DISTRIBUTIONALLY ROBUST CHANCE-CONSTRAINED MINIMUM VARIANCE BEAMFORMING

Xiao Zhang[⋆], *Qiang Feng*[†], *Ning Ge*[⋆], *and Jianhua Lu*[⋆†]

Tsinghua National Laboratory for Information Science and Technology (TNList)
[⋆]Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China
[†]Tsinghua Space Center, Tsinghua University, Beijing, 100084, China
Emails: z-xiao11@mails.tsinghua.edu.cn, {fengq802, gening, lhh-dee}@tsinghua.edu.cn

ABSTRACT

This paper studies distributionally robust chance-constrained minimum variance beamforming. In contrast to deterministic modeling of the steering vector, our approach models the uncertainty statistically via distributions. We select the weights that minimize the combined output power subject to the distributionally robust chance constraint that for all distributions in the uncertainty set, the gain should exceed unity with high probability. Our discussion begins with the simplest case where the distributional set contains only Gaussian distribution; then we derive the robust weights for three distributional sets, namely, the set of (central symmetric) distributions with known mean and covariance, and a distributional model where the mean is known, the components are independent and belong to some known bounded intervals. It can be seen that these four robust beamformers provide statistical interpretation for the deterministic counterpart. Finally, we demonstrate the performance of these robust beamformers via several numerical examples.

Index Terms— Minimum variance beamforming, chance constraint, distributionally robust optimization

1. INTRODUCTION

Minimum variance beamforming (MVB) [1] is a powerful technique, which finds applications in many areas [4], [5]. Many approaches have been proposed to improve its robustness against small sample size and model errors. Traditional approaches towards this end include [2]–[4]. More recent robust approaches include [6]–[10]. Generally, there are two approaches to designing these robust minimum variance beamformers (RMVBs). The first models the underlying array steering vector as a deterministic vector lying in some bounded sets, which results in the robust designs [6]–[8]. The second models the uncertainty statistically via distributions: the authors in [9] derived the robust beamformers for two types of distributional sets, namely, Gaussian distribution and a set of distributions with known mean and covariance, providing statistical interpretation for [7]; in [10], the authors developed the distributionally RMVB under first-order moment uncertainty, extending the uncertainty set from an ellipsoid [6] to a more general one. As noted, the chance-constrained version of [7] was studied in [9], whereas the chance-constrained version of [6] has not been investigated yet.

Inspired by the results on distributionally robust optimization [12]–[15], this paper studies the distributionally robust chance-constrained minimum variance beamforming, specifically, the chance-constrained version of [6]. Beginning with the simplest case where the distributional set contains only Gaussian distribution, we derive the robust beamformers for three distributional sets, providing statistical interpretation for [6]. Finally, we compare the performance of these robust beamformers via several numerical examples.

2. BACKGROUND

Consider a generic array of N sensors from where K snapshots are obtained. Let $\mathbf{y}(k) \in \mathbb{C}^N$ be the snapshot obtained at the kth sample instant ($k = 1, \ldots, K$). Each of these snapshots can be written as

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

where $\mathbf{a}(\theta) \in \mathbb{C}^N$ denotes the steering vector of the desired signal $s(k)$ impinging from direction θ, and $\mathbf{e}(k) \in \mathbb{C}^N$ models the effect of both interference and noise. Let $\mathbf{w}$ be the weight vector of the beamformer; then its combined output can be expressed as follows:

$$y_c(k) = \mathbf{w}^*\mathbf{y}(k) = \mathbf{w}^*\mathbf{a}(\theta)s(k) + \mathbf{w}^*\mathbf{e}(k). \tag{2}$$

If $\mathbf{a}(\theta)$ and $\mathbf{R}_y \triangleq E\{\mathbf{y}(k)\mathbf{y}(k)^*\}$ are known, the beamformer

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

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.

is the optimal linear combiner that maximizes the output SINR. However, in practice, $\mathbf{R}_y$ is rarely known, and is replaced by $\hat{\mathbf{R}}_y = \frac{1}{K}\sum_{k=1}^{K}\mathbf{y}(k)\mathbf{y}(k)^*$. The beamformer using $\hat{\mathbf{R}}_y$ is referred to as the MVB [1], and is given by

