ROBUST MINIMUM VARIANCE BEAMFORMING UNDER DISTRIBUTIONAL UNCERTAINTY

Xiao Zhang, Yang Li, Ning Ge, and Jianhua Lu

Tsinghua National Laboratory for Information Science and Technology (TNList)
Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China
Email: {z-xiao11,liy-11}@mails.tsinghua.edu.cn, {gening,lhh-dee}@tsinghua.edu.cn

ABSTRACT

This paper investigates distributionally robust minimum variance beamforming under first-order moment uncertainty. In contrast to deterministic modeling of the array response, our approach employs a distributional set to describe the uncertainty. The distributional set we introduce consists of two constraints: the probability measure constraint and a first-order moment constraint. The weights are selected to minimize the combined output power, subject to the modified distortionless response constraint that the expected real part of the array gain exceeds unity for all distributions in the uncertainty set. We begin our discussion by revealing the intrinsic connection between the distributionally robust minimum variance beamformers (DRMVB) and the robust minimum variance beamformer (RMVB). Then for the sample space described by a union of ellipsoids, the DRMVB is reformulated as the optimal solution of a semidefinite program (SDP). Finally, we demonstrate the performance of the DRMVB via several numerical examples.

***Index Terms*—** Minimum variance beamforming, distributionally robust optimization, strong duality, semidefinite programming

1. INTRODUCTION

The minimum variance beamformer (MVB) [1] is an optimal linear processor that maximizes the output signal to interference-plus-noise ratio (SINR), provided that the statistical covariance matrix and the array response are known. However, the statistical covariance matrix is rarely known in practice, and the use of the sample covariance matrix instead is known to degrade the performance dramatically, especially in the case in which the array response is imperfect as well [2]. One simple approach to improve the robustness of the MVB is the diagonal loaded (DL) beamformer [3], [4]. Unfortunately, there are two limitations of the DL approach: first, it is unclear how to select the loading factor properly; second, it ignores the knowledge one may have on the uncertainty. To overcome such limitations, several RMVBs have been derived in [5]–[8].

Inspired by the work [9], this paper investigates distributionally robust beamforming under first-order moment uncertainty. The DRMVB is derived for the sample space described by a union of ellipsoids. Simulations show that the DRMVB under first-order moment uncertainty provides better average output SINR and power estimate than the RMVB [5].

2. BACKGROUND

Consider an antenna array of N elements. Let $\mathbf{a}(\theta) \in \mathbb{C}^N$ be the array response to a plane wave of unit amplitude from direction θ. Assuming that a desired source $s(t)$ is impinging on the array from θ, we can write the complex array observation as

$$\mathbf{y}(t) = s(t)\mathbf{a}(\theta) + \mathbf{e}(t), \tag{1}$$

where $\mathbf{e}(t)$ is an additive term that captures the effect of both interference and noise. Let $\mathbf{y}(k)$ be the sampled array output; we can express the combined output as

$$y_c(k) = \mathbf{w}^*\mathbf{y}(k) = s(k)\mathbf{w}^*\mathbf{a}(\theta) + \mathbf{w}^*\mathbf{e}(k) \qquad k = 1, \ldots, K, \tag{2}$$

where $\mathbf{w} \in \mathbb{C}^N$ is the array weight vector and K is the number of snapshots. If $\mathbf{a}(\theta)$ and $\mathbf{R}_y \triangleq E\{\mathbf{y}(k)\mathbf{y}(k)^*\}$ are known, the beamformer

$$\mathbf{w}_{\text{opt}} = \frac{\mathbf{R}_y^{-1}\mathbf{a}(\theta)}{\mathbf{a}(\theta)^*\mathbf{R}_y^{-1}\mathbf{a}(\theta)}. \tag{3}$$

is the optimal linear combinator that maximizes the output SINR. In practice, $\mathbf{R}_y$ is rarely known, and is instead replaced by $\hat{\mathbf{R}}_y = \frac{1}{K}\sum_{k=1}^{K}\mathbf{y}(k)\mathbf{y}(k)^*$, and the solution to (3) using $\hat{\mathbf{R}}_y$ is referred to as the MVB, and is given by

$$\mathbf{w}_{\text{MV}} = \frac{\hat{\mathbf{R}}_y^{-1}\mathbf{a}(\theta)}{\mathbf{a}(\theta)^*\hat{\mathbf{R}}_y^{-1}\mathbf{a}(\theta)}. \tag{4}$$

