

Solução Fechada para o Filtro Multicanal de Wiener com Preservação de Função de Transferência Interauricular

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

Resumo— O filtro multicanal de Wiener com preservação da função de transferência interaural (MWF-ITF) é o método estado da arte para redução de ruído com preservação das pistas biauriculares de fontes sonoras pontuais. Entretanto, a determinação dos seus coeficientes envolve a solução de um problema de otimização não-convexo, levando a um custo computacional elevado. Este trabalho apresenta uma solução fechada para o MWF-ITF obtida através da relaxação dual do problema de otimização original. A propriedade de dualidade forte é empregada para garantir a otimalidade global da solução encontrada. Os resultados indicam que a solução proposta atinge um desempenho equivalente ao apresentado pelo MWF-ITF convencional, porém, com um tempo de processamento até 1300 vezes mais baixo.

Palavras-Chave— Aparelho auditivo biauricular, MWF, ITF, ILD, IPD, solução fechada.

Abstract— The multichannel Wiener filter with interaural transfer function preservation (MWF-ITF) is the state-of-the-art technique for noise reduction with preservation of the binaural cues of pointwise acoustic sources. However, the computation of its filter coefficients involves solving a non-convex optimization problem, leading to a high computational cost. This work presents a closed-form solution to the MWF-ITF, which is obtained using the dual relaxation approach for the original optimization problem. The strong duality property is employed to guarantee the global optimality of the solution. Results indicate that the proposed solution achieves performance equivalent to that presented by the traditional MWF-ITF, but with a processing time up to 1300 times lower.

Keywords— Binaural hearing aids, MWF, ITF, ILD, IPD, closed-form solution.

I. INTRODUÇÃO

Aparelhos auditivos biauriculares são pares de aparelhos auditivos que possuem um canal de comunicação entre si para troca de sinais e parâmetros de controle. Para esses aparelhos, é importante que técnicas de redução de ruído atendam a dois importantes critérios de projeto [1], quais sejam: (i) reduzir a potência do ruído aditivo e (ii) preservar a percepção original da posição espacial dos componentes sonoros captados pelos microfones. Para atingir esses dois objetivos conjuntamente, o projeto da técnica de redução de ruído deve levar em conta as informações de localização espacial utilizadas pelo sistema auditivo humano: a diferença de tempo/fase interauricular (*interaural time/phase difference* – ITD/IPD) e a diferença de nível interauricular (*interauricular level difference* – ILD).

Diego M. do Carmo e Márcio H. Costa – Programa de Pós-Graduação em Engenharia Elétrica – Universidade Federal de Santa Catarina. E-mails: diego.carmoh@gmail.com, costa@eel.ufsc.br. Ricardo A. Borsoi - CNRS, CRAN, Université de Lorraine, F-54000, Nancy, France. Email: raborsoi@gmail.com. Este trabalho foi parcialmente financiado pela CAPES.

Como regra geral, a IPD é a informação biauricular utilizada pelo sistema auditivo humano para localizar fontes sonoras emitindo sons com frequências abaixo de 1500 Hz, enquanto que a ILD é utilizada para localizar fontes sonoras com frequências iguais ou acima de 1500 Hz [2]. Portanto, para que fontes sonoras presentes no ambiente sejam corretamente localizadas pelo usuário de aparelhos auditivos biauriculares é necessário que as técnicas de redução de ruído garantam que a ILD e a IPD dos sinais processados pelos aparelhos (que são enviados aos alto-falantes) sejam iguais às captadas pelos microfones. Além da ILD e da IPD, uma outra informação biauricular importante é a coerência interauricular (*interaural coherence* – IC). Perceptualmente, a IC está associada ao tamanho percebido das fontes sonoras [1].

Dentre as técnicas de redução biauricular de ruído, as baseadas no filtro multicanal de Wiener (*multichannel Wiener filter* – MWF) são bastante estudadas e utilizadas [3]. O MWF determina um estimador para um sinal coerente de interesse (tipicamente um sinal de fala) presente nos microfones de referência. Como consequência, ele preserva perfeitamente a ILD e a IPD da fonte de fala. Contudo, altera a ILD e a IPD do ruído, o que modifica a percepção acústica do ambiente e dificulta, por exemplo, a identificação de outras fontes sonoras [3]. Por isso, uma abordagem comumente empregada para obter filtros biauriculares a partir do MWF baseia-se em adicionar à função custo que determina o MWF um termo associado à preservação da ILD e da IPD do sinal interferente. Esse termo adicional tem por objetivo penalizar soluções que distorcem as pistas biauriculares do ruído.

Dentre os termos para preservação de fontes pontuais que são adicionados à função custo do MWF, destacam-se o termo relacionado à preservação da ILD [4] e os associados à preservação da IPD [4], [5]. Um terceiro termo que pode ser adicionado à função custo do MWF é associado à função de transferência interauricular (*interaural transfer function* – ITF) [6]. A ITF carrega informação sobre a ILD e a IPD de um sinal gerado por uma fonte pontual. Dessa forma, é possível controlar o erro na ILD e na IPD com um único termo.

