

Novo Algoritmo Proporcional com Reúso de Coeficientes e Seleção de Dados

Rodrigo M. S. Pimenta^{1,2}, Leonardo C. Resende^{2,3}, Diego B. Haddad⁴, Mariane R. Petraglia², Newton N. Siqueira^{2,5}

Resumo— Algoritmos adaptativos devem idealmente apresentar uma alta taxa de convergência, um bom desempenho em regime permanente e um baixo custo computacional. Poucos algoritmos conseguem atender simultaneamente a estes requisitos. Este artigo propõe um problema de otimização local e determinístico cuja solução dá origem a um algoritmo adaptativo que apresenta uma maior taxa de convergência na identificação de sistemas esparsos, devido ao concurso da técnica de adaptação proporcional. Ademais, uma menor complexidade computacional é angariada pela seleção de dados (ensejada pela técnica *Set-Membership*), a qual evita atualizações quando o sinal de entrada não disponibiliza uma quantidade de informação considerada relevante. Por fim, a adoção do reúso de coeficientes permite obter um bom desempenho de identificação em regime permanente, notadamente em cenários de baixa SNR.

Palavras-Chave— Filtragem Adaptativa, Adaptação Proporcional, Reúso de Coeficientes, *Set-Membership*.

Abstract— Adaptive algorithms should ideally present a high convergence rate, a good steady state performance, and a low computational cost. Few algorithms can simultaneously meet these requirements. This paper proposes a local and deterministic optimization problem whose solution gives rise to an adaptive algorithm that presents a higher convergence rate in the identification of sparse systems due to the proportionate adaptation technique. In addition, a lower computational complexity is raised by the selection of data (obtained by the *Set-Membership* technique), which avoids updates when the input signal does not provide an amount of information considered relevant. Finally, the adoption of the reuse of coefficients allows to obtain a good performance of identification in steady-state conditions, notably in low SNR scenarios.

Keywords— Adaptive Filtering, Proportionate Adaptation, Coefficients Reuse, *Set-Membership*.

I. INTRODUÇÃO

O advento das técnicas de processamento digital de sinais permitiu a implementação de algoritmos de filtragem adaptativa os mais diversos, capazes de responder a exigentes requisitos de diversas aplicações desafiadoras, tais como cancelamento de eco acústico, equalização, predição, redução de interferências e identificação de sistemas [1], sendo esta última o foco deste artigo.

¹Coordenação de Telecomunicações, CEFET-RJ campus Maracanã, Rio de Janeiro-RJ, Brasil. ²Programa de Engenharia Elétrica (PEE), COPPE/UFRJ, Rio de Janeiro-RJ, Brasil. ³Instituto Federal do Rio de Janeiro. ⁴Coordenação de Engenharia de Computação, CEFET-RJ campus Petrópolis, Petrópolis-RJ, Brasil. ⁵Coordenação de Telecomunicações, CEFET-RJ campus Nova Iguaçu, Nova Iguaçu, Brasil. E-mails: rodrigo.pimenta@cefet-rj.br, leonardo.resende@ifrrj.edu.br, diego@pads.ufrj.br, mariane@pads.ufrj.br, newton.siqueira@cefet-rj.br.

Seja $w^0 \in \mathbb{R}^N$ um vetor¹ que contenha os coeficientes de uma função de transferência a identificar, a cuja saída $d(k)$ temos acesso, a menos de um ruído de medição imprevisível $\nu(k)$:

$$d(k) = x^T(k)w^0 + \nu(k), \quad (1)$$

onde $x(k) \in \mathbb{R}^N$ é um vetor que coleta N amostras consecutivas do sinal $x(k)$, de sorte que (1) reflete, a menos do ruído $\nu(k)$, uma filtragem empreendida por uma estrutura transversal [2].

Comumente, algoritmos de filtragem adaptativa, tendo acesso ao par de informações $\{x(k), d(k)\}$, implementam uma atualização recursiva do vetor adaptativo $w(k)$, de modo que $w(k) \in \mathbb{R}^N$ se aproxime² cada vez mais de w^0 . As equações de atualização destes algoritmos não raro se valem do erro

$$e(k) \triangleq d(k) - y(k) = d(k) - \overbrace{w^T(k)x(k)}^{=y(k)}, \quad (2)$$

o qual, refletindo a discrepância entre a saída do filtro $y(k)$ e o sinal de referência $d(k)$, fornece uma indicação sobre o quão longe o filtro encontra-se de seu estado ideal. A Fig. 1 apresenta o diagrama de blocos de um algoritmo de identificação de sistemas.

