

Filtragem Adaptativa em Subbandas para Separação Supervisionada de Fontes

Bruna Luisa Ramos Prado Vasques, Jéssica Bartholdy Sanson, Mariane Rembold Petraglia

Resumo—Este artigo investiga técnicas de filtragem adaptativa MIMO (*multiple-input multiple-output*) para a aplicação de separação supervisionada de fontes no contexto de misturas convolutivas. Partindo da observação de que existe correlação entre os sinais das misturas, é proposta uma modificação no algoritmo NSAF (*Normalized Subband Adaptive Filter*) com o objetivo de acelerar sua velocidade de convergência. Resultados de simulações com misturas de sinais de voz em ambientes reverberantes mostram o desempenho superior do algoritmo proposto em relação aos desempenhos dos algoritmos NLMS (*Normalized Least-Mean-Square*) e NSAF convencionais considerando tanto a velocidade de convergência quanto a SIR (*Signal-to-Interference Ratio*) após a convergência.

Palavras-Chave—Filtragem adaptativa, Filtro em subbandas adaptativo normalizado, Separação de fontes.

Abstract—This paper investigates adaptive filtering MIMO (*multiple-input multiple-output*) techniques for the application of supervised source separation in the context of convolutive mixtures. From the observation that there is correlation among the signals of the different mixtures, it is proposed a modification in the NSAF (*Normalized Subband Adaptive Filter*) algorithm in order to accelerate its convergence rate. Simulation results with mixtures of speech signals in reverberant environments show the superior performance of the proposed algorithm in relation to the performances of the NLMS (*Normalized Least-Mean-Square*) and conventional NSAF, considering both the convergence speed and SIR (*Signal-to-Interference Ratio*) after convergence.

Keywords—Adaptive filtering, Normalized Subband Adaptive Filter (NSAF), Source separation.

I. INTRODUÇÃO

A evolução das telecomunicações tem impulsionado o desenvolvimento de técnicas eficientes para processamento digital de sinais. Em particular, técnicas de filtragem adaptativa têm atraído um grande interesse. Devido ao bom desempenho, baixa complexidade computacional e alta robustez, essas técnicas têm sido amplamente utilizadas em diversas aplicações, tais como identificação de sistemas, equalização de canais, cancelamento de eco e separação de fontes [1],[2],[3],[4], sendo esta última aplicação a que abordaremos neste artigo.

A maioria dos sinais de áudio aquisitados correspondem a misturas de sinais de várias fontes, como por exemplo fala, música e ruídos ambientes e de equipamentos. Separação de fontes consiste na recuperação dos sinais originais das fontes de interesse a partir de um ou mais sinais de mistura.

Bruna Luisa Ramos Prado Vasques pertence ao IPqM, Marinha do Brasil, Rio de Janeiro-RJ, Brasil. Jéssica Bartholdy Sanson e Mariane Rembold Petraglia pertencem ao Programa de Engenharia Elétrica da COPPE/UFRJ, Rio de Janeiro-RJ, Brasil. E-mails: brunalrpv@gmail.com, jessikbs.37@hotmail.com, mariane@pads.ufrj.br. Este trabalho foi parcialmente financiado pelo CNPq e pela FAPERJ.

Aplicações diretas incluem palestras em tempo real com tradução simultânea e amostragem de sons para composição de música eletrônica. Já muitas aplicações derivadas visam modificar o sinal de mistura, por exemplo, no aprimoramento do sinal da fala dentro de aparelhos auditivos, cancelamento de voz para karaokês e renderização de áudio para dispositivos multicanal a partir de gravações em CD estéreo. Em algumas aplicações, são conhecidos trechos de sinais originais presentes nas misturas antes de se fazer a separação [5]. Nesses casos, pode-se usar algoritmos adaptativos supervisionados, como os que serão abordados neste trabalho, para obter os coeficientes do sistema de separação.

Já a separação cega de fontes de áudio (*Blind Audio Source Separation* - BASS) tem sido um tema de intensa investigação durante os últimos anos. Vários métodos de sucesso surgiram, tais como ICA (*Independent Component Analysis*) [6], CASA (*Computational Auditory Scene Analysis*) [7] e SD (*Sparse Decomposition*) [8]. No entanto, ainda é difícil de se avaliarem as características e limitações de algoritmos de separação de fontes devido à falta de medidas de desempenho adequadas, principalmente no caso desafiador de misturas convolutivas. A separação de fontes supervisionada pode ser uma boa ferramenta para auxiliar no aperfeiçoamento das técnicas de separação cega de fontes.

