





Amplitude Matching for Multizone Sound Field Control

Takumi Abe , Shoichi Koyama , *Member, IEEE*, Natsuki Ueno , *Member, IEEE*,
and Hiroshi Saruwatari , *Member, IEEE*

Abstract—A multizone sound field control method, called **amplitude matching**, is proposed. The objective of **amplitude matching** is to synthesize a desired amplitude (or magnitude) distribution over a target region with multiple loudspeakers, whereas the phase distribution is arbitrary. Most of the current multizone sound field control methods are intended to synthesize a specific sound field including phase or to control acoustic potential energy inside the target region. In **amplitude matching**, a specific desired amplitude distribution can be set, ignoring sound propagation directions. Although the optimization problem of **amplitude matching** does not have a closed-form solution, our proposed algorithm based on the alternating direction method of multipliers (ADMM) allows us to accurately and efficiently synthesize the desired amplitude distribution. We also introduce the differential-norm penalty for a time-domain filter design with a small filter length. The experimental results indicated that the proposed method outperforms current multizone sound field control methods in terms of accuracy of the synthesized amplitude distribution.

Index Terms—Multizone sound field control, personal audio, pressure matching, amplitude matching.

I. INTRODUCTION

SYNTHESIZING a desired sound field inside a target region using multiple loudspeakers (secondary sources) has various applications, such as virtual/augmented reality, spatial noise cancellation, and personal sound-zone generation. Sound field synthesis/control methods can be classified into two categories. One includes methods based on boundary integral equations analytically derived from the Helmholtz equation, such as wave field synthesis and higher-order ambisonics [1], [2], [3], [4], [5], [6]. The other includes methods based on numerical optimization to minimize error between synthesized and desired sound fields inside the target region, such as pressure matching (PM) and (weighted) mode matching [2], [7], [8], [9], [10]. Since the numerical-optimization-based methods enable us to

generate complex sound fields with a flexible array geometry of loudspeakers, they have a broad range of practical applications.

Multizone sound field control [11], [12], [13], [14], [15], which is one particular problem in sound field control, aims at generating different sound fields inside multiple target regions for personal audio applications [16], [17]. For example, plane-wave fields with different propagation angles are synthesized inside two regions. However, this problem is sometimes impractical because its physical feasibility becomes significantly low depending on the desired propagation angles of plane-waves.

Meanwhile, in some applications, it is desired to generate sound fields of certain acoustic power levels rather than synthesizing specific sound fields such as plane-wave fields of particular propagation angles. For example, the acoustic power should be high in one region, but suppressed in another region, focusing on whether the sound is audible or not. In previous studies, this type of problem has been addressed as an acoustic contrast control (ACC) problem [18], [19], where the ratio of acoustic potential energy in one region to that in another region is maximized to generate *acoustic bright* and *dark zones*. However, the power distribution inside the target region cannot be controlled by ACC because only the total energy is taken into account. Furthermore, ACC suffers from undesired distortion of filters due to discontinuities between frequencies when applied in the frequency domain. When ACC is applied in the time domain, a flat frequency response cannot be guaranteed.

We consider the problem of synthesizing specific amplitude (or magnitude) distributions inside the target regions by using secondary sources, which is referred to as *amplitude matching*. For example, a flat amplitude distribution over a target region is set as a desired distribution but its phase distribution is left arbitrary. Such an optimization problem does not have closed-form solution; therefore, iterative algorithms are necessary [20]. However, because of the indifferentiability of the cost function of amplitude matching, general gradient methods, such as gradient descent and (quasi-)Newton's methods, cannot be applied in a strict sense. Even if these algorithms work in practice, putting the theoretical validity aside, they basically require a high computational load for computing the inverse of a Hessian matrix (or its proxy) or searching step size parameter at each iteration.

In our previous study [21], we proposed amplitude matching based on the majorization–minimization (MM) algorithm [22], [23], where the optimization problem of amplitude matching is efficiently solved in the frequency domain with a guarantee

Manuscript received 4 September 2022; revised 30 November 2022; accepted 16 December 2022. Date of publication 23 December 2022; date of current version 3 January 2023. This work was supported in part by Japan Science and Technology Agency (JST) PRESTO under Grant JPMJPR18J4 and in part by Japan Science and Technology Agency (JST) FOREST Program, Japan, under Grant JPMJFR216M. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefan Bilbao. (Corresponding author: Shoichi Koyama.)

The authors are with the Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan (e-mail: takumi-abe@g.ecc.u-tokyo.ac.jp; koyama.shoichi@ieee.org; natsuki.ueno@ieee.org; hiroshi_saruwatari@ipc.i.u-tokyo.ac.jp).

Digital Object Identifier 10.1109/TASLP.2022.3231715

of a monotonic nonincrease in the cost function. However, this algorithm tends to be stagnated at local minima or saddle points. Furthermore, solving the amplitude matching problem at each frequency can lead to discontinuities in phase between frequency bins as in ACC. As a result, in a broadband case, the time-domain filter to obtain driving signals of the secondary sources can be unnecessarily long.

We propose amplitude matching based on the alternating direction method of multipliers (ADMM) [24] to achieve a more accurate synthesis of the amplitude distribution with an efficient procedure. We also introduce the differential-norm penalty to prevent discontinuities between frequency bins in the broadband case. The proposed amplitude matching method is evaluated by numerical experiments and compared with current multizone sound field control methods and amplitude matching based on MM algorithm. In addition, results of experiments using real data obtained in a practical environment are also presented.

The rest of this paper is organized as follows. In Section II, prior works on multizone sound field control methods are briefly summarized. The problem statement on amplitude matching is described with current algorithms for solving it in Section III. In Section IV, the proposed ADMM-based algorithm is developed. Differential-norm penalty for a time-domain filter design is introduced in Section V. Experimental results in the narrow-band and broadband cases are presented in Section VI. Finally, Section VII concludes this paper.

A. Notation

Italic letters denote scalars, lower case boldface italic letters denote vectors, and upper case boldface italic letters denote matrices. The sets of real and complex numbers are denoted by \mathbb{R} and \mathbb{C} , respectively. Subscripts of scalars, vectors, and matrices indicate their indexes. To illustrate, $x_{i,j}$ is the (i,j) th entry of the matrix \mathbf{X} . The imaginary unit and Napier's constant are denoted by j ($j^2 := -1$) and e , respectively. The complex conjugate, transpose, conjugate transpose, and inverse are denoted by superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$, respectively. The absolute value of a scalar x is denoted by $|x|$. The argument of complex-valued scalar x is denoted by $\arg(x)$. The signs $|\cdot|$ and $\arg(\cdot)$, and exponential e are also applied to vectors for each element. For example, $|\mathbf{x}| \odot e^{j\arg(\mathbf{y})}$ means the vector consisting of $|x_i|e^{j\arg(y_i)}$ with the Hadamard product \odot . The Euclidean norm of a vector \mathbf{x} is denoted by $\|\mathbf{x}\|$.

The angular frequency, sound velocity, and wavenumber are denoted by ω , c , and $k = \omega/c$, respectively. The harmonic time dependence $e^{-j\omega t}$ with the time t is assumed according to conventions.

II. PRIOR WORKS ON SOUND FIELD CONTROL

Sound field control is aimed at synthesizing a desired sound field inside a target region. Driving signals of L secondary sources are controlled so that the synthesized sound field corresponds to the desired one inside the target region Ω . The secondary sources are arbitrarily placed in the region outside Ω , as shown in Fig. 1. The target region Ω can consist of

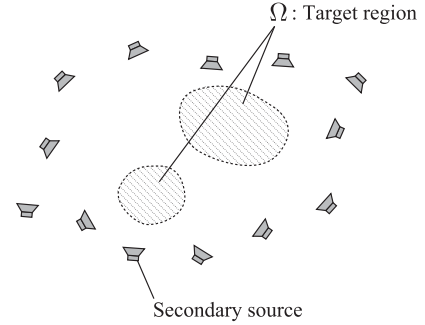


Fig. 1. Sound field control in multiple target regions.

multiple non-overlapped regions. The synthesized sound pressure $u^{\text{syn}}(\mathbf{r}, \omega)$ of angular frequency ω at position $\mathbf{r} \in \Omega$ is represented by a linear combination of transfer functions of the secondary sources as

$$u^{\text{syn}}(\mathbf{r}, \omega) = \sum_{l=1}^L d_l(\omega) g_l(\mathbf{r}, \omega), \quad (1)$$

where $d_l(\omega)$ and $g_l(\mathbf{r}, \omega)$ are the driving signal and the transfer function (Green's function) at \mathbf{r} for the l th secondary source ($l \in \{1, \dots, L\}$), respectively. Hereafter, the angular frequency ω is omitted for notational simplicity.

The objective of general sound field control is to obtain the driving signals $\mathbf{d} = [d_1, \dots, d_L]^T \in \mathbb{C}^L$ so that the following cost function $F(\mathbf{d})$ is minimized:

$$F(\mathbf{d}) := \int_{\Omega} \left| \sum_{l=1}^L d_l g_l(\mathbf{r}) - u^{\text{des}}(\mathbf{r}) \right|^2 d\mathbf{r}, \quad (2)$$

where $u^{\text{des}}(\mathbf{r})$ is the desired sound field. When Ω is separated into multiple regions, and different $u^{\text{des}}(\mathbf{r})$ values are set inside each region, this minimization problem is specifically called multizone sound field control. Since minimizing (2) is difficult to solve owing to the regional integral, several approximation methods have been proposed.

A. Pressure Matching

Pressure matching is a widely used sound field control method since its practical implementation is simple and flexible. The cost function $F(\mathbf{d})$ is approximated by discretizing the target region Ω into control points [7], [10], [11]. The number of control points and the position of the m th control point are denoted by M and $\mathbf{r}_m \in \Omega$ ($m \in \{1, \dots, M\}$), respectively. We define the vector $\mathbf{u}^{\text{des}} \in \mathbb{C}^M$ consisting of $u^{\text{des}}(\mathbf{r}_m)$ and the matrix $\mathbf{G} \in \mathbb{C}^{M \times L}$ consisting of $g_l(\mathbf{r}_m)$. Thus, the optimization problem for the pressure matching is formulated as

$$\underset{\mathbf{d}}{\text{minimize}} F_{\text{PM}}(\mathbf{d}) := \|\mathbf{G}\mathbf{d} - \mathbf{u}^{\text{des}}\|^2 + \gamma\|\mathbf{d}\|^2, \quad (3)$$

where γ is a constant parameter. The first term is the square error between the synthesized and desired sound pressures at the control points, and the second term is the regularization term to prevent an excessive amplitude of \mathbf{d} . To make the problem balanced or overdetermined, the number of control points M is

generally set to be larger than or equal to L . This optimization problem is solved in a closed form as the regularized least-squares solution:

$$\hat{\mathbf{d}} = (\mathbf{G}^H \mathbf{G} + \gamma \mathbf{I})^{-1} \mathbf{G}^H \mathbf{u}^{\text{des}}. \quad (4)$$

The pressure-matching method can be directly applied to the multizone sound field control.

Instead of discretizing the target region Ω , the sound field is represented by spherical wave function expansion in *mode-matching methods* [2]. The driving signal \mathbf{d} is obtained so that the expansion coefficients of $u^{\text{syn}}(\mathbf{r})$ correspond to those of $u^{\text{des}}(\mathbf{r})$. The weighting factors for each expansion coefficient are introduced in the weighted mode-matching [9] for accurate approximation of the original cost function (2) without empirical truncation of the expansion order. The mode-matching methods are also extended to multizone sound field control [12], [13].

