ROBUST ADAPTIVE BEAMFORMING WITH IMPRECISE STEERING VECTOR AND NOISE COVARIANCE MATRIX DUE TO FINITE SAMPLE SIZE

Mengyi Zhang[1], *Francisco Rubio*[2], *Daniel P. Palomar*[1], *and Xavier Mestre*[3]

[1]Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong
[2]Genetic Finance, Ltd., Hong Kong
[3] Centre Tecnològic de Telecomnicacions de Catalunya, Spain

ABSTRACT

Abstract— Minimum variance beamformers are widely used for array signal processing. It is known that the diagonal loading method can improve the robustness against mismatches caused by the imprecise steering vector (or the channel vector) and the noise covariance matrix. Instead of concentrating on one aspect of the mismatches and assuming perfect knowledge of the other, we handle both estimation error in the steering vector and the noise covariance matrix caused by the finite sample size simultaneously. We employ high-dimensional asymptotics to reflect the finite sample size, and estimate the optimal loading factor based on random matrix theory. In an asymptotic setting where the number of samples is comparable to the array dimension, we obtain a beamformer that is as good as the beamformer with optimal diagonal loading. Monte Carlo simulations show the advantage of our beamformer in the finite sample size regime.

***Index Terms*—** Diagonal loading, minimum variance beamformer, finite sample size, imprecise steering vector, random matrix theory.

1. INTRODUCTION

Robust adaptive beamforming is a key issue in array applications such as communications, sonar, and radar [1], [2]. If an infinite number of samples are available and the signal steering vector or the channel vector is perfectly known, the minimum variance (MV) beamformer is optimal since it maximizes the signal-to-noise ratio (SNR) at the output of the beamformer.

However, in practice, the number of samples available at the receiver is not sufficiently high. The traditional MV beamformer based on sample matrix inversion (SMI) is known to have a detrimental effect on the performance since there is estimation error between the sample correlation matrix and the true correlation matrix of observations. Moreover, traditional MV beamforming lacks robustness against even small mismatches in the desired signal steering vector. One of the most popular approaches to improve the SMI-based MV technique is the diagonal loading method. It has been shown that the diagonal loading method can improve the robustness against mismatches cause by imprecise knowledge of the steering vector and finite sample size [3], [4]; see also references therein.

There is plenty of work aiming at solving these mismatches, but most of them focused on only one aspect. For example, the robust MV beamformers proposed in [5] and [6] incorporated uncertainty constraints on the steering vector but did not consider the finite sample size effect. The presence of random steering vector was considered in [7] and a generalized loading of the covariance matrix was applied. In other papers such as [4], the main focus was to deal with the finite sample size effect, and a perfect steering vector was assumed. Both types of mismatches were considered and handled in a deterministic way [8], i.e., the worst-case design. However, the performance could be affected by improper uncertainty set modeling and the choice of some parameters.

In this paper, we handle both types of mismatches in the steering vector and the correlation matrix of the observations (or equivalently, the noise covariance matrix) simultaneously. Instead of using a presumed steering vector or channel vector, we use a pilot-assisted approach: in the training period, we estimate the steering vector and the noise covariance matrix with the pilot and the observations, and construct the beamformer using the diagonal loading method. Then in the evaluation period, we apply this beamformer to the observations that contain the data. We select the diagonal loading factor based on random matrix theory: to reflect the fact that the sample size is comparable to the dimension of the array, we employ high-dimensional asymptotics where both of them go to infinity. We derive the convergence of the two types of coupled errors and then correct them in the asymptotic regime. Our main contribution is to obtain an MV beamformer which is asymptotically as good as that with an optimal diagonal loading factor.

