SPLITTING THE RECURSIVE LEAST-SQUARES ALGORITHM

T. Magesacher, S. Haar*, R. Zukunft*, P. Ödling*, T. Nordström*, and P. O. Börjesson*

Telecommunications Research Center Vienna (ftw)
Maderstraße 1, A-1040 Vienna, Austria
magesacher@ftw.at

ABSTRACT

Exponentially weighted recursive least-squares (RLS) algorithms are commonly used for fast adaptation. In many cases the input signals are continuous-time. Either a fully analog implementation of the RLS algorithm is applied or the input data are sampled by analog-to-digital (AD) converters to be processed digitally. Although a digital realization is usually the preferred choice, it becomes unfeasible for high-frequency input signals since fast AD converters are needed. This paper proposes a hybrid analog/digital approach essentially allowing the AD conversion rate to be as low as the update-rate of the RLS algorithm. This is basically accomplished by sampling exponentially weighted correlation products instead of the input signals. Furthermore, we propose a mixed-signal filter exactly realizing the exponential weighting according to the cost function. Applying this approach to an interference cancelling problem demonstrates its performance and attractiveness regarding implementation.

1. INTRODUCTION

Adaptive algorithms are predominantly discussed in their discrete-time versions in literature, see [1] and its extensive list of references. However, there are cases where the observable data are continuous-time and furthermore of high-frequency. Recursive least-squares (RLS) lattice and fast transversal filters for continuous-time signal processing have already been proposed in [2] and [3], respectively. Hybrid analog/digital signal processing is known to have the potential of combining the best of both analog and digital worlds [4]. Analog hardware can handle high-frequency signals more efficiently but is limited, mostly, to simple and preferably linear operations. Digital signal processing can easily deal with nonlinear operations but is limited to comparatively low operating rates.

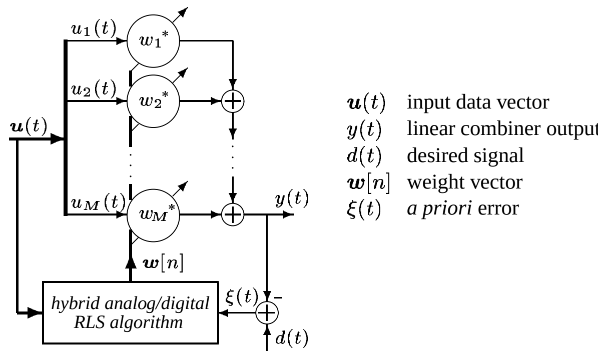

Figure 1: Continuous-time multiple-input adaptive linear combiner. The weights are updated by a hybrid analog/digital RLS algorithm.

In this paper we focus on splitting the RLS algorithm, applied to the multiple-input adaptive linear combiner operating on continuous-time input data (Figure 1), into an analog and a digital part. In Section 2 the derivation of the RLS algorithm is modified accordingly. An input signal weighting-filter that is optimum in the sense of minimizing the error function is derived. Section 3 proposes a mixed-signal realization of the filter. In Section 4, we demonstrate the feasibility of our approach applying it to an interference cancellation problem. Finally, Section 5 concludes the work.

2. RLS ADAPTATION OF THE CONTINUOUS-TIME LINEAR COMBINER

Figure 1 shows the continuous-time multiple-input linear combiner [5] adapted by an RLS algorithm. Unlike previous work, where both signal path and weight updating was considered to be either discrete-time or continuous-time, we intend to perform the weight update digitally while the signal path is processed in analog domain. Thus the weights

* Thomas Magesacher is with Infineon Technologies, Austria. He is also a PhD candidate at BridgeLab, LIS, Munich Techn. Univ., Germany.
† Sven Haar and Roland Zukunft are with BridgeLab, LIS, Munich Techn. Univ., Germany.
‡ Dr. Per Ödling is a Senior Researcher with the Telecommunications Research Center Vienna (ftw), Austria and Ass. Prof. with Lund Institute of Technology, Department of Electroscience, Sweden.
§ Dr. Tomas Nordström is a Senior Researcher with ftw, Austria.
¶ Prof. Per Ola Börjesson is with Lund Institute of Technology, Department of Electroscience, Sweden.

