

# Estimação de Sinais de Fontes a partir de Misturas Convolutivas

Carlos Vinícius C. Campos, Mariane R. Petraglia, José Gabriel R. C. Gomes

**Resumo**—A separação de sinais de voz em ambientes com alta reverberação continua sendo um grande desafio. Este artigo apresenta um algoritmo offline capaz de reconstruir os sinais das fontes a partir dos sinais gravados, utilizando os coeficientes dos filtros que modelam os caminhos desde as fontes até os microfones. Os resultados de simulações demonstraram que o algoritmo proposto é mais robusto em ambientes com tempo de reverberação elevado em comparação com abordagens tradicionais, obtendo valores superiores em todas as métricas de qualidade comumente utilizadas.

**Palavras-Chave**—Separação de fontes, Ambientes reverberantes, Processamento de voz.

**Abstract**—Separating voice signals in environments with high reverberation remains a major challenge. This article presents an offline algorithm capable of reconstructing source signals from recorded signals, using filter coefficients that model the paths from the sources to the microphones. The simulation results demonstrated that the proposed algorithm is more robust in environments with high reverberation time compared to traditional approaches, obtaining superior values in all commonly used quality metrics.

**Keywords**—Source separation, Reverberant environments, Speech processing.

## I. INTRODUÇÃO

A captação de sinais de áudio emitidos por diversas fontes em ambientes reverberantes, como auditórios e salas de reuniões, por meio de microfones, resulta em sinais que contêm misturas variadas das fontes, dependendo do trajeto percorrido pelas ondas sonoras de cada uma delas. Esses sinais dos microfones contêm somas de versões atenuadas e atrasadas dos sinais das fontes, devido ao tempo necessário para percorrer os caminhos diretos até os microfones, bem como devido às reflexões ocorridas na sala.

Ao processar esses sinais para separá-los utilizando filtros digitais lineares, os efeitos de reverberação na sala geralmente persistem, prejudicando tanto a inteligibilidade dos sinais de fala quanto o desempenho das métricas de avaliação da separação. Isso é especialmente evidente no algoritmo TRINICON (do inglês *TRIPle-N Independent component analysis for CONvulsive mixtures*) [3], que procura estimar os sinais das fontes eliminando as correlações cruzadas entre os sinais capturados pelos microfones, resultando em sinais de saída ainda contendo reverberação. Outras abordagens propostas para a obtenção dos sinais das fontes a partir de misturas convolutivas

incluem a Análise de Componentes Independentes no domínio da frequência (FD-ICA) e a Análise de Vetores Independentes (IVA). Os desempenhos dos algoritmos TRINICON, FD-ICA e IVA são comparados em [1] no contexto da separação de sinais de voz, revelando baixos valores de métricas de desempenho em ambientes com alta reverberação.

Neste trabalho é apresentado um método iterativo *offline* para obter estimativas dos sinais das fontes a partir dos sinais e sistemas de mistura. Tendo os coeficientes dos filtros FIR (*Finite Impulse Response*) que modelam os caminhos das fontes até os microfones, o algoritmo aqui apresentado tem por objetivo obter os sinais das fontes sem reverberação, uma vez que procura reconstruir os sinais originais de forma iterativa, a partir dos sinais captados pelos microfones, sem realizar a filtragem pelo sistema de separação, como os algoritmos tradicionais.

## II. DESCRIÇÃO DO PROBLEMA DE SEPARAÇÃO DE MISTURAS CONVOLUTIVAS

Nesta seção, descreveremos o modelo do sistema de mistura convolutiva abordado neste trabalho, e discutiremos a diferença entre o método tradicional e o proposto para obter os sinais separados. Em seguida, serão apresentadas as definições das métricas de qualidade de separação utilizadas.

