# Chapter 6
# CMS: Increasing Size by 2 Orders of Magnitude

The **C**ompact **M**uon **S**olenoid CMS operates at the **L**arge **H**adron **C**ollider LHC at CERN. It was conceived to study proton–proton (and nucleon–nucleon) collisions at a centre-of-mass energy of 14 TeV (5.5 TeV per nucleon for Pb-Pb) and at luminosities up to $\mathcal{L} = 10^{34}$ cm$^{-2}$s$^{-1}$ ($10^{27}$ cm$^{-2}$s$^{-1}$). The detector is characterized by high hermeticity and compactness. A 3.8 T superconducting solenoid surrounds a silicon tracker, an **e**lectromagnetic **cal**orimeter ECAL and a **h**adron **cal**orimeter HCAL. The return yoke is instrumented with muon detectors covering most of the $4\pi$ solid angle. Forward sampling calorimeters extend the CMS pseudo-rapidity coverage to high $\eta$ values. The detector is fully described in [4, 290, 292, 299]. Its layout can be seen in Fig. 6.1. The CMS detector was fully constructed on the surface and was lowered 100 m down to the cavern in 11 slices plus 2 very forward calorimeters (**H**adron **F**orward HF). Individual slices weighted up to 2500 t summing up to a weight of 14,000 t. The inner detectors like calorimeters and tracking detectors were inserted and instrumented later down in the cavern. A schematic how CMS identifies (IDs) basic particles is placed in the Appendix in Fig. B.3. In the following the CMS silicon tracker (Phase 0 until 2016) will be described with an update section describing the pixel upgrade installed in Spring 2017.

December 1999, on the eve of the last millennium, the CMS Tracker Collaboration decided to construct its tracker fully out of silicon, summing up to 206 m$^2$ silicon surface.[1] The design goal of the tracker is to operate with a maximum of 1 – 3% occupancy at all layers and a good impact parameter resolution in $R\phi$ and $Rz$.

[1] Prior to this meeting about 70 m$^2$ silicon plus ∼200 m$^2$ **M**icro**S**trip **G**as **C**hamber MSGC were specified in the initial **T**echnical **D**esign **R**eport TDR [290].



F. Hartmann, *Evolution of Silicon Sensor Technology in Particle Physics*,
Springer Tracts in Modern Physics 275, DOI 10.1007/978-3-319-64436-3_6

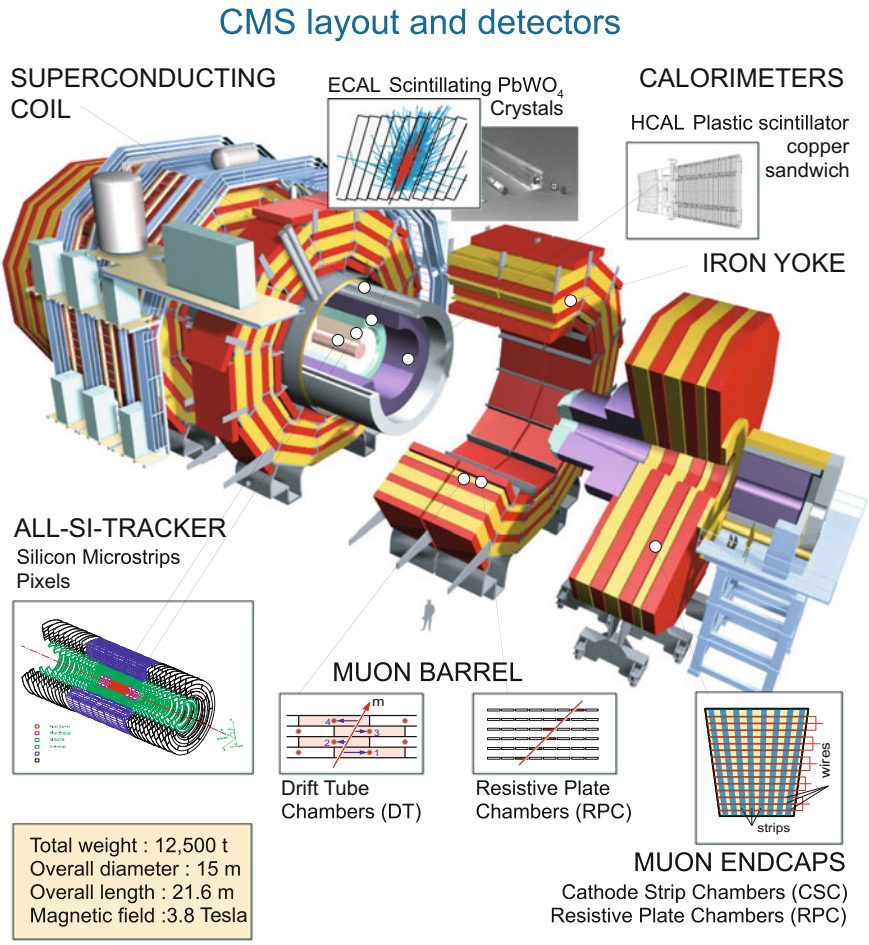



**Fig. 6.1** The CMS detector and its components [Courtesy of CERN, 333]

The requirements are

- efficient and robust pattern recognition
- fine granularity to resolve nearby tracks; e.g. from 3-prong $\tau$-lepton decays
- fast response time to resolve bunch crossings
- ability to reconstruct tracks from very low $p_T$ in the order of 100 MeV up to 1 TeV
- ability to reconstruct heavy objects translating to $1-2\%$ $p_T$ resolution at $\sim$100 GeV
- ability to operate in a crowded environment at 10 cm from primary vertex
- ability to tag b/$\tau$ through secondary vertices
- reconstruction efficiency: 95% for hadronic isolated high $p_T$ tracks and 90% for high $p_T$ tracks inside jets
- ability to operate in a very high radiation environment
- efficient and precise tracking in the pseudo-rapidity range $|\eta| \leq 2.5$.

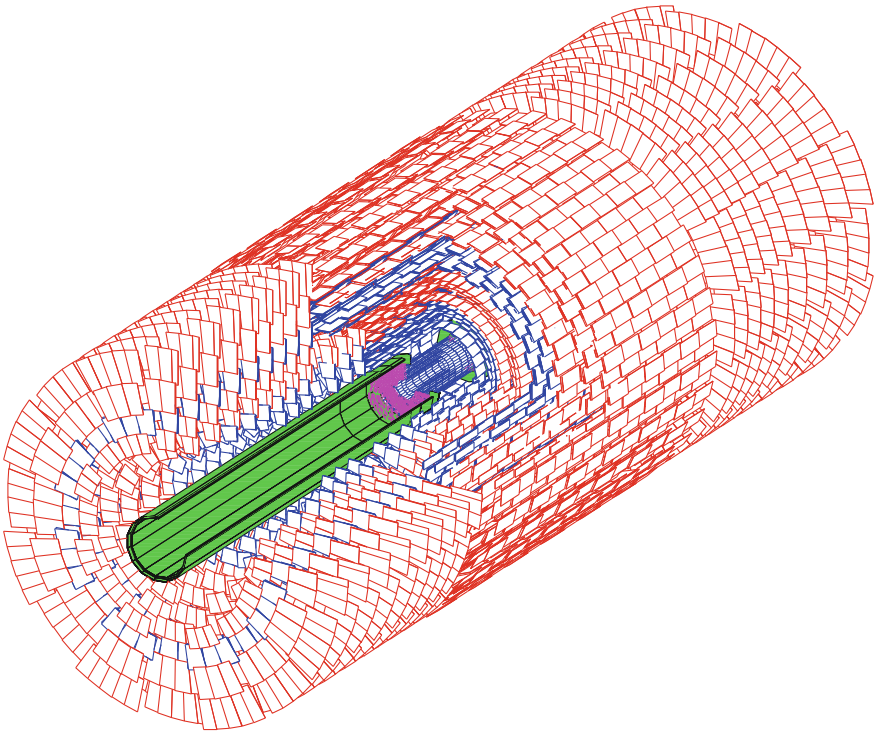


**Fig. 6.2** The CMS all-silicon tracker layout (figure displays the detector throughout 2008–2016) [129]. The pixel detector is located in the centre. It is surrounded by 10 strip barrel layers supplemented by 3 small forward and 9 large forward disks. The barrel layer 1, 2, 5, 6 and endcap rings 1, 2, 5 consist of double modules, mounted back-to-back with a stereo angle of 100 mrad. The overall length and diameter is 5.4 and 2.4 m, respectively. 1632 pixel (1440 in phase 0) and 15148 strip detector modules build up the tracker. All in all 127 million pixel cells (66 million in phase 0) and 10 million strips are available. The tracker directly surrounds the beam pipe and is enclosed by the electromagnetic calorimeter; no additional inner tracking device like a drift chamber or transition radiation tracker is installed [128]

The result is a small pixelated detector in the centre with a large volume strip detector surrounding it. All in all a cylinder of 2.5 m diameter and 5.4 m length is instrumented with 206 $m^2$ of active silicon detectors. The general 3D layout can be seen in Fig. 6.2. At the design luminosity of $10^{34}\,cm^{-2}s^{-1}$ the tracker will see above 500 particles from about 20 overlapping proton–proton collisions per bunch crossing, every 25 ns. With a homogeneous field of 3.8 T, low momentum particles spiral at low radii and therefore the occupancy decreases faster than $1/r^2$ as it would without a magnetic field. Fast readout and high granularity are needed to achieve the low design occupancy, get a reasonable track separation, good primary vertex resolution and good impact parameter resolution $\sigma_{d_0}$.

The CMS tracker as of the year 2017, is composed of a pixel detector with 4 barrel[2] layers and a silicon strip tracker with 10 barrel detection layers extending outwards to a radius of 1.1 m. Each system is completed by endcaps which consist of three disks in the pixel detector and three small plus nine large disks in the strip tracker on each side of the barrel extending the acceptance of the tracker up to a pseudo-rapidity of $|\eta| < 2.5$. A schematic with detailed individual module positions can be found in the Appendix in Fig. B.5.

## 6.1 The CMS Pixel Detector – Phase 0 – 2008 – 2016

In order to keep the occupancy at or below 1%, pixelated detectors have to be used at radii below 10 – 20 cm. For a pixel size of $100 \times 150\,\mu\text{m}^2$ in $R\phi$ and $Rz$, respectively, which is driven by the desired impact parameter resolution, the occupancy is of the order of $10^{-4}$ per pixel at LHC bunch crossing rate. The pixel sensors are discussed in Sect. 6.4.2. With analogue[3] signal interpolation, a hit resolution of 10 $(R\phi) \times 20\,\mu\text{m}(Rz)$ can be achieved (see also Sect. 6.6.2). The magnetic field of CMS, inducing a Lorentz angle, enhances the charge spread to more than one cell, thereby increasing resolution by interpolation of charges. While the magnetic field is naturally perpendicular in the barrel region, the forward region is built in a turbine-like geometry to profit from the effect (see Figs. 6.3 and 6.4). Due to the very high number of channels, the pixel readout is zero-suppressed (sparsified) and only for pixels being hit, analogue data along with the pixel cell address is being sent to the data acquisition. The data rate is dynamic, depending on beam background and instantaneous luminosity. With limited bandwidth (inside the chips themselves) at very high rates not all information can be sent out and an inefficiency in the order of a couple of percent is expected at nominal LHC luminosity ($\mathcal{L} = 10^{34}\ \text{s}^{-1}\text{cm}^{-2}$) and the maximum Level-1 trigger rate of L1 = 100 kHz in the innermost layer.

Three pixel barrel layers are located at radii 4.4, 7.3 and 10.2 cm. The pixel barrel is 53 cm long. Two pixel forward disks at $z = \pm 34.5$ and $z = \pm 46.5$ cm guarantee 3D space points coverage up to $|\eta| < 2.5$. With a total area larger than 1 $\text{m}^2$ the CMS pixel detector is as large as the complete DELPHI silicon vertex detector. The pixel detector layout is displayed in Fig. 6.3. It is instrumented with 768 hybrid pixel modules in the barrel plus 672 modules (plaquettes) in the forward disks for a grand total of 66 million pixels.

Figure 6.5 shows a very interesting photo demonstrating the complexity of tracking detectors in the very tight spaces of a **H**igh **E**nergy **P**hysics HEP Experiment. It displays the so-called supply tube since it supplies the detector with all necessary services: cooling pipes, control fibres, power for electronics, and sensor high voltage, also the electrical to optical converters are situated here. The readout fibres are prominently visible since the other parts are buried beneath. A very complex object!

[2] It was 3 barrel layers and 2 forward disks from 2008 to 2016.

[3] Analogue data represents pulse height to allow a centre-of-charge analysis later.

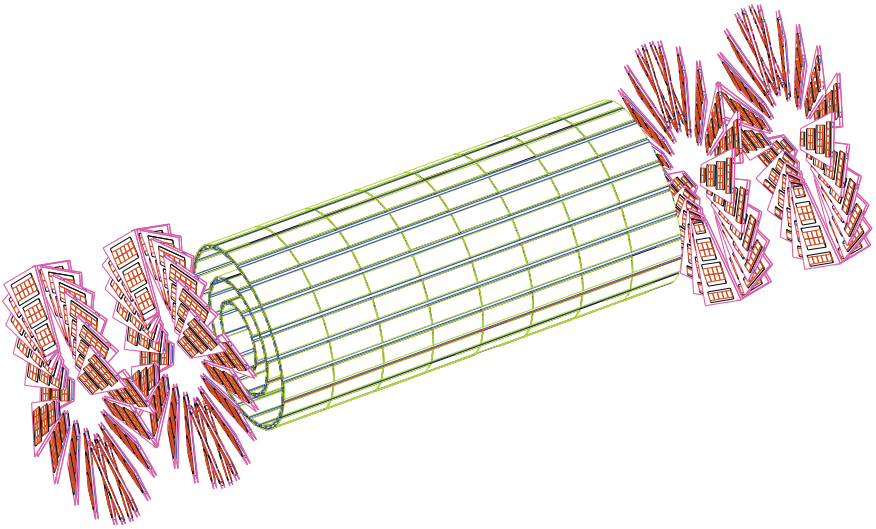


**Fig. 6.3** The CMS pixel detector layout. Three 53 cm long barrel layers at radii 4.4, 7.3 and 10.2 cm plus two forward disks at z = ±34.5 and z = ±46.5 cm guarantee the 3D space points coverage up to $|\eta| < 2.5$. The barrel (forward) pixel detector contains 768 (672) pixel modules (plaquetts) with 48 million (18 million) pixel cells covering 0.78 (0.28) $m^2$ [87]

**Fig. 6.4** One of the forward half-disks. Please note the turbine configuration emphasising the Lorentz angel spread. To achieve a wedge geometry seven different sensor geometries, assembled on so-called panels, haven been produced. Eight half-disks, forming 4 full disks, 2 at each end, finally surround the beam pipe [Courtesy of CMS/CERN/FNAL]

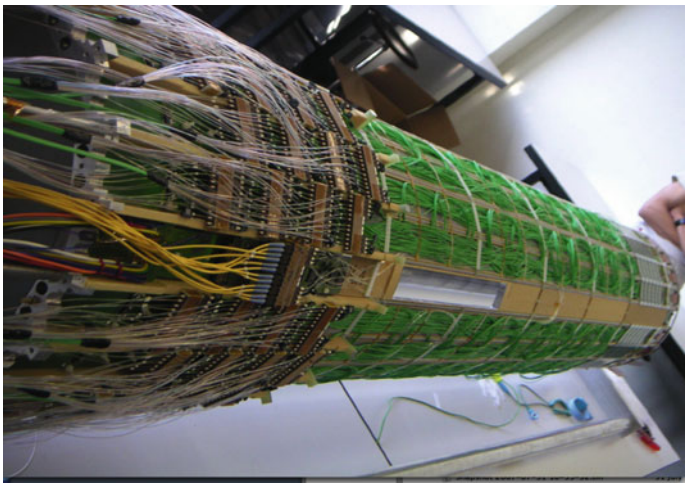


**Fig. 6.5** The supply tube of the CMS Pixel Detector (2008 – 2016) [Courtesy of CMS/CERN/PSI]

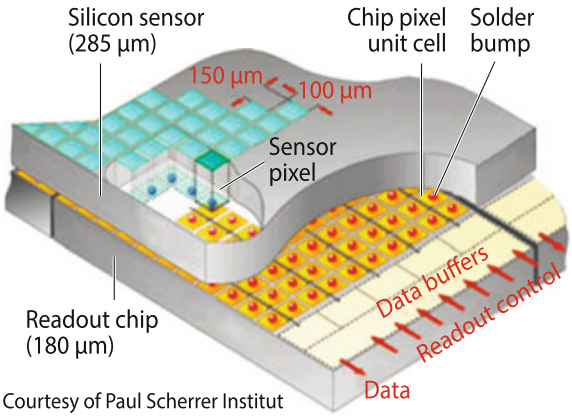



**Fig. 6.6** Schematic of the CMS pixel with 100 $\times$150 $\mu m^2$ pixels per cell. A good example of a hybrid active pixel sensor *HAPS* providing intrinsic two-dimensional position information. The electronics amplifier chip has the same size and channel pattern as the sensor and is bump bonded (flip-chip bonded) to the sensor. A more detailed schematic can be found in the next Fig. 6.7. A photo of the bump lithography is presented in Fig. 1.73 on p. 108

To quote my colleague and CMS Pixel coordinator Roland Horisberger:

> The supply tube is the place where all troubles meet.

The very reliable and radiation tolerant Hybrid Active Pixel Sensor HAPS technology has been chosen. The schematic of the electronic-sensor stack can be seen in Fig. 6.6 and a full pixel module including a module assembly picture is shown in Fig. 6.7.

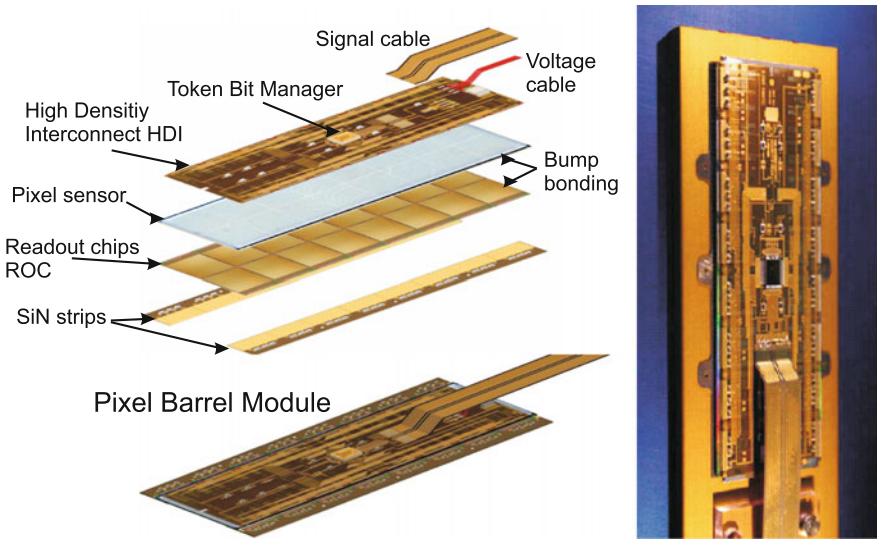



**Fig. 6.7** CMS Pixel module – barrel type [Courtesy Paul Scherrer Institute, CERN]

The pixel detector is essential for $b$ & $\tau$-tagging and primary vertex reconstruction. It also determines the track seeds towards the outer strip tracker. The detector is designed and built in half-shells allowing installation with the central LHC beam pipe[4] in place. Also the services are done with fast extraction and installation in mind. The reason is to allow maintenance and repair in a typical LHC **Y**ear-**E**nd-**T**echnical **S**top YETS and the replacement of the innermost layer (of the phase 0 detector) was initially foreseen 2 – 3 years into the run due to high radiation damage at low radius. In reality, the detector had been removed and repaired in the YETS 2009 and also in long shutdown 1 2013/2014. This is illustrated impressively with the photos Figs. 6.8 and 6.9 of the barrel and forward pixel detector installed inside CMS. Instead of replacing the innermost layer, a completely new detector (Phase I, see later) has been installed in spring 2017.

## 6.2 The Pixel Phase I Upgrade – Installed February/March 2017

As indicated in the previous section, the pixel detector design allows a replacement of its innermost layer during the lifetime of CMS and the initial plan was to replace just the innermost layer when radiation renders it dis-functional. In the end, the reason to replace the full detector was not driven by radiation damage but by the

[4]Installation of the beam pipe requires a high temperature bake-out to allow a good final vacuum which takes several weeks.

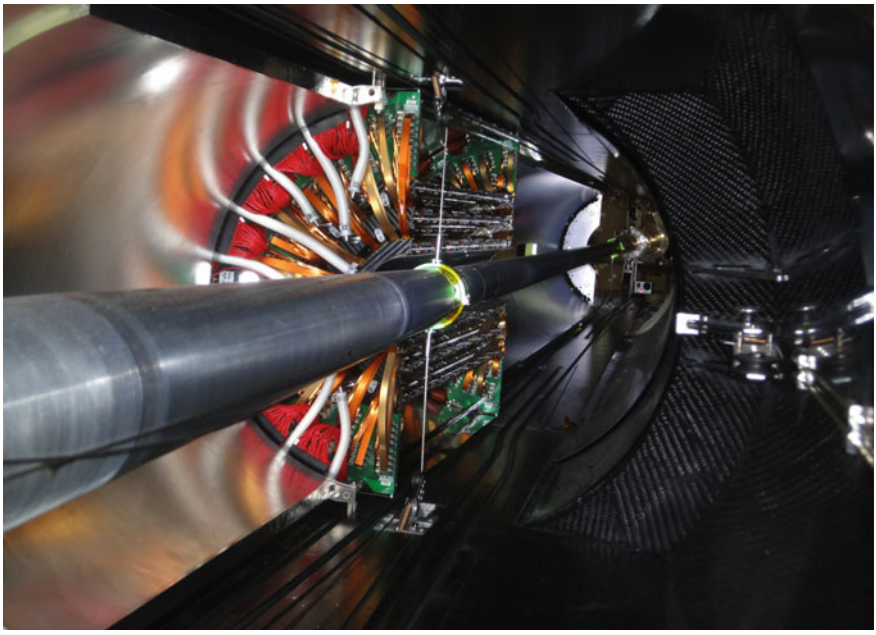


**Fig. 6.8** Half of the Barrel Pixel detector just placed into the heart of CMS – December 2014 at the end of Long Shutdown 1. One can see the three layers and the end flange where the cooling pipes fan out. The half-shell concept allows fast extraction and insertion WITH the central beam pipe of the LHC in place [Courtesy of CMS/CERN]

"beyond-LHC-design" increase of instantaneous luminosity achieved by the LHC. As described earlier, the numbers of on-chip buffers is limited and would lead to too high inefficiencies at very high luminosities, and also the data transfer from detector to the back-end is strictly limited. Due to these reasons, CMS decided to replace the full pixel detector instead of only the innermost layer. At the time of writing, the new detector has just been installed.

This is done in a conservative way but still using optimised design and readout electronics. The newly optimised readout chips are performing much better and are also more radiation tolerant – even the innermost layer will withstand an integrated luminosity of $\mathcal{L}$ =500 fb$^{-1}$ instead of the originally anticipated 250 fb$^{-1}$. The full information can be found in [309] and [304] while only a very brief summary is presented here, focussing on interesting facts demonstrating the evolution in the field. With a new smaller beam pipe[5] in place, it was possible to move the inner layer closer to the interaction point. A fourth layer has been added at higher radius improving the track link towards the outer strip tracker. Layer 2 and 3 radii are almost unchanged. The four barrel layers (BPIX) are now located at radii 3.0, 6.8, 10.9 and 16.0 cm; they span ±274.4 mm. The forward disks (FPIX) are moved further in

[5]New outer diameter D = 45 mm with respect to old D = 59.6 mm.

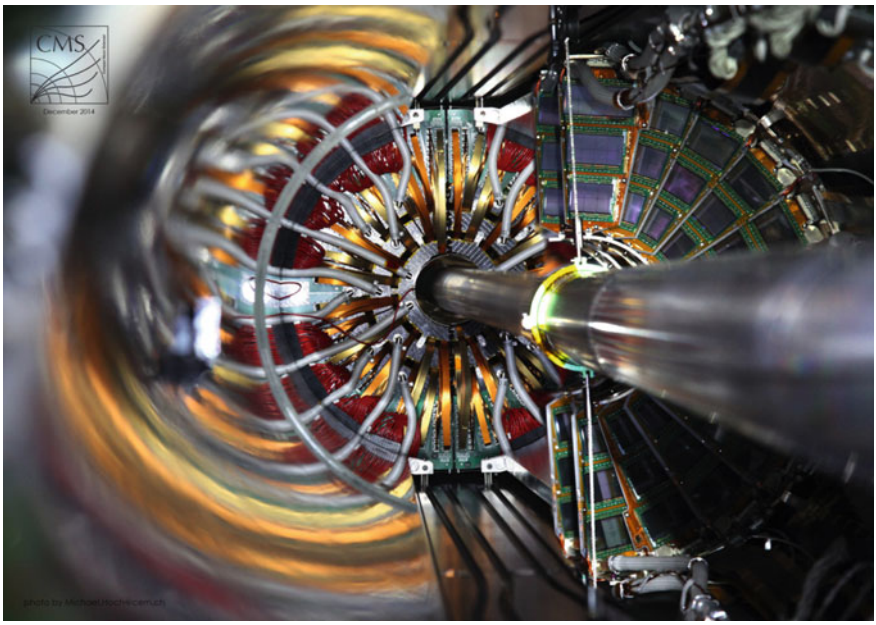



**Fig. 6.9** The full Barrel Pixel detector is in place and half of one end of the forward pixel detector is in place (December 2014). The individual pixel modules are visible. The module wedge form has been achieved using sensors of different sizes, housing a different number of ROCs (6, 8 or 10) [Courtesy of CMS/CERN]

and are now located at z = ±29.1, ±39.6 and ±51.6 cm; they span the radius from 4.5 to 16.1 cm. The whole detector (barrel and forward) features now one single module geometry for simplicity. The forward modules are still placed into a turbine structure to profit from the Lorentz angle. Due to this, the two mechanical structures for the positive and negative detector end, are mirrored instead of simply rotated, otherwise the Lorentz angle would shift in the wrong direction in one end. The pixel cell count increases from 18 to 48 and 48 to 79 million in the FPIX and BPIX respectively. Table 6.1 gives details about barrel and endcap configurations also compared to the first CMS pixel detector – Phase 0.

