# Análise de Robustez de um Método de Redução de Ruído para Implantes Cocleares

Rafael Attili Chiea e Márcio Holsbach Costa

*Resumo*— Recentemente, um novo método de redução de ruído, denominado C2F, foi desenvolvido especificamente para implantes cocleares (IC). Esse método resulta em maior inteligibilidade da fala em usuários de IC, em comparação com o filtro de Wiener (WF), quando os coeficientes são calculados de forma ideal. Em aplicações reais, os coeficientes são obtidos a partir de estimativas da razão sinal-ruído (SNR) e, portanto, acarretam perda do desempenho ótimo. Neste trabalho é realizada uma análise preliminar da robustez do método C2F a erros de estimação da SNR. Simulações numéricas corroboram a análise teórica mostrando que, apesar de apresentar diminuição de desempenho na redução de ruído, o C2F é menos sensível que o WF a erros de estimação da SNR.

## Palavras-Chave— Redução de ruído, Robustez, Implante coclear.

*Abstract*— Recently, a new noise reduction method, named C2F, was specifically developed for cochlear implants (IC). It results in greater speech intelligibility for IC users, as compared to the Wiener filter (WF), when the coefficients are ideally calculated. In real applications the coefficients are calculated from signal-to-noise ratio (SNR) estimates and, therefore, result a loss in the optimum performance. In this work, a robustness analysis of the C2F with respect to SNR errors was performed. Numerical simulations corroborate the derived theoretical analysis showing that, despite a performance decrease, the C2F is less sensitive than WF to SNR estimation errors.

#### Keywords—Noise reduction, Robustness, Cochlear implant.

#### I. INTRODUÇÃO

Implantes cocleares (IC) são as mais bem-sucedidas interfaces máquina-cérebro já desenvolvidas [1]. São próteses auditivas que possibilitam que pessoas com perda auditiva severa ou profunda percebam sons complexos e apresentem elevada inteligibilidade da fala humana [2]. Aparelhos auditivos têm como saída um sinal acústico. Por sua vez, nos implantes cocleares, o nervo auditivo é diretamente estimulado com pulsos de corrente elétrica aplicados através de um conjunto de 16 a 24 eletrodos posicionados no interior da cóclea [1] [2].

Na maior parte dos implantes cocleares contemporâneos, as estratégias de estimulação são baseadas em *vocoders* [1] [3]. De maneira simplificada, estratégias desse tipo processam o sinal o sinal captado pelo microfone de aquisição da seguinte maneira: o sinal de áudio é processado por um banco de filtros, resultando em distintas saídas, ou canais, associados a diferentes bandas de frequência. Em cada canal, são calculadas as envoltórias temporais dos sinais que modulam a amplitude de trens de pulsos, geralmente com taxa constante. Estes, por sua vez, acionam fontes de corrente elétrica conectadas aos diversos eletrodos. Vale notar que, consequentemente, não é utilizada a informação de fase de cada canal, também chamada de estrutura temporal fina. Dessa forma, há restrições significativas nas informações disponibilizadas ao usuário do dispositivo, em relação a normouvintes.

Mesmo com essas limitações de resolução temporal e espectral, em cenários sem a presença de ruído, a inteligibilidade da fala por parte de usuários de IC se aproxima à de ouvintes normais [4]. Contudo, na presença de ruído e reverberação, esse desempenho cai significativamente [5] [6], o que motiva a utilização de técnicas de redução de ruído. As técnicas de redução de ruído propostas para ICs baseiam-se, em maior parte, em máscaras tempo-frequência, dentre as quais se destacam a máscara binária (BM) [7] e o filtro de Wiener (WF) [8]. Mesmo proporcionando ganhos substanciais de inteligibilidade da fala em usuários de IC [9], essas máscaras não foram desenvolvidas especificamente para essa aplicação, e, assim, em seu desenvolvimento teórico, não levam em consideração as particularidades do dispositivo.

Recentemente foi proposto um filtro de redução de ruído aditivo para implantes cocleares, baseado na minimização do erro quadrático médio entre os quadrados das envoltórias do sinal estimado e da fala desejada [10]. Este filtro, chamado doravante de filtro de Chiea-Costa (C2F), é definido no domínio do tempo e aplicado em cada canal do banco de filtros, podendo ser interpretado como uma máscara tempofrequência. Simulações numéricas e um experimento psicoacústico piloto indicaram que o C2F resulta em melhores índices de inteligibilidade da fala por usuários de IC quando comparado com o WF, cujo objetivo é minimizar o erro quadrático médio entre o sinal estimado e a fala desejada.

