# Coherence Based Trajectory Optimization for Compressive Sensing of Sound Fields

Fabrice Katzberg, Marco Maass, and Alfred Mertins
*University of Lübeck*
*Institute for Signal Processing*
D-23562 Lübeck, Germany
{f.katzberg, marco.maass, alfred.mertins}@uni-luebeck.de

***Abstract*—Moving microphones allow for the fast acquisition of spatially dense sound-field data. The dynamic samples encode parameters that describe the particular sound field. For parameter decoding, a deconvolution problem in the time dimension and an interpolation problem in the spatial dimension must be solved simultaneously. The corresponding system of linear equations tends to be ill-posed and underdetermined unless an excessive number of samples is provided. Therefore, sparse recovery according to the compressed-sensing paradigm is required, which is achievable by exploiting the sparsity of sound fields in frequency domain. At this, stability and robustness depend on the sensing matrix and can be indicated by its coherence. For optimizing the coherence, mathematical tools that operate directly on the sensing matrix are impractical, as the dynamic-sensing matrix possesses structured entries constrained by the measurement trajectory. In this paper, we present an efficient update scheme that allows for the direct manipulation of the microphone trajectory for improving the coherence of the sensing matrix and, thus, reducing error bounds of the sparse recovery in frequency domain.**

***Index Terms*—Compressed sensing, dynamic sound-field measurements, trajectory optimization, room impulse responses.**

## I. Introduction

The estimation of sound fields from spatio-temporal samples poses a common inverse problem that is crucial in many audio applications dealing with reverberant environments. Sound fields within listening areas are described by spatio-temporal room impulse responses (RIRs), and the signal received at a particular listening position is modeled as the convolution result between the source signal and the RIR for the corresponding emitter-receiver configuration.

In stationary setups, RIRs are typically measured by using deterministic excitation signals that allow for straightforward deconvolution, such as maximum length sequences (MLSs) [1] and sine sweeps [2]. In order to reduce hardware effort and sampling time for acquiring spacious sound-field data, several methods have been proposed that exploit sparse signal structures according to the principle of compressed sensing (CS) [3] and allow for qualified sound-field estimates from spatially undersampled measurements [4], [5].

In dynamic setups, one moving microphone is sufficient for capturing entire sound-field information. An analytical approach for the reconstruction of RIRs along linear and circular trajectories, sampled at constant speed, has been presented in [6]. Beyond that, dynamic techniques are typically based on estimates from systems of linear equations that encode the convolutions of RIRs and source signals, and relate them to the measured samples. In one group of such methods, impulse responses at steadily changing positions are modeled in terms of time-varying systems, whose coefficients are tracked by adaptive filtering concepts [7], [8].

More recently, another group of dynamic methods has evolved, relying on a time-invariant spatio-temporal measurement model [9]–[12]. Here, the reconstruction from dynamic sound-field samples is interpreted as spatial interpolation task which is supplementally incorporated into the linear measurement equations by using positional information of the microphone. In [9], perfect-sequence excitation is used for the orthogonal expansion of impulse responses in time domain, in order to describe the dynamic spatio-temporal sampling by notional static sampling processes of single expansion coefficients. The method in [10] uses a spatial Fourier basis for the angular reconstruction of head-related impulse responses from continuous-azimuth recordings. In [11], a dynamic framework has been presented that allows for RIR reconstruction within cubical volumes. To achieve this, the sound field is parameterized by modeling notional grid points in space, and dynamic samples are understood as the result of bandlimited interpolation on that grid using sampled sinc-function approximations. The solution to the resulting linear system leads to RIRs on a virtual grid that enable spatial reconstruction. In practice, this inverse problem is most likely ill-posed or even underdetermined, thus, a CS based strategy has been proposed in [12] using sparse Fourier representations.

According to the CS paradigm, stable and robust parameter recovery is guaranteed for sensing matrices with low coherence. Lowering the coherence of the sensing matrix leads to the reduction of upper bounds for the recovery error [13]. Randomly generated sensing matrices were proven to be good choices for obtaining incoherent measurements [14], [15]. Nevertheless, it has also been shown that several iterative strategies often lead to well-designed deterministic matrices with optimized coherence and performance [16]–[20]. Most of these methods operate directly on the sensing matrix or, respectively, on the resulting Gram matrix. However, in many applications, entries of the matrix are not the design variables themselves (e.g., sampling positions), but rather result from

This work has been supported by the German Research Foundation under Grant No. ME 1170/10-2.

particular relationships. For cases where the sensing matrix is some nonlinear function of the model parameters for the CS problem, an iterative scheme is given in [20]. That algorithm performs an alternating minimization procedure and solves augmented Lagrangian subproblems, in order to improve the coherence subject to the design parameters.

