

Análise Estocástica do Algoritmo LMS no Domínio Transformado para Ambiente Não-estacionário

Javier Ernesto Kolodziej, Orlando José Tobias e Rui Seara

Resumo—Este trabalho apresenta um modelo estatístico para o algoritmo LMS no domínio transformado operando em um ambiente não-estacionário obtido de uma planta variante no tempo. O caso estacionário é também considerado aqui como sendo um caso particular do caso não-estacionário. A potência do sinal de entrada, requerida para a derivação do modelo estocástico do algoritmo, é estimada usando uma janela deslizante de comprimento fixo. Adicionalmente, para a derivação do modelo, são assumidos sinal de entrada gaussiano e passo de adaptação pequeno. Utilizando as expressões do modelo proposto, são determinados alguns parâmetros do algoritmo, tais como erro em excesso, desajuste e passo de adaptação para o mínimo erro em excesso. Através de simulações numéricas, a precisão do modelo proposto é avaliada.

Palavras-chave—Algoritmo TDLMS, caso não-estacionário, erro em excesso, desajuste, modelagem estatística, passo ótimo.

Abstract—This paper presents an improved statistical model for the transform-domain LMS algorithm working in a non-stationary environment arisen from a time-varying plant. The stationary case is also considered as being a particular case of the non-stationary one. The derived model considers a fixed-length sliding window for estimating the transformed input signal power. In addition, small step-size condition and Gaussian input data are assumed. By using the model expression, algorithm parameters such as excess error, misadjustment, and step size for minimum excess error are obtained. Through numerical simulations, the accuracy of the proposed model is assessed.

Keywords—TDLMS algorithm, non-stationary case, excess error, misadjustment, statistical modeling, optimum step size.

I. INTRODUÇÃO

O algoritmo LMS no domínio transformado (TDLMS) foi proposto por Narayan et al. [1] objetivando melhorar as características de convergência do algoritmo LMS padrão. O algoritmo TDLMS é similar ao LMS, porém tendo seu sinal de entrada pré-processado por uma transformação ortogonal de N sub-bandas, seguido de um processo de normalização que equaliza o conteúdo energético em cada sub-banda. Na prática, para implementar o processo de normalização, são necessárias estimativas da potência do sinal de entrada em cada sub-banda.

Javier E. Kolodziej e Rui Seara, LINSE – Laboratório de Circuitos e Processamento de Sinais, Depto. de Eng. Elétrica, Universidade Federal de Santa Catarina, Florianópolis, SC, E-mails: {javier, seara}@linse.ufsc.br.

Orlando J. Tobias, LINSE/UFSC e Depto. Eng. Elétrica e Telecom., Universidade Regional de Blumenau, SC, E-mail: tobias.oj@ieee.org.

Este trabalho foi parcialmente financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Para tal, é utilizada a medida da variância do sinal de entrada através de uma janela deslizante de comprimento fixo (JDCF).

Com vistas à modelagem, devido ao processo de normalização, uma importante dificuldade é, especificamente, o cálculo de valores esperados, tais como

$$E[\hat{\mathbf{D}}^{-1}(n)\hat{\mathbf{R}}(n)] \quad (1)$$

onde $\hat{\mathbf{D}}(n)$ e $\hat{\mathbf{R}}(n)$ são matrizes que dependem dos dados observados. Para resolver (1), várias aproximações simplificativas são geralmente consideradas [2]-[6]. Em particular, destaca-se o Princípio da Média (PM), utilizado em [5] para aproximar (1). Assim,

$$E[\hat{\mathbf{D}}^{-1}(n)\hat{\mathbf{R}}(n)] \approx E[\hat{\mathbf{D}}^{-1}(n)]E[\hat{\mathbf{R}}(n)]. \quad (2)$$

O PM resulta em expressões simples, apresentando precisão satisfatória se grandes janelas de observação forem utilizadas. No entanto, tal condição pode não ser encontrada em situações práticas, sendo que ainda muitos algoritmos apresentam restrições em relação ao custo computacional. Assim, procurando por um modelo preciso, independentemente do comprimento da janela de observação, o uso do PM não é o mais recomendado.

Em particular, o modelo proposto neste trabalho considera os seguintes aspectos:

- i) Valores esperados semelhantes à (1) são calculados sem invocar o PM.
- ii) Objetivando generalidade, o modelo proposto é derivado considerando ambientes não-estacionários (planta variante no tempo). Essa situação tem grande importância prática e na literatura técnica só umas poucas e breves análises têm sido encontradas. Além do mais, o caso estacionário pode ser diretamente obtido a partir das expressões propostas.

Expressões para o erro em excesso, desajuste e passo de adaptação para o mínimo erro em excesso são também derivadas a partir da análise proposta. Este trabalho está organizado como segue. A Seção II apresenta o modelo utilizado para caracterizar a planta variante no tempo. Na Seção III, expressões do modelo do algoritmo são derivadas. Resultados de simulação numérica são apresentados na Seção IV, demonstrando a validade do modelo proposto. Finalmente, na Seção V, são apresentadas as conclusões deste trabalho.

