Towards Gridless Sound Field Reconstruction

Ids van der Werf*†, Pablo Martínez-Nuevo*, Martin Møller*, Richard Hendriks† and Jorge Martinez†
*Bang & Olufsen, †Delft University of Technology

***Abstract*—The sound field in a room can be represented by a weighted sum of room modes. To estimate the room modes, current literature uses on-the-grid, sparse reconstruction methods. However, these on-the-grid methods are known to suffer from basis mismatch. In this work, we investigate the use of a gridless framework for estimating room modes using atomic norm minimization, a gridless method. The advantage of this approach would be that it does not suffer from this basis mismatch problem. We derive a bound for the sound field reconstruction problem in a one-dimensional room with rigid walls and relate this to the frequency separation that is required by the atomic norm. We conclude that for perfect reconstruction based on the investigated gridless approach, additional prior knowledge about the signal model is required. We show how recovery is possible in a one-dimensional setting by exploiting both the structure of the sound field and the acquisition method.**

***Index Terms*—atomic norm, sparse recovery, (spatial) frequency estimation, room acoustics, sound field reconstruction**

I. Introduction

Knowing how the sound pressure varies over space and time has many applications, e.g., room compensation [1], dereverberation [2], and sound zone reconstruction [3]. Reconstructing sound fields inside enclosures comes with extra challenges as the surroundings, such as the geometry of the enclosure and the materials used, influence the sound field. Reconstructing a satisfying sound field in the whole enclosure by extrapolating from few measurements is thus not an obvious task.

For ease of illustration, we focus on rectangular rooms in this paper (see Fig. 1). However, the principles here can be extended to any enclosure. Typically, microphones indicated in Fig. 1 by the red dots, are used to measure the sound pressure. However, the microphones cannot be placed arbitrarily across the room, but are typically placed upon the physical objects inside the room. This means that in practice only a small number of microphones can be used. From the sound pressure measured at the microphone locations, the sound pressure in the whole room must be estimated, as shown in the lower half of Fig. 1.

Several solutions for sound field reconstruction have been proposed in the past, e.g. [4]–[7]. Assuming that the room modes can be expanded into plane waves, current literature estimates the corresponding modal frequencies [5]–[7]. Knowledge about the shape of the room modes and the modal frequencies of a room is very useful, as it allows to calculate the sound field resulting from any source receiver pair.

In the low frequency range, the sound field can be represented by a small number of room modes. Therefore, estimating room modes is done by the use of compressive sensing techniques and convex optimization (e.g. [6], [7]), due to its versatility to include prior knowledge about the setting. These approaches use methods based on the lasso problem [8], which use a sparsity promoting ℓ_1-norm. Such methods are often referred to as 'on-the-grid' methods, as they use a grid to form a basis. However, they suffer from what is called basis mismatch [9], because the assumed basis never exactly matches the actual basis of the signal. As a result on-the-grid methods will always make an approximation of the room modes, and will never produce an exact reconstruction.

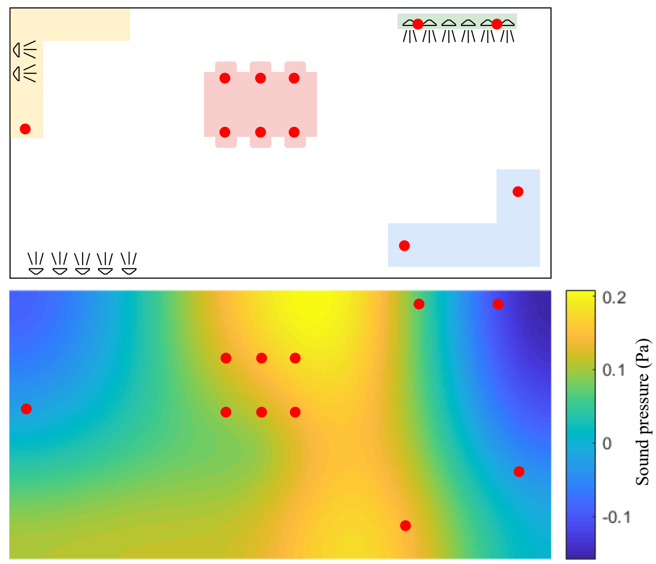

Fig. 1. Living room and corresponding sound field at 150 Hz.

