Loudspeaker Array to Maximize Acoustic Contrast Using Proximal Splitting Method

Kenta Imaizumi ⋆ Kenta Niwa † Kimitaka Tsutsumi ⋆
⋆ *NTT Corporation, NTT Service Evolution Laboratories, Kanagawa, Japan*
† *NTT Corporation, NTT Communication Science Laboratories, Kyoto, Japan*
kenta.imaizumi.uy@hco.ntt.co.jp

***Abstract*—We propose a method of creating acoustically bright and dark zones using a circular loudspeaker array. Practical reproduction must maximize acoustic contrast with a limited range of loudspeaker weights to ensure the quality of reproduced sounds. Therefore, we consider a constrained pressure-matching method that uses Tikhonov regularization. Furthermore, we impose additional sparse regularization to describe the ideal reproduced sound field as a simple model with a sparse structure. We obtain optimal loudspeaker weights by using an acoustic maximization algorithm based on a proximal splitting method. Numerical simulation results show that our method achieves a higher acoustic contrast than existing pressure-matching-based methods with a constraint of limited loudspeaker weights.**

***Index Terms*—acoustic contrast maximization, loudspeaker array, beamforming, sparse representation, convex optimization**

I. Introduction

Acoustic contrast maximization is a method of reproducing sound only in a target area (acoustically bright zone) while attenuating it in other areas (acoustically dark zones) [1]. This method enables listeners to enjoy audio content through audio interfaces without annoying other people nearby. Acoustic contrast maximization is implemented based on energy-based methods and sound field control methods. Energy-based methods maximize the ratio between the radiated power in a bright zone and that in a dark zone [2]–[4]. Sound field control methods reproduce a target sound pressure distribution defined at a set of control points. We propose a sound field control based method that can accurately reproduce the target sound field in the bright zone.

Sound field control methods are roughly divided into either pressure-matching based methods [5] or analytical methods [6]–[8]. Pressure-matching-based methods minimize error distribution between target sound fields and reproduced sound fields. These methods must solve linear inverse problems with a spatial correlation matrix of Green's functions of ill-determined rank. To obtain stable solutions, such linear inverse problems are replaced by nearby linear systems that are less sensitive to perturbations in a process called "regularization" [9]. One of the most commonly used regularization methods is Tikhonov regularization, which reduces the effect of instability in the inverse matrix of a spatial correlation matrix [10]. Many of the pressure-matching-based methods are based on this regularization [1]. Analytical methods can compute loudspeaker weights as solutions of a Helmholtz equation by comparing coefficients in a frequency [8] or a wavenumber domain [7].

Pressure-matching-based methods with regularization can be viewed as constrained convex optimization problems that minimize simple cost functions comprising multiple convex functions. More flexible "proximal splitting algorithms" [11], [12] have been introduced as a solution to such problems. Proximal splitting algorithms can deal with cost functions in which differentiable smooth functions like l_2 norm and some non-smooth functions like l_1 norm are mixed. These methods are used in applications such as signal and image processing, especially in sparse signal processing [13]. Similarly, in acoustic signal processing, algorithms have proven effective for sound field recording and reproduction [14], room impulse estimation [15], and multizone sound field reproduction [16].

In this study, we propose a method based on a constrained pressure-matching method. The conventional pressure-matching method minimizes the squared l_2 norm of the error distribution between target sound fields and reproduced sound fields. However, only minimizing the squared l_2 norm of the error distribution causes sound to leak into the dark zones. In addition to minimizing the l_2 norm, our method maximizes acoustic contrasts by using a sparsity of spatial energy distribution: Non-zero components reside only in a bright zone. To evaluate the sparsity of the spatial energy distribution, we use the mixed $l_{1,2}$ norm, which is often used to evaluate sparsity with a group structure. Furthermore, our method limits the loudspeaker weights by Tikhonov regularization. We formulate an algorithm based on the proximal splitting method to solve the complicated optimization problem that minimizes a cost function that consists of the sum of three convex functions that includes the above non-smooth functions.

