

Circuits for On-Chip Learning in Neuro-Fuzzy Controllers

Fernando Vidal-Verdú*, Rafael Navas* and Angel Rodríguez-Vázquez**

*Dept. de Electrónica. Universidad de Málaga,
Complejo Tecnológico sn, 29071-Málaga, SPAIN
FAX #34 5 2132781, email: vidal@ctima.uma.es

**Dept. of Analog Design, CNM, Edificio CICA,
C/ Tarfia sn, 41012-Sevilla, SPAIN
FAX #34 5 4624506, email: angel@cnm.us.es

Abstract

Learning algorithms have become of great interest to be applied not only to neural or hybrid neuro-fuzzy systems, but also as a tool to achieve a fine tuning of analog circuits, whose main drawback is their lack of precision. This paper presents accurate, discrete-time CMOS building blocks to implement learning rules on-chip. Specifically, a voltage mode high precision comparator as well as an absolute value circuit. These blocks, plus multiplexing in time techniques, are used to build a circuit to determine the polarity of the learning increments. Compactness and low power consumption have been considered main requirements, since they are essential to increase the complexity of the neural systems. An example circuit has been simulated with HSPICE with the parameters of a 1 μ m CMOS technology. Statistical variations of technological parameters were considered. The results show that all curves from 30 runs of a Monte Carlo analysis behave as expected, and at least 8 bits of resolution are achieved by the proposed techniques.

1. Introduction

Fuzzy systems are useful to solve control problems where the plant is ill-defined or very complex to be modeled by means of its related mathematical equations. Instead, a model-free estimation is made to capture the intuitive human knowledge expressed through if-then rules. Another widespread used model-free paradigms are the neural networks, which try to solve the problem through learning procedures. A great effort

is being done currently to join both approaches into the so-called neuro-fuzzy systems. Such approach produces systems that are transparent to the human thinking and, at the same time, are able to evolve under learning algorithms. Fig. 1(a) shows a particular case of the ANFIS neuro-fuzzy architecture, that corresponds to a 'singleton' implementation, while Fig. 1(b) shows the microphotography of a chip designed in a 1 μ m CMOS technology that implements it [1]. Inside this architecture, the first and fourth layers contain programmable blocks, while the remaining are formed by fixed nodes. The system global response $y = F(\mathbf{x}, \mathbf{w})$ for a given input $\mathbf{x} = (x_1, x_2, \dots, x_M)$ is determined by the set of programming parameters \mathbf{w} in the adaptive layers. Fig. 1(c) shows a set of curves that correspond to an implementation of the adaptive nodes in a 1.6 μ m CMOS standard technology [1].

This ability to accept learning procedures compensates the lack of structured knowledge or allows to correct errors due to the hardware implementation in the analog approach, when the learning rules are implemented on-chip or with the chip in-the-loop. The Fig. 1(d) depicts a typical supervised learning loop, where the learning rules act as a teacher that changes the system response, but also tunes the circuits that implement the algorithm. Therefore, we should be very careful in designing the circuits that implement them, because they are supposed to be more accurate than the underlying error-prone circuitry.

Implementations based on a pure analog approach obtain up to 9 bits of resolution in CMOS digital standard technologies [9]. This is achieved at the expense of a very high area consumption. Another way to warrant precision consists in implementing the learning circuitry with digital techniques, and interface with the analog system through A/D and D/A converters of the required resolution. Obviously, this strategy involves also large circuitry. Thus, both previous approaches are not suitable for on-chip implementation of learning, specially in the case of parallel learning rules, where compactness is essential. However, it is possible to use mixed signal techniques, which are used in the implementation of analog to digital converters, to reduce the area and power consumption. Very accurate circuits for updating as well as to store the weights have already been proposed [3][4]. In this paper, we discuss strategies to design precise circuits to implement the learning rule. Specifically, a compact and precise circuit to evaluate the polarity of the learning increments, which is the most crucial part for a successful learning [3][5], is proposed.

