

# Análise de Transiente do Algoritmo $\ell_0$ -sign-LMS

Leonardo O. Santos<sup>1</sup>, Diego B. Haddad<sup>2</sup>, Mariane R. Petraglia<sup>1</sup>

**Resumo**—Sistemas esparsos são comuns nas mais diversas áreas de aplicação da filtragem adaptativa, dentre elas controle, engenharia biomédica e cancelamento de eco. Algoritmos capazes de usar essa propriedade para acelerar a convergência e/ou melhorar o erro em regime permanente vêm ganhando cada vez mais destaque na comunidade acadêmica. Motivado por tais experiências, este trabalho apresenta uma análise teórica dos desempenhos em transiente e em regime permanente do algoritmo  $\ell_0$ -sign-LMS, o qual apresenta as propriedades de convergência desejadas para sistemas esparsos, além de robustez a ruído impulsivo e reduzida complexidade computacional. A análise é efetuada através do cálculo recursivo da matriz de autocorrelação dos desvios dos coeficientes, sendo assumidas algumas suposições comuns na literatura, mas não a hipótese muito frequente de que o sinal de entrada é branco. Ao final, são apresentados os resultados de diversas simulações para avaliar a acurácia do modelo estocástico proposto.

**Palavras-Chave**—filtragem adaptativa, cálculo recursivo da matriz de autocorrelação dos desvios,  $\ell_0$ -sign-LMS, identificação de sistemas esparsos, análise de transiente.

**Abstract**—Sparse systems are common in the most diverse areas of application of adaptive filtering, including control, biomedical engineering and echo cancellation. Algorithms capable of using this property to accelerate convergence and/or improve steady state error are gaining increasing prominence in the academic community. Motivated by these experiences, this work presents a theoretical analysis of the transient and steady-state performances of the  $\ell_0$ -sign-LMS algorithm, which presents the desired convergence properties for sparse systems, as well as robustness to impulsive noise and reduced computational complexity. The analysis is performed through the recursive calculation of the autocorrelation matrix of the coefficient deviations, adopting some assumptions commonly used in the literature, but not the very frequent hypothesis that the input signal is white. At the end, several simulation results are presented to evaluate the performance of the proposed stochastic model.

**Keywords**—Adaptive filtering, recursive autocorrelation matrix calculation,  $\ell_0$ -sign-LMS, sparse system identification, transient analysis.

## I. INTRODUÇÃO

As técnicas de filtragem adaptativa apresentam diversas aplicações, dentre as quais importa destacar: engenharia biomédica, equalização de canais, cancelamento de eco e identificação de sistemas [1]. Filtros adaptativos apresentam duas grandes vantagens com relação a filtros com

coeficientes fixos, a saber: não é necessário um conhecimento prévio das características do sinal de entrada (ou do sistema envolvido) e também podem ser empregados para modelagem ou equalização de sistemas variantes no tempo.

A Figura 1 mostra a estrutura de um algoritmo adaptativo (aplicado à identificação de sistemas, foco deste artigo), cujos coeficientes do filtro de comprimento  $N$  podem ser coletados no vetor  $\mathbf{w}(k)$  definido por:

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

sendo a saída do filtro adaptativo no instante  $k$  definida por  $y(k) \triangleq \mathbf{w}^T(k)\mathbf{x}(k)$ , com o vetor de entrada  $\mathbf{x}(k)$  definido por  $\mathbf{x}(k) = [x(k) \ x(k-1) \ \dots \ x(k-N+1)]^T$ . Após um número suficiente de iterações, é esperado que  $y(k)$  seja similar ao valor de referência  $d(k) = (\mathbf{w}^o)^T \mathbf{x}(k) + \nu(k)$ , sendo  $\mathbf{w}^o$  o vetor que contém os coeficientes a serem identificados e  $\nu(k)$  um ruído somado ao sinal de referência, geralmente atribuído a erros de medição e/ou de modelagem.

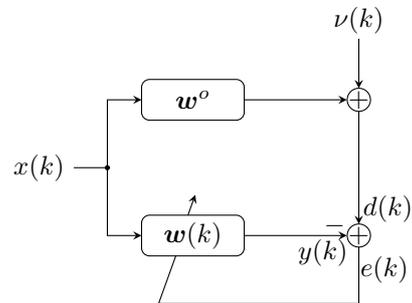


Fig. 1. Diagrama de blocos de um filtro adaptativo.

