

# Análise de Transiente do Algoritmo $\ell_2$ -norm Feature LMS Submodelado

Leonardo O. Santos, Diego B. Haddad, Mariane R. Petraglia

**Resumo**— Sistemas esparsos ou dotados de determinada estrutura, alguma característica intrínseca ao sistema que possa ser explorada, são comumente encontrados em aplicações de filtragem adaptativa. Algoritmos capazes de empregar essa propriedade para acelerar a convergência e/ou melhorar o desempenho em regime permanente são objeto de intenso estudo por parte da comunidade acadêmica nos últimos anos. Este artigo propõe um arcabouço matemático capaz de gerar algoritmos adaptativos, sendo eles conscientes de esparsidade ou não, através do método de multiplicadores de Lagrange. Utilizando este arcabouço matemático, é derivado um modelo estocástico do recém-proposto algoritmo  $\ell_2$ -norm Feature LMS para o caso em que o filtro adaptativo possui um comprimento subótimo. Os resultados dos experimentos apresentados confirmam a aderência das curvas preconizadas pelo modelo teórico com as derivadas de simulações.

**Palavras-Chave**— Filtragem Adaptativa, Sistemas Estruturados, Modelos Estocásticos.

**Abstract**— Sparse or structured systems are common in several adaptive filtering applications. Algorithms that employ this property to accelerate convergence and/or improve steady-state performance have attracted attention from the academic community in recent years. This paper presents a mathematical framework capable of generating structure-aware adaptive algorithms by the Lagrange multipliers method. From such framework, a stochastic model of the  $\ell_2$ -norm Feature LMS for suboptimal adaptive filter length is derived. The experimental results presented confirm the adherence of the curves obtained by the theoretical model with those derived from the simulations.

**Keywords**— Adaptive Filtering, Structured Systems, Stochastic Model.

## I. INTRODUÇÃO

Dada a sua versatilidade e devido à presença ubíqua de sistemas desconhecidos e/ou variantes no tempo, técnicas de filtragem adaptativa vêm sendo empregadas em diversas aplicações, tais como: cancelamento de eco acústico, equalização de canal, controle de ruído ativo e identificação de sistemas [1], [2]. Com o crescente número de pesquisas na área, foi observado que sistemas reais recorrentemente apresentam respostas ao impulso esparsas [1]–[5], que é a concentração de energia em, relativamente, poucos coeficientes. A esparsidade, ou outra característica estruturada, da resposta ao impulso é um tipo de conhecimento *a priori* sobre o sistema a ser identificado o qual pode ser incorporado ao algoritmo adaptativo, fazendo com que este apresente menor erro em regime permanente

e/ou maior taxa de convergência [3], [6], [7]. O advento de informações *a priori* acerca do sistema que cumpre ao filtro adaptativo emular pode, a princípio, ser empregado para melhorar suas características de aprendizado.

No contexto de identificação de sistemas, os coeficientes de um algoritmo de filtragem adaptativa são atualizados recursivamente de sorte que, após um número suficiente de iterações, o filtro possua uma estimativa adequada do sistema de referência. Tal atualização costuma ser regida pelo cálculo do gradiente de uma função custo estocástica, o qual é capaz de fornecer a direção de atualização engendrada pelo algoritmo adaptativo. Isto se deve ao fato de uma estratégia muito comum para a identificação adaptativa de um sistema de referência consistir na minimização de uma função custo, procedimento que não raro demanda o cálculo do gradiente estocástico dessa função. Cabe notar que o gradiente aponta para a direção de maior crescimento da função, e a atualização é feita no sentido contrário ao do gradiente, já que, afinal, o objetivo consiste em minimizar uma função custo<sup>1</sup>.

A Fig. 1 mostra a estrutura de um algoritmo adaptativo (aplicado à identificação de sistemas), cujos coeficientes do filtro de comprimento  $N$  são coletados no vetor  $\mathbf{w}(k) \in \mathbb{R}^N$  definido por:

$$\mathbf{w}(k) \triangleq [w_0(k) \quad w_1(k) \quad \dots \quad w_{N-1}(k)]^T, \quad (1)$$

sendo a saída do filtro no instante  $k$  dada por

$$y(k) \triangleq \mathbf{w}^T(k)\mathbf{x}(k), \quad (2)$$

com o vetor de entrada  $\mathbf{x}(k)$  descrito por:

$$\mathbf{x}(k) \triangleq [x(k) \quad x(k-1) \quad \dots \quad x(k-N+1)]^T, \quad (3)$$

onde  $k$  representa o índice da iteração. Após um número suficiente de iterações, é esperado que  $y(k)$  seja similar ao sinal de referência

$$d(k) = (\mathbf{w}^*)^T \mathbf{x}(k) + \nu(k), \quad (4)$$

sendo  $\mathbf{w}^* \in \mathbb{R}^N$  o vetor que contém os coeficientes a serem identificados:

$$\mathbf{w}^* \triangleq [w_0^* \quad w_1^* \quad \dots \quad w_{N-1}^*]^T, \quad (5)$$

e  $\nu(k)$  denotando ruído branco com variância  $\sigma_\nu^2(k)$  somado ao sinal de referência, cuja origem é normalmente atribuída a erros de medição e/ou de modelagem. Denotamos por  $e(k) \triangleq d(k) - y(k)$  o sinal de realimentação, o qual consiste no erro entre o sinal de referência e a saída do filtro adaptativo.

