

Análise do algoritmo LMS por difusão em redes heterogêneas

Amanda de Paula e Cristiano M. Panazio

Resumo— Nesse artigo é apresentada uma análise do algoritmo LMS por difusão no contexto específico em que cada nó na rede possui uma SNR diferente. Através de uma análise de autovalores, é mostrada qual condição os passos de adaptação de cada nó da rede deve obedecer, a fim de garantir que o algoritmo apresente uma taxa de convergência fixada. Sob essa condição, é mostrado como o passo de adaptação de cada nó na rede deve ser definido de modo que o mínimo MSD seja obtido. Os resultados teóricos são também ilustrados com exemplos de simulação.

Palavras-Chave— algoritmo adaptativo, LMS por difusão, análise de autovalores.

Abstract— In this article, we analyze the diffusion least mean square algorithm considering that the nodes in the network have different SNR values. Firstly, through an eigenvalue analysis, we show which condition on the algorithm step-size values provides a given convergence rate. Then, it is shown how to set the algorithm step-size in each node in order to achieve the minimum mean square deviation. Finally, simulation results are provided in order to assess our theoretical analysis.

Keywords— adaptive algorithm, diffusion LMS, eigenvalue analysis

I. INTRODUÇÃO

A estimação de um parâmetro desconhecido em uma rede de sensores (WSN, do inglês, *wireless sensor network*) é um tópico amplamente estudado na literatura. Suas aplicações se estendem a uma variedade de contextos, tais como cenários de agricultura, em que é necessário estimar parâmetros tais como temperatura, umidade e velocidade dos ventos.

A maioria das implementações de WSN necessita de um centro de fusão para a estimação dos parâmetros. Os sistemas descritos em [1] e [2] são exemplos desse tipo de WSN. Em geral, esses sistemas implementados a partir de um centro de fusão são caracterizados pelo alto consumo de energia. Para combater esse problema, implementações que não necessitam de um centro de fusão vêm ganhando atenção recentemente [3], [4].

Nesse contexto, uma proposta interessante é a classe de filtros adaptativos distribuídos [5], [6]. A maioria dos algoritmos adaptativos clássicos podem ser também aplicados no contexto distribuído. Por exemplo, uma versão distribuída do algoritmo mínimos quadrados médios (LMS, do inglês, *least mean square*) pode ser encontrada em [5], e do mínimos quadrados recursivos (RLS, do inglês, *recursive least squares*) pode ser encontrado em [7].

Amanda de Paula e Cristiano Panazio, Departamento de Engenharia de Telecomunicações e Controle (PTC), Escola Politécnica, Universidade de São Paulo, SP, Brazil, e-mails: {amanda, cpanazio}@lcs.poli.usp.br

Esse trabalho foi parcialmente financiado pela FAPESP

Um outro aspecto que merece atenção é a análise de algoritmos distribuídos no caso em que a observação em cada nó da rede adaptativa não apresenta a mesma confiabilidade em relação à estimativa do parâmetro desconhecido. Nessa situação, as estimativas produzidas pelos nós mais ruidosos tendem a comprometer a estimativa global da rede. Esse cenário heterogêneo pode ser encontrado em uma variedade de contextos, entre os quais podemos destacar o sensoriamento espectral para rádio cognitivo em que o canal entre cada um dos usuários secundários e o usuário primário pode apresentar diferentes valores de relação sinal-ruído (SNR, do inglês, *signal-to-noise ratio*) [8].

A contribuição desse artigo é encontrar a configuração ótima do passo de adaptação de cada nó para uma dada taxa de convergência com o objetivo de minimizar o desvio médio quadrático (MSD, do inglês, *mean-square deviation*) dos parâmetros estimados. Com esse objetivo, primeiramente é mostrado que a taxa de convergência é a mesma desde que a média aritmética entre todos os passos de adaptação seja a mesma. Em seguida, é mostrado através de minimização por multiplicadores de Lagrange qual a configuração ótima de passos de adaptação em termos de MSD.

