

Um Novo Algoritmo Adaptativo Baseado em Projeções com Restrição de Norma ℓ_0

Eduardo Beck, Eduardo Luiz Ortiz Batista e Rui Seara

Resumo— Este artigo apresenta uma nova abordagem para a derivação de algoritmos adaptativos com restrição de norma para a identificação de sistemas esparsos. A abordagem proposta é focada principalmente em manter o erro *a posteriori* igual a zero (como no caso do algoritmo NLMS), enquanto busca satisfazer, simultaneamente, uma restrição de norma apropriada. A partir de tal abordagem, um algoritmo baseado na norma ℓ_0 é desenvolvido. Uma versão desse algoritmo não requerendo o ajuste do parâmetro de norma é também derivada. Resultados de simulação são apresentados visando confirmar a eficácia do algoritmo proposto.

Palavras-Chave— Filtragem adaptativa, identificação de sistemas, método de projeções, otimização com restrição de norma, resposta ao impulso esparsa.

Abstract— This paper introduces a novel approach to derive norm-constrained adaptive algorithms for sparse system identification. The proposed approach is focused primarily on keeping the *a posteriori* error equal to zero (as in the case of the NLMS algorithm), while seeking to satisfy a norm constraint simultaneously. By using such an approach, an algorithm based on the ℓ_0 norm is developed. A version of this algorithm not requiring the adjustment of the norm parameter is also derived. Simulation results are shown aiming to confirm the effectiveness of the proposed algorithm.

Keywords— Adaptive filtering, system identification, projection algorithms, norm-constrained optimization, sparse impulse response.

I. INTRODUÇÃO

Muitas áreas de aplicação de processamento de sinais, tais como comunicações, acústica, antenas, sismologia, envolvem sistemas que são inerentemente esparsos [1]-[4]. Esses sistemas são caracterizados por respostas ao impulso que possuem apenas alguns coeficientes diferentes de zero, enquanto a maioria deles são iguais a zero ou têm magnitude desprezível [2]. Frequentemente, a identificação em tempo real desse tipo de resposta é requerida, a qual pode ser feita usando filtros adaptativos. Nesse contexto, uma área de pesquisa que tem atraído atenção nas últimas décadas é a de algoritmos de filtragem adaptativa que exploram a esparsidade de respostas ao impulso visando superar em desempenho os algoritmos adaptativos clássicos.

A maioria dos algoritmos adaptativos concebidos para identificação de sistemas esparsos é do tipo proporcional [3]. Nesses algoritmos, ganhos de adaptação distintos são atribuídos para cada coeficiente do filtro com o objetivo de acelerar

a convergência a partir da alocação de ganhos maiores aos coeficientes mais significativos. Um dos primeiros algoritmos a utilizar tal estratégia é o algoritmo NLMS proporcional (*proportionate normalized least-mean-square* - PNLMS) [5], que vem servindo de inspiração para o desenvolvimento de uma série de outros algoritmos do tipo proporcional com melhor desempenho [6]-[12]. Mais recentemente, uma nova classe de algoritmos adaptativos, baseada no uso de normas vetoriais, vem ganhando importância [13]-[15]. A ideia atrás desses algoritmos é considerar alguma função de norma relacionada ao grau de esparsidade de um vetor (por exemplo, normas ℓ_0 e ℓ_1) [2], com vistas a impor restrições ao processo de atualização dos coeficientes e buscar soluções esparsas. Em um primeiro grupo de algoritmos baseados em normas, a norma vetorial é incluída diretamente na função custo a ser minimizada com o objetivo de “penalizar” o processo adaptativo. Exemplos de algoritmos desse grupo são o ZA-LMS (*zero-attracting least-mean-square*) [13], o RZA-LMS (*reweighted ZA-LMS*) [13], o LMS com restrição de norma ℓ_0 (ℓ_0 -LMS) [14] e suas versões normalizadas (ZA-NLMS, RZA-NLMS e ℓ_0 -NLMS) [13], [14]. Um segundo grupo é composto por algoritmos baseados na teoria dos conjuntos [16] juntamente com o método de projeções descrito em [17]. Ao utilizar esta última abordagem, o vetor de coeficientes é projetado (em cada iteração) primeiramente visando obter um erro *a posteriori* próximo a zero e, subsequentemente, em uma superfície de soluções que possuam valores de norma limitados. Os principais algoritmos desse grupo são a *adaptive projection-based algorithm using ℓ_1 balls* (APL1) [15] e a sua versão baseada em projeções de norma ℓ_1 ponderada (*adaptive projection-based algorithm using weighted ℓ_1 balls* - APWL1) [15]. Outros algoritmos, como os apresentados em [18] e [19], foram desenvolvidos combinando tanto estratégias utilizadas nos algoritmos proporcionais quanto nos algoritmos baseados em normas.