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

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

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

Unfortunately, the major difficulty in implementing (5) is selection of the parameter μ. To alleviate this drawback, more theoretically rigorous RMVBs are derived respectively in [6]–[8]. The key idea of the RMVB derived in [6] is to model the uncertainty via a $2N$-dimensional real ellipsoid $\mathcal{E}(\mathbf{c},\mathbf{P}) = \{\mathbf{x} \,|\, (\mathbf{x}-\mathbf{c})^T\mathbf{P}^{-1}(\mathbf{x}-\mathbf{c}) \le 1\}$ where $\mathbf{P} \succeq \mathbf{0}$, and to design the weight vector as the solution of

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

Let $\mathbf{P} = (\mathbf{P}^{\frac{1}{2}})^2$, where $(\mathbf{P}^{\frac{1}{2}})^T = \mathbf{P}^{\frac{1}{2}}$; then $\mathcal{E}(\mathbf{c},\mathbf{P})$ can be reparameterized as $\{\mathbf{P}^{\frac{1}{2}}\mathbf{u}+\mathbf{c} \,|\, \|\mathbf{u}\|_2 \le 1\}$. By introducing

$$\mathbf{a} = \begin{bmatrix}\mathbf{Re}\,\mathbf{a}(\theta)\\ \mathbf{Im}\,\mathbf{a}(\theta)\end{bmatrix} \; \mathbf{x} = \begin{bmatrix}\mathbf{Re}\,\mathbf{w}\\ \mathbf{Im}\,\mathbf{w}\end{bmatrix} \; \hat{\mathbf{R}} = \begin{bmatrix}\mathbf{Re}\,\hat{\mathbf{R}}_y & -\mathbf{Im}\,\hat{\mathbf{R}}_y\\ \mathbf{Im}\,\hat{\mathbf{R}}_y & \mathbf{Re}\,\hat{\mathbf{R}}_y\end{bmatrix},$$

(6) can be rewritten into the real-valued form:

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \min_{\mathbf{a}\in\mathcal{E}(\mathbf{c},\mathbf{P})} \mathbf{x}^T\mathbf{a} \ge 1. \end{aligned} \tag{7}$$

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

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \|\mathbf{P}^{\frac{1}{2}}\mathbf{x}\|_2 = \mathbf{c}^T\mathbf{x} - 1, \end{aligned} \tag{8}$$

which, as shown in [6], can be solved in $\mathcal{O}\big((2N)^3\big)$ by Lagrange multiplier methods.

3. DISTRIBUTIONALLY ROBUST CHANCE-CONSTRAINED BEAMFORMING

In this section, we will model $\mathbf{a}$ statistically via distributions, and study distributionally robust chance-constrained beamforming for four distributional uncertainty sets.

3.1. Gaussian Distribution

We first consider the following beamforming problem

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \,\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \ge 1\} \ge 1-\epsilon, \end{aligned} \tag{9}$$

where $\mathbf{a}$ is Gaussian, i.e., $\mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{a}}, \mathbf{C}_{\mathbf{a}})$ and $1 > \epsilon > 0$.

Proposition 1: Suppose $\mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{a}}, \mathbf{C}_{\mathbf{a}})$. For $1 > \epsilon > 0$, the chance constraint

$$\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \ge 1\} \ge 1-\epsilon \tag{10}$$

holds if and only if

$$-\Phi^{-1}(\epsilon)\|\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}\mathbf{x}\|_2 \le \boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} - 1, \tag{11}$$

holds, where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-\frac{u^2}{2})\,du$.

Proof. Our proof is tailored from the proof of theorem 10.4.1 in [11]. Note that $\mathbf{x}^T\mathbf{a} \sim \mathcal{N}(\mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}, \mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x})$. For $\mathbf{x}$ with $\mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x} > 0$, then we have

$$\begin{aligned} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \ge 1\} =&1 - \mathbf{Pr}\{\frac{\mathbf{x}^T\mathbf{a} - \mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}}{\sqrt{\mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x}}} \le \frac{1-\mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}}{\sqrt{\mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x}}}\} \\ =&1 - \Phi(\frac{1-\mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}}{\sqrt{\mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x}}}) \ge 1-\epsilon. \end{aligned} \tag{12}$$

