simulations, we can, however, point to the improvements obtained in our physics performance by segmenting the HCAL into three parts, as in the Phase 1 upgrade, as an indicator of the benefits of a highly granular calorimeter in CMS.

2 Detector Concept

The main idea of the calorimeter design is shown in Figure [1]. Moving outwards from the interaction region, the volume begins at the front face of the current EE with an EM calorimeter with 25 X_0 lead-silicon sampling calorimeter with a depth of $\sim 1\lambda$. This is followed by a 4λ brass-silicon hadron calorimeter, the Front HCAL, which is followed by a 5λ brass-scintillator calorimeter, the Backing HCAL, to bring the total number of interaction lengths at 10. The gaps between the different calorimeter sections should be minimized, with the most significant one being between the Front and Backing HCALs where there will be a thermal screen.

The EM section and the Front HCAL will use cooled planes of silicon as the active medium, while the Back-HCAL, where the radiation levels are low, can be constructed with either plastic scintillator, as with the current HE, or with GEM or Micromegas gas chambers.

In this structure there will be only a small gap between the EM and Hadron calorimeters, which we expect will improve the jet energy measurements. Additionally, by bringing the hadron calorimeter into the space currently used for the EE electronics, there will be space behind the Back-HCAL where the radiation levels will be sufficiently low so that not only electronics stations (as with the HE-RBX boxes) and other services, but also additional muon chambers could be situated.

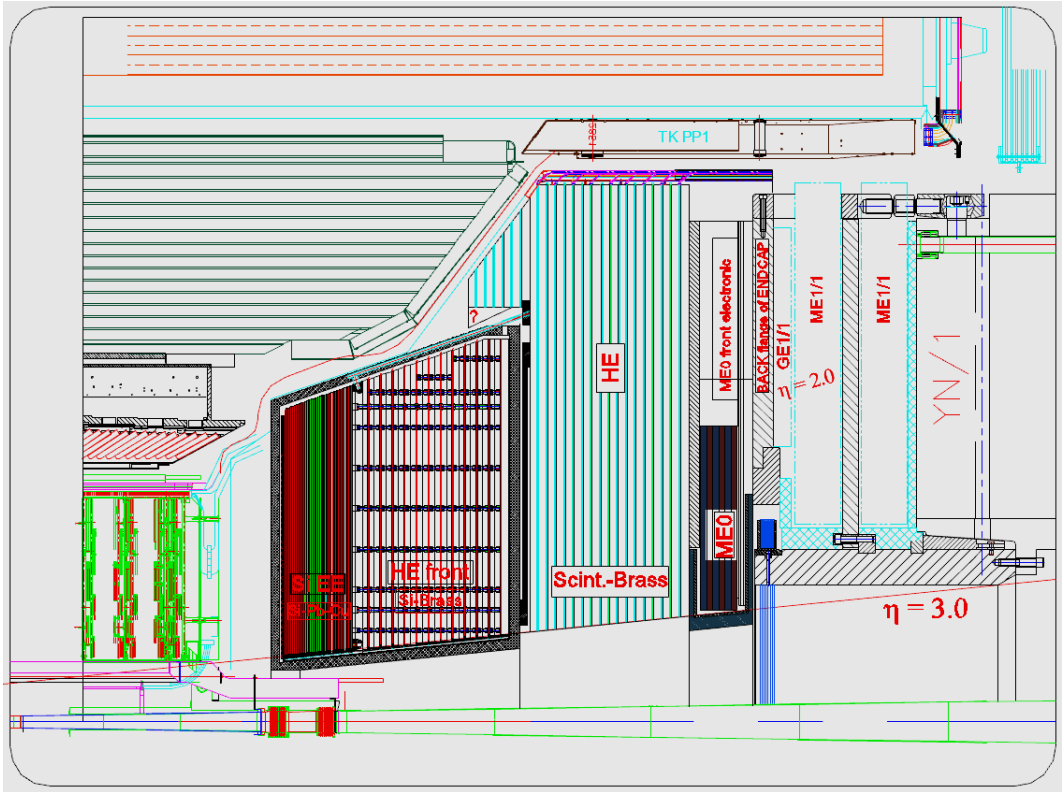

Figure 1: Sketch of a possible structure. The electromagnetic calorimeter has its front face at the same location as the front face of the current EE. Directly behind it there is the Front-HCAL, which a 4λ silicon-brass calorimeter. Behind that is a 5λ hadron calorimeter, which, since the radiation levels are low, could use the same technology as the current HE.

2.1 The Electromagnetic Calorimeter

For the EM calorimeter a variable sampling is proposed, determined by the thickness of lead plates in the electromagnetic calorimeter. The longitudinal sampling, determined by the absorber thickness is:

- 11 planes of silicon separated by $0.5X_0$ of lead/Cu,
- 10 planes of silicon separated by $0.8X_0$ of lead/Cu,
- 10 planes of silicon separated by $1.2X_0$ of lead/Cu.

A similiar configuration of absorbing (W) material has been studied in test beams by the CALICE Collaboration and, when instrumented with 500μm thick Silicon sensors, gave an energy resolution of around 16.5%/E with a linear energy response[3]. If shown to lead to tangible cost-effective benefits tungsten could be considered as absorber for the EE.

The scheme for the lateral granularity (pad structure on the silicon planes) is as follows:

- 21 planes of silicon with a pad size of 0.9 cm^2 followed by
- 10 planes of silicon with a pad size of 1.8 cm^2

These are indicative sizes and the scheme needs further optimization for physics performance, but has been used to estimate the total channel count, power and data handling requirements. The outermost "tower" ($1.44 < |\eta| < 1.57$) will have only alternate layers equipped with active longitudinal sampling. This results in $\sim$ 3.7 M readout channels. This structure has an effective radiation length is 13.5 mm, and we find in simulations that the Molière radius is 25 mm, which can compared with the value for $PbWO_4$ of 22 mm.

The total surface area of silicon is estimated to be 420 m^2 for both ECAL endcaps.

2.2 The Front Hadron Calorimeter

It is proposed to split the hadron calorimeter into two parts. The first half has silicon as the active medium and a thickness of 4λ with 0.33λ sampling, for a total thickness of about 70 cm, which allows it to approximately fit into the present ECAL volume. The absorber material would be brass[1]. Together with the EM part this would give a total depth (EM + hadronic) of $\sim 5\lambda$. The lateral granularity is 1.8 cm^2. This gives a total channel count of 1.4 M channels and a total surface area of 250 m^2 for both endcaps.

The option of a replaceable high eta nose is being considered, in order to have the ability to change only the highest eta part (e.g. $\sim 2.6 \leq |\eta| \leq 4$), which could allow for possible unanticipated issues arising, but is not discussed here.

2.3 The Backing Hadron Calorimeter

The backing hadron calorimeter has a depth of 5λ, to give a total depth for the whole calorimeter of 10λ. The aim is to arrive at a maximum radiation dose at the front face and at the highest $|\eta|$ of $\sim$5 MRads for an integrated luminosity of 3000fb^{-1}, i.e. comparable with that in the barrel region. This would enable the deployment of various technologies including plastic scintillator. The backing calorimeter is proposed to have 22 planes in the same geometry as the current HE (giving effectively 11 each with a thickness of $\sim 5\lambda$). The lateral granularity has still to be defined, but will be finer than the current HE.

As discussed above, this Endcap calorimeter system would be shorter than the present EE and HE, and the space made available behind the new calorimeter could be used in a variety of ways to improve the performance of the experiment. For example, an early measuring station for muons could be installed. GEM detectors are the natural choice for these detectors, to match to GE1/1, but Micromegas could also work in this region.

Table 1: Calorimeter Parameters

	EE	FH	Total
Area of silicon (m^2)	420	250	670
Channels	3.7M	1.4M	5.1M
Detector Modules	19,000	11,000	30,000
Weight One Endcap (tonnes)	16	63	79
Number of Si planes	31	12	

[1]Steel is also option, it is less expensive, more costly to machine and has a slightly larger interaction length.

3 Mechanical and Thermal Engineering

The layout of the endcap calorimeters in this option is shown in Fig. [1] Two variants of the mechanical design and two for the cooling design are being considered. Here we consider one pair of variants (mechanical and thermal) for the ECAL part and the other pair for the Front HCAL. We expect to eventually use the same pairing for both the ECAL and Front HCAL. Also shown is the Back HCAL using plastic scintillator.

3.1 Mechanical and Thermal Design for the ECAL

It is intended that the composite metal plates serve the functions of absorber, support and cooling. The calorimeter is constructed by stacking the composite metal plates on which are mounted the sensors/PCBs+electronics. The variant of the composite metal plate being considered for the ECAL is to sandwich lead plates of a size of 25×29 cm^2 of appropriate thickness in between two 1mm thick Cu plates.

The lead plates are glued/bonded/braised to the copper plates. The lead plates are arranged such that a series of channels having an equivalent hydraulic diameter of about 3mm are created for the cooling fluid to pass through. The hexagonal silicon modules are then tiled onto one side of the composite plate. The corners of some of the hexagonal modules will be cut to allow for string tie-rods to pass, through spacers, from a strong back plate to the front plate. The layout of the sensor, the cooled composite plate and the spacer-tie rods are shown in Fig. 2. The tie-rod has a diameter of 2mm and can be made out of high-strength wires. The spacers through which the rod passes make sure that no force is transmitted to the sensors through the composite plates and at the same time support the shear force. The diameter of the spacers would be about 4mm and the dead area allowed would have a diameter of 5mm. The longitudinal layout of a unit cell is shown in Fig. 3 and is as follows: 1mm Cu, Pb absorber, 1mm Cu, thin kapton insulating layer, 300um Si, 1.5-2mmPCB, 2mm air gap.

As one variant we are investigating a cooling by nucleate boiling circulating by thermosiphon, not dissimilar to the way the CMS coil is cooled The envisaged liquid is C_3F_8 like in the present evaporative cooling system of the ATLAS tracker, or a mixture like C_3F_8-C_2F_6, that would have the coolant evaporation temperature of -30°C. With this liquid in this nucleate boiling mode the exchange coefficient can reach 2 to 4 W/cm^2. In the choice of the liquid, the effect on the environment must also be considered. It is likely that liquids developed for the cold warehouses can replace, if necessary, the presently envisaged liquids. This first thermosiphon system allows the cooling fluid to circulate naturally without any mechanical pumping component in the endcap calorimeter circuit. The cold liquid starts from a phase-separation reservoir, situated on the top behind the cold endcap calorimeter and is fed at the bottom of the plates. The heat generated by the electronics warms up and creates bubbles in the cooling liquid that by gravity creates the pressure difference that drives the movement of the fluid. The mixture of gas/liquid is returned to the phase separator reservoir. The liquid fraction is recirculated, while the gas fraction is warmed up by an electric heater situated behind the cold endcap calorimeter and returned to the cooling plant as warm gas.

