An Efficient Crosstalk Cancellation Algorithm for Headrest-Based Spatial Audio Rendering

Nefeli Aikaterini Dourou
Dept. of Information Engineering
Università Politecnica delle Marche
Ancona, Italy
n.a.dourou@pm.univpm.it

Valeria Bruschi
Dept. of Information Engineering
Università Politecnica delle Marche
Ancona, Italy
v.bruschi@staff.univpm.it

Alessandro Terenzi
Dept. of Information Engineering
Università Politecnica delle Marche
Ancona, Italy
a.terenzi@staff.univpm.it

Stefania Cecchi
Dept. of Information Engineering
Università Politecnica delle Marche
Ancona, Italy
s.cecchi@staff.univpm.it

***Abstract*—Immersive audio systems aim to recreate three-dimensional soundscapes. A key component of such systems, when they are based on binaural reproduction, is the crosstalk cancellation (CTC), which minimizes signal leakage between channels to preserve spatial accuracy. This paper presents an enhanced crosstalk cancellation algorithm for headrest-based spatial audio rendering systems. The proposed method is based on a space average CTC algorithm, known for its robustness to small head movements, and introduces a causality constraint during the filter design phase to improve efficiency. Applied to a headrest-mounted loudspeaker configuration, the algorithm is evaluated using both objective acoustic metrics and subjective listening tests. Results demonstrate that the proposed algorithm delivers effective crosstalk suppression.**

***Index Terms*—crosstalk cancellation, audio headrest reproduction, spatial sound, immersive audio systems, 3D audio.**

I. Introduction

During the past years, spatial audio rendering has attracted significant attention from both the scientific and industrial communities due to its potential to improve sound and music perception in various applications, including virtual/augmented reality [1], [2], video conferencing [3], gaming [4], and also in the biomedical field [5]. These systems use binaural audio rendered through headphones, delivering spatially encoded signals to simulate virtual sound sources. However, headphones are often impractical in many scenarios, making loudspeakers the preferred alternative. A key limitation of loudspeakers is crosstalk, where each ear receives signals from both channels due to the sound field they produce. To address this issue, crosstalk cancellation (CTC) processing becomes essential. Crosstalk cancellation filters are fundamentally based on inverting a matrix of head-related transfer functions (HRTFs), which represent how sound propagates from loudspeakers to the listener's ears.

Frequency-domain inversion is generally preferred over time-domain methods due to its lower computational cost. A widely used technique, introduced in [6], employs FFT-based deconvolution to design finite impulse response (FIR) inverse filters via a least-squares solution with zeroth-order regularization. However, these filters often contain anti-causal components that, because of the periodic nature of the discrete Fourier transform, wrap around in time and may cause audible artifacts. A common approach to mitigate this is to introduce additional delay and apply time-windowing to the crosstalk filters, though the introduction of delay can increase overall system latency [7]. To address this issue, [7] introduces a causality constraint into the conventional FFT-based deconvolution method described in [6]. The algorithm employs Wiener-Hopf decomposition to factorize the determinant of the matrix into its minimum stable causal and stable, anti-causal components. More recently, another method incorporating a causality constraint directly into the frequency-domain least-squares optimization framework has been developed for crosstalk cancellation filter design [8].

With modern lifestyles increasingly involving extended time in transportation systems, binaural headrest audio systems are emerging as a promising solution. These systems incorporate loudspeakers integrated into seat headrests, enabling binaural audio reproduction for seated users in multiple contexts, such as in cars [9], [10], helicopters/aircrafts [11], [12] and yachts environments [13]. While existing research has implemented headrest applications for active noise control [12], [13] and personalized audio zones [14]–[16], relatively few studies have specifically investigated binaural audio reproduction in this context.

Specifically, Lundkvist et al. [10] developed and evaluated a crosstalk cancellation algorithm for binaural audio headrest systems. The authors tested three loudspeakers configurations and demonstrated that positioning the loudspeakers close to the listener's head can yield effective CTC performance at certain frequencies. The study further revealed that at certain frequencies, loudspeaker placement significantly influences matrix conditioning [6], while head shadowing naturally enhances

This work was supported by the financial program DM MiSE 3370 27 October 2023, project "Epignosis" F/340008/01/X59.

channel separation.

Stanhope et al. [17] investigated the feasibility of binaural reproduction using headrest-mounted loudspeakers without crosstalk cancellation. Their research focused on configurations in which the loudspeakers were positioned in direct contact with the listener's occipital region, demonstrating that broadband binaural rendering could be achieved in anechoic conditions solely through natural channel separation. However, the study omitted two critical considerations. The one is that the practical use introduces a gap between the head and the headrest, which can change as the listener moves. The other is that environmental reverberation can degrade binaural reproduction quality, since it may interfere with the perception of sound directionality and spatial cues, making the audio less immersive or harder to interpret.

