

Modelagem Estocástica do Algoritmo LMS para Formas Bilineares

Khaled Jamal Bakri, Eduardo Vinicius Kuhn e Rui Seara

Resumo—Este artigo trata da modelagem estocástica do algoritmo adaptativo LMS-BF (*least-mean-square for bilinear forms*), o qual é caracterizado a partir das respostas ao impulso temporal e espacial de um modelo espaço-temporal para sistemas MISO (*multiple-input/single-output*). Especificamente, considerando um problema de identificação de sistema com planta estacionária e sinais de entrada gaussianos reais, expressões de modelo são derivadas caracterizando o comportamento médio dos vetores de coeficientes dos filtros temporal, espacial e espaço-temporal, a curva de aprendizagem e recursões para computar as matrizes de correlação requeridas. Resultados de simulação para diferentes cenários de operação são apresentados, confirmando que o modelo proposto prevê satisfatoriamente o comportamento do algoritmo tanto na fase transitória quanto em regime permanente.

Palavras-chave—Algoritmo LMS-BF, filtragem adaptativa, formas bilineares, modelagem estocástica.

Abstract—This paper deals with the stochastic modeling of the least-mean-square algorithm for bilinear forms (LMS-BF), which is characterized from the temporal and spatial impulse responses of a spatiotemporal model for multiple-input/single-output (MISO) systems. Specifically, considering a system identification problem with stationary plant and real-valued Gaussian input data, model expressions are derived characterizing the mean weight behavior of the temporal, spatial, and spatiotemporal adaptive filters, learning curve, and recursions for computing the correlation matrices required. Simulation results for different operation scenarios are presented, confirming that the proposed model predicts satisfactorily the algorithm behavior for both transient and steady-state phases.

Keywords—LMS-BF algorithm, adaptive filtering, bilinear forms, stochastic modeling.

I. INTRODUÇÃO

Na literatura, formas bilineares vêm sendo utilizadas em aplicações relacionadas à obtenção de representações aproximadas de sistemas não-lineares [1]. Tais representações buscam caracterizar o comportamento desses sistemas através de sistemas lineares, fornecendo assim um adequado compromisso entre a precisão da representação obtida e a simplicidade do tratamento matemático. Para tal, diferentes algoritmos adaptativos envolvendo formas bilineares vêm sendo desenvolvidos visando aplicações de identificação de sistema [2]–[4], equalização de canal [5], cancelamento de eco [6], controle ativo de ruído [7], [8], redes neurais [9], dentre outras. Contudo, os algoritmos discutidos até então têm considerado que o termo bilinear seja definido com respeito aos dados, isto é, com vistas à relação de entradas e saídas do sistema. Em contraste, [10] introduz uma nova abordagem em que o termo bilinear é definido em relação às respostas ao impulso temporal e espacial de um modelo espaço-temporal para sistemas MISO (*multiple-input/single-output*). A partir

Khaled Jamal Bakri, Eduardo Vinicius Kuhn e Rui Seara, LINSE–Laboratório de Circuitos e Processamento de Sinais, Departamento de Engenharia Elétrica e Eletrônica, Universidade Federal de Santa Catarina (UFSC), Florianópolis-SC, Brasil, (e-mails: khaled@linse.ufsc.br, kuhn@linse.ufsc.br e seara@linse.ufsc.br).

Eduardo Vinicius Kuhn, LAPSE–Laboratório de Processamento de Sinais e Eletrônica, Departamento de Engenharia Eletrônica, Universidade Tecnológica Federal do Paraná (UTFPR), Toledo-PR, Brasil (kuhn@utfpr.edu.br).

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

dessa premissa, são revisados algoritmos adaptativos já consagrados, tais como o LMS (*least-mean-square*) [11], o NLMS (*normalized LMS*) [12] e o RLS (*recursive-least-square*) [13] (veja uma discussão mais detalhada em [1]).

Tendo em vista a aplicabilidade prática bem como o seu caráter inovador, torna-se imperativo compreender de maneira mais abrangente o funcionamento de algoritmos adaptativos desenvolvidos para operar com formas bilineares. Nesse sentido, [1] trata de alguns aspectos fundamentais relacionados ao desempenho desses algoritmos. Todavia, a discussão conduzida lá se baseia em resultados advindos de simulações. Além disso, os resultados teóricos apresentados focam apenas sobre o comportamento dos algoritmos em regime permanente. Consequentemente, o entendimento do comportamento dos algoritmos bem como o impacto dos parâmetros sobre seu desempenho tornam-se um tanto limitados, o que justifica a importância de uma modelagem estocástica mais abrangente. Tal modelagem busca derivar expressões analíticas que possibilitem prever o comportamento do algoritmo adaptativo frente a diferentes cenários de operação. A partir de tais expressões, relações de compromisso entre os parâmetros do algoritmo e seu desempenho podem ser estabelecidas, evitando assim a necessidade de extensivas simulações de Monte Carlo (MC).

Neste contexto, focando sobre o algoritmo adaptativo LMS operando com formas bilineares (introduzido em [11]), o presente trabalho tem os seguintes objetivos:

- i) desenvolver um modelo estocástico considerando sinais de entrada gaussianos reais não-correlacionados e correlacionados; e
- ii) obter expressões de modelo descrevendo o comportamento médio dos vetores de coeficientes dos filtros temporal, espacial e espaço-temporal, a curva de aprendizagem e recursões para computar as matrizes de correlação requeridas.

Vale salientar que o desenvolvimento apresentado aqui versa sobre o comportamento do algoritmo na fase transitória.

Este artigo está organizado como segue. A Seção II descreve o modelo espaço-temporal considerado e introduz o algoritmo LMS para formas bilineares (LMS-BF). A Seção III apresenta as expressões caracterizando o modelo estocástico proposto. Na Seção IV, resultados de simulação são mostrados, confirmando a precisão das expressões de modelo obtidas. Finalmente, a Seção V traz as considerações finais do presente trabalho de pesquisa.

