**SCIENCE CHINA**
Information Sciences

• RESEARCH PAPER •

February 2016, Vol. 59 022309:1–022309:13
doi: 10.1007/s11432-015-5342-6

# Robust secure transmission for multiuser MISO systems with probabilistic QoS constraints

Lijian ZHANG[1], Liang JIN[1*], Wenyu LUO[1], Chunming WANG[1] & Yanqun TANG[2]

[1]*National Digital Switching System Engineering & Technological Research Center, Zhengzhou* 450002*, China;*
[2]*School of Information Systems Engineering , Zhengzhou Information Science and Technology Institute, Zhengzhou* 450001*, China*

Received October 22, 2014; accepted March 18, 2015; published online January 4, 2016

**Abstract** This paper considers a unicast multiuser multiple-input single-output (MISO) downlink system overheard by multiple single-antenna eavesdroppers. The objective is to jointly design the beamforming vectors and the artificial noise (AN) covariance matrix with imperfect channel state information (CSI) at the transmitter, such that the total transmit power is minimized while satisfying probabilistic quality of service (QoS) constraints at legitimate users and the eavesdroppers. Using Bernstein-type inequalities and the $\mathcal{S}$-procedure, we recast the non-convex power minimization problem as two different convex semidefinite programs (SDPs) which can be solved using interior-point methods. Simulation results show that the proposed methods outperform a non-robust method and the ones using the isotropic AN.

**Keywords** physical layer security, artificial noise, imperfect CSI, robust beamforming, semidefinite program, outage probability

**Citation** Zhang L J, Jin L, Luo W Y, et al. Robust secure transmission for multiuser MISO systems with probabilistic QoS constraints. Sci China Inf Sci, 2016, 59(2): 022309, doi: 10.1007/s11432-015-5342-6

## 1 Introduction

With the rapid advances of wireless communications, the conventional cryptographic encryption [1, 2] used in the upper layers of communication networks has encountered more challenges in areas such as key distribution and management. For this reason, physical layer security that exploits the characteristics of wireless channels to guarantee secure transmission has recently become an active research area. The groundwork for physical layer security was laid by Wyner [3].

As a promising technology for high data rate communications, multiple antenna systems have attracted considerable attention [4]. It has been shown that the spatial degrees of freedom provided by multiple antennas can be used to enhance the physical layer security [5, 6]. Much existing work [7, 8] on physical layer security is based on the assumption of perfect channel state information (CSI). However, CSI error may be inevitable due to many reasons such as estimation error, quantization error, feedback delay, and time delays between reciprocal channels [9]. As such, robust transmit design has recently become an active topic to mitigate the impacts of imperfect CSI. Under the deterministic model, the worst-case

* Corresponding author (email: liangjin@263.net)

 

approach aims to satisfy the constraints for all the channel realizations [5, 10, 11]. Under the stochastic model, the transmit design is usually developed based on the ergodic (average) performance or the outage performance in various wiretap scenarios, for example, relay channels [12,13], single-input multiple-output (SIMO) channels [14], multiple-input multiple-output (MIMO) channels [15, 16], multiple-input single-output (MISO) channels [17–19], and interference channels [20].

In [16, 18, 19], robust wiretap transmission problem was investigated from the quality of service (QoS) perspective. The average minimum mean square error (MMSE) and the average signal-to-interference-and-noise ratio (SINR) are used as the security performance metrics in [16] and [18], respectively. However, in practice the wireless channels are usually not ergodic, and the instant QoS requirement may be violated with a high probability. Hence, an outage-based performance metric is more realistic. In [19], the authors proposed an SINR outage-based approach to secure the message delivery, where only one legitimate user was considered. Also, the perfect CSI of legitimate user is assumed to be available. For analyzing the statistical QoS constraints, an effective capacity region was derived and a method to obtain it was proposed for the two-user opportunistic spectrum access (OSA) system in [21], while the security issue was not considered. To the best of our knowledge, there are few studies on physical layer security considering the unicast multiuser scenario with probabilistic QoS constraints at both the legitimate users and the eavesdroppers. Besides, in many existing studies [3, 16, 19, 22, 23], the artificial noise (AN) is imposed on the null space of the legitimate channel, called isotropic AN, to confuse the eavesdroppers. Actually, this method may not be optimal when the CSI of eavesdroppers can be obtained partially or fully. This is because the CSI of both the legitimate users and the eavesdroppers can be simultaneously exploited to improve the security performance.

In this paper, we address the physical layer security for the unicast multiuser MISO downlink systems. Only the imperfect CSI of both the legitimate users and the eavesdroppers is available at the transmitter. The channel error vectors are zero-mean circularly symmetric complex Gaussian with known covariance matrix. To enhance the security, a robust AN-aided beamforming scheme is proposed, where the structure of AN is non-isotropic. Specifically, we aim to minimize the total transmit power under probabilistic QoS constraints by the joint optimization of beamforming vectors and the AN covariance matrix. Each legitimate receiver's probability of experiencing SINR outage should be below a preset level. On the other hand, the probability that each eavesdropper's SINR exceeds the given threshold should fall below a predefined value. The challenge in robust transmit design lies in the lack of close-form expressions of the probabilistic QoS constraints. Also, the considered problem formulations are non-convex in general and hard to solve optimally. To overcome these difficulties, we adopt the semidefinite relaxation (SDR) technique [24] and the approximation techniques, that is, the Bernstein-type inequalities [25] and the $\mathcal{S}$-procedure [26], to recast the original optimization problem as two different semidefinite programs (SDPs) which can be efficiently solved using interior-point methods [26]. Finally, the performance of the proposed scheme is evaluated through simulation studies.

Notations: Bold uppercase letters and bold lowercase letters denote matrices and vectors, respectively. $\mathrm{Tr}(\cdot)$, $(\cdot)^{\mathrm{H}}$, $\mathbb{C}^N$, $\mathbb{H}^N$, $\mathbb{C}^{N\times M}$, $|\cdot|$, $\|\cdot\|$, $\Pr\{\cdot\}$, $\mathrm{E}\{\cdot\}$, and $\mathrm{Re}(\cdot)$ denote the trace of a matrix, Hermitian transpose, the space of $N \times 1$ complex vector, the space of $N \times N$ Hermitian matrix, the space of $N \times M$ complex matrix, the absolute value, the Euclidean norm, the probability operator, the expectation operator, and the real part, respectively. $\mathbb{R}$ denotes the set of all real numbers. $\boldsymbol{I}_N$ is the identity matrix of size $N \times N$. $\boldsymbol{Q} \succeq \mathbf{0}$ means that $\boldsymbol{Q}$ is a positive semidefinite (PSD) matrix, and $\mathrm{vec}(\boldsymbol{Q})$ denotes a column vector by stacking all the elements of $\boldsymbol{Q}$. $\boldsymbol{x} \sim \mathcal{CN}(\boldsymbol{c}, \boldsymbol{Q})$ means that $\boldsymbol{x}$ is a complex circular Gaussian random vector with mean $\boldsymbol{c}$ and covariance $\boldsymbol{Q}$.

## 2 System model and problem formulation

We consider a multiuser secure communication system with an $N$-antenna transmitter (Alice), $M$ single-antenna legitimate receivers (Bobs), and $K$ single-antenna eavesdroppers (Eves). The Eves are non-concluding. This model corresponds to the following scenario: all the legitimate users and the eavesdroppers belong to the same system, but each one subscribes to different services. When subscribing to no services, the user will be regarded as an eavesdropper and kept from receiving any useful information. Let $\boldsymbol{h}_m \in \mathbb{C}^N$, $\forall m \in \mathcal{M} \triangleq \{1, \ldots, M\}$, denote the channel vector from Alice to the $m$th Bob, and $\boldsymbol{g}_k \in \mathbb{C}^N$, $\forall k \in \mathcal{K} \triangleq \{1, \ldots, K\}$, denote the channel vector from Alice to the $k$th Eve. It is assumed that all the channels are independent of each other and undergo slow frequency-flat fading. Let $\boldsymbol{x} \in \mathbb{C}^N$ denote the AN-assisted transmit signal vector. The received signals at the $m$th Bob and $k$th Eve are, respectively, given by

$$y_{b,m} = \boldsymbol{h}_m^{\mathrm{H}} \boldsymbol{x} + n_{b,m}, \ \forall m \in \mathcal{M}, \tag{1a}$$

$$y_{e,k} = \boldsymbol{g}_k^{\mathrm{H}} \boldsymbol{x} + n_{e,k}, \ \forall k \in \mathcal{K}, \tag{1b}$$

where $n_{b,m} \sim \mathcal{CN}(0, \sigma_{b,m}^2)$ and $n_{e,k} \sim \mathcal{CN}(0, \sigma_{e,k}^2)$ are additive white complex Gaussian noises at the $m$th Bob and the $k$th Eve, respectively. The transmitted signal vector $\boldsymbol{x}$ has the following structure here

$$\boldsymbol{x} = \sum_{m=1}^{M} \boldsymbol{w}_m s_m + \boldsymbol{z}, \tag{2}$$

where $s_m \in \mathbb{C}$ is the confidential and independent symbol transmitted from Alice to the $m$th Bob, with $\boldsymbol{E}\{|s_m|^2\} = 1$; $\boldsymbol{w}_m \in \mathbb{C}^N$ is the transmit beamforming vector corresponding to $s_m$; and $\boldsymbol{z} \in \mathbb{C}^N$ is the AN vector generated by Alice to interfere Eves. We assume that $\boldsymbol{z} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} \succeq \boldsymbol{0}$ being the AN covariance matrix.

