

# Redução de Ruído em Implantes Cocleares utilizando Arranjos de Microfones

Luciana Menezes Xavier de Souza, Márcio Holsbach Costa e Renata Coelho Borges

**Resumo**— Implantes cocleares (IC) são dispositivos eletrônicos implantáveis que têm como finalidade a restauração da audição em pessoas com perda auditiva profunda. Este trabalho apresenta uma comparação entre duas estratégias de redução de ruído para IC. É analisado o caso de um arranjo de microfones sujeito à técnica de conformação de feixe baseada no método de variação mínima com resposta sem distorção e uma modificação, na qual é preservada apenas a potência da fala de interesse. Simulações computacionais associadas a critérios objetivos indicam que a técnica de preservação da potência resulta em maior inteligibilidade e redução de ruído, mostrando-se mais apropriada para IC.

**Palavras-Chave**— Arranjos de microfones, Redução de ruído, Implante coclear.

**Abstract**— Cochlear implants (CI) are surgically implanted electronic devices that aim to restore hearing in people with profound hearing loss. This work presents a comparison between two noise reduction strategies for CI. A microphone array subjected to the minimum variance distortionless response beamformer technique and a modification in which only the power of the speech of interest is preserved are analyzed. Computer simulations associated with objective criteria indicate that the power preservation technique results in greater intelligibility and noise reduction, showing to be more appropriate for CI.

**Keywords**— Microphone array, Noise reduction, Cochlear implant.

## I. INTRODUÇÃO

O sistema auditivo humano é responsável pela captura, transmissão e processamento de informação sonora. Essa informação é interpretada pelo cérebro, permitindo a compreensão da fala [1]. Pessoas que apresentam redução da capacidade auditiva a partir de aproximadamente 41 dB de nível de pressão sonora (*Sound Pressure Level* - SPL), em relação à audição normal (25 dB SPL), são consideradas pessoas com deficiência auditiva [2] [3].

A redução na capacidade de ouvir, ou perda auditiva, tem sido vista cada vez mais como um problema de saúde de grande importância, uma vez que pode acarretar no isolamento social do indivíduo [4], na perda da autonomia [5], em depressão [6] e até mesmo em disfunções neurocognitivas [2].

Segundo relatório da Organização Mundial de Saúde (OMS) [7], estima-se que até 2050 cerca de 2,5 bilhões de pessoas terão algum grau de perda auditiva e pelo menos 700 milhões precisarão de reabilitação. Nos casos em que o grau de perda auditiva é classificado entre severo e profundo a compensação da audição é realizada através de implantes cocleares (IC) [8].

O IC é uma prótese auditiva implantável desenvolvida com o objetivo de restaurar a percepção sonora. No entanto, o

implante só tem efeito se o nervo auditivo de pelo menos uma das orelhas do indivíduo estiver preservado e responder a estímulos elétricos [8]. Apesar de ser uma solução bem-sucedida em termos gerais, os ICs apresentam restrições práticas importantes, como o tamanho, consumo de energia e limitações computacionais [9]. Experimentos psicoacústicos têm revelado que as principais queixas dos usuários de IC estão relacionadas ao ruído externo, que dificulta a comunicação.

Os métodos mais abordados para a redução de ruído em ICs consistem em máscaras tempo-frequência monocanais, com destaque para a máscara binária [10] e filtro de Wiener [11]. Entretanto, arranjos de microfones (AM) têm sido cada vez mais utilizados para implementar filtros espaciais, de modo a enfatizar sinais sonoros provenientes de uma determinada direção, em geral da região frontal do usuário [12] [13].

Esse trabalho apresenta a comparação entre duas estratégias de redução de ruído em IC baseadas no conformador de feixe de variação mínima com resposta sem distorção (*Minimum Variance Distortionless Response* - MVDR), amplamente utilizado em aparelhos auditivos [14]-[17]. A implementação é realizada no domínio tempo, de acordo com a arquitetura convencional de ICs. São utilizadas duas estratégias de restrição: (i) a correlação entre o componente de fala na saída do IC e o componente de fala no microfone de referência é idêntica à autocorrelação do componente de fala no microfone de referência (estratégia convencional) [13]; e (ii) a autocorrelação do componente de fala na saída do IC é igual à autocorrelação do componente de fala no microfone de referência. O desempenho de ambas as técnicas é avaliado por meio de critérios objetivos de inteligibilidade e da razão sinal-interferência mais ruído.

O restante desse artigo está estruturado como segue: na Seção II os sinais e sistemas envolvidos são apresentados. Os métodos de redução de ruído abordados e os resultados das simulações computacionais são descritos nas Seções III e IV, respectivamente. Por fim, na Seção V são apresentadas as conclusões decorrentes deste trabalho.

