

Isolation performance metrics for personal sound zone reproduction systems

Yue Qiao,^{1,a)}  Léo Guadagnin,² and Edgar Choueiri¹

¹3D Audio and Applied Acoustics Laboratory, Princeton University, Princeton, New Jersey 08544, USA

²FOCAL-JMlab, La Talaudière, France

yqiao@princeton.edu, l.guadagnin@focal.com, choueiri@princeton.edu

Abstract: Two isolation performance metrics, inter-zone isolation (IZI) and inter-program isolation (IPI), are introduced for evaluating personal sound zone (PSZ) systems. Compared to the commonly used acoustic contrast metric, IZI and IPI are generalized for multichannel audio and quantify the isolation of sound zones and of audio programs, respectively. The two metrics are shown to be generally non-interchangeable and suitable for different scenarios, such as generating dark zones (IZI) or minimizing audio-on-audio interference (IPI). Furthermore, two examples with free-field simulations are presented and demonstrate the applications of IZI and IPI in evaluating PSZ performance in different rendering modes and PSZ robustness. © 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

[Editor: David C. Swanson]

<https://doi.org/10.1121/10.0014604>

Received: 29 June 2022 Accepted: 13 September 2022 Published Online: 4 October 2022

1. Introduction

Personal sound zone (PSZ)¹ reproduction aims to deliver, using loudspeakers, individual audio programs to listeners in the same space with minimum interference between programs. In PSZ reproduction, given a specific audio program, a *bright zone* (BZ) refers to the area where the program is rendered for the listener, while a *dark zone* (DZ) refers to the area where the program is attenuated. For a specific listener, the audio programs are categorized into either *target program*, which is rendered for the intended listener with best possible audio quality, or *interfering program*, which is delivered to a different listener but may interfere with the target program for the intended listener.

Over the past two decades, various PSZ reproduction systems have been implemented with different loudspeaker configurations and sound zone specifications depending on the application scenario. The loudspeakers used in a PSZ system have been configured as linear,^{2–5} circular,^{6–11} or arc-shaped^{12,13} arrays in the far field or in the near field (e.g., head-rest loudspeakers in automotive cabins^{14,15}). The sound zones have been designed to be as large as a few meters^{2,3,6–9} or as small as a region that only includes the listener's heads or ears.^{4,5,7,10,12–15}

When evaluating the performance of PSZ systems, a commonly adopted metric is the so-called acoustic contrast (AC), defined by Choi and Kim¹⁶ as the ratio of the acoustic energy in BZ to that in DZ. Although the AC metric gives a measure of the separation between two sound zones, there are several limitations associated with its definition. First, AC is calculated using the sound pressure resulting from a *single-channel* input that corresponds to rendering *mono* audio programs. As will be shown, it does not reflect the isolation performance of other rendering modes, where *multichannel* programs with different inter-channel correlations (e.g., stereo and binaural programs¹⁷) are specified as input. In addition, while AC represents the isolation between sound zones, it does not give a measure of the level difference between programs in the same zone, which is more relevant to the listener's perception of audio-on-audio interference^{10,11,18} (excluding psychoacoustic masking effects).

In this paper, we introduce separate PSZ performance metrics for the sound zone and audio program isolation, which are independent of both the number of the channels in a program and the correlation between channels. With these defined metrics, it is possible to evaluate the isolation performance given arbitrary program specifications from two complementary perspectives (sound zone isolation and audio program isolation). We also present examples that utilize the defined metrics to illustrate (1) the effects of different rendering modes (e.g., mono/stereo programs or binaural programs with built-in cross talk cancellation) on the isolation performance and (2) the effective physical boundaries of a PSZ within which the isolation performance is preserved.

The rest of the paper is organized as follows: a mathematical definition of the PSZ system is presented in Sec. 2; the new metrics are defined and discussed in Sec. 3; and free-field simulations are used to illustrate two example use cases of the metrics in Sec. 4, followed by the conclusion and further discussion on the metrics and their applications in Sec. 5.

^{a)} Author to whom correspondence should be addressed.