In this paper, we propose a simple and powerful tool for optimizing the coherence of the dynamic-sampling problem presented in [12]. At that, there is no need to explicitly set up the sensing matrix, which is favorable for sound-field sampling at high frequencies involving a high number of unknown variables. Iteratively, a low-complexity objective function describing maximum correlation in frequency domain is designed, and a gradient descent step with respect to trajectory positions is performed. In numerical experiments we show that just slight adjustments of the microphone trajectory may lead to lower coherence and reduced recovery error.

## II. Spatio-Temporal Sampling Model

In a closed-room scenario, the propagation of the sound signal $s(t)$ originating at fixed source position can be modeled as linear time-invariant (LTI) system described by the spatio-temporal RIR $h(\boldsymbol{x}, t)$ that depends on the listener position $\boldsymbol{x} = [x, y, z]^T$ and time $t \in \mathbb{R}$.

For a signal having maximum frequency $f_{\text{max}}$, the sampling at $f_\text{s} > 2f_{\text{max}}$ yields uniform samples at equidistant time points $t_n = n/f_\text{s}$ ($n \in \mathbb{Z}$) that allow for aliasing-free reconstruction in time. Supposing that the amplitude of $h(\boldsymbol{x}, t)$ vanishes into the noise level for $t_n > t_{L-1}$, we model the temporal sampling of the sound-pressure field $p(\boldsymbol{x}, t)$ according to

$$p(\boldsymbol{x}, n) = \sum_{m=0}^{L-1} s(n-m)\, h(\boldsymbol{x}, m) + \eta(\boldsymbol{x}, n), \tag{1}$$

with $\eta(\boldsymbol{x}, n)$ being the measurement noise.

### A. Spatial Model for RIR Parameterization

Microphones generate samples at uniform points in time, but, for general setups, at non-uniform spatial positions inside the volume of interest $\Omega$. In order to describe spatio-temporal sound-field measurements, we model an appropriate parameterization of the continuous-space and discrete-time RIR within bounds of $\Omega$. A straightforward parameterization is given by the bandlimited interpolation in space on a three-dimensional grid of size $X \times Y \times Z$,

$$h(\boldsymbol{x}, n) \approx \sum_{p=1}^{G} h(\boldsymbol{g}_p, n)\, f_p(\boldsymbol{x}) \quad \forall\, \boldsymbol{x} \in \Omega, \tag{2}$$

where $\boldsymbol{g}_p \in \left\{\boldsymbol{x}_0 + [g_x\Delta_x, g_y\Delta_y, g_z\Delta_z]^T\right\}$ constitutes $G = XYZ$ grid positions at discrete grid points $[g_x, g_y, g_z]^T \in \mathbb{Z}^3$ with spacing $\Delta_\xi < c/(2f_{\text{max}})\, \forall\, \xi \in \{x, y, z\}$ [21] for speed of sound $c$ and grid origin $\boldsymbol{x}_0$, $h(\boldsymbol{g}_p, n)$ represents grid RIRs (parameters), and $f_p(\boldsymbol{x}) = f_p(x)\, f_p(y)\, f_p(z)$ are separable basis functions approximating the sinc functions centered at $\boldsymbol{g}_p$ and evaluated at point $\boldsymbol{x}$. The time dimension of the RIR is encoded in the parameterization itself which leads to $P = LG$ sound-field variables required for describing $h(\boldsymbol{x}, n)$ inside $\Omega$.

Using (2), the spatial sampling of (1) can be modeled as

$$\boldsymbol{p} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{\eta}, \tag{3}$$

where the measurement vector $\boldsymbol{p} \in \mathbb{R}^{MQ}$ contains the stack of $Q$ microphone signals of lengths $M$, $\boldsymbol{A} \in \mathbb{R}^{MQ \times P}$ is the sampling matrix, $\boldsymbol{h} \in \mathbb{R}^P$ encapsulates $P$ sound-field parameters, and $\boldsymbol{\eta}$ is the perturbation vector compensating for errors due to both sampling and parameterization. Of course, other parameterization models are possible, however, using bandlimited interpolation provides a capable framework due to its simplicity with separable lowpass filters that allow for error analyses at low computational cost [11], [12].

