

# Modelagem da Limitação da Taxa de Comunicação em Aparelhos Auditivos Biauriculares

Vitor P. Curtarelli e Márcio H. Costa

**Resumo**— Este trabalho apresenta uma nova proposta de modelagem dos efeitos de quantização e saturação decorrentes da redução da taxa de comunicação em aparelhos auditivos biauriculares. É analisado o comportamento do filtro de Wiener multicanal em aplicações de redução de ruído acústico. Comparações com o modelo clássico utilizado na literatura são apresentadas em termos de razão sinal-ruído, qualidade e inteligibilidade da fala processada nas saídas dos aparelhos auditivos. Os resultados obtidos indicam que o modelo proposto apresenta aumento da acurácia de predição, em especial em relação à razão sinal-ruído e à qualidade da fala.

**Palavras-Chave**— *Aparelho auditivo, Redução de ruído, Taxa de comunicação.*

**Abstract**— This work presents a new proposal for modeling the quantization and saturation effects resulting from the reduction of the communication rate in binaural hearing aids. The behavior of the multichannel Wiener filter noise-reduction method is approached. Comparisons with the classic model used in the literature are presented in terms of signal-to-noise ratio, quality, and intelligibility of the processed speech at the hearing aid outputs. The obtained results indicate that the proposed model presents increased prediction accuracy, especially for signal-to-noise ratio and speech quality.

**Keywords**— *Hearing aids, Noise reduction, Communication rate.*

## I. INTRODUÇÃO

O sistema auditivo humano tem como função principal o processamento da informação sonora, permitindo não apenas a comunicação, como também a percepção do cenário acústico. A perda total ou parcial da audição pode ser um fator limitante para a vida do indivíduo [1]. Segundo a OMS, 5% da população mundial sofre algum problema auditivo e se prevê que esse percentual poderá aumentar para 10% até 2050 [2]. Estima-se que cerca de 77% dessas pessoas poderiam ser beneficiadas pelo uso de aparelhos auditivos [3].

Aparelhos auditivos são dispositivos cujo objetivo é a compensação de limitações na audição, podendo até mesmo restaurar por completo a capacidade de comunicação do usuário, na ausência de ruído acústico no ambiente. Entretanto, grande parte das situações diárias não é caracterizada dessa forma, o que exige a utilização de técnicas de redução de ruído para assegurar uma inteligibilidade adequada.

Atualmente, uma das tecnologias mais avançadas de aparelhos auditivos é chamada de biauricular. Nela, são utilizados dois aparelhos auditivos, um em cada orelha, que se comunicam através de transmissão sem fio [4] [5]. Apesar de

apresentarem maior eficiência na redução de ruído aditivo e na capacidade de preservação do cenário acústico original, possuem como grande desvantagem um maior consumo de energia devido ao processo de transferência de dados, diminuindo a duração das baterias e conseqüentemente o tempo entre recargas [6] [7], o que acaba sendo inconveniente para o usuário.

A principal estratégia utilizada para minimizar o gasto de energia em aparelhos biauriculares é a redução da taxa de comunicação entre os aparelhos. Entretanto, essa estratégia pode diminuir o desempenho do processo de redução de ruído. Dessa forma, o desenvolvimento de modelos teóricos para a predição de desempenho do processo de redução de ruído, em função da taxa de comunicação, é de grande interesse.

A principal forma de modelagem para os erros de quantização foi proposta originalmente por Bernard Widrow [8], baseando-se em [9], e posteriormente refinada em [10], [11] e [12]. Nela, a quantização é aproximada por um ruído aditivo uniforme. Contudo, as condições para que essa modelagem seja bem-sucedida são bastante restritas e não são adequadas para sistemas de redução de ruído com limitação de comunicação, resultando em erros.

Neste trabalho é apresentada uma nova proposta de modelagem dos efeitos de quantização e seu desempenho na predição do comportamento do filtro de Wiener multicanal (MWF) biauricular. Como resultado, é apresentado um novo modelo para a predição dos coeficientes do MWF e realizada sua comparação com o modelo clássico utilizado na literatura.

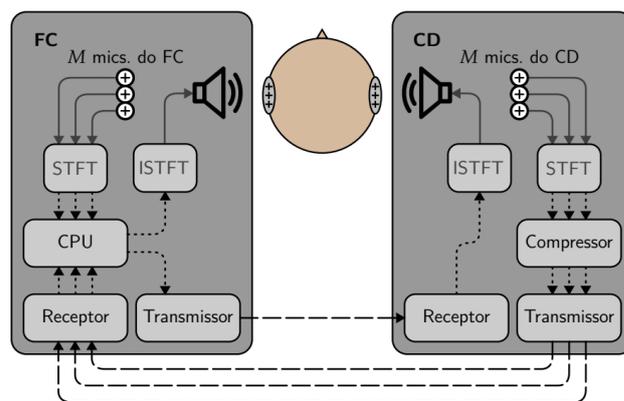


Fig. 1. Diagrama em blocos de um aparelho auditivo biauricular.

