ROBUST PRESSURE MATCHING WITH ATF PERTURBATION CONSTRAINTS FOR SOUND FIELD CONTROL

Junqing Zhang[1,2]*, Liming Shi*[2]*, Mads G. Christensen*[2]*, Wen Zhang*[1]*, Lijun Zhang*[1]*, and Jingdong Chen*[1]

[1]Center of Intelligent Acoustics and Immersive Communications, School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China
[2]Audio Analysis Lab, CREATE, Aalborg University, 9000 Aalborg, Denmark

ABSTRACT

Sound field control systems deployed in room acoustic environments require knowing the acoustic channel impulse responses between the loudspeakers and matching microphones, which are challenging to estimate accurately due to perturbations caused by such factors as temperature changes and sensors' position mismatches. To deal with this issue, a robust pressure matching algorithm is developed in this work where a perturbation term of the acoustic transfer function (ATF) is modeled as a Gaussian process, based on which an uncertainty constraint is applied to limit the impact of perturbation on pressure matching. This constrained problem is formulated as one of biconvex optimization, and a coordinate descent algorithm is adopted to estimate the optimal control filter. Simulations are performed and results show that the proposed method is able to achieve more accurate control as compared to the standard pressure matching algorithm in the presence of ATF perturbations.

Index Terms— sound field control, robustness, pressure matching, biconvex optimization, personal sound zones

1. INTRODUCTION

Sound field control aims to physically synthesize the desired sound field over a target region using a loudspeaker array [1], which can be used in a wide range of audio applications, such as sound field reproduction, personal sound zone generation [2], and spatial active noise control [3]. Representative methods for sound field control include the pressure matching (PM) method [4, 5], mode matching method [6, 7], amplitude matching method [8], subspace-based method [9], and deep learning-based method [10].

The performance of the sound field control system relies on the accurate knowledge of the acoustic transfer functions (ATFs) between the loudspeakers and matching microphones [11]. Commonly, the ATFs are obtained by pre-calibration or mathematical modeling [12]. Unfortunately, ATF perturbations are inevitable in practical applications due to such physical and environmental factors as sensor position mismatches, transducers' electro-acoustic response variations [13], and the sound speed variation due to room temperature variation [14]). These perturbations can lead to significant performance degradation or even control failure (e.g., failure of the PM method [15]). It was shown that the temperature-induced perturbations may cause more severe impact since the length of the acoustic impulse responses changes especially at higher frequencies and for longer propagation distances as the speed of sound changes [14].

Several works have been proposed to improve the robustness of sound field control systems against the ATF perturbations, which can be broadly divided into two categories. The first one is to add a regularization term with a constraint on the loudspeaker weights energy [16]. The regularization term is further extended to a general form to improve the robustness of the array against the changes of the acoustic environment and driver responses [17]. The results reported in [15] showed that the performance of the PM method can be significantly improved by using a well-selected regularization parameter. Recently, a geometry optimization method based on singular value decomposition of the ATF matrix has been proposed to further improve the robustness of the regularization parameter [18]. The second category is based on the *a priori* information of ATF perturbations. With the assumption that the marginal distribution of the ATF perturbations is accessible, a probability model-based optimization approach is applied to improve the robustness in personal audio reproduction [19, 20].

In this work, a robust pressure matching algorithm with ATF constraints is presented. In order to improve the system robustness against ATF perturbations, an ellipsoidal constraint of the ATF estimates is added to the conventional PM method. The biconvex optimization is then used to solve the reformulated optimization problem, and the coordinate descent algorithm is adopted to obtain the optimal ATF and loudspeaker control filter estimates in an iterative manner.

