

# Solução semifechada do conformador de feixe WCO-BMVDR para aparelhos auditivos biauriculares

Wilmer Lobato e Márcio Costa

**Resumo**— Este trabalho apresenta uma solução alternativa do conformador robusto WCO-BMVDR para aparelhos auditivos biauriculares. A metodologia proposta calcula o multiplicador de Lagrange de forma iterativa, mediante o método de Newton-Raphson, para obter a solução ótima. Simulações computacionais indicam a possibilidade de aumento da qualidade da fala em até 1,1 MOS-WPESQ e da razão sinal-interferência mais ruído biauricular em até 6,2 dB. Adicionalmente, possui um tempo de execução até 15,3 vezes menor que a forma convencional. O método proposto é especialmente efetivo para uma razão sinal-interferência de entrada entre 0 dB e 15 dB.

**Palavras-Chave**— *Aparelho auditivo, conformação de feixe, otimização de desempenho do pior caso.*

**Abstract**— This work presents an alternative solution to the WCO-BMVDR robust beamformer in hearing aid applications. The proposed methodology iteratively computes the Lagrange multiplier through the Newton-Raphson method to find the optimal solution. Computational simulations indicate significant improvements of up to 1.1 MOS-WPESQ in speech quality and up to 6.2 dB in terms of the binaural signal to interference plus noise ratio. In addition, it is up to 15.3 times faster than the conventional way. The proposed method is especially effective for input signal to interference ratios from 0 dB to 15 dB.

**Keywords**— *Hearing aids, beamforming, worst-case performance optimization.*

## I. INTRODUÇÃO

O conformador de feixe biauricular de mínima variância com resposta sem distorção (BMVDR, *binaural minimum variance distortionless response*) é uma das técnicas de redução de ruído mais exploradas em aplicações de aparelhos auditivos. Seu objetivo é a minimização da potência do ruído com a preservação do sinal de interesse [1].

O projeto do conformador BMVDR requer o conhecimento da matriz de coerência do ruído e do vetor de direção da fala [2]. Na prática, estes parâmetros devem ser estimados, introduzindo erros devido a diferentes fatores, tais como: desajuste de ganho e fase dos microfones; calibração imperfeita do arranjo; e acoplamento de microfones [3]. Diversos estudos têm demonstrado que o BMVDR é sensível a erros de estimação, acarretando na degradação do seu desempenho em termos de qualidade e inteligibilidade da fala processada [4].

A degradação de desempenho do BMVDR pode ser mitigada através de formulações robustas, como é o caso do conformador baseado na otimização do desempenho do pior caso (WCO, *worst case optimization*) [5], o qual minimiza a variância de saída impondo uma magnitude de resposta maior ou igual à unidade, dentro de uma região ao redor da localização esperada

Wilmer Lobato e Márcio Costa, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis-SC, E-mails: wilmer.lobato@gmail.com; costa@eel.ufsc.br. Este trabalho foi financiado pelo CNPq, projetos 162490/2015-1 e 315020/2018-0.

da fonte de interesse. Em [5], o problema de minimização é tratado como cone de segunda ordem (SOCP, *second-order cone program*), o qual é resolvido de forma eficiente através de métodos de pontos interiores. A adequação deste método para aparelhos auditivos biauriculares foi proposta em [6] [7], sendo chamada de conformador WCO-BMVDR. No entanto, o método de pontos interiores possui as seguintes desvantagens: alto custo computacional e ausência de solução analítica. A falta de uma solução analítica dificulta a análise do comportamento do conformador nas diferentes condições de aplicação.

A partir do exposto, o presente trabalho tem como objetivo a proposição de uma solução alternativa para o conformador robusto WCO-BMVDR proposto em [6]. As principais contribuições são: (a) derivação de uma solução semifechada do conformador WCO-BMVDR; e (b) simulações computacionais que demonstram os mesmos ganhos de qualidade e razão sinal ruído obtidos com o método convencional. A proposta apresentada reduz de forma significativa o tempo de processamento.

Ao longo do texto, letras minúsculas em itálico representam escalares. Letras minúsculas e maiúsculas em negrito denotam, respectivamente, vetores e matrizes. Os símbolos  $\{\cdot\}^T$  e  $\{\cdot\}^H$  são, respectivamente, o operador transposto e Hermitiano.

## II. FORMULAÇÃO DO PROBLEMA

Considere um par de aparelhos auditivos biauriculares que possui  $M$  microfones em cada orelha. Os sinais captados pelo arranjo de microfones são transformados para o domínio da frequência mediante a transformada de Fourier de tempo curto (STFT, *short-time Fourier transform*), de forma que:

$$y_{L,m}(p,k) = s(p,k)a_{L,m}(p,k) + v(p,k)b_{L,m}(p,k) + n(p,k), \quad (1)$$

em que  $y_{L,m}(p,k)$  é a STFT da fala contaminada para  $m = 0, 1, \dots, M-1$  microfones na orelha esquerda (L, *left*), para cada índice de quadro de tempo  $p$  e *bin* de frequência  $k$ ;  $s(p,k)$ ,  $v(p,k)$  e  $n(p,k)$  são, respectivamente, a STFT da fala desejada, do sinal de interferência e do ruído de fundo;  $a_{L,m}(p,k)$  e  $b_{L,m}(p,k)$  são as funções de transferência acústicas (ATF, *acoustic transfer function*), respectivamente, desde a fonte de fala e interferência, até o  $m$ -ésimo microfone da orelha esquerda. A forma vetorial dos sinais recebidos no lado esquerdo é expressa como:

