

Digital LMS Adaptation of Analog Filters Without Gradient Information

Anthony Chan Carusone, *Member, IEEE*, and David A. Johns, *Fellow, IEEE*

Abstract—The least mean square (LMS) algorithm has practical problems in the analog domain mainly due to dc offset effects. If digital LMS adaptation is used, a digitizer (analog-to-digital converter or comparator) is required for each gradient signal as well as the filter output. Furthermore, in some cases the state signals are not available anywhere in the analog signal path necessitating additional analog filters. Here, techniques for digitally estimating the gradient signals required for the LMS adaptation of analog filters are described. The techniques are free from dc offset effects and do not require access to the filter's internal state signals. Digitizers are required only on the input and error signal. The convergence rate and misadjustment are identical to traditional LMS adaptation, but an additional matrix multiplication is required for each iteration. Hence, analog circuit complexity is reduced but digital circuit complexity is increased with no change in overall performance making it an attractive option for mixed-signal integrated systems in digital CMOS. Signed and subsampled variations of the adaptive algorithm can provide a further reduction in analog and digital circuit complexity, but with a slower convergence rate. Theoretical analyses, behavioral simulations, and experimental results from an integrated filter are all presented.

Index Terms—Adaptive filters, continuous-time filters, gradient methods, ladder filters, least mean square methods, mixed analog-digital integrated circuits.

I. INTRODUCTION

ANALOG adaptive filters can offer many advantages over their digital counterparts in integrated communication systems [1]. At the receiver, the resolution and linearity of the analog-to-digital converter (ADC) can generally be reduced if preceded by an analog equalizer or echo canceler [2], [3]. In a full duplex transmitter, the line driver requirements can be relaxed if followed by an analog adaptive hybrid [4]. Unfortunately, the least mean square (LMS) algorithm, which is usually used for integrated adaptive filters, has practical problems in the analog domain due to dc offset effects [5], [6]. Digital implementations of the algorithm are possible, even with an analog signal path. However, they require access to digital gradient information which in turn must be produced by additional high-speed ADCs and may even require additional analog filters [5], [7]. This paper describes techniques for obtaining the digital gradient signals required for LMS adaptation without access to the analog filter's internal state signals. Previous work in this area has resulted in complicated algorithms which are

Manuscript received August 29, 2003; revised February 19, 2003. This paper was recommended by Associate Editor A. Petraglia.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: tcc@eecg.toronto.edu).

Digital Object Identifier 10.1109/TCSII.2003.815021

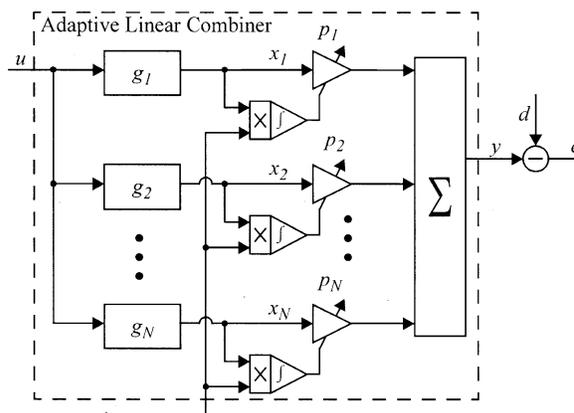


Fig. 1. Block diagram of an adaptive linear combiner with N parameters p_i adapted via the LMS algorithm.

difficult to implement [8]. Therefore, emphasis is placed upon reducing the adaptation hardware requirements.

First, some background on stochastic gradient adaptation in general and the LMS algorithm in particular is provided in Section II. Then, in Sections III and IV, two techniques are proposed to overcome the shortcomings of the LMS algorithm for analog adaptive filters. A theoretical analysis of the proposed techniques' convergence and misadjustment is performed in Section V. Behavioral simulations are used to verify the analytical results in Section VI. In Section VII, signed variations of the adaptation are considered to simplify their implementation. Finally, in Section VIII experimental results are provided for a fifth-order integrated analog filter with three adapted parameters.

II. BACKGROUND

Stochastic gradient adaptation takes the following general form:

$$\mathbf{p}(k+1) = \mathbf{p}(k) - \mu \cdot \hat{\nabla}_{\varepsilon} \mathbf{p}(k). \quad (1)$$

In (1), \mathbf{p} is the vector of filter parameters to be adapted $[p_1 \cdots p_N]^T$, $e = (d - y)$ is the error in the filter output y with respect to the desired output d , μ is a constant that determines the rate of adaptation, and $\hat{\nabla}_{\varepsilon} \mathbf{p}(k)$ is an estimate of the gradient of \mathbf{p} with respect to the mean squared error, $\varepsilon = E[e^2]$. Equation (1) attempts to increment the filter parameter vector by small steps in the direction of decreasing mean squared error. Stochastic gradient adaptation proceeds by iterating (1) until the mean squared error is minimized.

The method used to obtain the gradient estimate $\hat{\nabla}_{\varepsilon} \mathbf{p}(k)$ will depend upon the structure of the adapted filter. Since adapting

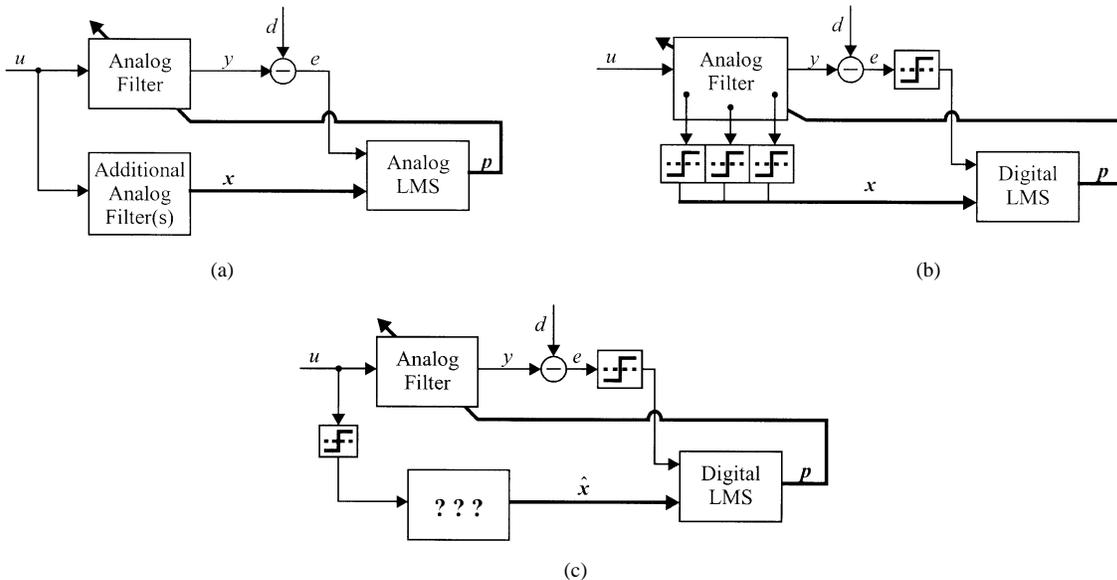


Fig. 2. Implementation of the LMS algorithm for an analog adaptive filter. (a) Analog implementation. (b) Digital implementation. (c) Digital adaptation without access to the filter state signals (proposed).

the poles of a filter can cause instability both in the signal path and in the adaptation process, it is usual to adapt only the zeros. Therefore, this work considers only analog filters with fixed poles and adapted zeros. Any such filter can be modeled by the adaptive linear combiner (ALC) shown in Fig. 1. The N signal path filters, g_i , are fixed and determine the locations of the ALC poles. (The pole locations for analog integrated filters are often chosen either heuristically, perhaps to obtain an equiripple and flat group delay response as in [9], or using numerical optimizations [3], [10].) By adapting the N parameters p_i of an ALC, the location of N filter zeros are optimized.

If the expected value of the gradient estimate equals the actual gradient, $E[\hat{\nabla}_{\mathbf{p}} \varepsilon(k)] = \nabla_{\mathbf{p}} \varepsilon(k)$, stochastic gradient adaptation will converge to a local minimum in the performance surface for small μ [11]. This is the case in the standard LMS algorithm, which uses the following simple gradient estimate:

$$\hat{\nabla}_{\mathbf{p}} \varepsilon(k)_{\text{LMS}} = -2e(k)\mathbf{x}(k). \quad (2)$$

In (2), \mathbf{x} is the vector of state signals $[x_1 \cdots x_N]^T$. The resulting iterative update rule is

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu e(k)\mathbf{x}(k). \quad (3)$$

Notice that in Fig. 1 the required state signals \mathbf{x} are available at the outputs of the fixed filters g_i .

There are two major challenges to performing the LMS adaptation on analog integrated filters. First, the state signals \mathbf{x} are sometimes difficult to obtain. Integrated analog filters with a ladder structure [12], [10] and a cascade of biquads [9], [13], [14] often use programmable feedforward gains to adjust the location of transfer function zeros. Although only the filter zeros are adjusted, LMS adaptation is complicated for these structures since, unlike Fig. 1, the state signals are not available at any internal nodes. Additional analog filters are required just to generate them [5]. A block diagram of this approach is shown in Fig. 2(a). The resulting complexity and power-consumption are prohibitive for most practical applications.

Second, even when the state signals are available, dc offsets on the state and error signals (always present in analog integrated filters) lead to inaccurate convergence of the LMS algorithm [5], [6]. Although much work has been done to mitigate these dc offset effects (e.g., [15]–[19]), the most common approach is to use digital circuitry to implement the LMS multiply and accumulate operations (e.g., [2] demonstrates this for one adapted zero). Digital implementations of (3) also allow the adaptation to be easily initialized and frozen. However, in order to maintain a high-speed analog signal path, the error signal and all of the state signals must be digitized by either ADCs or comparators (sign-sign LMS) as shown in Fig. 2(b). The digitizers can be area and power hungry, as well as loading speed-critical nodes internal to the filter. Furthermore, this approach is still only applicable when the analog filter has the required state signals available at internal nodes.

These two problems are addressed in this paper by performing digital LMS adaptation on an analog integrated filter without access to the filter's internal state signals. The state signals are estimated digitally by observing only the filter input, as shown in Fig. 2(c). This requires less analog hardware than the fully analog approach in Fig. 2(a) and dc offset effects are eliminated. Unlike the digital LMS adaptation in Fig. 2(b), this approach can be used on any analog filter structure with programmable zeros and requires fewer digitizers. Although digital complexity is increased, trading off analog circuit complexity for digital circuit complexity is generally desirable in mixed-signal deep-submicron CMOS.

