Iterative analysis of the steady-state weight fluctuations in LMS-type adaptive filters

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

[11] ——, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE Trans. Signal Processing*, vol. 39, pp. 92–114, Jan. 1991.

[12] S. Ljung and L. Ljung, "Error propagation properties of recursive least-squares adaptation algorithms," *Automatica*, vol. 21, no. 2, pp. 157–167, 1985.

[13] H. Lev-Ari, T. Kailath, and J. Cioffi, "Least squares adaptive lattice and transversal filters: A unified geometric theory," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 222–236, Mar. 1984.

[14] W. D. Lin, "On digital implementation of the fast Kalman algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, Apr. 1984.

[15] G. V. Moustakides, "Correcting the instability due to finite precision of the fast Kalman identification algorithms," *Signal Process.*, vol. 18, no. 1, Sept. 1989.

[16] P. Fabre and C. Gueguen, "Improvement of the fast recursive least-squares algorithms via normalization: A comparative study," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 296–308, Apr. 1986.

[17] A. D. Pouliezos and G. S. Stavrakakis, *Real Time Fault Monitoring of Industrial Processes*. Boston, MA: Kluwer, Apr. 1994.

[18] S. H. Ardalan and S. T. Alexander, "Fixed-point roundoff error analysis of the exponentially windowed RLS algorithm for time-varying systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 770–783, June 1987.

[19] J. M. Cioffi, "Limited-precision effects in adaptive filtering," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 821–833, July 1987.

[20] G.-O. A. Glentis and N. Kalouptsidis, "Fast adaptive algorithms for multichannel filtering and system identification," *IEEE Trans. Signal Processing*, vol. 40, pp. 2433–2458, Oct. 1992.

[21] J. G. Proakis and D. G. Manolakis, *Introduction to Digital Signal Processing*. New York: Macmillan.

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

[23] C. Papaodysseus, E. Koukoutsis, C. Triantafyllou, and C. Vassilatos, "Exact monitoring of the numerical error in various speech algorithms," in *Proc. Eurospeech*, 1991, vol. 3, pp. 1073–1076.

[24] C. Papaodysseus, E. Koukoutsis, and C. Vassilatos, "Accurate prediction of the numerical error generated in the forward linear prediction algorithms," *ISMM Int. J. Microcomput. Appl.*, vol. 12, no. 1, pp. 1–6, 1993.

[25] C. Papaodysseus, C. Triantafyllou, E. Koukoutsis, and G. Carayannis, "Error propagation and numerical recovery for a class of parametric DSP algorithms," in *Proc. ICASSP*, Apr. 1990, pp. 1349–1352.

[26] C. Papaodysseus, G. Carayannis, E. Koukoutsis, and E. Kayafas, "Comparing l-step ahead prediction and LS FIR filtering," *IEEE Trans. Signal Processing*, vol. 41, pp. 768–780, Feb. 1993.

[27] C. Papaodysseus, E. Koukoutsis, and C. Triantafyllou, "Error sources and error propagation in the Levinson-Durbin algorithm," *IEEE Trans. Signal Processing*, vol. 41, pp. 1635–1651, Apr. 1993.

[28] C. Papaodysseus, E. Koukoutsis, and C. Vassilatos, "Error sources, error propagation and methods of error correction in LS FIR filtering and l-step ahead prediction," *IEEE Trans. Signal Processing*, vol. 42, pp. 1097–1108, May 1994.

Iterative Analysis of the Steady-State Weight Fluctuations in LMS-Type Adaptive Filters

H. J. Butterweck

Abstract—An iterative method is proposed for the analysis of the steady-state weight fluctuations in an LMS-type adaptive FIR filter. Without the widely used independence assumption, a power series of the weight-error correlation matrix is derived in terms of the stepsize. Some new effects are observed, e.g., a decrease of the weight fluctuations along the tapped-delay line.

Index Terms—Independence assumption, LMS adaptive filters, steady-state analysis, weight-order correlation matrix.