O restante desse artigo se organiza da seguinte forma. Na Seção II é descrito o algoritmo de difusão utilizado ao longo desse artigo. Na Seção III, é apresentada uma abordagem para a escolha dos parâmetros do processamento local com o objetivo de manter uma taxa de convergência constante, e na Seção IV mostra-se a configuração de passos de adaptação que minimiza a MSD. Já na Seção V são mostrados os resultados de simulação. Finalmente, a Seção VI apresenta as principais conclusões do artigo.

II. ESTIMAÇÃO DISTRIBUÍDA ADAPTATIVA

Nesse artigo, considera-se que a cooperação entre os nós se dá através de uma regra de difusão [6]. Nesse modo de cooperação, cada nó troca sua estimativa com um subconjunto dos nós da rede.

Considera-se que os caminhos da rede são canais ideais e que se trata de uma rede bidirecional. Ainda em relação à topologia de rede, define-se a vizinhança do nó k , \mathcal{N}_k , como o subconjunto de nós da rede que estão conectados a ele, incluindo ele próprio.

No nó local k , assume-se um modelo linear. Assim, cada nó $k \in 1, \dots, N$ tem acesso ao conjunto de medidas $\{\mathbf{u}_k(i), d_k(i)\}$, que se relacionam da seguinte forma:

$$d_k(i) = \mathbf{u}_k(i)^T \mathbf{w}_{opt} + \eta_k(i) \quad (1)$$

em que \mathbf{w}_{opt} é o vetor desconhecido determinístico ótimo a

ser estimado e $\eta_k(i)$ é um ruído de medida gaussiano de média nula e variância $\sigma_{\eta,k}^2$.

No algoritmo de estimação adaptativa distribuída, além do par $\{\mathbf{u}_k(i), d_k(i)\}$, cada nó na rede também tem acesso às estimativas geradas em sua vizinhança, definidas como $\psi_l, l \in \mathcal{N}_k$. Então, uma estimativa agregada ϕ_k é obtida através da fusão das estimativas locais. Considerando combinações lineares, ϕ_k é dado por:

$$\phi_k(i) = \sum_{l \in \mathcal{N}_k} c_{k,l} \psi_l(i). \quad (2)$$

Os coeficientes $c_{k,l}$ podem ser escolhidos de várias formas. Em [5], é dito que a matriz de coeficientes \mathbf{C} deve ser tal que a soma dos elementos de suas linhas seja unitária.

Uma possível regra de escolha para seus coeficientes é dada através da regra de Metropolis [5]:

$$c_{k,l} = \begin{cases} \frac{1}{\max\{n_k, n_l\}}, & k \in \mathcal{N}_l \\ 0, & k \notin \mathcal{N}_l \\ 1 - \sum_{l \in \mathcal{N}_k, l \neq k} c_{k,l} & k = l \end{cases} \quad (3)$$

em que n_k e n_l representam o número de nós conectados ao nó k e ao nó l , respectivamente.

Localmente, cada nó na rede pode aplicar vários algoritmos adaptativos para atualizar suas estimativas. Com o objetivo de manter a complexidade computacional baixa, é comum utilizar versões do algoritmo LMS no processamento local. Neste caso, as estimativas locais $\phi_k(i)$ são atualizadas como descrito na Tabela 1, em que μ_k é o passo de adaptação do LMS relativo ao nó k .

Tabela 1

Algorithm 1 LMS

$$\begin{aligned} \phi_k(i) &= \sum_{l \in \mathcal{N}_k} c_{k,l} \psi_l(i) \\ e(i) &= d(i) - \phi_k(i)^T \mathbf{u}_k(i) \\ \psi_k(i+1) &= \phi_k(i) + \mu_k e(i) \mathbf{u}_k(i) \end{aligned}$$

III. PASSO DE ADAPTAÇÃO DO LMS EM REDES ADAPTATIVAS HETEROGRÉNEAS

Uma questão que surge na análise de redes heterogêneas é como escolher adequadamente os passos de adaptação de cada nó na rede. Para tanto, analisaremos uma rede completamente conectada. Nesse caso, a matriz de combinação \mathbf{C} deve ser escolhida como:

$$\mathbf{C} = \frac{1}{N} \mathbf{1} \mathbf{1}^T \quad (4)$$

em que N é o número de nós na rede e $\mathbf{1}$ é um vetor coluna N -dimensional com todos os elementos iguais a um.

