Mixed Signal CMOS High Precision Circuits for on Chip Learning

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. An exemplary circuit has been simulated with HSPICE with the parameters of a 1μm CMOS technology. Statistical variations of technological parameters were considered.The results show that all curves from 30 runs of a montecarlo analysis behave as expected, and at least 8 bits of resolution are achieved by the proposed techniques .

1. INTRODUCTION

During the last decade, an increasing interest in algorithms that do not require high precision, like neural, fuzzy, or neuro-fuzzy systems, has opened a new field of application to the analog approach. On the other hand, learning rules, that are anyway inherent to the neural systems, can be used to correct errors in analog circuits due to temperature variations, device mismatching, etc. Thus, such 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 and because learning processes need at least seven bits of resolution [1] (this resolution requirement depends on the learning task).

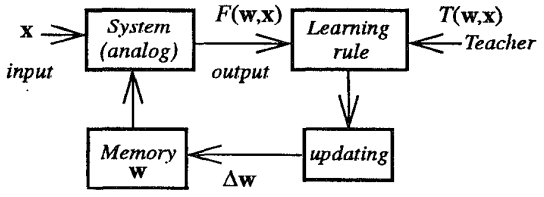

Fig. 1 Typical supervised learning loop

The Fig. 1 depicts a typical supervised learning loop. Implementations based on a pure analog approach obtain up to 9 bits of resolution in CMOS digital standard technologies [8]. 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.

2. POLARITY CIRCUIT ARCHITECTURE

In the system of Fig. 1, the global response is determined by a set of parameters $\mathbf{w} = \{w_1, w_2, \ldots, w_N\}$. Most supervised learning rules are based on a gradient-descent approach to change properly **w**. However, on-chip implementations of derivatives involves error-prone and large circuitry. Finite differences are used instead to calculate Δw_i (for $i = 1 \ldots N$) in the perturbative algorithms [2][3] as,

$$\Delta w_i = -\zeta [E(w_i) - E(w_i + pert)] \qquad (1)$$

where E, in an incremental process (the parameters **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

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

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]. Since $|z|^v$ is a monotone increasing function of z $\forall v \geq 1$, (2) can be calculated with the architecture in Fig. 2. In the follow-

0-7803-4455-3/98/$10.00 © 1998 IEEE

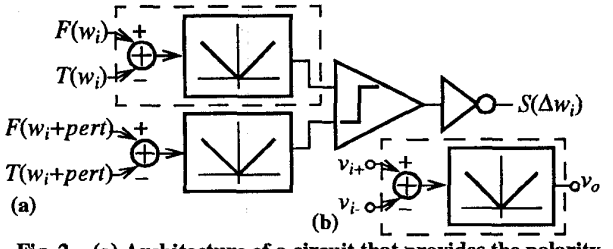

Fig. 2 (a) Architecture of a circuit that provides the polarity of the learning increment; (b) Adder-plus-absolute value block.

ing, we will propose strategies to implement the building blocks in Fig. 2 with mixed signal techniques to get a precise, compact circuit.

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,

$$v_o = \begin{cases} (v_{i+} - v_{i-}) & \text{if } v_{i+} \geq v_{i-} \\ -(v_{i+} - v_{i-}) & \text{if } v_{i+} \leq v_{i-} \end{cases} \quad (3)$$

A straightforward approach to implement (3) consists in connecting a differential amplifier and an absolute value circuit in cascade. This strategy computes first the difference $v_i = v_{i+} - v_{i-}$ and then calculates the absolute value of v_i with a full-wave rectifier like that depicted in Fig. 3(a). However, full-wave rectification should provide a very good matching between the positive ($p+$ in Fig. 3(a)) and negative ($p-$ in Fig. 3(a)) pieces of the output curve, in the sense that they should be identical, but with opposite first derivatives. Note that otherwise, precision of further comparison in Fig. 2(a) would be severely degraded. Reported full-wave rectifiers in voltage and current mode usually use different signal paths for positive and negative inputs, thus matching between $p+$ and $p-$ depends strongly on device matching.

