Nuclear Instruments and Methods in Physics Research A305 (1991) 527–532
North-Holland

Section II. Circuit design

Pixel detectors with local intelligence: an IC designer point of view

F. Krummenacher [1]

Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland, and
Smart Silicon Systems, CH-1012 Lausanne, Switzerland

Scaling laws for the analog front end and related problems are discussed for detectors in the range from microstrips to pixel detectors. Design strategies for fast- and low-power building blocks (charge-sensitive preamplifier, shaper, discriminator and analog storage) are looked into. Merging of functions for minimal transistor counts, local analog storage versus digital-only output (trade-offs and limitations) and precision of and matching between readout elements are also discussed.

1. Introduction

True two-dimensional detector arrays, commonly designated as pixel detectors, are likely to play an increasingly important role in several fields of science where position-sensitive detection of charged ionizing particles of X-rays is required. This paper focuses on the circuit design aspects of direct readout for pixel detectors including local data processing such as amplification, shaping, discrimination and local data storage.

2. From microstrips to pixel detectors

In order to evaluate the scaling laws for the analog front end, we first consider the typical circuit depicted in fig. 1. The charge sensitive preamplifier characteristics may be formulated as follows:

- feedback factor

$$h_{fb} = \frac{C_f}{C_d + C_i + C_f},$$

- total load capacitance

$$C_L = C_o + h_{fb}(C_d + C_f),$$

- integrating time constant

$$\tau_a = \frac{C_L}{h_{fb} g_m}$$

- charge-to-voltage gain

$$A = \frac{1}{C_f}.$$

where C_d is the detector capacitance, C_i the equivalent input capacitance of the amplifier, C_f the feedback capacitor and C_o the capacitance associated with the amplifier output.

Assuming that the series thermal noise is dominant (this is usually the case for short shaping time constants), the equivalent input noise charge can be written as

$$\mathrm{ENC}^2 = \gamma k T \frac{(C_d + C_i)^2}{g_m \tau_s},$$

where γ and g_m represent the noise factor and the transconductance of the preamplifier, respectively, and τ_s is the shaping time constant.

Suppose now that the segmentation of both the detector and the electronic is increased by a factor n. In a first-order approximation, this yields for each channel:

- for the detector

$$C_d \to \frac{C_d}{n},$$

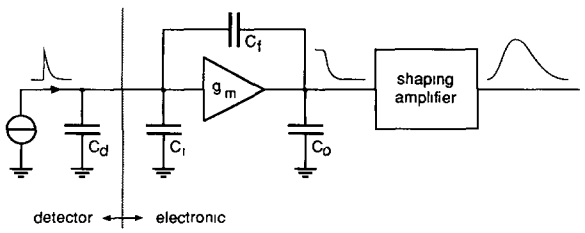

Fig. 1. Typical analog front end for direct readout of microstrip detectors.

[1] Present address: Electronics Lab., LEG/EL Ecublens, CH-1015 Lausanne, Switzerland.

0168-9002/91/$03.50

– for the electronic

$$C_i \rightarrow \frac{C_i}{n},$$

$$C_f \rightarrow \frac{C_f}{n},$$

$$C_o \rightarrow \frac{C_o}{n},$$

$$g_m \rightarrow \frac{g_m}{n}.$$

As a result, both the power dissipation and the speed are not modified but the gain is improved by n;

$$A \rightarrow An$$

and the noise power is reduced by n also;

$$\mathrm{ENC}^2 \rightarrow \frac{\mathrm{ENC}^2}{n}$$

It can therefore be concluded that in addition to improving the spatial resolution, increasing the segmentation yields a higher sensitivity (gain) and a lower noise for the same power dissipation or equivalently a lower power consumption for the same noise performance. This means that detector segmentation is a key design issue for low-noise/low-power front end.

There are, however, several practical limitations to segmentation:

1) The capacitance between adjacent detector elements can be larger than the capacitance of one element to ground when the pitch becomes comparable to or smaller than the detector thickness; in such a case C_d no longer scales as $1/n$.

2) The stray capacitance associated with the connection of the detector to the electronic (wire bonding in the case of microstrips, bump bonding for pixels) is relatively independent of the segmentation and results in an artificial increase of the equivalent input capacitance of the preamplifier, which means that above a certain limit (depending on the bonding technique), C_i does not scale as $1/n$ either.