Last decade a 'gridless' solution to this basis mismatch has been introduced, known as the atomic norm [10], [11]. In this work we investigate whether a gridless framework using the atomic norm can be used as a replacement for the on-the-grid methods for the estimation of modal frequencies. We start by studying a model of a room that has been simplified significantly, in order to reduce the complexity of the problem. However, its analysis gives insights into the challenges ahead, even for more complicated scenarios.

The remainder of this paper is organized as follows. First, in Section II, the signal model is introduced, as well as the on-the-grid and proposed gridless approach. Then in Section III we derive a theoretical bound for the sound field reconstruction problem in a boxed-shaped room and compare this with the frequency separation required by the atomic norm for a successful signal recovery. In Section IV we discuss how to include prior knowledge about the signal model and exploit the acquisition method, finally we validate the derived bound with numerical simulations in Section V.

II. Modal Frequency Estimation

A. Signal Model

For simplicity we will look at a one-dimensional room with rigid walls. We assume that the room has length L_x and is excited by a point source located at $x = x_0$. The Green's function in this setting is defined as

$$G(x, x_0, \omega) = -\frac{1}{L_x} \sum_{n=0}^{\infty} \frac{\psi_n(x_0)}{(\frac{\omega}{c})^2 - k_n^2} \psi_n(x), \tag{1}$$

where c is the speed of sound and $\psi_n(x)$ is the n'th room mode (eigenfunction) with corresponding n'th modal frequency k_n (eigenfrequency). Let us assume the source emits L temporal frequencies indicated by ω_l. The sound field is then defined as

$$p(x, \omega) = -\frac{1}{L_x} \sum_{l=1}^{L} C_l \delta(\omega - \omega_l) \sum_{n=0}^{\infty} \frac{\psi_n(x_0)}{(\frac{\omega}{c})^2 - k_n^2} \psi_n(x), \tag{2}$$

where $C_l \in \mathbb{R}$ is a constant, which allows for excitation frequencies with varying amplitudes. For a room with rigid walls, $\psi_n(x) = \sqrt{2}\cos(k_n x)$ and $k_n = n\frac{\pi}{L_x}$ [12].

In total, we define M_t potential microphone positions inside the room on a uniform grid, i.e., $x_m = \frac{m}{F}$ where $m \in \mathcal{J}_t$ for $\mathcal{J}_t = \{1, 2, \ldots, M_t\}$ and $\frac{1}{F}$ is the distance between successive positions. Note that $x_m \in (0, L_x)$. Similarly, we denote by $\mathcal{J}_o$ the subset of indexes corresponding to the observed measurements, i.e., $\mathcal{J}_o \subseteq \mathcal{J}_t$. We place $M_o \leq M_t$ microphones at positions chosen uniformly at random from the index set $\mathcal{J}_t$, to form the 'observed' set. The situation is illustrated in Fig. 2.

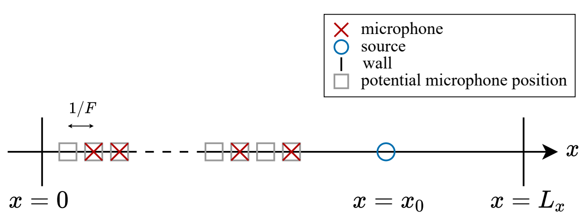

Fig. 2. Measurement setup in a 1D room.

