

Filtro Multicanal de Wiener com Restrição Quadrática para Preservação de Pistas Biauriculares em Aparelhos Auditivos

Diego M. do Carmo, Ricardo A. Borsoi, Márcio H. Costa

Resumo— A redução de ruído em aparelhos auditivos biauriculares é um tema de grande interesse prático. Até o momento, a maioria das técnicas existentes são baseadas na inclusão de termos de penalização à função custo do filtro multicanal de Wiener (MWF). Neste trabalho, propõe-se um novo método de redução de ruído baseado na otimização da função custo do MWF sujeito a restrições quadráticas de preservação da informação espacial biauricular. Simulações computacionais indicam que o método proposto preserva a percepção biauricular da fala, assim como as de campos acústicos difusos ou originários de fontes pontuais. A redução de ruído obtida é comparável com a do MWF convencional.

Palavras-Chave— Aparelho auditivo, pistas biauriculares, redução de ruído.

Abstract— Noise reduction in binaural hearing aids is a problem of great practical interest. Methods available in the literature are based on adding penalty terms to the multichannel Wiener filter (MWF) cost function. In this work, we propose a new method of noise reduction based on the optimization of the MWF cost function subject to quadratic restrictions for the preservation of binaural spatial information. Computational simulations indicate that the proposed method preserves spatial characteristics of speech as well as of noise originated from either a pointwise source or diffuse field. The noise reduction performance is similar to the conventional MWF method.

Keywords— hearing aids, binaural cues, noise reduction.

I. INTRODUÇÃO

Aparelhos auditivos biauriculares são a opção mais avançada disponível para a compensação da capacidade auditiva de um indivíduo com perdas leves a moderadas em ambas as orelhas. Esses dispositivos utilizam um canal de transmissão de dados para compartilhar sinais e parâmetros entre os dois aparelhos [1]. Os aparelhos auditivos modernos embarcam algoritmos de redução de ruído com o objetivo de prover níveis adequados de conforto acústico e inteligibilidade da fala na presença de ruído interferente. Nesse quesito, técnicas biauriculares apresentam grande vantagem em comparação ao processamento monoauricular, pois além de maior eficiência na redução de ruído, permitem a preservação da percepção original do campo acústico, por meio de controle da distorção das chamadas *pistas biauriculares* dos sinais captados [2].

Três pistas acústicas biauriculares normalmente são consideradas nessa área de aplicação. A diferença de nível interaural

(*interaural level difference* – ILD) e a diferença de tempo interaural (*interaural time difference* – ITD) são determinantes na localização de fontes acústicas pontuais no plano horizontal [3]. Por outro lado, a coerência interaural (*interaural coherence* – IC), definida como a correlação espacial entre os sinais nas orelhas, é a informação determinante para percepção de campos acústicos difusos, em que não há uma localização espacial definida para a fonte sonora [2][4].

Uma das principais abordagens para redução de ruído biauricular é o filtro de Wiener multicanal (*multichannel Wiener filter* – MWF), o qual proporciona excelente desempenho em termos de redução da potência do ruído com baixa distorção da fala de interesse [5]. Porém, o MWF distorce a percepção espacial do ruído, o qual passa a ser percebido na mesma localização da fonte de interesse. Para contornar esse problema, estratégias recentes propuseram filtros biauriculares obtidos através da otimização sem restrições de funções custo contendo: i) a função custo original do MWF, a qual permite a redução de ruído e a manutenção das pistas acústicas biauriculares originais da fala de interesse, e ii) termos adicionais que penalizam a distorção das pistas biauriculares do ruído [4][5]. Exemplos incluem a preservação da ILD/ITD, da IC, e da ITF¹, que permitem a preservação da percepção espacial biauricular tanto de sinais gerados por campos acústicos originados de fontes pontuais quanto por sinais decorrentes de campos acústicos difusos [4]–[6]. Em especial, destaca-se a função custo do MWF-ILD-IC proposta recentemente em [8].

Apesar de propiciarem redução de ruído e adequada preservação da percepção espacial de componentes do cenário acústico, as técnicas baseadas em funções custo aumentadas possuem desvantagens importantes. Primeiro, elas requerem a seleção de um ou mais parâmetros de ponderação para estabelecer a importância de cada termo na função custo global. Em geral, esses parâmetros não possuem relação direta com variáveis físicas relacionadas ao problema, dificultando seu projeto. Segundo, as funções custo existentes são não-convexas, não-suaves e de difícil solução, o que leva a uma solução sem garantia de otimalidade e possivelmente de elevado custo computacional; o que configura uma significativa desvantagem para aplicações embarcadas de tempo real.

Uma estratégia computacionalmente simples, baseada em um problema de otimização com restrições lineares (chamada

Diego M. do Carmo, Ricardo A. Borsoi e Márcio H. Costa – Programa de Pós-Graduação em Engenharia Elétrica – Universidade Federal de Santa Catarina. E-mails: diego.carmoh@gmail.com, raborsoi@gmail.com, costa@eel.ufsc.br. Este trabalho foi parcialmente financiado por CAPES e CNPq (315020/2018-0).

