

**TRANSMIT PRECODING AND BAYESIAN DETECTION
FOR COGNITIVE RADIO NETWORKS
WITH LIMITED CHANNEL STATE INFORMATION**

A Dissertation presented to
the Faculty of the Graduate School
at the University of Missouri

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
MOHANNAD HAMED AL-ALI
Dr. Dominic K. C. Ho, Dissertation Supervisor

MAY 2017

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

TRANSMIT PRECODING AND BAYESIAN DETECTION
FOR COGNITIVE RADIO NETWORKS
WITH LIMITED CHANNEL STATE INFORMATION

presented by Mohannad Hamed Al-Ali,
a candidate for the degree of Doctor of Philosophy and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Dominic K. C. Ho

Dr. Naz Islam

Dr. Justin Legarsky

Dr. Athanasios C. Micheas

Dedicated to

*The awaited Saviour,
Mohammad Ibn Al-Hasan Al-Mahdi*

ACKNOWLEDGMENTS

First of all, I would like to thank my creator, my Lord, Almighty Allah for his blessings and benefactions. Without his everlasting support and custody, I would never complete my PhD degree successfully. O' my Lord, I cannot achieve thanksgiving as my thanking You lacks thanksgiving!

I would like to express my gratitude toward my sponsor, the Higher Committee for Education Development (HCED) in Iraq. HCED has granted me a more-than-five-years scholarship to conduct my PhD study in US, and provided me consistent fund despite economical, political and military crises there in Iraq. Thanks HCED for making my dream of studying abroad come true!

I would like to express my sincere gratitude and appreciation to my friend, research advisor and mentor Prof. Dominic Ho. Working with Dr. Ho has proven to be one of the best things I have ever made. Dr. Ho has believed in me and considered my success (or fail) is his success (or fail). He has helped me to strive harder to develop better ideas and to become an outstanding researcher. He has always been available to sit with me and listen to my academic and personal issues. Being under his supervision will be one of the things that I shall always be proud of. Thanks Dr. Ho!

I would like to thank Prof. Naz Islam, Prof. Justing Legarsky, and Prof. Athanasios Micheas for serving in my PhD committee. Special thanks go to the Electrical and Computer Engineering department and the nice faculty and staff there who make my PhD studying at MU fruitful and delightful. Thanks Mizzou!

During my staying at MU, I have come across a lot of colleagues and met real

friends who have provided me support and created a very warm and pleasant community around me. I am very indebted to my amazing friends in Kumal's and Imam Hasan's groups who have shared with me my happiness, sadness and frequent occasions. Participation in their weekly gathering has relieved my studying stress and indeed sharpened my personal traits. Big thank goes to them!

Finally, the words cannot ever express my feelings toward my big family: father, mother, brothers and sisters, and my small family: my wife and kids. Their love, care, understanding, patience, and unlimited support have made my doctoral course of study much easier and successful. Special thank goes to my life partner, Amal, who walks down beside me in my tough road!

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	ix
LIST OF FIGURES	x
ABSTRACT	xiv
CHAPTER	
1 Introduction	1
1.1 Notation and Acronyms	1
1.2 Cognitive Radio Network	4
1.3 Robust Precoding for Underlay CR	7
1.3.1 CSI Acquisition	7
1.3.2 CSI Uncertainty Model	9
1.3.3 Interference Constraints	10
1.3.4 Robust Precoder Design	11
1.4 Spectrum Sensing	13
1.5 Research Contribution	14
1.6 Dissertation Organization	17
2 Transmit Precoding in Underlay MIMO CR with Unavailable or Imperfect Knowledge of Primary Interference Channel	18
2.1 Introduction	18
2.1.1 Background	19

2.1.2	Chapter Organization	21
2.2	System Model and Problem Formulation	21
2.2.1	System Model	21
2.2.2	Design Objective	23
2.2.3	Interference CSI	25
2.3	Unavailable CSI scenario	26
2.3.1	Proposed Solution	26
2.3.2	Solution Analysis	30
2.4	Inaccurate CSI	31
2.4.1	Modelling the Uncertainty Set	32
2.4.2	Robust Solution for IT Metric	34
2.4.3	Robust Solution Based Linearization for LR Metric	36
2.4.4	Robust Solution at Low INR for LR Metric	39
2.5	Selection Between Unavailable and Inaccurate CSI Solutions	39
2.5.1	Condition for the Rate of SU	40
2.5.2	Condition for the Power of SU	42
2.6	Numerical Results	43
2.6.1	Results for IT and LR Metrics	43
2.6.2	Results for LR Metric	52
2.7	Conclusion	55
2.8	Appendix	56
2.8.1	Optimal Solution of P1 with Perfect CSI	56

2.8.2	Proof of Proposition 1	58
2.8.3	Proof of Proposition 2	59
2.8.4	Proof of Proposition 3	60
2.8.5	Waterfilling Solutions	63
3	Precoding for MIMO Channels in Cognitive Radio Networks With CSI Uncertainties and for Compound Capacity	64
3.0.1	Introduction	64
3.0.2	Background	65
3.0.3	Chapter Organization	67
3.1	System Model and Assumptions	67
3.2	Cognitive Radio Problem	71
3.2.1	Solution for Δ_{21}	71
3.2.2	Solution for Δ_{22}	74
3.2.3	Solution for \mathbf{Q}	78
3.3	Compound Capacity Problem	83
3.4	Particular Solutions	84
3.4.1	Particular Solution for P-I	84
3.4.2	Particular Solution for P-II	86
3.5	Numerical Results	88
3.6	Conclusion	93
3.7	Appendix	95
3.7.1	Upper Bound for MI	95
3.7.2	Proof of Proposition 1	97

3.7.3	Proof of Theorem 1	97
3.7.4	Evaluation of the Matrix \mathbf{F}^*	100
3.7.5	Convergence Analysis	101
3.7.6	Proof of Proposition 2	103
3.7.7	Proof of Theorem 2	104
4	Bayesian Multi-antenna Sensing in Cognitive Radio Networks Using Fractional Bayes Factor	106
4.1	Introduction	106
4.1.1	Background	107
4.1.2	Chapter Organization	109
4.2	Problem Formulation	109
4.2.1	Problem Setup	109
4.2.2	Bayesian Detection	111
4.3	Proposed FBF Detector	112
4.3.1	FBF	112
4.3.2	Prior Distribution under H_0	113
4.3.3	Prior Distributions under H_1	114
4.3.4	Bayes Factor Evaluation	116
4.4	Numerical Results	117
4.5	Conclusion	118
4.6	Appendix	119
4.6.1	Bayes Factor for the Conjugate Prior	119
4.6.2	GLRT Formula	119

5	Summary and Future Work	121
5.1	Research Summary	121
5.2	Future Work	123
5.2.1	Stochastic Modeling	124
5.2.2	Bayesian Model Selection	125
5.2.3	Game Theory Based CR Design	125
	BIBLIOGRAPHY	126
	VITA	142

LIST OF TABLES

Table		Page
2.1	The average number of iterations for Algorithm 2 to converge at a certain leakage rate limit R_L and rate resolution θ	49
2.2	The actual value $\hat{\epsilon}_R$, the sufficient value ϵ_R^* and its percent deviation at different levels of the noise power σ_s^2	51
2.3	The value ϵ_P^* at different levels of the noise power σ_s^2	52

LIST OF FIGURES

Figure	Page
1.1 Schematic diagram for underlay CR network.	6
1.2 An example for spectrum holes or white space in different dimensions.	7
2.1 An underlay CR user (SU) coexists with a single PU where the transmitter and receiver of each user have multiple antennas. The notations of the various channels are shown.	23
2.2 Rates for SU from the proposed precoder, the one from [71] for unavailable interference CSI and a lower bound versus the interference limit (I_T), $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are obtained from 100 realizations of \mathbf{H}_s and \mathbf{R}_G	45
2.3 SU rate reduction factor versus the Doppler frequency of the time-varying channel \mathbf{G}_s at different levels of the interference limit I_T , $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are the average over 100 realizations of \mathbf{G}_s at a given \mathbf{H}_s	45
2.4 Rates for SU versus I_T for the IT metric, $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are generated by the average of 100 realizations for \mathbf{H}_s and $\Delta\mathbf{R}_G$ with $w = 0.1$ at a given \mathbf{G}_s	47

2.5	Rates for SU versus R_L for the LR metric, $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are generated by the average of 100 realizations for \mathbf{H}_s and $\Delta\mathbf{R}_G$ with $w = 0.1$ at a randomly generated \mathbf{G}_s	49
2.6	Leakage rates (defined below (2.4)) for PU versus the achievable rate for SU for the IT and LR metrics, $M_s = N_s = 4$ and $M_p = N_p = 4$. Results shown are the average of 100 realizations of \mathbf{H}_s and $\Delta\mathbf{R}_G$ at a randomly generated \mathbf{G}_s	50
2.7	Rates for SU versus the normalized uncertainty radius w , $M_s = M_p = 5$ and $N_s = N_p = 4$. The results shown are the averages over 100 realizations of \mathbf{H}_s and $\Delta\mathbf{R}_G$ at a randomly generated \mathbf{G}_s	50
2.8	Power of SU versus the normalized uncertainty radius w , $M_s = N_s = 5$ and $M_p = N_p = 4$. The results shown are the averages of 100 realizations for \mathbf{H}_s and $\Delta\mathbf{R}_G$ at a randomly generated \mathbf{G}_s	51
2.9	Rates for SU from the design using low INR approximation versus the leakage rate limit R_L . The results for imperfect interference CSI at $p = 1, 2$, and ∞ are shown. The settings of antennas are $M_s = N_s = 3$ and $M_p = N_p = 6$. The results are generated by the average of 1000 realizations for \mathbf{H}_s and \mathbf{G}_s for different values of w	53
2.10	Rates for SU from the iterative linearization solution for different leakage rate limit R_L and amount of interference CSI uncertainty w . The results are generated by the average of 1000 realizations for \mathbf{H}_s and \mathbf{G}_s . The settings for antennas are $M_s = N_s = 3$ and $M_p = N_p = 6$. . .	54

2.11	Rates for SU from the iterative linearization solution for different uncertainty levels $w = \epsilon/\ \mathbf{R}_G\ _{S\infty}$ and leakage rate limit R_L . The results are generated by the average of 1000 realizations for \mathbf{H}_s and \mathbf{G}_s . The settings for antennas are $M_s = N_s = 3$ and $M_p = N_p = 6$	55
3.1	An underlay MIMO CR network with single SU and single PU. CSI for the SU direct link and for the interference link from SU to PU is not accurate.	69
3.2	Rate for SU from the proposed CR solution and from [2] for different values of r_L (nats/s/Hz), SNR = 5dB, $M_2 = N_2 = N_1 = 4$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$ and $\hat{\mathbf{R}}_{21}$	91
3.3	Rate for SU from the proposed CR solution versus the uncertainty error w_1 , $r_L = 3$ (nats/s/Hz), $w_2 = 0.5$, SNR = 5dB, $M_2 = N_2 = N_1 = 2$ or 4. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$ and $\hat{\mathbf{R}}_{21}$	92
3.4	Rate for SU from the proposed CC solution, the beamforming and the equal power solutions versus SNR (dB), $M_2 = N_2 = 4$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$	93
3.5	Rate for SU from the proposed CC solution, the beamforming and the equal power solutions at different values of w_2 , $M_2 = N_2 = 4$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$	94
3.6	Rate for SU from the particular CR solution versus r_L , $M_2 = N_2 = N_1 = 5$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$ and $\hat{\mathbf{R}}_{21}$	95

3.7	Rate for SU from the particular CC solutions versus SNR (dB), $M_2 = 6, N_2 = 3$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$	96
4.1	An interweave CR model between PU and SU, where each user has multiple antennas and the channel \mathbf{H} is not known.	112
4.2	Performance of the proposed FBF, the conjugate prior [54] and the GLRT test [87] detector for PU active; the settings are $l = 4, p = 4, P_{\text{FA}} = 10^{-3}$ and $b = 0.1$	120
4.3	Comparison of ROC curves from different detectors for detecting PU active; the settings are $l = 5, p = 5, \text{SNR} = -4$ dB and $b = 0.12$	120

ABSTRACT

Cognitive radio (CR) represents a recent direction for enabling coexistence among heterogeneous networks. It can be a potential solution for the problem of scarce spectrum available for wireless communication systems. The study here investigates the underlay and interweave paradigms for the coexistence of CR network of secondary users (SUs) with a primary network of primary users (PUs). Under underlay mode, both networks communicate concurrently using the same resources. With interweave, SU is able to communicate as long as (some) PUs are not active. Usually, underlay or interweave employs multiple antennas at SU to use the spectral resources better and manage the interference towards the primary network. Performance of the CR network under either paradigm depends largely on the amount and quality of channel state information (CSI) available about the different communication links. In practical systems, often CSI at SU has uncertainty since it is deviated from the true one or is not known at all. This uncertainty should be accounted when designing the precoding schemes for SU or otherwise the interference impact on primary networks would violate the quality of service (QoS) requirements for PUs. This dissertation considers two cases regarding to the availability of CSI, the first one is when CSI is imperfect and the second is when CSI is completely not known.

For the underlay mode, we investigate two manifolds. The first one addresses the problem of maximizing the throughput of a multiple-input multiple-output (MIMO) SU when CSI of the interference link to PU is completely unknown or partially known. We study the achievable rates for SU under two different QoS requirements for the PU: the conventional interference temperature and leakage rate metrics. When CSI

is unavailable, we develop an iterative adaptation algorithm that satisfies the QoS constraint through exploiting the side-information in the primary communication network. When CSI is inaccurate, we model the uncertainty deterministically such that the uncertainty error belongs to a convex compact set defined by the Schatten norm. We design the precoder by following the worst case formulation. We further investigate the relation between the unknown and the inaccurate CSI cases when using the interference temperature metric, and reveal that the performance of the latter is not necessarily better than the former.

The second manifold assumes there is uncertainty in the SU intended link for communication as well as in the interference link from SU to PU. Similar to the first manifold, we follow the deterministic modelling using Schatten norm for the uncertainty and apply the worst case philosophy. For a given precoder matrix, we find the worst uncertainty error in the set that describes the uncertainty in each link. We further develop an iterative numerical algorithm for the precoder. Simpler solutions for the precoder and the uncertainty errors are derived under some special instances of the Schatten norm and certain requirement of transmission power.

For the interweave mode, we assume there is no CSI available at SU and derive a Bayesian detector for the proposed binary hypothesis problem. For the null or noise model, we propose a conjugate prior for the unknown spatial covariance matrix. For the alternative or data model, we propose a new class of improper priors for the covariance matrix. We introduce the fractional Bayes factor (FBF) approach to enhance the detection capability of the Bayes factor. The developed FBF is compared with those using the conjugate priors for both hypotheses and generalized likelihood ratio test (GLRT), and it yields significant improvement.

Chapter 1

Introduction

1.1 Notation and Acronyms

Throughout the dissertation, bold upper-case letters denote matrices and bold lower-case letters represent column vectors. Further notational symbols are

$\mathbb{C}^{m \times n}, \mathbb{H}^n$	The complex space of $m \times n$ matrices and the space of Hermitian $n \times n$ matrices
$\mathbb{R}^{m \times n}, \mathbb{R}_+^{m \times n}$	The real space of $m \times n$ matrices and of $m \times n$ matrices with non-negative entries
$(\cdot)^\dagger$	The Hermitian transpose
$(\cdot)^t$	The regular transpose
$ \mathbf{A} $	Determinant of the square matrix \mathbf{A}
$\text{Tr}(\mathbf{A})$	Trace of the matrix \mathbf{A}
$[\mathbf{A}]_{ij}$	The ij -th element of \mathbf{A}

$\lambda_i(\mathbf{A})$	The i th eigenvalue of \mathbf{A} such that $\lambda_1(\mathbf{A})$ denotes the maximum eigenvalue
$\mathbf{A} \succeq \mathbf{B}, \mathbf{A} \succ \mathbf{B}$	The matrix $\mathbf{A} - \mathbf{B}$ is positive semi-definite or positive-definite
$\mathbf{A} \preceq \mathbf{B}$	The matrix $\mathbf{A} - \mathbf{B}$ is negative semi-definite
$\text{diag}(\mathbf{a})$	Diagonal matrix formed by the elements of the vector \mathbf{a}
$\text{Diag}(\mathbf{A})$	Forms a diagonal matrix from the diagonal elements of \mathbf{A}
$\mathbf{U}_{\mathbf{A}} \mathbf{\Lambda}_{\mathbf{A}} \mathbf{U}_{\mathbf{A}}^\dagger$	Eigenvalue decomposition, $\mathbf{U}_{\mathbf{A}} \in \mathbb{C}^{n \times n}$ is a unitary matrix and $\mathbf{\Lambda}_{\mathbf{A}} = \text{diag}([\lambda_1(\mathbf{A}) \cdots \lambda_n(\mathbf{A})]^t)$ with the eigenvalues arranged in non-increasing order
$\text{vec}(\mathbf{A})$	Stacks the columns of \mathbf{A} into a single column
$\ \mathbf{a}\ , \ \mathbf{a}\ _{\mathbf{X}}$	The Euclidean norm $\sqrt{\mathbf{a}^\dagger \mathbf{a}}$ and the weighted version $\sqrt{\mathbf{a}^\dagger \mathbf{X} \mathbf{a}}$ for a Hermitian and positive semi-definite matrix \mathbf{X} of an appropriate size
$\mathbb{E}[\cdot]$	The statistical expectation
$\ A\ _p$	Schatten norm with order p
$ a $	The absolute value of the scalar a
$\log(\cdot)$	The natural logarithm
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	The distribution of a circularly symmetric complex Gaussian random vector with mean \mathbf{m} and covariance \mathbf{C}
\mathbf{I}	The identity matrix of an appropriate size
a^+	The maximum between a and 0
$\lceil a \rceil$	denotes the least integer $\geq a$

Acronyms

Some of the abbreviations and acronyms we used in the dissertation are summarized as follows:

ACK/NAK	ACKnowledgment and Negative AcKnowledgegment signaling protocol
AI	Alternate-Iterate optimization
AI-FDPS	Alternate-Iterate Feasible Direction Projected Subgradient
CC	Compound Capacity
CDMA	Code-Division Multiple Access
CR	Cognitive Radio
CSCG	Circularly Symmetric Complex Gaussian distribution
CSI	Channel State Information
DSA	Dynamic Spectrum Access
EVD	Eigenvalue Decomposition
FBF	Fractional Bayes Factor
FDD	Frequency-Division Duplexing
GLRT	Generalized Likelihood Ratio Test
HSDPA	High Speed Downlink Packet Access communication system
IT	Interference Temperature
INR	Interference-to-Noise Ratio
LMI	Linear Matrix Inequality
LR	Leakage Rate
LTE-A	Long Term Evolution-Advanced communication system
MI	Mutual Information

MIMO	Multiple-Input Multiple-Output
NMSPG	Non-Monotone Spectral Projected Gradient
OFDM	Orthogonal Frequency-Division Multiplexing
PU	Primary User
QoS	Quality of Service
ROC	Receiver Operating Characteristic
SDP	Semi-Definite Programming
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
SU	Secondary User
TDD	Time-Division Duplexing

1.2 Cognitive Radio Network

The increasing demand for reliable high-speed data services and the scarce of wireless spectrum have created many challenges to meet the quality of service (QoS) requirements of the end users. One enthused approach is to allow several heterogeneous networks to operate together while using dynamic spectrum access (DSA) techniques to minimize their undesirable interactions [32, 94, 100]. Unlike incumbent users of legacy communication systems, users that employ DSA require some sort of cognition on how to use the spectrum in an efficient and adaptive manner. Those users are usually called cognitive radio (CR) users and formally defined as reconfigurable

devices via software that can operate over multiple air interfaces and communication protocols [32, 46].

DSA provides a method to use spectrum through hierarchical spectrum access [94]. This access technique considers two (or more) disparate networks such that one of them has priority over the other. Users of the network with high precedence is known as primary users (PUs) and of the other network as the secondary users (SUs). SU is capable of using the spectrum provided that the amount of interference induced on PU is under an allowable limit. The availability of channel state information (CSI) at SU plays a dominant role in determining how and when to access or to release the physical channels of PUs. Based on the amount of CSI, there are two main paradigms for spectrum sharing under the hierarchical access: The underlay paradigm and the interweave paradigm [32].

The underlay paradigm provides a concept for concurrent operations of SU with PU. SU can exist as long as the interference introduced to the PU is below an acceptable limit [32, 46, 100, 110]. Figure 1.1 depicts the concept of underlay or concurrent spectrum access. The underlay CR requires CSI about the interference link from the transmitter of SU to the receiver of PU.

We should point out that the underlay concept can be extended to the case where PU and SU belong to the same network. For example in cellular network, we can consider the femtocell as SU and the macrocell as PU where both radio cells share the same frequency band [98, 101].

The interweave paradigm offers an opportunistic communication for SU through exploiting the spectrum holes that are not occupied by the PU. Figure 1.2 shows an example for the spectrum holes in the spectral resources. This paradigm is feasible

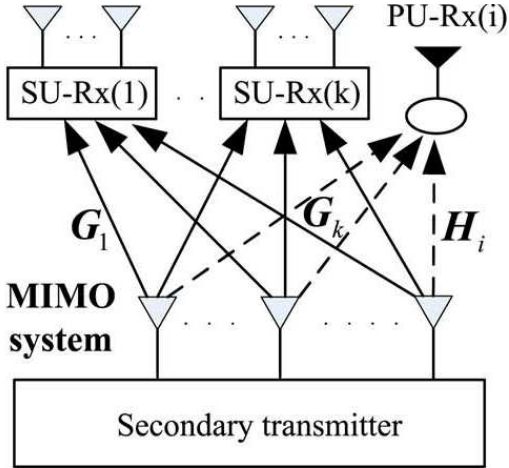


Figure 1.1: Schematic diagram for underlay CR network.

as there exists a significant amount of spectrum that is used infrequently by PU. In this paradigm, CSI about the activity of PU of whether it is utilizing the spectrum is crucial. The SU should use advanced signal processing algorithms to be able to sense, track and release any spectrum holes. Please refer to [6, 114] for a recent review of the spectrum sensing algorithms in CR networks.

Utilizing multiple antennas provides significant enhancement for the spectrum sharing in CR networks. In particular, using multi-antenna at SU in the interweave mode can obtain reliable spectrum sensing through exploiting the spatial domain. Moreover, the multiple antennas can reduce CSI demands as SU may not require prior knowledge about the PU signals or the interference channel from PU to SU [6, 84, 90]. In underlay CR, multiple antenna offers the flexibility of precoding, which can substantially reduce the interference from SU to PU and maintain QoS for PU [32, 77, 88]. Back to Figure 1.1, multiple-input multiple-output (MIMO) SU and MIMO PU users are illustrated. The precoding benefit offered by MIMO antennas in an underlay CR network relies on the availability of CSI for the connections from

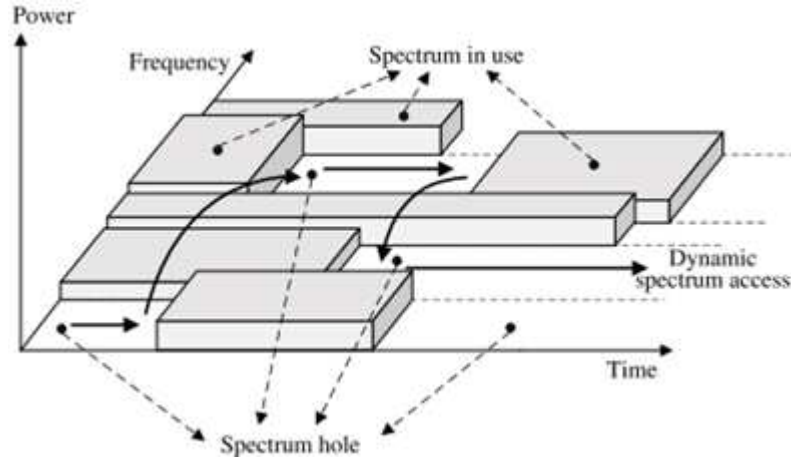


Figure 1.2: An example for spectrum holes or white space in different dimensions.