Thus, we have $\Phi(\frac{1-\mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}}{\sqrt{\mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x}}}) \le \epsilon$, and this is equivalent to (11). On the other hand, if $\mathbf{x}^T\mathbf{C}_{\mathbf{a}}\mathbf{x} = 0$, for $1 > \epsilon > 0$, the equality $\mathbf{x}^T\mathbf{a} = \mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}$ holds with probability 1, and evidently, (10) and (11) are equivalent. □

According to Proposition 1, (9) can be rewritten as

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. -\Phi^{-1}(\epsilon)\|\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}\mathbf{x}\|_2 \le \boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} - 1, \end{aligned} \tag{13}$$

which is equivalent to

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. -\Phi^{-1}(\epsilon)\|\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}\mathbf{x}\|_2 = \boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} - 1. \end{aligned} \tag{14}$$

The proof of their equivalence here is simple: for any given feasible $\mathbf{x}$ with $-\Phi^{-1}(\epsilon)\|\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}\mathbf{x}\|_2 < \boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} - 1$, a point

$$\bar{\mathbf{x}} = \frac{\mathbf{x}}{\boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} + \Phi^{-1}(\epsilon)\|\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}\mathbf{x}\|_2} \tag{15}$$

can be constructed such that $-\Phi^{-1}(\epsilon)\|\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}\bar{\mathbf{x}}\|_2 = \boldsymbol{\mu}_{\mathbf{a}}^T\bar{\mathbf{x}} - 1$ and meanwhile $\bar{\mathbf{x}}^T\hat{\mathbf{R}}\bar{\mathbf{x}} < \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x}$ holds, thus establishing their equivalence. To solve (14), three cases of ϵ are considered: a) if $1 > \epsilon > 0.5$, then $-\Phi^{-1}(\epsilon) < 0$; (13) is equivalent to a problem of minimizing a convex quadratic function subject to a nonconvex quadratic constraint and an affine constraint, which seems to be intractable; b) if $\epsilon = 0.5$, (14) is an MVB pointed at $\boldsymbol{\mu}_{\mathbf{a}}$; c) if $0.5 > \epsilon > 0$, (14) is exactly (8) with $\mathbf{P}^{\frac{1}{2}} = -\Phi^{-1}(\epsilon)\mathbf{C}_{\mathbf{a}}^{\frac{1}{2}}$ and $\mathbf{c} = \boldsymbol{\mu}_{\mathbf{a}}$.

3.2. Distributions with Known Mean and Covariance

Here we consider the situation where $\mathbf{a}$ belongs to the family of distribution with known mean $\boldsymbol{\mu}_\mathbf{a}$ and covariance $\mathbf{C}_\mathbf{a}$, i.e., $\mathbf{a} \sim (\boldsymbol{\mu}_\mathbf{a}, \mathbf{C}_\mathbf{a})$. The corresponding beamforming problem can be written down as follows:

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T \hat{\mathbf{R}} \mathbf{x} \\ &s.t. \inf_{\mathbf{a}\sim(\boldsymbol{\mu}_\mathbf{a},\mathbf{C}_\mathbf{a})} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \geq 1\} \geq 1-\epsilon. \end{aligned} \tag{16}$$

Proposition 2: Suppose $\mathbf{a} \sim (\boldsymbol{\mu}_\mathbf{a}, \mathbf{C}_\mathbf{a})$ and $1 > \epsilon > 0$. The set of $\mathbf{x}$ vector satisfying

$$\inf_{\mathbf{a}\sim(\boldsymbol{\mu}_\mathbf{a},\mathbf{C}_\mathbf{a})} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \geq 1\} \geq 1-\epsilon \tag{17}$$

is the same as those satisfying

$$\sqrt{(1-\epsilon)/\epsilon}\|\mathbf{C}_\mathbf{a}^{\frac{1}{2}}\mathbf{x}\|_2 \leq \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} - 1. \tag{18}$$

Proof. Since $\mathbf{a} \sim (\boldsymbol{\mu}_\mathbf{a}, \mathbf{C}_\mathbf{a})$, we have $\mathbf{x}^T\mathbf{a} \sim (\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}, \mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})$, and the chance constraint can be rewritten as

$$\sup_{\mathbf{x}^T\mathbf{a}\sim(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a},\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \leq \epsilon. \tag{19}$$

