Automatic Generalized Loading for Robust Adaptive Beamforming

Jun Yang, Xiaochuan Ma, Chaohuan Hou, *Fellow, IEEE*, and Yicong Liu

***Abstract*—The goal of this letter is to derive robust adaptive beamformers via generalized loading. In the proposed methods, Hermitian matrices are loaded on sample covariance matrix, and this is different from those methods based on the well-known diagonal loading approach. Furthermore, the computation of the loaded matrix is fully automatic, which is scarce in the literature. Numerical examples show that our methods are more robust to errors on array steering vector and sample covariance matrix than other tested parameter-free methods.**

***Index Terms*—Adaptive beamforming, generalized loading, minimum variance beamforming, robust beamforming.**

I. Introduction

Many approaches have been proposed during the past three decades to improve the robustness of adaptive beamformers. In the last few years, some methods with a clear theoretical background which make explicit use of an uncertainty set of the array steering vector have been proposed, such as [1]–[6]. However, the performance of the aforementioned methods depends on the choice of user parameters which are often related to, for example, the knowledge of errors on steering vector or variances of the signal-of-interest (SOI). Thus, the user parameters may be difficult to estimate in practice.

Fully parameter-free robust adaptive beamformers are scarce. One example is the HKB [7] via ridge regression (RR) based on the generalized sidelobe canceller (GSC) parameterization of the standard Capon beamformer (SCB). This approach can be extended to other well-investigated methods in the literature such as principal component regression (PCR) [8] and partial least squares (PLS) [9]. Another example is the method via the general linear combination (GLC) shrinkage-based covariance matrix estimation [10] which has been demonstrated to be useful in the case of small sample sizes. Note that both HKB and GLC can be seen as diagonal loading algorithms.

In this letter, we formulate a general form of the loading matrix with automatic parameter determination, which can be seen as a generalized loading approach in which a Hermitian matrix is loaded on sample covariance matrix instead of diagonal matrix. All the existing techniques such as diagonal loading methods, PCR, and PLS can be found as special cases of the general form. Based on the formulation, we also propose two special generalized loading algorithms. Simulation results show that the proposed methods are more robust to errors on steering vector and sample covariance matrix than other tested parameter-free methods.

Manuscript received August 28, 2008; revised October 20, 2008. Current version published February 11, 2009. This work was supported by the Commission of Science Technology and Industry for National Defence, China, under Grant No. A1320070067. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yimin Zhang.

J. Yang is with the Institute of Acoustics, Chinese Academy of Sciences (CAS) and Graduate University of CAS, Beijing 100190, China (e-mail: yangjun@mail.ioa.ac.cn).

X. Ma and C. Hou are with Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China (e-mail: maxc@mail.ioa.ac.cn; hch@mail.ioa.ac.cn).

Y. Liu is with the State Key Laboratory of Information Security, Graduate University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: ycliu@is.ac.cn).

Digital Object Identifier 10.1109/LSP.2008.2010807

II. Problem Formulation

Assume that there is an m element array of sensors. Let $\boldsymbol{R}$ and $\boldsymbol{a}$ denote the covariance matrix of the array output vector and the array steering vector. Without loss of generality, we suppose that the steering vectors are normalized such that $\|\boldsymbol{a}\|^2 = m$. The SCB is formulated so as to select a weight vector that minimizes the array output power by using the following linearly constrained quadratic problem:

$$\min_{\boldsymbol{w}} \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w}, \qquad \text{subject to } \boldsymbol{w}^{\mathrm{H}}\boldsymbol{a} = 1. \tag{1}$$

In practical applications, $\boldsymbol{R}$ is the finite sample covariance matrix $\boldsymbol{R} = \sum_{n=1}^{N} \boldsymbol{y}(n)\boldsymbol{y}^{\mathrm{H}}(n)/N$, where N is the number of snapshots and $\boldsymbol{y}(n)$ is the sample data. The solution to (1) is well known to be

$$\boldsymbol{w}_{\mathrm{SCB}} = \frac{\boldsymbol{R}^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{a}}. \tag{2}$$