<sup>1</sup>Uma exceção a essa regra reside nos algoritmos que se valem de funções objetivo baseadas na correntopia, as quais cumpre maximizar [8].

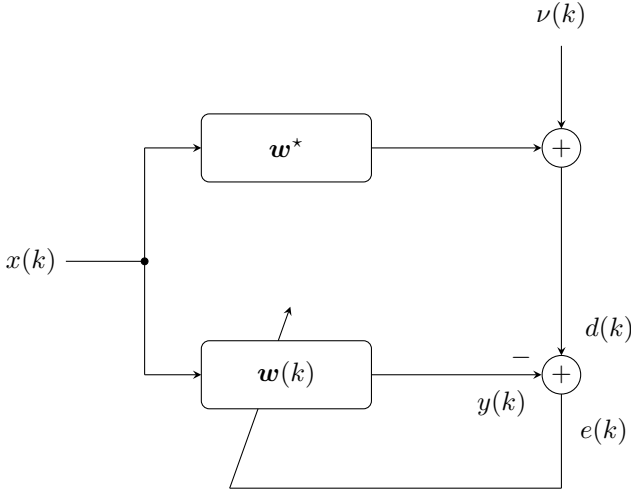


Fig. 1. Diagrama de blocos de um algoritmo adaptativo, voltado para a identificação de sistemas.

Este artigo é estruturado do seguinte modo: a Seção II descreve o algoritmo  $\ell_2$ -norm Feature LMS ( $\ell_2$ -F-LMS), recentemente publicado em [9]. Embora uma análise estocástica do desempenho do algoritmo tenha sido apresentada em [9], o caso em que o comprimento do filtro é deficiente não é contemplado. Tal configuração sub-ótima pode ocorrer na prática quando o sistema a identificar apresenta uma função de transferência longa, quando o projetista pretende lidar com limitações computacionais, ou quando um filtro adaptativo com comprimento variante no tempo é empregado [10]. A Seção III detalha o modelo estocástico proposto, o qual é cotejado com resultados simulados na Seção IV. Por fim, a Seção V elenca os comentários conclusivos do artigo.

## II. ALGORITMO $\ell_2$ -NORM FEATURE LMS

Nesta seção, é feita a derivação da equação de atualização do algoritmo  $\ell_2$ -F-LMS através da técnica dos multiplicadores de Lagrange. Note que, diferentemente da abordagem tradicional - a qual adota o método do gradiente estocástico para derivação de novos algoritmos - a equação de atualização do algoritmo é obtida resolvendo-se um problema de otimização com restrição.

O método dos multiplicadores de Lagrange permite encontrar uma solução para um problema de otimização determinístico. Portanto, considere o seguinte problema de otimização:

$$\min \frac{1}{2} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 + \kappa \|\mathbf{F}\mathbf{w}(k+1)\|^2 \quad (6)$$

sujeito a  $e_p(k) = [1 - \gamma(k)] e(k)$ ,

onde  $\kappa$  é um parâmetro ajustável para controlar o módulo da penalização,  $\gamma(k)$  é um parâmetro escolhido pelo projetista [11],  $e_p(k)$  é o erro *a posteriori* definido por:

$$e_p(k) \triangleq d(k) - \mathbf{w}^T(k+1)\mathbf{x}(k). \quad (7)$$

e a matriz de pesos  $\mathbf{F}$  penaliza soluções cuja energia do vetor  $\mathbf{F}\mathbf{w}(k+1)$  é muito grande [2], [6], [9]. Note que se for de conhecimento do projetista que o sistema a ser

identificado possuir uma resposta ao impulso que favorece baixas frequências, uma escolha possível para a matriz  $\mathbf{F}$  seria

$$\mathbf{F}_\ell = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{bmatrix}. \quad (8)$$

Já se for um sistema cuja resposta ao impulso acentua altas frequências, a seguinte matriz pode ser empregada:

$$\mathbf{F}_h = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \end{bmatrix}. \quad (9)$$

Para resolver este problema de otimização, a Eq. (6) deve ser reescrita no seguinte problema de otimização sem restrição:

$$\min \mathcal{F}[\mathbf{w}(k+1)] = \frac{1}{2} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 + \kappa \|\mathbf{F}\mathbf{w}(k+1)\|^2 + \lambda \{e_p(k) - [1 - \gamma(k)] e(k)\}, \quad (10)$$

onde  $\lambda$  é o multiplicador de Lagrange.