2. PROBLEM FORMULATION

Let us consider the M-dimensional antenna array with a sample size of N snapshots. Let $\mathbf{y}(n) \in \mathbb{C}^{M\times 1}$, $n = 1, ..., N$, denote a collection of received signal. At snapshot n, it can be expressed as

$$\mathbf{y}(n) = \mathbf{h}s(n) + \mathbf{n}(n). \tag{1}$$

Here, $\mathbf{h} \in \mathbb{C}^{M\times 1}$ is the steering vector or the channel vector and $s(n)$ is the transmitting signal. Without loss of generality, we can assume $\mathsf{E}(|s(n)|^2) = 1$. The vector $\mathbf{n}(n) \in \mathbb{C}^{M\times 1}$ is the noise which is independent and identical distributed (i.i.d.) Gaussian with mean zero and a covariance matrix $\mathbf{R}_n$. Note that $\mathbf{h}$ is not restricted to be the conventional deterministic steering vector pointing at a certain direction, but can also be the channel vector with arbitrary structure. In the training period, we assume $s(n)$ is known to us.

We let $\mathbf{w} \in \mathbb{C}^{M\times 1}$ denote the beamformer weights, then the output of the array at the nth snapshot can be expressed as

$$x(n) = \mathbf{w}^H\mathbf{y}(n). \tag{2}$$

It is shown in e.g. [2] that the MV beamformer is the optimal solution to the following problem:

$$\begin{array}{ll} \underset{\mathbf{w}}{\mathsf{minimize}} & \mathbf{w}^H\mathbf{R}_y\mathbf{w} \\ \mathsf{subject\ to} & \mathbf{w}^H\mathbf{h} = 1 \end{array} \tag{3}$$

This work was supported by the Hong Kong RGC 617911 research grant.

where $\mathbf{R}_y = \frac{1}{N}\mathsf{E}\left[\mathbf{y}(n)\mathbf{y}(n)^H\right]$ is the correlation matrix of the observations. In (3), the correlation matrix of the observations can also be replaced with the noise covariance matrix $\mathbf{R}_n$ [9] due to the identity $\mathbf{R}_y = \mathbf{R}_n + \mathbf{h}\mathbf{h}^H$, and the formulation is equivalent to the original one. In the following, we mainly focus on the noise covariance matrix $\mathbf{R}_n$. This problem admits a closed form solution, which is the clairvoyant (since it contains unknown $\mathbf{R}_n$ and $\mathbf{h}$), namely,

$$\mathbf{w}_{\mathsf{clv}} = \frac{\mathbf{R}_n^{-1}\mathbf{h}}{\mathbf{h}^H\mathbf{R}_n^{-1}\mathbf{h}}. \tag{4}$$

The clairvoyant beamformer maximizes the output SNR, which is the objective of the MV beamformer:

$$\mathsf{SNR} = \frac{|\mathbf{w}^H\mathbf{h}|^2}{\mathbf{w}^H\mathbf{R}_n\mathbf{w}}. \tag{5}$$

In practice, $\mathbf{R}_n$ and $\mathbf{h}$ are not known so neither is the optimal beamformer, therefore, we estimate all of them with the training sequence $s(n)$ and the observations $\mathbf{y}(n)$. We define the following matrix notations for simplicity: let $\mathbf{Y} = [\mathbf{y}(1), ..., \mathbf{y}(N)]$, $\mathbf{s} = [s(1), ..., s(N)]^T$ and $\mathbf{N} = [\mathbf{n}(1), ..., \mathbf{n}(N)]$, then (1) can be written in a matrix form:

$$\mathbf{Y} = \mathbf{h}\mathbf{s}^T + \mathbf{N}, \tag{6}$$

without loss of generality, we can assume $\frac{1}{N}||\mathbf{s}||_2^2 = 1$. (If it is not, we can absorb that constant in $\mathbf{h}$). Traditionally, we estimate the steering vector using

$$\hat{\mathbf{h}} = \frac{1}{N}\mathbf{Y}\mathbf{s}^*, \tag{7}$$

and estimate the covariance matrix of the noise using the sample covariance matrix (SCM)

$$\hat{\mathbf{R}}_{n,\mathsf{SCM}} = \frac{1}{N}(\mathbf{Y} - \hat{\mathbf{h}}\mathbf{s}^T)(\mathbf{Y} - \hat{\mathbf{h}}\mathbf{s}^T)^H. \tag{8}$$