B. Acoustic Contrast Control

Generating regions of high- and low-acoustic power distributions is one particular example of the multizone sound field control. In ACC [18], [19], the target region Ω is separated into two regions Ω_B and Ω_D ($\Omega_B \cup \Omega_D = \Omega$ and $\Omega_B \cap \Omega_D = \emptyset$). The regions Ω_B and Ω_D are the so-called (acoustic) bright and dark zones, respectively. The goal is to synthesize high acoustic potential energy inside Ω_B and low acoustic potential energy inside Ω_D .

The optimization problem of ACC is defined as the maximization of the ratio of acoustic potential energy inside Ω_B to that inside Ω_D as

$$\text{maximize}_{\mathbf{d}} Q(\mathbf{d}) := \frac{\|\mathbf{G}_B \mathbf{d}\|^2}{\|\mathbf{G}_D \mathbf{d}\|^2} = \frac{\mathbf{d}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{d}}{\mathbf{d}^H \mathbf{G}_D^H \mathbf{G}_D \mathbf{d}}, \quad (5)$$

where \mathbf{G}_B and \mathbf{G}_D are the transfer function matrices at the control points inside Ω_B and Ω_D , respectively. The optimal \mathbf{d} can be obtained as the eigenvector that corresponds to the largest eigenvalue of the matrix

$$(\mathbf{G}_D^H \mathbf{G}_D + \mu \mathbf{I})^{-1} \mathbf{G}_B^H \mathbf{G}_B, \quad (6)$$

where μ is the regularization parameter. This solution has indefiniteness for the scaling and phase rotation of \mathbf{d} . Since only the total acoustic energy inside the target region is considered in (5), it is not possible to synthesize specific power distributions. Therefore, a large variation of acoustic power inside the bright zone can be generated. Furthermore, solving (5) for each frequency leads to discontinuities in phase between frequency bins.

III. AMPLITUDE MATCHING

Pressure matching is applicable to synthesizing various desired sound fields inside the target region. However, in some applications, it is necessary to synthesize a sound field of the desired amplitude inside the target region, whereas the phase of the desired sound field is arbitrary. For example, the sound field of high- and low-acoustic-power distributions should be synthesized similarly to that in ACC, which will be useful in car audio and sound navigation systems. We formulate the

optimization problem to achieve sound field control only with the amplitude constraint, i.e., amplitude matching, as

$$\text{minimize}_{\mathbf{d}} J(\mathbf{d}) := \left\| |\mathbf{G} \mathbf{d}| - |\mathbf{u}^{\text{des}}| \right\|^2 + \lambda \|\mathbf{d}\|^2, \quad (7)$$

where λ is the constant parameter. The first term is the square error between the synthesized and desired amplitude distributions at the control points, and the second term is the regularization term of \mathbf{d} . In contrast to the cost function of ACC (5), (7) is designed to measure the error of spatial amplitude distributions. Thus, various amplitude distributions can be set as the desired sound field in the amplitude matching.

Since the optimization problem of the amplitude matching does not have closed-form solution owing to the nonconvexity and indifferentiability of the cost function $J(\mathbf{d})$, iterative algorithms are necessary to solve (7). Gradient methods, such as gradient descent and (quasi-)Newton's method, are widely used nonlinear optimization algorithms. Since $J(\mathbf{d})$ is not differentiable at an arbitrary point, the gradient methods are not applicable to solving (7) in a strict sense, but they can work in practice. The gradient of $J(\mathbf{d})$ can be obtained as

$$\nabla J(\mathbf{d}) = \mathbf{G}^H (|\mathbf{G} \mathbf{d}| - |\mathbf{u}^{\text{des}}|) \odot e^{j \arg(\mathbf{G} \mathbf{d})} + \lambda \mathbf{d}, \quad (8)$$

where $\mathbf{G} \mathbf{d}$ is assumed not to contain zero-valued components. By using (8), we can construct the gradient methods for solving (7), such as gradient descent and quasi-Newton's methods. In addition to the lack of theoretical validity, the computational complexity of these algorithms can be high, as discussed later.

A. MM Algorithm

An amplitude matching method based on the MM algorithm has been proposed by Koyama et al. [21]. In the MM algorithm, a surrogate function for a nonconvex objective function, which can be simply minimized, is constructed. The monotonic nonincrease in the objective function can be guaranteed by alternately updating the variable of the surrogate function and the variable to be optimized. The resulting update rules can be computationally efficient by appropriate construction of the surrogate function.

In [21], the surrogate function of $J(\mathbf{d})$ is obtained as

$$J^+(\mathbf{d}|\mathbf{v}^{(i)}) := \|\mathbf{G} \mathbf{d} - \mathbf{v}^{(i)}\|^2 + \lambda \|\mathbf{d}\|^2, \quad (9)$$

where $\mathbf{v}^{(i)}$ is the auxiliary variable defined with the iteration index i as

$$\mathbf{v}^{(i)} := |\mathbf{u}^{\text{des}}| \odot e^{j \arg(\mathbf{G} \mathbf{d}^{(i)})}. \quad (10)$$

Here, $\mathbf{d}^{(i)}$ is \mathbf{d} at the i th iteration. This surrogate function satisfies

$$J(\mathbf{d}) \leq J^+(\mathbf{d}|\mathbf{v}^{(i)}), \quad (11)$$

where the equality holds for $\mathbf{d} = \mathbf{d}^{(i)}$. See Appendix A for the proof. Thus, the driving signal \mathbf{d} is obtained by alternately updating $\mathbf{v}^{(i)}$ and $\mathbf{d}^{(i)}$ as

$$\mathbf{v}^{(i)} = |\mathbf{u}^{\text{des}}| \odot e^{j \arg(\mathbf{G} \mathbf{d}^{(i)})}, \quad (12)$$

$$\begin{aligned} \mathbf{d}^{(i+1)} &= \arg \min_{\mathbf{d}} J^+(\mathbf{d}|\mathbf{v}^{(i)}) \\ &= (\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{v}^{(i)}. \end{aligned} \quad (13)$$

Algorithm 1: MM Algorithm for Amplitude Matching.

Input: $\mathbf{d}^{(0)}, |\mathbf{u}^{\text{des}}|, \mathbf{G}, \lambda$
 1: $\mathbf{v}^{(0)} = |\mathbf{u}^{\text{des}}| \odot e^{j \arg(\mathbf{G}\mathbf{d}^{(0)})}$
 2: $i = 0$
 3: **while** stopping criterion not satisfied **do**
 4: $\mathbf{d}^{(i+1)} = (\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{v}^{(i)}$
 5: $\mathbf{v}^{(i+1)} = |\mathbf{u}^{\text{des}}| \odot e^{j \arg(\mathbf{G}\mathbf{d}^{(i+1)})}$
 6: $i = i + 1$
 7: **end while**
 8: **return** $\mathbf{d}^{(i)}$

The monotonic nonincrease in the objective function can be confirmed as

$$J(\mathbf{d}^{(i+1)}) \leq J^+(\mathbf{d}^{(i+1)}|\mathbf{v}^{(i)}) \leq J^+(\mathbf{d}^{(i)}|\mathbf{v}^{(i)}) = J(\mathbf{d}^{(i)}). \quad (14)$$

The iterative update rules are computed until a stopping condition is satisfied, for example, by setting a threshold for the variation of $J(\mathbf{d}^{(i)})$ or $\mathbf{d}^{(i)}$. The MM algorithm for the amplitude matching is summarized in Algorithm 1. Note that $(\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H$ in the fourth line can be calculated before the iteration. Thus, the computational cost for each iteration is $O(LM)$. This algorithm corresponds to the Griffin–Lim algorithm [25] and Gerchberg–Saxton algorithm [26] for the phase retrieval problem, whose cost function is similar to that in (7).

IV. ALTERNATING DIRECTION METHOD OF MULTIPLIERS FOR AMPLITUDE MATCHING

The MM algorithm presented in Section III-A is computationally efficient and simple for implementation. However, in some cases, the cost function does not sufficiently decrease, as later shown in the experiments. We propose the amplitude matching algorithm based on ADMM. For convex optimization problems, the convergence to the optimal solution can be guaranteed by ADMM [24]. It has been experimentally validated that ADMM effectively works for some nonconvex optimization problems [24], [27], [28]. In particular, Liang et al. [27] have shown that ADMM performs well in phase retrieval problems.

Since ADMM cannot be directly applied to solve (7), we reformulate the optimization problem by using auxiliary variables of amplitude and phase defined as $\mathbf{G}\mathbf{d} = \mathbf{a} \odot e^{j\theta}$ as

$$\begin{aligned} & \underset{\mathbf{d}, \mathbf{a} \geq 0, \theta}{\text{minimize}} \quad \|\mathbf{a} - |\mathbf{u}^{\text{des}}|\|^2 + \lambda \|\mathbf{d}\|^2 \\ & \text{subject to} \quad \mathbf{G}\mathbf{d} = \mathbf{a} \odot e^{j\theta}, \end{aligned} \quad (15)$$

where $\mathbf{a} \geq 0$ means that each element of \mathbf{a} is equal to or larger than 0. Then, the augmented Lagrangian function for (15) can be defined as

$$\begin{aligned} L_\rho(\mathbf{d}, \mathbf{a}, \theta, \mathbf{w}) := & \|\mathbf{a} - |\mathbf{u}^{\text{des}}|\|^2 + \lambda \|\mathbf{d}\|^2 \\ & + \Re[\mathbf{w}^H (\mathbf{G}\mathbf{d} - \mathbf{a} \odot e^{j\theta})] \\ & + \frac{\rho}{2} \|\mathbf{G}\mathbf{d} - \mathbf{a} \odot e^{j\theta}\|^2, \end{aligned} \quad (16)$$

where $\mathbf{w} \in \mathbb{C}^M$ is the Lagrange multiplier, $\Re[\cdot]$ represents the real part of complex value, and $\rho > 0$ is the penalty parameter. In ADMM, each variable is alternately updated, starting with initial values as

$$\{\mathbf{a}^{(i+1)}, \theta^{(i+1)}\} = \arg \min_{\mathbf{a} \geq 0, \theta} L_\rho(\mathbf{d}^{(i)}, \mathbf{a}, \theta, \mathbf{w}^{(i)}), \quad (17)$$

$$\mathbf{d}^{(i+1)} = \arg \min_{\mathbf{d}} L_\rho(\mathbf{d}, \mathbf{a}^{(i+1)}, \theta^{(i+1)}, \mathbf{w}^{(i)}), \quad (18)$$

$$\mathbf{w}^{(i+1)} = \mathbf{w}^{(i)} + \rho \left(\mathbf{G}\mathbf{d}^{(i+1)} - \mathbf{a}^{(i+1)} \odot e^{j\theta^{(i+1)}} \right). \quad (19)$$

The augmented Lagrangian function for each update is minimized for one or two variables while fixing the other variables.

