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

Citation for published version (APA):

Document status and date:
Published: 01/01/1996

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

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

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Iterative Analysis of the Steady-state Weight Fluctuations in LMS-type Adaptive Filters

by H.J. Butterweck

Report EUT 96-E-299
ISBN 90-6144-299-0
June 1996
ITERATIVE ANALYSIS
OF THE STEADY-STATE WEIGHT FLUCTUATIONS
IN LMS-TYPE ADAPTIVE FILTERS

by

H.J. Butterweck

EUT Report 96-E-299
ISBN 90-6144-299-0

Eindhoven
June 1996
ISBN 90-6144-299-0
NUGI 832
Trefw.: adaptieve filters / kleinste-kwadratenmethode / iteratie.
Subject headings: adaptive filters / FIR filters / LMS methods.
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 a fully adapted LMS transversal filter. Without the widely used independence assumption a power series of the weight-error correlation matrix is derived in terms of the adaptation constant. In this series the leading term satisfying a Lyapounov equation has a rather general structure that stands for a set of weight fluctuations with various amplitudes and mutual correlations. For a white reference signal it degenerates into a scalar representing mutually uncorrelated fluctuations with equal amplitudes. A study of the next term of the series then yields small off-diagonal matrix elements and thus weak mutual correlations. If also the input signal is white, the first two terms assume a scalar character and a weak correlation is now predicted by the third term. No correlation is found between even- and odd-numbered weights. Further, a power decrease along the delay line is observed such that the first weight fluctuates noticeably stronger than the last one. Computer simulations provide experimental evidence for the above-mentioned higher-order effects which are not found with the aid of the independence assumption.

Keywords: independence assumption, LMS-algorithm, adaptive filters, misadjustment.

Butterweck, H.J.

Address of the author:
Faculty of Electrical Engineering
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
E-mail: H.J.Butterweck@ele.tue.nl
Contents

1. Introduction 1
2. Basic dynamics and small-signal approximation 3
3. Iterative solution of the equations of motion 7
4. Series expansion of the weight-error correlation matrix 9
5. The zero-order solution for the WECM 12
6. The output and error signal in the zero-order approximation 15
7. The first-order solution for the WECM 17
8. The second-order solution for the WECM 20
9. Simulations 27
10. Conclusions 30
    Appendix 31
    References 32
    Figure Captions 34
    Figures 35
Iterative analysis of the steady-state weight fluctuations in LMS-type adaptive filters

H.J. Butterweck, Eindhoven University of Technology
Faculty of Electrical Engineering, Signal Processing Group

Acknowledgment

The assistance of H.J.W. Belt, H.A. Langeler, and A.W.J. Oomen during the computer simulations is gratefully acknowledged. The author is particularly indebted to W.F.G. Mecklenbräufer for his encouragement during the early stages of the research described in this report.

1. 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 most applications these fluctuations are wanted to be kept small, but this requirement conflicts with the obvious desideratum of a fast adaptation. To a certain extent one of these qualities has to be sacrificed for the other. This trade-off is settled through a proper choice of the "adaptation constant" $\mu$. With increasing $\mu$ the adaptation rate (the inverse of the time constant of adaptation) is raised at the cost of an increasing weight fluctuation.

With a stationary random excitation and under steady-state conditions the weight fluctuations can be characterized by a correlation matrix. Its diagonal terms stand for the powers of the pertinent fluctuations, whereas the off-diagonal entries represent the correlations between pairs of weight fluctuations. In the current literature, the correlation matrix of the weight fluctuations (henceforth denoted by "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 be justified for a "true" vector signal like that emerging from a sensor array. In the case of a tapped-delay line fed by a scalar input signal, as studied in this paper, the input vector is made up of successive values of that signal and, as such, has a more artificial nature. Within an updating cycle all vector components are shifted to the next place with the last component removed and the first component renewed. Clearly, the independence assumption conflicts with such a strong deterministic coherence between successive input vectors. Justified by a lack of competitive methods and encouraged by a fair agreement with measured results, it is, nevertheless, common practice to use the assumption also in such a situation [1-7,19].

Besides its use for the analysis of the weight error correlation matrix the independence assumption has found widespread application to related problems. Among these we note the joint-probability densities under Gaussian excitation [8-9], the normalized least-mean square (NLMS) algorithm [10-11] and nonlinear effects [18]. However, several publications [3,12] and textbooks [5,22] are aware of the dilemma associated with the carefree use of the independence assumption and point out the pertinent didactic difficulties. On the other hand, Gardner [4] states that "in order for such a relatively comprehensive analysis to be tractable, there is one simplifying assumption that cannot be removed" (the independence assumption is meant) thus articulating a general feeling that the independence assumption is indispensable for any analytic approach of the LMS algorithm. This feeling is confirmed by recent work [9,13-18], which avoids the independence assumption but does not yield analytic expressions for the WECM.

In this paper we describe an iterative approach without the independence assumption which finally leads to a power series for the WECM in terms of the adaptation constant. Only the first few terms of the series have a sufficiently simple form so that an efficient use can only be made of the truncated series. This amounts to a confinement to small adaptation constants, far below the value for which the algorithm becomes unstable. This confinement is compensated by a lack of any restrictions concerning the spectral distributions of the exciting signals. No signal needs to be white, and no signal needs to be Gaussian although in case one or more signals are white and/or Gaussian, the final results are considerably simplified.

Although many results admit an elegant frequency-domain interpretation, the analysis is throughout carried out in the time domain. In that domain we distinguish two time scales: one pertaining to the exciting signals and one pertaining to the weight fluctuations. For a small value of the adaptation constant the latter becomes large corresponding to slow fluctuations. This implies that for an experimental verification the averaging times have to be
rather long. In our analysis, all random processes are assumed to be stationary and ergodic implying that (i) the tracking problem is not addressed, (ii) the adaptive transients are not considered, and (iii) the calculated ensemble averages are verified by measuring time averages.

The emphasis is on the WECM, but in due course also the filter's terminal behaviour becomes involved. This implies that besides the scalar input signal and the "reference" signal also the output signal (= inner product of the input vector and the weight vector) is considered. Anticipating the final results we observe major deviations from what is found with the aid of the independence assumption. Particularly if the exciting signals are not white, some surprising insights are gained. As an example we mention the considerable correlation that can occur between the output and the reference signal such that under suitable conditions the power of the "error signal" is lower than that of the reference signal and, consequently, the "misadjustment" can become negative.

In this context we wish to refer to recent work of Solo [13-15] that is closely related to ours and in fact provides the mathematical background to the present, more intuitive approach. Solo uses the white-noise excitation (synonymous with the independence assumption) as a preparatory step towards more general admitted excitations and develops methods similar to ours (Section 2) for their treatment. Also the Lyapounov equation (34) is mentioned, but without drawing further conclusions concerning the character of its solutions and its impact upon the error signal.

2. Basic dynamics and small-signal approximation

With reference to Fig. 1, an adaptive filter with the time-varying weight vector \( \mathbf{w}_n \) (length \( M \)) tries to imitate a fixed filter with weight vector \( \mathbf{h} \) (for sake of convenience also of length \( M \)). The input signal \( x_n \) and the reference signal \( n_n \) are assumed to be sample functions of statistically independent, stationary, zero-mean random processes. If at some instant these random signals are applied to the (stable) system and we have \( \mathbf{w}_n \neq \mathbf{h} \) an adaptation process is initiated which, in global terms, directs \( \mathbf{w}_n \) towards \( \mathbf{h} \). However, \( \mathbf{w}_n \) does not reach \( \mathbf{h} \) asymptotically as a limiting value, but with some random fluctuations superimposed. In this "steady state" we have

\[
\mathbf{w}_n = \mathbf{h} + \mathbf{v}_n,
\]

where, like the excitations \( x_n \) and \( n_n \), also the \( M \)-dimensional \( \mathbf{v}_n \) is a real-valued, stationary, zero-mean random signal. The statistics of the fluctuations \( \mathbf{v}_n \) are the subject of this paper.
Specifically we derive analytic expressions for the $M \times M$ weight-error correlation matrix \( \text{WECM} \)

\[ V = E \{ v_k v'_k \}. \]  

(2)

Notice that the diagonal element \( V_{mm} \) denotes the power of the \( m \)'th weight fluctuation, while the off-diagonal element \( V_{mn} \) stands for the correlation between the fluctuations of the weights \( m \) and \( n \). Like any other correlation matrix, \( V \) is symmetric \((V = V')\) and positive (semi-) definite \((V \succeq 0)\).

For further use we define the $M \times 1$ input vector

\[ \mathbf{x}_k = (x_k, x_{k-1}, \ldots, x_{k-M+1})' \]  

(3)

made up of the scalar input signal and its \((M - 1)\) past values. The output signal is defined as the inner product

\[ y_k = w'_k x_k = h'_k x_k + y'_k x_k, \]  

while the error signal is given by

\[ e_k = n_k + h'_k x_k - y_k = n_k - y'_k x_k. \]  

(5)

Now we discuss the peculiar properties of the adaptive mechanism. These are expressed in the weight updating rule, which for the LMS algorithm reads as

\[ v_{k+1} = v_k + 2\mu e_k x_k = v_k + 2\mu (n_k x_k - x'_k x_k), \]  

(6)

where \( \mu \) is the adaptation constant. Further analysis is eased by making use of the normalized signals \( \sqrt{2\mu} n_k \) and \( \sqrt{2\mu} x_k \), which, for the sake of simplicity, are again denoted by \( n_k \) and \( x_k \), respectively. Formally this amounts to the requirement\(^1\) \( 2\mu = 1 \). The statement "\( \mu \) is small", as often tacitly presupposed in the sequel, is then phrased as "the power of \( x_k \) is small". (The concomitant statement concerning \( n_k \) is of minor importance because of the linear dependence of \( v_k \) on \( n_k \).) After normalization (6) passes into

\(^1\) In fact, we question the need for the use of an adaptation constant, which makes sense only in connection with some normalization of the input power. This remark does not apply to the NLMS algorithm, which cannot go without an adaptation constant.
This relation defines a deterministic operator \((n_k, x_k) \rightarrow (v_k)\) such that \(v_k\) is uniquely determined by the past values of \(n_k\) and \(x_k\). Again we mention the coherence in \(x_k\) due to the fact that only the scalar \(x_k\) can be prescribed, cf. (3).