Máscaras tempo-frequência, como a BM, o WF e o C2F necessitam da informação da razão sinal-ruído (SNR) *a priori* instantânea para a estimação de seus coeficientes de supressão [11]. Para a demonstração do limite superior de desempenho de cada técnica é normal a realização de simulações e/ou experimentos controlados nos quais exista o conhecimento prévio perfeito da SNR [7] [9] [10]. Entretanto, em aplicações reais, o conhecimento perfeito da SNR não é passível de ser obtido e os coeficientes dos filtros devem ser estimados, por exemplo, através de métodos de aprendizagem de máquina [12]. Alternativamente, podem ser utilizadas estimativas das potências individuais da fala e do ruído através do método de decisão-direta [13] [14] ou de métodos baseados na coerência entre dois microfones próximos [15].

Neste trabalho, é apresentada uma primeira abordagem para a caracterização da robustez do C2F a erros de estimação da SNR *a priori*. Como elemento de comparação será utilizado o WF. A seção II contém uma análise teórica da sensibilidade das funções de supressão associadas ao WF e ao C2F. Na seção III

Rafael Attili Chiea e Márcio Holsbach Costa, Universidade Federal de Santa Catarina, e-mails: rafaelchiea@gmail.com, costa@eel.ufsc.br. Este trabalho foi parcialmente financiado pelo CNPq, projeto 315020/2018-0.

são apresentadas simulações numéricas demonstrando o impacto de erros na estimação da SNR *a priori* no cálculo dos coeficientes dos filtros e, consequentemente, na inteligibilidade de usuários de IC. Finalmente, a seção IV apresenta a conclusão deste trabalho.

#### II. ANÁLISE TEÓRICA

A fala contaminada por ruído aditivo é representada na forma y(n) = x(n)+v(n), em que x(n) é o sinal de fala de interesse, v(n) é o ruído e *n* é o tempo discreto. Ambos x(n) e v(n) são considerados como não-observáveis individualmente e descorrelacionados entre si. O sinal y(n) é processado por um banco de k = 1,...,M (em que *M* é o número de canais) filtros, apresentando na saída de cada canal o sinal  $y_k(n) = x_k(n)+v_k(n)$ limitado em banda.

O sinal de fala em cada canal é estimado através de um processo de filtragem definido por:

$$\hat{x}_{k}(n) = c_{k}(n)y_{k}(n)$$
, (1)

em que  $c_k(n)$  é o coeficiente do filtro de estimação e  $\hat{x}_k(n)$  é a estimativa do sinal de interesse de  $x_k(n)$ .

O coeficiente  $c_k(n)$  pode ser interpretado como uma máscara tempo-frequência, com valores no intervalo [0,1], que tem como função atenuar canais com baixa SNR ( $c_k(n) \le 1$ ) e preservar canais com alta SNR ( $c_k(n) \ge 1$ )

Os coeficientes do filtro podem ser calculados de diferentes maneiras, de acordo com critérios de projeto. No caso do WF, o objetivo é minimizar o erro quadrático entre o sinal estimado e a fala de interesse. Seus coeficientes são dados por:

$$c_k(n) = w_k(n) = \frac{\xi_k(n)}{\xi_k(n) + 1} , \qquad (2)$$

em que  $\xi_k(n) = \sigma_{xk}^2/\sigma_{vk}^2$  é a SNR *a priori* local (no canal índice *k*); e  $\sigma_{xk}^2 = E\{x_k^2(n)\}$  e  $\sigma_{vk}^2 = E\{v_k^2(n)\}$  são, respectivamente, as variâncias dos sinais de fala e de ruído em cada um dos *k* canais.