Para minimizar a função custo da Eq. (10), cabe calcular o gradiente da função com respeito a  $\mathbf{w}(k+1)$  e igualá-lo a zero. Aplicando a aproximação  $\mathbf{F}^T \mathbf{F}\mathbf{w}(k+1) \approx \mathbf{F}^T \mathbf{F}\mathbf{w}(k)$  [12], é possível mostrar que:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \kappa \mathbf{F}^T \mathbf{F}\mathbf{w}(k) + \lambda \mathbf{x}(k). \quad (11)$$

Fica claro que a Eq. (11) ainda não é a equação de atualização dos coeficientes do filtro adaptativo, uma vez que cabe ainda calcular o multiplicador de Lagrange  $\lambda$ . Para isso, é necessário substituir as Eqs. (6) e (11) na restrição da Eq. (6) e assim obter:

$$e(k) - \gamma(k)e(k) = d(k) - \mathbf{w}^T(k)\mathbf{x}(k) - \lambda \mathbf{x}^T(k)\mathbf{x}(k) - \kappa \mathbf{w}^T(k) \mathbf{F} \mathbf{F}^T \mathbf{x}(k). \quad (12)$$

Note que  $d(k) - \mathbf{w}^T(k)\mathbf{x}(k) = e(k)$  e que  $\mathbf{x}^T(k)\mathbf{x}(k) = \|\mathbf{x}(k)\|^2$ . Portanto, a Eq. (12) pode ser reescrita da seguinte forma:

$$\lambda = \frac{\gamma(k)e(k)}{\|\mathbf{x}(k)\|^2} - \kappa \frac{\mathbf{w}^T(k) \mathbf{F} \mathbf{F}^T \mathbf{x}(k)}{\|\mathbf{x}(k)\|^2}. \quad (13)$$

Desprezando-se o termo  $\kappa \frac{\mathbf{w}^T(k) \mathbf{F} \mathbf{F}^T \mathbf{x}(k)}{\|\mathbf{x}(k)\|^2}$  [8], temos:

$$\lambda = \frac{\gamma(k)e(k)}{\|\mathbf{x}(k)\|^2}. \quad (14)$$

Uma vez calculado  $\lambda$ , basta substituir a Eq. (14) na Eq. (11). Note que todo o desenvolvimento até aqui não faz distinção quanto à normalização do algoritmo. Porém, como o foco deste artigo é o algoritmo  $\ell_2$ -F-LMS, iremos doravante assumir que  $\gamma(k) \triangleq \beta \|\mathbf{x}(k)\|^2$  [11]. Assim, obtemos a seguinte equação de atualização do filtro adaptativo:

$$\mathbf{w}(k+1) = [\mathbf{I} - \kappa \mathbf{F}^T \mathbf{F}] \mathbf{w}(k) + \beta \mathbf{x}(k) e(k). \quad (15)$$

### III. ANÁLISE

Os resultados obtidos em análises que consideram que o filtro adaptativo possui tamanho suficientemente grande não necessariamente refletem o comportamento do filtro em aplicações em que não é possível usar, por razões práticas, um número de coeficientes adequado [13]. No entanto, apesar da grande importância para problemas reais de se prever o comportamento dos algoritmos adaptativos nessas situações, não é comum encontrar análises que considerem o filtro adaptativo de tamanho menor que o sistema a ser identificado [10], [13]–[18]. Este artigo, apresenta uma análise abrangente de um algoritmo da família *Feature LMS*, que contempla tanto o regime de transiente quanto o permanente, e engloba o caso sub-modelado na aplicação de identificação de sistemas.

Na Seção I, foram definidos os vetores  $\mathbf{x}(k)$ ,  $\mathbf{w}(k)$  e  $\mathbf{w}^*$ , contendo as amostras do sinal de entrada, os coeficientes do filtro adaptativo e os coeficientes do filtro a ser identificado, respectivamente. Em um cenário de sub-modelagem, faz-se necessário acrescentar as seguintes definições:

$$\overline{\mathbf{w}}^* \triangleq [w_N^* \ w_{N+1}^* \ \dots \ w_{N+P-1}^*]^\top, \quad (16)$$

$$\overline{\mathbf{x}}(k) \triangleq [x(k-N) \ \dots \ x(k-N-P+1)]^\top, \quad (17)$$

onde os vetores  $\overline{\mathbf{x}}(k)$ ,  $\overline{\mathbf{w}}^* \in \mathbb{R}^P$  contêm, respectivamente, os coeficientes do sistema que o filtro adaptativo não é capaz de modelar e as amostras correspondentes do sinal de entrada. Note que essas definições acarretam em mudanças no cálculo do sinal de referência, que passa a ser escrito como:

$$d(k) = \mathbf{x}^\top(k)\mathbf{w}^* + \overline{\mathbf{x}}^\top(k)\overline{\mathbf{w}}^* + \nu(k) \quad (18)$$

onde  $\nu(k)$  representa um ruído aditivo com as mesmas características definidas na Seção I.