Uma grande desvantagem das técnicas baseadas em função custo aumentada está no fato de não haver uma solução fechada para o cálculo dos coeficientes para nenhuma das opções descritas anteriormente. Isso leva à necessidade de empregar algoritmos iterativos de otimização, o que torna o tempo de busca de uma solução bastante elevado em comparação a técnicas que admitem soluções fechadas. Visando solucionar esse problema, este trabalho apresenta uma solução fechada

para a determinação dos coeficientes que minimizam a função custo do MWF com o termo para preservação da ITF (MWF-ITF). Primeiramente, o problema do MWF-ITF é reformulado como um problema de otimização quadrático, o qual é solucionado utilizando resultados recentes de otimização não-convexa, os quais demonstram que a solução globalmente ótima do problema em questão pode ser encontrada através de sua relaxação dual e de suas condições de Karush-Kuhn-Tucker, as quais admitem uma solução eficiente baseada em operações simples de álgebra linear. Simulações indicam que o método proposto apresenta um desempenho equivalente ao MWF-ITF tradicional em termos de redução de ruído e preservação de pistas espaciais, contudo com um tempo de processamento até 1300 vezes menor.

II. MODELO DE SINAL E REDUÇÃO DE RUÍDO

O vetor de fala contaminada $\mathbf{y}_\ell(\lambda, k) \in \mathbb{C}^{M_\ell \times 1}$ de dimensão $M_\ell \times 1$ captado no aparelho auditivo do lado $\ell = \{L, R\}$ no domínio tempo-frequência é dado por

$$\mathbf{y}_\ell(\lambda, k) = [y_{\ell,1}(\lambda, k) y_{\ell,2}(\lambda, k) \cdots y_{\ell,M_\ell}(\lambda, k)]^\top, \quad (1)$$

em que $y_{\ell,m_\ell}(\lambda, k)$ é o sinal captado pelo microfone $m_\ell = 1, \dots, M/2$, para $M = 2M_L = 2M_R$; λ e k são os índices de tempo e frequência e $(\bullet)^\top$ é o operador de transposição. O sinal em (1) pode ser decomposto da seguinte forma:

$$\mathbf{y}_\ell(\lambda, k) = \mathbf{x}_\ell(\lambda, k) + \underbrace{\mathbf{u}_\ell(\lambda, k) + \mathbf{n}_\ell(\lambda, k)}_{\mathbf{v}_\ell(\lambda, k)}, \quad (2)$$

em que $\mathbf{x}_\ell(\lambda, k)$ é o vetor de fala, $\mathbf{v}_\ell(\lambda, k) = \mathbf{u}_\ell(\lambda, k) + \mathbf{n}_\ell(\lambda, k)$ é o vetor de ruído aditivo, que pode ser decomposto em um componente coerente $\mathbf{u}_\ell(\lambda, k)$ (sinal de interferência), e um componente não-coerente $\mathbf{n}_\ell(\lambda, k)$ (ruído ambiente e/ou do sensor). O número total de microfones em ambos os aparelhos auditivos é dado por $M = M_L + M_R$. Os sinais captados por esses microfones são conhecidos por ambos os aparelhos auditivos através de um canal de comunicação *full duplex* e são representados conjuntamente por:

$$\mathbf{y}(\lambda, k) = [\mathbf{y}_L^\top(\lambda, k) \mathbf{y}_R^\top(\lambda, k)]^\top \in \mathbb{C}^{M \times 1}. \quad (3)$$

Da mesma maneira que $\mathbf{y}_\ell(\lambda, k)$, o vetor em $\mathbf{y}(\lambda, k)$ pode ser decomposto em componentes de fala e ruído, isto é, $\mathbf{y}(\lambda, k) = \mathbf{x}(\lambda, k) + \mathbf{v}(\lambda, k)$, em que $\mathbf{v}(\lambda, k) = \mathbf{u}(\lambda, k) + \mathbf{n}(\lambda, k)$.

Um microfone em cada aparelho auditivo é selecionado como microfone de referência. Essa seleção é feita a partir de um vetor $\mathbf{q}_\ell \in \mathbb{R}^{M \times 1}$, o qual contém o valor 1 na posição do microfone de referência e 0 nas demais. O sinal captado nesse microfone é dada por $\mathbf{q}_\ell^\top \mathbf{y}(\lambda, k) = y_\ell(\lambda, k) = x_\ell(\lambda, k) + v_\ell(\lambda, k)$.

O sinal processado nos aparelhos auditivos, no domínio tempo-frequência, é determinado por $z_L(\lambda, k) = \mathbf{w}_L^H(\lambda, k) \mathbf{y}(\lambda, k)$ e $z_R(\lambda, k) = \mathbf{w}_R^H(\lambda, k) \mathbf{y}(\lambda, k)$, em que $\mathbf{w}_\ell(\lambda, k) \in \mathbb{C}^{M \times 1}$ é o vetor de coeficientes do filtro de redução de ruído no lado ℓ . Os vetores $\mathbf{w}_L(\lambda, k)$ e $\mathbf{w}_R(\lambda, k)$ e os vetores de seleção $\mathbf{q}_L(\lambda, k)$ e $\mathbf{q}_R(\lambda, k)$ são agrupados nos vetores $\mathbf{w}(\lambda, k)$ e $\mathbf{q}(\lambda, k)$, respectivamente, i.e.,

