Status of the CMS pixel detector

Viktor Veszpremi\textsuperscript{a,1,*}

\textsuperscript{a}Wigner Research Centre for Physics,
P.O.Box 49, 1525 Budapest, Hungary
E-mail: veszpremi.viktor@wigner.hu

The tracking detector of the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) is an all-silicon device. It is comprised of two sub-detectors. The pixel detector is the inner one, which is surrounded by the strip detector. The pixel detector provides seeds for charged particle tracking and measures the impact parameter of the reconstructed tracks. The impact parameter is essential in the reconstruction of primary interaction and secondary decay vertices. The pixel detector was upgraded in the beginning of 2017, during Run 2 of the LHC. Various interventions have been performed on the detector since then, the latest refurbishment taking place during Long Shutdown 2 (LS2) between 2019 and 2022, right after Run 2. The expected fluence in the innermost layer reaches the expected limit for the sensors after about 250 fb\textsuperscript{-1} of integrated luminosity; therefore, this layer was also scheduled to be replaced during LS2. In this paper, we describe the successful refurbishment and recommissioning program and the following relatively smooth data-taking period in the first year of Run 3. Preliminary studies of the performance will be presented along with the verification of the new layer 1 modules in which several weaknesses that were revealed during Run 2 have been fixed.

10th International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (Pixel2022)
12-16 December 2022
Santa Fe, New Mexico, USA

\textsuperscript{1}For the CMS collaboration
\textsuperscript{*}Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). https://pos.sissa.it/
1. Introduction

The tracking detector of the Compact Muon Solenoid (CMS) experiment [1] is an all-silicon device. Its modules are arranged in concentric cylinders around the interaction region of the LHC beams and are situated inside a 3.8 Tesla uniform magnetic field. The tracker is comprised of two sub-detectors with independent cooling, powering, and readout schemes. The inner sub-detector, the pixel detector, is segmented into 124 million strongly n-doped (n+) pixels of size 100 $\mu$m by 150 $\mu$m implanted into n-type bulk material with an active thickness of 285 $\mu$m and p-type back side. A reverse bias is applied across the sensor such that the implants collect the electrons created by the passage of charged particles. The pixel is surrounded by the silicon strip detector. The tracking detector provides measurement points in three dimensions along the curved trajectories of charged particles with high precision. The pixel detector provides seeds for an iterative, predominantly inside-out tracking algorithm and measures the impact parameter of the reconstructed tracks. The impact parameter is essential in the reconstruction of primary interaction and secondary decay vertices. The strip detector provides measurement points for the trajectory building and dominates the $p_T$ resolution of the reconstructed tracks.

The CMS experiment uses a right-handed coordinate system with the origin at the nominal collision point. The $x$-axis points towards the center of the LHC ring, the $y$-axis upward, perpendicularly to the LHC plane, and the $z$-axis along the counterclockwise beam direction. The polar angle $\theta$ is measured from the positive $z$-axis and the azimuthal angle $\phi$ from the positive $x$-axis in the $x$–$y$ plane. The radius, denoted by $r$, is the distance from the $z$-axis, and $\eta$ is the pseudo-rapidity computed from $\theta$.

The tracker used until 2016 measured charged particles within the range $|\eta| < 2.5$. For non-isolated particles of $1 \text{ GeV} < p_T < 10 \text{ GeV}$ and $|\eta| < 1.4$, the track parameter resolutions were typically $1.5\%$ in $p_T$ and $25$–$90$ ($45$–$150$) $\mu$m in the transverse (longitudinal) impact parameter [2]. At the start of 2017, a new pixel detector was installed [3]; the upgraded tracker measures charged particles up to $|\eta| < 3.0$ with typical resolutions of $1.5\%$ in $p_T$ and $20$–$75$ $\mu$m in the transverse impact parameter [4] for non-isolated particles of $1 \text{ GeV} < p_T < 10 \text{ GeV}$.