### B. Moving Microphones

In dynamic setups, sound-field information is acquired at non-uniform and time-varying positions inside $\Omega$. In this case, only one microphone ($Q = 1$) moving along the sampled trajectory $\boldsymbol{x}(n)$ is sufficient for gathering entire sound-field information. This setup is considered in the following descriptions. The expansion to $Q$ dynamic microphones is straightforward and may reduce acquisition time.

In order to fit the dynamic setup into the sampling and parameterization models according to (1), (2) and (3), respectively, we need to extend the position vector to its temporal dependency and interrelate the measured sound pressure on the trajectory and the unknown RIR on the trajectory subject to $p(\boldsymbol{x}(n), n) = h(\boldsymbol{r}, n) * s(n)|_{\boldsymbol{r}=\boldsymbol{x}(n)}$ [11]. Altogether, the incorporation of instantaneously varying microphone positions leads to the measurement matrix

$$\boldsymbol{A} = \left[\boldsymbol{D}_1\boldsymbol{S}, \boldsymbol{D}_2\boldsymbol{S}, \ldots, \boldsymbol{D}_G\boldsymbol{S}\right], \tag{4}$$

with $\boldsymbol{D}_p = \operatorname{diag}\{f_p(\boldsymbol{x}(0)), \ldots, f_p(\boldsymbol{x}(M-1))\}$ and $\boldsymbol{S} \in \mathbb{R}^{M \times L}$ being the convolution matrix of the source signal $s(n)$.

## III. Sparse Recovery

In practice, the linear system (3) easily tends to be underdetermined due to several reasons [12]. For larger bandwidths, the modeled grid in space demands a high number of unknown variables. Sampling constraints often lead to $M < P$ and, thus, to an infinite number of least-squares solutions. Nevertheless, the spatio-temporal spectrum of $h(\boldsymbol{x}, t)$ is ideally measurable on a sparsely occupied hypercone along the temporal frequency axis [21]. Thus, the principle of CS can be used to identify a stable and robust estimate for $\boldsymbol{h}$ living in a subspace of dimension $K \ll M < P$.

Let the matrix $\boldsymbol{F} \in \mathbb{C}^{U \times U}$ perform the discrete Fourier transform (DFT) of length $U$. Exploiting the separability on the uniform grid and using the Kronecker product $\otimes$, we define $\boldsymbol{\Psi} = \boldsymbol{F}_Z \otimes \boldsymbol{F}_Y \otimes \boldsymbol{F}_X \otimes \boldsymbol{T}_L$ and represent the sound-field parameters by $\boldsymbol{c} = \boldsymbol{\Psi}\boldsymbol{h}$ , which corresponds to the four-dimensional DFT of the virtual-grid RIR $h(\boldsymbol{g}_p, n)$. The least-squares problem may be constrained to the $K$-sparse problem

$$\underset{\boldsymbol{c} \in \mathbb{C}^P}{\arg\min}\ \|\boldsymbol{p} - \boldsymbol{\mathcal{A}}\boldsymbol{c}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{c}\|_0 \leq K, \tag{5}$$

with the sensing matrix $\boldsymbol{\mathcal{A}} = \boldsymbol{A}\boldsymbol{\Psi}^{-1}$, and $\|\boldsymbol{c}\|_0 = |\{i : c_i \neq 0\}|$ quantifying the support of $\boldsymbol{c}$. The problem (5) is NP-hard [22] and is solved in practice by using greedy algorithms [23], [24] or applying convex optimization to corresponding $\ell_1$-minimization problems [3], [25].

