

Análise Estatística do Algoritmo LMS Largamente Linear e Cálculo do Atraso Ótimo para Equalização Largamente Linear.

Francisco J. A. de Aquino, Carlos A. F. da Rocha e Leonardo S. Resende

Resumo – Este artigo apresenta uma análise estatística do algoritmo LMS largamente linear. Modelos analíticos para o comportamento do vetor de coeficientes e do erro quadrático médio são derivados. Estes novos modelos permitem a determinação do comportamento transitório e o estado permanente do algoritmo que podem ser úteis para propósitos de projeto. Além disso, são desenvolvidas expressões para o cálculo do atraso ótimo no caso de uma equalização largamente linear de um canal de comunicação. Simulações de Monte Carlo mostram uma excelente concordância com os resultados teóricos na predição do comportamento do algoritmo.

Palavras-Chave – Algoritmo LMS, processamento largamente linear, filtragem adaptativa, equalização de canal.

Abstract — This paper presents a statistical analysis of the widely-linear LMS algorithm. Analytical models are derived for the mean weight-error vector and the mean square error behaviors. These new models allow the determination of the transient behavior and steady-state results that should prove useful for design purposes. Furthermore, expressions are developed for calculating the optimum delay in the case of a widely linear communication channel equalization. Simulation experiments show excellent agreement with theoretical results in predicting the convergence behavior of the algorithm.

Keywords— LMS algorithm, widely linear processing, adaptive filtering, channel equalization.

I. INTRODUÇÃO

Em algumas aplicações de filtragem adaptativa, tais como a transmissão de dados reais (M -PAM ou BPSK) ou OQAM (*offset quadrature amplitude modulation*) por um canal complexo, o sinal recebido é impróprio [1].

Um vetor de sinal complexo \mathbf{x} é impróprio (rotacionalmente variante ou não circular) se a sua matriz de pseudo-covariância $E_{\mathbf{x}} = E\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^T\}$ é não nula [2, 3], onde $(\cdot)^T$ denota a operação de transposição e $E\{\cdot\}$ denota a operação de esperança matemática. Quando o sinal é impróprio, os filtros largamente lineares podem alcançar um desempenho significativamente maior que uma filtragem linear convencional [1, 2, 3, 4]. Nestes casos, o algoritmo LMS (*least mean square*) largamente linear (LMS-LL) é atualmente o método mais popular para ajustar os coeficientes do filtro [1, 4, 5, 6].

F. J. A. de Aquino é professor no CEFET-Ce e aluno de doutorado na UFSC (GPqCom), e-mail: fcoalthes_aq@cefet-ce.br. C. A. F. da Rocha e Leonardo S. Resende são professores no Departamento de Engenharia Elétrica da UFSC, laboratório GPqCom. E-mails: {aurelio, leonardo}@eel.ufsc.br.

O projeto de um filtro adaptativo é facilitado pela disponibilidade de modelos analíticos que são capazes de prever o comportamento do algoritmo utilizado na atualização dos parâmetros do filtro.

Este artigo provê uma análise estatística do algoritmo LMS-LL. Um modelo simples é obtido, permitindo a determinação das condições necessárias para a estabilidade do algoritmo, o seu comportamento transiente e de regime permanente. Estes resultados teóricos novos podem servir, por exemplo, como diretrizes fidedignas para o projeto de equalizadores largamente lineares. Adicionalmente, apresenta-se uma análise do atraso ótimo para o caso de equalização largamente linear usando o algoritmo LMS-LL.

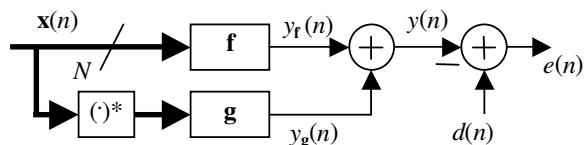


Fig. 1. Filtragem transversal largamente linear.

II. FILTRAGEM LARGAMENTE LINEAR

A Figura 1 ilustra o esquema de filtragem largamente linear não adaptativa, onde \mathbf{f} e \mathbf{g} são filtros transversais representados por vetores de coeficientes $N \times 1$, $(\cdot)^*$ indica a operação de conjugação complexa e $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$ o vetor de entrada. O sinal de entrada $x(n)$ e o sinal desejado $d(n)$ são modelados como processos estocásticos estacionários, gaussianos e de média nula, com variância σ_x^2 e σ_d^2 , respectivamente.

O erro de estimação é dado por:

$$\begin{aligned} e(n) &= d(n) - y(n) \\ &= d(n) - y_{\mathbf{f}}(n) - y_{\mathbf{g}}(n) \\ &= d(n) - \mathbf{w}^H \mathbf{u}(n), \end{aligned} \quad (1)$$

onde

$$\mathbf{u}(n) = \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{x}^*(n) \end{bmatrix}_{2N \times 1}, \quad (2)$$

$$\mathbf{w} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}_{2N \times 1}, \quad (3)$$

e H denota transposição Hermitiana (i.e., transposição combinada com conjugação complexa).