I. INTRODUCTION

In a stationary environment and with all adaptive transients died out, an LMS-type adaptive filter performs random fluctuations of its weighting coefficients around the optimal "Wiener solution," viz. the set of coefficients of some (actual or imaginary) filter that the adaptive filter attempts to imitate. In current literature, the weight error correlation matrix (WECM) is throughout determined with the aid of an "*independence assumption*" stating statistical independence of successive input vectors. Such an assumption can convincingly only be justified for a true vector signal like that emerging from a sensor array but not for a *tapped-delay line* (TDL) with a strong deterministic coherence between the input vectors. In that situation, the use of the assumption can be justified by a fair agreement between theoretical and experimental results.

It appears that Florian and Feuer [9] were the first to deliberately abandon the assumption for a TDL filter of length 2 and analyzing it for a white input signal and a white noise. Subsequently, Douglas *et al.* [4], [10] extended this work for longer filters and a colored input signal. Their exact computer-aided analysis applies for all values of the stepsize so that also the convergence issue can be addressed. However, the analysis becomes computationally burdensome for "long" filters (in fact, only a few taps can be handled).

In this correspondence, we describe an iterative approach without the independence assumption which leads to a power series for the WECM in terms of the stepsize, cf. also [5] and [1, App. I]. Since only the first few terms of the series are simple enough, we have to confine ourselves to *small stepsizes*, sufficiently below the stability bound. While independence theory presupposes white Gaussian output noise, our analysis does not require such a limitation. It is only for convenience, that we here adopt the *whiteness* assumption for the output noise, too.

II. ITERATIVE SOLUTION OF THE UPDATING EQUATION

An adaptive TDL filter with the time-varying $M \times 1$ weight vector $\boldsymbol{w}_k$ tries to imitate a fixed TDL filter with the $M \times 1$ weight vector $\boldsymbol{h}$. The input signal x_k and the additive noise n_k at the filter output are sample functions of statistically independent, real-valued, stationary,

zero-mean random processes. Under *steady-state conditions*, i.e., after completion of the adaptive process, the $M \times 1$ "weight error vector" $\boldsymbol{v}_k = \boldsymbol{w}_k - \boldsymbol{h}$ becomes stationary as well, with zero mean. Let the adaptive mechanism be governed by the LMS algorithm, with the *updating rule*

$$\boldsymbol{v}_{k+1} = (I - 2\mu \boldsymbol{x}_k \boldsymbol{x}_k^t)\boldsymbol{v}_k + 2\mu \boldsymbol{f}_k$$
$$\boldsymbol{f}_k = n_k \boldsymbol{x}_k \qquad (1)$$

where $\boldsymbol{x}_k$ denotes the $M \times 1$ "input vector" $(x_k, x_{k-1}, \cdots, x_{k-M+1})^t$, and μ denotes the "stepsize." Our aim is to determine the $M \times M$ WECM $V = E\{\boldsymbol{v}_k \boldsymbol{v}_k^t\}$, whose diagonal elements V_{mm} denote the powers of the weight fluctuations, whereas the off-diagonal element V_{mn} stands for their mutual correlations.

What occurs for sufficiently small μ values? Then, $\boldsymbol{v}_k$ varies so slowly that it can be approximated by some $\boldsymbol{\alpha}_k$ satisfying the difference equation

$$\boldsymbol{\alpha}_{k+1} = (I - 2\mu R)\boldsymbol{\alpha}_k + 2\mu \boldsymbol{f}_k$$
$$R = E\{\boldsymbol{x}_k \boldsymbol{x}_k^t\}. \qquad (2)$$