3) Increasing the segmentation means increasing the number of separate data paths sharing a common supply, clocks and biasing circuitry, resulting in an increasing importance of the connectivity (electrical routing) within the chip. Data management and readout strategy also become critical tasks for which special techniques (such as pipelining and/or sparse data scan) must be developed.

4) For a given process, the matching between channels is reduced as the elements become smaller in size.

For the above reasons, short-term implementations of two-dimensional detector arrays with local intelligence are likely to be limited to "short strips" (30 to 50 μm wide by 200 to 500 μm long) or large pixels (e.g. 100×100 μm^2 to 200×200 μm^2). Future realizations should benefit from SOI and 3D technologies [1] to reach finer granularities.

3. Design strategies for fast- and low-power analog building blocks

3.1. Charge sensitive preamplifier (CSA)

The input amplifier basically consists of a transconductance amplifier g_m with an integrating capacitance C_f in feedback (see fig. 1). The sensitivity and speed of the CSA have been evaluated in section 2. Assuming again that the thermal series noise is dominant, the noise performance of the CSA alone is derived hereafter.

The equivalent input noise resistance of the transconductance amplifier g_m is given by

$$R_n = \gamma / g_m,$$

where γ is the noise factor of the transconductor and usually ranges from 1/2 to 1. The output noise voltage of the CSA is then found to be

$$V_{ns}^2 = \frac{\gamma}{h_{fb}} \frac{kT}{C_L}$$

yielding an equivalent input noise charge

$$\mathrm{ENC}^2 = V_{ns}^2 C_f^2 = \gamma kT (C_d + C_i + C_f) \frac{C_f}{C_L}$$

or equivalently by definition of τ_a

$$\mathrm{ENC}^2 = \gamma kT \frac{(C_d + C_i + C_f)^2}{g_m \tau_a}.$$

If, for example, $C_d = C_i = C_0 = 100$ fF, $C_f = 10$ fF and $\gamma = 1$, we obtain

ENC = 56 e$^-$ rms at room temperature.

The latter result shows that smart pixels are indeed promising candidates for very low noise applications. It is also worth nothing that although output stray capacitance C_0 is to be minimized in order to optimize the speed/power trade-off, any nonminimum power dissipation due to the presence of C_0 is not wasted since it helps reducing the noise.

The simple amplifier configuration of fig. 1 can be directly coupled to the detector providing that an appropriate dc feedback is applied either periodically with a MOS reset switch [2] or continuously with a MOS resistor of high value [3]. The first approach is limited to synchronous applications (events are expected at precisely predictable time slots) whereas the second is more general and can be implemented so as to automatically compensate for the detector leakage current. A possible CMOS realization is depicted in fig. 2. One output of the P-channel differential pair is connected to the amplifier input: this first feedback is equivalent to a

resistor of value $R_f = 1/g_{m1}$ in parallel with C_f. The current of the second diff pair output (drain of M1b) is integrated onto capacitor C and the resulting voltage controls the gate of the N-channel MOS transistor M2. The action of this second feedback path is equivalent to an inductor connected in parallel with C_f and therefore the detector leakage current (dc component) flows into M2 rather than into the equivalent feedback resistance R_f. The main advantage of this configuration is that since M2 can sink a total current which is not limited by the value of the bias current I_p, the leakage current of the detector may largely exceed the value of I_p without compromising the circuit operation.

Bias current I_p (typically between 10 pA and 1 nA) and equivalent feedback resistance R_f (in the 100 MΩ–10 GΩ range) contribute to the "parallel" noise as well as the detector leakage current I_{leak}; they should therefore be chosen in accordance with the noise specifications. The transconductance g_{m1} of the differential pair M1a–M1b must also satisfy the condition

$$\tau_f = R_f C_f = \frac{C_f}{g_{m1}} \gg \tau_a = \frac{C_L}{h_{fb} g_{m1}}.$$

To ensure a correct damping of the feedback loop, the transconductance g_{m2} of transistor M2 must then be chosen so that

$$\frac{C}{g_{m2}} > 2\frac{C_f}{g_{m1}}.$$

Since g_{m2} will depend on the detector leakage current, the above condition must be satisfied for the maximum value of I_{leak}.

