A Fast Readout using Switched Current Techniques for a DEPFET-Pixel Vertex Detector at TESLA

M. Trimpl[a*], L. Andricek[b], P. Fischer[c], G. Lutz[b], R.H. Richter[b], L. Strüder[b], J. Ulrici[a], N. Wermes[a]

[a]Physikalisches Institut der Universität Bonn, Germany

[b]Halbleiterlabor des MPI München, Germany

[c]Lehrstuhl für Schaltungstechnik und Simulation, Universität Mannheim, Germany

A fully depleted silicon detector with a first amplifying transistor integrated in every pixel (DEPFET) is a promising proposal for the pixel-based vertex detector at TESLA. The DEPFET offers good spatial resolution, an excellent signal to noise ratio and low power consumption in a row-wise operation mode. A readout concept for a DEPFET pixel array matching the requirements at TESLA is described. In order to meet the operation specifications at TESLA (50 MHz row rate), a readout architecture based on current mode techniques (Switched Current) is presented. It contains stand alone zero suppression offering a triggerless operation. The core of the readout chip, a fast operating current memory cell, is discussed in detail. The results of a first prototype chip show that the requirements for TESLA are achievable.

Keywords: readout electronics, Switched Current, Linear Collider, TESLA, Active Pixel Sensor, DEPFET

1. Introduction

One of the important aims to be addressed by the detectors of a future Linear Collider like TESLA are an excellent efficiency and purity in the flavor identification of hadronic jets. The separation between b and c decay tracks and the distinction between their primary, secondary and tertiary vertices should become possible. The required impact parameter resolution leads to a detector scenario where the innermost layer of the vertex detector is placed as close as possible to the interaction point. In the present design of TESLA the first layer of the vertex detector is situated at $r = 15$ mm [1]. Due to the extremely focussed beam and the high bunch charge the e^+e^--background induced by beamstrahlung becomes severely high in this region (0.03/0.05 hits per mm^2 and bunch crossing for $\sqrt{s} = 500/800$ GeV and a magnetic field of 4T [1]). The bunch structure with a long bunch train of 950 μs at TESLA imposes the necessity to read out the detector several times during the train in order to keep the detector occupancy at a level that does not compromise the track reconstruction. At present the readout time of the whole vertex detector is fixed to 50 μs corresponding to an estimated occupancy below 1%.

In order to minimize the multiple scattering contribution to the impact parameter resolution, the reduction of any material in the detector area to a minimum is mandatory. This strongly requires close attention to the cooling and, therefore, power consumption in the sensor area. This disfavors a conventional hybrid pixel solution as used in the LHC-experiments in which the readout electronics is mounted on top of the detector and every pixel is processed in parallel. On the other hand, a row-wise readout of the detector, where the readout chips are placed at the end of the sensor, imposes a row rate of 50 MHz in order to meet a frame rate of 50 μs. To further reduce the amount of material in the detector area, a material budget of less than 0.1 % radiation length per layer is envisaged, the sensor itself is going to be thinned to ~ 50 μm. Consequently, the significant reduction of the resulting signal demands a

*Corresponding author. Tel.: +49-228-73-3239; fax: +49-228-73-3220. E-mail address: trimpl@physik.uni-bonn.de

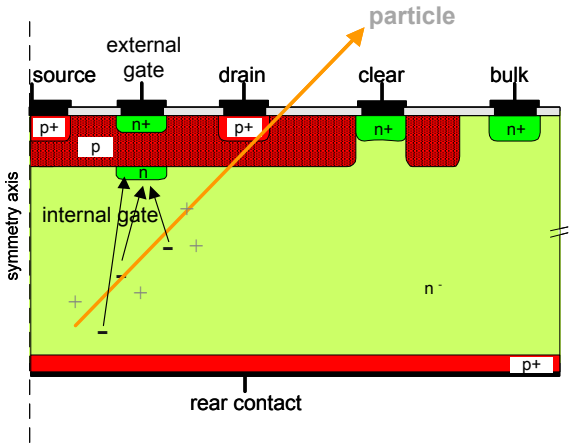

Figure 1. Cross sectional view of a DEPFET pixel showing the principle of operation.

low noise detector and readout to keep a sufficient signal to noise ratio. Furthermore, the sensor as well as the readout electronics has to tolerate the radiation exposure at TESLA.