## II. REDUÇÃO DE RUÍDO EM APARELHOS AUDITIVOS BIAURICULARES

Nesta seção, é apresentada a descrição do problema abordado e sua solução a partir do modelo clássico. Vetores e

matrizes são representados, respectivamente, por letras minúsculas e maiúsculas em negrito, e escalares em itálico.

### A. Aparelho Auditivo Biauricular

A Fig. 1 apresenta o diagrama em blocos de um aparelho auditivo biauricular. Considera-se que um dos aparelhos auditivos é chamado de centro de fusão (FC), sendo responsável pelo processamento dos sinais, enquanto o outro é denominado de dispositivo contralateral (CD) [13], sendo sua função transmitir os sinais acústicos captados por seus microfones para o FC, e receber (e reproduzir) o sinal enviado pelo FC. Cada aparelho possui  $M$  microfones para aquisição dos sinais sonoros.

Os sinais acústicos captados pelos microfones são transformados para o domínio frequência, de forma que  $y_m(\lambda, k) = x_m(\lambda, k) + r_m(\lambda, k)$ ; em que  $\lambda$  representa a janela de tempo;  $k$  o bin de frequência;  $m$  o índice do microfone;  $r$  o ruído global;  $y$  a fala contaminada; e  $x$  a fala de interesse, definida como  $x_m(\lambda, k) = a_m(k)s(\lambda, k)$ , em que  $s(\lambda, k)$  é a fala na posição da fonte geradora e  $a_m(k)$  é a função transferência relativa à cabeça (HRTF) entre a fonte geradora da fala e o  $m$ -ésimo microfone.

Em um sistema sem limitações na taxa de comunicação, os sinais adquiridos podem ser organizados na forma vetorial como  $\mathbf{y}(\lambda, k) = [y_1(\lambda, k) \dots y_M(\lambda, k) y_{M+1}(\lambda, k) \dots y_{2M}(\lambda, k)]^T$  e, de forma semelhante, para  $\mathbf{x}(\lambda, k)$  e  $\mathbf{r}(\lambda, k)$ , em que  $[\cdot]^T$  é o operador transposto. A primeira metade dos sinais de  $\mathbf{y}(\lambda, k)$  corresponde aos sinais do FC, e a segunda aos do CD. Neste trabalho considera-se que  $\mathbf{x}(\lambda, k)$  e  $\mathbf{r}(\lambda, k)$  são descorrelacionados e de média nula [14].

### B. Redução da Taxa de Comunicação

A taxa de comunicação entre os aparelhos auditivos depende do número de bits transmitidos por segundo. Mantendo a frequência de amostragem constante, a diminuição da taxa é realizada através da redução do número de bits usado para quantizar os dados.

Neste trabalho assume-se o uso da quantização uniforme *midread quantization* [11]. Assume-se que para um número pré-determinado de  $N$  bits utiliza-se um nível para 0,  $2^{N-1}-1$  para números positivos e o mesmo para números negativos. Dessa forma, tem-se  $2^N-1$  possíveis níveis. Embora um dos níveis seja desperdiçado, isso garante a presença do nível 0 e da existência de simetria no processo de saturação. Denomina-se  $\mathcal{N} = 2^{N-1}-1$  o número de níveis positivos/negativos e  $\Delta$  o passo de quantização. O valor de saturação do quantizador é dado por  $V = \Delta \cdot \mathcal{N}$ . Dessa forma, o processo de quantização de um valor real arbitrário  $v$  pode ser descrito como

$$Q_{\Delta, \mathcal{N}}(v) = \Delta \sum_{i=1}^{\mathcal{N}} \left[ u\left(v - \left(i - \frac{1}{2}\right)\Delta\right) - u\left(-v - \left(i - \frac{1}{2}\right)\Delta\right) \right], \quad (1)$$

onde  $u(\cdot)$  é a função degrau unitário. No caso de valores complexos ( $v = v^{\mathbb{R}} + jv^{\mathbb{I}}$ ) utiliza-se a quantização complexa  $Q_{\Delta, \mathcal{N}}^{\mathbb{C}}(\cdot)$ , definida por  $Q_{\Delta, \mathcal{N}}^{\mathbb{C}}(v) = Q_{\Delta, \mathcal{N}}(v^{\mathbb{R}}) + jQ_{\Delta, \mathcal{N}}(v^{\mathbb{I}})$ , em que  $(\cdot)^{\mathbb{R}}$  e  $(\cdot)^{\mathbb{I}}$  correspondem às partes real e imaginária do argumento, respectivamente. Neste trabalho considera-se que todos os quantizadores do CD têm os mesmos  $\Delta$  e  $N$ , embora essa restrição seja facilmente removível.