II. FORMULAÇÃO DO PROBLEMA

Considerando uma aplicação de identificação de sistemas MISO envolvendo formas bilineares (conforme ilustrado na Fig. 1) [1], o sinal desejado $d(n)$ pode ser expresso como¹

$$d(n) = \mathbf{h}^T \mathbf{X}(n) \mathbf{g} + w(n) \quad (1)$$

onde $\mathbf{h} = [h_1 \ \dots \ h_L]^T$ e $\mathbf{g} = [g_1 \ \dots \ g_M]^T$ denotam vetores contendo os coeficientes das respostas ao impulso temporal e espacial de um modelo espaço-temporal, respectivamente,

$$\mathbf{X}(n) = [\mathbf{x}_1(n) \ \dots \ \mathbf{x}_M(n)] \quad (2)$$

é uma matriz de entrada cujas colunas são vetores contendo as amostras mais recentes do m -ésimo sinal de entrada, dado por

¹ Note que $y(n) = \mathbf{h}^T \mathbf{X}(n) \mathbf{g}$ é um termo bilinear em \mathbf{g} e \mathbf{h} , visto que $y(n)$ é uma função linear de \mathbf{g} para um dado \mathbf{h} fixo e vice-versa.

$$\mathbf{x}_m(n) = [x_m(n) \ \cdots \ x_m(n-L+1)]^T, \quad m = 1, \dots, M \quad (3)$$

e $w(n)$ caracteriza o ruído aditivo de medição. Ainda, definindo um vetor contendo as colunas de $\mathbf{X}(n)$ empilhadas, isto é,

$$\tilde{\mathbf{x}}(n) = \text{vec}[\mathbf{X}(n)] \quad (4)$$

o sinal de saída da planta $y(n)$ pode ser expresso como

$$y(n) = \mathbf{f}^T \tilde{\mathbf{x}}(n) \quad (5)$$

com \otimes representando o produto de Kronecker e

$$\mathbf{f} = \mathbf{g} \otimes \mathbf{h} \quad (6)$$

um vetor de dimensão ML contendo os coeficientes da resposta ao impulso do modelo espaço-temporal.

Considere agora que \mathbf{h} e \mathbf{g} possam ser identificados através de filtros adaptativos com vetores de coeficientes $\hat{\mathbf{h}}(n)$ e $\hat{\mathbf{g}}(n)$, respectivamente. Então, o algoritmo LMS-BF pode ser utilizado, cujas equações de adaptação são [11]

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu_{\hat{\mathbf{h}}} \tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n) e_{\hat{\mathbf{g}}}(n) \quad (7)$$

e

$$\hat{\mathbf{g}}(n) = \hat{\mathbf{g}}(n-1) + \mu_{\hat{\mathbf{g}}} \tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) e_{\hat{\mathbf{h}}}(n) \quad (8)$$

onde $\mu_{\hat{\mathbf{h}}}$ e $\mu_{\hat{\mathbf{g}}}$ são os passos de adaptação,

$$\begin{aligned} e_{\hat{\mathbf{h}}}(n) &= d(n) - \hat{\mathbf{h}}^T(n-1) \tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n) \\ &= (\mathbf{g} \otimes \mathbf{h})^T \tilde{\mathbf{x}}(n) - \hat{\mathbf{h}}^T(n-1) \tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n) + w(n) \end{aligned} \quad (9)$$

e

$$\begin{aligned} e_{\hat{\mathbf{g}}}(n) &= d(n) - \hat{\mathbf{g}}^T(n-1) \tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) \\ &= (\mathbf{g} \otimes \mathbf{h})^T \tilde{\mathbf{x}}(n) - \hat{\mathbf{g}}^T(n-1) \tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) + w(n) \end{aligned} \quad (10)$$

denotam os sinais de erro com

$$\tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n) = [\hat{\mathbf{g}}(n-1) \otimes \mathbf{I}_L]^T \tilde{\mathbf{x}}(n) \quad (11)$$

e

$$\tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) = [\mathbf{I}_M \otimes \hat{\mathbf{h}}(n-1)]^T \tilde{\mathbf{x}}(n). \quad (12)$$

Vale salientar que \mathbf{I}_L e \mathbf{I}_M em (11) e (12) representam matrizes identidade de dimensão $L \times L$ e $M \times M$, respectivamente. Desse modo, uma estimativa para o sinal de saída do modelo espaço-temporal é obtida como

$$\begin{aligned} \hat{y}(n) &= \hat{\mathbf{h}}^T(n-1) \mathbf{X}(n) \hat{\mathbf{g}}(n-1) \\ &= [\hat{\mathbf{g}}(n-1) \otimes \hat{\mathbf{h}}(n-1)]^T \tilde{\mathbf{x}}(n) \\ &= \hat{\mathbf{f}}^T(n-1) \tilde{\mathbf{x}}(n) \end{aligned} \quad (13)$$

onde $\hat{\mathbf{f}}(n)$ caracteriza uma estimativa da resposta ao impulso do modelo espaço-temporal. Embora $\hat{\mathbf{h}}(n)$ e $\hat{\mathbf{g}}(n)$ possam ser versões escaladas de \mathbf{h} e \mathbf{g} , a estimativa da resposta ao impulso espaço-temporal

$$\hat{\mathbf{f}}(n) = \eta \hat{\mathbf{g}}(n) \otimes \frac{1}{\eta} \hat{\mathbf{h}}(n) \quad (14)$$

é identificada sem ambiguidade para $\eta \neq 0$ [1]. Por fim, o sinal de erro pode então ser expresso por

$$e(n) = d(n) - \hat{y}(n) \quad (15)$$

com $e(n) = e_{\hat{\mathbf{h}}}(n) = e_{\hat{\mathbf{g}}}(n)$ (conforme mostrado em [1]).

