A Riemannian Distance For Robust Downlink Beamforming

A RIEMANNIAN DISTANCE FOR ROBUST DOWNLINK BEAMFORMING

BY

LIJIN XU, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

© Copyright by Lijin Xu, September 2013

All Rights Reserved

Master of Applied Science (2013) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	A Riemannian Distance For Robust Downlink Beamforming
AUTHOR:	Lijin Xu B.Sc., (Electrical Engineering) Beijing University of Posts and Telecommunications, Beijing, China
SUPERVISOR:	Dr. K. M. Wong
NUMBER OF PAGES:	xi, 83

Dedications

To my parents and my friends

Abstract

We examine the robust downlink beamforming design from the point of outage probability constraint. We further reason that since the estimated downlink channel correlation (DCC) matrices form a manifold in the signal space, the estimation error should be measured in terms of Riemannian distance (RD) instead of the commonly used Euclidean distance (ED). Applying this concept of measure to our design constraint, we establish approximated outage probability constraints using multidimensional ball set and multidimensional cube set. We transform the design problem into a convex optimization problem which can be solved efficiently by standard methods. Our proposed methods apply to both Gaussian distribution assumption and uniform distribution assumption. Simulation results show that the performance of our design is superior to those of other robust beamformers recently developed.

Acknowledgements

I would like to express my sincere gratitude to all those people who gave me great help to complete this thesis. First and foremost, I deeply appreciate to my supervisor Dr. Kon Max Wong for his ingenious commitment, encouragement and valuable advice. I would like to thank Dr. Jian-Kang Zhang for his endless patience and helpful discussions, suggestions and guidance during my study. In addition, I would like to thank Dr. Tim Davidson, Dr. Jun Chen for their help in my research and study. I would also like to thank my lab mates and staff members at McMaster University who have contributed a lot to providing me with a great research atmosphere.

Finally, I wish to thank my friends who have always been a constant source of inspiration to me.

Notation and abbreviations

Notation

$\mathbf{A}$	matrix
$\mathbf{a}$	column vector
$(\cdot)^*$	the conjugate operator
$(\cdot)^T$	the transpose of a vector or a matrix
$(\cdot)^H$	the Hermitian transpose of a vector or a matrix
$(\cdot)^{-1}$	the reciprocal or the inversion operation
$(\hat{\cdot})$	the estimate of a parameter
$(\cdot)_j$	the jth element of a vector
$[\cdot]_{ij}$	the ijth element of a matrix
$\lvert \cdot \rvert$	the magnitude of a complex quantity
$\lVert \cdot \rVert$	the Euclidean norm of a vector or a matrix
$\otimes$	the Kronecker product
$\mathbb{E}\{\cdot\}$	the statistical expectation operator
$\mathrm{Tr}\{\cdot\}$	the trace of a matrix
$\mathrm{vec}\{\cdot\}$	the operator stacking the columns of a matrix on top of each other
$\mathcal{M}$	manifold

$\mathcal{B}$,$\mathcal{C}$	set
λ	eigenvalue
$\mathbf{I}$	identity matrix
$\mathbb{R}$	the field of real numbers
$\mathbb{C}$	the field of complex numbers
$\mathrm{p}(\cdot)$	the probability density function
$\mathrm{Pr}(\cdot)$	the probability operator
$\mathrm{erf}(\cdot)$	the error function
$\mathrm{Re}\{\cdot\}$	the real part of a vector or a matrix
$\mathrm{Im}\{\cdot\}$	the imaginary part of a vector or a matrix

Abbreviation

BS	Base Station
QoS	Quality of Service
CSI	Channel State Information
DCC	Downlink Channel Correlation
SDP	Semidefinite Program
MSE	Mean-square Error
SINR	Signal to Interference Noise Ratio
ED	Euclidean Distance
PSD	Positive Semidefinite
RD	Riemannian Distance
SOCP	Second-roder Cone Program
SDR	Semidefinite Relaxation
LMI	Linear Matrix Inequality

Contents

Abstract **iv**

Acknowledgements **v**

Notation and abbreviations **vi**

1 Introduction **1**

1.1 Overview of beamforming 1

1.2 Contribution of our work 5

1.3 Structure of the thesis 6

2 Beamforming methods **8**

2.1 Non-robust beamforming method 8

2.2 Robust beamforming method with absolute error bound 11

2.3 Probabilistic robust beamforming method 19

3 Probabilistic robust beamforming with Riemannian Distance **30**

3.1 Introduction to Riemannian distance 31

3.2 Worst-case robust beamforming method with Riemannian distance . . 32

3.3 Probabilistic robust beamforming problem formulation 36

3.4 Reformulation of design constraint 40
3.4.1 Application of Lemma 1 40
3.4.2 Choice of set $\mathcal{B}$ 42
3.4.3 Gaussian Assumption for vec($\tilde{\mathbf{\Delta}}_k$) 42
3.5 Convex design problem 46
3.6 Comparison of absolute error bound approach and probabilistic of outage approach 49

4 Alternative formulation for probabilistic robust beamforming 51
4.1 Approximation of set $\mathcal{B}$ 51
4.2 Uniform Distribution Assumption for vec($\tilde{\mathbf{\Delta}}_k$) 54
4.3 Reformulation 56
4.4 Convex design problem 57

5 Numerical Experiments 62
5.1 Simulations of Methods in Chapter 3 64
5.2 Simulations of Methods in Chapter 4 75

6 Conclusion and Future work 78

List of Figures

1.1 Illustration of downlink beamforming 2

2.1 Total transmitted power versus required SINR [13] 15

2.2 Total transmitted power versus SINR [14] 19

2.3 Feasibility and transmit power performance of the various methods [20] 29

5.1 Direction of arrival 63

5.2 Variation of transmission power with the separation angle, $\sigma_e^2 = 0.02$. 65

5.3 Histogram of weighted SINR for $\phi = 7°$, $\sigma_e^2 = 0.02$ 66

5.4 Variation of transmission power with QoS requirement, $\sigma_e^2 = 0.02$ . . 67

5.5 Variation of transmission power with the separation angle, $\sigma_e^2 = 0.025$ 68

5.6 Histogram of weighted SINR for $\phi = 7°$, $\sigma_e^2 = 0.025$ 69

5.7 Variation of transmission power with QoS requirement, $\sigma_e^2 = 0.025$ . . 70

5.8 Variation of transmission power with QoS requirement 72

5.9 Variation of transmission power with the separation angle, correlated error matrix 73

5.10 Variation of transmission power with QoS requirement, correlated error matrix 74

5.11 Comparison between $\mathcal{B}_k$ and $\mathbf{C}_k$ 76

5.12 Power variation based on uniform distribution 77

Chapter 1

Introduction

1.1 Overview of beamforming

In a wireless communication system, quality of service and capacity are of great importance. To ensure the communication reliability for each user, a system need to overcome interference and multipath fading. Antenna arrays can improve reliability and capacity. A beamformer is a processor with an array of sensors to provide a versatile form of spatial filtering. As shown in Fig. 1.1, the base station (BS) is equipped with multiple antennas so that beamforming can be carried out providing a satisfactory quality of service (QoS) to each user [1]. To change the directionality of the array when transmitting, a beamformer controls the phase and relative amplitude of the signal at each transmitter. Each beampattern has a mainlobe whose direction is designed to face the corresponding user so that the SINR can be maximized by means of enhancing the desired signal power and rejecting the interferences by beampattern nulls.

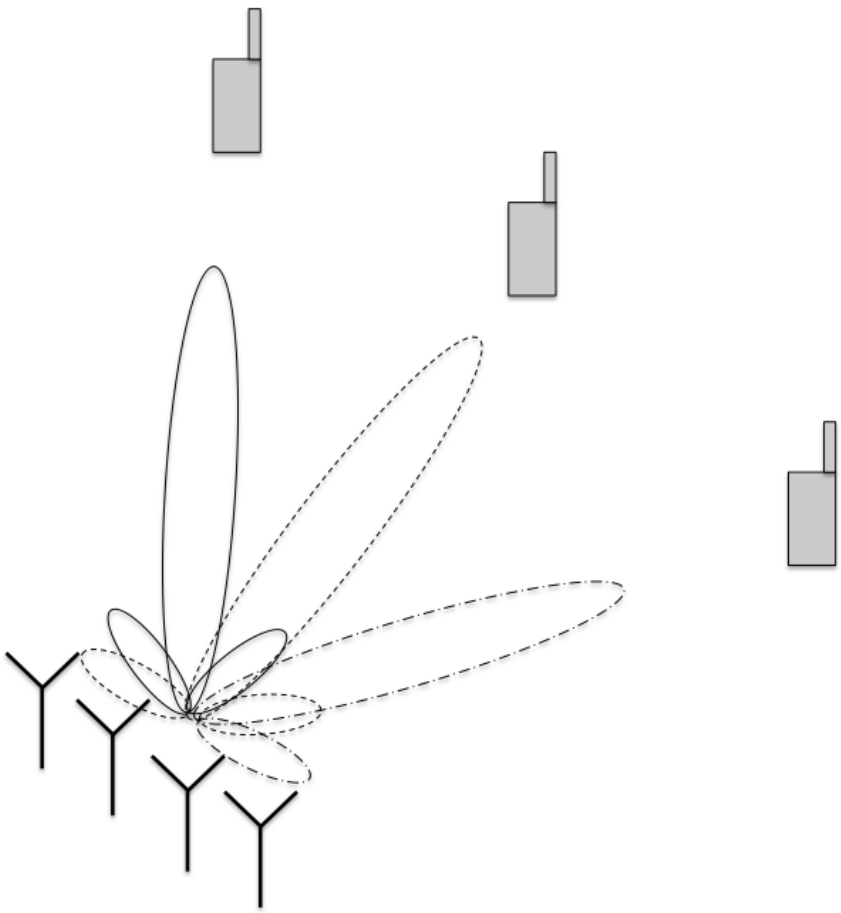


Figure 1.1: Illustration of downlink beamforming

Beamforming usually assumes that the second order statistics of the channel (represented by the downlink channel correlation matrices) are exactly known. Methods based on the fully known channel information are called non-robust beamforming methods. This assumption is difficult to satisfy in practice because the knowledge of the downlink channel correlation (DCC) matrices depends on the accuracy of the channel state information (CSI) available at the transmitter which, in turn, depends on the channel estimation errors caused by channel variability, array calibration, etc., as well as feedback quantization errors and feedback delay resulting in serious degradation in the performance of such techniques [2]−[5].

In recent years, consideration of the imperfection of CSI gave rise to the design algorithms of *robust beamforming* which is a class of beamforming techniques aiming to ensure the satisfaction of the QoS requirements of the users in the case of CSI mismatch. Various CSI mismatch models have been considered that lead to different robust beamforming approaches. In the worst case approach the concept is to design beamformers such that the QoS constraints are satisfied for all the channels defined in a bounded uncertainty set around the presumed CSI which is often the instantaneous channel estimates or the estimated second-order statistics of the channels [6]−[13]. In [6], three conservative design approaches that yield convex and computationally-efficient restrictions of the original design problem are derived. The three approaches yield semidefinite program (SDP) formulations that offer different trade-offs between the degree of conservatism and the size of the SDP. In [7] and [8], a MIMO system is considered under error uncertainties where an absolute error bound is given. In [7], the errors of the channel are bounded within an ellipsoid and a convex reformulation can be derived as semidefinte programs. The result shows that the optimal transmission

directions are just the right singular vectors of the nominal channel under some mild conditions. And this result reduces the complicated matrix-valued problem to a scalar power allocation problem. A convergent iterative algorithm is proposed in [8] to obtain the robust solution for multiuser MIMO systems with imperfect CSI at the receiver to guarantee the worst-case SINR. The positive semi-definiteness of the mismatched channel covariance matrices are taken into account in [12]. In [13], an exact representation of the worst-case solution is obtained using Lagrange duality.

Others developed algorithms by expressing QoS constraints in terms of the mean-square errors (MSE) instead of SINR [15]−[17]. When the channel vectors are exactly know, an MSE constraint is equivalent to an SINR constraint. Under the consideration of partial CSI, MSE constraints provide a conservative lower bound on the SINR. Such formulations, in general, offer no guarantee on outage performance.

A different view to robust beamforming is motivated by the random nature of signal-to-interference-noise ratio (SINR) caused by the channel mismatch. This led to the probabilistic approach which guarantees the QoS for the users with a predefined non-outage probability and which relies on the statistical models of the channel mismatches [18, 19, 20]. The difficulty of the problem stands in the fact that in general there are no closed form expressions for the probabilistic SINR constraints. The mismatch matrix between the real and estimated channel covariance matrix is assumed to have Gaussian distribution in [18] and the initially non-convex optimization problem is reformulated into tractable convex optimization problem. A relaxation-restriction approach is proposed to approximate the outage probability constraint optimization problem in [20] and this approach shows a significant improvement in computational complexity. Under different assumptions of CSI error distributions,

convex approximations to the initial robust beamforming problem were derived.

1.2 Contribution of our work

In the above developments of beamforming algorithms, the treatment of the errors, especially the errors in the estimation of the channels or the errors in the estimation of the DCC matrices is usually in terms of Euclidean distance (ED) (The equivalent inner product norm for matrices is often called the Frobenius norm.) While the use of the ED is generally well motivated if the mismatch is applied on the instantaneous channel estimates, it is generally not appropriate for modeling the mismatch on the DCC. Our reason is that since DCC matrices are not freely structured, but are *Hermitian* and *positive semi-definite* (PSD), therefore, they form a *manifold* in the signal space. Thus, instead of using the Frobenius norm, the distance between the true and the estimated DCC matrices $\mathbf{R}$ and $\hat{\mathbf{R}}$, should be measured by the *Riemannian distance* (RD) along the surface of the manifold [21]. This concept is akin to finding the distance between two cities on earth: The Euclidean distance between two cities is neither informative nor accurate.

The application of RD to robust beamforming design was proposed in [22] in which a worst-case approach to downlink robust beamforming with covariance based CSI feedback at the transmitter was taken. It has been shown that bounding the uncertainty set of the DCC mismatch with a RD yields a convenient convex reformulation and significantly improved beamforming performance as compared to previous approaches based on the Frobenius norm. In this thesis, we also employ the RD for measuring the estimation error in the DCC matrices. However, we take the outage probabilistic approach in robust downlink beamforming. The goal of our robust

beamforming design is to attain an SINR non-outage probability larger than a certain imposed threshold for each receiver in the system. To facilitate a solution, we define a convex restriction to the beamforming problem such that the probabilistic non-outage constraints are satisfied for all the errors in this set. We show that this goal can be achieved by imposing that the errors in the DCC matrices are bounded based on the RD.

During the process of converting the non-convex problem into convex one, the relaxation-restriction approach is used to approximate the original problem. The error model is established under both Gaussian distribution assumption and uniform distribution assumption. We further derive the relation between the SINR outage threshold and the RD-based bound on the covariance mismatch set. In both cases, we can find multidimensional ball set and multidimensional cube set to help turn the non-deterministic outage probability constraint into a deterministic one. The results of the reformulation show that different error matrix distribution assumptions have different effects on the computation complexity of finding the corresponding alternative set. However, the difference of the formation of the convex reformulation only depends on the expression of the different set. Simulations show the improved performance of our approach in terms of transmission power and beamformer accuracy with respect to previous methods.

1.3 Structure of the thesis

In this thesis, we propose a convex design of outage probabilistic robust downlink beamforming using RD. In Chapter 1, the background knowledge of beamforming

and the contribution of our work are introduced. Then, in Chapter 2, we introduce some existing beamforming methods in details. These existing methods can be classified into non-robust beamforming methods, robust beamforming methods with absolute error bound and probabilistic robust beamforming methods. Several popular and classical approaches are discussed in details in each section of Chapter 2. Our proposed method using RD as an error rmeasure is introduced in Chapter 3. A brief introduction of RD is provided and then a discussion on the derivation of our problem and the comparison with the worst-case approach is presented. An important step in the process to turn our problem into a convex one is to find a set such that the problem can be restricted into a deterministic one and one way of finding such a set is given in Chapter 3. However, due to the fact that there are many choices of such a set, an alternative formulation for probabilistic robust beamforming based on a different set is discussed in Chapter 4. We also make another distribution assumption and derive the corresponding convex reformulation in this chapter. Numerical results are shown in Chapter 5, where the simulations of Chapter 3 and Chapter 4 demonstrate the better performance of our proposed robust beamforming method compared to other referred methods. Finally, conclusion of the thesis and our future work are presented in Chapter 6.

Chapter 2

Beamforming methods

Beamforming methods can be applied to both the transmitter and the receiver of a wireless system. In this thesis, we focus our attention on downlink beamforming. This technique is relatively new. For downlink beamforming methods, the goal of design beamforming vectors is to make sure that each user would satisfy a predefined QoS threshold and at the same time, that the transmission power is as small as possible. In this chapter, we will introduce three kinds of well-known downlink beamforming methods, non-robust beamforming, robust beamforming with absolute error bound and a proposed probabilistic robust beamforming.

2.1 Non-robust beamforming method

For simplicity, we consider that there is one single base station equipped with N antennas serving K users. Each user has single antenna. The transmitted signal at BS can be expressed as

$$\mathbf{x}(t) = \sum_{k=1}^{K} \mathbf{w}_k s_k(t) \tag{2.1}$$

where $\mathbf{w}_k \in \mathbb{C}^N$ is the beamformer vector and $s_k(t)$ is the data stream transmitted for User k. Hence, the received signal for the kth user is given by

$$y_k(t) = \mathbf{h}_k^H \mathbf{x}(t) + n_k(t) \tag{2.2}$$

where $\mathbf{h}_k \in \mathbb{C}^N$ denotes the downlink channel vector from BS to User k, and $n_k(t)$ is the additive noise with variance σ_k^2.

In non-robust beamforming, we assume there is no error in the estimation of the channel vector or the channel covariance matrix. A basic goal of the downlink beamforming is to minimize the total transmission power subject to the chosen QoS constraints. We often choose the received SINR to be the QoS and give a predefined threshold on it. The problem formulation can be written as

$$\begin{aligned} \min_{\{\mathbf{w}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \|\mathbf{w}_k\|^2 \\ \text{s.t.} \quad & \frac{|\mathbf{w}_k^H \mathbf{h}_k|^2}{\sum_{\ell \neq k}^K |\mathbf{w}_\ell^H \mathbf{h}_k|^2 + \sigma_k^2} \geq \gamma_k, \ k = 1, \ldots, K \end{aligned} \tag{2.3}$$

where γ_k is the threshold of SINR for User k. One way of solving this problem is to rewrite it as a second-order cone program (SOCP) problem [1]

$$\begin{aligned} \min_{\{\mathbf{w}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \|\mathbf{w}_k\|^2 \\ \text{s.t.} \quad & (\mathbf{w}_k^H \mathbf{h}_k)^2 \geq \gamma_k \sum_{\ell \neq k}^K |\mathbf{w}_\ell^H \mathbf{h}_k|^2 + \gamma_k \sigma_k^2, \ k = 1, \ldots, K \end{aligned} \tag{2.4}$$

In Eq. (2.3), we employ instantaneous CSI $\mathbf{h}_k$, such that the expression of SINR is a measure of the instantaneous SINR. However, instantaneous SINR is hard to obtain

in practice. It is, therefore, more often to assume that the second order statistics of the channel is known and apply that to our problem. Thus, denoting the channel covariance matrix as $\mathbf{R}_k = \mathbb{E}\{\mathbf{h}_k\mathbf{h}_k^H\}$, the problem can be reformulated as

$$\begin{aligned} \min_{\{\mathbf{w}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \|\mathbf{w}_k\|^2 \\ \text{s.t.} \quad & \frac{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}{\sum_{\ell \neq k}^K \mathbf{w}_\ell^H \mathbf{R}_k \mathbf{w}_\ell + \sigma_k^2} \geq \gamma_k, \ k = 1, \ldots, K \end{aligned} \tag{2.5}$$

where we define the average SINR as

$$\text{SINR} = \frac{\text{average signal power}}{\text{average interference} + \text{noise power}} \tag{2.6}$$

and γ_k is the threshold of average SINR for User k.

