Optimization of a 65 nm CMOS imaging process for monolithic CMOS sensors for high energy physics


*CERN, Esplanade des Particules 1, CH-1211 Geneva 23, Switzerland
†Tower Partners Semiconductor Co. Ltd., Uozu city, Toyama, Japan
‡Tower Semiconductor Ltd., Migdal Haemek, Israel
§Etesian Semiconductor Ltd., Ramat Yishai, Israel
¶University and INFN Bari, Bari, Italy
‖University and INFN Catania, Catania, Italy
¶¶University and INFN Cagliari, Cagliari, Italy
††University and INFN Torino, Torino, Italy
‡‡University and INFN Trieste, Trieste, Italy
§§University of Salerno and Bologna, Italy
∗∗∗Speaker and corresponding author: W. Snoeys, CERN, Geneva, Switzerland, walter.snoeys@cern.ch
The long term goal of the CERN Experimental Physics Department R&D on monolithic sensors is the development of sub-100nm CMOS sensors for high energy physics. The first technology selected is the TPSCo 65nm CMOS imaging technology. A first submission MLR1 included several small test chips with sensor and circuit prototypes and transistor test structures. One of the main questions to be addressed was how to optimize the sensor in the presence of significant in-pixel circuitry. In this paper this optimization is described as well as the experimental results from the MLR1 run confirming its effectiveness. A second submission investigating wafer-scale stitching has just been completed. This work has been carried out in strong synergy with the ITS3 upgrade of the ALICE experiment.
1. State of the art

![Figure 1](image1.png)

**Figure 1:** (a) A schematic view of the ITS2 detector [1, 2] with about 12.5 Gpixels over about 10 m$^2$, (b) a picture during its installation in the ALICE experiment, and pictures of (c) inner and (d) outer half-barrels during construction, the latter covering about 2 m$^2$.

![Figure 2](image2.png)

**Figure 2:** Events taken by the ITS2 detector in the ALICE experiment. On the right all seven layers are clearly visible, the three inner layers are to be replaced for the ITS3 upgrade (see below).

CMOS monolithic active pixel sensors (MAPS) revolutionized visible imaging, and now they are making their way into high energy physics (HEP). Most detectors in HEP are still implemented in hybrid form with a readout chip and a silicon sensor chip, but monolithic sensors integrating readout and sensor in the same chip offer easier detector assembly, lower material budget, and lower cost. The largest system to date with monolithic CMOS sensors is the ALICE ITS2 detector covering a 10 m$^2$ area with 3 x 1.5 cm$^2$ ALPIDE CMOS sensor chips for about 12.5 billion pixels [1, 2]. The ALPIDE sensor [2, 3] contains full CMOS readout circuitry in the pixel, a pixel front end [4] with amplifier and discriminator and was the first monolithic sensor with a sparse readout
as in hybrid sensors. The sensor development was carried out from 2012 to 2016, the detector was installed in 2021 and is now taking data. Figure 1 shows pictures of the detector assembly, the ALPIDE chip, and the detector installed in the experiment. Figure 2 shows some events taken in the experiment.

Figure 3: MLR1 submission in the TPSCo 65 nm ISC technology: (a) Layout and (b) picture of one reticle field on the wafer containing 70 test chips of 1.5 mm x 1.5 mm or 3 mm x 1.5 mm (55 different ones, 15 chips printed twice), (c) a diced quarter wafer on tape, and (d) transistor test chips in a gelpack.

The success of the ALPIDE and the ITS2 raised interest by several other HEP experiments. Also in ALICE a further upgrade is planned with the ITS3 [5–7], aimed to further reduce material budget by replacing the inner layers with wafer-scale stitched MAPS sensors, thinned to a few tens of microns to make them flexible and bend them around the beam pipe[8]. R&D work started at CERN in collaboration with several other institutes [9] to develop MAPS sensors in sub-100nm technologies, also with specific interest for the ALICE ITS3 upgrade, not only for better density and performance but also because 300 mm wafers give access to large stitched sensors. A first technology identified is the TPSCo 65 nm ISC imaging CMOS technology, and a first Multi Layer per Reticle run (MLR1) was completed (see Figure 3). It included sensor optimization based on general principles as already applied and summarized below in the 180 nm CMOS imaging technology before going into detail on the 65 nm technology.

