PAPER

Efficient array design algorithm for wide-band application of the MULTIPLE Signal Classification algorithm

Ahmed Desoki^{1,*}, Jun-ichi Takada^{2,†} and Ichiro Hagiwara^{1,‡}

¹*Graduate School of Science and Technology, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8552 Japan*

²*Department of International Development Engineering, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8552 Japan*

(Received 9 October 2007, Accepted for publication 9 October 2008)

Abstract: This paper analyzes the error in MUSIC results due to the effect of finite precision arithmetic. Thus, relation of this error to sources correlation level and array and sources configuration parameters is clearly identified. As a result efficient array design algorithm suitable for acoustic environments is derived. This algorithm is efficient in the sense that it can determine minimum number of sensors. This algorithm is quite general as it includes the effect of all parameters such as number of sources, sources correlation level, maximum resolution, maximum source angle, number of sensors, sensor spacing and arithmetic precision. Also this algorithm is shown to be seamlessly applicable in realistic environments where many additional effects and sources of error often exist. During this paper it is shown that this algorithm is indispensable for DOA estimation in wide-band and reverberant environments.

Keywords: MUSIC, Wide-band, Precision, Resolution, Array design, Sensor spacing, Quantization

PACS number: 43.60.Gk [doi:10.1250/ast.30.187]

1. INTRODUCTION

Many practical applications rely on microphone arrays to detect and localize multiple acoustic sources. Direct examples include automated systems for surveillance, video conferencing and hands free speech acquisition. Other industrial applications include acoustic mapping of machines and vehicles for identification of noisy regions or parts. Localization data is also used indirectly for other tasks such as dereverberation of speech, fault prediction and analysis of machinery, cuing and tracking of television cameras, speaker verification, etc.

Compared to beamformer based [1] and time difference of arrival (TDOA) based [2] methods, eigenvector based methods are the only direction of arrival (DOA) estimation methods designed for multiple sources and are characterized by their super-resolution [1]. Eigenvector based methods have become very popular in narrow-band array processing [1] and led to a family of well known algorithms for narrow band direction finding like MUSIC [3], ESPRIT

[4] and MIN-NORM [5]. However, these methods are still not well applied in acoustical applications.

The basic assumption of eigenvector based methods is that the rank of the noise free array output covariance matrix (Eq. (8)) equals to the number of sources [1]. This assumption implies two subsequent assumptions, namely non coherent source signals and independent direction vectors associated with true angles of arrivals. Since the emergence of these methods, many researches concerned the case when the first assumption is violated [1]. However, for most narrow-band applications, the second assumption is an always valid fact.

Trying to apply techniques like MUSIC to realistic acoustic scenarios, two major problems are faced, which are reverberation and wide-band nature of sound signals. Reverberation not only yields the problem of coherent sources but also increases the number of sources to be resolved. On the other hand, wide-band nature of sound signals led to the development of some wide-band generalizations as incoherent averaging [6], coherent averaging [7] and focusing techniques [8–12]. Most of these methods can also handle the coherent sources problem and rely on applying eigenvector based methods on an array whose sensor spacing is considerably less than half wave length.

*e-mail: desoki.a.aa@m.titech.ac.jp

†e-mail: takada@ide.titech.ac.jp

‡e-mail: hagiwara.i.aa@m.m.titech.ac.jp

Considering the increase of the number of sources and the decrease of sensor spacing, violation of the second assumption of eigenvector based methods began to appear as a problem on computers with finite precision arithmetic. That is to say, for some number of sources, sensor spacing and sources separation, independence of direction vectors associated with the true DOA's may be lost within the finite precision arithmetic. In this case, the rank of the noise free covariance matrix will seem less than the number of sources and the hence obtained results will be erroneous.

In this paper, this phenomenon is analyzed for the MUSIC method and an efficient array design algorithm that is guaranteed for accurate results for any sources configuration and arithmetic precision is derived. This array design algorithm is quite general in the sense that it includes the effect of all parameters such as number of sources, sources correlation level, maximum resolution, maximum source angle, number of sensors, sensor spacing and arithmetic precision. Also this algorithm is shown to be seamlessly applicable in realistic environments where many additional effects and sources of error often exist. During this paper it is shown that this algorithm is indispensable for DOA estimation in wide-band and reverberant environments.

Kaveh in [13] derived a statistical expression for the deviation of MUSIC results from true values at the true DOA's. This quite complex expression was approximated for the special case of two plane waves to obtain MUSIC resolution threshold. However, Kaveh did not consider the effect of the finite precision arithmetic. This appears clearly from his expression which claims that infinitesimally separated sources can be resolved when the exact covariance matrix is available. It is however shown in this paper that there is still a resolution limit even when the exact covariance matrix is available.

This paper is organized as follows. In section 2 the signal propagation model is formulated and the MUSIC method is presented. In section 3 the condition number is introduced as a numerical measure of the rank of the covariance matrix. Then error in MUSIC results is derived in terms of this condition number and arithmetic precision. Relation between this condition number and array and sources configuration parameters is clearly identified in section 3.1. In section 3.2 a global array design algorithm is derived which guarantees accurate MUSIC results for any sources configuration and arithmetic precision. In section 4 the proposed array design algorithm is verified through simulations applied to an ideal wide-band DOA estimation problem. Applicability of the proposed algorithm in realistic environments is examined in section 4.1. In section 4.2 it is shown how the proposed array design algorithm can still be used in realistic experimentations where many additional effects and sources of error often

exist. Throughout this paper, plane wave propagation model is assumed and the number of sources is assumed known beforehand.

2. MODEL FORMULATION

Consider a uniform linear array (ULA) of M identical sensors receiving the wave signal generated by K narrow-band sources in the presence of an arbitrary noise signal. At time instant t , the received signal at the m -th sensor; $m = 1, 2, \dots, M$, can be written as [1]

$$p_m(t) = \sum_{k=1}^K u_k(t - \tau_{mk}) + n_m(t) \quad (1)$$

where u_k is k -th source signal and n_m is noise signal generated at the m -th sensor. Assuming plane wave propagation model, the delay τ_{mk} is expressed as $\tau_{mk} = -x_m \sin \phi_k / c$ where x_m is the m -th sensor position, ϕ_k is the angle of arrival of the k -th signal measured from the normal to the line joining the sensors, and c is propagation speed. If the center frequency of the received signals, f , is fairly large compared to the frequency band-width, narrow-band signals can be assumed. In this case (Eq. (1)) reduces to [1]

$$p_m(t) = \sum_{k=1}^K u_k(t) e^{j\pi \tilde{x}_m \sin \phi_k} + n_m(t) \quad (2)$$

where $\tilde{x}_m = 2fx_m/c = x_m/(\lambda/2)$ is the m -th sensor position normalized by half wavelength $\lambda/2$. During this paper, we will use the tilde “~” to denote normalization by half wavelength.

Based on this formulation, several narrow-band techniques for DOA estimation were developed [1]. Among them, is the Multiple Signal Classification (MUSIC) technique described in the following subsection.