We initially consider the case where $\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x} \neq 0$. According to the generalized Chebyshev inequality [12], [13], we have

$$\begin{aligned} &\sup_{\mathbf{x}^T\mathbf{a}\sim(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a},\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \\ &= \begin{cases} \mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x}/\big(\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x} + (\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}-1)^2\big), & \text{if } \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} \geq 1, \\ 1, & \text{if } \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} < 1. \end{cases} \end{aligned} \tag{20}$$

Under the assumption $1 > \epsilon > 0$, (19) reduces to two constraints $\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x}/\big(\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x} + (\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}-1)^2\big) \leq \epsilon$ and $\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} \geq 1$, or equivalently, $\sqrt{(1-\epsilon)/\epsilon}\|\mathbf{C}_\mathbf{a}^{\frac{1}{2}}\mathbf{x}\|_2 \leq \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} - 1$.

When $\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x} = 0$, the equality $\mathbf{x}^T\mathbf{a} = \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}$ holds with probability 1. Thus for $1 > \epsilon > 0$, we simply have

$$\sup_{\mathbf{x}^T\mathbf{a}\sim(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a},\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} = 0 \leq \epsilon \iff \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} \geq 1,$$

which completes our proof. ☐

Consequently, (16) is equivalent to

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T \hat{\mathbf{R}} \mathbf{x} \\ &s.t. \sqrt{(1-\epsilon)/\epsilon}\|\mathbf{C}_\mathbf{a}^{\frac{1}{2}}\mathbf{x}\|_2 = \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} - 1. \end{aligned} \tag{21}$$

3.3. Centrally Symmetric Distributions with Known Mean and Covariance

Consider the beamforming problem

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T \hat{\mathbf{R}} \mathbf{x} \\ &s.t. \inf_{\mathbf{a}\sim(\boldsymbol{\mu}_\mathbf{a},\mathbf{C}_\mathbf{a})_\mathcal{S}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \geq 1\} \geq 1-\epsilon, \end{aligned} \tag{22}$$

where, in addition to $\boldsymbol{\mu}_\mathbf{a}$ and $\mathbf{C}_\mathbf{a}$, the distribution of $\mathbf{a}$ is known to be centrally symmetric, i.e., $\mathbf{a} \sim (\boldsymbol{\mu}_\mathbf{a}, \mathbf{C}_\mathbf{a})_\mathcal{S}$. The distribution $P_\mathbf{a}$ is said to be centrally symmetric, if $P_\mathbf{a}(\boldsymbol{\mu}_\mathbf{a} + \mathcal{B}) = P_\mathbf{a}(\boldsymbol{\mu}_\mathbf{a} - \mathcal{B})$ for all Borel sets $\mathcal{B} \in \mathbb{R}^{2N}$.

Proposition 3: Suppose $\mathbf{a} \sim (\boldsymbol{\mu}_\mathbf{a}, \mathbf{C}_\mathbf{a})_\mathcal{S}$. If $0.5 > \epsilon > 0$, the distributionally robust chance constraint

$$\inf_{\mathbf{a}\sim(\boldsymbol{\mu}_\mathbf{a},\mathbf{C}_\mathbf{a})_\mathcal{S}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \geq 1\} \geq 1-\epsilon \tag{23}$$

is equivalent to

$$\sqrt{1/(2\epsilon)}\|\mathbf{C}_\mathbf{a}^{\frac{1}{2}}\mathbf{x}\|_2 \leq \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} - 1. \tag{24}$$

For $1 > \epsilon \geq 0.5$, (23) reduces to $\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} \geq 1$.

Proof. Since $\mathbf{a} \sim (\boldsymbol{\mu}_\mathbf{a}, \mathbf{C}_\mathbf{a})_\mathcal{S}$, we have $\mathbf{x}^T\mathbf{a} \sim (\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}, \mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})_\mathcal{S}$, and the chance constraint can be rewritten as

$$\sup_{\mathbf{x}^T\mathbf{a}\sim(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a},\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})_\mathcal{S}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \leq \epsilon. \tag{25}$$

For $\mathbf{x}$ with $\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x} > 0$, applying the generalized Chebyshev inequality for symmetric distribution [13], [15], we have

$$\begin{aligned} &\sup_{\mathbf{x}^T\mathbf{a}\sim(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a},\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x})_\mathcal{S}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \\ &= \begin{cases} 0.5\min\{1,\ \mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x}/(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}-1)^2\}, & \text{if } \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} > 1, \\ 0.5, & \text{if } \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} = 1, \\ 1, & \text{if } \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} < 1. \end{cases} \end{aligned} \tag{26}$$