Based on the SCB, the HKB can be formulated as a diagonal loading algorithm

$$\boldsymbol{w}_{\mathrm{HKB}} = \frac{(\boldsymbol{R}+\rho\boldsymbol{I})^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R}+\rho\boldsymbol{I})^{-1}\boldsymbol{a}} \tag{3}$$

where ρ is automatically computed using the Hoerl–Kernard–Baldwin formula (see [7] for details). The GLC which is based on shrinkage-based covariance matrix estimation can also be formulated as a diagonal algorithm

$$\boldsymbol{w}_{\mathrm{GLC}} = \frac{(\rho_1\boldsymbol{R}+\rho_2\boldsymbol{I})^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\rho_1\boldsymbol{R}+\rho_2\boldsymbol{I})^{-1}\boldsymbol{a}} = \frac{\left(\boldsymbol{R}+\frac{\rho_2}{\rho_1}\boldsymbol{I}\right)^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\left(\boldsymbol{R}+\frac{\rho_2}{\rho_1}\boldsymbol{I}\right)^{-1}\boldsymbol{a}} \tag{4}$$

in which ρ_1 and ρ_2 are chosen by minimizing the MSE of the shrinkage-based covariance matrix estimator (see [10] and [11]).

In this letter, we focus on the generalized loading approach which can be formulated as

$$\boldsymbol{w} = \frac{(\boldsymbol{R}+\Delta)^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R}+\Delta)^{-1}\boldsymbol{a}} \tag{5}$$

where $\boldsymbol{\Delta} = \boldsymbol{\Delta}^{\mathrm{H}}$ is the loaded matrix. We note that the generalized loading approach has more degrees of freedom than diagonal loading; thus, we can expect it to have great potential.

III. Shrinkage Estimator Based on GSC Parametrization

A. Definition of Shrinkage Estimator

In this section, we proposed a shrinkage estimator based on the GSC parameterization of SCB. The weight vector $\boldsymbol{w}$ in (1) can be reparameterized by a new parameter vector $\boldsymbol{\beta}$ according to

$$\boldsymbol{w} = \frac{\boldsymbol{a}}{m} - \boldsymbol{Q}\beta \tag{6}$$

where $\boldsymbol{Q} \in \mathbb{C}^{m\times(m-1)}$, $\boldsymbol{\beta} \in \mathbb{C}^{m-1}$, $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{a} = 0$, and $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q} = \boldsymbol{I}$. In this way, the formulation of (1) can be written as

$$\min_{\boldsymbol{\beta}} \left(\boldsymbol{Q}\beta - \frac{\boldsymbol{a}}{m}\right)^{\mathrm{H}} \boldsymbol{R}\left(\boldsymbol{Q}\beta - \frac{\boldsymbol{a}}{m}\right) = \min_{\boldsymbol{\beta}} \left\| \boldsymbol{R}^{1/2}\boldsymbol{Q}\boldsymbol{\beta} - \boldsymbol{R}^{1/2}\frac{\boldsymbol{a}}{m}\right\|^2 \triangleq \sigma^2 \tag{7}$$

with $\boldsymbol{R}^{1/2}$ to be the positive definite Hermitian square root of $\boldsymbol{R}$. The SCB is obtained as the standard least squared (LS) estimate of the theoretical $\tilde{\boldsymbol{\beta}}$ in a linear regression problem as

$$\boldsymbol{\beta}_{\mathrm{SCB}} = (\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q})^{-1}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\frac{\boldsymbol{a}}{m}. \tag{8}$$

Suppose that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{m-1}$ are the eigenvalues of $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q}$. Let the SVD decomposition of $\boldsymbol{R}^{1/2}\boldsymbol{Q}$ be

$$\boldsymbol{R}^{1/2}\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{H}} \tag{9}$$

where $\boldsymbol{\Sigma}^2 = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{m-1})$. Assuming $\boldsymbol{U} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{m-1})$ and $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m-1})$, we can rewrite $\boldsymbol{\beta}_{\mathrm{SCB}}$ as

$$\boldsymbol{\beta}_{\mathrm{SCB}} = \sum_{i=1}^{m-1}\left(\frac{1}{\sqrt{\lambda_i}}\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{R}^{1/2}\frac{\boldsymbol{a}}{m}\right)\boldsymbol{v}_i \triangleq \sum_{i=1}^{m-1}\boldsymbol{\beta}_i \tag{10}$$

