MEAN-SQUARE ERROR ANALYSIS OF THE LMS-RVSS ADAPTIVE ALGORITHM

Márcio H. Costa, and José C. M. Bermudez

Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina
Centro Tecnológico, Cidade Universitária, 88040-900, Florianópolis-SC, Brazil
phone: +(55)(48) 3721-9506, fax: +(55)(48) 3721-9280, email: costa@.eel.ufsc.br; j.bermudez@ieee.org

ABSTRACT

This work presents the mean square error analysis of the Robust Variable Step Size (RVSS), which has been recently proposed. An analytical model is derived for the excess mean square error for white Gaussian input signals and slow adaptation. As a side benefit of the analysis, a new theoretical model is also obtained for the classical VSS algorithm. Monte Carlo simulations verify the accuracy of the analytical model. A practical example with real nonstationary biomedical signals illustrates the applicability of the algorithm.

1. INTRODUCTION

Adaptive filtering has been extensively employed in many practical applications. Important results have been obtained, for instance, in noise and interference canceling for biomedical applications [1]. The Least Mean Square (LMS) adaptive filter family is very attractive for implementation of low-cost real-time systems due to its low computational complexity and robustness [2]. Thus, LMS is still the golden standard for performance evaluation of existing and new adaptive processing strategies.

It is well known that LMS-based algorithms' performance is highly dependent on the choice of the step size parameter. Larger step sizes tend to 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. Ideally, the step size should be large during the early adaptation, and have its value progressively reduced as the algorithm approaches steady-state. Thus, variable step size strategies are common solutions for obtaining both fast tracking and good steady-state performance.

Several variable step size LMS-type algorithms have been proposed. Two particularly interesting strategies were introduced in [3] and in [4]. The performances of these algorithms are largely insensitive to the power and to the statistics of the measurement noise. The price for such robustness is an increase in LMS complexity that is proportional to the number of adaptive filter coefficients.

To remain attractive for some demanding real-time applications such as anesthesiology monitoring [5], variable step size strategies should impose a minimal computational penalty to the basic LMS adaptive filter. The most promising low cost step size adjustment criteria are based on the instantaneous squared error [6], [7], on the frequency of gradient estimation signal changes [8], and on the correlation between input and error signals [9]. However, experimental results show that the steady-state performances provided by these techniques can be highly dependent on the measurement noise power level. This sensitivity can be explained by a mean steady-state step size bias which is dependent on the noise power. As a result, the performances of these algorithms tend to reduce with the signal-to-noise ratio (SNR). To overcome this problem, some algorithms incorporate measurement noise variance estimators to lessen the performance losses [10].

The variable step size (VSS) algorithm [6] provided a very effective strategy for LMS step size adjustment. Later on, authors of alternative variable step size algorithms have claimed better performances than VSS [7], [11]. More recently, the work in [12] demonstrated that VSS provides the closest to the optimum step size sequence when properly designed. This result revived the interest on VSS. So far, VSS appears to lead to the best tradeoff between convergence speed and steady-state misadjustment, even considering its intrinsic large sensitivity to the noise power.

Recently, a modified version of the VSS algorithm has been introduced [13]. The Robust Variable Step Size (RVSS) presents a smaller sensitivity to the measurement noise than VSS, at the price of a small computational cost increase. It can be applied whenever the input signal is not a noiseless constant modulus signal, and has been successfully applied to interference cancellation in biomedical applications [14].

This work presents a mean-square error (MSE) analysis of the RVSS algorithm. A recursive analytical model is derived for white Gaussian signals and slow adaptation. Monte Carlo simulations show excellent agreement with the theoretical model during transient. In steady-state, the model's accuracy reduces with the SNR. An example compares VSS and RVSS performances using real biomedical signals.