2. DEPFET Concept for TESLA

The concept of the DEPFET [2] as shown in Fig. 1 is the integration of a first amplifying transistor (JFET or MOSFET) in each pixel. Electrons generated by a traversing particle are accumulated in the internal gate by means of sidewards depletion and additional implants modulating the transistor current. The external gate can be used to switch on the transistor in order to probe the device current. Note that for the mere accumulation of charge itself the transistor does not need to be switched on. Due to the non-destructive readout of the pixel device the charge in the internal gate has to be removed from time to time. This reset procedure is also called CLEAR and is realized by an extra contact at the border area of each pixel. A particle can be detected by comparing the transistor current after an integration time with the pedestal current after the reset. DEPFET devices have already been produced that demonstrate a very low noise performance of less than ENC=5 e^- with single devices [3]. For a large DEPFET matrix operated at a readout speed as required for TESLA, a total noise figure in the order 50-100 e^-

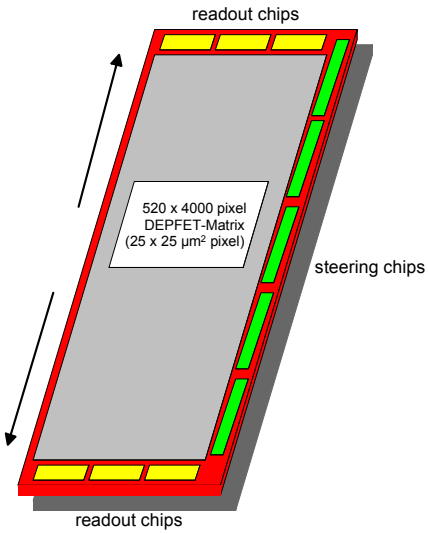

Figure 2. Sketch of one ladder of the DEPFET based vertex detector for TESLA.

or less is aimed for. A more detailed description of the DEPFET principle as well as the present level of technological development is given in [4].

Several DEPFET devices can be arranged in an array (see upper part of Fig. 3) to cover a large sensor area. The external gate contacts as well as the CLEAR contacts of each row are merged. The drains of the amplifying transistors that provide the device current are connected column-wise. A readout chip placed at the bottom of the columns processes the signals of one selected row by switching off the rest of the transistors in the array. This way of controlling a DEPFET array has been successfully used in the BIOSCOPE-System [5] designed for biomedical applications. A possible arrangement of a DEPFET-based ladder for TESLA [6] is shown in Fig. 2. The sensor area itself is thinned to ~ 50 μm leaving a thicker frame for mechanical stability. Steering chips for row selection and reset as well as readout chips are placed onto the frame, the steering chips at the ladder side and the readout chips outside the sensor area at the ladder end. The sensor is read out in both directions thus improving the readout speed by a factor of 2.

To meet the features of the DEPFET sensor, a fast readout with moderate noise performance has to be developed. The design concept and re-

alization of such a readout chip is the subject of the following sections.

3. Architecture of the Readout Chip

The basic architecture of the readout chip is presented in Fig. 3. The DEPFET-Matrix is read out via the drains which are connected column-wise with the chip. As the signal of the DEPFET device is a current and a fast readout is needed, the architecture is completely based on current mode signal processing. The readout chip consists of two major parts.

In the first part of the chip, a regulated cascode keeps the potential of the input node constant, eliminating the influence of the large sensor capacitance and the effect of the limited output conductance of the pixel transistor itself. After a row in the DEPFET array has been selected for readout, the DEPFET current is stored in a current memory cell. The same row is then reset and the pedestal is subtracted from the buffered current automatically at the output node of the memory cell due to its inverting property. The so-derived signal current is stored in a FIFO-like analog memory structure. This procedure guarantees a continuous readout of the sensor with a constant readout speed independent of the occupancy. As the two measurements of the signal and pedestal are performed shortly after each other, the 1/f noise of the system is intrinsically reduced. Furthermore, noise originated by leakage current in the sensor is reduced by the short frame time that is used.

