Chapter 7
The Design of the CMS Upgrade Tracker and the CMS High Granularity Forward Calorimeter Equipped with Silicon Sensors for the HL-LHC

The CMS detector has been designed and realized to cope with an integrated luminosity of about $\mathscr{L} = 300\ \text{fb}^{-1}$ with some margin and an instantaneous luminosity of $\mathscr{L} \sim 10^{34}\ \text{cm}^{-2}\text{s}^{-1}$. Subsequently, for the inner detectors: Tracker and forward calorimeter this also marks the end of life due to detrimental radiation damage. Both detectors could operate beyond $\mathscr{L} = 500\ \text{fb}^{-1}$ but by far not up to integrated luminosity $\mathscr{L} = 3000\ \text{fb}^{-1}$ – the goal of the **H**igh **L**uminosity LHC HL-LHC operation.

For the next stage – Phase II, planned start 2026 – the HL-LHC, the peak luminosity will be $\sim 0.5 \cdot 10^{35}\ \text{cm}^{-2}\text{s}^{-1}$, 5 – 10 times the design peak luminosity of the LHC, and even higher values are under consideration. This has the following implications for the CMS upgrades: (1) increased radiation and (2) higher **p**ile-**u**p PU[1], therefore increased occupancy. Figure 7.1 presents the expected integrated fluence Φ_{eq} in $n_{1\text{MeV}}/\text{cm}^2$ fluence after an integrated luminosity of $\mathscr{L} = 3\ \text{ab}^{-1}$ at the end of HL-LHC operation. The particle track density scales similarly for different radii.

As seen in Fig. 7.1 the radiation levels at $R = 4$ cm (inner most pixel) will be around $\Phi_{eq} = 2 \cdot 10^{16}\ n_{1\ \text{MeV}}/\text{cm}^2$, decreasing to around $\Phi_{eq} = 10^{15}\ n_{1\ \text{MeV}}/\text{cm}^2$ at $R = 20$ cm (Inner radius of **O**uter **T**racker OT) while at $R = 50$ cm still around $\Phi_{eq} = 3 \cdot 10^{14}\ n_{1\ \text{MeV}}/\text{cm}^2$ has to be endured. The increase in number of events/tracks due to pile-up for the different operation stages of LHC to HL-LHC is illustrated in Fig. 7.2.

This, of course, defines/limits possible technologies versus radii. Average pile-up can be up to PU = 200 events per crossing at the HL-LHC. To achieve the goals defined for the HL-LHC phase, a detector performance at least as good as the one of the current CMS detector is required. Given these requirements CMS opted for an all-silicon tracker plus a tungsten/copper-silicon sampling forward electromagnetic calorimeter, followed by a stainless steel-silicon sampling hadron calorimeter. The **H**igh **G**ranularity **C**alorimeter HGC or HGCAL is based on silicon pad detectors. In the Phase II Tracker and HGC cases, the radiation hardness and technical

[1] Pile-up: number of proton-proton collisions in a single bunch crossing: $\langle PU \rangle = 20$ (LHC design); >50 (LHC reality); 200 (HL-LHC expectation).

© Springer International Publishing AG 2017

F. Hartmann, *Evolution of Silicon Sensor Technology in Particle Physics*,
Springer Tracts in Modern Physics 275, DOI 10.1007/978-3-319-64436-3_7

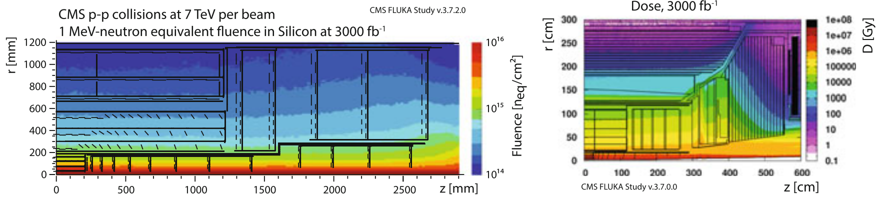

Fig. 7.1 Maps of expected particle fluence corresponding to an integrated luminosity of $\mathcal{L} = 3000\ \text{fb}^{-1}$, expressed in terms of 1 MeV *neutron equivalent fluence* for the Tracker on the *left* or dose for the HGC case on the *right*. *Left* Tracker volume. The expected fluence has a strong dependence on radius, while it is almost independent of the z-coordinate [303]. *Right* High Granularity Calorimeter [308]. Radiation from charged particles increases much faster than $1/r^2$; this behaviour derives from the magnetic field responsible for coiled tracks – this is especially significant at lower radii. At *lower radii*, the radiation is dominated by charged particles while at higher radii neutrons, backscattered contribution from the calorimeters, dominate. The break-even is around $r = 50$ cm. To achieve a reasonable occupancy it can be imagined to instrument 4 cm $< r <$ 20 cm with micro-pixels, 20 cm $< r <$ 50 cm with macro-pixels and $r > 50$ cm with single strip sensor modules

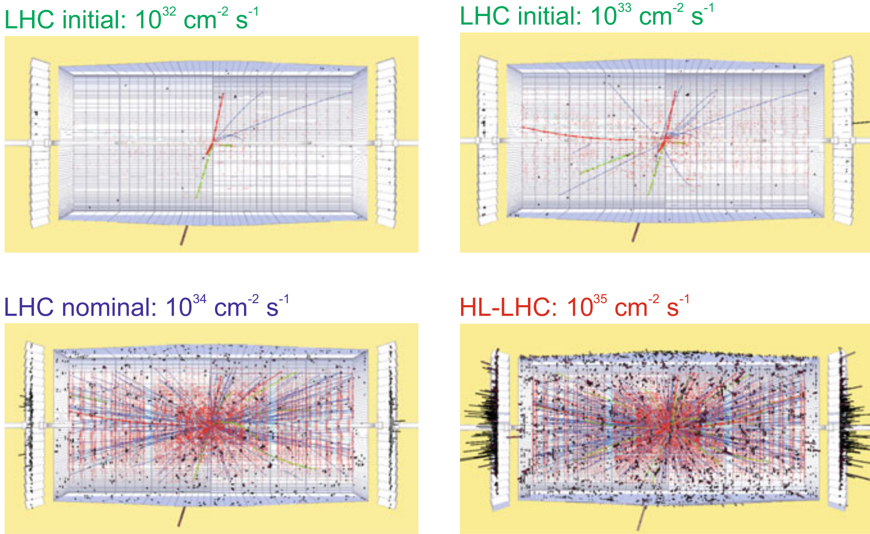

Fig. 7.2 Event pile-up versus luminosity. The current plan for the HL-LHC peak luminosity is $\mathcal{L} \sim 0.5 \cdot 10^{35}\ \text{cm}^{-2}\text{s}^{-1}$, the *last* figure shows about twice this value [346]

structurability of sensors with respect to cell-size is key. These new detectors will largely strengthen the **P**article **F**low PF concept where the particles and their property are being followed/reconstructed along the entire path in all sub-systems – the flow of particles. Initial recipes of ten times more radiation tolerant sensors came from RD50 but extensive R&D campaigns have been conducted within the CMS collaboration as well. Shorter strips and smaller pixels with respect to the current tracker and smaller pads with respect to the current crystal forward calorimeter cells will cope with the much denser particle environment.

In addition, with the ∼10-fold increase of luminosity, a simple increase of existing trigger thresholds (muons and calorimeter) to keep trigger rates below CMS capability is not an option because limits are determined by physics needs and higher thresholds (e.g. on energy) would cut directly into physics performance. A direct implication is the need of a high p_T track trigger at Level-1 for the HL-LHC operation – a complete novelty.

The forward silicon sampling calorimeter concept (HGC) has been inspired by the CALICE [288, 320, 335] idea (**CA**lorimeter for **LI**near **C**ollider **E**xperiment). In addition to strengthening the **p**article **f**low PF concept, the HGC will add timing information for charged and neutral particles.

As of 2017, both designs are in a mature state but changes/improvements are still expected. A snapshot of both, the design of the future Tracker and the HGC, will be described in the next sections. A much more detailed description of the full CMS upgrade plans can be found here [308]. Both detectors will be installed in long shutdown 3 LS3, scheduled for 2024 – 2026.

7.1 The CMS Tracker Upgrade for the HL-LHC – Phase II

2017, after publishing the Technical Proposal TP [308] plus scope document [302] and just before the **T**echnical **D**esign **R**eport TDR [303], the Phase II Tracker design is fairly advanced but details might still change. This section will therefore concentrate on the challenges and the solutions and sketches the earlier evolution to reach the current state. All numbers of e.g. number of modules or layer radius are only good approximations of the future final implementation.

The requirements for the CMS Phase II Tracker at the HL-LHC are:

- 10 times the radiation tolerance of the current detector, namely integrated luminosity $\mathscr{L} = 3000\ \text{fb}^{-1}$ resulting in close to $\Phi_{eq} = 1 \cdot 10^{14}$ to $2 \cdot 10^{16} n_{1\,\text{MeV}}/\text{cm}^2$ in the outer Strip layer and innermost Pixel layers accounting for about 1 GRad of dose.
  - the possibility to replace the innermost pixel layer in a Year-End-Technical-Stop
- increased cell granularity to maintain an occupancy $<1\%$ at 140 – 200 PU
- contribution to the Level-1 trigger (L1) – **O**uter **T**racker OT only
  - p_T-modules, which can intrinsically derive if a traversing particle's momentum is above 2 GeV, in the Outer Tracker
- much longer Level-1 trigger latency: 12.8 μs (today 6.4 μs)
- much higher readout rate: 750 kHz – 1 MHz (today 100 kHz)
- extended tracking acceptance: up to about $|\eta| \leq 4$ – concerns mostly the pixel detector (today $|\eta| \leq 2.5$)
- systematic minimization of material budget in the tracking volume

- robust pattern recognition – 'macro-pixels' in Outer Tracker and smaller pixels at low radii
  - improved two-track separation for improved track finding in high energetic jets
- supply much higher power, handle higher data rate and allow for lower temperatures with about the same volume of services (cable channels).

In order to maintain adequate track reconstruction performance at the much higher pile-up levels of the HL-LHC, the granularity of both Outer and Inner Tracker will be increased by roughly a factor four to six. The pitch in the outer strips section will stay roughly the same while the strip length will be reduced (20 to 5 cm and 10 to 2.5 cm and even to long macro-pixels of 1.5 mm length). The inner pixel cell size will be reduced by a factor six from 15000 μm^2 (today – LHC era) to 2500 μm^2. Especially the later reduction will improve two-track separation in high PU environment and impact parameter resolution. Many novel design choices ensure a significant reduction in material budget leading to reduced multiple scattering thus significantly improved p_T-resolution and a lower rate of γ conversion and less Bremsstrahlung[2] for electrons.

With the help of a dedicated tool, "TKlayout"[3] [38, 124, 162], many different layouts have been evaluated and the Tracker layout has been optimized. TKlayout calculates tracking parameters (p_T and impact parameter resolution, number of hits, number of Level-1 stubs, etc.) for offline as well as for L1 tracking. It calculates material budget, radiation lengths and projects leakage current and power. Many of the later design choices and optimizations, e.g. sensor spacing, module locations, usage of square modules in the endcaps, are based on TKlayout modelling and have finally been cross-checked with full simulation.

The perfect Tracker layout would be a sphere, where particles traverse the sensors always perpendicular passing the least amount of material (maybe with a slight angle to hit 2 cells) and would be fully hermetic. The future CMS Tracker is approaching this concept as much as possible in an overall cylindrical envelope. The ingredients, like 2S and PS[4] modules will be introduced in the next section. The full optimization is being done by several people and lots of input from all members of CMS during several years. Figure 7.3 demonstrates the design evolution starting from the current CMS Silicon Strip Tracker design – a small subset of design steps are being described. Outer Barrel configurations from 4 to 16 layers have been studied. Early on studies demonstrated that rectangular instead of wedge-shaped modules can be used in the forward direction with no significant penalty in mass and resolution but large reduction in system complexity (number of module types). An early proposed design featured a barrel-only layout, but as mentioned earlier a sphere would be the perfect design, subsequently the usual barrel-endcap design has been adopted. In a

[2]Often simply called Brems.

[3]3D tool adopting a non-simulative, parametrization-driven approach to tracker performance evaluation, taking e.g. multiple scattering as measurement error into account. It has been wetted against full simulation but it always provides the best case performance being independent of specific tracking algorithms. TKLayout runs more than 100 times faster than full simulation.

[4]2S = **S**trips+**S**trips sensors; PS = **P**ixel+**S**trips sensors.

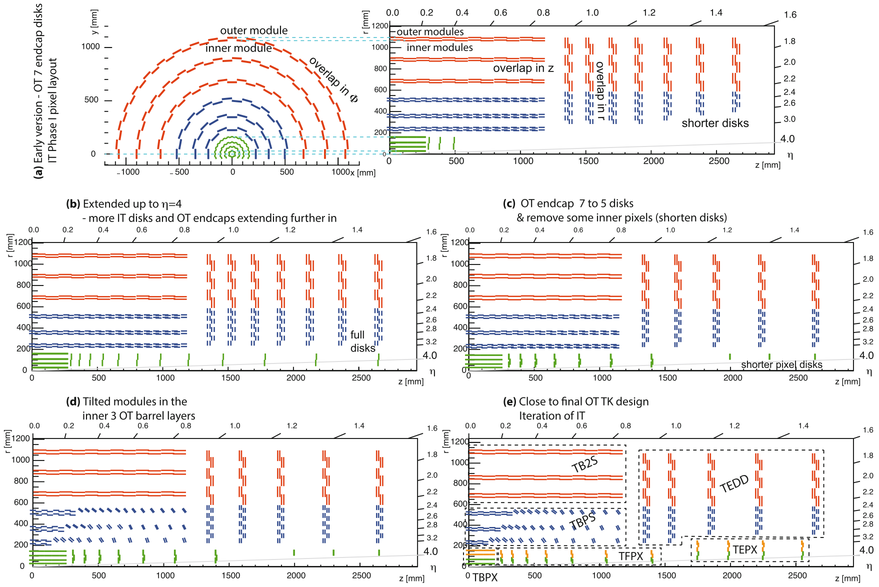

Fig. 7.3 The figures show the evolution of the Phase 2 Tracker layout. Figure **a** shows an early design in $r\Phi$ and rz view where also the module overlaps of inner and outer modules, inside one layer, are discernible. 2S modules are represented by red lines and PS modules by blue ones (each line represents one module) while micro-pixels layers are shown in *green* and *orange*. Initially 3+3 barrel layers and 7 endcaps plus the phase I pixel configuration – 4 barrel plus 3 endcaps have been evaluated. Also 4(2S)+2(PS) (without explicit picture) have been evaluated earlier. As a next step (figure **b**) 9 pixel full-size endcap disks were added (totalling 12 disks) to cover ranges up to $|\eta| = 4$ to have a first glimpse about performance. Consequently also the last three OT endcaps where extended to lower *radius*. In a later step (figure **c**) two OT endcap disks were removed without penalty in performance, as well as some pixel disks were shrunk to adapt to reality, namely a conical beam-pipe otherwise clashing with the detector. As further major step (figure **d**), and after many studies of mechanical feasibility, the larger part of the PS modules have been tilted. This saves a lot of modules thus material and monetary budget. The tilt also improves trigger performance at larger z – more in Sect. 7.1. Figure **e** shows the almost final layout where the IT tracker has been rationalized to allow insertion with beam pipe present – more in Sect. 7.1. The final TDR layout is presented in detail in Fig. 7.4

next step several additional pixel disks have been introduced allowing instrumentation up to $|\eta| = 4$. Several optimization steps later the number of outer endcap disks were reduced from seven to five while it became clear that the OT barrel section needs 3 pixelated plus three strip layers (initially only two PS but four 2S layers had been evaluated). A large TEDD endcap disk has 15 rings (six 2S plus nine PS). The PS module concept, with a very good intrinsic z-resolution due to the short 1.5 mm macro-pixels, also avoids any necessity to have stereo modules, as present in the current tracker. As a side remark, with the p_T module concept, stereo modules are anyhow not applicable since strips/macro-pixels need to be parallel. The sphere concept came back to mind and the modules in the inner three layers have been tilted within mechanical possibilities by 40 to 74°. This concept saves about 1300 PS

