

# Current-Mode Analog Adaptive Mechanism for Ultra-Low-Power Neural Networks

Rafal Dlugosz, Tomasz Talaska, and Witold Pedrycz, *Fellow, IEEE*

**Abstract**—Neural networks (NNs) implemented at the transistor level are powerful adaptive systems. They can perform hundreds of operations in parallel but at the expense of a large number of building blocks. In the case of analog realization, an extremely low chip area and low power dissipation can be achieved. To accomplish this, the building blocks should be simple. This brief presents a new current-mode low-complexity flexible adaptive mechanism (ADM) with a strongly reduced leakage in analog memory. Input signals ranging from 0.5 to 20  $\mu\text{A}$  are held for 10–50 ms, with the leakage rate from 0.2%/ms to 0.04%/ms, respectively, depending on temperature. A small storage capacitor of 200 fF enables a short write time ( $< 100$  ns). A single ADM cell occupies 1400  $\mu\text{m}^2$  when realized in the Taiwan Semiconductor Manufacturing Company Ltd. CMOS 0.18- $\mu\text{m}$  technology. The potential application of this NN is envisioned in a mobile platform based on a wireless sensor network to be used for online analysis of electrocardiography signals.

**Index Terms**—Adaptive mechanism (ADM), analog memory (AM), current mode, hardware neural networks (NNs), leakage compensation.

## I. INTRODUCTION

ANALOG memory (AM) elements are used in various applications, where different features are important and, thus, different optimization techniques are required. For instance, for analog-to-digital converters (ADCs), features such as speed, write precision that allows for high signal resolutions, and linearity are the key design requirements [1]. Since only one analog cell is typically used in this case, even a relatively complex circuit is acceptable. Analog filters operating in the megahertz range require even several dozen AM cells, and therefore, simpler circuits have to be used, whereas a long storage time  $T_{\text{ST}}$  is not critical. A different situation arises in hardware-realized neural networks (NNs) that contain even hundreds of weights (connections). Since the weights are seldom updated, a low  $LR$  is required, whereas power dissipation

Manuscript received April 19, 2010; revised August 24, 2010; accepted October 22, 2010. Date of current version January 19, 2011. This paper was recommended by Associate Editor R. Genov.

R. Dlugosz is with the Faculty of Telecommunication and Electrical Engineering, University of Technology and Life Sciences in Bydgoszcz, 85-796 Bydgoszcz, Poland, with the Department of Computer Science, Poznan University of Technology, 60-965 Poznan, Poland, and also with the Institute of Microtechnology, Swiss Federal Institute of Technology in Lausanne, 2000 Neuchâtel, Switzerland (e-mail: rafal.dlugosz@epfl.ch).

T. Talaska is with the Faculty of Telecommunication and Electrical Engineering, University of Technology and Life Sciences in Bydgoszcz, 85-796 Bydgoszcz, Poland (e-mail: talaska@utp.edu.pl).

W. Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: pedrycz@ece.ualberta.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSII.2010.2092827

and chip area must be low as well, particularly in the case of applications of NNs to the analysis of biomedical signals. For example, electrocardiography (ECG) signals are sampled at low  $f_S$ , usually well below 2 kHz [2].

Different AM cells have been proposed in the literature so far, but these realizations are not suitable for the application in large NNs. For example, an ultralong  $T_{\text{ST}}$  (in minutes) has been achieved in the AM proposed in [3], but a very complex structure, and thus a chip area larger than 0.05  $\text{mm}^2/\text{cell}$  and low  $f_S$  of 100 Hz, makes the use of this circuit inconvenient. In the current-mode circuit reported in [4], for the signals varying from 0.25 to 0.75  $\mu\text{A}$ ,  $LR = 1.56$  nA/ms, i.e., from 0.2%/ms to 0.6%/ms, but a large  $C_{\text{ST}}$  of 3.5 pF is unacceptable in large NNs.

Our prototype NN with built-in analog ADM blocks, which is the subject of this brief, is 90% composed of analog circuits arranged into 12 channels working in parallel [7]. The underlying idea is to use as many channels working in parallel as possible while simplifying the structure of a single channel. As a result, we arrived at a structure that, in comparison with a similar network recently realized by us on the microcontroller, offers the performance about 1000 times better. The analog NN sampled at 1 MHz dissipates the power of 700  $\mu\text{W}$ , whereas the  $\mu\text{C}$ -based network dissipates 380 mW at an  $f_S$  of 490 kHz.