An alternative configuration with loudspeakers positioned laterally on a test rig attached to an aircraft seat was investigated in [11]. The study implemented optimized crosstalk cancellation filters designed to minimize the error within a limited spatial region around the listener's head, referred to as space average (SA). Although this approach demonstrated an improved tolerance to listener displacement, the lateral loudspeaker arrangement presents practical implementation challenges for real-world applications such as automotive audio systems.

A recent study by the authors of the present paper on binaural audio headrests [18] found that crosstalk cancellation was less effective in headrest configurations compared to the conventional frontal setup. Moreover, the space average methodology delivered better performance than the well-known fast deconvolution method [6] when listeners moved laterally around the central position. Interestingly, subjective experiments demonstrated a high spatial impression even in the absence of crosstalk cancellation for a lateral setup similar to that in [11].

As a follow-up to the study in [18], in this paper, crosstalk cancellation is investigated for headrest setup, employing the previously mentioned space average methodology combined with the causality constraint method. The causality constraint, based on inverting the minimum-phase component of the matrix of the HRTFs and windowing the residual anti-causal, is explored as a method to calculate a good crosstalk filter with zero modeling delay. This approach aims to concentrate the filter's energy in the early samples of the crosstalk filter, potentially reducing the overall filter length. Objective and subjective results are presented.

Section II reviews crosstalk cancellation algorithms previously proposed in the literature, while Section III introduces the algorithm proposed in this work. Section IV desribes the experimental setup. Section V presents both objective and subjective evaluations of the crosstalk cancellation algorithms, and Section VII concludes the paper.

II. Reference crosstalk cancellation algorithms

This section presents the two reference crosstalk cancellation algorithms, which are analyzed in II-A and II-B.

A. Fast Deconvolution with Regularization

The fast deconvolution method with regularization, introduced by Kirkeby et al. [6], is an established technique for crosstalk cancellation that relies on measurements taken at a single position. The regularization term constrains the power output of the optimal filters while controlling their temporal decay. This mitigates artifacts arising from circular convolution (i.e., wrapped around effect), although at the cost of reduced inversion accuracy. The crosstalk matrix is computed as:

$$\begin{aligned}\mathbf{W}(k) &= \left[\mathbf{H}^H(k)\mathbf{H}(k) + \beta\mathbf{I}\right]^{-1}\mathbf{H}^H(k)\, e^{-z\Delta} = \\ &= \left[\frac{\det(\mathbf{H}(\mathbf{k}))^*}{\det(\mathbf{H}(\mathbf{k}))^*\det(\mathbf{H}(\mathbf{k})) + \beta}\right]\mathrm{adj}(\mathbf{H}(\mathbf{k}))\, e^{-z\Delta},\end{aligned} \tag{1}$$

where $\mathbf{H}(k)$ is the HRTF matrix, i.e., a 2×2 matrix of complex frequency responses, defined as follows:

$$\mathbf{H}(k) = \begin{bmatrix} H_{\mathrm{LL}}(k) & H_{\mathrm{RL}}(k) \\ H_{\mathrm{LR}}(k) & H_{\mathrm{RR}}(k) \end{bmatrix}. \tag{2}$$

$\mathbf{H}(k)^H$ is the Hermitian (i.e., conjugate transpose) of the matrix $\mathbf{H}(k)$, $\mathbf{I}$ is the identity matrix, and the regularization factor $\beta = 0.05$ is determined to be suitable for the experimental head-related transfer functions. The value of the regularization factor has been empirically chosen to guarantee reasonable loudspeaker levels and good perceived audio quality, as determined through subjective evaluation.

B. Space Average with Regularization

The space average algorithm was previously introduced in [11], [18]. The crosstalk filter is derived from an average of N different HRTF matrices $\mathbf{H}(k, p_n)$, each measured at the listener position p_n, for $n = 1, \ldots, N$. The resulting crosstalk filters, denoted as $\mathbf{W}_{\mathrm{SA}}(k)$, are computed as follows:

$$\begin{aligned}\mathbf{W}_{\mathrm{SA}}(k) &= [\mathbf{A}_{\mathrm{HH}} + \beta\mathbf{I}]^{-1}\mathbf{A}_{\mathrm{H}} = \\ &= \frac{\mathrm{adj}(\mathbf{A}_{\mathrm{HH}} + \beta\mathbf{I})\mathbf{A}_{\mathrm{H}}}{\det(\mathbf{A}_{\mathrm{HH}} + \beta\mathbf{I})}\, e^{-z\Delta},\end{aligned} \tag{3}$$

where

$$\mathbf{A}_{\mathrm{HH}} = \sum_{n=1}^{N} \mathbf{H}^H(k, p_n)\mathbf{H}(k, p_n), \tag{4}$$

and

$$\mathbf{A}_{\mathrm{H}} = \sum_{n=1}^{N} \mathbf{H}^H(k, p_n). \tag{5}$$

For the experiments, a regularization factor of $\beta = 1$ was selected based on empirical evaluation, following the approach used in the fast deconvolution method. The regularization factor β has the same effect as that described for the fast deconvolution in Sec. II-A.

