DOI: 10.61782/fa.2025.0817

FORUM ACUSTICUM EURONOISE 2025

SPATIALLY SELECTIVE ACTIVE NOISE CONTROL FOR OPEN-FITTING HEARABLES WITH ACAUSAL OPTIMIZATION

Tong Xiao* **Simon Doclo**
Department of Medical Physics and Acoustics and Cluster of Excellence Hearing4all,
Carl von Ossietzky Universität Oldenburg, Germany

ABSTRACT

Recent advances in active noise control have enabled the development of hearables with spatial selectivity, which actively suppress undesired noise while preserving desired sound from specific directions. In this work, we propose an improved approach to spatially selective active noise control that incorporates acausal relative impulse responses into the optimization process, resulting in significantly improved performance over the causal design. We evaluate the system through simulations using a pair of open-fitting hearables with spatially localized speech and noise sources in an anechoic environment. Performance is evaluated in terms of speech distortion, noise reduction, and signal-to-noise ratio improvement across different delays and degrees of acausality. Results show that the proposed acausal optimization consistently outperforms the causal approach across all metrics and scenarios, as acausal filters more effectively characterize the response of the desired source.

Keywords: active noise control, spatial selectivity, causality, beamforming

1. INTRODUCTION

Active noise control (ANC) hearables create a quiet environment by using secondary sources to generate anti-noise, which minimizes sound at specific positions when superimposed on the primary noise (leakage) [1, 2]. Based on their fit, hearables can be categorized as closed-fitting (completely occluding the ear), open-fitting (partially occluding the ear), and open-ear (no occlusion, e.g., smart glasses). Open-fitting and open-ear designs, in particular, reduce the occlusion effect and improve the physical comfort. Recent research focuses on designing intelligent ANC hearables with spatial selectivity, particularly for complex acoustic environments like cocktail-party scenarios involving multiple sound sources from different directions [3–5]. In these scenarios, listeners may wish to focus on desired sounds from a specific direction (e.g., the front) while blocking out undesired sounds from other directions.

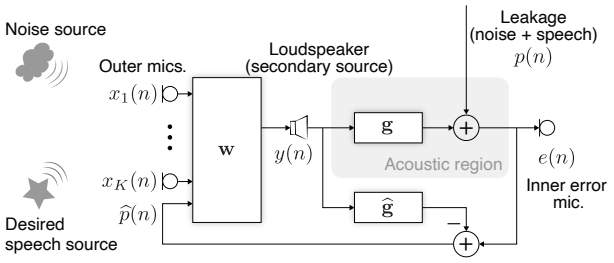


Figure 1: Block diagram of an ANC system with K outer microphones, one inner error microphone and one loudspeaker (i.e., secondary source). The control filter is denoted by $\mathbf{w}$, the secondary path is denoted by $\mathbf{g}$, and its estimate by $\widehat{\mathbf{g}}$.

Modern ANC hearables are commonly equipped with multiple microphones, i.e., microphones on the exterior of the hearable and error microphones in the interior close to the eardrum. On the one hand, conventional ANC algorithms suppress all leakage measured by the inner error microphones, including desired speech. On the other hand, multi-microphone noise reduction algorithms perform spatial filtering by reducing all noise while preserving speech from the desired direction [6–8]. However, these algorithms typically ignore the leakage and do not

*Corresponding author: Tong.Xiao@uni-oldenburg.de.

exploit the inner error microphones. Recent advancements have proposed integrating spatial filtering into an ANC system so that sounds from a desired direction are preserved and sounds from undesired directions are actively suppressed [9–17]. A recent study demonstrated that, when the speech component from an outer reference microphone signal is desired, a small delay is required to satisfy causality. The optimal delay is the acoustic delay between the reference microphone and the inner error microphone for the desired source [17]. These findings were derived using a system with *causal* relative impulse responses (ReIRs). However, under more flexible design constraints, *acausal* ReIRs may give a better performance.

In this paper, we present an improved approach to spatially selective ANC (SSANC) that leverages *acausal* ReIRs in the optimization process, thereby substantially improving performance over the causal optimization approach. We analyze the signal distortion, noise reduction, and signal-to-noise ratio (SNR) improvement of the inner error microphone signal when the speech component of an outer microphone signal is desired. We identify both the optimal range of delays and the optimal degree of acausality based on these evaluation metrics. These insights guide the design of a more effective SSANC for open-fitting hearables.

2. SIGNAL MODEL

As shown in Fig. 1, we consider a hearable with K outer microphones. Without loss of generality, we consider one loudspeaker as the secondary source and one inner error microphone, resulting in a total of $K+1$ microphones. We assume that the acoustic feedback paths between the loudspeaker and the outer microphones are known, such that acoustic feedback can be canceled. In addition, we assume that an estimate of the secondary path between the loudspeaker and the inner error microphone is available. Subscripts $(\cdot)_s$ and $(\cdot)_v$ denote the speech and noise components in signals, respectively.