SU-Transmit (SU-Tx) to SU-Receive (SU-Rx) and from SU-Tx to PU-Receive (PU-Rx).

In this dissertation, we shall address the precoding design for underlay MIMO CR networks under the assumption that there is a limited knowledge about CSI at SU. We shall also investigate the spectrum sensing problem assuming there is no CSI about PU available.

1.3 Robust Precoding for Underlay CR

1.3.1 CSI Acquisition

CSI of the link SU-Tx to SU-Rx can be obtained at SU-Tx through feedback [47]. Specifically, SU-Tx can send training or pilot signals, which can be used by SU-Rx to estimate the corresponding channel, and then feedback the estimated channel to SU-Tx.

CSI of the interference link from SU-Tx to PU-Rx can be made available at the SU-Tx through adopting different approaches. When PU follows frequency-division duplexing (FDD) transmission technique and there is cooperation between PU and SU, CSI of the interference link is obtained through feedback. However, in practice the two users are seldom cooperated. An alternative approach is to use blind null space learning for slow fading channels whenever the PU adapts its transmitted power as proposed in [63]. When PU follows time-division duplexing (TDD), SU can obtain CSI by exploiting reciprocity and listening to the receiver of PU [72]. Otherwise, blind techniques can be invoked to obtain the interference channel [90, 112]. It is worth to mention that current communication systems such as LTE-A, CDMA, and HSDPA, and WiFi networks usually exchange some control signals to maintain the quality of the link at a certain level [85, 106]. These signals can be exploited by SU to extract some related CSI to design the precoding scheme in CR network [63, 91].

CSI can be estimated instantaneously so long as the time for estimating CSI is small compared to the coherence time of the channel [39, 69, 80, 97, 101], see also [22, 51, 72, 99, 105]. Otherwise, statistical representations for CSI can be used especially for fast fading propagation environments, where the channel information is averaged over sufficiently long time interval [9, 43, 65, 97, 98].

Regardless of how to make the CSI available at SU-Tx, the obtained CSI is often far from perfect in practice [28, 55, 63, 112]. Channel estimation errors, limited and outdated feedback, and short coherence time of the physical channels are among the main factors that cause imperfect CSI. In some cases, the SU may not have any information about the interference channel or its distribution; making the precoding design problem more challenging [71].

This dissertation considers CSI of the direct link from SU-Tx to SU-Rx is inaccurate, and of the interference link from SU-Tx to PU-Rx is inaccurate or unknown. The assumption that the interference link has some channel errors is typical in underly CR, while relatively very few studies in the literature address the unknown CSI situation, see for example [71]. The consideration that the direct link can encounter CSI errors has been recently adopted for multiuser scheduling in MIMO CR network with space-time block coding in [111]. We should mention that our treatments are valid for instantaneous or statistical CSI.

1.3.2 CSI Uncertainty Model

The uncertainty in CSI can be taken into account to improve the precoder design. It can be modelled by following a deterministic or stochastic representation. The former assumes the error is within some uncertainty set whose shape and size are known. The design seeks the precoder that meets the stringent QoS of PU at the worst interference scenario over all CSI possibilities within the uncertainty set. The stochastic model characterizes the uncertainty error through a specific statistical distribution and the precoder is obtained based on the outage probability [101,110]. However, for analysis tractability, Gaussian or exponential distribution is assumed to model the uncertainty error, which may not resemble the true distribution. In this work, we shall follow the deterministic modelling for CSI uncertainty.

In the literature, there are different deterministic models for the uncertainty error. Particularly, the polyhedron based set is used to model the quantization error due to limited feedback as in [70]. Ellipsoid, Frobenius or spectral norm based set is usually used to model the errors due to the additive noise such as in [8,51,69,102,105] or [25,

36, 78, 89, 93, 95] . The authors in [35, 44] bounded the Kullback-Leibler divergence to model the uncertainty in probability distribution. Recently, uncertainty sets defined by unitarily-invariant and Schatten norm are used for the precoding design in [102]. For other uncertainty models that appear in CR literature, we refer the interested reader to [110].

The focus here is to provide a unified solution for the MIMO precoding design in the underlay CR network under different uncertainty sets. We shall represent the uncertainty errors associated with the SU-Tx to SU-Rx and SU-Tx to PU-Rx links using some generic matrix norm. In particular, we shall follow [102] and use the Schatten norm.

1.3.3 Interference Constraints

There are several types of interference constraints that have been proposed to satisfy the QoS for PU. [94] provides a nice discussion on how to choose the interference constraints in underly CR networks. Interference temperature (IT) is the conventional metric for interference measure at the receiver of PU, and it quantifies the interference power at PU-Rx. The IT metric is linear in the precoder matrix and mathematically tractable, and hence it has been used extensively in the precoding design of CR [8, 22, 36, 51, 72, 93, 105]. This metric requires CSI of the interference link from SU-Tx to PU-Rx only.

In [38], a rate loss constraint (RLC) that is related to the maximum rate loss of PU was imposed on SU. The RLC metric achieves a higher information rate for SU than that using the IT metric. However, this metric requires CSI for the interference link from SU-Tx to PU-Rx and for the PU links. The work in [23] confined the

interference amount through proposing a precoder design that satisfies a constraint on the PU information rate. Similar to the RLC metric, this design methodology assumes CSI for PU links is readily available at the SU. The authors in [71] suggested that the mutual information measure of the interference link from SU-Tx to PU-Rx is a better metric than the corresponding power measure, i.e., IT metric.

This dissertation adopts the metric suggested in [71] as well as the traditional IT metric to limit the interference on PU-Rx.

1.3.4 Robust Precoder Design

The approach for the MIMO precoder design under the worst case scenario of deterministic CSI uncertainty is mostly based on robust beamforming technique developed in the signal processing community [9, 97]. For point-to-point MIMO channels, the worst case approach was applied to solve the compound capacity problem and obtain a robust linear precoder [48, 49, 65, 82, 104]. The compound capacity is a suitable information rate metric when the transmitter observes inaccurate CSI knowledge and the CSI error lies in a deterministic set. It seeks the optimal precoder that maximizes the worst case mutual information (MI). When the error set of MIMO channels is isotropic, i.e., eigenvalue constraint based uncertainty set, [65] showed that under total power budget at the transmitter the compound capacity is achieved using uniform power allocation. In [104], the compound capacity of a nominal rank one Ricean MIMO channel with an ellipsoidal uncertainty set and total power constraint was analyzed. The works in [49, 82] considered an unitarily-invariant power set and a spectral norm defined channel uncertainty set., They unveiled through proposing new matrix inequalities that the compound capacity has an optimal channel-diagonalizing solu-

tion. That is, the precoder design complexity reduces from an optimization problem with matrix variables to a scalar power allocation problem.

The worst case robust beamforming design for CR network was used in [8,36,43,98] to minimize the transmit power and limit the signal-to-interference-plus-noise ratio (SINR) of the SU under interference power constraint. The works in [89,95] consider a robust beamforming design through optimizing SINR and the information rate metrics, respectively, with interference power constraints. In [93], the robust design for MIMO CR nodes is performed through minimizing the sum mean square error of SU with interference power constraints.

Due to the intractability caused by placing the uncertainty set into the optimization problem, most studies use approximations or relaxation along with semi-definite programming (SDP) to obtain a solution [8, 36, 39, 72]. Relatively few works focus on producing an equivalent and yet convex formulation of the original problem to yield more accurate and less conservative solutions. Among them, [51, 105] used the S-procedure and [43, 95, 98] utilized the Lagrange duality to obtain the equivalent formulations. However, they have limited applicability where the former is for the Euclidean norm and the latter the Frobenius and trace norms uncertainty sets.

We shall follow the trend and employ the worst cast robust optimization to obtain the precoder matrix that would optimize the worst case performance. We shall consider an information theoretic criterion in terms of the MI to define the performance metric.

1.4 Spectrum Sensing

Spectrum sensing for CR networks is a binary hypothesis testing problem that looks for distinguishing reliably between the null (or noise) hypothesis and the alternative (or data plus noise) hypothesis. Once the detection problem is solved, SU becomes aware of the PU activity and can access the physical channels when they are available.

It is well-known that according to the Neyman–Pearson lemma that the likelihood ratio test is optimal when the probability distribution under each hypothesis is known without uncertainty. Usually in practice, the distribution is not known perfectly due to the lack of knowledge about some of its parameters. There are two standard approaches to take care of the unknown parameters in hypothesis testing: Frequentist and Bayesian frameworks. The former method estimates the unknown parameters from the likelihood and plug back estimated quantities in the probability distributions, which would yield the generalized likelihood ratio test (GLRT). The Bayesian method estimates the unknown parameters by introducing prior distributions for them and then uses the posterior distributions to perform inference.

We can classify the developed detectors in the interweave multiantenna CR literature into deterministic and Bayesian detectors. For the first category, the test statistic can be based on energy [41], multivariate cyclostationary [73], eigenvalues of the sample covariance matrix [45, 79, 86, 87], or GLRT [16, 50, 90]. Bayesian detectors has be proposed in [5, 21, 40, 54]. In this work, we shall regard the Bayesian detection for the spectrum sensing.

1.5 Research Contribution

For the underlay CR paradigm, we have considered two scenarios: CSI for the SU-Tx to PU-Rx link is not accurate, and CSI for SU-Tx to SU-Rx and SU-Tx to PU-Rx links is not accurate. The contributions for the first scenario are as follows

- A precoder design without requiring CSI. The proposed design approach obtains the maximum SU rate while maintaining the QoS for PU without coordination between PU and SU and without channel estimation. The design exploits channel side-information usually broadcasted from PU to optimize the precoder. We show that this precoder is asymptotically optimal as the number of antennas at PU-Rx goes to infinity;
- A robust precoder design for SU with imprecise CSI modelled by the Schatten norm deterministic uncertainty. The previous techniques such as the S-lemma and differentiable Lagrange function are unable to handle the Schatten norm uncertainty. We propose a new approach based on the Lagrange dual and Hölder inequality to obtain a precoder using the IT metric for QoS. For the leakage rate (LR) metric, we develop an iterative linearization method for the design that ensures global convergence;
- The conditions on the amount of CSI uncertainty under which the robust design of imprecise CSI will not perform better than the one without CSI, under the IT metric.

The contributions for the second scenario are

- New algorithms for solving the CR problem, including the CC problem as well. Most of the related works from the literature in robust optimization are for

unconstrained maximin problems. A recent method took the interior-point methods with semi-infinite programming iterative steps and subgradient descent directions to address the constrained minimax problem. However, the iterative algorithm is applicable for equality constrained minimax problems and requires the functions in the equality constraints twice differentiable. Furthermore, it assumes there is sufficient number of solutions at each iteration for the inner non-concave maximization to obtain a descent direction;

- Most studies in the literature uses a specific norm, such as the spectral norm, to define the CSI uncertainty set for simplifying the solution finding, and consequently yielding a limited result to the problem. The work here uses the general matrix norm called the Schatten norm that comprises a number of frequently used matrix norms such as the spectral and Frobenious norms. We also do not restrict the transmit power constraint to be on the total power. As a result the proposed solution is more general and applicable to wider situations;
- Through the Lagrange dual and the Hölder inequality, we obtain a suboptimal solution for the worst case direct link CSI uncertainty in the alternate-iterate optimization process. This solution is more attractive than solving the constrained problem directly using the interior point methods, while maintaining comparable performance from our observations;
- We derive the optimal structure of the CR precoder matrix when the uncertainty sets are defined by the spectral norm and the transmit power requirement is always fulfilled. Furthermore, we prove for the CC problem that the optimal precoder has eigen-directions equal to those of the available CSI matrix for the

uncertainty set defined by the Schatten norm, when the power set is unitarily-invariant;

For the interweave CR paradigm, we assume that SU does not know CSI about PU or noise signals other than their probability distribution. Our contribution for this research problem is to use the fractional Bayes factor approach to produce a well-defined Bayes factor when the priors are improper. We also propose a new class of priors for the unknown parameters in the alternative hypothesis. Previous attempts in the literature are limited to the conjugate priors or conventional priors such as Jeffreys and g-priors. Furthermore, we have derived closed-form expressions for the marginal likelihoods and the test statistic.

The materials of this dissertation are based on the following papers:

[1] M. H. Al-Ali and K. C. Ho, “Robust transmit precoding for underlay MIMO cognitive radio with interference leakage rate limit,” in *Proc. IEEE ICASSP*, Shanghai, China, Mar. 2016, pp. 3001-3005.

[2] M. H. Al-Ali and K. C. Ho, “Transmit precoding in underlay MIMO cognitive radio with unavailable or imperfect knowledge of primary interference channel,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5143-5155, Aug. 2016.

[3] M. H. Al-Ali and K. C. Ho, “Bayesian multi-antenna sensing in cognitive radio networks using fractional Bayes factor,” in *Proc. IEEE ICASSP*, New Orleans, Mar. 2017.

[4] M. H. Al-Ali and K. C. Ho, “Precoding for MIMO channels in cognitive radio networks with CSI uncertainties and for compound capacity,” under 2nd review, *IEEE Trans. Signal Process.*

1.6 Dissertation Organization

The subsequent chapters of the dissertation are organized as follows. In chapter 2, we design the SU precoder by maintaining the QoS of PU for either IT or LR interference metric under unknown or imperfect CSI between SU-Tx and PU-Rx. Chapter 3 considers CSI of the links from SU-Tx to SU-Rx and from SU-Tx to PU-Rx is inaccurate. We design the worst case precoder matrix for the CR problem under the Schatten norm modelling for the uncertainty sets and a general convex power set. Chapter 4 addresses the spectrum sensing problem such that there is no CSI about the covariance matrices of the received signals. We follow the Bayesian philosophy to marginalize the unknown parameters and derive the detectors. Finally, chapter 5 summarizes the dissertation and discusses some possible future directions to develop the current work.

Chapter 2

Transmit Precoding in Underlay MIMO CR with Unavailable or Imperfect Knowledge of Primary Interference Channel

2.1 Introduction

This chapter addresses the problem of precoder design that maximizes the throughput of an SU in an underlay MIMO CR network, where the CSI from the SU to the PU is unavailable or inaccurate. The design maintains the QoS for the PU through an interference amount measure in terms of the interference temperature or the leakage rate. For the case of unknown CSI, we propose an iterative adaptation algorithm by exploiting the side-information in the primary communication network. For the case of imperfect CSI, we model the amount of uncertainty to be within a convex set defined by the Schatten norm and apply the maximin optimization to obtain the solution. To

complete the study in this chapter, we derive the conditions on the CSI uncertainty radius under which the robust design with imperfect CSI would not perform better than the one with unknown CSI, when using the interference temperature metric. The proposed techniques are supported by numerical simulations.

2.1.1 Background

The precoding benefit offered by MIMO antennas in a CR network relies on the availability of CSI for the connections from SU-Tx to SU-Rx and from SU-Tx to PU-Rx. Perfect CSI knowledge enables optimal precoding to meet the QoS requirement of PU by limiting the leakage power or leakage rate of the interference caused by SU [32, 88]. The obtained CSI is often far from perfect in practice [28, 63, 112]. Channel estimation errors, limited or outdated feedback, and short coherence time of the physical channels are among the main factors that cause imperfect CSI. SU-Tx is not even able to acquire the CSI when there is no co-operation from the PU network or when fast fading occurs [55]. These practical limitations make the design problem very challenging. This chapter proposes solutions for the design of MIMO precoders at the SU-Tx cognitive node that maximizes the throughput of SU while maintaining the QoS of PU, under the situation in which the interference CSI from SU-Tx to PU-Rx is unavailable or imperfect.

Although gaining more attention recently in CR, relatively few studies in literature address the unknown CSI situation. When neither the interference channel nor its distribution is accessible by the SU, Pei *et al.* [71] proposed a minimum rank precoding design that achieves a target information rate for SU. However, this method has no guarantee of the required QoS for PU.

The uncertainty in CSI can be taken into account to improve the precoder design. It can be modelled in a deterministic or stochastic form. The former assumes the error is within some uncertainty set whose shape and size are known. The design seeks the precoder that meets the stringent QoS of PU at the worst interference scenario over all the CSI possibilities within the uncertainty set. The stochastic model characterizes the uncertainty through a statistical distribution and the precoder is obtained based on the outage probability [67,80,101]. Apart from the models, CSI can appear instantaneous for slow fading, see [39,69,80,97,101] and [22,51,72,99,102,105], or long-term [9,43,95,98,101] for fast fading propagation environment. The study in this chapter uses the deterministic uncertainty model, and the proposed design is applicable for instantaneous or long-term CSI.

The approach for the precoder design under the worst case scenario of deterministic CSI uncertainty is mostly based on robust beamforming technique developed in the signal processing community [22,43,51,72,95,97,98,105]. Due to the intractability caused by placing the uncertainty set into the optimization problem, most studies use approximations or relaxation along with SDP to obtain a solution [22,39,69,72,97]. Relatively few works focus on producing an equivalent and yet convex formulation of the original problem to yield more accurate and less conservative solutions. Among them, [51,105] used the S-procedure and [43,95,98] utilized the Lagrange duality to obtain the equivalent formulations. They have limited applicability where the former is for the Euclidean norm and the latter the Frobenius and trace norms uncertainty sets. The study here provides an accurate design without limiting to a certain norm uncertainty set through the Schatten norm formulation.

2.1.2 Chapter Organization

This chapter is organized as follows. The system model and problem formulation are introduced in Section 2.2, where the interference metrics, objective function along with the constraints are discussed. Section 2.3 develops the precoder solution when no CSI is available, derives a lower bound for the SU rate performance without CSI, and analyzes its asymptotic performance. Section 2.4 provides the design for the robust precoder using imprecise CSI. Section 2.5 derives the conditions such that the robust solution with partial CSI will not be better than the other without CSI. Section 2.6 contains the simulation results. Section 2.7 is the conclusion and Section 2.8 illustrates the derivations developed in this chapter.

2.2 System Model and Problem Formulation

2.2.1 System Model

We shall begin our development with a scenario containing two transmit-receive pairs as shown in Figure 2.1, one for the PU and the other for the SU network. The PU pair has precedence of the spectrum resources and the SU pair seeks to communicate over the same physical resources by exploiting the underlay CR paradigm. The PU pair has M_p transmit and N_p receive antennas, and the MIMO channel is denoted by $\mathbf{H}_p \in \mathbb{C}^{N_p \times M_p}$. Similarly, the MIMO channel for SU is $\mathbf{H}_s \in \mathbb{C}^{N_s \times M_s}$ with M_s transmit and N_s receive antennas.

To simplify the illustration, we assume the transmission of each user is point-to-point over narrowband flat fading channels. Let $\mathbf{x}_s(n) \in \mathbb{C}^{M_s}$ be the zero-mean

transmitted data vector of SU at time n . The observed signal at SU-Rx is

$$\mathbf{y}_s(n) = \mathbf{H}_s(n)\mathbf{x}_s(n) + \mathbf{z}_s(n) \quad (2.1)$$

where $\mathbf{z}_s(n) \sim \mathcal{CN}(0, \sigma_s^2 \mathbf{I}_{N_s})$ is the additive noise of power σ_s^2 . The time index is added to the channel to signify it is time-varying.

We shall follow [19, 56] and assume the interference caused by PU on SU is negligible. If the interference from PU to SU is significant, the developments and precoder designs in this chapter remain valid by replacing the SU communication channel \mathbf{H}_s with one that contains the equivalent interference effect as detailed after (2). On the other hand, the presence of SU creates non-negligible interference to PU through the channel $\mathbf{G}_s \in \mathbb{C}^{N_p \times M_s}$ giving the interference $\mathbf{G}_s(n)\mathbf{x}_s(n)$. The interference should be kept to a certain level to maintain the QoS of the PU.

The development here is focused on the SU where \mathbf{H}_s is known while \mathbf{G}_s is unavailable or partially known to SU-Tx. The objective is to seek a linear precoder $\mathbf{T} \in \mathbb{C}^{M_s \times r}$ that applies to the data streams of SU, $\mathbf{s}(n) \in \mathbb{C}^{r \times 1}$, to form the transmit vector $\mathbf{x}_s(n) = \mathbf{T}\mathbf{s}(n)$ such that it will not violate the QoS of PU while the throughput of SU is maximized. The transmitted signal $\mathbf{x}_s(n)$ is modelled by a zero-mean complex Gaussian distribution. The design of the linear precoder is equivalent to determining the covariance matrix or codebook $\mathbf{Q}_s = \mathbb{E} [\mathbf{x}_s(n)\mathbf{x}_s^\dagger(n)]$, where $\mathbf{Q}_s = \mathbf{T}\mathbf{T}^\dagger$ [88]. The codebook should satisfy the power constraint $\text{Tr}(\mathbf{Q}_s) \leq P_T$, where P_T is the maximum allowable transmission power of SU-Tx.

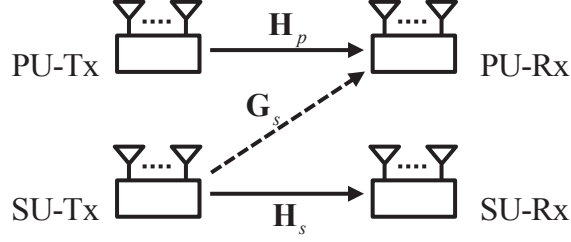


Figure 2.1: An underlay CR user (SU) coexists with a single PU where the transmitter and receiver of each user have multiple antennas. The notations of the various channels are shown.

2.2.2 Design Objective

We shall use an information-theoretic criterion to design \mathbf{Q}_s by maximizing the mutual information of SU. Under the Gaussian transmitted signal model, it is given by [88]

$$C_s(\mathbf{Q}_s) = \log |\mathbf{I}_{N_s} + \mathbf{H}_s \mathbf{Q}_s \mathbf{H}_s^\dagger / \sigma_s^2| \quad (\text{nats/s/Hz}). \quad (2.2)$$

$C_s(\mathbf{Q}_s)$ is called the rate of SU for the rest of the chapter. In (2), we can include the interference from PU to SU by replacing \mathbf{H}_s with $\tilde{\mathbf{H}}_s = \sigma_s \mathbf{R}_S^{-1/2} \mathbf{H}_s$ if it is not negligible, where $\mathbf{R}_S^{-1/2}$ is the matrix square root of \mathbf{R}_S . The matrix \mathbf{R}_S is defined as $\mathbf{R}_S \triangleq \sigma_s^2 \mathbf{I}_{N_s} + \mathbf{G}_p \mathbf{Q}_p \mathbf{G}_p^\dagger$, where $\mathbf{G}_p \in \mathbb{C}^{N_s \times M_p}$ is the interference channel from PU-Tx to SU-Rx and \mathbf{Q}_p is the precoder used by PU.

We shall use $\Gamma(\mathbf{Q}_s, \mathbf{R}_G)$ to denote the interference measure at SU-Tx for the purpose to maintain the QoS of PU. Apart from the precoder \mathbf{Q}_s , it also depends on the covariance matrix \mathbf{R}_G , where for slow fading environments [51, 89, 99, 102, 105] and fast fading propagations [9, 43, 95, 98] it is defined respectively as $\mathbf{R}_G = \mathbf{G}_s^\dagger \mathbf{G}_s$ and $\mathbf{R}_G = \mathbb{E} [\mathbf{G}_s^\dagger \mathbf{G}_s]$. The function $\Gamma(\mathbf{Q}_s, \mathbf{R}_G)$ can appear in several forms as described below.

IT Metric: The interference measure using the IT metric is [32, 88],

$$\Gamma(\mathbf{Q}_s, \mathbf{R}_G) = \text{Tr}(\mathbf{R}_G \mathbf{Q}_s) - I_T \quad (2.3)$$

where I_T is the maximum interference level allowed at the receiver.

LR Metric: It has been shown by Monte Carlo simulations [71] that the leakage rate of the interference link between SU-Tx and PU-Rx could be a better measure of the QoS for PU than IT which is essentially the leakage power. That is, instead of considering the power of the quantity $\mathbf{R}_G \mathbf{Q}_s$ we assess its information rate. We define the LR metric as

$$\Gamma(\mathbf{Q}_s, \mathbf{R}_G) = \log |\sigma_p^2 \mathbf{I}_{M_s} + \mathbf{R}_G^{1/2} \mathbf{Q}_s \mathbf{R}_G^{1/2}| - R_L. \quad (2.4)$$

The first term on the right side of (2.4) is the interference leakage rate encountered by PU and R_L is the maximum interference leakage rate at PU-Rx. σ_p^2 is the power of the Gaussian IID noise at PU-Rx and $\mathbf{R}_G^{1/2}$ is the matrix square root of \mathbf{R}_G .