The layout of the new detector is displayed in Fig. 6.10. A photo of the forward and barrel part can be seen in Fig. 6.11.

The new detector design addresses the following points:

- increase data transmission by changing from 40 MHz analogue transmission to 320 MHz digital transmission
- increase on-chip buffers to limit the inefficiency and therefore to cope with $\mathcal{L} = 2 \cdot 10^{34}\ \mathrm{s}^{-1}\mathrm{cm}^{-2}$
- improve radiation tolerance of the readout chips (ROCs)

**Table 6.1** The table expresses the main difference between the Phase 0 and the Phase I detector. Refer also to Figs. 6.3 and 6.10 displaying the corresponding layouts. The 2nd number in the Phase 0 module column refers to the half-modules. Impressively, the material budget at $\eta = 0$ is lower for phase I despite having one more layer

| | Phase I, since Spring 2017 | | | Phase 0, 2008 – 2016 | | |
|---|---|---|---|---|---|---|
| Cooling | $CO_2$ (bi-phase) | | | $C_6F_{14}$ monophase | | |
| Powering | DC-DC powering | | | direct power | | |
| Readout | 320 MB/sec (digital coded) | | | 40 MHz (analogue coded) | | |
| Pixel size | $100 \times 150\,\mu m$ | | | $100 \times 150\,\mu m$ | | |
| Material $\eta = 0$ | $X/X_0 = 5.5\%$ | | | $X/X_0 = 6\%$ | | |
| Material $\lvert\eta\rvert = 1.6$ | $X/X_0 = 20\%$ | | | $X/X_0 = 40\%$ | | |
| Layer | Radius (mm) | Facets | # Modules | Radius | Facets | # Modules |
| Beampipe | 22.5 | | | 29.8 | | |
| 4 | 160 | 64 | 512 | – | – | – |
| 3 | 109 | 44 | 352 | 102 | 46 | 320/32 |
| 2 | 68 | 28 | 224 | 73 | 30 | 224/32 |
| 1 | 30 | 12 | 96 | 44 | 18 | 128/32 |
| total | | | 1184 | | | 672/96 |
| Disk | Radius 45 – 161 mm | | | Radius 60 – 150 mm | | |
| | $z$ pos (mm) | Mod outer | Mod inner | $z$ pos | Sensors/blade | # Blades |
| 1 | $\pm$291 | 2*34 | 2*22 | $\pm$345 | 7 | 2*24 |
| 2 | $\pm$396 | 2*34 | 2*22 | $\pm$465 | 7 | 2*24 |
| 3 | $\pm$516 | 2*34 | 2*22 | – | – | – |
| Total | | 448 | | | | 96 |

- improve the hermiticity of the detector and establish an overall 4-hit coverage to increase redundancy and allow for a quadruplet track finding algorithms and therefore better track seeding capability. This highly improves the track quality/resolution and pattern recognition
- reduce mass wherever possible especially in the tracking volume $|\eta| \leq 2.5$
- reduce the innermost radius to improve impact parameter resolution $\sigma_{d_0}$
- add a layer at higher radius to significantly improve the impact parameter resolution in $z$ direction due to the larger lever arm (the strip detector at higher radii has only a good resolution in $\varphi$ not in $z$)

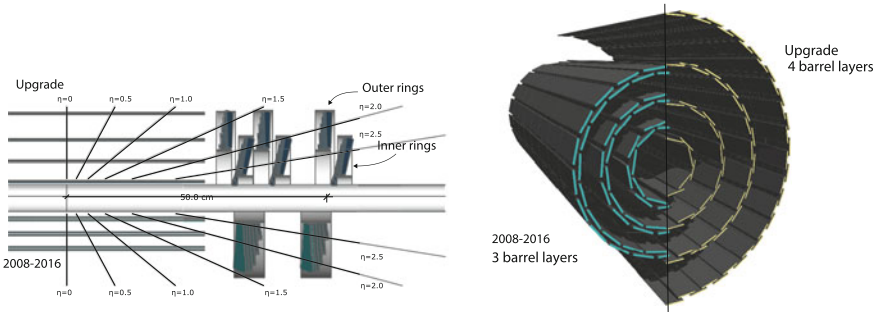



**Fig. 6.10** Layout of the CMS phase I pixel detector. One can see the increase from 3 to 4 layers in the barrel and from 2 to 3 disks in each endcap. In addition the inner layer moves *lower* in radius thus closer to the primary interaction point [304]

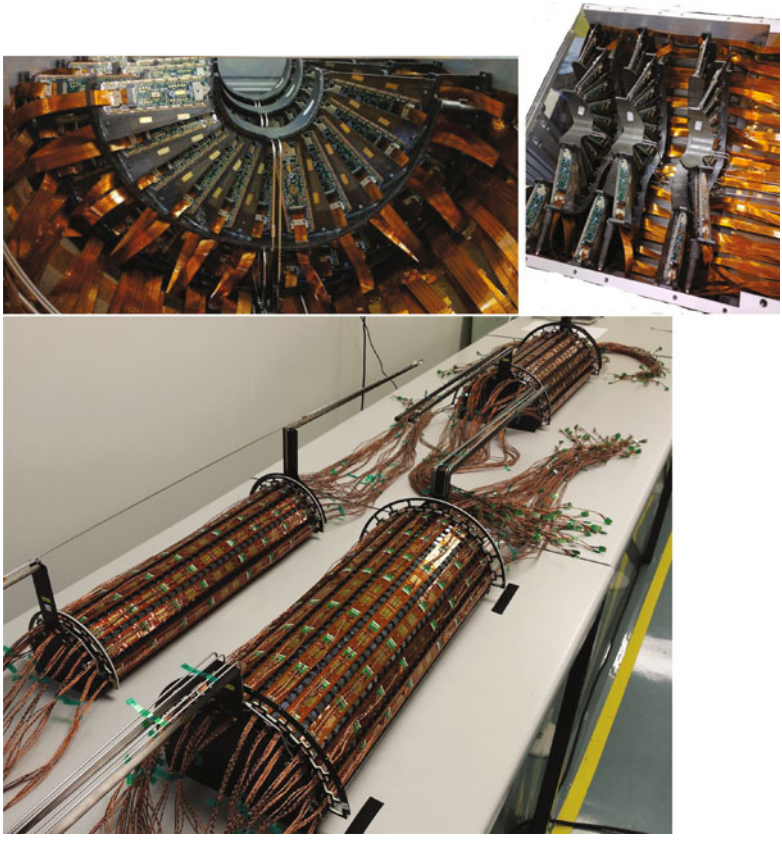


**Fig. 6.11** The *top left* photo shows the first assembled forward disks of the Phase I FPIX system. The turbine arrangement as in the first detector has been kept but the modules are of one single type; compare with Fig. 6.9. The *top right* photo displays three half-disks installed in a half-cylinder. In the *lower* part, you can see three out of four layers of the barrel detector during their assembly [Courtesy of CERN, Fermilab and PSI]

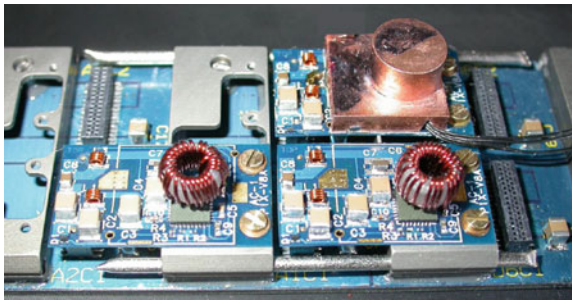


**Fig. 6.12** Photo of a DC-DC converter of the CMS phase I pixel detector [304]. The Figure displays the DC-DC with final shield (*upper right*) and with bare air coil (*lower part*) for illustration

The challenge is to send more than double the amount of data through the same number of optical fibres and to supply more power through the same number of power cables – it was not possible to lay more services.[6] The data throughput challenge is solved by changing from 40 MHz analogue transmission to 320 MHz digital transmission. The new chip features a fast 8-bit ADC retaining the ability to send pulse-height information to allow centre-of-charge calculation offline. As said earlier, the data and time stamp buffers have been increased to drastically decrease inefficiencies. DC-DC power converters are located in the supply tube. They allow higher input voltages, thereby increasing power without increasing electric current thus avoiding a higher cross-section of the cables [100]. The concept is that of buck converter[7] but dedicated radiation tolerant chips have been developed by CERN. A picture of the DC-DC converters (prototypes) with and without shielding the air coil is presented in Fig. 6.12.

The main ingredients to reduce mass, thus minimizing multiple scattering, are:

- move all service electronics, e.g. opto-hybrids out of the tracking volume into the supply tube
- use ultra low mass mechanics
- use bi-phase $CO_2$ cooling utilising smaller pipes with thinner pipe-walls

As result the total mass of the pixel barrel is halved despite placing about twice the amount of modules. The result of all mass reduction efforts can be seen in Fig. 6.13.

The $n$-in-$n$ DOFZ sensors, used already by the first CMS pixel detector, are radiation tolerant enough – the exact same concept has been re-used. Interestingly, the 4 in. masks for the barrel sensors are physically *the same* as for the first production and have simply been re-used at the company. The forward sensor cell and full sensor design did also not change conceptually but the production was done using 6 in. instead of 4 in. technology; fitting eight instead of three sensors onto one wafer. The decision to have a single sensor/module geometry simplified the system a lot.

[6] New cooling pipes, compatible with the high pressures of a $CO_2$ cooling system (triple vacuum jacketed steal pipes) have been installed during long shutdown 1; two years in advance of the pixel installation.

[7] A buck converter is a voltage step down converter. It is a switching converter storing the energy in an inductor. Due the CMS magnetic field only air coils can be used.

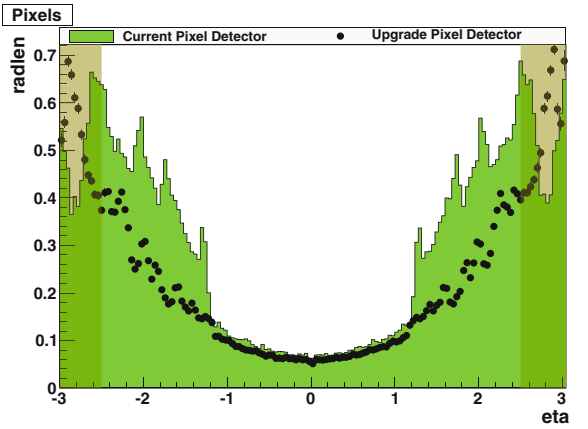



**Fig. 6.13** Radiation Lengths of the CMS phase I pixel detector. Clear reduction of mass in the relevant (*non-shaded*) region below $|\eta| = 2.5$ [304]

Section 6.3 gives more details about the sensors, where also Fig. 6.25 shows a photo of the real wafers scaled down by a factor of two.

## 6.3 The CMS Silicon Strip Tracker – SST

The CMS **S**ilicon **S**trip **T**racker SST spanning radii 20 cm $< r <$ 110 cm plus a substantial forward region is subdivided into **T**racker **I**nner **B**arrel TIB, **T**racker **I**nner **D**isk TID, **T**racker **O**uter **B**arrel TOB and **T**racker **E**ndcap TEC. The sub detector designs are different to match the requirements at the different radii—a bit like the manifold solutions as for DELPHI and CDF II. The different sub detector locations can be seen in Fig. 6.14, where a quarter section of the detector is presented.

*Tracker Inner Barrel TIB*

The **T**racker **I**nner **B**arrel TIB has four barrel layers assembled in shells, layers 1 and 2 are double-sided DS plus three inner Disks on both detector ends. The TIB is complemented by two disk-shaped forward detectors TID each composed of three small disks. The TIB/TID spans the radii 20 cm $< r <$ 55 cm with a length of 2.2 m($|z| <$ 110 cm). The TIB alone is 1.4 m long ($|z| <$ 65 cm). All modules are mounted on "strings" which are directly and individually placed on the structures. The reduced particle flux density allows the use of strip sensors with a typical cell size of 10 cm $\times$ 80 µm and 10 cm $\times$ 120 µm resulting in an occupancy of up to 1 – 3% per strip at LHC bunch crossing at nominal instantaneous luminosity. Modules are limited to one single sensor to accommodate for the maximum strip length. A photo of the inside of TIB can be found in Fig. 6.15 and a picture of a TID part in Fig. 6.16.

*Tracker Outer Barrel TOB*

The **T**racker **O**uter **B**arrel TOB structure where the modules are assembled in six concentric layers (TOB layers 1 and 2 are DS; globally Layer 5 and 6) surrounds

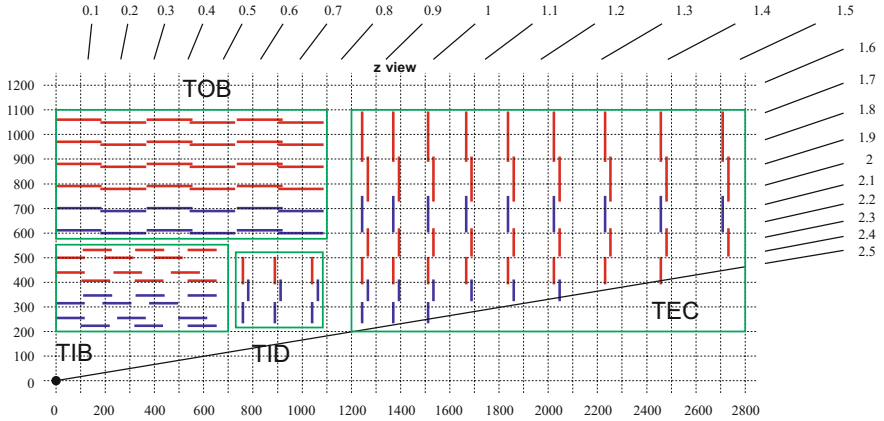



**Fig. 6.14** The CMS strip tracker layout. At intermediate radii (20 cm $<$ r $<$ 55 cm) the reduced particle flux allows the use of silicon microstrip detectors with a typical cell size of 10 cm $\times$ 80 $\mu$m, resulting in an occupancy of up to 1 – 3% per strip at LHC bunch crossing. In the outer region (55 cm $<$ r $<$ 110 cm) the strip pitch and length can be further increased to have cell size 20 cm $\times$ 180 $\mu$m. The detector is consequently instrumented with barrel geometry (*TIB* and *TOB*) including some stereo modules, complemented with the forward inner disks (TID) and the large endcaps on both sides (TEC). Barrel layer *1*, *2*, *5*, *6*, TID ring *1*, *2* and TEC ring *1*, *2* and *5* are instrumented with stereo double-sided modules

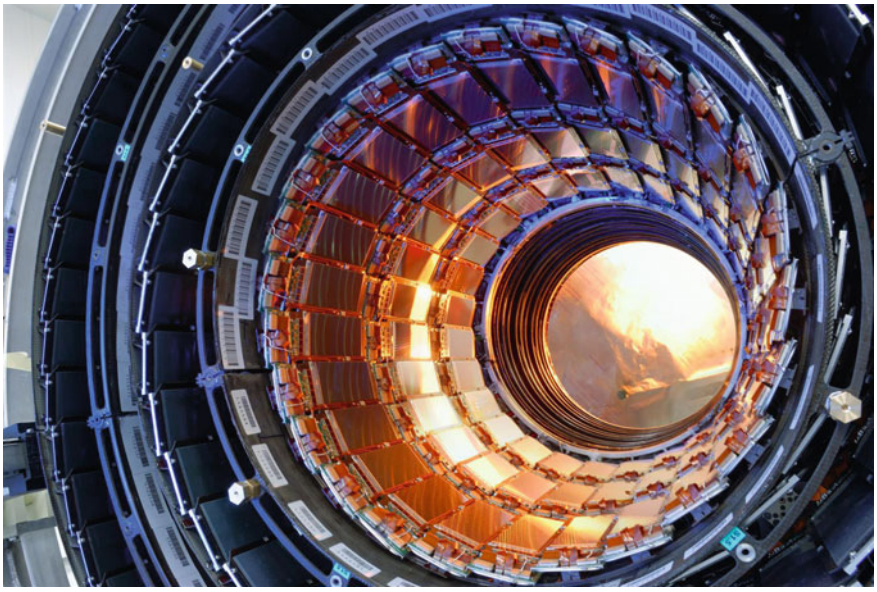


**Fig. 6.15** **T**racker **I**nner **B**arrel *TIB*. Three layers of tracker inner barrel detector are visible. In Fig. 6.18 it is inserted into the TOB. Finally, it will house the pixel detector. In the background the inner edges of the three **T**racker **I**nner **D**iscs can be seen [Courtesy of CERN]

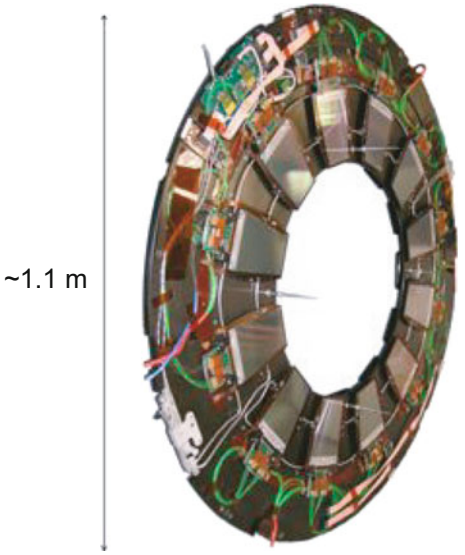



**Fig. 6.16** One disk of the **T**racker **I**nner **D**isk TID. Three of these disks on both TIB sides complete the inner detector fully embedded in the TOB [269]

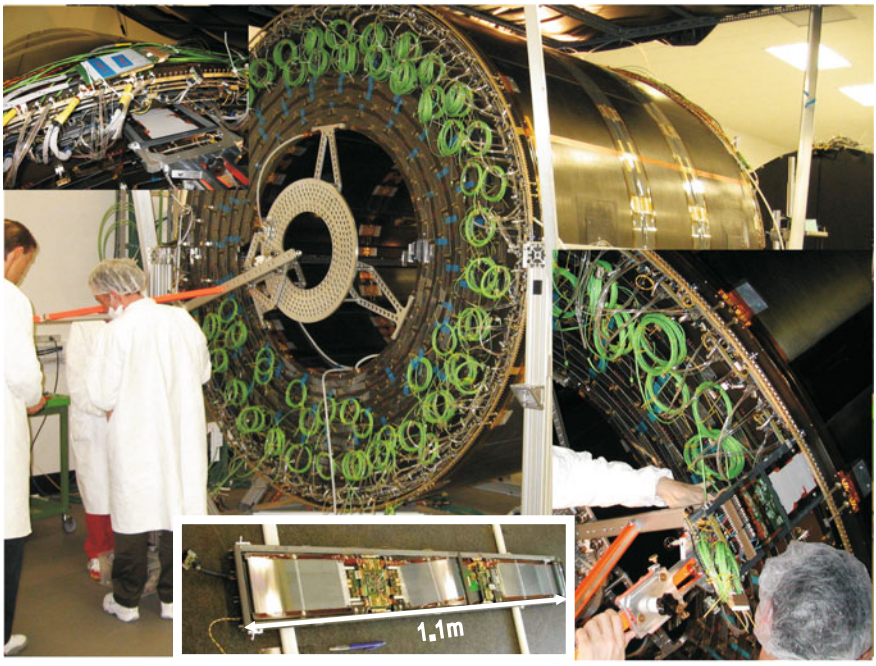



**Fig. 6.17** Rod insertion into TOB. The main picture gives a good overview of the size and complexity of the TOB. A special tool guarantees a safe insertion. A rod is displayed at the *bottom* [128]

**Fig. 6.18** TIB inserted in TOB, a rare view. The four layers of the tracker inner barrel are visible inside the tracker outer barrel. The TOB layers 5 – 10 are completely hidden by the necessary cabling. This is a rare view during the time of construction

the TIB. It is located inside the electromagnetic calorimeter. The TOB spans radii 55 cm $< r <$ 110 cm and a length of 2.2 m($|z| <$ 110 cm). For a detector of the size of the TOB an individual module placement is not practical anymore. A structure was designed where 688 one metre long "rods" equipped with modules are inserted. At these radii the pitches and strip lengths can be increased, which helps to moderate the number of readout channels, given the large area. The cell size is 20 cm $\times$ 122 $\mu$m and 20 cm $\times$ 183 $\mu$m with two sensors daisy-chained to one readout hybrid. The daisy-chaining has obvious advantages but also increases the channel capacitance load, and thus the noise. This is compensated by increasing the thickness from 320 $\mu$m ("thin") for TIB/TID to 500 $\mu$m ("thick") for TOB sensors. The same argument and design holds for TEC where inner rings 1 – 4 are "thin" and ring 5 – 7 sensors are "thick". An important detail is the smaller pitch in the two outermost layers reflecting the need to have a precision space point at a large lever arm (Fig. 5.3 on p. 199) and a precision link point to the calorimeters. For TIB and TOB, having a barrel-like structure, sensors are rectangular and strips run parallel to the beam $z$-axis. A double-sided module is composed out of two single-sided sensors mounted back-to-back where one is tilted by an angle of 100 mrad with respect to the $z$-axis. A photo of the TOB and a rod can be seen in Fig. 6.17. Figure 6.18 displays how the TIB is embedded in the TOB.

*Tracker End Cap TEC*

Two **t**racker **e**nd**c**aps TEC ensure a pseudo-rapidity coverage of $|\eta| \leq 2.5$. The strip orientation is radial, to achieve a continuity of all strips running radially through all rings from ring 1 to 7-sensors at different radii. Seven different types are necessary. All sensors are wedge type and for modules with two daisy-chained sensors, the outer pitches of e.g. sensor W5A must match inner strip pitches of sensor W5B for all individual strips, refer to Table 6.2. The detectors of ring 1, 2, 5 are made of double-sided modules. Modules of ring 1 – 4 contain a single sensor while ring 5 – 7 have a larger pitch and two sensors are daisy-chained together. As for the TOB the modules are arranged on substructures – on wedge-shaped "petals". Petals hold up to 7 rings of modules with the different geometries. One TEC consists of 9 disks populated with petals, each covering 1/16 of $2\pi$. There are eight different petal types varying in radial length, hence number of populated rings, to adapt for the location in $z$. Petal types differ for disks 1 – 3, 4 – 6, 7 and 8 and 9, the arrangement can be seen in Fig. 6.14. A total of 288 petals form both endcaps. The detector spans $120\,\text{cm} < |z| < 280\,\text{cm}$ and $20\,\text{cm} < r < 110\,\text{cm}$. An impressive photo of the TEC can be seen in Fig. 6.19.

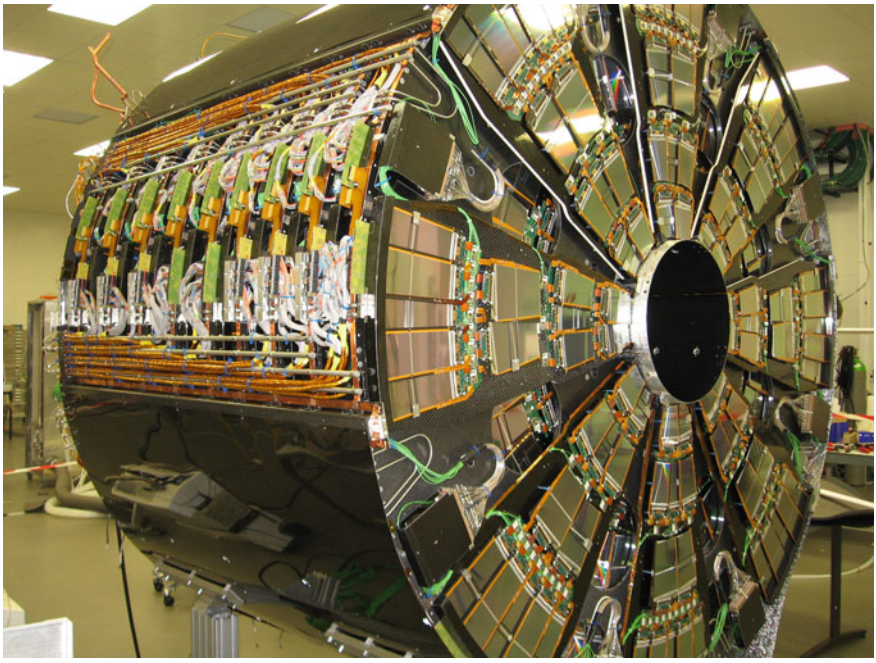


**Fig. 6.19** The **t**racker **e**nd**c**ap TEC, fully equipped and open, a rare view. Eight petals are visible on the front, the space in-between is covered with petals on the other non-visible side of the disk. Nine similar disks form one tracker endcap. The 96 visible wedge-shaped sensor modules placed in a wedge geometry on the petals guarantee a continuous-radial strip orientation. On the *left* an opening shows all the services for one sector, out of eight. The diameter of one TEC is 2.5 m while the length stretches to 1.7 m. One TEC holds 144 petals. All electrical and optical cables and cooling pipes were installed in autumn 2005. Both TECs were fully ready, with all petals installed and inserted into the tracker support tube early 2007 [Courtesy of CMS]

**Table 6.2** The different sensor types

| Sensor type | @ Layer | Trapezoidal a b h (cm) | Stereo module | Pitch (μm) | No. of strips | Thick (μm) | Resistivity (kΩcm) | Sensors module |
|---|---|---|---|---|---|---|---|---|
| IB1 | 1, 2 | 63.4 63.4 119.2 | YES | 80 | 768 | 320 | 1.5 – 3 | 1 |
| IB2 | 3, 4 | 63.4 63.4 119.2 | | 120 | 512 | 320 | 1.5 – 3 | 1 |
| OB1 | 9, 10 | 96.4 96.4 94.5 | | 122 | 768 | 500 | 3.5 – 7.5 | 2 |
| OB2 | 5, 6, 7, 8 | 96.4 96.4 94.5 | L5,L6 | 183 | 512 | 500 | 3.5 – 7.5 | 2 |
| W1TID | | 63.6 93.8 112.9 | | 80 – 119 | 768 | 320 | 1.5 – 3 | 1 |
| W1 | | 64.1 88.1 89.5 | YES | 81 – 112 | 768 | 320 | 1.5 – 3 | 1 |
| W2 | | 88.2 112.4 90.3 | YES | 113 – 143 | 768 | 320 | 1.5 – 3 | 1 |
| W3 | | 65.0 83.2 112.8 | | 124 – 158 | 512 | 320 | 1.5 – 3 | 1 |
| W4 | | 59.9 73.4 117.4 | | 113 – 139 | 512 | 320 | 1.5 – 3 | 1 |
| W5A | | 99.0 112.4 84.0 | YES | 126 – 143 | 768 | 500 | 3.5 – 7.5 | +W5B |
| W5B | | 112.4 123.0 66.1 | YES | 143 – 156 | 768 | 500 | 3.5 – 7.5 | +W5A |
| W6A | | 86.1 97.5 99.0 | | 163 – 185 | 512 | 500 | 3.5 – 7.5 | +W6B |
| W6B | | 97.5 107.6 87.8 | | 186 – 205 | 512 | 500 | 3.5 – 7.5 | +W6A |
| W7A | | 74.1 82.9 109.8 | | 140 – 156 | 512 | 500 | 3.5 – 7.5 | +W7B |
| W7B | | 82.9 90.0 98.8 | | 157 – 172 | 512 | 500 | 3.5 – 7.5 | +W7A |

The IB stands for **i**nner **b**arrel, the OB for **o**uter **b**arrel and the W for **w**edge, the sensors in the endcap. a, b defines the width and h the height of the sensors. Note that the OB1 sensors with the narrower pitch are located in the outermost layers to have a precise measurement with a long lever arm. All OB and W5–W7 sensors are 500 μm thick while all IB and W1–W4 are 320 μm thick

Detailed descriptions of the SST are also to be found in [128, 129, 292, 315]. Table 6.2 informs about all important sensor geometry parameters.