O vetor de coeficientes \mathbf{w} é escolhido de tal forma que o erro quadrático médio (MSE) seja minimizado:

$$J(\mathbf{w}) = E\left\{e(n)^2\right\} = \sigma_d^2 - \mathbf{v}^H \mathbf{w} - \mathbf{w}^H \mathbf{v} + \mathbf{w}^H \mathbf{U} \mathbf{w}, \quad (4)$$

onde $\sigma_d^2 = E\{|d(n)|^2\}$ é a variância de $d(n)$,

$$\mathbf{U} = E\{\mathbf{u}(n)\mathbf{u}^H(n)\} = \begin{bmatrix} \mathbf{R} & \mathbf{S} \\ \mathbf{S}^* & \mathbf{R}^* \end{bmatrix}_{2N \times 2N}, \quad (5)$$

$\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^H(n)\}$ é a matriz de autocorrelação $N \times N$ de $\mathbf{x}(n)$, $\mathbf{S} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$ é a matriz de pseudoautocorrelação $N \times N$ de $\mathbf{x}(n)$,

$$\mathbf{v} = E\{\mathbf{u}(n)d^*(n)\} = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix}_{2N \times 1}, \quad (6)$$

$\mathbf{p} = E\{\mathbf{x}(n)d^*(n)\}$ é vetor $N \times 1$ de correlação cruzada entre $\mathbf{x}(n)$ e $d(n)$, e $\mathbf{q} = E\{\mathbf{x}^*(n)d^*(n)\}$ é o vetor $N \times 1$ de pseudo correlação cruzada entre o vetor $\mathbf{x}(n)$ e $d(n)$. O vetor ótimo é dado por [1, 2]:

$$\mathbf{w}_{opt} = \mathbf{U}^{-1} \mathbf{v}, \quad (7)$$

assumindo que \mathbf{U} é não singular. Substituindo (7) em (4) chega-se ao MSE mínimo:

$$J_{min} = \sigma_d^2 - \mathbf{v}^H \mathbf{w}_{opt} = \sigma_d^2 - \mathbf{v}^H \mathbf{U}^{-1} \mathbf{v}. \quad (8)$$

III. ALGORITMO LMS LARGAMENTE LINEAR

O algoritmo adaptativo para atualizar o vetor de coeficientes \mathbf{w} é muito similar ao algoritmo LMS convencional. O gradiente instantâneo do vetor de erro quadrático $J(n) = |e(n)|^2$ é dado por:

$$\begin{aligned} \nabla J(n) &= -\mathbf{u}(n)d^*(n) + \mathbf{u}(n)\mathbf{x}^H(n)\mathbf{w}(n) \\ &= -\mathbf{u}(n)(d^*(n) - \mathbf{u}^H(n)\mathbf{w}(n)) \\ &= -\mathbf{u}(n)e^*(n), \end{aligned} \quad (9)$$

e a relação recursiva para atualização do vetor de coeficientes do algoritmo LMS-LL é dada como segue [1,7]:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu e^*(n)\mathbf{u}(n), \quad (10)$$

onde $e(n) = d(n) - \mathbf{w}(n-1)^H \mathbf{u}(n)$ é o erro de estimação e μ é o passo de adaptação do algoritmo.

IV. ANÁLISE ESTATÍSTICA

Nesta seção é detalhada a análise estatística do algoritmo LMS-LL. Em [7], é feita a mesma análise para o algoritmo LMS, mas considerando que o passo de adaptação é “pequeno”. Aqui, esta hipótese não será considerada, podendo o passo de adaptação assumir valores “grandes”, ou seja, próximo a um valor máximo.

A. Comportamento médio do vetor de coeficientes

Usando (1) e calculando o valor esperado em ambos os lados de (10) encontra-se

$$\begin{aligned} E\{\mathbf{w}(n)\} &= E\{\mathbf{w}(n-1)\} + \mu E\{\mathbf{u}(n)e^*(n)\} \\ &= E\{\mathbf{w}(n-1)\} + \mu E\{\mathbf{u}(n)d^*(n)\} \\ &\quad - \mu E\{\mathbf{u}(n)\mathbf{u}^H(n)\mathbf{w}(n-1)\}. \end{aligned} \quad (11)$$

Desprezando a dependência estatística entre o vetor de coeficientes e o vetor de sinal, de (11) obtém-se:

$$\begin{aligned} E\{\mathbf{w}(n)\} &= E\{\mathbf{w}(n-1)\} + \mu \mathbf{v} - \mu \mathbf{U} E\{\mathbf{w}(n-1)\} \\ &= (\mathbf{I} - \mu \mathbf{U}) E\{\mathbf{w}(n-1)\} + \mu \mathbf{v}, \end{aligned} \quad (12)$$

onde \mathbf{I} é uma matriz identidade. A Equação (12) é uma recursão determinística para o comportamento médio do vetor de coeficientes do algoritmo LMS-LL. Fazendo $n \rightarrow \infty$, o valor em regime permanente deste vetor é dado por:

$$\mathbf{w}_\infty = \lim_{n \rightarrow \infty} E\{\mathbf{w}(n)\} = \mathbf{U}^{-1} \mathbf{v} = \mathbf{w}_{opt}, \quad (13)$$

o qual corresponde ao filtro ótimo em (7). Definindo o vetor de erro como

$$\boldsymbol{\varepsilon}(n) = \mathbf{w}(n) - \mathbf{w}_{opt}, \quad (14)$$

e substituindo por $\mathbf{w}(n)$ em (10) e (12) chega-se a

$$\boldsymbol{\varepsilon}(n) = \boldsymbol{\varepsilon}(n-1) + \mu e^*(n)\mathbf{u}(n) \quad (15)$$