A atualização dos coeficientes de um filtro adaptativo costuma ser regida pelo cálculo do gradiente de uma função custo estocástica, o qual é capaz de fornecer a direção de atualização. Portanto, podemos escrever a equação de atualização de diversos algoritmos de filtragem adaptativa na seguinte forma:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \beta \frac{\partial \mathcal{F}[e(k)]}{\partial \mathbf{w}(k)}, \quad (2)$$

onde  $\frac{\partial \mathcal{F}[e(k)]}{\partial \mathbf{w}(k)}$  é o vetor gradiente da função  $\mathcal{F}[e(k)]$  e  $\beta$  é o passo de aprendizagem empregado para evitar saltos muito grandes durante o processo de aprendizagem do filtro, fazendo com que o algoritmo convirja sem experimentar oscilações capazes de degradar seu desempenho.

Vale notar que para se chegar na equação do algoritmo LMS, a partir da Equação (2), basta tomar como função custo o erro instantâneo quadrático:  $\mathcal{F}[e(k)] = e^2(k)$ . É

<sup>1</sup>CEFET-RJ. <sup>2</sup>PEE/COPPE/UFRJ. E-mails: leo.08@poli.ufrj.br, diego@pads.ufrj.br, mariane@pads.ufrj.br

sabido, entretanto, que na presença de ruído impulsivo, os coeficientes adaptativos costumam sofrer grandes perturbações, desviando-se da solução ótima. Uma solução para reduzir este problema é tomar como função custo o módulo do erro ( $\mathcal{F}[e(k)] = |e(k)|$ ), e não seu valor quadrático, obtendo-se a seguinte equação de atualização dos coeficientes:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \text{sign}[e(k)]\mathbf{x}(k), \quad (3)$$

a qual implementa o algoritmo conhecido como sign-LMS [2], [3].

## II. IDENTIFICAÇÃO DE SISTEMAS ESPARSOS

Desde o início dos anos 2000, muitos autores notaram que sistemas reais, recorrentemente, apresentam respostas ao impulso esparsas (ou seja, possuem muitos coeficientes de magnitude nula ou quase nula), sendo tais respostas ao impulso comuns em aplicações como canais de transmissão de televisão digital e cancelamento de eco [4].

Por se tratar de uma propriedade recorrente, a capacidade de incorporá-la no processo de aprendizagem do filtro adaptativo vem ganhando notoriedade por acelerar a convergência dos algoritmos e/ou diminuir o erro quadrático médio em regime permanente. Como nenhum dos algoritmos tradicionais da família LMS fazia uso dessa propriedade na identificação de sistemas, algoritmos conscientes da esparsidade foram desenvolvidos. Tais algoritmos são capazes de aproveitar a esparsidade dos sistemas de diferentes formas, seja identificando e não atualizando regiões inativas (isto é, com coeficientes muito próximos de zero) [5], seja propondo uma atualização específica para cada coeficiente com base em sua magnitude (algoritmos proporcionais) [6]–[8], ou ainda aplicando uma penalização aos coeficientes dentro de uma determinada faixa, de forma a “empurrá-los” para zero [4], [9], [10], dentre outras formas propostas na literatura [11].

Uma das primeiras derivações realizadas para se utilizar a propriedade de coeficientes esparsos foi a atualização proporcional dos coeficientes (PNLMS), que consiste na atualização proporcional dos coeficientes consoante a sua magnitude, de acordo com a seguinte equação de atualização:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\beta}{\mathbf{x}^T(k)\mathbf{\Lambda}(k)\mathbf{x}(k) + \delta} \mathbf{\Lambda}(k)\mathbf{x}(k)e(k), \quad (4)$$

onde a matriz diagonal  $\mathbf{\Lambda}(k)$  é responsável pela distribuição da energia de atualização ao longo dos coeficientes do filtro adaptativo.

Embora o algoritmo PNLMS apresente rápida convergência inicial quando comparado ao NLMS, essa taxa sofre grande redução durante o processo de convergência. Além disso, para o caso de respostas ao impulso não-esparsas, a convergência do PNLMS é mais lenta que a do NLMS [7]. Para compensar esses problemas, outros algoritmos proporcionais a partir do PNLMS foram propostos [12], [13]. Uma alternativa aos algoritmos proporcionais são os algoritmos que aplicam uma regularização dos coeficientes durante o processo de atualização. As penalizações mais comuns na literatura são regidas pela norma  $\ell_1$  ou pela norma  $\ell_0$ , sendo esta última o foco deste trabalho.