in which $\boldsymbol{\beta}_i$ is the component of $\boldsymbol{\beta}_{\mathrm{SCB}}$ along $\boldsymbol{v}_i$. Since the LS estimate is well known to have no bias, according to the properties of LS estimate from the literature on linear model, we have

$$\mathrm{E}(\boldsymbol{\beta}_{\mathrm{SCB}}) = \tilde{\boldsymbol{\beta}},\ \mathrm{Cov}(\boldsymbol{\beta}_{\mathrm{SCB}}) = \sigma^2(\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q})^{-1} = \sigma^2\boldsymbol{V}\boldsymbol{\Sigma}^{-2}\boldsymbol{V}^{\mathrm{H}} \tag{11}$$

where $\mathrm{E}(\cdot)$ is the expectation operator and $\mathrm{Cov}(\cdot)$ denotes the covariance matrix.

Thus, we can define the shrinkage estimator of $\tilde{\boldsymbol{\beta}}$ as

$$\hat{\boldsymbol{\beta}} = \sum_{i=1}^{m-1}\xi_i\boldsymbol{\beta}_i \tag{12}$$

where $\xi_1, \xi_2, \ldots, \xi_{m-1}$ are shrinkage factors. Similar to (11), we can get

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{V}\mathrm{diag}\left(\frac{\xi_1^2}{\lambda_1}, \frac{\xi_2^2}{\lambda_2}, \ldots, \frac{\xi_{m-1}^2}{\lambda_{m-1}}\right)\boldsymbol{V}^{\mathrm{H}}. \tag{13}$$

The shrinakge factors in (12) can be chosen by minimizing the MSE of $\hat{\boldsymbol{\beta}}$, which is denoted by

$$\begin{aligned}\mathrm{MSE}(\hat{\boldsymbol{\beta}}) &\triangleq \mathrm{E}\|\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\|^2 = \mathrm{E}[(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})^{\mathrm{H}}(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})]\\ &= \mathrm{E}\{[\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}})]^{\mathrm{H}}[\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}})]\} + [\mathrm{E}(\hat{\boldsymbol{\beta}}) - \tilde{\boldsymbol{\beta}}]^{\mathrm{H}}[\mathrm{E}(\hat{\boldsymbol{\beta}}) - \tilde{\boldsymbol{\beta}}]\\ &= \mathrm{tr}\{\mathrm{E}[(\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}}))(\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}}))^{\mathrm{H}}]\} + \|\mathrm{E}(\hat{\boldsymbol{\beta}}) - \tilde{\boldsymbol{\beta}}\|^2\\ &= \mathrm{tr}(\mathrm{Cov}(\hat{\boldsymbol{\beta}})) + \|\mathrm{E}(\hat{\boldsymbol{\beta}}) - \tilde{\boldsymbol{\beta}}\|^2\end{aligned} \tag{14}$$

in which $\mathrm{tr}(\cdot)$ is the trace operator.

B. Minimized MSE Shrinkage Estimate

In this section, we consider the MSE minimization problem of (12). Denote $\boldsymbol{V}^{\mathrm{H}}\boldsymbol{\beta}_{\mathrm{SCB}}$ by $\boldsymbol{\alpha}_{\mathrm{SCB}}$ and $\boldsymbol{V}^{\mathrm{H}}\hat{\boldsymbol{\beta}}$ by $\hat{\boldsymbol{\alpha}}$, respectively. Since $\boldsymbol{V}$ is an orthonormal matrix, we get $\mathrm{MSE}(\hat{\boldsymbol{\alpha}}) = \mathrm{MSE}(\hat{\boldsymbol{\beta}})$. The minimization problem can be formulated by