To obtain the update rules of \mathbf{a} and θ , we modify the subproblem (17) as

$$\begin{aligned} & \{\mathbf{a}^{(i+1)}, \theta^{(i+1)}\} \\ & = \arg \min_{\mathbf{a} \geq 0, \theta} L_\rho(\mathbf{d}^{(i)}, \mathbf{a}, \theta, \mathbf{w}^{(i)}) \\ & = \arg \min_{\mathbf{a} \geq 0, \theta} \|\mathbf{a} - |\mathbf{u}^{\text{des}}|\|^2 + \lambda \|\mathbf{d}^{(i)}\|^2 \\ & \quad + \frac{\rho}{2} \|\mathbf{G}\mathbf{d}^{(i)} + \frac{\mathbf{w}^{(i)}}{\rho} - \mathbf{a} \odot e^{j\theta}\|^2 - \frac{\|\mathbf{w}^{(i)}\|^2}{2\rho} \\ & = \arg \min_{\mathbf{a} \geq 0, \theta} \|\mathbf{a} - |\mathbf{u}^{\text{des}}|\|^2 + \frac{\rho}{2} \|\mathbf{h}^{(i)} - \mathbf{a} \odot e^{j\theta}\|^2, \end{aligned} \quad (20)$$

where $\mathbf{h}^{(i)} = \mathbf{G}\mathbf{d}^{(i)} + \mathbf{w}^{(i)}/\rho$. In the second line, the terms involving $\mathbf{a} \odot e^{j\theta}$ are combined into one term. The terms not related to the optimization with respect to \mathbf{a} and θ are omitted in the last line. The phase θ is included only in the second term, and choosing $\theta^{(i+1)}$ according to (21) is optimal, regardless of \mathbf{a} .

$$\theta^{(i+1)} = \arg(\mathbf{h}^{(i)}). \quad (21)$$

Then, \mathbf{a} is updated by minimizing the quadratic of \mathbf{a} as

$$\begin{aligned} \mathbf{a}^{(i+1)} & = \arg \min_{\mathbf{a} \geq 0} \|\mathbf{a} - |\mathbf{u}^{\text{des}}|\|^2 + \frac{\rho}{2} \|\mathbf{h}^{(i)} - \mathbf{a}\|^2 \\ & = \frac{\rho |\mathbf{h}^{(i)}| + 2|\mathbf{u}^{\text{des}}|}{\rho + 2}. \end{aligned} \quad (22)$$

The update rule of \mathbf{d} is obtained by solving

$$\begin{aligned} \mathbf{d}^{(i+1)} & = \arg \min_{\mathbf{d}} L_\rho(\mathbf{d}, \mathbf{a}^{(i+1)}, \theta^{(i+1)}, \mathbf{w}^{(i)}) \\ & = \arg \min_{\mathbf{d}} \lambda \|\mathbf{d}\|^2 + \frac{\rho}{2} \left\| \mathbf{G}\mathbf{d} + \frac{\mathbf{w}^{(i)}}{\rho} - \mathbf{a}^{(i+1)} \odot e^{j\theta^{(i+1)}} \right\|^2, \end{aligned} \quad (23)$$

where the same modification as that in (20) is used. This quadratic of \mathbf{d} can also be simply minimized as

$$\begin{aligned} \mathbf{d}^{(i+1)} & = \left(\frac{2\lambda}{\rho} \mathbf{I} + \mathbf{G}^H \mathbf{G} \right)^{-1} \\ & \quad \mathbf{G}^H \left(\mathbf{a}^{(i+1)} \odot e^{j\theta^{(i+1)}} - \frac{\mathbf{w}^{(i)}}{\rho} \right). \end{aligned} \quad (24)$$

Algorithm 2: ADMM Algorithm for Amplitude Matching.

Input: $\mathbf{d}^{(0)}, \mathbf{w}^{(0)}, |\mathbf{u}^{\text{des}}|, \rho, \lambda, \mathbf{G}$

- 1: $i = 0$
- 2: **while** stopping criterion not satisfied **do**
- 3: $\mathbf{h}^{(i)} = \mathbf{G}\mathbf{d}^{(i)} + \mathbf{w}^{(i)}/\rho$
- 4: $\theta^{(i+1)} = \arg(\mathbf{h}^{(i)})$
- 5: $\mathbf{a}^{(i+1)} = (\rho|\mathbf{h}^{(i)}| + 2|\mathbf{u}^{\text{des}}|)/(\rho + 2)$
- 6: $\mathbf{d}^{(i+1)} = ((2\lambda/\rho)\mathbf{I} + \mathbf{G}^H\mathbf{G})^{-1}$
- 7: $\quad \cdot \mathbf{G}^H \left(\mathbf{a}^{(i+1)} \odot e^{j\theta^{(i+1)}} - \mathbf{w}^{(i)}/\rho \right)$
- 8: $\mathbf{w}^{(i+1)} = \mathbf{w}^{(i)} + \rho \left(\mathbf{G}\mathbf{d}^{(i+1)} - \mathbf{a}^{(i+1)} \odot e^{j\theta^{(i+1)}} \right)$
- 9: $i = i + 1$
- 10: **end while**
- 11: **return** $\mathbf{d}^{(i)}$

Note that the inverse matrix in (24) can be computed before the iteration.

We summarize the proposed ADMM-based amplitude matching algorithm in Algorithm 2. The computational cost of the proposed algorithm for each iteration is $O(LM)$, which is identical to that of the MM-based algorithm. Among the gradient methods, the computational cost of the gradient descent method is relatively small, which requires $O(LM)$ for computing the gradient (8) and the cost for the line search of the step-size parameter. However, its convergence speed is generally low. The quasi-Newton's methods, such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, are practically useful because of their fast convergence and low computational cost by approximating the inverse of the Hessian matrix. The computational cost for each iteration is $O(L^2 + LM)$ with the line search of the step-size parameter. Therefore, the proposed method is computationally efficient as the MM-based algorithm. Further investigation is given in the experiments.

V. DIFFERENTIAL-NORM PENALTY FOR TIME-DOMAIN FILTER DESIGN

In the previous section, the optimal driving signals for amplitude matching are obtained in the frequency domain. In practical implementations, the frequency-domain driving signals of target frequency bins are transformed into the time domain by inverse Fourier transform to construct a single-input-multiple-output (SIMO) finite impulse response (FIR) filter, where the input is the source signal for playback and the output is the driving signals of the secondary sources. The procedure to obtain broadband driving signals is equivalent to solving the following optimization problem:

$$\underset{\{\mathbf{d}_k\}_{k=1}^K}{\text{minimize}} \sum_{k=1}^K \left\| |\mathbf{G}_k \mathbf{d}_k| - |\mathbf{u}^{\text{des}}| \right\|^2 + \lambda \sum_{k=1}^K \|\mathbf{d}_k\|^2, \quad (25)$$

where the index of the frequency bin k ($\in \{1, \dots, K\}$) is introduced in the subscript for each frequency-dependent variable. Here, $|\mathbf{u}^{\text{des}}|$ is assumed to be independent of frequency to obtain a flat amplitude response filter for each control point. Since the loss term and the regularization term of ℓ_2 -norm for

each frequency are simply added, (25) can be solved at every frequency as in Section IV.

However, since the driving signals are independently determined for each frequency, the resulting solution can have discontinuities between frequencies. These discontinuities can lead to an unnecessarily large FIR filter length. To achieve amplitude matching with a small FIR filter length, we propose the use of a penalty term for the discontinuities of \mathbf{d}_k , which is called the differential-norm penalty. We define the differential norm as

$$D(\mathbf{d}_k) := \sum_{k=2}^K \|\mathbf{d}_k - \mathbf{d}_{k-1}\|^2 = \|\bar{\mathbf{F}}\bar{\mathbf{d}}\|^2, \quad (26)$$

where $\bar{\mathbf{d}} = [\mathbf{d}_1^T, \dots, \mathbf{d}_K^T]^T \in \mathbb{C}^{LK}$ and $\bar{\mathbf{F}} \in \mathbb{C}^{L(K-1) \times LK}$ is defined as

$$\bar{\mathbf{F}} := \begin{bmatrix} -1 & 0 & \dots & 0 & 1 & & \\ & -1 & 0 & & 0 & 1 & \mathbf{O} \\ & & \ddots & \ddots & & \ddots & \\ & & & \mathbf{O} & -1 & 0 & \dots & 0 & 1 \end{bmatrix}. \quad (27)$$

Thus, the optimization problem of the amplitude matching with the differential-norm penalty is described as

$$\underset{\{\mathbf{d}_k\}_{k=1}^K}{\text{minimize}} \sum_{k=1}^K \left\| |\mathbf{G}_k \mathbf{d}_k| - |\mathbf{u}^{\text{des}}| \right\|^2 + \lambda \|\bar{\mathbf{F}}\bar{\mathbf{d}}\|^2. \quad (28)$$

The optimization problem (28) can be solved by ADMM as in the single-frequency case. First, (28) is reformulated as

$$\underset{\bar{\mathbf{a}}, \bar{\mathbf{a}} \geq \mathbf{0}, \bar{\theta}}{\text{minimize}} \|\bar{\mathbf{a}} - |\bar{\mathbf{u}}^{\text{des}}|\|^2 + \lambda \|\bar{\mathbf{F}}\bar{\mathbf{d}}\|^2$$

subject to $\bar{\mathbf{G}}\bar{\mathbf{d}} = \bar{\mathbf{a}} \odot e^{j\bar{\theta}}, \quad (29)$

where $\bar{\mathbf{a}} = [\mathbf{a}_1^T, \dots, \mathbf{a}_K^T]^T \in \mathbb{R}_{\geq 0}^{MK}$, $\bar{\theta} = [\theta_1^T, \dots, \theta_K^T]^T \in \mathbb{R}^{MK}$, $|\bar{\mathbf{u}}^{\text{des}}| = [|\mathbf{u}^{\text{des}}|^T, \dots, |\mathbf{u}^{\text{des}}|^T]^T \in \mathbb{R}_{\geq 0}^{MK}$, and $\bar{\mathbf{G}} \in \mathbb{C}^{MK \times LK}$ is the block diagonal matrix of $\{\mathbf{G}_k\}_{k=1}^K$ as

$$\bar{\mathbf{G}} := \begin{bmatrix} \mathbf{G}_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \mathbf{G}_K \end{bmatrix}. \quad (30)$$

The augmented Lagrangian function is defined with the Lagrange multiplier $\bar{\mathbf{w}} = [\mathbf{w}_1^T, \dots, \mathbf{w}_K^T]^T \in \mathbb{C}^{MK}$ as

$$\begin{aligned} \bar{L}_\rho(\bar{\mathbf{d}}, \bar{\mathbf{a}}, \bar{\theta}, \bar{\mathbf{w}}) := & \|\bar{\mathbf{a}} - |\bar{\mathbf{u}}^{\text{des}}|\|^2 + \lambda \|\bar{\mathbf{F}}\bar{\mathbf{d}}\|^2 \\ & + \Re \left[\bar{\mathbf{w}}^H \left(\bar{\mathbf{G}}\bar{\mathbf{d}} - \bar{\mathbf{a}} \odot e^{j\bar{\theta}} \right) \right] \\ & + \frac{\rho}{2} \left\| \bar{\mathbf{G}}\bar{\mathbf{d}} - \bar{\mathbf{a}} \odot e^{j\bar{\theta}} \right\|^2. \end{aligned} \quad (31)$$

