forum acusticum 2026
8-12 September · Graz, Austria

SOUND FIELD ESTIMATION IN CONTINUOUS TIME USING KERNEL RIDGE REGRESSION

Jesper Brunnström

Department of Information Technology, Uppsala University, Sweden

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

Kernel ridge regression (KRR) has emerged as an effective approach to sound field estimation (SFE). By constructing the kernel and its associated reproducing kernel Hilbert space (RKHS) appropriately, the optimal KRR solution can be guaranteed to satisfy the wave equation. Such kernels and RKHSs have so far only been derived in the frequency domain and discrete-time domain. Therefore, in this paper a continuous-time kernel and associated RKHS is proposed, which enforces the wave equation by construction. The kernel function has a simple closed-form expression which is shown to have lower computational cost compared to similar discrete-time methods. A KRR-based method for SFE with the proposed model is then presented. The proposed SFE method is flexible, naturally supporting moving microphones and non-uniform sampling. The proposed method is evaluated using simulated data with both stationary and moving microphones, demonstrating the value of the approach in terms of both estimation performance and computational cost.

Keywords: *sound field estimation, moving microphone, kernel ridge regression, time domain*

1. Introduction

Sound field estimation (SFE) is of central importance for many acoustic signal processing problems such as sound field reproduction [1, 2], source characterization [3], and room acoustics modelling [4]. The task is to reconstruct the sound pressure in a continuous region from a discrete set of measurements. It is an ill-posed inverse problem, where prior knowledge must be leveraged for accurate estimates. The sound field is often modelled as a superposition of solutions to the wave equation [5], with which the inverse problem can be solved using optimization-based [6] [7, Ch. 2] or Bayesian methods [8, 9].

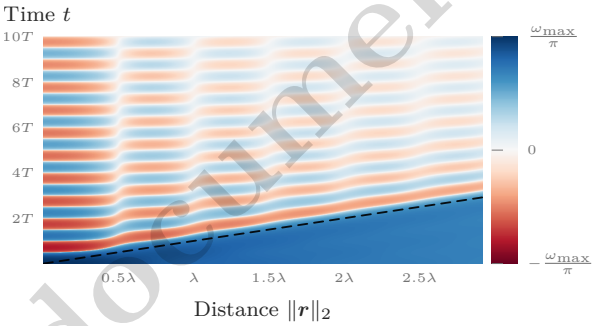

Figure 1: Depiction of the continuous-time kernel $\kappa(\boldsymbol{r}, t, \boldsymbol{0}, 0)$. The wavelength at the chosen bandlimit $\omega_{\max}$ is denoted by λ, and its period by T. The dashed line has a slope of c, the speed of sound. The data is scaled by the inverse hyperbolic sine to aid visibility.

Kernel ridge regression (KRR) [10] has emerged as an effective approach to SFE [11]. The approach allows for computationally tractable optimization over the infinite-dimensional space of solutions to the governing differential equation. By carefully constructing the reproducing kernel Hilbert space (RKHS) and its unique kernel function, the estimates can be guaranteed to satisfy the wave equation. The approach also allows for different types of prior knowledge to be used to improve the estimates [12–15]. KRR is closely related to infinite-dimensional Bayesian inference [16] and Gaussian process regression [17].

KRR for SFE have mostly been formulated in the frequency domain, which limits the types of measurements and prior knowledge that can be used. A continuous-time kernel is derived in [18], which relies on deep kernel learning to enforce the wave equation. The kernel in [15, 19] enforces the wave equation by construction, but is only defined in discrete time. In this paper, a continuous-time kernel is derived, illustrated in Fig. 1, which enforces the wave equation by construction and can be computed in closed form.

An advantage over the frequency-domain formulation

***Corresponding author**: jesper.brunnstrom@it.uu.se.*

is that moving microphones can be straightforwardly used. Methods for SFE with moving microphones that require estimates to satisfy the wave equation have been formulated in [19–22], among which the KRR method [19] has been shown to be a robust and effective estimator [23]. An advantage over the discrete-time formulation is that measurements with non-uniform sampling can be straightforwardly used. In addition, due to the simple closed-form expression of the kernel, the computational cost can be significantly reduced, while providing essentially identical estimation performance.