Let us define for simplicity the set of all feasible values of \mathbf{Q}_s to be

$$\mathcal{Q} = \{\mathbf{Q}_s : \mathbf{Q}_s \succeq 0, \text{Tr}(\mathbf{Q}_s) \leq P_T\}. \quad (2.5)$$

In mathematical form, the problem we would like to address is

$$(P1): \quad \max_{\mathbf{Q}_s \in \mathcal{Q}} C_s(\mathbf{Q}_s) \quad (2.6a)$$

$$\text{s.t. } \Gamma(\mathbf{Q}_s, \mathbf{R}_G) \leq 0. \quad (2.6b)$$

The problem P1 is convex in \mathbf{Q}_s under the IT metric, and it is not when the LR

metric (2.4) is used in (2.6b).

2.2.3 Interference CSI

If \mathbf{R}_G is perfectly known at SU-Tx, obtaining the solution to the problem P1 is straightforward by using a numerical SDP optimization package such as CVX [34]. Indeed, for interference-limited systems where $\text{Tr}(\mathbf{Q}_s) \leq P_T$ is inactive the optimal rate of SU from P1 under the IT metric has an algebraic form as derived in Appendix 2.8.1. The work in [88] derived closed-form solutions that are optimal when both PU-Rx and SU-Rx have a single antenna only, i.e., $N_p = N_s = 1$.

Due to the practical limitation that the PU and SU networks are seldom coordinated or have limited interaction, the CSI is often inaccurate and we shall model the available $\hat{\mathbf{R}}_G$ as

$$\mathbf{R}_G = \hat{\mathbf{R}}_G + \Delta\mathbf{R}_G. \quad (2.7)$$

In (2.7), $\Delta\mathbf{R}_G \in \mathbb{H}^{M_s}$ represents the CSI uncertainty. Obviously, obtaining the transmit precoder using P1 by pretending $\hat{\mathbf{R}}_G$ as the true value will not guarantee (2.6b) is fulfilled.

This chapter considers two cases about the interference CSI. The first case is that $\hat{\mathbf{R}}_G$ is not available at SU-Tx and we propose a new algorithm to solve P1. The second case is when an estimate $\hat{\mathbf{R}}_G$ is available [28, 63, 112], and we take into account the uncertainty $\Delta\mathbf{R}_G$ to derive a robust solution. In particular, we model the uncertainty within a certain set and optimize P1 based on the worst case concept [22, 69, 72, 97, 105] to reach a robust solution.

The precoder design without using interference CSI is suitable for use in FDD

systems with slow fading situation in general. The robust design precoder using inaccurate CSI is applicable to both FDD and TDD systems and can be applied to fast or slow fading scenarios.

2.3 Unavailable CSI scenario

This section addresses the SU precoder design in the situation where the interference channel covariance matrix \mathbf{R}_G is not available at SU-Tx. We shall propose an alternative formulation to P1 through the adaptation of the available power and rank of the precoder at SU such that the QoS for PU is maintained. The proposed solution assumes the availability of some side-information in the primary network to indicate if the PU transmission is successful. Such side-information can be easily assessable in modern communication systems as will be elaborated.

2.3.1 Proposed Solution

It is not possible to obtain the optimal precoder for P1 by a typical numerical optimization algorithm without knowing \mathbf{R}_G that appears in (2.6b). An alternative to the algebraic evaluation of (2.6b) is to examine if the PU transmission is successful when SU is operating with a certain precoder. This seems impractical since there are infinitely number of possible choices for \mathbf{Q}_s in \mathcal{Q} , and many of them could harm the PU transmission. We shall propose a suboptimal solution for \mathbf{Q}_s by limiting it to a certain structure so that the evaluation of (2.6b) through transmission is kept to a reasonable level.

Let us denote $\mathbf{R}_H = \mathbf{H}_s^\dagger \mathbf{H}_s$ that has a rank $d = \min(M_s, N_s)$ and eigenvalue

decomposition (EVD) $\mathbf{U}_{\mathbf{H}_s} \mathbf{\Lambda}_{\mathbf{H}_s} \mathbf{U}_{\mathbf{H}_s}^\dagger$, where $\mathbf{U}_{\mathbf{H}_s} \in \mathbb{C}^{M_s \times d}$ is a unitary matrix and the diagonal matrix $\mathbf{\Lambda}_{\mathbf{H}_s}$ has the eigenvalues arranged in non-increasing order. In the absence of the QoS constraint (2.6b), it is direct to show that the optimum precoder $\bar{\mathbf{Q}}_s$ of P1 follows the waterfilling solution

$$\bar{\mathbf{Q}}_s = \mathbf{U}_{\mathbf{H}_s} \bar{\mathbf{\Lambda}}_{\mathbf{Q}_s} \mathbf{U}_{\mathbf{H}_s}^\dagger, \quad \bar{\mathbf{\Lambda}}_{\mathbf{Q}_s} = (\nu^{-1} \mathbf{I}_d - \sigma_s^2 \mathbf{\Lambda}_{\mathbf{H}_s}^{-1})^+ \quad (2.8)$$

where ν is the water-level chosen such that $\text{Tr}(\bar{\mathbf{\Lambda}}_{\mathbf{Q}_s}) = P_T$. The solution in (2.8) would likely violate (2.6b) and a more conservative precoder is needed. The proposed suboptimal precoder has a structure similar to that in (2.8), but with a rank r and a certain transmission power P ,

$$\mathbf{Q}_s = \mathbf{U}_{\mathbf{H}_{s,r}} \mathbf{\Lambda}_{\mathbf{Q}_s} \mathbf{U}_{\mathbf{H}_{s,r}}^\dagger, \quad \mathbf{\Lambda}_{\mathbf{Q}_s} = (\nu^{-1} \mathbf{I}_r - \sigma_s^2 \mathbf{\Lambda}_{\mathbf{H}_{s,r}}^{-1})^+ \quad (2.9)$$

where $\mathbf{\Lambda}_{\mathbf{H}_{s,r}}$ is the upper left $r \times r$ block of $\mathbf{\Lambda}_{\mathbf{H}_s}$ and contains the largest r eigenvalues of \mathbf{R}_H , and $\mathbf{U}_{\mathbf{H}_{s,r}}$ is the matrix with the corresponding eigenvectors. The water-level ν in (2.9) is found from $\text{Tr}(\mathbf{\Lambda}_{\mathbf{Q}_s}) = P$. When putting (2.9) into (2.2), the problem P1 becomes

$$\text{(P2):} \quad \max_{P \leq P_T, r \leq d} C_s = \log |\mathbf{I}_r + \mathbf{\Lambda}_{\mathbf{H}_{s,r}} \mathbf{\Lambda}_{\mathbf{Q}_s} / \sigma_s^2| \quad (2.10a)$$

$$\text{s.t.} \quad \Gamma(\mathbf{\Lambda}_{\mathbf{Q}_s}, \mathbf{R}_{G,H}) \leq 0, \quad (2.10b)$$

where $\mathbf{R}_{G,H} = \mathbf{U}_{\mathbf{H}_{s,r}}^\dagger \mathbf{R}_G \mathbf{U}_{\mathbf{H}_{s,r}}$. The solution from P2 is dependent on the SU-Tx power and the precoder rank as implicitly indicated in $\mathbf{\Lambda}_{\mathbf{Q}_s}$ and $\mathbf{R}_{G,H}$. Note that we do not evaluate (2.10b) explicitly since \mathbf{R}_G is not known. We shall instead use the

side-information associated with the PU transmission to test if the QoS requirement is satisfied.

In this design, the precoder steers the SU signal in the directions of \mathbf{R}_H eigenvectors to maximize the SU rate while optimizing its power and rank to keep the interference to PU to an acceptable level by maintaining (2.10b). The optimization over the rank of \mathbf{Q}_s affects the spatial dimension of the interference, which translates to limiting the interference to PU [68, 71]. The proposed formulation and solution of P2 are different from [71] that minimizes the rank at a preselected SU rate without considering PU QoS and from [68] that solves the feasibility problem for interference alignment. Moreover, our work is different from [63] that learns iteratively the null space of \mathbf{R}_G by observing the transmitted power of PU-Tx, it also requires the number of deployed antennas at SU-Tx and PU-Rx are different.

The proposed problem P2 has lower complexity than the matrix optimization problem in P1 as it enables optimization with respect to two scalars only. Consequently, we limit the interference induced on PU-Rx by adapting the power and rank of the precoder starting with increasing values during optimization of the rate of SU.

The proposed solution for P2 meets (2.10b) through the use of the side-information from PU [63, 85, 106] that indicates PU's transmission performance. The operation of P2 depends on the form side-information is present in the primary network.

Continuous transmission: This approach requires PU-Rx to be able to broadcast an alert signal when the transmission quality is not acceptable (implying interference measure $\Gamma(\mathbf{A}_{\mathbf{Q}_s}, \mathbf{R}_{G,H})$ is positive). It also requires SU-Tx to initiate its transmission with a unit rank solution, i.e., beamforming, along with a precoder of low power. The power can be gradually increased as long as the QoS of PU is not violated.

This process can be repeated for higher rank solutions until $r = d$ or the SU rate improvement due to rank increasing is not significant. Similar concept of power adaptation scheme has been used in [63].

Packet transmission: For packet network with ACK/NAK signaling, outage caused by interference can be tolerated as long as it is below a certain percentage. The QoS requirement (2.10b) is maintained so long as the amount of outage caused by SU-Tx does not exceed that percentage limit.

The algorithm for optimizing P2 under continuous transmission is summarized in Algorithm 1. In step 2 the algorithm evaluates $\Lambda_{\mathbf{Q}_s}^{(k)}$ in P2 at the k th iteration. Whenever the SU rate obtained from (2.10a) at $\Lambda_{\mathbf{Q}_s}^{(k)}$ is larger than some value R_o , construct the precoding matrix \mathbf{T} based on $\mathbf{Q}_s^{(k)}$ in step 3 and generate the precoded data $\mathbf{x}_s(n) = \mathbf{T}\mathbf{s}(n)$ for transmission in step 4, where n is the time sample count local to a transmission cycle. Next, update R_o as long as the constraint (2.10b) is not violated at the given rank. We should notice that during the optimization process SU transmits data signals using different \mathbf{T} only when the calculated rate at the k iteration is larger than the previous best rate R_o . Once the maximum rate from P2 is achieved, the corresponding precoder solution is \mathbf{Q}_s^* . The SU continues sending data using \mathbf{Q}_s^* as long as \mathbf{G}_s remains the same. We shall only need to adapt the rank and power near those for the solution found previously if \mathbf{G}_s varies gradually.

In Algorithm 1, we can adjust the step size of the search in the loop over power to improve the resolution of the power value. Alternatively, it would be more efficient to apply a coarse search in the loop over power and followed with the bi-section technique to reach a fine resolution. Let P_o be the lower limit of the power of \mathbf{Q}_s , μ be the coarse step-size and ν be the desired resolution of the power. Typically μ

is much larger than ν . The total number of iterations to reach the power level value at a given rank is no more than $\lceil (P_T - P_o)/\mu + \log_2(\mu/\nu) + 1 \rceil$. In addition to the bi-section method, we can step from the initial power P_o to P_T using a grid search with an adaptive step size that can be set according to the channel coherence time or the side-information broadcast interval.

Algorithm 1 Iterative algorithm for solving P2

Initialization: Set the minimum SU power P_o and counter $k = 0$

Requirement: P_T, d, I_T or R_L

For $r = 1, \dots, d$

For $P = P_o, \dots, P_T$

 1. $k = k + 1$.

 2. Compute $\mathbf{\Lambda}_{\mathbf{Q}_s}^{(k)}$ from (2.9) and $C_s^{(k)}$ from (2.10a).

If $C_s^{(k)} > R_o$

 3. Obtain $\mathbf{Q}_s^{(k)}$ from (2.9) and construct \mathbf{T} .

 4. Transmit $\mathbf{x}_s(n) = \mathbf{T}\mathbf{s}(n)$ over n in a transmission cycle.

If PU transmission acceptable (no alert over the cycle)

 5. Set $R_o = C_s^{(k)}$ and $\mathbf{Q}_s^* = \mathbf{Q}_s^{(k)}$.

Else

 6. Break updating the power.

EndIf

EndIf

EndFor

EndFor

2.3.2 Solution Analysis

We would like to find an algebraic lower bound for the maximum rate obtained from P2 in which the CSI is not known. To simplify the illustration and gain insight the analysis is restricted to interference-limited systems in which the power constraint $\text{Tr}(\mathbf{Q}_s) \leq P_T$ is inactive.

Proposition 1: The rate of SU achieved from P2 has a lower bound that corre-

sponds to the solution when replacing the constraint (2.10b) by

$$\text{Tr}(\mathbf{\Lambda}_{\mathbf{R}_G, r} \mathbf{\Lambda}_{\mathbf{Q}_s}) \leq I_{TL} \quad (2.11)$$

where r is the rank of \mathbf{Q}_s , $\mathbf{\Lambda}_{\mathbf{R}_G, r}$ is a diagonal matrix that contains the largest r eigenvalues of \mathbf{R}_G , and I_{TL} is a constant that can refer to the IT or LR metric limit, i.e., I_T or R_L . The associated optimization problem can be solved numerically.

Proof: Please see Appendix 2.8.2. ■

The lower bound provides a guarantee of the SU rate achievable when using the proposed design from P2.

We have further investigated the asymptotic performance of the solution from P2 as $N_p \rightarrow \infty$. The main result is summarized in the following proposition.

Proposition 2: The solution of P2 can yield a performance approaching that of P1 as the number of antennas at PU-Rx tends to infinity.

Proof: Please see Appendix 2.8.3. ■

The performance in such a case will be the same as when the SU-Tx to PU-Rx CSI is exactly known.

2.4 Inaccurate CSI

We assume that an inaccurate copy of the channel covariance matrix $\hat{\mathbf{R}}_G$ is available and according to (2.7) it is different from the actual by $\Delta\mathbf{R}_G \triangleq \mathbf{R}_G - \hat{\mathbf{R}}_G$. Let the feasible set to which $\Delta\mathbf{R}_G$ belongs be

$$\mathcal{U} = \{\Delta\mathbf{R}_G : \Delta\mathbf{R}_G \in \xi, \hat{\mathbf{R}}_G + \Delta\mathbf{R}_G \succeq 0\} \quad (2.12)$$

where ξ is the uncertainty set to be defined. The design seeks the worst uncertainty matrix that gives the largest interference measure $\Gamma(\mathbf{Q}_s, \mathbf{R}_G)$ when solving P1, i.e., the maximin solution [18]. The problem P1 becomes

$$(P3): \quad \max_{\mathbf{Q}_s \in \mathcal{Q}} C_s(\mathbf{Q}_s) \quad (2.13a)$$

$$\text{s.t. } \Gamma_{\max} \leq 0 \quad (2.13b)$$

where Γ_{\max} corresponds to the value of the following optimization subproblem

$$\max_{\Delta \mathbf{R}_G \in \mathcal{U}} \Gamma(\mathbf{Q}_s, \hat{\mathbf{R}}_G + \Delta \mathbf{R}_G). \quad (2.14)$$

Before continuing our proposed solution further, let us elaborate on the uncertainty set since it affects the solution of the worst $\Delta \mathbf{R}_G$.

2.4.1 Modelling the Uncertainty Set

It is common in literature to define the uncertainty set ξ using some matrix norms [9, 51, 69, 89, 95, 99, 105], see in particular [99] and the references therein. Instead of dealing with each norm one by one, we follow [102] and use the Schatten norm $\|\cdot\|_{S_p}$ of certain order p to define the uncertainty set. The Schatten norm is defined as follows

Definition 1 [13, Proposition 9.2.3]: Let $\mathbf{W} \in \mathbb{C}^{n \times m}$ be a matrix whose i th singular value is $\sigma_i \in \mathbb{R}_+$, $i = 1, 2, \dots, \min(n, m)$. The maximum singular value is denoted by

σ_1 . The Schatten norm of \mathbf{W} is

$$\|\mathbf{W}\|_{S_p} = \begin{cases} \left(\sum_{i=1}^{\min(n,m)} \sigma_i^p\right)^{1/p}, & 1 \leq p < \infty \\ \sigma_1, & p = \infty. \end{cases} \quad (2.15)$$

It is easy to verify that by choosing p equal to 1, 2, or ∞ the Schatten norm becomes the Nuclear norm that characterizes the uncertainty in the matrix rank, the Frobenius norm that limits the error power or the Spectral norm that defines the eigenmode uncertainty [102].

The uncertainty set ξ_{S_p} for the channel covariance matrix under the Schatten p -norm is

$$\xi_{S_p} = \{\Delta\mathbf{R}_G : \|\Delta\mathbf{R}_G\|_{S_p} \leq \epsilon\} \quad (2.16)$$

where ϵ is the uncertainty radius. It is straightforward to verify that

$$\|\Delta\mathbf{R}_G\|_{S_\infty} \leq \|\Delta\mathbf{R}_G\|_{S_q} \leq \|\Delta\mathbf{R}_G\|_{S_p} \leq \|\Delta\mathbf{R}_G\|_{S_1} \quad (2.17)$$

for $1 \leq p \leq q \leq \infty$.

The following two subsections give the general framework to solve the subproblem in (2.14) for the two interference metrics IT and LR.

2.4.2 Robust Solution for IT Metric

At a given \mathbf{Q}_s and using (2.3) along with (2.16), (2.14) becomes

$$(P3-I): \quad \max_{\Delta \mathbf{R}_G} \text{Tr} \left((\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G) \mathbf{Q}_s \right) - I_T \quad (2.18a)$$

$$\text{s.t. } \Delta \mathbf{R}_G \in \mathcal{U}. \quad (2.18b)$$

The above subproblem is convex in $\Delta \mathbf{R}_G$. We shall provide a new solution to P3-I based on the Schatten p -norm and the Lagrange dual of P3-I. The Lagrange dual has been used before to solve a constrained problem [95, 98] and [43], but these previous works are on different optimization objectives and their solutions apply to the uncertainty sets modeled by the Frobenius and Nuclear norms only. Using the Lagrange dual has the benefit of avoiding the evaluation of $\Delta \mathbf{R}_G$ explicitly for the optimization.

Proposition 3: The subproblem P3-I can be compactly expressed in terms of its Lagrange dual as

$$\inf_{\mathbf{Y} \succeq 0} \text{Tr} (\hat{\mathbf{R}}_G (\mathbf{Q}_s + \mathbf{Y})) + \epsilon \|\mathbf{Q}_s + \mathbf{Y}\|_{S_q} - I_T \quad (2.19)$$

where \mathbf{Y} is a PSD matrix and $1/p + 1/q = 1$. The minimum of (2.19) is reached when $\mathbf{Y} = 0$.

Proof: Please see Appendix 2.8.4. ■

Interestingly enough, Proposition 3 can be utilized for solving the problem of robust transmit beamforming in a single- or multi-group multi-casting networks [43, 72, 98] when redefining the variables. It can also yield the robust design solutions that are derived from a different approach in [69] for coexisting a single antenna OFDM

based SU in a PU network under the general ellipsoidal norm. Moreover, our proposed solution encompasses the ones in [95,98] that account for the CSI uncertainty under the Frobenius norm.

Since the minimum of (2.19) is achieved when $\mathbf{Y} = 0$, we have

$$\Gamma_{\max} = \text{Tr}(\hat{\mathbf{R}}_{\mathbf{G}}\mathbf{Q}_s) + \epsilon\|\mathbf{Q}_s\|_{S_q} - I_T. \quad (2.20)$$

Comparing with (2.3), the robust design uses the dual norm of the Schatten p -norm scaled by the uncertainty radius to form the regularization term $\epsilon\|\mathbf{Q}_s\|_{S_q}$ for ensuring robustness. Indeed, one may infer from (2.17) that choosing the uncertainty set defined by the Spectral norm ξ_{∞} results in the most conservative robust solution. In this case we can rewrite (2.20) as

$$\Gamma_{\max} = \text{Tr}(\tilde{\mathbf{R}}_{\mathbf{G}}\mathbf{Q}_s) - I_T, \quad \tilde{\mathbf{R}}_{\mathbf{G}} = \hat{\mathbf{R}}_{\mathbf{G}} + \epsilon\mathbf{I}_{M_s}. \quad (2.21)$$

The equivalent CSI covariance matrix $\tilde{\mathbf{R}}_{\mathbf{G}}$ has the same eigenvectors as the available one but its eigenvalues are increased by the amount ϵ .

If the optimal precoder solution is of rank one, i.e., the transmission strategy is simply beamforming, (2.20) is the same and the robust solution to P3 is identical regardless of the choice of p in the Schatten norm CSI uncertainty. This indicates that beamforming is robustly optimal for the various norms of uncertainty despite its simplicity. Similar conclusion was also made in [102] when the objective function to be optimized in (2.13a) is $\text{Tr}(\mathbf{H}_s\mathbf{Q}_s\mathbf{H}_s^{\dagger})$ for \mathbf{Q}_s under the set \mathcal{Q} only.

In general, solving P3 using (2.20) for the left side of (2.13b) is not a difficult task by using a convex optimization software package. Nevertheless, for interference-

limited systems where the interference constraint (2.13b) dominates such that the transmit power constraint (2.5) is inactive, the solution to P3 under (2.21) is in the form of waterfilling, please see Appendix 2.8.5 for details.

2.4.3 Robust Solution Based Linearization for LR Metric

When using the LR interference metric defined in (2.4), (2.14) can be casted as

$$\begin{aligned}
 \text{(P3-II): } \quad & \max_{\Delta \mathbf{R}_G} \log |\sigma_p^2 \mathbf{I}_{M_s} + (\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G)^{1/2} \mathbf{Q}_s (\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G)^{1/2}| - R_L \quad (2.22a) \\
 & \text{s.t. } \Delta \mathbf{R}_G \in \mathcal{U}. \quad (2.22b)
 \end{aligned}$$

Note that \mathbf{Q}_s is fixed and $\Delta \mathbf{R}_G$ is the variable in (2.22). The resulting Γ_{\max} for (2.13b) after optimizing P3-II remains to be non-convex with respect to \mathbf{Q}_s and it will be challenging to solve the problem (2.13). The proposed approach handles the non-convexity by linearizing the log-det objective function in a local neighborhood and tightening the LR requirement a little to enable a simple solution for $\Delta \mathbf{R}_G$. Stimulated by the application of log-det function as a smooth surrogate for the rank function [27], we use a local minimization approach [23]. From [13, Proposition 8.6.13] and the fact that $\log(x)$ for $x > 0$ is an increasing function, we can deduce that the log-det function is strictly increasing with respect to \mathbf{Q}_s and as such a local linear approximation will represent its upper bound, which is achievable when the solution is reached.

The linear approximation of the objective function by the first-order Taylor series

expansion at a certain $\mathbf{Q}_s^{(l)}$, where l is the iteration index, is

$$\log |\sigma_p^2 \mathbf{I}_{M_s} + \mathbf{R}_G^{1/2} \mathbf{Q}_s \mathbf{R}_G^{1/2}| \approx \log |\sigma_p^2 \mathbf{I}_{M_s} + \mathbf{R}_G^{1/2} \mathbf{Q}_s^{(l)} \mathbf{R}_G^{1/2}| + \text{Tr}(\mathbf{A}^{(l)}(\mathbf{Q}_s - \mathbf{Q}_s^{(l)})), \quad (2.23a)$$

$$\mathbf{A}^{(l)} = \mathbf{R}_G^{1/2} (\sigma_p^2 \mathbf{I}_{M_s} + \mathbf{R}_G^{1/2} \mathbf{Q}_s^{(l)} \mathbf{R}_G^{1/2})^{-1} \mathbf{R}_G^{1/2} \quad (2.23b)$$

where $\mathbf{R}_G = \hat{\mathbf{R}}_G + \Delta \mathbf{R}_G$. Substituting (2.23a) into (2.22a) gives

$$\max_{\Delta \mathbf{R}_G} \text{Tr}(\mathbf{A}^{(l)} \mathbf{Q}_s) - R_L^{(l)} \quad (2.24)$$

where

$$R_L^{(l)} = R_L - \log |\sigma_p^2 \mathbf{I}_{M_s} + \mathbf{R}_G^{1/2} \mathbf{Q}_s^{(l)} \mathbf{R}_G^{1/2}| + \text{Tr}(\mathbf{A}^{(l)} \mathbf{Q}_s^{(l)}). \quad (2.25)$$

Replacing (2.22a) by (2.24) makes the optimization of P3 easier to handle with respect to \mathbf{Q}_s . However, finding the value of (2.24) under \mathcal{U} remains challenging.

