Nuclear Instruments and Methods in Physics Research A 731 (2013) 125–130

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

# Monolithic pixel detectors for high energy physics

W. Snoeys *

*CERN, CH-1211, Geneva 23, Switzerland*

ARTICLE INFO

Available online 28 May 2013

*Keywords:*
Particle detectors
Pixel detectors
Monolithic detectors
Tracking detectors

ABSTRACT

Monolithic pixel detectors integrating sensor matrix and readout in one piece of silicon have revolutionized imaging for consumer applications, but despite years of research they have not yet been widely adopted for high energy physics. Two major requirements for this application, radiation tolerance and low power consumption, require charge collection by drift for the most extreme radiation levels and an optimization of the collected signal charge over input capacitance ratio ($Q/C$). It is shown that monolithic detectors can achieve $Q/C$ for low analog power consumption and even carryout the promise to practically eliminate analog power consumption, but combining sufficient $Q/C$, collection by drift, and integration of readout circuitry within the pixel remains a challenge. An overview is given of different approaches to address this challenge, with possible advantages and disadvantages.

Published by Elsevier B.V.

## 1. Monolithic detectors for high energy physics

Monolithic detectors integrating sensor matrix and readout in one piece of silicon are standard for the detection of visible light. In high energy physics they offer advantages in terms of detector assembly, production cost and detector capacitance, and are in principle promising not only for pixel detectors but also for full tracking detectors. However, despite these advantages the default in high energy physics is still the hybrid approach where sensor matrix and readout electronics are implemented in two separate pieces of silicon, connected by either flip-chip bump bonding or by wire bonding.

Detectors for high energy physics have somewhat different requirements from those of traditional CMOS imagers for visible light. For visible light the signal is generated within a depth of a few microns. In physics single high energy particle traversals have to be detected, instantaneously generating charge over the full thickness of the silicon (about 80 electrons per micron traversed). It is therefore not important in physics if the signal generated very near ($\sim$1 μm) the surface is not collected, while for visible light a very significant fraction of the signal would be lost. The hit pixels (typically maximum a few percent) receive their charge in a very short time with respect to the readout time. Therefore the integration time can be limited to eliminate or reduce the effect of slow phenomena like dark current, which is a very critical parameter in imagers for visible light.

Several functional monolithic detectors on high resistivity silicon were developed for high energy physics but often required exotic fabrication steps (e.g. double sided processing [1]), incompatible with high volume manufacturing in standard semiconductor foundries. Recently devices have been manufactured in more standard CMOS technologies, but preserving the low capacitance and efficiency over the full surface for more complex readout circuitry has remained a challenge, especially if collection by drift is required. This is also the case for Monolithic Active Pixel Sensors (MAPS) [2,3]: they often use a "rolling shutter" architecture with very simple in-pixel circuitry, but sometimes contain more complex circuitry [4]. These MAPS often collect the signal charge by diffusion over a very limited depth, which makes them sensitive to radiation damage and results in relative slow signal collection as explained in Section 2.

So far monolithic pixel detectors have only been adopted in two high energy physics experiments: the Depleted P-Channel Field Effect Transistor (DEPFET [5]) pixels in Belle-II [6] and MAPS in STAR [2]. In both cases the readout is relatively slow (row by row), which is not always applicable. Monolithic detectors are not yet installed in the Large Hadron Collider or LHC, but they are considered for upgrades and for detectors in new accelerators like the Compact Linear Collider (CLIC) and the International Linear Collider (ILC).

To preserve good data quality, material in the inner layers has to be minimized to limit the probability for the particles to be scattered. This has really been a challenge: for all LHC experiments power consumption of these inner layers is so high that the copper cables to bring in the power and the cooling to extract it represent very substantial fractions of the total material budget. This is true especially for the zone between barrel and endcaps where cooling pipes and cables cause an undesirable peak of 1.5 to 2 radiation lengths, for both CMS [7] and ATLAS [8], in the moderately forward regions. Power consumption has therefore severely impacted the

* Tel.: +41 22767 6639.
*E-mail address:* walter.snoeys@cern.ch

0168-9002/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.nima.2013.05.073

amount of material in current experiments. Physicists would like to reduce this for future detector upgrades, but they would also like better performing detectors, capable of disentangling more complex events with more functionality. This is equivalent to pushing for lower power consumption and enhanced performance at the same time.

Clearly there is not much margin on power consumption: as an example, the ~100 million channels in the CMS silicon tracker consume about 33 kW in total (for 2.5 and 1.25 V supplies for the chips), but a similar amount is consumed in the cables providing the current [9].

There are different approaches to reduce material: reduce power consumption per channel, reduce the material per detector layer, use part of the mechanical structure to bring in the power, or use special powering schemes to reduce the copper cross-section: DC–DC converters, powering several modules in series. Monolithic detectors can reduce the material per layer, and as will be shown provide prospects to lower the power consumption.

The main challenges for pixel and tracking detectors in high energy physics are radiation tolerance and power consumption. The key parameters influencing these are respectively the charge collection mechanism (drift versus diffusion) and the collected charge over the input capacitance ratio. These will be discussed in the next paragraphs.