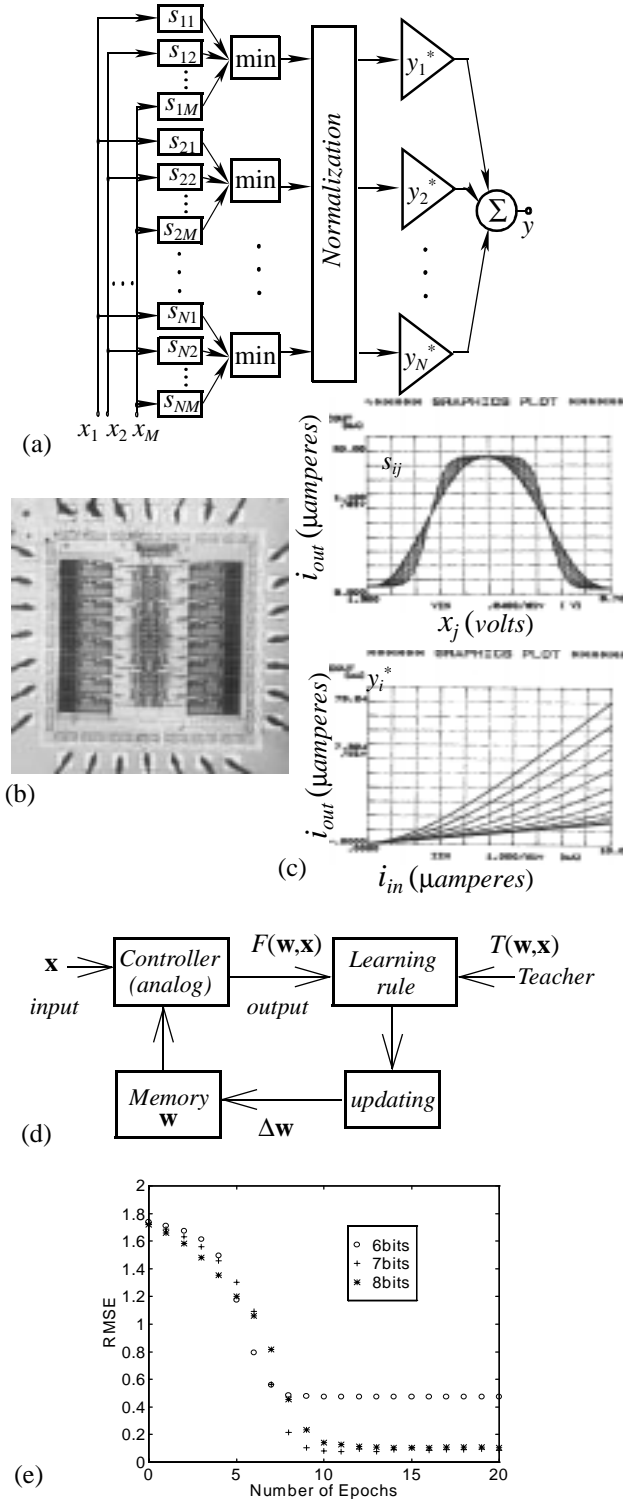


Fig. 1 (a) Neuro-Fuzzy controller architecture; (b) microphotography of an implementation; (c) curve families associated to the adaptive nodes; (d) supervised learning loop and (e) associated RMSE curves.

2. Polarity circuit architecture

$$\mathbf{w} = \{w_1, w_2, \dots, w_N\}$$

$$\Delta w_i$$

(for $i = 1 \dots N$) in the perturbative algorithms [2][3] as,

$$\Delta w_i = -\zeta [E(w_i) - E(w_i + \text{pert})] \quad (1)$$

where E , in an incremental process (the parameters \mathbf{w} are updated each time a new input is presented) is usually

$E = |F(\mathbf{w}, \mathbf{x}) - T(\mathbf{w}, \mathbf{x})|^2$ and ζ is a constant. The strongest restriction for successful learning is the computation of the sign of (1) [3], because an error in the sign will force an increment of w_i in the wrong direction. Let us define the step function

$$S(\Delta w_i) = \begin{cases} 1 & \text{if } \Delta w_i > 0 \\ 0 & \text{if } \Delta w_i < 0 \end{cases} \quad (2)$$