The system under consideration belongs to a more general class governed by the difference equation

\[
v_{k+1} = v_k - R_k v_k + f_k,
\]

where \(R_k\) is a symmetric, positive (semi-)definite time-varying \(M \times M\) matrix and \(f_k\) is an \(M \times 1\) excitation vector. \(R_k\) and \(f_k\) are assumed to be uncorrelated (not necessarily independent) stationary random signals, with \(f_k\) having zero mean. Indeed, with our special choices \(R_k = x_k x'_k\), \(f_k = n_k x_k\) and the assumed statistical independence of \(n_k\) and \(x_k\), we have

\[
E(f_k) = E(n_k)E(x_k) = 0 \quad \text{and} \quad E(\langle x_{k-i} x_{k-j} \rangle_n n_{k-j}) = E(\langle x_{k-i} x_{k-j} \rangle_n)E(n_{k-j}) = 0 \quad \text{for all} \ j.
\]

Moreover, the dyadic product \(R_k = x_k x'_k\) is positive semidefinite. As discussed above, we are interested in the steady state, where, like \(R_k\) and \(f_k\), also \(v_k\) is a stationary random signal.

Notice that, unlike \(f_k\), the time-varying \(R_k\) has a non-zero mean value (the "input correlation matrix" in our special case)

\[
R = E(R_k) \geq 0
\]

which, like \(R_k\), is positive (semi-)definite. It is important to recognize that, due to \(R_k \geq 0\), the term \((-R_k v_k)\) in (8) represents a time-dependent system damping. Without that term, (8) would represent an unstable system which, in continuous-time terms, behaves as an integrator. The zero-frequency contribution in the power spectrum of \(f_k\) (to be distinguished from the mean value which was assumed to be zero) would give rise to an infinite zero-frequency contribution in \(v_k\).

What occurs when \(R_k\) and herewith \(R\) is small (compared to the unit matrix)? Then the ideal "integrator" is slightly attenuated and the solution \(v_k\) contains very strong (but finite) low-frequency components. In other words, \(v_k\) is a slowly varying function. More precisely, the

\[2\] A vector \(f_k\) and a matrix \(R_k\) are uncorrelated if there is no correlation between any element of \(f_k\) and any element of \(R_k\).
slowly varying part of \( v_k \) becomes more and more dominant for \( R \to 0 \), while the superimposed high-frequency fluctuations remain unaffected by the system damping and maintain the relatively small amplitudes predicted by the ideal integrator.

The slowness of the \( v_k \) variations for \( R \to 0 \) leads to another observation of basic importance [13, Section 7.7]. Let us formulate (8) for \( n \) consecutive time instants and add these equations. Then we obtain a large-scale dynamic equation for the time interval of length \( n \) reading as

\[
v_{k,n} = v_k - \sum_{j=0}^{n-1} R_{k+j} v_{k+j} + \sum_{j=0}^{n-1} f_{k+j}.
\]

(9)

In the first sum we replace the slowly varying \( v_{k+j} \), viewed as a function of \( j \), by the constant \( v_k \) at the beginning of the time interval and obtain

\[
\left( \sum_{j=0}^{n-1} R_{k+j} \right) v_k.
\]

For \( n \) sufficiently large (i.e., for \( R \) sufficiently small) the sum approximates \( n R \), due to the ergodicity of the process \( R_k \). Inserting this approximation into (9) we obtain a large-scale difference equation for \( v_k \) that would also have been obtained if at the very outset the \( R_k \) in (8) was replaced by the constant average \( R \). In other words, the difference equation

\[
\alpha_{k+1} = \alpha_k - R \alpha_k + f_k
\]

(10)

has a solution \( \alpha_k \) that for \( R \to 0 \) approaches \( v_k \) satisfying (8). Thus for \( R \to 0 \) we can solve the simple difference equation (10) with constant coefficients instead of the difficult equation (8) with time-varying coefficients. In our situation where \( R_k = x_k x_k' \) the condition of small \( R \) amounts to a small power of \( x \). In the original notation we then deal with a small adaptation constant \( \mu \).

In the next section the solution \( \alpha_k \) of the simplified difference equation will be the starting point of a convergent iteration, leading to a series expansion for \( v_k \) of which \( \alpha_k \) is the leading term. This fact supports the correctness of the statement \( v_k = \alpha_k \) for \( R \to 0 \). (However, such arguments have an heuristic character. A rigorous mathematical proof requires instruments beyond the scope of this paper).
3. Iterative solution of the equations of motion

We now write

$$v_t = \alpha_k + \beta_k + \gamma_k + \ldots,$$

where the additive corrections $\beta_k$, $\gamma_k$, ... of the simple solution $\alpha_k$ will be determined by iteration. Insertion into (8) reads as

$$\alpha_k, \beta_k, \gamma_k, \ldots = (I - R)(\alpha_k + \beta_k + \gamma_k + \ldots) - P_k(\alpha_k + \beta_k + \gamma_k + \ldots) + f_k,$$

where

$$P_k = R_k - R; \ E(P_k) = 0.$$  \hspace{1cm} (12)

The iterative solution of (12) proceeds as follows:

$$\alpha_{k+1} = (I - R)\alpha_k + f_k,$$

$$\beta_{k+1} = (I - R)\beta_k - P_k\alpha_k,$$

$$\gamma_{k+1} = (I - R)\gamma_k - P_k\beta_k.$$ \hspace{1cm} (14)

The first equation determines $\alpha_k$ from $f_k$, then $\beta_k$ follows from $\alpha_k$, $\gamma_k$ follows from $\beta_k$, etc. Thus we proceed according to $f_k \rightarrow \alpha_k \rightarrow \beta_k \rightarrow \gamma_k \rightarrow \ldots$, where for sufficiently small $R$ the terms in the chain decrease to any wanted degree. Observe that the same operator $\mathcal{S}$ applies in all steps of the above scheme:

$$\alpha_k = \mathcal{S}\{f_k\},$$

$$\beta_k = \mathcal{S}\{-P_k\alpha_k\},$$

$$\gamma_k = \mathcal{S}\{-P_k\beta_k\},$$ \hspace{1cm} (15)

$^3$Use of the finite Greek alphabet for the infinite series (11) does not pose any problem if only the first few terms are actually evaluated.
It represents a simple linear time-invariant filtering of the low-pass type, explicitly governed by the convolutional relation

\[ a_k = \mathcal{G}\{f_k\} = \sum_{j=-\infty}^{\infty} H_j f_{k-j} \]  

(16)

with

\[ H_j = u_{j-1}(I-R)^j = H_j', \quad u_j = \begin{cases} 0 & \text{for } j < 0 \\ 1 & \text{for } j \geq 0 \end{cases}. \]

(17)

Notice that the matrix impulse response \( H_j \) "begins" at \( j = 1 \) so that only past values of \( f_k \) contribute to the present \( \mathcal{G}\{f_k\} \).

The above iteration is attractive in that it transforms the time-varying system parameter \( R_k \) in (8) into a set of excitation functions \( (-P_k \alpha_k), (-P_k \beta_k), \ldots \) serving as source terms in simple constant-coefficient equations of motion. 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 \( R \to 0 \), cf. (17).

We illustrate the low-pass behaviour of \( \mathcal{G}\{\cdot\} \) for the simple case \( M = 1 \) where \( R \) and \( H_j \) become scalars. There we find the frequency-domain system function

\[ \bar{H}(z) = \sum_j H_j z^{-j} = [z-(1-R)]^{-1} \]  

(18)

with a real pole at \( z = 1 - R \), just inside the unit circle. Extremely low frequencies are passed, whereas higher frequencies are attenuated as if the system was an ideal integrator with \( \bar{H}(z) = (z-1)^{-1} \).

With such a behaviour we have to proceed with caution if the input signal of \( \mathcal{G}\{\cdot\} \) has a nonzero mean. Fortunately, this situation does not occur for the input signals listed in (15). For \( f_k \), the zero-mean condition was part of the definition and for \( (-P_k \alpha_k) \) and the following terms it is a consequence of the requirement that \( f_k \) and \( R_k \) are uncorrelated:

\[ E\{-P_k \alpha_k\} = E\{(-R_k + R) \sum_j H_j f_{k-j}\} = \sum_j E\{-R_k H_j f_{k-j} + RH_j f_{k-j}\} = 0, \]

where each term in the sum vanishes. Not only are our input signals free from a d.c. component, but compared to the narrow-band (i.e. slowly varying) output signals they are rather broadband. For \( f_k = n_k \bar{x}_k \) the spectral distribution is determined by that of \( n_k \) and \( \bar{x}_k \) (the \( n_k \) is often assumed to be white!), for \( -P_k \alpha_k \) etc. it follows from that of \( P_k = \bar{x}_k \bar{x}_k' - R \) (note that \( \alpha_k \) is narrow-band around zero).
Due to the extreme low-pass character of $\mathcal{G}(t)$ only the zero-frequency part of the input power spectrum is passed. Since the pole of $\tilde{H}(z)$ is very near the unit circle a substantial amplification effect can be expected. We confirm this conjecture by studying the special case of a one-dimensional system $(M = 1)$ excited by a white signal. Then (16) implies

$$E[\alpha_k^2] = (\sum_j H_j^2)E[f_k^2]$$

$$\sum_j H_j^2 = \frac{1}{1-(1-R^2)} = \frac{1}{2R}$$

(19)

corresponding to an amplification factor $1/(2R)$, which becomes infinite for $R \to 0$ (for other values of $M$ and a more general spectral distribution of the input signal the $R^{-1}$ dependence is maintained with a modified prefactor). Remembering that $R = E\{x^2\}$ we arrive at the following table of power levels:

| Signal | $n_k$ | $f_k$ | $\alpha_k$ | $P_k\alpha_k$ | $\beta_k$ | $P_k\beta_k$ | $\gamma_k$ | ...
|--------|-------|-------|------------|--------------|-----------|--------------|------------|
| Power  | 0(1)  | 0(R)  | 0(1)       | 0(R²)        | 0(R)      | 0(R²)        | 0(R²)      | ...
|        | amplification | amplification | amplification | amplification | ...

