Electronics for pixel detectors.

M. Campbell
CERN, 1211 Geneva 23, Switzerland
Michael.Campbell@cern.ch

Abstract
Most modern HEP experiments use pixel detectors for vertex finding because these detectors provide clean and unambiguous position information even in a high multiplicity environment. At LHC three of the four main experiments will use pixel vertex detectors. There is also a strong development effort in the US centred around the proposed BTeV experiment. The chips being developed for these detectors will be discussed giving particular attention to the architectural choices of the various groups. Radiation tolerant deep sub-micron CMOS is used in most cases. In light of predicted developments in the semiconductor industry and bearing in mind some fundamental limits it is possible to foresee the trends in pixel detector design for future experiments.

I. INTRODUCTION

Pixel detectors are now important components in modern High Energy Physics (HEP) experiments. The initial development work for SSC and LHC was first applied in a heavy ion experiment [1] and in LEP [2]. This provided a basis for confidence for use in the future p-p experiments at LHC [3, 4, 5], the Alice heavy ion experiment [6] as well as to the proposed BTeV experiment at the Tevatron [7]. An extensive overview of the history of the development of pixel detectors for HEP is given in [8]. The present paper takes a closer look at the design of the electronics for the different present-day systems. Some basic concepts relevant to pixel electronics are reviewed. Then the work in progress for the new experiments is discussed with particular reference to the chosen readout architectures. Finally, an attempt is made to explore the way ahead for such detectors in future.

II. DEFINITIONS AND BASIC CONCEPTS

A pixel detector is a 2-dimensional matrix of microscopic (<< 1 mm) sensitive elements each of which is connected to its own pulse processing readout electronics. The detecting elements may be in the same substrate as the electronics (monolithic) or on a separate substrate and bump-bonded to the electronics (hybrid). While the monolithic pixel detector has the advantage that bump-bonding is not required, the hybrid detector allows the separate optimisation of the detector material and the electronics circuit.

The purpose of the detector is to provide unambiguous coordinates in two dimensions along with precise time stamp information with a very high signal to noise ratio. Pixel detectors are most useful in environments where the particle multiplicity in an event excludes, or makes difficult, the use of projective detectors. This explains why pixel detectors were first adopted in heavy ion experiments.

By nature pixel detectors have many channels and normally a threshold in charge is used to record the position of impinging particles. Therefore, one of the most important expressions governing pixel detector design is the general expression for noise hit rate, \( f_n \), in a single channel binary system [9]:

\[
f_n = \frac{1}{\sqrt{3}} \sqrt{f_b} \exp\left(\frac{-Q_{th}^2}{2\sigma_n^2}\right)
\]

where \( f_b \) is the bandwidth of the system, \( Q_{th} \) is the threshold in electrons and \( \sigma_n \) is the noise in rms electrons. In pixel detector systems \( f_n \) has to be multiplied by the number of channels to obtain the total number of noisy hits in the system. As pixel detectors are mostly used for pattern recognition in complicated events it is important to keep a large separation between threshold and noise in such systems in order to optimise use of the readout bandwidth. It can be shown that threshold variation adds quadratically to the noise and must also therefore be minimised. Fig. 1 (taken from [10]) shows the Landau distribution of charge deposited in a 300\( \mu \)m planar Si detector along with the noise and threshold in a typical readout system. In pixel detectors the charge deposited is shared between a number of pixels and experience has shown that full detection efficiency is obtained by placing the threshold at around one third of the most probable peak.

![Fig 1 The Landau distribution from a planar Si detector 300\( \mu \)m thick (top), the threshold distribution (middle) and noise (bottom), taken from [10]. Note the logarithmic y-axes.](image-url)
compromise between the required timing resolution and noise. The equations for series, ENC_s, and parallel noise, ENC_p, can be simplified to:
\[ ENC_s^2 \propto \frac{C_s^2}{g_m \tau_s} \]
\[ ENC_p^2 \propto I_o \tau_s \]
where \( C_s \) is the total capacitance on the input node of one channel, \( g_m \) is the transconductance of the input transistor, \( \tau_s \) is the shaping time and \( I_o \) is the leakage current of the sensor element. Other parallel noise sources have to be added to \( I_o \). In modern pixel detector systems the parallel noise is much lower than the series noise due to a fast shaping time and the small leakage current (fA-pA) from the tiny sensor volume. The rise time, \( t_r \), of the front end amplifier is given by:
\[ t_r \propto \frac{C_f (C_L + C_f)}{g_m C_f} \]
where \( C_f \) is the load on the output of the front-end amplifier and \( C_s \) is the amplitude of the feedback capacitance which is inversely proportional to the voltage gain of the system. From these expressions one can observe that \( C_s \) and \( C_f \) should be minimised and \( g_m \) maximised to obtain high speed. \( C_s \) may indeed be minimised by careful detector design, \( C_f \) by reducing the load on the preamplifier output, but maximising \( g_m \) implies increased power consumption.