2. Problem definition

Without loss of generality, we consider a PSZ system (illustrated in Fig. 1) with L loudspeakers and two sound zones, Z_A and Z_B . A total of K control points are defined in two zones, with the subset in each zone denoted as $\{K_A, K_B\}$, respectively. A total of I channels of signals represent the inputs to the system and are divided into two subsets corresponding to the two audio programs for the two zones, denoted as $\{I_A, I_B\}$. We use the term *channel* to refer to the particular input audio signal and the term *program* to refer to the unified audio content, which may contain one or more channels, as in mono or stereo/binaural programs. All subsequent quantities are implicitly dependent on frequency ω ; plain font is used for scalars and bold font for vectors (in lowercase) and matrices (in uppercase).

Given the complex loudspeaker gains $\mathbf{g} \in \mathbb{C}^{L \times 1}$, the sound pressure vector $\mathbf{p} \in \mathbb{C}^{K \times 1}$ at the control points is determined by

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

where $\mathbf{H} = \{H_{kl}\} \in \mathbb{C}^{K \times L}$ is the acoustic transfer function matrix between the loudspeakers and the control points. Furthermore, the loudspeaker gains for a channel s_i of an audio program are given by

$$\mathbf{g} = \mathbf{c}_i s_i, \tag{2}$$

where $\mathbf{c}_i \in \mathbb{C}^{L \times 1}$ corresponds to the PSZ filters designed for the channel s_i . Combining Eqs. (1) and (2), the input channels and the control point pressure are related by

$$\mathbf{p} = \mathbf{H}\mathbf{C}\mathbf{s}, \tag{3}$$

where $\mathbf{s} \in \mathbb{C}^{I \times 1}$ denotes a vector of input channels, and $\mathbf{C} = \{C_{li}\} \in \mathbb{C}^{L \times I}$ denotes the matrix containing PSZ filters for each channel. We further introduce the *system transfer function matrix*, \mathbf{M} , as the product of transfer function and filter matrices,

$$\mathbf{M} = \mathbf{H}\mathbf{C}, \tag{4}$$

where $\mathbf{M} = \{M_{ki}\} \in \mathbb{C}^{K \times I}$ and relates the input channels to the pressure at the control points and will be mainly used in the subsequent metric definition. We also note that the size of \mathbf{M} is independent of the number of loudspeakers.

3. Metric definition

Two aspects of the isolation performance of PSZ systems need to be considered: the isolation between two zones given a specific program and the isolation between target and interfering programs for a specific zone. Therefore, we define two separate isolation metrics, which are complementary for evaluating the overall system performance.

3.1 Inter-zone isolation (IZI)

To evaluate the isolation performance between the previously defined zones Z_A and Z_B , we define the IZI metric as the ratio of the averaged acoustic power spectra in the two zones. Since the input channels in an audio program are not necessarily correlated (for example, in a stereo program, both correlated and uncorrelated components generally exist),

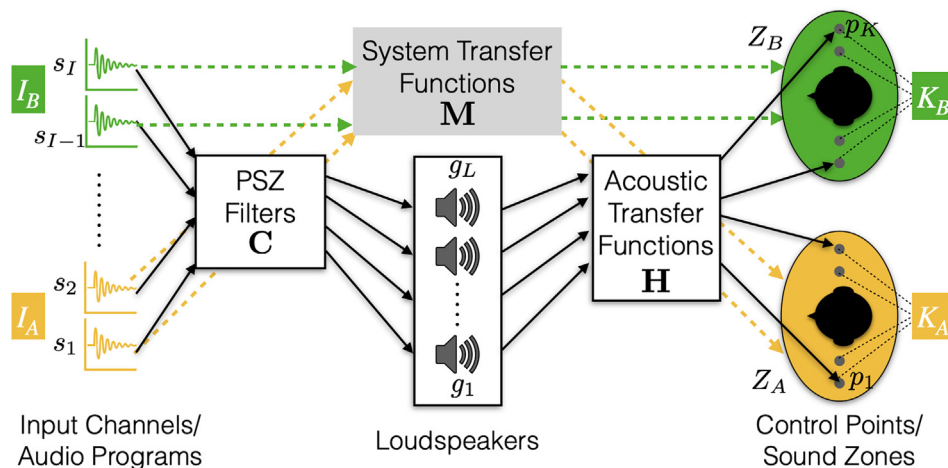


Fig. 1. Signal flow diagram of a general PSZ system consisting of I input audio programs (represented by virtual sources), L loudspeakers, and K control points in two sound zones, K_A and K_B . The other symbols are defined in the text above.

two terms are considered, which represent two extreme cases where all channels are correlated or none of them are correlated. IZI is then determined by the minimum of the two terms.