#### A. Análise de primeira ordem

Definindo-se o vetor de desvios dos coeficientes como

$$\tilde{\mathbf{w}}(k+1) \triangleq \mathbf{w}^* - \mathbf{w}(k+1), \quad (19)$$

é possível mostrar que para o algoritmo  $\ell_2$ -F-LMS a equação de atualização de  $\tilde{\mathbf{w}}(k)$  para o caso sub-modelado pode ser escrita como

$$\begin{aligned} \tilde{\mathbf{w}}(k+1) = & [\mathbf{I} - \beta \mathbf{x}(k)\mathbf{x}^\top(k) - \kappa \mathbf{F}^\top \mathbf{F}] \tilde{\mathbf{w}}(k) \\ & + \kappa \mathbf{F}^\top \mathbf{F} \mathbf{w}^* - \beta \mathbf{x}(k)\nu(k) \\ & - \beta \mathbf{x}(k)\overline{\mathbf{x}}^\top(k)\overline{\mathbf{w}}^*. \end{aligned} \quad (20)$$

Para realizar a análise estocástica de primeira ordem, aplica-se o operador de valor esperado em ambos os lados da Eq. (20), obtendo-se:

$$\mathbb{E}[\tilde{\mathbf{w}}(k+1)] = [\mathbf{I} - \beta \mathbf{R}_{\mathbf{x}} - \kappa \mathbf{F}^\top \mathbf{F}] \mathbb{E}[\tilde{\mathbf{w}}(k)] - \beta \mathbf{R}_{\overline{\mathbf{x}}}\overline{\mathbf{w}}^* + \kappa \mathbf{F}^\top \mathbf{F} \mathbf{w}^*, \quad (21)$$

onde  $\mathbf{R}_{\mathbf{x}} \triangleq \mathbb{E}[\mathbf{x}(k)\mathbf{x}^\top(k)]$  e  $\mathbf{R}_{\overline{\mathbf{x}}} \triangleq \mathbb{E}[\mathbf{x}(k)\overline{\mathbf{x}}^\top(k)]$  são as matrizes de autocorrelação e correlação cruzada dos vetores de entrada, respectivamente. Note que a ubíqua hipótese da independência (IA, do inglês *independence assumption* [19], [20]) foi usada no desenvolvimento da Eq. (21). Como resultado, a análise tende a apresentar melhor aderência às curvas empíricas quando o fator de aprendizagem apresenta um valor

relativamente baixo. Caso seja necessário obter maior precisão em configurações nas quais  $\beta$  assume valores altos, poder-se-ia empregar a análise de valor esperado exato [21].

É possível observar que o estimador proposto para o algoritmo apresenta um viés, o qual pode ser calculado aplicando-se o operador de limite, com  $k \rightarrow \infty$ , à Eq. (21), obtendo-se:

$$\lim_{k \rightarrow \infty} \{\mathbb{E}[\tilde{\mathbf{w}}(k)]\} = \left\{ \beta \mathbf{R}_{\mathbf{x}} + \kappa \mathbf{F}^\top \mathbf{F} \right\}^{-1} \left\{ \kappa \mathbf{F}^\top \mathbf{F} \mathbf{w}^* - \beta \mathbf{R}_{\overline{\mathbf{x}}}\overline{\mathbf{w}}^* \right\} \quad (22)$$

Vale a pena ressaltar que se o sinal de entrada for branco, temos  $\mathbf{R}_{\overline{\mathbf{x}}} = 0$ , e o viés do filtro adaptativo tende para o mesmo viés encontrado em [9], caso em que o filtro adaptativo possui o mesmo comprimento que o filtro ótimo. Já quando o sinal é colorido, aparece, somado a esse mesmo viés, um termo de perturbação proveniente da parte do sistema que o filtro adaptativo não é capaz de modelar.

#### B. Análise de Segunda Ordem

Apesar da análise de primeira ordem fornecer informações sobre o comportamento do algoritmo em regime permanente, tal análise não é capaz de oferecer predições acerca da performance dos erros quadráticos ou da estabilidade do algoritmo, fazendo-se necessária também uma análise quadrática média, na qual se está interessado na evolução média da matriz  $\mathbf{\Gamma}(k) = \tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^\top(k)$ . A partir da Eq. (20), é possível mostrar que:

$$\begin{aligned} \mathbf{\Gamma}(k+1) = & \mathbf{A}(k, \beta, \kappa)\mathbf{\Gamma}(k)\mathbf{A}(k, \beta, \kappa) + \mathcal{O}[\nu(k)] \\ & + \kappa \mathbf{A}(k, \beta, \kappa)\tilde{\mathbf{w}}(k)(\mathbf{w}^*)^\top \mathbf{F}^\top \mathbf{F} \\ & + \kappa \mathbf{F}^\top \mathbf{F} \mathbf{w}^* \tilde{\mathbf{w}}^\top(k)\mathbf{A}(k, \beta, \kappa) \\ & - \beta \mathbf{A}(k, \beta, \kappa)\tilde{\mathbf{w}}(k)(\overline{\mathbf{w}}^*)^\top \overline{\mathbf{x}}(k)\mathbf{x}^\top(k) \\ & - \beta \mathbf{x}(k)\overline{\mathbf{x}}^\top(k)\overline{\mathbf{w}}^* \tilde{\mathbf{w}}^\top(k)\mathbf{A}(k, \beta, \kappa) \\ & - \kappa \beta \mathbf{F}^\top \mathbf{F} \mathbf{w}^* (\overline{\mathbf{w}}^*)^\top \overline{\mathbf{x}}(k)\mathbf{x}^\top(k) \\ & - \kappa \beta \mathbf{x}(k)\overline{\mathbf{x}}^\top(k)\overline{\mathbf{w}}^* (\mathbf{w}^*)^\top \mathbf{F}^\top \mathbf{F} \\ & + \beta^2 \mathbf{x}(k)\overline{\mathbf{x}}^\top(k)\overline{\mathbf{w}}^* (\overline{\mathbf{w}}^*)^\top \overline{\mathbf{x}}(k)\mathbf{x}^\top(k) \\ & + \kappa^2 \mathbf{F}^\top \mathbf{F} \mathbf{w}^* (\mathbf{w}^*)^\top \mathbf{F}^\top \mathbf{F} \\ & + \beta^2 \mathbf{x}(k)\mathbf{x}^\top(k)\nu^2(k), \end{aligned} \quad (23)$$

