DOA Estimation Performance with Circular Arrays in Sound Fields with Finite Rate of Innovation

Falk-Martin Hoffmann*, Philip Arthur Nelson†, Filippo Maria Fazi‡,
*F.Hoffmann@soton.ac.uk, †P.A.Nelson@soton.ac.uk, ‡Filippo.Fazi@soton.ac.uk

Abstract—A novel Direction-of-Arrival (DOA) estimation method based on a plane wave sound field model was recently proposed for circular microphone arrays [1]. This work presents a detailed theoretical analysis of the method that relies on a Finite Rate of Innovation (FRI) assumption, investigating the impact of different theoretical acoustic source models. The method's estimation accuracy, and robustness against noisy measurement data and deviations from the model are investigated. The estimation performance is validated and assessed on the basis of results obtained from both simulations and experimental data.

I. INTRODUCTION

Direction-of-Arrival (DOA) estimation with sensor arrays is a problem of continued relevance, independent of the kind of wave field under consideration. Whether the objective is underwater source localisation or the optimisation of the beam pattern of a controllable source or receiver (e.g. wireless LAN router, teleconference systems with microphone arrays, etc.), a high accuracy DOA estimate for a given source has become a key element in many transmission systems.

In applications where the receiver is in the far field of the source, often the *exact* source position (i.e. direction *and* distance) is irrelevant and only the DOA is of interest. In that case, it is common to model the wave field in which the receiver is immersed as a number of plane waves [2]–[6]. A Plane Wave (PW) is fully described, for a given frequency, by its complex amplitude and DOA, yielding four parameters in 3D space, thus making a wave field composed of a finite number of PWs eligible for a sparse representation [7], [8]. Sparse wave fields are particularly suitable for DOA estimation problems as they do not require an infinite resolution of the estimator.

The DOA estimator presented in [1] belongs to the group of parametric approaches. Parametric DOA estimation methods rely on a sound field model, and prominent examples of well-established parametric methods are MUSIC [9] and ESPRIT [2]. Both provide a high-resolution DOA estimation performance yet only for a sound field generated by narrowband sources. To overcome this limitation and enable DOA estimation with wideband sources, Teutsch and Kellermann derived the EB-ESPRIT algorithm [4]. This evolution of the ESPRIT algorithm takes the latter's subspace nature and instead of applying it to the sensor signals directly, it is applied to the modal domain of a circular microphone array. Using the EB-ESPRIT algorithm, Teutsch and Kellermann reported estimates with less than three degrees error for a single source, setting the benchmark for this work.

In [1], the authors found that a sparse sound field with a finite number of plane waves travelling in the same plane can be described through a periodic function with a *finite rate of innovation* [10], [11]. It was shown that the sampling method for signals consisting of a finite number of Dirac pulses proposed by Vetterli et al. [10] can be directly transferred to solve the problem of DOA estimation with circular arrays in the frequency domain. The work presented in [1] was based on a perfect observation of an ideal sound field and the proposed method requires further investigation to assess its realistic performance, particularly the impact of noisy measurements. In that context, the findings of Blu et al. [11] on how to improve robustness of Finite Rate of Innovation (FRI) signal analysis against noise and model imperfections are eligible for adaption to the proposed DOA estimation method.

This work presents a complete description of sound fields with plane waves travelling in the same plane and provides a detailed investigation of the robustness of the proposed method. Perturbations introduced by measurement noise, sound field imperfections (i.e. point sources and plane waves that travel in a different plane) and properties of the circular measurement array are considered. The validation was carried out on the basis of measured data. The remainder of this work is organised in six further sections. The following section introduces the sound field model. Section III derives expressions for the data obtained from a circular microphone array while observing different sound field types. Section IV recapitulates the robust FRI signal analysis methods proposed by Blu et al. [11] and presents a noise model based on the nature of the sensor array observation. A simulation study of the DOA estimation performance and its results with respect to the estimation error are described and discussed in Section V. The results of the experimental validation are presented and discussed in the Section VI, followed by the concluding remarks in Section VII.

II. SOUND FIELD MODEL

This work uses predominantly the cylindrical coordinate system, but sometimes also the elevation angle θ. These and a diagram of the geometric entities relevant for this work are depicted in Fig. 1. All considerations hereunder pertain to the frequency domain and assume a (quasi-)stationary field. The factor $e^{-i\omega t}$ and the various quantities' explicit dependency on the angular frequency ω are omitted for reasons of brevity.

The model assumes that the sound field within a bounded area under consideration, V, is in the far field of isolated point sources and satisfies the homogeneous wave equation, so that the pressure is fully defined through a bounded Herglotz Density (HD) [12], [13]. An expression for the sound field of a number of plane waves is derived hereunder.
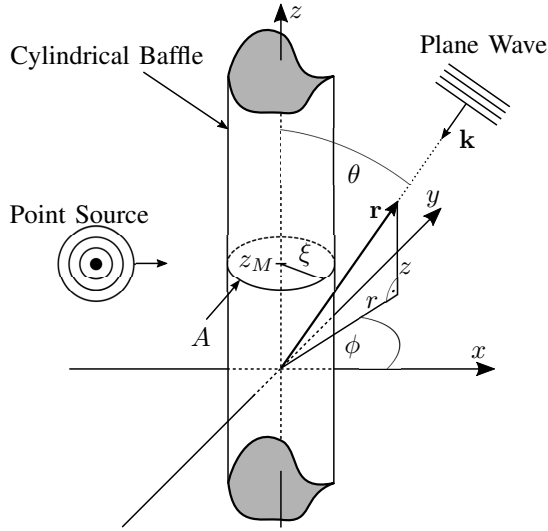
Fig. 1: Cylindrical coordinate system with $\mathbf{r} = (r, \phi, z)^T$, the elevation angle θ and other entities as they are used in this work.

A. Plane Wave Expansion in Cylindrical Coordinates

A single PW of unitary magnitude and zero-phase is defined by its wave number vector $\mathbf{k}$ (see Fig. 1) that describes the direction of travel. In cylindrical coordinates, the latter can be expressed in the form

$$\mathbf{k} = - \begin{bmatrix} k_r \cos \varphi \\ k_r \sin \varphi \\ k_z \end{bmatrix} \tag{1}$$

where the azimuth angle, φ, as well as k_r and k_z are associated with the DOA[1] of the PW. The l_2-norm of $\mathbf{k}$ provides the wave number

$$k = \|\mathbf{k}\|_2 = \sqrt{k_r^2 + k_z^2} = \frac{\omega}{c} \tag{2}$$

at the angular frequency ω and the speed of sound c. The series expansion of a single PW in cylindrical basis functions at the point $\mathbf{r} = (r, \phi, z)^T$ is given through (see Appendix A)

$$e^{i\mathbf{k} \cdot \mathbf{r}} = e^{-ik \cos \theta z} \sum_{n=-\infty}^{\infty} i^{-n} J_n(kr \sin \theta) e^{in(\phi - \varphi)}. \tag{3}$$

The above expression represents the sound field of a PW through the product of angular and axial modes, where only one of the latter is required for a single PW. It represents the kernel for the Herglotz Wave Function (HWF) [14] that expresses a sound field as the superposition of PWs.

B. The Herglotz Wave Function

The PWs forming the sound field can travel in any direction associated with the points on the sphere

$$\Omega := \{ \mathbf{r} : \|\mathbf{r}\|_2 = 1 \} \tag{4}$$

[1]The minus sign in (1) converts the vector pointing in the DOA to a vector pointing in the direction of travel instead.

around the origin. Each PW has an individual complex amplitude $q(\mathbf{k})$, $\frac{\mathbf{k}}{k} \in \Omega$. The sound field is then given through the HWF [14]

$$p(r, \phi, z) = \int_\Omega e^{i\mathbf{k} \cdot \mathbf{r}} q(\mathbf{k}) d\Omega \left(\frac{\mathbf{k}}{k} \right), \tag{5}$$

where $q(\mathbf{k})$ is known as the aforementioned HD. The HWF poses an expression for any sound field that satisfies the assumptions made at the beginning of this section.

C. Sound Field of L Plane Waves

Let the sound field consist of a finite set of L distinct far-field sources, i.e. PWs. Each PW is specified by a complex amplitude b_l, and a DOA (φ_l, θ_l), and the HD is of the form

$$q(\phi, \theta) = \sum_{l=1}^{L} \frac{b_l}{\sin \theta} \delta(\phi - \varphi_l) \delta(\theta - \theta_l). \tag{6}$$

Using the PW series expansion in Eq. (3), the integral in (5) then becomes the sum over all L PWs with the expression for the pressure

$$p(r, \phi, z) = \sum_{l=1}^{L} b_l \underbrace{\sum_{n=-\infty}^{\infty} R_n(k_r^{(l)} r) e^{-ik_z^{(l)} z} \Phi_n^*(\varphi_l) \Phi_n(\phi)}_{e^{i\mathbf{k}_l \cdot \mathbf{r}}}, \tag{7}$$

where the notation

$$k_r^{(l)} = k \sin \theta_l \tag{8}$$

$$k_z^{(l)} = k \cos \theta_l \tag{9}$$

$$R_n(x) = 2\pi i^{-n} J_n(x) \tag{10}$$

and

$$\Phi_n(\phi) = \frac{e^{in\phi}}{\sqrt{2\pi}} \tag{11}$$

was introduced for reasons of brevity and reasoning in subsequent sections. $R_n(x)$ and $\Phi_n(\phi)$ are commonly referred to as radial functions and angular basis functions (or modes), respectively [4], [15], [16].