e

$$E\{\boldsymbol{\varepsilon}(n)\} = (\mathbf{I} - \mu \mathbf{U}) E\{\boldsymbol{\varepsilon}(n-1)\} = (\mathbf{I} - \mu \mathbf{U})^n \boldsymbol{\varepsilon}(0), \quad (16)$$

que é uma recursão para o comportamento médio do vetor de erro. Usando a transformação de similaridade para a matriz \mathbf{U} , obtém-se:

$$\begin{aligned} E\{\boldsymbol{\varepsilon}(n)\} &= (\mathbf{I} - \mu \mathbf{Q} \boldsymbol{\Lambda} \mathbf{Q}^H)^n \boldsymbol{\varepsilon}(0) \\ &= \mathbf{Q} (\mathbf{I} - \mu \boldsymbol{\Lambda})^n \mathbf{Q}^H \boldsymbol{\varepsilon}(0), \end{aligned} \quad (17)$$

onde $\boldsymbol{\Lambda}$ é uma matriz diagonal consistindo dos autovalores de \mathbf{U} , que estão associados aos autovetores da matriz unitária \mathbf{Q} .

Do equacionamento acima, segue que a condição para a convergência do algoritmo LMS-LL é

$$0 < \mu < \frac{2}{\lambda_{max}}, \quad (18)$$

onde λ_{max} é o maior autovalor da matriz \mathbf{U} .

B. Comportamento médio do erro quadrático

Usando (14), o erro de estimação pode ser descrito como

$$e(n) = d(n) - (\boldsymbol{\varepsilon}(n-1) + \mathbf{w}_{opt})^H \mathbf{u}(n). \quad (19)$$

A partir da equação acima, pode ser mostrado que

$$\begin{aligned}
 |e(n)|^2 &= d(n)d^*(n) - d(n)\mathbf{u}^H(n)\boldsymbol{\varepsilon}(n-1) - d(n)\mathbf{u}^H(n)\mathbf{w}_{opt} \\
 &\quad - \boldsymbol{\varepsilon}^H(n-1)\mathbf{u}(n)d^*(n) + \boldsymbol{\varepsilon}^H(n-1)\mathbf{u}(n)\mathbf{u}^H(n)\boldsymbol{\varepsilon}(n-1) \\
 &\quad + \boldsymbol{\varepsilon}^H(n-1)\mathbf{u}(n)\mathbf{u}^H(n)\mathbf{w}_{opt} - \mathbf{w}_{opt}^H\mathbf{u}(n)d^*(n) \\
 &\quad + \mathbf{w}_{opt}^H\mathbf{u}(n)\mathbf{u}^H(n)\boldsymbol{\varepsilon}(n-1) + \mathbf{w}_{opt}^H\mathbf{u}(n)\mathbf{u}^H(n)\mathbf{w}_{opt}.
 \end{aligned} \tag{20}$$

Calculando o valor esperado em ambos os lados de (20) e desprezando as dependências estatísticas de $d(n)$ e $\mathbf{u}(n)$ com respeito a $\boldsymbol{\varepsilon}(n-1)$, obtém-se

$$\begin{aligned}
 E\{|e(n)|^2\} &= \sigma_d^2 - \mathbf{v}^H E\{\boldsymbol{\varepsilon}(n-1)\} - \mathbf{v}^H \mathbf{w}_{opt} - E\{\boldsymbol{\varepsilon}^H(n-1)\} \mathbf{v} \\
 &\quad - E\{\boldsymbol{\varepsilon}^H(n-1)\} \mathbf{v} + tr(\mathbf{U}\mathbf{K}(n-1)) + E\{\boldsymbol{\varepsilon}^H(n-1)\} \mathbf{U} \mathbf{w}_{opt} \\
 &\quad - \mathbf{w}_{opt}^H \mathbf{v} + \mathbf{w}_{opt}^H \mathbf{U} E\{\boldsymbol{\varepsilon}(n-1)\} + \mathbf{w}_{opt}^H \mathbf{U} \mathbf{w}_{opt}.
 \end{aligned} \tag{21}$$

onde $\mathbf{K}(n-1) = E\{\boldsymbol{\varepsilon}(n-1)\boldsymbol{\varepsilon}^H(n-1)\}$ é a matriz de correlação do vetor de erro $\boldsymbol{\varepsilon}(n-1)$. Pós-multiplicando (15) por sua transposta e calculando o valor esperado tem-se:

$$\begin{aligned}
 \mathbf{K}(n-1) &= \mathbf{K}(n-2) + \mu E\{\boldsymbol{\varepsilon}(n-2)\mathbf{u}^H(n-1)e(n-1)\} \\
 &\quad + \mu E\{e^*(n-1)\mathbf{u}(n-1)\boldsymbol{\varepsilon}^H(n-2)\} \\
 &\quad + \mu^2 E\{\mathbf{u}(n-1)e^*(n-1)\mathbf{u}^H(n-1)e(n-1)\}.
 \end{aligned} \tag{22}$$