These observations make the NN presented here suitable for the application in portable ECG systems [2] that can be realized as a wireless sensor network (WSN). The ECG signals are considered to be nonlinear and nonstationary, with a large variation in morphologies of different patients. Therefore, such signals require advanced analysis and classification techniques. In the literature, one can find many approaches of employing NNs to the analysis of such signals [5], [6]. To the best of our knowledge, the existing solutions are always software based. Our objective is to use the realized NN in ultra-low-power portable diagnostics devices. The network will be used directly in particular sensors of the WSN for the analysis of the ECG complexes, thus significantly reducing the amount of data being transferred between the sensors and the base station to save energy.

The prototype NN is composed of several functional blocks. One of them is a distance calculation circuit (DXW) that is used to determine the distance between the vector of the input signals  $X$  and the weight vectors  $W$  of the individual neurons. The winner-takes-all (WTA) circuit is then used to detect the winning neuron, i.e., the neuron with the smallest distance. The conscience mechanism [7] reduces the quantization error, thus improving the overall learning properties of the NN.

## II. IMPLEMENTATION ISSUES OF AM CELLS

As has been discussed, different tradeoffs exist in the AM design. One of them is the tradeoff between the data write time

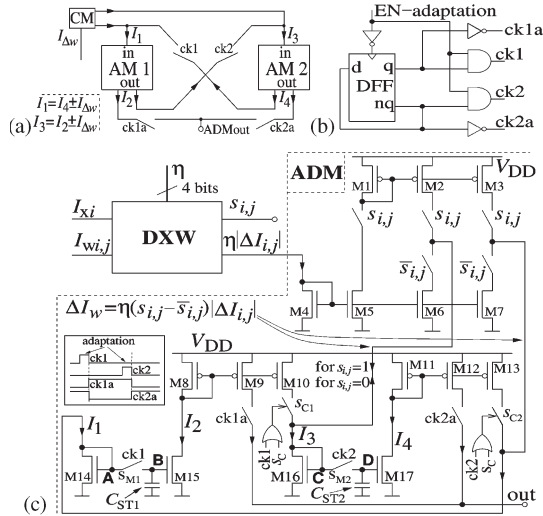


Fig. 1. Proposed ADM. (a) The idea. (b) Internal clock circuit activated by an adaptation enable (EN) signal. (c) Detailed diagram.

$T_W$  and the  $LR$ . A short  $T_W$  is achievable for small values of both the storage capacitor  $C_{ST}$  and the  $ON$ -resistance  $R_{ON}$  of the switches associated with  $C_{ST}$ . On the other hand, a low  $LR$  requires larger values of  $C_{ST}$ , but this reduces the data rate and enlarges the chip area [3], [4].

Various techniques have been proposed to overcome  $LR$ , including refreshing the memory [8] by the use of an ADC, storing weights as digital signals and then converting them to analog signals [9], [10], or just by increasing the value of  $C_{ST}$  [3], [4]. All these techniques require additional circuitry that increases both power dissipation and chip area and, therefore, are inconvenient in the case of a large number of AM cells. One of the possibilities here is adapt the weights in the NN with  $f_S$  high enough to compensate for the losses of information, but this solution is not efficient in low-data-rate systems. In light of this, we propose a new ADM circuit with a significantly reduced  $LR$  even for small values of  $C_{ST}$  and, thus, being very fast and chip area efficient.

### III. PROPOSED ADM

The proposed ADM is shown in Fig. 1. The underlying adaptation process is described as follows:

$$\underbrace{w_{i,j}(n+1)}_{I_{w_{i,j}}(n+1)} = \underbrace{w_{i,j}(n)}_{I_{w_{i,j}}(n)} + \eta \cdot (s_{i,j} - \bar{s}_{i,j}) \cdot \underbrace{|x_i(n) - w_{i,j}(n)|}_{|\Delta I_{i,j}|}. \quad (1)$$

This circuit cooperates with the DXW block used to calculate the current proportional to an absolute value of a term  $x_i - w_{i,j}$ , in which  $x_i$  is the  $i$ th network input, whereas  $w_{i,j}$  is the  $i$ th weight of a winning  $j$ th neuron. In this case, both  $x$  and  $w$  signals are represented as the corresponding currents. The current mode simplifies the overall NN structure since the summation is performed at junctions. A comparator built-in in the DXW block returns the logical signal  $s_{i,j} = \text{sign}(x_i - w_{i,j})$ . The resulting current  $|I_{\Delta w}| = \eta \cdot |\Delta I_{i,j}|$  is then used to update the value of the corresponding weight.