To proceed further, let us notice that the true CSI matrix, \mathbf{R}_G , is upper bounded by

$$\tilde{\mathbf{R}}_G = \hat{\mathbf{R}}_G + \epsilon \mathbf{I}_{M_s}. \quad (2.26)$$

This can be shown by realizing that $\tilde{\mathbf{R}}_G - \mathbf{R}_G = \epsilon \mathbf{I}_{M_s} - \Delta \mathbf{R}_G$ is Hermitian and has non-negative eigenvalues and hence PSD, since the maximum eigenvalue of $\Delta \mathbf{R}_G$ is no larger than ϵ under the uncertainty defined by (2.16). Thus (2.24) is ensured to be non-positive when

$$\tilde{\Gamma}_{\max}^{(l)} = \text{Tr}(\tilde{\mathbf{A}}^{(l)} \mathbf{Q}_s) - \tilde{R}_L^{(l)} \quad (2.27)$$

is non-positive, where $\tilde{\mathbf{A}}^{(l)}$ and $\tilde{R}_L^{(l)}$ are given in (2.23b) and (2.25) with \mathbf{R}_G replaced

by $\tilde{\mathbf{R}}_G$ defined in (2.26). Thus P3 becomes solvable by some SDP package after using $\tilde{\Gamma}_{\max}^{(l)}$ as Γ_{\max} in (2.13b). If we start with $\mathbf{Q}_s^{(1)} = 0$, the first iteration of (2.27) has a similar expression as (2.21) provided that $\sigma_p^2 = 1$. However, as l increases the value of (2.27) approaches 0, causing reduction in the rate of SU. Let

$$C_{sp}(\mathbf{Q}_s) = \log |\sigma_p^2 \mathbf{I}_{M_s} + \tilde{\mathbf{R}}_G^{1/2} \mathbf{Q}_s \tilde{\mathbf{R}}_G^{1/2}|. \quad (2.28)$$

Algorithm 2 summarizes the required steps to solve P3 under the LR metric. The algorithm happens off-line in the SU-Tx processing unit.

If the power constraint in (2.5) is inactive, it can be verified that the solution of P3 with the LR metric follows the waterfilling power allocation strategy at each iteration (see Appendix 2.8.5).

In Section 2.4.4, we shall propose another precoder design method when the interference-to-noise ratio (INR) is low.

Algorithm 2 Iterative robust solution for P3 under the LR metric

Initialization: Setting $\mathbf{Q}_s^{(1)} = \gamma \mathbf{I}_{M_s}$ counter $l = 0$, $\gamma \approx 0$ [27], and obtaining $C_{sp}(\mathbf{Q}_s^{(1)})$

Requirement: P_T , R_L , ϵ , and rate accuracy θ

Repeat

1. $l = l + 1$.
2. Evaluate $\tilde{\mathbf{A}}^{(l)}$ using (2.23b) and $\tilde{R}_L^{(l)}$ using (2.25) by replacing \mathbf{R}_G with $\tilde{\mathbf{R}}_G$ defined in (2.26).
3. Obtain $\tilde{\Gamma}_{\max}^{(l)}$ from (2.27).
4. Solve P3 by setting $\Gamma_{\max} = \tilde{\Gamma}_{\max}^{(l)}$ and obtain $\mathbf{Q}_s^{(l+1)}$.
5. Compute $C_{sp}(\mathbf{Q}_s^{(l+1)})$ using (2.28).

Until $|C_{sp}(\mathbf{Q}_s^{(l+1)}) - C_{sp}(\mathbf{Q}_s^{(l)})| \leq \theta$

2.4.4 Robust Solution at Low INR for LR Metric

We shall provide another tractable formulation for the optimization subproblem in the left side of (2.4). The LR metric can be expanded as

$$\Gamma(\mathbf{Q}_s, \mathbf{R}_G) = \frac{1}{\sigma_p^2} \text{Tr}(\mathbf{R}_G \mathbf{Q}_s) + o\left(\frac{1}{\sigma_p^2} \|\mathbf{R}_G^{1/2} \mathbf{Q}_s \mathbf{R}_G^{1/2}\|\right) - R_L \quad (2.29)$$

where $\|\cdot\|$ is a norm measure [99]. We shall define INR at PU-Rx as $\text{Tr}(\mathbf{R}_G \mathbf{Q}_s)/\sigma_p^2$. At low INR regimes where $\text{INR} \leq 0$ dB, the second term on the right side of (2.29) can be ignored. Hence, for a given precoder \mathbf{Q}_s and under low INR condition, the left side of the interference constraint (2.4) can be casted into the following subproblem

$$\text{(P3-III):} \quad \max_{\Delta \mathbf{R}_G} \text{Tr}\left(\frac{1}{\sigma_p^2} (\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G) \mathbf{Q}_s\right) - R_L \quad (2.30)$$

$$\text{s.t.} \quad \|\Delta \mathbf{R}_G\|_{sp} \leq \epsilon \quad (2.31)$$

$$\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G \succeq 0. \quad (2.32)$$

The problem P3-III has a similar format to P3-I and hence we can use Proposition 3 to find the optimal solution for (2.30).

2.5 Selection Between Unavailable and Inaccurate CSI Solutions

We have developed two solutions for the precoder matrix, one corresponds to the unavailable CSI situation P2 and the other the inaccurate CSI situation P3. Both designs have lower performance than the case when the CSI of the interference channel

is exactly known. However, their relative performance depends on how large the uncertainty radius ϵ is. While it is anticipated that the P3 design would always yield better SU rate than P2, we shall show that this is not necessarily the case. In particular, we are going to find the range of the uncertainty radius at which the robust design from P3 is guaranteed to be worse than the CSI absent solution from P2 for the IT metric. Two conditions on it will be derived, one is for the SU rate that is sufficient and the other is for the SU power that is exact. Since the conditions we are going to determine are sufficient or exact, it will ensure that there will be no degradation in, indeed there will be better, performance when we switch from the P3 precoder to the P2 precoder. We would like to clarify that some side-information regarding the successful transmission in the PU is still needed to obtain the P2 precoder solution.

2.5.1 Condition for the Rate of SU

We would like to determine the sufficient condition of the uncertainty radius ϵ at which the rate resulted from the robust design P3 is going to be no better than that from the design P2 under the IT metric.

The rate achievable from the proposed solution of P2, denoted R_s^* , is guaranteed

to be larger than that of P3, represented by $C_s(\mathbf{Q}_s^o)$, if

$$C_s(\mathbf{Q}_s^o) = \mathbb{E} \left[\log \left(1 + \lambda_1(\mathbf{R}_H \mathbf{Q}_s^o) / \sigma_s^2 \right) \right] + \underbrace{\sum_{i=2}^d \mathbb{E} \left[\log \left(1 + \lambda_i(\mathbf{R}_H \mathbf{Q}_s^o) / \sigma_s^2 \right) \right]}_{R_{\text{minor}}} \quad (2.33a)$$

$$\leq \mathbb{E} \left[\log \left(1 + \lambda_1(\mathbf{R}_H) \lambda_1(\mathbf{Q}_s^o) / \sigma_s^2 \right) \right] + R_{\text{minor}} \quad (2.33b)$$

$$\leq \mathbb{E} \left[\log \left(1 + \lambda_1(\mathbf{R}_H) \|\mathbf{Q}_s^o\|_{S_q} / \sigma_s^2 \right) \right] + R_{\text{minor}} \quad (2.33c)$$

$$\leq \log \left(1 + \lambda_1(\mathbf{R}_H) \mathbb{E} \left[\|\mathbf{Q}_s^o\|_{S_q} \right] / \sigma_s^2 \right) + R_{\text{minor}} \leq R_s^*, \quad (2.33d)$$

where the expectation is taken with respect to $\Delta \mathbf{R}_G$ and \mathbf{Q}_s^o is the optimal codebook from P3. The inequality in (2.33b) comes from the fact that $\lambda_1(\mathbf{A}\mathbf{B}) \leq \lambda_1(\mathbf{A})\lambda_1(\mathbf{B})$ for Hermitian PSD $n \times n$ matrices \mathbf{A} and \mathbf{B} [13, Fact 8.19.17]. The inequality in (2.33c) is from (2.17) and it is valid for $q \in [1, \infty]$ and we use Jensen's inequality to obtain (2.33d). The constraint (2.13b) is active at the optimum, and based on (2.20) it becomes

$$\text{Tr}(\hat{\mathbf{R}}_G \mathbf{Q}_s^o) + \epsilon \|\mathbf{Q}_s^o\|_{S_q} = I_T. \quad (2.34)$$

Hence, substituting (2.34) into (2.33d) and solving give explicitly the sufficient condition for uncertainty radius:

$$\epsilon \geq \epsilon_R^* = \frac{\lambda_1(\mathbf{R}_H)(I_T - \mathbb{E}[\text{Tr}(\hat{\mathbf{R}}_G \mathbf{Q}_s^o)])}{\sigma_s^2 [e^{(R_s^* - R_{\text{minor}})} - 1]}. \quad (2.35)$$

A valid value of ϵ_R^* requires both the numerator and the denominator be positive, which implies that $I_T > \mathbb{E}[\text{Tr}(\hat{\mathbf{R}}_G \mathbf{Q}_s^o)]$ and $R_s^* > R_{\text{minor}}$.

The expectations in (2.35) and R_{minor} defined in the second term of (2.33a) need

to be computed numerically since they do not have analytical explicit expressions.

2.5.2 Condition for the Power of SU

We are interested in obtaining the condition in which the power of the robust precoder, P_T^o , is not going to be larger than that of the CSI absent precoder, P_T^* . The main result of this subsection is summarized in the following proposition.

Proposition 4: The exact condition for uncertainty radius at which the power $\mathbb{E}[P_T^o]$ from P3 is not larger than P_T^* from P2 is

$$\epsilon \geq \epsilon_P^* = \mathbb{E} \left[(I_T - \text{Tr}(\hat{\mathbf{R}}_G \mathbf{Q}_s^o)) \text{Tr}(\mathbf{Q}_s^o) / (P_T^* \|\mathbf{Q}_s^o\|_{S_q}) \right] \quad (2.36)$$

for $q \in [1, \infty]$, where the expectation is taken over $\Delta \mathbf{R}_G$ which can be evaluated numerically.

Proof: For any uncertainty set with $p \in [1, \infty]$, $P_T^o = \text{Tr}(\mathbf{Q}_s^o)$ can be found from (2.13b) when Γ_{\max} has the value in (2.20). Hence when $P_T^o \leq P_T^*$, (2.36) is established.

■

Although the bounding value in (2.35) is more computationally complex than that in (2.36), the two conditions correspond to different operating requirements of SU. Specifically, SU may use the former if the rate is of primary interest, otherwise it can use the latter.

We should emphasize that (2.35) and (2.36) are the sufficient and exact conditions on the CSI uncertainty radius for the precoder of P3 performing worse than that of P2 and consequently we can ignore the imprecise CSI information. Thus, we would not expect to have any performance degeneration, indeed would have better

performance, when SU switches its operation from P3 to P2 when those conditions are met. In practice, the critical value ϵ_R^* in (2.35) or ϵ_P^* in (2.36) is computed numerically. During transmission, SU-Tx compares the uncertainty radius ϵ with them to see if the conditions are satisfied to decide using the precoder from P2 instead.

2.6 Numerical Results

2.6.1 Results for IT and LR Metrics

We provide numerical simulations to examine the several precoder designs proposed in this chapter. The total power limit of SU-Tx is $P_T = 20$. For the channel matrices \mathbf{H}_s and \mathbf{G}_s , the elements are independently drawn from a CSCG distribution with zero mean and unit variance. The PSD inaccurate CSI matrix $\hat{\mathbf{R}}_G$ is generated through

$$\hat{\mathbf{R}}_G = (\mathbf{G}_s - \delta\mathbf{G}_s)^\dagger (\mathbf{G}_s - \delta\mathbf{G}_s) \quad (2.37a)$$

$$= \underbrace{\mathbf{G}_s^\dagger \mathbf{G}_s}_{\mathbf{R}_G} - \underbrace{(\delta\mathbf{G}_s^\dagger \mathbf{G}_s + \mathbf{G}_s^\dagger \delta\mathbf{G}_s - \delta\mathbf{G}_s^\dagger \delta\mathbf{G}_s)}_{\Delta\mathbf{R}_G} \quad (2.37b)$$

where $\delta\mathbf{G}_s$ is the uncertainty whose elements are also drawn independently from a zero mean circularly symmetric complex Gaussian (CSCG) distribution. The variance is chosen so that the Schatten norm of the second term in (2.37b) is not larger than ϵ . Note that we require $\epsilon < \|\mathbf{R}_G\|_{S_p}$ to ensure a valid $\hat{\mathbf{R}}_G$ and we set $\epsilon = w\|\mathbf{R}_G\|_{S_\infty}$ with $w \in (0, 1)$ [102]. The results presented are the average over several realizations of \mathbf{H}_s , \mathbf{R}_G , $\Delta\mathbf{R}_G$ or a combination of them to obtain the statistical meaningful results. The value of P_o and initial value for R_o in Algorithm 1 are set to be 0.01 and 10^{-3}

nats/s/Hz. The antenna settings for the SU and PU are $M_s = N_s = 4$ and $M_p = N_p = 4$. The noise levels σ_s^2 and σ_p^2 are set to unity.

We first examine the behavior of the proposed precoder when no CSI is available from Algorithm 1, where the IT metric is used. We also include the precoder design from [71] for which we randomly select the target SU rate such that it satisfies the required I_T for PU. The results shown are the averages over 100 independent realizations of \mathbf{H}_s and \mathbf{G}_s . Figure 2.2 indicates that the SU rate of the proposed precoder (shown in solid line) follows the trend as in the perfect CSI case when solving (2.6) (shown in asterisk-symbol) as the interference limit (power) I_T increases. Using the studies in Section 2.3.2, the lower bound of the SU rate from the proposed algorithm is shown with plus-symbol. There is obvious advantage of the proposed design compared to the design from [71] (shown in triangle-symbol) in improving the rate. The design in [71] requires a line search, such as bi-section, for each rank $r \leq d$. Our proposed design as shown in Algorithm 1 needs a similar line search in each transmission cycle and therefore is more computationally demanding.

We have examined in Figure 2.3 the performance of the proposed algorithm, Algorithm 1, under the situation of time-varying interference channel \mathbf{G}_s using the model from [107] where each element of the channel follows a Rayleigh distribution. With the same antenna settings as in Figure 2.2, $P_o = 0.01$, $10 \log_{10}(I_T) = 0$, $\mu = 9.995$, and $\nu = 10^{-3}$, the degradation of the SU rate is about 4% at 20 Hz Doppler frequency when the side-information is broadcasted at every 0.01 msec. The degradation increases to 21% when the broadcast interval becomes 0.1 msec. The percentage degradation is insensitive to I_T .

We next examine the performance of the proposed robust precoders with imperfect

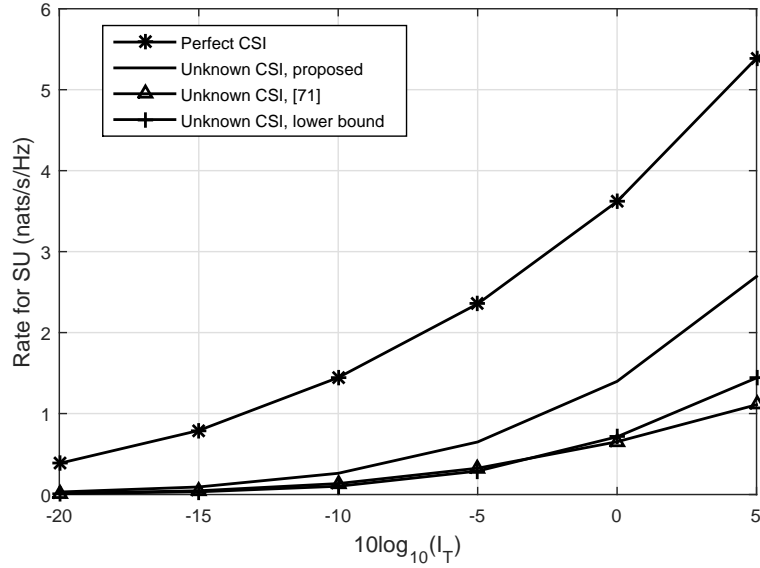


Figure 2.2: Rates for SU from the proposed precoder, the one from [71] for unavailable interference CSI and a lower bound versus the interference limit (I_T), $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are obtained from 100 realizations of \mathbf{H}_s and \mathbf{R}_G .

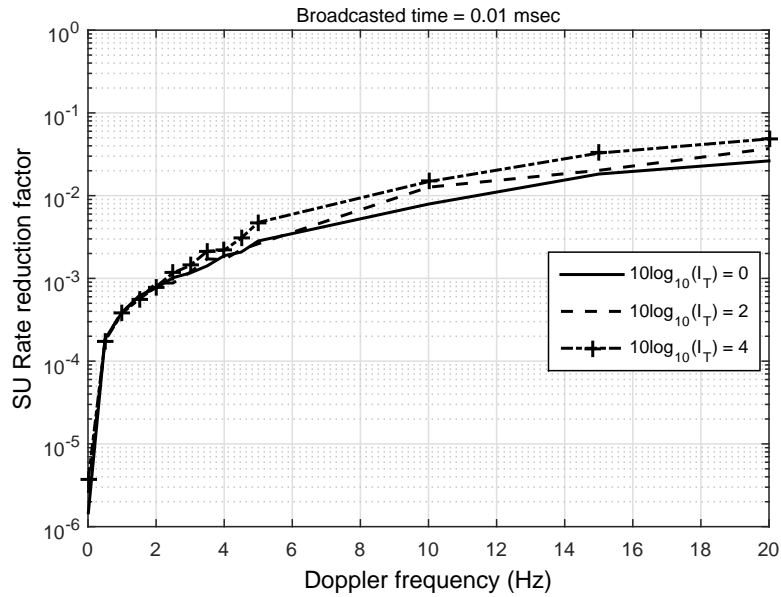


Figure 2.3: SU rate reduction factor versus the Doppler frequency of the time-varying channel \mathbf{G}_s at different levels of the interference limit I_T , $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are the average over 100 realizations of \mathbf{G}_s at a given \mathbf{H}_s .

CSI along with the one corresponds to unavailable CSI for the IT as well as LR metric. Figure 2.4 shows the achieved rate for SU at different values of the IT metric limit I_T . The uncertainty radius ϵ is set with $w = 0.1$, i.e., $\|\Delta\mathbf{R}_G\|_{S_p} \leq 0.1\|\mathbf{R}_G\|_{S_\infty}$. The channel \mathbf{G}_s is a single realization and is held fixed. We show the result for $p = 2$ in the Schatten norm using the proposed design from Section 2.4.2. For comparison purpose we also include the design from [22] that uses the Frobenius norm, i.e., $p = 2$, to model the CSI uncertainty. The results shown are averaged over 100 independent realizations of \mathbf{H}_s and $\Delta\mathbf{R}_G$. As expected, the SU rate for the robust design with imperfect CSI lies between the perfect CSI and the unknown CSI cases. The proposed robust design with $p = 2$ has better performance than the one from [22]. The work of [22] addresses the design for the CSI uncertainty defined by the Frobenius norm and uses the spectral norm as an upper bound to obtain a suboptimal solution. Both design methods may produce similar solutions in the special case where \mathbf{Q}_s has unity rank. These designs belong to the class of determinant maximization with linear matrix inequality (LMI) constraints that can be efficiently solved using the interior-point method [18]. Given M_s transmit antennas in SU, the proposed algorithm with $p = 2$ has computational complexity in terms of the number of Newton steps [96] as $O(\sqrt{4M_s})$, while that of [17] is $O(\sqrt{3M_s})$.

Figure 2.5 illustrates the results when the LR metric is used where the simulation setting is the same as that in Figure 2.4. The robust solution is obtained using Algorithm 2. Corresponding to the results in Figure 2.5, Table 2.1 shows the average number of iterations required for Algorithm 2 to converge at different rate limit R_L and SU rate resolution θ , where $w = 0.1$. It achieves better performance than the unknown CSI case as expected. We do not have the results from [22] for using the

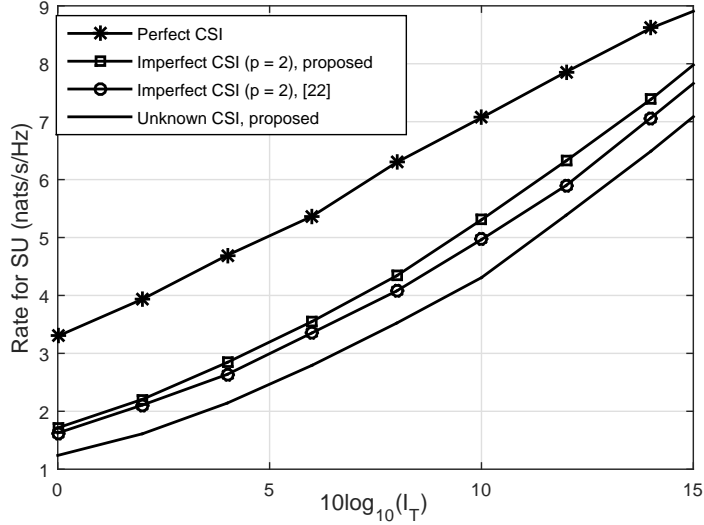


Figure 2.4: Rates for SU versus I_T for the IT metric, $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are generated by the average of 100 realizations for \mathbf{H}_s and $\Delta \mathbf{R}_G$ with $w = 0.1$ at a given \mathbf{G}_s .

LR metric since it is for the IT metric only. In both Figure 2.4 and Figure 2.5, the differences in the SU rates between the robust and the unknown CSI solutions will be larger if w is smaller. In addition, when the interference limit I_T or R_L becomes too large, the interference constraint of problems P1 to P3 would become inactive. As such, the power constraint would dominate performance, yielding comparable results for the perfect, imperfect, or unknown CSI solutions.

To gain some insight on the difference between the two interference metrics on PU when they are active, Figure 2.6 depicts the leakage rate for PU at different achievable rate values of SU for the same settings as that in Figure 2.4. The results of the proposed robust design for different p values have similar trend and only the one with $p = 2$ is shown. In general, one can notice for a certain rate of SU the corresponding leakage rate for PU due to the leakage metric is smaller, which is

more obvious for the imperfect CSI case or larger rate of SU. From the throughput perspective, the leakage metric has less impact on PU while the IT metric is more beneficial to SU. Also, the more uncertain is the interference CSI the higher the leak in the rate for PU.

Lastly, we investigate the performance of the robust precoder as the uncertainty radius increases at different p values. Figures 2.7 and 2.8 illustrate the achieved rate and power for SU at different values of the normalized uncertainty radius w . The settings are $10 \log_{10}(I_T) = 0$, $M_s = M_p = 5$ and $N_s = N_p = 4$. As expected the performance of the robust solutions starts at the same values as the perfect CSI case, it then deteriorates as w (or ϵ) increases and eventually becomes worse than the performance of the unknown CSI case. Due to the large uncertainty region of ξ_{S_∞} it has the largest performance degradation compared to the other norms while ξ_{S_1} yields the smallest degradation. The cross-over point between imperfect and unknown CSI cases can be viewed as the critical rate (or power) for the robust solution beyond which the available imperfect CSI is not useful.

