Spatial Active Noise Control Based on Kernel Interpolation of Sound Field

Shoichi Koyama, *Member, IEEE*, Jesper Brunnström, Hayato Ito, Natsuki Ueno, *Student Member, IEEE*, and Hiroshi Saruwatari, *Member, IEEE*

***Abstract*—An active noise control (ANC) method to reduce noise over a region in space based on kernel interpolation of sound field is proposed. Current methods of spatial ANC are largely based on spherical or circular harmonic expansion of the sound field, where the geometry of the error microphone array is restricted to a simple one such as a sphere or circle. We instead apply the kernel interpolation method, which allows for the estimation of a sound field in a continuous region with flexible array configurations. The interpolation scheme is used to derive adaptive filtering algorithms for minimizing the acoustic potential energy inside a target region. A practical time-domain algorithm is also developed together with its computationally efficient block-based equivalent. We conduct experiments to investigate the achievable level of noise reduction in a two-dimensional free space, as well as adaptive broadband noise control in a three-dimensional reverberant space. The experimental results indicated that the proposed method outperforms the multipoint-pressure-control-based method in terms of regional noise reduction.**

***Index Terms*—Adaptive filter, kernel interpolation, spatial active noise control, sound field control.**

I. Introduction

THE aim of active noise control (ANC) or active noise cancellation is to suppress unwanted noise by generating antinoise with secondary sources, i.e., loudspeakers. ANC techniques have been extensively investigated [1]–[4] along with their practical applicability, especially for canceling low-frequency noise [5]–[8].

Achieving effective ANC over a three-dimensional (3D) space generally requires multiple microphones and loudspeakers to capture and synthesize a 3D sound field. Adaptive filters are typically used to minimize the cost function, which in conventional multipoint pressure control is defined as the power of the residual noise signals captured by error microphones [3], [9]. Thus, these methods primarily reduce noise local to error microphone positions. Virtual sensing techniques attempt to estimate the noise at locations remote from the physical error microphones [10]–[12]; however, the cost function is still constructed as the power of the residual noise at discrete positions including virtual ones. Therefore, the noise field between the (physical/virtual) error microphone positions is not taken into consideration. Furthermore, most of the virtual sensing techniques require an identification stage to estimate a filter for predicting error signals at the virtual microphones.

In recent years, spatial ANC, which aims to control noise over a continuous 3D target region, has attracted attention [13]–[19] owing to advancements in sound field recording and reproduction methods [20]–[23]. Most spatial ANC methods are based on the expansion of the captured sound field into wave-domain basis functions, such as spherical and circular harmonics [24]. Consequently, the methods can only be applied to a simple array geometry, such as sphere, circle, or its combination [15], [18], [19]. Since the secondary sources and error microphones have to be placed surrounding the target region to suppress the interior noise field, it is preferable that their placement is flexible. Moreover, adaptive filtering algorithms for spatial ANC are derived in the frequency domain with a few exceptions [25], [26]. Implementing time-domain algorithms, especially for broadband noise, from these frequency-domain algorithms is not straightforward. Therefore, it is necessary to develop practical time-domain adaptive filtering algorithms for spatial ANC.

We propose a spatial ANC approach based on the kernel interpolation of a sound field. The cost function is defined as the acoustic potential energy inside the target region. To minimize this cost function with an adaptive filter, it is necessary to estimate the noise field from discrete microphone measurements. We apply kernel interpolation [27], which makes it possible to estimate a continuous sound field in the frequency domain under the constraint of the Helmholtz equation. This interpolation method is equivalent to spherical/circular harmonic analysis of infinite order [28], [29] when using pressure microphones. The sound field can also be estimated from arbitrarily placed microphones. Therefore, the proposed method can meet the demand for flexible microphone and loudspeaker arrangements. Two time-domain algorithms based on kernel interpolation are developed. An extension of the well-known filtered-x least mean squares (FxLMS) algorithm [2], [3] for minimizing the regional noise power is derived, which is called kernel-interpolation-based FxLMS (KI-FxLMS). We also develop the fast block KI-FxLMS, which is a more computationally efficient algorithm using blockwise operations and the fast Fourier transform (FFT).

Manuscript received May 1, 2021; revised August 2, 2021; accepted August 17, 2021. Date of publication August 26, 2021; date of current version October 2, 2021. This work was supported by JST PRESTO Grant JPMJPR18J4. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefan Bilbao. *(Corresponding author: Shoichi Koyama.)*

Shoichi Koyama, Hayato Ito, Natsuki Ueno, and Hiroshi Saruwatari are with the Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan (e-mail: koyama.shoichi@ieee.org; h1y1t52t5@gmail.com; natsuki.ueno@ieee.org; hiroshi_saruwatari@ipc.i.u-tokyo.ac.jp).

Jesper Brunnström is with the Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8654, Japan, and also with the Department of Electrical Engineering (ESAT), Stadius Center for Dynamical Systems, Signal Processing and Data Analytics, Katholieke Universiteit Leuven, 3001 Leuven, Belgium (e-mail: jesper.brunnstroem@kuleuven.be).

Digital Object Identifier 10.1109/TASLP.2021.3107983

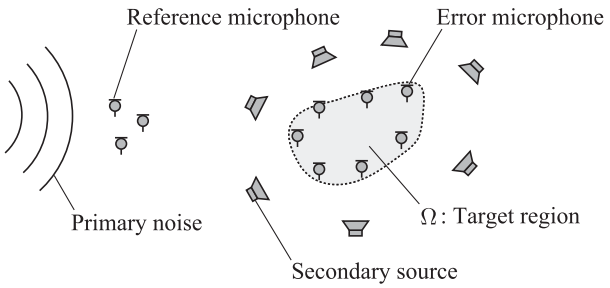


Fig. 1. Conceptual diagram of feedforward spatial active noise control.

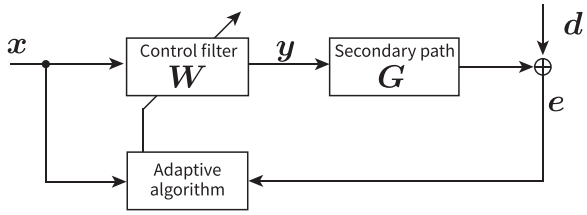


Fig. 2. Block diagram of a multichannel feedforward ANC system.

Although preliminary results of this study are presented in [17], [30], we here describe in detail the derivations of the algorithms with a newly obtained optimal frequency-domain filter. In addition, extensive experimental results including 3D spatial ANC comparing with the virtual sensing technique and that using real data obtained in a practical environment are also presented.

The rest of this paper is organized as follows. In Section II, the problem statement of spatial ANC is described. The kernel interpolation method is introduced in Section III. In Section IV, the proposed spatial ANC algorithms are developed. Experimental results compared with those of the multipoint-pressure-control-based ANC method and the virtual sensing technique are reported in Section V. Finally, Section VI concludes this paper.

A. Notation

Italic letters denote scalars, lower case boldface italic letters denote vectors, and upper case boldface italic 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. The unit sphere in $\mathbb{R}^3$ is denoted by $\mathbb{S}_2$. Subscripts of scalars, vectors, and tensors indicate their indexes. To illustrate, $x_{i,j}$ is the (i,j)th entry of the matrix $\boldsymbol{X}$. The identity matrix of size $N \times N$ is denoted by $\boldsymbol{I}_N$.

The imaginary unit is denoted by $\mathrm{j} := \sqrt{-1}$. The complex conjugate, transpose, conjugate transpose, and inverse are denoted by superscripts $(\cdot)^*$, $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $(\cdot)^{-1}$, respectively. The absolute value of a scalar x is denoted by $|x|$. The ℓ_p-norm of a vector $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|_p$.

Indexes of discrete time and filter coefficients are denoted by n and i, respectively, and angular frequency is denoted by ω, in arguments of variables or functions. For example, $\boldsymbol{x}(n)$ is the vector-valued signal at the time index n. The sound velocity and wave number are denoted by c and $k = \omega/c$, respectively. Harmonic time dependence $e^{\mathrm{j}\omega t}$ with the time t is assumed according to conventions in the signal processing literature.

II. Problem Statement

Suppose that a target region Ω is set in the 3D space $\mathbb{R}^3$ and that undesired noise arrives from outside of Ω. The objective of spatial ANC is to reduce incoming noise over the region Ω by controlling the output of the placed loudspeakers. The incoming noise and the loudspeakers are called the primary noise and secondary sources, respectively. As shown in Fig. 1, error microphones are placed inside Ω to measure the error of the adaptation process. Although the error microphones can be arbitrarily placed inside Ω, it is usually desirable that the error microphones are placed near the boundary of Ω to save space for ANC users. It is also known that the microphones should be placed near the boundary to estimate an interior sound field, but arranging the pressure microphones to be completely on the boundary can lead to the forbidden frequency problem [31]. The secondary sources are placed around Ω and are used to generate an antinoise field of the primary noise from the signals captured by the reference microphones.

The signals of the reference microphones and secondary sources in the discrete time domain are denoted by $\boldsymbol{x}(n) \in \mathbb{R}^R$ and $\boldsymbol{y}(n) \in \mathbb{R}^L$ with the time index n, where the numbers of reference microphones and secondary sources are R and L, respectively. The error microphone signals and primary noise at the error microphone positions are denoted by $\boldsymbol{e}(n)$ and $\boldsymbol{d}(n)$ $(\in \mathbb{R}^M)$, respectively. The number of error microphones is M. The total pressure field at the position $\boldsymbol{r} \in \Omega$ and time n is denoted by $u(\boldsymbol{r}, n)$. The pressure field at the position of the mth microphone $\boldsymbol{r}_m$ is equal to the mth element of $\boldsymbol{e}(n)$, meaning that $e_m(n) = u(\boldsymbol{r}_m, n)$. The signals in the frequency domain are distinguished by the angular frequency ω as an argument, for example, as $\boldsymbol{x}(\omega)$ $(\in \mathbb{C}^R)$.