¹ITF é acrônimo de *interaural transfer function* que significa função de transferência interauricular, em português.

de MWF-RTF²) foi proposta para a preservação de fontes interferentes pontuais [9]. Porém, seu adequado desempenho depende em grande parte da acurácia na estimação da função de transferência relativa entre a fonte interferente e os microfones do aparelho auditivo. Adicionalmente, este método não possibilita a preservação da percepção espacial de campos acústicos difusos, i.e., com coerência interaural menor que a unidade. Isso limita a sua aplicabilidade prática em comparação a estratégias como o MWF-ILD-IC [5][8].

Neste trabalho, propõe-se resolver o problema de redução de ruído com preservação de pistas binaurais utilizando um conjunto não-convexo de restrições quadráticas, problema conhecido como programação quadrática com restrições quadráticas (*quadratically constrained quadratic program – QCQP*), o qual é profundamente estudado na literatura especializada. Diferentemente dos métodos existentes, a distorção na percepção espacial do cenário acústico é controlada diretamente através das restrições, o que permite relacionar diretamente os parâmetros de projeto com as tolerâncias aceitáveis de distorção das pistas biauriculares. Simulações computacionais mostram que a solução proposta preserva as pistas biauriculares de campos acústicos difusos e decorrentes de fontes pontuais, sem perda significativa de desempenho (em termos de redução de ruído) em relação ao MWF.

Ao longo deste texto, letras minúsculas em itálico são variáveis escalares, maiúsculas em itálico são constantes; e letras maiúsculas e minúsculas em negrito representam, respectivamente, matrizes e vetores.

II. SINAIS E SISTEMA

Esta seção descreve os estágios de aquisição e processamento de sinais de um aparelho auditivo biauricular, assim como apresenta a definição das pistas biauriculares.

Os sinais de entrada são captados por $M/2$ microfones nos aparelhos auditivos da esquerda (*left – L*) e da direita (*right – R*), sendo M o número total de microfones do sistema biauricular. Considerando um segmento fixo de tempo indexado por λ , a representação na frequência (ω) do sinal contaminado $y_{\ell,m}(\lambda, \omega)$, captado pelo microfone $m \in \{1, 2, \dots, M/2\}$ no aparelho auditivo indexado por $\ell \in \{L, R\}$, é dada por

$$y_{\ell,m}(\lambda, \omega) = x_{\ell,m}(\lambda, \omega) + v_{\ell,m}(\lambda, \omega), \quad (1)$$

em que $x_{\ell,m}(\lambda, \omega)$ é a componente da fala e $v_{\ell,m}(\lambda, \omega)$ é a componente de ruído, que inclui sinais gerados por fontes interferentes pontuais, sinais associados a campos acústicos difusos e/ou ruído elétrico do microfone. Para simplificar a notação, os índices λ e ω serão omitidos no restante do texto.

O conjunto captado de amostras em um determinado instante é representado pelo vetor $\mathbf{y} \in \mathbb{C}^M$, conhecido por ambos os aparelhos auditivos devido a um canal de comunicação *full duplex*, definido como

$$\mathbf{y} = [\mathbf{y}_L^\top, \mathbf{y}_R^\top]^\top, \quad (2)$$

²RTF denomina *relative transfer function* ou função de transferência relativa, em português

em que $\mathbf{y}_\ell = [y_{\ell,1}, y_{\ell,2}, \dots, y_{\ell,M/2}]^\top$ é o vetor de amostras captadas no aparelho auditivo $\ell \in \{L, R\}$. O vetor \mathbf{y} pode ser representado pelo seguinte modelo aditivo [1][2][4]-[8]:

$$\mathbf{y} = \mathbf{x} + \mathbf{v}, \quad (3)$$

em que \mathbf{x} e \mathbf{v} são os vetores de amostras da fala e do ruído, definidos de maneira similar a \mathbf{y} em (2).

As matrizes de coerência do sinal contaminado, da fala e do ruído são definidas como $\Phi_{\mathbf{y}} = \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}$, $\Phi_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$, e $\Phi_{\mathbf{v}} = \mathbb{E}\{\mathbf{v}\mathbf{v}^H\}$, respectivamente, em que $\mathbb{E}\{\cdot\}$ representa o valor esperado em relação a \mathbf{x} e \mathbf{v} , e $(\cdot)^H$ é o operador Hermitiano. Definem-se também as matrizes $\Phi_{\mathbf{y}\mathbf{y}}$ e $\Phi_{\mathbf{x}\mathbf{x}}$, dadas por

$$\Phi_{\mathbf{y}\mathbf{y}} = \begin{bmatrix} \Phi_{\mathbf{y}} & \mathbf{0}_M \\ \mathbf{0}_M & \Phi_{\mathbf{y}} \end{bmatrix} \quad \text{e} \quad \Phi_{\mathbf{x}\mathbf{x}} = \begin{bmatrix} \Phi_{\mathbf{x}} & \mathbf{0}_M \\ \mathbf{0}_M & \Phi_{\mathbf{x}} \end{bmatrix}, \quad (4)$$

em que $\mathbf{0}_M$ é uma matriz quadrada de dimensão $M \times M$ composta por zeros.

