# Boundary-Informed Sound Field Reconstruction

David Sundström[†], Filip Elvander[*], and Andreas Jakobsson[†]
[*]Dept. of Information and Communications Engineering, Aalto University, Finland
[†]Dept. of Mathematical Sciences, Lund University, Sweden

***Abstract*—We consider the problem of reconstructing the sound field in a room using prior information of the boundary geometry, represented as a point cloud. In general, when no boundary information is available, an accurate sound field reconstruction over a large spatial region and at high frequencies requires numerous microphone measurements. On the other hand, if all geometrical and acoustical aspects of the boundaries are known, the sound field could, in theory, be simulated without any measurements. In this work, we address the intermediate case, where only partial or uncertain boundary information is available. This setting is similar to one studied in virtual reality applications, where the goal is to create a perceptually convincing audio experience. In this work, we focus on spatial sound control applications, which in contrast require an accurate sound field reconstruction. Therefore, we formulate the problem within a linear Bayesian framework, incorporating a boundary-informed prior derived from impedance boundary conditions. The formulation allows for joint optimization of the unknown hyperparameters, including the noise and signal variances and the impedance boundary conditions. Using numerical experiments, we show that incorporating the boundary-informed prior significantly enhances the reconstruction, notably even when only a few hundreds of boundary points are available or when the boundary positions are calibrated with an uncertainty up to 1 dm.**

***Index Terms*—Sound field reconstruction, spatial audio modelling, Bayesian estimation**

## I. Introduction

Applications involving sound field control in a defined region, such as spatial active noise control [1] and sound zone generation [2], depend on accurately reconstructing the sound field from nearby microphone measurements. In general, this is a challenging problem, in particular when a large region and high frequencies are of interest. In such settings, the number of parameters required to accurately represent the sound field is typically orders of magnitude greater than the number of available microphone measurements, necessitating the use of some form of regularization to allow for a unique solution.

Given the importance of the problem, various ways to introduce regularizing prior information has been studied in the recent literature [3]. For example, when using a linear setting with plane wave basis functions [4], free-field Greens functions [5], and/or spatial Fourier basis functions [6], the most common approach, by far, is using Tikhonov regularization [7], corresponding to an assumption of i.i.d. normal priors on the coefficients. While these assumptions are well suited for a diffuse sound field [8], sparse priors has also been used to promote a directional sound field [4], [9]. Additionally, for nonlinear models or linear models with basis functions that by construction do not satisfy the Helmholtz or wave equation, regularizers have been introduced to enforce pointwise adherence to the differential equation [10], [11]. Regardless of the type of regularization, the information provided by the microphone measurements is limited. In contrast, in an idealized scenario where the source position, source signal, speed of sound, boundary geometry, and acoustic properties are all known, the sound field can be computed as a forward problem, which is, for instance, exploited in simulation-based reconstruction methods [12].

Such methods have been extended to both time- and frequency domain formulations for the settings where the boundary conditions are either unknown [13], [14], partially unknown [15], the source positions are unknown [16], or for when both the source position and boundary conditions are unknown [17]. Regrettably, such simulation-based formulations require complete and precise calibration of the boundary geometry for the forward-problem to be meaningful. Even if possible to measure, it would typically be a both expensive and time-consuming process to obtain this information.

In contrast, partial information of the boundary-surfaces can easily be obtained using consumer-grade devices, such as a smartphone camera capturing just a few photos of a room [18]. Although the reconstructed point cloud may be coarse, uncertain, or incomplete, it can still provide valuable information that may be exploited for sound field reconstruction. The use of such partial boundary data has recently been explored in virtual reality applications [19]–[22], demonstrating successful reconstruction of perceptually convincing binaural signals when evaluated using metrics such as reverberation time, speech intelligibility, and early decay time. However, in sound field control applications, as studied here, precise reconstruction of the sound field is essential.

Aiming to improve the accuracy of the sound field reconstruction, this work proposes to regularize the sound field reconstruction problem using a three-dimensional (3D) point cloud that partially represents the boundaries of the room. To do so, we formulate the problem in a linear Bayesian framework, deriving a prior distribution for the linear coefficients from impedance boundary conditions [23]. Unknown hyperparameters, such as the signal and noise variances and the boundary-related parameters, are jointly optimized. When the hyperparameters are known, the method has similar computational complexity as other Gaussian process-based estimators [3], and can thus be used in downstream applications related to sound field control (see for example [1], [24]).