II. Acoustic contrast maximization

A. Green's function for circular loudspeaker array

Assuming that the target field is the free field in a 2-dimensional space, the sound field created by a point source located at the position $\mathbf{r}'$ is modeled on Green's function [17]. Considering a plane wave impinging on a rigid cylinder and its scattering on the surface, the Green's function for the l-th loudspeaker of a rigid circular array with radius a can be derived as follows [18]:

$$G_{2D}^{CLA}(\mathbf{r}|\mathbf{r}'_l, k) = \sum_{\mu=-\infty}^{\infty} -\frac{H_\mu(kr)}{2\pi ka H'_\mu(ka)} e^{j\mu(\phi-\phi'_l)}, \quad (1)$$

where $\mathbf{r} = (r, \phi)$ is an arbitrary point in the polar coordinate system. $j = \sqrt{-1}$ is the imaginary unit and k is the wave number. H_μ is the μ-th order Hankel function of the second kind, and H'_μ is its first order derivative. Hereinafter, the wave number k is omitted for simplicity.

B. Problem formulation

Acoustic contrast maximization is composed of a loudspeaker array and the microphones that constitute the bright and dark zones. The method reproduces sound only in the bright zone and suppresses it in the dark zone. The bright and dark zones can be defined arbitrarily by changing the number and position of the microphones. In this study, we consider the acoustic contrast maximization of the exterior sound field. The exterior sound field $p(\mathbf{r})$ at a control point $\mathbf{r}$ reproduced by a circular loudspeaker array is expressed as follows:

$$p(\mathbf{r}) = \sum_{l=1}^{L} G_{2D}^{CLA}(\mathbf{r}|\mathbf{r}'_l) \cdot w_l, \tag{2}$$

where L is the number of loudspeakers. w_l denotes l-th loudspeaker weight. The sound fields at the control points are defined as $\mathbf{p}_B = [p_B^1, p_B^2, \ldots, p_B^{M_B}]^{\mathrm{T}}$ for the bright zone and $\mathbf{p}_D = [p_D^1, p_D^2, \ldots, p_D^{M_D}]^{\mathrm{T}}$ for the dark zone. M_B is the number of control points in the bright zone, and M_D is the number of control points in the dark zone. The superscript $^{\mathrm{T}}$ denotes transposition. The sound fields of the bright zone $\mathbf{p}_B$ are written as the following matrix-vector product:

$$\mathbf{p}_B = \mathbf{H}_B \mathbf{w}, \tag{3}$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_L]^{\mathrm{T}}$ is the vector of the loudspeaker weights. $\mathbf{H}_B$ is the matrix of the Green's functions between the loudspeaker array and the control points in the bright zone. The sound fields of dark zone $\mathbf{p}_D$ and the Green's functions between the loudspeaker array and the control points of dark zone $\mathbf{H}_D$ can be defined by equivalent notation.

C. Conventional methods

The pressure-matching method is based on the sound field control method. The pressure-matching method minimizes the squared sum of error distributions between the target sound field and the reproduced sound field. The cost function of the pressure-matching method with a weighted factor is as follows [19]:

$$\inf_{\mathbf{w}} (\mathbf{p}^{\mathrm{des}} - \mathbf{H}\mathbf{w})^{\mathrm{H}} \mathbf{A} (\mathbf{p}^{\mathrm{des}} - \mathbf{H}\mathbf{w}) + \lambda_{\mathbf{WPM}} \|\mathbf{w}\|^2, \tag{4}$$

where the superscript $^{\mathrm{H}}$ denotes the Hermitian transpose. $\mathbf{H} = (\mathbf{H}_B^{\mathrm{T}} \; \mathbf{H}_D^{\mathrm{T}})^{\mathrm{T}}$ is a Green's function matrix between all the control points and the loudspeakers. $\mathbf{p}^{\mathrm{des}} = \left(\mathbf{p}_B^{\mathrm{des}\,\mathrm{T}} \mathbf{p}_D^{\mathrm{des}\,\mathrm{T}}\right)^{\mathrm{T}}$ denote desired sound pressure distribution at all the control points. $\mathbf{p}_B^{\mathrm{des}}$ denotes the desired sound field in the bright zone, and $\mathbf{p}_D^{\mathrm{des}}$ denotes the desired sound field in the dark zones. $\|\cdot\|$ means the l_2 norm. $\lambda_{\mathbf{WPM}}$ is the regularization parameter. $\mathbf{A}$ is the diagonal matrix, where the diagonal elements store the weight parameters that correspond to each control point. An analytical solution to this problem can be derived by setting the first order derivative of (4) with respect to $\mathbf{w}$ to zero as:

$$\mathbf{w} = \left(\mathbf{H}^{\mathrm{H}} \mathbf{A} \mathbf{H} + \lambda_{\mathbf{WPM}} \mathbf{I}\right)^{\dagger} \left(\mathbf{H}^{\mathrm{H}} \mathbf{A} \mathbf{p}^{\mathrm{des}}\right), \tag{5}$$

where $\mathbf{I}$ denotes the identity matrix. † is the (pseudo) inverse of a matrix. If $\mathbf{A}$ is a unit matrix, (5) is the loudspeaker weight vector of the pressure-matching method.

