



HAL
open science

Optimizing Source and Sensor Placement for Sound Field Control: An Overview

Shoichi Koyama, Gilles Chardon, Laurent Daudet

► **To cite this version:**

Shoichi Koyama, Gilles Chardon, Laurent Daudet. Optimizing Source and Sensor Placement for Sound Field Control: An Overview. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2020, 10.1109/TASLP.2020.2964958 . hal-02459046

HAL Id: hal-02459046

<https://centralesupelec.hal.science/hal-02459046>

Submitted on 29 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimizing Source and Sensor Placement for Sound Field Control: An Overview

Shoichi Koyama, *Member, IEEE*, Gilles Chardon, *Member, IEEE*, and Laurent Daudet, *Senior Member, IEEE*

Abstract—In order to control an acoustic field inside a target region, it is important to choose suitable positions of secondary sources (loudspeakers) and sensors (control points/microphones). This paper provides an overview of state-of-the-art source and sensor placement methods in sound field control. Although the placement of both sources and sensors greatly affects control accuracy and filter stability, their joint optimization has not been thoroughly investigated in the acoustics literature. In this context, we reformulate five general source and/or sensor placement methods that can be applied for sound field control. We compare the performance of these methods through extensive numerical simulations in both narrowband and broadband scenarios.

Index Terms—source and sensor placement, sound field control, sound field reproduction, subset selection, interpolation.

I. INTRODUCTION

THE aim of sound field control is to synthesize a desired sound field inside a target region. It can be applied to various settings including high-fidelity audio, virtual/augmented reality, and noise cancellation systems. A typical strategy for sound field control is to control sound pressures at multiple discrete positions, i.e., control points inside the target region, using multiple secondary sources, i.e., loudspeakers [1]–[7]. In general, the inverse of a given transfer function matrix between secondary sources and control points is calculated in a (regularized) least-squares-error sense; this is referred to as the *pressure matching method*. The positions of secondary sources and control points have a great effect not only on the control accuracy but also on the stability of the inverse filter. Therefore, these positions must be carefully chosen. In practical applications, it is often necessary to cover the entire target region with high accuracy while using the smallest possible number of secondary sources and control points, since the measurement of transfer functions is costly and time-intensive. Unstable inverse filters increase sensitivity to perturbations in transfer functions and can lead to extremely large loudspeaker outputs. Moreover, when the control points of the sound pressures are arranged on the boundary of an

enclosed region, it is known that the sound field inside this region cannot be uniquely determined at several frequencies, leading to significant deterioration of the control accuracy [8]. This nonuniqueness property is called the *forbidden frequency problem*.

In the context of sound field recording and reproduction, a similar problem has been addressed by modeling a sound field in a continuous setting, usually under the free-field assumption [9]–[11]. The positions of the loudspeakers for reproduction and microphones for recording are generally determined by regularly discretizing the continuous surface of the array. The forbidden frequency problem is typically avoided by using microphones mounted on an acoustically rigid object, directional microphones, or multiple layers of microphone arrays [10], [12]–[15]. This regular placement performs well in a free field and when the array has a simple shape, such as a sphere, plane, circle, or line. However, it may be suboptimal for the more complicated geometries used in practical situations and in a reverberant environment, as addressed in the pressure-matching-based sound field control.

As discussed above, to the best of our knowledge, joint optimization of the source and sensor placement for sound field control has not been thoroughly investigated, especially for an arbitrary shape of the target region. The loudspeaker placement problem has been tackled in the context of active noise and structural acoustic control [16]–[21]. Recent studies on loudspeaker placement methods for sound field reproduction were compared in [22]. Most work related to control-point placement in the literature has been about microphone array design [23], [24]. Nevertheless, a large number of sensor placement algorithms have been investigated in the context of sensor networks and machine learning. Most of them solve a selection problem from predefined candidate positions, and efficient algorithms for solving this combinatorial problem have been derived on the basis of various criteria from experimental design [25]–[28] or information-theoretic measures [29]–[31]. A method based on *frame potential* [32] has been proposed for both source and sensor placement [33], [34]. Finally, we recently proposed a joint source and sensor placement method for sound field control that was based on the empirical interpolation method [35].

The objectives of this paper are to give an overview of the state-of-the-art algorithms for source and/or sensor placement in the sound field control problem and to report the results of comparing five of these methods in terms of performance and computational complexity. As mentioned above, some of these methods applied to this problem have been proposed in various contexts; however, their application and experimental compari-

S. Koyama is with the Graduate School of Information Science and Technology, the University of Tokyo, Tokyo 113-8656, Japan (e-mail: koyama.shoichi@ieee.org).

S. Koyama is also with JST, PRESTO, Saitama 332-0012, Japan.

G. Chardon is with the Laboratoire des Signaux et Systèmes (L2S), CentraleSupélec-CNRS-Université Paris-Sud, Université Paris-Saclay, 3, rue Joliot Curie, 91192, Gif-sur-Yvette, France (email: gilles.chardon@centralesupelec.fr).

L. Daudet is with Institut Langevin, ESPCI Paris, CNRS, PSL University, Paris Diderot University, 1 rue Jussieu, Paris, 75005 France (e-mail: laurent.daudet@espci.fr).

Part of this work was carried out while S. Koyama was visiting Institut Langevin, Paris Diderot University as a JSPS Overseas Research Fellow.

Manuscript received April XX, 20XX; revised August XX, 20XX.

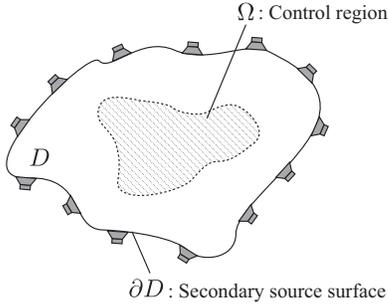


Fig. 1. Controlling sound field inside Ω with secondary sources placed on ∂D .

son have not been previously reported. In particular, one must reformulate some of these methods for the sensor selection problem when applying them to the broadband scenario. We describe, in a unified manner, five representative methods for this problem and provide the corresponding algorithms. These methods are compared in terms of control accuracy and filter stability by performing numerical simulations. All the code for reproducing the experimental results is available online [36].

The paper is organized as follows. The sound field control problem is defined and related works are overviewed in Sect. II. In Sect. III, the five source and/or sensor placement methods are introduced. The results of numerical simulations carried out to compare these methods are reported in Sect. IV. Finally, Sect. V concludes this paper.

A. Notation

Italic letters denote scalars, lower case boldface letters denote vectors, and upper case boldface letters denote tensors of order two or more, including matrices. The sets of real and complex numbers are denoted by \mathbb{R} and \mathbb{C} , respectively.

Subscripts of scalars, vectors, and tensors indicate their indexes. For example, $x_{i,j}$ is the (i,j) th entry of matrix \mathbf{X} , and \mathbf{x}_i is the i th vector extracted from \mathbf{X} . The submatrix extracted from \mathbf{X} using sets of indexes \mathcal{I} and \mathcal{J} is denoted as $\mathbf{X}_{\mathcal{I},\mathcal{J}}$.

The imaginary unit is denoted by j . The complex conjugate, conjugate transpose, and inverse are denoted by superscripts $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^{-1}$, respectively. Δ represents the Laplace operator. The absolute value of a scalar x is denoted as $|x|$. The ℓ_p -norm of a vector \mathbf{x} is denoted as $\|\mathbf{x}\|_p$.

The sound velocity, angular frequency, and wave number are denoted as c , ω , and $k = \omega/c$, respectively.

II. SECONDARY SOURCE AND SENSOR PLACEMENT FOR SOUND FIELD CONTROL

A. Sound Field Control Problem

We here define the sound field control problem. Suppose that region Ω is included in or coincides with region D , as shown in Fig. 1. We assume that the sound field inside Ω is synthesized by secondary sources placed on the boundary of D , ∂D . Region Ω and surface ∂D are referred to as the control region and secondary source surface, respectively.

Sound field $u(\mathbf{x}, \omega)$ satisfies the Helmholtz equation in D as

$$\Delta u(\mathbf{x}, \omega) + k^2 u(\mathbf{x}, \omega) = 0, \quad (1)$$

where \mathbf{x} is the position vector. On the basis of the single-layer boundary integral [37], $u(\mathbf{x}, \omega)$ can be represented using a continuous distribution of sources on ∂D . This integral is the theoretical foundation of sound field synthesis methods [38], such as wave field synthesis [9], [39], [40], higher-order ambisonics [10], [41], [42], and the equivalent source method for sound field reconstruction [43], [44]. Using the free-field Green's function $G_m(\cdot, \omega)$ on ∂D , $u(\mathbf{x}, \omega)$ for $\mathbf{x} \in D$ is represented as

$$u(\mathbf{x}, \omega) = \int_{\mathbf{y} \in \partial D} \varphi(\mathbf{y}) G_m(\mathbf{x}|\mathbf{y}, \omega) d\mathbf{y}, \quad (2)$$

where $\varphi(\cdot)$ is referred to as the source density and $G_m(\cdot, \omega)$ is defined as

$$G_m(\mathbf{x}|\mathbf{y}, \omega) = \frac{\exp(jk\|\mathbf{x} - \mathbf{y}\|_2)}{4\pi\|\mathbf{x} - \mathbf{y}\|_2}, \quad (3)$$

which corresponds to the transfer function of a monopole located at \mathbf{y} . Hereafter, ω is omitted for notational simplicity. This equation means that an arbitrary homogeneous sound field inside D can be synthesized by continuously distributed monopoles on ∂D , which are driven by $\varphi(\mathbf{y})$, assuming free-field propagation. When D is a simple shape such as a plane or sphere, $\varphi(\mathbf{y})$ can be analytically derived [38]. However, it is difficult to obtain an analytical form of $\varphi(\mathbf{y})$ when D has an arbitrary geometry or the transfer function of the secondary source cannot be approximated by a monopole, such as when it includes reverberation. An additional problem is the nonuniqueness of the determination of $\varphi(\mathbf{y})$ through the integral equation at forbidden frequencies, i.e., eigenfrequencies of domain Ω with Dirichlet boundary conditions [8].

In practice, a continuous distribution of secondary sources must be approximated by a discrete set of loudspeakers. Synthesized sound pressure $u_{\text{syn}}(\mathbf{x})$ at position \mathbf{x} is represented by a linear combination of transfer functions of the loudspeakers:

$$u_{\text{syn}}(\mathbf{x}) = \sum_{l=1}^L d_l g_l(\mathbf{x}), \quad (4)$$

where d_l and g_l are the driving signal and the transfer function (Green's function) at \mathbf{x} for the l th loudspeaker, respectively.

The driving signals d_l can be determined either by approximating integral (2) with a finite sum or by minimizing the objective function

$$J = \int_{\mathbf{x} \in \Omega} \left| \sum_{l=1}^L d_l g_l(\mathbf{x}) - u_{\text{des}}(\mathbf{x}) \right|^2 d\mathbf{x}, \quad (5)$$

which aims to minimize the mean square error of the reconstruction in the domain Ω of interest. The *secondary source placement* problem is twofold.

- The discrete set of secondary sources should be able to generate desired sound fields, such as plane waves and monopoles, in Ω .
- The sound field should be synthesized in a stable manner such that the amplitudes of the required driving signals of the secondary sources are bounded.

The minimization of J is difficult to solve directly because (5) includes an integral with respect to \mathbf{x} . Moreover, it is

difficult to measure or estimate $g_l(\mathbf{x})$ in the entire Ω for continuous \mathbf{x} , especially in a reverberant environment. A typical strategy for minimizing J is to discretize \mathbf{x} inside Ω and consider the following criterion:

$$J_d = \|\mathbf{u}^{\text{des}} - \mathbf{G}\mathbf{d}\|_2^2, \quad (6)$$

where $\mathbf{u}^{\text{des}} \in \mathbb{C}^M$ is the vector of the desired pressures at M discrete positions inside Ω , i.e., control points, $\mathbf{G} \in \mathbb{C}^{M \times L}$ is the matrix of the transfer functions between secondary sources and control points, and $\mathbf{d} \in \mathbb{C}^L$ is the vector of the driving signals. Then, \mathbf{d} can be obtained by

$$\mathbf{d} = \mathbf{G}^\dagger \mathbf{u}^{\text{des}}, \quad (7)$$

where $(\cdot)^\dagger$ represents the Moore–Penrose pseudoinverse. This strategy is referred to as pressure matching. Since the calculation of the inverse of \mathbf{G} frequently becomes unstable, it is usually necessary to regularize (7). For instance, the use of Tikhonov regularization for the case of $M \geq L$ is represented as

$$\mathbf{d} = (\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{u}^{\text{des}}, \quad (8)$$

where λ is the regularization parameter.

The *sensor or control point placement* problem is to build a set of control points such that the error J_d computed on the control points is an appropriate proxy for the continuous error J , or equivalently, a set of sampling points allowing accurate reconstruction of the sound field in Ω . As pointed out above, an inadequate choice of sampling points will not allow a fine control on the sound field. In particular, at the forbidden frequencies, eigenmodes of Ω with Dirichlet boundary conditions will not be observable by control points located on the boundary of Ω .

The *secondary source and control point placement* problem is to design sets of both sources and control points, possibly jointly, such that sound fields in the continuous domain Ω can be synthesized by a discrete set of secondary sources and controlled at a discrete set of points. A simple solution would be to finely discretize ∂D and Ω to approximate integrals (2) and (5). Apart from the obvious cost and the physical constraints preventing the use of a large number of loudspeakers and microphones, excessively dense sampling leads to extremely high redundancy of the transfer function matrix \mathbf{G} , especially at low frequencies. Furthermore, since it is necessary to measure the transfer functions by using loudspeakers and microphones to obtain \mathbf{G} , it is preferable that the numbers of secondary sources and control points are as small as possible. Their positions also have a great effect on the control accuracy and filter stability.

Finally, this placement problem poses the challenge of computational complexity. Indeed, the size of the sets of possible secondary sources and control points grows rapidly with the size of the control domain and the frequency. For example, in numerical analysis using the boundary element method, a common criterion for discretization is 6 nodes per wavelength [45]. For a cubic control region of 2.0 m length with secondary sources on its surface at a frequency of 1.7 kHz, there are 21,600 and 216,000 possible positions

for loudspeakers and control points, respectively. Quadratic or even cubic complexities in time and memory will limit the application of an algorithm to small control domains and low frequencies.