The inner error microphone signal $e(n)$, with n the discrete time index, is given by

$$e(n) = p(n) + (\mathbf{G}\mathbf{w})^{\mathcal{T}}\mathbf{x}(n), \tag{1}$$

where $(\cdot)^{\mathcal{T}}$ denotes the transpose operator. The leakage (including noise and desired speech) at the inner error microphone is denoted by $p(n)$, and the anti-noise component at the inner error microphone is given by $(\mathbf{G}\mathbf{w})^{\mathcal{T}}\mathbf{x}(n)$, where $\mathbf{w}$ is the stacked control filter, $\mathbf{x}(n)$ is the stacked input vector, and $\mathbf{G}$ represents the secondary path convolution matrix. The stacked control filter $\mathbf{w}$ is defined as

$$\mathbf{w} = [\mathbf{w}_1^{\mathcal{T}} \quad \mathbf{w}_2^{\mathcal{T}} \quad \dots \quad \mathbf{w}_{K+1}^{\mathcal{T}}]^{\mathcal{T}} \in \mathbb{R}^{(K+1)L_w}, \tag{2a}$$

$$\mathbf{w}_k = [w_{k,0}\ w_{k,1}\ \dots\ w_{k,L_w-1}]^{\mathcal{T}} \in \mathbb{R}^{L_w}, \tag{2b}$$

where L_w denotes the control filter length for each channel. The convolution matrix of the secondary path is defined as

$$\mathbf{G} = \text{blkdiag}\,(\mathbb{G}\dots\mathbb{G}) \in \mathbb{R}^{(K+1)L\times(K+1)L_w}, \tag{3a}$$

$$\mathbb{G} = \begin{bmatrix} g_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ g_{L_g-1} & \ddots & g_0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g_{L_g-1} \end{bmatrix} \in \mathbb{R}^{L\times L_w}, \tag{3b}$$

where $L = L_g + L_w - 1$, and L_g denotes the secondary path filter length. As input signals to the control filter we consider the K outer microphone signals $\mathbf{x}_k(n)$, $k = 1,\dots,K$, and an estimate of the leakage $\widehat{\mathbf{p}}(n)$, i.e., the stacked input vector $\mathbf{x}(n)$ is defined as

$$\mathbf{x}(n) = [\mathbf{x}_1^{\mathcal{T}}(n)\ \dots\ \mathbf{x}_K^{\mathcal{T}}(n)\ \widehat{\mathbf{p}}^{\mathcal{T}}(n)]^{\mathcal{T}} \in \mathbb{R}^{(K+1)L}, \tag{4}$$

with

$$\mathbf{x}_k(n) = [x_k(n)\ \dots\ x_k(n-L+1)]^{\mathcal{T}}, \tag{5a}$$

$$\widehat{\mathbf{p}}(n) = [\,\widehat{p}(n)\ \dots\ \widehat{p}(n-L+1)]^{\mathcal{T}}. \tag{5b}$$

The estimated leakage $\widehat{p}(n)$ can be computed from the inner error microphone signal $e(n)$, the loudspeaker signal vector $\mathbf{y}(n) = [y(n)\dots y(n-L_g+1)]^{\mathcal{T}}$, and an estimate of the secondary path $\widehat{\mathbf{g}}$ as

$$\widehat{p}(n) = e(n) - \widehat{\mathbf{g}}^{\mathcal{T}}\mathbf{y}(n). \tag{6}$$

Assuming a perfect estimate of the secondary path to be available, i.e., $\widehat{\mathbf{g}} = \mathbf{g} = [g_0\ g_1\ \dots\ g_{L_g-1}]^{\mathcal{T}}$, the leakage can be written as $p(n) = \widehat{p}(n) = \mathbf{q}^{\mathcal{T}}\mathbf{x}(n)$, with

$$\mathbf{q} = [\mathbf{0}^{\mathcal{T}}\ \dots\ \mathbf{0}^{\mathcal{T}}\ \boldsymbol{\delta}^{\mathcal{T}}]^{\mathcal{T}} \in \mathbb{R}^{(K+1)L}, \tag{7a}$$

$$\boldsymbol{\delta} = [1\quad 0\quad \dots\quad 0\,]^{\mathcal{T}} \in \mathbb{R}^{L}. \tag{7b}$$

Hence, the inner error microphone signal in (1) can be rewritten as

$$e(n) = (\mathbf{q} + \mathbf{G}\mathbf{w})^{\mathcal{T}}\mathbf{x}(n). \tag{8}$$