Em geral, um microfone em cada aparelho auditivo é definido como *microfone de referência*, e o sinal captado nele é dado por $\mathbf{q}_\ell^\top \mathbf{y}_\ell = y_\ell = x_\ell + v_\ell$, em que \mathbf{q}_ℓ é um vetor de seleção de amostra que possui valor 1 na posição do microfone de referência e zero nas demais posições.

O sinal processado nos aparelhos auditivos é obtido da seguinte forma $z_L = \mathbf{w}_L^H \mathbf{y}$ e $z_R = \mathbf{w}_R^H \mathbf{y}$, em que \mathbf{w}_L e \mathbf{w}_R são vetores de dimensão $M \times M$ com elementos complexos e representam os filtros redução de ruído de cada aparelho auditivo.

Os filtros de redução de ruído \mathbf{w}_L e \mathbf{w}_R , e os vetores de seleção \mathbf{q}_L e \mathbf{q}_R são agrupados nos vetores \mathbf{w} e \mathbf{q} , i.e.,

$$\mathbf{w} = [\mathbf{w}_L^\top \mathbf{w}_R^\top]^\top, \quad \mathbf{q} = [\mathbf{q}_L^\top \mathbf{q}_R^\top]^\top. \quad (5)$$

Por fim, definem-se matematicamente as informações biauriculares da fonte de ruído. As ITFs dos sinais de ruído recebido e processado são dadas, respectivamente, por [5]

$$ITF_{\mathbf{v}}^{\text{in}} = \frac{\mathbf{q}_L^\top \Phi_{\mathbf{v}} \mathbf{q}_R}{\mathbf{q}_R^\top \Phi_{\mathbf{v}} \mathbf{q}_R}, \quad ITF_{\mathbf{v}}^{\text{ou}}(\mathbf{w}) = \frac{\mathbf{w}_L^H \Phi_{\mathbf{v}} \mathbf{w}_R}{\mathbf{w}_R^H \Phi_{\mathbf{v}} \mathbf{w}_R}. \quad (6)$$

As ILDs dos sinais de ruído recebido e processado são dadas, respectivamente, por [10]

$$ILD_{\mathbf{v}}^{\text{in}} = \frac{\mathbf{q}_L^\top \Phi_{\mathbf{v}} \mathbf{q}_L}{\mathbf{q}_R^\top \Phi_{\mathbf{v}} \mathbf{q}_R}, \quad ILD_{\mathbf{v}}^{\text{ou}}(\mathbf{w}) = \frac{\mathbf{w}_L^H \Phi_{\mathbf{v}} \mathbf{w}_L}{\mathbf{w}_R^H \Phi_{\mathbf{v}} \mathbf{w}_R}. \quad (7)$$

As ICs dos sinais de ruído recebido e processado são dadas, respectivamente, por [4]

$$IC_{\text{in}}^{\mathbf{v}} = \frac{\mathbf{q}_L^\top \Phi_{\mathbf{v}} \mathbf{q}_R}{\sqrt{\mathbf{q}_L^\top \Phi_{\mathbf{v}} \mathbf{q}_L \cdot \mathbf{q}_R^\top \Phi_{\mathbf{v}} \mathbf{q}_R}}, \quad (8)$$

$$IC_{\text{ou}}^{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{w}_L^H \Phi_{\mathbf{v}} \mathbf{w}_R}{\sqrt{\mathbf{w}_L^H \Phi_{\mathbf{v}} \mathbf{w}_L \cdot \mathbf{w}_R^H \Phi_{\mathbf{v}} \mathbf{w}_R}}. \quad (9)$$

III. MWF E PRESERVAÇÃO DE PISTAS BIAURICULARES

O MWF determina estimadores da componente de fala captadas pelos microfones de referência, i.e., x_L e x_R , a partir da função custo [5]

$$J_{\text{MWF}}(\mathbf{w}) = \mathbb{E}\{|x_L - \mathbf{w}_L^H \mathbf{y}|^2\} + \mathbb{E}\{|x_R - \mathbf{w}_R^H \mathbf{y}|^2\}, \quad (10)$$

em que $|\cdot|$ é a operação módulo. Ao manipular-se (10), considerando \mathbf{x} e \mathbf{v} descorrelacionados e representando o resultado em função de (4) e (5), obtém-se

$$J_{\text{MWF}}(\mathbf{w}) = \mathbf{w}^H \Phi_{yy} \mathbf{w} - \mathbf{w}^H \mathbf{p}_{xx} - \mathbf{p}_{xx}^H \mathbf{w} + p_{xx}, \quad (11)$$

em que $\mathbf{p}_{xx} = \Phi_{xx} \mathbf{q}$ é um vetor que contém as correlações entre as componentes de fala captadas e a fala no microfone de referência, i.e., $\mathbb{E}\{\mathbf{x}x_\ell^*\}$; e $p_{xx} = \mathbf{q}^T \Phi_{xx} \mathbf{q} = \mathbb{E}\{|x_L|^2\} + \mathbb{E}\{|x_R|^2\}$ é a potência da fala nos microfones de referência.