We propose instead the strategy in Fig. 3(b), that uses a fully differential full-wave rectifier as front end circuit, followed by the adder at output. In order to implement the absolute value block in the shaded area of Fig. 3(b), let as

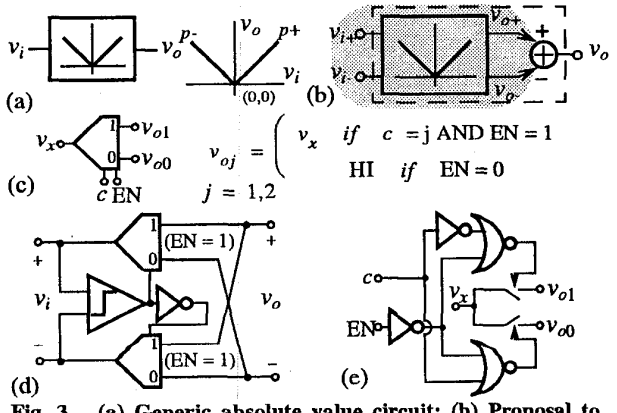

Fig. 3 (a) Generic absolute value circuit; (b) Proposal to implement (3); (c) Analog demultiplexor;(d)Fully differential absolute value circuit; (e) Implementation of the analog demultiplexor.

define an analog demultiplexor as in Fig. 3(c). Two analog demultiplexors like this and one comparator can be used to build the desired block as Fig. 3(d) 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. 3(e) 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 [6]. In the following section, we propose a novel voltage comparator to implement that in Fig. 3(d). Note that this comparator determines the resolution of the circuit.

As regards the adder, a very simple way to implement it is proposed in Fig. 4(a). Since a subtraction is required, a differential amplifier with unity gain can be used. Fig. 4(a) 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.

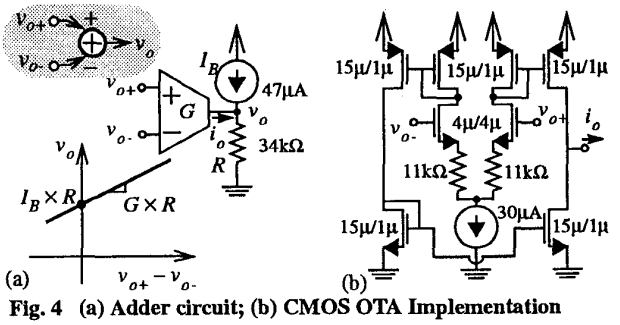

Fig. 4 (a) Adder circuit; (b) CMOS OTA Implementation

Fig. 4(b) shows the OTA implementation of the exemplary circuit in this paper with transistor sizes and resistor and current source values. The sources of the transistors in the differential pair of Fig. 4(b) are degenerated with resistors to enhance the linearity of the response curve. This is important because precision for further comparison is limited by the slope of the curve at bottom of Fig. 4(a), and regions with low first derivative degrade the overall performance. The resistors in Fig. 4 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 exemplary 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.

4. COMPARATOR CIRCUIT

As said above, the comparator determines the resolution of Fig. 3(d). Thus, accurate comparators are needed in Fig. 3 and Fig. 2 in order to get a successful learning. Open loop 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

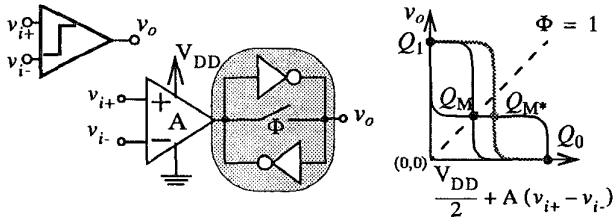

Fig. 5 Voltage comparator based on a latch.

implementation of such circuit uses a *latch* and a differential amplifier as front-end circuit to get a differential input [7]. This circuit is depicted in Fig. 5, where a digital signal Φ is used to reset the latch. In a perfect matching situation, for Φ=1 the latch is forced to be in the meta-stable state Q_M in Fig. 5. 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. 5. 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,

$$V_{off} = \frac{V_{off,\,\mathrm{LATCH}}}{A} + V_{off,\,\mathrm{AMPLIFIER}} \tag{4}$$