III. MODELO PROPOSTO

O modelo estocástico proposto é composto por um conjunto de expressões matemáticas descrevendo o comportamento médio dos vetores de coeficientes dos filtros temporal, espacial e espaço-temporal, curva de aprendizagem e recursões para computar as matrizes de correlação requeridas. Para derivar tais expressões, as seguintes hipóteses simplificativas são estabelecidas:

H1) Os sinais de entrada são obtidos através de processos gaussianos estacionários independentes entre si, tendo média zero, variância $\sigma_x^2 = E[x_m^2(n)]$ e matriz de autocorrelação $\mathbf{R} = E[\mathbf{x}_m(n) \mathbf{x}_m^T(n)]$ [1], [14], [15].

H2) O ruído de medição $w(n)$ tem média nula, variância σ_w^2 e é descorrelacionado de qualquer outro sinal envolvido no processo [14], [15].

H3) Os vetores de coeficientes $\hat{\mathbf{f}}(n)$, $\hat{\mathbf{h}}(n)$ e $\hat{\mathbf{g}}(n)$ são assumidos estatisticamente independentes entre si, bem como de qualquer outra variável do processo [1].

Note que as Hipóteses H1)-H3) vêm sendo comumente utilizadas na modelagem estocástica de diversos algoritmos adaptativos e são requeridas para tornar o desenvolvimento matematicamente tratável (conforme discutido em [14] e [15]).

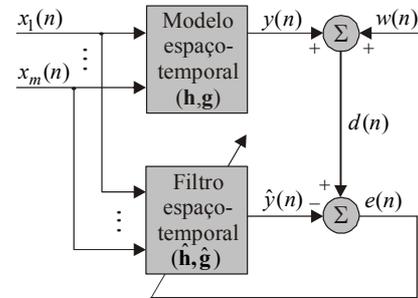


Fig.1. Diagrama de blocos de um problema de identificação de sistema considerando um filtro espaço-temporal.

A. Comportamento Médio dos Vetores de Coeficientes

O comportamento médio do vetor de coeficientes do filtro temporal $\hat{\mathbf{h}}(n)$ é obtido substituindo (10) em (7), tomando o valor esperado de ambos os lados da expressão resultante e utilizando as Hipóteses H1)-H3). Assim,

$$E[\hat{\mathbf{h}}(n)] = [\mathbf{I}_L - \mu_{\hat{\mathbf{h}}} \mathbf{R}_{\hat{\mathbf{g}}1}(n)] E[\hat{\mathbf{h}}(n-1)] + \mu_{\hat{\mathbf{h}}} \mathbf{R}_{\hat{\mathbf{g}}2}(n) \mathbf{h} \quad (16)$$

com

$$\mathbf{R}_{\hat{\mathbf{g}}1}(n) = E[\tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n) \tilde{\mathbf{x}}_{\hat{\mathbf{g}}}^T(n)] \quad (17)$$

e

$$\mathbf{R}_{\hat{\mathbf{g}}2}(n) = E[\tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n) \tilde{\mathbf{x}}^T(n)] [\mathbf{g} \otimes \mathbf{I}_L]. \quad (18)$$

Então, utilizando (11), considerando as propriedades do produto de Kronecker [16], e levando em conta as Hipóteses H1) e H3), (17) e (18) são simplificadas para

$$\mathbf{R}_{\hat{\mathbf{g}}1}(n) = \sum_{i=1}^M \hat{G}_{i,i}(n-1) \mathbf{R} \quad (19)$$

com

$$\hat{\mathbf{G}}(n) = E[\hat{\mathbf{g}}(n) \hat{\mathbf{g}}^T(n)] \quad (20)$$

e

$$\mathbf{R}_{\hat{\mathbf{g}}2}(n) = \sum_{i=1}^M G_{i,i}(n-1) \mathbf{R} \quad (21)$$

sendo

$$\mathbf{G}(n) = \mathbf{g} E[\hat{\mathbf{g}}^T(n)]. \quad (22)$$

Similarmente, substituindo (9) em (8), uma expressão que descreve o comportamento médio do vetor de coeficientes do filtro espacial $\hat{\mathbf{g}}(n)$ é obtida como

$$E[\hat{\mathbf{g}}(n)] = [\mathbf{I}_M - \mu_{\hat{\mathbf{g}}} \mathbf{R}_{\hat{\mathbf{h}}1}(n)] E[\hat{\mathbf{g}}(n-1)] + \mu_{\hat{\mathbf{g}}} \mathbf{R}_{\hat{\mathbf{h}}2}(n) \mathbf{g} \quad (23)$$

com

$$\mathbf{R}_{\hat{\mathbf{h}}1}(n) = E[\tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) \tilde{\mathbf{x}}_{\hat{\mathbf{h}}}^T(n)] \quad (24)$$

e

$$\mathbf{R}_{\hat{\mathbf{h}}2}(n) = E[\tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) \tilde{\mathbf{x}}^T(n)] [\mathbf{I}_M \otimes \mathbf{h}]. \quad (25)$$

Por sua vez, utilizando (12), levando em conta as propriedades do produto de Kronecker e do traço de uma matriz [16], como também considerando as Hipóteses H1) e H3), (24) e (25) são reescritas como

$$\mathbf{R}_{\hat{\mathbf{h}}1}(n) = \lambda^T \hat{\mathbf{h}}'(n-1) \mathbf{I}_M \quad (26)$$

e

$$\mathbf{R}_{\hat{\mathbf{h}}2}(n) = \lambda^T \mathbf{h}'(n-1) \mathbf{I}_M \quad (27)$$

onde λ denota um vetor contendo os autovalores de \mathbf{R} , enquanto $\hat{\mathbf{h}}'(n)$ e $\mathbf{h}'(n)$ são vetores contendo os elementos da diagonal de

$$\hat{\mathbf{H}}'(n) = \mathbf{Q}^T \hat{\mathbf{H}}(n) \mathbf{Q} \quad (28)$$

e

$$\mathbf{H}'(n) = \mathbf{Q}^T \mathbf{H}(n) \mathbf{Q} \quad (29)$$

sendo \mathbf{Q} uma matriz de autovetores oriunda da autodecomposição da matriz \mathbf{R} (isto é, $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ [15]). Note em (28) e (29) que $\hat{\mathbf{H}}(n) = E[\hat{\mathbf{h}}(n)\hat{\mathbf{h}}^T(n)]$ e $\mathbf{H}(n) = \mathbf{h}E[\hat{\mathbf{h}}^T(n)]$ caracterizam matrizes de correlação relacionadas ao filtro temporal.