2. Sensor optimization in the TowerJazz 180nm CMOS technology

Figure 4: ALPIDE pixel cross-section (left), and die picture (right)[3].

Figure 4 shows a cross-section of the ALPIDE pixel and a die picture of the ALPIDE chip. The charge collection electrode in the center is formed by an nwell diffusion. The in-pixel circuitry is placed in a deep pwell, which shields nwells within the circuitry from the epitaxial layer preventing
them to collect charge. The depletion layer is ‘balloon’-shaped, it extends from the junction at the
collection electrode, and does not reach the pixel edges. Charge is still collected from outside the
depletion layer by diffusion, but this collection is relatively slow and subject to charge trapping after
non-ionizing radiation. The dependence of the extension of the depletion layer on the reverse bias
is smaller for a spherical junction (as an approximation for the junction at the collection electrode)
than for a one-sided planar junction:

\[-V = \frac{qN_A}{2\varepsilon} \Delta^2 \quad (a) \]
\[\Delta = \frac{2\varepsilon|V|}{qN_A} \quad (b)\]

\[-V = \frac{qN_A}{3\varepsilon} \left[ \frac{3R_2^3}{2} + \frac{R_2^2}{R_1} + \frac{R_2^2}{2} \right] \quad (c) \]
\[R_2 = \frac{2\varepsilon|V|}{3R_1} \quad (d)\]

Figure 5: (left) Depleted volume versus epitaxial layer thickness for two different epitaxial resistivities.
(right) Lateral electric field (red and blue for opposite sign) illustrating the extension of the depletion layer
for the 10 \( \mu \)m and 25 \( \mu \)m points in the left plot. For the thin 10 \( \mu \)m epitaxial layer (top) increasing the
resistivity from 500 (a) to 5000 \( \Omega \)cm (b) does not significantly increase the depletion volume, for the thicker
25 \( \mu \)m epitaxial layer (bottom, c and d) it does.

Formulas (a) and (b) show that for an abrupt planar junction the depletion layer width \( \Delta \) is
proportional to the square root of the reverse bias, or more precisely the sum of the externally applied
reverse bias and the built-in voltage. Formula (c) shows reverse bias as a function of inner and
outer radius \( R_1 \) and \( R_2 \) of the depletion layer for an abrupt spherical junction. \( R_1 \) is approximately
equal to the radius of the collection electrode, small for a small collection electrode. In case \( R_2 \) is
significantly larger than \( R_1 \), the relation simplifies to formula (d), illustrating a much weaker
dependence of the depletion on the reverse bias with a cubic root instead of a square root and a term
equal to 1.5 times \( R_1 \) under the root. After the depletion reaches the substrate or the deep pwells,
further extension of the depletion is significantly reduced as both can provide significant depletion
charge over a small volume to balance the additional reverse bias. This is further illustrated in
Figure 5 (a), which shows the simulated depleted volume as a function of epitaxial layer thickness
for two epitaxial resistivity values for an example pixel of 15 \( \mu \)m pitch. For low thickness a higher
resistivity does not provide an advantage, but for higher thicknesses it does. Figure 5 (right), shows
the simulated lateral electric field in a sensor, significantly different from zero only in the depletion layer (red and blue represent the opposite sign). This further illustrates that a higher resistivity results in a larger depleted volume only for a thicker epitaxial layer.

Figure 6: (a) Introduction of a blanket deep low dose n-type implant creating a planar junction to achieve full depletion of the epitaxial layer including the low dose implant [10](b) Detection efficiency degradation in the pixel corners after 5 x 10^{14} 1 MeV n_{eq}/cm^2[13].

As a first step in sensor optimization, the 180 nm TowerJazz process was modified in a side development of the ALPIDE sensor [10]: a deep blanket low dose n-type implant was introduced (Figure 6) to create a planar junction deep in the sensor separated from the collection electrode. This achieves full depletion of the epitaxial layer with some reverse bias, and signal charge collection by drift, without any layout changes in the sensor design except the definition of the region of the implant over the pixel matrix, hence allowing the same designs to be processed in the standard and the modified process for comparison. The process modification improved the tolerance to NIEL from 10^{13} 1 MeV n_{eq}/cm^2 by an order of magnitude.