Neste artigo investigaremos a separação de fontes a partir de misturas convolutivas de forma supervisionada, através de filtragem adaptativa, utilizando dois algoritmos: o NLMS (*Normalized Least-Mean-Square*) [1] e o NSAF (*Normalized Subband Adaptive Filters*) [9]. O algoritmo NSAF decompõe os sinais de entrada e desejados, gerando sinais de erro em subbandas que são empregados na adaptação dos coeficientes do filtro aplicado em toda a faixa de frequências. Esse procedimento faz com que o NSAF difira de seu algoritmo antecessor SAF (*Subband Adaptive Filter*) [10], que emprega adaptação e subfiltros distintos nas diferentes subbandas. A partir da observação de que existe correlação entre os sinais das misturas, é proposta uma modificação no algoritmo NSAF para a aplicação de separação de fontes supervisionadas, introduzindo-se as matrizes de correlação dos sinais das misturas em subbandas na equação de atualização dos coeficientes a fim de acelerar a convergência do algoritmo.

II. O PROBLEMA DE SEPARAÇÃO DE FONTES

Misturas lineares convolutivas de sinais podem ser definidas de acordo com a equação:

$$\mathbf{x}(n) = \mathbf{H}(n) * \mathbf{s}(n) \quad (1)$$

em que “ $\ast\Delta$ ” é o operador convolução, $\mathbf{H}(n)$ é a matriz de mistura de dimensões $P \times Q$, sendo Q o número de fontes e P o número de sensores, contendo as respostas impulsivas dos filtros de mistura da j -ésima fonte ao i -ésimo sensor $h_{ij}(n)$, i.e.,

$$\mathbf{H}(n) = \begin{bmatrix} h_{11}(n) & h_{12}(n) & \dots & h_{1Q}(n) \\ h_{21}(n) & h_{22}(n) & \dots & h_{2Q}(n) \\ \vdots & \vdots & \ddots & \vdots \\ h_{P1}(n) & h_{P2}(n) & \dots & h_{PQ}(n) \end{bmatrix} \quad (2)$$

$\mathbf{s}(n)$ é o vetor composto pelos sinais das fontes, i.e.,

$$\mathbf{s}(n) = [s_1(n) \quad s_2(n) \quad \dots \quad s_Q(n)]^T \quad (3)$$

e $\mathbf{x}(n)$ é o vetor formado pelos sinais que chegam aos sensores, dado por

$$\mathbf{x}(n) = [x_1(n) \quad x_2(n) \quad \dots \quad x_P(n)]^T \quad (4)$$

O problema de separação de fontes convolutivas lineares consiste em encontrar a matriz de separação $\mathbf{W}(n)$ capaz de obter estimativas dos sinais das fontes a partir dos sinais de mistura, ou seja,

$$\mathbf{y}(n) = \mathbf{W}(n) \ast \mathbf{x}(n) \quad (5)$$

em que

$$\mathbf{y}(n) = [y_1(n) \quad y_2(n) \quad \dots \quad y_Q(n)]^T \quad (6)$$

e a matriz de separação $\mathbf{W}(n)$ de dimensões $Q \times P$ é definida por

$$\mathbf{W}(n) = \begin{bmatrix} w_{11}(n) & w_{12}(n) & \dots & w_{1P}(n) \\ w_{21}(n) & w_{22}(n) & \dots & w_{2P}(n) \\ \vdots & \vdots & \ddots & \vdots \\ w_{Q1}(n) & w_{Q2}(n) & \dots & w_{QP}(n) \end{bmatrix} \quad (7)$$

Quando o número de sensores P é igual ao número de fontes Q , o problema de separação é denominado determinado, sendo esse o caso considerado neste trabalho.

III. ALGORITMOS ADAPTATIVOS

Nesta seção serão apresentados os algoritmos adaptativos utilizados, o NLMS e o NSAF. Também será descrito como cada algoritmo foi adaptado para o problema de separação supervisionada de fontes no contexto de misturas convolutivas.