The update rule of each variable is obtained as

$$\bar{\theta}^{(i+1)} = \arg(\bar{\mathbf{h}}^{(i)}), \quad (32)$$

$$\bar{\mathbf{a}}^{(i+1)} = \frac{\rho|\bar{\mathbf{h}}^{(i)}| + 2|\bar{\mathbf{u}}^{\text{des}}|}{\rho + 2}, \quad (33)$$

$$\bar{\mathbf{d}}^{(i+1)} = \mathbf{M}^{-1} \bar{\mathbf{G}}^H \left(\bar{\mathbf{a}}^{(i+1)} \odot e^{j\bar{\theta}^{(i+1)}} - \frac{\bar{\mathbf{w}}^{(i)}}{\rho} \right), \quad (34)$$

$$\bar{\mathbf{w}}^{(i+1)} = \bar{\mathbf{w}}^{(i)} + \rho \left(\bar{\mathbf{G}}\bar{\mathbf{d}}^{(i+1)} - \bar{\mathbf{a}}^{(i+1)} \odot e^{j\bar{\boldsymbol{\theta}}^{(i+1)}} \right), \quad (35)$$

where $\bar{\mathbf{h}}^{(i)} = \bar{\mathbf{G}}\bar{\mathbf{d}}^{(i)} + \bar{\mathbf{w}}^{(i)}/\rho$ ($\in \mathbb{C}^{MK}$), and $\mathbf{M} \in \mathbb{C}^{LK \times LK}$ is defined as

$$\mathbf{M} = \begin{bmatrix} M_{11} & \dots & M_{K1} \\ \vdots & \ddots & \vdots \\ M_{1K} & \dots & M_{KK} \end{bmatrix}, \quad (36)$$

with

$$M_{ij} = \begin{cases} \mathbf{G}_i^H \mathbf{G}_i + \frac{2\lambda}{\rho} \mathbf{I} & i = j = 1, K \\ \mathbf{G}_i^H \mathbf{G}_i + \frac{4\lambda}{\rho} \mathbf{I} & i = j \neq 1, K \\ -\frac{2\lambda}{\rho} \mathbf{I} & i = j + 1, j - 1 \\ \mathbf{O} & \text{otherwise} \end{cases}. \quad (37)$$

The update rules of $\bar{\boldsymbol{\theta}}$, $\bar{\mathbf{a}}$, and $\bar{\mathbf{w}}$ are efficiently computed. On the other hand, the update of $\bar{\mathbf{d}}$ in (34) requires large computational complexity owing to \mathbf{M}^{-1} . Since the computation of $\mathbf{M}^{-1}\bar{\mathbf{G}}^H$ remains unchanged for each iteration, which can be performed before the iteration, which requires $O((M+L)L^2K^3)$, then the computational cost of the update of $\bar{\mathbf{d}}$ becomes $O(LMK^2)$. However, these computational complexities can be reduced by using the property of the block tridiagonal matrix of \mathbf{M} . We apply the block birecurrence method [29] in updating $\bar{\mathbf{d}}$, which reduces the computational cost to $O(L^3K)$ before the iteration and $O((M+L)LK)$ for each iteration. Furthermore, the memory requirements can also be reduced from $O((M+L)LK^2)$ to $O((M+L)LK)$ before the iteration and from $O(MLK^2)$ to $O((M+L)LK)$ at each iteration. Since K is generally much larger than L and M , these reductions of temporal and spatial complexities are significant. See Appendix B for the detailed procedure of the block birecurrence method.

VI. EXPERIMENTS

We conducted experiments to evaluate our proposed method. First, experimental results in the narrowband case are shown to compare the performance of the amplitude matching algorithms. Next, the broadband case is investigated to validate the effectiveness of the differential norm. Finally, we demonstrate experimental results using real data.

A. Narrowband Case

We conducted numerical experiments using 2D free-field simulations in the frequency domain, focusing on the comparison of the amplitude matching (AM) algorithms described in Section III with the conventional pressure matching (PM) and acoustic contrast control (ACC). As shown in Fig. 2, a circular loudspeaker array was placed, whose radius was 1.5 m and the center was at the origin. The array consisted of 48 loudspeakers that were equiangularly aligned. Each loudspeaker was assumed to be the 2D point source, whose transfer function is defined as

$$g_l(\mathbf{r}) = \frac{j}{4} H_0^{(1)}(k\|\mathbf{r} - \mathbf{r}_l\|), \quad (38)$$

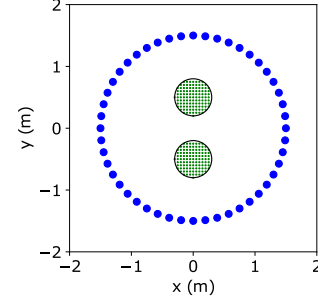


Fig. 2. Experimental setting for narrowband and broadband cases. Blue and green dots indicate loudspeakers and control points, respectively.

where $H_0^{(1)}(\cdot)$ is the 0th-order Hankel function of the first kind and \mathbf{r}_l is the position of the l th loudspeaker. Two circular target regions were set inside the loudspeaker array. The radii of the two regions were both 0.3 m, and their centers were at (0.0, 0.5) m and (0.0, -0.5) m. The upper and lower target regions are denoted as Ω_U and Ω_L , respectively. The target regions were discretized into 111 control points at intervals of 0.05 m. For the desired amplitude distributions, we set 1.0 in Ω_U and 0.0 in Ω_L ; therefore, Ω_U and Ω_L are supposed to be acoustic bright and dark zones, respectively. Hence, ACC was applied to synthesizing bright and dark zones in each region. The amplitude of the driving signal for ACC was determined so that the average amplitude inside Ω_U becomes 1.0.

To solve the optimization problem of AM, we compared the proposed ADMM-based algorithm (AM-ADMM), MM-based algorithm (AM-MM), and gradient method using BFGS (AM-BFGS). PM and ACC were also evaluated. Since the experiments were conducted in the frequency domain, we used the frequency-domain ADMM algorithm (Algorithm 2) in this section. The initial value of \mathbf{d} for the three AM algorithms was the solution of PM. The regularization parameter λ in (7) for the AM algorithms and γ in (5) for ACC were set as $\sigma_{\max}(\mathbf{G}) \times 10^{-3}$, where $\sigma_{\max}(\mathbf{G})$ means the largest singular value of \mathbf{G} . The parameter ρ in Algorithm 2 was set to 1.0.

As an evaluation measure, we define the mean square error (MSE) between synthesized and desired amplitude distributions as

$$\text{MSE}(\omega) = 10 \log_{10} \left(\frac{1}{M} \|\mathbf{u}^{\text{syn}}(\omega) - \mathbf{u}^{\text{des}}(\omega)\|^2 \right), \quad (39)$$

where $|\mathbf{u}^{\text{syn}}(\omega)|$ is the amplitude distribution of the synthesized sound field. We also evaluate with acoustic contrast (AC) defined as

$$\text{AC}(\omega) = 10 \log_{10} \frac{\|\mathbf{u}_{\Omega_U}^{\text{syn}}(\omega)\|^2}{\|\mathbf{u}_{\Omega_L}^{\text{syn}}(\omega)\|^2}, \quad (40)$$

where $\mathbf{u}_{\Omega_U}^{\text{syn}}$ and $\mathbf{u}_{\Omega_L}^{\text{syn}}$ are the synthesized pressure in Ω_U and Ω_L , respectively. The experiments were conducted using a 3.00 GHz 8-Core Intel Core i7-9700F processor and implemented with Python.

An appropriate setting of \mathbf{u}^{des} in (3), i.e., the desired phase distribution, for PM is not obvious, especially when the desired

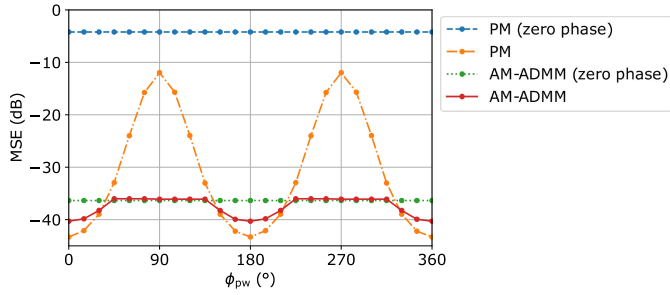


Fig. 3. MSE with respect to plane wave direction ϕ_{pw} of the desired field in Ω_U at 1400 Hz for narrowband case. Each PM solution was used as the initial value of AM-ADMM. The case when \mathbf{u}^{des} is set as the desired amplitude distribution with zero phase is also shown, which is denoted as PM (zero phase) and AM-ADMM (zero phase).

amplitude distribution and arrangement of the target regions are complicated, e.g., the case of non-uniform $|\mathbf{u}^{des}|$. However, in this setting, it can be inferred that the desired sound field in Ω_U for PM should be a plane wave propagating in the direction of 0° or 180° . The setting of \mathbf{u}^{des} for PM also has an impact on the AM methods because the solution of PM is used for the initial value in this experiment. Fig. 3 shows the MSE of PM with respect to the plane wave direction ϕ_{pw} of the desired sound field and AM-ADMM with the initial value of each PM solution. The case that \mathbf{u}^{des} is set as the desired amplitude distribution with zero phase, i.e., 1.0 in Ω_U and 0.0 in Ω_L , was also investigated, which is denoted as PM (zero phase) and AM-ADMM (zero phase). Although synthesizing a uniform pressure distribution of zero phase is impractical, this setting will be useful for initial value of the AM methods when an appropriate setting of \mathbf{u}^{des} is difficult. Note that the results of PM (zero phase) and AM-ADMM (zero phase) are not dependent on ϕ_{pw} . As expected, the lowest MSE of PM was achieved at $\phi_{pw} = 0^\circ$ and 180° , whose value was -43.3 dB. The MSE of PM (zero phase) was -4.2 dB, which was larger than the highest MSE of PM (-12.0 dB at $\phi_{pw} = 90^\circ$ and 270°). In AM-ADMM, the lowest MSE was -40.2 dB at $\phi_{pw} = 0^\circ$ and 180° , which was slightly larger than that of PM at the same ϕ_{pw} . However, the MSE of AM-ADMM was maintained at around -36 dB for $\phi_{pw} \in [45^\circ, 135^\circ]$ and $[225^\circ, 315^\circ]$, which was close to the MSE of AM-ADMM (zero phase), i.e., -36.4 dB. AM-ADMM attained low MSE even when the initial value was the PM solution with the high MSE value; therefore, PM (zero phase) is a reasonable compromise for the initial value of the AM methods.