2. PROBLEM STATEMENT

With the discrete-Fourier-transform (DFT) domain approach for sound field control [21], the sound pressure produced at the mth control point (the matching microphone position) can be represented by the ATF $g_{ml}(k)$ between the lth loudspeaker and mth control point, and the loudspeaker control filter $d_l(k)$[1], i.e.,

$$p_m(k) = \sum_{l=1}^{L} g_{ml}(k) d_l(k), m = 1, ..., M, \tag{1}$$

where $k = 2\pi f/c$ is the wavenumber with f being the frequency, c is the speed of sound in air, and L and M denote, respectively, the number of loudspeakers and the number of control microphones within the control region. In a vector/matrix form, (1) can be rewritten as

$$p_m(k) = \mathbf{g}_m^T(k)\mathbf{d}(k), \tag{2}$$

e-mail: junqingzhang@mail.nwpu.edu.cn

This work was supported in part by the National Key Research and Development Program of China under Grant No. 2018AAA0102200 and in part by the Key Program of National Science Foundation of China (NSFC) under Grant No. 61831019. This work was also supported by China Scholarship Council.

[1]Note that the input signal to each of the filters is assumed to be unity (e.g., sinusoidal wave with constant amplitude), and it is discarded.

where the superscripts ^{T} denotes the transpose operator, $\mathbf{d}(k) = [d_1(k)\ d_2(k)\ \cdots\ d_L(k)]^T$, and $\mathbf{g}_m(k) = [g_{m1}(k)\ g_{m2}(k)\ \cdots\ g_{mL}(k)]^T$.

Putting the sound pressure at all the M control points into a vector form, we obtain the following system:

$$\mathbf{p}(k) = \mathbf{G}(k)\mathbf{d}(k), \tag{3}$$

where $\mathbf{p}(k) = [p_1(k)\ p_2(k)\ \cdots\ p_M(k)]^T$ denotes the reproduced sound pressure vector, and $\mathbf{G}(k) = [\mathbf{g}_1(k)\ \mathbf{g}_2(k)\ \cdots\ \mathbf{g}_M(k)]^T$ is the matrix consisting of all the ATF vectors. To simplify the notation, the index k will be omitted in the rest of this paper.

In sound field control systems, the accurate knowledge of the ATFs between the loudspeakers and the matching microphones is important for achieving the desired performance. Unfortunately, such accurate information is not available in practical applications as explained in the introduction and ATF perturbations due to reasons such as sensor and source position mismatches are inevitable in practice, which often results in significant performance degeneration in sound field control [20]. In order to take into account the ATF perturbations, we model the measured ATF matrix as

$$\widehat{\mathbf{G}} = \mathbf{G}_{\mathrm{o}} + \Delta\mathbf{G}, \tag{4}$$

where $\widehat{\mathbf{G}}$, $\mathbf{G}_{\mathrm{o}}$, and $\Delta\mathbf{G}$ are the measured, true, and perturbation ATFs matrices, as defined in [17].

3. PROPOSED ROBUST CONTROL ALGORITHM

3.1. Pressure matching with ATF constraint

Let us first review the PM method in [4], which can be formulated as one of minimizing the mean square error (MSE) between the desired sound field $\mathbf{p}^{\mathrm{d}}$ and reproduced sound field within the control region subject to the loudspeaker array energy constraint. However, the traditional PM method is sensitive to changes and variations in ATFs. To circumvent this issue, we present in this work a robust control method against ATF perturbation by adding an ellipsoidal constraint with uncertainty on the ATFs. With the assumption that the perturbations, in terms of ATF variations, for each loudspeaker/microphone pair are independent, we vectorize the ATF matrix into a column vector. Thus, the general form of the optimization problem can be represented as

$$\min_{\mathbf{G},\mathbf{d}} \ \|\mathbf{G}\mathbf{d} - \mathbf{p}^{\mathrm{d}}\|^2 \tag{5}$$

$$\text{s. t. } \ \|\mathbf{d}\|^2 \le e_{\mathrm{h}}, \tag{5a}$$

$$\mathrm{vec}(\mathbf{G}) \in \mathcal{E}\left[\mathrm{vec}(\widehat{\mathbf{G}}), \mathbf{C}^{-1}\right], \tag{5b}$$

