
Volume 88 Issue 3 March 2008 ISSN 0165-1684
ELSEVIER
SIGNAL
PROCESSING
An International Journal
A publication of the European Association for Signal Processing (EURASIP)
EURASIP

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Signal Processing 88 (2008) 733–748
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

A noise resilient variable step-size LMS algorithm

Márcio Holsbach Costa*, José Carlos Moreira Bermudez

Department of Electrical Engineering, Federal University of Santa Catarina, 88040-900 Florianópolis, SC, Brazil

Received 17 March 2007; received in revised form 17 August 2007; accepted 20 September 2007
Available online 2 October 2007

Abstract

This work presents a modified version of the variable step size (VSS) least mean square (LMS) algorithm originally proposed by Kwong and Johnston [IEEE Trans. Signal Process. 40(7) (July 1992) 1633–1642]. The performance of the new algorithm, called noise resilient variable step size (NRVSS), is less sensitive than VSS to the power of the measurement noise. Its implementation requires only a very small increase in the computational complexity. Analytical models are derived for both NRVSS and VSS algorithms for Gaussian signals and small step-size fluctuations. Simulation results show that the NRVSS algorithm has approximately the same transient behavior as VSS but leads to lower steady-state excess mean-square errors as the signal-to-noise ratio (SNR) decreases. The NRVSS algorithm is specially indicated for adaptive interference reduction in biomedical applications.

Keywords: Adaptive filter; LMS; Variable step size; VSS; Noise suppression.

1. Introduction

Adaptive filtering is now ubiquitous in many signal processing application areas such as system identification, control, and communications. It is required, for instance, in hands-free telephony, hearing aids, biomedicine, audio, and videoconference [1,2]. Among the existing adaptive algorithms, those belonging to the family of least mean square (LMS) algorithms are particularly attractive for low-cost real-time implementations because of their robustness and low computational complexity [3,4].

It is well known that the performance of LMS-based algorithms depends directly on the choice of the step-size parameter. Larger step-sizes speed up convergence at the expense of a larger steady-state misadjustment. Smaller step-sizes tend to improve steady-state performance at the cost of a slower adaptation.

Variable step-size (VSS) strategies are frequently sought after to provide both fast convergence and good steady-state performances [5–14]. In general, the step-size should be large in the early adaptation, and have its value progressively reduced as the algorithm approaches steady-state. The rate at which the step-size is reduced depends on the strategy employed and on the system variables that control such strategy. Different strategies usually lead to distinct performance levels.

Several VSS LMS-type algorithms have been proposed in the literature. Two particularly interesting ones were introduced in [5,9]. The performances of these algorithms are largely insensitive to

*Corresponding author. Tel.: +55 48 3721 9506; fax: +55 48 3721 9280.
E-mail addresses: costa@eel.ufsc.br (M.H. Costa), j.bermudez@ieee.org (J.C.M. Bermudez).

0165-1684/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2007.09.015

the power and to the statistics of the measurement noise, which is a very desirable property. Such robustness, however, comes at the price of a significant increase in computational complexity (number of multiplications) when compared to LMS. The increase over LMS complexity is proportional to the number of adaptive filter coefficients.[1]

To remain attractive for real-time applications, step-size adjustment strategies should require minimal additions to the basic LMS computational cost. Among the several low-cost step-size adjustment criteria, the most promising ones are based on the instantaneous squared error [10], on the error correlation [11], on signal changes of successive gradient estimations [12], and on the correlation between input and error signals [13]. Nevertheless, experimental results show that the steady-state performances provided by these techniques can be highly dependent on the measurement noise power level. This high sensitivity can be explained by a driving term in the mean step-size update equation that is proportional to the noise power. As a result, these algorithms tend to yield poor performance for low signal-to-noise ratios (SNR). A practical example where low SNR occurs is network echo cancellation, which is usually subjected to severe double-talk [2,11]. To mitigate the algorithm's sensitivity, double-talk detectors are used and adaptation is interrupted in low SNR situations. Some algorithms incorporate measurement noise variance estimators to lessen the performance losses [15–17]. However, such strategy is not recommended in applications in which both the desired signal and the noise are always present. Such strategy can be very sensitive to the choice of the power estimator, as the noise signal is not independently accessible.

The VSS algorithm developed by Kwong and Johnston [10] provided an interesting strategy for LMS step-size adjustment. Later on, some authors proposed alternative VSS algorithms which were shown to perform better than VSS [11,18]. More recently, the work in [19] demonstrated that VSS provides the step-size sequence that is the closest to the optimum sequence when properly designed. This result revived the interest in VSS. So far, VSS appears to lead to the best tradeoff between convergence speed and steady-state misadjustment among the low-complexity algorithms, even considering its intrinsic large sensitivity to the noise power.

This work proposes a modified version of the VSS algorithm for applications using real signals that is less sensitive to the measurement noise, at the price of a small increase in computational cost. The new algorithm is called noise resilient variable step size (NRVSS). Since its optimal design requires a good estimation of the reference signal power, this algorithm is specially indicated for applications such as adaptive cancellation of power-line interference in biomedical measurements [20,21].

The paper is organized as follows. Section 2 presents a brief review of the VSS algorithm. Section 3 introduces the NRVSS algorithm. Section 4 provides an analysis of the mean NRVSS step-size behavior and compares VSS and NRVSS performances for white and correlated input signals. Section 5 presents the expression for the excess mean-square error (EMSE) for the NRVSS algorithm under slow adaptation conditions. Section 6 presents a closed formula to predict the algorithm's misadjustment. Section 7 presents simulations using synthetic signals and practical examples with real life signals in biomedical applications. The simulation results corroborate the main properties derived in the theory. Finally, Section 8 presents the main conclusions. Preliminary results of this work were presented in [22].

2. The VSS algorithm

The basic adaptive system block diagram is shown in Fig. 1. Here, n is the discrete time, $x(n)$ is a real input signal with variance r_x, $d(n)$ is the desired signal, $y(n)$ the output of the adaptive filter, $e(n)$ the error signal, and $z(n)$ the measurement noise with variance r_z. $\mathbf{w}(n) = [w_0(n)\ w_1(n)\ \ldots\ w_{N-1}(n)]^{\mathrm{T}}$ is

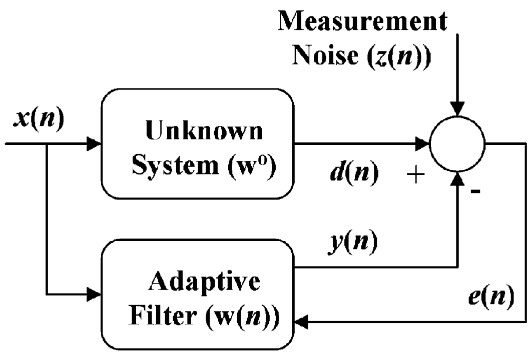

Fig. 1. Adaptive system.

[1]For step-size updating, [5] requires $2N+2$ and [9] requires $3N+3$ multiplications (where N is the number of coefficients).

the adaptive weight vector and $\mathbf{w}^{\mathbf{o}}(n) = [w_0^{\mathrm{o}}(n) w_1^{\mathrm{o}}(n) \ldots w_{N-1}^{\mathrm{o}}(n)]^{\mathrm{T}}$ is the impulse response of the unknown system. In this work $\mathbf{w}^{\mathbf{o}}(n)$ is assumed to be a first order random walk model according to

$$\mathbf{w}^{\mathbf{o}}(n+1) = \mathbf{w}^{\mathbf{o}}(n) + \mathbf{q}(n), \tag{1}$$

where $\mathbf{q}(n)$ is an i.i.d. random perturbation vector, with zero mean and variance r_q and independent of $x(n)$ and $z(n)$. The system's degree of nonstationarity (DN) is given by [23, p. 396]

$$\mathrm{DN} = \sqrt{\mathrm{tr}\{\mathbf{R}\} r_q / r_z}. \tag{2}$$

If $\mathrm{DN} \ll 1$, then the adaptive filter would be able to track variations in the weight vector. Results for the stationary case can be obtained by making $r_q = 0$ and $\mathbf{w}^{\mathbf{o}}(0) = \mathbf{w}^{\mathbf{o}}$ in (1).

The error signal is given by

$$e(n) = z(n) - \mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n), \tag{3}$$

where $\mathbf{x}(n) = [x(n)\ \ x(n-1)\ \ \ldots\ \ x(n-N+1)]^{\mathrm{T}}$ is the input signal vector and $\mathbf{v}(n) = \mathbf{w}(n) - \mathbf{w}^{\mathrm{o}}(n)$ the weight error vector ($\mathbf{v}(n) = [v_0(n)\ v_1(n)\ \ldots\ v_{N-1}(n)]^{\mathrm{T}}$).

The weight vector update equation for a VSS LMS algorithm is [10]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)e(n)\mathbf{x}(n). \tag{4}$$

For the VSS algorithm, the step-size at iteration n is given by

$$\mu(n) = \mu_{\mathrm{VSS}}(n) = \begin{cases} \beta_{\mathrm{MIN}} & \text{if } \beta_{\mathrm{VSS}}(n) < \beta_{\mathrm{MIN}}, \\ \beta_{\mathrm{VSS}}(n) & \text{if } \beta_{\mathrm{MIN}} \leqslant \beta_{\mathrm{VSS}}(n) \leqslant \beta_{\mathrm{MAX}}, \\ \beta_{\mathrm{MAX}} & \text{if } \beta_{\mathrm{VSS}}(n) > \beta_{\mathrm{MAX}}, \end{cases} \tag{5}$$

where $0 < \beta_{\mathrm{MIN}} < \beta_{\mathrm{MAX}}$ are limits set to guarantee tracking capability and stability, respectively, and $\beta_{\mathrm{VSS}}(n)$ is recursively determined from

$$\beta_{\mathrm{VSS}}(n+1) = \alpha_{\mathrm{VSS}}\beta_{\mathrm{VSS}}(n) + \gamma_{\mathrm{VSS}}e^2(n), \tag{6}$$

where α_{VSS} and γ_{VSS} are the control parameters.