Using the complex exponential expansion for the cosine function, and assuming we can satisfyingly represent the sound field with a finite number of room modes, we can define a system of equations over the set $\mathcal{J}_t$ as

$$\mathbf{P} = \mathbf{AS}, \tag{3}$$

where the potential measurements are stacked in the matrix $\mathbf{P} \in \mathbb{R}^{M_t \times L}$, with columns $\mathbf{p}_l = [p(x_1, \omega_l), \ldots, p(x_{M_t}, \omega_l)]^T$, $\mathbf{A} = [\mathbf{a}(k_1), \ldots, \mathbf{a}(k_N)] \in \mathbb{C}^{M_t \times N}$ is the steering matrix containing the set of modal frequencies, $\mathbf{a}(k_n) = [e^{jk_n x_1}, e^{jk_n x_2}, \ldots]^T$, with $n \in \{-\frac{N-1}{2}, \ldots, +\frac{N-1}{2}\}$, $\mathbf{S} \in \mathbb{R}^{N \times L}$ is the source matrix containing the weights of each modal frequency, with elements $s_{n,l} = -\frac{C_l}{L_x}\frac{\cos(k_n x_0)}{(\frac{\omega_l}{c})^2 - k_n^2}$. In general this system is underdetermined, $N > M_t$. However, the columns of the source matrix are approximately sparse, as only the room modes with modal frequencies close to $\frac{\omega_l}{c}$ get excited significantly. This allows for compressive sensing techniques to solve the problem.

B. Existing On-the-grid Method

In order to determine the modal frequencies and corresponding amplitudes, prior art (e.g. [6], [7]) minimizes an ℓ_1-norm of a weighted (sparse) vector $\mathbf{b}_l$,

$$\begin{aligned} \min_{[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_L] \in \mathbb{C}^{R \times L}} \quad & \textstyle\sum_l^L ||\mathbf{L}_l \mathbf{b}_l||_1 \\ \text{s.t.} \quad & \mathbf{P}_i - \mathbf{D}[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_L] = 0, \forall i \in \mathcal{J}_o \end{aligned} \tag{4}$$

where the basis $\mathbf{D}$ is formed by a dictionary of R spatial frequencies $v_r \in [-\pi F, \pi F]$ on a uniform grid and weighting $\mathbf{L}_l = \text{diag}((\frac{\omega_l}{c})^2 - \mathbf{v}^2)$, $\mathbf{v} = [v_1, \ldots, v_R]^T$.

The basis mismatch problem results from the discrete dictionary $\mathbf{D}$. One assumes that the modal frequencies are inside this dictionary, but in practice this is never exactly the case resulting in a mismatch between the actual basis $\mathbf{A}$ and the assumed basis $\mathbf{D}$. If the size of the dictionary is increased, the mismatch decreases, however this comes with computational costs and higher coherence between the columns of $\mathbf{D}$ [13]. If the dictionary is too small, the modal frequencies might not be in the dictionary, decreasing the accuracy of the reconstruction. If an on-the-grid method is used, one thus has to tackle this trade-off and will inherently make an error due to the fact that a grid is used.

C. Proposed Method

To circumvent the basis mismatch problem we investigate the use of a gridless framework, using the atomic norm. This method makes use of a set of atoms as dictionary,

$$\mathcal{A} = \{\mathbf{a}(f, \boldsymbol{\phi}) = \mathbf{a}(f)\boldsymbol{\phi} \,:\, f \in [0, F), \boldsymbol{\phi} \in \mathbb{C}^{1 \times L}, ||\boldsymbol{\phi}||_2 = 1\}, \tag{5}$$

where $\mathbf{a}(f) = [e^{j2\pi f x_1}, \ldots, e^{j2\pi f x_{M_t}}]^T \in \mathbb{C}^{M_t}$. Note that the frequency f is defined over a continuous interval, therefore the set defines a continuous dictionary. The atomic (ℓ_1) norm is defined as

$$||\mathbf{P}||_{\mathcal{A}} = \inf\left\{\sum_{k \in \mathcal{K}} c_k : \mathbf{P} = \sum_{k \in \mathcal{K}} c_k \mathbf{a}(f_k, \boldsymbol{\phi}_k), \mathbf{a}(f_k, \boldsymbol{\phi}_k) \in \mathcal{A}\right\}, \tag{6}$$