The problem of (2.5) is non-convex and difficult to solve. A popular method to convert it into a convex problem is semidefinite relaxation [27]. Letting $\mathbf{W}_k = \mathbf{w}_k\mathbf{w}_k^H$, problem (2.5) can be rewritten as

$$\begin{aligned} \min_{\{\mathbf{W}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \text{Tr}(\mathbf{W}_k) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{R}_k\mathbf{W}_k) - \gamma_k \sum_{\ell \neq k}^K \text{Tr}(\mathbf{R}_k\mathbf{W}_\ell) \geq \gamma_k \sigma_k^2 \\ & \mathbf{W}_k \succeq \mathbf{0}, \ \text{rank}(\mathbf{W}_k) = 1, \ k = 1, \ldots, K \end{aligned} \tag{2.7}$$

The objective function is linear in terms of $\mathbf{W}_k$, and the constraints are convex except the rank-one condition. Dropping the rank-one condition, the problem becomes an SDP problem which is convex and easy to solve. This relaxed problem is equivalent to

the original problem when the solutions of $\mathbf{W}_k$ are of rank one. It has been shown [1] that simulation results to such kind of problem invariably yield rank-one matrices $\mathbf{W}_k$ when solving the relaxed problem with the rank-one constraint dropped. Therefore, semidefinite relaxation is an effective way to convert the non-convex problem into a convex one and can make sure that the two problems yield the same solution.

In the non-robust beamforming method, we assume that we know the channel information. We design the beamforming vector $\mathbf{w}_k$ to match the corresponding signal steering vector $\mathbf{h}_k$ and make sure that their inner product is large, while the inner product between $\mathbf{w}_k$ and other users' steering vectors $\mathbf{h}_\ell(\ell \neq k)$ is as small as possible. However, the true channel vector or channel covariance matrix is not known in practice, and there is always error when we do estimation. Under such situations, the non-robust beamforming design may not have good performance in practice. Therefore, we introduce robust beamforming methods in the next sections where estimation error is taken into account when we design the beamforming vectors.

2.2 Robust beamforming method with absolute error bound

Robust beamforming algorithms become popular in recent years when taking into account of consideration of the imperfection of CSI. Many papers [13][14] have proposed approaches to solve beamforming problems under the consideration of uncertainties in the channel knowledge.

Most of the papers [13][14] apply worst-case robust beamforming methods in which the absolute error bound is given. For example, errors in the estimation of DCC

matrices in [13] are taken such that the true covariance matrix $\mathbf{R}_k$ is modeled as $\hat{\mathbf{R}}_k + \mathbf{\Delta}_k$. The estimation error $\mathbf{\Delta}_k$ is upper bounded by a known constant ϵ_k such that $\|\mathbf{\Delta}_k\| \leq \epsilon_k$. The worst-case robust beamforming problem can then be formulated as

$$\begin{aligned} &\min_{\{\mathbf{w}_k\}_{k=1}^K} \sum_{k=1}^K \|\mathbf{w}_k\|^2 \\ &\text{s.t.} \quad \min_{\|\mathbf{\Delta}_k\| \leq \epsilon_k} \frac{\mathbf{w}_k^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_k}{\sum_{\ell \neq k}^K \mathbf{w}_\ell^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_\ell + \sigma_k^2} \geq \gamma_k, \ k = 1, \ldots, K \end{aligned} \tag{2.8}$$

The constraint in (2.8) includes the worst-case SINR. In [1], the minimum value of the constraint can be approximated by the known error bound of $\mathbf{\Delta}_k$. Then the problem can be expressed as

$$\begin{aligned} &\min_{\{\mathbf{w}_k\}_{k=1}^K} \sum_{k=1}^K \|\mathbf{w}_k\|^2 \\ &\text{s.t.} \quad \frac{\mathbf{w}_k^H(\hat{\mathbf{R}}_k - \epsilon_k \mathbf{I})\mathbf{w}_k}{\sum_{\ell \neq k}^K \mathbf{w}_\ell^H(\hat{\mathbf{R}}_k + \epsilon_k \mathbf{I})\mathbf{w}_\ell + \sigma_k^2} \geq \gamma_k, \ k = 1, \ldots, K \end{aligned} \tag{2.9}$$

This problem can be converted into a convex problem by applying semidefinite relaxation similar to the process in Section 2.1. However, this approximation is too conservative in that it may have excluded many feasible solutions. A less conservative way is to solve the sub-optimization problem in the constraint first [13]. Denoting

$$\mathbf{A}_k = \gamma_k \sum_{\ell \neq k}^K \mathbf{w}_\ell \mathbf{w}_\ell^H - \mathbf{w}_k \mathbf{w}_k^H \tag{2.10}$$

The constraint in (2.8) can be rewrite as

$$\min_{\|\mathbf{\Delta}_k\|\leq\epsilon_k} \mathrm{Tr}\{(\hat{\mathbf{R}}_k+\mathbf{\Delta}_k)\mathbf{w}_k\mathbf{w}_k^H\} \geq \gamma_k \sum_{\ell\neq k}^{K} \mathrm{Tr}\{(\hat{\mathbf{R}}_k+\mathbf{\Delta}_k)\mathbf{w}_\ell\mathbf{w}_\ell^H\} + \gamma_k\sigma_k^2 \tag{2.11}$$

which is equivalent to

$$\min_{\|\mathbf{\Delta}_k\|\leq\epsilon_k} -(\mathrm{Tr}\{\mathbf{\Delta}_k\mathbf{A}_k\} + \mathrm{Tr}\{\hat{\mathbf{R}}_k\mathbf{A}_k\} + \sigma_k^2\gamma_k) \geq 0 \tag{2.12}$$

Considering the fact that the true covariance matrix $\mathbf{R}_k$ is positive semidefinite, we formulate the constraint as the following optimization problem

$$\begin{aligned} \min_{\mathbf{\Delta}_k} \ & -(\mathrm{Tr}\{\mathbf{\Delta}_k\mathbf{A}_k\} + \mathrm{Tr}\{\hat{\mathbf{R}}_k\mathbf{A}_k\} + \sigma_k^2\gamma_k) \\ \text{s.t.} \ & \|\mathbf{\Delta}_k\| \leq \epsilon_k \\ & \hat{\mathbf{R}}_k + \mathbf{\Delta}_k \succeq \mathbf{0} \end{aligned} \tag{2.13}$$

This problem can be solved by solving its Lagrange dual problem. By introducing Lagrange multiplier $\mathbf{Z}_k$, the Lagrange dual problem is

$$\begin{aligned} \max_{\mathbf{Z}_k} \ & -\epsilon_k\|\mathbf{A}_k+\mathbf{Z}_k\|^2 - \mathrm{Tr}\{\mathbf{R}_k(\mathbf{Z}_k+\mathbf{A}_k)\} - \sigma_l^2\gamma_k \\ \text{s.t.} \ & \mathbf{Z}_k \succeq \mathbf{0} \end{aligned} \tag{2.14}$$

Therefore, the minimization problem in (2.8) can be replaced by the maximization problem shown in Eq. (2.14). Therefore, the original problem can now be rewritten

as

$$\begin{aligned}
&\min_{\{\mathbf{w}_k,\mathbf{Z}_k\}} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \\
&\text{s.t.} \quad \max_{\mathbf{Z}_k \succeq \mathbf{0}} (-\epsilon_k \|\mathbf{A}_k + \mathbf{Z}_k\| - \mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \sigma_k^2 \gamma_k) \geq 0, \ k = 1, \ldots, K
\end{aligned} \tag{2.15}$$

The constraint can be satisfied when there exists a positive semidefinite $\mathbf{Z}_k$, so that problem (2.15) is equivalent to

$$\begin{aligned}
&\min_{\{\mathbf{w}_k,\mathbf{Z}_k\}} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \\
&\text{s.t.} \quad -\epsilon_k \|\mathbf{A}_k + \mathbf{Z}_k\| - \mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \sigma_k^2 \gamma_k \geq 0 \\
&\qquad \mathbf{Z}_k \succeq \mathbf{0}, \ k = 1, \ldots, K
\end{aligned} \tag{2.16}$$

Letting $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$ and dropping the rank-one condition, the problem becomes

$$\begin{aligned}
&\min_{\{\mathbf{W}_k,\mathbf{Z}_k\}} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \\
&\text{s.t.} \quad -\epsilon_k \|\mathbf{A}_k + \mathbf{Z}_k\| - \mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \sigma_k^2 \gamma_k \geq 0 \\
&\qquad \mathbf{Z}_k \succeq \mathbf{0}, \mathbf{W}_k \succeq \mathbf{0} \ k = 1, \ldots, K
\end{aligned} \tag{2.17}$$

Eq. (2.17) shows a relaxed convex SDP problem and can be solved using interior-point algorithms. Simulation results in Fig. 2.1 (copied from the original figure in [13]) show that all the solutions are rank-one matrices and the performance is better than that in (2.9).

Another robust downlink beamforming method is proposed in [14] where a trace bound to estimation error matrix is used. The difference between these two bounds

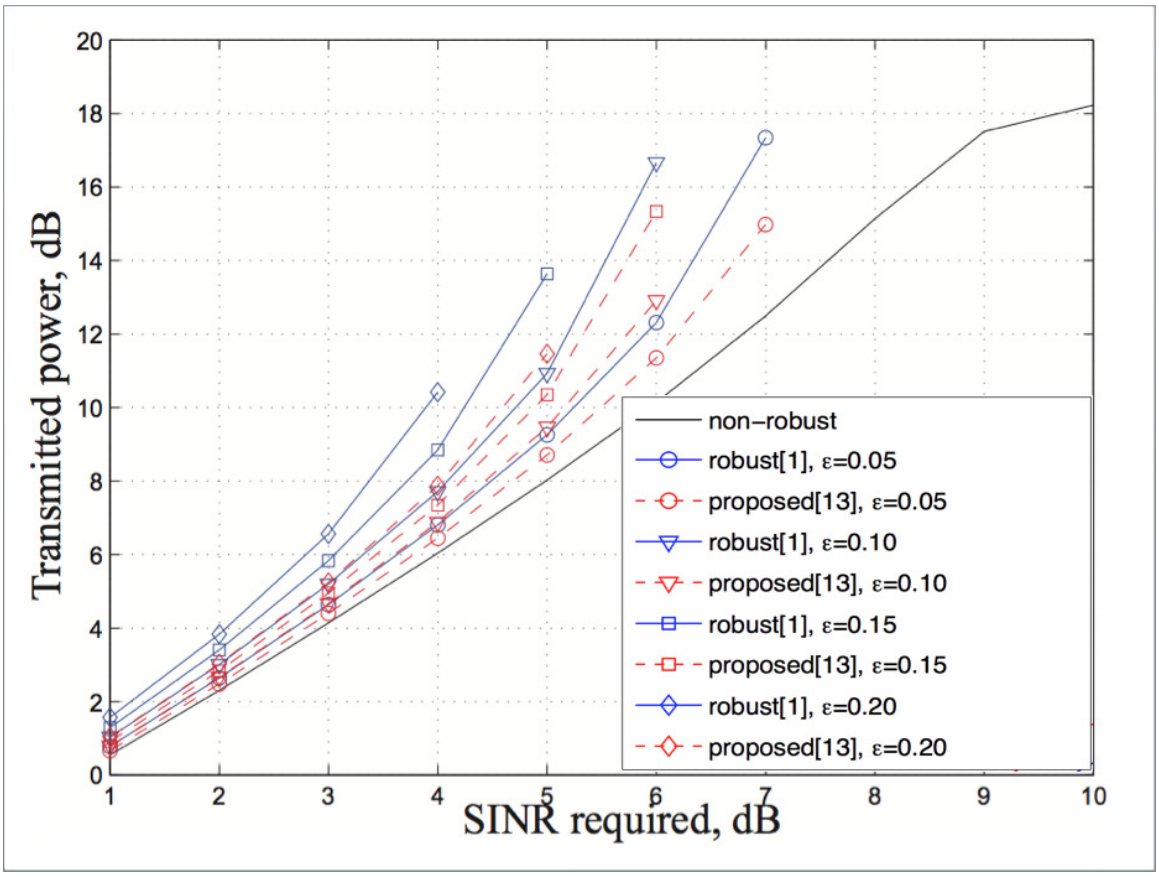


Figure 2.1: Total transmitted power versus required SINR [13]

is that in [13], the Frobenius norm bound restricts all entries of $\mathbf{\Delta}_k$, while in [14], the trace bound only constraints the diagonal elements of the mismatch matrix. The uncertainty sets are given by

$$\mathrm{Tr}\{\mathbf{\Delta}_k\} \leq \mu_k,\ \mathbf{\Delta}_k \succeq \mathbf{0},\ k = 1, \ldots, K \tag{2.18}$$

where μ_k is the predefined trace bound for User k.

Hence, the worst-case robust beamforming problem under trace bound can be formulated as

$$\begin{aligned} &\min_{\mathbf{w}_k} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \\ &\text{s.t.} \quad \min_{\mathrm{Tr}\{\mathbf{\Delta}_k\} \leq \mu_k} \frac{\mathbf{w}_k^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_k}{\sum_{\ell \neq k}^{K} \mathbf{w}_\ell^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_\ell + \sigma_k^2} \geq \gamma_k \\ &\qquad \hat{\mathbf{R}}_k + \mathbf{\Delta}_k \succeq \mathbf{0},\ \mathbf{\Delta}_k \succeq \mathbf{0} \end{aligned} \tag{2.19}$$

Following similar procedure as used in [13], we deal with the sub-optimization problem first. Using the same definition in Eq. (2.10), the constraint can be written as

$$\begin{aligned} &\min_{\mathbf{\Delta}_k} \ -\mathrm{Tr}\{(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{A}_k\} \\ &\text{s.t.} \quad \hat{\mathbf{R}}_k + \mathbf{\Delta}_k \succeq \mathbf{0},\ \mathbf{\Delta}_k \succeq \mathbf{0},\ \mathrm{Tr}\{\mathbf{\Delta}_k\} \leq \mu_k \end{aligned} \tag{2.20}$$

This optimization problem can be converted to its Lagrange dual problem equivalently by introducing Lagrange multipliers β_k and $\mathbf{Z}_k$. The simplified dual problem can be

expressed as

$$\begin{aligned} &\max_{\beta_k, \mathbf{Z}_k} \; -\mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \beta_k \mu_k \\ &\text{s.t.} \quad \beta_k \mathbf{I} + \mathbf{A}_k + \mathbf{Z}_k \succeq \mathbf{0}, \; \mathbf{Z}_k \succeq \mathbf{0}, \; \beta_k \geq 0 \end{aligned} \tag{2.21}$$

The strong duality holds between (2.20) and (2.21). We can solve the above problem in two steps. First, consider the maximization with respect to β_k only. From [14], the solution of $\min_x x$ s.t. $x\mathbf{I} - \mathbf{M} \succeq 0$ is $\lambda_{max}(\mathbf{M})$. Thus, the optimum value of the following problem

$$\begin{aligned} &\max_{\beta_k} \; -\beta_k \\ &\text{s.t.} \; \beta_k \mathbf{I} + \mathbf{A}_k + \mathbf{Z}_k \succeq \mathbf{0} \end{aligned}$$

is equivalent to $\lambda_{max}(-\mathbf{A}_k - \mathbf{Z}_k)$. The other part involving the maximization with respect to $\mathbf{Z}_k$ can now be combined with the maximized -β_k above so that we can rewrite the problem (2.21) as

$$\max_{\mathbf{Z}_k \succeq \mathbf{0}} \; -\mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \lambda_{\max}(-\mathbf{A}_k - \mathbf{Z}_k)\mu_k \tag{2.22}$$

where the variable β_k is replaced by the largest eigenvalue of $-(\mathbf{A}_x + \mathbf{Z}_k)$. Hence, we just have $\mathbf{Z}_k$ as our variable in this problem. Applying this optimization problem as

the constraint, the original problem in (2.19) can be expressed as

$$\begin{aligned}
&\min_{\mathbf{w}_k} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \\
&\text{s.t.} \quad \max_{\mathbf{Z}_k \succeq \mathbf{0}} \; -\mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \lambda_{\max}(-\mathbf{A}_x - \mathbf{Z}_k)\mu_k \geq \sigma_k^2 \gamma_k \\
&\qquad k = 1, \ldots, K
\end{aligned} \tag{2.23}$$

The constraint can be satisfied when there exists some positive semidefinite $\mathbf{Z}_k$ for which

$$-\mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \lambda_{\max}(-\mathbf{A}_x - \mathbf{Z}_k)\mu_k \geq \sigma_k^2 \gamma_k \tag{2.24}$$

Letting $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$ and applying semidefinite relaxation approach, the relaxed problem can be rewritten as

$$\begin{aligned}
&\min_{\{\mathbf{W}_k,\ \mathbf{Z}_k\}} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \\
&\text{s.t.} \quad -\mathrm{Tr}\{\hat{\mathbf{R}}_k(\mathbf{A}_k + \mathbf{Z}_k)\} - \lambda_{\max}(-\mathbf{A}_k - \mathbf{Z}_k)\mu_k \geq \sigma_k^2 \gamma_k \\
&\qquad \mathbf{Z}_k \succeq \mathbf{0}, \mathbf{W}_k \succeq \mathbf{0},\ k = 1, \ldots,\ K
\end{aligned} \tag{2.25}$$

The original non-convex problem has been converted into a convex problem by some reformulations and SDR technique. The idea of reformulation used in [14] is similar to that in [13], the approach proposed in both papers is that we first deal with the constraint of the original problem, and the constraint can be considered as a sub-optimization problem which can be solved by solving its Lagrange dual problem. Finally, applying SDR technique, the original problem can be converted into a convex problem which can be solved by software such as CVX. Both of them choose

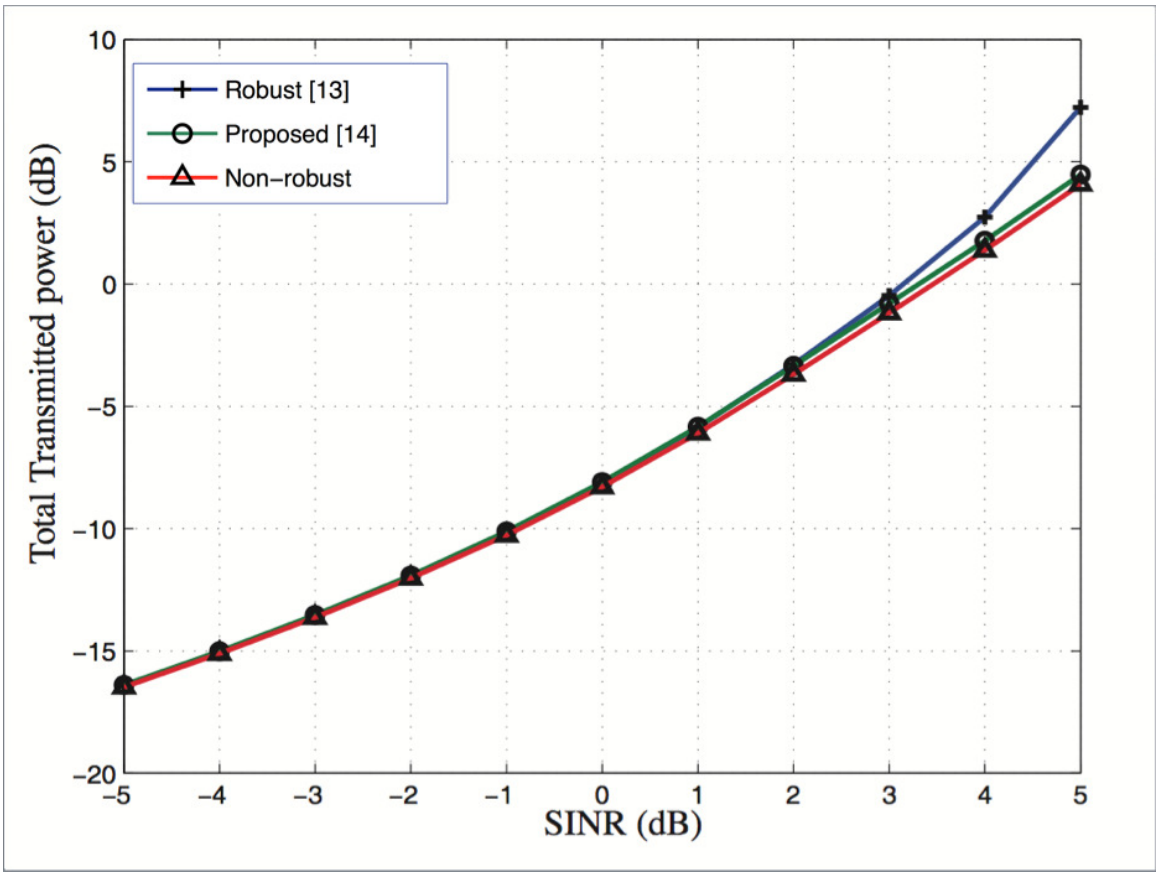


Figure 2.2: Total transmitted power versus SINR [14]

absolute error bound, Frobenius norm adds constraint to each element of $\mathbf{\Delta}_k$, while the alternative trace bound just limits the diagonal elements. Simulation results in Fig. 2.2 (copied from the original figure in [14]) show a small difference in performance between the two proposed methods.