2. VSS-RVSS UPDATE EQUATIONS

The weight-error vector update equation of the variable step size LMS algorithm is given by

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \mu(n)e(n)\mathbf{x}(n) \quad (1)$$

where n is the discrete time and $\mathbf{v}(n)=\mathbf{w}(n)-\mathbf{w}^{\mathbf{o}}$ is the weight-error vector, where $\mathbf{w}(n)=[w_0(n)\ w_1(n)\ \dots\ w_{N-1}(n)]^T$ is the adaptive weight vector and $\mathbf{w}^{\mathbf{o}}=[w^{\mathrm{o}}_0\ w^{\mathrm{o}}_1\ \dots\ w^{\mathrm{o}}_{N-1}]^T$ is the minimum MSE weight vector. $\mu(n)$ is the variable step size and $\mathbf{x}(n)=[x(n)\ x(n-1)\ \dots\ x(n-N+1)]^T$ is the input signal vector. The error signal is defined as

$$e(n) = z(n) - \mathbf{v}^T(n)\mathbf{x}(n) \quad (2)$$

where $z(n)$ is the measurement noise, which is assumed i.i.d., with power r_z and independent of $x(n)$. In the following, $x(n)$ is considered to be a zero-mean, white Gaussian signal with power r_0.

RVSS [13] is a modification of VSS [6]. Its update equation is given by (1) with

$$\mu(n) = \begin{cases} \beta_{MIN} & \text{if } |\beta(n)| < \beta_{MIN} \\ |\beta(n)| & \text{if } \beta_{MIN} \le |\beta(n)| \le \beta_{MAX} \\ \beta_{MAX} & \text{if } |\beta(n)| > \beta_{MAX} \end{cases} \quad (3)$$

with

$$\beta(n+1) = \alpha\beta(n) + \gamma\left[k\, \mathbf{x}^T(n)\mathbf{x}(n) - 1\right]e^2(n) \quad (4)$$

where k, α and γ are the control parameters and $\mathbf{x}^T(n)\mathbf{x}(n)$ is evaluated recursively. The original VSS update equation is obtained for k=0, α=α_{VSS} and γ =-γ_{VSS}. The bounding of $\mu(n)$ in (3) prevents unstable behavior and maintains the algorithm's tracking capability.

It has been shown in [13] that the measurement noise influence on the RVSS performance is minimized for

$$k = 1/(r_0 N) \quad (5)$$

Thus, a properly designed RVSS is less sensitive to the measurement noise than VSS, at the price of a small increase in the computational cost. In nonstationary applications, r_0 can be estimated recursively ($r(n)$=$r(n$-1)-$x^2(n$-K-1)+$x^2(n)$ – see Section 5).

3. LMS EXCESS MEAN SQUARE ERROR

The Excess Mean-Square Error (EMSE) of the algorithm (1) with white inputs can be approximated by [2]

$$EMSE = E\{e^2(n)\} - r_z = r_0 T_K(n) \quad (6)$$

where $E\{\cdot\}$ means statistical expectation and $T_K(n)$=$E\{\mathbf{v}^T(n)\,\mathbf{v}(n)\}$ is the mean-square deviation. It is shown in [6] that $T_K(n)$ can be approximated by

$$\begin{aligned} T_K(n+1) = &\left[1 - 2r_0 E\{\mu(n)\} + (N+2)r_0^2 E\{\mu^2(n)\}\right]T_K(n) \\ &+ E\{\mu^2(n)\}Nr_0 r_z \end{aligned} \quad (7)$$

Eq. (7) was derived assuming statistical independence between $\mu(n)$ and $x(n)$. This assumption is valid for slow adaptation (α close to 1 and small γ).

4. VSS-RVSS MEAN SQUARE STEP SIZE

To obtain analytical models for (6) and (7) it is necessary to evaluate $E\{\mu(n)\}$=$E\{|\beta(n)|\}$ and $E\{\mu^2(n)\}$= $E\{\beta^2(n)\}$. The former can be approximated by $E\{\mu(n)\} \cong E\{\beta^2(n)\}^{1/2}$. The evaluation of $E\{\beta^2(n)\}$, however, is a very difficult task. To make this problem mathematically tractable, we follow [6] and assume mutual statistical independence of $\mu(n)$, $\mathbf{x}(n)$ and $\mathbf{w}(n)$. Squaring (4) and taking its expected value we obtain

