CHARGE-BASED CMOS FIR ADAPTIVE FILTER

Milutin Stanacevic and Gert Cauwenberghs

Department of Electrical and Computer Engineering
and Center for Language and Speech Processing
Johns Hopkins University, Baltimore MD 21218
E-mail: {milutin,gert}@bach.ece.jhu.edu

ABSTRACT

A low-power, high-density CMOS design for an LMS adaptive FIR filter is presented. A combination of charge-based and subthreshold MOS circuits is used for filtering and adaptation. Accurate analog linear filtering is achieved through pulse-width modulation of input signal. A wide dynamic range is obtained by biasing the MOS switched current sources in the subthreshold regime, with the filter coefficients stored on the gates and adapted through a pilot version of the LMS rule. Simulation results with *SpectreS* verify the linearity and simplicity of the approach. The design allows integration of two 64-tap audio adaptive filters on a single 1 mm^2 0.5 μm CMOS chip.

1. INTRODUCTION

Applications of adaptive filtering in portable audio systems call for special-purpose micropower and high-density hardware implementing the filtering and adaptive functions. While digital signal processing (DSP) solutions provide adequate levels of power dissipation for many applications, a micropower approach is needed for applications such as hearing aids and MEMS sensors. This can be achieved using dedicated analog circuits with MOS transistors in subthreshold [1, 2].

Several analog implementations of adaptive filters exist in the literature [3, 4, 5, 6]. The filtering process itself involves linear multiplication of filter coefficients with a set of time-delayed inputs. Analog multiplication is often implemented using the square-law characteristic of a MOS transistor above threshold [7, 8]. Alternative implementations with the subthreshold MOS operation yield potentially lower power dissipation [9, 10, 11, 12], but are inherently nonlinear in the voltage domain.

We propose an alternative pulse-width modulation scheme with wide-range linear voltage inputs and MOS switched current sources, biased in the subthreshold region for a large (exponential) dynamic range of weight coefficients. A pulsed signal internal representation is also attractive from the prospective of neural models of information and signal processing [13] with area efficient implementation in VLSI [14, 15, 16].

For adaptation, the least-mean-square (LMS) algorithm, is widely used, due to its simplicity [17]. Analog implementation of the LMS rule requires four-quadrant outer product multiplication of the input and error vectors, typically implemented using Gilbert multipliers. We propose LMS adaptation using pulse-arithmetic and charge-based updates, that is less sensitive to analog effects and transistor mismatch.

2. CIRCUIT ARCHITECTURE

The input-output relationship of an adaptive filter is defined by

$$y[n] = \sum_k h_k x[n-k] \tag{1}$$

with filter coefficients adapted through LMS learning rule:

$$h_k[n+1] = h_k[n] - \eta x[n-k]e[n] \tag{2}$$

where η is the adaptation rate for learning. In the following we will describe in detail each of the building blocks of an adaptive filter: delay line, adaptation cell and multiply-and-accumulate circuit with output driver.

2.1. Delay Line

Typical audio applications require a large number of taps. In our system, a 64-tap analog delay line is realized. The cumulative effect of offset and linear gain errors in each stage of the delay line results in a sizable offset and scaling at the output. However, offset and gain errors do not disturb the linear filtering operation, and only contribute to the DC component of the output signal and modified filter coefficients. Indeed, assume an additive offset o_k and non-unity gain a_k for each stage. Then, the resulting output of the filter is

$$y[n] = \sum_k h'_k x[n-k] + \sum_k h'_k o_k \tag{3}$$

where

$$h'_k = \prod_{l=1}^{k} a_l h_k \tag{4}$$

which still implements a linear filter with an additive DC offset. Therefore, stringent design constraints on the offset and gain specifications of the delay element can be avoided, and a standard switched capacitor (SC) design can be used.