## II. Problem formulation

Consider a sound field $u : \mathbb{R}^3 \to \mathbb{C}$, for a given frequency $\omega$, that is measured using $M$ microphones, such that

$$\mathbf{y} = \mathbf{u} + \varepsilon, \tag{1}$$

where $\mathbf{u} = [u(\mathbf{r}_1), \dots, u(\mathbf{r}_M)]$ denotes the sound field at the microphone positions $\mathbf{r}_m \in \Omega \subseteq \mathbb{R}^3$, for $m = 1, \dots, M$, and $\varepsilon = [\varepsilon_1, \dots, \varepsilon_M]$ is the measurement noise, here assumed to be well modelled as an independent circularly symmetric zero-mean Gaussian noise, i.e.,

$$\varepsilon_i \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right), \tag{2}$$

where $\mathbf{I}$ denotes the identity matrix of appropriate dimension. The region $\Omega$ is assumed to be source-free, such that the sound field satisfies the homogeneous Helmholtz equation

$$\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}) = 0 \quad \forall \quad \mathbf{r} \in \Omega, \tag{3}$$

where $k = \omega/c$ denotes the wave number and $c$ the speed of sound. The sound field is here represented using a superposition of plane wave functions, also known as the Herglotz wave functions [25], defined as

$$u(\mathbf{r}) = \int_{\boldsymbol{\eta} \in \mathcal{S}_2} \tilde{u}(\boldsymbol{\eta}) e^{ik\boldsymbol{\eta}^T \mathbf{r}} d\boldsymbol{\eta}, \tag{4}$$

where $\tilde{u} : \mathcal{S}_2 \to \mathbb{C}$ is the complex source distribution on the unit sphere $\mathcal{S}_2$ in $\mathbb{R}^3$. The representation in (4) satisfies (3) and any solution to (3) can be arbitrarily well approximated by the Herglotz wave functions [26].

Furthermore, we assume a subset of the boundary of the region, denoted $\mathcal{B} \subseteq \delta\Omega$, to consist of a locally reacting surface with acoustical properties independent of the angle of incidence. Under these assumption, the acoustical properties of the boundaries can be characterized in terms of the specific impedance $\beta : \mathcal{B} \to \mathbb{C}$ using the impedance boundary conditions [23]

$$\beta(\mathbf{r})\mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) + iku(\mathbf{r}) = 0 \quad \forall \quad \mathbf{r} \in \mathcal{B}, \tag{5}$$

where $\mathbf{n} : \mathbb{R}^3 \to \mathcal{S}_2$ is the outward normal to the surface. In this work, we consider the problem of reconstructing the sound field in the region $\Omega$, given the measurements in (1), the microphone positions $\{\mathbf{r}_m\}_{m=1,\dots,M}$, the wave number $k$, and the set of boundary points $\mathcal{B}$ with corresponding normals.

## III. Proposed method

In order to introduce the partial boundary information in the sound field reconstruction problem, we here employ a Bayesian problem formulation [3], [27]. Let

$$\mathbf{y} = \boldsymbol{\Phi}\boldsymbol{\alpha} + \varepsilon, \tag{6}$$

be the discretized Herglotz representation of the measurement model in (1), where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_P]^T \in \mathbb{C}^P$ denotes the unknown vector of coefficients and $[\boldsymbol{\Phi}]_{m,p} = \phi_p(\mathbf{r}_m)$, where $\phi_p(\mathbf{r}) = e^{i\mathbf{k}_p \mathbf{r}}$ is a plane wave with $\mathbf{k}_p = k\boldsymbol{\eta}_p$, with $\boldsymbol{\eta}_p \in \mathcal{S}_2$ denoting the direction of arrival. The likelihood of observing $\mathbf{y}$, given the parameters, is given by $\mathbf{y}|\sigma^2, \boldsymbol{\alpha} \sim \mathcal{CN}\left(\boldsymbol{\Phi}\alpha, \sigma^2\mathbf{I}\right)$. A naive approach to estimate the sound field would be to maximize the likelihood function of $\boldsymbol{\alpha}$, which corresponds to solving the problem

$$\underset{\boldsymbol{\alpha} \in \mathbb{C}}{\text{minimize}} \quad \frac{1}{2\sigma^2} ||\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\alpha}||_2^2. \tag{7}$$