II. AMBIENTE NÃO-ESTACIONÁRIO

Uma característica fundamental dos filtros adaptativos é a sua habilidade de rastrear as variações do sinal. Para realizar uma análise tratável matematicamente, é usual assumir que os dados estatísticos (para o nosso caso, planta variante) variam de acordo com uma dada regra. Dessa forma, o sinal desejado é modelado como

$$d(n) = \mathbf{x}_T^T(n)\mathbf{w}_T^o(n) + \vartheta(n) \quad (3)$$

com a seguinte regra para descrever a evolução da planta:

$$\mathbf{w}_T^o(n+1) = a\mathbf{w}_T^o(n) + \boldsymbol{\eta}(n) \quad (4)$$

onde $\mathbf{x}_T(n) = \mathbf{T}\mathbf{x}(n) = [x_{T,0}(n) \ x_{T,1}(n) \ \dots \ x_{T,N-1}(n)]^T$ é o vetor do sinal de entrada no domínio transformado, \mathbf{T} denota uma matriz de transformação ortogonal, $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-N+1)]^T$ é o vetor do sinal de entrada, assumindo $\{x(n)\}$ um processo estacionário gaussiano de média zero e variância σ_x^2 , $\mathbf{w}_T^o(n) = \mathbf{T}\mathbf{w}^o(n) = [w_{T,1}^o(n) \ w_{T,2}^o(n) \ \dots \ w_{T,N-1}^o(n)]^T$ é o vetor da planta variante no tempo no domínio transformado e $\mathbf{w}^o(n) = [w_1^o(n) \ w_2^o(n) \ \dots \ w_{N-1}^o(n)]^T$, o vetor da planta variante no tempo. O ruído de medição $\vartheta(n)$ tem média zero e variância σ_ϑ^2 . A variável $a \in [0, 1]$. O vetor $\boldsymbol{\eta}(n)$ representa o processo de perturbação da planta, tendo média zero e matriz de autocorrelação Φ , e $\mathbf{w}_T^o(0)$ é um valor arbitrário.

Usualmente, análises de algoritmos adaptativos em ambientes não-estacionários disponíveis na literatura utilizam $a=1$, e/ou $\mathbf{w}_T^o(0) = \mathbf{0}$. Assim, para ambos os casos, tem-se um modelo restrito, sendo que o valor médio da planta não se modifica com o tempo. Além do mais, o modelo dado em (4) para $a=1$ é irreal, visto que ele apresenta um crescimento irrestrito da variância [7]. Neste trabalho, as derivações obtidas consideram um valor genérico para a , resultando em um modelo mais geral.

III. ANÁLISE

A. Formulação do Problema

Nesta seção, são derivadas expressões analíticas para os momentos de primeira e segunda ordem do vetor de pesos adaptativos. Para tal, iniciamos a análise considerando a equação de atualização do vetor de pesos no domínio transformado dada por

$$\mathbf{w}_T(n+1) = \mathbf{w}_T(n) + 2\mu\hat{\mathbf{D}}^{-1}(n)e(n)\mathbf{x}_T(n) \quad (5)$$

onde $\mathbf{w}_T(n) = [w_{T,0}(n) \ w_{T,1}(n) \ \dots \ w_{T,N-1}(n)]^T$ é o vetor de pesos do filtro adaptativo e $\hat{\mathbf{D}}(n) = \text{diag}[\hat{\sigma}_0^2(n) \ \hat{\sigma}_1^2(n) \ \dots \ \hat{\sigma}_{N-1}^2(n)]$ é a matriz de normalização do passo de adaptação, com $\hat{\sigma}_i^2(n)$ sendo a estimativa instantânea da variância de uma sub-banda. Na prática, uma JDCF é utilizada para estimar a variância de cada sub-banda. Assim,

$$\hat{\sigma}_i^2(n) = \frac{1}{M} \sum_{k=0}^{M-1} x_{T,i}^2(n-k) \quad (6)$$

onde M é o comprimento da janela de observação. O sinal de erro é

$$e(n) = d(n) - y(n) \quad (7)$$

com $y(n) = \mathbf{x}_T^T(n)\mathbf{w}_T(n)$. Agora, substituindo (3) em (7) e a expressão resultante em (5), utilizando o vetor de erro nos coeficientes no domínio transformado, definido como $\mathbf{v}_T(n) = \mathbf{w}_T(n) - \mathbf{w}_T^o(n)$, obtém-se

$$\begin{aligned} \mathbf{v}_T(n+1) = & [\mathbf{I} - 2\mu\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\mathbf{v}_T(n) \\ & + 2\mu\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\vartheta(n) - \mathbf{w}_T^o(n+1) + \mathbf{w}_T^o(n). \end{aligned} \quad (8)$$

Agora, substituindo (4) em (8), a equação de atualização em termos do $\mathbf{v}_T(n)$ é dada por

$$\begin{aligned} \mathbf{v}_T(n+1) = & [\mathbf{I} - 2\mu\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\mathbf{v}_T(n) - \boldsymbol{\eta}(n) \\ & + 2\mu\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\vartheta(n) + (1-a)\mathbf{w}_T^o(n). \end{aligned} \quad (9)$$

O passo seguinte é então determinar os momentos de primeira e segunda ordem de (9).

B. Considerações de Análise

Para realizar a análise estocástica, as seguintes considerações são assumidas:

- i) $\boldsymbol{\eta}(n)$ e $\boldsymbol{\eta}(m)$, para $m \neq n$, são descorrelacionados.
- ii) $\mathbf{v}_T(n)$ e $\mathbf{x}_T(n)$ são estatisticamente independentes entre si, visto que o passo de adaptação é assumido pequeno.
- iii) $\boldsymbol{\eta}(n)$ e $\vartheta(n)$ são descorrelacionados com qualquer outro sinal no sistema.

C. Momento de Primeira Ordem de $\mathbf{v}_T(n)$

Tomando o valor esperado de ambos os lados de (9) e levando em conta as suposições (i)-(iii), obtém-se

$$\begin{aligned} E[\mathbf{v}_T(n+1)] \cong & \{ \mathbf{I} - 2\mu \underbrace{E[\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]}_{\mathbf{P}} \} E[\mathbf{v}_T(n)] \\ & + (1-a)a^n \mathbf{w}_T^o(0) \end{aligned} \quad (10)$$

onde o último termo do lado direito representa a média do processo AR(1) dado por (4). Neste trabalho, os elementos da matriz \mathbf{P} são determinados sem utilizar o PM como é usualmente considerado na literatura [5], para detalhes veja o Apêndice I.

