RANDOM MATRIX-OPTIMIZED HIGH-DIMENSIONAL MVDR BEAMFORMING

Liusha Yang⋆*, Matthew McKay*⋆*, Romain Couillet*†

⋆Dept. of ECE, Hong Kong University of Science and Technology, Hong Kong
†CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France

ABSTRACT

A new approach to minimum variance distortionless response (MVDR) beamforming is proposed under the assumption of simultaneously large numbers of array sensors and observations. The key to our method is the design of an inverse covariance estimator which is appropriately optimized for the MVDR application. This is obtained by exploiting spectral properties of spiked covariance models in random matrix theory. Our proposed solution is simple to implement and is shown to yield performance improvements over competing approaches.

***Index Terms*—** MVDR beamforming, spiked covariance model, random matrix theory.

1. INTRODUCTION

Adaptive beamforming is widely used for adaptively steering a beam towards a desired signal while suppressing noise and interference at the output of a sensor array. One popular approach is the minimum variance distortionless response (MVDR) adaptive beamformer [1] designed to minimize the array output power while maintaining a distortionless response towards the signal of interest (SoI). The performance of the MVDR beamformer depends on the estimation accuracy of the inverse covariance matrix of the received signals, which is involved in the beamformer's construction. The "sample matrix inversion" (SMI) beamformer is a standard approach, which uses the inverse of the sample covariance matrix (SCM). However, while it performs well when the number of observations greatly exceeds the number of sensors of the receive array, in modern data-limited scenarios or high-dimensional applications (with large numbers of sensors), performance degradation occurs because of the increased estimation errors in the SCM, especially when the samples used for estimation include the SoI [2, 3].

Various approaches have been proposed to design more robust beamforming solutions which aim to overcome this problem. The most common are diagonal loading and eigen-subspace techniques [4–8]. As we will demonstrate, for the former approach, even with the theoretically-optimal diagonal loading factor for the MVDR beamformer, chosen to minimize the array output power, it exhibits suboptimal performance under data-limited or high-dimensional scenarios; for the latter, it becomes ineffective at low signal-to-noise ratios (SNRs) or when the dimension of the signal-plus-interference subspace is large.

In this paper, we propose a novel MVDR beamforming solution under high-dimensional settings and in the face of data limitations. Our proposed method is based on the assumption that the receive covariance matrix has a specific structure, referred to as a "spiked model" in random matrix theory (RMT). Specifically, this construction comprises a low-rank perturbation of a scaled identity matrix. Inspired by the recent work [9] which proposed optimized spike-model-based covariance estimators for a range of different objective functions, we exploit properties from RMT to identify the MVDR-optimal inverse covariance estimator, and consequently obtain our proposed MVDR beamforming solution. It is shown through numerical examples to yield exceptional beamformer performance for high-dimensional and data-limited scenarios.

Yang and McKay's work was supported by the Hong Kong RGC General Research Fund under grant numbers 16206914 and 16203315. Couillet's work was supported by the ANR Project RMT4GRAPH (ANR-14-CE28-0006).

2. BACKGROUND

2.1. Signal model and optimal MVDR beamforming

We consider a uniform linear array with N sensors, receiving $m < N$ narrow band signals. At snapshot $j \in \{1, \cdots, n\}$, the received observation vector $\mathbf{x}(j) \in \mathbb{C}^N$ can be represented by

$$\mathbf{x}(j) = \sqrt{p_1}\mathbf{a}(\theta_1)\phi_1(j) + \sum_{i=2}^{m} \sqrt{p_i}\mathbf{a}(\theta_i)\phi_i(j) + \mathbf{n}(j)\,, \quad (1)$$

where for $i = 1, \ldots, m$, signal $\phi_i(j) \in \mathbb{C}$ is independent complex Gaussian with zero mean and variance one, $p_i \in \mathbb{R}^+$ is the corresponding signal power, and $\mathbf{a}(\theta_i) \in \mathbb{C}^N$ is the unit norm steering vector that is a function of the direction of arrival (DoA) $\theta_i \in (-\pi, \pi]$ of the i-th source. The first term in (1) corresponds to the SoI and the second term to $m-1$ interferers. We assume $\mathbf{n}(j) \in \mathbb{C}^N$ is additive Gaussian white noise with zero mean and variance σ^2. The covariance matrix of the observations takes the following form:

$$\mathbf{C}_N = p_1\mathbf{a}(\theta_1)\mathbf{a}^H(\theta_1) + \sum_{i=2}^{m} p_i\mathbf{a}(\theta_i)\mathbf{a}^H(\theta_i) + \sigma^2\mathbf{I}_N. \quad (2)$$