where $\mathrm{vec}(\cdot)$ denotes the vectorization operator[2]. The sizes of the vector $\mathrm{vec}(\widehat{\mathbf{G}})$ and matrix $\mathbf{C}$ are $(M \times L) \times 1$ and $(M \times L) \times (M \times L)$, respectively. We assume that $\mathrm{vec}(\mathbf{G})$ belongs to an uncertainty ellipsoid $\mathcal{E}[\mathrm{vec}(\widehat{\mathbf{G}}), \mathbf{C}^{-1}]$ [22]. This ellipsoidal constraint is used to model the possible range of uncertainty, which can be parameterized by the center $\mathrm{vec}(\widehat{\mathbf{G}})$ and the symmetric positive semi-definite configuration matrix $\mathbf{C}$ [23]. When the matrix $\mathbf{C}$ is invertible, the uncertainty set can be represented as

$$\mathcal{E}\left[\mathrm{vec}(\widehat{\mathbf{G}}), \mathbf{C}^{-1}\right] = \left\{\mathrm{vec}(\mathbf{G}) | \left[\mathrm{vec}(\mathbf{G} - \widehat{\mathbf{G}})\right]^H \mathbf{C}^{-1} \left[\mathrm{vec}(\mathbf{G} - \widehat{\mathbf{G}})\right] \le \epsilon\right\}, \tag{6}$$

[2]The operator $\mathrm{vec}(\cdot)$ stacks the columns vectors of the matrix into a long column vector.

where the superscripts ^{H} denotes the conjugate-transpose operator, the non-negative parameter ϵ is used to describe the uncertainty about $\mathrm{vec}(\mathbf{G})$ and limits the size of the uncertainty set. The ellipsoidal parameters within matrix $\mathbf{C}$ can be obtained by using the additional *a priori* information about the ATF, i.e., the range or the marginal distribution of the perturbations [19, 24]. Note that (6) degenerates to a sphere constraint when $\mathbf{C}$ is the identity matrix.

The proposed method aims to find an optimal solution for (5) within the uncertainty set $\mathcal{E}$. The value of ϵ should be large enough, otherwise, the solution may not be included in the uncertainty set. This means that the estimated $\mathbf{G}$ can be close to the $\widehat{\mathbf{G}}$ but far away from $\mathbf{G}_{\mathrm{o}}$ when severe perturbation occurs. Therefore, a well-selected value of ϵ can provide a proper feasible set of the problem, which is beneficial for searching for a better-estimated $\mathbf{G}$[3].

3.2. Biconvex optimization for robust pressure matching

The problem given in (5) includes two parts: (i) calculating the control filter $\mathbf{d}$ and (ii) computing the $\mathbf{G}$ estimate, which is a biconvex optimization problem [25]. As the coordinate descent algorithm does, the most common way of solving such a problem is alternatively updating two variables by fixing one of them and solving the corresponding convex optimization problem.

3.2.1. Step1: computing the control filter $\mathbf{d}$

With the $\mathbf{G}$ estimate obtained from the previous step, the optimization problem (5) is simplified as

$$\min_{\mathbf{d}} \ \|\mathbf{G}\mathbf{d} - \mathbf{p}^{\mathrm{d}}\|^2 \qquad \text{s. t. } \ \|\mathbf{d}\|^2 \le e_{\mathrm{h}}. \tag{7}$$

Using a Lagrange multiplier, one we can rewrite (7) as a Lagrangian cost function, i.e.,

$$\mathcal{L}(\mathbf{d}) = \|\mathbf{G}\mathbf{d} - \mathbf{p}^{\mathrm{d}}\|^2 + \xi \|\mathbf{d}\|^2, \tag{8}$$

where ξ is the positive Lagrange multiplier. It follows then that the interior point method [26], i.e., the primal-dual method, can be used to obtain the solution.