Equation (7) is used to describe the measurement signals for a circular microphone array. In order to test the robustness of the FRI method against non-plane wave components in the sound field, the expression in Eq. (7) can be complemented by adding the series expansion of the Green function for the pressure field of a Point Source (PS) (see Appendix B).

III. CIRCULAR MICROPHONE ARRAY MEASUREMENT

Using the developed model, it is now possible to provide expressions for the signals obtained from circular microphone array measurements. This is a well-studied field that involves the spatial sampling of a sound field, with the known limitations associated to this process. This section introduces only the most relevant aspects of the technique and the interested reader is referred to the literature [4], [15], [16].

A. General PW Sound Field Measurement

Let a circular microphone array sense the pressure of the wave field continuously in the plane $z = z_M$ on the circle

$$A := \{\mathbf{r} : r = \xi, \phi \in [0, 2\pi], z = z_M\}, \quad (12)$$

depicted in Fig. 1. The pressure measured on A can then be expressed as a Fourier Series [1], [16]

$$p(\xi, \phi, z_M) = \sum_{n=-\infty}^{\infty} p_n \Phi_n(\phi) \quad (13)$$

with the coefficients p_n uniquely defining the sensed pressure. Multiplying both sides of (7) with $\Phi_{n'}^*(\phi)$ and integrating over A w.r.t. ϕ yields

$$p_n = \sum_{l=1}^{L} b_l R_n(k_r^{(l)} \xi) e^{-ik_z^{(l)} z_M} \Phi_n^*(\varphi_l). \quad (14)$$

In practice, the integration over A can only be approximated based on spatial sampling of the pressure through a microphone array (see. Section III-E).

The definition of the factors $R_n(k_r^{(l)} r)$ in (10) differs for a microphone array with a rigid scattering body [1], [4], [17]. For an infinite cylindrical scatterer of radius ξ (see Fig. 1), these radial functions become a combination of Bessel functions, $J_n(\cdot)$, Hankel functions of the first kind, $H_n(\cdot)$, and their derivatives:

$$R_n(k_r^{(l)} r) = \frac{2\pi}{i^n} \left(J_n(k_r^{(l)} r) - \frac{J_n'(k_r^{(l)} \xi)}{H_n'(k_r^{(l)} \xi)} H_n(k_r^{(l)} r) \right). \quad (15)$$

It can now be shown how the HD can be recovered from the pressure coefficients p_n.

B. HD Coefficients of PWs travelling parallel to $z = z_M$

In the case where all L PWs are travelling parallel to the plane $z = z_M$, i.e. $\{\theta_l\}_{l=1}^{L} = \frac{\pi}{2}$, the expression for the pressure coefficients in (14) can be further simplified to ($n = n'$)

$$p_n = R_n(k\xi) \sum_{l=1}^{L} b_l \Phi_n^*(\varphi_l) = R_n(k\xi) q_n. \quad (16)$$

The right hand side of the above equation indicates that the expression for p_n can be separated into a radial function, $R_n(k\xi)$, and the HD coefficients, q_n, describing the incident plane waves

$$q_n = \sum_{l=1}^{L} b_l \frac{e^{-in\varphi_l}}{\sqrt{2\pi}}. \quad (17)$$

These HD coefficients fully define the sound field with the corresponding HD obtained through Fourier synthesis

$$q(\phi) = \sum_{n=-\infty}^{\infty} q_n \frac{e^{in\phi}}{\sqrt{2\pi}} = \sum_{l=1}^{L} b_l \delta(\phi - \varphi_l), \quad (18)$$

where the dependency on θ is henceforth omitted for the sake of brevity (compare to Eq. (4)). An example of a HD with
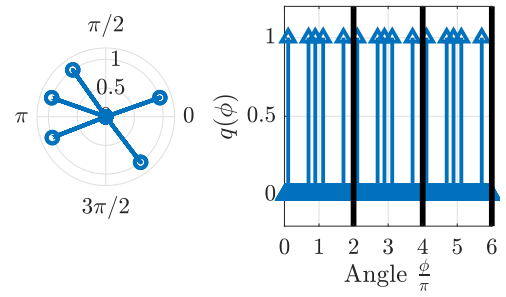

Fig. 2: HD $q(\phi)$ for $L = 5$ distinct plane waves, where the HD is periodic with $q(\phi) = q(\phi + u2\pi), \forall u \in \mathbb{Z}$.

$L = 5$ distinct plane waves of unit magnitude is shown in Fig. 2 and it can be seen that $q(\phi)$ must be periodic with a period of 2π. Such a HD is fully defined by the set of tupels $\{(\varphi_l, b_l)\}_{l=1}^{L}$. The extraction of the q_n from the array signals is crucial for the application of the FRI DOA estimation method.

At a later point it will be shown that a wave field specified by the HD coefficients in (17) is ideal for the application of the FRI DOA estimation method with circular arrays. The next two subsections present expressions for the HD coefficients of two source types that are likely to corrupt the performance of the DOA estimation method in combination with the measurement aperture under consideration.

C. HD Coefficients of a Single PW with $\theta_l \neq \frac{\pi}{2}$

For a single plane wave impinging on the measurement aperture at an angle $\theta_l \neq \frac{\pi}{2}$, the corresponding pressure coefficients are given through (14), yielding

$$p_n = b_l R_n(k_r^{(l)} \xi) e^{-ik_z^{(l)} z_M} \Phi_n^*(\varphi_l) \quad (19)$$

A factor separation as shown in Equation (16) of the result in (19) yields HD coefficients of the form

$$q_n = b_l \underbrace{\frac{R_n(k_r^{(l)} \xi)}{R_n(k\xi)} e^{-ik_z^{(l)} z_M}}_{\Psi_n(k, k_r^{(l)}, \xi)} \frac{e^{-in\varphi_l}}{\sqrt{2\pi}} \quad (20)$$

for the sound field of the single PW with $\theta_l \neq \frac{\pi}{2}$. The main difference to Eq. (17) is the additional factor $\Psi_n(k, k_r^{(l)}, \xi)$.

D. HD Coefficients of a Single Point Source in the Array Plane

The derivation of the HD coefficients for the sound field of a single PS located at $\mathbf{r}' = (r', \phi', z_M)^T$

$$q_n = b' \underbrace{\frac{i^{n+1}}{8\pi} \frac{\int_{-\infty}^{\infty} R_n(k_r \xi) H_n^{(1)}(k_r r') dk_z}{R_n(k\xi)}}_{\zeta_n(r')} \frac{e^{-in\phi'}}{\sqrt{2\pi}} \quad (21)$$

is given in Appendix C. Similar to the source type discussed in the previous subsection, the key difference between Eqs. (17) and (21) is an additional factor, namely $\zeta_n(r')$.

The introduced additional factors $\Psi_n(k, k_r^{(l)}, \xi)$ and $\zeta_n(r')$ are later shown to negatively affect a DOA estimation using the FRI method in combination with circular measurement apertures (see Sections V-D and V-E, respectively, and also Sec. IV-F).

E. Practical Limitations of Microphone Arrays

In practice, microphone arrays can only spatially sample the pressure on A and therefore cannot recover all coefficients q_n needed to fully reconstruct the HD. However, a uniformly distributed circular microphone array with M sensors can recover the coefficients q_n for $|n| \leq N$, where $N = \lfloor \frac{M-1}{2} \rfloor$, if no angular modes $\Phi_n(\phi)$ with $n > N$ contribute to the pressure field on A (i.e. no spatial aliasing). This is the case (to a good approximation) for all frequencies that satisfy $N \geq k\xi$ [16]–[18]. At higher frequencies, the recovered coefficients q_n are corrupted by spatial aliasing [16], [18]. It shall for now be assumed that the coefficients $q_n, |n| \leq N$ can be recovered perfectly from the array measurement, and that the frequency range in question ensures that no spatial aliasing occurs.

At this point, the models for the description and analysis of the proposed method are all defined.

IV. FINITE RATE OF INNOVATION ANALYSIS

The form of the HD given in (18) represents a set of distinct Dirac-pulses, for which the FRI theory was originally developed [10], [11]. Vetterli et al. originally addressed the problem of sampling non-bandlimited time domain signals, e.g. such as distinct occurrences of Dirac-pulses. The key difference between their and other previous work, and that presented here is that the observed signal, i.e. the HD, is not in the time domain but in the spatial domain. This section serves to put the theory into the perspective of a DOA estimation problem.

A. The System-Inherent Sampling Kernel

While the method proposed in [10] requires a sampling kernel (i.e. a lowpass filter) for the observed time domain signal, it must be said that this is not always required when sampling a pressure field in the spatial domain. The following paragraph shall briefly explain the reason that supports this statement.

It is known that modes of orders higher than a frequency dependent threshold N_A hardly contribute to the pressure observed on the measurement aperture A [4], [19]. Their corresponding radial functions $R_n(kr)$ act as a system-inherent lowpass filter, so that $p_n \approx 0, |n| > N_A$. Therefore, at least for frequencies at which the array does not suffer from spatial aliasing, the coefficients $\{q_n\}_{n=-N}^{N}$ can be recovered practically uncompromised, so that a spatially lowpass filtered, i.e. order limited, version of the HD can be reconstructed from them.

B. The Rate of Innovation of the HD

The HD coefficients q_n of the form given by (17) describe a (periodic) HD with finite rate of innovation. Every PW is defined by its DOA, φ_l, and its complex amplitude, b_l. In other words, every PW is an entity with two 'innovations' for the HD. It follows that a HD with L PWs incorporates $2L$ innovations. Due to its 2π periodicity, the HD $q(\phi)$ has the *rate of innovation* [10]

$$\rho = \frac{2L}{2\pi} = \frac{L}{\pi}$$

that is *finite* for $L < \infty$ (see also Fig. 2). The value of ρ or rather the number of innovations within the period, $2L$, determines how many consecutive coefficients q_n are needed to apply the Annihilating Filter (AF) method [10] [11] that provides the estimates for the DOAs, φ_l. The exact relation is derived in the following subsection.