D. Momento de Segunda Ordem de $\mathbf{v}_T(n)$ e Curva de Aprendizagem

O momento de segunda ordem do vetor de erro nos pesos no domínio transformado é obtido a partir de $\mathbf{K}(n) = E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]$. Assim, transpondo ambos os lados de (9), efetuando o produto $\mathbf{v}_T(n)\mathbf{v}_T^T(n)$, tomando o valor esperado de ambos os lados da expressão resultante, e utilizando as suposições (i)-(iii), obtém-se

$$\begin{aligned} \mathbf{K}(n+1) = & \mathbf{K}(n) - 2\mu\mathbf{K}(n)\mathbf{P}^T - 2\mu\mathbf{P}\mathbf{K}(n) + 4\mu^2\mathbf{S}\sigma_\vartheta^2 + (1-a)^2\mathbf{K}^o(n) \\ & + (1-a)E[\mathbf{w}_T^o(n)]E[\mathbf{v}_T^T(n)](\mathbf{I} - 2\mu\mathbf{P}^T) + 8\mu^2\mathbf{P}\mathbf{K}(n)\mathbf{P}^T + \Phi \\ & + (1-a)(\mathbf{I} - 2\mu\mathbf{P})E[\mathbf{v}_T(n)]E[\mathbf{w}_T^{oT}(n)] + 4\mu^2\text{tr}[\mathbf{K}(n)\mathbf{R}_T] \end{aligned} \quad (11)$$

com

$$\mathbf{K}^o(n) = E[\mathbf{w}_T^o(n)\mathbf{w}_T^{oT}(n)] = (a^2)^n \mathbf{K}^o(0) + \frac{1-(a^2)^n}{1-a^2} \Phi \quad (12)$$

e

$$\mathbf{S} = E[\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\hat{\mathbf{D}}^{-1}(n)]. \quad (13)$$

Um método de cálculo aproximado para \mathbf{S} é apresentado no Apêndice II. Tal cálculo é mais preciso do que outras abordagens encontradas na literatura [5].

Finalmente, (11) é utilizada para determinar a curva de aprendizagem, dada por [7]

$$E[e^2(n)] = \sigma_\vartheta^2 + \text{tr}\{\mathbf{R}_T\mathbf{K}(n)\}. \quad (14)$$

E. Erro em Excesso

Nesta seção, uma expressão para o erro em excesso é obtida. Assim, transpondo ambos os lados de (9), efetuando o produto $\hat{\mathbf{D}}(n)\mathbf{v}_T(n+1)\mathbf{v}_T^T(n+1)$, tomando o valor esperado da expressão resultante, assumindo $n \rightarrow \infty$ (e portanto as aproximações $\mathbf{K}(n+1) \cong \mathbf{K}(n)$, $E[\mathbf{v}_T(n)] \rightarrow 0$ e $\mathbf{K}^o(n) \rightarrow (1-a^2)^{-1}\Phi$ podem ser utilizadas) e, finalmente, tomando o traço da matriz resultante, obtém-se

$$\begin{aligned} \text{tr}\{\mathbf{K}(n)\mathbf{R}_T\} &= \mu\sigma_\vartheta^2 \text{tr}\{\mathbf{P}\} + \frac{1}{2\mu(1+a)} \text{tr}\{E[\hat{\mathbf{D}}(n)]\Phi\} \\ &+ \underbrace{\mu \text{tr}\{\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\hat{\mathbf{D}}^{-1}(n)\}}_{\gamma}. \end{aligned} \quad (15)$$

Agora, rearranjando γ como

$$\begin{aligned} \gamma &= \mathbf{x}_T^T(n)\hat{\mathbf{D}}^{-1}(n)\mathbf{x}_T(n)\text{tr}\{\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\} \\ &= \text{tr}\{\mathbf{P}\}\text{tr}\{\mathbf{R}_T\mathbf{K}(n)\} \end{aligned} \quad (16)$$

e considerando a definição do erro em excesso [7]

$$\xi_{\text{exc}}(n) = \text{tr}\{\mathbf{R}_T\mathbf{K}(n)\} \quad (17)$$

então, para $n \rightarrow \infty$, obtém-se

$$\xi_{\text{exc}} = \frac{1}{1-\mu\text{tr}\{\mathbf{P}\}} \left\{ \mu\sigma_\vartheta^2 \text{tr}\{\mathbf{P}\} + \frac{1}{2\mu(1+a)} \text{tr}\{E[\hat{\mathbf{D}}(n)]\Phi\} \right\}. \quad (18)$$

F. Passo de Adaptação para o Mínimo Erro em Excesso

O passo de adaptação que minimiza o erro quadrático médio (EQM) em excesso é obtido derivando (18) em relação ao passo μ e igualando o resultado a zero. Dessa forma, tem-se

$$\frac{\partial \xi_{\text{exc}}}{\partial \mu} = \frac{\sigma_\vartheta^2 c_1 \mu^2 + 2c_1 c_2 \mu - c_2}{(1-\mu c_1)^2} = 0 \quad (19)$$

onde $c_1 = \text{tr}\{\mathbf{P}\}$ e $c_2 = \frac{1}{2(1+a)} \text{tr}\{E[\hat{\mathbf{D}}(n)]\Phi\}$. Resolvendo (19)

para μ , obtém-se

$$\mu_{\text{opt}} = \frac{-c_1 c_2 + \sqrt{c_1^2 c_2^2 + \sigma_\vartheta^2 c_1 c_2}}{\sigma_\vartheta^2 c_1}. \quad (20)$$

G. Desajuste

O desajuste \mathcal{M} é obtido a partir de (18), o qual é dado por

$$\mathcal{M} = \frac{\xi_{\text{exc}}}{\xi_{\text{min}}} = \frac{1}{1-\mu\text{tr}\{\mathbf{P}\}} \left\{ \mu\text{tr}\{\mathbf{P}\} + \frac{\sigma_\vartheta^{-2}}{2\mu(1+a)} \text{tr}\{E[\hat{\mathbf{D}}(n)]\Phi\} \right\} \quad (21)$$

onde $\xi_{\text{min}} = \sigma_\vartheta^2$. Note que (21) [também (18)] é composto pela soma de dois termos. O primeiro é igual ao desajuste no caso estacionário (Φ é uma matriz nula), o segundo é o desajuste originado da não-estacionaridade. Assim, o desajuste para o caso não-estacionário é maior do que aquele observado no caso estacionário.