The acoustic contrast control method is based on an energy-based method. The acoustic contrast control method maximizes the ratio of spatially averaged sound pressure levels between the bright zone and the dark zone [4].

$$\inf_{\mathbf{w}} \mathbf{p}_D^{\mathrm{H}} \mathbf{p}_D + \lambda_{\mathbf{ACC},1} (\mathbf{p}_B^{\mathrm{H}} \mathbf{p}_B - B) + \lambda_{\mathbf{ACC},2} (\mathbf{w}^{\mathrm{H}} \mathbf{w} - E), \tag{6}$$

where B is the desired sound pressure in the bright zone. E is the maximum allowed power of loudspeaker weight. $\lambda_{\mathbf{ACC},1}$ and $\lambda_{\mathbf{ACC},2}$ are the Lagrange multipliers. By setting the first derivative of (6) with respect to w to zero and rearranging the equation, we obtain the following equation:

$$\lambda_{\mathbf{ACC},1} \mathbf{w} = -(\mathbf{H}_B^{\mathrm{H}} \mathbf{H}_B)^{\dagger} (\mathbf{H}_D^{\mathrm{H}} \mathbf{H}_D + \lambda_{\mathbf{ACC},2} \mathbf{I}) \mathbf{w}, \tag{7}$$

(7) can be considered an eigenvalue problem. Therefore, selecting the maximum eigenvalue of $(\mathbf{H}_D^{\mathrm{H}} \mathbf{H}_D + \lambda_{\mathbf{ACC}} \mathbf{I})^{\dagger} (\mathbf{H}_B^{\mathrm{H}} \mathbf{H}_B)$ gives the loudspeaker weights of the acoustic contrast control method. $\lambda_{\mathbf{ACC}}$ is the regularization parameter of the acoustic contrast control method.

Conventional methods can create bright and dark zones. However, the pressure-matching method causes sound to leak into the dark zone because it only minimizes the squared sum of error distributions between the target sound field and the reproduced sound field. Moreover, the acoustic contrast control method cannot control the sound field in the bright zone arbitrarily. Our method imposes an additional constraint on the pressure-matching method to obtain a higher acoustic contrast than conventional methods.

III. Proposed Method

Sparse regularization is used in various acoustic signal processing applications. Sound field estimation in a near field assumes that there are only a few sources in the estimated sound field [20]. In addition, methods that use a loudspeaker array also assume that only a few loudspeaker weights have significant values [16]. Sparse regularization can describe the estimated sound field and loudspeaker weights as a simple model. Thus, we propose a novel acoustic contrast maximization method that uses sparse regularization. The acoustic contrast maximization method creates acoustically bright and dark zones. In an ideal reproduced sound field, non-zero components reside only in bright zones while the power of an ideal reproduced sound field in a dark-zone is zero. Based on previous studies that used sparse regularization [13]–[16], [20], we consider that sparse regularization could effectively describe the reproduced sound field as a simple model. This sparse distribution of reproduced sound field power can be defined explicitly with the following mixed $l_{1,2}$ norm [12]:

$$\|\mathbf{x}\|_{1,2} := \sum_{\mathfrak{g} \in \mathfrak{G}} \|\mathbf{x}_{\mathfrak{g}}\|, \tag{8}$$

Algorithm 1 Proximal splitting method for (10)

Input: $\gamma_1,\ \gamma_2.\ \lambda_1,\ \lambda_2,\ \mathbf{H}$
Initialize: $\mathbf{w}^{(0)},\ \mathbf{q}^{(0)},\ i \Leftarrow 0$
  while $i < T_{\max}$ **do**
    $\mathbf{w}^{(i+1)} \Leftarrow \mathrm{prox}_{\gamma_1\lambda_1 g}\left(\mathbf{w}^{(i)} - \gamma_1\left(\nabla f(\mathbf{w}^{(i)}) + \mathbf{H}^{\mathrm{H}}\mathbf{q}^{(i)}\right)\right)$
    $\tilde{\mathbf{q}}^{(i+1)} \Leftarrow \mathbf{q}^{(i)} + \gamma_2\mathbf{H}(2\mathbf{w}^{(i+1)} - \mathbf{w}^{(i)})$
    $\mathbf{q}^{(i+1)} \Leftarrow \tilde{\mathbf{q}}^{(i+1)} - \gamma_2\mathrm{prox}_{\frac{\lambda_2}{\gamma_2}h}\left(\frac{1}{\gamma_2}\tilde{\mathbf{q}}^{(i+1)}\right)$
    $i \Leftarrow i+1$
  end while
Output: $\mathbf{w} \Leftarrow \mathbf{w}^{(i+1)}$