where $c_k > 0$ and $\mathcal{K}$ is the set containing the indices of the atoms. Alternatively to (4), we use the atomic norm to promote a sparse set of modal frequencies:

$$\min_{\tilde{\mathbf{P}} \in \mathbb{C}^{M_t \times L}} ||\tilde{\mathbf{P}}||_{\mathcal{A}}, \text{ s.t. } \tilde{\mathbf{P}}_i = \mathbf{P}_i, \quad \forall i \in \mathcal{J}_o, \tag{7}$$

here $\mathbf{P}_i$ denotes the i'th row of $\mathbf{P}$. The atomic norm can be cast into an SDP [11]. Therefore the optimization problem is reformulated as

$$\begin{aligned} \min_{\mathbf{W}\in\mathbb{C}^{L\times L},\mathbf{u}\in\mathbb{C}^{M_t},\tilde{\mathbf{P}}\in\mathbb{R}^{M_t\times L}} \quad & \mathrm{Tr}(\mathrm{Toep}(\mathbf{u}))+\mathrm{Tr}(\mathbf{W}) \\ \text{s.t.} \quad & \begin{bmatrix}\mathrm{Toep}(\mathbf{u}) & \tilde{\mathbf{P}} \\ \tilde{\mathbf{P}}^H & \mathbf{W}\end{bmatrix} \succcurlyeq 0 \\ & \tilde{\mathbf{P}}_i=\mathbf{P}_i, \quad \forall i\in\mathcal{J}_o. \end{aligned} \tag{8}$$

Using the matrix $\mathrm{Toep}(\mathbf{u})$, the atomic norm tries to find a matrix related to the covariance matrix of $\mathbf{P}$. From the optimal solution $\mathrm{Toep}(\mathbf{u}^*)$, we can therefore retrieve the estimated frequencies using any subspace method. Assume that the $\boldsymbol{\phi}_k$'s are independent random variables with $\mathbb{E}[\boldsymbol{\phi}_k]=\mathbf{0}$. Then, if the modal frequencies k_n adhere to a certain frequency separation Δ_k, which is the minimum pairwise distance,

$$\Delta_k \geq 2\pi F\frac{\alpha}{M_t-1}, \tag{9}$$

for $\alpha>0$, there exist a numerical constant C, such that

$$M_o \geq C\max\Big\{\log^2\frac{M_t\sqrt{L}}{\delta}, N\log\frac{N}{\delta}\log\frac{M_t\sqrt{L}}{\delta}\Big\}. \tag{10}$$

is sufficient to guarantee that, with probability at least $1-\delta$, the atomic norm will exactly recover the original signal [11]. Tang *et al.* [11] proved that successful recovery of the (modal) frequencies is guaranteed if $\alpha=4$ in (9). However in practice its value can be lower [14], and depends on the parameters of the model. Note that our model only approximately meets all the assumptions on $\boldsymbol{\phi}_k$, as the rows of $\mathbf{S}$ are not completely independent. It is therefore of great interest to look at the performance of the atomic norm for our signal model, and to derive corresponding bounds on the value of α.

III. Bounds For Sound Field Reconstruction

A. Frequency Separation of Modal Frequencies

The modal frequencies are given by $k_n=n\frac{\pi}{L_x}$ [12], and thus are separated by $\Delta_k=\frac{\pi}{L_x}$. Using this, we can rewrite (9) and derive that we need at least

$$\begin{aligned} M_t &\geq 2\pi F\frac{\alpha}{\Delta_k}+1 \\ &= 2\cdot F\cdot\alpha\cdot L_x+1. \end{aligned} \tag{11}$$

Additionally, the size of the room puts a limit to the number of possible measurement locations M_t; the measurements must be inside the room, thus $0<x_m<L_x, \forall m\in\mathcal{J}_t$. From this we must have that $\min_m(x_m)=\frac{1}{F}>0$, which is satisfied, and that

$$\max_m(x_m)=\frac{M_t}{F}<L_x, \tag{12}$$

and thus

$$M_t < F\cdot L_x. \tag{13}$$

Now we combine (11) and (13), to write

$$F\cdot\alpha\cdot 2L_x+1 < F\cdot L_x. \tag{14}$$

The inequality in (14) will be satisfied for practical situations ($L_x>0, F>0$) if and only if $0<\alpha<\frac{1}{2}$.