Let the actual ϵ at which the cross-over occurs in the rate case be denoted as ϵ_R . Table 2.2 summarizes ϵ_R at different SU noise power $10 \log_{10}(\sigma_s^2)$ for $p = 1, 2$, and ∞ . They are obtained using the same I_T value and antenna setting as in Figure 2.7, where \mathbf{H}_s and \mathbf{R}_G are created randomly at particular realizations and the averaging is over $\Delta \mathbf{R}_G$. The ϵ_R value is small at high noise power, implying that very accurate CSI is needed for the robust solution of P3 to be better than the unknown CSI solution of P2. Also listed are the boundary ϵ_R^* values from the sufficient condition (2.35) at which the robust design P3 is not better than the P2 solution. The ϵ_R^* values match with the actual cross-over points ϵ_R very well for $p = 1$, as can be seen from the

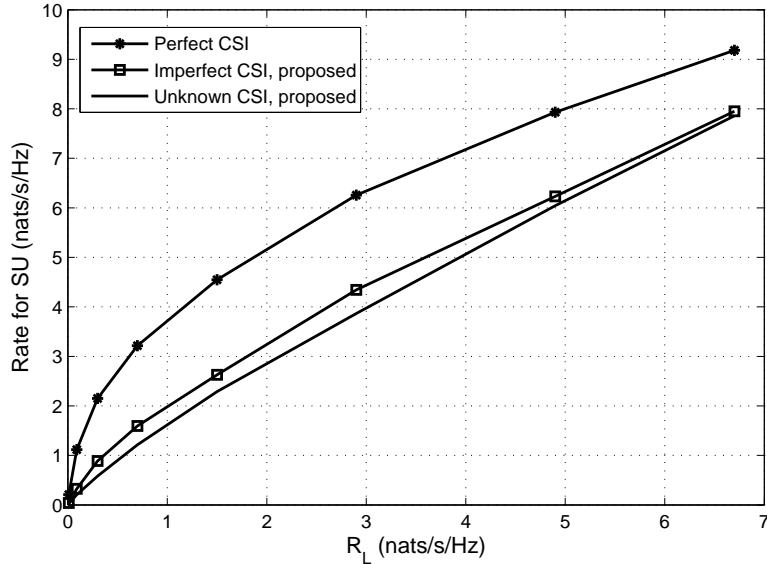


Figure 2.5: Rates for SU versus R_L for the LR metric, $M_s = N_s = 4$ and $M_p = N_p = 4$. The results are generated by the average of 100 realizations for \mathbf{H}_s and $\Delta\mathbf{R}_G$ with $w = 0.1$ at a randomly generated \mathbf{G}_s .

relative deviations. The deviations are larger for the other p values. Table 2.3 shows the critical uncertainty radius ϵ_p^* values for the power case evaluated from (2.36) at different noise power for $p = 1, 2$, and ∞ . We have verified that they are the values obtained from the simulations.

Table 2.1: The average number of iterations for Algorithm 2 to converge at a certain leakage rate limit R_L and rate resolution θ

LR limit		$R_L = 1$	$R_L = 2$	$R_L = 3$	$R_L = 4$
[Num. of Iterations]	$\theta = 10^{-3}$	5	6	6	6
	$\theta = 10^{-5}$	8	14	19	21

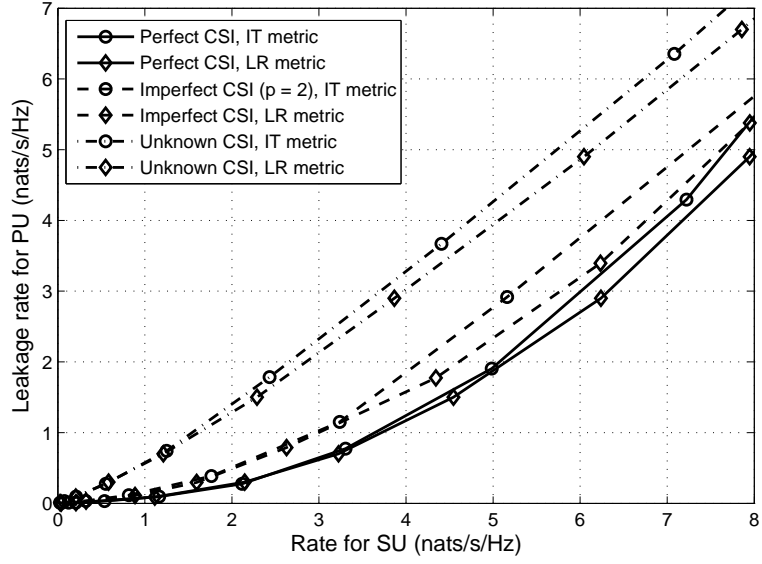


Figure 2.6: Leakage rates (defined below (2.4)) for PU versus the achievable rate for SU for the IT and LR metrics, $M_s = N_s = 4$ and $M_p = N_p = 4$. Results shown are the average of 100 realizations of \mathbf{H}_s and $\Delta\mathbf{R}_G$ at a randomly generated \mathbf{G}_s .

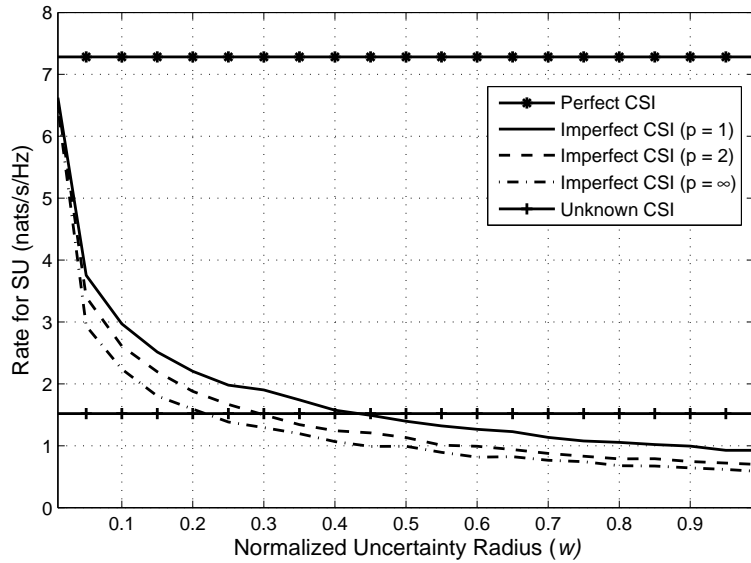


Figure 2.7: Rates for SU versus the normalized uncertainty radius w , $M_s = M_p = 5$ and $N_s = N_p = 4$. The results shown are the averages over 100 realizations of \mathbf{H}_s and $\Delta\mathbf{R}_G$ at a randomly generated \mathbf{G}_s .

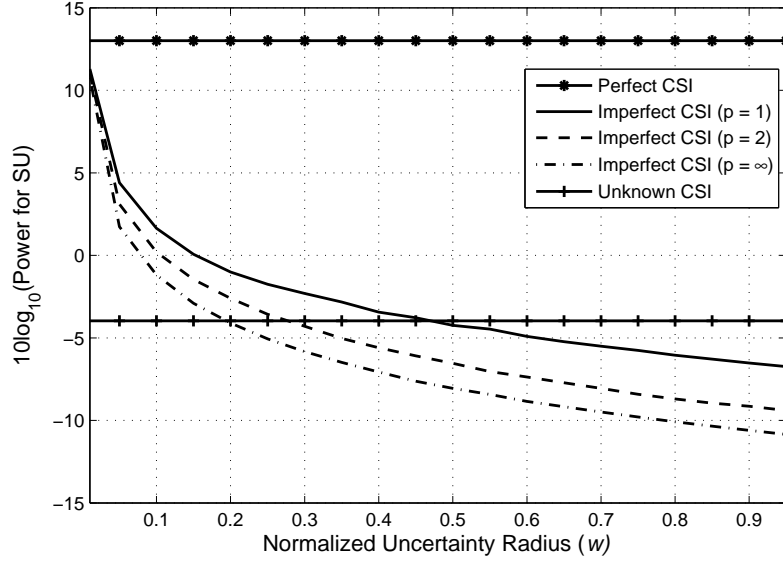


Figure 2.8: Power of SU versus the normalized uncertainty radius w , $M_s = N_s = 5$ and $M_p = N_p = 4$. The results shown are the averages of 100 realizations for \mathbf{H}_s and $\Delta\mathbf{R}_G$ at a randomly generated \mathbf{G}_s .

Table 2.2: The actual value $\hat{\epsilon}_R$, the sufficient value ϵ_R^* and its percent deviation at different levels of the noise power σ_s^2

$10\log_{10} \sigma_s^2$		-10	-5	0	5	10
$p = 1$	$\hat{\epsilon}_R$	2.816	2.0158	1.417	0.886	0.6033
	ϵ_R^*	2.8544	2.0503	1.4459	0.9084	0.6121
	% dev. ϵ_R^*	1.3613	1.7101	2.036	2.5273	1.4549
$p = 2$	$\hat{\epsilon}_R$	1.8796	1.4236	1.0256	0.624	0.5368
	ϵ_R^*	2.2536	1.679	1.1849	0.7528	0.57
	% dev. ϵ_R^*	19.8985	17.9377	15.5291	20.6358	6.1859
$p = \infty$	$\hat{\epsilon}_R$	1.2006	0.9986	0.7401	0.5427	0.4868
	ϵ_R^*	1.7217	1.3838	1.0186	0.6146	0.5132
	% dev. ϵ_R^*	43.4016	38.5786	37.636	13.2362	5.4091

Table 2.3: The value ϵ_P^* at different levels of the noise power σ_s^2

$10 \log_{10} \sigma_s^2$		-10	-5	0	5	10
$p = 1$	ϵ_P^*	2.022	2.1001	0.6195	0.6206	0.6236
$p = 2$	ϵ_P^*	1.3872	1.4899	0.5655	0.5671	0.5699
$p = \infty$	ϵ_P^*	0.9675	1.0912	0.5075	0.5157	0.5369

2.6.2 Results for LR Metric

We shall model the power characteristics of \mathbf{Q}_s using a general convex set \mathcal{Q}_s as [66]

$$\begin{aligned} \mathcal{Q}_s = \{ \mathbf{Q}_s \succeq 0 : \text{Tr}(\mathbf{Q}_s) \leq P_T, \lambda_{\max}(\mathbf{Q}_s) \leq P^{\max}, \\ [\mathbf{Q}]_{s,qq} \leq P_q^{\text{ant}}, q = 1, \dots, M_s \} \end{aligned} \quad (2.38)$$

where P^{\max} is the maximum average power and P_q^{ant} is the average power of the q th antenna.

We shall provide some numerical simulations that illustrate the performance of the proposed designs in Sections 2.4.3 and 2.4.4 for the LR metric. The antenna settings for the SU and PU are $M_s = N_s = 3$ and $M_p = N_p = 6$. The noise power values are $\sigma_s^2 = \sigma_p^2 = 1$. The results shown are the averages over 1000 independent realizations of \mathbf{H}_s and \mathbf{G}_s .

We first examine the performance of the precoder design under the low INR condition presented in Section 2.4.4. The power settings in the set \mathcal{Q}_s are $P_T = \sigma_p^2 = 1$, $P^{\max} = P_1^{\text{ant}} = 0.6$, and $P_q^{\text{ant}} = 0.3$ for $q = 2, \dots, M_s$. Figure 2.9 shows the achieved SU rate as the leakage rate limit R_L in nats/s/Hz increases. The result when the interference channel is perfectly known is shown as a reference. When the amount of

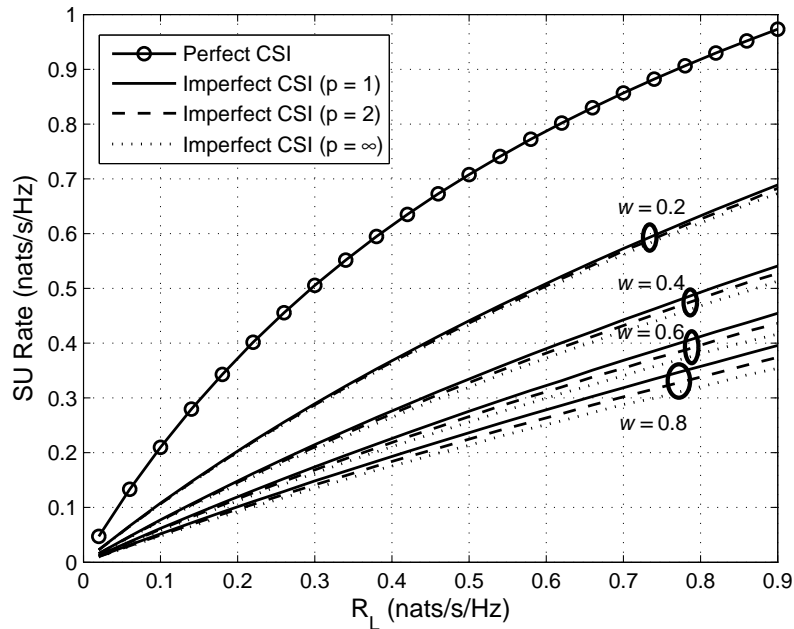


Figure 2.9: Rates for SU from the design using low INR approximation versus the leakage rate limit R_L . The results for imperfect interference CSI at $p = 1, 2$, and ∞ are shown. The settings of antennas are $M_s = N_s = 3$ and $M_p = N_p = 6$. The results are generated by the average of 1000 realizations for \mathbf{H}_s and \mathbf{G}_s for different values of w .

uncertainty in interference CSI increases (increasing w), the performance is further away from the perfect CSI scenario as expected. The reduction in performance seems to be more sensitive at small uncertainty than large. With respect to different p values for the Schatten norm used in defining the uncertainty set, the difference is more obvious at large amount of uncertainty and high leakage rate limit.

Next, we look at the behavior of the precoder design using the iterative linearization method summarized in Algorithm 2. The SU-Tx power settings are $P_T = 10$, $P^{max} = P_1^{ant} = 6$, and $P_q^{ant} = 5$ for $q = 2, \dots, M_s$. Figure 2.10 illustrates the achieved SU rate at different values of R_L and several levels of the interference CSI uncertainty, with the order of the Schatten norm for CSI uncertainty set fixed to $p = \infty$. We also

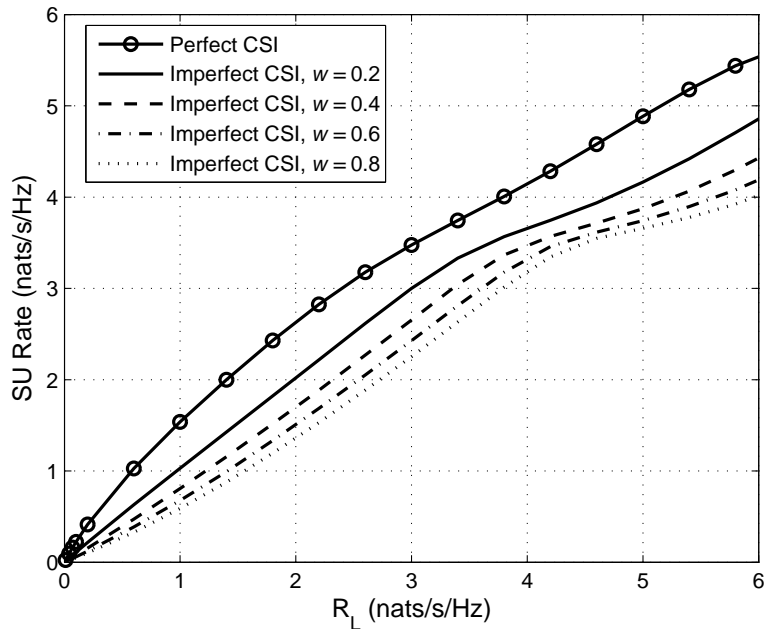


Figure 2.10: Rates for SU from the iterative linearization solution for different leakage rate limit R_L and amount of interference CSI uncertainty w . The results are generated by the average of 1000 realizations for \mathbf{H}_s and \mathbf{G}_s . The settings for antennas are $M_s = N_s = 3$ and $M_p = N_p = 6$.

include the result when the interference CSI is perfectly known for reference purpose. The proposed design yields a solution that follows very well with the ideal solution with perfect CSI. Similar to Figure 2.9, higher amount of CSI uncertainty (increasing w) would reduce the performance of SU.

Figure 2.11 illustrates the behavior of the SU rate for a given R_L as the amount of CSI uncertainty increases. The simulation setting is similar to that in Figure 2.10. The SU rate is more sensitive to the CSI uncertainty level when it is smaller. The SU rate increases with R_L as expected.

The solution using the low INR approximation has similar complexity as the algorithm in [22] for the uncertainty set defined by $p = 2$. At $R_L = 0.9$ nats/s/Hz

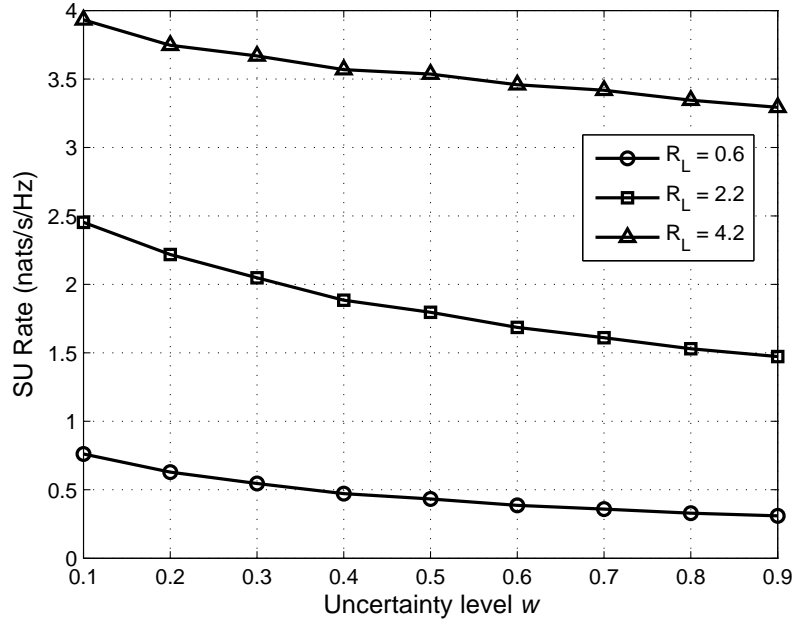


Figure 2.11: Rates for SU from the iterative linearization solution for different uncertainty levels $w = \epsilon/\|\mathbf{R}_G\|_{S\infty}$ and leakage rate limit R_L . The results are generated by the average of 1000 realizations for \mathbf{H}_s and \mathbf{G}_s . The settings for antennas are $M_s = N_s = 3$ and $M_p = N_p = 6$.

and $w = 0.2$ the computation time for the proposed algorithm is 1.5 times higher. The required time for the iterative linearization approach is 6.5 times higher for $\theta = 10^{-4}$, where the power settings are similar to that in Figure 2.9.

2.7 Conclusion

This chapter investigates the cognitive radio design of optimizing the rate for SU equipped with MIMO antennas when the instantaneous or statistical CSI from SU-Tx to PU-Rx is unavailable or imperfect, under the metric of interference temperature IT or the leakage rate LR to ensure the QoS of PU. When no CSI is available an iterative

adaptation algorithm is proposed that uses side-information in the primary network for the guarantee of QoS in reaching a solution. For the imperfect CSI scenario, we have utilized the maximin strategy to obtain a robust solution through the dual norm of the Schatten norm that defines the amount of CSI uncertainty. Further, we have shown that the leakage metric favors a higher rate for PU while the IT metric for SU. In addition to the design solutions, we have evaluated a lower bound for the rate of SU when the interference CSI is completely unavailable to the SU. We have also completed the study by deriving the conditions on the uncertainty radius for which the imperfect CSI solution is not better than the unavailable CSI one for the IT metric. Computer simulations verified and supported the proposed techniques and the theoretical developments.

2.8 Appendix

2.8.1 Optimal Solution of P1 with Perfect CSI

For interference-limited systems, the power constraint in P1 is inactive and can be ignored. [113, Theorem 1] gives the general form of the optimal covariance matrix $\check{\mathbf{Q}}_s$ for P1, which is

$$\check{\mathbf{Q}}_s = \mathbf{R}_H^{-1/2} \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\dagger \mathbf{R}_H^{-1/2} \quad (2.39)$$

where $\mathbf{R}_H = \mathbf{H}_s^\dagger \mathbf{H}_s$ is a full rank matrix, $\mathbf{U} \in \mathbb{C}^{M_s \times M_s}$ is a unitary matrix, and $\mathbf{\Sigma}$ is an $M_s \times M_s$ matrix whose diagonal entries are arranged in non-increasing order. Furthermore, according to [113, Theorem 2] \mathbf{U} is the eigenvectors matrix of the EVD of $\mathbf{R}_H^{-1/2} \mathbf{R}_G \mathbf{R}_H^{-1/2} = \mathbf{U}_R \mathbf{\Lambda}_R \mathbf{U}_R^\dagger$, where the diagonal elements of $\mathbf{\Lambda}_R$ are arranged in

increasing order. Thus, P1 under the IT metric can be formulated to be

$$\max_{\mathbf{\Sigma} \succeq \mathbf{0}} \log |\mathbf{I}_{M_s} + \mathbf{\Sigma}/\sigma_s^2| \quad (2.40a)$$

$$\text{s.t. } \text{Tr}(\mathbf{\Lambda}_{\mathbf{R}}\mathbf{\Sigma}) \leq I_T. \quad (2.40b)$$

Accordingly, the i th diagonal element of $\mathbf{\Sigma}$ can be found using waterfilling as

$$\Sigma(i, i) = \left(\frac{1}{\beta \mathbf{\Lambda}_{\mathbf{R}}(i, i)} - \sigma_s^2 \right)^+ \quad (2.41)$$

where β is the water-level chosen such that (2.40b) is satisfied with equality. From (2.40b) we can have a closed form expression for β ,

$$I_T = \sum_{j=1}^{\check{r}} \Sigma(j, j) \mathbf{\Lambda}_{\mathbf{R}}(j, j) = \sum_{j=1}^{\check{r}} \left[\left(\frac{1}{\beta \mathbf{\Lambda}_{\mathbf{R}}(j, j)} - \sigma_s^2 \right) \mathbf{\Lambda}_{\mathbf{R}}(j, j) \right] \quad (2.42)$$

$$\beta = \frac{\check{r}}{I_T + \sigma_s^2 \sum_{j=1}^{\check{r}} \mathbf{\Lambda}_{\mathbf{R}}(j, j)} \quad (2.43)$$

where (2.41) is used to get (2.42) and \check{r} is the optimal rank. Plugging (2.43) in (2.41) and substituting the result into (2.40a) yield the optimal rate \check{C}_s

$$\check{C}_s = \sum_{i=1}^{\check{r}} \log \left[\frac{I_T + \sigma_s^2 \sum_{j=1}^{\check{r}} \mathbf{\Lambda}_{\mathbf{R}}(j, j)}{\check{r} \sigma_s^2 \mathbf{\Lambda}_{\mathbf{R}}(i, i)} \right] \quad (2.44)$$

The optimal rank can be calculated by applying the following notion

$$\check{r} = \arg \max_m \Sigma(m, m), \quad (2.45)$$

provided that $\check{r} > 0$. Substituting (2.41) and (2.43) into (2.45) yields

$$\check{r} = \arg \max_m \left[\frac{I_T + \sigma_s^2 \sum_{j=1}^m \Lambda_{\mathbf{R}}(j, j)}{m \Lambda_{\mathbf{R}}(m, m)} - \sigma_s^2 \right]. \quad (2.46)$$

2.8.2 Proof of Proposition 1

Lemma 1 [42]: Let $\mathbf{A}, \mathbf{B} \in \mathbb{H}^N$ having eigenvalues $\gamma_1 \geq \dots \geq \gamma_N$ and $\delta_1 \geq \dots \geq \delta_N$. Then, $\text{Tr}(\mathbf{A}\mathbf{B}) \leq \sum_{i=1}^N \gamma_i \delta_i$ with the equality holds if \mathbf{A} and \mathbf{B} are diagonal.