where $\mathbf{x}_{\mathfrak{g}}$ denotes a vector comprising elements divided from $\mathbf{x}$ without overlapping each other. $\mathfrak{G}$ is an index set of $\mathbf{x}$. The mixed $l_{1,2}$ norm is often used to evaluate sparsity with a group structure [20]. By using the mixed $l_{1,2}$ norm (8), the sparse distribution of the reproduced sound field is defined by the following function:

$$\begin{aligned} h(\mathbf{p}) &= \|\mathbf{p}_B\|_{1,2} + \|\mathbf{p}_D\|_{1,2} \\ &= \|\mathbf{H}_B\mathbf{w}\|_{1,2} + \|\mathbf{H}_D\mathbf{w}\|_{1,2} \\ &= \|\mathbf{H}\mathbf{w}\|_{1,2}. \end{aligned} \tag{9}$$

In addition, loudspeaker weights must be limited so that the loudspeakers can reproduce sounds in proper quality. Many existing methods use the Tikhonov regularization that uses the l_2 norm as a constraint on the loudspeaker weights. We propose a method based on the pressure-matching method that includes the constraints on the sparse structure of the reproduced sound field by function (9) and Tikhonov regularization. However, this method includes a non-smooth function in the cost function, so it is difficult to obtain the loudspeaker weights directly like with a conventional pressure-matching method. Therefore, we propose an algorithm based on the proximal splitting method. The method is an iterative optimization method that can deal with a cost function that includes non-smooth functions [11], [12]. The optimization problem is defined as the sum of three convex functions as follows:

$$\begin{aligned} &\inf_{\mathbf{w},\mathbf{u}} \quad f(\mathbf{w}) + g(\mathbf{w}) + h(\mathbf{u}), \\ &f(\mathbf{w}) = (\mathbf{p}^{\mathrm{des}} - \mathbf{H}\mathbf{w})^{\mathrm{H}}\mathbf{A}(\mathbf{p}^{\mathrm{des}} - \mathbf{H}\mathbf{w}), \\ &g(\mathbf{w}) = \lambda_1\|\mathbf{w}\|^2, \\ &h(\mathbf{u}) = \lambda_2\|\mathbf{u}\|_{1,2}, \quad \text{s.t.}\ \mathbf{u} = \mathbf{H}\mathbf{w}, \end{aligned} \tag{10}$$

where λ_1 and λ_2 are non-negative balancing parameters that control the weight of each constraint. The optimization problem (10) is formulated into the following Lagrangian using the dual variable $\mathbf{q}$:

$$\inf_{\mathbf{w},\mathbf{u}} \sup_{\mathbf{q}} \quad f(\mathbf{w}) + g(\mathbf{w}) + h(\mathbf{u}) + \langle \mathbf{q}, \mathbf{H}\mathbf{w} - \mathbf{u}\rangle, \tag{11}$$

where $\langle , \rangle$ denotes the inner product of the vectors. Although there are methods of solving the primal problem directly, solving the primal-dual problem is easier because it is an unconstrained convex minimization problem. The primal-dual problem of (11) becomes a convex minimization problem with respect to $\mathbf{w}$ when $\mathbf{q}$ is fixed, and it becomes a concave function minimization problem with respect to $\mathbf{q}$ when $\mathbf{w}$ is fixed [21]. Therefore, we can consider the optimization problem separately for the primal variable $\mathbf{w}$ and the dual variable $\mathbf{q}$. Similarly, we can also separate them with respect to the constraints $g(\mathbf{w})$ and $h(\mathbf{u})$. The primal-dual problem of (11) can be expressed by the following equation:

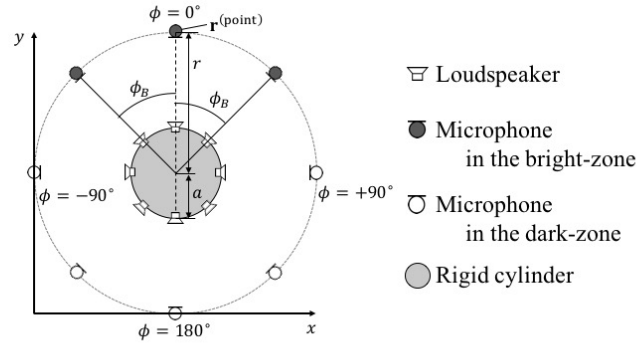

Fig. 1. Experimental setup of our personal sound field system. A circular loudspeaker array with radius a located at the center reproduces sound only in the bright zone defined by ϕ_B and suppresses sound in the dark zone.

$$\inf_{\mathbf{w}} \sup_{\mathbf{q}} \quad f(\mathbf{w}) + g(\mathbf{w}) - h^\star(\mathbf{q}) + \langle \mathbf{q}, \mathbf{H}\mathbf{w}\rangle, \tag{12}$$

where $h^\star(\cdot)$ is the convex conjugate of $h(\cdot)$.

$$h^\star(\mathbf{q}) = \sup_{\mathbf{u}} \left(\langle \mathbf{u}, \mathbf{q}\rangle - h(\mathbf{u})\right) \tag{13}$$

Finding the fixed point of (12) yields the solution to (12). To solve the convex minimization problem with respect to the primal variable $\mathbf{w}$ and the concave maximization problem with respect to the dual variable $\mathbf{q}$, we derive the subdifferential of (12):

$$\mathbf{0} \in \partial\left(f(\mathbf{w}) + g(\mathbf{w}) - h^\star(\mathbf{q}) + \langle \mathbf{q}, \mathbf{H}\mathbf{w}\rangle\right), \tag{14}$$

where ∂ denotes the subdifferential operator, $\in$ denotes the output of (11) as multi-valued when functions $g(\mathbf{w})$ and $h^\star(\mathbf{q})$ include non-smooth points. By using the proximal operator prox [22], (11) can be written as

$$\begin{aligned} &\mathbf{0} \in \nabla f(\mathbf{w}) + \partial g(\mathbf{w}) + \mathbf{H}^{\mathrm{H}}\mathbf{q}, \\ &\mathbf{w} \in (\mathrm{Id} + \gamma_1\partial g)(\mathbf{w}) + \gamma_1(\nabla f(\mathbf{w}) + \mathbf{H}^{\mathrm{H}}\mathbf{q}), \\ &\mathbf{w} \in (\mathrm{Id} + \gamma_1\partial g)^{-1}(\mathbf{w} - \gamma_1(\nabla f(\mathbf{w}) + \mathbf{H}^{\mathrm{H}}\mathbf{q})), \\ &\mathbf{w} = \mathrm{prox}_{\gamma_1\lambda_1 g}(\mathbf{w} - \gamma_1(\nabla f(\mathbf{w}) + \mathbf{H}^{\mathrm{H}}\mathbf{q})), \end{aligned} \tag{15}$$

$$\begin{aligned} &\mathbf{0} \in -\partial h^\star(\mathbf{q}) + \mathbf{H}\mathbf{w}, \\ &\mathbf{q} \in (\mathrm{Id} + \gamma_2\partial h^\star)(\mathbf{q}) - \gamma_2\mathbf{H}\mathbf{w}, \\ &\mathbf{q} \in (\mathrm{Id} + \gamma_2\partial h^\star)^{-1}(\mathbf{q} + \gamma_2\mathbf{H}\mathbf{w}), \\ &\mathbf{q} = \mathrm{prox}_{\gamma_2\lambda_2 h^\star}(\mathbf{q} + \gamma_2\mathbf{H}\mathbf{w}), \end{aligned} \tag{16}$$

where ∇ denotes the gradient for the differentiable function. Id is the identity, and $(\cdot)^{-1}$ is the inverse operator. γ_1 and γ_2 are step sizes. Since $\mathrm{prox}_{\gamma_2\lambda_2 h^\star}$ contains a subdifferential of a convex conjugate function and is complex, it can be obtained by Moreau's identity as follows [23]:

$$\mathrm{prox}_{\gamma_2\lambda_2 h^\star}(\mathbf{q}) = \mathbf{q} - \gamma_2\mathrm{prox}_{\frac{\lambda_2}{\gamma_2}h}\left(\frac{1}{\gamma_2}\mathbf{q}\right). \tag{17}$$