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

As matrizes de coerência da fala contaminada ($\Phi_y(k)$), do componente de fala ($\Phi_x(k)$), do componente interferente ($\Phi_u(k)$) e do componente de ruído incoerente ($\Phi_n(k)$) podem ser genericamente definidos como:

$$\Phi_d(k) = \mathbb{E}\{\mathbf{d}(k, \lambda) \mathbf{d}^H(k, \lambda)\} \in \mathbb{C}^{M \times M}, \quad (5)$$

em que $\mathbf{d} = \{\mathbf{y}, \mathbf{u}, \mathbf{n}, \mathbf{v}\}$, $(\bullet)^H$ representa o operador transposto conjugado e $\mathbb{E}\{\bullet\}$ é o operador valor esperado.

Define-se matematicamente a ILD, a IPD e a ITF do sinal interferente na entrada, respectivamente por:

$$\text{ILD}_{\text{in}}^u = (\mathbf{q}_L^\top \Phi_u \mathbf{q}_L) (\mathbf{q}_R^\top \Phi_u \mathbf{q}_R)^{-1}, \quad (6)$$

$$\text{IPD}_{\text{in}}^u = \angle(\mathbf{q}_L^H \Phi_u \mathbf{q}_R), \quad (7)$$

$$\text{ITF}_{\text{in}}^u = (\mathbf{q}_L^\top \Phi_u \mathbf{q}_R) (\mathbf{q}_R^\top \Phi_u \mathbf{q}_L)^{-1}, \quad (8)$$

em que $\angle(\bullet)$ é a fase desenrolada de um número complexo. Define-se também a ILD, a IPD e a ITF do sinal interferente processado, respectivamente, por:

$$\text{ILD}_{\text{ou}}^u = (\mathbf{w}_L^H \Phi_u \mathbf{w}_L) (\mathbf{w}_R^H \Phi_u \mathbf{w}_R)^{-1}, \quad (9)$$

$$\text{IPD}_{\text{ou}}^u = \angle(\mathbf{w}_L^H \Phi_u \mathbf{w}_R), \quad (10)$$

$$\text{ITF}_{\text{ou}}^u = (\mathbf{w}_L^H \Phi_u \mathbf{w}_R) (\mathbf{w}_R^H \Phi_u \mathbf{w}_L)^{-1}. \quad (11)$$

Define-se a seguinte matriz real paramétrica:

$$\mathbf{G}(\mathbf{A}) = \begin{bmatrix} \Re\{\mathbf{A}\} & -\Im\{\mathbf{A}\} \\ \Im\{\mathbf{A}\} & \Re\{\mathbf{A}\} \end{bmatrix} \quad (12)$$

em que $\Re\{\bullet\}$ e $\Im\{\bullet\}$ são as partes real e imaginária dos seus argumentos. Essa matriz é utilizada para facilitar a representação de problemas de otimização com variáveis complexas em termos de variáveis reais.

Para simplificar a notação, os índices λ e k serão omitidos das representações matemáticas posteriores.

A. Filtro de Wiener multicanal com termo adicional para preservação de pistas binauriculares

Uma das abordagens mais utilizadas para determinar os filtros \mathbf{w}_L e \mathbf{w}_R para aparelhos auditivos binauriculares é baseada na minimização da função custo do erro quadrático médio (*mean squared error* - MSE) entre os componentes de fala no microfone de referência ($\mathbf{q}_L^\top \mathbf{x}$ e $\mathbf{q}_R^\top \mathbf{x}$) e o sinal contaminado processado ($\mathbf{w}_L^H \mathbf{y}$ e $\mathbf{w}_R^H \mathbf{y}$). Sabendo da relação entre os vetores \mathbf{w}_R e \mathbf{w}_L e o vetor \mathbf{w} e assumindo-se \mathbf{x} e \mathbf{v} como vetores de variáveis aleatórias, decorrelacionados e de média zero, pode-se obter a expressão para a função custo:

$$J_{\text{MWF}}(\mathbf{w}) \triangleq \mathbf{w}^H \Phi_{yy} \mathbf{w} - \mathbf{w}^H \mathbf{p}_{xx} - \mathbf{p}_{xx}^H \mathbf{w} + p_{xx}, \quad (13)$$

em que $\mathbf{p}_{xx} = \Phi_{xx} \mathbf{q}$; $p_{xx} = \mathbf{q}^\top \Phi_{xx} \mathbf{q}$; Φ_{yy} e Φ_{xx} são definidas como

$$\Phi_{yy} = \begin{bmatrix} \Phi_y & \mathbf{O} \\ \mathbf{O} & \Phi_y \end{bmatrix} \text{ e } \Phi_{xx} = \begin{bmatrix} \Phi_x & \mathbf{O} \\ \mathbf{O} & \Phi_x \end{bmatrix}, \quad (14)$$

sendo \mathbf{O} a matriz de zeros de dimensão $M \times M$. Utilizando a função custo em (13), o projeto de \mathbf{w} pode ser feito a partir do seguinte problema de minimização [3]:

$$\mathbf{w}_{\text{MWF}} = \underset{\mathbf{w}}{\text{arg.min.}} J_{\text{MWF}}(\mathbf{w}). \quad (15)$$

A solução de (15) é dada por:

$$\mathbf{w}_{\text{MWF}} = \Phi_{yy}^{-1} \mathbf{r}_{xx}. \quad (16)$$

O MWF definido em (16) reduz a potência do ruído aditivo ao mesmo tempo que preserva as pistas binauriculares da fala [3]. Porém, ele altera a percepção espacial da fonte interferente [3] de maneira que seu ângulo de chegada passa a ser percebido como sendo o mesmo ângulo da fonte de fala. Essa mudança na percepção do componente interferente ocorre devido a diferenças entre a ILD/IPD de entrada e a ILD/IPD de saída.

