Optimal Spatial Sampling of Plant Transfer Functions for Head-Tracked Personal Sound Zones

Yue Qiao[1] and Edgar Choueiri[1]

[1]*3D Audio and Applied Acoustics Laboratory, Princeton University, Princeton, New Jersey, 08544, USA*

Correspondence should be addressed to Yue Qiao (`yqiao@princeton.edu`)

ABSTRACT

The implementation of head tracking in personal sound zone (PSZ) reproduction was investigated in terms of the optimal spatial resolution required for sampling the plant transfer functions, which results from a trade-off between the measurement effort and the robustness of isolation performance against head movements. The plant transfer functions of an experimental PSZ system were densely measured along translational moving trajectories of a dummy head, and then downsampled to different resolutions at which the PSZ filters were computed and the isolation performance was numerically simulated. By analyzing the variation in the isolation performance, the optimal sampling resolution, above which a given minimum level of isolation can be maintained over the reproduction area, was determined as a function of head position and frequency for two separate zones. It was found that the optimal spatial sampling resolution is in general inversely proportional to the distance between the two listeners, and to that between the moving listener and the loudspeaker array. Moreover, the high-frequency part of the plant transfer functions was found to require a higher sampling resolution than the low-frequency part, while a moving *bright* zone requires a lower sampling resolution than a moving *dark* zone.

1 Introduction

Personal sound zone (PSZ) [1, 2] reproduction is one sub-area of sound field control [3] that has received wide attention from the audio and acoustics community over the past two decades, with applications for mobile devices [4], automotive cabins [5, 6], and outdoor spaces [7]. Using loudspeaker arrays and digital signal processing techniques, two listening zones, often known as bright zone (*BZ*) and dark zone (*DZ*), are rendered in the same physical space, with the aim of providing listeners with minimally-interfering individual audio programs (e.g., speech/music), or generating quiet area for one listener while delivering audio to the other. Depending on the preferred/selected perfor-

mance metric (e.g., isolation between zones or reproduced audio quality at each zone), various filter design methods have been proposed, such as Pressure Matching (PM) [8, 9, 10], Acoustic Contrast Control (ACC) [11, 12, 13], and Variable Span Trade-Off Filtering [14, 15].

Up until recently, most PSZ studies have focused on reproducing static sound zones (i.e., they are not compensated for time-varying changes in the environment), in which case the system performance can easily degrade when the actual acoustic transfer functions (ATFs) significantly deviate from those used in plant modeling (i.e., estimating the plant ATFs based on previous observations) and filter design. Although regularization

can be applied to the filter design for mitigating the effects of such ATF mismatches [16, 17, 18], the uncertainties allowed in plant modeling can lead to suboptimal performance. This is especially true for systems that directly control the pressure around or at the listeners' ears [10, 19] (as opposed to those targeting larger spaces), as the plant ATFs can be easily altered by head movements. In order to improve the robustness against head movements without sacrificing the isolation performance, recent studies have proposed solutions that apply head tracking to PSZ reproduction [20, 21]; similar approaches with head tracking can also be found in other applications, e.g., loudspeaker crosstalk cancellation (XTC) [22, 23, 24, 25] and loudspeaker equalization [26]. Other studies addressed the general case of moving sound zones, using adaptive techniques such as moving horizon framework [27], Filtered-x Least-Mean-Square algorithm [6], and Recursive Least Squares algorithm [28] to update the PSZ filters and/or estimate the plant ATFs in real-time.

Although methods have been proposed for plant modeling as well as filter generation for adaptive PSZ reproduction, the practical implementation of such a system remains challenging due to the following reasons: 1) Unlike applications that target a single listener (e.g., XTC), the presence of more than one listener in PSZ systems significantly increases the amount of plant measurements required for adapting the reproduction to listeners' simultaneous movements; 2) Compared to XTC systems where filters are also generated based on the knowledge of the plant, PSZ systems usually need to achieve a higher level of performance index for perceptual acceptance [29], suggesting a finer spatial grid of plant measurements for head tracking in PSZ applications. Given these demanding requirements on the amount and resolution of plant measurements for PSZ reproduction, it is necessary to optimize the plant spatial sampling process, such that a desired level of isolation can be preserved over a large reproduction area with least possible number of plant measurements.

In this paper, using an experimental PSZ system that directly targets listeners' ears, we sought to investigate the effects of various system parameters (e.g., frequency, the relative position of a listener in the system, and the target listener) on the optimal spatial sampling resolution and derive the rules for minimizing the number of plant measurements required for achieving a certain isolation level. Through measuring the system plant ATFs along various listener moving trajectories

and simulating the isolation performance at different plant sampling resolutions, we determined the optimal resolution for the current system separately for the cases of a moving *BZ* and a moving *DZ*, and also in terms of the low and high frequency bands.