This work is supported by the National Basic Research Project of China (973 program) under grant number 2013CB329001 and the National Natural Science Foundation of China (NSFC) under grant numbers 61132002 and 61321061.

Unfortunately, the use of $\hat{\mathbf{R}}_y$ instead of $\mathbf{R}_y$ in (3) is known to dramatically degrade the performance, especially in the case when the knowledge of $\mathbf{a}(\theta)$ is imperfect as well [2]. A simple approach to improve its robustness is the DL beamformer [3]

$$\mathbf{w}_{\mathrm{DL}} = \frac{(\hat{\mathbf{R}}_y + \mu\mathbf{I})^{-1}\mathbf{a}(\theta)}{\mathbf{a}(\theta)^*(\hat{\mathbf{R}}_y + \mu\mathbf{I})^{-1}\mathbf{a}(\theta)}. \tag{5}$$

which amounts to replacing $\hat{\mathbf{R}}_y$ in (4) by $\hat{\mathbf{R}}_y + \mu\mathbf{I}$, where μ is the loading factor. The DL beamformer (5) is known to suffer from two limitations. First, it is unclear how to choose μ properly. Second, the approach ignores the knowledge one may have on the uncertainty. To avoid the aforementioned drawbacks of the DL beamformer (5), more theoretically rigorous RMVBs have been derived in [5]–[7].

The essence of the RMVB derived in [5] is to model the uncertainty via a $2N$-dimensional real ellipsoid, and to design the beamformer that minimizes the combined output power subject to the constraint that the real part of the array gain exceeds unity for all array responses in the ellipsoid, *i.e.*,

$$\begin{aligned} &\min_{\mathbf{w}} \mathbf{w}^*\hat{\mathbf{R}}_y\mathbf{w} \\ &s.t. \min_{\mathbf{a}\in\mathcal{E}(\mathbf{c},\mathbf{P})} \mathbf{Re}\,\mathbf{w}^*\mathbf{a}(\theta) \geq 1, \end{aligned} \tag{6}$$

where $\mathcal{E}(\mathbf{c},\mathbf{P}) = \{\mathbf{x} \,|\, (\mathbf{x}-\mathbf{c})^T\mathbf{P}^{-1}(\mathbf{x}-\mathbf{c}) \leq 1\}$. Let $\mathbf{P} = \mathbf{A}\mathbf{A}^T$, where $\mathbf{A} \in \mathbb{R}^{2N\times 2N}$; then we can express $\mathcal{E}(\mathbf{c},\mathbf{P})$ equivalently as $\mathcal{E} = \{\mathbf{A}\mathbf{u}+\mathbf{c} \,|\, \|\mathbf{u}\|_2 \leq 1\}$. Introducing

$$\mathbf{x} = \begin{bmatrix}\mathbf{Re}\,\mathbf{w}\\ \mathbf{Im}\,\mathbf{w}\end{bmatrix} \quad \hat{\mathbf{R}} = \begin{bmatrix}\mathbf{Re}\,\hat{\mathbf{R}}_y & -\mathbf{Im}\,\hat{\mathbf{R}}_y\\ \mathbf{Im}\,\hat{\mathbf{R}}_y & \mathbf{Re}\,\hat{\mathbf{R}}_y\end{bmatrix}$$

we can rewrite (6) into the real-valued form:

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \min_{\mathbf{a}\in\mathcal{E}} \mathbf{x}^T\mathbf{a} \geq 1. \end{aligned} \tag{7}$$

Applying the Cauchy-Schwartz inequality, we can simplify (7) as

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \|\mathbf{A}^T\mathbf{x}\|_2 \leq \mathbf{c}^T\mathbf{x} - 1. \end{aligned} \tag{8}$$

As mentioned in [5], the constraint in (8) is always tight for the optimal weights, and therefore (8) is equivalent to

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \|\mathbf{A}^T\mathbf{x}\|_2 = \mathbf{c}^T\mathbf{x} - 1, \end{aligned} \tag{9}$$

which can be solved efficiently by Lagrange multiplier methods [5].

3. ROBUST MINIMUM VARIANCE BEAMFORMING UNDER DISTRIBUTIONAL UNCERTAINTY

3.1. Robust Beamforming Problem

Assuming that we have collected some data $\{\mathbf{a}_i(\theta)\}_{i=1}^m$ from measuring the array response, we consider the following beamforming problem

$$\begin{aligned} &\min_{\mathbf{w}} \mathbf{w}^*\hat{\mathbf{R}}_y\mathbf{w}, \\ &s.t. \min_{P_{\mathbf{a}}\in\mathcal{D}(\mathcal{S},\hat{\mathbf{a}},\hat{\boldsymbol{\Sigma}},\gamma)} E_{\mathbf{a}}\{\mathbf{Re}\,\mathbf{w}^*\mathbf{a}(\theta)\} \geq 1. \end{aligned} \tag{10}$$

with $\mathcal{D}(\mathcal{S},\hat{\mathbf{a}},\hat{\boldsymbol{\Sigma}},\gamma) =$

$$\left\{ P_{\mathbf{a}} \in \mathcal{M}^+ \,\middle|\, \begin{array}{l} E\{1_{\mathcal{S}}(\mathbf{a})\} = 1 \\ (E\{\mathbf{a}\}-\hat{\mathbf{a}})^T\hat{\boldsymbol{\Sigma}}^{-1}(E\{\mathbf{a}\}-\hat{\mathbf{a}}) \leq \gamma \end{array} \right\},$$

where $\mathbf{a}_i = [\mathbf{Re}\,\mathbf{a}_i(\theta)^T, \mathbf{Im}\,\mathbf{a}_i(\theta)^T]^T, i = 1,\ldots,m$, $\hat{\mathbf{a}} = \frac{1}{m}\sum_{i=1}^m \mathbf{a}_i$, and $\hat{\boldsymbol{\Sigma}} = \frac{1}{m}\sum_{i=1}^m(\mathbf{a}_i-\hat{\mathbf{a}})(\mathbf{a}_i-\hat{\mathbf{a}})^T$; $\gamma > 0$ describes the size of $\mathcal{E}(\hat{\mathbf{a}},\gamma\hat{\boldsymbol{\Sigma}})$, and $\mathcal{M}^+$ denotes the set of all measures on $(\mathbb{R}^{2N},\mathfrak{B})$. The set $\mathfrak{B}$ represents the Borel σ-algebra on $\mathbb{R}^{2N}$. We first transform (10) into the real-valued form

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x}, \\ &s.t. \min_{P_{\mathbf{a}}\in\mathcal{D}(\mathcal{S},\hat{\mathbf{a}},\hat{\boldsymbol{\Sigma}},\gamma)} E\{\mathbf{x}^T\mathbf{a}\} \geq 1. \end{aligned} \tag{11}$$

Note that $\mathcal{D}(\mathcal{S},\hat{\mathbf{a}},\hat{\boldsymbol{\Sigma}},\gamma)$ is equivalent to

$$\left\{ P_{\mathbf{a}} \in \mathcal{M}^+ \,\middle|\, \begin{array}{l} E\{1_{\mathcal{S}}(\mathbf{a})\} = 1 \\ E\{\mathbf{a}\} \in \mathcal{E}(\hat{\mathbf{a}},\gamma\hat{\boldsymbol{\Sigma}}) \cap \mathbf{conv}\{\mathcal{S}\} \end{array} \right\}. \tag{12}$$

Now we present the following proposition to reveal the intrinsic connection between the RMVB (7) and the DRMVB (11).