O C2F, por sua vez, minimiza o erro quadrático médio entre os quadrados das envoltórias temporais do sinal estimado e da fala de interesse. Seus coeficientes são calculados por

$$c_k(n) = g_k(n) = \sqrt{\frac{2\xi_k^2(n) + \xi_k(n)}{2\xi_k^2(n) + 4\xi_k(n) + 2}} \quad . \tag{3}$$

A equação (3) pode ser reescrita da seguinte forma:

$$g_{k}(n) = \sqrt{\frac{1}{2} \left[ w_{k}^{2}(n) + w_{k}(n) \right]} .$$
 (4)

As equações (2) e (3) são referidas como funções de supressão. É comum representar tanto a SNR *a priori* local  $\zeta_k(n)$  quanto as funções de supressão em escala de decibel, definidas como (o índice (*n*) foi omitido por clareza):

$$\xi_k^{(\rm dB)} = 10 \log_{10} \xi_k \ , \tag{5}$$

$$w_k^{(dB)} = \xi_k^{(dB)} - 10\log_{10}\left(\xi_k + 1\right) , \qquad (6)$$

$$g_{k}^{(\text{dB})} = \frac{1}{2} \left[ \xi_{k}^{(\text{dB})} + 10 \log_{10} \frac{2\xi_{k} + 1}{\left(\xi_{k} + 1\right)^{2}} - 3,01 \right] .$$
(7)

A Figura 1 mostra as funções de supressão do C2F ( $g_k^{(dB)}$ ) e do WF ( $w_k$ ), em função da SNR *a priori* local ( $\xi_k^{(dB)}$ ).



Fig. 1 – Funções de supressão dos métodos (a) C2F  $(g_k^{(dB)})$  e (b) WF  $(w_k^{(dB)})$ , em função da SNR local  $(\xi_k^{(dB)})$ . Para fins de análise, as curvas podem ser divididas em 3 regiões (I, II e III), separadas arbitrariamente pelas linhas pontilhadas.



Fig. 2 – Gráficos das derivadas das funções de supressão do (a) C2F  $(g_k^{(dB)})$  e do (b) WF  $(w_k^{(dB)})$  em relação à  $\xi_k^{(dB)}$ . As regiões I, II e III estão separadas por linhas verticais pontilhadas.

TABELA I. APROXIMAÇÕES DAS FUNÇÕES DE SUPRESSÃO EM CADA REGIÃO.

		Região	
	(I)	(II)	(III)
	$\zeta_k >> 0 \text{ dB}$	$\zeta_k \ll 0 \; \mathrm{dB}$	$\xi_k \cong 0 \text{ dB}$
Wk <sup>(dB)</sup>	0 dB	$\xi k^{(\mathrm{dB})}$	Eq. (6)
$g_{k^{(\mathrm{dB})}}$	0 dB	$(\xi_k^{(dB)}-3,01)/2$	Eq. (7)

Na Figura 1, pode-se ver que as funções de supressão são contínuas e podem ser modeladas por partes, permitindo a identificação de três regiões: Para  $\zeta_k >> 0$  dB, na região I, o ganho é de aproximadamente 0 dB para ambos os métodos. Para  $\zeta_k << 0$  dB, na região II, as funções de ganho podem ser modeladas por retas com inclinações diferentes, sendo que o C2F apresenta ganhos mais elevados que o WF. A região III, centrada em  $\zeta_k = 0$  dB, é caracterizada como uma região de transição. A Tabela I apresenta aproximações das funções de supressão do WF e do C2F para cada uma das três regiões identificadas.

A Figura 2 apresenta as derivadas das funções de supressão dos métodos WF e C2F, equações (6) e (7), em relação à  $\xi_k^{(dB)}$ . As regiões I, II e III estão indicadas da mesma forma que na Fig. 1.

Através da Tabela I, e mais explicitamente da Figura 2, verifica-se que que em todo o domínio de  $\zeta_k^{(dB)}$ , a derivada de  $g_k^{(dB)}$  é menor que a derivada de  $w_k^{(dB)}$ . Visto que a sensibilidade de uma função à sua variável independente pode ser mensurada através de sua derivada, verifica-se que o C2F é menos sensível, e consequentemente mais robusto, a variações de  $\zeta_k^{(dB)}$  que o WF.

Na região I, as funções de supressão de ambos os métodos são aproximadamente constantes e iguais a 1 (0 dB) e variações de  $\xi_k$  não resultam em mudanças significativas no ganho. Por outro lado, na região II, as derivadas das funções de supressão se distinguem entre si de maneira aproximadamente constante, resultando em uma diferença máxima de 0,5 (3 dB) em favor (maior robustez) do C2F.