onde

$$\mathbf{A}(k, \beta, \kappa) \triangleq [\mathbf{I} - \beta \mathbf{x}(k)\mathbf{x}^\top(k) - \kappa \mathbf{F}^\top \mathbf{F}] = \mathbf{A}^\top(k, \beta, \kappa) \quad (24)$$

Infelizmente, a Eq. (23) não representa uma recursão implementável para a matriz  $\mathbf{\Gamma}(k)$ . De maneira a contornar tal adversidade, é empregado o operador  $\text{vec}(\cdot)$ , que é um operador onde as colunas de uma determinada matriz são empilhadas de forma que se tenha um vetor coluna como resultado, e o qual possui a seguinte propriedade muito conveniente:

$$\text{vec}(\mathbf{X}\mathbf{Y}\mathbf{Z}) = [\mathbf{Z}^\top \otimes \mathbf{X}] \text{vec}(\mathbf{Y}). \quad (25)$$

Usando a Eq. (25) na Eq. (23), pode ser provado que:

$$\begin{aligned} \text{vec}[\mathbf{\Gamma}(k+1)] = & \mathbf{\Theta}(k, \kappa, \beta) \text{vec}[\mathbf{\Gamma}(k)] + \text{vec}\{\mathcal{O}[\nu(k)]\} \\ & + \kappa \mathbf{\Lambda}_1(k) \text{vec}[\tilde{\mathbf{w}}(k)(\mathbf{w}^*)^\top] \\ & + \kappa \mathbf{\Lambda}_2(k) \text{vec}[(\mathbf{w}^*)^\top \tilde{\mathbf{w}}^\top(k)] \\ & - \beta \mathbf{\Lambda}_3(k) \text{vec}[\tilde{\mathbf{w}}(k)(\overline{\mathbf{w}}^*)^\top] \\ & - \beta \mathbf{\Lambda}_4(k) \text{vec}[\overline{\mathbf{w}}^* \tilde{\mathbf{w}}^\top(k)] \\ & - \kappa \beta \mathbf{\Lambda}_5(k) \text{vec}[\mathbf{w}^* (\overline{\mathbf{w}}^*)^\top] \\ & - \kappa \beta \mathbf{\Lambda}_6(k) \text{vec}[\overline{\mathbf{w}}^* (\mathbf{w}^*)^\top] \\ & + \beta^2 \mathbf{\Lambda}_7(k) \text{vec}[\overline{\mathbf{w}}^* (\overline{\mathbf{w}}^*)^\top] \\ & + \beta^2 \nu^2(k)\mathcal{X}(k) + \kappa^2 \mathbf{\Psi}, \end{aligned} \quad (26)$$

onde

$$\Theta(k, \kappa, \beta) \triangleq \mathbf{A}(k, \kappa, \beta) \otimes \mathbf{A}(k, \kappa, \beta), \quad (27)$$

$$\Lambda_1(k) \triangleq [\mathbf{F}^T \mathbf{F}] \otimes \mathbf{A}(k, \kappa, \beta), \quad (28)$$

$$\Lambda_2(k) \triangleq \mathbf{A}(k, \kappa, \beta) \otimes [\mathbf{F}^T \mathbf{F}], \quad (29)$$

$$\Lambda_3(k) \triangleq [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)] \otimes \mathbf{A}(k, \kappa, \beta), \quad (30)$$

$$\Lambda_4(k) \triangleq \mathbf{A}(k, \kappa, \beta) \otimes [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)], \quad (31)$$

$$\Lambda_5(k) \triangleq [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)] \otimes [\mathbf{F}^T \mathbf{F}], \quad (32)$$

$$\Lambda_6(k) \triangleq [\mathbf{F}^T \mathbf{F}] \otimes [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)], \quad (33)$$

$$\Lambda_7(k) \triangleq [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)] \otimes [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)], \quad (34)$$

$$\mathcal{X}(k) \triangleq \text{vec} [\mathbf{x}(k) \bar{\mathbf{x}}^T(k)], \quad (35)$$

$$\Psi \triangleq [(\mathbf{F}^T \mathbf{F}) \otimes (\mathbf{F}^T \mathbf{F})] \text{vec} \{ \mathbf{w}^* [\mathbf{w}^*]^T \}. \quad (36)$$