Proposition 1: The distributionally robust beamforming problem (11) is equivalent to the following robust beamforming problem

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x}, \\ &s.t. \min_{\mathbf{a}\in\mathcal{E}(\hat{\mathbf{a}},\gamma\hat{\boldsymbol{\Sigma}})\cap\mathbf{conv}\{\mathcal{S}\}} \mathbf{x}^T\mathbf{a} \geq 1 \end{aligned} \tag{13}$$

Proof. To establish the (11) and (13), we only need to establish the equivalence of

$$\min_{P_{\mathbf{a}}\in\mathcal{D}(\mathcal{S},\hat{\mathbf{a}},\hat{\boldsymbol{\Sigma}},\gamma)} E\{\mathbf{x}^T\mathbf{a}\} \tag{14}$$

and

$$\min_{\mathbf{a}\in\mathcal{E}(\hat{\mathbf{a}},\gamma\hat{\boldsymbol{\Sigma}})\cap\mathbf{conv}\{\mathcal{S}\}} \mathbf{x}^T\mathbf{a}. \tag{15}$$

We first assume that $P_{\mathbf{a}} \in \mathcal{D}(\mathcal{S}, \hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}, \gamma)$, then $E_{P_{\mathbf{a}}}\{\mathbf{a}\}$ is feasible in (15), with the same objective function $E_{P_{\mathbf{a}}}\{\mathbf{x}^T\mathbf{a}\}$. Conversely, for any given $\mathbf{a} \in \mathcal{E}(\hat{\mathbf{a}}, \gamma\hat{\mathbf{\Sigma}}) \cap \mathbf{conv}\{\mathcal{S}\}$, a distribution $\sum_{i=1}^{L} \eta_i \delta_{\mathbf{s}_i} \in \mathcal{D}(\mathcal{S}, \hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}, \gamma)$ can be constructed with the same objective value $\mathbf{x}^T\mathbf{a}$, where $\mathbf{s}_i \in \mathcal{S}, \forall i$, $\sum_{i=1}^{L} \eta_i = 1$, and $\eta_i \geq 0, \forall i$. This completes our proof. □

In the light of Proposition 1, the feasibility of $\mathcal{D}(\mathcal{S}, \hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}, \gamma)$ can be easily established, i.e., $\mathbf{conv}\{\mathcal{S}\} \cap \mathcal{E}(\hat{\mathbf{a}}, \gamma\hat{\mathbf{\Sigma}}) \neq \emptyset$. We will focus on the cases where $\mathbf{conv}\{\mathcal{S}\} \cap \mathcal{E}(\hat{\mathbf{a}}, \gamma\hat{\mathbf{\Sigma}}) \neq \emptyset$ and $\mathcal{E}(\hat{\mathbf{a}}, \gamma\hat{\mathbf{\Sigma}}) \not\supseteq \mathbf{conv}\{\mathcal{S}\}$.

To tackle (11), we use duality theory [10] to reformulate the maximization problem in (11). Putting the maximization problem into the integral form

$$\max_{P_{\mathbf{a}}} \int_{\mathcal{S}} -\mathbf{x}^T\mathbf{a}\, dP_{\mathbf{a}} \tag{16a}$$

$$s.t. \int_{\mathcal{S}} dP_{\mathbf{a}} = 1 \tag{16b}$$

$$\int_{\mathcal{S}} \begin{bmatrix} \hat{\mathbf{\Sigma}} & (\mathbf{a}-\hat{\mathbf{a}}) \\ (\mathbf{a}-\hat{\mathbf{a}})^T & \gamma \end{bmatrix} dP_{\mathbf{a}} \succeq \mathbf{0} \tag{16c}$$