3.2.2. Step2: Computing the $\mathbf{G}$ *estimate*

Similarly, the original optimization problem (5) degenerates to a concise form when the control filter $\mathbf{d}$ is fixed, i.e.,

$$\min_{\mathbf{G}} \ \|\mathbf{G}\mathbf{d} - \mathbf{p}^{\mathrm{d}}\|^2 \tag{9}$$

$$\text{s. t. } \ \left[\mathrm{vec}(\mathbf{G} - \widehat{\mathbf{G}})\right]^H \mathbf{C}^{-1} \left[\mathrm{vec}(\mathbf{G} - \widehat{\mathbf{G}})\right] \le \epsilon, \tag{9a}$$

which can be equivalently written as

$$\min_{\mathbf{G}} \ \|\mathbf{G}\mathbf{d} - \mathbf{p}^{\mathrm{d}}\| \tag{10}$$

$$\text{s. t. } \ \left\|\mathbf{C}^{-1/2} \left[\mathrm{vec}(\mathbf{G} - \widehat{\mathbf{G}})\right]\right\| \le \sqrt{\epsilon}, \tag{10a}$$

where $\mathbf{C}^{1/2}$ is the square root matrix of $\mathbf{C}$ with $\mathbf{C}^{1/2}(\mathbf{C}^{1/2})^H = \mathbf{C}$. The robust constraint in (10a) is a second-order cone constraint. It follows then that the entire problem in (10) can be formulated as one of semidefinite programming, which can be handled conveniently using readily available software packages, e.g., CVX [27]. The entire flow of the proposed algorithm is summarized in Algorithm 1.

[3]In real applications, the parameter ϵ should be chosen according to the actual acoustic environment.

Algorithm 1 Biconvex optimization for robust pressure matching

Input: e_{h}, ϵ, $\widehat{\mathbf{G}}$, $\mathbf{C}$, $\mathbf{p}^{\mathrm{d}}$
for step $u = 1, 2, \ldots, U$ **do**
if $u = 1$
· Calculate $\mathbf{d}(1)$ using $\widehat{\mathbf{G}}$ by solving (8)
· Calculate $\mathbf{G}(1)$ estimate using $\mathbf{d}(1)$ by solving (10)
else
· Calculate new optimal control filter $\mathbf{d}(u)$ using the $\mathbf{G}(u-1)$ estimate from the previous step by solving (8)
· Calculate new $\mathbf{G}(u)$ estimate using $\mathbf{d}(u-1)$ by solving (10)
end for
until Stop condition satisfied
Output: $\mathbf{d}$ and $\mathbf{G}$

3.3. Robust pressure matching for acoustic bright and dark zones generation

One of the particular applications of the proposed method is the sound zones generation [2], which aims to generate several independent listening zones within a room using a loudspeaker array. The listeners can enjoy an immersive listening experience in the bright zone, and meanwhile other listeners are not disturbed by the interference sounds generated by the loudspeaker array in the dark zone, in which the sound pressure level is low. In such a case, the formulation of the PM method can be modified as one of minimizing the MSE between the desired sound field $\mathbf{p}_{\mathrm{B}}^{\mathrm{d}}$ and reproduced sound field in the bright zone, with additional constraints on the dark zone energy, the loudspeaker weights energy, and the ATF perturbation terms, i.e.,

$$\min_{\mathbf{G}_{\mathrm{B}},\mathbf{d}} \quad \|\mathbf{G}_{\mathrm{B}}\mathbf{d} - \mathbf{p}_{\mathrm{B}}^{\mathrm{d}}\|^2 \tag{11}$$

$$\text{s. t.} \quad \|\mathbf{G}_{\mathrm{D}}\mathbf{d}\|^2 \leq e_{\mathrm{d}}, \tag{11a}$$

$$\|\mathbf{d}\|^2 \leq e_{\mathrm{h}}, \tag{11b}$$

$$\mathrm{vec}(\mathbf{G}_{\mathrm{B}}) \in \mathcal{E}\left[\mathrm{vec}(\widehat{\mathbf{G}}_{\mathrm{B}}), \mathbf{C}^{-1}\right], \tag{11c}$$

where the subscripts $_{\mathrm{B}}$ and $_{\mathrm{D}}$ denote the bright and dark zones, respectively. The bounds of the energy constraints, i.e., e_{d} and e_{h}, are usually chosen according to the perceptual requirements [28]. In this work, we only consider imposing the uncertainty constraint on the bright zone ATF perturbations for simplification as the focus of the proposed algorithm is on minimization of the reproduction error in the bright zone.