C. Recovering the DOA Information

The exact procedure for the DOA estimation with ideal coefficients q_n is described in [1]. The key step of the method is the identification of an AF defined by the sequence $\{a_k\}_{k=0}^{K}$, $K \leq N$, that, when convolved with the sequence of HD coefficients q_n, $n \in \mathbb{Z}$, yields a zero result (hence "annihilating"). It will become clear at the end of this subsection that the the DOAs of the L PWs can be extracted from the AF. In theory, a perfect AF can be calculated [1], [10]. In practice, however, acoustical array data used to obtain the HD coefficients may be compromised due to imperfections of the measurement system (e.g. manufacturing tolerances, sensor mismatches, etc.), model mismatches (e.g. point source vs. plane wave) and sensor noise [19], [20], which ultimately affects the estimation of the AF.

To make the estimator more robust against such, Blu et al. [11] proposed the use of a 'total least-squares approach', which was straightforwardly adopted for the proposed method described hereunder.

The global objective of the DOA estimator mechanism is to find an AF that satisfies

$$q_n * a_k = 0, \tag{22}$$

denoting the discrete convolution of the AF with the sequence of HD coefficients, i.e.

$$\sum_{k=0}^{K} a_k q_{n-k} = 0, \ (-N + K) \leq n \leq N. \tag{23}$$

The above equation defines a system of $2N + 1 - K$ equations that can be expressed in the algebraic form

$$\mathbf{Qa} = 0, \tag{24}$$

with the Toeplitz matrix

$$\mathbf{Q} = \underbrace{\begin{bmatrix} q_{-N+K} & q_{-N+K+1} & \cdots & q_{-N} \\ q_{-N+K+1} & q_{-N+K} & \cdots & q_{-N+1} \\ \vdots & \ddots & \ddots & \vdots \\ q_{-N+2K} & \ddots & q_{-N+K+1} & q_{-N+K} \\ \vdots & \ddots & \ddots & \vdots \\ q_N & q_{N-1} & \cdots & q_{N-K} \end{bmatrix}}_{(2N+1-K) \times (K+1)} \tag{25}$$

and $\mathbf{a} = [a_0, a_1, \ldots, a_K]^T$. For an ideal sound field with no measurement noise or other perturbations, the system is overdetermined for $N > K$ and the rank of $\mathbf{Q}$ cannot exceed L. This is due to the nature of the coefficients q_n. They can only lead to $K - L + 1$ independent, non-trivial solutions of the AF $\mathbf{a}$, one of which has zeros for the last $K - L$ entries [11]. It follows that

$$\text{rank}(\mathbf{Q}) = L, \qquad (26)$$

i.e. the number of PWs in the sound field can be identified from the rank of the Toeplitz matrix $\mathbf{Q}$.

When the coefficients q_n are corrupted by noise, a non-trivial solution to Eqs. (22) and (24) may not exist. As an alternative, Blu et al. [11] proposed to identify a Total Least-Squares (TLS) estimate for the AF $\mathbf{a}$ by solving the optimisation problem

$$\tilde{\mathbf{a}} = \underset{\mathbf{a}}{\text{argmin}} \|\mathbf{Q}\mathbf{a}\|_2^2, \; \|\mathbf{a}\|_2 = 1. \qquad (27)$$

With the constraint for $\mathbf{a}$, the solution to the above problem can be found from a Singular Value Decomposition (SVD) of $\mathbf{Q} = \mathbf{U}\mathbf{S}\mathbf{V}^H$. This provides the right-singular vectors $\{\mathbf{v}_k\}_{k=0}^{L}$ (i.e. the columns of $\mathbf{V}$) that all satisfy $\|\mathbf{v}_k\|_2 = 1$. The right-singular vector $\mathbf{v}_{\min}$ corresponding to the smallest non-zero singular value $\sigma_{\min}$ then provides a solution to the total least-squares problem that yields

$$\mathbf{Q}\tilde{\mathbf{a}} \approx 0, \; \text{with} \; \|\tilde{\mathbf{a}}\|_2 = 1, \qquad (28)$$

where $\tilde{\mathbf{a}} = \mathbf{v}_{\min}$. Unlike the AF presented in [1], the one obtained from the TLS method can have a leading factor ψ [11], yielding

$$\tilde{a}_n = \psi(\delta_n + \sum_{k=1}^{K} \alpha_k \delta_{k-n}).$$

Given that filter coefficients $\tilde{a}_n$ are a sequence (mathematically identical to a discrete time-domain signal), it is possible to calculate its z-transform, which can be expressed in the form

$$\tilde{A}(z) = \psi \prod_{k=1}^{K} (1 - u_k z^{-1}), \; z = Ae^{-i\phi}, \; A \in \mathbb{R}. \qquad (29)$$

It is known from [1] that the roots $\{u_k\}_{k=1}^{K}$ of the characteristic polynomial of $\tilde{A}(z)$ are of the form $u_k = e^{-i\varphi_k}$. Thus, the DOA can be calculated through

$$\varphi_k = i \ln(u_k), \forall k = 1 \ldots K, \qquad (30)$$

where $\ln(\cdot)$ denotes the principle value of the complex logarithm. The full process for the DOA estimation is summarised below.

DOA Estimation Procedure

1) Choose a value $K \leq N$. If L is known, $K = L$,
2) construct $\mathbf{Q}$ according to Eq. (25),
3) solve the TLS problem in (27) to estimate the AF $\tilde{\mathbf{a}}$,
4) calculate the roots, $\{u_k\}_{k=1}^{K}$, of the z-transform of $\tilde{\mathbf{a}}$,
5) convert the K roots into DOA estimates using Eq. (30),

6) use $\{\tilde{\varphi}_k\}_{k=1}^{K}$ to calculate the complex amplitudes, $\{\tilde{b}_k\}_{k=1}^{K}$, by solving the linear equation system defined by (17) for a set of coefficients q_n.

With respect to this work, the described procedure is the most general approach to the FRI DOA estimation method, yielding estimates for exactly K PWs, even if $L < K$ or $L > N$. This uncertainty about the number of PWs to be identified, K, is clearly not ideal and is further investigated in Sections IV-E and V-A.

Note that there are no restrictions on the complex values b_l, which uniquely define the source signals, hence the method works for both incoherent and coherent sources. This is an important distinction from other methods, e.g. MUSIC, ESPRIT, EB-ESPRIT.

For particularly noisy data, additional Cadzow-Denoising can be applied to the data before the AF method [11] to fit the data closer to the model. This was done for all applications of the DOA estimation method presented in this work, where the steps of the Cadzow-Denoising were iterated 20 times.

With the method fully described, it can now be tested by means of simulation. To make the latter's sound field data a better representative of practical measurements, the model developed in Section III is extended by a suitable noise model in the following.

D. A Measurement Noise Model

The applied noise model is a spatial transform of the transducer noise model for microphone arrays used in [16]. A pressure measurement on a circular aperture around the coordinate origin would lead to a noisy observation

$$\tilde{p}(\phi) = p(\phi) + d(\phi),$$

where d denotes the measurement noise. In that case, the HD coefficients become (compare Eq. (16))

$$\tilde{p}_n = p_n + \sigma_\epsilon \frac{e^{i2\pi\mu}}{\sqrt{2\pi}} = R_n(k\xi)q_n + \sigma_\epsilon \frac{e^{i2\pi\mu}}{\sqrt{2\pi}}. \qquad (31)$$

Division by $R_n(k\xi)$ provides the result for the HD coefficients obtained from a simulated circular microphone array measurement with noise,

$$\tilde{q}_n = q_n + \epsilon_n, \; \forall |n| \leq N. \qquad (32)$$

Each of the ideal coefficients q_n can be corrupted by the additive component

$$\epsilon_n = \frac{\sigma_\epsilon}{R_n(k\xi)} \frac{e^{i2\pi\mu}}{\sqrt{2\pi}} \qquad (33)$$

that reflects the statistical properties of the measurement noise in microphone arrays. σ_ϵ denotes the standard deviation of the noise in the observed pressure signal and μ is a random variable that is *uniformly* distributed on the interval $[0, 1]$.

The expected value for the magnitude of ϵ_n, $E\{|\epsilon_n|\}$, can be useful to estimate the number of PWs for the case that $L < N$. The model assumes that the noise distribution in the pressure domain is Gaussian, therefore the expected value of
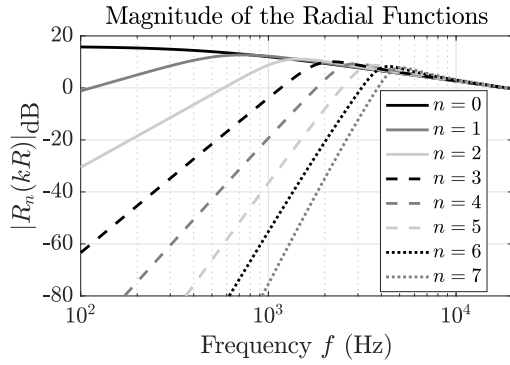
Fig. 3: Radial functions for a microphone array with radius $\xi = 7.5\,\text{cm}$ and $f_c \approx 5.1\,\text{kHz}$ ($N = 7$, $c = 343\,\frac{\text{m}}{\text{s}}$).

the magnitude of the measurement noise in the pressure domain, $|d(\phi)|$, must be

$$E\{|d(\phi)|\} = \frac{\sigma_\epsilon}{\sqrt{2\pi}}. \tag{34}$$

To obtain a useful expected value for the measurement noise in the HD coefficient domain, it is necessary to consider the influence of the radial functions $R_n(k\xi)$. From the literature on microphone arrays, it is known that the magnitude of the radial functions as given by (15) for $k_r^{(l)} = k$ differs significantly for different orders $|n| \leq N$ at frequencies lower than

$$f_c = \frac{Nc}{2\pi\xi}$$

[4], [16], [17]. This is rather unfavourable, since this property of the radial functions would affect the coefficients ϵ_n as a consequence of Eq. (33). The components of q_n that are associated with the individual PWs are constant in magnitude across the range $|n| \leq N$. Hence it would be difficult to define a clear 'noise floor' if the expected value for the noise magnitude is a function of n. However, at frequencies higher than f_c, the magnitude of the radial functions can be approximated by $|R_0(k\xi)|$, which can be seen in Fig. 3. It follows that the expected value for the magnitude of ϵ_n can be approximated by

$$E\{|\epsilon_n|\} \approx \frac{\sigma_\epsilon}{\sqrt{2\pi}|R_0(k\xi)|}, \quad n \leq N, \tag{35}$$

for frequencies higher than $f_c \approx 5.1\,\text{kHz}$ with $\xi = 7.5\,\text{cm}$. This value provides an estimate of the noise floor.

