

NANYANG
TECHNOLOGICAL
UNIVERSITY

COGNITIVE BEAMFORMING FOR SECONDARY SPECTRUM ACCESS

CAI SHENGMING

School of Electrical & Electronic Engineering

Nanyang Technological University

2014

COGNITIVE BEAMFORMING FOR SECONDARY SPECTRUM ACCESS

CAI SHENGMING

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2014

Summary

It has been widely accepted that cognitive radio can substantially improve the spectrum utilization compared with traditional spectrum allocation strategies. Cognitive radio technology enables secondary users to coexist with primary users opportunistically or concurrently. In the latter scenario, the primary user allows the secondary user to access its licensed spectrum, as long as the quality of service (QoS) of the primary transmission is guaranteed. Beamforming is a practical way to achieve this goal by utilizing multiple antennas to limit the interference at the primary receiver (PR) to a tolerable level. With transmit beamforming and receive beamforming, a multi-antenna secondary transmitter (ST) with the knowledge of cross channel state information (CSI) can project signals to appropriate directions such that no harmful interference is received at PR, while the secondary receiver (SR) can separate and recover the intended signals reliably. Perfect or imperfect knowledge of cross CSI at ST results in different interference conditions as well as different transmit strategies. The interference received at PR in the perfect cross CSI case is controllable, and hence ST can precisely determine the transmit beamforming directions to gain the maximum benefit under the primary interference constraint. However, such a strategy is not always suitable in the imperfect cross CSI case, due to the uncertainty of CSI error directions and amplitudes. In this case, ST needs to ensure that the interference constraint would not be violated even with the worst CSI error.

In this thesis, we first study cognitive beamforming subject to the secon-

dary power constraint and the primary interference power constraint. With bounded error model for the cross CSI, we rewrite the error boundary expressions as a set of linear matrix equalities, and then convert the whole problem into a convex optimization problem to calculate the optimal solution. When no cross CSI is available at ST, we aim at maximizing the secondary sum throughput subject to the interference expectation constraint and interference outage probability constraint at PR, respectively. We prove that under the former constraint, the beamforming problem can be formulated as a conventional MIMO transmission problem, which can be easily solved via the water-filling method. Under the latter constraint, we find that the interference power is related to the eigenvalues of the covariance matrix of the secondary signals, and thus it can be expressed as a linear combination of multiple chi-square distributed variables. By exploiting the impact of the eigenvalues on the interference outage probability and establishing the relationship between the vibrations of secondary covariance matrix and its eigenvalues, we design an algorithm for this beamforming problem.

In the second part of this thesis, we focus on beamforming in a multi-antenna cognitive radio system with transmission of multiple secondary data streams subject to the individual signal-to-noise ratio constraint per secondary data stream and the primary interference power constraint. The objective is to minimize the secondary transmit power consumption. Both perfect and imperfect cross CSI cases are considered and in both cases, the beamforming feasibility is tested. By exploiting the individual SNR constraints, we formulate the cognitive beamforming problem as an optimization problem on the Stiefel manifold. For zero forcing beamforming, we derive a closed form beamforming solution. For nonzero forcing beamforming, we prove that the strong duality holds for the nonconvex primal problem and thus the optimal solution can be easily obtained by solving the dual problem. In the case of imperfect cross CSI, we apply the S-procedure method to reformulate

the problem and provide an algorithm to obtain the beamforming solution. Finally, we raise some interesting open problems for the future work.

Contents

List of Figures	7
Conventions and List of Symbols	9
1 Introduction	11
1.1 Motivation	11
1.2 Objective	14
1.3 Major Contribution of the Thesis	14
1.4 Organization of the Thesis	16
2 Background and Preliminaries	18
2.1 Background	18
2.2 Convex Optimization	23
2.3 Logarithmic Barrier Function	24
2.4 Inner Product on the Tangent Space of Stiefel Manifold	26
2.5 Matrix Computation	26
2.6 S-procedure	28
3 Throughput Maximization: Partial CSI	29
3.1 Review of Existing Results	30
3.1.1 Perfect CSI Knowledge	30
3.1.2 Partial CSI Knowledge	32
3.2 System Model	33

3.3	One Single-Antenna PR	35
3.3.1	Interference Constraint Transformation and Solution	36
3.3.2	Optimality Conditions	38
3.4	Multiple Single-Antenna PRs	39
3.5	Simulation Results	40
3.6	Summary	42
4	Throughput Maximization: No Instantaneous CSI	44
4.1	System Model	47
4.2	Interference Expectation Constraint	49
4.3	Interference Outage Constraint	50
4.3.1	Problem Reformulation	50
4.3.2	Algorithm for the Local Optimum	52
4.4	Simulation Results	55
4.5	Summary	57
5	Power Minimization: Full CSI	59
5.1	System Model and Problem Formulation	61
5.1.1	System Model	61
5.1.2	Problem Formulation	64
5.1.3	Access Feasibility	65
5.2	The Case of Single Data Stream	67
5.3	The Case of Multiple Data Streams	70
5.3.1	Zero Forcing Beamforming	70
5.3.2	Nonzero Forcing Beamforming	74
5.4	Simulation Results	77
5.5	Summary	84
6	Power Minimization: Partial CSI	86
6.1	System Model	87

6.2	Maximization of Secondary Access Probability	89
6.2.1	S-procedure Transformation	89
6.2.2	Suboptimal Method	92
6.2.3	Optimal Method	94
6.3	Transmit Power Minimization	99
6.4	Simulation Results	102
6.5	Summary	105
7	Conclusion and Future Work	106
7.1	Conclusion	106
7.2	Future Work	108
	Appendix A Proof of Theorem 4.1	111
	Appendix B Proof of SNR Expression	113
	Appendix C Proof of Lemma 5.1	114
	Appendix D Proof of Lemma 5.3	115
	Appendix E Proof of Lemma 5.4	116
	Appendix F Derivation of D_T and D_τ	120
	Appendix G Derivation of D'_β	122
	Appendix H Derivation of D_v	123
	Author's Publications	125
	Bibliography	126

List of Figures

2.1	Approximation accuracy of the logarithmic barrier function with respect to t	25
3.1	Achievable rate versus transmit power at $\xi = 4$	40
3.2	Achievable rate versus interference threshold.	42
4.1	Secondary throughput subject to interference expectation constraint.	56
4.2	Secondary throughput subject to interference outage constraint.	57
5.1	Secondary access probability, i.e., the probability of $\xi_0 \leq \xi$. $m = n = 5, p = q = 2$ and $\rho = \rho_1 = \dots = \rho_d$	67
5.2	Power saving of optimal beamforming over feasible beamforming. $d = 1$	78
5.3	The impact of secondary SNR requirement. $d = 1$	79
5.4	The impact of primary interference constraint. $d = 2$ and $\rho_1 = \rho_2 = \rho$	80
5.5	The impact of secondary SNR requirement. $d = 2$ and $\rho_1 =$ $\rho_2 = \rho$	81
5.6	Minimum required transmit power per secondary data stream.	82
5.7	Minimum required transmit power with different number of secondary data streams.	83

5.8	Performance of distinct and identical secondary SNR requirements.	84
6.1	Outage probability versus interference power constraint.	103
6.2	Average transmit power versus ξ under spectrum access.	104

Conventions and List of Symbols

1. \mathbb{N} , \mathbb{R} and \mathbb{C} denote the set of all integers, real numbers and complex numbers, respectively. $\mathbb{R}^{m \times n}$ and $\mathbb{C}^{m \times n}$ denote the set of all $m \times n$ matrices whose entries belong to \mathbb{R} and \mathbb{C} , respectively.
2. Scalar is denoted by a lower-case letter, while vector is denoted by bold-face lower-case letter and matrix is denoted by bold-face upper-case letter.
3. Superscript H, T, C denote the Hermitian transpose, transpose and conjugate of a matrix, e.g., \mathbf{A}^H , \mathbf{A}^T , and \mathbf{A}^C .
4. Superscript -1 denotes the inverse of a full rank square matrix, e.g., \mathbf{A}^{-1} .

2. List of Symbols

- Re the real part of a complex number, vector or matrix.
- $\det(\cdot)$ determinant of matrix
- $\text{vec}(\cdot)$ vectorization of matrix
- \mathbf{a}_i the i -th element of vector \mathbf{a}
- \mathbf{A}_{ij} the entry of matrix \mathbf{A} located at the i -th row and j -th column

- $\mathbf{A}_{(i)}$ the i -th column of matrix \mathbf{A}
- $RS_{m \times n}(\mathbf{x})$ reshape from a vector $\mathbf{x} \in \mathbb{C}^{mn \times 1}$ to a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$, where $\mathbf{X}_{ij} = \mathbf{x}_{(j-1)m+i}$. It is the inverse operation of $\text{vec}(\cdot)$
- \triangleq equal by definition
- $\text{tr}(\cdot)$ trace of a square matrix
- element-wise product. $\mathbf{A} \bullet \mathbf{B} = \text{tr}(\mathbf{A}^T \mathbf{B})$, $\mathbf{A} \in \mathbb{C}^{m \times n}$, $\mathbf{B} \in \mathbb{C}^{m \times n}$
- \mathbf{I} identify matrix
- \mathbf{I}_m $m \times m$ identify matrix
- $O(\cdot)$ big O notation
- \otimes Kronecker product
- $\langle \cdot, \cdot \rangle$ inner product
- $E(\cdot)$ expectation
- $St(m, n)$ Stiefel manifold, the set $St(m, n) = \{\mathbf{X} \in \mathbb{C}^{m \times n} : \mathbf{X}^H \mathbf{X} = \mathbf{I}\}$
- $\pi(\mathbf{X})$ projection of \mathbf{X} onto the Stiefel manifold. The operation $\pi : \mathbb{C}^{m \times n} \rightarrow St(m, n)$ is defined to be $\pi(\mathbf{X}) = \arg \min_{\mathbf{Q} \in St(m, n)} \|\mathbf{X} - \mathbf{Q}\|^2$
- $\mathbf{A} \succ \mathbf{B}$ $\mathbf{A} - \mathbf{B}$ is positive definite
- $\mathbf{A} \succeq \mathbf{B}$ $\mathbf{A} - \mathbf{B}$ is positive semi-definite
- $\mathcal{CN}(\mu, \sigma^2)$ complex Gaussian distribution with mean μ and variance σ^2

Chapter 1

Introduction

1.1 Motivation

Cognitive radio (CR) has received considerable attention over the past few years because of its potential to ease the current overcrowded frequency spectrum. Based on the current spectrum allocation policy, most frequency bands are allocated exclusively to specified services. However, such policy results in underutilization of precious spectrum resources. In a report released by the Federal Communications Commission (FCC), it is stated that the usage of allocated spectrum varies from 15% to 85% [1]. In the meantime, the demand of extra spectrum is increasing with the rapid growth of wireless applications. As a result, it is worth considering the idea of allowing other users to use the allocated spectrum while guaranteeing the priority of authorized users.

In a CR network, the spectrum can be shared by the unauthorized users, which are called the secondary users (SUs), provided that they do not cause harmful interference to the authorized users, which are called the primary users (PUs) [2, 3]. SUs may transmit when they detect a spectrum hole in either time or frequency domain [4]-[10]. Such schemes work only when the spectrum is severely underutilized, otherwise SUs would not have sufficient

opportunities to gain channel access. Therefore, the secondary throughput would be significantly constrained and the secondary system would suffer from a long latency.

Multiple-input multiple-output (MIMO) technology provides extra spatial dimensions for transmissions. Multiple antennas can be used to reduce the interference at the primary receivers (PRs) and satisfy the demand for high data rate in the secondary wireless application by carefully designing the transmit and receive beamforming matrices [5]-[7]. Even if PUs are also using the spectrum at the same time, SUs may share the spectrum with PUs without causing harmful interference, and thus the secondary throughput can be greatly increased. As the cost, channel state information (CSI) is required to design the beamforming matrix. In some scenarios, since SUs can obtain CSI from PUs, the spectrum sharing scheme of CR communication can be performed to improve spectrum utilization.

In spite of the wide recognition of the advantages of spectrum sharing scheme, most existing results concentrate on the scenarios of single data stream per secondary receivers (SRs). As the mobile communication technology evolve toward 4G and 5G, user equipments will be equipped with more than one antenna and thus can receive multiple data streams in parallel [11, 12]. Many multiple data streams scenarios in cognitive radio, which can help increase the secondary throughput, are still unsolved due to the mathematical complexity [13]. The potentials of MIMO have been utilized at secondary transmitters (STs) but not SRs. It is very interesting to have an insight into problems of supporting multiple data streams per SR and exploit the performance potential of multiple data streams communication.

Both the areas of interference alignment and conventional MIMO communications consider the support of multiple data streams for one receiver. However, existing results of these areas cannot be directly applied to cognitive radio or are not optimal for cognitive radio due to the following reasons:

1. Interference alignment: First, in conventional interference alignment, all transmitters are assumed to share the spectrum with the same priority. They cooperatively design transmit beamforming matrices/vectors to eliminate interference at receivers. As PUs have no obligation to help SUs, SUs should fulfill all the calculation to guarantee QoS of PUs and maximize secondary performance.

Second, interference alignment aims at perfect alignment and interference cancellation. As a result, it has good performance in the high SNR regime where slight interference does not provide significant performance improvement. On the contrary, since it does not concern how much interference at PR can be accepted and how to utilize the interference limit, the performance in low and medium SNR regimes will be poor. This makes interference alignment not suitable for spectrum sharing as usually secondary devices operates in low and medium SNR regimes.

2. conventional MIMO communications: Similar to the first problem of interference alignment, transmitters in MIMO communications have the same priority to access channel, and they can share CSI information to jointly design optimal beamforming matrices. This is not possible in cognitive radio. SUs must estimate all the necessary information by themselves, which makes the behaviors of cognitive radio different from MIMO communications.

Moreover, in order to support multiple data streams for SRs, optimal algorithm designs for secondary QoS requirements and CSI conditions are very important but have not been well studied due to the high complexity of theoretical derivation. In this thesis, we will provide insights as well as solutions of multiple data streams transmission in cognitive radio under different scenarios.

1.2 Objective

The main objective of this thesis is to solve the mathematical problems of multiple secondary data streams communication and accordingly design feasible methods to make use of the potentials provided by this technology. To this end, we discuss two fundamental and important issues. Firstly, how the CSI knowledge errors would affect the secondary performance. Secondly, with different transmission targets, i.e., maximizing throughput and satisfying the individual signal-to-noise ratio (SNR) requirements, what the differences of the ST behaviors will be. Specifically, our goals to be achieved are listed as follows:

1. Guaranteeing that the given requirements and constraints, i.e., individual SNR requirements, interference constraint, transmit power constraint, can be satisfied in a CR network.
2. Finding out the differences in system designs and derivations under different transmission schemes, and providing the global or local optimal conditions to help determining whether an analytical or numerical solution can achieve a good performance.
3. Proposing feasible methods to exploit the advantages and potentials provided by multiple data streams transmission.

1.3 Major Contribution of the Thesis

In this thesis, we discuss some schemes for point-to-point transmission in a CR network where secondary transceivers are equipped with multiple antennas. We evaluate the performance boundaries for some aspects such as transmit power, interference power and access probability, and design practical methods to approach the optimal or suboptimal results. Here are the contributions of this thesis:

1. For the scenario where partial ST-PR CSI knowledge with element-wise bounded errors is available at ST and secondary throughput is to be maximized, we equivalently reformulate the problem based on the S-procedure technology. In the new problem, the channel errors have been removed, which means that it is unnecessary to compute the worst case channel error for the maximum possible interference power, and hence the computational complexity would be reduced. Besides, we show that the new problem is convex. Therefore, we can efficiently obtain the global optimal solution via convex optimization algorithm. Optimality conditions are provided for the analysis of physical meanings in the single carrier case and multiple carriers case, respectively.
2. For the scenario where ST-PR CSI knowledge is unavailable at ST and secondary throughput is to be maximized, we consider two types of primary interference constraints: one is on the expectation of interference at PR, referred to as interference expectation constraint, and the other is on the interference outage probability (i.e., the probability of interference at PR exceeding a predetermined threshold), referred to as interference outage constraint. We find that there is an analytical solution for beamforming subject to interference expectation constraint, while for beamforming subject to interference outage constraint, we find that the interference power is related to the eigenvalues of the secondary covariance matrix and can be expressed as a linear combination of multiple chisquare distributed variables. By exploiting the impact of the eigenvalues on the interference outage probability and establishing the relationship between the vibrations of secondary covariance matrix and its eigenvalues, we design an algorithm for this beamforming problem.
3. For the transmit power minimization problem subject to individual

secondary SNR requirements with perfect CSI knowledge, we use zero forcing beamforming (ZFBF) to represent the case when no interference is allowed at PU. If a positive interference limit is allowed, we call this case nonzero forcing beamforming (NZFBF) since the interference power constraint is positive. First, we formulate the secondary power minimization problem. After that, we formulate the problem as a minimization problem on Stiefel manifold [14]-[16]. For the ZFBF case, we derive a closed form solution for the optimal result. For the NZFBF case, we analyze the dual problem, and prove that global optimum can always be achieved by solving the dual problem using an efficient algorithm, although the primal problem is nonconvex.

4. For the transmit power minimization problem with individual secondary SNR requirements and imperfect CSI knowledge where the CSI errors are bounded by an ellipse expression, we propose an algorithm to test whether the current channel condition is feasible to support the worst interference power constraint. Then we apply logarithmic barrier function to rewrite the problem as an unconstrained problem and use gradient descent algorithm to get a local optimal solution after passing the feasibility test. Since the variable domain lies on the Stiefel manifold, which is only a subset of the complex domain, we project the descent direction to the tangent space of the Stiefel manifold, derive the optimality condition and stop criterion of the algorithm based on projected descent direction.

1.4 Organization of the Thesis

In Chapter 2, we offer some background knowledge on cognitive radio and MIMO communications. Some preliminary mathematical results are also included in this Chapter. Chapter 3 is devoted to the study of the partial

CSI scenario where element-wise bounded channel errors are considered. In Chapter 4, after briefly introducing the existing results for multiple data streams communication in a CR network, we focus on the scenario where ST has no instantaneous CSI knowledge of ST-PR channel. Both Chapter 3 and Chapter 4 are for the throughput maximization problem. In Chapter 5 and Chapter 6, we move to the power minimization problem with individual SNR requirements. In Chapter 5, perfect CSI knowledge is assumed. The results for the imperfect CSI knowledge scenario is provided in Chapter 6. The conclusion is given in Chapter 7, along with some possible future works related to transmission with multiple data streams in a CR network.

Chapter 2

Background and Preliminaries

2.1 Background

Wireless technologies have advanced rapidly over the last two decades and more and more spectrum resources are required to satisfy the demands from various wireless applications. However, according to present spectrum management strategy, frequency bands are exclusively allocated to specific applications and hence unlicensed applications have no access to them. This strategy results in the shortage of unallocated spectrum and very low spectrum utilization as well [1]. Such a spectrum authorization policy leads to the following contradictions. On one hand, the amount of unallocated spectrum is becoming less and less, and the price of applying for licensed frequency bandwidth is getting more and more. On the other hand, current spectrum utilization situation varies in different time, regions and frequency [2]. Roughly speaking, most allocated spectrum resources are under-utilized. As a result, the low spectrum efficiency implies a great potential in tackling the emerging demands on spectrum resources.

Cognitive radio, which allows other applications to use authorized spectrums without causing harmful interference to PUs, is a promising solution

to tackling these problems and has attracted many attentions in recent years [2]. Cognitive radio communication is conducted by SUs in the presence of PUs. Among these two types of users, PUs have higher priority in using spectrum resources and their performance should not be harmed even if SUs are also using spectrum at the same time [2]-[3].

It is estimated that the aggregate data rate or area capacity will increase by about 1000-fold from 4G to 5G for 2020 and beyond [11], [17]-[19]. To meet this requirement and support mobile communications for both indoors and outdoors, a heterogeneous networks architecture is necessary and multi-tier will be present. As cognitive radio is able to flexibly share spectrum among different tiers of networks and control interference to increase spectrum efficiency, it will be one of the candidate technologies in 5G [11], [12], [20], [21].

SUs can access the spectrum while PUs are idle, or share spectrum without causing unacceptable interference. In the former case, SUs should keep sensing till the spectrum is idle, and return the right of using spectrum to PUs as long as a new primary communication is detected [4]-[7], [22]-[24]. Such schemes work only when the spectrum is severely underutilized, otherwise SUs would not have sufficient opportunities to gain channel access. Therefore, the secondary throughput would be significantly constrained and the secondary system would suffer from a long latency. In the later case, SUs may work simultaneously with PUs but has to limit the performance downgrade experienced by PRs. ST should be able to detect the signals from primary transmitter (PT) or cooperate with PUs because ST requires channel state information to compute the exact or possible interference at PRs [25]-[30].

There are two main technologies in spectrum sharing for cognitive radio: one is overlay and the other is underlay [8]. In overlay schemes, SU allocates part of its power to help transmit the primary signal in order to compensate

for the additional interference from the secondary transmission [8], [34]-[37]. However, the problem is how ST can get *a priori* primary message in advance.

On the other hand, underlay allows parallel transmission of both primary and secondary messages, provided that the interference experienced at the PR is not harmful [8], [28]-[30], [35, 38, 40]. If there is a single antenna at ST, one can choose to do power control to handle the interference power. But if there are multiple antennas at ST, it is possible to make use of the spatial diversities and hence dynamically change the direction and power for each data stream, so as to control the overall additional interference at PR under given threshold and maximize the secondary benefit [39]-[46].

Both the overlay and underlay schemes in spectrum sharing require CSI knowledge to design beamforming matrices. When the primary communication system are operating in the time-division-duplex (TDD) mode, by exploiting the channel reciprocity, ST can periodically sense the channel and detect the ACK or pilot signal sent by PRs to estimate CSI knowledge before transmitting. In the scenario where the ST-PR channel changes slowly, this instantaneous CSI knowledge can be considered with high accuracy [47]-[51]. In a fading environment, it is very difficult for ST to estimate the instantaneous information perfectly, and thus ST will have partial CSI with different types of error, e.g., estimation error and delayed error [26]-[27], [30]-[33]. Therefore, results derived from perfect instantaneous CSI scenarios can be considered as upper bounds for fading channel scenarios.

It is widely accepted that in 5G, massive MIMO will be one of the key technologies, and multiple antennas will be installed at user equipments. It means that both the overlay and underlay schemes may be considered in 5G. In this case, multiple data streams at one SR can be supported to increase secondary capacity. However, most existing results for cognitive radio consider the scenarios where SRs have a single antenna each, and thus at most one data stream can be received by a SR. As a result, it is desirable

to exploit the spatial diversities provided by multiple antennas to improve performance at the receiver end.

If the number of ST antennas is larger than that of PR antennas, cognitive beamforming at ST is able to eliminate interference at PR via projecting the data streams to a suitable signal space and computing the respective power for each data stream with perfect CSI. At the same time, the SR may adopt the receive beamforming technology to separate each data stream and then recover them. When the CSI knowledge is imperfect, although the interference at PR cannot be precisely controlled, we can still limit the interference power within an acceptable region via signal subspace projection. For example, if the channel errors are upper bounded, the worst case interference power can be calculated to ensure that the interference constraint would not be violated; if the channel errors are unbounded with statistical information available, the interference expectation or the violation probability of interference constraint can be restricted below the harmful level. Thus MIMO technology has great advantages in cognitive radio communication.

Throughput maximization and power minimization are two important issues in the research of MIMO transmission. Throughput maximization problems in which each SR receives one data stream have been widely studied. In many scenarios with multiple ST-SR pairs, semi-definite programming (SDP) relaxation would be used to convert the original problem to a convex optimization problem by dropping the rank constraint and generate a local optimum [52]-[59]. It is shown in [58] that under certain conditions, some techniques can be used to generate a new solution from the one obtained by SDP relaxation without ruining the constraints or changing the objective function, and hence the solution is optimal. However, in general scenarios, the obtained local optimum may not be feasible for the original problem because usually its rank does not meet the requirement. As a result, approximation approaches such as the randomization procedure are used to

generate a feasible solution [59]-[60]. As to the cases of multiple data streams transmitted by one ST, the current results are mainly on the downlink or relay transmission with multiple single-antenna SRs in the presence of single-antenna or multi-antenna PRs [16],[61]-[65]. There is not too much work on the transmission to multi-antenna SRs because of the high complexity caused by the additional interference constraints comparing to the conventional MIMO communications.

When MIMO technology is adopted in CR communications, the expressions of transmit power and SNR usually results in quadratically constrained quadratic programming (QCQP) problems, which cannot be directly solved via convex optimization tools [51]-[68]. The feasibility of SDP relaxation technology relies on the objective and constraints of a particular problem, and it is not always applicable for all kinds of MIMO communications in a CR network. It is shown in [16] that the secondary transmit beamforming problem can be converted to an optimization problem with unitary constraint and an algorithm is proposed therein to compute the beamforming matrix that maximizes the secondary throughput. In this thesis, we are to apply this technology to solve the problem of power minimization with individual secondary SNR constraint.

Given perfect CSI knowledge, there exists a closed form solution for the zero interference constraint problem and a numerical solution for the nonzero interference constraint problem to obtain the system capacity [28]. From a practical point of view, the transmitter usually does not have perfect CSI. The transmitter may get CSI feedback with quantization error from the receiver in block fading channel because of limited feedback channel capacity. Quantization error is often considered to be upper bounded and the upper bound is determined by the method that the receiver uses to quantize CSI. Besides, it is possible that the received CSI is a delayed version because of fading. When an underlay cognitive radio communication system with multiple

antennas faces the upper bounded cross CSI error, ST needs to ensure that its transmit beamforming matrix does not violate the interference power constraint over all possible values of the actual CSI [69]-[72]. A bounded channel error model is introduced in [31]-[33], which is described by the expression of ellipsoids. This model is not only suitable for block fading channel, but also applicable to channels that vary slowly within a block. The expressions of ellipsoids describing the channel error determine the uncertainty region and the extent to which SU can protect PU. The single secondary data stream scenario is discussed in [69, 31]. For the schemes of multiple secondary data streams, an algorithm is proposed to calculate the system capacity given an ellipsoid cross CSI error model [33]. We will focus on the multiple secondary data streams communications and extend the existing results to different schemes in this thesis.

2.2 Convex Optimization

Consider the following optimization problem

$$\begin{aligned} \min_x \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \leq 0, i = 1, \dots, m \end{aligned}$$

where $f(x)$ and $g_i(x)$ are assumed convex. The Lagrangian associated with the original problem, or call primal problem, is given by

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) \quad (2.1)$$

where $\lambda_i \geq 0$ is referred to as the Lagrange multiplier associated with the i -th inequality constraint $g_i(x) \leq 0$. The optimization problem

$$\min_x \max_{\lambda} L(x, \lambda), \forall \lambda_i \geq 0 \quad (2.2)$$

is equivalent to the primal problem. That is because with all the constraints satisfied, both problems share the same optimal result; if any constraint is

unsatisfied, then the corresponding Lagrange multiplier can be chosen $+\infty$, which makes the optimization problem invalid.

The dual problem of (2.2) is defined as

$$\max_{\lambda} \min_x L(x, \lambda), \forall \lambda_i \geq 0. \quad (2.3)$$

The difference between the primal optimal solution and dual optimal solution is called duality gap. Obviously, the dual optimal solution serves as the lower bound of the primal optimal solution, so usually the duality gap is nonzero. In some particular scenarios, e.g., the primal problem is convex, the duality gap may be zero, then we can obtain the primal optimal solution via solving the dual problem instead of the primal one. In many cases, the dual problem is easier to solve and requires less computations. However, when the duality gap is nonzero, it is infeasible to solve the dual problem for the primal optimal solution.

2.3 Logarithmic Barrier Function

The logarithmic barrier function method is used to solve the optimization problem with inequality constraints [73]:

$$(P2.1) \quad \begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) \leq 0, i = 1, \dots, m \end{aligned}$$

where $f_0, \dots, f_m : \mathbb{C}^n \rightarrow \mathbb{R}$ are convex and twice continuously differentiable, and the problem is solvable, i.e., an optimal \mathbf{x}^* exists. If we can construct a barrier function $I_- : \mathbb{C}^n \rightarrow \mathbb{R}$ such that

$$I_-(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases}, \quad (2.4)$$

then (P2.1) can be rewritten as

$$\min_{\mathbf{x}} \quad f_0(\mathbf{x}) + \sum_{k=1}^m I_-(f_k(\mathbf{x})).$$

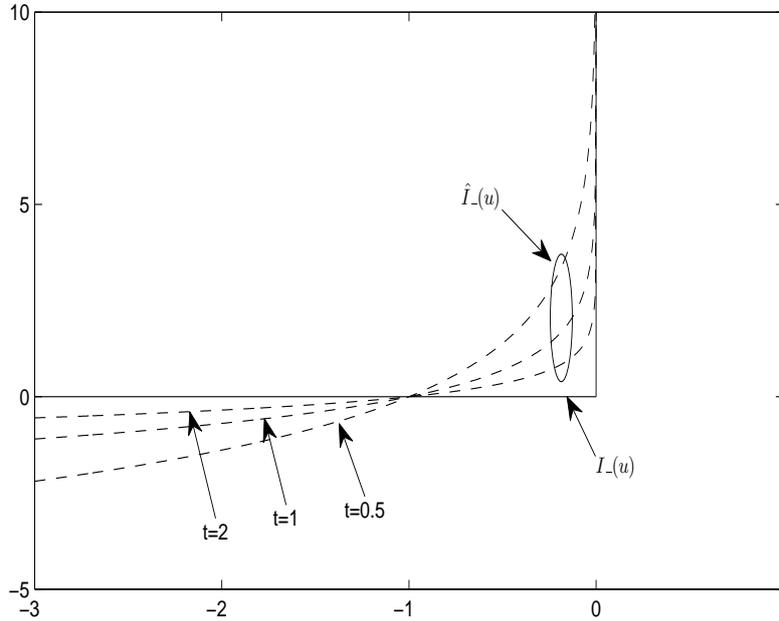


Figure 2.1: Approximation accuracy of the logarithmic barrier function with respect to t

In this problem \mathbf{x} must be chosen to make $f_k(\mathbf{x}) \leq 0$ for all k , otherwise the objective function value becomes positive infinity. Note that the barrier function is not differentiable so we cannot apply Newton's method or similar algorithms [73]-[74] to solve it.

A logarithmic barrier function is a barrier function constructed by using the log function to approximate $I_-(u)$. Letting

$$\hat{I}_-(u) \triangleq -\frac{1}{t} \log(-u), \quad (2.5)$$

we can see that $\hat{I}_- \rightarrow \infty$ as $u \rightarrow 0$. Furthermore, with the increase of t , the approximation becomes more accurate as shown in Fig. 2.1. Therefore we may reformulate the original problem with logarithmic barrier function \hat{I}_- :

$$\min_{\mathbf{x}} \quad f_0(\mathbf{x}) + \left(-\frac{1}{t}\right) \sum_{k=1}^m \log(-f_k(\mathbf{x})). \quad (2.6)$$

2.4 Inner Product on the Tangent Space of Stiefel Manifold

Suppose $\mathbf{X} \in St(m, n)$, then the tangent space of \mathbf{X} is defined by

$$\mathbf{T}_{\mathbf{x}} = \{\mathbf{D} \in \mathbb{C}^{m \times n} : \mathbf{D} = \mathbf{X}\mathbf{A} + \mathbf{X}_{\perp}\mathbf{B}\} \quad (2.7)$$

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a Hermitian matrix, $\mathbf{X}_{\perp} \in St(m, m - n)$ is orthogonal to \mathbf{X} , and $\mathbf{B} \in \mathbb{C}^{(m-n) \times n}$ is arbitrary [14]-[15].

For any matrix $\mathbf{X} \in St(m, n)$ and $\mathbf{D} \in \mathbb{C}^{m \times n}$, we can find a matrix $\mathbf{Z} \in \mathbf{T}_{\mathbf{x}}$ such that $\lim_{t \rightarrow 0} \pi(\mathbf{X} + t\mathbf{D}) = \lim_{t \rightarrow 0} \pi(\mathbf{X} + t\mathbf{Z}) + O(t^2)$.

