A Fully Integrated CMOS Readout Circuit for Particle Detectors

Zhang Yacong†, Chen Zhongjian, Lu Wengao, Zhao Baoying, and Ji Lijiu

(*Institute of Microelectronics, Peking University, Beijing* 100871, *China*)

Abstract: Novel schemes for a charge sensitive amplifier (CSA) and a CR-(RC)^n semi-Gaussian shaper in a fully integrated CMOS readout circuit for particle detectors are presented. The CSA is designed with poly-resistors as feedback components to reduce noise. Compared with conventional CSA, the input referred equivalent noise charge(ENC) is simulated to be reduced from 5036e to 2381e with a large detector capacitance of 150pF at the cost of 0.5V output swing loss. The CR-(RC)^n semi-Gaussian shaper uses MOS transistors in the triode region in series with poly-resistors to compensate process variation without much linearity reduction.

Key words: charge sensitive amplifier; shaper; readout circuit; noise optimization
EEACC: 1285; 7420
CLC number: TN432 **Document code**: A **Article ID**: 0253-4177(2007)02-0182-07

1 Introduction

Particle detectors are widely used in high-energy physics, nuclear physics, and radiation detection applications. To read out the charges generated, three approaches have been proposed[1]. The first is to integrate the input transistor of the preamplifier into the detector, to reduce the parasitic capacitance associated with the connection. However, the input transistor must be fabricated with the same process as the detector. Hence, it is hard to optimize the input transistor with reference to noise. The second is to use a JFET as the input transistor for its low noise property. The third is to integrate all the components required into one chip, which is called a fully integrated readout circuit[2]. With the increase of the number of channels and the rapid development of CMOS technology, the last approach is getting more and more attention because of the high density and low cost. However, the noise performance of CMOS transistors is worse than that of BJT and JFET. The total charges generated by one particle are so few that noise performance is the vital factor that must be considered. A charge sensitive amplifier (CSA) followed by a CR-(RC)^n shaping amplifier (shaper) is the most common readout circuit structure[1,3,8].

Fabricating high-accuracy passive components is another issue in fully integrated CMOS readout circuits. The value of the capacitors and resistors may vary by as much as 10% and 30%, respectively[4]. This leads to imprecisely controlled time constant of the shaper. In addition, the various parasitic capacitances can introduce undesired poles and zeros, affecting the frequency band of interest.

In this paper, the topology of one readout channel is described. Then, the novel CSA scheme is presented, which is designed according to the noise analysis[5] and results in less equivalent noise charge (ENC). The superiority of the proposed circuit is especially distinct for large detector capacitance. Next, the novel scheme of a CR-(RC)^n semi-Gaussian shaper is presented, using MOS resistors to compensate the process variation. The complete fully integrated readout circuit is verified by simulation to be satisfactory and feasible. It is ready to be taped out.

2 Circuit description

The topology of one channel is shown in Fig. 1[6]. The charges generated in the particle detector can be considered as a current impulse, which is identified as I_{in} in the figure. The charges are integrated onto a small integrating capacitor C_f, giving rise to a voltage step at the output of the CSA. R_f is used to discharge C_f in order to

† Corresponding author. Email: zhangyc@pku.edu.cn
Received 19 July 2006, revised manuscript received 31 August 2006

prevent the CSA from saturating. C_p and R_p are used for pole-zero cancellation, with the requirement that $C_pR_p = C_fR_f$. The components in the dashed box compose an RC cell, and one channel can contain n such cells. Together with C_p and R_p, this circuit is known as a CR-$(RC)^n$ semi-Gaussian shaper or pulse-shaping amplifier, which filters and shapes the CSA output signal. C_c is for test.