To acquire a value of σ_ϵ for a specific array, which may also be frequency dependent [16], one could simply perform a measurement without active sources in an anechoic environment.

E. Choosing K with Noisy Coefficients q_n

If the number L of PWs is known a priori, then the value for K should be simply chosen to be L. Otherwise it can always be chosen to be the maximum number of DOAs that can be estimated from the available array data, i.e. $K = N$. Three

cases can then be distinguished for a spatially stationary PW distribution when the individual PWs' energy exceeds that of the background noise:

1) $\mathbf{L < N}$: The DOAs corresponding to the L largest values of $|\tilde{b}_k|$ will be consistent irrespective of the point in time, where the $N - L$ smallest estimates must be expected to vary with every FFT frame due to noise.
2) $\mathbf{L = N}$: All DOA estimates describe an individual PW.
3) $\mathbf{L > N}$: The method provides N least-squares DOA estimates.

In the case of $L < N$, the number of PWs, L, can be estimated from the number of magnitudes $|\tilde{b}_k|$ that exceed the estimate of the noise floor given by Eq. (35). This is however only possible for frequencies higher than f_c.

In the case of $L > N$, the rate of innovation is too high for the available range of coefficients $|n| \leq N$, i.e. the sampling of the sound field is insufficient. The estimator can only provide $K = N$ DOA estimates, which will be shown to become increasingly inaccurate as L approaches N. One explanation for this is that the TLS solution for the AF 'groups' the $L - N$ surplus PWs together with the closest neighbours within the set of the remaining L PWs. This is more thoroughly investigated in Section V-A.

F. Interpretation of the Annihilating Filter

The z-transform of the sequence of HD coefficients

$$Q(z) = \sum_{n=-\infty}^{\infty} q_n z^{-n},$$

evaluated on the unit circle, $z = e^{-i\phi}$, is the HD weighted with $\sqrt{2\pi}$, so that

$$Q(z) = \sqrt{2\pi}q(\phi), \quad |z| = 1.$$

The z-transform of the AF in Eq. (29), evaluated on the unit circle, provides a complex-valued directivity pattern, the annihilating pattern $a(\phi)$ (depicted in Fig. 4), that has nulls in the directions φ_l, corresponding to the source directions. Hence, finding the DOAs is equivalent to finding the zeros of $a(\phi)$.

In this light, it can now be understood why both a PS and a PW with $\theta \neq \frac{\pi}{2}$ are likely to compromise the performance of the FRI method. The sequences of the HD coefficients given in (20) and (21) both describe an amplitude modulation of the carrier sequence $e^{-in\phi'}/\sqrt{2\pi}$ with the sequence $\Psi(k, k_r^{(l)}, \xi)$ and $\zeta_n(r')$, respectively. From the theory of the spectra of AM signals, it follows that the corresponding HD coefficient sequences contain more than one distinct component in the z-domain. The difference between the $Q(z)$ of a PW and that of a PS is conceptually depicted in Fig. 4, along with a possible corresponding annihilating pattern (dashed line) with its nulls located at $\phi = \varphi_{\text{PW}}$ and $\phi = \varphi_{\text{PS}}$. While the PW is responsible for just a single component, the PS has a 'carrier' component that refers to the position of the PS **and** modulation components resulting from the term $\zeta_n(r')$ described in Eq. (21).
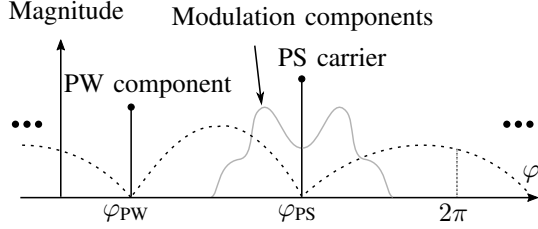
Fig. 4: Conceptual z-domain spectrum ($z = e^{-i\varphi}$) of the HD coefficients for the sound field of a PW and a PS, and the magnitude response of the annihilating pattern $a(\phi)$ (dashed line).

The same can be shown for the PW with $\theta \neq \frac{\pi}{2}$, where the term $\Psi_n(k, k_r^{(l)}, \xi)$ is responsible for the modulation components.

Nevertheless, depending on the modulating sequence, the carrier sequence $e^{-in\phi'}/\sqrt{2\pi}$ may still be the dominant component. In that case, the FRI method may still be able to detect the DOA of these types of sources. The TLS solution would then provide a filter (indicated as a dashed line in Fig. 4) that attempts to suppress the strongest components in the sequence, i.e. those of the individual carriers. The DOAs can then be identified as usual from the roots of the AF's z-transform (see Section IV-C). However, it must be expected that the estimation accuracy is adversely affected for these types of sources, e.g. if the different sources significantly differ in magnitude or for very close PSs. This is confirmed by means of simulation results given in Section V.

V. SIMULATIONS

For all simulations discussed hereunder, the DOA estimation method was set to identify $K = N$ plane waves (unless mentioned otherwise) by choosing the matrix $\mathbf{Q}$ accordingly. When the number of plane waves L is known a priori, only those DOA estimates associated with the largest magnitudes $|\tilde{b}_k|$ were taken into account. For those cases when $L > N$, only N DOA estimates are provided due to the inherent limitation of the method. Note that in this section, a DOA estimate $\tilde{\varphi}_k$ refers to a value for φ_l.

All sound fields were calculated directly in the cylindrical harmonics domain for $f = 1\,\text{kHz}$, using the equations presented in the preceding sections. That way the estimation method's accuracy and robustness can be studied without the influence of effects that stem from the microphone array (e.g. spatial aliasing, ill-conditioning etc.).
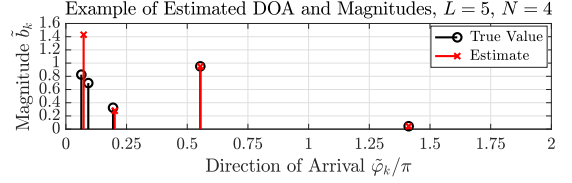
Before looking more closely at the simulation performance for the unrestricted case, $L \leq N$, the estimation behaviour for $L > N$ is investigated by means of a number of selected samples.

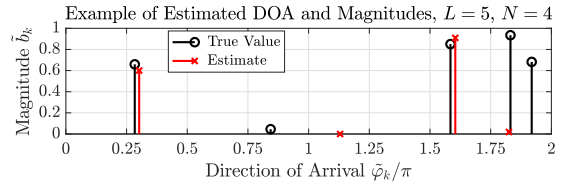### A. Estimation Behaviour for $L > N$

The effects occurring when estimating $K = N$ directions of arrival from the HD coefficients of a sound field with $L > N$
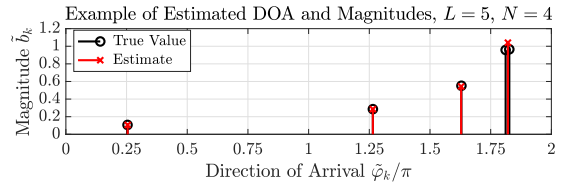
(a) Two PWs grouped (Example I)

(b) Two PWs grouped (Example II)

(c) Interpolation

Fig. 5: Results of different FRI estimation performances for $L = 5$ and $N = K = 4$.

PWs were briefly mentioned in Section IV-E. Figure 5 shows a selection of estimation results for $L = 5$ and $N = K = 4$.

Figure 5a shows an example of the aforementioned grouping behaviour. The estimates of the plane waves with a DOA at around 0.25, 1.25 and 1.625 (on the abscissa) are very accurate, as are the estimated magnitudes. Only for the second estimate on the right it can be noticed that its magnitude is slightly lower than the true value. The estimate beyond 1.75 on the abscissa 'groups' two true PWs with very small difference in their individual DOA. This is likely to explain why its estimated magnitude is larger than that of the two true plane waves.

A similar behaviour of the method can be observed from the result shown in Fig. 5b. The estimates of magnitude and DOA beyond 0.5 on the abscissa show good accuracy. The leftmost estimate 'groups' two plane waves, not quite as closely aligned as those in the previous example, and shows a significantly higher magnitude. In this case, the estimate to the right of 0.25 not only shows a magnitude mismatch but also a DOA estimate deviation from the true value.

The result in Fig. 5c is very different from the others. A grouping behaviour cannot be detected. Only three of the DOA estimates are close to true plane waves, but the magnitude estimate of one of them is significantly low. The fourth estimate is significantly inaccurate.