Apesar de prover adequado desempenho em termos de redução de ruído, psicoacusticamente a solução obtida a partir da minimização de (11) apresenta duas características importantes: i) o azimute percebido de uma fonte pontual interferente é, em condições ideais, igual ao azimute percebido da fonte de fala; e ii) uma fonte difusa passa a ser percebido como uma fonte pontual com azimute percebido igual ao da fonte de fala.

A fim de solucionar a distorção da percepção espacial causada pelo MWF foi proposto em [8] um termo de penalização para preservação de informação biauricular tanto de uma fonte pontual quanto uma fonte com campo acústico difuso, cuja definição matemática é dada por [8]

$$J(\mathbf{w}, \alpha) = J_{\text{MWF}}(\mathbf{w}) + \alpha_1 J_{\text{ILD}}(\mathbf{w}) + \alpha_2 J_{\text{IC}}(\mathbf{w}), \quad (12)$$

em que o conjunto de parâmetros $\alpha = \{\alpha_1, \alpha_2\}$ determina o compromisso entre a redução do ruído e a preservação de suas pistas acústicas; e J_{ILD} e J_{IC} são termos de penalização associados à preservação da ILD e da IC, nessa ordem, e são definidos, respectivamente, como [4][10]

$$J_{\text{ILD}}(\mathbf{w}) = [ILD_{\text{ou}}^v(\mathbf{w}) - ILD_{\text{in}}^v]^2, \quad (13)$$

$$J_{\text{IC}}(\mathbf{w}) = |IC_{\text{ou}}^v(\mathbf{w}) - IC_{\text{in}}^v|^2. \quad (14)$$

O projeto dos filtros de redução de ruído a partir de (12) é feito através do seguinte problema de otimização

$$\mathbf{w}_{\text{MWF-ILD-IC}}(\alpha) = \arg. \min_{\mathbf{w}} J(\mathbf{w}, \alpha). \quad (15)$$

Apesar de resultar em um desempenho de preservação biauricular adequado, o projeto dos filtros de redução de ruído a partir de (15) possui dois problemas práticos importantes: i) a dificuldade de projetar os parâmetros α_1 e α_2 , os quais não possuem relação matemática clara com o erro nas pistas acústicas de interesse; e ii) a impossibilidade de determinação de uma solução fechada ou computacionalmente eficiente, visto que o problema de otimização é não-convexo e não-suave.

IV. MWF COM RESTRIÇÃO QUADRÁTICA PARA PRESERVAÇÃO DE PISTAS BIAURICULARES

Neste trabalho propõe-se reformular a equação (15) como um problema com restrições quadráticas de preservação de informação espacial biauricular. Para tanto, reescreve-se (15) da seguinte forma

$$\mathbf{w}_{\text{MWF-RC}}(\delta) = \arg. \min_{\mathbf{w}} J_{\text{MWF}}(\mathbf{w}) \quad (16a)$$

$$\text{sujeito a (s.a.) } J_{\text{ILD}}(\mathbf{w}) \leq \delta_{\text{ILD}}^2, \quad (16b)$$

$$J_{\text{IC}}(\mathbf{w}) \leq \delta_{\text{IC}}^2, \quad (16c)$$

em que $\delta = \{\delta_{\text{ILD}}^2, \delta_{\text{IC}}^2\}$ representa o conjunto de parâmetros associados diretamente aos erros nas pistas acústicas e, por isso, têm uma relação com parâmetros psicoacústicos: erros na ILD estão associados à sensibilidade de variação do azimute da fonte sonora, enquanto erros na IC podem ser ajustados conforme o tipo de ruído (pontual ou difuso).

Para reformular (16) na forma de um problema com restrições quadráticas é necessário reescrever a região definida pelas restrições em (16b) e (16c) como a interseção de formas quadráticas em relação a \mathbf{w} , e.g., formas do tipo $\mathbf{w}^H \Phi \mathbf{w}$. Para tanto, nota-se que a restrição em (16b) é verdadeira se e somente se ela determina a mesma região factível de soluções da interseção das seguintes formas quadráticas

$$\mathbf{w}^H \Phi_1 \mathbf{w} \leq 0, \quad \mathbf{w}^H \Phi_2 \mathbf{w} \geq 0. \quad (17)$$

em que $\Phi_1 = \Phi_a(-\delta_{\text{ILD}})$ e $\Phi_2 = \Phi_a(+\delta_{\text{ILD}})$ para

$$\Phi_a(\delta_a) = \begin{bmatrix} \Phi_v & \mathbf{0}_M \\ \mathbf{0}_M & -(ILD_{\text{in}}^v - \delta_a) \cdot \Phi_v \end{bmatrix} \quad (18)$$

e $\delta_a \in \{-\delta_{\text{ILD}}, +\delta_{\text{ILD}}\}$.