The following equations show the two terms and the final IZI when Z_A is considered as BZ:

$$IZI_A^{corr} = \frac{\sum_{k \in K_A} \left| \sum_{i \in I_A} M_{ki} \right|^2 / |K_A|}{\sum_{k \in K_B} \left| \sum_{i \in I_A} M_{ki} \right|^2 / |K_B|}, \tag{5}$$

$$IZI_A^{uncorr} = \frac{\sum_{k \in K_A} \sum_{i \in I_A} |M_{ki}|^2 / |K_A|}{\sum_{k \in K_B} \sum_{i \in I_A} |M_{ki}|^2 / |K_B|}, \tag{6}$$

$$IZI_A = \min \{ IZI_A^{corr}, IZI_A^{uncorr} \}, \tag{7}$$

where $|K_A|$ and $|K_B|$ denote the number of control points in each set. The definitions imply an even spacing of the defined control points, as is commonly adopted in the PSZ literature. In the case of Z_B as BZ, the sets K_A and K_B are interchanged, and I_A is replaced by I_B .

It is worth noting that in the simplest case where only *single-channel* programs are considered, IZI is equivalent to AC. Assuming the input signal to be the Dirac delta, which in the frequency domain is represented as a constant, the loudspeaker gains are given by

$$\mathbf{q} = \mathbf{C}\mathbf{s} = \mathbf{c}_I s_1 = \mathbf{c}_I, \tag{8}$$

and as there is no distinction between the correlated and uncorrelated cases, IZI can be expressed as

$$IZI_A = \frac{\sum_{k \in K_A} |M_{k1}|^2 / |K_A|}{\sum_{k \in K_B} |M_{k1}|^2 / |K_B|} = \frac{|\mathbf{H}_A \mathbf{q}|^2 / |K_A|}{|\mathbf{H}_B \mathbf{q}|^2 / |K_B|}, \tag{9}$$

where \mathbf{H}_A and \mathbf{H}_B are the sub-matrices of \mathbf{H} corresponding to the two zones. The latter form is equivalent to the AC definition.¹⁶

3.2 Inter-program isolation (IPI)

We define the IPI metric as the ratio of the two averaged acoustic power spectra, in the same zone, corresponding to the two different audio programs. IPI therefore quantifies the isolation of the target program from the other program in a particular zone. As for the case of IZI, we compute both the correlated and uncorrelated components of IPI and adopt their minimum. Referring to the previously defined system, the IPI for Z_A is expressed as

$$IPI_A^{corr} = \frac{\sum_{k \in K_A} \left| \sum_{i \in I_A} M_{ki} \right|^2 / |I_A|}{\sum_{k \in K_A} \left| \sum_{i \in I_B} M_{ki} \right|^2 / |I_B|}, \tag{10}$$

$$IPI_A^{uncorr} = \frac{\sum_{k \in K_A} \sum_{i \in I_A} |M_{ki}|^2 / |I_A|}{\sum_{k \in K_A} \sum_{i \in I_B} |M_{ki}|^2 / |I_B|}, \tag{11}$$

$$IPI_A = \min \{ IPI_A^{corr}, IPI_A^{uncorr} \}, \tag{12}$$

where $|I_A|$ and $|I_B|$ denote the number of input channels in the audio program for each zone. For the IPI associated with Z_B , I_A and I_B are interchanged, and K_A is replaced by K_B .

As a special case, IPI can also be used to evaluate the isolation performance at a single control point by choosing a particular point k in the set of K_A (or K_B). We show in Sec. 4 that this single-point IPI is by itself useful in evaluating the robustness property of a generated PSZ.

4. Applications using the pressure matching method

In this section, we show two example use cases for the complementary metrics, IZI and IPI, using sound pressure level calculated from the free-field numerical simulation of a PSZ system. These examples are not able to be evaluated with the AC metric due to its limited definition. In the first example (Sec. 4.3), IZI and IPI are used to evaluate the effects of

different rendering modes on the isolation performance of the same PSZ system; in the second example (Sec. 4.4), the single-point IPI is calculated to determine the effective physical boundaries of sound zones in a two-listener PSZ system. All simulated PSZ filters in the examples are designed using the standard pressure matching (PM) method.¹⁹ As the PM method is usually described by vectors of loudspeaker gains and pressure at control points in most of the literature, we first re-express the method in terms of filter and transfer function matrices for the case of multichannel programs.