III. Proposed crosstalk cancellation algorithm

The proposed algorithm integrates a causality constraint into the space average framework (Eq. (3)). This constraint was originally introduced in [7] for the fast deconvolution method (Eq. (1)), to reduce wrap-around effects and echo artifacts. In this case, the crosstalk filters $\mathbf{W}_{\mathrm{SA,CC}}(k)$ are obtained using the following formula:

$$\mathbf{W}_{\mathrm{SA,CC}}(k) = \\ = \frac{1}{\det(\mathbf{A}_{\mathrm{HH}} + \beta\mathbf{I})^{+}} \left[\frac{\mathrm{adj}(\mathbf{A}_{\mathrm{HH}} + \beta\mathbf{I})\mathbf{A}_{\mathrm{H}}\, e^{-z\Delta}}{\det(\mathbf{A}_{\mathrm{HH}} + \beta\mathbf{I})^{-}} \right]_{+}, \quad (6)$$

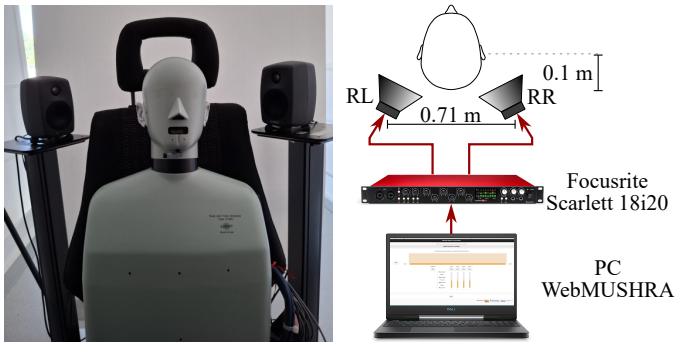
where $\mathbf{A}_{\mathrm{HH}}$ and $\mathbf{A}_{\mathrm{H}}$ are given in Eq. (4) and Eq. (5), respectively. As presented in Eq. (6), the causality constraint is achieved by first performing spectral factorization to decompose the minimum stable causal and stable anti-causal components, denoted as $(\cdot)^{+}$ and $(\cdot)^{-}$, respectively, using the Wiener–Hopf decomposition. Adopting the approach as described in the original work, the spectral factorization is implemented through the cepstral domain. The minimum-phase component is then inverted to obtain a minimum-phase inverse. To isolate the causal part of the term within the square brackets, denoted as $[\cdot]_{+}$, an asymmetric Tukey window is applied as in the original work, effectively suppressing the second half of the impulse response by setting it to zero. A regularization factor $\beta = 1$ was chosen, consistent with the value used in the space average method without the causality constraint (Eq. (3)).

IV. Experimental setup

The experimental setup used in this work is shown in Figure 1. The listener was positioned in a car seat with two Genelec 6010A loudspeakers managed by a Focusrite Scarlett 18i20 sound card. The loudspeakers are positioned at azimuthal angles of $\pm 105°$, with $0°$ defined as directly in front of the listener. This corresponds to a lateral separation of 71 cm, with the inter-speaker axis located 10 cm behind the center of the listener's head, replicating a typical headrest configuration. Their on-axis responses are directed toward the center of the listener's head. The HRTFs were measured using the binaural mannequin Brüel & Kjær Type 4128C and NU-Tech software [19], using a sweep excitation signal and a sampling rate of 48 kHz. Head-related impulse responses were captured over 256 samples, equivalent to 5.3 ms. Measurements were taken along a linear path by shifting the mannequin laterally to both the left and right, resulting in a total of seven positions.

The minimum-phase components of the impulse responses were extracted, with smoothing applied above 200 Hz, as described in [20], [21]. The impulse responses were normalized such that the average magnitude of the ipsilateral signals, calculated over the frequency range from 200 Hz to 16.5 kHz and across all measured positions, was adjusted to 0 dB.

The FFT length and the crosstalk filter length were chosen to be 2048 samples, that is, eight times the length of the measured impulse responses, as recommended in [6] for a 2×2 CTC system, and similarly adopted by [8]. For the crosstalk filters, a modeling delay of half the FFT length was applied to the fast deconvolution (Sec. II-A) and the space average (Sec. II-B), as suggested in [6]. In contrast, no modeling delay was introduced for the causality-constrained case. Table I provides an overview of the crosstalk cancellation algorithms along with their corresponding parameters.