In the second part of the chip, the FIFO structure is emptied successively. Immediately after storing the currents in a FIFO row, a parallel current compare for each column generates a digital hit pattern indicating signals above the threshold (hits). If there is no hit found in the complete FIFO row or if the amount of hits exceeds a programmable maximum value, the row is skipped and overwritten by succeeding cycles. According to the digital hit pattern, all analog cells in the row containing no hits are turned off in order to save static power. A digital hit finder [7] identifies the hits in the digital pattern of a row and multiplexes the corresponding analog values to an

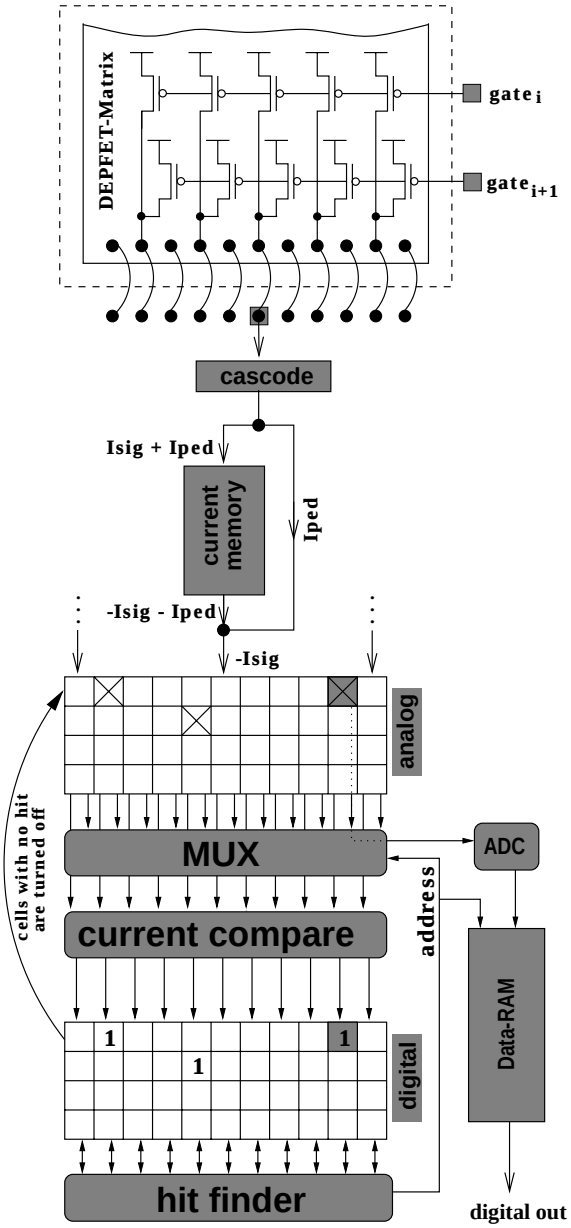

Figure 3. Basic Architecture of the readout chip. In part, only one channel is shown for simplicity.

algorithmic ADC. The hit finder finds up to 2 hits per cycle and processes a row until no more hits are found. The FIFO row is then available again for new data. Note that the FIFO structure's purpose is to derandomize the number of hits occurring in the sensor rows. Therefore a constant readout rate performed by the hit finder and the ADC can be used to read out the sensor. It is essential that the average hit occupancy does not exceed the readout rate of the hit finder. For the occupancy expected at TESLA, one hit finder is sufficient. To cope with larger occupancies more hit finders can be used in parallel. The maximum number of hits allowed in a row before it is skipped by the algorithm can be influenced by the FIFO depth.

Due to the immense data reduction, only one fast ADC is needed, thereby improving the power consumption of the readout chip. After the analog to digital conversion, the hit amplitude and its address are stored in a RAM for later readout. As the readout architecture offers a stand alone hit detection, no hardware trigger is needed.