2.1. Mean VSS step-size behavior

Assuming that the vector input process $\mathbf{x}(n)$ is zero-mean Gaussian,[2] that $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$ is statistically independent of $\mathbf{v}(n)$ [24], and that $\beta_{\mathrm{VSS}}(n)$ is independent of $\mathbf{x}(n)$ and $z(n)$,[3] the mean behavior of $\beta_{\mathrm{VSS}}(n)$ can be approximated by [10]

$$E\{\beta_{\mathrm{VSS}}(n+1)\} = \alpha_{\mathrm{VSS}}E\{\beta_{\mathrm{VSS}}(n)\} + \gamma_{\mathrm{VSS}}\,\mathrm{tr}\{\,\mathbf{R}\mathbf{K}_{\mathrm{VSS}}(n)\} + \gamma_{\mathrm{VSS}}r_z, \tag{7}$$

where $E\{\cdot\}$ means statistical expectation. $\mathbf{K}_{\mathrm{VSS}}(n) = E\{\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\}$ and $\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\}$ are the weight error vector and the input signal vector correlation matrices, respectively. This analysis also assumes that $\beta_{\mathrm{VSS}}(n)$ remains naturally bounded in $[\beta_{\mathrm{MIN}}, \beta_{\mathrm{MAX}}]$.

Note that the measurement noise influences the VSS step-size behavior through the two rightmost terms in (7). The term $\mathrm{tr}\{\mathbf{R}\mathbf{K}_{\mathrm{VSS}}(n)\}$ is determined by the steady-state misadjustment, which is usually designed to be small. Thus, this term is not the main performance degradation factor. The term $\gamma_{\mathrm{VSS}}r_z$, however, is proportional to the measurement noise power, which is independent of the adaptation process. It adds a bias equal to $\gamma_{\mathrm{VSS}}r_z/(1-\alpha_{\mathrm{VSS}})$ to the steady-state value of $E\{\beta_{\mathrm{VSS}}(n)\}$ for the white input signal case. This bias limits the step-size reduction, and thus the minimum achievable steady-state misadjustment.

3. NRVSS update equation

The steady-state bias of the mean step-size in (7) can be reduced if (5) is modified to

$$\mu(n) = \begin{cases} \beta_{\mathrm{MIN}} & \text{if } |\beta(n)| < \beta_{\mathrm{MIN}}, \\ |\beta(n)| & \text{if } \beta_{\mathrm{MIN}} \leqslant |\beta(n)| \leqslant \beta_{\mathrm{MAX}}, \\ \beta_{\mathrm{MAX}} & \text{if } |\beta(n)| > \beta_{\mathrm{MAX}}, \end{cases} \tag{8}$$

with $\beta(n)$ updated according to

$$\beta(n+1) = \alpha\beta(n) + \gamma[k\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n) - 1]e^2(n), \tag{9}$$

where k, α, and γ are the control parameters and $0 < \beta_{\mathrm{MIN}} < \beta_{\mathrm{MAX}}$. Note that taking the magnitude of $\beta(n)$ in (8) does not involve mathematical operations, and avoids problems that might accrue from negative values obtained from (9). The effect of these modifications on the algorithm behavior will become apparent after the analysis in the next section.

Comparing the computational complexities of (6) and (9), the latter requires only three extra multiplications per iteration, assuming that $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$ is evaluated recursively. For normalized algorithms

[2]The Gaussian input signal assumption has been widely used in adaptive filter analysis. The results obtained with Gaussian signal models are suggestive of the fundamental analysis and design issues, while this assumption keeps the mathematical analysis problem tractable.

[3]From (6), this assumption is approximately true for small values of γ_{VSS} [10].

such as NLMS (which already requires the evaluation of $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$), the computational cost increases only by two multiplications and one addition.

The NRVSS algorithm described by (8) and (9) could be considered a form of noise constrained VSS algorithm as defined in [15]. The VSS algorithm is a particular case of the NRVSS when $k = 0$, $\gamma = -\gamma_{\mathrm{VSS}}$, and $\alpha = \alpha_{\mathrm{VSS}}$. Another case for which RVSS degenerates into VSS is when the input signal is a noiseless constant modulus signal. In this particular case, $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$ is constant and (9) becomes equivalent to (6). Such input signals occur theoretically in some communication systems. However, in most practical situations the input signal will be contaminated by some sort of noise and $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$ will be random.

4. Mean NRVSS step-size behavior

In designing the NRVSS algorithm (9), the parameter k can be used to compensate for the influence of the measurement noise in the mean step-size behavior whenever the input signal $x(n)$ is not a noiseless constant modulus signal. Taking the expectation of (9) we obtain

$$E\{\beta(n+1)\} = \alpha E\{\beta(n)\} - \gamma E\{e^2(n)\} + k\gamma E\{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)e^2(n)\}, \tag{10}$$

where two expected values must be evaluated. Using the same statistical assumptions used in Section 2.1, the term $E\{e^2(n)\}$ is given by [10]

$$E\{e^2(n)\} = \mathrm{tr}\{\mathbf{RK}(n)\} + r_z, \tag{11}$$

where $\mathbf{K}(n) = E\{\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\}$ for $\mathbf{v}(n)$ updated using the NRVSS algorithm.

Squaring (3), multiplying the result by $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$ and neglecting the statistical dependence of $\mathbf{v}(n)$ and $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$, the last expected value in (10) can be evaluated using the Gaussian moment factoring theorem [1], resulting in

$$E\{e^2(n)\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)\} = 2\mathrm{tr}\{\mathbf{RRK}(n)\} + r_x N\,\mathrm{tr}\{\mathbf{RK}(n)\} + r_x N r_z. \tag{12}$$

Substituting (11) and (12) in (10) we obtain

$$\begin{aligned}E\{\beta(n+1)\} &= \alpha E\{\beta(n)\} + 2k\gamma\,\mathrm{tr}\{\mathbf{RRK}(n)\} + \gamma(kr_xN-1)E\{e^2(n)\}\\ &= \alpha E\{\beta(n)\} + 2k\gamma\,\mathrm{tr}\{\mathbf{RRK}(n)\} + \gamma(kr_xN-1)\mathrm{tr}\{\mathbf{RK}(n)\} + \gamma(kr_xN-1)r_z.\end{aligned} \tag{13}$$

4.1. Compensation of the noise influence

Examination of (13) shows that the last term on the right is the main source of noise power interference in the mean step-size behavior. Differently from VSS, this effect can be minimized for NRVSS through the appropriate choice of the free control parameter k. Using

$$k = 1/(r_x N) \tag{14}$$

in (13) we obtain

$$E\{\beta(n+1)\} = \alpha E\{\beta(n)\} + \frac{2\gamma}{r_x N}\mathrm{tr}\{\mathbf{RRK}(n)\}. \tag{15}$$

For white input signals, $\mathbf{R} = r_x\mathbf{I}$ with $\mathbf{I}$ the identity matrix, Eq. (15) simplifies to

$$E\{\beta(n+1)\} = \alpha E\{\beta(n)\} + \frac{2\gamma r_x}{N}\mathrm{tr}\{\mathbf{K}(n)\}. \tag{16}$$

In applications with nonstationary input signals, the parameter k can be periodically updated using estimations of r_x in time windows small enough for quasi-stationarity (see Example 6).

Eqs. (15) and (16) show that the proposed strategy is able to significantly reduce the direct influence of the measurement noise power on the mean behavior of the NRVSS step-size. The steady-state value of $E\{\beta(n)\}$ from (15) is now free from the noise-driven bias. This allows $E\{\beta(n)\}$ to be reduced to lower steady-state levels than in (7).

4.2. Comparison between VSS and NRVSS

In order to provide a meaningful comparison between the VSS and NRVSS steady-state performances, the following framework is assumed:

- white input signal;
- $\beta_{\mathrm{VSS}}(n)$ and $\beta(n)$ remain naturally bounded in $[\beta_{\mathrm{MIN}}, \beta_{\mathrm{MAX}}]$;
- $\alpha = \alpha_{\mathrm{VSS}}$ and $\gamma = N\gamma_{\mathrm{VSS}}/2$;
- the mean step-size converges to a constant steady-state value.

Under these assumptions, (7) and (16) turn, respectively, to

$$\begin{cases}\lim_{n\to\infty} E\{\beta_{\mathrm{VSS}}(n)\} = \dfrac{\gamma_{\mathrm{VSS}}}{(1-\alpha_{\mathrm{VSS}})} \times (r_x \lim_{n\to\infty}\mathrm{tr}\{\mathbf{K}_{\mathrm{VSS}}(n)\} + r_z),\\ \lim_{n\to\infty} E\{\beta(n)\} = \dfrac{\gamma_{\mathrm{VSS}}}{(1-\alpha_{\mathrm{VSS}})} r_x \lim_{n\to\infty}\mathrm{tr}\{\mathbf{K}(n)\}.\end{cases} \tag{17}$$

Comparison of Eqs. (17) (VSS and NRVSS, respectively) shows that the steady-state mean VSS step-size is directly proportional ($\gamma_{\rm VSS}/(1-\alpha_{\rm VSS})$) to the steady-state mean-square error ($\text{MSE}_{\rm SS} = r_x \lim_{n\to\infty} \text{tr}\{\mathbf{K}_{\rm VSS}(n)\} + r_z$) while in NRVSS, it is proportional (with the same factor of VSS) to the steady-state excess mean-square error ($\text{EMSE}_{\rm SS} = r_x \lim_{n\to\infty} \text{tr}\{\mathbf{K}(n)\}$)[4]. It is important to note here that the second order moments of both algorithms are different. However, for practical applications $\lim_{n\to\infty} \text{tr}\{\mathbf{K}_{\rm VSS}(n)\} \ll r_z$ and $\lim_{n\to\infty} \text{tr}\{\mathbf{K}(n)\} \ll r_z$. In this way, (17) demonstrates the NRVSS smaller steady-state sensitivity to the measurement noise power when compared to the VSS.