Appreciation of the full CMS detector size and the relation to the tracker can be gained by looking on Fig. 6.20 where the tracker is on the way to its final position in the heart of CMS – December 2007.

As a result from all necessary services, like cables, cooling pipes, support structures and electronic boards, the material budget is much higher than optimal and it is one of the issues which needs to be improved for any future detector. Figure 6.21 shows the material budget of the CMS tracker in units of radiation length. It increases from 0.3 $X_0$ at $\eta \approx 0$ to about 1.7 $X_0$ at $|\eta| \approx 1.4$, beyond which it falls to about 1 $X_0$ at $|\eta| \approx 2.5$. It is quite obvious that cables and support structures carry most of the weight, also the interface region between sub detectors, between barrel and forward region are prominently too massive.

Nevertheless, the designs described above lead to excellent performance underlined here with the track resolution plots for single traversing muon tracks in Fig. 6.22.

The issues of radiation and appropriate designs are discussed in the next sections. The front-end chip technology choice is presented in Sect. 6.4.1. The sensor design and strategies to achieve the necessary radiation tolerances are expressed in Sect. 6.4.2.

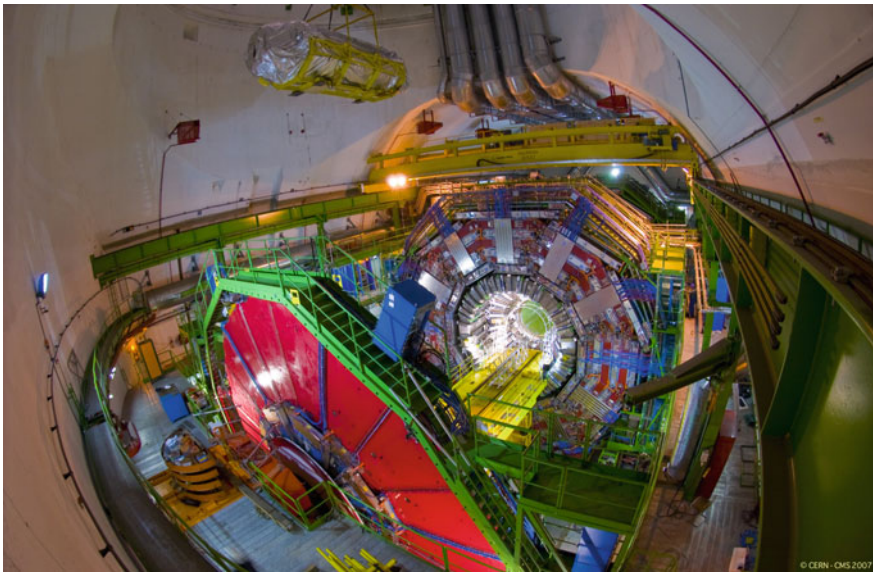



**Fig. 6.20** The CMS tracker on its way to the heart of CMS – 2000 man years on a hook. In December 2007 the CMS tracker finally descended down the 100 m deep shaft to be inserted into the heart of the CMS experiment to measure the particle tracks at between 15 and 22 positions in the detector with a precision of better than 20 μm [Courtesy of CERN]

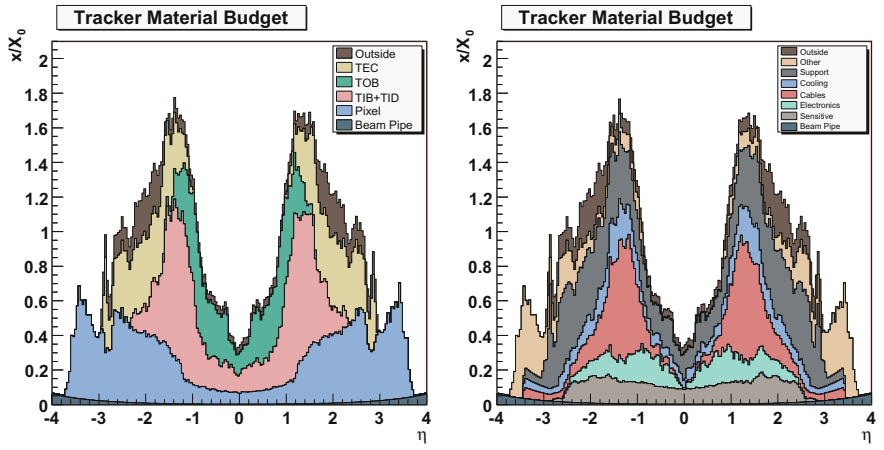



**Fig. 6.21** Material budget is displayed in units of radiation length as a function of pseudo-rapidity $\eta$ for the different sub detectors (*left*) and broken down into the functional contributions (*right*). [292]. A lot of design effort was spent to achieve a "light" tracker, Kapton cables were introduced directly on the flex hybrid with no extra connector; thin-walled cooling pipes were used; low mass power cables (aluminium) run inside the active CMS tracker volume, while outside low resistivity copper cables are used; structural elements are made of light weight carbon fibre and honeycomb structures. The beam pipe is made of beryllium. But despite all these efforts, at the interface of TIB and TOB to TEC all inner cables and pipes have to pass resulting in the two "too" high material peaks at both detector ends at $|\eta| = 1.4$

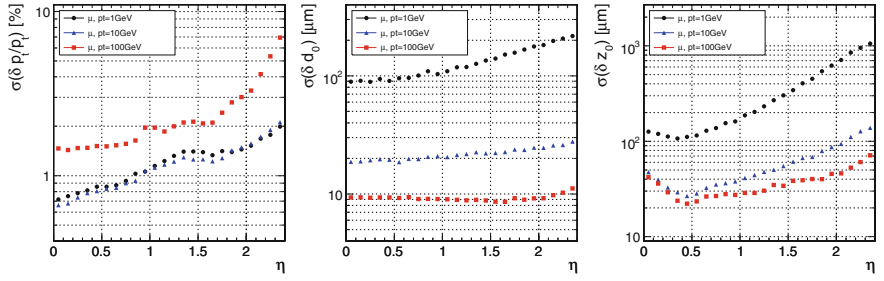



**Fig. 6.22** Resolution of several track parameters for single muons with transverse momenta of 1, 10 and 100 GeV: transverse momentum (*left*), transverse impact parameter (*middle*) and longitudinal impact parameter (*right*) – simulation [292]

Different from CDF II the CMS tracker is not participating in the hardware-level trigger but tracking information plays a major role in the **H**igh-**L**evel software **T**rigger HLT (computer farm). The Level-1 hardware trigger reduces the 40 MHz event rate to 100 kHz. The HLT further reduces the data rate to 100 Hz to be stored[8] – refer to Sect. 6.7 later.

[8] 100 Hz HLT rate is the original design, today 1 kHz rate is standard while several kHz can be stored for dedicated calibration studies, e.g. timing scans.

## 6.4 Design, How to Survive 10 Years in the Radiation Environment of LHC

The radiation environment at luminosities of $\mathcal{L} = 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ can be translated to integrated fluences[9] on the order of $\Phi_{eq} = 10^{13}\ n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$ in the outer strip tracker, $\Phi_{eq} = 1 \cdot 10^{14}\ n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$ in the inner strip tracker volume for the full operation period of 10 years. The pixel detectors expect to be subjected to even $\Phi_{eq} = 3 \cdot 10^{14}\ n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$ per year. These numbers are derived from simulations. Detailed numbers per volume cell including the real charge particle/neutron ratio are calculated. This radiation affects the gates in the ASIC transistors as well as the crystal lattice of the sensors. Early ASIC chips of the LEP or CDF chips would not have survived in the LHC environment. The basics of bulk and surface radiation-induced damages are discussed in Sect. 2. In the next sections the dedicated survival strategies of the CMS tracker are discussed.

### *6.4.1 Electronics – Quarter Micron Technology*

In the late 1990s radiation studies on chips proved that the thinner the gate oxides the more radiation tolerant the transistor, thus the electronic device. This was already seen for the different chip versions of LEP and CDF where the decrease in feature size exhibited an increased radiation tolerance with respect to ionising radiation levels. Within a "thin" oxide there is simply not enough lattice space for an effective defect concentration of holes to form and accumulate. In addition, the probability of electrons "tunnelling" into the oxide from the silicon bulk beneath is higher for a thinner oxide layer. These electrons can then recombine with the trapped holes near the interface. The detailed effect and technology exploitation is described in [155, 156]. The technology is also called "deep sub-micron" or "quarter micron[10]" and is much more radiation tolerant[11] than any former chip generation. Unfortunately the use of standard ASIC design libraries from the manufacturer is not sufficient to achieve the necessary radiation tolerance. The feature size defines the gate oxide thickness but there are always neighbouring oxides where surface damage can form and the resulting increasing surface leakage current again renders the transistors useless. The missing link is the full enclosure of all $n^+$ diffusions at different potentials including $N$-wells – fully enclosed NMOS transistors. A schematic is shown in Fig. 6.23. For CMS any larger feature size than 250 nm would not have been radiation

[9] Reminder: fluence numbers are always given in 1 MeV neutron equivalent/$\mathrm{cm}^2$ $[\Phi_{eq}] = n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$.

[10] At a feature size of 0.25 μm and below, chips start to be enough radiation tolerant for LHC operation.

[11] Another radiation-tolerant technology called DMILL existed but there were some production yield problems.

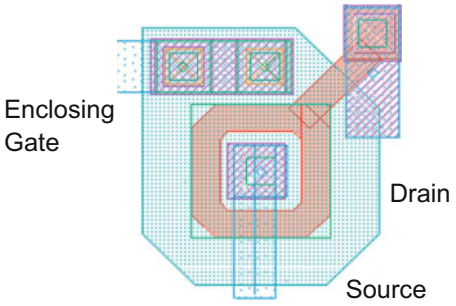



**Fig. 6.23** Radiation-tolerant transistor, 0.25 μm, enclosed layout. The gate encloses all $n^+$ regions completely avoiding any thick transistor-relevant oxide structures. This method prevents any radiation surface damage in parasitic thicker side oxides. The thin formed gate oxide is intrinsically radiation tolerant

tolerant enough. Hybrids equipped with CMS readout chip APV25[12] were irradiated to fluences of more than 30 years of LHC equivalent and exhibit no degradation.

Gigi Rolandi, the CMS Tracker Project Manager (2000 – 2006) said,

> The Quarter Micron technology arrived just in time, without it the realization of our plans would not have been possible!

Both, the pixel readout chip (ROC) and the strip detector chip (APV25) are produced in quarter micron technology.

In a high radiation environment there is also a chance of a state change in the chip microprocessor, the memory cells or the transistors caused by an ionising particle. These single event upsets have only minor effects in the front-end electronics used LHC experiments and are beyond the scope of this book. They were investigated thoroughly and are considered "soft", while their effects are of temporary nature. They do not permanently degrade the functionality of chips.

### 6.4.2 Silicon Sensors

The pixel detector, being subjected to fluences up to $\Phi_{eq} = 3 \cdot 10^{14}\, n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$ per year[13] at nominal LHC luminosity, utilizes different sensor designs and technologies than the strip detector, facing "only" a fluence of $\Phi_{eq} = 1 \cdot 10^{14}\, n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$ for the full 10 years of operation. In reality, the LHC is exceeding design luminosity; still the margins of the Tracker system and its sensors, especially with the Pixel Phase I upgrade (2017), are comfortable.

[12] 0.25 μm version of the APV chip – final version.

[13] With an initially lower luminosity an integral fluence of $\Phi_{eq} = 6 \cdot 10^{14}\, n_{1\,\mathrm{MeV}}/\mathrm{cm}^2$ was expected after 4 – 5 years of initial running.

**Pixel Sensors**

The pixel detector's sensor and chip design is described in detail in [87, 292]. The sensor technology did basically not change from Phase 0 to Phase I, only the sensor geometry did for the forward part. The CMS pixel sensors are designed and processed in $n$-in-$n$ technology on high resistivity oxygenated Float-Zone silicon (Diffused Oxygenated FZ – DOFZ), the most radiation hard concept available at the time (pre-LHC), and as proposed at the time by RD50. With the $n$-in-$n$ concept ($n^+$ pixelated implants on $n$-bulk) the depletion volume after type inversion develops from the segmented pixel face thus allowing operation at bias voltages lower than full depletion voltage (under-depleted operation[14]) and still sufficient charge (electron) collection. The high oxygen concentration also reduces acceptor creation during irradiation thereby reducing final depletion voltage.

With an initial resistivity of about 3.7 k$\Omega$ and a thickness of d = 285 $\mu$m full depletion voltages are about 50 – 60 V. Guard rings on the edge of the sensors, being kept on relative ground, guarantee a stable running up to bias voltages of 600 V without breakdown. A perfect homogeneity for the barrel part has been achieved by having all wafers out of one single ingot. The $n$-in-$n$ concept requires a double-sided process where the multi-guard ring structure are situated on the backside of the sensor, cascading the high negative voltage down to zero towards the side of the sensor. Historically, the barrel and forward pixel consortia acquired sensors from two different vendors with different design rules. Quality and performance are equally excellent but pixel cell isolation is solved with $p^+$-spray technique and $p^+$-stops for barrel [168] and forward [44] sensors, respectively. Photographs of the different pixel cells are shown in Fig. 6.24. Cells are DC-coupled; $n$-implants are connected by "vias" to the metal bump bonding pads and then connected with flip-chip technology to the chip cells (see Fig. 6.6 and Fig. 1.73). Both types are using punch-through as biasing techniques allowing IV characteristics at testing stage.[15] In the forward pixels, a simple opening in the $p^+$-stop presents a low resistance path until full depletion is reached changing to a high resistance path at over-depletion (volume below pixel cells depleted) to assure pixel isolation. The barrel pixel sensors, using moderated $p^+$-spray, utilize a dedicated punch-through structure with a well-defined bias dot for every pixel (small distance to GND potential and maximal area).

For Phase 0, two different sensor geometries have been realized for the barrel part; the standard rectangle corresponding to $2 \times 8$ read-chips and a smaller number of half-modules only $1 \times 8$ chips worth and necessary to instrument the connecting part of the two half-shells. The forward part is more complicated, where 7 different sensor geometries $1 \times 2$ up to $2 \times 5$ chips have been implemented to achieve a wedge geometry. Sensors plus chips sandwiches are called plaquettes; several plaquettes (3 or 4) form a module called a panel; two panels form a blade. The inner panel

[14] In $n$-in-$n$ under-depletion mode, the non-depleted zone stays on the backside only reducing the active volume but with depletion zone at the segmented pixel face.

[15] GND potential is finally connected via the individual chip cells but for sensor testing purposes a bias grid has been implemented.

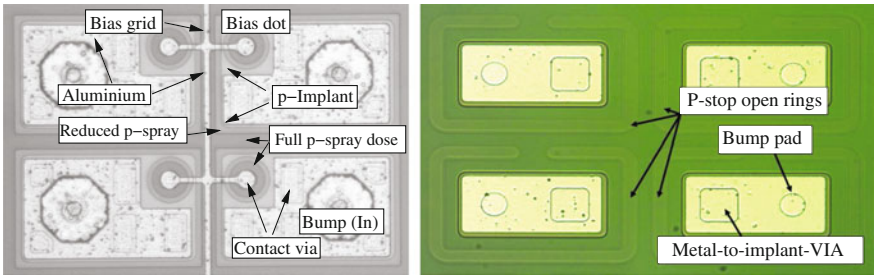



**Fig. 6.24** Barrel and forward pixel sensor cells. The differences of pixel isolation and biasing methods are shown. Indium bumps on the barrel one are already deposited but not yet reflown [87, 292]

features $1 \times 2$, $2 \times 3$, $2 \times 4$, and $1 \times 5$ type plaquettes, the outer one $2 \times 3$, $2 \times 4$ and $2 \times 5$. The outer panel can be checked in Fig. 6.4. The forward system has a total of 672 plaquettes on 96 blades. All sensors have been processed on 4 inch wafers; clearly with several different geometries on one wafer.

For Phase I, all sensors are rectangular, also for the forward disks. Figure 6.25 shows a photo of the real wafers of the Phase I detector scaled down by a factor of two. The barrel pixel sensor cells are exactly the same as shown in Fig. 6.24. Interestingly, the 4 inch masks for the barrel sensors are physically *the same* as for the first production and have, several years later, simply been re-used at the company. The forward pixel cells underwent a small optimization from Phase 0 to Phase I.

With the initial PSI46 analogue chip, test beam runs [258] with un-irradiated and irradiated detectors have shown that the barrel sensors still have 99.0% efficiency for tracking charged particles after an exposure of $\Phi_{eq} = 3 \cdot 10^{14}\ n_{1\mathrm{MeV}}/\mathrm{cm}^2$ and the forward sensors maintained an efficiency of 98.8% after $\Phi_{eq} = 8 \cdot 10^{14}\ n_{1\mathrm{MeV}}/\mathrm{cm}^2$. With the Phase I, PSIDIG and PROC600 digital chips with much lower threshold capability, the system becomes much more radiation tolerant and will be able to operate for the full until Long Shutdown 3 foreseen integrated luminosity of $\mathcal{L} = 500\ fb^{-1}$ corresponding to $\Phi_{eq} = 3 \cdot 10^{15}\ n_{1\mathrm{MeV}}/\mathrm{cm}^2$ accumulated in the innermost pixel layer at $r = 3$ cm. Barrel Layer 2 gets four times less. The system, as extra precaution, allows to extract the pixel detector and exchange the innermost layer in a normal Year-End-Technical-Stop. The strategy and mechanical structure allows, though, to replace layer 1 easily after about half the expected fluence.

### Strip Sensors

All silicon strip sensors are single-sided $p$-in-$n$ with AC-coupled readout and $p^+$ strips biased through polysilicon resistors, based on the planar process. Pitches range from 80 to 183 μm (up to 205 μm in the end of a wedge-shape sensor) without any intermediate strips. The substrate is non-oxygenated Float-Zone $n$-type silicon made from 6 in. diameter wafers. All barrel sensors are consequently rectangular with strips running parallel to the beam axis, a stereo angle for layers 1, 2, 5 and 6 is achieved by rotating the complete sensor inside the module by 100 mrad. Double-sided modules are composed of a normal and a stereo module back-to-back.

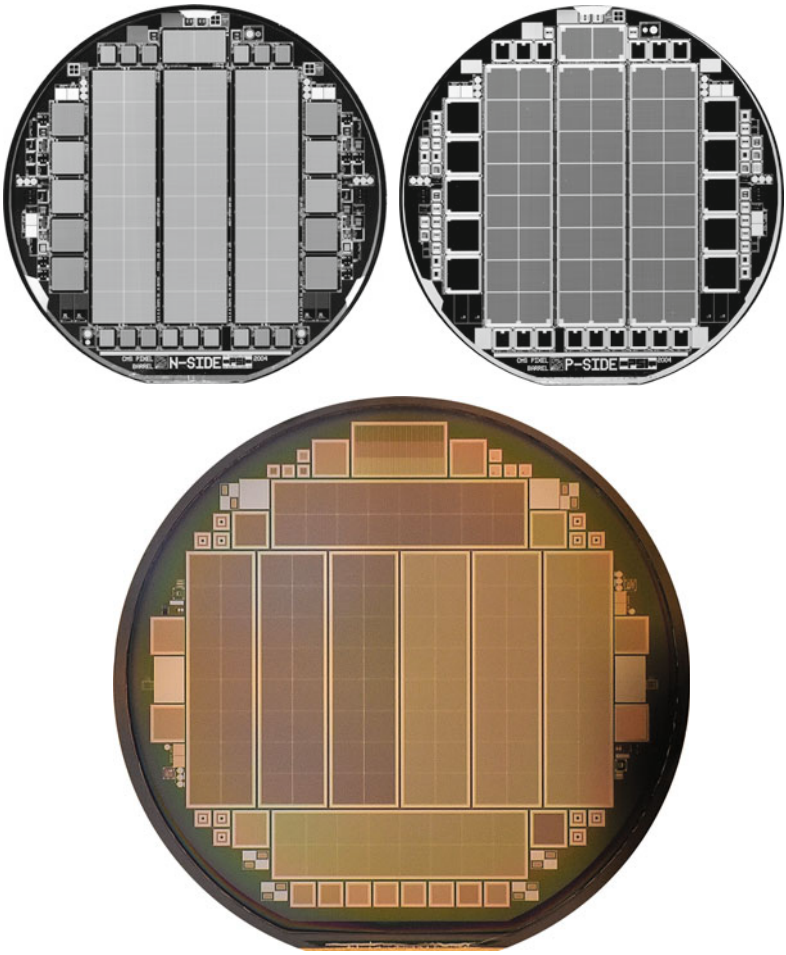



**Fig. 6.25** The photos show the wafers of the pixel sensors. All pictures are scaled down by a factor of two. The *first* shows the $n$-side (pixel cell) side of the barrel sensors. $n$-in-$n$ sensors need double-sided processing featuring guard rings on the backside and full backplane metallisation – see second (*right*) picture. The *last* photo shows the $n$-side of a forward pixel sensor wafer processed on 6 in. The resulting individual sensors have the same size but eight instead of three sensors fit. Photos of the individual pixel cells can be seen in Fig. 6.24 [Courtesy of Paul Scherrer Institute and Fermilab]

All sensors in the forward part are trapezoidal with strips running in the radial detector direction. As discussed earlier outer modules contain two daisy-chained sensors. In order to maintain a good signal-to-noise ratio of well above 10 (after irradiation) CMS uses thicker silicon sensors for the outer tracker region (500 μm thickness as opposed to the 320 μm in the inner tracker) with correspondingly higher signal. These thicker sensors would in principle have a higher depletion voltage, since $V_{FD} \propto D^2$. But since the radiation levels in the outer tracker are smaller, a higher

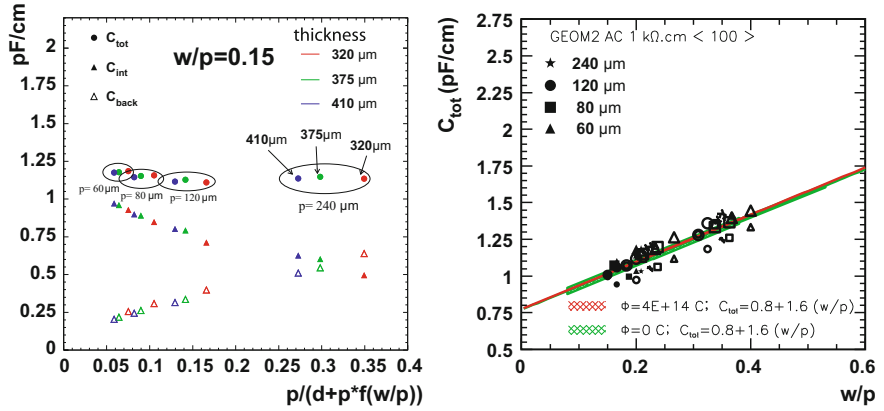



**Fig. 6.26** Varying width and pitch of the microstrip sensor, CMS finds a constant total capacitance for a constant width/pitch (w/p) ratio. The *left* plot shows the constancy of $C_{tot}$ together with its two components $C_{int}$ and $C_{back}$ for a constant w/p = 0.15, capacitances are plotted versus formula (6.2). The *right* plot shows the linear increase of $C_{tot}$ with increasing w/p. The test holds true for several sensor thicknesses and for ⟨111⟩ as well as ⟨100⟩, but only ⟨100⟩ shows also robustness versus radiation, refer also to Fig. 6.30 and [34, 47, 49, 77]

initial resistivity can be chosen such that the initial depletion voltages of thick and thin sensors are in the same range of 100–300 V, since $V_{FD} \propto N_{eff}$. The different sensor parameters, electrical as well as geometrical, are presented in Table 6.2. The load capacitance is the main source of noise for fast signal shaping, therefore detailed studies were conducted to establish a correlation between pitch p, strip width w, sensor thickness $D$ and load capacitance composed of inter-strip and backplane capacitance. Multi-geometry sensors were processed in the R&D framework of CMS together with the vendor where several combinations of w/p were tested on ⟨111⟩ silicon crystal orientation as well as ⟨100⟩ silicon, as well as for different resistivities. The result proved that for a constant w/p the total capacitance $C_{tot}$ stays constant, inter-strip capacitance decreases in the same order as backplane capacitance increases (see Fig. 6.26). Furthermore, to achieve robust coupling capacitors a multi-layer dielectric was used. A good process stability of polysilicon resistors was achieved by saturating by heavily doping the sheet resistance. Further information about the CMS sensor design can be found in [34, 47, 49, 77]. Formulas (6.1) – (6.3) parameterize the results

$$C_{int} = \left[0.03 + 1.62\frac{\mathrm{w} + 20\,\mu\mathrm{m}}{\mathrm{p}}\right]\frac{\mathrm{pF}}{\mathrm{cm}} \tag{6.1}$$

$$C_{back} = \varepsilon_0\varepsilon_{\mathrm{Si}}\frac{p}{D}\frac{1}{1 + \frac{p}{D}f\left(\frac{\mathrm{w}}{\mathrm{p}}\right)} \tag{6.2}$$

$$C_{tot} = \left(0.83 + 1.67\frac{\mathrm{w}}{\mathrm{p}}\right)\frac{\mathrm{pF}}{\mathrm{cm}} \tag{6.3}$$

A constant w/p = 0.25 is used for all CMS sensors resulting in a load capacitance of 1.2 pF/cm strip length.