2. Frequency-domain sound field model

In this section, a frequency-domain sound field model used previously [11–13] is summarized. It will be central to defining the time-domain model in Section 3. The complex sound pressure at angular frequency $\omega \in \mathbb{R}_{\geq 0}$ can be represented by a function $u_\omega : \Omega \to \mathbb{C}$ defined on the simply connected region $\Omega \subset \mathbb{R}^3$. Assuming that Ω is source-free, u_ω satisfies the homogenous Helmholtz equation

$$\Big(\Delta + \frac{\omega^2}{c^2}\Big) u_\omega = 0. \tag{1}$$

where Δ is the Laplace operator, and c is the speed of sound. The solutions to (1) can be well represented by Herglotz wave function [24, Def. 3.26]

$$u_\omega(\boldsymbol{r}) = \int_{\mathbb{S}_2} e^{-j\frac{\omega}{c}\boldsymbol{r}^\top \boldsymbol{d}} \mathring{u}_\omega(\boldsymbol{d}) ds(\boldsymbol{d}), \tag{2}$$

where the integral is with regard to the natural spherical measure ds. The coefficient of each plane wave is described by $\mathring{u}_\omega \in L_2(\mathbb{S}_2)$, a square-integrable function [25, Section 2.4.6] on the unit sphere $\mathbb{S}_2$.
A RKHS $\mathcal{H}_\omega$ with functions constrained by (1) can be defined as functions of the form (2) with inner product

$$\langle u_\omega, v_\omega \rangle_{\mathcal{H}_\omega} = \int_{\mathbb{S}_2} \mathring{u}_\omega(\boldsymbol{d}) \overline{\mathring{v}_\omega(\boldsymbol{d})}\, ds(\boldsymbol{d}). \tag{3}$$

The RKHS has a unique kernel function $\kappa_\omega : \Omega \times \Omega \to \mathbb{C}$, which is

$$\kappa_\omega(\boldsymbol{r}, \boldsymbol{r}') = j_0\Big(\frac{\omega}{c}\|\boldsymbol{r} - \boldsymbol{r}'\|_2\Big), \tag{4}$$

where j_0 is the zeroth order spherical Bessel function of the first kind. The kernel function satisfies the reproducing property $u_\omega(\boldsymbol{r}) = \langle u_\omega, \kappa_\omega(\cdot, \boldsymbol{r})\rangle_{\mathcal{H}_\omega}$.

3. Time-domain sound field model

In this section, a model for spatio-temporal sound pressure functions $u : \Omega \times \mathbb{R} \to \mathbb{R}$ is derived. With this model, a RKHS and its associated kernel function can be constructed, which will later be used for KRR. The function u is related to the frequency-domain function u_ω through the Fourier transform $u(\boldsymbol{r}, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u_\omega(\boldsymbol{r}) e^{j\omega t}\, d\omega$. That u is real-valued implies that u_ω is conjugate symmetric, i.e. $u_\omega = \overline{u_{-\omega}}$. The function u is also assumed to be temporally bandlimited to $\omega_{\max}$, which means that $u_\omega = 0$ for all $\omega \geq \omega_{\max}$. Then, u can be expressed as

$$u(\boldsymbol{r}, t) = \frac{1}{\pi}\mathfrak{Re}\Big[\int_0^{\omega_{\max}} u_\omega(\boldsymbol{r}) e^{j\omega t}\, d\omega\Big]. \tag{5}$$

The RKHS $\mathcal{H}$ for time-domain sound fields is defined as functions of the form (5) with inner product

$$\langle u, v\rangle_{\mathcal{H}} = \frac{1}{\pi}\mathfrak{Re}\Big[\int_0^{\omega_{\max}} \langle u_\omega, v_\omega\rangle_{\mathcal{H}_\omega}\, d\omega\Big]. \tag{6}$$