Ao longo do texto, variáveis escalares são representadas por letras minúsculas em itálico, vetores por letras minúsculas em negrito, matrizes por letras maiúsculas em negrito e os símbolos  $(\cdot)^T$  e  $(\cdot)^{-1}$  representam as operações de transposição e inversão, respectivamente.

## II. SINAIS E SISTEMAS

O cenário abordado neste trabalho refere-se a um campo acústico monitorado por um conjunto de  $M$  microfones. Os sinais envolvidos originam-se por emissões de fontes sonoras pontuais e por ruído de fundo. A fala de interesse, captada em cada um dos microfones (Fig. 1), é representada por  $x_m(t)$ , em que  $m \in \{1, 2, \dots, M\}$ . O ruído aditivo  $v_m(t)$  é composto pelo ruído de fundo (elétrico e/ou decorrente de campo acústico difuso)  $r_m(t)$  e ruídos provenientes de fontes acústicas pontuais

Luciana Souza e Márcio Costa, PPGEEL-UFSC, Florianópolis-SC, e-mail: lucianamxs19@gmail.com, costa@eel.ufsc.br. Renata Borges, Depto de Engenharia Elétrica, UTFPR, Cornélio Procopio-PR, e-mail: recoelhoborges@gmail.com. Este trabalho foi financiado pelo CNPq, projetos 315020/2018-0, 142341/2019-3 e 302492/2021-6.

$f_{p,m}(t)$ , para  $p = 1, 2, \dots, P$ , em que  $p$  indica a fonte,  $P$  é o número de fontes interferentes e  $t$  é a variável de tempo contínuo.

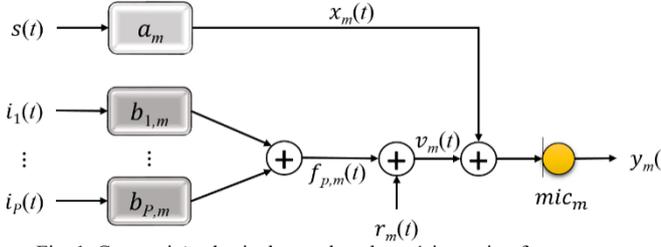


Fig. 1. Composição do sinal captado pelo  $m$ -ésimo microfone.

O sinal captado em cada um dos microfones é definido por:

$$y_m(t) = x_m(t) + v_m(t), \quad (1)$$

em que:

$$v_m(t) = \sum_{p=1}^P f_{p,m}(t) + r_m(t). \quad (2)$$

Os componentes da fala e das fontes interferentes em cada um dos microfones são definidos como:

$$x_m(t) = a_m(t) * s(t), \quad (3)$$

e

$$f_{p,m}(t) = b_{p,m}(t) * i_p(t). \quad (4)$$

em que  $*$  é o operador de convolução,  $s(t)$  é sinal de fala na posição da fonte geradora e  $a_m(t)$  e  $b_{p,m}(t)$  representam as respostas ao impulso associadas aos caminhos de propagação entre as fontes pontuais e os  $M$  microfones, respectivamente. As respostas ao impulso são conhecidas como resposta ao impulso relacionada à cabeça (*Head Related Impulse Response - HRIR*) e  $i_p(t)$  é o sinal da  $p$ -ésima fonte interferente (Fig. 1).

Neste trabalho é assumido que o IC está implantado na orelha direita do usuário, possuindo  $M/2$  microfones. Na orelha contralateral (esquerda) é colocado um molde auditivo com outros  $M/2$  microfones. Este molde é similar a um aparelho auditivo e tem como função apenas a captação do som. Os sinais dos microfones posicionados na orelha contralateral são enviados (por transmissão sem fio) para o processador posicionado na orelha do implante, como apresentado na Fig. 2. Em geral, o microfone frontal no lado implantado (direita neste caso) é chamado de microfone de referência.

O sinal de cada microfone é amostrado e processado por um banco de filtros para obtenção das componentes em  $K$  sub-bandas (Fig. 3). O sinal de entrada em cada um dos  $M$  bancos de filtros é definido por:

$$y_m(n) = x_m(n) + v_m(n), \quad (5)$$

em que  $n = t f_s$  é o índice de tempo discreto; e  $f_s$  é a frequência de amostragem. O processo de filtragem referente a cada sub-banda pode ser representado por um filtro de resposta finita ao impulso definido pelos coeficientes  $\{g_{k,0} \ g_{k,1} \ \dots \ g_{k,N-1}\}$ , em que  $k$  refere-se à sub-banda e  $N$  é o número de amostras com amplitude não desprezível [18]. Dessa forma, a saída de cada uma das sub-bandas, para cada um dos microfones, é dada por:

$$\mathbf{u}_{m,k}(n) = \mathbf{G}_k \mathbf{y}_m(n), \quad (6)$$

em que  $\mathbf{u}_{m,k}(n) = [u_{m,k}(n) \ u_{m,k}(n-1) \ \dots \ u_{m,k}(n-C+1)]^T$  possui dimensão  $C \times 1$ ;  $\mathbf{y}_m(n) = [y_m(n) \ y_m(n-1) \ \dots \ y_m(n-N-C+2)]^T$  possui dimensão  $(N+C-1) \times 1$ ;  $\mathbf{G}_k$  é uma matriz com os coeficientes do

banco de filtros, com dimensão  $C \times (N+C-1)$ . Dessa forma, a saída do arranjo de microfones é definida, a cada instante de tempo  $n$ , como:

$$z_k(n) = \sum_{m=1}^M \mathbf{u}_{m,k}^T(n) \mathbf{w}_{m,k} = \sum_{m=1}^M \mathbf{y}_m^T(n) \mathbf{G}_k^T \mathbf{w}_{m,k}. \quad (7)$$

em que  $\mathbf{w}_{m,k} = [w_{m,k,0} \ w_{m,k,1} \ \dots \ w_{m,k,C-1}]^T$  possui dimensão  $C \times 1$ .

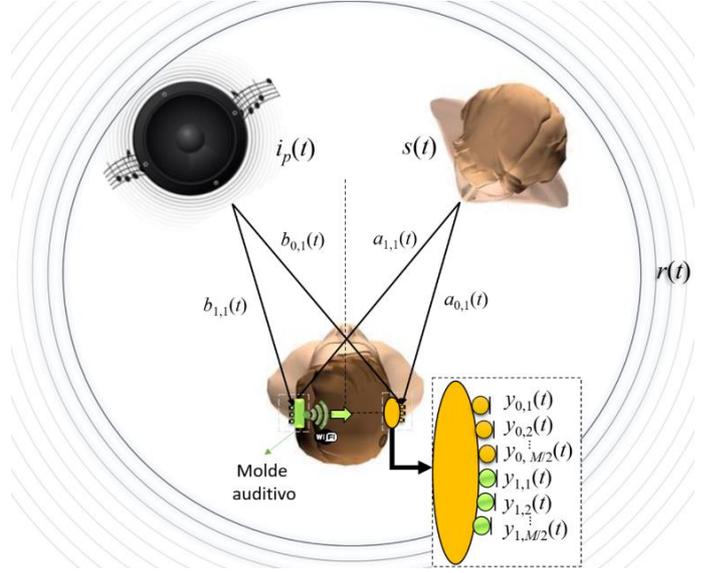


Fig. 2. Cenário acústico analisado. Considera-se o sinal de fala  $s(t)$ , o ruído interferente proveniente de uma fonte pontual  $i_p(t)$  e o ruído difuso  $r(t)$ . Assume-se que o implante esteja na orelha direita e a existência de um molde para captação e transmissão de sinais sonoros na orelha esquerda.

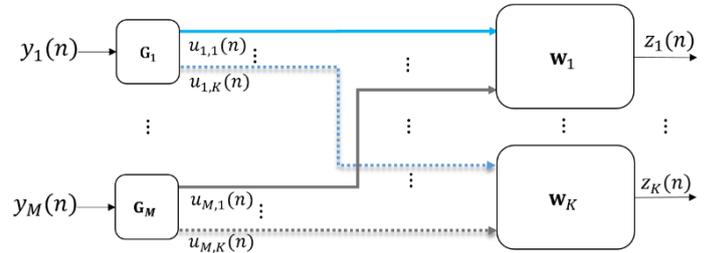


Fig. 3. Diagrama em blocos do arranjo de microfones e sistema de redução de ruído. São utilizados  $M$  microfones e processamento de  $K$  sub-bandas para estimulação.

A equação (7) pode ser expressa vetorialmente como:

$$z_k(n) = \mathbf{u}_k^T(n) \mathbf{w}_k, \quad (8)$$

em que  $\mathbf{w}_k = [\mathbf{w}_{1,k}^T \ \mathbf{w}_{2,k}^T \ \dots \ \mathbf{w}_{M,k}^T]^T$  e  $\mathbf{u}_k(n) = [\mathbf{u}_{1,k}^T(n) \ \mathbf{u}_{2,k}^T(n) \ \dots \ \mathbf{u}_{M,k}^T(n)]^T$  possuem dimensão  $MC \times 1$  e são, respectivamente, formados a partir do empilhamento dos vetores  $\mathbf{u}_{m,k}$  e  $\mathbf{w}_{m,k}$  para  $m = 1, 2, \dots, M$ . Portanto, o vetor  $\mathbf{u}_k(n)$  é definido como:

$$\mathbf{u}_k(n) = \mathbf{\Gamma}_k \mathbf{y}(n), \quad (9)$$

em que  $\mathbf{y}(n) = [\mathbf{y}_1^T(n) \ \mathbf{y}_2^T(n) \ \dots \ \mathbf{y}_M^T(n)]^T$  possui dimensão  $M(N+C-1) \times 1$  e  $\mathbf{\Gamma}_k$  é uma matriz bloco-diagonal de dimensão  $MC \times M(N+C-1)$  constituída em sua diagonal principal pela matriz  $\mathbf{G}_k$ . Assim:

$$z_k(n) = \mathbf{y}^T(n) \mathbf{\Gamma}_k^T \mathbf{w}_k. \quad (10)$$

### III. REDUÇÃO DE RUIDO

O sinal de saída do IC pode ser expresso em termos dos sinais de fala e ruído através da substituição de (5) em (10):

$$z_k(n) = \mathbf{x}^T(n) \mathbf{\Gamma}_k^T \mathbf{w}_k + \mathbf{v}^T(n) \mathbf{\Gamma}_k^T \mathbf{w}_k, \quad (11)$$

em que  $\mathbf{x}(n) = [\mathbf{x}_1^T(n) \mathbf{x}_2^T(n) \dots \mathbf{x}_M^T(n)]^T$  e  $\mathbf{v}(n) = [\mathbf{v}_1^T(n) \mathbf{v}_2^T(n) \dots \mathbf{v}_M^T(n)]^T$  possuem dimensão  $M(N+C-1) \times 1$ ; e  $\mathbf{x}_m(n) = [x_m(n) x_m(n-1) \dots x_m(n-N-C+2)]^T$  e  $\mathbf{v}_m(n) = [v_m(n) v_m(n-1) \dots v_m(n-N-C+2)]^T$  possuem dimensão  $(N+C-1) \times 1$ . A potência total na saída de cada sub-banda é dada por:

$$E\{z_k^2(n)\} = E\left\{\left[\mathbf{x}^T(n) \mathbf{\Gamma}_k^T \mathbf{w}_k + \mathbf{v}^T(n) \mathbf{\Gamma}_k^T \mathbf{w}_k\right]^2\right\}. \quad (12)$$

Considerando que a fala e o ruído são variáveis aleatórias independentes e de média zero, então  $E\{\mathbf{x}(n)\mathbf{v}^T(n)\} = E\{\mathbf{v}(n)\mathbf{x}^T(n)\} = \mathbf{0}$ , que leva (12) a:

$$E\{z_k^2(n)\} = \mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{w}_k + \mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{vv}} \mathbf{\Gamma}_k^T \mathbf{w}_k, \quad (13)$$

em que  $\mathbf{R}_{\mathbf{xx}}(n) = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$  e  $\mathbf{R}_{\mathbf{vv}}(n) = E\{\mathbf{v}(n)\mathbf{v}^T(n)\}$ .

A estratégia geral de redução de ruído em IC consiste na minimização da potência associada ao componente de ruído em (13), ou seja:

$$E\{z_k^2(n)\}_{|x_k(n)=0} = \mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{vv}} \mathbf{\Gamma}_k^T \mathbf{w}_k. \quad (14)$$

Visto que a solução trivial  $\mathbf{w}_k = \mathbf{0}$  também minimiza a potência da fala, é necessário estabelecer restrições de forma a preservar o sinal de interesse.

#### A. Preservação da Fala

A forma convencional utilizada para a preservação da fala é assumir, idealmente, que a saída do sistema associada ao componente de fala (como o sistema é linear então admite o princípio de superposição) deve corresponder ao sinal de fala no microfone de referência sem distorções, de forma que  $z_k^{fala}(n) = x_{\text{REF},k}(n)$ , portanto:

$$\mathbf{x}^T(n) \mathbf{\Gamma}_k^T \mathbf{w}_k = \mathbf{x}^T(n) \mathbf{\Gamma}_k^T \mathbf{q}_{\text{REF}}, \quad (15)$$

em que  $\mathbf{q}_{\text{REF}}$  é um vetor de dimensão  $MC \times 1$  cujos elementos são zeros, com exceção do primeiro. Como o sinal de interesse  $x(n)$  não é conhecido, é necessário estabelecer um artifício para superar tal limitação. Uma possibilidade é a utilização de momentos estatísticos da fala. Pré-multiplicando (15) por  $\mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{x}(n)$  e tomando-se o valor esperado obtém-se:

$$\mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{w}_k = \mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{q}_{\text{REF}}. \quad (16)$$

A equação (16) pode ser interpretada como a imposição de que a correlação entre o componente de fala na entrada e na saída do AM deve ser igual à autocorrelação do componente de fala na entrada. Como resultado, os coeficientes podem ser obtidos por um processo de otimização, de tal forma que:

$$\begin{aligned} \mathbf{w}_k^{\text{opt}} &= \arg \min_{\mathbf{w}_k} \mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{vv}} \mathbf{\Gamma}_k^T \mathbf{w}_k \\ \text{s.t. } &\mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{w}_k = \mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{q}_{\text{REF}}, \end{aligned} \quad (17)$$

em que ‘argmin’ significa ‘argumento que minimiza’ e ‘s.t.’ significa ‘sujeito a’.