4. EVALUATION

In this section, we verify the effectiveness of the proposed method through simulations and compare it with the conventional PM method. Specifically, the comparison involves the following methods: (i) PM (true) and (ii) PM (measured) correspond to the cases that the PM method is evaluated by the true ATFs and measured ATFs, respectively; (iii) RPM (sphere) and (iv) RPM (ellipsoid) correspond to the case of the proposed robust pressure matching (RPM) method with a sphere constraint and an ellipsoidal constraint, respectively.

4.1. Performance evaluation metrics

The performance evaluation metrics adopted are as follows:
• Normalized Signal Distortion Energy (NSDE) in the bright zone [21]:

$$\varepsilon_{\mathrm{NSDE}}(\mathbf{d}) = 10\log_{10}\frac{\left\|\mathbf{p}_{\mathrm{B}} - \mathbf{p}_{\mathrm{B}}^{\mathrm{d}}\right\|^2}{\left\|\mathbf{p}_{\mathrm{B}}^{\mathrm{d}}\right\|^2}, \tag{12}$$

where $\mathbf{p}_{\mathrm{B}}^{\mathrm{d}}$ denotes the desired sound field in the bright zone.
• Acoustic Contrast (AC): the energy ratio between the bright and dark zones after being controlled, which is defined as

$$\kappa_{\mathrm{AC}}(\mathbf{d}) = 10\log_{10}\left(\frac{M_{\mathrm{D}}\mathbf{d}^H\mathbf{G}_{\mathrm{B}}^H\mathbf{G}_{\mathrm{B}}\mathbf{d}}{M_{\mathrm{B}}\mathbf{d}^H\mathbf{G}_{\mathrm{D}}^H\mathbf{G}_{\mathrm{D}}\mathbf{d}}\right), \tag{13}$$

where M_{B} and M_{D} are the numbers of the control points within the bright zone and dark zone, respectively.

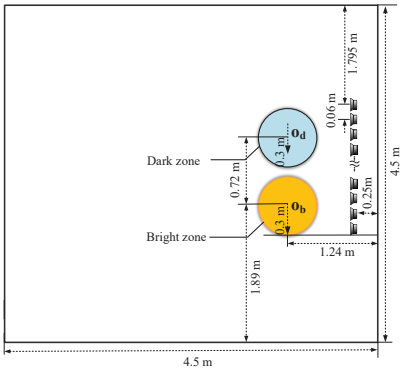

Fig. 1: Plan view of the system geometry within a room.

4.2. Simulation Setup

The evaluations are performed in a simulated room environment using the RIR Generator toolbox [29], and the simulation setup is configured according to [21]. As depicted in Fig. 1, the size of the room is 4.5 m $\times$ 4.5 m $\times$ 2.2 m (length$\times$width$\times$height), and the center of the speaker/microphone array and virtual sources are located in the same plane at the height of 1.2 m. The reverberation time T_{60} is controlled to be approximately 300 ms. The speed of sound is assumed to be $c = 343$ m/s, and the room temperature is assumed $T_0 = 26^{\circ}$C. A 16-element uniform line array with an interelement spacing of $d = 0.06$ m is used for reproduction. The bright and dark zones are in front of the loudspeaker array as illustrated in Fig. 1, and for each zone there are $M = 37$ control points uniformly distributed within the zone. All signals are sampled with a sampling frequency of 16 kHz, and the number of frequency bins is 401. The length of the control filters is set to 800 samples, which corresponds to a frequency grid of 20 Hz. The measured ATFs between the 8th loudspeaker and the matching points in the bright zone (i.e., 8th column of $\widehat{\mathbf{G}}_{\mathrm{B}}$) are used to generate the desired sound field $\mathbf{p}_{\mathrm{B}}^{\mathrm{d}}$. A modeling delay (i.e., 15 ms) is imposed to assure causality [30].