H. Grau de Não-estacionaridade

O grau de não-estacionaridade, denotado por $\psi(n)$, é definido como [7]

$$\psi(n) \triangleq \left\{ \frac{E\left[|y_{o,\text{inc}}(n)|^2\right]}{E\left[|\vartheta(n)|^2\right]} \right\}^{1/2} \quad (22)$$

onde

$$y_{o,\text{inc}}(n) = [\mathbf{w}_T^o(n+1) - \mathbf{w}_T^o(n)]^T \mathbf{x}_T(n) \quad (23)$$

é a saída devido à diferença $\mathbf{w}_T^o(n+1) - \mathbf{w}_T^o(n)$ (filtro incremental). O numerador de (22) representa a potência introduzida pela variação do filtro ótimo e o seu denominador denota o EQM mínimo. Utilizando (4) em (23), obtém-se

$$\begin{aligned} y_{o,\text{inc}}(n) &= [(a-1)\mathbf{w}_T^o(n) + \boldsymbol{\eta}(n)]^T \mathbf{x}_T(n) = (a-1)[a^n \mathbf{w}_T^o(0) \\ &+ \sum_{k=0}^{n-1} a^k \boldsymbol{\eta}(n-1-k)]^T \mathbf{x}_T(n). \end{aligned} \quad (24)$$

Levando em conta a independência entre $\boldsymbol{\eta}(n)$ e $\mathbf{x}_T(n)$ e a assunção (i), pode-se então escrever

$$\begin{aligned} E\left[|y_{o,\text{inc}}(n)|^2\right] &= (1-a)^2 a^{2n} \text{tr}\{\mathbf{K}^o(0)\mathbf{R}_T\} \\ &+ \frac{2-a^{2n}+a^{2n+1}}{1+a} \text{tr}\{\Phi\mathbf{R}_T\}. \end{aligned} \quad (25)$$

Agora, substituindo (25) em (22) e lembrando que σ_ϑ^2 é a variância de $\vartheta(n)$, tem-se

$$\psi(n) = \left\{ \frac{(1-a)^2 a^{2n}}{\sigma_\vartheta^2} \text{tr}\{\mathbf{K}^o(0)\mathbf{R}_T\} + \frac{2-a^{2n}+a^{2n+1}}{\sigma_\vartheta^2(1+a)} \text{tr}\{\Phi\mathbf{R}_T\} \right\}^{1/2}. \quad (26)$$

De (26), pode-se concluir que $\psi(n)$ apresenta uma evolução exponencial, iniciando pelo valor máximo

$$\psi(0) = \left\{ \frac{(1-a)^2}{\sigma_\vartheta^2} \text{tr}\{\mathbf{K}^o(0)\mathbf{R}_T\} + \frac{1}{\sigma_\vartheta^2} \text{tr}\{\Phi\mathbf{R}_T\} \right\}^{1/2} \quad (27)$$

e convergindo assintoticamente para

$$\psi(\infty) = \frac{\sqrt{2}}{\sigma_\vartheta \sqrt{1+a}} \left\{ \text{tr}\{\Phi\mathbf{R}_T\} \right\}^{1/2}. \quad (28)$$

Então, se o grau de não-estacionaridade é maior do que a unidade, as variações estatísticas do vetor de pesos ótimo são muito rápidas para o filtro poder rastreá-las (e o desajuste será grande); caso contrário, o filtro adaptativo terá habilidade para seguir as variações da planta. Neste trabalho, é avaliado o comportamento do algoritmo para a última situação, ou seja, quando o rastreamento é possível.

IV. RESULTADOS DE SIMULAÇÃO

Nesta seção, o modelo proposto é aplicado a um problema de identificação de sistema visando avaliar sua precisão para sinais de entrada gaussiano correlacionado, obtido a partir de um processo auto-regressivo AR(2) dado por

$$x(n) = a_1 x(n-1) + a_2 x(n-2) + v(n) \quad (29)$$

onde $v(n)$ é um ruído branco com variância θ_v^2 , o qual é ajustado para que a variância de $x(n)$ seja unitária, $a_1 = 0,18$ e $a_2 = -0,85$ são os coeficientes do processo AR(2). A dispersão

dos autovalores do sinal de entrada para $N=8$ é $\chi=46,1$. Os pesos da planta variante são obtidos de acordo com (4). Seus valores iniciais são selecionados como

$$\mathbf{w}_T^o(0) = \frac{\mathbf{w}_{aux}}{\sqrt{\mathbf{w}_{aux}^T \mathbf{w}_{aux}}} \quad (30)$$

com o vetor $\mathbf{w}_{aux} = \mathbf{T}[\text{sinc}(0) \text{ sinc}(1/N) \dots \text{sinc}(N-1/N)]^T$. Os elementos de $\boldsymbol{\eta}(n)$ são amostras de um ruído branco com matriz de autocorrelação $\boldsymbol{\Phi} = \sigma_\eta^2 \mathbf{I}$. Nos exemplos, são usados $a=0,999$ e dois níveis de ruído aditivo, $\theta_\eta^2 = -40\text{dB}$ e $\theta_\eta^2 = -60\text{dB}$, resultando, respectivamente, em um grau de não-estacionaridade de $\psi(0)=0,2075$ e $\psi(0)=0,588$, para cada nível de ruído considerado. Além disso, $\psi(n)=0,2$ para $n \rightarrow \infty$ nas duas situações. A transformada do cosseno discreta (TCD) é a transformada ortogonal aqui utilizada. Simulações Monte Carlo (MC) são obtidas considerando 500 realizações independentes.