$$\mathbf{y}_L(p,k) = s(p,k)\mathbf{a}_L(p,k) + v(p,k)\mathbf{b}_L(p,k) + \mathbf{n}(p,k), \quad (2)$$

em que  $\mathbf{y}_L(p,k)$  é o vetor da fala contaminada; e  $\mathbf{a}_L(p,k)$  e  $\mathbf{b}_L(p,k)$  são, respectivamente, os vetores de direção da fala e da interferência. Uma notação similar é aplicada para a orelha direita (R, *right*). Definem-se ainda os vetores empilhados  $\mathbf{y}(p,k) = [\mathbf{y}_L^T(p,k) \mathbf{y}_R^T(p,k)]^T$ ;  $\mathbf{a}(p,k) = [\mathbf{a}_L^T(p,k) \mathbf{a}_R^T(p,k)]^T$ , e  $\mathbf{b}(p,k) = [\mathbf{b}_L^T(p,k) \mathbf{b}_R^T(p,k)]^T$ , com dimensão  $2M \times 1$ .

Os sinais de saída (esquerdo e direito) do conformador biauricular são obtidos por filtragem linear dada por:

$$\begin{aligned} \mathbf{z}_L(p, k) &= \mathbf{w}_L^H(p, k) \mathbf{y}(p, k) \\ \mathbf{z}_R(p, k) &= \mathbf{w}_R^H(p, k) \mathbf{y}(p, k) \end{aligned} \quad (3)$$

em que  $\mathbf{w}_L(p, k)$  e  $\mathbf{w}_R(p, k)$  são os vetores de coeficientes do lado esquerdo e direito, que visam a reduzir o ruído total (interferência e ruído de fundo) [8].

#### A. Conformador BMVDR

O conformador de feixe biauricular de mínima variância sem distorção (BMVDR) proposto em [1], minimiza a potência total do ruído, e.g. interferência mais ruído de fundo, e preserva os componentes da fala original, aplicando uma restrição linear. A formulação do conformador BMVDR no domínio da frequência é dada por:

$$\begin{aligned} \min_{\mathbf{w}(p, k)} \mathbf{w}^H(p, k) \mathbf{\Phi}(p, k) \mathbf{w}(p, k) \\ \text{sujeito a } \mathbf{A}^H(p, k) \mathbf{w}(p, k) = \mathbf{1}_{2 \times 1} \end{aligned} \quad (4)$$

$$\text{em que: } \mathbf{A}(p, k) = \begin{bmatrix} \bar{\mathbf{a}}_L(p, k) & \mathbf{0}_{2M \times 1} \\ \mathbf{0}_{2M \times 1} & \bar{\mathbf{a}}_R(p, k) \end{bmatrix}; \quad (5)$$

$$\mathbf{\Phi}(p, k) = \begin{bmatrix} \mathbf{\Phi}_{vv}(p, k) & \mathbf{0}_{2M \times 2M} \\ \mathbf{0}_{2M \times 2M} & \mathbf{\Phi}_{vv}(p, k) \end{bmatrix}; \quad (6)$$

$\mathbf{w}(p, k) = [\mathbf{w}_L^T(p, k) \ \mathbf{w}_R^T(p, k)]^T$  é o vetor de coeficientes do conformador;  $\mathbf{\Phi}_{vv}(p, k)$  é a matriz de coerência do ruído total [1];  $\mathbf{1}_{2 \times 1}$  é um vetor coluna preenchido com 1, de dimensão  $2 \times 1$ ;  $\mathbf{0}_{i \times j}$  é uma matriz de zeros com dimensão  $i \times j$ ;  $\bar{\mathbf{a}}_L(k)$  e  $\bar{\mathbf{a}}_R(k)$  são, respectivamente, os vetores de direção da fala no lado esquerdo e direito, obtidos a partir das funções de transferência relativas aos microfones de referência frontais  $a_{L,0}(k)$  e  $a_{R,0}(k)$ , associados às orelhas esquerda e direita [2].

O conformador BMVDR definido na eq. (4) possui solução fechada em [1], dada por:

$$\mathbf{w}_{\text{BMVDR}}(p, k) = \frac{\mathbf{\Phi}^{-1}(p, k) \mathbf{A}(p, k) \mathbf{1}_{2 \times 1}}{\mathbf{A}^H(p, k) \mathbf{\Phi}^{-1}(p, k) \mathbf{A}(p, k)} \quad (7)$$

em que  $[\mathbf{A}^H(p, k) \mathbf{\Phi}^{-1}(p, k) \mathbf{A}(p, k)]^{-1} \mathbf{1}_{2 \times 1}$  é um vetor de dimensão  $2 \times 1$  que contém os fatores de escala dos filtros  $\mathbf{w}_L(p, k)$  e  $\mathbf{w}_R(p, k)$ , os quais não afetam a razão sinal-interferência mais ruído (SINR) de saída [9]. Este conformador é chamado de BMVDR ideal (I-BMVDR) devido ao conhecimento prévio de  $\mathbf{A}(p, k)$  e  $\mathbf{\Phi}(p, k)$ , que produz o filtro  $\mathbf{w}(p, k)$  ótimo [6].