4.1 PSZ filter generation

The general idea of the PM method¹⁹ is to minimize the difference between the specified target sound field with ideal zone/program isolation and the actual sound field generated by the loudspeakers. The target sound field is usually specified with target pressure \mathbf{p}_T at the control points, where in DZ it is set to zero, and in BZ it is set based on program-specific transfer functions. The cost function to be minimized is constructed as

$$J = \|\mathbf{p} - \mathbf{p}_T\|^2 + \beta\|\mathbf{g}\|^2 = \|\mathbf{H}\mathbf{g} - \mathbf{p}_T\|^2 + \beta\|\mathbf{g}\|^2, \tag{13}$$

where the latter term is introduced as Tikhonov regularization to improve both matrix conditioning and the robustness of the resulting PSZ filters. Taking a single input channel to be the Dirac delta, the above function can be rewritten by replacing \mathbf{g} and \mathbf{p}_T with filter \mathbf{c} and target transfer function \mathbf{m}_T , respectively,

$$J = \|\mathbf{H}\mathbf{c} - \mathbf{m}_T\|^2 + \beta\|\mathbf{c}\|^2. \tag{14}$$

The corresponding optimal solution \mathbf{c}^* is given by

$$\mathbf{c}^* = (\mathbf{H}^H\mathbf{H} + \beta\mathbf{I})^{-1}\mathbf{H}^H\mathbf{m}_T, \tag{15}$$

where $(\cdot)^H$ denotes taking the conjugate transpose. Due to added regularization, the solution applies to all three cases of $K < L$, $K = L$, and $K > L$. This solution is suitable only when a single-channel program is considered and the corresponding BZ has been specified. For the case of multichannel programs, multiple targets and filters are required. Therefore, assuming I channels, we modify the cost function to be the sum of the costs from minimizing the errors in all channels,

$$J = \sum_{i=1}^I J_i = \sum_{i=1}^I \|\mathbf{H}\mathbf{c}_i - \mathbf{m}_{T,i}\|^2 + \sum_{i=1}^I \beta\|\mathbf{c}_i\|^2 = \|\mathbf{H}\mathbf{C} - \mathbf{M}_T\|_F^2 + \beta\|\mathbf{C}\|_F^2, \tag{16}$$

where $\mathbf{C} = [\mathbf{c}_1 \cdots \mathbf{c}_I]$, $\mathbf{M}_T = [\mathbf{m}_{T,1} \cdots \mathbf{m}_{T,I}]$ are the filter and system transfer function matrices, respectively, and the subscript F denotes the Frobenius norm. The resulting optimal filter matrix \mathbf{C}^* is given by

$$\mathbf{C}^* = (\mathbf{H}^H\mathbf{H} + \beta\mathbf{I})^{-1}\mathbf{H}^H\mathbf{M}_T, \tag{17}$$

which has a similar form as Eq. (15) for a single program except that the target vector is replaced by the target matrix. Similar solutions can also be found in the literature on cross talk cancellation systems (such as Bai and Lee²⁰).

4.2 Simulation setup

The PSZ system adopted for the free-field simulation consists of a linear array of eight loudspeakers and two zones, as illustrated in Fig. 2. The loudspeakers are modeled as circular baffled pistons in the far field, with a spacing of 25 cm

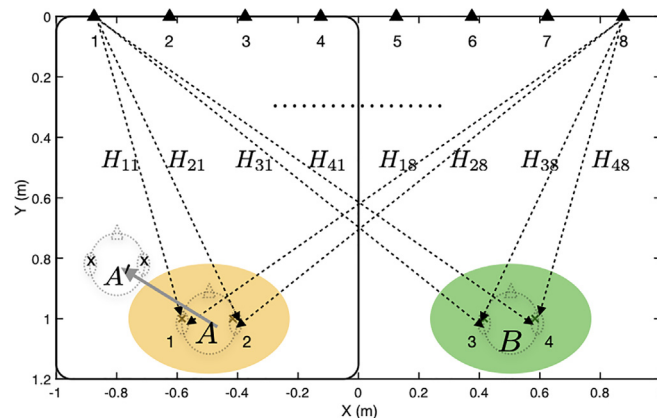


Fig. 2. Illustration of the simulated PSZ system. The black triangles on the top and the cross-markers at the ear positions of two listeners represent the loudspeakers and control points, respectively, and the sound zones are indicated by two oval shadows. The dashed lines illustrate the indexed loudspeaker and control point for the corresponding transfer function. A and B represent the two centered listeners, and A' represents the moved left listener. The left half of the sound field is used to show the results in Fig. 4.