B. Related Work on Secondary Source and Sensor Placement

In the context of sound field reproduction, the concept of a continuous secondary source distribution has been employed. As an example, we assume that Ω and D are spherical with radii R_Ω and R_D , respectively. Position vectors $\mathbf{x} \in \Omega$ and $\mathbf{y} \in \partial D$ are represented in spherical coordinates as $\mathbf{x} = (r, \theta, \phi)$ and $\mathbf{y} = (R_D, \theta_D, \phi_D)$, respectively. Then, the desired sound field $u_{\text{des}}(\mathbf{x})$ can be represented by the spherical wave function expansion as [10], [46], [47]

$$u_{\text{des}}(\mathbf{x}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \tilde{u}_{\text{des},\nu}^\mu j_\nu(kr) Y_\nu^\mu(\theta, \phi), \quad (9)$$

where $j_\nu(\cdot)$ is the ν th-order spherical Bessel function of the first kind and $Y_\nu^\mu(\cdot)$ is the spherical harmonic function of ν th order and μ th degree.

$$Y_\nu^\mu(\theta, \phi) = \sqrt{\frac{2\nu+1}{4\pi} \frac{(\nu-\mu)!}{(\nu+\mu)!}} P_\nu^\mu(\cos \theta) \exp(j\mu\phi) \quad (10)$$

Here, $P_\nu^\mu(\cdot)$ are the associated Legendre functions. Since the secondary source distribution on ∂D is assumed to be continuous, the expansion coefficients of the driving signals and transfer functions can also be defined as \tilde{d}_ν^μ and $\tilde{g}_\nu^\mu(r - R_D)$, respectively. We here assume that the transfer function of each secondary source is monopole; therefore, $\tilde{g}_\nu^\mu(r - R_D)$ is approximated using the expansion coefficients of $G_m(\mathbf{x}|\mathbf{y})$ as

$$\tilde{g}_\nu^\mu(r - R_D) \approx jk j_\nu(kr) h_\nu(kR_D) Y_\nu^\mu(\theta_D, \phi_D)^*, \quad (11)$$

where $h_\nu(\cdot)$ is the ν th-order spherical Hankel function of the first kind. It is also possible to take the directivity of the secondary sources into consideration if all the secondary sources have the same directivity, and their directional characteristics are known [48], [49]. Then, the expansion coefficients of the synthesized sound field $\tilde{u}_{\text{syn},\nu}^\mu$ can be represented by the product of \tilde{d}_ν^μ and $\tilde{g}_\nu^0(r - R_D)$ as [50]

$$\tilde{u}_{\text{syn},\nu}^\mu = 2\pi R_D^2 \sqrt{\frac{4\pi}{2\nu+1}} \tilde{d}_\nu^\mu \tilde{g}_\nu^0(r - R_D). \quad (12)$$

Since the synthesized field should be coincident with the desired field, i.e., $\tilde{u}_{\text{syn},\nu}^\mu = \tilde{u}_{\text{des},\nu}^\mu$, the driving signals of the secondary sources can be obtained as [10]

$$\begin{aligned} \tilde{d}_\nu^\mu &= \sqrt{\frac{2\nu+1}{4\pi}} \frac{j_\nu(kr)}{2\pi R_D^2 \tilde{g}_\nu^0(r - R_D)} \tilde{u}_{\text{des},\nu}^\mu \\ &= \frac{1}{2\pi j k R_D^2 h_\nu(kR_D)} \tilde{u}_{\text{des},\nu}^\mu. \end{aligned} \quad (13)$$

In practice, the secondary source distribution must be discretized as an array of loudspeakers. Platonic solids can be used for uniform sampling on the sphere; however, the number of samples is constrained, i.e., 4, 6, 8, 12, and 20 samples, corresponding to vertexes of the platonic solids.

Nearly uniform sampling schemes [51], such as t -design [52], allow the use of a wider range of sampling sets. The advantage of the (nearly) uniform sampling scheme is the small number of samples for a certain maximum expansion order ν_{\max} , which determines the frequency and range of the region of accurate reproduction [23], [53]. When secondary sources are distributed within a 2D plane, such as in a circular arrangement, some reasonable approximations for the 3D field are possible in this analytical approach, which is referred to as *2.5D reproduction*. It is in general difficult to apply such approximation for the pressure matching method.

To estimate the expansion coefficients of the desired field $\tilde{u}_{\text{des},\nu}^{\mu}$, the sound field is usually captured on the spherical surface of Ω , which also requires a sampling scheme for the microphone array design. Furthermore, the pressure distribution on the sphere is not sufficient to uniquely determine the sound field inside Ω at several frequencies because of the forbidden frequency problem. This nonuniqueness is caused by the existence of eigenmodes of the Helmholtz equation with Dirichlet boundary conditions. At an eigenfrequency, since an eigenmode has nonzero values inside the sphere but zero values on its boundary, it makes the estimation of its amplitude impossible using microphones on this boundary only. Uniqueness can be obtained by using multiple layers, with measurements inside the sphere where the eigenmodes have nonzero values, or with directional microphones, making the forbidden eigenfrequencies complex [12], [54]. Alternatively, microphone arrays mounted on an acoustically rigid object also prevent this nonuniqueness issue [10], [13], [15], [55], [56]. Such recording and reproduction methods based on spatial Fourier analysis of the sound field can only be applied for simple array geometries [11], [14], [46], [47], [57], where the regular sampling of the array geometry is usually applied. Even though this constraint on the array geometry can be relaxed by constructing a linear equation of the expansion coefficients of synthesized and desired sound fields, as in the *mode matching method* [10], [58]–[61], it is not a trivial task to obtain its optimal sampling scheme as in the pressure matching method.

For active noise and structural acoustic control, several attempts have been made at optimizing secondary source placement. The optimal positions are selected from the candidates on the basis of the mean-square-error criterion with the given primary source field, i.e., the desired sound field. Nonlinear optimization techniques such as sequential quadratic programming were applied in [17], [18]. The use of the correlation coefficients in multiple linear regression is suggested in [16], [19]. Heuristic algorithms such as genetic and simulated annealing algorithms are proposed for solving the combinatorial selection problem [20]. On the basis of the Gram–Schmidt orthogonalization of \mathbf{G} , Asano et al. [21] proposed an efficient algorithm that does not strongly depend on the desired sound field. The details of this method are described in Sect. III-A. In recent studies, sparse approximation algorithms [62], [63] have been applied to the loudspeaker placement in sound field reproduction [64]–[66]. A singular-value-decomposition-based algorithm was proposed in [67], and a comparative study of these recent loudspeaker placement methods was presented

in [22]. Again, however, these methods strongly depend on the specific desired sound field even though it will vary in practical applications. On the other hand, control-point placement has not been specifically addressed in the literature with a few exceptions [17], [68]. The authors recently proposed a method for joint source and sensor placement for the sound field control problem, which is independent of the desired field, on the basis of the empirical interpolation method [35] (see Sect. III-E for details).

On the other hand, general sensor placement as well as subset selection has been addressed in many studies on sensor networks and machine learning. The linear measurement model is usually considered as in (6) by regarding \mathbf{u}^{des} as the observation and \mathbf{d} as the parameter to be estimated. The sensor positions are chosen from the candidate positions on the basis of some performance measures. Instead of the mean-square-error criterion in the secondary source placement methods for sound field control [69], [70], measures on the Gram matrix $\mathbf{T} = \mathbf{G}^H \mathbf{G}$ from the experimental design [71] are usually used. For example, the cost function can be formulated as the minimization of the sum of eigenvalues of \mathbf{T}^{-1} (*A-optimality*), the minimization of the maximum eigenvalue of \mathbf{T}^{-1} (*E-optimality*), and the minimization of the log determinant of \mathbf{T}^{-1} (*D-optimality*) [25]–[28], [72]. Information-theoretic measures such as entropy [29], [30] and mutual information [31] have also been applied. In [33], the frame potential of the measurement matrix \mathbf{G} was used for the optimization criterion, which has also been applied to the source placement problem [34]. The optimization algorithms based on these criteria are classified into three categories: convex optimization [27], [72], greedy algorithms [28], [31], [33], [69], [70], and heuristics [25], [26], [29], [30]. These sensor placement methods based on the linear measurement model can be applied to control-point placement for the sound field control problem; however, to the best of our knowledge, such investigations have not been carried out. Moreover, since these methods were proposed for general sensor networks, they must be extended to treat broadband acoustic signals.

III. SECONDARY SOURCE AND SENSOR PLACEMENT METHODS

In this section we describe five source and/or sensor placement methods for sound field control, focusing on the practicality of generic methods, i.e., those that do not use the knowledge of a single specific desired sound field. Since it is difficult to directly solve this placement problem, it is relaxed to the problem of selecting a subset of secondary sources and sensors from the candidate positions, as described in Sect. II. We denote the sets of candidate source and sensor locations as \mathcal{L} ($|\mathcal{L}| = L$) and \mathcal{M} ($|\mathcal{M}| = M$), respectively, and assume that the transfer function matrix of candidate locations $\mathbf{G} \in \mathbb{C}^{L \times M}$ is known, e.g., obtained through numerical simulation. The goal is to choose an appropriate set of source and sensor locations from \mathcal{L} and \mathcal{M} in terms of control accuracy and filter robustness, where the selected sets of the source and sensor locations are denoted as $\mathcal{S} \subseteq \mathcal{L}$ and $\mathcal{T} \subseteq \mathcal{M}$, respectively. Thus, a submatrix of \mathbf{G} of the selected locations, i.e., $\mathbf{G}_{\mathcal{S},\mathcal{T}}$,

Algorithm 1 Source Placement Based on Gram–Schmidt Orthogonalization

Require: A set of candidate source locations \mathcal{L} , a transfer function matrix \mathbf{G} , and number of sources K

Ensure: Source locations $\mathcal{S} \subseteq \mathcal{L}$

Set $k = 1$ and $\mathcal{S} = \emptyset$

Initialize l_1 by using the desired pressure \mathbf{u}^{des} as

$$l_1 = \arg \min_{l \in \mathcal{L}} \frac{\|\mathbf{g}_l - \mathbf{p}\|_2}{\|\mathbf{g}_l\|_2},$$

where

$$\mathbf{p} = \frac{\mathbf{g}_l^H \mathbf{u}^{\text{des}}}{(\mathbf{u}^{\text{des}})^H \mathbf{u}^{\text{des}}} \mathbf{u}^{\text{des}}$$

Set the first orthonormal basis $\mathbf{v}_1 = \mathbf{g}_{l_1} / \|\mathbf{g}_{l_1}\|_2$

Update the set of selected and available locations

$$\mathcal{S} = l_1 \quad \text{and} \quad \mathcal{L} = \mathcal{L} \setminus l_1$$

for $k = 2$ to K **do**

Select the source index so that the angle between the transfer function vector and the subspace spanned by the previously identified sources \mathcal{S} is maximal

$$l_k = \arg \max_{l \in \mathcal{L}} \|\mathbf{e}_l\|_2,$$

where

$$\mathbf{e}_l = \mathbf{g}_l - \sum_{j=1}^{k-1} (\mathbf{v}_j^H \mathbf{g}_l) \mathbf{v}_j$$

Set the k th orthonormal basis $\mathbf{v}_k = \mathbf{e}_{l_k} / \|\mathbf{e}_{l_k}\|_2$

Update the set of selected and available locations

$$\mathcal{S} = \mathcal{S} \cup l_k \quad \text{and} \quad \mathcal{L} = \mathcal{L} \setminus l_k$$

end for

can be constructed. Since candidates \mathcal{L} and \mathcal{M} should be composed by finely discretizing Ω and ∂D , the exhaustive search of the subsets from the candidates is impractical.

In addition, a particular set of loudspeakers and control points should be able to synthesize and control sound fields in a wide range of frequencies, ideally most of the human hearing range. Subset selection over a broad frequency range is obviously more difficult than in the narrowband case as robustness must be jointly ensured for all frequencies by the sets of chosen loudspeakers and control points. The set of target frequency bins is denoted as \mathcal{F} ($|\mathcal{F}| = F$) and its index is denoted as f for the broadband case.

Two methods proposed for sound field control are introduced; they are based on Gram–Schmidt orthogonalization [21] and the empirical interpolation method [35]. Three general placement methods for sources or sensors based on D -optimal design [27], mutual information [31], and frame potential [33], [34], are applied to sound field control.

A. Secondary Source Placement Based on Gram–Schmidt Orthogonalization

Asano et al. [21] proposed a secondary source selection method based on Gram–Schmidt orthogonalization. The linear

Algorithm 2 Sensor Placement Based on Determinant

Require: A set of candidate sensor locations \mathcal{M} , a transfer function matrix \mathbf{G} , and number of sensors K

Ensure: Sensor locations $\mathcal{T} \subseteq \mathcal{M}$

Initialize $\mathbf{z} = (K/M)\mathbf{1}$

while $(-\nabla\psi^H \Delta\mathbf{z})^{1/2} > \epsilon$ **do**

Compute Newton search step $\Delta\mathbf{z}$

$$\Delta\mathbf{z} = -(\nabla^2\psi)^{-1} \nabla\psi + \left(\frac{\mathbf{1}^T (\nabla^2\psi)^{-1} \nabla\psi}{\mathbf{1}^T (\nabla^2\psi)^{-1} \mathbf{1}} \right) (\nabla^2\psi)^{-1} \mathbf{1}$$

Backtracking line search to compute the step size $\alpha \in (0, 1]$

Update \mathbf{z}

$$\mathbf{z} = \mathbf{z} + \alpha \Delta\mathbf{z}$$

end while

Set \mathcal{T} as the set of K largest coefficients of \mathbf{z}

Local optimization by swapping one of the K chosen sensors (\mathcal{T}) with one of the $M - K$ sensors not chosen ($\mathcal{M} \setminus \mathcal{T}$)

independence of the transfer functions is employed for the optimization criterion. As in Algorithm 1, in the k th step, one secondary source l_k is selected from the unused candidates so that the angle between its transfer function and the subspace spanned by the set of the transfer functions of the sources selected until the $k - 1$ th step is maximal. The l th transfer function vector \mathbf{g}_l is projected onto the subspace spanned by the orthonormal basis \mathbf{v}_j ($j \in \{1, \dots, k - 1\}$). The k th selection l_k is chosen so that the ℓ_2 -norm of the residual between \mathbf{g}_l and its projection is maximized:

$$l_k = \arg \max_{l \in \mathcal{L}} \|\mathbf{e}_l\|_2, \quad (14)$$

where

$$\mathbf{e}_l = \mathbf{g}_l - \sum_{j=1}^{k-1} (\mathbf{v}_j^H \mathbf{g}_l) \mathbf{v}_j. \quad (15)$$

Then, the k th orthonormal basis is set as $\mathbf{v}_k = \mathbf{e}_{l_k} / \|\mathbf{e}_{l_k}\|_2$ and \mathcal{S} and \mathcal{L} are updated. The initial secondary source l_1 is chosen so that the transfer function vector \mathbf{g}_1 is close to the vector of the desired sound pressures \mathbf{u}^{des} in [21]. In other words, l_1 minimizes the residual between \mathbf{g}_l and the projection of \mathbf{g}_l onto \mathbf{u}^{des} . Therefore, only the initial selection depends on the specific desired sound field. The computational cost of Algorithm 1 is $\mathcal{O}(MLk)$ for the k th iteration and the number of iterations is the required number of sources K . Therefore, the total computational cost amounts to $\mathcal{O}(MLK^2)$ in time. Memory cost is dominated by the storage of \mathbf{G} .

Although Algorithm 1 can only be applied for the narrowband case, an extension to the broadband case is also described in [21]. It is simply achieved by evaluating the (weighted) average of $\|\mathbf{e}_l\|_2$ over multiple frequency bins after calculating \mathbf{e}_l for all \mathcal{F} .

B. Sensor Placement Based on Determinant

A method for sensor selection by convex optimization has been proposed by Joshi and Boyd [27]. The method is based on D -optimal design, where the objective is to select sensors such that the determinant of the inverse of the Fisher matrix for parameter estimation is minimized. In the case where the measurement errors are uncorrelated normal variables, this is equivalent to minimizing the volume of the confidence region of the estimation of the parameters.