The upgraded pixel detector has four layers in the barrel region at radii of 2.9 cm, 6.8 cm, 10.9 cm, and 16 cm, and three disks spanning the radius between 4.5 cm and 16.1 cm on each side of the barrel, in the endcap or forward pixel regions, positioned along the $z$-axis at 32.4 cm, 39.9 cm, and 49.4 cm away from the origin of the CMS coordinate system. The pixels are read out in 52 by 80 pixel arrays by 18944 readout chips (ROC) in the barrel and 10752 ROCs in the endcap. An array of 8 by 2 ROCs form a module connected by high-density interconnect (HDI) flex printed circuit boards and are bump-bonded to the sensor. Data transmission from a module at 160 MHz is organized by the Token Bit Manager (TBM) chip into 8 channels in layer 1, 4 channels in layer 2, and 2 channels in layers 3 and 4. Two channels are merged by the TBM into a single signal and transmitted in a single line at 400 MHz to the back-end system [5]. Low voltage digital and analog power is provided by 1216 DC-DC power converter modules.

The local coordinate system of the modules is defined as follows: the axis pointing along the shorter edge of the module is denoted $x$, the one pointing along the longer edge as $y$. The $z$-axis is perpendicular to the sensor plane and is approximately parallel with the direction of the reverse bias. The orientation of the local $x$-axis in the barrel is fixed such that it always points in the direction of
the Lorentz-drift induced charge sharing. The direction of the local $y$-axis always points along the global $z$-axis of the CMS reference frame, i.e. in the direction of one of the beams and also in the direction of the magnetic field.

2. **Refurbishment during LS2**

The pixel was designed to be one of the last devices to be inserted during the construction of the CMS detector and to be removable within the time-frame of an extended technical stop or a long shut-down of the LHC. Various interventions have been performed on the detector since the start of the LHC operation. The latest one took place during the Long Shutdown 2 (LS2) between 2019 and 2022 following Run 2 which took place between 2015 and 2018. The expected fluence in the innermost layer is $3.6 \times 10^{15} \text{neq/cm}^2$ (for 1 MeV neutrons) after an integrated luminosity of 500 fb$^{-1}$. This is about twice of the expected limit for the sensors. Therefore, layer 1 was also scheduled to be replaced during LS2.

Along with exchanging the innermost layer, several refurbishment work items have been carried out, while the detector was kept dry and cold in the remaining time [6]. A few damaged modules were replaced in the inner face of layer 2 which became accessible after the removal of layer 1.

In the forward pixel detector, the CO2 cooling connections were consolidated. The power filter boards have also been replaced with a design that provides improved granularity for the high voltage (HV) distribution in order to better match the mapping of the low voltage distribution. Such an opportunistically executed redesign was useful because damages to ROCs that did not receive low voltage power while being connected to biased sensors were diagnosed to develop damages during Run 2 when several DC-DC converter modules became non-functional (see its effect on active detector fraction in Section 6.1).

All the DC-DC converters were replaced by modules equipped with an updated version of the FEAST chip that is protected against failure in the disabled output state after irradiation [7].

The HV power supplies were upgraded to be able to supply 800 V bias voltage albeit with a lower maximum current limit of 15 mA reduced from 20 mA.

3. **Reinsertion and positioning**

The pixel detector was reinserted in June 2021. Special care was taken in the positioning of the detector inside the support tube in order to center the innermost layer to the expected beam crossing region. Since the beginning of Run 1, CMS has been observing a slow upward drift in the beam spot position [8], due to a rising of the LHC magnets in the long straight section 5 (LSS5). The observed rate is in the order of 0.2 mm per year which was last confirmed by measurements of the beam spot between 2021 and 2022. This movement was corrected by LHC magnetically in 2017 and 2018 in two steps by a total amount of 1.8 mm. During LS2, the LSS5 was lowered vertically by 3 mm allowing for the removal of the magnetic correction. As a result of the careful positioning of the pixel detector during the reinsertion, a less than 0.2 mm beam offset has been achieved with respect to the barycenter of the barrel at the start of Run 3. However, a new horizontal offset was observed in the beam spot position by 1.1 mm, which lead to an uneven radiation damage in layer 1 during 2022, manifesting as a modulation in the leakage current as function of $\phi$ by more than
Figure 1: Distribution of the on-track cluster charge in the barrel pixel normalized to track impact angle before type inversion of layer 1 (left) [9] and the average of the normalized on-track cluster charge in the barrel pixel as function of the clock phase delay in the beginning of Run 3 (right) [11]. The most probable values (MPV) of the Landau distributions increase with the radius due to decreasing fluence except for layer 1, which has the largest MPV, as these modules have just been replaced with new ones.