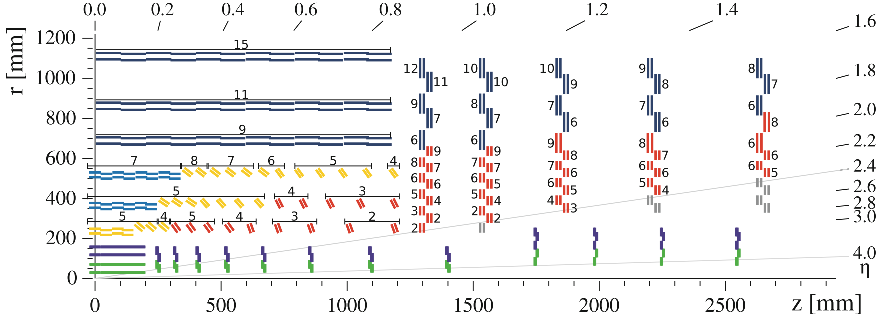

Fig. 7.4 The CMS Tracker TDR layout with emphasis on sensor spacing and track trigger 'acceptance windows'; a quadrant is shown. All OT modules are p_T modules with 2S at $r > 60$ cm and PS modules at $20 < r < 60$ cm. The sensor spacing is colour coded with *dark blue* for 1.8 mm, *light blue* for 1.6, *yellow* for 2.6 and *red* for 4 mm spacing; *grey modules* ($|\eta| > 2.5$) do not participate in the trigger. The numbers represent the 'number of strips in the acceptance windows', these are tunable during running (values here only exemplary or considered as starting point). Comparing the TEDD shown here with Fig. 7.3**e**, some optimization can be spotted resulting in better PS coverage and about 260 modules savings – short and long disks at different z now follow slightly different designs to allow for more space of the Inner Tracker IT forward pixels (*TEPX*). TEDD double-disk 1 moved closer to the barrel – optimise spacing for hermeticity. The IT micro-pixels consist of 4 barrel layers, 8 forward plus 4 extended forward disks covering up to $|\eta| > 4$. The *light green* pixel layers/disks consist of 1×2 chips modules while the violet ones consist of 2×2 chip modules. The tracker, all in all, has seven pixel layers – 4 micro- and 3 macro-pixels [303]

modules thus reduces material budget significantly[5] and a good amount of money. It will be later demonstrated that the change also has a very significant and positive impact on the Level-1 track trigger. There were more subtle changes to optimise the tracking and hermeticity, e.g. moving the first (at low z) TEDD double-disc closer to the barrel section or changing which modules of the TEDD or TB2S are located closer in and which further out (orientation) or overlaps of the modules. Figure 7.4 presents the layout at the time of the Technical Design Report TDR, still with some potential for future tuning. The necessity of having tracking information at the Level-1 Trigger also influenced the design strongly; more details later.

Outer Tracker – Strips and Macro-Pixels

The requirement to contribute to the trigger at Level-1 has a dramatic impact on the future CMS Tracker design, strongly defining the module concept and partially the full layout of tracker. The L1 trigger functionality was initially considered 'THE' challenge, but in the end allowed a very elegant, advantageous and powerful tracker design.

Dedicated p_T-modules in the **O**uter **T**racker OT have been developed, being able to discriminate intrinsically between high ($p_T > 2$ GeV) and low momentum traversing particles. Only the information from high p_T tracks, largely reduced with respect to the full data, is being transmitted for each crossing (40 MHz) and being used for the L1-Trigger – only from the Outer Tracker. The later offline tracking

[5] In the order of 0.2 x/X_0 in the range $1 < |\eta| < 2.5$.

exploits **I**nner **T**racker IT (Pixels) for track seeding and then utilizes mainly the large lever arm (large radius) of the Outer Tracker for momentum resolution; here the number of layers in the OT is less important, as long as it is above the critical number plus one for redundancy. The minimum number of OT layers is therefore defined by the L1 functionality but, on the other hand, the extra novel pixelation in the first three OT layers makes a fifth pixel layer unnecessary.

Being the basic ingredient, the modules and their p_T threshold capability will be introduced first. Every module in the Outer tracker consists of two narrowly spaced sensors with parallel strips(pixel) where the front-end chips receive signals from both sensors thus being able to correlate signals. High p_T tracks, bent in the strong magnetic field of CMS, are straighter and therefore pass in a narrower band of strips (macro-pixels) of the upper sensor – in the 'selection window', while lower momentum particles, bent stronger, miss the window. A pair of hits in the two sensors from one particle passing the correlation criteria is called a 'stub'. Figure 7.5 illustrates the concept. The 3.8 T field allows a relatively close spacing with 1.6, 1.8 and 2.6 mm distance and 4 mm. The centre and width of the 'selection window' is programmable and therefore adaptable/tunable after installation, for different locations in the tracker, adjusting for e.g. different z-position thus longer traversing path length or higher radius r. The 'stub' information is then transmitted for all bunch crossings at 40 MHz; about 100,000 stubs per bunch crossing corresponding to about 3% of all hits. The high magnetic field at 3.8 T of CMS, and consequent strong bending power, is a real advantage allowing p_T discrimination with narrow sensor spacing.

The module concepts are displayed in Fig. 7.6. The two module types are **S**trip-**S**trip 2S modules assembled from two strip sensors and **P**ixel-**S**trip PS modules

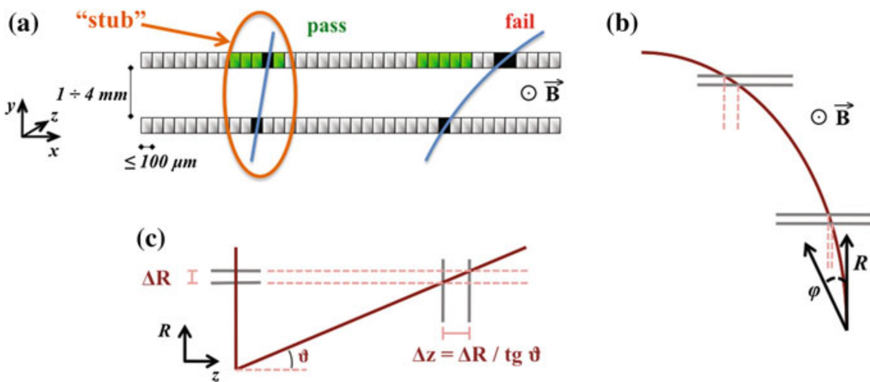

Fig. 7.5 **(a)** Correlation of signals in closely-spaced sensors enables rejection of low-p_T particles; the channels shown in *light green* represent the 'selection window' to define an accepted 'stub'. **(b)** The same transverse momentum corresponds to a larger displacement between the two signals at large radii for a given sensor spacing. **(c)** For the endcap disks, a larger spacing between the sensors is needed to achieve the same discriminating power as in the barrel at the same radius. The 'selection window' can be tuned along with the sensor spacing to achieve the desired p_T filtering in different regions of the detector [308]

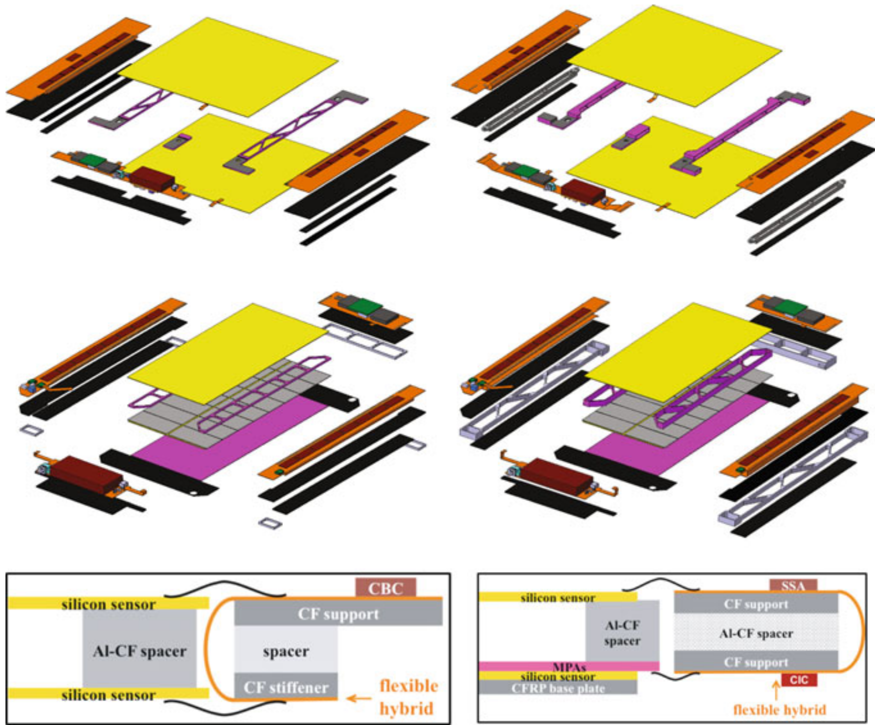

Fig. 7.6 The figures show the component view of modules with 1.8 (1.6) and 4 mm sensor spacings; the *upper* ones 2S and the *lower* ones PS modules (the 2.6 mm PS is not shown) [303]. The *yellow* rectangles are the strip sensors, the *violet* the pixel sensors with their *grey* MPA ASICS on *top*. The spacers are of Al-CF (**A**luminium **C**arbon **F**iber) a new material for stability, machinable and very good isotropic heat transfer – mind the different designs to allow different spacings with lowest material budget (alternatives are still under investigation). The CF (*black*) is for mechanical stability and thermal management, especially necessary as base plate below the pixel-ASIC sandwich. The readout hybrids are located on the *left* and *right* with eight CBC ASICs each plus a **c**oncentrator **c**hip (CIC); the whole ASICS-flex, 'CF – Al-CF-spacer – CF' sandwich including the flex bending will be fully delivered from industry. The service hybrid is situated on the *lower end* for the 2S module and split in two for the PS module. Realized in the technology as the readout hybrids, it contains DC-DC converter for powering and the optical readout. The *lower* figure shows a cut-view illustrating how the flexes are bent, the signals are routed and where the ASICs are located (*left* 2S module and *right* PS module). One can see where and how these modules are placed in Fig. 7.7

assembled of a pixelated (macro-pixel) and a strip sensor allowing for a decent z-resolution in the barrel. The 2S modules come with 1.8 and 4.0 mm sensor spacings and are mounted at $r > 60$ cm. They feature 2 times 8 chips with 254 channels each connected to 2 sensors with 2 times 1016 strips of length 5 cm and pitch 90 μm. The even chip channels are connected to the upper and the odd channels to the lower sensor allowing the above described in-chip correlation logic. The signals are routed through the flex hybrid being bent around the hybrid body. The total power of this

module type after 3000 fb^{-1} will be around 5 W. The readout chip is called **C**MS **B**inary **C**hip CBC, a subsequent continued development from the current APV-ASIC.

The PS modules are located at 20 cm $< r <$ 60 cm and come with sensor spacings of 1.6, 2.6 and 4 mm. The strip sensor features 2 × 960 strips with 2.5 cm length and a pitch of 100 µm connected via wire-bonding at the end to the **S**hort **S**trip **ASIC** SSA[6]. 16 **M**acro-**P**ixel-**A**SICs MPA with 16 × 120 cells are bump bonded to the macro-pixel sensor with 2 times 8 times 16 × 120 (30720) pixels of length ∼1.5 mm and a matching pitch of 100 µm. Processed signals from the SSA are routed via the bent flex hybrid to the MPA where the p_T correlation (stub finding) is handled. Being a 'pixel' module the power dissipation is higher but still remains at 8 W. The pixelation of the PS modules allows z-vertex resolution of the order of ∼1 mm at L1 in the central region. Due to the higher power dissipation, cooling pipes run along a base plate under the pixel sensor while the 2S modules have simply 5 individual cooling contacts.

In the end, 2S modules consist of two sensors, two readout hybrids plus one service hybrid plus 3 spacers and is planned to be mounted manually with dedicated jigs, different to the current tracker where modules have been assembled by robotic gantries. Similar for PS modules where industry delivers two spacers, four hybrids, one sensor and one bump bonded sensor-ASICs sandwich.

Sensor technology will be discussed in Sect. 7.1.1.

It is worth to note for future discussion that the module is a complete electrical entity with all necessary components; there are NO additional boards with additional data treatment in the tracker; all cables and optical fibres are directly routed outwards to power and readout. This reduces the complexity and material budget significantly. It also facilitates testing during assembly. Also data-out (trigger and full readout after L1-accept) and control-in signals share the same fibre pair – one fibre IN one fibre OUT per individual module. In the 2S modules binary strip data is fully synchronously readout to the **C**oncentrator **A**SIC (CIC) which sorts, sparsifies, formats and buffers the data before optical transmission. The **S**hort **S**trip **A**SIC SSA transmits the strip data to the **M**acro-**P**ixel-**A**SIC MPA responsible for the p_T selection L1-correlation. Trigger & full readout data, after a L1-accept signal, is then transmitted to the CIC to be sorted and shipped to the optical transmitters. All channels store events individually up to a latency of 12.8 µs, waiting for a trigger Level-1 accept and then transmit sparsified data with up to 750 kHz L1 rate.