The set of selected sensors is characterized by a vector of M Boolean values. As the direct optimization of the set is a combinatorial problem, it is replaced by its convex relaxation as

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} \quad \log \det \left(\sum_{m=1}^M z_m \mathbf{g}_m^* \mathbf{g}_m^T \right) \\ & \text{subject to} \quad \mathbf{1}^T \mathbf{z} = K, \quad 0 \leq z_m \leq 1, \end{aligned} \quad (16)$$

with variable $\mathbf{z} \in \mathbb{R}^M$. With the objective function and the constraints on \mathbf{z} being convex, this optimization problem can be efficiently solved using, for example, interior-point methods [73]. The indexes of the K largest elements of \mathbf{z} are taken as the set of selected sensors.

In this relaxed formulation, D -optimality is equivalent to G -optimality, i.e., the minimization of the maximal variance of the reconstruction of the acoustic field over the region of interest [74].

Since it is not necessary to solve (16) with high accuracy in practice, the following approximate problem, using a log-barrier to account for the constraints, is suggested in [27]:

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} \quad \psi(\mathbf{z}) = \log \det \left(\sum_{m=1}^M z_m \mathbf{g}_m^* \mathbf{g}_m^T \right) \\ & \quad + \kappa \sum_{m=1}^M (\log(z_m) + \log(1 - z_m)) \\ & \text{subject to} \quad \mathbf{1}^T \mathbf{z} = K, \end{aligned} \quad (17)$$

where κ is a positive parameter that controls the quality of the approximation. Since ψ is concave and smooth, Newton's method is applied to solve (17) as in Algorithm 2. The computational complexity for computing Newton's step $\Delta \mathbf{z}$ is $\mathcal{O}(M^3)$, which is dominated by the computation of the inverse of the Hessian $\nabla^2 \psi \in \mathbb{R}^{M \times M}$. The total computational cost depends on the number of iterations. Memory usage is governed by the Hessian matrix with $\mathcal{O}(M^2)$ coefficients.

The solution of the approximate relaxed problem (17) is improved by a local optimization method. Starting with the vector $\hat{\mathbf{z}}$ comprising the K largest values of \mathbf{z} of the solution of (17), the determinant of the error covariance $\det \hat{\mathbf{\Xi}}$ is evaluated by swapping one of the K chosen sensors with one of the $M - K$ sensors not chosen, where

$$\hat{\mathbf{\Xi}} = \left(\sum_{m=1}^M \hat{z}_m \mathbf{g}_m^* \mathbf{g}_m^T \right)^{-1}. \quad (18)$$

The computation of $\det \hat{\mathbf{\Xi}}$ for the swapped sensor positions can be reduced from $\mathcal{O}(L^3)$ to $\mathcal{O}(L^2)$ by using the low-rank update formula for the determinant of the matrix. The sensor selection $\hat{\mathbf{z}}$ is greedily updated if an increase in $\det \hat{\mathbf{\Xi}}$

Algorithm 3 Sensor Placement Based on Mutual Information

Require: Covariance matrix $\Sigma \in \mathbb{C}^{M \times M}$, number of sensors K , set of candidates \mathcal{M} ($|\mathcal{M}| = M$)

Ensure: Sensor selection $\mathcal{T} \subseteq \mathcal{M}$

$\mathcal{T} = \emptyset$

for $k = 1$ to K **do**

 Select the sensor index m

$$\hat{m} = \arg \max_{m \in \mathcal{M} \setminus \mathcal{T}} \frac{\sigma_m^2 - \Sigma_{m, \mathcal{T}} \Sigma_{\mathcal{T}, \mathcal{T}}^{-1} \Sigma_{\mathcal{T}, m}}{\sigma_m^2 - \Sigma_{m, \bar{\mathcal{T}}} \Sigma_{\bar{\mathcal{T}}, \bar{\mathcal{T}}}^{-1} \Sigma_{\bar{\mathcal{T}}, m}}$$

 Update the selected locations \mathcal{T}

$$\mathcal{T} = \mathcal{T} \cup \hat{m}$$

end for

is encountered while attempting all possible $K(M - K)$ swaps. By limiting the number of steps of the local optimization to M^3/L^2 , its total computational cost is evaluated as $\mathcal{O}(M^3)$, which is the same as the cost for solving (17).

This algorithm can be extended to the broadband case by solving the approximate problem formulated as

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} \quad \frac{1}{F} \sum_{f=1}^F \log \det \left(\sum_{m=1}^M z_m \mathbf{g}_{m,f}^* \mathbf{g}_{m,f}^T \right) \\ & \quad + \kappa \sum_{m=1}^M (\log(z_m) + \log(1 - z_m)) \\ & \text{subject to} \quad \mathbf{1}^T \mathbf{z} = K. \end{aligned} \quad (19)$$

This can also be solved by Newton's method as in the narrowband case. The computational cost for each iteration is $\mathcal{O}(M^3 F)$. The local optimization method may also be applied after obtaining the approximate solution.

C. Sensor Placement Based on Mutual Information

Krause et al. [31] have proposed a sensor placement method based on the maximization of the mutual information between the selected and unselected locations. Since this maximization problem is NP-hard, a greedy algorithm exploiting the submodularity of mutual information is used to find an approximate solution in polynomial time. At each iteration, the sensor that provides the maximum increase in the mutual information is selected and added to the list. More specifically, m that maximizes the increase in the mutual information from \mathcal{T} to $\mathcal{T} \cup \{m\}$, which is represented as $J_{\text{MI}} = \text{MI}(\mathcal{T} \cup \{m\}) - \text{MI}(\mathcal{T})$, is chosen as the next sensor \hat{m} (see Algorithm 3). By assuming a Gaussian process as the probabilistic model for spatial phenomena, this increase in the mutual information can be computed by

$$\begin{aligned} J_{\text{MI}} &= \text{MI}(\mathcal{T} \cup \{m\}) - \text{MI}(\mathcal{T}) \\ &= \frac{1}{2} \log \frac{\sigma_m^2 - \Sigma_{m, \mathcal{T}} \Sigma_{\mathcal{T}, \mathcal{T}}^{-1} \Sigma_{\mathcal{T}, m}}{\sigma_m^2 - \Sigma_{m, \bar{\mathcal{T}}} \Sigma_{\bar{\mathcal{T}}, \bar{\mathcal{T}}}^{-1} \Sigma_{\bar{\mathcal{T}}, m}}, \end{aligned} \quad (20)$$

where $\bar{\mathcal{T}}$ is $\mathcal{M} \setminus (\mathcal{T} \cup \{m\})$, Σ is the covariance matrix of the sensor observations at the candidate locations, and σ_m^2

Algorithm 4 Sensor Placement Based on FrameSense

Require: A set of candidate sensor locations \mathcal{M} , a transfer function matrix \mathbf{G} , and number of sensors K

Ensure: Sensor locations $\mathcal{T} \subseteq \mathcal{M}$

Set $\mathcal{T} = \emptyset$

Find the first two rows to eliminate

$$\mathcal{U} = \arg \min_{m, m' \in \mathcal{M}} |\langle \mathbf{g}_m, \mathbf{g}_{m'} \rangle|^2$$

Update the available locations

$$\mathcal{T} = \mathcal{M} \setminus \mathcal{U}$$

while $|\mathcal{U}| < |\mathcal{M}| - K$ **do**

Find the optimal row

$$\hat{m} = \arg \min_{m \in \mathcal{T}} J(\mathcal{U} \cup m)$$

Update the set of removed and available locations

$$\mathcal{U} = \mathcal{U} \cup \hat{m} \quad \text{and} \quad \mathcal{T} = \mathcal{T} \setminus \hat{m}$$

end while

is the diagonal element of Σ for the m th sensor. Since the logarithmic function is a monotonically increasing function, J_{MI} can be simplified to

$$J_{\text{MI}} = \frac{\sigma_m^2 - \Sigma_{m, \mathcal{T}} \Sigma_{\mathcal{T}, \mathcal{T}}^{-1} \Sigma_{\mathcal{T}, m}}{\sigma_m^2 - \Sigma_{m, \mathcal{T}} \Sigma_{\mathcal{T}, \mathcal{T}}^{-1} \Sigma_{\mathcal{T}, m}}. \quad (21)$$

The Gaussian kernel is used for modeling Σ in [31]; however, the covariance matrix Σ calculated by the predicted transfer function matrix can be used in this scenario.

The computational cost for calculating (21) is $\mathcal{O}(M^3)$ for each k and m because (21) includes the computation of the inverse matrices of $\Sigma_{\mathcal{T}, \mathcal{T}} \in \mathbb{C}^{(M-k+1) \times (M-k+1)}$ and M is much larger than k . The exhaustive search of \hat{m} requires $M - k + 1$ calculations of (21), i.e., $\mathcal{O}(M^4)$, for each k . Then, the total computational cost can be estimated as $\mathcal{O}(M^4 K)$, and the memory requirement necessary to store covariance matrix Σ is $\mathcal{O}(M^2)$.

In the broadband case, by assuming statistical independence of the spatial phenomena for frequency bins, the summation of the increase in the mutual information (20) for the target frequency bins can be used as

$$J_{\text{MI}} = \frac{1}{2} \sum_{f=1}^F \log \frac{\sigma_{m, f}^2 - \Sigma_{m, \mathcal{T}, f} \Sigma_{\mathcal{T}, \mathcal{T}, f}^{-1} \Sigma_{\mathcal{T}, m, f}}{\sigma_{m, f}^2 - \Sigma_{m, \mathcal{T}, f} \Sigma_{\mathcal{T}, \mathcal{T}, f}^{-1} \Sigma_{\mathcal{T}, m, f}}. \quad (22)$$

This algorithm requires the computation of the inverses of $\Sigma_{\mathcal{T}, \mathcal{T}, f}$ and $\Sigma_{\mathcal{T}, \mathcal{T}, f}$ for each frequency at each k , which results in a total computational cost of $\mathcal{O}(M^4 F K)$.

D. Source and Sensor Placement Based on FrameSense

FrameSense, a sensor placement method based on the minimization of the frame potential, was proposed by Ranieri et al. [33]. It was later extended to the source placement problem [34]. It is based on the minimization of the frame potential of a set of vectors from the candidate transfer function matrix.

Indeed, for a set of unit-norm vectors, the frame potential is minimized by unit-norm tight frames [32], which also minimize the mean square error of the estimation of the vector coefficients over this frame. The frame potential of \mathbf{G} is defined as

$$\text{FP}(\mathbf{G}) = \sum_{m, m'} |\langle \mathbf{g}_m, \mathbf{g}_{m'} \rangle|^2. \quad (23)$$

The frame potential acts here as a proxy for the mean square error of the estimation of the coefficients, as it can be used to bound the mean square error, with a sharpness depending on the norms of the rows of \mathbf{G} and its spectral properties.

A greedy worst-out algorithm is proposed in [33] for the sensor placement as in Algorithm 4. At each iteration, the row of \mathbf{G} that maximizes the increase in the frame potential, i.e., $J_{\text{FP}} = \text{FP}(\mathbf{G}) - \text{FP}(\mathbf{G}_{\mathcal{M} \setminus \mathcal{U}})$, is removed, where \mathcal{U} is the set of locations that are not suitable for sensing. The main contribution to the computational cost is the calculation of the frame potential of \mathbf{G} for various index sets; however, the element-wise squared Gram matrix $|(\mathbf{G}^H \mathbf{G})_{m, m'}|^2$ can be calculated first and reused at each iteration, which requires $\mathcal{O}(LM^2)$ time and $\mathcal{O}(M^2)$ memory.

Algorithm 4 can also be applied to the source placement problem by evaluating the frame potential of the columns of \mathbf{G} [34]. However, the source and sensor positions must be separately determined. The computational cost can be estimated as $\mathcal{O}(ML^2)$ in a similar manner to that in sensor selection.

Although Algorithm 4 as such can only be applied to the narrowband case, it can be extended to the broadband case by evaluating the frame potential for all the frequency bins as

$$\text{FP}(\mathbf{G}) = \sum_f \sum_{m, m'} |\langle \mathbf{g}_{m, f}, \mathbf{g}_{m', f} \rangle|^2. \quad (24)$$

Again, the greedy worst-out algorithm can be applied to minimize (24). An algorithm for the source placement can also be derived in a similar manner. The computational cost can be estimated as $\mathcal{O}(LM^2 F)$ and $\mathcal{O}(ML^2 F)$ for sensor and source selection, respectively.

E. Joint Source and Sensor Placement Based on Empirical Interpolation Method

The authors of the present contribution have recently proposed a joint source and sensor placement method based on the empirical interpolation method (EIM) [35]. EIM was first introduced in [75] in the context of numerical analysis of partial differential equations by the reduced basis method [76]. EIM was presented as a general interpolation procedure in [77] and applied to microphone array design in [24].

EIM jointly selects K interpolation functions and K sampling points from candidates such that the linear system to be solved for the interpolation remains stable when K grows. By regarding the transfer functions of the secondary sources and control points inside Ω as the interpolation functions and sampling points to approximate a desired sound field inside Ω , EIM can be applied to the source and sensor placement problem.

Algorithm 5 Joint Source and Sensor Placement Based on EIM

Require: A set of candidate points \mathcal{L} and \mathcal{M} , transfer function matrix $\mathbf{G} \in \mathbb{C}^{M \times L}$, and target error tolerance ϵ_{tol}

Ensure: Source and sensor locations $\mathcal{S} \subseteq \mathcal{L}$ and $\mathcal{T} \subseteq \mathcal{M}$, and their required number K

Set $k = 1$, $\mathcal{S} = \emptyset$, and $\mathcal{T} = \emptyset$

while $\epsilon > \epsilon_{\text{tol}}$ **do**

 Select the secondary source index

$$l_k = \arg \max_{l \in \mathcal{L} \setminus \mathcal{S}} \|\mathbf{G}_{\mathcal{M},l} - I_{k-1}[\mathbf{G}_{\mathcal{T},l}]\|_{\infty}$$

 and the corresponding index of the control point

$$m_k = \arg \max_{m \in \mathcal{M} \setminus \mathcal{T}} |\mathbf{G}_{m,l_k} - (I_{k-1}[\mathbf{G}_{\mathcal{T},l_k}])_m|.$$

 Define the next basis function by

$$\mathbf{h}_k = \frac{\mathbf{G}_{\mathcal{M},l_k} - I_{k-1}[\mathbf{G}_{\mathcal{T},l_k}]}{\mathbf{G}_{m_k,l_k} - (I_{k-1}[\mathbf{G}_{\mathcal{T},l_k}])_{m_k}}$$

 Update the set of selected locations

$$\mathcal{S} = \mathcal{S} \cup l_k \quad \text{and} \quad \mathcal{T} = \mathcal{T} \cup m_k$$

 Define the error by

$$\epsilon = \max_{l \in \mathcal{L}} \|\mathbf{G}_{\mathcal{M},l} - I_{k-1}[\mathbf{G}_{\mathcal{T},l}]\|_2.$$

 Set $k = k + 1$

end while

EIM is described in Algorithm 5; at each step, a new source and a new control point are identified. At iteration k , the new source is chosen as the one maximizing the ℓ_{∞} -norm between its transfer function $\mathbf{G}_{\mathcal{M},l}$ and its interpolation $I_{k-1}[\mathbf{G}_{\mathcal{T},l}] \in \mathbb{C}^M$ using a linear combination of the $k-1$ previously identified sources, fitted on set \mathcal{T} of $k-1$ control points. The point where this interpolation error is maximal is added to set \mathcal{T} . This procedure is iterated until the ℓ_2 -norm of the error between each column of the transfer function matrix and its interpolation becomes smaller than the predefined target error tolerance ϵ_{tol} . It is also possible to stop the iteration when k reaches a predefined number of sources and sensors K .