O foco deste trabalho é o desenvolvimento de algoritmos adaptativos para identificação de sistemas esparsos que possuam melhor desempenho, baixa complexidade computacional e reduzido número de parâmetros definidos pelo usuário. De forma semelhante aos algoritmos APL1 e APWL1, a proposta aqui apresentada é também fundamentada na teoria dos conjuntos e utiliza um método de projeções para atualizar o vetor de coeficientes. No entanto, a restrição de erro *a posteriori* igual a zero (característica do algoritmo NLMS - *normalized LMS*) e a restrição que limita a solução a um determinado valor de norma são aplicadas simultaneamente ao processo de otimização, buscando uma solução que atenda a ambas as restrições a cada iteração. Tal abordagem difere da abordagem apresentada em [15], na qual as restrições são

Eduardo Beck, Eduardo Luiz Ortiz Batista e Rui Seara, LINSE - Laboratório de Circuitos e Processamento de Sinais, Departamento de Engenharia Elétrica e Eletrônica, Universidade Federal de Santa Catarina, Florianópolis-SC, Brasil, E-mails: ebeck@linse.ufsc.br, ebatista@linse.ufsc.br, seara@linse.ufsc.br.

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

tomadas separadamente, realizando as projeções de forma sequencial. A partir da abordagem proposta, um novo algoritmo com restrição de norma ℓ_0 é desenvolvido, como também uma versão que não requer o ajuste do parâmetro de norma.

II. DEFINIÇÕES INICIAIS

Filtros e algoritmos adaptativos são geralmente desenvolvidos considerando uma estrutura de filtragem FIR cuja saída na iteração n é dada por $y(n) = \mathbf{w}^T(n) \mathbf{x}(n)$ com o vetor de coeficientes definido como $\mathbf{w}(n) = [w_0(n) \ w_1(n) \ \cdots \ w_{N-1}(n)]^T$ e o vetor de entrada como $\mathbf{x}(n) = [x(n) \ x(n-1) \ \cdots \ x(n-N+1)]^T$. Considerando tal estrutura de filtragem, o objetivo de um algoritmo adaptativo é minimizar uma função custo baseada no sinal de erro, o qual é definido como

$$e(n) = d(n) - \mathbf{w}^T(n) \mathbf{x}(n) \quad (1)$$

com $d(n)$ caracterizando o sinal desejado [20]. Em identificação de sistemas, $d(n)$ é tipicamente o sinal de saída da planta (sistema a ser modelado) somado a um ruído de medição. Um dos principais algoritmos adaptativos utilizados nesse tipo de aplicação é o NLMS [21], o qual tem por objetivo encontrar, a cada iteração, o vetor de coeficientes *a posteriori* $\mathbf{w}(n+1)$ mais próximo (em termos de distância euclidiana) do vetor de coeficientes atual $\mathbf{w}(n)$ que torna o erro *a posteriori*

$$\varepsilon(n) = d(n) - \mathbf{w}^T(n+1) \mathbf{x}(n) \quad (2)$$

igual a zero. Tal objetivo é buscado realizando uma projeção ortogonal do vetor $\mathbf{w}(n)$ em um conjunto de soluções para as quais $\varepsilon(n) = 0$, sendo esse conjunto caracterizado pelo hiperplano $S_\varepsilon(n) = \{\mathbf{w}(n+1) \in \mathbb{R}^N : \varepsilon(n) = 0\}$. Como resultado, a seguinte equação de atualização é obtida [20]:

$$\mathbf{w}_{\text{NLMS}}(n+1) = \mathbf{w}(n) + \mu \frac{e(n)}{\mathbf{x}^T(n) \mathbf{x}(n)} \mathbf{x}(n) \quad (3)$$

onde μ é um parâmetro de controle de passo.

Por outro lado, o algoritmo APL1 [15], o qual é dedicado à identificação de sistemas esparsos, acrescenta ao processo adaptativo uma segunda projeção do vetor de coeficientes em uma *norm-ball*, em sequência a projeções que buscam obter erro *a posteriori* igual a zero. A *norm-ball* é definida como

$$S_{\ell_1}(n) = \{\mathbf{w}(n+1) \in \mathbb{R}^N : \|\mathbf{w}(n+1)\| \leq \delta_{\ell_1}\} \quad (4)$$

com $\|\mathbf{w}(n+1)\|$ denotando a norma ℓ_1 do vetor $\mathbf{w}(n+1)$ e δ_{ℓ_1} , o limite de norma atribuído à restrição. Além disso, o algoritmo APWL1 [15] considera uma versão ponderada da norma ℓ_1 para definição da *norm-ball* considerada.

