



# A Convex Framework for the Multichannel Wiener Filter with Preservation of Binaural in Hearing Aids

---

Diego Carmo, Ricardo Borsoi and Marcio Costa

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

October 6, 2022

# Uma Abordagem Convexa para o Filtro de Wiener Multicanal com Preservação de Pistas Biauriculares em Aparelhos Auditivos

Diego M. do Carmo<sup>1</sup>, Ricardo A. Borsoi<sup>2</sup> and Márcio H. Costa<sup>1</sup>

<sup>1</sup> Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, Brasil.

<sup>2</sup> Centre de Recherche en Automatique de Nancy, Université de Lorraine, CNRS, Nancy, France.

**Resumo**— A redução de ruído é uma etapa fundamental no tratamento de sinais em aparelhos auditivos biauriculares. De forma geral, esse procedimento é caracterizado por um algoritmo de otimização que minimiza uma função custo composta por uma combinação do filtro multicanal de Wiener e de restrições ou termos de penalização que visam a garantir a preservação da informação espacial biauricular. Porém, a função custo é não-convexa, o que faz com que sua solução demande um alto custo computacional. Neste trabalho propomos uma abordagem computacionalmente eficiente para a redução de ruído com preservação da informação espacial biauricular, baseada em uma relaxação convexa do problema de otimização original. Experimentos indicam que a técnica proposta atinge desempenho comparável ao de algoritmos do estado da arte em termos de redução de ruído e preservação da percepção da localização espacial de fontes pontuais, mas com um tempo de processamento extremamente reduzido.

**Palavras chave**— Aparelho auditivo, pistas biauriculares, redução de ruído.

## I. INTRODUÇÃO

Aparelhos auditivos biauriculares têm como objetivo a compensação de perdas auditivas leves a moderadas, através do compartilhamento de sinais e parâmetros entre os aparelhos em ambas as orelhas [1]. Esses dispositivos utilizam algoritmos de redução de ruído de modo a melhorar o conforto acústico e a inteligibilidade da fala na presença de ruído interferente. Métodos biauriculares de redução de ruído possuem vantagens significativas sobre os monoauriculares, incluindo a possibilidade de preservar a percepção original do campo acústico, através da limitação da distorção das chamadas *pistas biauriculares* dos sinais processados [2]. As pistas biauriculares mais utilizadas consistem na diferença de nível interaural (ILD), na diferença de tempo interaural (ITD), e na coerência interaural (IC) [2, 3, 4]. A ILD e a ITD são determinantes na localização de fontes acústicas pontuais no plano horizontal [3], enquanto a IC é determinante na percepção de campos acústicos difusos (em que a fonte sonora não possui localização espacial definida) [2, 4].

O filtro de Wiener multicanal (MWF) é um dos principais

métodos de redução de ruído em aparelhos auditivos biauriculares. O MWF permite a obtenção de um alto desempenho em termos de redução de ruído, com baixa distorção do sinal de interesse [5]. Entretanto, o MWF distorce a percepção espacial do ruído interferente, levando o usuário a percebê-lo na mesma posição do sinal de interesse. Trabalhos recentes propuseram estratégias que contornam essa limitação calculando filtros biauriculares através da otimização de funções custo compostas pela soma da função custo originalmente utilizada pelo MWF com termos que têm por objetivo penalizar a distorção espacial do ruído [4, 5, 6]. Abordagens existentes consideram a preservação de pistas biauriculares através de funções custo baseadas na ILD, ITD, função de transferência interauricular, e IC [4, 7]. Recentemente, foi demonstrado em [8] e [9] que a preservação da ILD e da IC é o estado da arte em termos de preservação de cenários acústicos contendo tanto fontes pontuais quanto difusas.

Técnicas de redução de ruído baseadas em funções custo aumentadas permitem redução de ruído e a preservação espacial em níveis satisfatórios. Porém, requerem a otimização de funções custo não-convexas e não-suaves, cuja solução é computacionalmente complexa. Dessa forma, esse procedimento requer um custo computacional elevado, o que é incompatível com as limitações inerentes à sua aplicação em sistemas embarcados e em tempo real, como é o caso de aparelhos auditivos. Recentemente a função custo do método MWF-ILD-IC foi reformulada na forma de um problema de otimização quadrático com restrições não-convexas (QCQP) [9]. Apesar disto, esse tipo de problema também é de difícil solução, pertencendo à classe de problemas conhecidos como *NP-hard*, o que na prática também exige um custo computacional elevado [10].