The impulse response between the secondary sources and the error microphones (secondary paths) is represented by a multiple-input multiple-output (MIMO) finite impulse response (FIR) filter $\boldsymbol{G}(i) \in \mathbb{R}^{M \times L}$ of length J, where each i indicates a matrix of filter coefficients. The error microphone signal $\boldsymbol{e}(n)$ can be expressed as a superposition of the primary noise and the secondary source signals as

$$\boldsymbol{e}(n) = \boldsymbol{d}(n) + \sum_{i=0}^{J-1} \boldsymbol{G}(i)\boldsymbol{y}(n-i). \tag{1}$$

We assume that the true secondary paths are known, for instance, by measuring them offline. The control filter is a MIMO FIR filter $\boldsymbol{W}(i) \in \mathbb{R}^{L \times R}$ of length I. The driving signals of the secondary sources, $\boldsymbol{y}(n)$, are obtained by filtering the reference signal $\boldsymbol{x}(n)$ with the control filter $\boldsymbol{W}(i)$ as

$$\boldsymbol{y}(n) = \sum_{i=0}^{I-1} \boldsymbol{W}(i)\boldsymbol{x}(n-i). \tag{2}$$

It is assumed that the effect of the secondary source signals is negligible at the reference microphone positions.

The above-described ANC setup is a typical multichannel feedforward ANC system, whose block diagram is depicted in Fig. 2. Although we focus only on feedforward ANC in this study, feedback ANC is also useful to suppress periodic noises [3]. One of the benefits of feedback ANC is its lack of reference microphones, leading to simple system configurations.

Feedback ANC algorithms can be developed from the following derivations by constructing virtual reference signals as

$$\hat{\boldsymbol{x}}(n) = \boldsymbol{e}(n) - \sum_{j=0}^{J-1} \boldsymbol{G}(j)\boldsymbol{y}(n-j). \tag{3}$$

Our objective is to reduce the power of $u(\boldsymbol{r}, n)$ over the region Ω by adapting the control filter $\boldsymbol{W}(i)$. Since the error microphones are placed inside Ω in discrete locations, the continuous sound field $u(\boldsymbol{r}, n)$ has to be predicted from $\boldsymbol{e}(n)$.

III. Kernel Interpolation of Sound Field

We here introduce the kernel interpolation of a sound field [27], with the goal of estimating the sound field inside the source-free target region Ω from microphone observations. The kernel interpolation method is formulated in the frequency domain; hence, the signals are dependent on ω, which is omitted for notational simplicity in this section.

The sound pressure $u(\boldsymbol{r})$ at angular frequency ω is observed by M pressure microphones, i.e., the error microphones. The observed value of the mth microphone in the frequency domain is $e_m \in \mathbb{C}$. The interpolation problem is formulated as

$$\underset{u \in \mathcal{H}}{\text{minimize}} \sum_{m=1}^{M} |u(\boldsymbol{r}_m) - e_m|^2 + \lambda \|u\|_{\mathcal{H}}^2. \tag{4}$$

Here, $\mathcal{H}$ is a function space for which we seek a solution, which, therefore, should be designed for representing a sound field, $\|\cdot\|_{\mathcal{H}}$ is a norm on $\mathcal{H}$, and λ is a real non-negative parameter. The first term of (4) is a loss term between u and $e_1, \ldots, e_M$, and the second term is a regularization term. First, we introduce kernel ridge regression for solving (4) in Section III-A. Next, we define $\mathcal{H}$ for sound field interpolation in Section III-B.

A. Kernel Ridge Regression

When $\mathcal{H}$ is a function space called a *reproducing kernel Hilbert space*, the optimization problem (4) corresponds to kernel ridge regression, for which a closed-form solution can be obtained [32]. First, $\mathcal{H}$ is assumed to be a reproducing kernel Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and positive-definite reproducing kernel $\kappa : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$. On the basis of the representer theorem [33], the solution of (4) can be expressed as

$$u(\boldsymbol{r}) = \sum_{m=1}^{M} \alpha_m \kappa(\boldsymbol{r}, \boldsymbol{r}_m), \tag{5}$$

where $\alpha_m \in \mathbb{C}$ for $m = 1, \ldots, M$. By substituting (5) into the objective function of (4), we obtain

$$\boldsymbol{\alpha}^{\mathsf{H}}(\boldsymbol{K}^{\mathsf{H}}\boldsymbol{K} + \lambda \boldsymbol{K})\boldsymbol{\alpha} - \boldsymbol{\alpha}^{\mathsf{H}}\boldsymbol{K}^{\mathsf{H}}\boldsymbol{e} - \boldsymbol{e}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{\alpha} + \boldsymbol{e}^{\mathsf{H}}\boldsymbol{e}, \tag{6}$$

where

$$\boldsymbol{\alpha} := \begin{bmatrix} \alpha_1, & \ldots, & \alpha_M \end{bmatrix}^{\mathsf{T}}, \tag{7}$$

$$\boldsymbol{K} := \begin{bmatrix} \kappa(\boldsymbol{r}_1, \boldsymbol{r}_1) & \ldots & \kappa(\boldsymbol{r}_1, \boldsymbol{r}_M) \\ \vdots & \ddots & \vdots \\ \kappa(\boldsymbol{r}_M, \boldsymbol{r}_1) & \ldots & \kappa(\boldsymbol{r}_M, \boldsymbol{r}_M) \end{bmatrix}. \tag{8}$$

Since $\boldsymbol{K}$ is Hermitian and positive-definite, $\boldsymbol{\alpha}$ is obtained by minimizing (6) as

$$\boldsymbol{\alpha} = (\boldsymbol{K} + \lambda \boldsymbol{I}_M)^{-1} \boldsymbol{e}. \tag{9}$$

Therefore, $u(\boldsymbol{r})$ is estimated as

$$u(\boldsymbol{r}) = \boldsymbol{\kappa}(\boldsymbol{r})^{\mathsf{T}} (\boldsymbol{K} + \lambda \boldsymbol{I}_M)^{-1} \boldsymbol{e}, \tag{10}$$

where

$$\boldsymbol{\kappa}(\boldsymbol{r}) = \begin{bmatrix} \kappa(\boldsymbol{r}, \boldsymbol{r}_1), & \ldots, & \kappa(\boldsymbol{r}, \boldsymbol{r}_M) \end{bmatrix}^{\mathsf{T}}. \tag{11}$$

Thus, the closed-form solution of (4) is (10). Next, it is necessary to define appropriate $\mathcal{H}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, as well as κ.

B. Reproducing Kernel for Sound Field Interpolation

The pressure field u inside the source-free simply connected region Ω can be modeled as a solution of the following (homogeneous) Helmholtz equation [24]:

$$(\nabla^2 + \grave{k}^2) u = 0, \tag{12}$$

where ∇^2 is the Laplacian. Any solution of (12) can be well approximated by the superposition of plane waves, i.e., the Herglotz wave function [29], [34], as

$$u(\boldsymbol{r}) = \frac{1}{4\pi} \int_{\mathbb{S}_2} \tilde{u}(\boldsymbol{\xi}) e^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}\boldsymbol{r}} \mathrm{d}\boldsymbol{\xi}, \tag{13}$$

where $\tilde{u}$ is the (square-integrable) complex amplitude of the plane wave of arrival direction $\boldsymbol{\xi} \in \mathbb{S}_2$, and $\boldsymbol{k} := -\grave{k}\boldsymbol{\xi}$ is the wave vector. Using this representation, we define the inner product and norm over the Hilbert space $\mathcal{H}$ as

$$\langle u_1, u_2 \rangle_{\mathcal{H}} = \frac{1}{4\pi} \int_{\mathbb{S}_2} \frac{1}{\gamma(\boldsymbol{\xi})} \tilde{u}_1(\boldsymbol{\xi})^* \tilde{u}_2(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}, \tag{14}$$

$$\|u\|_{\mathcal{H}} = \sqrt{\langle u, u \rangle_{\mathcal{H}}}, \tag{15}$$

respectively. Here, $\gamma(\boldsymbol{\xi})$ is a directional weighting function, which is introduced to incorporate prior knowledge on source directions.

We set the kernel function $\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2)$ as

$$\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{4\pi} \int_{\mathbb{S}_2} \gamma(\boldsymbol{\xi}) e^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}(\boldsymbol{r}_1 - \boldsymbol{r}_2)} \mathrm{d}\boldsymbol{\xi}. \tag{16}$$

When we denote $\kappa(\boldsymbol{r}, \boldsymbol{r}')$ by $\kappa_{\boldsymbol{r}'}(\boldsymbol{r})$, $\langle \kappa_{\boldsymbol{r}'}, u \rangle$ is obtained as

$$\begin{aligned} \langle \kappa_{\boldsymbol{r}'}, u \rangle_{\mathcal{H}} &= \frac{1}{4\pi} \int_{\mathbb{S}_2} \gamma(\boldsymbol{\xi}) \frac{1}{\gamma(\boldsymbol{\xi})} e^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}\boldsymbol{r}'} \tilde{u}(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \\ &= \frac{1}{4\pi} \int_{\mathbb{S}_2} \tilde{u}(\boldsymbol{\xi}) e^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}\boldsymbol{r}'} \mathrm{d}\boldsymbol{\xi} \\ &= u(\boldsymbol{r}'). \end{aligned} \tag{17}$$

Therefore, we can confirm that $\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the reproducing kernel of $\mathcal{H}$. In this study, we set the directional weighting function to be uniform, i.e., $\gamma(\boldsymbol{\xi}) = 1$. Thus, the specific kernel function $\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is obtained from (49) as

$$\begin{aligned} \kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \frac{1}{4\pi} \int_{\mathbb{S}_2} e^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}(\boldsymbol{r}_1 - \boldsymbol{r}_2)} \mathrm{d}\boldsymbol{\xi} \\ &= j_0(\grave{k}\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|_2), \end{aligned} \tag{18}$$

where $j_0(\cdot)$ is the 0th-order spherical Bessel function of the first kind. See Appendix for the kernel function when using a directional weighting. The Gram matrix $\boldsymbol{K}$ when using this kernel function (18) corresponds to the normalized spatial covariance matrix in a diffuse sound field [35]. When a two-dimensional (2D) sound field is assumed instead, the corresponding kernel function becomes

$$\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) = J_0(\grave{k}\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|_2), \tag{19}$$

where $J_0(\cdot)$ is the 0th-order Bessel function of the first kind.

The kernel interpolation of a sound field from the microphone measurements $\boldsymbol{e}$ is achieved by using (10) with the kernel function (18) or (19). Depending on the setting of the number and placement of microphones, the inversion of the Gram matrix $\boldsymbol{K}$ sometimes becomes numerically unstable. One simple strategy to prevent this problem is to appropriately set the regularization parameter λ. Another well-known strategy is low-rank approximation of $\boldsymbol{K}$ [36], [37].