Within the high radiation environment of LHC all sensors will undergo type inversion, as discussed in Sect. 2. The main strategies of CMS to ensure radiation hardness of silicon sensors consist of

1. delaying the bulk type inversion and achieve $V_{FD}$ always below 400 – 500 V
2. the use of stable sensors with respect to high voltage and
3. the reduction of surface damage

*Delaying Type Inversion – Guarantee* $V_{FD} < 400 - 500\ V$

A tailoring of sensors and of the operation plus maintenance scenario is possible to withstand a substantial amount of radiation by taking into account all theory (see also Sect. 2) of radiation damage. Especially important is the evolution of full depletion voltage versus fluence, time and temperature. The three main strategies are the following:

- freeze out all reverse annealing by maintaining all sensors at sub-zero temperatures during operation and as much as possible during maintenance
- exploit beneficial annealing with a controlled warm up during the maintenance periods
- start with relatively low resistivity silicon to reach the inversion time late and end with about the same depletion voltage as the initial one

The employment of low resistivity silicon and therefore high initial $V_{FD}$ delays the type inversion point in time resulting in an operational depletion voltage after 10 years of LHC operation. Figure 6.27 (left part) shows the simulated evolution of $V_{FD}$ with time including all operation and maintenance periods, hence radiation and annealing periods. The different time constants for beneficial and reverse annealing have to be exploited. During operation at sub-zero temperature all defect annealing processes are practically suppressed. Figure 6.50 on p. 278 shows the individual sensor temperatures in situ compared to the original design value $T_{silicon} = -10\,^{\circ}\mathrm{C}$; there is margin to reduce temperatures by another 10 degrees. Operating at low temperature is also necessary to limit leakage current and avoid thermal runaway. During maintenance periods the tracker will be "warmed up" to benefit from the beneficial annealing component. Figure 6.27 (right part) shows the resulting depletion voltages depending on the "warm up" temperatures and it clearly shows that these periods are needed, but with meticulous temperature control. During the whole production period, sensors and test structures have been irradiated and results were compared to the Hamburg model – Fig. 6.28.

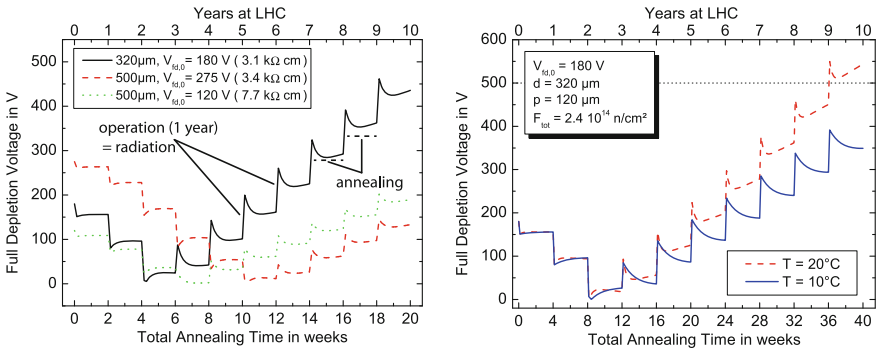



**Fig. 6.27** The CMS sensors are subjected to a high radiation field. To freeze the radiation damage and to decrease leakage current the operation temperature is −15 °C. During shutdown and maintenance periods the tracker will be "warmed up" for required interventions or to make controlled use of the beneficial annealing effect (see Sect. 2.1) [79]

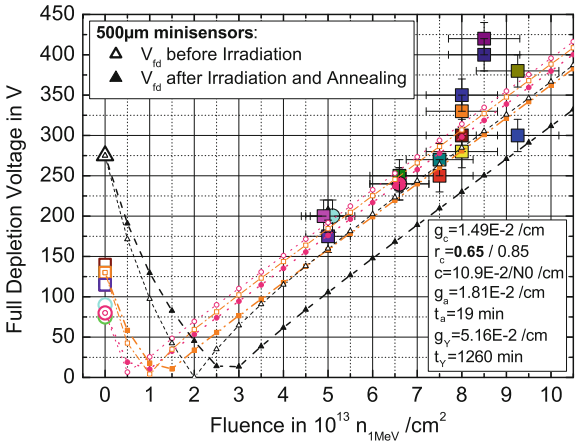



**Fig. 6.28** Depletion voltages of several CMS sensors for several fluences. These measurements were used to determine the Hamburg model (see Sect. 2.1) parameters for the CMS sensors and to evaluate operation parameters during the 10 years running period [79]

*High Voltage Robustness*

As a second measure, the sensors have to be extremely robust with respect to operation voltage. For several regions, bias voltages of around 500 V will be necessary after 10 years of LHC operation. The basic strategy is to shift high fields into the isolation oxide and consequently avoid high local fields:

- all implants are deep and round shaped
- no sharp corner on any metal structure, only rounded shapes
- active $n^{++}$ edge to have a defined homogenous field at the outer region ($p^{++}$ for $n$-in-$p$)
- no chips/breaks larger than ±20 µm at the cut edges, not penetrating the active edge
- metal overhang to distribute field to the $SiO_2$–metal region for strips, guard and bias ring

- implementation of a floating guard ring with a metal overhang to continuously smooth the high potential (HV) from the edges to GND applied on the inner bias ring. This is an improved version of a multi-guard ring which levels the potential discretely
- after type inversion, high field region is located at the sensor backplane

The use of round structures and avoidance of breaks and chips together with an active edge[16] are more or less standard methods to avoid any high local field. The implementation of a metal overhang is a CMS innovation and improves voltage robustness significantly. High fields[17] at the $p^+$ implants were shifted towards the metal strips into the $SiO_2$ insulation where the breakdown voltage is 20 times higher ($V_{break}(Si) = 30\ \mathrm{V/\mu m};\ V_{break}(SiO_2) = 600\ \mathrm{V/\mu m}$). In addition, it was possible to reduce a multi-guard ring to a single one with an overhang operating with a continuous potential shift of the high voltage at the sensor edge to GND at the bias ring. Many layout simulations and final testing of the implementation were necessary. The metal overhang ranges between 4 and 8 μm. At best, the aluminium electrode should even have a slightly more negative potential than the $p^+$ implant. All sensors were tested during standard quality control up to 550 V, many up to 800 V and some sensors were tested even up to 1500 V without breakdown.

*Surface Damage Reduction*

As a third measure CMS uses silicon with ⟨100⟩ crystal orientation, with less dangling bonds than standard ⟨111⟩ silicon (Fig. 6.29). Formula (6.4) lists the surface density of dangling bonds.

$$\text{Surface density} \begin{matrix} 11.8 \\ 9.6 \times 10^{11}\ \frac{\text{atoms}}{\text{cm}^2} \\ 6.8 \end{matrix} \begin{matrix} \langle 111 \rangle \\ \langle 110 \rangle \\ \langle 100 \rangle \end{matrix} \tag{6.4}$$

The use of ⟨100⟩ with the lowest number of dangling bonds naturally leads to a suppression of surface damage resulting in reduced increase of inter-strip capacitance after irradiation. Figure 6.30 illustrates the capacitance differences of radiated ⟨111⟩ and ⟨100⟩ sensors. In addition the use of ⟨100⟩ also decreases the undesired oxide charges therefore decreasing micro-discharge probability (Sect. 1.6.3).

The interested reader will find more exhaustive information about the CMS sensors and their quality control program in [13, 129, 183]. The process control, a vital part of the quality assurance, is described in detail in [35, 197].

[16] Was already implemented in CDF II, refer to Sects. 1.6.2 and 5.3.

[17] Main field from the high bias voltage applied to reverse bias the *pn*-junction.

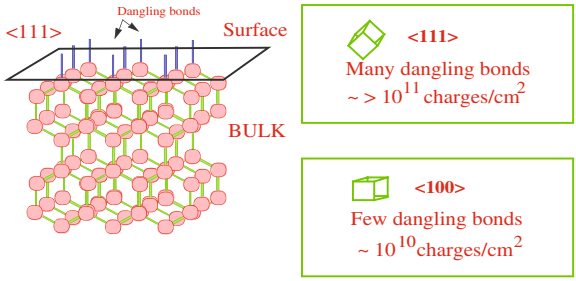



**Fig. 6.29** CMS uses ⟨100⟩ instead of ⟨111⟩ crystal orientation, exploiting the smaller number of dangling bonds at the Si – $SiO_2$ interface allowing less charge accumulation due to irradiation

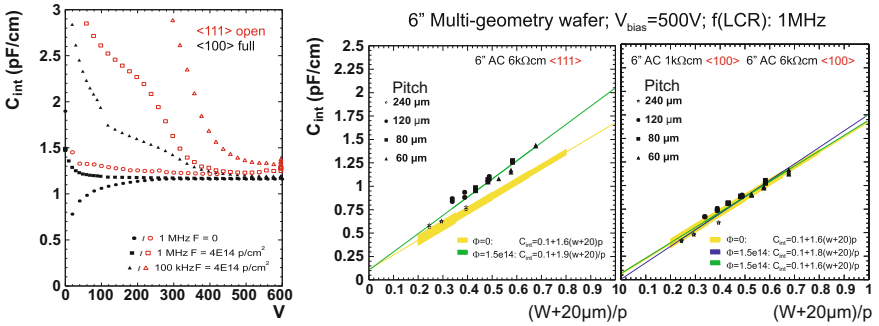



**Fig. 6.30** In the *left part*, the full bullets show the beneficial effect of ⟨100⟩ crystal orientation. Differences at over-depletion are better seen in the *right part*. The study is consistent for several different pitches p and strip widths w. A clear beneficial effect of the ⟨100⟩ orientation is shown [34, 47, 49, 77]

The following items summarize the CMS tracker strategies to guarantee a safe 10 year operation:

- relatively "low" silicon resistivity is used resulting in high initial depletion voltage shifting type inversion to a later time ending up with voltages that are not too high
- operation at low temperatures limits leakage currents, avoids thermal runaway and freezes reverse annealing
  - to freeze the reverse annealing, the detector must remain cold, also in the maintenance periods
- the beneficial annealing effect is exploited by a controlled temperature rise during shutdown periods
- voltage robustness of the sensors is improved to allow for high voltage running especially with the use of a metal overhang design
- ⟨100⟩ silicon was chosen to minimize Si – $SiO_2$ interface deterioration

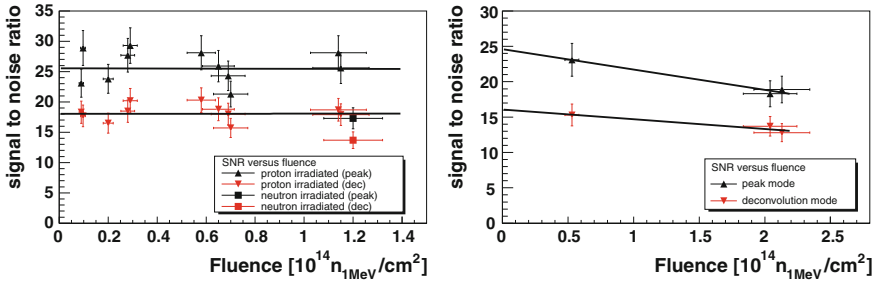



**Fig. 6.31** Signal/noise of full CMS modules after neutron and proton irradiation (*right plot*: protons only). The values far exceed the expectation and technical design specification [351]. A substantial number of full modules were irradiated to several fluences to guarantee full functionality for the 10 years of operation with a sufficient signal/noise value

Figure 6.31 depicts the estimated functionality over the complete operation period of the CMS tracker with higher than marginal signal/noise ratio. All sensors are operational at 400 V and above, a clear sign of a successful design. A number of modules were thoroughly tested for beam incident robustness which they passed successfully, refer to [98, 99, 130].

*The Thermal Screen*

To guarantee the sub-zero temperature described above during the whole operation and maintenance period of 10 years, an active thermal screen was designed and built. This is an innovative and unprecedented feature of a tracking detector. The thermal screen is an "intelligent" insulation, responsible to keep the inside below $-10\,^{\circ}\mathrm{C}$ and the outside at $18 \pm 1\,^{\circ}\mathrm{C}$. This is achieved with 32 cold panels inside a Rohacell insulation layer and several dedicated heating foils outside. Two dedicated and redundant cooling stations guarantee the cold operation, while the heating foils are controlled by a dedicated system via PID[18] feedback loops. This engineering masterpiece is one of the keys for the silicon sensor to survive in the harsh environment. The thermal screen is located inside the tracker support tube surrounding the TOB and TEC sub detectors.

## 6.5 Construction Issues for Large Detector Systems with Industry Involvement

Describing all the logistics, quality assurance steps, technology choices, failure types, etc. of the CMS tracker could fill a book by itself and is the subject of a larger number of publications and conference presentations, e.g. [12, 72, 128, 130, 269]. The challenge to equip 206 $\mathrm{m}^2$ with silicon detectors up to radii of $R = 110$ cm is unprecedented. The achievement was made possible by adopting and developing new production and quality assurance methods and by adapting industrial techniques.

[18] **P**roportional plus **I**ntegral plus **D**erivative PID control, an industry control standard.

**Table 6.3** Number of individual components built into the CMS strip tracker

| | |
|---|---|
| Thick sensors | 18,192 |
| Thin sensors | 6052 |
| Front-end hybrids/full modules | 15,148 |
| FE ASICs | 75,376 |
| Electronic channels | 93,16,352 |
| Petals | 288 |
| Rods | 688 |
| Analogue optical links for 65 m transm. | >40,000 |
| **F**ront-**E**nd **D**rivers (ADC boards) | 440 |
| Peak module assembly per day | 90 |
| Optical fibres | 3000 km |
| Voltage channels (LV+HV) | 8000 |
| Temperature and humidity sensors | 1000 |
| Slow control computer | 10 |

To stress the fact even a bit more, a petal consists of about 400 sub parts each, a module has about 10 items with about 1000 – 3000 wire-bond connections. A hybrid has 4 – 6 chips in combination with three additional ASICs and about 40 passive components

These procedures were novel in the field of particle physics detectors. A large collaboration of 51 institutes[19] with almost 500 physicists and engineers succeeded over a period of 12 – 15 years to design, develop and build this unique device.

Table 6.3 presents an example of the numbers of components and Fig. 6.32 visualizes the different locations, distributed all over the world, of the different assembly levels. In-between every shipment there are defined acceptance protocols, e.g. visual inspection and/or electrical testing. Also in-between each assembly step a quality assurance test is mandatory. A consequence of the increase in detector size is a change in the personal attitude towards the individual sensors and modules. With the number of sensors exceeding several thousand elements these have to be tested, handled and used in a nearly industrial manner. The time is over when each individual silicon sensor received special attention and was "brought to life" in many hours of careful handling. The silicon detector community had to develop standard procedures to systematically test sensors, efficiently perform the module construction and to use them in large detector systems.

### *6.5.1 Quality Assurance and Problems During the Process*

A detector of this size and complexity calls for unprecedented meticulous quality control to identify and solve problems as early as possible in the production. Some

[19] Today 2017, there are 69 Tracker Institutes.

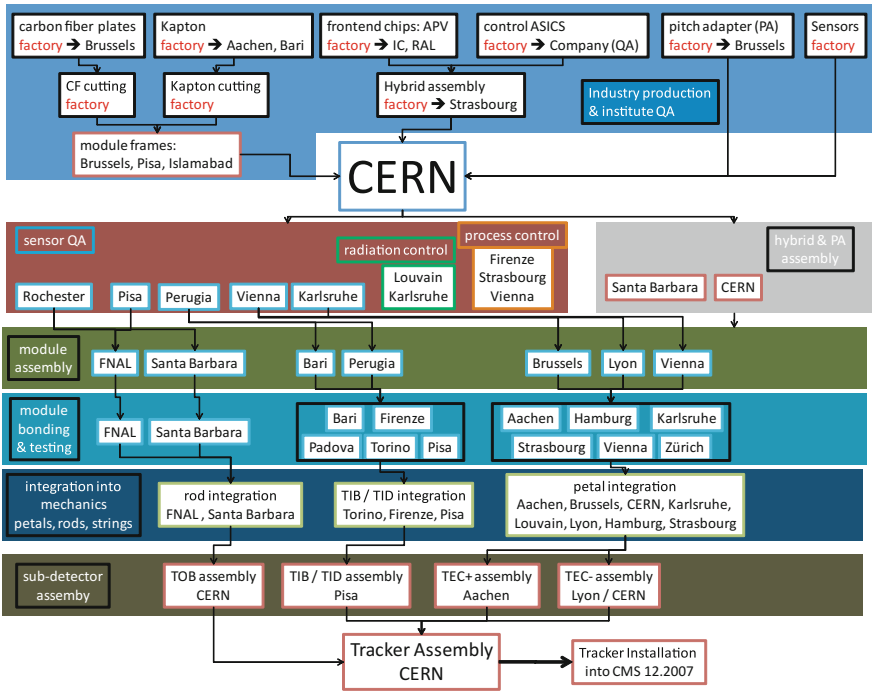



**Fig. 6.32** The CMS tracker logistics

examples are listed. Basic **Q**uality **A**ssurance QA strategies for silicon sensors are described in Sect. 1.7.1 including the steps introduced for CMS. In addition to the standard sensor quality control, CMS introduced a special "process monitoring" necessary for the long procurement time of 3 years for the sensors. Due to the stringent radiation requirements a fraction of sensors and test structures were irradiated to 1.5 times the nominal fluence levels. The result is an overall percentage of good strips in the accepted sensors of 99.94%. The early qualification steps are described in [13, 129]. Along the way all kinds of faults were encountered and all were subsequently solved, mostly in close collaboration with the vendor. In an early stage of delivery coupling capacitors consisted only of $SiO_2$. This insulation layer technique proved to be fragile and several pinholes developed during initial quality control. The issue was solved with the introduction of additional layers of nitride [183]. The stringent requirement of cut edge breaks/chips below 20 μm required several visits to one company to sensitize them to the problem. The thin sensor handling plus requested measurements "after" cutting was not a common standard for the manufacturer. Also scratches occurring on certain areas of the sensor was a frequent fault. On a wafer with 100 ASICs a scratch decreases the yield by a percentage but with a single sensor per wafer the yield drops to zero. Testing personnel were finally specially trained for CMS requirements, see also. Section 1.14 and [60, 130, 183]. In Sect. 1.14 a short

description of the low inter-strip resistance problem is presented. The problem was identified due to vigilance and the stringent process monitoring established, because inter-strip resistance was not a standard QA item of the manufacturer.

For a period of about 2 months a high flat-band voltage was experienced due to oxide contamination, oxide traps and/or interface states for a large quantity of ∼1000 sensors. An immediate irradiation campaign showed a resulting increase of inter-strip capacitance with radiation different to the non-contaminated ⟨100⟩ sensors [197].

Very prominent was the finding of corrosion of the aluminium bias and guard rings, described in Sect. 1.14 and elaborated in detail in [139].

Even after sensor QA individual channels exhibited an influence on common mode noise due to micro-discharges, observed in the module long-term testing. A single strip, inconspicuous during sensor QA, was able to significantly increase the noise of a full chip [60]. An additional voltage[20] had to be applied to the AC pad, while current was measured on the DC pad to mimic the operation condition in a module. In this configuration a strip vulnerable to micro-discharge exhibited a significantly higher current.

Glass pitch adapter series were delivered with varying metal line quality and bonding was not always possible. The ability of sensors and pitch adapters to be bonded was constantly monitored during the whole production. The final bonding failure rate stayed below 0.01%.

Another serious problem affected a series of hybrids. Vias were unstable under thermal stress. A cleaning step during processing was also etching away part of the glue between the Kapton layers. The metal conductor path was therefore stretching too far into the grooves and was breaking under thermal expansion. As a solution, an additional Kapton layer was introduced and the via size was increased, see also [85, 128, 181].

A new problem relevant only for sensors in a very high-radiation environment was detected during module production. The backside module contact was made with conductive glue between aluminium sensor backplane and metallized Kapton. Varying and ageing conductivity of the conductive glue was found between several Ω and several kΩ. Oxygen was diffused through the glue oxidating the aluminium surface and thus degrading conductivity. For a sensor with bulk resistance in the order of several GΩ, an additional series resistance of some kΩ is of no consequence at all. Since the bulk resistance drops significantly with radiation down to several kΩ, the voltage drop on the glue would be in the same order as for the sensor bulk and depletion could no longer be guaranteed [85, 128, 181]. All modules were therefore retrofitted with backplane bonding.

The prior mentioned active $n^{++}$ edge introduces the feature that the edge is at a high potential and therefore wire-bonding with a low loop height resulted in a voltage spark. Module quality control detected this problem and bonding loop parameters had to be adapted accordingly.

[20] During operation together with the hybrid, the potential on the chip connections are +0.85 V higher than the sensor ground.

All components, hybrids, modules, rods, petals and sub detectors, were long-term tested and all went through several thermal cycles. The level of quality assurance was new for the HEP community but ultimately with zero fault tolerance and rigid vigilance, the final strip inefficiency remains below 0.5%. This includes all sensor, chip, hybrid and module faults and all missing connections from readout and power towards the tracker.

### *6.5.2 Assembly*

To realize a tracker composed of 1440 pixel and 15148 strip detector modules, industrial quality control and assembly methods and logistic structures are necessary as well as a modular design. The smallest assembled element is a module, an example is shown in Fig. 6.33. Modules were assembled in six institutes worldwide using automated robotic "gantry" systems. A gantry system consists of a camera surveying the various components, a robotic arm which can hold vacuum pickup tools for placing sensors and hybrids and syringes to precisely dispense silicone epoxies. Pitch adapter to hybrid assembly was also fully automated on a gantry. A CMS gantry is shown in Fig. 6.34. With pattern recognition a precision placement of 10 μm and 0.1 mrad was achieved. A removable assembly plate holds components for three or four modules. After the assembly, the plates were moved, with module components fixed by vacuum, to complete the curing of the glue, while the gantry was available for further assembly work. With this procedure a worldwide peak assembly rate of 90 modules/day was achieved. Modules were then sent to bonding centres with high-throughput industrial bonding machines. Bond pictures are displayed in Fig. 6.35; 25 million custom wire-bonds were necessary to finish the module construction. As the next step, modules were assembled into TIB strings, TID disks (Fig. 6.16), TOB rods (Fig. 6.17) and TEC petals (Fig. 6.36) – in the shell, rod, and petal integration centres.

Overall, 7 sensor quality assurance centres, 3 hybrid production sites, 6 module gantry centres, 2 hybrid assembly centres, 13 bonding and module testing centres, 2 rod integration and 7 petal integration centres and 7 shell integration centres struggled through construction and logistics, however they were very successful. CMS has gained experience in logistics of a large production which is widely distributed over several institutes and countries, refer to Fig. 6.32. The petal integration center of Karlsruhe is presented in Fig. 6.37.

The remaining steps were (1) insertion of sub structures into the sub detectors (2) inserting sub detectors into tracker support tube (3) installation of the tracker into the CMS detector plus (4) the connection of all power cables, optical fibres and cooling pipes.

Rods were integrated into the TOB structure in a drawer-like fashion, see Fig. 6.17. The TOB was then inserted into the tracker support tube, resting on a dedicated rail system.