In practical scenarios, Alice may not have perfect CSI in general. Following the complex Gaussian stochastic error model [27], the true channels can be expressed as

$$\boldsymbol{h}_m = \hat{\boldsymbol{h}}_m + \boldsymbol{\Delta}_{b,m}, \ \forall m \in \mathcal{M}, \tag{3a}$$

$$\boldsymbol{g}_k = \hat{\boldsymbol{g}}_k + \boldsymbol{\Delta}_{e,k}, \ \forall k \in \mathcal{K}, \tag{3b}$$

where $\hat{\boldsymbol{h}}_m \in \mathbb{C}^N$ and $\hat{\boldsymbol{g}}_k \in \mathbb{C}^N$ are the estimated CSI vectors at Alice; $\boldsymbol{\Delta}_{b,m} \in \mathbb{C}^N$ and $\boldsymbol{\Delta}_{e,k} \in \mathbb{C}^N$ are the independent CSI error vectors. We assume that $\boldsymbol{\Delta}_{b,m} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q}_{b,m})$, $\boldsymbol{Q}_{b,m} \succ \boldsymbol{0}$ and $\boldsymbol{\Delta}_{e,k} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q}_{e,k})$, $\boldsymbol{Q}_{e,k} \succ \boldsymbol{0}$.

According to (1) and (2), the SINRs of the $m$th Bob and the $k$th Eve which eavesdrops the $m$th data stream can be, respectively, expressed as

$$\mathrm{SINR}_{b,m} = \frac{|\boldsymbol{h}_m^{\mathrm{H}} \boldsymbol{w}_m|^2}{\sum_{i=1, i\neq m}^{M} |\boldsymbol{h}_m^{\mathrm{H}} \boldsymbol{w}_i|^2 + \boldsymbol{h}_m^{\mathrm{H}} \boldsymbol{\Sigma} \boldsymbol{h}_m + \sigma_{b,m}^2}, \tag{4a}$$

$$\mathrm{SINR}_{e,k}^{(m)} = \frac{|\boldsymbol{g}_k^{\mathrm{H}} \boldsymbol{w}_m|^2}{\sum_{i=1, i\neq m}^{M} |\boldsymbol{g}_k^{\mathrm{H}} \boldsymbol{w}_i|^2 + \boldsymbol{g}_k^{\mathrm{H}} \boldsymbol{\Sigma} \boldsymbol{g}_k + \sigma_{e,k}^2}. \tag{4b}$$

In this paper, the problem of interest is to design the beamforming vectors and the AN covariance matrix, such that the total transmit power at Alice is minimized while satisfying the probabilistic QoS constraints at Bobs and Eves. Mathematically, the optimization problem can be formulated as

$$\min_{\{\boldsymbol{w}_m\}_{m=1}^{M}, \boldsymbol{\Sigma}} \ \sum_{m=1}^{M} \|\boldsymbol{w}_m\|^2 + \mathrm{Tr}(\boldsymbol{\Sigma}) \tag{5a}$$

$$\text{s.t.} \quad \Pr\Big\{\mathrm{SINR}_{b,m} \leqslant \gamma_{b,m}\Big\} \leqslant \rho_{b,m}, \ \forall m \in \mathcal{M}, \tag{5b}$$

$$\Pr\Big\{\mathrm{SINR}_{e,k}^{(m)} \geqslant \gamma_{e,k}^{(m)}\Big\} \leqslant \rho_{e,k}^{(m)}, \ \forall m \in \mathcal{M}, \ \forall k \in \mathcal{K}, \tag{5c}$$

$$\boldsymbol{\Sigma} \succeq \boldsymbol{0}, \tag{5d}$$

where $\gamma_{b,m}$ and $\gamma_{e,k}^{(m)}$ are the predefined SINR thresholds of the $m$th Bob and the $k$th Eve eavesdropping the $m$th data stream, respectively; $\rho_{b,m} \in (0,1]$ and $\rho_{e,k}^{(m)} \in (0,1]$ are the maximum tolerable probability values at the $m$th Bob and the $k$th Eve eavesdropping the $m$th data stream, respectively. The system's reliability is assured when the constraints in (5b) are satisfied, while the system's security is guaranteed if the constraints in (5c) are satisfied.

The optimization problem (5) is challenging to solve due to the fact that probabilistic functions have no closed-form solutions generally. Also, such problem is non-convex and hard to solve. In Section 3, we will propose two different methods to handle this optimization problem.

## 3 Robust optimization

In this section, the optimization problem (5) is solved by taking the following two steps: SDR and conservative reformulation. The first step will be discussed in Subsection 3.1. For the second step, two different conservative formulations will be obtained in Subsection 3.2. The key idea of the conservative reformulation is to develop a convex and tractable expressions of the upper bounds of the probabilistic QoS constraints in (5b) and (5c).

### 3.1 Semidefinite relaxation

We first define $\boldsymbol{W}_m \triangleq \boldsymbol{w}_m\boldsymbol{w}_m^{\mathrm{H}}$, and $\mathrm{rank}(\boldsymbol{W}_m) = 1$, $\forall m \in \mathcal{M}$. Then, according to (3) and (4), we can reformulate the probabilistic constraints in (5b) and (5c), respectively, as

$$\Pr\Big\{\boldsymbol{\Delta}_{b,m}^{\mathrm{H}}\boldsymbol{B}_{b,m}\boldsymbol{\Delta}_{b,m} + 2\mathrm{Re}\big\{\boldsymbol{\Delta}_{b,m}^{\mathrm{H}}\boldsymbol{B}_{b,m}\hat{\boldsymbol{h}}_m\big\} + \hat{\boldsymbol{h}}_m^{\mathrm{H}}\boldsymbol{B}_{b,m}\hat{\boldsymbol{h}}_m \leqslant \sigma_{b,m}^2\Big\} \leqslant \rho_{b,m},\ \forall m \in \mathcal{M}, \tag{6a}$$

$$\Pr\Big\{\boldsymbol{\Delta}_{e,k}^{\mathrm{H}}\boldsymbol{B}_{e,k}^{(m)}\boldsymbol{\Delta}_{e,k} + 2\mathrm{Re}\big\{\boldsymbol{\Delta}_{e,k}^{\mathrm{H}}\boldsymbol{B}_{e,k}^{(m)}\hat{\boldsymbol{g}}_k\big\} + \hat{\boldsymbol{g}}_k^{\mathrm{H}}\boldsymbol{B}_{e,k}^{(m)}\hat{\boldsymbol{g}}_k \geqslant \sigma_{e,k}^2\Big\} \leqslant \rho_{e,k}^{(m)},\ \forall m \in \mathcal{M},\ \forall k \in \mathcal{K}, \tag{6b}$$

where

$$\begin{cases} \boldsymbol{B}_{b,m} = \frac{1}{\gamma_{b,m}}\boldsymbol{W}_m - \sum_{i=1,i\neq m}^{M}\boldsymbol{W}_i - \boldsymbol{\Sigma}, \\ \boldsymbol{B}_{e,k}^{(m)} = \frac{1}{\gamma_{e,k}^{(m)}}\boldsymbol{W}_m - \sum_{i=1,i\neq m}^{M}\boldsymbol{W}_i - \boldsymbol{\Sigma}. \end{cases} \tag{7}$$