## III. ALGORITMO $\ell_0$ -SIGN-LMS

O algoritmo  $\ell_0$ -sign-LMS (uma das contribuições deste artigo) visa acelerar a convergência da adaptação para sistemas esparsos, apresentando simultaneamente robustez a ruído impulsivo de alta variância. Para isso, a penalização norma  $\ell_0$  é imposta ao algoritmo adaptativo sign-LMS (Eq. (3)) adicionando-a à função custo, como mostrado a seguir:

$$\mathcal{F}_{\ell_0\text{-SIGN}}(k) = |e(k)| + \frac{\kappa}{\beta} F_\rho[\mathbf{w}(k)], \quad (5)$$

onde  $\kappa$  é um parâmetro empregado para controlar a penalização imposta a soluções pouco esparsas e a função  $F_\rho[\mathbf{w}(k)]$  é uma função diferenciável por partes que aproxima a norma  $\ell_0$ . Uma escolha comum na literatura para essa função, e que será usada também nesse artigo, é [4], [9]:

$$F_\rho[\mathbf{w}(k)] = \sum_{n=0}^{N-1} \left(1 - e^{-\rho|w_n(k)|}\right), \quad (6)$$

onde  $\rho \in \mathbb{R}_+$  é um parâmetro arbitrável, o qual controla a aproximação da norma- $\ell_0$ .

Uma vez definida a função custo, é possível escrever a equação de atualização do  $\ell_0$ -sign-LMS da seguinte forma:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \mathbf{x}(k)\text{sign}[e(k)] + \kappa \mathbf{f}_\rho[\mathbf{w}(k)], \quad (7)$$

onde  $\mathbf{f}_\rho[\mathbf{w}(k)]$  é uma aproximação do negativo do gradiente de  $\mathcal{F}_\rho[\mathbf{w}(k)]$ .

Vale notar que  $\frac{\partial \mathcal{F}_\rho[\mathbf{w}(k)]}{\partial \mathbf{w}(k)}$  apresenta um custo computacional elevado por ter-se que calcular o resultado de uma exponencial a toda iteração. Por isso, será empregada uma aproximação de baixo custo desse gradiente:  $f_\rho[w_i(k)] \approx -\frac{\partial \mathcal{F}_\rho[\mathbf{w}(k)]}{\partial w_i(k)}$ . Tal aproximação é efetuada por meio de um truncamento dos dois primeiros termos da série de Taylor. O resultado pode ser visto como um termo de atração para zero, podendo ser descrito por [4]:

$$f_\rho[w_i(k)] = \begin{cases} \rho^2 w_i(k) + \rho, & -\frac{1}{\rho} \leq w_i(k) < 0 \\ \rho^2 w_i(k) - \rho, & 0 < w_i(k) \leq \frac{1}{\rho} \\ 0, & \text{no resto.} \end{cases} \quad (8)$$

Vale notar que a função  $f_\rho(x)$  implementa uma função que atrai valores de  $x$ , dentro de uma determinada faixa controlada por  $\rho$ , para 0 [4].

## IV. ANÁLISE DE TRANSIENTE

### A. Métodos de análise

A análise teórica de algoritmos adaptativos é de grande importância, pois cabe a ela fornecer de antemão informações sobre as características e o comportamento do algoritmo ao projetista. Diversas técnicas de análises foram empregadas na literatura ao longo dos anos, dentre as quais cumpre destacar: análise de esperança estatística exata [14], balanceamento de energia [15] e cálculo recursivo da matriz de autocorrelação dos desvios dos coeficientes [4]. A maior parte das técnicas tentam estimar de maneira confiável a evolução de duas métricas principais de desempenho em algoritmos adaptativos, o MSE (do

inglês, *Mean Square Error*) e o MSD (do inglês, *Mean Square Deviation*).

Este artigo efetua uma análise do algoritmo  $\ell_0$ -sign-LMS, empregando a técnica do cálculo recursivo da matriz de autocorrelação dos desvios. Tal técnica consiste em determinar equações a diferenças que consigam prever o comportamento de  $\mathbf{R}_{\tilde{\mathbf{w}}}(k) \triangleq \mathbb{E}[\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^T(k)]$ , onde  $\tilde{\mathbf{w}}(k) = \mathbf{w}^o - \mathbf{w}(k)$  é o vetor de desvio dos coeficientes adaptativos na  $k$ -ésima iteração.

Para que a análise seja possível, é necessário que algumas hipóteses sejam feitas:

**H1:**  $\mathbf{x}(k)$  e  $\tilde{\mathbf{w}}(k)$  são independentes.

**H2:** O ruído aditivo  $\nu(k)$ , de média zero e variância finita, é branco e independente de  $\mathbf{x}(k)$ . Sua variância no instante  $k$  é descrita por  $\sigma_\nu^2(k)$ .