5. NRVSS EMSE

The EMSE of the NRVSS algorithm can be estimated from

$$\text{EMSE}(n) = E\{e^2(n)\} - r_z = \text{tr}\{\mathbf{RK}(n)\}. \tag{18}$$

The evolution of the second order moments ($\mathbf{K}(n)$) can be obtained by subtracting (1) from (4), using (3), post-multiplying the result by its transpose, and taking its expectation in the same way as in [10], resulting in

$$\begin{aligned} E\{\mathbf{v}(n+1)\mathbf{v}^{\rm T}(n+1)\} = {} & E\{\mathbf{v}(n)\mathbf{v}^{\rm T}(n)\} \\ & + E\{\mu^2(n)z^2(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\} \\ & - E\{\mu(n)\mathbf{v}(n)\mathbf{v}^{\rm T}(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\} \\ & - E\{\mu(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\mathbf{v}(n)\mathbf{v}^{\rm T}(n)\} \\ & + E\{\mu^2(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\mathbf{v}(n)\mathbf{v}^{\rm T}(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\} \\ & + E\{\mathbf{q}(n)\mathbf{q}^{\rm T}(n)\}. \end{aligned} \tag{19}$$

To proceed with the analysis we assume that the adaptive step-size $\mu(n)$ is statistically independent of $\mathbf{v}(n)$ and $\mathbf{x}(n)\mathbf{x}^{\rm T}(n)$ [10,24]. This assumption is approximately valid if the term multiplying $e^2(n)$ in (9) is very small. Considering that k is given by (14), this term will be very small for a large number of taps and ergodic input signal $x(n)$. Then, the recursive equation for $\mathbf{K}(n)$ can be approximated by

$$\begin{aligned} \mathbf{K}(n+1) = {} & \mathbf{K}(n) - E\{\mu(n)\}[\mathbf{K}(n)\mathbf{R} + \mathbf{RK}(n)] \\ & + 2E\{\mu^2(n)\}\mathbf{RK}(n)\mathbf{R} \\ & + E\{\mu^2(n)\}[\text{tr}\{\mathbf{RK}(n)\} + r_z]\mathbf{R} + \mathbf{R_q}, \end{aligned} \tag{20}$$

[4]The EMSE represents the level above the minimum mean-square error (MSE) that would be obtained if the filter tap weights were fixed at their optimal values.

which is a function of the first two moments of $\mu(n)$. In the following, we derive approximate expressions for these moments.

5.1. Mean behavior of $\mu(n)$

Assuming small fluctuations of $\mu(n)$ about its mean and that $\beta_{\rm MIN} \leqslant \beta(n) \leqslant \beta_{\rm MAX}$, we approximate $E\{\mu(n)\}$ by

$$E\{\mu(n)\} = E\{|\beta(n)|\} \cong \sqrt{E\{\beta^2(n)\}}. \tag{21}$$

This approximation is more accurate during the transient phase of adaptation, when $E\{\beta(n)\}^2 \gg \text{Var}\{\beta(n)\}$, where $\text{Var}\{\cdot\}$ stands for variance.

5.2. Second order moment of $\mu(n)$

Squaring (9), taking its expected value using the same assumption that led to (20), and applying (11) and (12) we obtain

$$\begin{aligned} E\{\beta^2(n+1)\} = {} & \alpha^2 E\{\beta^2(n)\} + 4\alpha k\gamma E\{\beta(n)\}\text{tr}\{\mathbf{RRK}(n)\} \\ & + 2\alpha\gamma E\{\beta(n)\}(kNr_x - 1)E\{e^2(n)\} \\ & + k^2\gamma^2 E\{e^4(n)\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\} \\ & - 2k\gamma^2 E\{e^4(n)\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\} \\ & + \gamma^2 E\{e^4(n)\}. \end{aligned} \tag{22}$$

The last term in (22) is calculated by raising (3) to the fourth power, ignoring the variance of the weights, and applying the property of the moments of a Gaussian variable [25]. This procedure results in

$$E\{e^4(n)\} \cong 3r_z^2 + 6r_z\,\text{tr}\{\mathbf{RK}(n)\} + 3\text{tr}\{\mathbf{RK}(n)\}^2. \tag{23}$$

The two remaining expected values in (22) can be simplified using the Gaussian Moment Factoring Theorem [1], resulting in

$$\begin{aligned} E\{e^4(n)\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\} = {} & 3Nr_xr_z^2 + 12r_z\,\text{tr}\{\mathbf{RRK}(n)\} \\ & + 6Nr_xr_z\,\text{tr}\{\mathbf{RK}(n)\} \\ & + E\{(\mathbf{v}^{\rm T}(n)\mathbf{x}(n))^4\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\} \end{aligned} \tag{24}$$

and

$$\begin{aligned} & E\{e^4(n)\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\mathbf{x}^{\rm T}(n)\mathbf{x}(n)\} \\ & \quad = 3r_z^2\left(N^2r_x^2 + 2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} r_{j-i}^2\right) \\ & \qquad + 6r_zE\{(\mathbf{v}^{\rm T}(n)\mathbf{x}(n))^2(\mathbf{x}^{\rm T}(n)\mathbf{x}(n))^2\} \\ & \qquad + E\{(\mathbf{v}^{\rm T}(n)\mathbf{x}(n))^4(\mathbf{x}^{\rm T}(n)\mathbf{x}(n))^2\}. \end{aligned} \tag{25}$$

The evaluations of the three expectations in (24) and (25) are detailed in Appendices A and B. Here only the final results are presented for conciseness:

$$E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^4\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)\} = 3Nr_x\,\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RK}(n)\} + 12\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RRK}(n)\}, \tag{26}$$

$$E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^2(\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n))^2\} = 8\mathrm{tr}\{\mathbf{RRRK}(n)\} + 4Nr_x\,\mathrm{tr}\{\mathbf{RRK}(n)\} + \left(N^2r_x^2 + 2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right)\mathrm{tr}\{\mathbf{RK}(n)\}, \tag{27}$$

and

$$E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^4(\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n))^2\} = 24\mathrm{tr}\{\mathbf{RRK}(n)\mathbf{RRK}(n)\} + 48\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RRRK}(n)\} + 24r_xN\,\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RRK}(n)\} + 3\left(N^2r_x^2 + 2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right)\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RK}(n)\}. \tag{28}$$

Substituting (23)–(28) in (22) leads to

$$\begin{aligned}E\{\beta^2(n+1)\} &= \alpha^2E\{\beta^2(n)\} + 4\alpha k\gamma E\{\beta(n)\}\mathrm{tr}\{\mathbf{RRK}(n)\} \\ &+ 6\gamma^2k^2r_z(r_z + 2\mathrm{tr}\{\mathbf{RK}(n)\})\left(\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right) \\ &+ 3\gamma^2\mathrm{tr}\{\mathbf{RK}(n)\}^2 + 48\gamma^2k^2r_z\mathrm{tr}\{\mathbf{RRRK}(n)\} \\ &+ 3\gamma^2k\left(2k\left(\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right) - Nr_x\right)\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RK}(n)\} \\ &+ 24\gamma^2k^2\mathrm{tr}\{\mathbf{RRK}(n)\mathbf{RRK}(n)\} \\ &+ 48\gamma^2k^2\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RRRK}(n)\} \\ &+ 2\alpha\gamma E\{\beta(n)\}E\{\mathrm{e}^2(n)\}(kNr_x - 1) \\ &+ 3\gamma^2kNr_x\,\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RK}(n)\}(kNr_x - 1) \\ &+ 24\gamma^2k(\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RRK}(n)\} \\ &+ r_z\,\mathrm{tr}\{\mathbf{RRK}(n)\})(kNr_x - 1) \\ &+ 3\gamma^2r_z(r_z + 2\mathrm{tr}\{\mathbf{RK}(n)\})(1 - 2kNr_x \\ &+ k^2N^2r_x^2).\end{aligned} \tag{29}$$

Making $\alpha = \alpha_{\mathrm{VSS}}$, $\gamma = -\gamma_{\mathrm{VSS}}$, and $k = 0$, (29) becomes the recursion for $E\{\beta_{\mathrm{VSS}}^2(n)\}$.

$$\begin{aligned}E\{\beta_{\mathrm{VSS}}^2(n+1)\} &= \alpha_{\mathrm{VSS}}^2E\{\beta_{\mathrm{VSS}}^2(n)\} + 3\gamma_{\mathrm{VSS}}^2E\{e^2(n)\}^2 \\ &+ 2\alpha_{\mathrm{VSS}}\gamma_{\mathrm{VSS}}E\{\beta_{\mathrm{VSS}}(n)\}E\{e^2(n)\},\end{aligned} \tag{30}$$

where $E\{e^2(n)\}$ is given by (11). Eqs. (7), (18), (20), and (30) constitute a theoretical model to the VSS algorithm [10]. Using (14) in (29) it comes to

$$\begin{aligned}E\{\beta^2(n+1)\} &= \alpha^2E\{\beta^2(n)\} + \frac{6\gamma^2r_z^2}{N^2r_x^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2 \\ &+ 3\gamma^2\left(\mathrm{tr}\{\mathbf{RK}(n)\} + \frac{4r_z}{N^2r_x^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right)\mathrm{tr}\{\mathbf{RK}(n)\} \\ &+ \frac{4\alpha\gamma}{Nr_x}E\{\beta(n)\}\mathrm{tr}\{\mathbf{RRK}(n)\} \\ &+ \frac{48\gamma^2r_z}{N^2r_x^2}\mathrm{tr}\{\mathbf{RRRK}(n)\} \\ &+ 3\gamma^2\left(\frac{2}{N^2r_x^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2 - 1\right)\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RK}(n)\} \\ &+ \frac{24\gamma^2}{N^2r_x^2}\mathrm{tr}\{\mathbf{RRK}(n)\mathbf{RRK}(n)\} \\ &+ \frac{48\gamma^2}{N^2r_x^2}\mathrm{tr}\{\mathbf{RK}(n)\mathbf{RRRK}(n)\}.\end{aligned} \tag{31}$$

Eq. (31) models the behavior of the second moment of the VSS for the NRVSS algorithm. For white input signals (31) becomes

$$\begin{aligned}E\{\beta^2(n+1)\} &= \alpha^2E\{\beta^2(n)\} + \frac{6\gamma^2r_z^2}{N} + 3\gamma^2r_x^2\mathrm{tr}\{\mathbf{K}(n)\}^2 \\ &+ \frac{4\gamma r_x}{N}\left(\alpha E\{\beta(n)\} + 3\gamma r_z\left(1 + \frac{4}{N}\right)\right)\mathrm{tr}\{\mathbf{K}(n)\} \\ &- 3\gamma^2r_x^2\left(1 - \frac{2}{N} - \frac{24}{N^2}\right)\mathrm{tr}\{\mathbf{K}(n)\mathbf{K}(n)\}.\end{aligned} \tag{32}$$

Eqs. (18), (20), (21), and (31) allow predictions about the EMSE behavior for the NRVSS algorithm.