Uma abordagem para analisar a taxa de convergência consiste em introduzir o vetor de erro de coeficientes da seguinte forma:

$$\tilde{\psi}(i) = \left[\tilde{\psi}_1^T(i) \ \tilde{\psi}_2^T(i) \ \cdots \ \tilde{\psi}_N^T(i) \right]^T \quad (5)$$

em que

$$\tilde{\psi}_k = \psi_k - \mathbf{w}_{opt}. \quad (6)$$

Pode ser mostrado que a seguinte recursão é válida [5]:

$$\mathbb{E} \{ \tilde{\psi}(i) \} = (\mathbf{I}_{MN} - \mathbf{D} \mathbf{R}_{\mathbf{u}}) \mathbf{G} \mathbb{E} \{ \tilde{\psi}(i-1) \} \quad (7)$$

em que $\mathbf{G} = \mathbf{C} \otimes \mathbf{I}_M$, $\mathbf{D} = \mu \otimes \mathbf{I}_M$, $\mathbf{R}_{\mathbf{u}} = \mathbf{I}_N \otimes \mathbf{R}$, \mathbf{I}_N é a matriz identidade de dimensão N e \mathbb{E} denota o operador de esperança estatística. O símbolo \otimes denota o produto de Kronecker, a matriz μ é uma matriz diagonal cujos elementos são os passos de adaptação em cada nó e a matriz $\mathbf{R} = E \{ \mathbf{u}_k \mathbf{u}_k^T \}$ é a matriz de autocorrelação da entrada que é considerada a mesma para todo nó k .

A taxa de convergência é determinada pelo raio espectral [5], *i.e.*, a máxima magnitude dos autovalores da matriz $\mathbf{A} = (\mathbf{I}_{MN} - \mathbf{D} \mathbf{R}_{\mathbf{u}}) \mathbf{G}$.

Aplicando algumas propriedades do produto de Kronecker, pode ser mostrado que a matriz \mathbf{A} pode ser reescrita como:

$$\mathbf{A} = \mathbf{C} \otimes \mathbf{I}_M - \mu \mathbf{C} \otimes \mathbf{R}. \quad (8)$$

A. Entrada branca

Modelando o vetor regressor como um processo estocástico branco de potência unitária, a matriz \mathbf{A} pode ser reescrita como:

$$\mathbf{A} = (\mathbf{I}_N - \mu) \mathbf{C} \otimes \mathbf{I}_M. \quad (9)$$

Definindo $\mathbf{A}_1 = (\mathbf{I}_N - \mu) \mathbf{C}$, pode-se notar que os autovalores de \mathbf{A} são iguais aos autovalores de \mathbf{A}_1 . Portanto, os raios espectrais de \mathbf{A} e \mathbf{A}_1 são iguais [9].

A matriz \mathbf{A}_1 aplicada ao vetor \mathbf{v} resulta em:

$$\mathbf{A}_1 \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} = \frac{1}{N} \sum_{i=1}^N v_i \begin{bmatrix} 1 - \mu_1 & 0 & \cdots & 0 \\ 0 & 1 - \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 - \mu_N \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \quad (10)$$

Portanto:

$$\mathbf{A}_1 \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} = \frac{1}{N} \sum_{i=1}^N v_i \underbrace{\begin{bmatrix} 1 - \mu_1 \\ 1 - \mu_2 \\ \vdots \\ 1 - \mu_N \end{bmatrix}}_{\mathbf{v}_0}. \quad (11)$$