Na prática,  $\mathbf{\Phi}(p, k)$  e  $\mathbf{A}(p, k)$  são desconhecidos e devem ser estimados utilizando detectores de fala, métodos de estimação recursiva e de mínima distorção, como em [10] [11] [12]. A implementação prática do conformador BMVDR utiliza as estimativas  $\hat{\mathbf{\Phi}}_{vv}(p, k)$ ,  $\hat{\mathbf{a}}_L(p, k)$  e  $\hat{\mathbf{a}}_R(p, k)$  que formam  $\hat{\mathbf{\Phi}}(p, k)$  e  $\hat{\mathbf{A}}(p, k)$ , os quais substituem  $\mathbf{\Phi}(p, k)$  e  $\mathbf{A}(p, k)$  ao longo das eq. (4) -(7). Este método é chamado de conformador BMVDR estimado (E-BMVDR) em [6].

De acordo com [5] e [9], o desempenho do BMVDR é significativamente impactado por erros em  $\hat{\mathbf{\Phi}}(p, k)$  e  $\hat{\mathbf{A}}(p, k)$ .

#### B. Método de otimização de desempenho do pior caso

Dentre os conformadores robustos mais utilizados, destaca-se o método de otimização de desempenho do pior caso, o qual é uma abordagem baseada na minimização da variância de saída do conformador, restrita pela imposição de uma resposta de

magnitude do arranjo maior que a unidade, dentro de uma região de incerteza desejada [5].

No contexto dos aparelhos auditivos biauriculares, considere os vetores  $\mathbf{c}_L(p, k)$  e  $\mathbf{c}_R(p, k)$  definidos como:  $\mathbf{c}_L(p, k) = \hat{\mathbf{a}}_L(p, k) + \mathbf{e}_L(p, k)$  e  $\mathbf{c}_R(p, k) = \hat{\mathbf{a}}_R(p, k) + \mathbf{e}_R(p, k)$ ; em que  $\mathbf{e}_L(p, k)$  e  $\mathbf{e}_R(p, k)$  são os vetores de erro, os quais são desconhecidos, porém possuem normas quadráticas máximas esperadas, representadas por  $\|\mathbf{e}_L(p, k)\| \leq \zeta_L(p, k)$  e  $\|\mathbf{e}_R(p, k)\| \leq \zeta_R(p, k)$ , em que  $\zeta_L(p, k)$  e  $\zeta_R(p, k)$  são os valores limites superiores. Note que este modelamento do erro corresponde geometricamente a uma hipersfera [5].

A formulação biauricular deste conformador robusto é dada pelo seguinte problema de minimização ( $p$  e  $k$  foram omitidos):

$$\begin{aligned} \min_{\mathbf{w}_L, \mathbf{w}_R} \mathbf{w}_L^H \hat{\mathbf{\Phi}}_{vv} \mathbf{w}_L + \mathbf{w}_R^H \hat{\mathbf{\Phi}}_{vv} \mathbf{w}_R \\ \text{sujeito a } |\mathbf{w}_L^H \mathbf{c}_L| \geq 1, |\mathbf{w}_R^H \mathbf{c}_R| \geq 1, \\ \forall \mathbf{c}_L, \mathbf{c}_R \in \Xi(\zeta_L, \zeta_R) \end{aligned} \quad (8)$$

em que  $\Xi(\zeta_L, \zeta_R) = \{ \mathbf{c}_L, \mathbf{c}_R \mid \mathbf{c}_L = \hat{\mathbf{a}}_L + \mathbf{e}_L, \mathbf{c}_R = \hat{\mathbf{a}}_R + \mathbf{e}_R, \|\mathbf{e}_L\| \leq \zeta_L, \|\mathbf{e}_R\| \leq \zeta_R \}$ . De acordo com [5], o problema de minimização na eq. (8) possui uma forma equivalente dada por:

$$\begin{aligned} \min_{\mathbf{w}_L, \mathbf{w}_R} \mathbf{w}_L^H \hat{\mathbf{\Phi}}_{vv} \mathbf{w}_L + \mathbf{w}_R^H \hat{\mathbf{\Phi}}_{vv} \mathbf{w}_R \\ \text{sujeito a } \zeta_L \|\mathbf{w}_L\| \leq \mathbf{w}_L^H \hat{\mathbf{a}}_L - 1, \text{Im}\{\mathbf{w}_L^H \hat{\mathbf{a}}_L\} = 0, \\ \zeta_R \|\mathbf{w}_R\| \leq \mathbf{w}_R^H \hat{\mathbf{a}}_R - 1, \text{Im}\{\mathbf{w}_R^H \hat{\mathbf{a}}_R\} = 0. \end{aligned} \quad (9)$$

O problema de minimização em (9) pode ser transformado em um programa de cone de segunda ordem (SOCP) e resolvido de forma eficiente através de métodos de pontos interiores [5]. Em [6], a equação (9) foi chamada de conformador biauricular de mínima variância sem distorção baseado no desempenho do pior caso (WCO-BMVDR).

Os parâmetros que controlam a robustez do conformador,  $\zeta_L$  e  $\zeta_R$ , estabelecem um compromisso entre a distorção da fala e a redução do ruído. Em [6], estimativas de  $\zeta_L$  e  $\zeta_R$  foram definidas através das variações da diferença de nível interauricular (ILD, *interaural level difference*) do sinal de fala para cada *bin* de frequência. A ILD é uma pista biauricular associada à diferença entre a potência acústica em ambas as orelhas [8].