6. Steady-state misadjustment

Assuming white Gaussian input signals in a stationary environment and algorithm convergence,

(20) results in

$$\lim_{n\to\infty} T_K(n) = r_z \Bigg/ \left(\frac{2}{N} \frac{\lim_{n\to\infty} E\{\mu(n)\}}{\lim_{n\to\infty} E\{\mu^2(n)\}} - r_x \frac{N+2}{N} \right), \quad (33)$$

where $T_K(n) = \text{tr}\{\mathbf{K}(n)\}$ and $\text{tr}\{\cdot\}$ is the trace operator. Using (8) and (21) in (33) we obtain

$$\lim_{n\to\infty} T_K(n) = r_z \Bigg/ \left(\frac{2}{N\sqrt{\lim_{n\to\infty} E\{\beta^2(n)\}}} - r_x \frac{(N+2)}{N} \right). \quad (34)$$

Assuming convergence in (32) ($E\{\beta^2(n+1)\} \cong E\{\beta^2(n)\}$) we obtain

$$\begin{aligned}\lim_{n\to\infty} E\{\beta^2(n)\} = {} & \frac{6\gamma^2}{N(1-\alpha^2)} [r_z + r_x \lim_{n\to\infty} T_K(n)]^2 \\ & + \frac{48\gamma^2 r_x r_z}{N^2(1-\alpha^2)} \lim_{n\to\infty} T_K(n) \\ & + \frac{72\gamma^2 r_x^2}{N^2(1-\alpha^2)} \lim_{n\to\infty} T_K^2(n) \\ & + \frac{8\alpha\gamma^2 r_x^2}{N^2(1-\alpha^2)(1-\alpha)} \lim_{n\to\infty} T_K^2(n) \\ & + \frac{3\gamma^2 r_x^2}{(1-\alpha^2)} \left[1 - \frac{2}{N} - \frac{24}{N^2}\right] \\ & \times \sum_{i=0}^{N-1} \sum_{j\neq 1}^{N-1} \lim_{n\to\infty} k_i(n) \lim_{n\to\infty} k_j(n), \end{aligned} \quad (35)$$

where $k_i(n)$ ($i = 0,1, \ldots, N-1$) are the main diagonal elements of $\mathbf{K}(n)$. Assuming the usual practical situation of $\text{EMSE}_{\text{SS}} = r_x \lim_{n\to\infty} \text{tr}\{\mathbf{K}(n)\} \ll r_z$ (steady-state EMSE) and using (35) in (34) we obtain

$$\text{EMSE}_{\text{SS}}^2 - \frac{N}{N+2} \left[\frac{1}{\gamma r_x} \sqrt{\frac{2(1-\alpha^2)}{3N}} + 2r_z \left(\frac{N+1}{N}\right) \right] \times \text{EMSE}_{\text{SS}} + \frac{N}{N+2} r_z^2 = 0. \quad (36)$$

Solving (36) we finally obtain

$$\text{EMSE}_{\text{SS}} = \frac{N}{2(N+2)} \left(v - \sqrt{v^2 - 4\frac{N+2}{N} r_z^2} \right), \quad (37)$$

where

$$v = \frac{1}{\gamma r_x} \sqrt{\frac{2(1-\alpha^2)}{3N}} + 2r_z \left(\frac{N+1}{N}\right). \quad (38)$$

The misadjustment can be theoretically predicted by $M = \text{EMSE}_{\text{SS}}/r_z$.

7. Simulation results

This section presents six representative examples to illustrate the properties of the NRVSS algorithm. All examples compare the NRVSS and VSS performances to demonstrate that NRVSS can lead to a better steady-state performance for the same initial transient behavior. Such methodology uses the VSS as a golden standard, since it has been demonstrated in [19] that VSS can produce a close-to-optimal step-size sequence, outperforming algorithms such as those presented in [11,18]. The first four examples compare plots of VSS and NRVSS theoretical EMSE and mean step-size behaviors with Monte Carlo simulations for different environment conditions. The NRVSS sensitivity to errors in estimating the free parameter k in Eq. (14) is demonstrated in Example 2. The last two examples present a comparison between VSS and NRVSS algorithms when applied to biomedical applications using real data. Unless stated otherwise, all examples present the following common characteristics: the input signal is zero-mean Gaussian with unit power ($r_x = 1$). The additive measurement noise is zero-mean, independent and identically distributed, Gaussian, and uncorrelated with the input signal. The plant impulse response is a ten-tap ($N = 10$) Hanning window with unit norm ($\mathbf{w}^{\mathbf{o}\text{T}}\mathbf{w}^{\mathbf{o}} = 1$). $\alpha_{\text{VSS}} = 0.9997$; $\gamma_{\text{VSS}} = 2 \times 10^{-5}$; $\alpha = \alpha_{\text{VSS}}$; $\gamma = N\gamma_{\text{VSS}}/2$[5]; $\mathbf{w}(0) = [0\,0\,0 \;\ldots\; 0]^{\text{T}}$; $\beta_{\text{VSS}}(0) = \beta(0) = 0.01$; 500 runs. Only one in every ten samples is plotted in order to obtain smooth simulation curves. The signal-to-noise ratio is defined as $\text{SNR} = 10\log_{10}(r_x/r_z)$.

1. *Example* 1: NRVSS and VSS algorithms, correlated input signal generated by a first order autoregressive filter ($x(n) = a_1 x(n-1) + u(n)$) with $a_1 = 0.7$. SNR = 20 dB. Nonstationary plant generated by a first order random walk model

[5]This choice of γ equates the steady-state values of $E\{\beta(n)\}$ and $E\{\beta_{\text{VSS}}(n)\}$ for a noiseless environment and white input (see Eq. (17)). For large SNR, $\lim_{n\to\infty} E\{\beta_{\text{VSS}}(n)\}$ will exceed $\lim_{n\to\infty} E\{\beta(n)\}$ by a bias approximately equal to $\gamma\, r_z/(1-\alpha)$.

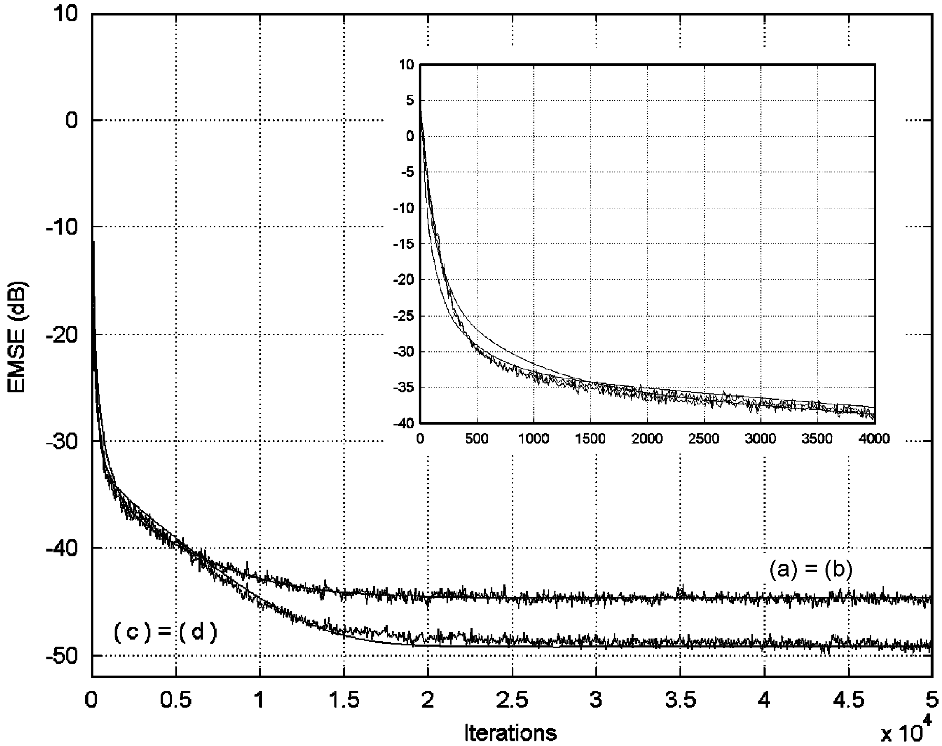

Fig. 2. Excess mean-square error (EMSE) for Example 1. Comparisons between (a) VSS model; (b) VSS simulation; (c) NRVSS model; and (d) NRVSS simulations. Inset shows a detail of the first 4000 iterations.

as in Eq. (1) with $r_q = 10^{-10}$ and $\mathrm{DN} = 3 \times 10^{-4}$. $\mathbf{w^o}(0)$ is a 10-tap normalized Hanning window.