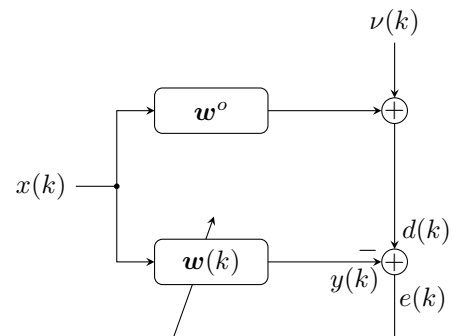


Fig. 1. Diagrama em blocos da estrutura de um algoritmo de filtragem adaptativa aplicado à identificação de sistemas.

Este artigo estrutura-se do seguinte modo: na Seção II, os algoritmos clássicos de filtragem adaptativa são apresentados. A Seção III descreve sucintamente alguns algoritmos cujas propriedades serão aproveitadas no algoritmo proposto, cuja derivação se dá na Seção IV. Por fim, a Seção V (de resultados) precede a Seção VI, que contempla as conclusões do artigo.

¹Neste artigo, todos os vetores são do tipo coluna.

²Doravante suporemos que o comprimento de $w(k)$ equivale ao de w^0 .

II. ALGORITMOS CLÁSSICOS

O primeiro algoritmo da família é o LMS (*Least-Mean Squares* [3]) nascido em 1960 pela pesquisa de Bernard Widrow e seu aluno Marcian E. Hoff [4]. A regra de aprendizado de Widrow-Hoff foi introduzida em conjunto com a rede linear Adaline (*Adaptive Linear Neuron*) e é similar à do *Perceptron* [4], mas com a habilidade de permitir saídas não limitadas somente a zeros e uns. A equação de atualização do LMS, descrita por

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

implica uma baixa complexidade computacional do algoritmo, o qual entretanto apresenta algumas desvantagens, dentre as quais um limite superior (que evita o fenômeno da divergência) para o fator de aprendizagem β extremamente dependente das propriedades estatísticas do sinal de entrada $\mathbf{x}(k)$ [1].

O algoritmo NLMS (*Normalized Least-Mean Squares*) (4) é uma variação do algoritmo LMS que implementa uma normalização pela energia do vetor de entrada que, pode-se demonstrar, garante a estabilidade caso tenhamos $\beta \in [0, 1]$ (cabendo ressaltar que β encontra-se sob o arbítrio do projetista). Sua equação de atualização padrão

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \frac{\mathbf{x}(k)e(k)}{\|\mathbf{x}(k)\|^2} \quad (4)$$

pode por vezes apresentar instabilidade numérica, caso o vetor de entrada $\mathbf{x}(k)$ apresente uma energia muito baixa. Normalmente, tal problema é contornado mediante a inserção do fator de regularização δ , resultando no algoritmo δ -NLMS, descrito por:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \frac{\mathbf{x}(k)e(k)}{\|\mathbf{x}(k)\|^2 + \delta}. \quad (5)$$

O emprego da técnica dos multiplicadores de Lagrange permite demonstrar que os algoritmos LMS, NLMS e δ -NLMS podem ser obtidos mediante a otimização de um problema local e determinístico com restrição:

$$\min_{\mathbf{w}(k+1)} \mathcal{F}_{\text{MD}}[\mathbf{w}(k+1)] \quad (6)$$

$$\text{sujeito a } e_p(k) = f[\mathbf{x}(k), \beta, e(k)],$$

onde diz-se que $\mathcal{F}_{\text{MD}}[\mathbf{w}(k+1)] \triangleq \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2$, por penalizar soluções $\mathbf{w}(k+1)$ distantes de $\mathbf{w}(k)$, implementa o princípio da *distorção mínima* e $e_p(k)$, conhecido como *erro a posteriori*, calcula a diferença entre o sinal de referência $d(k)$ e a saída do filtro para a entrada $\mathbf{x}(k)$ após tais informações terem sido empregadas para a atualização do filtro, ou seja:

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

Por fim, diferentes funções $f(\cdot)$ em (6) dão origem a diferentes algoritmos, de sorte que

$$f_{\text{LMS}}[\mathbf{x}(k), \beta, e(k)] \triangleq (1 - \beta\|\mathbf{x}(k)\|^2) e(k), \quad (8)$$

$$f_{\text{NLMS}}[\mathbf{x}(k), \beta, e(k)] \triangleq (1 - \beta) e(k) \quad (9)$$

e

$$f_{\delta\text{-NLMS}}[\mathbf{x}(k), \beta, e(k)] \triangleq \left(1 - \beta \frac{\|\mathbf{x}(k)\|^2}{\|\mathbf{x}(k)\|^2 + \delta}\right) e(k) \quad (10)$$

engendram como resultado as Equações (3), (4) e (5), respectivamente [5].

III. ALGUNS ALGORITMOS ADAPTATIVOS COM PROPRIEDADES DE INTERESSE

O algoritmo a ser proposto neste artigo conjuga vantagens de três diferentes algoritmos adaptativos, os quais serão sucintamente descritos a seguir.

A. Algoritmo RC (*Reusing Coefficient*)

Diversos algoritmos adaptativos (como o *Affine Projection Algorithm* [1]) aumentam a taxa de convergência por meio do reúso dos dados de entrada. Tais algoritmos costumam apresentar como desvantagem uma degradação do desempenho em regime permanente. Tais características podem ser consideradas duais às apresentadas pelo algoritmo RC (sigla para “reúso de coeficientes”) [6], o qual implementa um reúso dos últimos L vetores de coeficientes adaptativos. Os algoritmos da família RC aliam um melhor desempenho em regime permanente a uma perda na taxa de convergência [7]; por essa razão, [8] propõe e analisa o emprego conjunto de ambas as estratégias de reúso.

O algoritmo RC-NLMS pode ser descrito como um solucionador do seguinte problema de otimização:

$$\min_{\mathbf{w}(k+1)} \mathcal{F}_{\text{RC}}[\mathbf{w}(k+1)] \quad (11)$$

$$\text{sujeito a } e_p(k) = (1 - \beta)e'(k),$$

onde

$$\mathcal{F}_{\text{RC}}[\mathbf{w}(k+1)] \triangleq \sum_{l=0}^{L-1} \rho^l \|\mathbf{w}(k+1) - \mathbf{w}(k-l)\|^2 \quad (12)$$

e $\rho \in (0, 1]$ é um parâmetro ao arbítrio do projetista, com

$$e'(k) \triangleq d(k) - \frac{\rho - 1}{\rho^L - 1} \sum_{l=0}^{L-1} \rho^l \mathbf{w}^T(k-l)\mathbf{x}(k). \quad (13)$$

A solução de (11) dá origem à seguinte equação de atualização:

$$\mathbf{w}(k+1) = \frac{\rho - 1}{\rho^L - 1} \sum_{l=0}^{L-1} \rho^l \mathbf{w}(k-l) + \beta \frac{\mathbf{x}(k)e'(k)}{\|\mathbf{x}(k)\|^2}, \quad (14)$$

de sorte que $\mathbf{w}(k+1)$ depende de uma soma ponderada dos últimos L vetores $\mathbf{w}(k-l)$ (com $l \in \{0, 1, \dots, L-1\}$), a qual suaviza as oscilações do filtro. Tal propriedade é responsável pela melhora de desempenho em regime permanente, notadamente em configurações de baixa razão sinal-ruído [6].

B. Algoritmo SM (*Set-Membership*)

A técnica *Set-Membership* apresenta uma nova abordagem de estimação baseada num limiar referente à magnitude do erro, a partir do qual faz-se necessária a atualização dos

coeficientes do filtro. Para os casos em que o erro não ultrapassa o limiar pré-definido, os coeficientes adaptativos são mantidos sem alteração, o que propicia a redução do custo computacional. O problema de otimização que cumpre ao algoritmo *Set Membership* resolver pode ser descrito por:

$$\begin{cases} \min_{\mathbf{w}(k+1)} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 \\ \text{s.t. } |e_p(k)| \leq \bar{\gamma}, \end{cases} \quad (15)$$

onde $\bar{\gamma} = \sqrt{\tau}\sigma_v$, dando origem à equação de atualização

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \Delta \mathbf{w}(k), \quad (16)$$