For any function $f : St(m, n) \rightarrow \mathbb{R}$, we denote by $\mathbf{D}_{\mathbf{x}}$ the derivative of $f(\mathbf{X})$. The descent direction $\mathbf{Z}_{\mathbf{x}}$ in Newton's method [73] should belong to the tangent space of \mathbf{X} and can be calculated with the following formula:

$$\mathbf{Z}_{\mathbf{x}} = \mathbf{X}\mathbf{D}_{\mathbf{x}}^H\mathbf{X} - \mathbf{D}_{\mathbf{x}}. \quad (2.8)$$

The inner product $\langle \mathbf{D}_{\mathbf{x}}, \mathbf{D}_{\mathbf{x}} \rangle$ is given by [14]-[15]

$$\langle \mathbf{D}_{\mathbf{x}}, \mathbf{D}_{\mathbf{x}} \rangle \triangleq \text{tr}(\mathbf{D}_{\mathbf{x}}^H(\mathbf{I} - \frac{1}{2}\mathbf{X}\mathbf{X}^H)\mathbf{D}_{\mathbf{x}}). \quad (2.9)$$

2.5 Matrix Computation

1. The eigenvalue decomposition (EVD) of a Hermitian matrix $\mathbf{M} \in \mathbb{C}^{n \times n}$ is defined as

$$\mathbf{M} = \mathbf{U}^H\mathbf{A}\mathbf{U} \quad (2.10)$$

where $\mathbf{U} \in \mathbb{C}^{n \times n}$ is a unitary matrix, $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a diagonal matrix whose diagonal entries are the eigenvalues of \mathbf{M} . Without loss of generality, we assume $\mathbf{A}_{11} \geq \mathbf{A}_{22} \geq \dots \geq \mathbf{A}_{nn}$. $\mathbf{M}^{1/2} \in \mathbb{C}^{n \times n}$ denotes the square root of the Hermitian matrix \mathbf{M} with $\text{rank}(\mathbf{M}) = r \leq n$, i.e., $(\mathbf{M}^{1/2})^H\mathbf{M}^{1/2} = \mathbf{M}$, which is obtained from the following decomposition of \mathbf{M} : If \mathbf{M} is written as $\mathbf{M} = \mathbf{U}^H\mathbf{A}\mathbf{U}$, where $\mathbf{U}^H \in St(n, r)$,

$\mathbf{A} \in \mathbb{C}^{r \times r}$ is a diagonal matrix, then $\mathbf{M}^{1/2} = \mathbf{U}^H \mathbf{A} \mathbf{U}$. Likewise, $\mathbf{M}^{-1/2}$ is found as $\mathbf{M}^{-1/2} = \mathbf{U}^H \mathbf{A}^{-1/2} \mathbf{U}$.

2. The derivative of vector function $\mathbf{y}(\mathbf{x}) : \mathbb{C}^{m \times 1} \rightarrow \mathbb{C}^{n \times 1}$ is defined as

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \dots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}^H. \quad (2.11)$$

3. *Chain rule*: suppose \mathbf{z} is a function of \mathbf{y} , which is also a function of \mathbf{x} , then

$$\frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \frac{\partial \mathbf{z}}{\partial \mathbf{y}}.$$

This thesis chooses to express the first-order Taylor series approximation of an arbitrary function $f : \mathbb{C}^{m \times n} \rightarrow \mathbb{R}$ in the form of

$$\begin{aligned} f(\mathbf{X} + t\mathbf{Z}) &= f(\mathbf{X}) + t \operatorname{Re}\{\operatorname{tr}(\mathbf{D}^H \mathbf{Z})\} + O(t^2) \\ &= f(\mathbf{X}) + t \operatorname{Re}\{\operatorname{vec}^T(\mathbf{D}^C) \operatorname{vec}(\mathbf{Z})\} + O(t^2). \end{aligned} \quad (2.12)$$

where $\mathbf{D} = \frac{\partial f}{\partial \mathbf{X}} \in \mathbb{C}^{m \times n}$ denotes the derivative of f evaluated at \mathbf{X} .

4. The differential of some particular functions are given as follows:

$$d \det(\mathbf{X}) = \operatorname{Re}\{\operatorname{tr}((\det(\mathbf{X}) \mathbf{X}^{-1})^T d\mathbf{X})\} \quad (2.13)$$

$$d\mathbf{X}^{-1} = -\mathbf{X}^{-1} d\mathbf{X} \mathbf{X}^{-1} \quad (2.14)$$

$$\operatorname{vec}(\mathbf{A} d\mathbf{X} \mathbf{B}) = (\mathbf{B}^T \otimes \mathbf{A}) \operatorname{vec}(d\mathbf{X}) \quad (2.15)$$

5. Let \mathbf{M} be a square matrix and \mathbf{A} be its Adjugate matrix. Then we have the following result [75]:

$$\det(\mathbf{M} + d\mathbf{M}) = \operatorname{tr}(\mathbf{A} d\mathbf{M}) \quad (2.16)$$

2.6 S-procedure

We briefly introduce S-procedure [73] in this section. Let $\mathbf{V} \in \mathbb{C}^{n \times 1}$ be a subset of vector space $\mathbb{C}^{n \times 1}$ and $\sigma_k : \mathbf{V} \rightarrow \mathbb{R}, k = 0, 1, \dots, N$ be the functions defined on \mathbf{V} . Consider these two statements:

- (a) $\forall \mathbf{v} \in \mathbf{V}$, if $\sigma_k(\mathbf{v}) \geq 0, k = 1, 2, \dots, N$, then $\sigma_0(\mathbf{v}) \geq 0$.
- (b) There exist $\tau_k \geq 0, k = 1, \dots, N$ such that

$$\sigma_0(\mathbf{v}) - \tau_k \sum_{k=1}^N \sigma_k(\mathbf{v}) \geq 0, \forall \mathbf{v} \in \mathbf{V}.$$

Obviously, (b) is a sufficient condition of (a) since $\tau_k \geq 0$. The S-procedure is a method of using (b) to verify if (a) holds. In fact, it is a Lagrange relaxation technique [76] and the solution is lossless provided that $N \leq 2$ [77]. Particularly, if $\sigma_k(\mathbf{v})$ is expressed as

$$\sigma_k(\mathbf{v}) = \mathbf{v}^H \mathbf{Q}_k \mathbf{v} + \text{Re}\{\text{tr}(\mathbf{s}_k^H \mathbf{v})\} + r_k, \quad (2.17)$$

where $\mathbf{Q}_k \in \mathbb{C}^{n \times n}$ is a positive semi-definite Hermitian matrix, $\mathbf{s}_k \in \mathbb{C}^{n \times 1}$ and $r_k \in \mathbb{R}$, then (b) holds if and only if there exist $\tau_k \geq 0, k = 1, 2, \dots, N$ such that

$$\begin{bmatrix} \mathbf{Q}_0 & \mathbf{s}_0 \\ \mathbf{s}_0^H & r_0 \end{bmatrix} - \sum_{k=1}^N \tau_k \begin{bmatrix} \mathbf{Q}_k & \mathbf{s}_k \\ \mathbf{s}_k^H & r_k \end{bmatrix} \succeq \mathbf{0}. \quad (2.18)$$

Interested readers can refer to [73] for detailed proof.

Chapter 3

Throughput Maximization:

Partial CSI

It is widely known that in a CR network, beamforming technology can dynamically adjust the transmitted secondary signals, so as to limit the primary interference power and maximize the secondary throughput. One of the key requirements on beamforming is the CSI knowledge. When the obtained CSI perfectly matches the instantaneous actual CSI, we call it perfect or full CSI; when the obtained CSI is not completely the same as the actual CSI, we call it partial CSI. The more precise the CSI knowledge is, the better primary interference control and the higher secondary throughput can be achieved. In the scenarios where primary users can help ST get perfect or partial CSI knowledge between the primary system and secondary system, ST can figure out the possible region of interference power given a particular transmit beamforming matrix, and design the optimal one to limit the interference power while maximizing its throughput. In recent years, the point-to-point multiple data streams throughput maximization problem in a CR network have attracted some researchers' interests. Before starting to discuss the problems considered in this thesis, we will first briefly introduce the existing

results on point-to-point cognitive transmission with multiple data streams at one SR with perfect CSI and partial CSI with errors bounded by an ellipsoid model, and then propose an element-wise channel errors model to describe the channel errors and derive the optimal solution. The reason we propose an element-wise channel errors model is that the ellipsoid model implies the dependence of channel errors. When the channel errors are independent of each other, the ellipsoid channel errors model cannot describe the independence and will result in performance loss.

3.1 Review of Existing Results

Researchers mainly focus on two types of scenarios on this topic. The first one is with perfect CSI knowledge, and the second one is with partial CSI knowledge bounded by a particular ellipsoid region. We take two fundamental scenarios as examples here. Interested readers may refer to papers like [32, 78] for the details of extensions to different scenarios.

3.1.1 Perfect CSI Knowledge

The system model considered in [28] is with a ST-SR pair and K PRs. The number of antennas equipped at ST, SR and PR are denoted by m , n , and q , respectively. The ST-SR channel is denoted by $\mathbf{H} \in \mathbb{C}^{n \times m}$. The secondary transmission is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$$

where \mathbf{x} is the secondary transmitted signal symbols, \mathbf{n} is additive complex Gaussian noise and assumed that $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Suppose that the covariance of transmit beamforming matrix is $\mathbf{Q} \in \mathbb{C}^{m \times m}$. Because \mathbf{Q} is calculated as $\mathbf{Q} = \text{E}(\mathbf{x}\mathbf{x}^H)$, where \mathbf{x} usually contains several data streams precoded by a transmit beamforming matrix, \mathbf{Q} must be a positive semi-definite matrix

and its rank is determined by the number of data streams. Since the channel capacity is to be derived, it is assumed that Gaussian code-book with infinite number of codeword symbols is used.

Supposing that the number of antennas equipped at the k -th PR is denoted by q_k , $k = 1, \dots, K$, the ST-PR channels are denoted by $\mathbf{G}_k \in \mathbb{C}^{q \times m}$, $k = 1, \dots, K$. There is a sum interference power constraint for each PR. The interference power constraint is given by

$$\text{tr}(\mathbf{G}_k \mathbf{Q} \mathbf{G}_k^H) \leq \xi_k, \forall k = 1, \dots, K. \quad (3.1)$$

Besides, the transmit power of the beamforming matrix should be restricted by the ST power constraint P , i.e.,

$$\text{tr}(\mathbf{Q}) \leq P.$$

The problem is formulated as

$$\begin{aligned} \max_{\mathbf{Q} \succeq \mathbf{0}} \quad & \log \det(\mathbf{I} + \mathbf{H} \mathbf{Q} \mathbf{H}^H) \\ \text{s.t.} \quad & \text{tr}(\mathbf{Q}) \leq P \\ & \text{tr}(\mathbf{G}_k \mathbf{Q} \mathbf{G}_k^H) \leq \xi_k, \forall k = 1, \dots, K. \end{aligned}$$

Note that the function $\log \det(\cdot)$ is convex, so are the constraints and the feasible domain of \mathbf{Q} . Therefore, this problem can be solved by the standard convex optimization algorithms, such as the interior point algorithm [73].

Particularly, if each PR has only one antenna, or has multiple antennas but per antenna interference power constraint is considered, we can replace \mathbf{G}_k by the vector $\mathbf{g}_{k,l} \in \mathbb{C}^{1 \times m}$, where $\mathbf{g}_{k,l}$ represents the l -th antenna of the k -PR. The interference power constraint can be accordingly changed as follows:

$$\mathbf{g}_{k,l} \mathbf{Q} \mathbf{g}_{k,l}^H \leq \xi_k, \forall k = 1, \dots, K, l = 1, \dots, q_k. \quad (3.2)$$

Such a change does not affect the convexity of the interference constraint functions, so in this system model, the global optimum can be achieved regardless of the number of antennas at all the transceivers.

3.1.2 Partial CSI Knowledge

The system model in [33] is similar to the previous scenario, except that the ST-PR channel is not completely available at ST. For simplicity, we introduce the single carrier case here. The actual channel coefficient matrix between ST and the k -th PR is given by

$$\mathbf{G}_k = \hat{\mathbf{G}}_k + \mathbf{E}_k \quad (3.3)$$

where $\hat{\mathbf{G}}_k$ is the estimate of \mathbf{G}_k and \mathbf{E}_k is the corresponding estimation error matrix. It is supposed that the uncertainty is restricted by the following region:

$$U_k = \left\{ \mathbf{G}_k \left| \begin{array}{l} \mathbf{G}_k = \hat{\mathbf{G}}_k + \mathbf{E}_k, \\ \text{tr}(\mathbf{E}_k \mathbf{R}_k \mathbf{E}_k^H) \leq \epsilon_k \end{array} \right. \right\}, k = 1, \dots, K \quad (3.4)$$

where $\mathbf{R}_k \in \mathbb{C}^{m \times m}$ is a Hermitian positive definite matrix, $\epsilon_k \in \mathbb{R}$ is a nonnegative real number. Particularly, $\epsilon_k = 0$ means that the error matrix $\mathbf{E}_k = \mathbf{0}$, i.e., the CSI knowledge is perfect. For example, if \mathbf{E}_k is a row vector, then this shape is an ellipse. The directions of the region is determined by the eigenvectors of \mathbf{R}_k , while the size of each direction is determined by the corresponding eigenvalue. A larger eigenvalue of \mathbf{R}_k results in a smaller uncertainty region at the corresponding direction.

The problem considered in [33] is formulated as

$$\begin{aligned} \max_{\mathbf{Q} \succeq \mathbf{0}} \quad & \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \\ \text{s.t.} \quad & \text{tr}(\mathbf{Q}) \leq P, \\ & \text{tr}(\mathbf{G}_k \mathbf{Q} \mathbf{G}_k^H) \leq \xi_k, \forall \mathbf{G}_k \in U_k, k = 1, \dots, K. \end{aligned}$$

The main advantage of this channel model is that S-procedure can be directly used to obtain the solution. By applying S-procedure to the interference power constraint, this optimization problem can be equivalent trans-

formed to the following problem

$$\begin{aligned}
& \max_{\mathbf{Q}, \mathbf{S}_k, \mu_k} && \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \\
& \text{s.t.} && \text{tr}(\mathbf{Q}) \leq P, \\
& && \text{tr}(\hat{\mathbf{G}}_k \mathbf{Q} \hat{\mathbf{G}}_k^H) + \text{tr}(\hat{\mathbf{G}}_k \mathbf{S}_k \hat{\mathbf{G}}_k^H) + \mu_k \epsilon_k \leq \xi_k, k = 1, \dots, K, \\
& && \begin{bmatrix} \mathbf{S}_k & \mathbf{Q} \\ \mathbf{Q} & \mu_k \mathbf{R}_k - \mathbf{Q} \end{bmatrix} \succeq \mathbf{0}, \\
& && \mu_k \geq 0, k = 1, \dots, K, \\
& && \mathbf{Q} \succeq \mathbf{0}.
\end{aligned}$$

The above problem is convex, and hence it can be efficiently solved via convex optimization algorithms.

3.2 System Model

In this chapter, we consider another partial channel error model different from the ellipsoid one. The ellipsoid model implies that every channel error may depend on all the rest. It is not only related to the sum power gain of the channel errors, but also to the product of two instantaneous channel errors when the off-diagonal entries of \mathbf{R}_k are not all zero. For an arbitrary bounded channel error model, we can find an ellipsoid model to cover the whole possible region. It is desirable if the channel errors region perfectly match an ellipsoid one. If the actual channel errors region is not entirely of the same shape, e.g., the errors are independent of each other, the ellipsoid model can be regarded as defining a loose bound to contain the actual region. Therefore, while the ellipsoid models bring the convenience of obtaining a feasible solution, it may cause a performance loss due to the channel error model mismatch.

By applying the ellipsoid model, we do not have to know the distribution type of channel errors and the corresponding parameters. Instead, what we

need is the boundary of errors. Therefore, the ellipsoid model enables us to calculate the upper bound of the worst case interference power even without accurate information of channel errors distribution. On the other hand, it is impossible to calculate the expectation and variance of interference power with the ellipsoid model because relevant information is missing.

We consider a cognitive radio network which has a multi-antenna secondary user pair (ST-SR) and q single-antenna PRs. Suppose that the number of antennas at ST and SR are denoted by m and n , respectively. All the channel coefficient matrices are assumed constant during one symbol period and different from one period to another. The channel from ST to SR is $\mathbf{H} \in \mathbb{C}^{n \times m}$, which is perfectly known at ST. Suppose that ST has partial CSI knowledge of all the ST-PR channels, which is given by

$$\mathbf{g}_i = \bar{\mathbf{g}}_i + \mathbf{e}_i \quad (3.5)$$

where $\mathbf{g}_i \in \mathbb{C}^{1 \times m}$ represents the actual value of the channel from ST to the i -th PR, $\bar{\mathbf{g}}_i \in \mathbb{C}^{1 \times m}$ is the estimate of \mathbf{g}_i at ST, $\mathbf{e}_i = [e_{i,1}, \dots, e_{i,m}] \in \mathbb{C}^{1 \times m}$ is the channel error. We assume that \mathbf{e}_i is bounded and falls in the following uncertainty set:

$$\mathbf{V}_i = \{\mathbf{e}_i \mid |e_{i,k}| \leq \gamma_k, \forall k = 1, \dots, m\} \quad (3.6)$$

The channel error model (3.6) means that every entry of \mathbf{e}_i is restricted by a given individual bound, and may not be related to the other entries. It is practical since usually the entries of a wireless channel matrix and the channel errors are considered to be independent variables [79]-[83]. We call this error model element-wise channel error model.

Let $\mathbf{x} \in \mathbb{C}^{m \times 1}$ denote the secondary signals and $\mathbf{Q} \in \mathbb{C}^{m \times m}$ denote their covariance matrix, i.e., $\mathbf{Q} = \mathbf{E}(\mathbf{x}\mathbf{x}^H)$. It is required that for each PR, the primary interference threshold ξ_i should not be violated even with the worst \mathbf{e}_i . As a result, the interference constraint is given by

$$\text{tr}((\bar{\mathbf{g}}_i + \mathbf{e}_i)\mathbf{Q}(\bar{\mathbf{g}}_i + \mathbf{e}_i)^H) \leq \xi_i, \forall \mathbf{e}_i \in \mathbf{V}_i, i = 1, \dots, q.$$

The transmit power in each block is restricted by the maximum power of ST represented as P , i.e., it holds that $\text{tr}(\mathbf{Q}) \leq P$. The secondary throughput maximization problem is formulated as

$$(P3.1) \quad \max_{\mathbf{Q} \succeq \mathbf{0}} \quad \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H), \quad (3.7a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{Q}) \leq P, \quad (3.7b)$$

$$\begin{aligned} & \text{tr}((\bar{\mathbf{g}}_i + \mathbf{e}_i)\mathbf{Q}(\bar{\mathbf{g}}_i + \mathbf{e}_i)^H) \leq \xi_i, \\ & \forall \mathbf{e}_i \in \mathbf{V}_i, i = 1, \dots, q. \end{aligned} \quad (3.7c)$$

This problem can also be considered as a scenario where each PR has multiple antennas with individual antenna interference constraints, so it is applicable regardless of the actual antenna number at PRs.

3.3 One Single-Antenna PR

First, we consider the case where there is only one single-antenna PR, i.e., $q = 1$. We can simplify (P3.1) to the following problem

$$(P3.2) \quad \max_{\mathbf{Q} \succeq \mathbf{0}} \quad \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \quad (3.8a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{Q}) \leq P, \quad (3.8b)$$

$$\begin{aligned} & (\bar{\mathbf{g}}_1 + \mathbf{e}_1)\mathbf{Q}(\bar{\mathbf{g}}_1 + \mathbf{e}_1)^H \leq \xi_1, \\ & \forall \mathbf{e}_1 \in \mathbf{V}_1. \end{aligned} \quad (3.8c)$$

Since for particular $\bar{\mathbf{g}}_1$ and \mathbf{Q} , the constraint (3.8c) must be satisfied for all $\mathbf{e}_1 \in \mathbf{V}_1$. Therefore, (P3.2) is actually a semi-infinite optimization problem, i.e., with finite number of variables and infinite number of constraints. One method for such semi-infinite optimization problems is to use the S-procedure technique, which transforms the original infinite constraints to finite constraints and remains other constraints and objective function unchanged. The original constraint set is a subset of the new one, so the solution of the new problem can satisfy the original constraints. This solution

is feasible but usually suboptimal. In this problem, since only the ST-PR channel is imperfect, the channel error uncertainty will just have impact on the interference power, and the objective function will not be affected. We will also use S-procedure to solve this problem.

3.3.1 Interference Constraint Transformation and Solution

Let Υ_i denote a matrix in which all the entries are zero except the one located at the i -th row and i -th column. The domain \mathbf{V}_1 can be expressed as

$$\begin{aligned}
& |e_{1,k}| \leq \gamma_{1,k} \\
\Leftrightarrow & e_{1,k}e_{1,k}^H \leq \gamma_{1,k}^2 \\
\Leftrightarrow & \Upsilon_k \bullet (\mathbf{e}_1^H \mathbf{e}_1) \leq \gamma_{1,k}^2 \quad (3.9) \\
\Leftrightarrow & \mathbf{e}_1 \Upsilon_k \mathbf{e}_1^H \leq \gamma_{1,k}^2, \forall k = 1, \dots, m. \quad (3.10)
\end{aligned}$$

The Υ_k in (3.9) is used to select the entry of $\mathbf{e}_1^H \mathbf{e}_1$ at the k -th row and the k -th column. Since \mathbf{e}_1 is a vector, the entry of $\mathbf{e}_1^H \mathbf{e}_1$ at the k -th row and the k -th column is $|e_{1,k}|^2$, then we can use the element-wise product to select this value with Υ_k . Inequality (3.10) holds because $\Upsilon_k \bullet (\mathbf{e}_1^H \mathbf{e}_1) = \text{tr}(\Upsilon_k^T (\mathbf{e}_1^H \mathbf{e}_1)) = \mathbf{e}_1 \Upsilon_k \mathbf{e}_1^H$. Note that this transformation is not applicable for sum interference power calculation with multiple receive antennas, because for an arbitrary matrix \mathbf{Z} , the k -th diagonal entry of $\mathbf{Z}\mathbf{Z}^H$ is determined by all the entries at the k -th row of \mathbf{Z} instead of any single entry.

Now the element-wise constraints of channel error have been represented with respect to the whole vector \mathbf{e}_1 instead of its elements $e_{1,k}$. Assuming that $(\bar{\mathbf{g}}_1 + \mathbf{e}_1)\mathbf{Q}(\bar{\mathbf{g}}_1 + \mathbf{e}_1)^H \leq \xi_1$ is satisfied given $\mathbf{e}_1 \in \mathbf{V}_1$, there must exist

$\tau_k \geq 0$, $k = 1, \dots, m$, such that

$$S(\mathbf{Q}, \tau, \mathbf{e}_1) = \xi_1 - (\bar{\mathbf{g}}_1 + \mathbf{e}_1)\mathbf{Q}(\bar{\mathbf{g}}_1 + \mathbf{e}_1)^H - \sum_{k=1}^m \tau_k (\gamma_{1,k}^2 - \mathbf{e}_1 \mathbf{\Upsilon}_k \mathbf{e}_1^H) \geq 0, \forall \mathbf{e}_1 \in \mathbf{V}_1. \quad (3.11)$$

Therefore, we can set τ_k as an arbitrary nonnegative number if $\mathbf{e}_1 \mathbf{\Upsilon}_k \mathbf{e}_1^H = \gamma_{1,k}^2$, and set $\tau_k = 0$ if $\mathbf{e}_1 \mathbf{\Upsilon}_k \mathbf{e}_1^H < \gamma_{1,k}^2$. Actually τ_k is the Lagrange multiplier associated with the constraint $\mathbf{e}_1 \mathbf{\Upsilon}_k \mathbf{e}_1^H \leq \gamma_{1,k}^2$.

This inequality can be expressed in the form of matrix computation:

$$\begin{aligned} & \xi_1 - (\bar{\mathbf{g}}_1 + \mathbf{e}_1)\mathbf{Q}(\bar{\mathbf{g}}_1 + \mathbf{e}_1)^H - \sum_{k=1}^m \tau_k (\gamma_{1,k}^2 - \mathbf{e}_1 \mathbf{\Upsilon}_k \mathbf{e}_1^H) \geq 0 \\ \Leftrightarrow & \begin{bmatrix} 1 & \mathbf{e}_1 \end{bmatrix} \begin{bmatrix} \xi_1 - \sum_{k=1}^m \tau_k \gamma_k^2 - \bar{\mathbf{g}}_1 \mathbf{Q} \bar{\mathbf{g}}_1^H & -\bar{\mathbf{g}}_1 \mathbf{Q} \\ -\mathbf{Q} \bar{\mathbf{g}}_1^H & \sum_{k=1}^m \tau_k \mathbf{\Upsilon}_k - \mathbf{Q} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{e}_1^H \end{bmatrix} \geq 0, \\ & \forall \mathbf{e}_1 \in \mathbf{V}_1. \end{aligned} \quad (3.12)$$

Since the channel uncertainty is defined in a convex set and the interference power expression is convex, it has been proved in [84] that (3.12) can be equivalently transformed to the following inequality:

$$\begin{bmatrix} \xi_1 - \sum_{k=1}^m \tau_k \gamma_k^2 - \bar{\mathbf{g}}_1 \mathbf{Q} \bar{\mathbf{g}}_1^H & -\bar{\mathbf{g}}_1 \mathbf{Q} \\ -\mathbf{Q} \bar{\mathbf{g}}_1^H & \sum_{k=1}^m \tau_k \mathbf{\Upsilon}_k - \mathbf{Q} \end{bmatrix} \succeq \mathbf{0}, \exists \tau_k \geq 0, \forall k = 1, \dots, m. \quad (3.13)$$

Therefore, the problem can be formulated as

$$\begin{aligned} \max_{\mathbf{Q} \succeq \mathbf{0}} & \quad \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \\ \text{s.t.} & \quad \text{tr}(\mathbf{Q}) \leq P, \\ & \quad \begin{bmatrix} \xi_1 - \sum_{k=1}^m \tau_k \gamma_k^2 - \bar{\mathbf{g}}_1 \mathbf{Q} \bar{\mathbf{g}}_1^H & -\bar{\mathbf{g}}_1 \mathbf{Q} \\ -\mathbf{Q} \bar{\mathbf{g}}_1^H & \sum_{k=1}^m \tau_k \mathbf{\Upsilon}_k - \mathbf{Q} \end{bmatrix} \succeq \mathbf{0}, \\ & \quad \exists \tau_k \geq 0, \forall k = 1, \dots, m. \end{aligned} \quad (3.14)$$

It is a convex problem, and hence standard convex optimization algorithms can be adopted to solve the solution.

3.3.2 Optimality Conditions

Although we can obtain the optimal numerical solution, the optimality conditions and the physical meanings behind are not clear yet.

There are two constraints in (P3.2), one is transmit power constraint (3.8b), the other is worst case interference power constraint (3.8c). The worst case constraint (3.8c) is equivalent to (3.13), and can be decomposed as two inequalities according to the Schur complement formula:

$$\begin{aligned} \sum_{k=1}^m \tau_k \mathbf{r}_k - \mathbf{Q} & \succeq \mathbf{0}, \\ (\xi_1 - \sum_{k=1}^m \tau_k \gamma_k^2 - \bar{\mathbf{g}}_1 \mathbf{Q} \bar{\mathbf{g}}_1^H) - \bar{\mathbf{g}}_1 \mathbf{Q} (\sum_{k=1}^m \tau_k \mathbf{r}_k - \mathbf{Q})^{-1} \mathbf{Q} \bar{\mathbf{g}}_1^H & \geq 0. \end{aligned}$$

If the interference power constraint is active, then equality holds at least in one of the above inequalities. Note that in this case, the maximum possible interference power is exactly equal to ξ_1 , and ξ_1 only exists in the second inequality. Obviously, the second inequality must be zero, otherwise we can just reduce the value of ξ_1 without changing other parameters, and the interference constraint is not ruined. In another word, we can achieve the same throughput with a smaller ξ_1 , which contradicts to the assumption that the worst case interference constraint is active at $\mathbf{Q} = \mathbf{Q}^*$, where \mathbf{Q}^* denotes the optimal \mathbf{Q} . Therefore, $(\xi_1 - \sum_{k=1}^m \tau_k \gamma_k^2 - \bar{\mathbf{g}}_1 \mathbf{Q} \bar{\mathbf{g}}_1^H) - \bar{\mathbf{g}}_1 \mathbf{Q} (\sum_{k=1}^m \tau_k \mathbf{r}_k - \mathbf{Q})^{-1} \mathbf{Q} \bar{\mathbf{g}}_1^H = 0$ always holds whenever (3.13) is active at $\mathbf{Q} = \mathbf{Q}^*$, where τ_k^* means the corresponding Lagrangian multiplier.

Now we will discuss in which optimality condition $\sum_{k=1}^m \tau_k \mathbf{E}_k - \mathbf{Q}$ would be a positive semi-definite matrix but not a positive definite matrix. Supposing that the worse case \mathbf{e}_1 at \mathbf{Q}^* is denoted by \mathbf{e}_1^* . This value can be obtained by calculating the derivative of (3.11) with respect to \mathbf{e}_1 and setting it to zero:

$$\begin{aligned} \frac{\partial S(\mathbf{Q}^*, \tau^*, \mathbf{e}_1^*)}{\partial \mathbf{e}_1^*} &= -2(\bar{\mathbf{g}}_1 + \mathbf{e}_1^*) \mathbf{Q}^* + 2\mathbf{e}_1^* \left(\sum_{k=1}^m \tau_k^* \mathbf{r}_k \right) = 0 \\ \Leftrightarrow \left(\sum_{k=1}^m \tau_k^* \mathbf{r}_k - \mathbf{Q}^* \right) \mathbf{e}_1^* &= \bar{\mathbf{g}}_1 \mathbf{Q}^*. \end{aligned} \quad (3.15)$$

If $\sum_{k=1}^m \tau_k^* \mathbf{r}_k - \mathbf{Q}^* \succ \mathbf{0}$, we have $\mathbf{e}_1^* = \bar{\mathbf{g}}_1 \mathbf{Q}^* (\sum_{k=1}^m \tau_k^* \mathbf{r}_k - \mathbf{Q}^*)^{-1}$, which means that the worst \mathbf{e}_1^* is unique. Otherwise, since $\sum_{k=1}^m \tau_k^* \mathbf{r}_k - \mathbf{Q}^*$ is not of full rank, there are multiple \mathbf{e}_1^* satisfying (3.15), and thus there are multiple \mathbf{e}_1^* leading to the worst interference power at $\mathbf{Q} = \mathbf{Q}^*$.

If P is so small that only (3.8b) is active, then actually ST transmits as if the PR is absent, so the waterfilling algorithm can provide the analytical solution. If P is so large that the transmit power constraint can be ignored, the problem will be limited to the channel conditions and the error range, then P will have no impact on the throughput.

3.4 Multiple Single-Antenna PRs

The solution of the throughput maximization problem with multiple single-antenna PRs are similar to the one PR case, although the formulation and the optimality conditions are more complicated. By applying S-procedure, the problem (P3.1) can be reformulated as

$$(P3.3) \quad \max_{\mathbf{Q} \succeq \mathbf{0}} \quad \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H), \quad (3.16a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{Q}) \leq P, \quad (3.16b)$$

$$\begin{bmatrix} \xi_i - \sum_{k=1}^m \tau_{i,k} \gamma_{i,k}^2 - \bar{\mathbf{g}}_i \mathbf{Q} \bar{\mathbf{g}}_i^H & -\bar{\mathbf{g}}_i \mathbf{Q} \\ -\mathbf{Q} \bar{\mathbf{g}}_i^H & \sum_{k=1}^m \tau_{i,k} \mathbf{r}_k - \mathbf{Q} \end{bmatrix} \succeq \mathbf{0},$$

$$\exists \tau_{i,k} \geq 0, \forall k = 1, \dots, m, \forall i = 1, \dots, q. \quad (3.16c)$$

At the optimal \mathbf{Q} , not all the constraints in (3.16c) are active, therefore only the active interference constraints satisfy the optimality conditions discussed in Section 3.3.2.