For convenience, we first write (2.6b) using a unified expression for both IT and LR metrics. Beginning from (2.4), the leakage rate metric is upper bounded by

$$\log |\sigma_p^2 \mathbf{I}_{M_s} + \mathbf{R}_G^{1/2} \mathbf{Q}_s \mathbf{R}_G^{1/2}| - R_L \leq K \log(\sigma_p^2 + \text{Tr}(\mathbf{R}_G \mathbf{Q}_s)/K) - R_L \quad (2.47)$$

using the Jensen's inequality, where $K = \min(M_s, N_p)$. If we set the right side of (2.47) to be no bigger than 0, i.e.,

$$\text{Tr}(\mathbf{R}_G \mathbf{Q}_s) \leq K(e^{R_L/K} - \sigma_p^2), \quad (2.48)$$

(2.6b) will be guaranteed. From (2.48), we can obtain similar expression for the IT metric.

Assume \mathbf{Q}_s has a rank r , then applying Lemma 1 on the left side of (2.48) gives (2.11), where I_{TL} represents either I_T or the right side of (2.48). From the waterfilling solution of $\Lambda_{\mathbf{Q}_s}$ in (2.9), (2.11) can be written as

$$\text{Tr}(\Lambda_{\mathbf{R}_G, r}(\nu^{-1} \mathbf{I}_r - \sigma_s^2 \Lambda_{\mathbf{H}_s, r}^{-1})^+) \leq I_{TL}, \quad (2.49)$$

where $\Lambda_{\mathbf{R}_{G,r}}$ is assumed to be known, then P2 can be solved numerically under the interference-limited scenario. The r value needs to be chosen such that ν achieves (2.11) or (2.49) with equality.

The expressions in (2.10b) and (2.11) have to meet the same interference level I_{TL} . The constraint (2.11), however, is a sufficient condition for (2.10b) and would reduce the feasible set of the solution. Thus, it becomes clear that the rate corresponding to the precoder solution of P2 under (2.11) is no more than that from P2 under (2.10b).

2.8.3 Proof of Proposition 2

When the number of receiving antennas at PU-Rx N_p grows large while M_s is being fixed, we have from the asymptotic theory of complex Gaussian random matrix that

$$\lim_{N_p \rightarrow \infty} \mathbf{G}_s^\dagger \mathbf{G}_s / N_p = \mathbf{I}_{M_s}. \quad (2.50)$$

Using (2.50), the interference constraint (2.6b) in P1 is reduced to $\text{Tr}(\mathbf{Q}_s) \leq I_T / N_p$ for the IT metric or $\log |\sigma_p^2 \mathbf{I}_{M_s} + N_p \mathbf{Q}_s| \leq R_L$ for the LR one. Then the solution of P1 under known \mathbf{G}_s in Appendix 2.8.1 reduces to the one using the waterfilling technique as described in (2.8) for P2. In such a scenario, P1 and P2 will yield the same precoder solution and the P2 precoder is optimal.

2.8.4 Proof of Proposition 3

Lemma 2 (Hölder inequality) [13, Proposition 9.1.6]: Let $p, q \in [1, \infty]$ such that $1/p + 1/q = 1$, and let $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ with elements a_1, \dots, a_n and b_1, \dots, b_n . Then,

$$|\mathbf{a}^\dagger \mathbf{b}| \leq \|\mathbf{a}\|_p \|\mathbf{b}\|_q, \quad (2.51)$$

where $\|\cdot\|_p$ is the Hölder norm [13, Proposition 9.1.4]

$$\|\mathbf{a}\|_p = \begin{cases} (\sum_{i=1}^n |a_i|^p)^{1/p}, & 1 \leq p < \infty \\ \max_{i \in \{1, \dots, n\}} |a_i|, & p = \infty. \end{cases} \quad (2.52)$$

Equality in (2.51) holds if and only if

$$\begin{cases} |a_i| |b_i| = \|\mathbf{b}\|_\infty |a_i|, & p = 1, & (2.53a) \\ \frac{|a_1|^{1/q}}{|b_1|^{1/p}} = \dots = \frac{|a_n|^{1/q}}{|b_n|^{1/p}}, & 1 < p < \infty, & (2.53b) \\ |a_i| |b_i| = \|\mathbf{a}\|_\infty |b_i|, & p = \infty. & (2.53c) \end{cases}$$

We first rewrite P3-I as follows

$$\max_{\Delta \mathbf{R}_G} \text{Tr}((\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G) \mathbf{Q}_s) - I_T \quad (2.54a)$$

$$\text{s.t. } \|\Delta \mathbf{R}_G\|_{S_p} \leq \epsilon \quad (2.54b)$$

$$\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G \succeq 0. \quad (2.54c)$$

Then, we form the Lagrangian for P3-I and find the dual function $\mathcal{G}(\cdot)$ as follows

(ignore the constant I_T)

$$\mathcal{G}(\mathbf{Y}, \eta) = \sup_{\Delta \mathbf{R}_G \in \mathbb{H}^{M_s}} \text{Tr}((\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G) \mathbf{Q}_s) - \eta(\|\Delta \mathbf{R}_G\|_{S_p} - \epsilon) + \text{Tr}((\hat{\mathbf{R}}_G + \Delta \mathbf{R}_G) \mathbf{Y}) \quad (2.55)$$

$$= \sup_{\Delta \mathbf{R}_G \in \mathbb{H}^{M_s}} \mathcal{L}_{\Delta \mathbf{R}_G} + \text{Tr}(\hat{\mathbf{R}}_G(\mathbf{Q}_s + \mathbf{Y})) + \eta\epsilon \quad (2.56)$$

where $\eta \geq 0$ and $\mathbf{Y} \succeq 0$ are the Lagrange multipliers associated with the constraints (2.54b) and (2.54c) and $\mathcal{L}_{\Delta \mathbf{R}_G} \triangleq \text{Tr}(\Delta \mathbf{R}_G(\mathbf{Q}_s + \mathbf{Y})) - \eta\|\Delta \mathbf{R}_G\|_{S_p}$. To find the supremum of $\mathcal{L}_{\Delta \mathbf{R}_G}$ in (2.56), let the EVDs of $\Delta \mathbf{R}_G$ and $\mathbf{J} \triangleq \mathbf{Q}_s + \mathbf{Y}$ be $\mathbf{U}_\Delta \mathbf{\Lambda}_\Delta \mathbf{U}_\Delta^\dagger$ and $\mathbf{U}_J \mathbf{\Lambda}_J \mathbf{U}_J^\dagger$, respectively. Using Lemma 1, we obtain

$$\sup_{\Delta \mathbf{R}_G \in \mathbb{H}^{M_s}} \mathcal{L}_{\Delta \mathbf{R}_G} = \sup_{\mathbf{\Lambda}_\Delta} \text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) - \eta\|\mathbf{\Lambda}_\Delta\|_{S_p} \quad (2.57)$$

since $\text{Tr}(\Delta \mathbf{R}_G \mathbf{J}) \leq \text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J)$ and equality is achieved when $\mathbf{U}_\Delta = \mathbf{U}_J$.

Next let $\mathcal{L}_{\mathbf{\Lambda}_\Delta} = \text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) - \eta\|\mathbf{\Lambda}_\Delta\|_{S_p}$ and consider the following two cases:

1. $\text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) \leq 0$: from (2.57) we can conclude that $\sup_{\mathbf{\Lambda}_\Delta} \mathcal{L}_{\mathbf{\Lambda}_\Delta} \leq 0$ and equality is achieved when $\mathbf{\Lambda}_\Delta = 0$.

2. $\text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) > 0$: we apply Lemma 2 and the Schattern norm definition on $\mathcal{L}_{\mathbf{\Lambda}_\Delta}$ to obtain,

$$\mathcal{L}_{\mathbf{\Lambda}_\Delta} \leq \left(\sum_{i=1}^{M_s} |\mathbf{\Lambda}_\Delta(i, i)|^p \right)^{1/p} \left(\sum_{i=1}^{M_s} \mathbf{\Lambda}_J^q(i, i) \right)^{1/q} - \eta \left(\sum_{i=1}^{M_s} |\mathbf{\Lambda}_\Delta(i, i)|^p \right)^{1/p} \quad (2.58a)$$

$$= \|\mathbf{\Lambda}_\Delta\|_{S_p} (\|\mathbf{Q}_s + \mathbf{Y}\|_{S_q} - \eta) \quad (2.58b)$$

The constraint in (2.54b) is always active at the optimal solution for P3-I since we use the maximin principle [18] that seeks the most conservative precoder design from P1 at the worst uncertainty matrix from P3-I. Thus in (2.57), the value of the second

term is $\eta\|\mathbf{\Lambda}_\Delta\|_{S_p} = \eta\|\Delta\mathbf{R}_G\|_{S_p} = \eta\epsilon$ and would not be affected by bounding the first term. The equality in (2.58a) holds provided that the eigenvalues of $\Delta\mathbf{R}_G$ must satisfy (2.53) for given eigenvalues of \mathbf{J} . To this end, optimizing $\mathcal{L}_{\mathbf{\Lambda}_\Delta}$ in (2.57) would produce the following values:

$$\sup_{\Delta\mathbf{R}_G \in \mathbb{H}^{M_s}} \mathcal{L}_{\Delta\mathbf{R}_G} = \begin{cases} 0, & \|\mathbf{Q}_s + \mathbf{Y}\|_{S_q} \leq \eta \\ \infty, & \|\mathbf{Q}_s + \mathbf{Y}\|_{S_q} > \eta. \end{cases} \quad (2.59)$$

The dual function $\mathcal{G}(\mathbf{Y}, \eta)$ in (2.56) is then bounded from above only when $\|\mathbf{Q}_s + \mathbf{Y}\|_{S_q} \leq \eta$. Consequently, it can be written as

$$\mathcal{G}(\mathbf{Y}, \eta) = \begin{cases} \text{Tr}(\hat{\mathbf{R}}_G(\mathbf{Q}_s + \mathbf{Y})) + \eta\epsilon, & \text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) \leq 0 \\ \text{Tr}(\hat{\mathbf{R}}_G(\mathbf{Q}_s + \mathbf{Y})) + \eta\epsilon, & \text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) > 0. \end{cases} \quad (2.60)$$

To proceed further, we shall minimize the dual function with respect to the Lagrange multipliers. That is,

$$\inf_{\mathbf{Y} \succeq 0, \eta \geq 0} \mathcal{G}(\mathbf{Y}, \eta). \quad (2.61)$$

Clearly, the solution of (2.61) when $\text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) \leq 0$ is trivial as $\Delta\mathbf{R}_G = 0$. For the case $\text{Tr}(\mathbf{\Lambda}_\Delta \mathbf{\Lambda}_J) > 0$, the complementary slackness for the constraint (2.54b), i.e., $\eta(\|\Delta\mathbf{R}_G\|_{S_p} - \epsilon) = 0$, forces η to be positive with a minimum value of $\|\mathbf{Q}_s + \mathbf{Y}\|_{S_q}$ as (2.59) indicates. Thus, we obtain (2.19) as an alternative expression for P3-I.

In what follows we shall prove that $\mathbf{Y} = 0$ is the optimal solution for the problem defined in (2.19). Since the $\text{Tr}(\cdot)$ function is linear, the first term in (2.19) will be at the minimum when $\mathbf{Y} = 0$. Also, as $\mathbf{Q}_s + \mathbf{Y} \succeq \mathbf{Q}_s$, then from Definition 8.6.12

and Proposition 8.6.13 in [13] we obtain $\|\mathbf{Q}_s + \mathbf{Y}\|_{S_q} \geq \|\mathbf{Q}_s\|_{S_q}$, and again $\mathbf{Y} = \mathbf{0}$ minimizes the second term of (2.19).

The subproblem P3-I and (2.19) are dual and solving either one will yield the same optimal solution. The strong duality between them can be established by applying the Slater's condition [18]. To show this it is sufficient to find a value of $\Delta\mathbf{R}_G$ that makes P3-I strictly feasible, i.e., $\|\Delta\mathbf{R}_G\|_{S_p} < \epsilon$ and $\hat{\mathbf{R}}_G + \Delta\mathbf{R}_G \succ \mathbf{0}$. Following a similar approach as in [43], let $\Delta\mathbf{R}_G$ be having the same subspace as $\hat{\mathbf{R}}_G$ whose EVD is represented by $\hat{\mathbf{R}}_G = \mathbf{U}_{\hat{\mathbf{R}}_G} \mathbf{\Lambda}_{\hat{\mathbf{R}}_G} \mathbf{U}_{\hat{\mathbf{R}}_G}^\dagger$. Also, let the diagonal eigenvalues matrix of $\Delta\mathbf{R}_G$ be $\mathbf{\Lambda}_\Delta$. Now since $(\mathbf{\Lambda}_{\hat{\mathbf{R}}_G}(i, i) + \mathbf{\Lambda}_\Delta(i, i))$ has to be positive for $i = 1, 2, \dots, K = \min(M_s, N_p)$, the subproblem P3-I would be strictly feasible when $|\mathbf{\Lambda}_\Delta(i, i)| = \min(c\mathbf{\Lambda}_{\hat{\mathbf{R}}_G}(i, i), \alpha)$ where $c < 1$ and $0 < \alpha < \epsilon/\sqrt[3]{K}$.

2.8.5 Waterfilling Solutions

For the IT metric P3 has a channel-diagonalizing solution when the power constraint is inactive and Γ_{\max} in (2.13b) is equal to (2.21). Using the procedure of Appendix A, the waterfilling solution can be derived similarly by applying Theorem 1 [113] to obtain the general structure of the precoder and Theorem 2 to find the \mathbf{U} matrix.

When using the LR metric, we simply need to use the same procedure as in Appendix A to obtain the waterfilling solution for each iteration.

Chapter 3

Precoding for MIMO Channels in Cognitive Radio Networks With CSI Uncertainties and for Compound Capacity

3.0.1 Introduction

Inaccurate knowledge of CSI limits the performance offered by MIMO communications. The design of a MIMO precoder should take the CSI uncertainty into consideration to mitigate its effects. This chapter proposes a method to design the precoder for a secondary user in an underlay cognitive radio framework that maximizes its transmission performance, where the direct link to its receiver and the interference link to a primary user have CSI uncertainties. We model the CSI uncertainties through the Schatten norm. The proposed method solves iteratively a minimax problem by deriving the optimal solution for the worst case interference CSI uncertainty, applying

the alternate-iterate technique for the worst case direct link CSI deviation, and developing a feasible direction projected subgradient technique for the precoder. Simpler solutions for the precoder are also derived under some specific norm measure of CSI uncertainty and certain requirement of transmission power. Simulations corroborate the expected performance of the proposed design.

3.0.2 Background

This chapter considers an underlay CR point-to-point MIMO system in which an SU seeks communication in the presence of a PU where its QoS should be assured. The available CSI of the direct link between SU-Tx and SU-Rx and the interference link between SU-Tx and PU-Rx is not accurate. The CSI error in the direct link is attributed to the short coherence time, limited feedback or significant demand of spectrum resources by PUs [24]. The uncertainty in interference CSI results from the non-cooperative nature of PU with SU since both users often belong to different networks [8, 24–26, 52, 76, 92, 93, 103, 109]. The objective is to design a linear precoder for SU-Tx to improve the SU transmission performance that is robust to the CSI uncertainties, while ensuring the QoS for PU. The underlay CR paradigm is less demanding in managing interferences than some other approaches, such as interference alignment [26, 109] which requires the availability of CSI for all link connections and a precoder at each transmitter and a combiner at each receiver ends.

This work uses a deterministic model for the uncertainties and provides a novel solution to the precoder. The precoder design is however rather involved as it requires the solution of a maximin (or the equivalent minimax) optimization problem.

Essentially, we are interested in searching a precoder for Gaussian modelled data

that maximizes MI of SU, under its most conservative performance caused by the channel uncertainties in the SU intended connection and in the SU to PU interference link. The optimization is non-concave over the maximization on the precoder and non-convex over the minimization of the SU's MI, although it is concave over the maximization of the amount of interference to PU. Such a mixed optimization problem does not lend itself to apply the existing techniques in [65,104] or the matrix inequalities introduced in [49,82]. This chapter formulates the problem as a continuous maximin optimization [74] and solves one variable at a time while holding the others fixed. We first derive an explicit solution of the worst case CSI deviation that would cause the largest amount of interference from SU to PU. Next, by introducing a nuisance parameter, we obtain the CSI uncertainty value that yields the least MI at a given precoder through the alternate-iterate (AI) technique [33]. Finally, the precoder is then updated according to the feasible direction determined by the projection of a subgradient [14]. The optimization process repeats until a stationary solution is reached. We support the proposed method by providing the convergence analysis for reaching a stationary point.

It is critical to maintain the QoS of PU in CR. Such a constraint, nevertheless, creates a non-convex feasible set for the precoder that would complicate the optimization. To handle this difficulty, we express the QoS measure in the interference constraint through the Taylor-series and maintain up to the first order. Such an expansion provides more conservative requirement on the QoS. As the optimization cycles through, the linear approximation would be more accurate and eventually represents the actual constraint at a stationary solution.

The CR problem we are addressing encompasses another interesting problem in

the literature in which PU is absent that the QoS constraint is not needed. This is termed as the compound capacity (CC) problem [48, 49, 65, 82, 104]. The proposed solution is applicable to the CC problem as well except the optimization of CSI deviation on the interference link is not necessary. The optimization of the precoder is in fact simpler as its feasible set is convex.

3.0.3 Chapter Organization

Chapter three is organized as follows. The system model and problem formulation are introduced in Section 3.1. Section 3.2 develops an algorithm to solve the CR problem and proves the convergence of the proposed solution method. Section 3.3 briefly summarizes the reduced solution to the CC problem. In Section 3.4, particular solutions for the CR and CC problems are established. Section 3.5 provides the numerical results and Section 3.6 gives the conclusion.

3.1 System Model and Assumptions

Let us consider the scenario as shown in Figure 3.1 in which an SU communicates through the MIMO channel $\mathbf{H}_{22} \in \mathbb{C}^{N_2 \times M_2}$ in the presence of a PU, where M_2 and N_2 are the number of SU transmit and receive antennas. The signal from SU interferes with PU through the channel $\mathbf{H}_{21} \in \mathbb{C}^{N_1 \times M_2}$, where N_1 is the number of PU-Rx antennas. The channels are not known exactly due to limited feedback [24, 48] and the lack of coordination between PU and SU [1, 2, 36, 92, 93, 103]. The objective is to obtain a linear precoder for SU that maximizes the MI for its transmission where the available channel information has uncertainty, subject to the requirement

of maintaining the QoS of PU.

It is more direct here to represent the CSI in terms of the covariance matrices \mathbf{R}_{2j} , $j = 1, 2$. It is related to the MIMO channels by $\mathbf{R}_{2j} = \mathbf{H}_{2j}^\dagger \mathbf{H}_{2j}$ for slow fading and $\mathbf{R}_{2j} = \mathbb{E} \left[\mathbf{H}_{2j}^\dagger \mathbf{H}_{2j} \right]$ for fast fading. The available CSI at SU-Tx is $\hat{\mathbf{R}}_{2j}$. It is not accurate and is different from the actual by

$$\hat{\mathbf{R}}_{2j} = \mathbf{R}_{2j} + \mathbf{\Delta}_{2j}, \quad j = 1, 2. \quad (3.1)$$

$\mathbf{\Delta}_{2j} \in \mathbb{C}^{M_2 \times M_2}$ represents the error in acquiring CSI.

We model the channel acquisition error as deterministic using the non-empty compact and convex set \mathcal{U}_j , $j = 1, 2$, whose shape is defined by the Schatten norm¹ [2, 102],

$$\mathcal{U}_j = \{ \mathbf{\Delta}_{2j} : \|\mathbf{\Delta}_{2j}\|_{S_{p_j}} \leq \epsilon_j, \hat{\mathbf{R}}_{2j} - \mathbf{\Delta}_{2j} \succeq 0 \}, j = 1, 2. \quad (3.2)$$

The parameter $\epsilon_j > 0$ controls the size of the uncertainty set and $\|\cdot\|_{S_{p_j}}$ denotes the Schatten norm of order $p_j \in [1, \infty]$.

In (3.2), the requirement $\hat{\mathbf{R}}_{2j} - \mathbf{\Delta}_{2j} \succeq 0$ ensures the actual CSI \mathbf{R}_{2j} is PSD and it implicitly restricts $\mathbf{\Delta}_{2j}$ to be a Hermitian matrix.

The linear precoder we would like to obtain is represented by the matrix $\mathbf{Q} \in \mathbb{C}^{M_2 \times M_2}$. It processes the SU signal before transmission to improve the performance.

¹The developed techniques and solutions in this work are not limited to the Schatten norm based uncertainty set. They can be easily extended to any non-empty compact and convex uncertainty set.

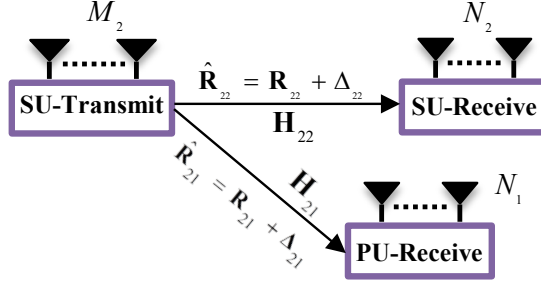


Figure 3.1: An underlay MIMO CR network with single SU and single PU. CSI for the SU direct link and for the interference link from SU to PU is not accurate.

The metric for maximization is

$$\phi(\mathbf{Q}, \mathbf{\Delta}_{22}) \triangleq \log |\mathbf{I} + \mathbf{R}_{22}\mathbf{Q}| \quad (3.3a)$$

$$= \log |\mathbf{I} + (\hat{\mathbf{R}}_{22} - \mathbf{\Delta}_{22})\mathbf{Q}| \quad (3.3b)$$

where $\log(\cdot)$ is the natural logarithm. (3.3b) resembles the MI of SU. It is indeed the MI for slow fading, and is the upper bound for the MI under fast fading where \mathbf{R}_{22} is defined with expectation, i.e., $\mathbf{R}_{22} = \mathbb{E} [\mathbf{H}_{22}^\dagger \mathbf{H}_{22}]$. Please refer to Appendix 3.7.1 for the details. Precoding optimization based on a tractable bound of a performance metric is common in CR with single [52, 76] or multiple antennas [28, 56]. The metric (3.3b) defines the information rate over the SU transmission link and henceforth we shall call $\phi(\cdot, \cdot)$ as the SU rate. The SU rate in (3.3) assumes the SU signal follows a Gaussian distribution [31] so that it is accurate and has a simple form. In practice the distribution of SU signal may be different. For finite-alphabet signals such as those created from quadrature amplitude modulation (QAM), the rate expression should be defined using the constellation-constraint MI instead [108]. In such a case, it will lead to a different solution for the precoder.

The precoder matrix \mathbf{Q} is not a free variable. It should be Hermitian and PSD,

and should satisfy the power budget defined by its elements or eigenvalues. We shall use a convex set \mathcal{Q} to represent these requirements together.

The maximization of the SU rate is under the constraint that the interference caused by SU-Tx to PU-Rx is under a certain limit. The interference level is quantified by

$$g(\mathbf{Q}, \mathbf{\Delta}_{21}) \triangleq \log |\mathbf{I} + \mathbf{R}_{21}\mathbf{Q}| \quad (3.4a)$$

$$= \log |\mathbf{I} + (\hat{\mathbf{R}}_{21} - \mathbf{\Delta}_{21})\mathbf{Q}|. \quad (3.4b)$$

The measure (3.4b) is often referred as the leakage rate [71] and was recently investigated in [1,2] for the CR design.

It would be more convenient to use $-\phi(\cdot, \cdot)$ instead of $\phi(\cdot, \cdot)$ for the development in the sequel. In mathematical form, the problem we would like to address can be posed as

$$(P-I): \min_{\mathbf{Q} \in \mathcal{Q}} \max_{\mathbf{\Delta}_{22} \in \mathcal{U}_2} -\phi(\mathbf{Q}, \mathbf{\Delta}_{22}) \quad (3.5a)$$

$$\text{s.t.} \quad \max_{\mathbf{\Delta}_{21} \in \mathcal{U}_1} g(\mathbf{Q}, \mathbf{\Delta}_{21}) \leq r_L. \quad (3.5b)$$

Note that (3.5a) is equivalent to $\max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{\Delta}_{22} \in \mathcal{U}_2} \phi(\mathbf{Q}, \mathbf{\Delta}_{22})$. We seek a conservative precoder that guarantees performance improvement as much as possible under the worst transmission scenario for SU. (3.5b) ensures the interference tolerance limit $r_L > 0$ is not exceeded for all possible CSI uncertainties expected in the interference link.

