

ROBUST ADAPTIVE BEAMFORMING USING A FULLY DATA-DEPENDENT LOADING TECHNIQUE

C.-C. Huang¹ and J.-H. Lee^{2,*}

¹Graduate Institute of Communication Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

²Department of Electrical Engineering, Graduate Institute of Communication Engineering, and Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

Abstract—This paper deals with adaptive array beamforming in the presence of errors due to steering vector mismatch and finite sample effect. Diagonal loading (DL) is one of the widely used techniques for dealing with these errors. However, the main drawback of DL techniques is that there is not an easy and reliable manner to determine the required loading factor. Recently, several DL approaches proposed the so-called automatic scheme on computing the required loading factor. In this paper, we propose a fully data-dependent loading to overcome the difficulties. The novelty is that the proposed method does not require any additional sophisticated scheme to choose the required loading. The loading factor can be completely obtained from the received array data. Analytical formulas for evaluating the performance of the proposed method under random steering vector error are further derived. Simulation results are provided to confirm the validity of the proposed method and make comparison with the existing DL methods.

1. INTRODUCTION

For the well-known Capon beamformer or minimum power distortionless response (MPDR) beamformer, the adaptive weights are calculated by minimizing the beamformer's output power subject to the constraint that forces the array to make a constant response in the steering direction [1–3]. When the ensemble data correlation matrix and the

Received 4 November 2011, Accepted 26 December 2011, Scheduled 31 December 2011

* Corresponding author: Ju-Hong Lee (juhong@cc.ee.ntu.edu.tw).

actual steering vector of desired signal are available, the solution to the constrained minimization problem is the optimal one that maximizes the array output signal-to-interference-plus-noise ratio (SINR). In practice, the ensemble correlation matrix is unavailable, we resort to using a sample matrix inversion (SMI) approach which solves the constrained minimization problem using a sample correlation matrix instead of the ensemble one. However, the sample correlation matrix may be inaccurately estimated or even ill-conditioned due to finite data samples [4, 5]. This leads to performance degradation of the SMI. Moreover, the SMI is very sensitive to the accuracy of the steering vector of the desired signal. It has been shown that even a small mismatch between the presumed steering vector and the actual one causes significant performance degradation [6, 7]. Therefore, adaptive beamforming with robustness against the errors mentioned above becomes popular.

Diagonal loading (DL) is one of the widely used techniques to improve robustness of the SMI against the errors [4, 8–12], where a scaled identity matrix is added to the sample correlation matrix. Although the DL is effective, the main drawback is that choosing the required loading factor is not an easy task. For example, in [10], the loading factor is found by using particle filters. The particle who has the highest posterior probability is chosen as the optimal loading factor. In [11], the loading factor is obtained by controlling the peak location of the main beam. However, the loading factors of those methods cannot be obtained analytically and have to be solved numerically. Furthermore, in [12], an analytical expression of the optimal loading factor is derived by maximizing the output SINR in the presence of random steering vector error, the main disadvantage is that the obtained negative loading factor may lead to a rank-deficient problem if it equals an eigenvalue of the correlation matrix. On the other hand, it follows from [13–15] that the loading factor can be calculated based on the uncertainty set of the steering vector. However, one still needs to specify the parameter related to the size of the uncertainty set, and it may be difficult to choose the parameter in practice. Unlike the conventional DL, an approach called the variable loading (VL) has been considered in [16, 17], and has shown its advantage over the conventional DL due to using the variable loading, where different loading factor is added to each eigenvalue of the sample correlation matrix instead of a fixed loading factor for all of the eigenvalues. However, one still needs to determine the variable loading and the optimal choice of this loading seems improbable. Recently, several novel loading techniques were interested in automatically computing the required loading [18–20]. For instance, a ridge regression based

method named HKB is proposed by [18] and the required loading is automatically computed by the available data. However, it has been shown in [19] that the HKB may have an inherent problem in choosing the required loading factor, which may be very large for a relatively large data snapshots. Then, an alternative parameter-free method is presented by [19]. This method uses a general linear combination (GLC) shrinkage-based correlation matrix instead of the original one. The required loading factor is automatically chosen by minimizing the mean squared error (MSE) of the shrinkage-based correlation matrix. However, the loading factor decreases or the effect of the robustness diminishes as the number of data snapshots increases. This may incur some problems in the presence of steering vector error. Note that both HKB and GLC are only efficient in the situation of small data snapshots. To generalize the concept of the HKB and GLC, an automatic generalized loading (AGL) method is presented in [20] by using a generalized Hermitian matrix which can be automatically computed by minimizing the MSE of a generalized sidelobe canceler (GSC) reparameterized vector. The AGL has been shown to be more robust to steering vector error over GLC and HKB. However, the computational complexity of the AGL is normally increased due to using a generalized loading matrix. Besides, theoretical analysis for each of the above methods is not available in the literature. Moreover, a comprehensive review of several parameter-free robust adaptive beamforming algorithms is provided by [21].

In this paper, we propose a fully data-dependent approach to deal with steering vector error and finite sample effect. The proposed method generates a novel loaded sample correlation matrix. Instead of determining the optimal loading matrix, we compute the loading matrix directly from the sample correlation matrix of the received array data. This leads to that the resulting loading factors are related only to the eigenvalues of the sample correlation matrix. Moreover, we show that the influence of the least significant eigenvalues which are due to noise can be effectively diminished, whereas the influence of the significant eigenvalues which are mainly contributed by signal sources can remain almost the same. As a result, the proposed method automatically provides large loading factors for the least significant eigenvalues and small ones for the significant eigenvalues. This is a significant advantage over the conventional DL techniques and achieves significant robustness against steering vector error and finite sample effect because of using a different loading for a different eigenvalue. To evaluate the performance of the proposed method, analytical formulas for the array output SINR under random steering vector error are derived. Finally, we provide several simulation examples to confirm

the validity of the proposed method and make comparison with the existing methods.

This paper is organized as follows. In Section 2, we briefly describe the performance degradation due to finite sample effect and the principle of the conventional DL techniques. Section 3 presents the fully data-dependent loading method. Analytical formulas of array output SINR for the proposed method are derived in Section 4. Simulation examples for confirming the validity and effectiveness of the proposed method are provided in Section 5. Finally, we conclude the paper in Section 6.

2. PROBLEM FORMULATION

Consider that there are K far-field signal sources including a desired signal and $K - 1$ interferers impinging on an M -element antenna array. The received data vector $\mathbf{x}(t)$ can be expressed as

$$\mathbf{x}(t) = \mathbf{a}_0 s_0(t) + \sum_{k=1}^{K-1} \mathbf{a}_k s_k(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{x}(t) = [x_1(t) \ x_2(t) \ \dots \ x_M(t)]^T \in \mathbb{C}^{M \times 1}$, $x_m(t)$, $m = 1, 2, \dots, M$, is the output of the m th antenna element; $s_k(t)$ denotes the k th signal with zero mean and variance σ_k^2 ; $\mathbf{a}_k = [a_1(\phi_k) \ a_2(\phi_k) \ \dots \ a_M(\phi_k)]^T \in \mathbb{C}^{M \times 1}$ represents the $M \times 1$ steering vector from angle ϕ_k off array broadside; $\mathbf{n}(t) \in \mathbb{C}^{M \times 1}$ is an additive white Gaussian noise vector with zero mean and covariance matrix $\sigma_n^2 \mathbf{I}$. In the paper, we consider the case that each of antenna elements within the array has an isotropic response. For the case of an antenna array with non-isotropic elements, synthesis of a desirable beam pattern is more complicated, however the techniques presented by [25, 26] are applicable.