4. Analysis of the Current Memory Cell

The most challenging part of the readout architecture in section 3 is the fast and precise buffering of the DEPFET-current for the pedestal subtraction and the storage in the analog FIFO. The design of such a current memory cell is described in this section. The basic principle of a current memory stage [8] is shown in Fig. 4, where C_G is the gate capacitance of the transistor M1. The sample and hold process is divided into three phases:

1. S1 and S2 are closed, S3 open. The gate capacitance of the transistor M1 is charged until the device provides the combined input and bias current ($I_{M1} = I_{in} + I_B$).
2. S2 is opened. The gate voltage and therefore the transistor current ideally remain unchanged.
3. Immediately after sampling, S1 is opened and S3 closed. As the current through M1 is still $I_{M1} = I_{in} + I_B$, I_{in} must be delivered by the output node.

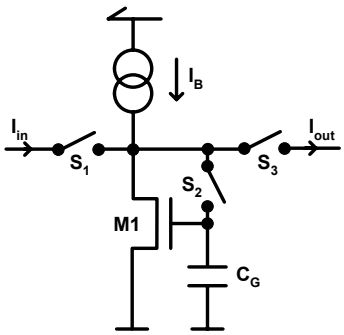

Figure 4. Basic principle of the current memory stage.

Thus, in the ideal case $I_{out} = -I_{in}$. However, this simple circuit suffers from several non-ideal effects like charge-injection of the sampling switch S2 and the limited output conductance of the transistor M1 and the biasing current source. Therefore, in the real case, $I_{out} = -I_{in} + \delta I$ where δI indicates the error made by the sampling process. A lot of techniques to cope with these deficiencies have been treated in literature [9]. In this design cascode techniques [10] have been used for the sampling transistor and the current source to decrease the output conductance. Other limitations and their treatment are described in the following.

4.1. Linearity of the current memory cell

The influence of charge injection of the sampling switch is inherent in any sampling system. The contribution of charge injection to an error of the sampled output current can be divided into a constant offset and a signal dependent part: $\delta I = \delta I^{const} + \delta I^{sig}$. In the readout architecture discussed in section 3, the constant offset becomes non-relevant because it can be compensated by an adjusted threshold in the current compare. Contrary, the signal dependent charge injection δI^{sig} will cause a nonlinearity in the transfer function of the memory cell and should, therefore, be reduced to a minimum.

To reduce δI^{sig}, the circuit in Fig. 5 is used. The sampling process is realized by two successive memory stages, a coarse and a fine one [11]. At any time switches S2 and S3 as well as switches S4 and S5 are driven complementary as indicated

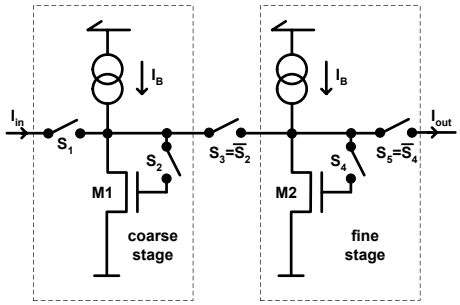

Figure 5. Basic principle of the two stage current memory cell.

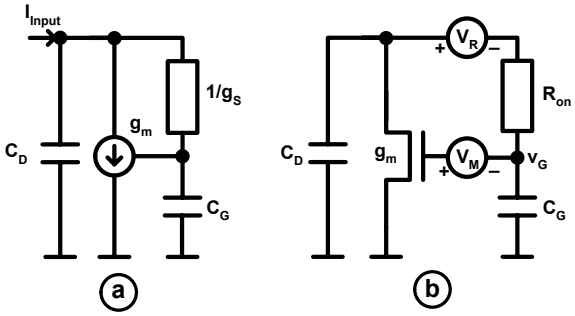

Figure 6. (a) Equivalent circuit of the current memory cell during the charging phase, (b) Circuit used for the noise analysis.