Table 1: Scheme of power levels in the different stages of iteration.

Clearly, the inherent amplification in $\mathcal{G}(t)$ is overcompensated by the $P_k$ factors each contributing an $R^2$ attenuation, so that per iteration cycle a net attenuation by a factor $R$ results. This statement establishes the convergence of the iteration procedure.

4. Series expansion of the weight-error correlation matrix

With the expansion (11) we are prepared to determine the WECM in the form

$$V = E\{y'_k\nu'_k\} = E\{(\alpha'_k + \beta'_k + \gamma'_k + \delta'_k + \epsilon'_k + \ldots)$$

$$\cdot (\alpha_k' + \beta_k' + \gamma_k' + \delta_k' + \epsilon_k' + \ldots)\}.$$  (20)

Elaborating (20) we find $V$ as a sum of partial correlations between the various signals $\alpha_k'$, $\beta_k'$, ..., determined by the iteration procedure. In order to find which combinations significantly contribute to $V$ we expand it into a power series in terms of the (scalar) input power

$$P = E\{x_k^2\}$$  (21)
reading as

\[ V = V_0 + V_1 P + V_2 P^2 + \ldots \]  

(22)

Notice that for a vanishing input signal \((P \to 0)\) a nonzero \(V = V_0\) is found. In accordance with Table 1 we then have a small input signal \(f_k = n_k x_k\) but a high amplification factor so that the first filter yields some finite \(a_k\) and thus a finite \(V\).

Further the reference signal \(n_k\) occurs linearly in the difference equation (8) so that \(\nu_k\) including the iteration terms \(\alpha_k, \beta_k, \ldots\) are linearly dependent on \(n_k\). With

\[ N = E \{n_k^2\} \]  

(23)

denoting the power of the reference signal, the matrix \(V\) is then proportional to \(N\) and the same is true for all \(V_i\) \((i = 0, 1, 2, \ldots)\) in the expansion (22). So we can write

\[ V = N(V_0' + V_1' P + V_2' P^2 + \ldots), \]  

(24)

where the primed correlations \(V_0', V_1', V_2' \ldots\) are dependent only on the spectral shapes of \(n_k\) and \(x_k\) and not on their powers. This notation enables us to formulate the dependence of \(V\) on the unnormalized excitations which, after multiplication by \((2\mu)^h\) passed into the normalized quantities. With \(\bar{P}\) and \(\bar{N}\) denoting the unnormalized powers we then have

\[ P = 2\mu \bar{P} \]  

(25)

\[ N = 2\mu \bar{N} \]

so that

\[ V = 2\mu \bar{N}(V_0' + V_1' P + V_2' P^2 + \ldots). \]  

(26)

Thus the expansion (22) in terms of the normalized input power is equivalent to a power series in terms of the adaptation constant. This series begins with a term linearly dependent on \(\mu\) so that \(V \to 0\) for \(\mu \to 0\).

Returning to the normalized quantities and working out (20) we can write each partial contribution to \(V\) as a power series expansion of the form (22). In general, such a series does not begin with the absolute term \(V_0\) but with some power term \(V_b P^b\). The expression in question is then said to be \(O(P^b)\) where \(b\) determines the order of magnitude deciding about the necessity to take the expression into further consideration. As an example let \(V\) be
studied up to the term $V_2P^2$ (no higher-order terms are considered in this paper), then only combinations in (20) with $b \leq 2$ need to be taken into account.

The assignment of $b$ to the various combinations in (20) can be obtained from Table 2. On the diagonal we find $E(\alpha_k\alpha_k') = 0(P^0) = 0(1)$, $E(\beta_k\beta_k') = 0(P)$, $E(\gamma_k\gamma_k') = 0(P^2)$ etc., in agreement with the one-dimensional results of Table 1 (where we had $R = P$). Further we observe that the order of magnitude of the combination $E(\alpha_k\beta_k')$ is surprisingly low (the same as $E(\beta_k\beta_k')$), which has obviously to be ascribed to a low degree of correlation between $\alpha_k$ and $\beta_k$. The same is true for the combinations $E(\alpha_k\delta_k')$ and $E(\gamma_k\beta_k')$. Finally, we see that the number of combinations with a given $b$ grows linearly with $b$ according to $(2b+1)$, where e.g. $E(\alpha_k\beta_k')$ and $E(\beta_k\alpha_k')$ as each other's transpose are counted as one combination.

In following sections we derive the results compiled in Table 2. The most general case ($n_k$ and $x_k$ arbitrarily coloured) will only be considered for $b=0$, cf. Section 5. Case $b=1$ is studied in Section 7 for a white process $n_k$. Finally, in Section 8 the combinations with $b=2$ are investigated under the more stringent restriction that $n_k$ and $x_k$ are white.

Under the last-mentioned assumption the terms $V_0$ and $VP$ are scalar matrices (unit matrices multiplied by scalar constants) corresponding to uncorrelated, equal-energy weight fluctuations. Deviations from this (trivial) behaviour are then represented by the more complicated quadratic term $V_2P^2$. Anticipating the results of Section 8, we find that $V_2P^2$ contains zeros at places where the difference between the row number and the column number is odd. This implies that no correlation exists between even and odd weights.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha_k'$</th>
<th>$\beta_k'$</th>
<th>$\gamma_k'$</th>
<th>$\delta_k'$</th>
<th>$\epsilon_k'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_k$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$\beta_k$</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>•</td>
</tr>
<tr>
<td>$\gamma_k$</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>$\delta_k$</td>
<td>2</td>
<td>2</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>$\epsilon_k$</td>
<td>2</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

Table 2. The order of magnitude $b$ for the various combinations of iterational solutions ($\ast$ means $b>2$).
For the weaker restriction of only a white $n_k$ we find in Section 7 that $V_0$ is again a scalar matrix, while $V_1 P$ is allowed to have nonzero off-diagonal elements. These appear to be proportional to those of $R$ so that $V_1 P$ has a Toeplitz structure. Finally, in the most general case (arbitrary colouring of $n_k$ and $x_k$) to be treated in Section 5, also $V_0$ can have off-diagonal elements arranged symmetrically with respect to both matrix diagonals.

5. The zero-order solution for the WECM

First we examine the zero-order term $V_0$ in the series expansion (22), without any assumption regarding the spectral distributions of $n_k$ and $x_k$. In accordance with Table 1, the term $V_0$ is solely determined by $\alpha_{k_0}$, i.e. by the result of the first iteration cycle. In view of the expansion (20), only the expression $E\{\alpha_{k_0} \alpha_{k_0}'\}$ contributes to $V_0$, but since $E\{\alpha_{k_0} \alpha_{k_0}'\}$ contributes also to $V_1, V_2,$ etc. in (22), we have to write

$$V_0 = \text{leading term of } E\{\alpha_{k_0} \alpha_{k_0}'\} = \lim_{P \to 0} E\{\alpha_{k_0} \alpha_{k_0}'\}. \tag{27}$$

With (16) and the understanding that all summations run from $-\infty$ to $+\infty$ we have

$$E\{\alpha_{k_0} \alpha_{k_0}'\} = E\sum_{i} \sum_{j} H_i f_{i-k_0} f_{i-k_0}^t H_j =$$

$$= \sum_{i} \sum_{j} H_i F^{(i-j)} H_j = \sum_{i} \sum_{j} H_i F^{(i)} H_{i-j} = \sum_{i} T^{(i)}, \tag{28}$$

where

$$F^{(i)} = E\{f_{i-k_0} f_{i-k_0}'\}, \tag{29}$$

$$T^{(i)} = \sum_{i} H_i F^{(i)} H_{i-j}. \tag{30}$$

In an attempt to sum up the series (30) we encounter the difficulty that the matrices $H_j$ and $F^{(i)}$ in general do not commute thus prohibiting the extraction of $F^{(i)}$ from the sum. In fact (30) does not admit an explicit summation. Instead we show that in the limit $P \to 0$ (which is of interest here) $T^{(i)}$ satisfies a Lyapunov equation, for which standard methods of solution exist. To this end we write (17) in the recursive form (where the Dirac function $\delta_i$ equals unity for $i = 0$ and zero elsewhere)

$$H_{i+1} = (I - R) H_i + \delta_i I = H_i (I - R) + \delta_i I. \tag{31}$$
This enables us to rewrite (30) as
\[
T^{(i)} = \sum_i H_i F^{(i)} H_{i+1} = \sum_i H_{i+1} F^{(i)} H_{i+1} \\
= \sum_i [(I-R)H_i + \delta I] F^{(i)} [(I-R)H_{i+1} + \delta_i I] \\
= (I-R) T^{(i)} (I-R) + F^{(i)} (I-R) H_i + (I-R) H_{i+1} F^{(i)} + \delta_i F^{(i)}.
\]

For \( P \to 0 \), i.e. \( R \to 0 \), we can neglect the terms \( RT^{(i)} R \), \( -F^{(i)} RH_i \), \( -RH_i F^{(i)} \) and approximate \( F^{(i)} H_i \) and \( H_{i+1} F^{(i)} \) by \( u_i H \) and \( u_{i+1} F^{(i)} \), respectively. Thus we arrive at the Lyapunov equation
\[
RT^{(i)} + T^{(i)} R = F^{(i)},
\]
valid in the limit \( P \to 0 \). Finally, summing up (33) over all \( I \) and using (28) and (27) we find
\[
RV_0 + V_0 R = \sum_i F^{(i)}.
\]

Here we have the main result of the present section stating that \( V_0 \) satisfies a Lyapunov equation [13, Section 9.6]. Notice that no longer the restriction \( P \to 0 \) need to be explicitly mentioned since this is part of the definition of \( V_0 \), cf. (27).