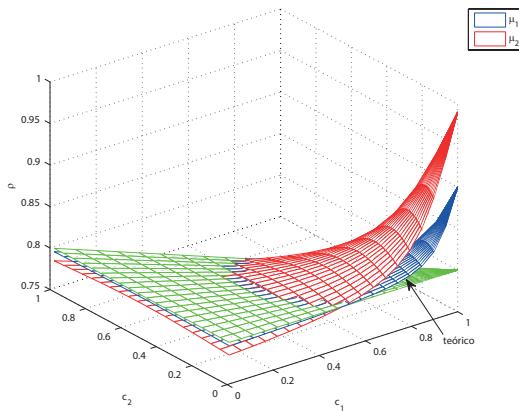
Logo, o vetor \mathbf{v}_0 é um autovetor de \mathbf{A}_1 associado ao autovalor $\lambda = 1 - \frac{1}{N} \sum_{i=1}^N \mu_i$. Dessa forma, o raio espectral de \mathbf{A} é dado por:

$$\rho = 1 - \frac{1}{N} \sum_{i=1}^N \mu_i. \quad (12)$$

Finalmente, podemos concluir que se a média dos passos for constante, a taxa de convergência do algoritmo também será constante.

Apesar da demonstração acima ser válida para a rede completamente conectada, em que a matriz de conectividade é dada por (4), o resultado em (12) é ainda uma boa aproximação para casos mais gerais, em que \mathbf{C} é uma matriz estocástica qualquer. A fim de ilustrar isso, consideraremos um exemplo com uma rede de 3 nós em que a matriz de combinação é dada por:

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & 1 - c_1 - c_2 \\ 1 - c_1 - c_2 & c_1 & c_2 \\ c_2 & 1 - c_1 - c_2 & c_1 \end{bmatrix}. \quad (13)$$


 Fig. 1. Raio espectral para as configurações de passo μ_1 e μ_2

em que $0 \leq c_1 \leq 1$ e $0 \leq c_2 \leq 1 - c_1$. Duas configurações de passo de adaptação com a mesma média são analisadas: $\mu_1 = [0, 30, 10, 2]$ e $\mu_2 = [0, 390, 0, 010, 2]$. A Fig. 1 mostra o raio espectral relativo a matrizes de combinação da forma (13). Pode-se notar que (12) é uma boa aproximação para o raio espectral para vários valores de c_1 e c_2 . A expressão teórica perde sua validade se $c_1 \approx 1$, caso em que a matriz de combinação tende à identidade, o que leva o algoritmo a se aproximar do caso não-cooperativo. Nessa situação, o raio espectral será dado por $\rho = 1 - \mu_{\min}$.

B. Entrada colorida

Nesse caso, consideramos uma entrada colorida com matriz de correlação dada por: $\mathbf{R} = \mathbf{U}\Lambda\mathbf{U}^H$, em que \mathbf{U} é uma matriz unitária e Λ é uma matriz diagonal com elementos $\Lambda(k, k) = \lambda_k$.

O raio espectral da matriz \mathbf{A} , nesse contexto, é dado por:

$$\rho = 1 - \lambda_{\min} \frac{1}{N} \sum_{i=1}^N \mu_i \quad (14)$$

A fim de provar (14), note que a matriz \mathbf{R}_U pode ser decomposta em $\mathbf{R}_U = (\mathbf{I}_N \otimes \mathbf{U}) (\mathbf{I}_N \otimes \Lambda) (\mathbf{I}_N \otimes \mathbf{U}^H)$. Portanto, podemos escrever \mathbf{A} como:

$$\mathbf{A} = (\mathbf{I}_{MN} - \mathbf{D} (\mathbf{I}_N \otimes \mathbf{U}) (\mathbf{I}_N \otimes \Lambda) (\mathbf{I}_N \otimes \mathbf{U}^H)) \mathbf{G}. \quad (15)$$

Nesse ponto, utilizaremos a seguinte propriedade relativa ao produto de Kronecker:

$$(\mathbf{M}_1 \otimes \mathbf{M}_2) (\mathbf{M}_3 \otimes \mathbf{M}_4) = \mathbf{M}_1 \mathbf{M}_3 \otimes \mathbf{M}_2 \mathbf{M}_4 \quad (16)$$