±10% [9]. This misalignment is planned to be mitigated during the year end technical stop between 2022 and 2023.

4. Time alignment with the first collisions

The ROCs measure the charges generated by the passing of charged particles originating from the bunch crossing interactions of the LHC beams in 25 ns time slices. It is therefore important to properly adjust the clock phase of the internal clocks of the ROCs to the collisions. The available granularity of this adjustment is 0.5 ns. The optimal phase is found by scanning the 25 ns window in steps and measuring the efficiency and cluster properties of each control group. The cluster charge in such a scan is shown in Fig. 1 where the white part of the plot on the right represents a full clock-cycle.

The pixel detector is constructed using two ROC variants. The forward disks and the barrel layers 2 to 4 are built with the PSI46dig chip, which has a pixel hit detection efficiency better than 98% up to 150 MHz/cm\(^2\) hit rate. The barrel layer 1 is built with the PROC600 with efficiency better than 97% up to 580 MHz/cm\(^2\). In both chips, the pixels unit cells (PUC), responsible for the charge amplification and zero-suppression, are arranged in double-columns. The registered pixel hits, charge deposits in the sensor surpassing the comparator threshold, are transferred to the double column peripheries and get buffered along with a time-stamp for the duration of the trigger latency. Hits in the different ROC variants receive time-stamps with different delays. However, due to space constraints in the barrel supply tube, layer 1 and layer 2 modules share the same clock-tree. During Run 2, the clock phase was adjusted for the best performance of the layer 1 modules, while the layer 2 modules were operating still at high hit-finding efficiency but at suboptimal resolution. The resolution of the cluster position is determined by the small pixel charges in the shoulders of the clusters and is independent of the pixels in the cores of the clusters. Because smaller charges pass the comparator threshold later due to time-walk effects, the optimal clock phase setting is the one where the leading edge of the clock arrives to the ROC as late as possible compared to the
arrival of the particles from the collisions in the LHC such that all pixel hits are still assigned to the correct bunch crossing. The new layer 1 modules received an updated TBM chip with an extra feature to adjust the clock phase arriving to the module. Using this feature, the timing of layer 1 was successfully separated from that of layer 2 making it possible to optimize the timing for all layers individually. The plot in Fig. 1 on the right shows the results of a phase delay scan of the master clock after all the adjustments. The optimal setting corresponds to the 0 ns delay point.

5. Improvements in the new layer 1

As it was alluded in the previous section, the planned replacement was an opportunity to fix a few undesired features in the layer 1 modules that were discovered during Run 2.

The previous version of the PROC600 had a dynamic efficiency loss at very low and very high particle rates (figure 7 in [3]). In particular, at the highest luminosities reached in 2018, the cluster finding efficiency dropped by several per cents. The cluster finding efficiency is defined here as the fraction of globally reconstructed muon tracks crossing the mid-plane of the sensor for which a cluster is found in the vicinity of the crossing point. The efficiency has a gradual decrease as function of the instantaneous luminosity up to a certain luminosity value after which it has a sharp turn-down (such as in Fig. 47 in [3]). This value has recently been measured to be at around $1.6 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ in the 2018 data. The dominant source of inefficiency is the overflow of the hit buffers in the double-columns. The new version of the chip has better than 97% cluster finding efficiency up to $2 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ as shown in Fig. 2 on the right. The plot was made using data collected in a few typical fills after 13.8 fb$^{-1}$ of delivered integrated luminosity, where the efficiency loss connected to any loss of charges due to the sensor not being fully depleted or incomplete readout of the charge is expected to be minimal. The LHC delivered a very high luminosity test fill where the instantaneous luminosity reached beyond the range of the plot, and hints have been seen that the turn-down starts at around $2.1 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$.