III. ALGORITMOS ADAPTATIVOS PROPOSTOS

Nesta seção, a abordagem proposta para o desenvolvimento de algoritmos adaptativos é primeiramente introduzida, seguida pela descrição da aproximação da norma ℓ_0 considerada e do desenvolvimento de um algoritmo baseado em tal aproximação. Além disso, um algoritmo com parâmetro de restrição de norma variável é também obtido.

A. Abordagem Proposta

A abordagem para o desenvolvimento de algoritmos adaptativos proposta neste trabalho baseia-se na busca (a cada iteração) por um vetor de coeficientes que mantém o erro *a posteriori* igual a zero (como no algoritmo NLMS) enquanto procura satisfazer uma restrição de norma (como nos algoritmos APL1 e APWL1). A norma escolhida aqui é a ℓ_0 , que corresponde ao número de elementos de um vetor que são diferentes de zero [22]. Portanto, a norma ℓ_0 é de fato uma métrica intimamente relacionada à definição estrita de esparsidade, o que motiva seu uso no desenvolvimento de algoritmos para identificação de sistemas esparsos. Assim, a ideia atrás da abordagem proposta é realizar, a cada iteração do algoritmo, uma única projeção ortogonal do vetor de coeficientes do filtro sobre a interseção entre um hiperplano $S_\varepsilon(n)$ de erro *a posteriori* igual a zero e outro conjunto S_{ℓ_0} de soluções possuindo valores determinados de norma ℓ_0 . Para tal, uma projeção euclidiana [23] necessita ser realizada através da solução do seguinte problema de otimização:

$$\begin{aligned} & \text{minimizar } 1/2 \|\mathbf{w}(n+1) - \mathbf{w}(n)\|_2^2 \\ & \text{sujeito a } \begin{cases} \varepsilon(n) = d(n) - \mathbf{w}^T(n+1) \mathbf{x}(n) = 0 \\ \|\mathbf{w}(n+1)\|_0 = \delta_{\ell_0} \end{cases} \quad (5) \end{aligned}$$

com δ_{ℓ_0} representando o parâmetro de restrição de norma. A obtenção de uma solução para (5) não é uma tarefa fácil, especialmente devido ao fato de que a norma ℓ_0 é uma função não escalável, não convexa e não diferenciável. De fato, essa função não pode sequer ser considerada uma norma vetorial no sentido estrito, visto que ela não atende a todos os requisitos necessários para tal. Para contornar os problemas gerados pelo uso da norma ℓ_0 , a estratégia adotada aqui é a de utilizar, ao invés da definição clássica de norma ℓ_0 , uma versão aproximada que seja uma função afim [23]. Com isso, o problema de otimização de (5) pode ser transformado em um problema convexo, facilitando a sua solução.

B. Aproximação para Norma ℓ_0

Visando obter uma aproximação afim para a norma ℓ_0 , primeiramente a função de aproximação descrita em [14] é considerada. Tal função é definida como

$$\|\mathbf{w}(n+1)\|_{0_L} = \sum_{k=0}^{N-1} 1 - z_k(n+1) \quad (6)$$

com

$$z_k(n) = \begin{cases} 1 - \beta |w_k(n)|, & |w_k(n)| \leq 1/\beta \\ 0, & |w_k(n)| > 1/\beta \end{cases} \quad (7)$$

onde β é um parâmetro de aproximação (se $\beta \rightarrow \infty$, $\|\mathbf{w}(n+1)\|_{0_L} \rightarrow \|\mathbf{w}(n+1)\|_0$). A expressão (6) pode ainda ser reescrita como

$$\|\mathbf{w}(n+1)\|_{0_L} = N - \mathbf{s}_z^T(n+1) \mathbf{s}_z(n+1) + \beta \mathbf{s}_z^T(n+1) \mathbf{w}(n+1) \quad (8)$$

onde $\mathbf{s}_z(n) = [s_{z0}(n) \ s_{z1}(n) \ \dots \ s_{z(N-1)}(n)]^T$ é um vetor cujos elementos são dados por

$$s_{zk}(n) = \begin{cases} \text{sgn}[w_k(n)], & |w_k(n)| \leq 1/\beta \\ 0, & |w_k(n)| > 1/\beta. \end{cases} \quad (9)$$