Um método bastante utilizado para reduzir as distorções causadas pelo MWF na ILD e na IPD do componente interferente é o MWF com termo de penalização da ITF. Neste método, os coeficientes são calculados a partir do seguinte problema de minimização [6]:

$$\mathbf{w}_{\text{MWF-ITF}} = \arg.\min_{\mathbf{w}} J_{\text{MWF}}(\mathbf{w}) + \alpha J_{\text{ITF}}(\mathbf{w}), \quad (17)$$

em que o termo $J_{\text{ITF}}(\mathbf{w})$, ponderado por $\alpha \geq 0$, penaliza distorções na ILD e na IPD do sinal interferente processado, sendo definido como [6]:

$$J_{\text{ITF}}(\mathbf{w}) \triangleq (\mathbf{w}^H \Phi_{aa} \mathbf{w})(\mathbf{w}^H \Phi_{bb} \mathbf{w})^{-1}, \quad (18)$$

em que

$$\Phi_{aa} = \begin{bmatrix} \Phi_u & -\beta^* \Phi_u \\ -\beta \Phi_u & |\beta|^2 \Phi_u \end{bmatrix} \text{ e } \Phi_{bb} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \Phi_u \end{bmatrix}, \quad (19)$$

sendo $\beta = \text{ITF}_u^{\text{in}}$ e $(\bullet)^*$ o conjugado de um número complexo.

III. PROPOSTA DE SOLUÇÃO FECHADA PARA O MWF-ITF

Para determinar uma solução fechada para o MWF-ITF, retira-se o termo $J_{\text{ITF}}(\mathbf{w})$ da função custo (17), reintegrando-o na forma de uma restrição do problema de minimização, resultando em:

$$\mathbf{w}_{\text{CB-MWF-ITF}} = \arg.\min_{\mathbf{w}} \mathbf{w}^H \Phi_{yy} \mathbf{w} - \mathbf{w}^H \mathbf{r}_{xx} - \mathbf{r}_{xx}^H \mathbf{w} \quad (20a)$$

$$\text{s.a. } (\mathbf{w}^H \Phi_{aa} \mathbf{w})(\mathbf{w}^H \Phi_{bb} \mathbf{w})^{-1} \leq \delta, \quad (20b)$$

em que $\delta > 0$ determina o erro entre a ITF_u^{ou} e ITF_n^{in} . Em seguida, escreve-se (20) utilizando explicitamente vetores e matrizes reais, o que resulta em:

$$\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}} = \arg.\min_{\tilde{\mathbf{w}}} \tilde{\mathbf{w}}^H \tilde{\Phi}_{yy} \tilde{\mathbf{w}} - 2\tilde{\mathbf{r}}_{xx}^T \tilde{\mathbf{w}} \quad (21a)$$

$$\text{s.a. } (\tilde{\mathbf{w}}^H \tilde{\Phi}_{aa} \tilde{\mathbf{w}})(\tilde{\mathbf{w}}^H \tilde{\Phi}_{bb} \tilde{\mathbf{w}})^{-1} \leq \delta, \quad (21b)$$

em que $\tilde{\Phi}_{yy} = \mathbf{G}(\Phi_{yy})$, $\tilde{\Phi}_{aa} = \mathbf{G}(\Phi_{aa})$, $\tilde{\Phi}_{bb} = \mathbf{G}(\Phi_{bb})$, $\tilde{\mathbf{w}} = [\Re\{\mathbf{w}\} \Im\{\mathbf{w}\}]^T$ e $\tilde{\mathbf{r}}_{xx} = [\Re\{\mathbf{r}_{xx}\}^T \Im\{\mathbf{r}_{xx}\}]^T$. A relação entre (20) e (21) é dada por:

$$\mathbf{w}_{\text{CB-MWF-ITF}} = [\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}}]_{1:M} + j [\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}}]_{M+1:2M},$$

em que $[\bullet]_{1:M}$ e $[\bullet]_{M+1:2M}$ denotam os primeiros e os últimos M elementos de seu argumento, respectivamente.