In order to achieve stable and robust recovery, the sensing matrix must satisfy the so-called restricted isometry property (RIP) with an RIP constant being as small as possible [26]. This constant determines upper bounds for recovery errors induced by measurement noise, the $K$-sparse signal approximation, and a mismatch of $\boldsymbol{\mathcal{A}}$ that in our case may occur due to impreciseness in position tracking. Verifying the RIP of a matrix is an NP-hard problem [27]. Sensing matrices generated by independent and identically distributed random processes [14] and random ensembles [15], respectively, meet the RIP with very high probability for a wide range of matrix sizes. However, in practice, totally random setups are often impossible. The setup may underlie certain restrictions and points on the trajectory are highly dependent on each other as the speed of a moving microphone is limited. Then, an indicator for RIP guarantees is given by the coherence measure

$$\mu(\boldsymbol{\mathcal{A}}) = \max_{1\leq u\neq v\leq P} \frac{|\langle \boldsymbol{a}_u, \boldsymbol{a}_v\rangle|}{\|\boldsymbol{a}_u\|_2 \, \|\boldsymbol{a}_v\|_2}, \tag{6}$$

with $\boldsymbol{a}_u$ being the $u$-th column of $\boldsymbol{\mathcal{A}}$. The error bounds of sparse recovery improve for a smaller coherence [13].

## IV. Proposed Coherence Optimization

In the following, we present a fast algorithm allowing for adjustments of given microphone positions, in order to reduce the coherence and improve the accuracy of the recovered sound-field signal. By assuming ideal lowpass filters in the parameterization, the origin of high correlations $|\langle \boldsymbol{a}_u, \boldsymbol{a}_v\rangle|$ is identified efficiently. Based on that, we propose to adapt the microphone trajectory by gradient descent steps, in order to approach a local minimum of maximum correlation and, thus, to obtain a sensing matrix better suited to the CS paradigm.

### A. Influence of Varying Microphone Positions

First, consider the measurement space to be confined to a line in $x$-dimension. Then, $\boldsymbol{\Psi}$ performs the two-dimensional DFT with frequency variables $k_x \in \{-\frac{X-1}{2}, \ldots, \frac{X-1}{2}\}$ and $l \in \{-\frac{L-1}{2}, \ldots, \frac{L-1}{2}\}$ for $X, L$ being odd. Let us define the signal $s_n(g_x, m) = s(n-m) f_{g_x}(\boldsymbol{x}(n))$ as the spatio-temporal, spectrally flat excitation on the uniformly sampled line and its discrete Fourier spectrum $\mathcal{S}_n(k_x, l)$ forming the $LX$ columns of $\boldsymbol{\mathcal{A}}$: $\boldsymbol{a}_{(k_x,l)} = [\mathcal{S}_0(k_x, l), \ldots, \mathcal{S}_{M-1}(k_x, l)]^T$. Then, using the grid-related trajectory $r_x(n) = (x(n) - x_0)/\Delta_x$, it can be shown [12] that the moving of the microphone from point $x(n)$ to $x(n+d)$ manipulates $\boldsymbol{\mathcal{A}}$ according to

$$\mathcal{S}_{n+d}(k_x, l) = \mathcal{S}_n(k_x, l)\, e^{-2\pi j d \frac{l}{L}}\, e^{-2\pi j (r_x(d) - r_x(n)) \frac{k_x}{X}}. \tag{7}$$

### B. Efficient Coherence Calculations

For measurements in three-dimensional space, we can use (7) and exploit the separability on a multidimensional grid, in order to describe the $LG$ columns of $\boldsymbol{\mathcal{A}}$ by structured spatial and temporal phase terms of different frequency pairs. This structure in $\boldsymbol{\mathcal{A}}$ allows for calculating the coherence (6) directly in respect of the grid-related trajectory $\boldsymbol{r}(n) = [r_x(n), r_y(n), r_z(n)]^T$ at low complexity $\mathcal{O}(LG)$: for finding the maximum correlation between columns, only combinations of frequency distances $\bar{l} = l' - l''$, $\bar{k}_\xi = k'_\xi - k''_\xi$ are relevant:

$$\tilde{\mu}(\boldsymbol{r}(\cdot)) = \max_{\bar{\boldsymbol{f}}\neq \boldsymbol{0}} \frac{1}{M} \left| \sum_{n=0}^{M-1} e^{-j\mathcal{C}_{\bar{\boldsymbol{f}}}(\boldsymbol{r}(n))} \right|, \tag{8}$$

with $\bar{\boldsymbol{f}} = (\bar{l}, \bar{\boldsymbol{k}})$, $\bar{\boldsymbol{k}} = (\bar{k}_x, \bar{k}_y, \bar{k}_z)$, $\mathcal{C}_{\bar{\boldsymbol{f}}}(\boldsymbol{r}(n)) = \mathcal{T}_{\bar{l}}(n) + \mathcal{X}_{\bar{\boldsymbol{k}}}(\boldsymbol{r}(n))$, involving the temporal relationship $\mathcal{T}_{\bar{l}}(n) = 2\pi n \bar{l}/L$ and the positional dependency

$$\mathcal{X}_{\bar{\boldsymbol{k}}}(\boldsymbol{r}(n)) = 2\pi \left( r_x(n)\frac{\bar{k}_x}{X} + r_y(n)\frac{\bar{k}_y}{Y} + r_z(n)\frac{\bar{k}_z}{Z} \right). \tag{9}$$