2.1. Narrow-Band MUSIC Technique [3]

Writing (Eq. (2)) more compactly in vector form yields

$$\mathbf{p}(t) = \mathbf{A}\mathbf{u}(t) + \mathbf{n}(t) \quad (3)$$

where $\mathbf{p}(t) = [p_1(t), p_2(t), \dots, p_M(t)]^T$, $\mathbf{A} = \sqrt{M}[\mathbf{a}(\phi_1), \mathbf{a}(\phi_2), \dots, \mathbf{a}(\phi_K)]$, $\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_K(t)]^T$, $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T$ and the superscript “ T ” denotes matrix transpose. The vector $\mathbf{a}(\phi_k)$ is the normalized direction vector associated with the arrival angle ϕ_k defined as

$$\mathbf{a}(\phi_k) = \frac{1}{\sqrt{M}} [e^{j\pi \tilde{x}_1 \sin \phi_k}, e^{j\pi \tilde{x}_2 \sin \phi_k}, \dots, e^{j\pi \tilde{x}_M \sin \phi_k}]^T \quad (4)$$

Assuming noises are uncorrelated between themselves and with source signals, it follows that the array output covariance matrix defined by

$$\mathbf{R} = E[\mathbf{p}(t)\mathbf{p}^\dagger(t)] \quad (5)$$

has the form

$$\mathbf{R} = \mathbf{A}\mathbf{R}_u\mathbf{A}^\dagger + \sigma^2\mathbf{I} \quad (6)$$

where $E[\]$ denotes statistical expectation, the superscript “ $\dagger$ ” denotes hermitian transpose, $\mathbf{R}_u$ is the source covariance matrix defined as

$$\mathbf{R}_u = E[\mathbf{u}(t)\mathbf{u}^\dagger(t)] \quad (7)$$

and σ^2 represents the noise power common to all sensors. The first term in (Eq. (6)) expresses the noise free covariance matrix. This term will be denoted by $\widehat{\mathbf{R}}$

$$\widehat{\mathbf{R}} = \mathbf{A}\mathbf{R}_u\mathbf{A}^\dagger \quad (8)$$

Assuming independent source signals, $\mathbf{R}_u$ is full rank. On the other hand, for $M > K$, due to the linear independence of the direction vectors $\mathbf{a}(\phi_k) \forall k$, $\mathbf{A}$ is also of rank K . Consequently the matrix $\widehat{\mathbf{R}}$ is also of rank K . Based on these assumptions, peaks of the function

$$S(\phi) = \frac{1}{\mathbf{a}^\dagger(\phi)\mathbf{B}_n\mathbf{B}_n^\dagger\mathbf{a}(\phi)} \quad (9)$$

will correspond to the true DOA's [3], where $\mathbf{B}_n$ will be defined later in (Eq. (12)) in section 3.

However, sometimes the computer feels the linearly independent columns of the matrix $\mathbf{A}$ as dependent. This can happen when the true angles of arrival are very close or nearly collinear. In this case, due to the finite precision of the computer and the finite length of the direction vectors, M , the full rank matrix $\mathbf{A}$ will seem rank deficient to the computer. Accordingly, the rank of the matrix $\widehat{\mathbf{R}}$ and hence the distinct eigenvalues of $\mathbf{R}$ will seem less than K . This falseness not only causes wrong prediction of the number of sources, in case it is unknown beforehand, but also MUSIC results of (Eq. (9)) will be wrong, even if the correct number of sources is substituted. This led the authors to inspect which sources and array configurations are sufficient for maintaining the full rank property of $\mathbf{A}$ for finite precision calculations, as detailed in section 3.

2.2. Application of MUSIC for Wide-Band

For this case, the narrow band assumption used to derive (Eq. (2)) is no longer valid. Applying Fourier transformation to (Eq. (1)) yields

$$P_m(f) = \sum_{k=1}^K U_k(f) e^{j\pi\tilde{x}_m(f)\sin\phi_k} + N_m(f) \quad (10)$$

Equation (10) is identical to (Eq. (2)) but in the frequency domain. Please note that $\tilde{x}_m(f)$ defined in (Eq. (2)) is now frequency dependent. In other words, although the physical sensor location x_m is constant, the normalized sensor location $\tilde{x}_m(f)$ does change with frequency. This in turn means that covariance matrices at different frequencies cannot be averaged as they are considered for different sensor arrays. This is the source of

pain for which narrow-band DOA estimation techniques cannot be directly applied to wide-band signals. Among wide-band generalizations of are incoherent and coherent averaging methods. They are formulated in appendices A and B respectively.

3. EFFECT OF FINITE PRECISION ARITHMETIC ON MUSIC RESULTS

Let $\mu_1 \geq \mu_2 \geq \dots \geq \mu_K$ denote the K nonzero eigenvalues of the non-negative definite matrix $\widehat{\mathbf{R}}$. Then for the ideal model described by (Eq. (6)), eigenvalue decomposition of $\mathbf{R}$ yields the M eigenvalues

$$\lambda_m = \begin{cases} \mu_k + \sigma^2 & m = k = 1, 2, \dots, K \\ \sigma^2 & m = K + 1, \dots, M \end{cases} \quad (11)$$

shown in Fig. 1 and their associated eigenvectors $\beta_1, \beta_2, \dots, \beta_K, \beta_{K+1}, \dots, \beta_M$.

Defining $\mathbf{B}_n$ as the matrix whose columns are eigenvectors associated with the noise eigenvalues,

$$\mathbf{B}_n = [\beta_{K+1}, \dots, \beta_M] \quad (12)$$

MUSIC spectrum can be calculated as in (Eq. (9)) [3].

For the eigen decomposition problem, as any finite precision arithmetic operation, numerical error between calculated and true eigenvalues, eigenvectors and subspaces does exist. Fortunately the perturbation theory [15,16] provides us with bounds for these errors. Defining the noise subspace as the subspace spanned by the columns of matrix $\mathbf{B}_n$, it becomes clear from (Eq. (9)) that MUSIC spectrum is affected only by the noise subspace calculation error. In [17] Davis *et al.* gave a general error bound for subspaces and later in [18] Stewart gave a more restricted one. During this section, bounds given in [17] and applied in [16] and [19] will be used. Readers interested in differences between these bounds may refer to [18].