Para os casos não-estacionários, os pesos são inicializados como

$$\mathbf{w}_T(0) = \mathbf{w}_T^o(0) \quad (31)$$

visando avaliar o comportamento de rastreamento do algoritmo. No caso estacionário, os pesos são inicializados com zeros para se avaliar o comportamento do algoritmo no modo de aquisição [7]. Para propósitos de comparação, o comportamento do algoritmo previsto por um modelo baseado no PM é também apresentado. Esse modelo é obtido generalizando (para o caso não-estacionário) o procedimento apresentado em [5].

A. Exemplo 1

Neste exemplo, os seguintes valores para os parâmetros são utilizados: $N=8$, $M=8$ e $\mu=0,1\mu_{opt}=0,00257$. Na Fig. 1, o momento de primeira ordem do vetor de erro nos coeficientes para $\theta_\eta^2 = -40\text{dB}$ [veja Fig. 1(a)] e $\theta_\eta^2 = -60\text{dB}$ [veja Fig. 1(b)] é mostrado. Para uma melhor visualização, somente quatro coeficientes são mostrados. A Fig. 2(a) apresenta a curva de aprendizagem (EQM) para os dois níveis de ruído considerados. Para esse exemplo, verifica-se um casamento muito bom entre simulação MC e modelo proposto.

B. Exemplo 2

Neste exemplo, a maior parte dos parâmetros do Exemplo 1 são mantidos, exceto M , que é aumentado para $M=64$, visando obter uma melhor estimativa da variância. Os resultados obtidos para a evolução do EQM são mostrados na Fig. 2(b). Note que, nesse caso, o modelo baseado no PM e o modelo proposto apresentam quase o mesmo comportamento, visto que agora o valor de M favorece a aproximação do PM.

C. Exemplo 3

Neste exemplo, o caso estacionário é considerado como caso particular do caso não-estacionário ($a=1$ e grau de não-estacionaridade $\psi=0$). Os parâmetros utilizados são: $N=8$, $M=8$ e $\mu=0,00257$. Os resultados obtidos são mostrados na Fig 2(c). Note que agora existe uma redução no valor do EQM, visto que o termo correspondente a não-estacionaridade em (18) é zero.

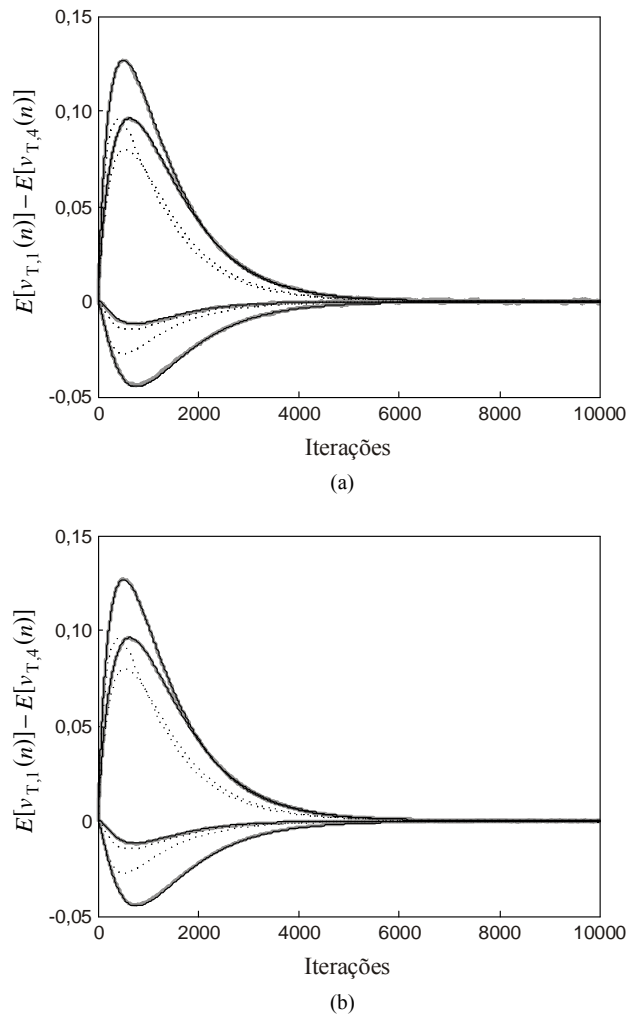


Fig. 1. Exemplo 1. Curvas do comportamento médio do vetor de erro $E[v_{T,i}(n)]$, para $i=0, 1, 2, 3$: (a) $\sigma_\eta^2 = -40\text{dB}$, (b) $\sigma_\eta^2 = -60\text{dB}$. (Linhas cinza) simulação MC. (Linhas tracejadas escuras) modelo baseado no PM. (Linhas contínuas escuras) modelo proposto.

D. Exemplo 4

Neste exemplo, resultados de simulação MC para o erro em excesso são comparados com os previstos a partir de (18). A mesma entrada correlacionada dos exemplos anteriores, com $\sigma_\eta^2 = -40\text{dB}$ e $M=8$, é utilizada. A Fig. 3 ilustra os resultados obtidos, mostrando a boa predição encontrada a partir de (18).

V. CONCLUSÕES

Este trabalho apresentou um modelo estatístico para o algoritmo TDLMS operando em ambiente não-estacionário. A análise realizada é independente da ordem do filtro e do comprimento da janela de observação usada para estimar a potência do sinal em cada sub-banda. Expressões para o erro em excesso, desajuste e passo de adaptação para o mínimo erro em excesso foram também derivadas. Adicionalmente, o caso estacionário também pode ser obtido a partir do modelo desenvolvido. Resultados de simulação MC confirmam uma muito boa precisão do modelo proposto.