A partir dos problemas descritos, este trabalho propõe uma abordagem convexa para a solução do problema de redução de ruído com preservação de pistas biauriculares em cenários acústicos contendo fontes pontuais e campos difusos. Baseando-se na formulação do MWF-ILD-IC proposta em [9], a relaxação dual do problema de otimização com restrições é derivada, a qual é sempre côncava. De modo a lidar com a não-linearidade da função custo, esta é convertida na forma de um problema de programação semide-

finida (SDP), o qual é convexo e pode ser resolvido em tempo polinomial. Apesar da formulação proposta se basear em uma relaxação (a qual não garante a obtenção de um ótimo global), simulações computacionais mostram que a abordagem proposta obtém desempenho satisfatório tanto em termos de redução de ruído quanto na preservação da localização de fontes acústicas pontuais. No caso de ruído difuso, a preservação das características espaciais ocorre de forma moderada. Como resultado, o tempo de cálculo dos coeficientes do filtro de redução de ruído diminuem consideravelmente (em torno de duas ordens de grandeza) em relação à solução do problema não-convexo original.

## II. SINAIS E SISTEMA

Esta seção descreve os estágios de aquisição e processamento de sinais de um aparelho auditivo biauricular, como também define e caracteriza as pistas espaciais biauriculares.

Os sinais acústicos são captados por um arranjo de  $M/2$  microfones em cada um dos aparelhos auditivos. O número total de microfones no sistema biauricular é  $M$ . A representação tempo-frequência de um segmento de fala contaminada captada pelo microfone  $m$  no aparelhos auditivo indexado por  $\ell$  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 o sinal captado de fala e  $v_{\ell,m}(\lambda, \omega)$  é o ruído aditivo. As amostras captadas pelos  $M$  microfones em um determinado instante de tempo são conhecidas por ambos os aparelhos auditivos através de um canal de comunicação *full duplex*. A representação desse conjunto de amostras é dada por:

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

em que vetor  $\mathbf{y}_\ell = [y_{\ell,1}, y_{\ell,2}, \dots, y_{\ell,M/2}]^\top$  contém as amostras captadas no lado  $\ell$ . Destaca-se ainda que o vetor  $\mathbf{y}$  pode ser caracterizado da seguinte forma [1, 2, 4, 8]:

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

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

Denotam-se por  $\Phi_x = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ ,  $\Phi_v = \mathbb{E}\{\mathbf{v}\mathbf{v}^H\}$  e  $\Phi_y = \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}$  as matrizes de coerência da fala, do ruído e da fala contaminada, respectivamente; onde  $(\cdot)^H$  representa o operador Hermitiano (conjugado transposto) e  $\mathbb{E}\{\cdot\}$  é o operador estatístico valor esperado.

Um microfone em cada aparelho auditivo é selecionado como *microfone de referência*. A fala contaminada captada

nesse microfone é dada por  $\mathbf{q}_\ell^\top \mathbf{y}_\ell = y_\ell = x_\ell + v_\ell$ , em que  $\mathbf{q}_\ell$  é um vetor de seleção de microfone, que contém o valor 1 na posição do microfone de referência e 0 nas demais posições.

O sinal processado nos aparelhos auditivos, no domínio tempo-frequência, é determinado por  $z_L = \mathbf{w}_L^H \mathbf{y}$  e  $z_R = \mathbf{w}_R^H \mathbf{y}$ , em que  $\mathbf{w}_L$  e  $\mathbf{w}_R$  caracterizam os filtros de redução de ruído de dimensão  $M \times 1$ . Os vetores  $\mathbf{w}_L$  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.,

$$\mathbf{w} = [\mathbf{w}_L^\top \ \mathbf{w}_R^\top]^\top, \quad \mathbf{q} = [\mathbf{q}_L^\top \ \mathbf{q}_R^\top]^\top. \quad (4)$$

Por fim, definimos as pistas biauriculares utilizadas neste trabalho. Elas são definidas para a um sinal genérico  $d \in \{n, u\}$ , em que in e ou representam as pistas de entrada e de saída [11]:

$$\text{ILD}_d^{\text{in}} = \frac{\mathbf{q}_L^\top \Phi_d \mathbf{q}_L}{\mathbf{q}_R^\top \Phi_d \mathbf{q}_R}, \quad \text{ILD}_d^{\text{ou}}(\mathbf{w}) = \frac{\mathbf{w}_L^H \Phi_d \mathbf{w}_L}{\mathbf{w}_R^H \Phi_d \mathbf{w}_R}, \quad (5)$$