Note que (8) ainda não é uma função afim. Então, uma segunda aproximação é adotada considerando uma condição de adaptação lenta: $\mathbf{s}_z(n+1) \cong \mathbf{s}_z(n)$. Assim, é possível obter a seguinte aproximação afim para a norma ℓ_0 de $\mathbf{w}(n+1)$:

$$f_{\ell_0}[\mathbf{w}(n+1)] = N - \mathbf{s}_z^T(n) \mathbf{s}_z(n) + \beta \mathbf{s}_z^T(n) \mathbf{w}(n+1). \quad (10)$$

C. Algoritmo Baseado na Norma ℓ_0

Considerando (10), torna-se possível definir uma versão convexa do problema de otimização descrito em (5). Assim,

$$\text{minimizar } 1/2 \|\mathbf{w}(n+1) - \mathbf{w}(n)\|_2^2 \quad (11)$$

$$\text{sujeito a } \begin{cases} \varepsilon(n) = d(n) - \mathbf{w}^T(n+1) \mathbf{x}(n) = 0 \\ f_{\ell_0}[\mathbf{w}(n+1)] = \delta_{\ell_0}. \end{cases}$$

Agora, utilizando o método dos multiplicadores de Lagrange [23], [24], o problema de (11) pode ser reescrito como um problema sem restrições com um lagrangiano dado por

$$\begin{aligned} \mathcal{L}[\mathbf{w}(n+1)] = & \frac{1}{2} [\mathbf{w}(n+1) - \mathbf{w}(n)]^T [\mathbf{w}(n+1) - \mathbf{w}(n)] \\ & + \lambda_1(n) [d(n) - \mathbf{w}^T(n+1) \mathbf{x}(n)] \\ & + \lambda_{\ell_p}(n) \{f_{\ell_p}[\mathbf{w}(n+1)] - \delta_{\ell_p}\} \end{aligned} \quad (12)$$

onde $\lambda_1(n)$ e $\lambda_{\ell_p}(n)$ são os multiplicadores de Lagrange correspondentes à primeira e à segunda restrições de (11), respectivamente. Então, fazendo o gradiente de (12) com respeito a $\mathbf{w}(n+1)$ igual a zero, obtém-se

$$\begin{aligned} \nabla \mathcal{L}[\mathbf{w}(n+1)] = & \mathbf{w}(n+1) - \mathbf{w}(n) \\ & - \lambda_1(n) \mathbf{x}(n) + \lambda_{\ell_0}(n) \mathbf{g}_{\ell_0}(n+1) = \mathbf{0} \end{aligned} \quad (13)$$

onde $\mathbf{g}_{\ell_0}(n+1) = \nabla f_{\ell_0}[\mathbf{w}(n+1)]$ é o gradiente de (10) com respeito a $\mathbf{w}(n+1)$. Resolvendo (13) para $\mathbf{w}(n+1)$, é possível obter

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \lambda_1(n) \mathbf{x}(n) - \lambda_{\ell_0}(n) \mathbf{g}_{\ell_0}(n+1). \quad (14)$$

O multiplicador de Lagrange $\lambda_1(n)$ é determinado substituindo (14) na primeira restrição de (11), resultando em

$$\begin{aligned} d(n) - \mathbf{w}^T(n) \mathbf{x}(n) - \lambda_1(n) \mathbf{x}^T(n) \mathbf{x}(n) \\ + \lambda_{\ell_0}(n) \mathbf{g}_{\ell_0}^T(n+1) \mathbf{x}(n) = 0. \end{aligned} \quad (15)$$

Dessa forma, considerando (1) e resolvendo (15) para $\lambda_1(n)$, obtém-se

$$\lambda_1(n) = \alpha_1(n) + \alpha_{\ell_0}(n) \lambda_{\ell_0}(n) \quad (16)$$

com

$$\alpha_1(n) = e(n) / [\mathbf{x}^T(n) \mathbf{x}(n)] \quad (17)$$

e

$$\alpha_{\ell_0}(n) = \mathbf{g}_{\ell_0}^T(n+1) \mathbf{x}(n) / [\mathbf{x}^T(n) \mathbf{x}(n)]. \quad (18)$$