in the figure. The initial state of the switches are: S1, S2 and S4 closed. After the first sampling step (S2 is opened) the resulting current through M1 is $I_{M1} = I_{in} + I_{Bias} - \delta I_c$, where $\delta I_c = \delta I_c^{const} + \delta I_c^{sig}$ indicates the total error (including charge injection) made by the coarse stage. As the input current is still connected to the circuit, the resulting output current of the coarse stage is δI_c, which becomes the input current of the succeeding fine stage. After the second sampling step (S4 is opened), the current through M2 is $I_{M2} = \delta I_c + I_{Bias} - \delta I_f$, where $\delta I_f = \delta I_f^{const} + \delta I_f^{sig}$ is the fine stage's error. Finally, the input switch S1 is opened resulting in an output current $I_{out} = -I_{in} + \delta I_f$.

The error of the coarse stage δI_c does no longer contribute to the sampled output. Even though the constant parts of the errors δI_c and δI_f are similar their signal dependent parts are different because the input range of the fine stage δI_c is much smaller than the input range of the coarse stage I_{in}. Since the signal dependent part of charge injection causes the nonlinearity of the cell, it can be reduced by an order of magnitude by the double stage sampling.

Due to the inverting feature of the memory cell itself, a total cancellation of charge injection can be achieved if two successive memory cells are used, like the proposed readout architecture in section 3. (The first for the pedestal subtraction and the second in the FIFO memory). The output current of this package is $I_{out} = -(-I_{in} + \delta I_{f1}) + \delta I_{f2}$. Assuming that the charge injection of both fine stages are the same ($\delta I_{f1} = \delta I_{f2}$), a total cancellation of charge injection is achieved. This is possible because the charge injection of the fine stages are dominated by the constant fraction and not by the signal dependent part and should, therefore, be similar. However, a total cancellation is not crucial for the readout architecture, as mentioned before.

4.2. Settling behavior

For a high speed application, the settling behavior of the circuit has to be considered as well. During the sampling process, a sufficient time is required to charge the gate capacitance of the storage transistor. Otherwise the residual voltage difference will cause a settling error in the sampled output. Fig. 6a shows the small signal equivalent circuit of the current memory cell during the charging phase. C_D is the capacitive load at the input node (usually dominated by a bus capacitance), C_G the gate capacitance of the memory transistor, $1/g_s$ the on-resistance of the sampling switch and g_m the transconductance of the memory transistor. The transfer function of this circuit is of second order:

$$H(s) = \left(1 + s\frac{C_D + C_G}{g_m} + s^2\frac{C_G C_D}{g_m g_s}\right)^{-1}. \qquad (1)$$

The influence of the memory transistor channel resistance has been neglected as it is greatly reduced by a cascode circuit. Depending on

$$Q = \frac{\sqrt{\frac{g_m}{g_s} C_D C_G}}{C_D + C_G},$$

the time response of the circuit can either be overdamped ($Q < 0.5$), underdamped ($Q > 0.5$) or

critically damped ($Q = 0.5$). Even if the fastest settling time can be achieved with the underdamped case, the parameters in the design were chosen for an overdamped case in order to avoid oscillation in the circuit.

Contrary to a conventional voltage sampling circuit, the charging time does not only depend on the switch resistance $1/g_s$ and the capacitive load (C_G and C_D) but also on the g_m of the storage transistor. Therefore a faster settling time can be obtained by a high transistor g_m while keeping a large storage capacitance for high accuracy. Of course, this scenario is limited by the transistor geometry which links g_m and C_G together. For the two stage sampling process mentioned in section 4.1, two settling times have to be considered. However, t_{fine} can be chosen much larger than t_{coarse}, as the error of the coarse stage does not directly contribute to the output.