$$\begin{aligned} E\{\beta^2(n+1)\} = &\alpha^2 E\{\beta^2(n)\} + 4\alpha k\gamma r_0^2 E\{\beta(n)\}T_K(n) \\ &+ 2\alpha\gamma E\{\beta(n)\}\cdot(kNr_0 - 1)\cdot E\{e^2(n)\} \\ &+ k^2\gamma^2 E\{e^4(n)\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)\} \\ &- 2k\gamma^2 E\{e^4(n)\mathbf{x}^T(n)\mathbf{x}(n)\} + \gamma^2 E\{e^4(n)\} \end{aligned} \quad (8)$$

where

$$\begin{cases} E\{e^2(n)\} = r_0 T_K(n) + r_z \\ E\{e^2(n)\mathbf{x}^T(n)\mathbf{x}(n)\} = (N+2)r_0^2 T_K(n) + Nr_0 r_z \end{cases} \quad (9)$$

were already evaluated in [13].

The last term in (8) is evaluated by rising (2) to the fourth power, taking initially the expectation conditioned on the coefficients, using the properties of Gaussian variables [15] and, finally, averaging over the coefficients. This procedure results in

$$E\{e^4(n)\} = 3r_z^2 + 6r_0 r_z T_K(n) + 3r_0^2 T_K^2(n) \quad (10)$$

Using the Gaussian Moment Factoring Theorem [2], the two remaining expected values in (8) can be simplified to:

$$\begin{aligned} E\{e^4(n)\mathbf{x}^T(n)\mathbf{x}(n)\} = &3Nr_0 r_z^2 + 12r_0^2 r_z T_K(n) \\ &+ 6Nr_0^2 r_z T_K(n) + E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^4 \mathbf{x}^T(n)\mathbf{x}(n)\right\} \end{aligned} \quad (11)$$

and

$$\begin{aligned} E\{e^4(n)\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)\} = &3N(N+2)r_0^2 r_z^2 \\ &+ 6r_z E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^2\left(\mathbf{x}^T(n)\mathbf{x}(n)\right)^2\right\} \\ &+ E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^4\left(\mathbf{x}^T(n)\mathbf{x}(n)\right)^2\right\} \end{aligned} \quad (12)$$

The three remaining expectations in (11) and (12) are evaluated in Appendices A, B and C, and are given by:

$$E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^2\left(\mathbf{x}^T(n)\mathbf{x}(n)\right)^2\right\} = (N^2 + 6N + 8)r_0^3 T_K(n), \quad (13)$$

$$E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^4 \mathbf{x}^T(n)\mathbf{x}(n)\right\} = (3N + 12)r_0^3 T_K^2(n) \quad (14)$$

and

$$E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^4\left(\mathbf{x}^T(n)\mathbf{x}(n)\right)^2\right\} = (3N^2 + 30N + 72)r_0^4 T_K^2(n) \quad (15)$$

Substituting (10) to (15) in (8) leads to

$$\begin{aligned} E\{\beta^2(n+1)\} = &\alpha^2 E\{\beta^2(n)\} + 24\gamma^2 k^2 r_0^2 r_z^2 \\ &+ 4\alpha\gamma k r_0 E\{\beta(n)\}\left(E\{e^2(n)\} - r_z\right) \\ &+ 2\gamma\left[\alpha(kNr_0 - 1)E\{\beta(n)\} + 36\gamma k^2 r_0^2(1 - r_z)\right]E\{e^2(n)\} \\ &- 24\gamma^2 k r_0 r_z\left[kr_0(N+1) - 1\right]E\{e^2(n)\} \\ &+ 3\gamma^2\left[1 + N(N+10)k^2 r_0^2 - 2kr_0(N+4)\right]E\{e^2(n)\}^2 \end{aligned} \quad (16)$$

Making k=0, α=α_{VSS}, γ=-γ_{VSS} and β=β_{VSS} in (16) yields

$$\begin{aligned} E\{\beta_{VSS}^2(n+1)\} = &\alpha_{VSS}^2 E\{\beta_{VSS}^2(n)\} + 3\gamma_{VSS}^2 E\{e^2(n)\}^2 \\ &+ 2\alpha_{VSS}\gamma_{VSS} E\{\beta_{VSS}(n)\}E\{e^2(n)\} \end{aligned} \quad (17)$$

which is a new recursive model for the evolution of the mean squared step-size of the VSS algorithm.