Assumindo que $\mathbf{u}(n-1)$ e $e(n-1)$ são conjuntamente gaussianos de média zero, a última parcela do valor esperado em (22) pode ser calculado usando o teorema da fatoração dos momentos de variáveis conjuntamente gaussianas [8]. Negligenciando novamente a dependência estatística de $e(n-1)$ e $\mathbf{u}(n-1)$ com $\boldsymbol{\varepsilon}(n-2)$, obtém-se:

$$\begin{aligned}
 \mathbf{K}(n-1) &= \mathbf{K}(n-2) - \mu \mathbf{K}(n-2) \mathbf{U} - \mu \mathbf{U} \mathbf{K}(n-2) \\
 &\quad + \mu^2 \left(\mathbf{U} E\{\boldsymbol{\varepsilon}(n-2)\} E\{\boldsymbol{\varepsilon}^H(n-2)\} + E\{|e(n-1)|^2\} \mathbf{U} \right)
 \end{aligned} \tag{23}$$

onde

$$\begin{aligned}
 E\{|e(n-1)|^2\} &= J_{min} + tr(\mathbf{U}\mathbf{K}(n-2)) \\
 &= J_{min} + J_{ex}(n-1),
 \end{aligned} \tag{24}$$

$J_{ex}(n-1)$ corresponde ao erro quadrático médio em excesso no instante $n-1$ e $tr(\cdot)$ denota o traço de (\cdot) . Segue que (24) pode ser escrita como:

$$\begin{aligned}
 \mathbf{K}(n-1) &= \mathbf{K}(n-2) - \mu (\mathbf{K}(n-2) \mathbf{U} + \mathbf{U} \mathbf{K}(n-2)) \\
 &\quad + \mu^2 \left(\mathbf{U} E\{\boldsymbol{\varepsilon}(n-2)\} E\{\boldsymbol{\varepsilon}^H(n-2)\} \mathbf{U} \right) \\
 &\quad + \mu^2 \left(J_{min} + tr(\mathbf{U}\mathbf{K}(n-2)) \right) \mathbf{U}.
 \end{aligned} \tag{25}$$

Então, (25) é uma recursão para matriz de correlação do vetor de erro de coeficientes.

Fazendo $n \rightarrow \infty$ em (24), obtém-se o erro quadrático médio em regime permanente, que é dado por

$$\begin{aligned}
 J_{\infty} &= \lim_{n \rightarrow \infty} E\{e^2(n-1)\} \\
 &= J_{min} + tr(\mathbf{U}\mathbf{K}_{\infty}) \\
 &= J_{min} + J_{ex}(\infty),
 \end{aligned} \tag{26}$$

onde $\mathbf{K}_{\infty} = \lim_{n \rightarrow \infty} \mathbf{K}(n-1)$. De (25), obtém-se então que o erro quadrático médio em excesso é

$$\begin{aligned}
 J_{ex}(\infty) &= tr(\mathbf{U}\mathbf{K}_{\infty}) \\
 &= \frac{\mu tr(\mathbf{U})}{2 - \mu tr(\mathbf{U})} J_{min}
 \end{aligned} \tag{27}$$

e

$$J_{\infty} = J_{min} \left(1 + \frac{\mu tr(\mathbf{U})}{2 - \mu tr(\mathbf{U})} \right). \tag{28}$$

Finalmente, o desajuste é dado por

$$\xi = \frac{J_{ex}(\infty)}{J_{min}} = \frac{\mu tr(\mathbf{U})}{2 - \mu tr(\mathbf{U})}. \tag{29}$$

Desde que ξ precisa ser um valor finito, o limitante superior de μ para a estabilidade do algoritmo LMS LL é

$$\mu_{max} < \frac{2}{tr(\mathbf{U})}. \tag{30}$$

A Tabela I apresenta o sumário do modelo desenvolvido.

Tabela I. Modelo Largamente Linear

Inicialização:

$$\mathbf{w}(0) = \mathbf{0}, \boldsymbol{\varepsilon}(0) = -\mathbf{w}_{opt}, 0 < \mu < \frac{2}{tr(\mathbf{U})}, \mathbf{K}(0) = \mathbf{w}_{opt} \mathbf{w}_{opt}^H$$

Regime permanente:

$$\mathbf{w}_{\infty} = \mathbf{w}_{opt}, J_{\infty} = J_{min} + J_{ex}(\infty), \xi = \frac{\mu tr(\mathbf{U})}{2 - \mu tr(\mathbf{U})}, \boldsymbol{\varepsilon}(\infty) = \mathbf{0}$$

Comportamento transiente:

$$\begin{aligned}
 E\{\mathbf{w}(n)\} &= (\mathbf{I} - \mu \mathbf{U}) E\{\mathbf{w}(n-1)\} + \mu \mathbf{v}, \\
 E\{\boldsymbol{\varepsilon}(n)\} &= (\mathbf{I} - \mu \mathbf{U})^n \boldsymbol{\varepsilon}(0), \\
 \mathbf{K}(n-1) &= \mathbf{K}(n-2) - \mu (\mathbf{K}(n-2) \mathbf{U} + \mathbf{U} \mathbf{K}(n-2)) \\
 &\quad + \mu^2 \left(\mathbf{U} E\{\boldsymbol{\varepsilon}(n-2)\} E\{\boldsymbol{\varepsilon}^H(n-2)\} \mathbf{U} \right) \\
 &\quad + \mu^2 \left(J_{min} + tr(\mathbf{U}\mathbf{K}(n-2)) \right) \mathbf{U}, \\
 E\{|e(n-1)|^2\} &= J_{min} + J_{ex}(n-1)
 \end{aligned}$$

Quando aplicado à equalização com treinamento, o sinal desejado $d(n)$ é, em geral, uma versão atrasada do sinal transmitido $a(n)$. O desempenho do equalizador é fortemente influenciado por este atraso [9]. Por este motivo, na próxima seção, é feita uma análise do atraso ótimo para a equalização adaptativa usando um algoritmo LMS LL. Essa análise é inspirada em [10], onde é encontrada uma expressão fechada para o cálculo do atraso ótimo no caso de um receptor usando equalização cega fracionalmente espaçada.