A. Algoritmo NLMS

O algoritmo NLMS é um dos algoritmos de filtragem adaptativa mais populares devido à sua simplicidade computacional, convergência comprovada em ambientes estacionários com ruídos gaussianos, e comportamento robusto quando implementado com aritmética de precisão finita [11]. O algoritmo NLMS atualiza os coeficientes do filtro utilizando o erro entre a saída desejada e a obtida pelo filtro. Os vetores de entrada e de coeficientes no instante n são denotados, respectivamente, por $\mathbf{x}(n) = [x(n) \quad x(n-1) \quad \dots \quad x(n-N+1)]^T$ e $\mathbf{w}(n) = [w_0(n) \quad w_1(n) \quad \dots \quad w_{N-1}(n)]^T$, sendo N o

tamanho do filtro adaptativo. Portanto, temos as seguintes equações de saída e erro:

$$y(n) = \mathbf{w}^T(n)\mathbf{x}(n) \quad (8)$$

$$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n) \quad (9)$$

onde $d(n)$ é o o sinal de referência.

A atualização dos coeficientes é dada por [12]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\delta + \mathbf{x}^T(n)\mathbf{x}(n)}e(n)\mathbf{x}(n) \quad (10)$$

sendo o passo de adaptação μ introduzido para controlar o desajuste dos coeficientes após a convergência e o parâmetro de regularização δ para evitar passos muito grandes quando $\mathbf{x}^T(n)\mathbf{x}(n)$ se torna muito pequeno. A faixa de valores adequada para escolha de μ é [12]

$$0 \leq \mu \leq 2 \quad (11)$$

B. Algoritmo NLMS para o problema de separação de fontes

Para adequar o algoritmo NLMS ao problema de separação supervisionada de fontes (considerando-se por simplicidade o caso de duas fontes e dois sensores), obtêm-se os sinais de saída, $y_1(n)$ e $y_2(n)$, e de erro, $e_1(n)$ e $e_2(n)$, através das equações

$$y_i(n) = \mathbf{w}_i^T(n)\mathbf{x}(n) \quad (12)$$

$$e_i(n) = d_i(n) - y_i(n) \quad (13)$$

onde o vetor de entrada $\mathbf{x}(n)$ é composto pelos dois vetores dos sinais de mistura

$$\mathbf{x}_i(n) = [x_i(n) \quad x_i(n-1) \quad \dots \quad x_i(n-N+1)]^T \quad (14)$$

para $i = 1, 2$, ou seja,

$$\mathbf{x}(n) = [\mathbf{x}_1^T(n) \quad \mathbf{x}_2^T(n)]^T \quad (15)$$

e os vetores de coeficientes, $\mathbf{w}_1(n)$ e $\mathbf{w}_2(n)$, contêm os coeficientes dos dois filtros que geram cada saída, dados por

$$\mathbf{w}_1(n) = [\mathbf{w}_{11}^T(n) \quad \mathbf{w}_{12}^T(n)]^T \quad (16)$$

$$\mathbf{w}_2(n) = [\mathbf{w}_{21}^T(n) \quad \mathbf{w}_{22}^T(n)]^T \quad (17)$$

onde

$$\mathbf{w}_{ij}(n) = [w_{ij,0}(n) \quad w_{ij,1}(n) \quad \dots \quad w_{ij,N-1}(n)]^T \quad (18)$$

sendo i e j os índices relativos às fontes e misturas, respectivamente, e N o número de coeficientes de cada filtro de separação. Portanto, a cada iteração, haverá a atualização dos vetores de coeficientes, $\mathbf{w}_1(n)$ e $\mathbf{w}_2(n)$, de acordo com a equação

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \frac{\mu}{\delta + \mathbf{x}^T(n)\mathbf{x}(n)}e_i(n)\mathbf{x}(n) \quad (19)$$

C. Algoritmo NSAF

O algoritmo NLMS converge lentamente quando os sinais de entrada são coloridos. Para resolver este problema, uma abordagem atraente é utilizar filtragem em subbandas, na qual o sinal de entrada colorido é decomposto em faixas de frequência quase mutuamente exclusivas e o sinal decimado de cada subbanda é aproximadamente branco [13].