2.3 Probabilistic robust beamforming method

In Section 2.2, we introduce robust downlink beamforming methods in which the error is bounded under a given constant such that the QoS is satisfied for each user. In these

proposed worst-case robust downlink beamforming methods[1, 13, 14], it is required that even the worst SINR must satisfy the predefined QoS threshold. However, in some cases, it seems to be good enough that the QoS constraint for each user can be satisfied with a high probability. Under this consideration, probabilistic robust beamforming methods have been proposed in many previous works, such as [18], [19] and [20]. In these papers, a stochastic model for the uncertainty is established, and the transmission power is minimized under an outage probabilistic QoS constraint. In [18], the mismatch matrix between the true and estimated covariance matrices is modeled as a Hermitian matrix with Gaussian distribution. While in [19] and [20], a stochastic model is applied to the channel estimation error with Gaussian distribution and uniform distribution, respectively. In all the three papers, a convex approximation is employed to solve the robust downlink beamforming problem.

The basic signal model in [18] is the same as what we introduced before. It is assumed that there is error in estimating the channel covariance matrix such that $\mathbf{R}_k = \hat{\mathbf{R}}_k + \mathbf{\Delta}_k$, where the entries of $\mathbf{\Delta}_k$ are assumed to be zero-mean, independent Gaussian with variance $\sigma^2_{e_k}$. Instead of using the QoS constraint in the worst-case robust downlink beamforming approach, an outage probability constraint is applied. If we define γ_k to be the threshold of SINR and ϵ_k to be the outage probability for User k, then the original problem formulation can be expressed as

$$\min_{\mathbf{w}_k} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \tag{2.26a}$$

$$\text{s.t. } \Pr\{\rho_k \geq \gamma_k\} \geq 1 - \epsilon_k, \ k = 1, \ldots, K \tag{2.26b}$$

where $\Pr(\cdot)$ denotes the probability and ρ_k is the average SINR as defined in Eq. (2.6)

$$\rho_k = \frac{\mathbf{w}_k^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_k}{\sum_{\ell \neq k}^{K} \mathbf{w}_\ell^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_\ell + \sigma_k^2} \tag{2.27}$$

Letting $\mathbf{W}_k = \mathbf{w}_k\mathbf{w}_k^H$ and introducing $\mathbf{Z}_k = \mathbf{W}_k - \gamma_k \sum_{\ell \neq k}^{K} \mathbf{W}_\ell$, the constraint in Eq. (2.26b) can be written as

$$\Pr\{\mathrm{Tr}\{(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{Z}_k\} \geq \gamma_k\sigma_k^2\} \geq 1 - \epsilon_k \tag{2.28}$$

SDR technique is applied in this problem, but the remaining outage probability constraint is still non-convex and difficult to be solved. With the assumption that the entries of $\mathbf{\Delta}_k$ are zero-mean Gaussian variables, the distribution of $\mathrm{Tr}\{(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{Z}_k\}$ can be obtained. Define a random variable as

$$y_k = \mathrm{Tr}\{(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{Z}_k\} \tag{2.29}$$

Therefore, $y_k \sim \mathcal{N}(\mathrm{Tr}\{\hat{\mathbf{R}}_k\mathbf{Z}_k\}, \sigma_{e_k}^2 \mathrm{Tr}\{\mathbf{Z}_k\mathbf{Z}_k^H\})$, where $\mathcal{N}(\ \cdot,\ \cdot)$ denotes Gaussian distribution. With known probability density function(pdf) of variable, the probability of the region that $y_k \geq \gamma_k\sigma_k^2$ can be calculated.

The result of the probability can be expressed using the Gaussian error function $\mathrm{erf}(\cdot)$

$$\text{Probability} = \begin{cases} \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\mu_k - \gamma_k\sigma_k^2}{\sqrt{2}\sigma_{e_k}\|\mathbf{Z}_k\|}\right), & \text{for } \gamma_k\sigma_k^2 \leq \mu_k \\ \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{\gamma_k\sigma_k^2 - \mu_k}{\sqrt{2}\sigma_{e_k}\|\mathbf{Z}_k\|}\right), & \text{for } \gamma_k\sigma_k^2 \geq \mu_k \end{cases} \tag{2.30}$$

where $\mu_k = \mathrm{Tr}\{\hat{\mathbf{R}}_k\mathbf{Z}_k\}$. Among the two equations shown in Eq. (2.30), we should choose the upper one since we need the probability being close to one. Therefore, the

outage probability constraint can be expressed as

$$\mathrm{erf}\left(\frac{\mu_k - \gamma_k \sigma_k^2}{\sqrt{2}\sigma_{e_k}\|\mathbf{Z}_k\|}\right) \geq 1 - 2\epsilon_k \tag{2.31}$$

After some straightforward transformation, we have

$$\mathrm{Tr}\{\hat{\mathbf{R}}_k \mathbf{Z}_k\} - \gamma_k \sigma_k^2 \geq \sqrt{2}\sigma_{e_k}\mathrm{erf}^{-1}(1 - 2\epsilon_k)\|\mathbf{Z}_k\| \tag{2.32}$$

where ϵ_k is a small number and $\mathrm{erf}^{-1}(\cdot)$ is the inverse error function. The outage probabilistic constraint has been converted into a deterministic constraint using the inverse error function together with the outage probability. Therefore, the original optimization problem becomes

$$\begin{aligned} \min_{\mathbf{W}_k} \ & \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \\ \text{s.t.} \ & \|\mathbf{Z}_k\| \leq \frac{1}{\sqrt{2}\sigma_{e_k}\mathrm{erf}^{-1}(1 - 2\epsilon_k)}(\mathrm{Tr}\{\hat{\mathbf{R}}_k \mathbf{Z}_k\} - \gamma_k \sigma_k^2) \\ & \mathbf{W}_k \succeq \mathbf{0}, \ k = 1, \ldots, K \end{aligned} \tag{2.33}$$

where we drop the rank-one condition to make the optimization problem convex. We can introduce a slack variable t_k to rewrite the constraint, so that the problem can be clearer and more tractable. The reformulated and relaxed optimization problem

can be written as

$$\begin{aligned} &\min_{\mathbf{W}_k} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \\ &\text{s.t.} \quad \|\mathbf{Z}_k\| \leq t_k,\ t_k = \frac{1}{\sqrt{2}\sigma_{e_k}\mathrm{erf}^{-1}(1-2\epsilon_k)}(\mathrm{Tr}\{\hat{\mathbf{R}}_k\mathbf{Z}_k\} - \gamma_k\sigma_k^2) \\ &\qquad \mathbf{W}_k \succeq \mathbf{0},\ k = 1, \ldots, K \end{aligned} \tag{2.34}$$

which is a convex optimization problem with linear objective function along with SDP and SOCP constraints. It can be solved by optimization tool such as CVX. Simulation shows that the results always satisfy rank-one condition, so the relaxed optimization problem actually yields the same solution for the original problem. What is more, the performance of the optimum beamformer in this method is better compared to that in the worst-case robust downlink beamforming approaches [1, 13, 14].

Another robust downlink beamforming method is proposed in [20], in which a stochastic model in channel estimation error vector is established and an outage probability constraint is considered on instantaneous SINR. Again, the base station is equipped with N antennas serving K users. The basic signal model is the same as before. The imperfect CSI is modeled as

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{e}_k,\ k = 1, \ldots, K \tag{2.35}$$

where $\mathbf{h}_k$ stands for the real channel vector from the BS to the kth user and the corresponding estimated channel vector is represented by $\hat{\mathbf{h}}_k$ with error denoting by $\mathbf{e}_k$. It is assumed that $\mathbf{e}_k \in \mathbb{C}^N$ is complex Gaussian CSI error with covariance matrix $\mathbf{C}_k$.

The QoS is represented by the instantaneous SINR for each user with the definition as follows:

$$\mathrm{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{\ell \neq k}^{K} |\mathbf{h}_k^H \mathbf{w}_\ell|^2 + \sigma_k^2} \tag{2.36}$$

For User k, the minimum SINR threshold is defined as γ_k, so that the outage probability robust downlink beamforming problem can be formulated as

$$\begin{aligned} &\min_{\mathbf{w}_k} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \\ &\text{s.t.} \quad \Pr\{\mathrm{SINR}_k \geq \gamma_k\} \geq 1 - \epsilon_k, \ k = 1, \ldots, K \end{aligned} \tag{2.37}$$

where ϵ_k is the outage probability.

First, the problem can be relaxed using semidefinite relaxation approach. Letting $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$, the optimization problem above can be expressed as

$$\min_{\mathbf{W}_k} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \tag{2.38a}$$

$$\text{s.t.} \quad \Pr\left\{ (\hat{\mathbf{h}}_k + \mathbf{e}_k)^H \left(\frac{1}{\gamma_k} \mathbf{W}_k - \sum_{\ell \neq k} \mathbf{W}_\ell \right) (\hat{\mathbf{h}}_k + \mathbf{e}_k) \geq \sigma_k^2 \right\} \geq 1 - \epsilon_k \tag{2.38b}$$

$$\mathbf{W}_k \succeq \mathbf{0}, \ \mathrm{rank}(\mathbf{W}_k) = 1, \ k = 1, \ldots, K \tag{2.38c}$$

where the rank-one condition can be dropped. The next step is to find a convex approximation of (2.38b). The constraint in the bracket of $\Pr(\cdot)$ can be expressed as a quadratic form in terms of $\mathbf{e}_k$, so that (2.38b) can be written as

$$\Pr\{\mathbf{e}^H \mathbf{Q} \mathbf{e} + 2\mathrm{Re}\{\mathbf{e}^H \mathbf{r}\} + s \geq 0\} \geq 1 - \epsilon \tag{2.39}$$

where $\mathbf{e}$ represents the whitened standard complex Gaussian vector with identity covariance matrix. The corresponding $\mathbf{Q}$, $\mathbf{r}$ and s are defined as

$$\mathbf{Q} = \mathbf{C}_k^{1/2}\left(\frac{1}{\gamma_k}\mathbf{W}_k - \sum_{\ell\neq k}^{K}\mathbf{W}_\ell\right)\mathbf{C}_k^{1/2},\ \mathbf{r} = \mathbf{C}_k^{1/2}\left(\frac{1}{\gamma_k}\mathbf{W}_k - \sum_{\ell\neq k}^{K}\mathbf{W}_\ell\right) \tag{2.40a}$$

$$s = \hat{\mathbf{h}}_k^H\left(\frac{1}{\gamma_k}\mathbf{W}_k - \sum_{\ell\neq k}^{K}\mathbf{W}_\ell\right)\hat{\mathbf{h}}_k - \sigma_k^2,\ \epsilon = \epsilon_k \tag{2.40b}$$

The outage probability constraint is non-convex, so that a convex approximation to this non-convex problem is needed. Due to the fact that the part in the bracket is quadratic function in terms of variable $\mathbf{e}$, we can apply the following lemma to convert the outage probability constraint into a deterministic one.

Lemma 1. *Suppose that we have a set $\mathcal{B}$ that satisfies the outage probability constraint, and the quadratic constraint*

$$\mathbf{e}^H\mathbf{Q}\mathbf{e} + 2Re\{\mathbf{e}^H\mathbf{r}\} + s \geq 0$$

holds for all $\mathbf{e} \in \mathcal{B}$, then we have $Pr\{\mathbf{e}^H\mathbf{Q}\mathbf{e} + 2Re\{\mathbf{e}^H\mathbf{r}\} + s \geq 0\} \geq 1 - \epsilon$

The lemma above suggests the replacement of the outage probability constraint by a deterministic worst-case constraint. The outage probability constraint is satisfied by finding the set $\mathcal{B}$. There are many choices of such a set, however, a natural one is a spherical set which gives a bound to the uncertainty just like what we use in worst-case approach. The set $\mathcal{B}$ can be defined as

$$\mathcal{B} = \{\mathbf{e} \in \mathbb{C}^n \mid \|\mathbf{e}\| \leq d\} \tag{2.41}$$

where d is the error bound and it can be calculated according to the outage probability ϵ. In worst-case approach, for example in [13], we consider the robust problem in the situation that the estimation error is upper bounded by $\|\mathbf{e}\| \leq d$, where d is a given constant. If the bounds of worst-case method and outage probabilistic method are chosen the same, then the problem formulations of the two methods would be the same. Therefore, we approximate the outage probability problem in a conservative step by a worst-case deterministic constraint as shown in (2.41). We have assumed that $\mathbf{e}$ is standard complex Gaussian, so that the real part and imaginary part of $\mathbf{e}$ are independent Gaussian variables with covariance matrix $\frac{1}{2}\mathbf{I}$. The Frobenius norm square of $\mathbf{e}$ can be expressed as

$$\|\mathbf{e}\|^2 = \frac{1}{2}\sum_{i=1}^{2n} z_i^2 \tag{2.42}$$

where z_i are independent, standard normal variables. Therefore, $2\|\mathbf{e}\|^2$ is distributed according to the chi-squared distribution with $2n$ degrees of freedom. If we define $\Phi^{-1}_{\chi^2_{2n}}(\cdot)$ as the inverse cumulative function of Chi-square distribution, then it is easy to get the value of d.

$$d = \sqrt{\frac{\Phi^{-1}_{\chi^2_{2n}}(1-\epsilon)}{2}} \tag{2.43}$$

After finding the set $\mathcal{B}$, the outage probability constraint can be approximated by the following two constraints

$$\mathbf{e}^H\mathbf{Q}\mathbf{e} + 2\mathrm{Re}\{\mathbf{e}^H\mathbf{r}\} + s \geq 0 \tag{2.44a}$$

$$\|\mathbf{e}\| \leq d \tag{2.44b}$$

The constraints shown in (2.44) are infinite constraints due to the random variable $\mathbf{e}$. To convert these infinite constraints into finite ones, $\mathcal{S}$-lemma is applied.

Lemma 2. *($\mathcal{S}$-lemma) Let $f_i(\mathbf{x}) = \mathbf{x}^H\mathbf{Q}_i\mathbf{x} + 2Re\{\mathbf{x}^H\mathbf{r}_i\} + s_i$ for $i = 0, 1$, where $\mathbf{x} \in \mathbb{C}^n$. Suppose that there exists an $\hat{\mathbf{x}} \in \mathbb{C}^n$ satisfying $f_1(\hat{\mathbf{x}}) < 0$. Then, the following statements are equivalent:*

1. $f_0(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{C}^n$ satisfying $f_1(\mathbf{x}) \leq 0$.

2. There exists a $t \geq 0$ such that

$$\begin{bmatrix} \mathbf{Q}_0 & \mathbf{r}_0 \\ \mathbf{r}_0^H & s_0 \end{bmatrix} + t \begin{bmatrix} \mathbf{Q}_1 & \mathbf{r}_1 \\ \mathbf{r}_1^H & s_1 \end{bmatrix} \succeq \mathbf{0} \tag{2.45}$$

According to $\mathcal{S}$-lemma, the corresponding $\mathbf{Q}_i,\ \mathbf{r}_i,\ s_i$ can be chosen as $(\mathbf{Q}_0,\ \mathbf{r}_0,\ s_0) = (\mathbf{Q},\ \mathbf{r},\ s)$ and $(\mathbf{Q}_1,\ \mathbf{r}_1,\ s_1) = (\mathbf{I}_n,\ \mathbf{0},\ -d^2)$.

Therefore, the optimization problem can be converted into the following convex problem:

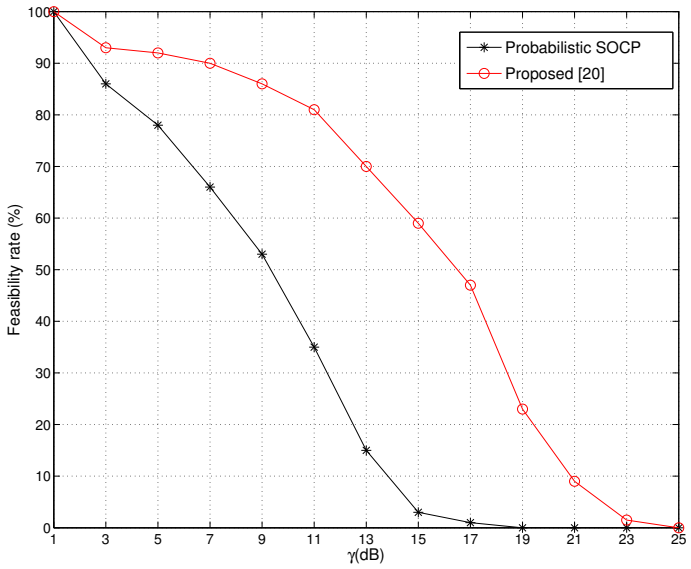
$$\min_{\mathbf{W}_k} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \tag{2.46a}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{Q}_k + t_k\mathbf{I} & \mathbf{r}_k \\ \mathbf{r}_k^H & s_k - t_k d_k^2 \end{bmatrix} \succeq \mathbf{0} \tag{2.46b}$$

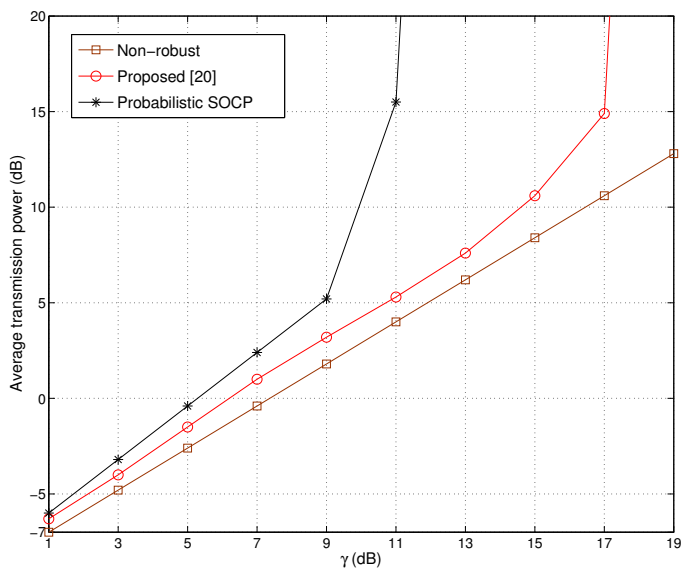
$$\mathbf{W}_k \succeq \mathbf{0},\ t_k \geq 0,\ k = 1, \ldots, K \tag{2.46c}$$

This is a convex approximation to the original outage probability constraint problem, and it can be solved by using standard optimization tool. In the simulation part of [20], the authors compare the performance of the proposed method to that using probabilistic SOCP method. The simulation results in Fig. 2.3 (copied from the

original figure in [20]) show better performances in the comparison of both feasibility rate and transmission power.



(a)



(b)

Figure 2.3: Feasibility and transmit power performance of the various methods [20]

Chapter 3

Probabilistic robust beamforming with Riemannian Distance

As we have mentioned before, probabilistic robust beamforming method has been proposed in [18][19]. In these former works, the Euclidean distance is used to bound the uncertainty of the estimation error. In this chapter, we will introduce a new metric, the Riemannian distance (RD), as the measure of estimation error. Due to the special structure of the PSD matrix, it has been proved that RD is a more precise way to calculate the distance between two such matrices. With the use of this error measure for the covariance matrices, we then proposed a probabilistic robust beamforming method and showed that we could convert this problem into a convex optimization problem.

3.1 Introduction to Riemannian distance

In robust beamforming, the Frobenius norm $\|\mathbf{\Delta}_k\|_{\mathrm{F}}$ is usually used to measure the distance between $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$, such that

$$\|\mathbf{\Delta}_k\|_F = \sqrt{\mathrm{trace}(\mathbf{\Delta}_k^H \mathbf{\Delta}_k)} \tag{3.1}$$

It is induced by the inner product of $\langle \mathbf{\Delta}_k, \mathbf{\Delta}_k \rangle$, so it can be considered as the Euclidean distance between $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$. While the Euclidean metric is very useful and popular in most applications, it may not be the most appropriate measure for the covariance matrix of the channel. In our paper, the DCC matrices have special structures. They are Hermitian symmetric and positive semidefinite. As discussed in [21], these properties describe a hyper-surface, called a manifold, in the signal space on which these points of PSD matrices are located. Therefore, we can say that our DCC matrices describe a manifold, a Riemannian manifold, to be more specific. The definition of RD has been given in [21], three kinds of which are derived in [21]. Among those, the second version of RD thought to be convenient for robust beamforming.