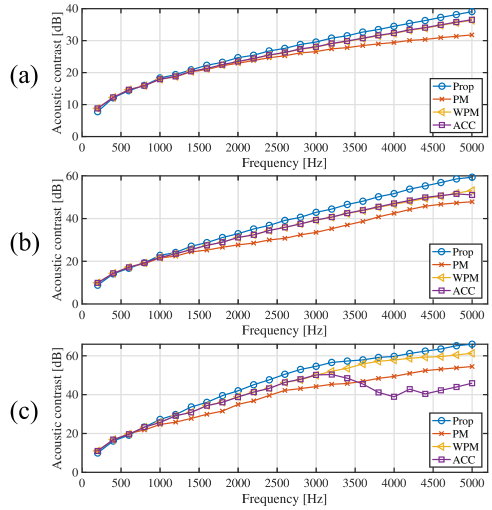

Fig. 2. Acoustic contrasts of our method and conventional methods. (a), (b), and (c) show the results when ϕ_B was set to $15^\circ, 30^\circ$, and 45°, respectively. The blue line, red line, yellow line, and purple line show the respective results of our method, **Prop**, **PM**, **WPM**, and **ACC**.

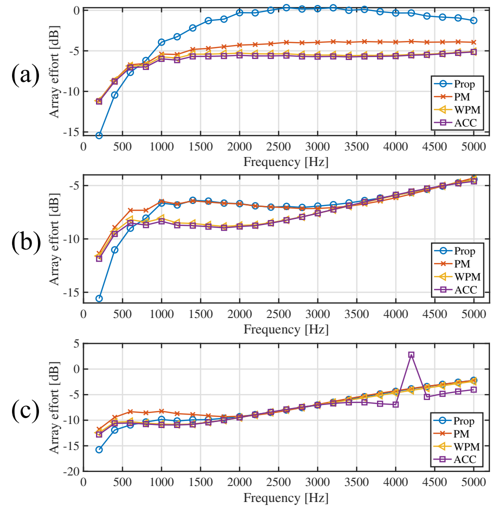

Fig. 3. Array efforts of our method and conventional methods. (a), (b), and (c) show the results when ϕ_B was set to $15^\circ, 30^\circ$, and 45°, respectively. The blue line, red line, yellow line, and purple line show the respective results of our method, **Prop**, **PM**, **WPM**, and **ACC**.

The proximity operators in $\mathrm{prox}_{\gamma_1 \|\cdot\|^2}$ and $\mathrm{prox}_{\frac{\lambda_2}{\gamma_2}\|\cdot\|_{1,2}}$ can be computed as follows:

$$\mathrm{prox}_{\lambda_1 g}(\mathbf{w}) = \frac{1}{1+\lambda_1}\mathbf{w}, \tag{18}$$

$$\left[\mathrm{prox}_{\frac{\lambda_2}{\gamma_2}\|\cdot\|_{1,2}}\left(\frac{1}{\gamma_2}\mathbf{q}\right)\right]_{\mathfrak{g}} = \max\left\{1 - \frac{\lambda_2}{\|\mathbf{q}_{\mathfrak{g}}\|}, 0\right\}\frac{1}{\gamma_2}\mathbf{q}_{\mathfrak{g}}. \tag{19}$$

Considering Nesterov's accelerated method [24] with respect to $\mathbf{w}$, the problem (10) can be solved by Algorithm 1.

IV. Performance Evaluation

A. Experimental setups

To investigate our method's performance, we conducted a numerical simulation of acoustic contrast maximization. We compared our method with the pressure-matching method (**PM**), the pressure-matching method with weighted factors (**WPM**), and the acoustic contrast control method (**ACC**).

Figure 1 shows the experimental setup. In all the simulations, a two-dimensional free field was assumed, and the speed of sound was set to $343.36\,\mathrm{m/s}$. The rigid circular loudspeaker array with radius $a = 0.10\,\mathrm{m}$ was located on the x-y plane. The number of loudspeakers L was set to 32. The control points were located on a circle with radius $r = 1.0\,\mathrm{m}$ (the control points were set to 360 points arranged in the ϕ direction at intervals of 1°). The Green's functions were calculated using (1) in which μ was truncated to 100. The area of $\phi = \pm\phi_B$ was defined as the bright zone, and the other areas were defined as the dark zone. A Hanning window was used as the desired sound field $\mathbf{p}^{\mathrm{des}}$ so that the sound field smoothly decayed from the bright zone to the dark zone [25].