However, this problem is typically ill-posed since $M \ll P$ in general. Instead, prior information about the sound field can be included by maximizing the *a posteriori* distribution, given by $p(\boldsymbol{\alpha}|\mathbf{y}, \sigma^2) \propto p(\mathbf{y}|\boldsymbol{\alpha}, \sigma^2)p(\boldsymbol{\alpha})$. When the prior distribtuion is a zero-mean circularly symmetric complex normal distribution, i.e.,

$$\boldsymbol{\alpha} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_\alpha), \tag{8}$$

the maximum *a posteriori* estimate is given as the solution to

$$\boldsymbol{\alpha}_{\text{MAP}} = \underset{\boldsymbol{\alpha} \in \mathbb{C}^P}{\arg\min} \quad \frac{1}{2\sigma^2} ||\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\alpha}||_2^2 + \boldsymbol{\alpha}^H \boldsymbol{\Sigma}_\alpha^{-1} \boldsymbol{\alpha}, \tag{9}$$

which has the closed form expression

$$\boldsymbol{\alpha}_{\text{MAP}} = \frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \boldsymbol{\Phi}^H \boldsymbol{\Phi} + \boldsymbol{\Sigma}_\alpha^{-1} \right)^{-1} \boldsymbol{\Phi}^H \mathbf{y}. \tag{10}$$

Using (10), the predictive distribution of the sound field at position $\mathbf{r}$ is given by $\hat{u}(\mathbf{r}) \sim \mathcal{CN}\left(\tilde{u}(\mathbf{r}), \tau^2(\mathbf{r})\right)$, where [27]

$$\tilde{u}(\mathbf{r}) = \boldsymbol{\phi}^H(\mathbf{r}) \boldsymbol{\Sigma}_\alpha \boldsymbol{\Phi} \mathbf{Q}^{-1} \mathbf{y}, \tag{11}$$

$$\tau^2(\mathbf{r}) = \boldsymbol{\phi}^H(\mathbf{r}) \boldsymbol{\Sigma}_\alpha \boldsymbol{\phi}(\mathbf{r}) - \boldsymbol{\phi}^H(\mathbf{r}) \boldsymbol{\Sigma}_\alpha \boldsymbol{\Phi} \mathbf{Q}^{-1} \boldsymbol{\Phi}^H \boldsymbol{\Sigma}_\alpha \boldsymbol{\phi}(\mathbf{r}), \tag{12}$$

where $\boldsymbol{\phi}(\mathbf{r}) = [\phi_1(\mathbf{r}), \dots, \phi_P(\mathbf{r})]^T$ and

$$\mathbf{Q} = \left(\sigma^2 \mathbf{I} + \boldsymbol{\Phi} \boldsymbol{\Sigma}_\alpha \boldsymbol{\Phi}^H\right). \tag{13}$$

From (11) and (12), one may note that the reconstructed sound field depends on the choice of $\boldsymbol{\Sigma}_\alpha$, with the well-known Tikhonov estimator corresponding to the choice $\boldsymbol{\Sigma}_\alpha = \sigma_\alpha^2 \mathbf{I}$. As an alternative, sound fields that consists of a sparse directivity pattern has also been modelled using a covariance structure on the form $\boldsymbol{\Sigma}_\alpha = \text{diag}(\left[\sigma_{\alpha,1}^2, \sigma_{\alpha,2}^2, \dots, \sigma_{\alpha,P}^2\right])$ [28].

### A. Introducing a boundary-informed prior distribution

Although the reconstruction in (11) inherently satisfies (3) due to the plane wave model, we further incorporate the remaining boundary information from (5) by formulating a prior distribution for $\boldsymbol{\alpha}$. In this preliminary work, for clarity of presentation, we assume all boundary points to share the same specific impedance, i.e., $\beta(\mathbf{r}) \equiv \beta$, $\forall \mathbf{r} \in \mathcal{B}$. By differentiating the plane-wave model along the normal, (5) can be reformulated as

$$\left(\beta \boldsymbol{\Psi} + \tilde{\boldsymbol{\Phi}}\right) \boldsymbol{\alpha} = 0 \tag{14}$$

where $[\boldsymbol{\Psi}]_{b,p} = i\mathbf{k}_p^T \mathbf{n}(\mathbf{r}_b) e^{i\mathbf{k}_p^T \mathbf{r}_b}$ and $[\tilde{\boldsymbol{\Phi}}]_{b,p} = ike^{i\mathbf{k}_p^T \mathbf{r}_b}$, for $p = 1, \dots, P$ and $b = 1, \dots, B$, where $B$ is the cardinality of the set $\mathcal{B}$. In principle, the boundary constraint (14) could be included as a linear constraint in (7). However, in practice the resulting solution will only approximately satisfy the constraint in (14), due to, for example, uncertainties in the 3D point cloud detailing the boundary. Therefore, we instead let