Using (5) in (16) yields

$$E\{\beta^2(n+1)\} = \alpha^2 E\{\beta^2(n)\} + \frac{6\gamma^2 r_0^2}{N}\left(\frac{12}{N}+1\right)T_K^2(n) + \frac{4\alpha\gamma}{N}E\{\beta(n)\}T_K(n) + \frac{12\gamma^2 r_0 r_z}{N}\left(\frac{4}{N}+1\right)T_K(n) + \frac{6\gamma^2 r_z^2}{N} \quad (18)$$

Eq. (18) is the desired model for the behavior of the mean-square step size value for the optimized RVSS algorithm.

Eqs. (6), (7) and (17) constitutes a new model for the conventional VSS algorithm, while Eqs. (6), (7) and (18) can be used to predict the behavior of the RVSS.

5. SIMULATIONS

In this section we present analytical and simulation results to verify the accuracy of the VSS and RVSS theoretical models, and to compare the performances of the two algorithms.

Three examples are presented. The aim of the first two is to demonstrate the validity and limitation of the used theoretical assumptions in the analysis and the resulting accuracy of the models. The last example demonstrates the good qualities of the RVSS when compared to the VSS when applied to real nonstationary biomedical signals.

The first example considers a SNR of 60 dB and the second example a SNR of 20 dB. The second example permits to visualize the impact of the measurement noise influence on the steady-state of the VSS algorithm. Both examples have the following common characteristics: the input signal is white Gaussian with unity power (r_0=1). The additive measurement noise is white, Gaussian and uncorrelated with the input signal. The plant is a ten-tap (N=10) Hanning window with unity norm ($\mathbf{w^{o}}^T\mathbf{w^o}$=1). α=0.9997; γ=10^{-4}; $\mathbf{w}(0)$=[0 0 0 … 0]^T and $\beta(0)$=0.01. The simulation results were averaged over 500 runs. The following parameters are used for the VSS algorithm for a fair comparison of convergence speeds: α_{VSS}=α and γ_{VSS} =2γ/N [13].

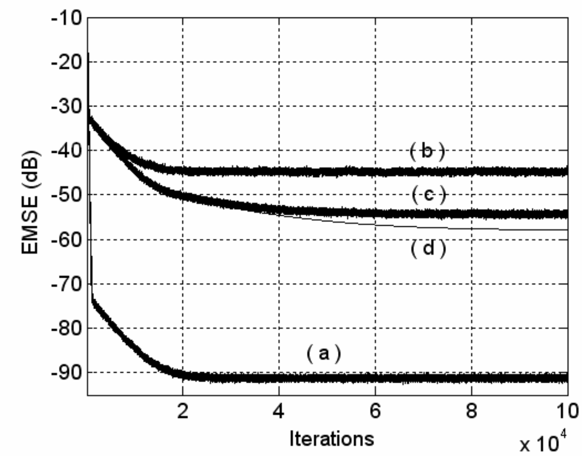

Figure 1 – Excess mean square error (EMSE) for Examples 1 and 2. (a) Analytical models and Monte Carlo simulations for VSS and RVSS are overlapped for Example 1 (SNR=60 dB). (b) Analytical model and simulation for VSS in Example 2 (SNR=20dB). (c) Simulations for the RVSS in Example 2. (d) Analytical model for the RVSS in Example 2.

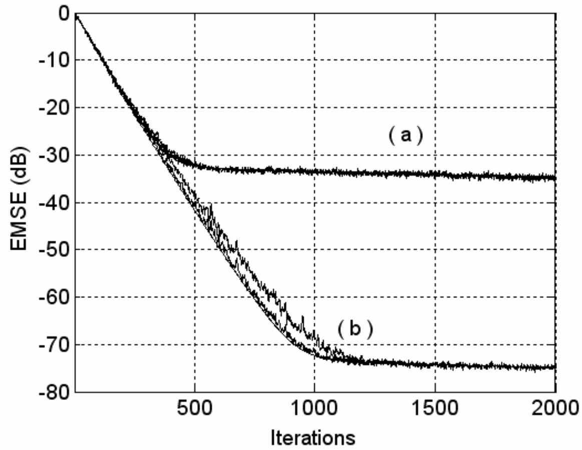

Figure 2 – Excess mean square error (EMSE) for Examples 1 and 2. First 2000 iterations of Fig. 1. (a) RVSS and VSS simulations and analytical models for Example 2 (SNR=20dB). (b) RVSS and VSS simulations and analytical models for Example 1 (SNR=60dB).