### A. Modelos de Sistemas de Mistura Convolutiva e de Separação Tradicional

Dado que se tem  $P$  fontes e os sinais  $x_q$  de  $Q$  microfones, queremos determinar estimativas  $\hat{s}_p$  para os sinais  $s_p$  das  $P$  fontes. Considera-se que o caminho do sinal das fontes até os microfones pode ser modelado por um conjunto de filtros FIR, com respostas ao impulso  $h_{pq}$  de comprimento  $L$ , chamado de Sistema de Mistura. Os sinais misturados, capturados pelos  $Q$  sensores, são dados por

$$x_q[n] = \sum_{p=1}^P \sum_{l=0}^{L-1} h_{pq}[l] s_p[n-l]. \quad (1)$$

Em algoritmos tradicionais, como o TRINICON [4], a estimação dos sinais das fontes é realizada por meio da filtragem dos sinais  $x_q$  dos microfones, usando-se normalmente filtros FIR, também de comprimento  $L$ , com coeficientes  $w_{pq}$ , sendo os sinais de saída  $y_p$  dados por

$$y_p[n] = \sum_{q=1}^Q \sum_{l=0}^{L-1} w_{pq}[l] x_q[n-l]. \quad (2)$$

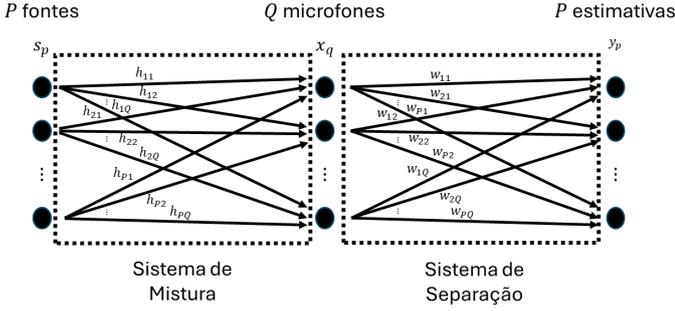


Fig. 1. Representação dos Sistemas de Mistura Convolutiva e de Separação Tradicional.

Esse modelo para os sistemas de mistura e separação é ilustrado pelo diagrama da Fig. 1.

Essa abordagem apresenta facilidades, como baixa complexidade de implementação, uma vez que a saída é obtida por simples filtragens FIR dos sinais captados pelos microfones, sendo utilizadas técnicas bastante diretas de filtragem adaptativa para a obtenção dos coeficientes dos filtros que constituem o sistema de separação. Essa abordagem, no entanto, proverá estimativas contaminadas pela reverberação da sala, o que dificulta tanto a avaliação objetiva da qualidade de separação das fontes por métricas convencionais, que serão melhor descritas a seguir, quanto a inteligibilidade do que está sendo dito, prejudicando o desempenho de equipamentos que utilizam comandos de voz.

Neste contexto, o algoritmo de estimação dos sinais apresentado neste trabalho procura obter diretamente os  $P$  sinais que, ao serem filtrados pelos filtros FIR que modelam o Sistema de Mistura, nos forneçam os  $Q$  sinais captados pelos microfones, não utilizando, portanto, o Sistema de Separação mostrado na Fig. 1.

### B. Métricas de Qualidade de Separação

As métricas SIR, SAR e SDR se baseiam em modelar a estimativa do sinal da  $p$ -ésima fonte  $\hat{s}_p$  como [2]

$$\hat{s}_p = s_p^{fonte} + e_p^{interf} + e_p^{artef}, \quad (3)$$

onde  $s_p^{fonte}$  é o sinal original da  $p$ -ésima fonte,

$$e_p^{interf} = \sum_{p' \neq p} \frac{\langle \hat{s}_p, s_{p'} \rangle s_{p'}}{\|s_{p'}\|^2} \quad (4)$$

é a deformação provocada pela interferência das outras fontes e  $e_p^{artef}$  contém os artefatos introduzidos pelos algoritmos de separação.

Assim, a qualidade de separação pode ser medida por meio das seguintes métricas médias, considerando-se as  $P$  fontes:

- **Signal to Interference Ratio (SIR):** quantifica a deformação média provocada pela interferência das outras fontes, e é calculada por:

$$\text{SIR}(\text{dB}) = 10 \log_{10} \left( \frac{1}{P} \sum_{p=1}^P \frac{\|s_p^{fonte}\|^2}{\|e_p^{interf}\|^2} \right); \quad (5)$$