Due to the fact that $\boldsymbol{\Delta}_{b,m} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q}_{b,m})$ and $\boldsymbol{\Delta}_{e,k} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q}_{e,k})$, we can rewrite the CSI error vectors as

$$\boldsymbol{\Delta}_{b,m} = \boldsymbol{Q}_{b,m}^{1/2}\boldsymbol{r}_{b,m},\ \forall m \in \mathcal{M}, \tag{8a}$$

$$\boldsymbol{\Delta}_{e,k} = \boldsymbol{Q}_{e,k}^{1/2}\boldsymbol{r}_{e,k},\ \ \forall k \in \mathcal{K}, \tag{8b}$$

where $\boldsymbol{r}_{b,m} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N)$; $\boldsymbol{r}_{e,k} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N)$; $\boldsymbol{Q}_{b,m}^{1/2}$ and $\boldsymbol{Q}_{e,k}^{1/2}$ are the PSD square roots of $\boldsymbol{Q}_{b,m}$ and $\boldsymbol{Q}_{e,k}$, respectively, that is, $\boldsymbol{Q}_{b,m} = \boldsymbol{Q}_{b,m}^{1/2}\boldsymbol{Q}_{b,m}^{1/2}$ and $\boldsymbol{Q}_{e,k} = \boldsymbol{Q}_{e,k}^{1/2}\boldsymbol{Q}_{e,k}^{1/2}$. After substituting (8a) and (8b) into (6a) and (6b), respectively, we can equivalently rewrite the probabilistic constraints in (6) as follows:

$$\Pr\Big\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{D}_{b,m}\boldsymbol{r}_{b,m} + 2\mathrm{Re}\big\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{d}_{b,m}\big\} \leqslant c_{b,m}\Big\} \leqslant \rho_{b,m},\ \forall m \in \mathcal{M}, \tag{9a}$$

$$\Pr\Big\{\boldsymbol{r}_{e,k}^{\mathrm{H}}\boldsymbol{D}_{e,k}^{(m)}\boldsymbol{r}_{e,k} + 2\mathrm{Re}\big\{\boldsymbol{r}_{e,k}^{\mathrm{H}}\boldsymbol{d}_{e,k}^{(m)}\big\} \geqslant c_{e,k}^{(m)}\Big\} \leqslant \rho_{e,k}^{(m)},\ \forall m \in \mathcal{M},\ \forall k \in \mathcal{K}, \tag{9b}$$

where

$$\begin{cases} \boldsymbol{D}_{b,m} = \boldsymbol{Q}_{b,m}^{1/2}\boldsymbol{B}_{b,m}\boldsymbol{Q}_{b,m}^{1/2},\ \ \boldsymbol{d}_{b,m} = \boldsymbol{Q}_{b,m}^{1/2}\boldsymbol{B}_{b,m}\hat{\boldsymbol{h}}_m,\ c_{b,m} = \sigma_{b,m}^2 - \hat{\boldsymbol{h}}_m^{\mathrm{H}}\boldsymbol{B}_{b,m}\hat{\boldsymbol{h}}_m, \\ \boldsymbol{D}_{e,k}^{(m)} = \boldsymbol{Q}_{e,k}^{1/2}\boldsymbol{B}_{e,k}^{(m)}\boldsymbol{Q}_{e,k}^{1/2},\ \ \boldsymbol{d}_{e,k}^{(m)} = \boldsymbol{Q}_{e,k}^{1/2}\boldsymbol{B}_{e,k}^{(m)}\hat{\boldsymbol{g}}_k,\ \ c_{e,k}^{(m)} = \sigma_{e,k}^2 - \hat{\boldsymbol{g}}_k^{\mathrm{H}}\boldsymbol{B}_{e,k}^{(m)}\hat{\boldsymbol{g}}_k. \end{cases} \tag{10}$$

Considering $\boldsymbol{W}_m \triangleq \boldsymbol{w}_m\boldsymbol{w}_m^{\mathrm{H}}$, $\forall m \in \mathcal{M}$, and substituting (9a) and (9b) into the problem (5), we can obtain

$$\min_{\{\boldsymbol{W}_m\}_{m=1}^M, \boldsymbol{\Sigma}} \ \sum_{m=1}^{M}\mathrm{Tr}(\boldsymbol{W}_m) + \mathrm{Tr}(\boldsymbol{\Sigma}) \tag{11a}$$

$$\text{s.t.} \quad \Pr\Big\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{D}_{b,m}\boldsymbol{r}_{b,m} + 2\mathrm{Re}\big\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{d}_{b,m}\big\} \leqslant c_{b,m}\Big\} \leqslant \rho_{b,m},\ \forall m \in \mathcal{M}, \tag{11b}$$

$$\Pr\Big\{\boldsymbol{r}_{e,k}^{\mathrm{H}}\boldsymbol{D}_{e,k}^{(m)}\boldsymbol{r}_{e,k}+2\mathrm{Re}\{\boldsymbol{r}_{e,k}^{\mathrm{H}}\boldsymbol{d}_{e,k}^{(m)}\}\geqslant c_{e,k}^{(m)}\Big\}\leqslant\rho_{e,k}^{(m)},\ \forall m\in\mathcal{M},\ \forall k\in\mathcal{K},\tag{11c}$$

$$\boldsymbol{W}_m\succeq\mathbf{0},\ \mathrm{rank}(\boldsymbol{W}_m)=1,\ \forall m\in\mathcal{M},\tag{11d}$$

$$\boldsymbol{\Sigma}\succeq\mathbf{0},\tag{11e}$$

where the constraints in (11d) are equivalent to the definition $\boldsymbol{W}_m\triangleq\boldsymbol{w}_m\boldsymbol{w}_m^{\mathrm{H}}$. Following the idea of SDR [24], we drop the non-convex constraint $\mathrm{rank}(\boldsymbol{W}_m)=1$, $\forall m\in\mathcal{M}$. Then, the rank-relaxed counterpart of the problem (11) can be expressed as

$$\min_{\{\boldsymbol{W}_m\}_{m=1}^M,\boldsymbol{\Sigma}}\ \sum_{m=1}^{M}\mathrm{Tr}(\boldsymbol{W}_m)+\mathrm{Tr}(\boldsymbol{\Sigma})\tag{12a}$$

$$\text{s.t.}\quad \Pr\Big\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{D}_{b,m}\boldsymbol{r}_{b,m}+2\mathrm{Re}\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{d}_{b,m}\}\leqslant c_{b,m}\Big\}\leqslant\rho_{b,m},\ \forall m\in\mathcal{M},\tag{12b}$$

$$\Pr\Big\{\boldsymbol{r}_{e,k}^{\mathrm{H}}\boldsymbol{D}_{e,k}^{(m)}\boldsymbol{r}_{e,k}+2\mathrm{Re}\{\boldsymbol{r}_{e,k}^{\mathrm{H}}\boldsymbol{d}_{e,k}^{(m)}\}\geqslant c_{e,k}^{(m)}\Big\}\leqslant\rho_{e,k}^{(m)},\ \forall m\in\mathcal{M},\ \forall k\in\mathcal{K},\tag{12c}$$

$$\boldsymbol{\Sigma}\succeq\mathbf{0},\ \boldsymbol{W}_m\succeq\mathbf{0},\ \forall m\in\mathcal{M}.\tag{12d}$$

Generally speaking, the problem (12) is a relaxed approximation to (11) due to the fact that the optimal solution $\boldsymbol{W}_m^\star$, $\forall m\in\mathcal{M}$, to the problem (12) may not satisfy the rank-one constraints in (11d). The procedure of getting the suboptimal beamforming vector from $\boldsymbol{W}_m^\star$ whose rank is not one will be developed in the next subsection.

### 3.2 Conservative reformulations

Because no closed-form expressions for the left-hand sides of the inequalities in (12b) and (12c) exist in general, it is still difficult to handle the relaxed optimization problem (12). In this subsection, the probabilistic constraints in (12b) and (12c) will be converted into two different deterministic ones that are *conservative* and tractable. If the deterministic constraints are satisfied, then the corresponding probabilistic constraints must be satisfied, which is the key idea of the conservative reformulation.

#### 3.2.1 *Conservative reformulation using Bernstein-type inequalities*

To convert the probabilistic constraints in (12b) and (12c) into deterministic ones, the following Lemma will be useful.

**Lemma 1** ([25]). For any $\boldsymbol{D}\in\mathbb{H}^N$, $\boldsymbol{d}\in\mathbb{C}^N$, $\boldsymbol{r}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{I}_N)$, and $\alpha\geqslant0$, we have

$$\Pr\Big\{\boldsymbol{r}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{r}+2\mathrm{Re}\{\boldsymbol{r}^{\mathrm{H}}\boldsymbol{d}\}\leqslant\mathrm{Tr}(\boldsymbol{D})-\sqrt{2\alpha(\|\mathrm{vec}(\boldsymbol{D})\|^2+2\|\boldsymbol{d}\|^2)}-\alpha s^-(\boldsymbol{D})\Big\}\leqslant\exp(-\alpha),\tag{13a}$$

$$\Pr\Big\{\boldsymbol{r}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{r}+2\mathrm{Re}\{\boldsymbol{r}^{\mathrm{H}}\boldsymbol{d}\}\geqslant\mathrm{Tr}(\boldsymbol{D})+\sqrt{2\alpha(\|\mathrm{vec}(\boldsymbol{D})\|^2+2\|\boldsymbol{d}\|^2)}+\alpha s^+(\boldsymbol{D})\Big\}\leqslant\exp(-\alpha),\tag{13b}$$

where $s^-(\boldsymbol{D})=\max\{\lambda_{\max}(-\boldsymbol{D}),0\}$, and $s^+(\boldsymbol{D})=\max\{\lambda_{\max}(\boldsymbol{D}),0\}$ with $\lambda_{\max}(\boldsymbol{D})$ representing the maximum eigenvalue of the matrix $\boldsymbol{D}$. The inequalities in (13) are called Bernstein-type inequalities.