2 Methods

We consider a general PSZ system that consists of an array of L loudspeakers and M control points (for the system that targets the ears of two listeners, $M = 4$). Each loudspeaker l has a complex gain of $g_l(\omega), l = 1, \cdots, L$, and the resulting sound pressure at each control point m is $p_m(\omega), m = 1, \cdots, M$, where ω denotes the frequency. The ATF corresponding to the loudspeaker l and the control point m is denoted as H_{ml}, which, in matrix form (known as the plant matrix) is

$$\mathbf{p} = \mathbf{H}\mathbf{g}, \tag{1}$$

where $\mathbf{p} = [p_1, \cdots, p_M]^T \in \mathbb{C}^{M \times 1}, \mathbf{H} = (H_{ml}) \in \mathbb{C}^{M \times L}$, and $\mathbf{g} = [g_1, \cdots, g_L]^T \in \mathbb{C}^{L \times 1}$. All quantities hereafter are implicitly dependent on the frequency ω.

2.1 Pressure Matching with Probabilistic Plant Modeling

We use the Pressure Matching (PM) method formulated in the frequency domain [8, 9, 10] to generate the PSZ filters as PM has control over the phase of target audio programs, compared to other methods such as ACC. Given a specified target pressure vector $\mathbf{p}_T \in \mathbb{C}^{M \times 1}$ at the control points, the original cost function J in PM is constructed as

$$J = \|\mathbf{p} - \mathbf{p}_T\|^2 = \|\mathbf{H}\mathbf{g} - \mathbf{p}_T\|^2, \tag{2}$$

and by minimizing J, the optimal loudspeaker gains $\mathbf{g}^*$ are given by

$$\mathbf{g}^* = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{p}_T, \tag{3}$$

where the $(\cdot)^H$ denotes taking the conjugate transpose. It should be noted that this form of solution only applies to overdetermined problems where $L < M$. In order to regularize the solution to ensure its robustness against a certain degree of ATF uncertainties, we adopt a probabilistic approach [18] by assuming a random distribution of the modeled ATF H_{ml} and optimizing the expectation of the resulting cost function

$$J_{prob} = \mathbb{E}\{\|\mathbf{p} - \mathbf{p}_T\|^2\}. \tag{4}$$

The corresponding optimal solution is given by

$$\mathbf{g}_{prob}^* = (\hat{\mathbf{H}}^H \hat{\mathbf{H}} + \Sigma_{m=1}^M \Sigma_m)^{-1} \hat{\mathbf{H}}^H \mathbf{p}_T, \qquad (5)$$

where $\hat{\mathbf{H}}$ contains all the expected values of $\{H_{ml}\}$, and Σ_m is expressed as

$$\Sigma_m = diag\{\sigma_{A,m1}^2, \cdots, \sigma_{A,mL}^2\}, \qquad (6)$$

where $\sigma_{A,ml}^2$ is the amplitude variance of H_{ml} and can be determined experimentally [18, 19].

2.2 Evaluation Metrics

We adopted two metrics to evaluate the isolation performance of the PSZ system, as introduced in [30]: Inter-Zone Isolation (IZI), which represents the isolation of the sound *zones* for an audio program, and Inter-Program Isolation (IPI), which represents the isolation of the target *program* from the interfering *program* in the same sound zone. We denote the two zones as Z_1, Z_2, the sub-matrices (i.e., the top/bottom two rows) of the ATF matrix $\mathbf{H}$ corresponding to $Z_{1,2}$ as $\mathbf{H}_{1,2}$, and the PSZ filters corresponding to Z_1 (or Z_2) being the *BZ* as $\mathbf{g}_1^*$ (or $\mathbf{g}_2^*$). In the current system, we focus on single-channel audio programs (i.e., a single vector $\mathbf{p}_T$ for given *BZ* and *DZ*), in which case the definition of IZI is expressed as

$$IZI_1 = \frac{\|\mathbf{H}_1 \mathbf{g}_1^*\|^2}{\|\mathbf{H}_2 \mathbf{g}_1^*\|^2}, \qquad IZI_2 = \frac{\|\mathbf{H}_2 \mathbf{g}_2^*\|^2}{\|\mathbf{H}_1 \mathbf{g}_2^*\|^2}, \qquad (7)$$

where the subscripts 1,2 refer to the target program for $Z_{1,2}$, or in other words, whether Z_1 or Z_2 is treated as *BZ*. In this case IZI is also equivalent to the commonly-used Acoustic Contrast (AC) metric [16]. Correspondingly, IPI is expressed as

$$IPI_1 = \frac{\|\mathbf{H}_1 \mathbf{g}_1^*\|^2}{\|\mathbf{H}_1 \mathbf{g}_2^*\|^2}, \qquad IPI_2 = \frac{\|\mathbf{H}_2 \mathbf{g}_2^*\|^2}{\|\mathbf{H}_2 \mathbf{g}_1^*\|^2}, \qquad (8)$$

where the subscripts 1,2 refer to the two zones, respectively.