O algoritmo NSAF, proposto em [9], decompõe o sinal de entrada e o sinal desejado em subbandas, permitindo explorar as propriedades de cada sinal resultante. Este algoritmo utiliza o princípio da perturbação mínima, em que, de uma iteração para a seguinte, os coeficientes do filtro adaptativo devem ser alterados de forma mínima, sujeito a restrições impostas aos erros em subbandas após a atualização (erros *a posteriori* em subbandas nulos).

Uma característica única do algoritmo NSAF é que o erro do sinal é calculado em subbandas, porém a atualização dos coeficientes do filtro adaptativo é feita em banda completa, conforme ilustrado na Fig. 1. No algoritmo NSAF, o sinal

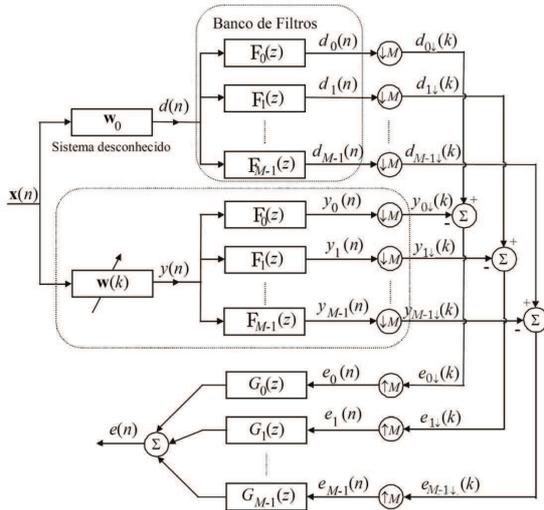


Fig. 1. Estrutura do algoritmo NSAF

desejado $d(n)$ e a saída do filtro $y(n)$ são decompostos em M subbandas pelos filtros de análise $F_0(z), \dots, F_{M-1}(z)$ e criticamente decimados (fator de decimação igual ao número de subbandas). Assim, o sinal de saída decimado da m -ésima subbanda é dado por

$$y_{m\downarrow}(k) = \sum_{l=0}^{N-1} \hat{w}_l(k) x_m(kM - l) = \hat{\mathbf{w}}^T(k) \mathbf{x}_m(k) \quad (20)$$

onde

$$\mathbf{x}_m(k) = [x_m(kM) \quad x_m(kM - 1) \quad \dots \quad x_m(kM - N + 1)]^T \quad (21)$$

$$\hat{\mathbf{w}}(k) = [\hat{w}_0(k) \quad \hat{w}_1(k) \quad \dots \quad \hat{w}_{N-1}(k)]^T \quad (22)$$

O sinal de erro correspondente é dado por

$$e_{m\downarrow}(k) = d_{m\downarrow}(k) - \hat{\mathbf{w}}^T(k) \mathbf{x}_m(k) \quad (23)$$

onde $d_{m\downarrow}(k)$ é o sinal desejado decimado da m -ésima subbanda.

A equação de atualização dos coeficientes do algoritmo NSAF é dada por

$$\hat{\mathbf{w}}(k+1) = \hat{\mathbf{w}}(k) + \mu \sum_{m=0}^{M-1} \frac{\mathbf{x}_m(k)}{\delta + \|\mathbf{x}_m(k)\|^2} e_{m\downarrow}(k) \quad (24)$$

onde os parâmetros μ e δ foram introduzidos com os mesmos objetivos descritos para o algoritmo NLMS.

Como a atualização dos coeficientes é feita em uma taxa mais baixa, decimando-se os sinais de erro, o aumento no custo computacional é muito pequeno [13] quando comparado ao do algoritmo NLMS, mas a convergência dos coeficientes se dá de forma mais rápida para sinais de entrada coloridos.