The right-hand term of (34) is the sum over all autocorrelation matrices of \( \ell_k \) as defined by (29). As such it equals the associated spectral density at zero frequency. Thus the statistical properties of the weight fluctuations are completely determined by the low-frequency part of the exciting signal, which agrees with our previous interpretation of the operator \( \alpha_k = \mathcal{H} \{ \ell_k \} \) as a low-pass filter with an extremely low cut-off frequency.

As yet, we have not used the specific form of the excitation \( \ell_k = n_k x_k \), where \( n_k \) and \( x_k \) are statistically independent. Using the notation
\[
N^{(i)} = E[n_k n_{k+1}] \quad \text{with} \quad N = N^{(0)} \quad \text{and} \\
R^{(i)} = E[x_k x_{k+1}^T] \quad \text{"Toeplitz" with} \quad R = R^{(0)} \quad \text{and} \quad P = P
\]
we obtain from (29)
\[ F^{(0)} = N^{(0)} R^{(0)}. \]  

(36)

Thus the autocorrelation of \( f_k \) equals the product of those of \( n_k \) and \( x_k \). Due to Parseval's theorem the sum over all \( F^{(0)} \), as required in (34), equals the average of the product of the pertinent power spectra. This implies that all frequencies contained in \( n_k \) and \( x_k \) equally contribute to the right-hand term of (34) representing the zero-frequency spectral component of \( f_k \).

There are only few cases, which admit a closed-form solution of (34). One such case occurs for a white process \( n_k \), which is extensively studied in Section 7. Then the right-hand sum of (34) consists of the single term \( N^{(0)} R^{(0)} = NR \), which leads to the well-known solution [20]

\[ V_0 = \frac{1}{2} N I. \]  

(37)

Thus the weight fluctuations are uncorrelated and have equal power \( N/2 \). Notice that the power has a finite limit for \( R \to 0 \), due to the amplifying property of the low-pass filter. Another explicitly solvable case is found for a white \( x_k \). Then \( R = PI \) and \( V_0 \) becomes a Toeplitz matrix with \( (V_0)_{mn} = \frac{1}{2} N^{(m-n)} \).

Returning to the more general case we observe that the correlation matrices \( R, \sum F^{(0)} \) and also \( V_0 \) are symmetric and positive definite. With regard to \( V_0 \), it is a basic property of the Lyapounov equation that it passes on these properties from the given matrices \( (R, \sum F^{(0)}) \) to the unknown \( V_0 \). Another question is how the Toeplitz structure of \( R \) and \( F^{(0)} \) (and that of \( \sum F^{(0)} \)) finds expression in \( V_0 \). Without trying to give a complete answer we can easily prove a specific property of \( V_0 \), viz. a double symmetry of the matrix with respect to both diagonals, as illustrated in Fig. 2.

Geometrically this property is interpreted as a symmetry with respect to the centre of the delay line. The proof is facilitated through introducing the matrix

\[
K = \begin{pmatrix}
0 & 0 & \ldots & 1 \\
0 & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots \\
1 & \ldots & 0 \\
1 & \ldots & 0 & 0
\end{pmatrix}, \quad K^2 = I,
\]  

(38)

where the 1's are positioned on the side-diagonal. Simultaneous pre- and post-multiplication
of a matrix by $K$ mirrors it with respect to that side-diagonal. Matrices like $R$ and $\sum F^{(t)}$ having Toeplitz structures remain unchanged under such a transformation and the same is true for $V_0$ (although in general not of the Toeplitz type) as follows from a pre- and post-multiplication of (34) by $K$:

$$K RV_0 K + KV_0 RK = \sum F^{(t)}. \tag{39}$$

left-hand side $= (KRK) (KV_0 K) + (KV_0 K) (KRK) = R (KV_0 K) + (KV_0 K) R$.

Clearly, $KV_0 K = V_0$, since both matrices satisfy the same Lyapounov equation, q.e.d.

6. The output and error signal in the zero-order approximation

Although our primary interest is focused on the weight fluctuations, some brief consideration of the output signal $y_k$ and the error signal $e_k$ is appropriate for the further understanding of the LMS algorithm. With reference to Fig. 1 we have

$$y_k = \frac{w'}{w} x_k = y_{k,w} + y_{k,f}, \tag{40}$$

where

$$y_{k,w} = h' x_k, \tag{41}$$

$$y_{k,f} = v' x_k. \tag{42}$$

denote the partial output signal due to the Wiener coefficients $h$ (the coefficients of the reference filter) and due to the fluctuations $v$, respectively. Here only the latter part is of real interest, because only this part contributes to the error signal, as follows from

$$e_k = n_k + h' x_k - y_k = n_k - y_{k,f}. \tag{43}$$

In this section we are interested in the mean-square error

$$MSE = E(e_k^2) = E((n_k - y_{k,f})^2) = N + E(y_{k,f}^2) - 2E(n_k y_{k,f}), \tag{44}$$

where $N$ is the power of the reference signal, while the remaining two terms are caused by
the weight fluctuations. The last term reflecting a correlation between the output and the reference signal deserves our particular interest. Only if it vanishes, a superposition of powers applies so that, compared to the basic value $N$, the MSE increases by the output power $E[y_{k,f}^2]$.

This power is now first determined, using the zero-order results:

$$E[y_{k,f}^2] = E[(\alpha_k^* x_k)^2] = E[\alpha_k^* x_k^t \alpha_k^*] = E[\alpha_k^* \alpha_k^t] = E[\alpha_k^* R \alpha_k^t].$$

(45)

Here we have exploited the fact that $x_k$ and $\alpha_k$ fluctuate on extremely different time scales. During some time interval of length $N$ the $\alpha_k$ can be considered as constant, whereas $N$ is sufficient to evaluate the time average with respect to $x_k$ (which equals the ensemble average, due to ergodicity). This leads to a local $E[y_{k,f}^2]$ pertinent to the "frozen" $\alpha_k$. In a second averaging operation with respect to $\alpha_k$ the final $E[y_{k,f}^2]$ is determined. This runs as follows:

$$\frac{\alpha_k^* R \alpha_k}{\alpha_k^*} = tr \{ R \alpha_k \alpha_k^t \} = tr \{ \alpha_k \alpha_k^t R \} = \frac{1}{2} tr(R \alpha_k \alpha_k^t + \alpha_k \alpha_k^t R).$$

With (27) and (34) we find

$$E[y_{k,f}^2] = \frac{1}{2} tr(R V_0 + V_0 R) = \frac{1}{2} \sum_i tr F^{(i)} = \frac{1}{2} \sum_i E[y_{k,f}^2] .$$

(46)

Next we determine the correlation term in (44):

$$-2E[n_k y_{k,f}] = -2E[n_k x_k^t \alpha_k] = -2E[\alpha_k^* x_k] = -2E\sum_i f_i^t H_i f_{k-i}$$

$$= -2\sum_{i=0}^{\infty} E[f_i^t f_{k-i}] = -\sum_{i=0}^{\infty} E[f_i^t f_{k-i}] .$$

(47)

This expression admits the following conclusion. For a white process $n_k$ we have

$E[f_i^t f_{k-i}] = E[n_k n_{k-i}] E[x_k^t x_{k-i}] = 0$ for all $i \neq 0$ so that $y_k$ and $n_k$ are uncorrelated and the exceptional situation of superposable partial powers arises. For more general spectral distributions we combine the expressions (46) and (47) under the name "excess mean-square error" (if normalized with respect to the power of the reference signal it is generally referred to as "misadjustment"):
\[ \text{MSE}_{\text{exc}} = E\{ \sum_i I_i I_i^* \} = \frac{1}{2} \sum_i E\{ \sum_i I_i I_i^* \} \]

\[ = M \left( E\{ n_k^2 \} E\{ x_k^2 \} - \frac{1}{2} \sum_i E\{ n_k^* n_{k'} x_k' \} E\{ x_k x_{k'} \} \right). \]  

(48)

Note that the last expression is formulated in terms of the scalar input signal. The following remarks regard the interpretation of (48):

1. The \( \text{MSE}_{\text{exc}} \) is symmetric with respect to \( n_k \) and \( x_k \); any statement implies another statement with \( n_k \) and \( x_k \) interchanged.

2. \( \text{MSE}_{\text{exc}} \) is proportional to the power \( N \) of the reference signal and can therefore easily be compared with the first term \( N \) in (44). As it is also proportional to the power \( P \) of the input signal it is much smaller than \( N \), because \( P \) has to be kept much smaller than unity to guarantee the validity of the zero-order solution.

3. In terms of the power spectra \( \tilde{N}(\Omega) \) and \( \tilde{P}(\Omega) \) of the reference and (scalar) input signal, respectively, (48) reads as

\[ \text{MSE}_{\text{exc}} = M \left( \tilde{N}(\Omega) \tilde{P}(\Omega) - \frac{1}{2} \tilde{N}(\Omega) \tilde{P}(\Omega) \right), \]  

(49)

where the upper bar denotes averaging over the entire frequency band \( -\pi \leq \Omega \leq \pi \).