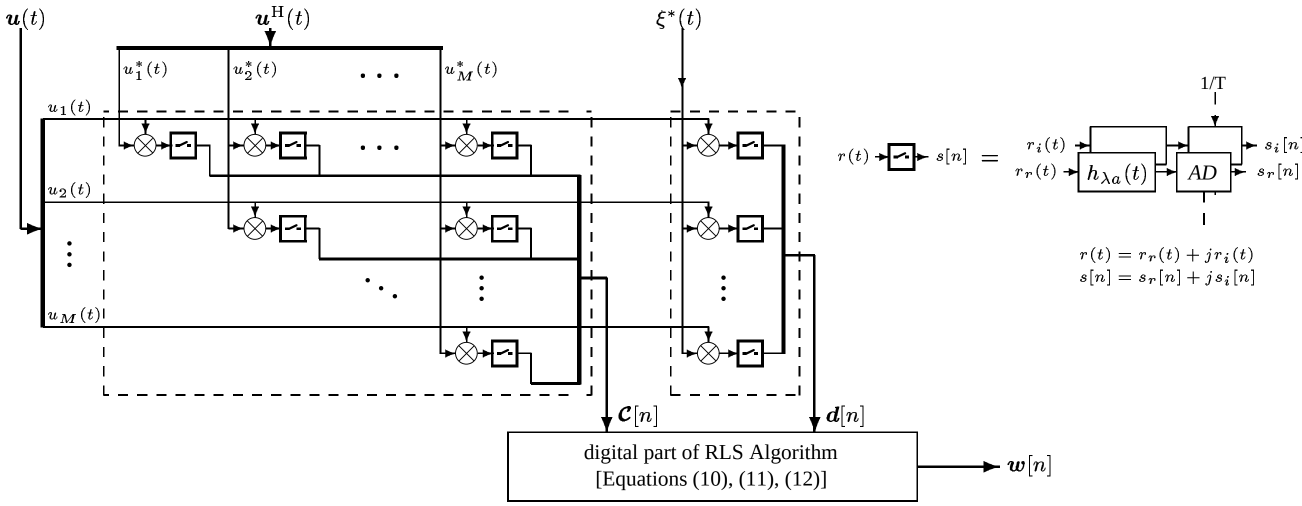

Figure 2: Hybrid analog/digital RLS algorithm. The correlation products are sampled after exponential weighting.

$\boldsymbol{w}(t)$ are piecewise-constant functions. Deriving the RLS algorithm for continuous-time input data, we will follow [1] using the notation introduced in Figure 1.

We base our analysis on a multiple linear regression model for the desired signal, that is,

$$d(t) = \boldsymbol{w}_o^{\mathrm{H}}(t)\,\boldsymbol{u}(t) + e_o(t), \tag{1}$$

where $\boldsymbol{w}_o(t)$ is the optimum weight vector, $e_o(t)$ is the measurement error, and $(\cdot)^{\mathrm{H}}$ denotes the Hermitian transpose. Under the assumption that the expectation of the unobservable measurement error $E[e_o(t)] = 0$, we choose the linear adaptive combiner with its output signal $y(t) = \boldsymbol{w}^{\mathrm{H}}(t)\,\boldsymbol{u}(t)$ as the model of interest. The estimation error $e(t)$ is given by

$$\begin{aligned} e(t) &= d(t) - y(t) \qquad\qquad (2)\\ &= d(t) - \boldsymbol{w}^{\mathrm{H}}(t)\,\boldsymbol{u}(t)\\ &= d(t) - \boldsymbol{w}^{\mathrm{H}}[n]\,\boldsymbol{u}(t), \qquad (n-1)T \le t < nT. \end{aligned}$$