Figure 7: (a) Addition of an additional deep pwell or (b) a gap in the low dose n-type implant to increase the lateral field near the pixel edges. This strongly accelerates the signal charge towards the collection electrode [14] and fixed the detection efficiency loss in the corners [15].

This early result triggered a development in this modified process for the outer pixel layer of the ITk upgrade in ATLAS, also funded by the European Commission’s Horizon 2020 Programme under contract number 675587, the STREAM project, with the design of two pixel sensors, the MALTA [11] and TJ MONOPIX [12], with very similar pixel sensor layout but different readout architectures. However, beam tests on irradiated devices from a first iteration showed significant detection efficiency degradation in the pixel corners after 5 x 10^{14} 1 MeV n_{eq}/cm^2 due to a very low lateral field near the pixel edges allowing charge to be prone to trapping by radiation induced defects[13].
To improve the lateral field near the pixel edges, an additional deep pwell implant or a gap in the deep low dose n-type implant was introduced (Figure 7) [14]. This accelerates the charge collection for charges generated near the pixel edges and yielded the desired improvement for fluencies up to $10^{15}$ 1 MeV n$_{eq}$/cm$^2$[15] with an additional improvement from a higher gain and better uniformity of the pixel front end [16]. After this heavy irradiation the devices needed cooling to -20 $^\circ$C to reduce the radiation-induced leakage and the noise.

3. Applying the same principles in 65 nm.

To optimize the sensor in the presence of significant in-pixel circuitry and also include standard digital gates in the pixel matrix, the same principles were applied to the 65 nm technology. As discussed before, without modifications the thinner epitaxial layer (10 $\mu$m instead of 25 $\mu$m in 180 nm) limits the extension of the depletion, and made these modifications more needed and more beneficial. A deep nwell was added to the pixel and its dose adjusted to form the low dose n-type implant. Also the deep pwell was modified to improve the isolation between the sensor and the circuitry, eliminate punch-through between the deep nwell implant and the circuitry, and prevent local potential wells retaining the signal charge. Four process splits were done, gradually moving from the standard process to the most optimized:

Split 1: Standard process without modifications.
Split 2: First modification of the deep pwell to improve isolation between circuitry and sensor and prevent punch-through between deep nwell and circuitry.
Split 3: Adding to split 2 the deep nwell adjustment in the pixel to allow full depletion.
Split 4: Adding to split 3 an additional deep pwell modification to prevent potential wells created by the additional in-pixel circuitry.

![Figure 8](attachment:image.png)

**Figure 8:** Simulated detection efficiency for minimum ionizing particles as a function of charge threshold for different process and pixel design variants for a pixel pitch of 15 $\mu$m. The second deep pwell modification, the difference between split 3 and split 4, eliminates the efficiency loss by potential wells introduced by the additional in-pixel circuitry for the modified pixel design with blanket low dose n-type implant and brings its performance for the process of split 4 to what is expected.

Three different pixel designs were implemented for all splits: the standard pixel design without deep n-well implant, the modified with a blanket deep nwell implant, and the modified with a gap...
Optimization of a 65 nm CMOS imaging process for monolithic CMOS sensors

W. Snoeys

in the deep nwell implant near the pixel border. No version with an additional deep pwell was implemented. Figure 8 shows the simulated detection efficiency for minimum ionizing particles as a function of the charge threshold for different process and pixel design variants. The most probable value for a MIP in 10 μm is about 600 electrons with an onset of the Landau distribution well below. If charge sharing is more pronounced, the seed pixel collects less charge and detection efficiency degrades for lower charge thresholds. The figure shows that, contrary to the standard pixel design, the pixel variant with gap and all process modifications remains efficient for higher thresholds as it concentrates the charge much more on a single pixel. The pixel variant with a blanket low dose n-type implant exhibits a very low lateral field and is very sensitive to potential wells retaining signal charge. Figure 8 compares the detection efficiency for this pixel variant for process split 3 and split 4, and illustrates the effect of the second pwell modification to eliminate the impact of potential wells introduced by the additional in-pixel circuitry causing charge loss. After this split 4 gives the expected result that the detection efficiency is worst for higher thresholds in the standard pixel design without low dose n-type implant, intermediate for the modified pixel with blanket low dose n-type implant, and best for the modified pixel with gap in the implant due to the impact of the pixel design on the charge sharing.