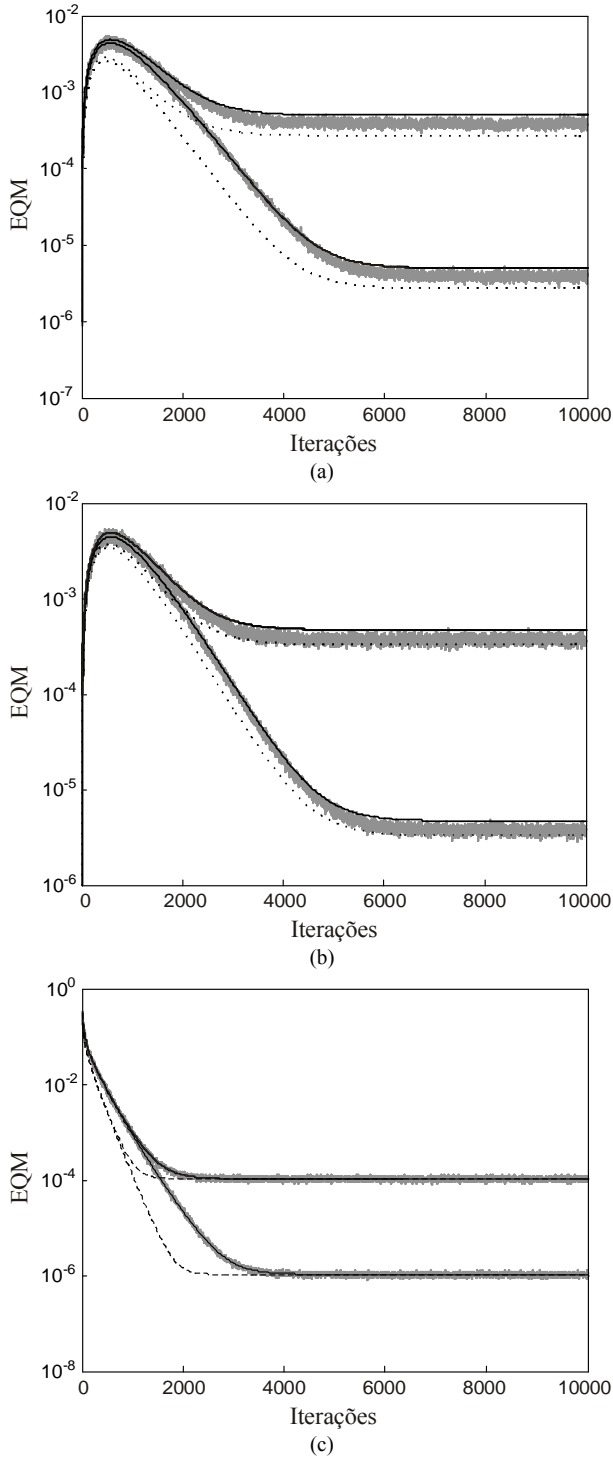


Fig. 2. Resultados de comparação para um ambiente estacionário. Curvas do EQM para $\sigma_{\theta}^2 = -40\text{dB}$ e $\sigma_{\theta}^2 = -60\text{dB}$. (a) Exemplo 1. (b) Exemplo 2. (c) Exemplo 3. (Linhas cinza) simulação MC. (Linhas tracejadas escuras) modelo baseado no PM. (Linhas contínuas escuras) modelo proposto.

APÊNDICE I - CÁLCULO DE \mathbf{P}

Para determinar \mathbf{P} , é definido um vetor estendido dado por $\mathbf{x}_{e,i,j}(n) = [x_{T,i}(n) \ x_{T,i}(n-1) \ \dots \ x_{T,i}(n-M+1) \ x_{T,j}(n)]^T$, tal que

$$\sum_{k=0}^{M-1} x_{T,i}^2(n-k) = \mathbf{x}_{e,i,j}^T(n) \mathbf{I}_s \mathbf{x}_{e,i,j}(n) \quad (32)$$

onde $\mathbf{I}_s = \text{diag}[1 \ \dots \ 1 \ 0]$. Considerando processos conjuntamente gaussianos, \mathbf{P} é calculada como

$$\{\mathbf{P}\}_{i,j} = \frac{M}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{[\det(\mathbf{R}_{e,i,j})]^{1/2}} \frac{x_{T,i} x_{T,j}}{\mathbf{x}_{e,i,j}^T \mathbf{I}_s \mathbf{x}_{e,i,j}} e^{-\frac{1}{2} \mathbf{x}_{e,i,j}^T \mathbf{R}_{e,i,j}^{-1} \mathbf{x}_{e,i,j}} d\mathbf{x}_{e,i,j} \quad (33)$$

onde $\mathbf{R}_{e,i,j} = E[\mathbf{x}_{e,i,j} \mathbf{x}_{e,i,j}^T]$ é a matriz de autocorrelação do vetor estendido $\mathbf{x}_{e,i,j}$. Para determinar (33), utiliza-se um procedimento similar ao apresentado em [8]. Assim, uma função auxiliar $f_{i,j}(\omega)$ é definida como

$$f_{i,j}(\omega) = \frac{M}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{[\det(\mathbf{R}_{e,i,j})]^{1/2}} \frac{x_{T,i} x_{T,j}}{\mathbf{x}_{e,i,j}^T \mathbf{I}_s \mathbf{x}_{e,i,j}} e^{-\frac{1}{2} \mathbf{x}_{e,i,j}^T \mathbf{L}_{i,j}^{-1}(\omega) \mathbf{x}_{e,i,j}} d\mathbf{x}_{e,i,j} \quad (34)$$

com

$$\mathbf{L}_{i,j}(\omega) = (2\omega \mathbf{I}_s + \mathbf{R}_{e,i,j}^{-1})^{-1}.$$