Let λ , β and N be a true eigenvalue, its associated eigenvector and the noise subspace for the complex hermitian matrix $\mathbf{R}$ defined by (Eq. (6)). Also consider their associated calculated versions as $\hat{\lambda}$, $\hat{\beta}$ and $\hat{N}$. The error E in $\hat{N}$ is bounded by [16, Theorem (11-7-11)]

Fig. 1 Eigenvalues of $\mathbf{R}$.

$$E \triangleq |\sin \angle(N, \hat{N})| \leq \frac{\|\mathbf{Res}\|_2}{\text{gap}} \quad (13)$$

where $\angle(N, \hat{N})$ is the largest angle between N and $\hat{N}$ [15], $\|\cdot\|_2$ is the Euclidian norm, $\mathbf{Res}$ (called the residual) denotes the error in the associated calculated eigenvalues and is defined by $\mathbf{Res} = \mathbf{R}\hat{\mathbf{B}}_n - \hat{\mathbf{B}}_n \text{diag}(\lambda_{K+1}, \dots, \lambda_M)$ and “gap” is the minimum absolute gap between signal subspace and noise subspace eigenvalues as shown in Fig. 1. In general, $\mathbf{Res}$ is bounded by the empirical relation [16, sect. 2-3-2], [19]

$$\|\mathbf{Res}\|_2 \leq f_n(M)\varepsilon\|\mathbf{R}\|_2 \quad (14)$$

where ε is the calculation precision and $f_n(M)$ is a modestly growing function of M which depends on the particular numerical algorithm. Substituting (Eq. (14)) in (Eq. (13)) yields

$$E \leq \frac{f_n(M)\varepsilon\|\mathbf{R}\|_2}{\text{gap}} \quad (15)$$

The reciprocal of “gap” can be thought of as a measure of the sensitivity of the subspace N to calculation errors [15]. Thus the cluster of noise eigenvalues which is far away from the signal eigenvalues, may have well determined noise subspace $\hat{N}$ even if its individual eigenvectors are ill-conditioned [19].

Considering (Eq. (15)), substituting $f_n(M) = 1$ is usually quite acceptable [19]. From its definition, $\text{gap} = \lambda_K - \sigma^2 = \mu_K$. On the other hand, $\|\mathbf{R}\|_2$ in (Eq. (15)) is strictly evaluated as $\|\mathbf{R}\|_2 = \lambda_1 = \mu_1 + \sigma^2$. However, accurate calculation of eigenvectors, and hence subspaces, of $\mathbf{R}$ defined by (Eq. (6)) is independent on σ^2 and is dependant only on $\hat{\mathbf{R}}$. Hence and since the goal of this paper is to keep accurate noise subspace calculation, σ^2 can be omitted when substituting for $\|\mathbf{R}\|_2$ in (11) [20]. Thus, substituting $f_n(M)$, “gap” and $\|\mathbf{R}\|_2$ in (Eq. (15)) yields

$$E \leq \varepsilon \frac{\lambda_1 - \sigma^2}{\lambda_K - \sigma^2} = \varepsilon \frac{\mu_1}{\mu_K} \quad (16)$$

Since $\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u$ has the same nonzero eigenvalues as $\hat{\mathbf{R}}$, we can substitute $\mu_1/\mu_K = \kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u)$ in (Eq. (16)) to obtain

$$E \leq \varepsilon \kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u) \quad (17)$$

where $\kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u)$ denotes the condition number for the matrix $\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u$ with respect to Euclidian norm and is defined by $\kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u) = \sigma_{\max}(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u) / \sigma_{\min}(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u)$ [15], where $\sigma_{\max}$ and $\sigma_{\min}$ are the maximal and minimal singular values respectively. By definition the condition number satisfies $1 \leq \kappa \leq \infty$. Since the most reliable measure of matrix rank is the number of non-zero singular values, the condition number κ is a very reliable numerical measure of

matrix rank [21]. For any matrix $\mathbf{C}$, $\kappa(\mathbf{C})$ values close to 1 indicate that columns of $\mathbf{C}$ are very independent (orthogonal). The higher $\kappa(\mathbf{C})$ the weaker the independence until $\mathbf{C}$ columns become dependent (rank deficient) for $\kappa(\mathbf{C}) = \infty$ [22].

By constraining the right hand side of (Eq. (17)) not to exceed some threshold $E_{\max}$ yields the condition

$$\kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u) \leq \frac{E_{\max}}{\varepsilon} \quad (18)$$

For calculations with precision ε , (Eq. (18)) gives the maximum allowable threshold for $\kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u)$. If this threshold is exceeded, the rank of the matrix $\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u$ will seem less than K to the computer. Consequently the matrix $\mathbf{R}$ will have less than K distinct eigenvalues and erroneous MUSIC spectrum will result.

Substituting the identity $\kappa(\mathbf{A}^\dagger \mathbf{A} \mathbf{R}_u) \leq \kappa(\mathbf{A}^\dagger \mathbf{A})\kappa(\mathbf{R}_u) \equiv \kappa^2(\mathbf{A})\kappa(\mathbf{R}_u)$ [21] in (Eq. (17)), we can similarly obtain the more conservative condition

$$\kappa^2(\mathbf{A}) \leq \frac{E_{\max}}{\varepsilon \kappa(\mathbf{R}_u)} \quad (19)$$

Since $\kappa(\mathbf{A})$ represents the dependence of direction vectors $\mathbf{a}(\phi_k) \forall k$ and $\kappa(\mathbf{R}_u)$ determines the correlation level between source signals, then (Eq. (19)) gives the maximum allowable $\kappa(\mathbf{A})$, or the maximum attainable resolution, for a specific level of correlation of source signals.

In the following subsection, the condition number $\kappa(\mathbf{A})$ will be related to array and sources configuration parameters. This relation will then be used to design an array such that $\kappa(\mathbf{A})$ never exceeds the limit set by (Eq. (19)) for certain sources configuration.

3.1. $\kappa(\mathbf{A})$ Relation to Array and Sources Configuration Parameters

By definition $\kappa^2(\mathbf{A}) = \kappa(\mathbf{A}^\dagger \mathbf{A})$. For a ULA symmetric with respect to the origin and with normalized sensor spacing $\tilde{d}$, substituting $\tilde{x}_m = l\tilde{d}$ in (Eq. (4)), where $l = m - 1 - (M - 1)/2$, and making some mathematical simplifications, we obtain

$$\frac{\mathbf{A}^\dagger \mathbf{A}}{M} = \begin{bmatrix} 1 & & & \\ & 1 & \mathbf{a}_j^\dagger \mathbf{a}_i & \\ & \mathbf{a}_i^\dagger \mathbf{a}_j & \ddots & \\ & & & 1 \end{bmatrix}$$

where

$$\mathbf{a}_i^\dagger \mathbf{a}_j = \frac{\sin(M\eta_{ij})}{M \sin(\eta_{ij})}, \quad (20)$$

$$\eta_{ij} = \pi \tilde{d} \frac{\sin \phi_j - \sin \phi_i}{2} \quad (21)$$

and $\mathbf{a}_i \triangleq \mathbf{a}(\phi_i) \forall i : 1, \dots, K$.

Fig. 2 $|\mathbf{a}_i^\dagger \mathbf{a}_j|$ versus $|\eta_{ij}|$ for $M = 7$ and the two worst cases for sources.