Nota-se que a restrição (21b) pode ser reescrita na forma de uma restrição quadrática, uma vez que:

$$J_{\text{ITF}}(\tilde{\mathbf{w}}) \leq \delta \iff \tilde{\mathbf{w}}^T \tilde{\Phi}_{cc} \tilde{\mathbf{w}} \leq 0, \quad (22)$$

em que $\tilde{\Phi}_{cc} = \tilde{\Phi}_{aa} - \delta \tilde{\Phi}_{bb}$. Isso permite rescrever o problema (21) de maneira equivalente ao seguinte problema de

minimização de uma função custo quadrática com restrição também quadrática:

$$\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}} = \arg.\min_{\tilde{\mathbf{w}}} \tilde{\mathbf{w}}^H \tilde{\Phi}_{yy} \tilde{\mathbf{w}} - 2\tilde{\mathbf{r}}_{xx}^T \tilde{\mathbf{w}} \quad (23a)$$

$$\text{s.a. } c_{\text{ITF}}(\tilde{\mathbf{w}}) \triangleq \tilde{\mathbf{w}}^H \tilde{\Phi}_{cc} \tilde{\mathbf{w}} \leq 0. \quad (23b)$$

Como o problema em (23) possui uma função custo quadrática e uma única restrição quadrática, se existir um $\tilde{\mathbf{w}}$ tal que $c_{\text{ITF}}(\tilde{\mathbf{w}}) < 0$ (i.e., se a condição de Slater [8] for satisfeita), então $\tilde{\mathbf{w}}^*$ é uma solução global de (23) se e somente se existe um τ^* tal que as condições de primeira ordem de Karush–Kuhn–Tucker [8] são satisfeitas (veja o Teorema 1 de [7]):

$$(\tilde{\Phi}_{yy} + \tau^* \tilde{\Phi}_{cc}) \tilde{\mathbf{w}}_{\text{CB-MWF-ITF}} = \tilde{\mathbf{r}}_{xx}, \quad (24a)$$

$$c_{\text{ITF}}(\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}}) \leq 0, \quad (24b)$$

$$\tau^* c_{\text{ITF}}(\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}}) = 0, \quad (24c)$$

$$\tilde{\Phi}_{yy} + \tau^* \tilde{\Phi}_{cc} \succeq 0, \quad (24d)$$

em que $\tau^* \geq 0$ é o multiplicador de Lagrange do problema (23) [8]. As condições (24a)-(24c) formam um conjunto de (in)equações não-lineares que pode ser resolvido de forma analítica conforme demonstrado em [7]. Assim, ao encontrar o multiplicador de Lagrange ótimo, τ^* , a solução fechada para $\mathbf{w}_{\text{CB-MWF-ITF}}$ é dada resolvendo o sistema linear em (24a) considerando a condição em (24d), o que resulta em:

$$\tilde{\mathbf{w}}_{\text{CB-MWF-ITF}}(\tau^*) = (\tilde{\Phi}_{yy} + \tau^* \tilde{\Phi}_{cc})^\dagger \tilde{\mathbf{r}}_{xx}. \quad (25)$$

em que $(\bullet)^\dagger$ representa a pseudo-inversa à esquerda.

A. Valor ótimo do multiplicador de Lagrange

Para calcular o multiplicador de Lagrange ótimo, o Algoritmo 3.2 proposto em [7] pode ser utilizado, o qual propõe resolver o sistema em (24), isto é, encontrar os valores de τ^* e $\tilde{\mathbf{w}}^*$ resolvendo um problema de autovetor-autovalor generalizado [9]. Esse algoritmo envolve duas etapas. Primeiramente, necessita-se de uma solução inicial $\hat{\tau} \geq 0$ tal que $\tilde{\Phi}_{yy} + \hat{\tau} \tilde{\Phi}_{cc} \succ 0$. No caso do MWF-ITF a matriz $\tilde{\Phi}_{yy}$ pode ser assumida positiva definida sem perda de generalidade. Consequentemente, qualquer valor de $\hat{\tau}$ menor que um dado limiar positivo garantirá $\tilde{\Phi}_{yy} + \hat{\tau} \tilde{\Phi}_{cc} \succ 0$.

Para prosseguir, definem-se as matrizes \mathbf{M}_0 , \mathbf{M}_1 , e $\widehat{\mathbf{M}}$ como:

$$\mathbf{M}_0 = \begin{bmatrix} 0 & \mathbf{0}^\top & \tilde{\mathbf{r}}_{xx}^\top \\ \mathbf{0} & \tilde{\Phi}_{cc} & -\tilde{\Phi}_{yy} \\ \tilde{\mathbf{r}}_{xx} & -\tilde{\Phi}_{yy} & \mathbf{O} \end{bmatrix}, \quad (26)$$

$$\mathbf{M}_1 = \begin{bmatrix} 0 & \mathbf{0}^\top & \mathbf{0}^\top \\ 0 & \mathbf{O} & -\tilde{\Phi}_{cc} \\ \mathbf{0} & -\tilde{\Phi}_{cc} & \mathbf{O} \end{bmatrix}, \quad (27)$$

$$\widehat{\mathbf{M}} = \mathbf{M}_0 + \hat{\tau} \mathbf{M}_1, \quad (28)$$

em que $\mathbf{0}$ é um vetor de zeros de dimensão $M \times 1$. Essas matrizes são importantes, pois, como mostrado no Teorema 3 de [7], o multiplicador de Lagrange ótimo $\tau^* \geq 0$ em (24) é um autovalor generalizado do par de matrizes em (28), isto é, existe um vetor $\tilde{\mathbf{z}}$ diferente de zero tal que:

$$\widehat{\mathbf{M}} \tilde{\mathbf{z}} = (\mathbf{M}_0 + \tau^* \mathbf{M}_1) \tilde{\mathbf{z}} = \mathbf{0} \iff \mathbf{M}_0 \tilde{\mathbf{z}} = \tau^* \mathbf{M}_1 \tilde{\mathbf{z}}. \quad (29)$$