The quantity (8) is equivalent to (6) for spectrally flat behavior of the measuring process in any dimension, which can be met by selecting appropriate design parameters. For the temporal dimension, the choice of $L$-shift cross orthogonal excitation sequences is suitable (e.g., MLSs). For the spatial dimensions, using higher order maximally flat fractional delay filters and modeling the grid spacing according to (at least) twofold spatial oversampling $\Delta_\xi \leq c/(4 f_{\text{max}})$ is appropriate. Then, the bandlimited signal is located at the lower halfband where the interpolator approaches ideal response, and the model of $\boldsymbol{\mathcal{A}}$ (and, thus, (6) and (8)) can be reduced to these low spatial frequencies. For a non-ideal design, the measure (8) is an efficient approximation of (6) (cf. [12]).

### C. Fast Update Scheme for Trajectory Adjustments

A simple procedure is proposed for updating the trajectory $\boldsymbol{r}(n)$ subject to the minimization of the maximum correlation between two columns of $\boldsymbol{\mathcal{A}}$. Exploiting the low-complexity measure (8), we define the objective function

$$J(\boldsymbol{r}(\cdot)) = \left| \sum_{n=0}^{M-1} e^{-j\mathcal{C}_{\bar{\boldsymbol{f}}'}(\boldsymbol{r}(n))} \right|, \tag{10}$$

where the frequency distances in $\bar{\boldsymbol{f}}'$ are selected in accordance with the highest spectral correlation for the current trajectory,

$$\bar{\boldsymbol{f}}' = \operatorname*{arg\,max}_{\bar{\boldsymbol{f}}\neq \boldsymbol{0}} \left| \sum_{n=0}^{M-1} e^{-j\mathcal{C}_{\bar{\boldsymbol{f}}}(\boldsymbol{r}(n))} \right|. \tag{11}$$

In order to minimize (10), different scenarios can be considered by adapting either one single position (e. g., to find optimal direction of future movement), multiple, or even all points on the trajectory simultaneously at iteration $i$. The latter case emphasizes the need for a cost-effective optimization, as the number of measuring positions in dynamic setups is, in general, as high as the number of provided samples.

Updates for one particular position variable $r_\xi(n^*)$ are performed following the gradient descent scheme

$$r_\xi^{[i+1]}(n^*) = r_\xi^{[i]}(n^*) - \alpha \; \partial J(\boldsymbol{r}(\cdot))/\partial r_\xi^{[i]}(n^*), \tag{12}$$

with $\alpha$ being a small step size. Each iteration goes along with a redesign of the objective function in a greedy fashion: if the

origin of the maximum correlation, i.e. (11), relocates, then $J(\boldsymbol{r}(\cdot))$ is adapted accordingly.

By using Euler's formula and defining the summations $\sigma_{\cos}(\boldsymbol{r}(\cdot)) = \sum_{n=0}^{M-1} \cos(\mathcal{C}_{\bar{\boldsymbol{f}}'}(\boldsymbol{r}(n)))$ and $\sigma_{\sin}(\boldsymbol{r}(\cdot))$, analogously, the objective function can be rewritten according to $J(\boldsymbol{r}(\cdot)) = (\sigma^2_{\cos}(\boldsymbol{r}(\cdot)) + \sigma^2_{\sin}(\boldsymbol{r}(\cdot)))^{1/2}$, in order to deduce simple expressions for the partial derivatives composing the considered gradient. Applying the chain rule several times and using trigonometric identities, the partial derivative subject to the specific position variable reads

$$\frac{\partial J(\boldsymbol{r}(\cdot))}{\partial r_\xi(n^*)} = \gamma_\xi \sum_{n=0}^{M-1} \sin\left(\mathcal{C}_{\bar{\boldsymbol{f}}'}(\boldsymbol{r}(n) - \boldsymbol{r}(n^*))\right), \qquad (13)$$