III. LMS ADAPTATION WITH A COORDINATE TRANSFORM

This section will describe a simple technique for digitally estimating the analog filter states from the sampled filter input. If the filter input is digitized at the Nyquist rate, each of the analog filters g_1 – g_N can be emulated by digital filters \hat{g}_1 – \hat{g}_N . The outputs of these filters provide digital estimates of the state signals

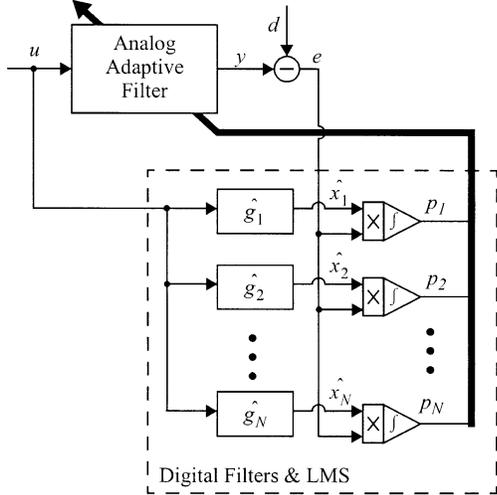


Fig. 3. Estimating the state signals of an analog adaptive linear combiner by emulating the signal path filters digitally.

$\hat{\mathbf{x}}$. These estimates can then be used in place of the actual analog filter states \mathbf{x} for adaptation as follows:

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu e(k)\hat{\mathbf{x}}(k). \quad (4)$$

A block diagram of this approach is shown in Fig. 3.

In general, the digital filters required to emulate g_1 – g_N will have an infinite length impulse response. However, for the remainder of this paper, it is assumed that the tails of the impulse responses have been truncated so that finite impulse response (FIR) filters of length M may be used for \hat{g}_1 – \hat{g}_N . For stable filters, the error incurred by the truncation decreases as M increases and, hence, can be made arbitrarily small by increasing the filter length.

If transversal filters of length M are used for the digital filters $\hat{g}_1 \cdots \hat{g}_N$, (4) can be written in terms of the sampled input vector $\mathbf{u}(k) = [u(k) \ u(k-1) \ \cdots \ u(k-M+1)]^T$ as follows:

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu e(k) \cdot \mathbf{G}^T \cdot \mathbf{u}(k). \quad (5)$$

In (5), \mathbf{G} is an $M \times N$ matrix whose columns are the finite-length impulse responses of the transversal filters $\hat{g}_1 \cdots \hat{g}_N$

$$\mathbf{G} = \begin{bmatrix} \hat{g}_1(1) & \cdots & \hat{g}_N(1) \\ \vdots & & \vdots \\ \hat{g}_1(M) & \cdots & \hat{g}_N(M) \end{bmatrix}. \quad (6)$$

A block diagram implementation of (5) for an analog adaptive filter is shown in Fig. 4. This approach will be referred to as LMS adaptation utilizing a coordinate transformation (LMS-CT). The “coordinate transformation” in (5) is from the input vector \mathbf{u} to the state estimates $\hat{\mathbf{x}}$ by the matrix \mathbf{G}^T . This should not be confused with transform domain or filter bank adaptive filtering [20], [21] where a matrix transformation is applied digitally in the signal path. Here, it is assumed that the main signal path must be analog, so digital signal processing is performed on input samples just to obtain the gradient information required for adaptation. The matrix transformation \mathbf{G} is determined by the structure of the analog filter which is

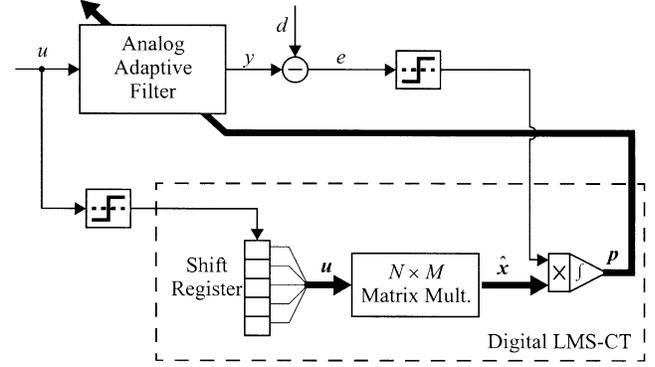


Fig. 4. Block diagram of digital LMS-CT and LMS-ICT adaptation for an analog filter requiring access to only the filter input and error signal.

likely to be dictated by circuit design considerations, whereas in [20] and [21], the matrix transformation is designed to improve certain convergence properties of the adaptation.

If the adaptive filter has a transversal structure, the \mathbf{G} -matrix will be a $N \times N$ identity matrix, and the matrix multiplication in Fig. 4 is trivial. This approach has already been used in combination with sign-sign LMS adaptation for switched-capacitor analog adaptive transversal filters [22], [23]. Introducing the matrix multiplication in (5) generalizes the LMS-CT approach to permit adaptation of any analog filter with programmable zeros.

If slower adaptation can be tolerated, (5) need not be iterated at the Nyquist rate. So, the matrix multiplication, the sampling of the error signal e , and the multiply/accumulate operations can be performed at a decreased rate. However, it is still necessary to sample the filter input u at the full Nyquist rate in order to avoid aliasing in the state estimates $\hat{\mathbf{x}}$.

If the filter input must be sampled and digitized at the Nyquist rate anyway, it might seem natural to implement the signal path using digital filters and eliminate the analog filter completely. However, this would require a high-speed ADC at the input and wide multipliers and adders in the filter; far fewer bits are required if digital signal processing is used only to obtain the state estimates for adaptation. In fact, Section VII verifies that one-bit samples of the input and trivial one-bit multipliers produce state estimates of sufficient accuracy for LMS adaptation.

IV. LMS ADAPTATION WITH AN INVERSE COORDINATE TRANSFORM

Although LMS-CT adaptation obviates the need for sampling the filter state signals directly, it does require the filter input to be sampled at the Nyquist rate. Since analog adaptive filters are often used in high-speed signal paths, this digitizer would have to be clocked at a very fast rate. In this section, a technique is introduced that allows the input digitizer to be subsampled. The technique is described by an iterative update equation with the exact same form as (5) except that a different matrix is used in place of \mathbf{G}^T .

First, consider the LMS algorithm for adapting the tap weights $\mathbf{q} = [q_1 \cdots q_M]^T$ of a length- M transversal filter

$$q_i(k+1) = q_i(k) + 2\mu e(k)u(k-i+1). \quad (7)$$

If slower adaptation can be tolerated, subsampled versions of the input and error signals can be used in (7). Specifically, u and e can be subsampled by $V\times$ and $W\times$ respectively, as long as V and W are relatively prime. For instance, it may be convenient to subsample the input of an M -tap filter by $M\times$ and the output error by $(M+1)\times$. Each parameter is then updated every $M(M+1)$ samples. With $M=5$, the l th parameter update would be given by

$$\begin{aligned} \mathbf{q}(30(l+1)) &= \mathbf{q}(30l) + 2\mu \begin{bmatrix} e(30l+1)u(30l) \\ e(30l+7)u(30l+5) \\ e(30l+13)u(30l+10) \\ e(30l+19)u(30l+15) \\ e(30l+25)u(30l+20) \end{bmatrix} \\ &\equiv \mathbf{q}(30l) - \mu \cdot \hat{\nabla}_\varepsilon \mathbf{q}(30l). \end{aligned} \quad (8)$$

Note that only every fifth sample of u and every sixth sample of e are used in (8), so the digitizers on those signals may be clocked at one-fifth and one-sixth the Nyquist rate respectively. Since the parameters are updated every 30th sample, the algorithm will converge $30\times$ slower.

We now wish to use similarly subsampled versions of e and u to update the parameters of an arbitrary analog filter with programmable zeros. In order to do this, the gradient estimate $\hat{\nabla}_\varepsilon \mathbf{q} \in \mathbb{R}^M$ used in (8) must be projected onto \mathbb{R}^N so that it can be used to update the N -dimensional analog filter parameter vector, \mathbf{p} . A projection method for linearly constrained optimization problems may be used for this purpose [24]. The linear constraint is that the ALC impulse response must stay within the N -dimensional column space of \mathbf{G} since only impulse responses of the form $\mathbf{q} = \mathbf{G}\mathbf{p}$ are possible. This is equivalent to enforcing the following condition:

$$\begin{aligned} \mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{q} &= \mathbf{q} \\ \Leftrightarrow (\mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T - \mathbf{I})\mathbf{q} &= \mathbf{0}. \end{aligned} \quad (9)$$

As shown in [24], this condition can be enforced during adaptation by using $\mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \cdot \hat{\nabla}_\varepsilon \mathbf{q}$ as the gradient estimate instead of $\hat{\nabla}_\varepsilon \mathbf{q}$ resulting in the following update rule:

$$\mathbf{q}(k+1) = \mathbf{q}(k) - \mu \cdot \mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \cdot \hat{\nabla}_\varepsilon \mathbf{q}. \quad (10)$$

Since $\mathbf{q} = \mathbf{G}\mathbf{p}$, (10) may be rewritten in terms of the actual ALC parameters $\mathbf{p}(k)$ by omitting the left-hand matrix multiplication of \mathbf{G}

$$\begin{aligned} \mathbf{p}(k+1) &= \mathbf{p}(k) - \mu \cdot (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \cdot \hat{\nabla}_\varepsilon \mathbf{q}(k) \\ &= \mathbf{p}(k) - \mu \cdot \mathbf{K} \cdot \hat{\nabla}_\varepsilon \mathbf{q}(k) \end{aligned} \quad (11)$$

where the $N \times M$ matrix \mathbf{K} is the pseudo-inverse of \mathbf{G} :

$$\mathbf{K} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T. \quad (12)$$

An intuitive explanation for this choice of \mathbf{K} follows. Recall that an ALC with parameters \mathbf{p} is equivalent to a transversal filter with parameters $\mathbf{q} = \mathbf{G}\mathbf{p}$ (after sampling and truncation). Therefore, the matrix \mathbf{G} maps ALC parameter vectors to transversal filter parameter vectors, $\mathbf{p} \rightarrow \mathbf{q}$. On the other hand, the matrix \mathbf{K} must map the parameter update vector for a

transversal filter, $\Delta \mathbf{q}(k) = -\mu \hat{\nabla}_\varepsilon \mathbf{q}(k)$, to a parameter update vector for the ALC, $\Delta \mathbf{p}(k)$. Hence, \mathbf{K} must perform the inverse mapping of \mathbf{G} . But \mathbf{G} is a rectangular matrix with $M > N$, so an exact inverse for \mathbf{G} will generally not exist. Instead, the pseudo-inverse of \mathbf{G} is used since it provides the inverse mapping with the smallest squared error [25].