2. *Example* 2: NRVSS and VSS algorithms, stationary case, white input signal, and SNR = 60, 20, 15, 10, 5, and 0 dB.
3. *Example* 3: NRVSS and VSS algorithms, white input signal, and measurement noise with abrupt power variation (at iteration 100,000 the additive noise power changes from $r_z = 10^{-6}$ to $r_z = 10^{-2}$).
4. *Example* 4: NRVSS and VSS algorithms, white input signal and plant with abrupt signal change (at iteration 100,000 the plant changes from $\mathbf{w^o}$ to $-\mathbf{w^o}$). $[\beta_{\mathrm{MIN}}, \beta_{\mathrm{MAX}}] = [0, 0.09]$; SNR = 60 dB.
5. *Example* 5: In this example a real epoch of a power-line contamined electrocardiographic (ECG) signal is processed by both VSS and NRVSS algorithms. ECG signal was acquired by a four-channel biomedical acquisition system [26] from a 21 years old normal subject. Laboratory acquisition used high impedance electrode coupling and 30 kΩ unbalanced impedances between differential amplifier inputs in order to simulate a signal highly contaminated due to manipulation of a patient during surgery or due to severe influence of electric equipment. The sampling frequency was 1 kHz. The VSS and NRVSS parameters were: 30 coefficients; $\beta_{\mathrm{VSS}}(0) = \beta(0) = 0.03$; $[\beta_{\mathrm{MIN}}, \beta_{\mathrm{MAX}}] = [0, 0.03]$. The reference signal was obtained from the power-line transformer equipment and contains an almost pure 60 Hz sinusoidal signal. The ECG signal is

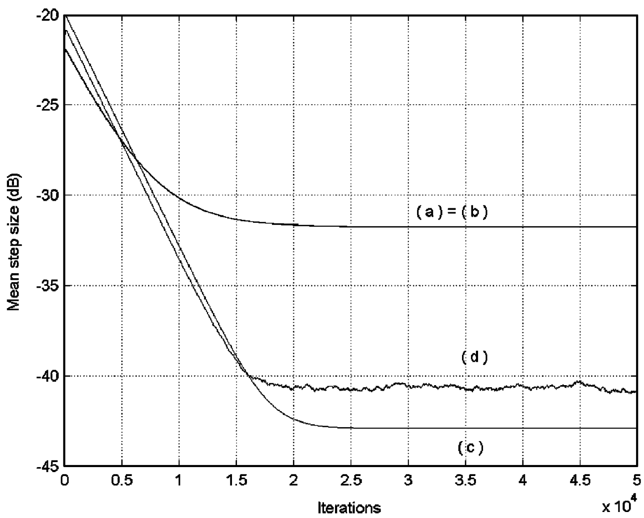

Fig. 3. Mean behavior of the step-size ($\beta(n)$) for Example 1. Comparisons between (a) VSS model; (b) VSS simulation; (c) NRVSS model; and (d) NRVSS simulations.

contamined by the 60 Hz interference and high order harmonics (nonlinear contamination). The k parameter in Eq. (14) can be accurately estimated since power-line amplitude and frequency are very stable. The SNR is −9.2 dB.

6. *Example* 6: In this example a real electroencephalographic (EEG) signal is artificially contaminated with a real electro-oculographic (EOG) interference [27] through a one tap fixed filter with values 1, 2.5, 5, and 10 (corresponding to SNR = 0, −8, −14, and −20 dB, respectively) [28], and a delay of three samples [29]. This example demonstrates the usefulness of the NRVSS in the case of nonstationary reference signals. The adaptive filter has ten taps and the reference signal power is recursively estimated at each sample using a 300 tap moving average filter $r_x(n) = r_x(n-1) - x^2(n-K-1) + x^2(n)$, for $K = 300$. $\gamma_{VSS} = 5 \times 10^{-7}$; $\beta(0) = 0.016$; $[\beta_{MIN}, \beta_{MAX}] = [0, 0.032]$.

Table 1
Steady-state EMSE for VSS and NRVSS algorithms under different SNR conditions

SNR	VSS			NRVSS		
	[10, Eq. (43)]	[10, Eq. (20)]	Simulation	Eq. (37)	Eq. (18)	Simulation
20	−44.74	−44.73	−44.75	−58.01	−58	−54.31
15	−34.68	−34.67	−34.67	−48.01	−47.98	−44.60
10	−24.46	−24.43	−24.43	−38.03	−37.92	−34.93
5	−13.52	−13.52	−13.51	−28.06	−27.71	−25.17
0	Diverges	Diverges	Diverges	−18.16	−16.87	−14.92

Comparisons between theoretical models and Monte Carlo simulations. All results are in dB.

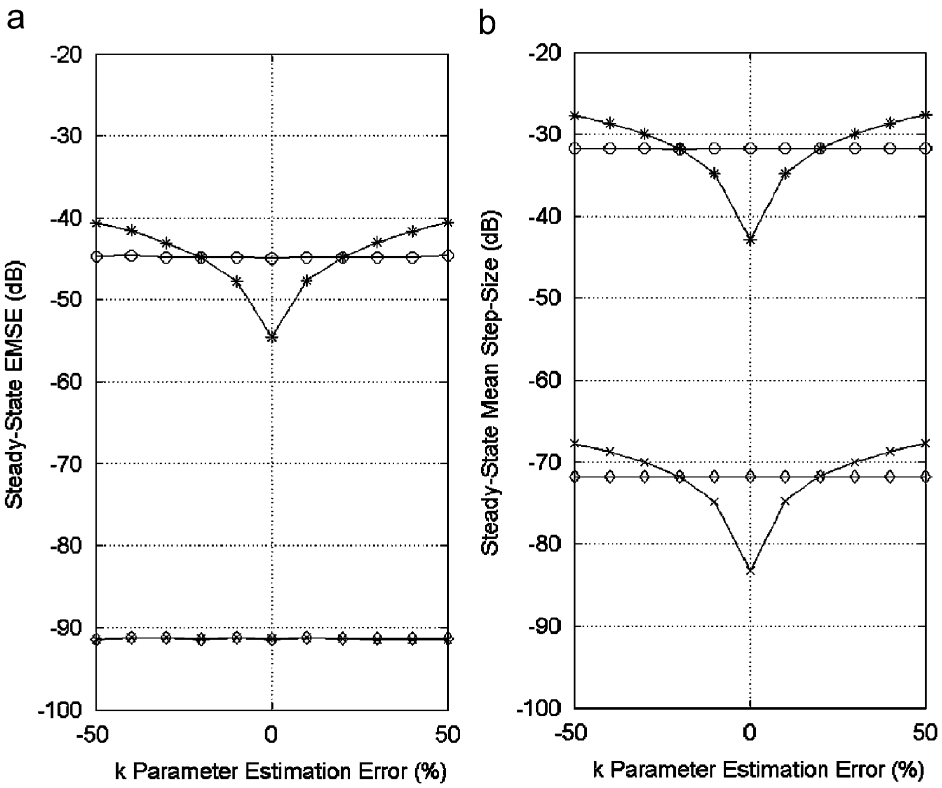

Fig. 4. Steady-state performance of the NRVSS algorithm, according to Example 2, with estimation errors of k parameter in Eq. (14). (a) Steady-state EMSE for SNR = 60 dB (◇—VSS algorithm and ×—NRVSS algorithm) and for SNR = 20 dB (○—VSS algorithm and ∗—NRVSS algorithm); (b) steady-state mean step size for SNR = 60 dB (◇—VSS algorithm and ×—NRVSS algorithm) and for SNR = 20 dB (○—VSS algorithm and ∗—NRVSS algorithm). Horizontal axis presents percentual error of the estimation of k parameter in Eq. (14). Vertical axis is in dB.

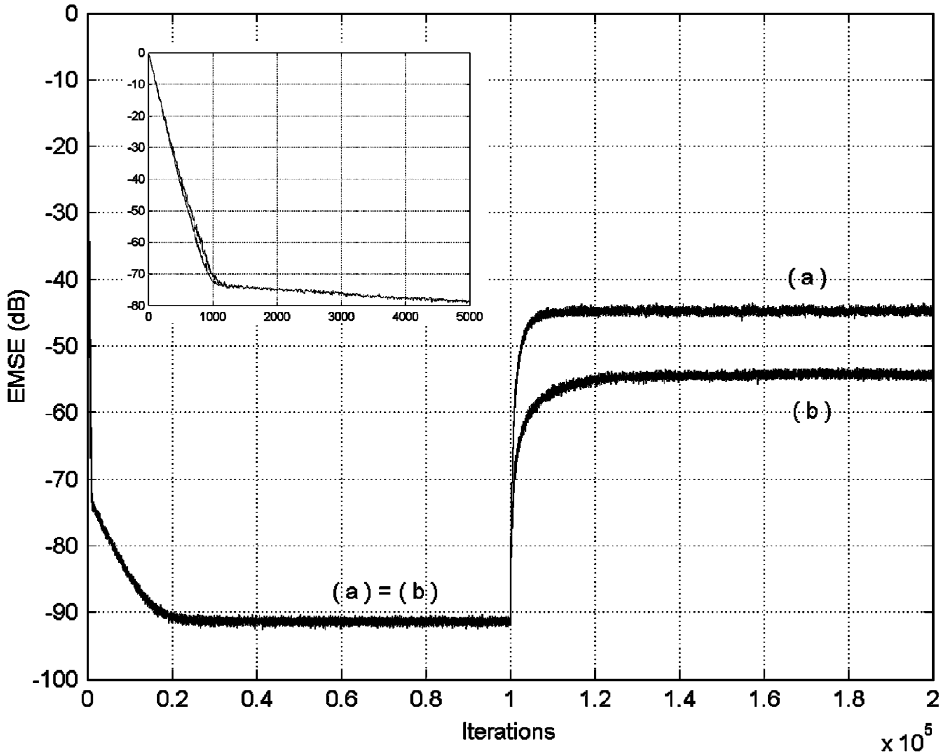

Fig. 5. Excess mean-square error (EMSE) for Example 3. Comparisons between (a) VSS and (b) NRVSS simulations. The inset shows a detail of the first 2000 iterations.

These examples illustrate the ability of the NRVSS algorithm to achieve higher cancellation levels as the SNR decreases. This occurs due to the low algorithm sensitivity to the noise power.