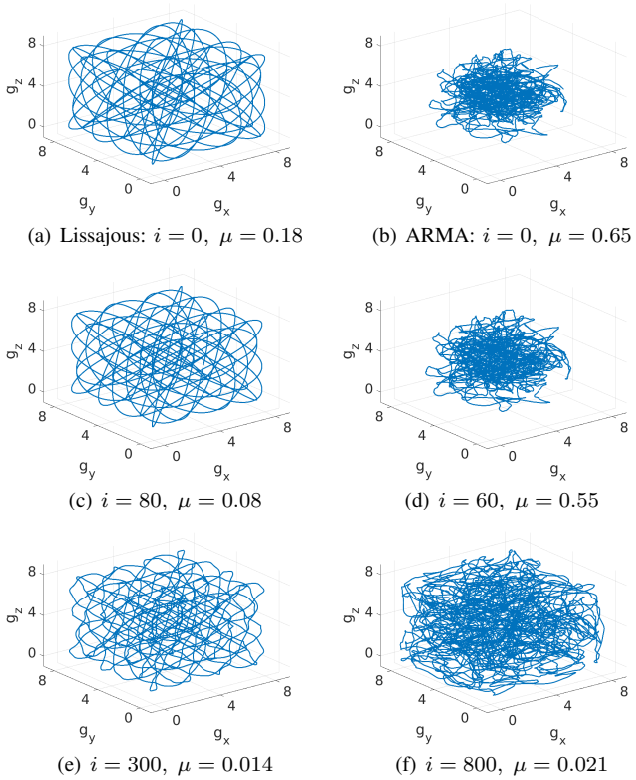
where the factor $\gamma_\xi = (2\pi \bar{k}'_\xi)/(DJ(\boldsymbol{r}(\cdot)))$ depends on the particular dimension $\xi \in \{x, y, z\}$ and $D \in \{X, Y, Z\}$. All in all, efficient adjustments of trajectory positions are performed by iteratively finding maximum spectral correlation according to (11), calculating the gradient of the free variables in terms of (13), and updating subject to (12), until some predefined exit conditions are reached. These could be, for example, constraints respecting limitations of the measurement setup, i.e., boundaries of the spatial grid, maximum distance between positions, or maximum microphone speed.

### *D. Adaptation to Sensor Placement Problems*

The presented model design and technique of position optimization is actually adaptable to any sampling problem requiring the recovery of bandlimited, spectrally sparse functions on a multidimensional Cartesian grid. For example, we can also consider a stationary sensor setup of $Q$ microphones at grid-related locations $\boldsymbol{r}_q$ and obtain a conformable and even more structured sensing matrix where the time-varying component dissolves. Then, the coherence based objective function simplifies to $J(\boldsymbol{r}.) = |\sum_{q=1}^{Q} e^{-j\mathcal{X}_{\bar{\boldsymbol{k}}'}(\boldsymbol{r}_q)}|$. The corresponding analysis is similar and leads to position updates by analogy with (12) and (13) setting the temporal dependency to $\bar{l}' = 0$ and, thus, $\mathcal{T}_{\bar{l}'}(n) = 0$. Also, the same formulas are obtained for the pure sound-field interpolation task presented in [28], where RIRs are directly available at several locations.

## V. Experiments and Results

It has been shown in [11] that for sufficiently long sampling the minimum mean squared error of estimates from the inverse problem (3) with (4) becomes smaller for trajectory positions sampled closer to the notional grid points. For the sparse recovery in frequency domain, no such general statement on the trajectory is applicable. Nevertheless, in [12] the use of a Lissajous trajectory covering the entire volume of interest has been experimentally shown to be a good choice, performing even better than totally random dynamic sampling positions. Here, we use the presented optimization method to modify the points on a Lissajous trajectory (Fig. 1(a)). Also, we test the algorithm on a random trajectory resulting from an autoregressive moving average process (ARMA) in each dimension (Fig. 1(b)). The trajectories are sampled at $4 \cdot 10^5$ positions and scaled to directly fit into the cubical volume $\Omega$.