IV. Spatial ANC Based on Kernel Interpolation

The objective of spatial ANC is to reduce the power of the sound field inside the target region Ω. Therefore, we define the cost function for adapting the control filter as the acoustic potential energy inside Ω,

$$\mathcal{L} := \mathbb{E}\left[\int_{\Omega} u(\boldsymbol{r},n)^2 \mathrm{d}\boldsymbol{r}\right], \tag{20}$$

where $\mathbb{E}[\cdot]$ represents the expectation with respect to the time n. By assuming a stationary sound field, we can describe the frequency-domain cost function at ω as

$$\bar{\mathcal{L}}(\omega) := \int_{\Omega} |u(\boldsymbol{r},\omega)|^2 \mathrm{d}\boldsymbol{r}. \tag{21}$$

In both cases, the pressure distribution must be obtained. We apply kernel interpolation of the sound field to estimate $u(\boldsymbol{r},n)$ or $u(\boldsymbol{r},\omega)$ from the error microphone signals $\boldsymbol{e}(n)$ or $\boldsymbol{e}(\omega)$. First, we formulate the optimal filter in the frequency domain. Then, two adaptive filtering algorithms in the time domain are derived. We also discuss the difference of these algorithms from non-spatial ANC algorithms, i.e., multipoint pressure control.

A. Optimal Filter in Frequency Domain

The relationship between the principal signals, given in the time domain in (1) and (2), can be similarly obtained in the frequency domain as

$$\boldsymbol{e}(\omega) = \boldsymbol{d}(\omega) + \boldsymbol{G}(\omega)\boldsymbol{y}(\omega), \tag{22}$$

$$\boldsymbol{y}(\omega) = \boldsymbol{W}(\omega)\boldsymbol{x}(\omega). \tag{23}$$

The pressure distribution $u(\boldsymbol{r},\omega)$ can be estimated from the error signals $\boldsymbol{e}(\omega)$ as

$$u(\boldsymbol{r},\omega) = \boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r},\omega)\boldsymbol{e}(\omega) \tag{24}$$

$$\boldsymbol{z}(\boldsymbol{r},\omega) = \left[\left(\boldsymbol{K}(\omega) + \lambda \boldsymbol{I}_M\right)^{-1}\right]^{\mathsf{T}} \boldsymbol{\kappa}(\boldsymbol{r},\omega). \tag{25}$$

Here, $\boldsymbol{z}(\boldsymbol{r},\omega)$ is the interpolation filter in the frequency domain, which is directly obtained from the kernel ridge regression formulation in (10). The kernel function is given in (18) for the 3D case and (19) for the 2D case. By substituting (24) into (21), we can rewrite the frequency-domain cost function $\bar{\mathcal{L}}(\omega)$ as

$$\bar{\mathcal{L}}(\omega) = \boldsymbol{e}^{\mathsf{H}}(\omega)\boldsymbol{A}(\omega)\boldsymbol{e}(\omega), \tag{26}$$

where

$$\begin{aligned}\boldsymbol{A}(\omega) &:= \int_{\Omega} \boldsymbol{z}^*(\boldsymbol{r},\omega)\boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r},\omega)\mathrm{d}\boldsymbol{r} \\ &= \boldsymbol{P}^{\mathsf{H}}(\omega)\left[\int_{\Omega} \boldsymbol{\kappa}^*(\boldsymbol{r},\omega)\boldsymbol{\kappa}^{\mathsf{T}}(\boldsymbol{r},\omega)\mathrm{d}\boldsymbol{r}\right]\boldsymbol{P}(\omega).\end{aligned} \tag{27}$$

Here, $\boldsymbol{P}(\omega) := (\boldsymbol{K}(\omega) + \lambda \boldsymbol{I}_M)^{-1}$ is defined. Thus, the cost function $\bar{\mathcal{L}}(\omega)$ is the power of the error signals subject to the weighting matrix $\boldsymbol{A}(\omega)$. The elements of $\boldsymbol{A}(\omega)$ can be determined from the positions of the error microphones alone. Numerical integration must be performed to calculate the integral in (27) in most cases; however, it can be avoided for some simple shapes of Ω [17], [38].

The gradient of $\bar{\mathcal{L}}(\omega)$ with respect to $\boldsymbol{W}^*(\omega)$ is derived as

$$\begin{aligned}\frac{\partial \bar{\mathcal{L}}}{\partial \boldsymbol{W}^*} &= \frac{\partial}{\partial \boldsymbol{W}^*}\boldsymbol{e}^{\mathsf{H}}\boldsymbol{A}\boldsymbol{e} \\ &= \boldsymbol{G}^{\mathsf{H}}\boldsymbol{A}\left(\boldsymbol{G}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{d}\right)\boldsymbol{x}^{\mathsf{H}},\end{aligned} \tag{28}$$

where ω is omitted. The optimal control filter $\boldsymbol{W}(\omega)$ satisfies $\partial\bar{\mathcal{L}}(\omega)/\partial\boldsymbol{W}^*(\omega) = 0$ as

$$\boldsymbol{W}(\omega) = -(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{A}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathsf{H}}\boldsymbol{A}\boldsymbol{R}_{dx}\boldsymbol{R}_{xx}^{-1}, \tag{29}$$

where $\boldsymbol{R}_{dx} = \boldsymbol{d}\boldsymbol{x}^{\mathsf{H}}$ and $\boldsymbol{R}_{xx} = \boldsymbol{x}\boldsymbol{x}^{\mathsf{H}}$. This control filter $\boldsymbol{W}(\omega)$ is noncausal and includes the unobservable value $\boldsymbol{d}$, but can still be useful to investigate fundamental properties. Although steepest descent algorithms for minimizing $\bar{\mathcal{L}}(\omega)$ can be derived from the gradient $\partial\bar{\mathcal{L}}/\partial\boldsymbol{W}^* = \boldsymbol{G}^{\mathsf{H}}\boldsymbol{A}\boldsymbol{e}\boldsymbol{x}^{\mathsf{H}}$, the direct use of such frequency-domain adaptive filters suffers from a significant delay.

B. Kernel-Interpolation-Based FxLMS Algorithm

In this section, we develop a practical time-domain algorithm based on the frequency domain formulation given in Section IV-A. The algorithm is referred to as kernel-interpolation-based FxLMS (KI-FxLMS). The time-domain kernel-interpolation filter is obtained by inverse discrete-time Fourier transform of $\boldsymbol{z}(\boldsymbol{r},\omega)$ in (25) as

$$\boldsymbol{z}(\boldsymbol{r},i) = \mathcal{F}^{-1}\left[\boldsymbol{z}(\boldsymbol{r},\omega)\right]. \tag{30}$$

Then, the pressure distribution in the time domain can be estimated as

$$u(\boldsymbol{r},n) = \sum_{i=-\infty}^{\infty} \boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r},i)\boldsymbol{e}(n-i). \tag{31}$$

The time-domain cost function (20) can thus be reformulated as

$$\mathcal{L} = \mathbb{E}\left[\sum_{i,j=-\infty}^{\infty} \boldsymbol{e}^{\mathsf{T}}(n-i)\boldsymbol{\Gamma}(i,j)\boldsymbol{e}(n-j)\right], \tag{32}$$

where $\boldsymbol{\Gamma}(i,j)$ is the coefficient matrix of the interpolation filter defined as

$$\boldsymbol{\Gamma}(i,j) := \int_{\Omega} \boldsymbol{z}(\boldsymbol{r},i)\boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r},j)\mathrm{d}\boldsymbol{r}. \tag{33}$$

The gradient of $\mathcal{L}$ with respect to the filter coefficients $\boldsymbol{W}(i)$, denoted by $\boldsymbol{\Delta}(i)$, is obtained as

$$\begin{aligned}\boldsymbol{\Delta}(i) &:= \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}(i)} \\ &= 2\sum_{\nu,\eta=-\infty}^{\infty}\sum_{j=0}^{J-1} \boldsymbol{G}^{\mathsf{T}}(j)\boldsymbol{\Gamma}^{\mathsf{T}}(\nu,\eta) \\ &\qquad \cdot \mathbb{E}\left[\boldsymbol{e}(n-\nu)\boldsymbol{x}^{\mathsf{T}}(n-i-j-\eta)\right],\end{aligned} \tag{34}$$

where the symmetric property of $\boldsymbol{\Gamma}(i,j)$ derived from its definition (33), specifically $\boldsymbol{\Gamma}(i,j) = \boldsymbol{\Gamma}^{\mathsf{T}}(j,i)$, is used. Note that $\boldsymbol{G}$ and $\boldsymbol{\Gamma}$ are here considered deterministic.

To obtain a practical algorithm, we further simplify the gradient. We apply a change in variables $\eta = \nu + k$ to (34). If the signals $\boldsymbol{x}(n)$ and $\boldsymbol{e}(n)$ are assumed to be wide-sense stationary in a local time window, the cross correlation $\mathbb{E}[\boldsymbol{e}(n)\boldsymbol{x}^{\mathsf{T}}(n-i)]$ is dependent only on the time difference i. Then, the gradient $\boldsymbol{\Delta}(i)$ can be represented as

$$\boldsymbol{\Delta}(i) = 2\sum_{k=-\infty}^{\infty}\sum_{j=0}^{J-1}\boldsymbol{G}^{\mathsf{T}}(j)\boldsymbol{A}^{\mathsf{T}}(k) \cdot \mathbb{E}\left[\boldsymbol{e}(n)\boldsymbol{x}^{\mathsf{T}}(n-i-j-k)\right], \quad (35)$$

where $\boldsymbol{A}(k)$ is the simplified weighting filter matrix, which can be obtained from $\boldsymbol{A}(\omega)$ in (27) as

$$\begin{aligned}\boldsymbol{A}(k) &= \sum_{\nu=-\infty}^{\infty}\boldsymbol{\Gamma}(\nu,\nu+k)\\ &= \int_{\Omega}\sum_{\nu=-\infty}^{\infty}\boldsymbol{z}(\boldsymbol{r},\nu)\boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r},\nu+k)\mathrm{d}\boldsymbol{r}\\ &= \mathcal{F}^{-1}\left[\boldsymbol{A}(\omega)\right](k). \end{aligned} \quad (36)$$

The relationship between the correlation operation in the time and frequency domains is used in the last line.