Nesta seção verificou-se que o C2F é menos suscetível a erros na estimativa de  $\zeta_k^{(dB)}$  em relação ao WF. A seguir, avalia-se o impacto dessa diferença na inteligibilidade de usuários de IC através de simulações numéricas.

#### III. SIMULAÇÕES NUMÉRICAS

Para realizar uma comparação entre as sensibilidades dos métodos C2F e WF foram realizadas simulações numéricas utilizando todas as 720 frases das listas foneticamente balanceadas do banco de dados IEEE corpus [16]. Como ruído aditivo foi utilizado um ruído de balbuciação, formado artificialmente pela concatenação de quatro sequências do sinal *"International Speech Test Signal"* (ISTS) [17], aleatoriamente deslocadas no tempo.

O banco de filtros utilizado é do tipo *gammatone* [18] com 22 canais. O sinal em cada saída do banco de filtros foi tratado de forma separada e independente. A cada janela de 1 ms, são atualizadas as  $\zeta_k(n)$  e os respectivos ganhos da máscara tempofrequência,  $g_k(n)$  ou  $w_k(n)$ . Nesse experimento, em função da contaminação artificial da fala pelo ruído, tem-se acesso individual aos sinais  $x_k(n) e v_k(n)$ , o que permite o cálculo exato de  $\zeta_k(n)$ . Erros na estimação de  $\zeta_k(n)$  foram introduzidos artificialmente de acordo com a seguinte equação:

$$\widetilde{\xi}_{k}^{(\mathrm{dB})}(n) = \xi_{k}^{(\mathrm{dB})}(n) + z(n) , \qquad (8)$$

em que z(n) é um processo gaussiano de média zero e variância  $\sigma_z^2 = \{5, 10, 50, 100\}.$ 

Os coeficientes dos filtros WF e C2F foram calculados de acordo com as equações (2) e (3), substituindo  $\xi_k(n)$  por sua versão corrompida descrita em (8). Finalmente, a estimativa da fala em cada banda k é calculada por (1).

#### A. Influência dos erros de SNR nos coeficientes dos filtros

Em um primeiro momento avaliou-se o impacto do erro de estimação de  $\zeta_k(n)$  nos valores dos coeficientes dos filtros. Para tal, foi utilizado o erro percentual do valor do coeficiente calculado com  $\tilde{\zeta}_k(n)$  em relação ao coeficiente ideal, dado por:

$$\varepsilon_{c}(k) = \frac{1}{N} \sum_{n=1}^{N} \left| \frac{\tilde{c}_{k}(n) - c_{k}(n)}{c_{k}(n)} \right| \times 100\% , \qquad (9)$$

em que o índice *k* refere-se ao canal em questão;  $c_k = \{w_k, g_k\}$  é o valor ideal do coeficiente;  $\tilde{c}_k = \{\tilde{w}_k, \tilde{g}_k\}$  é o valor do coeficiente calculado a partir de  $\xi_k(n)$ ; e *N* é o número total de janelas para as quais  $c_k(n)$  é não-nulo, de forma a evitar a divisão por zero. Valores de  $\varepsilon_c(k)$  foram calculados para todos os sinais de fala contaminados.

A Figura 3 apresenta um gráfico de barras com a média e desvio padrão de  $\varepsilon_c(k)$ , considerando todos os arquivos de áudio e canais, para diferentes níveis de interferência,  $\sigma_z^2$ , com SNR = 0 dB. Em azul estão os resultados para os sinais processados com o WF e em vermelho para o C2F.

Na Figura 3 fica claro que o erro percentual do valor do coeficiente do filtro para o C2F é consideravelmente menor que para o WF, sobretudo quando  $\sigma_z^2 > 10$ . A diferença é notória para  $\sigma_z^2 = 100$ , em que o erro absoluto médio no cálculo dos coeficientes de WF é superior a 400%, enquanto que para o C2F, é inferior a 90%. Esses resultados corroboram a análise apresentada na seção II, comprovando que o WF é mais sensível que o C2F a erros na estimação de  $\zeta_k$ .