Uma vez encontrada a equação recursiva para atualização da matriz de desvios, é necessário aplicar o operador do valor esperado à Eq. (26), obtendo-se:

$$\begin{aligned} \mathbf{v}(k+1) = & \mathbb{E}[\Theta(k, \kappa, \beta)] \mathbf{v}(k) + \sigma_v^2 \beta^2 \mathbb{E}[\mathcal{X}(k)] + \kappa^2 \Psi \\ & + \kappa \mathbb{E}[\Lambda_1(k)] \text{vec} \{ \mathbb{E}[\tilde{\mathbf{w}}(k)] (\mathbf{w}^*)^T \} \\ & + \kappa \mathbb{E}[\Lambda_2(k)] \text{vec} \{ \mathbf{w}^* \mathbb{E}[\tilde{\mathbf{w}}^T(k)] \}, \\ & - \beta \mathbb{E}[\Lambda_3(k)] \text{vec} \{ \mathbb{E}[\tilde{\mathbf{w}}(k)] (\bar{\mathbf{w}}^*)^T \} \\ & - \beta \mathbb{E}[\Lambda_4(k)] \text{vec} \{ \bar{\mathbf{w}}^* \mathbb{E}[\tilde{\mathbf{w}}^T(k)] \} \\ & - \kappa \beta \mathbb{E}[\Lambda_5(k)] \text{vec} \{ \mathbf{w}^* (\bar{\mathbf{w}}^*)^T \} \\ & - \kappa \beta \mathbb{E}[\Lambda_6(k)] \text{vec} \{ \bar{\mathbf{w}}^* (\mathbf{w}^*)^T \} \\ & + \beta^2 \mathbb{E}[\Lambda_7(k)] \text{vec} \{ \bar{\mathbf{w}}^* (\bar{\mathbf{w}}^*)^T \}, \end{aligned} \quad (37)$$

onde  $\mathbf{v}(k) \triangleq \mathbb{E}[\text{vec}[\Gamma(k)]]$ . Note que como  $\text{vec}(\mathbf{X})$  é um operador bijetor, a matriz  $\mathbf{R}_{\tilde{\mathbf{w}}}(k)$ , onde  $\mathbf{R}_{\tilde{\mathbf{w}}}(k) \triangleq \mathbb{E}[\tilde{\mathbf{w}}(k) \tilde{\mathbf{w}}^T(k)]$ , pode ser facilmente recuperada a partir de  $\mathbf{v}(k)$ . Observe que as hipóteses IA e *noise assumption* (NA, [1]) foram usadas na derivação da Eq. (37). Note também que a Eq. (37) depende do termo  $\mathbb{E}[\tilde{\mathbf{w}}(k)]$ , o qual pode ser encontrado pela Eq. (21).

Com a equação estatística acerca da evolução da matriz de autocorrelação dos desvios, é possível estimar as evoluções do MSE e do MSD ao longo do tempo pelas expressões

$$\text{MSE}(k) = \sigma_v^2 + \text{Tr}[\mathbf{R}_{\tilde{\mathbf{w}}} \mathbf{R}_{\tilde{\mathbf{w}}}(k)] + (\bar{\mathbf{w}}^*)^T \mathbf{R}_{\tilde{\mathbf{w}}} \bar{\mathbf{w}}^* + 2 \mathbb{E}[\tilde{\mathbf{w}}(k)] \mathbf{R}_{\tilde{\mathbf{w}}} \bar{\mathbf{w}}^*, \quad (38)$$

e

$$\text{MSD}(k) = \text{Tr}[\mathbf{R}_{\tilde{\mathbf{w}}}(k)], \quad (39)$$

onde  $\mathbf{R}_{\tilde{\mathbf{w}}} = \mathbb{E}[\tilde{\mathbf{x}}(k) \tilde{\mathbf{x}}^T(k)]$  é a matriz de autocorrelação de dimensão  $(P) \times (P)$  do sinal de entrada.

#### IV. RESULTADOS

Nesta seção, são feitas simulações acerca da análise do algoritmo  $\ell_2$ -F-LMS para o caso submodelado. Tais simulações consistem em identificar um sistema FIR passa-baixas com 20 coeficientes dados por:

$$w_{i,\ell}^* = \begin{cases} 0, & \text{if } 0 \leq i \leq 9 \\ 0.05(i-9), & \text{if } 10 \leq i \leq 14 \\ 0.3, & \text{if } 15 \leq i \leq 19 \end{cases}, \quad (40)$$

Como o filtro ótimo escolhido para realizar as simulações é um filtro passa-baixas, a matriz de pesos escolhida é  $\mathbf{F}_l$ . Note que essa matriz penaliza coeficientes vizinhos cuja diferença é grande, que é uma característica contrária a filtros passa-baixas

que possuem variações suaves entre coeficientes próximos. O sinal de entrada é obtido filtrando-se um sinal branco, gaussiano, e de variância unitária por  $H(z) = 1 + 0.8z^{-1} - 0.2z^{-2}$ . As curvas de aprendizado são obtidas fazendo a média dos resultados provenientes de 1000 simulações independentes de Monte Carlo.