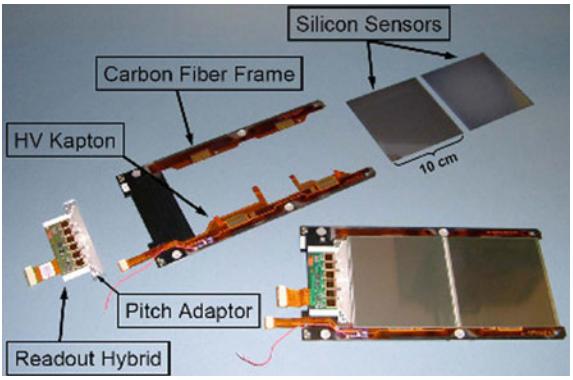



**Fig. 6.33** A CMS module. The different parts forming a module are the frame of carbon fibre and Kapton, the hybrid with the front-end electronics and pitch adapter and the silicon sensors. Courtesy of colleagues from UCSB Santa Barbara, California

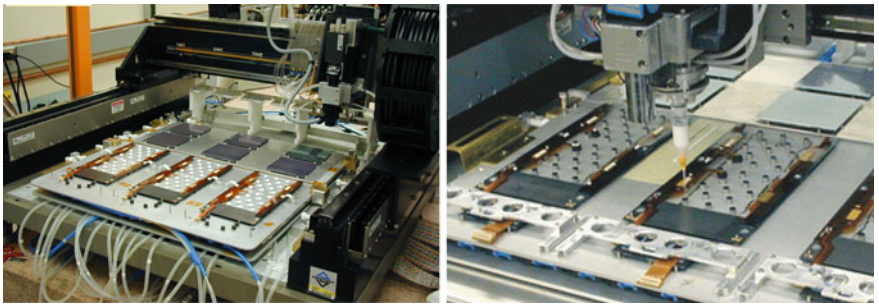


**Fig. 6.34** The gantry is a robotic assembly for modules. The gantry systems consist of a camera surveying the various components, a robotic arm which can hold vacuum pickup tools for placing sensors and hybrids and syringes to precisely dispense silicone epoxies [Courtesy of CMS]

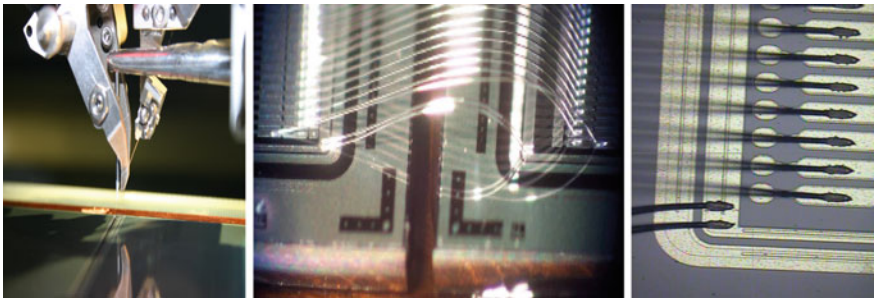


**Fig. 6.35** CMS module bonding

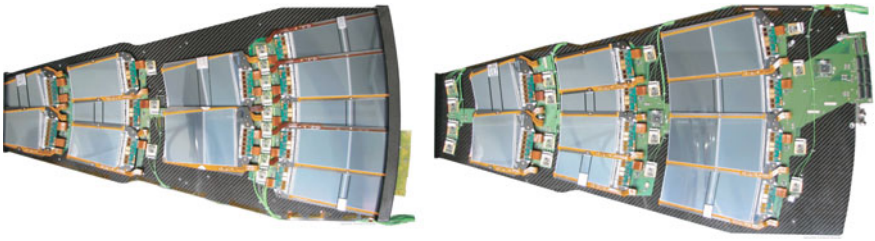


**Fig. 6.36** The inner disk petals are the largest of the eight petal types, equipped with 28 modules and more than 1 m long. The pictures [292] show both sides of a petal. All modules are assembled with an overlap and space seen taken by electronics is covered on the other side with modules. A 100% coverage plus overlap is established. All strips are following a full radial path, but some modules are rotated by 100 mrad to have two-dimensional point information. Eight different petal types are needed to cover all cases on the endcap volume. One TEC is composed of 144 petals

TIB and TID modules were mounted directly and manually onto the different TIB layers and TID disks. TIB and TID were assembled together with cables, pipes and fibres, integrated then tested and inserted into the TOB, resting on another dedicated rail system, see Fig. 6.18. The picture also shows a fraction of the amount of necessary services to be connected.

Petals were mounted on the front and backside of the nine TEC disks. Insertion happened sectorwise on one disk face only, then the whole TEC was rotated to get easy access to the other disk face. For the TEC all services like fibres, cooling and power cables were integrated even before the first petal was inserted. A sector cabling and petal placement from the front can be seen in Fig. 6.19. The TECs were then inserted into the tracker support tube on the same rail system as the TOB.

Late December 2007, 8 years after the decision to construct an all-silicon tracker, it was finally inserted into CMS, inside the barrel electromagnetic calorimeter. Figure 6.38 catches the historic moment.

The CMS spokesperson (2008 – 2009), Tejinder Virdee about the CMS tracker:

> Constructing a scientific instrument of this size and complexity, designed to operate at the LHC without intervention for more than ten years, is a major engineering and scientific achievement. More than five hundred scientists and engineers from fifty-one research institutions world-wide have contributed to the success of the project.

The meticulous design and quality assurance paid off. The CMS tracker was installed inside CMS; it was connected and is ready to operate for 10 years. August 2008 (publishing date of the first edition), about 99.5% of the tracker was working properly.

**Fig. 6.37** Petal integration in Karlsruhe. Three technicians plus two physicists were assembling two petals simultaneously. The petals were mounted on so-called petal grills which enable easy turnaround and thus access to both sides. After assembly, petals were long-term tested in fridges with additional active petal cooling

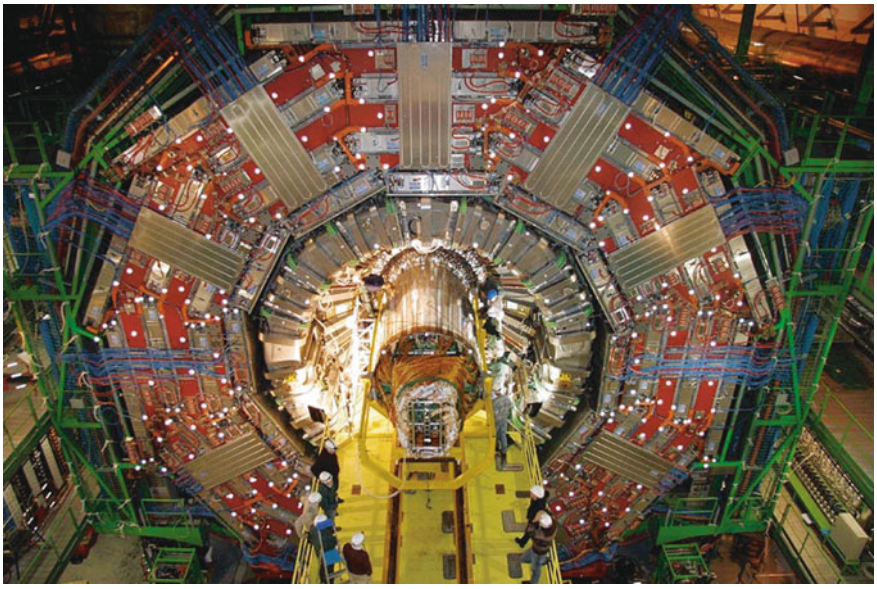


**Fig. 6.38** Tracker insertion. After assembly and testing, the CMS tracker arrives in the heart of CMS. From this moment on another 3 months were needed to connect all fibres, cooling pipes and power cables before commissioning and data taking was possible. Around 130 km of cables and 18 km of cooling pipes were installed and tested [Courtesy of CERN]

## 6.6 Tracker Operation and Performance

During RUN I and RUN II of the LHC, the CMS Tracker and CMS in general showed an excellent performance. Several papers have been published and a collection can be found here [52, 75, 103, 228, 295, 306, 344, 359]. The uptime was always excellent $\gg$99% in 2012. The SST had an uptime of 100% in 197 out of 302 so-called physics fills of the LHC machine! The main reasons for Tracker specific downtimes were failures of the cooling system, power racks or primary power supply failures, problems with individual **F**ront **E**nd **D**river boards FEDs (and here specifically the service infrastructure; failing crate power supplies). In the early days of operation and during the increase of instantaneous luminosity firmware upgrades of the FEDs were necessary to cope with rare occurrences of problems inside the detector and data hiccups.

### *6.6.1 Lessons Learned from Operation and Maintenance*

The key of successfully operating a complex and large detector such as the CMS Tracker is careful planning and early automation of regular tasks. Detector Control, Safety and Data Acquisition systems have been built defined by databases, meticulously filled with info about cabling, connection and sensor map, parameters, limits, etc. . The software then scanned all hardware items to determine the baseline parameters such as pedestals, noise, gain of laser, dead or flaky channels, etc. For the pixel detector several iterations of chip parameter tuning are necessary regularly to achieve the optimal thresholds for every individual pixel. Despite full automation, this procedure took several weeks and is being repeated after every LHC **Y**ear-**E**nd-**T**echnical-**S**top YETS and some partial aspects weekly. In general, all systems are behaving very well but some surprises and difficulties were encountered along the way.

Safety is essential for a detector, basically inaccessible and therefore not really repairable. All safety systems ran from day 0 and never failed. All relevant values like temperature, humidity, current, voltages are constantly monitored and many participate in a majority logic to interlock the detector in case of values exceeding predefined threshold values. In reality, warnings including SMSs are issued long before and a software system analyses trends and shuts down smaller parts of the system before any hardware safety interlock kicks in. With this compromise, safety is guaranteed but in case of over-temperature in a small section, the major fraction of the detector remains available for physics data taking.

In my personal experience

> It is always the low-tech that fails. It's the plumbing and cable connections. We are very good in getting all complex parts, e.g. ASICs, sensors, control and readout boards going but fail in the simple ones.

Four[21] dedicated cooling stations operate with $C_6F_{14}$ in the experimental[22] cavern serve 180 cooling loops of the SST and 2 times 8 for the Thermal Screen. These are served by the so-called "brine circuit" via heat exchangers. In 2008 we experienced a rupture of one of those heat exchangers ending up having some brine in one of the cooling plants (fortunately not in the detector). As follow up, brine has been replaced with $C_6F_{14}$ adding much to robustness but also reducing the cooling performance[23] making a large refurbishment of the plants in the first long shutdown necessary.

Later, an overpressure[24] incident during a maintenance operation weakened some joints and some of the over-pressured cooling lines started to leak a substantial amount of $C_6F_{14}$; subsequently, today, 5/180 lines are closed and the corresponding detector parts and are being only cooled by neighbouring lines via the carbon fibre structure. Luckily the margin in the system allows to operate them during the full design integrated luminosity despite changes due to radiation. And even without them the tracking efficiency would only be marginally affected – still something to avoid for future detectors. The built-in redundancy paid off. In the endcap, two cooling circuits serve alternatively the odd and even layers of the disks in one sector and simulation prove that one missing circuit would not jeopardize tracking/physics performance significantly. In some cases the overpressure moved the pipes slightly and in the very dense environment, they now touch the sensor backplanes short-circuiting the sensor HV voltage potential. Also the cooling contacts of some modules degraded resulting now in a spectrum of temperatures across the whole detector – temperature maps will be presented in Fig. 6.50 in the section about sensor evolution with radiation.

To reduce sensor leakage current and prevent reverse annealing the system was designed to keep sensor temperatures during operation at $T = -10\,°C$ and consequently the dew points below $T = -30\,°C$ to avoid any risk of condensation. These dew point specifications were met from day 0 inside the detector but not in the interface volume towards the detector where some bare pipework[25] was in contact with higher humidity values. In short, for the first years the CMS Tracker was operated about 25 °C higher than specified. Final operation conditions were only met in Run II, which started 2015 after two years of refurbishment [52]. Today, the environment conditions (dryness and cooling power) allow operation at $-25\,°C$ (coolant temperature). With these conditions the Strip Tracker is operational until long shutdown 3 LS3 (2024 – 2026) with respect to degradation due to radiation damage. The ingredients of success were (a) largely refurbished cooling plants, (b) an engineered vapour-sealing concept, (c) heater elements outside the cold volumes (bulkhead and

[21] Until 2017, an additional $C_6F_{14}$ one was serving 18 lines for the pixel detector but has been replaced with a $CO_2$ system for the Pixel Phase I detector. The $CO_2$ is situated in the service cavern and has only dedicated passive manifolds on the experimental cavern.

[22] In CMS magnetic fringe field and radiation environment; with about 60 m long transfer lines.

[23] $C_6F_{14}$ is much denser/heavier than water.

[24] Pipes had been closed on both sides with the liquid not yet thermalised to room temperature thus pressure increased during warm up. Later the possibility to close the return end was abandoned.

[25] With 180 pipes in a very small and thus crowded environment, thermal insulation was not possible.

cooling ducts), (d) highly granular[26] and precise temperature and dew point sensors also remote at the end of a gas extraction system[27], and (e) most importantly a very high flux of dry gas (installation of a dedicated membrane plant[28]. With the initial relatively low radiation levels during RUN I, no significant additional degradation has been encountered due to the elevated temperature – values have been closely monitored and some results will be presented in Sect. 6.6.2. In the first month of operating cold another unexpected phenomenon was encountered. The valves of the cooling lines, located in the experimental cavern thus inside the magnetic fringe field, are piloted by pressure of dry air. Despite a gas dryness with dew points below $-60\,^{\circ}\mathrm{C}$, after several weeks ice clogging was encountered inside pipes touching structural elements of temperature $\mathrm{T} = -15\,^{\circ}\mathrm{C}$. Water from the scarcely humid cavern air diffused through the plastic pipes! The then formed ice blocked the gas flow/pressure thereby closing the valve thus switching off cooling for parts of the detector. The solution is to avoid stagnant air by adding some micro openings in the pipes flushing out any water content.

Erik Butz, CMS SST Operations Manager and Technical Field Manager

> Seeing this detector in operation, one cannot help but marvel at how well this large and complex system works. On the other hand also after a number of years, it still manages to surprise us with new and unexpected problems

As for operation, in Run I, 97.7% of the barrel pixel and 92.2% of the forward pixel were operational. During LS1 the pixel detector had been extracted and repaired (see Figs. 6.8 and 6.9). The forward detector was fully recovered 100% minus one single chip. In the barrel part, all reasonably accessible modules have been treated arriving at 98.9% functional units. The detector has been repaired in LS1[75]. A small number of defective detector modules have been replaced. But the main problems were missing or flaky connections or some connections with inferior continuity thus showing signal degradations and slow signals. Finally one half of one layer in one sector of the barrel did not work with magnetic field, due to a tripping power channel, reducing the barrel efficiency to 98.33% in RUN II, a very respectable value.

About 97% of the SST channels are reliably operational; reason for failure are the previously described short-circuited sensor backplanes, malfunctioning sensor and electronics voltage connections, but mainly failure of the control chain circuits, called control rings.

The silicon systems are only switched-ON with so-called “STABLE BEAM” to avoid losses in the unlikely case of a beam accident. To maximise the collected integrated luminosity, every effort is undertaken to reduce the switch-ON-time. In principle two control mainframes can control the 3888 SST HV power lines but the read-back of the status values takes time thus initially four and then eight units have been installed achieving switch-ON-times of initially about 4 min and finally

[26] Several hundreds of temperature and dew point sensors have been installed to have good local understanding of the environment; something missing during RUN I.

[27] Most often called sniffer pipes.

[28] System reducing the oxygen content in air to a non-flammable value, often used in oil tankers to inert the environment preventing fire.

one minute. Several beam parameter need checking before the high voltages are raised. In the first months this was under strict expert[29] control, delegating then to regular Tracker shift experts then to the CMS central crew shifter, the so-called CMS Technical Shifter and finally, after 1.5 years of operations, all checks have been fully automated. On average the automatic mechanism is 60 s faster and prevents incidents like five minutes human reaction times. Still the system prevents switch-ON in bad or not fully understood conditions.

Another interesting surprise, during RUN I was the roughly 30% increase of radiation damage in one quadrant of the innermost pixel layer due to the local radial displacement of the detector with respect to the primary beam by a couple of millimetres. More about radiation dependence on radius will be discussed in the Sect. 6.6.2.

Despite all challenges and surprises the system is robust and redundant against failures. The meticulous quality assurance during construction and design paid off and the detector will operate longer than initially designed for and thought of.

### *6.6.2 Signal Processing, Some Key Figures and Tracking with the CMS Tracker*

This section discusses the essential Tracking and Vertexing performance of the CMS Tracker in an exemplary, educational way focussing on dependencies of design and reality – how it all fits together. It is not meant as an exhaustive description of the CMS performance giving all or even the most recent details. These are shown in much greater details here [295, 310, 312]. The next paragraphs will, therefore give a descriptive overview of signal processing, essential parameters like position resolution, Signal/Noise S/N, tracking efficiency and momentum resolution. It will also introduce the alignment challenge, detector parameter evolution with radiation during operation plus give an insight of the role of the Tracker in the CMS High Level Trigger. It should be mentioned that in the end, all results are compared with dedicated detector simulation. Basic particle interaction with the detector material plus the individual sensor element response including particle drift, induction – signal formation on strips and pixels, Lorentz angle, etc., are simulated, taking into account also alignment, bad channel list, calibration values, etc. Without having the same level of precision in simulation and detector performance no analysis is possible since the data is not well understood. These simulation are called “MonteCarlo”.

*Signal Processing*

As a first step, the detector needs to be timed-in with respect to the LHC clock to ensure having all signals of all pixels and strips in the optimal range of the preamplifier. Due to cable signal paths (cable lengths, positions within the detector, time-of-flight, back-end readout sequences) every single module has its own time

[29] Two individuals took 14 h-shifts at the experiment seven days a week for 6 weeks.

delay. With cosmics particles and even more important with collision data the delays are being scanned until the optimised/maximised signal has been found.

In addition, despite only small fluctuations in number of bad channels (strips/pixels), these are automatically identified on a run by run basis thanks to a dedicated software. The corresponding channels are masked in the offline reconstruction database, to optimise offline reconstruction and tracking. The bad components are also treated in the same way in MonteCarlo samples to reflect reality and to allow the best MonteCarlo/Data matching.

As described earlier, the full CMS data can neither be read out nor stored, and only 100 – 1000 Hz of events[30] are finally saved for further physics analysis. In addition only '*zero-suppressed ZS*' data is being stored, where values below threshold are suppressed. This sparsification in the pixels happens already in-situ on chip level while the SST implements this step in the back-end hardware.

In general for the SST, the analogue signal from the silicon sensors is preamplified, shaped and deconvoluted (rf. Sect. 1.10) by the front-end chip APV25; then is amplified, multiplexed, converted to light by the **A**nalog **O**pto **H**ybrid AOH and sent via optical fibres to the FEDs. Clock, trigger signals and slow control communication with the front-end electronics are managed by the **F**ront **E**nd **C**ontroller FEC boards and transmitted via optical fibres to the **D**igital **O**pto **H**ybrid DOH for each control ring of the Tracker: signals are distributed by the DOH to all **C**ommunication **C**ontrol **U**nits CCU. Finally each CCU sends signals to the chips of a set of Tracker modules, in particular: clock and trigger; it receives the data from detector slow control measurements provided by each module by the **D**etector **C**ontrol **U**nit DCU[31]. The DCU chip also contains a unique electronic identifier prohibiting later ambiguities or wrong assigning.

For the SST, the full pulse-height data for every strip is transferred to the back-end FED boards where the zero-suppression happens – this is often being called synchronous read out. Every individual chip channel has its 192 capacitive buffer cells. The dedicated zero suppression algorithm runs in the **F**ield-**P**rogrammable **G**ate **A**rray FPGAs of off-detector FEDs allowing more complex and adaptable comparisons. So-called "*pedestals*" (the baseline signal level of a given strip when no particle is present) and "*common mode*" noise (event-by-event fluctuations in the baseline within each tracker front-end readout chip) are subtracted suppressing any 'zero-event'. Zero-suppression accepts a strip if its charge exceeds the expected channel noise by at least a factor of five, or if both the strip and one of its neighbours have a charge exceeding twice the channel noise. As a result, information for only a small fraction of the channels in any given event is being sent to the CMS central **D**ata **Acq**uisition DAQ and finally, after **H**igh **L**evel **T**rigger HLT selection of events retained for offline storage. This data treatment is illustrated in Fig. 6.39. In special occasions the data from all strips inside the SST can be stored. This mode is called **V**irgin **R**aw VR and is used for calibration and during the **H**eavy **I**on HI periods,

[30] Design was 100 Hz, since 2012 more than 1 kHz events of data are being saved.

[31] DCU; situated on each hybrid, the DCU measures temperatures of hybrid and sensor, several voltages plus the sensor leakage current.

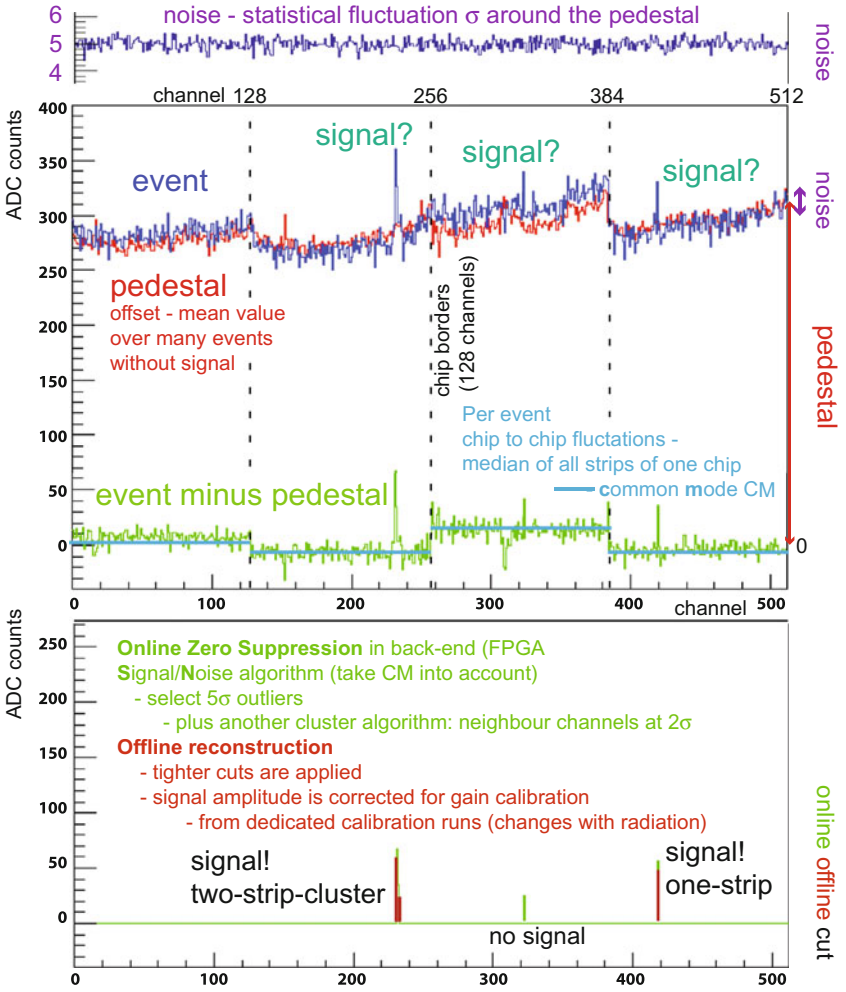



**Fig. 6.39** Each strip/channel has a "*pedestal*", the mean value of the offset with a statistical fluctuation "*noise*" around it. The noise is displayed in the *uppermost part*. Below, the plot shows the raw analogue signal (dark blue), the "*pedestals*" (*red line*) and the signal after pedestal subtraction (*light green*). After pedestal subtraction, each chip has still an additional global 'common' offset. The *lower plot* then shows the result of the online (*light green*) zero suppression a cluster algorithm running on FPGAs on the back-end electronics. This is the signal information sent to offline reconstruction where gain calibration and tighter cuts, e.g. via a more refined "clusterizer" algorithm, are applied (*red*). Lastly the information is being translated into position, taking Lorentz angle and centre-of-charge into account

where the multiplicity is extremely high but the initial collision rate is strongly reduced. One SST event in VR sums up to 14 MB just for the SST while with ZS it depends on the occupancy/luminosity and sums to about 0.5 MB at design luminosity.

Regular calibration runs are necessary to determine pedestals and noise. Further cuts happen later and positions are Lorentz-Angle corrected, 10 μm (20 μm) in the TIB (TOB). The position of the hit corresponding to each cluster is then determined from the charge-weighted average of its strips positions. The uncertainty in the hit position is usually parametrized as a function of the expected width of the cluster obtained from the track angle.

With comparably low radiation levels in the SST, no sensor signal degradation is expected and noise degradation due to increase of currents can be factored in – signal-to-noise values will remain high throughout the lifetime of the detector. Still the effects need to be followed and corrected for in the offline reconstruction stages. The optical links, on the other hand, change with radiation and the edge-emitting laser diodes have four gain stages to compensate for non-uniformities and radiation effects. In addition the laser gain is temperature dependent. Also differences in voltages and temperature in the front-end have an offset effect. The SST front-end chips send a so-called digital *tick mark*, a normed signal, allowing to adjust the laser gains to achieve a homogeneous *tick mark* at the back-end – such a calibration run is taken about every other month. Altogether the gain of every single APV chip[32] is then a parameter in the offline reconstruction allowing also dE/dx measurements and to determine if a small signal came "*out-of-time*", e.g. from a previous bunch crossing.

In the Pixel system, zero-suppression happens already at the front-end electronics sending only pulse-height information and pixel hit addresses to the back-end ady only pulse height information of a 'hit' pixel is being sent to the chip's periphery and buffered there, waiting for a L1-Trigger signal (rf. Sect. 1.11). The front-end single-pixel-threshold is about 3200 electrons (∼1500 electrons for Phase I). The system needs regular, about monthly, calibration runs due to radiation but also for environment changes, e.g. temperature. Charges are being injected into each pixel, and the efficiency is being determined. Also the FED optical receiver offset needs adjusting to keep the baseline level of the signal in the middle of the ADC range. The hit pixels signals are later correlated to pixel weighted clusters and the cluster shape determines the final hit location and resolution. In first order a centre-of-charge corrected for the Lorentz-Angle gives the location, this is used mainly in the High-Level-Trigger stage. In the full offline reconstruction so-called "*pixel templates*" are being used. These templates "know" a priori the expected pixel cluster shape depending on the spacial sensor placement with respect to the primary interaction zone; e.g. "knowing" the principle incident angle of the particle but also taking into account the Lorentz angle, the operation voltage (thus field configuration and thus drift behaviour) and the possible partial depletion configuration, etc. into account, all following the evolution with radiation.