4.3. Simulation results

In this subsection, we show the comparison results of the conventional PM method and the proposed RPM method after 100 iterations. Among different causes of ATF perturbations, we simulate the case of matching position mismatch and investigate the reproduction performance over a wide band of operating frequencies with different levels of position mismatch. Random perturbation within a defined range is added to the positions of the matching microphones, which generates an uncertainty term $\Delta\mathbf{G}_{\mathrm{B}}$ as given in (4). We use the term $\|\Delta\mathbf{G}_{\mathrm{B}}\|_{\mathrm{F}}$, i.e., the Frobenius norm of the matrix $\Delta\mathbf{G}_{\mathrm{B}}$,

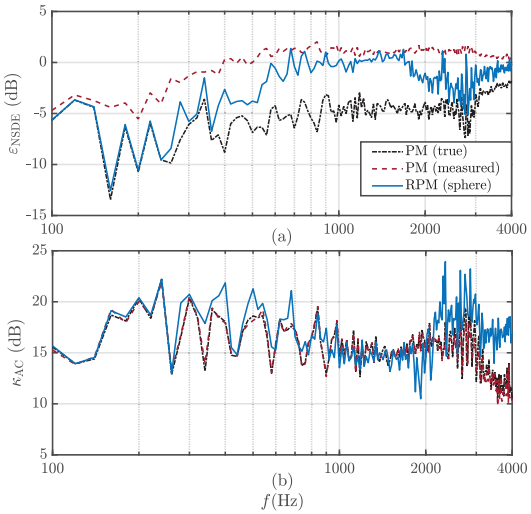

Fig. 2: Evaluation results along with frequency.

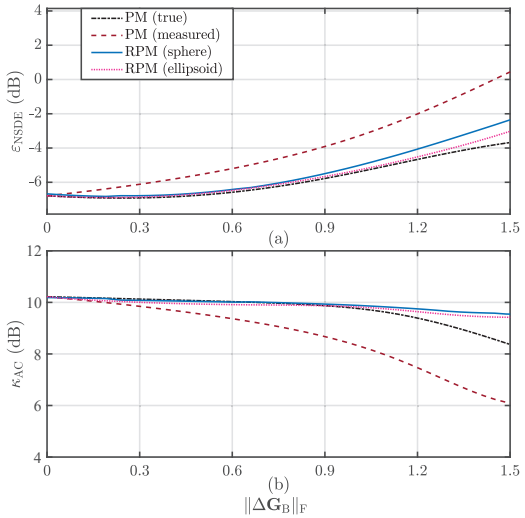

Fig. 3: Evaluation results along with the perturbation level.

to quantify the level of perturbation. The parameter ϵ is chosen as $\epsilon = \|\mathbf{C}^{-1/2}[\mathrm{vec}(\mathbf{G}_{\mathrm{Bo}} - \widehat{\mathbf{G}}_{\mathrm{B}})]\|^2$ [31], and $\mathbf{G}_{\mathrm{Bo}}$ is the true ATF matrix within bright zone. The black, red, blue, and pink lines in Figs. 2 and 3 correspond to the method of PM (true), PM (measured), RPM (sphere), and RPM (ellipsoid), respectively.

4.3.1. Performance versus frequency

Firstly, we evaluate the system performance over a broadband frequency range of $[100, 4000]$ Hz. In order to simulate the sensor position mismatch, we considered adding uniformly distributed random perturbation within the range $[0, 0.2]$ m to the microphone positions in the bright zone at the same height of the microphone array, which corresponds to the perturbation level of $\|\Delta \mathbf{G}_{\mathrm{B}}\|_{\mathrm{F}} = 1.0689$. The simulation results are averaged over 100 Monte-Carlo trials. By referring to (11a) and (11b), the dark zone energy constraint bound e_{d} and array energy constraint bound e_{h} are chosen as -30 dB and 13 dB, respectively. As shown in Fig. 2 (a), the PM with true ATF, i.e., PM (true), has the lowest NSDE. The PM (measured) method, which does not consider the perturbation term, has the largest NSDE as the sensor position mismatch degrades system performance. In contrast, the proposed RPM method has reasonably good NSDE performance over the entire frequency band. Especially, within the low-frequency range $[100, 300]$ Hz, the control effect of RPM is more obvious. The results in Fig. 2 (b) shows that the PM (true) and PM (measured) have about the same AC performance over the entire frequency range, while the proposed RPM achieves slightly better AC performance over the frequency ranges of $[280, 700]$ Hz and $[2200, 4000]$ Hz with the highest AC reaching approximately 24 dB at 2340 Hz.