A circuit that implements (2) can be used in learning circuitry whose weight update building block uses as input a digital signal that provides the polarity of the increments [3][5]. Note that

$$\begin{aligned} S(\Delta w_i) &= S[E(w_i + \text{pert}) - E(w_i)] = \\ &= S[|F_p - T_p|^2 - |F - T|^2] = S[|F_p - T_p| - |F - T|] \end{aligned} \quad (3)$$

where the index p in (3) means that F or T have been evaluated with perturbed weights. The last equality in (3) is due to $|z|^v$ as a monotone increasing function of $z \forall v \geq 1$, thus $|z_1|^v \geq |z_2|^v \forall v \geq 1$ if $z_1 \geq z_2$, and $S(|z_1|^v - |z_2|^v)$ does not depend on v . As a consequence, S in (2) can be calculated with the architecture in Fig. 2. In the following, we will propose strategies to implement the building blocks in Fig. 2 with mixed signal techniques to get a precise, compact circuit. Fig. 1(e) shows some learning curves corresponding to the controller in Fig. 1(a) in a learning loop [1]. The curves were obtained by modelling the controller at transistor level and simulating the whole loop in a computer. Successful con-

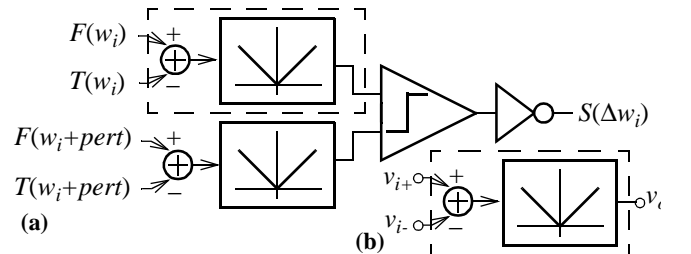


Fig. 2 Architecture of a circuit that provides the polarity of the learning increment (a); adder-plus-absolute value block (b).

vergence is achieved for resolutions in the parameters of up to seven bits. In the following, we will propose circuitry to implement Fig. 2 that is able to cope with this requirement besides of consuming low area and power.

3. Adder plus absolute value circuitry

The first block to implement in Fig. 2(a) is the adder plus absolute value block in Fig. 2(b), which computes v_o as,

$$v_o = \begin{cases} (v_{i+} - v_{i-}) & \text{if } v_{i+} \geq v_{i-} \\ -(v_{i+} - v_{i-}) & \text{if } v_{i+} \leq v_{i-} \end{cases} \quad (4)$$

This corresponds to a full-wave rectification operator that can be built like Fig. 3 depicts, where half-wave rectifiers are

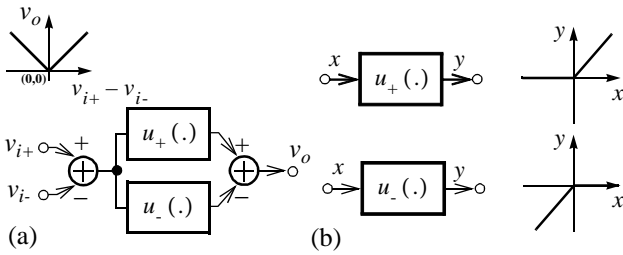


Fig. 3 Full-wave rectifier (a) and half-wave rectifier blocks (b).

defined as,

$$y = u_+(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & , \text{otherwise} \end{cases} \quad (5)$$

and

$$y = u_-(x) = \begin{cases} x & \text{if } x \leq 0 \\ 0 & , \text{otherwise} \end{cases} \quad (6)$$

Usual mechanisms to implement such operators exploit that the output current in diodes and current mirrors is negligible for positive (or negative) input currents. For implementations where the output voltage carries information, the voltage drop in the diodes introduces an error that is divided by the gain of an amplifier in a feed-back loop in the so-called superdiodes. Another approach exploits the large resistance and zero offset voltage of an analog switch in the off state, and the use of comparators to encode digitally the sign of the input signal. An example of the first strategy is depicted in Fig. 4(a) for currents [6], while Fig. 4(b) illustrates the second strategy for a transresistance circuit.