D. Algoritmo NSAF para o problema de separação de fontes

Para adequar o algoritmo NSAF ao problema de separação de fontes supervisionada, foi empregado um procedimento semelhante ao do algoritmo NLMS. Assim, as equações (23) e (24) foram modificadas para gerar os sinais de erro em subbandas e atualizar os coeficientes dos filtros de separação. Considerando o caso de duas fontes e dois sensores, os vetores com os coeficientes dos filtros que geram as estimativas das fontes são definidos como

$$\hat{\mathbf{w}}_1(k) = [\hat{w}_{11}^T(k) \quad \hat{w}_{12}^T(k)]^T \quad (25)$$

$$\hat{\mathbf{w}}_2(k) = [\hat{w}_{21}^T(k) \quad \hat{w}_{22}^T(k)]^T \quad (26)$$

sendo

$$\hat{w}_{ij}(k) = [\hat{w}_{ij,0}(k) \quad \hat{w}_{ij,1}(k) \quad \dots \quad \hat{w}_{ij,N-1}(k)]^T \quad (27)$$

o vetor com os N coeficientes do filtro $W_{ij}(z)$ que gera a parcela da estimativa da i -ésima fonte a partir da j -ésima mistura. Definindo o vetor com as amostras dos sinais de mistura na m -ésima subbanda como

$$\mathbf{x}_m(k) = [\mathbf{x}_{1,m}^T(k) \quad \mathbf{x}_{2,m}^T(k)]^T \quad (28)$$

onde

$$\mathbf{x}_{i,m}(k) = [x_{i,m}(kM) \quad x_{i,m}(kM - 1) \quad \dots \quad x_{i,m}(kM - N + 1)]^T \quad (29)$$

o sinal de erro em subbanda decimado e a equação de atualização dos filtros do sistema de separação relativos à estimativa da i -ésima fonte são dados, respectivamente, por

$$e_{i,m\downarrow}(k) = d_{i,m\downarrow}(k) - \hat{\mathbf{w}}_i^T(k) \mathbf{x}_m(k) \quad (30)$$

$$\hat{\mathbf{w}}_i(k+1) = \hat{\mathbf{w}}_i(k) + \mu \sum_{m=0}^{M-1} \frac{\mathbf{x}_m(k)}{\delta + \|\mathbf{x}_m(k)\|^2} e_{i,m\downarrow}(k) \quad (31)$$

E. Modificação no algoritmo NSAF para o problema de separação

Partindo da observação de que existe correlação entre os sinais das misturas, para acelerar a convergência do algoritmo NSAF foram introduzidas estimativas das matrizes correlação dos sinais de mistura em subbandas $\mathbf{R}_m(k) =$

$E[\mathbf{x}_m(k)\mathbf{x}_m^T(k)]$ na equação de atualização dos coeficientes, obtendo-se

$$\hat{\mathbf{w}}_i(k+1) = \hat{\mathbf{w}}_i(k) + \mu \sum_{m=0}^{M-1} \hat{\mathbf{R}}_m^{-1}(k) \frac{\mathbf{x}_m(k)}{\delta + \|\mathbf{x}_m(k)\|^2} e_{i,m\downarrow}(k) \quad (32)$$

onde

$$\hat{\mathbf{R}}_m(k) = \mathbf{x}_m(k)\mathbf{x}_m^T(k) = \begin{bmatrix} \hat{\mathbf{R}}_{m,11}(k) & \hat{\mathbf{R}}_{m,12}(k) \\ \hat{\mathbf{R}}_{m,21}(k) & \hat{\mathbf{R}}_{m,22}(k) \end{bmatrix} \quad (33)$$

Considerando que os sinais de entrada decimados nas subbandas são aproximadamente brancos, as matrizes $\hat{\mathbf{R}}_{m,ij}(k)$ podem ser aproximadas pelas matrizes diagonais

$$\hat{\mathbf{R}}_{m,ij}(k) \approx \sigma_{m,ij}^2(k)\mathbf{I} \quad (34)$$

onde

$$\sigma_{m,ij}^2(k) = \mathbf{x}_{i,m}^T(k)\mathbf{x}_{j,m}(k) \quad (35)$$

Definindo a matriz correlação inversa dos sinais de misturas na m -ésima subbanda como

$$\hat{\mathbf{R}}_m(k)^{-1} = \begin{bmatrix} \hat{\mathbf{S}}_{m,11}(k) & \hat{\mathbf{S}}_{m,12}(k) \\ \hat{\mathbf{S}}_{m,21}(k) & \hat{\mathbf{S}}_{m,22}(k) \end{bmatrix} \quad (36)$$