The proposed ADM consists of two current-mode AM cells that operate alternately [19]. Depending on the value of  $s_{i,j}$ , the

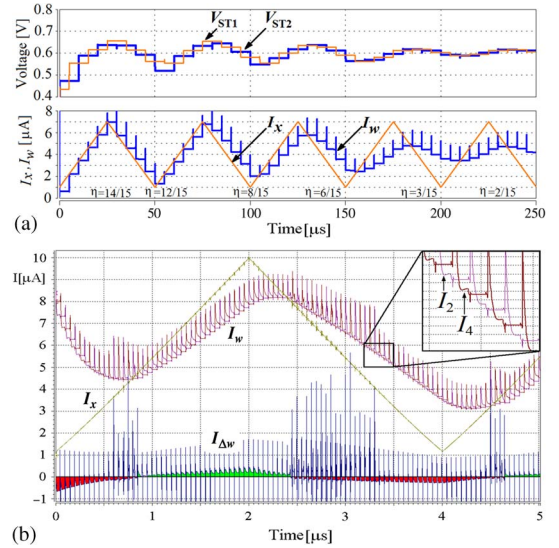


Fig. 2. Performance of the proposed circuit (transistor-level simulations). The adaptation process (a) for different values of the learning rate  $\eta$  and the sampling frequency  $f_S = 200$  kHz and (b) for  $f_S = 20$  MHz and  $\eta = 0.1$ .

current  $\eta \cdot |\Delta I|$  is either added or subtracted from the previous value of the weight stored in one cell. The resulting current is then stored in the second cell. Each ADM block is controlled by an individual clock, as shown in Fig. 1(b), and activated only during the adaptation. After the adaptation, this block switches the output of the second cell to the output of the entire ADM. If a given neuron wins again, the role of both cells is reversed. As aforementioned, low chip area was one of the main requirements, and therefore, simple memory cells have been used, although  $S^2I$  and  $S^3I$  solutions were considered as well.

### IV. NEGATIVE EFFECTS IN THE PROPOSED ADM BLOCK

#### A. Charge Injection Effect

One of the ways to make the ADM faster is to minimize the time constant for charging the storage capacitor  $C_{ST}$ , which, in practice, means the necessity of minimizing the  $ON$ -resistance of the switches  $S_M$  by increasing their channel width, but this increases the charge injection effect across the  $C_{ST}$  capacitors. In the proposed ADM, this negative effect has been strongly minimized by the use of dummy switches DS (not shown in Fig. 1), which are controlled by clock signals of opposite polarity.

The proposed circuit operates in the current mode, but the signals are stored across  $C_{ST}$  capacitors as voltages. In the AM working in the voltage mode, a  $\Delta V_{ST}$  error introduced by the charge injection effect has different values depending on the value of  $V_{ST}$  and, therefore, is difficult to be compensated in a wide range of  $V_{ST}$ . A different situation arises in the case of the current-mode memory cell. Fig. 2 (top) shows the adaptation process realized for different values of the learning rate  $\eta$ .

The input current  $I_x$  (a triangle waveform) vary in between 1 and 8  $\mu A$  in this case, whereas the resultant  $V_{ST}$  voltage varies only moderately between 0.45 and 0.65 V (see the first 100  $\mu s$ ). This case occurs only at the beginning of the learning process for the values of  $\eta$  being close to 1. Then, as  $\eta$  becomes smaller, the variation of  $V_{ST}$  voltage becomes small as well and is much easier to compensate. This feature can be considered as an advantage of the current-mode approach visible in this case.

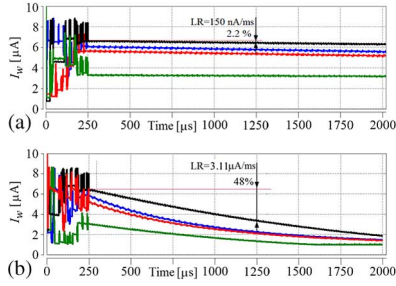


Fig. 3. Simulations of selected weights for the uncompensated ADM circuit for (a)  $T_E = 0$  °C and (b)  $T_E = 80$  °C. The adaptation process ends at 250  $\mu$ s.

### B. Current Leakage Effect

Another negative effect is the leakage current. This current is composed of two components. One of them is the current  $I_{DB}$  that flows through the drain-to-bulk diodes of turned-off transistors used in the switches. We have reduced this effect by using transmission gates in the switches, in which the  $I_{DB,p}$  and  $I_{DB,n}$  currents for the PMOS and NMOS transistors, respectively, are counterbalanced [3]. The problem with this technique is that a given junction area ratio between both types of transistors minimizes the leakage for only one particular value of  $V_{ST}$ , i.e., for one particular value of the input current. To make the proposed circuit flexible and suitable for different currents, n-wells of the PMOS transistors in the switches are supplied separately, which enables the control of this effect. Transistor sizes have been carefully selected to minimize the variation of  $V_{ST}$  even for large  $\Delta I_x$ , which significantly minimizes this component of the leakage current. This issue is studied later in more detail (see Fig. 7).