Substituting the standard gradient estimate for an LMS adaptive transversal filter, $\hat{\nabla}_\varepsilon \mathbf{q}(k) = -2e(k)\mathbf{u}(k)$, into (11) gives the following iterative update rule:

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu e(k) \cdot \mathbf{K} \cdot \mathbf{u}(k). \quad (13)$$

Adaptation described by (12) and (13) will be referred to as LMS adaptation using an inverse coordinate transform (LMS-ICT). The computations required for each iteration of LMS-ICT adaptation are exactly the same as LMS-CT adaptation; both require the product $e(k) \cdot \mathbf{u}(k)$ to be multiplied by a constant $N \times M$ matrix prior to integration, as shown in Fig. 4. The major advantage of LMS-ICT adaptation is that both u and e may be subsampled, whereas only e can be subsampled using LMS-CT adaptation. Again, the subsampling factors must be chosen relatively prime. Equation (14) shows LMS-ICT adaptation for the case $M=5$ with u and e subsampled by $5\times$ and $6\times$, respectively

$$\mathbf{p}(30(l+1)) = \mathbf{p}(30l) + 2\mu \mathbf{K} \begin{bmatrix} e(30l+1)u(30l) \\ e(30l+7)u(30l+5) \\ e(30l+13)u(30l+10) \\ e(30l+19)u(30l+15) \\ e(30l+25)u(30l+20) \end{bmatrix}. \quad (14)$$

Again, other (relatively prime) subsampling factors V and W may be chosen. Subsampling by $M\times$ and $(M+1)\times$ leads to a straightforward hardware implementation with $M(M+1)\times$ slower convergence. Note that the impulse responses $\hat{g}_1 \cdots \hat{g}_N$ must still be sampled at the full Nyquist rate in order to create the \mathbf{G} -matrix (6) and, hence, the \mathbf{K} -matrix (12). However, since \mathbf{G} and \mathbf{K} remain fixed throughout adaptation, they can be measured just once *a priori*, then stored in a memory.

V. CONVERGENCE AND MISADJUSTMENT ANALYSIS

Like the traditional LMS algorithm, LMS-CT and LMS-ICT adaptation can be considered special cases of (1) with gradient estimates defined as follows:

$$\hat{\nabla}_\varepsilon \mathbf{p}(k)_{\text{LMS-CT}} = -2e(k) \cdot \mathbf{G}^T \cdot \mathbf{u}(k) \quad (15)$$

$$\hat{\nabla}_\varepsilon \mathbf{p}(k)_{\text{LMS-ICT}} = -2e(k) \cdot \mathbf{K} \cdot \mathbf{u}(k). \quad (16)$$

Assuming the errors introduced by sampling and truncating the ALC impulse responses $\hat{g}_1 \cdots \hat{g}_N$ in (6) are negligible, the state estimates used for LMS-CT adaptation are equal to the actual filter state signals, $\mathbf{G}^T \mathbf{u}(k) = \mathbf{x}(k)$. Therefore, $\hat{\nabla}_\varepsilon \mathbf{p}(k)_{\text{LMS-CT}} = \hat{\nabla}_\varepsilon \mathbf{p}(k)_{\text{LMS}}$ and LMS-CT adaptation will have the same stability and misadjustment properties as the full LMS algorithm.

If $E[\hat{\nabla}_\varepsilon \mathbf{p}(k)]$ is not parallel to the actual gradient, $\nabla_\varepsilon \mathbf{p}(k)$, there is said to be some ‘‘gradient misalignment’’ in the adaptation. It will now be shown that this is the case for LMS-ICT

adaptation. The expected value of the LMS-ICT gradient estimate is

$$\begin{aligned} E[\hat{\nabla}_{\varepsilon} \mathbf{p}_{\text{LMS-ICT}}(k)] &= E[\mathbf{K} \cdot 2e(k)\mathbf{u}(k)] \\ &= \mathbf{K} \cdot E[2e(k)\mathbf{u}(k)] \\ &= \mathbf{K} \cdot \nabla_{\varepsilon} \mathbf{q}(k). \end{aligned} \quad (17)$$

In (17), we have used the fact that $E[2e\mathbf{u}] = \nabla_{\varepsilon} \mathbf{q}$, since $2e\mathbf{u}$ is the unbiased gradient estimate used by LMS adaptive transversal filters. Since $\mathbf{K} \cdot \nabla_{\varepsilon} \mathbf{q}$ is not necessarily parallel to $\nabla_{\varepsilon} \mathbf{p}$, there is some gradient misalignment in LMS-ICT adaptation.

Fortunately, it is possible for the adaptation to converge in spite of this gradient misalignment.¹ Next, it will be shown that using LMS-ICT adaptation, the expected value of the ALC parameter vector converges to the optimal value, \mathbf{p}^* , as $k \rightarrow \infty$. Taking the expectation of both sides of the parameter update rule in (13) yields

$$E[\mathbf{p}(k+1)] = E[\mathbf{p}(k)] + 2\mu\mathbf{K} \cdot E[e(k)\mathbf{u}(k)]. \quad (18)$$

The term $E[e(k)\mathbf{u}(k)]$ can be rewritten in terms of the optimal transversal filter parameters \mathbf{q}^* , the input autocorrelation matrix $\mathbf{R} \equiv E[\mathbf{u}\mathbf{u}^T]$, and $E[\mathbf{q}(k)]$ using the Wiener–Hopf equation and assuming $\mathbf{q}(k)$ is independent of $\mathbf{u}(k)$:²

$$\begin{aligned} E[e(k)\mathbf{u}(k)] &= E[d(k)\mathbf{u}(k)] - E[y(k)\mathbf{u}(k)] \\ &= \mathbf{R}\mathbf{q}^* - E[\mathbf{u}(k)\mathbf{u}^T(k)\mathbf{q}(k)] \\ &= \mathbf{R}\mathbf{q}^* - \mathbf{R}E[\mathbf{q}(k)]. \end{aligned} \quad (19)$$

Substituting (19) into (18) yields

$$\begin{aligned} E[\mathbf{p}(k+1)] &= E[\mathbf{p}(k)] + 2\mu\mathbf{K} \cdot (\mathbf{R}\mathbf{q}^* - \mathbf{R}E[\mathbf{q}(k)]) \\ &= E[\mathbf{p}(k)] + 2\mu\mathbf{K} \cdot (\mathbf{R}\mathbf{q}^* - \mathbf{R}GE[\mathbf{p}(k)]) \\ &= (\mathbf{I} - 2\mu\mathbf{K}\mathbf{R}G)E[\mathbf{p}(k)] + 2\mu\mathbf{K}\mathbf{R}\mathbf{q}^*. \end{aligned} \quad (20)$$

Using the Wiener–Hopf equation again allows us to relate \mathbf{p}^* and \mathbf{q}^*

$$\begin{aligned} E[\mathbf{x}\mathbf{x}^T]\mathbf{p}^* &= E[d\mathbf{x}] \\ \mathbf{G}^T E[\mathbf{u}\mathbf{u}^T]\mathbf{G}\mathbf{p}^* &= \mathbf{G}^T E[d\mathbf{u}] \\ \mathbf{G}^T \mathbf{R}\mathbf{G}\mathbf{p}^* &= \mathbf{G}^T \mathbf{R}\mathbf{q}^* \\ (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}\mathbf{G}\mathbf{p}^* &= (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}\mathbf{q}^* \\ \mathbf{K}\mathbf{R}\mathbf{G}\mathbf{p}^* &= \mathbf{K}\mathbf{R}\mathbf{q}^*. \end{aligned} \quad (21)$$

Substituting (21) into (20) gives

$$E[\mathbf{p}(k+1)] = (\mathbf{I} - 2\mu\mathbf{K}\mathbf{R}G)E[\mathbf{p}(k)] + 2\mu\mathbf{K}\mathbf{R}\mathbf{G}\mathbf{p}^*. \quad (22)$$

After performing a coordinate transformation to a principal axis system [11], (22) can be rewritten in terms of a transformed weight-error vector, $\mathbf{w} = \mathbf{Q}^{-1} \cdot (\mathbf{p} - \mathbf{p}^*)$ where \mathbf{Q} is the eigenvector matrix of $\mathbf{K}\mathbf{R}G$

$$E[\mathbf{w}(k)] = (\mathbf{I} - 2\mu\Lambda)^k \mathbf{w}(0). \quad (23)$$

In (23), Λ is the diagonal eigenvalue matrix of $\mathbf{K}\mathbf{R}G$. An equation similar to (23) also describes the convergence of the LMS algorithm, except that Λ is the diagonal eigenvalue matrix of

¹Gradient misalignment has been demonstrated in the sign-sign LMS algorithm [26], yet many practical adaptive integrated filters employ it [23], [27], [28].

²The independence assumption is often invoked in statistical analyses of the LMS algorithm [11], [29], [30] and leads to reliable theoretical predictions of performance, even when there is some dependence between $\mathbf{q}(k)$ and $\mathbf{u}(k)$.

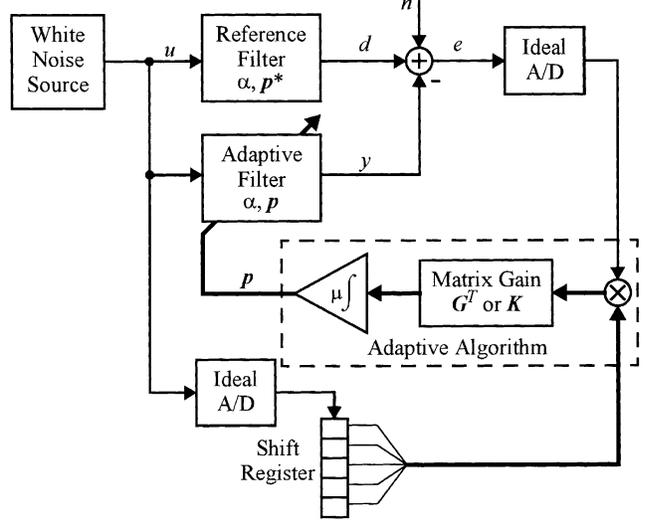


Fig. 5. Simulated adaptive filter model-matching system.

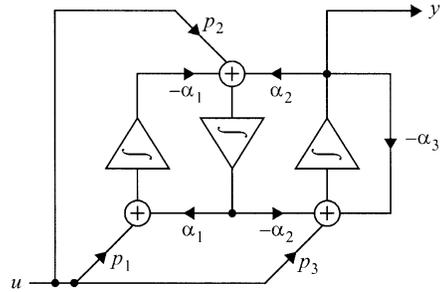


Fig. 6. Third-order orthonormal ladder filter using multiple feed-ins of the input signal.