Table I shows the MSE and AC of each method at 1400 Hz. PM ($\phi_{pw} = 0^\circ$) and AM-ADMM ($\phi_{pw} = 0^\circ$) indicate PM with the desired field of the plane wave with the optimal angle $\phi_{pw} = 0^\circ$ and AM-ADMM with the initial value of the solution of PM ($\phi_{pw} = 0^\circ$), respectively. In the other AM methods, PM (zero phase) was used for the initial value. The synthesized pressure distribution of each method is shown in Fig. 4. In ACC, the highest AC was achieved with a low acoustic potential energy in Ω_L , but the amplitude distribution in Ω_U was not uniform. Thus, even though the MSE of ACC was very low in Ω_L , that of ACC was high in Ω_U . The plane-wave-like pressure distribution was synthesized by AM-ADMM. In AM-MM and AM-BFGS,

TABLE I
MSE AND AC IN dB IN NARROWBAND CASE AT 1400 HZ

	MSE	MSE in Ω_U	MSE in Ω_L	AC
PM ($\phi_{pw} = 0^\circ$)	-43.3	-41.0	-48.8	48.8
AM-ADMM ($\phi_{pw} = 0^\circ$)	-40.3	-37.5	-50.3	50.3
PM (zero phase)	-4.2	-1.2	-30.6	14.5
ACC	-6.4	-3.4	-67.9	67.9
AM-BFGS	-16.0	-13.1	-35.3	35.1
AM-MM	-15.3	-12.3	-35.5	35.2
AM-ADMM	-36.4	-34.6	-39.4	39.3

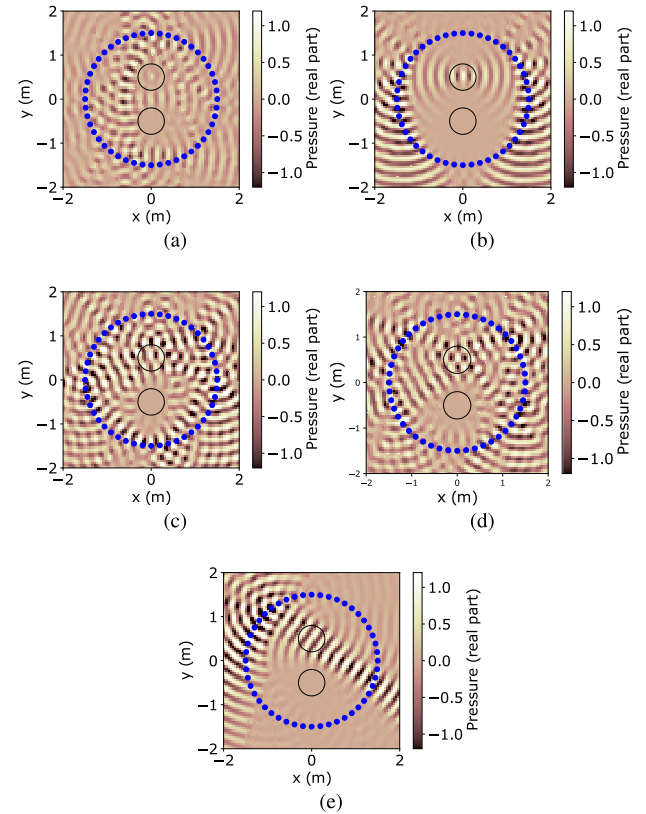


Fig. 4. Synthesized pressure distribution in narrowband case at 1400 Hz. (a) PM (zero phase); (b) ACC; (c) AM-BFGS; (d) AM-MM; (e) AM-ADMM.

the synthesized pressure distributions were still close to that in PM (zero phase), i.e., the initial value. The total MSE was the lowest in AM-ADMM among the methods not using the optimal plane wave angle $\phi_{pw} = 0^\circ$.

Figs. 5 and 6 show the MSE and AC of each method with respect to the frequency from 100 to 4000 Hz at intervals of 50 Hz. Since the optimization problem of AM was solved for each frequency, the MSEs of the AM algorithms largely varied at each frequency. However, they were significantly lower than those of PM (zero phase) and ACC. In contrast, the highest AC was achieved by ACC because this measure corresponds to the objective function of ACC. Among the three AM algorithms, the MSEs of AM-MM were relatively higher than those of AM-ADMM and AM-BFGS, especially at the frequencies below 1400 Hz. This means that AM-MM did not reach the optimal

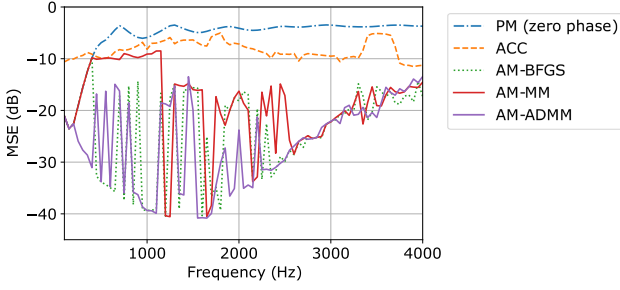


Fig. 5. MSE with respect to frequency for narrowband case.

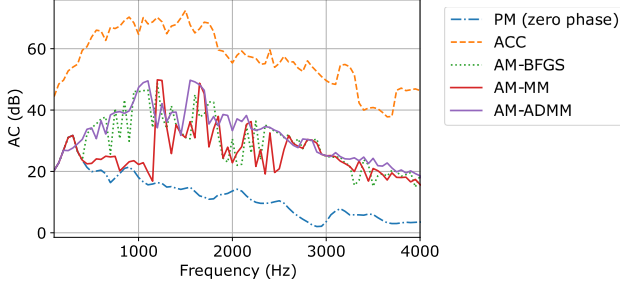
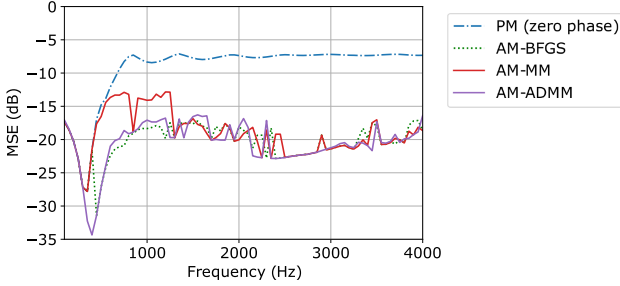


Fig. 6. AC with respect to frequency for narrowband case.


 Fig. 7. MSE with respect to frequency for narrowband case when truncated Gaussian function is set for the desired amplitude distribution in Ω_{UL} .

solution owing to its initial value dependence at these frequencies. The performance of AM-ADMM was comparable to that of AM-BFGS, but high MSEs appeared at several frequencies. Nevertheless, their general performance was better than that of the other methods.

Next, we investigated the case when the desired amplitude distribution is not uniform. The desired amplitude $|\mathbf{u}^{\text{des}}|$ in Ω_U is set to be a truncated Gaussian function with its mean at the center of Ω_U and variance 0.04. The desired field \mathbf{u}^{des} for PM is difficult to set, and ACC is not applicable to synthesize a specific phase distribution; therefore, PM (zero phase) and AM methods were compared. In PM (zero phase), the desired pressure distribution was set to be $|\mathbf{u}^{\text{des}}|$ with zero phase. Fig. 7 shows the MSE of each frequency. The desired amplitude distribution was accurately synthesized by the AM methods although their general MSEs were relatively high, compared to those in Fig. 5.

Finally, we demonstrate the case that the desired amplitude distributions are more complex. We set three circular target regions of 0.3 m radius with their centers at $(-0.2\sqrt{3}, 0.2)$ m, $(0.2\sqrt{3}, 0.2)$ m, and $(0.0, -0.4)$ m. We denote these upper left,

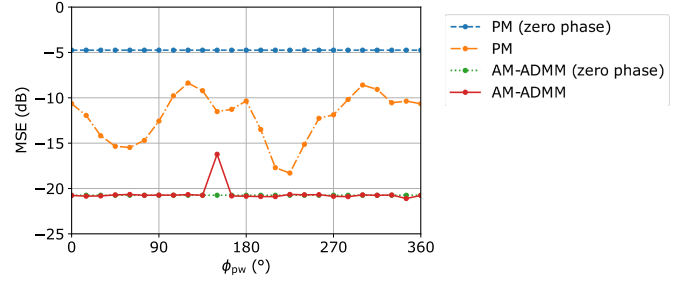

 Fig. 8. MSE with respect to ϕ_{pw} at 1400 Hz for narrowband case when three target regions were set, where ϕ_{pw} denotes the plane wave angle of the desired field in Ω_{UL} . Each PM solution was used as the initial value of AM-ADMM. The case when \mathbf{u}^{des} is set as the desired amplitude distribution with zero phase is also shown, which is denoted as PM (zero phase) and AM-ADMM (zero phase).

 TABLE II
 MSE IN dB FOR NARROWBAND CASE AT 1400 HZ WHEN THREE TARGET REGIONS WERE SET

	MSE	MSE in Ω_{UL}	MSE in Ω_{UR}	MSE in Ω_{LC}
PM ($\phi_{pw} = 225^\circ$)	-18.3	-23.1	-17.7	-16.5
AM-ADMM ($\phi_{pw} = 225^\circ$)	-20.7	-22.4	-18.0	-23.5
PM (zero phase)	-4.7	-0.9	-7.0	-27.0
ACC	-6.6	-2.7	-8.9	-51.7
AM-BFGS	-21.5	-21.3	-20.8	-22.4
AM-MM	-17.2	-13.8	-20.4	-21.6
AM-ADMM	-20.7	-22.7	-18.0	-24.0

upper right, and lower center target regions as Ω_{UL} , Ω_{UR} , and Ω_{LC} , respectively. The desired amplitude distributions were uniform: 1.0 in Ω_{UL} , 0.5 in Ω_{UR} , and 0.0 in Ω_{LC} . ACC was applied by separately obtaining the driving signals for synthesizing a single bright zone in Ω_{UL} or Ω_{UR} ; then, these signals were normalized and summed. Since the desired amplitude distributions are uniform, plane wave fields will be a reasonable setting for the desired field of PM; however, it is difficult to infer physically feasible plane wave angles because of complex arrangement of the target regions. Fig. 8 shows the MSE with respect to the plane wave angle ϕ_{pw} in Ω_{UL} at 1400 Hz when the desired field in Ω_{UR} was fixed to be a plane wave field of $\phi_{pw} = 210^\circ$. The direction $\phi_{pw} = 210^\circ$ in Ω_{UR} was taken from the propagation direction of the sound field synthesized by AM-ADMM. The MSE of AM-ADMM was lower than that of PM for all ϕ_{pw} . AM-ADMM (zero phase) attained almost the same or even lower MSE of AM-ADMM. The MSE of each method at 1400 Hz is shown in Table II. The MSE with respect to frequency is shown in Fig. 9. Since AC cannot be defined for three regions, we only evaluated with MSE. Again, the lowest MSE was achieved by AM-ADMM and AM-BFGS at most of frequencies. The synthesized pressure distributions at 1400 Hz are shown in Fig. 10. These results indicate that our proposed algorithm performs well even when the desired amplitude distributions are complex.

Fig. 11 shows the cost function value of the AM algorithms in terms of computation time. The decreasing speed of the cost function was high in AM-MM and AM-ADMM, but it stagnated at around 6.7 in AM-MM. The cost function decreased to 3.2 in

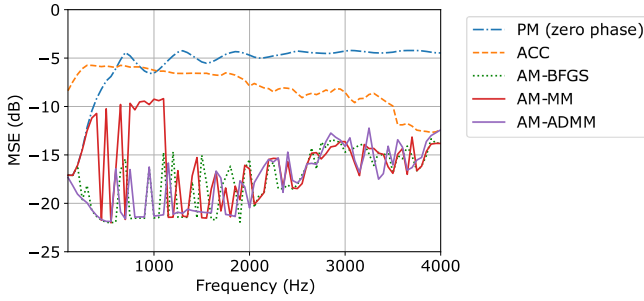


Fig. 9. MSE with respect to frequency for narrowband case when three target regions were set.