III. PIXEL DETECTOR SYSTEMS FOR HEP

There are 4 major developments for HEP experiments underway at present. The Atlas and CMS vertex detectors are aimed towards the high event rate p-p experiments at LHC. The common Alice/LHCb development has two quite different aims: the rather low event rate but very high multiplicity Alice vertex detector and the LHCb RICH detector readout. At the proposed BTeV experiment in Fermilab the chip should provide information for the first level trigger. The intention in what follows is to highlight the similarities and differences between them.

A. Atlas and CMS

Atlas and CMS have very similar environments. Both vertex detectors must withstand neutron fluences of up to \( 10^{15} \text{n}_\text{f/cm}^2 \) and each has to provide unambiguous 2-dimensional hit information every 25 ns. Both have the usual requirements of minimal power consumption and material. As the total neutron fluence is well above the value where the n-type bulk material behaves as p-type, n⁰ on n detectors are used. Atlas uses a p-spray as isolation between pixels [12] and CMS uses two concentric broken p⁺ guard rings [13]. Both experiments expect to operate the detectors not fully depleted near the end of the lifetime because of reverse annealing. Therefore the most probable peak energy deposited by particles will be strongly attenuated. There may also be charge collection time issues. As mentioned above a threshold must be set in each pixel to keep the data volume from the detector manageable and this should be around one third of the most probable peak. This implies that the required minimum operating threshold be around 2 - 2.5 keV. The Atlas group has decided for rectangular pixels (50 µm x 400 µm) giving optimum spatial resolution in r-φ while CMS planned to have square pixels (at present 150 µm x 150 µm) making use of the Lorentz angle of the charge drift in the sensor produced by the 4 T magnetic field at CMS. Both experiments elected to make first full scale prototype chips in the DMILL technology [14] and this has had a strong influence on the choice of layout and readout architecture.

The Atlas chip is organised as a matrix of 18 x 160 pixels. The layout of the cell is such that the columns are grouped as pairs enabling two columns to use the same readout bus. This means that the layout is flipped from one column to its neighbour, a practice which is not normally recommended. Each pixel comprises a preamplifier-shaper (\( t_{peak} = 27 \text{ ns} \)) with a transistor in feedback which is biased by a diode connected transistor connected to the preamplifier output followed by a discriminator, see Fig. 2 taken from [15]. There is a 3-bit register which permits threshold adjustment pixel-by-pixel and two further bits which control masking and testing operations. The power consumption per channel is expected to be 40 µW. As the feedback is a constant current in the linear range of the amplifier the discriminator provides Time-over-Threshold (ToT) information.

When a pixel is hit, it sends a Fast-OR signal to the End of Column (EoC) logic. A token is sent up the column pair and the first hit pixel encountered puts its address on the bus along with a timestamp. It does the same for the trailing edge of the hit. The token then drops to the next hit pixel. There is an 16-cell deep hit buffer at the bottom of the column which stores the addresses, timestamps and ToT values for the hit pixels. There is also logic at the End of Column (EoC) which compares the timestamps of the hits with the external trigger input. Extra peripheral logic is used to package the hit information into a serial stream for readout. Although results on test chips were encouraging [16] the first results from the bump-bonded full-scale prototypes were disappointing. Some of the problems were traced to yield issues associated with the very high component density of the design. A new full-scale prototype is under development using deep sub-micron technology.

The CMS chip is organised as a matrix of 53 x 52 pixels. Also in this case alternating columns are mirrored and grouped as pairs, two columns sharing the same readout bus. The front-end is a classic preamplifier-shaper circuit (\( t_{peak} = 27 \text{ ns} \)). The schematic of the front-end is shown in Fig. 3 taken from [17]. The buffered output of the shaper is sent to a discriminator which has a 3-bit trim DAC. At the
output of the discriminator a pulse stretching circuit is used to produce a signal which sample-and-holds the analog value of the buffered shaper output. The pulse width is tuned to sample the analog signal on or near to the peak. The power consumption of one pixel is around 40 \( \mu \text{W} \).

The largest prototype to date has 18 x 160 pixels and the pixel cell size is 50 \( \mu \text{m} \times 400 \mu \text{m} \). Each cell has a preamplifier followed by a shaper \( t_{\text{peak}} = 150 \text{ ns} \). The feedback of the preamplifier has two branches: a very low bandwidth feedback amplifier which drives a current source at the input which compensates for detector leakage current, and a simple NMOS transistor which provides the fast return to zero, see Fig. 4 taken from [20]. This dual feedback system was needed as the W/L ratio of the enclosed gate NMOS cannot be made lower than 2.3 [21]. Another unique feature of the pixel cell is the 3-bit ADC which has been implemented at the output of the shaper.