Some more features of the pixel system: With sparsification in the front-end chips, the limited size of the internal buffer of the readout chips cause a dynamic inefficiency, where hits are lost, increasing with instantaneous luminosity and with trigger rate. At design luminosity the dynamic inefficiencies is around a couple of percent in layer 1

[32] It showed that a gain per individual channel was not necessary.

**Table 6.4** Measured Hit Resolution of the CMS SST with analogue readout compared to the binary case. The ranges of resolutions reflect the difference in cluster sizes (how many strips flag a signal), in r$\phi$ but mainly in pseudorapidity $\eta$. With larger $\eta$ thus larger angle thus larger path length in the silicon, the net deposited energy from a traversing charged particle increases. In first order, larger clusters (more strips) allow a better centre-of-charge determination, even more true for thicker sensors. This effect is much better illustrated in the next Fig. 6.40 (Pixel case) where the angle translates directly to cluster size and resolution [295]

| Sensor layer | Pitch ($\mu$m) | Resolution ($\mu$m) | Binary resolution ($\mu$m) |
|---|---|---|---|
| TIB 1 – 2 | 80 | 10.1 – 19.1 | 23.1 |
| TIB 3 – 4 | 120 | 20.8 – 29.5 | 34.6 |
| TOB 1 – 4 | 183 | 16.9 – 42.3 | 52.8 |
| TOB 5 – 6 | 122 | 11.8 – 26.6 | 35.2 |

while being too high for twice the design instantaneous luminosity.[33] This bottleneck has been resolved by the Pixel Phase I upgrade (see Sect. 6.2). Single-event upsets temporarily cause loss of information at a rate of approximately two readout chips per hour – an automatic software recovery mechanism has been established. Finally, readout errors signalled by the FEDs depend on the rate of beam induced background.

The whole process of operation and calibration is more complex than described above and requires regular but not constant effort and presence of experts. All processes are highly automated to allow also inexperienced shifters to monitor data taking and take calibration runs, while the full detector analysis is still in the regime of experts.

*The Essentials—Signal/Noise and Point Resolution*

During LHC operation, in the SST, the APV25 runs in deconvolution mode, where a weighted average of three samples is formed, effectively shortening the pulse to 25 ns and containing the full signal within a single LHC bunch crossing. This shorter pulse shape comes at the cost of a noise increase of 30%. The resulting Landau distribution, after correcting for the track incidence angle (path length corrected signal-to-noise ratio), exhibits a **M**ost **P**robable **V**alue: MPV (S/N) 18 for the thin sensors at the inner radii and MPV (S/N) 22 for the thick sensors at the outer radii. At first order, the point resolution in an analogue system is $\sigma_x \propto \frac{\text{pitch}}{\text{signal/noise}}$ compared to $\sigma_x = \frac{\text{pitch}}{\sqrt{12}}$ in a binary system. In a real large system, also angle and path length, and corresponding cluster size (signal distribution over several strips), thus location of the module with respect to the primary vertex, play a role. The point resolution, an important factor to measure the sagitta (see definition Appendix A), defining the momentum resolution, of the CMS Strip Tracker are given in Table 6.4.

The correlation of resolution to angle and cluster size is even clearer visible in the Pixel system with a pixel cell length of 150 $\mu$m. Figure 6.40 shows the spatial $z$-resolution versus incident angle in the Pixel Barrel part as derived by the "template" method; the plot also shows the corresponding cluster sizes (number of pixels

[33] It should be mentioned that the LHC is performing much better than the design specification – about twice in 2017.

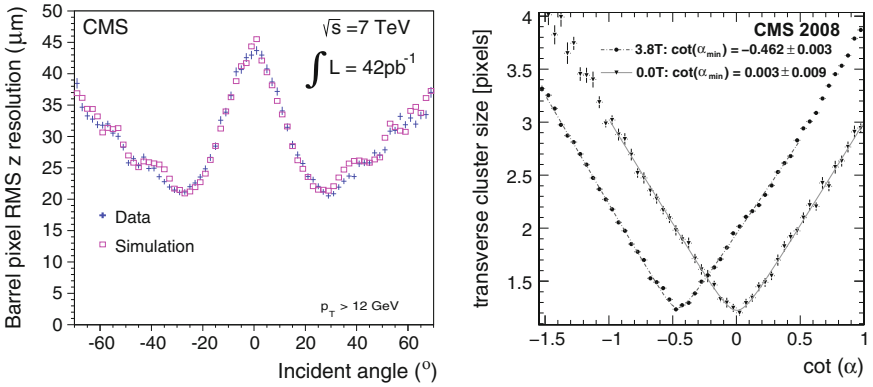



**Fig. 6.40** The $z$-resolution of the Barrel Pixel Detector versus incident angle with ionising particles originating at the interaction point z = 0 [295]. The *right* corresponding plot shows the "cluster sizes versus angle" leading to the resolutions in the *left* plot; optimal at $\cot(30\,^\circ) = 0.5 \mathrel{\hat{=}} 2$ pixel cluster. This plot is also called Happy Face and used to monitor the detector health, meaning when the plot 'smiles', everything is in order. The *right* plot also nicely demonstrates the shift of cluster due to the Lorentz angle / Lorentz shift in the magnetic field [294]

above threshold). With increased angle the path length increases thus the amount of ionisation (charge/signal) increases but also the number of pixel collecting the charge changes. At incident angle single pixels are hit and the resolution is not much better than binary while the best resolution is achieved when "a couple" of pixels participate (two-pixel cluster). At a more shallow angles many pixels see charges below threshold, part of the charge is thus "lost" and the resolution worsens. With even more shallow angle also the multiple scattering increases.

The presented resolution is about optimal thanks to the "template" approach, improvement would theoretically be possible at the detector design state to optimize pixel (sensor cells in general) cells depending more on location. But in all practicalities, chip cells are what they are and the more differences the less practical the realization of the detector. Changes of cell geometries are normally only done for different layers and barrel and forward partitions or in more small "boutique" like detectors like DELPHI, featuring changes even inside layers.

*Alignment*

As prerequisite for tracking the knowledge of the precise position in 3D space of the sensors down to the individual cells with respect to each other and also with respect to the global CMS coordinate system is of utmost importance – the so-called alignment. The paper [313] gives a very deep insight on the matter while this paragraph only gives a simple description.

Figure 6.41 figuratively illustrates the issue and since it is impossible to construct the tracker mechanically with a ∼10 μm or better precision, which is necessary to

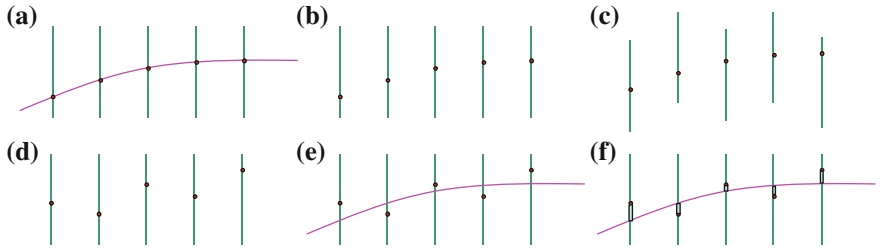



**Fig. 6.41** A charged particle traverses the detector (bent by the magnetic field) (**a**) and produces hits in the sensor planes (**b**). The sensors itself might be displaced with respect to the expected ideal geometry (**c**), and without that knowledge or correction the tracking algorithm would see (**d, e**) a zigzag around the real track and thus reconstruct a wrong track. With a good statistics of high purity reconstructed tracks, the deviations from the reality – the residuals – can be determined **(f)** thus understanding the correct geometry and feeding it back into the reconstruction algorithms – telling the database the real position of the sensor cells [10]

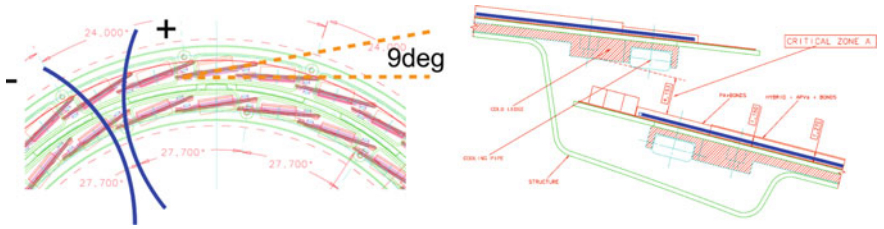



**Fig. 6.42** Modules are arranged with a dedicated overlap to ensure hermeticity but also allowing for module to module alignment inside layers, since some tracks passes both modules here. Given the space constraints this is mechanically challenging

achieve the tracking design performance, we have to "align" it later in the geometry description database. The built-geometry already takes the need of module to module alignment into account by implementing an overlap of sensors inside the individual layers – this is illustrated in Fig. 6.42.

For decent track reconstruction, one needs to achieve $\sigma(\text{align}) - \sigma(\text{intrinsic hit}) = 10 - 30\,\mu\text{m}$. Especially in the pixel, a miss-alignment of a few tenths of micrometers would seriously degrade the $b$-tagging performance. 15,148 SST plus 1440 pixel modules need to be aligned corresponding to 16,588 * (3 positions + 3 orientations + 3 deformations) individual parameters. The large number of individual modules and their arrangement over a large volume with some sensors as far as $\approx$6 m apart takes the alignment challenge to a new stage compared to earlier experiments. All in all, on the order of $\mathcal{O}(200{,}000)$ parameters are being simultaneously determined, together with O(10M) track parameters treated as nuisance parameters in the fit, with the help of dedicated computers with 256 GB of RAM each. To achieve this several millions of high purity tracks of different track topologies are necessary. The module parameters are then updated by minimizing the $\chi^2$ of the residuals (deviation from

correct hit position, see also Fig. 6.41). The final complete set of derived parameters is then simply called "Tracker Geometry".

The "Tracker Geometry" takes the following items into account:

1. internal alignment, mainly adjusting the positions and angles of the modules relative to each other
2. absolute position and orientation of the tracker including "Weak Modes" (see below)
3. bow intrinsic to sensors
4. kinks between sensors of a two-sensor module
5. stability: movements over time, especially at certain operational transients, like thermal cycles or magnet ramps

The high number of collision tracks (isolated muons and minimum bias events[34]) are perfect for the internal module to module alignment and to determine the sensor bow but fail to ascertain the absolute position and global orientation of the tracker. Tracks from the origin only traverse one side of the tracker thus do not connect left/right or up/down. They also have no means to constrain the distance of the endcaps or the distance between endcap and barrel. Any coherent change of alignment parameters for many modules, e.g. of one full sub-detector, cannot be understood by a single kinematic topology. These coherent global changes are called "Weak Modes". Examples of Weak Modes are global twists, skews, barrel detector being off-centre with respect to the endcaps, global bow, z-expansion of endcaps, etc.

A kinematic track diversity is necessary to address all aspects. Especially cosmic ray particles in jargon *cosmics*, having a completely different topology, e.g. breaking the cylindrical symmetry, connecting top/bottom and determine endcap distances, are of great significance. Beam halo tracks can also add important information to constrain the endcaps. Straight muons without magnetic field are useful but with magnetic field the particle momentum is known thus a handle on the multiple scattering is given. Muons from $Z \rightarrow \mu^+\mu^-$ or objects with invariant mass of a resonance give additional constraints.

The intrinsic bow of the sensors introduces a position error up to 10 μm and is compensated in the "Tracker Geometry".

The SST has proven to be mechanically very stable over time; with temperature ramps from RT to −15 °C, only changes of ≈50 μm have been observed relaxing back to the original position at RT – no hysteresis. More information about the measuring method can be found in [297]. Intrinsically also the Pixel detector is very stable but with thermal or magnet cycles the two mechanically unconstrained barrel half-shells can move several tens of micrometres up to a maximum of 100 μm in $z$. Collisions and cosmics are registered continuously and the full module level alignment is being tuned about 1 – 3 times per year while the pixel global structure position is being checked for every LHC fill and corrected for when necessary. In addition at the start of the year a so-called CRAFT and CRUZET campaigns (**C**osmic

[34] **M**inimum **B**ias MB events with inelastic scattering but minimum detector activity, not the high momentum or high multiplicity part, a special MB trigger exist.

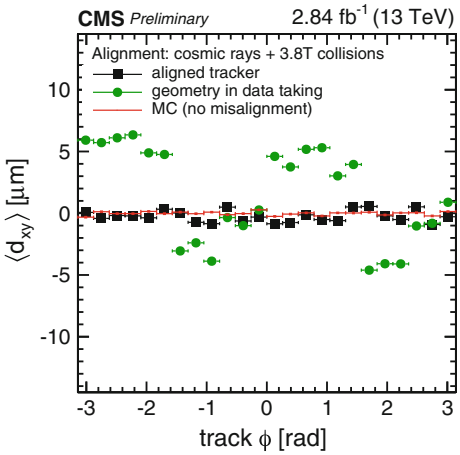



**Fig. 6.43** At construction, sensors are placed on modules with a certain precision, modules are placed on bigger structures with another certain precision and structures are assembled themselves with yet another certain precision. The whole structure then experiences magnet-, power-, cooling cycles thus the absolute position in space is, to a certain extent, an unknown. Fitting numerous tracks from cosmic rays and collisions allows a virtual repositioning of all active elements. This mathematical, statistical procedure is called alignment. The plot shows the mean of the track-vertex residuals for the case of ideal positions (*red line*); the geometrical real positions derived by data taking (*green*) and final 'reality' after software alignment process (*black*). The strips of the outer Tracker are aligned to better than 10 μm while the pixel barrel even to about 2 μm. The figure shows the ideal positions; the geometrical knowledge (*green dots*) at the beginning of 2015 and the geometrical alignment at the end of 2015 (*black squares*) [217, 300]

**R**un at **F**our **T**esla and **C**osmic **R**un at **Ze**ro **T**esla) are being done collecting a statistically significant sample of cosmics. In addition, cosmics are registered during inter-fill periods.

Figure 6.43 demonstrates the final alignment in 2015 being very close to optimal. It is clear that the initial precision, as built, need not be much better then 0.5 – 1 mm and even larger deviations[35] can be aligned during the commissioning period, given enough time and a statistically significant number of tracks.

*Tracking at CMS – an Example*

Basically, tracking is a clever "road search". The first space-point plus an assumption of its origin (beam spot) gives a first possible track direction (vector). One now 'searches' the next space-point within a 'search window'. This 'search window' is predicted by taking detector geometry, magnetic field, possible momentum and potential combinatorics into account. The first combination of points gives the 'track seed'. With a seed, the position in the next plane is "predicted"; the "measurement"

[35] One extreme example: in 2015, one pixel forward detector was globally located 3 – 5 mm away from its foreseen place, given that it was not properly fixed and the final cable connection pulled it slightly. This was full recognised and mitigated by software alignment.

is then considered and, with a weighted average, merged with the "prediction" – filtered. The prediction/filtering improves with each step and the search window size shrinks. The algorithms are based on the so-called **K**almann **F**ilter KF[36] [39, 40, 110, 206]. The KF also takes potential multiple scattering and energy loss as stochastic deviation from the straight path into account. Material description and inhomogeneities of the B-field information are also included.

In general it is a local process, but still, the reconstruction of trajectories of many charged particles is a computationally challenging task. This is especially true for large combinatorics in high occupancy environments, where more than one hit candidate is statistically found in the "predicted" location/window. Due to the high occupancy thus high combinatorics CMS utilises an *Iterative Tracking* approach, where the initial iterations search for tracks that have less combinatorics (tracks with many pixel hits, with a relatively high momentum to minimise the multiple scattering, with a strong compatibility with the beamspot and a rather small search window).

With design luminosity at the LHC the detector encounters 20 interactions resulting in a total average of above 500 tracks. After each iteration, hits associated with tracks are removed, thereby reducing the remaining combinatorial complexity, and simplifying subsequent iterations in a search for more difficult classes of tracks (e.g., low-$p_T$, or greatly displaced tracks).
In general each iteration consists of five steps:

1. seed generation with an initial track candidate estimating the initial trajectory and uncertainty
2. track finding utilising KF as described above
3. track fitting with best possible estimate of the parameters of each trajectory
4. track selection sets quality flags, and discards tracks that fail specified criteria
5. remove hits associated with identified tracks

The occupancies in the Pixels, with 0.02%, are one to two orders or magnitude lower than in the strips system with 1 – 3%. With a much higher particle density nearer to the interaction point this is counter-intuitive, but Pixel cell sizes are much smaller: $100\,\mu\text{m} \times 150\,\mu\text{m}$ compared to $80\,\mu\text{m} \times 10$ cm. The track seeds are therefore found and reconstructed in the Pixel system. High channel granularity is the key for good tracking but it comes with high sensor and electronics channel number thus higher power consumption due to larger cable mass and cooling circuits thus larger material budget thus more multiple scattering – a balance is necessary. In dense environments, like within jets, the correct hit-to-track association is even more challenging.

The iterations work inside out (1) pixel triplets then (2) pixel pairs, recovering gaps or non-working modules or dynamic inefficiency, (3) non-pointing (displaced) tracks with pixel triplets, (4) mixed pixel/strip triplets; then inner strips system (5) TIB/inner TEC and outer strips system (6) TOB/ outer TEC. This is a rough basic

[36] Kalman Filters have a broad use-case namely navigation, radar tracking, stock market prediction. Deviations as from wind, money exchange rates, *multiple scattering* can be processed as stochastic noise.

concept since the algorithms are evolving. Generally, iterations beyond the first one improve the acceptance either in $p_T$ or in displacement from primary vertex or take nuclear interactions into account. They also recover the full hermeticity of the detector by considering tracks where the particle passed through a crack or non-working element.

In a nutshell, some contradicting ingredients for high quality tracking and excellent momentum resolution are (rf. also Appendix A on p. 343):

- **key ingredient: good channel granularity** thus low occupancy thus low combinatorics
  - best would be to only have one track in the search window
  - low combinatorics especially important in the seeding layers
- no deviation from track – no multiple scattering – no mass
  - algorithm compensates
- start at lowest possible radius
- have a large lever arm – high radius
- very good point resolution
- any additional constraint helps, e.g. vertex; info from other sub-detectors

The schematic Fig. 6.44 illustrates track paths for different $p_T$ and different $\eta$ inside the 4 T solenoid field of the CMS Tracker. With this, it becomes clear that tracks below $p_T \sim 1$ GeV 'loop' inside the detector (depending on their initial $\eta$ direction. All tracks with the same transverse momentum naturally have the same track radius $R$ and sagitta $s$ (rf. formula A.1 in the Appendix). Only the transverse, orthogonal track component with respect to the solenoid B-field contributes to the momentum measurement (bending in $\phi$).

Figure 6.45 illustrates the transverse momentum resolution $p_T$ for muons with fixed $p_T$ against pseudorapidity $\eta$ (left) and against momentum (right). Figures for pions look similar, resolution for electrons is worse, for more details refer to [295]. The degradation of $p_T$ resolution for higher $\eta$ stems from the decreasing effective lever arm $L$ (see schematic Fig. 6.44). This effect is even more pronounced for high $p_T$ tracks. The step at around $|\eta| = 1$ is due to the gap between Tracker barrel and endcap disks and different resolutions in the disks itself. Note the extra complication for the endcap (TEC and TID), where the accurate measured coordinates are $z$ and $\phi$ and with wedge shape sensors also the pitches thus point resolutions change within an individual sensor, compared to $r$ and $\phi$ measurement with fixed pitch for the barrel configuration. The optimum resolution is achieved for particles around 3 GeV. At lower momentum, the resolution is fully dominated by multiple scattering and its value reflects the amount of material traversed by the track (more at higher $\eta$). At higher momentum the point resolution and even small miss-alignments play a role – the track is simply straight and a sagitta is difficult to measure precisely. For high $p_T$ muons the larger lever arm taking the muon detector into account helps a lot – see Fig. 6.55.

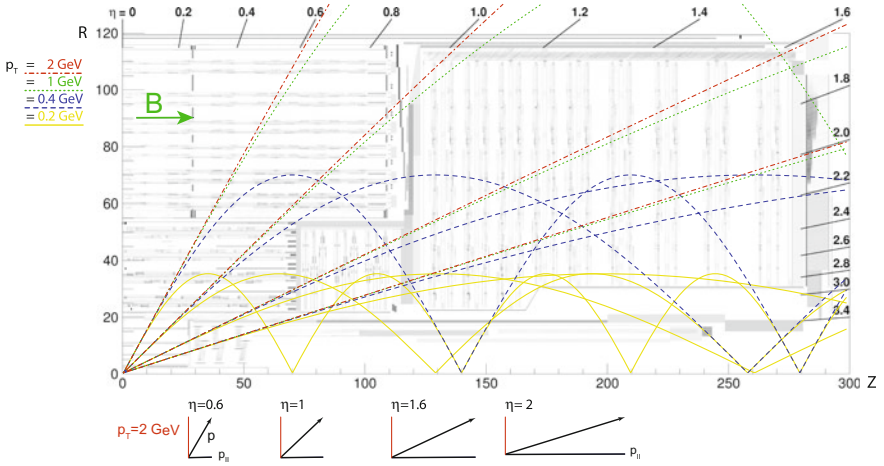



**Fig. 6.44** Tracks in CMS – a schematic. For $\eta = 0.6$, 1, 1.6 and 2 the plot features four tracks with the same fixed $p_T$: 0.2 (*yellow*), 0.4 (*blue*), 1 (*green*) and 2 (*red*). All tracks of the same transverse momentum $p_T$ have the same track radius $R$ and sagitta $s$ in the solenoid B-field; their momentum $p$ though changes with initial direction. Tracks, leaving the tracker at $|\eta| > 1.6$ (in the endcap not in the outermost radius), have a shorter effective lever arm $L$. For example the *red* 2 GeV track has an $L = 1.1$ m for $\eta = 0.6$, 1 and 1.6 but only about 70 cm for the $\eta = 2$ direction. The 'looper' 0.4 GeV (0.2 GeV) track does not reach the Tracker radius anyhow thus sees only $L = 70$ cm ($L = 35$ cm) or $L = 60$ cm for the $\eta = 2$ case. Not demonstrated here but the low energetic tracks are also affected more by multiple scattering. The cartoons on the *bottom show* the $p$ components ($p_T$ and $p_\parallel$) for the $p_T = 2$ GeV tracks for the different $\eta$ directions. Figure from [260]

In short

- versus $\eta$: path length (multiple scattering), switch from barrel to endcap configuration and effective lever arm $L$ defines the $p_T$ resolution
- versus $p_T$
  - at low $p_T$ multiple scattering dominantly degrades the $p_T$ resolution
  - at high $p_T$ point resolution and large lever arm are the important factors

Another complication is that particles of different type interact differently with the detector materials and their paths are different. Expressing these differences, tracking efficiencies for muons, pions and electrons are displayed in Fig. 6.46.

The efficiency for muons is basically 100%, with no strong interaction and basically no electromagnetic radiation (Bremsstrahlung), they traverse the full volume and the multiple Coulomb scattering and energy loss is effectively included in the KF. Tracking information from the outer muon detectors give a long lever arm improving momentum resolution significantly, especially at high momenta (straight tracks).

Charged pions, additionally, are subject to elastic and inelastic nuclear interactions, deviating grossly from the original path. Tracks can be interrupted or two tracks with less hits or no tracks at all are being identified. They are also stopped in the calorimeter (giving another crude space point).

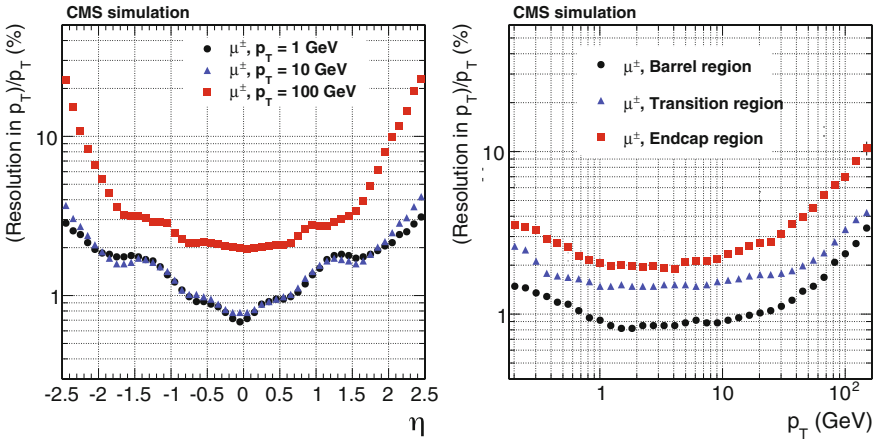



**Fig. 6.45** $p_T$ resolution of muons for different energies versus $\eta$ on the *left* and versus transverse momentum $p_T$ on the *right* [295]

Electrons are much more complex due to their additional energy loss due to Bremsstrahlung, a highly non-Gaussian process and therefore non-standard for the KF. About 35% of electrons radiate more than 70% of their initial energy before reaching the **e**lectromagnetic **cal**orimeter ECAL that surrounds the tracker. Sometimes they do not even reach the outer tracker layers thus produce less hits. In general, their path is less stringent and with momentum loss along the way, changes bending thus changing estimated 'search window' position. To compensate, the energy and position measured in the electromagnetic calorimeter plus the assumption that the electron originated from the beam spot, constrains the trajectory further [305].

Another interesting case are very displaced tracks (identified in later iterations) from photon conversion where the vertex ($e^+e^-$) can be anywhere with material present. Radiated photons (e.g. from electron Bremsstrahlung) can convert to electron-positron pairs producing new tracks in the middle of the tracking volume, complicating the situation even further. The effect is enhanced in the endcap section due to higher material budget especially in the interface region. For electrons, the highest quality information comes from the pixel region before Bremsstrahlung kicks in.