Devido à dependência da J_{IC} em relação a IC_{ou}^v , a restrição (16c) depende da raiz quadrada de uma função dos coeficientes. Isso impossibilita que ela seja escrita como a interseção de formas quadráticas. Contudo, pode-se obter uma aproximação para região factível determinada por (16c), cujo qualidade depende da restrição (16b).

Para entender o desenvolvimento da aproximação obtida, define-se o argumento da operação módulo presente nessa restrição (16c) como $e_{\text{IC}} \in \mathbb{C}$, i.e.,

$$\begin{aligned} e_{\text{IC}}(\mathbf{w}) &= IC_{\text{ou}}^v(\mathbf{w}) - IC_{\text{in}}^v \\ &= \frac{\mathbf{w}_L^H \Phi_v \mathbf{w}_R}{\sqrt{\mathbf{w}_L^H \Phi_v \mathbf{w}_L \cdot \mathbf{w}_R^H \Phi_v \mathbf{w}_R}} - IC_{\text{in}}^v. \end{aligned} \quad (19)$$

Sendo e_{IC} um número complexo, pode-se escrever (16c) de forma equivalente, como

$$J_{\text{IC}}(\mathbf{w}) = \Re\{e_{\text{IC}}(\mathbf{w})\}^2 + \Im\{e_{\text{IC}}(\mathbf{w})\}^2 \leq \delta_{\text{IC}}^2, \quad (20)$$

em que $\Re\{\cdot\}$ e $\Im\{\cdot\}$ são, respectivamente, as partes reais e imaginárias de um número complexo.

Analisando-se (19), fica mais evidente a dificuldade para escrever (16c) como a interseção de formas quadráticas devido à presença da raiz quadrada da IC_{ou}^v . Para contornar esse problema, note que

$$\sqrt{\mathbf{w}_L^H \Phi_v \mathbf{w}_L \cdot \mathbf{w}_R^H \Phi_v \mathbf{w}_R} = \sqrt{ILD_{\text{ou}}^v(\mathbf{w})} \cdot \mathbf{w}_R^H \Phi_v \mathbf{w}_R. \quad (21)$$

Como a restrição em (16b) é atendida para valores suficientemente pequenos de δ_{ILD}^2 , i.e. valores que garantem a preservação das pistas acústicas de interesse, pode-se assumir que $ILD_{\text{ou}} \approx ILD_{\text{in}}$ e, portanto, assume-se que a raiz quadrada no lado direito de (21) pode ser aproximada da seguinte forma

$$\sqrt{ILD_{\text{ou}}^v(\mathbf{w})} \cdot \mathbf{w}_R^H \Phi_v \mathbf{w}_R \approx \sqrt{ILD_{\text{in}}^v} \cdot \mathbf{w}_R^H \Phi_v \mathbf{w}_R. \quad (22)$$

Portanto, considerando (21) e (22), pode-se aproximar (19) pelo valor \tilde{e}_{IC} , i.e.,

$$e_{\text{IC}} \approx \tilde{e}_{\text{IC}} = \frac{1}{\sqrt{ILD_{\text{in}}^v}} \cdot \frac{\mathbf{w}_L^H \Phi_v \mathbf{w}_R}{\mathbf{w}_R^H \Phi_v \mathbf{w}_R} - IC_{\text{in}}^v. \quad (23)$$

Portanto, considerando (23), pode-se aproximar a região de soluções definidas por (20) da seguinte forma

$$\tilde{J}_{IC}(\mathbf{w}) = \Re\{\tilde{\epsilon}_{IC}(\mathbf{w})\}^2 + \Im\{\tilde{\epsilon}_{IC}(\mathbf{w})\}^2 \leq \delta_{IC}^2. \quad (24)$$

Nota-se que a qualidade dessa aproximação depende diretamente do valor de δ_{IC}^2 : quanto menor esse valor melhor é a aproximação considerada.

Utilizando as variáveis adicionais ε_{\Re} e $\varepsilon_{\Im} \in \mathbb{R}$, a restrição (24) pode ser representada de forma equivalente como

$$\varepsilon_{\Re}^2 + \varepsilon_{\Im}^2 \leq \delta_{IC}^2, \quad (25)$$

$$-\varepsilon_{\Re} \leq \Re\{\tilde{\epsilon}_{IC}(\mathbf{w})\} \leq \varepsilon_{\Re}, \quad (26)$$

$$-\varepsilon_{\Im} \leq \Im\{\tilde{\epsilon}_{IC}(\mathbf{w})\} \leq \varepsilon_{\Im}. \quad (27)$$