$E[\mathbf{x}\mathbf{x}^T]$. Therefore, many results for LMS adaptive filters can be applied here by substituting $\mathbf{K}\mathbf{R}G$ for $E[\mathbf{x}\mathbf{x}^T]$. Specifically, if

$$\mu < \frac{1}{\lambda_{\max}} \quad (24)$$

where λ_{\max} is the largest eigenvalue of $\mathbf{K}\mathbf{R}G$, then $(\mathbf{I} - 2\mu\Lambda)^k \rightarrow 0$ as $k \rightarrow \infty$. Hence, $E[\mathbf{w}(k)] \rightarrow 0$ and $E[\mathbf{p}(k)] \rightarrow \mathbf{p}^*$, proving that the mean value of the parameter vector will converge to its optimal value using LMS-ICT adaptation, as long as μ satisfies (24). Furthermore, the time constant of decay of the MSE (in terms of the sampling time) and the steady-state misadjustment are

$$\tau_{\text{MSE}} = \frac{1}{4\mu} \cdot \left(\frac{1}{\lambda} \right)_{\text{avg}} \quad (25)$$

$$\text{Misadjustment} = \mu \sum_i \lambda_i = \mu \cdot \text{tr}(\mathbf{K}\mathbf{R}G). \quad (26)$$

VI. SIMULATION RESULTS

Model-matching experiments were used to verify the LMS-CT and LMS-ICT approaches on continuous-time filters. All of the simulations described in this section use the block diagram shown in Fig. 5. An independent additive noise source n is introduced to control the steady-state error after convergence. The time scale is normalized to a sampling

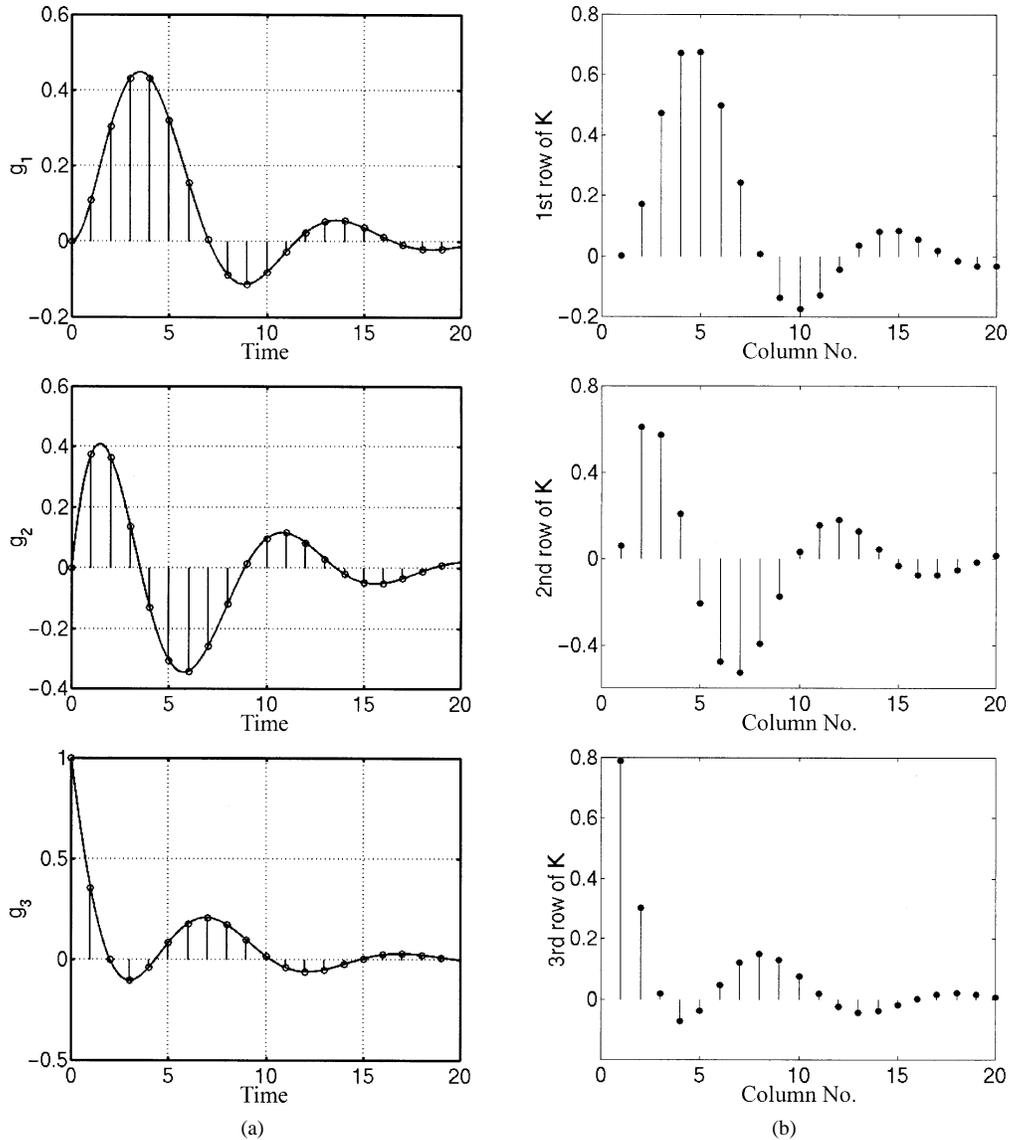


Fig. 7. (a) Impulse responses for the third-order orthonormal ladder filter, sampled to obtain the rows of matrix G . (b) The rows of the pseudo-inverse K .

rate of $f_S = 1$. The reference filter is a third-order elliptic low-pass transfer function with 0-dB dc gain, 0.5 dB of ripple in the passband extending to $0.1f_S$, and 40 dB of stopband attenuation. The filter input is white-noise bandlimited by an eighth-order elliptic filter with 0.1 dB of passband ripple to $0.4f_S$ and 60 dB of stopband attenuation beyond $0.5f_S$.

A. Orthonormal Ladder Filter

Interestingly, when the impulse responses $\hat{g}_i(k)$ are orthonormal, the LMS-CT and LMS-ICT adaptations become identical. This can be seen by arranging the impulse responses into column vectors, $\mathbf{g}_i = [\hat{g}_i(1) \hat{g}_i(2) \cdots \hat{g}_i(N)]^T$. Since the vectors are orthonormal, $\mathbf{g}_i^T \cdot \mathbf{g}_k = 0$ for $i \neq k$ and $\mathbf{g}_i^T \cdot \mathbf{g}_i = 1$. By substituting $\mathbf{G} = [\mathbf{g}_1 \mathbf{g}_2 \cdots \mathbf{g}_N]$ into (12), it is easily verified that $\mathbf{K} = \mathbf{G}^T$, as follows:

$$\begin{aligned} \mathbf{K} &= ([\mathbf{g}_1 \mathbf{g}_2 \cdots \mathbf{g}_N]^T \cdot [\mathbf{g}_1 \mathbf{g}_2 \cdots \mathbf{g}_N])^{-1} \cdot \mathbf{G}^T \\ &= (\mathbf{I}_N)^{-1} \cdot \mathbf{G}^T \\ &= \mathbf{G}^T. \end{aligned} \quad (27)$$

Orthonormal ladder filters have this property [31]. A third-order continuous-time orthonormal ladder structure is shown in Fig. 6. By making the feed-in parameters \mathbf{p} adaptive, the structure becomes an adaptive linear combiner. As mentioned earlier, filters with adaptive feed-ins are of particular interest because the state signals required for traditional LMS adaptation are not available anywhere in the system. In order to perform LMS adaptation, it would be necessary to operate a second filter in parallel with the first just to obtain the gradient signals [7]. In addition to the extra complexity and power consumption which this implies, mismatches between the two filters result in dc offsets that limit the accuracy of the adaptation. Fortunately, the LMS-CT and LMS-ICT adaptations can be used without access to the filter's internal states.

In order to achieve the desired fixed pole locations, the feedback parameters for both the adaptive filter and reference filter were fixed at $\alpha = [0.4905 \ 0.6025 \ 0.7789]$. The feed-in parameters for the reference filter were fixed at $\mathbf{p}^* = [0.5940 \ 0 \ 0.0493]^T$.

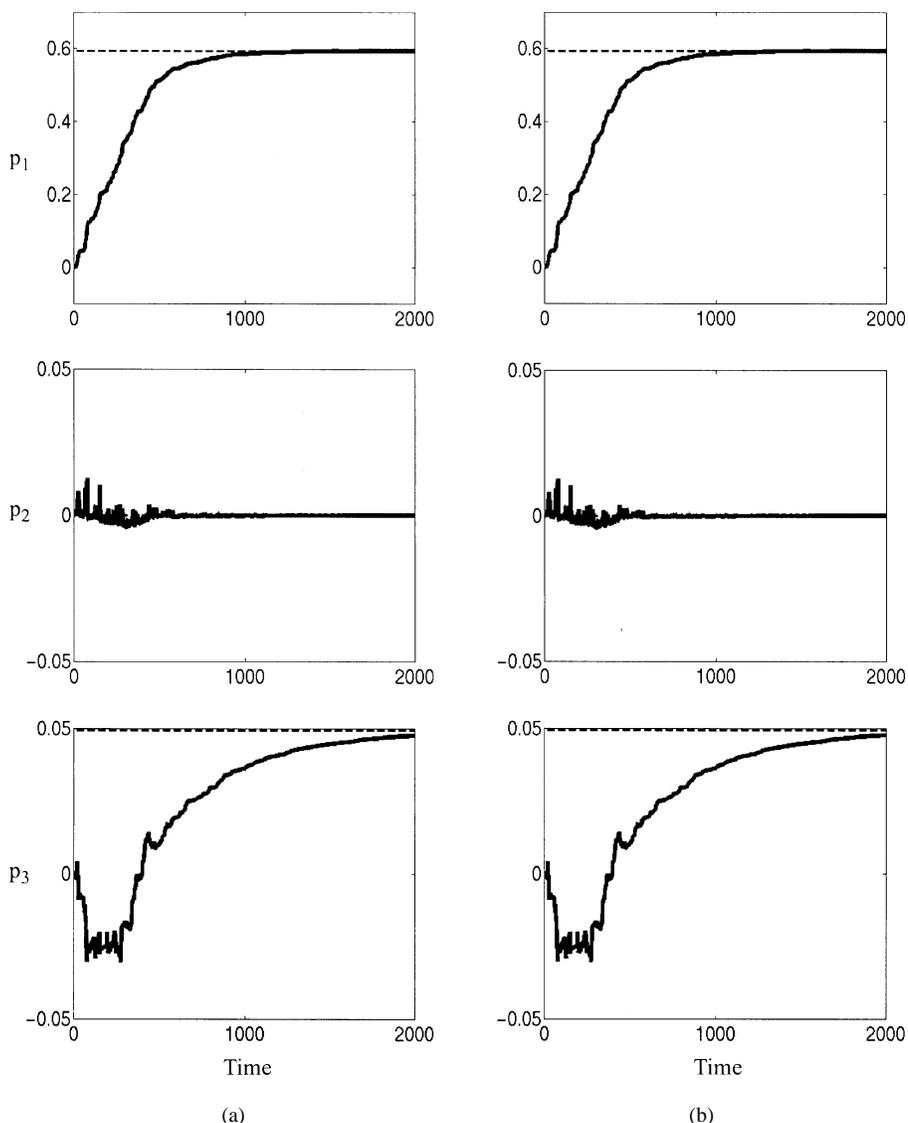


Fig. 8. Sample learning curves for a third-order orthonormal ladder model-matching system using (a) LMS-CT and (b) LMS-ICT adaptation. In both cases, no steady-state error is introduced and μ is chosen for a misadjustment of 1%.