A Fig. 2 mostra as evoluções médias de três coeficientes obtidas a partir da análise teórica de primeira ordem e das simulações. Nesse experimento, foram utilizados os parâmetros  $\beta = 0.015$ ,  $\kappa = 0.0316$ ,  $\sigma_v^2 = 0.1$ . A Fig. 2(a) mostra o resultado obtido com  $P = 1$ , enquanto que a Fig. 2(b) mostra o resultado obtido com  $P = 4$ . Comparando os resultados teóricos e experimentais da Fig. 4, pode-se concluir que a análise de primeira ordem apresentada é capaz de prever a evolução média dos coeficientes de forma acurada para o caso sub-modelado de identificação de sistemas.

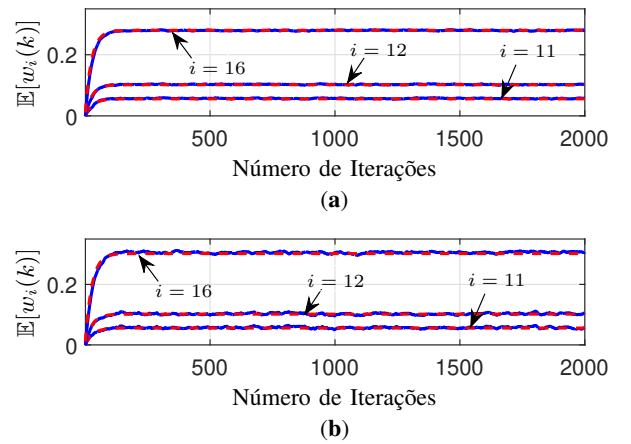


Fig. 2. Evoluções médias teóricas (em vermelho) e simuladas (em azul) dos coeficientes  $w_i(k)$ ,  $i \in \{11, 12, 16\}$ : (a)  $P = 1$ ; (b)  $P = 4$ .

A Fig. 3 mostra a evoluções teóricas e experimentais do MSE ao longo das iterações. Nessas simulações, foram utilizados os parâmetros  $P = 1$ ,  $\kappa = 0.0316$ ,  $\sigma_v^2 = 0.1$  e ótimo  $w_\ell^*$  em ambos os casos. A Fig. 3(a) mostra o resultado obtido com  $\beta = 0.015$ , enquanto que a Fig. 3(b) mostra o resultado obtido com  $\beta = 0.0015$ . A Fig. 4 mostra as evoluções do MSE ao longo das iterações, com os parâmetros  $P = 4$ ,  $\kappa = 0.0316$ ,  $\sigma_v^2 = 0.1$  e ótimo  $w_\ell^*$ . A Fig. 4(a) mostra os resultados obtidos com  $\beta = 0.015$ , enquanto que a Fig. 4(b) mostra os resultados obtidos com  $\beta = 0.0015$ .

A partir dos resultados mostrados nas Figs. 3 e 4, pode-se concluir que o modelo teórico obtido da análise de segunda ordem para o caso sub-modelado é capaz de prever a curva de aprendizado de maneira fidedigna com os resultados experimentais.

#### V. CONCLUSÕES

Neste trabalho, foi proposto um modelo estocástico com objetivo de estudar o comportamento do algoritmo quando há restrições quanto ao comprimento do filtro adaptativo, não sendo este capaz de ter o mesmo comprimento do filtro ótimo. Pelos experimentos realizados neste trabalho, é possível

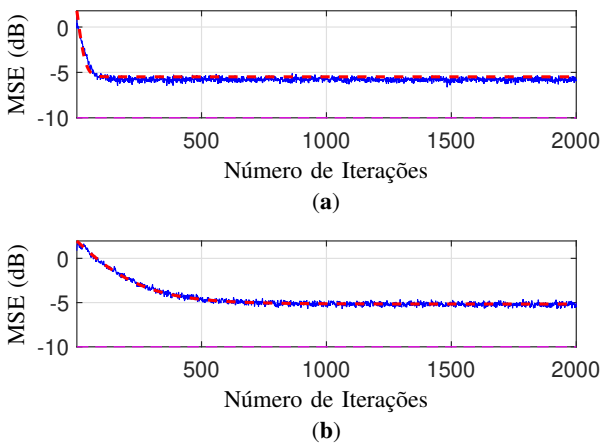


Fig. 3. Variância do ruído aditivo (linha magenta tracejada). Evoluções teóricas (em vermelho) e simuladas (em azul) do MSE (em dB) com  $P = 1$ : (a)  $\beta = 0.015$ ; (b)  $\beta = 0.0015$ .