Theorem 3. *$\mathcal{M}$ denotes the manifold of PSD matrices. If $\mathbf{P}_1 \in \mathcal{M}$ and $\mathbf{P}_2 \in \mathcal{M}$, then the geodesic distance between $\mathbf{P}_1$ and $\mathbf{P}_2$ on $\mathcal{M}$ is*

$$d_{R_2}(\mathbf{P}_1, \mathbf{P}_2) = \sqrt{\mathrm{Tr}\mathbf{P}_1 + \mathrm{Tr}\mathbf{P}_2 - 2\mathrm{Tr}\mathbf{P}_1^{1/2}\mathbf{P}_2^{1/2}} \tag{3.2}$$

The distance between the two matrices should be measured along the shortest path on the manifold between the points. The Euclidean distance shown in Eq.(3.1) measures the straight line distance between the two points may be less appropriate

and shorter than Riemannian distance. This concept is akin to finding the distance between two cites on earth. As mentioned in the Introduction, $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$ are Hermitian and PSD matrices forming a manifold in the signal space. Thus, we should use the RD to measure the difference between $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$ on the DCC matrix manifold.

3.2 Worst-case robust beamforming method with Riemannian distance

In [22], the application of RD to robust beamforming was proposed to deal with a worst-case problem with imperfect second order CSI at the transmitter. A RD is used to bound the uncertainty of estimation error between the real and estimated covariance matrices. The worst-case robust downlink beamforming problem has been first formulated and then a tractable convex optimization problem is derived. The simulation results show an improved performance of the proposed method comparing to previous methods.

The authors consider the senario of a BS with N antennas serving K users. The transmitted signal at the BS is given by

$$\mathbf{x}(t) = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{u}_k s_k(t) \tag{3.3}$$

where p_k is the transmission power, $\mathbf{u}_k \in \mathbb{C}^N$ is the normalized beamforming weight vector and $s_k(t)$ is the transmitted data stream with unit power (i.e., $\mathbb{E}\{|s_k(t)|^2\} = 1$) sampled at time t. The received signal for User k can be modeled as

$$y_k(t) = \mathbf{h}_k^H \mathbf{x}(t) + n_k(t),\ \ k = 1, \ldots, K \tag{3.4}$$

where $\mathbf{h}_k \in \mathbb{C}^N$ denotes the downlink channel vector from the BS to the kth user, and $n_k(t)$ is the assumed additive complex circular Gaussian noise having zero mean and variance σ_k^2. The real and estimated channel covariance matrices are denoted by $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$ respectively with the mismatch matrix $\mathbf{\Delta}_k = \mathbf{R}_k - \hat{\mathbf{R}}_k$ for User k. As discussed in the last section, the two covariance matrices lie on a Riemannian manifold, so that we can express the RD between the two as

$$d_{Rk}(\mathbf{R}_k, \hat{\mathbf{R}}_k) = \sqrt{\mathrm{Tr}\mathbf{R}_k + \mathrm{Tr}\hat{\mathbf{R}}_k - 2\mathrm{Tr}\mathbf{R}_k^{1/2}\hat{\mathbf{R}}_k^{1/2}} \tag{3.5}$$

Therefore, the worst-case robust downlink beamforming problem can be formulated as

$$\min_{\{\mathbf{w}_k\}} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \tag{3.6a}$$

$$\text{s.t.} \quad \min_{\substack{\hat{\mathbf{R}}_k + \mathbf{\Delta}_k \succeq \mathbf{0} \\ d_{Rk}^2 \leq \alpha_k^2}} \frac{\mathbf{w}_k^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_k}{\sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{w}_\ell^H(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{w}_\ell + \sigma_k^2} \geq \gamma_k, k = 1, \ldots, K \tag{3.6b}$$

where α_k is the bound on the RD and γ_k is the predefined SINR threshold. The sub-optimization problem Eq. (3.6b) can be rewritten by letting $\mathbf{A}_k = \gamma_k \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{w}_\ell \mathbf{w}_\ell^H - \mathbf{w}_k \mathbf{w}_k^H$

$$\min_{\mathbf{\Delta}_k} - (\mathrm{Tr}\{\mathbf{\Delta}_k \mathbf{A}_k\} + \mathrm{Tr}\{\hat{\mathbf{R}}_k \mathbf{A}_k\} + \sigma_k^2 \gamma_k) \geq 0 \tag{3.7a}$$

$$\text{s.t. } d_{Rk}^2(\hat{\mathbf{R}}_k, \hat{\mathbf{R}}_k + \mathbf{\Delta}_k) \leq \alpha_k^2 \tag{3.7b}$$

$$\hat{\mathbf{R}}_k + \mathbf{\Delta}_k \succeq \mathbf{0} \tag{3.7c}$$

Therefore, the robust beamforming problem can be expressed as

$$\begin{aligned}
&\min_{\{\mathbf{w}_k\}} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \\
&\text{s.t. The sub-optimizaiton problem Eq. (3.7) is satisfied} \\
&\quad \text{for } k = 1, \ldots, K
\end{aligned} \tag{3.8}$$

The difficulty to solve this problem comes from the sub-optimization problem of Eq. (3.7). To make the constraints satisfied, the authors proposed the following:

Proposition 1: A sufficient condition for the worst case SINR constraints to be satisfied is that there exists a set of non-negative λ_k such that

$$\mathbf{X}_k = \begin{pmatrix} -\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2} & \mathbf{b}_k(\lambda_k) \\ \mathbf{b}_k^H(\lambda_k) & c_k(\lambda_k) \end{pmatrix} \succeq \mathbf{0} \tag{3.9}$$

where

$$\mathbf{b}_k(\lambda_k) = -\lambda_k \mathrm{vec}(\hat{\mathbf{R}}_k^{1/2}) \tag{3.10a}$$

$$c_k(\lambda_k) = \lambda_k \mathrm{Tr}\{\hat{\mathbf{R}}_k\} - \sigma_k^2 \gamma_k - \lambda_k \alpha_k^2 \tag{3.10b}$$

The proof of Proposition 1 has been given in [22] by letting $\mathbf{q}_k = \mathrm{vec}((\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)^{1/2})$ and solving the sub-optmization problem due to the strong duality property. First, Eq. (3.7) can be rewritten as

$$\begin{aligned}
&\min_{\mathbf{q}_k} -\mathbf{q}_k^H (\mathbf{I}_N \otimes \mathbf{A}_k) \mathbf{q}_k - \sigma_k^2 \gamma_k \\
&\text{s.t. } \mathbf{q}_k^H \mathbf{q}_k - 2\mathrm{Re}\{\mathrm{vec}^H(\hat{\mathbf{R}}_k^{1/2} \mathbf{q}_k)\} + \mathrm{Tr}\{\hat{\mathbf{R}}_k\} - \alpha_k^2 \leq 0
\end{aligned} \tag{3.11}$$

With the notations defined in Eq. (3.10), the dual problem can be expressed as

$$\begin{aligned} \max_{\lambda_k} \ & c_k(\lambda_k) - \mathbf{b}_k^H(\lambda_k)(-\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2})\mathbf{b}_k(\lambda_k) \\ \text{s.t.} \ & -\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2} \succeq \mathbf{0} \\ & \mathbf{b}_k(\lambda_k) \in \mathcal{R}(-\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2}) \end{aligned} \tag{3.12}$$

The constraint of the worst-case robust beamforming problem has been converted to the dual problem as shown above. To satisfy this constraint, we need to find a non-negative λ_k that can make the object value of the dual problem greater or equal to zero. Furthermore, it has been proved that it is sufficient to find one λ_k in the feasible set of (3.12) to make sure that the objective value is non-negative.

For some λ_k that we find, we always have

$$c_k(\lambda_k) - \mathbf{b}_k^H(\lambda_k)(-\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2})\mathbf{b}_k(\lambda_k) \geq 0 \tag{3.13}$$

and it is equivalent to express it using Schur complement as what is showed in Proposition 1.

The worst-case SINR constraints in Eq. (3.7) have been replaced by Proposition 1. Letting $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H, \ k = 1, \ldots, K$ and applying semidefinite relaxation to drop the rank one condition first, then we can express the original optimization problem

by a tractable convex optimization problem which can be expressed as

$$\begin{aligned} &\min_{\mathbf{W}_k,\lambda_k} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} \\ &\text{s.t.} \quad \begin{pmatrix} -\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2} & \mathbf{b}_k(\lambda_k) \\ \mathbf{b}_k^H(\lambda_k) & c_k(\lambda_k) \end{pmatrix} \succeq \mathbf{0} \\ &\qquad \lambda_k \geq 0, k = 1, \ldots, K \end{aligned} \tag{3.14}$$

where

$$\mathbf{b}_k(\lambda_k) = -\lambda_k \mathrm{vec}(\hat{\mathbf{R}}_k^{1/2}) \tag{3.15a}$$

$$c_k(\lambda_k) = \lambda_k \mathrm{Tr}\{\hat{\mathbf{R}}_k\} - \sigma_k^2 \gamma_k - \lambda_k \alpha_k^2 \tag{3.15b}$$

Eq. (3.14) shows that it has a different formation of worst-case robust downlink beamforming problem from conventional methods due to the application of Riemannian distance as the measurement between two DCC matrices. Simulation results in Chapter 5 show that the proposed method applying the RD has dramatically improved performance compared to other robust beamforming methods employing the Frobenius distance.

3.3 Probabilistic robust beamforming problem formulation

In this section, we develop the approach to robust beamforming based on the outage probability using the RD as a measure of the covariance matrix estimation error. We

consider the same scenario as that in Section 3.2. Denote the signal transmitted from the BS as $\mathbf{x}(t)$, and assuming that there are K users in the cell for which the kth user is receiving signal $s_k(t)$, then

$$\mathbf{x}(t) = \sum_{k=1}^{K} \mathbf{w}_k s_k(t) \tag{3.16}$$

where $\mathbf{w}_k \in \mathbb{C}^N$ is the transmitter beamforming vector for User k, and $s_k(t)$, the data stream for User k, is assumed to have zero mean and unit power (i.e., $\mathbb{E}\{|s_k(t)|^2\} = 1$). Hence, the received signal for User k can be modeled as

$$y_k(t) = \mathbf{h}_k^H \mathbf{x}(t) + n_k(t) \tag{3.17}$$

where $\mathbf{h}_k \in \mathbb{C}^N$ is the downlink channel vector from the BS to User k, and $n_k(t)$ is the additive Gaussian noise having zero mean and variance σ_k^2.

The commonly used worst case based downlink beamforming approach requires that all the QoS constraints must satisfy the predefined threshold, but this kind of method is sometimes too conservative. So, instead of using that, we replace the QoS constraints which are of worst case based by the more relaxed probabilistic constraints. Thus, our algorithm aims at minimizing the transmission power while keeping the SINR outage probability of each user under a specified value. The problem can be formulated as

$$\min_{\mathbf{w}_1,\cdots,\mathbf{w}_K \in \mathbb{C}^N} \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \tag{3.18a}$$

$$\text{s.t.} \quad \Pr\{\rho_k \geq \gamma_k\} \geq 1 - \epsilon_k, \quad k = 1, \ldots, K \tag{3.18b}$$

where γ_k is the SINR threshold for User k and ϵ_k is the corresponding outage probability.

Eq. (3.18b) is the outage probability constraint that we choose. The outage probability ϵ_k adds the flexibility to the problem, and we can choose different value of ϵ_k according to the different requirement of the service fidelity to make sure that for each user, the QoS constraint can be satisfied no less than the threshold at a percentage of $(1-\epsilon_k)\times 100\%$. As we can see that the outage probability constraint has useful meaning, but it also increases the complexity of the optimization problem because the probabilistic constraint makes the problem non-convex.

To convert this problem into a convex optimization problem, we can first use semidefinite relaxation (SDR) to rewrite the formulation. Let $\mathbf{W}_i = \mathbf{w}_i\mathbf{w}_i^H$ and denote

$$\mathbf{D}_k = \left(\frac{1}{\gamma_k}\mathbf{w}_k\mathbf{w}_k^H - \sum_{\ell\neq k}^{K}\mathbf{w}_\ell\mathbf{w}_\ell^H\right) = \frac{1}{\gamma_k}\mathbf{W}_k - \sum_{\ell\neq k}^{K}\mathbf{W}_\ell \tag{3.19}$$

Then the problem of Eq. (3.18) can be rewritten as

$$\min_{\mathbf{W}_1,\cdots,\mathbf{W}_K\in\mathbb{H}^{N\times N}} \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{W}_k) \tag{3.20a}$$

$$\text{s.t.} \quad \Pr\left\{\mathrm{Tr}[(\hat{\mathbf{R}}_k+\mathbf{\Delta}_k)\mathbf{D}_k]\geq\sigma_k^2\right\}\geq 1-\epsilon_k \tag{3.20b}$$

$$\mathrm{rank}(\mathbf{W}_k)=1,\ \mathbf{W}_k\succeq\mathbf{0},\ k=1,\cdots,K \tag{3.20c}$$

where $\mathbb{H}^{N\times N}$ denotes the set of $N\times N$ Hermitian matrices. This is a non-convex problem due to the non-convex constraints. After relaxing the rank constraint of

Eq. (3.20c) to a PSD constraint, the relaxed problem can be written as

$$\min_{\mathbf{W}_1,\cdots,\mathbf{W}_K\in\mathbb{H}^{N\times N}} \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k) \tag{3.21a}$$

$$\text{s.t.} \quad \Pr\left\{\mathrm{Tr}[(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{D}_k] \geq \sigma_k^2\right\} \geq 1 - \epsilon_k \tag{3.21b}$$

$$\mathbf{W}_k \succeq \mathbf{0},\ k = 1, \cdots, K \tag{3.21c}$$

We have to remember that the solutions of $\mathbf{W}_k$ in this relaxed problem may have rank higher than one, so we need to check the solution afterward to see whether solutions to the two problems are identical. However, the non-convex outage probabilistic constraint of Eq. (3.21b) may still present us with difficulties. Examination of Eq. (3.21b) shows that the outage probabilistic constraint is now expressed in terms of the random error matrix $\mathbf{\Delta}_k$, which is the error matrix between $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$. Thus the constraint will depend on the way this difference is measured. As we have discussed in the last two sections, RD is a more precise way to measure the error between two DCC matrices with PSD structure. Therefore, we apply RD as the measurement in our work. The RD between $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$ can be expressed as

$$d_{R_2}(\mathbf{R}_k, \hat{\mathbf{R}}_k) = \sqrt{\mathrm{Tr}\mathbf{R}_k + \mathrm{Tr}\hat{\mathbf{R}}_k - 2\mathrm{Tr}\mathbf{R}_k^{1/2}\hat{\mathbf{R}}_k^{1/2}} \tag{3.22}$$

From Eq. (3.22), we can see that the expression of RD involves terms of square root of $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$, to rewrite this equation, we can assume that $\mathbf{R}_k^{\frac{1}{2}} = \hat{\mathbf{R}}_k^{\frac{1}{2}} + \tilde{\mathbf{\Delta}}_k$, where $\tilde{\mathbf{\Delta}}_k$ can be considered as the mismatch matrix between the square root of DCC matrices. $\tilde{\mathbf{\Delta}}$ should be chosen to guarantee that both $\mathbf{R}_k^{\frac{1}{2}}$ and $\hat{\mathbf{R}}_k^{\frac{1}{2}}$ are PSD matrices. Thus, we

can express Eq. (3.22) as follows:

$$\begin{aligned} d_{\mathrm{R}}^2(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k, \hat{\mathbf{R}}_k) &= \mathrm{Tr}(2\hat{\mathbf{R}}_k + \mathbf{\Delta}_k) - 2\mathrm{Tr}[(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)^{\frac{1}{2}}\hat{\mathbf{R}}_k^{\frac{1}{2}}] \\ &= \mathrm{Tr}[(\mathbf{R}_k^{\frac{1}{2}} - \hat{\mathbf{R}}_k^{\frac{1}{2}})(\mathbf{R}_k^{\frac{1}{2}} - \hat{\mathbf{R}}_k^{\frac{1}{2}})^H] = \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \end{aligned} \tag{3.23}$$

We now apply Eq. (3.23) to reformulate the constraint in (3.21b).

3.4 Reformulation of design constraint

Our objective function is linear in terms of variables $\mathbf{W}_k$. To simplify the problem formulation, we need to transform the constraints further. The constraint we have is a probabilistic constraint which is not convex. The expression of RD involves random variable $\tilde{\mathbf{\Delta}}_k$ which has probabilistic property. We can make reasonable assumptions on these distribution of these random variables and arrive at a probabilistic constraint on the outage. We can then relates the outage probability to the uncertainty bound using Lemma 1. Finally, we can reformulate the constraint by a set of deterministic constraint.

3.4.1 Application of Lemma 1

We first introduce the following lemma [20]

<u>Lemma 1</u>: Given a set $\mathcal{B} \subset \mathbb{C}^{N^2}$ with $\Pr\{\mathbf{y} \in \mathcal{B}\} \geq 1 - \epsilon$ such that $\forall\ \mathbf{y} \in \mathcal{B}$, $\mathbf{y}^H\mathbf{Q}\mathbf{y} + \mathbf{y}^H\mathbf{b} + \mathbf{b}^H\mathbf{y} + c \geq 0$, we have

$$\Pr\{\mathbf{y}^H\mathbf{Q}\mathbf{y} + \mathbf{y}^H\mathbf{b} + \mathbf{b}^H\mathbf{y} + c \geq 0\} \geq 1 - \epsilon \qquad \square$$

Proof. Assume the probability density function of $\mathbf{y}$ is $p(\mathbf{y})$, and $\mathbf{y}^H\mathbf{Q}\mathbf{y} + \mathbf{y}^H\mathbf{b} +$

$\mathbf{b}^H\mathbf{y} + c \geq 0$ is satisfied, for all $\mathbf{y} \in \mathcal{B}$, then

$$\begin{aligned}
&\Pr\{\mathbf{y}^H\mathbf{Q}\mathbf{y} + \mathbf{y}^H\mathbf{b} + \mathbf{b}^H\mathbf{y} + c \geq 0\} \\
&= \int_{\mathbf{y}^H\mathbf{Q}\mathbf{y}+\mathbf{y}^H\mathbf{b}+\mathbf{b}^H\mathbf{y}+c\geq 0} p(\mathbf{y})d\mathbf{y} \\
&\geq \int_{\mathbf{y}\in\mathcal{B}} p(\mathbf{y})d\mathbf{y} \\
&\geq 1 - \epsilon
\end{aligned}$$

The reverse statement in Lemma 1, however, may not be true in general. □

The argument in the outage probability constraint is quadratic form in Lemma 1, so we can rewrite constraint in Eq. (3.21b) to turn it into quadratic form.

$$\begin{aligned}
&\mathrm{Tr}[(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k)\mathbf{D}_k] = \mathrm{Tr}\{(\hat{\mathbf{R}}_k^{\frac{1}{2}} + \tilde{\mathbf{\Delta}}_k)^H\mathbf{D}_k(\hat{\mathbf{R}}_k^{\frac{1}{2}} + \tilde{\mathbf{\Delta}}_k)\} \\
&= \mathrm{Tr}\{\tilde{\mathbf{\Delta}}_k^H\mathbf{D}_k\tilde{\mathbf{\Delta}}_k + \tilde{\mathbf{\Delta}}_k^H\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}} + \hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{D}_k\tilde{\mathbf{\Delta}}_k + \hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}\} \\
&= \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{D}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) + \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) \\
&\quad + \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) + \mathrm{Tr}(\hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})
\end{aligned} \tag{3.24}$$

with $\hat{\mathbf{R}}_k^{\frac{1}{2}H}$ denoting the Hermitian conjugate of $\hat{\mathbf{R}}_k^{\frac{1}{2}}$. Expressing Eq. (3.24) using the $\mathrm{vec}(\cdot)$ function, the outage probability constraint of Eq. (3.21b) can be re-written as

$$\begin{aligned}
&\Pr[\mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{D}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) + \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) \\
&\quad + \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) + \mathrm{Tr}(\hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) - \sigma_k^2 \geq 0] \geq 1 - \epsilon_k
\end{aligned} \tag{3.25}$$

The argument of the probability constraint in Eq. (3.25) is in the same quadratic form as that in Lemma 1, with $\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)$ standing for $\mathbf{y}$.