usando a aproximação (34) e observando que $\sigma_{m,12}^2(k) = \sigma_{m,21}^2(k)$, obtemos

$$\hat{\mathbf{S}}_{m,11}(k) = (\sigma_{m,11}^2(k) - \sigma_{m,12}^4(k)\sigma_{m,22}^{-2}(k))^{-1}\mathbf{I} \quad (37)$$

$$\hat{\mathbf{S}}_{m,22}(k) = (\sigma_{m,22}^2(k) - \sigma_{m,12}^4(k)\sigma_{m,11}^{-2}(k))^{-1}\mathbf{I} \quad (38)$$

$$\hat{\mathbf{S}}_{m,12}(k) = -\sigma_{m,11}^{-2}(k)\sigma_{m,12}^2(k)(\sigma_{m,22}^2(k) - \sigma_{m,12}^4(k)\sigma_{m,11}^{-2}(k))^{-1}\mathbf{I} \quad (39)$$

$$\hat{\mathbf{S}}_{m,21}(k) = -\sigma_{m,22}^{-2}(k)\sigma_{m,12}^2(k)(\sigma_{m,11}^2(k) - \sigma_{m,12}^4(k)\sigma_{m,22}^{-2}(k))^{-1}\mathbf{I} \quad (40)$$

O aumento na complexidade computacional resultante da introdução das inversas das matrizes de correção na equação de atualização dos coeficientes (31) é de $3M^2 + 8M$ multiplicações. Como, em geral, o número de subbandas M é bem menor do que o comprimento dos filtros adaptativos N , esse aumento não é significativo.

IV. SINAL DESEJADO PARA A AVALIAÇÃO DA SEPARAÇÃO CEGA DE FONTES

No problema de separação cega de sinais de áudio em misturas convolutivas, em geral não é possível recuperar os sinais originais das fontes, sendo permitido obter como solução válida os sinais originais filtrados. Denotando por $H_{ij}(z)$ a função de transferência correspondente da i -ésima fonte ao j -ésimo sensor, pode-se escrever, no domínio Z , os sinais das misturas $X_i(z)$ e de saída do sistema de separação $Y_i(z)$ para o problema de duas fontes e dois sensores como

$$X_1(z) = H_{11}(z)S_1(z) + H_{12}(z)S_2(z) \quad (41)$$

$$X_2(z) = H_{21}(z)S_1(z) + H_{22}(z)S_2(z) \quad (42)$$

$$Y_1(z) = W_{11}(z)X_1(z) + W_{12}(z)X_2(z) \quad (43)$$

$$Y_2(z) = W_{21}(z)X_1(z) + W_{22}(z)X_2(z) \quad (44)$$

Substituindo as equações (41) e (42) em (43) e (44), tem-se

$$Y_1(z) = (W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z))S_1(z) + (W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z))S_2(z) \quad (45)$$

$$Y_2(z) = (W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z))S_1(z) + (W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z))S_2(z) \quad (46)$$

Para encontrarmos uma solução possível para o problema foi imposto ao sistema que o sinal desejado de saída $D_1(z)$ deveria representar o sinal original $S_1(z)$ filtrado e o sinal desejado de saída $D_2(z)$, o sinal $S_2(z)$ filtrado. Para isso, basta igualarmos a zero a parcela do sinal correspondente a $S_2(z)$ na equação de $Y_1(z)$ e a parcela do sinal correspondente a $S_1(z)$ na equação de $Y_2(z)$, ou seja,

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (47)$$

$$W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (48)$$

As equações acima possuem infinitas soluções e a solução adotada foi $W_{11}(z) = H_{22}(z)$, $W_{12}(z) = -H_{12}(z)$, $W_{21}(z) = H_{21}(z)$ e $W_{22}(z) = -H_{11}(z)$. Assim, os sinais $D_1(z)$ e $D_2(z)$ passam a ser:

$$D_1(z) = (W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z))S_1(z) = (H_{22}(z)H_{11}(z) - H_{12}(z)H_{21}(z))S_1(z) \quad (49)$$

$$D_2(z) = (W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z))S_2(z) = (H_{12}(z)H_{21}(z) - H_{11}(z)H_{22}(z))S_2(z) \quad (50)$$