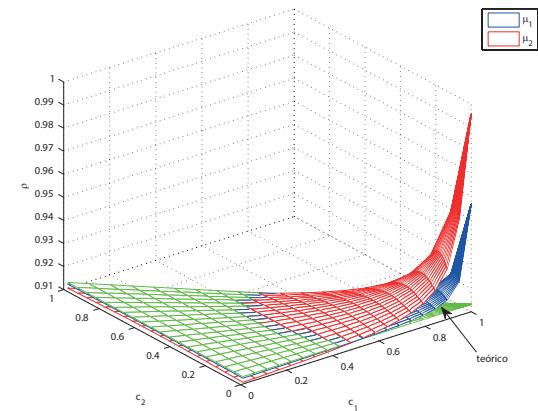
em que \mathbf{M}_1 , \mathbf{M}_2 , \mathbf{M}_3 e \mathbf{M}_4 são matrizes com dimensões que permitem os produtos $\mathbf{M}_1 \mathbf{M}_3$ e $\mathbf{M}_2 \mathbf{M}_4$. Aplicando a propriedade (16), e lembrando que $\mathbf{D} = \mu \otimes \mathbf{I}_M$ e $\mathbf{G} = \mathbf{C} \otimes \mathbf{I}_M$, as seguintes igualdades são válidas:

$$\mathbf{D} (\mathbf{I}_N \otimes \mathbf{U}) = (\mathbf{I}_N \otimes \mathbf{U}) \mathbf{D} \quad (17)$$

$$(\mathbf{I}_N \otimes \mathbf{U}^H) \mathbf{G} = \mathbf{G} (\mathbf{I}_N \otimes \mathbf{U}^H) \quad (18)$$

Portanto, a matriz \mathbf{A} pode ser escrita como:

$$\mathbf{A} = (\mathbf{I}_N \otimes \mathbf{U}) \mathbf{E} \mathbf{G} (\mathbf{I}_N \otimes \mathbf{U})^H \quad (19)$$


 Fig. 2. Raio espectral para as configurações de passo μ_1 e μ_2 para o caso de entrada colorida.

em que \mathbf{E} é uma matriz diagonal dada por:

$$\mathbf{E} = \mathbf{I}_{NM} - \mathbf{D} (\mathbf{I}_N \otimes \Lambda). \quad (20)$$

Pode-se notar que os vetores da forma $\mathbf{v} = [\mathbf{0}_{k-1} \ 1 - \mu_1 \lambda_k \ \mathbf{0}_{M-k} \ \dots \ \mathbf{0}_{k-1} \ 1 - \mu_N \lambda_k \ \mathbf{0}_{M-k}]^T$ são autovetores associados à matriz $\mathbf{E} \mathbf{G}$ e que seus respectivos autovalores são dados por $1 - \lambda_k \frac{1}{N} \sum_{i=1}^N \mu_i$. Os demais $(N-1)M$ autovalores de $\mathbf{E} \mathbf{G}$ são nulos, uma vez que $\mathbf{C} = \frac{1}{N} \mathbf{1} \mathbf{1}^T$ apresenta $N-1$ autovalores nulos. Portanto, o raio espectral da matriz \mathbf{A} é de fato dado por (14).

A eq. (14) mostra que, de modo análogo ao caso de entrada branca, a taxa de convergência permanece a mesma se a média dos passos de adaptação de cada filtro da rede permanecer constante. Para ilustrar esse resultado, considere um exemplo em que a primeira linha da matriz de autocorrelação é dada por $\mathbf{r} = [1 \ 0, 2 \ 0, 4 \ 0, 1 \ 0]$.

Assim como no caso de entrada branca, iremos analisar uma rede de três nós e matriz de combinação dada por (13). A Fig. 2 mostra o raio espectral em função dos parâmetros da matriz de combinação. Pode-se notar que a expressão teórica em (14) é uma boa aproximação para várias escolhas da matriz de combinação.

IV. CONFIGURAÇÃO ÓTIMA DE PASSO DE ADAPTAÇÃO

O desempenho do algoritmo adaptativo é frequentemente avaliado em termos do MSD, definido como:

$$\text{MSD} = \frac{1}{N} \sum_{k=1}^N \mathbb{E} \left\{ \left\| \tilde{\psi}_k \right\|^2 \right\}. \quad (21)$$

Na seção anterior, foi mostrado que se a média dos passos for constante, a taxa de convergência é também constante. Nessa seção, será apresentado um método para designar o passo de adaptação de cada nó com o objetivo de minimizar o MSD.