(a) (b)

Fig. 1: (a) Headrest setup used for the HRTF measurements and (b) scheme of the hardware setup used for the listening tests.

V. Results

This section presents objective and subjective evaluations of the three crosstalk cancellation methods. The objective evaluation examines the resulting crosstalk filters in terms of their time-domain response (Figure 2) and magnitude response (Figure 3), as well as the overall responses at the central (Figure 4) and an off-cental listening position (Figure 5). The subjective evaluation assesses the perceptual performance of the filters, focusing on localization accuracy (Table II) and spatial impression (Figure 8).

A. Objective evaluation

Figure 2 presents the crosstalk cancellation filters derived using the three approaches: (2a) fast deconvolution applied to the central position measurement, and the space average algorithm, both (2b) without and (2c) with the causality constraint. For the sake of demonstration, only the left ipsilateral filter is presented, since a symmetric configuration is assumed. The filter derived through the fast deconvolution algorithm and the space average without causality constraint methods concentrate their energy near the midpoint of the FFT length, corresponding to the chosen modeling delay of half the FFT length. In contrast, the filter derived through the space average with causality constraint shifts this energy to the initial samples, reflecting a zero modeling delay, with subsequent samples decaying toward zero. This results in a filter with smaller delay and shorter effective length, which is advantageous for dynamic crosstalk cancellation systems in terms of computational complexity. Figure 3 shows the magnitude responses of the crosstalk cancellation filters computed using the three algorithms. A slight asymmetry in the setup

Crosstalk algorithm	abbr.	β	modeling delay
Fast Deconvolution with regularization	FD	0.05	Half FFT length
Space Average with regularization	SA	1	Half FFT length
Space Average with causality constraint and regularization	SA CC	1	Zero

TABLE I: Overview of the crosstalk cancellation algorithms used in the experimental procedure, their abbreviations, regularization parameters β, and modeling delays.

is reflected in the magnitude responses, particularly in those derived using the fast deconvolution (Figure 3a) and the space average with a causality constraint (Figure 3c) methodologies.

Figures 4 and 5 present the overall responses at the assumed central position and a non-central position, respectively. The non-central position corresponds to one of the measured locations used in deriving the space average crosstalk filter, specifically representing a moderate displacement from the central point relative to the measured positions. An ideal response exhibits a flat magnitude at 0 dB for the ipsilateral responses (LL and RR) while achieving maximum possible attenuation of the contralateral responses (LR and RL). Across all three algorithms and listening positions, the ipsilateral responses (LL and RR) exhibit a consistent trend of low-frequency attenuation, beyond the minor variations among them. This attenuation is likely attributed to the strong regularization applied during the crosstalk filter design, which predominantly affects the low-frequency range for the specific set of HRTFs used. The best performance is achieved at the assumed central position when the fast deconvolution algorithm is applied (Figure 4a). In this case, the non-flat ipsilateral response highlights the effect of regularization. It is worth emphasizing that regularization was essential to ensure filters capable of delivering perceptually good audio quality. Comparing the space average responses with and without the causality constraint for the two positions, it is evident that causality enhances channel separation, as can be seen in Figures 4b and 4c for the central position and Figures 5b and 5c for the off-center position.

Moreover, it is very important to highlight that the objective metrics assess the performance of the inversion process rather than the overall performance of the crosstalk cancellation system. This is because they are based on preprocessed impulse responses, rather than the actual measured ones, particularly considering that the preprocessing maintained only the minimum-phase component of the impulse response. In contrast, the perceptual assessment reflects the true performance of the crosstalk cancellation system.

B. Subjective evaluation

A panel of 12 listeners (25 to 45 years, M=27, SD=6) participated in the experiment. The listening test included a localization task and an evaluation of spatial impression. The maximum angular error recorded was 45° for all participants; therefore, no participants were excluded from the analysis. Participants evaluated the following audio material:

- binaural signal with crosstalk algorithm developed with fast deconvolution applied to the central position measurement;
- binaural signal with crosstalk algorithm developed with space average methodology;
- binaural signal with crosstalk algorithm developed with space average with causality constraint methodology

1) Localization test: The first part assessed the localization accuracy for azimuthal sound sources. Figure 6 illustrates the graphical user interface (GUI) developed for the localization test. Binaural stimuli were generated by convolving three monaural tracks with HRTFs from the MIT Media Lab database [22] corresponding to different azimuthal positions at 0° elevation. The tested sound source locations corresponded to azimuth angles of 90°, 135°, 180°, 225°, and 270°, with 90° indicating the right, 180° the rear, and 270° the left of the listener.