$$\mathbf{p}^{\mathrm{des}}(\phi) = \begin{cases} 0.5 \times (1 - \cos\frac{\phi_B + \phi}{\phi_B}\pi) & (|\phi| \le \phi_B), \\ 0 & (\text{otherwise}). \end{cases} \tag{20}$$

The regularization parameters were set to $\lambda_{\mathbf{PM}} = \sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H}) \times 10^{-2}, \lambda_{\mathbf{WPM}} = \sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H})$, and $\lambda_{\mathbf{ACC}} = \sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H}) \times 10^{-2}$. $\sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H})$ denotes the maximum singular value of a $(\mathbf{H}^{\mathrm{H}}\mathbf{H})$. All the regularization parameters were set experimentally. To further suppress the sound in the dark zone, the weighting parameters of $\mathbf{A}$ of **WPM** and our method were set to 1 in the bright zone and 100 in the dark zone. The balancing parameters λ_1 and λ_2 and the step sizes γ_1 and γ_2 of our method were set to $\lambda_1 = \sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H}) \times 10^{-3}$, $\lambda_2 = 0.10$, $\gamma_1 = 2/\kappa \times 0.75 \times 10^{-2}$, and $\gamma_2 = (1/\gamma_1 - \kappa/2)/\sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H}) \times 0.75 \times 10^{-2}$, respectively. $\kappa = \sigma_{\max}(\mathbf{H}^{\mathrm{H}}\mathbf{H})$ satisfies the Lipschitz constant requirement.

To evaluate the acoustic contrast maximization methods, we defined the acoustic contrast and the array effort as performance measures. Acoustic contrast is the ratio between the mean square pressure in a bright zone and that in a dark zone [3]. Acoustic contrast is defined as follows [4]:

$$\mathrm{AC} = 10\log_{10}\frac{M_D\|\mathbf{p}_B(r_{ref}, \phi_b)\|^2}{M_B\|\mathbf{p}_D(r_{ref}, \phi_d)\|^2}. \tag{21}$$

The radial distance r_{ref} used in the evaluation is set to $r - 0.2 \le r_{ref} \le r + 0.2$ and the widths of the bright zone and the dark zone are $\phi_b \le |\phi_B|$ and $\phi_d > |\phi_B|$. The array effort is defined as the sum of the squared source strength normalized by the source strength of a single monopole at the center of a loudspeaker array that produces the same mean square pressure at an arbitrary point $\mathbf{r}^{(\mathrm{point})}$ in the bright zone [4].

$$\mathrm{AE} = 10\log_{10}\frac{\mathbf{w}^{\mathrm{H}}\mathbf{w}}{|w_0|^2}, \tag{22}$$

where w_0 is the source strength of a single monopole at the center of a loudspeaker array that produces the same mean square pressure at the point $\mathbf{r}^{(\mathrm{point})}$.

$$w_0 = \frac{\mathbf{H}_B^{(\mathrm{point})}\mathbf{w}}{H_0}, \tag{23}$$

where $\mathbf{H}_B^{(\mathrm{point})}$ is the vector of Green's functions between the loudspeaker array and the point $\mathbf{r}^{(\mathrm{point})}$ in the bright zone. H_0 is the Green's function between the monopole located at the center of the loudspeaker array and the point $\mathbf{r}^{(\mathrm{point})}$. We set point $\mathbf{r}^{(\mathrm{point})}$ to the control point at the $\phi = 0°$.

B. Results

Fig.2 shows the acoustic contrast. In Figs.2, (a), (b), and (c) show the results when ϕ_B was set to $15°, 30°$, and $45°$, respectively. The acoustic contrast was confirmed to increase as the frequency went higher in almost all the methods independent of the bright zone's range. However, the acoustic contrast of **ACC** at high frequencies reduced when the bright zone's range widened. Our method obtained higher acoustic contrast than the conventional methods. Since **PM** and **WPM** are methods based on pressure-matching, the acoustic contrast does not appear to increase because the generated sound field is focused on matching the desired sound field. Although our method is also based on pressure-matching, the acoustic contrast is seemingly high because the sparseness of the reproduced sound field is evaluated by the mixed $l_{1,2}$ norm. By modeling the reproduced sound field using sparse regularization, the acoustic contrast of our method is higher than the **ACC**, which aims to maximize the acoustic contrast.