Para analisar essa questão, será considerada uma rede com três nós, assim como feito na seção anterior. Como parâmetros de simulação, a potência vetor de entrada é mantida constante para todos os nós e a potência de ruído de medida em cada nó é dada por: $\sigma_\eta^2 = [0, 20, 0, 04, 0, 06]$. Fixando o passo médio em

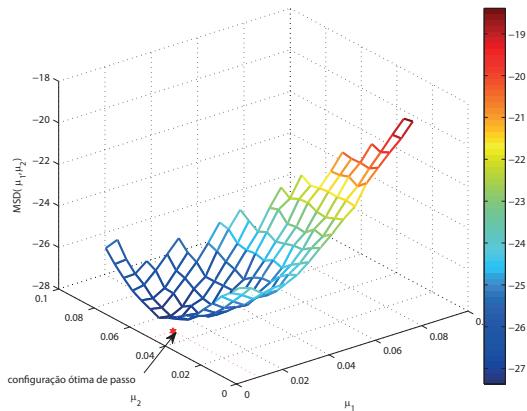


Fig. 3. MSD para diferentes configurações de passo de adaptação

$\bar{\mu} = 0,03$, variou-se os passos de adaptação para os nós 1 e 2. Os valores de MSD obtidos para cada configuração de passo está indicado na Fig. 3, pode-se notar que o valor mínimo de MSD é alcançado para a configuração de passo $\mu_1 \approx 0,01$, $\mu_2 \approx 0,05$ (e, consequentemente, $\mu_3 \approx 0,03$.) A partir desse resultado, conjectura-se que a configuração ótima de passos de adaptação deve obedecer a seguinte condição:

$$\sigma_{\eta,1}^2 \mu_1 = \sigma_{\eta,2}^2 \mu_2 = \cdots = \sigma_{\eta,N}^2 \mu_N. \quad (22)$$

A condição (22) e a condição $\bar{\mu} = \frac{1}{N} \sum_{k=1}^N \mu_k$, levam à elaboração da seguinte proposição:

Proposição 1 (Distribuição ótima de passo de adaptação): Para um dado passo de adaptação médio $\bar{\mu}$, o passo de adaptação em cada nó deve ser dado por:

$$\mu_i = \frac{\bar{\mu}}{\sigma_{\eta,i}^2} \mathcal{H}(\boldsymbol{\sigma}_\eta^2) \quad (23)$$

em que $\mathcal{H}(\sigma_\eta^2)$ representa a média harmônica do vetor de potência de ruído:

$$\mathcal{H}(\sigma_\eta^2) = \frac{1}{\frac{1}{N} \sum_{k=1}^N \frac{1}{\sigma_k^2}}. \quad (24)$$

Demonstração: O MSD pode ser escrito como:

$$\text{MSD} = \frac{1}{N^2} \left[\sum_{i=1}^M \frac{\lambda_i}{(1 - (1 - \lambda_i \bar{\mu})^2)} \right] \left[\sum_{k=1}^N \mu_k^2 \sigma_{\eta, k}^2 \right]. \quad (25)$$

Eliminando o termo constante em (25) relativo ao passo de adaptação, a configuração ótima de passo de adaptação será dada pela solução do seguinte problema de otimização:

$$\begin{aligned} & \underset{\mu = [\mu_1 \ \mu_1 \cdots \ \mu_N]}{\text{minimize}} && \sum_{k=1}^N \mu_k^2 \sigma_{\eta, k}^2 \\ & \text{subject to} && \frac{1}{N} \sum_{k=1}^N \mu_k = \bar{\mu} \end{aligned}$$

O Lagrangiano associado a esse problema é dada por:

$$L(\boldsymbol{\mu}, l) = \sum_{k=1}^N \mu_k^2 \sigma_{\eta, k}^2 + l \left(\frac{1}{N} \sum_{i=1}^N \mu_i - \bar{\mu} \right). \quad (26)$$