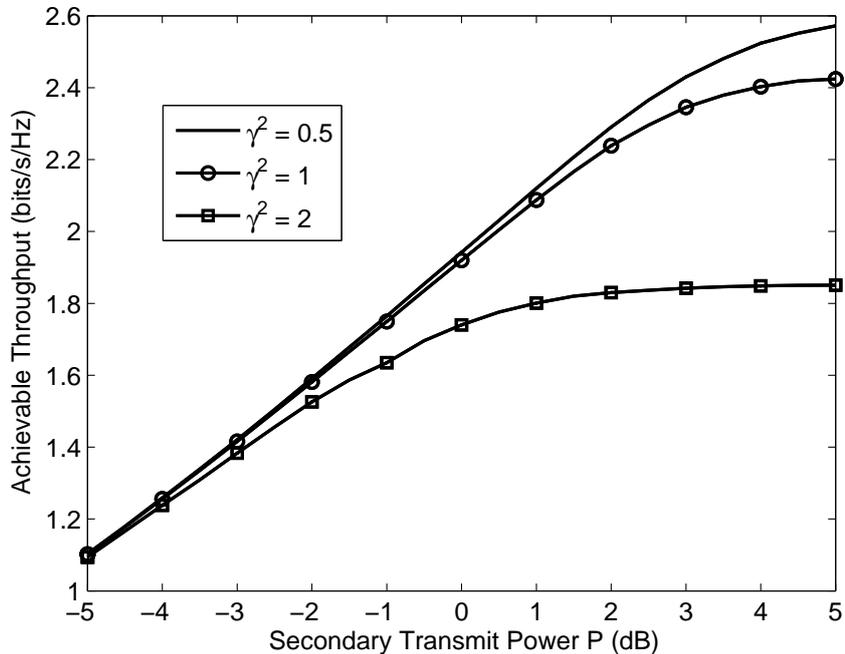


Figure 3.1: Achievable rate versus transmit power at $\xi = 4$.

3.5 Simulation Results

This section provides some simulation results of the proposed algorithm for performance evaluation. There is one single-antenna PR in the cognitive network. Both ST and SR are equipped with 2 antennas. All entries of \mathbf{H} and $\bar{\mathbf{g}}_1$ are i.i.d random variables following the distribution $\mathcal{CN}(0, 1)$. For simplicity, all the entries of \mathbf{e}_1 have the same channel error bounds, i.e., $|e_{1,1}| = |e_{1,2}| = \gamma$.

Fig. 3.1 plots the secondary throughput against the maximum transmit power P . It shows that there is almost no difference among the throughput corresponding to different values of γ^2 at $P = -5$ dB. This is because under the given γ^2 values, the interference constraint is hardly active and the channel uncertainty has little impact during the optimization of \mathbf{Q} due to the small P . As P increases from -5 dB to -1 dB, the performance gaps

between the system of $\gamma^2 = 2$ and the other two systems become larger and larger, while the systems of $\gamma^2 = 0.5$ and $\gamma^2 = 1$ still almost share the same throughput. At the region $P > -1$ dB, we can see an obvious gap between the performance of $\gamma^2 = 1$ and $\gamma^2 = 0.5$, and the gap grows with the increase of P . The performance difference mainly comes from the channel uncertainty. The element-wise channel error model defines a rectangle region centered at $\bar{\mathbf{g}}_1$. A higher γ^2 means a larger uncertainty region and the interference constraint must be satisfied at every point inside, so larger performance loss will be caused. When P is large enough corresponding to a particular nonzero value of γ^2 , the interference constraint will be dominant to the optimization of \mathbf{Q} , then the increase of P will nearly not affect the system performance. Note that when the antenna number of ST is larger than the sum of antenna number at PRs, as P approaches ∞ , there is not an upper bound for the system throughput with perfect knowledge of \mathbf{g}_1 at ST. That is because by using zero forcing beamforming, the secondary signals can be projected to the null space to eliminate interference at PRs, then the transmit power can be infinite to achieve arbitrary large throughput. As a result, a more precise channel estimation leads to a higher throughput. But if the antenna number of ST is equal to or less than the sum of antenna number at PRs, the interference is unavoidable, therefore the throughput will be bounded at large P .

The secondary throughput versus ξ is demonstrated in Fig. 3.2. Both the perfect CSI and partial CSI cases are listed for comparison. For the partial CSI case, we set $\gamma^2 = 1$. Given a P , the throughput will increase with the increase of ξ , and finally reach upper limit when the interference constraint can be ignored with a sufficiently large ξ . The bound can be calculated by the conventional MIMO transmission with power constraint only. As ξ approaches 0, the throughput in the perfect CSI scenario is nonzero because zero forcing can be used to cancel the interference power, but the throughput

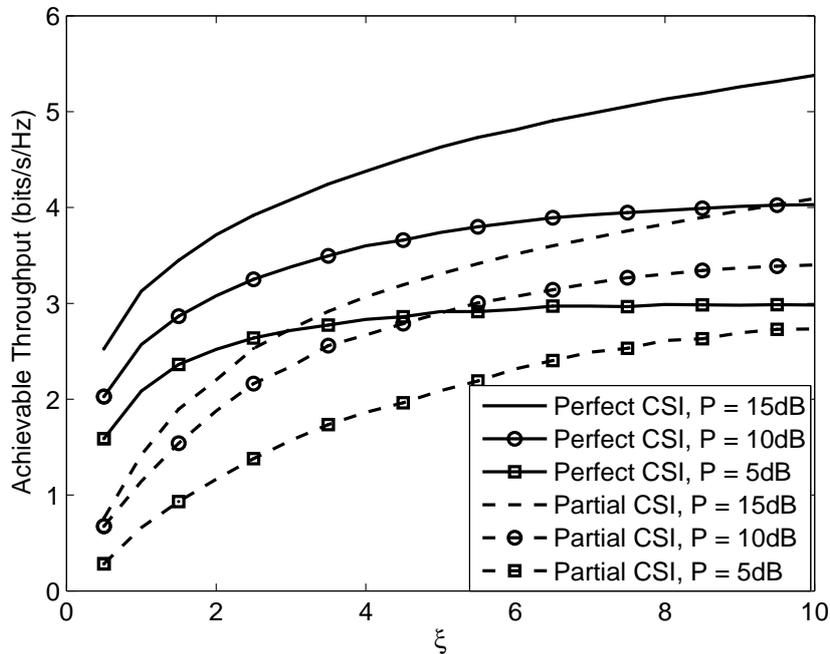


Figure 3.2: Achievable rate versus interference threshold.

in the partial CSI scenario would reduce to zero due to the existence of channel uncertainty. It is shown that at $\xi = 0.5$, the curves of $P = 15$ dB and $P = 10$ dB in the partial CSI case become very close. The reason is that ST has to restrict its transmit to satisfy the interference constraint in the low ξ region, and thus the benefit of a higher transmit power is limited.

3.6 Summary

In this chapter, we studied the element-wise channel error model in cognitive transmission with multiple data streams, where all PRs are equipped with single antenna and the secondary throughput is to be maximized. We proved that the channel uncertainty region can be equivalently transformed to a new formulation with finite number of constraints. Since the new problem is

convex, the global optimal solution can be obtained with existing algorithms. The optimality conditions were discussed as well. Finally, we provided the simulation results to demonstrate the impacts of different variables on the system performance.

Chapter 4

Throughput Maximization: No Instantaneous CSI

As discussed in Chapter 3, CSI knowledge is critical for cognitive beamforming. However, not all kinds of existing primary systems are designed to provide opportunities for SUs to obtain the CSI knowledge. As a result, if a secondary system is to be deployed in an area where there are such PUs sharing the same frequency spectrum, it is a must to estimate the potential interference to the PUs by the owner of the secondary system. For example, we may assume that the statistical CSI information of the PUs can be collected by the secondary system owner's investigation before the deployment, so we do not know the instantaneous CSI but know the statistical CSI.

In this chapter, we consider a spectrum sharing scenario where secondary CSI (ST-SR channel) is available but the instantaneous cross CSI (ST-PR channel) is unavailable at ST. This assumption is practical since most current primary devices are not designed to help ST estimate or obtain cross CSI from PUs. When only statistical cross CSI is available, authors of [85] consider the estimation of CSI knowledge in a blind manner, and design the pre- and post-coding matrices based on interference alignment for SUs to access

the spectrum. However, the interference alignment based channel estimation and beamforming design are not optimal in cognitive radio. The reasons are listed as below:

1. In conventional interference alignment, all transmitters are assumed to share the spectrum with the same priority. They cooperatively design transmit beamforming matrices/vectors to eliminate interference at receivers.

In cognitive radio, PUs have no obligation to cooperate with SUs. Instead, since PUs have paid for the spectrum license, they just transmit without considering the existence of SUs and their QoS must be guaranteed. SUs should carefully design their transmit policies to control the interference power at PRs below a certain threshold while maximizing secondary performance. It means that the requirements of transmit beamforming design of cognitive radio will be different from conventional interference alignment.

2. Interference alignment seeks perfect alignment and interference cancellation. Zero forcing beamforming is an important technique to achieve this goal. The optimality of this approach can be guaranteed in the high SNR regime, but the performance loss is significant in the regime of low or medium SNR.

Spectrum sharing in cognitive radio does not try to totally eliminate interference. In contrast, SUs prefer causing slight interference at PR to gain significant performance improvement. Consequently, SUs usually do not consider the simple zero forcing beamforming method, but wish to apply some more complicated methods to fully utilize spectrum access opportunities. To this end, power control will be very important in spectrum sharing.

3. Because interference alignment is optimal in the high SNR regime,

while secondary systems usually operate in the low or medium SNR regime, using interference alignment will not be an optimal option in cognitive radio.

Comparing with interference alignment and conventional MIMO communications, the novelty of this chapter is that when multiple data streams are received by one SR, we theoretically derive the expression of primary interference power, and thus design algorithms to fully exploiting the primary interference constraint to maximize secondary throughput. In the area of interference alignment, interference is to be canceled after pre- and post-coding, therefore the primary interference threshold ξ is not consider and the existing results do not offer any idea of utilizing the primary interference constraint. In the area of conventional MIMO communications, the transmitter is required to transmit without instantaneous CSI between the receiver. The transmitter analyzes the statistical CSI to maximize transmission performance over this channel. However, in cognitive radio, the ST-SR channel is usually assumed available while the ST-PR channel is not easy to obtain. As a result, the uncertainties of the ST-PR channel and the primary interference power serve as constraints, and ST needs to maintain the interference within a certain region rather than maximizing or minimizing it. Since the channel uncertainties in MIMO communications and cognitive communications play different roles, problem formulation and theoretical derivation are different.

In this chapter, we study cognitive beamforming for secondary throughput maximization by exploiting the primary interference constraints. Two types of primary interference constraints are considered: one is on the expectation of interference at PR, referred to as interference expectation constraint, and the other is on the interference outage probability (i.e., the probability of interference at PR exceeding a predetermined threshold), referred to as interference outage constraint. We find that there is a closed-form solution for beamforming subject to interference expectation constraint, while

for beamforming subject to interference outage constraint, the interference power is related to the eigenvalues of the secondary covariance matrix and can be expressed as a linear combination of multiple chi-square distributed variables. By exploiting the impact of the eigenvalues on the interference outage probability and establishing the relationship between the vibrations of secondary covariance matrix and its eigenvalues, we design an algorithm for this beamforming problem. Numerical results are provided to demonstrate the performance of the proposed beamforming solutions subject to both interference expectation and interference outage constraints.

4.1 System Model

We consider a cognitive radio network which consists of a primary user pair (PT-PR) and a secondary user pair (ST-SR). The signals received at PR and SR are respectively given by

$$\begin{aligned}\mathbf{y}_p &= \Xi \mathbf{x}_p + \mathbf{G} \mathbf{x}_s + \mathbf{n}_p \\ \mathbf{y}_s &= \mathbf{H} \mathbf{x}_s + \mathbf{n}_s\end{aligned}$$

where $\mathbf{x}_p \in \mathbb{C}^{p \times 1}$ and $\mathbf{x}_s \in \mathbb{C}^{m \times 1}$ denote the primary and secondary transmitted signal, respectively; \mathbf{n}_p and \mathbf{n}_s are the additive complex gaussian noise with zero mean and unit variance at PR and SR, respectively. Let $\Xi \in \mathbb{C}^{q \times p}$ denote the PT-PR channel and $\mathbf{G} \in \mathbb{C}^{q \times m}$ denote the ST-PR channel. All the channel coefficient matrices Ξ , \mathbf{G} , and \mathbf{H} are assumed constant during one symbol period and different from one period to another. Suppose that the channel coefficient matrices are normalized, and thus all the entries of Ξ , \mathbf{H} and \mathbf{G} are independent and identically distributed (i.i.d.) and follow complex normal Gaussian distribution, i.e., $\Xi_{ij}, \mathbf{H}_{ij}, \mathbf{G}_{ij} \sim \mathcal{CN}(0, 1)$. We assume that the secondary CSI (\mathbf{H}) is known at ST, while the cross CSI (\mathbf{G}) is unknown at ST except its second-order statistics.

Suppose that there are d secondary data streams, with $d \leq \min(m, n)$ where m and n represent the number of antennas at ST and SR, respectively. Letting $\mathbf{s}_s \in \mathbb{C}^{d \times 1}$ denote the secondary information to be transmitted with $\mathbb{E}(\mathbf{s}_s \mathbf{s}_s^H) = \mathbf{I}$, and $\mathbf{T} \in \mathbb{C}^{m \times d}$ denote the secondary transmit beamforming matrix, \mathbf{x}_s can be expressed as $\mathbf{x}_s = \mathbf{T} \mathbf{s}_s$. The covariance matrix of the secondary transmitted signals \mathbf{Q} follows that $\mathbf{Q} = \mathbf{T} \mathbf{T}^H$.

To protect primary transmission, ST should not cause harmful interference to PR. As mentioned earlier, we consider two types of interference constraints in this chapter. The interference expectation constraint is given by

$$\mathbb{E}_{\mathbf{G}}\{\text{tr}(\mathbf{G} \mathbf{Q} \mathbf{G}^H)\} \leq \xi \quad (4.1)$$

where ξ is the interference threshold at PR. The interference outage constraint is given by

$$\mathcal{P}(\text{tr}(\mathbf{G} \mathbf{Q} \mathbf{G}^H) \geq \xi) \leq \mathcal{P}_0 \quad (4.2)$$

where \mathcal{P}_0 is the maximum allowable interference outage probability at PR. Note that since \mathbf{G} is unknown at ST and unbounded, \mathbf{Q} can only be set to $\mathbf{0}$ for any $\xi \geq 0$ if $\mathcal{P}_0 = 0$.

The objective of this chapter is to find the optimal secondary covariance matrix maximizing the secondary throughput.

Subject to the interference expectation constraint, our problem is formulated as

$$\begin{aligned} \text{(P4.1)} \quad & \max_{\mathbf{Q} \succeq \mathbf{0}} && \log \det(\mathbf{I} + \mathbf{H} \mathbf{Q} \mathbf{H}^H) \\ & \text{s.t.} && \text{tr}(\mathbf{Q}) \leq P \\ & && \mathbb{E}_{\mathbf{G}}(\text{tr}(\mathbf{G} \mathbf{Q} \mathbf{G}^H)) \leq \xi. \end{aligned}$$

where P is the secondary power constraint and $\mathbf{Q} \succeq \mathbf{0}$ denotes that \mathbf{Q} is a positive semi-definite matrix.

Subject to the interference outage constraint, the problem is formulated

as

$$\begin{aligned}
\text{(P4.2)} \quad & \max_{\mathbf{Q} \succeq \mathbf{0}} \quad \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \\
& \text{s.t.} \quad \text{tr}(\mathbf{Q}) \leq P \\
& \mathcal{P}(\text{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^H) \geq \xi) \leq \mathcal{P}_0.
\end{aligned}$$

4.2 Interference Expectation Constraint

The interference expectation constraint in (4.1) can be interpreted as a long term interference constraint. Since the channel \mathbf{G} is unavailable at ST, the interference at PR can be rewritten as

$$\begin{aligned}
& \text{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^H) \\
& = \text{tr}(\mathbf{U}_\mathbf{Q}\mathbf{A}_\mathbf{Q}\mathbf{U}_\mathbf{Q}^H\mathbf{G}^H\mathbf{G}) \\
& = \text{tr}(\mathbf{A}_\mathbf{Q}\hat{\mathbf{G}}^H\hat{\mathbf{G}}) \\
& \triangleq \text{tr}(\mathbf{A}_\mathbf{Q}\mathbf{M}_1)
\end{aligned} \tag{4.3}$$

where $\mathbf{Q} = \mathbf{U}_\mathbf{Q}\mathbf{A}_\mathbf{Q}\mathbf{U}_\mathbf{Q}^H$ represents the eigenvalue decomposition of \mathbf{Q} , and $\mathbf{U}_\mathbf{Q} \in \mathbb{C}^{m \times d}$, $\mathbf{U}_\mathbf{Q}^H\mathbf{U}_\mathbf{Q} = \mathbf{I}$, $\mathbf{A}_\mathbf{Q} \in \mathbb{C}^{d \times d}$ is diagonal, $\hat{\mathbf{G}} \triangleq \mathbf{G}\mathbf{U}_\mathbf{Q}$. As a result, entries of $\hat{\mathbf{G}}$ are i.i.d. and follow complex normal Gaussian distribution. Then \mathbf{M}_1 is a matrix following Wishart distribution. The diagonal entries of \mathbf{M}_1 are chi-square distributed with $2q$ degrees of freedom. Since $\text{tr}(\mathbf{A}_\mathbf{Q}\mathbf{M}_1)$ is only related to the diagonal entries of \mathbf{M}_1 , (4.3) can be rewritten as

$$\text{tr}(\mathbf{A}_\mathbf{Q}\mathbf{M}_1) = \sum_{k=1}^d a_k w_k, 2w_k \in \mathcal{X}^2(2q), \forall k = 1, \dots, d \tag{4.4}$$

where a_k and w_k denote the k -th diagonal entries of $\mathbf{A}_\mathbf{Q}$ and \mathbf{M}_1 , respectively. This reflects that when the instantaneous ST-PR channel is unknown, the interference at PR is equivalent to a linear combination of chi-square distributed variables.

By substituting (4.4) into (4.1), we have

$$\begin{aligned}
& \mathbb{E}_{\mathbf{G}}(\text{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^H)) \\
&= \mathbb{E}_{w_k} \left(\sum_{k=1}^d a_k w_k \right) \\
&= \sum_{k=1}^d a_k \mathbb{E}(w_k) = q \sum_{k=1}^d a_k \stackrel{(a)}{=} q \text{tr}(\mathbf{Q}) \leq \xi
\end{aligned}$$

where (a) is because the trace of a matrix is equal to the sum of its eigenvalues. As a result, (P4.1) can be reformulated as

$$\begin{aligned}
& \max_{\mathbf{Q} \succeq \mathbf{0}} \quad \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \\
& \text{s.t.} \quad \text{tr}(\mathbf{Q}) \leq \min\left(\frac{\xi}{q}, P\right).
\end{aligned}$$

Obviously, (P4.1) has been reduced to a conventional MIMO throughput maximization problem, which can be easily solved via the water-filling method. Note that there exists a critical point at $P = \xi/q$. If $P < \xi/q$, the throughput is determined by P only, otherwise it is determined by ξ only.

4.3 Interference Outage Constraint

4.3.1 Problem Reformulation

We consider the case of $m > n$ first. Let \mathbf{V}_{\perp} denote a $m \times (m - n)$ complex matrix such that $\mathbf{V}_{\perp}^H \mathbf{H} = \mathbf{0}$ and $\mathbf{V}_{\perp}^H \mathbf{V}_{\perp} = \mathbf{I}$. Let $\mathbf{V}_{//} \in \mathbb{C}^{m \times n}$ and $[\mathbf{V}_{\perp} \ \mathbf{V}_{//}][\mathbf{V}_{\perp} \ \mathbf{V}_{//}]^H = \mathbf{I}$. Note that such \mathbf{V}_{\perp} and $\mathbf{V}_{//}$ are not unique and any matrices satisfying these conditions will work. Let us decompose the transmit beamforming matrix \mathbf{T} as

$$\mathbf{T} = \mathbf{T}_{//} + \mathbf{T}_{\perp} \tag{4.5}$$

where $\mathbf{V}_{//}^H \mathbf{T}_{\perp} = \mathbf{0}$ and $\mathbf{V}_{\perp}^H \mathbf{T}_{//} = \mathbf{0}$. It means that \mathbf{T} is decomposed as two parts, in which $\mathbf{T}_{//}$ lies in the same signal subspace as \mathbf{H} while \mathbf{T}_{\perp} lies in

the null signal subspace of \mathbf{H} . Therefore, we have $\mathbf{T}_\perp^H \mathbf{T}_{//} = \mathbf{0}$ and

$$\begin{aligned} \mathbf{Q} &= \mathbf{T}\mathbf{T}^H \\ &= \mathbf{T}_{//}\mathbf{T}_{//}^H + (\mathbf{T}_{//}\mathbf{T}_\perp^H + \mathbf{T}_\perp\mathbf{T}_{//}^H + \mathbf{T}_\perp\mathbf{T}_\perp^H). \end{aligned} \quad (4.6)$$

We define $\mathbf{Q}_1 \triangleq \mathbf{T}_{//}\mathbf{T}_{//}^H$ and $\mathbf{Q}_2 \triangleq \mathbf{T}_{//}\mathbf{T}_\perp^H + \mathbf{T}_\perp\mathbf{T}_{//}^H + \mathbf{T}_\perp\mathbf{T}_\perp^H$. Clearly, \mathbf{Q}_1 is a positive semi-definite matrix but \mathbf{Q}_2 may not be. It is easy to derive that

$$\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) = \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_1\mathbf{H}^H) \quad (4.7)$$

and thus \mathbf{Q}_1 determines the system throughput while \mathbf{Q}_2 makes no difference to it.

Theorem 4.1. For any solution \mathbf{Q} , we have

$$\text{tr}(\mathbf{Q}) \geq \text{tr}(\mathbf{Q}_1) \quad (4.8a)$$

$$\mathcal{P}(\text{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^H) \geq \xi) \geq \mathcal{P}(\text{tr}(\mathbf{G}\mathbf{Q}_1\mathbf{G}^H) \geq \xi). \quad (4.8b)$$

The equalities hold if and only if $\mathbf{Q}_2 = \mathbf{0}$.

Proof. See Appendix A. □

Remarks: Theorem 4.1 suggests that a nonzero \mathbf{Q}_2 not only consumes more transmit power but also causes a higher outage probability as compared to the zero \mathbf{Q}_2 . Only \mathbf{Q}_1 contributes to the optimal solution and \mathbf{Q}_2 can be neglected when solving the problem. Note that this conclusion only establishes when \mathbf{G} is completely unavailable. For a particular realization of \mathbf{G} , regardless of whether it is perfectly or partially known at ST, usually the optimal beamforming matrix \mathbf{T} contains a nonzero part in the signal subspace orthogonal to \mathbf{H} .

For optimal beamforming, we see that equalities in (4.8) hold, then \mathbf{T} can be rewritten as $\mathbf{T} = \mathbf{V}_{//}\mathbf{P}$ and \mathbf{Q} is given by

$$\mathbf{Q} = \mathbf{V}_{//}\mathbf{X}\mathbf{V}_{//}^H \quad (4.9)$$

where $\mathbf{X} = \mathbf{P}\mathbf{P}^H$ is a positive semi-definite Hermitian matrix. The eigenvalues of \mathbf{X} are exactly equal to those of \mathbf{Q} , and the transmit power $\text{tr}(\mathbf{Q}) = \text{tr}(\mathbf{X}) = \sum_{k=1}^d \lambda_k(\mathbf{X})$. By substituting (4.9) into (P4.2) and defining $\mathbf{M}_2 \triangleq \mathbf{V}_{//}^H \mathbf{H}^H \mathbf{H} \mathbf{V}_{//}$, (P4.2) is reformulated as

$$(P4.3) \quad \max_{\mathbf{X} \succeq \mathbf{0}} \quad \log \det(\mathbf{I} + \mathbf{M}_2 \mathbf{X}) \quad (4.10a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{X}) \leq P \quad (4.10b)$$

$$\mathcal{P}\left(\sum_{k=1}^d \lambda_i(\mathbf{X}) w_k \geq \xi\right) \leq \mathcal{P}_0. \quad (4.10c)$$

Note that since \mathbf{H} is random, it is reasonable to assume that \mathbf{M}_2 is of full rank in (P4.3). For the case of $m \leq n$, \mathbf{V}_\perp does not exist, then we simply let $\mathbf{M}_2 \triangleq \mathbf{H}^H \mathbf{H} \in \mathbb{C}^{m \times m}$, which is also a full-rank matrix. Therefore, (P4.3) is applicable regardless of the exact values of m and n .

4.3.2 Algorithm for the Local Optimum

The constraint (4.10c) represents a linear combination of chi-square distributed variables. For a particular \mathbf{X} , we need the cumulative distribution function (CDF) or probability density function (PDF) of $\sum_{k=1}^d \lambda_i(\mathbf{X}) w_k$ to check if the given \mathbf{X} is feasible. The CDF of such expression has been provided in [86] for the case of $d = 2, 3$ and in [87] and [88] for $d = 4, 5$. For example, when $d = 2$, the CDF is given by

$$\begin{aligned} & \mathcal{P}(b_1 w_1 + b_2 w_2 \leq \xi) \\ &= \frac{2}{\sqrt{\frac{1}{b_1} + \frac{1}{b_2}}} \int_0^{\frac{\xi}{4}(\frac{1}{b_1} + \frac{1}{b_2})} e^{-x} I_0\left(\frac{b_2 - b_1}{b_2 + b_1} x\right) dx \end{aligned} \quad (4.11)$$

where $b_1 \leq b_2$ and $b_1 + b_2 = 1$, $I_0(\cdot)$ is the Bessel function of the first kind with order zero. For the case $b_1 + b_2 \neq 1$, the corresponding CDF expression can be easily obtained from (4.11). To the best of our knowledge, there is no simple expression for the PDF or CDF of $\sum_{k=1}^d \lambda_i(\mathbf{X}) w_k$ for all $d \geq 0$, and the CDF is very complicated to derive for $d \geq 6$.

Generally speaking, it is unlikely to have a closed-form optimal solution for (P4.3). Although having not been mathematically proved, intuitively (4.10c) is not a convex/concave function of \mathbf{X} . Therefore it is unsuitable to solve the Lagrangian dual problem to obtain the global optimum [73]. In this chapter, we propose an algorithm to obtain the local optimum for the problem.

We can see that the objective function (4.10a) and constraint (4.10b) are related to \mathbf{X} but the constraint (4.10c) are only related to its eigenvalues. Before developing the algorithm, let us show how the change of \mathbf{X} affects its eigenvalues first. Let $\mathbf{X} = \mathbf{U}_x \mathbf{A}_x \mathbf{U}_x^H$ denote the eigenvalue decomposition of \mathbf{X} . Assuming that there is a vibration direct $\Delta \mathbf{X}$ ($\Delta \mathbf{X}$ is Hermitian) over \mathbf{X} , the corresponding change directions of the \mathbf{U}_x and \mathbf{A}_x are denoted by $\Delta \mathbf{U}_x$ and $\Delta \mathbf{A}_x$, respectively. For a sufficiently small t , we have

$$\begin{aligned} & \mathbf{X} + t\Delta \mathbf{X} \\ = & (\mathbf{U}_x + t\Delta \mathbf{U}_x)(\mathbf{A}_x + t\Delta \mathbf{A}_x)(\mathbf{U}_x + t\Delta \mathbf{U}_x)^H + O(t^2). \end{aligned}$$

From [15], we know that $\Delta \mathbf{U}_x$ is in the form of $\Delta \mathbf{U}_x = \mathbf{U}_x \mathbf{C}$ where \mathbf{C} is a skew-Hermitian matrix ($\mathbf{C}^H = -\mathbf{C}$). The difference of eigenvalues is given by

$$\begin{aligned} t\Delta \mathbf{A}_x &= (\mathbf{U}_x + t\Delta \mathbf{U}_x)^H (\mathbf{X} + t\Delta \mathbf{X}) (\mathbf{U}_x + t\Delta \mathbf{U}_x) \\ &\quad - (\mathbf{U}_x + t\Delta \mathbf{U}_x)^H \mathbf{X} (\mathbf{U}_x + t\Delta \mathbf{U}_x) + O(t^2) \\ &= t(\mathbf{C}^H \mathbf{U}_x^H \mathbf{X} \mathbf{U}_x + \mathbf{U}_x^H \mathbf{X} \mathbf{U}_x \mathbf{C}) + t(\mathbf{U}_x^H \Delta \mathbf{X} \mathbf{U}_x) \\ &\quad + O(t^2). \end{aligned}$$

Clearly, $t\Delta \mathbf{A}_x$ is a real diagonal matrix. Since \mathbf{C} is skew-Hermitian and $\mathbf{U}_x^H \mathbf{X} \mathbf{U}_x$ is Hermitian, it can be shown that the real parts of the diagonal entries of $\mathbf{C}^H \mathbf{U}_x^H \mathbf{X} \mathbf{U}_x + \mathbf{U}_x^H \mathbf{X} \mathbf{U}_x \mathbf{C}$ are all zeros, which means the vibration direction of eigenvalues is determined by the diagonal entries of $\mathbf{U}_x^H \Delta \mathbf{X} \mathbf{U}_x$. In the following, we show how to choose an appropriate $\Delta \mathbf{X}$.

Problem (P4.3) can be rewritten as an unconstrained problem by using the logarithmic barrier function method. This method combines objective and constraint functions together to generate a new expression, in which logarithm operation is applied to ensure all the original constraints being satisfied. Then we can use the gradient or subgradient algorithms to obtain a local optimal solution. This unconstrained problem is given by

$$(P4.4) \quad \max_{\mathbf{X} \succeq \mathbf{0}} f(\mathbf{X}) \triangleq \log \det(\mathbf{I} + \mathbf{M}_2 \mathbf{X}) \\ + \mu [\log(P - \text{tr}(\mathbf{X})) + \\ \log(\mathcal{P}_0 - \mathcal{P}(\sum_{k=1}^m \lambda_i(\mathbf{X}) w_k \geq \xi))]]$$

where $\mu \geq 0$ is a parameter that controls the approximation accuracy. It should be sufficiently small when the algorithm terminates.

The algorithm requires descent direction \mathbf{D} and step size γ in each iteration to gradually approach the local optimum. For the case of $m \leq 5$, we can set $\mathbf{D} = \partial f / \partial \mathbf{X}$, the derivative of f with respect to \mathbf{X} , as the descent direction. When $m > 6$, we should select a subgradient instead of gradient as the descent direction because the expression of gradient becomes unlikely to obtain. The step size can be chosen according to the Armijo rule [74], i.e.,

$$f(\mathbf{X} + \gamma \mathbf{D}) - f(\mathbf{X}) \geq \frac{1}{2} \gamma \text{tr}(\mathbf{D}^H \mathbf{D}) \\ f(\mathbf{X} + 2\gamma \mathbf{D}) - f(\mathbf{X}) \leq \gamma \text{tr}(\mathbf{D}^H \mathbf{D}).$$

We provide an algorithm below that uses the logarithmic barrier function to obtain the local optimal solution.

Algorithm 4.1

1. Set initial $\mathbf{X} := \mathbf{I}$, $\mu := 1$.
2. If μ is sufficiently small, stop. Otherwise go to Step 3.

3. Compute descent direction \mathbf{D} . If $\text{tr}(\mathbf{D}^H \mathbf{D})$ is sufficiently small, set $\mu := \mu/2$ and go to Step 2. Otherwise go to Step 4.
 4. Compute step size γ according to the Armijo rule.
 5. Update $\mathbf{X} := \mathbf{X} + \gamma \mathbf{D}$. Go to Step 3 to repeat.
-

4.4 Simulation Results

In this section, we present some numerical results to evaluate the performance of our proposed algorithm. We consider a CR system where ST is equipped with 3 antennas while SR and PR are equipped with 2 antennas, i.e., $m = 3$, $n = 2$ and $q = 2$.

Firstly we consider the interference expectation constraint. Fig. 4.1 illustrates the average secondary throughput versus secondary power constraint P and interference constraint ξ . We observe that at the low secondary power region, the secondary throughput grows as the secondary power increase, and the considered 3 different ξ values give exactly the same performance, because in this case, it depends on the secondary power only. When the secondary transmit power becomes larger than the critical point ($P = \xi/q$), the throughput is determined by ξ only and keeps constant as long as $P \geq \xi/q$. In this case, the critical point is larger with a larger ξ , and thus leading to a higher corresponding throughput as shown in the figure.