e para a IC [4]:

$$\text{IC}_{\text{in}}^d = \frac{\mathbf{q}_L^\top \Phi_d \mathbf{q}_R}{\sqrt{\mathbf{q}_L^\top \Phi_d \mathbf{q}_L \cdot \mathbf{q}_R^\top \Phi_d \mathbf{q}_R}}, \quad (6)$$

$$\text{IC}_{\text{ou}}^d(\mathbf{w}) = \frac{\mathbf{w}_L^H \Phi_d \mathbf{w}_R}{\sqrt{\mathbf{w}_L^H \Phi_d \mathbf{w}_L \cdot \mathbf{w}_R^H \Phi_d \mathbf{w}_R}}. \quad (7)$$

## III. FILTRO DE WIENER MULTICANAL COM PRESERVAÇÃO DE PISTAS BIAURICULARES

Assumindo que  $\mathbf{x}$  e  $\mathbf{v}$  são descorrelacionados, o MWF calcula um estimador do sinal de fala adquirido pelos microfones de referência ( $x_L$  e  $x_R$ ) minimizando a função custo [5]:

$$J_{\text{MWF}}(\mathbf{w}) = \mathbf{w}^H \Phi_{yy} \mathbf{w} - \mathbf{w}^H \mathbf{p}_{xx} - \mathbf{p}_{xx}^H \mathbf{w} + p_{xx}, \quad (8)$$

onde  $\mathbf{p}_{xx} = \Phi_{xx} \mathbf{q}$ ,  $p_{xx} = \mathbf{q}^\top \Phi_{xx} \mathbf{q}$ , e as matrizes  $\Phi_{yy}$  e  $\Phi_{xx}$  são definidas como

$$\Phi_{yy} = \begin{bmatrix} \Phi_y & \mathbf{0}_M \\ \mathbf{0}_M & \Phi_y \end{bmatrix} \quad \text{e} \quad \Phi_{xx} = \begin{bmatrix} \Phi_x & \mathbf{0}_M \\ \mathbf{0}_M & \Phi_x \end{bmatrix}, \quad (9)$$

onde  $\mathbf{0}_M$  é uma matriz de zeros de dimensão  $M \times M$ . O MWF provê uma adequada redução da potência do ruído interferente. Porém, o MWF altera a percepção da fonte interferente, deslocando o azimuth percebido para a mesma posição da fala. Esta distorção também ocorre em campos acústicos difusos.

Para mitigar as distorções causadas pelo MWF na percepção espacial de um sinal proveniente de uma fonte pontual interferente ou de um campo acústico difuso, uma função custo aumentada foi proposta em [8]:

$$J(\mathbf{w}, \alpha) = J_{\text{MWF}}(\mathbf{w}) + \alpha_1 J_{\text{ILD}}(\mathbf{w}) + \alpha_2 J_{\text{IC}}(\mathbf{w}), \quad (10)$$

onde os termos  $J_{ILD}$  e  $J_{IC}$  penalizam, respectivamente, distorções na ILD e na IC do sinal interferente processado, e são dados por [4, 11]:

$$J_{ILD}(\mathbf{w}) = [\text{ILD}_{\text{ou}}^d(\mathbf{w}) - \text{ILD}_{\text{in}}^d]^2, \quad (11)$$

$$J_{IC}(\mathbf{w}) = |\text{IC}_{\text{ou}}^d(\mathbf{w}) - \text{IC}_{\text{in}}^d|^2. \quad (12)$$

em que  $d$  representa a componente de ruído a ser preservada (pontual ou difusa). Os parâmetros  $\alpha = \{\alpha_1, \alpha_2\}$  controlam o compromisso entre os objetivos de redução de ruído e preservação das pistas binauriculares na função custo; os filtros são obtidos através da minimização de  $J(\mathbf{w}, \alpha)$ :

$$\mathbf{w}_{\text{MWF-ILD-IC}}(\alpha) = \arg. \min_{\mathbf{w}} J(\mathbf{w}, \alpha). \quad (13)$$