The maximum heat dissipation will be on the plate that has sensors, each with 256-channels. The current estimate for this is 1.75kW leading to around 300W/m^2. Initial calculations show that 25-capillaries of 3mm equivalent hydraulic diameter should be sufficient to remove the heat. As said, the fluid would be fed from the bottom of the plates. At the top of the plate it is anticipated that

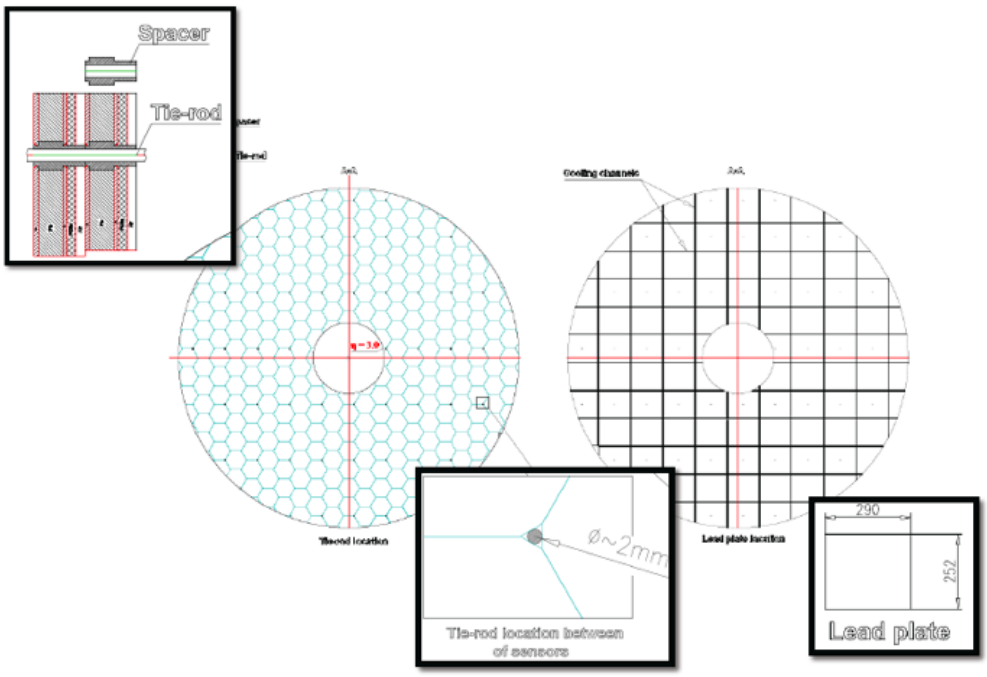

Figure 2: The detailed layout of the sensor, the cooled composite plate and the spacer-tie rods.

there would be 70:30 mix of liquid and gas.

The ATLAS thermosiphon system, Fig. 4, provides warm (20°C), high-pressure liquid perfluoropropane (C_3F_8) at the distribution point. The liquid expands inside capillaries and evaporates at -25°C and 1.67 bar at the detectors liquid distribution point, in our case this would be the phase-separation reservoir. The warm (20°C) gas collected at the detector exhaust is taken to the thermosiphon condenser located on the roof of the SH1 building at Point 1. Here, the gas is liquefied at 0.3 bar (-60°C). The 92-m-high liquid column creates up to 16.5 bar of hydrostatic pressure at the detectors liquid distribution point. The thermosiphon system will circulate 1.2 kg of perfluoropropane per second to remove up to 62.4 kW of heat dissipated by the ATLAS inner silicon detector electronics

3.2 Mechanical and Thermal Design for the Front HCAL

As variants the design for FH could either follows the mechanical structure of the current HE (i.e. a bolted structure) or the one prototyped and described in the ILD TDR, shown in Fig. 5 with carbon fibre structure that integrates every alternate absorber plate, and creates drawers for the insertion of two active planes plus another absorber plate. We would use a composite absorber plate that integrates the cooling function.

The active planes thus have a shape of a “petal” (Fig. 6) or a “drawer”. Both designs can be accomplished with three sensor geometries: full- or two half-hexagons. The sensors would be put into intimate contact with (glued to?) composite absorber plate.

The installation of the ILD-type of modules would be off of the front plate of BH (Backing HCAL) using horizontal rails

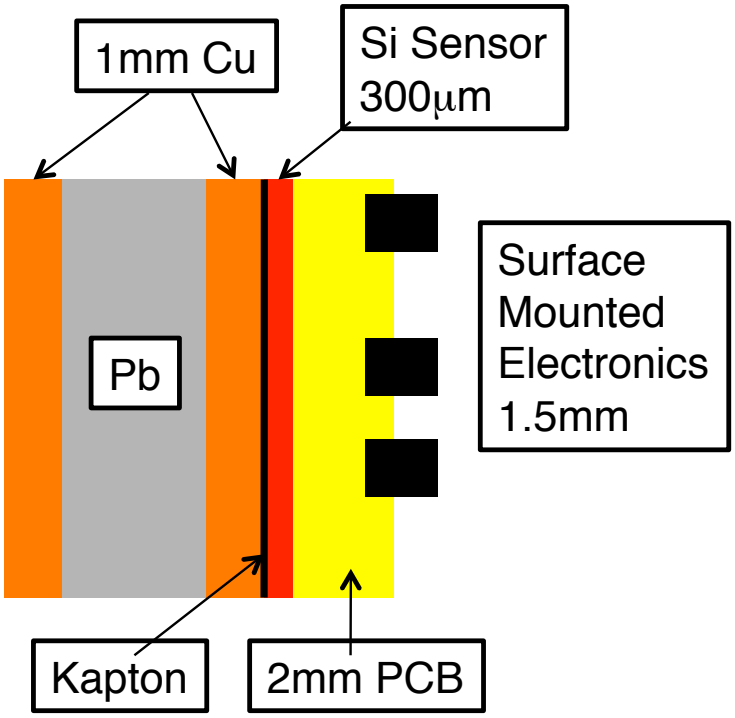

Figure 3: Sketch of the longitudinal arrangement in a single layer.

For this variant we shall consider the use of evaporative CO^2 cooling described in the Phase I pixels TDR and currently being designed for the Phase II CMS Tracker. We are in contact with the engineers responsible for the design of the CMS Phase II Tracker. This variant would use SS pipes through which the cooling fluid would pass. The SS pipes would be embedded in a copper plate of a thickness of approximately 3mm. In CO^2 based cooling evaporation takes place at much higher pressures than other two-phase refrigerants. In general, the volume of vapour created stays low while it remains compressed, which means that it flows more easily through small channels. The evaporation temperature of high-pressure CO^2 in small cooling lines is also more stable because the pressure drop has a limited effect on the boiling pressure. Hence it is possible to use smaller cooling pipes (see Fig. 7).

Additional benefits of using CO_2, a natural gas, are low cost and environmentally friendliness. CO_2 evaporates from its liquid phase between -56°C and +31°C and a practical range of application is from -45°C to +25°C.

Whatever fluid is used care will have to taken about purity as far as radio-activation and venting is concerned.

3.3 Mechanical Structure - Supports

The draft detailed longitudinal segmentation of the three parts of the calorimeter is shown in Fig. 8. Also shown are 5cm thick heat screens. It is anticipated that active screens will enable a reduction in this thickness.

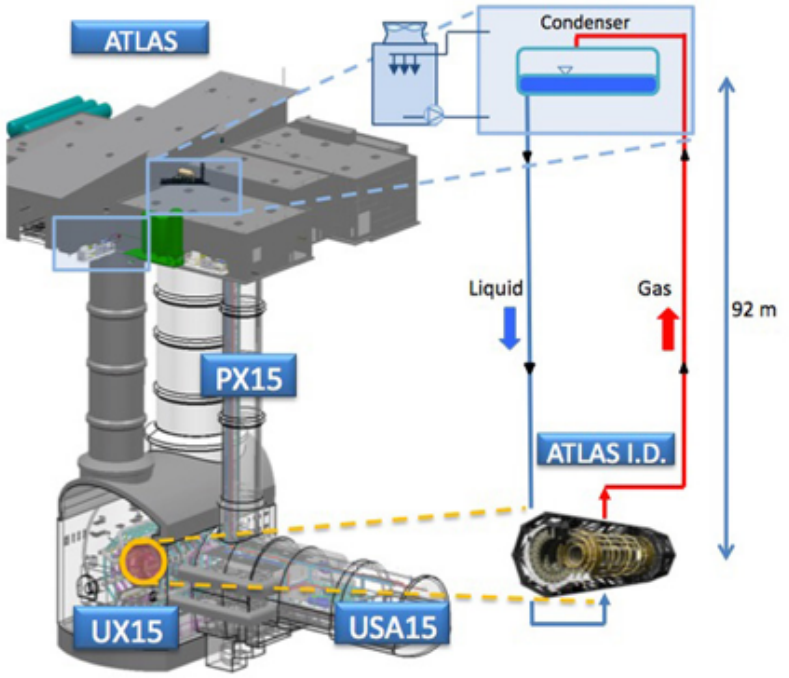

Figure 4: There ATLAS Thermosyphon system. The plant is situated on the surface giving a head of 100 m to create the pressure difference that drives the movement of the fluid without any pump. The ATLAS thermosiphon system provides warm (20°C), high-pressure liquid perfluoropropane (C_3F_8) at the distribution point.

3.4 Services

For the variants being discussed here, with reference to Fig 2, one proposal is to support the ECAL section using tie-rods (stainless steel or titanium) attached to the Front HCAL (see detail in Fig. 9). Since the bolted structure for the Front and Back HCAL is similar to the current CMS-HCAL design, the support also is envisaged to be similar to the current one.

We have made an estimate of the services that need to taken in/out from the densest active plane for the sector geometry. At the periphery we have a space of about 80cm×2mm. Assuming a sector design, there are approximately 25 modules to be serviced per 30° sector. The services comprise:

Low Voltage: A section of about 50mm^2 of copper is needed. Small cables or a 200μm kapton with a 100μm Cu sheet with a width of 2cm can be considered. Taking the kapton cables leads to 100 cables with 2 cables in and 2 out. We envisage arranging these cables in 5 stacks of 5, each stack with a section of 4cm×1mm.

High Voltage: It is estimated that we will need to supply 10 mA/module at a voltage of 900V. One mm diameter cables are used in our current tracker modules to supply the HV. We

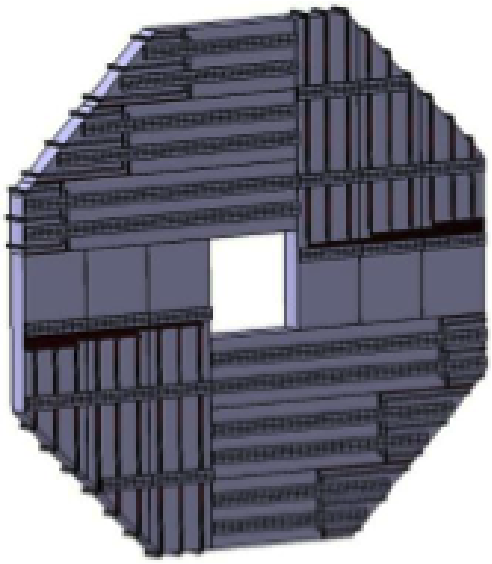

Figure 5: ILD endcap structure with "drawers".

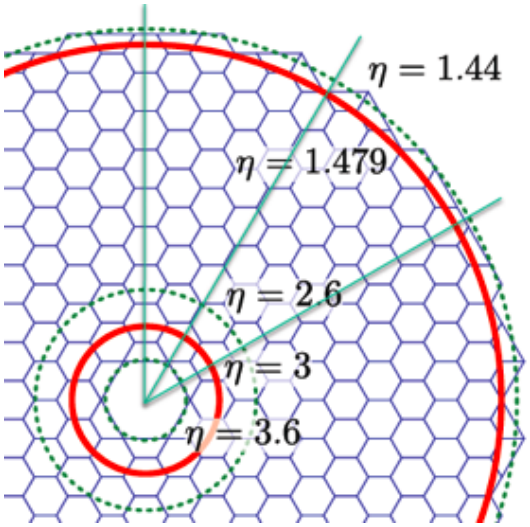

Figure 6: "Petals" geometry for the active layers showing use of full- and half-hexagons.

envisage arranging these cables at 5 locations, each set with a section of 1cm×1mm.

Data Links: We envisage 3-4 links per module. Using 4 Twinax links per module, leads to 100 link-cables per sector. Each cable has a section of 1.25mm×1.95mm. We envisage arranging these cables in sets of 20 at 5 locations, each with a section of 4cm×1.25mm.

Hence the section used for cables is 45cm×1.25mm giving the fraction used as 35%. It is assumed that the cooling pipes are accounted within the thickness of the composite metal plates.

4 Module Design

4.1 Sensors

The silicon sensors for the HGC will be simple, large area, single-sided, DC-Coupled with pad read-out; they will have an active thickness of 200μm, except towards the inner radius where this will be reduced to 100μm, with physical thickness can be set at 320μm or 500μm to allow production on high volume commercial lines. The preferred sensor type if p-on-n, as this minimises the processing costs, but n-on-p remains a possible alternative should it prove to be more radiation tolerant. The sensors will be hexagonally shaped, so as to make best use of the wafer surface (a factor 1.3 gain with respect to a square sensor), while providing a conveniently tile-able geometry. For the sensor production we assume the timely availability of 8" production lines, so that a full size hexagonal sensor will cover about 230 cm^2.