The classical MVDR beamformer [1] seeks the beamformer weight vector $\mathbf{h} = \mathbf{h}(\theta_1) \in \mathbb{C}^N$ that minimizes the output power $P(\mathbf{h}) = \mathbf{h}^H\mathbf{C}_N\mathbf{h}$ while ensuring a distortionless response towards the DoA of the SoI θ_1. It is the solution to the following linearly constrained quadratic problem:

$$\min_{\mathbf{h}\in\mathbb{C}^N} P(\mathbf{h}) = \mathbf{h}^H\mathbf{C}_N\mathbf{h}, \quad \text{s.t. } \mathbf{h}^H\mathbf{a}(\theta_1) = 1.$$

The well-known solution to this constrained optimization is

$$\mathbf{h}_{\text{MVDR}} = \frac{\mathbf{C}_N^{-1}\mathbf{a}(\theta_1)}{\mathbf{a}^H(\theta_1)\mathbf{C}_N^{-1}\mathbf{a}(\theta_1)} \quad (3)$$

and the corresponding output power is

$$P(\mathbf{h}_{\text{MVDR}}) = \frac{1}{\mathbf{a}^H(\theta_1)\mathbf{C}_N^{-1}\mathbf{a}(\theta_1)}$$

where $\mathbf{a}(\theta_1)$ is assumed known. In the following, for convenience, we will consider the normalized total output power, $\rho(\mathbf{h}) = P(\mathbf{h})/\sigma^2$.

From the matrix inversion lemma [10], an equivalent representation of the optimal beamformer is $\mathbf{h}_{\mathrm{MVDR}} = \frac{\mathbf{C}_{\mathrm{i+n}}^{-1}\mathbf{a}(\theta_1)}{\mathbf{a}^H(\theta_1)\mathbf{C}_{\mathrm{i+n}}^{-1}\mathbf{a}(\theta_1)}$ where $\mathbf{C}_{\mathrm{i+n}} = \mathbf{C}_N - p_1\mathbf{a}(\theta_1)\mathbf{a}(\theta_1)^H$ is the interference-plus-noise covariance matrix. Here we assume that the SoI is present in the observations, which makes it difficult to estimate $\mathbf{C}_{\mathrm{i+n}}$. Indeed, the presence of the SoI in the training data has been shown to dramatically reduce the convergence rate of beamforming algorithms [2, 5].

We note that, in addition to minimizing the output power, the MVDR beamformer also maximizes the output SINR [11], defined as $\mathrm{SINR} = \frac{p_1\left|\hat{\mathbf{h}}^H\mathbf{a}(\theta_1)\right|^2}{\hat{\mathbf{h}}^H\mathbf{C}_{\mathrm{i+n}}\hat{\mathbf{h}}}$. Hence, we can also measure the beamformer performance in terms of its achieved SINR.

2.2. Sample-based implementation of MVDR beamforming

In reality, $\mathbf{C}_N^{-1}$ is unknown and instead we form an estimate, denoted by $\hat{\mathbf{C}}_N^{-1}$. Thus the MVDR beamformer based on any given plug-in estimator $\hat{\mathbf{C}}_N^{-1}$ is constructed as

$$\hat{\mathbf{h}}_{\mathrm{MVDR}} = \frac{\hat{\mathbf{C}}_N^{-1}\mathbf{a}(\theta_1)}{\mathbf{a}^H(\theta_1)\hat{\mathbf{C}}_N^{-1}\mathbf{a}(\theta_1)}. \tag{4}$$