The unique kernel function $\kappa : \Omega^2 \times \mathbb{R}^2 \to \mathbb{R}$ for $\mathcal{H}$ that satisfies the reproducing property $u(\boldsymbol{r}, t) = \langle u, \kappa(\cdot, \cdot, \boldsymbol{r}, t)\rangle_{\mathcal{H}}$ is

$$\kappa(\boldsymbol{r}, t, \boldsymbol{r}', t') = \frac{1}{\pi}\mathfrak{Re}\Big[\int_0^{\omega_{\max}} \kappa_\omega(\boldsymbol{r}, \boldsymbol{r}') e^{j\omega(t-t')}\, d\omega\Big]. \tag{7}$$

The kernel in (7) has a closed-form expression in terms of the sine integral function $\mathrm{Si}(t) = \int_0^t \frac{\sin t}{t} dt$, which is

$$\begin{aligned}\kappa(\boldsymbol{r}, t, \boldsymbol{r}', t') = \frac{1}{2\pi\tau}\Big(&\mathrm{Si}(\omega_{\max}(t - t' + \tau)) \\ &- \mathrm{Si}(\omega_{\max}(t - t' - \tau))\Big),\end{aligned} \tag{8}$$

where $\tau = \frac{\|\boldsymbol{r} - \boldsymbol{r}'\|_2}{c}$ is the propagation time between two points in space. When $\boldsymbol{r} = \boldsymbol{r}'$, which implies that $\tau = 0$, (8) can be identified as $\kappa(\boldsymbol{r}, t, \boldsymbol{r}', t') = \frac{\sin(\omega_{\max}(t-t'))}{\pi(t-t')}$. When both $\boldsymbol{r} = \boldsymbol{r}'$ and $t = t'$, then $\kappa(\boldsymbol{r}, t, \boldsymbol{r}', t') = \frac{\omega_{\max}}{\pi}$.

4. Sound field estimation with stationary microphones

The proposed sound field model in Section 3 provides a natural setting for analyzing and solving problems involving spatio-temporal sound fields. The problem considered in this section is the estimation of room impulse responses (RIRs) in a continuous region using a set of stationary microphones.

4.1. Problem statement

A single sound source of interest emits the known source signal s. The acoustic paths from the source to each point $\boldsymbol{r} \in \Omega$ are assumed to be well modelled as linear and time-invariant, and can therefore be represented by the RIR function $h : \Omega \times \mathbb{R} \to \mathbb{R}$. The sound pressure $p : \Omega \times \mathbb{R} \to \mathbb{R}$ generated by the source can then be described by the convolution

$$p(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} h(\boldsymbol{r}, t') s(t - t')\, dt'. \tag{9}$$

Omnidirectional microphones are placed at M positions $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_M$ with which the signal (9) is sampled synchronously at a sampling rate of f_s, i.e. at the time steps $t_l = \frac{l}{f_s}$. Measurements are collected for enough time that deconvolution is possible, which leads to the

data model

$$h_{ml} = h(\boldsymbol{r}_m, t_l) + \epsilon_m(t_l), \quad \begin{matrix} l \in \{0, \dots, L-1\} \\ m \in \{1, \dots, M\} \end{matrix}, \quad (10)$$

where ϵ is additive noise dependent on noise sources, deconvolution method, and the acoustic environment. It is assumed that the RIR function is effectively non-zero only where $t \in [0, T]$, due to causality and the exponential decay naturally exhibited by RIRs. Therefore, $L \geq T f_s$ samples are sufficient to represent each RIR. The SFE task is to estimate h over the continuous region Ω from data of the form (10).