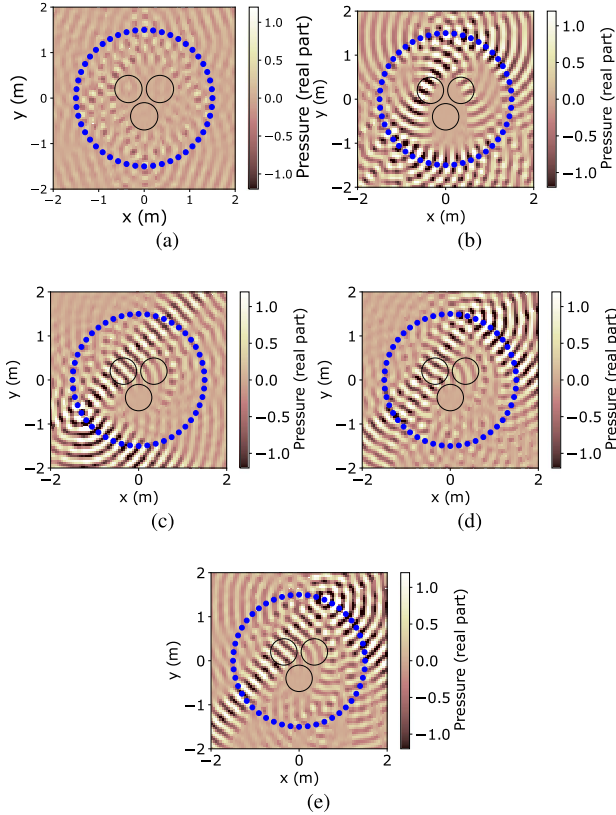


Fig. 10. Synthesized pressure distribution of three AM methods in narrowband case at 1400 Hz when three target regions were set. (a) PM (zero phase); (b) ACC; (c) AM-BFGS; (d) AM-MM; (e) AM-ADMM.

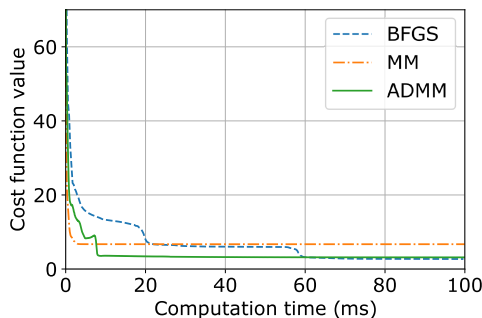


Fig. 11. Cost function value with respect to computation time. Three AM algorithms were compared.

TABLE III
MSE AND AC IN dB IN BROADBAND CASE

	MSE	MSE in Ω_U	MSE in Ω_L	AC
AM w/ Diff.-norm	-13.2	-10.9	-18.1	17.3
AM w/ ℓ_2 -norm	-9.3	-6.6	-17.6	14.6

AM-ADMM and 2.7 in AM-BFGS. The computation time until the cost function reached 3.5 was 56.8 ms in AM-BFGS and 7.9 ms in AM-ADMM. A similar tendency can be seen at other frequencies. Therefore, the proposed ADMM algorithm is effective both in control accuracy and computational time.

B. Broadband Case

Next, we show experimental results in the broadband case, focusing on the effectiveness validation of the differential-norm penalty for AM. Again, numerical simulation is conducted under the 2D free-field assumption. The placements of loudspeakers, target regions, and control points, and the settings of the desired amplitude were the same as those in Section VI-A (see Fig. 2).

We compare AM with the differential-norm penalty (28) and ℓ_2 -norm regularization (25), which are denoted as AM w/ Diff.-norm and AM w/ ℓ_2 -norm, respectively. The parameter λ in (28) was set to 10.0, and the other parameters were the same as those in AM-ADMM in the narrowband case. After generating the frequency-domain driving signals using the ADMM-based algorithm, time-domain FIR filters were obtained by their inverse fast Fourier transform (FFT). The sampling frequency was 16 kHz, and the number of frequency bins, i.e., FFT length, was 32768. For evaluation measures in the broadband case, we define average MSE and AC for all the frequency bins as

$$\overline{\text{MSE}} = 10 \log_{10} \left(\frac{1}{F} \sum_f \frac{1}{M} \left\| |\mathbf{u}^{\text{syn}}(\omega_f)| - |\mathbf{u}^{\text{des}}(\omega_f)| \right\|^2 \right) \quad (41)$$

$$\overline{\text{AC}} = 10 \log_{10} \left(\frac{\sum_f \|\mathbf{u}_{\Omega_B}^{\text{syn}}(\omega_f)\|^2}{\sum_f \|\mathbf{u}_{\Omega_D}^{\text{syn}}(\omega_f)\|^2} \right), \quad (42)$$

where F is the number of frequency bins, f is its index, and $\mathbf{u}_{\Omega_B}^{\text{syn}}$ and $\mathbf{u}_{\Omega_D}^{\text{syn}}$ are synthesized pressure in bright and dark zones, respectively. The synthesized pressures in the time domain at the evaluation points were transformed into the frequency domain to compute these evaluation measures.

Fig. 12 shows the FIR filters of AM w/ Diff.-norm and ℓ_2 -norm before truncating the filter length. Apparently, the decrease in the filter amplitude from the peak of the AM w/ Diff.-norm is much sharper than that of ℓ_2 -norm. When the optimal filter length is defined so that the decrease from the peak amplitude is 30 dB, that of AM w/ ℓ_2 -norm was 12888. On the other hand, the optimal filter length of AM w/ Diff.-norm was 162. The magnitude and phase responses of the time-domain filters of AM w/ ℓ_2 -norm and AM w/ Diff.-norm are shown in Figs. 13 and 14, respectively. The discontinuities in the frequency response were suppressed by the differential norm penalty.

Table III shows the $\overline{\text{MSE}}$ and $\overline{\text{AC}}$ of each method when the filter length is truncated to 220. This filter length is defined so

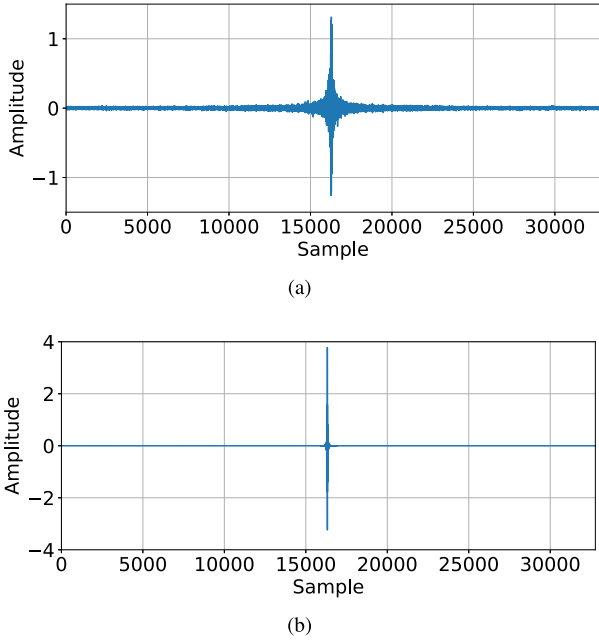


Fig. 12. Time-domain filters generated by (a) AM w/ ℓ_2 -norm and (b) AM w/ Diff.-norm.

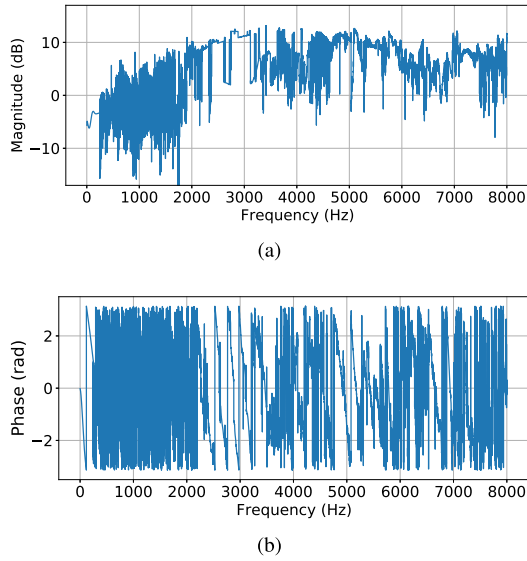


Fig. 13. (a) Magnitude and (b) phase responses of time-domain filter of AM w/ ℓ_2 -norm.

that the decrease in filter amplitude from the peak is 40 dB in AM w/ Diff.-norm. MSE in the frequency domain is also plotted in Fig. 15. MSE of AM w/ Diff.-norm was particularly lower than that of ℓ_2 -norm inside Ω_U . The frequency-domain MSE of AM w/ ℓ_2 -norm largely fluctuated for each frequency bin at high frequencies, whereas that of Diff.-norm was relatively smooth.

C. Experiments Using Real Data

To investigate the performance of the proposed method in a practical environment, we conducted experiments using the

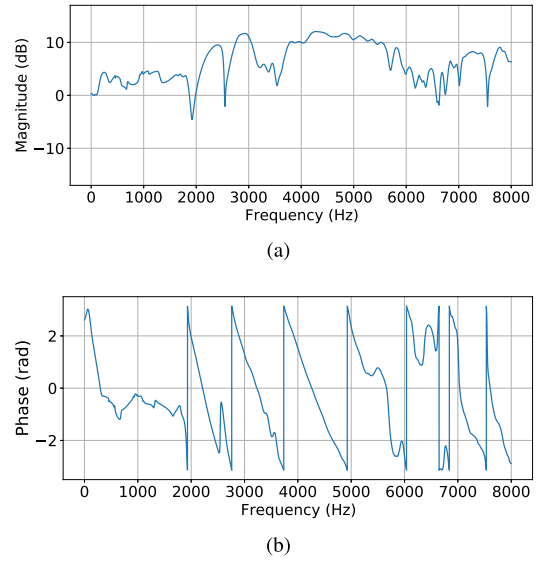


Fig. 14. (a) Magnitude and (b) phase responses of time-domain filter of AM w/ Diff.-norm.

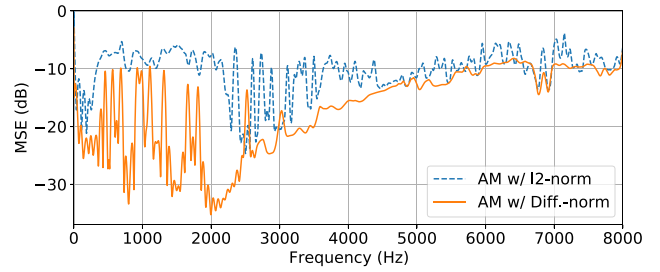


Fig. 15. MSE with respect to frequency in broadband case.