The performance of this beamformer in terms of its normalized total output power is given by

$$\rho(\hat{\mathbf{h}}_{\mathrm{MVDR}}) = \frac{P(\hat{\mathbf{h}}_{\mathrm{MVDR}})}{\sigma^2} = \frac{1}{\sigma^2}\frac{\mathbf{a}^H(\theta_1)\hat{\mathbf{C}}_N^{-1}\mathbf{C}_N\hat{\mathbf{C}}_N^{-1}\mathbf{a}(\theta_1)}{\left(\mathbf{a}^H(\theta_1)\hat{\mathbf{C}}_N^{-1}\mathbf{a}(\theta_1)\right)^2}. \tag{5}$$

It reaches $\rho_{\min} = \rho(\mathbf{h}_{\mathrm{MVDR}})$ only when $\hat{\mathbf{C}}_N^{-1} = \mathbf{C}_N^{-1}$, otherwise it is larger because of the imperfect inverse covariance estimation.

We define the SCM as $\mathbf{S}_N = \frac{1}{n}\sum_{j=1}^n \mathbf{x}(j)\mathbf{x}^H(j)$. It is widely known that the SMI beamformer $\hat{\mathbf{h}}_{\mathrm{SMI}} = \frac{\mathbf{S}_N^{-1}\mathbf{a}(\theta_1)}{\mathbf{a}^H(\theta_1)\mathbf{S}_N^{-1}\mathbf{a}(\theta_1)}$ can have poor performance due to the finite-sampling effect, yielding much higher normalized total output power than the theoretical minimum $\rho_{\min}$, especially when N and n have a similar order of magnitude, and for the case of interest with the SoI present in the observations [2, 8, 12]. In this work, by utilizing prior knowledge of the structure of $\mathbf{C}_N$, we will propose an optimized $\hat{\mathbf{C}}_{\mathrm{MVDR}}^{-1}$, and consequently an optimized $\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$, which is designed to minimize (5).

3. OPTIMIZED HIGH-DIMENSIONAL MVDR DESIGN

3.1. MVDR beamforming based on spiked covariance models

We aim to construct a MVDR-optimized sample-based estimate of the precision matrix $\mathbf{C}_N^{-1}$, with $\mathbf{C}_N$ defined in (2). We note that, through eigen-decomposition, the matrix $\mathbf{C}_N$ can be expressed as

$$\mathbf{C}_N = \sigma^2\left(\mathbf{I}_N + \sum_{i=1}^m t_i\mathbf{v}_i\mathbf{v}_i^H\right), \tag{6}$$

which has eigenvalues $(\sigma^2(t_1+1), \dots, \sigma^2(t_m+1), \sigma^2, \dots, \sigma^2)$, where $t_i > 0$ for $i = 1, \dots, m$, and with $\mathbf{v}_1, \dots, \mathbf{v}_m$ denoting the eigenvectors corresponding to the largest m eigenvalues. Because $\mathbf{C}_N$ is a low rank perturbation of $\sigma^2\mathbf{I}_N$, the structure of $\mathbf{C}_N$ is called a *spiked* covariance model [13]. For simplicity, we assume that σ^2 and m are known, though these may be estimated using standard methods [14].

Denoting the eigenvalues and associated eigenvectors of the SCM $\mathbf{S}_N$ by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ and $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N$, in this paper, we look for estimators of the form

$$\hat{\mathbf{C}}_N^{-1}(\mathbf{S}_N) = \sum_{i=1}^N \eta_i\mathbf{u}_i\mathbf{u}_i^H$$

where $\eta_i > 0$ are eigenvalue "shrinkage" functions to be suitably designed for the MVDR application. With prior knowledge of the spiked covariance structure of $\mathbf{C}_N$, it is natural to apply "hard clipping" to the smallest $N-m$ sample eigenvalues, so that $\eta_{m+1} = \cdots = \eta_N = 1/\sigma^2$. This, in effect, decreases the estimation error in the SCM referred to as the *eigenvalue spreading* phenomenon [15]. Thus, we construct the estimator as

$$\hat{\mathbf{C}}_N^{-1}(\mathbf{S}_N) = \frac{1}{\sigma^2}\left(\mathbf{I}_N + \sum_{i=1}^m w_i\mathbf{u}_i\mathbf{u}_i^H\right) \tag{7}$$

where $w_i = \sigma^2\eta_i - 1$. Our aim is to find the optimal $\mathbf{w}^* = [w_1^*, \dots, w_m^*]^T$ which forms the optimized $\hat{\mathbf{C}}_{\mathrm{MVDR}}^{-1}$ that minimizes ρ in (5).