A cross-talk effect between signal lines internal to the ROCs was also identified during Run 2 causing spurious extra pixel hits. As a consequence, the ROCs had to be tuned to higher thresholds, well above 3000 e$^-$. This problem has been fixed in the new chips, and a low in-time threshold around 2000 e$^-$ has been achieved as seen on the left in Fig. 2. The threshold is defined as the

---

**Figure 2:** In-time threshold measured by internal charge injection (left) and cluster finding efficiency as function of the instantaneous luminosity (right)[9].
amplitude of the charge injected internally into the PUCs for which the hit detection efficiency reaches 50\%, and is found by scanning a range of charge injection amplitudes. The unit of the injected charge is VCal, where one VCal corresponds to about 50 e\(^{-}\). It is possible to set lower absolute thresholds; however, due to the time-walk effects, low charges would be associated to the next bunch crossing. The in-time threshold describes the minimum pixel charge still falling into the correct bunch crossing. As a result of the improvement in hit efficiency and the lower threshold, a great improvement is seen in the lower tail of the cluster charge distribution in Fig. 1.

Beyond the new timing feature that was described in Section 4, two additional improvements have been introduced into the new TBM chips. The previous version was exposed to a certain type of single even upset (SEU) that blocked the data-transmission by the TBM to the back-end and caused it to be unresponsive to programming signals. The only possibility to bring the TBM out of such a “stuck” state was by power-cycling the module. Such an intervention was made difficult by the weakness of the DC-DC converter mentioned earlier. The affected detector modules were, and still are, power-cycled between LHC fills. The underlying reason for this vulnerability has been identified and corrected in the new chip design. In order to provide a second “safety net”, the TBM chips have received also a reset mechanism.

Finally, towards the very end of Run 2 a module was found which had developed a short in its HV bias line. As a precaution, the bias voltage in layer 1 was not raised beyond 450 V. The vulnerability was identified in the HDI design and fixed for the new modules. The new modules also received new cables. The HDIs have been subsequently tested up to 1 kV and the module cables up to 1.1 kV. At the end of 2022, layer 1 was operated up at 550 V for a short time without any problems.

6. Performance evaluation

6.1 Active detector fraction

Additionally to the few layer 2 modules that were replaced during the refurbishment, more modules have been recovered by replacing components on the supply tube, where the services are provided to the modules, and in the following commissioning period by adjustments in the calibration methods. The active detector fractions were 93.5\% in the barrel and 96.7\% in the forward pixel at the end of Run 2. In the first year of the Run 3 data-taking, these figures increased to 98.4\% in the barrel and 98\% in the forward pixel. The cluster occupancy is represented in Fig. 3 for layer 1 and for inner rings of the forward disks. The white vertical and horizontal areas in the center of the plots correspond to invalid module coordinates. The white areas in the colored part of the plot represent permanently bad ROCs. In addition, there are examples for ROCs, for example in module 1 in ladder 6 of layer 1, which have lower occupancy due to an SEU that occurred during the period when data was accumulated into the plot and became bad temporarily. The modules in the forward disks where the occupancy is significantly lower than in the modules with adjacent blade-coordinates are those that were damaged by not receiving low voltage while the sensors were under high voltage bias (see also in Section 2). These modules also produce more noise hits in general.

Channels reading temporarily bad modules are automatically masked during data-taking in response to frequent readout errors. When the number of masked channels passes a threshold, an
Figure 3: Cluster occupancy in layer 1 of the pixel barrel (left) and in the inner rings of the forward disks (right) [11].

Figure 4: Number of readout channels masked in the back-end in response to frequent readout errors in each layer and disk [9].

automatic mechanism attempts to recover them by reprogramming the modules. The number of masked channels is seen in Fig. 4 as a function of the time that passed since the start of the run. Not all modules are successfully recovered. The number of those increases over time in layers 2 to 4 and in the forward disks. These are predominantly due to the TBMs getting “stuck”. These modules are recovered between two LHC fills. In layer 1, instead, the fraction of masked channels decreases with the decaying luminosity, as the TBMs are protected against the same SEU mechanisms.

6.2 Radiation damage effects

Radiation damage leads to charge trapping and change in the effective doping of the sensors, among other changes. Both affect the charge collected in the clusters. In order to verify the performance of the new layer 1, compare its behavior to the previous version, and systematically study the radiation effects, the regular collision data and dedicated high voltage bias scan runs are
periodically evaluated. We use these results also in deciding when to change the bias voltage for physics data-taking.