Figure 6.47 shows the space points where photon conversion happened, clearly reconstructing the modules and mechanical supports as well as cooling and cables. Especially impressive is the correlation of photo conversion and cooling pipe presence.

Finally all tracks are extrapolated towards the primary or secondary vertices. Figure 6.48 shows the resolution of the primary vertex. The more tracks, or better said richness in kinematic diversity and higher integrated momentum, the easier to constrain the vertex and thus the better the resolution. Even with the worst resolution, primary vertex miss-identification is practically excluded.

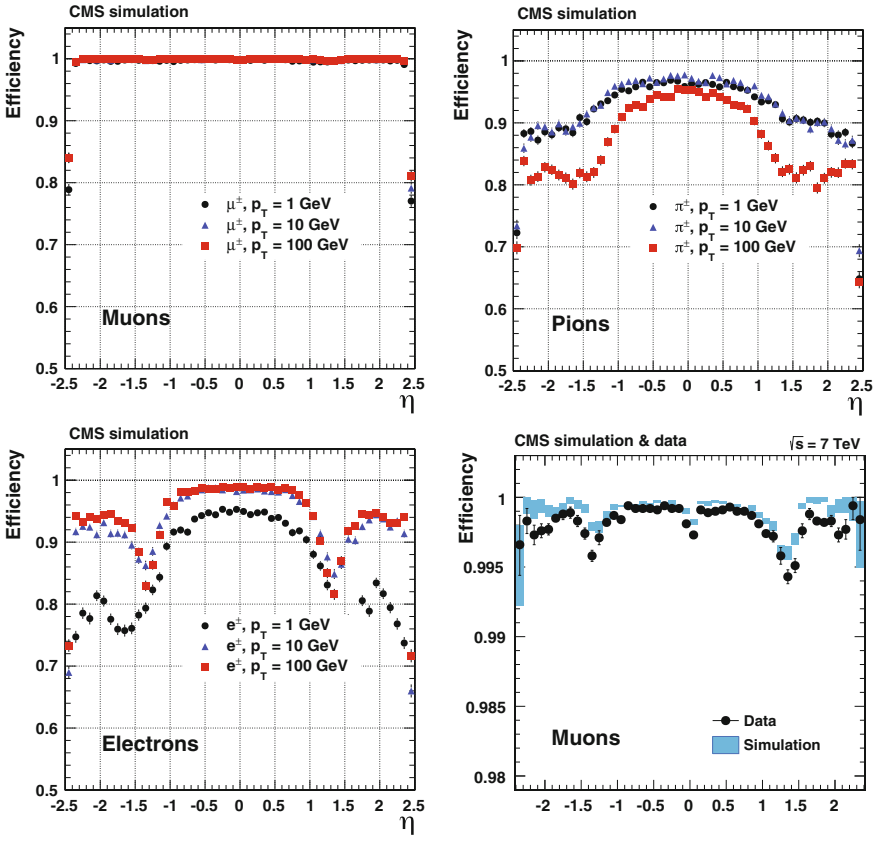



**Fig. 6.46** Simulated tracking efficiency for muons, pions, electron and muon efficiency compared with data. The *lower right* plots display a zoom of the first plot and compares data to simulation [295]

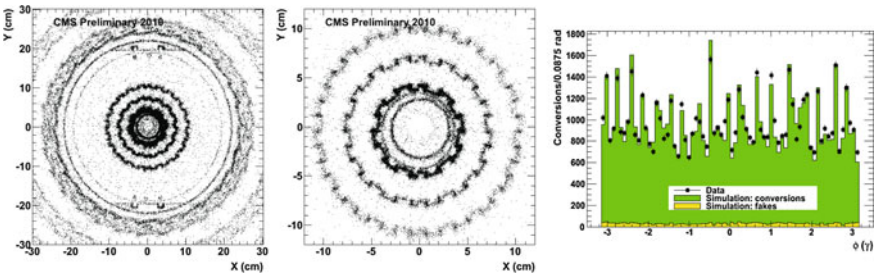



**Fig. 6.47** Photo conversion as reconstructed with data. The spacial occurrences clearly depict the support structure module and services. The *middle* figure is a zoom of the *left* one. Also the beam pipe is clearly discernible; it is slightly displaced with respect to the pixel envelope. The *right* figure shows the correlation of photo conversion and cooling pipes locations [307]

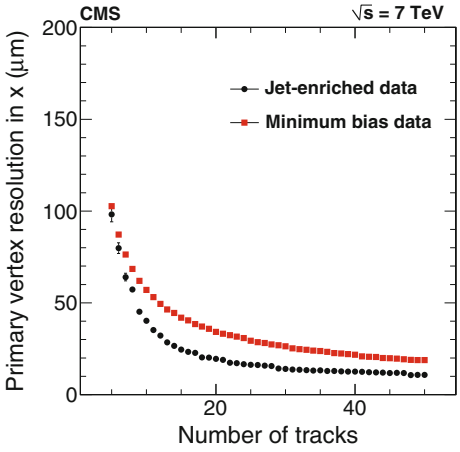



**Fig. 6.48** The jet-enriched sample has significantly higher mean $p_T$, resulting in better resolution in the track impact parameter and consequently better vertex resolution, approaching 10 μm in x for primary vertices using at least 50 tracks. The resolution of the reconstructed vertex position is driven by the pixel detector, since it is the sub-structure that is closest to the interaction point and has the best hit resolution [295]

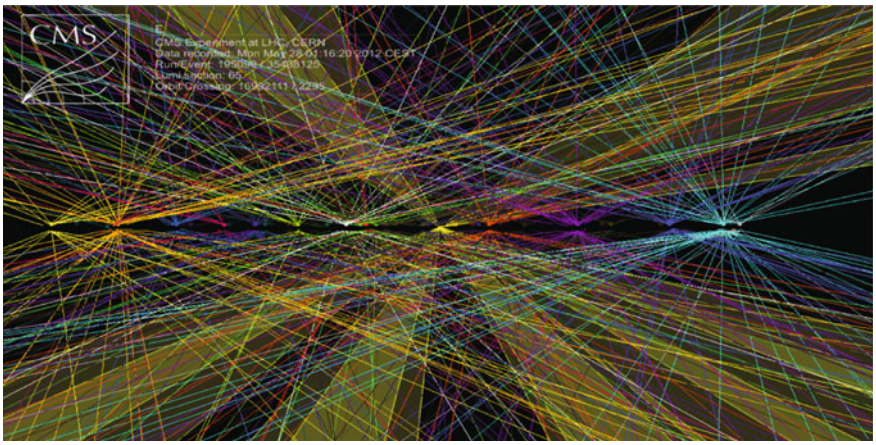



**Fig. 6.49** The situation at high pile-up of about PU = 45 with about 1500 tracks cross the CMS tracker every 25 ns. The spacing between individual vertices is about 1 – 2 mm [Courtesy of CMS/CERN]

Figure 6.49 gives an impression of the real environment – a zoom into the primary interaction zone depicting the multiple reconstructed/interpolated tracks at about nominal luminosity. More event displays with lower and higher **p**ile-**u**p PU can be visited in Fig B.6 on p. 355 in the Appendix.

*High-Level Trigger*

Another key role is the Tracker's extensive usage in the **H**igh **L**evel **T**rigger HLT for all physics event streams [274]. The Level-1 trigger system is based on custom electronics and reduces the input rate of 40 MHz bunch crossing[37] by a factor of 400 down to 100 kHz. The Level-1 decision must be made in 4.8 μs,[38] corresponding to the maximum storage capacity on the front-end chips. The HLT, a software trigger, relies on commercial processors and reduces the stored data by another factor of 1000 down to 100 Hz (since RUN II 2015 more like a factor of 100 down to 1 kHz). While the tracker is not participating in the Level-1 decision, it plays an important role in the HLT. Simplified and much faster versions of the full track reconstruction algorithms, with reduced accuracy, are used in the HLT.

The three basic principles are

- pixel only tracking
- "*regional tracking*": The HLT track reconstruction is restricted to the **R**egion(s) **o**f **I**nterest RoI, which is defined by the information from the Level-1 triggers from the muon and calorimeter systems.
- "*conditional tracking*": Cuts that stop reconstruction during event selection are introduced, like "momentum is definitively above or below threshold", "N good hits found". Often four or six space points are enough for track reconstruction. The transverse momentum $p_T$ can be calculated with only a partial track, since it does not significantly[39] change up to the end of the track – good enough for triggering using strip layers after pixels.

These fast reconstructions subsequently allow general $p_T$ cuts, defined by physics requirements (momentum cuts, keeping events most relevant for physics studies, are for example 0.9 GeV for $b$-tagging or up to 20 GeV for muons). Low momentum coiled tracks are not reconstructed.

As an example a $b$ jet HLT trigger takes a Level-1 trigger from the calorimeter (jet), the track seed for conditional track reconstruction is created around the Level-1 jet direction. Tracks are solely reconstructed in a cone (RoI) of $\Delta R < 0.15$ around the jet direction. This makes a jet direction refinement possible, to distinguish if it comes from the primary vertex or not. In short, the number of tracks with transverse impact parameter statistically incompatible with a track originating from the beam-line are counted – thus originating from a secondary vertex.

A second example is the $\tau$ trigger. A narrow jet in the calorimeter defines the RoI – cone around a jet. The $\tau$ hypothesis requires an isolated track associated to the jet. Conditional and regional tracking in the RoI is performed. The $\tau$ hypothesis is falsified if no track is found in a narrow cone around the jet (signal cone) or an additional track is found in a second broader cone (isolation cone) – *regional tracking*. To save CPU time, the track is not fully reconstructed and only a loose primary vertex association is requested – *conditional tracking*. Later at the offline reconstruction stage,

[37] With a pile-up of PU = 25 one has about $10^9$ interactions every second.

[38] 192 cells times 25 ns.

[39] Except for electrons – Bremsstrahlung.

the full track and decay vertices are reconstructed. To reconstruct a full "good" track, the track seed plus five helix parameters (two positions, two angles plus curvature) have to be determined.

In general not all stages of iterative tracking are executed and good use is being made of pixel hits with intrinsic 3D information. For example seed tracks consists of only pairs[40] of hits (not triples) and in addition the seed track has to originate from within few millimetres of the primary vertex.

Tracking uses only about 20% of HLT resources, due to the fact that track reconstructions is only executed after other requirements have been satisfied, reducing the rate significantly, e.g. jets, or high energy in the calorimeter or tracks in the muon system. Further steps then give the 3D primary vertex and impact parameter to have a handle on *b*-tagging.

*Evolution of Sensor Parameters with Radiation at CMS*

Are the parametrization models describing the evolution of sensors parameters as presented in Chap. 2 valid for long term operations and do they have projective power for the future? Some early publications about how detector evolution with radiation is monitored in the LHC detectors can be found here [24, 133, 345].

Radiation levels span more than three orders of magnitude within the tracker volume. With mainly charged particle irradiation in the centre and backscattered neutrons from the calorimeter in the outer radius, basically every particle mix and level presents itself – several thousand radiation campaigns in one go. Figure 7.1 on p. 292 gives an impression of the radiation level distribution although the quantitative levels show the **H**igh **L**uminosity LHC HL-LHC case. In addition, temperatures are different during maintenance and operation periods thus also annealing processes can be investigated. There is also a broad spread in sensor temperatures in the SST, due to some strongly degraded cooling contacts and some inoperable cooling lines (see Fig. 6.50a). In CMS, every SST module features an ASIC (DCU) measuring sensor and hybrid temperatures plus sensor leakage currents and hybrid low voltages during operation. This allows a fine granular monitoring for the different cases of radiation and temperature (annealing).

As a reminder and to put the following plots and results in perspective:

- the Tracker is *being operated* cold ($T_{silicon} = -10\,^{\circ}\mathrm{C}$) to lower the leakage current of the moment. Reminder: $I(T) \sim T_{silicon}^2 \cdot e^{-\frac{E_{eff}}{2 \cdot T_{silicon} \cdot k_B}}$
- the Tracker is *being operated* cold to "freeze" diffusion processes thereby preventing reverse annealing of $N_{eff}$ (depletion voltage); $T_{silicon} = 0\,^{\circ}\mathrm{C}$ or even $+5\,^{\circ}\mathrm{C}$ would be sufficient
- the Tracker is *maintained cold* to "freeze" diffusion/annealing thus preventing reverse annealing (depletion voltage)
- the Tracker is being warmed up *during maintenance* for some period to profit from beneficial annealing (depletion voltage) and current annealing

[40] Worth to note: Since 2017 with the phase I pixel detector, the pixel seeding is composed of quadruplets with triplets as backup instead of triplets with pairs as backup.

- annealing of leakage current always decreases the current and the higher the temperature the faster the decrease – thus higher temperatures are beneficial but the requirement is contradicting the requirements above to prevent reverse annealing of $N_{eff}$ (depletion voltage)

Figure 6.50 shows so-called Tracker maps for different parameters (temperature and leakage current increase with radiation normalised to volume or to volume and temperature). Every small square, wedge and triangle (in case of stereo modules – two sensor closely spaced back-to-back) represent a single module and the colour shows the parameter value (°C or $\frac{\mu A}{fb^{-1}cm^3}$). The SST barrel layers are simply unfolded flat and endcap disks are displayed separately. For example "TOB1" shows the innermost TOB barrel layer which has stereo-modules thus triangles inside the squares representing the inner (triangle) or outer (triangle) module of the sandwich (square). All black fields are 'bad components', i.e. modules not participating in the readout or showing false DCU readout (see first map for details).

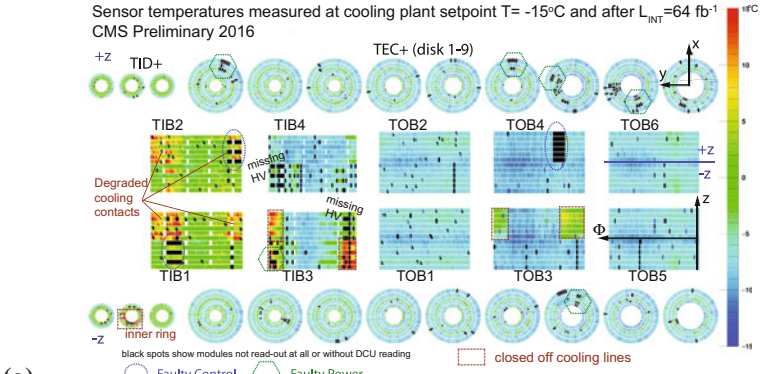
The first plot **(a)** presents the individual sensor temperatures for a cooling plant set temperature of −15 °C illustrating a large temperature spread. The double module layers (TIB1,2 & TOB1,2) are hotter per se, while the individual hot modules or regions within layers represent degraded cooling contacts or non-operational cooling loops. For higher operating temperatures and the same radiation exposure, we expect higher operation current but smaller current increase due to simultaneous annealing.

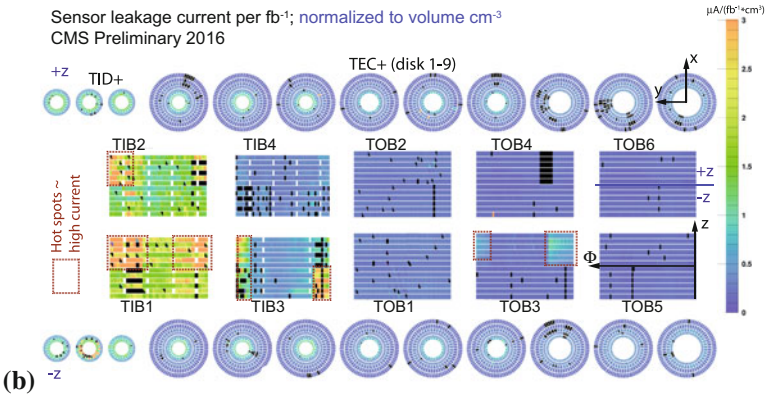
The second map **(b)** shows the corresponding leakage current increase per $fb^{-1}$ normalized to volume[41] after several years of operation thus significant radiation exposure. The plot is not normalized to temperature thus the warmer TIB layers show a higher leakage current and the hot spots within a layer exhibit significantly higher currents, demonstrating the need to *operate* cold; about every 7 K the current doubles/halves. At any given moment, it is possible to "dial" the desired current by changing the temperature within the limits of cooling power and ambient humidities. Due to the exponential current-temperature dependence 'subtleties' as current-radiation dependence, meaning layer-to-layer dependence, are not discernible here.

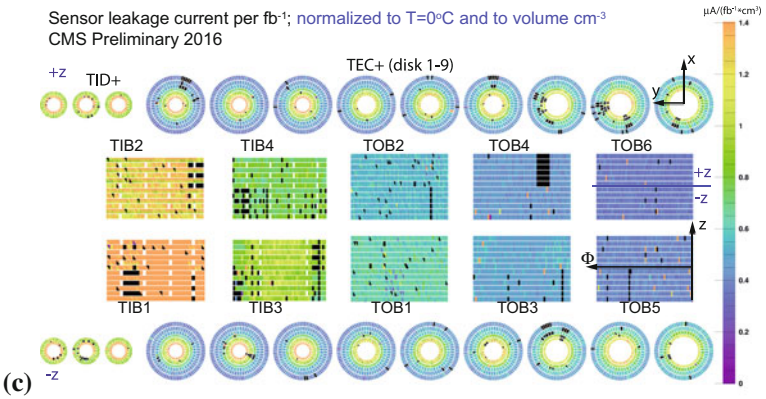
With the next map **(c)**, normalised to volume and temperature (T = 0 °C), a clear layer-to-layer dependence becomes visible corresponding to the different radiation exposure levels. The radial radiation dependence is also nicely visible in the endcap disk spanning over the full radius. Looking closer, one discerns higher currents for disks at higher $z$ position (disk 8&9) meaning closer to the endcap calorimeter thus being effected by neutron backscattering[42]. Still the differences in currents presented do not reflect solely the difference in radiation exposure but the whole thermal history. As discussed in Sect. 2.1.2 the effective $\alpha$-parameter in $\frac{\Delta I}{V} = \alpha \Phi_{eq}$ is temperature dependent; thus the evolution with radiation differs for different temperatures thus here differs for the different modules. The evolution of currents is a convolution of crystal damage induced by radiation and the parallel annealing process – the higher the temperature, the more annealing (self-repairing), the smaller is $\alpha$. TIB Layer 1

[41] E.g. TOB has 2 modules and each TOB sensor is 5/3 times thicker than a TIB sensor.

[42] A moderator has been installed instrumented between endap calorimeter and Tracker to limit the neutron exposure.



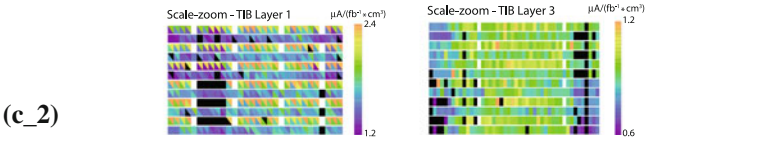




**Fig. 6.50** (Continued)

◀ **Fig. 6.50** The figures show so-called Tracker maps where the squares, wedges and triangles represent individual modules and their parameter values are *colour* coded – mind the different scales and units. The larger rectangles represent the barrel layers 1–10 while the large circles show the endcaps spanning the full radius of the tracker. As an example TOB L1 means Tracker Outer Barrel – Layer 1; the black spots show "bad" modules as detailed in plot (**a**). TIB L1+2, TOB L1+2, TID R1+2, TEC R1+2+5 feature double modules and therefore have a higher temperature. The features of the four plots are described in the text above. TEC disks of higher number are located at higher $z$ thus closer to the endcap calorimeter. In order, the maps show (**a**) silicon temperatures for cooling plant set temperatures of $-15\,^{\circ}\mathrm{C}$, (**b**) sensor leakage currents ($\Delta I$) increase per $\mathrm{fb}^{-1}$ normalized to volume for a cooling plant set temperature of $-15\,^{\circ}\mathrm{C}$, (**c**) leakage current increase, normalized to volume and temperature ($\mathrm{T}=0\,^{\circ}\mathrm{C}$) and **(c_2)** a 'scale-zoom' into TIB L1 and L3 to overcome the saturation effect of (**c**). All values were measured after an integrated luminosity of $\mathcal{L} = 64\ \mathrm{fb}^{-1}$ and with cooling plant temperature of $\mathrm{T} = -15\,^{\circ}\mathrm{C}$ [Courtesy of CMS]

and Layer 3 of map **(c)** give a faint hint about this feature. The difference of $\Delta I$ within a layer (especially TIB1 and TIB3) is purely defined by temperature; the hotter the module the lower the current increase. Tracker map **(c_2)** shows a 'scale-zoom' into TIB L1 and L3 where the hot regions exhibit a significant lower current (more bluish), normalized to $\mathrm{T} = 0°\mathrm{C}$ – compare with hot regions in map **(a)**, mind also the different scale which, again, represents the different levels of radiation exposure.

Figure 6.51a shows the current evolution versus integrated luminosity for the pixel detector layers; the currents are normalised to $0\,^{\circ}\mathrm{C}$ emphasising the behaviour with radiation and annealing. The predicted linear dependence with radiation fluence is given; layers at lower radii see higher fluence for the same luminosity thus exhibit a steeper slope. The distinctive drops represent maintenance periods, **Y**ear-**E**nd-**T**echnical-**S**tops YETS (a couple of months in time) where the temperature was partially at room temperature. The **H**eavy **I**on HI operation period, about 1 month long, before each YETS features a comparably very low radiation fluence thus no rise in current and already some annealing. The big drop in the middle represents the **l**ong **s**hutdown **1** LS1 when the detector experienced room temperature for more extended periods.[43] The increase in slope after LS1 is due to the decrease in operation temperature from cooling set point $0\,^{\circ}\mathrm{C}$ to $-10\,^{\circ}\mathrm{C}$; the simultaneous annealing during operation is largely reduced at the lower temperature ($\sim 0\,^{\circ}\mathrm{C}$ on sensor). Due to initial lower instantaneous luminosity the equivalent radiation damage (first $5\ \mathrm{fb}^{-1}$) has been accumulated over a longer period thus more time was available for annealing thus the initial slope is less steep.

Figure 6.51b presents the leakage current of a representative single module in the SST versus time, neither normalised to volume nor to temperature. The measurements, versus time, are compared to the model prediction. The model works iteratively, each day the increase with fluence and/or the decrease with annealing is being separately calculated and then summed up. The FLUKA simulation [102, 200] provides a conversion factor between luminosity and radiation fluence for the different locations in space thus for each individual module. For the annealing component the actual sensor temperature is taken into account – measured daily. The periods of

[43] The extracted pixel detector was still kept cold most of the LS1 period.

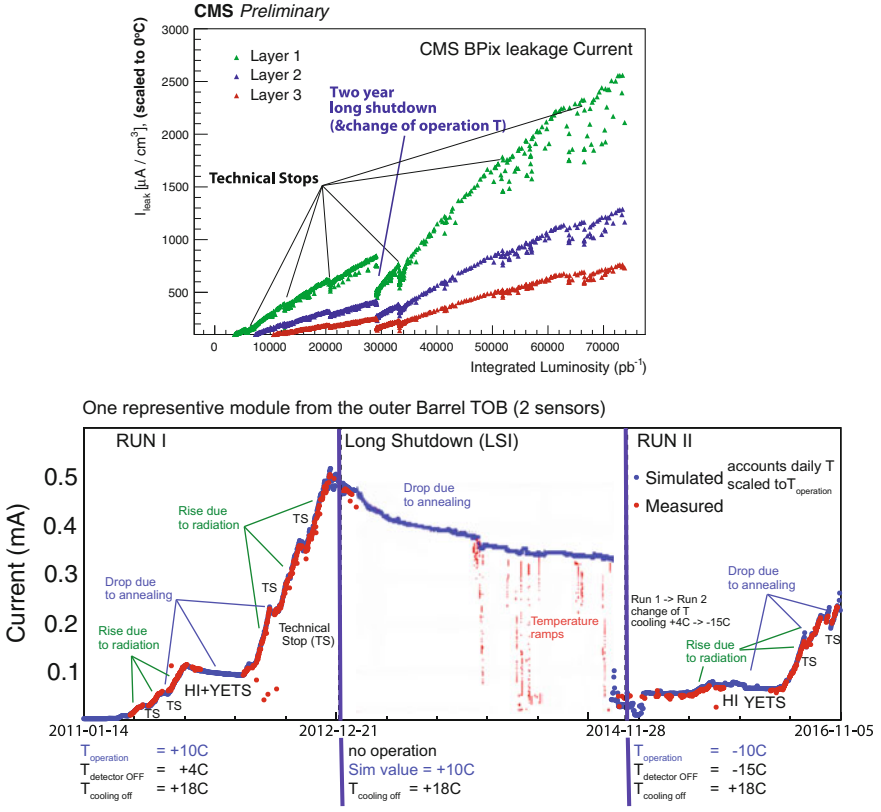



**Fig. 6.51** Leakage Current versus luminosity or versus time. (**a**) Shows the mean leakage currents of the three pixel layers normalised to 0 °C exhibiting different slopes corresponding to the fluence at the given layer(radius). The drops are due to enhanced annealing during extended periods at room temperature (Technical Stops TS or YETS). More in the text description. (**b**) Shows the evolution of leakage current of a single SST module with time. It is neither normalised to temperature nor to volume. The *red dots* show the measured current, the *blue dots* the modelled current assuming normal operation temperature. The measurements rely on the full **d**ata **acq**uisition DAQ system thus is not available at all times, e.g. mostly not during LS1. The current drops during maintenance/annealing periods are clearly visible and are modelled correctly as well as increases during the data taking periods where the radiation damage induced current increase dominates but the annealing component is still present. Beginning as of 2015, the operation temperature has been reduced by almost 20 °C reflected in the measured and simulated currents [Courtesy of CMS]

data collection (radiation exposure) and technical stops (annealing only) are clearly visible.