Finalmente, o comportamento médio do vetor de coeficientes do filtro espaço-temporal $\hat{\mathbf{f}}(n)$ é determinado tomando o valor esperado de ambos os lados de (14) e considerando a Hipótese H3). Dessa forma,

$$E[\hat{\mathbf{f}}(n)] = E[\hat{\mathbf{g}}(n)] \otimes E[\hat{\mathbf{h}}(n)]. \quad (30)$$

Portanto, assumindo que expressões recursivas caracterizando a evolução de (20), (22), (28) e (29) possam ser determinadas [veja mais à frente (46), (52), (53) e (56)], as matrizes (19), (21), (26) e (27) podem ser computadas; conseqüentemente, é possível prever o comportamento médio do vetor de coeficientes do filtro temporal, espacial e espaço-temporal através de (16), (23) e (30), respectivamente.

B. Curva de Aprendizagem

Uma expressão que descreve a curva de aprendizagem do algoritmo [erro quadrático médio (EQM)] pode ser determinada substituindo (5) e (13) em (15) e reescrevendo a expressão resultante em função do vetor de erro nos coeficientes do filtro espaço-temporal $\mathbf{c}_f(n) = \mathbf{f} - \hat{\mathbf{f}}(n)$. Desse modo,

$$e(n) = \mathbf{c}_f^T(n-1)\tilde{\mathbf{x}}(n) + w(n). \quad (31)$$

Então, calculando $e^2(n)$ a partir de (31), tomando o valor esperado de ambos os lados da expressão resultante, e considerando a Hipótese H2), tem-se

$$J(n) = J_{\min} + J_{\text{ex}}(n) \quad (32)$$

onde

$$\begin{aligned} J_{\min} &= E[w^2(n)] \\ &= \sigma_w^2 \end{aligned} \quad (33)$$

representa o EQM mínimo obtido quando $\hat{\mathbf{f}}(n) = \mathbf{f}$, enquanto

$$J_{\text{ex}}(n) = E[\mathbf{c}_f^T(n-1)\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}^T(n)\mathbf{c}_f(n-1)] \quad (34)$$

corresponde ao EQM em excesso (EQME) introduzido pelo algoritmo.

Agora, utilizando as Hipóteses H1) e H3), juntamente com as propriedades do produto de Kronecker e do traço de uma matriz [16], (34) pode ser reescrita como

$$J_{\text{ex}}(n) = \text{tr}[\mathbf{R}_1 \mathbf{K}(n-1)] \quad (35)$$

onde

$$\mathbf{R}_1 = \mathbf{I}_M \otimes \mathbf{R} \quad (36)$$

denota a matriz autocorrelação de $\tilde{\mathbf{x}}(n)$ e $\mathbf{K}(n)$ caracteriza a matriz de correlação do vetor de erro nos coeficientes do filtro espaço-temporal, isto é,

$$\begin{aligned} \mathbf{K}(n) &= E[\mathbf{c}_f(n)\mathbf{c}_f^T(n)] \\ &= \mathbf{f}\mathbf{f}^T - 2\mathbf{F}(n) + \hat{\mathbf{F}}(n) \end{aligned} \quad (37)$$

com

$$\begin{aligned} \mathbf{F}(n) &= \mathbf{f}E[\hat{\mathbf{f}}^T(n)] \\ &= \mathbf{G}(n) \otimes \mathbf{H}(n) \end{aligned} \quad (38)$$

e

$$\begin{aligned} \hat{\mathbf{F}}(n) &= E[\hat{\mathbf{f}}(n)\hat{\mathbf{f}}^T(n)] \\ &= \hat{\mathbf{G}}(n) \otimes \hat{\mathbf{H}}(n). \end{aligned} \quad (39)$$

Por fim, a partir de (35) e (36), e da autodecomposição da matriz \mathbf{R} , a expressão que descreve a evolução do EQME é simplificada para

$$\begin{aligned} J_{\text{ex}}(n) &= \text{tr}[\mathbf{\Lambda}_1 \mathbf{Q}_1^T \mathbf{K}(n-1) \mathbf{Q}_1] \\ &= \lambda_1^T \mathbf{k}'(n-1) \end{aligned} \quad (40)$$

com λ_1 e $\mathbf{k}'(n)$ contendo os elementos da diagonal de

$$\mathbf{\Lambda}_1 = \mathbf{I}_M \otimes \mathbf{\Lambda} \quad (41)$$

e

$$\mathbf{K}'(n) = \mathbf{Q}_1^T \mathbf{K}(n) \mathbf{Q}_1 \quad (42)$$

para

$$\mathbf{Q}_1 = \mathbf{I}_M \otimes \mathbf{Q}. \quad (43)$$

Portanto, levando em conta (32), (33) e (40), a curva de aprendizagem do algoritmo é completamente caracterizada se a potência do ruído de medição, os autovalores da matriz \mathbf{R} e o vetor $\mathbf{k}'(n)$ forem conhecidos.