$$\begin{aligned}&\min\nolimits_{\xi_1,\xi_2,\ldots,\xi_{m-1}}\mathrm{MSE}(\hat{\boldsymbol{\beta}})\\ &\quad= \min_{\xi_1,\xi_2,\ldots,\xi_{m-1}}\mathrm{MSE}(\hat{\boldsymbol{\alpha}})\\ &\quad= \min_{\xi_1,\xi_2,\ldots,\xi_{m-1}}\{\mathrm{tr}(\mathrm{Cov}(\hat{\boldsymbol{\alpha}})) + \|\mathrm{E}(\hat{\boldsymbol{\alpha}}) - \boldsymbol{V}^{\mathrm{H}}\tilde{\boldsymbol{\beta}}\|^2\}\\ &\quad= \min_{\xi_1,\xi_2,\ldots,\xi_{m-1}}\{\mathrm{tr}(\mathrm{Cov}(\hat{\boldsymbol{\alpha}})) + \|\mathrm{E}(\hat{\boldsymbol{\alpha}} - \boldsymbol{V}^{\mathrm{H}}\boldsymbol{\beta}_{\mathrm{SCB}})\|^2\}.\end{aligned} \tag{15}$$

According to (13), we get that $\mathrm{tr}(\mathrm{Cov}(\hat{\boldsymbol{\alpha}})) = \mathrm{tr}(\mathrm{Cov}(\hat{\boldsymbol{\beta}})) = \sigma^2\sum_{i=1}^{m-1}\xi_i^2/\lambda_i$. Then (15) equals to

$$\begin{aligned}&\min_{\xi_1,\xi_2,\ldots,\xi_{m-1}}\left\{\sigma^2\sum_{i=1}^{m-1}\frac{\xi_i^2}{\lambda_i} + \|\mathrm{E}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_{\mathrm{SCB}})\|^2\right\}\\ &\quad= \min_{\xi_1,\xi_2,\ldots,\xi_{m-1}}\sum_{i=1}^{m-1}\left\{\frac{\sigma^2}{\lambda_i}\xi_i^2 + (1-\xi_i)^2|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2\right\}\end{aligned} \tag{16}$$

in which $\boldsymbol{\alpha}_{\mathrm{SCB}} = (\boldsymbol{\alpha}_{\mathrm{SCB}1}, \boldsymbol{\alpha}_{\mathrm{SCB}2}, \ldots, \boldsymbol{\alpha}_{\mathrm{SCB}m-1})^{\mathrm{T}}$. Thus, the minimized MSE shrinkage factors are chosen by

$$\xi_i = \frac{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2 + \frac{\sigma^2}{\lambda_i}} = \frac{\lambda_i}{\lambda_i + \frac{\sigma^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2}}. \tag{17}$$

In practice, since both $|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2$ and σ^2 are unknown, we use $|\boldsymbol{\alpha}_{\mathrm{SCB}i}|^2$ instead of $|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2$ and

$$\hat{\sigma}^2 = \left\|\boldsymbol{R}^{1/2}\boldsymbol{Q}\boldsymbol{\beta}_{\mathrm{SCB}} - \boldsymbol{R}^{1/2}\frac{\boldsymbol{a}}{m}\right\|^2 \tag{18}$$

as an estimate of σ^2.

IV. Automatic Generalized Loading Method

A. Generalized Loading Robust Adaptive Beamformer

Using the shrinkage factors of (17), we can get a parameter-free beamformer. Denoting the diagonal matrix

$$\boldsymbol{\Xi} = \mathrm{diag}\left\{\frac{\sigma^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}1})|^2}, \frac{\sigma^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}2})|^2}, \ldots, \frac{\sigma^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}m-1})|^2}\right\} \tag{19}$$

we can compute the weight vector $\hat{\boldsymbol{w}}$ as

$$\hat{\boldsymbol{w}} = \frac{\boldsymbol{a}}{m} - \boldsymbol{Q}\hat{\boldsymbol{\beta}} = \frac{\boldsymbol{a}}{m} - \boldsymbol{Q}\sum_{1}^{m-1}\frac{\xi_i}{\sqrt{\lambda_i}}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{R}^{1/2}\frac{\boldsymbol{a}}{m}$$

$$= \frac{\boldsymbol{a}}{m} - \boldsymbol{Q}\sum_{1}^{m-1} \frac{\sqrt{\lambda_i}}{\lambda_i + \frac{\sigma^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2}} \boldsymbol{v}_i \boldsymbol{u}_i^{\mathrm{H}} \boldsymbol{R}^{1/2} \frac{\boldsymbol{a}}{m}$$
$$= [\boldsymbol{I} - \boldsymbol{Q}(\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q} + \boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}})^{-1}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}]\frac{\boldsymbol{a}}{m}. \quad (20)$$