A função custo em (10) é considerada o estado da arte em termos de preservação espacial de sinais provenientes de fontes pontuais e de campos acústicos difusos [8]. Porém, uma desvantagem inerente ao uso de funções custo aumentadas baseadas no MWF para redução de ruído, como em [8], consiste na necessidade de ajuste dos parâmetros de ponderação  $\alpha_i$ , o que não é intuitivo. Além disso, resultam em problemas de otimização cuja solução é computacionalmente complexa. Em [9], propomos uma alternativa para este problema de otimização baseada em restrições quadráticas. Esse algoritmo representa o estado da arte para preservação de um campo acústico genérico (difuso ou decorrente de uma fonte pontual), conforme demonstrado experimentalmente em [9] e [8]. A derivação da proposta em [9] pode ser realizada partindo-se do problema de otimização apresentado em (10) na forma de um problema com restrições:

$$\mathbf{w}_{\text{MWF-RC}}(\delta) = \arg. \min_{\mathbf{w}} J_{\text{MWF}}(\mathbf{w}) \quad (14a)$$

$$\text{sujeito a (s.a.) } J_{ILD}(\mathbf{w}) \leq \delta_{ILD}^2, \quad (14b)$$

$$J_{IC}(\mathbf{w}) \leq \delta_{IC}^2, \quad (14c)$$

onde os valores  $\delta = \{\delta_{ILD}^2, \delta_{IC}^2\}$  possuem uma relação direta com parâmetros psicoacústicos do problema, estando diretamente associados à variação do azimute e à característica espacial do ruído (pontual ou difuso).

Manipulando-se as restrições (14b) e (14c), o problema (14) foi convertido na forma de um problema com restrições quadráticas [9]. Mais precisamente, se o valor de  $\delta_{ILD}$  for suficientemente pequeno de modo que seja possível assumir que  $\text{ILD}_{\text{ou}} \approx \text{ILD}_{\text{in}}$  (i.e., a preservação da ILD é garantida), então utilizando algumas aproximações é possível re-escrever as restrições (14b) e (14c) como a intersecção de restrições quadráticas (não necessariamente convexas). Isso torna possível escrever (14) na forma do seguinte problema

de otimização [9]:

$$\mathbf{w}_{\text{QCQP-MWF-ILD-IC}}(\delta) = \arg. \min_{\mathbf{w}} J_{\text{MWF}}(\mathbf{w}) \quad (15a)$$

$$\text{s.a. } (-1)^q \mathbf{w}^H \Phi_q \mathbf{w} \leq 0, \quad q = 1, 2, \dots, 6. \quad (15b)$$

onde:

$$\Phi_1 = \Phi_a(-\delta_{ILD}), \quad \Phi_2 = \Phi_a(+\delta_{ILD}), \quad (16)$$

$$\Phi_3 = \Phi_b(-\delta_{ILD}), \quad \Phi_4 = \Phi_b(+\delta_{ILD}), \quad (17)$$

$$\Phi_5 = \Phi_c(-\delta_{ILD}), \quad \Phi_6 = \Phi_c(+\delta_{ILD}), \quad (18)$$

e as matrizes  $\Phi_a$ ,  $\Phi_b$  e  $\Phi_c$  são dadas por:

$$\Phi_a(\delta_a) = \begin{bmatrix} \Phi_v & \mathbf{0}_M \\ \mathbf{0}_M & -(\text{ILD}_v^{\text{in}} - \delta_a) \cdot \Phi_v \end{bmatrix}, \quad (19)$$

$$\Phi_b(\delta_b) = \begin{bmatrix} \mathbf{0}_M & \Phi_v \\ \Phi_v & -2(\Re\{\text{ITF}_v^{\text{in}}\} - \delta_b \sqrt{\text{ILD}_v^{\text{in}}}) \cdot \Phi_v \end{bmatrix}, \quad (20)$$

$$\Phi_c(\delta_c) = \begin{bmatrix} \mathbf{0}_M & j\Phi_v \\ -j\Phi_v & 2(\Im\{\text{ITF}_v^{\text{in}}\} + \delta_c \sqrt{\text{ILD}_v^{\text{in}}}) \cdot \Phi_v \end{bmatrix}, \quad (21)$$

para  $\delta_a, \delta_b, \delta_c \in \{-\delta_{ILD}, +\delta_{ILD}\}$ , em que  $\Re\{\cdot\}$  e  $\Im\{\cdot\}$  determinam as partes real e imaginária de seus argumentos. Derivações detalhadas podem ser encontrados em [9].

#### IV. UMA SOLUÇÃO CONVEXA

A reformulação quadrática do problema de redução de ruído com preservação de pistas apresentada na seção anterior demonstra que o problema original estudado em [8] pode ser aproximado de maneira acurada por um QCQP, o qual é mais conhecido na literatura de otimização [10]. Porém, devido ao número elevado de restrições, o QCQP da equação (15) é um problema não-convexo e de difícil solução computacional. Especificamente, QCQPs em variáveis complexas com mais do que duas restrições são problemas *NP-hard* no caso geral; isto é, não é possível encontrar a solução ótima em tempo polinomial [10]. Esta é uma limitação considerável para aplicações de aparelhos auditivos, as quais necessitam de algoritmos computacionalmente eficientes.