**H3:** As seguintes aproximações são necessárias para que se tenha uma solução fechada:

$$\begin{aligned} \mathbb{E}\{\tilde{w}_i(k)f_\rho[w_j(k)]\} &\approx \mathbb{E}\{\tilde{w}_i(k)\}\mathbb{E}\{f_\rho[w_j(k)]\} \\ \mathbb{E}\{f_\rho[w_i(k)]f_\rho[w_j(k)]\} &\approx \mathbb{E}\{f_\rho[w_i(k)]\}\mathbb{E}\{f_\rho[w_j(k)]\} \end{aligned} \quad (9)$$

**H4:** A distribuição dos coeficientes  $\tilde{w}_i(k)$  é Gaussiana.

**H5:** O sinal de entrada é estacionário, gaussiano e todos os autovalores da sua matriz de autocorrelação são não nulos.

Com as hipóteses H1 e H2, podemos calcular o MSE segundo [16]:

$$\xi(k) \triangleq \mathbb{E}[e^2(k)] \approx \sigma_\nu^2(k) + \text{Tr}\{\mathbf{R}_{\tilde{\mathbf{w}}}\mathbb{E}[\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^T(k)]\}, \quad (10)$$

onde  $\mathbf{R}_{\tilde{\mathbf{w}}} = \mathbb{E}[\mathbf{x}(k)\mathbf{x}^T(k)]$  é a autocorrelação do sinal de entrada,  $\text{Tr}\{\mathbf{X}\}$  é o traço da matriz  $\mathbf{X}$  e  $\sigma_\nu^2(k)$  é a variância do ruído no instante  $k$ .

Já o MSD pode ser calculado pela seguinte equação:

$$\text{MSD}(k) \triangleq \mathbb{E}[\|\mathbf{w}^o - \mathbf{w}(k)\|^2] = \text{Tr}\{\mathbb{E}[\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^T(k)]\}. \quad (11)$$

## B. Descrição dos momentos de primeira e segunda ordens

A análise adotada neste artigo consistirá no acompanhamento da matriz de autocorrelação dos desvios. Portanto, devemos manipular a equação de atualização do algoritmo, substituindo  $\mathbf{w}(k)$  por  $\mathbf{w}^o - \tilde{\mathbf{w}}(k)$ , de modo a encontrar uma atualização recursiva dos desvios. Para que a análise teórica possa ser feita de maneira mais simples, doravante trabalharemos com a equação de atualização na sua forma escalar, como a seguir:

$$\tilde{w}_i(k+1) = \tilde{w}_i(k) - \beta x(k-i)\text{sign}[e(k)] - \kappa f_\rho[w_i(k)]. \quad (12)$$

Para que possamos inferir as estatísticas de primeira e segunda ordem, devemos primeiro escrever o erro  $e(k)$  da seguinte forma:

$$\begin{aligned} e(k) &= d(k) - y(k) = (\mathbf{w}^o)^T \mathbf{x}(k) + \nu(k) - \mathbf{w}^T(k)\mathbf{x}(k) \\ &= \sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k) \end{aligned} \quad (13)$$

Substituindo (13) na equação (12), temos:

$$\begin{aligned} \tilde{w}_i(k+1) &= \tilde{w}_i(k) - \kappa f_\rho[w_i(k)] \\ &\quad - \beta x(k-i)\text{sign}\left[\sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k)\right]. \end{aligned} \quad (14)$$

Uma vez definida a equação de atualização dos desvios na sua forma escalar, devemos propor uma equação estatística que descreva o comportamento do algoritmo na média,

uma vez que (14) é uma equação determinística. Cabe aplicar o operador do valor esperado à equação (14), mas para que seja possível prosseguir a análise com o valor esperado, é necessário aplicar o teorema de Price [17], [18]. Tal teorema permite que escrevamos:

$$\mathbb{E}[\text{sign}(A)B] = \sqrt{\frac{2}{\pi\sigma_A^2}}\mathbb{E}[AB], \quad (15)$$

onde  $A$  e  $B$  são duas variáveis aleatórias conjuntamente Gaussianas e de média zero, e  $\sigma_A^2 = \mathbb{E}[A^2]$ . Portanto, o teorema de Price permite que obtermos a seguinte expressão para o momento de primeira ordem:

$$\begin{aligned} \mathbb{E}[\tilde{w}_i(k+1)] &= \mathbb{E}[\tilde{w}_i(k)] - \beta \sqrt{\frac{2}{\pi\sigma_e^2}} \sum_{n=0}^{N-1} \mathbb{E}[\tilde{w}_n(k)]r_{in} \\ &\quad - \kappa \mathbb{E}\{f_\rho[w_i(k)]\}, \end{aligned} \quad (16)$$

onde  $\sigma_e^2$  é a variância do erro e  $r_{in} \triangleq \mathbb{E}[x(k-i)x(k-n)]$ . Para estimar a evolução da matriz de autocorrelação dos desvios, devemos estimar a evolução dos coeficientes de  $\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^T(k)$  na média. Com base na equação (14), a equação do cálculo recursivo para um coeficiente genérico dessa matriz é dada por:

$$\begin{aligned} \tilde{w}_i(k+1)\tilde{w}_j(k+1) &= \tilde{w}_i(k)\tilde{w}_j(k) \\ &\quad - \beta \tilde{w}_i(k)x(k-j)\text{sign}(\sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k)) \\ &\quad - \kappa \tilde{w}_i(k)f_\rho[w_j(k)] \\ &\quad - \beta \tilde{w}_j(k)x(k-i)\text{sign}(\sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k)) \\ &\quad + \beta^2 x(k-i)x(k-j)\text{sign}^2(\sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k)) \\ &\quad + \kappa \beta x(k-i)\text{sign}(\sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k))f_\rho[w_j(k)] \\ &\quad - \kappa \tilde{w}_j(k)f_\rho[w_i(k)] \\ &\quad + \kappa \beta x(k-j)\text{sign}(\sum_{n=0}^{N-1} \tilde{w}_n(k)x(k-n) + \nu(k))f_\rho[w_i(k)] \\ &\quad + \kappa^2 f_\rho[w_i(k)]f_\rho[w_j(k)]. \end{aligned} \quad (17)$$

Como a matriz de autocorrelação dos desvios é calculada via  $\mathbb{E}[\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^T(k)]$ , devemos aplicar o operador do valor esperado na equação (17), a qual passa a ser escrita de acordo com (18).

$$\begin{aligned} \mathbb{E}[\tilde{w}_i(k+1)\tilde{w}_j(k+1)] &= \mathbb{E}[\tilde{w}_i(k)\tilde{w}_j(k)] \\ &\quad - \beta \sqrt{\frac{2}{\pi\sigma_e^2}} \sum_{n=0}^{N-1} \mathbb{E}[\tilde{w}_i(k)\tilde{w}_n(k)]r_{jn} - \kappa \mathbb{E}[\tilde{w}_i(k)]\mathbb{E}[f_\rho[w_j(k)]] \\ &\quad - \beta \sqrt{\frac{2}{\pi\sigma_e^2}} \sum_{n=0}^{N-1} \mathbb{E}[\tilde{w}_j(k)\tilde{w}_n(k)]r_{in} + \beta^2 r_{ij} \\ &\quad + \kappa \beta \sqrt{\frac{2}{\pi\sigma_e^2}} \sum_{n=0}^{N-1} \mathbb{E}[\tilde{w}_n(k)]r_{in}\mathbb{E}[f_\rho[w_j(k)]] \\ &\quad - \kappa \mathbb{E}[\tilde{w}_j(k)]\mathbb{E}[f_\rho[w_i(k)]] + \kappa^2 \mathbb{E}[f_\rho[w_i(k)]]\mathbb{E}[f_\rho[w_j(k)]] \\ &\quad + \kappa \beta \sqrt{\frac{2}{\pi\sigma_e^2}} \sum_{n=0}^{N-1} \mathbb{E}[\tilde{w}_n(k)]r_{jn}\mathbb{E}[f_\rho[w_i(k)]]. \end{aligned} \quad (18)$$

Uma vez definidas as equações recursivas para os momentos de primeira e segunda ordem, cabe agora empreender o cálculo do termo  $\mathbb{E}[f_\rho[w_i(k)]]$  aplicando a hipótese H4:

$$\begin{aligned} \mathbb{E}[f_\rho(w_i(k))] &= \frac{1}{\sqrt{2\pi\sigma_{i,k}}} \int_{-\infty}^{\infty} f_\rho[w_i(k)] e^{-\frac{(w_i(k)-\mu_{i,k})^2}{2\sigma_{i,k}^2}} dw_i(k) \\ &= \frac{\rho^2 \sigma_{i,k}}{\sqrt{2\pi}} \left[ e^{\frac{-(\mu_{i,k} + \frac{1}{\rho})^2}{2\sigma_{i,k}^2}} - e^{\frac{-(\mu_{i,k} - \frac{1}{\rho})^2}{2\sigma_{i,k}^2}} \right] \\ &\quad + \frac{\rho^2 \mu_{i,k}}{2} \left[ \text{erf}\left(\frac{\mu_{i,k} + \frac{1}{\rho}}{\sqrt{2}\sigma_{i,k}}\right) - \text{erf}\left(\frac{\mu_{i,k} - \frac{1}{\rho}}{\sqrt{2}\sigma_{i,k}}\right) \right] \\ &\quad + \frac{\rho}{2} \left[ \text{erf}\left(\frac{\mu_{i,k} + \frac{1}{\rho}}{\sqrt{2}\sigma_{i,k}}\right) + \text{erf}\left(\frac{\mu_{i,k} - \frac{1}{\rho}}{\sqrt{2}\sigma_{i,k}}\right) - 2\text{erf}\left(\frac{\mu_{i,k}}{\sqrt{2}\sigma_{i,k}}\right) \right], \end{aligned} \quad (19)$$

onde  $\text{erf}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ .

Para tanto, sejam  $\mu_{i,k}$  e  $\sigma_{i,k}^2$  a média e a variância de  $\tilde{w}_i(k)$ , respectivamente. Como  $w_i(k) = w_i^o - \tilde{w}_i(k)$ , e  $w_i^o$  é determinístico, podemos observar que a variância de  $w_i(k)$  também é dada por  $\sigma_{i,k}^2$ , e a média de  $w_i(k)$  é dada por  $\bar{\mu}_{i,k} = w_i^o - \mu_{i,k}$ .

## V. SIMULAÇÕES

Para avaliar a adequação do modelo estocástico proposto neste trabalho ao caso real, foram feitas diversas simulações, nas quais um filtro adaptativo tenta identificar uma dentre as funções de transferência descritas pelos primeiros  $N$  coeficientes de [19]. Algumas dessas funções de transferência possuem coeficientes muito pequenos, então um fator de escalonamento  $\alpha$  foi utilizado. O filtro adaptativo possui o mesmo comprimento  $N$  e o sinal de entrada é gerado passando um ruído branco e gaussiano por um filtro de colorimento com função de transferência:  $H(z) = 1 - 0,6z + 0,2z^{-1} - 0,3z^{-2}$ .

### A. Comportamento em Regime Transiente

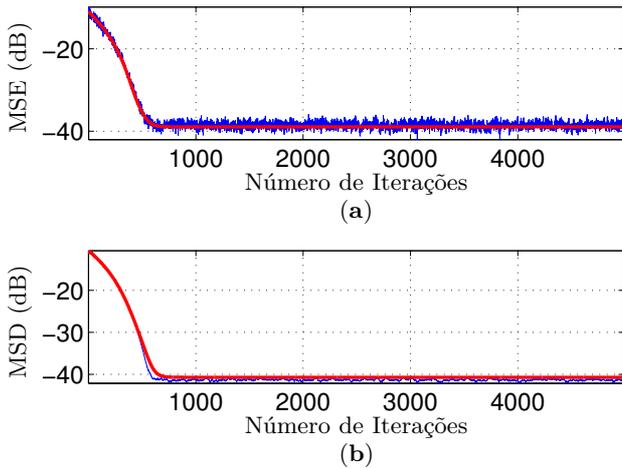


Fig. 2. Desempenho do algoritmo  $\ell_0$ -sign-LMS. Simulações de Monte Carlo (em azul) e modelo proposto (em vermelho). (a) MSE (dB); (b) MSD (dB).

Para um primeiro exemplo, foi utilizado o Modelo 1 de [19] com  $N = 12$ ,  $\alpha = 1$ , 100 médias de Monte Carlo,  $\beta = 10^{-3}$ ,  $\delta = 10^{-6}$ ,  $\sigma_v^2 = 10^{-6}$ ,  $\kappa = 10^{-9}$  e  $\rho = 20$ . A Figura 2 mostra a evolução do algoritmo  $\ell_0$ -sign-LMS, sendo possível perceber que o modelo proposto consegue rastrear o comportamento do filtro real tanto no caso do MSE quanto no caso do MSD.

Em um segundo exemplo, foi observado o comportamento do sistema quando  $\sigma_v^2$  varia a cada iteração. Foi utilizado o Modelo 8 de [19] com  $N = 12$ ,  $\alpha = 1$ , 100 médias de Monte Carlo,  $\beta = 10^{-3}$ ,  $\delta = 10^{-6}$ ,  $\kappa = 10^{-9}$  e  $\rho = 20$ . A Figura 3 mostra que o modelo estocástico proposto é capaz de prever com precisão a evolução do MSE ao longo das iterações.

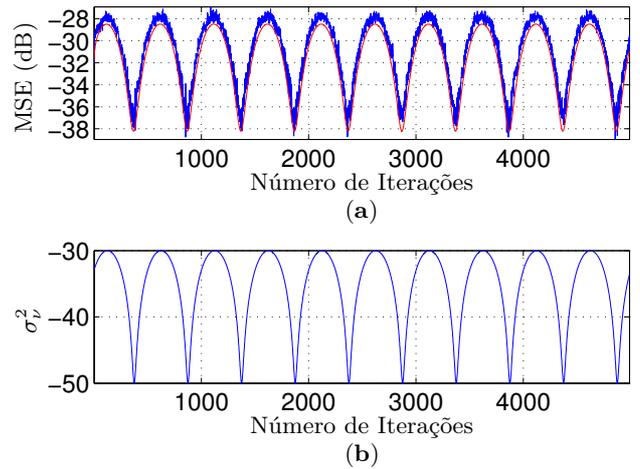


Fig. 3. (a) Evolução do MSE para o algoritmo  $\ell_0$ -sign-LMS com  $\sigma_v^2$  variando. Simulações de Monte Carlo (em azul) e modelo proposto (em vermelho). (b) Comportamento de  $\sigma_v^2$  ao longo das iterações.