4.2. Kernel ridge regression

KRR can be used to estimate the RIR function using the data in (10), which can be stated as the optimization problem

$$\underset{u \in \mathcal{H}}{\text{minimize}} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \big(h_{ml} - u(\boldsymbol{r}_m, t_l)\big)^2 + \lambda \|u\|_{\mathcal{H}}^2. \quad (11)$$

The optimal solution of (11) is characterized as [10]

$$u(\boldsymbol{r}, t) = \sum_{m=1}^{M} \sum_{l=0}^{L-1} \kappa(\boldsymbol{r}, t, \boldsymbol{r}_m, t_l) a_{ml}, \quad (12)$$

where $a_{ml} \in \mathbb{R}$ is an a priori unknown parameter. Inserting (12) into (11), the closed-form solution is

$$\boldsymbol{a} = (\boldsymbol{K} + \lambda \boldsymbol{I})^{-1} \boldsymbol{h}, \quad (13)$$

where $\boldsymbol{a} = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_M) \in \mathbb{R}^{ML}$, $\boldsymbol{a}_m = (a_{m1}, \dots, a_{mL}) \in \mathbb{R}^L$, and $\boldsymbol{h} \in \mathbb{R}^{ML}$ is constructed from h_{ml} analogously. The row and column described by ml and $m'l'$ respectively of $\boldsymbol{K} \in \mathbb{R}^{ML \times ML}$ is $\boldsymbol{K}_{mlm'l'} = \kappa(\boldsymbol{r}_m, t_l, \boldsymbol{r}_{m'}, t_{l'})$. The parameters $\boldsymbol{a}$ can then be used in (12) to evaluate the estimated RIR function at arbitrary points in time and space.

5. Sound field estimation with moving microphones

The proposed sound field model is also naturally suited to SFE with moving microphones, which is considered in this section.

5.1. Problem statement

The sound field within Ω is generated by a single source of interest as described by (9). Measurements are collected with a continuously moving microphone, sampling the signal at N points $(\boldsymbol{r}_1, t_1), \dots, (\boldsymbol{r}_N, t_N)$. The measurements can be described as

$$p_n = p(\boldsymbol{r}_n, t_n) + \epsilon(t_n), \quad n \in \{0, \dots, N-1\}. \quad (14)$$

where ϵ represents additive noise.

It is now assumed that the source emits a periodic signal with period T, equal to the RIR length. This is common in acoustic measurements when the source is a loudspeaker with a predetermined training signal, such as a periodic sweep [26, 27]. This implies that the sound field can be modelled as periodic, which leads to $t_n = \frac{n \bmod L}{f_s}$, where $L = T f_s$ is for simplicity assumed to be an integer. The sound field is only strictly periodic in the noise-free case, which can be relevant if the temporal correlation of the noise is to be exploited, which is not the case here.

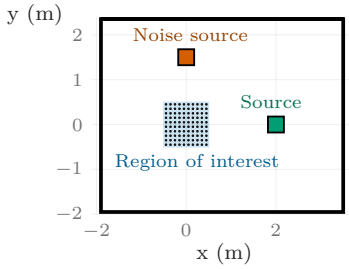

Figure 2: Positions of the region of interest, source, noise source, evaluation points, and simulated room walls.

5.2. Kernel ridge regression

It is difficult to directly deconvolve the measured sound pressure with a moving microphone. Instead, the periodic sound pressure function p can be estimated, and then deconvolved to h for chosen positions $\boldsymbol{r} \in \Omega$. Using the data model (14), the KRR optimization problem to estimate p can be stated as

$$\underset{u \in \mathcal{H}}{\text{minimize}} \sum_{n=0}^{N-1} \big(p_n - u(\boldsymbol{r}_n, t_n)\big)^2 + \lambda \|u\|_{\mathcal{H}}^2. \quad (15)$$

As in the stationary microphone case, the optimal estimate can be characterized as [10]

$$u(\boldsymbol{r}, t) = \sum_{n=0}^{N-1} \kappa(\boldsymbol{r}, t, \boldsymbol{r}_n, t_n) a_n, \quad (16)$$

where $a_n \in \mathbb{R}$ is an a priori unknown parameter. The closed-form solution is

$$\boldsymbol{a} = (\boldsymbol{K} + \lambda \boldsymbol{I})^{-1} \boldsymbol{p}, \quad (17)$$

where the parameters and data are structured into vectors as $\boldsymbol{a} = (a_0, \dots, a_{N-1}) \in \mathbb{R}^N$ and $\boldsymbol{p} = (p_0, \dots, p_{N-1}) \in \mathbb{R}^N$. The kernel matrix $\boldsymbol{K} \in \mathbb{R}^{N \times N}$ is defined as $\boldsymbol{K}_{nn'} = \kappa(\boldsymbol{r}_n, t_n, \boldsymbol{r}_{n'}, t_{n'})$. The parameters from (17) can then be inserted into (16) for an estimate of p at an arbitrary point in time and space, which can be deconvolved for an estimate of h.