3 Evaluation Setup

We first measured the plant Binaural Room Transfer Function (or BRTF, as the representation of ATF in the context of the current system) with a fine resolution along different head moving trajectories. Then, we downsampled the original plant BRTF grid to sparser

resolutions and used the remaining BRTFs to generate the filter sets. Finally, we simulated the system performance by convolving the full BRTF grid with each of the generated filter sets and calculating the IZI and IPI metrics. By analyzing the variation in IZI and IPI due to head movements, we determined the lowest sampling resolution for achieving a specified isolation threshold. For simplicity, we only implemented the case where only one listener moves and the other listener is fixed as reference; however, we will show that the findings can be applied to the case of two moving listeners.

3.1 System Implementation

The experimental PSZ system, which is identical to the one studied in [19], comprises a linear loudspeaker array with 8 mid-range transducers and two listeners represented by two B&K Head and Torso Simulators (HATS) with in-ear binaural microphones (Theoretica Applied Physics BACCH-BM Pro), placed in a typical listening room ($RT_{60} \approx 0.24$ s in the range 1300-6300 Hz). The translational head movements were realized with a custom made mechanical translation stage (see Fig. 1). We used synchronized exponential sine sweep (ESS) signals [31] (a variant of the traditional ESS [32] that correctly estimates higher harmonic frequency responses) to measure the BRTFs at 48 kHz sampling frequency, with each sweep having a duration of 0.5 second. All measured BRIRs (Binaural Room Impulse Responses, the time-domain counterpart of BRTFs) were truncated to the first 8192 samples for subsequent processing (it was verified that further increasing the truncation window length has no noticeable effect on the evaluation results).

3.2 Filter Generation

The PSZ filters used in the evaluation were generated with the aim of maximizing the isolation performance at matched listener positions while preserving the robustness against minor head misalignments, similar to the approach in [19]. For the the definition of target pressure and the choice of the variance matrix in Eq. 6, we refer the readers to the aforementioned reference for details. As an improvement to the original approach, all the plant BRIRs used for filter generation (i.e., the plant matrix $\mathbf{H}$ and the target pressure $\mathbf{p}_T$) were windowed to the first 4096 samples (late reverb tails discarded) to improve the robustness, as suggested

Fig. 1: Photo of the implemented system setup for head translations in the Y direction (front/back).

Fig. 2: Diagram of the measurement setup. The dotted lines around the left HATS represent the moving trajectories along which the BRTFs were captured. The black triangles represent the loudspeakers.
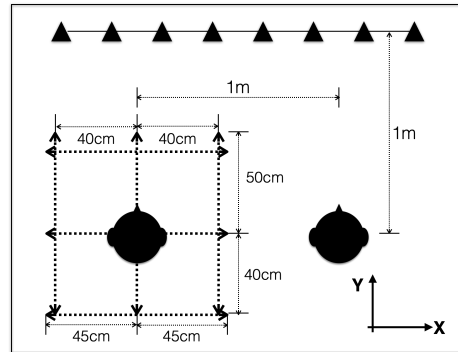
in [33, 34]. In addition, when the filter is generated for a new head position, the target pressure is re-selected corresponding to that position, as re-applying the old target might degrade the isolation performance. The filters were computed with 8192-size FFT and exported to 8192 taps, in order to avoid any discrepancy in the performance due to using short filters designed with frequency- and time-domain methods [12].

3.3 Experimental Design

In this study, we focus on translational head movements and determine the optimal spatial sampling resolution for the translation in two orthogonal directions: X (left/right) and Y (front/back), assuming height changes are less common in practical applications. The originally measured BRTF grid has a resolution of 1 cm between two adjacent sampling points. Fig. 2 shows the moving trajectories of the left HATS along which the BRTFs of both HATS were measured: three 90-cm-long lines for the X and Y directions, respectively. In the evaluation, the grid was downsampled to the resolutions of 3, 5, 10 cm, and the PSZ filters generated from these sparser grids were "assigned" to the original sampling points in a nearest-neighbor manner, i.e., no interpolation was performed between the filters when there is a mismatch between the plant BRTF and its assigned filter.

4 Simulation Results

We show the simulated IZI and IPI for the cases where the PSZ filters were generated with the full and downsampled plant BRTF grids measured along the specified

trajectories. Recall in Sec. 2.2 that there are two sets of IZI and IPI metrics defined for the two HATS and their corresponding target programs; here we only present the results for IZI_2 and IPI_2 (the subscripts of which are thereafter neglected) as the position-dependent target pressure for the left *BZ* also affects the sound pressure level (SPL) in *BZ*, and therefore complicates the analysis of the isolation performance. In this context, *IZI* and *IPI* can be regarded as the isolation performance for a moving *DZ* (with a static *BZ*), and that for a moving *BZ* (with a static *DZ*), respectively. All results to be presented were processed with 1/3-octave smoothing [35] for better visualization, and band-limited to 200-7000 Hz due to Signal-to-Noise Ratio limitations and the working range of the transducers.