Consequentemente, conforme demonstrado em [7], a partir do valor da solução inicial $\hat{\tau}$ e do sinal da restrição $c_{\text{ITF}}(\tilde{\mathbf{w}}(\hat{\tau}))$, é possível determinar o valor de τ^* analiticamente utilizando a solução de um problema de autovalores generalizados [9], o que leva a três casos diferentes:

- primeiro, se $c_{\text{ITF}}(\tilde{\mathbf{w}}(\hat{\tau})) = 0$, ocorre o caso mais simples em que o multiplicador de Lagrange pode ser obtido diretamente como $\tau^* = \hat{\tau}$.
- se $c_{\text{ITF}}(\tilde{\mathbf{w}}(\hat{\tau})) > 0$, então precisa-se do autovalor generalizado finito mais à direita (maior) ξ do par de matrizes $(\mathbf{M}_1 + \xi \widehat{\mathbf{M}})\mathbf{z} = \mathbf{0}$, de forma que o multiplicador de Lagrange será dado por $\tau^* = \hat{\tau} + \xi^{-1}$.
- se $c_{\text{ITF}}(\tilde{\mathbf{w}}(\hat{\tau})) < 0$, então precisa-se do autovalor generalizado finito mais à esquerda (menor) ξ do par de matrizes $(\mathbf{M}_1 + \xi \widehat{\mathbf{M}})\mathbf{z} = \mathbf{0}$. Agora há mais dois casos para tratar: se $\hat{\tau} = 0$, ou se $\Re(\hat{\tau} + \xi^{-1}) \leq 0$, então o multiplicador de Lagrange é dado por como $\tau^* = 0$, caso contrário, é dado por $\tau^* = \hat{\tau} + \xi^{-1}$.

Após obter τ^* , os coeficientes ótimos podem ser calculados utilizando (25).

IV. SIMULAÇÕES COMPUTACIONAIS

Esta seção compara o desempenho do método CB-MWF-ITF, proposto na Seção III, com o do MWF, dado por (15), e com o do MWF-ITF convencional, dado por (17).

A. Cenário acústico simulado

O componente de fala e o componente interferente foram gerados convoluindo-se sinais de fala com repostas ao impulso de um par de aparelhos auditivos retroauriculares, cuja base de dados é descrita em [11]. Os aparelhos auditivos possuem 3 microfones cada, totalizando $M = 6$ microfones. As repostas ao impulso foram obtidas em uma sala com tempo de reverberação $T_{60} \approx 300$ ms. Foram utilizados áudios de fala da base de dados em [12], dos quais quinze foram utilizados para gerar os componentes de fala, e outros quinze os componentes interferentes. A duração média dos sinais é de 2 segundos. A fonte da fala localiza-se no azimute 0° (frente), a uma distância de 3 m do aparelho auditivo. O componente de ruído foi gerado considerando o método descrito em [13]. O sinal contaminado foi gerado somando-se o componente de fala, o componente interferente e o componente de ruído ambiente. Essa combinação foi feita de maneira que a razão sinal-ruído (*signal-to-noise ratio* – SNR) e a razão sinal-interferência (*signal-to-interference ratio* – SIR) fossem iguais a 0 dB, o que resultou em um razão sinal-interferência-mais-ruído (*signal-to-interference-plus-noise* – SINR) igual a -3 dB [12]. Um total de 285 simulações (15 conjuntos de sinais e 19 azimutes para a fonte interferente) foram realizadas para cada técnica de redução de ruído.

B. Parâmetros e algoritmos de otimização

Os coeficientes do MWF foram obtidos a partir de (16). Os coeficientes do MWF-ITF convencional, $\mathbf{w}_{\text{MWF-ITF}}$, foram obtidos a partir do algoritmo definido no Apêndice A, utilizando a rotina 'fmincon' do Matlab. O gradiente e a Hessiana do termo J_{ITF} definido em (18) foram fornecidos. Os coeficientes da técnica proposta foram obtidos a partir de (25), juntamente

com o τ^* obtido utilizando a descrição da seção III-A. O valor inicial $\hat{\tau}$ foi definido como 10^{-1} . O erro na ITF, $\delta(\lambda, k)$, foi definido arbitrariamente pequeno com valor de 10^{-4} .