We define the cost function

$$\mathcal{E}[n] = \frac{1}{T}\int_{t=0}^{nT} \sigma\lambda^{\frac{nT-t}{T}}\,\left|e(t)\right|^2\,\mathrm{d}t, \tag{3}$$

where λ is a forgetting factor weighting recent data higher than older data, and σ is a scalar constant. Note that the coefficients are held constant during the observation interval $0 \le t \le nT$. Minimizing the cost function we arrive at

$$\hat{\boldsymbol{w}}[n] = \boldsymbol{\Phi}^{-1}[n]\,\boldsymbol{z}[n]. \tag{4}$$

The vector $\boldsymbol{z}$[n] contains the time-averaged cross-correlation functions between the input data and the desired response, that is,

$$z_k[n] = \frac{1}{T}\int_{t=0}^{nT} \sigma\lambda^{\frac{nT-t}{T}}\,d^*(t)\,u_k(t)\,\mathrm{d}t. \tag{5}$$

The elements of the time-averaged correlation matrix $\boldsymbol{\Phi}[n]$ are given by

$$\Phi_{k,l}[n] = \frac{1}{T}\int_{t=0}^{nT} \sigma\lambda^{\frac{nT-t}{T}}\,u_k(t)\,u_l^*(t)\,\mathrm{d}t. \tag{6}$$

Now we want to find a recursive expression for $\boldsymbol{z}[n]$. This can be accomplished by splitting the integral in (5), that is,

$$\begin{aligned} \boldsymbol{z}[n] &= \lambda\,\frac{1}{T}\int_{t=0}^{(n-1)T} \sigma\lambda^{\frac{nT-t-T}{T}}\,d^*(t)\,\boldsymbol{u}(t)\,\mathrm{d}t\\ &\quad + \frac{1}{T}\int_{t=(n-1)T}^{nT} \sigma\lambda^{\frac{nT-t}{T}}\,d^*(t)\,\boldsymbol{u}(t)\,\mathrm{d}t\\ &= \lambda\,\boldsymbol{z}[n-1] + \frac{1}{T}\int_{t=(n-1)T}^{nT} \sigma\lambda^{\frac{nT-t}{T}}\,d^*(t)\,\boldsymbol{u}(t)\,\mathrm{d}t. \end{aligned} \tag{7}$$

We may as well express the entire correlation matrix recursively:

$$\underbrace{\boldsymbol{\Phi}[n]}_{\boldsymbol{\mathcal{A}}} = \underbrace{\lambda\,\boldsymbol{\Phi}[n-1]}_{\boldsymbol{\mathcal{B}}^{-1}} + \underbrace{\frac{1}{T}\int_{t=(n-1)T}^{nT} \sigma\lambda^{\frac{nT-t}{T}}\,\boldsymbol{u}(t)\,\boldsymbol{u}^{\mathrm{H}}(t)\,\mathrm{d}t}_{\boldsymbol{\mathcal{C}}}. \tag{8}$$

In order to find a recursive update rule for the inverse of the correlation matrix, we introduce a modified version of the matrix inversion lemma:

$$\begin{aligned} \boldsymbol{\mathcal{A}} &= \boldsymbol{\mathcal{B}}^{-1} + \boldsymbol{\mathcal{C}}\\ \boldsymbol{\mathcal{A}}^{-1} &= \boldsymbol{\mathcal{B}} - \frac{1}{1 + \mathbf{1}^{\mathrm{T}}\left(\boldsymbol{\mathcal{B}}^{\mathrm{T}} \circ \boldsymbol{\mathcal{C}}\right)\mathbf{1}}\,\boldsymbol{\mathcal{B}}\,\boldsymbol{\mathcal{C}}\,\boldsymbol{\mathcal{B}}. \end{aligned} \tag{9}$$

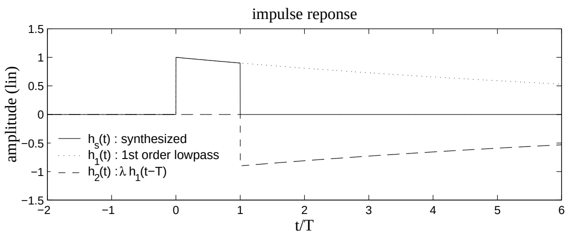

Figure 3: Synthesis of the optimum weighting filter.