Since $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{a} = 0$, we add

$$-\boldsymbol{Q}(\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q} + \boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}})^{-1}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}}\frac{\boldsymbol{a}}{m} \equiv 0 \quad (21)$$

to (20) and define $\hat{\boldsymbol{R}} \triangleq \boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}}$. Then (20) equals to

$$[\boldsymbol{I} - \boldsymbol{Q}(\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q} + \boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}})^{-1}\boldsymbol{Q}^{\mathrm{H}}(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}})]\frac{\boldsymbol{a}}{m}$$
$$= \hat{\boldsymbol{R}}^{-1/2}[\boldsymbol{I} - \hat{\boldsymbol{R}}^{1/2}\boldsymbol{Q}(\boldsymbol{Q}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{Q})^{-1}\boldsymbol{Q}^{\mathrm{H}}\hat{\boldsymbol{R}}^{1/2}]\hat{\boldsymbol{R}}^{1/2}\frac{\boldsymbol{a}}{m}. \quad (22)$$

We can see that $\boldsymbol{I} - \hat{\boldsymbol{R}}^{1/2}\boldsymbol{Q}(\boldsymbol{Q}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{Q})^{-1}\boldsymbol{Q}^{\mathrm{H}}\hat{\boldsymbol{R}}^{1/2}$ is the orthogonal projection matrix onto the complement of the column space of $\hat{\boldsymbol{R}}^{1/2}\boldsymbol{Q}$. Note that $\hat{\boldsymbol{R}}^{1/2}\boldsymbol{Q}$ has rank $m-1$ and

$$(\hat{\boldsymbol{R}}^{1/2}\boldsymbol{Q})^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1/2}\boldsymbol{a} = 0. \quad (23)$$

Consequently

$$\hat{\boldsymbol{w}} = \hat{\boldsymbol{R}}^{-1/2}\frac{\hat{\boldsymbol{R}}^{-1/2}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1/2}}{\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}}\hat{\boldsymbol{R}}^{1/2}\frac{\boldsymbol{a}}{m} = \frac{\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}}$$
$$= \frac{(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}})^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}})^{-1}\boldsymbol{a}} \triangleq \frac{(\boldsymbol{R} + \boldsymbol{\Delta}_1)^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R} + \boldsymbol{\Delta}_1)^{-1}\boldsymbol{a}}. (24)$$

B. Relations With Other Parameter-Free Methods

From (24), we note that the generalized loading algorithm can be seen as an extension to diagonal loading. Take the HKB, for example: the HKB of (3) can be obtained as a special case of (12) when $\xi_i = \lambda_i/(\lambda_i + \rho)$ for $i = 1, 2, \ldots, m-1$.

Furthermore, the methods based on PCR and PLS [8], [9] which are of the same type of RR can also be seen as special cases of (12). The PCR-based method can be obtained by setting $\xi_1 = \xi_2 = \cdots = \xi_d = 1$ and $\xi_{d+1} = \xi_{d+2} = \cdots = \xi_{m-1} = 0$, where d is the number of principal components. For the PLS-based method, we assume that p is the number of extracted score vectors. The space spanned by the columns of $[\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{a}/m, (\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q})\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{a}/m, \ldots, (\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q})^{p-1}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{a}/m]$ is called the p-dimensional Krylov space, which is denoted by $\mathcal{K}$. Then the $\boldsymbol{\beta}_{\mathrm{PLS}}$, which denotes the new parameter vector for PLS-based beamformer, is the solution of the optimization problem

$$\min_{\boldsymbol{\beta}} \left\| \boldsymbol{R}^{1/2}\frac{\boldsymbol{a}}{m} - \boldsymbol{R}^{1/2}\boldsymbol{Q}\boldsymbol{\beta} \right\|^2, \qquad \text{subject to } \boldsymbol{\beta} \in \mathcal{K}. \quad (25)$$