$$P_{\mathbf{a}} \in \mathcal{M}^+, \tag{16d}$$

we can derive its dual problem as follows:

$$\min_{\lambda, \mathbf{Z}, \mathbf{z}, \nu} \lambda + tr\{\mathbf{Z}\hat{\mathbf{\Sigma}}\} - 2\hat{\mathbf{a}}^T\mathbf{z} + \gamma\nu, \tag{17a}$$

$$s.t. -\lambda + 2\mathbf{a}^T\mathbf{z} - \mathbf{x}^T\mathbf{a} \leq 0, \quad \forall \mathbf{a} \in \mathcal{S}, \tag{17b}$$

$$\begin{bmatrix} \mathbf{Z} & \mathbf{z} \\ \mathbf{z}^T & \nu \end{bmatrix} \succeq \mathbf{0}, \tag{17c}$$

where $\lambda \in \mathbb{R}$ is the dual variable for (16b); $\mathbf{Z} \in \mathbb{R}^{2N\times 2N}$, $\mathbf{z} \in \mathbb{R}^{2N}$, and $\nu \in \mathbb{R}$ together are the dual variables for (16c). Note that (17) can be further simplified by solving analytically for $(\mathbf{Z}, \nu)$, while keeping $(\mathbf{z}, \lambda)$ fixed. We consider two cases of $\mathbf{z}$: either $\mathbf{z} = \mathbf{0}$ or $\mathbf{z} \neq \mathbf{0}$. We first assume that $\mathbf{z} = \mathbf{0}$, then $\nu^* = 0$ and $\mathbf{Z}^* = \mathbf{0}$ minimize the dual objective function (17a). If $\mathbf{z} \neq \mathbf{0}$, it must be true that $\nu > 0$, otherwise, we can construct a vector $\left[\mathbf{z}^T, g\right]^T$ such that

$$\left[\mathbf{z}^T, -g\right] \begin{bmatrix} \mathbf{Z} & \mathbf{z} \\ \mathbf{z}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ -g \end{bmatrix} = \mathbf{z}^T\mathbf{Z}\mathbf{z} - 2g\mathbf{z}^T\mathbf{z} < 0,$$

for $g > \frac{\mathbf{z}^T\mathbf{Z}\mathbf{z}}{2\mathbf{z}^T\mathbf{z}}$, which contradicts (17c). Therefore, we have $\nu > 0$. Applying Schur's complement, we can rewrite the constraint (17c) as $\mathbf{Z} \succeq \frac{\mathbf{z}\mathbf{z}^T}{\nu}$. Obviously, the dual objective function (17a) is minimized with

$$\mathbf{Z}^* = \frac{\mathbf{z}\mathbf{z}^T}{\nu}. \tag{18}$$

Substituting (18) into the dual objective function (17a), we obtain

$$\min_{\nu > 0} \frac{1}{\nu}\mathbf{z}^T\hat{\mathbf{\Sigma}}\mathbf{z} + \gamma\nu. \tag{19}$$

Since (19) is convex in ν, by setting the first-order derivative of (19) to zero, we immediately obtain

$$\nu^* = \sqrt{\frac{1}{\gamma}\mathbf{z}^T\hat{\mathbf{\Sigma}}\mathbf{z}}. \tag{20}$$

Combining both cases of $\mathbf{z}$, we obtain the optimal $\mathbf{Z}$ and ν given by (18) and (20). Then (17) can be simplified as

$$\begin{aligned} &\min_{\lambda, \mathbf{z}} \lambda + 2\sqrt{\gamma}\|\hat{\mathbf{\Sigma}}^{\frac{1}{2}}\mathbf{z}\|_2 - 2\hat{\mathbf{a}}^T\mathbf{z} \\ &s.t. -\lambda + 2\mathbf{a}^T\mathbf{z} - \mathbf{x}^T\mathbf{a} \leq 0, \quad \forall \mathbf{a} \in \mathcal{S}. \end{aligned} \tag{21}$$

To establish strong duality for the primal-dual pair (16) and (21), we have the following proposition.