Example 1 (Figs. 2 and 3) shows comparisons between VSS and NRVSS for correlated input signals and a time-varying channel with DN equal to 3×10^{-4}. Both algorithms present approximately the same initial transient behavior but NRVSS achieves lower steady-state EMSE (difference of 4.3 dB). The theoretical models produced good predictions for both algorithms. This example illustrates that NRVSS inherits the VSS good performance characteristics discussed in [19]. The inset in Fig. 2 shows that the transient behavior is basically the same for both algorithms. Simulations and analytical models are almost superimposed. Fig. 3 shows the evolution of the mean step-size for both algorithms. Since the NRVSS update Eq. (9) permits the step-size to fluctuate about zero (positive and negative amplitudes) it can achieve lower values than VSS. Note, however, that Eq. (8) avoids the use of negative step sizes by the adaptive algorithm. The improvement in steady-state EMSE is about 4 dB relative to VSS, while maintaining the same initial transient behavior. As the DN decreases the NRVSS performance improves, compared to VSS. For $\mathrm{DN}\geqslant3\times10^{-3}$ NRVSS presents worse performance than VSS and for $\mathrm{DN}\geqslant3\times10^{-2}$ both algorithms become unstable.

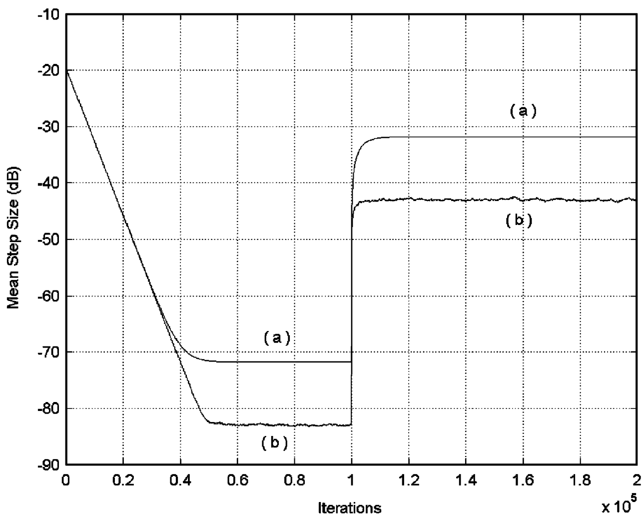

Fig. 6. Mean behavior of the step-size ($\beta(n)$) for Example 3. Comparisons between (a) VSS and (b) NRVSS simulations.

Example 2 verifies the performance of the NRVSS under low SNR and with respect to errors

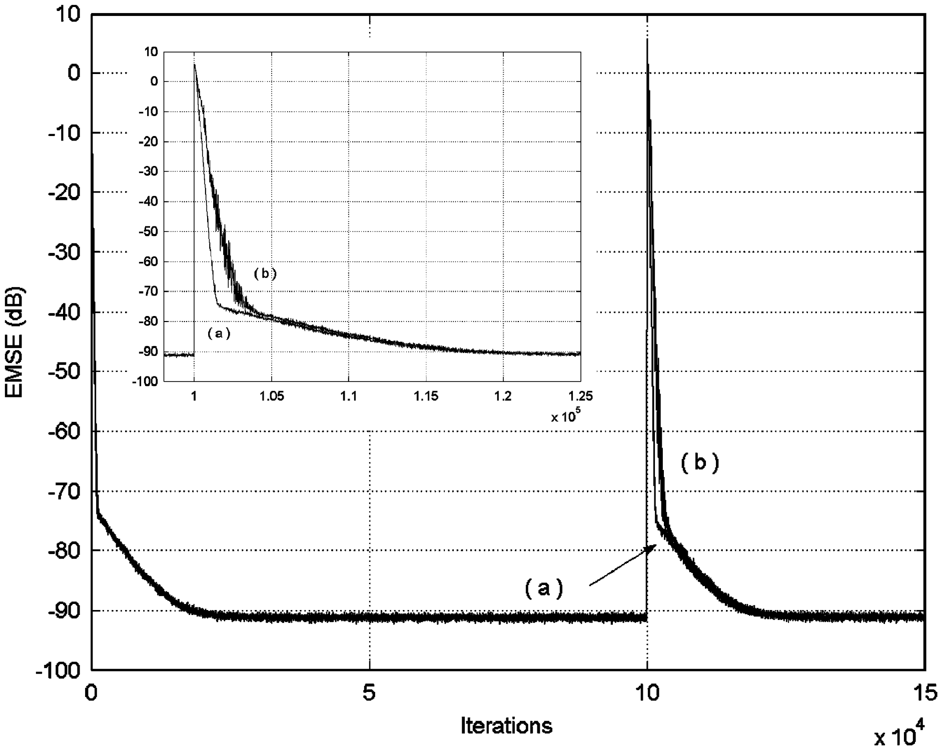

Fig. 7. Excess mean-square error (EMSE) for Example 4. Comparisons between (a) VSS and (b) NRVSS simulations. Inset details recovery behavior.

in estimating the parameter k in Eq. (14). Table 1 presents the steady-state EMSE for VSS and RVSS for a SNR ranging from 20 to 0 dB. Simulation results are compared with theoretical predictions obtained from Eqs. (18) and (37). The main causes of mismatch between the analytical and simulation results seem to be the following: (a) the neglected fourth order moments in Eqs. (26) and (28); (b) the assumption of independence of the step-size, $\mathbf{x}(n)$ and $z(n)$; (c) the assumption of independence between $\mathbf{v}(n)$ and $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$; and (d) the assumption of $E\{\beta(n)\} \cong E\{\beta^2(n)\}^{1/2}$. Assumption of a Gaussian distribution to the weight error vector [30] does not result in improvements of the theoretical results. Fig. 4 shows that estimation errors in the range of 20% still lead to a better performance of NRVSS when compared with VSS. Note that, differently from other VSS algorithms [15–17] that assume knowledge of the additive noise power (hard to accurately estimate), the NRVSS uses estimations of the reference signal, which is always available and noise free. In fact, in practical systems, errors in estimating the reference signal power can be considered negligible for stationary input signals.

Example 3 (Figs. 5 and 6) and Example 4 (Figs. 7 and 8) show that the VSS original recovery ability for abrupt changes is retained by the NRVSS. Under lower SNR conditions (Figs. 5 and 6 after noise power change) the NRVSS shows a better performance in steady-state conditions.

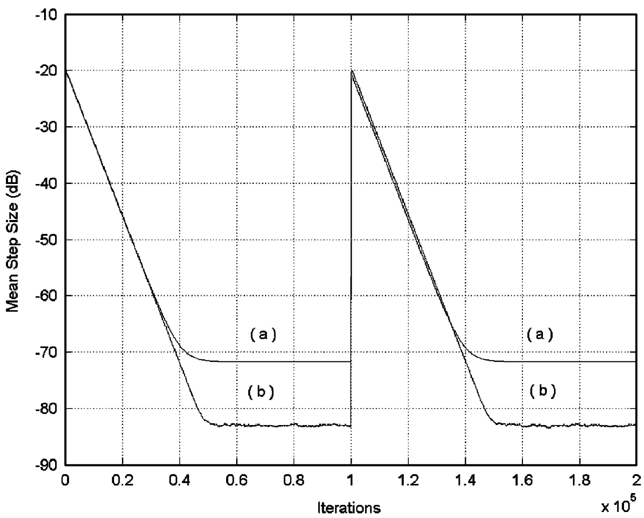

Fig. 8. Mean behavior of the step-size ($\beta(n)$) for Example 4. Comparisons between (a) VSS and (b) NRVSS simulations.

Example 5 presents a comparison between the VSS and NRVSS algorithms in a real application of interference suppression in bioelectric signals. The

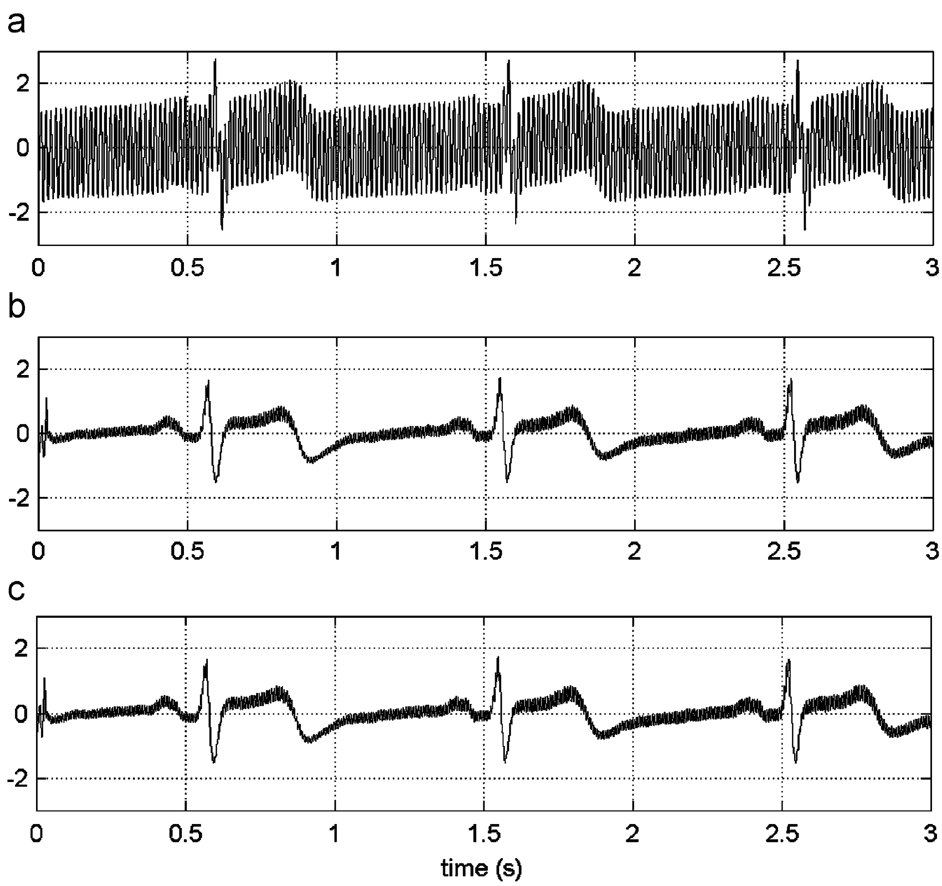

Fig. 9. Comparisons between VSS and NRVSS algorithms applied to a biomedical interference suppression application. (a) Power-line contaminated real ECG signal; (b) VSS processed ECG; and (c) NRVSS processed ECG.