We first examine the best case scenario where the PSZ filters are generated with highest resolution, matching the full grid of plant BRTFs, along the two trajectories that intersect at the center position of the left HATS (see Fig. 2). Fig. 3 shows the corresponding IZI and IPI with respect to frequency and the relative head position, with the black contour line indicating 20 dB of isolation. It can be seen that IZI and IPI achieve above 20 dB irrespective of the HATS position at most frequencies above 500 Hz. This implies that, ideally, a desired level of isolation can be preserved over a relatively large listening area. At lower frequencies, however, the isolation threshold is not always achieved due to the limitations of the system (e.g., room effects, loudspeaker array layout, and number of loudspeakers), and we also observe a dependency of IZI and IPI on

the head position. Specifically, IZI and IPI levels start to decrease as the left HATS moves closer to the right HATS in the X direction, likely due to the occlusion and scattering effects of the other listener; however, such a trend is not clearly observed for the Y translations, meaning that front/back movements have a lower impact on the low-frequency isolation performance. At frequencies roughly above 1 kHz, we notice some interference-like patterns for the IPI corresponding to X translations and a monotonic decrease in IPI as the left HATS moves away from the loudspeakers in the Y direction. This is potentially due to the changes in the target pressure for the left *BZ* as the left HATS moves in both directions.

Next, we show how the expected isolation level is affected by applying filter sets that were generated from the downsampled plant BRTF grids. Figs. 4 and 5 show the IZI and IPI maps with sparse filter sets (updated every 3, 5, and 10 cm) for the same two trajectories as in the previous case. We see that as the spatial sampling becomes sparser, the isolation level drops more rapidly in between the two matched filters. We note that IPI is generally more robust than IZI, as the former is associated with the static *DZ*, which is indirectly affected by head movements of the moving *BZ*, and the latter corresponds to the moving *DZ* and therefore is less robust against head misalignments. Comparing X and Y translations, we find a lower robustness in the IZI for Y translations than that for X translations (e.g., by comparing the size of areas above 20 dB for the plots with 10-cm resolution), but the opposite is true for IPI. From these observations we conclude that Y translations can have a higher impact on the isolation of the moving *DZ* (corresponding to IZI), but a lower impact on that of the static *DZ* (corresponding to IPI), compared to X translations. However, this conclusion may no longer hold for PSZ systems with different array layouts.

The optimal resolution, which leads to a minimum of 20 dB isolation above 500 Hz for most head positions, not only depends on the target zone/listener, but is also strongly affected by the head position as well as frequency. For example, from Fig. 4 we note that, while a resolution of 5 cm is sufficient for sampling both a moving *DZ* (indicated by IZI) and a moving *BZ* (indicated by IPI) in the X direction, for frequencies between 500 and 1500 Hz (marked by the dashed line in the plots), it does not apply to higher frequencies above 1500 Hz, unless the resolution is increased to 3 cm or higher. Particularly for IZI, the required resolution to maintain

the isolation level increases as the left HATS moves closer to the right one, meaning that the optimal spatial sampling grid would be non-uniformly distributed along the X direction. We also find, for Y translations (Fig. 5), the difference between the optimal sampling resolutions for IZI and IPI: 5 cm is sufficient for IPI at most frequencies, while 3 cm or higher is required by IZI. Compared to X translations, however, the spatial non-uniformity is not observed along the Y direction.

Lastly, we compare the isolation performance resulting from the same spatial sampling resolution at parallel trajectories. Figs. 6 and 7 show the results for the three parallel trajectories in each moving direction, with a 3-cm resolution of plant spatial sampling. We note that although the robustness at frequencies below 1500 Hz remains mostly unchanged for different moving trajectories, the higher-frequency robustness shows a clear dependency on the offset positions. In Fig. 6, we see that as the left HATS moves away from the loudspeakers, the robustness regarding both IZI and IPI increases, meaning that the optimal resolution decreases with increasing the distance to the loudspeakers. This is expected as near-field BRTFs generally have more variations with head movements than those in the far field. In Fig. 7, a similar trend appears as the distance between the two HATS decreases. These observations further corroborate the finding from above that the optimal spatial sampling resolution is strongly dependent on the relative position of the listener.