4.3. Noise Analysis

Another limitation of the current memory cell's performance is the sampling noise. It originates from a variation of the gate voltage of the memory transistor during the sampling phase caused by different noise sources in the circuit. At the end of the sampling phase, the noise voltage is stored at the gate capacitance and results in a current noise at the output. Fig. 6b shows the different noise sources considered in this analysis. The transistor noise itself is modelled by the voltage source V_M and the noise contribution of the switch resistance is given by V_R. This analysis can focus on thermal noise only for the following reasons. First, large area transistors (A$\sim$100 μm^2) were used for sampling, secondly the large bandwidth of the circuit emphasizes the thermal noise contribution and finally, 1/f noise is suppressed by the inherent correlated double sampling performed by the memory cell [12]. Therefore, a white noise spectrum is assumed for the noise sources modelled by $\langle V_R^2 \rangle = 4kTR_{on}$ for the switch resistance and $\langle V_M^2 \rangle = \frac{8}{3}kTg_m^{-1}$ for the sampling transistor itself. Calculating the amplitude of the filtered gate voltage leads to a total rms noise for the output current of

$$\langle i_{out}^2 \rangle = g_m^2 \langle v_G^2 \rangle = g_m^2 \frac{kT}{C_G} \left(\frac{\frac{2}{3}C_G + C_D}{C_G + C_D} \right). \qquad (2)$$

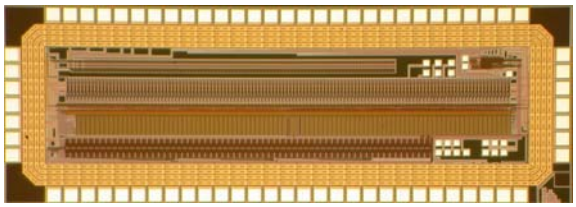

Figure 7. Microphotograph of the 1.5 x 4 mm^2 prototype chip in 0.25 μm radiation tolerant CMOS technology.

Like in a voltage sample and hold, the switch resistance R_{on} does not influence the sampling noise. For all possible drain and gate capacitances, the bracket in (2) is between $\frac{2}{3}$ and 1. Neglecting the bracket, the sampling noise becomes approximately $\langle i_{out}^2 \rangle \approx g_m^2 \frac{kT}{C_G}$ and is almost independent of C_D. With the design parameters $g_m \approx 370$ μS and $C_G \approx 620$ fF, a sampling noise of 24 nA - 30 nA (depending on the drain capacitance) is expected. Note that even in the case of the double staged cell mentioned in section 4.1, the sampling noise is still given by (2) because the error (including the sampling noise) of the coarse stage is corrected by the fine stage and therefore does not contribute to the output. The noise contribution of the biasing current source has been neglected in this calculation. It can simply be added to the sampling noise if it becomes considerably high.

5. Results

A first prototype chip implementing all major building blocks for a fast DEPFET readout scheme for TESLA has been manufactured using a 5 metal 0.25 μm CMOS technology with a radiation tolerant design. Note that the radiation tolerant layout imposes several constraints on the transistor parameters [13]. Fig. 7 shows a microphotograph of the 1.5 x 4 mm^2 prototype chip. The current comparator and the digital hit finder have been successfully tested up to 50 MHz.

The measured linearity of the analog frontend, consisting of 2 successive current memory cells with the regulated input cascode, is given in Fig. 8. The linearity decreases for currents larger

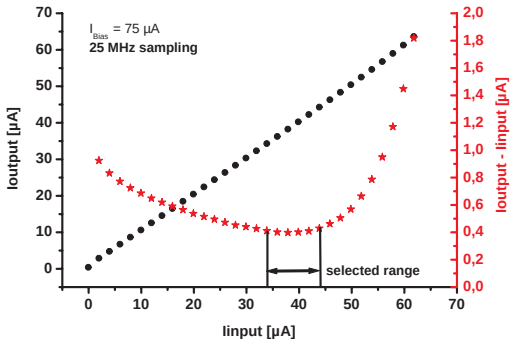

Figure 8. Linearity of the complete analog front-end. The points indicate the output current, the stars indicate the deviation between input and output current (differential non-linearity).