A equação (17) corresponde ao conformador de feixe MVDR implementado no domínio tempo [19].

#### B. Preservação da Potência da Fala

Uma possibilidade alternativa de definição de restrição do processo de minimização de (14) é estabelecer a igualdade entre a autocorrelação do componente de fala na entrada e na saída do AM. Para tanto, eleva-se ao quadrado ambos os lados de (15) e toma-se seu valor esperado, de forma que:

$$\mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{w}_k = \mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{q}_{\text{REF}}. \quad (18)$$

A equação (18) tem como objetivo preservar a potência do componente de fala na saída do AM, ou seja,  $E\{[z_k^{\text{fala}}(n)]^2\} = E\{x_{\text{REF},k}^2(n)\}$ . Utilizando-se (14) e (18), pode-se definir o problema de otimização dos coeficientes do sistema de redução de ruído como:

$$\begin{aligned} \mathbf{w}_k^{\text{opt}} &= \arg \min_{\mathbf{w}_k} \mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{vv}} \mathbf{\Gamma}_k^T \mathbf{w}_k \\ \text{s.t. } &\mathbf{w}_k^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{w}_k = \mathbf{q}_{\text{REF}}^T \mathbf{\Gamma}_k \mathbf{R}_{\mathbf{xx}} \mathbf{\Gamma}_k^T \mathbf{q}_{\text{REF}}. \end{aligned} \quad (19)$$

Diferentemente de (17), a identidade em (19) não impõe restrições de fase ao vetor de coeficientes, aumentando as possibilidades para minimização da função custo. Essa metodologia (amplamente utilizada em outras aplicações [14] [19] [20] [21] [22] [23]) vem ao encontro das características intrínsecas do implante coclear que despreza a estrutura fina do sinal, entregando ao usuário apenas informações relativas à envoltória do sinal acústico [9][10]. Surpreendentemente, uma extensa revisão da literatura indica que esse enfoque não tem sido utilizado na área de arranjos de microfones em aplicações de IC, apesar de sua aderência à aplicação em questão.

### IV. SIMULAÇÕES COMPUTACIONAIS

Nesta seção é realizada uma comparação entre o método convencional de redução de ruído em arranjos de microfones, descrito na equação (17), e o método descrito em (19). Os resultados apresentados na forma gráfica obedecem à seguinte regra de cores: sinal contaminado em cinza; método convencional (equação (17)) em azul; e método descrito na equação (19) em verde.

As simulações foram realizadas utilizando respostas ao impulso do tipo HRIR relativas a um escritório do banco de dados descrito em [24] (ambiente ‘Office I’, com tempo de reverberação de 300 ms). As informações foram obtidas a partir de um manequim simulador de torso e cabeça (Bruel & Kjaer type 4128-C) com um aparelho auditivo de três microfones em cada orelha. Arbitrou-se a fonte de fala no azimute de  $0^\circ$  (em frente à linha de visada do manequim) e a fonte de ruído em  $-60^\circ$  (lado esquerdo). Foram utilizados 20 áudios de fala humana foneticamente balanceadas do banco de dados em português [25] e um ruído interferente do tipo fala artificial (*artificial speech-like noise* ICRA-1) [26].

Os parâmetros de simulação utilizados foram:  $M = 6$ ,  $C = 1$ ,  $N = 200$ , razão sinal-ruído (SNR) de 30 dB, e razão sinal-interferência (SIR) entre  $-10$  dB e 10 dB, variando de 5 em 5 dB. Estimativas das matrizes de correlação  $\mathbf{R}_{\mathbf{xx}}(n)$  e  $\mathbf{R}_{\mathbf{vv}}(n)$  foram obtidas diretamente a partir dos sinais de fala e ruído de forma a obter um limite de desempenho superior e permitir comparações com outras referências da literatura [27] [28]. Em situações práticas é utilizado um detector de fala para estimação de  $\mathbf{R}_{\mathbf{vv}}(n)$  e então  $\mathbf{R}_{\mathbf{xx}}(n) = \mathbf{R}_{\mathbf{yy}}(n) - \mathbf{R}_{\mathbf{vv}}(n)$

Para a avaliação dos resultados foi utilizada a razão sinal-interferência mais ruído (*Signal to Interference plus Noise Ratio* - SINR) [29] para estabelecer uma relação entre as potências da

fala e do ruído. A razão de energia de modulação entre fala e reverberação para implantes cocleares (*Speech to Reverberation Modulation Energy Ratio for Cochlear Implants* - SRMR-CI) [30] e a medida de inteligibilidade objetiva de tempo curto (*Short Term Objective Intelligibility* - STOI) [31] foram utilizadas para quantificar a inteligibilidade.