### C. Filtro de Wiener Multicanal

O MWF é um estimador ótimo amplamente estudado na redução de ruído em aparelhos auditivos [15]. Ele consiste no processo de filtragem linear dos sinais captados pelos microfones, resultando nos seguintes sinais de excitação dos alto-falantes (antes do processo de reconstrução no domínio tempo):

$$\hat{x}_{\text{FC}}(\lambda, k) = \mathbf{w}_{\text{FC}}^H(\lambda, k) \mathbf{y}_q(\lambda, k), \quad (2)$$

$$\hat{x}_{\text{CD}}(\lambda, k) = \mathbf{w}_{\text{CD}}^H(\lambda, k) \mathbf{y}_q(\lambda, k), \quad (3)$$

em que  $\hat{x}_{\text{FC}}(\lambda, k)$  e  $\hat{x}_{\text{CD}}(\lambda, k)$  são os sinais de saída do FC e CD, respectivamente;  $\mathbf{w}_{\text{FC}}$  e  $\mathbf{w}_{\text{CD}}$  são os respectivos vetores de coeficientes;  $(\cdot)^H$  é o operador conjugado-transposto; e  $\mathbf{y}_q(\lambda, k)$  é o vetor de sinais de entrada quantizados disponível no FC. Os coeficientes ótimos do MWF são obtidos através do processo de otimização dado por

$$[\mathbf{w}_{\text{FC}}^{\text{otim}} \ \mathbf{w}_{\text{CD}}^{\text{otim}}] = \min_{\mathbf{w}_{\text{FC}} \ \mathbf{w}_{\text{CD}}} J(\lambda, k), \quad (4)$$

em que

$$J(\lambda, k) = E\{ |x_{\text{FC}}(\lambda, k) - \hat{x}_{\text{FC}}(\lambda, k)|^2 \} + E\{ |x_{\text{CD}}(\lambda, k) - \hat{x}_{\text{CD}}(\lambda, k)|^2 \} \quad (5)$$

e  $E\{\cdot\}$  é o operador valor esperado. Assume-se que a transmissão de  $\hat{x}_{\text{CD}}(\lambda, k)$  do FC ao CD (para reconstrução no domínio tempo e excitação do alto-falante) não sofre efeitos relevantes de quantização. Pode ser mostrado que a solução do problema descrito em (4) é

$$\mathbf{w}_{\text{FC}}^{\text{otim}}(\lambda, k) = \mathbf{R}_{\mathbf{y}_q, \mathbf{y}_q}^{-1}(\lambda, k) \mathbf{R}_{\mathbf{x}_q, \mathbf{y}_q}(\lambda, k) \mathbf{q}_{\text{FC}}, \quad (6)$$

$$\mathbf{w}_{\text{CD}}^{\text{otim}}(\lambda, k) = \mathbf{R}_{\mathbf{y}_q, \mathbf{y}_q}^{-1}(\lambda, k) \mathbf{R}_{\mathbf{x}_q, \mathbf{y}_q}(\lambda, k) \mathbf{q}_{\text{CD}}, \quad (7)$$

em que  $\mathbf{R}_{\mathbf{y}_q, \mathbf{y}_q}(\lambda, k) = E\{\mathbf{y}_q(\lambda, k) \mathbf{y}_q^H(\lambda, k)\}$  é a matriz de coerência do vetor  $\mathbf{y}_q$ ;  $\mathbf{R}_{\mathbf{x}_q, \mathbf{y}_q}(\lambda, k) = E\{\mathbf{x}(\lambda, k) \mathbf{y}_q^H(\lambda, k)\}$  é a matriz de coerência cruzada entre  $\mathbf{x}(\lambda, k)$  e  $\mathbf{y}_q(\lambda, k)$ ;  $\mathbf{q}_{\text{FC}}$  é o vetor seletor do sinal de referência do FC ( $x_{\text{FC}}(\lambda, k) = \mathbf{q}_{\text{FC}}^T \mathbf{x}(\lambda, k)$ ), composto por um único valor igual a 1, e os demais nulos; e  $\mathbf{q}_{\text{CD}}$  é definido de forma semelhante à  $\mathbf{q}_{\text{FC}}$  (com  $x_{\text{CD}}(\lambda, k) = \mathbf{q}_{\text{CD}}^T \mathbf{x}(\lambda, k)$ ).

### D. Modelagem das Matrizes de Coerência

Considerando  $\Delta$  suficientemente pequeno [10] ou o uso de *dithering* aditivo [11] [12] e  $N$  tal que não ocorra saturação [10], é possível assumir que  $\mathbf{y}_q(\lambda, k) \cong \mathbf{y}(\lambda, k) + \mathbf{e}(\lambda, k)$ , em que  $\mathbf{e}(\lambda, k) = [0 \dots 0 \ e_{M+1}(\lambda, k) \dots e_{2M}(\lambda, k)]^T$  é o vetor erro de quantização, cujos  $M$  últimos elementos (referentes ao CD) são variáveis aleatórias com distribuição uniforme, branca, não-correlacionadas com  $\mathbf{y}(\lambda, k)$  e variância  $\sigma_e^2(\lambda, k) = E\{|e_c(\lambda, k)|^2\} \cong \Delta^2/6$ . Assim,

$$\mathbf{R}_{\mathbf{y}_q, \mathbf{y}_q}(\lambda, k) \cong \mathbf{R}_{\mathbf{y}, \mathbf{y}}(\lambda, k) + \mathbf{R}_{\mathbf{e}, \mathbf{e}}(\lambda, k), \quad (8)$$

$$\mathbf{R}_{\mathbf{x}_q, \mathbf{y}_q}(\lambda, k) \cong \mathbf{R}_{\mathbf{x}, \mathbf{y}}(\lambda, k), \quad (9)$$

visto que  $\mathbf{R}_{\mathbf{x}_q, \mathbf{y}_q}(\lambda, k) = \mathbf{R}_{\mathbf{x}, \mathbf{y}}(\lambda, k)$  em função da inexistência de correlação entre  $\mathbf{x}(\lambda, k)$  e  $\mathbf{r}(\lambda, k)$ .