between two adjacent units. The two zones are separated by 1 m and are also 1 m from the array. Two control points are defined in each zone, representing the ear locations of a listener with a spacing of 16.8 cm (the listeners' heads are not included in the simulation). In addition to the case of both listener *A* and *B* located in the zone center, a moved listener (illustrated as *A'* in the figure) with a displacement of $(x, y) = (-0.3, -0.2)m$ is also simulated. Three cases of target matrices are considered, which correspond to three rendering modes for *mono* programs, *stereo* programs without cross talk cancellation (XTC), and *binaural* programs with XTC (equivalent to a multi-listener transaural system²¹), respectively, and are given by

$$M_{mono} = \begin{bmatrix} \frac{H_{11} + H_{14}}{2} & 0 \\ \frac{H_{21} + H_{24}}{2} & 0 \\ 0 & \frac{H_{35} + H_{38}}{2} \\ 0 & \frac{H_{45} + H_{48}}{2} \end{bmatrix}, \quad M_{stereo} = \begin{bmatrix} H_{11} & H_{14} & 0 & 0 \\ H_{21} & H_{24} & 0 & 0 \\ 0 & 0 & H_{35} & H_{38} \\ 0 & 0 & H_{45} & H_{48} \end{bmatrix}, \quad M_{XTC} = \begin{bmatrix} H_{11} & 0 & 0 & 0 \\ 0 & H_{24} & 0 & 0 \\ 0 & 0 & H_{35} & 0 \\ 0 & 0 & 0 & H_{48} \end{bmatrix}, \quad (18)$$

where the transfer functions of loudspeakers 1, 4, 5, and 8 (denoted in Fig. 2) are chosen as those between the input channels and the control points. To simulate realistic uncertainties in the transfer functions due to factors such as loudspeaker position inaccuracies or response/gain variances, the transfer functions are sampled from a complex Gaussian distribution for each transfer function H_{kl} modeled as

$$H_{kl} = A_{kl}e^{i\phi_{kl}}, \quad A_{kl} \sim N(\hat{A}_{kl}, \sigma_A^2), \quad \phi_{kl} \sim N(\hat{\phi}_{kl}, \sigma_\phi^2), \quad (19)$$

where A, ϕ denote the amplitude and phase of the transfer function, $N(\cdot, \cdot)$ denotes the normal distribution, the hat symbol denotes the value obtained from the free-field simulation, and $\sigma_A^2, \sigma_\phi^2$ denote the variances. In the simulation, we simply choose $\sigma_A^2 = \sigma_\phi^2 = \sigma^2 = 10^{-4}$ at all frequencies and set the constant regularization as $\beta = K \cdot \sigma^2 = 4 \times 10^{-4}$ to minimize the expected cost in Eq. (14), following a probabilistic approach.²² From the same specified distribution $N(0, \sigma^2)$, we sample two sets of transfer functions independently for filter generation and performance evaluation, and each set is averaged across ten trials to represent the procedure in actual experiments. The corresponding PSZ filters are computed using Eq. (17).

4.3 Evaluating PSZ performance with different rendering modes

In the literature, the performance of PSZ systems is usually evaluated with a single set of target pressure,^{2,6,8,15} which corresponds to a fixed rendering mode for mono programs in each zone. While most existing systems are capable of delivering multichannel programs, the impact of their corresponding rendering modes on the isolation performance has not been studied. By using the new IZI and IPI metrics, such potential impact can be explicitly evaluated for various choices of target matrices.