4. Contrary to common belief, \( \text{MSE}_{\text{exc}} \) can become negative. This occurs e.g. if both \( n_k \) and \( x_k \) are narrow-band processes of the low-pass type resulting in a strong dominance of the second term in (49). The "normal" situation with \( \text{MSE}_{\text{exc}} > 0 \) occurs e.g. if \( \tilde{N}(\Omega) \) and/or \( \tilde{P}(\Omega) \) are constant (white processes) or if \( \tilde{N}(\Omega) \) and \( \tilde{P}(\Omega) \) do not overlap each other.

7. The first-order solution for the WECM

In this section we extend our study of the series expansion (22) of \( V \) to the first two terms. More specifically, we determine the sum \( (V_0 + V_1) \) and, in order to render the problem manageable, we assume that the reference signal \( n_k \) is white. In accordance with Table 2, the partial correlations \( E\{ a_{k} \alpha_{k} \} \), \( E\{ b_{k} \alpha_{k} \} \), \( E\{ a_{k} ^* \beta_{k} \} \), \( E\{ \gamma_{k} \gamma_{k} ^* \} \), \( E\{ \alpha_{k} \gamma_{k} \} \), and \( E\{ \beta_{k} \beta_{k} ^* \} \) contribute to the above sum in the sense that the first-mentioned term contributes to \( V_0 \) and \( V_1 \) while the remaining terms contribute only to \( V_1 \) (of course, all partial correlations can also contribute to \( V_2 \), \( V_3 \), etc., cf. Section 8).

\[ \text{As an example, } \tilde{N}(\Omega) \propto \cos^2 n \Omega / 2 \text{ and } \tilde{P}(\Omega) \propto \cos^2 n \Omega / 2 \text{ leads to a negative } \text{MSE}_{\text{exc}} \text{ for } n \geq 3. \text{ For } n = 1, 2 \text{ we obtain } \text{MSE}_{\text{exc}} \geq 0. \]
We begin with $E\{\alpha_k\alpha'_j\}$ for which the leading term has already been determined in (37). The complete power series is found with (28), where $T^{(k)} = 0$ for $k \neq 0$, due to the whiteness assumption for $n_k$. Due to the commutativity $R H_i = H_i R$ we obtain

$$E\{\alpha_k\alpha'_j\} = \sum_i H_i F^{(0)} H_i = N R \sum_i H_i^2 =$$

$$= \frac{1}{2} N (I - \frac{1}{2} R)^{-1} = \frac{1}{4} N I + \frac{1}{4} N R + O(P^2).$$

In the last expression the first term represents the contribution to $V_0$, while the second term is the contribution to $V_1 P$.

For the five remaining correlations (four of which occur in pairs of each other's transpose) only the leading terms contributing to $V_1 P$ need to be considered:

$$E\{\beta_k\alpha'_j\} = E \sum_i H_i P_{k+i} \alpha_{k+i-j} \alpha'_j =$$

$$E \sum_i \sum_j \sum_k H_i P_{k+i} H_j f_{k+i-j}^{(k)} H_k =$$

$$= N E \sum_i \sum_j H_i P_{k+i} H_j R_{k+i-j} H_{i,j}$$

$$= N E \sum_i \sum_j H_i P_{k+i} H_j P_{k+i-j} H_{i,j}$$

$$= N \sum_i \sum_j H_i E \{P_{k+i} H_j P_{k+i-j}\} H_{i,j},$$

$$E\{\gamma_k\alpha'_j\} = E \sum_i -H_i P_{k+i} \beta_{k+i} \alpha'_j =$$

$$E \sum_i \sum_j H_i P_{k+i} H_j f_{k+i-j}^{(k)} \alpha_{k+i-j} \alpha'_j =$$

$$= \sum_i \sum_j H_i E \{P_{k+i} H_j P_{k+i-j}\} E \{\alpha_{k+i-j} \alpha'_j\}$$

$$= \frac{1}{2} N \sum_i \sum_j H_i E \{P_{k+i} H_j P_{k+i-j}\} H_{i,j},$$

(52)
\[
E \left\{ \mathbf{\bar{B}}_k \mathbf{\bar{B}}_k' \right\} = E \sum_{i} \sum_{j} H_i P_{k-i} \mathbf{\alpha}_{k-i} \mathbf{\alpha}_{k-j}^t P_{k-j} \mathbf{H}_j \\
= E \sum_{i} \sum_{j} H_i P_{k-i} E \left\{ \mathbf{\alpha}_{k-i} \mathbf{\alpha}_{k-j}^t \right\} P_{k-j} \mathbf{H}_j \\
= E \sum_{i} \sum_{j} H_i P_{k-i} E \left\{ \mathbf{\alpha}_{k-i} \mathbf{\alpha}_{k-j}^t \right\} P_{k-i} \mathbf{H}_{i-j} \\
= \frac{1}{2} N \sum_{i} \sum_{j} H_i E \left\{ P_{i} \left( H_j + H_{-j} + \delta_j \right) P_{k-j} \right\} H_{i-j}, \\
E \left\{ \mathbf{\alpha}_{k} \mathbf{\bar{B}}_k' \right\} = E \left\{ \mathbf{\bar{B}}_k \mathbf{\alpha}_{k}^t \right\} = \frac{1}{2} N \sum_{i} \sum_{j} - H_i E \left\{ P_{k} H_j P_{k-j} \right\} H_{i-j}, \\
E \left\{ \mathbf{\alpha}_{k} \mathbf{\bar{Y}}_k' \right\} = E \left\{ \mathbf{\bar{Y}}_k \mathbf{\alpha}_{k}^t \right\} = \frac{1}{2} N \sum_{i} \sum_{j} H_i E \left\{ P_{k} H_j^2 P_{k-j} \right\} H_{i-j}.
\]

In (52) and (53) we have again exploited the factorization property of the expectation operator when applied to the product of two random signals one of which varies extremely slowly. Furthermore, use has been made of the identity

\[
E \left\{ \mathbf{\alpha}_{k} \mathbf{\alpha}_{k}^t \right\} = E \sum_{i} \sum_{j} H_i E \left\{ P_{k} H_j \mathbf{P}_{k-j} \right\} = \frac{1}{2} N \sum_{i} \sum_{j} H_i R H_{i-j} \\
= \frac{1}{2} N \left( I - R \right)^{11} \left( I - \frac{1}{2} R \right)^{-1} = \frac{1}{2} N \left( I - R \right)^{11} = \frac{1}{2} N \left( H_i + H_{-i} + \delta_i \right).
\]

Notice that the = signs can be throughout interpreted as equality signs, if only the leading terms (= contributions to \( V_i P \)) are considered.

The total contribution of the terms (51) - (55) to \( V_i P \) reduces to

\[
N \sum_{i} \sum_{j} H_i \left[ - E \left\{ P_{k} H_j P_{k-j} \right\} - E \left\{ P_{k} H_j P_{k-j} \right\} + \frac{1}{2} E \left\{ P_{k} H_j P_{k-j} \right\} \\
+ \frac{1}{2} E \left\{ P_{k} H_j P_{k-j} \right\} + \frac{1}{2} E \left\{ P_{k} H_j P_{k-j} \right\} + \frac{1}{2} E \left\{ P_{k} H_j P_{k-j} \right\} \\
+ \frac{1}{2} E \left\{ P_{k} \delta_j P_{k-j} \right\} \right] H_{i-j} = \frac{1}{2} N \sum_{i} \sum_{j} H_i E \left\{ P_{k}^2 \right\} H_i
\]

so that we obtain

\[
V_0 + V_i P = \frac{1}{2} N I + \frac{1}{2} N \left\{ \frac{1}{2} R + \sum_{i} \sum_{j} H_i E \left\{ P_{k}^2 \right\} H_i \right\}.
\]

Further elaboration is possible only with specific assumptions about the joint probability density of \( x_i \). If this is Gaussian, the expectation \( E \left\{ P_{k}^2 \right\} \) can be rewritten as
because $E\{x_k x'_k\} = 2 R^2 + R \text{tr} R$ for a Gaussian signal. With (50) the sum in (57) then passes into

$$\sum_{i} H_i (R^2 + R \text{tr} R) H_i = (R^2 + R \text{tr} R) \sum_{i} H_i^2 = \frac{1}{2} (R + I \text{tr} R).$$

Finally, with $\text{tr} R = MP$ we arrive at

$$V_0 + V_1 P = \frac{1}{2} NI + \frac{1}{2} N (R + \frac{1}{2} I \text{tr} R) = \frac{1}{2} N \{I(1 + \frac{1}{2} MP) + R\}. \quad (59)$$

Thus, in the first-order 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 $R$. Geometrically this implies that all pairs of neighbouring taps have the same correlation and that this statement also holds for all pairs with some given distance on the delay line, cf. Fig. 3.

Moreover, the correlation between two weight fluctuations is proportional to that between the pertinent "tap signals" with a universal proportionality factor $\frac{1}{2} N$ common to all such pairs. It should be observed, however, that the two types of random fluctuations occur on completely different time scales. Contrary to the tap signals the low-frequency weight fluctuations are associated with large time scales which increase with decreasing amplitude of the input signal or, in unnormalized terms, with decreasing adaptation constant.

For the autocorrelations (= powers) of the tap weights we have to modify these statements in the sense that only the equality of the powers is maintained. The above proportionality factor with respect to the tap signal power does not apply here, but has to be replaced by a much larger factor. Hence, the degree of correlation between two weight fluctuations, defined as the ratio of the crosscorrelation and the autocorrelation is extremely small. Together with the slowness of the weight fluctuations this results in considerable difficulties to verify the derived crosscorrelations experimentally with sufficient accuracy (cf. Section 9).