Utilizando-se (8) e (9) em (6) e (7) obtêm-se estimativas dos coeficientes ótimos do MWF e, portanto, predições do desempenho do processo de redução de ruído para diferentes taxas de comunicação. A partir desses resultados o projetista pode realizar inferências sobre o compromisso desejado entre consumo de energia (número de bits) e qualidade/inteligibilidade do sinal processado.

### III. PROPOSTA DE UM NOVO MODELO PARA O PROCESSO DE REDUÇÃO DA TAXA DE COMUNICAÇÃO

Esta seção apresenta um novo modelo para geração das matrizes de coerência  $\mathbf{R}_{y_q}(\lambda, k)$  e  $\mathbf{R}_{y_q y_q}(\lambda, k)$  necessárias para o cálculo dos vetores de coeficientes descritos em (6) e (7).

#### A. Aproximação da Quantização e Saturação

Dada a característica descontínua de  $u(\cdot)$  em (1), o que inviabiliza a obtenção de suas derivadas, utilizou-se a aproximação

$$u(v) = \lim_{\sigma^2 \rightarrow 0} [1 + f_\sigma(v)] / 2, \quad (10)$$

em que  $f_\sigma(\cdot)$  é a função erro modificada, dada por [16]

$$f_\sigma(v) = \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \int_0^v \exp\{-t / (2\sigma^2)\} dt, \quad (11)$$

sendo  $v$  o valor a ser quantizado e  $t$  uma variável auxiliar. A Fig. 2 apresenta uma comparação entre (1) original e (1) utilizando a aproximação descrita em (10) para  $\sigma = 0.05$ ,  $\Delta = 1/3$  e  $N = 3$ .

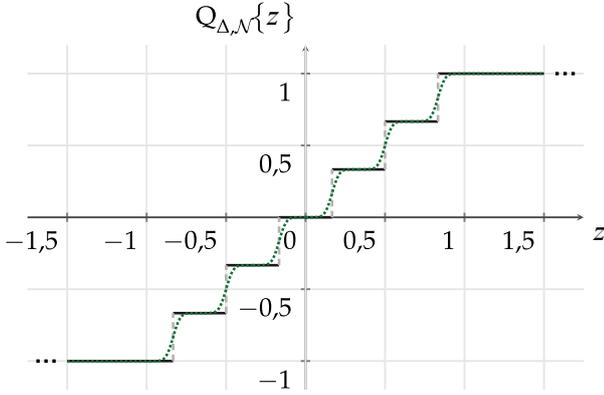


Fig. 2. Comparação entre (1) e sua aproximação utilizando (10) e (11) para  $\sigma = 0.05$ ,  $\Delta = 1/3$  e  $N = 3$ .

#### B. Estimação de $\mathbf{R}_{y_q}$

Considerando duas variáveis aleatórias complexas arbitrárias  $v$  e  $z$ , pela definição de  $Q_{\Delta, N}^C(\cdot)$  pode-se escrever que

$$\begin{aligned} E\{v[Q_{\Delta, N}^C(z)]^*\} &= E\{v^{\mathbb{R}} Q_{\Delta, N}(z^{\mathbb{R}})\} + jE\{v^{\mathbb{I}} Q_{\Delta, N}(z^{\mathbb{R}})\} \\ &\quad - jE\{v^{\mathbb{R}} Q_{\Delta, N}(z^{\mathbb{I}})\} + E\{v^{\mathbb{I}} Q_{\Delta, N}(z^{\mathbb{I}})\} \end{aligned} \quad (12)$$

Assumindo que as partes real e imaginária de  $v$  e  $z$  possuem distribuição Gaussiana de média nula e utilizando a aproximação do processo de quantização definida pelas equações (1), (10) e (11), os termos do lado direito de (12) podem ser calculados através do Teorema Modificado de Price [17], de forma que [16]:

$$E\{v^{\mathbb{A}} Q_{\Delta, N}(z^{\mathbb{B}})\} = \frac{\sqrt{2\Delta} E\{v^{\mathbb{A}} z^{\mathbb{B}}\}}{\sigma_{z^{\mathbb{B}}} \sqrt{\pi}} \sum_{i=1}^N \exp\left\{-\frac{\left(i - \frac{1}{2}\right)^2 \Delta^2}{\sigma_{z^{\mathbb{B}}}^2}\right\} \quad (13)$$