Proposition 2: For $\gamma > 0$, strong duality holds for the primal-dual pair (16) and (21).

Proof. Since $\hat{\mathbf{a}} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{a}_i$, where $\mathbf{a}_i \in \mathcal{S}, \forall i$, a probability distribution

$$P_{\mathbf{a}} = \frac{1}{m}\sum_{i=1}^{m} \delta_{\mathbf{a}_i} \tag{22}$$

can be constructed to lie in the relative interior of the set $\mathcal{D}(\mathcal{S}, \hat{\mathbf{a}}, \hat{\mathbf{\Sigma}}, \gamma)$. According to the weaker version of Proposition 3.4 in [10], strong duality holds for the primal-dual pair, and that if the optimal value of problem (16) is finite, then the set of optimal solutions of problem (16) is nonempty [9]. □

3.2. Tractable Beamformer Design

Since strong duality has been established, replacing the maximization problem in (11) by (21), we can reformulate (11) as

$$\min_{\mathbf{x}, \lambda, \mathbf{z}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \tag{23a}$$

$$s.t. \lambda + 2\sqrt{\gamma}\|\hat{\mathbf{\Sigma}}^{\frac{1}{2}}\mathbf{z}\|_2 - 2\hat{\mathbf{a}}^T\mathbf{z} \leq -1 \tag{23b}$$

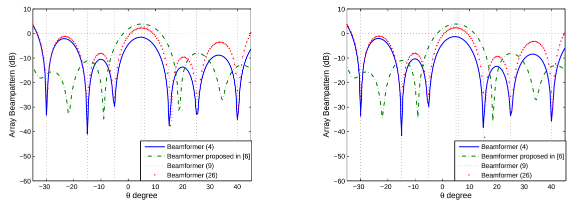
$$-\lambda + 2\mathbf{a}^T\mathbf{z} - \mathbf{x}^T\mathbf{a} \leq 0, \quad \forall \mathbf{a} \in \mathcal{S}, \tag{23c}$$

which is a convex programming problem. Nevertheless, problem (23) becomes tractable only if we choose $\mathcal{S}$ properly. In the following discussion, we will derive the DRMVBs for $\mathcal{S} = \cup_{i=1}^{n}\mathcal{E}_i$. Suppose that the ellipsoids $\mathcal{E}_1, \ldots, \mathcal{E}_n$ are described by the following n convex quadratic inequalities:

$$\mathbf{a}^T\mathbf{Q}_i\mathbf{a} + 2\mathbf{q}_i^T\mathbf{a} + r_i \leq 0, \quad i = 1, \ldots, n. \tag{24}$$

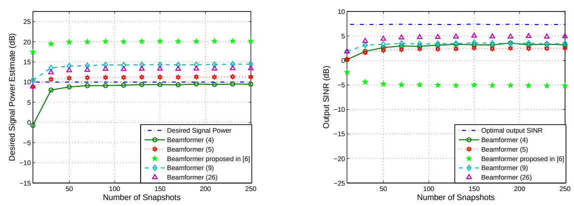
If (24) is strictly feasible, then by S-lemma [11], the condition that $-\lambda + 2\mathbf{a}^T\mathbf{z} - \mathbf{x}^T\mathbf{a} \leq 0$, $\forall \mathbf{a} \in \cup_{i=1}^{n}\mathcal{E}_i$ can be equivalently expressed as: there exist $\tau_1, \ldots, \tau_n$ such that

$$\begin{bmatrix} 0 & (\frac{1}{2}\mathbf{x} - \mathbf{z}) \\ (\frac{1}{2}\mathbf{x} - \mathbf{z})^T & \lambda \end{bmatrix} + \tau_i \begin{bmatrix} \mathbf{Q}_i & \mathbf{q}_i \\ \mathbf{q}_i^T & r_i \end{bmatrix} \succeq \mathbf{0}, \quad i = 1, \ldots, n. \tag{25}$$