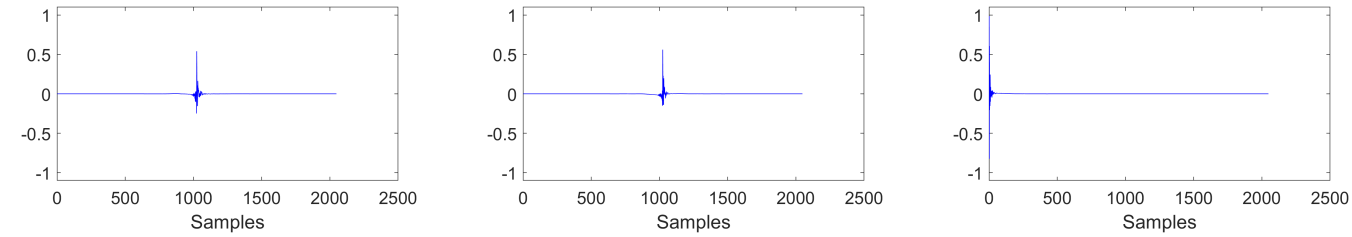
Each trial evaluated a stimulus at three fixed azimuthal angles, and each crosstalk algorithm was evaluated in separate trials. For each binaural stimulus, listeners identified the perceived source direction by selecting among the five possible discrete angles. Before the localization test, participants completed a training phase with stimuli presented at all experimental azimuthal angles.

As shown in Table II, the three algorithms demonstrated comparable localization performance, with the fast deconvolution and space average with causality constraint methods showing a slight advantage. Notably, the space average method exhibited the lowest localization accuracy across all tracks. The superior performance of the fast deconvolution method may be attributed to the listener's possible positioning at the location where the central position measurement was taken.

Focusing on the individual tracks, the classical track consistently resulted in lower percentages of correct angle identification across all crosstalk cancellation algorithms, particularly in comparison to the pop and rock tracks. This outcome indicates a potential dependency of localization performance on the characteristics of the audio material.

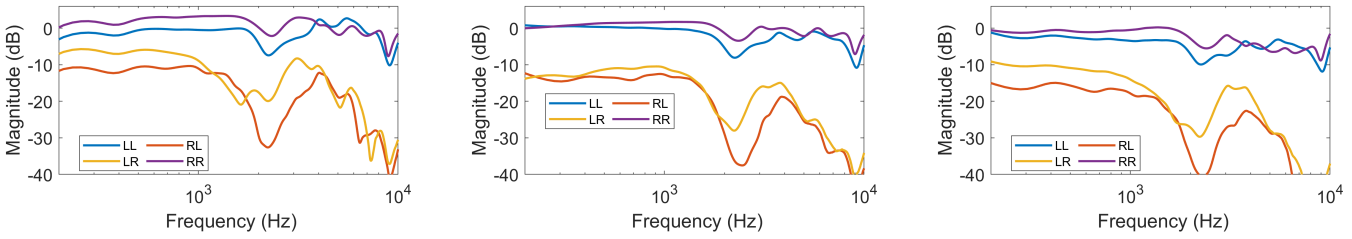
TABLE II: Percentage of correct answers in the localization experiment.

Algorithm	Classical	Pop	Rock	Mean
Fast Deconvolution	**58**%	**83**%	**89**%	**77**%
Space Average	56%	75%	75%	69%
Space Average with causality constraint	**56**%	**83**%	**86**%	**75**%



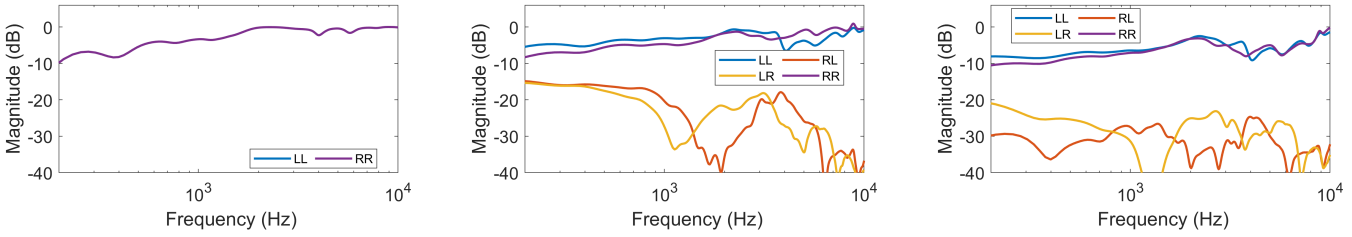
(a) (b) (c)

Fig. 2: Left ipsilateral (LL) crosstalk filter derived via the (a) fast deconvolution, (b) space average, and (c) space average with causality constraint method.



(a) (b) (c)

Fig. 3: Magnitude responses of the crosstalk filter derived via the (a) fast deconvolution, (b) space average, and (c) space average with causality constraint method.