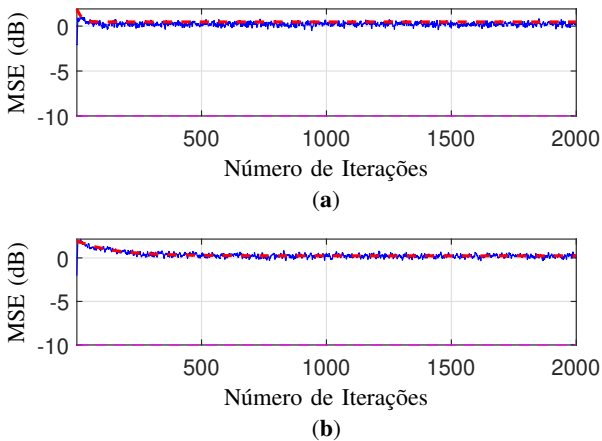


Fig. 4. Variância do ruído aditivo (linha magenta tracejada). Evoluções teóricas (em vermelho) e simuladas (em azul) do MSE (em dB) com  $P = 4$ : (a)  $\beta = 0.015$ ; (b)  $\beta = 0.0015$ .

concluir que o modelo teórico é capaz de prever o comportamento do algoritmo de maneira satisfatória e que as curvas de aprendizado são fidedignas às experimentais.

#### REFERÊNCIAS

[1] K. da S. Olinto, D. B. Haddad, and M. R. Petraglia, “Transient analysis of  $l_0$ -LMS and  $l_0$ -NLMS algorithms,” *Signal Processing*, vol. 127, pp. 217 – 226, Jan 2016.

[2] P. S. R. Diniz, H. Yazdanpanah, and M. V. S. Lima, “Feature LMS algorithms,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4144–4148, April 2018.

[3] D. B. Haddad and M. R. Petraglia, “Transient and steady-state MSE analysis of the IMPNLMS algorithm,” *Digital Signal Processing*, vol. 33, pp. 50 – 59, July 2014.

[4] Y. Gu, J. Jin, and S. Mei, “ $l_0$ -norm constraint LMS algorithm for sparse system identification,” *IEEE Signal Processing Letters*, vol. 16, pp. 774–777, Sept 2009.

[5] N. Kalouptsidis, G. Mileounis, B. Babadi, and V. Tarokh, “Adaptive algorithms for sparse system identification,” *Signal Processing*, vol. 91, pp. 1910 – 1919, Feb 2011.

[6] H. Yazdanpanah, J. Apolinario, P. Diniz, and M. Lima, “ $l_0$ -norm Feature LMS algorithms,” in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 311–315, IEEE, Nov 2018.

[7] G. Su, J. Jin, Y. Gu, and J. Wang, “Performance analysis of  $l_0$ -norm constraint least mean square algorithm,” *IEEE Transactions on Signal Processing*, vol. 60, pp. 2223–2235, May 2012.

[8] D. B. Haddad, M. R. Petraglia, and A. Petraglia, “A unified approach for sparsity-aware and maximum correntropy adaptive filters,” in *2016 24th European Signal Processing Conference (EUSIPCO)*, pp. 170–174, Aug 2016.

[9] D. Haddad, L. Santos, L. Almeida, G. Santos, and M. Petraglia, “ $l_2$ -norm Feature LMS algorithm,” *Electronics Letters*, Feb 2020.

[10] P. Lara, L. D. Tarrataca, and D. B. Haddad, “Exact expectation analysis of the deficient-length LMS algorithm,” *Signal Processing*, vol. 162, pp. 54 – 64, Apr 2019.

[11] A. H. Sayed, *Adaptive Filters*. John Wiley & Sons, 1 ed., 2011.

[12] M. Lima, I. Sobron, W. Martins, and P. Diniz, “Stability and MSE analyses of affine projection algorithms for sparse system identification,” *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 6399–6403, 05 2014.

[13] K. Mayyas, “Performance analysis of the deficient-length LMS adaptive algorithm,” *Signal Processing, IEEE Transactions on*, vol. 53, pp. 2727 – 2734, Sep 2005.

[14] R. Bilcu, P. Kuosmanen, and K. Egiazarian, “A new variable length LMS algorithm: Theoretical analysis and implementations,” in *9th International conference on Electronics, Circuits and Systems*, vol. 3, pp. 1031 – 1034 vol.3, IEEE, Feb 2002.

[15] Y. Gu, K. Tang, H. Cui, and W. Du, “Convergence analysis of a deficient-length LMS filter and optimal-length sequence to model exponential decay impulse response,” *Signal Processing Letters, IEEE*, vol. 10, pp. 4 – 7, Feb 2003.

[16] M. Wu and J. Yang, “A step-size control method for deficient length FBLMS algorithm,” *IEEE Signal Processing Letters*, vol. 22, pp. 1–1, Jan 2014.

[17] Y. Xia, S. Douglas, and D. Mandic, “Performance analysis of the deficient-length augmented CLMS algorithm for second order noncircular complex signals,” *Signal Processing*, vol. 144, Oct 2017.

[18] C. Paleologu, S. Ciochina, and J. Benesty, “Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation,” *Signal Processing Letters, IEEE*, vol. 15, pp. 5 – 8, Feb 2008.

[19] P. S. R. Diniz, “Adaptive filtering: Algorithms and practical implementation,” *The international series in Engineering and Computer Scienc*, 2008.

[20] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2008.

[21] P. Lara, F. Igreja, L. D. T. J. Tarrataca, D. B. Haddad, and M. R. Petraglia, “Exact expectation evaluation and design of variable step-size adaptive algorithms,” *IEEE Signal Processing Letters*, vol. 26, pp. 74–78, Jan 2019.