Fig. 4.2 compares the throughput subject to the interference outage constraint under the same ξ values. The secondary power constraint is varied from 0 dB to 20 dB while the interference outage constraint \mathcal{P}_0 is chosen to be 5%. It is shown that the throughput increases with the increase of secondary power constraint and finally reaches a limit when the transmit power is sufficiently large. Basically, Fig. 4.2 shows a similar behavior as that in

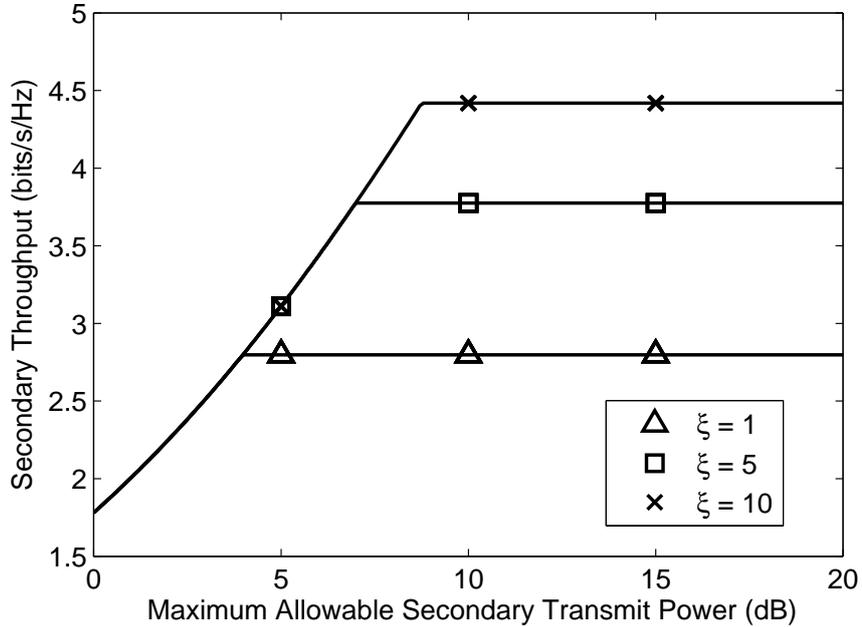


Figure 4.1: Secondary throughput subject to interference expectation constraint.

Fig. 4.1. The difference is that in Fig. 4.1, the throughput performance depends on either P or ξ only (there is a sharp edge on the curve due to the critical point), while in Fig. 4.2, we are able to observe the joint effect of them (the curve is smooth because of no critical point): at small P values, P is dominant; when P increases, we see that P and ξ jointly affect the achievable throughput; when P further increases, ξ becomes dominant. The main reason of the difference is that the interference expectation constraint consider expectation only, while the interference outage constraint is more related to the distribution of the cross channel, and it enables ST to use the cross channel statistics to adjust \mathbf{T} and reduce the interference outage probability.

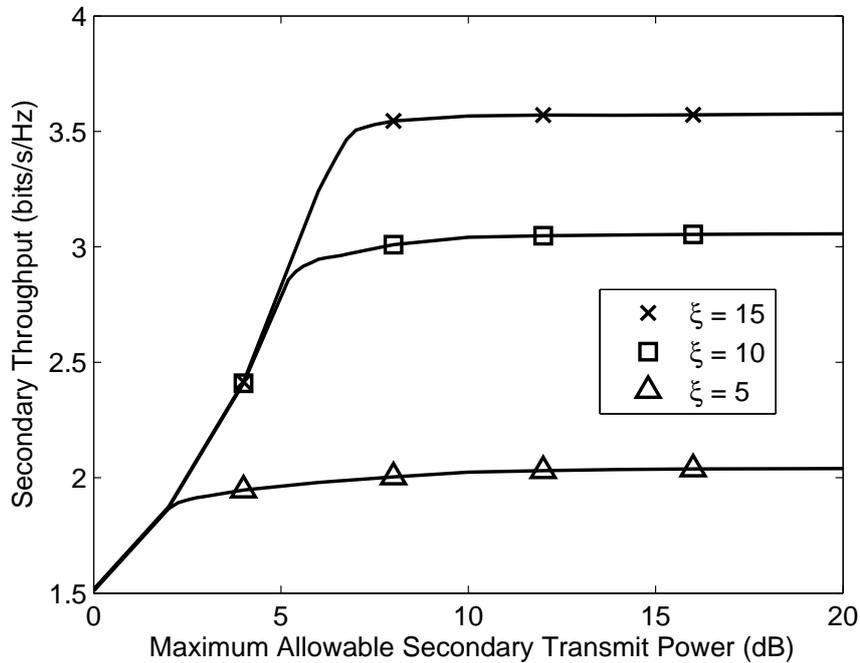


Figure 4.2: Secondary throughput subject to interference outage constraint.

4.5 Summary

In this chapter, we studied cognitive beamforming when the ST-SR channel is known but the ST-PR channel is unknown at ST. Two interference constraints at PR, namely interference expectation constraint and interference outage constraint are considered. To ensure the interference expectation to be below the interference threshold at PR, we transformed the problem to a new one that can be easily and optimally solved via the water-filling algorithm, and showed that there exists a critical point related to the secondary transmit power and the primary interference threshold. When the secondary transmit power is less than the critical point, the secondary throughput is determined by its transmit power only, otherwise it is determined by the primary interference threshold only. For the interference outage constraint, we proved that the optimal beamforming is parallel to the signal subspace

of the ST-SR channel, and that the interference is only related to the eigenvalues of the secondary covariance matrix. We proposed an algorithm to obtain the local optimal beamforming solution by exploiting the relationship between the secondary covariance matrix and its eigenvalues. Numerical results demonstrated the performance of the proposed cognitive beamforming solutions subject to both constraints.

Chapter 5

Power Minimization: Full CSI

When a cognitive beamforming optimization problem does not impose any constraint on the data streams, e.g., the number and SNR requirements, we can first derive the optimal covariance \mathbf{Q} of ST beamforming matrix, and then decompose \mathbf{Q} to obtain the optimal beamforming solution. The problem with variable \mathbf{Q} is a relaxed problem, and this problem shares the same optimal results with the original one if the original problem is convex. Usually, the relaxation cannot lead to the global optimal solution of ST beamforming matrix for a nonconvex original problem. When the result of the relaxed convex optimization problem cannot satisfied all the nonconvex constraints, a randomization technology is required to generate a feasible solution. The randomized solution will approach to the global optimum when the number of randomization samples are sufficiently large. This is a widely used method for beamforming optimization problems. In this chapter and the next chapter, we will consider the case when each data stream has an individual SNR constraint, and design another method to obtain the global or local optimum of the problems.

The individual SNR constraint of each data stream is useful for multiple data streams transmission. In a practical MIMO communication system, a data stream can only be transmitted with finite transmission mode. Speci-

fications like data rate and error rate of each mode are determined, which means that each mode has its minimum SNR requirement to support required transmission quality. According to the current channel conditions, one may dynamically select transmission modes for all the data streams to maximize throughput. When the transmission modes combination have been determined, the transmitter can minimize transmit power without violating QoS requirements, especially if power consumption is important for the communication system.

Under individual SNR constraints, the downlink transmission where each user has a single data stream is studied in [61]. In [62], the study is extended to multiple data streams. However, there exists interference between any two data streams even if they are for the same user. In [63], the authors studied the transmit power minimization problem with individual SNR requirements and used joint decoding to remove the interference. An iterative algorithm is proposed therein to solve the problem, but it is not clear whether the iterative algorithm can converge to the global/local optimum. The cognitive transmission with multiple antennas equipped at ST and SR is studied in [16], where the secondary throughput is maximized subject to the secondary power constraint and primary interference power constraint. It is shown therein that the secondary transmit beamforming problem can be converted to an optimization problem with unitary constraint, and then an algorithm is proposed to compute the beamforming matrix such that a local optimum can be obtained.

For the case of nonzero interference power constraint, the expressions of the secondary SNR and the interference power received at PU usually result in QCQP problems and these problems may not be directly solved by convex tools, especially when there is a rank constraint. SDP relaxation can be used to convert such a problem to a convex optimization problem by dropping the rank constraint and generate a local optimum [57]. It is shown in [59]

that under certain conditions, a new solution can be generated from the one obtained by SDP relaxation without ruining the constraints or changing the objective function, and hence the solution is optimal. However, in general scenarios, the obtained local optimum may not be feasible for the original problem because usually its rank does not meet the requirement. As a result, approximation approaches such as the randomization procedure are used to generate a feasible solution [58, 60].

In this chapter, we study the power minimization problem of multiple data streams transmission per SR with full cross CSI available at ST. Individual SNR requirement of each secondary data stream is considered so that our work is more practical to implement than existing results on cognitive beamforming with multiple data streams. On the other hand, we provide a pre- and post-coding method to help receiver decode signals without performance loss caused by interference between secondary data streams. With individual SNR requirements, although the power minimization problem becomes nonconvex, we prove that strong duality holds, and thus global optimal solution can be obtained by solving the dual problem.

5.1 System Model and Problem Formulation

5.1.1 System Model

We consider a multi-antenna CR system where an ST-SR pair and a PT-PR pair coexist and share the same spectrum in an underlay approach. Multiple data streams will be supported in the secondary system. In particular, both the interference caused by secondary transmission experienced at PR and the interference caused by primary transmission experienced at SR are considered in this chapter. Like Chapters 3 and 4, we consider the scenario where multiple antennas are equipped at the PT, PR, ST and SR. Let $\mathbf{G}_x \in \mathbb{C}^{n \times p}$ and $\mathbf{H}_x \in \mathbb{C}^{q \times m}$ denote the channel matrices of the PT-SR link and ST-

PR link, respectively. It is assumed that the ST-SR link $\mathbf{H} \in \mathbb{C}^{n \times m}$ and ST-PR link $\mathbf{H}_x \in \mathbb{C}^{q \times m}$ are known at ST, and the PT-SR link $\mathbf{G}_x \in \mathbb{C}^{n \times p}$ is known at SR. Under this assumption, subject to its own SNR constraint at each data stream, ST is able to adjust its beamforming matrix based on the channel knowledge so as to optimally balance between minimizing its own transmit power and avoiding interferences at PR. In practice, the channel from the ST (PT) to PR (SR) can be estimated at ST (SR) by, e.g., periodically sensing the pilot signal from PR (PT) provided that time-division-duplexing (TDD) is employed by the primary transmission. In a fading environment, there are cases where it is difficult for ST/SR to perfectly estimate instantaneous channels. In such cases, the results obtained in this chapter provide a performance upper-bound for the considered secondary transmit beamforming problem in a CR network.

We assume frequency flat block fading, where the channel matrices remain constant during one transmission block and vary independently from one block to another. Letting $\mathbf{x}_p \in \mathbb{C}^{p \times 1}$ denote the transmitted primary signal with zero mean and variance P_p , the received signal at SR can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x}_s + \mathbf{G}_x\mathbf{x}_p + \mathbf{n}_s \quad (5.1)$$

The second term on the right-hand side of (5.1) represents the interference from the primary transmission. Therefore, the interference-plus-noise covariance matrix at SR is given by

$$\mathbf{W} = \text{E}[(\mathbf{G}_x\mathbf{x}_p + \mathbf{n}_s)(\mathbf{G}_x\mathbf{x}_p + \mathbf{n}_s)^H] = P_p\mathbf{G}_x\mathbf{G}_x^H + \mathbf{I}_n.$$

The transmitted secondary signal can be represented as

$$\mathbf{x}_s = \mathbf{T}\mathbf{s}_s$$

where $\mathbf{s}_s \in \mathbb{C}^{d \times 1}$ denotes the secondary data, modeled as a random vector with $d \leq \min(m, n)$ denoting the number of secondary data streams and $\text{E}[\mathbf{s}_s^H\mathbf{s}_s] = \mathbf{I}_d$.

It can be easily shown that the eigenvalues of $\mathbf{T}^H \mathbf{H}^H \mathbf{W}^{-1} \mathbf{H} \mathbf{T}$ represent the SNR values of secondary data streams at SR, after proper receive beamforming that maximizes the SNR of each data stream, as shown in Appendix B. Letting $\mathbf{M} \triangleq \mathbf{H}^H \mathbf{W}^{-1} \mathbf{H}$, we can use EVD to decompose $\mathbf{T}^H \mathbf{M} \mathbf{T}$ as $\mathbf{T}^H \mathbf{M} \mathbf{T} = \mathbf{U}^H \mathbf{\Sigma} \mathbf{U}$, where the diagonal entries of $\mathbf{\Sigma}$ now represent the SNR of each secondary data stream. We can always choose a unitary matrix and post-multiply it to \mathbf{T} to get a new \mathbf{T} such that $\mathbf{T}^H \mathbf{M} \mathbf{T}$ is diagonal with the same eigenvalues, i.e.

$$\mathbf{T}^H \mathbf{M} \mathbf{T} = \mathbf{\Sigma}. \quad (5.2)$$

In order to protect the primary communication, the interference power experienced at PR should not exceed a certain threshold. The peak interference power constraint can then be written as

$$\text{tr}(\mathbf{T}^H \mathbf{H}_x^H \mathbf{H}_x \mathbf{T}) \leq \xi.$$

As ξ increases, ST has higher flexibility to design the transmit beamforming matrix. If ξ is sufficiently large, ST can communicate to SR as if PR is absent. For a certain ξ , it is possible that the underlying channel conditions fail to support the secondary quality of service (QoS) requirement with a certain d . In this case, ST may have to reduce d or relax its QoS requirement to be able to transmit. Given the secondary QoS requirement and primary interference power constraint, we can test if there is a feasible secondary beamforming solution.

Clearly, the number of secondary data streams d should not be greater than $\min(m, n)$. If the number of ST antennas, m , is strictly larger than the number of PR antennas, q , then there are $m - q$ available degrees of freedom or spatial dimensions for secondary transmission without causing any interference at PR, which can be realized by placing \mathbf{T} in the null space of \mathbf{H}_x . On the other hand, if the PR can tolerate a nonzero interference (i.e., $\xi > 0$), the number of supported secondary data streams can be greater than $m - q$,

depending on the value of ξ and the underlying channel condition. Therefore, the secondary system can support at least $m - q$ secondary data streams. In this chapter, we consider individual SNR requirements for all the secondary data streams. Let ρ_i denote the i -th data stream's SNR requirement, where $i = 1, \dots, d$. Without loss of generality, we assume $\rho_1 \geq \dots \geq \rho_d$.

5.1.2 Problem Formulation

Our objective is to minimize the secondary sum transmit power while satisfying both the secondary SNR constraint and the primary interference power constraint. Such problem is formulated as

$$(P5.1) \quad \min_{\mathbf{T}} \text{tr}(\mathbf{T}^H \mathbf{T}) \quad (5.3a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{T}^H \mathbf{H}_x^H \mathbf{H}_x \mathbf{T}) \leq \xi \quad (5.3b)$$

$$\mathbf{\Sigma} \succeq \text{diag}(\rho_1, \rho_2, \dots, \rho_d). \quad (5.3c)$$

Lemma 5.1. Any \mathbf{T} satisfying (5.2) can be expressed as

$$\mathbf{T} = \mathbf{M}^{-1/2} \mathbf{V} \mathbf{\Sigma}^{1/2} \quad (5.4)$$

where $\mathbf{V} \in St(m, d)$.

Proof. Please refer to Appendix C. □

Lemma 5.2. The inequality constraint in (5.3c) can be replaced with its equality constraint, i.e., $\mathbf{\Sigma} = \text{diag}(\rho_1, \dots, \rho_d)$.

Proof. With (5.4), the objective function in (5.3a) and the constraint in (5.3b) can be respectively rewritten as

$$\begin{aligned} \text{tr}(\mathbf{T}^H \mathbf{T}) &= \text{tr}((\mathbf{\Sigma}^{1/2})^H \mathbf{V}^H (\mathbf{M}^{-1/2})^H \mathbf{M}^{-1/2} \mathbf{V} \mathbf{\Sigma}^{1/2}) \\ &= \text{tr}(\mathbf{\Sigma} \mathbf{V}^H \mathbf{M}^{-1} \mathbf{V}) \end{aligned} \quad (5.5)$$

and

$$\text{tr}(\mathbf{T}^H \mathbf{H}_x^H \mathbf{H}_x \mathbf{T}) = \text{tr}(\boldsymbol{\Sigma} \mathbf{V}^H \mathbf{M}_x \mathbf{V}) \leq \xi \quad (5.6)$$

where $\mathbf{M}_x \triangleq \mathbf{M}^{-1/2} \mathbf{H}_x^H \mathbf{H}_x \mathbf{M}^{-1/2}$.

Since it always holds that

$$\begin{aligned} \text{tr}(\boldsymbol{\Sigma} \mathbf{V}^H \mathbf{M}^{-1} \mathbf{V}) &\geq \text{tr}(\text{diag}(\rho_1, \dots, \rho_d) \mathbf{V}^H \mathbf{M}^{-1} \mathbf{V}) \\ \text{tr}(\boldsymbol{\Sigma}' \mathbf{V}^H \mathbf{M}_x \mathbf{V}) &\geq (\text{tr}(\text{diag}(\rho_1, \dots, \rho_d) \mathbf{V}^H \mathbf{M}_x \mathbf{V})) \end{aligned}$$

we can set $\boldsymbol{\Sigma} = \text{diag}(\rho_1, \dots, \rho_d)$ without affecting the optimal solution of (P5.1). \square

With (5.5), (5.6), and Lemma 5.2, the original problem (P5.1) can be reformulated as

$$(P5.2) \quad \min_{\mathbf{V} \in St(m,d)} \text{tr}(\boldsymbol{\Sigma} \mathbf{V}^H \mathbf{M}^{-1} \mathbf{V}) \quad (5.7a)$$

$$\text{s.t. } \text{tr}(\boldsymbol{\Sigma} \mathbf{V}^H \mathbf{M}_x \mathbf{V}) \leq \xi \quad (5.7b)$$

5.1.3 Access Feasibility

Before solving (P5.2), we need to perform a feasibility test. Note that by expressing \mathbf{T} in the form of (5.4), we have already met the secondary SNR requirement. However, for a given primary interference power constraint (ξ) and underlying channel condition (\mathbf{M}_x), (5.7b) is not always satisfied. If no $\mathbf{V} \in St(m, d)$ satisfies the interference constraint (5.7b), we say that the secondary transmission is not *feasible*.

We define

$$\xi_0 \triangleq \min_{\mathbf{V} \in St(m,d)} \text{tr}(\boldsymbol{\Sigma} \mathbf{V}^H \mathbf{M}_x \mathbf{V}). \quad (5.8)$$

If the actual interference power constraint $\xi \geq \xi_0$, we can find a feasible \mathbf{V} to generate the secondary transmit beamforming matrix to satisfy both

the secondary SNR and primary interference power constraint, otherwise ST should keep silent.

In order to find ξ_0 , we provide the following lemma first.

Lemma 5.3. Given a diagonal matrix $\mathbf{\Delta} = \text{diag}(\delta_1, \dots, \delta_u) \in \mathbb{C}^{u \times u}$ ($\delta_1 \geq \dots \geq \delta_u$) and a Hermitian matrix $\mathbf{\Omega} \in \mathbb{C}^{v \times v}$ ($v \geq u$) with $\omega_i, i = 1, \dots, v$, being its eigenvalues ($\omega_1 \geq \dots \geq \omega_v$) and $\boldsymbol{\varphi}_i$ being its eigenvector corresponding to ω_i , for a matrix $\mathbf{\Theta} \in St(v, u)$, we have the following inequality:

$$\text{tr}(\mathbf{\Delta} \mathbf{\Theta}^H \mathbf{\Omega} \mathbf{\Theta}) \geq \sum_{i=1}^u \delta_i \omega_{v-i+1} \quad (5.9)$$

where the equality holds if $\mathbf{\Theta}$ is constructed as $\mathbf{\Theta} = [\boldsymbol{\varphi}_v, \dots, \boldsymbol{\varphi}_{v-u+1}]$.

Proof. Please see Appendix D. □

With Lemma 5.3, we can show that ξ_0 is given by

$$\xi_0 = \sum_{i=1}^d \rho_i x_{m-i+1} \quad (5.10)$$

where x_i denotes the i -th eigenvalue of \mathbf{M}_x and $x_1 \geq \dots \geq x_m$. Since $\text{rank}(\mathbf{M}_x) = q$, we have

$$\begin{cases} x_i > 0, & i \leq q \\ x_i = 0, & i > q. \end{cases} \quad (5.11)$$

It is clear from (5.10) and (5.11) that when $m-q \geq d$, $\xi_0 = 0$ and therefore secondary access (via ZFB or NFB) is always possible; when $m-q < d$, ξ_0 will always be greater than zero and therefore only NFB is possible.

In Fig. 5.1, we illustrate the secondary access probability, i.e., the probability of $\xi_0 \leq \xi$, with different number of secondary data streams, where identical SNR requirements are considered for multiple-stream cases. It is assumed that all the channels involved are i.i.d. block fading Rayleigh channels with channel variance of one. With $m = n = 5$ and $p = q = 2$, Fig. 5.1

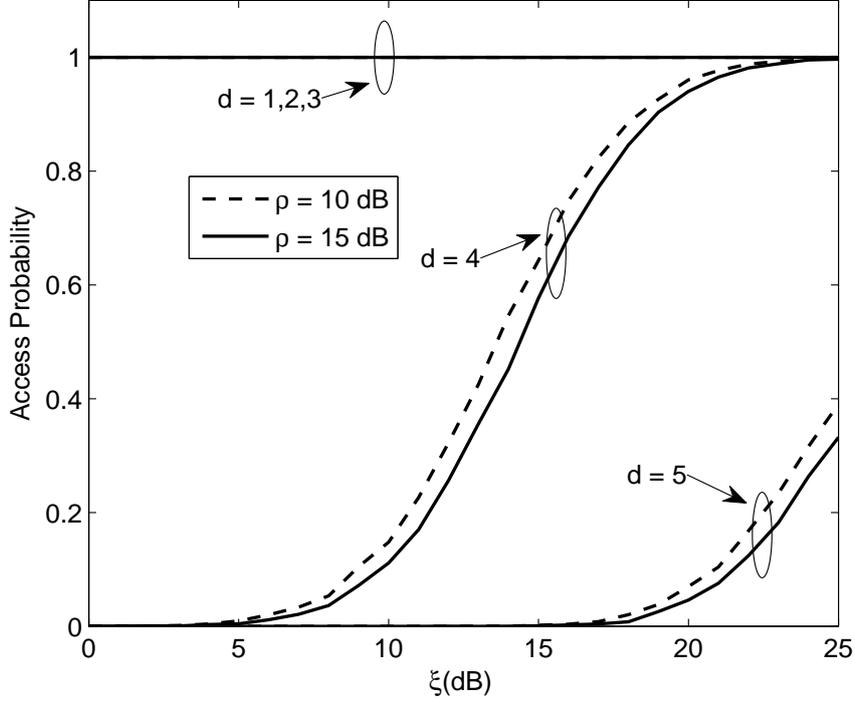


Figure 5.1: Secondary access probability, i.e., the probability of $\xi_0 \leq \xi$. $m = n = 5$, $p = q = 2$ and $\rho = \rho_1 = \dots = \rho_d$.

shows that the cases of $d = 1, 2$ and 3 lead to 100% access probability since $m - q \geq d$ in these cases. On the other hand, when $d > 3$, the access probability, increasing with a lower secondary SNR requirement, heavily depends on the number of secondary streams and the primary interference constraint ξ . It is obvious that in this case more data streams or lower primary interference constraint results in a lower secondary access probability.

5.2 The Case of Single Data Stream

In this section, we show that SDP relaxation can be used to find the optimal secondary beamforming solution for the single data stream case.

By defining $\mathbf{X} \triangleq \mathbf{T}\mathbf{T}^H$ and dropping the rank constraint ($\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{T}) = d$), (P5.1) can be reformulated as a relaxed problem

$$(P5.3) \quad \min_{\mathbf{X} \succeq \mathbf{0}} \text{tr}(\mathbf{X}) \quad (5.12a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{H}_x^H \mathbf{H}_x \mathbf{X}) \leq \xi \quad (5.12b)$$

$$\det(\mathbf{X}\mathbf{M} - \rho_i \mathbf{I}_n) = 0, \forall i = 1, \dots, d \quad (5.12c)$$

The secondary SNR requirement is reflected in (5.12c), since $\mathbf{X}\mathbf{M}$ shares the same eigenvalues with $\mathbf{T}^H \mathbf{M} \mathbf{T}$ and these eigenvalues denote the required secondary SNR values.

When there is a single data stream, the constraint (5.12c) can be equivalently rewritten as $\text{tr}(\mathbf{X}\mathbf{M}) = \rho_1$. It turns out that all the constraints together with the objective function are convex and thus we can apply any convex optimization algorithm to solve the problem. Let \mathbf{X}^* denote the solution for (P5.3). Since the problem is relaxed, \mathbf{X}^* leads to the optimal transmit beamforming vector if $\text{rank}(\mathbf{X}^*) = 1$ or another rank-one Hermitian matrix can be generated from \mathbf{X}^* with all the optimization objective function and constraints unchanged [58]. Otherwise there will be a nonzero gap between the solutions of the relaxed problem (P5.3) and the original problem (P5.1).

Theorem 5.1. When $d = 1$, there exists a rank-one solution of the relaxed problem that is optimal.

Proof. The dual of the relaxed single data stream problem is given by

$$\min_{\mu_1, \mu_2 \in \mathbb{R}} \mu_1 \xi + \mu_2 \rho_1 \quad (5.13a)$$

$$\text{s.t.} \quad (\mathbf{I} + \mu_1 \mathbf{H}_x^H \mathbf{H}_x + \mu_2 \mathbf{M}) \succeq \mathbf{0} \quad (5.13b)$$

$$\mu_1 \geq 0 \quad (5.13c)$$

where μ_1 and μ_2 are the Lagrange multipliers associated with (5.12b) and $\text{tr}(\mathbf{X}\mathbf{M}) = \rho_1$, respectively. Since the relaxed problem is convex, strong

duality holds, i.e., the original and dual problems lead to the same solution. As a result, the following complementary conditions are established

$$\text{tr}(\mathbf{X}^*(\mathbf{I} + \mu_1^* \mathbf{H}_x^H \mathbf{H}_x + \mu_2^* \mathbf{M})) = 0 \quad (5.14a)$$

$$\mu_1^*(\text{tr}(\mathbf{X}^* \mathbf{H}_x^H \mathbf{H}_x) - \xi) = 0 \quad (5.14b)$$

$$\mu_2^*(\text{tr}(\mathbf{X}^* \mathbf{M}) - \rho_1) = 0 \quad (5.14c)$$

where \mathbf{X}^* and (μ_1^*, μ_2^*) are the primal and dual optimal solutions, respectively [73].

If $\text{rank}(\mathbf{X}^*) = 1$, then \mathbf{X}^* can be decomposed as $\mathbf{X}^* = \mathbf{t}\mathbf{t}^H$, and hence \mathbf{t} is the optimal transmit beamforming vector.

If $\text{rank}(\mathbf{X}^*) = R > 1$, we can find a vector $\mathbf{t} \in \mathbb{C}^{m \times 1}$ such that the following equations

$$\text{tr}(\mathbf{t}\mathbf{t}^H) = \text{tr}(\mathbf{X}^*), \quad (5.15a)$$

$$\text{tr}(\mathbf{t}\mathbf{t}^H \mathbf{H}_x^H \mathbf{H}_x) = \text{tr}(\mathbf{X}^* \mathbf{H}_x^H \mathbf{H}_x) \quad (5.15b)$$

$$\text{tr}(\mathbf{t}\mathbf{t}^H \mathbf{M}) = \text{tr}(\mathbf{X}^* \mathbf{M}), \quad (5.15c)$$

hold simultaneously [59, 89].

It can be directly concluded from (5.15b) and (5.15c) that

$$\mu_1^*(\text{tr}(\mathbf{t}\mathbf{t}^H \mathbf{H}_x^H \mathbf{H}_x) - \xi) = 0,$$

$$\mu_2^*(\text{tr}(\mathbf{t}\mathbf{t}^H \mathbf{M}) - \rho_1) = 0.$$

Because

$$\begin{aligned} & \text{tr}(\mathbf{t}\mathbf{t}^H (\mathbf{I} + \mu_1^* \mathbf{H}_x^H \mathbf{H}_x + \mu_2^* \mathbf{M})) \\ &= \text{tr}(\mathbf{t}\mathbf{t}^H) + \mu_1^* \text{tr}(\mathbf{t}\mathbf{t}^H \mathbf{H}_x^H \mathbf{H}_x) + \mu_2^* \text{tr}(\mathbf{t}\mathbf{t}^H \mathbf{M}) \\ &= \text{tr}(\mathbf{X}^*) + \mu_1^* \text{tr}(\mathbf{X}^* \mathbf{H}_x^H \mathbf{H}_x) + \mu_2^* \text{tr}(\mathbf{X}^* \mathbf{M}) \\ &= \text{tr}(\mathbf{X}^* (\mathbf{I} + \mu_1^* \mathbf{H}_x^H \mathbf{H}_x + \mu_2^* \mathbf{M})) = 0, \end{aligned}$$

matrix $\mathbf{t}\mathbf{t}^H$ satisfies condition (5.14a). Therefore, $\mathbf{t}\mathbf{t}^H$ satisfies all the complementary conditions of (5.14), and it is the optimal rank-one solution of

(P5.3), and \mathbf{t} is the optimal secondary transmit beamforming vector resulting in zero duality gap. \square

For the multiple data streams case, since constraint (5.12c) is nonconvex, to the best of our knowledge, no applicable reformulation or relaxation on this constraint set can be found. As a result, the SDP relaxation might not be feasible for the considered multiple data streams case. In the next section, we reformulate the cognitive beamforming problem for multiple secondary data streams to a new problem on the Stiefel manifold and solve it effectively.

5.3 The Case of Multiple Data Streams

5.3.1 Zero Forcing Beamforming

In the ZFB scenario, no interference is allowed at PR. According to (5.10) and (5.11), this scenario is possible only when $m - q \geq d$. Therefore, the primary interference constraint (5.7b) is rewritten as

$$\mathbf{H}_x \mathbf{T} = \mathbf{0}. \quad (5.16)$$

The ZFB constraint (5.16) requires that the transmit beamforming matrix \mathbf{T} should be projected to the null space of \mathbf{H}_x . By substituting (5.4) and the singular value decomposition (SVD) of $\mathbf{H}_x = \mathbf{U}_1 \mathbf{A}_1 \mathbf{V}_1$ into (5.16), where \mathbf{U}_1 is a $q \times q$ unitary matrix, \mathbf{A}_1 is a positive definite diagonal matrix, and $\mathbf{V}_1^H \in St(m, q)$, we have

$$\mathbf{H}_x \mathbf{T} = \mathbf{U}_1 \mathbf{A}_1 \mathbf{V}_1 \mathbf{M}^{-1/2} \mathbf{V} \mathbf{\Sigma}^{1/2} = \mathbf{0}. \quad (5.17)$$

It is clear in (5.17) that matrices \mathbf{U}_1 and \mathbf{A}_1 are all full-rank square matrices. We remove them by left multiplying the corresponding inverse matrices and drop the SNR requirement matrix $\mathbf{\Sigma}$. It thus follows that

$$\mathbf{V}_1 \mathbf{M}^{-1/2} \mathbf{V} = \mathbf{0}. \quad (5.18)$$

Now we denote the SVD of $\mathbf{V}_1\mathbf{M}^{-1/2}$ as $\mathbf{V}_1\mathbf{M}^{-1/2} = \mathbf{U}_2\mathbf{A}_2\mathbf{V}_2$, where \mathbf{U}_2 is a $q \times q$ unitary matrix, \mathbf{A}_2 is a positive definite diagonal matrix, and $\mathbf{V}_2^H \in St(m, q)$. Similarly, (5.18) establishes when

$$\mathbf{V}_2\mathbf{V} = \mathbf{0}. \quad (5.19)$$

The m -dimensional space where \mathbf{V}_2 and \mathbf{V} lie in can be separated into two subspaces via the projected-channel SVD [15]: one is perpendicular to \mathbf{V}_2 (by multiplying $(\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2)$) and the other is parallel to \mathbf{V}_2 (by multiplying $\mathbf{V}_2^H\mathbf{V}_2$). Because $\text{rank}(\mathbf{V}_2^H\mathbf{V}_2) = \text{rank}(\mathbf{V}_2) = q$ and $\mathbf{V}_2(\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2) = \mathbf{0}$, we have $\text{rank}(\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2) = m - q$. By applying the subspace separation on \mathbf{V} , i.e.,

$$\begin{aligned} \mathbf{V} &= (\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2 + \mathbf{V}_2^H\mathbf{V}_2)\mathbf{V} \\ &= (\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2)\mathbf{V} \end{aligned} \quad (5.20)$$

we can rewrite the ZFB problem as

$$(P5.4) \quad \min_{\mathbf{V} \in St(m, d)} \text{tr}(\mathbf{\Sigma}\mathbf{V}^H\mathbf{M}^{-1}\mathbf{V}) \quad (5.21a)$$

$$\text{s.t. } \mathbf{V} = (\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2)\mathbf{V}. \quad (5.21b)$$

The problem is now reduced to finding a Stiefel manifold matrix \mathbf{V} with a signal subspace constraint.