If $1 > \epsilon \geq 0.5$, according to (26), (25) reduces to $\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} \geq 1$. For $0.5 > \epsilon > 0$, the chance constraint is equivalent to $0.5\min\{1,\ \mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x}/(\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}-1)^2\} \leq \epsilon$ and $\mathbf{x}^T\boldsymbol{\mu}_\mathbf{a} > 1$, or equivalently, $\sqrt{1/(2\epsilon)}\|\mathbf{C}_\mathbf{a}^{\frac{1}{2}}\mathbf{x}\|_2 \leq \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} - 1$.

When $\mathbf{x}^T\mathbf{C}_\mathbf{a}\mathbf{x} = 0$, for $1 > \epsilon > 0$, $\mathbf{x}^T\mathbf{a} = \mathbf{x}^T\boldsymbol{\mu}_\mathbf{a}$ holds with probability 1, and thus (24) and (25) are equivalent. ☐

Therefore, for $0.5 > \epsilon > 0$, (22) is equivalent to

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T \hat{\mathbf{R}} \mathbf{x} \\ &s.t. \sqrt{1/(2\epsilon)}\|\mathbf{C}_\mathbf{a}^{\frac{1}{2}}\mathbf{x}\|_2 = \boldsymbol{\mu}_\mathbf{a}^T\mathbf{x} - 1. \end{aligned} \tag{27}$$

When $1 > \epsilon \geq 0.5$, (22) is exactly the MVB point at $\boldsymbol{\mu}_\mathbf{a}$.

3.4. Variations in independent and bounded intervals

Here we consider the following uncertainty model: $\mathbf{a} = \boldsymbol{\mu}_\mathbf{a} + \boldsymbol{\omega}$, where the mean $\boldsymbol{\mu}_\mathbf{a}$ is known and $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_{2N}$ are independent random variables such that $\forall i$, $\boldsymbol{\omega}_i$ takes its values in $[l_i, u_i]$ with probability 1. Note that $\forall i$, $u_i \geq 0 \geq l_i$ Let $(\boldsymbol{\mu}_\mathbf{a}, \mathbf{D})_\mathcal{I}$ be the family of distributions of $\mathbf{a}$ satisfying the above condition, where $\mathbf{D} = \mathbf{diag}\{u_1 - l_1, \ldots, u_{2N} - l_{2N}\}$. The robust beamformng problem can be formulated as

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{x}^T \hat{\mathbf{R}} \mathbf{x} \\ &s.t. \inf_{\mathbf{a}\sim(\boldsymbol{\mu}_\mathbf{a},\mathbf{D})_\mathcal{I}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} \geq 1\} \geq 1-\epsilon. \end{aligned} \tag{28}$$

Proposition 4: For $1 > \epsilon > 0$, the robust chance constraint

$$\sup_{\mathbf{a}\sim(\boldsymbol{\mu}_{\mathbf{a}},\mathbf{D})_{\mathcal{I}}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \leq \epsilon \tag{29}$$

holds if

$$\sqrt{-0.5\ln(\epsilon)}\|\mathbf{D}\mathbf{x}\|_2 \leq \boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} - 1. \tag{30}$$

Proof. We first write the minimization in the constraint as

$$\sup_{\mathbf{a}\sim(\boldsymbol{\mu}_{\mathbf{a}},\mathbf{D})_{\mathcal{I}}} \mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \leq \epsilon. \tag{31}$$

For $\mathbf{x}$ with $\mathbf{D}\mathbf{x} \neq 0$, applying the Hoeffding's inequality [15], [16], we have that

$$\begin{aligned}\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} =& \mathbf{Pr}\{-\mathbf{x}^T\boldsymbol{\omega} > \mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}} - 1\} \\ \leq& \exp\{-\frac{2(\mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}-1)^2}{\sum_{i=1}^{2N}\mathbf{x}_i^2(u_i-l_i)^2}\}.\end{aligned} \tag{32}$$