In conclusion, a frequency separation, shown in (9), with at least $\alpha<\frac{1}{2}$ is sufficient, in order to be able to solve the modal frequency estimation problem exactly with high probability. Now, before we look at the performance of the atomic norm in practice, we first derive a lower bound that is inherent to the problem itself.

B. Knowns vs. Unknowns

Due to the fact that only finite number of measurements are available in practice, the estimated modal frequencies cannot be arbitrarily close. We are interested in a lower bound on the frequency separation, regardless of the method one is using to solve the problem. We can then relate this bound to the frequency separation required by the atomic norm, to get an idea of the performance of the atomic norm.

From (3) we know that we have M_tL knowns, the number of (real) elements in the measurement matrix $\mathbf{P}$. On the contrary, the steering matrix $\mathbf{A}$ and the source matrix $\mathbf{S}$ are unknown, resulting in $N(1+L)$ unknowns in total; N unknown frequencies in $\mathbf{A}$ (not M_tN because the structure of $\mathbf{A}$ is assumed to be known), and NL unknowns in $\mathbf{S}$ (of which the structure is not known). If no other prior knowledge is available, the number of unknown variables cannot be lower than the number of known ones. From this it follows that,

$$M_tL \geq N(1+L), \tag{15}$$

and thus,

$$\frac{1}{N} \geq \frac{(1+L)}{M_tL}. \tag{16}$$

Note that if N (spatial) frequencies are to be fit uniformly on a interval of length $2\pi F$, then the maximum frequency separation that can be reached is $\frac{2\pi F}{N}$, thus $\frac{2\pi F}{N}=2\pi F\frac{\alpha}{M_t-1}$. As a result, and using (16),

$$\begin{aligned} \alpha &= (M_t-1)\frac{1}{N} \geq (M_t-1)\frac{(1+L)}{M_tL} \\ &= (1+\frac{1}{L})(1-\frac{1}{M_t}). \end{aligned} \tag{17}$$

From (17) it is clear that α can never be lower than $\frac{1}{2}$ provided that $L>0$ and $N>0$. Thus, we can conclude that the frequency separation required by the atomic norm can not be reached, by construction of the problem. This means that more prior knowledge about the problem has to be included to lower the bound derived in (17).

IV. Exploiting prior knowledge

We have shown that atomic norm minimization without further exploiting the structure of the problem is not enough for recovery. Thus, we need to incorporate more prior knowledge to meet the bound. In particular, we describe here two strategies used in conjunction to realize successful recovery.

A. Mirror Image

In the case of rigid walls, the sound field is perfectly reflected by the walls. Therefore we can create “image microphones” on the other side of a wall, as illustrated in Fig. 3. For this operation, the location of the corresponding wall must be known. In order to avoid a trivial case, we assume the length of the room is not fully known, i.e., we only take as reference one of the walls and use it to duplicate the measurements by mirroring them. By including one reflection, the number of knowns is increased by a factor two, therefore the lower bound on α becomes

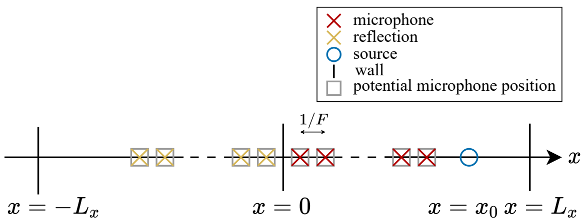

Fig. 3. 1D room with one reflection.

$$\alpha \geq \frac{1}{2}(1+\frac{1}{L})(1-\frac{1}{M_t}). \tag{18}$$

By adding these new measurements, we have managed to reduce the constraint on the number of measurements. However, we still need to further exploit the structure of the problem to meet the inequality in (14).