The beamformer $\mathbf{w}_{\mathsf{trd}}$ will be in the same form as in (4) while replacing $\mathbf{R}_n$ and $\mathbf{h}$ with $\hat{\mathbf{R}}_{n,\mathsf{SCM}}$ and $\hat{\mathbf{h}}$, respectively. If the sample size N is large enough, $\hat{\mathbf{R}}_{n,\mathsf{SCM}}$ and $\hat{\mathbf{h}}$ are good estimates of $\mathbf{R}_n$ and $\mathbf{h}$, and hence $\mathbf{w}_{\mathsf{trd}}$ is a good estimate of $\mathbf{w}_{\mathsf{clv}}$. However, in many scenarios, the number of available snapshots is not sufficiently high. In order to mitigate the finite-sample-size effects, the traditional solution is modified by the diagonal loading method:

$$\mathbf{w}_{\mathsf{dl}} = \frac{\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}}}{\hat{\mathbf{h}}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}}}. \tag{9}$$

The choice of the diagonal loading factor ρ is crucial and tunes (9) to provide a response between between the traditional SMI filter ($\rho = 0$) and the matched filter $\hat{\mathbf{h}}$ ($\rho \to \infty$). Our objective is to maximize the output SNR, and therefore we substitute (9) into the SNR expression in (5) and obtain

$$\mathsf{SNR}_{\mathsf{dl}} = \frac{|\mathbf{h}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}}|^2}{\hat{\mathbf{h}}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{R}_n\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}}}. \tag{10}$$

In (10), the precise knowledge of $\mathbf{h}$ is partially used since it contains the estimated $\hat{\mathbf{h}}$. The corresponding loading factor ρ maximizing (10) would be suboptimal compared to ρ maximizing the following expression, which assumes the steering vector is perfectly known:

$$\mathsf{SNR}_{\mathsf{dl-s}} = \frac{|\mathbf{h}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{h}|^2}{\mathbf{h}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{R}_n\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{h}}. \tag{11}$$

Ideally, if we knew the true covariance matrix $\mathbf{R}_n$ and the steering vector $\mathbf{h}$, we could directly select ρ which maximizes (10) and obtain the corresponding beamformer. However, in our case, $\mathbf{R}_n$ and $\mathbf{h}$ cannot be known. We tackle this problem using random matrix theory: we consider a more practical scenario where the sample size N is comparable to the array dimension M. Mathematically speaking, this is formulated as the asymptotic regime where M and N both go to infinity with certain ratio. We first derive a deterministic quantity that (10) converges to at the double limit, and then based on the deterministic quantity, we provide an estimator which also approaches to the deterministic quantity at the double limit. Since the consistent estimator of (10) depends only on the observations $\mathbf{y}(n)$, the transmitting signal $s(n)$, and the diagonal loading factor ρ, we can maximize it to obtain the optimal ρ and obtain $\mathbf{w}_{\mathsf{dl}}$.

In the following, we will provide the main results on the deterministic equivalent and consistent estimator of SNR in (10).

3. ASYMPTOTIC OUTPUT SNR WITH DIAGONAL LOADING

In this section we provide our theoretical results on the convergence results of (10). We first begin with technical hypotheses and some further definitions.

3.1. Assumptions and further definitions

The following set of assumptions will be maintained throughout the paper.

(A1) Let the spectral norm of $\mathbf{R}_n$ and the Euclidean norm of $\mathbf{h}$ be bounded uniformly in M and N.

(A2) Let $\mathbf{Z}$ be an $M \times N$ matrix whose elements $\mathbf{Z}_{ij}$ are i.i.d. standardized Gaussian random variables. Then the noise matrix can be written as $\mathbf{N} = \mathbf{R}_n^{1/2}\mathbf{Z}$.

We consider the limiting regime defined by both M and N growing large without bound at the same rate, i.e., $M, N \to \infty$ such that $0 < \mathsf{liminf}\ c \leq \mathsf{limsup}\ c < \infty$, with $c = M/N$. In this limiting regime, $a \asymp b$ denotes they are asymptotic equivalents, i.e., $|a - b| \to 0$ almost surely.