Based on the Capon beamformer, the optimal weight vector is obtained by minimizing the array output power subject to the main-beam constraint [1, 2].

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \quad \text{subject to} \quad \mathbf{w}^H \mathbf{a}_0 = 1. \quad (2)$$

The solution to (2) is given by

$$\mathbf{w}_o = \frac{\mathbf{R}_{xx}^{-1} \mathbf{a}_0}{\mathbf{a}_0^H \mathbf{R}_{xx}^{-1} \mathbf{a}_0} \quad (3)$$

where $\mathbf{R}_{xx} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\}$ is the ensemble correlation matrix of $\mathbf{x}(t)$. In practice, \mathbf{R}_{xx} is unavailable and the knowledge of \mathbf{a}_0 may be

inaccurate. A sample matrix inversion (SMI) approach is commonly used to solve the constrained minimization problem of (2) by using a sample correlation matrix instead of the ensemble one. Under the actual steering vector \mathbf{a} , the solution of (2) can be expressed as

$$\mathbf{w}_{\text{smi}} = \frac{\hat{\mathbf{R}}_{xx}^{-1} \mathbf{a}}{\mathbf{a}^H \hat{\mathbf{R}}_{xx}^{-1} \mathbf{a}}. \tag{4}$$

where $\mathbf{a} \neq \mathbf{a}_0$ due to steering vector error. The sample correlation matrix $\hat{\mathbf{R}}_{xx}$ is computed from the received data vector $\mathbf{x}(t)$ as follows:

$$\hat{\mathbf{R}}_{xx} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}(t_n) \mathbf{x}^H(t_n) \tag{5}$$

where N denotes the number of data snapshots and t_n the n th time instant. Without steering vector error, \mathbf{w}_{smi} converges to \mathbf{w}_o as N increases. To see the effect due to finite samples, we let $\hat{\mathbf{R}}_{xx}$ be decomposed as

$$\hat{\mathbf{R}}_{xx} = \sum_{i=1}^M \hat{\lambda}_i \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H \tag{6}$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_M$ are the eigenvalues, $\hat{\mathbf{e}}_i, i = 1, 2, \dots, M$, are the corresponding eigenvectors. The eigenvalues of (6) can be written as [2]

$$\hat{\lambda}_i = \begin{cases} \hat{\lambda}_i^s + \hat{\lambda}_{\min}, & i = 1, 2, \dots, K \\ \hat{\lambda}_i^c + \hat{\lambda}_{\min}, & i = K + 1, K + 2, \dots, M - 1 \\ \hat{\lambda}_{\min}, & i = M. \end{cases} \tag{7}$$

We note that the significant eigenvalues $\hat{\lambda}_i, i = 1, \dots, K$, are mainly contributed by the signal sources and the $M - K$ least significant eigenvalues are contributed by noise. $\hat{\lambda}_{\min}$ denotes the minimum eigenvalue and $\hat{\lambda}_i^s$ represent the estimated eigenvalues due to the signal sources. $\hat{\lambda}_i^c$ denote the differences between the other $M - K - 1$ noise eigenvalues and $\hat{\lambda}_{\min}$. The inverse of $\hat{\mathbf{R}}_{xx}$ can be written as

$$\hat{\mathbf{R}}_{xx}^{-1} = \sum_{i=1}^M \frac{1}{\hat{\lambda}_i} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H = \sum_{i=1}^K \frac{1}{\hat{\lambda}_i^s + \hat{\lambda}_{\min}} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H + \sum_{k=K+1}^{M-1} \frac{1}{\hat{\lambda}_i^c + \hat{\lambda}_{\min}} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H + \frac{1}{\hat{\lambda}_{\min}} \hat{\mathbf{e}}_M \hat{\mathbf{e}}_M^H. \tag{8}$$

In general, the noise eigenvalues (least significant ones) are relative small and thus the inverses of the noise eigenvalues are dominant in (8) as compared to the inverses of $\hat{\lambda}_i^s, i = 1, 2, \dots, K$, especially the inverse of $\hat{\lambda}_{\min}$ when N is very small. In the presence of steering vector error,

the inverses of the noise eigenvalues would result in a \mathbf{w}_{smi} with large norm or a high level of sidelobe.

In the literature, diagonal loading (DL) is the well-known technique to overcome the drawback. The principle of DL techniques is to add a scaled identity matrix to the sample correlation matrix to obtain

$$\mathbf{R}_{\text{dl}} = \hat{\mathbf{R}}_{xx} + \gamma \mathbf{I} \quad (9)$$

and the corresponding weight vector is given by

$$\mathbf{w}_{\text{dl}} = \frac{\mathbf{R}_{\text{dl}}^{-1} \mathbf{a}}{\mathbf{a}^H \mathbf{R}_{\text{dl}}^{-1} \mathbf{a}} = \frac{(\hat{\mathbf{R}}_{xx} + \gamma \mathbf{I})^{-1} \mathbf{a}}{\mathbf{a}^H (\hat{\mathbf{R}}_{xx} + \gamma \mathbf{I})^{-1} \mathbf{a}}. \quad (10)$$

The inverse of \mathbf{R}_{dl} can be expressed as

$$\mathbf{R}_{\text{dl}}^{-1} = \sum_{i=1}^K \frac{1}{\hat{\lambda}_i^s + \hat{\lambda}_{\min} + \gamma} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H + \sum_{k=K+1}^{M-1} \frac{1}{\hat{\lambda}_k^e + \hat{\lambda}_{\min} + \gamma} \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^H + \frac{1}{\hat{\lambda}_{\min} + \gamma} \hat{\mathbf{e}}_M \hat{\mathbf{e}}_M^H \quad (11)$$

where γ denotes the loading factor which should be appropriately determined. \mathbf{I} is the identity matrix with an appropriate size. Since the inverse of $(\hat{\lambda}_i^e + \hat{\lambda}_{\min} + \gamma)$ or $(\hat{\lambda}_{\min} + \gamma)$ is not larger than $1/\gamma$, the influence of the noise eigenvalues is restricted and the norm of \mathbf{w}_{dl} will not tend to be large. Although the DL can effectively cure the problem due to noise eigenvalue, the main shortcoming is that there is no easy and reliable scheme for choosing an appropriate γ .

3. PROPOSED METHOD

In this section, we propose a novel method with a fully data-dependent loading to tackle finite sample effect and steering vector error. Consider the following loading added to the sample correlation matrix $\hat{\mathbf{R}}_{xx}$

$$\mathbf{R}_{\text{gl}} = \hat{\mathbf{R}}_{xx} + \mathbf{Q} \quad (12)$$

where \mathbf{Q} is a loading matrix to be determined. The optimal weight vector based on \mathbf{R}_{gl} is therefore given by

$$\mathbf{w}_{\text{gl}} = \frac{\mathbf{R}_{\text{gl}}^{-1} \mathbf{a}}{\mathbf{a}^H \mathbf{R}_{\text{gl}}^{-1} \mathbf{a}} = \frac{(\hat{\mathbf{R}}_{xx} + \mathbf{Q})^{-1} \mathbf{a}}{\mathbf{a}^H (\hat{\mathbf{R}}_{xx} + \mathbf{Q})^{-1} \mathbf{a}}. \quad (13)$$

In the literature, several techniques require a *priori* information for choosing \mathbf{Q} . In the AGL of [20], \mathbf{Q} is set to a Hermitian matrix related to GSC parameterization of \mathbf{w}_{smi} . On the other hand, [22] chooses \mathbf{Q} according to the covariance matrix of steering vector error. Moreover, theoretical analyses on the performance of array beamformers using these techniques are not available.