The delay element, shown in Figure 1, is implemented by cascading two sample-and-hold circuits [18]. A cascoded inverter is used for the high-gain amplifier. This delay element has small chip area, is parasitic insensitive, and operates fast. The clock rate for audio applications is not high, so there are no problems with the slew-rate and settling-time.
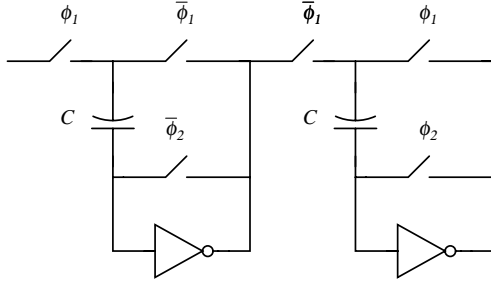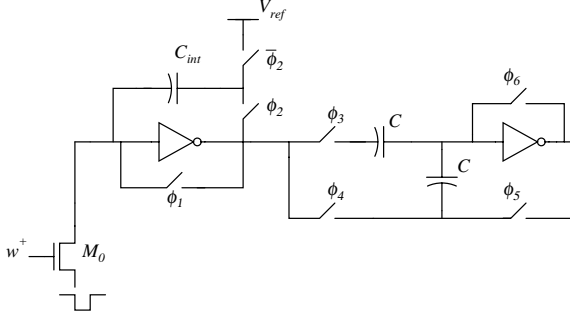
Figure 1: Delay element.

Figure 2: Integration of switched currents and weighted subtraction.

(a)

(b)

Figure 3: Pulse-width modulation: (a) Circuit implementation. (b) Timing diagram.

2.2. Multiply-Accumulate

The multiply-accumulate terms in (1) are implemented by integrating switched currents controlled by a pulse-width modulation of input signal and gate voltages of a pair of CMOS current sources, as shown in Figure 2. The realized multiplication is four-quadrant, with differential weights and bipolar input signal.

The source voltage of the multiplication transistor is pulsed, where the width of the pulse is proportional to absolute value of input signal, x. The polarity of the input signal, relative to reference voltage V_{ref}, controls the position of the pulse with respect to reference time t_0, counted as negative on one side, and positive on the other, as given in the Figure 3(a). The circuit used for pulse-width modulation and for determining the sign of the input signal is shown in Figure 3(b). The reference time point t_0 is used in the accumulate circuit to determine the sign of the contribution to the output. Active low source voltage controls the amount of transistor current.

The weights are stored differentially as voltages w^+ and w^- on two complementary switched current sources, each with the same pulsed input signal. In subthreshold, the current during activation of the sources is exponential in the weights, implementing a coefficient

$$h = h_0(\exp(\kappa\beta w^+) - \exp(\kappa\beta w^-)) \qquad (5)$$

where $\beta = \frac{q}{kT}$ and $h_0 = I_0 \exp(-\beta V_s)$. The advantage of this nonlinear transformation is that a wide dynamic range of coefficients is obtained over a limited linear range of voltages w^+ and w^-.

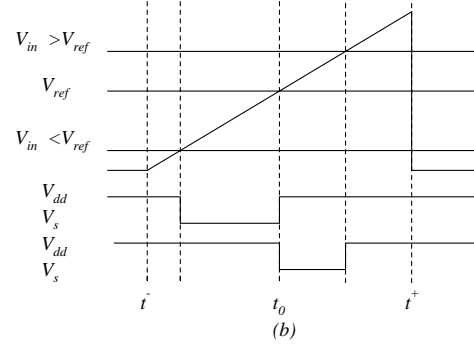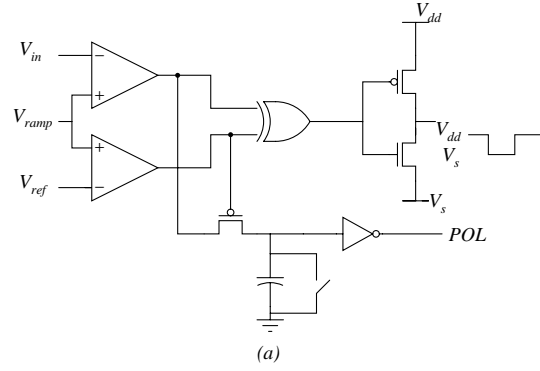Current pulses integrated on the output capacitor before the reference time t_0 are taken with negative sign, accounting for the negative sign of the input signal. The pulses integrated after t_0 are contributed by positive input signals and are taken with positive sign. The difference between positive and negative contributions is obtained by subtracting twice the negative term from the sum of positive and negative terms.

$$
\begin{aligned}
V^+ - V^- &= \int_{t_0}^{t^+} I_{out}dt - \int_{t^-}^{t_0} I_{out}dt \\
&= \int_{t^-}^{t^+} I_{out}dt - 2\int_{t^-}^{t_0} I_{out}dt \qquad (6)
\end{aligned}
$$

The circuit for implementing this weighted subtraction is similar to one used in the algorithmic A/D converter [19] and is given in Figure 2. The fully differential design is adopted throughout, with separate signal paths for w^+ and w^- contributes (each in turn with separate integrations of $V^+ - V^-$). Following the two differential integration stages is a standard SC subtraction stage.