- **Signal to Artifact Ratio (SAR):** quantifica a razão média entre a energia do sinal original e a energia do erro referente aos artefatos introduzidos pelos algoritmos de separação, e é obtida por:

$$\text{SAR}(\text{dB}) = 10 \log_{10} \left( \frac{1}{P} \sum_{p=1}^P \frac{\|s_p^{fonte}\|^2}{\|e_p^{artef}\|^2} \right); \quad (6)$$

- **Signal to Distortion Ratio (SDR):** quantifica a razão média entre a energia do sinal original da fonte sendo estimada e a energia da distorção, onde a distorção é o sinal resultante da soma entre os erros gerados pela interferência das outras fontes e dos erros introduzidos pelos algoritmos de separação, ou seja,

$$\text{SDR}(\text{dB}) = 10 \log_{10} \left( \frac{1}{P} \sum_{p=1}^P \frac{\|s_p^{fonte}\|^2}{\|e_p^{interf} + e_p^{artef}\|^2} \right). \quad (7)$$

## III. ALGORITMO DE ESTIMAÇÃO DAS FONTES

O algoritmo proposto para estimação dos sinais das fontes será apresentado a seguir em duas partes. Primeiramente, será mostrada a equação base do método e o processo iterativo realizado na sua execução e, em seguida, será detalhado o processo utilizado para a obtenção dos coeficientes, denotados como coeficientes ótimos, utilizados na modelagem do Sistema de Mistura.

### A. Equação-Base do Método Proposto

Pela Eq. (1) é possível observar que cada amostra do sinal  $x_q[n]$  é dada em função de  $L$  amostras de cada um dos sinais  $s_p[n]$ . Sendo

$$\mathbf{h}_{pq}(L) = [h_{pq}[0] \ h_{pq}[1] \ \dots \ h_{pq}[L-1]] \quad (8)$$

e reescrevendo a contribuição do sinal de cada fonte em forma matricial, obtém-se

$$x_q[n] = \sum_{p=1}^P \mathbf{h}_{pq}(L) \begin{bmatrix} s_p[n] \\ s_p[n-1] \\ \vdots \\ s_p[n-(L-1)] \end{bmatrix}. \quad (9)$$

A fim de se obter  $K$  amostras dos vetores dos microfones a partir dos sinais das fontes, pode-se definir

$$\mathbf{x}_q(n, K) = [x_q[n] \ x_q[n-1] \ \dots \ x_q[n-(K-1)]]^T, \quad (10)$$

$$\mathbf{s}_p(n, K, L) = [s_p[n] \ s_p[n-1] \ \dots \ s_p[n-K-L+2]]^T, \quad (11)$$

e a matriz circulante de dimensão  $K \times (K+L-1)$

$$\mathbf{H}_{pq}(K, L) = \begin{bmatrix} \mathbf{h}_{pq}(L) & \mathbf{0}_{1 \times L} & \dots & \mathbf{0}_{1 \times L} \\ \mathbf{0}_{1 \times L} & \mathbf{h}_{pq}(L) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{1 \times L} \\ \mathbf{0}_{1 \times L} & \dots & \mathbf{0}_{1 \times L} & \mathbf{h}_{pq}(L) \end{bmatrix}, \quad (12)$$

sendo  $\mathbf{0}_{1 \times L}$  o vetor nulo de dimensão  $1 \times L$ , de modo que o vetor com  $K$  amostras do sinal do microfone  $x_q[n]$  pode ser obtido por

$$\mathbf{x}_q(n, K) = \sum_{p=1}^P \mathbf{H}_{pq}(K, L) \mathbf{s}_p(n, K, L). \quad (13)$$

A Eq. (13) pode ser expandida para o cálculo de  $K$  amostras dos sinais de todos os microfones combinando-se as expressões de  $\mathbf{x}_q(n, K)$  referentes a  $q = 1, \dots, Q$  em um vetor e empilhando-se as amostras dos sinais das fontes  $\mathbf{s}_p[n]$ , para  $p = 1, \dots, P$ , em um outro vetor.