\section*{8. The second-order solution for the WECM}

Now consider the case of a white reference and a white input signal. Then (59) passes into
so that in the first-order theory $V$ remains a scalar matrix, corresponding to uncorrelated, equal-power weight fluctuations. Thus if there are any crosscorrelations these must be due to second- and higher-order effects. As far as second-order effects are concerned, we shall show in the present section that these will again lead to marginal correlations between the various weight fluctuations. The main effects are again of a scalar nature leading to a modification of (60) such that the expression between brackets becomes a second-order polynomial in $P$. Like the correlation predicted by (59) for a coloured input signal the effects are so weak that they are difficult to verify experimentally (cf. Section 9).

The second-order analysis is considerably more complicated than that for the first-order effects; mainly due to the greater number (15 in total) of relevant partial correlations but also because of the greater complexity of the higher-order correlations involved. Therefore, in this section, elaboration of various analytical details has to be left to the reader.

For a white input signal we have

$$R = P I,$$

$$H_i = u_{i-1} X^{-1} I, \xi = 1 - P,$$

satisfying the (useful) identity

$$\sum_i H_i H_{i,i} = \frac{1}{1 - \xi^2} (\xi H_i + \xi H_i + \delta_i).$$

In agreement with (20) the desired correlation $V$ is found as a sum of partial correlations

$$V = \sum_{q=1}^{15} V(q),$$

where $q$ is an unspecified summation index pertinent to all combinations yielding an $O(\xi^0)$, $O(\xi^1)$, and/or $O(\xi^2)$ contribution to $V$, as indicated in Table 2. Each of the $V(q)$ contributions can be expressed in terms of second-, third-, and fourth-order autocorrelations of $P$ as follows:

---

3 The order of the $P_k$ correlation does not always determine the order of the $V(q)$ contribution in a direct manner. The relation between the two orders is discussed further down.
\[ V(q) = \frac{N}{1 - \xi^2} \left[ A(q) + \sum_j B_j(q) E \{ P_k P_{k-j} \} + \sum_j \sum_{t} C_{jt}(q) E \{ P_k P_{k+j} P_{k+j-1} \} \right. \\
\left. \quad + \sum \sum \sum D_{jtm}(q) E \{ P_k P_{k-j} P_{k-j-1} P_{k-j-t-m} \} \right]. \]  

(64)

With \( q = 1 \) denoting the basic autocorrelation \( E \{ \alpha_k \alpha'_k \} \) we find \( A(1) = (1 - \xi) I \) and \( B_j(1) = C_{jt}(1) = D_{jtm}(1) = 0 \), while for all other combinations we have \( A(q) = 0 \) and coefficients \( B \)'s, \( C \)'s, and \( D \)'s as listed in Table 3.

From (63) we conclude that \( V \) can be written in the same form as (64) reading

\[ V = \frac{N}{1 - \xi^2} \left\{ (1 - \xi) I + \sum_j B_j E \{ P_k P_{k-j} \} + \sum_j \sum_{t} C_{jt} E \{ P_k P_{k+j} P_{k+j-1} \} \right. \\
\left. \quad + \sum \sum \sum D_{jtm} E \{ P_k P_{k-j} P_{k-j-1} P_{k-j-t-m} \} \right\}, \]

(65)

where (with \( f_{jtm} = \xi H_{j-t-m} + \xi H_{j-m-t} + \delta_{j-t-m} \))

\[ B_j = \sum_q B_j(q) = -\frac{1 - \xi}{1 + \xi} (\xi H_j^2 + \xi H_j') + \frac{1}{1 + \xi} \delta_j, \]

\[ C_{jt} = \sum_q C_{jt}(q) = \xi \frac{1 - \xi}{1 + \xi} (H_j H_{j,t} + H_{j,t} H_{j,t} + H_j H_{j,t} H_{j,t} + H_j H_{j,t} H_{j,t}) \]

\[ + \frac{1 - \xi}{1 + \xi} H_j^2 \delta_{j,t} - \frac{\xi}{1 + \xi} (H_j^2 \delta_j + H_j^2 \delta_t) \]

\[ = -\frac{\xi}{1 + \xi} (H_j^2 \delta_j + H_j^2 \delta_t) \].
<table>
<thead>
<tr>
<th>$a_k$</th>
<th>$\xi H_j H_{j,l}$</th>
<th>$\xi^2 (1+\xi)^{-1} H_j^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_k$</td>
<td>$-\xi H_j$</td>
<td>$(1+\xi)^{-1} (\xi^2 H_j^2 + \xi H_j + \delta_j)$</td>
</tr>
<tr>
<td>$\gamma_k$</td>
<td>$\xi^2 (1+\xi)^{-1} H_j^2$</td>
<td></td>
</tr>
<tr>
<td>$\delta_k$</td>
<td>$\xi H_j H_{j+l}$</td>
<td></td>
</tr>
<tr>
<td>$\epsilon_k$</td>
<td>$\xi H_j H_{j+l}$</td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>$\xi H_j H_{j,l}$</td>
<td>$\xi^2 (1+\xi)^{-1} H_j H_{j,l}$</td>
</tr>
<tr>
<td>$B_j$</td>
<td>$H_j H_{j,l} (\xi H_j + \xi H_{j,l} + \delta_{j+1})$</td>
<td>$-(1+\xi)^{-1} H_j (\xi H_j + \xi H_{j,l} + \delta_{j+1}) (\xi H_j + \xi H_{j,l} + \delta_{j+1})$</td>
</tr>
<tr>
<td>$C_j$</td>
<td>$-\xi^2 H_j H_{j,l}$</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Coefficients in eq.(64) for the various correlations. Empty places denote zero entries. Note the identity $f_{jlm} = \xi H_{j+1,m} + \xi H_{j,l,m} + \delta_{j+1,m}$. 
Insertion into (65) yields

\[ V = \frac{N}{1-\xi} \left\{ (1-\xi) I + \frac{1}{1+\xi} E \{ P_k^2 \} - \frac{1-\xi}{1+\xi} \sum_j (\xi H_j^2 + \xi H_j^2) E \{ P_k P_{k-j} \} \right\} \]

\[ = \frac{\xi}{1+\xi} \sum_j H_j E \{ P_k P_{k-j} \} \]

\[ + \frac{\xi}{1+\xi} \left[ \sum_j H_j E \{ P_k P_{k-j} \} \right] \]

\[ + \frac{\xi}{1+\xi} \sum_j H_j E \{ P_k P_{k-j} \} \]

\[ + \frac{1}{1+\xi} \sum_j H_j E \{ P_k^2 \} \frac{1}{1+\xi} \sum_j E \left\{ \frac{1}{2} \right\} \]

We now determine the order of the various terms in (67) as a function of \( P \):

\[ \frac{N}{1-\xi} (1-\xi) I = \frac{N}{1+\xi} I = \frac{N}{1+\xi} \left( 1 + P + \frac{P^2}{2} \right) I = 0(P^0), \]

\[ \frac{N}{1-\xi} \frac{1}{1+\xi} E \{ P_k^2 \} = \frac{N}{4P} (1 + P) E \{ P_k^2 \} = 0(P). \]
With the identity $E\{P_{k,j}^2\} = R_2 + R_{1n}R$ for a Gaussian input signal the sum of these two terms passes into the former result (60) in the first-order approximation. The remaining terms in (67) throughout yield second-order contributions, as follows from the following approximations:

$$- \frac{N}{1 - \varepsilon^2} \left[ \frac{1}{1 + \varepsilon} \sum_j (\xi H_j^2 + \xi H_j^3) E\{P_{k,j}^2 \} \right]$$

$$= - \frac{N}{4} \sum_{j=0}^{\infty} E\{P_{k,j}^2 \};$$

$$- \frac{N}{1 - \varepsilon^2} \left[ \frac{\varepsilon}{1 + \varepsilon} \sum_j H_j^2 E\{P_{k,j}^2 + P_{k,j}^2 P_k \} \right]$$

$$= - \frac{N}{4P} \sum_{j=1}^{\infty} E\{P_{k,j}^2 + P_{k,j}^2 P_k \}. $$

Although the individual terms in the remainder of (67) are of order $O(P^3)$ they sum up to an $O(P^3)$ contribution. As an example consider $E\{P_{k,j}^2 P_{k,j}^2 P_{k,j}^2 \}$ that asymptotically decreases to zero for $j \to \infty$ but not for $k \to \infty$ where it assumes the value $E\{P_{k,j}^2 \} E\{P_{k,j}^2 \}$. Convergence of the infinite sum is now guaranteed by the slowly decreasing $H$ functions. This results in

$$\sum_j H_j^2 H_{j+1}^2 E\{P_{k,j}^2 P_{k,j+1}^2 \} = \sum_j H_j^2 E\{P_{k,j}^2 \} E\{P_{k,j+1}^2 \} \sum_j H_j^2 H_{j+1}^2 ;$$

$$\frac{1}{1 - \varepsilon^2} \sum_{j=1}^{\infty} E\{P_{k,j}^2 \} E\{P_{k}^2 \} = \frac{1}{2} \sum_{j=1}^{\infty} E\{P_{k,j}^2 \} E\{P_{k}^2 \}. $$

Finally (67) passes into

$$V \approx \frac{N}{4} \left\{ (2 + P + \frac{1}{2} P^2) I + \left( \frac{1}{P} + 1 \right) E\{P_{k}^2 \} - \sum_{j=0}^{\infty} E\{P_{k,j}^2 \} \right\}$$

$$- \frac{1}{P} \sum_{j=1}^{\infty} E\{P_{k,j}^2 + P_{k,j}^2 P_k \} + \frac{1}{2P^2} \sum_{j=1}^{\infty} E\{P_{k,j}^2 \} E\{P_{k}^2 \}$$

$$+ \frac{1}{2P^2} \sum_{j=1}^{\infty} E\{P_{k,j}^2 \} E\{P_{k,j}^2 \} + \frac{1}{2P^2} \sum_{j=1}^{\infty} E\{P_{k,j}^2 \} E\{P_{k,j}^2 \} \}$$

(68)
The second-order and third-order autocorrelations of \( P_k \) needed in (68) are evaluated in the Appendix. We find

\[
E\{P_kP_{k-j}\} = \delta_j[(M-3)P^2 + Q]I + P^2 I^{(2j)}
\]

valid for all \( j \), while for \( j > 0 \) the following relation applies:

\[
E\{P_kP_{k-j}^2\} = (QP - P^3)E_j + [2QP + (M-4)P^3]I^{(2j)}.
\]

Here we have introduced \( Q = E\{x_i^4\} \) and the two matrices \( E_j \) and \( I^{(2j)} \), whose \( (mn) \)-elements are defined by

\[
(E_j)_{mn} = \delta_{m-n}u_{m-j-1},
\]

\[
I^{(2j)}_{mn} = \delta_{m-n-2j}.
\]

Thus \( E_j \) denotes a unit matrix of which the first \( j \) "1" entries have been replaced by "0" entries, while \( I^{(2j)} \) is a Toeplitz matrix with "1" entries on the \( (2j) \)th parallel to the diagonal and "0" entries elsewhere (for \( j \geq 0 \) the "1" entries are "below" or "above", respectively).