3.2. Shaper

An elementary shaper amplifier is shown in fig. 3a: the high-pass characteristic is implemented by means of bypass capacitor C_{HP} and the low-pass transfer function can be realized with (parasitic) capacitor C_{LP}. The advantages of such a configuration over a purely passive RC–CR implementation are the tunability (by means of bias current I_s), the possibility of higher gain (for $g_{mi} > g_{mo}$) and the high input impedance. When the asymmetric nonlinear behaviour of this simple circuit is

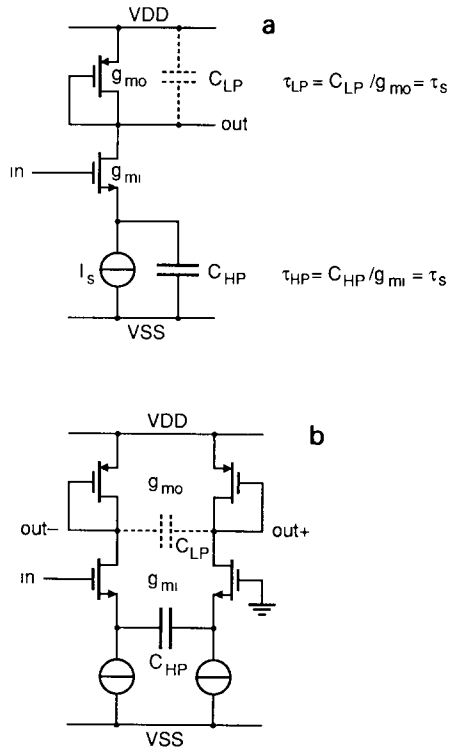

Fig. 3. Elementary shaper/amplifiers (a) single-ended and (b) differential implementations.

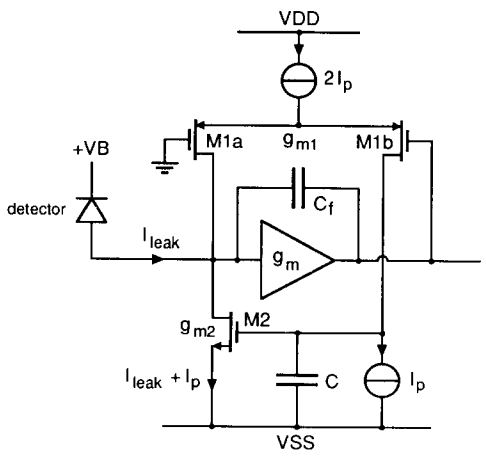

Fig. 2. Example of a CSA with a dc feedback circuitry which automatically compensates for the detector leakage current.

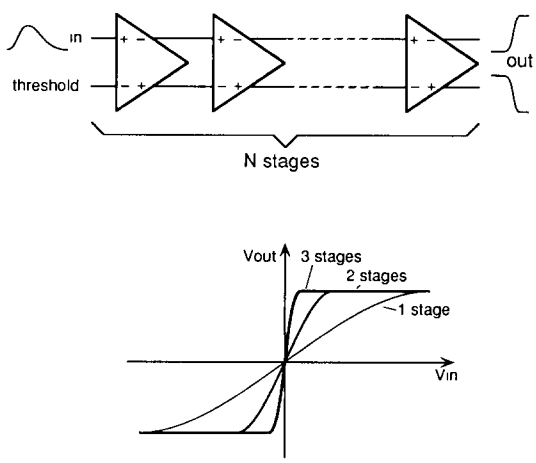

Fig. 4. Asynchronous comparator realized as a cascade of low-gain limiting amplifiers.

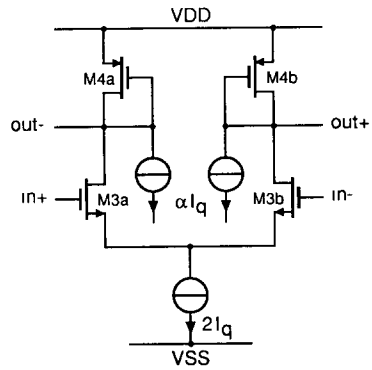

Fig. 5. Typical MOS implementation of a low-gain, low-power limiting amplifier.

a severe limitation, a balanced implementation (fig. 3b) can be used.