5 Discussion

It is useful to generalize the above findings to other PSZ systems. First, we expect different values of IZI and IPI if the filters were designed with methods other than PM, e.g., ACC, as the latter can lead to higher isolation than the former but at a cost of phase distortion [9, 2]. Besides, other system-specific factors that may affect the results include filter design parameters (e.g., regularization level and target pressure specification), loudspeaker array setup (e.g., layout and number of transducers), and room acoustics characteristics (e.g., Reverberation Time and Direct-to-Reverberant Ratio). Correspondingly, the filter robustness and the derived optimal spatial sampling resolution may be changed slightly. Despite these sources of discrepancy, we argue that the qualitative optimization rules, such as the nonuniform distribution of optimal spatial sampling, are applicable to any PSZ system of similar dimensions,
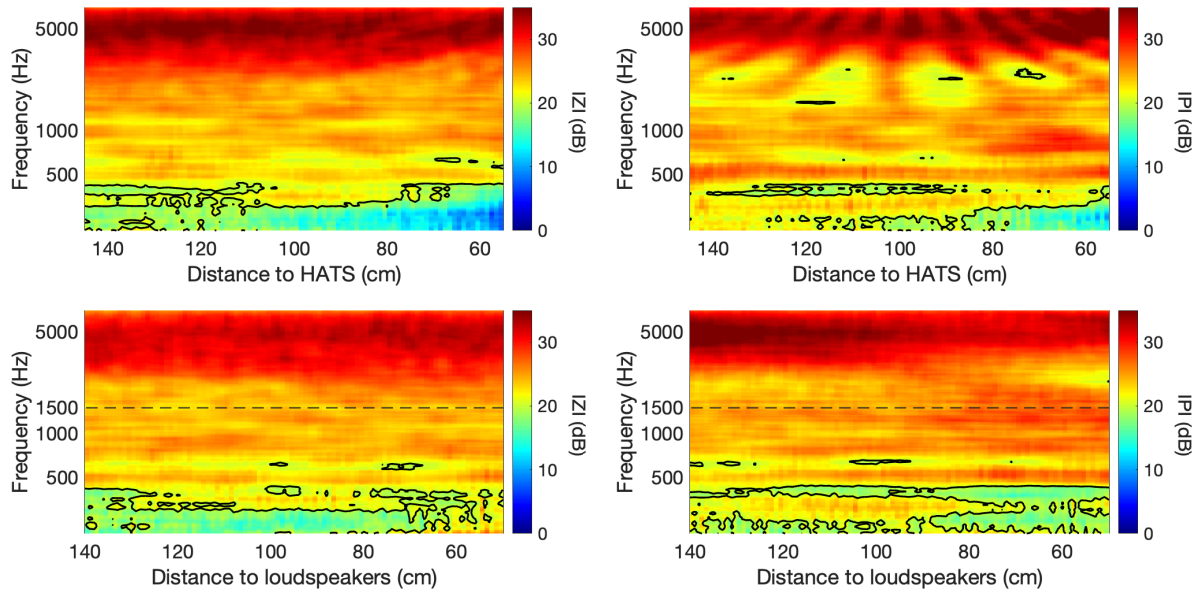
Fig. 3: Simulated IZI and IPI for the X (top row) and Y (bottom row) head translations with PSZ filters generated at highest resolution (1 cm).
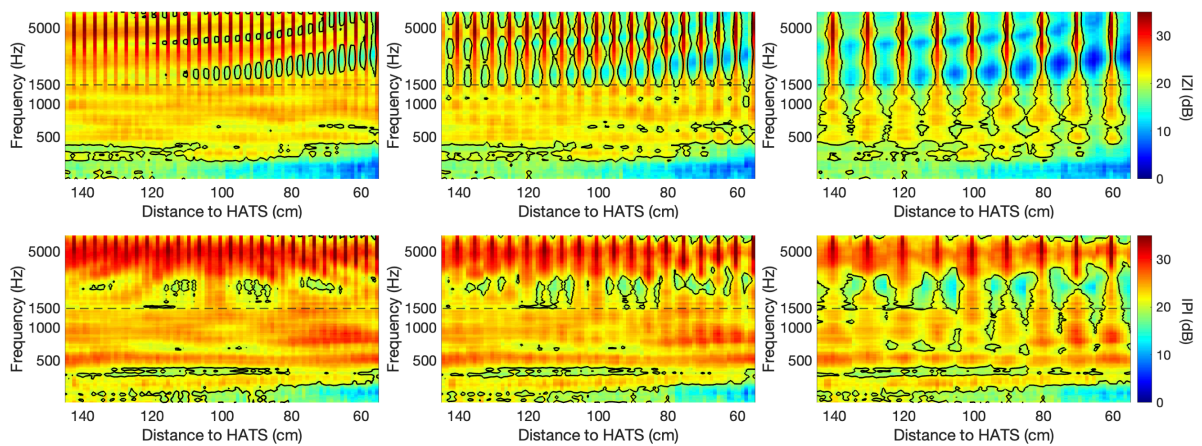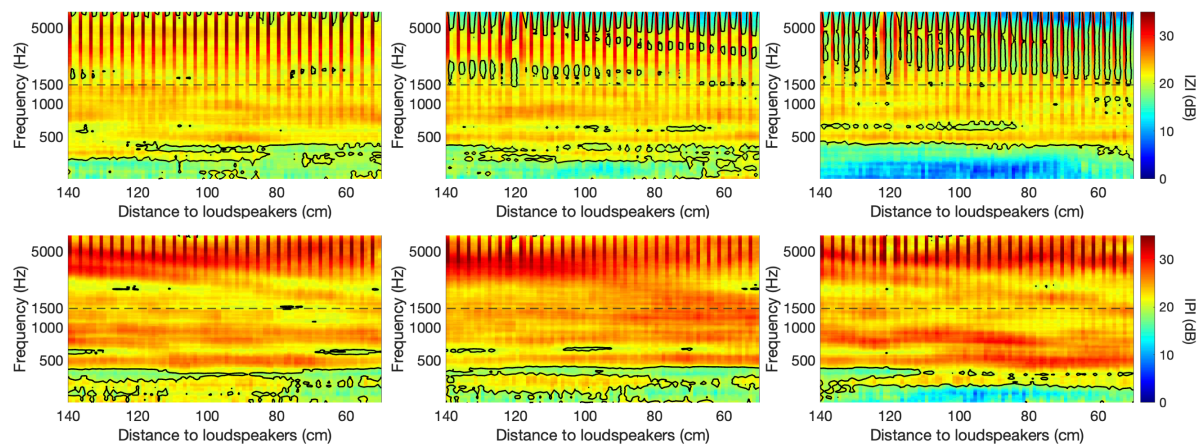