The impulse responses $g_1(t)$, $g_2(t)$, and $g_3(t)$ of the ALC were measured by setting $\mathbf{p} = [1 \ 0 \ 0]^T$, $[0 \ 1 \ 0]^T$ and $[0 \ 0 \ 1]^T$. The sampled impulse responses are plotted in Fig. 7(a), which shows that $M = 20$ samples are sufficient to capture at least 99.8% of the impulse response power. The matrix \mathbf{G} was then constructed using (6), as follows:

$$\mathbf{G}^T \begin{bmatrix} 0 & 0.110 & 0.304 & 0.432 & \cdots \\ 0 & 0.374 & 0.365 & 0.137 & \cdots \\ 1.000 & 0.355 & -0.002 & -0.104 & \cdots \end{bmatrix} \quad (28)$$

and the pseudo-inverse \mathbf{K} is calculated from (12)

$$\mathbf{K} = \begin{bmatrix} 0.001 & 0.171 & 0.473 & 0.673 & \cdots \\ 0.059 & 0.609 & 0.572 & 0.208 & \cdots \\ 0.790 & 0.303 & 0.020 & -0.074 & \cdots \end{bmatrix}. \quad (29)$$

The rows of \mathbf{K} are plotted in Fig. 7(b). Except for a scaling factor for normalization, the waveforms are similar to the columns of \mathbf{G} plotted in Fig. 7(a)³ as expected due to the orthonormal ladder structure.

First, simulations were performed with the noise source turned off ($n = 0$). As can be seen from Fig. 8, both LMS-CT and LMS-ICT adaptation converged to their optimal parameter values with zero steady-state error. The errors incurred by aliasing and truncating the impulse responses had no effect on the result.

Next, some finite steady-state error was introduced via n to examine the algorithms' misadjustment. A noise power of $\text{var}(n) = 0.01$ was used, which is about 3.5 dB less than the output power of the reference filter, $\text{var}(d)$. The input autocor-

³They would be identical if the columns of \mathbf{G} were perfectly orthogonal. However, the frequency response of g_3 extends beyond the Nyquist rate (only 12 dB of attenuation at $f_S/2$) and the resulting aliasing in $\hat{g}_3(k)$ causes the columns of \mathbf{G} to be not perfectly orthogonal.

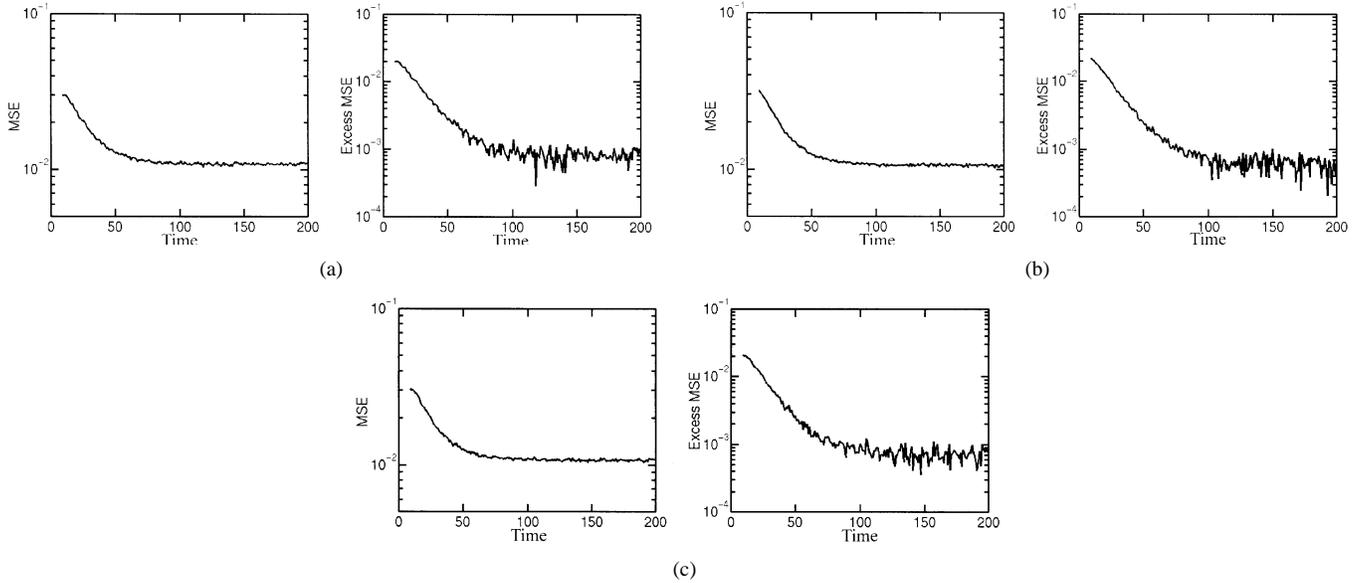


Fig. 9. Simulation of a third-order orthonormal ladder model-matching system using (a) LMS-CT, (b) LMS-ICT and (c) full LMS adaptation. In all cases, μ is selected for a misadjustment of 10% (corresponding to an excess MSE of 10^{-3}) and the results are averaged over an ensemble of 10 000 simulation runs.

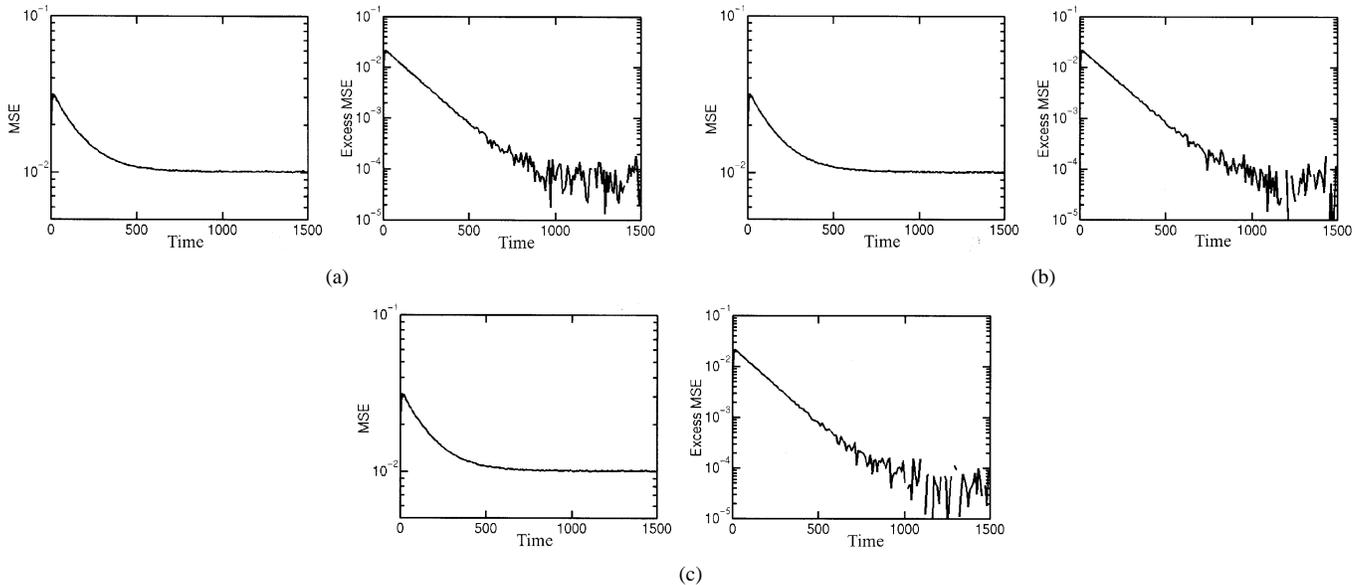


Fig. 10. Simulation of a third-order orthonormal ladder model-matching system using (a) LMS-CT, (b) LMS-ICT, and (c) full LMS adaptation. In all cases, μ is selected for a misadjustment of 1% (corresponding to an excess MSE of 10^{-4}) and each point is averaged over ten consecutive samples of the error signal and over an ensemble of 10 000 simulation runs.

relation matrix is a 20×20 matrix, $\mathbf{R} = E[\mathbf{u}\mathbf{u}^T]$, which can be calculated from a knowledge of the input statistics

$$\mathbf{R} = \begin{bmatrix} 0.0813 & 0.0166 & -0.0142 & \cdots & -0.0012 \\ 0.0166 & 0.0813 & 0.0166 & \cdots & 0.0008 \\ -0.0142 & 0.0166 & 0.0813 & \cdots & -0.0001 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -0.0012 & 0.0008 & 0.0001 & \cdots & 0.0813 \end{bmatrix}. \quad (30)$$

In general, the autocorrelation matrix \mathbf{R} will not be known *a priori*. However, its knowledge is assumed here to demon-

strate that the LMS, LMS-CT, and LMS-ICT adaptations all have the same performance, although different values of μ may be required for each.

The convergence properties of the LMS-ICT adaptation are determined by the eigenvalues of \mathbf{KRG}

$$\lambda_{\text{LMS-ICT}} = \text{eig}(\mathbf{KRG}) = 0.0993, 0.0988, 0.0967. \quad (31)$$

The value of μ required for a misadjustment of 10% is

$$\mu_{\text{LMS-ICT}} = \frac{0.1}{\sum \lambda_{\text{LMS-ICT}}} = 0.3453. \quad (32)$$

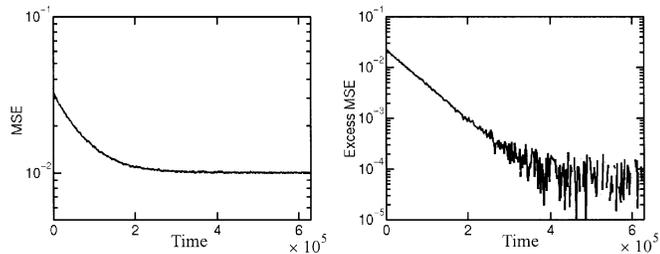


Fig. 11. Simulation results for the third-order orthonormal ladder model-matching experiment using subsampled LMS-ICT adaptation with a misadjustment of 1% (corresponding to an excess MSE of 10^{-4}). Each data point is averaged over 2000 consecutive data samples and 25 separate simulation runs.