P-I is a constrained continuous minimax problem and it is rather difficult to solve for several reasons. First, the objective function $-\phi(\mathbf{Q}, \mathbf{\Delta}_{22})$ is not concave in $\mathbf{\Delta}_{22}$.

Second, the interference constraint (3.5b) is not convex in \mathbf{Q} . Third, the optimization on the left side of (3.5b) does not have a closed form expression in terms of \mathbf{Q} , unlike the studies in [2, 95, 98].

In addition to the CR problem, P-I becomes the CC problem [65] if the constraint (3.5b) is absent. That is, we only consider the transmission of SU without worrying the QoS of PU. The CC problem has found interests in the research community as well [48, 49, 65, 82, 104].

3.2 Cognitive Radio Problem

We would like to solve P-I over the optimization variables Δ_{21} , Δ_{22} and \mathbf{Q} . They are coupled together and it is extremely difficult for joint optimization. Rather, we resort to the approach similar to coordinate descent by iteratively optimizing over the variables one by one in sequence; while optimizing in one variable the others are held fixed at their latest values. Let us assume that after the previous iteration we have obtained the precoder solution $\mathbf{Q}^{(l)}$, where l denotes the iteration count.

3.2.1 Solution for Δ_{21}

Utilizing the interference constraint (3.5b) requires the optimal value of the subproblem

$$(P-I.1): \max_{\Delta_{21} \in \mathcal{U}_1} g(\Delta_{21}, \mathbf{Q}). \quad (3.6)$$

P-I.1 is concave in Δ_{21} for a given precoder.

To proceed, we first introduce a proposition for the set of possible solutions for

Δ_{21} .

Proposition 1: For a given precoder matrix \mathbf{Q} the optimal uncertainty matrix Δ_{21} to the problem P-I.1 must be NSD.

Proof: The key idea for the proof is to show that the function $g(\cdot, \cdot)$ in P-I.1 would have a larger value when Δ_{21} is NSD than PSD. Please refer to Appendix 3.7.2 for the details. ■

As a result, $\hat{\mathbf{R}}_{21} - \Delta_{21} \succeq 0$ is automatically satisfied and does not need to be included in the set \mathcal{U}_1 .

We shall derive an analytic solution of Δ_{21} for the problem (3.6) at $\mathbf{Q} = \mathbf{Q}^{(l)}$. Note that Δ_{21} does not depend on Δ_{22} explicitly. Let us begin by defining the matrix $\mathbf{T}^{(l)} \triangleq \mathbf{Q}^{(l)}(\mathbf{I} + \hat{\mathbf{R}}_{21}\mathbf{Q}^{(l)})^{-1}$, which is PSD according to Lemma 2 in Appendix 3.7.3. It is direct to verify the subproblem P-I.1 in (3.6) at $\mathbf{Q} = \mathbf{Q}^{(l)}$ is equivalent to

$$\text{(P-I.2): } \max_{-\Delta_{21} \succeq 0} \log |\mathbf{I} - \Delta_{21} \mathbf{T}^{(l)}| \quad (3.7a)$$

$$\text{s.t. } \|\Delta_{21}\|_{S_{p_1}}^{p_1} \leq \epsilon_1^{p_1}. \quad (3.7b)$$

The solution to P-I.2 is obtained through the EVD representations of $\mathbf{T}^{(l)}$ and Δ_{21} . Let $\mathbf{T}^{(l)} = \mathbf{U}_{\mathbf{T}^{(l)}} \mathbf{\Lambda}_{\mathbf{T}^{(l)}} \mathbf{U}_{\mathbf{T}^{(l)}}^\dagger$ with eigenvalues $\lambda_1(\mathbf{T}^{(l)}) \geq \dots \geq \lambda_{M_2}(\mathbf{T}^{(l)})$ and $\Delta_{21} = \mathbf{U}_{\Delta_{21}} \mathbf{\Lambda}_{\Delta_{21}} \mathbf{U}_{\Delta_{21}}^\dagger$ with eigenvalues $-\lambda_1(\Delta_{21}) \geq \dots \geq -\lambda_{M_2}(\Delta_{21})$. From the Hadamard's inequality [57, Ch. 9, B.5] that states $|\mathbf{A}| \leq \prod_{i=1}^n [\mathbf{A}]_{ii}$ with $\mathbf{A} \in \mathbb{C}^{n \times n}$ being Hermitian and PSD, the objective function (3.7a) is maximized when $\mathbf{U}_{\Delta_{21}} = \mathbf{U}_{\mathbf{T}^{(l)}}$.

Introducing the Lagrange multiplier $\mu \geq 0$ for the constraint (3.7b), the Lagrange

dual for P-I.2 under $\mathbf{U}_{\Delta_{21}} = \mathbf{U}_{\mathbf{T}^{(l)}}$ is

$$\max_{[\mathbf{\Lambda}_{\Delta_{21}}]_{ii} \geq 0} \sum_{i=1}^{M_2} \log(1 + |\lambda_i(\Delta_{21})| |\lambda_i(\mathbf{T}^{(l)})|) - \mu \left(\sum_{i=1}^{M_2} |\lambda_i(\Delta_{21})|^{p_1} - \epsilon_1^{p_1} \right). \quad (3.8)$$

Taking derivative with respect to $\lambda_i(\Delta_{21})$ and setting it to zero give

$$\lambda_i(\Delta_{21}) = - \left(\frac{1}{\mu p_1 |\lambda_i(\Delta_{21})|^{p_1-1}} - \frac{1}{\lambda_i(\mathbf{T}^{(l)})} \right)^+ \quad (3.9)$$

for $p_1 \in [1, \infty)$, where the value $x^+ = \max(x, 0)$. For a given μ , we can solve (3.9) for $\lambda_i(\Delta_{21})$ through a simple line search. Applying another level of line search yields μ such that $\sum_{i=1}^{M_2} |\lambda_i^{p_1}(\Delta_{21})| = \epsilon_1^{p_1}$.

For $p_1 = \infty$, it is straightforward to validate that $\Delta_{21} = -\epsilon_1 \mathbf{I}$ is optimal since the constraint (3.7b) is equivalent to $-\Delta_{21} \preceq \epsilon_1 \mathbf{I}$.

Putting together the eigenvectors and the eigenvalues, we have the solution to P-I.2 and it is denoted as $\Delta_{21}^{(l)}$.

We next obtain the solutions for Δ_{22} and \mathbf{Q} . For ease of illustration, (3.5a) in the problem P-I is expressed as two sub-problems

$$(P-I.3): \quad \Psi(\mathbf{Q}) \triangleq \max_{\Delta_{22} \in \mathcal{U}_2} -\phi(\mathbf{Q}, \Delta_{22}), \quad (3.10)$$

$$(P-I.4): \quad \min_{\mathbf{Q} \in \mathcal{Q}} \Psi(\mathbf{Q}). \quad (3.11)$$

P-I.3 is not concave in Δ_{22} at a given \mathbf{Q} and P-I.4 is convex in \mathbf{Q} over the set \mathcal{Q} .

3.2.2 Solution for Δ_{22}

Using a similar procedure as for Proposition 1 in Appendix 3.7.2, we can show that Δ_{22} should be PSD for the solution to P-I.3. Incorporating this result, the uncertainty set for Δ_{22} in (3.2) becomes

$$\tilde{\mathcal{U}}_2 = \{\Delta_{22} : \|\Delta_{22}\|_{S_{p_2}} \leq \epsilon_2, \hat{\mathbf{R}}_{22} - \Delta_{22} \succeq 0, \Delta_{22} \succeq 0\}. \quad (3.12)$$

Given the intermediate solution for the precoder $\mathbf{Q}^{(l)}$, the objective is to find $\Delta_{22} \in \tilde{\mathcal{U}}_2$ that maximizes $-\phi(\mathbf{Q}^{(l)}, \Delta_{22})$. We shall make use of the following lemma to convert $-\phi(\cdot, \cdot)$ to a different and yet equivalent form to simplify the solution finding.

Lemma 1: Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be any PD matrix (not necessarily Hermitian). Consider the function $z(\mathbf{S}) = -\text{Tr}(\mathbf{S}\mathbf{A}) + \log |\mathbf{S}| + n$, where $\mathbf{S} \in \mathbb{C}^{n \times n}$ is any PD matrix. Then

$$\max_{\mathbf{S} \succ 0} z(\mathbf{S}) = -\log |\mathbf{A}|. \quad (3.13)$$

The optimization on the left is achieved when $\mathbf{S} = \mathbf{A}^{-1}$.

Proof: It is direct to show (3.13) as $z(\mathbf{S})$ is concave and differentiable in \mathbf{S} . ■

When using the fact that Δ_{22} is PSD and applying (3.13) to the function $-\phi(\cdot, \cdot)$ defined in (3.3b), (3.10) becomes

$$\max_{\Delta_{22} \in \tilde{\mathcal{U}}_2, \mathbf{W}_2 \succ 0} -\text{Tr} \left(\mathbf{Q}^{(l)} \mathbf{W}_2 (\hat{\mathbf{R}}_{22} - \Delta_{22}) \right) + f(\mathbf{W}_2) \quad (3.14)$$

where \mathbf{W}_2 is a nuisance variable that is PD and

$$f(\mathbf{W}_2) = -\text{Tr}(\mathbf{W}_2) + \log |\mathbf{W}_2| + M_2. \quad (3.15)$$

The optimization problem (3.14) is biconcave, i.e., it is concave in $\mathbf{\Delta}_{22}$ or \mathbf{W}_2 but not in both [33]. We shall solve it based on the AI approach using the alternate convex search (ACS) method [33]. In particular, each iteration of ACS consists of the following two steps:

- Fix $\mathbf{\Delta}_{22}$ and maximize (3.14) over \mathbf{W}_2 . The resulting solution for \mathbf{W}_2 is in closed-form

$$\mathbf{W}_2 = \left(\mathbf{I} + (\hat{\mathbf{R}}_{22} - \mathbf{\Delta}_{22})\mathbf{Q}^{(l)} \right)^{-1}. \quad (3.16)$$

- Fix \mathbf{W}_2 at the value obtained in (3.16) and maximize (3.14) over $\mathbf{\Delta}_{22}$, which is uniquely determined from the following subproblem

$$(P-I.5) \quad \max_{\mathbf{\Delta}_{22} \succeq 0} \text{Tr}(\mathbf{Q}^{(l)}\mathbf{W}_2\mathbf{\Delta}_{22}) \quad (3.17a)$$

$$\text{s.t.} \quad \|\mathbf{\Delta}_{22}\|_{S_{p_2}} \leq \epsilon_2 \quad (3.17b)$$

$$\hat{\mathbf{R}}_{22} - \mathbf{\Delta}_{22} \succeq 0. \quad (3.17c)$$

The subproblem (3.17) has a linear objective function and convex constraints in $\mathbf{\Delta}_{22}$. It can be solved using the interior-point methods [15] or some software optimization packages such as CVX [34]. Alternatively, we shall propose an attractive suboptimal solution for (3.17) and provide the corresponding structure of $\mathbf{\Delta}_{22}$.

Suboptimal Solution to P-I.5

We solve (3.17) by considering an upper bound of its Lagrange dual function and denote the resulting solution as Δ_{22}^* . The following theorem defines the structure of the solution.

Theorem 1: Given that the matrix $\mathbf{Q}^{(l)}\mathbf{W}_2$ in (3.17a) is Hermitian and PSD. Let $\mathbf{F}^* \in \mathbb{C}^{M_2 \times M_2}$ be a PSD matrix obtained through (3.20) and it is Hermitian. Define the matrix $\mathbf{E}_2 \triangleq \mathbf{Q}^{(l)}\mathbf{W}_2 - \mathbf{F}^*$ whose EVD is $\mathbf{E}_2 = \mathbf{U}_{\mathbf{E}_2}\mathbf{\Lambda}_{\mathbf{E}_2}\mathbf{U}_{\mathbf{E}_2}^\dagger$ where $\lambda_1(\mathbf{E}_2) \geq \dots \geq \lambda_{M_2}(\mathbf{E}_2)$. A suboptimal solution to (3.17) that optimizes an upper bound of its Lagrange dual function is $\Delta_{22}^* = \mathbf{U}_{\Delta_{22}^*}\mathbf{\Lambda}_{\Delta_{22}^*}\mathbf{U}_{\Delta_{22}^*}^\dagger$, where

1. The eigenvectors of Δ_{22}^* are the same as those of $\mathbf{U}_{\mathbf{E}_2}$, i.e., $\mathbf{U}_{\Delta_{22}^*} = \mathbf{U}_{\mathbf{E}_2}$.
2. The eigenvalues of Δ_{22}^* , $\lambda_1(\Delta_{22}^*) \geq \dots \geq \lambda_{M_2}(\Delta_{22}^*) \geq 0$, are:
 - i. If $p_2 \in (1, \infty)$

$$\lambda_i(\Delta_{22}^*) = (|\lambda_i(\mathbf{E}_2)|/|\lambda_1(\mathbf{E}_2)|)^{\frac{q_2}{p_2}} \lambda_1(\Delta_{22}^*), \quad i = 2, \dots, M_2 \quad (3.18)$$

where p_2 is the Schatten norm order in the set $\tilde{\mathcal{U}}_2$ and $q_2 = (1 - 1/p_2)^{-1}$. The maximum eigenvalue $\lambda_1(\Delta_{22}^*)$ is

$$\lambda_1(\Delta_{22}^*) = \epsilon_2 \left[\sum_{k=1}^{M_2} (|\lambda_k(\mathbf{E}_2)|/|\lambda_1(\mathbf{E}_2)|)^{q_2} \right]^{\frac{-1}{p_2}}. \quad (3.19)$$

- ii. If $p_2 = 1$, $\lambda_1(\Delta_{22}^*) = \epsilon_2$ and $\lambda_i(\Delta_{22}^*) = 0$ for $i = 2, \dots, M_2$.
- iii. If $p_2 = \infty$, $\Delta_{22}^* = \epsilon_2 \mathbf{I}$.

Proof: Please refer to Appendix 3.7.3. ■

Corollary 1: The matrix Δ_{22}^* from Theorem 1 is the optimal solution for (3.17) if

the matrix $\mathbf{F}^* = 0$.

Proof: Please see the end of Appendix 3.7.3. ■

Corollary 2: The matrix $\mathbf{F}^* \succeq 0$ presented in Theorem 1 represents the optimal solution of the following optimization subproblem

$$\min_{\mathbf{F} \succeq 0} h(\mathbf{F}) \triangleq \epsilon_2 \|\mathbf{Q}^{(l)} \mathbf{W}_2 - \mathbf{F}\|_{S_{q_2}} + \text{Tr}(\hat{\mathbf{R}}_{22} \mathbf{F}). \quad (3.20)$$

Proof: Please refer to Appendix 3.7.4. ■

The problem (3.20) is convex in \mathbf{F} . For $q_2 \in [1, \infty)$, the function $h(\mathbf{F})$ is differentiable and hence (3.20) can be solved efficiently using the non-monotone spectral projected gradient (NMSPG) method with superlinear convergence [17]. We have included the first derivative of the function $h(\mathbf{F})$ in Appendix 3.7.4 that will be useful in obtaining the numerical results presented in Section 3.5. For $q_2 = \infty$, $h(\mathbf{F})$ is not differentiable and we reformulate the problem (3.20) as linear SDP in Appendix 3.7.4, which can be solved in polynomial time using an interior-point based method [15].

The proposed suboptimal solution for P-I.5 through Theorem 1 and (3.20) is more computationally efficient than solving (3.17) directly using the interior-point methods. The NMSPG algorithm for (3.20) has a linear time complexity and involves $h(\mathbf{F})$ and its gradient only. The interior-point methods, on the other hand, have polynomial time complexity. Furthermore, Theorem 1 yields the optimal solution for P-I.5 when the constraint (3.17c) is always satisfied.

The computation of \mathbf{F}^* from (3.20) does not depend on the solution Δ_{22}^* . Hence, using \mathbf{F}^* in \mathbf{E}_2 yields Δ_{22}^* analytically.

The AI process starts with $\Delta_{22} = \gamma_{\Delta_{22}} \mathbf{I}$ where $\gamma_{\Delta_{22}} \approx 0$, and will reach an accumulation or a stationary point [33, Theorem 4.7]. The solution of \mathbf{W}_2 is discarded

and the solution of Δ_{22} is the intermediate solution $\Delta_{22}^{(l)}$ for use to update \mathbf{Q} .

3.2.3 Solution for \mathbf{Q}

We now update \mathbf{Q} , using $\mathbf{Q}^{(l)}$, $\Delta_{21}^{(l)}$ and $\Delta_{22}^{(l)}$. The corresponding optimization problem P-I.4 is solved using a feasible descent direction method. The method uses scaled gradient projection to generate the feasible direction for updating the precoder matrix [14, Sec. 2.3]. The gradient of Ψ with respect to \mathbf{Q} may not be unique and the descent direction is obtained by a subgradient instead [60]. For simplicity, let us use $\Psi^{(l)}(\mathbf{Q})$ to denote $-\phi(\mathbf{Q}, \Delta_{22}^{(l)})$.

A direction vector at $\mathbf{Q}^{(l)}$ is

$$\mathbf{d}(\mathbf{Q}) = \text{vec}(\mathbf{Q}) - \text{vec}(\mathbf{Q}^{(l)}). \quad (3.21)$$

The update equation for \mathbf{Q} in the next iteration is

$$\text{vec}(\mathbf{Q}^{(l+1)}) = \text{vec}(\mathbf{Q}^{(l)}) + \alpha_l \mathbf{d}^{(l)} \quad (3.22)$$

where the parameter $\alpha_l \in (0, 1]$ is the step-size at the l -th iteration that will be discussed later. The feasible direction vector is $\mathbf{d}^{(l)} = \mathbf{d}(\check{\mathbf{Q}}^{(l)})$. $\check{\mathbf{Q}}^{(l)} \in \mathcal{Q}$ is obtained through scaled subgradient projection by using a quadratic approximation of $\Psi^{(l)}(\mathbf{Q})$ around $\mathbf{Q}^{(l)}$, which is equivalent to solving the following Euclidean projection problem [14, Sec. 2.3]

$$\check{\mathbf{Q}}^{(l)} = \arg \min_{\mathbf{Q} \in \mathcal{Q}} \frac{1}{2} \left\| \mathbf{d}(\mathbf{Q}) - \bar{\mathbf{d}}^{(l)}(\mathbf{Q}^{(l)}, \Delta_{22}^{(l)}) \right\|_{\mathbf{N}^{-1(l)}}^2, \quad (3.23)$$

where $\bar{\mathbf{d}}^{(l)}$ is a descent direction that may not be feasible and is given by

$$\bar{\mathbf{d}}^{(l)} = -\mathbf{N}^{(l)} \text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)})), \quad (3.24)$$

$$\Psi_g^{(l)}(\mathbf{Q}^{(l)}) = -\left(\mathbf{I} + (\hat{\mathbf{R}}_{22} - \Delta_{22}^{(l)})\mathbf{Q}^{(l)}\right)^{-1} (\hat{\mathbf{R}}_{22} - \Delta_{22}^{(l)}). \quad (3.25)$$

The matrix $\mathbf{N}^{(l)} \in \mathbb{C}^{M_2^2 \times M_2^2}$ is PD symmetric that approximates the Hessian inverse with respect to \mathbf{Q} , and the choice of it will be discussed at the end of this Section. $\Psi_g^{(l)}(\mathbf{Q}^{(l)})$ represents a subgradient of Ψ with respect to \mathbf{Q} evaluated at $\mathbf{Q}^{(l)}$, please see Lemma 4 in Appendix 3.7.5. The projection problem (3.23) can be solved using the interior-point methods [15] or an optimization package [34].

The precoder \mathbf{Q} in P-I must be solved under the interference constraint (3.5b) to guarantee the QoS of PU. The constraint makes the feasible set of \mathbf{Q} non-convex. The non-convexity in the constraint is handled through the Taylor-series expansion up to the first order. The linearized constraint is more conservative than the original [2] and the two become closer and eventually equivalent as we approach the solution.

To ensure $\mathbf{d}^{(l)}$ is a feasible direction for the problem P-I, $\check{\mathbf{Q}}^{(l)}$ is determined by the Euclidean projection (3.23) together with the interference constraint,

$$\check{\mathbf{Q}}^{(l)} \in \arg \min_{\mathbf{Q} \in \mathcal{Q}} \frac{1}{2} \|\mathbf{d}(\mathbf{Q}) - \bar{\mathbf{d}}^{(l)}(\mathbf{Q}^{(l)}, \Delta_{22}^{(l)})\|_{\mathbf{N}^{-1(l)}}^2 \quad (3.26a)$$

$$\text{s.t.} \quad \log |\mathbf{I} + \mathbf{R}_{21}^{(l)}\mathbf{Q}| \leq r_L \quad (3.26b)$$

where $\mathbf{R}_{21}^{(l)} \triangleq \hat{\mathbf{R}}_{21} - \Delta_{21}^{(l)}$. The solution of the projection problem (3.26) exists as the feasible set is closed and it may not be unique due to the non-convex constraint (3.26b). We shall obtain one possible solution out of many by applying a local approximation through linearization to the constraint (3.26b) to obtain a tractable

formulation. Expanding the left side of (3.26b) around a nominal point $\tilde{\mathbf{Q}}^{(k)}$, (3.26) becomes

$$\tilde{\mathbf{Q}}^{(k+1)} = \arg \min_{\mathbf{Q} \in \mathcal{Q}} \frac{1}{2} \|\mathbf{d}(\mathbf{Q}) - \bar{\mathbf{d}}^{(l)}(\mathbf{Q}^{(l)}, \mathbf{\Delta}_{22}^{(l)})\|_{\mathbf{N}^{-1(l)}}^2 \quad (3.27a)$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{M}^{(k)} \mathbf{Q}) \leq r^{(k)} \quad (3.27b)$$

where the variables $\mathbf{M}^{(k)}$ and $r^{(k)}$ are defined as

$$\mathbf{M}^{(k)} = \left(\mathbf{I} + \mathbf{R}_{21}^{(l)} \tilde{\mathbf{Q}}^{(k)} \right)^{-1} \mathbf{R}_{21}^{(l)}, \quad (3.28a)$$

$$r^{(k)} = r_L - \log |\mathbf{I} + \mathbf{R}_{21}^{(l)} \tilde{\mathbf{Q}}^{(k)}| + \text{Tr}(\mathbf{M}^{(k)} \tilde{\mathbf{Q}}^{(k)}). \quad (3.28b)$$

The matrix $\mathbf{M}^{(k)}$ represents the partial derivative of the left side of (3.26b) with respect to \mathbf{Q} evaluated at $\tilde{\mathbf{Q}}^{(k)}$. Note that $k = 0, 1, \dots$ here represents the iteration index for the first order approximation and (3.27) is iterated over k to reach a solution.

The proposed formulation (3.27) would guarantee a feasible solution to the sub-problem (3.26) as the feasible set of (3.27) is a subset of that in (3.26). This means that (3.27b) provides a conservative approximation of (3.26b). Hence, as k increases the precoder matrix $\tilde{\mathbf{Q}}^{(k)}$ converges to a solution for (3.26). The iteration begins with the initial value $\tilde{\mathbf{Q}}^{(0)} = \gamma_{\text{CR}} \mathbf{I}$ where $\gamma_{\text{CR}} \approx 0$ and the converged value is $\check{\mathbf{Q}}^{(l)}$.

As the feasible region for \mathbf{Q} in P-I is not convex, direct update using (3.22) is no longer valid. The solution for \mathbf{Q} requires the optimization

$$\mathbf{Q}^{(l+1)} \in \arg \min_{\mathbf{Q} \in \mathcal{Q}} \|\mathbf{d}(\mathbf{Q}) - \alpha_l \mathbf{d}^{(l)}(\check{\mathbf{Q}}^{(l)})\|^2 \quad (3.29a)$$

$$\text{s.t.} \quad \log |\mathbf{I} + \mathbf{R}_{21}^{(l)} \mathbf{Q}| \leq r_L. \quad (3.29b)$$

There could be multiple solutions to (3.29) due to the non-convex optimization problem, and applying the approach as in (3.27)-(3.28) with linearization of (3.29b) would lead to a solution.