The M-by-M matrices $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ are positive-definite, $\boldsymbol{\mathcal{C}}$ is a Hermitian matrix of size M-by-M, $\mathbf{1}$ is the all-1-vector of length M, and $\circ$ denotes the element-wise multiplication of two matrices (*Schur product*). Thus the recursion for $\boldsymbol{\mathcal{P}}[n] = \boldsymbol{\Phi}^{-1}[n]$ results in

$$\boldsymbol{\mathcal{P}}[n]=\frac{1}{\lambda}\boldsymbol{\mathcal{P}}[n-1]-\frac{1}{\lambda}\boldsymbol{\mathcal{K}}[n]\underbrace{\left[\frac{1}{T}\int_{t=(n-1)T}^{nT}\sigma\lambda^{\frac{nT-t}{T}}\boldsymbol{u}(t)\boldsymbol{u}^{\mathrm{H}}(t)\mathrm{d}t\right]}_{\boldsymbol{\mathcal{C}}[n]}\boldsymbol{\mathcal{P}}[n-1] \tag{10}$$

with

$$\boldsymbol{\mathcal{K}}[n]=\frac{\frac{1}{\lambda}\boldsymbol{\mathcal{P}}[n-1]}{1+\frac{1}{\lambda}\mathbf{1}^{\mathrm{T}}\left(\boldsymbol{\mathcal{P}}^{\mathrm{T}}[n-1]\circ\boldsymbol{\mathcal{C}}[n]\right)\mathbf{1}}. \tag{11}$$

Substituting (7) and (10) in (4) yields, after a few manipulations, the recursive update rule for the weights:

$$\hat{\boldsymbol{w}}[n] = \hat{\boldsymbol{w}}[n-1] + \boldsymbol{\mathcal{K}}[n]\underbrace{\left(\frac{1}{T}\int_{t=(n-1)T}^{nT}\sigma\lambda^{\frac{nT-t}{T}}\,\boldsymbol{u}(t)\xi^*(t)\,\mathrm{d}t\right)}_{\boldsymbol{d}[n]}. \tag{12}$$

The *a priori* estimation error $\xi(t)$ is given by

$$\xi(t) = d(t) - \hat{\boldsymbol{w}}^{\mathrm{H}}[n]\,\boldsymbol{u}(t), \qquad nT \le t < (n+1)T. \tag{13}$$

The recursive update-rules (10), (11), and (12) constitute the hybrid analog/digital RLS algorithm. The integrals in $\boldsymbol{\mathcal{C}}[n]$ and $\boldsymbol{d}[n]$ correspond to the convolution of $\frac{1}{T}\,\sigma\lambda^{\frac{t}{T}}$ with $u_k(t)u_l^*(t)$ and $u_k(t)\xi^*(t)$, respectively, followed by sampling at time instants $t=nT$. Thus applying a filter with the impulse response

$$h_{\lambda a}(t) = \begin{cases} \frac{1}{T}\,\sigma\lambda^{\frac{t}{T}} & 0 \le t \le T \\ 0 & \text{else} \end{cases} \tag{14}$$

to the signals $\boldsymbol{u}(t)\boldsymbol{u}^{\mathrm{H}}(t)$ and $\boldsymbol{u}(t)\xi^*(t)$ and sampling the results every T seconds, as depicted in Figure 2, is optimum in the sense of minimizing $\mathcal{E}[n]$. Note that it is sufficient to process the products $u_k(t)u_l^*(t)$ for $l \ge k$ since the matrix $\boldsymbol{\mathcal{C}}[n]$ is Hermitian.

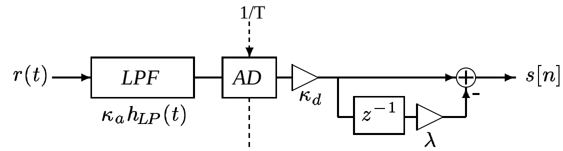

Figure 4: Mixed-signal realization of the optimum weighting filter: first-order analog lowpass filter, AD converter with sampling rate $1/T$, and digital highpass filter.