The subthreshold leakage is a dominant leakage component. It results from a finite OFF-resistance  $R_{OFF}$  of the switches  $S_{M1}$  and  $S_{M2}$  and strongly depends on the difference between the potentials at both ends of these switches. In the proposed circuit, an influence of this effect can be observed in the case when the compensation is turned off [19], i.e., for  $s_C = "0"$  (see Fig. 1). Some results concerning this case for selected environment temperatures  $T_E$  are shown in Fig. 3.

In this case, the configuration switches  $S_{C1,2}$  controlled by an external  $s_C$  signal are closed only during the adaptation, i.e., for either ck1 or ck2 signals equal to "1." As a result, after the adaptation (for ck1 = 0 and ck2 = 0), the currents  $I_1$  and  $I_3$  are small ( $\approx |I_{\Delta w}|$ ), and the potentials at points A and C are small as well. The resulting  $LR$  causes the losses of information in between 2.2%/ms and 48%/ms for  $T_E$  varying from 0 °C to 80 °C. To keep this error below some level (e.g., 2%), the ADM should refresh the information stored in the AM at  $f_S$  between 1.1 and 24 kHz. The refreshment rate for an example network with 30 neurons should be in between 33 and 720 kHz, depending on  $T_E$ . We note that the ECG signals require  $f_S$  from 1 to 2 kHz.

In the discussed circuit, this leakage component has been strongly reduced by minimizing the difference between the potentials across the switches  $S_{M1}$  and  $S_{M2}$ . To enable this, the switches  $S_{C1}$  and  $S_{C2}$  must be closed permanently, which is achieved for  $s_C = "1"$ . In this case, the currents  $I_1$  and  $I_3$  differ from the currents  $I_2$  and  $I_4$ , respectively, by 0%–10% only, depending on the value of the input current  $\Delta I_w$ . This current linearly depends on the value of the learning rate  $\eta$ , which is the highest at the beginning of the learning process but then

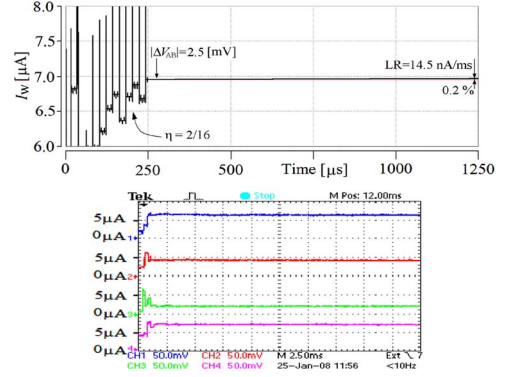


Fig. 4. Adaptation process for the compensated ADM. (a) Simulations and (b) measurements both for  $T_E = 60$  °C. The process has been turned off for  $\eta = 2/16$ . The difference of potentials  $\Delta V_{AB}$  across the switch  $S_{M1}$  of 2.5 mV enlarged the  $LR$  by 3.6 nA/ms, i.e., by 0.05%.

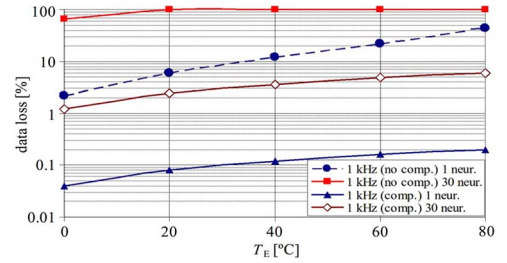


Fig. 5. Data loss in the proposed ADM versus the environment temperature  $T_E$  for the compensated and uncompensated circuits, for a single neuron, and for a network with 30 neurons.

monotonically decreases to zero. Finally, the currents  $I_1$  to  $I_4$  become almost equal, which levels the potentials in points A, B, C, and D. Due to the leakage effect compensation, it is possible to achieve a storage time up to 10–50 ms even for relatively small values of the  $C_{ST}$  capacitors of only 200 fF in this case.

Using small  $C_{ST}$  capacitors shortens  $T_W$  to below 100 ns, as shown in Fig. 2 (bottom). As a result, the achievable data rate of an entire NN can be as high as 4 MHz, taking into account the delay caused by the other blocks in the calculation process.