Neste artigo, consideraremos, sem perda de generalidade, o caso em que o número de fontes e o número de microfones é igual a 2 ( $P = Q = 2$ ). Assim, os vetores com as amostras de todos os sinais das misturas e das fontes são dados, respectivamente, por

$$\mathbf{X}_K[n] = \begin{bmatrix} \mathbf{x}_1(n, K)^T & \mathbf{x}_2(n, K)^T \end{bmatrix}^T, \quad (14)$$

$$\mathbf{S}_K[n, L] = \begin{bmatrix} \mathbf{s}_1(n, K, L)^T & \mathbf{s}_2(n, K, L)^T \end{bmatrix}^T, \quad (15)$$

e a nova matriz de coeficientes  $\mathbf{H}(K, L)$  tem a seguinte forma:

$$\mathbf{H}(K, L) = \begin{bmatrix} \mathbf{H}_{11}(K, L) & \mathbf{H}_{12}(K, L) \\ \mathbf{H}_{21}(K, L) & \mathbf{H}_{22}(K, L) \end{bmatrix}. \quad (16)$$

Com essas definições, a Eq. (13) pode ser reescrita como

$$\mathbf{X}_K[n] = \mathbf{H}(K, L) \mathbf{S}_K[n, L]. \quad (17)$$

Tendo em vista que o objetivo do algoritmo proposto é estimar os sinais  $\mathbf{s}_p$  das fontes com base nos sinais  $\mathbf{x}_q$  dos microfones, a Eq. (17) pode ser interpretada como um sistema linear subdeterminado com  $2K$  equações com  $2(K + L - 1)$  variáveis. É importante observar que as definições dadas pelas Eqs. (14), (15) e (16) podem ser facilmente expandidas para o caso genérico de  $P$  fontes e  $Q$  microfones.

### B. Algoritmo Iterativo para Estimação das Fontes

O sistema subdeterminado da Eq. (17) pode ser reescrito de modo a gerar a equação-base para um método iterativo. Uma possível abordagem é minimizar os erros entre os sinais dos microfones calculados a partir das estimativas das fontes na iteração  $i$  e os sinais  $\mathbf{x}_q$  dos microfones. O vetor de  $K + L - 1$  amostras estimadas dos sinais das fontes no índice  $n$  na  $i$ -ésima iteração será dado por

$$(\hat{\mathbf{S}}_K[n, L])_i = \begin{bmatrix} (\hat{s}_1[n])_i \\ (\hat{s}_1[n-1])_i \\ \vdots \\ (\hat{s}_1[n - (K-1) - (L-1)])_i \\ (\hat{s}_2[n])_i \\ (\hat{s}_2[n-1])_i \\ \vdots \\ (\hat{s}_2[n - (K-1) - (L-1)])_i \end{bmatrix}. \quad (18)$$

Com base nesse vetor e utilizando a matriz de coeficientes  $\mathbf{H}(K, L)$  da Eq. (16), pode-se obter o vetor  $(\hat{\mathbf{X}}_K[n])_i$  por meio da seguinte equação:

$$(\hat{\mathbf{X}}_K[n])_i = \mathbf{H}(K, L) (\hat{\mathbf{S}}_K[n, L])_i. \quad (19)$$

O objetivo do Algoritmo de Estimação das Fontes proposto será, então, encontrar os sinais  $(\hat{s}_p[n])_i$  que minimizam o erro médio quadrático

$$\text{MSE}_i = \frac{1}{K} \|\mathbf{X}_K[n] - (\hat{\mathbf{X}}_K[n])_i\|^2. \quad (20)$$

O gradiente dessa função,  $\text{MSE}_i$ , em relação a cada uma das amostras  $\mathbf{s}_p[m]$ , onde  $p = 1, 2$  e  $m = n, n-1, \dots, n - (K-1) - (L-1)$ , é dado por

$$\nabla \text{MSE}_i = -\frac{2}{K} \mathbf{H}(K, L)^T (\mathbf{X}_K[n] - (\hat{\mathbf{X}}_K[n])_i) \quad (21)$$

Aplicando-se o método do gradiente descendente, a atualização das amostras do sinal pode ser feita pela equação:

$$(\hat{\mathbf{S}}_K[n, L])_{i+1} = (\hat{\mathbf{S}}_K[n, L])_i + \alpha \mathbf{H}(K, L)^T (\mathbf{X}_K[n] - (\hat{\mathbf{X}}_K[n])_i) \quad (22)$$

onde  $\alpha$  é o passo de atualização do algoritmo.