Introducing auxiliary variable $\alpha_{b,m}=-\ln\rho_{b,m}$, and applying the Bernstein-type inequality in (13a), we have

$$\Pr\Big\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{D}_{b,m}\boldsymbol{r}_{b,m}+2\mathrm{Re}\{\boldsymbol{r}_{b,m}^{\mathrm{H}}\boldsymbol{d}_{b,m}\}\leqslant \mathrm{Tr}(\boldsymbol{D}_{b,m})-\sqrt{2\alpha_{b,m}(\|\mathrm{vec}(\boldsymbol{D}_{b,m})\|^2+2\|\boldsymbol{d}_{b,m}\|^2)}-\alpha_{b,m}s^-(\boldsymbol{D}_{b,m})\Big\}\leqslant\rho_{b,m},\ \forall m\in\mathcal{M}.\tag{14}$$

Comparing (12b) with (14), we know that if the following constraints

$$c_{b,m}\leqslant\mathrm{Tr}(\boldsymbol{D}_{b,m})-\sqrt{2\alpha_{b,m}(\|\mathrm{vec}(\boldsymbol{D}_{b,m})\|^2+2\|\boldsymbol{d}_{b,m}\|^2)}-\alpha_{b,m}s^-(\boldsymbol{D}_{b,m}),\ \forall m\in\mathcal{M}\tag{15}$$

hold true, then the constraints in (12b) are satisfied.

Similarly, applying the Bernstein-type inequality in (13b), we can obtain the conservative reformulations of the probabilistic constraints in (12c) as follows:

$$c_{e,k}^{(m)} \geqslant \operatorname{Tr}(\boldsymbol{D}_{e,k}^{(m)}) + \sqrt{2\alpha_{e,k}^{(m)}(\|\operatorname{vec}(\boldsymbol{D}_{e,k}^{(m)})\|^2 + 2\|\boldsymbol{d}_{e,k}^{(m)}\|^2)} + \alpha_{e,k}^{(m)} s^{+}(\boldsymbol{D}_{e,k}^{(m)}),\ \forall m \in \mathcal{M},\ \ \forall k \in \mathcal{K}, \tag{16}$$

where $\alpha_{e,k}^{(m)} = -\ln \rho_{e,k}^{(m)}$ is the auxiliary variable. Also, the constraints in (15) and (16) can be equivalently reformulated as [25]

$$\begin{cases} \operatorname{Tr}(\boldsymbol{D}_{b,m}) - \sqrt{2\alpha_{b,m}}\mu_{b,m} - \alpha_{b,m}\nu_{b,m} \geqslant c_{b,m}, \\ \left\| \begin{matrix} \operatorname{vec}(\boldsymbol{D}_{b,m}) \\ \sqrt{2}\boldsymbol{d}_{b,m} \end{matrix} \right\| \leqslant \mu_{b,m}, \\ \nu_{b,m}\boldsymbol{I}_N + \boldsymbol{D}_{b,m} \succeq \boldsymbol{0}, \\ \nu_{b,m} \geqslant 0. \end{cases} \quad \forall m \in \mathcal{M}, \tag{17a}$$

and

$$\begin{cases} \operatorname{Tr}(\boldsymbol{D}_{e,k}^{(m)}) + \sqrt{2\alpha_{e,k}^{(m)}}\mu_{e,k}^{(m)} + \alpha_{e,k}^{(m)}\nu_{e,k}^{(m)} \leqslant c_{e,k}^{(m)}, \\ \left\| \begin{matrix} \operatorname{vec}(\boldsymbol{D}_{e,k}^{(m)}) \\ \sqrt{2}\boldsymbol{d}_{e,k}^{(m)} \end{matrix} \right\| \leqslant \mu_{e,k}^{(m)}, \\ \nu_{e,k}^{(m)}\boldsymbol{I}_N - \boldsymbol{D}_{e,k}^{(m)} \succeq \boldsymbol{0}, \\ \nu_{e,k}^{(m)} \geqslant 0, \end{cases} \quad \forall m \in \mathcal{M},\ \forall k \in \mathcal{K}, \tag{17b}$$

respectively, where $\mu_{b,m}$, $\nu_{b,m}$, $\mu_{e,k}^{(m)}$, and $\nu_{e,k}^{(m)}$ are auxiliary variables. Note that the combination of the third and the fourth inequalities in (17a) is equivalent to the constraint $\nu_{b,m} \geqslant s^{-}(\boldsymbol{D}_{b,m}) = \max\{\lambda_{\max}(-\boldsymbol{D}_{b,m}), 0\}$, while the combination of the third and the fourth inequalities in (17b) is equivalent to the constraint $\nu_{e,k}^{(m)} \geqslant s^{+}(\boldsymbol{D}_{e,k}^{(m)}) = \max\{\lambda_{\max}(\boldsymbol{D}_{e,k}^{(m)}), 0\}$. Replacing (12b) and (12c) with (17a) and (17b), respectively, the optimization problem (12) can be rewritten as

$$\min_{\{\boldsymbol{W}_m\}_{m=1}^{M}, \boldsymbol{\Sigma}} \sum_{m=1}^{M} \operatorname{Tr}(\boldsymbol{W}_m) + \operatorname{Tr}(\boldsymbol{\Sigma}) \tag{18a}$$

$$\text{s.t.} \quad \text{(17a) and (17b) satisfied,} \tag{18b}$$

$$\boldsymbol{\Sigma} \succeq \boldsymbol{0},\ \boldsymbol{W}_m \succeq \boldsymbol{0},\ \ \forall m \in \mathcal{M}. \tag{18c}$$

Now, one can easily show that problem (18) is a SDP, and thus can be efficiently solved through efficient solvers, such as CVX [28]. If the solution gives rank-one $\boldsymbol{W}_m^*$, $\forall m \in \mathcal{M}$, the principal eigenvector corresponding to the only non-zero eigenvalue of $\boldsymbol{W}_m^*$ can be selected as the optimal beamforming vector $\boldsymbol{w}_m^*$. Otherwise, it is necessary to apply the rank-one approximation procedure [24] to get the approximate solution $\widehat{\boldsymbol{w}}_m^*$ to the problem (18). Here, we design one Gaussian randomization procedure provided in Table 1. This algorithm is custom-designed for (18), and follows the idea of [18].

$$\begin{aligned} \min_{\{\beta_m^{(f)}\}_{m=1}^{M}, \beta_z^{(f)}} \quad & \textstyle\sum_{m=1}^{M} \operatorname{Tr}(\boldsymbol{W}_m) + \operatorname{Tr}(\boldsymbol{\Sigma}) \\ \text{s.t.} \quad & \text{(17a) and (17b) satisfied.} \end{aligned} \tag{19}$$

### 3.2.2 *Conservative reformulation by $\mathcal{S}$-procedure*

In this part, we alternatively convert the probabilistic constraints in (12b) and (12c) into the worst-case deterministic constraints following the method proposed in [29]. Then, using the $\mathcal{S}$-Procedure, we transform the optimization problem (12) into a tractable SDP.

**Lemma 2** ([29])**.** Let $\boldsymbol{r} \sim \mathbb{C}^N$ be a continuous random vector following certain statistical distribution and $\boldsymbol{G}(\boldsymbol{r}) : \mathbb{C}^N \to \mathbb{R}$ be a function of $\boldsymbol{r}$. Then, the following implication holds

$$\boldsymbol{G}(\boldsymbol{r}) \geqslant 0,\ \forall \|\boldsymbol{r}\|^2 \leqslant R^2,\ \Pr\left\{\|\boldsymbol{r}\|^2 \leqslant R^2\right\} \geqslant 1-\rho \implies \Pr\left\{\boldsymbol{G}(\boldsymbol{r}) \leqslant 0\right\} \leqslant \rho, \tag{20}$$

**Table 1** Gaussian randomization procedure for the optimization problem (18)

**Initialization:** The number of randomizations $F$ and an optimal solution $\{\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_M^*,\ \boldsymbol{\Sigma}^*\}$ to the SDR problem (18) .

**for** $f=1,\ldots,F$

1: generate the set of beamforming vectors $\{\boldsymbol{w}_1^{(f)},\ldots,\boldsymbol{w}_M^{(f)}\},\ \boldsymbol{w}_m^{(f)}\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{W}_m^*),\ \forall m\in\mathcal{M}$;

2: Let $\beta_m^{(f)}\geqslant 0,\ \forall m\in\mathcal{M}$ and $\beta_z^{(f)}\geqslant 0$ denote the power scaling factors, substitute $\boldsymbol{W}_m=\beta_m^{(f)}\boldsymbol{w}_m^{(f)}(\boldsymbol{w}_m^{(f)})^{\rm H},\ \forall m\in\mathcal{M}$ and $\boldsymbol{\Sigma}=\beta_z^{(f)}\boldsymbol{\Sigma}^*$ into the optimization problem (19), solve it with respect to $\{\beta_1^{(f)},\ldots,\beta_M^{(f)},\beta_z^{(f)}\}$;

3: If the problem is feasible, then set the total power $P^{(f)}$ to be the optimal objective value, or set $P^{(f)}=+\infty$;

**end**

Let $f^*=\arg\min_{f=1,\ldots,F}P^{(f)}$;

**Output:** $\widehat{\boldsymbol{w}}_m^\star=\sqrt{\beta_m^{(f^*)}}\boldsymbol{w}_m^{(f^*)},\ \forall m\in\mathcal{M},\ \widehat{\boldsymbol{\Sigma}}^\star=\beta_z^{(f^*)}\boldsymbol{\Sigma}^*$ as an approximate solution to the optimization problem (18).

for some $R>0$, where $\rho\in(0,1]$.