Thus, in the limiting case $\mu \to 0$, the time-varying coefficient $\boldsymbol{x}_k \boldsymbol{x}_k^t$ in (1) can be replaced with its average: the "correlation matrix" R. This statement generally phrased as "direct averaging" [2], [6] is based on the insight that (1) and (2) describe the same large-scale behavior, which is found through averaging the difference equations over a sufficient number of consecutive time instants and replacing time averaging with ensemble averaging. For larger μ values, the *zero-order solution* $\boldsymbol{\alpha}_k$ can be corrected by an iteration. Writing $\boldsymbol{v}_k = \boldsymbol{\alpha}_k + \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k + \cdots$, insertion into (1) yields

$$\boldsymbol{\alpha}_{k+1} + \boldsymbol{\beta}_{k+1} + \boldsymbol{\gamma}_{k+1} + \cdots$$
$$= (I - 2\mu R)(\boldsymbol{\alpha}_k + \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k + \cdots)$$
$$- 2\mu P_k (\boldsymbol{\alpha}_k + \boldsymbol{\beta}_k + \cdots) + 2\mu \boldsymbol{f}_k$$
$$P_k = \boldsymbol{x}_k \boldsymbol{x}_k^t - R = P_k^t; \quad E\{P_k\} = 0. \qquad (3)$$

The iterative solution of (3) proceeds as follows:

$$\boldsymbol{\alpha}_{k+1} = (I - 2\mu R)\boldsymbol{\alpha}_k + 2\mu \boldsymbol{f}_k$$
$$\boldsymbol{\beta}_{k+1} = (I - 2\mu R)\boldsymbol{\beta}_k - 2\mu P_k \boldsymbol{\alpha}_k$$
$$\boldsymbol{\gamma}_{k+1} = (I - 2\mu R)\boldsymbol{\gamma}_k - 2\mu P_k \boldsymbol{\beta}_k \qquad (4)$$

etc. The zero-order solution $\boldsymbol{\alpha}_k$ is determined by $\boldsymbol{f}_k$, whereupon the "*first-order correction*" $\boldsymbol{\beta}_k$ follows from $\boldsymbol{\alpha}_k$, the "*second-order correction*" $\boldsymbol{\gamma}_k$ follows from $\boldsymbol{\beta}_k$, etc. Thus, we proceed according to $\boldsymbol{f}_k \to \boldsymbol{\alpha}_k \to \boldsymbol{\beta}_k \to \boldsymbol{\gamma}_k \to \cdots$, where for sufficiently small μ, the terms in the chain decrease to any wanted degree. Observe that the same operator $\mathcal{L}$ applies in all steps of the above scheme: $\boldsymbol{\alpha}_k = \mathcal{L}\{2\mu \boldsymbol{f}_k\}, \boldsymbol{\beta}_k = \mathcal{L}\{-2\mu P_k \boldsymbol{\alpha}_k\}, \boldsymbol{\gamma}_k = \mathcal{L}\{-2\mu P_k \boldsymbol{\beta}_k\}$, etc. It represents a simple linear time-invariant filtering of the low-pass type

that is explicitly governed by the convolutional relation

$$\boldsymbol{\alpha}_k = \mathcal{L}\{2\mu \boldsymbol{f}_k\} = \sum_{j=-\infty}^{\infty} H_j (2\mu \boldsymbol{f}_{k-j})$$
$$H_j = u_{j-1}(I - 2\mu R)^{j-1} = H_j^t$$
$$u_j = 0 \quad \text{for} \quad j < 0$$
$$u_j = 1 \quad \text{for} \quad j \geq 0. \qquad (5)$$

The above iteration is attractive in that it transforms the time-varying system parameter $\boldsymbol{x}_k \boldsymbol{x}_k^t$ in (1) into a set of excitation functions $(-P_k \boldsymbol{\alpha}_k), (-P_k \boldsymbol{\beta}_k), \cdots$ in (4) serving as source terms in simple constant-coefficient updating equations. Thus, the problem is reduced to a study of the passage of stationary stochastic signals through a low-pass system, whose cut-off frequency is extremely low for $\mu \to 0$.