em que  $\mathbb{A}, \mathbb{B} \in \{\mathbb{R}, \mathbb{I}\}$  e  $\sigma_{z^{\mathbb{B}}} = E\{(z^{\mathbb{B}})^2\}^{1/2}$ . Utilizando-se (13) em (12) chega-se em:

$$\mathbf{R}_{y_q}(\lambda, k) = \mathbf{\Gamma}(\lambda, k) \odot \mathbf{R}_{xx}(\lambda, k), \quad (14)$$

em que  $\odot$  representa o produto Hadamard,

$$\mathbf{\Gamma}(\lambda, k) = \begin{bmatrix} 1 & \cdots & 1 & \gamma_{M+1}(\lambda, k) & \cdots & \gamma_{2M}(\lambda, k) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & \gamma_{M+1}(\lambda, k) & \cdots & \gamma_{2M}(\lambda, k) \end{bmatrix}, \quad (15)$$

$$\gamma_c(\lambda, k) = \frac{2\Delta}{\sigma_{y_c}(\lambda, k) \sqrt{\pi}} \sum_{i=1}^N \exp\left\{-\frac{\left(i - \frac{1}{2}\right)^2 \Delta^2}{\sigma_{y_c}^2(\lambda, k)}\right\}, \quad (16)$$

$$\text{e } \sigma_{y_c}^2(\lambda, k) = E\{|y_c(\lambda, k)|^2\}.$$

#### C. Estimação de $\mathbf{R}_{y_q y_q}$

Procedendo-se de maneira semelhante à Seção III.B pode-se obter

$$\begin{aligned} E\{Q_{\Delta, N}^C(v) [Q_{\Delta, N}^C(z)]^*\} &= \\ E\{Q_{\Delta, N}(v^{\mathbb{R}}) Q_{\Delta, N}(z^{\mathbb{R}})\} &- jE\{Q_{\Delta, N}(v^{\mathbb{R}}) Q_{\Delta, N}(z^{\mathbb{I}})\} \\ + jE\{Q_{\Delta, N}(v^{\mathbb{I}}) Q_{\Delta, N}(z^{\mathbb{R}})\} &+ E\{Q_{\Delta, N}(v^{\mathbb{I}}) Q_{\Delta, N}(z^{\mathbb{I}})\} \end{aligned} \quad (17)$$

Realizando as mesmas considerações e suposições da Seção III.B e utilizando-se o Teorema Modificado de Price pode-se demonstrar que [18]

$$E\{Q_{\Delta, N}(v^{\mathbb{A}}) Q_{\Delta, N}(z^{\mathbb{B}})\} = \frac{\Delta^2 E\{v^{\mathbb{A}} z^{\mathbb{B}}\}}{\pi \sigma_v \sigma_z} \sum_{i=1}^N \sum_{j=1}^N \int_{-1}^1 \Xi_{i,j}(\Lambda_{ij}, \rho, \alpha) d\alpha, \quad (18)$$

em que

$$\Xi_{i,j}(\Lambda_{ij}, \rho, \alpha) = |\mathbf{K}_\alpha|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \Lambda_{ij}^T \mathbf{K}_\alpha^{-1} \Lambda_{ij}\right\}, \quad (19)$$

$$\mathbf{K}_\alpha = \begin{bmatrix} 1 & \rho\alpha \\ \rho\alpha & 1 \end{bmatrix}, \quad (20)$$

$$\Lambda_{i,j} = \begin{bmatrix} (2i-1)\Delta & (2j-i)\Delta \\ 2\sigma_v & 2\sigma_z \end{bmatrix}^T, \quad (21)$$

$$\rho = \frac{E\{v^{\mathbb{A}} z^{\mathbb{B}}\}}{\sigma_v \sigma_z}, \quad (22)$$

$\sigma_{v^{\mathbb{A}}} = E\{(v^{\mathbb{A}})^2\}^{1/2}$  (o mesmo para  $\sigma_{z^{\mathbb{B}}}$ ) e  $\rho = E\{v^{\mathbb{A}} z^{\mathbb{B}}\} / (\sigma_{v^{\mathbb{A}}} \sigma_{z^{\mathbb{B}}})$  é complexo com  $|\rho| \leq 1$ . Substituindo-se (18) em (17) e desenvolvendo, chega-se em:

$$\begin{aligned} \mathbf{R}_{y_q y_q}(\lambda, k) &= \mathbf{Y}^S(\lambda, k) \odot \mathbf{R}_{yy}(\lambda, k) \\ &\quad + \mathbf{Y}^D(\lambda, k) \odot \mathbf{R}_{yy}^H(\lambda, k) \end{aligned} \quad (23)$$

em que

$$\mathbf{Y}^S = \frac{\mathbf{Y}^{\mathbb{R}} + \mathbf{Y}^{\mathbb{I}}}{2}, \quad (24)$$

$$\mathbf{Y}^D = \frac{\mathbf{Y}^{\mathbb{R}} - \mathbf{Y}^{\mathbb{I}}}{2}, \quad (25)$$