onde verifica-se introdução de um fator de relaxamento $\beta \in (0, 1]$, cuja escolha deve contemplar o compromisso entre taxa de convergência e desempenho em regime permanente, e com $\Delta \mathbf{w}(k)$ sendo definido por:

$$\Delta \mathbf{w}(k) \triangleq \begin{cases} \frac{[1 - \frac{\bar{\gamma}}{|e(k)|}] \mathbf{x}(k) e(k)}{\|\mathbf{x}(k)\|^2}, & \text{para } |e(k)| < \bar{\gamma}, \\ 0, & \text{caso contrário.} \end{cases} \quad (17)$$

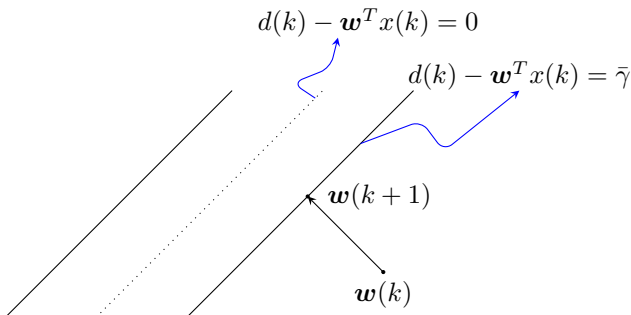


Fig. 2. Interpretação geométrica do algoritmo *Set-Membership*. Caso $\mathbf{w}(k)$ esteja na região onde $e(k) > \bar{\gamma}$, a atualização SM com $\beta = 1$ implica uma projeção ortogonal de $\mathbf{w}(k)$ no hiperplano no qual o erro *a posteriori* é igual a $\bar{\gamma}$.

A interpretação geométrica do processo de atualização do algoritmo SM-NLMS (Eq. (16)) é apresentada na Figura 2. A maior vantagem do algoritmo SM é a redução na complexidade computacional frente aos algoritmos clássicos como NLMS, RLS (*Recursive-Least-Squares*) [1] e AP (*Affine Projection*) [9]. Dada esta propriedade desejável da técnica SM, diversos autores propõem conjugá-lo com outras técnicas. Assim, temos por exemplo o algoritmo SM-BNDRLMS (*Set-Membership Binormalized Data-Reusing LMS*) [10], o qual reduzir a dependência da taxa de convergência com relação aos autovalores da matriz de correlação do sinal de entrada. Outro exemplo é o SM-AP (*Set-Membership Affine Projection*) [11] que se utiliza do reúso de dados do algoritmo AP conjugado a uma menor complexidade computacional (advinda do emprego do SM).

C. Algoritmo PNLMS (*Proportionate Normalized Least-Mean Squares*)

O algoritmo PNLMS [12] se vale da adaptação proporcional para aumentar a taxa de identificação de sistemas esparsos. Essa eficiência é possível porque a energia de atualização é

distribuída de modo aproximadamente proporcional à magnitude dos coeficientes, de sorte que coeficientes de maior magnitude tendam a sofrer atualizações mais pronunciadas. Os passos proporcionais (específicos para cada coeficiente adaptativo) são implementados por uma matriz diagonal $\Lambda(k)$ e a equação de atualização do algoritmo é dada por

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta \frac{\Lambda(k) \mathbf{x}(k) e(k)}{\|\mathbf{x}(k)\|_{\Lambda(k)}^2 + \delta}, \quad (18)$$

a qual pode ser compreendida como solucionadora do seguinte problema de otimização com restrições:

$$\min_{\mathbf{w}(k+1)} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|_{\Lambda^{-1}(k)}^2 \quad (19)$$

$$e_p(k) = \left(1 - \beta \frac{\|\mathbf{x}(k)\|_{\Lambda(k)}^2}{\|\mathbf{x}(k)\|_{\Lambda(k)}^2 + \delta} \right) e(k).$$