. Full-wave rectification should provide a very good matching between the positive $p+$ and negative $p-$ pieces of the output curve (see top of Fig. 4), in the sense that they should be identical, but with opposite first derivatives. Note that otherwise, the precision of the further comparison in Fig. 2(a) would be severely degraded. Reported full-wave rectifiers in voltage and current mode usually use different signal paths

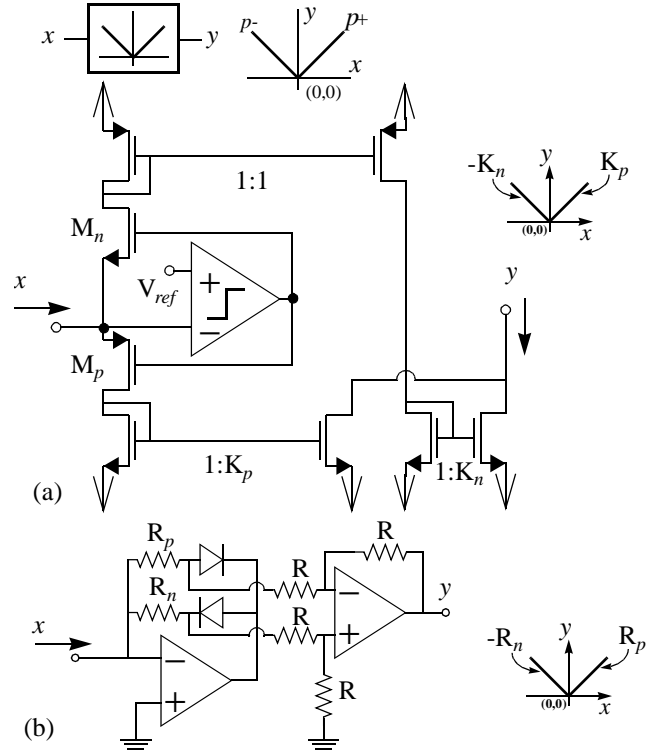


Fig. 4 Transresistance (a) and current mode (b) implementations of a full-wave rectifier.

for positive and negative inputs, thus matching between $p+$ and $p-$ depends strongly on device matching. For instance, $u_-(x)$ and $u_+(x)$ are obtained directly as drain currents of the transistors M_n and M_p respectively in Fig. 4(a), but further processing is necessary to provide the output current. This processing is readily carried out by means of current mirrors. However, each current mirror introduces an error due mainly to the finite output resistance and the mismatching between input and output transistors, which adds offset and gain errors in every reflection. Any other solution in current mode will present a similar drawback. Fig. 4(b) shows a transresistance circuit with diodes as rectification operators and OPAMPs. Here, the main error sources are the voltage offset of the amplifiers and the mismatching among the resistances.

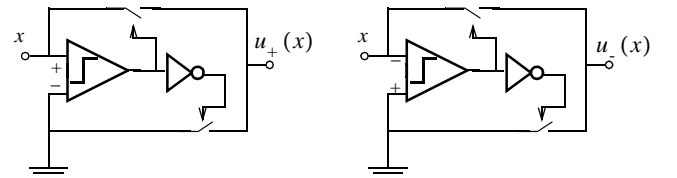


Fig. 5 Half-wave rectifiers with analog switches.

In order to reduce the error, the number of devices in the signal path should be as low as possible. Note that the implementation of $u_-(x)$ and $u_+(x)$ in Fig. 5 do not have any device but analog switches, which have zero offset

voltage, in the signal path. Hence, the only source of error is the offset of the comparator. To extend this strategy to a circuit that performs the full-wave rectification, a simple approach is depicted in Fig. 6, where the inversion of the input signal is made by providing proper control signals to the current switches, which is got by means of simple digital circuitry. Let us define an analog demultiplexor as in Fig. 6(a). Two analog demultiplexors like this and one comparator can be used to build the desired block as Fig. 6(b) depicts. The comparator provides a digital signal c whose value is 1 for positive and 0 for negative input values. This signal controls the two analog demultiplexors that create the proper signal paths to ensure that the output is always positive. Fig. 6(c) shows a very simple implementation of the analog demultiplexors with analog switches and digital gates. A similar strategy is followed for rectification in voltage-charge domain [7].

Note that an adder is necessary at output of Fig. 6(d) to provide a single-end output. Considerations to build this adder are closely related to the strategy followed in the design of the polarity circuit, thus it will be described in the section 5.. At this point, the comparator circuit is the only source of error in the form of an offset at input which equals the offset of the comparator. In the following section, we propose a voltage comparator to implement that in Fig. 6 and fulfils the requirements of high precision and compactness.