Since \mathbf{V} lies in the null space of \mathbf{V}_2 , we define the portion of \mathbf{M}^{-1} that is also in the null space of \mathbf{V}_2 as

$$\mathbf{R} \triangleq (\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2)^H\mathbf{M}^{-1}(\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2). \quad (5.22)$$

By substituting the constraint in (5.21b) into the objective function in (5.21a), we have

$$\begin{aligned} &\text{tr}(\mathbf{\Sigma}\mathbf{V}^H\mathbf{M}^{-1}\mathbf{V}) \\ &= \text{tr}(\mathbf{\Sigma}\mathbf{V}^H(\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2)^H\mathbf{M}^{-1}(\mathbf{I}_m - \mathbf{V}_2^H\mathbf{V}_2)\mathbf{V}) \\ &= \text{tr}(\mathbf{\Sigma}\mathbf{V}^H\mathbf{R}\mathbf{V}). \end{aligned} \quad (5.23)$$

It can be easily shown that $\text{rank}(\mathbf{R}) = \text{rank}(\mathbf{I}_m - \mathbf{V}_2^H \mathbf{V}_2) = m - q$, and hence \mathbf{R} can be decomposed as

$$\mathbf{R} = \mathbf{V}_R \mathbf{\Lambda}_R \mathbf{V}_R^H \quad (5.24)$$

where $\mathbf{V}_R \in St(m, m - q)$, $\mathbf{\Lambda}_R$ is an $(m - q) \times (m - q)$ diagonal positive definite matrix with entries in non-decreasing order. Since $\mathbf{V}_2 \mathbf{R} = \mathbf{0}$, we have $\mathbf{V}_2 \mathbf{V}_R = \mathbf{0}$. With \mathbf{V}_R , the optimal Stiefel manifold matrix \mathbf{V} is given in the following theorem.

Theorem 5.2. For the ZFB problem (P5.4), the solution is given by

$$\mathbf{V}^* = \mathbf{V}_R \begin{bmatrix} \mathbf{I}_d \\ \mathbf{0} \end{bmatrix}_{(m-q) \times d}.$$

Proof of Theorem 5.2

We first consider the case of $d = m - q$. In this case, \mathbf{V} is an $m \times (m - q)$ matrix on the Stiefel manifold. Since \mathbf{V} and \mathbf{V}_R are of the same size and perpendicular to \mathbf{V}_2 , \mathbf{V} and \mathbf{V}_R are in the same signal subspace and there exists a unitary matrix \mathbf{Q} that satisfies

$$\mathbf{V} = \mathbf{V}_R \mathbf{Q}. \quad (5.25)$$

By substituting (5.25) into (5.4), the transmit beamforming matrix is given by

$$\mathbf{T} = \mathbf{M}^{-1/2} \mathbf{V}_R \mathbf{Q} \mathbf{\Sigma}^{1/2}. \quad (5.26)$$

Therefore finding the optimal \mathbf{T} is equivalent to finding the optimal \mathbf{Q} . By substituting (5.23) and (5.25) into (P5.4), the problem can be reformulated as

$$\min_{\mathbf{Q} \in St(m-q, m-q)} \text{tr}(\mathbf{\Sigma} \mathbf{Q}^H \mathbf{\Lambda}_R \mathbf{Q}). \quad (5.27)$$

By applying Lemma 5.3 and considering the structure of $\mathbf{\Lambda}_R$, the optimal \mathbf{Q} can be easily constructed as $\mathbf{Q}^* = \mathbf{I}_{m-q}$.

If $d < m - q$, we can treat this case as if there are still $m - q$ secondary data streams, among which $m - q - d$ data streams have zero-valued SNR requirements. In other words, the SNR constraint matrix is still an $(m - q) \times (m - q)$ diagonal matrix, but given by $\text{diag}(\rho_1, \dots, \rho_d, 0, \dots, 0)$. The corresponding \mathbf{V} is still an $m \times (m - q)$ matrix on the Stiefel manifold. As a result, the result for the case of $d = m - q$ can be directly applied and the optimal transmit beamforming matrix can be shown to be

$$\begin{aligned} \mathbf{T}^* &= \mathbf{M}^{-1/2} \mathbf{V}_R \mathbf{I}_{m-q} \text{diag}(\rho_1^{1/2}, \dots, \rho_d^{1/2}, 0, \dots, 0) \\ &= \mathbf{M}^{-1/2} \mathbf{V}_R \begin{bmatrix} \mathbf{I}_d \\ \mathbf{0} \end{bmatrix}_{(m-q) \times d} \Sigma^{1/2}. \end{aligned} \quad (5.28)$$

This completes the proof for Theorem 5.2.

Remarks: In (5.28), $\Sigma^{1/2}$ is used to allocate power to the secondary data streams, \mathbf{V}_R projects the data streams to suitable signal subspaces in order to avoid interference at PR, $\mathbf{Q}^* = [\mathbf{I}_d \mathbf{0}]^T$ is used to select the optimal dimension among the possible ones, and $\mathbf{M}^{-1/2}$ is used to handle the interference together with noise at SR.

Other Choices of \mathbf{Q}^*

When the number of distinct SNR requirements is less than d , there exist multiple data streams that have identical SNR requirement, and therefore there will be multiple possible forms of \mathbf{Q}^* as well as \mathbf{V}^* . Suppose Σ has K distinct SNR values with

$$\Sigma = \text{diag}(\underbrace{\rho'_1, \dots, \rho'_1}_{n_1}, \underbrace{\rho'_2, \dots, \rho'_2}_{n_2}, \dots, \underbrace{\rho'_K, \dots, \rho'_K}_{n_K})$$

where $\sum_{k=1}^K n_k = d$ and $\rho'_1 > \rho'_2 \cdots > \rho'_{n_K}$. It turns out that the optimal \mathbf{Q} is a block diagonal matrix in the form of

$$\mathbf{Q}^* = \begin{bmatrix} \text{diag}(\mathbf{Q}_1^*, \dots, \mathbf{Q}_K^*) \\ \mathbf{0} \end{bmatrix} \in St(m - q, d) \quad (5.29)$$

where \mathbf{Q}_k^* , $k = 1, \dots, K$, can be any $n_k \times n_k$ unitary matrix. For example, if $d = m - q = 4$ and $\mathbf{\Sigma} = \text{diag}(a, a, b, b)$ ($a > b > 0$), the optimal \mathbf{Q}^* is given by

$$\mathbf{Q}^* = \begin{bmatrix} \mathbf{Q}_1^* & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^* \end{bmatrix}$$

where \mathbf{Q}_1^* and \mathbf{Q}_2^* can be any 2×2 unitary matrices. Specifically, if all ρ_i 's are distinct, \mathbf{Q}^* must be a diagonal matrix with entries of 1 or $e^{j\theta}$; if all ρ_i 's are identical, \mathbf{Q}^* can be an arbitrary unitary matrix.

5.3.2 Nonzero Forcing Beamforming

In the following, we focus on the NFB problem (P5.2) with $\xi > 0$ and solve it by examining its Lagrangian dual. Although problem (P5.2) is nonconvex, we provide a sufficient condition for the strong duality to hold, and therefore general convex optimization algorithms can be applied to solve the problem.

Dual Problem

As compared to ZFB, NFB relaxes the interference constraint at the PR by allowing a nonzero ξ , therefore it is possible for the secondary system to use fewer secondary transmit antennas or transmit more secondary data streams. Moreover, it provides more opportunities for the secondary system to access the channel.

The Lagrangian of (P5.2) is defined as

$$\begin{aligned} L(\mathbf{V}, y) &\triangleq \text{tr}(\mathbf{\Sigma}\mathbf{V}^H\mathbf{M}^{-1}\mathbf{V}) + y(\text{tr}(\mathbf{\Sigma}\mathbf{V}^H\mathbf{M}_x\mathbf{V}) - \xi) \\ &= \text{tr}(\mathbf{\Sigma}\mathbf{V}^H(\mathbf{M}^{-1} + y\mathbf{M}_x)\mathbf{V}) - y\xi \end{aligned} \quad (5.30)$$

where $y \geq 0$ is the Lagrange multiplier. Define the dual objective $g(\lambda)$ as an unconstrained minimization of the Lagrangian

$$g(y) = \min_{\mathbf{V} \in \text{St}(m,d)} L(\mathbf{V}, y). \quad (5.31)$$

Let λ_i , $i = 1, \dots, m$, denote the i -th eigenvalue of $\mathbf{M}^{-1} + y\mathbf{M}_x$ with $\lambda_1 \leq \dots \leq \lambda_m$. Given a y , we can derive from Lemma 5.3 that $L(\mathbf{V}, y)$ is minimized when \mathbf{V} is constructed as

$$\mathbf{V}_y = [\mathbf{v}_1, \dots, \mathbf{v}_d] \quad (5.32)$$

where \mathbf{v}_i , $i = 1, \dots, d$, denotes the eigenvector of $\mathbf{M}^{-1} + y\mathbf{M}_x$ corresponding to λ_i . Therefore, the matrix \mathbf{V}_y , as well as the corresponding sum transmit power $\text{tr}(\Sigma \mathbf{V}_y^H \mathbf{M}^{-1} \mathbf{V}_y)$ and interference power $\text{tr}(\Sigma \mathbf{V}_y^H \mathbf{M}_x \mathbf{V}_y)$, are functions of y . The Lagrange dual problem associated with (P5.2) is given by

$$(P5.5) \quad \max g(y) \quad (5.33a)$$

$$\text{s.t. } y \geq 0. \quad (5.33b)$$

Note that the original problem (P5.2) is nonconvex because the domain of \mathbf{V} is nonconvex. Therefore, we cannot directly tell if the strong duality holds or not. The following theorem provides a sufficient condition for strong duality between the primal problem (P5.2) and the dual problem (P5.5).

Theorem 5.3. The strong duality between (P5.2) and (P5.5) holds if the feasibility test is passed, i.e., $\xi \geq \xi_0$.

Remarks: The feasible domain is nonempty if and only if $\xi \geq \xi_0$. As a result, as long as the ST can access the channel, we can always solve the dual problem to obtain the optimal beamforming solution, which is of zero duality gap.

Proof of Theorem 5.3

To the best of our knowledge, although there are a number of constraint qualifications under which strong duality holds even if the primal problem is nonconvex [73, 90, 91], none of them can be directly applied to decide if the strong duality holds in our case.

For an arbitrary y , the difference between $g(y)$ and the corresponding primal objective function value $\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_y^H\mathbf{M}^{-1}\mathbf{V}_y)$ is

$$G(y) = y(\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_y^H\mathbf{M}_x\mathbf{V}_y) - \xi). \quad (5.34)$$

If there exists a y such that $G(y) = 0$, then the result of the dual problem (P5.5) becomes $\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_y^H\mathbf{M}^{-1}\mathbf{V}_y)$, which is exactly equal to the corresponding primal objective function value. In this case, y and \mathbf{V}_y are the optimal solution of the dual problem (P5.5) and the primal problem (P5.2), respectively, and zero duality gap holds. Otherwise the optimal primal solution \mathbf{V} is not related to any y , which means that the zero duality gap does not hold and we cannot obtain the optimal \mathbf{V} by solving the dual problem (P5.5).

It is known that the dual function $g(y)$ yields a lower bound of (P5.2) for all $y \geq 0$, and usually it is easier to solve the dual problem (P5.5). If there exists a y such that $G(y) = 0$, then the result of (P5.5) becomes $\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_y^H\mathbf{M}^{-1}\mathbf{V}_y)$, which is exactly equal to the corresponding primal objective function value, and it serves as the infimum of (P5.2). Since \mathbf{V}_y is determined by y , it is a subset of \mathbf{V} . If no y and the corresponding \mathbf{V}_y satisfy $G(y) = 0$, we cannot obtain the optimal \mathbf{V} by solving the dual problem (P5.5). As a result, $G(y)$ is a loose lower bound of (P5.2) for all $y \geq 0$, and the strong duality does not hold.

We will need the following lemma to prove Theorem 5.3.

Lemma 5.4. The transmit power $\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_y^H\mathbf{M}^{-1}\mathbf{V}_y)$ is a monotonically increasing function of y and the interference power $\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_y^H\mathbf{M}_x\mathbf{V}_y)$ is a monotonically decreasing function of y .

Proof. Please see Appendix E. □

Lemma 5.4 suggests that the minimum interference power is achieved at $y = \infty$. If the secondary transmission is feasible, i.e., $\xi \geq \xi_0$, it is easy to show that either of the following two conditions is satisfied:

$$\begin{cases} \text{tr}(\boldsymbol{\Sigma} \mathbf{V}_y^H \mathbf{M}_x \mathbf{V}_y) = \xi, \exists y > 0 \\ \text{tr}(\boldsymbol{\Sigma} \mathbf{V}_y^H \mathbf{M}_x \mathbf{V}_y) \leq \xi, y = 0. \end{cases}$$

Therefore, $y(\text{tr}(\boldsymbol{\Sigma} \mathbf{V}_y^H \mathbf{M}_x \mathbf{V}_y) - \xi) = 0$ is always satisfied, i.e., the strong duality holds. That is to say, passing the feasibility test is the sufficient condition for zero duality gap in our problem. Furthermore, the \mathbf{V}_y corresponding to the minimum feasible y of the dual problem is the optimal solution that minimizes the sum transmit power. General convex optimization algorithms can be applied to find the optimal y and consequently the optimal transmit beamforming matrix.

5.4 Simulation Results

Numerical results are provided in this section to illustrate the performance of the proposed beamforming solutions under various channel conditions and system requirements. We assume that each entry of channel matrices \mathbf{G}_x , \mathbf{H}_x , and \mathbf{H} is an i.i.d. random variable, distributed as $\mathcal{CN}(0, 1)$. The primary transmit power P_p is assumed to be 0 dB, unless otherwise specified. Throughout this section, we set $p = q = 2$ and $m = n = 5$, and d is selected to satisfy $m - q \geq d$ such that secondary access is possible. In the case of multiple secondary data streams (i.e., $d > 1$), equal SNR requirements are considered in all the figures except Fig. 8.

Firstly, we consider the case of single secondary data stream in Fig. 5.2 and Fig. 5.3. Fig. 5.2 shows the minimum required average secondary transmit power based on the derived optimal beamforming matrix for ZFB. As a

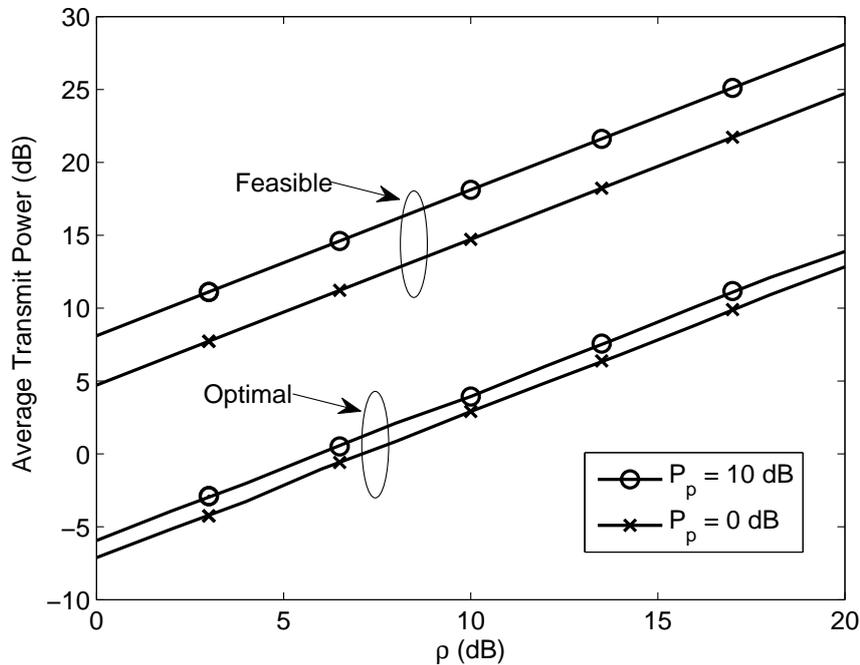


Figure 5.2: Power saving of optimal beamforming over feasible beamforming. $d = 1$.

comparison, the required secondary transmit power averaged over all feasible beamforming matrices is also shown in Fig. 5.2. Here, the feasible beamforming matrix refers to any beamforming matrix that satisfies the interference constraint in (3b) and SNR requirement in (3c). From (5), it is clear that the required secondary transmit power linearly increases with the secondary SNR requirement, as shown in Fig. 5.2. In Fig. 5.2, we observe that the optimal beamforming matrix brings significant power saving over the feasible beamforming matrix and that this power saving increases with the primary transmit power.

In Fig. 5.3 we compare the SDP relaxation approach in Section III and the Stiefel manifold transformation approach in Section IV for both ZFB and NFB cases. One secondary data stream is to be transmitted. Fig. 5.3 verifies

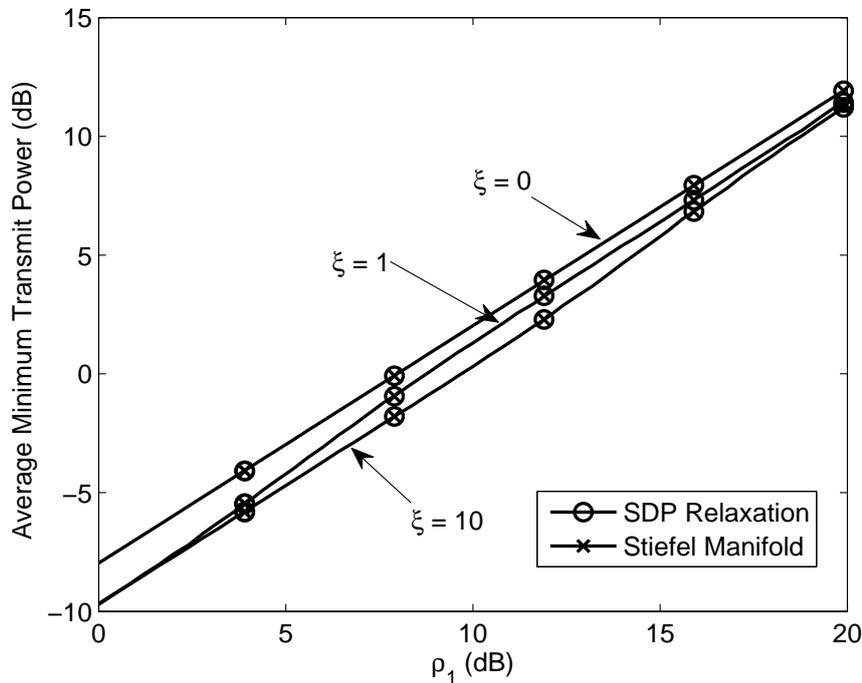


Figure 5.3: The impact of secondary SNR requirement. $d = 1$.

that the SDP relaxation approach and the Stiefel manifold transformation approach achieve the same optimal performance, as expected.

Next, we consider the case of multiple secondary data streams. For simplicity, only the optimal beamforming solutions are presented in the rest of this section. The impact of interference power constraint ξ and secondary SNR requirement ρ on the minimum required secondary transmit power is shown in Figs. 5.4 and 5.5, where $d = 2$ secondary data streams are considered. From Fig. 5.4, it is observed that for a fixed ρ , the required transmit power decreases significantly when ξ is slightly increased over the zero value (the beamforming is changed from ZFB to NFB) while the required power decreases slowly when the positive ξ further increases. This is not unexpected since ZFB significantly restricts the available beamforming dimensions as compared to NFB. In this case, there are only three dimensions available

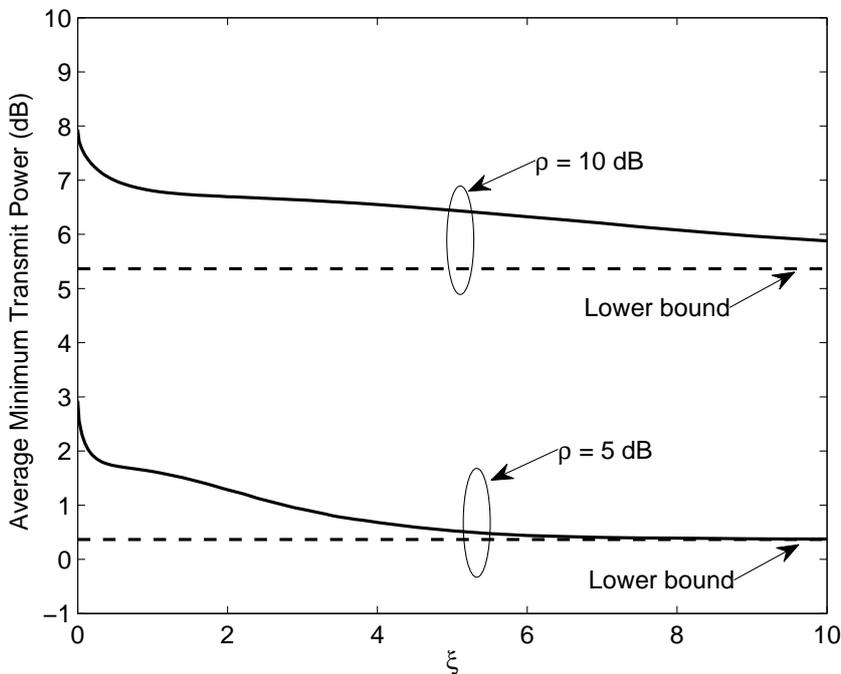


Figure 5.4: The impact of primary interference constraint. $d = 2$ and $\rho_1 = \rho_2 = \rho$.

for ZFB and the other two dimensions are not permitted to use due to the nature of ZFB, while NFB can use all the five dimensions for beamforming. On the other hand, for both ZFB and NFB (i.e., zero and nonzero ξ values), a lower secondary SNR requirement always brings a significant reduction in the required secondary transmit power. The lower bound of the secondary transmit power when $\xi \rightarrow \infty$ (it can be easily obtained by dropping constraint (3b) in (P5.1)) is also plotted in Fig. 5.4, where we observe that ρ is more dominant than ξ in determining the required secondary power. The similar finding can be observed in Fig. 5.5, as well as the relatively small difference among the nonzero ξ values. In particular, we observe that at small ρ values, there is almost no difference between the nonzero ξ values in Fig. 5.5. This is because when the secondary SNR requirement is small,

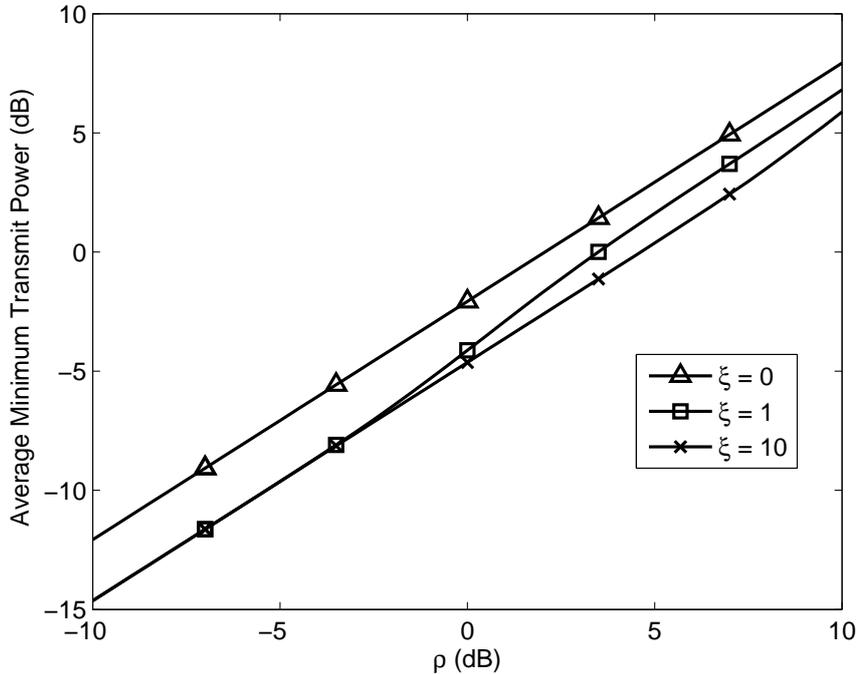


Figure 5.5: The impact of secondary SNR requirement. $d = 2$ and $\rho_1 = \rho_2 = \rho$.

the required secondary power is small and thus the interference caused to the primary receiver is also small and always below the underlying primary interference constraint ξ . When ρ increases, the required secondary transmit power needs to increase and therefore the caused interference may no longer be lower than ξ . For a more strict interference constraint (i.e., a lower ξ), the available secondary beamforming directions are subject to more restrictions and therefore more secondary transmit power is needed to satisfy this primary interference constraint as well as the secondary SNR constraint at the same time.

The minimum required secondary power per data stream is investigated in Fig. 5.6, where the SNR requirement of each secondary data stream is assumed to be equal to each other, denoted as ρ . Together with Fig. 5.3, it is

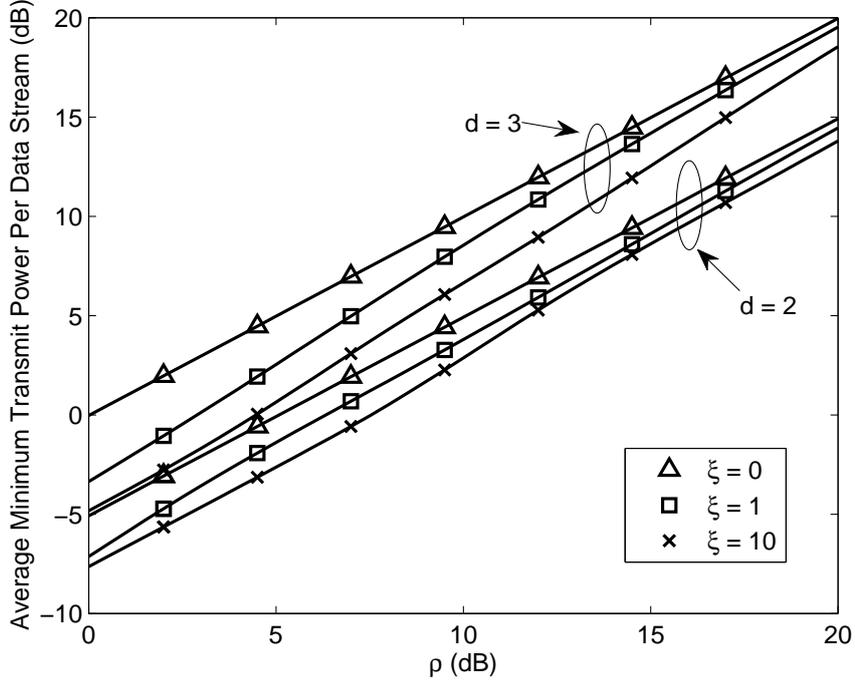


Figure 5.6: Minimum required transmit power per secondary data stream.

seen that when d is increased from 1 to 3, more power needs to be allocated to each data stream on average. When d is small, ST can choose channels with good conditions to transmit data streams and therefore the required power per data stream can be relatively small. As d increases, while satisfying the primary interference power constraint as a higher priority, ST may have to use channels with poor conditions and hence more power is required to transmit data streams with the same secondary SNR requirement per data stream. Based on Figs. 5.3 and 5.6, the total minimum required transmit power with different number of secondary data streams is shown in Fig. 5.7.

Fig. 5.8 illustrates the impact of distinct and identical secondary SNR requirements for both ZFB and NFB (zero and nonzero ξ values) with $d = 2$ data streams. Here, the SNR requirements have been chosen such that the achievable sum-rate is the same in both distinct and identical SNR require-

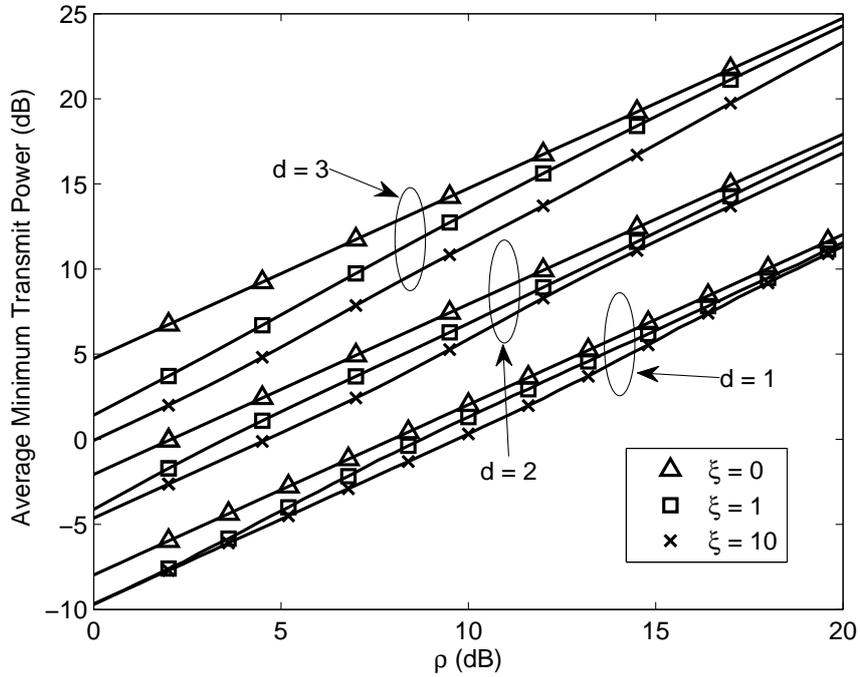


Figure 5.7: Minimum required transmit power with different number of secondary data streams.

ments cases. Fig. 5.8 shows that the case of distinct SNR requirements costs less power than the identical SNR requirements case. The reason is given as follows. MIMO provides parallel channels for the transmission of multiple data streams but the underlying conditions of different channels may not be the same. From the secondary system point of view, with distinct individual SNR requirements, we can place the secondary data stream with a higher SNR requirement to the channel with better condition to save the transmit power. This is why the case of distinct SNR requirements outperforms the case of identical SNR requirements in terms of lower secondary transmit power.

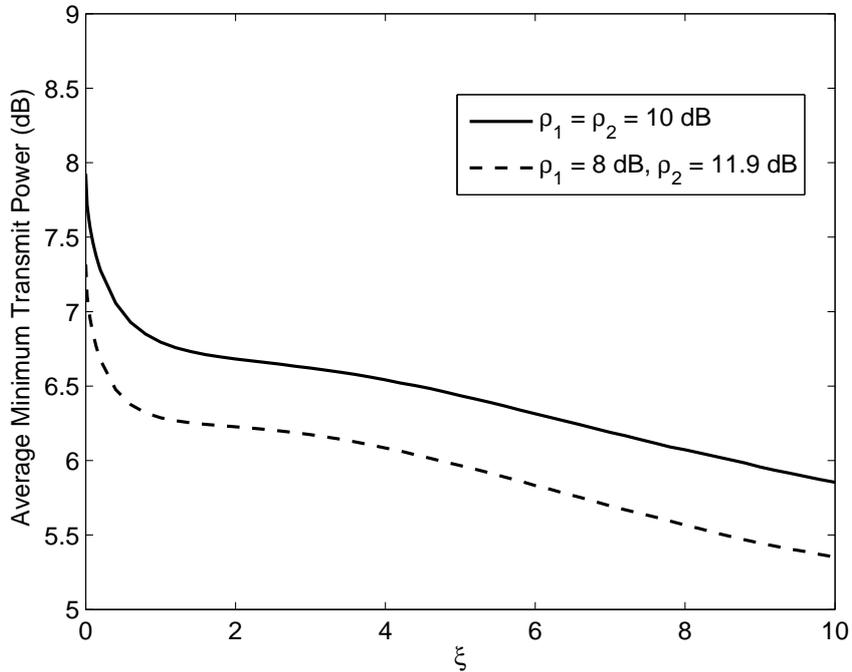


Figure 5.8: Performance of distinct and identical secondary SNR requirements.