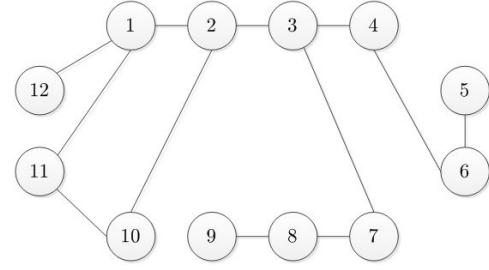


Fig. 4. Rede de difusão

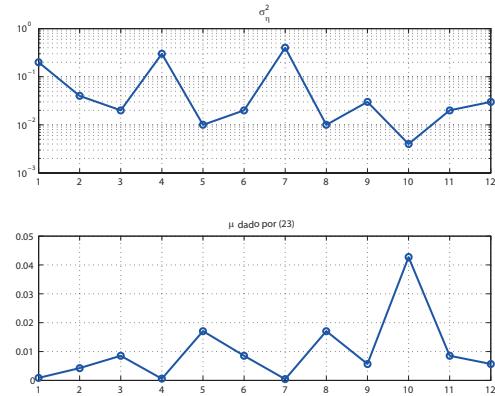


Fig. 5. Distribuição de potência de ruído em cada nó, e o respectivo passo de adaptação ótimo

Igualando o gradiente do Lagrangiano a zero, temos que a seguinte relação precisa ser obedecida:

$$\frac{\partial}{\partial \mu_k} L(\boldsymbol{\mu}, l) = 0 \rightarrow \mu_k = -\frac{l}{2N\sigma_k^2}. \quad (27)$$

Aplicando (27) na equação de restrição, resulta em (23). ■

É interessante observar que esse é um resultado que não depende da matriz de correlação do sinal de entrada.

V. RESULTADOS DE SIMULAÇÃO

Os resultados apresentados nas seções anteriores assumem que a rede de cooperação é completamente conectada. Entretanto, como também foi indicado nas seções anteriores, tais resultados podem ser estendidos para outras matrizes de combinação. Nesse seção, nós corroboraremos essa afirmação com alguns resultados de simulação.

Será considerada uma rede com 12 nós como indicado na Fig. 4. No primeiro cenário de simulação, é considerado que a entrada é branca de potência unitária. O vetor a ser estimado apresenta $M = 5$ coeficientes iguais e potência unitária, *i.e.*, $\mathbf{w}_{opt} = [1, 1, \dots, 1]/\sqrt{M}$.

A potência de ruído em cada um dos nós está indicada na Fig. 5. O passo de adaptação médio foi fixado em $\bar{\mu} = 0,01$ e cada passo individual foi calculado de acordo com (23), os valores de passo de adaptação em cada nó estão também mostrados na Fig. 5.

A Fig. 6 mostra o MSD para o primeiro cenário de simulação. O MSD foi estimado a partir da média de 1000

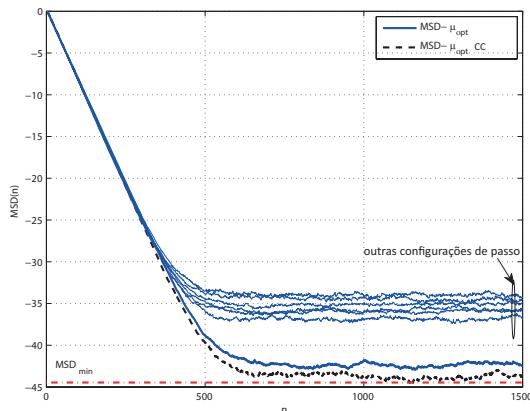


Fig. 6. MSD para diferentes configurações de passo fixando $\bar{\mu} = 0,01$ e entrada branca. O desempenho da rede completamente conectada (CC) está indicado na curva tracejada.

realizações independentes de medidas instantâneas de MSD:

$$\text{MSD}(i) = \frac{1}{N} \sum_{k=1}^N \|\psi_k(i) - \mathbf{w}_{\text{opt}}\|^2. \quad (28)$$