4.2 Module design

Each sensor will be assembled into a module. In our current design the first 20 layers of EC are equipped modules with 256 channels, while the back layers of EC and all of the layers in HC are

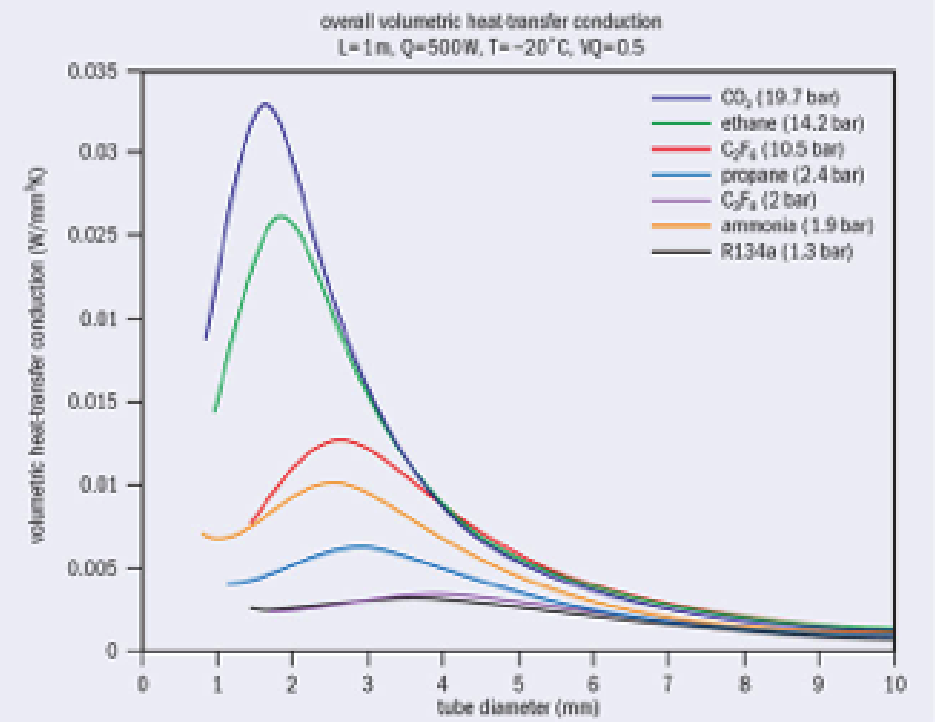

Figure 7: Overall volumetric heat transfer conduction coefficients for different refrigerants as a function of tube diameter.

equipped with modules with 128 channels.

The Front-End read chips are 64-channel wide, and include an amplifier, a 40MHz low power ADC as well as logic for digital data handling. There will be either 4 or 2 such chips on a module, according to the number of read-out channels. A large area multi-layer PCB covering most of the sensor surface will route signals from the sensor to the Front-End chips. The connections to the pads on the sensor will be made with wire bonding, through suitable openings in the PCB, while the Front-End chip will likely be flip-chip bonded to the PCB.

It is envisaged that the Trigger will be generated as the sum of 4 adjacent channels, and that both zero suppression as well as a simple data compression scheme will be applied to reduce both the L1 Trigger and full resolution read-out data rates. This functionality could be either included in the Front-End chips, or may be carried out in a separate data concentrator chip, which would then also provide the interface to the data links.

Each module will produce up to 7.5-Gbps of L1 Trigger data, and up to 3.2-Gbps of full resolution data, assuming a 1MHz L1 accept rate. The radiation levels at the inner radius of the EC will likely preclude the use of on-module optical links. As a result, the present design uses three separate 5-Gbps electrical links to transfer the data from the module to the back of the HCAL, about 4m away, where optical links will be deployed. A fourth link provides the clock and control function. The electrical links will use Twinax cables, arranged as a ribbon in a thin, flexible support strip. Very compact connectors exist for this type of cable, with a height of 2mm or less, which can fit within the available space inside the calorimeter.

Similarly, the radiation levels, and very stringent space constraints, will make it difficult to house

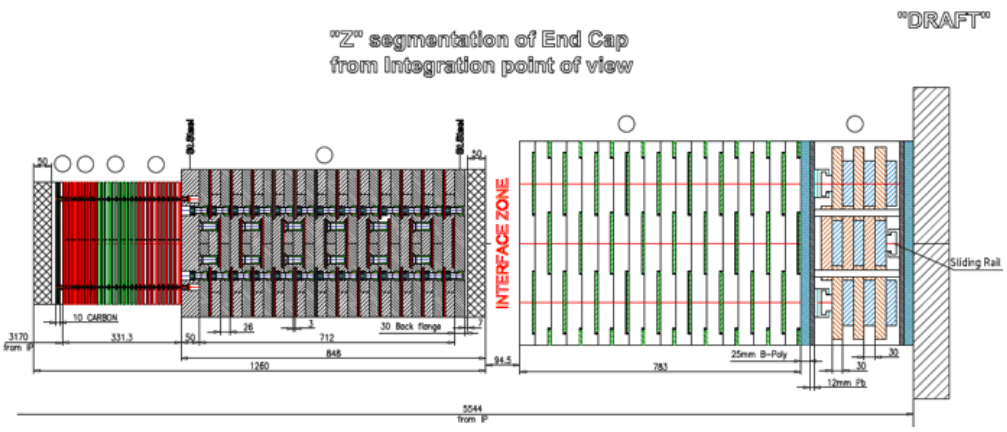

Figure 8: The details of the longitudinal segmentation.

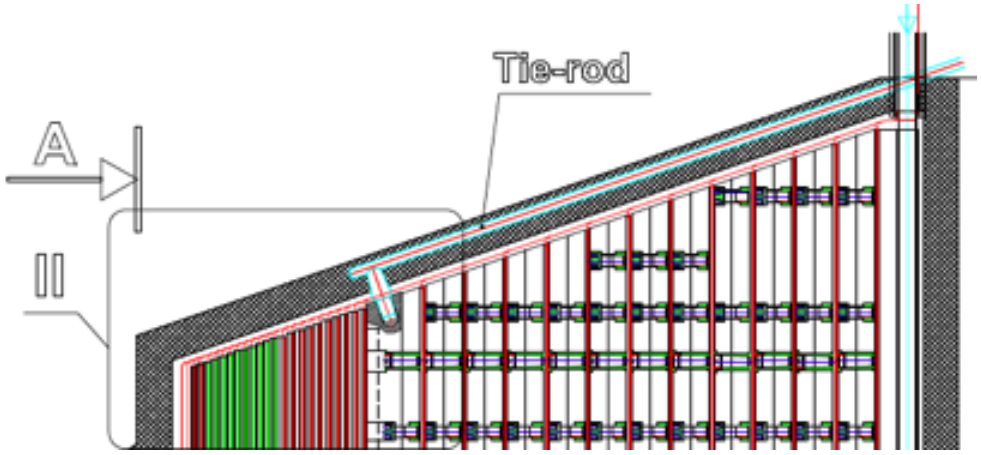

Figure 9: The details of the longitudinal segmentation.

a DC-DC converter directly on module. The present design therefore foresees the deployment of DC-DC converters at the back of the HCAL, so that a substantial copper cross-section is required to limit the power loss over the 4m cable length. The power cables will be low profile Kapton PCBs, with wide copper traces providing the required cross-section. The sensor bias will be provided with a small diameter high voltage cable (1kV). Various options for connecting the power and bias cables to the module are being explored. One possibility is to use a compact strain relief to mechanically secure the cables to the module, and then to wire-bond them to the module. In this scenario, the module would be connected to the cables just prior to being integrated into the mechanical structure. The use of wire-bonds allows the cables to be disconnected from the module, should this ever be required for repair and/or re-work. Alternatively, a low temperature solder connection could be used instead, which may prove more robust over the full life time of the calorimeter.

Cooling circuits are imbedded in a Copper support plate. Modules will be placed onto the cooled support plate such that the silicon is in good thermal contact over its full surface, with a thin Kapton layer ensuring high voltage insulation of the sensor back-plane (which may be biased up to 900V). Heat generated by the module electronics flows through the sensor, to the cooled support plate.

4.3 Module Power Requirements

The present power estimates from the front-end readout are based on the design target of ∼6mW/channel for the GDSP chip presently under development; for the links the power estimate is based on the CERN LpGBT, with 500mW for a bi-directional data/CTRL link running at 5Gbps, and 300mW for a unidirectional data only link.

In addition to the front-end amplifiers, the GDSP chip includes a 40MHz 9-bit effective ADC per channel, operating at 40MHz, as well as a sophisticated back-end for digital signal processing. The target power per channel for the GDSP is 1mW for front-end amplifiers, 4mW for the ADC and 1mW for the digital signal processing respectively.

In extrapolating this to the requirements for the HGCAL, we make the following assumptions:

The cell sizes are 0.9 and 1.8cm^2 with an active thickness of 100-200μm, the input capacitance will be in the range of 100 - 200pF, and the shaping time is (∼10ns) to minimize the effects of out-of-time pileup.

- We allow for double the power of each front-end amplifier, up to 2mW/channel
- The simplest way to handle the 12bit dynamic range requirement for the ECAL section is to have 3 front-end amplifiers with different gains: We allow for 6mW/channel for the 3 front end-amplifiers
- In order to avoid complications due to switching from one channel to another, we foresee that each front-end amplifier will be independently digitized; we note that the full digitization will only be required for one of three ADCs in any given bunch crossing: We allow for 6mW/channel for the ADC
- The present design calls for a 64-channel wide read-out chip, rather than the more usual 128 channels. Only some of the power for the on-chip digital processing scales with the number of channels while part remains constant. In addition, a data concentrator chip is foreseen which may perform further digital signal processing (trigger sums etc): We allow for double the power for the digital signal processing, up to 2mW/channel

Under these assumptions, the power for the read-out chips amounts to 14mW/channel for ECAL section. For the HCAL section, a dynamic range of 9bits is sufficient so that, with the same set of assumptions stated above, we arrive at an allowance of 8mW/channel for the read-out chips.

4.4 Links

The spectrum of the energy deposited in individual cells is very steeply falling, so that with even a simple data compression scheme only relatively few bits need to be transferred compared to the full 12/9 bit dynamic range for the ECAL/HCAL. Present estimates indicate that, at high η, the L1 trigger data can be accommodated within a 6.4 (3.2) Gbps bandwidth for an ECAL (HCAL) module, while the L1 read-out information will require less than 3.2 Gbps per module.

Accordingly, the present design uses 2 one-directional LPGBT links for the L1 Trigger data, and 1 bi-directional LPGBT link for the read-out and control functions.

4.5 Module Power

With these assumptions, the power dissipation for the ECAL modules is 4.1W for the 256 channel and 2.3W for the 128 channel modules, with an additional 0.6W for modules in Trigger layers; the power consumption of the HCAL modules is 1.8 W.

4.6 Production

There are approximately 30,000 8" modules to be produced. In comparison with the existing (and planned) tracker modules, the modules described here are simpler with fewer wire bonds, and individual parts, and do not have the tight specifications for mechanical precision that are essential for a precision tracker. The design of the detector module will be optimized for ease of assembly and with a possible industrialization in mind.

5 Detector Readout

5.1 Front-End Electronics

The front-end electronics will be similar to that which is conventionally used in silicon tracker systems. The technical parameters will be close to the following:

- Shaping time: 5-10ns (the signal from silicon sensors is fast and a small shaping time will minimize noise from pileup)
- Quantization error: $< 0.2\% \rightarrow$ 9-bit
- Dynamic range: 12 bits $\rightarrow$ 3 gain ranges with a 9-bit ADC (assuming saturation at 1 TeV, max. pulse height at shower maximum should correspond to 10k mips), and noise level set at $\sim$1 mip.
- Gain ranges (maximum of each range): 100 mips, 1000 mips, 10000 mips.
- ADC samples at 40 MHz.
- Pipeline (digital) length: 10-20 us
- Preamplifier: highest gain range : 1fC to 300fC

For each channel there are three amplifiers with three different gains coupled to three separate 40 MHz 10-bit (9-bit effective) ADC to produce 12 data bits. These are sent to a data concentrator where trigger primitives are formed and the data are stored for the 20 μsec interval for the Level 1 decision. On a level-1 accept the data concentrator forms the weighted sum of the data from adjacent bunch crossings (as is currently done in the CMS tracker APV chip[2]) and transmits it to the DAQ.