Motivated by Lemma 1, we define suitable convex sets $\mathcal{B}_k, (k = 1, \cdots, K)$ for which the bounds can be related to the non-outage thresholds in Eq. (3.21) and which allow us to reduce the probabilistic constraints Eq. (3.21b) to deterministic SINR constraints for all the mismatches in the newly defined set.

3.4.2 Choice of set $\mathcal{B}$

From Chapter 2, we know the expression of Riemannian distance d_{R}^2 in Eq. (3.23), according to which, a natural choice for $\mathcal{B}_k$ is a sphere:

$$\begin{aligned}\mathcal{B}_k =&\{\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \in \mathbb{C}^{N^2} \Big| \; d_R^2 \leq \alpha_k^2\} \\ =&\{\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \in \mathbb{C}^{N^2} \Big| \; \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \leq \alpha_k^2\}\end{aligned} \tag{3.26}$$

where the constant α_k^2 is the bound of RD. We need this set to satisfy the condition $\Pr(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \in \mathcal{B}_k) \geq 1 - \epsilon_k$. So α_k^2 should be chosen according to the probability density function of $\tilde{\mathbf{\Delta}}_k$ to make sure that the outage probability constraint is achieved. Therefore, α_k is a function of the outage probability ϵ_k.

3.4.3 Gaussian Assumption for $\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)$

To obtain the relationship between α_k and ϵ_k, we need to calculate the probability of $\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)$ appearing in the set $\mathcal{B}$. To do that, we first assume that $\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)$ is zero-mean complex Gaussian with known covariance matrix $\mathbf{C}_k$, that is to say $\mathrm{vec}(\tilde{\mathbf{\Delta}})$ is $\mathcal{CN}(\mathbf{0}, \mathbf{C})$. The covariance matrix could be defined correspondingly to approximate the channel correlation, no matter whether the error matrix is independent or not.

In the following, we show two cases. In the first case, we assume that $\mathrm{vec}(\tilde{\mathbf{\Delta}})$ is

independent which implies that the covariance matrix $\mathbf{C}_k$ would be a scaled identity matrix, i.e., $\mathbf{C}_k = \sigma_e^2\mathbf{I}$. Thus, we have

$$\|\text{vec}(\tilde{\mathbf{\Delta}}_k)\|^2 = \frac{1}{2\sigma_e^2}\sum_{i=1}^{2N^2} z_i^2 \tag{3.27}$$

where z_i are independent, standard normal variables. Therefore, it is easy to deduce that $2\sigma_e^2\|\text{vec}(\tilde{\mathbf{\Delta}}_k)\|^2$ is Chi-square distribution with $2N^2$ degrees of freedom. The bound α_k can be achieved as

$$\alpha_k = \sqrt{\frac{\Phi^{-1}_{\chi^2_{2N^2}}(1-\epsilon_k)}{2\sigma_e^2}} \tag{3.28}$$

where $\Phi^{-1}_{\chi^2_n}(\cdot)$ is the inverse cumulative distribution function of the Chi-square random variable with n degrees of freedom.

We then consider the second case when the variable $\text{vec}(\tilde{\mathbf{\Delta}}_k)$ is correlated. Here, we may not know the expression of cumulative distribution function directly. However, for a given the covariance matrix, we can apply eigenvalue decomposition to $\mathbf{C}_k$, so that $\mathbf{C}_k = \mathbf{U}_k^H\mathbf{\Lambda}_k\mathbf{U}_k$, where the diagonal elements of $\mathbf{\Lambda}_k$ are denoted by $\lambda_1, \lambda_2, \cdots, \lambda_q$ with $q \leq N^2$, q being the number of non-zero eigenvalues, and $\mathbf{U}_k$ is a unitary matrix.

Let $\mathbf{v}_k = \mathbf{U}_k\mathbf{C}_k^{-\frac{1}{2}}\text{vec}(\tilde{\mathbf{\Delta}}_k)$, we have $\mathbf{v}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ such that

$$\text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) = \mathbf{v}_k^H\mathbf{\Lambda}_k\mathbf{v}_k = \sum_{i=1}^{q}\lambda_i\zeta_i \leq \alpha_k^2 \tag{3.29}$$

where $\zeta_i = |v_{ki}|^2$, $i = 1, \cdots, q$ with v_{ki} being the ith element of $\mathbf{v}_k$. $\mathbf{v}_k$ is the whitened variable and is independently distributed. So that ζ_i is Chi-square distribution with

2 degrees of freedom which is an exponential distribution with probability density function $p(\zeta_i) = e^{-\zeta_i}$ [23]. We also know that ζ_i is independent, so we can express the joint probability density function by the product of each one of them, which is $p(\zeta_1, \zeta_2, \cdots, \zeta_q) = \prod_{i=1}^{q} p(\zeta_i)$. Therefore, the probability P of the condition shown in Eq. (3.29) can be evaluated as

$$\begin{aligned} P &= \int_0^{L_{u_q}} \cdots \int_0^{L_{u_1}} p(\zeta_1, \zeta_2, \cdots, \zeta_q)\ d\zeta_1 \cdots d\zeta_q \\ &= \int_0^{L_{u_q}} \cdots \int_0^{L_{u_1}} p(\zeta_1) \cdots p(\zeta_q)\ d\zeta_1 \cdots d\zeta_q \end{aligned} \tag{3.30}$$

where the lower and upper limits of integration are given by $0 \leq \zeta_i \leq L_{u_i}$.

Integrating consecutively following the index order, in each integral, we consider other variables as constant, so we obtain the expression of the upper bound each time as $L_{u_i} = \frac{1}{\lambda_i}(\alpha_k^2 - \sum_{j=i+1}^{q} \lambda_j \zeta_j)$. For example, we take the first integral in terms of ζ_1, then $L_{u_1} = \frac{1}{\lambda_1}(\alpha_k^2 - \sum_{j=2}^{q} \lambda_j \zeta_j)$, the result of the first integral would be

$$\begin{aligned} \int_0^{L_{u_1}} p(\zeta_1) d\zeta_1 &= \int_0^{L_{u_1}} e^{-\zeta_1} d\zeta_1 = 1 + (-1)e^{-\zeta_1}|_{L_{u_1}} \\ &= 1 + (-1)e^{-\frac{1}{\lambda_1}(\alpha_k^2 - \sum_{i=2}^{q} \lambda_i \zeta_i)} \end{aligned} \tag{3.31}$$

The result as shown in Eq. (3.31) would take part into the next integral in terms of ζ_2. After the ith integral, we will have the expression involving all $i+1, \ldots, q$th variables, and during the process of integral, we can find out that the integral results include terms $e^{-\frac{1}{\lambda_i}(\alpha_k^2 - \sum_{m=i+2}^{q} \lambda_m \zeta_m)}, i = 1, \ldots, q$ and the corresponding coefficients of these terms satisfy recursion. We let c_i represents the coefficient of $e^{-\frac{1}{\lambda_i}(\alpha_k^2 - \sum_{m=i+2}^{q} \lambda_m \zeta_m)}$

after the ith integration. We conclude that c_i satisfies the following regular pattern.

$$
\begin{aligned}
c_1 &= -1 \\
c_2 &= -1 - (-\frac{\lambda_1}{\lambda_2 - \lambda_1})c_1 \\
c_3 &= -1 - (-1)^2 \frac{\lambda_1}{\lambda_2 - \lambda_1} \frac{\lambda_1}{\lambda_3 - \lambda_1} c_1 - (-\frac{\lambda_2}{\lambda_3 - \lambda_2})c_2 \\
&\vdots \\
c_i &= -1 - \sum_{m=1}^{i-1} \left\{ (-1)^{i-m} \left[\prod_{n=m+1}^{i} \frac{\lambda_m}{\lambda_n - \lambda_m} \right] c_m \right\}
\end{aligned} \tag{3.32}
$$

After integrating all q integrations, the result contains terms $e^{-\frac{1}{\lambda_i}\alpha_k^2}$ $(i = 1, \cdots, q)$, the result can be expressed as

$$
P(\alpha_k) = 1 + a_1 e^{-\frac{1}{\lambda_1}\alpha_k^2} + a_2 e^{\frac{1}{\lambda_2}\alpha_k^2} + \cdots + a_q e^{-\frac{1}{\lambda_q}\alpha_k^2} = 1 + \sum_{i=1}^{p} a_i e^{-\frac{\alpha_k^2}{\lambda_i}} \tag{3.33}
$$

where a_i is the corresponding coefficient to each term, and it can be obtained from c_i, the relationship between a_i and c_i would be

$$
a_i = (-1)^{q-i} \prod_{m=i+1}^{q} \frac{\lambda_i}{\lambda_m - \lambda_i} c_i
$$

The probability is a function of α_k, now, according to the outage probability constraint, we have $P(\alpha_k) \geq 1 - \epsilon_k$. Hence, we can write $\alpha_k = P^{-1}(1 - \epsilon_k)$, where $P^{-1}(\cdot)$ is the inverse cumulative distribution function which can be obtained numerically from Eq. (3.33).

Now, we obtain the set $\mathcal{B}$ which satisfies the outage probability constraint, and the condition $\Pr\{\mathbf{y} \in \mathcal{B}\} \geq 1 - \epsilon_k$ can be replaced by the quadratic expression of

Eq. (3.29) with α_k obtained according to the distribution of $\text{vec}(\tilde{\mathbf{\Delta}}_k)$. From Lemma 1, we can conclude that the outage probability constraint of Eq. (3.21b) is satisfied if

$$\text{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{D}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) + \text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) +$$

$$\text{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\text{vec}(\tilde{\mathbf{\Delta}}_k) + \text{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}^{\frac{1}{2}}) - \sigma_k^2 \geq 0 \tag{3.34a}$$

$$\forall \ \ \text{vec}(\tilde{\mathbf{\Delta}}_k)\text{such that} \quad \text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) \leq \alpha_k^2 \tag{3.34b}$$

By applying Lemma 1, we change our outage probabilistic constraint into two deterministic constraints, and as we can see that the two reformulated constraints are both of quadratic form. However, $\text{vec}(\tilde{\mathbf{\Delta}}_k)$ is a random vector, thus, Eq. (3.34) contains an infinite number of constraints.

3.5 Convex design problem

We have reformulated the outage probabilistic constraint into two quadratic constraints for each User k. However, we mentioned that the reformulated constraints are actually infinite constraints which make the problem hard to solve. In this section, we introduce $\mathcal{S}$-lemma to turn the infinite constraints into tractable constraints.

After the reformulation in Section 3.4, our problem has been converted to:

$$\begin{aligned}
&\min_{\mathbf{W}_1,\cdots,\mathbf{W}_K\in\mathbb{H}^{N\times N}} \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k) \\
&\text{s.t. } \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N\otimes\mathbf{D}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)+\mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})+ \\
&\quad \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)+\mathrm{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}^{\frac{1}{2}})-\sigma_k^2\geq 0 \\
&\quad \forall\ \ \mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \text{ such that } \ \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)\leq\alpha_k^2 \\
&\quad \mathbf{W}_k\succeq\mathbf{0} \qquad \text{for } \ k=1,\cdots,K
\end{aligned}$$

In order to reduce the infinite number of constraints in Eq. (3.34) to a tractable set, we can apply the $\mathcal{S}$-lemma [24] here

<u>$\mathcal{S}$-lemma</u>: Let $f:\mathbb{C}^n\to\mathbb{R}$ and $g:\mathbb{C}^n\to\mathbb{R}$ be quadratic functions. We define $\mathcal{S}_g=\{\mathbf{z}|g(\mathbf{z})\geq 0\}$, then we have $f(\mathbf{z})\geq 0\ \forall\ \mathbf{z}\in\mathcal{S}_g$ if and only if there exists $t\geq 0$ such that $f(\mathbf{z})-tg(\mathbf{z})\geq 0\ \forall\ \mathbf{z}\in\mathbb{C}^n$. □

Applying the $\mathcal{S}$-lemma to Eq. (3.34), we let

$$\begin{aligned}
f(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)) =&\mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N\otimes\mathbf{D}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)+\mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})+ \\
&\mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)+\mathrm{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}^{\frac{1}{2}})-\sigma_k^2\geq 0 \\
\text{and } g(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)) = \alpha_k^2-&\mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)\geq 0
\end{aligned} \tag{3.35}$$

if there exist $t_k\geq 0$ such that

$$f(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k))-t_k g(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k))\geq 0$$

and using Schur complement [25], $\mathbf{x}^H\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x}$ iff $\mathbf{A} \succeq \mathbf{0}$, the constraints in Eq. (3.34) are equivalent to:

$$\begin{bmatrix} (\mathbf{I}_N \otimes \mathbf{D}_k) & \mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) \\ \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) & \mathrm{Tr}(\hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) - \sigma_k^2 \end{bmatrix} + t_k \begin{bmatrix} \mathbf{I}_{N^2} & 0 \\ 0 & -\alpha_k^2 \end{bmatrix} \succeq \mathbf{0}$$

$$\text{for} \quad t_k \geq 0 \tag{3.36}$$

Thus, relaxing the rank-one constraint to a PSD constraint together with the above linear matrix inequality, the optimization problem in (3.21) can be written as

$$\begin{aligned} &\min_{\mathbf{W}_1,\cdots,\mathbf{W}_K \in \mathbb{H}^{N\times N}} \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k) \\ &\text{s.t.} \begin{bmatrix} (\mathbf{I}_N \otimes \mathbf{D}_k) + t_k\mathbf{I}_{N^2} & \mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) \\ \mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})^H & \mathrm{Tr}(\hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) - \sigma_k^2 - t_k\alpha_k^2 \end{bmatrix} \succeq \mathbf{0} \\ &\quad \mathbf{W}_k \succeq \mathbf{0} \\ &\quad t_k \geq 0, \qquad \text{for} \quad k = 1, \cdots, K \end{aligned} \tag{3.37}$$

The design problem now becomes a convex optimization problem and can be solved by standard optimization tools (such as CVX [26]). If the solutions to this problem are rank-one, then they are also the solutions to (3.21). If the solutions obtained are of higher ranks, then a random search procedure [27] can be used to obtain the approximate solution. We note, however, that in our simulations, rank one solutions have been obtained in all cases.

3.6 Comparison of absolute error bound approach and probabilistic of outage approach

In the derivations in Sections 3.3 to 3.5, we choose the outage probabilistic robust beamforming approach, while in Section 3.2, a worst-case robust beamforming approach is applied. The two approaches are totally different from the stand point of view, however, we observe that the final convex formulations of the two problems are identical. This observation is shown in this section followed by discussion of the reasons for the two approaches yielding identical results.

As we can see in Eq. (3.14) and Eq. (3.37), the converted constraints in both problems are LMI inequalities with similar formation. In both approaches, due to the application of RD and probabilistic property, we introduce variable $\text{vec}(\tilde{\mathbf{\Delta}}_k)$ and square root of estimated covariance matrix $\hat{\mathbf{R}}_k$. Suppose now, we transform the result of Eq. (3.14) and express it in terms of $\hat{\mathbf{R}}_k^{\frac{1}{2}}$. Letting $\mathbf{q}_k = \text{vec}(\hat{\mathbf{R}}_k^{\frac{1}{2}}) + \text{vec}(\tilde{\mathbf{\Delta}}_k)$, Eq. (3.11) can be rewritten as

$$\begin{aligned}
\min_{\text{vec}(\tilde{\mathbf{\Delta}}_k)} \quad & -\text{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{A}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) - 2\text{Re}\{\text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\mathbf{A}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\} \\
& + \text{Tr}(\hat{\mathbf{R}}_k^{\frac{1}{2}H}\mathbf{A}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) - \sigma_k^2\gamma_k \\
\text{s.t.} \quad & \text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) - \alpha_k^2 \leq 0
\end{aligned} \tag{3.38}$$

(3.38) represents the transformed sub-optimization problem in [22], and by comparing to Eq. (3.34), it is obvious that they are identical if and only if the objective value of the sub-optimization problem is greater or equal to zero. Following the proof of Proposition 1 in Section 3.2, it can be proved that there exists one λ_k in the feasible

set that makes sure the objective value is non-negative. By writing the dual problem and using the Schur complement, the worst-case SINR constraints can be expressed as

$$\begin{bmatrix} -\mathbf{I}_N \otimes \mathbf{A}_k + \lambda_k \mathbf{I}_{N^2} & -\mathrm{vec}(\mathbf{A}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}) \\ -\mathrm{vec}^H(\mathbf{A}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}) & -\mathrm{Tr}\left(\hat{\mathbf{R}}_k^{\frac{1}{2}} \mathbf{A}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}\right) - \lambda_k \alpha_k^2 - \sigma_k^2 \gamma_k \end{bmatrix} \succeq \mathbf{0} \tag{3.39}$$

Comparing Eq. (3.39) to our proposed outage probabilistic constraint of Eq. (3.37) and noting the different notation of $\mathbf{A}_k = -\gamma_k \mathbf{D}_k$, we find that the two different methods lead to the same result. However, we have to mention that in the proposed outage probabilistic method, the error bound α_k is chosen according to the predefined outage probability which guarantees the probabilistic constraint. In worst-case approach, the outage probability has not been taken into account, and that is the main difference between the two methods.

Another reason that leads to the same result is that, in our outage probabilistic robust beamforming method, we use Lemma 1 to replace the outage probability constraint with two quadratic constraints. At the same time, we choose set $\mathcal{B}$ to be a multidimensional ball which makes the two formulations the same. From Lemma 1, it is evident that there might be other sets of different shapes that may also satisfy the outage probabilistic condition. If these sets of other shapes are chosen, then the problem formulation of the two methods would not be the same. We will show such cases in the next chapter.

Chapter 4

Alternative formulation for probabilistic robust beamforming

At the end of last chapter, we mentioned that one of the reasons leading to the similarity of our probabilistic robust beamforming method and worst-case robust beamforming method is that we have chosen a multidimensional ball $\mathcal{B}$ to be the set satisfying the outage probabilistic constraint. In this chapter, we will introduce an alternative set different from the one in last chapter. We can also show that under this new set the probabilistic robust downlink problem can still be converted into a convex problem having a totally different formulation from that in worst-case robust downlink method.

4.1 Approximation of set $\mathcal{B}$

In Section 3.4.2, we choose the set to be a spherical set with radius α_k^2 and make the probability of this set equal to the non-outage probability such that we could use this

set to simplify the constraint in our problem. However, according to Lemma 1, it only requires the chosen set to satisfy the outage probability constraint. Therefore, there are many choices of the set no matter what the assumption we make on distribution of the variable is. The difficulty comes from a good mathematical expression for the set. The spherical set we chose has a good expression and also shows a meaningful explanation that the error is bounded under a Riemannian distance. On the basis of that, we can make an approximation of the set $\mathcal{B}$ by other shapes. One simple approximation of the set $\mathcal{B}$ is a multidimensional cube. We simplify the problem by assuming that the cube is approximately the same as the probability of set $\mathcal{B}$ if the two sets have the same volume.

The equations representing an n-dimensional cube having half its side length a can be expressed as

$$\begin{aligned} -a \leq x_1 \leq a \\ -a \leq x_2 \leq a \\ \vdots \\ -a \leq x_n \leq a \end{aligned} \tag{4.1}$$

where x_i $(i = 1, \ldots, n)$ represents the variable in the ith dimension. As we can see, a multidimensional cube can be represented by a set of inequalities.