5.5 Summary

Secondary transmit beamforming with multiple secondary data streams was considered in this chapter. The optimal beamforming strategy was designed to minimize the secondary sum transmit power, subject to the individual secondary SNR requirements and primary interference power constraint. By exploiting the Stiefel manifold, we derived the closed form solution for zero forcing beamforming, while for nonzero forcing beamforming, we proved that the strong duality holds for the nonconvex primal problem and thus the optimal solution can be easily obtained by solving the dual problem. Numerical results were presented to demonstrate the performance of the proposed cognitive beamforming solutions. In the single data stream case, our beamforming

solution achieves the same result as the SDP relaxation as expected. For both single and multiple secondary data streams, our solution significantly outperforms the non-optimized beamforming solutions.

Chapter 6

Power Minimization: Partial CSI

In this chapter, we will extend the results on secondary power minimization with individual SNR constraints to the partial cross CSI scenario. Because SUs may not always be able to precise current CSI information due to reasons like delay and estimation errors, there will be errors between the available CSI and the current values. As a result, the uncertainties of channel errors must be considered to guarantee the QoS of primary communications. The employed CSI uncertainties will be formulated as an ellipsoidal uncertainty set. Such channel uncertainty models have been applied to solve problems with different kinds of objectives. In [33], throughput maximization problems have been discussed for the single carrier and multiple carriers CR communications. An overall mean-square error (MSE) minimization problem is considered in [32], where the channel errors are modeled as a ball set which is a special case of the ellipsoidal set. A secondary multicast system in which all the SRs and PRs are equipped with a single antenna is proposed in [69], where the channel uncertainties of both the ST-SR channels and ST-PR channels fall in ellipsoidal uncertainty sets, and the minimum secondary SNR

is to be maximized.

The above problems are either convex problems, or nonconvex problems solved by using the SDR and randomization techniques. As mentioned in Chapter 5, the individual SNR requirements involve the eigenvalues of the product of matrix variables. To the best of authors' knowledge, it is impossible to generate a solution satisfying constraints on eigenvalues from the randomization technique. Therefore, we will continue to use the matrix decomposition technique to derive solutions and analyze the performance. Because of the channel uncertainties, the problem becomes more complicated. We provide a local optimal solution and a suboptimal solution with less computational complexity. Performance will be evaluated in the simulation section.

6.1 System Model

The system model structure considered in this chapter is nearly the same as that used in Chapter 5 except PR is equipped with single antenna and the ST-PR channel is denoted by $\mathbf{h}_x \in \mathbb{C}^{m \times 1}$. The interference received at PR is expressed as $\mathbf{h}_x^H \mathbf{T}_s \mathbf{x}_s$. Assume that ST and SR know the ST-SR channel state information perfectly, while ST only has imperfect knowledge about the ST-PR channel state information. The actual ST-PR channel is given by

$$\mathbf{h}_x = \hat{\mathbf{h}}_x + \mathbf{e}_x \quad (6.1)$$

where $\hat{\mathbf{h}}_x$ is the erroneous estimate known by ST, and \mathbf{e}_x is the channel estimation error. Like [31], we assume that the possible values of \mathbf{e}_x are bounded by the following ellipsoid expression

$$\mathbf{e}_x^H \mathbf{W}_x \mathbf{e}_x \leq 1 \quad (6.2)$$

where $\mathbf{W}_x \succ \mathbf{0}$ determines the shape and amplitude of the channel error.

As mentioned in Section 3.1, the secondary transmit beamforming matrix can be expressed as $\mathbf{T}_s = \mathbf{M}^{-1/2} \mathbf{V}_s \mathbf{\Sigma}^{1/2}$. Then the interference power at PR is

$$\begin{aligned} E_{\mathbf{x}}(\mathbf{h}_x^H \mathbf{T}_s \mathbf{x} \mathbf{x}^H \mathbf{T}_s^H \mathbf{h}_x) &= \mathbf{h}_x^H \mathbf{T}_s \mathbf{T}_s^H \mathbf{h}_x \\ &= (\hat{\mathbf{h}}_x + \mathbf{e}_x)^H \mathbf{M}^{-1/2} \mathbf{V}_s \mathbf{\Sigma} \mathbf{V}_s^H \mathbf{M}^{-1/2} (\hat{\mathbf{h}}_x + \mathbf{e}_x) \\ &\triangleq (\hat{\mathbf{h}}_x + \mathbf{e}_x)^H \mathbf{F}(\mathbf{V}_s) (\hat{\mathbf{h}}_x + \mathbf{e}_x). \end{aligned} \quad (6.3)$$

Notice that $\mathbf{F}(\mathbf{V}_s)$ has contained the SNR requirements and information of other channels except \mathbf{h}_x .

If ST hopes to use a \mathbf{V}_s to form the transmit beamforming matrix and send data to SR, it should guarantee that the peak primary interference power constraint, ξ , is not violated over all \mathbf{e}_x . Let $I(\mathbf{V}_s)$ denote the maximum possible interference power, which can be formulated as

$$I(\mathbf{V}_s) = \max_{\mathbf{e}_x} (\hat{\mathbf{h}}_x + \mathbf{e}_x)^H \mathbf{F}(\mathbf{V}_s) (\hat{\mathbf{h}}_x + \mathbf{e}_x). \quad (6.4)$$

In the following, we use I to replace $I(\mathbf{V}_s)$ for the sake of simplicity. ST is allowed to access the channel with \mathbf{V}_s only when $I \leq \xi$. Due to the uncertainty of channel error, given SNR requirement together with the interference power constraint, the condition $I \leq \xi$ may not be always satisfied. For example, when ST-SR channel is under deep fading, ξ is small and possible channel errors along all space directions are large, ST must use low transmit power to keep interference acceptable but at the same time the power does not suffice to compensate the fading to achieve required SNR. Then ST should keep silent till channel condition becomes good enough.

There is no way to entirely avoid the occurrence of such situation unless $\hat{\mathbf{h}}_x = \mathbf{h}_x$ or $\xi = \infty$. We call the probability that $I \leq \xi$ secondary access probability. Intuitively, the increase of minimum eigenvalue of \mathbf{W}_x (implying $\hat{\mathbf{h}}_x$ getting more close to \mathbf{h}_x) or the increase of ξ would increase secondary access probability. Besides, different \mathbf{V}_s may lead to different secondary access pro-

bability. Higher secondary access probability means higher throughput for SUs, so a method for achieving high secondary access probability is required.

6.2 Maximization of Secondary Access Probability

Let us define I_{min} as

$$I_{min} = \min_{\mathbf{V}_s} I. \quad (6.5)$$

Obviously, ST gains the maximum probability for using channel if $I = I_{min}$. Thus the problem of maximizing secondary access probability is converted to minimizing I , i.e.,

$$(P6.1) \quad \min_{\mathbf{V}_s, I} I \quad (6.6a)$$

$$\text{s.t. } (\hat{\mathbf{h}}_x + \mathbf{e}_x)^H \mathbf{F}(\mathbf{V}_s) (\hat{\mathbf{h}}_x + \mathbf{e}_x) \leq I \quad (6.6b)$$

$$\mathbf{e}_x^H \mathbf{W}_x \mathbf{e}_x \leq 1 \quad (6.6c)$$

$$\mathbf{V}_s \in St(m, d). \quad (6.6d)$$

Because \mathbf{e}_x has infinite possible values, (P6.1) is a semi-infinite min-max problem (SMMP) [74]. General algorithms for solving a SMMP problem usually require exhaustive computation [74] and are not suitable for a wireless device without extremely powerful computational ability.

6.2.1 S-procedure Transformation

Referring to Section 2.4, \mathbf{e}_x in (P6.1) acts as \mathbf{v} in (2.14) and constraints (6.6b) and (6.6c) are in quadratic forms. Viewing (6.6b) as function σ_0 while (6.6c) as function σ_1 , by applying S-procedure, these two constraints can be

converted to

$$\begin{bmatrix} I - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau & -\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \\ -\mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x & -\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x \end{bmatrix} \succeq \mathbf{0} \quad (6.7)$$

where τ is a positive number depending on I and \mathbf{V}_s .

Without affecting the performance, we can replace the sign \succeq in (6.7) by $<$, i.e.,

$$(\hat{\mathbf{h}}_x + \mathbf{e}_x)^H \mathbf{F}(\mathbf{V}_s) (\hat{\mathbf{h}}_x + \mathbf{e}_x) < I, \quad (6.8)$$

then the matrix in (6.7) is strictly larger than $\mathbf{0}$. From the strict Schur complement formula, the constraints (6.6c) and (6.8) are equal to the following two expressions

$$-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x \succ \mathbf{0} \quad (6.9)$$

$$(I - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x > 0. \quad (6.10)$$

As a result, (P6.1) can be rewritten as

$$(P6.2) \quad \min_{\mathbf{V}_s, I, \tau} I \quad (6.11a)$$

$$\text{s.t. } (I - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x > 0 \quad (6.11b)$$

$$\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succ \mathbf{0} \quad (6.11c)$$

$$\mathbf{V}_s \in St(m, d), \tau > 0. \quad (6.11d)$$

To the best of our knowledge, there is no closed-form solution to express the relationship between I and \mathbf{V}_s . We need to develop an algorithm to obtain I_{min} . To minimize the objective function with constraint inequalities, we may use the logarithmic barrier function method to combine both the objective function and constraint functions to generate the result via a Newton's method based algorithm. However, this method requires that all the values of constraint functions are scalar, while the left hand size of (6.11c) is a matrix. The logarithmic barrier function method cannot be directly applied in

our problem. Notice that the feasible domains of variables, \mathbf{V}_s and τ , are continuous. After we find a suitable I , \mathbf{V}_s and τ as initial parameters which satisfy (6.11b) and (6.11c), then during the implementation of the algorithm, $\det(-\mathbf{F}(\mathbf{V}_s) + \tau\mathbf{W}_x) > 0$ is *almost surely* equivalent to (6.11c). Here we use *almost surely* because the eigenvalues of $-\mathbf{F}(\mathbf{V}_s) + \tau\mathbf{W}_x$ are all positive when (6.11c) satisfied, but there are only a few and finite gradient directions \mathbf{Z} that may lead to even number of eigenvalues of $-\mathbf{F}(\pi(\mathbf{V}_s) + \gamma\mathbf{Z}) + \tau\mathbf{W}_x$ converting from positive to negative at the same time with appropriate step size γ . It is almost impossible that the descent direction exactly coincides with any one of them. According to our simulation test, such situation did not happen. So we believe it is reasonable to ignore this problem.

Given a \mathbf{V}_s , we set τ_0 to be the largest number that makes $\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s)$ a singular matrix, i.e., $\det(\tau_0\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s)) = 0$ and $\det(\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s)) > 0, \forall \tau > \tau_0$. τ_0 always exists because when $\tau \rightarrow \infty$, $\tau\mathbf{W}_x - \mathbf{F} \approx \tau\mathbf{W}_x$ and $\mathbf{W}_x \succ \mathbf{0}$. It is easy to search the value of τ_0 , e.g., by using bisection method or some other algorithms.

By applying the logarithmic barrier function method, we obtain (P6.3) from (P6.2):

$$(P6.3) \quad \min_{\mathbf{V}_s, I, \tau} I - \frac{1}{t} \log((I - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x) - \frac{1}{t} \log(\det(\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))) \quad (6.12a)$$

$$\text{s.t. } \mathbf{V}_s \in St(m, d), \tau \geq \tau_0, I > 0. \quad (6.12b)$$

Let $\beta(\mathbf{V}_s, \tau) \triangleq \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x + \tau + \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x$. Calculating the derivative of (6.12a) with respect to I and setting the result equal to 0, the I that minimizes (6.12a) is $I = \frac{1}{t} + \beta(\mathbf{V}_s, \tau)$. The logarithmic barrier function method requires t to increase to a sufficiently larger number so as to accurately approximate the barrier function (2.4). As t approaches

the infinity, we have

$$I = \beta(\mathbf{V}_s, \tau). \quad (6.13)$$

Therefore, we need to consider $\beta(\mathbf{V}_s, \tau)$ only and ignore other terms in (6.12a).

6.2.2 Suboptimal Method

Notice the case $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) = \mathbf{0}$, i.e., $\hat{\mathbf{h}}_x^H \mathbf{M}^{-1/2} \mathbf{V}_s = \mathbf{0}$. $\beta(\mathbf{V}_s, \tau)$ is minimized when $\tau = \tau_0$ and $\beta(\mathbf{V}_s, \tau) = \tau_0$ since other terms of $\beta(\mathbf{V}_s, \tau)$ are zero. In the suboptimal algorithm, we only consider the case when \mathbf{V}_s lies in the null space of $\hat{\mathbf{h}}_x^H \mathbf{M}^{-1/2}$. Let $\mathbf{V}_\perp \in \mathbb{C}^{m \times (m-1)}$ denote a matrix satisfying $\hat{\mathbf{h}}_x^H \mathbf{M}^{-1/2} \mathbf{V}_\perp = \mathbf{0}$, then any \mathbf{V}_s that lies in the null space of $\hat{\mathbf{h}}_x^H \mathbf{M}^{-1/2}$ satisfies $\mathbf{V}_s = \mathbf{V}_\perp \mathbf{T}$ where $\mathbf{T} \in St(m-1, d)$. Since different \mathbf{V}_s has different related τ_0 , our target is to find a \mathbf{V}_s which minimizes τ_0 . Abbreviate $\mathbf{F}(\mathbf{V}_s)$ as \mathbf{F} for the sake of simplicity, the problem is formulated as:

$$(P6.4) \quad \min_{\mathbf{T}} \tau_0 \quad (6.14a)$$

$$\text{s.t. } \hat{\mathbf{h}}_x^H \mathbf{F} = \mathbf{0} \quad (6.14b)$$

$$\tau_0 \mathbf{W}_x - \mathbf{F} \succeq \mathbf{0} \quad (6.14c)$$

$$\mathbf{T} \in St(m-1, d). \quad (6.14d)$$

Let $\mathbf{B}_1 \triangleq \tau_0 \mathbf{W}_x - \mathbf{F}$. It is reasonable to assume that $\text{rank}(\mathbf{B}_1) = m-1$. Then its adjugate matrix is

$$\mathbf{Q}_1 = \prod_{k=1}^{m_1} \lambda_k \mathbf{u}_1 \mathbf{u}_1^H \quad (6.15)$$

where λ_k denotes the k -th non-zero eigenvalue of \mathbf{B}_1 and \mathbf{u}_1 denotes the eigenvector belonging to the eigenvalue 0. As mentioned in Chapter 2,

$\det(\mathbf{B}_1 + d\mathbf{B}_1) = \text{tr}(\mathbf{Q}_1 d\mathbf{B}_1)$, then we have

$$\begin{aligned}
\mathbf{0} &= \det(\mathbf{B}_1 + d\mathbf{B}_1) - \det(\mathbf{B}_1) \\
&= \text{tr}(\mathbf{Q}_1 d\mathbf{B}_1) \\
&= \text{tr}(\mathbf{Q}_1 (d\tau_0 \mathbf{W}_x - d\mathbf{F})) \\
&= d\tau_0 \text{tr}(\mathbf{Q}_1 \mathbf{W}_x) - \text{tr}(\mathbf{Q}_1 d\mathbf{F}).
\end{aligned} \tag{6.16}$$

Therefore,

$$\begin{aligned}
d\tau_0 &= \text{tr}(\mathbf{Q}_1 d\mathbf{F}) / \text{tr}(\mathbf{Q}_1 \mathbf{W}_x) \\
&= \frac{\text{vec}^T(\mathbf{Q}_1^T)}{\text{tr}(\mathbf{Q}_1 \mathbf{W}_x)} \text{vec}(d\mathbf{F})
\end{aligned} \tag{6.17}$$

where $d\mathbf{F}$ and the derivative of τ_0 with respect to \mathbf{T} , denoted by $\mathbf{D}_{\mathbf{T}}$, can be calculated as provided in Appendix F. The descent direction is given by $\mathbf{Z}_{\mathbf{T}} = \mathbf{T}\mathbf{D}_{\mathbf{T}}^H\mathbf{T} - \mathbf{D}_{\mathbf{T}}$. Step size γ can be determined by Armijo step size rule. We may use gradient descent method to achieve the minimum τ_0 . We take the square root of inner product of $\mathbf{Z}_{\mathbf{T}}$ as the stopping criterion. If the inner product is sufficiently small, the algorithm can be considered converged. Different from the definition of inner product of common vector, the inner product of derivative on Stiefel manifold is defined as (2.6). The algorithm is designed as follows:

Algorithm 6.1 Algorithm for computing suboptimal τ_0

Step 0. Choose \mathbf{V}_{\perp} and $\mathbf{T} \in St(m-1, d)$, compute τ_0 .

Step 1. Compute the descent direction $\mathbf{Z}_{\mathbf{T}}$ and the inner product $\mathbf{S} = \langle \mathbf{Z}_{\mathbf{T}}, \mathbf{Z}_{\mathbf{T}} \rangle$. If $\sqrt{\mathbf{S}}$ is sufficiently small, then stop. Else go to Step 2.

Step 2. Determine step size γ according to Armijo step size rule. Set $\mathbf{T} := \pi(\mathbf{T} + \gamma\mathbf{D}_{\mathbf{T}})$.

Step 3. Compute the new τ_0 . Go to Step 1.

6.2.3 Optimal Method

From (P6.3) and (6.13), we rewrite the problem of minimizing I , i.e., $\beta(\mathbf{V}_s, \tau)$, as

$$(P6.5) \quad \min_{\mathbf{V}_s, \tau} \beta(\mathbf{V}_s, \tau) \quad (6.18a)$$

$$\text{s.t. } \mathbf{V}_s \in St(m, d), \tau \geq \tau_0 \quad (6.18b)$$

Note that I_{min} is equal to the minimum value of $\beta(\mathbf{V}_s, \tau)$ and the τ_0 generated by Algorithm 4.1 serves as an upper bound of I_{min} . Below we use τ_u to represent it.

$\beta(\mathbf{V}_s, \tau)$ consists of two terms in addition of $\hat{\mathbf{h}}_x \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x$: one is τ , a monotonically increasing function of τ ; the other is $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x$, a monotonically decreasing function of τ provided that $\tau > \tau_0$ and $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \neq \mathbf{0}$. Since τ_u is an upper bound of I_{min} and $\beta(\mathbf{V}_s, \tau) \geq \tau$, then I_{min} can be achieved only when $\tau < \tau_u$. Note that for any $\tau < \tau_u$, there is no \mathbf{V}_s satisfying $\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succeq \mathbf{0}$ and $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) = \mathbf{0}$ at the same time, so $\beta(\mathbf{V}_s, \tau)$ achieves its minimum value at the point $\frac{\partial \beta}{\partial \tau} = 0$, i.e.,

$$\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{W}_x (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x = 1. \quad (6.19)$$

As τ is obtained by calculating (6.19), we cannot write an explicit expression for the derivative of I with respect to \mathbf{V}_s without τ . The minimum interference value should be computed via an iterative algorithm. We may fix a τ to compute the optimal \mathbf{V}_s , and then use this \mathbf{V}_s to find the corresponding optimal τ , and then fix τ to compute \mathbf{V}_s again, and so on and so

forth, till both τ and \mathbf{V}_s converge. This is the main technique we are to apply. However, if we choose a random \mathbf{V}_s as a starting point to implement iterative algorithm, it is likely to converge to a local optimum that $\mathbf{F}(\mathbf{V}_s)\hat{\mathbf{h}}_x = \mathbf{0}$ and $\min_{\mathbf{V}_s} \beta(\mathbf{V}_s, \tau) = \tau_0$.

In order to avoid the occurrence of this situation, we first need to find out the minimum τ satisfying (6.11c) and the related \mathbf{V}_s , represented as τ_l and $\mathbf{V}_{s,l}$, respectively. τ_l serves as a lower bound of I_{min} because any $\tau < \tau_l$ violates $\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succeq \mathbf{0}$ and $I = \beta(\mathbf{V}_s, \tau) \geq \tau \geq \tau_l, \forall \mathbf{V}_s \in St(m, d)$. And then we search for a τ from τ_l to τ_u such that when the iterative algorithm starts, the updated τ is still smaller than τ_u . A τ satisfying the above requirement does not always exist, depending on the channel condition and \mathbf{W}_x . If there is no such τ , then τ_u is the minimum achievable interference. If it exists, then after getting this τ and the corresponding $\mathbf{V}_{s,l}$, we can start the iterative algorithm with fixed $\mathbf{V}_{s,l}$.

The problem of finding τ_l is how to get the minimum τ such that there exists a \mathbf{V}_s making $\tau\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succeq \mathbf{0}$, which can be formulated as

$$\min_{\mathbf{V}_s, \tau_0} \tau_0 \quad (6.20a)$$

$$\text{s.t. } \tau_0\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succeq \mathbf{0} \quad (6.20b)$$

$$\mathbf{V}_s \in St(m, d), \tau_0 \geq 0. \quad (6.20c)$$

As mentioned before, given \mathbf{V}_s , we can find a τ_0 such that $\tau_0\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succeq \mathbf{0}$ and $\det(\tau_0\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s)) = 0$. Similar to Algorithm 4.1, let $\mathbf{B}_2 \triangleq \tau_0\mathbf{W}_x - \mathbf{F}(\mathbf{V}_s)$. The adjugate matrix of \mathbf{B}_2 is

$$\mathbf{Q}_2 = \prod_{k=1}^{m_1} \lambda_k \mathbf{u}_2 \mathbf{u}_2^H \quad (6.21)$$

where λ_k denotes the k -th non-zero eigenvalue of \mathbf{B}_2 and \mathbf{u}_2 denotes the

eigenvector belonging to the eigenvalue 0. We have

$$\begin{aligned}
\mathbf{0} &= \det(\mathbf{B}_2 + d\mathbf{B}_2) - \det(\mathbf{B}_2) \\
&= \text{tr}(\mathbf{Q}_2 d\mathbf{B}_2) \\
&= \text{tr}(\mathbf{Q}_2 (d\tau_0 \mathbf{W}_x - d\mathbf{F})) \\
&= d\tau_0 \text{tr}(\mathbf{Q}_2 \mathbf{W}_x) - \text{tr}(\mathbf{Q}_2 d\mathbf{F}).
\end{aligned} \tag{6.22}$$

Thus

$$\begin{aligned}
d\tau_0 &= \text{tr}(\mathbf{Q}_2 d\mathbf{F}) / \text{tr}(\mathbf{Q}_2 \mathbf{W}_x) \\
&= \frac{\text{vec}^T(\mathbf{Q}_2^T)}{\text{tr}(\mathbf{Q}_2 \mathbf{W}_x)} \text{vec}(d\mathbf{F})
\end{aligned} \tag{6.23}$$

The detailed derivation of the derivative of τ_0 with respect to \mathbf{V}_s , denoted by \mathbf{D}_τ , is provided in Appendix F. Then the descent direction is $\mathbf{Z}_\tau = \mathbf{V}_s \mathbf{D}_\tau^H \mathbf{V}_s - \mathbf{D}_\tau$. Step size γ can be determined by Armijo step size rule. We may use gradient descent method to achieve the minimum τ_0 .

After that we need to find out if there is a $\tau \in (\tau_l, \tau_u)$ and a \mathbf{V}_s resulting in $\beta(\mathbf{V}_s, \tau) < \tau_u$. Assuming such τ exists, the interference minimization problem can be written as:

$$(P6.6) \quad \min_{\mathbf{V}_s, \tau} \beta(\mathbf{V}_s, \tau) \tag{6.24a}$$

$$\text{s.t. } \tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s) \succeq \mathbf{0} \tag{6.24b}$$

$$\mathbf{V}_s \in St(m, d), \tau \in (\tau_l, \tau_u). \tag{6.24c}$$

Temporarily dropping the constraint $\tau \in (\tau_l, \tau_u)$, then \mathbf{D}_ξ , the derivative of $\beta(\mathbf{V}_s, \tau)$ with respect to \mathbf{V}_s , can be obtained as in Appendix C. The descent direction is $\mathbf{Z}_\xi = \mathbf{V}_s \mathbf{D}_\xi^H \mathbf{V}_s - \mathbf{D}_\xi$. The step size can be chosen from Armijo step size rule. The iterative algorithm can be developed as follows:

Algorithm 6.2 Generate τ and \mathbf{V}_s for computing I_{min} with an initial $\tau < \tau_u$ and the related \mathbf{V}_s

Step 0. Choose $\mathbf{V}_s = \mathbf{V}_{s,l}$ and $\tau \in (\tau_l, \tau_u)$.

Step 1. Compute the descent direction \mathbf{Z}_ξ and the inner product $\mathbf{S} = \langle \mathbf{Z}_\xi, \mathbf{Z}_\xi \rangle$. If $\sqrt{\mathbf{S}}$ is sufficiently small, then stop. Else go to Step 2.

Step 2. Determine step size γ according to Armijo step size rule. Set $\mathbf{V}_s := \pi(\mathbf{V}_s + \gamma \mathbf{V}_s)$.

Step 3. Compute the new τ using (6.19). Go to Step 1.

The Step 0 of Algorithm 4.2 requires a predetermined τ to start the algorithm. However, not all $\tau \in (\tau_l, \tau_u)$ are feasible to be selected as an initial τ . Sometimes given an initial $\tau \in (\tau_l, \tau_u)$, after the updated operation in the Step 3 of Algorithm 4.2, τ will become larger than τ_u . According to our test, if Algorithm 4.2 obtains a τ larger than τ_u during the implementation, then it can no longer obtain a τ small than τ_u and hence the result will converge to τ_u finally. We find that as long as $\tau \in (\tau_l, \tau_u)$ after the Step 3 of Algorithm 4.2 is carried out for the first time, the constraint $\tau \in (\tau_l, \tau_u)$ will not be ruined till the algorithm stops. Thus we only need to implement Algorithm 4.2 for one round to check the feasibility of an initial τ . If there exists feasible τ , then at least every τ in a continuous region within (τ_l, τ_u) is feasible. So we do not have to exhaustively search every $\tau \in (\tau_l, \tau_u)$. Instead, we may separate (τ_l, τ_u) into small pieces and pick up an arbitrary τ from each piece to check the feasibility. If a feasible region is found, we randomly choose a τ in the feasible region to implement Algorithm 4.2. The value of I_{min} is either τ_u or the result of Algorithm 4.2.

The complete algorithm is presented as follows:

Algorithm 6.3 Complete algorithm for generating I_{min}

Step 0. Choose a random \mathbf{V}_s , $N = 2$

Step 1. Implement Algorithm 4.1 and set $\tau_u := \tau_0$. If $\tau_u < \xi$, stop. Else go to Step 2.

Step 2. Implement Algorithm 4.2 and set $\tau_l := \tau_0$.

Step 3. Choose interval $\theta := (\tau_u - \tau_l)/N$, $k := 1$.

Step 4. Set $\tau := \tau_l + k\theta$. Fix τ , compute gradient descent direction \mathbf{Z}_ξ and step size γ according to Armijo step size rule till $\sqrt{\langle \mathbf{Z}_\xi, \mathbf{Z}_\xi \rangle}$ is sufficiently small.

Step 5. Compute τ from (6.19). if $\tau < \tau_u$, then go to Step 6. Else if $k < N - 1$, set $k := k + 2$ and go to Step 4; otherwise set $N := 2N$. If N is sufficiently large, set $\xi = \xi_u$ and stop; else go to Step 3.

Step 6. Compute τ such that $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x + \tau + \hat{\mathbf{h}}_x \mathbf{F}(\mathbf{V}_s) (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x$ is minimized.

Step 7. Fix τ , compute the descent direction \mathbf{Z}_ξ and inner product $\mathbf{S}_\xi, \triangleq \langle \mathbf{Z}_\xi, \mathbf{Z}_\xi \rangle$. If $\sqrt{\mathbf{S}_\xi}$ is sufficiently small, stop. Else go to Step 8.

Step 8. Determine step size γ according to Armijo step size rule.

Step 9. Set $\mathbf{V}_s := \pi(\mathbf{V}_s + \gamma \mathbf{Z}_\xi)$. Go to Step 6.

6.3 Transmit Power Minimization

After confirming the feasibility, we have obtained a \mathbf{V}_s which does not violate all the constraints. With permission to access channel, ST may prefer to transmit with the least power. Then the next thing to do is to derive the optimal \mathbf{V}_s such that transmit power is minimized. The problem can be formulated as

$$(P6.7) \quad \min_{\mathbf{V}_s, \tau} \text{tr}(\boldsymbol{\Sigma} \mathbf{V}_s^H \mathbf{M}^{-1} \mathbf{V}_s) \quad (6.25a)$$

$$\text{s.t. } (\hat{\mathbf{h}}_x + \mathbf{e}_x)^H \mathbf{F}(\mathbf{V}_s) (\hat{\mathbf{h}}_x + \mathbf{e}_x) \leq \xi \quad (6.25b)$$

$$\mathbf{e}_x^H \mathbf{W}_x \mathbf{e}_x \leq 1 \quad (6.25c)$$

$$\mathbf{V}_s \in St(m, d), \tau \geq 0. \quad (6.25d)$$

By applying S-procedure, (P6.7) can be rewritten as

$$(P6.8) \quad \min_{\mathbf{V}_s} \text{tr}(\boldsymbol{\Sigma} \mathbf{V}_s^H \mathbf{M}^{-1} \mathbf{V}_s) \quad (6.26a)$$

$$\text{s.t. } (\xi - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x \geq 0 \quad (6.26b)$$

$$-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x \succeq \mathbf{0} \quad (6.26c)$$

$$\mathbf{V}_s \in St(m, d), \tau \geq 0. \quad (6.26d)$$

We perform the logarithmic barrier function method to combine the objective function and constraint functions of (P6.8) to get (P6.9):

$$\begin{aligned}
\text{(P6.9)} \quad & \min_{\mathbf{V}_s, \tau} f(\mathbf{V}_s, \tau) \triangleq \\
& \text{tr}(\boldsymbol{\Sigma} \mathbf{V}_s^H \mathbf{M}^{-1} \mathbf{V}_s) + \\
& \left(-\frac{1}{t}\right) \log((\xi - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \\
& \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x) + \left(-\frac{1}{t}\right) \log \det(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x) \\
& \text{s.t. } \mathbf{V}_s \in St(m, d), \tau \geq 0.
\end{aligned}$$

In (P6.9), if \mathbf{V}_s is given, then there exists a τ that minimizes (6.27). Therefore, τ can be viewed as an implicit function with respect to \mathbf{V}_s . However, to the best of our knowledge, we cannot exactly write an expression to replace τ using \mathbf{V}_s in (6.27). In the following, we use an iterative algorithm to compute the optimal \mathbf{V}_s and τ . We firstly fix τ and generate the optimal \mathbf{V}_s with gradient descent method. Next we obtain the optimal τ with the generated \mathbf{V}_s and then substitute the new τ into the first step to compute the corresponding optimal \mathbf{V}_s . Both step run iteratively till \mathbf{V}_s and τ converge.

The derivative of (6.26b) with respect to τ is equal to $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{W}_x (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - 1$. We let it be equal to zero and thus it is the same as (6.19). We solve (6.19) and update τ by choosing the largest number among the solutions. Because $\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{W}_x (\tau \mathbf{W}_x - \mathbf{F}(\mathbf{V}_s))^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x$ is equal to the positive infinity for $\tau = \tau_0$ and equal to 0 for $\tau = \infty$, there exists a solution between τ_0 and ∞ such that (6.19) and the constraint (6.26c) are satisfied.