Diferentes escolhas dos elementos da diagonal de $\Lambda(k)$ dão lugar a diferentes algoritmos proporcionais. De modo a solucionar o problema de baixa convergência dos coeficientes adaptativos de baixa magnitude, foi proposto o algoritmo MPNLMS (μ -law PNLMS) [13], o qual advoga o emprego do logaritmo da magnitude como *passo de aprendizagem*. O algoritmo MPNLMS é mais robusto do que o PNLMS, sendo capaz de tratar coeficientes de elevada e de baixa magnitude, mas sua vantagem ocorre apenas quando as funções de transferência são esparsas. No caso de tais funções serem dispersivas, o algoritmo clássico NLMS converge mais rapidamente do que o MPNLMS. Na tentativa de solucionar o problema foi proposto o algoritmo IPNLMS (*Improved PNLMS*) [14], cuja escolha dos coeficientes da diagonal principal de $\Lambda(k)$ depende de um parâmetro α , o qual permite balancear o ajuste entre os algoritmos NLMS (caso em que $\alpha = -1$) e PNLMS (caso em que $\alpha \simeq 1$), sendo esses os limites inferior e superior da escolha de α . [14] relata que a escolha $\alpha = -0,5$ permite ao IPNLMS superar tanto o NLMS quanto o PNLMS original, independentemente da distribuição de energia da função de transferência.

Outra abordagem para solucionar a limitação do algoritmo MPNLMS a sinais esparsos é o algoritmo IMPNLMS (*Improved μ -law PNLMS*), proposto em [15] e analisado em [16]. O IMPNLMS é capaz de convergir tão rápido quanto os algoritmos PNLMS e MPNLMS para funções de transferência esparsas, tendo um desempenho similar ao do NLMS na identificação de funções de transferência dispersivas. A técnica adapta-se à esparsidade estimada do canal por meio de

$$\zeta_w(k) = \frac{N}{N - \sqrt{N}} + \left(1 - \frac{\|\mathbf{w}(k)\|_1}{\sqrt{N} \|\mathbf{w}(k)\|_2} \right), \quad (20)$$

onde ζ_w varia entre 0 (função de transferência dispersiva) e 1 (função de transferência esparsa), utilizando um fator de esquecimento λ' para estimar a esparsidade de forma adaptativa, sendo $\zeta(k) = (1 - \lambda')\zeta(k-1) + \lambda'\zeta_w(k)$, com $0 < \lambda' \leq 1$. O parâmetro de ajuste α agora é variável consoante a expressão

$$\alpha(k) = 2\zeta(k) - 1. \quad (21)$$

O i -ésimo elemento $g_i(k)$ da matriz diagonal $\Lambda(k)$ apre-

senta magnitude conforme

$$g_i(k) = \frac{1 - \alpha(k)}{2N} + \frac{(1 + \alpha(k))F(|w_i(k)|)}{2 \sum_{n=0}^{N-1} F(|w_n(k)|) + \epsilon}, \quad (22)$$

onde ϵ é um parâmetro de regularização e

$$F(|w_i(k)|) = \begin{cases} 400|w_i(k)|, & \text{para } |w_i(k)| < 0,005 \\ 8,51|w_i(k)| + 1,96, & \text{caso contrário.} \end{cases} \quad (23)$$

IV. ALGORITMO PROPOSTO

As vantagens dos algoritmos RC-NLMS, SM-NLMS e IMPNLMS podem ser combinadas, dando origem a um novo algoritmo, intitulado RC-SM-IMPNLMS (do inglês *Reusing Coefficient Set-Membership Improved μ -law Proportional Normalized Least-Mean Squares*). Tal algoritmo deve conjugar uma alta taxa de convergência na identificação de sistemas esparsos (propriedade herdada dos algoritmos proporcionais) a um desempenho melhorado em regime permanente (devido ao reúso de coeficientes) e a um menor custo computacional (oriundo do emprego do conceito de filtragem *Set-Membership*).

O problema de otimização que cumpre ao algoritmo RC-SM-PNLMS solucionar pode ser descrito por:

$$\begin{aligned} \min_{\mathbf{w}(k+1)} \mathcal{F}_{\text{RC-SMP}}[\mathbf{w}(k+1), e(k)] \quad (24) \\ \text{sujeito a } |e_p(k)| \leq \bar{\gamma}, \end{aligned}$$

onde

$$\mathcal{F}_{\text{RC-SMP}}[\mathbf{w}(k+1), e(k)] \triangleq \sum_{l=0}^{L\{e(k)\}-1} \rho^l \|\mathbf{w}(k+1) - \mathbf{w}(k-l)\|^2, \quad (25)$$

com o fator de reúso L sendo uma função de $e(k)$, por razões a serem posteriormente elucidadas.