Fig. 3 – Média e desvio padrão de  $\varepsilon_c(k)$ , considerando todos os arquivos de áudio e canais, para SNR = 0 dB e diferentes níveis de interferência  $\sigma_z^2$ . Em azul estão os resultados para os sinais processados com o WF e em vermelho para o C2F.

#### B. Inteligibilidade de usuários de IC

Nesta seção, é avaliado o impacto de erros de estimação de  $\xi_k$  na inteligibilidade de usuários de IC. As misturas de fala e ruído foram construídas para SNR = { -20, -15, -10, -5, 0, 10, 20 } dB.

O conjunto de sinais  $\hat{x}_k(n)$  é usado para o cálculo da métrica *Speech-to-reverberation modutation energy ratio for cochlear implants* (SRMR-CI) [19]. Essa métrica foi especificamente desenvolvida para apresentar alta correlação com a inteligibilidade de portadores de implantes cocleares em experimentos psicoacústicos e pode ser mapeada em valores de inteligibilidade da seguinte maneira [19]:

$$I_{\%} = \frac{1}{1 + e^{a_1 \text{SRMR-CI}_{norm} + a_2}} \times 100\% \quad , \tag{10}$$

em que I<sub>%</sub> é o percentual de inteligibilidade;  $a_1 = 12,1742$  e  $a_2 = -7,4535$  são coeficientes de ajuste [19]; e SRMR-CI<sub>norm</sub> é calculada por:



Fig. 4 – Diagramas de caixa representando a inteligibilidade de usuários de IC, de acordo com a métrica SRMR-CI, para cada nível de SNR. Em azul estão os resultados para os sinais processados com o WF; em vermelho para os processados pelo C2F; e em preto o sinal contaminado não processado. O símbolo + indica que as distribuições para WF e C2F não são estatisticamente distintas (teste de Wilcoxon, p < 0.05). Os gráficos A, B, C e D se referem, respectivamente a  $\sigma_z^2 = \{ 5, 10, 50, 100 \}$ .

$$SRMR_{norm} = \frac{SRMR_{t}}{SRMR_{y}}, \qquad (11)$$

em que SRMR<sub>t</sub> é a métrica do sinal em teste e SRMR<sub>x</sub> é o valor da métrica para o sinal desejado (fala).

Os resultados obtidos neste experimento estão dispostos na Figura 4, no formato de diagramas de caixa. Os gráficos A, B, C e D se referem, respectivamente, a  $\sigma_z^2 = \{5, 10, 50, 100\}$ .

Considerando as Figuras 4A e 4B, assim como os resultados obtidos em [10], verifica-se que o desempenho de usuários de IC em inteligibilidade da fala não deve sofrer impactos significativos quando a imprecisão na estimação de  $\zeta_k$  é pequena ( $\sigma_z^2 \le 10$ ). Por outro lado, quando  $\sigma_z^2 = 50$  (Figura 4C), a redução nos níveis de inteligibilidade começa a ser notada. Por exemplo, para SNR < -5dB  $e \sigma_z^2 = 50$  (Figura 4C), as medianas de inteligibilidade ficam abaixo de 60%, enquanto que quando  $\sigma_z^2 = 10$  (Figura 4B) encontram-se entre 70% e 80%. Para  $\sigma_z^2 = 100$  (Figura 4D), os erros de estimação de  $\zeta_k$  resultam em uma degradação significativa, resultando em níveis de inteligibilidade abaixo de 40% para SNR < -5 dB.

Em todos os casos (A, B, C e D) mostrados na Figura 4, a inteligibilidade se mantém acima de 80% para os casos de SNR = { 10, 20 } dB. Nessas situações, além do sinal não processado já apresentar níveis de inteligibilidade elevados (> 80%), a maior parte das  $\zeta_k$ , se encontra na região I (Figura 1), na qual a função de supressão é aproximadamente unitária e a derivada (Figura 2) é praticamente nula. Com a redução da SNR global,  $\zeta_k$  também diminui, consequentemente, erros na estimação implicam em quedas significativas nos níveis de inteligibilidade. De qualquer forma, em todas as situações estudadas, a inteligibilidade resultante do processamento com o C2F é sempre maior que a obtida com o WF.