To perform gradient descent method, we need to compute the derivative of $f(\mathbf{V}_s, \tau)$ with respect to \mathbf{V}_s with fixed τ , defined as \mathbf{D}_v . The expression of \mathbf{D}_v is given by:

$$\mathbf{D}_v = (2\boldsymbol{\Sigma} \mathbf{V}_s \mathbf{M}^{-1})^H + (RS_{d \times m}^H(\mathbf{A}_v \mathbf{D}_1)) + (RS_{d \times m}^H((\mathbf{A}_v \mathbf{D}_2)^C)) \quad (6.27)$$

where

$$\begin{aligned} \mathbf{A}_v = & \left(\frac{1}{t}\right) \frac{1}{(\xi - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x} \\ & \{\hat{\mathbf{h}}_x^T \otimes \hat{\mathbf{h}}_x^H + [(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x] \otimes \hat{\mathbf{h}}_x + \\ & [(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x] \otimes [(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}] + \\ & \hat{\mathbf{h}}_x^T \otimes [-\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}] + \\ & \mathbf{I} \otimes (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}\}, \end{aligned} \quad (6.28)$$

$$\mathbf{D}_1 = (\boldsymbol{\Sigma} \mathbf{V}_s^H \mathbf{M}^{-1/2})^T \otimes \mathbf{M}^{-1/2}, \quad (6.29)$$

and

$$\mathbf{D}_2 = [(\mathbf{M}^{-1/2})^T \otimes (\mathbf{M}^{-1/2} \mathbf{V}_s \boldsymbol{\Sigma})] \mathbf{P}^H. \quad (6.30)$$

\mathbf{P} is a permutation matrix only depending on m and d such that

$$\text{vec}(\mathbf{X}^H) = \mathbf{P} \text{vec}(\mathbf{X}^C), \forall \mathbf{X} \in \mathbb{C}^{m \times d}. \quad (6.31)$$

Please refer to Appendix H for the detailed derivation of \mathbf{D}_v . The descent direction is expressed as $\mathbf{Z}_v = \mathbf{V}_s \mathbf{D}_v^H \mathbf{V}_s - \mathbf{D}_v$. As to the step size γ , we choose Armijo step size rule to decide its value. The whole algorithm is listed as follows:

Algorithm 6.4 Power minimization algorithm

Step 0. Choose $\mathbf{V}_s \in St(m, d)$ and $\tau \geq 0$ obtained from the Algorithm 4.3. Set $t := 1$.

Step 1. Fix τ , compute the descent direction \mathbf{Z}_v . Determine step size γ according to Armijo step size rule. Set $\mathbf{V}_s := \pi(\mathbf{V}_s + \gamma \mathbf{Z}_v)$.

Step 2. Compute inner product $\mathbf{S} = \langle \mathbf{Z}_v, \mathbf{Z}_v \rangle$. If $\sqrt{\mathbf{S}}$ is sufficiently small, then go to Step 3. Else go to Step 1.

Step 3. Fix \mathbf{V}_s , compute the optimal τ^* . If $|\tau^* - \tau|$ is sufficiently small, then go to Step 4. Else set $\tau := \tau^*$ and go to Step 1.

Step 4. If t is sufficiently large, then exit. Else set $t := 2t$ and go to Step 1.

6.4 Simulation Results

In this section, we evaluate the performance of the proposed algorithms. The simulation parameters are set as follows each entry of channel coefficient matrices \mathbf{G} , \mathbf{H} , \mathbf{G}_x and vector $\hat{\mathbf{h}}_x$ are i.i.d. and follows the distribution $\mathcal{CN}(0, 1)$. The additive white Gaussian noise is of zero mean and unit variance. We set $m = n = 3$, $p = q = 2$, and $d = 2$. The secondary SNR requirements are 7dB and 3dB. The transmit power P_p of PT is 10dB. \mathbf{W}_x is set to be $\sigma_w^2 \mathbf{I}$ where various value of σ_w^2 will be evaluated.

Fig. 6.1 displays the secondary access probability versus the interference power constraint ξ with $\sigma_w^2 = 1, 5, 10$. As observed from Fig. 6.1, the outage probability decreases and approaches 0 as ξ increases. The value of σ_w^2 affects performance significantly. Larger σ_w^2 means higher accuracy of $\hat{\mathbf{h}}_x$, and hence the secondary access probability would increase more rapidly. That is to say, there is a higher probability to gain secondary spectrum access even with a relatively small ξ .

It is also shown that the I computed by Algorithm 4.1 is generally very close to I_{min} . We find that usually I_{min} is the same as the value obtained by Algorithm 4.1. That is because we either fail to find a $\tau < \tau_u$ in Steps 3-5 of Algorithm 4.3, or we obtain a $\tau < \tau_u$ but actually it converges to τ_u . Only when there exists a convergence region within (τ_l, τ_u) can we have $I_{min} < \tau_u$.

In Fig. 6.2, the average transmit power is obtained by calculating the mean of transmit power when the current channel condition can support both

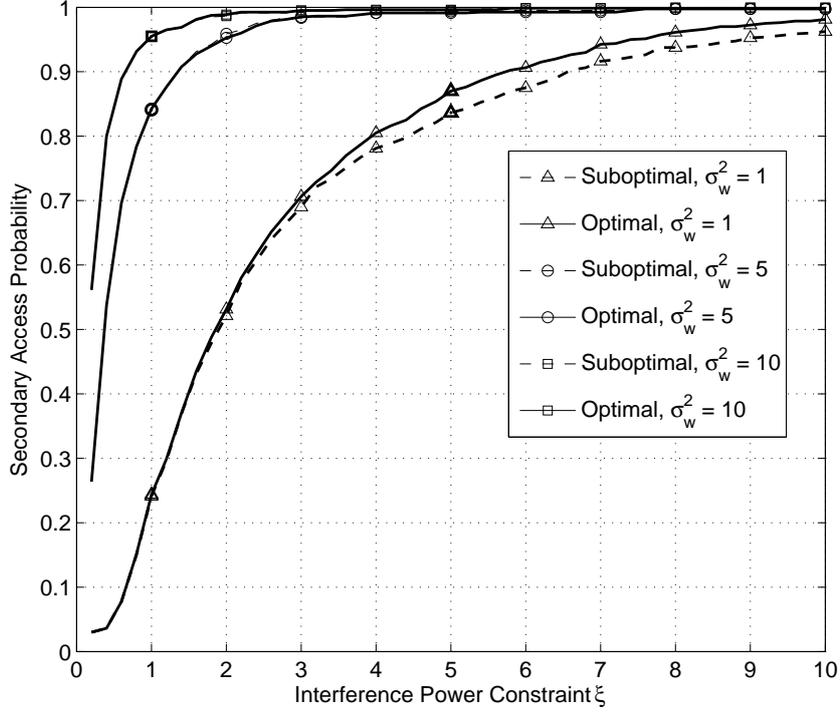


Figure 6.1: Outage probability versus interference power constraint.

the secondary SNR requirements and primary interference power constraint, i.e., the cases when the channel cannot support satisfactory transmission will not be calculated. This figure shows that as ξ increases, the average transmit power increases first, and then gradually falls to a floor. The reason why the required transmit power increases at the beginning is that, a small ξ restricts the upper bound of the transmit power, and ST must wait till ST-SR channel is good enough to achieve SNR requirement using very low power. Actually the low transmit power does not mean excellent performance. In this case ST has little opportunity to use the channel and must keep silent almost all the time. Since larger σ_w^2 implies better ability of ST to control the possible interference at PR, ST has more chances to transmit and hence the transmit power is higher. As ξ grows large enough, ST gains enough opportunities to

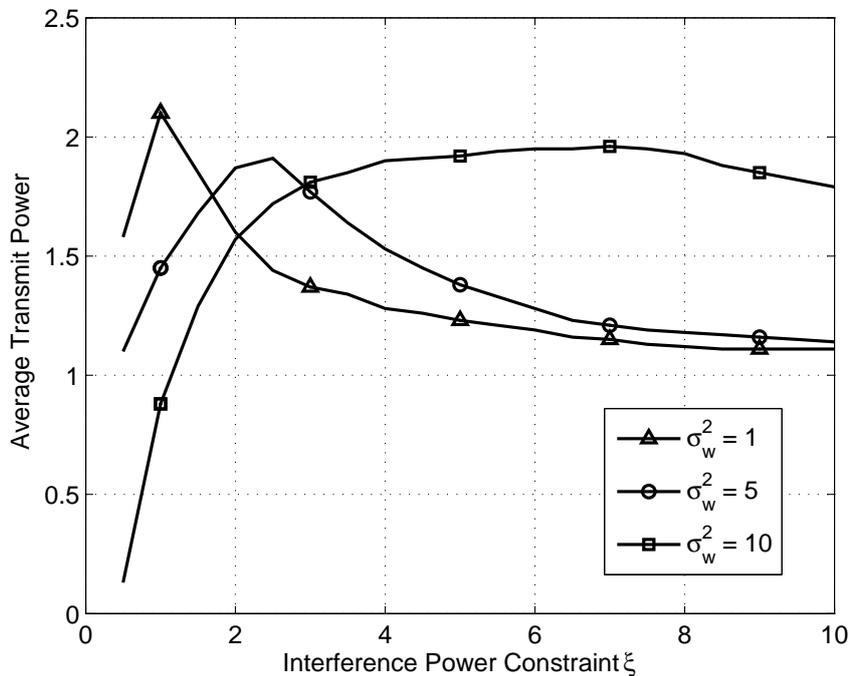


Figure 6.2: Average transmit power versus ξ under spectrum access.

transmit, then the curves reflect the relationship between σ_w^2 and transmit power. ST with smaller σ_w^2 should be more careful to select the beamforming direction such that I would not exceed ξ over all possible \mathbf{h}_x , so it is not able to allocate more power to some directions which may be more helpful to reduce sum transmit power but also may have higher interference increase rate. The ability of handling the impact of CSI error determines how well ST can reduce its transmit power. So ST with higher σ_w^2 can consume less power with a moderate ξ . When ξ is very large, the average transmit power gap between different values of σ_w^2 reduces with the increase of ξ , as shown in Fig. 6.2.

As ξ becomes larger (provided that there exists a feasible solution under the current channel conditions), the interference power constraint is relaxed, which means that the feasible domain of variable \mathbf{V}_s is enlarged. There-

fore, it is possible to find a new solution with smaller transmit power, and the average transmit power decreases. When ξ approaches infinity, ST can transmit without concerning about the interference power constraint. All the transmit power is used to satisfy the SINR requirements, that is why the graphs gradually fall to a floor when ξ increases, and this floor is unrelated to σ_w^2 .

6.5 Summary

In this chapter, the secondary power minimization problem with ellipsoid cross channel errors was studied. We decomposed the covariance of beamforming matrix to match the individual SNR requirements, then designed the optimal and suboptimal algorithms to calculate the minimum interference threshold value which can be achieved given the knowledge of a channel group realization $(\mathbf{H}, \hat{\mathbf{h}}_x)$. We proposed an algorithm to obtain the suboptimal solution for the power minimization problem if the actual interference threshold is achievable. System performance including channel access probability and transmit power were presented in the simulations. Although we focused on the one single-antenna PR scenario, actually the proposed method can be extended to the multiple single-antenna PRs scenario by modifying the interference constraints as well as the algorithms to adapt to the new environment.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this thesis, we studied the point-to-point communication in CR network with multiple data streams. We evaluated the performance boundaries and designed algorithms to increase spectrum efficiency and reduce resources cost in different communication scenarios, particularly with different extents of CSI knowledge. Throughput maximization and power minimization are the main targets in our consideration for the system design.

We first considered the throughput maximization problem. For the system model with partial CSI knowledge where the channel errors are element-wise bounded, we equivalently reformulated the channel error model and removed the uncertainties in the new formulation. And then we showed that the new problems were convex, therefore the optimal solution can be obtained via existing algorithms. The optimality conditions for both the single PR and multiple PRs cases were discussed. For the system model with no instantaneous CSI knowledge, by taking the interference power expectation and interference outage probability as the metrics to consider the interference constraint, we proposed different solutions to these problems ac-

cordingly. We equivalently rewrote the interference power in another form from the probability perspective. Then for the IE constraint problem, we showed that the problem can be formulated as a conventional point-to-point MIMO communication problem, which has an analytical solution; for the IO constraint problem, supposing that the channel errors follow complex Gaussian distribution, we provided the expression to calculate the cumulative density function (CDF) of interference power, and derived the algorithms corresponding to different CDFs to approach the local optimal solution. The proposed methods offer the opportunity to evaluate the potential impact on primary communications when a CR network is to be deployed in a primary communication environment that instantaneous CSI knowledge is unavailable.

Then we moved to the power minimization problem with individual SNR constraint for each secondary data stream. Similarly, we also considered the perfect CSI scenario and partial CSI scenario. For the perfect CSI scenario, we first decomposed the transmit beamforming matrix to three parts, then found out that among these parts, the diagonal matrix contributes to the received SNR values and the matrix on the Stiefel manifold affects the minimum transmit power. After that, we adjusted the diagonal matrix to make all the individual SNR constraints satisfied at SR, thus the problem became an optimization problem on the Stiefel manifold. Different from the throughput maximization problem, here the interference constraint and the individual SNR constraint may not be always satisfied at the same time. We proposed an efficient method to test the existence of a feasible solution. After passing the feasibility test, we derived the closed form solution for the zero interference constraint problem. For the nonzero interference constraint problem, although it is not convex due the nonconvexity of the variable domain, we found that global optimum can be obtained with a high-efficiency algorithm by analyzing the dual problem. The optimality holds for arbitrary

number of antennas at transceivers and arbitrary values of individual SNR requirements, as long as there is only one active PR in this CR network.

For the partial CSI problem, the channel error is bounded by an ellipsoid expression. Because the individual SNR requirements are nonconvex, this problem cannot be solved by convex optimization as the existing results on throughput maximization. We decomposed the transmit beamforming matrix like the perfect CSI scenario. In order to check whether the secondary transmission can be conducted in the current CSI condition, we used logarithmic barrier function to formulate the problem of minimizing interference power threshold and simplified the calculation by applying the optimal conditions. Besides, a more efficient solution is proposed for the suboptimal interference power threshold. Once the actual interference threshold is no smaller than the required value, i.e., the secondary transmission can be conducted, we proposed an algorithm to calculate the local optimal solution to minimize the secondary transmit power.

7.2 Future Work

In this thesis, we have considered cognitive beamforming for secondary spectrum access. With our results on throughput maximization, we may extend the work to cognitive broadcast and multicast transmission under various CSI conditions and different objectives. For example, with perfect or partial CSI knowledge, we can assume that in a multi-carrier communication scenario, there exists an ST transmitting to multiple SRs and each carrier is allocated to only one SR. Subject to the sum transmit power constraint, one may wish to maximize the minimum throughput in a broadcast scenario, and maximize the weighted sum throughput in a multicast scenario. These problems usually are not convex due to the exclusive carriers allocation. They are interesting open problems in transmissions with multiple data streams

in CR networks. Different types of CSI errors may lead to different interference constraints, as well as the methods to obtain solutions to achieve the required goals, especially in the case when only partial CSI is available. In this thesis, the element-wise bounded CSI error resulted from channel estimation is considered, but other types of CSI errors, e.g., unbounded error resulted from the feedback channel limit or feedback delay, are not discussed yet. Some statistical information of the CSI errors, e.g., expectation, covariance, distribution, is also not considered. Given these channel error models and statistical information, it is possible to exploit more features about the interference power. Moreover, we can consider other constraints about the statistical behaviours of the interference power.

In all the discussed schemes, we concentrated on the block fading channel model where the channel remains constant during one symbol block. Other types of channel fading are not considered yet. If the ST-SR channel experiences fast fading, the obtained capacity/throughput should be ergodic or robust to the imperfect channel information; if the ST-PR channel is fast fading, the interference power should be restricted within an acceptable region in the aspect of probability or maximum allowed interference. Besides, the multipath effect may be made use of to increase the performance. Under different channel conditions, as well as different number of antennas at receivers, it is possible to propose solutions to get the global optimal or local optimal results to match the transmission requirements.

Per transmit antenna power constraint is a practical subject for the transmit beamforming matrix design. Using OFDM to relieve the fast fading effect may impose strong requirement on the transmit power supply. Besides, beamforming causes unbalanced power assignment to each antenna. Therefore, with a RF chain limited multi-antenna ST, the sum power constraint and the per antenna power constraint are essential for designing a suitable transmit beamforming matrix. The per transmit antenna power constraint

itself is convex, so problems can be easily solved if it is convex without this constraint. However, when we are about to solve a problem consisting of a nonconvex objective function or other nonconvex constraints, the problem transformation and relaxation will be more complicated due to the additional per antenna power constraint. This is also one of the open problems we are interested to work on in the future for multiple data streams transmissions in CR networks.

Appendix A

Proof of Theorem 4.1

Firstly, we prove (4.8a). The transmit power can be decomposed as

$$\begin{aligned}\mathrm{tr}(\mathbf{Q}) &= \mathrm{tr}(\mathbf{Q}_1) + \mathrm{tr}(\mathbf{Q}_2) \\ &\stackrel{(a)}{=} \mathrm{tr}(\mathbf{Q}_1) + \mathrm{tr}(\mathbf{T}_\perp \mathbf{T}_\perp^H) \stackrel{(b)}{\geq} \mathrm{tr}(\mathbf{Q}_1)\end{aligned}\tag{A.1}$$

where (a) is due to $\mathrm{tr}(\mathbf{Q}_2) = 2\mathrm{Re}\{\mathrm{tr}(\mathbf{T}_{//} \mathbf{T}_\perp^H)\} + \mathrm{tr}(\mathbf{T}_\perp \mathbf{T}_\perp^H)$ and $\mathrm{tr}(\mathbf{T}_{//} \mathbf{T}_\perp^H) = \mathrm{tr}(\mathbf{T}_\perp^H \mathbf{T}_{//}) = 0$. As $\mathbf{T}_\perp \mathbf{T}_\perp^H$ is positive semi-definite, equality in (b) holds if and only if $\mathbf{T}_\perp = \mathbf{0}$. As a result, a nonzero \mathbf{Q}_2 requires a higher secondary transmit power.

Next, we show that $\mathbf{Q}_2 \neq \mathbf{0}$ leads to a higher interference outage probability. Before proving it, let us introduce the following lemma first.

Lemma A.1. Let $\lambda_k(\mathbf{Z})$ denote the k -th eigenvalue of Hermitian matrix \mathbf{Z} . Suppose that \mathbf{A} and \mathbf{B} are $a \times a$ matrices, and $\lambda_1(\mathbf{B}) \geq \dots \geq \lambda_a(\mathbf{B}) \geq 0$, $\lambda_1(\mathbf{A}) \geq \dots \geq \lambda_a(\mathbf{A})$. It follows that:

$$\lambda_k(\mathbf{A} + \mathbf{B}) \geq \lambda_k(\mathbf{A}) + \lambda_a(\mathbf{B}) \geq \lambda_k(\mathbf{A}).\tag{A.2}$$

Proof. The first inequality is obtained from the Weyl's inequality, and the second one holds because $\lambda_a(\mathbf{B}) \geq 0$. \square

Now we prove (4.8b). Because $\mathbf{T}_\perp \mathbf{T}_\perp^H$ is positive semi-definite, with Lemma A.1, we have

$$\begin{aligned}
\lambda_k(\mathbf{Q}) &= \lambda_k(\mathbf{T}\mathbf{T}^H) \\
&\stackrel{(c)}{=} \lambda_k(\mathbf{T}^H\mathbf{T}) \\
&\stackrel{(d)}{=} \lambda_k(\mathbf{Q}_1) + \lambda_k(\mathbf{T}_\perp^H\mathbf{T}_\perp) \geq \lambda_k(\mathbf{Q}_1)
\end{aligned} \tag{A.3}$$

where (c) is because for arbitrary $a \times b$ matrices \mathbf{A} and \mathbf{B} , the eigenvalues of $\mathbf{A}^H\mathbf{B}$ are the same as $\mathbf{B}\mathbf{A}^H$; (d) is because $\mathbf{T}_\perp^H\mathbf{T}_{//} = \mathbf{T}_{//}^H\mathbf{T}_\perp = \mathbf{0}$, then only \mathbf{Q}_1 and $\mathbf{T}_\perp^H\mathbf{T}_\perp$ are left. The equality in (A.3) holds if and only if $\mathbf{T}_\perp = \mathbf{0}$. As a result, the interference outage probability can be shown as

$$\begin{aligned}
&\mathcal{P}(\text{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^H) \geq \xi) \\
&= \mathcal{P}\left(\sum_{k=1}^m \lambda_k(\mathbf{Q})w_k \geq \xi\right) \\
&= \mathcal{P}\left(\sum_{k=1}^m \lambda_k(\mathbf{Q}_1 + \mathbf{T}_\perp^H\mathbf{T}_\perp)w_k \geq \xi\right) \\
&\geq \mathcal{P}\left(\sum_{k=1}^m \lambda_k(\mathbf{Q}_1)w_k \geq \xi\right) \\
&= \mathcal{P}(\text{tr}(\mathbf{G}\mathbf{Q}_1\mathbf{G}^H) \geq \xi).
\end{aligned} \tag{A.4}$$

Therefore, a nonzero \mathbf{T}_\perp always leads to a higher interference outage probability. This completes the proof.

Appendix B

Proof of SNR Expression

Letting $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_d]$ and denoting \mathbf{r}_i ($i = 1, \dots, d$) as the receive beamforming vector for the i th data stream, the SNR for the i th data stream is given by

$$\hat{\rho}_i = \frac{\mathbf{r}_i \mathbf{H} \mathbf{t}_i \mathbf{t}_i^H \mathbf{H}^H \mathbf{r}_i^H}{\mathbf{r}_i (\sum_{k \neq i} \mathbf{H} \mathbf{t}_k \mathbf{t}_k^H \mathbf{H}^H + \mathbf{W}) \mathbf{r}_i^H} = \frac{\mathbf{r}_i \mathbf{M}_1 \mathbf{r}_i^H}{\mathbf{r}_i \mathbf{M}_2 \mathbf{r}_i^H}. \quad (\text{B.1})$$

Equation (B.1) is a generalized Rayleigh quotient and the \mathbf{r}_i that maximizes $\hat{\rho}_i$ is the dominant eigenvector of $(\mathbf{M}_2^{-1/2})^H \mathbf{M}_1 \mathbf{M}_2^{-1/2}$ [93]. It thus follows that the maximum $\hat{\rho}_i$ is equal to the i -th eigenvalue of $\mathbf{T}^H \mathbf{H}^H \mathbf{W}^{-1} \mathbf{H} \mathbf{T}$.

Appendix C

Proof of Lemma 5.1

For any \mathbf{T} that satisfies (5.2), we define $\mathbf{P} \triangleq \mathbf{M}^{1/2}\mathbf{T} \in \mathbb{C}^{m \times d}$ and let \mathbf{p}_i denote the i th column of \mathbf{P} . Since $\mathbf{\Sigma}$ is a diagonal matrix, it follows from (5.2) that all \mathbf{p}_i 's are orthogonal to each other and the norm of \mathbf{p}_i is equal to $\Sigma_{ii}^{1/2}$. We normalize each \mathbf{p}_i to form a matrix $\mathbf{V} = [\mathbf{p}_1/|\mathbf{p}_1| \cdots \mathbf{p}_d/|\mathbf{p}_d|] = \mathbf{P}\mathbf{\Sigma}^{-1/2} \in St(m, d)$. This directly yields $\mathbf{T} = \mathbf{M}^{-1/2}\mathbf{P} = \mathbf{M}^{-1/2}\mathbf{V}\mathbf{\Sigma}^{1/2}$.

Appendix D

Proof of Lemma 5.3

It is shown in [92] that for any $w \times w$ Hermitian matrices \mathbf{A} and \mathbf{B} ,

$$\sum_{i=1}^w \lambda_i(\mathbf{A})\lambda_{w-i+1}(\mathbf{B}) \leq \text{tr}(\mathbf{A}\mathbf{B}) \quad (\text{D.1})$$

where $\lambda_i(\mathbf{X})$ denotes the i -th eigenvalue of matrix $\mathbf{X} \in \mathbb{C}^{w \times w}$ and $\lambda_1(\mathbf{X}) \geq \dots \geq \lambda_w(\mathbf{X})$.

Let $\Delta' \triangleq \text{diag}(\Delta \mathbf{0}) \in \mathbb{C}^{v \times v}$ and $\Theta' \triangleq [\Theta \ \Theta_{\perp}] \in \mathbb{C}^{v \times v}$ where $\text{rank}(\Theta_{\perp}) = v - u$ and $\Theta_{\perp}^H \Theta = \mathbf{0}$. Notice that $\lambda_{u+1}(\Delta') = \dots = \lambda_v(\Delta') = 0$ and the eigenvalues of $\Theta'^H \Omega \Theta'$ are equal to those of Ω . By using (D.1), we have

$$\begin{aligned} \sum_{i=1}^u \delta_i \omega_{v-i+1} &= \sum_{i=1}^v \lambda_i(\Delta') \lambda_{v-i+1}(\Omega) \\ &\leq \text{tr}(\Delta' \Theta'^H \Omega \Theta') \\ &= \text{tr}(\Delta \Theta^H \Omega \Theta). \end{aligned}$$

The equality holds when $\Theta'^H \Omega \Theta' = \text{diag}(\omega_v, \dots, \omega_1)$, i.e., $\Theta' = [\varphi_v, \dots, \varphi_1]$. It thus follows that $\Theta = [\varphi_v, \dots, \varphi_{v-u+1}]$. This completes the proof.

Appendix E

Proof of Lemma 5.4

Let $\Sigma' = \text{diag}(\Sigma, \mathbf{0}) \in \mathbb{R}^{m \times m}$, and $\mathbf{V}' = [\mathbf{V}_y \mathbf{V}_{y\perp}]$ where $\mathbf{V}_{y\perp} \in \mathbb{C}^{m \times (m-d)}$ and its i -th ($i = 1, \dots, m-d$) column is the eigenvector of $\mathbf{M}^{-1} + y\mathbf{M}_x$ corresponding to λ_{d+i} . The dual problem (P5.5) can be equivalently rewritten as

$$\max_{y \geq 0} \min_{\mathbf{V}' \in St(m, m)} \text{tr}(\Sigma' \mathbf{V}'^H (\mathbf{M}^{-1} + y\mathbf{M}_x) \mathbf{V}' - y\xi). \quad (\text{E.1})$$

Let $\pi : \mathbb{C}^{m \times m} \rightarrow St(m, m)$ be a projection operator mapping an $m \times m$ complex matrix to the closest point on the Stiefel manifold [15]. Suppose that \mathbf{Z} is the derivative of the \mathbf{V}' with respect to y . It means that given a sufficiently small increment Δy , the \mathbf{V}' related to $\mathbf{M}^{-1} + (y + \Delta y)\mathbf{M}_x$, denoted by \mathbf{V}'' , can be approximated as

$$\mathbf{V}'' \approx \pi(\mathbf{V}' + \Delta y \mathbf{Z}) = \mathbf{V}' + \Delta y \mathbf{Z} + O(\Delta y^2)$$

where \mathbf{Z} can be written as $\mathbf{Z} = \mathbf{V}' \mathbf{C}$ and $\mathbf{C} \in \mathbb{C}^{m \times m}$ is a skew-Hermitian matrix ($\mathbf{C}^H + \mathbf{C} = \mathbf{0}$) [15] to be determined. Let us define $\mathbf{\Lambda} \triangleq \text{diag}(\lambda_1, \dots, \lambda_m)$ and $\mathbf{F} \triangleq \mathbf{V}'^H \mathbf{M}_x \mathbf{V}'$. Since the columns of \mathbf{V}'' are the eigenvectors of $\mathbf{M}^{-1} + (y + \Delta y)\mathbf{M}_x$, the expression $\mathbf{V}''^H (\mathbf{M}^{-1} + (y + \Delta y)\mathbf{M}_x) \mathbf{V}''$ is a diagonal matrix consisting of the eigenvalues of $\mathbf{M}^{-1} + (y + \Delta y)\mathbf{M}_x$. As $\mathbf{V}''^H (\mathbf{M}^{-1} + (y +$

$\Delta y)\mathbf{M}_x)\mathbf{V}''$ can be decomposed as

$$\begin{aligned}
& \mathbf{V}''^H(\mathbf{M}^{-1} + (y + \Delta y)\mathbf{M}_x)\mathbf{V}'' \\
& \approx \pi(\mathbf{V}' + \Delta y\mathbf{Z})^H(\mathbf{M}^{-1} + (y + \Delta y)\mathbf{M}_x)\pi(\mathbf{V}' + \Delta y\mathbf{Z}) \\
& = \mathbf{V}'^H(\mathbf{M}^{-1} + y\mathbf{M}_x)\mathbf{V}' + \Delta y(\mathbf{Z}^H(\mathbf{M}^{-1} + y\mathbf{M}_x)\mathbf{V}' \\
& \quad + \mathbf{V}'^H(\mathbf{M}^{-1} + y\mathbf{M}_x)\mathbf{Z} + \mathbf{V}'^H\mathbf{M}_x\mathbf{V}') + O(\Delta y^2) \\
& = \mathbf{\Lambda} + \Delta y(-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C} + \mathbf{F}) + O(\Delta y^2), \tag{E.2}
\end{aligned}$$

the matrix $(-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C} + \mathbf{F})$ must be diagonal. By computing the expressions of the entries of $-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C} + \mathbf{F}$ and setting the off-diagonal entries to be zero, it turns out that

$$(\lambda_i - \lambda_j)C_{ij} = -F_{ij}, i \neq j. \tag{E.3}$$

Now let us consider the impact of positive Δy on the interference power. Let $\mathbf{C}^{(d)}$ and $\mathbf{C}^{(od)}$ denote the matrices that only contain the diagonal and off-diagonal entries of \mathbf{C} , respectively. The difference of interference power

at a sufficiently small increment $\Delta y > 0$ is given by

$$\begin{aligned}
& \text{tr}(\boldsymbol{\Sigma}'\pi(\mathbf{V}' + \Delta y\mathbf{Z})^H\mathbf{M}_x\pi(\mathbf{V}'^H + \Delta y\mathbf{Z})) - \text{tr}(\boldsymbol{\Sigma}'\mathbf{V}'^H\mathbf{M}_x\mathbf{V}') \\
& \approx 2\Delta y\text{Re}\{\text{tr}(\boldsymbol{\Sigma}'\mathbf{V}'^H\mathbf{M}_x\mathbf{Z})\} \\
& = 2\Delta y\text{Re}\{\text{tr}(\boldsymbol{\Sigma}'\mathbf{F}(\mathbf{C}^{(d)} + \mathbf{C}^{(od)}))\} \\
& \stackrel{(a)}{=} 2\Delta y\text{Re}\{\text{tr}(\boldsymbol{\Sigma}'\mathbf{F}\mathbf{C}^{(od)})\} \\
& = 2\Delta y\sum_{i=1}^d\rho_i\left(\sum_{j\neq i}^d F_{ij}C_{ji}^{(od)}\right) \\
& \stackrel{(b)}{=} 2\Delta y\sum_{i=1}^d\rho_i\left(\sum_{j\neq i}^d(\lambda_i - \lambda_j)|C_{ij}^{(od)}|^2\right) \\
& = 2\Delta y\left(\sum_{i=1}^{d-1}\sum_{j=i+1}^d(\rho_i - \rho_j)(\lambda_i - \lambda_j)|C_{ij}^{(od)}|^2\right. \\
& \quad \left. + \sum_{i=1}^d\sum_{j=d+1}^m\rho_i(\lambda_i - \lambda_j)|C_{ij}^{(od)}|^2\right) \\
& \stackrel{(c)}{\leq} 0.
\end{aligned} \tag{E.4}$$

In (E.4), (a) holds because $\text{Re}\{\text{tr}(\boldsymbol{\Sigma}'\mathbf{V}'^H\mathbf{M}_x\mathbf{V}'\mathbf{C}^{(d)})\} = \text{Re}\{\text{tr}((\mathbf{C}^{(d)}\boldsymbol{\Sigma}')(\mathbf{V}'^H\mathbf{M}_x\mathbf{V}'))\} = 0$, where $\mathbf{C}^{(d)}\boldsymbol{\Sigma}'$ is skew-Hermitian and $\mathbf{V}'^H\mathbf{M}_x\mathbf{V}'$ is Hermitian. (b) is obtained by substituting (E.3). (c) is because $\rho_i \geq \rho_j$ and $\lambda_i \leq \lambda_j$ for $i < j$, and the equality holds if and only if $\rho_i = \rho_j$, $\lambda_i = \lambda_j$, $\forall i \neq j$. Considering the random nature of the channel conditions, it is nearly impossible for the equality under (c) to hold. Therefore, we conclude that the interference power in the form of $\text{tr}(\boldsymbol{\Sigma}'\mathbf{V}'^H\mathbf{M}_x\mathbf{V}')$ is a monotonically decreasing function of y .