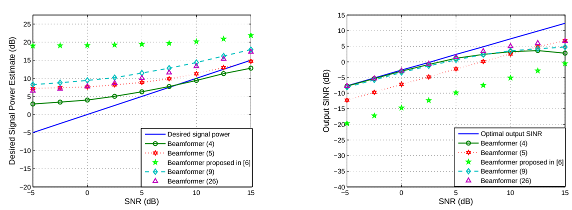
(a) (b)

Fig. 1. Comparison of Beampatterns with $K = 250$, SNR=10 dB, $\alpha = 5°$, $\beta = 6°$, $p = 0.8$, and $\sigma_{\mathrm{p}}^2 = (0.015)^2$: a) $\theta_1 = \alpha$, b) $\theta_1 = \beta$.



(a) (b)

Fig. 2. Performance comparison with SNR =10 dB: (a) desired signal power estimate, (b) SINR versus K.



(a) (b)

Fig. 3. Performance comparison with $K = 250$: (a) desired signal power estimate, (b) SINR versus SNR.

Then problem (23) can be reformulated as an SDP:

$$
\begin{aligned}
\min_{t,\mathbf{x},\lambda,\mathbf{z},\boldsymbol{\tau}} \quad & t \\
s.t. \quad & \|\hat{\mathbf{R}}^{\frac{1}{2}}\mathbf{x}\|_2 \leq t \\
& \lambda + 2\sqrt{\gamma}\|\hat{\boldsymbol{\Sigma}}^{\frac{1}{2}}\mathbf{z}\|_2 - 2\hat{\mathbf{a}}^T\mathbf{z} \leq -1 \\
& \tau_1 \geq 0, \ldots, \tau_n \geq 0 \\
& \begin{bmatrix} \mathbf{0} & (\frac{1}{2}\mathbf{x} - \mathbf{z}) \\ (\frac{1}{2}\mathbf{x} - \mathbf{z})^T & \lambda \end{bmatrix} + \tau_i \begin{bmatrix} \mathbf{Q}_i & \mathbf{q}_i \\ \mathbf{q}_i^T & r_i \end{bmatrix} \succeq \mathbf{0} \\
& \qquad i = 1, \ldots, n. \qquad (26)
\end{aligned}
$$