Assuming that the speech and the noise components are uncorrelated, then we have

$$e_s(n) = (\mathbf{q} + \mathbf{G}\mathbf{w})^{\mathcal{T}}\mathbf{x}_s(n), \tag{9a}$$

$$e_v(n) = (\mathbf{q} + \mathbf{G}\mathbf{w})^{\mathcal{T}}\mathbf{x}_v(n). \tag{9b}$$

3. ACAUSAL OPTIMIZATION

In the SSANC system, the objective is to minimize the power of the inner error microphone signal while preserving the delayed desired speech component of an outer reference microphone signal. This can be achieved by imposing the constraint

$$e_s(n) = x_{\mathrm{ref},s}(n-\Delta). \tag{10}$$

For each channel, the speech component can be represented as the convolution of the speech component of the reference microphone signal and an acausal ReIR filter, that is,

$$x_{k,s}(n) = \sum_{l=-L_a}^{L_h-1} h_k(l)\, x_{\mathrm{ref},s}(n-l), \tag{11}$$

where the ReIR has L_a-tap anti-causal part and L_h-tap causal part. The stacked input vector for all the channels can then be written as

$$\mathbf{x}_s(n) = \mathbf{H}^{\mathcal{T}}\, \mathbf{x}_{\mathrm{ref},s}(n), \tag{12}$$

where

$$\mathbf{H} = [\mathbf{H}_1 \dots \mathbf{H}_{K+1}] \in \mathbb{R}^{(L_a+L_h+L-1)\times(K+1)L}, \tag{13a}$$

$$\mathbf{H}_k = \begin{bmatrix} h_{k,-L_a} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ h_{k,L_h-1} & \ddots & 0 \\ 0 & \ddots & h_{k,-L_a} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{k,L_h-1} \end{bmatrix} \in \mathbb{R}^{(L_a+L_h+L-1)\times L}, \tag{13b}$$

$$\mathbf{x}_{\mathrm{ref},s}(n) = [x_{\mathrm{ref},s}(n+L_a) \dots x_{\mathrm{ref},s}(n-L_h-L+2)]^{\mathcal{T}} \in \mathbb{R}^{L_a+L_h+L-1}, \tag{13c}$$

and $\mathbf{x}_s(n) \in \mathbb{R}^{(K+1)L}$ is similarly defined as (4).

We rewrite (10) as

$$e_s(n) = \boldsymbol{\delta}_\Delta^{\mathcal{T}} \mathbf{x}_{\mathrm{ref},s}(n), \tag{14}$$

where

$$\boldsymbol{\delta}_\Delta = [\underbrace{0 \dots 0}_{L_a}\ \underbrace{0 \dots 0}_{\Delta}\ \underbrace{1\, 0 \dots 0}_{L_h+L-1-\Delta}]^{\mathcal{T}} \in \mathbb{R}^{L_a+L_h+L-1}. \tag{15}$$

Thus, using (9a), (12) and (14), we reformulate (10) as

$$\mathbf{H}(\mathbf{q}+\mathbf{G}\mathbf{w}) = \boldsymbol{\delta}_\Delta. \tag{16}$$

It is important to note that although the ReIRs, e.g., $\mathbf{H}$ and $\boldsymbol{\delta}_\Delta$, are modeled to be *acausal*, the SSANC control filter $\mathbf{w}$ for generating the anti-noise is still *causal*.

Therefore, the optimization problem for an SSANC system can be defined as

$$\underset{\mathbf{w}}{\text{minimize}} \quad \mathcal{E}\left\{e^2(n)\right\} + \beta \mathbf{w}^{\mathcal{T}}\mathbf{w} \tag{17a}$$

$$\text{subject to} \quad \mathbf{H}(\mathbf{q}+\mathbf{G}\mathbf{w}) = \boldsymbol{\delta}_\Delta, \tag{17b}$$

where $\mathcal{E}\{\cdot\}$ denotes the mathematical expectation operator. β denotes a regularization factor. The solution is found to be

$$\begin{aligned}\mathbf{w} = &-\left[\mathbf{I}-\boldsymbol{\Phi}_{\mathbf{rr}}^{-1}\mathbf{G}^{\mathcal{T}}\mathbf{H}^{\mathcal{T}}\left(\mathbf{H}\mathbf{G}\boldsymbol{\Phi}_{\mathbf{rr}}^{-1}\mathbf{G}^{\mathcal{T}}\mathbf{H}^{\mathcal{T}}+\rho\mathbf{I}\right)^{-1}\mathbf{H}\mathbf{G}\right]\boldsymbol{\Phi}_{\mathbf{rr}}^{-1}\boldsymbol{\phi} \\ &+\boldsymbol{\Phi}_{\mathbf{rr}}^{-1}\mathbf{G}^{\mathcal{T}}\mathbf{H}^{\mathcal{T}}\left(\mathbf{H}\mathbf{G}\boldsymbol{\Phi}_{\mathbf{rr}}^{-1}\mathbf{G}^{\mathcal{T}}\mathbf{H}^{\mathcal{T}}+\rho\mathbf{I}\right)^{-1}(\boldsymbol{\delta}_\Delta-\mathbf{H}\mathbf{q}),\end{aligned} \tag{18}$$