Now let us see how the value of y affects the required transmit power. Equation (5.30) can be rewritten as

$$\begin{aligned}
& L(\mathbf{V}', y) \\
& = (1 + y)(\text{tr}(\boldsymbol{\Sigma}'\mathbf{V}'^H(\alpha\mathbf{M}^{-1} + (1 - \alpha)\mathbf{M}_x)\mathbf{V}')) - y\xi
\end{aligned} \tag{E.5}$$

where $\alpha = 1/(1 + y) \in [0, \infty)$. We see that the selected \mathbf{V}' in the dual problem is determined by the weighted sum of \mathbf{M}^{-1} and \mathbf{M}_x . A larger y increases the weight of \mathbf{M}_x and at the same time decreases the weight of \mathbf{M}^{-1} . Let $\Delta\alpha$ be the change of α as y becomes $y + \Delta y$ and $\mathbf{E} \triangleq \mathbf{V}'^H \mathbf{M}^{-1} \mathbf{V}'$. Similar to (E.2), it can be derived that

$$\begin{aligned} & \mathbf{V}''^H ((\alpha + \Delta\alpha)\mathbf{M}^{-1} + (1 - \alpha - \Delta\alpha)\mathbf{M}_x) \mathbf{V}'' \\ &= \alpha\mathbf{\Lambda} - \frac{\Delta\alpha}{\alpha} [(-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C}) - \alpha(\mathbf{E} - \mathbf{F})] + O(\Delta\alpha^2) \\ &= \alpha\mathbf{\Lambda} - \frac{\Delta\alpha}{\alpha} [\alpha(-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C} + \mathbf{F}) + (1 - \alpha)(-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C} \\ & \quad - \frac{\alpha}{1 - \alpha}\mathbf{E})] + O(\Delta\alpha^2) \end{aligned}$$

where the matrix $-\mathbf{C}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{C} - \frac{\alpha}{1 - \alpha}\mathbf{E}$ should be diagonal. As a result, we have

$$(\lambda_i - \lambda_j)C_{ij} = \frac{\alpha}{1 - \alpha}E_{ij} = \frac{1}{y}E_{ij}. \quad (\text{E.6})$$

With (E.6), the difference of the required transmit power at $\Delta y > 0$ can be shown as

$$\begin{aligned} & \text{tr}(\mathbf{\Sigma}'\pi(\mathbf{V}' + t\mathbf{Z})^H \mathbf{M}^{-1} \pi(\mathbf{V}'^H + t\mathbf{Z})) - \text{tr}(\mathbf{\Sigma}'\mathbf{V}'^H \mathbf{M}^{-1} \mathbf{V}') \\ & \approx -\frac{2\Delta y}{y} \left(\sum_{i=1}^{d-1} \sum_{j=i+1}^d (\rho_i - \rho_j)(\lambda_i - \lambda_j) |C_{ij}^{(od)}|^2 \right. \\ & \quad \left. + \sum_{i=1}^d \sum_{j=d+1}^m \rho_i (\lambda_i - \lambda_j) |C_{ij}^{(od)}|^2 \right) \\ & \geq 0. \end{aligned} \quad (\text{E.7})$$

With a similar argument used for (E.4), we conclude that the required transmit power in terms of $\text{tr}(\mathbf{\Sigma}'\mathbf{V}'^H \mathbf{M}^{-1} \mathbf{V}')$ is a monotonically increasing function of y . The monotonicity of functions $\text{tr}(\mathbf{\Sigma}'\mathbf{V}'^H \mathbf{M}_x \mathbf{V}')$ and $\text{tr}(\mathbf{\Sigma}'\mathbf{V}'^H \mathbf{M}^{-1} \mathbf{V}')$ will not change if $\mathbf{\Sigma}'$ and \mathbf{V}' are replaced by $\mathbf{\Sigma}$ and \mathbf{V}_y , respectively. This completes the proof.

Appendix F

Derivation of \mathbf{D}_T and \mathbf{D}_τ

Let $d\mathbf{F}$ denote $\mathbf{F}(\mathbf{V}_s + t\mathbf{Z}) - \mathbf{F}(\mathbf{V}_s) - O(t^2)$ where t is sufficiently small. Note that

$$\begin{aligned}\mathbf{F}(\mathbf{V}_s + t\mathbf{Z}) &= \mathbf{M}^{-1/2}(\mathbf{V}_s + t\mathbf{Z})\boldsymbol{\Sigma}(\mathbf{V}_s + t\mathbf{Z})^H\mathbf{M}^{-1/2} \\ &= \mathbf{F}(\mathbf{V}_s) + \mathbf{M}^{-1/2}(t\mathbf{Z})\boldsymbol{\Sigma}\mathbf{V}_s^H\mathbf{M}^{-1/2} + \mathbf{M}^{-1/2}\mathbf{V}_s\boldsymbol{\Sigma}(t\mathbf{Z}^H)\mathbf{M}^{-1/2} + O(t^2).\end{aligned}$$

Therefore,

$$\begin{aligned}d\mathbf{F} &= \mathbf{M}^{-1/2}(t\mathbf{Z})\boldsymbol{\Sigma}\mathbf{V}_s^H\mathbf{M}^{-1/2} + \mathbf{M}^{-1/2}\mathbf{V}_s\boldsymbol{\Sigma}(t\mathbf{Z}^H)\mathbf{M}^{-1/2} \\ &= \mathbf{M}^{-1/2}(t\mathbf{Z}\boldsymbol{\Sigma}\mathbf{V}_s^H + \mathbf{V}_s\boldsymbol{\Sigma}t\mathbf{Z}^H)\mathbf{M}^{-1/2}\end{aligned}\tag{F.1}$$

$$\begin{aligned}\text{vec}(d\mathbf{F}) &= \text{vec}(\mathbf{M}^{-1/2}(t\mathbf{Z}\boldsymbol{\Sigma}\mathbf{V}_s^H + \mathbf{V}_s\boldsymbol{\Sigma}t\mathbf{Z}^H)\mathbf{M}^{-1/2}) \\ &= \text{vec}(\mathbf{M}^{-1/2}t\mathbf{Z}\boldsymbol{\Sigma}\mathbf{V}_s^H\mathbf{M}^{-1/2}) + \text{vec}(\mathbf{M}^{-1/2}\mathbf{V}_s\boldsymbol{\Sigma}t\mathbf{Z}^H\mathbf{M}^{-1/2}) \\ &= [(\boldsymbol{\Sigma}\mathbf{V}_s^H\mathbf{M}^{-1/2})^T \otimes \mathbf{M}^{-1/2}]\text{vec}(t\mathbf{Z}) + \\ &\quad [(\mathbf{M}^{-1/2})^T \otimes (\mathbf{M}^{-1/2}\mathbf{V}_s\boldsymbol{\Sigma})]\text{vec}(t\mathbf{Z}^H).\end{aligned}\tag{F.2}$$

There exists a permutation matrix \mathbf{P} , which only depends on the size of \mathbf{Z} , such that

$$\mathbf{P}\text{vec}(t\mathbf{Z}^H) = \text{vec}^C(t\mathbf{Z}).$$

Consequently, (F.2) can be rewritten as

$$\begin{aligned}\text{vec}(d\mathbf{F}) &= [(\boldsymbol{\Sigma}\mathbf{V}_s^H\mathbf{M}^{-1/2})^T \otimes \mathbf{M}^{-1/2}]\text{vec}(t\mathbf{Z}) + [(\mathbf{M}^{-1/2})^T \otimes (\mathbf{M}^{-1/2}\mathbf{V}_s\boldsymbol{\Sigma})] \\ &\quad \mathbf{P}^H\text{vec}^C(t\mathbf{Z}) \\ &\triangleq \mathbf{D}_1\text{vec}(t\mathbf{Z}) + \mathbf{D}_2\text{vec}^C(t\mathbf{Z}).\end{aligned}\quad (\text{F.3})$$

When $\mathbf{Z} = \mathbf{V}_\perp\mathbf{T}$ as shown in Section 4.2.2, we obtain that

$$\text{vec}(d\mathbf{F}) = \mathbf{D}_1\text{vec}(t\mathbf{V}_\perp\mathbf{T}) + \mathbf{D}_2\text{vec}^C(t\mathbf{V}_\perp\mathbf{T}).\quad (\text{F.4})$$

By substituting (F.4) into (6.17), we have

$$\begin{aligned}d\tau_0 &= \frac{1}{\text{tr}(\mathbf{Q}_1\mathbf{W}_\mathbf{x})}\text{vec}^T(\mathbf{Q}_1^T)\text{vec}(d\mathbf{F}) \\ &= \frac{1}{\text{tr}(\mathbf{Q}_1\mathbf{W}_\mathbf{x})}\text{Re}\{\text{vec}^T(\mathbf{Q}_1^T)\mathbf{D}_1 + (\text{vec}^T(\mathbf{Q}_1^T)\mathbf{D}_2)^C\}\text{vec}(d\mathbf{V}_s) \\ &= \frac{1}{\text{tr}(\mathbf{Q}_1\mathbf{W}_\mathbf{x})}\text{Re}\{(\text{vec}^T(\mathbf{Q}_1^T)\mathbf{D}_1 + (\text{vec}^T(\mathbf{Q}_1^T)\mathbf{D}_2)^C)\mathbf{I} \otimes \mathbf{V}_\perp\text{vec}(d\mathbf{T})\}.\end{aligned}\quad (\text{F.5})$$

The derivative of τ_0 with respect to \mathbf{T} is given by

$$\mathbf{D}_\mathbf{T} = \frac{1}{\text{tr}(\mathbf{Q}_1\mathbf{W}_\mathbf{x})}RS_{d \times m}^H(\text{vec}(\mathbf{Q}_1^T)^T\mathbf{D}_1 + (\text{vec}(\mathbf{Q}_1^T)^T\mathbf{D}_2)^C)\mathbf{I} \otimes \mathbf{V}_\perp.\quad (\text{F.6})$$

Similarly, when calculating τ_1 , we substituting (F.4) into (6.23) and have

$$\mathbf{D}_\tau = \frac{1}{\text{tr}(\mathbf{Q}_2\mathbf{W}_\mathbf{x})}RS_{d \times m}^H(\text{vec}(\mathbf{Q}_2^T)^T\mathbf{D}_1 + (\text{vec}(\mathbf{Q}_2^T)^T\mathbf{D}_2)^C).\quad (\text{F.7})$$

Appendix G

Derivation of \mathbf{D}'_{β}

Let $d\mathbf{F}$ denote $\mathbf{F}(\mathbf{V}_s + t\mathbf{Z}) - \mathbf{F}(\mathbf{V}_s) - O(t^2)$ where t is sufficiently small. With a fixed τ , the differential of β is given by

$$\begin{aligned}
 d\beta &= \{\hat{\mathbf{h}}_{\mathbf{x}}^T \otimes \hat{\mathbf{h}}_{\mathbf{x}}^H + [(-\mathbf{F}(\mathbf{V}_s) + \tau\mathbf{W}_{\mathbf{x}})^{-1}\mathbf{F}(\mathbf{V}_s)\hat{\mathbf{h}}_{\mathbf{x}}]^T \otimes \hat{\mathbf{h}}_{\mathbf{x}} + \\
 &\quad [(-\mathbf{F}(\mathbf{V}_s) + \tau\mathbf{W}_{\mathbf{x}})^{-1}\mathbf{F}(\mathbf{V}_s)\hat{\mathbf{h}}_{\mathbf{x}}]^T \otimes [(-\mathbf{F}(\mathbf{V}_s) + \tau\mathbf{W}_{\mathbf{x}})^{-1}] + \\
 &\quad \hat{\mathbf{h}}_{\mathbf{x}}^T \otimes [-\hat{\mathbf{h}}_{\mathbf{x}}^H \mathbf{F}(\mathbf{V}_s)(-\mathbf{F}(\mathbf{V}_s) + \tau\mathbf{W}_{\mathbf{x}})^{-1}]\text{vec}(d\mathbf{F}) \\
 &\triangleq \mathbf{D}'_{\beta}\text{vec}(d\mathbf{F}). \tag{G.1}
 \end{aligned}$$

As described in Appendix B, $\text{vec}(d\mathbf{F})$ can be rewritten as

$$\begin{aligned}
 \text{vec}(d\mathbf{F}) &= [(\boldsymbol{\Sigma}\mathbf{V}_s^H\mathbf{M}^{-1/2})^T \otimes \mathbf{M}^{-1/2}]\text{vec}(t\mathbf{Z}) + [(\mathbf{M}^{-1/2})^T \otimes (\mathbf{M}^{-1/2}\mathbf{V}_s\boldsymbol{\Sigma})]\mathbf{P}^H\text{vec}^C(t\mathbf{Z}) \\
 &\triangleq \mathbf{D}_1\text{vec}(t\mathbf{Z}) + \mathbf{D}_2\text{vec}^C(t\mathbf{Z}). \tag{G.2}
 \end{aligned}$$

Thus

$$d\beta = \text{Re}\{\text{tr}(((\mathbf{D}'_{\beta}\mathbf{D}_1) + (\mathbf{D}'_{\beta}\mathbf{D}_2)^C)\text{vec}(t\mathbf{Z}))\} \tag{G.3}$$

and the derivative of β with respect to \mathbf{V}_s is given by

$$\mathbf{D}_{\beta} = RS_{d \times m}^H((\mathbf{D}'_{\beta}\mathbf{D}_1) + (\mathbf{D}'_{\beta}\mathbf{D}_2)^C). \tag{G.4}$$

Appendix H

Derivation of $\mathbf{D}_\mathbf{v}$

Let $d\mathbf{F}$ denote $\mathbf{F}(\mathbf{V}_s + t\mathbf{Z}) - \mathbf{F}(\mathbf{V}_s) - O(t^2)$ where t is sufficiently small. We have

$$\begin{aligned}
& f(\mathbf{V}_s + t\mathbf{Z}) - f(\mathbf{V}_s) \\
= & 2\text{Re}\{\text{tr}(\boldsymbol{\Sigma}\mathbf{V}_s\mathbf{M}^{-1}t\mathbf{Z})\} + \\
& \left(-\frac{1}{t}\right) \frac{1}{(\xi - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x} \\
& \text{Re}\{-\hat{\mathbf{h}}_x^H (d\mathbf{F}) \hat{\mathbf{h}}_x + [-\hat{\mathbf{h}}_x^H (d\mathbf{F}) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x] + \\
& [-\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) ((-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} (d\mathbf{F}) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}) \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x] + \\
& [-\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} (d\mathbf{F}) \hat{\mathbf{h}}_x]\} + \\
& \left(\frac{1}{t}\right) \text{Re}\{\text{tr}[(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} (d\mathbf{F})]\} + O(t^2). \tag{H.1}
\end{aligned}$$

As described in Appendix B, $\text{vec}(d\mathbf{F})$ can be rewritten as

$$\begin{aligned}
\text{vec}(d\mathbf{F}) &= [(\boldsymbol{\Sigma}\mathbf{V}_s^H \mathbf{M}^{-1/2})^T \otimes \mathbf{M}^{-1/2}] \text{vec}(t\mathbf{Z}) + [(\mathbf{M}^{-1/2})^T \otimes (\mathbf{M}^{-1/2} \mathbf{V}_s \boldsymbol{\Sigma})] \\
& \quad \mathbf{P}^H \text{vec}^C(t\mathbf{Z}) \\
& \triangleq \mathbf{D}_1 \text{vec}(t\mathbf{Z}) + \mathbf{D}_2 \text{vec}^C(t\mathbf{Z}). \tag{H.2}
\end{aligned}$$

Examining terms in (H.1), they are either in the form of $\text{Re}\{\mathbf{v}_1^H d\mathbf{F} \mathbf{v}_2\}$ or $\text{Re}\{\text{tr}(\mathbf{Q}d\mathbf{F})\}$, where \mathbf{v}_1 and \mathbf{v}_2 are known vectors and \mathbf{Q} is a known matrix.

For the former case, we can transform it as

$$\begin{aligned}
\operatorname{Re}\{\mathbf{v}_1^H d\mathbf{F}\mathbf{v}_2\} &= \operatorname{Re}\{(\mathbf{v}_2^T \otimes \mathbf{v}_1^H)\operatorname{vec}(d\mathbf{F})\} \\
&= \operatorname{Re}\{(\mathbf{v}_2^T \otimes \mathbf{v}_1^H)[\mathbf{D}_1\operatorname{vec}(t\mathbf{Z}) + \mathbf{D}_2\operatorname{vec}^C(t\mathbf{Z})]\} \\
&= \operatorname{Re}\{[(\mathbf{v}_2^T \otimes \mathbf{v}_1^H)\mathbf{D}_1 + ((\mathbf{v}_2^T \otimes \mathbf{v}_1^H)\mathbf{D}_2)^C]\operatorname{vec}(t\mathbf{Z})\}. \quad (\text{H.3})
\end{aligned}$$

For the latter case,

$$\begin{aligned}
\operatorname{Re}\{\operatorname{tr}(\mathbf{Q}(d\mathbf{F}))\} &= \operatorname{Re}\{\operatorname{vec}^T(\mathbf{Q}^T)\operatorname{vec}(d\mathbf{F})\} \\
&= \operatorname{Re}\{\operatorname{vec}^T(\mathbf{Q}^T)[\mathbf{D}_1\operatorname{vec}(t\mathbf{Z}) + \mathbf{D}_2\operatorname{vec}^C(t\mathbf{Z})]\} \\
&= \operatorname{Re}\{[\operatorname{vec}^T(\mathbf{Q}^T)\mathbf{D}_1 + (\operatorname{vec}^T(\mathbf{Q}^T)\mathbf{D}_2)^C]\operatorname{vec}(t\mathbf{Z})\}. \quad (\text{H.4})
\end{aligned}$$

Since $\operatorname{Re}\{\operatorname{tr}(\mathbf{D}_v^H t\mathbf{Z})\} = \operatorname{Re}\{\operatorname{vec}^T(\mathbf{D}_v^C)\operatorname{vec}(t\mathbf{Z})\}$, according to (H.3) and (H.4), (H.1) can be represented in the form of

$$f(\mathbf{V}_s + t\mathbf{Z}) - f(\mathbf{V}_s) = \operatorname{Re}\{\mathbf{v}^T \operatorname{vec}(t\mathbf{Z})\} + O(t^2)$$

where \mathbf{v} is a vector. Obviously, \mathbf{v} is exactly equal to the vectorization of the conjugate of derivative \mathbf{D}_v . From (H.1), (H.2), (H.3) and (H.4), we obtain

$$\mathbf{D}_v = (2\Sigma\mathbf{V}_s\mathbf{M}^{-1})^H + RS_{d \times m}^H((\mathbf{A}_v\mathbf{D}_1)) + RS_{d \times m}^H(((\mathbf{A}_v\mathbf{D}_2)^C)) \quad (\text{H.5})$$

where

$$\begin{aligned}
\mathbf{A}_v &= \frac{1}{t} \frac{1}{(\xi - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x - \tau) - \hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x} \\
&\quad \{\hat{\mathbf{h}}_x^T \otimes \hat{\mathbf{h}}_x^H + [(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x]^T \otimes \hat{\mathbf{h}}_x + \\
&\quad [(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1} \mathbf{F}(\mathbf{V}_s) \hat{\mathbf{h}}_x]^T \otimes [(-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}] + \\
&\quad \hat{\mathbf{h}}_x^T \otimes [-\hat{\mathbf{h}}_x^H \mathbf{F}(\mathbf{V}_s) (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}] + \\
&\quad \mathbf{I} \otimes (-\mathbf{F}(\mathbf{V}_s) + \tau \mathbf{W}_x)^{-1}\}. \quad (\text{H.6})
\end{aligned}$$

Author's Publications

1. S.-M. Cai and Y. Gong, "Throughput maximization with statistical knowledge of cross channel state information," *IEEE Commun. Letters*, vol. 18, no. 11, pp. 2031-2034, Nov. 2014.
2. S.-M. Cai and Y. Gong, "Cognitive beamforming for multiple secondary data streams with individual SNR constraints," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4189-4198, Sept. 2013.
3. S.-M. Cai and Y. Gong, "Cognitive beamforming with unknown cross channel state information," in *Proc. IEEE Int. Conf. Communications (ICC) 2013*, June 2013, pp. 4931-4935.
4. S.-M. Cai and Y. Gong, "Cognitive beamforming exploiting Stiefel manifold for sum-transmit power minimization," in *Proc. IEEE Int. Conf. Information, Communications and Signal Processing (ICICS) 2011*, Dec. 2011, pp. 1-5.

Bibliography

- [1] FCC, “Spectrum policy task force report,” *ET Docket*, no. 02-155, Nov. 2002.
- [2] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.
- [3] J. Mitola III and G. Q. Maguire, “Cognitive radio: making software radios more personal,” *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13-18, Aug. 1999.
- [4] Y.-C. Liang, Y. Zeng; E.C.Y. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 1326-1337, Apr. 2008.
- [5] Z. Quan, S. Cui, and A. H. Sayed, “Optimal linear cooperation for spectrum sensing in cognitive radio networks,” *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 28-40, Feb. 2008.
- [6] H. Su and X. Zhang, “Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 118-129, Jan. 2008.
- [7] K. B. Letaief and W. Zhang, “Cooperative communications for cognitive radio networks,” *Proc. IEEE*, vol. 97, no. 5, pp. 878-893, May 2009.

- [8] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," in *Proc. IEEE*, vol. 97, no. 5, pp. 894-914, May 2009.
- [9] W.-Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845-3857, Oct. 2008.
- [10] S. Xie, Y. Liu, Y. Zhang, and R. Yu, "A parallel cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4079-4092, Oct. 2010.
- [11] J. G. Andrew, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, June 2014.
- [12] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 5, no. 2, pp. 122-130, Feb. 2014.
- [13] A. Goldsmith, "Wireless Communications," Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [14] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Applicat.*, vol. 20, no. 2, pp. 303-353, Oct. 1998.
- [15] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635-650, Mar. 2002.
- [16] H. Yu and V. K. N. Lau, "Rank-constrained Schur-convex optimization with multiple trace/log-det constraints," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 304-314, Jan. 2011.

- [17] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118-127, June 2014.
- [18] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile communication," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 36-43, May 2014.
- [19] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26-35, May 2015.
- [20] J. Mitola III, J. Guerci, J. Reed, Y.-D. Yao, Y. Chen, T. C. Clancy, J. Dwyer, H. Li, "Accelerating 5G QoE via public-private spectrum sharing," vol. 52, no. 5, pp. 77-85, May 2014.
- [21] X. Zhang, W. Cheng, and H. Zhang "Heterogeneous statistical QoS provisioning over 5G mobile wireless networks," *IEEE Network*, vol. 28, no. 6, pp. 46-53, Nov. 2014.
- [22] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128-1140, Mar. 2009.
- [23] R. Fan and H. Jiang, "Optimal multi-channel cooperative sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1128-1138, Mar. 2010.

- [24] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys and Tutorials*, vol. 11, no. 1, pp. 116-130, Mar. 2009.
- [25] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: a survey," *IEEE Commun. Surveys and Tutorials*, vol. 13, no. 3, pp. 372-395, Sept. 2011.
- [26] P. Setoodeh and S. Haykin, "Robust transmit power control for cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 915-939, May 2009.
- [27] J. Wang, G. Scutari, and D. P. Palomar, "Robust MIMO cognitive radio via game theory," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1183-1201, Mar. 2011.
- [28] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88-102, Feb. 2008.
- [29] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 38-51, Jan. 2008.
- [30] L. Zhang, Y.-C. Liang, Y. Xin and H. V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143-4153, Aug. 2009.
- [31] Y. Pei, Y.-C. Liang, K. C. Teh, and K. H. Li, "Secure communication in multiantenna cognitive radio networks with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1683-1693, Apr. 2011.

- [32] Y. Zhang, E. Dall'Anese, and G. B. Giannakis, "Distributed optimal beamformers for cognitive radios robust to channel uncertainties," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6495-6508, Dec. 2012
- [33] T. A.-Khasib, M. B. Shenouda, and L. Lampe, "Dynamic spectrum management for multiple-antenna cognitive radio systems: designs with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 2850-2859, Sept. 2011.
- [34] S. Srinivasa and S. A. Jafar, "The throughput potential of cognitive radio: a theoretical perspective," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 73-79, May 2007.
- [35] Q. Zhao and B.M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79-89, May 2007.
- [36] S. Sridharan and S. Vishwanath, "On the capacity of a class of MIMO cognitive radios," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 103-117, Feb. 2008.
- [37] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.
- [38] L. B. Le and E. Hossain, "Resource allocation for spectrum snderlay in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5306-5315, Dec. 2008.
- [39] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112-2120, Apr. 2009.

- [40] R. Zhang, S. Cui, and Y.-C. Liang, "On ergodic sum capacity of fading cognitive multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5161-5178, Nov. 2009.
- [41] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471-487, Feb. 2007.
- [42] L. Zhang, Y. Xin, and Y.-C. Liang, "Weighted sum rate optimization for cognitive radio MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2950-2959, June 2009.
- [43] S. Yiu, M. Vu, and V. Tarokh, "Interference and noise reduction by beamforming in cognitive networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3144-3153, Oct. 2009.
- [44] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust linear transceiver design in MIMO ad hoc cognitive radio networks with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1448-1457, May 2011.
- [45] F. Gao, R. Zhang, Y.-C. Liang, and X. Wang, "Design of learning-based MIMO cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1707-1720, May 2010.
- [46] H. Du, T. Ratnarajah, M. Pesavento, and C. B. Papadias, "Joint transceiver beamforming in MIMO cognitive radio network via second-order cone programming," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 781-792, Feb. 2012.
- [47] G. Scutari and D. P. Palomar, "MIMO Cognitive Radio: A Game Theoretical Approach," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 761-780, Feb. 2010.

- [48] S.-J. Kim and G. B. Giannakis, "Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117-3131, May 2011.
- [49] K. Lee, C.-B. Chae, J. R. W. Heath, and J. Kang, "MIMO transceiver designs for spatial sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3570-3576, Nov. 2011.
- [50] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and S. Cui, "On the relationship between the multi-antenna secrecy communications and cognitive radio communications," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1877-1886, June 2010.
- [51] R. Zhang, Y.-C. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks: A convex optimization perspective," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 102-114, May 2010.
- [52] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664-678, Feb. 2010.
- [53] S. W. Han, H. Kim, Y. Han, and J. M. Cioffi, "Efficient power allocation schemes for nonconvex sum-rate maximization on Gaussian cognitive MAC," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 753-757, Mar. 2010.
- [54] A. B. Gershman, N. D. Sidiropoulos, and S. Shahbazpanahi, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62-75, May 2010.
- [55] K. Huang and R. Zhang, "Cooperative feedback for multiantenna cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 747-758, Feb. 2011.

- [56] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 20-34, May, 2010.
- [57] S. Burer and R. D. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Mathematical Programming*, vol. 95, no. 2, pp. 329-357, Feb. 2003
- [58] Z. Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1426-1438, Aug. 2006.
- [59] Y. J. Zhang and A. M.-C. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 1683-1693, Feb. 2011.
- [60] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239-2251, June 2006.
- [61] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Vehicular Technology*, vol. 53, no. 1, pp. 18-28, Jan. 2004.
- [62] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve, "Linear processing for the downlink in multiuser MIMO systems with multiple data streams," in *Proc. IEEE Int. Conf. Commun.*, June 2006, vol. 9, pp. 4113-4118.
- [63] Y.-H. Yang, S.-C. Lin, and H.-J. Su, "Multiuser MIMO downlink beamforming design based on group maximum SINR filtering," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1746-1758, Apr. 2011.

- [64] J.-H. Noh and S.-J. Oh, "Beamforming in a multi-user cognitive radio system with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 616-625, Feb. 2013.
- [65] H. Islam, Y.-C. Liang, and A. T. Hoang, "Joint power control and beamforming for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2415-2419, July, 2008.
- [66] A. Tajer, N. Prasad, and X. Wang, "Beamforming and rate allocation in MISO cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 362-377, Jan. 2010.
- [67] Y. Shi, Y. T. Hou, S. Kompella, and H. D. Sherali, "Maximizing capacity in multihop cognitive radio networks under the SINR model," *IEEE Trans. Mobile Computing*, vol. 10, no. 7, pp. 954-967, July 2011.
- [68] M. F. Hanif, P. J. Smith, D. P. Taylor, and P. A. Martin, "MIMO cognitive radios with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3688-3699, Nov. 2011.
- [69] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871-4881, Dec. 2009.
- [70] H. A. Suraweera, P. J. Smith, and M. Shafi, "Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, May 2010.
- [71] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust downlink beamforming in multiuser MISO cognitive radio networks with imperfect channel-state information," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 2852-2860, Jul. 2010.

- [72] X. Kang, Y.-C. Liang, A. Nallanathan, H. K. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 940-950, Feb. 2009.
- [73] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [74] E. Polak, *Optimization Algorithms and Consistent Approximations*, New York: Springer-Verlag, 1997.
- [75] W. Kahan, "Jacobi's formula for the derivative of a determinant," [Online]. Available: <http://www.cs.berkeley.edu/~wkahan/MathH110/jacobi.pdf>
- [76] A. d'Aspremont and S. Boyd, "Relaxations and randomized methods for nonconvex QCQPs," [Online]. Available: <http://www.stanford.edu/class/ee392o/relaxations.pdf>
- [77] A. L. Fradkov and V. A. Yakubovich, "The S-procedure and a duality relations in nonconvex problems of quadratic programming," *Vestnik Leningrad Univ. Math.*, vol. 6, no. 1, pp. 101-109, 1979.
- [78] R. Zhang, F. Gao, and Y.-C. Liang, "Cognitive beamforming made practical: Effective interference channel and learning-throughput tradeoff," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 706-718, Feb. 2010.
- [79] T. R. Ramya and S. Bhashyam, "Using delayed feedback for antenna selection in MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6059-6067, Dec. 2009.
- [80] W. Xu and X. Dong, "Optimized one-way relaying strategy with outdated CSI quantization for spatial multiplexing," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4458-4464, Aug. 2012.

- [81] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multiaccess space-time block-coded MIMO systems: a probabilistically constrained approach," *IEEE J. Sel. Areas. Commun.*, vol. 24, no. 8, pp. 1560-1570, Aug. 2006.
- [82] S. Ye, R. S. Blum, and L. J. Cimini, "Adaptive OFDM systems with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3255-3265, Nov. 2006.
- [83] Z. Zhou and B. Vucetic, "Adaptive coded MIMO systems with near full multiplexing gain using outdated CSI," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 294-302, Jan. 2011.
- [84] V. A. Yakubovich, "S-procedure in nonlinear control theory," *Vestnik Leningrad Univ.*, vol. 4, no. 1, pp. 73-93, 1977.
- [85] C. G. Tsinos and K. Berberidis, "Blind opportunistic interference alignment in cognitive radio system," in *Prod. IEEE Int. Conf. Communications (ICC) 2013*, June 2013, pp. 626-639.
- [86] H. Solomon, "Distribution of quadratic forms-tables and application," *Applied Mathematics and Statistics Laboratories*, Technical Report 45, Stanford University, 1960.
- [87] N. L. Johnson and S. Kotz, "Tables of distributions of quadratic forms in central normal variable," *Institute of Statistics Mimeo. Series No. 543*, University of North Carolina, 1967.
- [88] N. L. Johnson and S. Kotz, "Tables of distributions of quadratic forms in central normal variable, II," *Institute of Statistics Mimeo. Series No. 557*, University of North Carolina, 1967.
- [89] Y. Huang, A. De Maio, and S. Zhang, "Semidefinite programming, matrix decomposition, and radar code design," in *Convex Optimization in*

- Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds., pp. 192-228. Cambridge Univ. Press, New York, 2010.
- [90] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.
- [91] A. E. Ozdaglar and D. P. Bertsekas, "The relation between pseudonormality and quasiregularity in constrained optimization," [Online]. Available: <http://web.mit.edu/dimitrib/www/Quasiregularity.pdf>.
- [92] J. B. Lasserre, "A trace inequality for matrix product," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1500-1501, Aug. 1995.
- [93] R. A. Horn and C. A. Johnson. *Matrix Analysis*, Cambridge, U.K.: Cambridge Univ. Press, 1985.