When a pixel is hit it pulls down a Fast-OR signal which indicates a hit to the EoC logic. The EoC logic notes the timestamp and sends a token up the column. When a hit pixel receives the token it sends its address and the contents of the ADC to the EoC. There is some core logic which packages the timestamp, address and amplitude information for immediate readout. In the case of BTeV this information is used in the generation of the first level trigger.

**B BTeV**

There is one large-scale development underway inside the HEP community but outside the LHC programme. This is the FNAL pixel development which aims towards the proposed BTeV experiment but which might also be useful for the upgrades of the other experiments at the Tevatron. The radiation environment at the BTeV experiment is similar to the LHC but the bunch crossing interval is 132 ns instead of 25 ns. \( n \) on \( n \) detectors will be used but the decision about \( p \)-spray or \( p \)-stop isolation is pending. The group chose to design the chip in deep sub-micron CMOS following the radiation tolerant design techniques used by the RD49 Collaboration [19]. Interestingly, the FNAL team developed a common rules file which allows them produce a design which is compatible with two different 0.25 \( \mu \text{m} \) processes.

When a pixel is hit it pulls down a Fast-OR signal which indicates a hit to the EoC logic. The EoC logic notes the timestamp and sends a token up the column. When a hit pixel receives the token it sends its address and the contents of the ADC to the EoC. There is some core logic which packages the timestamp, address and amplitude information for immediate readout. In the case of BTeV this information is used in the generation of the first level trigger.

**C Alice/LHCb**

The chip which has been developed for the Alice tracker and the LHCb RICH readout is probably the most complicated of the readout chips to date containing over 13 million transistors. Neither experiment expects a radiation dose even close to those of Atlas and CMS. Straightforward \( p \) on \( n \) detectors were chosen as these are cheaper and easier to obtain. The chip has a matrix of 256 x 32 cells and each cell measures 50 \( \mu \text{m} \times 425 \mu \text{m} \). The preamplifier is differential which should improve the power supply rejection ratio and limit the substrate induced noise at the expense of increased power consumption. As a fast return to zero was required for the LHCb application a new front-end architecture was used which uses the preamplifier along with two shaping stages, see Fig. 5 taken from [22]. The circuit is ready to accept another hit after 150 ns. The output of the second shaper is connected to a discriminator with 3-bits of threshold adjust. There are two readout modes of operation: Alice and LHCb.
In Alice mode, if the discriminator fires one of the two registers stores the timestamp in the block marked delay. The contents of the 8-bit timestamp bus are ramped up and down with a modulo determined by the Level 1 trigger latency. When there is a coincidence with a Level 1 trigger and a hit resulting from the positive comparison of one of the register contents with the timestamp, a 1 is put in a 4-bit FIFO. A Level 2 trigger accept triggers the transfer of the FIFO information to a 256-bit shift register for readout.

In LHCb mode, the outputs of the discriminators of 8 pixels are ORed together and the 16 registers are used sequentially to store timestamps. On Level 0 trigger the event is stored in a 16-bit FIFO. A Level 1 trigger initiates the readout of the matrix. 32 clock cycles are needed for full readout. A full description of the readout system is given in [23]. Results from this full scale prototype chip are presented in [24].

D General remarks
All of the above developments are either aiming at, or have achieved, noise levels of less than 200 e⁻ rms and a threshold uniformity at or below that level. The requirements in timing and power consumption are similar also. It is interesting to note that the developments using DMILL have chosen architectures which require < 100 transistors per pixel while the other developments use many hundreds of transistors with the attendant increase in functionality and pixel complexity.

A common issue however, is that the tracking precision of all of these detectors is not limited by pixel dimension but by material thickness. Thinner detectors and electronics are desirable but these are fragile. A main contributor to material is the cooling systems which are required to evacuate the heat dissipated by the electronics. At around 100 µW/pixel the total power consumption is around 0.5 W/cm². This gives some hints about where future technical efforts should focus.

IV. FUTURE TRENDS

Vertex detector development is now and forever intimately linked to developments in the electronics industry. Even the CERN pixel developments follow Moore's law of exponentially increasing component density with time [25]. Experience with the LHC pixel detectors has taught us that we must use as much as possible industry standard technology to be able to achieve our aims within a tolerable price range.