According to Lemma 2, we know that the constraints in (12b) are satisfied if

$$\boldsymbol{r}_{b,m}^{\rm H}\boldsymbol{D}_{b,m}\boldsymbol{r}_{b,m}+2{\rm Re}\{\boldsymbol{r}_{b,m}^{\rm H}\boldsymbol{d}_{b,m}\}-c_{b,m}\geqslant 0,\ \forall\|\boldsymbol{r}_{b,m}\|^2\leqslant(R_{b,m})^2,\ \forall m\in\mathcal{M} \tag{21}$$

holds true, where $R_{b,m}>0$ such that

$$\Pr\{\|\boldsymbol{r}_{b,m}\|^2\leqslant R_{b,m}^2\}\geqslant 1-\rho_{b,m},\ \forall m\in\mathcal{M}. \tag{22}$$

We here need to determine the value of $R_{b,m}$. Note that $2\|\boldsymbol{r}_{b,m}\|^2$ are Chi-square random variables with $2N$ degrees of freedom due to the fact that $\boldsymbol{r}_{b,m}\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{I}_N)$ [30]. Let ICDF$(\cdot)$ denote the inverse cumulative distribution function of the Chi-square random variable. Then, setting $R_{b,m}=\sqrt{{\rm ICDF}(1-\rho_{b,m})/2}$ is sufficient to guarantee the constraints in (22).

Similarly, we can set $R_{e,k}^{(m)}=\sqrt{{\rm ICDF}(1-\rho_{e,k}^{(m)})/2}$, which ensures that

$$\Pr\{\|\boldsymbol{r}_{e,k}\|^2\leqslant(R_{e,k}^{(m)})^2\}\geqslant 1-\rho_{e,k}^{(m)},\ \forall m\in\mathcal{M},\ \forall k\in\mathcal{K}. \tag{23}$$

Then, according to Lemma 2, the following worst-case deterministic formulations

$$-\boldsymbol{r}_{e,k}^{\rm H}\boldsymbol{D}_{e,k}^{(m)}\boldsymbol{r}_{e,k}-2{\rm Re}\{\boldsymbol{r}_{e,k}^{\rm H}\boldsymbol{d}_{e,k}^{(m)}\}+c_{e,k}^{(m)}\geqslant 0,\ \forall\|\boldsymbol{r}_{e,k}\|^2\leqslant(R_{e,k}^{(m)})^2,\ \forall m\in\mathcal{M},\ \forall k\in\mathcal{K} \tag{24}$$

mean that the constraints in (12c) are satisfied.

Replacing (12b) and (12c) with (21) and (24), respectively, the probabilistically constrained optimization problem (12) can be recast as the one with worst-case deterministic constraints as follows:

$$\min_{\{\boldsymbol{W}_m\}_{m=1}^M,\boldsymbol{\Sigma}}\ \sum_{m=1}^M{\rm Tr}(\boldsymbol{W}_m)+{\rm Tr}(\boldsymbol{\Sigma}) \tag{25a}$$

$$\text{s.t.}\quad \boldsymbol{r}_{b,m}^{\rm H}\boldsymbol{D}_{b,m}\boldsymbol{r}_{b,m}+2{\rm Re}\{\boldsymbol{r}_{b,m}^{\rm H}\boldsymbol{d}_{b,m}\}-c_{b,m}\geqslant 0,\ \forall\|\boldsymbol{r}_{b,m}\|^2\leqslant R_{b,m}^2,\ \forall m\in\mathcal{M}, \tag{25b}$$

$$\boldsymbol{r}_{e,k}^{\rm H}\boldsymbol{D}_{e,k}^{(m)}\boldsymbol{r}_{e,k}+2{\rm Re}\{\boldsymbol{r}_{e,k}^{\rm H}\boldsymbol{d}_{e,k}^{(m)}\}-c_{e,k}^{(m)}\leqslant 0,\ \forall\|\boldsymbol{r}_{e,k}\|^2\leqslant(R_{e,k}^{(m)})^2,\ \forall m\in\mathcal{M},\ \forall k\in\mathcal{K}, \tag{25c}$$

$$\boldsymbol{\Sigma}\succeq\boldsymbol{0},\ \boldsymbol{W}_m\succeq\boldsymbol{0},\ \forall m\in\mathcal{M}. \tag{25d}$$

Although the problem (25) is convex, there are semi-infinite constraints as seen in (25b) and (25c). To make the problem more tractable, we use the $\mathcal{S}$-procedure [26] to convert the constraints in (25b) and (25c) to finitely many linear matrix inequalities (LMIs).

**Lemma 3** ($\mathcal{S}$-Procedure [26])**.** Let $\boldsymbol{G}_k(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{D}_k\boldsymbol{x} + 2\mathrm{Re}\{\boldsymbol{d}_k^{\mathrm{H}}\boldsymbol{x}\} + c_k$, where $\boldsymbol{D}_k \in \mathbb{H}^N,\ \boldsymbol{d}_k \in \mathbb{C}^N,\ c_k \in \mathbb{R},\ k = 1, 2$. The implication $\boldsymbol{G}_1(\boldsymbol{v}) \leqslant 0 \Rightarrow \boldsymbol{G}_2(\boldsymbol{v}) \leqslant 0$ holds if and only if there exists $\lambda$ such that

$$\lambda \geqslant 0,\ \lambda \begin{bmatrix} \boldsymbol{D}_1 & \boldsymbol{d}_1 \\ \boldsymbol{d}_1^{\mathrm{H}} & c_1 \end{bmatrix} - \begin{bmatrix} \boldsymbol{D}_2 & \boldsymbol{d}_2 \\ \boldsymbol{d}_2^{\mathrm{H}} & c_2 \end{bmatrix} \succeq \boldsymbol{0},$$

provided that there exists a point $\bar{\boldsymbol{x}}$ such that $\boldsymbol{G}_1(\bar{\boldsymbol{x}}) < 0$.

Applying $\mathcal{S}$-procedure to the constraints in (25b) and (25c), we can get the following equivalent LMIs

$$\begin{bmatrix} \lambda_{b,m}\boldsymbol{I}_N + \boldsymbol{D}_{b,m} & \boldsymbol{d}_{b,m} \\ \boldsymbol{d}_{b,m}^{\mathrm{H}} & -\lambda_{b,m}R_{b,m}^2 - c_{b,m} \end{bmatrix} \succeq \boldsymbol{0},\ \forall m \in \mathcal{M}, \tag{26a}$$

$$\begin{bmatrix} \lambda_{e,k}^{(m)}\boldsymbol{I}_N - \boldsymbol{D}_{e,k}^{(m)} & -\boldsymbol{d}_{e,k}^{(m)} \\ -(\boldsymbol{d}_{e,k}^{(m)})^{\mathrm{H}} & -\lambda_{e,k}^{(m)}(R_{e,k}^{(m)})^2 + c_{e,k}^{(m)} \end{bmatrix} \succeq \boldsymbol{0},\ \forall m \in \mathcal{M},\ \forall k \in \mathcal{K}, \tag{26b}$$

respectively, for some $\lambda_{b,m} \geqslant 0$ and $\lambda_{e,k}^{(m)} \geqslant 0$. Substituting (26a) and (26b) back into (25), the alternative conservative reformulation of the relaxation problem (12) can be written as

$$\min_{\{\boldsymbol{W}_m\}_{m=1}^M, \boldsymbol{\Sigma}} \ \sum_{m=1}^{M} \mathrm{Tr}(\boldsymbol{W}_m) + \mathrm{Tr}(\boldsymbol{\Sigma}) \tag{27a}$$

$$\text{s.t.} \quad \text{(26a) and (26b) satisfied,} \tag{27b}$$

$$\lambda_{b,m} \geqslant 0,\ \lambda_{e,k}^{(m)} \geqslant 0,\ \forall m \in \mathcal{M},\ \forall k \in \mathcal{K}, \tag{27c}$$

$$\boldsymbol{\Sigma} \succeq \boldsymbol{0},\ \boldsymbol{W}_m \succeq \boldsymbol{0},\ \forall m \in \mathcal{M}. \tag{27d}$$

The problem (27) is a SDP which can be efficiently solved. The Gaussian randomization procedure of gaining the feasible beamforming vectors from the non-rank-one solution $\boldsymbol{W}_m^*,\ \forall m \in \mathcal{M}$, to the problem (27) is similar to that listed in Table 1, except that the optimization problem (28) should be solved instead of problem (19) .