The volume of an n-dimensional ball with radius of r can be expressed as $V_{ball} = \frac{2^{n-1}}{n}\pi r^n$ and for an n-dimensional cube, its volume is $V_{cube} = (2a)^n$. We let the two

sets have the same volume, then the relationship between r and a would be

$$V = \frac{2^{n-1}}{n}\pi r^n = (2a)^n$$
$$\text{so that} \quad a = \sqrt[n]{\frac{\pi}{2n}}r \tag{4.2}$$

Eq. (3.26) shows a spherical set of variable $\text{vec}(\tilde{\mathbf{\Delta}}_k)$ with radius α_k. For simplicity, we express the ith element of $\text{vec}(\tilde{\mathbf{\Delta}}_k)$ as ξ_{ki}, $i = 1, \ldots, N^2$. If we denote the half-length of the approximated cube as β_k, according to the relationship in Eq. (4.2), we have

$$\beta_k = \sqrt[N^2]{\frac{\pi}{2N^2}}\alpha_k \tag{4.3}$$

Therefore, the approximated multidimensional cube $\mathcal{C}_k$ can be expressed as a set of inequalities such as

$$\begin{aligned} -\beta_k \leq & \xi_{k1} \leq \beta_k \\ -\beta_k \leq & \xi_{k2} \leq \beta_k \\ & \vdots \\ -\beta_k \leq & \xi_{kN^2} \leq \beta_k \end{aligned} \tag{4.4}$$

The set $\mathcal{C}_k$ is an approximation to set $\mathcal{B}_k$ by assuming that they have the same volume. However, we have to notice that the probability of set $\mathcal{C}_k$ may not be exactly the same as that of set $\mathcal{B}_k$. We make an assumption that the probability of the two sets is almost the same, so that we can use this set $\mathcal{C}_k$ to simplify our problem later. It is reasonable to make such an assumption because there does exist a cube that satisfies the outage probability condition. Due to the distribution of the variable, finding

the exactly expression of the cube may be complicated. The difference between the two sets is small and it can be considered that such a difference would not introduce significant discrepancy on probability of the two sets. In fact, we can also approximate the set $\mathcal{B}$ by other polyhedrons. As an example, let us consider the two-dimensional case in which we can approximate a circle by square, hexagon or octagon and so on. With the increase of the number of edges, the approximated shape is closer and closer to the circle assuming that they have the same area. The same idea can be extended to the n-dimensional case, therefore we can actually find more precise approximations to set $\mathcal{B}$. However, in our work, we will only discuss the simplest case of a cube.

4.2 Uniform Distribution Assumption for vec($\tilde{\mathbf{\Delta}}_k$)

In Section 4.1, we assume that the distribution of $\tilde{\mathbf{\Delta}}_k$ is zero-mean Gaussian, i.e., being the same as what we assume in 3.4.3. Due to the fact that it is complicated to find an exactly set with probability of 1-ϵ_k, we make an approximation of the set $\mathcal{B}_k$ by the set $\mathcal{C}_k$ which has the same volume with $\mathcal{B}_k$. In this section, we discuss the same problem under the consideration of uniform distribution, in which case, no approximation is needed and $\mathcal{B}_k$ and $\mathcal{C}_k$ can have exactly the same probability.

We compare here, the worst-case method and the outage probability method. In the worst-case method, the error uncertainty is upper bounded by a given constant, i.e., $\text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) \leq \alpha_{kw}^2$, where α_{kw} denotes the bound of the error variable in the worst-case method. We assume that vec($\tilde{\mathbf{\Delta}}_k$) is uniformly distributed within this set. The problem reformulation is the same as shown in Eq. (3.14) with α_k replaced by α_{kw}. Now, we consider the second case. A spherical set having radius α_k can be found following the method that we have discussed in Section 3.4.2. The radius α_k

can be calculated according to the non-outage probability ϵ_k and the distribution of $\text{vec}(\tilde{\mathbf{\Delta}}_k)$. Because the distribution is uniform within the set $\text{vec}^H(\tilde{\mathbf{\Delta}}_k)\text{vec}(\tilde{\mathbf{\Delta}}_k) \leq \alpha_{kw}^2$, we can conclude that the ratio of the volume is the same as the ratio of the probability. Therefore, the relationship between α_k and α_{kw} can be derived as follows:

$$\frac{2^{n-1}}{n}\pi\alpha_{kw}^n \times (1-\epsilon_k) = \frac{2^{n-1}}{n}\pi\alpha_k^n \Longleftrightarrow \alpha_k = \sqrt[N^2]{1-\epsilon_k}\alpha_{kw} \tag{4.5}$$

Then, the spherical set $\mathcal{B}_k$ has been chosen with radius α_k of Eq. (4.5). Lemma 1 can be applied and the final problem formulation is the same as Eq. (3.37).

Now, we consider the alternative formulation when uniform distribution is applied. Different from what we discussed in Section 4.1, we can choose a multidimensional cube $\mathcal{C}_k$ which has the same volume and the same probability as the set $\mathcal{B}_k$ without approximation. The reason is that, under uniform distribution assumption, the same volume leads to the same probability no matter what shape we choose. For simplicity, we still choose the cube as an alternative for comparison. Without loss of generality, the relationship of Eq. (4.3) is still valid. Thus, the alternative set $\mathcal{C}_k$ can be expressed as a set of inequalities as shown in Eq. (4.4).

In summary, in Section 4.1, a Gaussian assumption is considered following what we derived in 3.4.3. We notice that under Gaussian distribution, it is complicated to find an exactly set which has the same probability as the original set $\mathcal{B}_k$, so an approximation of the set is applied. However, we can choose a simpler assumption of uniform distribution that makes the probability of the two sets exactly equal. We notice that under uniform distribution, the expression of the alternative set $\mathcal{C}_k$ remains the same, but the complexity of finding the exact set decreases. And we also think it is more convincing for us to use this result as a comparison of the performance of

the two sets $\mathcal{B}_k$ and $\mathcal{C}_k$.

4.3 Reformulation

We replace the set $\mathcal{B}_k$ by the approximated set $\mathcal{C}_k$. According to Lemma 1, the outage probability constraint can be converted into two parts:

$$\begin{aligned}
&\mathrm{vec}^H(\tilde{\boldsymbol{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{D}_k)\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) + \mathrm{vec}^H(\tilde{\boldsymbol{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})+ \\
&\quad \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) + \mathrm{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}^{\frac{1}{2}}) - \sigma_k^2 \geq 0 && (4.6a)\\
&\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) \in \mathcal{C}_k && (4.6b)
\end{aligned}$$

Eq. (4.6a) shows a quadratic constraint in terms of $\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k)$, while Eq. (4.6b) shows a set of inequalities in terms of each element ξ_{ki} of $\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k)$. We want to express the second constraint in quadratic form so that we can transform the two constraints together later on. As shown in Eq. (4.4), the variable $\boldsymbol{\xi}_k$ has N^2 dimensions, and for each dimension, the variable is bounded under a calculated constant. For the ith dimension, the following two expressions are identical for both real and imaginary part

$$-\beta_k \leq \xi_{ki} \leq \beta_k \iff \beta_k^2 - \xi_{ki}^2 \geq 0 \tag{4.7}$$

If we denote $\mathbf{e}_i$ as a vector with the ith element equal to unity and all the other elements equal to zero, we have $\xi_{ki} = \mathbf{e}_i^T\boldsymbol{\xi}_k$. Then, Eq. (4.7) can be expressed as

$$\beta_k^2 - \xi_{ki}^2 \geq 0 \iff \beta_k^2 - \boldsymbol{\xi}_k^H(\mathbf{e}_i\mathbf{e}_i^T)\boldsymbol{\xi}_k \geq 0 \tag{4.8}$$

Denoting $\mathbf{E}_i = \mathbf{e}_i\mathbf{e}_i^T$, the set of inequalities (4.4) can be written as

$$\beta_k^2 - \boldsymbol{\xi}_k^H \mathbf{E}_i \boldsymbol{\xi}_k \geq 0,\ i = 1, \ldots, N^2 \tag{4.9}$$

Therefore, the constraints shown in Eq. (4.4) can be reformulated as

$$\begin{aligned} &\mathrm{vec}^H(\tilde{\boldsymbol{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{D}_k)\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) + \mathrm{vec}^H(\tilde{\boldsymbol{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) + \\ &\quad \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) + \mathrm{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}^{\frac{1}{2}}) - \sigma_k^2 \geq 0 \qquad (4.10a) \\ &\forall\ \mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) \text{ such that } \beta_k^2 - \mathrm{vec}^H(\tilde{\boldsymbol{\Delta}}_k)\mathbf{E}_i\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k) \geq 0 \qquad (4.10b) \\ &\quad i = 1, \ldots, N^2 \end{aligned}$$

where $\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k)$ replaces $\boldsymbol{\xi}_k$. Eq. (4.10) shows a set of quadratic constraints in terms of $\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k)$, but there are still an infinite number of constraints.

4.4 Convex design problem

Now, we have converted the outage probability constraint into a set of quadratic constraints. The reason to express the constraints in terms of $\mathrm{vec}(\tilde{\boldsymbol{\Delta}}_k)$ is that this will give us added advantage when reformulating them into a convex design. In Section 3.5, we apply $\mathcal{S}$-lemma to do the reformulation such that we have two quadratic constraints and they can be combined together according to the lemma. In this section, we can apply the similar idea again by using $\mathcal{S}$-lemma. However, due to the fact that right now we have more than two quadratic constraints, we need to apply the extension of $\mathcal{S}$-lemma. Let us recall what $\mathcal{S}$-lemma says:

<u>$\mathcal{S}$-lemma</u>: Let $f : \mathbb{C}^n \to \mathbb{R}$ and $g : \mathbb{C}^n \to \mathbb{R}$ be quadratic functions. We define

$\mathcal{S}_g = \{\mathbf{z}|g(\mathbf{z}) \geq 0\}$, then we have $f(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathcal{S}_g$ if and only if there exists $t \geq 0$ such that $f(\mathbf{z}) - tg(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathbb{C}^n$.

Corollary (Extension of $\mathcal{S}$-lemma): Let $f : \mathbb{C}^n \to \mathbb{R}$ and $g_i(i = 1, \ldots, N) : \mathbb{C}^n \to \mathbb{R}$ be quadratic functions. We denote $\mathcal{S}_{g_i} = \{\mathbf{z}|g_i(\mathbf{z}) \geq 0\}$. Then we have $f(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \cap\mathcal{S}_{g_i}$ if and only if there exist $t_1, \ldots, t_N \geq 0$ such that $f(\mathbf{z}) - \sum_{i=1}^{N} t_i g_i(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathbb{C}^n$.

Proof. According to $\mathcal{S}$-lemma, we have $f(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathcal{S}_{g_1}$ if and only if there exists $t_1 \geq 0$ such that $f(\mathbf{z}) - t_1 g_1(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathbb{C}^n$. If we let $f_1(\mathbf{z}) = f(\mathbf{z}) - t_1 g_1(\mathbf{z})$, then we have $f_1(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathcal{S}_{g_2}$ if and only if there exists $t_2 \geq 0$ such that $f_1(\mathbf{z}) - t_2 g_2(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathbb{C}^n$. Following the same idea until $i = N$, if we let $f_N(\mathbf{z}) = f_{N-1}(\mathbf{z}) - t_N g_N(\mathbf{z})$, then we have $f_N(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathcal{S}_{g_N}$ if and only if there exists $t_N \geq 0$ such that $f_N(\mathbf{z}) - t_N g_N(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathbb{C}^n$. Therefore, if we denote $\mathcal{S} = \cap\mathcal{S}_{g_i}$, then we have $f(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathcal{S}$ if and only if there exist $t_1, \ldots, t_N \geq 0$ such that $f_{N-1}(\mathbf{z}) - t_N g_N(\mathbf{z}) = f(\mathbf{z}) - \sum_{i=1}^{N} t_i g_i(\mathbf{z}) \geq 0 \ \forall \ \mathbf{z} \in \mathbb{C}^n$.

□

Eq. (4.10) shows $N^2 + 1$ quadratic constraints in terms of $\mathbf{z}$, therefore we can consider them as

$$
\begin{aligned}
f_1(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)) &= \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)(\mathbf{I}_N \otimes \mathbf{D}_k)\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) + \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathrm{vec}(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}}) + \\
&\quad \mathrm{vec}^H(\mathbf{D}_k\hat{\mathbf{R}}_k^{\frac{1}{2}})\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) + \mathrm{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H}\mathbf{D}_k\hat{\mathbf{R}}^{\frac{1}{2}}) - \sigma_k^2 \geq 0 \\
g_i(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)) &= \beta_k^2 - \mathrm{vec}^H(\tilde{\mathbf{\Delta}}_k)\mathbf{E}_i\mathrm{vec}(\tilde{\mathbf{\Delta}}_k) \geq 0 \quad i = 1, \ldots, N^2
\end{aligned}
$$

Applying the Extension of $\mathcal{S}$-lemma, we have the recursive relationship $f_i(\mathrm{vec}(\tilde{\mathbf{\Delta}}_k)) =$

$f_{i-1}(\text{vec}(\tilde{\mathbf{\Delta}}_k)) - t_{k(i-1)}g_{i-1}(\text{vec}(\tilde{\mathbf{\Delta}}_k))$. Finally, we have quadratic function as

$$f_{N^2+1}(\text{vec}(\tilde{\mathbf{\Delta}}_k)) = f_1(\text{vec}(\tilde{\mathbf{\Delta}}_k)) - \sum_{i=1}^{N^2} t_{ki} g_i(\text{vec}(\tilde{\mathbf{\Delta}}_k)) \geq 0 \tag{4.11}$$

where t_{ki} can be found in every step when we apply the $\mathcal{S}$-lemma for User k.

Using Schur complement, Eq. (4.11) can be expressed as:

$$\begin{bmatrix} (\mathbf{I}_N \otimes \mathbf{D}_k) & \text{vec}(\mathbf{D}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}) \\ \text{vec}^H(\mathbf{D}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}) & \text{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H} \mathbf{D}_k \hat{\mathbf{R}}^{\frac{1}{2}}) - \sigma_k^2 \end{bmatrix} + \sum_{i=1}^{N^2} t_{ki} \begin{bmatrix} \mathbf{E}_i & \mathbf{0} \\ \mathbf{0}^T & -\beta_k^2 \end{bmatrix} \succeq \mathbf{0} \tag{4.12}$$

where $t_{ki} \geq 0$.

Thus, after dropping the rank-one condition, the optimization problem under an approximated cube can be written as

$$\begin{aligned} &\min_{\mathbf{W}_1, \cdots, \mathbf{W}_K \in \mathbb{H}^{N \times N}} \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k) \\ &\text{s.t.} \begin{bmatrix} (\mathbf{I}_N \otimes \mathbf{D}_k) & \text{vec}(\mathbf{D}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}) \\ \text{vec}^H(\mathbf{D}_k \hat{\mathbf{R}}_k^{\frac{1}{2}}) & \text{Tr}(\hat{\mathbf{R}}^{\frac{1}{2}H} \mathbf{D}_k \hat{\mathbf{R}}^{\frac{1}{2}}) - \sigma_k^2 \end{bmatrix} + \sum_{i=1}^{N^2} t_{ki} \begin{bmatrix} \mathbf{E}_i & \mathbf{0} \\ \mathbf{0}^T & -\beta_k^2 \end{bmatrix} \succeq \mathbf{0} \\ &\quad \mathbf{W}_k \succeq \mathbf{0} \\ &\quad t_{ki} \geq 0, \qquad \text{for} \quad k = 1, \cdots, K \end{aligned} \tag{4.13}$$

Eq. (4.13) shows a convex optimization problem with linear objective function and LMI constraints and PSD constraints. It can be solved by optimization tools, such as CVX.

At the end of last chapter, we have a discussion of the comparison between our proposed outage probabilistic downlink beamforming approach and the worst-case

downlink beamforming approach. Both of the two methods apply Riemannian distance as the measurement of DCC matrices, and, after reformulation into convex optimization forms, the two problems appeared to be identical, yielding identical performance beamformers. This, however, is not the case in general. The two problems started from two different standpoints, and their approaches are fundamentally different. In this chapter, we reveal the difference between the two approaches by choosing another set $\mathcal{C}_k$ which is different from the spherical set $\mathcal{B}_k$, and a convex optimization problem is developed under such a set. Comparing Eq. (4.13) to Eq. (3.37), we can observe that the two optimization problems are different even though we apply similar idea in the process of convex reformulation. Therefore, we have demonstrated that the outage probabilistic downlink beamforming problem is essentially different from the worst-case downlink beamforming problem. In this chapter, we first consider the Gaussian distribution case where the set $\mathcal{C}_k$ is an approximation to the set $\mathcal{B}_k$, but we have to notice that there exists such a multidimensional cube that exactly satisfies the outage probability constraint. Due to the reason that finding the exact half-length of the cube is complicated under the Gaussian variable assumption, we make an approximation that equal volumes of the hypersphere and the hypercube yield the same probability of outage. However, such an assumption would not change the formation of Eq. (4.13), the only difference is that the value of β_k contains a slight error. To make our statement more convincing, a uniform distribution is considered where the set $\mathcal{C}_k$ has the same volume and the same probability as the set $\mathcal{B}_k$. In this case, there is no approximation error in the value of β_k. The expressions of the set $\mathcal{B}_k$ and $\mathcal{C}_k$ remain the same as those we derived in the Gaussian assumption case. The problem reformulation in each set remains the same for different distribution

assumptions. A multidimensional cube is a simple approximation of set $\mathcal{B}_k$, besides that, there are many choices of such a set that could be made use of to turn the outage probabilistic constraint into a deterministic one. Different choices of sets will lead to different reformulations of the original optimization problem yielding different optimum beamforming algorithms.

Chapter 5

Numerical Experiments

In this chapter, we show several examples in our computer simulations illustrating the performance of our proposed robust downlink beamforming algorithms in Chapter 3 and Chapter 4. We consider two different assumptions in the simulation of Chapter 3. First, we assume that the elements of the error matrix are independent, complex Gaussian variables with covariance being the same scaled identity matrix. Then, we consider the case that the element of the error matrix is correlated having a given covariance matrix. Under both assumptions, the performances of our proposed algorithm are compared with those of (i) the non-robust ($\mathbf{R}_k$ is exactly known), (ii) robust beamformer (error bounded by Frobenius norm) [13], and (iii) robust outage probability beamformer (error measured in Euclidean metric) [18]. Simulations of the proposed alternative approach discussed in Chapter 4 are compared with those in Chapter 3.

In all cases, we choose the same scenario as that in [28] for which there are $K = 3$ users served by a single BS having $N = 6$ transmitter antennas. The users are located at $\theta_1 = 10°$, $\theta_2 = 10° + \phi$ and $\theta_3 = 10° + 2\phi$, where ϕ is the separation angle between

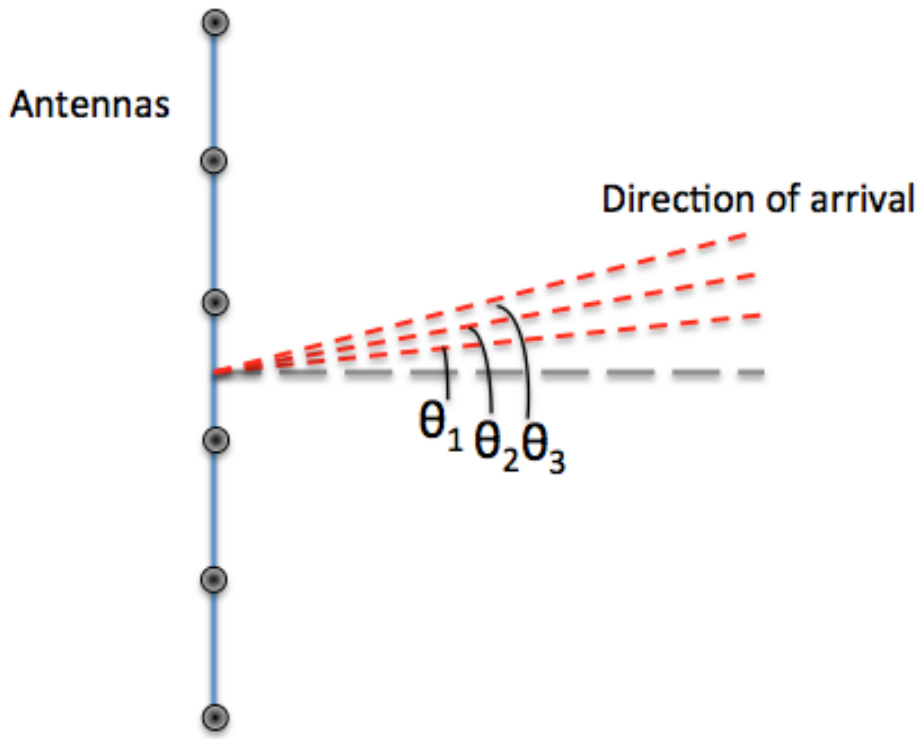


Figure 5.1: Direction of arrival

two users as shown in Fig. 5.1. Each user is surrounded by a large number of local scatterers corresponding to a standard deviation spread angle of $\sigma_\theta = 1°$. The channel covariance matrix perturbed by the dispersion is given by

$$[\mathbf{R}(\theta, \sigma_\theta)]_{kl} = e^{j\pi(k-l)\sin\theta} e^{-\frac{(\pi(k-l)\sigma_\theta \cos\theta)^2}{2}} \tag{5.1}$$

For convenience, we choose the SINR threshold to be the same for all users and the channel is infested with additive zero-mean white Gaussian noise with the same variance $\sigma_k^2 = 0.1$. The outage probability threshold is chosen as $\epsilon_k = 0.1$ for all $k = 1, \cdots, K$.