Uma das limitações da solução de (9), obtida através do método de pontos interiores, é o alto custo computacional decorrente da norma quadrática. Além disso, esta solução não permite a análise do comportamento do conformador em relação à métricas objetivas, tal como a SINR de saída.

### III. SOLUÇÃO ALTERNATIVA DO WCO-BMVDR

Esta seção apresenta uma solução alternativa para o conformador WCO-BMVDR. Inicialmente, apresenta-se a versão do WCO-BMVDR com restrições de igualdade proposto em [5]. Em sequência, aplica-se a técnica dos multiplicadores de Lagrange, e analisam-se as características da solução obtida.

#### A. Restrições de igualdade no WCO-BMVDR

Em [5] foi provado que o conformador da eq. (9) é satisfeito usando restrições de igualdade ao invés de desigualdade. Como consequência, são dispensadas restrições relacionadas à parte imaginária. Desta forma, o conformador WCO-BMVDR (na orelha esquerda) pode ser definido como:

$$\min_{\mathbf{w}_L} \mathbf{w}_L^H \hat{\mathbf{\Phi}}_{vv} \mathbf{w}_L \text{ sujeito a } \xi_L^2 \|\mathbf{w}_L\|^2 = (\mathbf{w}_L^H \hat{\mathbf{a}}_L - 1)^2. \quad (10)$$

O problema de minimização em (10) pode ser resolvido aplicando-se o método dos multiplicadores de Lagrange. Primeiro, denota-se a função lagrangiana  $L(\mathbf{w}_L, \lambda_L)$ , dada por:

$$L(\mathbf{w}_L, \lambda_L) = \mathbf{w}_L^H \hat{\mathbf{\Phi}}_{vv} \mathbf{w}_L + \lambda_L [\xi_L^2 \|\mathbf{w}_L\|^2 - (\mathbf{w}_L^H \hat{\mathbf{a}}_L - 1)^2]. \quad (11)$$

Em sequência, calcula-se o gradiente de  $L(\mathbf{w}_L, \lambda_L)$  em relação a  $\mathbf{w}_L^*$ , resultando na solução:

$$\mathbf{w}_L = \frac{(\hat{\mathbf{\Phi}}_{vv} + \lambda_L \xi_L^2 \mathbf{I})^{-1} \hat{\mathbf{a}}_L}{\hat{\mathbf{a}}_L^H (\hat{\mathbf{\Phi}}_{vv} + \lambda_L \xi_L^2 \mathbf{I})^{-1} \hat{\mathbf{a}}_L - \lambda_L^{-1}}. \quad (12)$$

A partir da eq. (12), decompõe-se a matriz de coerência estimada do ruído total  $\hat{\mathbf{\Phi}}_{vv}$ , aplicando a decomposição de valores singulares (SVD, *singular value decomposition*) e obtendo  $\hat{\mathbf{\Phi}}_{vv} = \hat{\mathbf{U}}^H \hat{\mathbf{\Sigma}} \hat{\mathbf{U}}$ , onde  $\hat{\mathbf{U}}$  é a matriz de autovetores e  $\hat{\mathbf{\Sigma}}$  é a matriz diagonal que contém os autovalores da matriz  $\hat{\mathbf{\Phi}}_{vv}$  em ordem decrescente,  $\hat{\mathbf{\Sigma}} = \text{diag}(\hat{\sigma}_0, \hat{\sigma}_1, \dots, \hat{\sigma}_{M-1})$ , de forma que:

$$\mathbf{w}_L = \frac{\hat{\mathbf{U}}(\hat{\mathbf{\Sigma}} + \lambda_L \xi_L^2 \mathbf{I})^{-1} \hat{\mathbf{U}}^H \hat{\mathbf{a}}_L}{\hat{\mathbf{a}}_L^H \hat{\mathbf{U}}(\hat{\mathbf{\Sigma}} + \lambda_L \xi_L^2 \mathbf{I})^{-1} \hat{\mathbf{U}}^H \hat{\mathbf{a}}_L - \lambda_L^{-1}}. \quad (13)$$

A solução alternativa na eq. (13) será chamada de WCO-EQ-BMVDR (WCO-BMVDR com restrições de igualdade). Esta solução foi inicialmente apresentada em [13].

### B. Cálculo do multiplicador de Lagrange

Para encontrar o valor de  $\lambda_L$ , primeiro deve-se calcular o gradiente de  $L(\mathbf{w}_L, \lambda_L)$  em relação a  $\lambda_L$ , obtendo:

$$\frac{\partial L(\mathbf{w}_L, \lambda_L)}{\partial \lambda_L} = \xi_L^2 \mathbf{w}_L^H \mathbf{w}_L - (\mathbf{w}_L^H \hat{\mathbf{a}}_L - 1)^2 = 0. \quad (14)$$

Ao substituir o vetor solução, dado na eq. (13), em (14), após algumas manipulações matemáticas, tem-se:

$$\lambda_L^2 \xi_L^2 \hat{\mathbf{z}}_L^H (\hat{\mathbf{\Sigma}} + \lambda_L \xi_L^2 \mathbf{I})^{-2} \hat{\mathbf{z}}_L = 1, \quad (15)$$

onde  $\hat{\mathbf{z}}_L = \hat{\mathbf{U}}^H \hat{\mathbf{a}}_L$ . A eq. (15) pode ser expressa através da soma de  $M$  termos, representada por  $f(\lambda_L)$ , de forma que:

$$f(\lambda_L) = \sum_{m=0}^{M-1} \frac{|\hat{z}_{L,m}|^2}{(\xi_L + \hat{\sigma}_m / \lambda_L \xi_L)^2} - 1 = 0. \quad (16)$$

No entanto, a equação característica em (16) pode resultar em várias raízes reais ou complexas para  $\lambda_L$ .