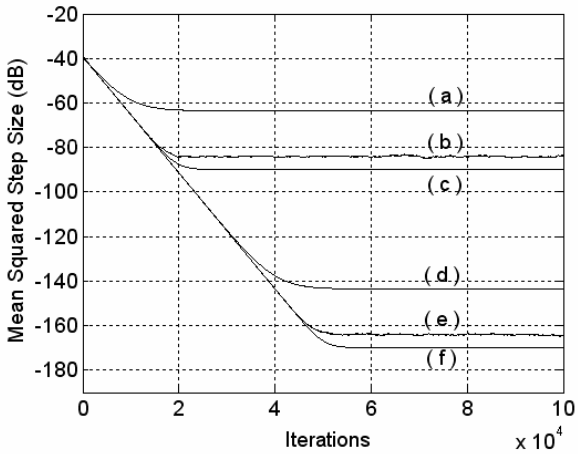

Figure 3 – Mean squared step size ($E\{\beta^2(n)\}$) evolution for Examples 1 (SNR=60dB) and 2 (SNR=20dB). (a) VSS model and simulations for Example 2. (b) RVSS simulation for Example 2. (c) RVSS model for Example 2. (d) VSS model and simulations for Example 1. (e) RVSS simulations for Example 1. (f) RVSS model for Example 1.

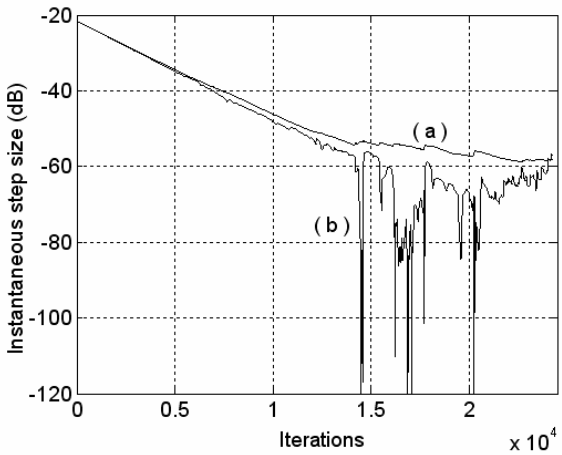

Figure 4 – Instantaneous step size for Example 3. (a) VSS algorithm. (b) RVSS algorithm.

Figs. 1 and 2 show the EMSE behavior for the VSS and RVSS algorithms for Examples 1 and 2. Both algorithms present similar behaviors for a 60 dB SNR (Example 1). Excellent matching can be verified between theoretical and simulated EMSE in both the transient and steady-state phases. For SNR=20dB, VSS and RVSS present basically the same transient but very distinct steady-state behaviors. The VSS theoretical model is able to accurately predict the algorithm's behavior while the RVSS model is conservative. This steady-state model deviation could be attributed to the neglected fourth order moments in (14) and (15). Comparison of Figs. 1b and 1c shows the smaller sensitivity of RVSS to the measurement noise influence.

Fig. 3 shows the evolution of the mean squared step size for Examples 1 and 2. The RVSS analytical model accurately predicts the transient behavior, but underestimates the steady-state behavior. Fig. 3 clearly demonstrates that RVSS can lead to lower steady-state step-size values than VSS. Several tests demonstrated that the prediction errors on the actual steady-state EMSE and on the mean squared step-size are bounded to less than -6 dB. Note that the use of the EMSE instead of the MSE clarifies the differences between theoretical and simulations curves. Thus, the assumptions and simplifications used in the analysis do not compromise the model's usefulness for most practical applications.

In the third example a real electroencephalographic (EEG) signal is artificially contaminated with a real electrooculographic (EOG) interference [16] through a one tap fixed filter with value 0.6 and a delay of five samples. The adaptive filter has ten taps and the reference signal power is estimated at each sample by a 200 tap moving average filter implemented recursively. γ_{VSS}=$5\cdot10^{-7}$; $\beta(0)$=0.0833; [β_{MIN}, β_{MAX}]=[0, 0.17]. Fig. 4 shows that both algorithms present approximately the same behavior during the transient period. After iteration 15,000 (steady-sate condition) the RVSS instantaneous step size achieves lower values.