4. Comparator circuit

As said above, the comparator determines the resolution of Fig. 6. Thus, accurate comparators are needed in Fig. 6 and Fig. 2 in order to get a successful learning. Open loop

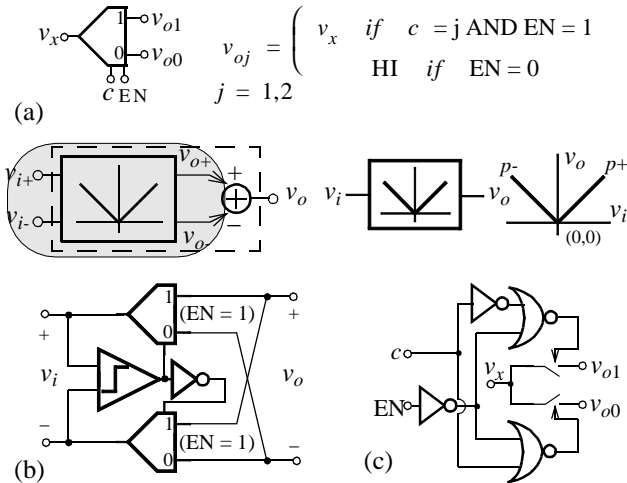


Fig. 6 Analog demultiplexor (a); fully-differential absolute value circuit (b) and implementation of the analog demultiplexor (c).

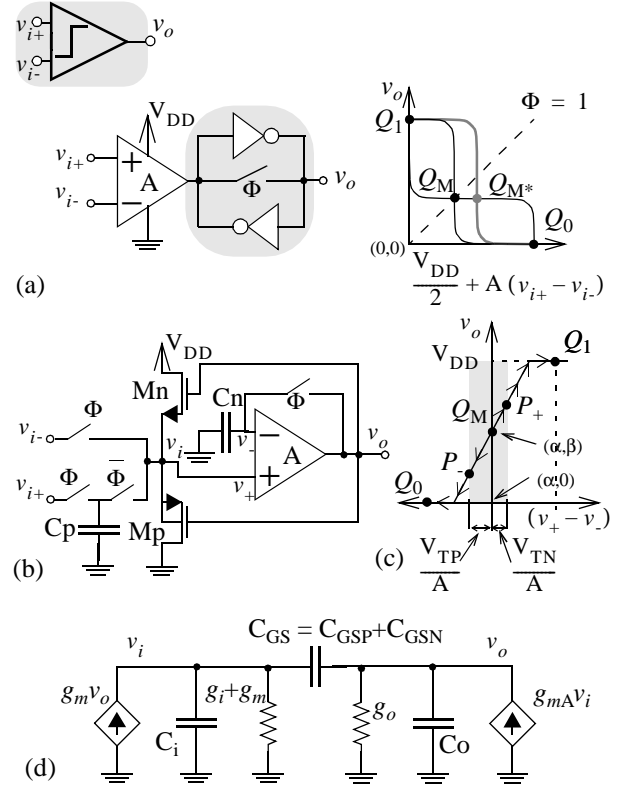


Fig. 7 Voltage comparator based on a latch (a); proposed comparator circuit (b); large signal behavior (c) and simplified small-signal model (d).

operational amplifiers can be used as voltage comparators. However, to enhance speed and facilitate output interfacing, a regenerative sense amplifier is a better option. A common implementation of such circuit uses a *latch* and a differential amplifier as front-end circuit to get a differential input [8]. This circuit is depicted in Fig. 7(a), where a digital signal Φ is used to reset the latch. In a perfect matching situation, for $\Phi=1$ the latch is forced to be in the meta-stable state Q_M . However, mismatches place this state in a point out of the *input=output* line (Q_{M^*}). This limits the achievable resolution to about 5 bits for the single latch shaded in Fig. 7(a). To improve the resolution, front-end amplifier gain is increased, thus the latch offset is divided by this gain. This approach has two main drawbacks:

- Large gains are needed for the front-end amplifier, thus high area and power consumption.
- The offset of the front-end amplifier remains, thus the final offset is,

$$V_{off} = \frac{V_{off, LATCH}}{A} + V_{off, AMPLIFIER} \quad (7)$$

As a consequence of both previous points, large area consumption is required to reduce the offset in (7). Fig. 7(b) presents a comparator based on a regenerative amplifier that overcomes the previous inconveniences. The circuit works as

follows. For $\Phi=1$, the amplifier acts as a voltage follower due to the negative feedback loop. Note that sources and gates of the transistors Mn and Mp are at the same voltage, thus the transistors are cut-off and the circuit has a high impedance input. The voltage v_{i-} is then presented at input, and due to the negative feedback loop the following value is stored in Cn,

$$v_{0-} = \frac{A(v_{i-} + E_{OS})}{A+1} \quad (8)$$

where A and E_{OS} are the gain and the voltage offset of the amplifier respectively. In addition, the input v_{i+} value is also stored in Cp. The circuit remains in Q_M (see Fig. 7(c)) as long as the voltage value v_{i-} remains at input. This point is defined by its coordinates,

$$\alpha = \frac{v_{i-} + E_{OS}}{A+1}, \quad \beta = \frac{A(v_{i-} + E_{OS})}{A+1} \quad (9)$$

When the phase signal changes to $\Phi=0$, the amplifier works in a *positive feedback loop* because of the gate to source capacitors associated to the transistors Mn and Mp, and the amplifier output changes in the sense of taking Mn or Mp out of the cut-off region. This conclusion is reached by performing small signal analysis in the simplified small signal model for Fig. 7(b) depicted in Fig. 7(d). Note that only one transistor, Mn or Mp, can be out of the cut-off region, thus g_m equals the small signal transconductance of this transistor. In the central region of Fig. 7(c), both transistors are cut-off, thus we can consider $g_m \approx 0$ and $g_i \approx 0$. Analysis on this circuit provides the following pole,

$$s = \frac{g_{mA}C_{GS} - g_o(C_i + C_{GS})}{C_o(C_i + C_{GS})} \quad (10)$$

Hence the circuit is not stable as long as $g_o(C_i + C_{GS}) < g_{mA}C_{GS}$.

The transistor Mn will enter in saturation for an increase in the non inverting input voltage of the amplifier, while the transistor Mp will do it for a decrease of this voltage. Such conditions are readily translated to $v_{i+} \geq v_{i-}$ for an increase $v_{i+} \leq v_{i-}$ for a decrease. The gain of the *positive feedback loop* is reinforced by Mn or Mp as soon as any of them enters the saturation region, and the circuit evolves quickly toward Q_1 in the former case and toward Q_0 in the latter, thus

$$Q_{final} = \begin{cases} Q_1 & \text{if } v_{i+} \geq v_{i-} \\ Q_0 & \text{if } v_{i+} \leq v_{i-} \end{cases} \quad (11)$$

Mismatching of transistors Mn and Mp with respect to ideal ones changes basically the width of the shaded region of Fig. 7(c). This does not affect the resolution of the circuit as long as Q_M is not a stable point.

Note that the circuit performance is not affected at first order by the offset of the amplifier. Limitations are due mainly to

the charge injected in the capacitors in the ON-OFF transition of the analog switches. This error is reduced by increasing the size of the capacitors, as well as by means of dummy transistors in the analog switches. On the other hand, special care must be taken to avoid overshooting at the non-inverting input of the amplifier at the beginning of the evaluation phase, because it could make that the wrong transistor (Mn or Mp) enters in saturation and the circuit evolves toward an incorrect final state. This effect is a consequence of the hysteresis associated to the positive feedback, and it is minimized by enlarging the size of the shaded region in Fig. 7(c). The example comparator of this paper is built with the amplifier, capacitors and analog switches depicted in Fig. 8. Despite small devices are used, the resolution of this comparator is more than 8 bits, measured from 30 runs of a Monte Carlo transitory analysis.

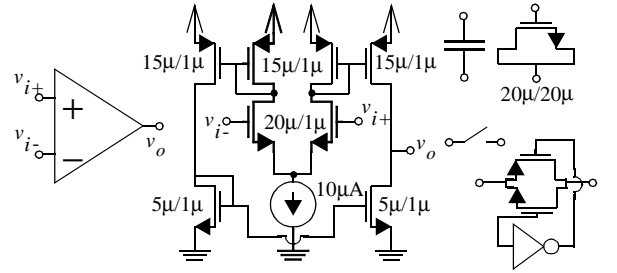


Fig. 8 Implementations of the Amplifier, the capacitors and the analog switches in Fig. 7.