By plugging (7) and (6) into (5), the normalized total output power ρ is seen as a function of $\mathbf{w} = [w_1, \dots, w_m]^T$, which we denote as $\rho(\mathbf{w})$ in the following. In particular, our optimization problem now becomes

$$\mathbf{w}^* = \operatorname*{argmin}_{\mathbf{w}\in\mathcal{L}^m} \rho(\mathbf{w}) \tag{8}$$

where $\rho(\mathbf{w})$ is defined in (9) at the top of the next page. The parameter range is specified as[1] $\mathcal{L}^m = [-1+\xi, q)^m$, for some small $\xi > 0$ and large $q > 0$.

It is challenging to solve the problem (8) and obtain the optimal $\mathbf{w}^*$ in closed-form directly. Even if such $\mathbf{w}^*$ was obtained, because of the unobservable quantities of t_i and $\mathbf{v}_i$ involved, it could not be used in reality. We tackle these problems by appealing to results on the asymptotic properties of spiked covariance matrices to give a simplified asymptotic representation $\bar{\rho}(\mathbf{w})$ of $\rho(\mathbf{w})$ as $N, n \to \infty$ and obtain the oracle $\bar{\mathbf{w}}^*$ that minimizes $\bar{\rho}(\mathbf{w})$ depending on $\mathbf{C}_N$. Then we provide a sample-based consistent estimator $\hat{\mathbf{w}}^*$ which constructs the optimal $\hat{\mathbf{C}}_{\mathrm{MVDR}}^{-1}$ and consequently $\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$. For our asymptotic analysis, we assume the following:

Assumption 1.

a. As $N, n \to \infty$, $N/n = c_N \to c$ for a certain $c > 0$.

b. The number of spikes m is fixed, independently of N and n, while $t_1 > \cdots > t_m$ with $t_m > \sqrt{c}$, for all large N.

Remark 1. *In Assumption 1.b, the quantity $\sqrt{c}$ represents a fundamental "phase transition" point [16, 17]; that is, for each $i \in \{1, \dots, m\}$ such that for large N, $t_i > \sqrt{c}$, there is a deterministic one-to-one mapping between t_i and λ_i under Assumption 1.a, which can be used to estimate t_i. As for $t_i \leq \sqrt{c}$ for large N, the relation no longer holds, and the value of t_i can no longer be estimated. A similar comment also applies for eigenvectors. Hence, we assume $t_i > \sqrt{c}$ for large N in order to estimate $\mathbf{C}_N$ fully.*

[1] Note that we restrict $\mathbf{w}$ to a bounded set $\mathcal{L}^m$, which is a technical condition employed for establishing the uniform convergence result in Theorem 1 presented later.

$$\rho(\mathbf{w}) = \frac{\mathbf{a}^H(\theta_1)\left(\mathbf{I}_N + \sum_{i=1}^m w_i \mathbf{u}_i \mathbf{u}_i^H\right)\left(\mathbf{I}_N + \sum_{j=1}^m t_j \mathbf{v}_j \mathbf{v}_j^H\right)\left(\mathbf{I}_N + \sum_{h=1}^m w_h \mathbf{u}_h \mathbf{u}_h^H\right)\mathbf{a}(\theta_1)}{\left(\mathbf{a}^H(\theta_1)\left(\mathbf{I}_N + \sum_{l=1}^m w_l \mathbf{u}_l \mathbf{u}_l^H\right)\mathbf{a}(\theta_1)\right)^2} \quad (9)$$

3.2. Deterministic equivalent $\bar{\rho}(\mathbf{w})$ and the optimal $\bar{\mathbf{w}}^*$

Define the deterministic quantities $k_i = \mathbf{a}^H(\theta_1)\mathbf{v}_i\mathbf{v}_i^H\mathbf{a}(\theta_1)$ and $s_i = \frac{1-c_N/t_i^2}{1+c_N/t_i}$, $i = 1, \dots, m$, with δ_{ij} the Kronecker-delta function. We have the following result:

Theorem 1. *[Deterministic equivalent] Let Assumption 1 hold. As* $N, n \to \infty$, $\sup_{\mathbf{w}\in\mathcal{L}^m} |\rho(\mathbf{w}) - \bar{\rho}(\mathbf{w})| \xrightarrow{\text{a.s.}} 0$ *where*

$$\bar{\rho}(\mathbf{w}) = \frac{\mathbf{w}^T\mathbf{B}\mathbf{w} + 2\mathbf{w}^T\mathbf{d} + a}{(1+\mathbf{w}^T\mathbf{e})^2}$$

with

$$\mathbf{B} = \operatorname{diag}\left[s_1k_1(1+t_1s_1), \dots, s_mk_m(1+t_ms_m)\right],$$
$$\mathbf{d} = \left[s_1k_1(1+t_1), \dots, s_mk_m(1+t_m)\right]^T,$$
$$\mathbf{e} = \left[s_1k_1, \dots, s_mk_m\right]^T, \quad a = 1 + \sum_{i=1}^m t_ik_i .$$

Proof: Details are provided in [14].

The following theorem provides the value $\bar{\mathbf{w}}^* = [\bar{w}_1^*, \dots, \bar{w}_m^*]^T$ that minimizes $\bar{\rho}(\mathbf{w})$.

Theorem 2. *[Optimal weights] Under the setting of Theorem 1,*

$$\bar{\mathbf{w}}^* = \underset{\mathbf{w}\in\mathcal{L}^m}{\operatorname{argmin}}\, \bar{\rho}(\mathbf{w})$$

where, for $i = 1, \dots, m$,

$$\bar{w}_i^* = \frac{t_i + c_N}{t_i^2 + t_i}(\psi - t_i), \quad \text{with } \psi = \frac{\sum_{j=1}^m \frac{k_j}{t_j}}{\sum_{j=1}^m \frac{k_j}{t_j^2}} .$$

Proof: Details are provided in [14].

3.3. Estimated optimal weights $\hat{\mathbf{w}}^*$ and proposed $\hat{\mathbf{h}}_{\text{MVDRopt}}$

The optimal weights $\bar{w}_i^*$, $i = 1, \dots, m$, present in Theorem 2 cannot be directly used in practice because of the involved unobservable quantities, i.e., t_i and $\mathbf{v}_i$. In the following we provide sample-based consistent estimators of these optimal weights to address this issue.

Theorem 3. *[Estimated optimal weights] Under the setting of Theorem 1, for all large n with probability one,* $\lambda_i > \sigma^2(1+\sqrt{c_N})^2$, $i = 1, \dots, m$, *and we have*

$$|\hat{w}_i^* - \bar{w}_i^*| \xrightarrow{\text{a.s.}} 0$$

where $\hat{w}_i^* = \frac{\hat{t}_i + c_N}{\hat{t}_i^2 + \hat{t}_i}\left(\hat{\psi} - \hat{t}_i\right)$, *in which* $\hat{\psi} = \frac{\sum_{j=1}^m \frac{\hat{k}_j}{\hat{t}_j}}{\sum_{j=1}^m \frac{\hat{k}_j}{\hat{t}_j^2}}$,

$$\hat{t}_i = \frac{\lambda_i/\sigma^2 + 1 - c_N + \sqrt{(\lambda_i/\sigma^2 + 1 - c_N)^2 - 4\lambda_i/\sigma^2}}{2} - 1,$$
$$\hat{k}_i = \frac{1 + c_N/\hat{t}_i}{1 - c_N/(\hat{t}_i)^2}\mathbf{a}^H(\theta_1)\mathbf{u}_i\mathbf{u}_i^H\mathbf{a}(\theta_1) .$$

Algorithm 1 Proposed MVDR beamformer construction

1. Compute the optimized shrinkage parameters $\hat{w}_i^*$, $i = 1, \dots, m$ according to Theorem 3.
2. Form the precision matrix estimator :
   $\hat{\mathbf{C}}_{\text{MVDR}}^{-1} = \frac{1}{\sigma^2}\left(\mathbf{I}_N + \sum_{i=1}^m \hat{w}_i^*\mathbf{u}_i\mathbf{u}_i^H\right)$.
3. Construct the MVDR beamformer:
   $\hat{\mathbf{h}}_{\text{MVDRopt}} = \frac{\hat{\mathbf{C}}_{\text{MVDR}}^{-1}\mathbf{a}(\theta_1)}{\mathbf{a}^H(\theta_1)\hat{\mathbf{C}}_{\text{MVDR}}^{-1}\mathbf{a}(\theta_1)}$.