The modules design was originally driven by the trigger requirements but has also very significant other benefits:

- it is an effective way of collecting two space-points Two sensors share mechanics, cooling and electronics resulting in significantly reduced mass.
- each layer (module) acts directly as track seed with **vector** (4D = 3D position plus bend) information instead of **space points** (3D).

---

[6]Interesting fact: The 'binary' SSA chip features two comparator thresholds, namely 0.4 and 1.5 MIPs, enabling to distinguish MIPs and highly ionising particles, a possible signature for a hypothetical long lived stable particle.

- in addition to the p_T discrimination at Level-1; the stub as well as full track information from Level-1 can enter the High Level Trigger stations. This should even speed-up later offline reconstruction.
- the self-contained design is very appealing. Only one data link connection, one power and one high voltage line. No other auxiliary electronics in the tracker volume, simplifying validation and construction.
- PS macro-pixel modules provide unambiguous 3D information facilitating pattern recognition in a high PU environment. This allows 3D track seeding.
- PS modules are more cost-effective than micro-pixel hybrid detectors due the larger cell size thus larger pitches allowing for more industrial cheaper bump bonding.

The mechanics for the outermost three OT barrel layers, consisting of 2S modules and called TB2S (**T**racker **B**arrel **2S**), will be very similar to the current tracker implementation (see previous chapter). In contrast, Fig. 7.7 depicts the novel mechanical structure of the inner three OT layers populated with PS modules. The inner (at small z) straight sections are built of so-called planks holding the PS modules. The PS modules in the tilted section (at higher z) are mounted on rings also holding the cooling loops. This part is called TBPS volume **T**racker **B**arrel **PS**.

The innovative decision to populate the endcaps with rectangular (square for 2S and rectangular for PS) instead of wedge shape sensor makes the current petal concept obsolete and bigger structures like full or half-disks are more adequate. To

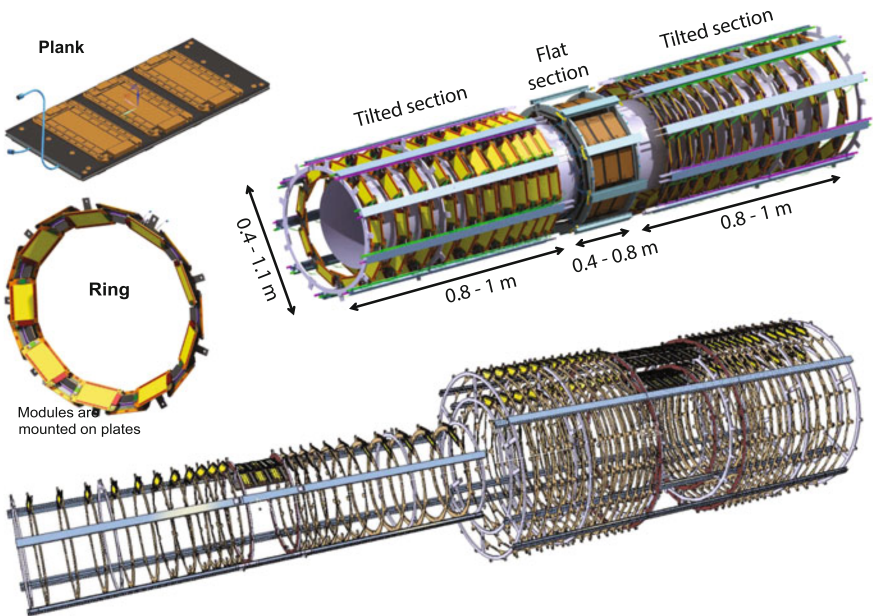

Fig. 7.7 The three TBPS layers are populated with PS modules on 'planks' for the inner flat section and on rings for higher z positions to achieve the tilt shown earlier. The *lower* part also illustrates how the final pieces will be put together [303]

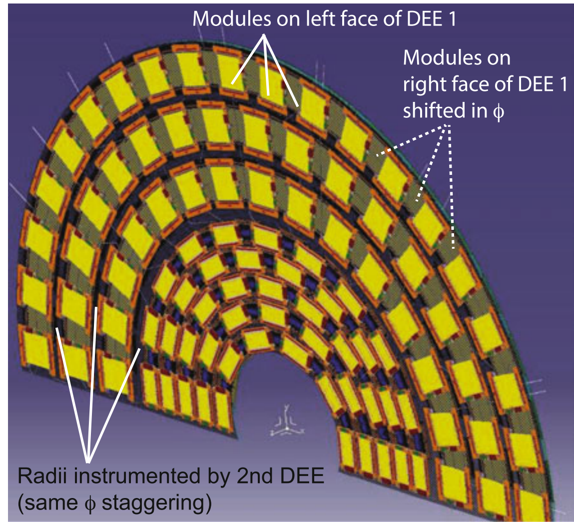

Fig. 7.8 The figures shows a single dee, to achieve full hermeticity without modules clashing, the other side of the dee needs to be populated with modules shifted in ϕ and the second dee needs to equip the locations in-between radii of the other dee [308]

facilitate handling half-disks called 'DEEs' are chosen and to guarantee hermeticity dees are assembled to double-Disks (four dees form one double-disk). Consequently the endcap volume is called TEDD **T**racker **E**ndcap **D**ouble **D**isk. One dee populated with modules is shown in Fig. 7.8. The dees support both 2S and PS modules and integrate the cooling loops. Even without additional electronics components the system is quite complex and routing of cables and fibres is a challenge.

Inner Tracker IT—Micro-Pixels

The system will consist out of 4 barrel layers (r = 3 to 16 cm) plus 8 regular sized end-disks plus 4 extended end-disks and will cover up to $|\eta| = 4$. The Outer/Inner Tracker boundary is located at r = 20 cm (r = 30 cm) $|z| < 160$ cm ($|z| > 160$ cm). The CMS central beam pipe deviates from cylindrical to conical shape at larger $|z|$. A full flat cylindrical pixel mechanics could not pass this 'obstacle' with straight insertion and therefore needs to be inserted from an angle. To achieve this the pixel detector is radially displaced further out at $|z| = 160$ cm. This is to allow pixel insertion/removal with beam pipe removal. This comes natural the current tracker, spanning only $|\eta| < 2.5$, and where the OT endcaps disks at higher z are naturally shorter. For an $|\eta| < 4$ coverage system, this volume has to be instrumented with pixel disks but at higher radius for higher z (see Fig. 7.9). CMS will keep the possibility to remove/repair the pixel detector or to replace the inner layer in a standard **Y**ear-**E**nd-**T**echnical-**S**top YETS. The extended disks are also perfectly suitable to accurately measure the luminosity (with background reduction taking tracks into account instead of simple hit counting).

It is fair to say that the layout, today 2017, is not final, e.g. the numbers of disks might change and even tilted designs are still under discussion. Even the pixel cell size might increase slightly depending on the ASICs development, its radiation hardness and number of features – more details on the RD53 ROC in Sect. 1.11.

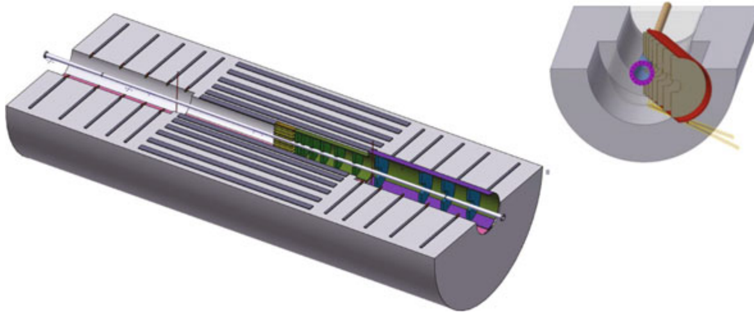

Fig. 7.9 CMS Pixel Phase II TDR layout and insertion studies

With respect to the current detector, the pixel cell size will be reduced by a factor six from 15000 to 2500 μm^2, while the total coverage will be much increased. The readout chip (ROC), to be bump bonded, will have cell sizes of $50 \times 50\,\mu m^2$ allowing easy mapping to pixel sensor cells of $50 \times 50\,\mu m^2$ or $25 \times 100\,\mu m^2$ for different locations, e.g. barrel and forward. Sensor technology will be discussed in Sect. 7.1.1. Also multiples like $100 \times 100\,\mu m^2$ or $50 \times 200\,\mu m^2$ are feasible if readout pattern in the chips can be appropriately configured to save power. The ASIC is a common ATLAS/CMS development driven by the RD53 collaboration (rf. also Sect. 1.11). Bump bonding of these small pitches will be a challenge, but seems doable and affordable.

Square ($50 \times 50\,\mu m^2$) pixels are better for z-resolution in the central region but would require lower detection threshold at higher z in the barrel region since, with the inclined track, charges are spread over many pixels; and this might get worse with radiation damage. Rectangular ($25 \times 100\,\mu m^2$) pixels are better in ϕ-resolution due to the smaller pitch (25 versus 50 μm), and clearly momentum resolution will benefit from this. The magnetic field and tracks are more or less aligned for disks at very high z (high η), thus the effective **B** is low thus the momentum resolution is pretty bad. The instrumentation at high η (forward region – disks) serves more the purpose of particle flow to reconstruct the full event and especially link all tracks to the individual vertex, especially in a high pile-up environment – also called pile-up mitigation. This might favour $50 \times 50\,\mu m^2$ pixels in the forward direction to emphasize z- over ϕ-resolution here.

As for the OT and for the current CMS Phase I pixel detector only rectangular modules are foreseen limiting the number of module types; here limited to two types with 2 or 4 ASICs.

Due to the high radiation, the very limited space and the very stringent requirement of low mass, services are moved out to higher radius.[7] Due to limited radiation tolerance, the optical links are located at higher radius and signals have to be driven

[7]Unfortunately, the trick, used in Phase I to move them out of the tracking volume does not work any more with tracking up to $|\eta| = 4$ – the whole volume is dedicated to tracking.

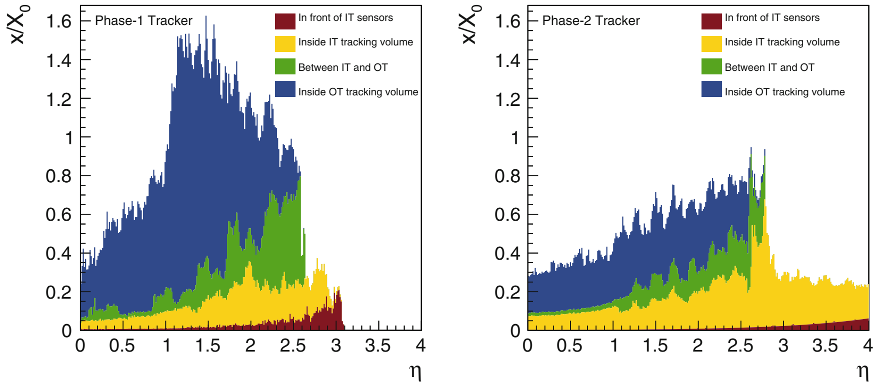

Fig. 7.10 Material budget distribution versus η for the current and future Tracker (TDR geometry). As an example, the move from flat to tilted geometry saves in the order of 0.2 x/X_0 in the range $1 < |\eta| < 2.5$ [303]

electrically over 1 to 2 m. The RD53 ROC will be the only active element on the modules! Direct powering and even DC-DC converters would result in many kilograms of cables thus serial-powering is being adopted.

CMS expects about 1 GRad of integrated ionising radiation and a hit rate of up to 3 GHz/cm^2 (~100 kHz/pixel) in the innermost layer. The chip is under development within the RD53 collaboration (joint venture of ATLAS and CMS). The chip will cope with a L1-accept data rate of about 750 kHz and features a trigger latency of 12.8 μs. Different to the Outer Tracker, no pixel information is available at Level-1. The high rate does not allow to ship data to the chip periphery for buffering but several cells store events locally. Instead of having a large analogue (pre-amplifier) section with digital logic 'just at the periphery', the new chips will have analogue 'islands' in a 'sea' of digital logic. Chip size will be 22 mm × 16.4 mm active area plus 22 mm × 2 mm periphery. Depending on locations modules will consist of 1 × 2 (rectangular-inner radii) or 2 × 2 (square – outer radius) chips with up to 3 low-mass data links per chip with about 1.3 Gbits/s (see also Fig. 7.4).

CMS Tracker Phase II System

A very important aspect is the reduction of mass inside the Tracker volume and in front of the electromagnetic calorimeter. The tracker material is a limiting factor for ECAL performance (energy corrections), increase Bremsstrahlung and photon conversion and the main source of track reconstruction inefficiency (nuclear interactions).

Figure 7.10 demonstrates the substantial reduction with respect to today.[8]

[8]Both plots do not account for the material of services, e.g. connectors or manifolds outside the TK volume.

The reasons for the massive reduction are manifold, the main saving lies in the reduced cable cross-section possible due to the DC-DC conversion at module level:

- less layers
- DC-DC converters will reduce the amount of copper (diameter of the power cables) drastically
- CO_2 cooling comes with thinner/smaller pipes and the liquid CO_2 per se is also lighter than C_6F_{14}
- reduce silicon thickness if affordable (although the Fig. 7.10 shows the standard sensor thickness of 300 μm)
- less additional boards and connectors
- avoid meander cabling
- optimise module design – save at every ‘corner’
  - no extra glass pitch adapter – integrated in flex hybrid, pushing the hybrid technology to its limit – down to 42 μm track width and 42 μm vias
  - two sensors in a single frame
  - balance thermal management – the margin is in the cooling plant capacity, not in massive thermal contacts
- The whole system is optimised/balanced for minimal power consumption. This sets limits on some parameters, e.g. longer latency (more buffers), faster readout, more channels would increase power thus higher current (more copper) and temperature (cooling infrastructure).
- serial powering in the Inner Tracker system.

Table 7.1 compares the current CMS Phase I and the future Phase II Tracker

Track Trigger Technique

The Track Trigger at Level-1 is based on the capability of the intrinsic low p_T discrimination power of the p_T-modules in the Outer Tracker and their capability to sent out ‘stub’ data for every event at 40 MHz. The concept works down to 20 – 25 cm in radius with a realistic 100 μm strip/pixel pitch and reasonable sensor to sensor spacing of 1.6 – 4 mm, thanks to the strong CMS magnetic field of 3.8 T. Around 3% of all tracks are above 2 GeV accounting for roughly 100,000 stubs at PU = 140. These 3% of tracks at 40 MHz use about 80% of the readout bandwidth leaving about 20% for the full data with a L1-accept rate of 750 kHz. The data path is illustrated in Fig. 7.11 where the 512 deep buffers guarantee a latency of 12.8 μs before event info is overwritten. The stub identification is handled by the CBC and MPA ASICs. It is a pure self-seeding detector-intrinsic measurement at a true 40 MHz level not depending on any pre-trigger info from the calorimeters or the muon detectors at L1.