### IV. CONCLUSÃO

Neste trabalho foi realizada uma primeira abordagem no estudo da robustez do método de redução de ruído C2F em comparação com o filtro de Wiener. Os resultados das simulações numéricas corroboram a teoria apresentada, indicando que o C2F possui menor sensibilidade a erros de estimação da razão sinal-ruído a priori. Essa menor sensibilidade se reflete em melhores índices de inteligibilidade (considerando a métrica SRMR-CI) para o C2F em comparação com o filtro de Wiener, em todas as situações de erro de estimação da SNR analisadas.

#### AGRADECIMENTOS

Os autores gostariam de agradecer à CAPES, ao CNPq e à Wavetech-ST.

#### Referências

- J. Wouters, H. J. McDermott, T. Francart, "Sound coding in cochlear implants", *IEEE Signal Processing Magazine*, v. 32, n. February, pp. 67-80, 2015.
- [2] J. Wouters et al., "Sound processing for better coding of monaural and binaural cues in auditory prostheses", *Proceeding of the IEEE*, v. 101, n. 9, pp. 1986-1997, 2013.
- [3] P. C. Loizou, "Speech processing in vocoder-centric cochlear implants", *Cochlear and Brainstem Implants*, v. 64, pp. 109-143, 2006.
- [4] B. S. Wilson, M. F. Dorman, "The surprising performance of presentday cochlear implants", *IEEE Transactions on Biomedical Engineering*, v. 54, n. 6, pp. 969-972, 2007.
- [5] A. Hast et al., "Speech perception of elderly cochlear implant users

under different noise conditions", Otology and Neurology, v. 36, n. 10, pp. 1638-1643, 2015.

- [6] F. Bergeron, M. Hotton, "Perception in noise with the Digisonic SP cochlear implant: Clinical trial of Saphyr processor's upgraded signal processing", *European Annals of Otorhinolaryngology, Head and Neck Diseases*, pp. 3-5, 2016.
- [7] D. Wang, G. J. Brown, Computational Auditory Scene Analysis: Principles, Algorithms, and Applications. Wiley/IEEE Press, 2006.
- [8] 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.
- [9] R. Koning, N. Madhu, J. Wouters, "Ideal time-frequency masking algorithms lead to different speech intelligibility and quality in normalhearing and cochlear implant listeners", *IEEE Transactions on Biomedical Engineering*, v. 62, n. 1, pp. 331-341, 2015.
- [10] R. A. Chiea et al., "Um estimador de envoltória temporal para redução de ruído em implantes cocleares", XXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, Petrópolis/RJ, 2019.
- [11] P. C. Loizou, *Speech enhancement Theory and practice* 2<sup>nd</sup>. CRC Press, 2013.
- [12] Y. Wang, A. Narayanan, and D. L. Wang, "On training targets for supervised speech separation", *IEEE/ACM Transactions on Audio*, *Speech, and Language Processing*, v. 22, n. 12, pp. 1849-1858, 2014.

- [13] M. J. Alam, D. O'Shaughnessy, S. A. Selouani, "Speech enhancement employing a sigmoid-type gain function with a modified a priori signalto-noise ratio (SNR) estimator", *Canadian Conference on Electrical and Computer Engineering*, v. 2, n. 6, pp. 631-635, 2008.
- [14] Y. Ephraim, D. Malah, "Speech enhancement using a minimum meansquare error short-time spectral amplitude estimator", *IEEE Transactions on Acoustics, Speech and Signal Processing*, n. 6, pp. 1109-1121, 1984.
- [15] N. Yousefian, P. C Loizou, "A Dual-Microphone Speech Enhancement Algorithm Based on the Coherence Function", *IEEE Transactions on Audio, Speech and Language Processing*, v. 20, n. 2, pp. 599-609, 2012.
- [16] IEEE recommended practice for speech quality measurements, *IEEE Transactions on Audio and Electroacoustics*, v.17 n. 3, pp. 225-246, 1969.
- [17] I. Holube, S. Fredelake, M. Vlaming, B. Kollmeier, "Development and analysis of an international speech test signal (ISTS)", *International Journal of Audiology*, v. 49, n. 12, pp. 891-903, 2010.
- [18] P. I. M. Johannesma, "The pre-response stimulus ensemble of neurons in the cochlear nucleus", *Symposium on Hearing Theory*, pp. 58-69, 1972.
- [19] J. F. Santos, 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.