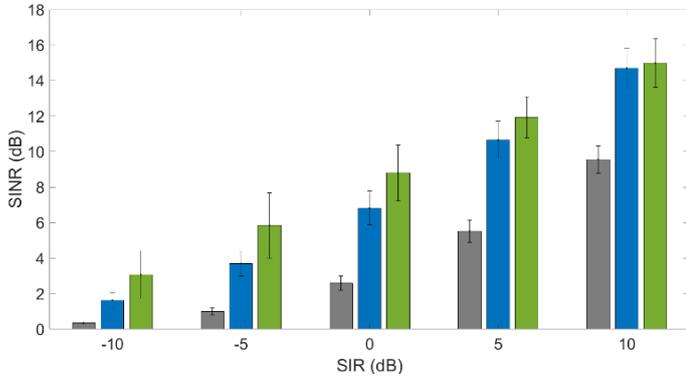


Fig. 4. SINR de saída. Sinal contaminado (cinza); equação (17) (azul); equação (19) (verde).

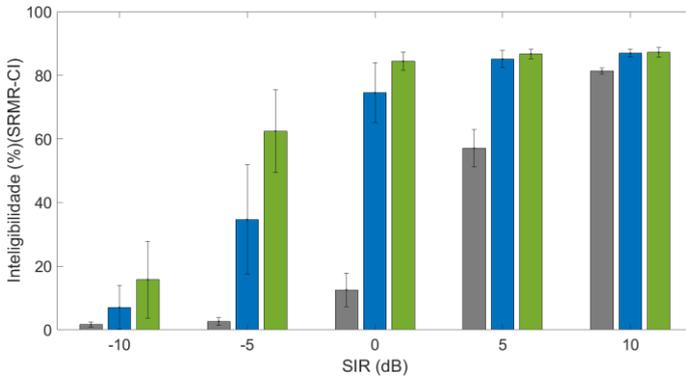


Fig. 5. SRMR-CI percentual. Sinal contaminado (cinza); equação (17) (azul); equação (19) (verde).

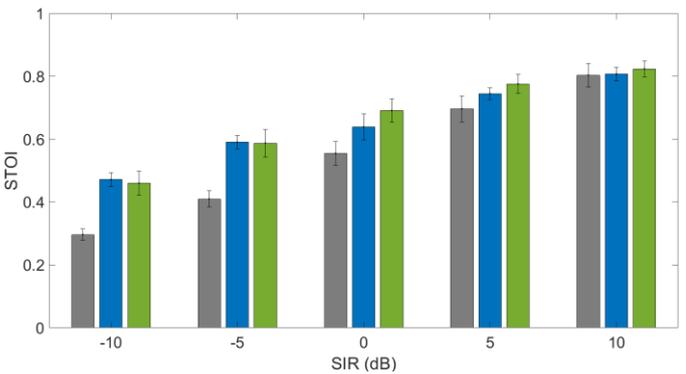


Fig. 6. STOI. Sinal contaminado (cinza); equação (17) (azul); equação (19) (verde).

As Figuras 4 a 6 apresentam gráficos de barras com os resultados médios para as três métricas indicadas. As hastes em preto apresentam a informação de desvio padrão.

A análise da SINR (Fig. 4) evidencia que o método apresentado em (19) permite uma maior redução de interferência/ruído para todas as SIR avaliadas. Esse resultado corrobora a expectativa inicial de que a liberação da fase do conjunto de restrições permitiria novas possibilidades para o conjunto de coeficientes de forma a obter melhor desempenho.

A inteligibilidade percentual obtida por meio da métrica SRMR-CI (Fig. 5) indica melhor desempenho do método apresentado na equação (19) para todas as SIRs, principalmente para valores menores ou iguais a 0 dB. A SRMR-CI possui um coeficiente de correlação com experimentos psicoacústicos maior que 0,9.

Os resultados obtidos pela métrica STOI são apresentados na Fig. 6. O método descrito na equação (19) apresenta maior inteligibilidade para  $0 \leq SIR \leq 10$ , com desempenho inferior (mas muito próximo) ao do método convencional apenas para SIR  $\{-10, -5\}$  dB. Para verificar tal afirmação aplicou-se o teste Kruskal-Wallis [32], a hipótese nula é que não existe diferença estatística para as duas metodologias nas SIRs negativas. Para um nível de significância de  $\alpha = 4,52 \cdot 10^{-6}$  nenhum dos grupos de dados tem classificações médias significativamente diferentes.