O problema de otimização com restrições (24) pode ser convertido, por meio da técnica dos multiplicadores de Lagrange, num problema equivalente³ sem restrições. Quando $|e(k)| \leq \bar{\gamma}$, cumpre tão somente minimizar (25), igualando-se seu gradiente ao vetor nulo. Tal procedimento dá origem a:

$$\mathbf{w}(k+1) = \Theta(\rho) \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{w}(k-l), \quad (26)$$

onde $\Theta(\rho) \triangleq (\rho - 1) / (\rho^{L\{e(k)\}} - 1)$. Observe que quando $L\{e(k)\} = 1$, (26) implica $\mathbf{w}(k+1) = \mathbf{w}(k)$, o que significa uma ausência de atualização. Dentro da filosofia de seleção de dados, descabe empreender atualizações desnecessárias (que implicam um aumento do custo computacional) quando o grau de novidade dos dados for considerado pouco relevante.

Quando $e(k) > \bar{\gamma}$, podemos converter - por meio da técnica dos multiplicadores de Lagrange - o problema (24) no problema de minimização da função $\mathcal{F}_{\text{U}}^+[\mathbf{w}(k+1)]$, sendo esta definida por

$$\mathcal{F}_{\text{U}}^+[\mathbf{w}(k+1)] \triangleq \mathcal{F}_{\text{RC-SMP}}[\mathbf{w}(k+1), e(k)] + \lambda [e_p(k) - \bar{\gamma}], \quad (27)$$

onde λ é o multiplicador de Lagrange. O gradiente de

$\mathcal{F}_{\text{U}}^+[\mathbf{w}(k+1)]$ com relação a $\mathbf{w}(k+1)$ é dado por

$$\begin{aligned} \frac{\partial \mathcal{F}_{\text{U}}^+[\mathbf{w}(k+1)]}{\partial \mathbf{w}(k+1)} = & 2 \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{\Lambda}^{-1}(k) \mathbf{w}(k+1) \quad (28) \\ & - 2 \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{\Lambda}^{-1}(k) \mathbf{w}(k-l) - \lambda \mathbf{x}(k) \end{aligned}$$

Igualando-se $\frac{\partial \mathcal{F}_{\text{U}}^+[\mathbf{w}(k+1)]}{\partial \mathbf{w}(k+1)}$ ao vetor nulo e isolando $\mathbf{w}(k+1)$, encontramos o seguinte formato da equação de atualização:

$$\mathbf{w}(k+1) = \Theta(\rho) \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{w}(k-l) + \frac{\lambda}{2} \Theta(\rho) \mathbf{\Lambda}(k) \mathbf{x}(k), \quad (29)$$

a qual exige a determinação de λ . Podemos calcular tal variável inserindo (29) na restrição $e_p(k) = \bar{\gamma}$, de modo a encontrar:

$$\lambda = \frac{2}{\Theta(\rho)} \frac{e'(k) - \bar{\gamma}}{\|\mathbf{x}(k)\|_{\mathbf{\Lambda}(k)}^2}. \quad (30)$$

Substituindo (30) em (29), encontramos a equação de atualização para o caso $e(k) > \bar{\gamma}$:

$$\mathbf{w}(k+1) = \Theta(\rho) \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{w}(k-l) + \frac{[e'(k) - \bar{\gamma}] \mathbf{\Lambda}(k) \mathbf{x}(k)}{\|\mathbf{x}(k)\|_{\mathbf{\Lambda}(k)}^2}. \quad (31)$$

Quando $e(k) < -\bar{\gamma}$, cumpre definir a função sem restrições $\mathcal{F}_{\text{U}}^-[\mathbf{w}(k+1)]$ a ser minimizada:

$$\mathcal{F}_{\text{U}}^-[\mathbf{w}(k+1)] \triangleq \mathcal{F}_{\text{RC-SMP}}[\mathbf{w}(k+1), e(k)] + \lambda [e_p(k) + \bar{\gamma}]. \quad (32)$$