6. CONCLUSION

This work presented a theoretical analysis of the mean-square error behavior of the LMS-RVSS algorithm. The analysis considered white Gaussian input signals and slow adaptation. As a side benefit, a new analytical model has also been obtained for the behavior of the well-known VSS algorithm. Comparisons between theoretical predictions and Monte Carlo simulations have shown excellent agreement during transient for both VSS and RVSS. In steady-state, the RVSS model's accuracy degrades for low SNR, but should still be useful for practical applications. The RVSS was shown to be an interesting choice for biomedical applications.

ACKNOWLEDGMENTS

The authors would like to thank Funpesquisa/UFSC and CNPq (Brazilian Ministry of Science and Technology) for the financial support.

APPENDIX A: $E\{(\mathbf{v}^T(n)\mathbf{x}(n))^2(\mathbf{x}^T(n)\mathbf{x}(n))^2\}$

Assuming statistical independence between $\mathbf{x}(n)$ and $\mathbf{v}(n)$ the following approximation can be obtained

$$\begin{aligned}&E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^2\left(\mathbf{x}^T(n)\mathbf{x}(n)\right)^2\right\}\\&=tr\left\{E\left\{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\right\}E\left\{\mathbf{v}(n)\mathbf{v}^T(n)\right\}\right\}\end{aligned} \tag{A1}$$

For white Gaussian input signals, the off-diagonal elements of the resulting matrix from the first expectation in the right hand side of (A1) are zero. The diagonal elements can be evaluated by the following expression

$$q_{i,i}=E\left\{x^2(n-i)\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}x^2(n-k)x^2(n-l)\right\} \tag{A2}$$

where i and j index the lines and columns, respectively ($q_{i,j}$=0 for $i\neq j$). As a result

$$E\left\{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\right\}=\left(N^2+6N+8\right)r_0^3\mathbf{I} \tag{A3}$$

Substituting (A3) in (A1) leads to (13)

APPENDIX B: $E\{(\mathbf{v}^T(n)\mathbf{x}(n))^4\mathbf{x}^T(n)\mathbf{x}(n)\}$

Assuming two Gaussian random variables y_1 and y_2 they can be expanded as an orthonormal series [17] given by

$$\begin{cases}y_1=k_0w_1\\y_2=a_0w_1+a_2w_2\end{cases} \tag{B1}$$

where $E\{w_kw_l\}$=0 for $k\neq l$ and $E\{w_kw_l\}$=1 for k=l. For slow adaptation conditions, it can be assumed that $\mathbf{v}(n)$ is slowly varying with relation to $\mathbf{x}(n)$, resulting in

$$E\left\{\left(\mathbf{v}^T(n)\mathbf{x}(n)\right)^4\mathbf{x}^T(n)\mathbf{x}(n)\right\}=E\left\{\sum_{p=0}^{N-1}E\left\{y_1^4y_{2p}^2\middle|\mathbf{v}(n)\right\}\right\} \tag{B2}$$

where y_1=$\mathbf{v}^T(n)\mathbf{x}(n)$ and y_{2p}=$x(n\text{-}p)$. Substituting (B1) in (B2) yields, after some manipulations,

$$\sum_{p=0}^{N-1}E\left\{y_1^4y_2^2\middle|\mathbf{v}(n)\right\}=15k_0^4\sum_{p=0}^{N-1}a_0^2+3k_0^4\sum_{p=0}^{N-1}a_1^2 \tag{B3}$$

The expansion parameters can be evaluated through the following relations

$$\begin{cases}k_0^2=E\left\{y_1^2\middle|\mathbf{v}(n)\right\}=r_0\mathbf{v}^T(n)\mathbf{v}(n)\\[4pt] a_1^2=\dfrac{E\left\{y_1y_2\middle|\mathbf{v}(n)\right\}^2}{E\left\{y_1^2\right\}}=v_p^2(n)r_0^2\Big/\left[r_0\mathbf{v}^T(n)\mathbf{v}(n)\right]\\[4pt] a_2^2=E\left\{y_2^2\middle|\mathbf{v}(n)\right\}-a_1^2=r_0\dfrac{\mathbf{v}^T(n)\mathbf{v}(n)-v_p^2(n)}{\mathbf{v}^T(n)\mathbf{v}(n)}\end{cases} \tag{B4}$$