While there are cases where some estimates may still be of value, it becomes clear that the estimation accuracy decreases when $L > N$. That makes the method unreliable in such cases. For predominantly isotropic sound fields with $L < N$ distinct sources of sufficiently high level, the method may prove more

robust, as will be seen in the next subsection.

B. The DOA Estimation Error (DEE)

As a measure to evaluate the quality of the estimated directions of arrival, $\{\tilde{\varphi}_k\}_{k=1}^{K}$, the DOA Estimation Error (DEE) in percent is used in this work. This is defined for a single plane PW by

$$\text{DEE} = (|\tilde{\varphi}_k - \varphi_k| \mod \pi) \cdot 100\,\%, \qquad (36)$$

where $\mod$ denotes the modulo operator. A DEE of $100\,\%$ corresponds to the maximum difference between estimate and true value of $\pi \equiv 180°$.

C. Mean and Median DEE

In order to assess the estimation accuracy of the FRI DOA estimation method, a simulation study was conducted for different Signal-to-Noise Ratio (SNR) values and numbers of PWs. For each combination of SNR and L, 10,000 samples of randomised PW fields were simulated and averaged. Note that the DEE does not depend on the PW magnitude $|b_l|$, so the results shown below may be different if the DEE was weighted by $|b_l|$.

The clean HD coefficients were calculated from Eq. (17). The complex amplitudes b_l of the individual plane waves were calculated from

$$b_l = (0.01 + 0.99\nu)e^{-i2\pi\mu}, \qquad (37)$$

where ν, μ are uniformly distributed on the interval $[0, 1]$. The maximum difference between the magnitudes of the individual plane waves is thus $40\,\text{dB}$. The corresponding DOA parameters, φ_l, are uniformly distributed on the interval $[0, 2\pi]$.

The noise was generated directly in the HD coefficient domain using

$$\epsilon_n = \sigma_\epsilon \xi e^{i2\pi\mu},$$

where ξ is a normally distributed random variable and μ is uniformly distributed on the interval $[0, 1]$. The value σ_ϵ was chosen so that the noisy coefficients $\tilde{q}_n = q_n + \epsilon_n$ have a specified SNR given by

$$\text{SNR} = 10 \log_{10}\left(\frac{\sum_{|n| \leq N} |q_n|^2}{\sum_{|n| \leq N} |\epsilon_n|^2}\right). \qquad (38)$$

Figure 6 shows the results of a simulation study for the mean/median DEE in percent. The algorithm was applied to the noisy coefficients $\tilde{q}_n = q_n + \epsilon_n$ within the range $|n| \leq 7$. This corresponds to the HD coefficient data that can be obtained from a microphone array measurement with $M = 15$ sensors. The maximum number of plane waves whose DOA can be estimated is $K = N = 7$. For the conducted simulations, the algorithm parameter for the number of plane waves was chosen so that $K = \min(L, N)$. Note that with the given dynamic range of the PW magnitudes specified in Eq. (37), some of them may be covered by the noise floor.

The result in Fig. 6a shows that the mean accuracy of the DOA estimates increases with the SNR and decreases with the number of sources. At 0 dB SNR the values for more than one
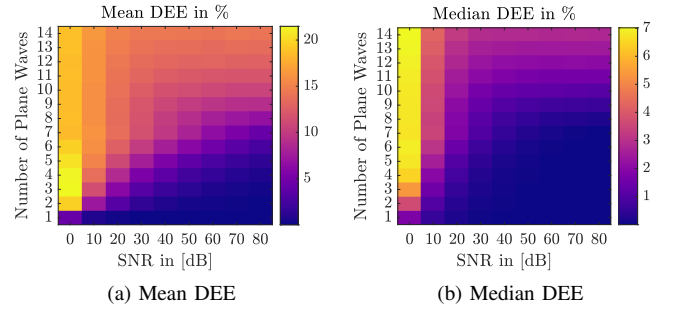

(a) Mean DEE   (b) Median DEE

Fig. 6: Performance of the FRI estimator averaged over 10,000 samples for different combinations of SNR and L.

PW range around a DEE of $20\,\%$, with a maximum at $L = 3$ PWs . At high SNR values, the DEE increases significantly for more than $L = 7$ plane waves, as a result of the limitation due to the available range of coefficients q_n.

The results for the median DEE given in Fig. 6b are similar w.r.t. the overall trend, however the maximum values are significantly lower than those shown in the mean results. This suggests that, for a given combination of SNR and value L, $50\,\%$ of the estimates' DEE are lower than the displayed value. Accordingly, the other $50\,\%$ of estimates may suffer from 'outliers' that are significantly wrong compared to the true value. These outliers cause the mean to be significantly higher than the median DEE values.

D. Sound Fields with PWs not travelling parallel to the Measurement Plane

The coefficients describing the sound field of a PW that is not travelling parallel to the measurement plane $z = z_M$ is given in Eq. (20). Such a plane wave is henceforth referred to as an Out-Of-Plane Plane Wave (OOP-PW). It was postulated in Section IV-F that the effect of the additional factor $\Psi_n(k, k_r^{(l)}, \xi)$ in the specific form of the coefficients q_n would have a negative effect on the accuracy of the DOA estimates for more than one PW. It was also argued that the FRI method may provide an adequate estimate of the DOA for a single such plane wave regardless.

1) A Single OOP-PW: This subsection investigates the influence of the elevation angle θ_l of a single OOP-PW on the DOA estimation performance. Figure 7 shows the mean DEE for a single OOP-PW for randomly chosen DOA component φ_l. The elevation angle associated with the OOP-PW was tested for $0.01\pi \leq \theta_l \leq 0.99\pi$. The DOA estimation would naturally fail for values $\theta_l = 0$ and $\theta_l = \pi$, which were thus not tested. The estimates were obtained from the coefficients $\{q_n\}_{|n| \leq 1}$, which suffice to estimate the DOA of $N = 1$ plane wave. The average result is below $1.6 \times 10^{-14}\,\%$.

The same simulation was repeated for a greater range of coefficients, $\{q_n\}_{|n| \leq 5}$, and the result is shown in Fig. 8. While for some values of θ_l there appears to be no difference, the mean DEE is significantly higher for angles $0.07\pi \leq \theta_l \leq 0.35\pi$ and $0.65\pi \leq \theta_l \leq 0.93\pi$. This suggests that the estimation
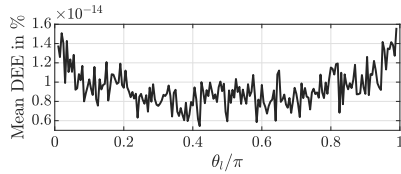
Fig. 7: Mean DEE for single OOP-PW for $0.01\pi \leq \theta_l \leq 0.99\pi$, averaged over 100 samples with randomly chosen φ_l, $N = 1$.
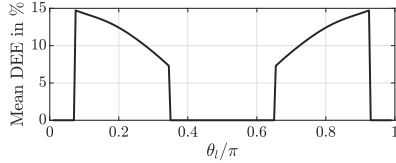

Fig. 8: Mean DEE for single OOP-PW for $0.01\pi \leq \theta_l \leq 0.99\pi$, averaged over 100 samples with randomly chosen DOA, as obtained with a greater coefficient range ($N = 5$).

corrupting effect of the additional factor $\Psi_n(k, k_r^{(l)}, \xi)$ increases with N.

2) One PW and one OOP-PW: The DEE performance for two PWs, where one is an OOP-PW, is shown in Fig. 9. Following the findings in Section V-D1, the coefficient range
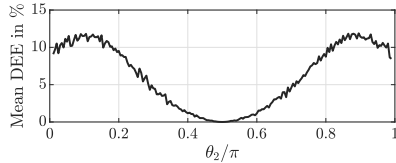

Fig. 9: Mean DEE for one OOP-PW with $\theta_2 \in [0.01\pi \ldots 0.99\pi]$ and one with $\theta_1 = \frac{\pi}{2}$, averaged over 100 samples with randomly chosen φ_1 and φ_2, $N = 2$.

was chosen to $|n| \leq 2 = L$ so that both PWs can be identified but the coefficient range is kept as low as possible. Both PWs had the same magnitude. One plane wave is impinging at $\theta_1 = \frac{\pi}{2}$ and the other is simulated at different angles θ_2. The two parameters φ_1 and φ_2 were randomly chosen for every iteration. It can be observed that the accuracy of the DOA estimates decreases up to a DEE of approximately $12\,\%$ as the parameter θ_2 deviates from $\frac{\pi}{2}$.

E. Sound Fields with Point Sources

As with the plane waves discussed in Section V-D, the specific form of the coefficients q_n describing the sound field of a single PS in the plane of the microphone array were predicted to have a negative effect on the accuracy of the DOA estimates associated with the waves impinging from more than one PS. This is now investigated further with a set of simulations.

1) Estimation Accuracy for a Single PS: Figure 10 shows the mean DEE for the DOA estimates obtained over 100 iterations of the sound field with a randomly positioned (w.r.t. the angle ϕ') single PS as a function of the position parameter r'. The
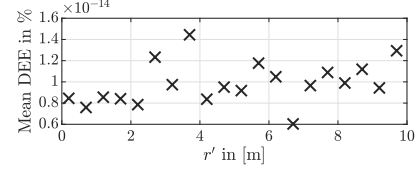

Fig. 10: Mean DEE for a randomly positioned PS as a function of r', averaged over 100 samples per evaluation radius, $N = 1$.

set of coefficients $\{q_n\}_{|n| \leq 1}$ were calculated from Eq. (21), i.e. the point sources lie in the same plane as the microphones. The simulated array radius was $a = 0.1\,\mathrm{m}$ and the FRI algorithm was set to detect one source only ($K = 1$). It can be seen that a very high estimation accuracy was obtained that supports the reasoning given in Section IV-F. Note that the simulation did not include any noise.