Aplicando em (32) os mesmos passos das Eqs. (28), (29), (30) e (31), temos que a equação de atualização para o caso $e(k) < -\bar{\gamma}$ pode ser descrita por:

$$\mathbf{w}(k+1) = \Theta(\rho) \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{w}(k-l) + \frac{[e'(k) + \bar{\gamma}] \mathbf{\Lambda}(k) \mathbf{x}(k)}{\|\mathbf{x}(k)\|_{\mathbf{\Lambda}(k)}^2} \quad (33)$$

Podemos unificar as Eqs. (31) e (33) numa única, para o caso $|e(k)| > \bar{\gamma}$:

$$\begin{aligned} \mathbf{w}(k+1) = & \Theta(\rho) \sum_{l=0}^{L\{e(k)\}-1} \rho^l \mathbf{w}(k-l) \quad (34) \\ & + \beta \frac{\left[\frac{e'(k)}{e(k)} - \frac{\bar{\gamma}}{|e(k)|} \right] e(k) \mathbf{\Lambda}(k) \mathbf{x}(k)}{\|\mathbf{x}(k)\|_{\mathbf{\Lambda}(k)}^2 + \delta}, \end{aligned}$$

onde um parâmetro de relaxação $\beta \in (0, 1]$ foi inserido de sorte a permitir um compromisso entre taxa de convergência e desempenho em regime permanente [17] e δ é um parâmetro de regularização capaz de evitar instabilidades numéricas na ocorrência da divisão por um número muito pequeno [5].

Ao adotar um fator de reúso de coeficientes igual a L no algoritmo proposto, é interessante que este não atualize quando a condição $|e(k)| < \bar{\gamma}$ ocorrer, de modo a tornar efetiva uma das principais vantagens da família SM, a redução de custo computacional. Isto pode ser facilmente implementado por

³Isto é, que preserva a solução do problema original.

meio de um fator de reuso dependente de $e(k)$:

$$L\{e(k)\} = \begin{cases} 1, & \text{para } |e(k)| < \bar{\gamma} \\ L, & \text{caso contrário,} \end{cases} \quad (35)$$

de modo que a atualização do algoritmo proposto RC-SM-IMPNLMS pode ser descrita como:

$$\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k), & \text{caso } |e(k)| \leq \bar{\gamma} \\ \text{calculado segundo (34),} & \text{caso } |e(k)| > \bar{\gamma} \end{cases} \quad (36)$$

V. RESULTADOS

Os algoritmos utilizados para a comparação com o algoritmo proposto foram o NLMS, o RC-NLMS, o IMPNLMS, o SM-NLMS e o ℓ_0 -NLMS, com parâmetros $L = 7$, $\rho = 0,9$, $\epsilon = 10^{-3}$, $\lambda' = 0,96$, $\kappa = 10^{-8}$ e $\tau = 1$. Foi adicionado um ruído branco e gaussiano de variância σ_v^2 ao sinal de referência, sendo todos os resultados médios oriundos de 1000 ensaios independentes de Monte Carlo. A função de transferência ideal apresenta 50 coeficientes (dos quais os 10 primeiros consistem de uns, sendo nulos os demais). A Figura 3 apresenta o MSE em regime permanente em função de β para $\sigma_v^2 = 10^{-2}$, demonstrando que o algoritmo proposto apresenta desempenho em regime permanente significativamente superior à maioria dos algoritmos concorrentes.

Já a Figura 4 apresenta a evolução do MSD dos diversos algoritmos (com os mesmos parâmetros empregados na geração da Figura 3) para uma taxa de convergência semelhante e $\sigma_v^2 = 10^{-1}$. O algoritmo proposto apresenta melhor desempenho que os demais em regime permanente, onde ele apresenta uma probabilidade de atualização de aproximadamente 32%, o que torna seu custo computacional significativamente mais baixo do que o do RC-NLMS, algoritmo cujo desempenho mais se aproxima do algoritmo proposto.

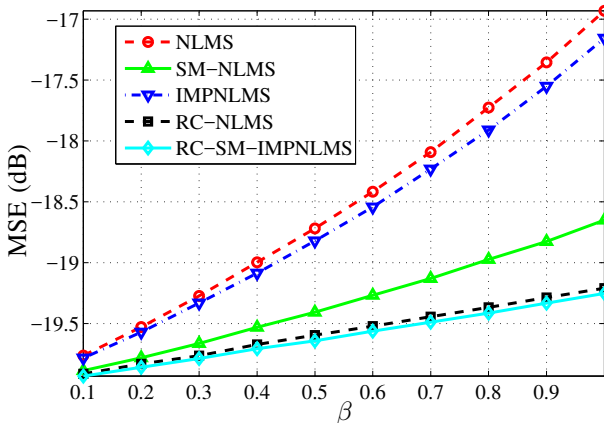


Fig. 3. MSE em regime permanente em função de β .