where

$$\boldsymbol{\Phi}_{\mathbf{xx}} = \mathcal{E}\left\{\mathbf{x}(n)\mathbf{x}^{\mathcal{T}}(n)\right\}, \tag{19a}$$

$$\boldsymbol{\Phi}_{\mathbf{rr}} = \mathbf{G}^{\mathcal{T}}\boldsymbol{\Phi}_{\mathbf{xx}}\mathbf{G} + \beta\mathbf{I}, \tag{19b}$$

$$\boldsymbol{\phi} = \mathbf{G}^{\mathcal{T}}\boldsymbol{\Phi}_{\mathbf{xx}}\mathbf{q}, \tag{19c}$$

and ρ denotes a small regularization factor, $\mathbf{I}$ denotes the identity matrix.

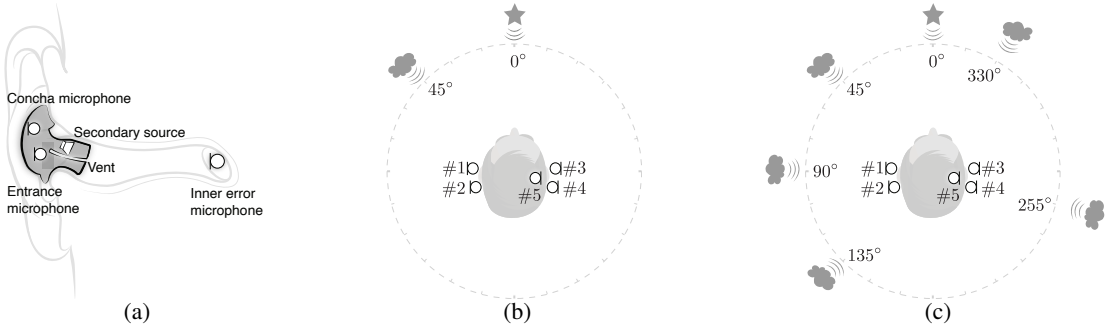
It should be noted that the solution is similar to the previous studies [15, 17]. The major difference is that the ReIR matrix $\mathbf{H}$ is acausal now containing L_a-tap anti-causal part of the ReIR for all channels. When $L_a = 0$, the solution becomes causal, as in the previous approach.

4. SIMULATIONS

4.1 Setup

For the simulation, we considered a pair of open-fitting hearables [18, 19] inserted into both ears of a GRAS 45BB-12 KEMAR Head & Torso simulator, as shown in Fig. 2(a). We used four outer microphones (entrance microphones and concha microphones at the left and right ears, labeled as #1– #4), one inner error microphone (located at the right ear, labeled as #5), and the outer receiver at the right ear as the secondary source. The inner error microphone was assumed to be at the eardrum.

Two acoustic scenarios were considered. In the first scenario, as shown in Fig. 2(b), we considered a desired speech source at 0° and an undesired speech source at 45°



(a) (b) (c)

Figure 2: (a) Illustration of the open-fitting hearable. (b) First acoustic scenario. A desired speech source is at $0°$, and one undesired speech source is at $45°$. (c) Second acoustic scenario. The desired speech source is at $0°$, and five babble noise sources are at $45°$, $90°$, $135°$, $255°$ and $330°$.

("p361_005" and "p243_018" from the VCTK dataset [20], respectively). In the second scenario, as shown in Fig. 2(c), the same desired speech source remained at $0°$, and five noise sources were placed at $45°$, $90°$, $135°$, $255°$ and $330°$ (babble noise from the NOISEX-92 database [21]). The desired speech and noise components in all microphone signals were generated by convolving source signals with measured anechoic impulse responses from the database [19]. All signals were 5 seconds in duration and sampled at 16 kHz. The desired speech and noise components were mixed such that the signal-to-noise ratio (SNR) of the leakage at the inner error microphone was set to -5 dB. In the second scenario, all noise sources contributed the same energy. As the secondary path estimate we used the measured impulse response between the outer receiver and the inner error microphone at the right ear (from [19]).