Proof: Details are provided in [14].

This leads to our proposed MVDR beamformer, described in Algorithm 1.

4. NUMERICAL SIMULATIONS

In our simulations, we consider a uniform linear array with N identical omnidirectional sensors located half a wavelength apart. Each data point is computed by averaging over 200 independent Monte-Carlo trials. The sensor array receives $m = 6$ uncorrelated narrowband signals from the far field. The DoA of the SoI is $\theta_1 = 0^o$, while the DoAs of the interferers are $\theta_2 = 5^o$, $\theta_3 = 10^o$, $\theta_4 = 30^o$, $\theta_5 = 50^o$ and $\theta_6 = 70^o$. The noise is complex Gaussian with mean zero and variance one.

4.1. ρ-Performance and the deterministic equivalent

We first numerically study the convergence of our proposed algorithm in terms of the function ρ. Define SNR $= \frac{p_1}{\sigma^2}$ and INR $= \frac{p_i}{\sigma^2}$ (taken to be the same for all $i = 2, \dots, m$). In Fig. 1, for SNR $= 5$ dB, INR $= 30$ dB and $n = 2N$, four quantities are presented: the expectation $\mathbb{E}[\rho(\hat{\mathbf{w}}^*)]$ (computed empirically) with our proposed $\hat{\mathbf{w}}^*$ in Theorem 3; the asymptotic deterministic equivalent $\bar{\rho}(\bar{\mathbf{w}}^*)$ based on Theorem 1; the theoretical minimum (oracle) $\rho_{\min} = 1/(\sigma^2\mathbf{a}^H(\theta_1)\mathbf{C}_N^{-1}\mathbf{a}(\theta_1))$ and the expectation $\mathbb{E}[\rho(\hat{\mathbf{h}}_{\text{SMI}})]$ with the SMI. As expected, $\mathbb{E}[\rho(\hat{\mathbf{w}}^*)]$ converges to $\bar{\rho}(\bar{\mathbf{w}}^*)$ with the increase of N and n. It is also demonstrated that $\bar{\rho}(\bar{\mathbf{w}}^*)$ is close to $\rho_{\min}$, indicating the near-optimal performance of our proposed approach. On the other hand, $\mathbb{E}[\rho(\hat{\mathbf{h}}_{\text{SMI}})]$ is larger than $\mathbb{E}[\rho(\hat{\mathbf{w}}^*)]$ over the entire range of N.

4.2. Beamformer performance and comparison against previous methods

We compare the performance of $\hat{\mathbf{h}}_{\text{MVDRopt}}$ against the optimal MVDR beamformer (3) with known $\mathbf{C}_N$, as well as $\hat{\mathbf{h}}_{\text{SMI}}$, the SMI beamformer. For each method, we present the beampattern and the output SINR. For these comparisons, we fix INR $= 30$ dB and $n = 2N = 200$. Shown in Fig. 2(a), our method $\hat{\mathbf{h}}_{\text{MVDRopt}}$ achieves significantly enhanced noise suppression compared with $\hat{\mathbf{h}}_{\text{SMI}}$, in addition to placing deeper interference nulls. As for the SINR performance, shown in Fig. 2(b), the proposed beamformer

$\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$ uniformly yields higher SINR than $\hat{\mathbf{h}}_{\mathrm{SMI}}$ and is quite close to the optimal beamformer when SNR is low.