Neste trabalho, propomos uma relaxação convexa do QCQP (15), a qual permite obter uma solução que apesar de sub-ótima, pode ser calculada de maneira muito mais eficiente e em tempo polinomial utilizando programação semidefinida [12]. Para isso, otimizamos a função custo dual de (15).

Primeiramente, calculamos o Lagrangeano de (15), o qual é dado por:

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\mu}) = \mathbf{w}^H \boldsymbol{\Phi}_{yq}(\boldsymbol{\mu}) \mathbf{w} - \mathbf{w}^H \boldsymbol{\Phi}_{xx} \mathbf{q} - \mathbf{q}^T \boldsymbol{\Phi}_{xx} \mathbf{w} + P_{xx}, \quad (22)$$

onde  $\boldsymbol{\mu} \in \mathbb{R}^6$  contém os multiplicadores de Lagrange, e:

$$\boldsymbol{\Phi}_{yq}(\boldsymbol{\mu}) = \boldsymbol{\Phi}_{yy} + \mu_q \sum_{q=1}^6 \boldsymbol{\Phi}_q. \quad (23)$$

A chamada *relaxação dual* do problema (15) é dada por:

$$\max_{\boldsymbol{\mu} \geq 0} \min_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\mu}). \quad (24)$$

A solução desse problema garante um limite inferior para o valor ótimo do problema original definido em (15). Além disso, o problema dual é sempre côncavo, mesmo quando o problema original é não-convexo.

Uma propriedade importante do Lagrangeano do problema em questão é que se considerarmos os multiplicadores de Lagrange fixos, podemos encontrar os coeficientes  $\mathbf{w}$  que minimizam  $\mathcal{L}(\mathbf{w}, \boldsymbol{\mu})$  analiticamente. Calculando o gradiente de  $\mathcal{L}(\mathbf{w}, \boldsymbol{\mu})$  com respeito a  $\mathbf{w}$ , obtém-se:

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\mu}) = \boldsymbol{\Phi}_{yq}(\boldsymbol{\mu}) \mathbf{w} - \boldsymbol{\Phi}_{xx} \mathbf{q}. \quad (25)$$

Igualando (25) a zero e resolvendo para  $\mathbf{w}$ , encontramos o mínimo:

$$\mathbf{w}_{\text{WLD-LG}} = \boldsymbol{\Phi}_{yq}(\boldsymbol{\mu})^{-1} \boldsymbol{\Phi}_{xx} \mathbf{q}, \quad (26)$$

onde assumimos  $\boldsymbol{\Phi}_{yq}(\boldsymbol{\mu}) \succeq 0$  para que o valor da função custo seja finito. Agora, o problema de maximização em (24) pode ser representado substituindo  $\mathbf{w}_{\text{WLD-LG}}$  em  $\mathcal{L}(\mathbf{w}, \boldsymbol{\mu})$ , levando a:

$$\max_{\boldsymbol{\mu} \geq 0} -\mathbf{r}^H \boldsymbol{\Phi}_{yq}(\boldsymbol{\mu}) \mathbf{r}, \quad \text{s.a. } \boldsymbol{\Phi}_{yq} \succeq 0. \quad (27)$$

O problema descrito em (27) pode ser representado de forma equivalente usando um variável auxiliar  $\gamma$  e o complemento de Schur [12], o que leva a:

$$\begin{aligned} \max_{\gamma, \boldsymbol{\mu}} \quad & \gamma \\ \text{s.a.} \quad & \boldsymbol{\mu} \geq 0, \\ & \begin{bmatrix} \boldsymbol{\Phi}_{yq}(\boldsymbol{\mu}) & \mathbf{r}_{xx} \\ \mathbf{r}_{xx}^H & P_{xx} - \gamma \end{bmatrix} \succeq 0. \end{aligned} \quad (28)$$

O problema descrito em (28) consiste em um problema de programação semidefinida (SDP) com variáveis  $\gamma$  e  $\boldsymbol{\mu}$ , o qual pode ser resolvido de maneira computacionalmente eficiente (em tempo polinomial) [12].