6. Evaluation

The proposed method, referred to as C-KRR, is evaluated using simulations. The experiments use RIRs generated with the image-source method [28, 29] at a sampling rate of $f_s = 1400\,\text{Hz}$ and reverberation time of 0.45 s. The bandlimit is then set according to the Nyquist frequency as $\omega_{\max} = \pi f_s$. The RIRs are high-pass filtered with cut-off frequency 20 Hz to simulate the low-frequency attenuation of a loudspeaker.

The geometry of the scenario is shown in Fig. 2. The

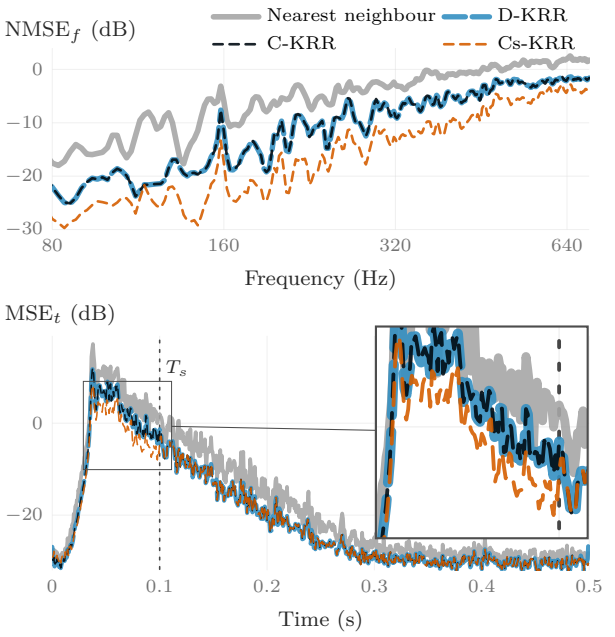

Figure 3: Estimation performance as a function of frequency (top) and time (bottom) using stationary microphones.

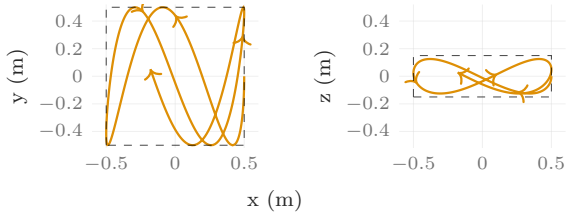

Figure 4: Microphone trajectory for the simulated moving microphone when $N = 22400$.

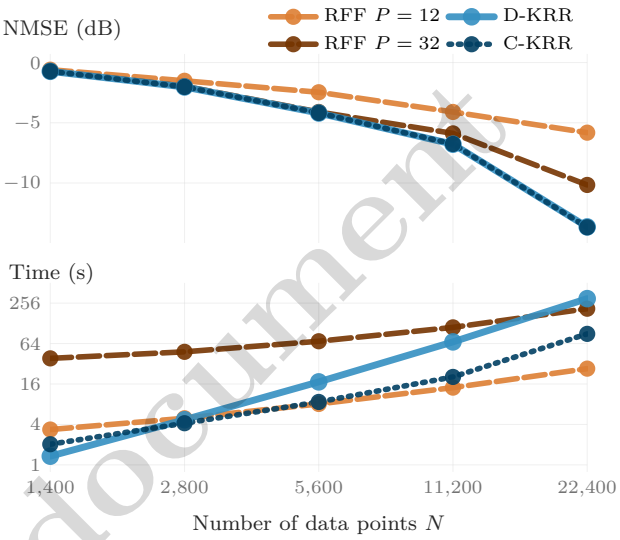

Figure 5: Estimation performance (top) and computation time (bottom) as a function of the amount of data N using a moving microphone.