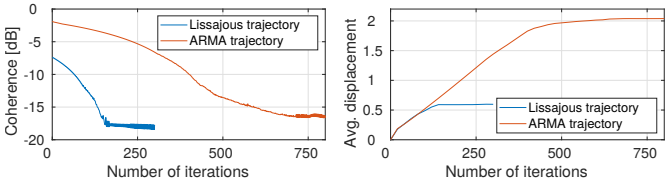



(a) Lissajous: $i = 0,\ \mu = 0.18$ (b) ARMA: $i = 0,\ \mu = 0.65$

(c) $i = 80,\ \mu = 0.08$ (d) $i = 60,\ \mu = 0.55$

(e) $i = 300,\ \mu = 0.014$ (f) $i = 800,\ \mu = 0.021$

Fig. 1. Optimized microphone trajectories. (a)-(b) Original states and (c)-(f) improved versions after $i$ iterations leading to lower coherence $\mu$.




(a) (b)

Fig. 2. (a) Reduction of coherence over the number of iterations $i$ and (b) the corresponding average displacement from the original state at $i = 0$.

For sound-field parameterization, we define $9^3$ notional grid points with spacings $\Delta_\xi = c/(4f_{\max})$, covering $\Omega$ entirely. The number of RIR taps is limited to $L = 2047$. To achieve spectral flatness of excitation, MLSs with period length $L$ are chosen, scaled to yield zero DC offset, and played over 200 periods during measurements. For interpolation in space, Lagrangian fractional delay filters of order 11 are used, ensuring (nearly) ideal frequency response at the lower halfband. This setup originally results in sensing matrices $\boldsymbol{\mathcal{A}}$ of size $4{\cdot}10^5 \times 1.5{\cdot}10^6$, having coherence $\mu = 0.18$ for the Lissajous trajectory and $\mu = 0.65$ for the trajectory based on ARMA processes.

Without building up large sensing matrices, the proposed optimization technique was capable of reducing the coherence by manipulating the pre-defined trajectory positions. Performing the update rule (12) with (13) on each trajectory point at step size $\alpha = 0.1/(2\pi)$ obtained setups where the coherence of the corresponding CS problem was significantly lowered by 0.1 after only a couple of iterations (Figs. 1(c) and 1(d)), and finally reached a minimum in $\mu = 0.014$ and $\mu = 0.021$,

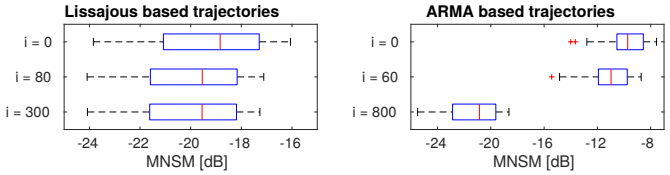



Fig. 3. Recovery error using original trajectories and the improved versions.

respectively (Figs. 1(e) and 1(f)). No constraints were made on the resulting microphone velocity. In Fig. 2(a), the decrease of coherence is shown subject to the number of iterations. The sporadic peaks in Fig. 2(a) result from the greedy-like adaptation of $J(\boldsymbol{r}(\cdot))$ according to the current maximum correlation. The average change of positions compared to the original state $\boldsymbol{r}^{[0]}(n)$, $\sum_{n=0}^{M-1} \|\boldsymbol{r}^{[0]}(n) - \boldsymbol{r}^{[i]}(n)\|/M$, is depicted in Fig. 2(b). For both types of trajectories, just little changes in $r_\xi(n)$ by $0.2$ on average led to the coherence improvement by $0.1$. At this, the ARMA based trajectory took a slightly more expansive configuration (Fig. 1(d)), and the Lissajous type was squeezed, having a bit more distance to the grid points at the boundary of $\Omega$ (Fig. 1(c)).

In numerical experiments, we tested the trajectories from Fig. 1 for sparse sound-field recovery in frequency domain. Using the image source method [29], we simulated dynamic measurements in fifty reverberant environments, randomly chosen according to the uniform distribution of room sizes $[2\,\mathrm{m}; 10\,\mathrm{m}]^3$, sampling frequencies $f_\mathrm{s} \in [6000; 16000]\,\mathrm{Hz}$, and reverberation times $T_{60} \in [0.05\,\mathrm{s}; L/f_\mathrm{s}]$. The source position and the origin of $\Omega$ were randomly selected in each experiment. White Gaussian measurement noise was added at a signal-to-noise ratio of $20\,\mathrm{dB}$. We applied the CS based reconstruction algorithm from [12]. To evaluate spatial RIR recovery, the MNSM error measure as in [11], [12] is used. In Fig. 3, the sound-field recovery errors based on the particular measurement trajectories are presented in box plots. The adjusted trajectories lead to better performance, also for the cases where positions are just slightly manipulated.