$$\min_{\{\beta_m^{(f)}\}_{m=1}^M, \beta_z^{(f)}} \ \sum_{m=1}^{M} \mathrm{Tr}(\boldsymbol{W}_m) + \mathrm{Tr}(\boldsymbol{\Sigma}) \tag{28a}$$

$$\text{s.t.} \quad \text{(26a) and (26b) satisfied,} \tag{28b}$$

$$\lambda_{b,m} \geqslant 0,\ \lambda_{e,k}^{(m)} \geqslant 0,\ \forall m \in \mathcal{M},\ \forall k \in \mathcal{K}. \tag{28c}$$

### 3.3 Complexity analysis

In this subsection, the complexity of various methods will be discussed. The complexity is mainly dominated by the part of handling the problems (18) and (27). According to [31], the interior-point methods will take $\mathcal{O}(\sqrt{\theta}\log(1/\epsilon))$ iterations to find an optimal solution, where $\theta$ represents the barrier parameter and $\epsilon$ is the accuracy of the solution. Let $\mathcal{O}(\mathcal{L})$ denote the arithmetic cost of each iteration. Then, the worst-case complexity of solving a given optimization problem is $\mathcal{O}(\sqrt{\theta}\mathcal{L}\log(1/\epsilon))$. We first determine the value of $\theta$. From [31], for the mixed second-order cone programming (SOCP) and SDP problem, the arithmetic barrier parameter is

$$\theta = \sum_{i=1}^{m_{\mathrm{sdp}}} k_{i,\mathrm{sdp}} + 2m_{\mathrm{socp}}, \tag{29}$$

where $m_{\mathrm{sdp}}$ is the number of the PSD constraints, $k_{i,\mathrm{sdp}}$ is the dimension of the $i$th semidefinite cone, and $m_{\mathrm{socp}}$ is the number of the second-order cone (SOC) constraints. The arithmetic cost of each iteration, including the operations of assembling and factorizing the system matrix, is on the order of

$$\mathcal{L} = \underbrace{n\sum_{i=1}^{m_{\mathrm{sdp}}} k_{i,\mathrm{sdp}}^3 + n^2\sum_{i=1}^{m_{\mathrm{sdp}}} k_{i,\mathrm{sdp}}^2 + n\sum_{i=1}^{m_{\mathrm{socp}}} k_{i,\mathrm{socp}}^2}_{\text{assembling cost}} + \underbrace{n^3}_{\text{factorizing cost}}, \tag{30}$$

**Table 2** Complexity comparison of various methods

| Method | Variables (size,number) | PSD constraints (size,number) | SOC constraints (size,number) | Barrier parameter $\theta$ |
|---|---|---|---|---|
| Bernstein | $(N, M+1)$<br>$(1, 2M(K+1))$ | $(N, MK+2M+1)$<br>$(1, 2M(K+1))$ | $(N^2+N+1, M(K+1))$ | $M(KN+2N+4K+4)+N$ |
| Worst-case | $(N, M+1)$<br>$(1, M(K+1))$ | $(N+1, M(K+1))$<br>$(N, M+1)$<br>$(1, M(K+1))$ | 0 | $M(KN+2N+2K+2)+N$ |
| Bernstein isotropic AN | $(N, M)$<br>$(1, 2M(K+1)+1)$ | $(N, MK+2M)$<br>$(1, 2M(K+1)+1)$ | $(N^2+N+1, M(K+1))$ | $M(KN+2N+4K+4)+1$ |
| Worst-case isotropic AN | $(N, M)$<br>$(1, M(K+1)+1)$ | $(N+1, M(K+1))$<br>$(N, M)$<br>$(1, M(K+1)+1)$ | 0 | $M(KN+2N+2K+2)+1$ |

where $n$ is the number of the decision variables and $k_{i,\text{socp}}$ is the dimension of the $i$th SOC.

The method discussed in Subsection 3.2.1 and the one in Subsection 3.2.2 are, respectively, referred to as the "Bernstein" method and the "Worst-case" method. We compare the proposed methods with the following ones: (1) the "Bernstein isotropic AN" method, which is the same as the "Bernstein" method, except that the AN is imposed in the null space of estimated main channels, called the isotropic AN [16, 23]; (2) the "Worst-case isotropic AN" method, which is the same as the "Worst-case" method, except that the AN is isotropic. Table 2 shows the size and number of both variables and constraints for various methods. Note that (1) the scalar linear constraints can be regarded as the PSD constraints of size 1; (2) the structure of the AN is fixed and only the AN power needs to be optimized while using the methods with isotropic AN. Also, as can be seen from the second column in Table 2, the number of decision variables $n$ of each method is on the order of $MN^2$. For large $N$, $M$, and $K$, all the methods have the same dominating terms in (30). Hence, we just compare the barrier parameters listed in the fifth column in Table 2. From Table 2, it can be seen that "Bernstein" method has the highest complexity, followed by "Bernstein isotropic AN" method, "Worst-case" method, and "Worst-case isotropic AN" method in sequence. Interestingly, in Section 4, we will find that this ranking is the same as that of the performances of various methods.

## 4 Simulation results

In this section, we carry out some simulations to evaluate the performances of the proposed methods. In each simulation, we assume the number of transmit antennas at Alice is $N = 5$, the number of Bobs is $M = 3$, and the number of Eves is $K = 2$. It is noteworthy that the relation $N > M$ should hold true for assuring the existence of the null space of the legitimate channels for the methods with isotropic AN [16, 23]. All the channel estimates are independent identically distributed Rayleigh flat fading, that is, $\hat{\boldsymbol{h}}_m \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_5)$, $\hat{\boldsymbol{g}}_k \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_5)$, $\forall m, k$. For simplicity, the CSI error vectors of Bobs have identical covariances, that is, $\boldsymbol{Q}_{b,m} = \varepsilon_b \boldsymbol{I}_N = 0.002\boldsymbol{I}_5$, $\forall m$. Similarly, $\boldsymbol{Q}_{e,k} = \varepsilon_e \boldsymbol{I}_N = 0.005\boldsymbol{I}_5, \forall k$, unless specified otherwise. Without loss of generality, we have the following parameter settings. The noise variances of both Bobs and Eves are assumed to be the same, that is, $\sigma_{b,m}^2 = \sigma_{e,k}^2 = \sigma^2 = 0.1$, $\forall m, k$. The predefined SINR thresholds of Bobs are set to be the same, that is, $\gamma_{b,m} = \gamma_b$, $\forall m$. We assume that $\gamma_{e,k}^{(m)} = \gamma_e = -5$ dB, $\forall m, k$. The probability values at both Bobs and Eves are set as $\rho_{b,m} = \rho_b$, $\forall m$ and $\rho_{e,k}^{(m)} = \rho_e, \forall m, k$.

In the simulations, the SDPs are solved using the optimization solver CVX [28]. We compare the performances of the following methods: the "Bernstein" method, the "Worst-case" method, the "Bernstein isotropic AN" method, the "Worst-case isotropic AN" method, and the "Non-robust" method which takes the estimated CSI as the true CSI.

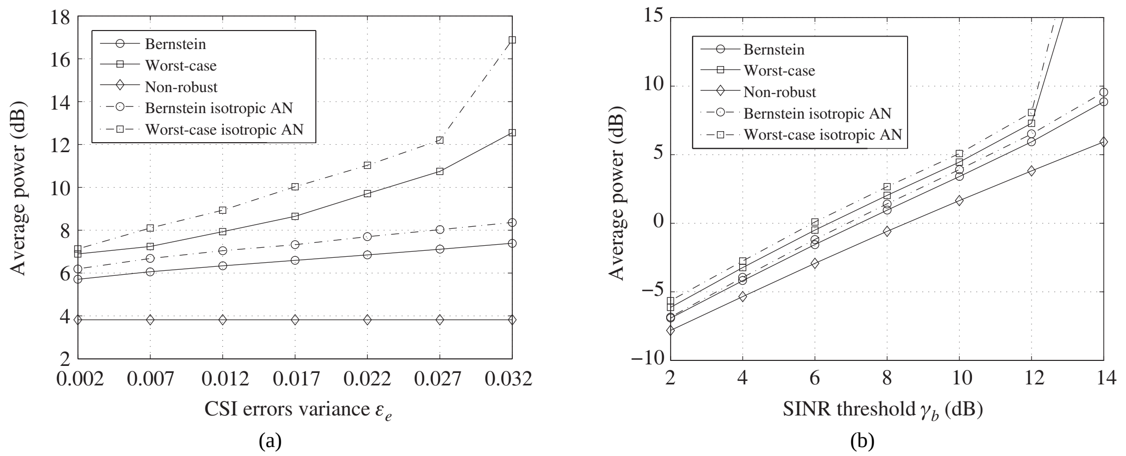


**Figure 1** Average power versus (a) the CSI errors variance of Eves with $\gamma_b = 12$ dB, $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\rho_b = \rho_e = 0.05$ and (b) the SINR threshold of Bobs with $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\varepsilon_e = 0.005$, $\rho_b = \rho_e = 0.05$.