The Track Trigger relies on six double-sensor layers giving six stubs (5 + 1 redundancy) with at least two (plus one for redundancy) measurement points with good z resolution from the PS modules (2 + 1 for redundancy). Hermeticity is a key requirement.

Table 7.1 The table compares the current and future CMS Tracker (TDR version, numbers to be fine-tuned in a future engineering design review). According to TDR plans the Phase II Tracker active surface will be slightly smaller than the first CMS Tracker - larger IT but smaller OT

	CMS Tracker (Phase I)	CMS Phase II Tracker
Outer tracker (OT) – Strips and Macro-pixels		
Silicon Surface	$\sim$206 m^2	$\sim$191 m^2
Strips	9.3M	42M
Macro-Pixels	0	172.5M
Modules	15148	$\sim$13300
Readout rate	100 kHz	750 kHz (on L1 accept) / 40 MHz (@L1)
Cooling	mono-phase C_6F_{14}	bi-phase CO_2
Cooling pipes	Al, stainless steal, Titanium	Titanium
Powering	direct	DC-DC
Inner Tracker (IT) – Micro-Pixels		
Pixel surface	$\sim$1.75 m^2	$\sim$4.9 m^2
Pixels	$\sim$127M	$\sim$2000M
Modules	1632	$\sim$4300
Readout rate	100 kHz	750 kHz
Cooling	bi-phase CO_2	bi-phase CO_2
Cooling pipes	stainless steal	Titanium
Powering	DC-DC	serial powering

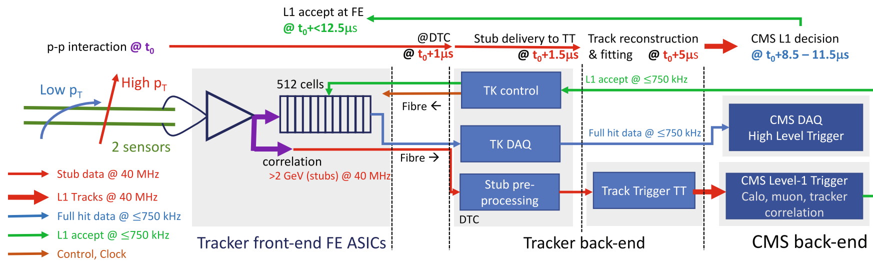

Fig. 7.11 The stub data (*red path*), identified in the front-end ASICs is sent with full 40 MHz to the back-end, where tracks are pre-processed and in dedicated boards identified, fitted and sent further to the CMS central trigger. The main data handling is done by the **D**ata, **T**rigger and **C**ontrol DTC boards. Evaluating the full information of Level-1-tracks, Calorimeter information and Muons in dedicated FPGAs farms, a L1-accept is being sent back to the front-ends where all data is buffered for the latency of 12.8 µs. The *upper* figure text 'repeats' the Track Trigger data flow with timing information. After the L1-accept, the full, although sparsified, data is being sent to the Tracker/CMS DAQ (data acquisition) with a rate of $\leq$750 kHz

As can be seen in Fig. 7.6, readout ASICS are situated on both ends of the sensors of PS and 2S modules increasing granularity/halving strips lengths. Due to this fact p_T correlation only works for the two halves of the modules individually with no

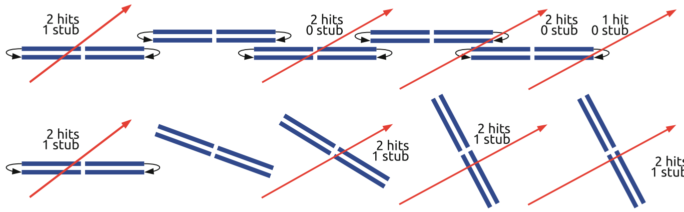

Fig. 7.12 The *upper* cartoon shows the flat geometry where tracks at L1 are lost when passing in the *middle* and at an angle where the *upper* sensor is being 'hit on the other end'. Readout and thus p_T correlation processing is only individually per end possible. The lower cartoon demonstrates how the tilt recovers the stub finding [301]

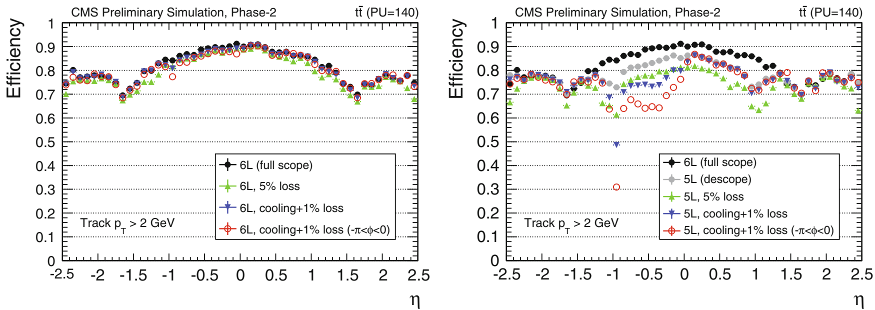

Fig. 7.13 Simulation of Track Trigger perforance. The *left* plot demonstrates the robustness and redundancy of a six-layer detector compared to a five-layer detector (*right*) where failures are largely detrimental [302]

communication between left to right end[9], thus tracks passing close to the middle and at an angle cannot be correlated by the front-ends thus no stub is being produced.

Tilting the sensors (modules) with respect to a beam-parallel orientation; turning them to face the track orthogonal remedies this and all stubs are found. Track Trigger efficiencies would be reduced at larger z by about 30% without tilting the modules. Figure 7.12 illustrates how the stubs are being recovered by the tilt. With the lower number of modules, the tilted layout also reduces the stub rate by about a factor of two compared to the flat geometry, decreasing data transmission rate and facilitating Level-1 track finding.

The simulation results shown in Fig. 7.13 demonstrate the necessity for a minimum of six barrel p_T-layers. It should be mentioned that with a high granular four-layer pixel detector, missing 1 to 2 outer layers would not be much detrimental for offline reconstruction but the Track Trigger is based solely on the Outer Tracker. For lower radii ($r < 20$ cm), tracks are all 'straight' within any mechanically reasonable sensor to sensor spacing (Δr) in an already very densely populated volume. A p_T based L1-

[9]Establishing a communication line between the two module/sensor ends would be an alternative but the corresponding R&D to establish this with **T**hrough **S**ilicon **V**ias TSV proved to be a not yet mature technology.

trigger is therefore not possible with the inner micro-pixel detector (Inner Tracker). The left plot, featuring six OT layers, demonstrates its robustness versus random and globally localised failures, e.g. like a cooling loop spanning a full hemisphere (180°). The five layer detector starts with a lower efficiency but suffers significantly more from failures, especially locally. Six layers provide the necessary redundancy.

The dedicated back-end electronics has about 3 to 4 μs for track identification and track fitting to deliver tracks to the CMS global trigger with further trigger recipes to allow final L1-accept signal arriving back at the front-end within the latency of 12.8 μs before the data is overwritten. Tracks above a defined momentum have a maximum bending thus are always confined in a defined detector segment/volume. Virtually segmenting the detector in such slices ensures that a high p_T track stays within such a segment or within the segment plus a neighbour segment. Taking the front-end to back-end cabling into account, time-multiplexing[10] can now deliver all data of such segments plus their neighbours for each individual bunch crossing to a single track-trigger processing unit. The track trigger system starts with pattern recognition of the already reduced set of stubs (compared with the full set of space points with full readout). The identified tracks are then being fitted, doublets are being removed and final information is being sent to the CMS trigger.

Three pattern recognition methods have been developed and demonstrated to work. More info can be found at [303] and [240].

- **A**ssociative **M**emories AM (custom ASICs) basically storing all possible a priori calculated patterns and then compare them with the found stubs
- the ‘Tracklet’ approach where stubs of one layer are combined with stubs in the second layer forming a so-called ‘Tracklet’ which is then followed to the next stubs/tracklet – all realized in FPGAs
- Patterns/Tracks are identified by Hough Transforming the stub information – all realized in FPGAs

The track fitting has also been fully realized in FPGAs ranging from χ^2 fit to a full Kalman filter. All methods have been fully demonstrated in hardware; all residing within the required 3 to 4 μs.

7.1.1 Sensors for the HL-LHC CMS Tracker

This section describes steps during the sensor development and evaluation and not the very final implementation.

As already discussed in Sect. 2.2.4, with the onset of trapping at high fluences, the concept of full depletion voltage becomes more and more abstract and less relevant. In principle we only need to guarantee high efficiency, high resolution and low power

[10]Data from multiple sources are routed through a multiplexing network which directs all the data from an individual bunch crossing to a single processor.

consumption, but we are also interested in current, CV[11] scans to evaluate N_{eff} and its profile, trapping, CCE, **s**ignal-to-**n**oise S/N and strip parameters (inter-strip resistance, inter-strip capacitance, etc.).

The CMS strip-, macro-pixel- and micro-pixel-sensor R&D is far advanced but some parameters, e.g. thickness, punch-through structures, lengths of polysilicon resistors, p^+-stop structures, doping concentrations, active edges, alignment markers, test structures are still under final refinement. Global parameters like bulk material (n-in-p), outer dimensions and, pitches are chosen.

To reach this point, CMS conducted a long and in depth R&D with the goal to evaluate (a) base material with respect to radiation tolerance and annealing, (b) geometry and (c) special designs, e.g. routing structures. Figure 7.14 shows the basic wafer layout containing diodes, several sizes of mini-sensors, dedicated test structures (improved version of the former CMS ones), large areas of 2S and PS sensors with different pitches and strip width-to-pitch versions and two different lengths of macro-pixels. Wafers have been procured in large variety of base materials and thicknesses from a single vendor to allow absolute comparison of results. CMS procured Epitaxial material **Epi**, standard floatzone **FZ**, deep diffused FZ **dd-FZ**, magnetic Czochralski **mCz** all in p-in-n & n-in-p in different active thicknesses ranging from 50 μm to 300 μm (thinning rf. also Sect. 1.9.3, p.93). Some of the p-in-n & n-in-p came with a double-metal layer to study routing options, mainly implementation of the **p**itch **a**dapter PA functionality in the sensor itself saving the glass PA, and thus its mass[12] (routing rf. also Sect. 1.4). For the macro-pixels (DC-coupled) punch-through as well as polysilicon biasing has been explored (biasing rf. also Fig. 1.27 on p.34). It is worth to mention that the wafer-backplane is an Al-grid (mm-scale) to allow laser-light injection from the back for CCE and TCT studies. The full granularity is listed in Table 7.2. The main results have been published here [20, 73, 82, 89, 146, 196, 222, 223, 243, 244, 308, 314, 328].

Learning from RD50 (rf. Sect. 2.2) that NIEL is violated and only the real mix/percentages of charged and neutral particle radiation at the given location fully validates the material, CMS irradiated the wafers to the more realistic radiation mix for several radii. The main important values are **s**ignal-to-**n**oise S/N or **C**harge **C**ollection **E**fficiency CCE but also IV, CV and all strip parameters have been investigated; e.g. to determine depletion voltage and power consumption. It is worth to mention that, in all cases, strip sensors plus diodes have been investigated and LHC electronics with correct sampling times and frequency have been used. Using strip sensors allowed full characterization of strip parameters, like inter-strip resistance, inter-strip capacitances, coupling capacitances, bias resistors, etc. Many structures, especially diodes, have also been evaluated with TCT, DLTS and TSC (rf. Sect. 1.8)

[11] At the relevant fluences, depletion voltage (V_{FD}) is a more abstract concept.

[12] The R&D was successful but it is even more economic to implement the PA routing into the hybrids.

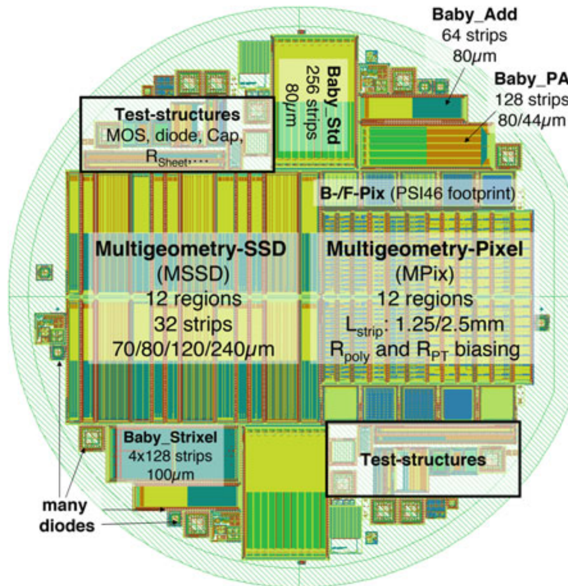

Fig. 7.14 R&D wafer with different structures, described in the text to evaluate base material, design rules and geometry with respect to radiation and subsequent annealing. The main work-horse for the radiation/annealing campaign are diodes, Baby_Std and Baby_Add but also the multi-geometry structures have partially been irradiated; plus some of the test-structures. Multi-geometry-strip detectors and multi-geometry macro-pixel have been evaluated for geometry and design optimization. Baby_Add was used to study Lorentz angle wrt. to radiation, temperature and voltages. The test-structures are being used for process evaluation [308]

Table 7.2 The table shows the variety of wafers composing the several-year long testing campaign of CMS. Some of the wafers have been processed with a second metal layer for routing studies. The structures on the wafers can be visited in Fig. 7.14

Material	Bulk doping	Isolation for n-in-p	Thickness [μm]
EPI	p-in-n, n-in-p	p^+-stop, p-spray	50, 70, 100
FZ physical thickness	p-in-n, n-in-p	p^+-stop, p-spray	200, 300
FZ physical on carrier wafer	p-in-n, n-in-p	p^+-stop, p-spray	120, 200
dd-FZ	p-in-n, n-in-p	p^+-stop, p-spray	120, 200
mCz[a]	p-in-n, n-in-p	p^+-stop, p-spray	200

[a]Deep diffusion is not available for mCz, due to the thermal donor creation during the high temperature process thus drastically changing the base material. Some discussion happened recently with a wafer provider, which believes it could be done with a "donor-kill" temperature cycle – very high temperature to break the oxygen clusters and then cool down fast

for different radiation levels. The measurements have been cross-checked with simulation and often led to a better tuning of the simulation itself (rf. also Sect. 1.8.5, p.80). Finally, after the mixed irradiations, the structures have been subjected to a longterm annealing campaign.