3.1. Concept of Fully Data-dependent Loading

Here, we present a novel method for determining the loading matrix \mathbf{Q} . First, consider the case of $\mathbf{Q} = \hat{\mathbf{R}}_{xx}^{-1}$. (12) becomes

$$\tilde{\mathbf{R}}_{gl} = \hat{\mathbf{R}}_{xx} + \hat{\mathbf{R}}_{xx}^{-1}. \tag{14}$$

Accordingly, \mathbf{w}_{gl} of (13) is given by

$$\mathbf{w}_{pr,1} = \frac{\tilde{\mathbf{R}}_{gl}^{-1} \mathbf{a}}{\mathbf{a}^H \tilde{\mathbf{R}}_{gl}^{-1} \mathbf{a}} = \frac{(\hat{\mathbf{R}}_{xx} + \hat{\mathbf{R}}_{xx}^{-1})^{-1} \mathbf{a}}{\mathbf{a}^H (\hat{\mathbf{R}}_{xx} + \hat{\mathbf{R}}_{xx}^{-1})^{-1} \mathbf{a}}. \tag{15}$$

The inverse of $\tilde{\mathbf{R}}_{gl}$ can be expressed as

$$\tilde{\mathbf{R}}_{gl}^{-1} = \sum_{i=1}^K \frac{1}{\hat{\lambda}_i^s + \hat{\lambda}_{\min} + \alpha_i} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H + \sum_{i=K+1}^{M-1} \frac{1}{\hat{\lambda}_i^s + \hat{\lambda}_{\min} + \alpha_i} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H + \frac{1}{\hat{\lambda}_{\min} + \alpha_M} \hat{\mathbf{e}}_M \hat{\mathbf{e}}_M^H \tag{16}$$

with

$$\alpha_i = \frac{1}{\hat{\lambda}_i} = \begin{cases} \frac{1}{\hat{\lambda}_i^s + \hat{\lambda}_{\min}}, & i = 1, 2, \dots, K \\ \frac{1}{\hat{\lambda}_i^s + \hat{\lambda}_{\min}}, & i = K + 1, K + 2, \dots, M - 1 \\ \frac{1}{\hat{\lambda}_{\min}}, & i = M. \end{cases} \tag{17}$$

Comparing (11) and (16), we note that this method belongs to a DL with the loading factor equal to α_i . However, the main advantage of (16) is that it does not need to choose α_i since α_i are completely obtained by the inverses of the eigenvalues of $\hat{\mathbf{R}}_{xx}$ as shown by (6). In the presence of a very small $\hat{\lambda}_{\min}$, α_M tends to be very large and the influence of $(\hat{\lambda}_{\min} + \alpha_M)$ significantly diminishes. For a very large $\hat{\lambda}_i^s$, α_i is relatively small and the influence of $(\hat{\lambda}_i^s + \alpha_i)$ is almost the same as that of $\hat{\lambda}_i^s$. This implies that α_i becomes large for the least significant eigenvalues and small for the significant ones automatically. Comparing (16) with (8), we can see that $\tilde{\mathbf{R}}_{gl}^{-1}$ of (16) is less sensitive to the noise eigenvalues than $\hat{\mathbf{R}}_{xx}^{-1}$ of (8). Hence, we would expect that $\mathbf{w}_{pr,1}$ can achieve better robustness than \mathbf{w}_{smi} . As compared with \mathbf{R}_{dl}^{-1} of (11), we observe that $1/(\hat{\lambda}_i + \alpha_i)$ is closer to $1/\hat{\lambda}_i$ than $1/(\hat{\lambda}_i + \gamma)$ when $\hat{\lambda}_i \geq \gamma$. $1/(\hat{\lambda}_i + \alpha_i)$ is less than $1/(\hat{\lambda}_i + \gamma)$ when $\hat{\lambda}_i < \gamma$. Therefore, the inverses of the significant eigenvalues will be more accurate than those of the DL techniques. Moreover, the inverses of the least significant eigenvalues will be condensed by the proposed method. Consequently, using $\mathbf{w}_{pr,1}$ of (15) leads to some advantages over using \mathbf{w}_{dl} of (10).

Finally, we note that $\mathbf{w}_{pr,1}$ is equivalent to the solution of the variable loading (VL) method of [16] or [17] when the loading factor δ

of the VL is equal to one. The VL has shown its better performance over the DL due to using a variable loading for each of the eigenvalues instead of a fixed γ for all of the eigenvalues as shown by (9). However, further endeavor on determining an appropriate δ is required.

3.2. General Formulation of Fully Data-dependent Loading

Based on the theoretical achievement of Section 3.1, we further consider a general loading matrix $\mathbf{Q} = \sum_{k=1}^p \hat{\mathbf{R}}_{xx}^{-k}$ and define a matrix $\bar{\mathbf{R}}_{\text{gl}} = \mathbf{R}_{\text{gl}} |_{\mathbf{Q}=\sum_{k=1}^p \hat{\mathbf{R}}_{xx}^{-k}}$ as follows:

$$\bar{\mathbf{R}}_{\text{gl}} = \hat{\mathbf{R}}_{xx} + \sum_{k=1}^p \hat{\mathbf{R}}_{xx}^{-k}. \quad (18)$$

The corresponding weight vector is given by

$$\mathbf{w}_{\text{pr}} = \frac{\bar{\mathbf{R}}_{\text{gl}}^{-1} \mathbf{a}}{\mathbf{a}^H \bar{\mathbf{R}}_{\text{gl}}^{-1} \mathbf{a}} = \frac{(\hat{\mathbf{R}}_{xx} + \sum_{k=1}^p \hat{\mathbf{R}}_{xx}^{-k})^{-1} \mathbf{a}}{\mathbf{a}^H (\hat{\mathbf{R}}_{xx} + \sum_{k=1}^p \hat{\mathbf{R}}_{xx}^{-k})^{-1} \mathbf{a}}. \quad (19)$$

We note that \mathbf{w}_{pr} is the same as \mathbf{w}_{smi} of (4) when $p = 0$. Otherwise, it follows from (6) and (18) that α_i can now be expressed as

$$\alpha_i = \sum_{k=1}^p \frac{1}{\hat{\lambda}_i^k} = \begin{cases} p, & \text{if } \hat{\lambda}_i = 1 \\ \frac{1 - \hat{\lambda}_i^{-p}}{\hat{\lambda}_i - 1}, & \text{if } \hat{\lambda}_i \neq 1. \end{cases} \quad (20)$$

It is clear that α_i provides a fixed loading factor when the eigenvalue $\hat{\lambda}_i$ is equal to one. In contrast, α_i is determined by $\hat{\lambda}_i$ when $\hat{\lambda}_i \neq 1$. In the presence of small eigenvalues $\hat{\lambda}_i$, α_i increases its effect as p increases. On the other hand, α_i decreases its effect for large $\hat{\lambda}_i$. This property provides some advantages over the conventional DL techniques with a fixed loading factor for each of the eigenvalues. Moreover, using (18) preserves the advantage of using (14) with more degrees of freedom in loading.