We simulate and compare the IZI and IPI performance of the system setup described above for three target matrices given by Eq. (18) and two cases of listener positions, (1) two listeners centered in both zones and (2) one listener moving away from the zone, with the results for IZI_A and IPI_A shown in Fig. 3. Figure 3(a) corresponds to case 1, where IZI_A and IPI_A are almost identical due to the symmetric setup and free-field assumptions. The *mono* mode yields the best performance, followed by the *stereo* mode, while the *XTC* mode has the worst isolation. In particular, the degradation in the *XTC* mode is more significant below 1 kHz, indicating a potential trade-off between program isolation and cross talk cancellation, due to the increased wavelength (and therefore weaker isolation between listeners) and the requirement in

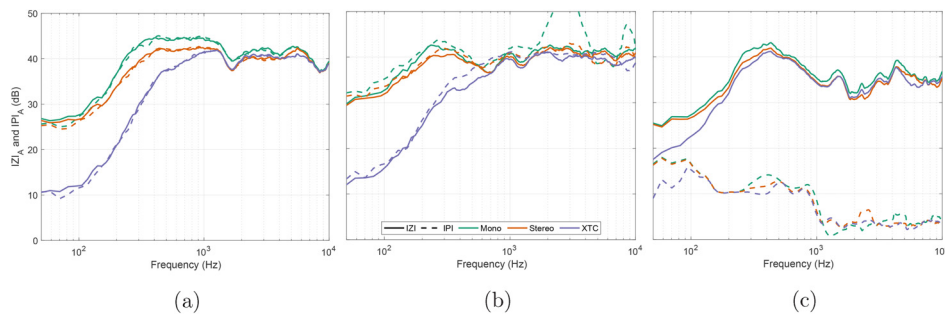


Fig. 3. Simulated IZI_A (solid lines) and IPI_A (dashed lines) for two evaluation setups of different listener positions: $\{A, B\}$ (a) and $\{A', B\}$ [(b) and (c)], as illustrated in Fig. 2. The three colors represent three different program specifications: mono, stereo, and XTC programs. One-third-octave logarithmic smoothing is applied to the spectra for better visualization.

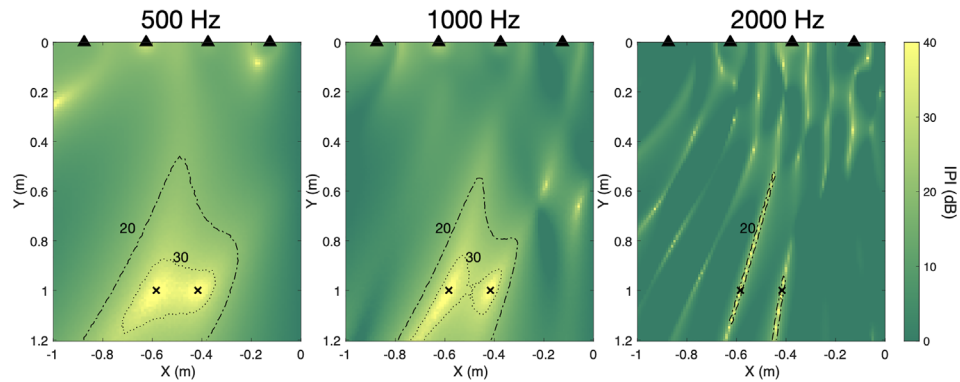


Fig. 4. Computed spatial maps of single-point IPI (truncated to 40 dB) for the left half of the specified sound field at frequencies of 0.5 (left), 1 (center), and 2 kHz (right). Two contour lines corresponding to two IPI levels are shown in each plot with different markers [20 dB (dashed-dotted line) and 30 dB (dotted line)].

the cost function for cancelling the cross talk. For PSZ systems rendering binaural audio with XTC, such trade-off can be further optimized based on established perceptual preferences of listeners,¹⁷ or one can simply downmix the audio as mono to trade-off spatial quality for better isolation. Figures 3(b) and 3(c) correspond to case 2, with the PSZ filters designed for the new and the centered position, respectively. We start to observe differences between the two metrics in Fig. 3(b) due to breaking of the symmetry; in Fig. 3(c), the differences become clearer as the PSZ filters are no longer optimal. IZI_A shows a minor degradation in isolation between two listeners, whereas IPI_A reflects severely degraded isolation between two programs for the moved listener, which is more indicative of the actual experience of the listener with unoptimized PSZ filters.