[2] As a reminder in the CMS tracker each microstrip (C$\sim$20-40 pF) is read out by a charge sensitive amplifier with a shaping time of $\sim$50ns. The output voltage is sampled at the beam crossing rate of 40 MHz. Samples are stored in an analog pipeline up to Level-1 latency (3.2 μs). Following a trigger a weighted sum of 3 samples is formed in an analog circuit. This confines signal to a single bx and gives pulse height information.

The proposed layout of the front-end readout is shown in Figure[10]. In this scheme the two Low-Power GBTXs (LpGBTX)[3] operate in separate ways: for the trigger, the data are sent up on two links operating each operating at an effective rate of 3.28 Gbps and the data are sent via the second LpGBTX at on one link.

In our current design the first 20 layers of EC are equipped with modules with 256 channels, while the back layers of EC and all of the layers in HC are equipped with modules with 128 channels. The modules with 256 channels will require four FE ASICS, while those with 128 will need only two. Both types of modules will require a data concentrator and two LpGBTXs, for the EE modules that contribute to the trigger, while those that do not will need only one.

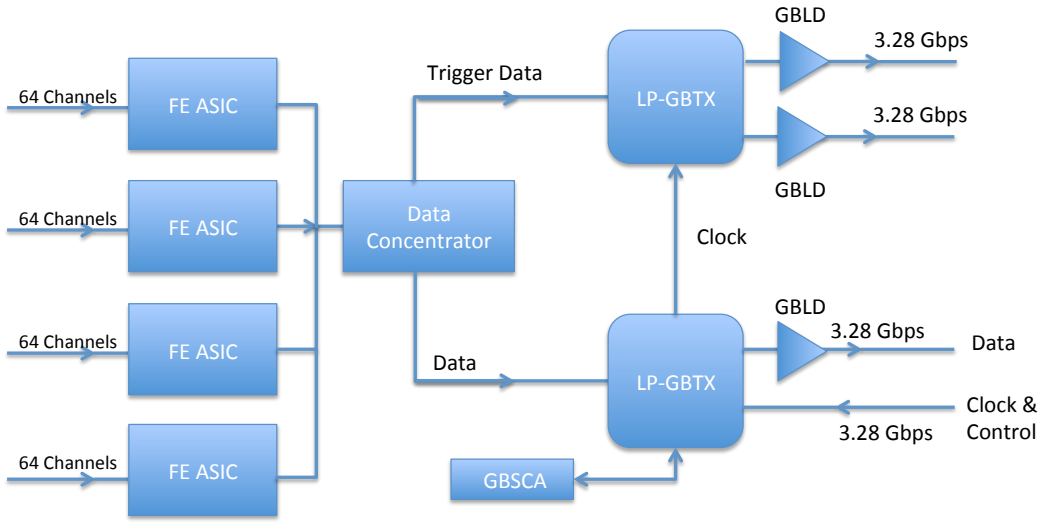

Figure 10: The readout of a 256-channel EE module with trigger data. It will consist of four 64-channel front-end ASICS connected to a data handing ASIC, similar to the ECAL FENIX chip. For every bunch crossing the trigger data will be sent on two TX links from two LpGBTXs, and on Level 1 accept data will be sent on a separate link.

5.2 Data Transfer

The data would first be transported to the back of the calorimeter on copper with $\sim$ Twinax cables. Twinax cables manufactured by TempFlex have been tested by members of the ATLAS pixel detector group who have shown that they can transfer transfer data over 5m at rates up to 8 Gbits/sec. The cross-section of the ATLAS Twinax cable is shown in Figure[11]. Micro-twisted pair cables could provide an alternative solution. The change from electrical to optical switch would take place behind the calorimeter using a rad hard FPGA, like the Igloo2 that will be used in the QIE.[4] The FPGA would be used only to deserialize the data, compress it, and to retransmit it with 10 Gbps optical links, to the service cavern.

[3] We plan to use the 'Tracker' version of the LpGBTX, which it well matched to the needs of the HGCal. This version is proposed to have only twenty 160 Mbps E-link I/O ports and requires 50% of the power of the 'Generic' version of the LpGBTX Details of the LpGBTX can be found at http://cern.ch/proj-gbt

[4] http://www.microsemi.com/products/fpga-soc/low-power

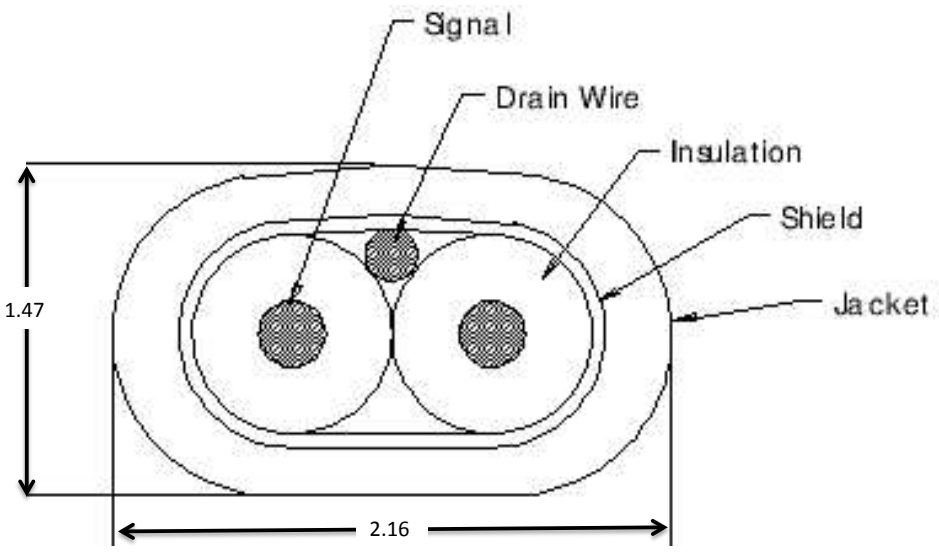

Figure 11: Shown here is the cross-section of the Twinax cable that is planned for the Phase 2 pixel detector of ATLAS.

6 Power Requirements Distribution

The low voltage power consumption for the present design of the HGC is 70kW for the ECAL and 20kW for the Front-Hcal, for a total of 90kW of front-end power at the modules. At the 1.2V level required by the electronics, this amounts to over 80kA of current. Delivering such a current over the approximately 14m long cable paths from the power supplies to the detector, while limiting the power loss along the cable to 50%, would require a copper cross-section in excess of 300cm^2. Using a DC-DC conversion as close as possible to the detector front-end can substantially reduce this.

Based on the ongoing development of DC-DC converters for the CMS Pixels and Outer Tracker, we assume a conversion ration of 8/1 with a 65% efficiency..

The very stringent space constraints within the calorimeter make it impossible to accommodate conventional transformer coils inside the calorimeter, directly on the modules. More compact PCB-based coil designs are being studied, but are not yet a proven technology. Similarly, placing the DC-DC converters on the outside surface of either the HGC and/or the Backing HCAL may be possible, but requires detailed study.

In the present design, we assume that the DC-DC converters are located behind the Backing HCAL, in the location of the present RBX boxes, and together with the optical links and other HGC services. In such a configuration, the cable runs from the power supply racks to the DC-DC converters are typically 10m long and the longest path to the HGC modules is about 4m.

Under these assumptions, and allowing for a power loss over the cables consistent with an overall 50% power delivery efficiency, the copper cross-section from the power supplies to the DC-DC converters and through the services choke points across the ME1/1 chambers can be reduced by about an order of magnitude, to around 30cm^2. On the other hand, in this configuration a copper cross-section of about 200cm^2 would be required for the cables from the DC-DC converters to the front-end modules (about 1mm^2 to 2mm^2 per module).

A serial-powering scheme would allow a substantial reduction in the copper cross-section of the cables from the DC-DC converters to the front-end modules, and is a possible option for further study.

7 Trigger

Every alternate active plane will be used for the Level-1 trigger with a granularity of 2×2 cm^2. The energy resolution of the trigger will thus be $\sqrt{2}$ worse, but still sufficient for the Level-1 trigger. For these channels the full resolution data are sent out at 40 MHz. With a granularity of 2×2cm^2 for the ECAL part, the total number of trigger channels is $\sim$ 540k channels, while for the HCAL part, with a granularity of 4×4cm^2 the trigger channel count is 137.5K.

The Level-1 accept rate is assumed to be 1 MHz. Upon receipt of the level 1 trigger 12-bit signal the weighted sum of the three adjacent bunch crossings are sent from every module to the FPGAs where they are compressed and sent to the off-detector electronics with commercial 10 Gbps optical links. The number of optical data links required to readout the detector 3.5k from each end assuming a compression factor of two and bandwidth usage of 60%.

8 Calibration and Monitoring

For silicon sensors it is standard to use m.i.p. deposits to obtain absolute and relative calibration. A study has been made where the inter-calibration of the individual pads has been randomly altered with a Gaussian width of 1, 2, 5, 10 and 20%. We are targeting a constant term smaller than 1%. In order to keep the contribution to the constant term below 0.5%, the inter-calibration error has to be kept below 5%. Silicon sensors have a uniform response by production and will be monitored during production at the wafer level. This would give the absolute and relative inter-calibration at startup. Furthermore we intend to take many tens of "longitudinal" towers into test beams to calibrate the responses to electrons and hadrons before startup.

During operation we intend to follow the inter-calibration by following the m.i.p. response. For this purpose each sensor will have a few small "inter-calibration" pads, of a size of a size of 0.20 cm^2 at varying eta, so as to give small low mip occupancy. These pads would also be coupled to a dedicated high-gain amplifier within the front-end readout chip, and read out through the normal chain. Events for inter-calibration would be selected by requiring that the surrounding cells are empty.

9 CALICE

The CALICE collaboration has been studying high granularity calorimetry for several years now and have made much progress in the understanding this type of calorimeter. While research in CALICE and the ILC calorimeter groups has been mainly directed towards design a calorimeter with excellent jet energy resolution for linear colliders. Although there are differences, for a calorimeter at the HL-LHC there are many, there are many areas where their experience can be applied in the forward calorimeter. The most relevant ones are mechanics and test beam results.

In ILD the mechanical structure of the ILD Si-W calorimeter is as shown in Fig. 13 and a detail

of the module design is shown in Fig. 14. Alternate layers of the tungsten absorber are embedded in the carbon-fiber support structure, and the detector module consists of an absorber plate with detectors mounted on either side.

The integrated circuits are recessed into the PCB to minimize the gap thickness.

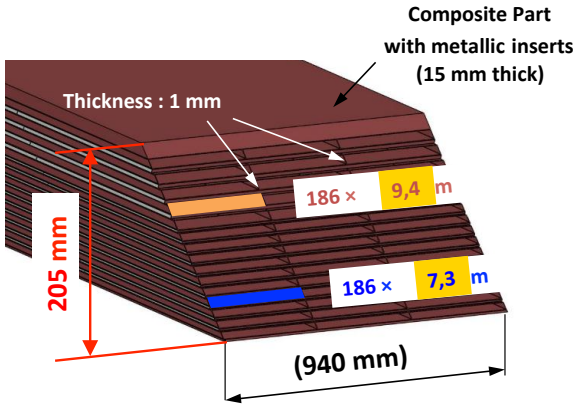

Figure 12: The ECAL of the ILD Si-W detector. The carbon fiber structure provides mechanical support to the modules and has embedded in each layer one plane of the tungsten absorber.