In practice, the weighting filter matrix $\boldsymbol{A}(k)$ has to be implemented as a causal FIR filter. We approximate $\boldsymbol{A}(k)$ by truncating it to a FIR filter of length $2K+1$ and adding a delay of *K* samples to make it causal. Then, the causal weighting filter matrix is defined as

$$\hat{\boldsymbol{A}}(i) := \begin{cases}\boldsymbol{A}(i-K), & 0 \le i \le 2K\\ \boldsymbol{0}, & \text{otherwise}\end{cases} \quad (37)$$

The same delay K must be added to the error signal to obtain the correct cross-correlation. Therefore, the gradient $\boldsymbol{\Delta}(i)$ becomes

$$\boldsymbol{\Delta}(i) \approx 2\sum_{k=0}^{2K}\sum_{j=0}^{J-1}\boldsymbol{G}^{\mathsf{T}}(j)\hat{\boldsymbol{A}}^{\mathsf{T}}(k) \cdot \mathbb{E}\left[\boldsymbol{e}(n-K)\boldsymbol{x}^{\mathsf{T}}(n-i-j-k)\right]. \quad (38)$$

When the secondary path is modeled offline, $\boldsymbol{G}$ and $\hat{\boldsymbol{A}}$ can be combined into a single filter in advance, reducing computational cost. We define the combination of these two FIR filters as

$$\boldsymbol{H}(i) := \sum_{j=0}^{2K}\hat{\boldsymbol{A}}(j)\boldsymbol{G}(i-j). \quad (39)$$

Note that $\hat{\boldsymbol{A}}$, and by extension $\boldsymbol{H}$, can be computed before the ANC process is started. Then, replacing the expected value of the gradient with the instantaneous value, the update rule of the KI-FxLMS is obtained as

$$\boldsymbol{W}_{n+1}(i) = \boldsymbol{W}_n(i) - \mu\sum_{j=0}^{J+2K-1}\boldsymbol{H}^{\mathsf{T}}(j)\boldsymbol{e}(n-K)\boldsymbol{x}^{\mathsf{T}}(n-i-j), \quad (40)$$

where the subscript of $\boldsymbol{W}(i)$ indicates the time index, μ is the step-size parameter, and the constant factor 2 in (40) is absorbed into μ. The step-size parameter is normalized at each time step by the power of the filtered reference signal, some variation of which is used in many descriptions of the multichannel FxLMS algorithm [39]–[41], as

$$\mu(n) = \frac{\mu_0}{\sum_i \left\|\sum_j \boldsymbol{G}(j)\otimes\boldsymbol{x}(n-i-j)\right\|_F^2 + \beta}, \quad (41)$$

where μ_0 is a constant parameter, $\otimes$ denotes the Kronecker product, and β is a parameter to avoid division by zero.

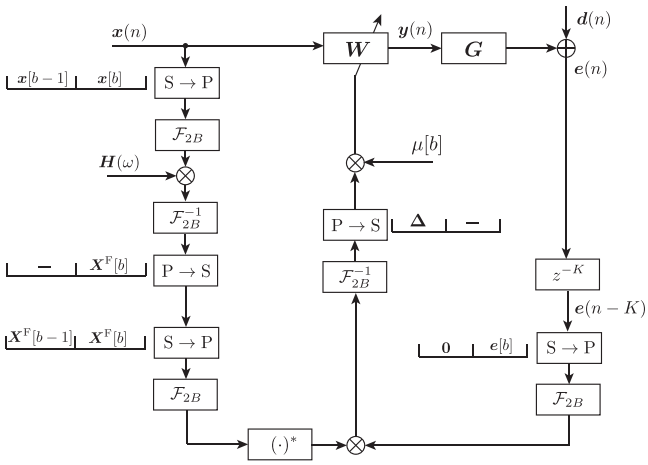


Fig. 3. Block diagram of fast block KI-FxLMS. The serial-to-parallel (S → P) block indicates whether to zero-pad the signal or concatenate the two last signal blocks. The parallel-to-serial (P → S) block indicates whether the first or second block should be kept.

C. Fast Block Kernel-Interpolation-Based FxLMS Algorithm

The computational cost of KI-FxLMS can be high when a large number of channels and long filters are used. By keeping the control filter constant during a finite block, we can employ a blockwise operation, which allows us to implement the algorithm with the FFT to reduce computational cost. Such a technique has been used for the standard FxLMS, which is called fast block FxLMS [42]. We here describe a blockwise implementation of KI-FxLMS, i.e., fast block KI-FxLMS.

The linear convolution and correlation operations in the filter update are performed with the overlap-save algorithm [43] using the FFT. To avoid causality problems that can be introduced by block-based processing, the control filter is still obtained in the time domain. The loudspeaker signals are then easily computed sample by sample by computing the convolution between $\boldsymbol{x}(n)$ and $\boldsymbol{W}(i)$ in the time domain.

The block diagram of the fast block KI-FxLMS is shown in Fig. 3. The block size is denoted by B, and the index of the block is denoted by b. Each block of signal is represented using square brackets; for example, $\boldsymbol{x}[b]$ indicates the bth reference signal block consisting of $\boldsymbol{x}(n)$ for $n \in [bB, (b+1)B-1]$. The length of the control filter I is set to be equal to the block length B. The FFT operation for sequences of two block lengths $2B$ is denoted by $\mathcal{F}_{2B}$. After the inverse FFT $\mathcal{F}_{2B}^{-1}$, one block of its output is discarded. The serial-to-parallel (S → P) block indicates whether to zero-pad the signal or concatenate the two latest signal blocks. Each parallel-to-serial (P → S) block indicates whether the first or second block should be kept. The reference signal $\boldsymbol{x}$ filtered by $\boldsymbol{H}$ is denoted by $\boldsymbol{X}^{\mathrm{F}} \in \mathbb{R}^{M\times L\times R}$. The frequency-domain filter $\boldsymbol{H}(\omega)$ is here obtained by the FFT

of $\boldsymbol{H}(i)$ after applying zero padding up to the length of $2B$ at the end of the filter. It is also possible to separately filter with $\boldsymbol{G}$ and $\hat{\boldsymbol{A}}$ in series, which is necessary when estimating the secondary path online [44]. To calculate the linear correlation, the filtered reference signal $\boldsymbol{X}^{\mathrm{F}}$ is multiplied by $\boldsymbol{e}$ in the frequency domain, and summed over the dimension of length M. Eventually, the gradient $\boldsymbol{\Delta}(i) \in \mathbb{R}^{L\times R}$ is obtained.

The fast block KI-FxLMS is computationally efficient in many cases owing to the FFT algorithm. The KI-FxLMS is often more computationally costly, but performs the filter update sample by sample. Their computational costs are evaluated in the next section.

D. Comparison With Multipoint Pressure Control

In ANC methods based on multipoint pressure control, adaptive filtering algorithms are derived with the aim of reducing the primary noise at the error microphone positions. Therefore, the cost function is typically formulated as the power of the error microphone signal as

$$\mathcal{J} = \mathbb{E}\left[\|\boldsymbol{e}(n)\|_2^2\right]. \tag{42}$$

In the frequency domain, the cost function for multipoint pressure control at ω is given by

$$\bar{\mathcal{J}}(\omega) = \|\boldsymbol{e}(\omega)\|_2^2. \tag{43}$$

By comparing (26) and (43), one can find that the difference between the cost functions in the frequency domain is only in the weighting matrix $\boldsymbol{A}(\omega)$ applied to the error signals. Therefore, the optimal filter for multipoint pressure control in the frequency domain can be obtained by replacing $\boldsymbol{A}(\omega)$ in (29) with the identity matrix $\boldsymbol{I}_M$, giving

$$\boldsymbol{W}(\omega) = -(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathsf{H}}\boldsymbol{R}_{dx}\boldsymbol{R}_{xx}^{-1}. \tag{44}$$

In other words, the proposed method achieves spatial interpolation solely through the multiplication with the weighting matrix $\boldsymbol{A}$ in (29).

Similar insights can also be applied to the time-domain algorithm. As in (32), the filter coefficient matrix $\boldsymbol{\Gamma}$, which is later simplified to the weighting filter matrix $\boldsymbol{A}(i)$ in (36), is the difference between the proposed cost function and (42). Thus, by skipping the convolution of $\hat{\boldsymbol{A}}(i)$ in KI-FxLMS, the FxLMS algorithm for multipoint pressure control can be obtained as

$$\boldsymbol{W}_{n+1}(i) = \boldsymbol{W}_n(i) - \mu\sum_{j=0}^{J-1}\boldsymbol{G}^{\mathsf{T}}(j)\boldsymbol{e}(n)\boldsymbol{x}^{\mathsf{T}}(n-i-j). \tag{45}$$

The fast block version of the FxLMS algorithm can also be derived for multipoint pressure control.

The difference in computational cost between KI-FxLMS and FxLMS stems from the weighting filter $\boldsymbol{A}$. The number of real-valued multiplications required for every sample of the input signals, not counting the update of the step-size parameter, is compared among FxLMS, fast block FxLMS, KI-FxLMS, and fast block KI-FxLMS in Table I. Since the filter $\boldsymbol{H}(i)$ of length $2K+J$ is applied in KI-FxLMS instead of $\boldsymbol{G}(i)$ of length J, the number of real-valued multiplication increases by $2RLMK$. The fast block implementation often has reduced computational cost both for FxLMS and KI-FxLMS owing to the FFT. We assume that an N-length FFT requires $N/2\log_2 N$ complex-valued multiplications, and each complex-valued multiplication corresponds to four real-valued multiplications. The difference in computational cost between the regular and fast block implementations primarily depends on the filter lengths. For instance, when the lengths of $\boldsymbol{W}(i)$ and $\boldsymbol{H}(i)$ are equal, $M = 48$, $R = 1$, and $L = 32$, the fast block KI-FxLMS is faster than KI-FxLMS for $I > 32$. On the other hand, a disadvantage of the fast block implementations is the processing delay caused by the accumulation of samples necessary before processing. Note that the numbers of real-valued multiplications of fast block FxLMS and KI-FxLMS algorithms are identical.