The charge collected in clusters after a certain amount of integrated delivered luminosity is plotted as a function of the production depth within the sensor in Fig. 5 for layer 1 and 2. The origin of the plot corresponds to the implant side where the electrons are collected. The grazing angle method [10] is followed in the measurement, where the production depth is computed from the displacement of the center of the pixels in long clusters with respect to the impact point of the track associated to the cluster and the impact angle of the track. The measurement also corrects for the Lorentz-drift of the charges. The measured average pixel charge is meant to have an arbitrary unit, because the gain calibration between the measured ADC values and the real charge deposits varies in time. It is possible, however, to infer from the plot the effect of countering charge deposits by raising the bias voltage at a given point in time, and also to identify when full depletion is no longer achieved. Equivalent information is used to derive the Lorentz-angle and serves as input for the production of the charge distribution templates from which the position of the clusters is determined.

Figure 5 shows that the 150 V bias voltage that was applied in layer 1 after 10.9 fb\(^{-1}\) was not sufficient to fully deplete the sensors, but the charge collection efficiency was nearly optimal after raising it to 300 V. A further increase to 400 V was applied in layer 1 towards the end of the 2022 data-taking, not shown in the plot. A change in the HV set point for layer 2 would have been due had the data-taking lasted for longer. A slight annealing effect in layer 2 is also seen between the measurements at 10.9 fb\(^{-1}\) and 11.2 fb\(^{-1}\), where there was a technical stop and the detector was brought to room temperature.

Figures 6 and 7 show the results of the bias scans that were performed in layer 1 in a single HV power group. A HV group contains three modules at three different z-positions (out of the four possible positions) in layer 1. Such “mini” scans are performed in all detector layers and disks without compromising data-quality as they cover small and non-overlapping solid angles and thus can be done almost transparently during data-taking. In the dedicated bias scan runs, performed less frequently, the bias voltage in the full layers is varied.
The hit efficiency plot demonstrates that nearly full efficiency is reached at a lower bias voltage than nearly maximal cluster charge – or full depletion voltage. The full depletion voltage can be determined from the cluster charge plot or from the plot of the cluster size in the local $y$-direction where no Lorentz-drift is present. The effect of the Lorentz-drift, or more precisely the decrease of the Lorentz-angle as function of the bias voltage, is seen in the cluster size along the local $x$-direction. The bias voltage is chosen in order to maximize hit efficiency and expected resolution. The transverse resolution, that corresponds to the $x$-direction in the local frame, is expected to be the most optimal when the clusters are two pixel wide on average.

### 6.3 Cluster properties and resolution

The cluster size in both local $x$- and $y$-direction is regularly monitored and evaluated as seen for the barrel pixel in Fig. 8 as a function of the delivered integrated luminosity. The initial large variation of the cluster size took place in a short period of time that covers the type inversion. It was not possible to follow the rapid change with the tuning of the bias voltage in such a compressed time scale. The bias voltage was first increased in the first technical stop of the LHC. Later the change slows down as the fraction of the integrated luminosity newly delivered by the LHC becomes small compared to the already integrated luminosity, making the cluster properties more uniform.
Figure 8: The average size of on-track clusters in the barrel measured along the local $x$-coordinate, in the direction of the the Lorentz-drift, (left) and along the local $y$-coordinate, in the direction of the pseudo-rapidity (right) as functions of the integrated luminosity delivered by the LHC in 2022 [9].

Figure 9: The dispersion of the hit-triplet residual distribution in the $r - \phi$ direction, in the transverse plane, (left) and in the global, longitudinal, $z$-direction (right). The quantities are proportional to the intrinsic resolutions. [9].

The size in the $x$-direction is determined by the Lorentz-angle. By comparing layer 1 to the other layers, the excellent performance of the new layer 1 is confirmed. It is also clear from Fig. 7 that the cluster size in $x$ will remain below two pixels in the future and the main goal of changing the bias voltage will be to optimize the charge collection efficiency. Even further on, when the bias voltage reaches the hardware limit, and the sensor will be operated “under-depleted”, the only way to achieve better resolution will be by updating the description of the cluster charge distribution in the offline reconstruction.