Before proceeding to the main theorems in this paper, we introduce some further definitions: Let $\mathbf{T} = \mathbf{I} - \frac{1}{N}\mathbf{s}^*\mathbf{s}^T$, and note that $\mathbf{T}\mathbf{T}^H = \mathbf{T}$. Therefore, (8) can be written as $\hat{\mathbf{R}}_{n,\mathsf{SCM}} = \frac{1}{N}\mathbf{Y}\mathbf{T}\mathbf{Y}^H$. Moreover, we define $\{\delta, \tilde{\delta}\}$ being the unique positive solutions to the following system of equations [10]:

$$\begin{cases} \tilde{\delta} & = \frac{1}{N}\mathsf{tr}\left[\mathbf{T}(\mathbf{I} + \delta\mathbf{T})^{-1}\right] \\ \delta & = \frac{1}{N}\mathsf{tr}\left[\mathbf{R}_n(\tilde{\delta}\mathbf{R}_n + \rho\mathbf{I})^{-1}\right], \end{cases} \tag{12}$$

and also

$$\begin{cases} \tilde{\gamma} & = \frac{1}{N}\mathsf{tr}\left[\left(\mathbf{T}(\mathbf{I} + \delta\mathbf{T})^{-1}\right)^2\right] \\ \gamma & = \frac{1}{N}\mathsf{tr}\left[\left(\mathbf{R}_n(\tilde{\delta}\mathbf{R}_n + \rho\mathbf{I})^{-1}\right)^2\right] \end{cases} \tag{13}$$

which are essential quantities for the asymptotic equivalent and estimator of SNR with diagonal loading.

We next decompose the numerator and denominator of $\mathsf{SNR}_{\mathsf{dl}}$ into the following components. Let $\mathbf{v} = \frac{1}{N}\mathbf{N}\mathbf{s}^*$ and recall that $\hat{\mathbf{h}} = \mathbf{h} + \mathbf{v}$, so that these components can be written as:

$$\xi_1 = \mathbf{h}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{h} \tag{14}$$

$$\xi_2 = \mathbf{h}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{v} \tag{15}$$

$$\xi_3 = \mathbf{v}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{v} \tag{16}$$

$$\xi_4 = \mathbf{h}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{R}_n \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{h} \tag{17}$$

$$\xi_5 = \mathbf{h}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{R}_n \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{v} \tag{18}$$

$$\xi_6 = \mathbf{v}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{R}_n \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \mathbf{v}. \tag{19}$$

With (14)-(19), the output SNR is given by:

$$\mathsf{SNR} = \frac{|\xi_1 + \xi_2|^2}{\xi_4 + \xi_5 + \xi_5^* + \xi_6}. \tag{20}$$

3.2. Asymptotic equivalent of the output SNR

The following theorem shows asymptotic convergence of (14)-(19).

Theorem 1. *Define the following deterministic quantities,*

$$\bar{\xi}_1 = \mathbf{h}^H(\tilde{\delta}\mathbf{R}_n + \rho\mathbf{I})^{-1}\mathbf{h} \tag{21}$$

$$\bar{\xi}_2 = 0 \tag{22}$$

$$\bar{\xi}_3 = \delta \tag{23}$$

$$\bar{\xi}_4 = \frac{1}{1-\gamma\tilde{\gamma}}\mathbf{h}^H(\tilde{\delta}\mathbf{R}_n + \rho\mathbf{I})^{-2}\mathbf{h} \tag{24}$$

$$\bar{\xi}_5 = 0 \tag{25}$$

$$\bar{\xi}_6 = \frac{\gamma}{1-\gamma\tilde{\gamma}} \tag{26}$$

Under Assumptions (A1)-(A2), we have $\xi_i \asymp \bar{\xi}_i,\ i = 1, ..., 6$, i.e., the random quantities in (14)-(19) behave as the deterministic quantities (21)-(26) in the double limit.

The proofs of Theorem 1 and the following Theorem 2 are omitted for lack of space. We mainly use convergence results in [10] and [11]. The details will be included in a full version of this paper which is now in preparation.