Agora, para calcular o multiplicador $\lambda_{\ell_0}(n)$, é primeiramente requerido calcular o gradiente de (10) com respeito a $\mathbf{w}(n+1)$,

o que resulta em

$$\mathbf{g}_{\ell_0}(n+1) = \beta \mathbf{s}_z(n). \quad (19)$$

Então, considerando (19) e (14) e incluindo um parâmetro μ para controlar a adaptação, a seguinte equação de atualização é obtida:

$$\begin{aligned} \mathbf{w}(n+1) = & \mathbf{w}(n) + \mu \{ [\alpha_1(n) + \alpha_{\ell_0}(n) \lambda_{\ell_0}(n)] \mathbf{x}(n) \\ & - \beta \lambda_{\ell_0}(n) \mathbf{s}_z(n) \} \end{aligned} \quad (20)$$

com $\alpha_1(n)$ dado por (17) e, a partir de (18) e (19),

$$\alpha_{\ell_0}(n) = \beta \mathbf{s}_z^T(n) \mathbf{x}(n) / [\mathbf{x}^T(n) \mathbf{x}(n)]. \quad (21)$$

Para determinar $\lambda_{\ell_0}(n)$, (20) é substituído em (10) e, fazendo $f_{\ell_0}[\mathbf{w}(n+1)] = \delta_{\ell_0}$, é possível obter

$$\begin{aligned} \lambda_{\ell_0}(n) = & \frac{1}{\beta} \{ \mathbf{s}_z^T(n) \mathbf{w}(n) + \alpha_1(n) \mathbf{s}_z^T(n) \mathbf{x}(n) \\ & + [N - \mathbf{s}_z^T(n) \mathbf{s}_z(n) - \delta_{\ell_0}] / \beta \} \\ & \times \frac{1}{\mathbf{s}_z^T(n) \mathbf{s}_z(n) - \alpha_{\ell_0}(n) \mathbf{s}_z^T(n) \mathbf{x}(n) / \beta}. \end{aligned} \quad (22)$$

Em resumo, o algoritmo baseado na norma ℓ_0 aqui obtido consiste na adaptação dos coeficientes do filtro usando (20) com o auxílio de (17), (21) e (22).

D. Algoritmo com Restrição Variável de Norma ℓ_0

Assim como no caso dos algoritmos APL1 e APWL1, o algoritmo de norma ℓ_0 obtido na seção anterior depende da escolha apropriada de um parâmetro de restrição de norma. A escolha de tal parâmetro, por sua vez, depende do conhecimento das características da planta a ser identificada. No entanto, essa não é uma característica desejável para um algoritmo adaptativo, uma vez que as informações da planta são usualmente desconhecidas *a priori*. Assim, o desenvolvimento de uma versão do algoritmo derivado na seção anterior com um parâmetro de norma variável que se ajuste automaticamente às condições do problema é de grande interesse. Para tal, o seguinte parâmetro variável de restrição de norma é inicialmente definido:

$$\delta_{v\ell_0}(n) = N - \mathbf{s}_z^T(n) \mathbf{s}_z(n) = N - N_z(n) \quad (23)$$

onde $N_z(n)$ representa o número de componentes com valores absolutos menores do que o limiar $1/\beta$ (componentes não significativos) no instante n . Utilizando tal parâmetro, verifica-se que o termo $\beta \mathbf{s}_z^T(n) \mathbf{w}(n+1)$ [em (8)] tende para zero, forçando os componentes não significativos do vetor de coeficientes para zero. Finalmente, substituindo, em (22), δ_{ℓ_0} por (23), considerando a expressão resultante juntamente com (20) e (21), e incluindo também o parâmetro de controle de passo μ no processo, obtém-se a seguinte regra de adaptação para o algoritmo com restrição variável de norma ℓ_0 :

$$\begin{aligned} \mathbf{w}(n+1) = & \mathbf{w}(n) + \mu \{ [\alpha_1(n) + \alpha_{v\ell_0}(n) \lambda_{v\ell_0}(n)] \mathbf{x}(n) \\ & - \lambda_{v\ell_0}(n) \mathbf{s}_z(n) \} \end{aligned} \quad (24)$$

com $\alpha_1(n)$ dado por (17),

$$\alpha_{v\ell_0}(n) = \mathbf{s}_z^T(n) \mathbf{x}(n) / [\mathbf{x}^T(n) \mathbf{x}(n)] \quad (25)$$

e

$$\lambda_{v\ell_0}(n) = \frac{\mathbf{s}_z^T(n) \mathbf{w}(n) + \alpha_1(n) \mathbf{s}_z^T(n) \mathbf{x}(n)}{\mathbf{s}_z^T(n) \mathbf{s}_z(n) - \alpha_{v\ell_0}(n) \mathbf{s}_z^T(n) \mathbf{x}(n)}. \quad (26)$$

É importante mencionar que o uso de $\delta_{v\ell_0}(n)$ no lugar de δ_{ℓ_0} não implica qualquer alteração sobre o efeito da primeira restrição de (11) no processo de adaptação e, conseqüentemente, a solução obtida com o algoritmo de passo variável também levará a $\varepsilon = 0$ para $\mu = 1$.