Letting $\boldsymbol{W}$ be an orthogonal basis of $\mathcal{K}$, the linear map $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q}$ restricted to $\mathcal{K}$ for an element $\boldsymbol{k} \in \mathcal{K}$ is defined as the orthogonal projection of $(\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q})\boldsymbol{k}$ onto the space $\mathcal{K}$. The map is represented by the $p \times p$ tridiagonal matrix $\boldsymbol{L} = \boldsymbol{W}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{Q}\boldsymbol{W}$. Letting the first p eigenvector-eigenvalue pairs (Ritz pairs) of $\boldsymbol{L}$ be $(\boldsymbol{r}_i, \mu_i)$, then the PLS-based method can be denoted by choosing shrinkage factors as $\xi_i = 1 - \prod_{j=1}^{p}(1 - \lambda_i/\mu_j), \quad i = 1, 2, \ldots, m-1$ (see [9] for more details).

V. MODIFIED GENERALIZED LOADING METHOD

As we cited before, we use $|\boldsymbol{\alpha}_{\mathrm{SCB}i}|^2$ instead of $|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2$ and $\hat{\sigma}^2$ in (18) as an estimate of σ^2 in practice. Unfortunately, when N is low, $\hat{\sigma}^2$ is usually underestimated because of the overhead errors on sample covariance matrix, which leads to limited effect of the proposed method. Meanwhile, $|\boldsymbol{\alpha}_{\mathrm{SCB}i}|^2$ may suffer from robustness problem in case of small sample sizes. In such a case, the condition number of $\boldsymbol{R} + \boldsymbol{\Delta}_1$ may not be evidently reduced compared to $\boldsymbol{R}$ since $\boldsymbol{\Delta}_1$ may be ill-conditioned. Note that the conventional diagonal loading algorithm has no such problem. Thus, we formulate the modified generalized loading method as

$$\breve{w} = \frac{(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}} + \tau\boldsymbol{I})^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}} + \tau\boldsymbol{I})^{-1}\boldsymbol{a}} = \frac{(\hat{\boldsymbol{R}} + \tau\boldsymbol{I})^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\hat{\boldsymbol{R}} + \tau\boldsymbol{I})^{-1}\boldsymbol{a}}. \quad (26)$$

Similar to the GLC, we can use $\hat{\boldsymbol{R}}$ instead of sample covariance matrix and the results in [11] to obtain the minimized MSE estimate of covariance matrix as $\hat{\rho}_1\hat{\boldsymbol{R}} + \hat{\rho}_2\boldsymbol{I}$ and choose $\tau = \hat{\rho}_2/\hat{\rho}_1$. From the results in [11], we know

$$\hat{\rho}_2 = \hat{\nu}\frac{\hat{\rho}}{\|\hat{\boldsymbol{R}} - \hat{\nu}\boldsymbol{I}\|^2}, \hat{\rho}_1 = 1 - \frac{\hat{\rho}_2}{\hat{\nu}} \quad (27)$$

where $\hat{\nu} = \mathrm{tr}(\hat{\boldsymbol{R}})/m$ and $\hat{\rho} = \sum_{n=1}^{N}\|\boldsymbol{y}(n)\|^4/N^2 - \|\hat{\boldsymbol{R}}\|^2/N$. Note that, when we use $\hat{\boldsymbol{R}}$ instead of sample covariance matrix, we should ensure that $\hat{\rho}$ is always no less than 0. However, $\hat{\rho} < 0$ happens when the elements of loaded matrix $\boldsymbol{\Delta}_1$ are high. To cope with this, we redefine $\hat{\rho}$ as

$$\hat{\rho} = \max\left\{\frac{1}{N^2}\sum_{n=1}^{N}\|\boldsymbol{y}(n)\|^4 - \frac{1}{N}\|\hat{\boldsymbol{R}}\|^2, 0\right\}. \quad (28)$$