3. MIXED-SIGNAL REALIZATION OF THE OPTIMUM WEIGHTING-FILTER

The above derivation of the mixed-signal RLS algorithm does not introduce any constraints concerning the scalar coefficient σ, except $\sigma \ge 0$. Thus σ may be chosen arbitrarily. However, regarding the implementation, it is reasonable to ensure proper scaling of the filter output signal. Setting $\sigma = \frac{ln(\lambda)}{\lambda-1}$, $(0<\lambda<1)$, assures a gain of 1 for low frequencies.

The exponential nature of (14) motivates the use of a continuous-time first-order lowpass filter whose impulse response is given by

$$h_{LP}(t) = \frac{1}{\tau}\,e^{-\frac{t}{\tau}}, \tag{15}$$

where τ is the filter's time constant. Comparing (14) and (15) yields $\tau = -\frac{T}{ln(\lambda)}$, $(0<\lambda<1)$. The desired impulse response is realized for $t<T$ by scaling $h_{LP}(t)$:

$$h_1(t) = \kappa\, h_{LP}(t), \qquad \kappa = \kappa_a\kappa_d = \sigma\frac{\tau}{T}. \tag{16}$$

The scaling constant may be split into an analog gain κ_a and a digital gain κ_d. To eliminate the tail of $h_1(t)$ for $t>T$ we combine it additively with a second impulse response $h_2(t) = -\lambda\, h_1(t-T)$ to yield the synthesized response $h_s(t) = h_1(t) + h_2(t)$, which is illustrated in Figure 3. Let $r(t)$ be the lowpass filter input signal. The corresponding filter output $s(t)$ is given by

$$\begin{aligned} s(t) &= r(t) * h_s(t) \\ &= \int_{\tau=-\infty}^{\infty} r(\tau)\, h_1(t-\tau)\, d\tau - \lambda\int_{\tau=-\infty}^{\infty} r(\tau)\, h_1(t-T-\tau)\, d\tau \\ &= \quad s_1(t) - \lambda\, s_1(t-T), \end{aligned} \tag{17}$$

where * denotes the convolution operator. Since the algorithm needs only the values of $s(t)$ at time instances $t=nT$, we sample $s(t)$, that is,

$$s(nT) = s[n] = s_1[n] - \lambda\, s_1[n-1]. \tag{18}$$

This result yields the mixed-signal realization of (14) shown in Figure 4.

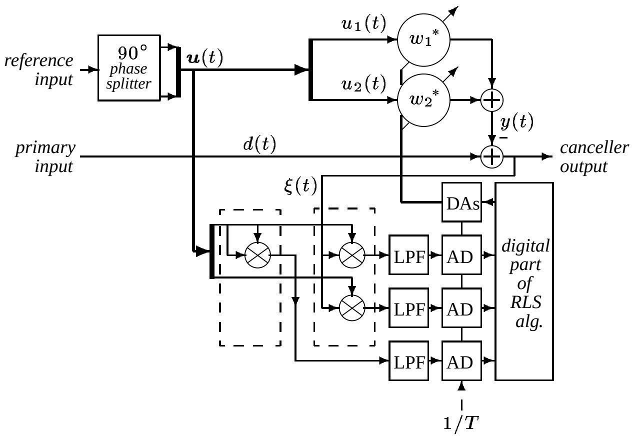

Figure 5: Application of the splitting approach to narrowband interference cancellation (real valued, orthogonal input signals; $M{=}2$).