C. Processamento

Para processamento, o sinal contaminado no tempo foi dividido em trechos de 128 amostras (ou 8 ms) ponderadas por uma janela de análise. A sobreposição entre trechos de amostras adjacentes foi de 50%. A representação tempo-frequência desses trechos foi obtida utilizando a transformada de Fourier de tempo curto com $K = 256$ pontos e preenchimento com zeros. As amostras no domínio do tempo-frequência foram processadas pelos filtros de redução de ruído, obtendo-se o sinal processado no domínio da frequência para ambos os aparelhos auditivos. Esses sinais foram transformados novamente para o domínio do tempo e ponderados por uma janela de síntese. Por fim, o trecho de sinal enviado aos alto-falantes dos aparelhos auditivos foi gerado fazendo-se a adição de trechos adjacentes de amostras processadas. As janelas de análise e síntese são ambas definidas pela raiz quadrada da janela de Hanning. A estimação das matrizes de coerência da fala Φ_x e do componente coerente do ruído Φ_u foi feita a partir do método de branqueamento das matrizes de coerência, para mais informações sobre esse processo, consulte a Seção V de [12].

D. Critérios objetivos

O desempenho dos métodos foi comparado a partir das seguintes métricas de desempenho: os erros na ILD e na IPD, respectivamente denominados de ΔILD e ΔIPD [12]; e o ganho de SINR (ΔSINR) [12].

E. Resultados e discussão

Variação de ILD e IPD: As Figuras 1(a) e (b) apresentam a ΔILD para os componentes de interferência (ΔILD_u) e de fala (ΔILD_x). Analisando-se as curvas de ΔILD_u , vê-se que o MWF clássico distorce severamente a ILD da fonte interferente em comparação às duas técnicas baseadas na ITF. Esse comportamento é conhecido e perceptualmente indica que a direção percebida de chegada do sinal da fonte interferente é a mesma da fonte de fala [3]. Analisando-se as curvas de ΔILD_x , percebe-se que o método CB-MWF-ITF proposto e o MWF-ITF convencional resultam em valores de ΔILD_u numericamente próximos. Indicando que ambos mantêm a localização percebida da fonte de fala em 0° . As Figuras 1(c) e (d) apresentam a ΔIPD para o componente interferente (ΔIPD_u) e de fala (ΔIPD_x). Em relação à ΔIPD para o componente de interferência (ΔIPD_u), o CB-MWF-ITF e o MWF-ITF apresentam um desempenho semelhante entre si e melhor que o obtido pelo MWF. Portanto, a partir dos resultados apresentados nesta seção, encontram-se evidências de que após processamento pelo CB-MWF-ITF e MWF-ITF, o azimute percebido da fonte interferente se mantém próximo do azimute original da fonte de interferência. Além disso, ambas as técnicas preservam a percepção original da posição da fonte de fala, uma vez que o desempenho delas é semelhante ao do MWF. **Redução de componentes indesejados:** A Figura

1(d) apresenta a ΔSINR . O método proposto apresenta um desempenho de redução de ruído equivalente ao do MWF clássico e ao do MWF-ITF convencional para todos os ângulos de interferência. **Tempo de busca de coeficientes:** A partir de simulações realizadas em um computador pessoal com processador Intel Core i7-3770, rodando em 3,40 GHz e o utilizando o programa Matlab, foram obtidos os seguintes tempos médios de cálculo dos coeficientes de redução de ruído: 7×10^{-5} segundos para o MWF; 1,3 segundos para o MWF-ITF convencional; e 1×10^{-3} segundos para o CB-MWF-ITF. Isto indica que o tempo para se determinar os coeficientes utilizando o método proposto CB-MWF-ITF é 1300 vezes mais rápido do que o do MWF-ITF convencional (i.e., solução de (17)) e 14 vezes mais lento do que o do MWF.

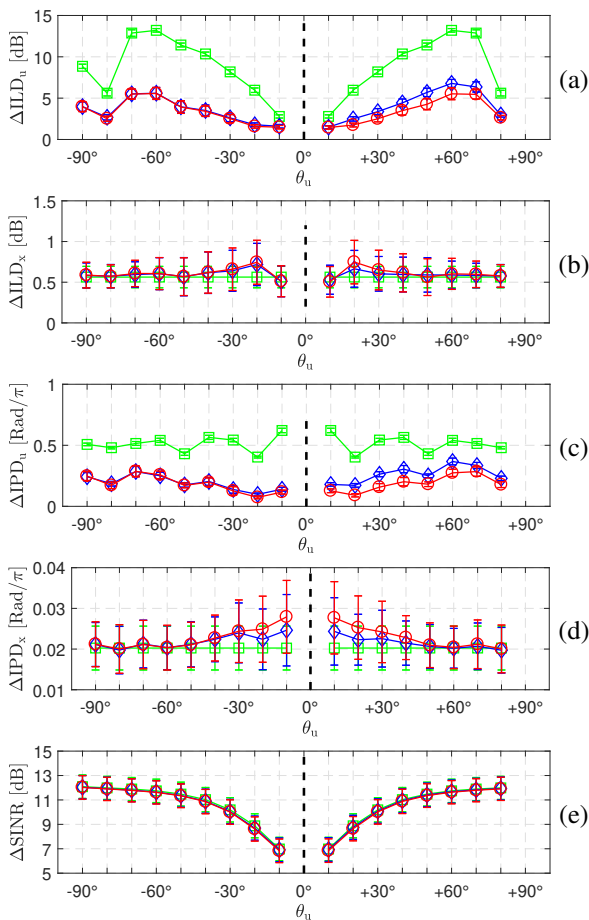


Fig. 1: Métricas de desempenho: (a) ΔILD_u , (b) ΔILD_x , (c) ΔIPD_u , (d) ΔIPD_x , (e) ΔSINR , para as seguintes técnicas: MWF ($-\square-$), MWF-ITF ($-\diamond-$) e CB-MWF-ITF ($-\circ-$).