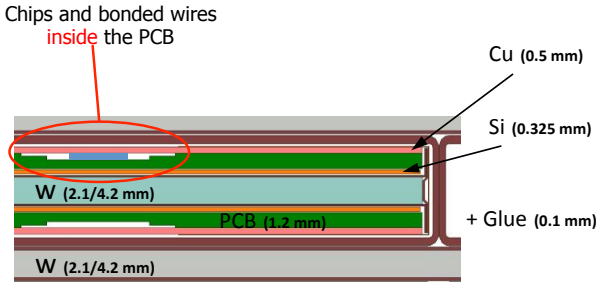

Figure 13: Detail of the module structure. The module is constructed with two plains of detectors mounted on either side of an absorber plate. The electronic components are recessed into PCB with wire-bond connections.

10 Simulations

10.1 Stand-Alone Simulations

Simulation setup

Two different simulation setups are used, in order to systematically provide an independent cross-check of the expected results for the performance of HGCal:

- stand-alone simulation of a simplified geometry using a calorimeter stack detector with a transverse size of $20 \times 20\,\mathrm{cm}^2$
- cylindrical endcap geometry integrated within an extended upgrade scenario within CMSSW.

In both cases alternative analysis are run in parallel and cross-check each other systematically. In the following sections we describe in more detail each simulation setup.

10.2 Stand-alone simulation setup

The stand-alone setup is used for three different purposes:

1. provide a quick handle to optimise the design parameters

2. extract performance results in ideal conditions

3. cross-check the performances in pileup conditions with the CMSSW setup

The implementation, based on GEANT 4 v9.6.2, is fully available. In our simulation we use the QGSP_FTFP_BERT physics list. We set the range cuts to be 10 μm for electrons, positrons and photons in all silicon volumes and 1 mm elsewhere. For reference, when considering 700 μm in Si, a 420 keV (5.8 keV) electron (photon) would deposit all its energy in one step. With the reduced 10 μm, the energy threshold lowers to 32 keV for electrons and 990 eV for photons.

Most materials are assumed to be default in GEANT 4 except the Printed Circuit Board (PCB) material which is defined as the G10 admixture composed of Si, O, C and H in the proportions 1:2:3:3 with a density of 1.7g/ cm ^{3}.

The baseline setup is described in Table 2 and it consists of 3 sampling sections, each composed of 10 layers with increasing absorber radiation lengths in the proportion 1:1.6:2.4. This setup is expected to provide fine grain sampling of the early stages of shower development, slightly coarser sampling near the shower maximum (for $E > 5$GeV) and coarser towards the end of the shower. We are still investigating whether this is optimal.

Table 2: Layout of the baseline sampling ECAL calorimeter. For each section the number of layers and material budget of each material per layer is quoted. The total material budget of a section is given in the rightmost column in units of radiation lengths.

Section	Layers	Material thickness/layer (mm)					Total budget
		Pb	Cu	Si	PCB	Air	($1/X_0$)
A	0-9	1.63	3	0.2	1	2	5
B	10-19	3.32	3	0.2	1	2	8
C	20-29	5.56	3	0.2	1	2	12

Figure 15 shows an event display of a 50GeVelectron shower transversing the baseline geometry detailed in Table 2.

10.3 Electromagnetic shower properties

Longitudinal evolution and energy estimation

In order to characterize the longitudinal evolution of the showers in our setup we sum the energy deposited by all hits in the Si layers. Energy is measured relative to the estimated MIP signal. In this study electron events are used: for incoming energies below 150GeV, 2500 events have been generated; for higher energies 1250 events are used.

Figure 16 shows the longitudinal shower evolution for single electron events as function of the X_0 transversed in the detector. It can be observed that the level of energy fluctuations for lower energy electrons is non-negligible, with δ's dominating the contribution to the total energy deposited in each Si layer. For sufficiently energetic electrons these contributions become of less important and the fluctuations are dominated by the statistical fluctuations in the number of particles comprising the shower. For very high energy showers (>150GeV) part of the energy can't be contained in

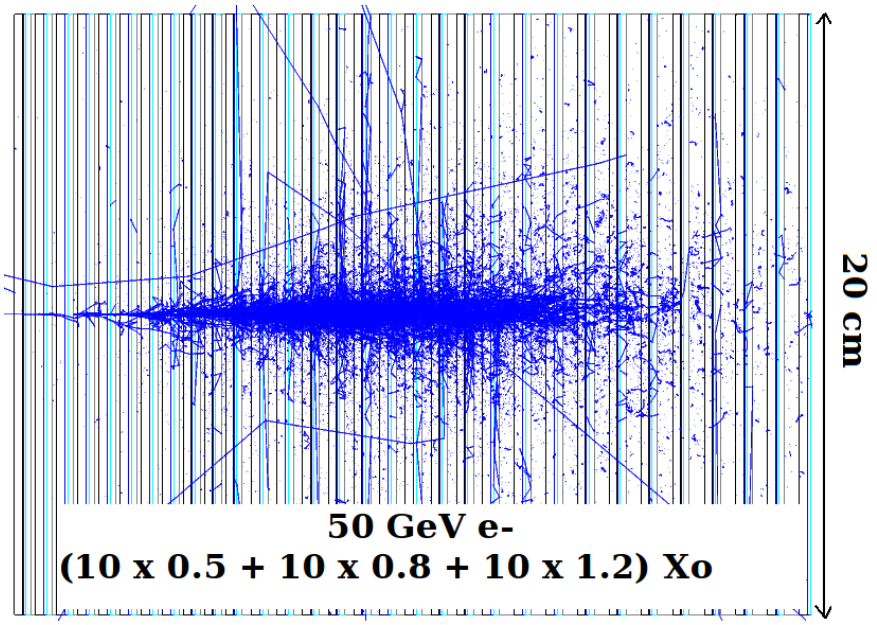

Figure 14: Event display of a 50 GeV electron showers in the standalone simulation. Thirty layers in three blocks of different thicknesses are used as the calorimeter model (cf. Table 2). Electron deposits are shown in blue. Photons have been hidden for better visualisation.

the detector setup using $25X_0$. Given that in this regime the showers are dominated by statistical fluctuations in its composition a fit to the longitudinal profile can be used safely and used to estimate the leakage fraction.

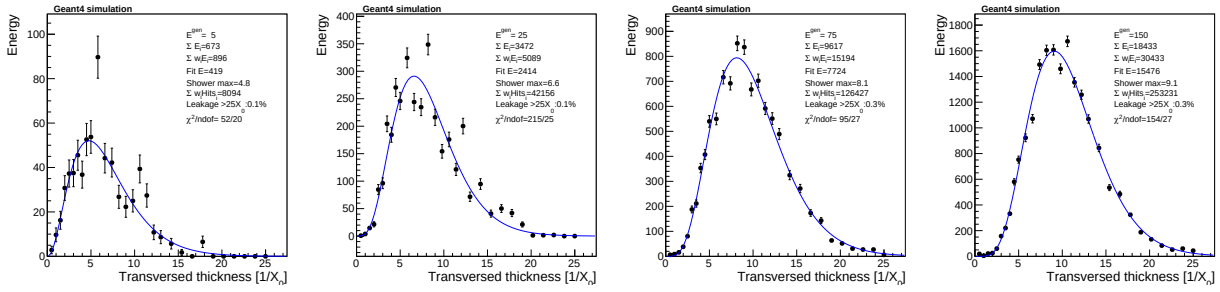

Figure 15: From *left* to *right*: energy deposits in different Si layers for single electron events with energies: 10, 25, 50, 100GeV. The deposits are shown as function of the transversed distance in X_0 units. A fit is overlaid using the functional form described in the text. The results obtained for the different energy estimators considered in the analysis, as well as for the shower leakage estimated from the fit, are shown in the caption

We sum the energy deposited in each Si layer normalized by the material overburden of the sampling section,i.e.:

$$\text{weighted E} = \sum_{i=1}^{N} \frac{X_0^i}{X_0^1} \cdot E_i \tag{1}$$

These weights correspond to 1 for section A, 1.6 for section B and 2.4 for section C. These weights can also be optimised to minimise the energy resolution for the incoming energy range of interest. The functional form used to adjust the measured energy deposits in each layer is:

$$\mathcal{E}(x) = \alpha \cdot x^a \cdot e^{-bx} \tag{2}$$

where x is the transversed material overburden measured in X_0 units. This approach is expected to recover the shower leakage for higher incident energies.

The distribution of the estimator is fitted with a Gaussian for each incoming generated beam energy. An unbinned-likelihood fit is used for this purpose. Figures 17 is an example of using the weighted energy.

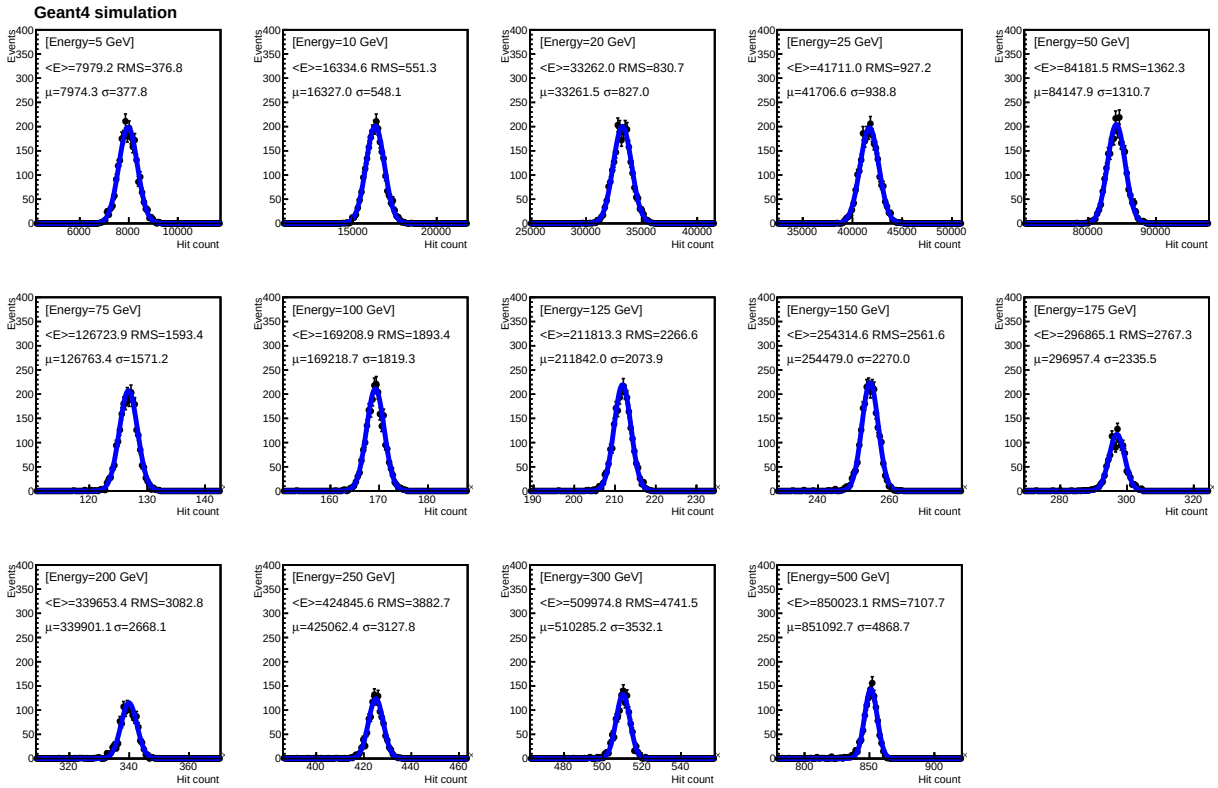

Figure 16: Weighted energy sum estimator distributions for different beam energies. The result of an unbinned-likelihood fit of a gaussian is superimposed. The mean and average of the distribution and the gaussian are compared in the caption.

The parameters of the fitted Gaussian can be used for two purposes: the mean is used to obtain the calibration (i.e. dependency of the energy estimator on the incoming energy) and the ratio of the width with respect to the mean as an estimator for the resolution. The calibration curves obtained

are shown in Fig. 18 *left* where linear behaviour is observed for all the estimators. Figure 18 *right* shows the energy resolution as a function of the incoming energy. A fit to a resolution model is super-imposed for each curve. The resolution model is based on the quadratic sum of a stochastic term (proportional to $1/\sqrt{E}$) with a constant term,i.e.:

$$\left(\frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\right)^2 = \left(\frac{\sigma_{\mathrm{stoch}}}{\sqrt{\mathrm{E}}}\right)^2 + \sigma^2_{\mathrm{cte}} \tag{3}$$

With this method a resolution of 20.9% is obtained and the residual constant term can be almost fully removed with a shower leakage recovery algorithm.