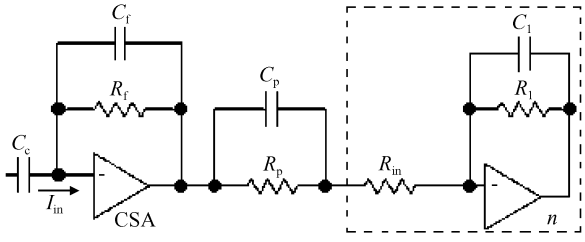

Fig. 1　Topology of one channel

3　Design considerations

3.1　Noise analysis

The noise performance of a detector readout circuit is usually described by the equivalent noise charge (ENC), which is defined as the ratio of the total integrated rms noise at the output of the pulse shaper to the signal amplitude due to one electron charge[5].

The ENC can be calculated using the noise model in Fig. 2, where V_{eqi}^2 is the equivalent input noise voltage of the CSA and I_d^2 is the shot noise contributed by R_f and the detector leakage current. V_{eqi}^2 is dominated by the thermal and flicker noise of the CSA input transistor.

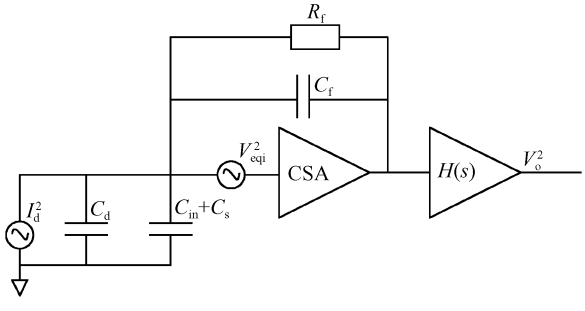

Fig. 2　Noise sources of the readout circuit

The ENCs are expressed as follows, assuming a semi-Gaussian pulse shaper is used[5]:

$$\mathrm{ENC}_d^2 = \frac{8}{3}kT\frac{1}{g_m}\times\frac{C_t^2 B\left(\frac{3}{2}, n-\frac{1}{2}\right)n}{4\pi q^2\tau_s}\left(\frac{n!^2 e^{2n}}{n^{2n}}\right) \quad (1)$$

$$\mathrm{ENC}_f^2 = \frac{K_f}{C_{ox}^2 WL}\times\frac{C_t^2}{2q^2 n}\times\left(\frac{n!^2 e^{2n}}{n^{2n}}\right) \quad (2)$$

$$\mathrm{ENC}_o^2 = \left(2qI_o + \frac{4kT}{R_f}\right)\frac{\tau_s B\left(\frac{3}{2}, n+\frac{1}{2}\right)}{4\pi q^2 n}\left(\frac{n!^2 e^{2n}}{n^{2n}}\right) \quad (3)$$

Here ENC_d^2, ENC_f^2, and ENC_o^2 are ENC^2 due to thermal, flicker, and shot noise, respectively. C_t is the total capacitance at the input node, including the detector capacitance C_d, the CSA input capacitance C_{in}, and the parasitic capacitance C_p. I_o is the sum of the detector leakage current and the equivalent noise current of the bias network. τ_s and n are the peaking time and the order of the shaper, respectively. $B(x, y)$ is the beta function. All the other symbols have their normal meaning. These formulas are the guides to the design of the CSA and shaper.

3.2　Design of CSA

The amplifier A in Fig. 1 is usually implemented with a single-ended folded-cascode amplifier because of its high DC gain and large bandwidth. Although the noise of the input transistor dominates due to its large transconductance[5], other transistors also generate noise. A new circuit is proposed to reduce the noise of the current source and load transistor. Compared to the conventional architecture, the resistors are added to the sources of these transistors, as can be seen in Fig. 3.