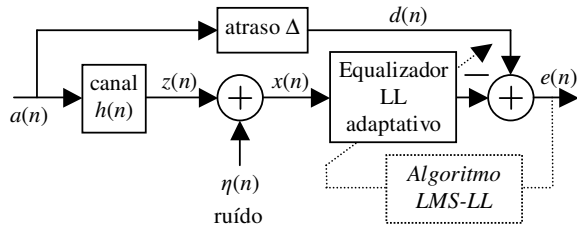


Fig. 2. Equalizador LL adaptativo treinado.

V. ATRASO ÓTIMO PARA UM EQUALIZADOR LL

Em sistemas de comunicação é de fundamental importância o combate à interferência entre símbolos (IES) provocada pelo canal de comunicação [11]. A redução da IES pode ser conseguida pelo emprego de equalizadores largamente lineares adaptativos [1].

Para o cálculo do atraso ótimo Δ do equalizador LL adaptativo, ver Fig. 2, será considerado um canal complexo de ordem L conhecido cuja resposta ao impulso é dada por $\mathbf{h} = [h_0, h_1, \dots, h_L]$. Por questão de simplicidade, na análise a seguir será considerado que o canal é fixo. Note que esse atraso ótimo leva em conta o atraso total do sistema canal + equalizador.

Pode-se expressar o sinal de entrada $x(n)$ do equalizador LL, ver Fig. 2, por:

$$x(n) = \sum_{k=0}^L h_k a(n-k) + \eta(n), \quad (31)$$

sendo $\eta(n)$ um ruído aditivo complexo, gaussiano, de média zero com variância σ_η^2 e $a(n)$ um sinal pertencente a uma constelação real e variância σ_a^2 . O vetor de dados $\mathbf{x}(n)$ pode ser expresso por

$$\mathbf{x}(n) = \mathbf{H}\mathbf{a}(n) + \boldsymbol{\eta}(n), \quad (32)$$

onde

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \dots & h_L & 0 & \dots & 0 \\ 0 & h_0 & h_1 & \dots & h_L & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_0 & h_1 & \dots & h_L \end{bmatrix} \quad (33)$$

é a matriz de convolução do canal com dimensão $N \times (L+N)$, $\mathbf{a}(n) = [a(n) \ a(n-1) \ \dots \ a(n-L-N+1)]^T$ é o vetor de dados com dimensão $(L+N+1) \times 1$ e $\boldsymbol{\eta}(n) = [\eta(n) \ \eta(n-1) \ \dots \ \eta(n-L-N+1)]^T$ é o vetor de ruído de mesma dimensão de $\mathbf{a}(n)$. É considerado que $N > L$. Note que, devido à operação de conjugação complexa, o sinal a ser processado pelo equalizador LL é $\mathbf{u}(n)$, já definido em (2).

A. Atraso ótimo para equalização ZF

Quando a relação sinal-ruído é alta, o ruído pode ser desprezado e o equalizador LL pode ser visto como um equalizador ZF (*zero forcing*). Neste caso, deseja-se que a convolução do canal com o equalizador resulte em um impulso atrasado. Sob algumas condições, é possível que o equalizador LL ZF equalize perfeitamente o canal [4, 12].

A convolução do canal com o equalizador LL pode ser expressa por:

$$\mathbf{H}_c \mathbf{w}_z = \mathbf{h}_\Delta, \quad (34)$$

onde $\mathbf{H}_c = [\mathbf{H}^T \ \mathbf{H}^H]$, com \mathbf{H} definido em (33), \mathbf{w}_z é um vetor coluna com os coeficientes do equalizador segundo o critério ZF e \mathbf{h}_Δ é um vetor coluna com um elemento unitário na posição Δ e zeros nas demais linhas, isto é, \mathbf{h}_Δ representa um atraso Δ . Note que \mathbf{H}_c é uma matriz retangular e o sistema (34) admite infinitas soluções. Nessa situação, para um certo atraso Δ , a solução procurada é a de norma mínima dada pelo cálculo da pseudo-inversa de \mathbf{H}_c :

$$\mathbf{w}_{z-\Delta} = (\mathbf{H}_c^H \mathbf{H}_c)^{-1} \mathbf{H}_c^H \mathbf{h}_\Delta = \mathbf{H}_c^+ \mathbf{h}_\Delta, \quad (35)$$

onde \mathbf{H}_c^+ representa a pseudo-inversa de \mathbf{H}_c . Na presença de ruído na entrada do equalizador LL, a variância do ruído na saída do sistema será dado por

$$\sigma_{ZF-\Delta}^2 = \sigma_n^2 \mathbf{w}_{z-\Delta}^H \mathbf{w}_{z-\Delta}. \quad (36)$$