The convergence results of ξ_1 and ξ_4 coincide with the results in [4]. Theorem 1 enables us to analyze the asymptotic output SNR in (10) because it is enough to analyze its asymptotic equivalent, which is easier to characterize because of its deterministic nature. We will next show that, Theorem 1 helps to derive the estimator of the output SNR which only depends on the observations, the training sequence and the diagonal loading factor, so that we can calibrate the diagonal loading factor to maximize the estimator of the output SNR.

4. CONSISTENT ESTIMATION OF THE OPTIMAL LOADING FACTOR

In order to calibrate the loading factor, we have to obtain an observable estimator of (10). We begin with the following lemma which provides a consistent estimator of δ.

Lemma 1. *([10]) Under Assumptions (As1) to (As3), a consistent estimator of δ, denoted by $\hat{\delta}$, is given by:*

$$\frac{1}{N}\mathsf{tr}\left[\hat{\mathbf{R}}_{n,\mathsf{SCM}}\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\right] = \hat{\delta}\frac{1}{N}\mathsf{tr}\left[\mathbf{T}(\mathbf{I} + \hat{\delta}\mathbf{T})^{-1}\right]. \tag{27}$$

Note that in [10] and [11], the above result holds for a general positive semidefinite $\mathbf{T}$ with bounded spectral norm. However in our case, it can be simplified further with the particular structure of $\mathbf{T}$. Recall that $\mathbf{T} = \mathbf{I} - \frac{1}{N}\mathbf{s}^*\mathbf{s}^T$, so that $\mathbf{T}$ has an eigenvalue 1 with multiplicity $N-1$ and an eigenvalue 0 with multiplicity 1. Therefore, letting $D = \frac{1}{N}\mathsf{tr}\left[\hat{\mathbf{R}}_{n,\mathsf{SCM}}\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\right]$, we can show that

$$\hat{\delta} = \frac{D/(1-1/N)}{\alpha(1 - D/(1-1/N))}. \tag{28}$$

We propose to estimate the numerator and denominator of (20) separately. We first deal with the numerator which is easier, because only $\mathbf{h}$ cannot be observed. Hence, we provide a candidate estimate of it, i.e., $\hat{\mathbf{h}}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}}$. Noting that

$$\begin{aligned}&\mathbf{h}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}} \\ &= \hat{\mathbf{h}}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}} - \xi_2 - \xi_3\end{aligned} \tag{29}$$

together with $\xi_2 \asymp 0$ (as stated in Theorem 1 (22)) and $\xi_3 \asymp \hat{\delta}$ (as stated in Theorem 1 (23) and Lemma 2 (27)), we can easily obtain a consistent estimator of the numerator.

Now we proceed to provide the estimator of the denominator of (20). We consider the following conventional (plug-in) estimator which replaces the unknown $\mathbf{R}_n$ with $\hat{\mathbf{R}}_{n,\mathsf{SCM}}$, i.e.,

$$\mathsf{DNM}_{\mathsf{cnv}} = \hat{\mathbf{h}}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{n,\mathsf{SCM}}\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{h}}. \tag{30}$$

Similarly, we can decompose $\mathsf{DNM}_{\mathsf{cnv}}$ into the following components:

$$\mathsf{DNM}_{\mathsf{cnv}} = \xi_{4,\mathsf{cnv}} + \xi_{5,\mathsf{cnv}} + \xi_{5,\mathsf{cnv}}^* + \xi_{6,\mathsf{cnv}} \tag{31}$$

where

$$\xi_{4,\mathsf{cnv}} = \mathbf{h}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{n,\mathsf{SCM}}\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{h} \tag{32}$$

$$\xi_{5,\mathsf{cnv}} = \mathbf{h}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{n,\mathsf{SCM}}\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{v} \tag{33}$$

$$\xi_{6,\mathsf{cnv}} = \mathbf{v}^H\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{n,\mathsf{SCM}}\left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1}\mathbf{v}, \tag{34}$$

Regarding the components (32)-(34), we have the following results:

Theorem 2. *Under Assumptions (A1)-(A2), the following convergence results hold true:*

$$a\xi_{4,\mathsf{cnv}} \asymp \bar{\xi}_4,\ \xi_{5,\mathsf{cnv}} \asymp 0,\ \textit{and}\ a\xi_{6,\mathsf{cnv}} \asymp \bar{\xi}_6 \tag{35}$$

where $a = \frac{1}{\frac{1}{N}\mathsf{tr}\left[\mathbf{T}(\mathbf{I}+\hat{\delta}\mathbf{T})^{-2}\right]}$.