4.4 Determining effective PSZ boundaries

In most PSZ systems, sound zones are specified with regular geometries (e.g., round^{7,16} or square^{4,10,11,15} shapes). However, due to the constraints of practical systems, such as the number and distribution of loudspeakers and control points, the isolation performance within the sound zone is often non-uniform, leading to a certain lack of robustness against possible listener movements within zones. To quantify such robustness, the effective “boundaries” of sound zones are defined by using the previously defined single-point IPI metric as the contour line for a certain IPI level (e.g., 20 dB). As a result, the robustness against listener movements can be evaluated by the area (or volume) within the boundaries. Furthermore, the dependency of that robustness on moving directions can be evaluated with the projection of the effective area/volume along the direction.

Figure 4 shows three computed spatial maps of simulated single-point IPI in two dimensions (2D) for the left half of the sound field (indicated by the thick solid line in Fig. 2) at frequencies 0.5, 1, and 2 kHz, for the rendering mode of a target *mono* program for the left listener. In each map plot, two contour lines are shown with different line types, with the outer and inner lines corresponding to 20 and 30 dB of IPI, respectively. It is clear that given the boundaries defined by those contour lines, the shapes of the sound zones are irregular. As PSZ filters are determined by both the exact transfer functions and the target matrices [see Eq. (17)], we also expect a great variability of the resulting PSZ boundaries with the choice of system configurations and rendering modes. Comparing the three map plots, a trend in the decrease in the effective size of sound zones is observed as frequency increases, which has been well recognized in the literature^{1,7,15} as reduced robustness but was difficult to study using the AC metric. It is worth noting that the simulation is aimed to only illustrate the trend, which is mainly due to the wavelength changes and can be observed with and without head presence.

5. Conclusion and final discussion

This paper introduces two metrics, IZI and IPI, for evaluating the isolation performance of generalized PSZ systems from two different aspects. The IZI metric, which reduces to the commonly used AC metric for the special case of rendering single-channel programs, represents the isolation of the sound zones for a (single-channel or multichannel) program, whereas the IPI metric quantifies the level of isolation of the target program from the interfering program in the same sound zone.

The two metrics, albeit defined by similar expressions, are generally non-interchangeable except for special cases where both the physical system setup and the program assignment are perfectly symmetric with respect to the two listeners. In addition, the different emphases of the two metrics make them suitable for different situations: IZI compares the acoustic energy between two sound zones and, therefore, is more suitable in the cases where a high contrast of sound energy between different regions is desired, such as creating a dark zone in which all audio programs are attenuated; and IPI is related to the acoustic energy of different programs rendered at the same zone and, therefore, is more applicable

when different programs are present concurrently and also more suitable for the objective evaluation of the audio-on-audio interference.

In Sec. 4, we present two examples of different applications of the IZI and IPI metrics, with implications for future work. In the first example, we show the potential trade-off between the isolation performance and cross talk cancellation at low frequencies for PSZ systems with cross talk-cancelled binaural content. This offers the potential to further improve the filter design method by optimizing the trade-off in accordance with subjective preferences. In the second example, we show that the single-point IPI metric can serve as a basis for evaluating the robustness of the generated sound zones against listener movements. This allows the definition of a new metric that specifically quantifies the sound zone robustness, which is beyond the scope of this work and will be presented in a future study.

Acknowledgments

The authors wish to thank R. Sridhar and J. Tylka for their foundational contributions to this work. This work was supported by a research grant from FOCAL-JMLab.