Por fim, note que apesar da relaxação convexa em (28) não possuir garantias para a obtenção da solução ótima global de (15), em casos práticos, observamos uma solução muito próxima em um tempo reduzido. Além disso, encontrar o ótimo global requer tempo de cálculo exponencial, o que não é aceitável (devido à característica de sistema embarcado e funcionamento em tempo real) em aplicações relacionadas a aparelhos auditivos.

## V. SIMULAÇÕES COMPUTACIONAIS

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

### A. Técnicas analisadas

Três técnicas de redução de ruído foram analisadas: i) o MWF convencional, definido pela minimização de (8); ii) a técnica apresentada em [9], denominada QCQP-MWF-ILD-IC, que representa o estado da arte dentre as técnicas de preservação de campos acústicos genéricos; e iii) a técnica proposta neste trabalho, descrita em (28) e denominada SDP-MWF-ILD-IC. Para estas duas últimas técnicas utilizou-se  $\delta_{\text{ILD}} = 10^{-4}$  e  $\delta_{\text{IC}} = 10^{-4}$ .

### B. Cenário acústico

Dois cenários acústicos simulados, denominados  $C_1$  e  $C_2$ , foram utilizados nos experimentos, ambos contendo uma fonte pontual de fala de interesse. No cenário acústico  $C_1$ , a componente de ruído é decorrente da existência de um campo acústico difuso (ruído  $n$ ), i.e.,  $v = n$ . No cenário acústico  $C_2$ , além do sinal  $n$ , a componente de ruído contém um sinal proveniente de uma fonte pontual interferente (sinal  $u$ ), i.e.,  $v = u + n$ . Nos cenários acústicos  $C_1$  e  $C_2$ , a fonte de fala localiza-se no azimute  $\theta_x = 15^\circ$  (hemisfério direito) e está a 3 metros distante do usuário de aparelhos auditivos. Em  $C_2$ , a fonte pontual interferente também está a 3 metros de distância e seu azimute ( $\theta_u$ ) foi variado de  $\theta_u = -90^\circ$  (esquerda) até  $\theta_u = +90^\circ$  (direita) com passos de  $15^\circ$ . Presente em  $C_1$  e  $C_2$ , o campo acústico difuso foi assumido como sendo cilíndrico e foi gerado segundo o algoritmo em [13]. As funções de transferência (*head related impulse responses* – HRIR) que determinam as características acústicas das fontes pontuais foram obtidas da base de dados descrita em [14]. Elas foram obtidas em ambiente reverberante, utilizando um par de aparelhos auditivos e um simulador de cabeça e tronco. Cada aparelho auditivo possui 3 microfones, totalizando  $M = 6$  microfones.

### C. Sinais de fala e ruído

Os sinais de fala e interferência gerados pelas fontes pontuais foram simulados utilizando áudios de aproximadamente 2 segundos de duração obtidos de [15]. A frequência de amostragem utilizada foi 16 kHz. As componentes de fala e interferência em cada microfone foram geradas convoluindo-se individualmente os sinais de áudio com as HRIRs associadas ao caminho acústico entre a fonte acústica e cada um dos microfones do aparelho auditivo. O sinal contaminado nos microfones foi gerado somando-se a componente de fala e a respectiva componente de ruído de cada cenário acústico. No cenário acústico  $C_1$ , a razão sinal-ruído (*signal-to-noise ratio* - *SNR*) entre a fala e o ruído difuso é 0 dB. No cenário acústico  $C_2$ , a *SNR* e a razão sinal-interferência (*signal-to-interference ratio* - *SIR*) são ambas 0 dB, resultando em uma razão sinal-ruído-mais-interferência (*signal-to-noise-plus-interference ratio* - *SINR*) de -3 dB.

### D. Processamento

A fala contaminada foi processada em janelas de 256 amostras (ou 16 ms), que foram ponderadas por uma janela de análise e transformadas para domínio da frequência utilizando a transformada de Fourier de tempo curto com 512 pontos. A sobreposição entre trechos de amostras adjacentes é de 50%. Após o processo de filtragem no domínio da frequência, utilizou-se a transformada de Fourier inversa para obter as amostras processadas no domínio do tempo. O sinal reconstruído foi obtido utilizando-se o algoritmo de *overlap-and-add* a partir da multiplicação das amostras processadas por uma janela de síntese. As janelas de análise e síntese são ambas iguais à raiz quadrada da janela de Hanning. A descrição detalhada do algoritmo para representação dos sinais no domínio tempo-frequência é apresentada em [16]. As matrizes de coerência  $\Phi_n$  e  $\Phi_y$  foram calculadas diretamente da componente de ruído difuso e da fala contaminada. As matrizes  $\Phi_x$  e  $\Phi_u$  foram determinadas utilizando-se o método de *branqueamento da matriz de coerência*, apresentado em [17]. Os coeficientes  $w_L$  e  $w_R$  da técnica QCQP-MWF-ILD-IC foram obtidos utilizando-se a função `fmincon`<sup>1</sup> do Matlab. Já os coeficientes da técnica SDP-MWF-ILD-IC foram obtidos utilizando-se o solver SeDuMi [18]. A modelagem de ambos os programas de otimização foi feita utilizando-se o toolbox Yalmip [19].