The results for the compensated circuit for  $T_E = 60$  °C are shown in Fig. 4. The value of the  $LR$  varies rather moderately in between 0.04%/ms and 0.2%/ms for  $T_E$  and in between 0 °C and 80 °C, respectively. Very similar results have been obtained in both the postlayout simulations and the measurements. In the latter case, the circuit was tested between 0 °C and 60 °C.

To demonstrate how the  $LR$  affects the NN performance, the results of the comparative study are shown in Fig. 5 for the compensated and uncompensated ADMs, for a single neuron, and for an NN with 30 neurons sampled at 1 kHz, which are treated as a function of the environment temperature  $T_E$ . The losses of information for the human body temperature are about 2%–3% for 30 neurons. This is acceptable in the classification of the ECG complexes performed by the use of the WTA NNs.

It is worth noting that, since the currents  $\eta \cdot \text{sign}(x - w) \cdot |\Delta I|$  in particular training epochs can be either positive or negative with a mean value equal to zero, the  $V_A - V_B$  and  $V_C - V_D$  terms can be either positive or negative, and thus, the signal stored in the AM either decreases or increases due to the subthreshold leakage. As a result, the resultant error does not accumulate along the adaptation process, and the problem becomes less significant.

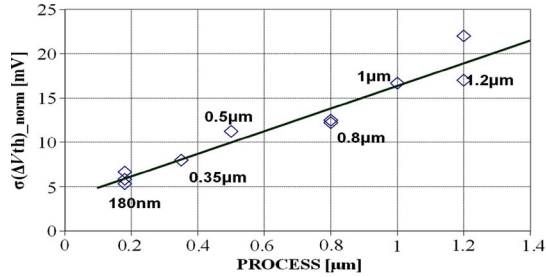


Fig. 6.  $\sigma(\Delta V_{TH})$  normalized for  $W \cdot L = 1 \mu\text{m}^2$  versus the CMOS technology.

### C. Influence of the Mismatch Effect on the Leakage Rate

The proposed ADM can operate with the supply voltage that is in the range of 1–1.8 V with the currents that are at the level of 0.5–20  $\mu\text{A}$ . Depending on the value of the input current, the corresponding gate potentials of transistors in the AM are either below or above the threshold voltage  $V_{TH}$ , which for the CMOS 0.18- $\mu\text{m}$  process is approximately equal to 0.45 V.

One of the important aspects in the overall design process of the NN was to select a suitable technology. As the network is mainly composed of analog current-mode circuits, whereas the ADM is based on the current mirrors (CMs), the mismatch between transistors must be taken into consideration.

The threshold voltage mismatch  $\Delta V_{TH}$  is usually assumed to be the main mismatch component, and therefore, in this brief, we mostly limit the discussion to this effect. An analysis of the current factor mismatch  $\Delta\beta$  shows that this component exhibits a relatively small influence on the behavior of the proposed ADM circuit. For transistor sizes that have been selected in the design, it contributes to the overall error at the level of 0.3% [21]. The influence of the channel length modulation effect is visible mostly at the beginning of the learning process when particular  $V_{DS}$  voltages have different values. In this case, the influence of this effect does not exceed 0.4% [22]. At the end of the learning process, this effect can be neglected as the currents  $I_1$ ,  $I_2$ ,  $I_3$ , and  $I_4$ , as well as the  $V_{DS}$  voltages, become equal.

It can be observed that the value of  $\Delta V_{TH}$  for given sizes of transistors depends on the process. This issue, which is called technology characterization, has been intensively studied for different CMOS technologies [11]–[14]. Both mismatch components are typically shown as the standard deviation versus the reverse of the square root of the transistor gate area [11], [21], i.e.,

$$\sigma(\Delta V_{TH})[\text{mV}] = f(1/\sqrt{W \cdot L})[1/\mu\text{m}]. \quad (2)$$

Since this function is approximately linear, it is possible to normalize the results for a given gate area, e.g.,  $W \cdot L = 1 \mu\text{m}^2$ , which allows for a direct comparison of different technologies. The comparative results are shown in Fig. 6.

Although some improvement is visible when moving toward newer processes, it is rather limited. For example, considering the 0.8- and 0.18- $\mu\text{m}$  processes, the difference is only 50%.

On the basis of the voltage and current values visible in the waveforms of Fig. 2, the signal distortion introduced by the mismatch in the proposed ADM can be evaluated.