Note that (30) implies $\exp\{-\frac{2(\mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}-1)^2}{\sum_{i=1}^{2N}\mathbf{x}_i^2(u_i-l_i)^2}\} \leq \epsilon$; then $\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \leq \epsilon$ holds for all $\mathbf{a} \sim (\boldsymbol{\mu}_{\mathbf{a}},\mathbf{D})_{\mathcal{I}}$. For $\mathbf{D}\mathbf{x} = 0$, from (30), we have $\boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} \geq 1$. We now show $\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} = 0$, and thus $\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} \leq \epsilon$. To show $\mathbf{Pr}\{\mathbf{x}^T\mathbf{a} < 1\} = 0$, or equivalently, $\mathbf{Pr}\{\mathbf{x}^T\boldsymbol{\omega} \geq 1 - \mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}\} = 1$, we prove $\mathbf{Pr}\{\mathbf{x}^T\boldsymbol{\omega} \geq 1 - \mathbf{x}^T\boldsymbol{\mu}_{\mathbf{a}}\} \geq \mathbf{Pr}\{\mathbf{x}^T\boldsymbol{\omega} = 0\} = 1$. Since $\mathbf{D}\mathbf{x} = 0$, we have $(u_i - l_i)\mathbf{x}_i = 0$, $\forall i$. Two cases of i are discussed: a) if $u_i - l_i > 0$, then $\mathbf{x}_i = 0$; b) if $u_i = l_i$, from $u_i \geq 0 \geq l_i$, then $\boldsymbol{\omega}_i = 0$ holds with probability 1. As a result, $\mathbf{Pr}\{\mathbf{x}^T\boldsymbol{\omega} = 0\} = 1$, which completes our proof. □

As a result, a tractable approximation of (28) is given by

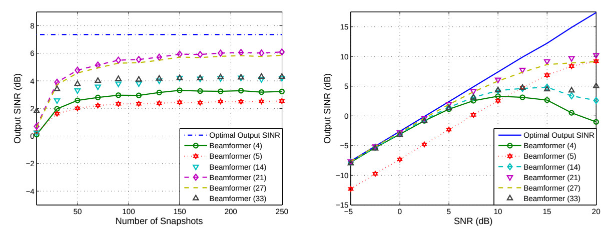
$$\begin{aligned}&\min_{\mathbf{x}} \mathbf{x}^T\hat{\mathbf{R}}\mathbf{x} \\ &s.t. \sqrt{-0.5\ln(\epsilon)}\|\mathbf{D}\mathbf{x}\|_2 = \boldsymbol{\mu}_{\mathbf{a}}^T\mathbf{x} - 1.\end{aligned} \tag{33}$$

When $\epsilon = 1$, (33) is the MVB pointed at $\boldsymbol{\mu}_{\mathbf{a}}$.

4. NUMERICAL RESULTS

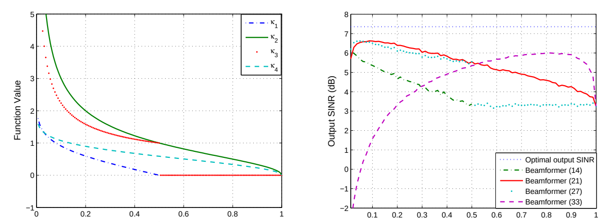
We present below three numerical examples to demonstrate the performance of (14), (21), (27), and (33). In all examples, a ten-sensor uniform linear array centered at the origin and spaced 0.4-wavelength apart is considered. Both the sensor position error and the angle of arrival (AOA) error are simulated: the position of each element is perturbed independently by a Gaussian vector $\mathcal{N}(\mathbf{0},(0.015\lambda)^2\mathbf{I}_{2\times2})$; θ_1 is modeled as a binary random variable with probability $\Pr\{\theta_1 = 5°\} = 0.8$ and $\Pr\{\theta_1 = 6°\} = 0.2$. The observations are generated as follows: $m_{\theta_1} = 100$ realizations of θ_1 are collected; for each value of θ_1, $m_{\mathrm{p}} = 1000$ samples are independently generated, i.e. $\{\mathbf{a}_i^1\}_i$ and $\{\mathbf{a}_i^2\}_i$. Then $\boldsymbol{\mu}_{\mathbf{a}}$ and $\mathbf{C}_{\mathbf{a}}$ can approximately determined from the data, i.e., $\boldsymbol{\mu}_{\mathbf{a}} = (m_1\hat{\mathbf{a}}_1 + m_2\hat{\mathbf{a}}_2)/m_{\theta_1}$, $\mathbf{C}_{\mathbf{a}} = \big(m_1\sum_{i=1}^{m_{\mathrm{p}}}(\mathbf{a}_i^1 - \boldsymbol{\mu}_{\mathbf{a}})^2 + m_2\sum_{i=1}^{m_{\mathrm{p}}}(\mathbf{a}_i^2 - \boldsymbol{\mu}_{\mathbf{a}})^2\big)/(m_{\theta_1}m_{\mathrm{p}})$, where m_1 and m_2 denote the numbers that $\{\theta_1 = 5°\}$ and $\{\theta_1 = 6°\}$ are observed.