III. SERIES EXPANSION OF THE WEIGHT-ERROR CORRELATION MATRIX

With the expansion $\boldsymbol{v}_k = \boldsymbol{\alpha}_k + \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k + \cdots$ the WECM can be written as

$$V = E\{\boldsymbol{v}_k \boldsymbol{v}_k^t\} = E\{(\boldsymbol{\alpha}_k + \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k + \cdots)$$
$$\cdot (\boldsymbol{\alpha}_k^t + \boldsymbol{\beta}_k^t + \boldsymbol{\gamma}_k^t + \cdots)\}$$
$$= E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^t\} + [E\{\boldsymbol{\alpha}_k \boldsymbol{\beta}_k^t\} + E\{\boldsymbol{\beta}_k \boldsymbol{\alpha}_k^t\}$$
$$+ E\{\boldsymbol{\gamma}_k \boldsymbol{\alpha}_k^t\} + E\{\boldsymbol{\alpha}_k \boldsymbol{\gamma}_k^t\} + E\{\boldsymbol{\beta}_k \boldsymbol{\beta}_k^t\}] + \cdots \qquad (6)$$

where the first term is $O(\mu^1)$ [cf. (8)], whereas the term between squared brackets is $O(\mu^2)$, cf. (10), and the omitted terms are of third (and higher) order. (Observe the relatively low value of $E\{\boldsymbol{\alpha}_k \boldsymbol{\beta}_k^t\}$ due to a small degree of correlation between $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ [8]). Obviously, (6) represents a Taylor series expansion of the form

$$V = V_1 \mu + V_2 \mu^2 + \cdots \qquad (7)$$

with $E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^t\}$ contributing to $V_1 \mu, V_2 \mu^2, V_3 \mu^3$ etc., the bracketed expression in (6) contributing to $V_2 \mu^2, V_3 \mu^3$ etc., and the remainder of (6) contributing to $V_3 \mu^3$ etc. The terms $V_1 \mu$ and $V_2 \mu^2$ will be determined below, followed by a brief discussion of the term $V_3 \mu^3$, which was evaluated elsewhere [8].

Under the assumption of *white additive noise* $E\{n_k n_{k-i}\} = N\delta_i$, the evaluation of the first term in (6) becomes simple. With (5) and (1), we obtain

$$E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^t\} = 4\mu^2 E \sum_i \sum_j H_i \boldsymbol{f}_{k-i} \boldsymbol{f}_{k-j}^t H_j$$
$$= 4\mu^2 N \sum_i H_i R H_i = 4\mu^2 N R \sum_i H_i^2$$
$$= \mu N[I - \mu R]^{-1} = \mu N I + \mu^2 N R + O(\mu^3). \qquad (8)$$

The first right-hand term $\mu N I$, which is equal to $V_1 \mu$, represents a set of uncorrelated weight fluctuations with equal power μN independent of the amplitude and spectral distribution of the input signal. In passing, we note that in case of colored output noise

[2], [7], where V_1 satisfies the Lyapounov equation $RV_1 + V_1 R = 2\sum_l E\{\boldsymbol{f}_k \boldsymbol{f}_{k-l}^t\}$, this independence is lost.

Next, we consider the bracketed expression in (6) contributing to $V_2 \mu^2, V_3 \mu^3$, etc. Only the leading terms of the five correlations (four of which occur in transpose pairs) need to be evaluated, which allows appropriate approximations:

$$
\begin{aligned}
E\{\boldsymbol{\beta}_k \boldsymbol{\alpha}_k^t\} \\
&= -2\mu E \sum_i H_i P_{k-i} \boldsymbol{\alpha}_{k-i} \boldsymbol{\alpha}_k^t \\
&= -8\mu^3 E \sum_i \sum_j \sum_\ell H_i P_{k-i} H_j \boldsymbol{f}_{k-i-j} \boldsymbol{f}_{k-\ell}^t H_\ell \\
&= -8\mu^3 N E \sum_i \sum_j H_i P_{k-i} H_j \boldsymbol{x}_{k-i-j} \boldsymbol{x}_{k-i-j}^t H_{i+j} \\
&= -8\mu^3 N E \sum_i \sum_j H_i P_{k-i} H_j P_{k-i-j} H_{i+j} \\
&= -8\mu^3 N \sum_i \sum_j H_i E\{P_k H_j P_{k-j}\} H_{i+j} \\
&= E\{\boldsymbol{\alpha}_k \boldsymbol{\beta}_k^t\}^t
\end{aligned}
$$