A plot for $|\mathbf{a}_i^\dagger \mathbf{a}_j|$ versus $|\eta_{ij}|$ is shown in Fig. 2. As shown in this plot, $|\mathbf{a}_i^\dagger \mathbf{a}_j|$ has two main lobes at $|\eta_{ij}| = 0$ and $|\eta_{ij}| = \pi$ and has other small oscillations in between. Now we define the safe zone in Fig. 2 as the zone which includes that in between small oscillations. The bound of this zone is denoted by $(\eta_{\text{limit}}, \mathbf{aa}_{\text{limit}})$. This zone is characterized by small values for $|\mathbf{a}_i^\dagger \mathbf{a}_j| \forall i, j : 1, \dots, K$. This indicates well independence of $\mathbf{a}_i$ and $\mathbf{a}_j$. Thus, according to Gershgorin circle theorem [15], eigenvalues of $\mathbf{A}^\dagger \mathbf{A}/M$ matrix will be close to 1 and hence $\kappa(\mathbf{A})$ will be small for the safe zone. The best case is when $|\mathbf{a}_i^\dagger \mathbf{a}_j| = 0 \forall i, j$. In this case, $\mathbf{a}_i$ and $\mathbf{a}_j$ are orthogonal and $\kappa(\mathbf{A}) = 1$. Conversely, we define the dangerous zone in Fig. 2 as the remaining zone which has higher values for $|\mathbf{a}_i^\dagger \mathbf{a}_j|$ and can cause high values for $\kappa(\mathbf{A})$. Now, let's make the following definitions. Define $\delta\phi$ as the minimum angle separating two sources, ϕ_{max} as the maximum angle of a source, $\mathbf{aa}_{\text{max } 1}$ as the maximum possible left hand side value for $|\mathbf{a}_i^\dagger \mathbf{a}_j|$ and $\mathbf{aa}_{\text{max } 2}$ as the maximum possible right hand side value for $|\mathbf{a}_i^\dagger \mathbf{a}_j|$ as shown in Fig. 2. Assuming $\mathbf{aa}_{\text{max } 1}$ occurs for $(i = l, j = m)$ and $\mathbf{aa}_{\text{max } 2}$ occurs for $(i = l, j = n)$, it can be shown that $\mathbf{aa}_{\text{max } 1}$ and $\mathbf{aa}_{\text{max } 2}$ occur for the respective pairs $(\phi_l = \phi_{\text{max}}, \phi_m = \phi_{\text{max}} - \delta\phi)$ and $(\phi_l = \phi_{\text{max}}, \phi_n = -\phi_{\text{max}})$ denoted in Fig. 2 by case (1) and case (2) respectively. Substituting these values in (Eq. (21)) yields the corresponding values of $|\eta_{ij}|$ as

$$\eta_1 = \frac{\pi \tilde{d}}{2} (\sin(\phi_{\text{max}}) - \sin(\phi_{\text{max}} - \delta\phi)) \quad (22)$$

and

$$\eta_2 = \pi \tilde{d} \sin(\phi_{\text{max}}) \quad (23)$$

In cases (1) and (2) of Fig. 2, s_i is the opposite of the DOA for angle ϕ_i and is corresponding to $\mathbf{a}_i$.

Fig. 3 Most dangerous scenarios for any K sources.

Based on the two worst cases of Fig. 2, it was noticed that an upper bound for $\kappa(\mathbf{A})$ can be calculated as stated in the following observation.

Observation (1): For any K sources with minimum separation $\delta\phi$ and maximum angle ϕ_{max} , i.e. $|\phi_i - \phi_j| \geq \delta\phi$ and $|\phi_i| \leq \phi_{\text{max}} \forall i, j : 1, \dots, K$, then the maximum value for $\kappa(\mathbf{A})$, denoted as $\kappa_{\text{max}}(\mathbf{A})$, occurs for either one of the two scenarios shown in Fig. 3 provided that either $|\mathbf{a}_1^\dagger \mathbf{a}_2|$ for Scenario (1) or $|\mathbf{a}_1^\dagger \mathbf{a}_K|$ for Scenario (2) or both of them fall in the dangerous zone shown in Fig. 2.

Proof: Theoretical proof of this observation is extremely difficult. However, this observation is found valid through extensive Monte Carlo simulations.

Observation (2): Substituting $l = 1, m = 2$, and $n = K$ in cases (1) and (2) of Fig. 2, if $\mathbf{aa}_{\text{max } 1} > \mathbf{aa}_{\text{max } 2}$, then scenario (1) of Fig. 3 has $\kappa(\mathbf{A})$ higher than scenario (2). Conversely, if $\mathbf{aa}_{\text{max } 2} > \mathbf{aa}_{\text{max } 1}$, then scenario (2) of Fig. 3 has $\kappa(\mathbf{A})$ higher than scenario (1). In other words, if case (1) is more dangerous than case (2), then scenario (1) of Fig. 3 is more dangerous than scenario (2). Conversely, if case (2) is more dangerous than case (1), then scenario (2) of Fig. 3 is more dangerous than scenario (1).

Proof: Also no theoretical proof is given for this observation. It is found always valid in our extensive Monte Carlo simulations.

Based on observation (1), $\kappa_{\text{max}}(\mathbf{A})$ can be calculated as the maximum of $\kappa(\mathbf{A})$ for scenarios (1) and (2). In Fig. 4, $\kappa_{\text{max}}^2(\mathbf{A})$ is calculated for $K = 5$ sources and $\phi_{\text{max}} = 30^\circ$ and its contours are plotted versus $M, \delta\phi$ and $\tilde{d}$. Lower contour limit of this plot is set to 2.5 instead of 1 to avoid non interesting damped oscillations found in this range $1 \leq \kappa_{\text{max}}(\mathbf{A}) < 2.5$.

3.2. Efficient Array Design Algorithm

We mean by array design how to determine array configuration parameters $M, \tilde{d}$, and array distance from sources plane (equivalent to determining ϕ_{max}) such that the array is capable of resolving any K sources whose arrival angles are within $[-\phi_{\text{max}}, \phi_{\text{max}}]$ and with maximum

Fig. 4 $\kappa_{\max}^2(\mathbf{A})$ contours for $K=5$ sources and $\phi_{\max}=30^\circ$; (a) versus M and $\delta\phi$ for $\tilde{d}=1$. (b) versus M and $\tilde{d}$ for $\delta\phi=5^\circ$.

resolution of $\delta\phi$. The proposed algorithm is efficient in the sense that it can determine minimum number of sensors.

Many design algorithms can be set to choose the above mentioned parameters for different scenarios. Two of such scenarios are shown in Fig. 5. As shown in this figure, the proposed design algorithm is quite general and can be used for any combination of input parameters.

4. VERIFICATION OF THE PROPOSED ARRAY DESIGN ALGORITHM

Example (1): Consider the problem of designing a

ULA for the very wide-band extending from 1 kHz to 10 kHz. Using the incoherent averaging approach, physical sensor spacing is set to half the wavelength of the maximum frequency. Thus $\tilde{d}(f)$ will range from 0.1 to 1, which is too severe variation. Hence application of the MUSIC algorithm over wide frequency bands gives us an excellent opportunity to verify our claimed array design algorithm for different values of $\tilde{d}$.