Manipulando-se a restrição em (26) para os limiares superiores e inferiores, verifica-se que ela é verdadeira se e somente se ela determina a mesma região da interseção das seguintes formas quadráticas

$$\mathbf{w}^H \Phi_3 \mathbf{w} \geq 0, \quad \mathbf{w}^H \Phi_4 \mathbf{w} \leq 0, \quad (28)$$

em que $\Phi_3 = \Phi_b(-\varepsilon_{\Re})$ e $\Phi_4 = \Phi_b(+\varepsilon_{\Re})$ para

$$\Phi_b(\varepsilon_b) = \begin{bmatrix} \mathbf{0}_M & \Phi_v \\ \Phi_v^H & -2(\Re\{ITF_v^{in}\} - \varepsilon_b \sqrt{ILLD_v^{in}}) \cdot \Phi_v \end{bmatrix}, \quad (29)$$

para $\varepsilon_b \in \{-\varepsilon_{\Re}, +\varepsilon_{\Re}\}$. De forma equivalente, a restrição em (27) é dada por

$$\mathbf{w}^H \Phi_5 \mathbf{w} \geq 0, \quad \mathbf{w}^H \Phi_6 \mathbf{w} \leq 0, \quad (30)$$

em que $\Phi_5 = \Phi_c(-\varepsilon_{\Im})$ e $\Phi_6 = \Phi_c(+\varepsilon_{\Im})$ para

$$\Phi_c(\varepsilon_c) = - \begin{bmatrix} \mathbf{0}_M & -j\Phi_v \\ j\Phi_v^H & -2(\Im\{ITF_v^{in}\} - \varepsilon_c \sqrt{ILLD_v^{in}}) \cdot \Phi_v \end{bmatrix}, \quad (31)$$

e $\varepsilon_c \in \{-\varepsilon_{\Im}, +\varepsilon_{\Im}\}$.

Utilizando as restrições em (17), (25), (28) e (30), pode-se obter uma solução aproximada para o problema em (16) a partir da solução de um problema com restrições quadráticas de preservação biauricular (PBQ) dado por

$$\mathbf{w}_{MWF-PBQ1}(\delta) = \arg. \min_{\mathbf{w}, \varepsilon_{\Re}, \varepsilon_{\Im}} J_{MWF}(\mathbf{w}) \quad (32a)$$

$$\text{s.a. } \varepsilon_{\Re}^2 + \varepsilon_{\Im}^2 \leq \delta_{IC}^2, \quad (32b)$$

$$\mathbf{w}^H \Phi_q \mathbf{w} \leq 0, \quad q = 1, 2, \dots, 6. \quad (32c)$$

Pode-se reduzir a dimensionalidade do espaço de busca da solução fazendo $\varepsilon_{\Re}^2 = \varepsilon_{\Im}^2 = \delta_{IC}^2$. Com isso, obtém-se o seguinte problema

$$\mathbf{w}_{MWF-PBQ2}(\delta) = \arg. \min_{\mathbf{w}} J_{MWF}(\mathbf{w}) \quad (33a)$$

$$\text{s.a. } \mathbf{w}^H \Phi_q \mathbf{w} \leq 0, \quad q = 1, 2, \dots, 6. \quad (33b)$$

V. SIMULAÇÕES COMPUTACIONAIS

Nesta seção são descritos os experimentos de avaliação de desempenho realizados.

1) *Técnicas analisadas*: Três técnicas de redução de ruído foram analisadas: i) o MWF, definido pela minimização de (11); ii) a técnica proposta na forma do problema em (32) e iii) na forma do problema (33), denominadas respectivamente de MWF-PBQ1 e MWF-PBQ2. Para estas duas últimas técnicas foi utilizado $\delta_{ILD} = 10^{-2}/ILLD_{in}^v$ e $\delta_{IC} = 10^{-3}$. A normalização em δ_{ILD} compensa variações de potência em cada frequência.

2) *Cenário acústico*: Dois cenários acústicos, denominados C1 e C2 foram considerados nos experimentos, ambos contendo uma fonte pontual de fala e uma componente de ruído. No cenário acústico C1, a potência do ruído é majoritariamente decorrente de uma fonte pontual e minoritariamente por um campo difuso. No cenário acústico C2, a situação é invertida. Nos cenários acústicos C1 e C2, a fonte de fala foi assumida a uma distância de 3 metros do aparelho auditivo localizada no azimute $\theta_x = 15^\circ$. A fonte de interferência foi assumida também a 3 metros de distância e seu azimute (θ_u) variou de $\theta_u = -90^\circ$ (esquerda) até $\theta_u = +90^\circ$ (direita), espaçado de 15 graus. O campo acústico difuso foi assumido como sendo cilíndrico. As funções de transferência (*head related impulse responses* – HRIR) da base de dados descrita em [11] foram utilizadas para a simulação dos cenários acústicos. Essas HRIRs foram obtidas utilizando um par de aparelhos auditivos e um simulador de cabeça e tronco, localizados em ambiente anecoico. Cada aparelho auditivo possui 3 microfones, totalizando $M = 6$ microfones no sistema biauricular.