$z(n)$ signal in Fig. 1 is the 60 Hz noise-free ECG and the reference signal is the main power-line signal. Fig. 9 shows the first 3 s from the beginning of the adaptive process. This figure illustrates the good performance of both algorithms in canceling power-line interference from ECG signals. Though Fig. 9 shows that both algorithms are capable of suppressing 60 Hz interference, analysis of the power spectrum obtained through the fast Fourier transform shows that NRVSS provides an extra 11 B of attenuation (-4.5 dB) compared to VSS (6.5 dB). Fig. 10 shows the instantaneous step-size behavior of both algorithms. The initial adaptation velocity is approximately equal for both algorithms while NRVSS achieves lower step-size values in steady state.

In Example 6 an EEG epoch is artificially contaminated by EOG activity. The VSS LMS algorithm is used to reduce the eye movement interference. The parameter k is estimated at each sample from the nonstationary EOG signal, in order to produce real-time estimations of its power. Figs. 11 and 12 present, respectively, the instantaneous squared error and the instantaneous step-size

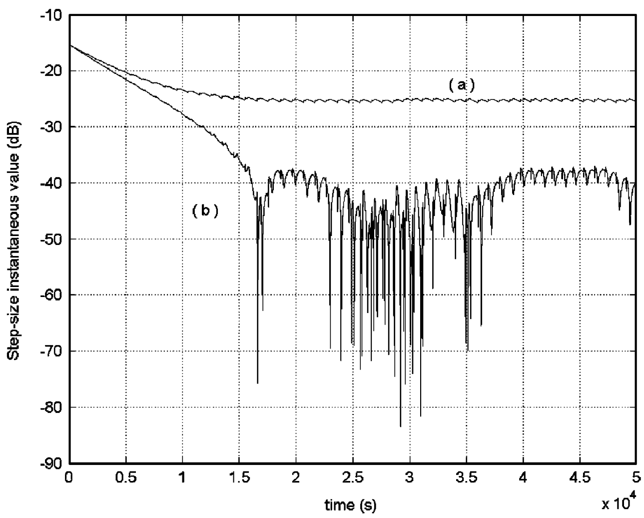

Fig. 10. Instantaneous value of VSS (a) and NRVSS (b) step-sizes for Example 5.

for NRVSS and VSS algorithms for the case of a SNR = -8 dB. This SNR is representative of electrodes placed at Cz-M1 (left mastoid) with reference at Fpz, according to the 10–20 International System.

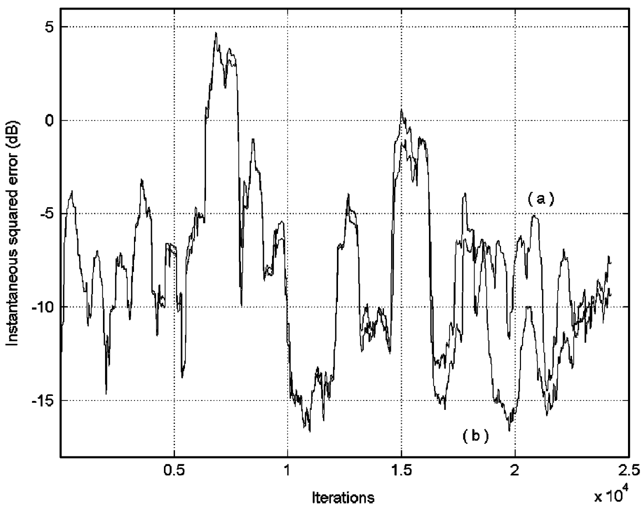

Fig. 11. Instantaneous squared error of VSS (a) and NRVSS (b) algorithms for Example 6 and SNR $= -8$ dB. One in every fifty samples is plotted to obtain smooth curves.

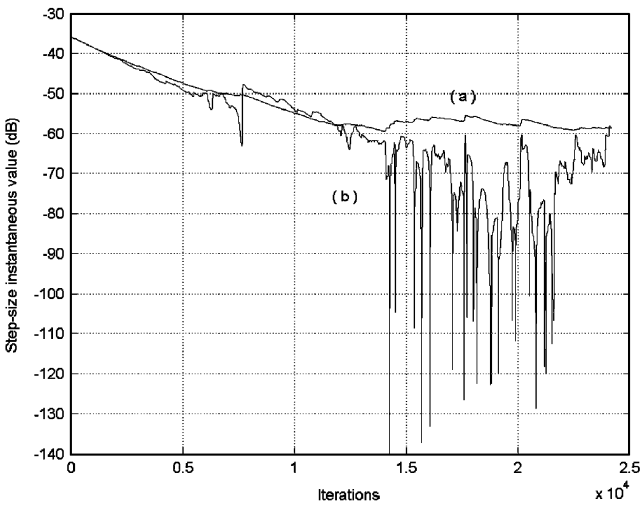

Fig. 12. Instantaneous value of VSS (a) and NRVSS (b) step-sizes for Example 6 and SNR $= -8$ dB. One in every fifty samples is plotted to obtain smooth curves.

Such derivation is characteristic of auditory evoked potential acquisitions [31]. Even in such unfavorable condition (nonstationary reference signal) NRVSS overcomes VSS in performance. Simulations with SNR $= 0$, -14, and -20 dB, representing electrodes placed at different scalp positions, presented approximately the same result.

In spite of the good characteristics of the NRVSS algorithm, extensive simulations have shown that stability bounds of Eq. (9), with relation to α_{VSS}, γ_{VSS}, and with the DN, are somewhat tighter than the equivalent bounds for VSS. However, when both algorithms converge the NRVSS presents a better steady-state performance than VSS. A bounded MSE can be always guaranteed by Eq. (8). The maximum allowed step size must be chosen in light of both the companion adaptive algorithm and the environment conditions.

8. Conclusions

This work presented a new VSS algorithm based on the VSS algorithm originally proposed by Kwong and Johnston. Analysis has demonstrated that the new algorithm is less sensitive to the power of the measurement noise when compared to VSS, at the price of a very small increase in the computational cost. Analytical models were developed for the new algorithm and for the conventional VSS for Gaussian input signals and small step-size fluctuations. Monte Carlo simulations illustrated the validity of the theoretical results and the properties of the new algorithm. The NRVSS algorithm constitutes an alternative to the conventional VSS, being especially attractive for applications with a SNR lower than 40 dB.

Acknowledgment

This work was supported by the Brazilian Ministry of Science and Technology (CNPq) under Grants 301866/2005-8 and 307024/2006-7.

Appendix A. Derivation of $E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^{2P}\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)\}$

Assuming two zero-mean Gaussian random variables y_1 and $\mathbf{y}_2$, they can be expanded as an orthonormal series [32] given by

$$\begin{cases} y_1 = k_0 w_1, \\ \mathbf{y}_2 = \mathbf{a}_1 w_1 + \mathbf{A}_2 \mathbf{w}_2, \end{cases} \tag{A.1}$$

where $\mathbf{w}_2 = [w_2\, w_3 \ldots w_{N+1}]^{\mathrm{T}}$, $E\{w_i w_j\}|_{i \neq j} = 0$, $E\{w_i^2\} = 1$, $\mathbf{a}_1$ is a vector, and $\mathbf{A}_2$ is a matrix, the last two with dimensions compatible with $\mathbf{y}_2$. Hence,

$$E\{(y_1)^{2P}\mathbf{y}_2^{\mathrm{T}}\mathbf{y}_2\} = \frac{\mathbf{a}_1^{\mathrm{T}}\mathbf{a}_1}{k_0^2} E\{y_1^{2(P+1)}\} + \mathrm{tr}\{\mathbf{A}_2^{\mathrm{T}}\mathbf{A}_2\} E\{(y_1)^{2P}\}. \tag{A.2}$$

The expansion parameters can be evaluated through the following relations:

$$\begin{cases} E\{y_1^2\} = k_0^2, \\ E\{y_1 \mathbf{y}_2\} = k_0 \mathbf{a}_1, \\ E\{\mathbf{y}_2^{\mathrm{T}}\mathbf{y}_2\} = \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_1 + \mathrm{tr}\{\mathbf{A}_2^{\mathrm{T}}\mathbf{A}_2\}. \end{cases} \tag{A.3}$$

Substituting (A.3) in (A.2) yields

$$E\{(y_1)^{2P}\mathbf{y}_2^{\mathrm{T}}\mathbf{y}_2\} = E\{(y_1)^{2P}\}E\{\mathbf{y}_2^{\mathrm{T}}\mathbf{y}_2\} + (E\{(y_1)^{2P}y_1^2\} - E\{(y_1)^{2P}\}E\{y_1^2\})E\{y_1\mathbf{y}_2^{\mathrm{T}}\}E\{y_1\mathbf{y}_2\}/E\{y_1^2\}^2. \quad \text{(A.4)}$$

If $y_1 = \mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n)$, $\mathbf{y}_2 = \mathbf{x}(n)$, $P = 2$, and the expected value is conditioned on $\mathbf{v}(n)$, (A.4) yields

$$E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^4\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)|\mathbf{v}(n)\} = 3Nr_0\,\mathrm{tr}\{\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\} + 12\mathrm{tr}\{\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\mathbf{R}\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\}, \quad \text{(A.5)}$$

where independence between $\mathbf{v}(n)$ and $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$ has been assumed [24]. Neglecting the weight error fourth order moments,

$$E\{\mathrm{tr}\{\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\mathbf{A}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\}\} \cong \mathrm{tr}\{\mathbf{R}\mathbf{K}(n)\mathbf{A}\mathbf{K}(n)\}, \quad \text{(A.6)}$$

(where $\mathbf{A}$ means $\mathbf{R}$ or $\mathbf{RR}$), (A.5) leads to (26). This approximation is better as the algorithm converges, i.e. in steady-state conditions. In general, good results are obtained with this expression whenever the variance of the weights is larger than their mean.