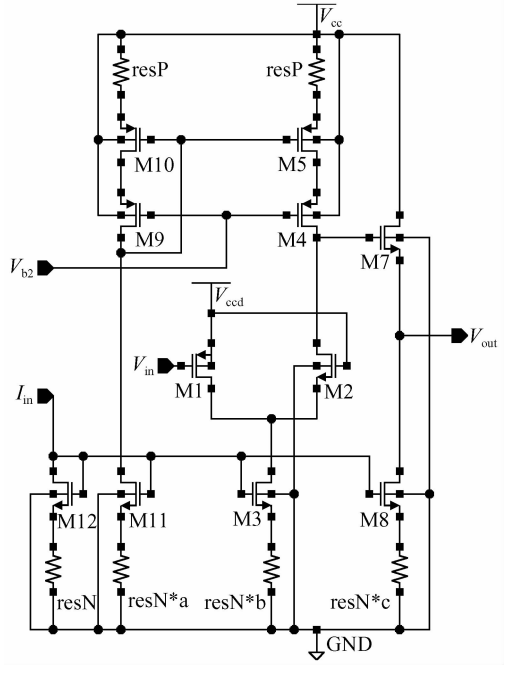

Fig. 3　Topology of the proposed amplifier

Besides M1, other important noise generators are M3 and M5. Without the resistors, the input referred equivalent noise due to Mi (i = 3, 5) can be expressed as

$$V_{n,i}^2 = g_{mi}^2 \times \left(\frac{8kT}{3g_{mi}} + \frac{K_f}{W_i L_i C_{ox} f}\right) / g_{m1}^2 \quad (4)$$

if the channel length modulation effect is ignored.

In CMOS technology, nMOS has much more flicker noise than pMOS. In addition, the current in M3 is even larger than that in M1. Therefore, both the thermal and flicker noise of M3 are not negligible.

The usual method to reduce the noise due to M3 and M5 includes increasing g_{m1}, increasing the gate area, and reducing W/L of M3 and M5, but they are not sufficient if extremely low ENC is required.

With the resistors added, the equivalent input noise due to Mi (i = 3 or 5) and its corresponding source resistance R_{si} is given by

$$V_{n,i}^2 = \left[\frac{g_{mi}^2}{(1 + (g_{mi} + g_{mbi})R_{si})^2} \times \left(\frac{8kT}{3g_{mi}} + \frac{K_f}{W_i L_i C_{ox} f} + 4kTR_{si}\right)\right] / g_{m1}^2 \quad (5)$$

As for the thermal noise, it can be verified that $f(R_{si}) = \frac{g_{mi}^2}{(1 + (g_{mi} + g_{mbi})R_{si})^2} \times \left(\frac{8kT}{3g_{mi}} + 4kTR_{si}\right)$ is a descent function of R_{si}. Hence, larger resistance is beneficial for noise suppression. Take M3 as an example. The parameters in the designed circuit are $I_{ds3} = 250\mu$A, $g_{m3} = 1.39$mS, and $g_{mb3} = 0.51$mS. It can be calculated that $f(R_{si} = 0) = 4kT \times 0.93\text{mA}^2$, then $f(R_{si} = 2000) = 4kT \times 0.208\text{mA}^2$. A reduction of 77.6% is achieved. $R_{si} = 2000$ means a 0.5V voltage drop ($R_{si} I_{ds3}$) on the resistor. Another advantage is that the resistors are free of flicker noise. Equation (5) indicates that the flicker noise is reduced by a factor of $(1 + (g_{mi} + g_{mbi})R_{si})$[7].

The coefficients of resN and resP in each branch of the circuit scale according to the current passing through it to guarantee equal voltage drop across each resistor.

The dimension and the bias current of the input transistor of the CSA should be optimized with the noise. For minimal ENC_d^2, g_m ought to be enlarged, requiring minimal gate length and maximal bias current. However, the gate width has a double effect. It enlarges g_m but also increases C_t. Referring to ENC_f^2, WL also has a double effect. Considering the thermal and flicker noise together, minimal L should be chosen whereas an optimal gate width exists for ENC_d and ENC_f, respectively[5].

$$W_{opt,d} = \frac{C_d + C_f}{2C_{ox}L}, \quad W_{opt,f} = \frac{3(C_d + C_f)}{2C_{ox}L} \quad (6)$$

Nevertheless, when the detector capacitance C_d is on the order of several tens of pF, the calculated W_{opt} is so huge that it is impractical for integration. For the case of $C_d = 150$pF, $W_{opt,d} = 49000\mu$m and $W_{opt,f} = 147000\mu$m are calculated in 0.5μm technology. A relatively smaller device should be used at the expense of larger ENC. In our design, W is set to be 1000μm. This transistor is biased at the edge of strong-inversion. The bias current is chosen to be 200μA to meet power constraints.