Now we apply our design algorithm of Fig. 5(b) on the DOA estimation problem defined by: $K=5$ white Gaussian noise sources, $\phi_{\max}=30^\circ$, $\delta\phi=5^\circ$ and $\text{SNR}=\infty$, where SNR is the input signal to noise ratio defined by $\text{SNR} = E[|U(f)|^2]/\sigma^2(f)$. Infinite SNR was chosen to express the difference of this work from that of Kaveh [13]. As Kaveh did not consider the effect of finite precision arithmetic, his expression showed that infinitely separated sources can be resolved for infinite SNR and number of averagings, i.e. when the exact covariance matrix is obtained. In contrast, we prove through this simulation that, due to finite precision arithmetic, MUSIC still has finite resolution even when the exact covariance matrix is available and the SNR is infinity.

Concerning the incoherent averaging approach, MUSIC results should be kept accurate all over this large $\tilde{d}(f)$ variation. From Fig. 4(b), it is clear that the smaller $\tilde{d}$ the higher $\kappa_{\max}^2(\mathbf{A})$ and hence the higher the MUSIC error. So, the sensor array is designed on the smallest sensor spacing, $\tilde{d}=0.1$.

Using 100 averagings or higher, $\kappa(\mathbf{R}_U(f))$ can be kept around 2. Following the algorithm of Fig. 5, for MUSIC error threshold $E_{\max}=10^{-3}$, we obtain $\kappa_{\max}^2(\mathbf{A}) \leq 10^{-3}/2\epsilon_{\text{single}}$ and $\kappa_{\max}^2(\mathbf{A}) \leq 10^{-3}/2\epsilon_{\text{double}}$ for single and double precision arithmetic respectively, where $\epsilon_{\text{single}}=2^{-23}$ and $\epsilon_{\text{double}}=2^{-52}$. Contour lines corresponding to these values are almost coincident with $\kappa_{\max}^2(\mathbf{A})=10^{-3}/\epsilon_{\text{single}}$ and $\kappa_{\max}^2(\mathbf{A})=10^{-3}/\epsilon_{\text{double}}$ contours of Fig. 4. Then following Fig. 5(b) and Fig. 4(b) for $\tilde{d}=0.1$, $M=97$ and $M=10$

Fig. 5 Efficient array design algorithms.

Fig. 6 Single precision MUSIC results for $M = 97$; (a) $\kappa^2(\mathbf{A})$ versus $\tilde{d}$. (b) E versus $\tilde{d}$. (c) MUSIC results for different $\tilde{d}$.

Fig. 7 Double precision MUSIC results for $M = 10$; (a) $\kappa^2(\mathbf{A})$ versus $\tilde{d}$. (b) E versus $\tilde{d}$. (c) MUSIC results for different $\tilde{d}$.

are obtained for single and double precision arithmetic respectively. These values of M are the minimum number of sensors required for MUSIC error $E \leq 10^{-3}$.

As was discussed in section 3.1, for the above defined problem, case (1) of Fig. 2 and hence scenario (1) of Fig. 3 is the most dangerous. This scenario is simulated for different $\tilde{d}$ values as shown in Figs. 6 and 7 for single and double precision arithmetic respectively.

As shown in Figs. 6 and 7, for the range $\tilde{d} \geq 0.1$, arrays designed using our algorithm keep $\kappa^2(\mathbf{A})$ and MUSIC error E lower than their respective maximum allowable thresholds. Hence accurate MUSIC results are obtained as in Figs. 6(c) and 7(c). In conclusion, these figures prove the accuracy and efficiency of the proposed array design algorithm.

4.1. Applicability of the Proposed Algorithm in Realistic Environments

So far the exact ensemble averages of the array output covariances, Eq. (6), were assumed available. In this case, MUSIC results are unaffected by the SNR and averaging is

not necessary. In reality however the exact covariance matrix is not available and is instead estimated from the array output. Considering realistic experimentations, additional effects that cause violations to the model of Eq. (6) often exist. Some example effects include measurement precision, noise signals correlated between themselves and/or with source signals, non-planer propagation wavefronts, sensors positioning error, sensors calibration error, etc. For such experimentations the estimated covariance matrix can be written as

$$\mathbf{R}^{\text{est}} = \mathbf{A}\mathbf{R}_u\mathbf{A}^\dagger + \mathbf{R}_{\text{errors}} = \hat{\mathbf{R}} + \mathbf{R}_{\text{errors}}, \quad (24)$$

where $\mathbf{R}_{\text{errors}}$ is an additional error term due to the effect of several imperfections and sources of error of the particular experiment. The resulting error in the calculated noise subspace due to the use of $\mathbf{R}^{\text{est}}$ instead of $\mathbf{R}$ of Eq. (6) can still be bounded using the perturbation theory [15,16]. Similar to Eq. (13) this error can be written as (keeping in mind that $\mathbf{R}$ and $\hat{\mathbf{R}}$ have eigenvalues related by $\lambda_m = \mu_m + \sigma^2$ and identical eigenvectors/subspaces)

$$E^{\text{est}} \triangleq |\sin \angle(N(\hat{\mathbf{R}}), N(\mathbf{R}^{\text{est}}))|$$

$$= |\sin \angle(N(\hat{\mathbf{R}}), N(\mathbf{R}^{\text{est}}))| \leq \frac{\|\mathbf{Res}^{\text{est}}\|_2}{\text{gap}}, \quad (25)$$

where $N(\hat{\mathbf{R}})$, $N(\mathbf{R})$ and $N(\mathbf{R}^{\text{est}})$ are the noise subspaces of $\hat{\mathbf{R}}$, $\mathbf{R}$ and $\mathbf{R}^{\text{est}}$ respectively. The “gap” in the above equation is the minimum absolute gap between the signal subspace and the noise subspace eigenvalues of the matrix $\hat{\mathbf{R}}$ and hence is still as defined in Fig. 1. This time however $\mathbf{Res}^{\text{est}}$ represents the deviation of $\mathbf{R}^{\text{est}}$ eigenvalues from those of $\hat{\mathbf{R}}$ and is bounded by [15, Fact 1–11]

$$\|\mathbf{Res}^{\text{est}}\|_2 \leq \|\mathbf{R}_{\text{errors}}\|_2. \quad (26)$$

As was the case for errors due to arithmetic precision, for a specific experimentation usually estimation of $\|\mathbf{R}_{\text{errors}}\|_2$ is not possible. Hence, without loss of generality, the following assumption is made

$$\|\mathbf{R}_{\text{errors}}\|_2 \propto \|\hat{\mathbf{R}}\|_2. \quad (27)$$

Thus using Eq. (26) we can write

$$\|\mathbf{Res}^{\text{est}}\|_2 \leq \|\mathbf{R}_{\text{errors}}\|_2 = fn'(\epsilon) \|\hat{\mathbf{R}}\|_2. \quad (28)$$

where $fn'()$ is some fussy function depending on the details of the particular experimentation. ϵ was introduced in Eq. (28) to make $fn'()$ independent on the calculation arithmetic precision and to allow the utilization of the proposed array design algorithm as explained in the following. Substituting Eq. (28), $\|\hat{\mathbf{R}}\|_2 = \mu_1$ and $\text{gap} = \mu_K$ in Eq. (25) yields

$$\frac{E^{\text{est}}}{fn'(\epsilon)} \leq \epsilon \frac{\mu_1}{\mu_K}. \quad (29)$$

Comparing Eq. (29) with Eq. (16) we can write

$$E = \frac{E^{\text{est}}}{fn'(\epsilon)}. \quad (30)$$

Thus the MUSIC error of a realistic experiment, E^{est} , and the realistic environment details, $fn'()$, can be represented by a single parameter E which denotes the MUSIC error due to finite precision arithmetic of a similar ideal experiment whose exact covariance matrix is available and given by Eq. (6). Thus substituting Eq. (30) in Eq. (29) we arrive at equation identical to Eq. (16). Hence all the subsequent discussion of section 3 and the proposed array algorithm still hold and can be systematically used as will be clarified subsequently.