An auxiliary basis $\{\mathbf{h}_1, \dots, \mathbf{h}_K\}$ of the space spanned by the identified sources is built iteratively to compute the interpolations, given by

$$I_k[\mathbf{G}_{\mathcal{T},l}] = \sum_{j=1}^k c_{lj} \mathbf{h}_j, \quad (25)$$

where the interpolation $I_k[\mathbf{G}_{\mathcal{T},l}]$ satisfies

$$(I_k[\mathbf{G}_{\mathcal{T},l}])_{\mathcal{T}} = \mathbf{G}_{\mathcal{T},l}. \quad (26)$$

Therefore, the coefficients c_{lj} in (25) are obtained by solving

$$\mathbf{G}_{\mathcal{T},l} = \mathbf{H}_{\mathcal{T}}^{(k)} \mathbf{c}_l \quad (27)$$

for \mathbf{c}_l , i.e., the vectors of coefficients c_{lj} . By construction, the matrix $\mathbf{H}_{\mathcal{T}}^{(k)}$ is lower triangular, and its inversion can be

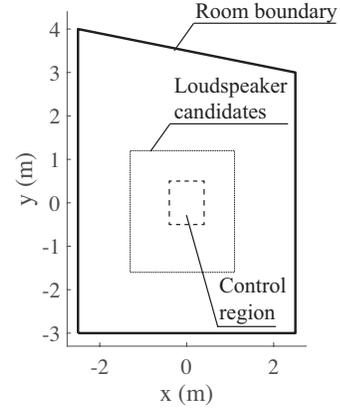


Fig. 2. Experimental setting of the 2D simulations. The absorption ratio of the room boundary for FEM is set at 0.50.

iteratively computed with a total cost of $\mathcal{O}(K^3)$. The computational cost for each iteration is dominated by that of (25), whose cost is $\mathcal{O}(kML)$. Therefore, the total computational cost for the joint source and sensor selection can be estimated as $\mathcal{O}(K^2ML)$.

When sound field control is necessary over a broad frequency band, the transfer function for the input of Algorithm 5 becomes a third-order tensor including the dimension of frequency, i.e., $\mathbf{G} \in \mathbb{C}^{M \times L \times F}$. Algorithm 5 can be applied by evaluating the ℓ_{∞} -norm for the matrix $\mathbf{G}_{\mathcal{M},l,\mathcal{F}}$ and vector $\mathbf{G}_{m,l_k,\mathcal{F}}$ in the first and second lines, respectively. The total computational cost becomes $\mathcal{O}(K^2MLF)$. Memory usage is dominated by \mathbf{G} .

IV. EXPERIMENTS

In this section, 2D numerical simulations are conducted for the comparative evaluation of the five methods introduced in Sect. III. A 2D trapezoidal room is assumed, which is depicted by the bold line in Fig. 2. Acoustic simulations were performed using FreeFem++ [78], a finite element method (FEM) solver. The size of the elements for FEM is determined to be 0.1 times the simulated wavelength. The specific acoustic impedance ratio of each wall is set at 5.83 for all the frequencies and corresponds to an absorption ratio of 0.50. Candidate loudspeaker positions are set along the boundary of a rectangular region of dimensions 2.4 m \times 2.8 m. The boundary is regularly discretized into 256 candidate positions. The control region is set as a rectangular region of dimensions 0.8 m \times 1.0 m, discretized with intervals of 0.04 m to construct candidate control points. The regions of candidate loudspeakers and control points are depicted by the dotted and dashed lines, respectively, in Fig. 2.

The source placement method based on Gram–Schmidt orthogonalization (Sect. III-A), the sensor placement methods based on the determinant (Sect. III-B) and mutual information (Sect. III-C), and the source and sensor placement methods based on FrameSense (Sect. III-D) and EIM (Sect. III-E) are denoted as **GSO**, **Det**, **MI**, **FS**, and **EIM**, respectively. We also consider both regular and random placement methods, which are denoted as **Reg** and **Rand**, respectively. In **Reg**, the

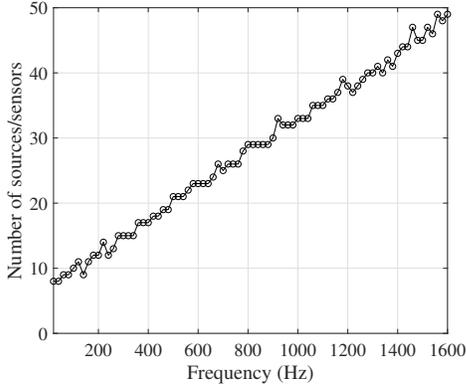


Fig. 3. Number of sources and sensors K at each frequency obtained by **EIM** with $\epsilon_{\text{tol}} = 1.0 \times 10^{-2}$ for narrowband case.

loudspeaker and control point positions are chosen from the candidates on the rectangular boundaries so that the intervals between elements are as equal as possible. In **Rand**, the loudspeaker locations are drawn from the uniform measure on ∂D , and the control points are drawn from the uniform measure in Ω . The control points for **GSO** are regularly set on the two-layer boundary of the control region to avoid the forbidden frequency problem. The interval between these two layers is 0.04 m. The loudspeaker placements in **Det** and **MI** are the same as that in **Reg**.

To evaluate the control accuracy of each placement, we define the signal-to-distortion ratio (SDR) as

$$\text{SDR}(\omega) = 10 \log_{10} \frac{\int_{\Omega} |u_{\text{des}}(\mathbf{x}, \omega)|^2 d\mathbf{x}}{\int_{\Omega} |u_{\text{syn}}(\mathbf{x}, \omega) - u_{\text{des}}(\mathbf{x}, \omega)|^2 d\mathbf{x}}, \quad (28)$$

where $u_{\text{des}}(\cdot)$ and $u_{\text{syn}}(\cdot)$ are the desired and synthesized pressure fields, respectively. The desired sound field is a plane-wave field and its arrival angle is varied from 0 to 359 deg at 1 deg intervals. Then, SDRs for all the plane-wave angles are averaged. We also evaluate the filter stability using the condition number of $\mathbf{G}_{\mathcal{S}, \mathcal{T}}$ as

$$\kappa(\mathbf{G}_{\mathcal{S}, \mathcal{T}}) = \frac{\sigma_{\max}(\mathbf{G}_{\mathcal{S}, \mathcal{T}})}{\sigma_{\min}(\mathbf{G}_{\mathcal{S}, \mathcal{T}})}, \quad (29)$$

where $\sigma_{\max}(\mathbf{G}_{\mathcal{S}, \mathcal{T}})$ and $\sigma_{\min}(\mathbf{G}_{\mathcal{S}, \mathcal{T}})$ denote the maximum and minimum singular values of $\mathbf{G}_{\mathcal{S}, \mathcal{T}}$, respectively.

We first compare the reproduction performance of plane waves in the narrowband case. A more realistic broadband scenario is then considered.

A. Narrowband case

In the narrowband case, the numbers and locations of the loudspeakers and control points are determined at each frequency. The number of elements K is determined by **EIM** with $\epsilon_{\text{tol}} = 1.0 \times 10^{-2}$ since all the other methods require K for the placement. The driving signals of the loudspeakers are obtained using (7) without regularization to avoid the determination problem of the regularization parameter and make the sensitivity of the solution clearer.

Fig. 3 shows the number of loudspeakers and control points K determined by **EIM** at each frequency. The SDR and

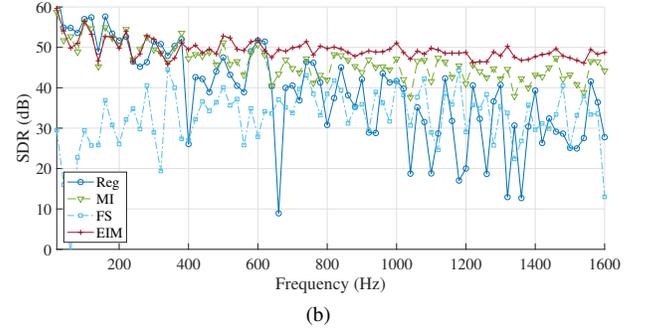
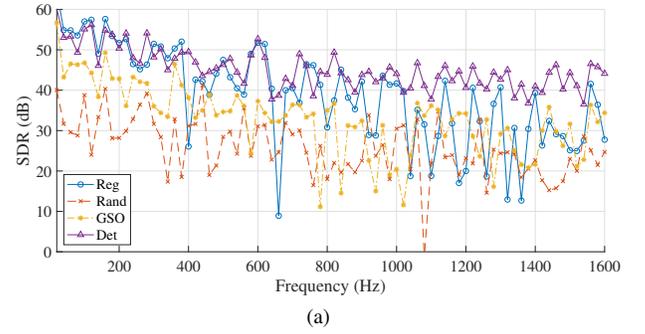


Fig. 4. SDR with respect to frequency for evaluating control accuracy in narrowband case.

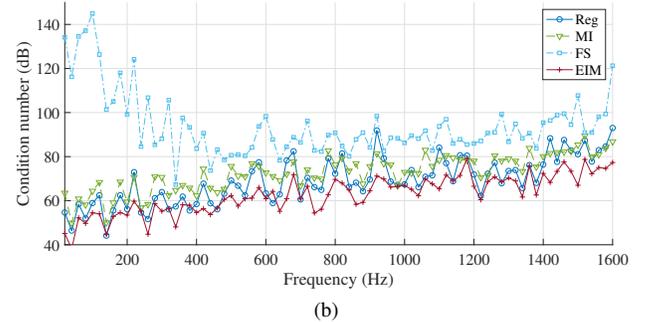
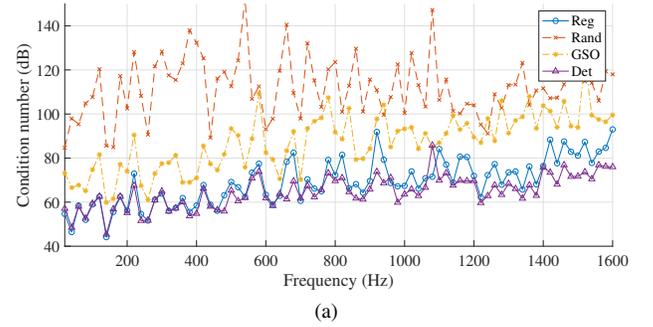


Fig. 5. Condition number with respect to frequency for evaluating filter stability in narrowband case.

condition number with respect to the frequency are plotted in Figs. 4 and 5, respectively. Although the SDRs of **Reg** are high at low frequencies, they have many dips, e.g., at 400 Hz and 660 Hz, owing to the forbidden frequency problem. The placements of **Rand** are randomly determined at each frequency; therefore, the SDRs of **Rand** greatly fluctuate and

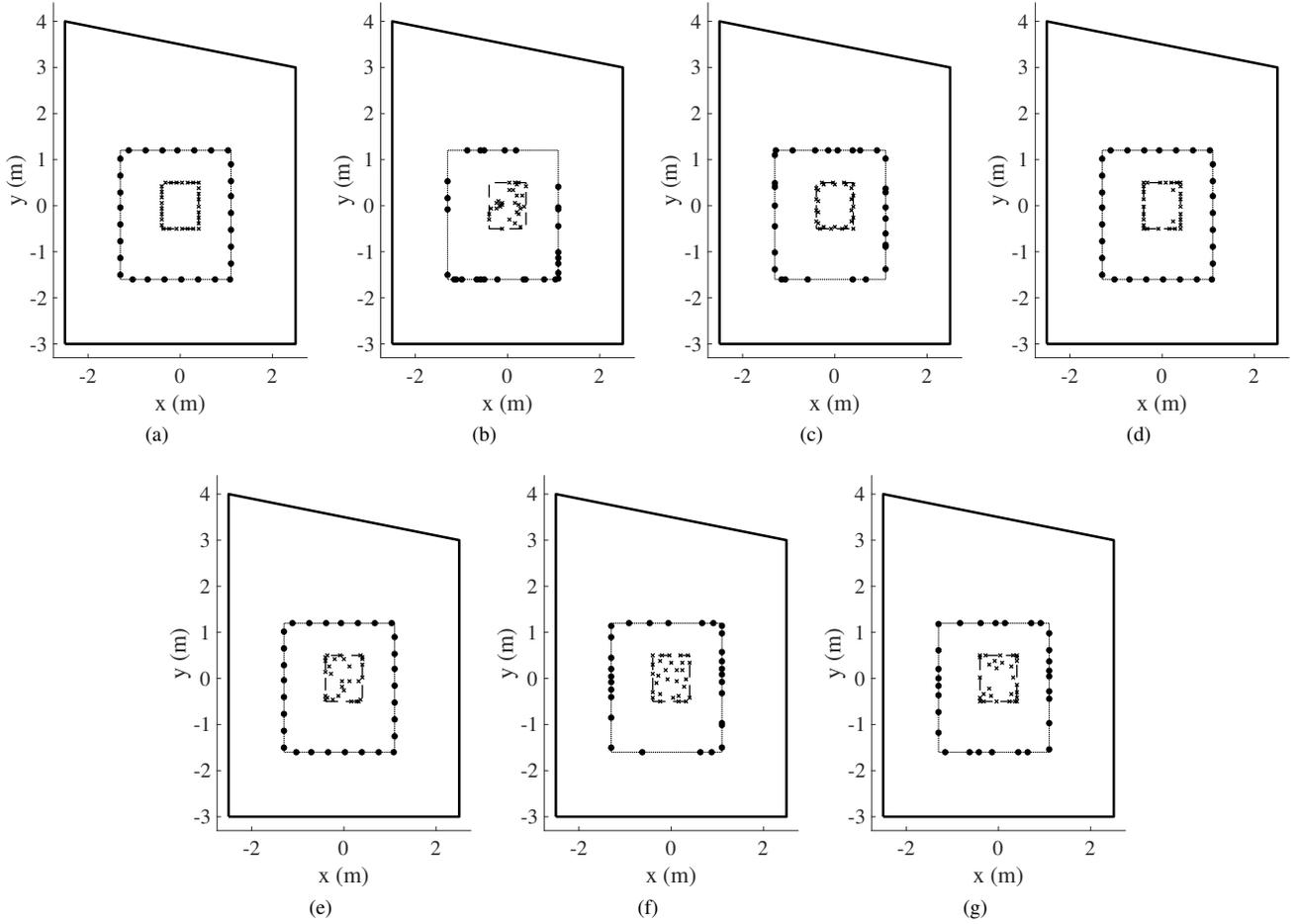


Fig. 6. Selected secondary source and sensor positions at 800 Hz in narrowband case. Black dots and crosses represent selected secondary sources and control points, respectively. (a) **Reg**; (b) **Rand**; (c) **GSO**; (d) **Det**; (e) **MI**; (f) **FS**; (g) **EIM**.

are generally low. Compared with the SDRs of **Reg** and **Rand**, those of **GSO** and **FS** are relatively high, but they still have several dips. Generally high SDRs are achieved by **Det**, **MI**, and **EIM**, and the SDRs of **EIM** are higher than those of **Det** and **MI**. The condition numbers of **Rand** are extremely high compared with those of the other methods. Those of **FS** are also high at low frequencies. The lowest condition numbers are achieved by **Det** and **EIM**.