Logo, o objetivo é encontrar um atraso Δ de tal forma que $\mathbf{w}_{z-\Delta}$ possua a menor norma entre todas as possíveis soluções que geram atrasos de equalização variando entre 0 e $N+L$ amostras. Portanto, considerando-se, uma equalização usando o critério ZF, o atraso ótimo é aquele que seleciona a coluna da matriz pseudo-inversa que possui a menor norma:

$$\Delta_{opt-ZF} = \arg \min_{\Delta} \left\{ \left\| \mathbf{H}_c^+ \mathbf{h}_\Delta \right\|_2 \right\}. \quad (37)$$

B. Atraso ótimo para equalização MSE

Será calculado agora o atraso Δ ótimo levando usando o critério MSE. Inicialmente, note que, para o caso do canal ser conhecido, o cálculo do filtro ótimo expresso por (7) pode ser calculado por [1]:

$$\mathbf{w}_{opt} = (\mathbf{H}_c^H \mathbf{H}_c + \zeta \mathbf{I})^{-1} \mathbf{H}_c^H \mathbf{h}_\Delta, \quad (38)$$

sendo $\zeta = \sigma_\eta^2 / \sigma_a^2$ a relação sinal-ruído.

Para um equalizador LL ótimo \mathbf{w}_{opt} , segundo o critério MSE, com um dado atraso Δ de equalização, o erro na saída do equalizador, considerando que o ruído é AWGN e a fonte de sinal é i.i.d., pode ser expresso por:

$$\begin{aligned} e(n) &= a(n-\Delta) - \mathbf{w}_{opt}^H \mathbf{u}(n) \\ e(n) &= a(n-\Delta) - \mathbf{w}_{opt}^H \mathbf{z}_e(n) + \mathbf{w}_{opt}^H \boldsymbol{\eta}_e(n) \\ e(n) &= \underbrace{a(n-\Delta) - \mathbf{w}_{opt}^H \mathbf{H}_c^T \mathbf{a}_e(n)}_{IES \text{ residual}} + \underbrace{\mathbf{w}_{opt}^H \boldsymbol{\eta}_e(n)}_{\text{ruído filtrado}}, \end{aligned} \quad (39)$$

sendo $\mathbf{z}_e(n) = \mathbf{H}_c \mathbf{a}_e(n)$, $\mathbf{a}_e(n) = [\mathbf{a}^T(n) \ \mathbf{a}^T(n-1)]^T$ e $\boldsymbol{\eta}_e(n) = [\boldsymbol{\eta}^T(n) \ \boldsymbol{\eta}^T(n-1)]^T$. Note que agora $\mathbf{H}_c \mathbf{w}_z - \mathbf{h}_\Delta \neq \mathbf{0}$. A IES residual em (39) pode ser escrita como

$$\begin{aligned} IES(n) &= \mathbf{h}_\Delta^T \mathbf{a}_e(n) - \mathbf{w}_{opt}^H \mathbf{H}_c^T \mathbf{a}_e(n) \\ &= (\mathbf{h}_\Delta^T - \mathbf{w}_{opt}^H \mathbf{H}_c^T) \mathbf{a}_e(n) \\ &= \mathbf{a}_e^T(n) (\mathbf{h}_\Delta - \mathbf{H}_c \mathbf{w}_{opt}^*). \end{aligned} \quad (40)$$

Considerando o ruído filtrado e a IES residual independentes, podemos escrever o MSE resultante como

$$\begin{aligned}
 J(\Delta) &= \sigma_a^2 (\mathbf{h}_\Delta - \mathbf{H}_c \mathbf{w}_{opt}^*)^H (\mathbf{h}_\Delta - \mathbf{H}_c \mathbf{w}_{opt}^*) + \sigma_\eta^2 \mathbf{w}_{opt}^T \mathbf{w}_{opt}^* \\
 &= \sigma_a^2 (\mathbf{h}_\Delta^H \mathbf{h}_\Delta - \mathbf{h}_\Delta^H \mathbf{H}_c \mathbf{w}_{opt}^* - \mathbf{w}_{opt}^T \mathbf{H}_c^H \mathbf{h}_\Delta + \mathbf{w}_{opt}^T \mathbf{H}_c^H \mathbf{H}_c \mathbf{w}_{opt}^*) \\
 &\quad + \sigma_\eta^2 \mathbf{w}_{opt}^T \mathbf{w}_{opt}^* \\
 &= \sigma_a^2 (\mathbf{h}_\Delta^H \mathbf{h}_\Delta - \mathbf{h}_\Delta^H \mathbf{H}_c \mathbf{w}_{opt}^* - \mathbf{w}_{opt}^T \mathbf{H}_c^H \mathbf{h}_\Delta) \\
 &\quad + \sigma_a^2 \mathbf{w}_{opt}^T (\mathbf{H}_c^H \mathbf{H}_c + \xi \mathbf{I}) \mathbf{w}_{opt}^*,
 \end{aligned} \tag{41}$$