5. The polarity circuit

The Fig. 9 depicts the final implementation of the polarity circuit in Fig. 2(a), where the absolute value building block at the input is implemented as explained in section 3.. Note that it has two inputs besides of the differential inputs. The input Φ corresponds to the phase signal of the comparator in the absolute value circuit of Fig. 6(b), because the comparator is implemented as explained in the previous section (see Fig. 7(b)). On the other hand, the enable input EN corresponds to that in the analog demultiplexors of Fig. 6. Signals at these inputs are depicted in Fig. 9. The computation is finished after 4Δ . For $0 \leq t < 2\Delta$, comparisons for the proper operation of the analog demultiplexors are made, but only the outputs of the top input block (T) is presented at the adder input, because the bottom block (B) has high impedance outputs ($EN_B = \overline{\Phi_2} = 0$). For $2\Delta \leq t < 3\Delta$, $|F - T|$ is stored in the capacitor Cn of the output comparator. For $\tau = 3\Delta$ the top input block outputs are disabled ($EN_T = \Phi_2 = 0$), while the bottom input block outputs are enabled ($EN_B = \overline{\Phi_2} = 1$), and $|F_p - T_p|$ is presented at the comparator input. Thus, the comparison of the two previously obtained absolute values

is carried out (note that multiplexing in time and enable signals allow to save the capacitor C_p and the analog switches in Fig. 7(a)).

As regards the adder, a very simple way to implement it is proposed in Fig. 9(a). Since a subtraction is required, a differential amplifier with unity gain can be used. Fig. 9(b) consists of an OTA loaded by a resistor and a current source. The resistor performs the I/V conversion and the current source shifts the output to adapt the output range to the input range of the following circuit.

Fig. 9(c) shows the OTA implementation of the example circuit in this paper with transistor sizes and resistor and current source values. The sources of the transistors in the differential pair of Fig. 9(g) are degenerated with resistors to enhance the linearity of the response curve. The resistors in Fig. 9 can be implemented in standard technologies with transistors or using polysilicon, diffusion or well sheets. Ideal resistors have been considered for the simulations of the example circuits of this paper, because the adder circuit is shared by both adder-plus-absolute value circuits in Fig. 2, thus mismatching is not going to affect the result. This strategy also allows the use of small transistors in the implementation of the OTA. Sharing of the adder circuit is possible by multiplexing the circuit in time.

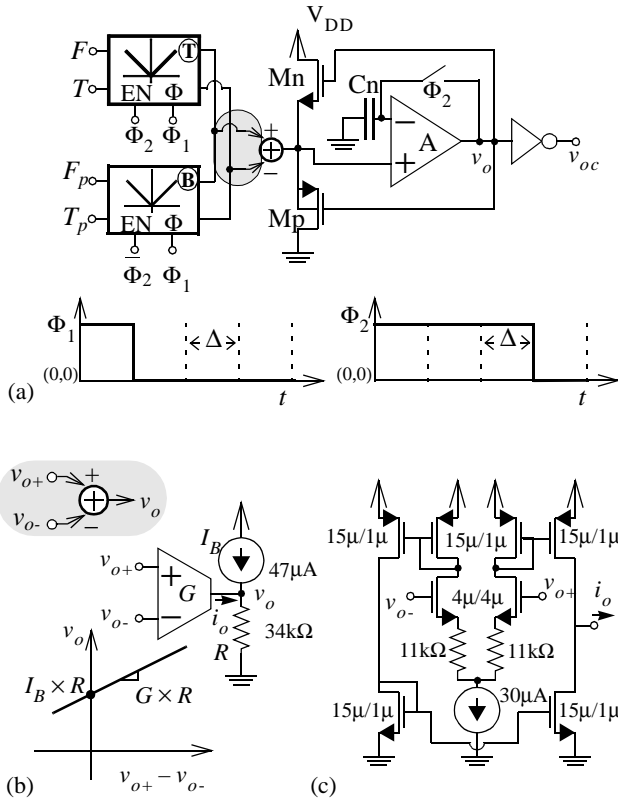


Fig. 9 The polarity circuit (a); adder circuit (b) and CMOS OTA Implementation (c).