TABLE I
COMPUTATIONAL COMPLEXITIES OF ALGORITHMS

Algorithm	Number of real-valued multiplications
FxLMS	$RL(2I + M(I + J))$
Fast block FxLMS	$RL(16M + I + 1)$ $+4(R + M + RL(2M + 1))\log_2 2I$
KI-FxLMS	$RL(2I + M(I + J + 2K))$
Fast block KI-FxLMS	$RL(16M + I + 1)$ $+4(R + M + RL(2M + 1))\log_2 2I$

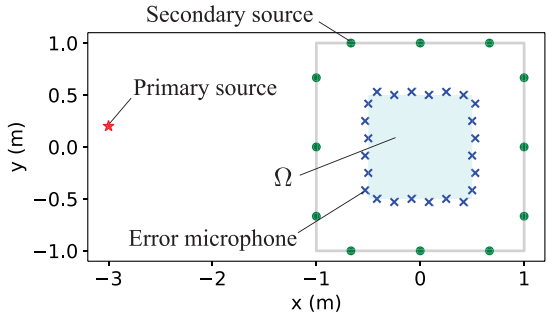


Fig. 4. Experimental setup for 2D numerical simulation.

V. EXPERIMENTS

Experiments were conducted to evaluate the proposed spatial ANC algorithms in comparison with the multipoint-pressure-control-based ANC. First, their fundamental performances were compared by numerical simulation in the frequency domain assuming a 2D sound field. Then, the time-domain algorithms were evaluated in a 3D reverberant field, including a comparison with a virtual sensing technique. Finally, experimental results using real data obtained by a practical array system are shown.

A. 2D Numerical Simulations in Frequency Domain

In a 2D free field, the target region Ω was chosen as a square of dimensions $1.0\,\mathrm{m} \times 1.0\,\mathrm{m}$ with its center at the origin. The numbers of error microphones and secondary sources were $M = 24$ and $L = 12$, respectively. As shown in Fig. 4, the error microphones were regularly placed along the boundary of Ω, but half of them were shifted 0.03 m outward to alleviate the effect of the forbidden frequency problem [31]. The secondary sources were regularly arranged along the square of dimensions $2.0\,\mathrm{m} \times 2.0\,\mathrm{m}$. The reference signal was directly obtained from

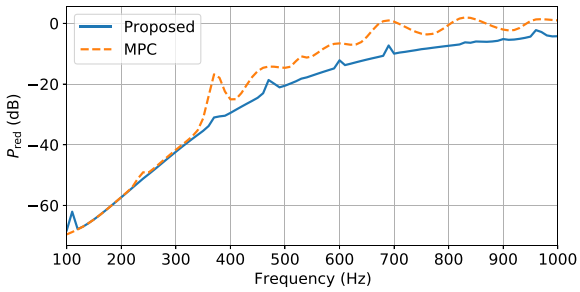


Fig. 5. Regional power reduction P_{red} of the proposed method (Proposed) and multipoint pressure control (MPC) with respect to frequency. The primary noise source was a single point source.

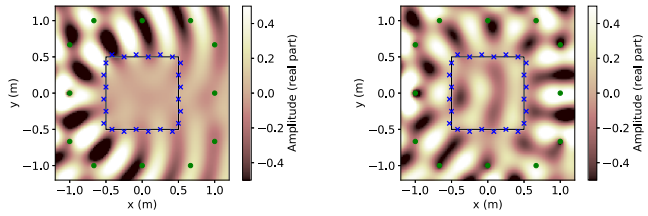


Fig. 6. Pressure distribution at 700 Hz. (a) Proposed; (b) MPC. The blue crosses and green dots indicate error microphones and secondary sources, respectively. The boundary of the target region Ω is indicated by the square.

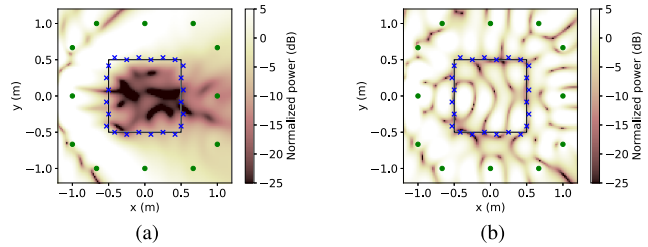


Fig. 7. Normalized power distribution at 700 Hz. (a) Proposed; (b) MPC. P_{red} was -9.97 dB for Proposed and 0.55 dB for MPC.

the primary noise source. We here investigate two cases for the primary noise source positions.

Assuming a stationary sound field, we compared the frequency-domain optimal filters of the proposed method (29) and multipoint pressure control (MPC) (44). The parameter λ in (25) was set to 10^{-4}.

As a performance measure for spatial ANC, we define the regional power reduction P_{red} as

$$P_{\mathrm{red}}(\omega) = 10\log_{10}\frac{\sum_j |u(\boldsymbol{r}_j,\omega)|^2}{\sum_j |u_{\mathrm{p}}(\boldsymbol{r}_j,\omega)|^2}, \tag{46}$$

where $u_{\mathrm{p}}(\cdot)$ denotes the pressure field of the primary noise source without ANC. The evaluation points $\boldsymbol{r}_j \in \Omega$ were obtained by discretizing the target region Ω every 0.02 m. The number of evaluation points was 50×50.

First, results for the case when the primary noise source was a single point source at $(-3.0, 0.2)$ m are shown. The amplitude of the primary source signal was 10.0. Fig. 5 shows the relationship between the frequency and the regional power reduction P_{red}. The investigated frequencies were from 100 to 1000 Hz at intervals of 10 Hz. At low frequencies, the noise power was reduced considerably by both methods, and the performance difference

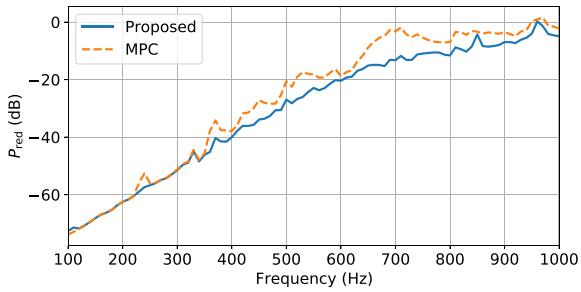


Fig. 8. Regional power reduction P_{red} of Proposed and MPC with respect to frequency. The primary noise source was 16 point sources surrounding the target region.

was small. The effect of the spatial interpolation in the proposed method can be seen at higher frequencies. Above 350 Hz, the P_{red} of MPC became larger than that of Proposed. The regional power reduction of the two methods gradually increased as the frequency increased, because the accuracy of capturing and synthesizing a sound field with a limited number of loudspeakers and microphones deteriorates at higher frequencies.

Fig. 6 shows the pressure distribution at 700 Hz. The power distribution normalized by the average power of the primary noise field inside Ω, i.e., $|u(\boldsymbol{r},\omega)|^2 / \int_\Omega |u_{\mathrm{p}}(\boldsymbol{r},\omega)|^2 \mathrm{d}\boldsymbol{r}$, is shown in Fig. 7. The acoustic power inside Ω was reduced by the proposed method, but remained for MPC. The regional power reduction P_{red} was -9.97 dB for Proposed and 0.55 dB for MPC.

Next, the case where multiple primary noise sources exist is investigated. Sixteen point sources were equiangularly placed along the circle of radius 3.0 m with its center at the origin. The amplitudes of the primary source signals were determined by the complex Gaussian distribution of mean 10.0 and variance 4.0. These settings were intended to simulate a diffuse field, which is the assumption used for deriving the kernel function (18), for the primary noise field. The relationship between the frequency and P_{red} is shown in Fig. 8. Again, the P_{red} of Proposed was lower than that of MPC above 350 Hz. Compared with the result of the single primary source case, the overall regional noise reduction improved, especially below 800 Hz.

B. Investigation of Interpolation Filter

As described in Section IV-B, the interpolation filter $\boldsymbol{A}(i)$ must be truncated to be usable in real-time systems. Therefore, the truncation parameter K defined in (37) has to be chosen appropriately for the KI-FxLMS algorithm. A short filter is desirable because it can be cheaper to compute and the control filter update has to be delayed for K samples in the KI-FxLMS algorithm. Conversely, the cost of a low K is the truncation error incurred by setting the filter coefficients $\boldsymbol{A}(i)$ to zero for $|i| > K$. To investigate the parameter, we define the truncation error of the filter $\boldsymbol{A}(i)$ as

$$E_{\boldsymbol{A}}(K) = 10\log_{10}\frac{\sum_{i=K+1}^{\infty} \|\boldsymbol{A}(i)\|_F^2 + \|\boldsymbol{A}(-i)\|_F^2}{\sum_{i=-\infty}^{\infty} \|\boldsymbol{A}(i)\|_F^2}. \tag{47}$$

The expression is equivalent to the normalized mean square error of the filtered signal, assuming that the signal before being

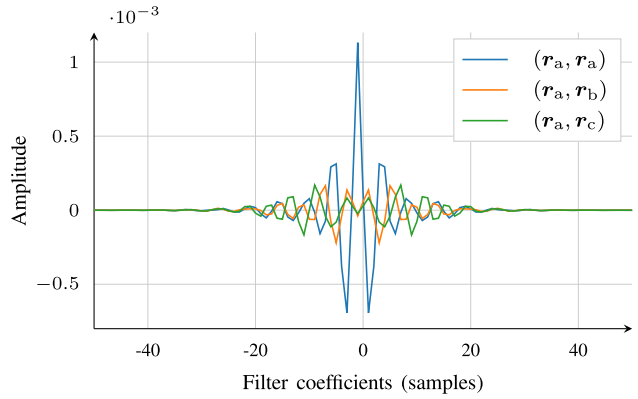


Fig. 9. Impulse response of the kernel interpolation filter $\boldsymbol{A}(i)$ for channels corresponding to different microphone pairs. The microphone positions are $\boldsymbol{r}_\mathrm{a} = (-0.275, -0.3, 0.05)\,\mathrm{m}$, $\boldsymbol{r}_\mathrm{b} = (0.3, -0.275, 0.05)\,\mathrm{m}$, and $\boldsymbol{r}_\mathrm{c} = (0.275, 0.3, 0.05)\,\mathrm{m}$.

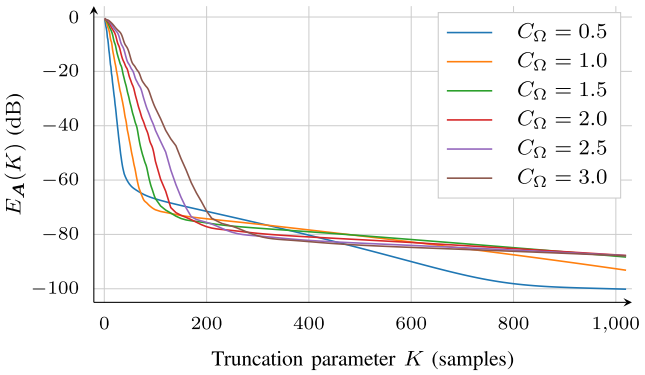


Fig. 10. Truncation error of the interpolation filter $\boldsymbol{A}(i)$ for a cuboid target region of size $C_\Omega\,\mathrm{m} \times C_\Omega\,\mathrm{m} \times 0.1\,\mathrm{m}$.

filtered is white with independent channels. This allows the error to be expressed solely in terms of the filter coefficients.