Fig. 4: Simulated IZI (top) and IPI (bottom) for X translations with filters generated using downsampled BRTFs at 3-cm (left column), 5-cm (middle column), and 10-cm (right column) resolutions.
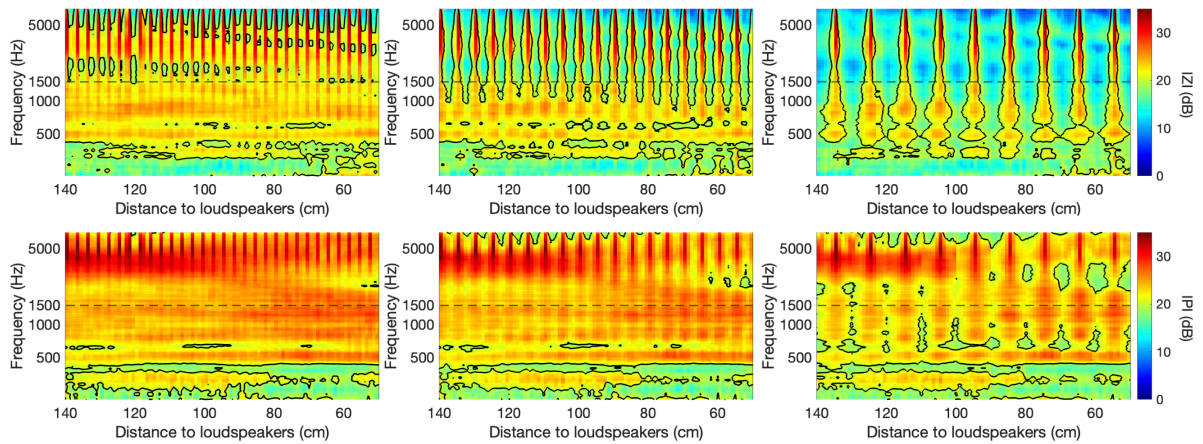
Fig. 5: Simulated IZI (top) and IPI (bottom) for Y translations with filters generated using downsampled BRTFs at 3-cm (left column), 5-cm (middle column), and 10-cm (right column) resolutions.
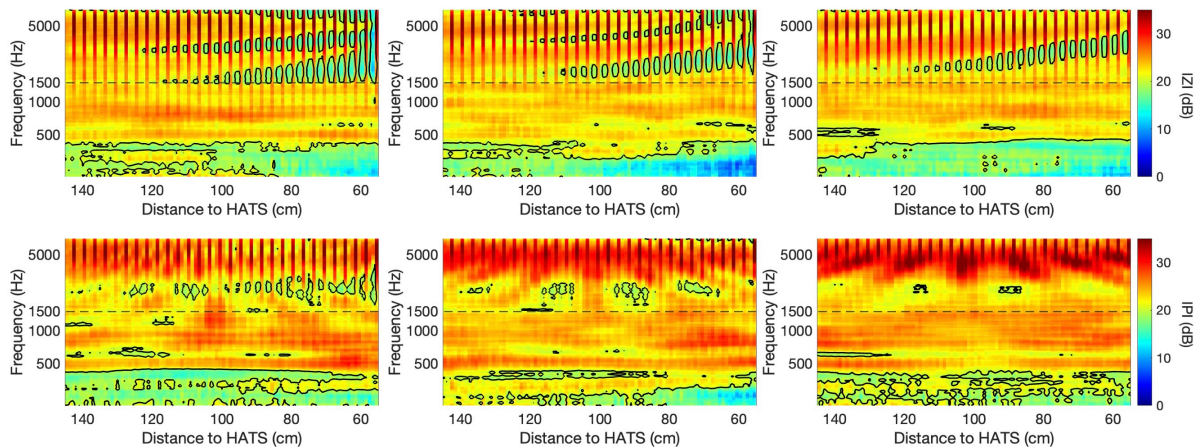
Fig. 6: Simulated IZI (top row) and IPI (bottom row) for X translations at three Y-direction offsets: 50 cm forward (left column), no offset (middle column), and 40 cm backward (right column), with 3-cm spatial sampling resolution.