Fig.3 shows the array efforts of our method and conventional methods. In Figs.3, (a), (b), and (c) show the results when ϕ_B was set to $15°, 30°$, and $45°$, respectively. The acoustic contrast of our method is slightly higher than that of the conventional methods. However, the array efforts of all the methods are low enough that the loudspeakers can reproduce proper quality. Since the array effort is lower than $0\ \mathrm{dB}$, our method can reproduce the sound more efficiently than a monopole.

V. Conclusion

We proposed a method of creating acoustically bright and dark zones by using a rigid circular loudspeaker array. Our method described the reproduced sound field using sparse regularization as a simple model. We formulated an algorithm based on a proximal splitting method to solve the complicated optimization problem with a cost function that consists of smooth and non-smooth functions. The results of numerical experiments indicated that the acoustic contrast of our method is higher than that of conventional methods. In addition, our method can reproduce processed sounds of proper quality by Tikhonov regularization.

Experimentally determined hyperparameters such as step size greatly affected our results. In future work, we will consider selecting step sizes and balancing the parameters of a proximal splitting method for optimal performance. In addition, we will conduct a detailed experiment using an actual machine in an anechoic chamber.

References

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

[2] W.F. Druyvesteyn and J. Garas, "Personal sound," *Journal of the Audio Engineering Society*, vol. 45, no. 9, pp. 685–701, 1997.

[3] J. W. Choi and Y. H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1695–1700, 2002.

[4] S. J. Elliott, J. Cheer, J. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 7, pp. 2123–2133, Sep. 2012.

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

[6] S. Spors K. Helwani and H. Buchner, "The synthesis of sound figures," *Multidimensional Systems and Signal Processing*, vol. 25, no. 2, pp. 379–403, Nov 2013.

[7] T. Okamoto, "Analytical methods of generating multiple sound zones for open and baffled circular loudspeaker arrays," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct 2015, pp. 1–5.

[8] K. Tsutsumi, K. Imaizumi, A. Nakadaira, and Y. Haneda, "Analytical method to convert circular harmonic expansion coefficients for sound field synthesis by using multipole loudspeaker array," in *European Signal Processing Conference (EUSIPCO)*, Sep. 2019, pp. 1–5.

[9] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, "Tikhonov regularization and the L-curve for large discrete ill-posed problems," *Journal of Computational and Applied Mathematics*, vol. 123, no. 1, pp. 423 – 446, 2000, Numerical Analysis 2000. Vol. III: Linear Algebra.

[10] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Dokl.*, vol. 4, pp. 1035–1038, 1963.

[11] P. L. Combettes and J. C. Pesquet, "Proximal splitting methods in signal processing," *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* Springer-Verlag, pp. 185–212, 2011.

[12] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends® in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[13] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[14] N. Murata S. Koyama and H. Saruwatari, "Sparse sound field decomposition for super-resolution in recording and reproduction," *The Journal of the Acoustical Society of America*, vol. 143, no. 6, pp. 3780–3795, 2018.

[15] R. Mignot, L. Daudet, and F. Ollivier, "Room reverberation reconstruction: Interpolation of the early part using compressed sensing," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 11, pp. 2301–2312, Nov 2013.

[16] N. Radmanesh and I. S. Burnett, "Generation of isolated wideband sound fields using a combined two-stage lasso-ls algorithm," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 2, pp. 378–387, Feb 2013.

[17] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Elsevier Science, 1999.

[18] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc*, vol. 53, no. 11, pp. 1004–1025, 2005.

[19] S. Doclo and M. Moonen, "Design of far-field and near-field broadband beamformers using eigenfilters," *Signal Processing*, vol. 83, no. 12, pp. 2641–2673, 2003.

[20] S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 1, pp. 172–184, March 2019.

[21] Y. Nesterov, "Primal-dual subgradient methods for convex problems," *Mathematical programming*, vol. 120, no. 1, pp. 221–259, 2009.

[22] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[23] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of Optimization Theory and Applications*, vol. 158, no. 2, pp. 460–479, 2013.

[24] Y. Nesterov, "A method of solving a convex programming problem with convergence rate $o(1/k^2)$," in *Sov. Math. Dokl*, vol. 27.

[25] K. Sato and Y. Haneda, "Sidelobe suppression by desired directivity pattern optimization for a small circular loudspeaker array," *Acoustical Science and Technology*, vol. 39, no. 3, pp. 243–251, 2018.