Inicia-se o processo de estimação atribuindo valores arbitrários para os sinais das fontes. Neste ponto, realiza-se uma atualização das estimativas das amostras para  $n = n_0$ , em seguida para  $n = n_0 + 1$ , e assim por diante, até  $n = N$ , onde  $N$  é o comprimento dos sinais dos microfones. Nessa abordagem, este é o fim da primeira iteração.

Repete-se, então, todo esse processo, revisando-se, a cada iteração, todas as amostras das estimativas dos sinais das fontes. Esse processo é ilustrado pelo diagrama da Fig. 2, dentro do bloco com entrada no índice  $n = n_0$  e saída no índice  $n = n_0 + 1$ .

Como essa implementação demanda a disponibilidade de todas as amostras dos sinais dos microfones desde a primeira iteração, esta é uma abordagem *offline* para a estimação dos sinais dos microfones. Cada vez que o bloco básico é executado,  $K + L - 1$  amostras dos sinais das fontes são atualizadas. Esse processo segue até que todas as amostras dos sinais das fontes sejam atualizadas pelo menos uma vez a cada iteração.

### C. Coeficientes Ótimos do Sistema de Separação

Conforme ideia mostrada em [7], pode-se deduzir os coeficientes ótimos do modelo de separação com base em coeficientes do modelo do Sistema de Mistura.

Para um sistema com duas fontes e dois sinais de saída, a Eq. (17), do Sistema de Mistura, pode ser reescrita por meio do sistema de equações matriciais<sup>1</sup>:

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{H}_{11} \mathbf{S}_1 + \mathbf{H}_{21} \mathbf{S}_2 \\ \mathbf{X}_2 &= \mathbf{H}_{12} \mathbf{S}_1 + \mathbf{H}_{22} \mathbf{S}_2 \end{aligned} \quad (23)$$

e o sistema de separação pode ser expresso como:

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{W}_{11} \mathbf{X}_1 + \mathbf{W}_{12} \mathbf{X}_2 \\ \mathbf{Y}_2 &= \mathbf{W}_{21} \mathbf{X}_1 + \mathbf{W}_{22} \mathbf{X}_2. \end{aligned} \quad (24)$$

Substituindo  $\mathbf{X}_1$  e  $\mathbf{X}_2$ , da Eq. (23), na Eq. (24), pode-se escrever os sinais das saídas em função dos sinais das fontes como

$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{11} \mathbf{H}_{11} + \mathbf{W}_{12} \mathbf{H}_{12} & \mathbf{W}_{11} \mathbf{H}_{21} + \mathbf{W}_{12} \mathbf{H}_{22} \\ \mathbf{W}_{21} \mathbf{H}_{11} + \mathbf{W}_{22} \mathbf{H}_{12} & \mathbf{W}_{21} \mathbf{H}_{21} + \mathbf{W}_{22} \mathbf{H}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \end{bmatrix}. \quad (25)$$

<sup>1</sup>Os índices  $K, L$  e  $n$  foram omitidos para simplificar a notação.