For traditional LMS and LMS-CT adaptation, the eigenvalues are all identically $\lambda = 0.0607$ owing to the orthonormal filter structure. The corresponding value for μ is

$$\mu_{\text{LMS and LMS-CT}} = \frac{0.1}{\sum \lambda} = 0.5490. \quad (33)$$

These values together with (26) predict that the MSE should decay with a time constant of approximately eight iterations for all three algorithms. The simulation results plotted in Fig. 9 indicate that the MSE does, indeed, decay to the noise floor provided by $\text{var}(n) = 0.01$ at the same rate in all three cases.

The ‘‘Excess MSE’’ is defined as the MSE observed in steady-state minus the minimum MSE, in this case $E[e^2] - \text{var}(n)$. It is related to the misadjustment by: Excess MSE = (Misadjustment) \times $\text{var}(n)$. The Excess MSE is also plotted in Fig. 9 showing that the misadjustment is 10% as expected in all three cases. Fig. 10 shows similar simulation results for a misadjustment of 1%.

Next, the same system was simulated using LMS-ICT adaptation with the input and error signals subsampled. Since $M = 20$, the input was subsampled by $20\times$ and the error signal was subsampled by $21\times$. The results for a misadjustment of 1% are plotted in Fig. 11. Comparing them with the simulation results in Fig. 10 shows the same misadjustment but with $M(M+1) = 420\times$ slower convergence, as expected.

B. Feedforward Companion Form Filter

In this section, model-matching simulations are performed using a different filter structure. The filter structure, shown in Fig. 12, is a third-order continuous-time companion form filter with variable feed-in coefficients, p_i . Again, since the feed-in parameters are adapted, the state signals required for traditional LMS adaptation are not available. Unlike the orthonormal ladder filter, the impulse responses are not orthogonal. As a result, the matrices \mathbf{G}^T and \mathbf{K} are quite different and there is gradient misalignment using LMS-ICT adaptation.

Fig. 13 shows behavioral simulation results with length $M = 20$ impulse responses and zero excess error added (i.e., $n = 0$) for both LMS-CT and LMS-ICT adaptation. Although both simulations converge with zero steady-state error, the parameter vector evolves along very different trajectories. The trajectories are projected onto the $p_2 = 0$ plane and plotted along with

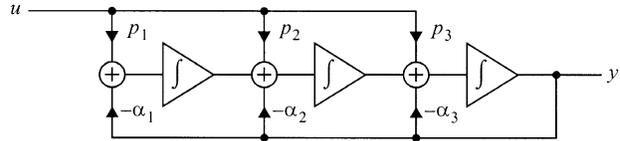


Fig. 12. Third-order feedforward companion form filter.

error surface contours in Fig. 14. In this plot, the gradient misalignment present in LMS-ICT adaptation is evident since the learning trajectory is not orthogonal to the MSE contours.

VII. SIGNED ALGORITHMS

It is possible to take the sign of the error signal or the gradient signal or both in (3) in order to simplify the implementation of the LMS algorithm [32]. Taking the sign of both results in the sign-sign LMS (SS-LMS) algorithm

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu \cdot \text{sgn}(e(k)) \cdot \text{sgn}(\mathbf{x}(k)). \quad (34)$$

The product $-\text{sgn}(e(k)) \cdot \text{sgn}(x_i(k))$ provides, on average, the correct sign of each gradient component. The SS-LMS algorithm proceeds by changing the parameters in fixed steps of size 2μ each iteration. The digital multiplication of the error and state signals is performed by a single exclusive-OR gate resulting in considerable hardware savings. Although it is true that the SS-LMS algorithm has demonstrated instability in certain circumstances [26], its simplified hardware has proved useful in numerous applications (e.g., [23], [27], [28]). Stability of the SS-LMS algorithm is usually verified for a particular application via extensive simulations.

The same approach can be used to simplify the hardware required for the LMS-CT and LMS-ICT adaptations. Taking the sign of the error and input data signals and of each entry in the matrices \mathbf{G}^T and \mathbf{K} results in the following update equations:

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu \text{sgn}(\mathbf{G}^T) \cdot \text{sgn}(e(k)) \cdot \text{sgn}(\mathbf{u}(k)) \quad (35)$$

$$\mathbf{p}(k+1) = \mathbf{p}(k) + 2\mu \text{sgn}(\mathbf{K}) \cdot \text{sgn}(e(k)) \cdot \text{sgn}(\mathbf{u}(k)). \quad (36)$$

Equation (35) will be used as the update rule for the sign-sign LMS-CT adaptation (SS-LMS-CT) and (36) for the sign-sign LMS-ICT adaptation (SS-LMS-ICT). This allows the two digitizers at the filter input and error signal to be implemented with simple comparators. The multiplication of the three signed quantities in both (35) and (36) can be performed by three-input XOR gates. The result is a significant decrease in circuit complexity and power consumption.

Behavioral simulations were performed using the same model-matching experiment as in Section VI-A to verify this approach. Simulation results are plotted in Fig. 15. There is no noticeable difference between the performance of the SS-LMS-CT and the SS-LMS-ICT adaptations. For the same misadjustment, the signed implementations converge slower than the full LMS-CT and LMS-ICT adaptations, but this is not surprising since it is well known that the SS-LMS algorithm is slower than the full LMS algorithm. Of course, by taking a

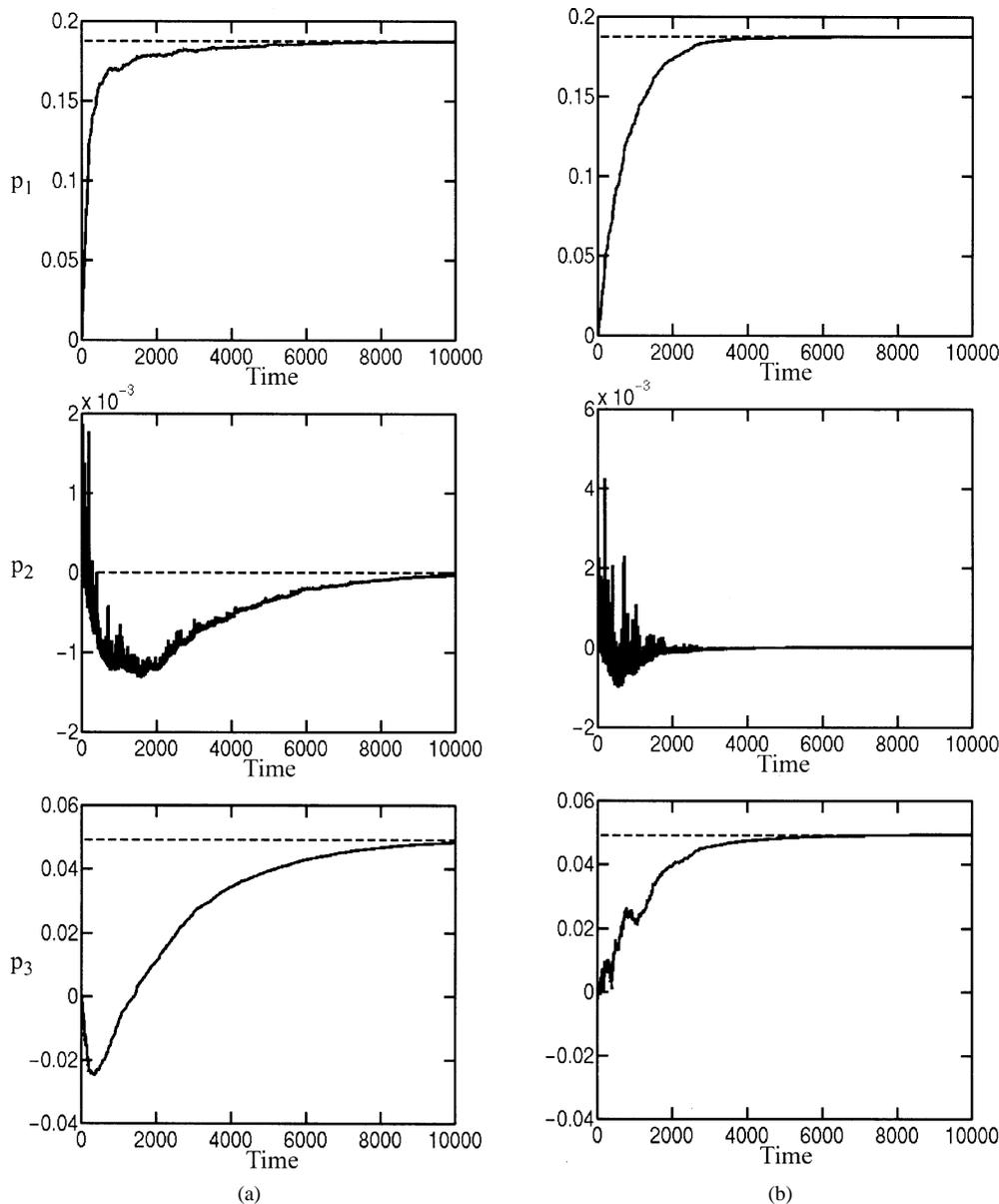


Fig. 13. Model-matching learning curves for a feedforward companion form filter using (a) LMS-CT and (b) LMS-ICT adaptation.

larger value of μ , the slower convergence can be traded off for increased misadjustment.

VIII. EXPERIMENTAL RESULTS

To verify the practicality of the LMS-CT and LMS-ICT adaptations in a real integrated system, model-matching experiments were performed using a fifth-order orthonormal ladder CMOS integrated analog filter [33]. The filter structure is shown in Fig. 16. Each of the three feed-in taps is digitally programmable with five bits of resolution. The filter is low pass with low linearity (25–30 dB total harmonic distortion at 200 mVpp depending upon the feed-in gains and input frequency) and a cutoff frequency programmable up to around 70 MHz. A die photo is shown in Fig. 17.

First, the required impulse responses were obtained by differentiating the step responses measured for each filter g_1 – g_3 on an oscilloscope. The results are plotted in Fig. 18. The waveforms

are messy due to noise and nonlinearity introduced by the filter and measurement equipment.

The experimental setup is diagrammed in Fig. 19. The same filter is used for the adapted and reference signal paths allowing the optimal parameter values, \mathbf{p}^* , to be known precisely. First, the oscilloscope digitizes the filter output with the filter's feed-in values programmed to their optimal values, \mathbf{p}^* . The digitized waveform is then stored by the PC for use as the desired signal, d . Then, the same input sequence is repeated with the feed-in parameters programmed to the current adapted values, $\mathbf{p}(k)$. This time, the digitized waveform is used as the output signal, y . The oscilloscope also digitizes the filter input, u , on a second channel. The error signal $e = d - y$ and the input u are then used to perform one iteration of the adaptive algorithm's parameter update equation in software.