Table 7.3 Integrated Radiation fluences expected after 3000 fb^{-1} according to FLUKA simulations at the time and the corresponding radius in the detector. The small radii 5 and 10 cm * are only relevant for *thin* micro-pixel structures, thus only diodes and a small number of Baby_Stds have been irradiated to that level. Structures have always been measured after individual neutron and proton fluence and with the mix in the end

Radius (cm)	Protons	Neutrons	Total	Ratio	Thickness (μm)	Sensor type
	Fluence in ($10^{14} n_{1\text{MeV}}/\text{cm}^2$)					
60	3	4	7	0.75	≥200	2S sensors
20	10	5	15	2.0	≥200	PS sensors
15	15	6	21	2.5	≥200	Micropixels
10*	30	7	37	4.29	≤200	Micropixels
5*	130	10	140	13	≤200	Micropixels

The main work-horse structures for the radiation tolerance studies have been implemented twice on the wafer (upper and lower half) allowing to irradiate one structure with neutrons and one with protons, measure these and then subject the neutron irradiated one to proton irradiation and vice versa resulting in two pieces irradiated to the correct fluence-mix. Table 7.3 shows the main fluences Baby_Std (rf. Fig. 7.14) and diodes have been submitted to.

All measurement results have been saved in a global database to allow full comparison and evaluation; all participating institutes went through a measurement cross-calibration early on.

In a nut-shell the campaign, with several tens of man-years of work, validated the routing concepts and identified the Phase II sensor baseline for the CMS Outer Tracker, namely n-in-p with an active thickness of 200 μm. The following radiation results of the R&D campaign allow also clear recommendation for the IT. As a result, thin (n-in-p) planar sensors are the baseline for most parts of the CMS micro-pixel volume, while 3D sensors (rf. also Sect. 1.12.7) are also candidates for the innermost layer. Final layout and even final cell geometries are still being discussed but are close to be frozen. Also the active thickness needs final evaluation.

Outer Tracker Sensors for the HL-LHC Tracker

The several year long CMS R&D campaign produced a very exhaustive set of data for different materials, thicknesses, geometries etc.. The next section gives an excerpt only, while a more detailed discussion can be found in [328]. The campaign concentrated on measurements of small strip sensors plus diodes. All strip parameters, also inter-strip resistance and capacitance stayed within limits for all levels of radiation independent of strip isolation strategies.

Figures 7.15 and 7.16 demonstrate the main results leading to the baseline sensor choice – n-in-p with an active thickness of 200 μm (or at least below 300 μm).

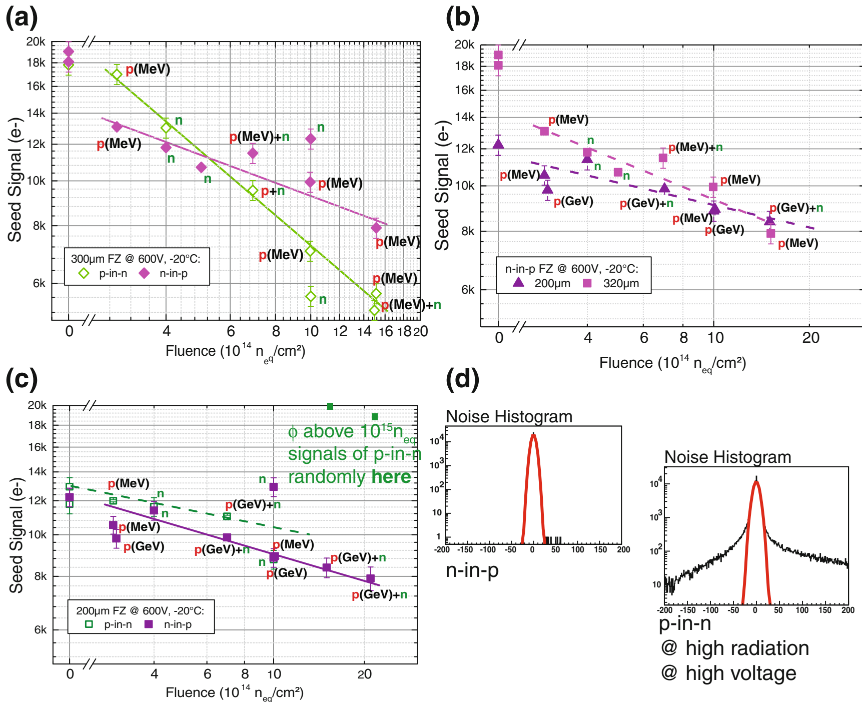

Fig. 7.15 Overview of the results of the CMS Phase II sensor campaign. 'Seed' signals – single strip signals are plotted and not the summed signal of strip clusters. Figure **(a) - (c)** show the 'seed signal' evolution versus radiation fluences for n- and p-bulk for thicknesses of 200 and 300 μm. The *small labels* indicate the type of radiation namely neutron ($\sim$1 MeV) or proton of different energies (23 MeV or 24 GeV) and the mixed cases (proton+ neutrons – labled as p+n) as described in the main text. Figure **(d)** show the noise histogram of n-in-p sensors (*left*) and for p-in-n (*right*) for higher radiation levels. The p-in-n show anomalous high non-Gaussian noise which is often higher then the signal itself; we see micro-discharges at high fields (or locations of charge amplification) [78, 308, 328]

The future CMS readout is binary[13] thus registers only above-threshold signals. Therefore, only 'seed' strip signals are evaluated and shown.

Figure 7.15a compares $D = 300\ \mu$m thick p-in-n (hole readout) with n-in-p strip sensors (electron readout). As expected, at a fluence level of $6 \cdot 10^{14}\ n_{1\text{MeV}}/\text{cm}^2$, where trapping is becoming the dominant detrimental effect, charge collection for holes is decreasing below that of electrons (slower drift of holes thus more affected by trapping, better electric and weighting field combination and also higher electric field at n-electrodes after irradiation; rf. also Sect. 2.2.4). For the 300 μm thick sensors, the hole signal clearly drops below the electron signal and also below the readout threshold.

[13] The on-chip comparator has no cluster algorithm but works individually for each channel. No centre-of-charge can be applied since the smaller signals are simply not recorded; resulting in binary position resolution $\sigma_x = \frac{pitch}{\sqrt{12}}$.

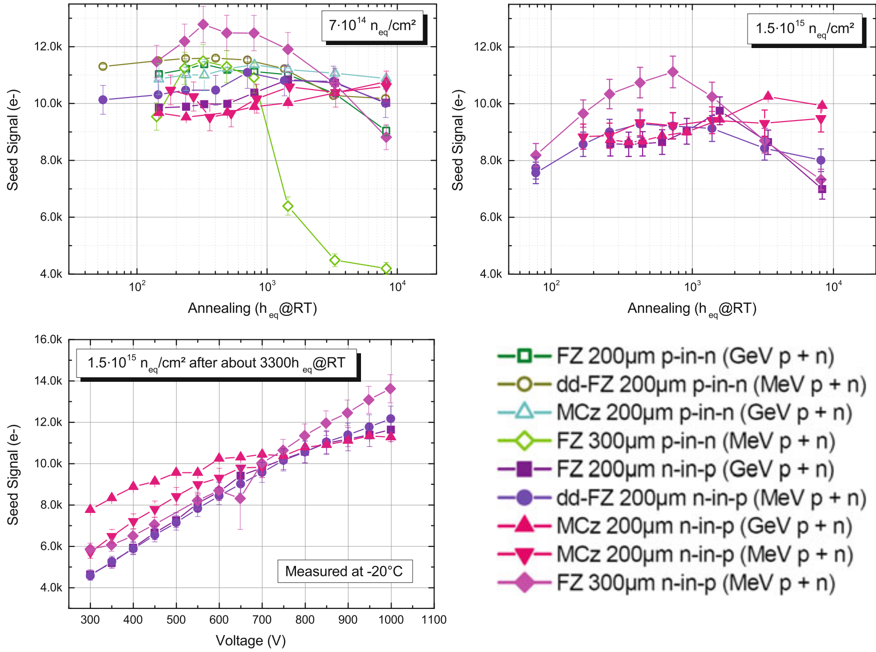

Fig. 7.16 The annealing for highly irradiated sensors are shown (n-type bulk only for *lower* fluences); all times normalised to hours at room temperature. Lines are drawn to guide the eye. The legend names the material, followed by sensor thickness, polarity and finally type of irradiations, e.g. "dd-FZ 200µm n-in-p (MeV/p+n)". FZ stands for floatzone, mCz for magnetic Czochralski and dd-FZ for **d**eep **d**iffused-FZ. The *upper* plots demonstrate clearly that signals of 200 µm thick sensors do not decrease with annealing time while 300 µm, especially p-in-n do. The *lower left* plot shows the signals versus bias voltage demonstrating that higher voltage improves the margin. The superior behaviour of mCz is explained in the main text and demonstrated further in Fig. 7.17 [308]

The advantage of higher primary charge generation for larger thicknesses D becomes marginal at high fluences when trapping dominates. Thinner sensors have less current ($I \sim D$) and much smaller depletion voltages ($V_{FD} \sim D^2$), meaning the electric fields across the sensor and at the electrodes are much higher! Thinner sensors also draw less power ($P \sim D^3$). We will revisit this effect when we discuss the annealing behaviour with Fig. 7.16. Figure 7.15b, comparing 200 and 300 µm thick n-in-p strip sensors, show the above described effect.

With $V_{\text{bias}} = 600$ V at $\Phi_{eq} = 10^{15}\ n_{1\text{MeV}}/\text{cm}^2$, 300 µm thick silicon does not provide more signal than a 200 µm one; a marginal improvement at higher voltage (e.g. 900 V) exists.

Figure 7.15c compares p-in-n with n-in-p, both 200 µm thick. The significant detrimental effect as seen for the D = 300 µm thick comparison is clearly gone and at first look, p-in-n seems even superior to n-in-p. As a matter of fact, field strengths at the electrodes for thin sensors are high (enough) for both types. To note, the 'higher' signal for n-in-p is present only for strip sensor and not for diodes thus is a consequence of the segmentation-design and not of the base material. Taking

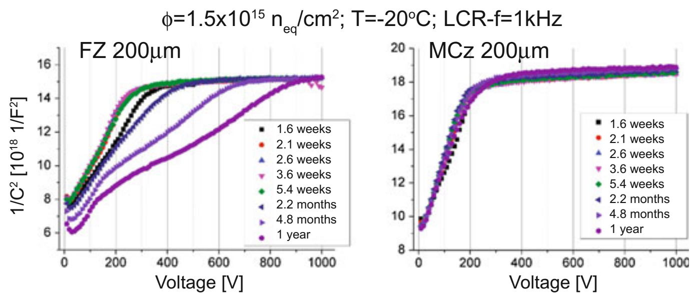

Fig. 7.17 The CV scans of mCz, compared to FZ, show basically no change in depletion voltage nor electric field strength [143]

not only signal but also noise into account the picture changes drastically. At higher fluences and/or higher voltages, p-in-n strip sensors exhibit large non-Gaussian[14] tails in the distribution expressed in Fig. 7.15d. These would result in fake hits. De facto, p-in-n sensors see charge amplification (rf. also Sect. 2.2.5). At high fluences (where the plot shows no green data points), non-physical signals up to 44k electrons and down to 5k electrons have been measured. Dedicated simulations[15] confirm very high fields at the p^+-electrodes for p-in-n strips sensors, while the fields are more 'distributed' for n-in-p sensors with additional p^+-spray or p^+-stops.

Figure 7.16 shows the annealing behaviour of 200 and 300 μm thick n- and p-type sensors. Signals of the 300 μm thick sensors (especially p-in-n) drop with annealing time, while signals from thin sensors stay basically constant. As shown earlier, N_{eff} changes with annealing thus field strengths changes and this affects thinner sensors simply much less than thicker ones, since electric fields always stay high enough.

This gives a fantastic outlook for future operation without the need to keep the detector cold during maintenance periods or even better to warm it up in a controlled fashion, benefiting from the leakage current annealing. This is already true for V_{bias}=600 V, and with even more margin towards higher bias voltages increasing field strengths. The full Tracker system is designed to cope with 800 V – ample margin.

Looking closely, mCz[16] is strikingly superior to FZ with even less changes during the annealing process. Figure 7.17 shows virtually no change in CV scans for all

[14]This non-Gaussian, also called anomalous, noise exists for the p-in-n sensors in this campaign. It is probably that this could be overcome with an adapted strips design but it still looks like, that n-in-p sensors with p^+-stops have more margin than p-in-n.

[15]Device simulations [316] have shown that irradiated p-in-n strip sensors develop high electric fields at the strip edges, intensifying with increasing accumulation of oxide charge. The electric fields around the n-strips in n-in-p sensors are instead reduced by higher oxide charge, which makes them more robust with respect to effects, such as breakdown, noise or micro-discharge after heavy irradiation with charged particles.

[16]The oxygen content of dd-FZ is already very high due to the high temperature treatment. Still, the oxygen concentration in mCz is one order of magnitude higher and due to the crystal growth process in the magnetic field more homogeneous.

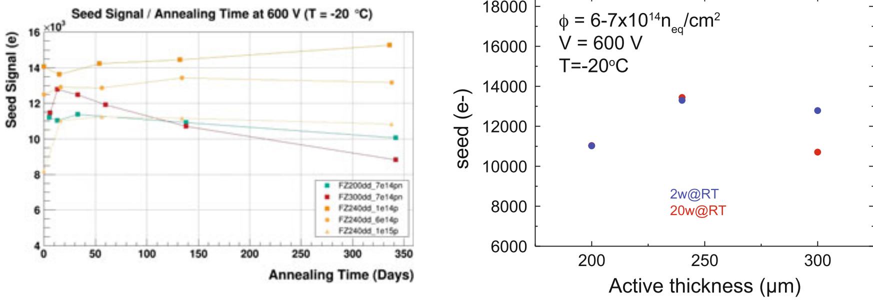

Fig. 7.18 The plots demonstrate the signal robustness of thin sensors against annealing. It seems that 240 is as robust as 200 µm with globally a larger signal. The *left* plot shows 'seed signal' versus annealing normalised to hours at room temperature. The *right* plot compares signals after irradiation for three thicknesses before and after annealing [78]

annealing times thus a low and constant depletion voltage, thus high electric field strength. The changes of electric field strength in FZ are inside working parameters, the E-fields of mCz do not even change.

Adding all observations up, the optimal choice would be mCz n-in-p with a thickness of $d = 200$ µm implemented in a system with enough margin with respect to high voltage robustness. Unfortunately, mCz might not be readily available in quantity in the high resistivity ρ regime[17] thus the next best choice is FZ or dd-FZ available with higher ρ. Now, the signal in 200 µm is a bit marginal, with respect to the noise in the front-end readout electronics, thus maybe the best choice lies in-between $d = 300$ and 200 µm. Full size prototypes (along with baby sensors on the wafer half-moons) have been procured and evaluated. The results can be visited in Fig. 7.18, demonstrating that an active thickness of 240 µm is as robust as 200 µm with globally a larger signal. Thickness can be further tuned; it can be chosen precisely by physical thinning or deep diffusion.