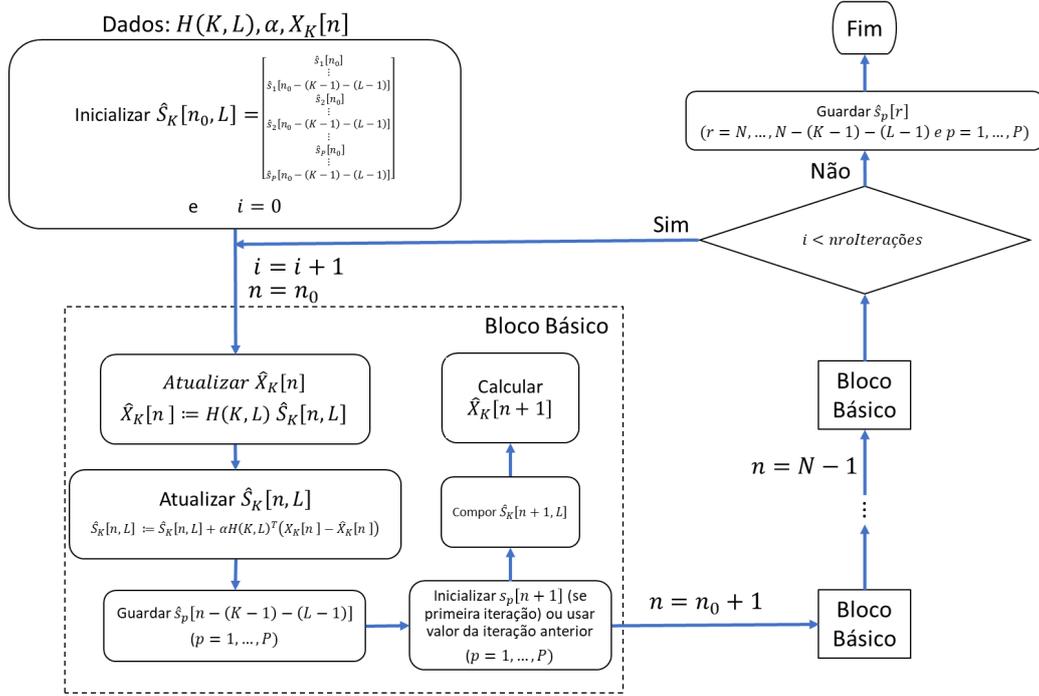


Fig. 2. Fluxograma do Algoritmo de Estimação das Fontes

Uma vez que o objetivo do algoritmo é separar os sinais das fontes, neste ponto, pode-se aplicar as condições  $(\mathbf{W}_{11}\mathbf{H}_{21} + \mathbf{W}_{12}\mathbf{H}_{22}) = \mathbf{0}$  e  $(\mathbf{W}_{21}\mathbf{H}_{11} + \mathbf{W}_{22}\mathbf{H}_{12}) = \mathbf{0}$ . De modo a atender essas condições, basta impor que

$$\begin{aligned} \mathbf{W}_{11} &= \mathbf{H}_{22} & , & & \mathbf{W}_{12} &= -\mathbf{H}_{21}, \\ \mathbf{W}_{21} &= \mathbf{H}_{12} & , & & \mathbf{W}_{22} &= -\mathbf{H}_{11}. \end{aligned} \quad (26)$$

Para determinar os coeficientes  $\mathbf{H}_{ii}$  são, então, utilizados dois sinais de ruído branco independentes no lugar das fontes originais e é aplicado um algoritmo de identificação de sistemas [5] que minimiza o erro médio quadrático entre os sinais  $x_q$  e as saídas dos filtros que modelam o sistema de mistura, fornecendo os coeficientes que denotaremos como coeficientes ótimos. Nesse ponto utiliza-se o fato desses coeficientes serem independentes dos sinais das fontes, dependendo apenas das características do ambiente. Em uma situação prática, essa substituição dos sinais das fontes pelos sinais de ruído branco poderia ser feita com a utilização de alto-falantes reproduzindo esses ruídos brancos nas posições em que as fontes de áudio estarão posicionadas posteriormente, funcionando como uma espécie de calibração do sistema. Outra possibilidade é utilizar os filtros de separação dos sinais das fontes obtidos por um algoritmo de separação cega.

Em posse dos coeficientes  $\mathbf{H}_{ii}$ , os coeficientes de separação ótimos  $\mathbf{W}_{ii}$  são determinados com base nas igualdades expressas na Eq. (26).

#### IV. RESULTADOS DE ESTIMAÇÕES DE FONTES ACÚSTICAS EM AMBIENTES REVERBERANTES

Para avaliar o desempenho do método proposto em comparação ao método de estimação de fontes tradicional, aplicamos os dois algoritmos a misturas de sinais de voz em ambientes com tempos de reverberação variados. Os sinais das

misturas foram obtidos filtrando-se sinais anecoicos de uma voz masculina e uma feminina, amostrados em 16 kHz, por respostas ao impulso de salas geradas usando o Modelo Fonte-Imagem (ISM) [6]. Em todas as simulações, foram utilizados dois microfones espaçados de 5 cm, sendo o par de fontes posicionado a 1 metro dos microfones, a  $\pm 45^\circ$  da direção perpendicular ao eixo central dos sensores. Os algoritmos foram implementados e os resultados obtidos por meio do *software* MATLAB® versão R2023a.