### E. Critérios objetivos

Os resultados da simulação foram avaliados utilizando-se cinco métricas: i) erro na ILD ( $\Delta ILD_d$ ), ii) erro na IPD

( $\Delta IPD_d$ ), iii) erro na MSC ( $\Delta MSC_d$ ) [4], iv) variação de SINR ( $\Delta SINR$ ) e v) variação de SNR ( $\Delta SNR$ ). Para  $C_1$ , as métricas de desempenho foram calculadas utilizando-se os sinais de fala e o sinal de campo difuso. Para  $C_2$ , essas métricas de desempenho foram calculadas utilizando-se os sinais de fala e o sinal de interferência (fonte pontual).

## VI. RESULTADOS

A Tabela 1 apresenta os resultados obtidos para o cenário acústico  $C_1$  em que  $v = n$ . Neste cenário acústico, a técnica proposta apresenta um desempenho muito próximo do apresentado pelo MWF, enquanto que o QCQP-MWF-ILD-IC consegue preservar muito bem as pistas espaciais originais do campo acústico difuso, i.e., fazendo com que  $\Delta ILD_n$ ,  $\Delta IPD_n$  e  $\Delta MSC_n$  sejam próximos de zero. Porém, o QCQP-MWF-ILD-IC tem uma perda de desempenho significativa em termos de redução de ruído quando comparado ao MWF e ao SDP-MWF-ILD-IC.

Tabela 1: Cenário acústico  $C_1$  – Variação das pistas biauriculares do sinal de campo acústico difuso para as técnicas MWF e MWF-ILD-IC (nas formas de QCQP e SDP).

	MWF	QCQP	SDP
$\Delta ILD_n$ [dB]	0,895	0,114	0,818
$\Delta IPD_n$ [rad/ $\pi$ ]	0,628	0,040	0,628
$\Delta MSC_n$	0,980	0,004	0,980
$\Delta SNR$ [dB]	11,47	7,42	11,77

A Figura 1 apresenta os erros de ILD (Figura 1(a)) e IPD (Figura 1(b)) da fonte interferente em função de  $\theta_u$ , assim como a  $\Delta SINR$  (Figura 1(c)), para o cenário acústico  $C_2$ . O MWF apresenta os maiores erros nas pistas biauriculares da fonte interferente em todas as situações analisadas. Esse resultado corrobora resultados teóricos e práticos já conhecidos e indica o deslocamento aparente da fonte interferente para posição da fonte de fala, i.e.,  $\theta_u = \theta_x = 0^\circ$ . O SDP-MWF-ILD-IC reduz significativamente os erros de ILD e de IPD em comparação ao MWF. No azimute  $\theta_u = -60^\circ$ , a diferença de  $\Delta ILD_u$  entre MWF e SDP-MWF-ILD e entre MWF e QCQP-MWF-ILD-IC é, respectivamente, de 9,48 e 7,68. Em comparação com o QCQP-MWF-ILD-IC, o SDP-MWF-ILD-IC apresenta uma maior redução nos erros de ILD e IPD. Analisando-se a  $\Delta SINR$ , a técnica proposta apresenta um desempenho aproximadamente igual ao MWF, enquanto o QCQP-MWF-ILD-IC atinge, na média, 1,5 dB a menos em comparação ao SDP-MWF-ILD-IC. O gráfico de  $\Delta SNR$  foi omitido dado que é similar ao gráfico de SINR e, portanto, as conclusões obtidas são as mesmas.

As simulações foram realizadas em um computador de mesa, com processador Intel Core i7-3770 (*clock* de 3,40

<sup>1</sup><https://www.mathworks.com/help/optim/ug/fmincon.html>

## AGRADECIMENTOS

Este trabalho foi parcialmente financiado pela CAPES e pelo CNPq (315020/2018-0, 302492/2021-6).