From (28) and (27), we note that when N is high, $\hat{\rho}$ approaches to 0 so that τ approaches to 0. Thus, the contribution of $\tau\boldsymbol{I}$ to the loaded matrix is neglectful. If $\hat{\rho} = 0$, the modified method is the same as the generalized loading method of (24). Subsequently, we get

$$\breve{w} = \frac{(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}} + \tau\boldsymbol{I})^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R} + \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{Q}^{\mathrm{H}} + \tau\boldsymbol{I})^{-1}\boldsymbol{a}} \triangleq \frac{(\boldsymbol{R} + \boldsymbol{\Delta}_2)^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{R} + \boldsymbol{\Delta}_2)^{-1}\boldsymbol{a}}. \quad (29)$$

Moreover, we note that the modified generalized loading method can also be seen as a special case of (12) by choosing shrinkage factors as

$$\xi_i = \frac{\lambda_i}{\lambda_i + \frac{\sigma^2}{|\mathrm{E}(\boldsymbol{\alpha}_{\mathrm{SCB}i})|^2} + \tau}, \quad i = 1, 2, \ldots, m-1. \quad (30)$$

VI. NUMERICAL EXAMPLES

In this section, we evaluate our approaches using Monte Carlo simulation. In all examples, we consider a half wavelength spacing ULA with $m = 10$ omnidirectional sensors. The assumed steering vectors are perturbed by white Gaussian noise such that the unknown steering vectors are $\hat{\boldsymbol{a}} = \boldsymbol{a} + \boldsymbol{\delta}$, where $\boldsymbol{\delta} \sim \mathcal{CN}(0, \gamma^2\boldsymbol{I})$. There are three temporally white complex Gaussian farfield signals impinging from the directions $\{-30^\circ, 0^\circ, 10^\circ\}$ with the powers $\{1, 1, s^2\}$. We consider the first and second sources as interferences and the third as our source of interest. The noise is spatially and temporally white and it has a complex Gaussian zero-mean distribution with variance 0.02. The methods we evaluate in all examples include 1) the SCB, 2) the GLC in [10], 3) the HKB in [7] , 4) the method in [8] based on PCR, 5) the method in [9] based on PLS, 6) the proposed method of (24) which is denoted by GL_1, and 7) the proposed method of (29) which is denoted by

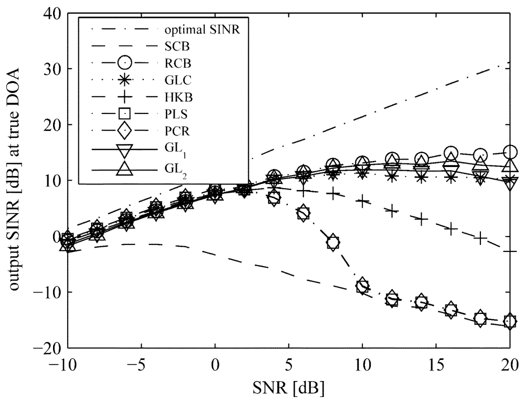

Fig. 1. Mean output SINRs for varying input SNR.

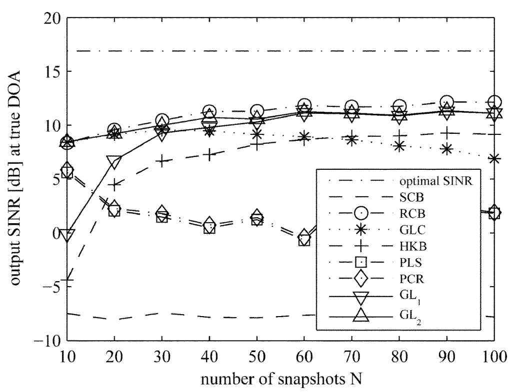

Fig. 2. Mean output SINRs for varying N.

GL_2. We also choose one of the robust methods that need user parameters, 8) the robust Capon beamformer (RCB), in order to compare the performance with that of parameter-free methods. We assume the knowledge of errors on array steering vector (covariance matrix $\gamma^2\boldsymbol{I}$ of $\boldsymbol{\delta}$) is known, and set the parameter ϵ (see [1] for details) of RCB such that $\mathrm{Prob}(\epsilon < \|\boldsymbol{\delta}\|^2) = 0.98$. In all the examples, we set $\gamma^2 = 0.2$ (about -7 dB). $M = 100$ Monte Carlo trials are performed.