(a) (b) (c)

Fig. 4: Overall responses derived via the (a) fast deconvolution, (b) space average, and (c) space average with causality constraint method, at the assumed central position. In (a), the contralateral responses (RL, LR) are omitted as they fall below -300 dB.

A noteworthy byproduct of the experiment was the higher accuracy observed in participants' identification of the rear angle at 180°, with 80% correct responses, and the angle at 135° (or 225°), with 77% correct responses. These results contrast with the lower accuracy for the lateral angle, which yielded only 64% correct responses.

2) Spatial impression evaluation: The second part evaluated the spatial impression, defined as "The subjective impression that the performance takes place in an appropriate spatial environment" [23]. The interface used to evaluate spatial impression is depicted in Figure 7. Participants evaluated three binaural stimuli on the bipolar discrete seven-grade scale [23], where the fast deconvolution algorithm served as the reference.

Figure 8 shows the evaluations on the bipolar scale across all three experimental tracks. The fast deconvolution algorithm consistently received a score of zero, indicating that participants reliably identified the reference condition in every trial. The statistical significance of the difference between the algorithms was evaluated by means of pairwise comparisons following a repeated measures ANOVA with Bonferroni correction. The space average algorithms consistently outperformed the fast deconvolution algorithm, except for a few minor instances ($p < 0.001$). Notably, a slight, though not statistically significant ($p < 0.325$), increase in spatial impression was observed when the causality constraint was applied.

VI. Discussion

The key innovation of this work lies in integrating the causality constraint, previously applied only to the fast deconvolution method, into the space average method, which aims to extend the listener's sweet spot. Applying this causality

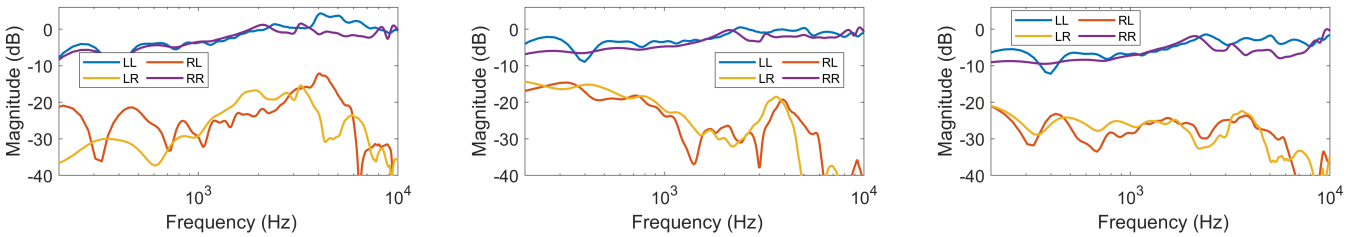


Fig. 5: Overall responses derived via the (a) fast deconvolution, (b) space average, and (c) space average with causality constraint method, at a non-central position.

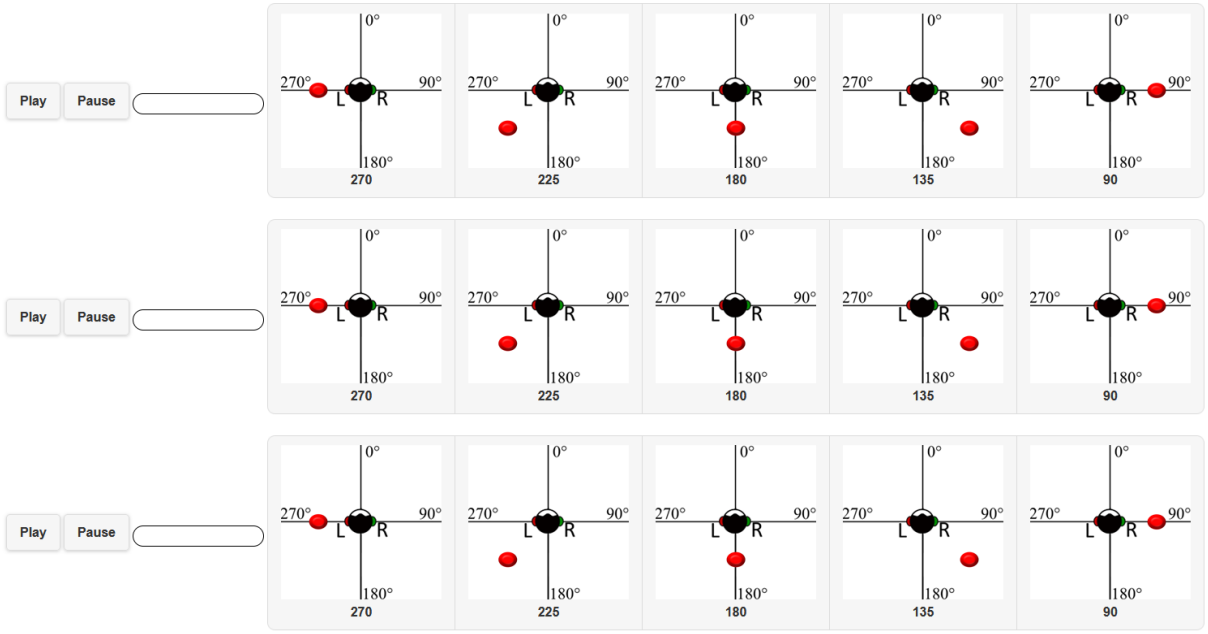


Fig. 6: Experimental GUI for the localization experiment.