### C. Prova de unicidade da solução

Analisando-se as condições limites da eq. (16), tem-se que:

$$\lim_{\lambda_L \rightarrow 0} f(\lambda_L) = -1 \quad (17)$$

$$\text{e } \lim_{\lambda_L \rightarrow \infty} f(\lambda_L) = \|\hat{\mathbf{a}}_L\|^2 / \xi_L^2 - 1. \quad (18)$$

Da eq. (17) verifica-se que  $\lambda_L = 0$  não é uma solução da equação característica. Para o caso de  $\lambda_L \rightarrow \infty$  a solução da eq. (18) seria  $\xi_L = \|\hat{\mathbf{a}}_L\|$ . Dessa forma, chega-se à conclusão de que a condição  $\xi_L < \|\hat{\mathbf{a}}_L\|$  é necessária para garantir uma solução real e positiva [5] [14].

Calculando-se a derivada da eq. (16) em relação a  $\lambda_L$ , chega-se em:

$$f'(\lambda_L) = \sum_{m=0}^{M-1} \frac{2\lambda_L \xi_L^2 |\hat{z}_{L,m}|^2 \hat{\sigma}_m}{(\lambda_L \xi_L^2 + \hat{\sigma}_m)^3}. \quad (19)$$

Na eq. (19),  $\xi_L$ ,  $\hat{\sigma}_m$  e  $|\hat{z}_{L,m}|$  são constantes reais e positivas e, portanto, conclui-se que a derivada  $f'(\lambda_L)$  também é real e positiva no intervalo  $\lambda_L \in (0, \infty)$ . Considerando estes fatos, e à luz da teoria do cálculo [15], verifica-se que  $f(\lambda_L)$  é monotonicamente crescente e, portanto, conclui-se que a equação característica apresentada em (16) possui apenas uma única solução  $\lambda_L$  real e positiva.

### D. Solução numérica: método de Newton-Raphson

Para encontrar a raiz da função  $f(\lambda_L)$ , um intervalo inicial de busca é definido como  $0 \leq \lambda_L \leq \lambda_{L,\text{sup}}$ , em que  $\lambda_{L,\text{sup}}$  corresponde ao limitante superior da função  $f(\lambda_{L,\text{sup}})$ , o qual é obtido quando  $\hat{\sigma}_m = \hat{\sigma}_{M-1}$ , de forma que:

$$f(\lambda_L)|_{\hat{\sigma}_m = \hat{\sigma}_{M-1}} \leq \frac{\|\hat{\mathbf{a}}_L\|^2}{(\xi_L + \hat{\sigma}_{M-1} / \lambda_L \xi_L)^2} - 1 = f(\lambda_{L,\text{sup}}). \quad (20)$$

A partir das eqs. (16) e (19), o intervalo inicial pode ser refinado utilizando-se o método de Newton-Raphson, conforme o pseudocódigo apresentado na Fig. 1.

```

1: PROCESSO NEWTON( f, f', λL,0 )
2: Raiz ← λL,0
3: ENQUANTO |f(λL,0)| > ε FAZER
4:     Raiz ← λL,0 - f(λL,0) / f'(λL,0)
5:     λL,0 ← Raiz
6: FIM ENQUANTO
7: DEVOLVE Raiz
8: FIM PROCESSO
    
```

Fig. 1. Pseudocódigo do método de Newton-Raphson. Adaptado de [16].

### E. Solução fechada para duas fontes pontuais

Considere o cenário acústico no qual o arranjo de microfones recebe sinais de duas fontes acústicas pontuais: uma fonte de fala e uma fonte de interferência. O ruído de fundo é desprezado, ou seja,  $\text{SNR} \rightarrow \infty$ . Dentro deste contexto, é possível simplificar a equação característica em (16) considerando somente os dois maiores autovalores da matriz de coerência do ruído,  $\hat{\sigma}_0$  e  $\hat{\sigma}_1$ , e desprezando os outros, isto é  $\hat{\sigma}_m = 0$  para  $m = 2, 3, \dots, M-1$ . Dessa forma, a eq. (16) resulta em:

$$\frac{|\hat{z}_{L,0}|^2}{\left(\xi_L + \frac{\hat{\sigma}_0}{\lambda_L \xi_L}\right)^2} + \frac{|\hat{z}_{L,1}|^2}{\left(\xi_L + \frac{\hat{\sigma}_1}{\lambda_L \xi_L}\right)^2} + \sum_{m=2}^{M-1} |\hat{z}_{L,m}|^2 - 1 = 0, \quad (21)$$