Assuming that both transistors in the CM have equal sizes, then according to  $\Delta V_{TH}$ ,  $\Delta\beta$ , and  $\Delta\lambda$ , the following formulas

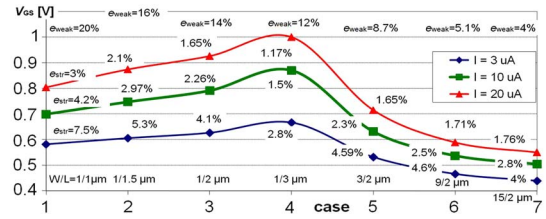


Fig. 7. Influence to the  $V_{TH}$  mismatch on the gain error of the CM with equal transistors.

can be provided for the weak and strong inversion regions [15]:

$$G_{\text{weak}} = \frac{I_2}{I_1} \approx \left(1 + \frac{\Delta\beta}{\beta}\right) \exp\left(\frac{-\Delta V_{TH}}{V_T}\right) \times \left(1 + \Delta\lambda \frac{V_{DS}}{1 + \lambda V_{DS}}\right) \quad (3)$$

$$G_{\text{str}} = \frac{I_2}{I_1} \approx \left(1 + \frac{\Delta\beta}{\beta}\right) \frac{(V_{GS} - V_{TH} - \Delta V_{TH})^2}{(V_{GS} - V_{TH})^2} \times \left(1 + \Delta\lambda \frac{V_{DS}}{1 + \lambda V_{DS}}\right). \quad (4)$$

The currents  $I_1$  and  $I_2$  are the input and output signals of the CM, respectively, whereas  $V_T$  is the thermal voltage (26 mV).

For the CMOS 0.18- $\mu\text{m}$  process, the average normalized value of  $\Delta V_{TH}$  is 5.8 mV. The case study for selected transistor sizes and several constant values of  $I_1$ , i.e., 3, 10, and 20  $\mu\text{A}$  for the NMOS-type CM, is shown in Fig. 7. For a given value of  $I_1$ , the value of the gate-to-source voltage  $V_{GS}$  strongly depends on transistor sizes.  $e_{\text{weak}}$  and  $e_{\text{str}}$  are resultant gain errors of the CM with theoretically equal transistors. In the strong inversion region, the improvement resulting from increasing transistor sizes is limited as the  $V_{GS}$  voltage decreases for the constant current. The minimum value of  $e_{\text{str}}$  is for  $W/L = 1/3 \mu\text{m}$ , but the variation of  $V_{GS}$ , i.e., the  $V_{ST}$  voltage, is the largest in this case, and the drain-to-bulk leakage is more difficult for the compensation. For this reason, the sizes of transistors in the CMs that carry the currents  $I_1$  to  $I_4$  have been chosen to be  $3/2 \mu\text{m}$ . In this case, for the input currents in between 3 and 10  $\mu\text{A}$ ,  $\Delta V_{ST} = 0.1$  V. The gain error of the PMOS-type CMs is smaller in this case since the  $V_{GS}$  voltage is larger (0.8–0.9 V). As a result, an average value of  $e_{\text{str}}$  in the loop  $I_1$  to  $I_4$  is equal to about 2.6%.

A different situation is for the CMs carrying the current  $|I_{\Delta w}|$ . These currents are relatively small (refer to Fig. 2, bottom), and therefore, these mirrors operate mainly in the weak inversion region. For this reason, to minimize the value of the gain error  $e_{\text{weak}}$ , transistor sizes have been chosen to be equal to  $9/2 \mu\text{m}$ .

Considering the presented values, we implemented the NN in the CMOS 0.18- $\mu\text{m}$  process and assumed that the network will operate in the strong inversion region, i.e., with the input currents larger than 3  $\mu\text{A}$ , as the expected improvement from using newer technologies is relatively small in this case.

The aforementioned loop is composed of four CMs; therefore, if all mirrors would introduce an error with the equal sign, as well as the maximum absolute value, which is a theoretical case, then for  $\eta = 0$ ,  $I_1/I_2 \approx G_{\text{str}}^4$ , and resultant error would be 10% for transistor sizes of  $3/2 \mu\text{m}$  working in the strong inversion. This error would cause a difference between the potentials at both ends of the switches  $S_{M1,2}$  of about 5 mV,

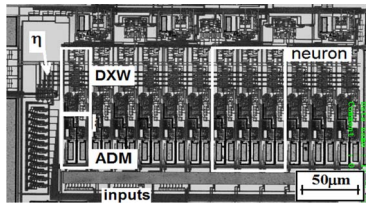


Fig. 8. Prototype self-organizing WTA NN with four neurons (four outputs) and three inputs implemented in the Taiwan Semiconductor Manufacturing Company Ltd. CMOS 0.18- $\mu\text{m}$  technology.