3.3. Discriminators

In asynchronous readout systems, discriminators (or comparators) are best realized as a cascade of low-gain limiting amplifiers (fig. 4), the small-signal gain and speed of which may be optimized as well as the number of stages of minimum power dissipation. A typical low gain stage is given in fig. 5. The input pair M3a–M3b is preferably operated in weak inversion where the transconductance is maximum at given drain current [4]. By preventing the drain current of diode-connected loads M4a–M4b from going down to zero when the input pair is overdriven, current sources αI_q ($\alpha = 0.05$–0.1 typically) reduce the time necessary to recover from overload conditions, thus increasing the speed. M4a–M4b are to be biased at moderate gate overdrive voltage ($V_{GS} - V_T = 0.4$ to 0.6 V) so as to realize an optimum small-signal gain of 3 to 5.

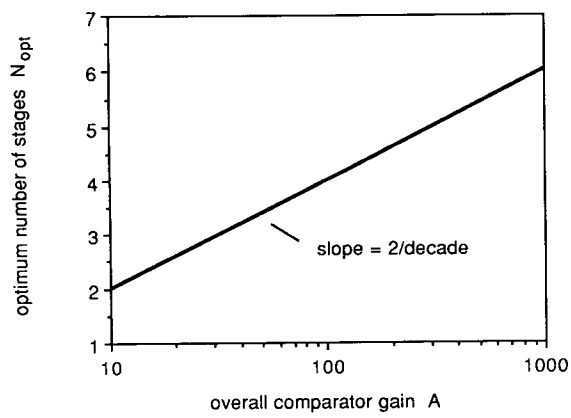

Fig. 6. Optimum number of stages for the asynchronous comparator for fig. 4 as a function of the overall small-signal gain.

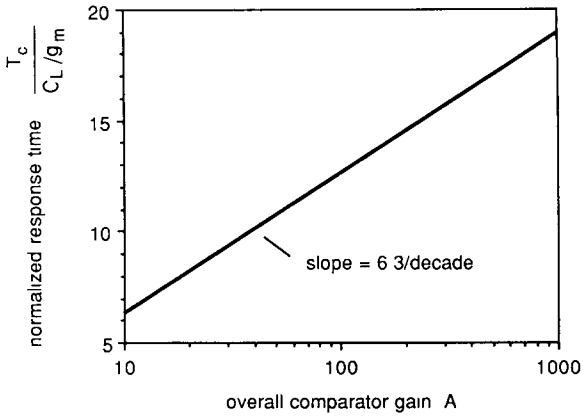

Fig. 7. Normalized total comparison time corresponding to the optimum defined in fig. 6.

Assuming N identical gain stages, the time T_c necessary for comparison is approximately

$$T_c = N\sqrt[N]{A}\,\frac{C_L}{g_m},$$

where A is the overall small-signal gain, C_L the equivalent load capacitance of one stage and g_m the transconductance of the differential pair M3a–M3b. The optimum number of stages which minimizes the power consumption is plotted in fig. 6 as a function of A. The gain per stage at optimum of found to be about 3.2, independently of the overall gain of the comparator, and the corresponding normalized total comparison time is given in fig. 7. Note that for correct operation of the discriminator, T_c must be much smaller than the shaping time constant.

When the operation of the discriminator can be precisely synchronized with the input signal, a regenerative (latched) comparator may be used. An example of

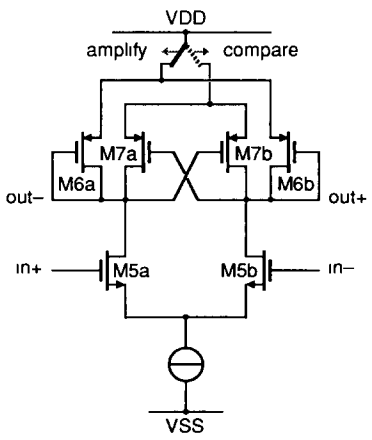

Fig. 8. Regenerative (latched) comparator for synchronous readout systems.