As an example, the selected positions of loudspeakers and control points at 800 Hz are shown in Fig. 6. In **Det**, **MI**, **FS**, and **EIM**, relatively large numbers of control points are placed on the boundary of the control region. This tendency is strong for **Det**. Indeed, the sound field inside a closed region is determined by the pressure on the boundary except at forbidden frequencies. It can be considered that the control points inside the control region have the effect of preventing nonuniqueness at these frequencies. This effect was demonstrated in [79] for a circular domain, where the stability of the reconstruction with a number of measurements linear in the number of parameters describing the sound field necessitates most of the samples to be on the border and some samples to be inside the domain. In this sense, the two layer placement of control points used in **GSO** is not suitable because the number of sampling points on the boundary becomes small. Furthermore, closely located control points can lead to an excessively high

condition number. The synthesized pressure and normalized error distributions of each method for the plane-wave arrival angle of 219 deg are shown in Figs. 7 and 8, respectively. The normalized error distribution is computed as

$$\text{Err}(\mathbf{x}, \omega) = 10 \log_{10} \frac{|u_{\text{syn}}(\mathbf{x}, \omega) - u_{\text{des}}(\mathbf{x}, \omega)|^2}{|u_{\text{des}}(\mathbf{x}, \omega)|^2}. \quad (30)$$

The frequency is close to the forbidden frequency of the target region; therefore, the control accuracy inside the region is low in **Reg** although that on the boundary is high. The region of high control accuracy is relatively broad in **EIM**.

The boxplot of the output power of the driving signals at 800 Hz is shown in Fig. 9. The maximum whisker length is set as 1.5 times the interquartile range. The red crosses denote outliers. When the filter stability is low, the amplitude of the loudspeakers can be very large and it can be impractical for output using ordinary loudspeakers. Compared with the output power of **Reg**, those of **Rand**, **GSO**, and **FS** are relatively large. Those of **Det** and **MI** are almost the same as that of **Reg**. A relatively small output power is achieved by **EIM**, which corresponds to the low condition number of the transfer function matrix.

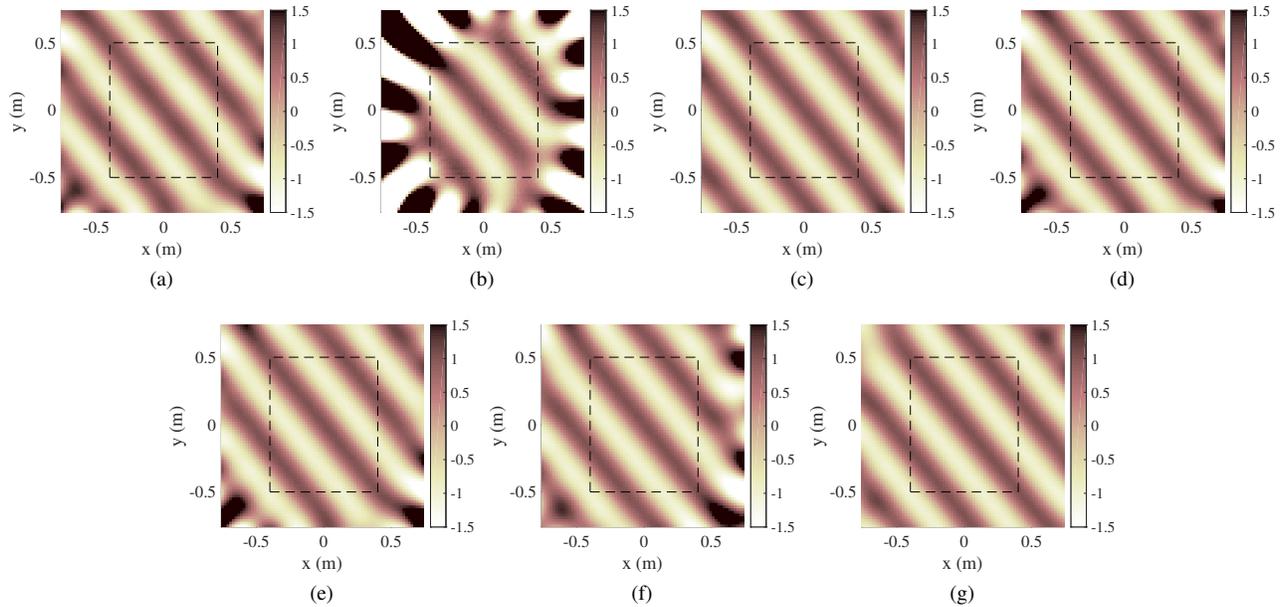


Fig. 7. Synthesized pressure fields of plane wave from 219 deg at 800 Hz in narrowband case. (a) **Reg**; (b) **Rand**; (c) **GSO**; (d) **Det**; (e) **MI**; (f) **FS**; (g) **EIM**.

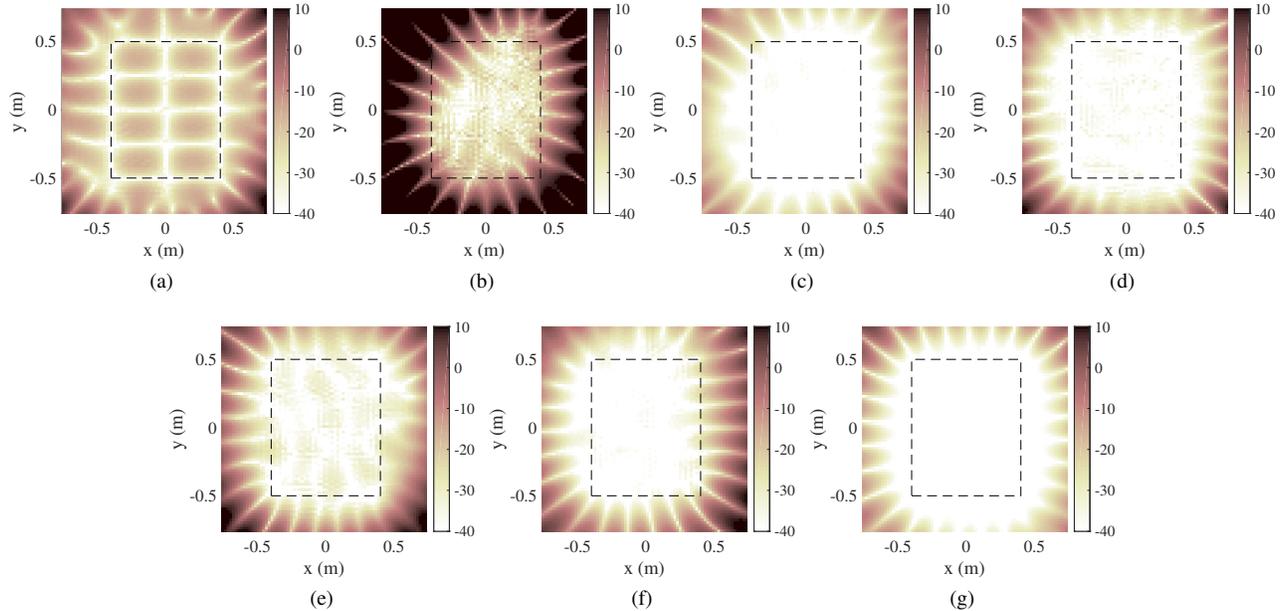


Fig. 8. Normalized error distributions at 800 Hz in narrowband case. (a) **Reg**; (b) **Rand**; (c) **GSO**; (d) **Det**; (e) **MI**; (f) **FS**; (g) **EIM**. SDRs are 21.8, 10.0, 43.9, 40.4, 35.4, 38.1, and 53.2 dB, respectively.

B. Further investigation of regular placement

Although the regular placement is simple and straightforward, from Fig. 4, it can be observed that **Reg** does not perform as well as **Det**, **MI**, or **EIM**, particularly at forbidden frequencies, e.g., at 660 Hz. The effect of forbidden frequencies can be alleviated by placing several control points inside the target region, and the general performance of **Reg** can be improved. However, no obvious solution exists for the choice of the number of internal control points and their positions. In this section, different strategies for the placement of additional internal points are explored in the regular case, in order to investigate the individual effects of the internal points and the overall placement of the control points.

Takane et al. [68] proposed several empirical strategies to place internal control points, which is inspired by CHIEF (combined Helmholtz integral equation formulation) points in the boundary element method [80], [81]. In the 2D sound field control, it is suggested to place three control points around the border or six control points apart from the border of the target region so that the sound intensity at an empirically chosen local position is controlled in addition to the pressure distribution on the boundary. Sensor placement methods introduced in Sect. III are also applicable to selecting additional control points. We here evaluate the performance of the method proposed by Takane et al. [68] (**Reg+CHIEF**) and **EIM** for selecting control points in addition to the regular

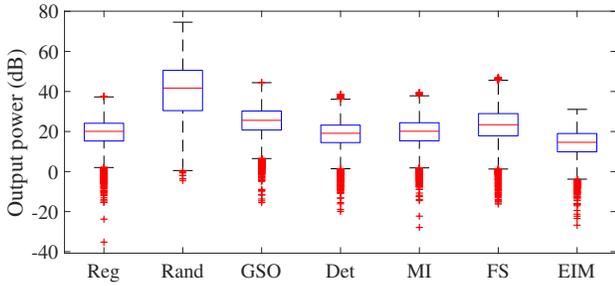


Fig. 9. Boxplot of output power of loudspeakers at 800 Hz in narrowband case. The red crosses denote outliers. The excessively large output is considered to originate from the unstable inverse filter.

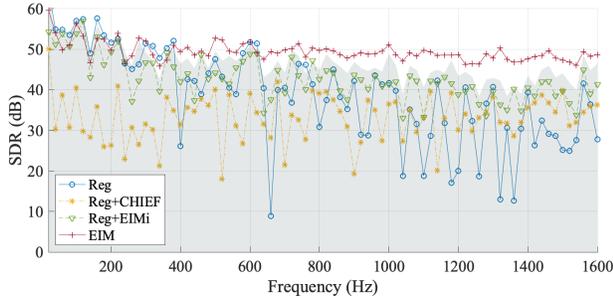


Fig. 10. Regular placement with six additional control points selected by **Reg+CHIEF** and **Reg+EIMi** is evaluated in narrowband case. SDR is plotted with respect to frequency. The gray area indicates approximate performance bound obtained by the random sampling of six additional control points with 10^6 realizations.

placement on the boundary (**Reg+EIMi**). The reason for the choice of **EIM** for selecting additional control points is its good performance in the narrowband case. The number of additional control points is set as $K_{\text{int}} = 6$ in accordance with one of the suggestions in [68]. In both methods, $K - K_{\text{int}}$ control points are placed at regular intervals along the border. As shown in Fig. 11(a), three pairs of control points are placed around the center of the target region at intervals of 0.08 m in **Reg+CHIEF**. The interval between a pair of control points is 0.04 m. In **Reg+EIMi**, K_{int} control points are placed by running **EIM** on the basis of the space of sound fields spanned by the sources that have a value of 0 on the regularly placed control points. To obtain approximate performance bounds for the regular placement with additional control points, regular placement on the border with random selections of $K_{\text{int}} = 6$ interior control points is tested and performed 10^6 times.

Fig. 10 shows the SDR with respect to the frequency of **Reg+CHIEF** and **Reg+EIMi** with the performance range of random sampling indicated by the gray area. The results of **Reg** and **EIM** are also shown for reference. Although several sharp dips of SDR are alleviated by **Reg+CHIEF**, the general performance of this method is lower than those of the other methods. The performance range of random sampling is significantly broad; therefore, the performance of the regular placement with additional control points highly depends on the placement of additional points. The SDR of **Reg+EIMi** is fairly close to the upper limit; therefore, **Reg+EIMi** is a reasonable choice for the placement of additional control

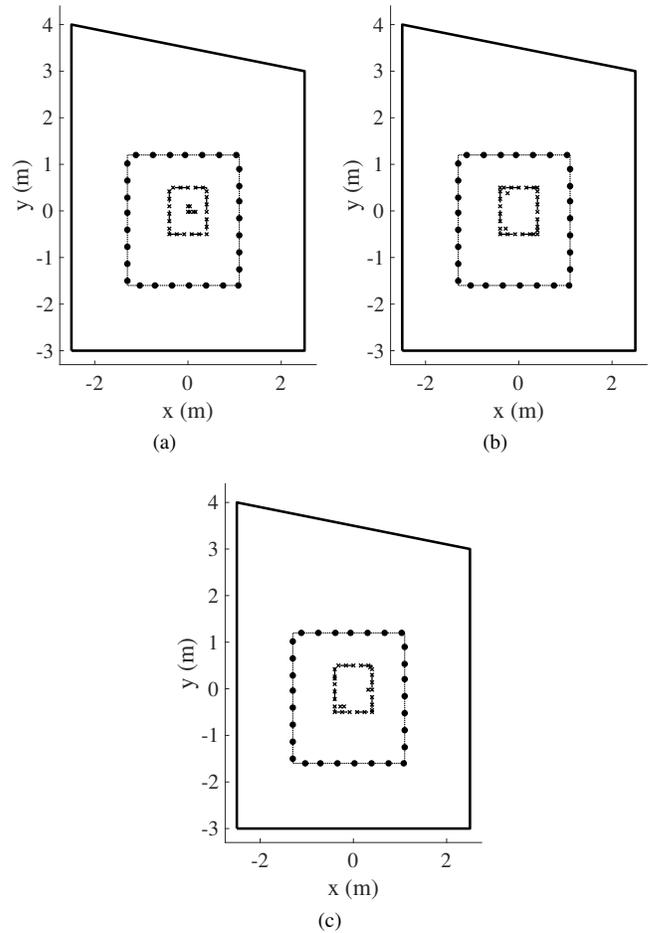


Fig. 11. Secondary source and sensor positions of **Reg+CHIEF**, **Reg+EIMi**, and the best placement of the random sampling for six additional control points at 800 Hz in narrowband case. Black dots and crosses represent selected secondary sources and control points, respectively. (a) **Reg+CHIEF**; (b) **Reg+EIMi**; (c) the best placement of random sampling of additional control points.

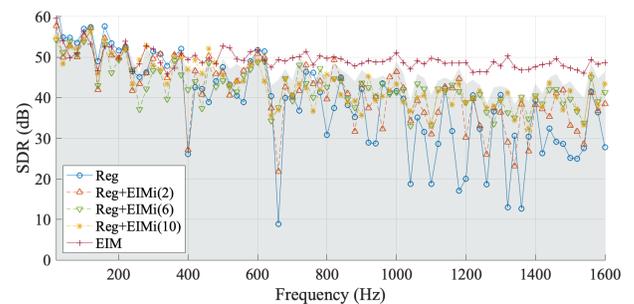


Fig. 12. SDRs of **Reg+EIMi** using 2, 6, and 10 additional control points are compared. The number shown in parentheses indicates the number of additional control points for **Reg+EIMi**.

points. However, the performance of **Reg+EIMi** is still lower than that of **EIM** in general. Note that the best placement of random sampling with respect to SDR does not necessarily indicate a low condition number. In Figs. 11(b) and (c), the placement of **Reg+EIMi** and the best placement of random sampling at 800 Hz are shown. It can be observed that, in both cases, the six additional control points selected by **EIM** are not

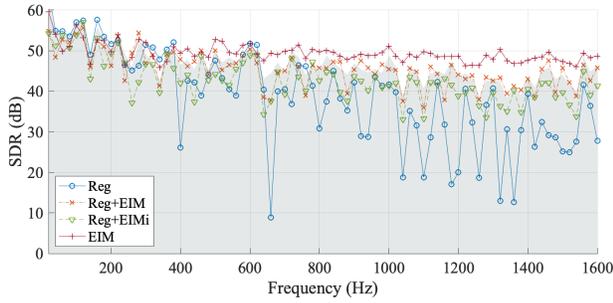


Fig. 13. SDRs of **Reg+EIM** and **Reg+EIMi** using 6 additional control points are compared.