3.3 Design of the shaper

The shaper is used for filtering and shaping. In essence, it is a bandpass filter that filters out the low- and high-frequency noise. At the same time, it changes the step voltage into a semi-Gaussian shape to suppress pileup and to facilitate the recording of the pulse height. The height of the shaper output pulse is a measure of the total charge generated in the detector. The transfer function from the output of the CSA to the output of the shaper can be expressed as

$$H(s) = \frac{1 + sC_pR_p}{1 + \frac{R_p}{R_{in}} + sC_pR_p} \times \left(\frac{R_1}{R_{in}(1 + sC_1R_1)}\right)^n \quad (7)$$

Note that if $C_pR_p/(1 + R_p/R_{in}) = R_1C_1$ is satisfied, then a uniform time constant of R_1C_1 is achieved for the denominator. Since normally R_p is much bigger than R_{in}, the above equation is simplified to $C_pR_{in} = C_1R_1$. In this formula, the required time constant $\tau_0 = R_1C_1$ and gain that is proportional to $\frac{1}{1 + R_p/R_{in}}\left(\frac{R_1}{R_{in}}\right)^n$ are the basis of component values choice.

From Eqs. (1～3), an important conclusion can be drawn that $\text{ENC}_d^2 \propto C_t^2/\tau_s$, $\text{ENC}_f^2 \propto C_t^2$, $\text{ENC}_o^2 \propto \tau_s$. This results in an optimum τ_s, corresponding with the time at which the contributions of ENC_d and ENC_o are equal[5]. It is easy to deduce that for detectors with little leakage current and large capacitance, such as 75～150pF, the optimum τ_s would be very large. In practice, the

counting rate requirement and pileup problem constrain the upper limit of τ_s. In our cases, the particle rate is about $10^5 s^{-1}$, which restricts τ_s to be 1μs or so. This small τ_s indicates that ENC_o has little effect on the total ENC. For the transfer function given in Eq. (7), the peaking time (also called the shaping time) τ_s is related to τ_0 by $\tau_s = n\tau_0$.

The number of RC cells n affects the noise performance, but the influence is not very significant as long as n is larger than 4[9]. n also changes the shape of the output voltage. As n increases, the curve becomes more symmetrical, which means it is restored back to the baseline more quickly[9]. Thus for a given counting rate, a larger shaping time can be chosen if more cells are cascaded. This reduces ENC_d^2 at the expense of the power and area.

The amplifier in the shaper is the same as that in the CSA to prevent dc current from flowing through R_{in}.

3.4 Realization of R_f and R_p

R_f is usually a big resistor on the order of 100MΩ～10GΩ because the time constant determined by R_f and C_f must be much larger than the shaping time. It is area-consuming to implement it with physical resistors in CMOS technology. Alternately, a MOS transistor in the triode region exhibits a resistance with the value of $R_{on} = \frac{1}{k'(W/L)(V_{GS} - V_T)}$. Hence it is possible to realize a large resistance by setting W/L and $V_{GS} - V_T$ small. However, a small overdrive voltage would lead to poor linearity and a small aspect ratio means large L, which introduces large parasitic capacitance. The CSA can be modeled by the circuit of Fig. 4 as a first order approximation, where R_1 and R_2 (both equal $R_f/2$) and C_1 to C_3 represent the MOS resistor with parasitic capacitance. Because the input of the amplifier is ac-grounded, C_1 has no effect. The transfer function is

$$\frac{V_o}{I_{in}} = -\frac{R_f(1 + sR_fC_2/4)}{1 + sR_fC_f(1 + sR_fC_2/4)} \qquad (8)$$