B. Spectral symmetry

For the rigid wall case, the room modes are cosines. All positive modal frequencies will also occur on the negative side of the spatial frequency spectrum. Hence, we only need to consider the positive side of the spectrum. To reduce the number of unknowns, we would thus like to remove the negative side of the spectrum in our signal. This can be done by using the Hilbert transform, if a continuous measurement is available. Instead of a static grid of microphones, one could use a moving microphone, (the variable x in the signal model in (2) is replaced by vt, where v is the speed of the microphone), to get a continuous measurement in time [15]. We assume that the Doppler effect, due to the moving microphone, can be removed [16].

First, the Hilbert transform is applied to the continuous signal, then the signal is discretized again, by $t_m = \frac{1}{v}x_m$, such that the notation is consistent. By performing the Hilbert transform, the number of unknowns is decreased by a factor two, therefore the lower bound on α becomes

$$\alpha \geq \frac{1}{4}(1+\frac{1}{L})(1-\frac{1}{M_t}). \tag{19}$$

Now the lower bound is smaller than $\frac{1}{2}$. Hence, the only question is whether the atomic norm attains a value lower than $\frac{1}{2}$ in practice. To ensure the framework finds a solution consisting of only positive frequencies, an extra constrained is added to the minimization problem, inspired by [17].

Assuming the negative frequencies are removed, the optimization problem becomes

$$\begin{aligned} \min_{\mathbf{W}\in\mathbb{C}^{L\times L},\mathbf{u}\in\mathbb{C}^{2M_t},\tilde{\mathbf{P}}\in\mathbb{R}^{2M_t\times L}} \quad & \mathrm{Tr}(\mathrm{Toep}(\mathbf{u}))+\mathrm{Tr}(\mathbf{W}) \\ \text{s.t.} \quad & \begin{bmatrix}\mathrm{Toep}(\mathbf{u}) & \tilde{\mathbf{P}} \\ \tilde{\mathbf{P}}^H & \mathbf{W}\end{bmatrix} \succcurlyeq 0 \\ & \tilde{\mathbf{P}}_i = \begin{bmatrix}\mathbf{\Pi}_{M_t}\mathbf{P}^* \\ \mathbf{P}\end{bmatrix}_i, \forall i \in \mathcal{J}_o \\ e^{-ja}\mathbf{F}\mathrm{Toep}(\mathbf{u})\mathbf{G}^H + e^{ja}\mathbf{G}\mathrm{Toep}(\mathbf{u})\mathbf{F}^H & \\ -2\cos(b)\mathbf{G}\mathrm{Toep}(\mathbf{u})\mathbf{G}^H & \succcurlyeq 0. \end{aligned} \tag{20}$$

Here $\mathbf{\Pi}_{M_t}$ is a permutation matrix with ones on its anti-diagonal, and zeros elsewhere, $a = b = \frac{\pi}{2}$ and

$$\mathbf{F} = \begin{bmatrix}\mathbf{0}_{(M_t-1),1} & \mathbf{I}_{M_t-1}\end{bmatrix}, \tag{21}$$

$$\mathbf{G} = \begin{bmatrix}\mathbf{I}_{M_t-1} & \mathbf{0}_{(M_t-1),1}\end{bmatrix}. \tag{22}$$

Note the slight abuse of notation in (20), as the index i indicates two rows, due to the construction of the mirrored measurements.

V. Numerical Simulations

In Section II we mentioned that in practice $\alpha < 4$. In Section III we showed that α should be lower than $\frac{1}{2}$ for perfect signal recovery, while a lower bound is given by (17). In this section, we verify the bounds with numerical simulations and show what happens if prior knowledge is included, as described in Section IV.