Theorem 2 shows that $a\mathsf{DNM}_{\mathsf{cnv}}$ is a consistent estimator of the denominator of (20) since they have the same asymptotic equivalent.

With the above theorems and observations on the estimators of numerator and denominator of (20), we can claim the following:

Claim 1. The following expression is a consistent estimator of the output SNR with diagonal loading in (10), as a function of the loading factor ρ:

$$\mathsf{SINR}_{\mathsf{dl-est}}(\rho) = \frac{|\hat{\mathbf{h}}^H \left(\hat{\mathbf{R}}_{n,\mathsf{SCM}} + \rho\mathbf{I}\right)^{-1} \hat{\mathbf{h}} - \hat{\delta}(\rho)|^2}{a(\rho)\mathsf{DNM}_{\mathsf{cnv}}(\rho)}. \tag{36}$$

where $\mathsf{DNM}_{\mathsf{cnv}}(\rho)$ is given in (30) and $a(\rho)$ is defined in Theorem 2.

Now we can use exhaustive search to find the optimal ρ that maximizes $\mathsf{SINR}_{\mathsf{dl-est}}(\rho)$ in (36).

5. MONTE CARLO SIMULATIONS

In the simulation, we show the advantage of our MV beamformer in finite sample size settings. We let the channel vector $\mathbf{h}$ have complex Gaussian distributed entries with mean zero and variance one, and assume that $\mathbf{s}$ is complex Gaussian distributed. The covariance matrix of noise vector is generated $\mathbf{R}_n(i,j) = \sigma_n^2 \cdot 0.7^{|i-j|}$, where transmitting SNR is $1/\sigma_n^2$. We compare our proposed RMT estimator with the following estimators:

(Clairvoyant): It is the upper bound for all types of beamformers, refer to (4).

(Optimal diagonal loading): It is the upper bound for all beamformers with a diagonal loading structure, whose loading factor optimizes (11). The true covariance matrix $\mathbf{R}_n$ is known and a perfect steering vector is assumed, but the covariance matrix estimator is restricted to be a linear combination of the sample covariance matrix and the identity.

(LSMI beamformer): The loading factor is chosen to be $10\sigma_n^2$ which has been empirically shown to be a suitable value in [5].

(Traditional SMI): It is the traditional sample covariance matrix inversion method.

In our proposed method, we replace $\mathbf{I}$ with $\frac{1}{M}\mathsf{tr}(\hat{\mathbf{R}}_{n,\mathsf{SCM}})\mathbf{I}$ in (36) to properly scale ρ and simplify the search process. Since the corrected noise covariance matrix is a linear combination of $\hat{\mathbf{R}}_{n,\mathsf{SCM}}$ and $\mathbf{I}$ with weights 1 and ρ, respectively, it is better to make $\hat{\mathbf{R}}_{n,\mathsf{SCM}}$ and $\mathbf{I}$ in similar scale (which could differ a lot due to different transmitting SNR).

Our experiments is conducted 200 times and the average is compared.

We plot the output SNR versus the number of training samples in Fig. 1. It can be seen that the output SNR of the clairvoyant is constant, since this method does not depend on the samples. The performance of our proposed beamformer is very close to that of optimal diagonal loading method, even when the sample size is small, the proposed method dominates the LSMI beamformer and the traditional SMI method. When N becomes larger and larger, the performance of the SMI method becomes better. The performance of LSMI does not change much, since it exploits a fixed and large diagonal loading factor. We plot the output SNR versus transmitting SNR in Fig. 2. It can be seen that the performance of our proposed beamformer is very close to that of optimal diagonal loading method under all transmitting SNRs. When transmitting SNR is low, the difference of our proposed beamformer is larger because of the error in estimating the steering vector. Moreover, in this scenario, our proposed method outperforms LSMI by approximately 2dB and SMI by nearly 3dB.