2.3. Adaptation Cell

The adaptation weight cell capable of providing fine weight changes with both positive and negative increments, is shown in Figure 4(a). Implementation of the learning rule (2) would require four-quadrant multiplication. A 'pilot' (i.e. bang-bang) version of LMS learning rule

$$h_k[n+1] = h_k[n] - \eta \operatorname{sgn}(x[n-k]) \operatorname{sgn}(e[n]) \qquad (7)$$

simplifies the architecture, so that multiplication reduces to XNOR operation. The XNOR is implemented using only one transistor
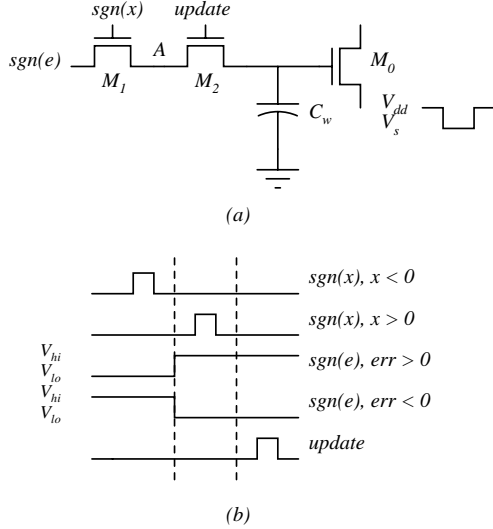
Figure 4: Adaptation cell: (a) Circuit implementation. (b) Timing diagram.

$M1$ with proper signal timing. The signals $\mathrm{sgn}(x)$ and $\mathrm{sgn}(e)$ are coded as a pulse and a two-level signal, respectively, as shown in Figure 4(b). The position of the pulse is determined by the sign of the input signal, while the sign of the error signal determines the order of the levels V_{lo} and V_{hi}. These voltage levels are applied externally, which control the value of adaptation rate η. During the $\mathrm{sgn}(x)$ pulse, the parasitic capacitor C_p at node A is charged to voltage V_{lo} or V_{hi}, determined by $\mathrm{sgn}(e)$.The voltage at this node is

$$V_{Ak}^{+}[n] = \mathrm{sgn}(x[n-k]) \ \mathrm{sgn}(e[n]) \ \delta V_A + V_{A0} \qquad (8)$$

for w^{+} cell and

$$V_{Ak}^{-}[n] = -\,\mathrm{sgn}(x[n-k]) \ \mathrm{sgn}(e[n]) \ \delta V_A + V_{A0} \qquad (9)$$

for w^{-} cell, with

$$V_{lo} = V_{A0} - \delta V_A, \qquad V_{hi} = V_{A0} + \delta V_A \qquad (10)$$

When $update$ goes high, the charge on C_p and C_w is shared. The resulting change on the capacitor is given by

$$w_k^{+}[n+1] = w_k^{+}[n] + \frac{C_p}{C_w + C_p}(V_{Ak}^{+}[n] - w_k^{+}[n]) \qquad (11)$$

and with similar equation for w_k^{-}. For $C_p \ll C_w$, small increments are obtained. To reduce the area of implementation C_w is implemented as a MOS capacitance and C_p is the small parasitic capacitance on the drain/source diffusion between transistors $M1$ and $M2$. The common mode component $\frac{1}{2}(w_k^{+} + w_k^{-})$ is regulated by the weight decay term on the right side of (11), pulling the values towards the center of the range V_{A0}.