IV. RESULTADOS DE SIMULAÇÕES

Nesta seção, resultados de simulação são apresentados com o objetivo de avaliar o desempenho do algoritmo com restrição variável de norma ℓ_0 proposto em comparação com outros algoritmos adaptativos voltados para identificação de sistemas esparsos encontrados na literatura. Tais resultados são obtidos a partir de simulações de Monte Carlo (MC) (média de 100 realizações independentes) que consideram problemas de modelagem de plantas FIR esparsas. As plantas consideradas são as seguintes:

- i) Planta \mathbf{h}_1 caracterizada por um vetor de coeficientes de comprimento 100 com apenas 4 coeficientes diferentes de zero [11]. Os coeficientes diferentes de zero têm valores dados por $\{0, 1, 1, 0, -0, 5, 0, 1\}$ e estão localizados nas posições $\{1, 30, 35, 85\}$. O grau de esparsidade [2] dessa planta é $\mathcal{S}(\mathbf{h}_1) = 0,9435$.
- ii) Planta \mathbf{h}_2 (mais realista) dada por um vetor de coeficientes de comprimento 512, cuja parte inicial (atraso) tem 100 coeficientes iguais a zero, seguida por uma parte central composta por um caminho de eco padrão (modelo 4, 128 coeficientes) extraída da Recomendação ITU-T G.168 [25], e uma parte de cauda com 284 coeficientes também iguais a zero {grau de esparsidade [2] $\mathcal{S}(\mathbf{h}_2) = 0,7253$ }.

Um ruído de medição gaussiano branco $r(n)$ com variância $\sigma_r^2 = 10^{-3}$ é adicionado à saída das plantas consideradas. A avaliação do desempenho é realizada em função do desalinhamento normalizado [26].

A. Exemplo 1

O algoritmo de norma ℓ_0 com restrição variável proposto possui apenas o parâmetro β a mais do que o algoritmo NLMS. Visando avaliar a sensibilidade desse algoritmo ao parâmetro β , um conjunto de simulações é realizado considerando o algoritmo para a identificação da planta \mathbf{h}_1 com sinal de entrada gaussiano branco. Nesse conjunto de simulações, o valor de β é variado de 5 até 1000, sendo os resultados de desalinhamento normalizado em regime permanente mostrados na Fig. 1. Nota-se que três faixas de valores de β com características distintas de desempenho destacam-se nessa figura. A primeira faixa, indicada por a), mostra que o desempenho do algoritmo é consideravelmente degradado se β é muito pequeno. Tal comportamento é devido ao fato de o limite $1/\beta$ ser muito grande e, conseqüentemente, alguns dos coeficientes diferentes de zero do filtro são tratados como não significativos e forçados para zero. A segunda faixa, indicada por b), corresponde aos valores de $70 \leq \beta \leq 550$ para os quais um desempenho superior do algoritmo é obtido (fatores como variância do ruído, grau de correlação do sinal de entrada e

magnitude dos coeficientes da planta podem alterar os limites dessa faixa). Finalmente, no terceiro intervalo [indicado por c) na Fig. 1], o limite definido por $1/\beta$ é tão pequeno que todos os coeficientes do filtro são significativos, tornando seu desempenho similar ao do algoritmo NLMS. A partir desses resultados, pode-se constatar que a sensibilidade do algoritmo proposto com respeito a β não é crítica, uma vez que um desempenho satisfatório é obtido para uma ampla faixa de valores de β .