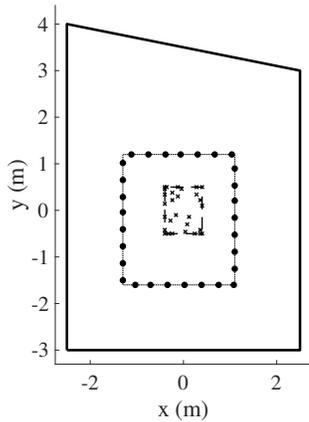


Fig. 14. Secondary source and sensor positions of **Reg+EIM** at 800 Hz in narrowband case.

necessarily placed inside the target region, which implies that the gap in performances between **Reg** and the other methods is not only caused by the absence of interior points in **Reg**.

We also investigate the effect of the number of additional control points. In Fig. 12, the SDR with respect to frequency for **Reg+EIMi** at $K_{\text{int}} = 2, 6,$ and 10 is plotted with the performance range of random sampling at $K_{\text{int}} = 6$. The number shown in parentheses indicates the number of additional control points. **Reg+EIMi** with two additional control points still suffers from the problem of forbidden frequencies, e.g., at 660 Hz. The sharp decrease in control accuracy is prevented by increasing the number of additional control points. The difference between the results of placing six and ten additional control points is small. Therefore, it is necessary to place a sufficient number of additional control points. As shown in Fig. 11(b), several control points can be placed at positions on the boundary.

From the results in Figs. 10 and 12, it can be confirmed that the performance of the regular control-point placement on the boundary of the target region can be significantly improved by placing additional control points by **EIM**. However, there is still a gap in the control accuracy between **Reg+EIMi** and **EIM**. To investigate where this gap comes from, we compare **EIM**, **Reg+EIMi** with six additional control points, and the regular loudspeaker placement with the sensor placement by **EIM** from all the candidates inside the target region (**Reg+EIM**). Note that as K_{int} approaches K ,

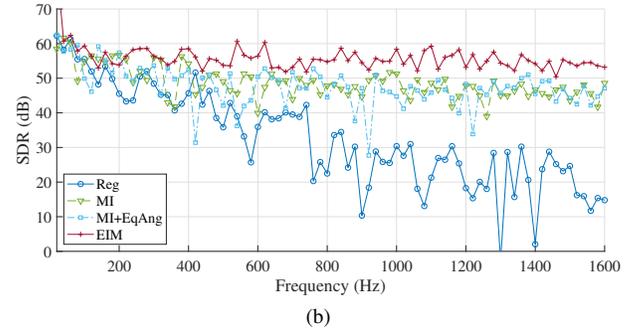
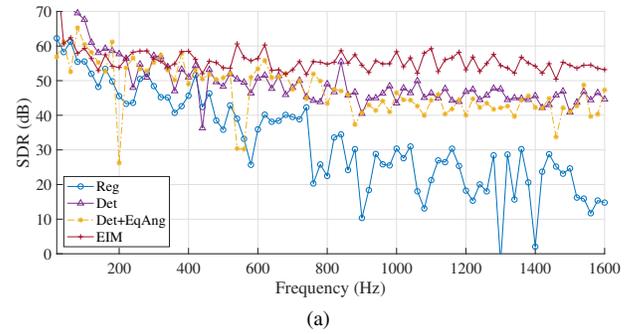


Fig. 15. SDR with respect to frequency for the case that the sizes of the target region and region of loudspeaker candidates are comparable. **Det+EqAng** and **MI+EqAng** denote the equiangular placement of loudspeakers with its origin at the center of the target region (**EqAng**) combined with the control-point placement of **Det** and **MI**, respectively.

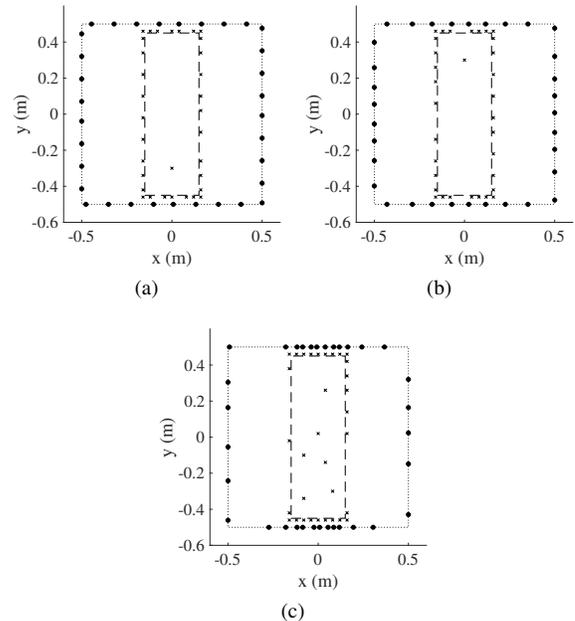


Fig. 16. Secondary source and sensor positions of **Det**, **Det+EqAng**, and **EIM** at 800 Hz in narrowband case. Black dots and crosses represent selected secondary sources and control points, respectively. The region of the candidates is enlarged. (a) **Det**; (b) **Det+EqAng**; (c) **EIM**.

Reg+EIMi becomes similar to **Reg+EIM**. Fig. 13 shows the SDRs of **Reg**, **Reg+EIM**, **Reg+EIMi**, and **EIM** with respect to frequency. Again, the approximate performance bound for the regular placement with $K_{\text{int}} = 6$ control points obtained by random sampling is indicated by the gray area. It can

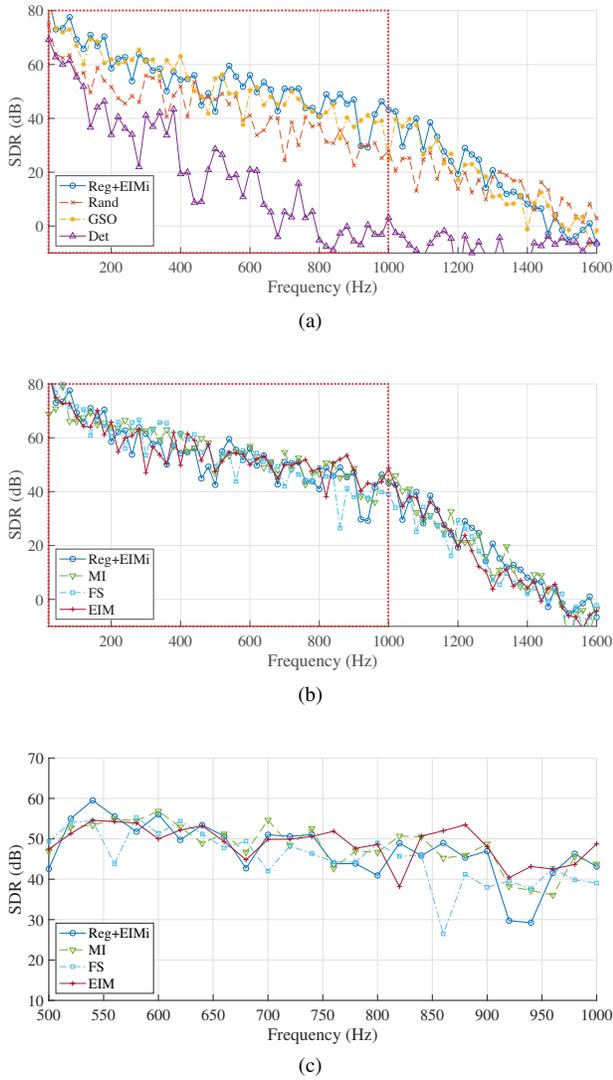


Fig. 17. SDR with respect to frequency in broadband case. The frequency range used in the source and/or sensor placement is indicated by red dashed line. Results of **Reg+EIM**, **MI**, **FS**, and **EIM** for 500–1000 Hz are enlarged in (c).

be observed that the SDR of **Reg+EIM** is generally higher than that of **Reg+EIMi**, and even higher than the upper limit of random sampling at several frequencies. Therefore, the control accuracy can be improved by optimizing the control-point placement on the border. The placement by **Reg+EIM** is shown in Fig. 14. The control accuracy can be significantly improved by jointly optimizing the placement of sources and sensors by **EIM**.

An alternative to the placement of secondary sources at regular distances along the border is placement at regular *angles* with respect to the center of the target region. These heuristics are based on a far-field approximation, where the sound field generated by a secondary source is similar to a plane wave. The regular placement in angle (i.e., equiangular placement) would minimize the maximal gap in the angle between a plane wave from a given direction and the closest source. Figs. 6(f) and (g) show that these heuristics are reasonable in the geometry considered in this case, where the target region is

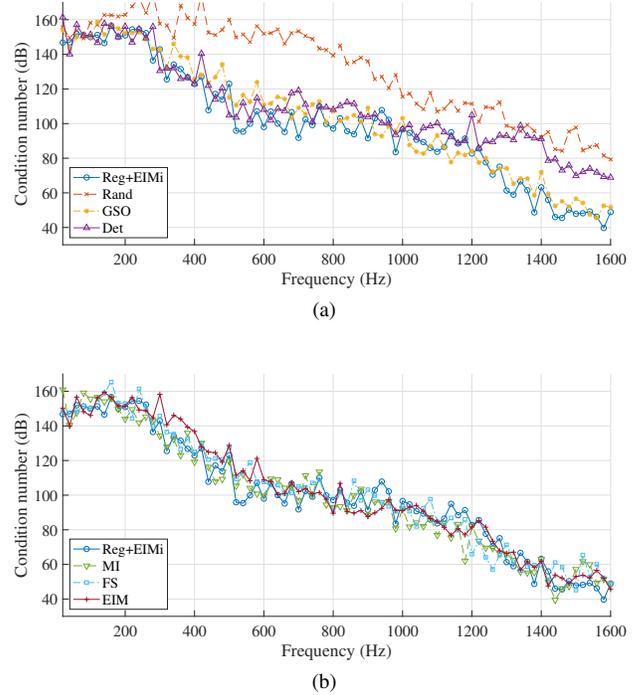


Fig. 18. Condition number with respect to frequency in broadband case. The frequency range used in the source and/or sensor placement is indicated by red dashed line.

smaller than the region of loudspeaker candidates. Indeed, the equiangular placement of the sources (**EqAng**) would yield a similar configuration to those of **FS** and **EIM**. However, in different geometries, performances of equiangular placement are not always close to those of optimized placement of the sources, or even of regular placement, especially when the sizes of the target region and region of the loudspeaker candidates are comparable. Fig. 15 is SDR with respect to frequency when the target region has dimensions of $0.3 \text{ m} \times 0.9 \text{ m}$, and the loudspeaker candidates are on the boundary of the rectangular region with dimensions of $1.0 \text{ m} \times 1.0 \text{ m}$. The target region and loudspeaker region share the same center. **EqAng** combined with the control-point placement by **Det** and **MI** (**Det+EqAng** and **MI+EqAng**) is compared with **Reg**, **Det**, **MI**, and **EIM**. The regular loudspeaker placement by **Det** and **MI** performs better than the equiangular loudspeaker placement by **Det+EqAng** and **MI+EqAng** at many frequencies. In addition, the jointly optimized placement of the sources and control points (**EIM**) performs better than regular and equiangular source placement with optimized control-point placement. The positions of loudspeakers and control points selected by **Det**, **Det+EqAng**, and **EIM** at 800 Hz are shown in Fig. 16. Here, the region of the candidates is enlarged.

Furthermore, in some practical situations, the target region can be separated into multiple regions, i.e., multizone control [7], [82], [83], and the loudspeaker candidates cannot fully enclose the target region, e.g., when loudspeakers cannot be installed on several walls. In such cases, the optimal loudspeaker placement can be complicated. Therefore, it is difficult to find a simple strategy of placing loudspeakers

without optimal placement methods.

C. Broadband case

In most practical applications, sound field control is needed over a broad frequency band. We compare the seven placement methods in the broadband case. The transfer functions from 20 to 1000 Hz at intervals of 20 Hz are used for the source and/or sensor placement. The numbers of loudspeakers and control points are both 35, as determined by **EIM** with $\epsilon_{\text{tol}} = 1.0 \times 10^{-2}$. The other experimental settings are the same as those in the narrowband case. We here replace **Reg** with **Reg+EIMi**, i.e., the regular placement with six additional control points selected by **EIM**, to avoid the severe effect of the forbidden frequency problem.

Figs. 17 and 18 show the SDRs and condition numbers of the methods with respect to frequency, respectively. The red dashed line indicates the frequency range used in the source and/or sensor placement. The SDRs of **Reg+EIMi**, **MI**, **FS**, and **EIM** are enlarged for frequencies of 500–1000 Hz in Fig. 17(c). Since six additional control points are placed in **Reg+EIMi**, large dips of the SDR are avoided. The performance of **Det** significantly deteriorates in the broadband case. Although the SDRs of **Rand** and **GSO** do not have large dips, they are generally low. The general performance of **FS** in the broadband case is improved compared with that in the narrowband case. Relatively high SDRs are achieved by **Reg+EIMi**, **MI**, **FS**, and **EIM**. The condition number is very high at low frequencies in all the methods because of high redundancy. Among them, **Rand** has an extremely high condition number. The condition numbers of the other methods are comparable, but those of **GSO** and **Det** are relatively high.

The selected positions of loudspeakers and control points are shown in Fig. 19. Similarly to the narrowband case, in **Det**, **MI**, **FS**, and **EIM**, relatively large numbers of control points are placed on the boundary of the target region. The number of control points in the interior target region in the broadband case is increased in **EIM** and decreased in **Det**, **MI**, and **FS** compared with that in the narrowband case. In **Det**, all the control points are placed on the boundary. As an example, the synthesized pressure and normalized error distributions for the plane-wave arrival angle of 219 deg are plotted in Figs. 20 and 21, respectively.

In practice, the actual transfer functions between loudspeakers and control points will be different from those measured or predicted by the acoustic numerical simulation used for source and/or sensor placement. To simulate this mismatch of the transfer functions and evaluate the robustness against it, we evaluate the control accuracy when the positions of the loudspeakers and control points include some errors. At the control stage, position errors are added, which are drawn from a Gaussian distribution with a standard deviation of 0.01 m. The inverse filters are calculated with regularization (8). The regularization parameter λ is chosen within the interval $[10^{-6}, 10]$, discretized into 100 values on a logarithmic scale, so that the highest SDR is achieved. The SDRs are averaged for 10 realizations of position errors. The average SDR with respect to frequency is plotted in Fig. 22. Again,

the red dashed line indicates the frequency range used in the source and/or sensor placement. The results of **Reg+EIMi**, **MI**, **FS**, and **EIM** are enlarged for frequencies of 500–1000 Hz in Fig. 22(c). Several dips of the SDR appear in **Reg+EIMi**. The SDRs of **Rand**, **GSO**, and **Det** are generally lower than those of the other methods. The results of the other methods are comparable, but the SDR of **EIM** is slightly higher than those of the other methods, especially at frequencies above 600 Hz.