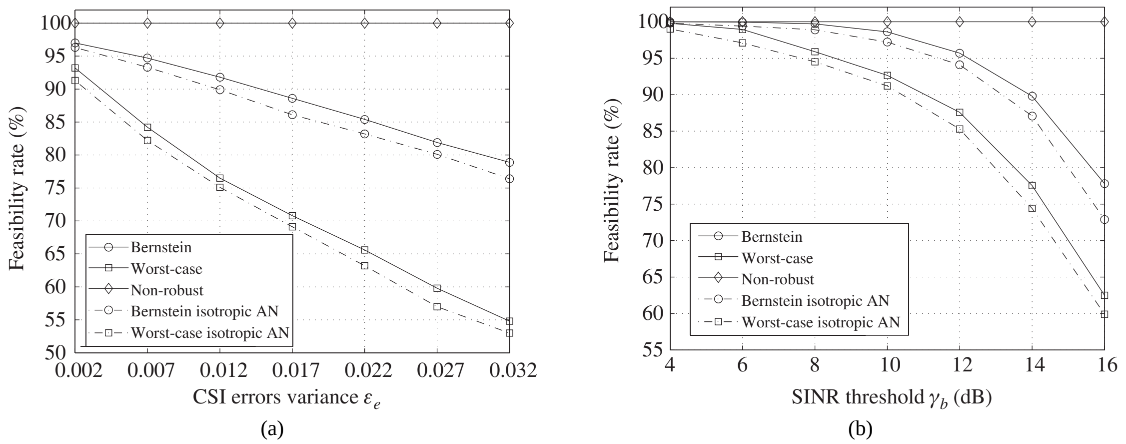


**Figure 2** Feasibility rate versus (a) the CSI errors variance of Eves with $\gamma_b = 12$ dB, $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\rho_b = \rho_e = 0.05$ and (b) the SINR threshold of Bobs with $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\varepsilon_e = 0.005$, $\rho_b = \rho_e = 0.05$.

The impact of the CSI errors variance $\varepsilon_e$ on the power consumptions of various methods is presented in Figure 1(a) with $\gamma_b = 12$ dB. The result was obtained based on 500 channel realizations for which all methods yield feasible solutions at $\varepsilon_e = 0.032$ and $\gamma_b = 12$ dB. It can be seen that as $\varepsilon_e$ increases from 0.002 to 0.032, the average powers consumed by various methods increase correspondingly. This phenomenon can be understood easily. As the channel errors variance $\varepsilon_e$ increases, more power should be allocated to the AN for keeping the probability that the SINR of Eves exceeds the predefined threshold below the maximum allowable probability value. We can see that the performances of "Bernstein" and "Bernstein isotropic AN" methods outperform the ones of the "Worst-case" and "Worst-case isotropic AN" methods in terms of the power consumption. This is due to the fact that the reformulations using the Bernstein-type inequalities are less conservative than using the $\mathcal{S}$-procedure. Also, we find that the power consumption of the "Bernstein" method is less than that of the "Bernstein isotropic AN" method under the same parameter settings. This verifies that the pattern of AN that we used is more efficient in jamming the eavesdroppers than the isotropic AN. Besides, we find that the power consumption of "Non-robust" method is constant. This is because this method does not take the channel error into consideration. Hence, the reliability (i.e., the constraints in (5b)) and security (i.e., the constraints in (5c)) are not assured, which will be seen in Figure 4.

Figure 1(b) depicts the average power transmitted at Alice versus the SINR threshold of Bobs with fixed CSI error variances, that is, $\varepsilon_b = 0.002$ and $\varepsilon_e = 0.005$, for various methods. The result was obtained based on the same channel realizations used in the last result in Figure 1(a). From Figure 1(b), we can

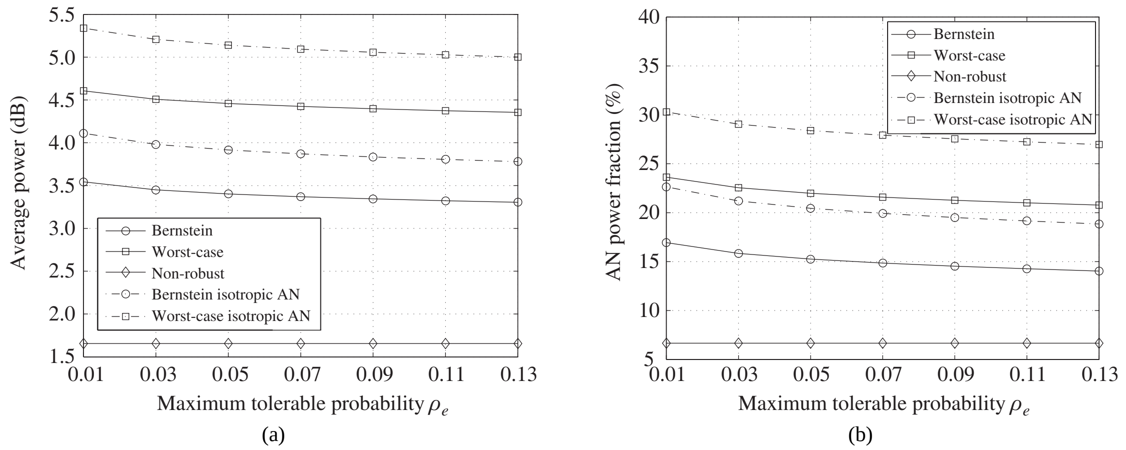


**Figure 3** (a) Average power versus maximum tolerable probability $\rho_e$ and (b) AN power fraction versus maximum tolerable probability $\rho_e$ with $\gamma_b = 10$ dB, $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\varepsilon_e = 0.005$, $\rho_b = 0.05$.

observe that the average power is an increasing function of the SINR threshold of Bobs. This is because more information-bearing power is necessary for keeping the SINR outage probability below the preset maximum allowable probability value as the SINR threshold of Bobs increases. Also, the comparison analysis of power performances between different methods is similar to that in Figure 1(a). The reason has been explored above. Note that the average power is set as infinity when the method is infeasible for at least one channel realization. The results of the "Non-robust" method are plotted as a benchmark to show how much additional power is necessary to assure the robustness.

Figure 2 (a) and (b) show the feasibility rates of various methods versus the channel errors variance $\varepsilon_e$ of Eves and the predefined SINR threshold $\gamma_b$ of Bobs, respectively. The parameter settings in Figure 2 (a) and (b) are the same as those in Figure 1 (a) and (b), respectively. It can be observed that the feasibility rates of robust methods decrease with the increasing channel errors variance $\varepsilon_e$ of Eves and the increasing SINR threshold of Bobs , while that of the "Non-robust" method is always one. Of all the robust methods, the "Bernstein" method shows the best feasibility rate performance. This confirms that the reformulation using the Bernstein-type inequality is less conservative. Also, the advantage of the non-isotropic AN over the isotropic AN is shown again when the same reformulation is adopted.

In Figure 3(a), we demonstrate how the Eves' maximum tolerable probability $\rho_e$ affects the average transmit power consumed by various methods with $\gamma_b = 10$ dB, $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\varepsilon_e = 0.005$, $\rho_b = 0.05$. As can be seen from Figure 3(a), the average transmit power decreases with the increasing maximum tolerable probability $\rho_e$, and the performance ranking is consistent with the analyses in Figures 1 and 2. To get more insights, in Figure 3(b) we plot the AN power fraction versus the maximum tolerable probability $\rho_e$. The parameter settings are the same as those in Figure 3(a). It can be found that the power fraction allocated to AN decreases with increasing maximum tolerable probability $\rho_e$. This is due to the fact that less AN power is needed to confuse Eves when the security requirements degrade. Also, it can be seen that more isotropic AN power is allocated than the non-isotropic AN power under the same conditions. This is because the pattern of the non-isotropic AN is designed by considering the estimated CSI of both Bobs and Eves, while that of the isotropic AN is fixed in the null space of the channels from Alice to Bobs.