4.3.2. Performance versus the perturbation level

We further evaluate the performance of those methods against different levels of positions mismatch. The energy constraints bounds e_{d} and e_{h} are chosen as -21.87 dB and -10 dB, respectively. The operating frequency is fixed at $f = 660$ Hz. As shown in Fig. 3 (a), the conventional PM method with true ATF has the smallest NSDE, while the PM (measured) method without considering the perturbation has the largest NSDE. The proposed RPM shows marginal improvement of NSDE within the entire perturbation level range. For a small perturbation, i.e., $\|\Delta \mathbf{G}_{\mathrm{B}}\|_{\mathrm{F}} = [0, 0.6]$, the NSDE of RPM is close to that of PM with true ATF. Even under severe perturbation conditions, e.g., $\|\Delta \mathbf{G}_{\mathrm{B}}\|_{\mathrm{F}} = 1.5$, the RPM (sphere) can still achieve approximately -2.4 dB NSDE performance, while the PM (measured) only can obtain approximately 0.3 dB. The results in Fig. 3 (b) show that the proposed RPM (sphere) has approximately the same AC performance as the PM with true ATF over the perturbation level range $[0, 0.9]$. Beyond this range, the proposed RPM obtains the best AC performance, even better than the PM (true) method. In case that perturbation is severe, e.g., $\|\Delta \mathbf{G}_{\mathrm{B}}\|_{\mathrm{F}} = 1.5$, the RPM (sphere) achieves approximately 10 dB AC performance, which is higher than that of the PM (true) method. These results demonstrate the advantage of the proposed method in the presence of severe perturbation, and also validates the trade-off relationship between the NSDE and AC performances.

Finally, we examine the effects of the spherical and ellipsoidal constraints on ATF estimation. As mentioned in Sec. 3.1, the ellipsoidal parameters, i.e., $[\mathrm{vec}(\widehat{\mathbf{G}}_{\mathrm{B}}), \mathbf{C}^{-1}]$ are usually decided according to the additional *a priori* information of ATF. We generate 600 ATFs equally spaced within the perturbation level $[0, 6]$, and the center $\mathrm{vec}(\widehat{\mathbf{G}}_{\mathrm{B}})$ and the matrix $\mathbf{C}$ denote the sample mean and the covariance matrix, respectively. Results show that the RPM (ellipsoid) achieves a better NSDE than that of the RPM (sphere) when the perturbation level is larger than 0.6, while the AC performances of the RPM (ellipsoid) and RPM (sphere) are almost the same.

5. CONCLUSION

In this work, a robust pressure matching method with an optimal ATF estimation was proposed to improve the robustness of sound field control systems. The proposed method imposes an uncertainty constraint on the ATF perturbation term to the PM method, leading to improved system robustness against sensor position mismatches. The biconvex optimization algorithm was applied to obtain the optimal ATF and loudspeaker control filter estimates. The developed algorithm was evaluated in a two-zone reproduction case and the results showed that the proposed approach achieved smaller signal distortion within the reproduction region. With the *a priori* information of ATF, the proposed method with a general ellipsoidal constraint achieved better normalized signal distortion energy performance than the proposed method with a sphere constraint.

6. REFERENCES

[1] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Speech, Audio Process.*, vol. 9, no. 6, pp. 697–707, 2001.

[2] T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, "Personal sound zones: Delivering interface-free audio to multiple listeners," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 81–91, 2015.

[3] D. Shi, B. Lam, S. Wen, and W.-S. Gan, "Multichannel active noise control with spatial derivative constraints to enlarge the quiet zone," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2020, pp. 8419–8423.

[4] T. Betlehem and P. D. Teal, "A constrained optimization approach for multi-zone surround sound," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 437–440.

[5] F. 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.*, vol. 25, no. 8, pp. 1579–1591, 2017.