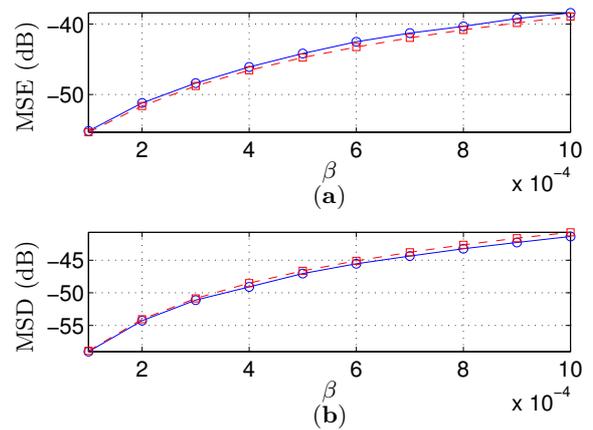


Fig. 4. Desempenho do algoritmo  $\ell_0$ -sign-LMS em regime permanente para diferentes valores de  $\beta$ . Simulações de Monte Carlo (em azul) e modelo proposto (em vermelho). (a) MSE (dB); (b) MSD (dB).

### B. Comportamento em Regime Permanente

Em um terceiro exemplo, foi observado o comportamento do sistema em regime permanente para diversos valores de  $\beta$ . Para essa simulação, foi utilizado o Modelo 4 de [19] com  $N = 12$ ,  $\alpha = 10$ , 10 médias de Monte Carlo,  $\delta = 10^{-6}$ ,  $\sigma_v^2 = 10^{-6}$ ,  $\kappa = 10^{-9}$  e  $\rho = 20$ . É possível perceber (vide Figura 4) que o modelo proposto consegue prever o MSE e o MSD com precisão maior quando  $\beta$  é pequeno, o que é coerente com o fato de que o teorema de Price é mais acurado quando o valor de  $\beta$  é baixo [20], [21].

Em um quarto exemplo, foi observado o comportamento do sistema em regime permanente para diversos valores de  $\kappa$  (vide Figura 5). Para essa simulação, foi utilizado o Modelo 5 de [19] com  $N = 12$ ,  $\alpha = 10$ , 10 médias de Monte Carlo,  $\delta = 10^{-6}$ ,  $\sigma_v^2 = 10^{-6}$ ,  $\beta = 10^{-3}$  e  $\rho = 20$ . Neste exemplo, é interessante notar que o modelo proposto

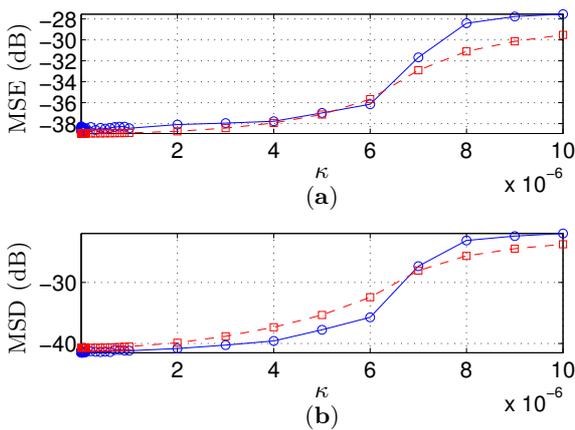


Fig. 5. Desempenho do algoritmo  $\ell_0$ -sign-LMS em regime permanente para diferentes valores de  $\kappa$ . Simulações de Monte Carlo (em azul) e modelo proposto (em vermelho). (a) MSE (dB); (b) MSD (dB).

se adequa melhor para valores de  $\kappa$  pequenos. A Figura 6 ilustra a robustez do algoritmo proposto perante ruído impulsivo.