Aplicando (38) em (41), obtém-se:

$$\begin{aligned}
 J(\Delta) &= \sigma_a^2 (\mathbf{h}_\Delta^H \mathbf{h}_\Delta - \mathbf{h}_\Delta^H \mathbf{H}_c \mathbf{w}_{opt}^* - \mathbf{w}_{opt}^T \mathbf{H}_c^H \mathbf{h}_\Delta + \mathbf{w}_{opt}^T \mathbf{H}_c^H \mathbf{H}_c \mathbf{w}_{opt}^*) \\
 &= \sigma_a^2 (\mathbf{h}_\Delta^H \mathbf{h}_\Delta - \mathbf{h}_\Delta^H \mathbf{H}_c (\mathbf{H}_c^H \mathbf{H}_c + \xi \mathbf{I})^{-1} \mathbf{H}_c^H \mathbf{h}_\Delta) \\
 &= \sigma_a^2 \mathbf{h}_\Delta^H (\mathbf{I} - \mathbf{H}_c (\mathbf{H}_c^H \mathbf{H}_c + \xi \mathbf{I})^{-1} \mathbf{H}_c^H) \mathbf{h}_\Delta.
 \end{aligned} \tag{42}$$

Como \mathbf{h}_Δ é um atraso puro, o produto $\mathbf{h}_\Delta^H \mathbf{M} \mathbf{h}_\Delta$, sendo \mathbf{M} uma matriz quadrada, resulta no elemento da diagonal principal correspondente à coluna e linha de índices Δ . Logo, para a minimização de $J(\Delta)$, segundo o critério MSE, o atraso ótimo Δ deve ser escolhido de tal forma que selecione o menor elemento da diagonal principal de $\mathbf{I} - \mathbf{H}_c (\mathbf{H}_c^H \mathbf{H}_c + \xi \mathbf{I})^{-1} \mathbf{H}_c^H$. Em termos matemáticos:

$$\Delta_{opt-MSE} = \arg \min_{\Delta} \left\{ \mathbf{h}_\Delta^H (\mathbf{I} - \mathbf{H}_c (\mathbf{H}_c^H \mathbf{H}_c + \xi \mathbf{I})^{-1} \mathbf{H}_c^H) \mathbf{h}_\Delta \right\}. \tag{43}$$

VI. RESULTADOS DE SIMULAÇÃO

Para verificar a precisão dos resultados teóricos apresentados nas últimas seções, aplica-se o modelo ao problema de equalização largamente linear de um canal complexo, como mostrado na Fig. 2.

Considere um sinal $a[k]$ pertencente a uma constelação 4-PAM, uma relação sinal-ruído de 30 dB e um canal cuja resposta ao impulso é dada por $\mathbf{h} = [0,519; 0,467+0,104j; 0,1413-0,4746j; -0,3174+0,1455j; 0,3612+0,0689j]$. Neste exemplo, estamos considerando que os filtros \mathbf{f} e \mathbf{g} têm seis coeficientes cada. A seqüência de treinamento é $d[k] = a[k-\Delta_1]$, sendo Δ_1 um atraso igual a 3 amostras.

A Fig. 3 apresenta as curvas teóricas e simuladas (média de 500 realizações) para o erro quadrático médio. Nota-se que o comportamento previsto e o experimental apresentam uma excelente concordância. Já a Fig. 4 mostra a evolução dos valores dos coeficientes para o filtro \mathbf{f} e os valores preditos pelo modelo, obtido com a média de apenas 5 realizações. Observa-se que existe uma boa concordância entre o modelo teórico e a simulação.

A Tabela II compara o MSE obtido na simulação com o resultado teórico esperado e o erro teórico de estado permanente, considerando-se os dois valores de μ simulados. Novamente, existe uma boa concordância entre os valores teóricos e os obtidos por simulação.

A Fig. 5 compara o desempenho do equalizador em função do atraso escolhido. Observe que atraso Δ_1 escolhido para a equalização do canal foi o atraso ótimo. Pela Fig. 5, também é possível perceber que o atraso ótimo, para uma relação sinal-ruído de 20 dB, difere do atraso ótimo, usando o critério MSE, quando a relação sinal-ruído é de 30 ou 40 dB. Como esperado, o atraso ótimo usando o critério ZF independe da relação sinal-ruído. Entretanto, o equalizador LL ZF apresenta um desempenho sempre inferior ao equalizador LL MSE.

Foi verificado por meio de várias simulações que o atraso ótimo para um equalizador LL MSE será o mesmo que para um equalizador LL ZF se ambos tiverem o mesmo número de coeficientes, a não ser que o nível de ruído seja muito alto (baixa relação sinal-ruído).

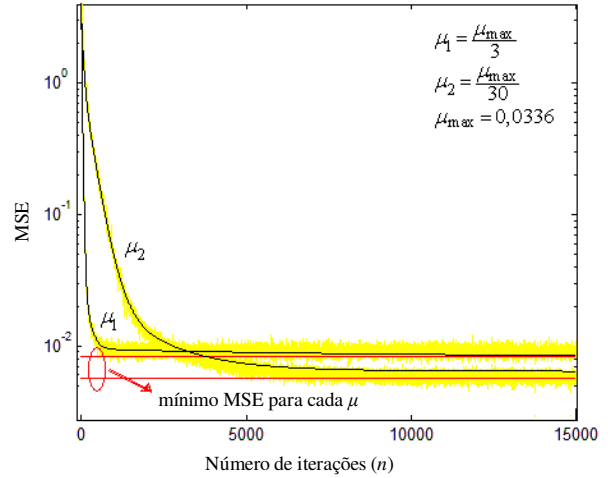


Fig. 3. Comportamento do MSE.