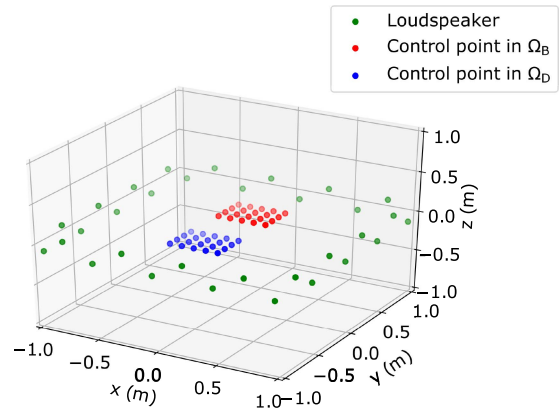


Fig. 16. Placements of loudspeakers and control points in the experiments using real data. Green, red, and blue dots indicate loudspeakers, control points in Ω_B , and those in Ω_D , respectively.

impulse response dataset presented in [30]. Configurations of loudspeakers and control points are shown in Fig. 16. We set two 2D square target regions of acoustic bright and dark zones, Ω_B and Ω_D , with dimensions of 0.4 m \times 0.3 m on the xy -plane at $z = 0$, whose centers were at (0.0 m, 0.4 m, 0.0 m) and (0.0 m, -0.4 m, 0.0 m), respectively. The loudspeakers were

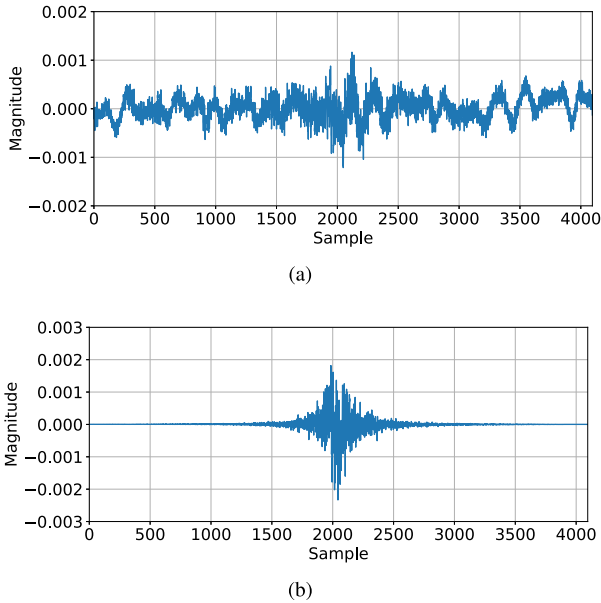


Fig. 17. Time-domain filters generated by (a) AM w/ ℓ_2 -norm and (b) AM w/ Diff.-norm when using real data.

TABLE IV
MSE AND AC IN dB FOR EXPERIMENTS USING REAL DATA

	MSE	MSE in Ω_B	MSE in Ω_D	AC
PM (zero phase)	-9.2	-9.0	-9.5	7.3
ACC	-7.9	-5.1	-17.7	13.8
AM w/ ℓ_2 norm	-12.9	-11.1	-15.8	10.6
AM w/ Diff. norm	-19.1	-18.4	-19.9	18.4

regularly arranged along the borders of two squares with dimensions of $2.0 \text{ m} \times 2.0 \text{ m}$ at $z = \pm 0.1 \text{ m}$. Twenty control points were obtained by discretizing each target region every 0.1 m . The desired amplitudes were 1.0 in Ω_B and 0.0 in Ω_D and the generated signal was a band-limited pulse signal from 100 Hz to 1000 Hz . $\overline{\text{MSE}}$ and $\overline{\text{AC}}$ were evaluated at the positions of 0.05 m intervals inside the target regions. The reverberation time T_{60} was 0.19 s .

First, the filters generated by AM w/ ℓ_2 -norm and Diff.-norm are compared. The parameters for ADMM-based algorithms were the same as those in Section VI-B. Fig. 17 shows the filters before truncation obtained by AM w/ ℓ_2 -norm and Diff.-norm. Since the impulse responses of the loudspeakers include reverberation, the difference in the necessary filter length for these two methods was much larger than that in Section VI-B.

Table IV shows $\overline{\text{MSE}}$ and $\overline{\text{AC}}$ of each method. The proposed AM w/ Diff.-norm achieved the lowest $\overline{\text{MSE}}$ and highest $\overline{\text{AC}}$ among the four methods. Figs. 18 and 19 show the MSE and AC of each frequency, where the plots were smoothed by moving average of 20 frequency bins. The magnitude distributions of the synthesized sound field are shown in Fig. 20. In Ω_D , the magnitude distribution of the proposed method was low even at the positions other than the control points, compared with those of the other methods.

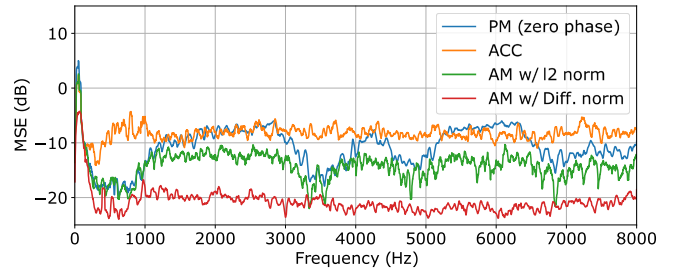


Fig. 18. MSE with respect to frequency when using real data.

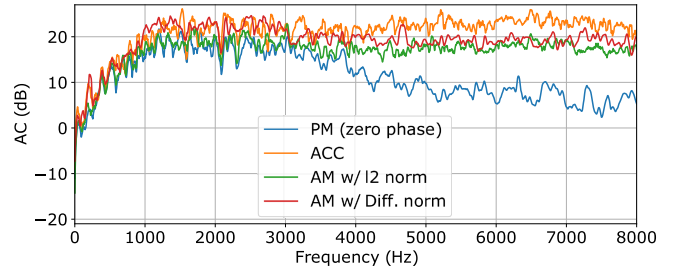


Fig. 19. AC with respect to frequency when using real data.

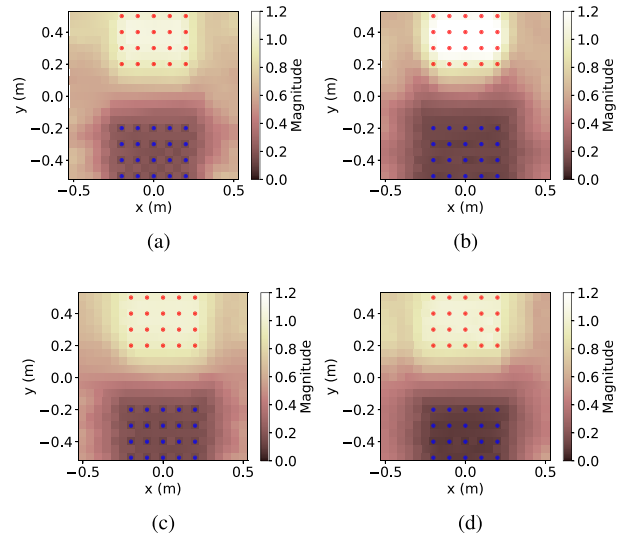


Fig. 20. Magnitude distributions of synthesized sound fields when using real data. Red and blue dots indicate control points in Ω_B and Ω_D , respectively. (a) PM (zero phase); (b) ACC; (c) AM w/ ℓ_2 -norm; (d) AM w/ Diff.-norm.

VII. CONCLUSION

We proposed the amplitude matching method for multizone sound field control. The amplitude matching aims to synthesize a desired amplitude distribution over a target region, whereas the phase distribution is arbitrary, using multiple secondary sources. Since the cost function of amplitude matching is neither linear, convex, nor differentiable, it is important to develop efficient algorithms for solving the amplitude matching problem. An algorithm based on ADMM is formulated both in the frequency and time domains. To avoid an unnecessarily large time-domain filter length in the broadband case, the differential-norm penalty for adjacent frequency bins is also introduced. In the experiments,

Algorithm 3: ADMM Algorithm for Amplitude Matching With Differential-Norm Penalty.

Input: $\bar{\mathbf{d}}^{(0)}$, $\bar{\mathbf{w}}^{(0)}$, $|\bar{\mathbf{u}}^{\text{des}}|$, $\bar{\mathbf{G}}$, ρ , λ , k_b
 1: Initialize \mathbf{A}_k and $\mathbf{b}_k^{(0)}$ by (49)–(54)
 2: $i = 0$
 3: **while** stopping criterion not satisfied **do**
 4: $\bar{\mathbf{h}}^{(i)} = \bar{\mathbf{G}}\bar{\mathbf{d}}^{(i)} + \bar{\mathbf{w}}^{(i)}/\rho$
 5: $\bar{\theta}^{(i+1)} = \arg(\bar{\mathbf{h}}^{(i)})$
 6: $\bar{\mathbf{a}}^{(i+1)} = (\rho|\bar{\mathbf{h}}^{(i)}| + 2|\bar{\mathbf{u}}^{\text{des}}|)/(\rho + 2)$
 7: $\mathbf{p}_k^{(i+1)} = \mathbf{G}_k^H (\mathbf{a}_k^{(i+1)} \odot e^{j\theta_k^{(i+1)}} - \mathbf{w}_k^{(i)}/\rho)$
 8: Update \mathbf{b}_k by (50), (52), and (54)
 9: Update \mathbf{d}_k by (55)–(58)
 10: $\bar{\mathbf{w}}^{(i+1)} = \bar{\mathbf{w}}^{(i)} + \rho (\bar{\mathbf{G}}\bar{\mathbf{d}}^{(i+1)} - \bar{\mathbf{a}}^{(i+1)} \odot e^{j\bar{\theta}^{(i+1)}})$
 11: $i = i + 1$
 12: **end while**
 13: **return** $\bar{\mathbf{d}}^{(i)}$

it was shown that the proposed method enables us to accurately and efficiently synthesize the desired amplitude distributions, compared with current multizone sound field control methods.

APPENDIX A

DERIVATION OF MM ALGORITHM FOR AMPLITUDE MATCHING

By using

$$(\mathbf{G}\mathbf{d})^H \mathbf{v}^{(i)} = |\mathbf{G}\mathbf{d}|^T |\mathbf{u}^{\text{des}}| \odot e^{j(\arg(\mathbf{G}\mathbf{d}^{(i)}) - \arg(\mathbf{G}\mathbf{d}))}, \quad (43)$$

we obtain the following inequality:

$$\begin{aligned} & (\mathbf{G}\mathbf{d})^H \mathbf{v}^{(i)} + \mathbf{v}^{(i)H} \mathbf{G}\mathbf{d} \\ &= 2|\mathbf{G}\mathbf{d}|^T |\mathbf{u}^{\text{des}}| \odot \cos(\arg(\mathbf{G}\mathbf{d}^{(i)}) - \arg(\mathbf{G}\mathbf{d})) \\ &\leq 2|\mathbf{G}\mathbf{d}|^T |\mathbf{u}^{\text{des}}|, \end{aligned} \quad (44)$$

where the equality holds for $\mathbf{d} = \mathbf{d}^{(i)}$. Therefore,

$$\begin{aligned} & \left\| |\mathbf{G}\mathbf{d}| - |\mathbf{u}^{\text{des}}| \right\|^2 = \|\mathbf{G}\mathbf{d}\|^2 - 2|\mathbf{G}\mathbf{d}|^T |\mathbf{u}^{\text{des}}| + \|\mathbf{u}^{\text{des}}\|^2 \\ &\leq \|\mathbf{G}\mathbf{d}\|^2 - \mathbf{d}^H \mathbf{G}^H \mathbf{v}^{(i)} - \mathbf{v}^{(i)H} \mathbf{G}\mathbf{d} + \|\mathbf{u}^{\text{des}}\|^2 \\ &= \left\| \mathbf{G}\mathbf{d} - \mathbf{v}^{(i)} \right\|^2. \end{aligned} \quad (45)$$

Thus, we obtain the surrogate function of $J(\mathbf{d})$ as

$$J^+(\mathbf{d}|\mathbf{v}^{(i)}) = \left\| \mathbf{G}\mathbf{d} - \mathbf{v}^{(i)} \right\|^2 + \lambda \|\mathbf{d}\|^2. \quad (46)$$