Nas 10 estimações de inteligibilidade realizadas (5 SIRs para STOI e 5 SIRs para SRMR-CI) 8 indicaram melhor desempenho para o método descrito na equação (19). Apesar da alta correlação dos critérios de STOI e SRMR-CI com a inteligibilidade em pessoas implantadas é importante lembrar que o primeiro foi criado originalmente para quantificar a inteligibilidade em normo-ouvintes enquanto que o segundo foi especificamente criado para implantados. Dessa forma, os resultados de SRMR-CI devem retratar como maior fidelidade o desempenho dos métodos estudados em testes psicoacústicos com pessoas implantadas.

A Fig. 7 apresenta diagramas bidimensionais de caixas mostrando a relação entre inteligibilidade percentual obtida via SRMR-CI e STOI para SIR = 0 dB. Nesses diagramas, os círculos preenchidos na cor preta indicam a mediana, e as barras se referem ao primeiro e terceiro quartis. Os valores discrepantes (*outliers*) foram omitidos por mais clareza. Nessa figura a vantagem da técnica descrita na equação (19), para SIR = 0 dB, é melhor visualizada. Futuramente, experimentos psicoacústicos devem ser realizados para corroborar as evidências obtidas neste trabalho.

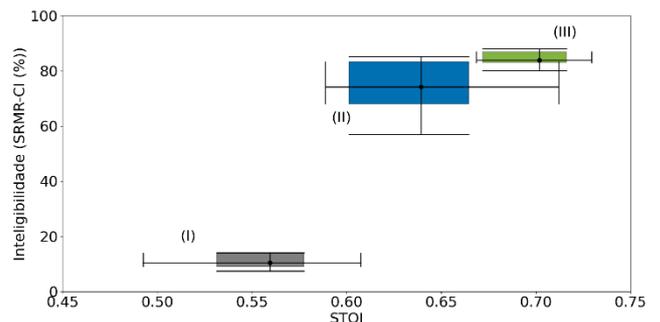


Fig. 7. Diagrama bidimensional de caixas para SRMR-CI percentual versus STOI. Sinal contaminado (cinza (I)); equação (17) (azul (II)); equação (19) (verde (III)). SNR = 30 dB e SIR = 0 dB.

## V. CONCLUSÕES

Este trabalho apresentou uma comparação de desempenho entre dois métodos de redução de ruído utilizando arranjos de microfones em implantes cocleares. Uma variação do método de conformação de feixe baseada na variação mínima com resposta sem distorção (MVDR), ainda não explorado em aplicações de implante coclear no domínio do tempo, resultou em maior razão sinal-interferência mais ruído e indícios de

maior inteligibilidade através do critério objetivo SRMR-CI. Os resultados sugerem que o método proposto apresenta maior aderência conceitual ao problema de implantes cocleares que o método MVDR convencional amplamente utilizado, em decorrência da falta de informação de estrutura fina dos sinais pelos sistema de estimulação do IC.