Dados os coeficientes  $h_{pq}$  dos filtros FIR que modelam o caminho desde as fontes até os microfones para diferentes comprimentos  $L$  desses filtros, a seguir são apresentados os resultados obtidos com a utilização desses coeficientes na construção das matrizes  $\mathbf{H}(K, L)$  de cada caso. Em todos os testes, utilizou-se  $K = 5$ ,  $N = 3000$  iterações e passo  $\alpha = 0.1$ .

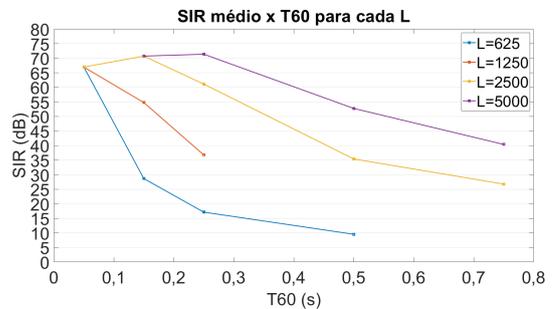


Fig. 3. SIR Médio das Estimativas obtidas pelo Algoritmo de Estimação das Fontes Proposto

Nas Figs. 3 e 4 são mostrados os valores médios da SIR obtidos com os algoritmos Proposto e Tradicional, respectivamente, para diferentes tempos de reverberação ( $T_{60}$ ) e

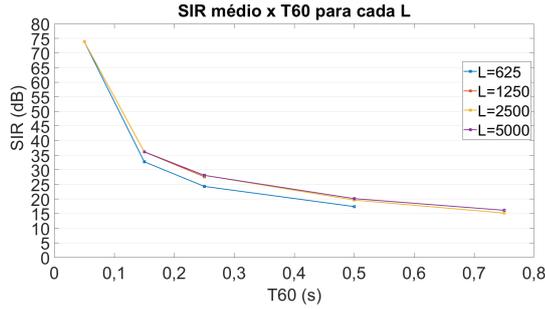


Fig. 4. SIR Médio das Estimativas obtidas pelo Modelo Tradicional

tamanhos de filtros  $L$ . Pode-se observar que para ambientes pouco reverberantes, com  $T_{60} = 50$  ms, tanto o modelo tradicional quanto o algoritmo de estimação de sinais apresentaram SIR na faixa de 70 dB, necessitando de menos coeficientes,  $L = 625$ , para atingir essa qualidade. Para ambientes mais reverberantes, a qualidade das estimativas obtidas pelo modelo tradicional decai rapidamente com o aumento do  $T_{60}$ , enquanto a qualidade das estimativas obtidas pelo algoritmo proposto se mantém mais alta, para um número suficiente de coeficientes. Por exemplo, utilizando-se  $L = 5000$  coeficientes para  $T_{60} = 750$  ms, o algoritmo proposto obtém uma estimativa com SIR de aproximadamente 40 dB, em comparação a 15 dB obtida com o modelo tradicional. Isso ocorre tanto para as métricas SIR quanto para as métricas SAR e SDR, como pode ser verificado nas Figs. 5 e 6. Estas figuras contêm os valores médios da SAR obtidos com o método proposto e com o tradicional, respectivamente. Os gráficos da SDR não foram incluídos por serem semelhantes aos da SAR, uma vez que os artefatos predominaram em relação às interferências nos testes realizados.

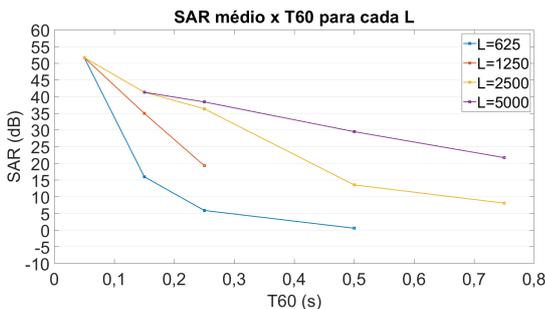


Fig. 5. SAR Médio das Estimativas obtidas pelo Algoritmo de Estimação das Fontes Proposto

Um aspecto que pode ser observado em relação ao algoritmo proposto é que o seu desempenho é bastante sensível à quantidade de coeficientes dos filtros. No entanto, ao se atingir uma quantidade de coeficientes suficiente para modelar a reverberação da sala, a adição de mais coeficientes não implicará ganhos de qualidade das estimativas dos sinais, como pode ser observado nas Figs. 3 e 5.