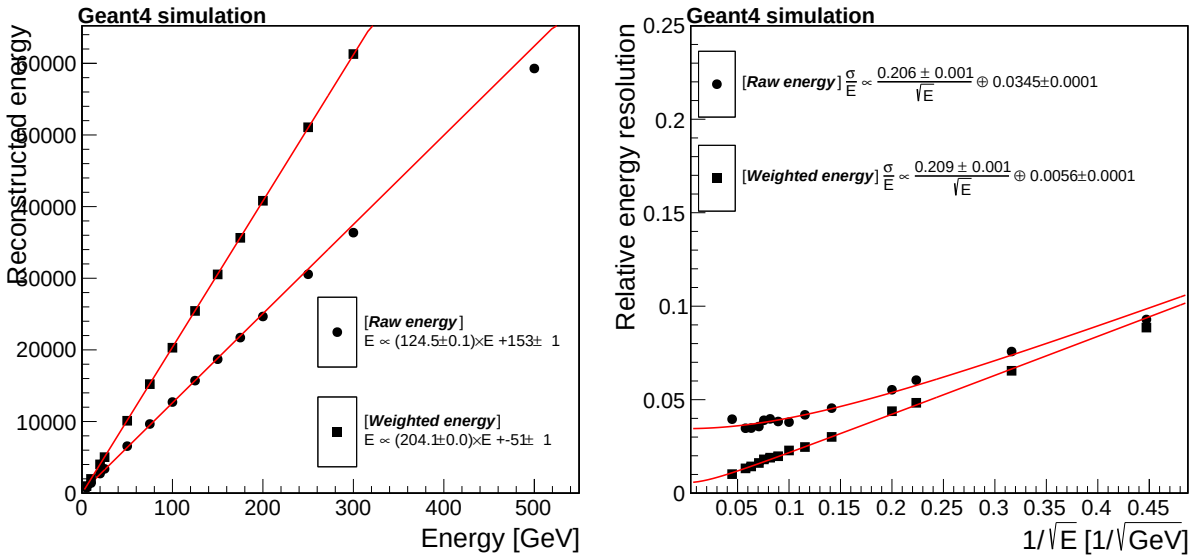

Figure 17: *Left*: reconstructed energy (in MIP units) as a function of the generated energy E. *Right*: energy resolution as a function of $\frac{1}{\sqrt{E}}$. In both cases single electron events are simulated.

Figure 19 shows the energy deposits distribution as function of the distance to the estimated shower maximum for different incident energies. The shower maximum is estimated on an event by-event-basis using the procedure described above. These distributions can be used to profile both the average energy deposit as well as the characteristic spread at each layer (or shower age).

Figure 20 summarizes the average energy profile (and energy fluctuation) of the showers for different incident energies. The plot on the *left* is shown as function of the thickness traversed. It can be observed that for incident energies below 5GeVthe shower maximum will occur on average in the first section of the detector. For energies up to 500GeVthe shower maximum will be contained in the second section of the detector.

The centered shower profile can, furthermore, be used to profile the average fluctuations. This is shown on Figure 21. The core of the shower is, as expected, less prone to statistical fluctuations with an intrinsic resolution of $< 20\%$ for $E > 20$GeV. The initial stage is prone to the largest fluctuations and the fine sampling at these earlier stages may therefore provide additional handle on the energy resolution. The last part of the shower, as it will be shown in the next Section, is expected to be dominated by the halo of the shower and, again, is prone to larger statistical

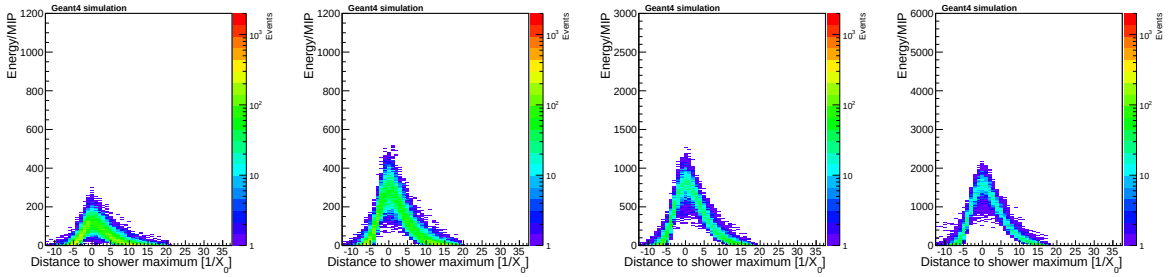

Figure 18: Energy distributions versus the distance in X_0 to the shower maximum for different electron energies. From *left* to *right* the incident electron energies are: 10GeV, 25GeV, 75GeVand 150GeV.

fluctuations with respect to the central region.

10.4 Occupancy

The occupancy has been studied by using pileup modeled with a Poisson-distribution with a mean of 140. The values obtained are illustrated in Fig 22 for a thresholds of 1 mip and 5 mips. It should be noted that 1 mip is equivalent to an energy of 5 MeV. We shall use such numbers to examine data transport from the detector, to the back of the HCAL and then to the surface. We are assessing different compression schemes. The occupancy per 1 cm^2 under conditions of pileup of a mean of 140 events for various eta values for a threshold of 1 mip (*Left*) and a threshold of 5 mips (*Right*).

10.5 Transverse shower evolution and shower containment

The generic properties of the evolution of the electromagnetic showers in the transverse plane can be studied by sampling uniformly each Si layer with a fine $2.5 \times 2.5\,\mathrm{mm}^2$ grid and summing the energy deposited in each cell. Figure 23 shows the evolution of two showers in the transverse plane for incoming energies of 50GeVand 150GeV. Using the radial distance $\rho = \sqrt{x^2 + y^2}$ with respect to the simulated electron axis we can examine the average energy deposited as function of the distance to the shower center as well as shower energy fraction contained within a given distance from the shower center. These results are illustrated, for the same incident energies, in Figures 24 and 25. One can observe that the showers are narrow; with 90% of the energy contained within 20 mm up to the shower maximum. After that the profile of the energy is less centered (halo-like) and the energy is spread over the plane. The 90% radius tends to increase exponentially after each layer.

The Molière radius can be extracted, i.e. the radius within which 90% of the shower energy is expected to be contained. Figure 26 summarizes the result obtained. We estimate the Molière radius to be 25.5 mm and the 68% containment radius to be 8.5 mm. If the air gap is increased to 4 mm (decreased to 1 mm) in the simulation the Molière radius is expected to increase to 31 mm (decrease to 22 mm) and the 68% containment radius is expected to increase to 10.5 mm (decrease to 7 mm).

We conclude this section with a study on the expected impact on resolution from summing the energy in cylinders of different sizes. For this purpose we sum energy which falls within a distance ρ

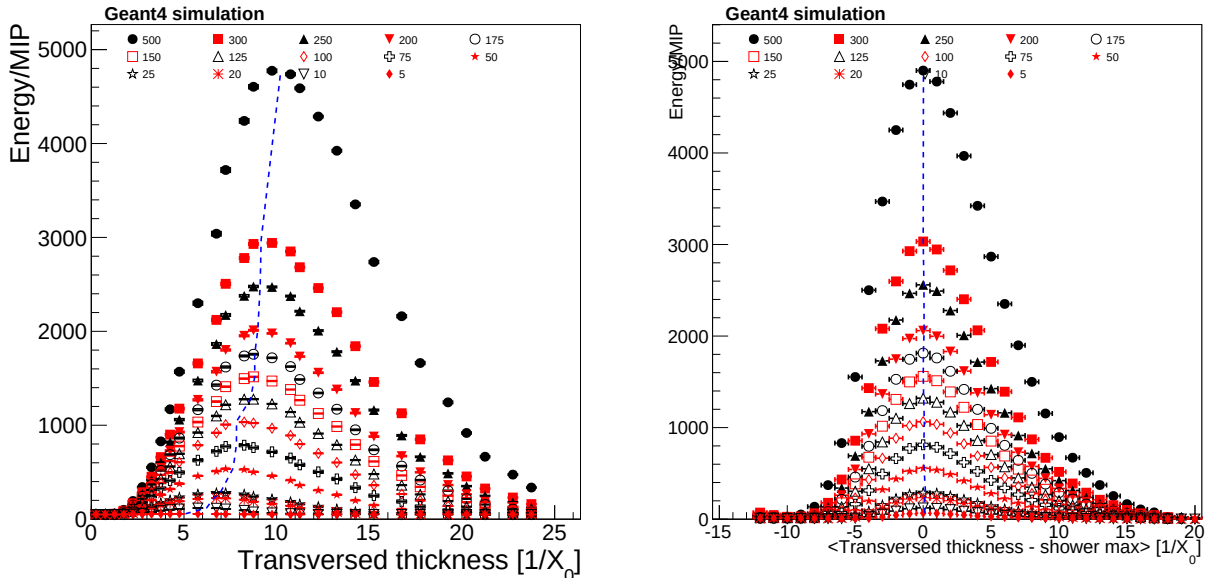

Figure 19: *Left*: average energy deposited as function of the transversed thickness in the calorimeter. A blue curve connects the estimated shower maximum. *Right*: average energy deposited as function of the distance to the shower maximum estimated on an event-per-event basis. *Right*: average relative width of the energy deposits as function of the distance to the shower maximum estimated on an event-per-event basis.

of the shower axis. The weighted energy estimator is obtained as described in the previous Section and the resolution of this estimator is evaluated in the same way. Table 3 gives an idea of the possibilities when different radii are used to collect the energy. An intermediate approach, using a dynamical range where the first layers are integrated with smaller cone sizes and later layers are integrated with larger cone, sizes following the expression $8.7 \times e^{0.064 \times \text{layer}}$ is leads to a degradation on the resolution of the order of 9%. These approaches will be explored in more detail in the presence of pileup.

Table 3: Electron energy resolution parameters for different radius of energy integration. See text for details.

ρ (mm)	$\left(\frac{\sigma_E}{E}\right)_{\text{stoch}}$	$\left(\frac{\sigma_E}{E}\right)_{\text{cte}}$
∞	0.209 ± 0.001	0.0056 ± 0.0001
7	0.237 ± 0.001	0.0090 ± 0.0001
25	0.216 ± 0.001	0.0069 ± 0.0001
dynamical	0.228 ± 0.001	0.0128 ± 0.0001

10.6 Variation of resolution with the silicon wafer thickness

In another study we investigated how varying the thickness of the silicon detector affects the resolution. The results are given in Table 4

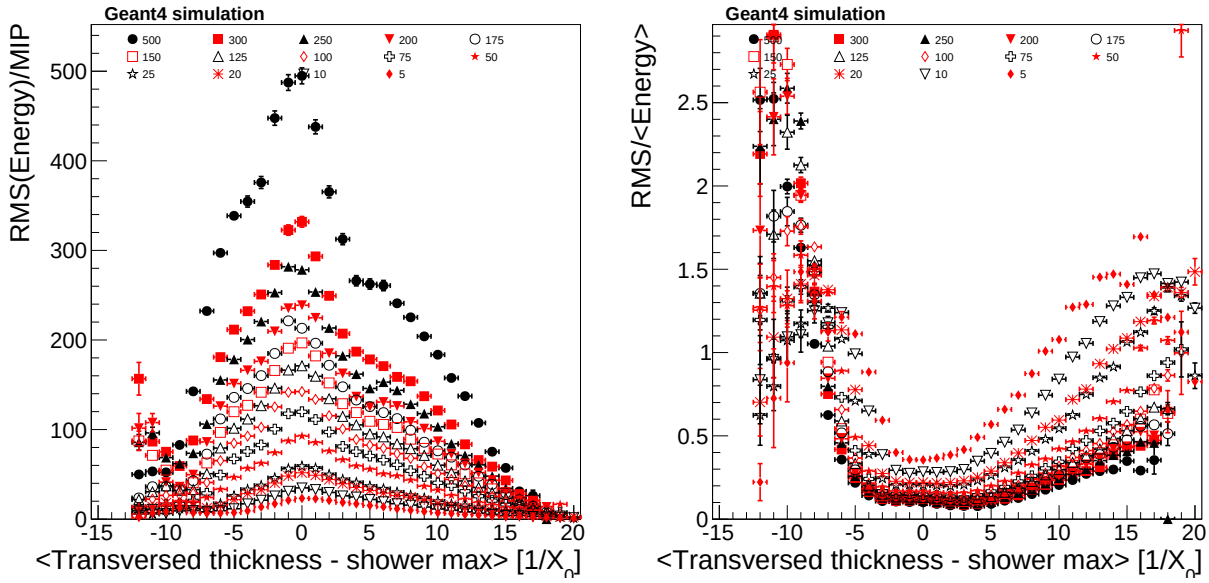

Figure 20: *Left*: average width of the energy deposits. *Right*: average relative width of the energy deposits. Both quantities are represented as function of the distance to the shower maximum estimated on an event-per-event basis.