Since α_i is fully dependent on $\hat{\lambda}_i$, the only assumption of the proposed method is that the eigenvalues related to the signal sources are required to be greater than one, i.e., $\hat{\lambda}_i > 1$ for $i = 1, 2, \dots, K$, and the eigenvalues related to noise are required to be not greater than one, i.e., $\hat{\lambda}_i \leq 1$ for $i = K + 1, K + 2, \dots, M$. Under this assumption, we can have that

$$\mathbf{w}_{\text{pr}} \propto \sum_{i=1}^K \frac{\mathbf{e}_i^H \mathbf{a}}{\hat{\lambda}_i + \Delta \hat{\lambda}_i} \mathbf{e}_i \quad (21)$$

when p approaches infinity, where $\Delta\hat{\lambda}_i = \frac{1}{\lambda_i - 1}$. We note that the components corresponding to the noise eigenvalues are completely eliminated by using α_i . Moreover, the summation term of (21) can be approximated by

$$\sum_{i=1}^K \frac{\mathbf{e}_i^H \mathbf{a}}{\hat{\lambda}_i} \mathbf{e}_i \tag{22}$$

when $\Delta\hat{\lambda}_i$ is small enough as compared with $\hat{\lambda}_i$. We note that (22) is the solution of the eigenspace-based beamformer (ESB) [7, 23]. Hence, the proposed method of (19) is similar to the ESB when $\hat{\lambda}_i$, $i = 1, \dots, K$, are significantly large. The ESB has been widely realized as one of the most powerful robust methods against arbitrary steering vector error [7]. Nevertheless, the proposed method dose not require any information regarding the signal subspace or noise subspace of $\hat{\mathbf{R}}_{xx}$. In the next section, we analyze the performance of the proposed method in the presence of random steering vector error.

4. PERFORMANCE ANALYSIS

In this section, we evaluate the performance of the proposed method in terms of array output SINR. For simplicity, we consider an M -element uniform linear array (ULA) excited by a desired signal and an interferer. The output SINR is defined as follows:

$$\text{SINR} = \frac{P_d}{P_o - P_d} = \frac{\mathbf{w}^H \mathbf{R}_d \mathbf{w}}{\mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} - \mathbf{w}^H \mathbf{R}_d \mathbf{w}} \tag{23}$$

where $P_d = \mathbf{w}^H \mathbf{R}_d \mathbf{w}$ denotes the output power of the desired signal with $\mathbf{R}_d = \sigma_{s_0}^2 \mathbf{a}_0 \mathbf{a}_0^H$, $P_o = \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w}$ represents the array output power. In the presence of random steering vector error, the actual steering vector \mathbf{a} is defined as

$$\mathbf{a} = \mathbf{a}_0 + \sigma_e \mathbf{\Delta} \tag{24}$$

where $\mathbf{\Delta}$ is a random error vector and σ_e is a proportional factor. Without loss of generality, we assume that the elements of $\mathbf{\Delta}$ are independent complex Gaussian with zero mean and unit variance, i.e., $E\{\mathbf{\Delta} \mathbf{\Delta}^H\} = \mathbf{I}$. Without considering finite-sample effect, we decompose the ensemble correlation matrix \mathbf{R}_{xx} as follows:

$$\mathbf{R}_{xx} = \sum_{i=0}^{M-1} \lambda_i \mathbf{e}_i \mathbf{e}_i^H = \mathbf{E}_s \mathbf{\Lambda}_s \mathbf{E}_s^H + \mathbf{E}_n \mathbf{\Lambda}_n \mathbf{E}_n^H \tag{25}$$

where $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{M-1}$ are the eigenvalues, \mathbf{e}_i , $i = 0, 1, \dots, M-1$, are the corresponding eigenvectors; $\mathbf{\Lambda}_s = \text{diag}\{\lambda_0 \lambda_1\}$; $\mathbf{\Lambda}_n = \sigma_n^2 \mathbf{I}$;

$\mathbf{E}_s = [\mathbf{e}_0 \ \mathbf{e}_1]$ is a basis matrix which spans the signal subspace of \mathbf{R}_{xx} ; $\mathbf{E}_n = [\mathbf{e}_2 \ \dots \ \mathbf{e}_{M-1}]$ is a basis matrix which spans the noise subspace. Based on (25), we let the inverse of the matrix $\bar{\mathbf{R}}_{\text{gl}}^{-1}$ of (18) be expressed as

$$\bar{\mathbf{R}}_{\text{gl}}^{-1} = \underbrace{\mathbf{E}_s \bar{\mathbf{\Lambda}}_s^{-1} \mathbf{E}_s^H}_{\mathbf{P}_s} + \underbrace{\mathbf{E}_n \bar{\mathbf{\Lambda}}_n^{-1} \mathbf{E}_n^H}_{\mathbf{P}_n} \quad (26)$$

where $\bar{\mathbf{\Lambda}}_s = \text{diag}\{\bar{\lambda}_0, \bar{\lambda}_1\}$ with $\bar{\lambda}_i = \lambda_i + \alpha_i$ and $\alpha_i = \frac{1-\lambda_i^{-p}}{\lambda_i^{-1}}$ for $i = 0, 1$, and $\bar{\mathbf{\Lambda}}_n = (\sigma_n^2 + \beta)\mathbf{I}$ with $\beta = \frac{1-\sigma_n^{-2p}}{\sigma_n^{-1}}$. Then, we can rewrite the weight vector \mathbf{w}_{pr} of (19) as

$$\mathbf{w}_{\text{pr}} = \mu_{\text{pr}}(\mathbf{P}_s + \mathbf{P}_n)(\mathbf{a}_0 + \sigma_e \mathbf{\Delta}) \quad (27)$$

where $\mu_{\text{pr}} = 1/\mathbf{a}^H \bar{\mathbf{R}}_{\text{gl}}^{-1} \mathbf{a}$. Due to the fact that $\mathbf{E}_s \mathbf{E}_s^H + \mathbf{E}_n \mathbf{E}_n^H = \mathbf{I}$, we have $\mathbf{E}_s \mathbf{E}_s^H \mathbf{\Delta} + \mathbf{E}_n \mathbf{E}_n^H \mathbf{\Delta} = \mathbf{\Delta}$, where $\mathbf{E}_s \mathbf{E}_s^H \mathbf{\Delta} = \eta_0 \mathbf{e}_0 + \eta_1 \mathbf{e}_1$ with $\eta_i = \mathbf{e}_i^H \mathbf{\Delta}$, $i = 0, 1$. Furthermore, since the range space of $\mathbf{E}_s = [\mathbf{e}_0 \ \mathbf{e}_1]$ is the same as that of $\mathbf{A} = [\mathbf{a}_0 \ \mathbf{a}_1]$, i.e., $\text{span}\{\mathbf{E}_s\} = \text{span}\{\mathbf{A}\}$, we have $\eta_0 \mathbf{e}_0 + \eta_1 \mathbf{e}_1 = h_0 \mathbf{a}_0 + h_1 \mathbf{a}_1$. Hence, the error vector $\mathbf{\Delta}$ can be rewritten as

$$\mathbf{\Delta} = h_0 \mathbf{a}_0 + h_1 \mathbf{a}_1 + \mathbf{E}_n \mathbf{E}_n^H \mathbf{\Delta}. \quad (28)$$