We set the filter lengths to $L_w = 600$, $L_g = 280$, and $L_h = 262$. The acausal ReIRs were computed using the least-mean-squares adaptive filter (after convergence) from microphone signals simulated with white noise at $0°$, using the entrance microphone #3 as the reference microphone. Similar to the previous studies [15, 17], we were only interested in the frequency range above 100 Hz. Therefore, a minimum-phase high-pass filter with a cut-off frequency at 100 Hz was applied to the desired signal $x_{\mathrm{ref},s}(n-\Delta)$. The effects of delay ($\Delta$) and acausality ($L_a$) on system performance were investigated.

4.2 Evaluation metrics

The following three metrics are used for evaluation. The intelligibility-weighted spectral distortion is used to assess the amount of speech distortion [9, 22, 23]. It is defined as

$$\mathrm{SD}_{\mathrm{intellig}}\,(\mathrm{dB}) = \sum_{b=1}^{B} I(\omega_b)\, 10\log_{10}\frac{\mathcal{P}_{\epsilon}(\omega_b)}{\mathcal{P}_{\mathrm{ref},s}(\omega_b)}, \quad (20)$$

where the band importance function $I(\omega_b)$ expresses the importance of the b-th one-third octave band for intelligibility [24], and B denotes the total number of bands. $\mathcal{P}_{\epsilon}(\omega_b)$ is the power spectral density of $\epsilon(n)$ in the b-th band, where $\epsilon(n) = e_s(n) - x_{\mathrm{ref},s}(n-\Delta)$. $\mathcal{P}_{\mathrm{ref},s}(\omega_b)$ is the power spectral density of $x_{\mathrm{ref},s}(n-\Delta)$ in the b-th band.

The noise reduction is defined as the difference between the power of the noise component of the leakage $p_v(n)$ (without control) and the noise component of the inner error microphone signal $e_v(n)$ (with control), i.e.,

$$\mathrm{NR}\,(\mathrm{dB}) = 10\log_{10}\sum_{n=1}^{N} p_v^2(n) - 10\log_{10}\sum_{n=1}^{N} e_v^2(n), \quad (21)$$

where N denotes the total signal length.

The SNR improvement (ΔSNR) is defined as

$$\Delta\mathrm{SNR}\,(\mathrm{dB}) = 10\log_{10}\frac{\sum_{n=1}^{N} e_s^2(n)}{\sum_{n=1}^{N} e_v^2(n)} - 10\log_{10}\frac{\sum_{n=1}^{N} p_s^2(n)}{\sum_{n=1}^{N} p_v^2(n)}. \quad (22)$$

4.3 Performance for different delays

First, we investigate the effect of delay Δ on performance. The speech distortion, the noise reduction and the SNR improvement are shown in Fig. 3 for different delays ranging from zero to 80 samples (5 ms) for the first scenario.