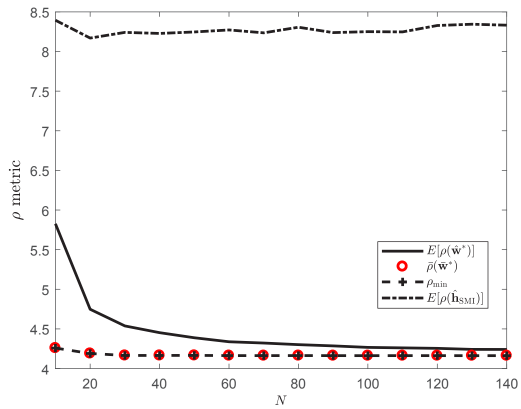
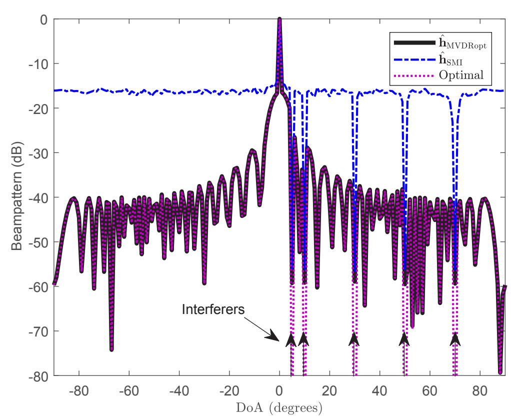


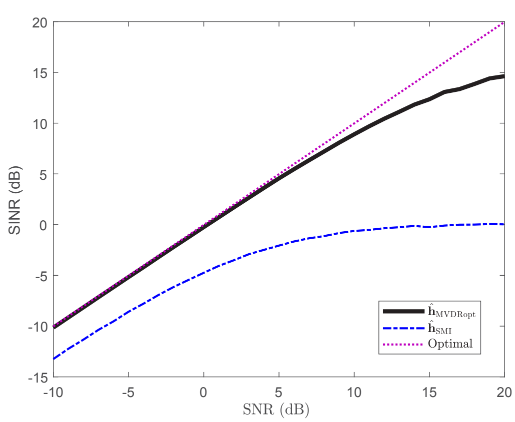
Fig. 1. ρ-performance when $n = 2N$.

Next, we compare with more advanced beamformers, which are designed to at least partially overcome the problem of sample insufficiency. One is the popular diagonal loading strategy $\hat{\mathbf{h}}_{\mathrm{DL}}$ (see, e.g., [6–8]), which employs the construction (4) with the covariance matrix estimator $\hat{\mathbf{C}}_N(\varphi) = (1-\varphi)\mathbf{S}_N + \varphi\mathbf{I}_N$ where $\varphi \in (0,1)$ is an empirically computed "oracle" solution, chosen to minimize the normalized total output power in (5). It provides the optimal performance achievable with the diagonal loading method. The other is the eigen-subspace beamformer $\hat{\mathbf{h}}_{\mathrm{Eigsub}} = \frac{\mathbf{S}_N^{-1}\mathbf{a}_{\mathrm{Sub}}}{\mathbf{a}_{\mathrm{Sub}}^H\mathbf{S}_N^{-1}\mathbf{a}_{\mathrm{Sub}}}$, where $\mathbf{a}_{\mathrm{Sub}} = \sum_{i=1}^m \mathbf{u}_i\mathbf{u}_i^H\mathbf{a}(\theta_1)$ [4, 5]. The beampatterns of these beamformers are shown in Fig. 3(a) for SNR $= 5$ dB. Despite that $\hat{\mathbf{h}}_{\mathrm{Eigsub}}$ exhibits slightly smaller "noise gain" than $\hat{\mathbf{h}}_{\mathrm{DL}}$, the proposed $\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$ has a significantly improved response, suppressing the noise the most. Fig. 3(b) shows the output SINR for different SNRs where $\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$ uniformly displays the highest SINR, performing comparable to $\hat{\mathbf{h}}_{\mathrm{DL}}$ and $\hat{\mathbf{h}}_{\mathrm{Eigsub}}$ at low SNR and high SNR respectively.

While not shown due to space constraint, we have conducted further comparison with MVDR beamformers constructed from alternative spike-model-based covariance matrix estimators [9]. This again revealed superior performance of our proposed method. Full details are presented in [14].