There is a zero located at $4/R_fC_2$ and two poles at $\frac{1 \pm \sqrt{1 - C_2/C_f}}{R_fC_2/2}$. Only when $C_2 \ll C_f$ can the desired result of $p_1 = 1/R_fC_f$ be achieved. If $C_2 > C_f$ the poles will become conjugated complex poles located at higher frequencies than the zero. The ac response with C_2 as a parameter is shown in Fig. 5. Since C_f is on the order of hundreds of fF or less, the area of the MOS transistor used as R_f is restricted. In our design, C_f is set to be 200fF to get considerable gain of the CSA, and R_f is realized by two 0.5μm/20μm transistors in series.

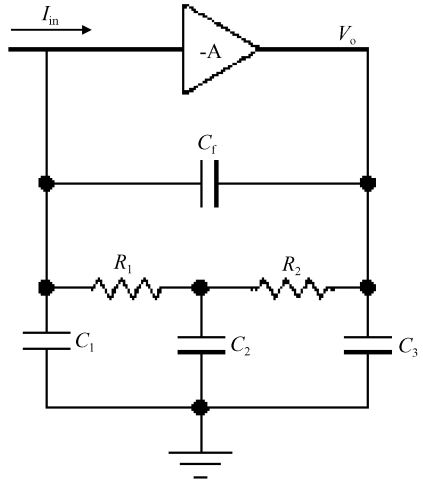

Fig. 4　Equivalent circuit of CSA with MOS R_f

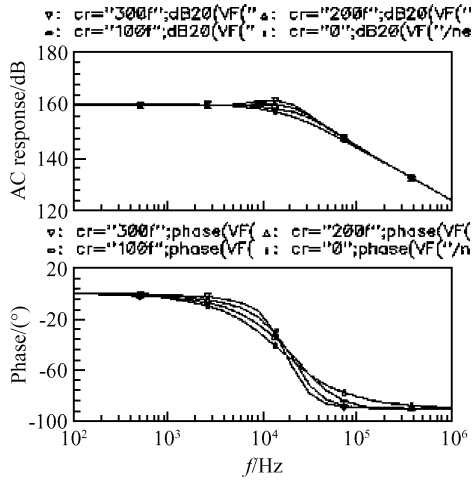

Fig. 5　AC response of Fig. 4

R_p and C_p provide the pole-zero cancellation to prevent the undershoot which decays back to the baseline with the long time constant of R_fC_f. Besides, the multiple between R_f and R_p provides a current gain of $N = R_f/R_p$. For match consideration, R_p is realized with N copies of R_f in parallel.

3.5 Realization of R_{in} and R_1

The shaper time constant $\tau_0 = R_1C_1$ is a strong function of process parameters, because it is related with the absolute value of passive com-

ponents. To compensate process variation, R_{in} and R_1 are implemented with high-resistance-poly resistors in series with triode-region MOS transistors. This is feasible because the voltage at the input of the amplifier is fixed. The circuit is shown in Fig. 6, where V_{c2} and V_{c3} are the tuning voltages. $V_{ds} \ll V_{gs} - V_{th}$ is required to keep enough linearity for the MOS to work in the triode region. This inequality gives two indications: (1) V_{c2} and V_{c3} must be high enough to ensure that V_{gs} is large, limiting the tuning range; (2) The MOS resistance should be small since the proportion of it to the total resistance determines V_{ds}. This is in conflict with the compensation capability. In this design, two MOS's are used in parallel to enlarge the tuning range. One exhibits big resistance while the other is relatively small. In case the fabricated resistors R_{ina} and R_{1a} come out to be much smaller than the designed value, the big resistance is used while the small is turned off by grounding its gate. The linearity is sacrificed, but the probability of this case happening is small. Normally, the MOS's are tuned to a moderate value to maintain strong linearity.