$$\mathbf{Y}^{\mathbb{A}} = \begin{bmatrix} 1 & \cdots & 1 & \gamma_{M+1} & \cdots & \gamma_{2M} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & \gamma_{M+1} & \cdots & \gamma_{2M} \\ \gamma_{M+1} & \cdots & \gamma_{M+1} & \nu_{M+1, M+1}^{\mathbb{A}} & \cdots & \nu_{M+1, 2M}^{\mathbb{A}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{2M} & \cdots & \gamma_{M+1} & \nu_{2M, M+1}^{\mathbb{A}} & \cdots & \nu_{2M, 2M}^{\mathbb{A}} \end{bmatrix}, \quad (26)$$

$$\nu_{i,c}^{\mathbb{A}} = \frac{2\Delta^2}{\pi \sigma_{v^{\mathbb{A}}} \sigma_{z^{\mathbb{C}}}} \sum_{i=1}^N \sum_{j=1}^N \int_{-1}^1 \Xi_{i,j}(\Lambda_{ij}, \rho^{\mathbb{A}}, \alpha) d\alpha. \quad (27)$$

#### IV. SIMULAÇÕES COMPUTACIONAIS

Nesta seção são apresentados os métodos utilizados para a avaliação do modelo desenvolvido.

##### A. Cenário acústico

O cenário acústico simulado foi construído emulando-se um ambiente anecoico, conforme descrito em [19]. Cada dispositivo do aparelho auditivo possui 3 microfones. Assumindo que o centro da cabeça do usuário está posicionado na origem do sistema de coordenadas, a fonte de interesse foi posicionada em (80 cm, 0°, 0°); ou seja, no plano horizontal médio, diretamente na frente do usuário. As falas utilizadas são provenientes de um indivíduo do gênero feminino [20]. A fonte interferente foi posicionada em (300 cm, 0°, -60°) (lado esquerdo). Esse azimute foi escolhido de forma que os sinais captados pelo FC apresentassem maior contaminação pela fonte interferente que os do CD. O sinal utilizado para a interferência foi um ruído ICRA [21], que apresenta características espectrais semelhantes às da voz, com razão sinal-interferência (SIR) de 0 dB (média entre os dois microfones frontais de referência).

Foi adicionado ruído branco Gaussiano a todos os microfones de forma a obter razão sinal-ruído (SNR – *signal to noise ratio*) de 27 dB, com o objetivo de emular a influência de ruído ambiental difuso e ruído eletrônico. Os sinais de fala contaminada, representados no domínio tempo, foram normalizados na faixa (-1,+1) e foram analisadas quantizações para  $3 \leq N \leq 9$  bits. Foi avaliado  $V \in \{2, 20\}$  (no domínio frequência), representando, respectivamente, casos com muito e pouco efeito da saturação.

Os sinais foram amostrados com frequência de amostragem de 16 kHz. A representação no domínio da frequência foi obtida através da transformada de Fourier de tempo curto (STFT) com uma janela de tempo de 256 amostras e sobreposição de 50%. A reconstrução foi realizada através da transformada inversa de Fourier de tempo curto (ISTFT) seguida do método de *overlap and add* [22].

##### B. Métricas de avaliação

A avaliação de desempenho do modelo proposto foi realizada utilizando-se três critérios objetivos: razão sinal-interferência-ruído-e-quantização (SINQR), qualidade da fala e inteligibilidade.

A avaliação da SINQR foi realizada através da estimação da SNR *a posteriori* [23]. A avaliação da qualidade foi realizada através da PESQ [24] e a da inteligibilidade utilizando a STOI [25]. Esses critérios foram escolhidos por apresentarem elevada correlação com experimentos psicoacústicos com voluntários [14].

Para cada um dos critérios objetivos, a métrica de avaliação apresentada nos gráficos é o erro absoluto médio entre o critério objetivo calculado para o sinal contaminado filtrado pelo filtro de Wiener utilizando: (a) matrizes de coerência provenientes do modelo proposto (equações (14) e (23)) ou clássico (equações (8) e (9)); e (b) matrizes de coerência estimadas a partir de simulações numéricas (160 realizações) com sinais sujeitos à redução da taxa de comunicação (quantização e saturação).

#### V. RESULTADOS

Os resultados das simulações descritas na seção anterior são apresentados na Fig. 3 até Fig. 5. Os eixos das abscissas indicam o número de bits da quantização, enquanto que os das ordenadas o erro absoluto médio para cada um dos critérios analisados. Linhas azuis com xis representam o erro associado ao modelo proposto neste trabalho, enquanto que linhas laranjas com triângulos o erro utilizando o modelo clássico de Widrow [8]. Linhas contínuas referem-se ao caso sem saturação ( $V = 20$ ) e tracejadas ao caso com saturação ( $V = 2$ ).