As a consequence of both previous points, large area consumption is required to reduce the offset in (4). Fig. 6(a) presents a novel comparator based on a regenerative amplifier that overcomes the previous inconveniences. The circuit works as follows. For Φ=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, thanks to the negative feedback loop, stored in Cn. In addition, the input v_{i+} value is also stored in Cp. The circuit remains in Q_M (see Fig. 6(b)) as long as the voltage value v_{i-} remains at input. When the phase signal changes to Φ=0, the amplifier works in open loop, and the previously stored value of v_{i+} is compared with that of v_{i-} stored in Cn, and the amplifier output changes in the sense of taking Mn or Mp out of the cut-off region. The transistor Mn will enter in saturation for positive differential inputs, while the transistor Mp will do it for negative ones. Note that a *positive feedback loop* is now created with the Mn or Mp transistor and the amplifier, and the circuit evolves toward Q_1 in the former case and toward Q_0 in the latter.

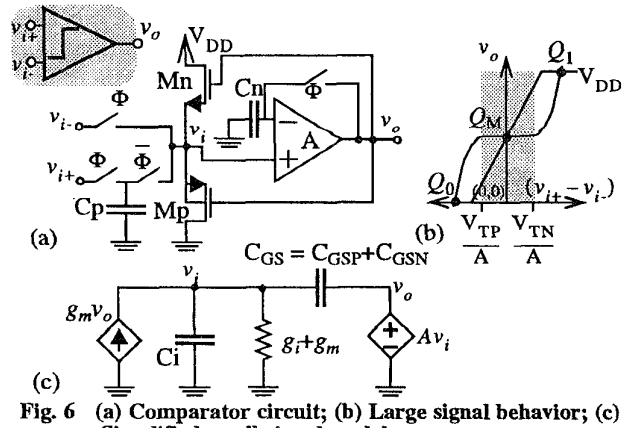

Fig. 6 (a) Comparator circuit; (b) Large signal behavior; (c) Simplified small-signal model.

Mismatching of transistors Mn and Mp with respect to ideal ones changes basically the width of the shaded region of Fig. 6(b). This does not affect the resolution of the circuit as long as Q_M is not a stable point. We reach this conclusion by performing small signal analysis of the circuit in Fig. 6(a). Fig. 6(c) depicts a simplified small signal model for Fig. 6(a). Note that only one transistor is out of the cut-off region, thus g_m equals the small signal transconductance of this transistor. Analysis on this circuit provides the following pole,

$$s = \frac{(A-1)\,g_m - g_i}{C_i - (A-1)\,(C_{GSP} + C_{GSN})} \tag{5}$$

In the central shaded region of Fig. 6(a), both transistors are cut-off, thus we can consider $g_m \approx 0$. The circuit is not stable as long as $C_i < (A-1)\,(C_{GSP} + C_{GSN})$. Thus, under this condition, the circuit will evolve out of the central region. Note that for increasing values of g_m, the circuit becomes stable, which corresponds to both stable points Q_0 and Q_1. Therefore, the circuit provides the right value as long as the charge transfer between the capacitor Cn and the parasitic capacitor Ci (which stores v_{i-}) provides an increment of the input voltage in the right direction. A simple analysis gives the following condition for that,

$$Q_{final} = \begin{cases} Q_1 & if \quad v_{i+} \geq v_{i-} + (\Delta Q / Cp) \\ Q_0 & if \quad v_{i+} \leq v_{i-} + (\Delta Q / Cp) \end{cases} \tag{6}$$

Where ΔQ is the charge pumped out of the channel of the analog current switch.

The previous discussion has not been taken into account the offset of the amplifier. The circuit in Fig. 6(a) has another interesting feature: for Φ=1, an offset cancelation is performed, thus the comparator offset is

$$V_{off} = \frac{V_{off,\,\mathrm{AMPLIFIER}}}{A} + \frac{\Delta Q}{ACn} + \frac{\Delta Q}{Cp} \tag{7}$$

Which is much smaller than that provided in (4) if small transistors are used. The exemplary comparator of this paper is built with the amplifier, capacitors and analog switches depicted in Fig. 7. Despite small devices are used, the resolution of these comparator is more than 8 bits, measured from 30 runs of a montecarlo transitory analysis.