The influence of the compromising factor $\zeta_n(r_l)$ in the coefficients q_n was found to increase with the range of coefficients used for the estimation, as it was observed for the OOP-PW in the previous subsection, thus negatively affecting the performance.

2) Estimation Accuracy for Two Point Sources: A similar simulation was conducted for two PSs, where the coefficients are given through Eq. (21) and superposition as

$$q_n = b_1 \frac{e^{-in\varphi_1}}{\sqrt{2\pi}} \zeta_n(r_1) + b_2 \frac{e^{-in\varphi_2}}{\sqrt{2\pi}} \zeta_n(r_2).$$

Three different scenarios were investigated using the parameters $b_1 = b_2 = 1$, $r_2 = 0.2 \ldots 9.7\,\mathrm{m}$ and:

1) First source located closely ($r_1 = 0.2\,\mathrm{m}$),
2) first source located distantly ($r_1 = 9.7\,\mathrm{m}$), and
3) both sources at the same distance, i.e. $r_1 = r_2$.

The FRI algorithm was set to detect two sources, i.e. $K = 2$, using the set of coefficients $\{q_n\}_{|n| \leq 2}$. The results are shown in Fig. 11. It can be observed that the average DEE has drastically increased just by adding a second PS. The severity however differs between the three scenarios.

In the first scenario, the results in Fig. 11a show that for all radii $r_2 > 1\,\mathrm{m}$ the mean DEE is higher than $20\,\%$. When $r_2 < 1\,\mathrm{m}$ then the accuracy for both sources remains below $10\,\%$.

In the second scenario, the results in Fig. 11b show that the DEE is larger than $10\,\%$ for radii smaller than $1\,\mathrm{m}$. For larger radii, the DEE drops significantly below $10\,\%$, reaching values less than $1\,\%$ for $r_2 > 3.2\,\mathrm{m}$. The mean DEE appears to have a decreasing trend as the sources move to larger radii. This can be expected since the wave field of a PS that is far away approximates that of a plane wave.

As in the second scenario, the results of the third scenario also describe a decreasing trend for the DEE as shown in
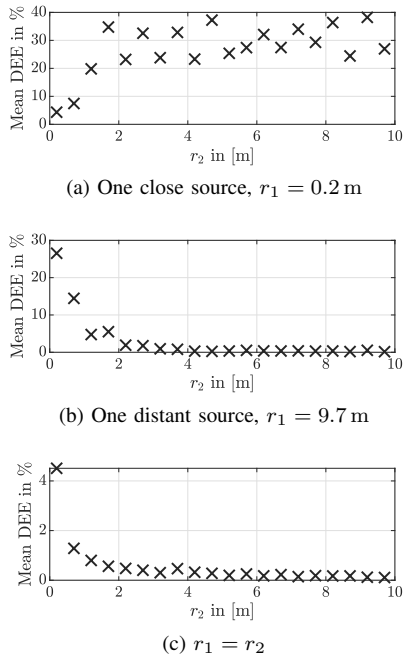
(a) One close source, $r_1 = 0.2\,\mathrm{m}$

(b) One distant source, $r_1 = 9.7\,\mathrm{m}$

(c) $r_1 = r_2$

Fig. 11: Mean DEE for two randomly positioned PSs as a function of r_2, averaged over 100 samples, $N = 2$.

Fig. 11c. At $r_1 = r_2 = 0.2\,\mathrm{m}$ the mean DEE is already clearly below $10\,\%$ and dropping further. The DEE drops significantly below $1\,\%$ for $r_1 = r_2 > 2\,\mathrm{m}$.

F. Discussion of Simulation Results

The simulation study of the DOA estimation performance for different values of SNR and numbers of PWs confirm the degrading influence of noise on the estimation accuracy. The number of sources in the field also affects the accuracy. As stated by Blu et al. [11], using a surplus of coefficients (i.e. $N > L$) increases the robustness against noise for an observation, i.e. the HD, with distinct Dirac pulses (compare Section V-F6). The significant difference between the mean and the median results suggest that $50\,\%$ of the estimates have an accuracy with less than $7\,\%$ error, while the other $50\,\%$ may suffer from outliers with DEE values larger than $20\,\%$.

The relatively high DEE values would be lower if the difference in magnitude between the individual PWs were smaller (it was simulated as $40\,\mathrm{dB}$). Due to the large differences in magnitude, some of the plane waves may either vanish below the noise floor or the estimates would be degraded due to a low SNR for that specific source.

1) Single OOP-PW: As it was argued in Section IV-F, the modulating sequence $\Psi_n(k, k_r^{(l)}, \xi)$ had little effect on the DOA estimation accuracy, but only when using a small range of coefficients, $|n| \leq 1$. When increasing the latter to $|n| \leq 5$, the performance degraded significantly for a distinct range of angles. This suggests that the overall modulation due to the factor $\Psi_n(k, k_r^{(l)}, \xi)$ is so strong that it overpowers the carrier

in the extended sequence $q_{-5}, \ldots, q_5$ and thus leads to a high DEE (see also Section V-F6).

2) One PW and one OOP-PW: In light of the theoretical considerations made in Section IV-F, the severe loss of accuracy with the obtained estimates suggests that the modulating sequence $\Psi_n(k, k_r^{(2)}, \xi)$ is very strong in magnitude and its contribution cloaks both carriers $e^{-in\varphi_1}/\sqrt{2\pi}$ and $e^{-in\varphi_2}/\sqrt{2\pi}$, and strongly alters the roots of the AF. The result is incorrect DOA estimates.

3) Single Point Source: Similar to the results for the single OOP-PW, the estimation performance for the single point source have shown results of high accuracy. Following the performance dependency on the range of coefficients q_n found for the single OOP-PW, it was found that the same applies for the field of a single point source, though the results were not presented for reasons of brevity. Accordingly, the modulation of the additional factor $\zeta_n(r')$ is also strong enough to affect the estimation accuracy when using a wider range of coefficients.

4) Two Point Sources: The scenario where the first PS is located at a small radius ($r_1 = 0.2\,\mathrm{m}$) and the second source is moved from $r_2 = 0.2\,\mathrm{m}$ to $r_2 = 9.7\,\mathrm{m}$ shows the worst performance among the results for two PSs. This suggests that the modulation component around the 'carrier' of the close PS (see Sec. IV-F) in the spatial frequency spectrum of q_n is cloaking the 'carrier' of the far PS, even at radii where r_2 is only slightly larger than or equal to r_1. The DOA estimate for φ_2 therefore becomes strongly inaccurate, leading to the comparably high mean DEE of more than $20\,\%$.

In the scenario where the first PS is located far away ($r_1 = 9.7\,\mathrm{m}$), the DEE improvement that occurs as the distance of the second source increases can be explained by looking at the source strength. As discussed in Section IV-F, a strong difference in strength of the PSs' components in q_n might cause the modulation contributions to mask the 'carrier' associated with the weaker source. That would lead to a bad DOA estimate for the more distant PS, hence causing an elevated mean DEE for $r_2 < 2\,\mathrm{m}$. Once the levels become more comparable, the mean DEE is similar to that obtained in the third scenario.

The third scenario had both PSs located at the same distance from the origin. While still above $4\,\%$ for $r_1 = r_2 = 0.2\,\mathrm{m}$, the mean DEE drops quickly towards very small values as the source radii increase. The field of the PSs eventually effectively resembles that of plane waves. Since their source strength is equal, their estimates would on average be of similar accuracy. The high mean DEE at very close range, $r_1 = r_2 < 1\,\mathrm{m}$, suggests that the modulation components occurring in the coefficient sequence are strong enough to cloak the dominance of their carriers, leading to elevated DEE results.

5) Practical Implications of the Results for Point Sources and OOP-PW: Point sources and OOP-PW are very relevant for any practical implementation. The point source can model any source whose dimensions are small compared to the wave length that is located within close distance to the measurement aperture. However, even if a sound source is sufficiently distant from the microphone array for its wave fronts to resemble those of a plane wave travelling parallel to the measurement

plane, there may be floor or ceiling reflections that lead to an OOP-PW in the sound field.

6) Influence of the Range of Coefficients: As stated by Blu et al. [11], using the highest possible number of coefficients as input to the estimation stage increases the robustness of the estimates against noise through redundancy. It was however found in this work that this only applies for sound fields with distinct plane waves that propagate in the plane of the measurement aperture. The simulations of both the point sources and the plane waves travelling out of the measurement plane have revealed that the imperfections introduced by the additional factors $\zeta_n(r')$ and $\Psi_n(k, k_r^{(l)}, \xi)$, respectively, affect the performance of the DOA estimates. For single sources, the influence of the additional factors can be overridden by keeping the range of coefficients q_n used in the DOA estimation algorithm as small as possible. As the number of sources increases, the estimator requires a higher number of coefficients (see Section IV-C). At the same time, however, the influence of the modulating factors compromises the performance of the estimator (see Section V-D1).

VI. Experimental Validation

This section investigates the DOA estimation accuracy based on measured sound fields. The sound fields were created using ten different sources, both individually and in various combinations.

One goal of the experiments was to test the dependency of the accuracy of the DOA estimates on the number of sources in the sound field. In order to do that, the data of the individual sources were linearly combined to generate arbitrary fields with multiple sources. The data that was measured with several active sources at the same time was used to validate the results obtained from the linear combination of the individual data.