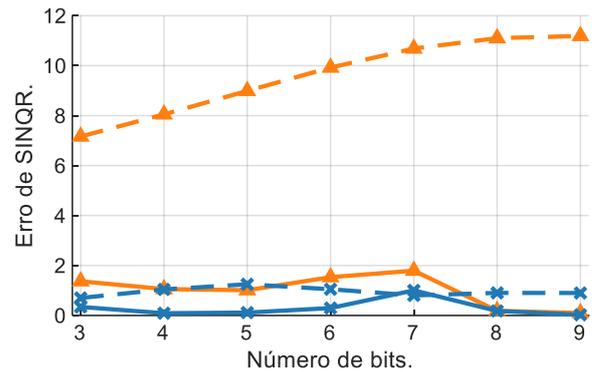


Fig. 3. Erro de SINQR para o modelo proposto (azul) e clássico (laranja).  $V = 20$  (linha contínua) e  $V = 2$  (linha tracejada).

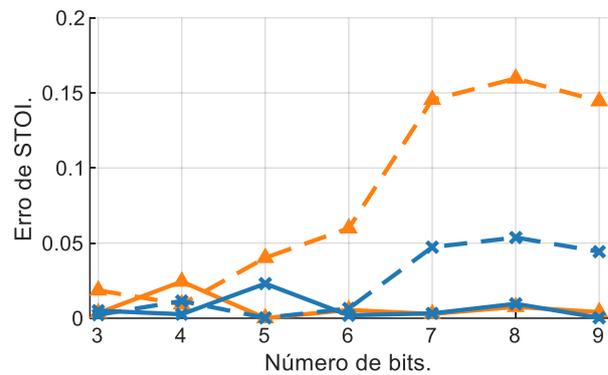


Fig. 4. Erro de STOI para o modelo proposto (azul) e clássico (laranja).  $V = 20$  (linha contínua) e  $V = 2$  (linha tracejada).

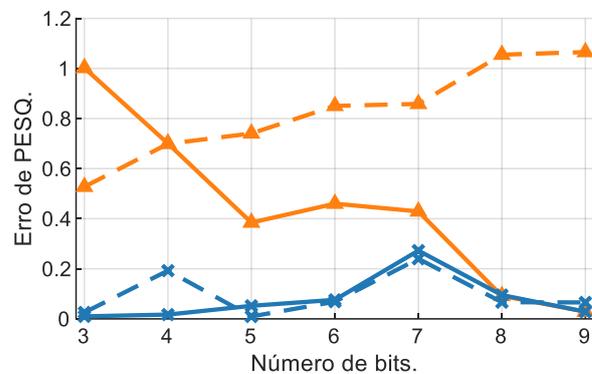


Fig. 5. Erro de PESQ para o modelo proposto (azul) e clássico (laranja).  $V = 20$  (linha contínua) e  $V = 2$  (linha tracejada).

### A. Resultados de SINQR

A Fig. 3 apresenta o erro de SINQR. O modelo proposto produz um erro menor que o resultante da aplicação do modelo clássico para ambos os valores do limiar de saturação. Para  $V=20$  ambos os modelos apresentam erros menores que 2 dB; entretanto, a diferença é muito maior para o caso  $V=2$ , visto que o modelo clássico não leva em consideração a possibilidade de saturação.

### B. Resultados de STOI

A Fig. 4 apresenta os resultados de STOI. Para a situação sem saturação ( $V=20$ ), o modelo clássico possui desempenho semelhante ao do modelo proposto e, portanto, o modelo proposto não apresenta vantagens. Entretanto, na existência de influência de saturação ( $V=2$ ) o modelo proposto apresenta menor erro.

### C. Resultados de PESQ

A Fig. 5 apresenta o resultado de PESQ. No caso desse critério são encontradas maiores discrepâncias. Considerando-se um pequeno número de bits ( $<8$ ), mesmo no caso sem saturação, o modelo clássico apresenta valores elevados de erro, visto que variações de PESQ maiores que 0,2 são perceptíveis em experimentos psicoacústicos com voluntários [26].

### D. Avaliação Global

Os resultados apresentados representam uma amostra de um conjunto amplo de simulações que evidenciam a vantagem da utilização do modelo proposto em relação ao modelo clássico, principalmente em relação ao critério PESQ.

Todas as simulações indicaram que na presença de saturação o modelo clássico apresenta significativa limitação na capacidade de predição do comportamento real do filtro de Wiener. No caso em que os efeitos de saturação são desprezíveis, a inteligibilidade pode ser predita com elevada acurácia por qualquer um dos modelos.

O modelo clássico se apresenta como o mais adequado ao uso pelo projetista de sistemas de redução de ruído em aparelhos auditivos apenas em situações em que o critério de interesse for a inteligibilidade, devido ao baixo custo computacional.

## VI. CONCLUSÕES

Neste trabalho foi apresentada uma nova proposta de modelagem dos efeitos da redução da taxa de comunicação sobre o desempenho do filtro de Wiener multicanal em aparelhos auditivos biaurais. A motivação deste estudo é o melhor entendimento dos efeitos da quantização e saturação no desempenho do processo de redução de ruído. Simulações computacionais indicam que no caso de ocorrência de saturação o modelo proposto apresenta aumento significativo da acurácia de predição em relação ao modelo clássico, em termos de razão sinal-ruído, qualidade da fala e inteligibilidade. No caso de inexistência de efeitos de saturação, o modelo proposto apresenta melhor capacidade de predição em termos de razão sinal-ruído e qualidade, tendo um desempenho semelhante ao do modelo clássico para a inteligibilidade.