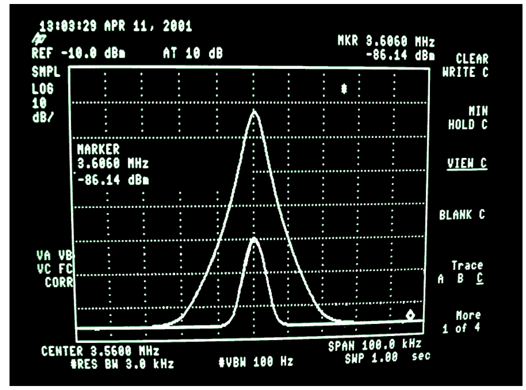

Figure 6: Measured steady state suppression (after $\approx$ 10 iterations) of a 12 MHz sinusoid, $1/T = 10\,kHz$.

4. APPLICATION TO INTERFERENCE CANCELLATION

The elimination of a single-frequency disturbance is a common problem in communications. In case a reference signal, somehow correlated to the disturbance, is available, cancellation is an efficient way to combat the interference [6]. The reference signal, which is essentially a time-shifted and scaled version of the disturbance, is split into two orthogonal components that constitute the input data, as shown in Figure 5. Accordingly, $M{=}2$ for our single-frequency case. Furthermore, all signals are real valued. However, since the two elements of the input vector are mutually orthogonal, we may introduce two assumptions. First, the time-averaged cross-products of $u_1(t)$ and $u_2(t)$ are zero, so all off-diagonal elements of $\mathcal{C}[n]$ vanish. Secondly, $u_1(t)$ and $u_2(t)$ have the same time-averaged power, thus all main diagonal elements of $\mathcal{C}[n]$ are equal. Hence, we just have to process one main-diagonal element of $\mathcal{C}[n]$.

The interference canceller shown in Figure 5 is applied in wireline communications [7] where narrowband interference of several MHz, caused by radio amateurs, may severely disturb data transmission. The choice of the update-rate $1/T$ depends mainly on the tracking requirements introduced by the application. For reference-based interference cancellation the changes of correlation between disturber and reference have to be tracked. In comparison to the high-frequency disturbance, the correlation between disturber and reference varies very slowly. Thus $1/T$ can be chosen to be orders of magnitude below the frequency of the disturber. Prototype measurements show a steady state interference suppression of $\approx$ 40 dB (Figure 6).

Our approach requires three AD converters for this application. Note that direct AD conversion of the input signals would also need three AD converters. Due to splitting the algorithm the sampling rate may be as low as $1/T$.

5. CONCLUSIONS

In this paper we split the RLS algorithm, applied to adapt the continuous-time multiple-input linear combiner, into an analog and a digital part. There are two major contributions: First, we demonstrate the feasibility of the hybrid analog/digital approach for high-frequency input data. The sampling rate of the converters is reduced to the algorithm's update-rate which may be orders of magnitude below the signal's frequency. Secondly, we propose a mixed-signal realization of the weighting-filter obtained by incorporating the time-continuity of the input signals in the derivation of the RLS algorithm. We show the elegance of this approach applying it to narrowband interference cancellation.

6. REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, ISBN 0-13-322760-X, third edition, 1996.

[2] H. Lev-Ari, “Continuous-time, discrete-order lattice filters,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, Apr. 1986, pp. 2947–2950.

[3] H. Lev-Ari, J. Cioffi, and T. Kailath, “Continuous-time least-squares fast transversal filters,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, Apr. 1987, pp. 415–418.

[4] H. Lev-Ari, T. Kailath, and J. Cioffi, “Adaptive Recursive-Least-Squares Lattice and Transversal Filters for Continuous-Time Signal Processing,” *IEEE Trans. Circuits Syst. II*, vol. 29, no. 2, pp. 81–89, Feb. 1992.

[5] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, 1985.

[6] B. Widrow, “Adaptive Noise Cancelling: Principles and Applications,” in *Proc. IEEE*, 1975, vol. 63, pp. 1692–1716.

[7] T. Magesacher, P. Ödling, T. Nordström, T. Lundberg, M. Isaksson, and P. O. Börjesson, “An Adaptive Mixed-Signal Narrowband Interference Canceller for Wireline Transmission Systems,” in *Proc. IEEE Int. Symp. Circuits and Systems*, May 2001.