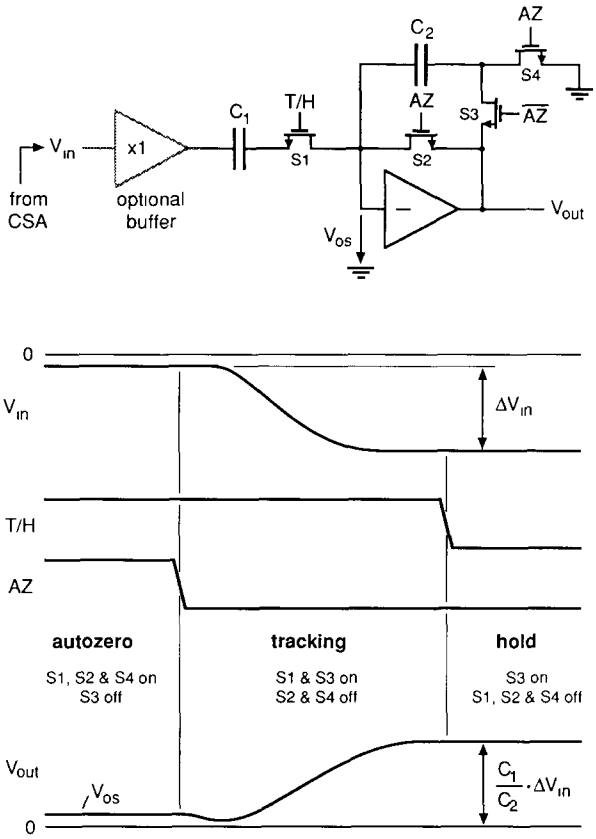

Fig. 9. Analog storage with correlated double sampling (CDS) for amplifier offset and noise reduction.

such a circuit is present in fig. 8 [5]. In the sketched position of the switch, the circuit behaves like a low-gain amplifier. When the input signal reaches its maximum, the switch is thrown into the other position and the output voltage grows rapidly thanks to the positive feedback induced by the cross connection of the gates of M7a and M7b. For this reason, the regenerative comparator is faster than a cascade of several limiting amplifiers and consumes less power.

3.4. Analog storage (sample-and-hold)

When analog information has to be stored, a switched-capacitor sample-and-hold circuit can be used. The circuit of fig. 9 compensates for the amplifier offset (the latter can be a simple inverter) and also implements the differentiation (high-pass) function so that a preliminary shaping of the output signal of the CSA is not necessary [2], thus reducing circuit complexity.

4. Circuit compaction

When implementing functional blocks, one usually attempts to realize the corresponding voltage-to-voltage transfer function or characteristics, as this was done in the previous section. However, the basic active elements, namely the MOS transistors, are by nature voltage-controlled current sources; consequently the circuit complexity could be reduced by alternating voltage-to-current and current-to-voltage transfer functions in cascaded implementations.

For example, in the circuit of fig. 10, the low-pass (RC) characteristic of the shaping is implemented by adding, if needed, a capacitor at the CSA output, whereas the high-pass (CR) function is performed at the comparator input by ac coupling of the differential pair. Since the correctly shaped signal now corresponds to the differential current $I_1 - I_2$, the discrimination threshold is adjusted by varying a current (I_{q1} and/or I_{q2}) rather than a voltage.

5. Local analog storage versus digital-only output

There is no doubt that adding local analog storage capability to a 1-bit digital output will increase both the circuit complexity, the pixel size and the power dissipation per pixel as well as the time necessary for readout. For example, there is a need for absolute values together with a wide dynamic range in calorimeters, but

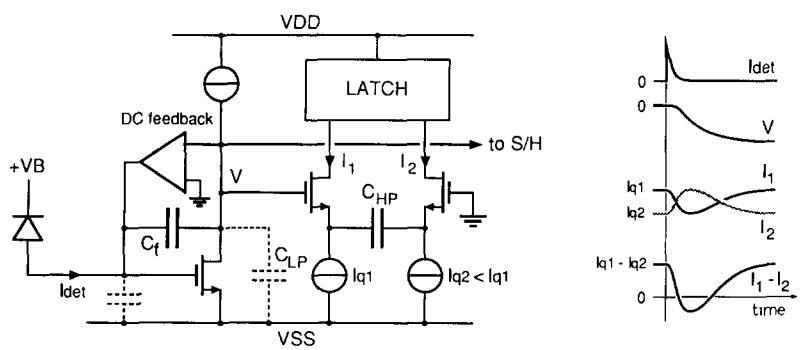

Fig. 10. Example of compacted analog front-end circuit for minimum transistor count and minimum power dissipation.

such applications can usually (and indeed will have to) accommodate a nonminimum pixel pitch.