Cada uma das várias curvas de MSD apresentadas na Fig. 6 corresponde a uma configuração de passo de adaptação diferente, sorteada a partir de uma distribuição uniforme e normalizada para garantir um passo de adaptação médio de $\bar{\mu} = 0,01$. O primeiro ponto importante a ser observado é que a taxa de convergência, de fato, se mantém constante, mesmo que a rede não seja completamente conectada. Pode-se também notar que a configuração de passo de adaptação que leva ao menor valor de MSD é aquela mostrada na Fig. 5. Além disso, é importante enfatizar que o valor mínimo de MSD dado por (25) é uma boa aproximação para o valor mínimo de MSD atingido. A diferença entre a expressão teórica dada por (25) e o valor obtido por simulação pode ser explicada pelo fato da rede não ser completamente conectada. Com o objetivo de ilustrar esse fato, foi feita uma simulação com os passos de adaptação ótimo e uma rede completamente conectada. O resultado está indicado pela curva pontilhada na Fig. 6, a partir da qual pode-se perceber que o valor teórico expresso por (25) é atingido.

Além do caso branco, também foram feitas simulações para o caso colorido. Nesse caso, considerou-se que o vetor regressor de entrada u_k foi modelado como um processo autoregressivo com função de autocorrelação dada por $r(n) = 0,2^n$.

Os resultados estão indicados na Fig. 7, a partir dos quais pode-se notar que a taxa de convergência é também constante para o caso de entrada colorida, mesmo que a rede não seja completamente conectada. Além disso, nota-se que o valor mínimo de MSD é alcançado quando o valor do passo de adaptação é também escolhido de acordo com (23). Assim como no caso de entrada branca, (25) é uma boa aproximação para o valor mínimo de MSD. Também mostra-se a simulação para o caso de rede completamente conectada em que o valor de MSD se aproxima ainda mais de (25).

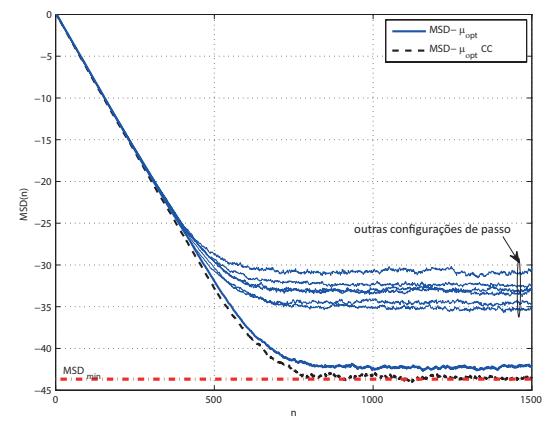


Fig. 7. MSD para diferentes configurações de passo fixando $\bar{\mu} = 0,01$ e entrada colorida. O desempenho da rede completamente conectada (CC) está indicado na curva tracejada.

VI. CONCLUSÕES

Nesse artigo foi mostrado que dado que o valor médio dos passos de adaptação é fixado, a taxa de convergência do algoritmo se mantém constante para o caso de redes completamente conectadas. Através de resultados de simulação, indicou-se que esse resultado pode ser estendido a topologias de rede mais gerais. Mostrou-se também que o valor mínimo de MSD é alcançado se o passo de adaptação de cada nó na rede é inversamente proporcional à potência de ruído de cada um dos nós. Um problema dessa abordagem é que é necessário um conhecimento *a priori* das potências de ruído de todos os nós na rede. Como trabalho futuro, pretende-se propor estratégias para difundir também as estimativas de potência de ruído na rede para superar esse problema de conhecimento *a priori* das potências de ruído.

REFERÊNCIAS

- [1] M. Guerriero, L. Svensson, and P. Willett, "Bayesian data fusion for distributed target detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3417–3421, June 2010.
- [2] R. Viswanathan and P.K. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
- [3] Qing Ling and Zhi Tian, "Decentralized sparse signal recovery for compressive sleeping wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3816–3827, July 2010.
- [4] S. Barbarossa and G. Scutari, "Decentralized maximum-likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3456–3470, July 2007.
- [5] C.G. Lopes and A.H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, July 2008.
- [6] C.G. Lopes and A.H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, Aug. 2007.
- [7] F.S. Cattivelli, C.G. Lopes, and A.H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1865–1877, May 2008.
- [8] F.S. Cattivelli and A.H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1917–1932, May 2011.
- [9] C.D. Meyer, *Matrix analysis and applied linear algebra*, SIAM, 2000.