3) *Sinais de fala e ruído*: Os sinais de fala e de interferência foram simulados a partir de sinais de fala de aproximadamente 2 segundos de duração obtidos de [12]. A frequência de amostragem utilizada foi 16 kHz. As componentes de fala e de interferência nos microfones foram geradas convoluindo-se individualmente os sinais com as HRIRs associados aos respectivos azimutes de chegada. O ruído de campo difuso foi gerado utilizando o método em [13], cuja potência foi determinada utilizando o sinal descrito em [14]. O sinal contaminado foi gerado somando-se as componentes de fala, de interferência e ruído difuso. No cenário acústico C1, a relação sinal-ruído (*signal-to-noise ratio* – SNR) – entre o sinal desejado e o ruído difuso – foi estabelecida em 40 dB, e a relação sinal interferência (*signal-to-interference ratio* – SIR) foi estabelecida em 0 dB. No cenário acústico C2, a SNR foi estabelecida em 0 dB, e a SIR em 40 dB. Em ambos os cenários acústicos a SINR resultante foi de 0 dB.

4) *Processamento*: O sinal contaminado foi processado em janelas de 256 amostras (ou 16 ms). Essas amostras foram ponderadas por uma janela de análise (raiz quadrada da janela de Hanning) [15], e a representação tempo-frequência do resultado foi obtida utilizando a transformada de Fourier de tempo curto com 512 pontos. A sobreposição entre trechos de amostras adjacentes foi de 50%. As matrizes de coerência Φ_v e Φ_y foram calculadas diretamente da componente de ruído e do sinal contaminado, enquanto que a matriz de coerência da fala foi aproximada por $\Phi_x = \Phi_y - \Phi_v$. Os coeficientes \mathbf{w}_L e \mathbf{w}_R das técnicas com restrições foram obtidos utilizando a função 'fmincon'³ do Matlab.

³<https://www.mathworks.com/help/optim/ug/fmincon.html>

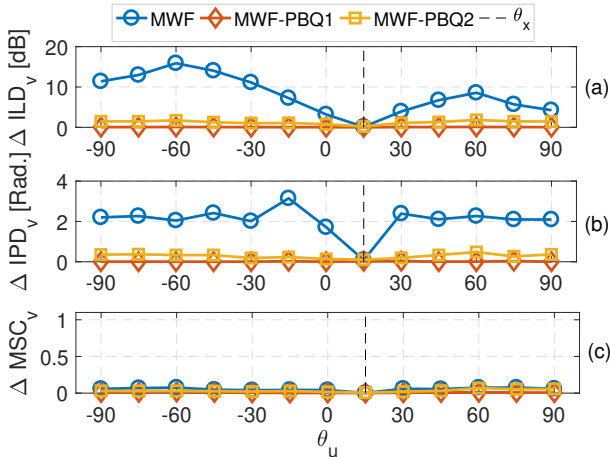


Fig. 1. Cenário acústico C1. Variação das pistas biauriculares do ruído: (a) ΔILD_v , (b) ΔIPD_v , (c) ΔMSC_v .

5) *Crítérios objetivos*: Os resultados da simulação foram avaliados de acordo com os erros na ILD (ΔILD_d), na IPD (ΔIPD_d), e o erro quadrático entre as MSCs de entrada e saída (ΔMSC_d) do ruído e da fala ($d \in \{v, x\}$) [4]. São apresentados resultados associados à diferença entre as SINRs globais de entrada e saída, representadas por ΔSINR , que mede a variação global SINR nas duas orelhas.

VI. RESULTADOS

A Figura 1 apresenta os erros de ILD, IPD e MSC do ruído em função de θ_u . O MWF distorce as pistas biauriculares do ruído em todas as situações em que $\theta_u \neq \theta_x$. As técnicas propostas mantêm a ΔILD_v ($< 0,5$ dB) e a ΔIPD_v ($< 0,2$ rad) próximas de zero para todos os valores de θ_u . Isso indica a preservação da percepção espacial original da componente de ruído. A ΔMSC_v se mantém próxima de zero para as três técnicas analisadas. Por outro lado, ambas as técnicas propostas distorcem as pistas biauriculares da fala. Entretanto, uma análise psicoacústica preliminar realizada pelos autores indica que os erros resultantes não são suficientes para afetar significativamente a percepção subjetiva do azimute da fala. Para as situações em que $\theta_u \neq \theta_x$, as três técnicas apresentam ΔSINR , aproximadamente constantes: $\Delta\text{SINR} = 19$ dB, para o MWF; $\Delta\text{SINR} = 24,2$ dB, para o MWF-PBQ1; e $\Delta\text{SINR} = 18,4$ dB, para o MWF-PBQ2. A Figura 2 apresenta os erros na ILD, IPD e MSC do ruído total em função de θ_u para o cenário acústico C2. O azimute da fonte interferente não afeta significativamente o MSC. O MWF apresenta as maiores distorções nas pistas acústicas, pois faz com que MSC_v seja próximo de 1. As técnicas propostas apresentam ΔMSC_v próximas de zero, indicando a preservação da percepção espacial original do campo acústico difuso dominante. Para todos os valores de θ_u , obteve-se a seguinte SINR para as três técnicas analisadas: 10,6 dB, para o MWF; 8,4 dB, para o MWF-PBQ1; e 8,2 dB, para o MWF-PBQ2.