A. Measurement Setup

The measurement array consists of 15 electret microphones arranged in a circle with uniform distribution on an acrylic cylindrical baffle with radius $\xi = 7.5\,\text{cm}$ (see Fig. 1). The array was set up in the anechoic chamber of the ISVR. Nine *Genelec 8020C* loudspeakers were arranged at different angles to the array at a distance larger than $r = 2.5\,\text{m}$, while one was deliberately put at a radial distance of $1\,\text{m}$ to simulate a point source. A diagram of the arrangement is shown in Fig. 12. The floor of the anechoic chamber consists of a steel grid over a volume of absorbing fibre glass wedges. To reduce the amount of reflections additional wedges were used to cover the floor between the loudspeakers and the microphone array.

Exponential sweeps of 10 seconds duration were used as measurement signals to exclude non-linear effects from the data and obtain a high measurement SNR, using the method proposed by Farina [21], [22]. The impulse responses obtained from the inverse filtering were then transformed into the frequency domain by using the *MATLAB* implementation of the Fast Fourier Transform (FFT). As an alternative, the same measurements were performed with band-limited noise (1 to $8\,\text{kHz}$) for comparison. All data was evaluated at the optimal frequency $f_c \approx 5.1\,\text{kHz}$ (compare Eq. IV-D).
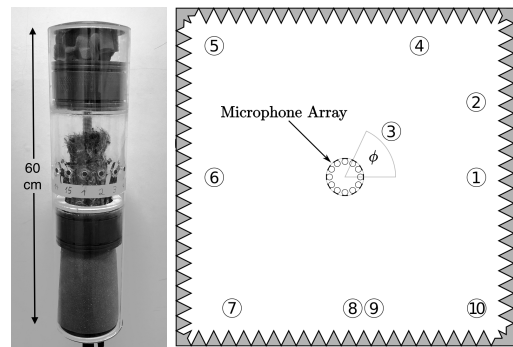

Fig. 12: (left) Microphone array and (right) diagram of the arrangement of array and sound sources inside the anechoic chamber.
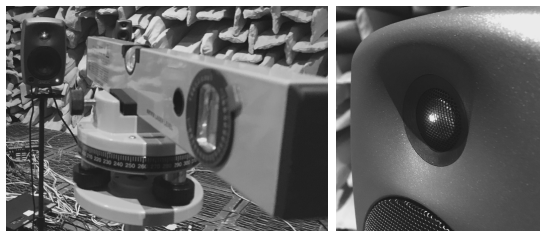

Fig. 13: (Left) Angle laser gauge with Genelec 8020C, (Right) Close-up view of tweeter with laser reflection on the diaphragm.

Source No.	Dist. to Origin	φ_l	$\tilde{\varphi}_k$	SNR/dB
1	2.88 m	$0°$	$2.4°$	73.0
2	3.47 m	$30°$	$33.9°$	72.9
3	1.04 m	$45°$	$43.8°$	80.6
4	3.45 m	$60°$	$58.2°$	69.1
5	4.20 m	$135°$	$138.3°$	65.9
6	3.08 m	$180°$	$182.7°$	61.1
7	3.24 m	$230°$	$233.7°$	73.4
8	2.52 m	$275°$	$276.6°$	68.0
9	2.56 m	$280°$	$282°$	77.3
10	3.35 m	$315°$	$314.1°$	72.1

TABLE I: Reference location information for all ten loudspeakers, estimated ($\tilde{\varphi}_l$) DOA given in degrees and SNR obtained from measurements of the individual sources at $f \approx 5.1\,\text{kHz}$, $N = 7$.

B. Measured DOA of the Sources

The DOAs of the ten loudspeakers were individually measured using a laser angle gauge, located at the origin of the coordinate system. The loudspeakers were adjusted so that the laser pointed at the middle of the tweeter (see Fig. 13), ensuring all tweeters were in the same plane as the 15 microphone capsules. The distance of each tweeter from the origin was obtained with a tape measure. The results are noted down in Table I. It shows that the error of the estimates is less than $4°$, and when ignoring the results for sources 2, 5 and 7, the errors are even less than 3 degrees, making the method's performance directly comparable to that of EB-ESPRIT [4].

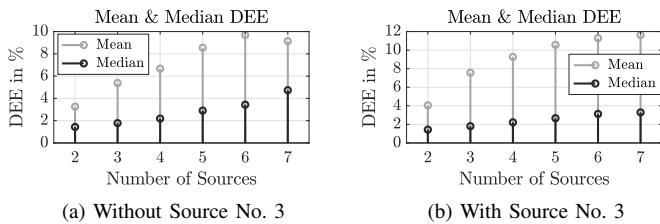The depicted values will be used as reference for the

(a) Without Source No. 3

(b) With Source No. 3

Fig. 14: Mean and median of the DEE in %, averaged over 1,000 randomised iterations for various numbers of sources.

Source	1	4	5	7	9	10
(est.) $\bar{\varphi}_l$ (deg)	3.7	60.6	126.7	230.9	284.4	309.8
(true) φ_l (deg)	0	60	135	230	280	315
(EB-ESPRIT) φ_l (deg)	4.4	70.5	136.8	231.4	282.9	315

TABLE II: Estimates from measurement signals obtained in a field with six noise sources active at the same time, with a mean DEE of $2.13\,\%$ (proposed method) and $1.94\,\%$ (EB-ESPRIT).

testing of the DOA estimation performance based on the linear combinations of individual sound fields. The SNR was estimated using Eq. (38), where the value of the measured data $\tilde{q}_n$ was used in the numerator. The noise coefficients ϵ_n were estimated from data taken from the tail of the estimated impulse responses as proposed by Stan et al. [23]. The SNR data suggests that the level difference between the strongest and weakest source is approximately $19.5\,\mathrm{dB}$.

C. Mean & Median DEE for Linear Combinations of Individual Measurements

The mean and median results for the DEE of DOA estimates obtained from measurements are shown in Fig. 14a. The results for each number of sources were averaged over 1,000 iterations, during which sources were randomly chosen, where the near-field source (No. 3) was excluded. The algorithm was set to detect $N = 7$ sources, but only the L sources with the largest magnitude $|b_l|$ were chosen for these results.

It can be seen that both the mean and the median have a rising trend as L increases. The mean rises to approx. $9.7\,\%$ for $L = 6$ and the median reaches slightly under $4.7\,\%$ for $L = 7$. That makes for an uncertainty of $\pm 17.46°$ and $8.5°$, respectively.

Figure 14b shows the results of the same experiment, only the near source was included in the pool of sources this time. Both mean and median have a rising trend as L increases. The mean peaks at approx. $11.6\,\%$ for $L = 7$ and the median reaches slightly under $3.3\,\%$. That makes for an uncertainty of $\pm 20.9°$ and $\pm 5.9°$, respectively.

The results of Table I indicate that the reference position of sources 4, 5 and 7 may suffer from setup imperfections. The influence of these were investigated by repeating the analysis using the estimated values $\tilde{\varphi}_l$ in Table I as reference. It was found that the mean and median DEE results then improve slightly for $L < 5$, compared to the results given in Fig. 14.

D. Example: Measurement of Six Sources

Measurement data obtained from six simultaneously active sources were analysed to confirm that the DOA estimation also works for fields with several sources playing (uncorrelated) noise at the same time. The results are given in Table II, along with the results obtained with EB-ESPRIT [4]. The mean accuracy of EB-ESPRIT is marginally higher than that of the

proposed method, however, the standard deviation of the error is $3.7°$ with EB-ESPRIT and $2.9°$ with the proposed method. EB-ESPRIT was implemented as described in [4], and 4096 time samples ($\approx 85\,\mathrm{ms}$) were used to obtain the result with both methods. Note that the proposed method is implemented in the frequency domain and was set to only evaluate data from the frequency bin at $5.1\,\mathrm{kHz}$, while another independent set of estimates could theoretically also be obtained for other frequency bands in a single FFT frame. While a comparison to other established methods (e.g. MUSIC, ESPRIT, etc.) may be possible in principle, it should be noted that, unlike EB-ESPRIT and the proposed method, they are not designed for wideband data [4] and they require interpretation of their associated angular spectrum. This underpins the potential of the proposed method.

E. Discussion

The values obtained for the mean and median of the DEE compare well to those obtained from the simulations with additional noise presented in Section V-C. The discrepancy between mean and median results suggest that the estimates' accuracy sometimes suffers from outliers. The effect of the near-field source was reflected in the slightly degraded mean performance, yet the median remains nearly entirely unaffected by it; it even improves at high values of L, probably due to a relatively high SNR. This suggests that, for the case under consideration, real sources are sufficiently different from ideal point sources to not entirely compromise the DOA estimation performance.

It must be expected that the small angular distance between sources 8 and 9 has a negative effect on the estimation performance. If the resolution of the method is not high enough to resolve the difference between the two sources, then this will potentially lead to worse DEE results whenever these two sources are randomly selected into the same set. The relatively good result of the multi-source measurement supports this hypothesis.

Another aspect that needs to be considered is the influence of the floor grid of the ISVR's anechoic chamber, which may have generated some unwanted reflections at high frequencies. This could then result in a wave field with plane waves that are not travelling parallel to the measurement plane z_M. It was shown in Section V that OOP-PW can degrade the performance of the DOA estimation mechanism due to the additional modulation of the sequence of coefficients q_n. It must thus be expected that the estimation performance with the measured data may have suffered from degradation due to grid floor reflections. While this type of vulnerability of the estimation method is

undesirable, it should be seen with relation to the fact that the measurement aperture is designed for two-dimensional measurements to begin with. It may be argued that this is actually not a fault of the method but of the two-dimensional measurement aperture.