To gain more insights into the impact of imperfect CSI on the reliability and the security of the system, the cumulative distribution functions (CDFs) of the SINR of the first Bob and the SINR of the first Eve eavesdropping the first data stream are plotted in Figure 4 (a) and (b), respectively. The SINR thresholds of Bobs and Eves are set as $\gamma_b = 8$ dB, $\gamma_e = -5$ dB, and the channel error variances are set as $\varepsilon_b = 0.002$, $\varepsilon_e = 0.005$. From Figure 4(a), we can find that at the threshold SINR = 8 dB the CDF values of robust methods are less than the predefined probability value $\rho_b = 0.05$, and close to 0. This means that nearly all the channel realizations satisfy the QoS constraints at Bobs. The reason behind this phenomenon is that all the robust methods are conservative. It can also be seen that the SINR of the "Non-robust"

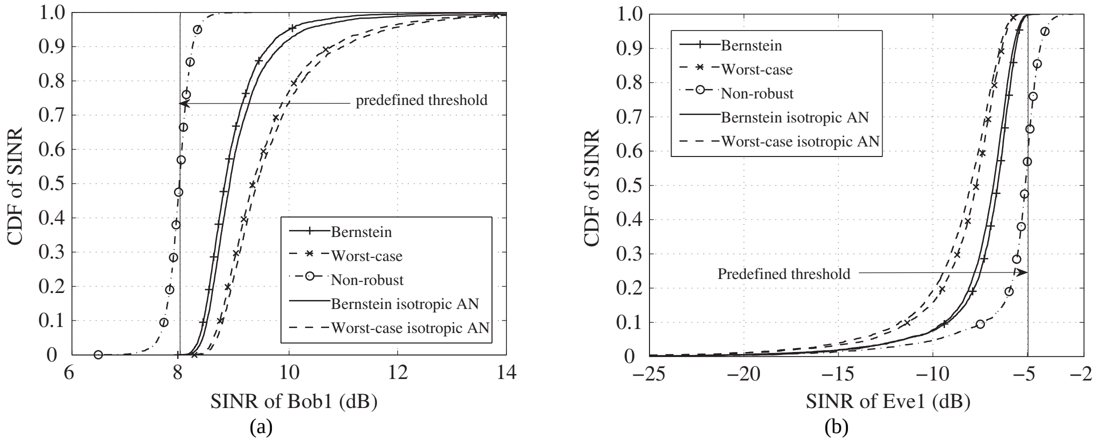


**Figure 4** CDF of achieved SINR (a) of the first Bob and (b) of the first Eve eavesdropping the first data stream. Both (a) and (b) have common parameters with $\gamma_b = 8$ dB, $\gamma_e = -5$ dB, $\varepsilon_b = 0.002$, $\varepsilon_e = 0.005$, $\rho_b = \rho_e = 0.05$.

method is below the SINR threshold for about 60% of the channel realizations. As can be seen from Figure 4(b), the probability that the first Eve's SINR exceeds the threshold value is kept below $\rho_e = 0.05$ when using the robust methods, while the QoS constraint is violated with the probability 40% for the "Non-robust" method. It can be concluded that although the "Non-robust" method consumes less power than the robust methods as shown in Figure 1, the former has the poorest performance in terms of the reliability and the security. This is because the "Non-robust" method ignores the channel error during the optimization. Both Figure 4 (a) and (b) show that the SINR of the "Bernstein" method is more closer to the predefined threshold compared with other robust methods and thus consumes less power. This is consistent with the results shown in Figure 1.

## 5 Conclusion

In this paper, we developed a robust AN-aided beamforming scheme for the unicast multiuser MISO downlink systems in the presence of multiple single-antenna eavesdroppers. We focus on minimizing the total transmit power subject to the probabilistic QoS constraints at the legitimate users and the eavesdroppers by optimizing the beamforming vectors and the AN covariance matrix jointly. After applying the SDR technique, the probabilistic QoS constraints are recast as two different deterministic ones using the Bernstein-type inequality and the $\mathcal{S}$-procedure. The resulting optimization problems are solved using interior-point methods. Simulations were carried out to show that the proposed robust AN-aided beamforming methods outperform a non-robust method and the ones using the isotropic AN.

**Acknowledgements** This work was supported in part by National Natural Science Foundation of China (Grant Nos. 61171108, 61401510, 61379006), and National High-tech R&D Program of China (863) (Grant No. SS2015AA011306).

**Conflict of interest** The authors declare that they have no conflict of interest.

### References


1 Jiang S Q. On $\tau$-time secure key agreement. Sci China Inf Sci, 2015, 58: 012110
2 Ni L, Chen G L, Li J H, et al. Strongly secure identity-based authenticated key agreement protocols in the escrow mode. Sci China Inf Sci, 2013, 56: 082113
3 Wyner A D. The wire-tap channel. Bell Syst Tech J, 1975, 54: 1355–1387
4 Zhang Z, Zhang W, Tellambura C. MIMO-OFDM channel estimation in the presence of frequency offsets. IEEE Trans Wirel Commun, 2008, 7: 2329–2339
5 Li Q, Ma W K. Spatially selective artificial-noise aided transmit optimization for MISO multi-Eves secrecy rate maximization. IEEE Trans Signal Process, 2013, 61: 2704–2717

6 Wang B, Mu P C, Yang P Z, et al. Two-step transmission with artificial noise for secure wireless SIMO communications. Sci China Inf Sci, 2015, 58: 042308

7 Wang H M, Liu F, Xia X G. Joint source-relay precoding and power allocation for secure amplify-and-forward MIMO relay networks. IEEE Trans Inf Forens Secur, 2014, 9: 1240–1250

8 Liu L, Zhang R, Chua K C. Secrecy wireless information and power transfer with MISO beamforming. IEEE Trans Signal Process, 2014, 62: 1850–1863

9 Wang J, Bengtsson M, Ottersten B, et al. Robust MIMO precoding for several classes of channel uncertainty. IEEE Trans Signal Process, 2013, 61: 3056–3070

10 Tang Y, Xiong J, Ma D, et al. Robust artificial noise aided transmit design for MISO wiretap channels with channel uncertainty. IEEE Commun Lett, 2013, 17: 2096–2099

11 Li J, Petropulu A P. Explicit solution of worst-case secrecy rate for MISO wiretap channels with spherical uncertainty. IEEE Trans Signal Process, 2012, 60: 3892–3895

12 Guan X, Cai Y, Yang W. On the reliability-security tradeoff and secrecy throughput in cooperative ARQ. IEEE Commun Lett, 2014, 18: 479–482

13 Cai Y, Guan X, Yang W. Secure transmission design and performance analysis for cooperation exploring outdated CSI. IEEE Commun Lett, 2014, 18: 1637–1640

14 Wang L, Yang N, Elkashlan M, et al. Physical layer security of maximal ratio combining in two-wave with diffuse power fading channels. IEEE Trans Inf Forens Secur, 2014, 9: 247–258

15 Li Q, Ma W K, So A M C. A safe approximation approach to secrecy outage design for MIMO wiretap channels. IEEE Signal Process Lett, 2014, 21: 118–121

16 Pei M Y, Wei J B, Wong K K, et al. Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI. IEEE Trans Wirel Commun, 2012, 11: 544–549

17 Lin P H, Lai S H, Lin S C, et al. On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels. IEEE J Sel Area Commun, 2013, 31: 1728–1740

18 Liao W C, Chang T H, Ma W K, et al. QoS-based transmit beamforming in the presence of eavesdroppers: an optimized artificial-noise-aided approach. IEEE Trans Signal Process, 2011, 59: 1202–1216

19 Romero-Zurita N, Ghogho M, McLernon DC. Outage probability based power distribution between data and artificial noise for physical layer security. IEEE Signal Process Lett, 2012, 19: 71–74

20 Fei Z S, Ni J Q, Zhao D, et al. Ergodic secrecy rate of two-user MISO interference channels with statistical CSI. Sci China Inf Sci, 2014, 57: 102302

21 Xu Y, Wang J, Wu Q. Effective capacity region of two-user opportunistic spectrum access. Sci China Inf Sci, 2011, 54: 1928–1937

22 Goel S, Negi R. Guaranteeing secrecy using artificial noise. IEEE Trans Wirel Commun, 2008, 7: 2180–2189

23 Zhou X, McKay M R. Secure transmission with artificial noise over fading channels: achievable rate and optimal power allocation. IEEE Trans Veh Technol, 2010, 59: 3831–3842

24 Luo Z Q, Ma W K, So A C, et al. Semidefinite relaxation of quadratic optimization problems. IEEE Signal Process Mag, 2010, 27: 20–34

25 Ma S, Sun D. Chance constrained robust beamforming in cognitive radio networks. IEEE Commun Lett, 2013, 17: 67–70

26 Boyd S, Vandenberghe L. Convex Optimization. Cambridge: Cambridge University Press, 2004

27 Li Q, Ma W K, So A C. Safe convex approximation to outage-based MISO secrecy rate optimization under imperfect CSI and with artificial noise. In: Proceedings of 2011 Conference Record of the 45th Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, 2011. 207–211

28 Grant M, Boyd S. CVX: Matlab software for disciplined convex programming. version 2.0. http://cvxr.com/cvx, 2012

29 Wang K Y, Chang T H, Ma W K, et al. A semidefinite relaxation based conservative approach to robust transmit beamforming with probabilistic SINR constraints. In: Proceedings of European Signal Processing Conference (EUSIPCO), Aalborg, 2010. 23–27

30 Leon-Garcia A. Probability, Statistics, and Random Processes for Electrical Engineering. Upper Saddle River: Pearson Education Inc., 2008

31 Ben-Tal A, Nemirovski A. Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. Philadelphia: SIAM, 2001