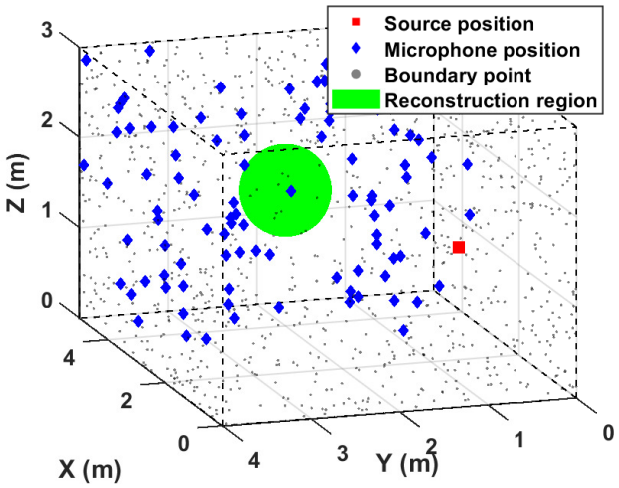


Fig. 1: Illustration of the experimental setup.

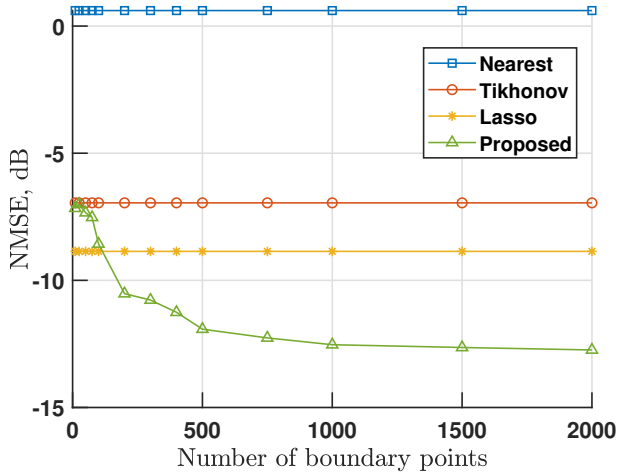


Fig. 2: Illustration of the NMSE as a function of the number of boundary points used to form the regularizer.

$\left(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}}\right)\boldsymbol{\alpha}\sim\mathcal{CN}(\mathbf{0},I\sigma_\alpha^2)$, which corresponds to assigning the prior

$$\mathcal{CN}\left(\mathbf{0},\sigma_\alpha^2\left(\left(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}}\right)^H\left(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}}\right)\right)^{-1}\right) \tag{15}$$

to $\boldsymbol{\alpha}$. However, when only a few boundary points are used, the matrix $(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}})^H(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}})$ will be rank deficient. Therefore, we propose to assign the prior distribution in (8) with the covariance defined as

$$\boldsymbol{\Sigma}_\alpha=\sigma_\alpha^2\left(\mathbf{I}+\mu\left(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}}\right)^H\left(\beta\boldsymbol{\Psi}+\tilde{\boldsymbol{\Phi}}\right)\right)^{-1}. \tag{16}$$

Note that although we in this work use a plane-wave model to represent the field, similar priors could be formulated for other basis functions using a similar approach. Also, note that the maximum a posteriori estimate in (10) formed using this prior coincides with the Tikhonov estimator when $\mu=0$.