Then, we have that $\mathbf{a}_0^H \mathbf{P}_s \mathbf{\Delta} \approx \mathbf{a}_0^H \mathbf{P}_s \mathbf{a}_0$ and (A1) of Appendix A.1 can be expressed as

$$P_d \approx \sigma_{s_0}^2 |\mu_{\text{pr}}|^2 |1 + h_0 \sigma_e|^2 (\mathbf{a}_0^H \mathbf{P}_s \mathbf{a}_0)^2 \quad (29)$$

On the other hand, using (28), we have an approximation of (A2) as follows:

$$P_o \approx |\mu_{\text{pr}}|^2 (|1 + h_0 \sigma_e|^2 \mathbf{a}_0^H \bar{\mathbf{P}}_s \mathbf{a}_0 + \sigma_e^2 |h_1|^2 \mathbf{a}_1^H \bar{\mathbf{P}}_s \mathbf{a}_1 + \sigma_e^2 \mathbf{\Delta}^H \bar{\mathbf{P}}_n \mathbf{\Delta}) \quad (30)$$

with the fact that $\mathbf{a}_i^H \bar{\mathbf{P}}_s \mathbf{a}_j \approx 0$ when $i \neq j$ and $\mathbf{a}_i^H \bar{\mathbf{P}}_n \mathbf{a}_j = 0$ for all i, j . Consider the case that there exists only a desired signal and background noise; it is easy to show from (25) that $\bar{\mathbf{E}}_s = \mathbf{e}_0 \propto \mathbf{a}_0 / \sqrt{M}$ and $\bar{\mathbf{\Lambda}}_s = \xi_0 = M\sigma_{s_0}^2 + \sigma_n^2$. Hence, we have that $\bar{\mathbf{\Lambda}}_s = \xi_0 = \xi_0 + \tilde{\alpha}_0$ and then $\mathbf{a}_0^H \mathbf{P}_s \mathbf{a}_0 \approx M\bar{\xi}_0^{-1}$, where $\tilde{\alpha}_0 = \frac{1-\xi_0^{-p}}{\xi_0^{-1}}$ plays the same role as α_0 . When there exists an additional interferer, we note that $\mathbf{a}_0^H \mathbf{P}_s \mathbf{a}_0 \approx M\bar{\xi}_0^{-1}$ is also valid if $|\mathbf{a}_0^H \mathbf{a}_i| \ll M$ for $i \neq 0$. As to $\mathbf{a}_i^H \bar{\mathbf{P}}_s \mathbf{a}_i$ of (30), similarly, we can have $\mathbf{a}_i^H \bar{\mathbf{P}}_s \mathbf{a}_i \approx M\bar{\xi}_i \bar{\xi}_i^{-2}$, where $\bar{\xi}_i = M\sigma_{s_i}^2 + \sigma_n^2$ and $\bar{\xi}_i = \xi_i + \tilde{\alpha}_i$ with $\tilde{\alpha}_i = \frac{1-\xi_i^{-p}}{\xi_i^{-1}}$, $i = 0, 1$. Accordingly, (29) and (30) can be further rewritten as

$$P_d \approx \sigma_{s_0}^2 |\mu_{\text{pr}}|^2 |1 + h_0 \sigma_e|^2 M^2 \bar{\xi}_0^{-2} \quad (31)$$

and

$$P_o \approx |\mu_{pr}|^2 (|1 + h_0\sigma_e|^2 M\xi_0\bar{\xi}_0^{-2} + \sigma_e^2|h_1|^2 M\xi_1\bar{\xi}_1^{-2} + \sigma_e^2\Delta^H\bar{\mathbf{P}}_n\Delta) \quad (32)$$

respectively. Then, the difference of (32) and (31) is given by

$$P_o - P_d \approx |\mu_{pr}|^2 \left[|1 + h_0\sigma_e|^2 \underbrace{(\xi_0 - M\sigma_{s_0}^2)}_{\sigma_n^2} M\bar{\xi}_0^{-2} + \sigma_e^2|h_1|^2 M\xi_1\bar{\xi}_1^{-2} + \sigma_e^2\Delta^H\bar{\mathbf{P}}_n\Delta \right]. \quad (33)$$

Hence, the output SINR of the proposed method is given by

$$\text{SINR}_{pr} \approx \frac{|1 + h_0\sigma_e|^2 M \cdot \text{SNR}\bar{\xi}_0^{-2}}{|1 + h_0\sigma_e|^2\bar{\xi}_0^{-2} + \sigma_e^2|h_1|^2(1 + M \cdot \text{INR})\bar{\xi}_1^{-2} + \frac{\sigma_e^2}{M\sigma_n^2}\Delta^H\bar{\mathbf{P}}_n\Delta} \quad (34)$$

where $\text{SNR} = \sigma_{s_0}^2/\sigma_n^2$ and $\text{INR} = \sigma_{s_1}^2/\sigma_n^2$ denote the signal-to-noise power ratio (SNR) and interference-to-noise power ratio (INR), respectively.

Since Δ is a random vector, we next evaluate the statistical expectation of SINR_{pr} . Using (A4) and (A5) of Appendix A.2, we can have from (34) that

$$E\{|1 + h_0\sigma_e|^2\} = 1 + \frac{\sigma_e^2}{M}. \quad (35)$$

Furthermore, the statistical expectation of $\Delta^H\bar{\mathbf{P}}_n\Delta$ can be expressed as

$$E\{\Delta^H\bar{\mathbf{P}}_n\Delta\} = \text{tr}\{\bar{\mathbf{P}}_n\} = \sigma_n^2(\sigma_n^2 + \beta)^{-2}(M - 2). \quad (36)$$

Accordingly, the statistical expectation of SINR_{pr} can be approximately expressed as [24]

$$E\{\text{SINR}_{pr}\} \approx \frac{\left(1 + \frac{\sigma_e^2}{M}\right) M \cdot \text{SNR}\bar{\xi}_0^{-2}}{\left(1 + \frac{\sigma_e^2}{M}\right)\bar{\xi}_0^{-2} + \frac{\sigma_e^2}{M}(1 + M \cdot \text{INR})\bar{\xi}_0^{-2} + \frac{\sigma_e^2}{M\sigma_n^4}\left(1 + \frac{\beta}{\sigma_n^2}\right)^{-2}(M - 2)}. \quad (37)$$

We note from (37) that the last term in denominator is more significant than the other two terms. This implies that the steering vector error is more significant in the noise components. Consider two extreme situations, namely $p = 0$ and $p = \infty$. For $p = 0$, we have that $\tilde{\alpha}_i = 0$ and $\beta = 0$, then $\bar{\xi}_i = \xi_i$ (recall that $\bar{\xi}_i = \xi_i + \tilde{\alpha}_i$). Thus, (37) can be written as

$$E\{\text{SINR}_{pr}\}|_{p=0} \approx \frac{M \cdot \text{SNR}}{1 + \frac{\sigma_e^2}{M + \sigma_e^2} \frac{(1 + M \cdot \text{INR})^{-1} + (M - 2)}{(1 + M \cdot \text{SNR})^{-2}}}. \quad (38)$$