Using (B4) in (B3), removing the conditioning and approximating $E\{\mathbf{v}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{v}(n)\}$ by $E\{\mathbf{v}^T(n)\mathbf{v}(n)\}\cdot E\{\mathbf{v}^T(n)\mathbf{v}(n)\}$ results in (14).

APPENDIX C: $E\{(\mathbf{v}^T(n)\mathbf{x}(n))^4(\mathbf{x}^T(n)\mathbf{x}(n))^2\}$

Assuming three Gaussian random variables y_1, y_2 and y_3, they can be expanded as an orthonormal series [17] as

$$\begin{cases}y_1=k_0w_1\\y_2=a_1w_1+a_2w_2\\y_3=b_1w_1+b_2w_2+b_3w_3\end{cases} \tag{C1}$$

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

$$E\left\{y_1^4 y_2^2 y_3^2\right\} = \frac{105 a_1^2 b_1^2}{k_0^4} E\left\{y_1^2\right\}^4 + 3\left(3a_2^2 b_2^2 + a_2^2 b_3^2\right) E\left\{y_1^2\right\}^2 + \frac{15}{k_0^2}\left(a_1^2 b_2^2 + a_1^2 b_3^2 + a_2^2 b_1^2 + 4 a_1 a_2 b_1 b_2\right) E\left\{y_1^2\right\}^3 \tag{C2}$$

The expansion parameter's can be evaluated through the following relations

$$\begin{cases} E\left\{y_1^2\right\} = k_0^2; & E\left\{y_1 y_2\right\} = k_0 a_1; \\ E\left\{y_2^2\right\} = a_1^2 + a_2^2; & E\left\{y_1 y_3\right\} = k_0 b_1; \\ E\left\{y_3^2\right\} = b_1^2 + b_2^2 + b_3^2; & E\left\{y_2 y_3\right\} = a_1 b_1 + a_2 b_2 \end{cases} \tag{C3}$$

Substituting (C3) in (C2) yields

$$E\left\{y_1^4 y_2^2 y_3^2\right\} = 105 E\left\{y_1 y_2\right\}^2 E\left\{y_1 y_3\right\}^2 + 3k_1 + \frac{15 k_2}{E\left\{y_1^2\right\} E\left\{y_2^2\right\} - E\left\{y_1 y_2\right\}^2} \tag{C4}$$

where

$$\begin{aligned} k_1 = {} & 2E\left\{y_1^2\right\}^2 E\left\{y_2 y_3\right\}^2 + 3E\left\{y_1 y_2\right\}^2 E\left\{y_1 y_3\right\}^2 \\ & - E\left\{y_1^2\right\} E\left\{y_3^2\right\} E\left\{y_1 y_2\right\}^2 - E\left\{y_1^2\right\} E\left\{y_2^2\right\} E\left\{y_1 y_3\right\}^2 \\ & + E\left\{y_1^2\right\}^2 E\left\{y_2^2\right\} E\left\{y_3^2\right\} \\ & - 4E\left\{y_1^2\right\} E\left\{y_1 y_2\right\} E\left\{y_1 y_3\right\} E\left\{y_2 y_3\right\} \end{aligned} \tag{C5}$$

and

$$\begin{aligned} k_2 = {} & E\left\{y_1^2\right\}^2 E\left\{y_2^2\right\} E\left\{y_3^2\right\} E\left\{y_1 y_2\right\}^2 \\ & + E\left\{y_1^2\right\}^2 E\left\{y_2^2\right\}^2 E\left\{y_1 y_3\right\}^2 \\ & + 4E\left\{y_1^2\right\}^2 E\left\{y_2^2\right\} E\left\{y_1 y_2\right\} E\left\{y_1 y_3\right\} E\left\{y_2 y_3\right\} \\ & - 7E\left\{y_1^2\right\} E\left\{y_2^2\right\} E\left\{y_1 y_2\right\}^2 E\left\{y_1 y_3\right\}^2 \\ & - E\left\{y_1^2\right\} E\left\{y_3^2\right\} E\left\{y_1 y_2\right\}^4 + 6E\left\{y_1 y_2\right\}^4 E\left\{y_1 y_3\right\}^2 \\ & - 4E\left\{y_1^2\right\} E\left\{y_1 y_2\right\}^3 E\left\{y_1 y_3\right\} E\left\{y_2 y_3\right\} \end{aligned} \tag{C6}$$