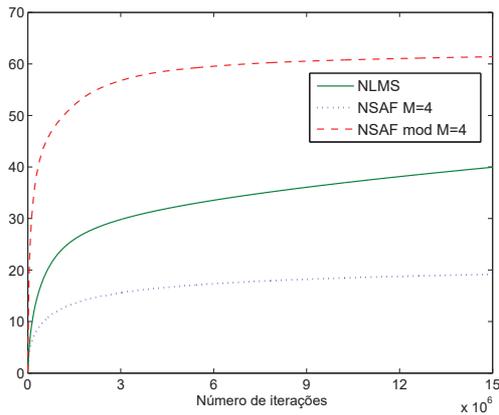
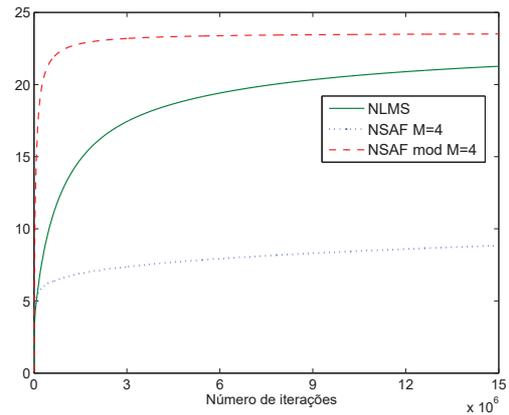
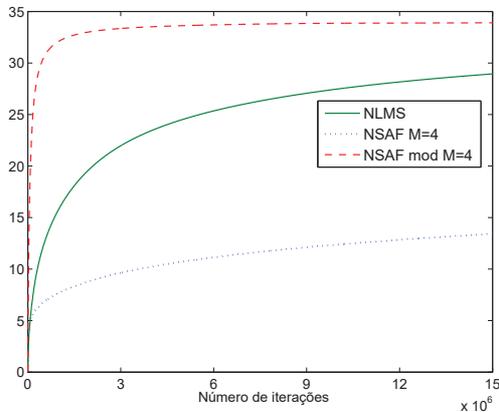
Os sinais desejados das equações acima foram usados nas simulações apresentadas a seguir.

V. RESULTADOS DAS SIMULAÇÕES

Os algoritmos adaptativos para separação supervisionada de fontes descritos na Seção III foram avaliados empregando-se duas misturas convolutivas de dois sinais de voz, correspondendo aos sinais aquisitados por dois microfones afastados de 5 cm em ambiente reverberante e amostrados em 8 kHz. As simulações da propagação acústica de salas foram desenvolvidas usando o modelo "Image-Source" descrito em [14]. Foram feitas simulações para três diferentes tempos de reverberação ($T60$): 0.1 s, 0.25 s e 0.5 s. A medida de avaliação utilizada foi a SIR (*Signal to Interference Ratio*). Para a decomposição em subbandas, empregou-se um banco de filtros modulado por cosseno [15] com $M = 4$ subbandas e filtro protótipo de comprimento 32. O passo de adaptação e parâmetro de regularização utilizados em todos os algoritmos foram $\mu = 0.5$ e $\delta = 0.1$.

Nas Figs. 2, 3 e 4 são mostradas as evoluções da SIR ao longo das iterações obtidas com os algoritmos NLMS, NSAF e NSAF modificado para $T60 = 0.1$ s, 0.25 s e 0.5 s, respectivamente, com filtros de separação de comprimentos $N = 623$, 1599 e 3227, iguais aos comprimentos dos filtros de mistura (denominados L).

Pode-se verificar nessas figuras que o algoritmo NSAF modificado apresenta taxa de convergência consideravelmente maior que as do NLMS e do NSAF convencional para o problema de separação supervisionada de fontes. Como esperado, a introdução da matriz de correlação dos sinais de mistura na equação de atualização dos coeficientes acelera a convergência do algoritmo NSAF, a qual é lenta devido à forte correlação entre os sinais das misturas. O processamento em subbandas explora as características dos sinais de mistura nas subbandas, através da normalização do passo de adaptação, resultando em

Fig. 2. Evolução da SIR(dB) para $T60 = 0.1$ s.Fig. 4. Evolução da SIR(dB) para $T60 = 0.5$ s.Fig. 3. Evolução da SIR(dB) para $T60 = 0.25$ s.

valores de SIR mais altos do que os obtidos com o algoritmo NLMS.