Under these conditions, it would be impossible to use a traditional LMS algorithm since the filter's state signals are completely unavailable. However, using the LMS-CT

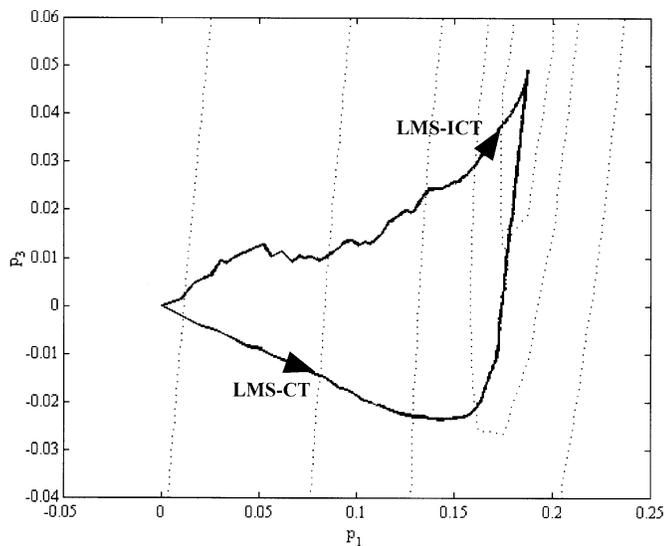


Fig. 14. Learning trajectories of model-matching experiments with MSE contours.

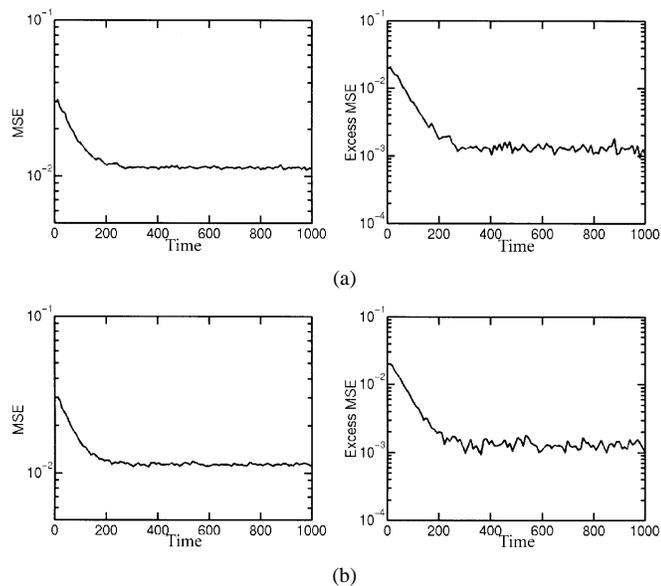


Fig. 15. Simulation results for the third-order orthonormal ladder model-matching experiment using (a) SS-LMS-CT and (b) SS-LMS-ICT adaptation. In both cases, μ is selected for a misadjustment of approximately 10% (corresponding to an excess MSE of 10^{-3}) and each point is averaged over ten consecutive samples of the error signal and over an ensemble of 1000 simulation runs.

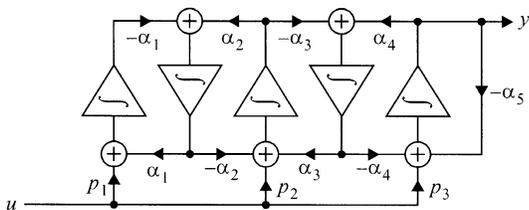


Fig. 16. Fifth-order orthonormal ladder filter with three programmable feed-ins.

and LMS-ICT adaptations, the model-matching experiment succeeds. The 5-bit parameter values and MSE are plotted over time in Figs. 20 and 21. Approximately 2000 iterations

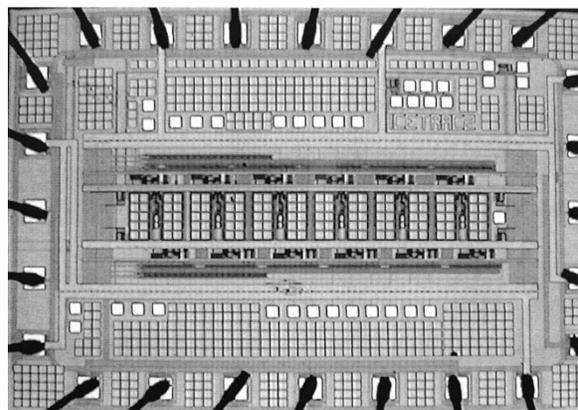


Fig. 17. Die photo of the fifth-order orthonormal ladder analog filter test chip.

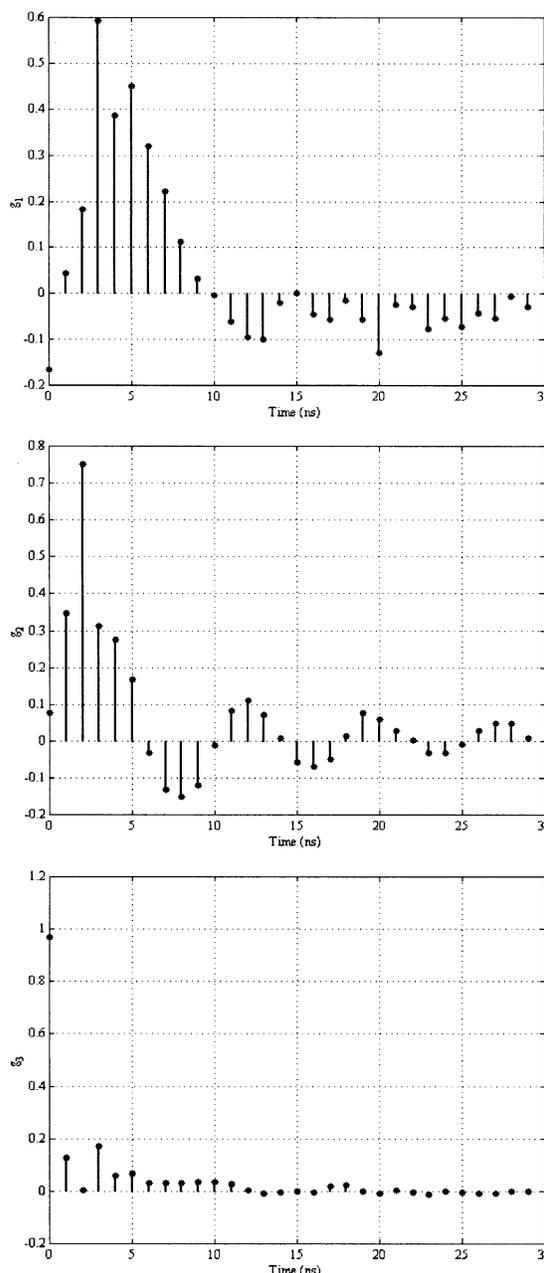


Fig. 18. Sampled and truncated impulse responses of the fifth-order integrated analog filter.

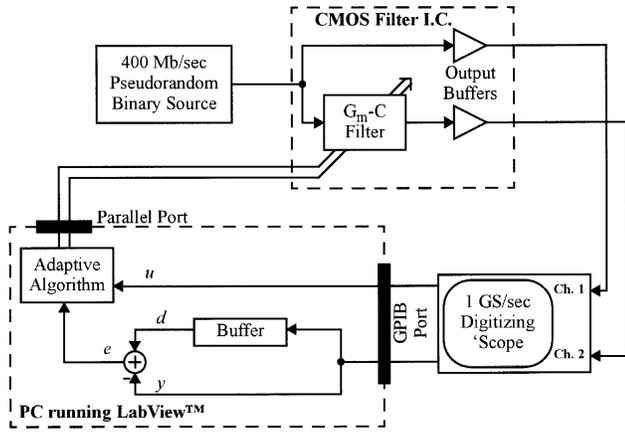


Fig. 19. Experimental setup for performing the LMS-CT and LMS-ICT adaptations on the integrated analog filter.

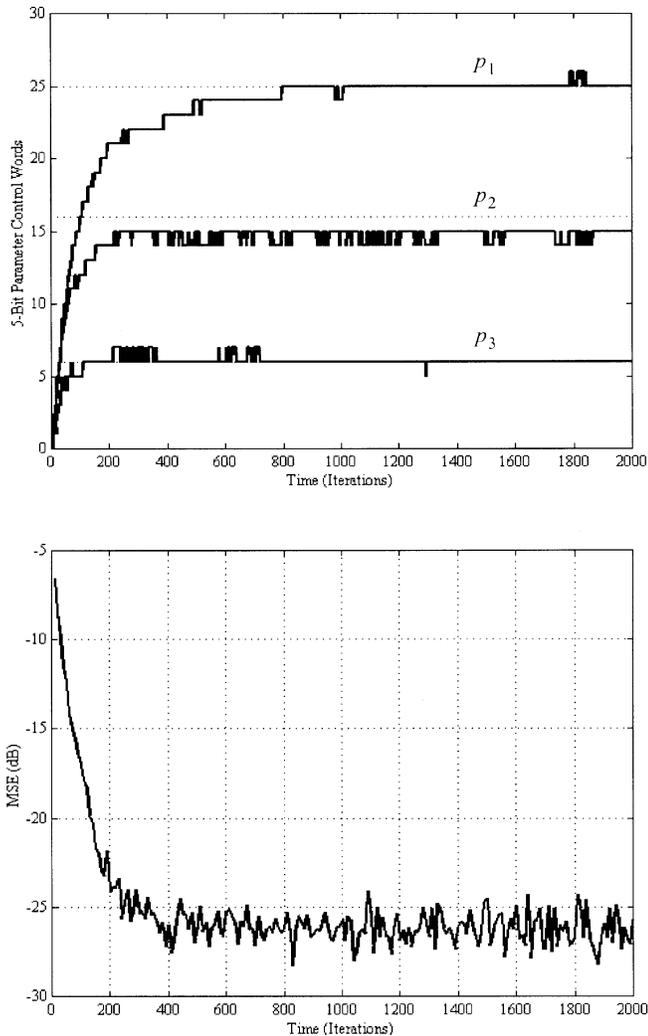


Fig. 20. Model-matching learning curves and MSE relative to desired output using the LMS-CT algorithm.

are required to obtain convergence. A steady-state error of 1 LSB persists on p_2 and the resulting steady-state MSE is 26 dB below the filter output power for both algorithms. These steady-state errors are comparable in magnitude to the filter's nonlinearity.