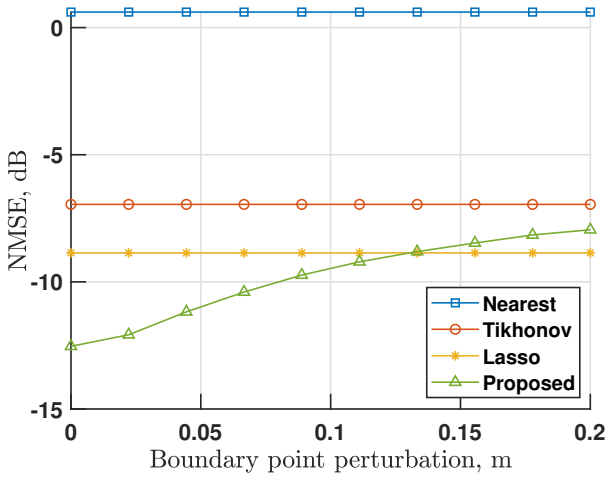


Fig. 3: Illustration of the robustness with respect to constant perturbations in the assumed boundary point positions.

### *B. Joint estimation of hyperparameters and impedance boundary conditions*

In this section, we consider a Bayesian approach for estimating the unknown hyperparameters [29]. The hyperparameters are estimated by maximizing their posterior distributions, which corresponds to maximizing the marginal likelihood (see for example [27] for a detailed derivation)

$$\underset{\beta\in\mathbb{C},\sigma^2,\sigma_\alpha^2,\mu\in\mathbb{R}_+}{\arg\min}\ \frac{1}{2}\mathbf{y}^H\mathbf{Q}^{-1}\mathbf{y}+\frac{1}{2}\log\left(|\mathbf{Q}|\right), \tag{17}$$

where $|\mathbf{Q}|$ denotes the determinant of $\mathbf{Q}$. To account for the non-negativity constraints and introduce the fact that the parameters can span several orders of magnitude, the variables are reparametrized as $\sigma^2=e^a$, $\sigma_\alpha^2=e^b$, $\mu=e^d$, and $\beta=e^\eta$, resulting in the unconstrained problem

$$J(\theta)=\underset{\eta\in\mathbb{C},a,b,d\in\mathbb{R}}{\arg\min}\ \frac{1}{2}\mathbf{y}^H\mathbf{Q}^{-1}\mathbf{y}+\frac{1}{2}\log\left(|\mathbf{Q}|\right), \tag{18}$$

where $\theta=\{a,b,d,\eta\}$. The non-convex problem in (18) may be solved using a conjugate-gradient based solver with Polak-Ribière search directions [30], as implemented in [31]. The gradients with respect to a parameter $\theta_i$ in the covariance matrix $\mathbf{Q}$ can efficiently be computed by straight-forward extension of the result in [27] for the complex-valued setting, resulting in

$$\frac{\partial}{\partial\theta_i}J(\theta)=-\frac{1}{2}\mathrm{trace}\left(\left(\boldsymbol{\xi}\boldsymbol{\xi}^H-\mathbf{Q}^{-1}\right)\frac{\partial\mathbf{Q}}{\partial\theta_i}\right), \tag{19}$$

where $\boldsymbol{\xi}=\mathbf{Q}^{-1}\mathbf{y}$. The partial derivatives of $\mathbf{Q}$ with respect to each hyperparameter are stated in the appendix.

## IV. Numerical Experiments

In the following, we want to understand the limits for when the use of a boundary-informed prior is useful and, as a result, give suggestions on the calibration accuracy that is required when collecting a real dataset for this purpose.

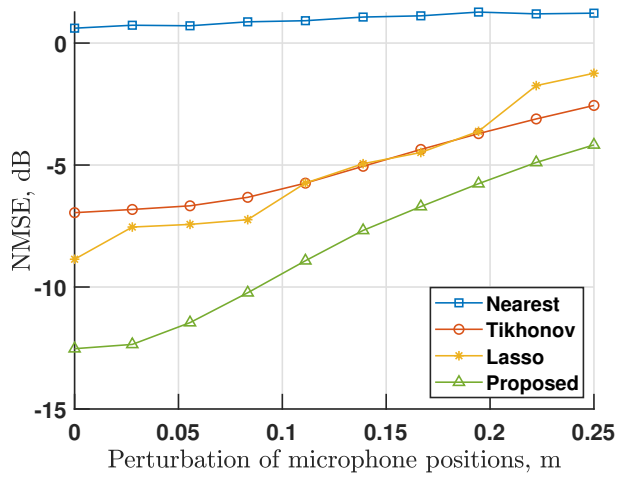


Fig. 4: Illustration of the robustness with respect to constant perturbations in the assumed microphone positions.