Table 4: Variation of the stochastic and constant terms with the silicon detector thickness.

Silicon thickness (μm)	Stochastic term	Constant Term
80	0.248±0.001	0.0060±0.0001
120	0.228±0.001	0.0059± 0.0001
200	0.209±0.001	0.0056± 0.0001
500	0.183±0.001	0.0056±0.0001

10.7 Fast- and Full-Sim

The process of including this geometry into CMSSW is already well underway. This will allow us to use both Fast- and Full-Sim to study the physics performance. The concept geometry XML files have been generated, the detector description (DetId and numbering scheme) is implemented in CMSSW and SimHits from the HGC calorimeter are being produced. Software validation of the GEN-SIM step is being performed.

The power of a high granularity ECAL and HCAL can be fully exploited using a full particle flow reconstruction technique. The particle flow (PF) reconstruction makes full use of all sub-detector parts, with the tracker and calorimeter information used synergistically, and yields a much improved energy resolution and performance. In CMS we have a custom made PF reconstruction package with an excellent performance so far. Currently the CMS PF package is very much designed and tuned around the current CMS detector. On the same time, the large community of experts working on high granularity calorimetry for a Linear Collider over the past years has gained much expertise on these issues, and has developed many tools and code for a full and generic particle flow reconstruction. Such a tool is the PandoraPFA C++ software development kit [4] providing a

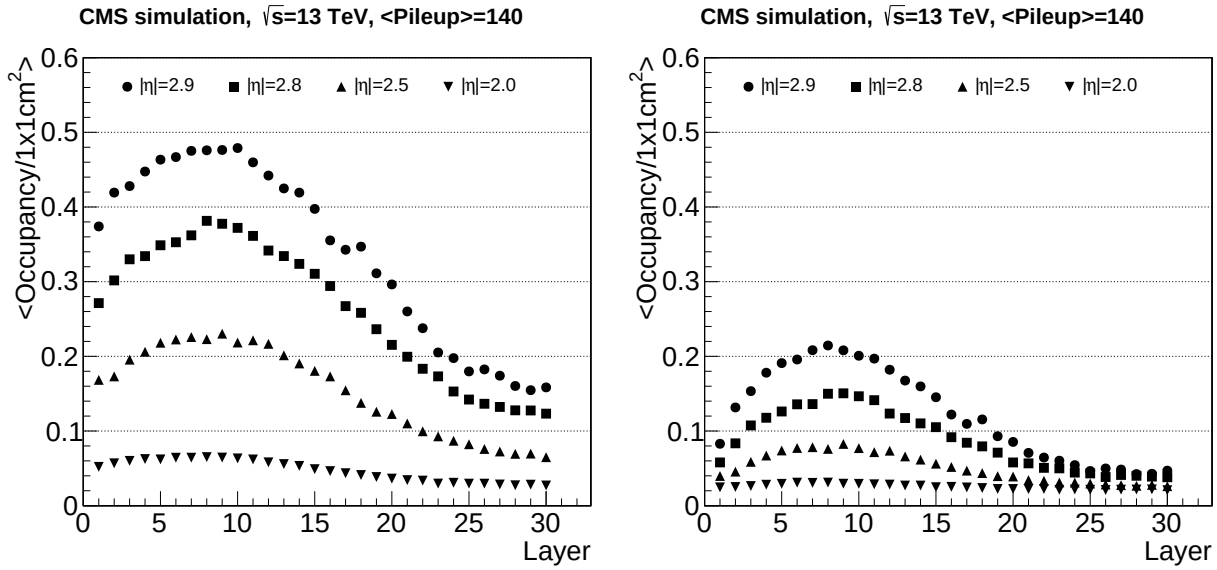

Figure 21: *Left*: Occupancy per unit cm^2 when threshold is 1 mip, *Right*: Occupancy per unit cm^2 when threshold is 5 mips

highly sophisticated PF reconstruction for high granularity detectors which is flexible and can be incorporated by various user applications and detector configurations. The PanforaPFA package is widely used by almost all ICL/CLIC studies, and by other high energy physics experiments as well (MIcroBOONE). We have already began the process of integrating a full particle flow reconstruction, interfacing PandoraPFA into the CMSSW. The needed information regarding calorimeter hits, tracks and the geometry information are already in place, and we currently have a fully running version using the current CMS detectors. We have produced very preliminary first results using single electron, pion and muons events with encouraging results: Particle flow objects have been successfully reconstructed linking tracks to calorimeter clusters composed of several calorimeter hits. We plan to fully debug and tune the PandoraPFA during the next months in order to be in a position to perform detailed physics studies in the Summer.

11 R&D Programme

Small high performance silicon-based calorimeter have been built. However, for operation at the LHC the major questions major questions to be addressed by an R&D program are those related to radiation hardness, engineering and system design. It will nevertheless be necessary to construct a prototype to begin to test out some of the main ideas by the middle of 2015.

There are several components to the R&D program that will need to be started before the TDR, these are:

- Thermal model: A thermal model to benchmark the detector cooling to be compared against the thermal FEA.
- Maquette or scale model to establish the minimum separation between plates of the absorber.

- Testing of prototype calorimeters to benchmark the detector simulations.
- Radiation tests of silicon sensors and if feasible installation after irradiation in the prototype to quantify the effects of radiation on the performance of the calorimeter.

11.1 Mechanical Models

As the thermal engineering is a critical component of this design we will need to model and measure the cooling and the heat flow in the detector. This can be done with detailed ANSYS calculations, but needs to be benchmarked with a thermal model built with the a CO_2 or a C_3F_8 cooling system and heating elements. The construction of a maquette to properly understand space limitations absorber gap will be a very useful step and can be made without significant resources.

11.2 Sensors

There are many data available on tests conducted on sensors using proton irradiations up to very high fluences. These need to be complemented with neutron irradiation up to the same high fluences. In collaboration with ongoing tests of Si sensors we plan to irradiate different samples (Si thickness, types e.g. epitaxial etc.) and to study any phenomena that might affect performance such as charge collection, multiplications, etc. We will use available diode test structure from the "Hamamatsu Campaign" in order to extend the ongoing Tracker sensor R&D, in particular to include the very high neutron fluences characteristic of the high eta region of the End-Cap Electromagnetic calorimeter, with a particular focus on those devices with depletion thickness of 200 μm and 100 μm.

Of particular relevance are the bias voltage dependence of the sensor leakage current and signal response: charge collection and any possible excess noise factor and/or non-Gaussian noise behaviour.

Should the onset of any impact charge ionisation effects, and the resulting charge multiplication be observed, these will be studied in detail.

11.3 Beam Tests

There are many data available from the tests conducted by linear collider collaborations (e.g. CALICE) against which we are benchmarking our simulations. We are investigating what tests are planned by these collaborations and may propose some dedicated tests.

11.4 Medium Scale Prototype

Another major step is the construction of a medium scale prototype, which will be used to study the combination of the EE and FE systems in a beam. Data from this test beam can be used to benchmark simulation results. For this a prototype calorimeter sufficiently large to contain hadronic showers will need to be constructed and tested in a high-energy beam to evaluate the detector performance. A useful prototype could consist of a small EE section and a larger FH section, with a backing calorimeter placed behind it. The EE section could be built with a layer of sensor hexagonal sensors cut from 6" wafers, with sampling layers as detailed in section 2.1. The FH section could be made with layers of seven tiled hexagonal sensors with the sampling fraction

as described in Section 2.2. The backing calorimeter one of several existing prototype calorimeters and would not need to built from scratch.

It will need to have a sufficiently fine granularity to demonstrate the event reconstruction and it should allow for several possible detector configurations and granularities to be tested in a medium scale system. The time-scale of this prototype is set by the date of the TDR, for which test results will be necessary.

To equip the prototype electronically one possibility would-be to design a readout using radiation hard components developed for CMS. Specifically, the PACE chip of the pre-shower detector and the AD1240, the 4 channel ADC developed for the ECAL, and the GOL. Sufficient components are available to equip the prototype calorimeter. The off-detector electronics can be made using commercial components.

12 HGCAL Collaboration

This collaboration has only been active since the CMS Upgrade Week in late October 2013, when the idea of using silicon sensors for a high granularity calorimeter system became a possibility following discussions with a major silicon sensor manufacturer. Since then, CMS members from the institutions listed on the first page, from several different subsystem groups within CMS, have worked together to define this option for CMS. In this we have been aided by many colleagues from the CALICE and ILC/CLIC communities, and by engineers at our own and other institutions, which has allowed us to develop this idea quickly. Moreover, we have had discussions with colleagues at several of the national labs, not listed on the first page, who have shown a strong level of interest in this project.

References

[1] https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome

[2] T. Behnke *et al. (ed.)*, Reference Design Report "Volume 4: Detectors" (2007), available at http://lcdev.kek.jp/RDR/

[3] C. Adloff *et al.* "Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons." Nucl. Instrum. Meth. A **608** (2009) 372 - 383.

[4] M. Thomson, "Particle flow calorimetry and the Pandora PFA algorithm," Nucl. Instrum. Meth. A **611** (2009) 25 - 40, and J.S. Marshall, A.Munnich and M.A.Thomson, "Performance of particle flow calorimetry at CLIC," Nucl. Instrum. Meth. A **700**(2013)153 - 162.