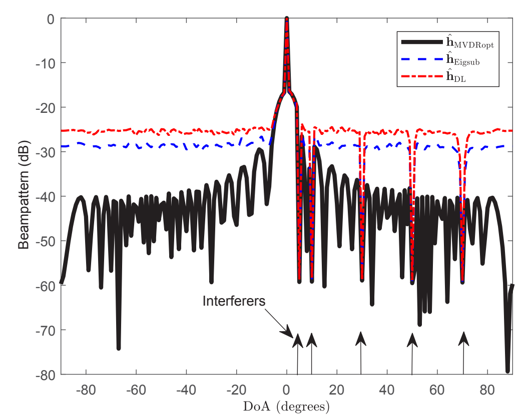


(a) Beampatterns (SNR $= 5$ dB)

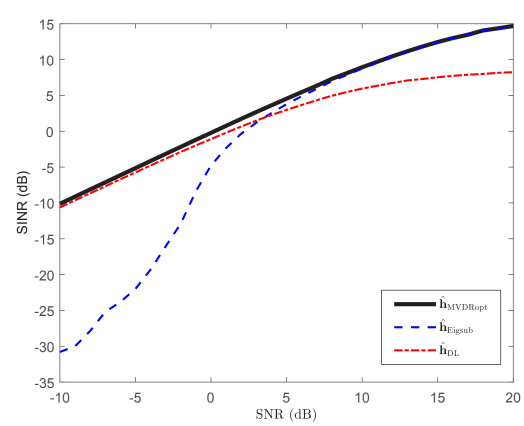


(b) Output SINR

Fig. 2. Beampattern and SINR performance comparison of $\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$ and $\hat{\mathbf{h}}_{\mathrm{SMI}}$ $(N = 100, n = 200)$.



(a) Beampatterns (SNR $= 5$ dB)



(b) Output SINR

Fig. 3. Beampattern and SINR performance comparison of $\hat{\mathbf{h}}_{\mathrm{MVDRopt}}$, $\hat{\mathbf{h}}_{\mathrm{Eigsub}}$ and $\hat{\mathbf{h}}_{\mathrm{DL}}$ $(N = 100, n = 200)$.

5. REFERENCES

[1] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[2] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.

[3] D. M. Boroson, “Sample size considerations for adaptive arrays,” *IEEE Trans. Aerosp. and Electron. Syst.*, vol. 24, no. 4, pp. 446–451, Jul. 1980.

[4] L. Chang and C.-C. Yeh, “Performance of DMI and eigenspace-based beamformers,” *IEEE Trans. Antennas Propag.*, vol. 40, no. 11, pp. 1336–1347, Nov. 1992.

[5] D. D. Feldman and L. J. Griffiths, “A projection approach for robust adaptive beamforming,” *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 867–876, Apr. 1994.

[6] B. D. Carlson, “Covariance matrix estimation errors and diagonal loading in adaptive arrays,” *IEEE Trans. Aerosp. and Electron. Syst.*, vol. 24, no. 4, pp. 397–401, July 1988.

[7] M. W. Ganz, R. Moses, and S. Wilson, “Convergence of the SMI and the diagonally loaded SMI algorithms with weak interference (adaptive array),” *IEEE Trans. Antennas Propagat.*, vol. 38, no. 3, pp. 394–399, Mar. 1990.

[8] X. Mestre and M. Lagunas, “Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays,” *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 69–82, Jan. 2006.

[9] D. L. Donoho, M. Gavish, and I. M. Johnstone, “Optimal shrinkage of eigenvalues in the spiked covariance model,” *to appear in Ann. Statist., arXiv preprint arXiv:1311.0851*, 2017.

[10] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: Wiley, 1999.

[11] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing*. New York: Wiley, 2002.

[12] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, “Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem,” *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[13] I. M. Johnstone, “On the distribution of the largest eigenvalue in principal components analysis,” *Ann. Statist.*, pp. 295–327, 2001.

[14] L. Yang, M. R. McKay, and R. Couillet, “High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models,” *to appear in IEEE Trans. Signal Process.*

[15] V. A. Marčenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices,” *Math. USSR-Sbornik*, vol. 1, no. 4, p. 457, 1967.

[16] J. Baik, G. B. Arous, and S. Péché, “Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices,” *Ann. Probab.*, vol. 33, no. 5, pp. 1643–1697, 2005.

[17] F. Benaych-Georges and R. R. Nadakuditi, “The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices,” *Adv. Math.*, vol. 227, no. 1, pp. 494–521, 2011.