First, we look at the value α that is attained in practice when no prior knowledge is included. We investigate the probability of successful recovery with respect to the frequency separation as a function of α that is required by the atomic norm (see (9)). We simulate the model described by (3), however, we change the set of modal frequencies to $k_n = n \cdot 2\pi F \frac{\alpha}{M_t-1}$. We consider the full data case, $M_o = M_t = 15$. The source emits L frequencies, which are i.i.d. from $\mathcal{U}[0, c\pi F)$. The modal frequency separation is changed by using different $\alpha \in \{0.5, 0.55, 0.60, \ldots, 2\}$. We solve the SDP in (8) and retrieve the frequencies by performing ESPRIT [18] on the the optimal Toeplitz matrix.

We say the signal is recovered with success if the number of retrieved modal frequencies $\hat{N}$ is equal to the actual number of modal frequencies in the signal N that lay in the Nyquist range $(-\pi F, \pi F)$, thus $\hat{N} = N$ and if additionally

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\hat{k}_n - k_n)^2} < 10^{-5} \tag{23}$$

where $\{\hat{k}_1, \ldots, \hat{k}_N\}$ denote the retrieved modal frequencies and $\{k_1, \ldots, k_N\}$ denote the actual modal frequencies present in the signal. In total 30 Monte Carlo runs are performed for each value of α. The successes are averaged over the runs to get an estimate of the probability of successful recovery. The result is shown in Fig. 4. We make the following observations:

- The derived bound in (17) is respected.

- The atomic norm attains values close to the bound and in practice it performs better than the $\alpha = 4$ that was theoretically proven for the general case (meeting the independence assumption on the ϕ_k's) by [11].
- As suggested by the derived bound in (17), increasing the number of excitation frequencies L leads to a smaller lower bound on α, and the atomic norm also attains lower values if more excitation frequencies are included.
- The atomic norm is not able to reach $\alpha < 0.5$ and is thus not able to perfectly reconstruct the modal frequencies of a boxed-shaped room.

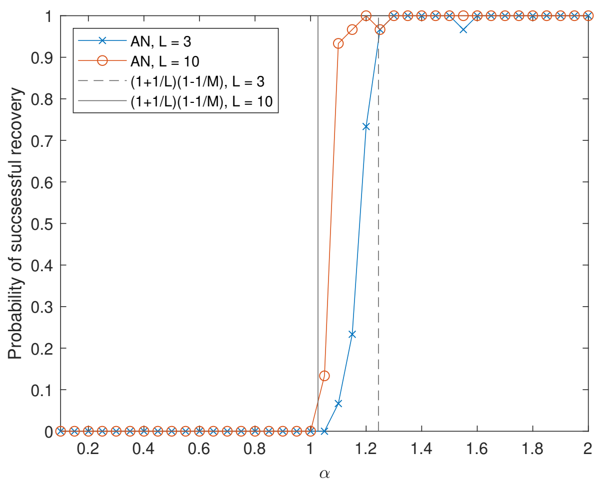

Fig. 4. Probability of successful recovery for varying frequency separation, averaged over 30 Monte Carlo runs, $M = M_o = M_t = 15$.

Finally we simulate the room with the correct modal frequencies, $k_n = n\frac{\pi}{L_x}$, while we also include more prior knowledge, as discussed in Section IV. We only simulate the positive frequencies, thus the first $\frac{N-1}{2}$ columns of $\mathbf{A}$ are omitted. We use $M_o = M_t = 15$, $L_x = 5$, $F = 3$, the source emits $L = 50$ frequencies, which are i.i.d., $\omega_l \sim \mathcal{U}[0, c\pi F)$. We solve the SDP in (20). The magnitude of the dual polynomial of the problem is shown in Fig. 5, which attains one exactly at the modal frequencies. The RMSE of the estimated frequencies is in the order of 10^{-8}.

VI. Conclusion

We have investigated if we can estimate the room modes using the atomic norm, a gridless method. We have derived a bound for the sound field reconstruction problem and relate this to the frequency separation that is required by the atomic norm for perfect reconstruction. We show that prior knowledge has to be included to be able to reach perfect reconstruction by the atomic norm. We describe two strategies that, in conjunction, can be used to realize successful recovery. Our findings are confirmed by numerical simulations.