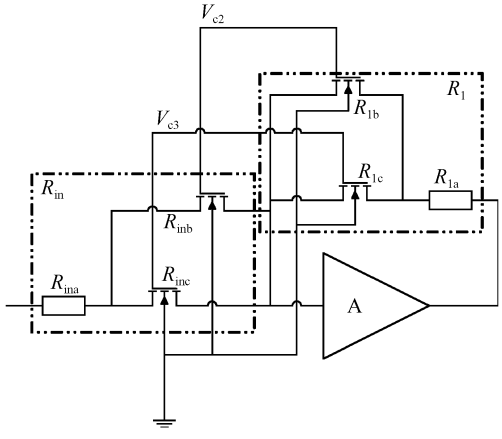

Fig. 6 Implementation of R_{in} and R_1

4 Simulation results

The readout circuit is designed with CSMC 0.5μm DPDM n-well technology. The layout without pads occupying 330μm × 300μm is shown in Fig. 7. The circuit is ready to be taped out. The transient response of a current pulse input is shown in Fig. 8. The output of the shaper exhibits semi-Gaussian shape. Compromising between noise and power, the peaking time is set to be 1μs or so and $n = 2$. Since the quality factor of the shaper is 0.5, the pulse is restored to the baseline without undershoot, as the figure shows. The gain of the complete circuit is about 79mV/fC, and the static power is 4.2mW per channel.

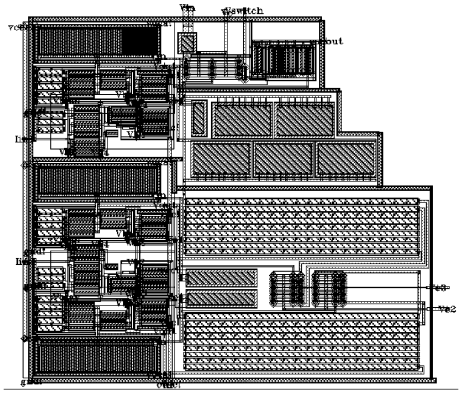

Fig. 7 Layout of the readout circuit

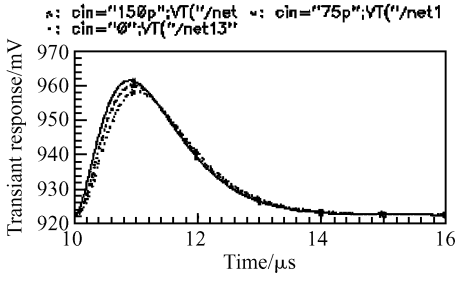

Fig. 8 Transient response of Fig. 1

ENC is calculated for Fig. 1 with the amplifier shown in Fig. 3 for three cases, which are resN = resP = 0 (reduced to the conventional circuit), resN = 5000 (the voltage drop on the resistor is 0.5V) and resP = 0, and resN = resP = 5000. According to the definition mentioned in Section 3.1, ENC is given by $\mathrm{ENC} = \sqrt{\overline{V_{on}^2}} \Big/ \frac{qV_{op}}{C_c V_s}$, where $\overline{V_{on}^2} = \int S_o(f)\mathrm{d}f$ is the integration of the output noise voltage spectral density across the bandwidth. The upper limit of the integration is set to be 1GHz, well above the bandwidth of the practical measurement system. $\overline{V_{on}^2}$ can be directly read in the noise summary window of the simulator. V_{op} is the output pulse height due to a block pulse with amplitude V_s applied on the test capacitance C_c. The results with $C_d = 0, 30, 60, 90, 120$, and 150pF are shown in Fig. 9. Apparently, the resistors reduce ENC dramatically, especially when C_d is large. This benefit is gained at the expense of a smaller voltage swing. Therefore the best choice is resN = 5000 while resP = 0, which means a swing loss of 0.5V but a large noise performance gain.

In this case, the output pulse height allowed is about 1.2V. resP = 5000 makes 0.5V more swing loss but the noise performance improvement is little because resP is at the source of pMOS, whose flicker noise is much smaller than nMOS. The relatively small swing is not a severe problem, because a variable gain amplifier follows the pulse shaper.