region of interest is a cuboid of size $1 \times 1 \times 0.25\,\mathrm{m}^3$. A noise source emits white Gaussian noise scaled such that the measured signal-to-noise ratio (SNR) is 30 dB. The source emits a periodic perfect sweep [26] of length $T = 0.5\,\mathrm{s}$, which implies $L = 700$. Each RIR is obtained from L samples, which is deconvolved by applying an orthogonal transform defined by the source signal [26]. This orthogonal deconvolution process preserves the SNR.

The estimated RIR function is evaluated at the times $t_l = \frac{l}{f_s}$ for $l \in \{0, \ldots, L-1\}$ and positions $\boldsymbol{r} \in \mathcal{E}$, where $\mathcal{E}$ is a set of 200 evaluation points placed on a 10 cm equidistant grid in the region of interest. The primary error metric is the normalized mean square error (NMSE). The alternative metric NMSE$_f$ is computed and normalized at each frequency individually. NMSE and NMSE$_f$ are computed using an L-length discrete Fourier transform (DFT), where the frequency bins below 20 Hz are excluded. MSE$_t$ is non-normalized, and computed for each time step t_l individually. All displayed results are the mean results from 10 random seeds. The regularization parameter for each method is chosen as 10^{α} according to the value in $\alpha \in \{-6, \ldots, 0\}$ that minimizes the NMSE.

6.1. Stationary microphones

First, SFE with stationary microphones as described in Section 4 is considered. RIRs are measured with $M = 8$ microphones placed randomly with uniform distribution in the region of interest. The discrete-time method [15] is considered, referred to as D-KRR, in this case equivalent to the frequency-domain method [11] as shown in [15, Appendix D]. Nearest neighbour interpolation is included as a baseline.

To demonstrate the value of non-uniform sampling, an additional $M_s = 8$ short measurements are added, which only contain data for the initial $T_s = 0.1\,\mathrm{s}$ of the RIR. Short measurement times have been shown to be practical for sound field control, where estimates of the initial part of the RIR are crucial [30]. This data can be straightforwardly used together with the data from the original M microphones using the proposed method, and this estimate is referred to as Cs-KRR. The NMSE$_f$ and MSE$_t$ are shown in Fig 3. The proposed C-KRR using the same data as D-KRR obtains effectively identical error. The proposed Cs-KRR has improved performance between 0 to 0.1 s and effectively identical performance elsewhere compared to D-KRR and C-KRR, owing to the additional data. All methods significantly outperform nearest neighbour.

6.2. Moving microphones

Next, SFE with moving microphones as described in Section 5 is considered. A moving microphone is simulated to move along a Lissajous curve at 0.5 m/s, depicted in Fig. 4.

The proposed C-KRR and the discrete-time KRR method [19] referred to as D-KRR, are compared

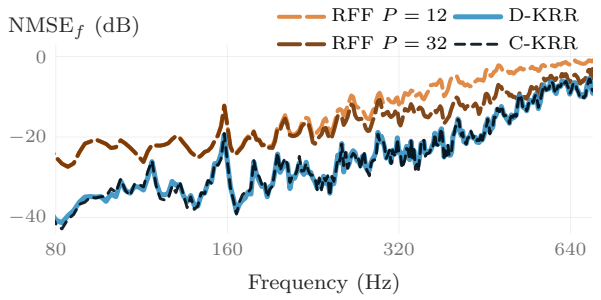

Figure 6: Estimation performance as a function of frequency, using a moving microphone and $N = 22400$.

against a SFE method using a finite-dimensional plane wave model, referred to as random Fourier features (RFF) [19, Section V.B][20]. RFF can be viewed as a finite-dimensional approximation of KRR [19]. The plane wave directions are chosen according to a spherical t-design [31], rotated by a random angle with uniform distribution on the sphere. A t-design of order 5 and 7 are used with $P = 12$ and $P = 32$ directions respectively.