C. Matrizes de Correlação dos Vetores de Coeficientes

Uma expressão recursiva descrevendo a evolução da matriz de correlação temporal $\hat{\mathbf{H}}(n)$ pode ser derivada substituindo (10) em (7), determinando o produto $\hat{\mathbf{h}}(n)\hat{\mathbf{h}}^T(n)$, tomando o valor esperado de ambos os lados da expressão resultante, e utilizando (17), (18) e (33) juntamente com as Hipóteses H1)-H3). Logo,

$$\begin{aligned} \hat{\mathbf{H}}(n) &= \hat{\mathbf{H}}(n-1) + \mu_{\hat{\mathbf{h}}} \mathbf{R}_{\hat{\mathbf{g}}_2}(n) \mathbf{H}(n-1) - \mu_{\hat{\mathbf{h}}} \mathbf{R}_{\hat{\mathbf{g}}_1}(n) \hat{\mathbf{H}}(n-1) \\ &\quad + \mu_{\hat{\mathbf{h}}} \mathbf{H}(n-1) \mathbf{R}_{\hat{\mathbf{g}}_2}(n) - \mu_{\hat{\mathbf{h}}} \hat{\mathbf{H}}(n-1) \mathbf{R}_{\hat{\mathbf{g}}_1}(n) + \mu_{\hat{\mathbf{h}}}^2 \mathbf{R}_{\hat{\mathbf{g}}_1}(n) J_{\min} \\ &\quad + \mu_{\hat{\mathbf{h}}}^2 \mathbf{R}_{\hat{\mathbf{g}}_3}(n) \end{aligned} \quad (44)$$

com

$$\mathbf{R}_{\hat{\mathbf{g}}_3}(n) = E[\tilde{\mathbf{x}}_{\hat{\mathbf{g}}}(n)\tilde{\mathbf{x}}^T(n)\mathbf{K}(n-1)\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}_{\hat{\mathbf{g}}}^T(n)]. \quad (45)$$

Então, considerando a autodecomposição de \mathbf{R} , pré- e pós-multiplicando (44) por \mathbf{Q}^T e \mathbf{Q} , respectivamente, utilizando as propriedades do produto de Hadamard [16] juntamente com as Hipóteses H1)-H3), obtém-se

$$\begin{aligned} \hat{\mathbf{H}}'(n) &= \hat{\mathbf{H}}'(n-1) \circ \{ \mathbf{1}_{L \times L} - \mu_{\hat{\mathbf{h}}} [\mathbf{u}(n) \mathbf{1}_{1 \times L} + \mathbf{1}_{L \times 1} \mathbf{u}^T(n)] \} \\ &\quad + \mu_{\hat{\mathbf{h}}} \mathbf{H}'(n-1) \circ [\mathbf{v}(n) \mathbf{1}_{1 \times L} + \mathbf{1}_{L \times 1} \mathbf{v}^T(n)] + \mu_{\hat{\mathbf{h}}}^2 \mathbf{U}(n) J_{\min} \\ &\quad + \mu_{\hat{\mathbf{h}}}^2 \mathbf{Z}(n) \end{aligned} \quad (46)$$

sendo $\mathbf{1}_{L \times L}$ uma matriz de 1's com dimensão $L \times L$, enquanto $\mathbf{u}(n)$ e $\mathbf{v}(n)$ denotam vetores contendo os elementos da diagonal de

$$\mathbf{U}(n) = \sum_{i=1}^M \hat{G}_{i,i}(n-1) \mathbf{\Lambda} \quad (47)$$

e

$$\mathbf{V}(n) = \sum_{i=1}^M G_{i,i}(n-1) \mathbf{\Lambda} \quad (48)$$

respectivamente. Note em (46) que

$$\mathbf{Z}(n) = \sum_{i=1}^M \sum_{j=1}^M \hat{G}_{i,j}(n-1) \mathbf{D}_{i,j}(n) \quad (49)$$

com

$$\mathbf{D}_{i,j}(n) = \begin{cases} 2[\lambda \lambda^T \circ \mathbf{K}'_{i,j}(n-1)] + \lambda^T \mathbf{k}'_{i,j}(n-1) \mathbf{\Lambda}, & i = j \\ \lambda \lambda^T \circ \mathbf{K}'_{i,j}(n-1), & i \neq j \end{cases} \quad (50)$$

onde $\mathbf{k}'_{i,j}(n)$ são vetores contendo os elementos da diagonal das matrizes $\mathbf{K}'_{i,j}(n)$ (com dimensão $L \times L$) advindas do particionamento de

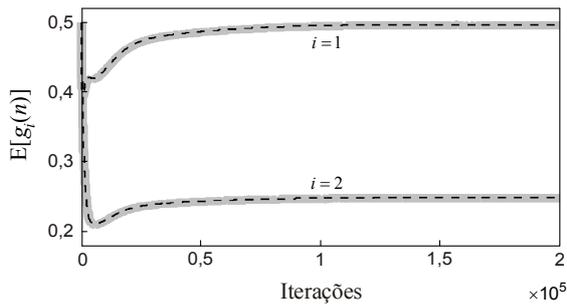
$$\mathbf{K}'(n) = \begin{bmatrix} \mathbf{K}'_{11}(n) & \cdots & \mathbf{K}'_{1M}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{K}'_{M1}(n) & \cdots & \mathbf{K}'_{MM}(n) \end{bmatrix}. \quad (51)$$

Similarmente ao realizado para obter (46), uma expressão recursiva caracterizando a evolução de (29) é derivada como