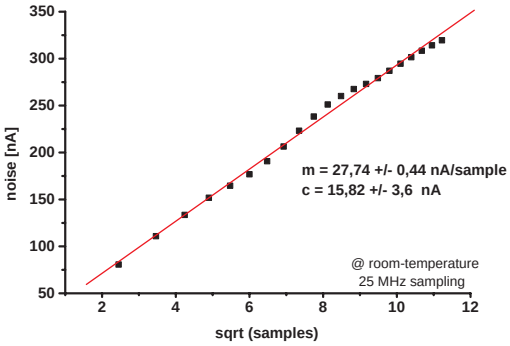

Figure 9. Total sampling noise as a function of the square root of the number of sampling steps.

than 50 μA, as the input becomes of the same order as the bias current used in the cell. Therefore, a tradeoff between static power consumption and dynamic range has to be made. For a DEPFET readout, a dynamic range of about 10 μA is sufficient. This range corresponds to 10000 electrons, assuming a DEPFET gain of 1 nA per collected electron in the internal gate. Fig. 8 shows an accuracy of the complete analog stage for the required dynamic range of better than 0.1% at 25 MHz sampling frequency.

To analyze the noise contribution of the memory cell, a queue of several successive cells has been implemented in the chip. Therefore the number of samples that are made before the output current is measured can be varied over a large range. Fig. 9 shows the measured total noise at the output as a function of the square root of the number of sampling steps. A linear behavior, as expected from independent statistical contributions, is observed. A fit extracts a sampling noise of less than 28 nA (ENC=28 e^-) per sample, which agrees well with the calculated value of 24-30 nA, hence confirming that the noise of the current memory cell is mainly given by the thermal noise contributions of the sampling process as assumed in the noise analysis. For the proposed readout at TESLA with two memory cells (one for the pedestal subtraction and one in the analog FIFO), a total noise of about 40 e^- would have to be added in quadrature to the sensor noise.

6. Conclusion

A readout concept for a DEPFET pixel-based vertex detector matching the high speed requirements at TESLA has been developed. The proposed concept is completely based on current mode techniques well suited to the current operation mode of the DEPFET sensor. The architecture of the readout chip offers correlated double sampling of the input signal and a triggerless zero suppression. The core of the readout architecture, a fast current memory cell, has been analyzed in detail. The results of a prototype chip using a radiation tolerant design show that the requirements are achievable. A full readout chip for a DEPFET array is being designed at present.

REFERENCES

1. TESLA - Technical Design Report, Part4: A detector for TESLA, ISBN 3-935702-00-0, 2001
2. J. Kemmer and G. Lutz, New detector concepts, Nucl. Inst. Meth. **A253** (1987) 365
3. J. Ulrici et al., Spectroscopic and imaging performance of DEPFET pixel sensors, Nucl. Inst. Meth. **A465** (2000) 247
4. R.H. Richter et al., Design and technology of

DEPFET pixel sensors for linear collider applications, these proceedings
5. W. Neeser et al., The DEPFET-Pixel Bioscope, IEEE Trans., Nucl.Sci 47 No.3 (2000)
6. A DEPFET based pixel vertex detector for the detector at TESLA LC-Note, LC-DET-2002-004, DESY April 2002
7. P. Fischer, H. Krüger, G. Gomes, Mephisto - a 128 channel strip readout chips with real-time data sparsification and multi-hit-capability, Nucl. Inst. Meth. **A431** (1999) 134
8. J. B. Hughes, et al., Switched Currents - A New Technique for Analog Sampled Data Signal Processing, Proc. of Int. Symp. Circuits and Systems (ISCAS), Portland, Oregon, May 1989, pp. 1584–1587
9. B.E. Jonsson, Switched-Current Signal Processing and A/D Conversion Circuits, ISBN 0-7923-7871-7
10. T. Loelinger, W. Guggenbühl, Cascode Configurations for Switched Current Copiers, Analog Integrated Circuits and Signal Processing, 19, 115-127 (1999)
11. C. Toumazou, S. Xiou, An N-Step charge injection cancellation scheme for switched current circuits, Electronics Letters, vol.30 pp.680-681 (1994)
12. J.B. Hughes, Analogue Techniques for Very Large Scale Integrated Circuits, PHD dissertation, Southampton 1992
13. G. Anelli et al., Radiation Tolerant VLSI Circuits in Standart Deep Submicron Technologies for the LHC Experiments: Practical Design Aspects, IEEE Trans., Vol.46 No.6 p.1690 (1999)