VII. CONCLUSIONS

A thorough analysis of a previously proposed DOA estimation method was presented. The work is based on theoretical considerations, simulation and experimental results. The DOA estimation method uses data obtained from circular arrays in sound fields with a finite rate of innovation.

A set of theoretical models for the array output in such sound fields was derived, considering both plane waves and point sources. It has been argued that the proposed method has a distinction over other methods (e.g. MUSIC, ESPRIT, EB-ESPRIT) in not requiring the sources to be incoherent. For example, the proposed method could be used to identify the early reflections from the spatial impulse response of a room. It was found through theoretical considerations that the DOA estimates may suffer from point sources located within close range of the array and plane waves that do not travel parallel to the plane of the measurement aperture (OOP-PW).

The sound source models were applied in a simulation study to investigate the robustness and accuracy of DOA estimates against measurement noise and model imperfections. Percentages of average estimation errors were given, indicating the dependency of the DOA accuracy on the SNR and the number of plane waves in the field. It was found that the DOA estimation error for ideal sound fields increases significantly for sound fields where the rate of innovation exceeds what can be captured by the array.

Experimental results confirmed the findings on the performance in ideal sound fields of the simulation study. It was observed in both simulation and experimental results, that the mean accuracy is notably higher than its median. This suggests that the estimation method produces outliers that deviate from the median by a significant step, while 50% of the time, the estimates have a much higher accuracy than is suggested by the mean accuracy. A direct comparison to EB-ESPRIT found that both methods yield a comparable accuracy.

Point sources and OOP-PW were shown to severely degrade the DOA estimation performance if their number in the field is larger than one. These types of sources were argued to be practically relevant and thus pose an important factor to consider in practical implementations. It can be concluded that, since it captures only a two-dimensional subset of the sound field, a circular array and the data obtained from it are particularly vulnerable to any sound field components that deviate from a two-dimensional scenario.

APPENDIX A
GENERAL PLANE WAVE EXPANSION IN CYLINDRICAL COORDINATES

Using Eq. (1), the propagation term of the PW is given through

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik_r r(\cos\varphi\cos\phi + \sin\varphi\sin\phi)}e^{-ik_z z}, \quad (39)$$

where $\mathbf{r} = (r\cos\phi, r\sin\phi, z)^T$. Using the product theorems of the sine and cosine functions [24] and exploiting the symmetry of the latter yields

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik_r r\cos(\phi-\varphi)}e^{-ik_z z}. \quad (40)$$

The z-component of $\mathbf{k}$ can alternatively be written as a function of the elevation angle θ associated with the DOA, so that $k_z = k\cos\theta$. It then follows from (2) that $k_r = k\sin\theta$.

Using the Jacobi-Anger expansion [14] to expand the first term of Eq. (40) provides the expression

$$e^{-ik\sin\theta\, r\cos(\phi-\varphi)} = \sum_{n=-\infty}^{\infty} i^n J_n(-kr\sin\theta)e^{in(\phi-\varphi)}. \quad (41)$$

It follows from the series expansion of the Bessel function $J_n(z)$ [17] that $J_n(-z) = (-1)^n J_n(z)$. The conjunction of this and the result in (41) applied to Eq. (40) leads to the series expansion of a single PW in cylinder basis functions [17]

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik\cos\theta z}\sum_{n=-\infty}^{\infty} i^{-n} J_n(kr\sin\theta)e^{in(\phi-\varphi)}. \quad (42)$$

APPENDIX B
SERIES EXPANSION FOR THE SOUND FIELD OF A POINT SOURCE

In cylindrical coordinates, the Green function of the Helmholtz equation can be expressed through cylinder basis functions [25], [26], yielding

$$G(\mathbf{r}|\mathbf{r}') = \frac{i}{4}\sum_{n=-\infty}^{\infty} i^n \Phi_n^*(\phi')\Phi_n(\phi)$$
$$\cdot \int_{-\infty}^{\infty} \frac{e^{ik_z(z-z')}}{2\pi} R_n(k_r r)H_n^{(1)}(k_r r')dk_z \quad (43)$$

for $r < r'$. It poses an expression for the pressure of a PS located at $\mathbf{r}' = (r', \phi', z')$, with $k_r = k\sin\theta'$ and $k_z = k\cos\theta'$, where $\theta' = \arctan\left(\frac{r'}{z'}\right)$.

Note that the integral in (43) is not trivial to solve and typically needs to be either calculated numerically or approximated with a stationary phase approximation [17], [27].

APPENDIX C
HD COEFFICIENTS FOR THE SOUND FIELD OF A POINT SOURCE

The pressure field of a single PS measured on the circle A in the plane $z = z_M$ yields the result of the form in Eq. (13).

A comparison of the factors with those in Eq. (43) yields the expression for the pressure coefficients

$$p_n = b' \frac{i^{n+1}}{4} \Phi_n^*(\phi') \int_{-\infty}^{\infty} \frac{e^{ik_z(z_M - z')}}{2\pi} R_n(k_r\xi) H_n^{(1)}(k_r r') dk_z.$$

If the PS is located within the plane, i.e. $z' = z_M$, this expression simplifies to

$$p_n = b' \Phi_n^*(\phi') \frac{i^{n+1}}{8\pi} \int_{-\infty}^{\infty} R_n(k_r\xi) H_n^{(1)}(k_r r') dk_z. \qquad (44)$$

Equating the right hand side expression with $R_n(k\xi)q_n$ (compare Eq. (16)) and rigorous reformulation of Eq. (44) yields the expression for the HD coefficients:

$$q_n = b' \frac{e^{-in\phi'}}{\sqrt{2\pi}} \frac{i^{n+1}}{8\pi} \frac{\int_{-\infty}^{\infty} R_n(k_r\xi) H_n^{(1)}(k_r r') dk_z}{R_n(k\xi)} \qquad (45)$$

REFERENCES

[1] F.-M. Hoffmann, F. M. Fazi, and P. Nelson, "Plane wave identification with circular arrays by means of a Finite Rate of Innovation Approach," in *140th Convention of the Audio Engineering Society*, Paris, France, May 2016, Student Technical Papers Award. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18220

[2] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[3] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *Journal of the Acoustical Society of America*, vol. 4, pp. 2149–2157, Oct. 2004.

[4] H. Teutsch and W. Kellermann, "Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays," *Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 2724–2736, Nov. 2006.

[5] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 1, pp. 2–16, Jan. 2010.

[6] N. Barrett and S. Berge, "High angular resolution planewave expansion (HARPEX)," in *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, 2010, pp. 4–9.

[7] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[8] A. Wabnitz, N. Epain, A. McEwan, and C. Jin, "Upscaling ambisonic sound scenes using compressed sensing techniques," in *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2011, pp. 1–4.

[9] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation,*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[10] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002, IEEE Signal Processing Society's 2006 Best Paper Award.

[11] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, Mar. 2008.

[12] F. M. Fazi, "Sound field reproduction," Ph.D. dissertation, University of Southampton, Faculty of Engineering, Science and Mathematics, Institute of Sound and Vibration Research, 2010.

[13] F. M. Fazi, M. Noisternig, and O. Warusfel, "Representation of sound fields for audio recording and reproduction," in *Acoustics 2012: 11ème Congrès Français d'Acoustique. Annual meeting of the Institute of Acoustics, Nantes, FR, 23 - 27 Apr 2012*, 2012.

[14] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer New York Heidelberg Dordrecht London, 1998.

[15] M. A. Poletti, "A unified theory of horizontal holographic sound systems," *Journal of the Audio Engineering Society*, vol. 48, no. 12, pp. 1155–1182, 2000. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=12033

[16] F.-M. Hoffmann and F. Fazi, "Theoretical study of acoustic circular arrays with tangential pressure gradient sensors," *IEEE/ACM Transactions on Audio, Speech, and Language Processing,*, vol. 23, no. 11, pp. 1762–1774, Nov. 2015.

[17] E. G. Williams, *Fourier Acoustic: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 24-28 Oval Road, London NW1 7DX, UK, 1999.

[18] D. L. Alon and B. Rafaely, "Spatial aliasing-cancellation for circular microphone arrays," in *4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, Nancy, France, May 2014, pp. 137–141.

[19] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, Jan. 2005.

[20] M. A. Poletti, "Effect of noise and transducer variability on the performance of circular microphone arrays," *Journal of the Audio Engineering Society*, vol. 53, no. 5, pp. 371–384, 2005.

[21] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *108th Convention of the Audio Engineering Society*, Feb. 2000. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=10211

[22] ——, "Advancements in impulse response measurements by sine sweeps," in *122nd Convention of the Audio Engineering Society*, Vienna, Austria, 2007.

[23] G.-B. Stan, J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *Journal of the Audio Engineering Society*, vol. 50, pp. 249–262, 2002.

[24] I. Bronstein, K. Semendjajew, G. Musiol, and H. Mühlig, *Taschenbuch der Mathematik*. Wissenschaftlicher Verlag Harri Deutsch GmbH, Frankfurt am Main, Germany, 2006.

[25] J. Asvestas, J. Bowman, P. Christiansen, O. Einarsson, R. Kleinman, D. Sengupta, T. Senior, F. Sleator, P. Uslenghi, and N. Zitron, *Electromagnetic and Acoustic Scattering by Simple Shapes*, J. Bowman, T. Senior, and P. Uslenghi, Eds. North-Holland Pub. Co. - Amsterdam, 1970. [Online]. Available: https://books.google.co.uk/books?id=ueENAQAAIAAJ

[26] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*, W. Perkins, Ed. John Wiley & Sons, 2003.

[27] L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of Acoustics*, 4th ed., B. Russiello, Ed. John Wiley & Sons, Inc., New York Chichester Weinheim Brisbane Singapore Toronto, 2000.