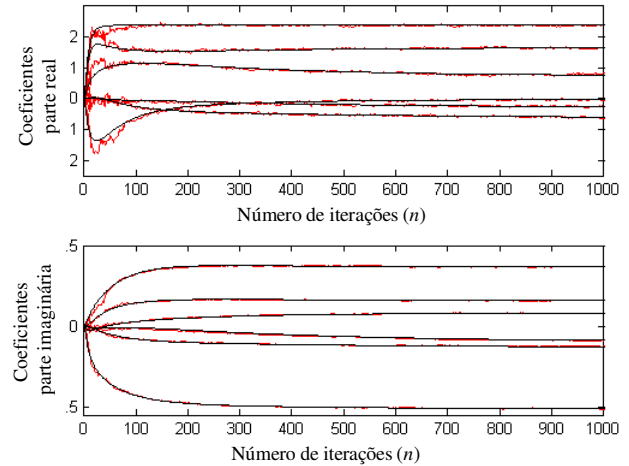


Fig. 4. Valores teóricos e simulados para o filtro \mathbf{f} .

Tabela II: comparações entre os MSE, $\Delta = 3$, $n = 15000$

	$J(n)$ Simulação	$J(n)$ teórico	J_∞ teórico
μ_1	0,00918	0,00844	0,00820
μ_2	0,00574	0,00640	0,00566

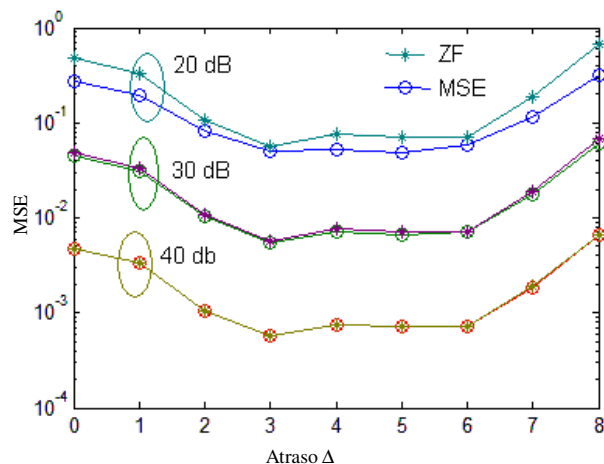


Fig. 5. Atrazo ótimo segundo os critérios MSE e ZF para várias relações de sinal e ruído.

VII. CONCLUSÃO

Neste artigo, apresentou-se uma análise estatística do algoritmo LMS largamente linear. Recursões determinísticas foram encontradas para o erro quadrático médio e para o comportamento médio do vetor de erro dos coeficientes. Além disso, foram encontradas as condições para a convergência do algoritmo e o desajuste esperado, trazendo um maior conhecimento sobre o comportamento do algoritmo. Também foram encontradas expressões fechadas para o cálculo do atrazo ótimo de equalização quando o canal e a relação sinal-ruído são conhecidos ou possuem boas estimativas. O atrazo ótimo, usando os critérios ZF ou MSE resultam no mesmo valor, a não ser que a relação sinal-ruído seja baixa.

AGRADECIMENTOS

Este trabalho foi parcialmente financiado pelo Centro Federal de Educação Tecnológica do Ceará (CEFET-Ce),

Universidade Federal de Santa Catarina (UFSC) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) processo número 484391/2006-2.

REFERÊNCIAS

- [1] W. H. Gerstacker, R. Schober and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. on Comm.*, vol. 51, no. 9, pp. 1512–1523, Sep. 2003.
- [2] B. Picinbono, and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Processing*, vol. 43, pp. 2030–2033, Aug. 1995.
- [3] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.
- [4] F.J.A de Aquino, C.A.F da Rocha, L.S. Resende, L., "Widely Linear Prediction for Blind Equalization", *IEEE International Conference on Communications, ICC'07*, pp. 2985-2990, Jun 2007.
- [5] R. Schober, W.H. Gerstacker, L.H.-J. Lampe, "A widely linear LMS algorithm for MAI suppression for DS-CDMA", *IEEE International Conference on Communications, ICC'03*, Vol. 4, pp. 2520-2525, May 2003.
- [6] Jae-Jin Jeon; J.G. Andrews, Koeng-Mo Sung, "Adaptive widely linear minimum output energy algorithm for DS-CDMA systems", *IEEE International Conference on Communications, ICC 2005*, Vol. 3, pp.2117-2121, May 2005.
- [7] S. Haykin, *Adaptive Filter Theory*, 4th edition, Prentice-Hall, New Jersey, 2002.
- [8] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd edition, McGraw-Hill, New York, 1991.
- [9] P.A. Voois; I. Lee, e J. M. Cioffi, "The effect of decision delay in finite-length decision feedback equalization", *IEEE Transactions on Information Theory*, Vol. 42, pp. 618-621, Mar. 1996.
- [10] T.J. Endres, S.D. Halford, C.R. Johnson, Jr., G.B. Giannakis, "Blind Adaptive Channel Equalization Using Fractionally-Spaced Receivers: A Comparison Study", *Proc. Conference on Information Sciences and Systems*, Princeton, N.J., março de 1996.
- [11] Proakis, J. G., "Digital Communications", McGraw-Hill, 3a. Edição, 1995.
- [12] F.J.A de Aquino, C.A.F da Rocha, e L.S. Resende, "Accelerating the Convergence of the Widely Linear LMS Algorithm for Channel Equalization," *VI International Telecommunications Symposium (ITS2006)*, Fortaleza-CE, Brazil, vol. 1, pp. 57-61, setembro de 2006