We investigate the effect of truncating the interpolation filter $\boldsymbol{A}$ for one choice of error microphone array and target region. The target region is defined as a cuboid of dimensions $C_\Omega\,\mathrm{m} \times C_\Omega\,\mathrm{m} \times 0.1\,\mathrm{m}$ with its center at the origin. The error microphones, 48 in total, are regularly placed along the squares of dimensions $C_\Omega\,\mathrm{m} \times C_\Omega\,\mathrm{m}$ on the two planes $z = 0.05\,\mathrm{m}$ and $-0.05\,\mathrm{m}$, but half of them shifted 0.03 m outward. The impulse response for a few channels of the interpolation filter, calculated for $C_\Omega = 0.6$, is shown in Fig. 9. The impulse responses correspond to the weighting between three microphone pairs at the positions $(\boldsymbol{r}_\mathrm{a}, \boldsymbol{r}_\mathrm{a})$, $(\boldsymbol{r}_\mathrm{a}, \boldsymbol{r}_\mathrm{b})$, and $(\boldsymbol{r}_\mathrm{a}, \boldsymbol{r}_\mathrm{c})$. The positions of the chosen microphones are $\boldsymbol{r}_\mathrm{a} = (-0.275, -0.3, 0.05)\,\mathrm{m}$, $\boldsymbol{r}_\mathrm{b} = (0.3, -0.275, 0.05)\,\mathrm{m}$, and $\boldsymbol{r}_\mathrm{c} = (0.275, 0.3, 0.05)\,\mathrm{m}$.

The interpolation filter was computed with a sampling rate of 4000 Hz, the regularization parameter in (25) was $\lambda = 10^{-3}$, and the integral in (27) was computed using a naive Monte Carlo integration procedure with 1000 samples [45]. The true filter in (47) is a noncausal IIR filter and cannot be computed exactly. As a substitute, a long FIR filter of length 16383 samples was generated.

The truncation error as a function of K for various values of C_Ω is shown in Fig. 10. The error decreases more slowly for larger values of C_Ω, suggesting a correlation between the size of the target region and the desired value of K. However, for all sizes of the target region, the truncation error decreases rapidly below a certain value of K, after which the decrease slows down dramatically. Attempting to select a value of K just past that point is recommended, as it provides most of the value for a small cost.

C. Time-Domain Simulations Comparing With Remote Microphone Technique

The proposed time-domain algorithm is evaluated in a 3D simulated reverberant environment. One current ANC method applicable to spatial ANC with arbitrarily placed microphones and loudspeakers is virtual sensing, by densely placing virtual microphones inside the target region. We here compare the proposed method with the remote microphone technique (RMT) [10], [46], which is a representative virtual sensing method, in addition to the MPC-based method. The RMT requires an identification stage to estimate a filter for predicting signals at the virtual microphone positions from the primary noise at a fixed position. At this stage, it is necessary to place microphones at both the physical and virtual positions, but they are removed from the virtual positions at the running stage. Note that the identification stage is not necessary for the proposed and MPC-based methods.

For the RMT with M_v virtual microphones, the virtual secondary paths $\boldsymbol{G}_\mathrm{v}(i) \in \mathbb{R}^{M_\mathrm{v} \times L}$, the acoustic paths from the secondary sources to the virtual microphone positions, must be measured. We here assume that they are perfectly estimated in advance. At the identification stage, a FIR observation filter for predicting the primary noise at the virtual microphone positions from the primary noise at the physical microphone positions is estimated. During the running stage, the virtual error signal is obtained by estimating $\boldsymbol{d}(n)$ from (1), applying the observation filter, and adding the secondary signals at the virtual microphone positions estimated from $\boldsymbol{G}_\mathrm{v}(i)$ and $\boldsymbol{y}(n)$. Then, the control filter is updated to reduce the physical and virtual error signals with the FxLMS algorithm.

We focus on the performance comparison between the proposed kernel-interpolation-based method and RMT; therefore, the fast block algorithms were used for all three time-domain algorithms. The reverberant environment was simulated using the image source method [47], [48]. The room dimension was $7.0\,\mathrm{m} \times 5.0\,\mathrm{m} \times 2.5\,\mathrm{m}$, and the reflection coefficients were set so that the reverberation time T_{60} became 0.35 s. For the RMT, the observation filter was estimated using the physical and virtual error signals of 40000 samples, i.e., 10 s, with the LMS algorithm until the normalized mean square error for the virtual error signal estimation became less than -40 dB. The primary noise source was placed at $(-2.8, 0.3, 0.0)\,\mathrm{m}$. The primary noise source emitted white Gaussian noise filtered by a bandpass filter with a passband of 50 to 900 Hz, which is hereafter referred to as **Noise**.

As a performance measure, we define the regional power reduction in the time domain as

$$P_\mathrm{red}(n) = 10 \log_{10} \frac{\sum_j \sum_\nu u(\boldsymbol{r}_j, n-\nu)^2}{\sum_j \sum_\nu u_\mathrm{p}(\boldsymbol{r}_j, n-\nu)^2}, \qquad (48)$$

where the signal power is calculated over a window in time, $\nu \in [0, T-1]$. The positions $\boldsymbol{r}_j$ are the evaluation points.

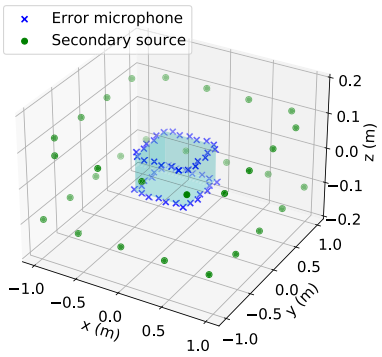


Fig. 11. Positions of error microphones and secondary sources for experiments using real data. The target region Ω is illustrated by the blue cuboid.

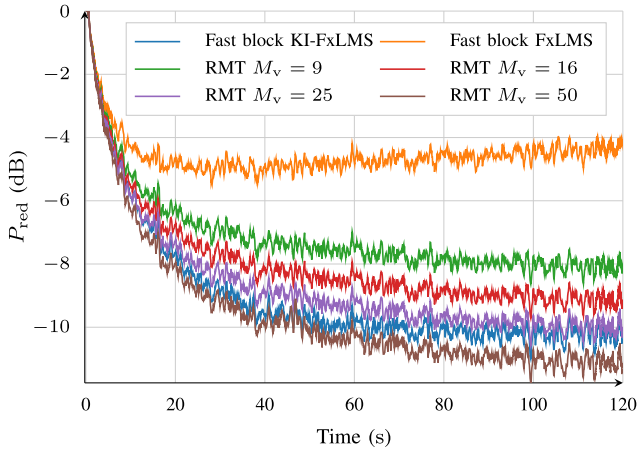


Fig. 12. Regional power reduction as a function of time, comparing fast block KI-FxLMS and fast block FxLMS to the RMT when different numbers of virtual microphones are used. The P_{red} for the last 60 s were $P_{\mathrm{red}} = -4.55$, $-10.14\,\mathrm{dB}$ for fast block FxLMS and fast block KI-FxLMS, and $P_{\mathrm{red}} = -7.91$, -8.93, -9.76, and $-10.86\,\mathrm{dB}$ for RMT with an increasing number of virtual microphones.

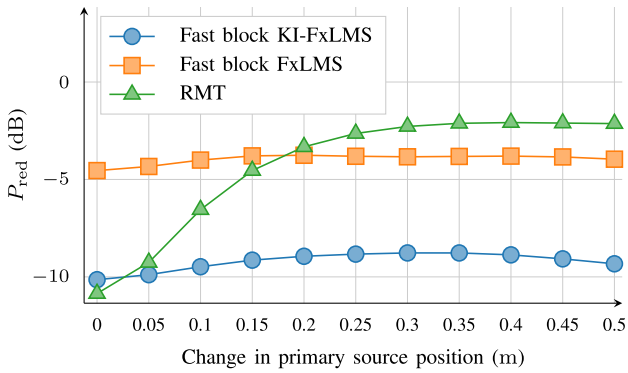


Fig. 13. Regional power reduction for the RMT with $M_{\mathrm{v}} = 50$ when the primary source position was changed in the positive y direction after the identification stage. The P_{red} is calculated in the interval 60 to 120 s.

As shown in Fig. 11, the target region Ω and the error microphone array were defined as in Section V-B with $C_\Omega = 0.6$. The numbers of error microphones and secondary sources were $M = 48$ and $L = 32$, respectively. The secondary sources were arranged along two squares of dimensions $2.0\,\mathrm{m} \times 2.0\,\mathrm{m}$ at the heights $z = 0.1\,\mathrm{m}$ and $-0.1\,\mathrm{m}$. The evaluation points were 972 equally spaced points $3.33 \times 10^{-2}\,\mathrm{m}$ apart inside the target region. The reference signal was directly obtained from the primary noise source.

The truncation length of the kernel interpolation filter was $K = 77$, chosen with some margin according to the heuristic presented in Section V-B. The sample rate, integration procedure, and regularization parameter were also the same as in Section V-B. The length of the control filter was $I = 2048$, and the block size B for the fast block algorithms was therefore also 2048. The observation filter length for the RMT was 256. The parameter β in the update of the step-size parameter (41) was 10^{-3}. The constant parameter μ_0 was chosen from $[10^1, 10^3]$, equally divided into 30 values in a logarithmic scale for the fast block KI-FxLMS. The same procedure was used for RMT and fast block FxLMS with the range of $[10^{-2}, 10^1]$.