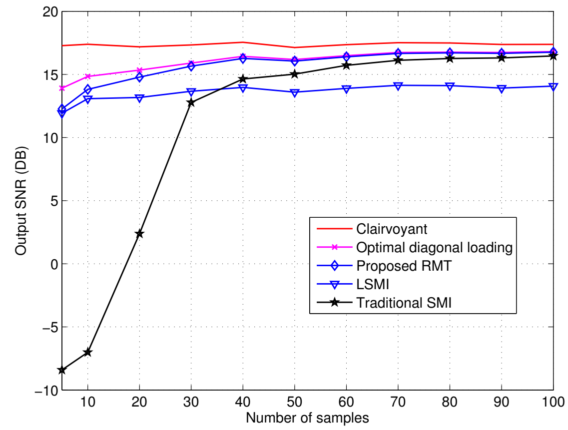

Fig. 1. Output SNR: transmitting SNR is 0 dB, $M = 20$, N varies from 5 to 100

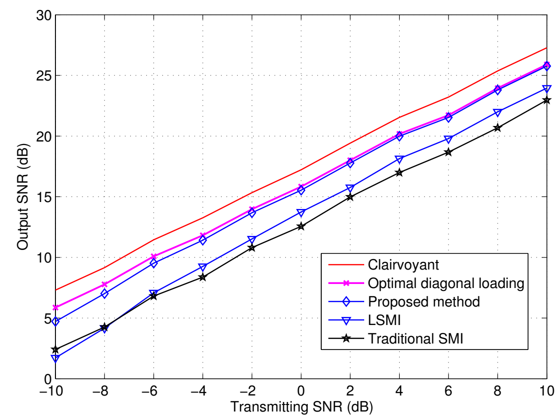

Fig. 2. Output SNR: $M = 20$, $N = 30$, transmitting SNR varies from -10 dB to 10dB

6. CONCLUSION

In this paper we have provided an MV beamformer with diagonal loading, which is robust against the imprecise steering vector and the insufficiency of the samples. We have employed high-dimensional asymptotics to reflect the fact that the number of samples is comparable to the array dimension, and estimate the optimal loading factor based on random matrix theory. Monte Carlo simulations have shown the competitive advantage of our proposed method under realistic finite sample-size settings.

7. REFERENCES

[1] H. Van Trees and J. Wiley, *Optimum array processing*. New York: Wiley, 2002.

[2] A. Gershman, N. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamform-

ing," *Signal Processing Magazine, IEEE*, vol. 27, pp. 62 –75, May 2010.

[3] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. on Signal Processing*, vol. 51, pp. 1702 – 1715, Jul. 2003.

[4] X. Mestre and M. Lagunas, "Finite sample size effect on minimum variance beamformers: optimum diagonal loading factor for large arrays," *IEEE Trans. on Signal Processing*, vol. 54, pp. 69 – 82, Jan. 2006.

[5] S. Vorobyov, A. 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, pp. 313 – 324, Feb. 2003.

[6] L. Lei, J. Lie, A. Gershman, and C. See, "Robust adaptive beamforming in partly calibrated sparse sensor arrays," *IEEE Trans. on Signal Processing*, vol. 58, no. 3, pp. 1661–1667, 2010.

[7] O. Besson 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, pp. 452–459, Feb. 2005.

[8] S. Shahbazpanahi, A. Gershman, Z. Luo, and K. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2257–2269, 2003.

[9] D. Johnson and D. Dudgeon, *Array signal processing: Concepts and techniques*. PTR Prentice Hall, Englewood Cliffs, NJ, 1993.

[10] F. Rubio, X. Mestre, and D. P. Palomar, "Performance analysis and optimal selection of large minimum variance portfolios under estimation risk," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, pp. 337 –350, Aug. 2012.

[11] M. Zhang, F. Rubio, and D. P. Palomar, "Improved calibration of high-dimensional precision matrices," *IEEE Trans. on Signal Processing*, vol. 61, no. 6, pp. 1509–1519, 2013.