VII. CONCLUSÕES

Este trabalho apresentou um novo método para redução de ruído e preservação das pistas biauriculares baseado no MWF. Assume-se a presença de uma fonte pontual de fala de interesse e de uma componente de ruído genérica (pontual/difusa).

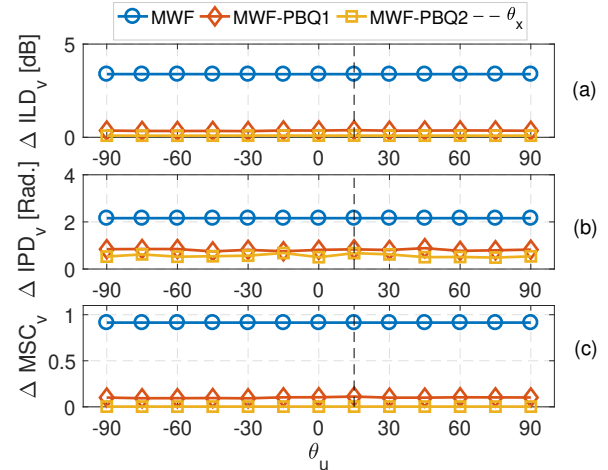


Fig. 2. Cenário acústico C2. Variação das pistas biauriculares do ruído: (a) ΔILD_v , (b) ΔIPD_v , (c) ΔMSC_v .

Os resultados obtidos indicam que a técnica proposta é capaz de preservar as pistas biauriculares tanto da fala quanto do ruído, mantendo aproximadamente a mesma capacidade de redução de ruído do MWF.

REFERÊNCIAS

- [1] V. Hamacher *et al.*, “Binaural signal processing in hearing aids: Technologies and algorithms”, *Advances in digital speech transmission*, v. 14, p. 401–429, 2008.
- [2] S. Doclo *et al.*, “Binaural speech processing with application to hearing devices”, *Audio Source Separation and Speech Enhancement*, cap. 18, p. 413–442, 2018.
- [3] J. Blauert, (1997), *The psychophysics of human sound localization*. em *Spatial Hearing*. MIT Press.
- [4] V. H. Marquardt e S. Doclo, “Interaural coherence preservation in multichannel Wiener filtering-based noise reduction for binaural hearing aids”, *IEEE Trans. Audio Speech Lang. Process.*, v. 23, n. 12, p. 2162–2176, 2015.
- [5] B. Cornelis *et al.*, “Theoretical analysis of binaural multimicrophone noise reduction techniques”, *IEEE Trans. Audio Speech Lang. Process.*, v. 18, n. 2, p. 342–355, 2009.
- [6] F. P. Itturiet e M. H. Costa, “Perceptually relevant preservation of interaural time differences in binaural hearing aids”, *IEEE Trans. Audio Speech Lang. Process.*, v. 27, n. 4, p. 753–764, 2019.
- [7] J. Werner e M. H. Costa, “A noise-reduction method with coherence enhancement for binaural hearing aids”, *J. Commun. Inf. Syst.*, v. 35, n. 1, p. 338–348, 2020.
- [8] D. Carmo, M. H. Costa e R. Borsoi, “Proposta de uma função custo para a preservação espacial de fontes sonoras em aparelhos auditivos biauriculares”, em *Simp. Bras. Telecom. Proc. Sinais*, 2020, p. 1–5.
- [9] D. Marquardt *et al.*, “Theoretical analysis of linearly constrained multichannel Wiener filtering algorithms for combined noise reduction and binaural cue preservation in binaural hearing aids”, *IEEE/ACM Trans. Audio Speech Lang. Process.*, v. 23, n. 12, p. 2384–2397, 2015.
- [10] M. H. Costa e P. A. Naylor, “ILD preservation in the multichannel Wiener filter for binaural hearing aid applications”, em *Europ. Signal Process. Conf.*, 2014, p. 636–640.
- [11] H. Kayser *et al.*, “Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses”, *EURASIP J. Adv. Signal Process.*, v. 2009, p. 1–10, 2009.
- [12] G. E. Henter *et al.*, “Repeated harvard sentence prompts corpus version 0.5”, 2014, disponível em: <https://datashare.is.ed.ac.uk/handle/10283/561?show=full>.
- [13] E. A. P. Habets, I. Cohen e S. Gannot, “Generating nonstationary multisensor signals under a spatial coherence constraint”, *J. Acoust. Soc. Am.*, v. 124, n. 5, p. 2911–2917, 2008.
- [14] W. A. Dreschler *et al.*, “ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment”, *Audiol.*, v. 40, n. 3, p. 148–157, 2001.
- [15] R. Crochiere, “A weighted overlap-add method of short-time Fourier analysis/synthesis”, *IEEE Trans. Audio Speech Lang. Process.*, vol. 28, no. 1, pp. 99–102, 28.