First, we investigate the effect of the number of virtual microphones. We compare $M_{\mathrm{v}} = 9$, 16, 25, and 50 virtual microphones. The first three are distributed on equally spaced grids inside the target region at the height $z = 0\,\mathrm{m}$, and the final set is placed on two equally spaced grids of 5×5 microphones, at the heights $z = -0.025\,\mathrm{m}$ and $0.025\,\mathrm{m}$. The resulting noise reduction over time can be seen in Fig. 12, comparing against fast block KI-FxLMS and fast block FxLMS. The signal power interval was set to $T = 2048$ samples. It can be seen that for fast block FxLMS with the RMT to reach a performance comparable to the proposed method, a higher number of virtual microphones are necessary. A single P_{red} value is calculated at the end of the 120 s with the signal power interval set to 60 s, i.e., $T = 240000$ samples. The obtained values were $P_{\mathrm{red}} = -4.55$ and $-10.14\,\mathrm{dB}$ for fast block FxLMS and fast block KI-FxLMS respectively. For the RMT, the obtained values were $P_{\mathrm{red}} = -7.91$, -8.93, -9.76, and $-10.86\,\mathrm{dB}$ for increasing number of virtual microphones.

Next, we evaluate the algorithms' robustness against changes in the position of the primary noise source. The primary noise source position was changed in the positive y direction after the identification stage. For this simulation, $M_{\mathrm{v}} = 50$ virtual microphones were used. The resulting noise reduction P_{red} of the three compared algorithms can be seen in Fig. 13, as a function of the distance the primary source is moved after the identification stage. The signal power interval T was again set to 60 s, calculated after 120 s adaptation. The figure shows that the RMT can achieve a similar level of performance as the proposed method under ideal circumstances, when the conditions present during the identification stage remains. However, because the primary noise must be predicted using the pre-calculated observation filter, the method becomes sensitive to changes in the primary noise. In this scenario, only a 0.2 m movement of the primary source is enough for the RMT to degrade instead of improve FxLMS performance.

D. Time-Domain Experiments Using Real Data

Another set of time-domain simulations were conducted using real impulse response data measured in a practical environment. The impulse responses between the primary and secondary sources and the error microphones and evaluation points were measured one by one at each point using swept-sine signals [49]. The room dimensions were approximately $7.0\,\mathrm{m} \times 6.4\,\mathrm{m} \times 2.7\,\mathrm{m}$, and the reverberation time T_{60} was

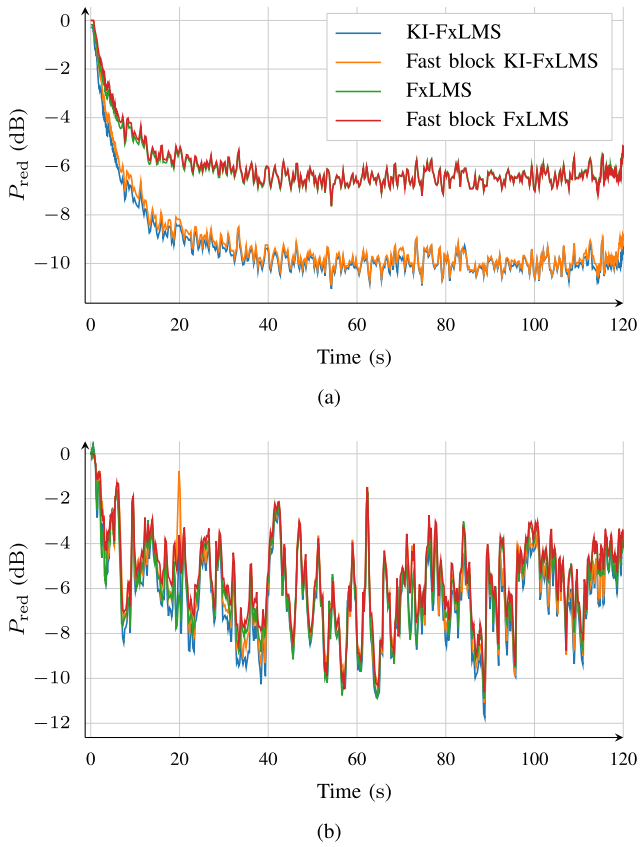


Fig. 14. Regional power reduction with respect to time. (a) **Noise**; (b) **Music**.

TABLE II
REGIONAL POWER REDUCTION IN dB FOR DIFFERENT NOISE SOURCES

	FxLMS	Fast block FxLMS	KI-FxLMS	Fast block KI-FxLMS
Noise	−6.41	−6.43	**−10.00**	−9.89
Music	−5.69	−5.34	**−6.20**	−5.78

about 0.38 s. The primary and secondary sources were ordinary closed loudspeakers, and the error microphone was omnidirectional.

The error microphones, target region and secondary sources were placed as described in Section V-C and shown in Fig. 11. We set 72 evaluation points inside Ω, where 36 points were regularly arranged at intervals of 0.1 m inside each of two squares of dimensions $0.6\,\mathrm{m} \times 0.6\,\mathrm{m}$ at the heights of $z = 0.025$ m and -0.025 m. The primary source was placed at $(-2.8, 0.3, 0.0)$ m. All relevant algorithm and simulation parameters were selected as in Section V-C. The stepsize for KI-FxLMS and FxLMS was determined from the same range as its fast block counterpart but divided by $B = 2048$.

We investigated two types of noise source for evaluation. The first is **Noise** as introduced in Section V-C. The other is the music signal *High Horse* by *Secret Mountains* from the dataset [50] (**Music**).

The regional power reduction P_{red} with respect to time is shown in Fig. 14, for which the signal power interval was chosen as $T = 2048$ samples. The P_{red} of each algorithm is shown in Table II, where the time interval for calculating P_{red} was set to 60 s, calculated after 120 s adaptation. For **Noise**, the resulting noise reduction by KI-FxLMS and its fast block version was much larger than that by FxLMS and its fast block version, because the proposed kernel-interpolation-based methods take the regional noise into consideration. KI-FxLMS was slightly better than fast block KI-FxLMS. FxLMS and fast block FxLMS differ less with regards to P_{red} as both are optimizing for the reduction of noise power at the error microphone positions instead. Since **Music** is nonstationary, the regional reduction varies significantly with time. However, as shown in Table II, the kernel-interpolation-based methods still outperformed the MPC-based methods. Since quick adaptation is necessary for nonstationary noises, the difference in P_{red} between KI-FxLMS and fast-block KI-FxLMS for **Music** was larger than that for **Noise**.

VI. CONCLUSION

We proposed a spatial ANC method based on kernel interpolation for the suppression of noise inside a continuous region. The noise between the error microphones is taken into account by estimating the sound field within the region using kernel interpolation. Whereas current spatial ANC methods based on spherical/circular harmonic expansion require a simple error microphone array geometry, the proposed method allows for arbitrary placement. A practical time-domain algorithm was derived, called the kernel-interpolation-based FxLMS (KI-FxLMS), which features only a marginal increase in computational complexity over FxLMS. We also presented a computationally efficient block-based implementation using FFT, the fast block KI-FxLMS. In the presented experiments, the proposed kernel-interpolation-based methods consistently outperformed the conventional multipoint-pressure-control-based methods with regards to regional power reduction. Compared with the virtual sensing technique, robustness against changes in the primary noise position was also shown. Although the fast block KI-FxLMS is more computationally efficient, KI-FxLMS demonstrated better performance, especially for nonstationary noise sources.

APPENDIX
KERNEL FUNCTION WITH DIRECTIONAL WEIGHTING

We here derive the kernel function incorporating directional weighting. Such kernel functions are applicable when prior information on the approximate source direction is available [29], [51]. First, we introduce the following relation:

$$\frac{1}{4\pi}\int_{\mathbb{S}_2} e^{-\mathrm{j}\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{v}}\mathrm{d}\boldsymbol{\xi} = j_0\left(\left(v_1^2+v_2^2+v_3^2\right)^{\frac{1}{2}}\right), \tag{49}$$

where $\boldsymbol{v} = [v_1, v_2, v_3]^{\mathsf{T}} \in \mathbb{C}^3$. The proof is given in [29].

We here consider the following directional weighting function:

$$\gamma(\boldsymbol{\xi}) := e^{\rho\boldsymbol{\xi}^{\mathsf{T}}\hat{\boldsymbol{r}}}, \tag{50}$$

where $\rho \geq 0$ is a constant parameter and $\hat{\boldsymbol{r}} \in \mathbb{S}_2$ represents the prior arrival direction of the source. This function is derived from the von Mises–Fisher distribution in directional statistics [52]. By using (50) for $\rho > 0$, one can find that the smaller $\rho\boldsymbol{\xi}^{\mathsf{T}}\hat{\boldsymbol{r}}$ is, the larger the norm $\|u\|_{\mathcal{H}}$ in (15) becomes, and vice versa. That is, the regularization term in (4) becomes larger when the

difference between the prior arrival direction $\hat{\boldsymbol{r}}$ and the direction of $\boldsymbol{\xi}$ becomes larger. The shape of $\gamma(\boldsymbol{\xi})$ becomes sharper with increasing ρ.

By substituting (50) into (16) and using (49), one can derive the kernel function with directional weighting as

$$\begin{aligned}\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \frac{1}{4\pi} \int_{\mathbb{S}_2} e^{\rho \boldsymbol{\xi}^{\mathsf{T}} \hat{\boldsymbol{r}}} \cdot e^{-\mathrm{j}\boldsymbol{k}^{\mathsf{T}}(\boldsymbol{r}_1 - \boldsymbol{r}_2)} \mathrm{d}\boldsymbol{\xi} \\ &= j_0\Big(\big[(\mathrm{j}\rho \sin\theta \cos\phi - \mathbb{k} x_{12})^2 \\ &\quad + (\mathrm{j}\rho \sin\theta \sin\phi - \mathbb{k} y_{12})^2 + (\mathrm{j}\rho \cos\theta - \mathbb{k} z_{12})^2 \big]^{\frac{1}{2}} \Big), \end{aligned} \tag{51}$$

where ϕ and θ are the azimuth and zenith angles of $\hat{\boldsymbol{r}}$, and $\boldsymbol{r}_1 - \boldsymbol{r}_2 := [x_{12}, y_{12}, z_{12}]^{\mathsf{T}}$. By setting $\rho = 0$, one can find that (51) corresponds to (18). In the 2D case, the kernel function with directional weighting is similarly derived as

$$\begin{aligned}\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) &= J_0\left(\big[(\mathrm{j}\rho \cos\phi - \mathbb{k} x_{12})^2 + (\mathrm{j}\rho \sin\phi - \mathbb{k} y_{12})^2 \big]^{\frac{1}{2}} \right), \end{aligned} \tag{52}$$

where ϕ is the azimuth angle of $\hat{\boldsymbol{r}}$. Again, for $\rho = 0$, (52) corresponds to (19). Although we focused on the kernel function of the uniform weighting function (18) and (19) in this paper, (51) and (52) are also applicable by replacing the kernel functions [51].