Using data from the first 1, 2, 4, 8 and 16 s of the trajectory, the resulting NMSE is shown in Fig. 5. The NMSE$_f$ is shown in Fig. 6 for $N = 22400$. Both KRR methods consistently outperform both RFF methods, especially when more data is available, i.e. when N is larger. The proposed C-KRR performs effectively identically to D-KRR across all frequencies and amounts of data.

The computational cost quickly becomes prohibitively large for larger sampling rates and longer measurement times, making computational efficiency important for SFE methods with moving microphones. The time required to compute the estimates using the CPU of a laptop with an Intel Core Ultra 7 165U processor is shown in Fig. 5. The methods are implemented in Python[1] as JIT-compiled code using JAX [32]. The proposed C-KRR is faster than D-KRR, due to the simple expression of the kernel. This demonstrates the usefulness of the proposed model, as the two methods provide effectively identical estimation performance. For very large N, the computational cost of both C-KRR and D-KRR is dominated by solving a linear system with N unknowns, at which point the computational cost can be expected to be more similar. As expected, the computational cost of RFF is less dependent on N, which makes it faster to compute for large N, at the cost of higher NMSE.

7. Conclusion

A continuous-time sound field model for KRR-based SFE have been proposed, with a kernel and RKHS that enforces the wave equation by construction. The approach have been shown to be effective and flexible, naturally supporting non-uniformly sampled data and moving microphones for SFE. The proposed kernel can be expressed in closed form, leading to lower computational cost compared to discrete-time methods for SFE with moving microphones.

[1] Code associated with the investigated methods are available at github.com/sounds-research/aspcol

References

[1] T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, "Personal sound zones: Delivering interface-free audio to multiple listeners," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 81–91, 2015. DOI: 10.1109/MSP.2014.2360707

[2] W. Zhang, P. N. Samarasinghe, H. Chen, and T. D. Abhayapala, "Surround by sound: A review of spatial audio recording and reproduction," *Appl. Sci.*, vol. 7, no. 5, 2017, Art. no. 532. DOI: 10.3390/app7050532

[3] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality* (RWTHedition). Springer, 2020. DOI: 10.1007/978-3-030-51202-6

[4] L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," *J. Acoust. Soc. Am.*, vol. 138, no. 2, pp. 708–730, 2015. DOI: 10.1121/1.4926438

[5] E. G. Williams and J. A. Mann, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999, vol. 108.

[6] N. Antonello, E. D. Sena, M. Moonen, P. A. Naylor, and T. van Waterschoot, "Room impulse response interpolation using a sparse spatio-temporal representation of the sound field," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 10, pp. 1929–1941, 2017. DOI: 10.1109/TASLP.2017.2730284

[7] V. Pulkki, S. Delikaris-Manias, and A. Politis, *Parametric Time-Frequency Domain Spatial Audio*. 2017. DOI: 10.1002/9781119252634

[8] D. Caviedes-Nozal and E. Fernandez-Grande, "Spatio-temporal bayesian regression for room impulse response reconstruction with spherical waves," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 31, pp. 3263–3277, 2023. DOI: 10.1109/TASLP.2023.3306708

[9] D. Sundström, F. Elvander, and A. Jakobsson, "Boundary-informed sound field reconstruction," in *Proc. European Signal Process. Conf. (EUSIPCO)*, 2025, pp. 81–85. DOI: 10.23919/EUSIPCO63237.2025.11226330

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

[11] 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. DOI: 10.1109/IWAENC.2018.8521334

[12] 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, 2021. DOI: 10.1109/TSP.2021.3070228

[13] J. Ribeiro, S. Koyama, and H. Saruwatari, "Sound field estimation based on physics-constrained kernel interpolation adapted to environment," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 32, pp. 4369–4383, 2024. DOI: 10.1109/TASLP.2024.3467951