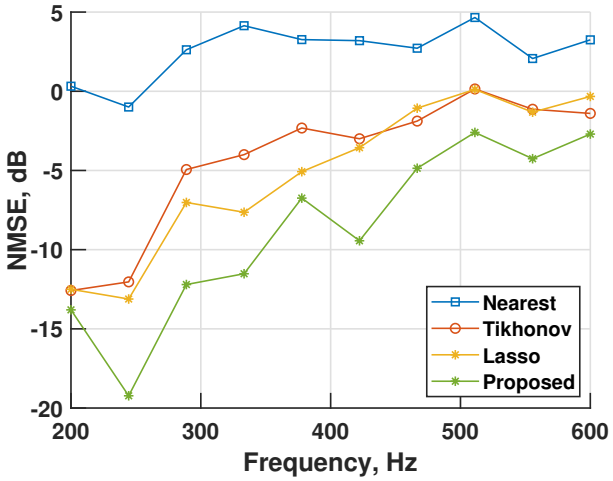


Fig. 5: Illustration of the NMSE of the reconstruction for various frequencies.

Therefore, sound fields are simulated in a controlled setting using the image source method (ISM) as implemented in [32] with the speed of sound 343m/s and the sampling frequency 8kHz. Since the ISM simulation only is guaranteed to satisfy the boundary conditions in (5) exactly as $\beta \to \infty$, that is, in the idealistic setting of lossless reflections [23], the reflection coefficients are set to 0.95 as an approximation. Note that this results in a slight mismatch between the model in Section III and the simulation, making the evaluation setting less favorable for the proposed method. The experimental setup is illustrated in Figure 1, including 100 microphones that are uniformly distributed in half of a shoebox room except in a spherical region of radius 0.5 m, where instead 20 validation positions are sampled uniformly random. Circularly symmetric Gaussian noise is added to the signals to obtain a signal-to-noise ratio of 20dB for a frequency of 300Hz. The reconstruction is measured in terms of the normalized mean squared error, defined as

$$\text{NMSE} = \frac{1}{JN} \sum_{j=1}^{J} \sum_{n=1}^{N} \frac{|u_{j,n} - \hat{u}_{j,n}|^2}{|u_{j,n}|^2}, \tag{20}$$

where $\hat{u}_{j,n}$ and $u_{j,n}$ denote the predicted and true signal for the $j$th validation point and $n$th simulation, for $N = 10$ Monte-Carlo simulations. To give the results more meaning, we include three benchmark methods. First, the naive reconstruction is formed using the microphone signal located nearest to the reconstruction position, denoted Nearest. Secondly, the Tikhonov estimator is constructed by setting $\mu = 0$ in (16), which, compared to the proposed method provides an interpretation of the impact of the boundary information in the prior covariance in (16). Finally, also the Lasso estimate is obtained by adding the regularizing term $\lambda||\boldsymbol{\alpha}||_1$ to (7). For all methods, 1000 plane waves distributed on a Fibonacci lattice are used [33]. Firstly, we study how the number of boundary points affects the reconstruction. The boundary points are uniformly sampled on the surface of the shoebox room with outward normals, as illustrated in Figure 1, with the results shown in Figure 2. Notable is that the reconstruction error decreases rapidly when including the first few hundreds of points and flats out when about 1000 boundary points are used, which therefore is used in the following experiments. Of practical interest is also studying the robustness with respect to the assumption of the boundary point and microphone positions. In order to do so, a constant perturbation of random direction is introduced to each boundary point and microphone position in Figures 3 and 4. In Figure 3, it is worth noting that Proposed obtains a lower reconstruction error as compared the other methods for boundary position errors up to as high as 1 dm. This result open up the possibility of being able to use a point cloud representation of the boundary obtained from everyday sensors such as, for example, the camera sensor in a smartphone, a case which will be examined in more detail in further work. Figure 4 illustrates that Proposed achieves a lower NMSE than the other methods even in the presence of large errors in the assumption of microphone position. This is not obvious, since Proposed includes an assumption of the microphone position in both the observation model and the prior, while the other models only include the assumption of the microphone positions in the observation model. However, the result indicate that the error due to the error in the observation model dominates the error of the reconstructed sound field in this case. In general, the results indicate that the boundary information is of significance value, even when only partial or uncertain information of the boundary is available, which is also confirmed for a wide frequency range in Figure 5.