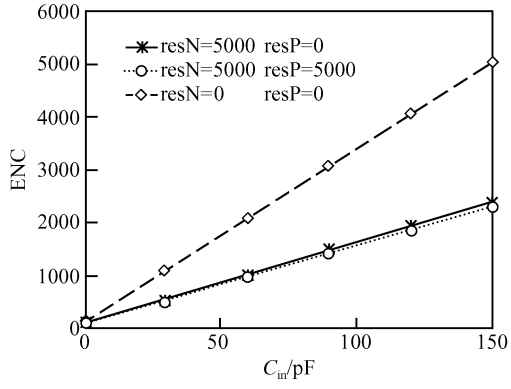

Fig. 9　Calculated ENC with different resistances

The circuit is also simulated using the ff and ss corner. Mainly due to the dependency of the transconductance on the process parameters, the gain and ENC also change. Take C_d = 150pF as an example. The gain is 72.7, 69.4, 62.5mV/fC and ENC is 2265e, 2381e, 2439e in ff, tt, and ss corner, respectively. The gain variation is not important as explained above, while the percent of ENC variation is about 10%.

The variation of resN with process affects ENC and the voltage swing. resN = 4000 and 6000 result in ENC = 2471e and 2315e, corresponding to a deviation of 3.8% and 2.8% from the case of resN = 5000, respectively. The influence on swing can be examined by $I_{ds} \times$ resN.

As explained above, ENC depends on the dimension of the CSA input transistor, the bias current, the peaking time and the process parameters. Thereby, it is not easy to compare different circuits. Some results in recent literature are listed in Table 1 for reference.

Table 1　Performance comparison

	Process (CMOS)	Peaking time/μs	Power /mW	ENC(e)
Geronimo[8]	0.25μm	5	-	300 + 7/pF
Geronimo[10]	0.25μm	0.6	1.25	100 + 25/pF
Grybos[11]	0.8μm	0.5～1	4	200@3pF
Yeom[12]	0.35μm	0.5	-	380 + 30/pF
This work	0.5μm	1	4.2	110 + 15/pF

If R_{ina} and R_{1a} are fixed, the shaping time τ_s can be tuned from 0.92 to 1.125μs by changing V_{c2} and V_{c3} from 5 to 3.5V (or 0V if the transistor is supposed to be off) in Fig. 6. This means that when R_{ina} and R_1 deviate from their designed value due to process variation, τ_s can be tuned back to its normal value.

The linearity of four typical cases is shown in Fig. 10. $V_{c2} = V_{c3} = 5$V corresponds to the minimal shaping time, which is 0.92μs. Because in this case, the resistance shown by the MOS's is the smallest, the linearity is the best. On the contrary, when $V_{c2} = 3.5$ and $V_{c3} = 0$V, corresponding to a shaping time of 1.125μs, the MOS resistance is the largest and the linearity is the worst. The sources of nonlinearity include that of the MOS transistor used as R_f, R_p, R_{in} and R_1 and the amplifiers. The linearity shown in the figure is calculated according to the following formula where the residual is the difference between the simulated result and the linear fit result.

$$\text{Linearity} = \left(1 - \left|\frac{\max(\text{residual})}{\text{swing}(V_{out})}\right|\right) \times 100\% \tag{9}$$

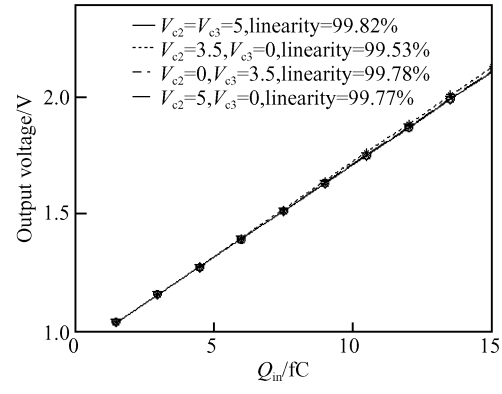

Fig. 10　Linear fitness of the gain

5　Summary

A fully integrated CMOS readout circuit is presented with detailed design considerations. A new amplifier scheme is proposed based on the noise analysis. The resistors at the sources of the current mirrors can improve the noise performance at the cost of reduced voltage swing. The problem of the relatively small swing can be resolved by the variable gain amplifier following the shaper, which is not included in this paper. MOS transistors working in the triode region in series

with poly-resistors provide tuning flexibilities for the shaper to compensate process variation without much linearity reduction.