Example (2): In this example the effect of measurement precision will be examined. This precision is governed by the weakest point in the measurement chain, which is the analog to digital converter (ADC) [23], and is determined by the number of bits used to digitize (quantize) the analog input signal [14]. To examine this effect simulations given in example (1) is repeated here for

Fig. 8 MUSIC results for different quantization levels and arithmetic precisions.

$\tilde{d} = 0.8$, $M = 10$ and different quantization levels and arithmetic precisions as shown in Fig. 8.

As shown in Fig. 8 for 16-bit ADC, which is much less than the 24-bit and 53-bit significant precision of single and double precision floating point [24], MUSIC results obtained by single and double precision arithmetic are very similar. On the other hand, using 24-bit ADC converter, whose precision is same as the single precision floating point arithmetic, double precision MUSIC results are considerably better than single precision results. This concludes that the arithmetic precision should be fairly, and not so much, higher than measurement precision.

Example (3): Consider the problem of designing a ULA for the same wide-band and source specifications considered in example (1). The following changes are however introduced to imitate realistic acoustic environments. To simulate reverberation the last four source signals are now exactly delayed versions of the first one. Delays are given by $T/8$, $3T/8$, and $5T/8$ where T is the measurement snapshot duration. Also contiguous non-overlapping intervals of the sources frequency spectrum, $\mathbf{U}(f)$, are intentionally removed (set to zero) to simulate signals like speech where the effective number of speakers may change with time and frequency. Widths of these intervals are fixed to $(f_{\text{max}} - f_{\text{min}})/K$. Due to these adverse conditions, the more robust coherent averaging method formulated in appendix B will be used in this example.

In contrast to the ideal example (1), the following effects and sources of error are also included in this example to imitate effects usually found in realistic experimentations. Noise signals are now correlated between themselves and with the source signals with $\text{SNR} = 30$ dB. Physical array output, $p_m(t)$, is also quantized using 16-bit ADC. In addition, using the approximate Eq. (B-9) for calculating the virtual array output, $P_m^v(f)$, does add a considerable amount of error [12]. Due to these effects the ideal model of Eq. (6) no longer holds and the realistic model of Eq. (24) holds instead.

Fast Fourier Transform (FFT) is performed with maximum frequency of f_{max} and $n_{\text{FFT}} = 101$ narrow-band frequency component per FFT (measurement) snapshot.

Fig. 9 Realistic MUSIC spectrum by the spatial resampling method, for $M = 97$ and different $\tilde{d}^v$.

This yields $n_f = 91$ narrow-band frequency component in the range $[f_{\min}, f_{\max}]$. Thus as shown from Eq. (B-6) every FFT snapshot yields 91 averagings. Hence only a single FFT snapshot is used in this example which yields $\kappa(\mathbf{R}_U^F)$ around 2. To avoid aliasing sensor spacing of the physical array is set to half wavelength corresponding to $f_{\max}$, as in example (1). By using $f_0 = f_{\min}$ [8], Eq. (B-1) yields $\tilde{d}^v = f_0/f_{\max} = 0.1$.

Thus we arrive at the same array design problem considered in example (1) for single precision arithmetic. In that example when the exact covariance matrix of Eq. (6) (or equivalently Eq. (A-3) for that example) was available, MUSIC error threshold of $E_{\max} = 10^{-3}$ was enough for satisfactory MUSIC spectrum as shown in Fig. 6. The ideal version of the current example can be obtained by substituting the same source signals and DOA's of the current example into the exact covariance matrix model of Eq. (6) (or equivalently Eq. (B-7) for the current example). Using the array designed in example (1), i.e., $M = 97$, results obtained from such ideal version was found to yield results completely identical to the whole Fig. 6 and hence they are not plotted here again. Results obtained for the current non-ideal example are however plotted in Fig. 9 for different $\tilde{d}^v$. As shown in this figure, and as expected, MUSIC spectrum is worse than that of Fig. 6(c). Thus it is clear that smaller $E_{\max}$ should be used when designing arrays for realistic experiments. Appropriate choice of $E_{\max}$ is the topic of the following subsection.

4.2. Quality of DOA Estimation and Suitable Choice of $E_{\max}$

So far, the MUSIC error threshold, $E_{\max}$, of an ideal experiment whose exact covariance matrix is available and given by Eq. (6), was arbitrarily chosen as 10^{-3} . This value was shown to yield satisfactory MUSIC spectrum. On the other hand for realistic experimentations where the exact covariance matrix is instead estimated, it was shown that more appropriate choice is needed. Moreover we are

usually interested in the quality of DOA estimation rather than the ad hoc parameter E . Quality of DOA estimation can be quantified using different measures depending on the particular application. Detection probability, dynamic range, bias, variance and root mean square error are some examples of these measures. Hence, a methodology for proper choice of $E_{\max}$ for some desired quality of DOA estimation is proposed in the following steps:

a) For whatever realistic DOA estimation environment, perform an experiment (will be called calibration experiment) using some reference source signals and DOA's. Using different frequency bandwidths plot different MUSIC spectrums versus $\tilde{d}$ as those displayed in Fig. 9 for the case of example (3).

Note that similar spectrums can be obtained versus other parameters like M , $\phi_{\max}$ or $\delta\phi$. However this will be much more expensive for the case of M as an increased number of sensors will be needed. Also it will be much more difficult for the case of $\phi_{\max}$ or $\delta\phi$ as for each $\phi_{\max}$ or $\delta\phi$ value a new experimental setup is needed.

b) Using the same source signals and DOA's of the calibration experiment, calculate the exact covariance matrix using Eq. (6) (or its equivalent variant). Using this exact covariance matrix, calculate the ideal MUSIC spectrums and plot figures similar to Figs. 6(b) and 6(c).

c) By comparing Fig. 6(b) with both Figs. 6(c) and 9, obtain the relation between the MUSIC error, E , and any desired measure of DOA estimation quality, for the calibration experiment (and its corresponding ideal experiment if needed).

In this demonstration we choose the detection minimum dynamic range as the measure of DOA estimation quality. Thus the relation between E and the detection dynamic range is obtained and plotted in Fig. 10.

d) Using the obtained relation, and for the desired DOA estimation quality, choose the suitable MUSIC error threshold $E_{\max}$ for use with the proposed array design algorithm.

Fig. 10 Detection dynamic range versus MUSIC error for a realistic and its corresponding ideal experiment.

Fig. 11 Realistic MUSIC spectrum by the spatial resampling method, for $M = 156$ and different $\tilde{d}^v$.