Figure 8: Figure 8 shows the simulated detection efficiency for minimum ionizing particles as a function of the charge threshold for different process and pixel design variants. The most probable value for a MIP in 10 μm is about 600 electrons with an onset of the Landau distribution well below. If charge sharing is more pronounced, the seed pixel collects less charge and detection efficiency degrades for lower charge thresholds. The figure shows that, contrary to the standard pixel design, the pixel variant with gap and all process modifications remains efficient for higher thresholds as it concentrates the charge much more on a single pixel. The pixel variant with a blanket low dose n-type implant exhibits a very low lateral field and is very sensitive to potential wells retaining signal charge. Figure 8 compares the detection efficiency for this pixel variant for process split 3 and split 4, and illustrates the effect of the second pwell modification to eliminate the impact of potential wells introduced by the additional in-pixel circuitry causing charge loss. After this split 4 gives the expected result that the detection efficiency is worst for higher thresholds in the standard pixel design without low dose n-type implant, intermediate for the modified pixel with blanket low dose n-type implant, and best for the modified pixel with gap in the implant due to the impact of the pixel design on the charge sharing.

Figure 9: Simulated detection efficiency for minimum ionizing particles for the three different pixel designs, standard without low dose n-type implant, modified with blanket low dose n-type implant, and gap with a gap in the low-dose n-type implant, for several pixel pitches for the process with all modifications (as split 4). Only the pixel variant with gap maintains good efficiency at larger pixel pitches.

Figure 9 shows the simulated detection efficiency for minimum ionizing particles for the three different pixel variants and different pixel pitches for the process with all modifications (as split 4). Charge sharing reduces the amount of signal charge on a single pixel and reduces efficiency for larger thresholds and larger pixel pitches. The increased lateral field generated by the gap in the low dose n-type implant is essential to reduce the charge sharing and concentrate an amount of signal charge on a single pixel sufficient to remain efficient for larger pixel pitches.
4. Experimental results.

Figure 10: Different sensor test chips implemented on the first MLR1 run (a) the analog pixel test structure (APTS) [17] (b) the digital pixel test structure (DPTS) [18] and (c) the CE65[19, 20].

Figure 10 shows different pixel sensor test chips implemented in the first MLR1. All of them measure 1.5 mm by 1.5 mm. The analog pixel test structure or APTS [17] features a 6 by 6 pixel matrix, of which the central 4 by 4 pixels are connected via an analog buffer chain to the output. There are two variants, in the source follower version the readout chain consists of 4 source follower stages, in the opamp version the output stage is based on an operational amplifier. The digital pixel test structure or DPTS [18] contains a 32 by 32 pixel matrix at 15 μm pitch with an amplifier and discriminator in each pixel and a time encoded digital readout. The CE65 contains a pixel matrix with analog readout in a rolling shutter configuration [19, 20].

Figure 11: Spectrum of the charge collected on the seed pixel in different variants of the APTS with source follower under irradiation by a radioactive $^{55}$Fe source[21].

Figure 11 shows the spectrum of the charge collected on the seed pixel from a $^{55}$Fe radioactive source for different pixel variants and process splits. For the standard variant without low dose n-type implant and hence only a somewhat different deep pwell, the sensor behaves very similarly for the standard process (split 1) and for the process with all modifications (split 4): the iron peaks are still visible but rather low, and a very large charge sharing peak is visible at much lower signal amplitudes. The situation is very different for split 4 with the low dose n-type implant with and
without gap: both variants suppress charge sharing significantly, the version with gap even more, practically eliminating the charge sharing peak and resulting in a much more pronounced iron peak.

Figure 12: Spectrum of the charge collected on the seed pixel under irradiation by a radioactive $^{55}$Fe source for various pitches for the standard pixel variant without deep low dose n-type implant in the APTS with source follower. Charge sharing increases with pixel pitch resulting in a larger charge sharing peak and less pronounced iron peaks [21].

Figure 13: Spectrum of the charge collected on the seed pixel under irradiation by a radioactive $^{55}$Fe source for various pitches in the APTS with source follower for the pixel variant with low dose n-type implant with gap [21].