We observe that the term $M - 2$ in the denominator is dominant and it significantly degrades the output SINR due to the steering vector

error. For $p = \infty$, we have that $\tilde{\alpha}_i = 1/(\xi_i - 1)$ and $\beta = \infty$. Then, (37) becomes

$$E\{\text{SINR}_{\text{pr}}\}|_{p=\infty} \approx \frac{M \cdot \text{SNR}}{1 + \frac{\sigma_e^2}{M + \sigma_e^2} \frac{(1 + M \cdot \text{INR})(\xi_0 + \tilde{\alpha}_0)^2}{(\xi_1 + \tilde{\alpha}_1)^2}}. \quad (39)$$

Further, we can approximate $\tilde{\alpha}_i$ by taking the first two terms of its Taylor series expression to obtain

$$\tilde{\alpha}_i \approx \frac{1}{\xi_i} \left(1 + \frac{1}{\xi_i} \right), \quad i = 0, 1, \quad (40)$$

when $1/\xi_i$ is small enough. Substituting $\tilde{\alpha}_i$ of (40) into (39), we obtain

$$E\{\text{SINR}_{\text{pr}}\}|_{p=\infty} \approx \frac{M \cdot \text{SNR}}{1 + \frac{\sigma_e^2}{M + \sigma_e^2} \frac{(1 + M \cdot \text{SNR})^2}{1 + M \cdot \text{INR}} \rho} \quad (41)$$

where $\rho = \frac{(1 + \xi_0^{-2} + \xi_0^{-3})^2}{(1 + \xi_1^{-2} + \xi_1^{-3})^2}$. Since ξ_0 and ξ_1 are related to the desired signal and the interferer, respectively, we can have

$$\rho > 1 \text{ when } \text{SNR} < \text{INR}, \quad (42)$$

$$\rho = 1 \text{ when } \text{SNR} = \text{INR}, \quad (43)$$

$$\rho < 1 \text{ when } \text{SNR} > \text{INR}, \quad (44)$$

It follows from (44) that ρ decreases when SNR increases and $\text{SNR} > \text{INR}$. However, (41) reveals that increasing SNR enhances the effect of the steering vector error because SNR is more dominant in the denominator of (41) than ρ . Although we do not prefer the case of $\rho > 1$, a small SNR will diminish the effect of the steering vector error. Nevertheless, when ξ_0 and ξ_1 are large, ρ is close to one no matter what the SNR and INR are. Thus, (41) has the following approximation

$$E\{\text{SINR}_{\text{pr}}\}|_{p=\infty} \approx \frac{M \cdot \text{SNR}}{1 + \frac{\sigma_e^2}{M + \sigma_e^2} \frac{(1 + M \cdot \text{SNR})^2}{(1 + M \cdot \text{INR})}}. \quad (45)$$

We observe from (41) or (45) that the denominator does not have the $M - 2$ term. Hence, the degradation due to the steering vector error can be efficiently alleviated by increasing p . Comparing (45) with (38), we can conclude that

$$E\{\text{SINR}_{\text{pr}}\}|_{p=\infty} > E\{\text{SINR}_{\text{pr}}\}|_{p=0}. \quad (46)$$

As a result, the output SINR of the proposed method is larger than that of the SMI which is the special case of the proposed method with $p = 0$.

5. SIMULATION EXAMPLES

In this section, we present several simulation examples by using the SMI [2], the DL [8], the HKB [18], the GLC [19], the AGL [20], and the proposed method for comparison. For all simulations, we use a ULA of $M = 10$ array elements with inter-element spacing equal to half-wavelength. We note that the ULA with inter-element spacing equal to half-wavelength is referred to as a standard linear array (SLA) [2, p.51]. We use the SLA to steer the angles between -90° and 90° . In general, the resolution of an antenna array increases as the array dimension or the inter-element spacing increases. High array resolution enhances the maximum output SINR [3, p.34]. The desired signal and the interference are binary phase-shift-keying (BPSK) signals. The desired signal impinges on the array from 0° off array broadside with SNR equal to 10 dB. Two interferes with INR equal to 20 dB impinge on the array from 30° and 60° off array broadside, respectively. The loading factor γ for the DL is set to $10\sigma_n^2$, where the noise variance $\sigma_n^2 = 0.4$. Moreover, all the simulation results are obtained by averaging 100 independent runs.

Example 1: Here, we present the output SINR versus the value of p for different N data snapshots when $\sigma_e^2 = 0.2$. We observe from Fig. 1 that the performance of the proposed method improves as p increases and achieves a steady state when p is larger than 5 for all N . Hence, we use $p = 5$ for the following simulations. Fig. 2 depicts the output SINR versus the number of data snapshots without steering vector error. We observe that the DL, the AGL, and the proposed method converge to the optimal SINR faster than the other methods. The SMI converges to the optimal SINR very slowly. The GLC provides the second best performance. However, the loading factor of the GLC

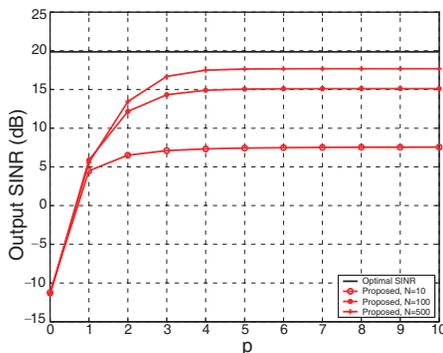


Figure 1. The output SINR versus p for *Example 1*, $\sigma_e^2 = 0.2$.

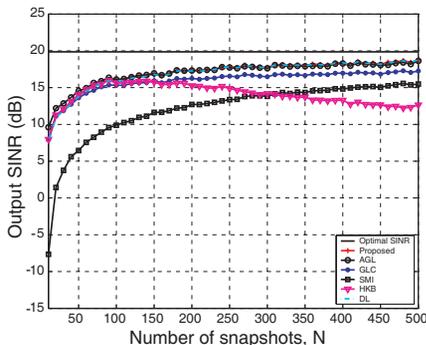


Figure 2. The output SINR versus N for *Example 1*, $\sigma_e^2 = 0$.

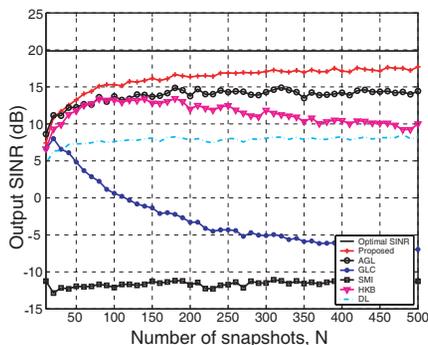


Figure 3. The output SINR versus N for *Example 1*, $\sigma_e^2 = 0.2$.

decreases when N increases. This indicates that the GLC converges to the SMI when N is large enough. On the other hand, although the HKB gives good performance for small N , it suffers from performance degradation when N increases. This is because the loading factors of the HKB may be very large when N is large. Fig. 3 plots the output SINR versus the number of data snapshots with $\sigma_e^2 = 0.2$. We observe that all of the methods suffer from performance degradation in this case. Moreover, the GLC degrades significantly due to the loading factor decreases when N increases. As in the case of $\sigma_e^2 = 0$, the HKB performs well when N is small and degrades when N increases. The AGL has the capability against the difficulty due to the steering vector error. However, the proposed method outperforms all of the other methods.