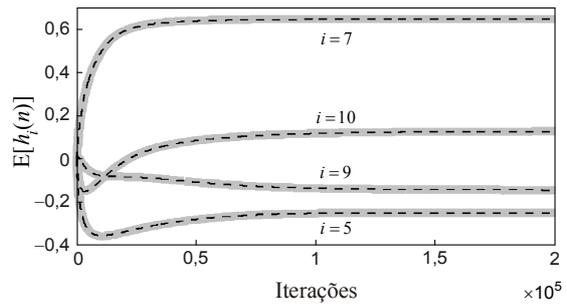
$$\mathbf{H}'(n) = [\mathbf{1}_{L \times L} - \mu_{\hat{\mathbf{h}}} \mathbf{u}(n) \mathbf{1}_{1 \times L}] \circ \mathbf{H}'(n-1) + \mu_{\hat{\mathbf{h}}} \mathbf{v}(n) \mathbf{1}_{1 \times L} \circ \mathbf{Q}^T \mathbf{h} \mathbf{h}^T \mathbf{Q}. \quad (52)$$

Por sua vez, uma expressão recursiva descrevendo a evolução da matriz de correlação espacial $\hat{\mathbf{G}}(n)$ é derivada substituindo (9) em (8), determinando o produto $\hat{\mathbf{g}}(n) \hat{\mathbf{g}}^T(n)$, tomando o valor esperado de ambos os lados da expressão resultante, utilizando (26), (27) e (33), como também as Hipóteses H1)-H3). Dessa forma,

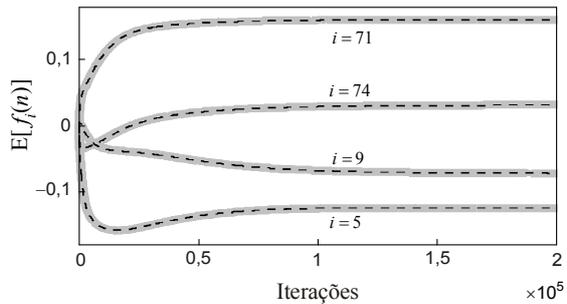
$$\begin{aligned} \hat{\mathbf{G}}(n) = & \hat{\mathbf{G}}(n-1) + \mu_{\hat{\mathbf{g}}} \mathbf{R}_{\hat{\mathbf{h}}_2}(n) \mathbf{G}(n-1) - \mu_{\hat{\mathbf{g}}} \mathbf{R}_{\hat{\mathbf{h}}_1}(n) \hat{\mathbf{G}}(n-1) \\ & + \mu_{\hat{\mathbf{g}}} \mathbf{G}(n-1) \mathbf{R}_{\hat{\mathbf{h}}_2}(n) - \mu_{\hat{\mathbf{g}}} \hat{\mathbf{G}}(n-1) \mathbf{R}_{\hat{\mathbf{h}}_1}(n) \\ & + \mu_{\hat{\mathbf{g}}}^2 \mathbf{R}_{\hat{\mathbf{h}}_1}(n) \mathbf{J}_{\min} + \mu_{\hat{\mathbf{g}}}^2 \mathbf{R}_{\hat{\mathbf{h}}_3}(n) \end{aligned} \quad (53)$$



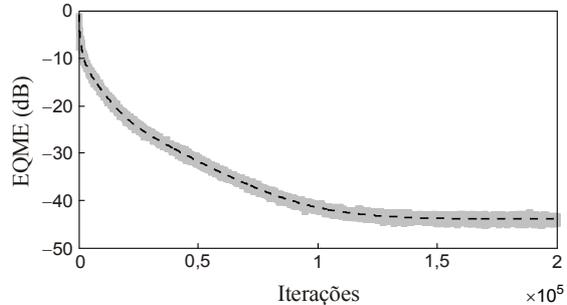
(a)



(b)



(c)



(d)

Fig. 2. Exemplo 1. Resultados obtidos por simulações de MC (linhas cinza) e predito a partir do modelo proposto (linhas escuras tracejadas). (a) Comportamento médio do vetor de coeficientes do filtro adaptativo espacial. (b) Comportamento médio de quatro coeficientes do filtro adaptativo temporal. (c) Comportamento médio de quatro coeficientes do filtro adaptativo espaço-temporal. (d) Curva de aprendizagem (EQME).

com

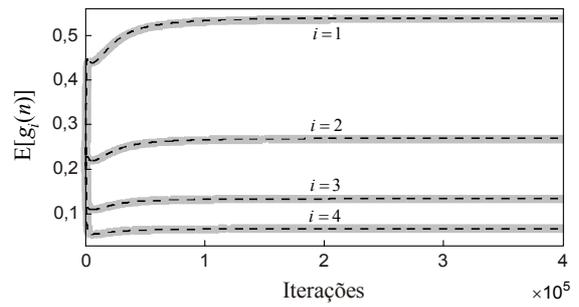
$$\mathbf{R}_{\hat{\mathbf{h}}_3}(n) = E[\tilde{\mathbf{x}}_{\hat{\mathbf{h}}}(n) \tilde{\mathbf{x}}_{\hat{\mathbf{h}}}^T(n) \mathbf{K}(n-1) \tilde{\mathbf{x}}(n) \tilde{\mathbf{x}}_{\hat{\mathbf{h}}}^T(n)]. \quad (54)$$

Note que os elementos de (54) podem ser determinados como $\mathbf{R}_{\hat{\mathbf{h}}_3}(n)$

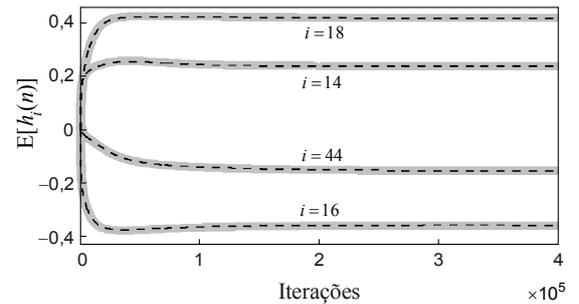
$$= \begin{cases} 2\lambda^T [\mathbf{K}'_{i,j}(n-1) \circ \hat{\mathbf{H}}'(n-1)] \lambda + \lambda^T \mathbf{k}'_{i,j}(n-1) \lambda^T \hat{\mathbf{h}}'(n-1), & i = j \\ \lambda^T [\mathbf{K}'_{i,j}(n-1) \circ \hat{\mathbf{H}}'(n-1)] \lambda, & i \neq j. \end{cases} \quad (55)$$