## 2. Radiation tolerance and charge collection: drift versus diffusion

Two mechanisms exist to collect charge generated in silicon by ionizing radiation: diffusion and drift. In the case of diffusion there is no strong force acting on the charge carriers: they "bounce around" and ultimately get collected on a charge collection electrode. This process takes time (up to several hundred ns) and therefore increases the chance for the signal charge to be captured by defects or traps in the material and be lost for readout. In the case of collection by drift, an electric field pushes the charge carriers to their destination, which reduces the time to collect the charge by at least an order of magnitude. This strongly reduces the chance for signal charge to be captured by defects or traps. Since the main component of radiation damage in the sensor is the creation of defects or traps, better radiation tolerance is achieved if signal charge is collected by drift. In traditional MAPS charge is collected primarily by diffusion, and radiation tolerance therefore it does not exceed fluences of $10^{12}$–$10^{13}$ neutrons/cm$^2$, often already with significant performance degradation at lower fluences. Recently some MAPS devices with a higher radiation tolerance have been reported with a higher resistivity epitaxial layer for which the drift component in the charge collection is more important [2].

Silicon sensors traditionally used in high-energy physics collect signal charge primarily by drift from a depletion layer of several hundred microns deep in a few nanoseconds. They show tolerance to fluences of $10^{15}$ neutrons/cm$^2$. To obtain a depletion layer of this thickness it requires a high resistivity substrate: the higher the resistivity the lower is the voltage which has to be applied. For detector upgrades at the LHC where fluences of $10^{16}$ neutron equiv./cm$^2$ are envisaged for the inner layers, the radiation induced leakage current and the bias required to still collect some charge out of the detector becomes so large that the detector power consumption becomes non-negligible compared to the consumption of the readout electronics. Therefore other schemes have been studied like 3D detectors [10] in which the depletion is generated laterally using pillar-shaped electrodes reaching very deep into the silicon substrate. This reduces the operating voltage and power consumption by at least an order of magnitude, but increases the detector capacitance.

The very high fluences of $10^{16}$ neutron equiv./cm$^2$ will impose the design of future monolithic detectors collecting signal charge by drift. Moreover, the severe ionizing radiation environment (up to several hundred of Mrad), typically affecting the CMOS readout circuitry more than the sensor, will push toward smaller linewidth technologies. It has been established that the use of thinner oxides in deep submicron CMOS technologies drastically improves their radiation tolerance [11,12].

Apart from signal size, timing and radiation tolerance, an often neglected difference between charge collection by drift and diffusion is the much larger cluster size for collection by diffusion [13]. Cluster size has a very significant effect on device performance: dividing the charge over more than a few pixels reduces the effective signal-to-noise and all the parameters depending on it like for instance position resolution. Having to deal with information from a larger number of pixels for a single particle hit also has a direct impact on the architecture: more memory space has to be provided to store the hit information, or dedicated cluster processing circuitry has to be foreseen to combine and reduce data from the different pixels in the cluster before storage.

## 3. Power consumption and the importance of the $Q/C$ ratio

As mentioned before low power consumption is the key to low mass detectors in high energy physics experiments. In the inner layers power consumption cannot easily be increased even for enhanced performance. At the LHC silicon trackers consume about 20 mW/cm$^2$ and pixel detectors consume several hundred mW/cm$^2$. The contributors are the power consumption of the analog circuits, the digital processing circuitry on chip, the power needed to transmit the data to the outside, and the power consumption in the detector. The analog power consumption is often determined by the required signal-to-noise ratio for a given bandwidth. The amplitude of the signal in this context can be expressed as a voltage: the ratio of the collected signal charge $Q$ divided by the equivalent pixel input capacitance $C$ or $Q/C$. To obtain the signal-to-noise ratio, the $Q/C$ ratio has to be compared to the noise expressed as an equivalent series voltage at the pixel input. In the following this is further explored and leads to the possibility of optimizing the $Q/C$ ratio to reduce analog power consumption. Afterward, the other contributors to the power consumption are briefly discussed.

In a simplified model, the noise of a MOS transistor can be represented by an equivalent voltage source in series with the gate contact, or by an equivalent current source in parallel with the transistor. Depending on whether the transistor is in weak or in strong inversion, the frequency spectrum of this voltage source equals one of the two following expressions [14]:

$$dv_{eq}^2 = \left(\frac{K_F}{WLC_{ox}^2 f^{\alpha}} + \frac{2kTn}{g_m}\right) df \quad \text{for weak inversion}$$

$$dv_{eq}^2 = \left(\frac{K_F}{WLC_{ox}^2 f^{\alpha}} + \frac{4kT\gamma}{g_m}\right) df \quad \text{for strong inversion}$$

The first term in each of the two expressions corresponds to a term decreasing with frequency, and is often referred to as $1/f$ noise, although the exponent $\alpha$ is not always exactly equal to one ($K_F$ is a proportionality constant, $W$ and $L$ are transistor width and length respectively, $C_{ox}$ is the gate oxide capacitance per unit area, and $f$ is the frequency). The second term represents the thermal noise and the power consumption is related to this noise via the transconductance $g_m$ of the transistor ($k$ is Boltzmann's constant,