V. CONCLUSÕES

Este trabalho apresentou uma solução fechada para o método de redução de ruído baseado no filtro de Wiener multicanal com preservação da ITF, para aparelhos auditivos binauriculares. A solução fechada foi encontrada através da relaxação dual de uma reformulação do problema de otimização original. Os resultados indicaram que o desempenho da solução proposta é equivalente ao apresentado pelo método MWF-ITF convencional, porém com um tempo de processamento 1300 vezes menor.

REFERÊNCIAS

- [1] Doclo, S., Gannot, S., Marquardt, D. e Hadad, E. “Binaural speech processing with application to hearing devices”. *Audio Source Separation and Speech Enhancement*, 413-442, 2018.
- [2] Blauert, J. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [3] Cornelis, B., Doclo, S., Bogaert, T., Moonen, M. e Wouters, J. “Theoretical analysis of binaural multimicrophone noise reduction techniques”. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(2), 342-355, 2009.
- [4] Doclo, S., Dong, R., Klasen, T.J., Wouters, J., Haykin, S., Moonen, M. “Extension of the multi-channel Wiener filter with localization cues for noise reduction in binaural hearing aids”. *IEEE Intl. Work. on Acoustic Echo and Noise Control (IWAENC)*, Sept. 2005, pp. 221-224.
- [5] Marquardt, D., Hohmann, V. e Doclo, S. “Coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids.” *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 8648-8652.
- [6] Klasen, T. J., et al. “Binaural multi-channel Wiener filtering for hearing aids: preserving interaural time and level differences”. *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. Vol. 5, 2006, pp. 145-148.
- [7] Adachi, S. e Nakatsukasa, Y. “Eigenvalue-based algorithm and analysis for nonconvex QCQP with one constraint.” *Mathematical Programming* 173.1-2, 2019, pp. 79-116.
- [8] Boyd, S. e Vandenberghe, L. “Convex optimization”. Cambridge University Press, 2004.
- [9] Golub, G. H. e Loan, C.F. “Matrix Computations”, terceira edição, Johns Hopkins University Press, 1996.
- [10] Vandenberghe, L. e Boyd, S. “Semidefinite programming.” *SIAM Review*, 1996, pp. 49-95.
- [11] Kayser, H., Ewert, S.D., Anemüller, J., Rohdenburg, T., Hohmann, V. e Kollmeier, B. “Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses”. *EURASIP Journal on Advances in Signal Processing*, 2009, pp. 1-10.
- [12] Henter, G.E., Merritt, T., Shannon, M., Mayo, C. e King, S. “Readme for REHASP 0.5: The repeated Harvard sentence prompts corpus version 0.5 [dataset]”. University of Edinburgh, 2014.
- [13] Hadad, E., Doclo, S. e Gannot, S. “The binaural LCMV beamformer and its performance analysis”. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(3), 543-558, 2016.
- [14] Habets, E.A., Cohen, I. e Gannot, S. “Generating nonstationary multi-sensor signals under a spatial coherence constraint”. *The Journal of the Acoustical Society of America*, 124(5), 2911-2917, 2008.

APÊNDICE

A. Algoritmo iterativo para obter valor de $\mathbf{w}_{\text{MWF-ITF}}$ e α .

O vetor $\tilde{\alpha} \in \mathbb{R}_{++}^G$ foi obtido para $G = 500$ elementos linearmente espaçados em uma escala logarítmica entre 10^{-10} e 1, de acordo com a Tabela I.

TABELA I: Algoritmo iterativo para cálculo de $\mathbf{w}_{\text{MWF-ITF}}$ e α .

Algoritmo: Busca iterativa para determinar α e \mathbf{w} para as frequências discretas $k = 1, \dots, N$.
Entrada: δ , ITF_{in}^u , $\tilde{\alpha}$, G , Φ_u , Φ_x , Φ_y .
Saída: $\mathbf{w}_{\text{MWF-ITF}}$.

Faça $e_{\text{ITF}} = \infty$
 Faça $i = 0$
 Inicialize $\mathbf{w} \in \mathbb{C}^{M \times 1}$ com valores aleatórios.
Enquanto $e_{\text{ITF}}^2 < \delta^2$ e $i < G$
 $i = i + 1$
 Fazendo $\alpha_i = \tilde{\alpha}(i)$, resolva o problema abaixo utilizando a rotina 'fmincon':
 $\mathbf{w}(\alpha_i, \lambda, k) = \arg.\min. J_{\text{MWF}}(\lambda, k) + \alpha_i J_{\text{ITF}}(\lambda, k)$
 $\mathbf{w}_{\text{UNC}} = \mathbf{w}(\alpha_i, \lambda, k)$
 $e_{\text{ITF}}^2 = J_{\text{ITF}}(\mathbf{w}_{\text{UNC}})$
Fim Enquanto
Retorne $\mathbf{w} = \mathbf{w}_{\text{UNC}}$