The impact of the process and pixel design optimization is further illustrated by measuring the spectrum of the charge collected by the seed pixel for different pixel pitches for the different variants under iron source exposure. Figure 12 shows this for the standard pixel variant without deep low dose n-type implant in the process of split 4. Charge sharing increases with increasing pixel pitch illustrated by the larger charge sharing peak and even less pronounced iron peaks. The contrast with the pixel variant with the deep low dose n-type implant with gap shown in Figure 13 is striking: charge sharing is suppressed with very pronounced $^{55}$Fe peaks. For this pixel variant
the iron peaks are practically identical indicating that charge sharing is not significantly affected for the various pixel pitches.

Figure 14: Signal fall time versus signal amplitude under $^{55}$Fe irradiation measured on the APTS with opamp for different cluster sizes, comparing the standard pixel variant without deep low dose n-type implant (left) and the modified variant with gap in this n-type implant (right).

Figure 14 compares signal fall time versus signal amplitude for the standard pixel variant without low dose n-type implant and the modified variant with gap in this implant for various cluster sizes. For the standard variant fall time and amplitude are rather broadly distributed. For the variant with gap the amplitude is concentrated at the $^{55}$Fe peak, and the fall time well below the ns. This also illustrates the impact of the process and pixel design modifications on the charge collection. For the variant with gap sensor timing studies are in progress. In 180 nm time resolutions better than 150 ps were achieved [22], there are first indications that this may be improved in 65 nm.

Figure 15 shows the detection efficiency and fake hit rate as a function of the charge threshold for the digital pixel test structure or DPTS with the pixel variant with low dose n-type implant with gap in the optimized process split (split 4). The efficiency is over 99 % and increasing the reverse bias brings the fake hit rate down to the detection limit, broadening the window of operation. The chip is irradiated up to 10 kGy and $10^{13}$ 1 MeV $n_{eq}$/cm$^2$, the requirement for the ITS3 upgrade. In fact, the same variant of the DPTS proved full detection efficiency in a test beam a few weeks after diced chips became available, validating the circuitry and the sensor with process modifications. In addition full efficiency was also reached after $10^{15}$ 1 MeV $n_{eq}$/cm$^2$, at room temperature[21]. Higher fluences are under investigation, but cooling will be necessary to reach full efficiency. Tolerance to ionizing radiation is in line with other deep submicron CMOS technologies[23, 24], and single event upset cross-sections are according to expectations.

MAPS do not necessarily require dedicated imaging technologies, but this section illustrates some flexibility on the foundry side has been very helpful to significantly influence the charge
collection in the sensor and enhance sensor performance. The modifications discussed here also improve sensor performance and operating margin before irradiation by concentrating more charge on a single pixel, this was experimentally confirmed also for the 180 nm in [25].

After these encouraging results, a new submission has just been completed with two stitched sensors, MOSS [26, 27] and MOST, and about 50 test chips, with as main goals to learn about stitching and continue to learn about the technology [28]. The two stitched chips have a different approach to powering and yield: the MOSS has fully separated powered sub-matrices [26, 27] to be powered down in case of defects, the MOST has a global power network, conservatively designed, to which small subcircuits are connected using switches. This allows a higher granularity in the powering, and subcircuits to be designed at full density. Figure 16 shows wafer and reticle.
5. Conclusions

To develop sub-100 nm CMOS monolithic sensors, also towards the ALICE ITS3 upgrade, the TPSCo 65 nm ISC CMOS imaging technology was selected as a technology candidate. A first run, MLR1, was submitted with some splits on the process and various pixel sensor designs to optimize the sensor in the presence of complex readout circuitry including digital gates. The optimization is carried out following general principles already applied to the TowerJazz 180 nm imaging technology. Measurements now validate the effectiveness of this optimization with a fully efficient sensor, analog front end and digital readout chain in a 15 \( \mu \text{m} \) by 15 \( \mu \text{m} \) pixel in the DPTS, and measurements on other test chips clearly illustrate the impact of the optimization on the charge collection. All these results now help to qualify this technology for use in high energy physics. A stitched engineering run has been submitted to learn about stitching and to continue to learn about the technology.

References


[14] M. Munker et al., "Simulations of CMOS pixel sensors with a small collection electrode, improved for a faster charge collection and increased radiation tolerance", *2019 JINST 14 C05013*, [https://doi.org/10.1088/1748-0221/14/05/C05013](https://doi.org/10.1088/1748-0221/14/05/C05013).