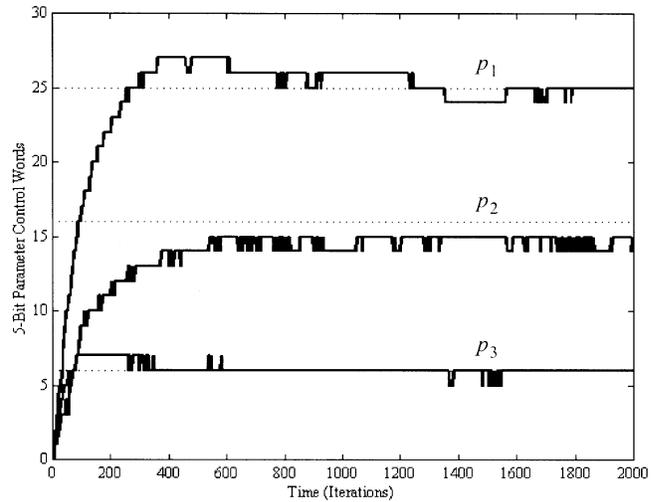


Fig. 21. Model-matching learning curves and MSE relative to desired output using the LMS-ICT algorithm.

IX. CONCLUSION

This paper has described and demonstrated techniques for obtaining digital estimates of the gradient signals required for LMS adaptation without access to a filter's internal state signals. The techniques are particularly useful for digitally adapting high-speed analog filters with several adapted parameters. Using traditional LMS adaptation, a digitizer (ADC or comparator) is required for each gradient signal as well as the filter output. Furthermore, in some popular filter structures, such as those with programmable feed-ins, the state signals are not available anywhere in the analog signal path so additional analog filters must be built to accommodate the LMS algorithm.

Using the LMS-CT or LMS-ICT adaptations, digitizers are required on the input and error signals only and digital signal processing is used to obtain estimates of the gradient signals. Compared to the traditional LMS algorithm, the convergence rate and misadjustment are identical. An additional matrix multiplication is required for each iteration of the algorithm. For the LMS-CT and LMS-ICT adaptations, knowledge of the filter's pole locations is required to create the matrices G^T and K respectively. The matrix entries can be measured once and then

stored in a memory. The adaptation was robust with respect to errors in the matrix entries, both in the simulations of the signed algorithms (Section VII), and due to noise in the experimental setup (Section VIII).

Overall, analog circuit complexity is reduced but digital circuit complexity is increased with little or no change in overall performance making it an attractive option for mixed-signal integrated systems in digital CMOS processes. The signed and subsampled variations of these algorithms allow a system designer to reduce the analog and digital circuit complexity even further, but with a slower convergence rate. This is likely to be a desirable tradeoff in applications such as wired communication where the adaptation rate is not a limiting factor. Combining these techniques, adaptation can be performed using just two comparators and relatively simple digital logic, all of which can be subsampled below the Nyquist rate.

REFERENCES

- [1] A. Carusone and D. A. Johns, "Analogue adaptive filters: Past and present," *Inst. Electr. Eng. Proc.*, vol. 147, pp. 82–90, Feb. 2000.
- [2] J. E. C. Brown, P. J. Hurst, B. C. Rothenberg, and S. H. Lewis, "A CMOS adaptive continuous-time forward equalizer, LPF, and RAM-DFE for magnetic recording," *IEEE J. Solid-State Circuits*, vol. 34, pp. 162–169, Feb. 1999.
- [3] P. K. D. Pai, A. D. Brewster, and A. Abidi, "A 160-MHz analog front-end IC for EPR-IV PRML magnetic storage read channels," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1803–1816, Nov. 1996.
- [4] F. Pecourt, J. Hauptmann, and A. Tenen, "An integrated adaptive analog balancing hybrid for use in (A)DSL modems," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 1999, pp. 252–253.
- [5] D. A. Johns, W. M. Snelgrove, and A. S. Sedra, "Continuous-time LMS adaptive recursive filters," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 769–777, July 1991.
- [6] A. Shoval, D. A. Johns, and W. M. Snelgrove, "Comparison of dc offset effects in four LMS adaptive algorithms," *IEEE Trans. Circuits Syst. II*, vol. 42, pp. 176–185, Mar. 1995.
- [7] K. A. Kozma, D. A. Johns, and A. S. Sedra, "Automatic tuning of continuous-time integrated filters using an adaptive filter technique," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1241–1248, Nov. 1991.
- [8] A. Carusone and D. A. Johns, "Obtaining digital gradient signals for analog adaptive filters," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 3, May 1999, pp. 54–57.
- [9] B. E. Bloodworth, P. P. Siniscalchi, G. A. De Veirman, A. Jezdic, R. Pierson, and R. Sundaraman, "A 450 Mb/s analog front end for PRML read channels," *IEEE J. Solid-State Circuits*, vol. 34, pp. 1661–1675, Nov. 1999.
- [10] S.-S. Lee and C. A. Laber, "A BiCMOS continuous-time filter for video signal processing applications," *IEEE J. Solid-State Circuits*, vol. 33, pp. 1373–1382, Sept. 1998.
- [11] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [12] S. Doshio, T. Morie, and H. Fujiyama, "A 200 MHz seventh-order equiripple continuous-time filter by design of nonlinearity suppression in 0.25 μm CMOS process," *IEEE J. Solid-State Circuits*, vol. 37, pp. 559–565, May 2002.
- [13] V. Gopinathan, M. Tarsia, and D. Choi, "Design considerations and implementation of a programmable high-frequency continuous-time filter and variable-gain amplifier in submicrometer CMOS," *IEEE J. Solid-State Circuits*, vol. 34, pp. 1698–1707, Dec. 1999.
- [14] G. Bollati, S. Marchese, M. Demicheli, and R. Castello, "An eighth-order CMOS low-pass filter with 30–120 MHz tuning range and programmable boost," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1056–1066, July 2001.
- [15] C.-P. J. Tzeng, "An adaptive offset cancellation technique for adaptive filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 799–803, May 1990.
- [16] U. Menzi and G. S. Moschytz, "Adaptive switched-capacitor filters based on the LMS algorithm," *IEEE Trans. Circuits Syst.*, vol. 40, pp. 929–942, Dec. 1993.
- [17] H. Qiuting, "Offset compensation scheme for analogue LMS adaptive FIR filters," *Electron. Lett.*, vol. 38, pp. 1203–1205, June 1992.
- [18] F. J. Kub and E. W. Justh, "Analog CMOS implementation of high frequency least-mean square error learning circuit," in *Proc. IEEE Int. Symp. Circuits and Systems*, Feb. 1995, pp. 74–75.
- [19] A. Shoval, D. A. Johns, and W. M. Snelgrove, "Median-based offset cancellation circuit technique," in *Proc. IEEE Int. Symp. Circuits and Systems*, May 1992, pp. 2033–2036.
- [20] S. S. Narayan, A. M. Peterson, and M. J. Narasimha, "Transform domain LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 609–615, June 1983.
- [21] B. E. Usevitch and M. T. Orchard, "Adaptive filtering using filter banks," *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 255–265, Mar. 1996.
- [22] J. Sonntag, O. Agazzi, P. Aziz, H. Burger, V. Comino, M. Heimann, T. Karanink, J. Khoury, G. Madine, K. Nagaraj, G. Offord, R. Peruzzi, J. Plany, N. Rao, N. Sayiner, P. Setty, and K. Threadgill, "A high speed, low power PRML read channel device," *IEEE Trans. Magn.*, vol. 31, pp. 1186–1195, Mar. 1995.
- [23] S. Kiriaki, T. L. Viswanathan, G. Feygin, B. Staszewski, R. Pierson, B. Krenik, M. de Wit, and K. Nagaraj, "A 160-MHz analog equalizer for magnetic disk read channels," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1839–1850, Nov. 1997.
- [24] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.
- [25] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. New York: Academic, 1980.
- [26] C. R. Rohrs, C. R. Johnson, and J. D. Mills, "A stability problem in sign-sign adaptive algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 4, Apr. 1986, pp. 2999–3001.
- [27] M. Q. Le, P. J. Hurst, and J. P. Keane, "An adaptive analog noise-predictive decision-feedback equalizer," *IEEE J. Solid-State Circuits*, vol. 37, pp. 105–113, Feb. 2002.
- [28] T.-C. Lee and B. Razavi, "A 125-MHz mixed-signal echo canceller for gigabit Ethernet on copper wire," *IEEE J. Solid-State Circuits*, vol. 36, pp. 366–373, Mar. 2001.
- [29] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.
- [30] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [31] D. A. Johns, W. M. Snelgrove, and A. S. Sedra, "Orthonormal ladder filters," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 337–343, Mar. 1989.
- [32] D. Hirsch and W. Wolf, "A simple adaptive equalizer for efficient data transmission," *IEEE Trans. Commun.*, vol. COM-18, p. 5, 1970.
- [33] A. Chan Carusone and D. A. Johns, "A fifth-order G_m - C filter in 0.25- μm CMOS with digitally programmable poles and zeroes," in *Proc. IEEE Int. Symp. Circuits and Syst.*, vol. IV, May 2002, pp. 635–638.



Anthony Chan Carusone (S'96–M'02) received the B.A.Sc. degree from the Engineering Science Division, University of Toronto, Toronto, ON, Canada, in 1997, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Toronto, in 2002.

Since 2001, he has been an Assistant Professor in the Department of Electrical and Computer Engineering, University of Toronto, where he is a member of the electronics research group and the Information Systems Laboratory. He has also been

an occasional consultant to industry for companies such as Analog Devices Inc., Snowbush Inc., and Gennum Corporation. He has several publications in the areas of analog filters, adaptation, and chip-to-chip communications, including a chapter on analog adaptive filters in *Design of High Frequency Integrated Analog Filters* (London, U.K.: IEE Press, 2002). His research is focused on integrated circuits for high-speed signal processing, both analog and digital.

In 1997, Dr. Chan Carusone received the Governor-General's Silver Medal awarded to the student graduating with the highest average from approved Canadian university programs, and from 1997 to 2001, he held Natural Sciences and Engineering Research Council Postgraduate Scholarships. In 2002, he was named Canada Research Chair in Integrated Systems and an Ontario Distinguished Researcher.



David A. Johns (S'81–M'89–SM'94–F'01) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Toronto, Toronto, ON, Canada, in 1980, 1983, and 1989, respectively.

Since 1998, he has been with the University of Toronto where he is currently a full professor. He has ongoing research programs in the general area of analog integrated circuits with particular emphasis on digital communications. His research work has resulted in more than 40 publications. He is co-author of a textbook entitled *Analog Integrated*

Circuit Design (New York: Wiley, 1997) and has given numerous industrial short courses. Together with academic experience, he also has spent a number of years in the semiconductor industry and is a co-founder of Snowbush Microelectronics.

Dr. Johns received 1999 IEEE Darlington Award. He served as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: ANALOG AND DIGITAL SIGNAL PROCESSING from 1993 to 1995, and for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: FUNDAMENTAL THEORY AND APPLICATIONS from 1995 to 1997 and was elected to Adcom for IEEE Solid-State Society (SSC-S) in 2002.