Fig. 7: Simulated IZI (top row) and IPI (bottom row) for Y translations at three X-direction offsets: 45 cm towards left (left column), no offset (middle column), and 45 cm towards right (right column), with 3-cm spatial sampling resolution.

as they are determined by the fundamental acoustic properties of the system, such as the coupling of two listeners' BRTFs in the near field [19]. Furthermore, in the case of two moving listeners, their corresponding plant BRTFs can also be sampled with different optimal resolutions as long as *BZ* and *DZ* are not exchanged.

In addition to optimizing plant spatial sampling, the results also shed light on improving the performance of head-tracked PSZ systems through other approaches. For example, the interpolation between plant ATFs, although not investigated here, can be optimized by posing position-dependent accuracy constraints. Moreover, the real-time plant modeling approaches that utilize adaptive filtering [28] can also be improved by defining the required convergence time based on the listener position. Lastly, the observed transitional frequency at around 1500 Hz can be used as a natural crossover, above which the isolation performance can be traded off for higher robustness by using more generalized plant modeling (e.g., the analytical plant models in [20, 21]).

6 Conclusion

Head tracking is critical to improving the robustness of the PSZ system. In order to minimize the number of plant measurements required for maintaining high isolation, we investigated the effects of different plant spatial sampling resolutions through measurements of the plant BRTFs of an experimental PSZ system along

translational trajectories of a moving dummy head, and numerical simulations of system performance regarding different sampling resolutions. The above results lead to the following optimization rules: In order to exceed a given level of isolation over a large reproduction area, the plant transfer functions of a head-tracked PSZ system should be spatially sampled more densely:

- as the distance between the two listeners decreases, and as the distance between the listener and the loudspeaker array decreases;

- for the reproduction of high-frequency audio content (roughly above 1500 Hz) than for reproducing lower frequencies;

- for capturing the plant ATFs associated with a moving *DZ* than for capturing those related to a moving *BZ*.

Such rules can be used as guidelines to the practical implementation of head-tracked PSZ systems. Future work will focus on optimizing the plant spatial sampling regarding head rotations and accelerating the general plant measurement process.

References

[1] Druyvesteyn, W. and Garas, J., "Personal sound," *Journal of the Audio Engineering Society*, 45(9), pp. 685–701, 1997.

[2] Betlehem, T., Zhang, W., Poletti, M. A., and Abhayapala, T. D., "Personal sound zones: Delivering interface-free audio to multiple listeners," *IEEE Signal Processing Magazine*, 32(2), pp. 81–91, 2015.

[3] Yang, J., Wu, M., and Han, L., "A Review of Sound Field Control," *Applied Sciences*, 12(14), p. 7319, 2022.

[4] Elliott, S. J., Cheer, J., Choi, J.-W., and Kim, Y., "Robustness and regularization of personal audio systems," *IEEE Transactions on Audio, Speech, and Language Processing*, 20(7), pp. 2123–2133, 2012.

[5] Cheer, J., Elliott, S. J., and Gálvez, M. F. S., "Design and implementation of a car cabin personal audio system," *Journal of the Audio Engineering Society*, 61(6), pp. 412–424, 2013.

[6] Vindrola, L., Melon, M., Chamard, J.-C., and Gazengel, B., "Use of the filtered-x least-mean-squares algorithm to adapt personal sound zones in a car cabin," *The Journal of the Acoustical Society of America*, 150(3), pp. 1779–1793, 2021.

[7] Heuchel, F. M., Caviedes-Nozal, D., Brunskog, J., Agerkvist, F. T., and Fernandez-Grande, E., "Large-scale outdoor sound field control," *The Journal of the Acoustical Society of America*, 148(4), pp. 2392–2402, 2020.

[8] Poletti, M., "An investigation of 2-d multizone surround sound systems," in *Audio Engineering Society Convention 125*, Audio Engineering Society, 2008.

[9] Chang, J.-H. and Jacobsen, F., "Sound field control with a circular double-layer array of loudspeakers," *The Journal of the Acoustical Society of America*, 131(6), pp. 4518–4525, 2012.

[10] Vindrola, L., Melon, M., Chamard, J.-C., and Gazengel, B., "Pressure matching with forced filters for personal sound zones application," *Journal of the Audio Engineering Society*, 68(11), pp. 832–842, 2020.

[11] Choi, J.-W. and Kim, Y.-H., "Generation of an acoustically bright zone with an illuminated region using multiple sources," *The Journal of the Acoustical Society of America*, 111(4), pp. 1695–1700, 2002.

[12] Gálvez, M. F. S., Elliott, S. J., and Cheer, J., "Time domain optimization of filters used in a loudspeaker array for personal audio," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(11), pp. 1869–1878, 2015.