Fig. 1 shows the output SINR as functions of the snapshot and the SNR. In this example, six interferers with powers all equal to 20 dB impinge from $\theta_1 = 15°$, $\theta_2 = 25°$, $\theta_3 = 40°$, $\theta_4 = -5°$, $\theta_5 = -15°$, and $\theta_6 = -30°$. From Fig. 1, we can see that (19) performs the most robust, yielding the highest SINR. Let $\kappa_1 = -\Phi^{-1}(\epsilon)$, $0.5 \geq \epsilon \geq 0$, $\kappa_2 = \sqrt{(1-\epsilon)/\epsilon}$, $\kappa_3 = \sqrt{1/(2\epsilon)}$, $0.5 \geq \epsilon \geq 0$, $\kappa_3 = 0$, $1 \geq \epsilon > 0.5$, and $\kappa_4 = \sqrt{-0.5\ln(\epsilon)}$. Fig. 2(a) plots κ_1, κ_2, κ_3, and κ_4. For (14), (21), (27), we can see that $\kappa_1 \geq \kappa_2 \geq \kappa_3$ for $0.5 \geq \epsilon \geq 0$ and $\kappa_2 \geq \kappa_3$ for $1 \geq \epsilon \geq 0.5$, indicating that (21) maintains the gain for the largest region. Fig. 2(b) shows the SINR as a function of ϵ, where $K = 250$ and SINR=10 dB. We note that (21) achieves its best SINR at about 0.1, whereas (33), as an approximation, achieves its best SINR at about 0.85.



(a) (b)

Fig. 1. SINR comparison with $\epsilon = 0.3$: (a) SINR versus K with SNR =10 dB; SINR versus SNR with $K = 250$.



(a) (b)

Fig. 2. (a) Plots of κ_1, κ_2, κ_3, and κ_4, (b) SINR versus ϵ.

5. CONCLUSIONS

This paper studies distributionally robust chance-constrained minimum variance beamforming. The essence of our approach is to employ the distributional sets and to use distributionally robust chance-constraints to design the weights. The beamformer is first derived for Gaussian model, and then further developed for more complex distributional sets, providing statistical interpretation to the deterministic counterpart. Finally, simulations are conducted to compare the performance of these beamformers.

6. REFERENCES

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

[2] 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.

[3] D. D. Feldman and L. J. Griffiths, "A projection approach to robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 42, pp. 867–876, Apr. 1994.

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

[5] 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.

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

[7] 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.

[8] 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.

[9] 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.

[10] X. Zhang, Y. Li, N. Ge, and J. Lu, "Robust minimum variance beamforming under distributional uncertainty," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2514–2518, 2015.

[11] A. Prekopa, *Stochastic Programming*. Dodrecht, N.L.: Kluwer Academic Publishers, 1995.

[12] D. Bertsimas and I. Popescu, "Optimal inequalities in probability theory: A convex optimization approach," *SIAM J. Optim.*, vol. 19, no. 3, pp. 780–804, 2005.

[13] I. Popescu, "An SDP approach to optimal moment bounds for convex classes of distributions," *Math. of Oper. Res.*, vol. 50, no. 3, pp. 632–657, 2005.

[14] L. El. Ghaoui, M. Oks, and F. Oustry, "Worst-case value-at-risk and robust portfolio optimization: A conic programming approach," *Oper. Res.*, vol. 51, no. 4, pp. 543–556, Jul.-Aug., 2003.

[15] G. C. Calafiore and L. El. Ghaoui, "On distributionally robust chance-constrained linear programs," *Journ. Optim. Theory and Appl.*, vol. 130, no. 1, pp. 1–22, 2006.

[16] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A nonasymptotic theory of independence*. Oxford, U.K.: Oxford Univ. Press, 2013.