$$
\begin{aligned}
E\{\boldsymbol{\gamma}_k \boldsymbol{\alpha}_k^t\} \\
&= -2\mu E \sum_i H_i P_{k-i} \boldsymbol{\beta}_{k-i} \boldsymbol{\alpha}_k^t \\
&= 4\mu^2 E \sum_i \sum_j H_i P_{k-i} H_j P_{k-i-j} \boldsymbol{\alpha}_{k-i-j} \boldsymbol{\alpha}_k^t \\
&\approx 4\mu^2 \sum_i \sum_j H_i E\{P_{k-i} H_j P_{k-i-j}\} E\{\boldsymbol{\alpha}_{k-i-j} \boldsymbol{\alpha}_k^t\} \\
&\approx 4\mu^3 N \sum_i \sum_j H_i E\{P_k H_j P_{k-j}\} H_{i+j} \\
&= E\{\boldsymbol{\alpha}_k \boldsymbol{\gamma}_k^t\}^t
\end{aligned}
$$

$$
\begin{aligned}
E\{\boldsymbol{\beta}_k \boldsymbol{\beta}_k^t\} \\
&= 4\mu^2 E \sum_i \sum_j H_i P_{k-i} \boldsymbol{\alpha}_{k-i} \boldsymbol{\alpha}_{k-j}^t P_{k-j} H_j \\
&\approx 4\mu^2 E \sum_i \sum_j H_i P_{k-i} E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_{k-(j-i)}^t\} P_{k-j} H_j \\
&= 4\mu^2 E \sum_i \sum_j H_i P_{k-i} E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_{k-j}^t\} P_{k-i-j} H_{i+j} \\
&\approx 4\mu^3 N \sum_i \sum_j H_i \\
&\quad \cdot E\{P_k (H_j + H_{-j} + \delta_j) P_{k-j}\} H_{i+j}
\end{aligned}
$$

where use has been made of the identity

$$
\begin{aligned}
E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_{k-\ell}^t\} &= 4\mu^2 E \sum_i \sum_j H_i \boldsymbol{f}_{k-i} \boldsymbol{f}_{k-\ell-j}^t H_j \\
&= 4\mu^2 N \sum_i H_i R H_{i-\ell} \\
&= \mu N (I - 2\mu R)^{|\ell|} (I - \mu R)^{-1} \\
&\approx \mu N (I - 2\mu R)^{|\ell|} \\
&\approx \mu N (H_\ell + H_{-\ell} + \delta_\ell).
\end{aligned}
$$

Evaluating the first two terms of (6) finally yields (all terms in the double sums $\sum_i \sum_j$ with $i \neq j$ cancel out)

$$
V_1 \mu + V_2 \mu^2 = \mu N I + \mu N \left[\mu R + 4\mu^2 \sum_i H_i E\{P_k^2\} H_i \right]. \quad (9)
$$

For a Gaussian input signal with

$$
\begin{aligned}
E\{P_k^2\} &= E\{(\boldsymbol{x}_k \boldsymbol{x}_k^t - R)^2\} \\
&= E\{\boldsymbol{x}_k \boldsymbol{x}_k^t \boldsymbol{x}_k \boldsymbol{x}_k^t\} - R^2 = R^2 + R \operatorname{tr} R
\end{aligned}
$$

we find the closed-form solution

$$
\begin{aligned}
V &\approx V_1 \mu + V_2 \mu^2 = \mu N \{I(1 + \mu M X) + 2\mu R\} \\
X &= E\{x_k^2\} = \operatorname{tr} R/M. \quad (10)
\end{aligned}
$$