Das Figs. 2, 3 e 4, pode-se também concluir que há uma limitação no modelo de separação de fontes que emprega filtros de resposta ao impulso finita (FIR), a qual tem efeito mais acentuado na SIR conforme o tempo de reverberação aumenta. Portanto, por melhor que seja o algoritmo de separação cega implementado, para os casos de alta reverberação, a SIR máxima a ser alcançada com filtros de separação FIR de comprimento $N = L$ está limitada pelo modelo do sistema de separação empregado (em torno de 23 dB, para $T60 = 0.5$ s), enquanto que para tempos de reverberação menores é possível obter altos valores de SIR (acima de 60 dB para $T60 = 0.1$ s).

VI. CONCLUSÕES

Neste trabalho foram empregados algoritmos adaptativos supervisionados para determinar as limitações do modelo do sistema de separação de misturas convolutivas composto por filtros FIR. Devido à correlação entre os sinais de mistura, os algoritmos adaptativos convencionais, tanto em banda cheia quanto em subbandas, apresentam convergência lenta e resultam em baixos valores de SIR. Para melhorar o desempenho do algoritmo NSAF, foram introduzidas na equação

de atualização dos coeficientes as matrizes de correlação dos sinais de mistura nas diferentes subbandas, as quais, utilizando a propriedade de branqueamento dos sinais decimados em subbandas, podem ser aproximadas por matrizes diagonais em blocos, resultando em um aumento pequeno na complexidade computacional em relação ao algoritmo NSAF convencional. Resultados de simulações com sinais de voz em ambientes com reverberação confirmaram o melhor desempenho do algoritmo proposto em relação aos algoritmos convencionais, e quantificaram as limitações na SIR obtidas com o modelo de sistema de separação que emprega filtros FIR.

REFERÊNCIAS

- [1] S. S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, 4th Ed., N.J: Prentice, 2002.
- [2] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Wiley, 2003.
- [3] A. H. Sayed, *Adaptive Filters*. Wiley, 2008.
- [4] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. Wiley, 1998.
- [5] P. Smaragdakis, B. Raj, and M. Shashanka, "Blind source separation: statistical principles," *Independent Component Analysis and Signal Separation: 7th International Conference*, pp.414–421, Sep. 2007.
- [6] J. F. Cardoso, "Supervised and Semi-supervised Separation of Sounds from Single-Channel Mixtures," in *Proceedings of the IEEE*, v. 9, no. 10, pp. 2009–2025, Oct. 1998.
- [7] D. Ellis, *Prediction-driven computational auditory scene analysis*. Ph.D. dissertation, MIT, Jun. 1998.
- [8] M. Zibulevsky and B. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Computation*, v. 13, no. 4, pp. 863–882, Apr. 2001.
- [9] K. A. Lee and W. S. Gan, "Improving Convergence of the NLMS Algorithm Using Constrained Subband Updates," *IEEE Signal Processing Letters*, v. 11, no. 9, pp. 736–739, 2004.
- [10] J. J. Shynk, "Frequency domain and multirate adaptive filtering," *IEEE Signal Processing Mag.*, v. 9, pp. 14–37, Jan. 1992.
- [11] C. Ruch Jr., *Implementação de Filtros Adaptativos em FPGA*. Tese de doutorado, Programa de Pós-Graduação em Engenharia Elétrica- PUC, 2006.
- [12] P. S. R. Dimiz, *Adaptive Filtering: Algorithms and Practical Implementation*. Springer US, 4th Ed., 2013.
- [13] K.A. Lee, W.S. Gan, S.M. Kuo *Subband adaptive Filtering: Theory and Implementation*. Wiley, Hoboken, NJ, 2009.
- [14] Lehmann, E., Johansson, A., Nordholm, S., " Reverberation-Time Prediction Method for Room Impulse Responses Simulated with the Image-Source Model," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2007, pp. 159–162, Oct 2007.
- [15] S. K. K. Mitra *Digital Signal Processing: A Computer-Based Approach*. McGraw-Hill Higher Education, 2nd Ed. 2000.