[6] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2100–2111, 2005.

[7] N. Ueno, S. Koyama, and H. Saruwatari, "Three-dimensional sound field reproduction based on weighted mode-matching method," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 12, pp. 1852–1867, 2019.

[8] S. Koyama, T. Amakasu, N. Ueno, and H. Saruwatari, "Amplitude matching: Majorization–minimization algorithm for sound field control only with amplitude constraint," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2021, pp. 411–415.

[9] L. Shi, T. Lee, L. Zhang, J. K. Nielsen, and M. G. Christensen, "A fast reduced-rank sound zone control algorithm using the conjugate gradient method," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2020, pp. 436–440.

[10] F. Lluís, P. Martínez-Nuevo, M. B. Møller, and E. S. Sven, "Sound field reconstruction in rooms: Inpainting meets super-resolution," *J. Acoust. Soc. Am.*, vol. 148, no. 2, pp. 649–659, 2020.

[11] M. 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.*, vol. 27, no. 9, pp. 1405–1418, 2019.

[12] W. Jin and W. B. Kleijn, "Theory and design of multizone soundfield reproduction using sparse methods," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 12, pp. 2343 – 2355, 2015.

[13] J.-Y. Park, J.-W. Choi, and Y.-H. Kim, "Acoustic contrast sensitivity to transfer function errors in the design of a personal audio system," *J. Acoust. Soc. Am.*, vol. 134, no. 1, pp. 112–118, 2013.

[14] T. Betlehem, L. Krishnan, and P. Teal, "Temperature robust active-compensated sound field reproduction using impulse response shaping," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2018, pp. 481–485.

[15] P. Coleman, P. J. Jackson, M. Olik, M. B. Møller, M. Olsen, and J. A. Pederson, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *J. Acoust. Soc. Am.*, vol. 135, no. 4, pp. 1929–1940, 2014.

[16] T. Betlehem and C. Withers, "Sound field reproduction with energy constraint on loudspeaker weights," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 8, pp. 2388–2392, 2012.

[17] S. J. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 7, pp. 2123–2133, 2012.

[18] Q. 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.*, vol. 27, no. 3, pp. 610–620, 2018.

[19] Q. Zhu, P. Coleman, M. Wu, and J. Yang, "Robust acoustic contrast control with reduced in-situ measurement by acoustic modelling," *J. Audio Eng. Soc.*, vol. 65, no. 6, pp. 460–473, 2017.

[20] C. Benker, "Sensitivity of the sound zones problem to sources of error," Master's thesis, Technical University of Denmark, 2021.

[21] T. Lee, L. Shi, J. K. Nielsen, and M. G. Christensen, "Fast generation of sound zones using variable span trade-off filters in the DFT-domain," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 363–378, 2020.

[22] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, 2003.

[23] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, 2005.

[24] Y. Huang, M. Zhou, and S. A. Vorobyov, "New designs on mv-dr robust adaptive beamforming based on optimal steering vector estimation," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3624–3638, 2019.

[25] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: a survey and extensions," *Math. Method. Oper. Res.*, vol. 66, no. 3, pp. 373–407, 2007.

[26] S. P. Boyd and L. Vandenberghe, *Convex optimization*. US, New York: Cambridge university press, 2004.

[27] M. Grant and S. P. Boyd, "Cvx: Matlab software for disciplined convex programming," 2014, [Online]. Available: http://cvxr.com/cvx/.

[28] J. Zhang, W. Zhang, T. D. Abhayapala, and L. Zhang, "2.5D multizone reproduction using weighted mode matching: Performance analysis and experimental validation," *J. Acoust. Soc. Am.*, vol. 147, no. 3, pp. 1404–1417, 2020.

[29] E. A. P. Habets, "Room impulse response generator," Tech. Rep., 2006, [Online]. Available: https://github.com/ehabets/RIR-Generator.

[30] V. Molés-Cases, G. Piñero, M. D. Diego, and A. Gonzalez, "Personal sound zones by subband filtering and time domain optimization," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 2684–2696, 2020.

[31] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, 2004.