Do mesmo modo utilizado para determinar (53), uma expressão recursiva descrevendo a evolução de (22) é obtida como

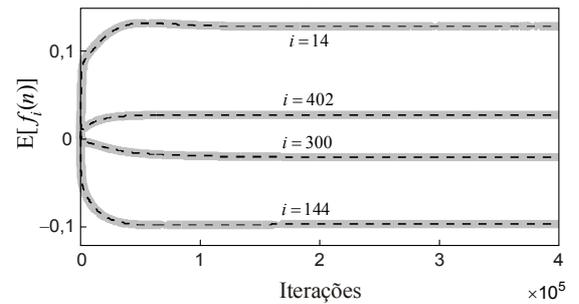
$$\mathbf{G}(n) = [\mathbf{I}_M - \mu_{\hat{\mathbf{g}}} \mathbf{R}_{\hat{\mathbf{h}}_1}(n)] \mathbf{G}(n-1) + \mu_{\hat{\mathbf{g}}} \mathbf{R}_{\hat{\mathbf{h}}_2}(n) \mathbf{g} \mathbf{g}^T. \quad (56)$$



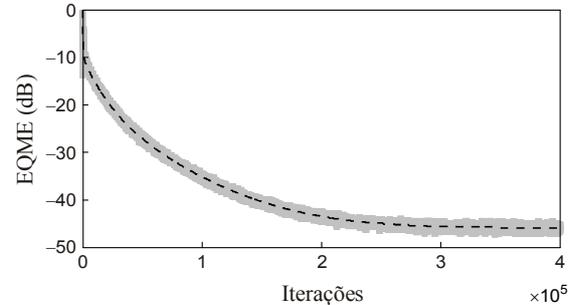
(a)



(b)



(c)



(d)

Fig. 3. Exemplo 2. Resultados obtidos por simulações de MC (linhas cinza) e predito a partir do modelo proposto (linhas escuras tracejadas). (a) Comportamento médio do vetor de coeficientes do filtro adaptativo espacial. (b) Comportamento médio de quatro coeficientes do filtro adaptativo temporal. (c) Comportamento médio de quatro coeficientes do filtro adaptativo espaço-temporal. (d) Curva de aprendizagem (EQME).

Finalmente, a partir de (28), (29) e (37)-(39), pode-se reescrever (42) como

$$\mathbf{K}'(n) = \mathbf{Q}_1^T \mathbf{f} \mathbf{f}^T \mathbf{Q}_1 - 2\mathbf{F}'(n) + \hat{\mathbf{F}}'(n) \quad (57)$$

onde

$$\mathbf{F}'(n) = \mathbf{G}(n) \otimes \mathbf{H}'(n) \quad (58)$$

e

$$\hat{\mathbf{F}}'(n) = \hat{\mathbf{G}}(n) \otimes \hat{\mathbf{H}}'(n). \quad (59)$$

Portanto, levando em conta (26), (27), (47)-(50) e (55), é possível prever o comportamento das matrizes de correlação dos vetores de coeficientes dos filtros temporal, espacial e espaço-temporal através de (44), (46), (53) e (56); consequentemente, a curva de aprendizagem do algoritmo, dada em (32), pode também ser predita.

IV. RESULTADOS DE SIMULAÇÃO

Visando verificar a precisão do modelo proposto, simulações de MC (considerando 200 realizações independentes) são efetuadas para um problema de identificação de sistema. Para tal, os sinais de entrada [veja a Hipótese H1)] são obtidos através de um processo autorregressivo de segunda ordem AR(2), dado por [15]

$$x_m(n) = -a_1 x_m(n-1) - a_2 x_m(n-2) + v(n) \quad (60)$$

onde a_1 e a_2 são os coeficientes do processo e $v(n)$ é um ruído branco cuja variância é ajustada de tal modo que $\sigma_x^2 = 1$. A razão sinal-ruído (SNR) é definida como [1]

$$\text{SNR} = 10 \log_{10} \left[\frac{(\mathbf{g} \otimes \mathbf{h})^T \mathbf{R}_1 (\mathbf{g} \otimes \mathbf{h})}{\sigma_w^2} \right]. \quad (61)$$

A resposta ao impulso espacial \mathbf{g} é dada por

$$\mathbf{g} = [1 \ 2^{-1} \ \dots \ 2^{-(M-1)}]^T. \quad (62)$$

Os parâmetros do algoritmo são: $\mu_{\hat{\mathbf{g}}} = \mu_{\hat{\mathbf{h}}} = 0,001$, $\hat{\mathbf{h}}(0) = [1 \ 0 \ \dots \ 0]^T$ e $\hat{\mathbf{g}}(0) = M^{-1} [1 \ \dots \ 1]^T$.

A. Exemplo 1

Neste exemplo, a precisão do modelo é verificada para sinais de entrada fracamente correlacionados e uma dada SNR. Especificamente, os sinais de entrada são obtidos de (60) para $a_1 = -0,6$ e $a_2 = 0,8$, resultando em uma dispersão de autovalores das matrizes de autocorrelação do sinal de entrada $\chi = 144,78$. A resposta ao impulso temporal \mathbf{h} (com dimensão $L = 64$) é obtida a partir da Recomendação ITU-T G.168 [17, Modelo 1], enquanto a resposta ao impulso espacial \mathbf{g} (com dimensão $M = 2$) advém de (62). Ainda, a variância do ruído de medição é $\sigma_w^2 = 2,89 \times 10^{-3}$, resultando em uma SNR de aproximadamente 20 dB.