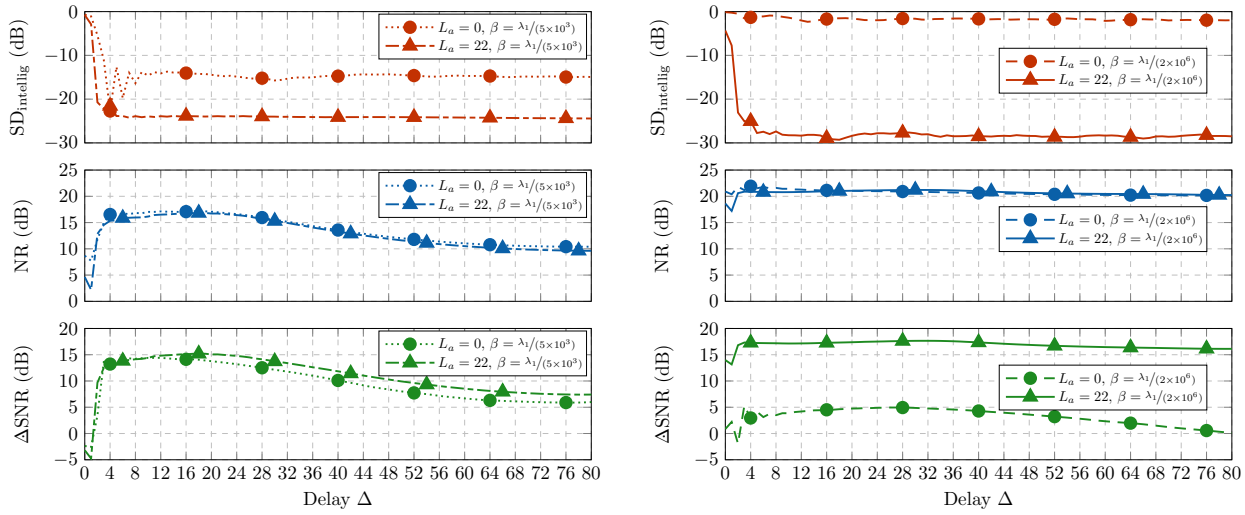


Figure 3: Speech distortion, noise reduction, and SNR improvement for the first simulation scenario for different delay Δ when (a) $L_a = 0$ and $L_a = 22$, with $\beta = \lambda_{\max}(\mathbf{G}^T\mathbf{\Phi_{xx}}\mathbf{G})/(5\times10^3) = \lambda_1/(5\times10^3)$, (b) $L_a = 0$ and $L_a = 22$, with $\beta = \lambda_1/(2\times10^6)$. In all cases, $\rho = \lambda_{\max}(\mathbf{HG}\mathbf{\Phi_{rr}}^{-1}\mathbf{G}^T\mathbf{H}^T)/(1\times10^5)$.

Four design choices are considered. When $L_a = 0$, $\beta = \lambda_{\max}(\mathbf{G}^T\mathbf{\Phi_{xx}}\mathbf{G})/(5\times10^3) = \lambda_1/(5\times10^3)$ and $\rho = \lambda_{\max}(\mathbf{HG}\mathbf{\Phi_{rr}}^{-1}\mathbf{G}^T\mathbf{H}^T)/(1\times10^5)$ as shown in Fig. 3(a), where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue, this design corresponds to the previous causal optimization [15, 17]. It can be observed that there is very high speech distortion, low noise reduction and low SNR improvement for $\Delta < 4$. This is due to the delays between the reference microphone and the inner error microphone from the desired source, which has a 4-sample delay. The causality constraint requires the delay to be at least four samples. For $4 \leq \Delta \leq 30$, the speech distortion is about -15 dB, the noise reduction has about 15–17 dB, and the SNR improvement is about 14 dB. However, further increases in delay result in degraded noise reduction, which in turn leads to reduced SNR improvement. These results agree with the previous findings.

When $L_a = 22$ with the same β and ρ as shown in Fig. 3(a), it can be observed that the speech distortion is reduced to a much lower level for $\Delta \geq 4$, reaching about -24 dB. However, the noise reduction and the SNR improvement do not show significant benefits compared to the previous results.

The main advantage of acausal optimization is its ability to achieve better performance with a smaller β value, i.e., when less constraint is imposed on the secondary source. The results are shown in Fig. 3(b) when $\beta = \lambda_1/(2\times10^6)$. In the causal case ($L_a = 0$), the system

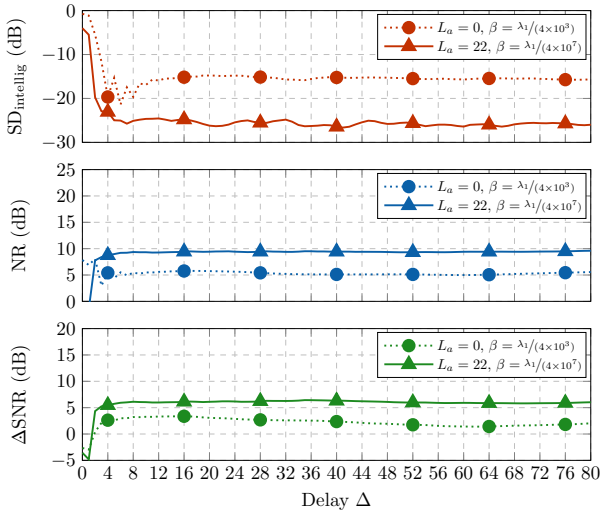


Figure 4: Speech distortion, noise reduction, and SNR improvement for the second simulation scenario for different delay Δ when $L_a = 0$, $\beta = \lambda_1/(4\times10^3)$, and when $L_a = 22$, $\beta = \lambda_1/(4\times10^7)$. In both cases, $\rho = \lambda_{\max}(\mathbf{HG}\mathbf{\Phi_{rr}}^{-1}\mathbf{G}^T\mathbf{H}^T)/(2\times10^5)$.

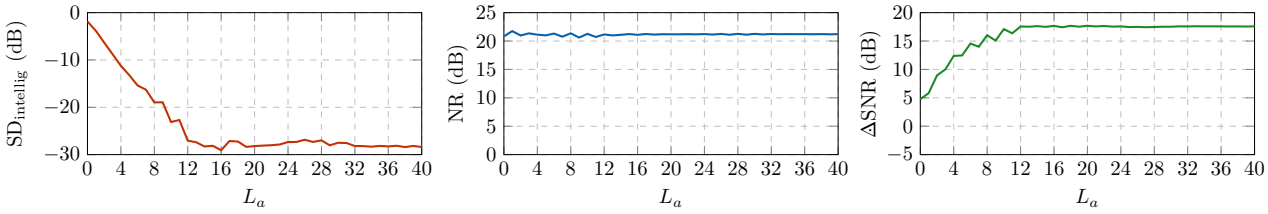


Figure 5: Speech distortion, noise reduction, and SNR improvement for the first simulation scenario for different L_a values when $\Delta = 32$, $\beta = \lambda_1/(2\times10^6)$ and $\rho = \lambda_{\max}(\mathbf{H}\mathbf{G}\mathbf{\Phi}_{\mathbf{rr}}^{-1}\mathbf{G}^{\mathcal{T}}\mathbf{H}^{\mathcal{T}})/(1\times10^5)$.