## V. Conclusion

We introduce a boundary-informed prior distribution for the reconstruction of a sound field from microphone measurements. Compared to simulation-based approaches, which require the full geometry, we show that boundary information

is useful also in settings where the point cloud representation is coarse or uncertain. Simulated numerical experiments in reverberant rooms both illustrate the great potential benefit of this information, and serves as a guideline for the required accuracy when collecting data for a real experiments.

## APPENDIX

### GRADIENTS OF $\mathbf{Q}$ WITH RESPECT TO HYPERPARAMETERS

The gradient of $\mathbf{Q}$ in (13) with respect to $a$ and $b$ are given by $\frac{\partial}{\partial a}\mathbf{Q} = \sigma^2 I$ and $\frac{\partial}{\partial b}\mathbf{Q} = \Phi\mathbf{\Sigma}_\alpha\Phi^H$, respectively. Furthermore, the gradients with respect to $d$ is given by $\frac{\partial}{\partial d}\mathbf{Q} = -\sigma_\alpha^2\mu\mathbf{B}^H\mathbf{B}$, where

$$\mathbf{B} = \left(\beta\mathbf{\Psi} + \tilde{\mathbf{\Phi}}\right)\left(I + \mu\left(\beta\mathbf{\Psi} + \tilde{\mathbf{\Phi}}\right)^H\left(\beta\mathbf{\Psi} + \tilde{\mathbf{\Phi}}\right)\right)^{-1}\mathbf{\Phi}^H.$$

To compute the gradient with respect to the complex-valued variable $\eta$, we note that $\mathbf{Q}$ is not a holomorphic function with respect to $\eta$ due to the conjugation of $\eta$ in (16). Therefore, the derivative does not exist in the classical sense, but may, using Wirtinger calculus, be expressed as the derivative with respect to $\eta^*$ (see, e.g., [34]), yielding

$$2\frac{\partial}{\partial\eta^*}\mathbf{Q} = -2\sigma_\alpha^2\mu\mathbf{C}^H\left(|\beta|^2\mathbf{\Psi}^H\mathbf{\Psi} + \beta^*\mathbf{\Psi}^H\tilde{\mathbf{\Phi}}\right)\mathbf{C} \qquad (21)$$

where $\mathbf{C} = \left(I + \mu(\beta\mathbf{\Psi} + \tilde{\mathbf{\Phi}})^H(\beta\mathbf{\Psi} + \tilde{\mathbf{\Phi}})\right)^{-1}\mathbf{\Phi}^H$.

## REFERENCES


[1] S. Koyama, J. Brunnström, H. Ito, N. Ueno, and H. Saruwatari, "Spatial active noise control based on kernel interpolation of sound field," *IEEE Trans. Audio Speech and Lang. Process.*, vol. 29, pp. 3052–3063, 2021.

[2] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, 2010.

[3] 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. Amer.*, vol. 149, no. 2, pp. 1107–1119, 2021.

[4] R. Horiuchi, S. Koyama, J.G.C. Ribeiro, N. Ueno, and H. Saruwatari, "Kernel learning for sound field estimation with l1 and l2 regularizations," in *IEEE Int. Workshop Appl. Signal Process. Audio Acoust.*, New Paltz, NY, USA, 2021, pp. 261–265.

[5] N. Antonello, E. De 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 Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 10, pp. 1929–1941, 2017.

[6] E.G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic press, 1999.

[7] A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-posed problems*, V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977.

[8] 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, 2017.

[9] S. A. Verburg and E. Fernandez-Grande, "Reconstruction of the sound field in a room using compressive sensing," *J. Acoust. Soc. Amer.*, vol. 143, no. 6, pp. 3770–3779, 2018.

[10] S. Damiano, F. Miotello, M. Pezzoli, A. Bernardini, F. Antonacci, A. Sarti, and T. v. Waterschoot, "A zero-shot physics-informed dictionary learning approach for sound field reconstruction," in *IEEE Int. Conf. Acoust., Speech, Sig. Process.*, Hyderabad, India, 4 2025.

[11] D. Sundström, S. Koyama, and A. Jakobsson, "Sound field estimation using deep kernel learning regularized by the wave equation," in *Int. Workshop Acoust. Signal Enhancement*, Aalborg, Denmark, 09 2024, IEEE, pp. 319–323.