V. CONCLUSION

For ideal sound field control, the discretization of the boundary integrals involved in the pressure matching method based on the single-layer potential equation necessitates fine discretization of the boundary of the domain, which, unfortunately, often requires a high number of sources (loudspeakers) and control points (microphones), and in practice one must deal with limited hardware resources. The challenge is to mitigate this reduction in the number of loudspeakers/microphones while maintaining good fidelity. Since the locations of both source and control points have a large impact on the performance, these positions must be optimized for specific environments and geometries. In this paper, we described, in a unified fashion, five source and/or sensor placement methods for sound field control and provided a source code for reproducibility.

Two-dimensional numerical simulation results indicate that, if internal control points are used, all the methods are able to avoid the so-called forbidden frequency problem in the narrowband case. The best performance was obtained by the joint selection of the sources and control points with the empirical interpolation method. The regular placement is simple, but necessitates placing the control points inside the target region in addition to the regular sensor placement on the boundary to deal with instability problems at forbidden frequencies. Even when instabilities are avoided, the reproduction performances of regular placements of control points and loudspeakers remain lower than those of optimized placements. Therefore, optimizing source and sensor placements in accordance with the specific geometry of the reproduction domain and the possible positions of the loudspeakers is surely a better strategy than empirical methods. Although the performances of the methods were comparable in the broadband case, the empirical interpolation method exhibited good robustness with respect to errors in the positions of loudspeakers and control points. Still, ensuring robustness over a wide frequency range remains a difficult problem that deserves further work. Regarding computational complexity, measured as how the number of operations and the memory requirements scale with the number of potential sources and control points, it can be noticed that the Gram–Schmidt orthogonalization and the empirical interpolation method scale linearly, whereas the other methods all have higher complexities. In practice, picking up sources and sensors one by one avoids the combinatorial search of the global optimum for all sources and sensors simultaneously, and all these methods can be run on standard workstations with a reasonable number of potential sources and control points, at least in the 2D case.

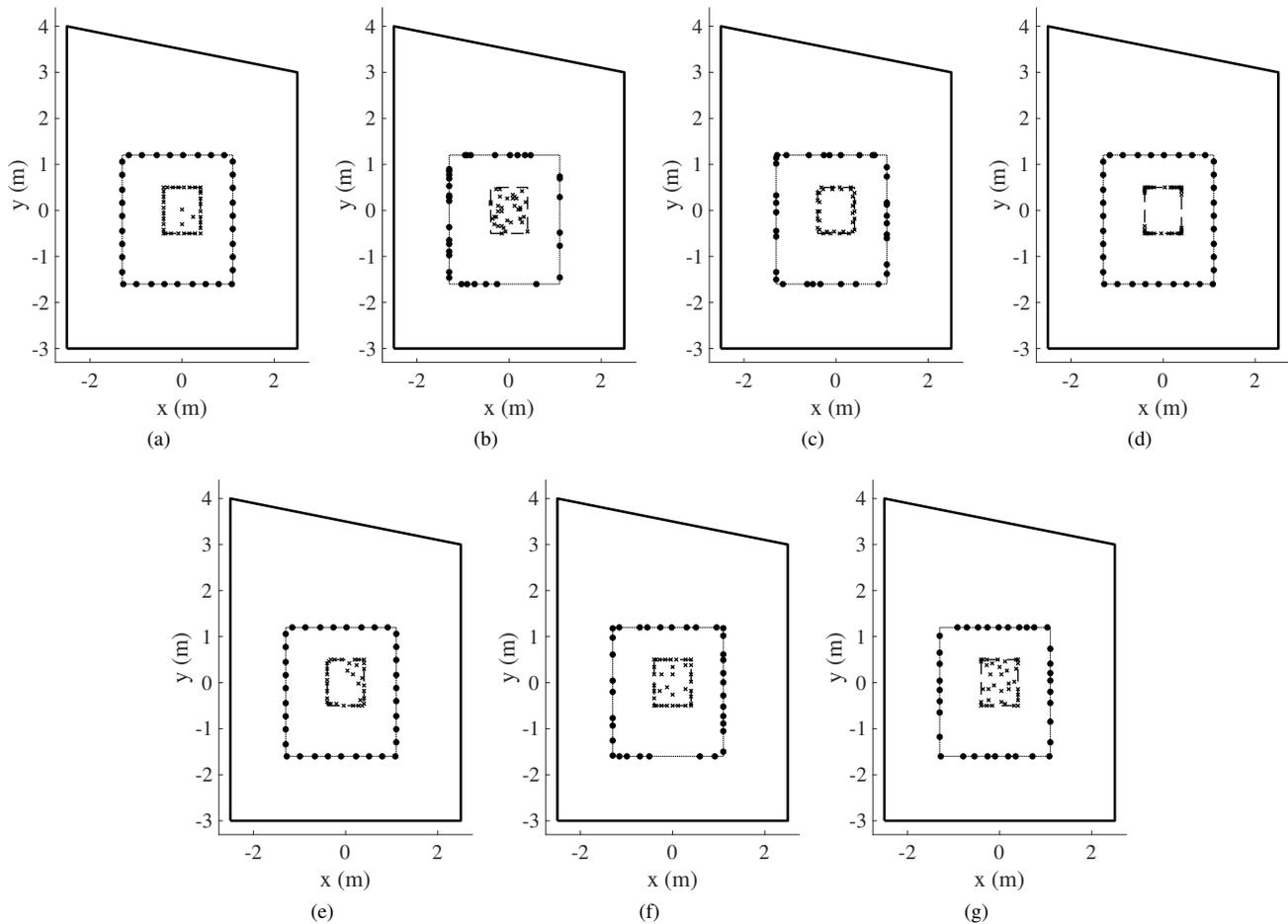


Fig. 19. Selected secondary source and sensor positions in broadband case. Black dots and crosses represent selected secondary sources and control points, respectively. (a) **Reg+EIMi**; (b) **Rand**; (c) **GSO**; (d) **Det**; (e) **MI**; (f) **FS**; (g) **EIM**.

Much work still remains to solve this problem in terms of theoretical problem setting, efficient algorithms, and practical engineering issues. On one hand, for sound field control, one could, for instance, replace the 2D FEM software employed here with specialized 3D acoustics simulation software that can model complex geometries, different boundary conditions (e.g., materials on the walls), and the directivity of sources. However, even for candidate positions restricted within a 2D plane, this would raise very challenging issues of computational complexity. On the other hand, it would be interesting to transfer some of these methods to settings other than acoustics, for instance, to temperature fields in chip design, to jointly optimize the positions of heat sources/sinks and temperature sensors.

ACKNOWLEDGMENT

This work was supported by JST, PRESTO Grant Number JPMJPR18J4, Japan, and LABEX WIFI (Laboratory of Excellence ANR-10-LABX-24) within the French Program “Investments for the Future” under reference ANR-10-IDEX-0001-02 PSL*.

REFERENCES

- [1] P. A. Nelson and S. J. Elliott, *Active Control of Sound*. London, UK: Academic Press, 1992.
- [2] M. Miyoshi and Y. Kaneda, “Inverse filtering of room acoustics,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 2, pp. 145–152, 1988.
- [3] P. A. Nelson, “Active control of acoustic fields and the reproduction of sound,” *J. Sound Vibr.*, vol. 177, no. 4, pp. 447–477, 1993.
- [4] O. Kirkeby and P. A. Nelson, “Reproduction of plane wave sound fields,” *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2992–3000, 1993.
- [5] O. Kirkeby, P. A. Nelson, F. O. Bustamante, and H. Hamada, “Local sound field reproduction using digital signal processing,” *J. Acoust. Soc. Am.*, vol. 100, no. 3, pp. 1584–1593, 1996.
- [6] P.-A. Gauthier and A. Berry, “Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain,” *J. Acoust. Soc. Am.*, vol. 117, no. 2, pp. 662–678, 2005.
- [7] T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, “Personal sound zones: Delivering interface-free audio to multiple listeners,” *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 81–91, 2015.
- [8] F. M. Fazi and P. A. Nelson, “Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control,” *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 1–14, 2012.
- [9] A. J. Berkhout, D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.
- [10] M. A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, 2005.
- [11] 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, 2013.
- [12] T. Betlehem and T. D. Abhayapala, “Theory and design of sound field reproduction in reverberant environment,” *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2100–2111, 2005.

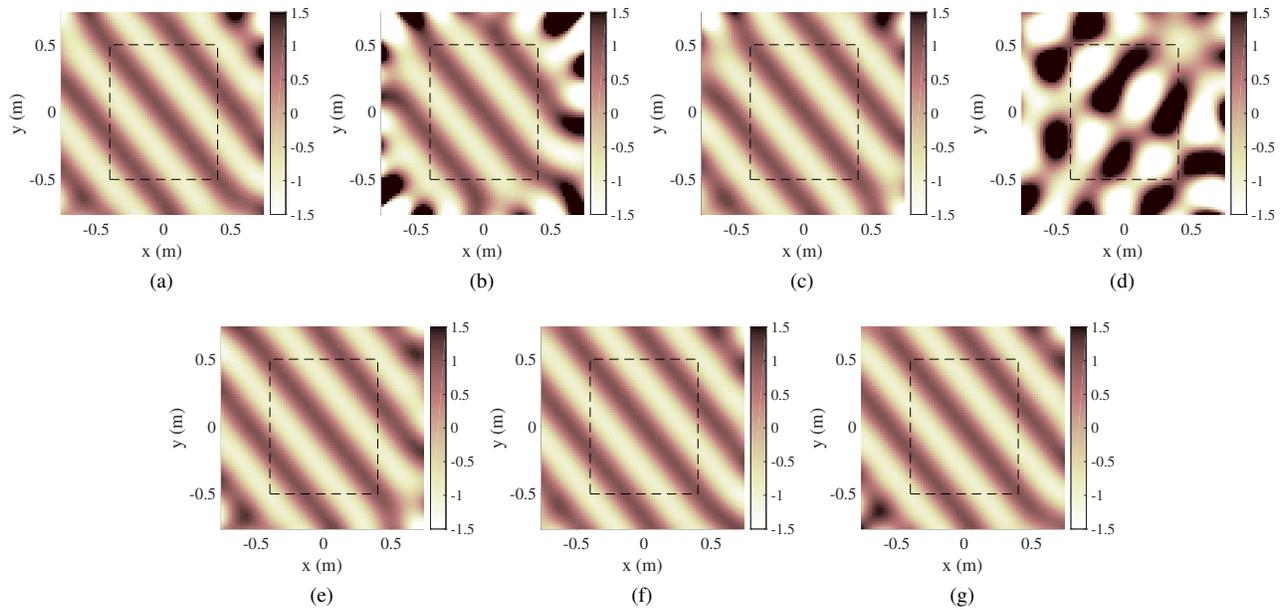


Fig. 20. Synthesized pressure fields of plane wave from 219 deg at 800 Hz in broadband case. (a) **Reg+EIMI**; (b) **Rand**; (c) **GSO**; (d) **Det**; (e) **MI**; (f) **FS**; (g) **EIM**.

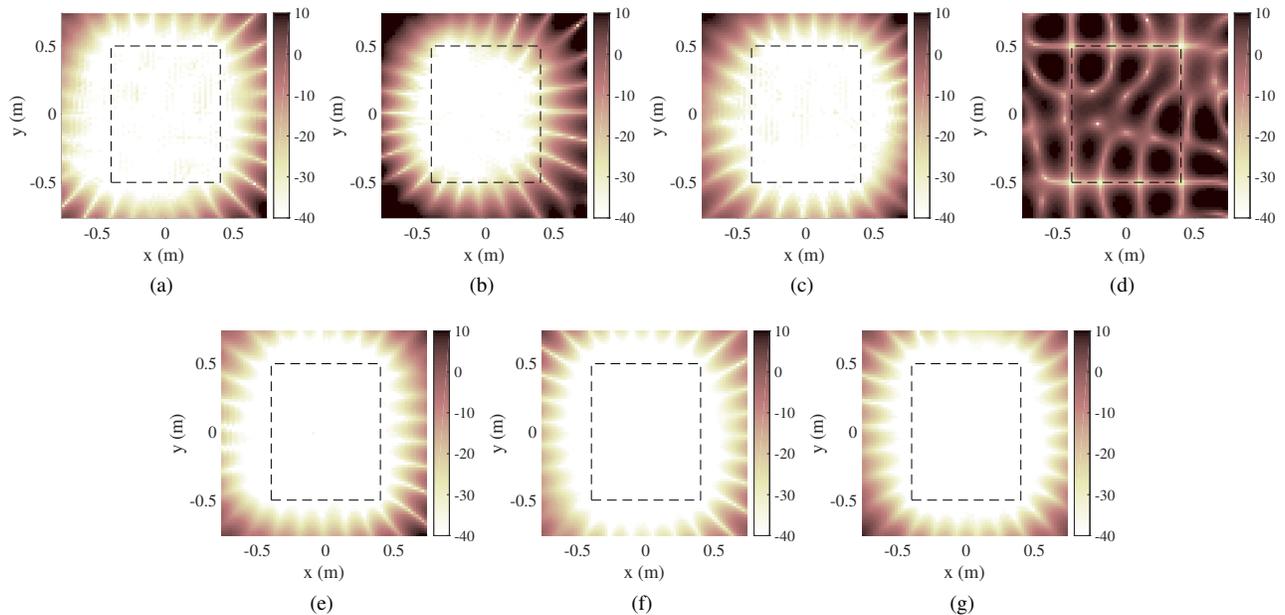


Fig. 21. Normalized error distributions at 800 Hz in broadband case. (a) **Reg+EIMI**; (b) **Rand**; (c) **GSO**; (d) **Det**; (e) **MI**; (f) **FS**; (g) **EIM**. SDRs are 41.2, 28.3, 41.7, -6.1 , 49.3, 54.4, and 52.3 dB, respectively.