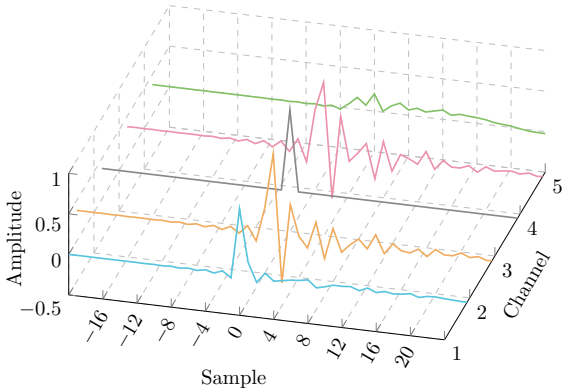


Figure 6: Acausal ReIRs for all five channels.

achieves a high level of noise reduction (21 dB), but at the cost of significant speech distortion (approximately -2 dB), resulting in an SNR improvement of just under 5 dB. In this configuration, the SSANC system fails to operate as intended and instead behaves similarly to conventional ANC, reducing both speech and noise components of the leakage. In contrast, the acausal design ($L_a = 22$) achieves much lower speech distortion (around -28 dB for most delays), maintains a similar level of noise reduction (21 dB), and delivers an SNR improvement exceeding 17 dB for most delays. Furthermore, the acausal approach yields consistent performance across varying delays, reducing the sensitivity to delay selection and thus imposing fewer constraints on the system compared to the causal design, which requires careful tuning of delay parameters.

The results for the second scenario are shown in Fig. 4. The performance reduces overall compared to the first scenario, which is expected since the scenario is much more challenging. However, the general trend remains the same. The speech distortion is reduced from about -15 dB to about -26 dB, the noise reduction increases from approximately 6 dB to around 9.5 dB, and the SNR improvement is increased from about 3 dB to about 6 dB.

4.4 Performance for different degrees of acausality

The results above show the performance for the causal design ($L_a = 0$) and one acausal design ($L_a = 22$). We now investigate the effect of different degrees of acausality, i.e., different lengths of anti-causal part L_a on performance. For this analysis, we use the first scenario and fix the delay to $\Delta = 32$ (2 ms), with $\beta = \lambda_1/(2\times10^6)$ and $\rho = \lambda_{\max}(\mathbf{H}\mathbf{G}\mathbf{\Phi}_{\mathbf{rr}}^{-1}\mathbf{G}^{\mathcal{T}}\mathbf{H}^{\mathcal{T}})/(1\times10^5)$. The results for different L_a values are shown in Fig. 5.

It can be observed that noise reduction remains consistent across all L_a values. However, speech distortion is high when L_a is small. As L_a increases, the speech distortion decreases and stabilizes when L_a reaches approximately 12. The SNR improvement follows a similar trend.

By examining the ReIRs for all five channels, as depicted in Fig. 6, it can be observed that the amplitude of the impulse responses also stabilizes around $L_a = 12$. Thus, using acausal filters to model the response from the desired source enables the SSANC system to achieve better performance than that of a purely causal design.

5. CONCLUSION

This paper has presented an improved approach to spatially selective active noise control that incorporates acausal relative impulse responses into the optimization process, leading to a substantial performance improvement over the causal optimization approach. Using a pair of open-fitting hearables in two acoustic scenarios, we demonstrated that the speech distortion under acausal optimization remains consistently lower than that of the causal optimization. Furthermore, both noise reduction and SNR improvement are significantly higher with acausal optimization. This indicates that acausal relative impulse responses offer a more accurate representation of the desired source, leading to improved noise control performance.

6. ACKNOWLEDGMENTS

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 352015383 – SFB 1330 C1.

7. REFERENCES

[1] S. J. Elliott, *Signal processing for active control*. Academic Press, 2000.

[2] C. Hansen, S. Snyder, X. Qiu, L. Brooks, and D. Moreau, *Active Control of Noise and Vibration*. CRC Press, 2 ed., Nov. 2012.