REFERENCES

[1] P. A. Nelson and S. J. Elliott, *Active Control of Sound*. London, U.K.: Academic Press, 1992.

[2] S. J. Elliott and P. A. Nelson, "Active noise control," *IEEE Signal Process. Mag.*, vol. 10, no. 4, pp. 12–35, Oct. 1993.

[3] S. M. Kuo and D. R. Morgan, "Active noise control: A tutorial review," *Proc. IEEE*, vol. 87, no. 6, pp. 943–973, Jun. 1999.

[4] Y. Kajikawa, W.-S. Gan, and S. M. Kuo, "Recent advances on active noise control: Open issues and innovative applications," *APSIPA Trans. Signal Inform. Process.*, vol. 1, no. e3, pp. 1–21, 2012.

[5] H. Sano, T. Inoue, A. Takahashi, K. Terai, and Y. Nakamura, "Active control system for low-frequency road noise combined with an audio system," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 7, pp. 755–763, Oct. 2001.

[6] S. M. Kuo, S. Mitra, and W.-S. Gan, "Active noise control system for headphone applications," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 2, pp. 331–335, Mar. 2006.

[7] S. Chakravarthy and S. M. Kuo, "Application of active noise control for reducing snore," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2006, pp. V. 305–308.

[8] G. Kumamoto, M. Kida, R. Hirayama, Y. Kajikawa, T. Tani, and Y. Kurumi, "Active noise control system for reducing MR noise," *IEICE Trans. Fundamentals*, vol. 19, no. 2, pp. 474–480, 2011.

[9] S. J. Elliott, I. M. Stothers, and P. A. Nelson, "A multiple error LMS algorithm and its application to the active control of sound and vibration," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1423–1434, Oct. 1987.

[10] D. J. Moreau, B. S. Cazzolato, A. C. Zander, and C. D. Petersen, "A review of virtual sensing algorithms for active noise control," *Algorithms*, vol. 1, no. 2, pp. 69–99, 2008.

[11] G. Bonito, S. J. Elliott, and C. C. Boucher, "Generation of zones of quiet using a virtual microphone arrangement," *J. Acoust. Soc. Amer.*, vol. 101, no. 6, pp. 3498–3516, 1997.

[12] D. J. Moreau, J. Ghan, B. S. Cazzolato, and A. C. Zander, "Active noise control in a pure tone diffuse sound field using virtual sensing," *J. Acoust. Soc. Amer.*, vol. 125, no. 6, pp. 3742–3755, 2009.

[13] S. Spors and H. Buchner, "An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering," in *Proc. IEEE Int. Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, 2007, pp. 171–174.

[14] P. N. Samarasinghe, W. Zhang, and T. D. Abhayapala, "Recent advances in active noise control inside automobile cabins: Toward quieter cars," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 61–73, Nov. 2016.

[15] J. Zhang, T. D. Abhayapala, W. Zhang, P. N. Samarasinghe, and S. Jiang, "Active noise control over space: A wave domain approach," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 4, pp. 774–786, Apr. 2018.

[16] J. Donley, C. Ritz, and W. B. Kleijn, "On the comparison of two room compensation / dereverberation method employing active acoustic boundary absorption," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2018, pp. 221–225.

[17] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Feedforward spatial active noise control based on kernel interpolation of sound field," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2019, pp. 511–515.

[18] N. Murata, J. Zhang, Y. Maeno, and Y. Mitsufuji, "Global and local mode-domain adaptive algorithms for spatial active noise control using higher-order sources," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2019, pp. 526–530.

[19] Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, N. Murata, and T. D. Abhayapala, "Spherical-harmonic-domain feedforward active noise control using sparse decomposition of reference signals from distributed sensor arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 656–670, 2020.

[20] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th AES Conv.*, Amsterdam, The Netherlands, 2008, Art. no. 7358.

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

[22] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 4, pp. 685–696, Apr. 2013.

[23] S. Koyama, K. Furuya, K. Wakayama, S. Shimauchi, and H. Saruwatari, "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," *J. Acoust. Soc. Amer.*, vol. 139, no. 3, pp. 1024–1036, 2016.

[24] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. London, U.K.: Academic Press, 1999.

[25] H. Sun, T. D. Abhayapala, and P. N. Samarasinghe, "Time domain spherical harmonic analysis for adaptive noise cancelation over a spatial region," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2019, pp. 516–539.

[26] F. Ma, W. Zhang, and T. D. Abhayapala, "Active control of outgoing broadband noise fields in rooms," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 529–539, Dec. 2019, doi: 10.1109/TASLP.2019.2960716.

[27] N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of helmholtz equation for sound field interpolation," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2018, pp. 436–440.

[28] 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, Jan. 2018.

[29] N. Ueno, S. Koyama, and H. Saruwatari, "Directionally weighted wave field estimation exploiting prior information on source direction," *IEEE Trans. Signal Process.*, vol. 69, pp. 2383–2395, Apr. 2021, doi: 10.1109/TSP.2021.3070228.

[30] J. Brunnström and S. Koyama, "Kernel-interpolation-based filtered-x least mean square for spatial active noise control in time domain," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2021, pp. 161–165.

[31] S. Koyama, G. Chardon, and L. Daudet, "Optimizing source and sensor placement for sound field control: An overview," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 686–714, Jan. 2020, doi: 10.1109/TASLP.2020.2964958.

[32] K. P. Murphy, *Machine Learning: A. Probabilistic Perspective*. Cambridge, U.K.: MIT Press, 2012.

[33] B. Schölkopf, R. Herbrich, and A. J. Smola, "A generalized representer theorem," in *Proc. Int. Conf. Comput. Learn. Theory (COLT)*, 2001, pp. 416–426.

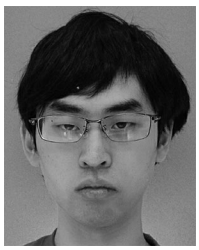
[34] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. New York, NY, USA: Springer, 2013.

[35] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson Jr., "Measurement of correlation coefficients in reverberant sound fields," *J. Acoust. Soc. Amer.*, vol. 27, no. 6, pp. 1072–1077, 1955.

[36] S. Fine and K. Scheinberg, "Efficient SVM training using low-rank kernel representations," *J. Mach. Learn. Res.*, vol. 2, pp. 243–264, 2001.
[37] C. K. I. Williams and M. Seeger, "Using the nyström method to speed up kernel machines," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2000, pp. 682–688.
[38] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Three-dimensional spatial active noise control based on kernel-induced sound field interpolation," in *Proc. Int. Congr. Acoust. (ICA)*, 2019, pp. 1101–1108.
[39] S. Johansson, S. Nordebo, T. Lagö, P. Sjösten, and I. Claesson, "A novel multiple-reference, multiple-channel, normalized filtered-x LMS algorithm for active control of propeller-induced noise in aircraft cabins," in *Proc. Int. Congr. Expo. Noise Cont. Eng. (INTER-NOISE)*, 1998, pp. 664–667.
[40] D. Shi, W.-S. Gan, B. Lam, W. Shulin, and X. Shen, "Active noise control based on momentum multichannel normalized filtered-x least mean square algorithm," in *Proc. Int. Congr. Expo. Noise Cont. Eng. (INTER-NOISE)*, Seoul, Korea, 2020, pp. 709–714.
[41] I. Chung, "Multi-channel normalized FxLMS algorithm for active noise control," *J. Acoust. Soc. Korea*, vol. 35, pp. 280–287, Jul. 2016.
[42] D. P. Das, G. Panda, and S. M. Kuo, "New block filtered-x LMS algorithms for active noise control systems," *IET Signal Process.*, vol. 1, no. 2, pp. 73–81, 2007.
[43] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.
[44] L. J. Eriksson and M. C. Allie, "Use of random noise for online transducer modeling in an adaptive active attenuation system," *J. Acoust. Soc. Amer.*, vol. 85, no. 2, pp. 797–802, 1989.
[45] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, *3rd ed.*, New York, NY, USA: Cambridge Univ. Press, 2007.
[46] S. J. Elliott, W. Jung, and J. Cheer, "Causality and robustness in the remote sensing of acoustic pressure, with application to local active sound control," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2019, pp. 8484–8488.
[47] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, 1979.
[48] R. Scheibler, E. Bezzam, and I. Dokmanić, "Pyroomacoustics: A python package for audio room simulation and array processing algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2018, pp. 351–355.
[49] Y. Suzuki, F. Asano, H.-Y. Kim, and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," *J. Acoust. Soc. Amer.*, vol. 97, no. 2, pp. 1119–1123, 1995.
[50] Z. Rafii, A. Liutkus, F. Stöter, S. I. Mimilakisand, and R. Bittner, "MUSDB-18HQ - an uncompressed version of MUSDB18," Dec. 2019. [Online]. Available: https://doi.org/10.5281/zenodo.3338373
[51] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Spatial active noise control based on kernel interpolation with directional weighting," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2020, pp. 8399–8403.
[52] K. V. Mardia and P. E. Jupp, *Directional Statistics*. Chichester, U.K.: Wiley, 2009.

Shoichi Koyama (Member, IEEE) received the B.E., M.S., and Ph.D. degrees from the University of Tokyo, Tokyo, Japan, in 2007, 2009, and 2014, respectively. In 2009, he joined Nippon Telegraph and Telephone Corporation, Tokyo, Japan, as a Researcher in acoustic signal processing. In 2014, he moved to the University of Tokyo and has been a 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 audio signal processing, acoustic inverse problems, and spatial audio. He was the recipient of Itakura Prize Innovative Young Researcher Award by the Acoustical Society of Japan in 2015, and the Research Award by Funai Foundation for Information Technology in 2018.

Jesper Brunnström received the M.Sc. degree in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden, in 2021. He is currently working toward the Ph.D. degree with KU Leuven, Leuven, Belgium, as an early stage Researcher in the Marie Curie European training network, Service-Oriented, Ubiquitous, and Network-Driven Sound. Between 2019 and 2020, he was a Research Student with the University of Tokyo, Tokyo, Japan. His research interests include active noise control, sound field reproduction, acoustic signal processing, and distributed signal processing.



Hayato Ito received the B.A. degree in liberal arts and M.S. degree in information science and technology from the University of Tokyo, Tokyo, Japan, in 2018 and 2020, respectively. His research interests include spatial audio and active noise control.

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

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