References and links

- ¹W. Druyvesteyn and J. Garas, "Personal sound," *J. Audio Eng. Soc.* **45**(9), 685–701 (1997).
- ²F. Olivieri, M. Shin, F. M. Fazi, P. A. Nelson, and P. Otto, "Loudspeaker array processing for multi-zone audio reproduction based on analytical and measured electroacoustical transfer functions," in *Proceedings of the Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*, Guildford, UK (September 2–4, 2013).
- ³T. Okamoto and A. Sakaguchi, "Experimental validation of spatial Fourier transform-based multiple sound zone generation with a linear loudspeaker array," *J. Acoust. Soc. Am.* **141**(3), 1769–1780 (2017).
- ⁴X. Ma, P. J. Hegarty, and J. J. Larsen, "Mitigation of nonlinear distortion in sound zone control by constraining individual loudspeaker driver amplitudes," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada (April 15–20, 2018), pp. 456–460.
- ⁵M. F. S. Gálvez, D. Menzies, and F. M. Fazi, "Dynamic audio reproduction with linear loudspeaker arrays," *J. Audio Eng. Soc.* **67**(4), 190–200 (2019).
- ⁶J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *J. Acoust. Soc. Am.* **131**(6), 4518–4525 (2012).
- ⁷P. Coleman, P. J. Jackson, M. Olik, M. Møller, M. Olsen, and J. Abildgaard Pedersen, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *J. Acoust. Soc. Am.* **135**(4), 1929–1940 (2014).
- ⁸F. Olivieri, F. M. Fazi, S. Fontana, D. Menzies, and P. A. Nelson, "Generation of private sound with a circular loudspeaker array and the weighted pressure matching method," *IEEE/ACM Trans. Audio Speech Lang. Process.* **25**(8), 1579–1591 (2017).
- ⁹K. Imaizumi, K. Niwa, and K. Tsutsumi, "Loudspeaker array to maximize acoustic contrast using proximal splitting method," in *Proceedings of the 2021 29th European Signal Processing Conference (EUSIPCO)*, Dublin, Ireland (August 23–27, 2021), pp. 91–95.
- ¹⁰K. Baykaner, P. Coleman, R. Mason, P. J. Jackson, J. Francombe, M. Olik, and S. Bech, "The relationship between target quality and interference in sound zone," *J. Audio Eng. Soc.* **63**(1/2), 78–89 (2015).
- ¹¹J. Rämö, S. Bech, and S. H. Jensen, "Validating a real-time perceptual model predicting distraction caused by audio-on-audio interference," *J. Acoust. Soc. Am.* **144**(1), 153–163 (2018).
- ¹²Q. Zhu, P. Coleman, M. Wu, and J. Yang, "Robust acoustic contrast control with reduced *in-situ* measurement by acoustic modeling," *J. Audio Eng. Soc.* **65**(6), 460–473 (2017).
- ¹³Q. Zhu, P. Coleman, X. Qiu, M. Wu, J. Yang, and I. Burnett, "Robust personal audio geometry optimization in the SVD-based modal domain," *IEEE/ACM Trans. Audio Speech Lang. Process.* **27**(3), 610–620 (2019).
- ¹⁴S. J. Elliott and M. Jones, "An active headrest for personal audio," *J. Acoust. Soc. Am.* **119**(5), 2702–2709 (2006).
- ¹⁵L. Vindrola, M. Melon, J.-C. Chamard, and B. Gazengel, "Use of the filtered-x least-mean-squares algorithm to adapt personal sound zones in a car cabin," *J. Acoust. Soc. Am.* **150**(3), 1779–1793 (2021).
- ¹⁶J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.* **111**(4), 1695–1700 (2002).
- ¹⁷N. Canter and P. Coleman, "Delivering personalised 3D audio to multiple listeners: Determining the perceptual trade-off between acoustic contrast and cross-talk," in *Proceedings of Audio Engineering Society Convention 150* (May 25–28, 2021).
- ¹⁸J. Francombe, R. Mason, M. Dewhirst, and S. Bech, "A model of distraction in an audio-on-audio interference situation with music program material," *J. Audio Eng. Soc.* **63**(1/2), 63–77 (2015).
- ¹⁹M. Poletti, "An investigation of 2-D multizone surround sound systems," in *Proceedings of Audio Engineering Society Convention 125*, San Francisco, CA (October 2–5, 2008).
- ²⁰M. R. Bai and C.-C. Lee, "Objective and subjective analysis of effects of listening angle on crosstalk cancellation in spatial sound reproduction," *J. Acoust. Soc. Am.* **120**(4), 1976–1989 (2006).
- ²¹C. House, S. Dennison, D. Morgan, N. Rushton, G. White, J. Cheer, and S. Elliott, "Personal spatial audio in cars: Development of a loudspeaker array for multi-listener transaural reproduction in a vehicle," in *Proceedings of Reproduced Sound 2017: Sound Quality by Design*, Nottingham, UK (November 21–23, 2017).
- ²²M. B. Møller, J. K. Nielsen, E. Fernandez-Grande, and S. K. Olesen, "On the influence of transfer function noise on sound zone control in a room," *IEEE/ACM Trans. Audio Speech Lang. Process.* **27**(9), 1405–1418 (2019).