B. Exemplo 2

Neste exemplo, o algoritmo com restrição variável de norma ℓ_0 é avaliado no contexto da modelagem da planta \mathbf{h}_2 , comparando-o com outros algoritmos voltados à identificação de sistemas esparsos. São consideradas aqui duas situações distintas, sendo a primeira usando um sinal de entrada gaussiano branco, com variância unitária, e a segunda, um sinal de entrada gaussiano colorido obtido a partir de um processo AR(2) [27], dado por $x(n) = 0,8x(n-1) - 0,8x(n-2) + v(n)$, onde $v(n)$ é um ruído gaussiano branco com variância $\sigma_v^2 = 0,29$ (resultando em uma dispersão de autovalores do sinal de entrada de $\chi = 211$). Os parâmetros dos algoritmos são obtidos através de busca exaustiva, sendo ajustados para obter a melhor taxa de convergência para um desalinhamento normalizado em regime permanente semelhante ao do algoritmo com restrição variável de norma ℓ_0 proposto com $\mu = 1$. O melhor valor de β do algoritmo proposto para essa configuração é $\beta = 2300$, para o caso de sinal de entrada gaussiano branco, e $\beta = 3000$ para sinal de entrada gaussiano colorido. Os algoritmos utilizados na comparação representam alguns dos melhores exemplares de cada categoria citada na Seção I. Para o caso de sinal de entrada gaussiano branco, os algoritmos considerados são configurados da seguinte forma: $\mu = 0,3$ para o NLMS [21]; $\mu = 0,3$, $\varrho = 5/N$ e $\delta_p = 0,01$ para o PNLMS [5]; $\mu = 0,3$, $\varrho = 5/N$ e $\delta_p = 0,01$ para o SPNLMS [8], [9]; $\mu = 0,55$, $\kappa = 10^{-8}$ e $\beta = 100$ para o ℓ_0 -NLMS [14]; e $\mu = 0,8$, $Q = 1$, $\phi = 0$ e $\delta_{W\ell_1} = 256$ para o APWL1 [15]. Para o caso de sinal de entrada gaussiano colorido, os parâmetros são: $\mu = 0,45$ para o NLMS [21]; $\mu = 0,38$, $\varrho = 5/N$ e $\delta_p = 0,01$ para o PNLMS [5]; $\mu = 0,27$, $\varrho = 5/N$ e $\delta_p = 0,01$ para o SPNLMS [8], [9]; $\mu = 0,55$, $\kappa = 10^{-9}$ e $\beta = 100$ para o ℓ_0 -NLMS [14]; e $\mu = 0,6$, $Q = 1$, $\phi = 0$ e $\delta_{W\ell_1} = 400$ para o APWL1 [15]. Os resultados obtidos

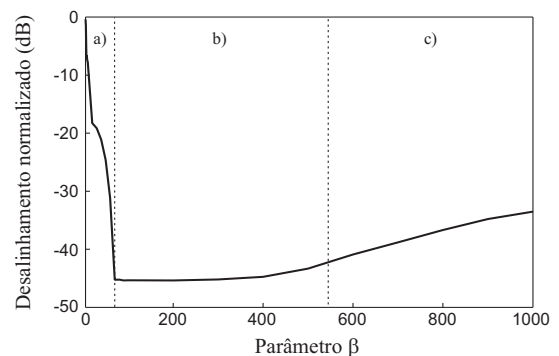


Fig. 1. Exemplo 1. Desempenho em regime permanente do algoritmo $V\text{-}\ell_0\text{NC}$ para diferentes valores de β .

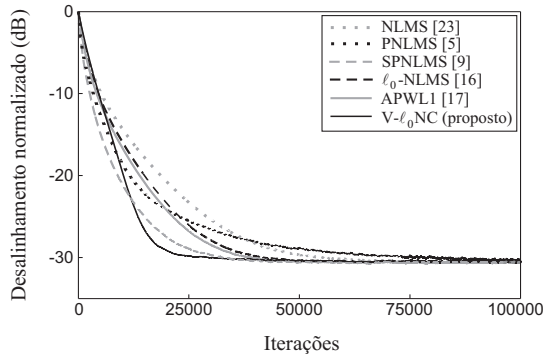


Fig. 2. Example 2. Comparação de desempenho de diferentes algoritmos para sinal de entrada gaussiano branco.

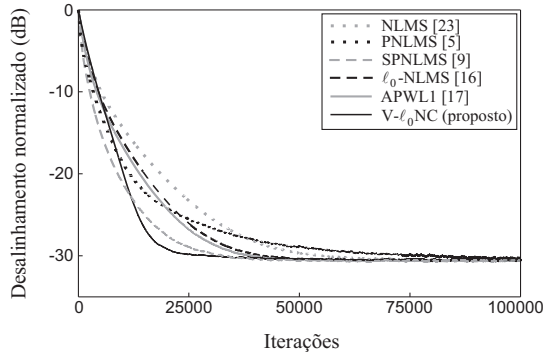


Fig. 3. Example 2. Comparação de desempenho de diferentes algoritmos para sinal de entrada gaussiano colorido.

para sinal de entrada gaussiano branco são apresentados na Fig. 2, enquanto os obtidos usando o sinal de entrada gaussiano colorido são mostrados na Fig. 3. Observa-se que, na simulação com sinal de entrada gaussiano branco, o algoritmo com melhor desempenho é o SPNLMS, seguido muito de perto pelo algoritmo com restrição variável de norma ℓ_0 proposto. Por outro lado, para o sinal de entrada gaussiano colorido, o melhor desempenho é obtido pelo algoritmo proposto seguido pelo SPNLMS.