At the beginning of 2014, in the middle of the long shutdown 1 the first temperature ramps had been done. This is visible by the "falling" red dots where the current follows the temperature.[44] From beginning of 2015 onwards the model takes the new temperature into account, see story in Sect. 6.6.1 on p. 257.

[44]For the simulation the temperature is only taken into account 2 times per day thus the simulated currents (blue) do not follow here.

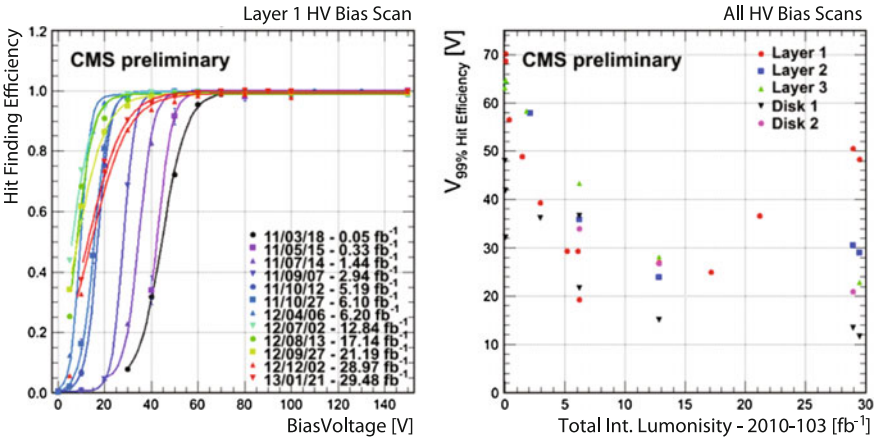



**Fig. 6.52** The *left* plot illustrates the Layer 1 CMS pixel HV bias scans, where the start of full efficiency marks the depletion voltage $V_{FD}$. The *right* figure plots the points of full efficiency for all scans and all pixel layers/disks. This illustrates a clear decrease and later rise after point of SCSI [75]

At regular intervals so-called high voltage HV bias scans are conducted, where the operation (reverse) voltage is being scanned and for different voltages, the corresponding hit finding efficiency is measured for the Pixels and the signal height for the SST. The full detector is normally scanned once or twice per year while a smaller representative part is scanned about monthly[45]. The plateaus in Fig. 6.52 show the point of full efficiency (or full signal) when the detector is fully depleted. Figure 6.52a nicely demonstrates with increasing integrated luminosity the shift of full efficiency to lower voltage values and from a certain value back up again. This point marks the point of **S**pace **C**harge **S**ign **I**nversion SCSI (rf. Sect. 2.1.1) when the initial $n$-type bulk ($n$-in-$n$ pixel sensor) changes to $p$-type. Figure 6.52b shows the points of full hit efficiencies for all Pixel barrel layers and endcap disks. Layer 1 and Layer 2 are already inverted from $n$- to $p$-bulk, i.e. overall positive to negative space charge while Layer 3 and all disks are still $n$-type. 2017, no SST sensor has yet reached the SCSI point but depletion values are decreasing, as expected. The presented situation is a convolution of radiation damage and annealing periods; see Sect. 2.1.

All in all, the Hamburg Model parametrises the situation for $n$-type Float-Zone sensors extremely well in case all parameters are taken correctly into account. Projections show full functionality for the design integrated fluence with good margin. All initial design parameters prove to be good, e.g. no signal decrease or increase in noise has been observed for the CMS SST (Spring 2017).

In addition to the values described above, also the Lorentz Angle, the laser gain, signal, signal-to-noise, ASIC low voltage currents, thresholds, etc. change.

[45]The CMS data during a full scan is not useful for physics analyses; during a small scan it is.

## 6.7 Physics with the CMS Tracker and High-Level Trigger

For the nominal luminosity operation, 500 tracks of leptons, heavy quarks, corresponding jets plus jets from additionally created mesons – mostly pions – are expected. A simple event is shown in Fig. 6.53.

Interesting physics signals are many orders of magnitude less frequent than background events! Cross-sections vary over many orders of magnitude where physics processes have to be extracted:

- inelastic: $10^9$ Hz
- $bb$ production: $10^6$ – $10^7$ Hz
- $W \rightarrow l\nu$ : $10^2$ Hz
- $tt$ production: 10 Hz
- Higgs (100 GeV/c$^2$): 0.1 Hz
- Higgs (600 GeV/c$^2$): 0.01 Hz

The Tracker is essential, not only to reconstruct charged particles and measure their momentum, but also for **H**igh **L**evel **T**rigger HLT, lepton identification ($\mu$ and $e$), $b$- and $\tau$-tagging; also particle flow (e.g. for jet reconstruction) in the full energy range. In combination with the muon detection system and the calorimeters, the CMS Tracker substantially improves momentum and energy resolution. Track isolation cuts can, for example, distinguish between muons coming from jets or from heavy objects. The CMS tracker is also ultimately necessary for all analyses to precisely reconstruct the primary vertex.

The CMS Tracker also plays a role in conjunction with the other sub detectors to identify the particle (particle ID); essential for physics object reconstruction. The calorimeter at higher radius is not able to determine if the energy depositing particle is neutral or has a charge, meaning the electromagnetic calorimeter cannot distinguish between electrons and photons nor can the hadron calorimeter distinguish between neutron and protons. The presence or absence of a pointing track (from a charged particle) arbitrates here. Especially in a dense environment it is important to show that the photon in the calorimeter has no track in the tracker, thus is no electron. As further step the full momentum and energy in all sub detectors are taken into account to form the full and most precise picture – this concept is called **P**article **F**low PF [293, 305]. The PF concept combines the whole information from CMS and afterwards 'forgets' about the detector and use the reconstructed particles as if they were coming from an event generator. The simple story of particle ID with different sub detectors is illustrated in the Appendix in Fig. B.3 on p. 352.

Figure 6.54 illustrates the flexibility of the CMS trigger where data saving on certain energy ranges can be enhanced – trigger selected. The Tracker, dominating the muon momentum resolution at lower energies, participates in selection here during the HLT stage.

### Tracker Contributions to Precision Measurements

Several parameter determinations, e.g. electron energy, cannot be attributed to a single CMS sub detector. The following paragraphs introduce the tracker contribution to

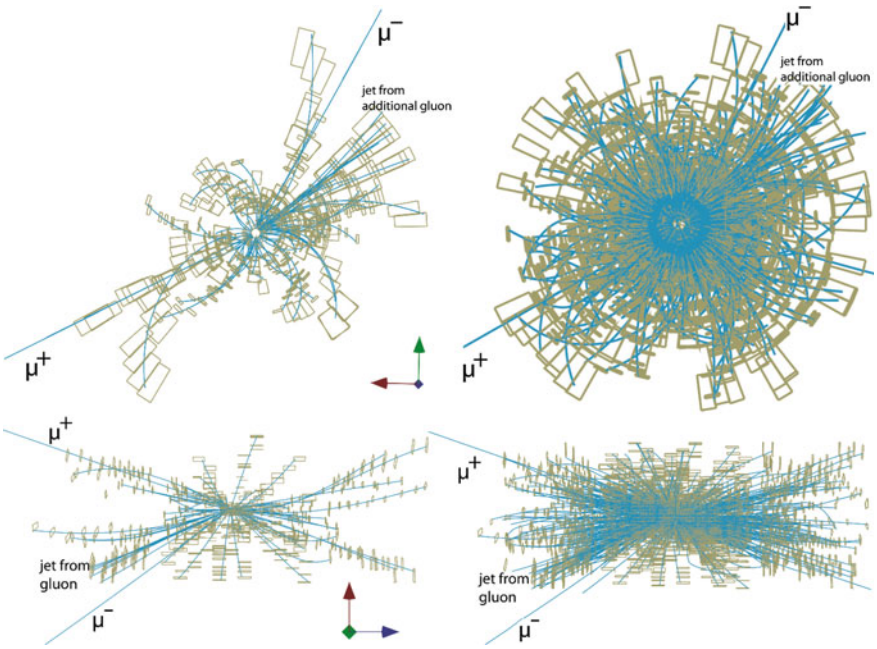



**Fig. 6.53** $p\bar{p} \to Z + X; \hookrightarrow Z \to \mu^+\mu^-$ (simulation). The event displayed is a very basic one. As a primary reaction a u quark plus a $\bar{u}$ quark of two protons annihilate to form a Z, subsequently decaying into $\mu^+$ and $\mu^-$. The additional visible tracks are particles (mostly pions) from additional soft processes of gluons and other quarks of the two protons. The *left* picture shows the bare event. The *right* displays the same event with five additional minimum bias events from soft processes of other proton pairs. *Below*, the same event is shown from another angle. At high luminosity operation about 20 minimum bias events are expected. About 500 additional tracks are expected

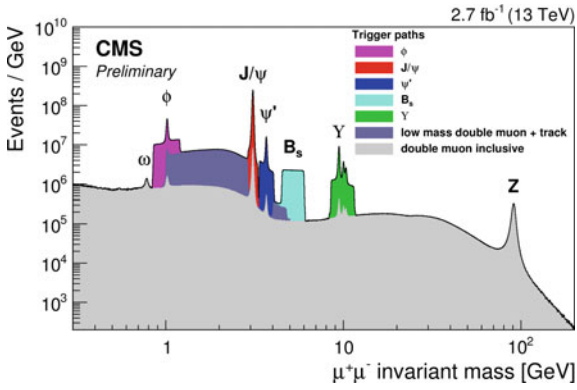



**Fig. 6.54** Muon Momentums at CMS and HLT. The *coloured* steps are energy ranges saved and reconstructed with higher priority defined by the trigger menu [296]

muon reconstruction and electron energy determination as examples. A full overview of the CMS detector's physics performance can be found in [298].

*Contribution to Muon Reconstruction*

A basic strength of the CM(uon)S detector is the reconstruction of muons, where the muon stations together with the tracker and calorimeter are combined in the energy flow concept. In the *global reconstruction* muon trajectories are extrapolated into the tracker volume, where the energy loss in the material plus multiple scattering is taken into account. The tracker measurements improve the transverse momentum resolution at lower energies. Figure 6.55 shows the global muon $p_T$ resolution.

A complementary algorithm to the *global reconstruction* is the *muon identification*. All tracks in the silicon tracker are quantified versus a muon compatibility hypothesis. All available information is taken into account including energy loss in the tracker and associated energy deposited in the calorimeters. This method allows the identification of low $p_T$ muons not reaching the outer muon layers. *Muon identification* adds about 10% to muon detection efficiency with respect to stand-alone *muon reconstruction*. The combination of both reconstruction algorithms enhances statistical samples of events containing multiple muons (e.g. $H \rightarrow ZZ^* \rightarrow \mu^+\mu^-\mu^+\mu^-$).

The tracker together with the calorimeters also significantly contributes to the muon isolation analysis for low $p_T$ muons. Muons originating from $b$, $c$, $K$ or $\pi$ decays are accompanied by particles from the associated jet. Muons from heavy objects like $W$, $Z$ are more isolated. If the deposited energy in a defined cone around the muon is lower than a certain threshold, the muon is considered isolated – the deposited energy is determined by transverse energy in the calorimeter or sum of transverse momenta of reconstructed particle tracks. The muon isolation (jet/no jet) helps to distinguish between muons of heavy objects or muons produced in jets.

*Contribution to Electron and Photon Reconstruction*

As for the muons, the tracker improves the ECAL selection of prompt electrons and photons by applying isolation criteria. To distinguish prompt photons and electrons from those produced in a jet, the summed energy deposited in a cone around the identified electron or photon must remain below a defined threshold. For an electron, in addition, the reconstructed energy in the ECAL must match the reconstructed track momentum of the tracker. The isolation algorithm improves background reduction for the channels $H \rightarrow ZZ^* \rightarrow e^+e^-e^+e^-$ and $H \rightarrow \gamma\gamma$, where isolated electrons or photons are expected.

The tracker also improves the precision on the electron energy measurement in a certain range, through the high precision $p_T$ measurement. Figure 6.56 shows the relative energy resolution for electrons measured by ECAL and tracker.

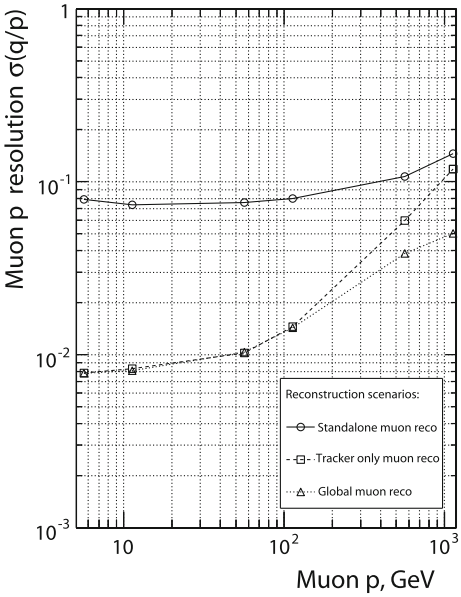



**Fig. 6.55** Global muon $p_\perp = p_T$ resolution. Up to energies of 1 TeV, the transverse momentum resolution of a muon is dominated by the tracker, while at higher energies, the multiple scattering term is less significant and the larger lever arm of the muon chambers improves resolution [298]

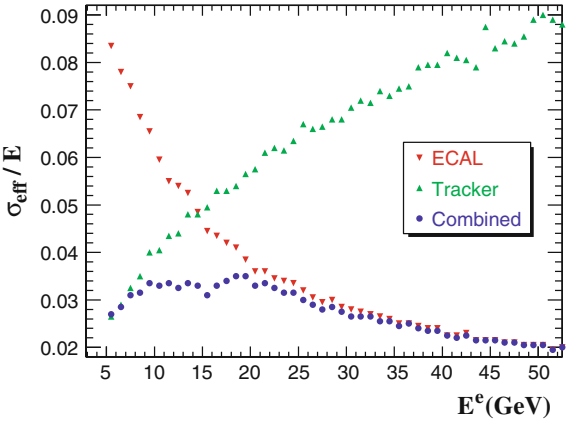



**Fig. 6.56** Electron energy measurement. The plot shows the relative electron energy resolution $\sigma_{\text{eff}}$ /E, as measured with the CMS tracker, ECAL and the combined value [298]

*Some Examples*

With introductory remarks some physics cases with significant CMS tracker impact are listed:

- the key component for the $B$ physics program is the ability of secondary vertex reconstruction also at HLT level with conditional tracking for processes like exclusive $B$-decays
  - $B_s \rightarrow \mu^-\mu^+$ and $B_s \rightarrow \mu^-\mu^+K^-K^+$
    flavour changing neutral current ($b \rightarrow s$ loop-level processes)
  - $B_s \rightarrow J\Psi\Phi$
    Gold-plated CP violation decay mode sensitive to new physics
- the LHC targeted Higgs discovery in the full mass range up to 1 TeV. Examining the expected decay products, one realizes immediately that lepton-$id$, $b$-, $\tau$-tagging and missing energy resolution are crucial. $b$-tagging is especially crucial to reduce background. The next paragraph with Fig. 6.57 and Fig. 6.58 gives some example event displays of the final discovery at 125 GeV and the tracker role
  - $ttH \rightarrow ttbb$: A very challenging topology where $b$ and lepton tagging is necessary plus a final full reconstruction of e.g. $bbbbqql\nu$
  - $H \rightarrow ZZ \rightarrow l^+l^-l^+l^-(l = e, \mu)$ tracker improves $\mu$ resolution significantly with respect to the muon chambers
  - $H \rightarrow \gamma\gamma$ tracker can help identify the vertex or reject $\pi^\circ$ (jet) but unfortunately for this analysis the high tracker material budget is more a hindrance
  - $H \rightarrow \tau\tau$ a very challenging channel with the $\tau$ decaying before reaching a detector element. All decay products need to be reconstructed. It decays in two thirds into hadrons plus a $\nu_\tau$.
- **M**inimal **S**uper **S**ymmetric **M**odel MSSM or **SU**per **SY**mmetry SUSY in general. $b$ and $\tau$-tagging play a key role as well as for full reconstruction as also for HLT
  - the neutral MSSM Higgs is predominantly produced via gluon fusion $gg \rightarrow bbH^0$. Besides $AH \rightarrow bb$ with an enormous background significant channels are $AH \rightarrow \tau\tau$, thus $\tau$-tagging is mandatory
  - charged MSSN Higgs production $gg \rightarrow tbH^+$ decays dominantly into $tb$ and $\tau\nu$ where again $t$, $b$- and $\tau$-tagging is important
  - SUSY signatures: Squark and gluino production leads to lepton(s) plus missing energy plus jets. Also here the tracker plays a key role

The CMS tracker plays an essential role to address the full range of physics which can be accessed at the LHC. Standard model physics like $b$ physics, $t$ physics, Standard model Higgs searches, MSSM Higgs searches and SUSY searches are planned. At 14 TeV jets are mostly expected but also the standard candles $W$, $Z$, $tt$. The full detector concept and especially its trigger flexibility must cover all eventualities, new physics possibly waits just around the corner.

*The Higgs Discovery a Success of the LHC, ATLAS and CMS*

At a CERN seminar on 4th July 2012, the ATLAS and CMS experiments at CERN's Large Hadron Collider announced the observation of a new particle in the mass region around 126 GeV consistent with the Higgs boson predicted by the Standard Model. The Higgs boson, as proposed within the Standard Model, is the simplest manifestation of the Brout–Englert–Higgs mechanism. On 8th October 2012, this discovery led to the Nobel prize in Physics for François Englert and Peter Higgs "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider." The full story is published in [311], the primary CMS Higgs publication, and in [285] where the first combined ATLAS/CMS result is shown. The two main discovery decay channels are shown exemplarily. The Figures 6.57 and 6.58 show the expected event displays of a simulated Higgs particle compared to a real events of a Higgs candidate followed by the statistics plots derived from simulation compared to the experimentally determined ones. Obviously the simulated ones show a wrong Higgs mass, since it was only measured with the discovery. The lower part of the figures show a 3D event display of a Higgs candidate.

At a first glance one could think that the $H \rightarrow \gamma\gamma$ events are fully reconstructed by the electromagnetic calorimeter or the $H \rightarrow \mu\mu$ channel by the muon system.

The role of the tracker in the $H \rightarrow \mu\mu$ case is an improved momentum and energy resolution. $\gamma$ identification is only possible with a combination of electromagnetic calorimeter and tracker. Especially in a dense high pile-up environment, a high granular tracker is important to make sure no track points to the $\gamma$; meaning the supposed $\gamma$ is not an electron.

These results and their precision underline the marvellous functionality of the CMS detector and how the different sub-detectors perform together.

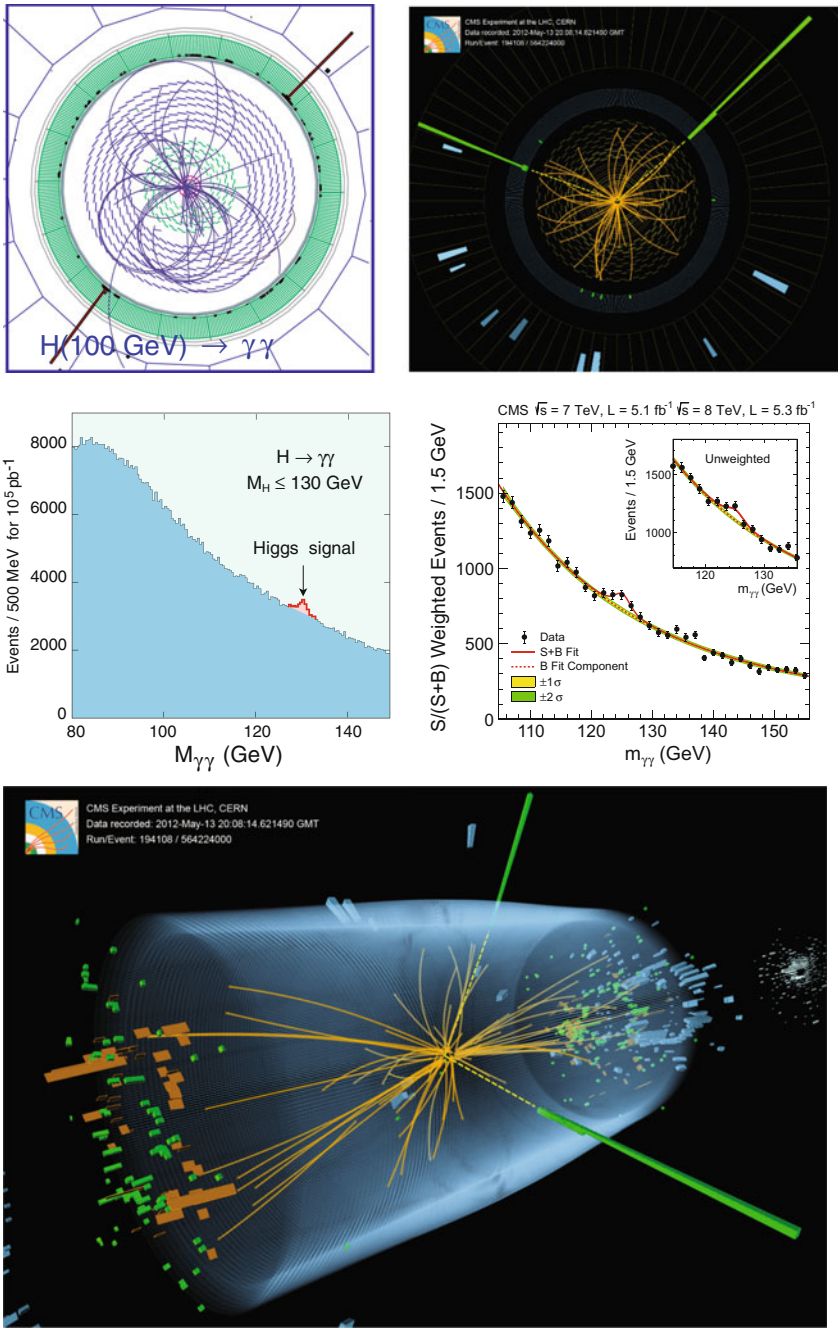



**Fig. 6.57** Higgs decaying into $\gamma\gamma$ [Courtesy of CMS/CERN] [291, 311]

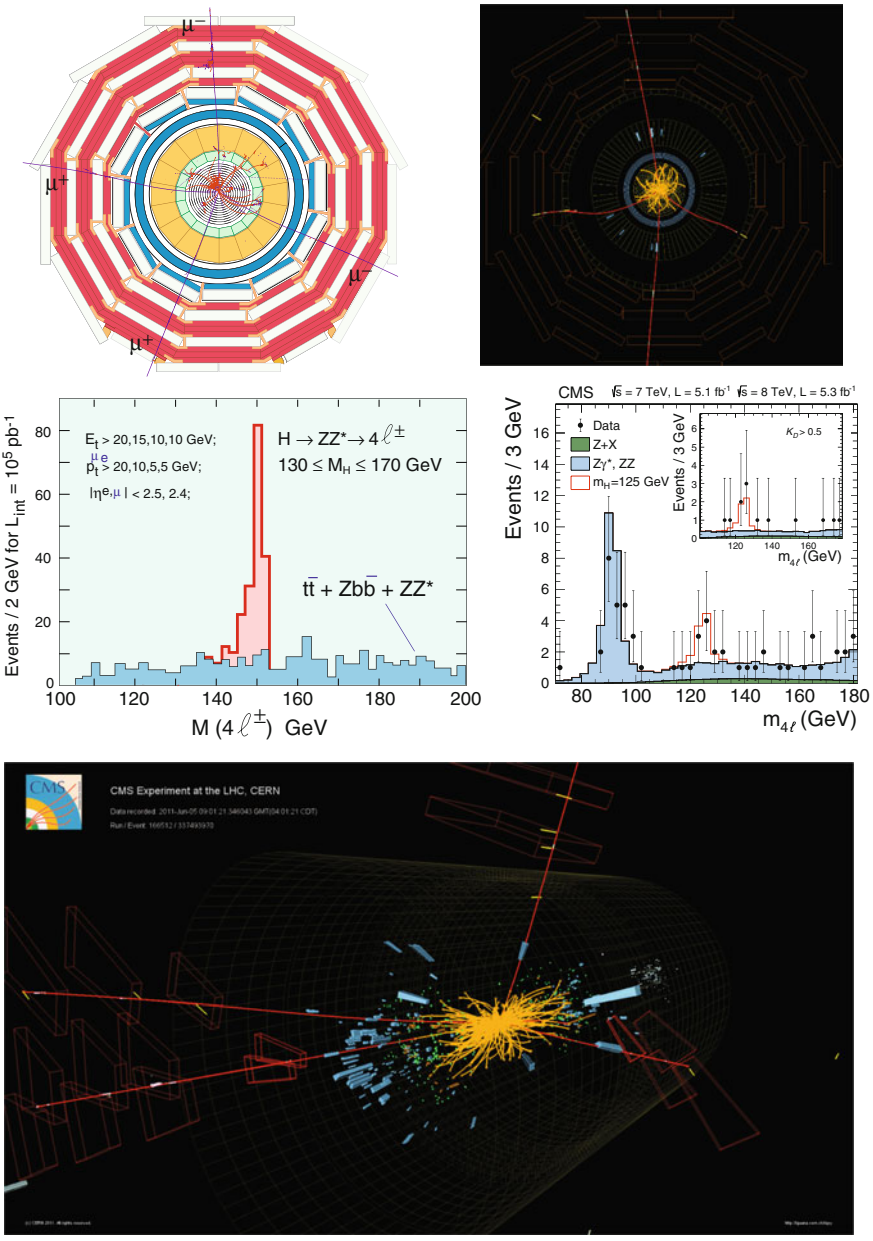



**Fig. 6.58** Higgs decaying into $ZZ \to \mu\mu$ [Courtesy of CMS/CERN] [291, 311]