Note que para $\omega = 0$, (34) é igual ao valor esperado procurado. Assim,

$$\{\mathbf{P}\}_{i,j} = f_{i,j}(0). \quad (35)$$

Esta função auxiliar é utilizada para transformar o problema de resolução das $M+1$ integrais de (33) em um problema visando resolver uma única integral definida. Assim, aplicando derivadas parciais em (34) com respeito a ω , o termo $\mathbf{x}_{e,i,j}^T \mathbf{I}_s \mathbf{x}_{e,i,j}$ no denominador de seu integrando é eliminado, resultando em

$$\frac{\partial f_{i,j}(\omega)}{\partial \omega} = -\frac{M \{\det[\mathbf{L}_{i,j}(\omega)]\}^{-1/2}}{[\det(\mathbf{R}_{e,i,j})]^{1/2}} \alpha(\omega) \quad (36)$$

com

$$\alpha(\omega) = \frac{1}{\sqrt{2\pi} \{\det[\mathbf{L}_{i,j}(\omega)]\}^{1/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_{T,i} x_{T,j} e^{-\frac{1}{2} \mathbf{x}_{e,i,j}^T \mathbf{L}_{i,j}^{-1}(\omega) \mathbf{x}_{e,i,j}} d\mathbf{x}_{e,i,j} \quad (37)$$

A expressão (37) pode ser entendida como sendo a correlação cruzada entre $x_{T,i}$ e $x_{T,j}$, quando $x_{T,i}$ e $x_{T,j}$ são variáveis conjuntamente gaussianas com matriz de covariância $\mathbf{L}_{i,j}(\omega)$. Assim, o fator $\alpha(\omega)$ representa o $(1, M+1)$ -ésimo elemento da matriz $\mathbf{L}_{i,j}(\omega)$. Assim,

$$\frac{\partial \{\mathbf{F}(\omega)\}_{i,j}}{\partial \omega} = -\frac{M[(2\omega \mathbf{R}_{e,i,j} \mathbf{I}_s + \mathbf{I})^{-1} \mathbf{R}_{e,i,j}]_{1,M+1}}{[\det(2\omega \mathbf{R}_{e,i,j} \mathbf{I}_s + \mathbf{I})]^{1/2}}. \quad (38)$$

Então, considerando

$$\int_0^{\infty} \frac{\partial}{\partial \omega} f_{i,j}(\omega) d\omega = f_{i,j}(\infty) - f_{i,j}(0) \quad (39)$$

e que $f_{i,j}(\infty) = 0$, obtém-se

$$\{\mathbf{P}\}_{i,j} = \left[M \mathbf{Q}_{s,i,j} \int_0^{\infty} \frac{(2\omega \mathbf{\Lambda}_{s,i,j} \mathbf{I}_s + \mathbf{I})^{-1}}{[\det(2\omega \mathbf{\Lambda}_{s,i,j} \mathbf{I}_s + \mathbf{I})]^{1/2}} d\omega \mathbf{Q}_{s,i,j}^{-1} \mathbf{R}_{e,i,j} \right]_{1,M+1} \quad (40)$$

onde $\mathbf{Q}_{s,i,j}$ é a matriz contendo os autovetores de $\mathbf{R}_{e,i,j}\mathbf{I}_s$ e $\mathbf{\Lambda}_{s,i,j}$ é uma matriz diagonal contendo seus autovalores. Da estrutura de $\mathbf{R}_{e,i,j}\mathbf{I}_s$, pode ser inferido que se tem um autovalor igual a zero e os outros M são coincidentes com os autovalores de $\mathbf{R}_{T,i} = E[\mathbf{x}_{T,i}(n)\mathbf{x}_{T,i}^T(n)]$, onde $\mathbf{x}_{T,i}(n) = [x_{T,i}(n) \ x_{T,i}(n-1) \ \dots \ x_{T,i}(n-M+1)]^T$. Assim, podemos definir uma matriz diagonal \mathbf{H}_i com elementos dados por

$$\{\mathbf{H}_i\}_{l,l} = \begin{cases} \int_0^\infty \frac{1}{(1+2\omega\lambda_{i,l})[\prod_{k=1}^M (2\omega\lambda_{i,k}+1)]^{1/2}} d\omega, & l \neq M+1 \\ \int_0^\infty \frac{1}{[\prod_{k=1}^M (2\omega\lambda_{i,k}+1)]^{1/2}} d\omega, & l = M+1 \end{cases} \quad (41)$$

tal que (33) pode ser reescrita como

$$\{\mathbf{P}\}_{i,j} = M \left[\mathbf{Q}_{s,i,j} \mathbf{H}_i \mathbf{Q}_{s,i,j}^{-1} \mathbf{R}_{e,i,j} \right]_{1,M+1}. \quad (42)$$

Em (41), $\lambda_{i,k}$, para $k=1, \dots, M$, são os autovalores da matriz $\mathbf{R}_{T,i}$. Para determinar (41), uma integral hiperelíptica de alta ordem deve ser resolvida (ver [9] por detalhes).

APÊNDICE II - DETERMINAÇÃO DE S

De acordo com o método proposto para determinar \mathbf{S} , são calculados de forma diferente os elementos da diagonal principal e os elementos fora da diagonal. Para calcular os elementos da diagonal de \mathbf{S} , um procedimento similar ao apresentado no Apêndice I é considerado. Assim, utilizando a definição formal do valor esperado de processos conjuntamente gaussianos, tem-se

$$\{\mathbf{S}\}_{i,i} = \frac{M^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{[\det(\mathbf{R}_{T,i})]^{1/2}} \frac{x_{T,i}^2}{(\mathbf{x}_{T,i}^T \mathbf{x}_{T,i})^2} e^{-\frac{1}{2} \mathbf{x}_{T,i}^T \mathbf{R}_{T,i}^{-1} \mathbf{x}_{T,i}} d\mathbf{x}_{T,i}. \quad (43)$$