## REFERÊNCIAS

- [1] R. Einhorn, "Hearing aid technology for the 21st Century: A proposal for universal wireless connectivity and improved sound quality", *IEEE Pulse*, v. 8, n. 2, pp. 25-28, 2017.

- [2] OMS, "Deafness and hearing loss", <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>. Acessado em 24/01/2023
- [3] NIDCD, "Quick statistics about hearing", <https://www.nidcd.nih.gov/health/cochlear-implants>. Acessado em 24/01/2023
- [4] G. A. Day, G. G. Browning e S. Gatehouse, "Benefit from binaural hearing aids in individuals with a severe hearing impairment", *British Journal of Audiology*, v. 22, pp. 273-277, 1988.
- [5] P. Avan, F. Giraudet e B. Buki, "Importance of binaural hearing", *Audiology & Neuro-otology*, v. 20, pp. 3-6, 2015.
- [6] S. H. Lotterman, R. N. Kasten e D. M. Majerus, "Battery life and nonlinear distortion in hearing aids", *Journal of Speech and Hearing Disorders*, v. 3, pp. 82-118, 1967.
- [7] F. Z. Sachs, P. K. Weathersby e J. Tufts, "Model for estimating life-cycle costs associated with noise-induced hearing loss", Naval Submarine Medical Research Laboratory Technical Report, n. 1248, pp. 1-62, 2007.
- [8] B. Widrow, "A study of rough amplitude quantization by means of Nyquist sampling theory", *IRE Transactions on Circuit Theory*, v. 3, n. 4, pp. 266-276, 1956.
- [9] W. R. Bennett, "Spectra of quantized signals", *The Bell System Technical Journal*, v. 27, n. 3, pp. 446-472, 1948.
- [10] A. Sripad e D. Snyder, "A necessary and sufficient condition for quantization errors to be uniform and white", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 25, n. 5, pp. 442-448, 1977.
- [11] S. P. Lipshitz, R. A. Wannamker e J. Vanderkooy, "Quantization and dither: A theoretical survey", *Journal of the Audio Engineering Society*, v. 40, n. 5, pp. 355-375, 1992.
- [12] R. A. Wannamaker et al. "A theory of nonsubtractive dither", *IEEE Transactions on Signal Processing*, v. 48, pp. 499-516, 2000.
- [13] J. Amini et al., "Rate-constrained noise reduction in wireless acoustic sensor networks", *IEEE Transactions on Audio, Speech, and Language Processing*, v. 28, pp. 1-12, 2020.
- [14] P. C. Loizou, *Speech enhancement: Theory and practice*, 2<sup>nd</sup> Ed. CRC Press, 2013.
- [15] J. Chen et al., "New insights into the noise reduction Wiener filter", *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 4, pp. 1218-1234, 2006.
- [16] J. Shynk e N. J. Bershad, "Steady-state analysis of a single-layer perceptron based on a system identification model with bias terms", *IEEE Transactions on Circuits and Systems*, v. 38, n. 9, pp. 1030-1042, 1991.
- [17] R. A. Pawula, "A modified version of Price's theorem", *IEEE Transactions on Information Theory*, v. 13, n. 2, pp. 285-288, 1967.
- [18] N. J. Bershad, pp. Celka e J.-M. Vesin, "Stochastic analysis of gradient adaptive identification of nonlinear systems with memory for Gaussian data and noisy input and output measurements", *IEEE Transactions on Signal Processing*, v. 47, n. 3, pp. 675-689, 1999.
- [19] H. Kayser et al. "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.
- [20] G. Henter et al., "Measuring the perceptual effects of modelling assumptions in speech synthesis using stimuli constructed from repeated natural speech", *Proceedings of the Interspeech*, pp. 1504-1508, 2014.
- [21] W. Dreschler et al. "ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing aid assessment", *Audiology*, v. 40, n. 3, pp. 148-157, 2001.
- [22] R. E. Crochiere, "A weighted overlap-add method of short-time Fourier analysis/synthesis", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. ASSP-28, n. 1, pp. 99-102, 1980.
- [23] M. Borga, Canonical correlation: A tutorial, [https://www.cs.cmu.edu/~tom/10701\\_sp11/slides/CCA\\_tutorial.pdf](https://www.cs.cmu.edu/~tom/10701_sp11/slides/CCA_tutorial.pdf). Acessado em 14/08/2022.
- [24] A. Rix et al., "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs", *Proceedings of ICASSP*, v. 2, pp. 749-752, 2001.
- [25] C. Taal et al., "A short-time objective intelligibility measure for time-frequency weighted noisy speech", *Proceedings of ICASSP*, pp. 4214-4217, 2010.
- [26] J. C. Servetti e J. C. Martin, "Error tolerant MAC extension for speech communications over 802.11 WLANs", *Proceedings of IEEE Vehicular Technology Conference*, v. 4, pp. 2330-2334, 2005.