[3] Y. Kajikawa, W.-S. Gan, and S. M. Kuo, "Recent advances on active noise control: open issues and innovative applications," *APSIPA Transactions on Signal and Information Processing*, vol. 1, e3, pp. 1–21, 2012.

[4] C.-Y. Chang, A. Siswanto, C.-Y. Ho, T.-K. Yeh, Y.-R. Chen, and S. M. Kuo, "Listening in a noisy environment: Integration of active noise control in audio products," *IEEE Consumer Electronics Magazine*, vol. 5, no. 4, pp. 34–43, 2016.

[5] R. Gupta, J. He, R. Ranjan, W.-S. Gan, F. Klein, C. Schneiderwind, A. Neidhardt, K. Brandenburg, and V. Välimäki, "Augmented/mixed reality audio for hearables: sensing, control, and rendering," *IEEE Signal Processing Magazine*, vol. 39, no. 3, pp. 63–89, 2022.

[6] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[7] S. Doclo, W. Kellermann, S. Makino, and S. E. Nordholm, "Multichannel signal enhancement algorithms for assisted listening devices: Exploiting spatial diversity using multiple microphones," *IEEE Signal Processing Magazine*, vol. 32, pp. 18–30, Mar. 2015.

[8] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 4, pp. 692–730, 2017.

[9] R. Serizel, M. Moonen, J. Wouters, and S. H. Jensen, "Integrated active noise control and noise reduction in hearing aids," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1137–1146, 2010.

[10] R. Serizel, M. Moonen, J. Wouters, and S. H. Jensen, "Output SNR analysis of integrated active noise control and noise reduction in hearing aids under a single speech source scenario," *Signal Processing*, vol. 91, no. 8, pp. 1719–1729, 2011.

[11] D. Dalga and S. Doclo, "Combined feedforward-feedback noise reduction schemes for open-fitting hearing aids," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, (New Paltz, NY, USA), pp. 185–188, 2011.

[12] D. Dalga and S. Doclo, "Theoretical performance analysis of ANC-motivated noise reduction algorithms for open-fitting hearing aids," in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, (Aachen, Germany), pp. 1–4, 2012.

[13] C.-Y. Ho, K.-K. Shyu, C.-Y. Chang, and S. M. Kuo, "Integrated active noise control for open-fit hearing aids with customized filter," *Applied Acoustics*, vol. 137, pp. 1–8, 2018.

[14] V. Patel, J. Cheer, and S. Fontana, "Design and implementation of an active noise control headphone with directional hear-through capability," *IEEE Transactions on Consumer Electronics*, vol. 66, pp. 32–40, Feb. 2020.

[15] T. Xiao, B. Xu, and C. Zhao, "Spatially selective active noise control systems," *The Journal of the Acoustical Society of America*, vol. 153, pp. 2733–2744, May 2023.

[16] H. Zhang, J. Zhang, F. Ma, P. N. Samarasinghe, and H. Sun, "A time-domain multi-channel directional active noise control system," in *Proc. European Signal Processing Conference (EUSIPCO)*, (Helsinki, Finland), pp. 376–380, 2023.

[17] T. Xiao and S. Doclo, "Effect of target signals and delays on spatially selective active noise control for open-fitting hearables," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Seoul, Republic of Korea), pp. 1056–1060, 2024.

[18] F. Denk, M. Lettau, H. Schepker, S. Doclo, R. Roden, M. Blau, J.-H. Bach, J. Wellmann, and B. Kollmeier, "A one-size-fits-all earpiece with multiple microphones and drivers for hearing device research," in *Proc. AES International Conference on Headphone Technology*, (San Francisco, USA), pp. 1–9, Aug. 2019.

[19] F. Denk and B. Kollmeier, "The hearpiece database of individual transfer functions of an in-the-ear earpiece for hearing device research," *Acta Acustica*, vol. 5, no. 2, pp. 1–16, 2021.

[20] C. Veaux, J. Yamagishi, and K. MacDonald, "CSTR VCTK corpus: English multi-speaker corpus for CSTR voice cloning toolkit," 2017.

[21] A. Varga and H. J. Steeneken, "Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems," *Speech Communication*, vol. 12, no. 3, pp. 247–251, 1993.

[22] A. Spriet, M. Moonen, and J. Wouters, "Spatially preprocessed speech distortion weighted multi-channel Wiener filtering for noise reduction," *Signal Processing*, vol. 84, no. 12, pp. 2367–2387, 2004.

[23] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for the speech distortion weighted multichannel Wiener filter for robust noise reduction," *Speech Communication*, vol. 49, no. 7, pp. 636–656, 2007.

[24] Acoustical Society of America (ASA), "Methods for Calculation of the Speech Intelligibility Index," ANSI/ASA S3.5-1997 Standard, American National Standards Institute (ANSI), 1997.