[13] Møller, M. B. and Olsen, M., "Sound zones: On performance prediction of contrast control methods," in *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*, Audio Engineering Society, 2016.

[14] Lee, T., Nielsen, J. K., and Christensen, M. G., "Signal-adaptive and perceptually optimized sound zones with variable span trade-off filters," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28, pp. 2412–2426, 2020.

[15] Brunnström, J., Koyama, S., and Moonen, M., "Variable Span Trade-Off Filter for Sound Zone Control with Kernel Interpolation Weighting," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1071–1075, IEEE, 2022.

[16] Coleman, P., Jackson, P. J., Olik, M., Møller, M., Olsen, M., and Abildgaard Pedersen, J., "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *The Journal of the Acoustical Society of America*, 135(4), pp. 1929–1940, 2014.

[17] Olsen, M. and Møller, M. B., "Sound zones: on the effect of ambient temperature variations in feed-forward systems," in *Audio Engineering Society Convention 142*, Audio Engineering Society, 2017.

[18] Møller, M. B., Nielsen, J. K., Fernandez-Grande, E., and Olesen, S. K., "On the influence of transfer function noise on sound zone control in a room," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(9), pp. 1405–1418, 2019.

[19] Qiao, Y. and Choueiri, E., "The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions," in *Audio Engineering Society Convention 152*, Audio Engineering Society, 2022.

[20] Gálvez, M. F. S., Menzies, D., and Fazi, F. M., "Dynamic audio reproduction with linear loudspeaker arrays," *Journal of the Audio Engineering Society*, 67(4), pp. 190–200, 2019.

[21] Qiao, Y. and Choueiri, E., "Real-time Implementation of the Spectral Division Method for Binaural Personal Audio Delivery with Head Tracking," in *Audio Engineering Society Convention 151*, Audio Engineering Society, 2021.

[22] Ma, X., Hohnerlein, C., and Ahrens, J., "Concept and Perceptual Validation of Listener-Position Adaptive Superdirective Crosstalk Cancellation Using a Linear Loudspeaker Array," *Journal of the Audio Engineering Society*, 67(11), pp. 871–881, 2019.

[23] Bruschi, V., Nobili, S., Bettarelli, F., and Cecchi, S., "Listener-position Sub-band Adaptive Crosstalk Canceller using HRTFs Interpolation for Immersive Audio Systems," in *Audio Engineering Society Convention 150*, Audio Engineering Society, 2021.

[24] Kabzinski, T. and Jax, P., "An adaptive crosstalk cancellation system using microphones at the ears," in *Audio Engineering Society Convention 147*, Audio Engineering Society, 2019.

[25] Masiero, B. and Vorländer, M., "A framework for the calculation of dynamic crosstalk cancellation filters," *IEEE/ACM transactions on audio, speech, and language processing*, 22(9), pp. 1345–1354, 2014.

[26] Lindfors, J., Liski, J., and Välimäki, V., "Loudspeaker Equalization for a Moving Listener," *Journal of the Audio Engineering Society*, 70(9), pp. 722–730, 2022.

[27] Møller, M. B. and Østergaard, J., "A moving horizon framework for sound zones," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28, pp. 256–265, 2019.

[28] Sipei, Z. and Burnett, I. S., "Adaptive personal sound zones systems with online plant modelling," in *Proceedings of the 24th International Congress on Acoustics (ICA 2022)*, 2022.

[29] Canter, N. and Coleman, P., "Delivering personalised 3D audio to multiple listeners: Determining the perceptual trade-off between acoustic contrast and cross-talk," in *Audio Engineering Society Convention 150*, Audio Engineering Society, 2021.

[30] Qiao, Y., Guadagnin, L., and Choueiri, E., "Isolation performance metrics for personal sound zone reproduction systems," *JASA Express Letters*, 2(10), p. 104801, 2022.

[31] Novak, A., Lotton, P., and Simon, L., "Synchronized swept-sine: Theory, application, and implementation," *Journal of the Audio Engineering Society*, 63(10), pp. 786–798, 2015.

[32] Farina, A., "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio engineering society convention 108*, Audio Engineering Society, 2000.

[33] Ebri, M., Strozzi, N., Fazi, F. M., Farina, A., and Cattani, L., "Individual Listening Zone with Frequency-Dependent Trim of Measured Impulse Responses," in *Audio Engineering Society Convention 149*, Audio Engineering Society, 2020.

[34] Molés-Cases, V., Elliott, S. J., Cheer, J., Piñero, G., and Gonzalez, A., "Weighted pressure matching with windowed targets for personal sound zones," *The Journal of the Acoustical Society of America*, 151(1), pp. 334–345, 2022.

[35] Tylka, J. G., Boren, B. B., and Choueiri, E. Y., "A Generalized Method for Fractional-Octave Smoothing of Transfer Functions that Preserves Log-Frequency Symmetry," *Journal of the Audio Engineering Society*, 65(3), pp. 239–245, 2017.