There are two physical mechanisms besides the adaptation that affect voltage on the weight capacitor. The first is charge-injection from transistor $M2$ and the second is charge leakage due to n drain p substrate junction of $M2$. Since both mechanisms affect voltage in the same direction, we need to compensate this bias, which is accomplished by applying voltages V_{lo} and V_{hi} that are biased in the opposite direction to V_{A0}. If a longer time of the weight storage is needed, dynamic refresh of the capacitor memory is necessary [20].
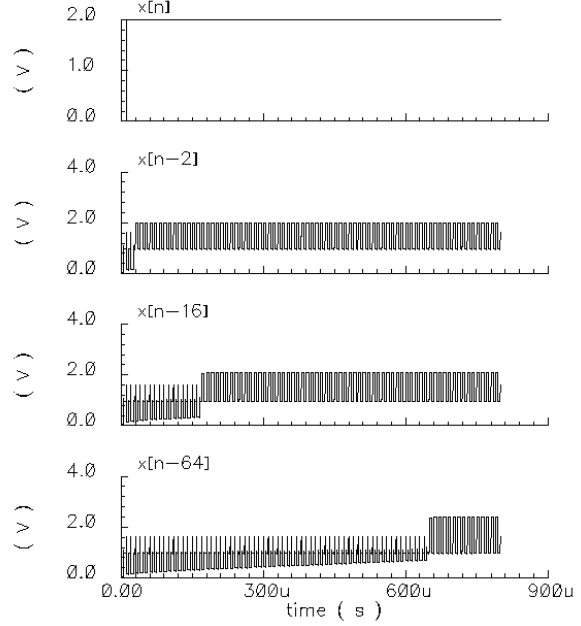

Figure 5: Step response at different tap positions (2, 16 and 64) in the 64-element delay line, for zero initial conditions in all taps.

3. SIMULATION RESULTS

The operation of the circuits was verified through *SpectreS* simulation in Cadence using parameters obtained from a 0.5 μm CMOS process. The step response of the delay line, at different tap positions, is shown in Figure 5. After the transients due to the (zero) initial conditions, the cumulative offset at each tap position settles to a constant over time. Figure 6 shows the trajectory of the differential weights with constant sign of input and error signal. Figure 7 shows linear characteristic of the multiplication. Estimated power dissipation for two 64-tap filters, at a 100 kHz sampling rate, is 200 μW and the energy dissipated per cell per clock cycle is 16 pJ.

4. CONCLUSION

An efficient, low-power and high-density analog realization of FIR adaptive filter is presented, making use of pulse-based charge-mode computation. The circuit operates with subthreshold MOS transistors and achieves a wide linear voltage range. The scheme extends to the design of neural filtering systems, including Independent Component Analysis (ICA) [21]. An efficient charge-mode implementation of the LMS rule is included in the architecture. The delay line can be replaced by more general elements such as all-pass filters for further enhancement.
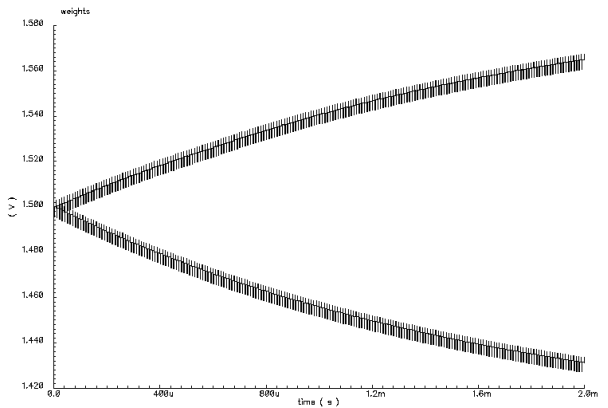
Figure 6: Trajectory of the differential weights over time, under adaptation with updates of constant polarity.
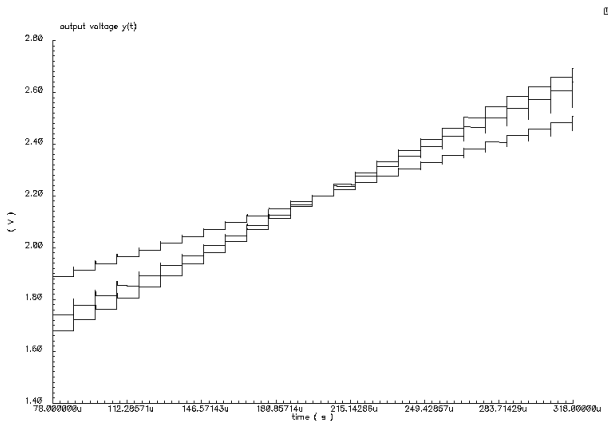

Figure 7: Multiply characteristics for three values of the weight.

5. REFERENCES

[1] C.A. Mead, *Analog VLSI and Neural Systems*, Reading, MA: Addison-Wesley, 1989.

[2] G. Cauwenberghs and M.A.Bayoumi, Eds., *Learning on Silicon*, Norwell, MA: Kluwer Academic, 1999.