6. Results

The Fig. 10 shows some results from HSPICE simulations that illustrate the performance of the presented circuits. The parameter Δ in Fig. 9 equals 100ns in these simulations. Thirty runs of a Monte Carlo analysis were done with an standard n-well CMOS $1\mu\text{m}$ technology. Parameter deviations were modeled as reported in [9], with the values for our technology in Table I. Note that the circuit provides the right value for the 30 Monte Carlo curves for signals to compare that differs in 4mV in a range of 1V. The circuits behave quite well also for smaller differences, and many curves still go on well. These results are obtained in spite of the small devices used, thus obtaining high resolution without degrading compactness.

7. References

- [1] F. Vidal-Verdú, M. Delgado-Restituto, R. Navas-González and A. Rodríguez-Vázquez: "A Building Block Approach to the Design of Analog Neuro-Fuzzy Systems in CMOS Digital Technologies". pp. 357-390 in *Fuzzy Hardware Architectures and Applications*. Kluwer Ac. Pub. 1998.
- [2] M. Jabri and B. Flower, "Weight perturbation: An optimal architecture and learning technique for analog VLSI feedforward and recurrent multilayered networks", *IEEE Trans. on Neural Networks*, Vol. 3, No. 1 pp. 154-157, 1992
- [3] Gert Cauwenberghs: "An Analog VLSI recurrent Neural Network Learning a Continuous-Time Trajectory". *IEEE Trans. on Neural Networks*, Vol. 7, No. 2 pp. 346-361, March 1996.
- [4] Gert Cauwenberghs: "Fault-Tolerant Dynamic Multilevel Storage in Analog VLSI". *IEEE Trans. on Circuits and Systems-II*, Vol. 41, No. 12 pp. 827-829, December 1994.
- [5] A. J. Montalvo, R. S. Gyurcsik and J.J. Paulos: "An Analog VLSI Neural Network with On-Chip Perturbation Learning". *IEEE Journal of Solid-State Circuits*, Vol. 32, No. 4, April 1997.
- [6] A. Rodríguez-Vázquez and M. Delgado-Restituto: "Generation of Chaotic Signals using Current-Mode Techniques". *Journal of Intelligent and Fuzzy Systems*, Vol. 2, pp. 15-37, 1994.
- [7] J.L. Huertas, A. Rodríguez-Vázquez and A. Rueda: "Low-Order Polynomial Curve Fitting using Switched-Capacitor Circuits", *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1123-1125. 1984
- [8] B. Nauta and A. G. W. Venes: "A 70-MS/s 110-mW 8-b CMOS Folding and Interpolating A/D Converter". *IEEE Journal of Solid-State Circuits*, Vol. 30, No. 12, December 1995.

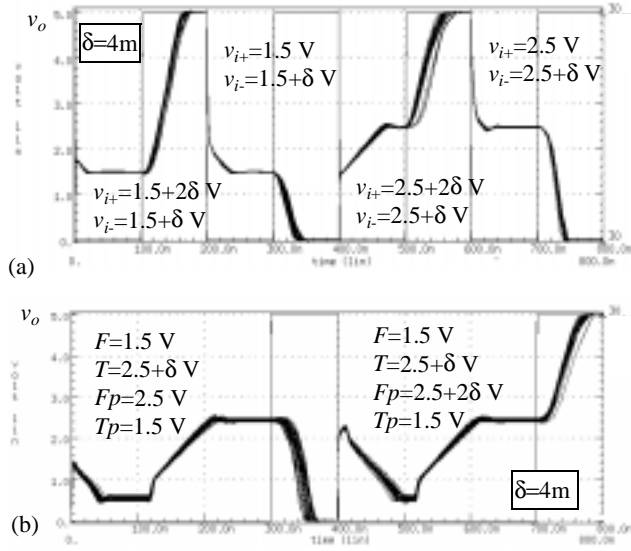


Fig. 10 Comparator output (a) and Non-inverted polarity circuit output (b).

- [9] M.J.M. Pelgrom et al.: "Matching Properties of MOS Transistors". *IEEE J. of Solid-State Circ.*, Vol. 39, pp. 1433-1440, June 1990.

A_{VT0n} (V μ m)	A_{VT0p} (V μ m)	$A_{\beta n}$ (μ m)	$A_{\beta p}$ (μ m)	$A_{\gamma n}$ (V $^{0.5}\mu$ m)	$A_{\gamma p}$ (V $^{0.5}\mu$ m)
12m	14.4m	3.3%	4.5%	6.4m	4.8m

Table I: Proportionality constants of the Pelgrom's model in the technology used.