[14] Z. Liang, W. Zhang, and T. D. Abhayapala, "Sound field reconstruction using neural processes with dynamic kernels," *EURASIP J. Audio Speech Music Process.*, 2024. DOI: 10.1186/s13636-024-00333-x

[15] J. Brunnström, M. B. Møller, J. Østergaard, S. Koyama, T. van Waterschoot, and M. Moonen, "Time-domain sound field estimation using kernel ridge regression," *IEEE Trans. Audio Speech Lang. Process.*, 2026. DOI: 10.1109/TASLPRO.2026.3662464

[16] 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. DOI: 10.1109/LSP.2017.2775242

[17] D. Caviedes-Nozal, N. A. B. Riis, F. M. Heuchel, J. Brunskog, P. Gerstoft, and E. Fernandez-Grande, "Gaussian processes for sound field reconstruction," *J. Acoust. Soc. Am.*, vol. 149, no. 2, pp. 1107–1119, 2021. DOI: 10.1121/10.0003497

[18] D. Sundström, S. Koyama, and A. Jakobsson, "Sound field estimation using deep kernel learning regularized by the wave equation," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2024, pp. 319–323. DOI: 10.1109/IWAENC61483.2024.10694575

[19] J. Brunnström, M. B. Møller, J. Østergaard, S. Koyama, T. van Waterschoot, and M. Moonen, *Sound field estimation with moving microphones using kernel ridge regression*, arXiv:2509.16358, 2025. DOI: 10.48550/arXiv.2509.16358

[20] S. Verburg, Y. Pene, and E. Fernandez-Grande, "Dynamic room impulse response measurements with a robotic arm," in *Proc. Forum Acusticum*, 2025. DOI: 10.61782/fa.2025.0316

[21] F. Katzberg, M. Maass, and A. Mertins, "Spherical harmonic representation for dynamic sound-field measurements," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2021, pp. 426–430. DOI: 10.1109/ICASSP39728.2021.9413708

[22] J. Brunnström, M. B. Møller, and M. Moonen, "Bayesian sound field estimation using moving microphones," *IEEE Open J. Signal Process.*, vol. 6, pp. 312–322, 2025. DOI: 10.1109/OJSP.2025.3526546

[23] J. Brunnström, M. B. Møller, T. van Waterschoot, M. Moonen, and J. Østergaard, "Experimental validation of sound field estimation methods using moving microphones," in *Proc. Forum Acusticum*, 2025. DOI: 10.61782/fa.2025.0379

[24] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering* (Applied Mathematical Sciences), 4th. 2019.

[25] R. A. Kennedy and P. Sadeghi, *Hilbert Space Methods in Signal Processing*. Cambridge University Press, 2013.

[26] C. Antweiler, A. Telle, and P. Vary, "NLMS-type system identification of MISO systems with shifted perfect sequences," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2008.

[27] N. Hahn and S. Spors, "Simultaneous measurement of spatial room impulse responses from multiple sound sources using a continuously moving microphone," in *Proc. European Signal Process. Conf. (EUSIPCO)*, 2018, pp. 2180–2184. DOI: 10.23919/EUSIPCO.2018.8553532

[28] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979. DOI: 10.1121/1.382599

[29] 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. DOI: 10.1109/ICASSP.2018.8461310

[30] J. Cadavid, M. B. Møller, C. S. Pedersen, S. Bech, T. van Waterschoot, and J. Østergaard, "Performance of low-frequency sound zones with very fast room impulse response measurements," *J. Acoust. Soc. Am.*, vol. 155, no. 1, pp. 757–768, 2024. DOI: 10.1121/10.0024519

[31] R. S. Womersley, "Efficient spherical designs with good geometric properties," in *Contemporary Computational Mathematics - A Celebration of the 80th Birthday of Ian Sloan*, J. Dick, F. Y. Kuo, and H. Woźniakowski, Eds., Springer International Publishing, 2018, pp. 1243–1285. DOI: 10.1007/978-3-319-72456-0_57

[32] J. Bradbury et al., *JAX: Composable transformations of Python+NumPy programs*, [Online]. Available: github.com/jax-ml/jax, 2018.