Appendix B. Derivation of $E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^{2P}\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)\}$

Assuming three Gaussian random variables y_1, y_2, and y_3, they can be expanded as an orthonormal series [32] given by

$$\begin{cases} y_1 = k_0w_1, \\ y_2 = a_1w_1 + a_2w_2, \\ y_3 = b_1w_1 + b_2w_2 + b_3w_3, \end{cases} \quad \text{(B.1)}$$

where $E\{w_kw_l\} = 0$ for $k \neq l$ and $E\{w_kw_l\} = 1$ for $k = l$; as a result

$$E\{(y_1)^{2P}y_2^2y_3^2\} = \frac{a_1^2b_1^2}{k_0^4}E\{(y_1)^{2P}y_1^4\} + (3a_2^2b_2^2 + a_2^2b_3^2)E\{(y_1)^{2P}\} + \frac{1}{k_0^2}(a_1^2b_2^2 + a_1^2b_3^2 + a_2^2b_1^2 + 4a_1a_2b_1b_2)E\{(y_1)^{2P}y_1^2\}. \quad \text{(B.2)}$$

The expansion parameters can be evaluated through the following relations:

$$\begin{cases} E\{y_1^4\} = 3k_0^4, & E\{y_1y_2\} = k_0a_1, \\ E\{y_1^2\} = k_0^2, & E\{y_1y_3\} = k_0b_1, \\ E\{y_2^2\} = a_1^2 + a_2^2, & E\{y_2y_3\} = a_1b_1 + a_2b_2, \\ E\{y_3^2\} = b_1^2 + b_2^2 + b_3^2. \end{cases} \quad \text{(B.3)}$$

Substituting (B.3) in (B.2) yields

$$E\{(y_1)^{2P}y_2^2y_3^2\} = \frac{1}{E\{y_1^2\}^4}E\{y_1y_2\}^2E\{y_1y_3\}^2E\{(y_1)^{2P}y_1^4\} + \frac{k_1E\{(y_1)^{2P}y_1^2\}}{E\{y_1^2\}^3(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)} + \frac{k_2E\{(y_1)^{2P}\}}{E\{y_1^2\}^2}, \quad \text{(B.4)}$$

where

$$k_1 = E\{y_1^2\}^2E\{y_2^2\}E\{y_3^2\}E\{y_1y_2\}^2 + E\{y_1^2\}^2E\{y_2^2\}^2E\{y_1y_3\}^2 + 4E\{y_1^2\}^2E\{y_2^2\}E\{y_1y_2\}E\{y_1y_3\}E\{y_2y_3\} - 7E\{y_1^2\}E\{y_2^2\}E\{y_1y_2\}^2E\{y_1y_3\}^2 - E\{y_1^2\}E\{y_3^2\}E\{y_1y_2\}^4 + 6E\{y_1y_2\}^4E\{y_1y_3\}^2 - 4E\{y_1^2\}E\{y_1y_2\}^3E\{y_1y_3\}E\{y_2y_3\} \quad \text{(B.5)}$$

and

$$k_2 = 2E\{y_1^2\}^2E\{y_2y_3\}^2 - 4E\{y_1^2\}E\{y_1y_2\}E\{y_1y_3\}E\{y_2y_3\} + 3E\{y_1y_2\}^2E\{y_1y_3\}^2 + E\{y_1^2\}^2E\{y_2^2\}E\{y_3^2\} - E\{y_1^2\}E\{y_3^2\}E\{y_1y_2\}^2 - E\{y_1^2\}E\{y_2^2\}E\{y_1y_3\}^2. \quad \text{(B.6)}$$

(1) *Case* 1: If $y_1 = \mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n)$, $y_2 = x(n-i)$, $y_3 = x(n-j)$, and $P = 1$, then (B.4) can be used to obtain the following conditioned expectation

$$E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^2(\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n))^2|\mathbf{v}(n)\} = \mathrm{tr}\left\{\left(8\mathbf{RRR} + 4Nr_x\mathbf{RR} + \left(N^2r_x^2 + 2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right)\mathbf{R}\right)\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\right\}. \quad \text{(B.7)}$$

Assuming independence between $\mathbf{v}(n)$ and $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$ [24] and averaging over $\mathbf{v}(n)$, (B.7) comes to (27).

(2) *Case* 2: If $y_1 = \mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n)$, $y_2 = x(n-i)$, $y_3 = x(n-j)$, and $P = 2$, then (B.4) can be used to obtain

$$E\{(\mathbf{v}^{\mathrm{T}}(n)\mathbf{x}(n))^4(\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n))^2|\mathbf{v}(n)\} = 24\mathrm{tr}\{\mathbf{R}\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\mathbf{R}\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\} + \mathrm{tr}\{\mathbf{R}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\mathbf{A}\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)\}, \tag{B.8}$$

where

$$\mathbf{A} = 48\mathbf{R}\mathbf{R}\mathbf{R} + 24r_xN\mathbf{R}\mathbf{R} + 3\left(N^2r_x^2 + 2\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}r_{j-i}^2\right)\mathbf{R}. \tag{B.9}$$

Assuming independence between $\mathbf{v}(n)$ and $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$ [24], averaging over $\mathbf{v}(n)$, and neglecting the fourth order moments of the weights leads to (28).

References

[1] S. Haykin, Adaptive Filter Theory, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1991.

[2] C. Breining, et al., Acoustic echo control. An application of very-high-order adaptive filters, IEEE Signal Process. Mag. 16 (4) (July 1999) 42–69.

[3] B. Widrow, S.D. Stearns, Adaptive Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[4] D.G. Manolakis, V.K. Ingle, S.M. Kogon, Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing, Artech House Publishers, 2000.

[5] S. Koike, A class of adaptive step-size control algorithms for adaptive filters, IEEE Trans. Signal Process. 50 (6) (June 2002) 1315–1326.

[6] M.A. Aldajani, Adaptive step-size sign least mean squares, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 2004, pp. 669–672.

[7] B. Krstajic, L.J. Stankovic, Z. Uskokovic, An approach to variable step-size LMS algorithm, Electron. Lett. 38 (16) (August 2002) 927–928.

[8] J.B. Evans, A new variable step-size method suitable for efficient VLSI implementation, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 1991, pp. 2105–2108.

[9] J. Okello, et al., A new modified variable step-size for the LMS algorithm, in: Proceedings of the International Symposium on Circuits and Systems, 1998, pp. 170–173.

[10] R.H. Kwong, E.W. Johnston, A variable step-size LMS algorithm, IEEE Trans. Signal Process. 40 (7) (July 1992) 1633–1642.

[11] T. Abousnar, K. Mayyas, A robust variable step-size LMS-type algorithm: analysis and simulation, IEEE Trans. Signal Process. 45 (3) (March 1997) 631–639.

[12] R.W. Harris, D. Chabries, F. Bishop, A variable step (VS) adaptive filter algorithm, IEEE Trans. Acoust. Speech Signal Process. ASSP-34 (2) (April 1986) 309–316.

[13] T.J. Shan, T. Kailath, Adaptive algorithms with automatic gain control feature, IEEE Trans. Circuits Syst. 35 (1) (January 1988) 122–127.

[14] H.C. Shin, A.H. Sayed, W.J. Song, Variable step-size NLMS and affine projection algorithms, IEEE Signal Process. Lett. 11 (2) (2004) 132–135.

[15] Y. Wei, S.B. Gelfand, J.V. Krogmeier, Noise-constrained Least Mean Square algorithm, IEEE Trans. Signal Process. 49 (9) (September 2001) 1961–1970.

[16] A. Mader, H. Puder, G.U. Schmidt, Step-size control for acoustic echo cancellation filters—an overview, Signal Process. 80 (2000) 1697–1719.

[17] J. Benesty, H. Rey, L.R. Vega, S. Tressens, A nonparametric VSS NLMS algorithm, IEEE Signal Process. Lett. 12 (10) (2006) 581–584.

[18] S. Koike, A novel adaptive step-size control algorithm for adaptive filters, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 1999, pp. 1845–1848.

[19] C.G. Lopes, J.C.M. Bermudez, Evaluation and design of variable step-size adaptive algorithms, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 2001, pp. 3845–3848.

[20] M.F. Chimeno, R. Pallàs-Areny, A comprehensive model for power line interference in biopotential measurements, IEEE Trans. Instrum. Meas. 49 (3) (June 2000) 535–540.

[21] V. Almenar, A. Albiol, A new adaptive scheme for ECG enhancement, Signal Process. 75 (1999) 253–263.

[22] M.H. Costa, J.C.M. Bermudez, A robust variable step-size algorithm for LMS adaptive filters, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 2006, pp. 93–96.

[23] A.H. Sayed, Fundamentals of Adaptive Filtering, Wiley-IEEE Press, 2003.

[24] J. Minkoff, Comment: On the unnecessary assumption of statistical independence between reference signal and filter weights in feedforward adaptive systems, IEEE Trans. Signal Process. 49 (5) (May 2001) 1109.

[25] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1991.

[26] M.C. Tavares, M.H. Costa, F.A. Costa, F.M. Azevedo, Microcontroller-based multiprocessor system for ECG, EEG and MLAEP registering during surgery, in: Proceedings of the World Congress on Medical Physics and Biomedical Engineering, 2003, p. 1.

[27] A.L. Goldberger, et al., PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.

[28] R.J. Croft, R.J. Barry, EOG correction: a new perspective, Electroencephalogr. Clin. Neurophysiol. 107 (1998) 387–394.

[29] P. He, G. Wilson, C. Russel, Removal of ocular artifacts from electro-encephalogram by adaptive filtering, Med. Biol. Eng. Comput. 42 (2004) 407–412.

[30] N.J. Bershad, Z.Q. Lian, On the probability density function of the LMS adaptive filter weights, Acoust. Speech Signal Process. 37 (1) (1989) 43–56.

[31] N. Takano, T. Maruyama, M. Tamagawa, K. Yana, EOG artifact canceling for EEG analysis using RLS adaptive filtering technique, in: Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, 1993, pp. 348–349.

[32] N.J. Bershad, P. Celka, J.M. Vesin, Stochastic analysis of gradient adaptive identification of nonlinear systems with memory for Gaussian data and noisy input and output measurements, IEEE Trans. Signal Process. 47 (3) (1999) 675–689.