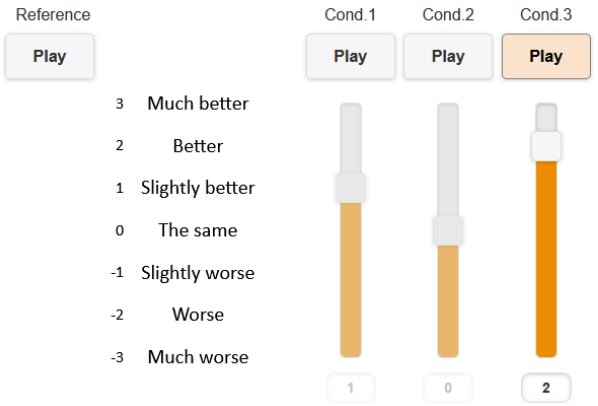


Fig. 7: Experimental GUI for the spatial impression evaluation.

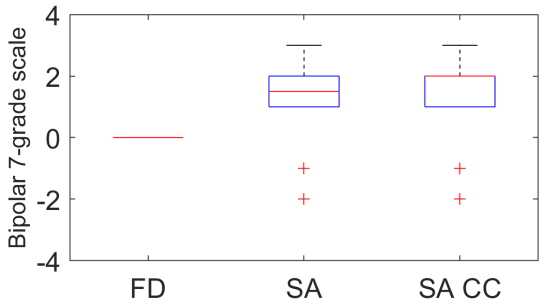


Fig. 8: Spatial Impression evaluations for the Fast Deconvolution (FD), Space Average (SA) and Space Average with causality constraint (SA CC).

constraint reduced both the filter length and the delay, while providing a slight improvement in localization and spatial impression.

Regarding the experimental setup and evaluation procedure, several considerations have to be taken into account. First, given the proximity of the loudspeakers to the listener, even slight positional shifts could significantly affect the results.

Small differences between the left and right sides of the crosstalk filters are presented, due to small asymmetries in the setup. Second, the measurements were conducted using a binaural mannequin and the objective evaluation was also based on measurements from this same mannequin, though head-related impulse responses are known to be highly individual. Third, the evaluation focused on the minimum-phase component of the measured impulse responses rather than the mixed-phase responses. Thus, the overall responses (Figures 4b and 5b) represent the performance of inversion, rather than the true responses at the listening point. Focusing on the localization task, informal participant feedback suggested that some sound sources appeared to deviate slightly from the predefined set of angles. However, since participants could only choose from these fixed options, subtle differences between the algorithms may not have been reported.

Future work will involve evaluating listeners at off-central positions to assess the enhancement offered by the space average method compared to the fast deconvolution applied at the central position. Furthermore, the experimental procedure will enable participants to select sound source positions across the continuous 360° azimuthal plane. Additionally, the proposed algorithm will be assessed in terms of computational complexity to evaluate its advantages for dynamic crosstalk cancellation systems.

VII. Conclusions

This work investigated crosstalk cancellation algorithms for a binaural audio headrest, with particular focus on maintaining effectiveness under small lateral displacements of the listener. Three approaches were evaluated: the traditional fast deconvolution method applied to the central listening position, a recently proposed space average method designed to expand the sweet spot area, and a novel method introduced in this work that incorporates a causality constraint into the space average approach.

The space average method, as a spatial extension of the traditional fast deconvolution approach, requires a modeling delay—typically half the FFT length—to ensure causality of the resulting filters. Although effective, this modeling delay increases the latency of the CTC system. The proposed method addresses this limitation by incorporating a causality constraint, previously applied to the fast deconvolution method, directly into the filter design process. In the context of binaural audio headrests, this allows the derivation of filters with zero modeling delay and reduced to half the filter length. Moreover, the perceptual evaluation indicated slightly better localization performance and spatial impression after the application of the causality constraint.

Compared to fast deconvolution, the space average method with and without the causality constraint delivered an enhanced sense of spatial impression. However, fast deconvolution showed slightly better performance in terms of localization accuracy, particularly compared to the space average without causality constraint method.