[3] J.G. Harris, J.K. Juan and H.C. Principe, "Analog Hardware Implementation of Adaptive Filter Structures," *Proc. of International Conference on Neural Networks*, Houston, 1997.

[4] F.J. Kub and E.W. Justh, "Analog CMOS Implementation of High-Frequency Least-Mean Square Error Learning Circuit," *IEEE Journal of Solid-State Circuits*, vol. **30** (12), pp 1391-1398, Dec 1995.

[5] U. Menzi and G.S. Moschytz, "Adaptive Switched-Capacitor Filters Based on the LMS Algorithm," *IEEE Transactions on Circuits and Systems I*, vol. **40** (12), pp 929-942, Dec 1993.

[6] T. Ritoniemi, T. Karema and H. Tenhunen, "A Sigma-Delta Modulation Based Analog Adaptive Filter," *Proc. IEEE International Symposium on Circuits and Systems*, vol. **6**, pp 2657-2660, 1992.

[7] G. Han and E. Sanchez-Sinencio, "CMOS Transconductance Multipliers: a Tutorial," *IEEE Transactions on Circuits and Systems II*, vol. **45** (12), pp 1550-1563, Dec 1998.

[8] F.J. Kub, K.K. Moon, I.A. Mack and F.M. Long, "Programmable Analog Vector-Matrix Multipliers," *IEEE Journal of Solid-State Circuits*, vol. **25** (1), pp 207-214, Feb 1990.

[9] A.G. Andreou, K.A. Boahen, P.O. Pouliquen, A. Pavasovic, R.E. Jenkins and K. Strohbehn, "Current-Mode Subthreshold MOS Circuits for Analog VLSI Neural Systems," *IEEE Transactions on Neural Networks*, vol. **2** (2), pp 205-213, Mar 1997.

[10] S.I. Liu and C.C. Chang, "CMOS Subthreshold Four-Quadrant Multiplier Based on an Balanced Source-Coupled Pairs," *Int. J. Electronics*, vol. **78**, pp 327-332, Feb 1995.

[11] R.T. Edwards and G. Cauwenberghs, "A Second-Order Log-Domain Bandpass Filter for Audio Frequency Applications," *Proc. IEEE International Symposium on Circuits and Systems*, vol. **3**, pp 651-654, 1998.

[12] C. Toumazou, J. Ngarmnil and T.S. Lande, "Micropower Log-Domain Filter for Electronic Cochlea," *Electronics Letters*, vol. **30** (22), pp 1839-1841, Oct 1994.

[13] W. Maass and C.M. Bishop, Eds., *Pulsed Neural Networks*, Cambridge, MA: MIT Press, 1999.

[14] A.F. Murray, A. Hamilton, H.M. Reekie and L. Tarassenko, "Pulse-Stream Arithmetic in Programmable Neural Networks," *IEEE International Symposium on Circuits and Systems*, vol. **2**, pp 1210-1212, 1989.

[15] M. Nagata, T. Yoneda, D. Nomasaki, M. Sano and A. Iwata, "A Minimum-Distance Search Circuit using Dual-Line PWM Signal Processing and Charge-Packet Counting Techniques," *Dig. IEEE Int. Solid-State Circuits Conf.*, pp 41-43, 1997.

[16] I.S. Han, "Neuro-Fuzzy Hybrid Hardware Implementation with Neural Chip URAN," *IEEE International Conference on Neural Networks*, vol. **6**, pp 3978-3981, 1994.

[17] S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1986.

[18] F.J. Wang and G.C. Temes, "A Fast Offset-Free Sample-and-Hold Circuit," *Proc. of the IEEE Custom Integrated Circuits*, pp 5.6/1-5.6/3, 1988.

[19] M.K. Mayes and S.W. Chin, "A 200 mW, 1 Msample/s, 16-b Pipelined A/D Converter with On-Chip 32-b Microcontroller," *IEEE Journal of Solid-State Circuits*, vol. **31** (12), pp 1862-1872, Dec 1996.

[20] J. Lubkin and G. Cauwenberghs, "A Micropower Learning Vector Quantizer for Parallel Analog-to-Digital Data Compression," *IEEE International Symposium on Circuits and Systems*, vol. **3**, pp 58-61, 1998.

[21] M. Cohen and G. Cauwenberghs, "Blind Separation of Linear Convolutive Mixtures through Parallel Stochastic Optimization," *IEEE International Symposium on Circuits and Systems*, vol. **3**, pp 17-20, 1998.