Além disso, a questão dos baixos valores das métricas SAR e SDR para os casos em que  $T_{60} \geq 150$  ms, que ocorre em outros algoritmos de separação, não ocorreu com os sinais estimados pelo algoritmo proposto, com os valores

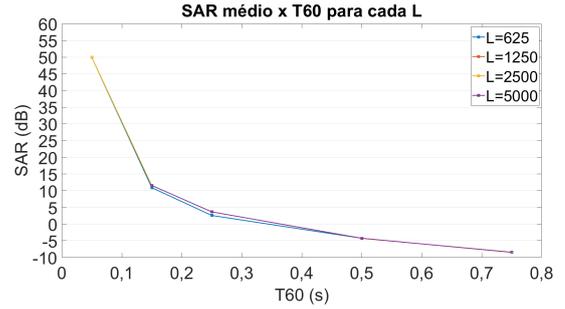


Fig. 6. SAR Médio das Estimativas obtidas pelo Modelo Tradicional

dessas duas métricas ficando bem próximos (menores que 0, 1 dB) nos casos em que  $L$  atingiu o número suficiente para o  $T_{60}$  considerado. Isso indica que os erros correspondentes à interferência do sinal de uma fonte na estimativa do sinal da outra estão bastante reduzidos, sendo esses erros causados principalmente pelo algoritmo de separação.

## V. CONCLUSÕES

O Algoritmo de Estimação de Fontes a partir de misturas convolutivas proposto neste artigo mostrou-se mais robusto para utilização em ambientes com alta reverberação. Como exemplo, em simulações realizadas em ambiente com tempo de reverberação de 250 ms, a reconstrução de sinais de voz pelo novo algoritmo apresentou razão sinal-interferência (SIR) média de 71,4 dB para  $L = 5000$ , enquanto que o arranjo tradicionalmente utilizado apresentou 28,1 dB. Os maiores valores obtidos para todas as métricas qualitativas avaliadas devem-se ao fato de que, idealmente, as estimativas dos sinais resultantes do algoritmo proposto não apresentam efeito de reverberação, ao contrário dos algoritmos tradicionais que estimam sinais provenientes das fontes com efeitos de filtragem resultantes do sistema de separação.

## AGRADECIMENTOS

Este trabalho é apoiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), Código de Financiamento 001, pelo CNPq, e pela FAPERJ.

## REFERÊNCIAS

- [1] T. Haubner A. Brendel and W. Kellermann. A unifying view on blind source separation of convolutive mixtures based on independent component analysis. *IEEE Transactions on Audio, Speech and Language Processing*, 71:816–830, January 2023.
- [2] P. B. Batalheiro. *Novos Algoritmos em Subbandas para Filtragem Adaptativa e Separação Cega de Fontes*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2009.
- [3] Herbert Buchner, Robert Aichner, and Walter Kellermann. Blind source separation for convolutive mixtures exploiting nongaussianity, nonwhiteness, and nonstationarity, 2003.
- [4] Herbert Buchner, Robert Aichner, and Walter Kellermann. *TRINICON-based Blind System Identification with Application to Multiple-Source Localization and Separation*, pages 101–147. Springer Netherlands, Dordrecht, 2007.
- [5] Paulo S. R. Diniz. *Introduction to Adaptive Filtering*, pages 1–8. Springer International Publishing, Cham, 2020.
- [6] P. Bofill E. Vincent, H. Sawada et al. First stereo audio source separation evaluation campaign: data, algorithms and results. *Proc. Int. Conf. on Independent Component Analysis and Signal Separation*, 2007.
- [7] B. L. R. P. Vasques. Limitações de algoritmos de separação cega de sinais de voz em ambientes reverberantes. Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2015.