4. NUMERICAL RESULTS

We present three numerical examples comparing the performance of the MVB, the DL beamformer (5), the RMVB [6], the RMVB (9), and the DRMVB (26). In all examples below, we assume a ten-sensor uniform linear array centered at the origin and spaced 0.4-wavelength apart. Two sources of errors, namely, the sensor position error and the AOA error are considered. To simulate the sensor position error, the position of each element is perturbed independently by a zero-mean Gaussian random vector with variance $\sigma_{\mathrm{p}}^2\mathbf{I}_{2\times 2}$. The AOA of the desired source is modeled as a binary random variable θ_1 with probability distribution $\Pr\{\theta_1 = \alpha\} = p$ and $\Pr\{\theta_1 = \beta\} = 1 - p$, where $p > 0$. The measurements are generated in the following manner: first, $m_{\mathrm{AOA}} = 100$ independent observations of θ_1 are collected, and then for each realization of θ_1, $m_{\mathrm{p}} = 1000$ samples are independently generated. Since θ_1 is binary, three ellipsoids $\mathcal{E}_s(\hat{\mathbf{a}}_s, \gamma_s\hat{\boldsymbol{\Sigma}}_s)$, $\mathcal{E}_1(\hat{\mathbf{a}}_1, \gamma_1\hat{\boldsymbol{\Sigma}}_1)$, and $\mathcal{E}_2(\hat{\mathbf{a}}_2, \gamma_2\hat{\boldsymbol{\Sigma}}_2)$ can be constructed to cover all the measurement and the two groups of measurements. After the ellipsoid is constructed, the uncertainty sets for the RMVB [5] and the sample spaces for the DRMVB can now be specified: the uncertainty set and the sample spaces for (9) and (26) are selected to be $\mathcal{E}_s(\hat{\mathbf{a}}_s, \gamma_s\hat{\boldsymbol{\Sigma}}_s)$ and $\mathcal{E}_1(\hat{\mathbf{a}}_1, \gamma_1\hat{\boldsymbol{\Sigma}}_1) \cup \mathcal{E}_2(\hat{\mathbf{a}}_2, \gamma_2\hat{\boldsymbol{\Sigma}}_2)$ respectively. We set $\gamma = 0.49\gamma_s$. Regarding the type-II R-MVB, the complex sphere is constructed with center $\hat{\mathbf{a}}(\theta) = \sum_{i=1}^m \mathbf{a}_i(\theta)$ and radius $\delta_{\min} = \sup_i \|\mathbf{a}_i(\theta) - \hat{\mathbf{a}}(\theta)\|_2$. The diagonal loading coefficient μ for (5) is chosen empirically to be $10\sigma_0^2$. All the convex programs are solved by CVX [12].

Fig. 1 shows the beampatterns of the beamformers, which are obtained from one Monte Carlo realization of the waveforms and the array response. In this example, six interferers with powers $\sigma_i^2 = 20$ dB, $i = 2, \ldots, 7$ impinge from $\theta_2 = -30°$, $\theta_3 = -15°$, $\theta_4 = -5°$, $\theta_5 = 15°$, $\theta_6 = 25°$, and $\theta_7 = 40°$. We can see from Fig. 1 that the DRMVB generates the beampattern with smaller main lobe than the R-MVB (9), providing better interference-plus-noise rejection capability than the RMVB (9). Fig. 2 and Fig. 3 respectively show the desired signal power estimates and the output SINRs as functions of the snapshot numbers and the SNRs. The setup here is similar to that in example 1, except that we vary K or SNR to evaluate the power estimates and the output SINRs. Each point in Fig. 2 – Fig. 9 are obtained from 200 Monte Carlo simulations. From Fig. 2 and 3, we note that by implementing a first-order moment constraint, the DRMVB can yield better power estimates and can provide sufficient robustness against the model error than the RMVB (9).

5. CONCLUSIONS

This paper studied distributionally robust beamforming under first-order moment uncertainty. The connection between the DRMVB and the RMVB are revealed. For the sample space described by a union of ellipsoids, the DRMVBs with two types of support uncertainty sets are derived. Simulations show that the DRMVB under first-order moment uncertainty provides superior performance over the RMVB in terms of average output SINR and power estimate.

6. REFERENCES

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

[2] A. B. Gershman, N. D. Sidropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, 2010.

[3] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 7, pp. 397–401, Jul. 1988.

[4] H. L. V. Trees, *Optimum Array Processing*. New York: Wiley, 2002.

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

[6] S. 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.

[7] 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.

[8] S. Vorobyov, H. Chen, and A. B. Gershman, "On the relationship between robust minimum variance beamformings with probablilistic and worse-case distortionless response constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5719–5724, Nov. 2008.

[9] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, vol. 58, no. 3, pp. 595–612, May-Jun, 2010.

[10] A. Shapiro, "On duality theory of conic linear problems," in *Semi-Infinite Programming: Recent Advances*, M. Goberna and A. López, Eds. Dordrecht, N.L.: Academic Publishers, 2001, pp. 135–165.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[12] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," ver. 2.0 beta. http://cvxr.com/cvx, Sep. 2013.