VI. CONCLUSÕES

Neste artigo, um novo algoritmo (RC-SM-IMPNLMS) foi proposto. O algoritmo conjuga diversas propriedades desejáveis, como uma alta taxa de convergência na identificação de sistemas esparsos, um bom desempenho em regime permanente e um custo computacional inferior, já que nem sempre requer o cálculo da atualização dos coeficientes adaptativos.

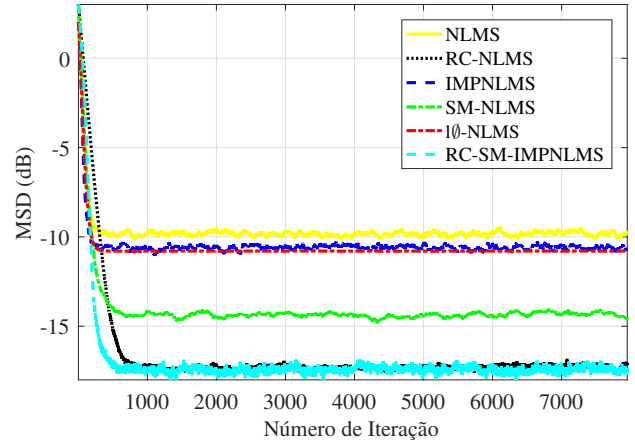


Fig. 4. Evolução do MSD (em dB) ao longo das iterações.

AGRADECIMENTOS

Os autores agradecem ao apoio dos órgãos de fomento (CNPq, CAPES e FAPERJ).

REFERÊNCIAS

- [1] H. Simon, "Adaptive filter theory," *Prentice Hall*, vol. 2, 2002.
- [2] A. Antoniou, *Digital signal processing*. McGraw-Hill Toronto, Canada:, 2006.
- [3] B. Widrow and M. E. Hoff, "Adaptive switching circuits," in *Neurocomputing: foundations of research*, pp. 123–134, MIT Press, 1988.
- [4] B. Widrow and M. A. Lehr, "30 years of adaptive neural networks: perceptron, madaline, and backpropagation," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1415–1442, 1990.
- [5] A. H. Sayed, *Fundamentals of adaptive filtering*. John Wiley & Sons, 2003.
- [6] H. Cho, C. W. Lee, and S. W. Kim, "Derivation of a new normalized least mean squares algorithm with modified minimization criterion," *Signal Processing*, vol. 89, no. 4, pp. 692–695, 2009.
- [7] S.-E. Kim, J.-W. Lee, and W.-J. Song, "Steady-state analysis of the nlms algorithm with reusing coefficient vector and a method for improving its performance," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 4120–4123, IEEE, 2011.
- [8] S.-E. Kim, J.-W. Lee, and W.-J. Song, "A noise-resilient affine projection algorithm and its convergence analysis," *Signal Processing*, vol. 121, pp. 94–101, 2016.
- [9] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan (Part I: Communications)*, vol. 67, no. 5, pp. 19–27, 1984.
- [10] P. S. Diniz and S. Werner, "Set-membership binormalized data-reusing lms algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 124–134, 2003.
- [11] S. Werner and P. S. Diniz, "Set-membership affine projection algorithm," *IEEE Signal Processing Letters*, vol. 8, no. 8, pp. 231–235, 2001.
- [12] D. L. Huttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 5, pp. 508–518, 2000.
- [13] H. Deng and M. Doroslovacki, "Improving convergence of the pnllms algorithm for sparse impulse response identification," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 181–184, 2005.
- [14] J. Benesty and S. L. Gay, "An improved pnllms algorithm," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, vol. 2, pp. II-1881, IEEE, 2002.
- [15] M. Fukumoto, S. Saiki, et al., "An improved mu-law proportionate nlms algorithm," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 3797–3800, IEEE, 2008.
- [16] 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, 2014.
- [17] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. New York, USA: Springer, 4th ed., 2013.