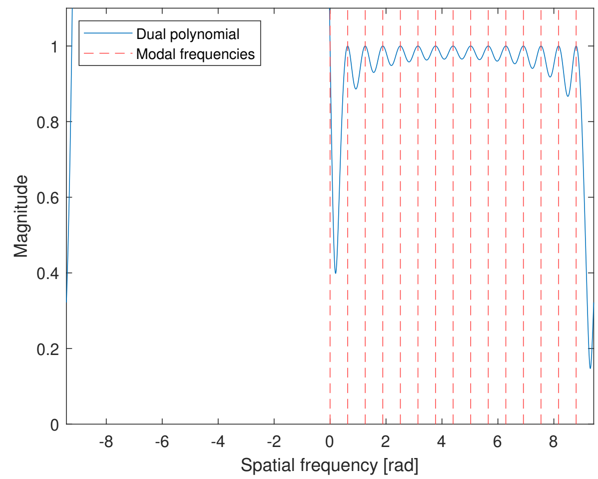

Fig. 5. Magnitude of dual polynomial, $M = 15, L = 50$.

References

[1] S. Cecchi, A. Carini, and S. Spors, "Room response equalization — A review," *Appl. Sci.*, vol. 8, no. 1, 2018, Art. no. 16.

[2] P. A. Naylor and N. D. Gaubitch, *Speech Dereverberation*. New York, NY, USA: Springer, 2010.

[3] T. Betlehem, W. Zhang, M. A. Poletti and T. D. Abhayapala, "Personal Sound Zones: Delivering interface-free audio to multiple listeners," in *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 81-91, March 2015.

[4] F. Lluís, P. Martínez-Nuevo, M. Møller and S. E. Shepstone, "Sound field reconstruction in rooms: inpainting meets super-resolution," *J. Acoust. Soc. Am.*, vol. 148, no. 2, pp. 649–659, 2020.

[5] T. P. Vu and H. Lissek, "Low frequency sound field reconstruction in a non-rectangular room using a small number of microphones," *Acta Acust.*, vol. 4, no. 2, 2020.

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

[7] E. F. Grande, "Sound field reconstruction in a room from spatially distributed measurements," in *Proc. 23rd Int. Congr. on Acoust.*, 2019, pp. 4961–4968.

[8] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *J. Roy. Stat. Soc. Ser. B*, vol. 58, no. 1, 1996, pp. 267–288.

[9] Y. Chi, L. L. Scharf, A. Pezeshki and A. R. Calderbank, "Sensitivity to Basis Mismatch in Compressed Sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182-2195, May 2011.

[10] E. Candès and C. Fernandez-Granda, "Towards a Mathematical Theory of Super-Resolution," in *Commun. Pure Appl. Math.*, vol. 67, no.6, pp. 906–956, 2014.

[11] G. Tang, B. N. Bhaskar, P. Shah and B. Recht, "Compressed Sensing Off the Grid," *IEEE Trans. Inform. Theory*, vol. 59, no. 11, pp. 7465-7490, Nov. 2013.

[12] H. Kuttruff, "The sound field in a closed space," in *Room acoustics*, 4th ed. London: Spon Press, 2005.

[13] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," in *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21-30, March 2008.

[14] Z. Yang and L. Xie, "Exact Joint Sparse Frequency Recovery via Optimization Methods," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5145-5157, Oct. 2016.

[15] J. Unnikrishnan and M. Vetterli, "Sampling and Reconstruction of Spatial Fields Using Mobile Sensors," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2328-2340, May 2013.

[16] T. Ajdler, L. Sbaiz and M. Vetterli, "Dynamic measurement of room impulse responses using a moving microphone," *J. Acoust. Soc. Am.*, vol. 122, no. 3, pp. 1636–1645, 2007.

[17] H. Chao and L. Vandenberghe, "Extensions of semidefinite programming methods for atomic decomposition," in *Proc. 2016 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2016, pp. 4757-4761.

[18] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984-995, Jul. 1989.