APPENDIX B

BLOCK BIRECURRENCE METHOD FOR AMPLITUDE MATCHING WITH DIFFERENTIAL-NORM PENALTY

We apply the block birecurrence method [29] to solve (34) in ADMM algorithm for amplitude matching with the differential-norm penalty. First, (34) is rewritten as the block tridiagonal

simultaneous linear equations as

$$\mathbf{M} \begin{bmatrix} \mathbf{d}_1^{(i+1)} \\ \mathbf{d}_2^{(i+1)} \\ \vdots \\ \mathbf{d}_K^{(i+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^{(i+1)} \\ \mathbf{p}_2^{(i+1)} \\ \vdots \\ \mathbf{p}_K^{(i+1)} \end{bmatrix}, \quad (47)$$

where $\mathbf{p}_k^{(i+1)}$ is defined as

$$\mathbf{p}_k^{(i+1)} := \mathbf{G}_k^H \left(\mathbf{a}_k^{(i+1)} \odot e^{j\theta_k^{(i+1)}} - \frac{\mathbf{w}_k^{(i)}}{\rho} \right). \quad (48)$$

We define $\{\mathbf{A}_k\}_{k=1}^K$ and $\{\mathbf{b}_k\}_{k=1}^K$, which are separately computed for $k = 1$, $k \in [2, k_b]$, $k \in [k_b + 1, K - 1]$, and $k = K$ with the integer k_b (*balancer*). Their update rules are written as

$$\mathbf{A}_k = \frac{2\lambda}{\rho} \left(\mathbf{G}_k^H \mathbf{G}_k + \frac{2\lambda}{\rho} \mathbf{I} \right)^{-1}, \quad (49)$$

$$\mathbf{b}_k^{(i+1)} = \left(\mathbf{G}_k^H \mathbf{G}_k + \frac{2\lambda}{\rho} \mathbf{I} \right)^{-1} \mathbf{p}_k^{(i+1)}, \quad (50)$$

for $k = 1$ and K ,

$$\mathbf{A}_k = \frac{2\lambda}{\rho} \left(\mathbf{G}_k^H \mathbf{G}_k + \frac{4\lambda}{\rho} \mathbf{I} - \frac{2\lambda}{\rho} \mathbf{A}_{k-1} \right)^{-1}, \quad (51)$$

$$\begin{aligned} \mathbf{b}_k^{(i+1)} &= \left(\mathbf{G}_k^H \mathbf{G}_k + \frac{4\lambda}{\rho} \mathbf{I} - \frac{2\lambda}{\rho} \mathbf{A}_{k-1} \right)^{-1} \\ &\times \left(\mathbf{p}_k^{(i+1)} + \frac{2\lambda}{\rho} \mathbf{b}_{k-1}^{(i+1)} \right), \end{aligned} \quad (52)$$

for $k = 2, \dots, k_b$, and

$$\mathbf{A}_k = \frac{2\lambda}{\rho} \left(\mathbf{G}_k^H \mathbf{G}_k + \frac{4\lambda}{\rho} \mathbf{I} - \frac{2\lambda}{\rho} \mathbf{A}_{k+1} \right)^{-1}, \quad (53)$$

$$\begin{aligned} \mathbf{b}_k^{(i+1)} &= \left(\mathbf{G}_k^H \mathbf{G}_k + \frac{4\lambda}{\rho} \mathbf{I} - \frac{2\lambda}{\rho} \mathbf{A}_{k+1} \right)^{-1} \left(\mathbf{p}_k^{(i+1)} + \frac{2\lambda}{\rho} \mathbf{b}_{k+1}^{(i+1)} \right), \end{aligned} \quad (54)$$

for $k = K - 1, K - 2, \dots, k_b + 1$. We set k_b to $K/2$ in the experiment.

Then, the recursive update rule of $\{\mathbf{d}_k\}_{k=1}^K$ is obtained as

$$\mathbf{d}_{k_b}^{(i+1)} = (\mathbf{I} - \mathbf{A}_{k_b} \mathbf{A}_{k_b+1})^{-1} \left(\mathbf{b}_{k_b}^{(i+1)} + \mathbf{A}_{k_b} \mathbf{b}_{k_b+1}^{(i+1)} \right), \quad (55)$$

for $k = k_b$,

$$\mathbf{d}_{k_b+1}^{(i+1)} = (\mathbf{I} - \mathbf{A}_{k_b+1} \mathbf{A}_{k_b})^{-1} \left(\mathbf{b}_{k_b+1}^{(i+1)} + \mathbf{A}_{k_b+1} \mathbf{b}_{k_b}^{(i+1)} \right), \quad (56)$$

for $k = k_b + 1$,

$$\mathbf{d}_k^{(i+1)} = \mathbf{A}_k \mathbf{d}_{k+1} + \mathbf{b}_k^{(i+1)}, \quad (57)$$

for $k = k_b - 1, k_b - 2, \dots, 1$, and

$$\mathbf{d}_k^{(i+1)} = \mathbf{A}_k \mathbf{d}_{k-1} + \mathbf{b}_k^{(i+1)}, \quad (58)$$

for $k = k_b + 2, k_b + 3, \dots, K$.

Since inverse matrices in the computation of $\{\mathbf{b}_k^{(i+1)}\}_{k=1}^K$, $\mathbf{d}_{k_b}^{(i+1)}$, $\mathbf{d}_{k_b+1}^{(i+1)}$, and $\{\mathbf{A}_k\}_{k=1}^K$ remain unchanged at each iteration, they can be computed before the iteration. The proposed ADMM algorithm for amplitude matching with the differential-norm penalty is summarized in Algorithm 3.

REFERENCES

- [1] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th AES Conv.*, 2008, Art. no. 7358. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14488>
- [2] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, 2005.
- [3] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica United Acustica*, vol. 94, pp. 988–999, 2008.
- [4] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 1, pp. 107–116, Jan. 2009.
- [5] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 4, pp. 685–696, Apr. 2013.
- [6] S. Koyama, K. Furuya, K. Wakayama, S. Shimauchi, and H. Saruwatari, "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," *J. Acoust. Soc. Amer.*, vol. 139, no. 3, pp. 1024–1036, 2016.
- [7] O. Kirkeby, P. A. Nelson, F. O. Bustamante, and H. Hamada, "Local sound field reproduction using digital signal processing," *J. Acoust. Soc. Amer.*, vol. 100, no. 3, pp. 1584–1593, 1996.
- [8] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant environment," *J. Acoust. Soc. Amer.*, vol. 117, no. 4, pp. 2100–2111, 2005.
- [9] 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, Dec. 2019.
- [10] S. Koyama, G. Chardon, and L. Daudet, "Optimizing source and sensor placement for sound field control: An overview," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 686–714, 2020.
- [11] M. Poletti, "An investigation of 2-D multizone surround sound systems," in *Proc. 125th AES Conv.*, 2008, Art. no. 7551. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14703>
- [12] W. Zhang, T. Abhayapala, T. Betlehem, and F. Fazi, "Analysis and control of multi-zone sound field reproduction using modal-domain approach," *J. Acoust. Soc. Amer.*, vol. 140, pp. 2134–2144, Sep. 2016.
- [13] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, Aug. 2011.
- [14] T. Betlehem and P. D. Teal, "A constrained optimization approach for multi-zone surround sound," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2011, pp. 437–440.
- [15] W. Jin and W. B. Kleijn, "Multizone soundfield reproduction in reverberant rooms using compressed sensing techniques," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2014, pp. 4728–4732.
- [16] W. F. Druyvesteyn and J. Garas, "Personal sound," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 685–701, 1997.
- [17] 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, Mar. 2015.
- [18] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Amer.*, vol. 111, no. 4, pp. 1695–1700, 2002.
- [19] Y. Cai, M. Wu, and J. Yang, "Sound reproduction in personal audio systems using the least-squares approach with acoustic contrast control constraint," *J. Acoust. Soc. Amer.*, vol. 135, no. 2, pp. 734–741, 2014.
- [20] J. Nocedal and S. J. Wright, *Numerical Optimization*. Berlin, Germany: Springer, 2006.
- [21] 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.
- [22] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *Amer. Statistician*, vol. 58, no. 1, pp. 30–37, 2004.
- [23] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2016.
- [24] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [25] D. Griffin and J. Lim, "Signal estimation from modified short-time fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.
- [26] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
- [27] J. Liang, P. Stoica, Y. Jing, and J. Li, "Phase retrieval via the alternating direction method of multipliers," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 5–9, Jan. 2018.
- [28] Z. Wen, C. Yang, X. Liu, and S. Marchesini, "Alternating direction methods for classical and ptychographic phase retrieval," *Inverse Problems*, vol. 28, 2012, Art. no. 115010.
- [29] T. Narutomi, "Parallel algorithms for large scale computation in science and technology," Ph.D. dissertation, Graduate Sch. Eng., Tohoku Univ., Sendai, Japan, 1993.
- [30] S. Koyama, T. Nishida, K. Kimura, T. Abe, N. Ueno, and J. Brunström, "MESHRIR: A dataset of room impulse responses on meshed grid points for evaluating sound field analysis and synthesis methods," in *Proc. IEEE Int. Workshop Appl. Signal Process. Audio Acoust.*, 2021, pp. 1–5.



Takumi Abe received the B.E. degree in mathematical engineering and information physics in 2021 from the University of Tokyo, Tokyo, Japan, where he is currently working toward the master's degree. His research interests include haptics and ultrasonic wave.



Shoichi Koyama (Member, IEEE) received the B.E., M.S., and Ph.D. degrees from the University of Tokyo, Tokyo, Japan, in 2007, 2009, and 2014, respectively. In 2009, he joined Nippon Telegraph and Telephone (NTT) Corporation, Tokyo, Japan, as a Researcher of acoustic signal processing. In 2014, he moved to the University of Tokyo, where he has been a Lecturer since 2018. From 2016 to 2018, he was a Visiting Researcher with Paris Diderot University (Paris7), Institut Langevin, Paris, France. His research interests include audio signal processing,

acoustic inverse problems, and spatial audio. He was the recipient of Itakura Prize Innovative Young Researcher Award by the Acoustical Society of Japan in 2015, and the Research Award by Funai Foundation for Information Technology in 2018.



Natsuki Ueno (Member, IEEE) received the B.E. degree in engineering from Kyoto University, Kyoto, Japan, in 2016, and the M.S. and Ph.D. degrees in information science and technology from the University of Tokyo, Tokyo, Japan, in 2018 and 2021, respectively. He is currently a Project Assistant Professor with Tokyo Metropolitan University, Tokyo, Japan. His research interests include spatial audio and acoustic signal processing.



Hiroshi Saruwatari (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Nagoya University, Nagoya, Japan, in 1991, 1993, and 2000, respectively. In 1993, he joined SECOM IS Laboratory, Tokyo, Japan, and in 2000, Nara Institute of Science and Technology, Ikoma, Japan. Since 2014, he has been a Professor with The University of Tokyo, Tokyo, Japan. His research interests include statistical audio signal processing, blind source separation, and speech enhancement. He has put his research into the world's first commercially available independent-component-analysis-based BSS microphone in 2007. He was the recipient of several paper awards from IEICE in 2001 and 2006, from TAF in 2004, 2009, 2012, and 2018, from IEEE-IROS2005 in 2006, and from APSIPA in 2013 and 2018, and also the DOCOMO Mobile Science Award in 2011, Ichimura Award in 2013, Commendation for Science and Technology by the Minister of Education in 2015, Achievement Award from IEICE in 2017, and Hoko-Award in 2018. He has been professionally involved in various volunteer works for IEEE, EURASIP, IEICE, and ASJ. Since 2018, he has been an APSIPA Distinguished Lecturer.