5.1 Simulations of Methods in Chapter 3

In this section, we first show the simulation results of the first case in Section 3.4.3. The error matrix $\tilde{\mathbf{\Delta}}_k$ is assumed to be zero-mean complex Gaussian with covariance $\tilde{\mathbf{C}}_k = \sigma_e^2 \mathbf{I}_N$, where σ_e^2 denotes the error variance. The bound α_k of the set $\mathcal{B}_k$ can be calculated according to Eq. (3.28). We choose error variance to be $\sigma_e^2 = 0.02$ and $\sigma_e^2 = 0.025$ respectively, so that we can compare the performances of our proposed method in the two cases.

Fig. 5.2 examines the variation of transmission power against the angle of separation between the users. The separation angle ϕ is varied from 7° to 10° while the required SINR level γ_k is set at 2dB. It can be observed that the transmission power of all methods decreases as the separation angle is increased. It can also be observed that the proposed robust design using RD requires lower transmission power than all the other robust beamforming methods.

We define the normalized QoS as

$$\eta_k = \frac{\mathbf{w}_k^H (\hat{\mathbf{R}}_k + \mathbf{\Delta}_k) \mathbf{w}_k}{\gamma_k \sum_{\ell=1, \ell \neq k}^{K} \mathbf{w}_\ell^H (\hat{\mathbf{R}}_k + \mathbf{\Delta}_k) \mathbf{w}_\ell + \gamma_k \sigma_k^2} \tag{5.2}$$

For a beamformer, satisfying the QoS constraint requires $\eta_k \geq 1$ whereas minimization of transmission power calls for η_k to be as close to unity as possible. Fig. 5.3 shows the distributions of the value of η_k for the different beamformers in 1500 trials in which γ_k is set at 2dB. It can be observed that in the case of the non-robust technique only about 50% of the trials have the QoS constraint satisfied, while the other three robust techniques have the QoS constraint satisfied in all the trials. Further, the robust Frobenius method [13] is comparatively conservative having all the

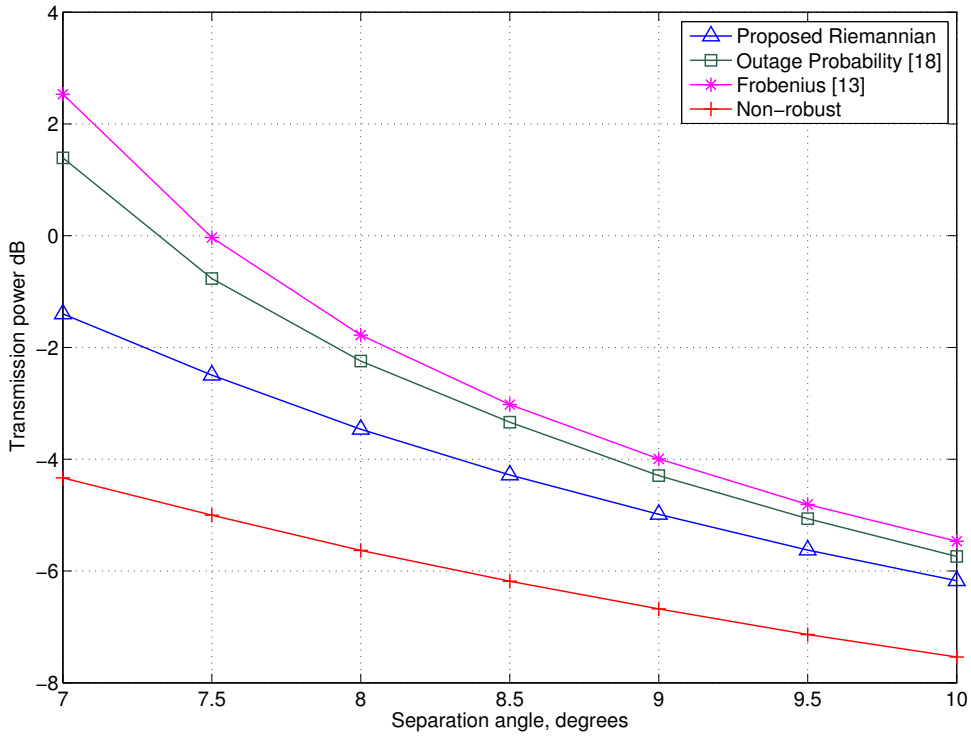


Figure 5.2: Variation of transmission power with the separation angle, $\sigma_e^2 = 0.02$

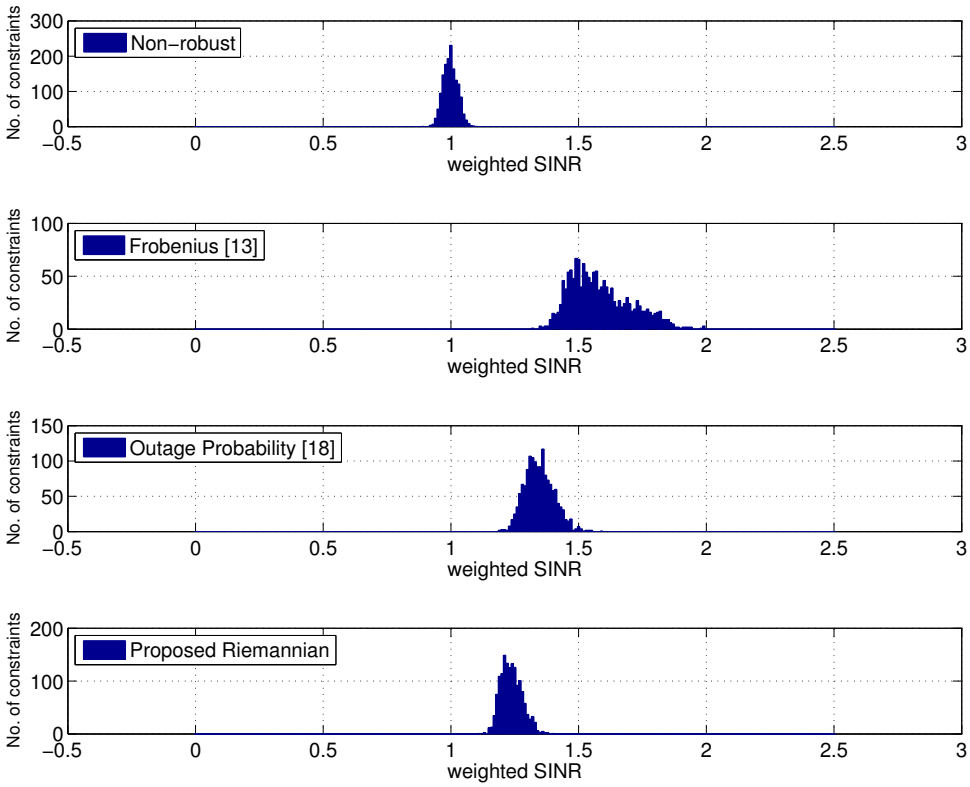


Figure 5.3: Histogram of weighted SINR for $\phi = 7°$, $\sigma_e^2 = 0.02$

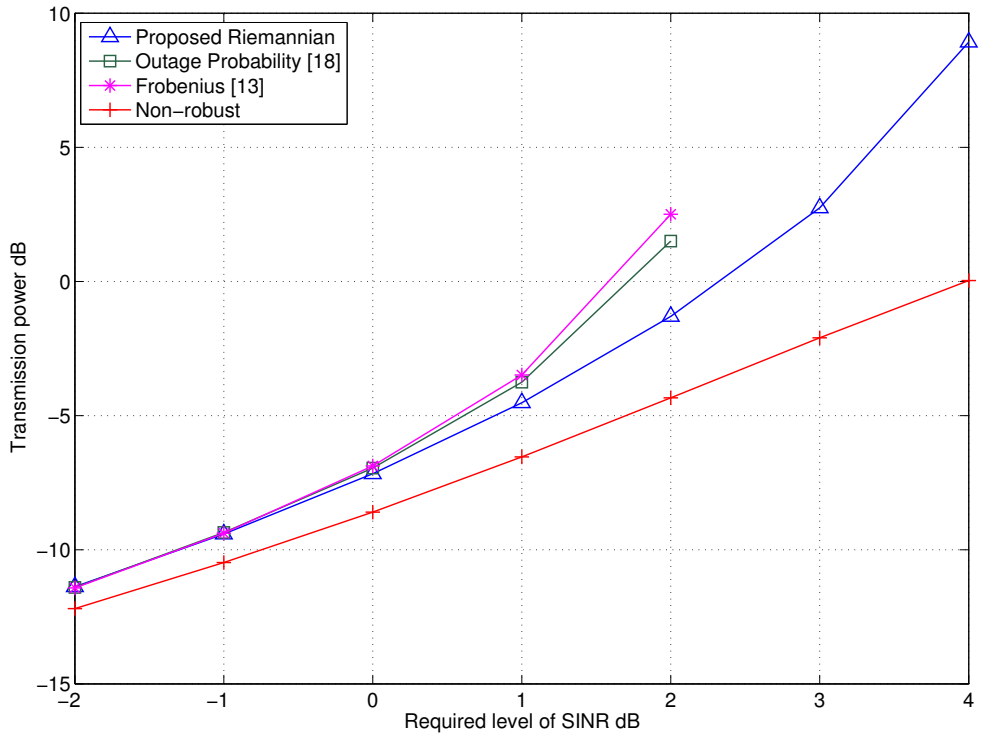


Figure 5.4: Variation of transmission power with QoS requirement, $\sigma_e^2 = 0.02$

values of η_k substantially larger than 1 while the proposed design using RD has the distribution closest to unity.

Fig. 5.4 depicts the variation of the transmission power with the required level of SINR. Here, all γ_k are kept equal in each trial but vary together from -2dB to 4dB from trial to trial. Again, it can be observed that the non-robust beamformer (assuming perfect CSI) requires the lowest transmission power. On the other hand, among the robust beamformers, our algorithm using RD yields the lowest power requirement, whereas the robust beamformer using Frobenius error bound requires the highest transmission power.

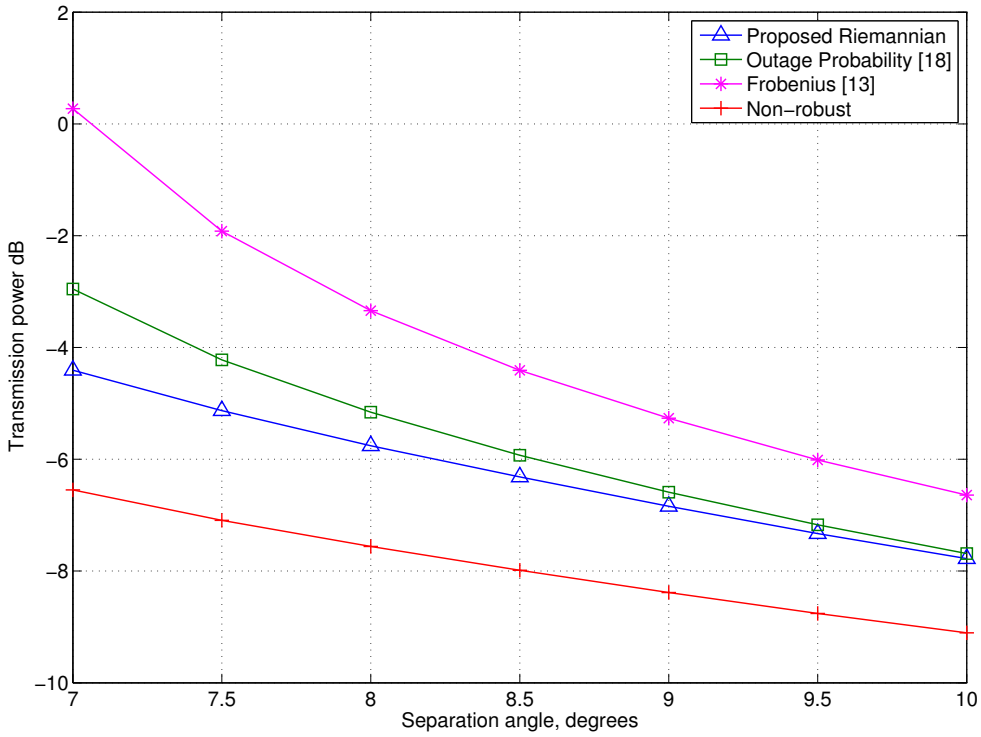


Figure 5.5: Variation of transmission power with the separation angle, $\sigma_e^2 = 0.025$

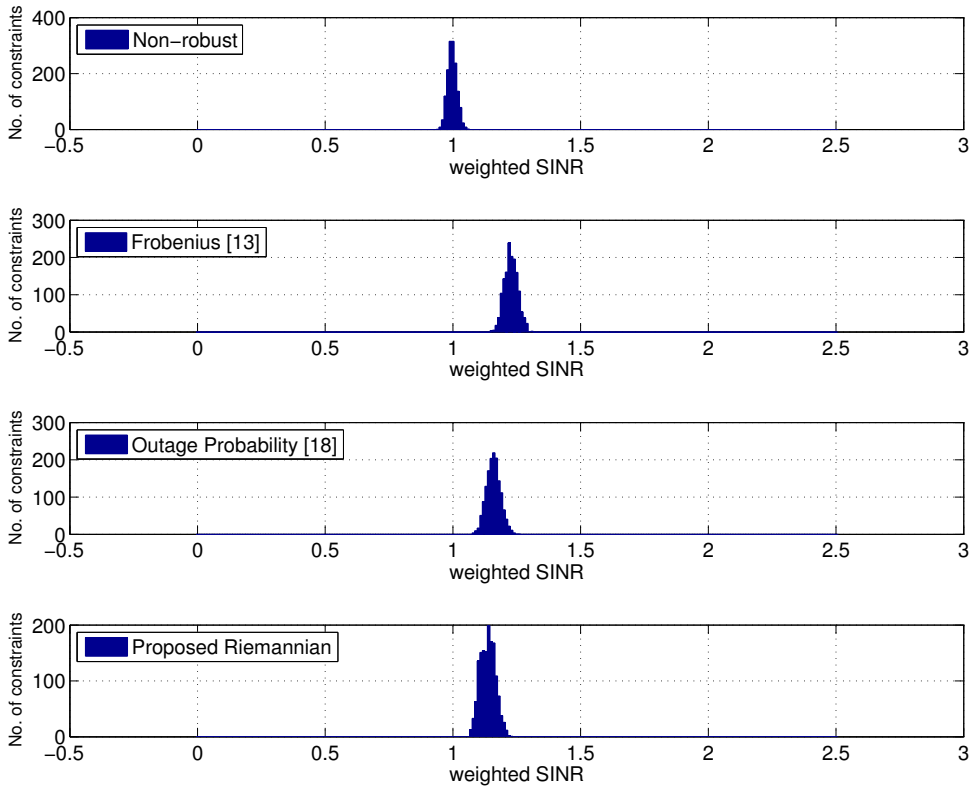


Figure 5.6: Histogram of weighted SINR for $\phi = 7°$, $\sigma_e^2 = 0.025$

We change the error variance to be $\sigma_e^2 = 0.025$ and do the same simulations. Fig. 5.5 shows the variation of the transmission power against the angle separation with the required SINR equal to 1dB. The separation angle is set from 7° to 10°. We can observe the same variation of the transmission power as what we observed in Fig. 5.2. The required transmission power for all methods decreases with the increase of separation angle. It can be also observed that in a higher error variance case, our proposed robust beamforming method still shows better performance in transmission power than other robust beamforming methods.

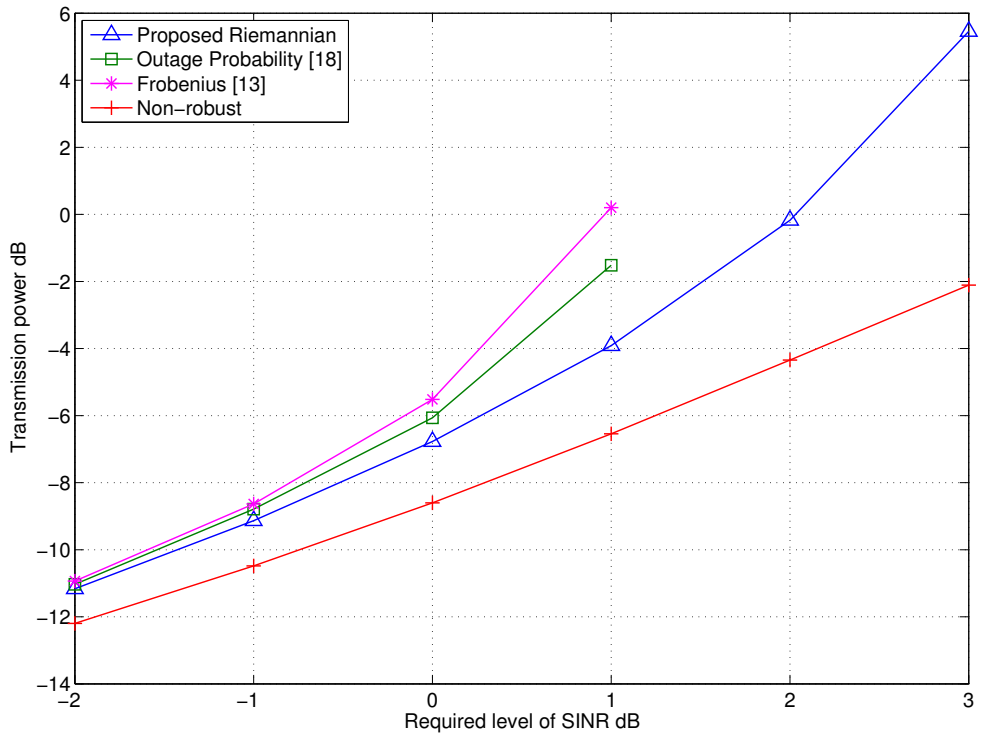


Figure 5.7: Variation of transmission power with QoS requirement, $\sigma_e^2 = 0.025$

Fig. 5.6 shows the histogram of the weighted SINR in which γ_k is set at 1dB. From the figure, it can be observed that the central value of our proposed method is closest to unity. The Frobenius method [13] is comparatively conservative with all the values larger than unity. In the case of non-robust method, only about 50% of the values satisfy the QoS constraint. Therefore, even though the non-robust method shows a minimum transmission power requirement in Fig. 5.5, it can not guarantee the QoS satisfaction of all the users.

Fig. 5.7 describes the variation of the transmission power against SINR threshold. The required SINR is varied from -2dB to 3dB, and in each trial, all γ_k are kept the

same. It can be observed that, similar to Fig. 5.4, except the non-robust method, our proposed method requires the lowest transmission power among the three robust beamforming methods. Compare to Fig. 5.4, we will find that for a fixed level of SINR, with the increase of error variance, the required transmission power increases for all methods. What is more, as the required SINR increases, the possibility that there is no feasible solution for Frobenius method [13] and outage probability method [18] increases. In Fig. 5.4, the maximum feasible level of SINR for Frobenius method [13] and outage probability method [18] is 2dB, while in Fig. 5.7, the level drops to 1dB due to the higher error variance. Both simulations (Fig. 5.4 and Fig. 5.7) show that our proposed robust design not only requires less transmission power but also works better at higher SINR requirement.

Fig. 5.8 shows a comparison of the performances of our proposed method under 4 given error variances. As the increasing of error variance, at a given SINR, it can be observed that the required transmission power increases. When the required level of SINR is low, the increase of error variance does not have a significant effect on the required transmission power. However, at a high level of SINR, we can observe a significant increase on transmission power. When the error variance is large enough, for example, 5dB as the Fig. 5.8 shows, there is no feasible solution in our proposed method.