Thus, in the quadratic approximation, the weight-error correlation matrix V has a Toeplitz structure, provided n_k is white and x_k is Gaussian. Outside the diagonal, it is proportional to R. Thus, the correlation between the (slow) weight fluctuations $v_{k,i}$ and $v_{k,j}$ at points i and j $(i \neq j)$ depends on the distance $|i - j|$ only and is proportional to the correlation between the (fast) "tap signals" x_{k-i+1}, x_{k-j+1} with a common proportionality factor $2\mu^2 N$. The "degree of correlation" between different weight errors defined as the ratio of the crosscorrelation and the autocorrelation approximately equals $R_{ij} = 2\mu E\{x_{k-i+1} x_{k-j+1}\} = 2\mu E\{x_k x_{k-(i-j)}\}$ and, as such, is small. Together with the slowness of the weight fluctuations, this leads to considerable difficulties to determine the crosscorrelations experimentally [8].

The analysis of the weight-error correlation can be extended to higher order terms of the Taylor expansion (7), but already the next term $V_3 \mu^3$ is rather complicated, mainly due to the greater number (15 in total) of relevant partial correlations. For the special case of a white Gaussian input signal, this term has been determined in [8], yielding the approximate WECM

$$
\begin{aligned}
V &\approx V_1 \mu + V_2 \mu^2 + V_3 \mu^3 \\
&= \mu N \{[1 + \mu X(M + 2) + \mu^2 X^2 (M + 2)^2] I - 4\mu^2 X^2 T \\
&\quad - 8\mu^2 X^2 S\}
\end{aligned}
$$

$$
T = \begin{pmatrix}
0 & 0 & 1 & 0 & 1 & \cdots \\
0 & 0 & 0 & 1 & 0 & \cdots \\
1 & 0 & 0 & 0 & 1 & \cdots \\
0 & 1 & 0 & 0 & 0 & \\
1 & 0 & 1 & 0 & 0 & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{pmatrix}
$$

$$
S = \begin{pmatrix}
0 & 0 & 0 & 0 & \cdots & & 0 \\
0 & 1 & 0 & 0 & \cdots & & 0 \\
0 & 0 & 2 & 0 & \cdots & & 0 \\
0 & 0 & 0 & 3 & \cdots & & 0 \\
& & & & & \ddots & \\
0 & 0 & 0 & 0 & \cdots & & (M-1)
\end{pmatrix}. \quad (11)
$$

Thus, V is a linear combination of the matrices I, T, S. The first contribution represents a set of uncorrelated, equal-power weight error fluctuations, the second contribution proportional to T describes a set of identical crosscorrelations among the family of even weights and a complementary set among the odd weights, whereas the term proportional to S is readily interpreted as a linear power decrease of the error fluctuations from the front of the delay line toward its end.

It is interesting to compare our result (valid for small stepsizes μ) with that of Douglas et al. [4], [10] (suitable for small filter lengths M). The small region, where μ and M are small, can be handled

by either theory. Without going into detail, we have found a perfect match in that region with respect to the amplitude of the weight fluctuations, particularly its decrease along the tapped-delay line in the third-order approximation.

IV. CONCLUSIONS

This correspondence deals with the correlation matrix V of the weight errors in an LMS-type adaptive TDL filter. Restricting ourselves to a white output noise and avoiding any independence assumption, we have determined the coefficients of a power series $V = V_1\mu + V_2\mu^2 + V_3\mu^3 + \cdots$ in terms of the stepsize μ. The first term $V_1\mu$ is a scalar matrix, representing a set of equal-power, uncorrelated weight fluctuations, in agreement with what is found with the aid of the independence assumption [1]. The quadratic approximation $V_1\mu + V_2\mu^2$ represents a set of weakly correlated equal-power weight fluctuations with a slightly increased common power level. In the third-order approximation, we observe a power decrease along the delay line. This effect can run up to several percent and is more easily observed than the second-order effects [8].

We expect that the proposed iterative method will also lend itself to the treatment of adjacent questions such as adaptation transients and filter tracking. In addition, it might be applicable to other adaptive algorithms like the normalized LMS type. We were able to show that an independence assumption is not required so that teaching adaptive filtering is released from an inconsistent tool [11].