After some elementary manipulations, insertion of (69) and (70) into (68) yields

\[
V = \frac{N}{4}(c_1I - c_2T - 2c_2S),
\]

where

\[
c_1 = 2 + P + \frac{1}{2}P^2 + (M - 2 + \frac{Q}{P^2})(P + MP^2 + \frac{Q}{2}),
\]

\[
c_2 = Q - P^2,
\]

\[
T = \sum_{j \neq 0} I^{(2j)},
\]

\[
S = \sum_{j \neq -1} E^{(2j)}.
\]
Clearly, $V$ is a linear combination of three basic matrices, viz. the unit matrix and the matrices $T$ and $S$ looking as follows:

$$
T = \begin{pmatrix}
0 & 0 & 1 & 0 & 1 & \ldots \\
0 & 0 & 1 & 0 & \ldots \\
1 & 0 & 0 & 0 & 1 & \ldots \\
0 & 1 & 0 & 0 & 0 & \ldots \\
1 & 0 & 1 & 0 & 0 & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}, \quad S = \begin{pmatrix}
0 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 2 & 0 & \ldots & 0 \\
0 & 0 & 0 & 3 & \ldots & 0 \\
0 & 0 & 0 & \ldots & (M-1)
\end{pmatrix}.
$$

(74)

The main contribution to $V$ is the scalar matrix $c_N M / 4$ representing a set of uncorrelated, equal-power weight error fluctuations. The second contribution proportional to $T$ describes a set of identical crosscorrelations between even weights and between odd weights, while the term proportional to $S$ is readily interpreted as a linear energy decrease of the error fluctuations from the front of the delay line towards its end. This typical second-order effect is easier to observe than the crosscorrelations represented by the $T$ matrix and also easier than those determined in the previous section for a coloured input signal. Notice that the second-order theory does not lead to any correlation for a pair formed by an even and an odd weight number.

The above expression for $V$ can be simplified for a Gaussian input signal for which $Q = 3P^2$. We then find the conveniently arranged expression

$$
V = \frac{N}{4} \left\{ \left[ 2 + P(M+2) + \frac{P^2}{2}(M+2)^2 \right] I - 2P^2 T - 4P^2 S \right\},
$$

(75)

whose linear part matches the previous result (60). Notice the energy increase of the weight fluctuations with increasing input power. Moreover, the functional dependence on the combination $P(M+2)$ is noticeable. It implies a strong increase of the weight fluctuations with increasing filter order $M$.

9. Simulations

A new theory involving various approximations calls for experimental support. To this end we have carried out a number of computer simulations. Anticipating their results the details of which are discussed in the present section we find a global agreement within the uncertainties set by inevitable statistical errors.
The easiest experimental situation arises when the input and the reference signal are coloured. Then we deal with various "strong" effects in the sense that they even occur in the limiting case $P \to 0$ yielding $V = V_0$, as determined by (34). Here the agreement between theory and experiment is such that a detailed discussion can be omitted. Also the expected effect that the error can have a lower power than the reference signal (corresponding to a negative misadjustment) is readily observed in the simulations. The only point to remain aware of in performing the simulations is the slowness of the weight fluctuations for low input powers $P$ requiring long observation times.

Unlike this situation the predicted first- and second-order effects are "weak" in the sense that they are hard to observe under normal laboratory conditions. Indeed, the observation of the (majority of the) effects described by (59) and (75) requires carefully prepared experiments. For small $P$ the results are obscured by inevitable statistical uncertainties, whereas for high $P$ values the small-signal assumptions are violated and higher-order effects disturb the results. The only practicable way towards experimental verification is to choose extremely large observation and averaging times and to work with a high arithmetic precision.

The typical problems will be discussed with the aid of characteristic examples. First we study a fifth-order filter ($M = 5$) excited by a high-pass filtered white Gaussian noise. The high-pass filter is of the FIR type with a single transmission zero at $z = 0.5$. The input power, equal to the diagonal entries of the input correlation matrix $R$, is chosen as $P = 3.125 \times 10^2$. Then the entries on the first parallel to the $R$ diagonal have the value $-1.25 \times 10^{-2}$, while all other matrix entries vanish. The reference power $N$ equals unity. With (59) we then determine the WECM in the first-order approximation, with results listed in Table 4.

In Table 4 also the experimental outcomes are listed with averaging times of $94 \times 10^6$ samples. The indicated tolerances have been determined by registrating intermediate readouts every $10^6$ samples (The simulations start with the initial condition $v_0 = 0$. Note that an intermediate readout must not be followed by a reset of the filter state). Interpreting the results we state a fair experimental support of the theoretical results. Anyway, the sign and the order of magnitude of the various crosscorrelations are correctly predicted. The remaining deviations have presumably to be ascribed to the relatively high input level, which tends to increase the diagonal elements and to decrease (in an absolute sense) the off-diagonal elements (experiments with still higher input levels support this hypothesis). Note that a desirable input power reduction would require still longer measuring times to achieve the same relative accuracy of the final results.
Table 4. Theoretical and measured matrix elements of the 5x5 weight-error correlation matrix for an excitation with a high-pass filtered white noise (transmission zero at $z = 0.5$). The unit-power reference signal is white.

As a second example, let us consider a third-order ($M = 3$) filter with a white Gaussian input signal. Here (75) predicts vanishing elements $V_{12}$ and $V_{23}$, while we have $V_{13} = -N/2 \times P^2$. The measured values of $V_{12}$ and $V_{13}$ are listed in Table 5, using $N = 1$ and an averaging time of $50 \times 10^6$ samples.

<table>
<thead>
<tr>
<th>$P$</th>
<th>$10V_{13}$ measured</th>
<th>$10V_{12}$ measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>.045</td>
<td>-1.0 $\pm$ .4</td>
<td>-.26 $\pm$ .3</td>
</tr>
<tr>
<td>.05</td>
<td>-1.2 $\pm$ .4</td>
<td>-.25 $\pm$ .4</td>
</tr>
<tr>
<td>.06</td>
<td>-1.8 $\pm$ .4</td>
<td>-.66 $\pm$ .4</td>
</tr>
<tr>
<td>.07</td>
<td>-2.4 $\pm$ .3</td>
<td>-.67 $\pm$ .4</td>
</tr>
<tr>
<td>.08</td>
<td>-3.2 $\pm$ .3</td>
<td>.00 $\pm$ .4</td>
</tr>
<tr>
<td>.09</td>
<td>-4.0 $\pm$ .4</td>
<td>-.12 $\pm$ .4</td>
</tr>
</tbody>
</table>

Table 5. Crosscorrelations in a third-order filter under white excitation.

For small $P$ we state a fair agreement with the theoretical results, within the statistical uncertainties. For higher $P$, the theory predicts crosscorrelations $V_{13}$ stronger than the measured values, but the statement $V_{12} = 0$ appears to remain valid.

Finally we study a fifth-order filter ($M = 5$) filter excited by a white Gaussian input signal. Now our interest is focused on the power decrease along the delay line, as represented by
the matrix $S$ in (75). Again using $N = 1$ we find a power decrease of $P^2$ per section, following (75). In the experiment we have chosen $P = 0.07$, yielding theoretical power steps of $.0049$. With a total time of $10^8$ samples, we measure the fluctuation powers listed in Table 6.

$$V_{11} \quad V_{22} \quad V_{33} \quad V_{44} \quad V_{55}$$

Table 6. Fluctuation powers in a fifth-order filter under white noise excitation.

The total theoretical power decrease equals $(M-1)P^2 = 0.0196$ and has to be compared with an actual decrease of 0.0154. In the first instance we thus have a fair agreement which, however, can be still improved by refining the analytic result. To this end, we have to go back to (67), yielding a total power decrease $(1 + \xi^2 + \xi^4 + \xi^6) P^2$, which (with $\xi = 1-P = 0.93$) has the value 0.0159 fitting the measured value with surprising accuracy. Clearly the effect under consideration is considerably stronger than the previously discussed crosscorrelations in the first- and second-order solutions. Particularly for large $M$ values, it is easily observed under normal laboratory conditions.

10. Conclusions

This paper deals with the correlation matrix $V$ of the weight-errors in an LMS-type adaptive FIR filter. Viewed as a function of the input power $P$, we have expanded $V$ in a power series $V = V_0 + V_1P + V_2P^2 + \ldots$. The leading term $V_0$ has been investigated for an arbitrary colouring of the input and the reference signal. It satisfies the Lyapounov equation (34) and has a rather general form in the sense that the diagonal terms representing the amplitudes of the weight fluctuations and the off-diagonal terms representing their mutual correlation can widely differ. In fact, only some symmetry conditions have to be met.