## 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, *The psychophysics of human sound localization*. em *Spatial Hearing*. MIT Press, 1997.
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. D. M. do Carmo, R. A. Borsoi, and M. H. Costa, “Robust parameter strategy for Wiener-based binaural noise reduction methods in hearing aids”, *Biomed. Signal Process. Control*, v. 74, p. 1–11, 2022.
7. F. P. Itturriet 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.
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. M. do Carmo, R. A. Borsoi, and M. H. Costa, “Filtro multicanal de Wiener com restrição quadrática para preservação de pistas biauriculares em aparelhos auditivos”, *Simp. Bras. Telecom. Proc. Sinais*, 2021, p. 1–5.
10. J. Linderoth, “A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs”, *Math. Program.*, v. 103, n. 2, p. 251–282, 2005.
11. 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.
12. L. Vandenberghe, and S. Boyd, “Semidefinite programming”, *SIAM Review*, v. 38, n. 1, p. 49–95, 1996.
13. E. A. P. Habets, I. Cohen e S. Gannot, “Generating nonstationary multi-sensor signals under a spatial coherence constraint”, *J. Acoust. Soc. Am.*, v. 124, n. 5, p. 2911–2917, 2008.
14. 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.
15. 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>.
16. R. Crochiere, “A weighted overlap-add method of short-time Fourier analysis/synthesis”, *IEEE Trans. Audio Speech Lang. Process.*, v. 28, n. 1, p. 99–102, 28.
17. S. Markovich-Golan, S. Gannot, and W. Kellermann, “Performance analysis of the covariance-whitening and the covariance-subtraction methods for estimating the relative transfer function,” in *Proc. Eur. Signal Process. Conf.*, 2018, p. 2499–2503.
18. J. F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones”. *Optimization methods and software*, v. 11, n. 1–4, p. 625–653, 1999.
19. J. Lofberg, “YALMIP: A toolbox for modeling and optimization in MATLAB”. *Proc. IEEE Int. Conf. Robot. Automat.*, 2004, p. 284–289.

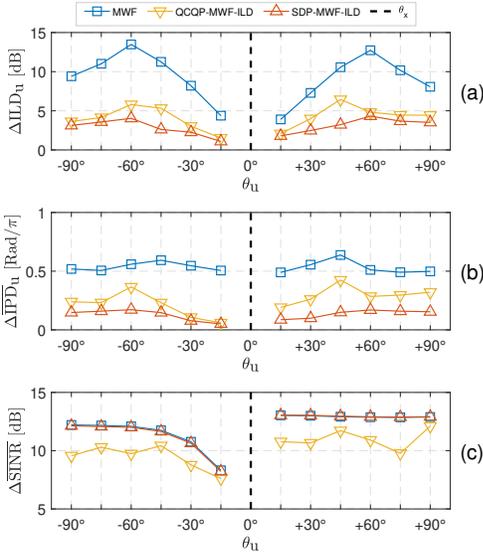


Figura 1: Cenário acústico  $C_2$  – Métricas analisadas: (a)  $\Delta IILD_u$ , (b)  $\Delta IPD_u$ , (c)  $\Delta SINR$ . Técnicas de processamento: (i) MWF ( $-\square-$ ), (ii) QCQP-MWF-ILD-IC ( $-\nabla-$ ), (iii) SDP-MWF-ILD-IC ( $-\triangle-$ ).

GHz) e utilizando Matlab™. A Tabela 2 apresenta os tempos de processamento das três técnicas analisadas. Verificamos que a técnica proposta apresenta uma redução significativa do tempo de processamento em comparação com o QCQP-MWF-ILD-IC.

Tabela 2: Tempo de otimização para as técnicas MWF, QCQP-MWF-ILD-IC e SDP-MWF-ILD-IC.

	MWF	QCQP	SDP
tempo [s]	$9 \times 10^{-5}$	1,319	0,074

## VII. CONCLUSÕES

Neste trabalho, uma abordagem convexa e computacionalmente eficiente para a redução de ruído, baseada no MWF com preservação de pistas biauriculares, foi proposta, considerando uma fonte pontual de interesse e ruído proveniente de uma fonte pontual ou campo difuso. Resultados experimentais de simulações indicam que o método proposto é capaz de preservar de forma acurada a percepção acústica de fontes pontuais de fala e ruído, com elevada eficiência computacional, mantendo um desempenho de redução de ruído comparável ao do MWF convencional. A preservação espacial de ruído difuso é semelhante à do MWF convencional.

## CONFLITOS DE INTERESSE

Os autores declaram não ter conflitos de interesse.