Due to different field strengths, leakage currents at same voltage are not scaling exactly with thickness; they are always a bit higher than expected but are still lower for thin sensors (200 µm thick sensor currents are about 85% of 320 µm).

Table 7.4 presents the different sensor dimensions of the Outer Tracker (TB2S and TBPS). The 2S and PS_strip sensors are common single-sided, AC-coupled, polysilicon biased n-in-p with atoll-p^+-stop isolation. The PS_pixel sensor is a single-sided, DC-coupled, punch-through biased[18] n-in-p with atoll-p^+-stop isolation. Dimensions make use of the maximal area within the vendor design rules of a 6 in. wafer. Edges, metal overhang, p^+-stop concentrations ([243]), etc. have been optimized with simulation and cross-checked on different process runs. The sensors

[17] As reminder, low resistivity ρ would result in a high starting depletion voltage which for n-in-p sensors only increase with radiation.

[18] Also polysilicon is an option.

Table 7.4 The recipe used for the current CMS sensor strip layouts (rf. Sect. 6.4.2) also holds for thinner n-in-p sensors: All sensors feature a width-to-pitch ratio of w/p=0.25 and a metal overhang of 5 µm for strips and 5 µm for the macro-pixels

Sensor type	Width (mm)	Length (mm)	Strip/Pixel		Number of strips/pixels
			Pitch	Length	
Strip-Strip (2S)	94.183	102.7	90 µm	5 cm	2 × 1016 = 2032
Pixel-Strip (PS_s)	98.74	49.16	100 µm	2.5 cm	2 × 960 = 1920
Pixel-Strip (PS_p)	98.74	49.16	100 µm	1446 µm	32 × 960 = 30720

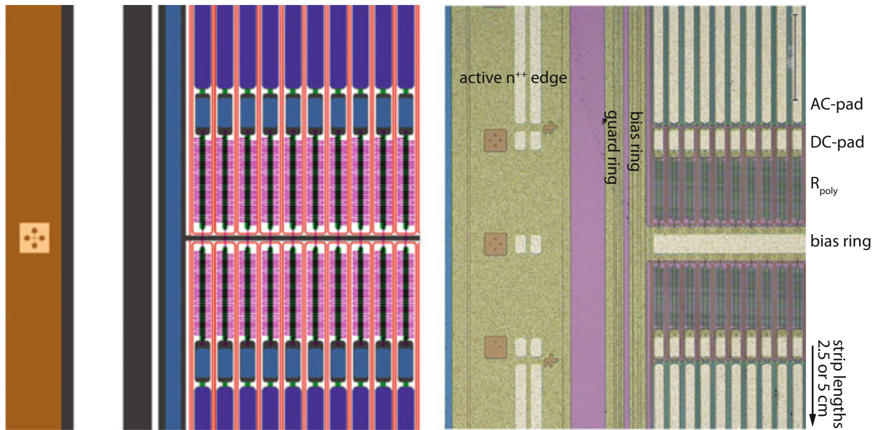

Fig. 7.19 Centre edge of a 2S sensor. The *left* picture shows the design while the *right* a real photo. The polysilicon bias resistors R_{poly}, DC and AC pads are all located in the sensor center to facilitate testing in on row instead of at the sensor edges. The picture also nicely shows the atoll p^{+}-stops guaranteeing the strips isolations in an n-in-p sensor. Looking closely, one can see the vias from DC pad to R_{poly} and back to the bias ring. Lighter parts on the pads and rings are openings in the passivation

will be procured thinned or utilising the deep diffusion[19] process which might save cost at some vendor. dd-FZ has the advantage of high mechanical stability and no need for thinning at the company but obviously introduces more mass to the Tracker with respect to a physically thinned sensor.

Figure 7.19 shows a potential sensor design which has also been prototyped in some varieties. The details of the centre of the strip sensors (2S and PS_strip) where the DC-pads and polysilicon resistors are located to allow testing all strips (both ends) in a single row. In addition, in the absence of bias resistors at the sensor ends, the AC pads could be moved further out decreasing wire-bond length.

[19]Only dd-FZ exists, the deep diffusion process is not available for mCz. Although some companies might explore dd for mCz.

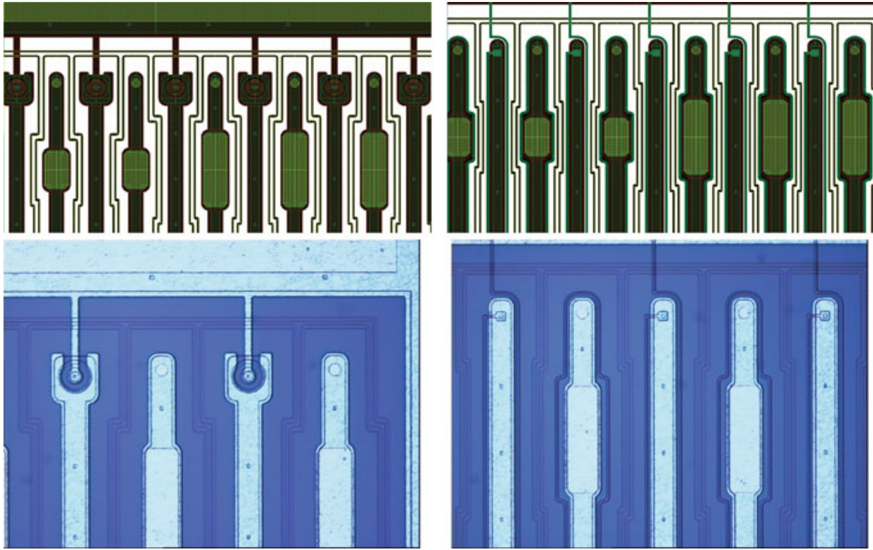

Fig. 7.20 The CMS PS_p sensor; layout and photo. The *left* shows the **p**unch-**t**hrough PT implementation and the *right* the polysilicon one, were the bias resistor runs around the full macro-pixel. All strips are surrounded by individual p^+-stops

Figure 7.20 shows the layouts and photos of early prototypes of the PS_pixel sensors. **P**unch-**T**hrough PT and polysilion biasing have been tried. Both are working and PT requires less processing steps. The p^+-stop design is also roughly exemplarily for 2S and the PS_strip sensor with an atoll for every single strip/macro-pixel plus a common one surrounding the combined strip area, isolating against the bias and guard rings.

The CMS Tracker community was also exploring, mainly for potential cost savings, the usage of 8 in. wafers despite the fact that everything was designed for sensor sizes optimised for 6 in. ones. For the macro-pixel ($5 \times 10\,\text{cm}^2$; DC-coupled) wafers, three sensors would fit on a wafer instead of two. One 2S ($10 \times 10\,\text{cm}^2$; AC-coupled) plus one PS_strip ($5 \times 10\,\text{cm}^2$; AC-coupled) sensor would fit but a smaller number of PS_strip sensors than 2S sensors are needed. One gains when the price increase from 6 to 8 in. is less then a factor 1.5. In earlier design studies, CMS explored and produced a 2S sensor with a special size of $16 \times 10\,\text{cm}^2$ with $D = 200\,\mu\text{m}$ thickness to make proper use of the 8 in. area, but dismissed it finally. The photo of this special 2S-long sensor, the first 8 in. wafer ever processed in HEP is shown in Fig. B.7 on p.356 in the Appendix, for more information see [36].

Inner Tracker IT Sensors for the HL-LHC Tracker

Being subjected to even higher radiation levels than the OT, the micro-pixel sensors will also be processed in n-in-p technology collecting electrons. The innermost layer will have to withstand up to $\Phi \approx 2 \cdot 10^{16}\ n_{1\text{MeV}}/\text{cm}^2$ of fluence or a dose of

0.5 – 1 GRad integrated after $\mathscr{L} = 3000$ fb^{-1}. CMS will keep the global detector design, allowing for access without beam pipe removal.

The sensor thickness will be reduced even more to achieve high enough fields after high radiation and annealing. Values like 100 – 150 μm of active thickness are being evaluated – physical thinning, active on handle wafer or deep diffused material (dd-FZ). With much smaller capacitances compared to strips thus much lower 'noise' smaller signals can be accommodated for. The specifications for the RD53 pixel readout chip target a threshold below 1000 e$^-$ down to 600 e$^-$ thus signals about 5000 to 6000 e$^-$ are ok. Indications are good, but due to the lack of radiation tolerance ASICs with small cell sizes (pitches), the usage of planar thin sensors have not been fully proven for the innermost layer. Therefore, radiation tolerant 3D sensors (rf. also Sect. 1.12.7) are also probable candidates for the innermost layer. But, the small pitch, thus 3D-column small **width to depth**, will be a technology challenge to establish in 3D technology.

With direct ASIC to pixel channel **b**ump **b**onding BB the sensor size has to be a multiple of the chip size. Modules will be only rectangular, no wedge shape is planned. With chip sizes about 2×2 cm^2, final BB yield will most probably define the maximum size module, while 2×1 and 2×2 versions are being evaluated. The electronic cell will be 50×50 μm^2 while sensor cells will be 25×100 μm^2, 50×50 μm^2 or potentially a multiple of this (50×200, 100×100) where only some electronics channels are being connected to save power in the outer layers. The BB pitch is defined by the electronics cell and the BB ball size by the smallest dimension of the sensor pixel – BB-pad on ASIC defines placement/routing on sensor. The BB pad spacings of sensor can be seen in the prototype layouts shown Fig. 7.21. Industry is setting-up for this feature size. With very small pitches like 25×100 μm^2 charge is basically always shared between two pixels, improving resolution.[20] As mentioned earlier 25×100 μm^2 sensor cells preliminary seems to be the best choice despite slightly degraded z-resolution. Another aspect of the thin sensor is the corresponding mechanical bow or warping resulting in extra complication for BB. Often edges lose BB contacts. Possibly deep diffused material with the good balance of active thickness[21] and mechanical stability could be the best choice; e.g. physical thickness of 200 μm and active of 130 μm. In general mechanical stability degrades with decreasing physical thickness and increasing surface.

Another challenge posed by the small pixel size is the pixel isolation and biasing strategy. The p^+-stops barely fit in between the cells, especially in the atoll-geometry, but it is doable. Common p^+-stops would give more margin but also more risk since a local defect could influence much larger areas.

In the final configuration, pixels are directly and individually connected to the pixel ASIC thus a biasing grid is per se not necessary. It serves for initial testing and could provide protection against large currents in the final system. The protection feature, though, could be implemented in the ASIC. With a very high processing yield, testing

[20] Time-over-threshold readout will be used and centre-of-charge calculation is possible; different to the binary readout in the Outer Tracker.

[21] Final active thickness still to be optimized – 130 μm is just an example.

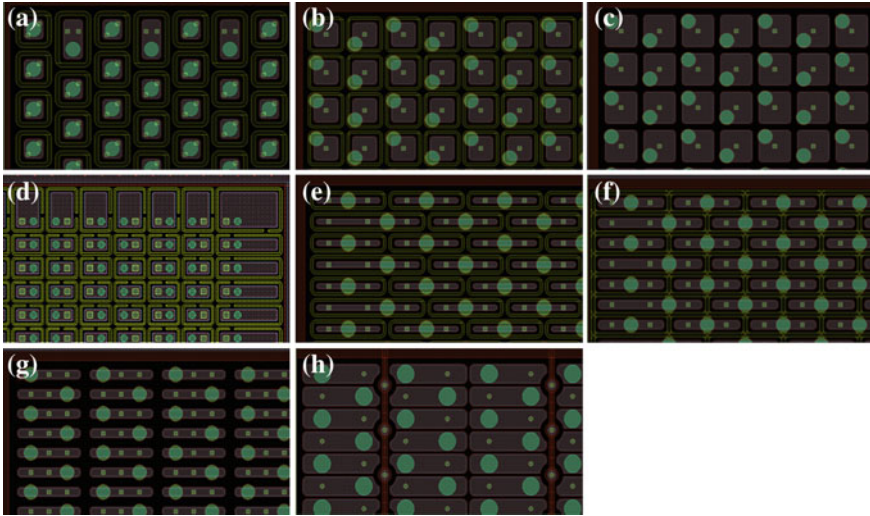

Fig. 7.21 Several potential small pixel cell layouts are shown. The two baseline cell dimensions $50 \times 50\ \mu m^2$ (*upper row*) and $25 \times 100\ \mu m^2$ (*lower rows*) are being studied. The *small squares* show the via connections between metal and implant, the *circles* the bump bond pad. Their positions are always staggered to maximise the bond pitches. Layouts **(a), (b), (d), (e)** feature individual p^+-stops, layout **(d)** has an opening in p^+-stop similar to the CMS forward pixel layout shown in Fig. 6.24. Layout **(f)** shows a common p^+-stop configuration. p^+-spray configurations are shown in **(c)**, **(g)**, **(h)**. With a pixel cell area of $50 \times 50\ \mu m^2$, space is a premium; p^+-spray implants can be larger due to the missing p^+-stops. It is worth to note that only **(d)** and **(h)** can be power tested at sensor level while all others connect to ground ONLY after bump bonding through the electronic chips. **P**unch-**T**hroughs PT connections for each individual cell are basically excluded and a common PT scheme, where one dot connects to four cells is presented in **(h)**

could be reduced to the minimum and a bias grid might not be necessary. This would also largely simplify the layout thus processing and thus probably yield. Another aspect is that every **p**unch-**t**hrough PT contact and the bias rails themselves are areas of slightly degraded efficiency.

Figure 7.21 shows potential layouts for the pixel upgrade. Cell dimensions are either $50 \times 50\ \mu m^2$ or $25 \times 100\ \mu m^2$. These are very challenging; implementations of p^+-stops and on sensor biasing schemes are demanding. Punch-through PT biasing per pixel would use up comparatively more space of the pixel cell and would render substantial amount of area potentially less effective for charge collection; a common PT connection for four pixels seems feasible. A probable scheme without any sensor bias scheme (no PT, polybias or even a bias grid) is envisaged where final connection only happens after bump bonding. In principle an extra metal layer, short-circuiting all pixel cells could be applied during production to make an IV-scan and then be removed before assembly. The figure is educational, showing many different layout possibilities. The only fix point in all configurations is to match the bump bonding pads of sensors and electronic chips. The cell geometries are largely independent

and even longer routings than shown here would be possible. It is important though to maximise space between bond pads. The figure shows a staggered **a, e, f** (also called bricked design) and non-staggered **b, c, d, g, h** cell placement.