Na sequência, a eq. (21) é manipulada matematicamente, resultando no polinômio de quarta ordem  $g(\lambda_L)$  na forma padrão:

$$g(\lambda_L) = d_0 \lambda_L^4 + d_1 \lambda_L^3 + d_2 \lambda_L^2 + d_3 \lambda_L + d_4 = 0, \quad (22)$$

onde os coeficientes  $d_0$  e  $d_1$  são dados por:

$$d_0 = (\|\hat{\mathbf{a}}_L\|^2 - \xi_L^2) \xi_L^8, \quad (23)$$

$$d_1 = 2[(\|\hat{\mathbf{a}}_L\|^2 - \xi_L^2)(\hat{\sigma}_0 + \hat{\sigma}_1) - \hat{\sigma}_0 |\hat{z}_{L,0}|^2 - \hat{\sigma}_1 |\hat{z}_{L,1}|^2] \xi_L^6;$$

e os coeficientes  $d_2$ ,  $d_3$  e  $d_4$  são expressos como:

$$\begin{aligned}
 d_2 &= [(\|\hat{\mathbf{a}}_L\|^2 - \xi_L^2)(\hat{\sigma}_0^2 + 4\hat{\sigma}_0\hat{\sigma}_1 + \hat{\sigma}_1^2) \\
 &\quad - 4\hat{\sigma}_0\hat{\sigma}_1(|\hat{z}_{L,0}|^2 + |\hat{z}_{L,1}|^2) - \hat{\sigma}_0^2|\hat{z}_{L,0}|^2 - \hat{\sigma}_1^2|\hat{z}_{L,1}|^2] \xi_L^4, \quad (24) \\
 d_3 &= 2(\|\hat{\mathbf{a}}_L\|^2 - |\hat{z}_{L,0}|^2 - |\hat{z}_{L,1}|^2 - \xi_L^2)(\hat{\sigma}_0\hat{\sigma}_1 + \hat{\sigma}_0^2\hat{\sigma}_1)\xi_L^2, \\
 d_4 &= (\|\hat{\mathbf{a}}_L\|^2 - |\hat{z}_{L,0}|^2 - |\hat{z}_{L,1}|^2 - \xi_L^2)\hat{\sigma}_0^2\hat{\sigma}_1^2.
 \end{aligned}$$

O polinômio em (22), portanto, possui quatro raízes que podem ser encontradas de forma analítica. Segundo demonstrado na Seção III.C, o valor de  $\lambda_L$  obtido será real e positivo.

#### IV. SIMULAÇÕES COMPUTACIONAIS

Nesta seção são apresentados resultados de simulações de forma a demonstrar o desempenho do conformador proposto.

##### A. Descrição do experimento

Foi utilizado um banco de respostas impulsivas obtido de [17], o qual considera um par de aparelhos auditivos do tipo BTE (*behind-the-ear*) colocados nas orelhas de um manequim do tipo cabeça e torso. Cada aparelho auditivo possui  $M = 3$  microfones, operando de forma biauricular. Duas fontes acústicas pontuais foram consideradas: uma fonte de fala alocada na frente do usuário (azimute  $0^\circ$ ), e uma fonte de interferência alocada no lado direito (azimute  $45^\circ$ ). A distância radial entre a fonte e o usuário é de 0,8 metros e o ângulo de elevação é de  $0^\circ$ .

Foram considerados 6 sinais de fala com duração média de 3,8 segundos, obtidos de [18] e ruído tipo ICRA obtido de [19]. Ambos os sinais foram filtrados pelas respostas impulsivas para SIR de  $-10$  dB até 30 dB em passos de 5 dB, considerando 30 realizações. O ambiente acústico considerado foi uma câmara anecoica, com tempo de reverberação menor que 50 ms.

##### B. Processamento no domínio da frequência

Os sinais foram amostrados a 16 kHz e sua representação no domínio da frequência foi obtida através da transformada de Fourier de tempo curto considerando *frames* de 16 ms de duração e 256 *bins* de frequência, janela de Hanning de 128 pontos e 50% de sobreposição.

Comparações foram realizadas entre os seguintes métodos:

1) O conformador I-BMVDR, descrito na eq. (7), utiliza informação perfeita das matrizes  $\Phi$  e  $\mathbf{A}$ . Para tanto, assume-se o conhecimento individual da fala e ruído, utilizando-se o método da covariância amostral [4] para obter a matriz de coerência do ruído total  $\Phi$ . Em sequência, a matriz de coerência da fala é obtida usando a subtração de covariâncias [20]. Finalizando, os vetores de direção esquerdo e direito ( $\mathbf{A}$ ), são calculados utilizando-se o método de mínima distorção [12];

2) O conformador E-BMVDR, descrito também na eq. (7), utiliza informação aproximada  $\hat{\Phi}$  e  $\hat{\mathbf{A}}$ . Para tanto, é utilizado o detector de voz proposto em [10]. Na ausência de fala, aplica-se [4] para estimar  $\hat{\Phi}$ . Na presença de fala, estima-se a matriz de coerência da fala mediante [20]. Os vetores de direção relativa ( $\hat{\mathbf{A}}$ ) são obtidos da mesma forma que em [12];

3) O conformador WCO-EQ-BMVDR, apresentado na eq. (13), utiliza as mesmas estimativas  $\hat{\Phi}$  e  $\hat{\mathbf{A}}$  utilizadas para o conformador E-BMVDR e estimativas de  $\xi_L$  e  $\xi_R$  descritas em [6]. Os valores de  $\lambda_L$  e  $\lambda_R$  são calculados de forma iterativa

mediante o método numérico de Newton-Raphson apresentado na Fig. 1.

4) O conformador robusto alternativo apresentado em [21] chamado S-BMVDR (*steering-vector robust-based BMVDR*).