V. CONCLUSÕES

Neste trabalho, uma nova abordagem para o desenvolvimento de algoritmos adaptativos para identificação de sistemas esparsos foi apresentada. Como na abordagem utilizada para a obtenção dos algoritmos APL1 e APWL1 [15], a abordagem aqui proposta baseia-se na busca de um vetor de coeficientes que atenda uma restrição de norma vetorial. A principal diferença da abordagem proposta com aquela usada em [15] para obtenção dos algoritmos APL1 e APWL1 diz respeito à estratégia utilizada para a busca do vetor de coeficientes, que é baseada na projeção direta na intersecção de um hiperplano de soluções de erro *a posteriori* igual a zero e um hiperplano de soluções que satisfaz uma restrição relaxada de norma. Os algoritmos aqui obtidos apresentam custo computacional e número de parâmetros a serem definidos pelo usuário que se aproximam aos do algoritmo NLMS, enquanto mantêm as propriedades de promoção de esparsidade utilizando norma vetorial, características dos algoritmos APL1 e APWL1. Resultados de simulação numérica confirmaram a eficácia da formulação proposta como também dos algoritmos propostos.

REFERÊNCIAS

- [1] P. Naylor, J. Cui, and M. Brookes, "Adaptive algorithms for sparse echo cancellation," *Signal Process.*, vol. 86, no. 6, pp. 1182–1192, Jun. 2006.
- [2] N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4723–4741, Oct. 2009.
- [3] C. Paleologu, J. Benesty, and S. Ciochina, *Sparse Adaptive Filters for Echo Cancellation*. San Rafael, CA: Morgan & Claypool, 2010.
- [4] E. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Analysis Applic.*, vol. 14, no. 5-6, pp. 877–905, Dec. 2008.
- [5] D. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sep. 2000.
- [6] J. Benesty and S. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Orlando, FL, May 2002, vol. 2, pp. 1881–1884.
- [7] J. Cui, P. Naylor, and D. Brown, "An improved IPNLMS algorithm for echo cancellation in packet-switched networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Montreal, Canada, May 2004, vol. 4, pp. 141–144.
- [8] M. Doroslovacki, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 181–184, Mar. 2005.
- [9] H. Deng and M. Doroslovacki, "Proportionate adaptive algorithms for network echo cancellation," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1794–1803, May 2006.
- [10] P. Loganathan, A. Khong, and P. Naylor, "A Class of sparseness-controlled algorithms for echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 17, no. 8, pp. 1591–1601, Nov. 2009.
- [11] F. C. Souza, O. J. Tobias, R. Seara, and D. R. Morgan, "A PNLMS algorithm with individual activation factors," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2036–2047, Apr. 2010.
- [12] F. C. Souza, R. Seara, and D. R. Morgan, "An enhanced IAF-PNLMS adaptive algorithm for sparse impulse response identification," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3301–3307, Jun. 2012.
- [13] Y. Chen, Y. Gu, and A. Hero, "Sparse LMS for system identification," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 3125–3128.
- [14] Y. Gu, J. Jin, and S. Mei, " ℓ_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.
- [15] Y. Kopsinis, K. Slavakis, and S. Theodoridis, "Online sparse system identification and signal reconstruction using projections onto weighted ℓ_1 balls," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 936–952, Mar. 2011.
- [16] P. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, no. 2, pp. 182–208, Feb. 1993.
- [17] S. Theodoridis, K. Slavakis, and I. Yamada, "Adaptive learning in a world of projections," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 97–123, Jan. 2011.
- [18] C. Paleologu, J. Benesty, and S. Ciochina, "An improved proportionate NLMS algorithm based on the ℓ_0 norm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, TX, Mar. 2010, pp. 309–312.
- [19] S. Chouvardas, K. Slavakis, Y. Kopsinis, and S. Theodoridis, "A sparsity promoting adaptive algorithm for distributed learning," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5412–5425, Oct. 2012.
- [20] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [21] J. Nagumo and A. Noda, "A learning method for system identification," *IEEE Trans. Autom. Control*, vol. 4, no. 3, pp. 282–287, Jun. 1967.
- [22] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ Pr, 2004.
- [24] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed. New York, NY: Springer Science, 2006.
- [25] ITU-T, *Recommendation G. 711, General Aspects of Digital Transmission Systems*. Geneva, Switzerland: Int. Telecomm. Union, 1993.
- [26] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO Signal Processing*. Berlin, Germany: Springer, 2006.
- [27] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice-Hall, 1993.