References

[1] O'Connor P, Gramegna G, Rehak P, et al. Ultra low noise CMOS preamplifier-shaper for X-ray spectroscopy. Nuclear Instruments and Methods in Physics Research A, 1998, 409: 315

[2] Vandenbussche J, Leyn F, Plas G V, et al. A fully integrated low-power CMOS particle detector front-end for space applications. IEEE Trans Nucl Sci, 1998, 45(4): 2272

[3] Hu Y, Deptuch G, Turchetta R, et al. A low-noise, low-power CMOS SOI readout front-end for silicon detectors leakage current compensation with capability. IEEE Trans Circuits Syst, 2001, 48: 1022

[4] Johns D, Martin K. Analog integrated circuit design. John Wiley, 1997

[5] Sansen W M C, Chang Z Y. Limits of low noise performance of detector readout front ends in CMOS technology. IEEE Trans Circuits Syst, 1990, 37: 1375

[6] Britton C L, Alley G T, Simpson M L, et al. Design and characterization of the BVX: an 8-channel CMOS preamplifier-shaper for silicon strips. IEEE Trans Nucl Sci, 1994, 41(1): 352

[7] Wei Lan, Gao Jun, Chen Zhongjian, et al. A new structure of low-noise CMOS differential amplifier. ASICON, 2005, 1: 360

[8] De Geronimo G, Carini G, Murray W S, et al. Front-end ASIC for co-planar grid sensors. IEEE Trans Nucl Sci, 2005, 52(5): 2003

[9] Chang Z Y, Sansen W M C. Low-noise wide-band amplifiers in bipolar and CMOS technologies. Kluwer Academic Publishers, 1991

[10] De Geronimo G, Fried J, O'Connor P, et al. Front-end ASIC for a GEM based time projection chamber. IEEE Trans Nucl Sci, 2004, 51(4): 1312

[11] Grybos P, Rodriguez A E C, Idzik M, et al. RX64DTH-a fully inte-grated 64-channel ASIC for a digital X-ray imaging system with energy window selection. IEEE Trans Nucl Sci, 2005, 52(4): 839

[12] Yeom J Y, Defendi I, Takahashi H, et al. A 12-channel CMOS preamplifier-shaper-discriminator ASIC for APD and gas counters. IEEE Nuclear Science Symposium Conference Record, 2005

一种粒子探测器的 CMOS 读出电路设计

张雅聪† 陈中建 鲁文高 赵宝瑛 吉利久

(北京大学微电子研究院，北京 100871)

摘要： 提出了新型的应用于粒子探测器 CMOS 读出电路中的电荷灵敏放大器和 CR-(RC)^n 半高斯整形器的结构.电荷灵敏放大器采用多晶硅电阻做反馈来减小噪声,仿真发现与传统结构相比,在探测器电容高达 150pF 时,输入等效噪声电荷数由 5036 个电子减小到 2381 个,代价是输出摆幅减小了 0.5V.在整形器中,MOS 管电阻与多晶硅电阻串联,通过调节 MOS 管的栅压来改变阻值,以补偿工艺的偏差,在不明显降低线性度的情况下保证了时间常数能够比较精确控制.

关键词： 电荷灵敏放大器；整形器；读出电路；噪声优化

EEACC： 1285；7420

中图分类号： TN432 **文献标识码：** A **文章编号：** 0253-4177(2007)02-0182-07

† 通信作者.Email:zhangyc@pku.edu.cn

2006-07-19 收到,2006-08-31 定稿