5. THE POLARITY CIRCUIT

The Fig. 8 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 input. The input Φ corresponds to the phase

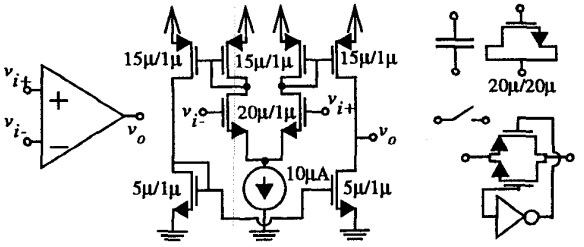

Fig. 7 Implementations of the OTA, the capacitors and the analog switches in Fig. 6.

signal of the comparator in the absolute value circuit of Fig. 3(d), because the comparator is implemented as explained in the previous section (see Fig. 6(a)). On the other hand, the enable input EN corresponds to that in the analog demultiplexors of Fig. 3. Signals at these inputs are depicted in Fig. 8. The computation is finished after 4Δ. For $0 \le t < 2\Delta$, 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 ($EN_B=\Phi_2=0$). For $2\Delta \le t < 3\Delta$, $|F-T|$ is stored in the capacitor Cn of the output comparator. For $t = 3\Delta$ the top input block outputs are disabled ($EN_T=\Phi_2=0$), while the bottom input block outputs are enabled ($EN_B=\Phi_2=1$), 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 Cp and the analog switches in Fig. 6(a)).

6. RESULTS

The Fig. 9 shows some results from HSPICE simulations that illustrate the performance of the presented circuits. The parameter Δ in Fig. 8 equals 100ns in these simulations. Thirty runs of a montecarlo analysis were done with an standard n-well CMOS 1μm technology. Parameter deviations were modeled as reported in [8], with the values for our technology in Table I. Note that the circuit provides the right value for the 30 montecarlo 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.

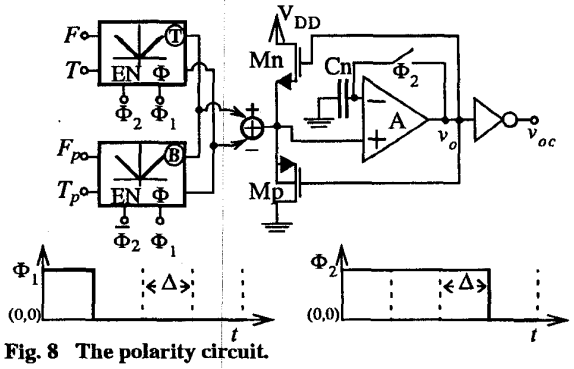

Fig. 8 The polarity circuit.

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] 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

[7] 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.

[8] 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} (Vμm)	A_{VT0p} (Vμm)	$A_{\beta n}$ (μm)	$A_{\beta p}$ (μm)	$A_{\gamma n}$ ($V^{0.5}$μm)	$A_{\gamma p}$ ($V^{0.5}$μm)
12m	14.4m	3.3%	4.5%	6.4m	4.8m

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

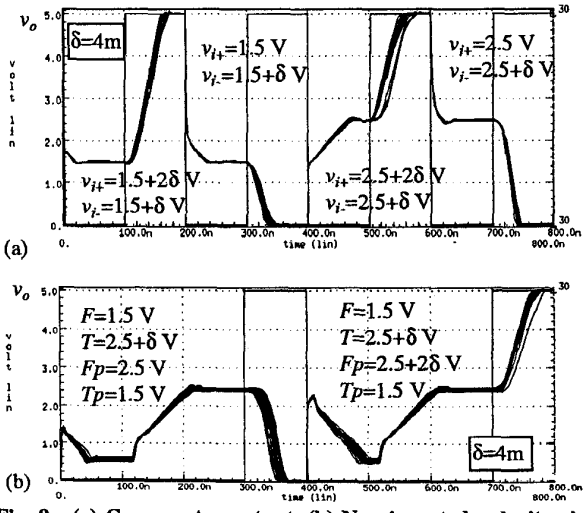

Fig. 9 (a) Comparator output. (b) Non-inverted polarity circuit output.