##### C. Medidas objetivas

Dois medidas objetivas de desempenho foram utilizadas: a) WPESQ (*Wideband Perceptual Evaluation of Speech Quality*, a qual utiliza a escala *Mean Opinion Score* (MOS), para avaliar a qualidade da fala; e a (b) razão sinal-interferência mais ruído biauricular (média de ambas as orelhas), para avaliação de redução de ruído. Segundo [22] e [23], incrementos menores que 0,2 MOS-WPESQ e 3 dB BSINR não são perceptíveis.

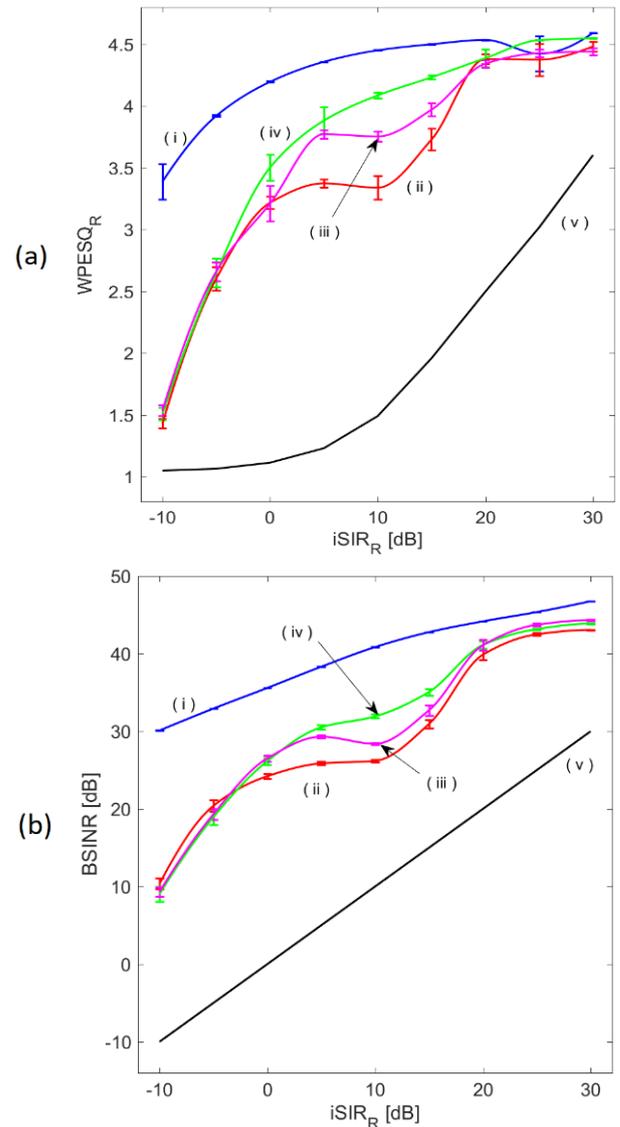


Fig. 2. Critérios objetivos MOS-WPESQ (a) e BSINR (b) para os conformadores: (i) I-BMVDR (azul); (ii) E-BMVDR (vermelho); (iii) S-BMVDR (magenta); (iv) WCO-EQ-BMVDR (verde). Sinais não-processados em preto (v). Assumiu-se uma SNR de entrada de 50 dB.

A Fig. 2 apresenta as métricas MOS-WPESQ e BSINR obtidas para os conformadores avaliados. Note-se que o desempenho do conformador I-BMVDR representa o limite

superior de desempenho, apresentando incrementos de até 3,3 MOS-WPESQ e 40 dB de BSINR em comparação ao sinal não processado. O conformador E-BMVDR atinge aumentos de 2,0 MOS-WPESQ e 25 dB de BSINR, apresentando desempenho adequado para aplicações com aparelhos auditivos em uma larga faixa de SIR de entrada, mesmo sujeito a erros de estimação. Os dois conformadores robustos avaliados, S-BMVDR e EQ-WCO-BMVDR apresentam desempenho superior ao E-BMVDR, porém o conformador WCO-EQ-BMVDR apresentou melhor desempenho, com aumentos adicionais de até 1,0 MOS-WPESQ e 6,2 dB BSINR, os quais são relevantes perceptualmente [22] [23]. Note-se ainda que a faixa na qual é apresentado aumento de desempenho,  $0 \leq \text{SIR} \leq 15$  dB é crucial em aplicações de aparelhos auditivos [8].

#### D. Tempo de execução

Nesta seção, é comparado o tempo de execução dos conformadores E-BMVDR, WCO-BMVDR em [6], e WCO-EQ-BMVDR (proposto). Foi utilizado o programa MatLab em um computador *Desktop* com processador Intel Xeon ES-2420 de 1,90 GHz. Foi calculada a média de 48 realizações, utilizando-se ruído tipo ICRA. Os resultados são apresentados na Tabela I.

TABELA I. TEMPO MÉDIO DE EXECUÇÃO DOS CONFORMADORES E-BMVDR, WCO-BMVDR E WCO-EQ-BMVDR.

	E-BMVDR	WCO-BMVDR	WCO-EQ-BMVDR
Tempo CPU	0,45 s	9,80 s	0,64 s

A Tabela I indica que o conformador WCO-EQ-BMVDR proposto possui um tempo de execução até 15,3 vezes menor que o do conformador WCO-BMVDR proposto em [6].