Another challenge on system level is the danger of sparking since n-in-p sensors have the high voltage at the sensor cut edges and the bumb-bonded chip on GND-potential nearby. Several options are being explored from dedicated extra passivation[22] on sensor level to under-filling between sensor and chips. This is a real challenge.

CMS will minimise the non-active edges as much as possible with planar technology but is not exploring slim edges with DRIE processes. In case, 3D technology will be used for the innermost layer, DRIE comes with the process thus can be envisaged as active edge processing.

7.2 The CMS Endcap Calorimeter Upgrade for the HL-LHC

The CMS **H**igh **G**ranularity **C**alorimeter HGC, replacing the Run I endcap electromagnetic and hadronic calorimeter, will be the largest silicon based detector, with a larger silicon surface than all existing silicon based tracking detectors altogether. Radiation levels are as challenging as in the innermost trackers and potentially even beyond.[23] According to simulations (see Fig. 7.1 right plot on p.292), radiation levels will reach $\Phi_{\rm eq} = 1.6 \cdot 10^{16}\ n_{\rm 1MeV}/{\rm cm}^2$ (at $|\eta| = 3$) very similar to the innermost layer of the Inner Tracker (pixel detector).

A tracker detects the individual charged particles themselves (even in dense jets) while the calorimeter detects electromagnetic[24] or hadronic showers resulting from the original particle and its interaction with the calorimeter absorber. The massive calorimeter stops the particle fully, absorbing and measuring its full energy while the light-weight Tracker only registers the small amount of ionisation energy of the traversing particle—a **m**inimum **i**onizing **p**article MIP (changing particle energy and momentum as little as possible). The shower energy is still distributed over many cells but each HGC cell gets a much higher signal than sensor cells in a Tracker *allowing quite novel and different design choices*. The HGC has only some sensitivity to MIPs for calibration reasons, but has not been designed to detect individual MIPs.

The details of the following chapter are mainly based on [308] to show the concept in a self-consistent way. In reality, several design choices are still in the flux these days and some improvements and probable (or decided) design changes will be listed in the last paragraph later.

[22] E.g. BCB **B**enzo**c**yclo**b**utene deposition and parylene coating is being investigated.

[23] In case the layout would be enlarged to cover up to $|\eta| = 4$ the expected levels would increase to even $\Phi_{\rm eq} = 10^{17}\ n_{\rm 1MeV}/{\rm cm}^2$.

[24] Electrons, positrons from pair production and γs from Bremsstrahlung.

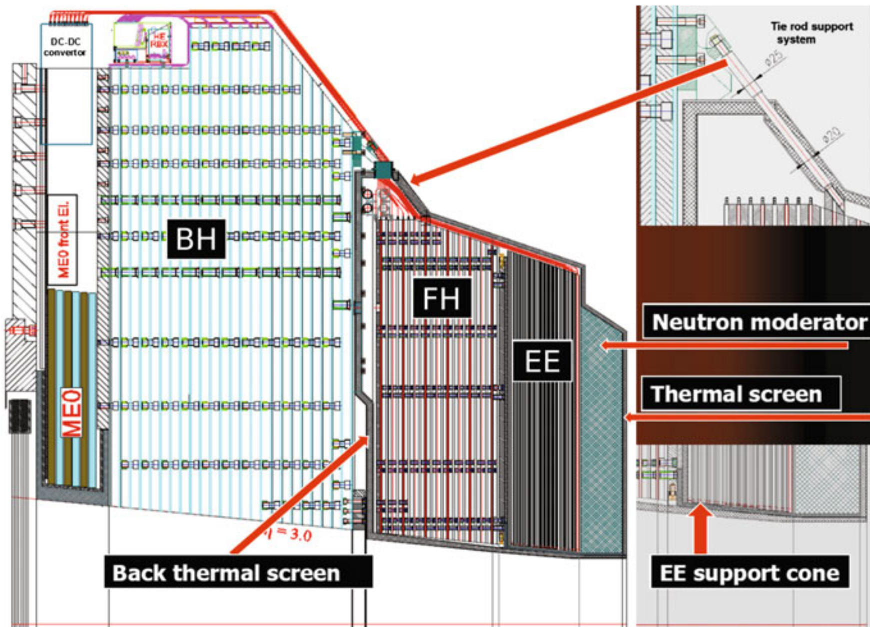

Fig. 7.22 The design layout of the HGC – CMS Technical Proposal [308]

The HGC is a new high-granularity sampling calorimeter, covering $1.5 < |\eta| < 3.0$ with a silicon/tungsten electromagnetic section followed by two hadronic sections, both using brass as the primary absorber material. In the front layers, the active material is composed of silicon pad sensors while the back layers will be instrumented with plastic scintillators. This chapter describes primarily the silicon instrumented sections and discusses the different and/or similar challenges compared to the tracker and the proposed optimized solutions.

The proposed layout is shown in Fig. 7.22 and more details can be found in [308]. The system is divided in three parts, the **E**lectromagnetic **E**ndcap Calorimeter EE, the **F**orward **H**adron calorimeter FH and the **B**acking **H**adron calorimeter BH. The global design parameters are listed in Table 7.5. Details about the silicon sensors can be found in Table 7.7 on p.325.

Inspired by CALICE [320, 335], the HGC approach majorily champions the **p**article **f**low PF concept, but the (a) high radiation tolerance of silicon sensors and (b) the possibility of mass production of highly, dedicatedly structured sensors, was the main reason for the choice. With the relatively high channel granularity, the HGC will be able to continue "*tracking*" inside the calorimeter. This will be especially important in the high pile-up environment, allowing particle separation and energy-particle linking (PF). The example of a simulated event is displayed in Fig. 7.23.

Table 7.5 The table shows the global design parameters as described here (Technical Proposal version, final numbers and design choices might change) [308]

	EE	FH	BH	Total
Layers	28	12	11	41
Active material	Silicon	Silicon	Scintillator	
Absorber material	tungsten/copper	brass	brass	
Channels	4.3M	1.8M		6.1M
Area of Silicon (m^2)	380	209		589
Hadronic interaction length[a] λ	1.3$\approx 25X_0$	3.5	5.5	10.3
Temperature	$-30°$C	$-30°$C	RT	

[a] λ Hadronic interaction length. Muon chambers are normally placed behind at least 10λ corresponding to $P = e^{-(x)/\lambda} = 4.5 \cdot 10^{-5}$ for $x = 10\lambda$, 'preventing' all particles but muons from reaching them

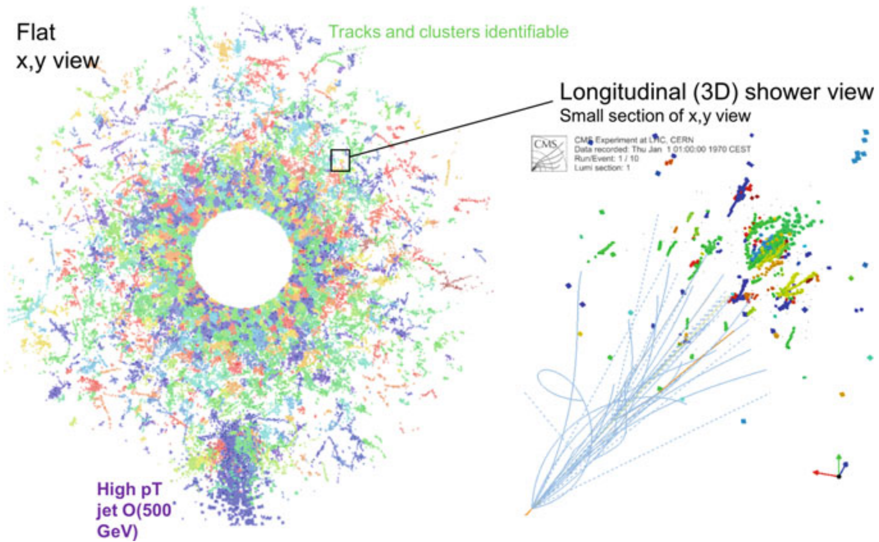

Fig. 7.23 A simulated 200 pile-up event of the CMS HGC. The high granularity allows, even in a high pile-up situation, to associate energy deposition to individual particles. This detailed topical information is ideally suited for the particle flow concept. The different colours aide, to see the energy deposition of different particles. The HGC "*tracks*" particles in the calorimeter [Courtesy CMS]

In EE, tungsten absorbers alternate with active layers. The active layers are composed of a complicated multi-sandwich. A 6 mm thick centre plate consists of Cu with embedded cooling pipes with modules placed on both faces. The situation is illustrated in Fig. 7.24. The FH structure is similar but the absorber will be made of brass with single plane structures in-between, instead of the double-module concept with Cu-absorber cooling plate.

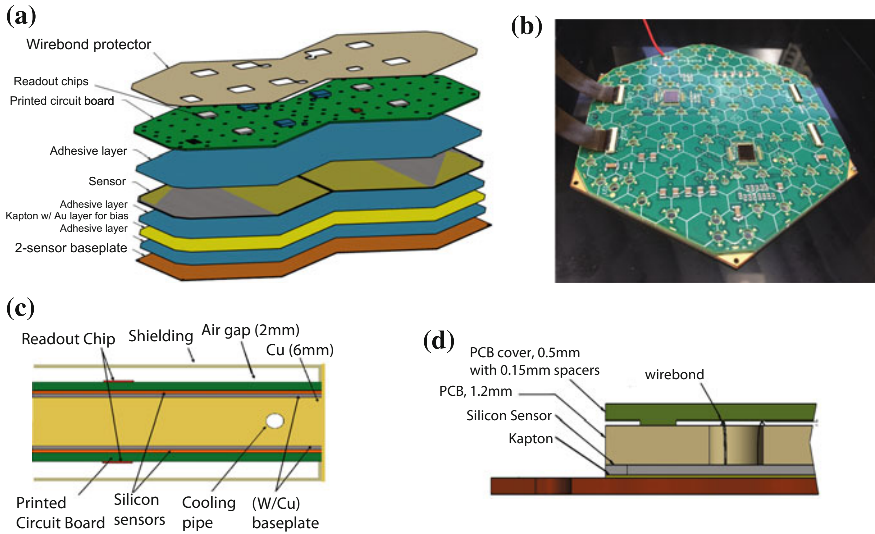

Fig. 7.24 HGC modules on the cooling plate **(a)** the module schematics showing the different layers (printed circuit board, silicon sensors, and W/Cu (75/25%)-baseplate); **(b)** photo of an early half-module prototype; **(c)** Multi-sandwich or double-module (HGC module with W/Cu-base plate – Cu with embedded CO_2 cooling pipe – HGC module with W/Cu-base plate), the copper serves as additional absorber; **(d)** illustrates the detail how the PCB – sensor contact is realized via wire-bonding through a hole in the PCB [Courtesy CMS][308]

The relatively low channel number per wafer, with relatively large pads of $A = 0.5$ and $1\ \text{cm}^2$, is manageable and allows easy wire-bonding from the upper PCB to the pads via holes. All services, a real challenge, have to be routed in the very small gap of 2 mm (see Fig. 7.24c): low voltage and high voltage power in and data out, while the cooling is comfortably located inside the absorber.

To limit the current and power in the silicon sensors after an integrated luminosity of $\mathscr{L} = 3000\ \text{fb}^{-1}$, the sensor temperature needs to be at $-30\,^\circ$C or lower during operation. Similar to the HL-Tracker bi-phase CO_2 will be used. Still the total sensor power is estimated to be ~25 kW. In contrast to the Tracker, material budget is not a problem. While cooling contacts in the Tracker have to be minimized, in the HGC they will be implemented as 6 mm thick copper plates with embedded CO_2 cooling pipes using the full sensor surface as thermal contact. The electronics adds about another 100 kW of power thus a total of 125 kW has to be handled inside the cold EE+FH volume.[25]

As in the CMS Tracker, the levels of radiation varies by three orders of magnitude along radius but also along the z-coordinate, illustrated in Fig. 7.1 on p.292. Accordingly sensors with different active thicknesses of 300, 200 and 100 μm are being considered – the thinner the active volume the higher the radiation tolerance,

[25]Potentially the BH part will be in the cold volume too, to ease thermal interfaces.

rf. Sect. 2.2.4. Electromagnetic and hadron showers deposit much more charge than a MIP, the fact that the signal also scales with thickness is less of an issue for a calorimeter. As a matter of fact, the front-end electronics must have a very large dynamic range as about 5000 MIPS are expected for a 1 TeV electromagnetic shower compared to 50 MIPs at 10 GeV. A charge of 10 pC would correspond to 3000/4500/9000 MIPs in a 300/200/100 μm thick silicon sensor – for physics signals, signal charge thus signal/noise is therefore less of an issue, neither is the total noise. MIP sensitivity is a desirable feature for the HGC, but not a must to detect individually; it will be used mainly to calibrate the detector. Simulations show that MIP calibration can be achieved by combining information of several layers along a MIP 'track' or by implementing dedicated smaller calibration cells in the detector to follow evolution of signal with radiation. This calibration is especially important since, different to a tracker, the calorimeter must be more sensitive to charge/energy deposition measurement than 'simple' 3D local position. The combined multi-layer information will also be useable for muon-tagging.

Large pads together with a thin active volume equals to a large cell capacitance of 40 or 60 pF, compare to about 20 pF for the largest CMS double sensor module or about 5 pF for the largest Phase II Tracker sensor. Together with a planned fast shaping time of 10 to 20 ns, noise will be significant (rf. Sect. 1.5) especially for a standard charge amplifier. The HGC will use a charge amplifier/shaper plus a **T**ime-**o**ver-**T**hreshold ToT circuit. This allows to span the necessary huge dynamic range. After the DC-coupled preamplifier and shaper, a 10-bit ADC is used for measurement of small pulses, and saturates for pulses of $\geq$100 fC. For signals $\geq$80 fC the ToT comparator, starts and stops a TDC, thus providing a digitization measuring the magnitude of large signals. Very simplified, by counting the time-over-threshold of the signal it doesn't matter if the signal saturates in between. S/N values for MIPs (small pulse) are expected to be between 13.7 down to 1.7 depending on sensor thickness and radiation level. Here it becomes clear again, that the individual detector cells are not suited to detect MIPs. Sparsified/zero-suppressed data will be transmitted to the back-end electronics.

The large signals (large gain/signal - fast rise time - low "jitter", the main contributor to the time resolution σ_t) opens up to the possibility to achieve very fast timing information; a precision of about 20 to 50 ps for individual sensor cells is envisaged. The publication [71] gives a first impression of what is achievable. Tests show timing resolution of about $\sigma_t \sim 20$ ps for ~ 5 MIPs in 300 μm thick silicon or 20 MIPs with 120 μm thick sensors. $\sigma_t \sim$700 ps/(S/N) seems a reasonable approximation. Thus precision timing of 50 ps for $Q_{deposited} > 60$ fC seems possible given adequate chip functionality (Fig. 7.25).