New particle accelerators beyond LHC and the Tevatron are being discussed. It may be that technical limits in the detectors should now be given serious consideration in the earliest stages of design of new machines. One of the major issues here is the problem of cooling and its influence on material and hence tracking precision. As the push towards smaller pixels continues the power consumption density (for the same time precision) increases. This is because the input capacitance is dominated by pixel-to-pixel capacitance and the total pixel-to-pixel capacitance per unit area increases with granularity. Increasing the bunch crossing frequency would lead to a further increase in power density.

Radiation tolerance of future CMOS technologies seems to be obtainable using the design techniques already mentioned. However, it will be necessary to monitor carefully each generation of technologies for unexpected phenomena. In all cases Single Event Upset will probably present the main challenge to designers.

There seem to be two kinds of machine emerging each having a distinctive physics reach and environment: the large linear electron-positron colliders, and the next generation of hadron colliders. The e-p machines will provide events which are essentially clean with low multiplicity and with event rates in the range of kHz. For these applications it may still be possible to use projective detectors or pixel-type detectors where every pixel is read out. CCD and APS sensors seem the most likely candidates here. Both detectors provide the very highest spatial resolution but the readout tends to be relatively slow. APS sensors based on standard CMOS technologies are being studied [26].

For the hadron machines the pattern recognition capability of pixel detectors is likely to still be the dominant requirement. The very tiny charge collection from standard APS detectors makes achieving good pattern recognition extremely difficult. An interesting modification to the APS idea is presented in [27]. In this development cooling is used to obtain a larger charge collection in the substrate of the standard CMOS components. Also an interesting detector biasing scheme is proposed. Cryogenic CMOS is discussed later. However, developing cryogenic mixed-mode electronics on top of the sensor volume remains a formidable challenge.

There are some developments which could be applied in many future pixel systems. MCM-D is a technology which is of great interest although it has only been studied so far by the Atlas pixel community [28] who had both the courage and the resources to investigate it. Here the detector is used as the substrate and the MCM layers, which are alternately BCB and metal, are grown on top. In this way each pixel is connected through the layers to its bump bond and all of the readout lines and power supply lines can be brought in on the same substrate. As high dielectric constant insulating layers may be used power supply de-coupling can be provided as well. This technology offers great promise as it leads to a reduction in the overall mass of the detector while offering good
mechanical rigidity and at the same time the possibility to map sensor elements to readout channels of different dimensions.

Some detector development work based at Stanford/Hawaii [29] uses reactive ion etching to make high aspect ratio holes in Si. These are subsequently filled with doped Si to make a detector which is made up of pillars of alternating p' and n' doping. This has the advantage that the voltage needed to deplete the detector is much reduced and may be of particular interest in detectors where radiation damage causes inverse annealing to take place. Of particular interest to the pixel community in general would be the possibility this technology might offer in reducing strongly the dead area which surrounds present pixel detectors for guard rings. One could imagine that the same technique is used to etch a very clean cut near the edge of the sensitive pixel matrix. This edge may have doped Si applied and this would limit the electric field laterally. Also the clean edge from the etch might reduce surface leakage currents.

Cryogenic operation of the readout electronics has the advantage that a better transconductance to drain current ratio is obtained even if the transistor thresholds are increased [30]. This may lead to the possibility of reducing the problem of power consumption density. In any case cryogenically cooled Si detectors, with or without defect engineering [30, 31], will probably be used in future experiments. There is a whole new field of mixed-mode, cryogenic deep sub-micron design opening up.

V. SUMMARY AND CONCLUSIONS

Pixel detectors are key components in most new large scale HEP experiments. The developments for the LHC experiments are well under way with most groups now focussing on deep sub micron CMOS. All of the present systems are limited in physics performance by materials budget which is strongly correlated to power consumption density and the subsequent cooling systems. Expected increases in granularity and speed in future systems will require very careful system optimisation. Novel circuit architectures should be developed and some technology advances may also help.

VI. ACKNOWLEDGEMENTS

Many people assisted me in preparing this manuscript and the associated talk. I am particularly indebted to Kevin Einsweiler of Atlas, Roland Horisberger of CMS and Abd al Mekkaoui and Ray Yarema of BTeV for providing me with material. I receive constant help and encouragement from my friends and colleagues in the CERN Microelectronics Group, the Alice and LHCb RICH pixel teams and the Medipix Collaboration. Dave Barney and Mike Lamont advised me on machine matters. Erik Heijne and Cinzia Da Via` contributed greatly with discussions about present and future detectors.

VII. REFERENCES

[7] Proposal for an Experiment to Measure Mixing, CP Violation and Rare Decays in Charm and Beauty Particle Decays at the FermiLab Collider - BTeV, FermiLab, May 2000.