Now, we consider the second assumption in Section 3.4.3, in which the error matrix $\tilde{\mathbf{\Delta}}_k$ is assumed to be correlated Gaussian with covariance matrix $\mathbf{C}_k$. The covariance matrix $\mathbf{C}_k$ should satisfy $\mathbf{C}_k = \mathbf{\Omega}_k \otimes \mathbf{\Sigma}_k$, where $\mathbf{\Omega}_k$ and $\mathbf{\Sigma}_k$ are positive definite matrices. The bound α_k can be obtained numerically according to Eq. (3.33). Fig. 5.9 depicts the variation of transmission power with the separation angle. Because

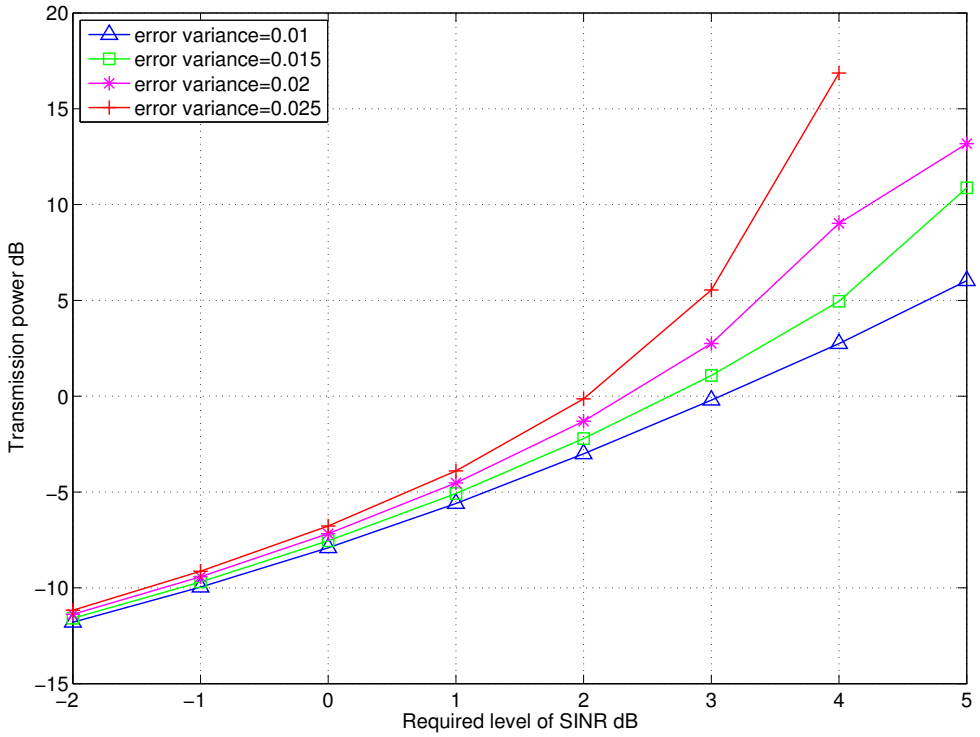


Figure 5.8: Variation of transmission power with QoS requirement

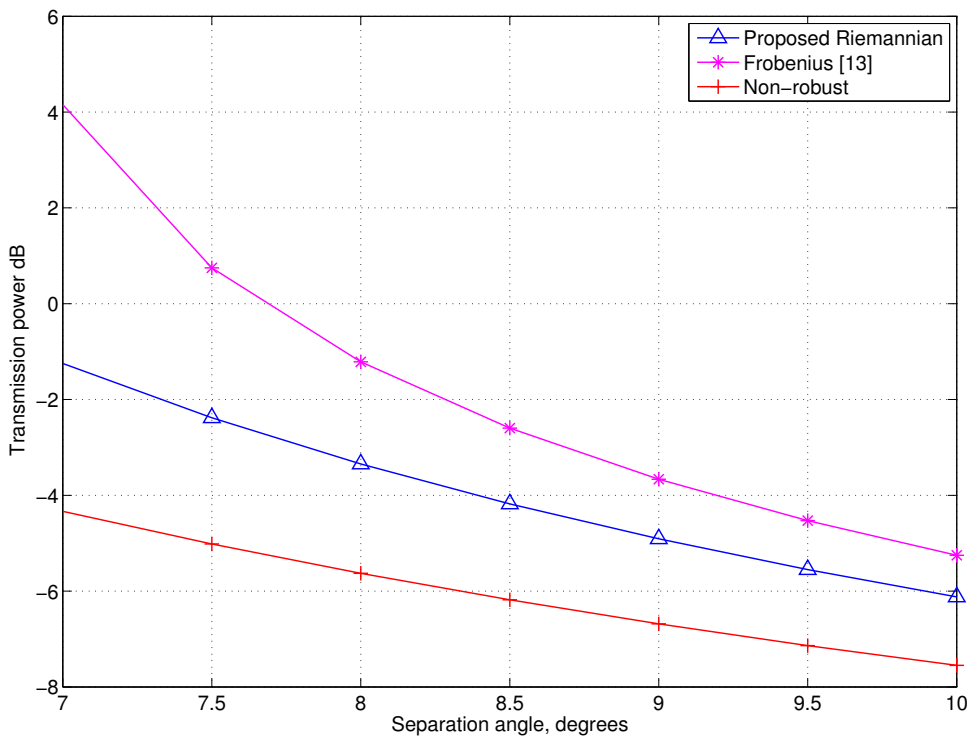


Figure 5.9: Variation of transmission power with the separation angle, correlated error matrix

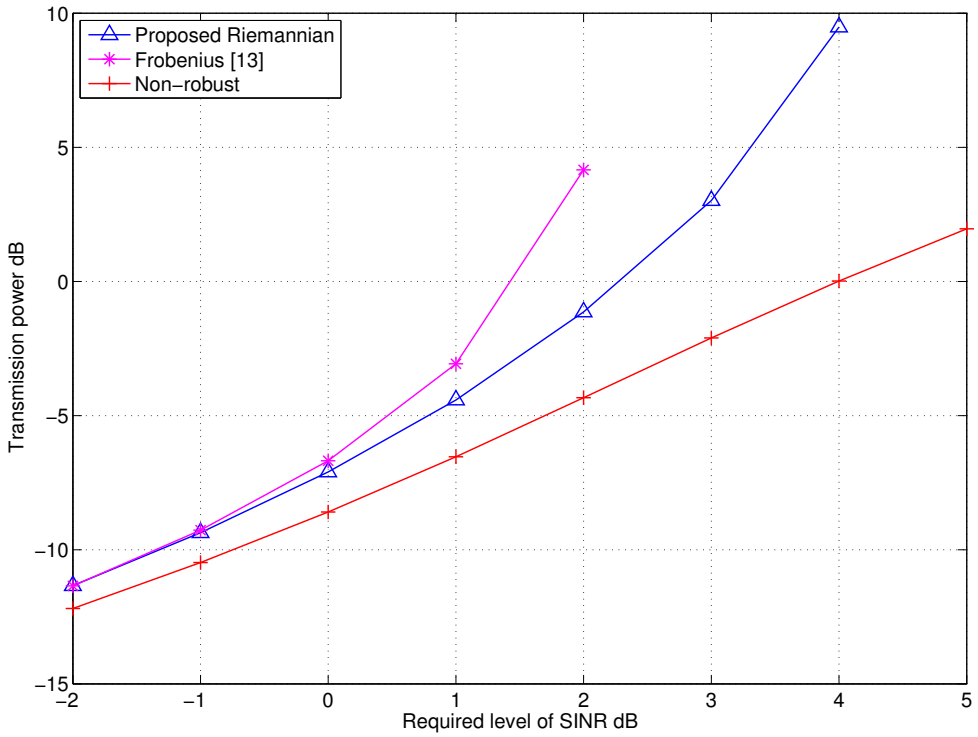


Figure 5.10: Variation of transmission power with QoS requirement, correlated error matrix

in [18], the error matrix is assumed to be independent, so here we just compare the performance of our proposed method and Frobenius method [13]. It can be observed that the transmission power of our proposed robust beamforming method decreases with the increase of separation angle. Fig. 5.10 shows the required transmission power against the SINR requirement. Comparing to the Frobenius method [13], our proposed method has a better performance.

5.2 Simulations of Methods in Chapter 4

In last section, the simulation results show the performances of the method we proposed in Chapter 3, where a multidimensional ball $\mathcal{B}_k$ is applied. In Chapter 4, we choose an alternative set $\mathcal{C}_k$ which is an approximation cube set of $\mathcal{B}_k$ to reformulate the robust beamforming problem. It turns out that the problem formulations under the two different sets are different. In this section, we will show some simulation results following the idea of Chapter 4 and compare the results with that in Chapter 3.

To make fair comparison, we generate the channel covariance matrix in the same way as shown in Eq. (5.1). The error matrix $\tilde{\mathbf{\Delta}}_k$ is assumed to be zero-mean complex Gaussian with covariance $\tilde{\mathbf{C}}_k = 0.02\mathbf{I}_N$. The half-length of the cube β_k can be calculated according to Eq.(4.3). The SINR requirement is varied from -5dB to 3dB, and the separation angle is set at $\phi = 7°$. Fig. 5.11 shows the variation of the transmission power against the required SINR for both sets. From the figure, it can be observed that the simulation results for the two sets have a small difference. The performance under set $\mathcal{B}_k$ is marginally better than that under set $\mathcal{C}_k$. And it can be also observed that under both sets, the performance of our proposed method is better than the performances of [13] and [18] in Fig.5.4.

Fig. 5.12 shows the performances of the power variation against required level of SINR. In this simulation, the error matrix $\tilde{\mathbf{\Delta}}_k$ is assumed to be uniformly distributed. We compare the performance of the worst-case approach proposed in [22] and our proposed method. It can be observed that the performances of our proposed methods using the multidimensional ball set and the multidimensional cube set are better than that of the worst-case approach. Similar result can be also observed that the

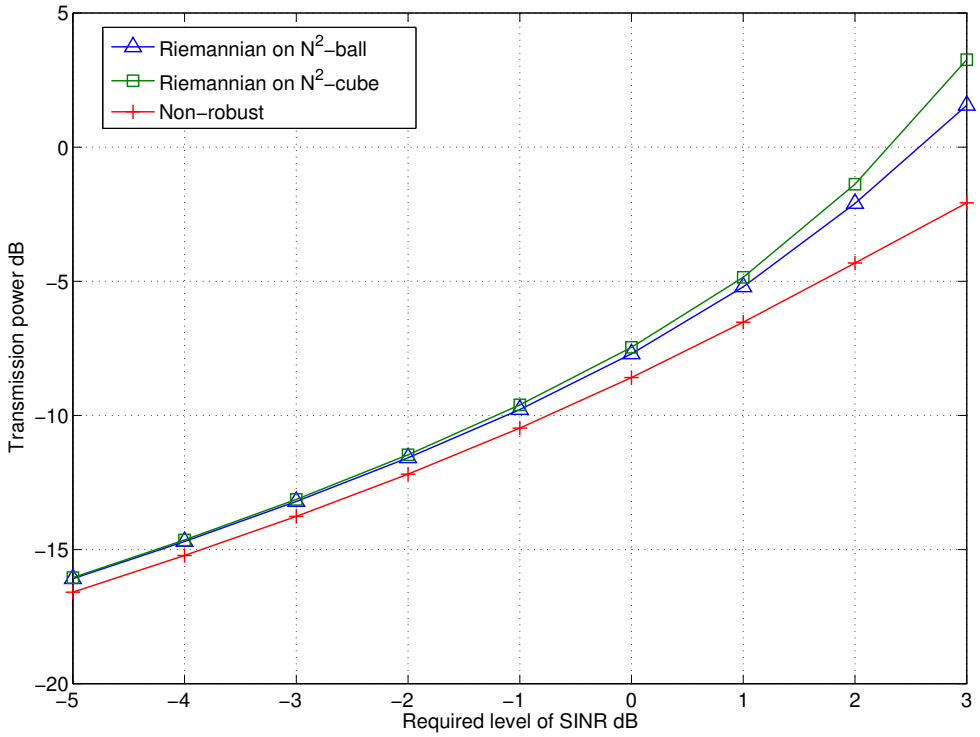


Figure 5.11: Comparison between $\mathcal{B}_k$ and $\mathbf{C}_k$

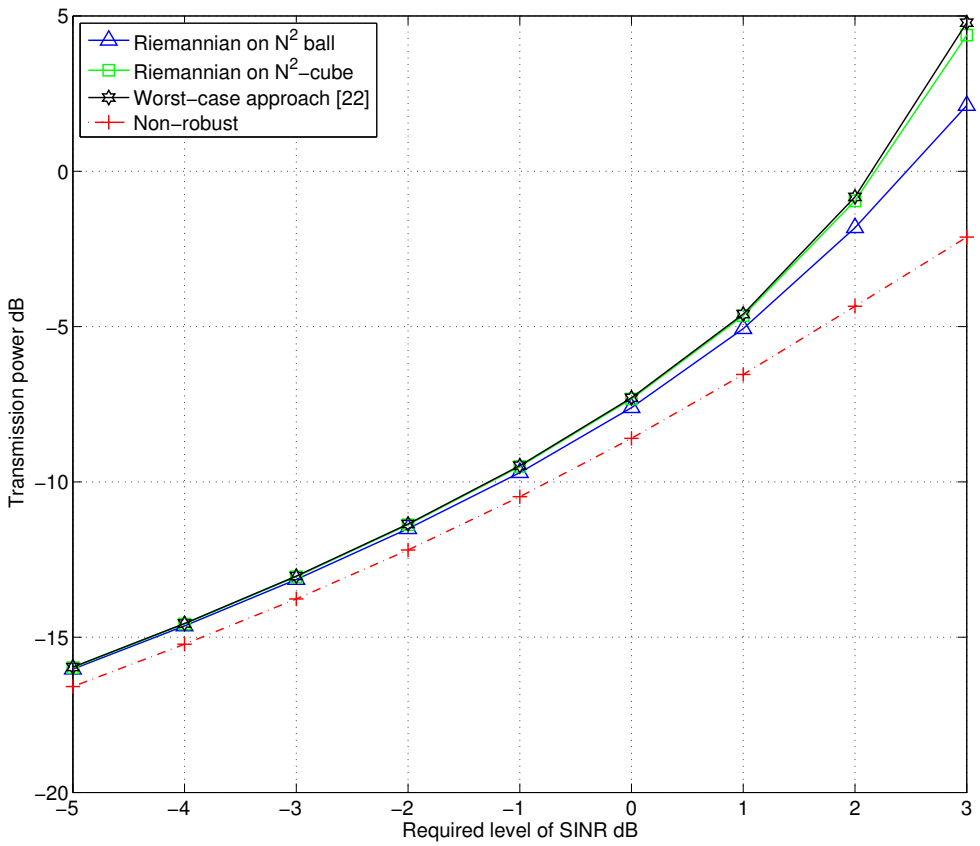


Figure 5.12: Power variation based on uniform distribution

performance of the ball set is better than that of the cube set even though under the uniform distribution the two sets are exactly the same in terms of both volume and probability.

Chapter 6

Conclusion and Future work

In this thesis we examine the robust downlink beamforming design from a statistical point of view guaranteeing that the QoS constraints are satisfied with an imposed non-outage probability. Furthermore, by reasoning that the estimated DCC matrices are Hermitian and PSD thereby forming a manifold in the signal space, we use the measure of RD for the estimation errors instead of the commonly used ED in the formulation of our problem. Applying this measure to our design constraint, we transform the design into a convex optimization problem the solution of which can be obtained efficiently using standard methods. We consider two different distribution assumptions for the error matrix, Gaussian distribution and uniform distribution. In both cases, we derive the convex reformulations under the multidimensional ball set and the multidimensional cube set. Simulation results show that the performance of our design is superior to those of other robust beamformers recently developed.

In this thesis we apply the second type of RD as our measure because of the convenience of computation. In the future, we can try to apply the first type of RD to solve our problem. In Chapter 4, we provide two distribution assumptions

and claim that different distribution assumptions will lead to the different accuracy of the new set. However, we do not know which kind of distribution is the best approximation to the real scenario. Thus, it would be a good consideration of our future research to find a distribution true to the real situation. We have proposed two sets in our thesis among many possible choices of the set. In our future work, we can try to find other sets which satisfy the outage probability constraint and derive the corresponding convex reformulations.

Bibliography

[1] M. Bengtsson and B. Ottersten, *Handbook of Antennas in Wireless Communications: Optimal and Suboptimal Transmit Beamforming*, Boco Raton, FL:CRC, 2001.

[2] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?", *IEEE Trans. on Inf. Th.*, vol. 49, no. 4, pp. 951-963, April 2003.

[3] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?", *IEEE Comm. Mag.*, vol. 42, no. 10, pp. 54-59, Oct. 2004.

[4] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aero. Electron. Syst.,* vol. 24, pp. 297-401, July 1998.

[5] A.B. Gershman, "Robust adaptive beamforming in sensor array," *Int. Journ. Elec. Comm.,* vol. 53, p. 395-409, Dec. 1999.

[6] M. B. Shenouda and T. N Davidson, "Conic convex formulations of robust downlink precoder designs with quality of service constraints", *IEEE J. Sel. Topics in Sig. Proc.*, vol. 1, pp. 714-724, Dec. 2007.

[7] J. Wang and D. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge", *IEEE Trans. in Sig. Proc.*, vol. 57, no. 8, Aug. 2009.

[8] G. Zheng, K. K. Wong and T. S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions", *EURASIP J. on Advances in Sig. Proc.*, pp. 1-15, Jun. 2008.

[9] E. Song, Q. Shi, M. Sanjabi, R. Sun, Z. Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?", *Proc. IEEE ICASSP'2011*, pp 3096-3099, May 2011.

[10] J.Wang and M. Payaro, "Is transmit beamforming robust?", *Proc. IEEE Int. Conf. on Comm*, pp 1-5, Jun 2011.

[11] J. Wang, M Bengtsson, B. Ottersten, D. Palomar, "Robust maximin MIMO precoding for arbitrary convex uncertainty sets, *Proc. IEEE ICASSP'2012*, pp. 3045-3048 , March 2012.

[12] V. Sharma, I. Wajid, A. Gershman, H. Chen and S. Lambotharan, "Robust downlink beamforming using positive semidefinite covariance constraints", *IEEE Int. ITG Workshop on Smart Antennas*, WSA, Feb. 2008.

[13] I. Wajid, Y. C Eldar, A. Gershman. "Robust downlink beamforming using covariance channel state information", *Proc. IEEE ICASSP'2009*, pp. 2285-2288 , April 2009.

[14] Law, K.L.; Wajid, I.; Pesavento, M., "Robust downlink beamforming in multi-group multicasting using trace bounds on the covariance mismatches," *Proc. IEEE ICASSP'2012*, pp.3229,3232, 25-30 March 2012

[15] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems", *IEEE Trans. in Sig. Proc.*, vol. 57, no. 2, Feb. 2009.

[16] M. B. Shenouda and T. N. Davidson, "On the design of linear transceivers for multiuser

systems with channel uncertainty", *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 10151024, Aug. 2008.

[17] M. B. Shenouda and T. N. Davidson, "Nonlinear and linear broadcasting with QoS requirements: Tractable approaches for bounded channel uncertainties", *IEEE Trans. Sig. Proc.*, vol. 57, pp. 19361947, May 2009.

[18] B.K. Chalise, S. Shahazpanahi, A. Czylwik, and A.B. Gershman, "Robust downlink beamforming based on outage probability specifications", *IEEE Trans. Wireless Comm.*, vol. 6, pp 3498-3505, Oct 2007.

[19] M. B. Shenouda and T. N. Davidson, "Outage-based designs for multi-user transceivers", *Proc. IEEE ICASSP 2009*, pp 2389-2392, April, 2009.

[20] K-Y. Wang, A.M-C. So, T-H. Chung, W-K. Ma, and C-Y. Chi, "Outage constrained robust transmit optimization for mulitiuser MISO downlinks: Tractable approxmations by conic optimization", arXiv.org>cs>arXiv:1108.0982.

[21] Y. Li and K.M. Wong, "A Riemannian distance for signal classification by power spectral density", to appear in *IEEE J. Selected Topics in Sig. Proc.*, 2013.

[22] D. Ciochina, M. Pesavento, and K.M. Wong, "Worst case robust downlink beamforming on the Riemannian manifold", *Proc. IEEE ICASSP*, Vancouver, Canada, May, 2013.

[23] R.J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, 1982.

[24] Z.-Q. Luo, J.F. Sturm, S.-Z. Zhang, "Multivariate non-negative quadratic mappings", *SIAM J. Optim.*, vol. 14, no. 4, pp. 1140-1162, 2004.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[26] M. Grant, S. Boyd, and Y. Y. Ye, "CVX: Matlab software for disciplined convex programming," . *http://cvxr.com/cvx*, Sept. 2012.

[27] Z.-Q. Luo, W.-K. Ma, A.M.-C. So, Y. Ye, S.-Z. Zhang, "Semidefinite relaxation of quadratic optimization problems", *IEEE Sig. Proc. Mag.*,vol. 27, pp. 20-34, May, 2010.

[28] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *37th Annual Allerton Conference on Communications, Control and Computing*, pp. 987-996, Sep. 1999.