Example 2: The beampatterns of using the aforementioned methods for $N = 500$ are shown in Figs. 4 and 5 without and with steering vector error, respectively. From Fig. 4, we note that all of the methods are capable of preserving the desired signal and suppressing the interference. However, the GLC has higher sidelobe levels than the other methods due to large N , whereas the HKB has lower sidelobe levels but it can not properly deal with interference because the required loading factor is quite large due to large N . In contrast, the proposed method works satisfactorily and has the performance similar to that of the AGL. From Fig. 5, it is observed that the GLC suppresses the desired signal due to its inherent drawback when N is large. As to the HKB, using a large loading factor reduces its capability in interference cancelation. Although the AGL works satisfactorily, the proposed method again outperforms all of the other methods.

Example 3: First, the output SINR versus σ_e^2 is plotted in Fig. 6, where the parameters are the same as those used by *Example 2*. We observe that the performance degradation of each method is more pronounced as σ_e^2 increases. However, the proposed method provides more robustness against steering vector error than each of the other methods. Next, we present the output SINR versus the input SNR for comparison, where $\sigma_e^2 = 0.2$. The input SNR varies from -10 dB to 10 dB with an increment equal to 2 dB. From the results shown in Fig. 7, we see that the performance of the proposed method is more acceptable under various values of SNR. Finally, to evaluate the influence of interference on the proposed method, a figure for the output SINR versus signal-to-interference ratio (SIR) is depicted in Fig. 8, where SNR is fixed and set to 0 dB and INR varies from 0 dB to 20 dB. From Fig. 8, the output SINR of the proposed method remains almost the same when $\sigma_e^2 = 0$. For the case of $\sigma_e^2 = 0.2$, the output SINR of the proposed method degrades as SIR increases or INR decreases. This can be expected from (45) in which the value of output SINR decreases as INR decreases when $\sigma_e^2 \neq 0$ and remains the same when $\sigma_e^2 = 0$.

Example 4: We present this example to confirm the validity of the theoretical results shown by (37), (38), and (45). Fig. 9 shows the simulation results with only one interferer from 30° and the ensemble correlation matrix \mathbf{R}_{xx} . We can see that the theoretical results are very close to the simulation results. We also note from Fig. 9 that the results of using the SMI are almost the same as those of the theoretical results shown by (37) or (38) when $p = 0$. Hence, the validity of the theoretical results is confirmed. Moreover, we observe from Fig. 9 that the proposed method with $p = 5$ is capable of providing performance almost close to (45).

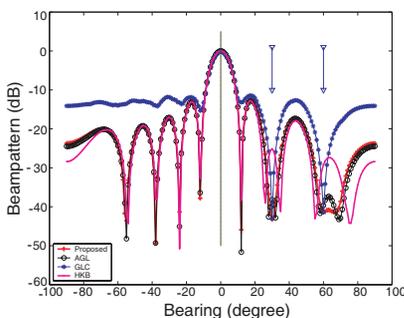


Figure 4. The beam pattern for *Example 2*, $\sigma_e^2 = 0$ and $N = 500$.

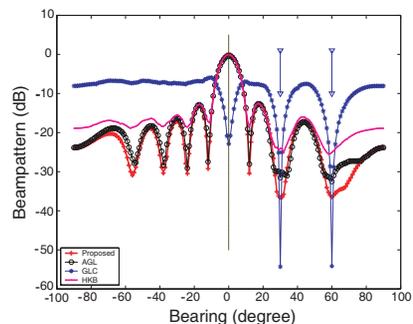


Figure 5. The beam pattern for *Example 2*, $\sigma_e^2 = 0.2$ and $N = 500$.

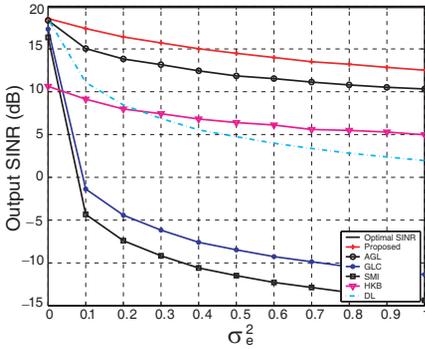


Figure 6. The output SINR versus σ_e^2 for *Example 3*, $N = 500$.

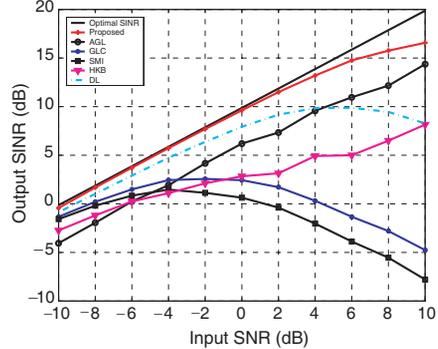


Figure 7. The output SINR versus the input SNR for *Example 3*, $\sigma_e^2 = 0.2$ and $N = 500$.

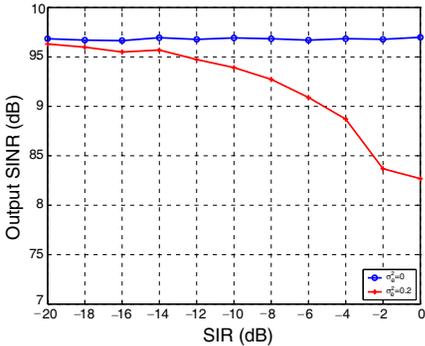


Figure 8. The output SINR versus SIR for *Example 3*, $N = 500$.

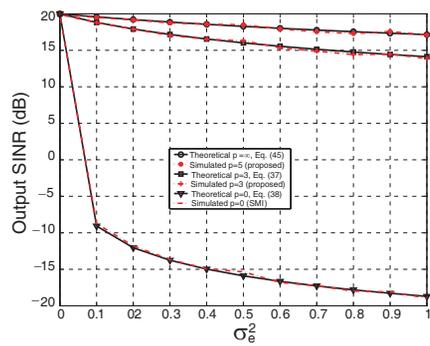


Figure 9. Confirmation of the theoretical results shown by (37), (38), and (45). *Example 4*.

6. CONCLUSION

An efficient method has been presented to deal with the performance deterioration under finite sample effect and steering vector error. The proposed method utilizes a fully data-dependent loading to overcome the difficulties due to these errors. Unlike the existing diagonal loading (DL) techniques, the proposed method does not require any additional sophisticated scheme to choose the required loading. The loading factor can be completely obtained from the received array data. Analytical formulas for evaluating the performance of the proposed method under random steering vector error have been derived. Simulation results have confirmed the validity of the proposed method and shown the

effectiveness of the proposed method as compared with the existing DL techniques.

ACKNOWLEDGMENT

This work was supported by the National Science Council of TAIWAN under Grant NSC97-2221-E002-174-MY3.

APPENDIX A.

A.1.