Aqui também é definida uma função auxiliar $g_i(\omega)$ como

$$g_i(\omega) = \frac{M^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{[\det(\mathbf{R}_{T,i})]^{1/2}} \frac{x_{T,i}^2}{(\mathbf{x}_{T,i}^T \mathbf{x}_{T,i})^2} e^{-\frac{1}{2} \mathbf{x}_{T,i}^T \mathbf{M}_i^{-1}(\omega) \mathbf{x}_{T,i}} d\mathbf{x}_{T,i} \quad (44)$$

com

$$\mathbf{M}_i(\omega) = (2\omega\mathbf{I} + \mathbf{R}_{T,i}^{-1})^{-1}.$$

Note que

$$\{\mathbf{S}\}_{i,i} = g_i(0). \quad (45)$$

Agora, aplicando derivadas parciais em (44) com respeito à variável ω duas vezes, o termo $(\mathbf{x}_{T,i}^T \mathbf{x}_{T,i})^2$ que é o denominador de $g_i(\omega)$ é eliminado. Em conseqüência, o fator $\beta(\omega)$ representa o $(1,1)$ -ésimo elemento da matriz $\mathbf{M}_i(\omega)$. Dessa forma,

$$\frac{\partial^2 \{\mathbf{g}(\omega)\}_i}{\partial \omega^2} = \frac{M^2 [(2\omega\mathbf{R}_{T,i} + \mathbf{I})^{-1} \mathbf{R}_{T,i}]_{1,1}}{[\det(2\omega\mathbf{R}_{T,i} + \mathbf{I})]^{1/2}}.$$

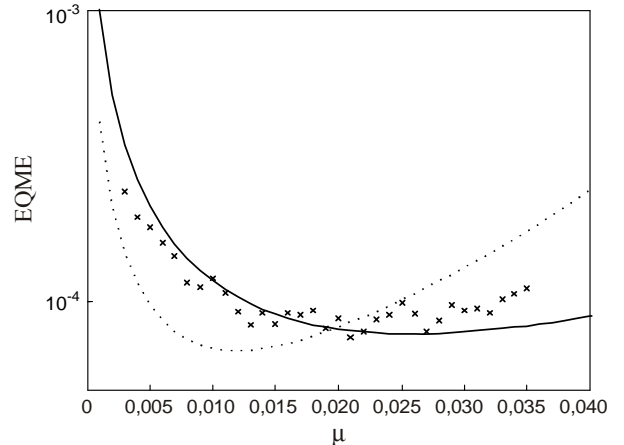


Fig. 3. Exemplo 4. Resultados de comparação para o erro em excesso. (Valores marcados com \times) simulação MC. (Linhas tracejadas escuras) modelo baseado no PM. (Linhas contínuas escuras) modelo proposto.

Assim, considerando

$$E \left[\frac{M^2 x_{T,i}^2}{(\mathbf{x}_{T,i}^T \mathbf{x}_{T,i})^2} \right] = \int_{\omega_2=0}^{\infty} \int_{\omega_1=0}^{\infty} \frac{M^2 [(2\omega_1 \mathbf{R}_{T,i} + \mathbf{I})^{-1} \mathbf{R}_{T,i}]_{1,1}}{[\det(2\omega_1 \mathbf{R}_{T,i} + \mathbf{I})]^{1/2}} d\omega_1 d\omega_2 \quad (46)$$

e utilizando uma matriz auxiliar $\mathbf{U}_{i,i}$, os elementos da diagonal de \mathbf{S} podem ser calculados como

$$\{\mathbf{S}\}_{i,i} = M^2 [\mathbf{Q}_{T,i} \mathbf{U}_{i,i} \mathbf{Q}_{T,i}^T]_{1,1} \quad (47)$$

onde

$$\{\mathbf{U}_{i,i}\}_{l,l} \approx \frac{-1}{4\sqrt{a_i}} \left\{ \sum_{q=1}^{M/2} \frac{A_{i,l,q}}{\lambda'_{i,q}} [1 + \ln(2\lambda'_{i,q})] + \frac{B_{i,l}}{\lambda_{i,l}} [1 + \ln(2\lambda_{i,l})] \right\}. \quad (48)$$

Finalmente, calculamos os elementos fora da diagonal principal de \mathbf{S} , considerando primeiro o PM e em seguida que as variâncias estimadas são não-correlacionadas umas com as outras. Então,

$$\{\mathbf{S}\}_{i,j} = E \left[\frac{x_{T,i} x_{T,j}}{\hat{\sigma}_i^2 \hat{\sigma}_j^2} \right] \approx E \left[\frac{1}{\hat{\sigma}_i^2} \right] E[x_{T,i} x_{T,j}] E \left[\frac{1}{\hat{\sigma}_j^2} \right]. \quad (49)$$

REFERÊNCIAS

- [1] S. S. Narayan, A. M. Peterson, and M. J. Narasimha, "Transform domain LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 3, pp. 609-615, Jun. 1983.
- [2] S. Hosur and A. H. Tewfik, "Wavelet transform domain adaptive FIR filtering," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 617-630, Mar. 1997.
- [3] S. J. Chern, J. C. Horng, and K. M. Wong, "The performance of the hybrid LMS adaptive algorithm," *Signal Processing*, vol. 44, no. 1, pp. 67-88, Jun. 1995.
- [4] D. I. Kim and P. De Wilde, "Performance analysis of the DCT-LMS adaptive filtering algorithm," *Signal Processing*, vol. 80, no. 8, pp. 1629-1654, Aug. 2000.
- [5] E. M. Lobato, O. J. Tobias, and R. Seara, "Stochastic modeling of the transform-domain ϵ LMS algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1840-1852, May 2008.
- [6] C. G. Samson and U. Reddy, "Fixed point error analysis of the normalized ladder algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 5, pp. 1177-1191, Oct. 1983.
- [7] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing*. New York: McGraw-Hill, 2000.
- [8] M. Rupp, "The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1149-1160, Mar. 1993.
- [9] J. E. Kolodziej, O. J. Tobias e R. Seara, "Um modelo preciso para o algoritmo NLMS," in *Anais do XVII Congresso Brasileiro de Automática*, Juiz de Fora, MG, Set. 2008, pp. 1-6.