B. Exemplo 2

Agora, a precisão do modelo proposto é verificada para sinais de entrada fortemente correlacionadas e uma SNR mais alta do que aquela do Exemplo 1. Para tal, os sinais de entrada são obtidos de (60) com $a_1 = -0,6$ e $a_2 = 0,93$, resultando em uma dispersão de autovalores das matrizes de autocorrelação do sinal de entrada $\chi = 1173,38$. A resposta ao impulso temporal \mathbf{h} (com dimensão $L = 128$) é obtida a partir da Recomendação ITU-T G.128 [17, Modelo 4], enquanto a resposta ao impulso espacial \mathbf{g} (com dimensão $M = 4$) advém de (62). Ainda, a variância do ruído de medição é $\sigma_w^2 = 8,48 \times 10^{-4}$, originando uma SNR de aproximadamente 30 dB.

As Figs. 2 e 3 mostram os resultados obtidos para os cenários descritos nos Exemplos 1 e 2, respectivamente. Observa-se, através dessas figuras, que o modelo proposto

prediz satisfatoriamente o comportamento do algoritmo LMS-BF, tanto na fase transitória quanto em regime permanente. Tal precisão pode ser verificada para diferentes dispersões de autovalores da matriz de autocorrelação do sinal de entrada, valores de SNR e ordens dos filtros temporal e espacial. Além disso, os resultados apresentados ratificam a precisão das hipóteses e aproximações utilizadas para o desenvolvimento do modelo.

V. CONSIDERAÇÕES FINAIS

Neste artigo, assumindo um número reduzido de hipóteses simplificativas, foi obtido um modelo estocástico do algoritmo LMS considerando formas bilineares. Particularmente, as expressões derivadas permitem prever o comportamento médio do vetor de coeficientes dos filtros temporal, espacial e espaço-temporal, a curva de aprendizagem e a evolução das matrizes de correlação requeridas no modelo. Os resultados de simulação ratificam a precisão das expressões de modelo obtidas frente a diferentes condições de operação. Por fim, destaca-se que o presente trabalho fornece algumas ferramentas para a análise de outros algoritmos adaptativos envolvendo formas bilineares.

REFERÊNCIAS

- [1] C. Paleologu, J. Benesty, and S. Ciochină, "Adaptive filtering for the identification of bilinear forms," *Digit. Signal Process.*, vol. 75, pp. 153-167, Apr. 2018.
- [2] P. Lopes dos Santos, J. A. Ramos, and J. L. Martins de Carvalho, "Identification of bilinear systems with white noise inputs: An iterative deterministic-stochastic subspace approach," *IEEE Trans. Control Syst. Technol.*, vol. 17, no. 5, pp. 1145-1153, Sep. 2009.
- [3] H. K. Baik and V. J. Mathews, "Adaptive lattice bilinear filters," *IEEE Trans. on Signal Process.*, vol. 41, no. 6, pp. 2033-2046, Jun. 1993.
- [4] M. Inagaki and H. Mochizuki, "Bilinear system identification by Volterra kernels estimation," *IEEE Trans. Autom. Control*, vol. 29, no. 8, pp. 746-749, Aug. 1984.
- [5] J. Lee and V. J. Mathews, "Adaptive bilinear predictors," in *Proc. Int. Conf. Acoust. Speech and Signal Process. (ICASSP)*, Adelaide, SA, Australia, Apr. 1994, vol. 3, pp. 489-492.
- [6] R. Hu and H. M. Ahmed, "Echo cancellation in high speed data transmission systems using adaptive layered bilinear filters," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 655-663, Feb/Apr 1994.
- [7] S.M. Kuo and H.-T. Wu, "Nonlinear adaptive bilinear filters for active noise control systems," *IEEE Trans. Circuits Syst. (Reg. Papers)*, vol. 52, no. 3, pp. 617-624, Mar. 2005.
- [8] L. Tan and J. Jiang, "Nonlinear active noise control using diagonal-channel LMS and RLS bilinear filters," in *Proc. Int. Midwest Symp. Circuits Syst. (MWSCAS)*, College Station, TX, USA, Aug. 2014, pp. 789-792.
- [9] H. Zhao, X. Zeng and Z. He, "Low-complexity nonlinear adaptive filter based on a pipelined bilinear recurrent neural network," *IEEE Trans. Neural Netw.*, vol. 22, no. 9, pp. 1494-1507, Sep. 2011.
- [10] J. Benesty, C. Paleologu, and S. Ciochină, "On the identification of bilinear forms with the Wiener filter," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 653-657, May 2017.
- [11] S. Ciochină, C. Paleologu, and J. Benesty, "Analysis of an LMS algorithm for bilinear forms," in *Proc. Int. Conf. Digital Signal Process. (DSP)*, London, UK, Aug. 2017, pp. 1-5.
- [12] C. Paleologu, J. Benesty, and S. Ciochină, "An NLMS algorithm for the identification of bilinear forms," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Kos, Greece, Aug./Sep. 2017, pp. 2620-2624.
- [13] C. Elisei-Ilieșcu, C. Paleologu, R. A. Dobre, S. Ciochina and J. Benesty, "An RLS algorithm for the identification of bilinear forms," in *Proc. Int. Symp. Design Technol. Electron. Packag. (SIITME)*, Constanta, Romania, Oct. 2017, pp. 292-295.
- [14] A. H. Sayed, *Adaptive Filters*. Hoboken, NJ: John Wiley & Sons, 2009.
- [15] S. Haykin, *Adaptive Filter Theory*, 5th ed. Upper Saddle River, NJ: Prentice Hall, 2014.
- [16] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed. Woodstock, OX: Princeton University Press, 2009.
- [17] ITU-T Recommendation G.168, *Digital Network Echo Cancellers*. Geneva, Switzerland: Int. Telecomm. Union, 2009.