References

[1] T. Potter, Z. Cvetković, and E. De Sena, "On the relative importance of visual and spatial audio rendering on VR immersion," *Frontiers in Signal Processing*, vol. 2, p. 904866, 2022.

[2] T. Langlotz, H. Regenbrecht, S. Zollmann, and D. Schmalstieg, "Audio stickies: visually-guided spatial audio annotations on a mobile augmented reality platform," in *Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration*, 2013, pp. 545–554.

[3] M. Wong and R. Duraiswami, "Shared-space: Spatial audio and video layouts for videoconferencing in a virtual room," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*. IEEE, 2021, pp. 1–6.

[4] J. Broderick, J. Duggan, and S. Redfern, "The importance of spatial audio in modern games and virtual environments," in *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 2018, pp. 1–9.

[5] D. Johnston, H. Egermann, and G. Kearney, "The use of binaural based spatial audio in the reduction of auditory hypersensitivity in autistic young people," *International journal of environmental research and public health*, vol. 19, no. 19, p. 12474, 2022.

[6] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast Deconvolution of Multichannel Systems using Regularization," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 2, pp. 189–194, Mar. 1998.

[7] B. Masiero and M. Vorländer, "A framework for the calculation of dynamic crosstalk cancellation filters," *IEEE/ACM transactions on audio, speech, and language processing*, vol. 22, no. 9, pp. 1345–1354, 2014.

[8] T. Kabzinski and P. Jax, "A causality-constrained frequency-domain least-squares filter design method for crosstalk cancellation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2942–2956, 2021.

[9] J. Kovačević, N. Kaprocki, and A. Popović, "Review of automotive audio technologies: Immersive audio case study," in *2019 Zooming Innovation in Consumer Technologies Conference (ZINC)*. IEEE, 2019, pp. 98–99.

[10] A. Lundkvist, A. Nykänen, and R. Johnsson, "3D-sound in car compartments based on loudspeaker reproduction using crosstalk cancellation," in *130th Convention of Audio Engineering Society*. Audio Engineering Society, 2011.

[11] S. Elliott, C. House, J. Cheer, and M. Simon-Galvez, "Cross-talk cancellation for headrest sound reproduction," in *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*. Audio Engineering Society, 2016.

[12] J. Buck and D. Sachau, "Active headrests with selective delayless subband adaptive filters in an aircraft cabin," *Mechanical Systems and Signal Processing*, vol. 148, p. 107164, 2021.

[13] H. Chen, D. Long, H. Zou, S. Wang, J. Tao, and X. Qiu, "Improving near-field spatial uniformity of secondary sources for active noise control headrests," *Applied Acoustics*, vol. 230, p. 110437, 2025.

[14] S. Goose, L. Riddle, C. Fuller, T. Gupta, and A. Marcus, "Paz: In-vehicle personalized audio zones," *IEEE MultiMedia*, vol. 23, no. 4, pp. 32–41, 2015.

[15] J. Linjama and V. Välimäki, "Immersive personal sound using a surface nearfield source," in *153rd Convention of Audio Engineering Society*. Audio Engineering Society, 2022.

[16] H. Oppermann and S. Checa, "Sonic opportunities presented by personalized sound zones," in *Audio Engineering Society Conference: AES 2022 International Automotive Audio Conference*. Audio Engineering Society, 2022.

[17] E. Stanhope, L. Hobden, and S. Oxnard, "Near-field binaural rendering: Evaluating the natural channel separation of loudspeakers mounted in a headrest," in *Audio Engineering Society Conference: AES 2023 International Conference on Spatial and Immersive Audio*. Audio Engineering Society, 2023.

[18] N. A. Dourou, V. Bruschi, and S. Cecchi, "Investigation on a crosstalk cancellation algorithm for binaural audio headrest," *Journal of the audio engineering society*, 2025, in press.

[19] A. Lattanzi, F. Bettarelli, and S. Cecchi, "Nu-tech: The entry tool of the hartes toolchain for algorithms design," in *Proc. 124th Audio Engineering Society Convention*, 2008, pp. 1–8.

[20] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *Journal of the audio engineering society*, vol. 48, no. 11, pp. 1011–1031, 2000.

[21] A. Farina, A. Bellini, E. Armelloni *et al.*, "Implementation of cross-talk canceling filters with warped structures-subjective evaluation of the loudspeaker reproduction of stereo recordings," *Proc. of SHARC2000, Boston*, pp. 11–13, 2000.

[22] B. Gardner and K. Martin, "HRTF measurements of a KEMAR dummy-head microphone," *MIT Media Lab Perceptual Computing Technical Report #280*, 1994 (update on October 13, 2006). [Online]. Available: http://www.media.mit.edu

[23] ITU-R BS. 1284-2, "General methods for the subjective assessment of sound quality ," 2019.