First, we set $N = 50$ and vary s^2 such that the input SNR goes from -10 dB to 20 dB. The mean SINRs can be seen in Fig. 1. According to the plots, we note that when SNR is high, PLS and PCR fail completely, and HKB degrades severely. The RCB, GL_1, GL_2, and GLC give good SINR performance. In parameter-free methods, the proposed GL_2 seems to perform best. Next, we set $s^2 = 0.08$ (an input SNR of about 6 dB) and vary N from 10 to 100. The mean SINRs are shown in Fig. 2. The simulation results show that PLS and PCR fail completely, GLC performs well in the case of small sample sizes but degrades as N increases, and HKB is not robust enough to errors on sample covariance matrix when N is low. However, both of the proposed methods give good SINR performance. In particular, we note that the GL_2 is more robust to errors than other parameter-free methods. It gives a similar performance to the RCB. Finally, we consider the SOI power estimates for varying N. Set $s^2 = 0.08$ (about -11 dB) and vary N from 10 to 100, and we get the mean SOI power estimates shown in Fig. 3. From the figure, we can see that our proposed methods, GL_1 and GL_2, outperform other tested methods, even including the RCB which gives an overestimated SOI power. Particularly, we can see that the GL_2 seems to perform best of all.

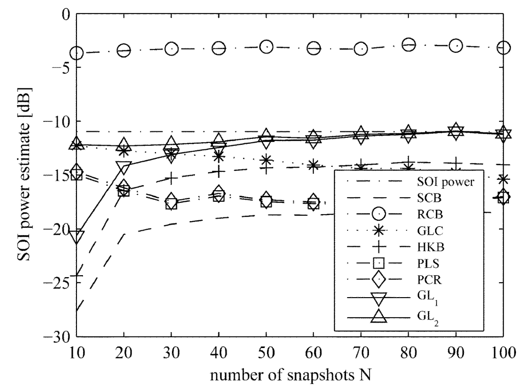

Fig. 3. Mean SOI power estimates for varying N.

VII. Conclusion

In this letter, we have proposed two robust adaptive beamformers based on generalized loading approach. Both methods are parameter-free, which means the loaded matrices are computed automatically without choice of user parameters. Numerical examples in terms of SINR and SOI power estimate show that the proposed methods are robust to errors on array steering vector and sample covariance matrix. However, we should note that the minimized MSE estimate is well known to be biased, which is different from the LS estimate (best linear unbiased estimate). Thus, the proposed beamformers based on the minimized MSE estimator cannot have good performance in all scenarios. For instance, when the power of signal is much lower than the noise level, SOI power may be overestimated by the proposed beamformers. That is to say, when used for energy detection, the proposed methods may perform poorly. Nevertheless, improving the performance of an energy detector is not the objective of a beamformer.

Acknowledgment

The authors would like to thank Dr. M. Zhou (Elsevier Beijing Office, China) for her valuable suggestions.

References

[1] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

[2] R. Lorenz and S. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

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

[4] Y. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 168–181, Jan. 2005.

[5] P. Stoica, Z. Wang, and J. Li, "Robust Capon beamforming," *IEEE Signal Process. Lett.*, vol. 10, no. 6, pp. 172–175, Jun. 2003.

[6] 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. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[7] Y. Selén, R. Abrahamsson, and P. Stoica, "Automatic robust adaptive beamforming via ridge regression," *Signal Process.*, vol. 88, no. 1, pp. 33–49, 2008.

[8] J. Yang, X. Ma, C. Hou, Y. Liu, and W. Li, "Fully automatic robust adaptive beamforming via principal component regression," in *Proc. Int. Conf. Signal Processing*, Oct. 2008, pp. 358–361.

[9] J. Yang, X. Ma, C. Hou, Y. Liu, and W. Li, "Robust adaptive beamforming using partial least squares," in *Proc. Int. Conf. Signal Processing*, Oct. 2008, pp. 362–365.

[10] J. Li, L. Du, and P. Stoica, "Fully automatic computation of diagonal loading levels for robust adaptive beamforming," in *Proc. Int. Conf. Acoustics, Speech, Signal Process.*, Apr. 2008, pp. 2325–2328.

[11] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2598–2602, Jun. 2008.