#### REFERÊNCIAS

- [1] J. Blauert. *The Technology of Binaural Listening*. Springer Science & Business Media, 2013.
- [2] N. S. Reed, A. Altan, J. A. Deal, C. Yeh, A. D. Kravetz, M. Wallhagen e F. R. Lin, “Trends in health care costs and utilization associated with untreated hearing loss over 10 years”. *JAMA Otolaryngology–Head & Neck Surgery*, v. 145, n. 1, pp. 27-34, 2019.
- [3] ANSI. American National Standard Institute. *American National Standard Specification for Audiometers (ANSI S3.6)*. New York: ANSI, 1989.
- [4] A. Maharani, N. Pendleton e I. Leroi, “Hearing impairment, loneliness, social isolation, and cognitive function: longitudinal analysis using English longitudinal study on ageing”, *The American Journal of Geriatric Psychiatry*, v. 27, n. 12, pp. 1348-1356, 2019.
- [5] M. H. Sonnet, B. Montaut-Verient, J. Y. Niemier, M. Hoen, L. Ribeyre e C. Parietti-Winkler, “Cognitive abilities and quality of life after cochlear implantation in the elderly”, *Otology & Neurotology*, v. 38, n. 8, pp. 296-301, 2017.
- [6] R. Manrique-Huarte, D. Calavia, A. H. Irujo, L. Girón e M. Manrique-Rodríguez, “Treatment for hearing loss among the elderly: auditory outcomes and impact on quality of life”, *Audiology and Neurotology*, v. 21, n. 1, pp. 29-35, 2016.
- [7] OMS. *World Report on hearing*. Disponível: <https://apps.who.int/iris/bitstream/handle/10665/339906/9789240021501-eng.pdf?>. Acesso em: 19 de Jan. 2023.
- [8] W. H. Shapiro e T. S. Bradham, “Cochlear implant programming”, *Otolaryngologic Clinics of North America*, v. 45, n. 1, pp. 111-127, 2012.
- [9] F. G. Zeng, S. Rebscher, W. Harrison, X. Sun e H. Feng, “Cochlear implants: System design, integration, and evaluation”, *IEEE Reviews in Biomedical Engineering*, vol. 1, pp. 115-142, 2008.
- [10] D. Wang e G. J. Brown, *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*. New York, NY: Wiley/IEEE press, 2006.
- [11] P. C. Loizou, *Speech Enhancement: Theory and practice*. Segunda edição. CRC Press, 2013.
- [12] B. D. Van Veen e K. M. Buckley, “Beamforming: A versatile approach to spatial filtering”, *IEEE ASSP Magazine*, v. 5, n. 2, pp. 4-24, 1988.
- [13] J. Benesty, J. Chen e Y. Huang. *Microphone Array Signal Processing*, Springer-Verlag, 2008.
- [14] J. Bourgeois e W. Minker, *Time-domain beamforming and blind source separation: Speech input in the car environment*, Springer, 2009.
- [15] Z. Sun, Y. Li, H. Jiang, F. Chen e Z. Wang, “A MVDR-MWF combined algorithm for binaural hearing aid system”, *IEEE Biomedical Circuits and Systems Conference*, pp. 1-4, 2018.
- [16] R. Ali, T. Van Waterschoot e M. Moonen, “Integration of a priori and estimated constraints into an MVDR beamformer for speech enhancement”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 27, n. 12, pp. 2288-2300, 2019.
- [17] W. J. L. Malaver e M. H. Costa, “Worst-case-optimization robust-MVDR beamformer for stereo noise reduction in hearing aids”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 28, pp. 2224-2237, 2020.
- [18] M. Slaney, “Auditory toolbox”, Interval Research Corporation, Technical Report, v. 10, n. 1998, pp. 1194, 1998.
- [19] S. Gannot, E. Vincent, S. Markovich-Golan e A. Ozerov, “A consolidated perspective on multi-microphone speech enhancement and source separation”, *IEEE/ACM Transactions on Audio, Speech and Language Processing*, v. 25, n. 4, pp. 692-730, 2017.
- [20] D. Ward, R. Kennedy e R. Williamson, “Theory and design of broadband sensor arrays with frequency invariant far field beam patterns”, *The Journal of the Acoustical Society of America*, v. 97, n. 2, pp. 1023-1034, 1995.
- [21] O. L. Frost, “An algorithm for linearly constrained adaptive array processing”, *Proceedings of the IEEE*, v. 60, n. 8, pp. 926-935, 1972.
- [22] S. Markovich, S. Gannot e I. Cohen, “Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals”, *IEEE Transactions on Audio, Speech and Language Processing*, v. 17, n. 6, pp. 1071-1086, 2009.
- [23] K. M. Buckley, “Spatial/spectral filtering with linearly constrained minimum variance beamformers”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 35, n. 3, pp. 249-266, 1987.
- [24] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann e B. Kollmeier. “Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses”, *EURASIP Journal on Advances in Signal Processing*, v. 2009, pp. 1-10, 2009.
- [25] A. Alcaim, J. A. Solewicz e J. A. Moraes, “Frequência de ocorrência dos fonemas e listas de frases foneticamente balanceadas no português falado no Rio de Janeiro”, *Revista da Sociedade Brasileira de Telecomunicações*, v. 7, n. 1, pp. 23-41, 1992.
- [26] W. A. Dreschler, H. Verschuure, C. Ludvigsen e S. Westermann, “ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment”, *Audiology*, v. 40, n. 3, pp. 148-157, 2001.
- [27] R. A. Chiea, M. H. Costa e J. A. Cordioli, “An optimal envelope-based noise reduction method for cochlear implants: An upper bound performance investigation”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 29, pp. 1729-1739, 2021.
- [28] L. M. X. Souza, M. H. Costa e R. C. Borges, “Noise reduction by beamforming in cochlear implants: A time domain study”, *CLAIB & CBEB*, pp. 1-6, 2022.
- [29] G. Zhang, Q. Li, Q. Zhang, J. Qin e L. Yang, “Signal-to-interference-plus-noise ratio-based multi-relay beamforming for multi-user multiple-input multiple-output cognitive relay networks with interference from primary network”, *IET Communications*, v. 9, n. 2, pp. 227-238, 2015.
- [30] J. F. Santos e T. H. Falk, “Updating the SRMR-CI metric for improved intelligibility prediction for cochlear implant users”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 22, n. 12, pp. 2197-2206, 2014.
- [31] C. H. Taal, R. C. Hendriks, R. Heusdens e J. Jensen, “An algorithm for intelligibility prediction of time-frequency weighted noisy speech”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 19, n. 7, pp. 2125-2136, 2011.
- [32] E. Ostertagova; O. Ostertag; J. Kováč. “Methodology and application of the Kruskal-Wallis test”, *Applied mechanics and materials*, v. 611, pp. 115-120, 2014.