For a white reference signal, $V_0$ degenerates into a scalar matrix, representing a set of equal-power, uncorrelated weight fluctuations. Now the interest is focused on the refined approximation $V_0 + V_1P$, as given by (59). This represents a set of weakly correlated equal-power weight fluctuations with a slightly increased common power level. If also the input signal is white, the refined approximation again becomes a scalar and the third term $V_2P^2$ becomes interesting. The result for the sum $V_0 + V_1P + V_2P^2$ is found in (75) and is again characterized by an increase in fluctuation power and weak mutual correlations (except for pairs formed by even- and odd-numbered weights). Another effect of particular significance for higher-order filters ($M > 1$) now comes on, viz. a power decrease along the delay line. This effect can run up to several percent and is easily observed under ordinary laboratory conditions.
Several peculiarities of our approach deserve special consideration. First, the eigenvalues and eigenvectors of $R$ do not explicitly enter the theory which therefore cannot be characterized as a modal analysis. Rather the correlation matrix $R$ occurs directly in the basic results (34) and (59). Further, some discrepancy between the simplicity of the results and the tediousness involved for their derivation might suggest the existence of another, easier way to the same results. Finally, we expect that the proposed iterative method will also lend itself to the treatment of adjacent questions such as adaptation transients and filter tracking. Also it might be applicable to other adaptive algorithms like the normalized LMS type.

The main value of our treatment seems to be of a didactic nature. We were able to show that an independence assumption is not required for the understanding of the LMS algorithm. This way teaching adaptive filtering is released from an inconsistent tool. In a basic course one can confine oneself to the first iteration step yielding $\alpha_0$ and herewith $V_0$, particularly in the simplest form (37) for a white reference signal [21].

APPENDIX

Determination of second- and third-order correlation functions for a white input signal.

1. The matrix $E\{R_k R_{k+j}\} = E\{x_k x'_j x_{k,j} x'_{k,j}\}$, whose $(mn)$-element equals

$$E \sum_p x_{k-m+1} x_{k-p+1} x_{k-j-p+1} x_{k-j+1}$$  \hspace{1cm} (A.1)

Due to the whiteness assumption a nonvanishing contribution to this sum occurs for each $p$ leading to a pairing of equally indexed $x$'s. To start with, consider the case $j\neq 0$. Here the second and third factor in the sum cannot pair so that only the pairings (1,2) (3,4) and (1,3) (2,4) are possible. The first requires $p=m=n$, and the second requires $p=j+n=m-j$. The total contribution to (A1) thus amounts to

$$P^2 (\delta_{m-n} + \delta_{m-n-2j}).$$

For $j=0$ the situation changes radically. The second and third term now pair for all $p$, while the first and fourth term pair for $m=n$. Thus the $M$ contributions to (A1) sum up to $MP^2 \delta_{m-n}$. However, for $p=m=n$ a correction is required, because then the two pairs merge into a quartet involving a fourth-order moment $Q=E\{x^4\}$. Summarizing, with (71) we have

$$E\{R_k R_{k+j}\} = (1-\delta)P^2 (I + I^{(20)}) + \delta_0 [(M-1) P^2 + Q]I,$$

from which (69) can be derived by setting $P_k = R_k - R$. 
2. The matrix \( E\{R_kR_k^2\} = E\{x_kx_k'x_{-k-j}x_{-k-j}'\} \), whose \((mn)\)-element equals

\[
E\sum_{p,q} x_{k,p}\mathbf{x}_{k-q-p+1} x_{k-j+q-p+1} x_{k-j-q+1} x_{k-j+n+1}.
\]  

(A2)

Following (68), this expression needs only to be determined for \( j > 0 \). Again, contributions to (A2) come about through pairing. The fourth and the fifth term always form a pair, while the second and the third term cannot form a pair. Thus only the combinations (1,2) (3,6) (4,5) and (1,3) (2,6) (4,5) have to be considered. The first combination requires \( p = m = n \), whereas the second occurs for \( p = j + n = m - j \) so that \( m - n = 2j \). Both combinations occur for all \( q \), but for certain \( q \)'s the pairs can merge into quartets. For the second combination two quartets can be formed: (1,3,4,5) occurs for \( q = p \), while (2,6,4,5) occurs for \( q = n \). For the first combination the quartet (3,6,4,5) again occurs for \( q = p \), but the other quartet (1,2,4,5) requires \( q = m - j \) which can only be realized for \( m > j \). This condition is reckoned with through introduction of the matrix \( E_j \), as defined by (71):

\[
E\{R_kR_k^2\} = [Q + (M - 1)P^2]\mathbf{P}I + [Q - P^2]\mathbf{P}E_j + [2Q + (M - 2)P^2]\mathbf{P}I^{(2)}.
\]

The ultimately required \( E\{P_kP_k^2\} \) as formulated in (70), is then found by writing \( P_k = R_k - R \) and using the previously found second-order correlation (69).

References


Figure Captions

Fig. 1. Basic adaptive system, in which the adaptive filter $\mathbf{w}$ tries to imitate the fixed filter $\mathbf{h}$.

Fig. 2. Double symmetry of the matrix $V_0$ and corresponding geometrical symmetry with respect to the center of the delay line ($\times$ denote equal matrix entries).

Fig. 3. For a white reference signal and a Gaussian input signal, the correlation between pairs of weight fluctuations is only determined by the tap distance.
Figure 1

Figure 2

Figure 3
(277) Smolders, A.B.
FINITE STACKED MICROSTRIP ARRAYS WITH THICK SUBSTRATES.

(278) Bollen, M.H.J. and M.A. van Houten
ON INSULAR POWER SYSTEMS: Drawing up an inventory of phenomena and research possibilities.

(279) Deursen, A.P.J. van
ELECTROMAGNETIC COMPATIBILITY: Part 5. Installation and mitigation guidelines. Section 3.
cabling and wiring.

(280) Bollen, M.H.J.
LITERATURE SEARCH FOR RELIABILITY DATA OF COMPONENTS IN ELECTRIC DISTRIBUTION NETWORKS.

(281) Weiland, Siep
A BEHAVIORAL APPROACH TO BALANCED REPRESENTATIONS OF DYNAMICAL SYSTEMS.

(282) Gorshkov, Yu.A. and V.I. Vladimirrov
LINE REVERSAL GAS FLOW TEMPERATURE MEASUREMENTS: Evaluations of the optical arrangements for
the instrument.

(283) Creyghton, Y.L.M. and W.R. Rutgers, E.H. van Veldhuizen
IN-SITU INVESTIGATION OF PULSED CORONA DISCHARGE.

(284) Li, H.Q and R.F.P. Smeets
GAP-LENGTH DEPENDENT PHENOMENA OF HIGH-FREQUENCY VACUUM ARCS.

(285) Deursen, A.P.J. and Jochen A.G. Jess
ON THE DEVELOPMENT OF A FAST AND ACCURATE BRIDGING FAULT SIMULATOR.

(286) Falkus, H.M. and A.A.H. Damen
MULTIVARIABLE H-INFINITY CONTROL DESIGN TOOLBOX User manual.

(287) Dong, X.Z. and J.G.J. Slout
THERMAL BUCKLING BEHAVIOUR OF FUSE WIRES.

(288) Bangert, A. van and J.P.M. Voster
CCSTOOL2: An expansion, minimization, and verification tool for finite state
CGS descriptions.

(289) Ruer, Th.G. van de
MODELING OF DOUBLE BARRIER RESONANT TUNNELING DIODES: D.C. and noise model.

(290) Dolmans, G.
ELECTROMAGNETIC FIELDS INSIDE A LARGE ROOM WITH PERFECTLY CONDUCTING WALLS.
(287) Liao, Boshu and P. Massee
RELIABILITY ANALYSIS OF AUXILIARY ELECTRICAL SYSTEMS AND GENERATING UNITS.

(288) Weiland, Siep and Anton A. Steurvogel
OPTIMAL HANKEL NORM IDENTIFICATION OF DYNAMICAL SYSTEMS.

(289) Konieczny, Pawel A. and Lech Józwieck
MINIMAL INPUT SUPPORT PROBLEM AND ALGORITHMS TO SOLVE IT.

(290) Voeten, J.P.M.
POOSL: An object-oriented specification language for the analysis and design of hardware/software systems.

(291) Smeets, B.H.T. and N.H.J. Bollen
STOCHASTIC MODELLING OF PROTECTION SYSTEMS: Comparison of four mathematical techniques

(292) Voeten, J.P.M. and A. van Rangelrooij

(293) Voeten, J.P.M.
SEMANTICS OF POOSL: An object-oriented specification language for the analysis and design of hardware/software systems.

(294) Osch, A.W.H. van
MODELLING OF PHASEODYMUM-DOPED FLUORIDE AND SULFIDE FIBRE AMPLIFIERS FOR THE 1.3 \mu M WAVELENGTH REGION.

(295) Bastiaans, Martin J.

(296) Blaschke, F. and A.J.A. Vandeput
REGELTECHNIEKEN VOOR DRAAIVELMACHINES. (Control of AC machines, in Dutch).

(297) Dolmans, G.
DIVERSITY SYSTEMS FOR MOBILE COMMUNICATION IN A LARGE ROOM.

(298) Mazak, J. and J.L.F. Balseiro
MODELING OF A FLUIDIZED BED REACTOR FOR ETHYLENE POLYMERIZATION.

(299) Butterweck, H.J.
ITERATIVE ANALYSIS OF THE STEADY-STATE WEIGHT FLUCTUATIONS IN LMS-TYPE ADAPTIVE FILTERS.