ACKNOWLEDGMENT

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 3rd ed. London, U.K.: Prentice-Hall, 1996.
[2] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[3] V. Solo, "The error variance of LMS with time-varying weights," *IEEE Trans. Signal Processing*, vol. 40, pp. 803–813, 1992.
[4] S. C. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. Signal Processing*, vol. 43, pp. 2863–2871, 1995.
[5] H. J. Butterweck, "A steady-state analysis of the LMS adaptive algorithm without use of the independence assumption," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Detroit, MI, May 9–12, 1995, pp. 1404–1407.
[6] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes With Applications to Stochastic System Theory*. Cambridge, MA: MIT Press, 1984.
[7] H. J. Butterweck, "An approach to LMS adaptive filtering without use of the independence assumption," in *Proc. Eur. Conf. Signal Processing*, G. Ramponi, Ed., Trieste, Italy, 1996, pp. 1223–1226.
[8] ———, "Iterative analysis of the steady-state weight fluctuations in LMS-type adaptive filters," Tech. Rep. 96-E-299, Eindhoven Univ. Technol., Eindhoven, The Netherlands, June 1996.
[9] S. Florian and A. Feuer, "Performance analysis of the LMS algorithm with a tapped delay line (two-dimensional case)," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1542–1549, 1995.
[10] S. C. Douglas and T. H.-Y. Meng, "Exact expectation analysis of the LMS adaptive filter without the independence assumption," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 1992, San Francisco, CA, vol. IV, pp. 61–64.
[11] H. J. Butterweck, "The independence assumption: A dispensable tool in adaptive filter theory," *Signal Process.*, vol. 57, pp. 305–310, 1997.

Sufficient Stability Bounds for Slowly Varying Direct-Form Recursive Linear Filters and Their Applications in Adaptive IIR Filters

Alberto Carini, V. John Mathews, and Giovanni L. Sicuranza

Abstract—This correspondence derives a sufficient time-varying bound on the maximum variation of the coefficients of an exponentially stable time-varying direct-form homogeneous linear recursive filter. The stability bound is less conservative than all previously derived bounds for time-varying IIR systems. The bound is then applied to control the step size of output-error adaptive IIR filters to achieve bounded-input bounded-output (BIBO) stability of the adaptive filter. Experimental results that demonstrate the good stability characteristics of the resulting algorithms are included. This correspondence also contains comparisons with other competing output-error adaptive IIR filters. The results indicate that the stabilized method possesses better convergence behavior than other competing techniques.

Index Terms—Adaptive IIR filter, time-varying recursive linear filter.

I. INTRODUCTION

Adaptive IIR filters have been the subject of active research over the last three decades [5], [9], [11], [12], [15]. Despite a large amount of work that has been done, some open issues still remain. One of these issues is that of ensuring the stability of the time-varying IIR filter that results from the identification process.

Researchers have attempted to derive adaptive IIR filters that operate in a stable manner in several different ways. One class of algorithms is obtained by means of the equation-error technique. In the equation-error technique, the IIR filter is identified by the use of a two-channel adaptive FIR filter that operates on samples of the input and the desired response signals. Since the system model employed in equation-error methods is not recursive, the adaptive filter can operate in a stable manner when the step size is properly selected. However, this fact does not ensure the stability of the resulting IIR filter. Moreover, it is well-known that equation-error adaptive algorithms give biased solutions when the desired response signal is corrupted by noise.

Output error algorithms have become popular in adaptive IIR filtering research in recent years. In output error techniques, the adaptive filter operates in a recursive manner on the input signal to provide an estimate of the desired response signal. A class of such methods requires a certain system transfer function to be strictly positive real (SPR) in order to avoid problems with instability and to ensure the convergence of the algorithm. This class of algorithms includes the pseudo-linear regression algorithm (PRA) [3], which is also known as Feintuch's algorithm, Landau's algorithm [7], the hyperstable adaptive recursive filter (HARF) [4], and the simplified