From (23)–(27), the output power of the desired signal is given by

$$\begin{aligned}
 P_d &= \sigma_{s_0}^2 |\mu_{\text{pr}} (\mathbf{a}_0 + \sigma_e \mathbf{\Delta})^H (\mathbf{P}_s + \mathbf{P}_n) \mathbf{a}_0|^2 \\
 &= \sigma_{s_0}^2 |\mu_{\text{pr}}|^2 \left[(\mathbf{a}_0^H \mathbf{P}_s \mathbf{a}_0)^2 + 2\Re \{ \sigma_e \mathbf{a}_0^H \mathbf{P}_s \mathbf{a}_0 \mathbf{a}_0^H \mathbf{P}_s \mathbf{\Delta} \} \right] + \sigma_e^2 |\mathbf{a}_0^H \mathbf{P}_s \mathbf{\Delta}|^2 \quad (\text{A1})
 \end{aligned}$$

where $\Re\{x\}$ denotes the real part of a complex value x . On the other hand, the corresponding output power can be written as

$$\begin{aligned}
 P_o &= \mathbf{w}_{\text{pr}}^H \mathbf{E}_s \mathbf{\Lambda}_s \mathbf{E}_s^H \mathbf{w}_{\text{pr}} + \sigma_n^2 \mathbf{w}_{\text{pr}}^H \mathbf{E}_n \mathbf{E}_n^H \mathbf{w}_{\text{pr}} \\
 &= |\mu_{\text{pr}}|^2 (\mathbf{a}_0^H \bar{\mathbf{P}}_s \mathbf{a}_0 + 2\Re \{ \sigma_e \mathbf{a}_0^H \bar{\mathbf{P}}_s \mathbf{\Delta} \} + \sigma_e^2 \mathbf{\Delta}^H \bar{\mathbf{P}}_s \mathbf{\Delta} + \sigma_e^2 \mathbf{\Delta}^H \bar{\mathbf{P}}_n \mathbf{\Delta}) \quad (\text{A2})
 \end{aligned}$$

where $\bar{\mathbf{P}}_s = \mathbf{E}_s \bar{\mathbf{\Lambda}}_s^{-1} \mathbf{\Lambda}_s \bar{\mathbf{\Lambda}}_s^{-1} \mathbf{E}_s^H$ and $\bar{\mathbf{P}}_n = \sigma_n^2 \mathbf{E}_n \bar{\mathbf{\Lambda}}_s^{-2} \mathbf{E}_n$.

A.2.

Assume that $|\mathbf{a}_i^H \mathbf{a}_j| \ll M$ for $i \neq j$ in which $\mathbf{a}_i^H \mathbf{a}_i = M$, h_i of (28) can be obtained as

$$h_i \approx \frac{\mathbf{a}_i^H \mathbf{\Delta}}{M}, \quad i = 0, 1. \quad (\text{A3})$$

Then, we have that

$$E\{h_i\} = 0 \quad (\text{A4})$$

due to $E\{\mathbf{\Delta}\} = 0$ and

$$E\{|h_i|^2\} \approx E\{|\mathbf{a}_i^H \mathbf{e}|^2\} / M^2 = \text{tr} \{ \mathbf{a}_i \mathbf{a}_i^H E\{\mathbf{\Delta} \mathbf{\Delta}^H\} \} / M^2 = 1/M \quad (\text{A5})$$

where $\text{tr}\{\mathbf{A}\}$ denotes the trace of the square matrix \mathbf{A} .

REFERENCES

1. Capon, J., "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, Vol. 57, 1408–1418, Aug. 1969.
2. Van Trees, H. L., *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, John & Wiley Sons, Inc., New York, USA, 2002.
3. Monzingo, R. A. and T.W. Miller, *Introduction to Adaptive Array*, Wiley, New York, 1980.
4. Carlson, B. D., "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 24, 397–401, Jul. 1988.
5. Wax, M. and Y. Anu, "Performance analysis of the minimum variance beamformer," *IEEE Trans. Signal Processing*, Vol. 44, 928–937, Apr. 1996.
6. Wax, M. and Y. Anu, "Performance analysis of the minimum variance beamformer in the presence of steering vector errors," *IEEE Trans. Signal Processing*, Vol. 44, 938–947, Apr. 1996.
7. Feldman, D. D. and L. J. Griffiths, "A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Processing*, Vol. 42, 867–876, Apr. 1994.
8. Cox, H., R. M. Zeskind, and M. H. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 35, 1365–1376, Oct. 1987.
9. Lee, C.-C. and J.-H. Lee, "Robust adaptive array beamforming under steering vector errors," *IEEE Trans. Antennas Propagat.*, Vol. 45, 168–175, 1997.
10. Li, Y., Y.-J. Gu, Z.-G. Shi, and K. S. Chen, "Robust adaptive beamforming based on particle filter with noise unknown," *Progress In Electromagnetics Research*, Vol. 90, 151–169, 2009.
11. Wang, W., R. Wu, and J. Liang, "A novel diagonal loading method for robust adaptive beamforming," *Progress In Electromagnetics Research C*, Vol. 18, 245–255, 2011.
12. Vincent, F. and O. Besson, "Steering vector errors and diagonal loading," *IEE Proc. Radar Sonar Navig.*, Vol. 151, No. 6, 337–343, Dec. 2004.
13. Li, J., P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. on Signal Processing*, Vol. 51, No. 7, 1702–1715, Jul. 2003.
14. Lorenz, R. G. and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. on Signal Processing*, Vol. 53, No. 5,

- 1684–1696, May 2005.
15. Vorobyov, S., A. B. Gershman, and Z.-Q. Luo, “Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem,” *IEEE Trans. on Signal Processing*, Vol. 51, 313–324, Feb. 2003.
 16. Gu, J., “Robust beamforming based on variable loading,” *Electronics Letters*, Vol. 41, No. 2, Jan. 2005.
 17. Gu, J. and P. J. Wolfe, “Robust adaptive beamforming using variable loading,” *Fourth IEEE Workshop on Sensor Array and Multichannel Processing*, 1–5, Jul. 2006.
 18. Selén, Y., R. Abrahamsson, and P. Stoica, “Automatic robust adaptive beamforming via ridge regression,” *Signal Processing*, Vol. 88, No. 1, 33–49, 2008.
 19. Li, J., L. Du, and P. Stoica, “Fully automatic computation of diagonal loading levels for robust adaptive beamforming,” *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 46, No. 1, 449–458, Jan. 2010.
 20. Yang, J., X. H. Ma, C. H. Hou, and Y. Liu, “Automatic generalized loading for robust adaptive beamforming,” *IEEE Signal Process. Lett.*, Vol. 16, No. 3, 219–222, Mar. 2009.
 21. Du, L., T. Yardibi, J. Li, and P. Stoica, “Review of user parameter-free robust adaptive beamforming algorithms,” *Digital Signal Processing*, Vol. 19, No. 4, 567–582, Jul. 2009.
 22. Besson, O. and F. Vincent, “Performance analysis of beamformers using generalized loading of the covariance matrix in the presence of random steering vector errors,” *IEEE Trans. on Signal Processing*, Vol. 53, 452–459, Feb. 2005.
 23. Chang, L. and C. C. Yeh, “Performance of DMI and eigenspace-based beamformers,” *IEEE Trans. Antennas Propagat.*, Vol. 40, No. 11, 1336–1347, Nov. 1992.
 24. Yu, L., W. Liu, and R. Langley, “SINR analysis of the subtraction-based SMI beamformer,” *IEEE Trans. on Signal Processing*, Vol. 58, No. 11, 5926–5932, Nov. 2010.
 25. Olen, C. A. and R. T. Compton Jr., “A numerical pattern synthesis algorithm for arrays,” *IEEE Trans. Antennas Propagat.*, Vol. 38, 1666–1676, Oct. 1990.
 26. Zhou, P. Y. and M. A. Ingram, “Pattern synthesis for arbitrary arrays using an adaptive array method,” *IEEE Trans. Antennas Propagat.*, Vol. 47, No. 5, 862–869, May 1999.