- [13] M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array," *J. Acoust. Soc. Am.*, vol. 118, no. 5, pp. 3094–3103, 2005.
- [14] S. Koyama, K. Furuya, Y. Hiwasaki, Y. Haneda, and Y. Suzuki, "Wave field reconstruction filtering in cylindrical harmonic domain for with-height recording and reproduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 10, pp. 1546–1557, 2014.
- [15] 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. Am.*, vol. 139, no. 3, pp. 1024–1036, 2016.
- [16] S. D. Snyder and H. Hansen, "Using multiple regression to optimize active noise control system design," *J. Sound Vibr.*, vol. 148, no. 3, pp. 537–542, 1991.
- [17] R. L. Clark and C. R. Fuller, "Optimal placement of piezoelectric actuators and polyvinylidene fluoride error sensors in active structural acoustic control approaches," *J. Acoust. Soc. Am.*, vol. 92, no. 3, pp. 1521–1533, 1992.
- [18] T. C. Yang, C. H. Tseng, and S. F. Ling, "Constrained optimization of active noise control system design," *J. Acoust. Soc. Am.*, vol. 95, no. 6, pp. 3390–3399, 1994.
- [19] C. E. Ruckman and C. F. Fuller, "Optimizing actuator locations in active noise control systems using subset selection," *J. Sound Vibr.*, vol. 186, no. 3, pp. 395–406, 1995.
- [20] K. H. Baek and S. J. Elliott, "Natural algorithms for choosing source locations in active control system," *J. Sound Vibr.*, vol. 186, no. 2, pp. 245–267, 1995.
- [21] F. Asano, Y. Suzuki, and D. C. Swanson, "Optimization of control source configuration in active control systems using Gram-Schmidt orthogonalization," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 2, pp. 213–230, 1999.
- [22] H. Khalilian, I. V. Bajić, and R. G. Vaughan, "Comparison of loud-speaker placement methods for sound field reproduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 8, pp. 1364–1379, 2016.

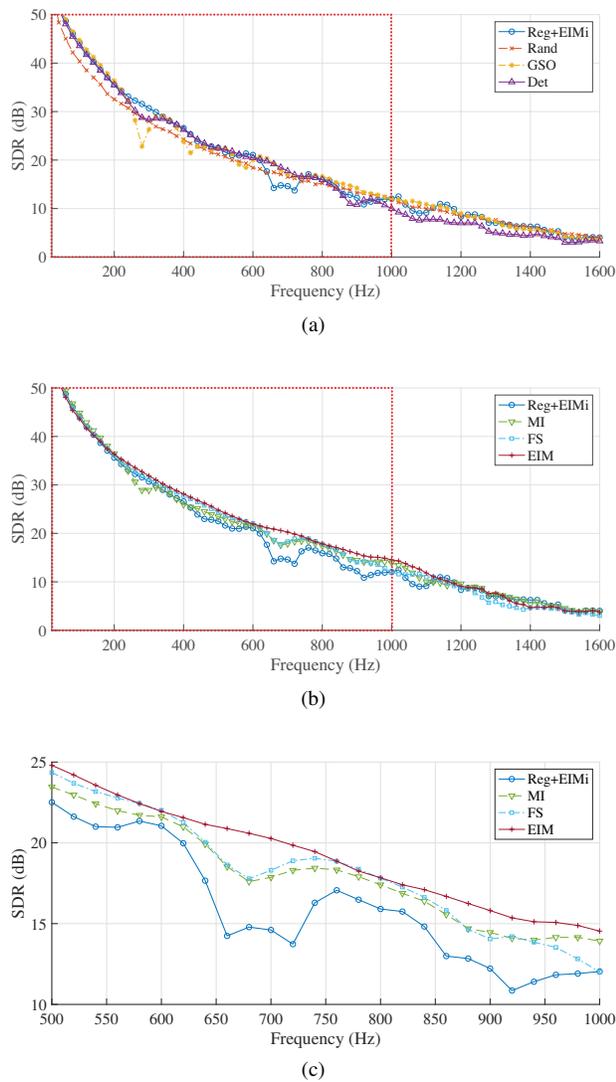


Fig. 22. SDR with respect to frequency in broadband case when position errors are added to secondary sources and control points using Gaussian distribution with a standard deviation of 0.01 m. The frequency range used in the source and/or sensor placement is indicated by red dashed line. Results of **Reg+EIMi**, **MI**, **FS**, and **EIM** for 500–1000 Hz are enlarged in (c).

2016.

[23] B. Rafaely, “Analysis and design of spherical microphone arrays,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 13, no. 1, pp. 135–143, 2005.

[24] G. Chardon, W. Kreuzer, and M. Noisternig, “Design of spatial microphone arrays for sound field interpolation,” *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 5, pp. 780–790, 2015.

[25] N. K. Nguyen and A. J. Miller, “A review of some exchange algorithms for constructing discrete D-optimal designs,” *Comput. Stat. Data Anal.*, vol. 14, no. 4, pp. 489–498, 1992.

[26] L. Yao, W. Sethares, and D. Kammer, “Sensor placement for on-orbit modal identification via a genetic algorithm,” *Am. Inst. Aeronaut. Astronaut. J.*, vol. 31, no. 10, pp. 1922–1928, 1993.

[27] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, 2009.

[28] M. Shamaiah, S. Benerjee, and H. Vikalo, “Greedy sensor selection: leveraging submodularity,” in *IEEE Conf. Decision Control (CDC)*, Atlanta, GA, USA, Dec. 2010, pp. 2572–2577.

[29] D. J. C. MacKay, “Information-based objective functions for active data selection,” *Neural Comput.*, vol. 4, no. 4, pp. 590–604, 1992.

[30] H. Wang, K. Yao, G. Pottie, and D. Estrin, “Entropy-based sensor selection heuristic for target localization,” in *Proc. Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Berkeley, CA, USA, Apr. 2004, pp. 36–45.

[31] A. Krause, A. Singh, and C. Guestrin, “Near-optimal sensor placement

in gaussian process: Theory, efficient algorithms and empirical studies,” *J. Mach. Learn. Res.*, vol. 9, pp. 235–284, 2008.

[32] J. J. Benedetto and M. Fickus, “Finite normalized tight frames,” *Adv. Comput. Math.*, vol. 18, no. 2–4, pp. 357–385, 2003.

[33] J. Ranieri, A. Cheoira, and M. Vetterli, “Near-optimal sensor placement for linear inverse problems,” *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1135–1146, 2014.

[34] J. Ranieri and M. Vetterli, “Near-optimal source placement for linear physical fields,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 6825–6829.

[35] S. Koyama, G. Chardon, and L. Daudet, “Joint source and sensor placement for sound field control based on empirical interpolation method,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Calgary, Apr. 2018, pp. 501–505.

[36] —, “Codes for reproducing the experimental results,” <https://github.com/sh01k/SourceSensorPlacementSFC>.

[37] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. New York, NY, USA: Springer, 2013.

[38] F. Fazi, P. Nelson, and R. Potthast, “Analogies and differences between three methods for sound field reproduction,” in *Proc. Ambisonics Symp.*, Graz, Austria, June 2009.

[39] S. Spors, R. Rabenstein, and J. Ahrens, “The theory of wave field synthesis revisited,” in *Proc. 124th AES Conv.*, Amsterdam, Netherlands, Oct. 2008.

[40] F. Winter, H. Wierstorf, C. Hold, F. Krüger, A. Raake, and S. Spors, “Colouration in local wave field synthesis,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 10, pp. 1913–1924, 2019.

[41] J. Daniel, S. Moureau, and R. Nicol, “Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging,” in *Proc. 114th AES Conv.*, Amsterdam, Netherlands, Mar. 2003.

[42] J. Daniel, “Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonics format,” in *Proc. 23rd AES Int. Conf.*, Copenhagen, Denmark, May 2003.

[43] G. H. Koopmann, L. Song, and J. B. Fahline, “A method for computing acoustic fields based on the principle of wave superposition,” *J. Acoust. Soc. Am.*, vol. 86, no. 5, pp. 2433–2438, 1989.

[44] M. E. Johnson, S. J. Elliott, K. H. Baek, and J. Garcia-Bontio, “An equivalent source technique for calculating the sound field inside an enclosure containing scattering objects,” *J. Acoust. Soc. Am.*, vol. 104, no. 3, pp. 1221–1231, 1998.

[45] S. Marburg, “Six boundary elements per wavelength: is that enough?” *J. Computat. Acoust.*, vol. 10, no. 01, pp. 25–51, 2002.

[46] J. Ahrens and S. Spors, “An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions,” *Acta Acustica united with Acustica*, vol. 94, pp. 988–999, 2008.

[47] 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, 2009.

[48] J. Ahrens and S. Spors, “An analytical approach to 3D sound field reproduction employing spherical distribution of non-omnidirectional loudspeakers,” *Limassol*, Mar. 2010.

[49] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, “Sound field reproduction method in spatio-temporal frequency domain considering directivity of loudspeakers,” in *Proc. 132nd AES Conv.*, Budapest, Apr. 2012.

[50] J. R. Driscoll and D. M. Healy, “Computing fourier transforms and convolutions on the 2-sphere,” *Adv. Appl. Math.*, vol. 15, no. 2, pp. 202–250, 1994.

[51] E. B. Saff and A. B. J. Kuijlaars, “Distributing many points on a sphere,” *Math. Intell.*, vol. 19, no. 1, pp. 5–11, 1997.

[52] R. H. Hardin and N. J. A. Sloane, “Mclaren’s improved snub cube and other new spherical designs in three dimensions,” *Discrete Comput. Geom.*, vol. 15, no. 4, pp. 429–441, 1996.

[53] B. Rafaely, *Fundamentals of Spherical Array Processing*. Berlin, Germany: Springer, 2015.

[54] S. Ise, “A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems,” *Acta Acustica united with Acustica*, vol. 85, no. 1, pp. 78–87, 1999.

[55] J. Meyer and G. Elko, “A highly scalable spherical microphone array based on an orthogonal decomposition of the soundfield,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2002, pp. II–1781–1784.

[56] T. D. Abhayapala and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone array,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2002, pp. II–1949–1952.

- [57] J. Ahrens and S. Spors, "Sound field reproduction using planar and linear arrays of loudspeakers," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 8, pp. 2038–2050, 2010.
- [58] A. Gupta and T. D. Abhayapala, "Three-dimensional sound field reproduction using multiple circular arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 5, 2011.
- [59] N. Ueno, S. Koyama, and H. Saruwatari, "Three-dimensional sound field reproduction based on weighted mode-matching method," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 12, pp. 1852–1867, 2019.
- [60] A. Laborie, R. Bruno, and S. Montoya, "A new comprehensive approach of surround sound recording," in *Proc. 114th AES Conv.*, Amsterdam, Netherlands, 2003.
- [61] N. Ueno, S. Koyama, and H. Saruwatari, "Sound field recording using distributed microphones based on harmonic analysis of infinite order," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 135–139, 2018.
- [62] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [63] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer, 2010.
- [64] G. N. Lilis, D. Angelosante, and G. B. Giannakis, "Sound field reproduction using Lasso," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 8, pp. 1902–1921, 2010.
- [65] H. Khalilian, I. V. Bajić, and R. G. Vaughan, "Loudspeaker placement for sound field reproduction by constrained matching pursuit," in *IEEE Int. Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, NY, USA, Oct. 2013, pp. 1–4.
- [66] N. Radmanesh and I. S. Burnett, "Generation of isolated wideband sound fields using a combined two-stage Lasso-LS algorithm," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 2, pp. 378–387d, 2013.
- [67] H. Khalilian, I. V. Bajić, and R. G. Vaughan, "Towards optimal loudspeaker placement for sound field reproduction," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Vancouver, Canada, May 2013, pp. 321–325.
- [68] S. Takane, Y. Suzuki, and T. Sone, "A new method for global sound field reproduction based on Kirchhoff's integral equation," *Acta Acustica united with Acustica*, vol. 85, no. 2, pp. 250–257, 1999.
- [69] A. Das and D. Kempe, "Algorithms for subset selection in linear regression," in *Proc. ACM Symp. Theory of Comput. (STOC)*, Victoria, Canada, May 2008, pp. 45–54.
- [70] —, "Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection," in *Int. Conf. Mach. Learn. (ICML)*, Bellevue, WA, USA, June 2011.
- [71] D. M. Steinberg and W. G. Hunter, "Experimental design: Review and comment," *Technometrics*, vol. 26, no. 2, pp. 71–97, 1984.
- [72] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, 2015.
- [73] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [74] J. Kiefer and J. Wolfowitz, "The equivalence of two extremum problems," *Can. J. Math.*, vol. 12, pp. 363–366, 1960.
- [75] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, "An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations," *C. R. Acad. Sci. Paris, Ser. I*, vol. 339, pp. 667–672, 2004.
- [76] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera, "Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations," *ESAIM: M2AN*, vol. 41, no. 3, pp. 575–605, 2007.
- [77] Y. Maday, N. C. Nguyen, A. T. Patera, and S. H. Pau, "A general multi-purpose interpolation procedure: the magic points," *Communications on Pure & Applied Analysis*, vol. 8, no. 1, pp. 383–404, 2009.
- [78] F. Hecht, "New development in FreeFem++," *J. Numer. Math.*, vol. 20, no. 3-4, pp. 251–265, 2012.
- [79] G. Chardon, A. Cohen, and L. Daudet, "Sampling and reconstruction of solutions to the Helmholtz equation," *Sampling Theory Signal Image Process.*, vol. 13, no. 1, pp. 67–89, 2014.
- [80] H. A. Schenck, "Improved integral formulation for acoustic radiation problems," *J. Acoust. Soc. Am.*, vol. 44, no. 1, pp. 41–58, 1968.
- [81] T. W. Wu and A. F. Seybert, "A weighted residual formulation for the CHIEF method in acoustics," *J. Acoust. Soc. Am.*, vol. 90, no. 3, pp. 1608–1614, 1991.
- [82] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695–1700, 2002.

- [83] 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, 2010.



Shoichi Koyama (M'10) received the B.E., M.S., and Ph.D. degrees from the University of Tokyo, Tokyo, Japan, in 2007, 2009, and 2014, respectively. He joined Nippon Telegraph and Telephone Corporation in 2009 as a Researcher in acoustic signal processing. He moved to the University of Tokyo in 2014 and is currently an Assistant Professor (Lecturer) since 2018. From 2016 to 2018, he was also a Visiting Researcher with Paris Diderot University (Paris7), Institut Langevin, Paris, France. His research interests include acoustic inverse problems,

sound field analysis and synthesis, and spatial audio.

He is a member of the Acoustical Society of America, the Audio Engineering Society, the Institute of Electronics, Information and Communication Engineers, and the Acoustical Society of Japan (ASJ). He was the recipient of Itakura Prize Innovative Young Researcher Award by ASJ in 2015, and the Research Award by Funai Foundation for Information Technology in 2018.



Gilles Chardon received the engineering degrees of Ecole Polytechnique and Telecom ParisTech in 2009, as well as the MSc ATIAM of Université Pierre et Marie Curie, Paris VI. After working towards his PhD at Institut Langevin in Paris, and a postdoctoral position with the Mathematics and Signal Processing group of the Acoustics Research Institute of the Austrian Academy of Sciences in Vienna, he is now Associate Professor with CentraleSupélec, Gif-sur-Yvette, France. His main research interests include sparse representation of acoustical fields, inverse

problems and numerical analysis in acoustics.



Laurent Daudet (M'04–SM'10) graduated in physics from Ecole Normale Supérieure (Paris, France) and holds a PhD in Applied Mathematics from Marseille University. He is currently on leave from his position of Professor of Physics at Paris Diderot University, Paris. Prior to that or in parallel, he has held various academic positions: fellow of the Institut Universitaire de France, associate professor at Université Pierre et Marie Curie, post-doctoral fellow and visiting senior lecturer at Queen Mary University of London, UK, visiting professor at the

National Institute for Informatics in Tokyo, Japan. Laurent has authored or co-authored nearly 200 scientific publications, mostly on signal processing applied to wave physics. He has been a consultant to various small and large companies, and is a co-inventor in several patents. He is now co-founder and CTO at LightOn, a startup developing optical co-processors for machine learning.