For example using Fig. 10, for a target dynamic range of approximately 40 dB., $E_{\max} \approx 10^{-5}$ is obtained for the environment considered in example (3). Using this value in the proposed array design algorithm (using the $10^{-5}/2/\varepsilon_{\text{single}}$ contour of Fig. 4(b)), $M = 156$ is obtained. Using this number of sensors in example (3) yields the results shown in Fig. 11. In contrast to Fig. 9, Fig. 11 shows accurate DOA estimation with minimum dynamic range of ≈ 44 dB. at $\tilde{d}^v = 0.1$.

5. CONCLUSION

The MUSIC method, as an example of eigenvector based DOA estimation methods, is known to have a number of limitations. For example it assumes narrow-band signal model and its performance deteriorates with increased correlation of source signals. Hence it cannot be directly applied in acoustical environments due to their wide-band and reverberant nature. Hence a number of generalizations were proposed to overcome these limitations. Most wide-band generalizations share in applying the plain MUSIC algorithm on an array whose normalized sensor spacing is considerably less than unity. In addition, reverberation not only produce highly correlated sources but also increase the number of sources and decrease their separation.

Considering these characteristics, applicability of the MUSIC method was considered in this paper. MUSIC error was expressed in terms of the number of sources, sources correlation level, maximum resolution, maximum source angle, number of sensors, sensor spacing and arithmetic precision. As a consequence, efficient array design algorithm, capable of determining minimum number of sensors, was established. This design algorithm was shown to be seamlessly applicable to realistic environments where many additional sources of error and violations to the ideal signal model usually exist. Hence this algorithm is indispensable for situations requiring many sensors such as acoustical environments. Also it was shown that the

arithmetic precision should be fairly, and not so much, higher than measurement precision.

ACKNOWLEDGMENT

The authors would like to express their deep gratitude for the referees for their valuable suggestions that much improved this paper.

REFERENCES

- [1] V. S. U. Pillai, *Array Signal Processing* (Springer Verlag, New York, 1989).
- [2] G. C. Carter, *Coherence and Time Delay Estimation* (IEEE Press, Piscataway, N.J., 1993).
- [3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *Proc. RADC Spectral Est. Workshop 1979*, pp. 243–258 (1979).
- [4] R. Roy and T. Kailath, "ESPRIT — Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, **37**, 984–995 (1989).
- [5] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp. Electron. Syst.*, **AES-19**, 134–139 (1983).
- [6] G. Su and M. Morf, "Signal subspace approach for multiple wide-band emitter location," *IEEE Trans. Acoust. Speech Signal Process.*, **31**, 1502–1522 (1983).
- [7] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust. Speech Signal Process.*, **33**, 823–831 (1985).
- [8] J. Krolik and D. Swingler, "Focused wide-band array processing by spatial resampling," *IEEE Trans. Acoust. Speech Signal Process.*, **38**, 356–360 (1990).
- [9] B. Friedlander and A. J. Weiss, "Direction finding for wide-band signals using an interpolated array," *IEEE Trans. Signal Process.*, **41**, 1618–1634 (1993).
- [10] M. Doron, E. Doron and H. Weiss, "Coherent wide-band processing for arbitrary array geometry," *IEEE Trans. Signal Process.*, **41**, 414–417 (1993).
- [11] K. M. Buckley and L. J. Griffiths, "Broad-band signal-subspace spatial-spectrum (BASS-ALE) estimation," *IEEE Trans. Acoust. Speech Signal Process.*, **36**, 953–964 (1988).
- [12] A. Desoki, J. Takada and I. Hagiwara, "Improved finite samples resampling for unbiased wide-band direction of arrival estimation," submitted to *Acoust. Sci. & Tech.*
- [13] M. Kaveh and A. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust. Speech Signal Process.*, **34**, 331–341 (1986).
- [14] S. W. Smith, *The Scientist's and Engineer's Guide to Digital Signal Processing* (California Technical Publishing, San Diego, Calif., 1999).
- [15] G. H. Golub and C. F. Van Loan, *Matrix Computations* (The John Hopkins University Press, Baltimore, 1996).
- [16] G. B. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall Inc., Englewood Cliffs, 1980).
- [17] C. Davis and W. Kahan, "The rotation of eigenvectors by a perturbation III," *SIAM J. Numer. Anal.*, **7**, 1–46 (1970).
- [18] G. W. Stewart, "Error and perturbation bounds for subspaces associated with certain eigenvalue problems," *SIAM Rev.*, **15**, 727–764 (1973).
- [19] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov and D. Sorensen, "Further details: Error bounds

- for the symmetric eigenproblem” in *LAPACK Users’ Guide* (SIAM, Philadelphia, 1999).
- [20] I. S. Dhillon and B. Parlett, “Orthogonal eigenvectors and relative gaps,” *SIAM J. Matrix Anal. Appl.*, **25**, 858–899 (2004).
 - [21] A. S. Deif, *Advanced Matrix Theory for Scientists and Engineers* (Abacus/Gordon & Breach, London, 1991).
 - [22] S. Leach, “Singular value decomposition—A primer,” *Unpublished Manuscript, Department of Computer science, Brown University, Providence, RI 02912*.
 - [23] O. T. Anderson and N. J. Jacobsen, “New technology increases the dynamic ranges of data acquisition systems based on 24-bit technology,” *Sound Vib.*, **39**(4), pp. 8–14 (2005).
 - [24] IEEE, *IEEE Standard for Binary Floating-Point Arithmetic. (ANSI/IEEE Std 754-1985)* (Institute of Electrical and Electronics Engineers, New York, 1985).
 - [25] G. Bienvenu, P. Fierxer, G. Vezzosi, L. Kopp and F. Florin, “Coherent wide band high resolution processing for linear array,” *Proc. Signal Proc., ICASSP 89*, pp. 2799–2802 (1989).
 - [26] H. Wang, C. C. Li and J. X. Zhu, “High-resolution direction finding in the presence of multipath: A frequency-domain smoothing approach,” *Proc. ICASSP 87*, pp. 2276–2279 (1987).
 - [27] J. X. Zhu and H. Wang, “A comparison of smoothing approaches for angle measurement,” *IEEE Trans. Aerosp. Electron. Syst.*, **25**, 529–535 (1989).
 - [28] J. Krolik and D. Swinger, “The performance of minimax spatial resampling filters for focusing wide-band arrays,” *IEEE Trans. Acoust. Speech Signal Process.*, **39**, 1899–1903 (1991).