In this context, next to the ToT circuit, the FE-ASIC will implement an ADC and **T**ime-**o**f-**A**rrival ToA mode. Timing information in the forward region is especially useful to disentangle primary vertex information derived from arrival time – important in the future high-pile-up environment. In principle the HGC will provide 4D coordinates plus energy information. Different to the tracker, the calorimeter also measures neutral particles (e.g. photons, neutrons, π°, etc.); the timing information here will allow to link these particles to the primary vertex.

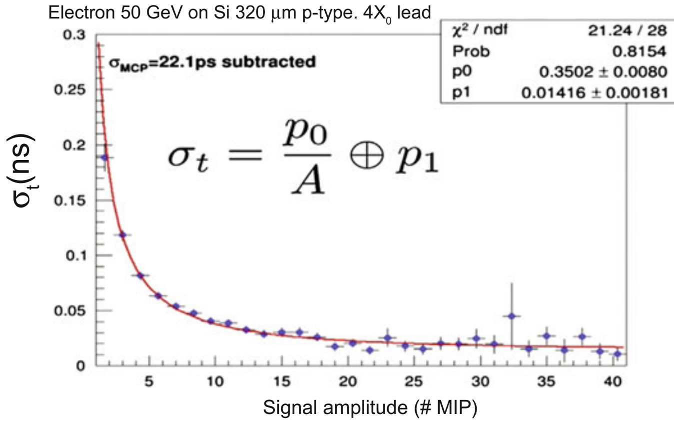

Fig. 7.25 Fast timing with silicon at large charge deposition. 300 μm thick silicon diodes have been tested in a test beam at CERN where 50 GeV electrons have been directed on a $4 \cdot X_o$ lead absorber. $\sigma_t \sim 700$ ps/(S/N) seems a reasonable approximation according to [71]

Table 7.6 The table shows the similar and different challenges of the CMS-HGC and CMS -HL-Tracking system

	HGC	Tracker
Radiaton tolerance	$1 \cdot 10^{16}\ n_{1\mathrm{MeV}}/\mathrm{cm}^2$ – more neutrons	$1.6 \cdot 10^{16}\ n_{1\mathrm{MeV}}/\mathrm{cm}^2$ – more charged particles
Occupancy	mid-size pads	thin pitch strips and/or pixels
Channel capacitance	up to 60 pF	up to 5 pF
Dynamic range	1–5000 MIPs (shower)	1 MIP
Calibration	MIP calibration is difficult	no problem
Material budget	high mass density necessary	$\sim$ largest challenge to keep mass low
Services	only 2 cm gap for routing	space and stringent mass constraints
Alignment	hundreds of micrometer	couple of micrometers
Data rate	similar challenge	
L1-trigger	similar challenge – large BE electronics system (combinatorics)	
Cooling	similar challenge CO_2	

A comparison of similarities and difference between HGC and the CMS HL-Tracker is given in Table 7.6.

As for the Tracker, due to bandwidth limitations, not all information from the HGC can be transmitted from the detector to the back-end electronics in the service cavern. Since the calorimeters are essential for the Level-1 trigger, part of the information must be summed up and congregated. 2×2 sensor cells from EE and FH will be summed up for every other sensitive plane and checked against a threshold; the

Table 7.7 Different Silicon sensors for different locations depending on the radiation environment (Technical Proposal version [308], final numbers might change)

Thickness (μm)	300	200	100
Maximum dose (Mrad)	3	20	100
Max fluence (1MeV$_{eq}$/cm^2)	$6 \cdot 10^{14}$	$2.5 \cdot 10^{15}$	$1 \cdot 10^{16}$
EE region	R > 120 cm	120 > R > 75 cm	R < 75cm
FH region	R > 100 cm	100 > R > 60 cm	R < 60cm
Si wafer area (m^2)	290	203	96
Cell size (cm^2)	1.05	1.05	0.53
Cell capacitance (pF)	40	60	60
Initial SN for MIP	13.7	7.0	3.5
S/N after $\mathscr{L}$=3000 fb^{-1}	6.5	2.7	1.7

congregated (also sparsified) info will be sent at the full rate of 40 MHz by the front-end electronics to the service cavern, where trigger primitives will be generated. The back-end electronics will then check for shower signatures taking the projection from the collision region into account. The complexity will be similar to the one of the CMS HL-Track Trigger.

HGC Sensors

The different sensor configurations and their expected signal/noise values before and after irradiation are listed in Table 7.7. Given the very large area, a large numbers of wafers is necessary and cost must be minimised. About 14000 EE plus 7500 FH modules will be installed where a module contains 2 sensors (rf. Figure 7.24) thus either 43000 6 in. or 21500 8 in. wafers are needed. The sensors are hexagonal shaped (see Figs. 7.26 and 7.27), which allows tiling as a honeycomb structure. This layout maximizes the surface-use of a circular wafer. On the other hand final offline reconstruction and especially fast trigger processing must be able to cope with this geometry and pattern matching, thus square cells (and square sensor) geometries had initially been evaluated as well. The studies proved that hexagonal cells or ok.

The sensors will be *n*-in-*p* as baseline (with *p*-in-*n* still an option) DC-coupled wafers with hexagonal pads of about A=1 or 0.5 cm^2. *n*-in-*p* would be even more cost effective since no extra processing for p^+-stops is necessary to guarantee pad isolation.

Active wafer thickness will be 300, 200 and 120 μm (rf. Table 7.7). While the first two feature 128 cells of size 1 cm^2, the 120 μm thick sensor has 256 cells of size 0.5 cm^2. The cell capacitance is large due to the big pad area plus the small thickness D of the sensor. To keep it under control, the pad size is halved for the 120 μm thin sensor. This limits the maximal capacitance to 60 pF. The individual cells are then directly wire-bonded to pads on a PCB and then routed to the readout chip located on the PCB.

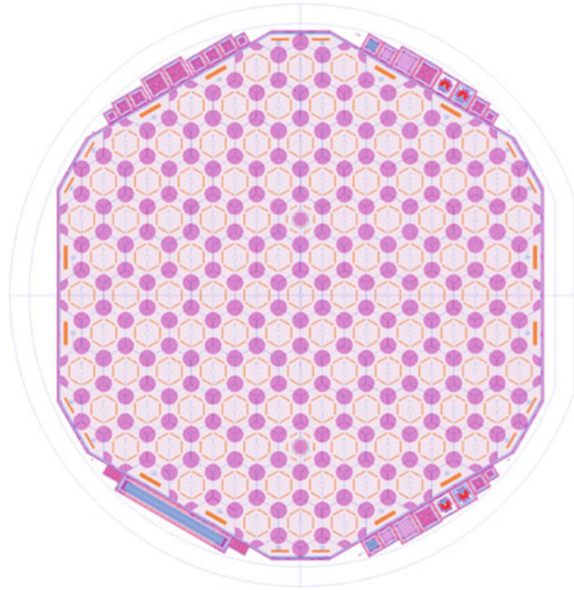

Fig. 7.26 The HGC hexagonal sensor design. The hexagonal shape allows to utilize 137 cm^2 of the wafer and the sensors can be tiled later to create a homogeneous area. The pads/cells are hexagonal too

On a 6 in. wafer, the sensor covers an area of 137 cm^2 compared to the nominal $10 \times 10\ cm^2$ of a square sensor thus a surface gain factor of 1.3 is achieved, reducing cost. The wafer layout and a photo of the first 6 in. prototype is shown in Figs. 7.26 and 7.27.

The active thickness will either be achieved by direct processing on physically thin sensors or by using deep-diffused wafers (rf. Sect. 1.9.3 on p.93 and Sect. 7.1.1). The use of high-volume commercial lines is a must and explorations are ongoing if 8 in. wafer production will be possible, while the baseline is 6 in. wafer processing. Deep-diffused material is only available on 6 in. and not for 8 in. for the moment, thus 8 in. processing would have to happen on thin wafers. One will see what the future holds.

Different to the tracker volume, the HGC radiation environment is strongly dominated by neutrons instead of charged particles, thus all defect engineering concepts based on compensation do not really hold (rf. for example Fig. 2.16 on p.155). The suppression of built up of negative space charge (acceptors) after charged hadron radiation in oxygen enriched silicon does not help (rf. Sect. 2.2.2 on p.154).

Still, radiation tolerance studies have shown adequate performance after the full expected fluence, but active thickness has to be adapted depending on the integrated fluence level thus location. The advantage of thin sensors for very high radiation levels has been discussed in Sect. 2.2.4. As long as the electric field strengths are high enough, charges are drifting fast enough to induce a sizeable signal in the readout electronics (remember, no charge amplification). Another advantage of pad detectors is, that both electrons and holes are contributing equally to the signal formation (rf. also Sect. 1.3.2). The sensors and modules must feature a very high voltage robust-

Fig. 7.27 Photo of an 8 in. HGC prototype wafer scaled down by a factor 2. The hexagonal structure optimizes the use of the wafer surface and can finally be tiled to cover a flat surface with minimal inactive areas. Below, a zoomed view of the right wafer corner is shown. All readout cells are hexagonal. The *circles* at the cell corners are openings in the passivation to allow wire-bonding. Two wafer cells have an extra smaller pad in the centre with smaller capacitance (thus lower noise) allowing for more precise calibration of MIP signals

ness, since sensor bias voltages up to 900 V after radiation are foreseen – a solvable challenge. Given early results of a radiation study, 100 μm might be just too thin and the optimum lies at a thickness of 120 μm. The Tracker is undergoing the same kind of optimization, especially since the deep-diffused material or thinned sensors allows to choose the active thickness quite precisely. In case the active/instrumented volume increases from $|\eta| \leq 3$ to $|\eta| \leq 4$, the highest integrated fluences would even reach

$\Phi_{eq} = 10^{17}\ n_{1\mathrm{MeV}}/\mathrm{cm}^2$. Then, even thinner sensors like 50 μm active thickness or 3D sensors would be taken into account.

The HGC is probably the most courageous endeavour due to its sheer size and complexity, but it will also lift the possibilities of particle detection to a whole new level.

Some Further Evolution

Usually, on the path from a technical proposal towards a Technical Design Report, several formerly open design choices are decided and new opportunities are exploited. Some HGC examples, expressing this kind of evolution, are briefly given here. Also the naming convention has been rationalized, FH+BH have been combined into EC-H or ECH **E**ndcap **C**alorimeter-**H**adron and EE is now named EC-E or ECE **E**ndcap **C**alorimeter-**E**lectromagnetic.

Interestingly, the hadron calorimeter absorber material has been changed from brass to stainless steel. Steel is much cheaper than brass and offers more engineering possibilities.

Initially mechanical alveolar structures, where wedge-shaped structures would have been slided-in, competed with more standard disk-structures, and finally the disk-concept was selected (stack of 30° or larger segmented cassettes and tungsten plates with spacers). Finally, the whole endcap calorimeter (EE, FH, BH) will reside in a single cold volume with a single thermal and gas/humidity barrier, while initially the BH could have stayed in a warm volume. This allows to have all feed-throughs at the back with more space available. The cold operation opens up several possibilities for the BH part, namely the instrumentation with silicon sensors at lower radius with higher radiation levels and restricting scintillators to lower irradiated volumes. The scintillators, in turn, can then use SiPMs directly on tile. Cold operation lowers the radiation induced dark noise of the SiPMs. The new considered layout is displayed in Fig. 7.28.

As mentioned earlier, the single module sensitivity is not really good enough for MIPs (with S/N down to $\sim$1.7) but combining information from several layers opens the door for muon-tagging (MIPS), besides the mentioned calibration. This is especially true for layers at larger z where the showers of other particles, for example pions, have already decayed. The new additional high precision silicon layers in the BH (at lower radius), therefore facilitate the muon-tagging significantly.

Most probably, 8 in. wafers will be used (now baseline), reducing the number of wafers significantly and increasing individual sensor surface largely. Given the increased sensor surface, single-sensor-modules are becoming attractive, being easier to manufacture and to handle with still a ‘reasonable’ total number of modules—ballpark: the use of 6 in. wafers would result in 40,000 *single*-sensor-modules or 20,000 *double*-sensor modules while the use of 8 in. wafers would reduce this to a reasonable number of 20,000 *single*-sensor-modules. Along the way, rectangular sensor cells instead of hexagonal ones had been discussed, ending up with

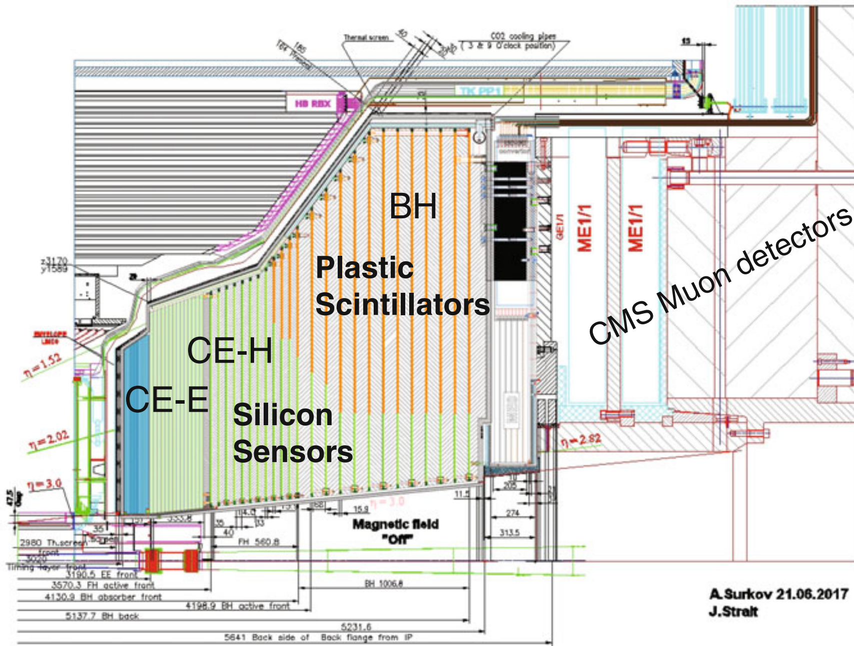

Fig. 7.28 The design layout of the HGC—version considered for the TDR, shown at the HGC Comprehensive Review June 2017. This layout, compared to Fig. 7.22 gives a nice impression of how such detectors evolve. The whole detector will reside in a cold volume with balanced instrumentation of plastic scintillator with SiPMs on tile (*orange*) at higher z and higher radius and silicon sensors (*green*) now also in the BH section at lower radius [Courtesy CMS]

the original hexagonal concept. Today 2017, no deep-diffused wafers exist in 8 in. The idea is to use physically thin FZ material for the D= 300 μm and 200 μm case and use EPI material for the D = 100 μm thin case.