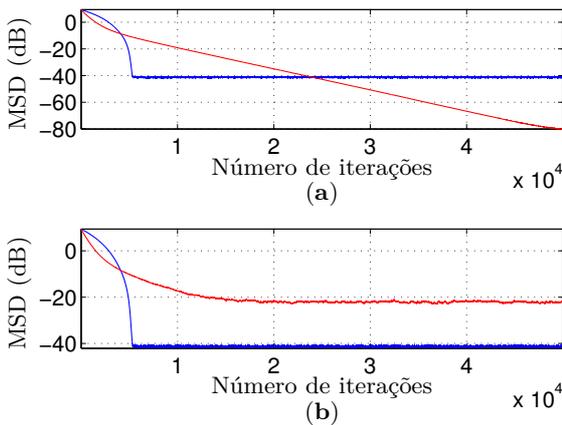


Fig. 6. Desempenho dos algoritmos  $\ell_0$ -sign-LMS (em azul) e  $\ell_0$ -LMS (em vermelho). (a) Sem a presença de ruído impulsivo. (b) Com a presença de ruído impulsivo, com distribuição gaussiana de média = 0 e  $\sigma^2 = 10$  e com probabilidade de ocorrência igual a 1%.

## VI. CONCLUSÃO

Neste artigo apresentamos uma análise teórica da convergência em transiente e em regime permanente do algoritmo  $\ell_0$ -sign-LMS. Com base nas simulações apresentadas, pôde-se observar que o modelo teórico mostrou-se capaz de prever o comportamento do algoritmo estudado ao longo deste trabalho de maneira satisfatória, prevendo a evolução das métricas de desempenho adequadamente, sem apresentar restrições quanto ao sinal de entrada ser branco e à variância do ruído ser constante.

## REFERÊNCIAS

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

- [2] M. S. Mohammed and K. Ki-Seong, "Sign least mean squares-based deconvolution technique for ultrasonic testing," *Russian Journal of Nondestructive Testing*, vol. 48, pp. 609–613, Oct 2012.
- [3] G. Gui and L. Xu, "Regularization parameter selection method for sign lms with reweighted  $\ell_1$ -norm constraint algorithm," *arXiv preprint arXiv:1503.03608*, Apr 2015.
- [4] K. da S. Olinto, D. B. Haddad, and M. R. Petraglia, "Transient analysis of  $\ell_0$ -lms and  $\ell_0$ -nlms algorithms," *Signal Processing*, vol. 127, pp. 217 – 226, Jan 2016.
- [5] R. K. Martin, W. A. Sethares, R. C. Williamson, and C. R. Johnson, "Exploiting sparsity in adaptive filters," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1883–1894, Aug 2002.
- [6] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 508–518, Sep 2000.
- [7] 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, Jul 2014.
- [8] L. R. Vega, H. Rey, J. Benesty, and S. Tressens, "A family of robust algorithms exploiting sparsity in adaptive filters," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, pp. 572–581, May 2009.
- [9] Y. Gu, J. Jin, and S. Mei, " $\ell_0$  norm constraint lms algorithm for sparse system identification," *IEEE Signal Processing Letters*, vol. 16, pp. 774–777, Sep 2009.
- [10] G. Su, J. Jin, Y. Gu, and J. Wang, "Performance analysis of  $\ell_0$  norm constraint least mean square algorithm," *IEEE Transactions on Signal Processing*, vol. 60, pp. 2223–2235, May 2012.
- [11] N. Kalouptsidis, G. Mileounis, B. Babadi, and V. Tarokh, "Adaptive algorithms for sparse system identification," *Signal Processing*, vol. 91, pp. 1910 – 1919, Feb 2011.
- [12] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Signals, Systems & Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on*, vol. 1, pp. 394–398, IEEE, Nov 1998.
- [13] H. Deng and M. Doroslovacki, "Proportionate adaptive algorithms for network echo cancellation," *IEEE Transactions on Signal Processing*, vol. 54, pp. 1794–1803, Apr 2006.
- [14] S. C. Douglas and W. Pan, "Exact expectation analysis of the lms adaptive filter," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2863–2871, Dec 1995.
- [15] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters with error nonlinearities," *IEEE Transactions on Signal Processing*, vol. 51, pp. 653–663, Mar 2003.
- [16] P. S. R. Diniz, *Adaptive filtering*. Springer, 1997.
- [17] R. Price, "A useful theorem for nonlinear devices having gaussian inputs," *IRE Transactions on Information Theory*, vol. 4, pp. 69–72, Jun 1958.
- [18] E. Eweda, "Analysis and design of a signed regressor lms algorithm for stationary and nonstationary adaptive filtering with correlated gaussian data," *IEEE Transactions on Circuits and Systems*, vol. 37, pp. 1367–1374, Nov 1990.
- [19] I. TSG, "15, digital network echo cancellers (recommendation)," tech. rep., Tech. Rep. G. 168, ITU-T, 2004.
- [20] E. V. Papoulis and T. Stathaki, "A normalized robust mixed-norm adaptive algorithm for system identification," *IEEE Signal Processing Letters*, vol. 11, pp. 56–59, Jan 2004.
- [21] V. Mathews and S. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 450–454, Apr 1987.