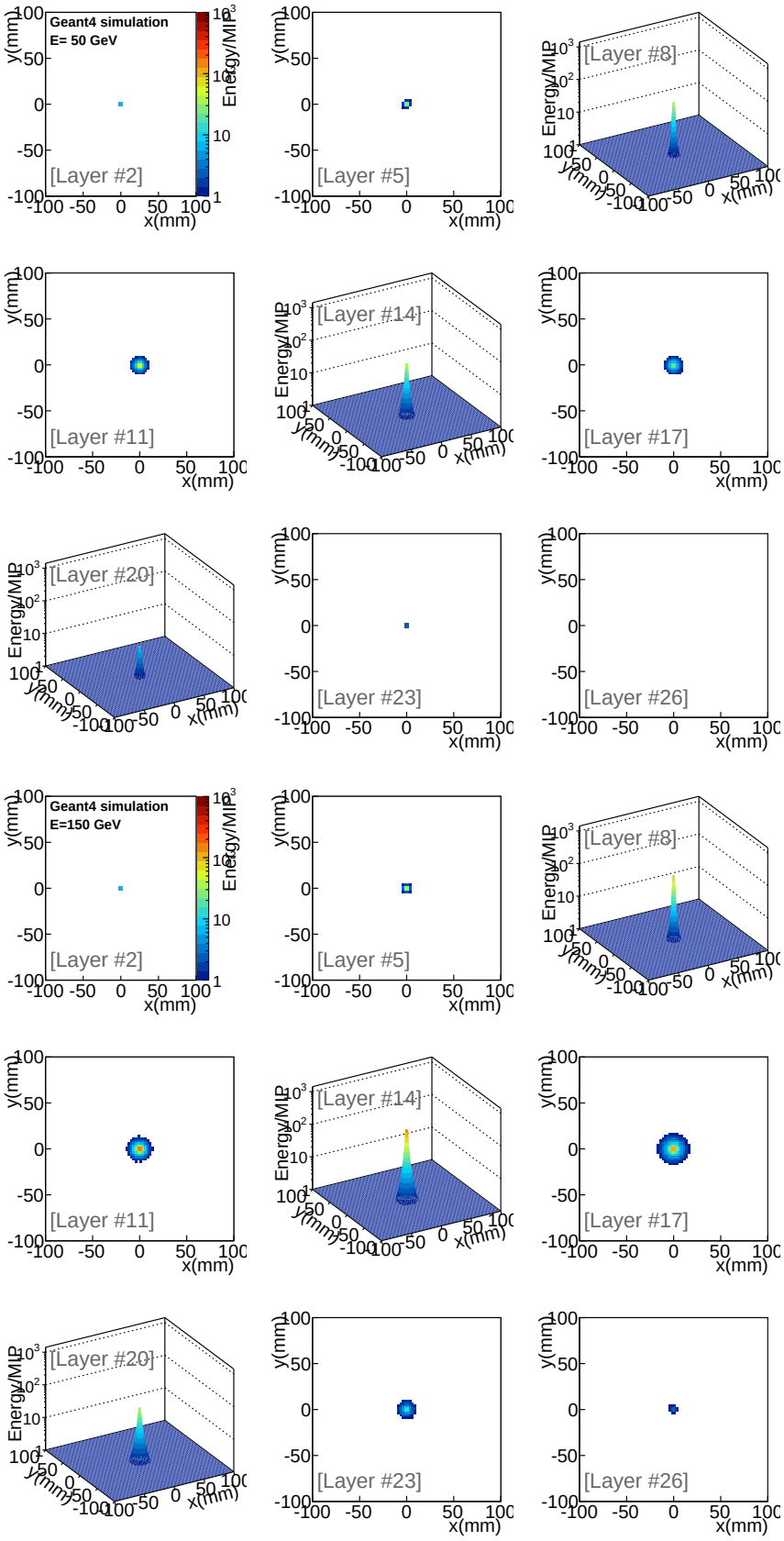

Figure 22: Average evolution of electromagnetic showers in the transverse plane for selected Si layers. The per-sector weights are applied to the energy measured in MIP equivalents. Incoming electrons with E=50GeV(150GeV) are shown on *top* (*bottom*).

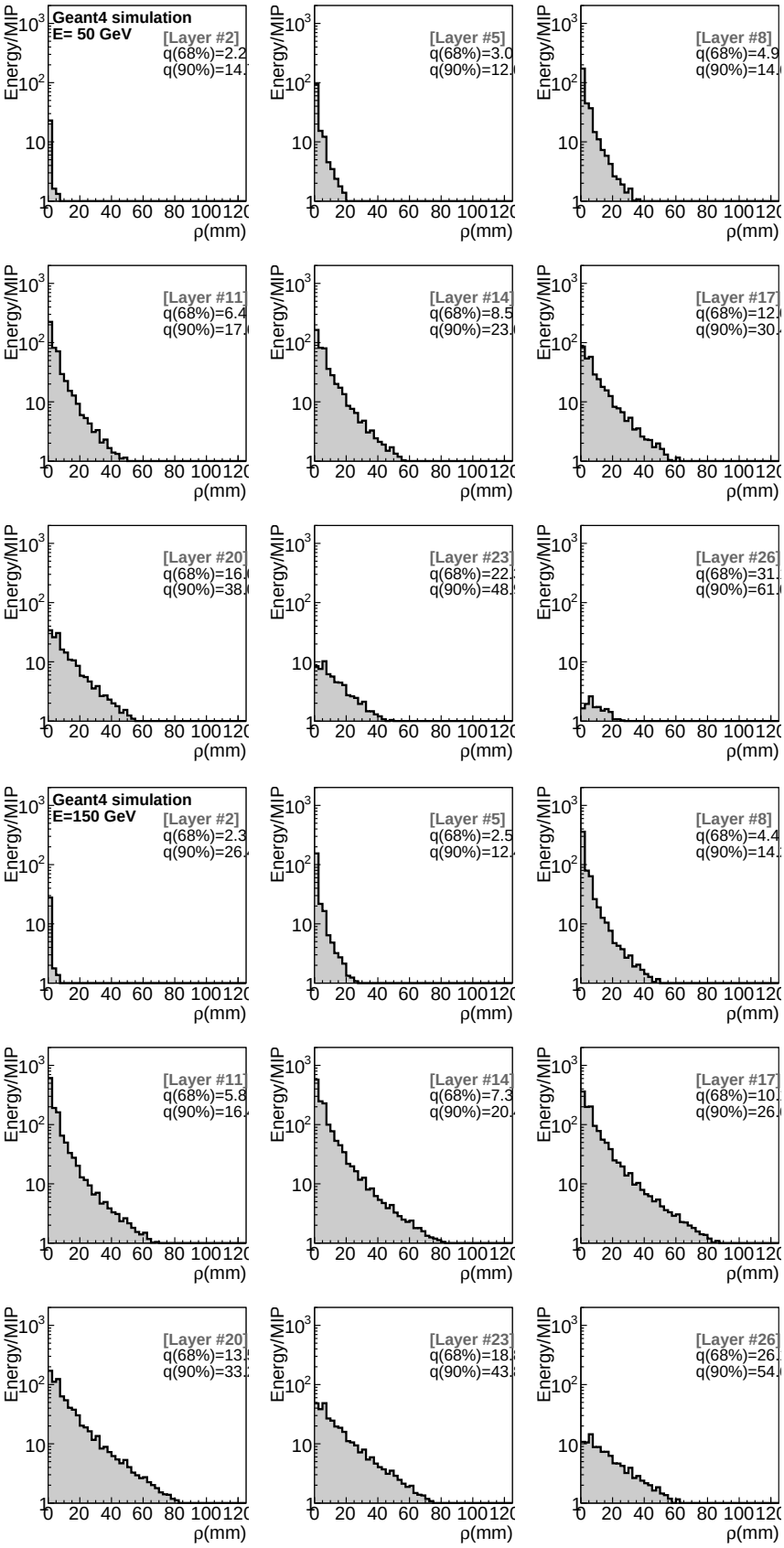

Figure 23: Average energy deposited as function of the distance to the shower center. Incoming electrons with E=50GeV(150GeV) are shown on top (*bottom*). In the captions, the distances corresponding to the 68% and 90% quantiles of the distributions are quoted.

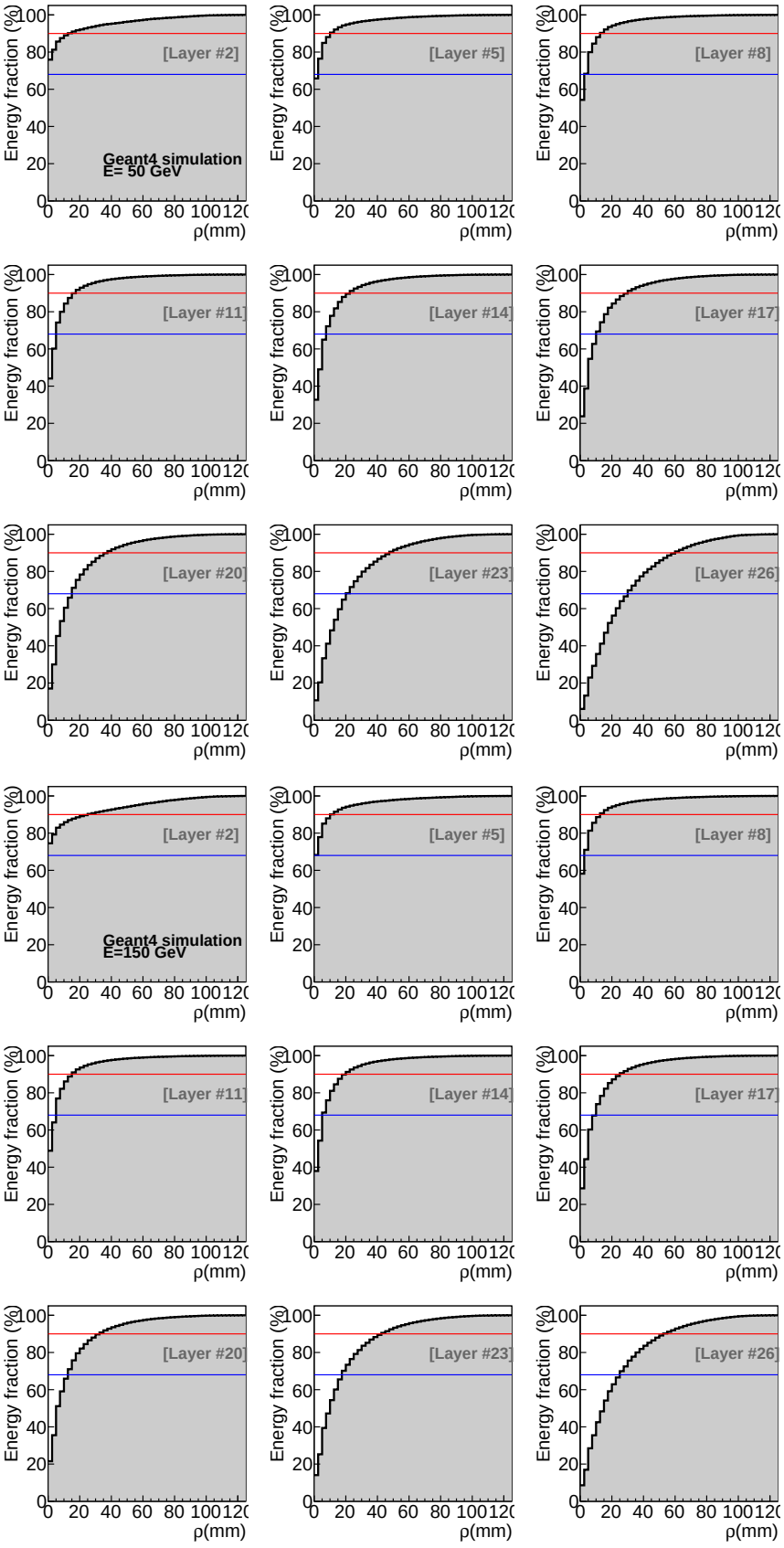

Figure 24: Average energy fraction contained within a given distance to the shower center. Incoming electrons with E=50GeV(150GeV) are shown on *top* (*bottom*). The blue and red lines mark the radius for which 68% and 90% of the energy are contained, correspondingly.

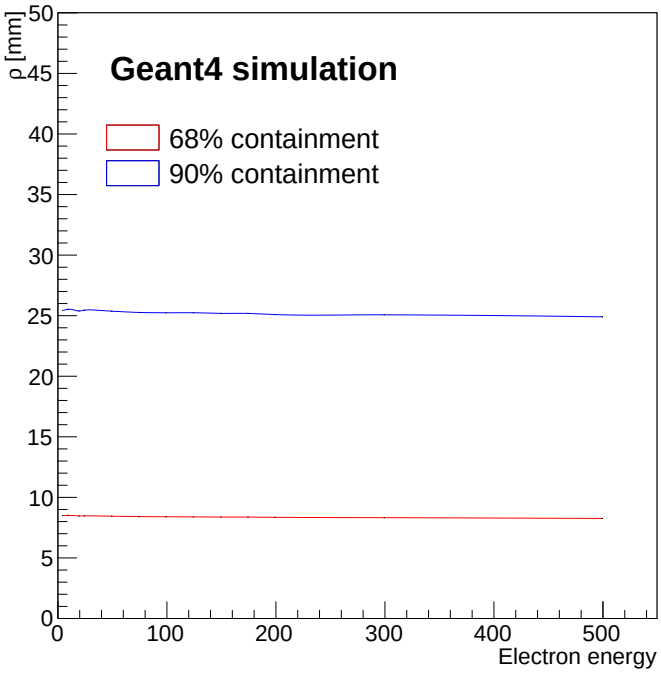

Figure 25: Estimated Moliére radius and radius for 68% containment of the electromagnetic shower energy.