TABLE I  
COMPARISON OF EXAMPLE AM CELLS

Ref	Tech. [μm]	$f_s$ [MHz]	$P$ [μW]	$C_{ST}$ [pF]	Area [mm <sup>2</sup> ]	LR [%/ms]	FOM $P/f_s$ [pJ]
[1]	0.18	165	600	ND	ND	N/D	3.6
[3]	1.5	100 [Hz]	0.01	6	0.0535	10 aA/s	100
[4]	0.8	0.025	6.75	3.5	0.045	0.6	270
[16]	0.6	1.28	1980	ND	0.167	N/D	1546
[17]	0.5	10	10	0.2	0.0016	0.006	1
[18]	0.35	30	3000	0.3	ND	N/D	100
This work (*)	0.18	20 MHz	2–100	0.2	0.0014	0.06 (at room $T_E$ )	2 (**)

(\*) data for entire adaptive mechanism (ADM) with two AM cells  
(\*\*) for  $f_s=15$  MHz,  $I_{\text{average}}=5$  μA,  $V_{DD}=1.5$  V, with the compensation

thus increasing the leakage current by about 40%. In practice, this error is smaller. In the measurements, we did not observe such high influence.

The prototype NN with four neurons and three inputs (12 ADM blocks) realized in the CMOS 0.18- $\mu\text{m}$  process is shown in Fig. 8. A single ADM, together with a DXW block, occupies an area of  $120 \times 25$   $\mu\text{m}$ . The power dissipation of a single ADM block is equal to about  $2 \cdot V_{DD} \cdot (I_1 + I_3)$  and varies in between 2 and 100  $\mu\text{W}$ . It depends on the values of the currents and the  $V_{DD}$  voltage. The current  $I_{Dw}$  is very small and does not contribute significantly in the overall power dissipation. In the case when the NN is sampled at a relatively high  $f_s$ , the compensation can be turned off. In this case, the currents  $I_1$  and  $I_3$  are almost zero, which reduces power dissipation by approximately half.

The comparison of the proposed ADM with other AM cells is shown in Table I. A direct comparison is not easy since particular solutions are addressed for different applications and come with different requirements. The figure-of-merit (FOM) has been defined as a ratio of the power dissipation  $P$  and the maximum sampling frequency, which is the reverse of the storage time. For a more complete comparison, the FOM should have included the signal resolution [20]. However, this information is not available in most of the cited papers.

The lowest leakage of  $1\text{E}-07\%/ms$  has been achieved in [3] but at the expense of a large chip area and low  $f_s$ . Our ADM offers a flexible solution with a wide range of  $f_s$ , whereas it offers a very simple structure with a low chip area. It should be noticed that we have presented the results not only for a single AM but also for the entire ADM with two cells.

## V. CONCLUSION

A new analog current-mode ADM with a leakage-compensated AM has been discussed in the context of the application to ultra-low-power NNs realized in hardware. In comparison with other solutions, the proposed circuit features a very small chip area, low power dissipation, and storage time sufficiently long for the application of the NN in the analysis of low-data-rate biomedical signals. The circuit exhibits very

high flexibility when it comes to the sampling frequency, which can vary from 20 Hz (200 Hz for temperatures above 80 °C) to about 20 MHz.