The main argument in favour of analog readout is that in tracking systems the spatial resolution can be improved by interpolation. However, the increase in pixel size owing to analog storage and readout circuits will reduce the effective final resolution.

If interpolation is to be an issue, the signal strength of adjacent pixels can best be compared locally in order to avoid an important loading of the peripheral data processing electronics. One could for example exploit the relation between the input overdrive and response time of the comparator to find out, among contiguous "fired" pixels, which one received the highest signal, and use extremely simple excitatory/inhibitory local interaction between pixels to suppress insignificant information. Another and complementary technique is to use a 3-level data encoding scheme (two comparators with two different thresholds).

In any case, it is likely that several different approaches, both for the pixel element and for the readout strategy, have to be evaluated and compared before one can decide of the best trade-off.

6. Precision and matching between elements

In IC technology, the *absolute accuracy* of the process parameters (such as mobility of carriers, oxide thickness, threshold voltage etc.) is not controlled to better than 10–30% typically, whereas the *matching accuracy* of identically laid out devices on the same chip is much better (0.1–1% typically). Good designs therefore rely on matching and minimize the need for absolute values. One must keep in mind, however, that both absolute accuracy and matching are reduced as the device geometry approaches the process minimum feature size and as the distance between the elements to be matched is increased.

It is unrealistic to tune independently the performance of every pixel, but the systematic errors due to the overall absolute inaccuracies can be removed thanks to calibration. Basically, the latter may be accomplished by adjusting a pilot current and/or voltage so that the average circuit performance (speed and discriminator level for example) of a group of pixels meets the specifications. The "master/slave" calibration technique can be used to automatically tune the circuit and compensate for parameter variations with temperature, radiation damage and aging. In this technique, the performance of a "dummy" pixel element (the "master"), undergoing the same operating conditions as the active pixels (the "slaves"), is continuously sensed by a dedicated on-chip circuitry and automatically readjusted to match an external reference frequency, amplitude (charge, current or voltage) or element (resistor or capacitor). As the "master/slave" approach relies on the matching between the reference master element and the many active slave pixels, the size of the devices to be matched should neither be too small in size nor too far from each other. However, the size/matching trade-off must be evaluated experimentally as it is likely to be process dependent.

References

[1] J.-P. Colinge, these Proceedings (2nd Int. Workshop on Silicon Pixel Detectors for Particles and X-Rays, Leuven, Belgium, 1990) Nucl. Instr. and Meth. A305 (1991) 615.
[2] W. Buttler, B.J. Hosticka and G. Lutz, Proc. 5th Europ. Symp. on Semiconductor Detectors, Nucl. Instr. and Meth. A288 (1990) 187.
[3] E. Beuville, K. Borer, E. Chesi, E.H.M. Heijne, P. Jarron, B. Lisowski and S. Singh, ibid., p. 157.
[4] E.A. Vittoz, IEEE J. Solid-State Circuits SC-20 (1985) 657.
[5] M. Campbell, E.H.M. Heijne, P. Jarron, F. Krummenacher, C.C. Enz, M. Declercq, E. Vittoz and G. Viertel, Nucl. Instr. and Meth. A290 (1990) 149.

Discussion

E. Heijne (CERN): What is the maximum value of detector leakage current that your proposed scheme can tolerate? Could one imagine another solution in which the leakage current is actively compensated by injecting current in the opposite direction?

F. Krummenacher: If the detector leakage current has increased after some time it can still be absorbed completely by the transconductance of the input transistor. Provided its operation is close to weak inversion this can go over several decades without problem. Some current noise will of course be added in comparison with the situation for minimum current. This scheme will only work for one direction of detector bias polarity.

M. Cuzin: Could you specify for which kind of pixel detector your noise calculation of 65 electrons rms has been made? You base this on a capacitance of 100 fF

F. Krummenacher: The 100 fF capacitance represents a conservative number for a large silicon pixel of several 100 μm, or one with a thin depleted region. The capacitance may also be increased depending on the connection, if this is for example a wire bonding.