#### V. CONCLUSÃO

O presente trabalho apresentou uma forma alternativa para o cálculo dos coeficientes do conformador WCO-BMVDR utilizado para aplicações em aparelhos auditivos. Como resultado obteve-se uma função custo analítica dependente de dois parâmetros ( $\lambda_L$  e  $\lambda_R$ ), que podem ser obtidos através de um processo numérico de otimização (Newton-Raphson). O procedimento proposto resulta no mesmo desempenho do conformador WCO-BMVDR, em termos das métricas MOS-WPESQ e BSINR, com um tempo de execução até 15,3 vezes menor, o que apresenta uma grande vantagem em aplicações de aparelhos auditivos.

#### REFERÊNCIAS

- [1] E. Hadad, D. Marquardt, S. Doclo e S. Gannot, "Theoretical analysis of binaural transfer function MVDR beamformers with interference cue preservation constraints", *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2449-2464, 2015.
- [2] E. Hadad, S. Doclo e S. Gannot, "The binaural LCMV beamformer and its performance analysis", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 3, pp. 543-558, 2016.
- [3] H. Chen, "Robustness analysis of nearfield subband beamformers in the presence of microphone gain and phase errors", *Digital Signal Processing: A Review Journal*, vol. 23, no. 5, pp. 1712-1719, 2013.
- [4] B. Cauchi, I. Kodrasi, R. Rehr, S. Gerlach, A. Jukic, T. Gerkmann, S. Doclo e S. Goetze, "Combination of MVDR beamforming and single-channel spectral processing for enhancing noisy and reverberant speech", *EURASIP Journal on Advances in Signal Processing*, vol. 2015, no. 61, pp. 1-12, 2015.
- [5] S. A. Vorobyov, A. B. Gershman e Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem", *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 313-324, 2003.
- [6] W. Lobato e M. Costa, "Worst-case optimization robust-MVDR beamformer for stereo noise reduction in hearing aids", *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2224-2237, 2020.
- [7] W. Lobato e M. Costa, "Conformador de feixe robusto MVDR baseado na otimização de desempenho do pior caso para aparelhos auditivos biauriculares", em *XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrt)*, Florianópolis, 2020, pp. 1-5.
- [8] S. Doclo, S. Gannot, M. Moonen e S. Spriet, "Acoustic beamforming for hearing aid applications", em *Handbook on Array Processing and Sensor Networks*, John Wiley & Sons, 2008, pp. 269-302.
- [9] S. Gannot, E. Vincent, S. Markovich-Golan e A. Ozerov, "A consolidated perspective on multi-microphone speech enhancement and source separation", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 4, pp. 692-730, 2017.
- [10] J. Sohn, N. S. Kim e W. Sung, "A statistical model-based voice activity detection", *IEEE Signal Processing Letters*, vol. 6, no. 1, pp. 1-3, 1999.
- [11] J. Marin-Hurtado e D. Anderson, "Practical MWF-based noise-reduction methods for binaural hearing aids", em *XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA)*, Medellín, 2012, pp. 84-91.
- [12] M. Taseska e E. Habets, "Relative transfer function estimation exploiting instantaneous signals and the signal subspace", em *Proceedings of the European Signal Processing Conference (EUSPICO)*, 2015, pp. 404-408.
- [13] L. Zhang e W. Liu, "Robust beamforming for coherent signals based on the spatial-smoothing technique", *Signal Processing*, vol. 92, no. 11, pp. 2747-2758, 2012.
- [14] J. Li, P. Stoica e Z. Wang, "On robust Capon beamforming and diagonal loading", *IEEE Transactions on Signal Processing*, vol. 5, no. 7, pp. 1702-1715, 2003.
- [15] J. Stewart, *Single Variable Calculus*, Belmont: CA: Cengage Learning, 2010.
- [16] A. Gershman, Z.-Q. Luo e S. Shahbazpanahi, "Robust adaptive beamforming based on worst-case performance optimization", em *Robust adaptive beamforming*, Hoboken, NJ, John Wiley & Sons, 2006, pp. 49-90.
- [17] H. Kayser, S. D. Ewert, J. Anemuller, T. Rohdenburg, V. Hohmann e B. Kollmeier, "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses", *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1-10, 2009.
- [18] Y. Hu e P. Loizou, "Subject evaluation and comparison of speech enhancement algorithm", *Speech Communication*, vol. 49, pp. 588-601, 2007.
- [19] W. A. Dreschler, H. Verschuure, C. Ludvigsen e S. Westermann, "ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment", *Audiology*, vol. 40, pp. 148-157, 2001.
- [20] E. A. Habets e J. Benesty, "A perspective on frequency-domain beamformers in room acoustics", *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 3, pp. 947-960, 2012.
- [21] F. Shen, F. Chen e J. Song, "Robust adaptive beamforming based on steering vector estimation and covariance matrix reconstruction", *IEEE Communications Letters*, vol. 19, no. 9, pp. 1636-1639, 2015.
- [22] A. Servetti e J. C. D. Martin, "802.11 MAC protocol with selective error detection for speech transmission", em *Quality of Service in Multiservice IP Networks*, Springer, 2005, pp. 509-519.
- [23] D. McShefferty, W. M. Whitmer e M. A. Akeroyd, "The just meaningful difference in speech-to-noise ratio", *Trends in Hearing*, vol. 20, pp. 1-11, 2016.