## REFERENCES

- [1] O. Rajaei, A. Jahani, and M. S. Bakhtiar, "A low voltage, high speed, high resolution class AB switched current sample and hold," in *Proc. IEEE Int. Symp. Circuits Syst.*, Kos, Greece, May 2006, pp. 1039–1042.
- [2] J. J. Segura-Juárez, D. Cuesta-Frau, L. Samblas-Pena, and M. Aboy, "A microcontroller-based portable electrocardiograph recorder," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1686–1690, Sep. 2004.
- [3] M. O'Halloran and R. Sarpeshkar, "A 10-nW 12-bit accurate analog storage cell with 10-aA leakage," *IEEE J. Solid-State Circuits*, vol. 39, no. 11, pp. 1985–1996, Nov. 2004.
- [4] J. A. De Lima and A. S. Cordeiro, "A low-voltage low-power analog memory cell with built-in 4-quadrant multiplication," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 4, pp. 191–195, Apr. 2003.
- [5] O. T. Inan, L. Giovannardi, and G. T. A. Kovacs, "Robust neural network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2507–2515, Dec. 2006.
- [6] L. He, W. Hou, X. Zhen, and C. Peng, "Recognition of ECG patterns using artificial neural network," in *Proc. Int. Conf. Intell. Syst. Des. Appl.*, Jinan, China, Oct. 2006, vol. 2, pp. 477–481.
- [7] R. Dlugosz, T. Talaska, W. Pedrycz, and R. Wojtyna, "Realization of the conscience mechanism in CMOS implementation of winner-takes-all self-organizing neural networks," *IEEE Trans. Neural Netw.*, vol. 21, no. 6, pp. 961–971, Jun. 2010.
- [8] D. Macq, M. Verleysen, P. Jespers, and J.-D. Legat, "Analog implementation of a Kohonen map with on-chip learning," *IEEE Trans. Neural Netw.*, vol. 4, no. 3, pp. 456–461, May 1993.
- [9] I. Baturone, S. Sanchez-Solano, and J. L. Huertas, "Self-checking current-mode analogue memory," *Electron. Lett.*, vol. 33, no. 16, pp. 1349–1350, Jul. 1997.
- [10] C.-Y. Wu and W.-K. Kuo, "A new analog implementation of the Kohonen neural network," in *Proc. Int. Symp. VLSI Technol. Syst. Appl.*, Taipei, Taiwan, 1993, pp. 262–266.
- [11] J. A. Croon, M. Rosmeulen, S. Decoutere, W. Sansen, and H. E. Maes, "An easy-to-use mismatch model for the MOS transistor," *IEEE J. Solid-State Circuits*, vol. 37, no. 8, pp. 1056–1064, Aug. 2002.
- [12] T. Serrano-Gotarredona and B. Linares-Barranco, "CMOS transistor mismatch model valid from weak to strong inversion," in *Proc. Eur. Solid-State Circuits Conf.*, Estoril, Portugal, 2003, pp. 627–630.
- [13] A. Van den Bosch, M. Steyaert, and W. Sansen, "The extraction of transistor mismatch parameters: The CMOS current-steering D/A converter as a test structure," in *Proc. IEEE Int. Symp. Circuits Syst.*, Geneva, Switzerland, May 2000, vol. 2, pp. 745–748.
- [14] S. J. Lovett, M. Welten, A. Mathewson, and B. Mason, "Optimizing MOS transistor mismatch," *IEEE J. Solid-State Circuits*, vol. 33, no. 1, pp. 147–150, Jan. 1998.
- [15] M. Conti, G. F. Dalla Betta, S. Orcioni, G. Soncini, C. Turchetti, and N. Zorzi, "Test structure for mismatch characterization of MOS transistors in subthreshold regime," in *Proc. IEEE Int. Conf. Microelectron. Test Struct.*, Monterey, CA, 1997, vol. 10, pp. 173–178.
- [16] E. de Lira Mendes, P. Loumeau, and J.-F. Naviner, "A switched-current sample and hold circuit for low frequency applications," in *Proc. IEEE Int. Symp. Circuits Syst.*, Sydney, Australia, May 2001, vol. 1, pp. 452–455.
- [17] R. Carmona, S. Espejo, R. Dominguez-Castro, A. Rodriguez-Vazquez, T. Roska, T. Kozek, and L. O. Chua, "A 0.5  $\mu\text{m}$  CMOS CNN analog random access memory chip for massive image processing," in *Proc. IEEE Int. Workshop Cellular Neural Netw. Appl.*, London, U.K., Apr. 1998, pp. 271–276.
- [18] Y. Sugimoto, "A realization of a below-1-V operational and 30-MS/s sample-and-hold IC with a 56-dB signal-to-noise ratio by applying the current-based circuit approach," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 1, pp. 110–117, Jan. 2004.
- [19] T. Talaska, R. Dlugosz, and W. Pedrycz, "Adaptive weights change mechanism for Kohonen's neural network implemented in CMOS 0.18  $\mu\text{m}$  technology," in *Proc. ESANN*, Bruges, Belgium, Apr. 2007, pp. 151–156.
- [20] F. Goodenough, "Analog technologies of all varieties dominate ISSCC," *Electron. Des.*, vol. 44, no. 4, pp. 96–111, Feb. 1996.
- [21] R. Difrenza, P. Linares, and G. Ghibaudo, "A new model for the current factor mismatch in the MOS transistor," *Solid State Electron.*, vol. 47, no. 7, pp. 1167–1171, Jul. 2003.
- [22] D. Binkley, *Tradeoffs and Optimization in Analog CMOS Design*. New York: Wiley, 2008, pp. 157–159.