[12] W.A. Veronesi and J.D. Maynard, "Digital holographic reconstruction of sources with arbitrarily shaped surfaces," *J. Acoust. Soc. Amer.*, vol. 85, no. 2, pp. 588–598, 1989.

[13] G. P. Nava, Y. Yasuda, Y. Sato, and S. Sakamoto, "On the in situ estimation of surface acoustic impedance in interiors of arbitrary shape by acoustical inverse methods," *Acoust. Sci. Tech.*, vol. 30, no. 2, pp. 100–109, 2009.

[14] N. Antonello, T. van Waterschoot, M. Moonen, and P. A. Naylor, "Identification of surface acoustic impedances in a reverberant room using the fdtd method," in *Int. Workshop Acoust. Signal Enhancement*, Juan-les-Pins, France, 09 2014, pp. 114–118.

[15] J. M. Schmid, M. Eser, and S. Marburg, "Bayesian approach for the in situ estimation of the acoustic boundary admittance," *J. of Theoretical and Computational Acoust.*, vol. 31, no. 04, 2023.

[16] B. Ahn, K. Yang, B. Hamilton, J. Sheaffer, A. Ranjan, M. Sarabia, O. Tuzel, and J.H.R. Chang, "Novel-view acoustic synthesis from 3d reconstructed rooms," *arXiv preprint arXiv:2310.15130*, 2023.

[17] N. Bertin, S. Kitić, and R. Gribonval, "Joint estimation of sound source location and boundary impedance with physics-driven cosparse regularization," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Shanghai, China, 03 2016, pp. 6340–6344.

[18] V. Leroy, Y. Cabon, and J. Revaud, "Grounding image matching in 3d with mast3r," in *European Conf. on Computer Vision*, Milan, Italy, 09 2024, Springer, pp. 71–91.

[19] M. Chen, K. Su, and E. Shlizerman, "Be everywhere-hear everything (bee): Audio scene reconstruction by sparse audio-visual samples," in *IEEE International Conf. on Computer Vision*, Paris, France, 10 2023, pp. 7853–7862.

[20] S. Liang, C. Huang, Y. Tian, A. Kumar, and C. Xu, "Av-nerf: Learning neural fields for real-world audio-visual scene synthesis," *Adv. in Neural Inf. Process. Syst.*, vol. 36, 12 2023.

[21] S. Liang, C. Huang, Y. Tian, A. Kumar, and C. Xu, "Neural acoustic context field: Rendering realistic room impulse response with neural fields," *arXiv preprint arXiv:2309.15977*, 2023.

[22] S. Majumder, C. Chen, Z. Al-Halah, and K. Grauman, "Few-shot audio-visual learning of environment acoustics," *Adv. in Neural Inf. Process. Syst.*, vol. 35, pp. 2522–2536, 12 2022.

[23] H. Kuttruff, *Room acoustics*, Crc Press, 2016.

[24] J. Brunnström, S. Koyama, and M. Moonen, "Variable span trade-off filter for sound zone control with kernel interpolation weighting," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Singapore, 05 2022, pp. 1071–1075.

[25] D. L. Colton, R. Kress, and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, vol. 93, Springer, 1998.

[26] 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.

[27] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, 11 2005.

[28] K. L. Gemba, S. Nannuru, P. Gerstoft, and W. S. Hodgkiss, "Multi-frequency sparse bayesian learning for robust matched field processing," *J. Acoust. Soc. Amer.*, vol. 141, no. 5, pp. 3411–3420, 2017.

[29] J Antoni, "A bayesian approach to sound source reconstruction: Optimal basis, regularization, and focusing," *J. Acoust. Soc. Amer.*, vol. 131, no. 4, pp. 2873–2890, 2012.

[30] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *Revue française d'informatique et de recherche opérationnelle. Série rouge*, vol. 3, no. 16, pp. 35–43, 1969.

[31] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (gpml) toolbox," *The Jour. of Mach. Learning Research*, vol. 11, pp. 3011–3015, 2010.

[32] E.A.P. Habets, "Room impulse response generator," *Technische Universiteit Eindhoven, Tech. Rep*, vol. 2, no. 2.4, pp. 1, 2006.

[33] Á. González, "Measurement of areas on a sphere using fibonacci and latitude–longitude lattices," *Mathematical geosciences*, vol. 42, pp. 49–64, 2010.

[34] Simon Haykin, "Adaptive filter theory," *Prentice Hall*, vol. 2, 2002.