## VI. Conclusions

In this paper, we presented a simple method of updating microphone positions on a pre-defined trajectory, in order to obtain measurements that lead to lower coherence of the sensing matrix. This makes the dynamic-sampling procedure more compatible with the CS principle and leads to reduced errors of the sparse recovery in frequency domain.

## References


[1] D. D. Rife and J. Vanderkooy, "Transfer-function measurement with maximum-length sequences," *J. Audio Eng. Soc.*, vol. 37, no. 6, pp. 419–444, 1989.

[2] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Proc. 122th Conv. Audio Eng. Soc.*, Conv. Paper 7121, 2007.

[3] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[4] A. Benichoux, L. Simon, E. Vincent, and R. Gribonval, "Convex regularizations for the simultaneous recording of room impulse responses," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 1976–1986, 2014.

[5] R. Mignot, L. Daudet, and F. Ollivier, "Room reverberation reconstruction: interpolation of the early part using compressed sensing," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 11, pp. 2301–2312, 2013.

[6] 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.

[7] J. Benesty, Y. Huang, J. Chen, and P. A. Naylor, "Adaptive algorithms for the identification of sparse impulse responses," in *Topics in Acoustic Echo and Noise Control*, E. Hänsler and G. Schmidt, Eds. Berlin: Springer, 2006, pp. 125–153.

[8] A. Carini, "Efficient NLMS and RLS algorithms for perfect and imperfect periodic sequences," *IEEE Trans. Signal Process.*, vol. 85, no. 4, pp. 2048–2059, 2010.

[9] N. Hahn and S. Spors, "Comparison of continuous measurement techniques for spatial room impulse responses," in *Europ. Signal Process. Conf.*, 2016, pp. 1638–1642.

[10] C. Urbanietz and G. Enzner, "Spatial-Fourier retrieval of head-related impulse responses from fast continuous-azimuth recordings in the time-domain," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2019.

[11] F. Katzberg, R. Mazur, M. Maass, P. Koch, and A. Mertins, "Sound-field measurement with moving microphones," *J. Acoust. Soc. Am.*, vol. 141, no. 5, pp. 3220–3235, 2017.

[12] ——, "A compressed sensing framework for dynamic sound-field measurements," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 11, pp. 1962–1975, 2018.

[13] T. Cai, G. Xu, and J. Zhang, "On recovery of sparse signals via $\ell_1$ minimization," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3388–3397, 2009.

[14] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.

[15] H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical foundations and numerical methods for sparse recovery*, ser. Radon Series Comp. Appl. Math. 9, M. Fornasier, Ed. Berlin: de Gruyter, 2010, pp. 1–94.

[16] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, 2007.

[17] V. Abolghasemi, S. Ferdowsi, and S. Sanei, "A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing," *Signal Process.*, vol. 92, pp. 999–1009, 2012.

[18] J. Pan and Y. Qiu, "An orthogonal method for measurement matrix optimization," *Circuits Syst. Signal Process.*, vol. 35, pp. 837–849, 2015.

[19] C. Lu, H. Li, and Z. Lin, "Optimized projections for compressed sensing via direct mutual coherence minimization," *Signal Process.*, vol. 151, pp. 45–55, 2018.

[20] R. Obermeier and J. Martinez-Lorenzo, "Sensing matrix design via mutual coherence minimization for electromagnetic compressive imaging applications," *IEEE Trans. Comput. Imag.*, vol. 3, no. 2, pp. 217–229, 2017.

[21] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3790–3804, 2006.

[22] B. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.

[23] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[24] T. Blumensath and M. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 629–654, 2008.

[25] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[26] E. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[27] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1248–1259, 2013.

[28] F. Katzberg, R. Mazur, M. Maass, M. Böhme, and A. Mertins, "Spatial interpolation of room impulse responses using compressed sensing," in *Proc. Int. Workshop Acoust. Signal Enhancement*, 2018, pp. 426–430.

[29] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979.