The size in the $y$-direction strongly depends on the radius of the layer, purely because of geometrical reasons. The inner layers have larger $\eta$ coverage. The size in $y$ determines the longitudinal resolution of the pixel detector, in the global $z$-axis. Its variation is largely similar to what is seen in the $x$-direction, but the resulting variation in the resolution is more significant.

Figure 9 shows the dispersion of the hit-triplet residuals in the local $x$- and $y$-directions, which is proportional to the intrinsic resolution of the detector, as a function of the delivered integrated luminosity. The residuals are computed by selecting high quality, high-$p_T$ tracks with hits on the layer under consideration and on the two adjacent layers. The two adjacent hits are refitted analytically with a helix using the $p_T$ of the original track, and the distance between the hit in the considered layer and the intersection of the propagated trajectory with the layer is computed. In the computation for layer 1, the hits from layers 2 and 3 are used. Since there is no fix point between the
primary interaction point and the layer 1 hit, the residuals in layer 1 are more weakly constrained.

The intrinsic transverse resolution of the module design is, therefore, best determined when measuring it in layer 2 or 3, and found to be better than 9.5 µm, demonstrating the beneficial effects of the charge sharing. The results for layer 1 were compared to the previous version of the layer and found to be significantly better, as expected from the lower threshold tuning made possible by the improvements in the chip design. Although, layer 2 has not been changed, its performance has also improved because of the better time-alignment that was made possible by the new timing functionality in the TBMs.

The two colors in Fig. 9 correspond to two different methods to compute the cluster position. The template algorithm is based on the detailed cluster shape simulation as it was mentioned before. The computation of the cluster position is time consuming, therefore this method is only used in offline track reconstruction during the final refit. The high-level trigger reconstruction and the trajectory building step in the offline tracking uses the generic algorithm. This algorithm determines the position of the cluster using the pixel charges in the shoulders of the clusters and the assumption of the Lorentz-angle. The template algorithm performs better as long as the charge distribution is well modelled, while the generic algorithm is more robust in terms of its input parameters. This is also demonstrated by the smaller difference at the measurement points approaching 11 fb$^{-1}$, when the sensors in layer 1 became under-depleted.

6.4 Hit efficiency monitoring

The hit efficiency was also monitored throughout the first year of Run 3. The analysis excludes known temporarily or permanently bad ROCs. The main source of inefficiency in the functioning ROCs is due to the hit buffering in the double-columns as mentioned in Section 5. The dependence of the efficiency on the instantaneous luminosity was also confirmed earlier to have improved according to expectations. The gradual degradation of the efficiency approaching 11 fb$^{-1}$ of integrated luminosity and the bulk of the sudden increase at 11 fb$^{-1}$, however, cannot be due to the ROCs as the instantaneous luminosity did not vary much. Therefore, it must be caused by the charge collection loss in the sensors due to gradually running under-depleted. After raising the bias voltage to 300 V and recovering the charge collection efficiency as shown in Fig. 5, the remaining inefficiency consequently is purely determined by the ROCs. The measurement in Fig. 10 was carried out right after this time.

7. Summary

The CMS pixel detector has undergone a successful refurbishment and recommissioning program followed by a relatively smooth data-taking period. The refurbishment involved the replacement of the innermost layer by an entirely new one with updated front-end chips. The detector received new DC-DC converters for the low voltage powering and upgraded power supplies for the high voltage bias. The cooling and services have been consolidated in the forward parts of the detector. Preliminary studies have been presented on the performance. The active fraction of the entire detector remains high. The new layer 1 modules have fulfilled our best hopes, several weaknesses that were revealed during Run 2 have been fixed. The radiation damage seen in layer 1 is largely compatible with expectations drawn from the earlier version. The lower threshold in layer 1
Status of the CMS pixel detector

Viktor Veszpremi

Figure 10: The cluster finding efficiency in the barrel layers as function of the integrated luminosity delivered by the LHC in 2022 [9].

and the higher maximum bias voltage will provide and extended period for high-quality data-taking. Studies to determine the compatibility of this period with the expected duration of Run 3 are being updated.

Acknowledgments

This work was supported by the Hungarian Scientific Research Fund under contract number OTKA K124850.

References


12