APPENDIX A: INCOHERENT AVERAGING

By writing (Eq. (10)) more compactly in vector form yields

$$\mathbf{P}(f) = \mathbf{A}(f)\mathbf{U}(f) + \mathbf{N}(f) \quad (\text{A.1})$$

where $\mathbf{P}(f)$, $\mathbf{U}(f)$ and $\mathbf{N}(f)$ are the Fourier transform of $\mathbf{p}(f)$, $\mathbf{u}(f)$ and $\mathbf{n}(f)$ of (Eq. (3)) respectively, and $\mathbf{A}(f)$ is exactly as defined in (Eq. (3)) but is now frequency dependent. As a consequence, the output covariance matrix becomes also frequency dependent as

$$\mathbf{R}(f) = E[\mathbf{P}(f)\mathbf{P}^\dagger(f)] \quad (\text{A.2})$$

Substituting (Eq. (A.1)) in (Eq. (A.2)) and assuming elements of $\mathbf{N}(f)$ are also uncorrelated between themselves and with source signals, the covariance matrix takes the form

$$\mathbf{R}(f) = \mathbf{A}(f)\mathbf{R}_U(f)\mathbf{A}^\dagger(f) + \sigma^2(f)\mathbf{I} \quad (\text{A.3})$$

with $\mathbf{R}_U(f)$ defined as

$$\mathbf{R}_U(f) = E[\mathbf{U}(f)\mathbf{U}^\dagger(f)] \quad (\text{A.4})$$

Using $\mathbf{R}(f)$ of (Eq. (A.3)) instead of $\mathbf{R}$ of (Eq. (6)), MUSIC results can similarly be calculated using (Eq. (9)) for any frequency f . These results can then be averaged for different frequencies for wide-band DOA estimation [6]. This incoherent averaging approach is however slow and fails for highly correlated signals and has some other disadvantages [7].

APPENDIX B: COHERENT AVERAGING

These methods are called coherent since they use all the frequency components of a wide-band signal coherently. As compared to incoherent methods, advantages of coherent averaging include the ability to handle highly correlated, and even coherent, signals, statistical stability and less measurement and computational time. The concept of coherent averaging was first introduced by Wang and Kaveh in [7]. Since then a number of different methods following the same concept were introduced as in [8–12]. In the following coherent averaging by the spatial resampling method [8] will be considered as an example of coherent averaging methods.

In this method a set of virtual arrays are used instead of the physical array. Positions of sensors of these virtual arrays, x_m^v , are chosen so that the normalized sensor position, $\tilde{x}_m^v$, is independent on frequency. That is

$$\tilde{x}_m^v = \frac{2f}{c}x_m^v(f) = \text{const} = \frac{2f_0}{c}x_m, \quad (\text{B.1})$$

where f_0 is some reference frequency called “focusing frequency” and x_m denotes the physical sensor position. From now on the superscript “v” will be used to denote properties related to the virtual arrays. Thus it is clear that virtual sensors’ positions are inversely proportional to frequency

$$x_m^v(f) = \frac{f_0}{f}x_m. \quad (\text{B.2})$$

Thus at a frequency f , the wave field received by the corresponding virtual array can be written similar to (Eq. (10)) as

$$\begin{aligned} P_m^v(f) &= \sum_{k=1}^K U_k(f) e^{j\pi \tilde{x}_m^v \sin \phi_k} + N_m^v(f) \\ &= \sum_{k=1}^K U_k(f) e^{2j\pi x_m f_0 \sin \phi_k / c} + N_m^v(f), \end{aligned} \quad (\text{B.3})$$

where $N_m^v(f)$ denotes the noise signal generated at the m -th virtual sensor in the frequency domain.

Since $\tilde{x}_m^v$ is independent on frequency, Eq. (B.3) is problem free as compared to Eq. (10). Vector form of Eq. (B.3) can be written as

$$\mathbf{P}^v(f) = \mathbf{A}^v(f_0)\mathbf{U}(f) + \mathbf{N}^v(f), \quad (\text{B.4})$$

where $\mathbf{P}^v(f) = [P_1^v(f), P_2^v(f), \dots, P_M^v(f)]^T$, $\mathbf{A}^v(f_0) = \sqrt{M}[\mathbf{a}^v(\phi_1), \mathbf{a}^v(\phi_2), \dots, \mathbf{a}^v(\phi_K)]$ and $\mathbf{N}^v(f) = [N_1^v(f), N_2^v(f), \dots, N_M^v(f)]^T$. The vector $\mathbf{a}^v(\phi_k)$ is as defined in Eq. (4) but with $\tilde{x}_m$ replaced with $\tilde{x}_m^v$ of Eq. (B.1), that is

$$\mathbf{a}^v(\phi_k) = \frac{[e^{2j\pi \tilde{x}_1^v f_0 \sin \phi_k / c}, e^{2j\pi \tilde{x}_2^v f_0 \sin \phi_k / c}, \dots, e^{2j\pi \tilde{x}_M^v f_0 \sin \phi_k / c}]^T}{\sqrt{M}}. \quad (\text{B.5})$$

Thus for different frequencies, covariance matrices of the corresponding virtual arrays can be coherently averaged (focused) to obtain the focused covariance matrix defined by

$$\mathbf{R}^F = E \left[\frac{1}{n_f} \sum_f \mathbf{P}^v(f) (\mathbf{P}^v(f))^\dagger \right], \quad (\text{B.6})$$

where the superscript “F” denotes focused covariance matrix and n_f is the number of narrow-band frequency components within the considered wide frequency band. Thus the focused covariance matrix can be written as

$$\mathbf{R}^F = \mathbf{A}^v(f_0) \mathbf{R}_U^F (\mathbf{A}^v(f_0))^\dagger + (\sigma^v(f))^2 \mathbf{I}, \quad (\text{B.7})$$

where $\mathbf{R}_U^F$ is the focused sources covariance in frequency domain defined by

$$\mathbf{R}_U^F = E \left[\frac{1}{n_f} \sum_f \mathbf{U}(f) (\mathbf{U}(f))^\dagger \right], \quad (\text{B.8})$$

and $(\sigma^v(f))^2 \mathbf{I}$ represents noise power common to all the virtual sensors in frequency domain. Note that the summation over frequency in Eq. (B.8) implicitly performs

frequency smoothing [26,27] for $\mathbf{R}_U^F$ and consequently for $\mathbf{R}^F$. This smoothing step is why coherent averaging methods can handle fully correlated signals and have the advantages of statistical stability, less measurement and computational time [7,26,27].

Finally by using $\mathbf{R}^F$ of Eq. (B.7), $\mathbf{A}^v(f_0)$ and $\mathbf{a}^v(\phi)$ of Eq. (B.5) instead of $\mathbf{R}$ of Eq. (6), $\mathbf{A}$ and $\mathbf{a}(\phi)$ of Eq. (4), MUSIC results can similarly be calculated using Eq. (9).

B.1. Calculation of Virtual Arrays' Output

Output of the virtual arrays, $P_m^v(f)$, can be calculated by interpolating (resampling) the output of the physical array [25]. Krolik and Swingler in [28] proposed an optimum linear shift-variant resampling filter for calculating $P_m^v(f)$. Using their resampling filter, $P_m^v(f)$ can be approximated as

$$P_m^v(f) \approx \sum_{m=-\lfloor M/2 \rfloor}^{\lfloor M/2 \rfloor} P_m(f) \frac{\sin(\pi(nf_0/f - m))}{\pi(nf_0/f - m)}, \quad (\text{B.9})$$

with $\lfloor M/2 \rfloor$ is the largest integer smaller than or equal to $M/2$.