Let $y_1=\mathbf{v}^T(n)\mathbf{x}(n)$, $y_2=x(n-i)$ and $y_3=x(n-j)$. Assuming independence between $\mathbf{v}(n)$ and $\mathbf{x}(n)$, the conditional expectation $E\{(\mathbf{v}^T(n)\mathbf{x}(n))^4\mathbf{x}^T(n)\mathbf{x}(n)|\mathbf{v}(n)\}$ can be evaluated through (C4). Removing the conditioning and approximating $E\{\mathbf{v}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{v}(n)\}$ by $E\{\mathbf{v}^T(n)\mathbf{v}(n)\}\cdot E\{\mathbf{v}^T(n)\mathbf{v}(n)\}$ leads to (15). Eq. (15) assumes a zero-mean weight error vector. This approximation is more valid as the algorithm approaches convergence.

REFERENCES

[1] N.V., Thakor, and Y.S. Zhu, "Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection," *IEEE Trans. on Biomedical Engineering*, vol. 38, no. 8, pp. 785-794, Aug. 1991.

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

[3] S. Koike, "A class of adaptive step-size control algorithms for adaptive filters," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1315-1326, Jun. 2002.

[4] J. Okello, Y. Itoh, Y. Fukui, I. Nakanishi, and M. Kobayashi, "A new modified variable step size for the LMS algorithm," in Proc. *Int. Simp. Circuits and Systems*, 1998, pp. 170-173.

[5] M.C. Tavares, M.C. Costa, F.A.O. Costa, and F.M. Azevedo, "Microcontroller-based multiprocessor system for ECG, EEG and MLAEP registering during surgery," in *Proc. World Congress on Medical Physics and Biomedical Engineering*, 2003, Sidney, Australia, pp. 1.

[6] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Trans. Signal Processing*, vol. 40, no. 7, pp. 1633-1642, Jul. 1992.

[7] T. Abousnar and K. Mayyas, "A robust variable step size LMS-type algorithm: analysis and simulation," *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 631-639, Mar. 1997.

[8] R. W. Harris, D. Chabries, and F. Bishop, "A variable step (VS) adaptive filter algorithm," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP34, no. 2, pp. 309-316, Apr. 1986.

[9] T. J. Shan and T. Kailath, "Adaptive algorithms with automatic gain control feature," *IEEE Trans. Circuits and Systems*, vol. 35, no. 1, pp. 122-127, Jan. 1988.

[10] Y. Wei, S.B. Gelfand, and J. V. Krogmeier, "Noise-constrained Least Mean Square algorithm," *IEEE Trans. Signal Processing*, vol. 49, no. 9, pp. 1961-1970, Sep. 2001.

[11] S. Koike, "A novel adaptive step size control algorithm for adaptive filters," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, 1999, pp. 1845-1848.

[12] C. G. Lopes and J. C. M. Bermudez, "Evaluation and design of variable step size adaptive algorithms," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, 2001, pp. 3845-3848.

[13] M. H. Costa and J. C. M. Bermudez, "A Robust variable step size algorithm for LMS adaptive filters," in *Int. Conf. Acoustics, Speech and Signal Processing*, 2006, Toulouse, France, pp. 93-96.

[14] M. H. Costa and M. C. Tavares, "Variable step size algorithm for interference canceling in biomedical signals," in *Brazilian Congress of Biomedical Engineering*, 2006, São Pedro, Brazil pp. 366-370 (*in Portuguese*).

[15] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1991.

[16] A. L.Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp.e215-e220, 2000.

[17] 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 Processing,* vol. 47, no. 3, pp. 675-689, 1999.