To complete the solution finding, the details for selecting the step-size α_l and the PD matrix $\mathbf{N}^{(l)}$ are as follows.

Step-Size

We shall use an exogenous step-size that can be fixed or diminishing [14, 60]. Such a step-size ensures the convergence of (3.22) to a stationary point, as shown in Appendix 3.7.5.

Hessian Inverse

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm can yield a good estimate of the Hessian inverse as the iteration proceeds [74, Ch. 4]. To simplify the algorithm and ensure stable behavior, we simply choose \mathbf{N} to be a scalar multiple of an identity matrix by following the approach from the spectral gradient method [17]

$$\mathbf{N}^{(l)} = s^{(l)} \mathbf{I}, \quad s^{(l)} = \min \left(s_{\max}, \max \left(s_{\min}, \rho^{(l-1)} \mathbf{x}^{(l-1)\dagger} \mathbf{x}^{(l-1)} \right) \right) \quad (3.30)$$

where $s^{(l)}$ is known as the spectral step-length, $0 \leq s_{\min} \leq s_{\max} \leq \infty$ are two pre-determined constants and

$$\mathbf{x}^{(l-1)} = \text{vec}(\mathbf{Q}^{(l)}) - \text{vec}(\mathbf{Q}^{(l-1)}), \quad (3.31a)$$

$$\mathbf{y}^{(l-1)} = \text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)})) - \text{vec}(\Psi_g^{(l-1)}(\mathbf{Q}^{(l-1)})), \quad (3.31b)$$

$$\rho^{(l-1)} = 1/(\mathbf{y}^{(l-1)\dagger} \mathbf{x}^{(l-1)}). \quad (3.31c)$$

Algorithm 1 illustrates the major steps for solving P-I, through the AI and fea-

sible direction projected subgradient (FDPS) techniques. In Algorithm 1, we use a diminishing step-size rule for α_l in (3.29). The algorithm would converge to a critical point and reach a solution for \mathbf{Q} . This is shown in Appendix 3.7.5 with the aid of the Danskin's min-max Theorem [15, Proposition 4.5.1].

Algorithm 1 Solving P-I Using the AI and FDPS Techniques

Input: Provide the initial matrices $\mathbf{Q}^{(1)}$, $\mathbf{N}^{(1)}$; the Schatten norm order p_j and the error size ϵ_j for $j = 1, 2$; the LR limit r_L . Define the step-size rule for α_l ; set the iteration counter $l = 0$.

Step 1. $l = l + 1$.

Step 2. If $\mathbf{Q}^{(l)}$ satisfies a certain stopping criterion: STOP.

Step 3. Find $\Delta_{21}^{(l)}$ by the explicit solution of (3.7).

Step 4. Obtain the solution $\Delta_{22}^{(l)}$ by iterating (3.16) and (3.17).

Step 5. Determine $\bar{\mathbf{d}}^{(l)}(\mathbf{Q}^{(l)}, \Delta_{22}^{(l)})$ using (3.24), solve (3.26) for $\check{\mathbf{Q}}^{(l)}$ through (3.27).

Step 6. Evaluate α_l from a diminishing step-size rule.

Step 7. Obtain the precoder $\mathbf{Q}^{(l+1)}$ by solving (3.29) and $\mathbf{N}^{(l+1)}$ using (3.30). Go back to **Step 1**.

A typical choice for initialization is $\mathbf{Q}^{(1)} = \gamma_{\mathbf{Q}}\mathbf{I}$ and $\mathbf{N}^{(1)} = \mathbf{I}$, where $\gamma_{\mathbf{Q}} \approx 0$. We would like to clarify that the optimization problem P-I.3 shown in (3.10) is non-concave in Δ_{22} and the feasible set of \mathbf{Q} in P-I is non-convex. There could be multiple stationary points and there is no guarantee the proposed algorithm will give the global optimum solution. For better result, one can use several different initializations for $\mathbf{Q}^{(1)}$ and select the solution that gives the best objective value.

It is straightforward to extend the proposed CR solution for the presence of multiple PUs. In such a case, a separate solution for the amount of interference CSI uncertainty for each PU is evaluated according to Section 3.2.1. Furthermore, the number of constraints is expanded for the PUs in (3.29b) when obtaining the precoder.

3.3 Compound Capacity Problem

When dropping (3.5b) in P-I, we have the CC problem

$$(P-II): \min_{\mathbf{Q} \in \mathcal{Q}} \max_{\mathbf{\Delta}_{22} \in \mathcal{U}_2} -\phi(\mathbf{Q}, \mathbf{\Delta}_{22}). \quad (3.32)$$

The CC problem is considered to be challenging to solve in the literature due to the inner maximization over $\mathbf{\Delta}_{22}$ before optimization in \mathbf{Q} can take place. Indeed, analytical solutions for P-II are known under very few cases: when \mathbf{H}_{22} belongs to an isotropic set [65], the available channel $\hat{\mathbf{H}}_{22}$ is rank one [104], and the uncertainty set is defined for the spectral norm only ($p_2 = \infty$) [49, 82]. In this work, the considered uncertainty set \mathcal{U}_2 is governed by the Schatten norm and the feasible set \mathcal{Q} is more general than the previous studies [48, 49, 65, 82, 104].

The proposed algorithm for CR is applicable to the CC problem as well. There are two main changes. The first one is that we do not need to update $\mathbf{\Delta}_{21}$ in each iteration. The second is that the feasible set of \mathbf{Q} is convex because of the absence of the constraint (3.5b), making it easier to solve.

Algorithm 2 summarizes the main framework to find a solution for the CC problem. It would converge to a critical point for $\mathbf{\Delta}_{22}$ and \mathbf{Q} . The proof of convergence follows a similar analysis as in Appendix 3.7.5.

Similar to Algorithm 1, the initialization is $\mathbf{Q}^{(1)} = \gamma_{\mathbf{Q}} \mathbf{I}$ and $\mathbf{N}^{(1)} = \mathbf{I}$ with $\gamma_{\mathbf{Q}} \approx 0$. The optimization problem P-I.3 in (3.10) is non-concave in $\mathbf{\Delta}_{22}$. The possibility remains of having multiple stationary points that some could be local optima. Using several different initializations for $\mathbf{Q}^{(1)}$ could improve the chance of obtaining the global optimum solution.

Algorithm 2 Solving P-II Using the AI and FDPS Approach

Input: Provide the initial matrices $\mathbf{Q}^{(1)}$ and $\mathbf{N}^{(1)}$; the Schatten norm order p_2 and the uncertainty size ϵ_2 . Set counter $l = 0$.

Step 1. $l = l + 1$.

Step 2. If $\mathbf{Q}^{(l)}$ satisfies a suitable stopping criterion: STOP.

Step 3. Obtain the solution $\Delta_{22}^{(l)}$ by iterating (3.16) and (3.17).

Step 4. Calculate $\Psi_g^{(l)}(\mathbf{Q}^{(l)})$ using (3.25) and $\bar{\mathbf{d}}^{(l)}$ using (3.24); solve the subproblem (3.23) for $\check{\mathbf{Q}}^{(l)}$ and obtain $\mathbf{d}^{(l)}$ from (3.21) at $\mathbf{Q} = \check{\mathbf{Q}}^{(l)}$.

Step 5. Determine α_l using a fixed or diminishing step-size rule.

Step 6. Obtain the precoder $\mathbf{Q}^{(l+1)}$ from (3.22) and $\mathbf{N}^{(l+1)}$ using (3.30). Go back to Step 1.

3.4 Particular Solutions

Under some specific conditions the proposed solutions for the CR and CC problems can be reduced to simpler forms for evaluation.

3.4.1 Particular Solution for P-I

We shall consider the solution of the CR problem P-I under the following two conditions.

- The uncertainty sets \mathcal{U}_1 and \mathcal{U}_2 for $\hat{\mathbf{R}}_{21}$ and $\hat{\mathbf{R}}_{22}$ are defined under $p_j = \infty$, $j = 1, 2$ in the Schatten norm (2.15), which is equivalent to the spectral norm. Using the spectral norm will lead to the most conservative solution among the class of unitarily-invariant norms [48];
- The power limitation constraints for the precoder in the feasible set \mathcal{Q} are inactive in the presence of the interference constraint (3.5b). Such situation occurs often in a CR network where the performance is interference limited rather than power-budget restricted.

Let us define the PSD matrix $\mathbf{R}_{22}^{\epsilon_2} = \mathbf{U}_{\hat{\mathbf{R}}_{22}}(\boldsymbol{\Lambda}_{\hat{\mathbf{R}}_{22}} - \epsilon_2\mathbf{I})^+\mathbf{U}_{\hat{\mathbf{R}}_{22}}^\dagger$. The inequality $\|\boldsymbol{\Delta}_{22}\|_{s_\infty} \leq \epsilon_2$ from the uncertainty set $\tilde{\mathcal{U}}_2$ is equivalent to $-\boldsymbol{\Delta}_{22} \succeq -\epsilon_2\mathbf{I}$ where $\boldsymbol{\Delta}_{22}$ is PSD. Adding both sides by $\hat{\mathbf{R}}_{22}$ gives

$$\hat{\mathbf{R}}_{22} - \boldsymbol{\Delta}_{22} \succeq \mathbf{R}_{22}^{\epsilon_2} \quad (3.33a)$$

$$\Rightarrow \log |\mathbf{I} + (\hat{\mathbf{R}}_{22} - \boldsymbol{\Delta}_{22})\mathbf{Q}| \geq \log |\mathbf{I} + \mathbf{R}_{22}^{\epsilon_2}\mathbf{Q}|. \quad (3.33b)$$

As a result, the solution of $\boldsymbol{\Delta}_{22}$ that yields the minimum MI is $\boldsymbol{\Delta}_{22} = \epsilon_2\mathbf{I}$.

Similarly, let $\mathbf{R}_{21}^{\epsilon_1} = \hat{\mathbf{R}}_{21} + \epsilon_1\mathbf{I}$. The inequality $\|\boldsymbol{\Delta}_{21}\|_{s_\infty} \leq \epsilon_1$ from the uncertainty set \mathcal{U}_1 is equivalent to $-\boldsymbol{\Delta}_{21} \preceq \epsilon_1\mathbf{I}$. Adding both sides by $\hat{\mathbf{R}}_{21}$ yields

$$\hat{\mathbf{R}}_{21} - \boldsymbol{\Delta}_{21} \preceq \mathbf{R}_{21}^{\epsilon_1} \quad (3.34a)$$

$$\Rightarrow \log |\mathbf{I} + (\hat{\mathbf{R}}_{21} - \boldsymbol{\Delta}_{21})\mathbf{Q}| \leq \log |\mathbf{I} + \mathbf{R}_{21}^{\epsilon_1}\mathbf{Q}|. \quad (3.34b)$$

The largest amount of interference to PU occurs when $\boldsymbol{\Delta}_{21} = -\epsilon_1\mathbf{I}$.

Using the uncertainty matrix solutions $\boldsymbol{\Delta}_{22} = \epsilon_2\mathbf{I}$ and $\boldsymbol{\Delta}_{21} = -\epsilon_1\mathbf{I}$, P-I in (3.5) becomes

$$(P\text{-III}): \max_{\mathbf{Q} \succeq 0} \log |\mathbf{I} + \mathbf{R}_{22}^{\epsilon_2}\mathbf{Q}| \quad (3.35a)$$

$$\text{s.t. } \log |\mathbf{I} + \mathbf{R}_{21}^{\epsilon_1}\mathbf{Q}| \leq r_L \quad (3.35b)$$

where the power limitation in the feasible set \mathcal{Q} is assumed satisfied. The following proposition shows that the optimal precoder should diagonalize the matrix $\mathbf{R}_{22}^{\epsilon_2}$.

Proposition 2: Let $\mathbf{U} \in \mathbb{C}^{M_2 \times r}$ with $r \leq M_2$ be (semi-)unitary and $\boldsymbol{\Sigma} \in \mathbb{R}^{r \times r}$ a diagonal matrix with non-negative non-increasing elements. The optimal precoder

solution for the problem P-III would diagonalize the covariance matrix $\mathbf{R}_{22}^{\epsilon_2}$ and has the structure

$$\mathbf{Q} = (\mathbf{R}_{22}^{\epsilon_2})^{-\frac{\dagger}{2}} \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\dagger (\mathbf{R}_{22}^{\epsilon_2})^{-\frac{\dagger}{2}} \quad (3.36)$$

where $(\mathbf{R}_{22}^{\epsilon_2})^{\frac{1}{2}}$ is the matrix square root of $\mathbf{R}_{22}^{\epsilon_2}$ and $(\cdot)^\dagger$ denotes the pseudo-inverse of (\cdot) .

Define the PSD matrix $\bar{\mathbf{R}} = (\mathbf{R}_{22}^{\epsilon_2})^{-\frac{\dagger}{2}} \mathbf{R}_{21}^{\epsilon_1} (\mathbf{R}_{22}^{\epsilon_2})^{-\frac{\dagger}{2}}$ whose EVD is $\bar{\mathbf{R}} = \mathbf{U}_{\bar{\mathbf{R}}} \mathbf{\Lambda}_{\bar{\mathbf{R}}} \mathbf{U}_{\bar{\mathbf{R}}}^\dagger$ where $\mathbf{U}_{\bar{\mathbf{R}}} \in \mathbb{C}^{M_2 \times r}$ and $\mathbf{\Lambda}_{\bar{\mathbf{R}}} \in \mathbb{R}^{r \times r}$ are (semi-)unitary and diagonal matrices. Then, the selection $\mathbf{U} = \mathbf{U}_{\bar{\mathbf{R}}} \mathbf{I}^a$ would optimize (3.35) where \mathbf{I}^a is the anti-identity matrix of appropriate size. The matrix $\mathbf{\Sigma}$ is obtained by solving the problem

$$\text{(P-IV): } \max_{\mathbf{\Sigma} \succeq 0} \log |\mathbf{I} + \mathbf{\Sigma}| \quad (3.37a)$$

$$\text{s.t. } \log |\mathbf{I} + \mathbf{I}^a \mathbf{\Lambda}_{\bar{\mathbf{R}}} \mathbf{I}^a \mathbf{\Sigma}| \leq r_L. \quad (3.37b)$$

Proof: Please refer to Appendix 3.7.6. ■

P-IV is not convex in $\mathbf{\Sigma}$ due to the constraint (3.37b). (3.37b) can become tractable by using the Taylor-series expansion and P-IV can be solved by the iterative linearization technique as presented in Section 3.2.3.

3.4.2 Particular Solution for P-II

When the power of the precoder \mathbf{Q} is restricted to be in a unitarily-invariant convex set \mathcal{Q}_u , for example, the total power restriction (2.5) or the maximum power limitation

$$\mathcal{Q}_u = \{\mathbf{Q} : \mathbf{Q} \succeq 0, \lambda_1(\mathbf{Q}) \leq P_{\max}\}, \quad (3.38)$$

the following theorem states the precoder for the CC problem has the same eigenvectors as the available channel covariance.

Theorem 2: Consider the EVD representation $\mathbf{Q} = \mathbf{U}_\mathbf{Q}\mathbf{\Lambda}_\mathbf{Q}\mathbf{U}_\mathbf{Q}^\dagger$ and $\hat{\mathbf{R}}_{22} = \mathbf{U}_{\hat{\mathbf{R}}_{22}}\mathbf{\Lambda}_{\hat{\mathbf{R}}_{22}}\mathbf{U}_{\hat{\mathbf{R}}_{22}}^\dagger$ where the elements along the diagonal of $\mathbf{\Lambda}_\mathbf{Q}$ and $\mathbf{\Lambda}_{\hat{\mathbf{R}}_{22}}$ are arranged in non-increasing order. For any $\mathbf{\Delta}_{22} \in \mathcal{U}_2$, define the matrix $\mathbf{\Delta}_{22}^o \triangleq \mathbf{J}\mathbf{U}_{\hat{\mathbf{R}}_{22}}^\dagger\mathbf{\Delta}_{22}\mathbf{U}_{\hat{\mathbf{R}}_{22}}\mathbf{J}$, where $\mathbf{J} \in \mathbb{R}^{n \times n}$ is a diagonal matrix with diagonal elements either +1 or -1. If $\mathbf{Q} \in \mathcal{Q}_u$ where \mathcal{Q}_u is a unitarily-invariant compact and convex set, the optimal precoder for the CC problem P-II has eigenvectors equal to those of the channel $\hat{\mathbf{R}}_{22}$ available, i.e., $\mathbf{U}_\mathbf{Q} = \mathbf{U}_{\hat{\mathbf{R}}_{22}}$. Additionally, P-II becomes

$$(P-V): \max_{\mathbf{\Lambda}_\mathbf{Q} \in \mathcal{Q}_u} \min_{\mathbf{\Delta}_{22}^o \in \mathcal{U}_2^o} \log |\mathbf{I} + (\mathbf{\Lambda}_{\hat{\mathbf{R}}_{22}} - \mathbf{\Delta}_{22}^o)\mathbf{\Lambda}_\mathbf{Q}| \quad (3.39)$$

where the uncertainty set \mathcal{U}_2^o is

$$\mathcal{U}_2^o = \{\|\mathbf{\Delta}_{22}^o\|_{S_{p_2}} \leq \epsilon_2, \mathbf{\Lambda}_{\hat{\mathbf{R}}_{22}} - \mathbf{\Delta}_{22}^o \succeq 0, \mathbf{\Delta}_{22}^o \succeq 0\}. \quad (3.40)$$

Proof: Please see Appendix 3.7.7. ■

The conclusion that the precoder has the same eigenmodes as the available channel has previously been demonstrated for the CSI uncertainty set defined by the spectral norm only under unitarily-invariant convex set for the transmission power [49, 82]. We generalize the results here and prove that the same conclusion carries for the CSI uncertainty set defined by the Schatten norm.

As apparent from (3.39), the outer maximization for the CC problem simplifies considerably since the unknowns are M_2 eigenvalues rather than $M_2(M_2 + 1)/2$ independent elements of \mathbf{Q} .

The modification to the FDPS method developed in Section 3.2.3 for P-V is straightforward. In particular, the matrices \mathbf{Q} and $\hat{\mathbf{R}}_{22}$ become diagonal counterparts $\Lambda_{\mathbf{Q}}$ and $\Lambda_{\hat{\mathbf{R}}_{22}}$ with diagonal elements arranged in non-increasing order. Furthermore, Δ_{22} is now Δ_{22}^o , and \mathcal{Q} and \mathcal{U}_2 turn into \mathcal{Q}_u and \mathcal{U}_2^o .

We can simplify P-V further by restricting the matrix Δ_{22}^o to be diagonal. Consequently, both \mathbf{Q} and Δ_{22} share the same eigenvectors as $\hat{\mathbf{R}}_{22}$, which leads to a channel-diagonalizing solution. In this case, we only need to obtain the eigenvalues of \mathbf{Q} and Δ_{22} . The restriction that Δ_{22}^o has diagonal structure will likely lead to a suboptimal solution, unless the channel uncertainty set \mathcal{U}_2 is defined by the spectral norm [49, 82].

3.5 Numerical Results

We shall provide numerical examples to illustrate the performance of proposed precoder solutions. The available CSI is set as $\hat{\mathbf{R}}_{2j} = \hat{\mathbf{H}}_{2j}^\dagger \hat{\mathbf{H}}_{2j}$ for $j = 1, 2$, where the elements of $\hat{\mathbf{H}}_{2j}$ are IID and follow a Gaussian distribution with zero mean and unit variance. The SU rate reported is the worst case rate by the design methodology and is the average over the randomly generated matrices $\hat{\mathbf{R}}_{2j}$. The uncertainty error size is $\epsilon_j = w_j \|\hat{\mathbf{R}}_{2j}\|_{S_\infty}$ where the normalized parameter $w_j \in (0, 1)$ is used to adjust the amount of uncertainty. The noise variance is fixed to unity and hence the total SU transmit power P_T is equal to the signal-to-noise ratio (SNR), i.e., $P_T = \text{SNR}$. In addition to the power sets in (2.5) and (3.38), we consider the per-transmit dimension

power set defined as

$$\mathcal{Q} = \{\mathbf{Q} : \mathbf{Q} \succeq 0, \mathbf{Q} \preceq \text{diag}([P_1, \dots, P_{M_2}]^T)\} \quad (3.41)$$

where P_m , $m = 1, \dots, M_2$, is the power limit for the m -th antenna. Unless otherwise stated, we simply set the power values as $P_{\max} = 0.7P_T$ and $P_m = 0.75^m P_T$ for illustration purpose.

For all the proposed precoder designs, we compute α_l in (3.22) by a diminishing step-size rule such that the rule $\alpha_{l+1} = 1/(l+1)$ is used for the CR problem while $\alpha_{l+1} = \alpha_l(1 - \beta\alpha_l)$ with $\beta = 10^{-2} \in (0, 1)$ and $\alpha_0 = 1$ for the CC problem. The proposed solutions shown in Algorithm 1 and Algorithm 2 are computed with one initialization of $\mathbf{Q}^{(1)}$ only. The initialization constants $\gamma_{\Delta_{22}}$, $\gamma_{\mathbf{Q}}$ and γ_{CR} are set to 10^{-4} and the constants s_{\min} and s_{\max} are 10^{-7} and 10^{20} . We shall denote the solution for Δ_{22} as AI(O) when solving P-I.5 optimally using CVX [34] and AI(SO) as the suboptimal solution.

We first examine Algorithm 1 and the proposed solutions for the CR problem in Section 3.2 under different values of the Schatten norm orders p_j and normalized uncertainty error w_j , $j = 1, 2$. Figure 3.2 illustrates the SU rate in (nats/s/Hz) as the leakage rate limit r_L (nats/s/Hz) increases. It includes the robust solution from [2] for comparison. The antenna setting is $M_2 = N_2 = N_1 = 4$ and SNR = 5dB. We use the AI(O) method to obtain the uncertainty element Δ_{22} . It is obvious from Figure 3.2 that the proposed solutions for \mathbf{Q} and Δ_{21} provide higher rate than that in [2]. It should be noted that the work [2] does not consider CSI uncertainty in the direct link of SU; therefore, we limit the comparison in Figure 3.2 for $w_2 = 0$ only. As expected, increasing the uncertainty error w_2 for fixed values of w_1 will deteriorate

the SU rate. Also, the proposed solution when $p_1 = p_2 = \infty$ produces the most conservative solution among the Schatten norm orders.

Figure 3.3 depicts the SU rate versus the uncertainty error w_1 of the interference link for antenna settings of $M_2 = N_2 = N_1 = 2$ or 4, where the leakage rate limit and the uncertainty error of the direct link are kept at $r_L = 3$ (nats/s/Hz) and $w_2 = 0.5$. The AI(SO) method is applied to obtain Δ_{22} . The results show that the SU rate decreases as the uncertainty error w_1 increases for a fixed value of w_2 , and the decrease is more significant for the four antenna than the two antenna setting. The uncertainty sets defined with $p_1 = p_2 = 1$ have the smallest size among all other Schatten norm orders and SU has the highest rate. The effect of the Schatten norm order on the SU rate is apparent.

We next consider Algorithm 2 to obtain the designs for the CC problem presented in Section 3.3 when the normalized uncertainty error $w_2 = 0.3$ or 0.5 and the Schatten norm order $p_2 = 2$. Both precoder solutions AI(O)-FDPS or AI(SO)-FDPS are provided. To illustrate the effectiveness of the proposed FDPS technique, we compare the proposed precoder solutions with the beamforming precoder that transmits signals over the principal eigenvalue of the worst channel and with the equal power allocation precoder that distributes power equally among the antennas. Both the beamforming and equal power precoders require Δ_{22} and it is taken from the proposed AI(O)-FDPS solution. Figure 3.4 illustrates the average SU rate as SNR increases. The antenna setting for SU is $M_2 = N_2 = 4$. Figure 3.4 indicates that the proposed precoder significantly outperforms the beamforming and equal power ones. The performance of the AI(O)-FDPS and AI(SO)-FDPS is almost identical at different SNR and w_2 values.

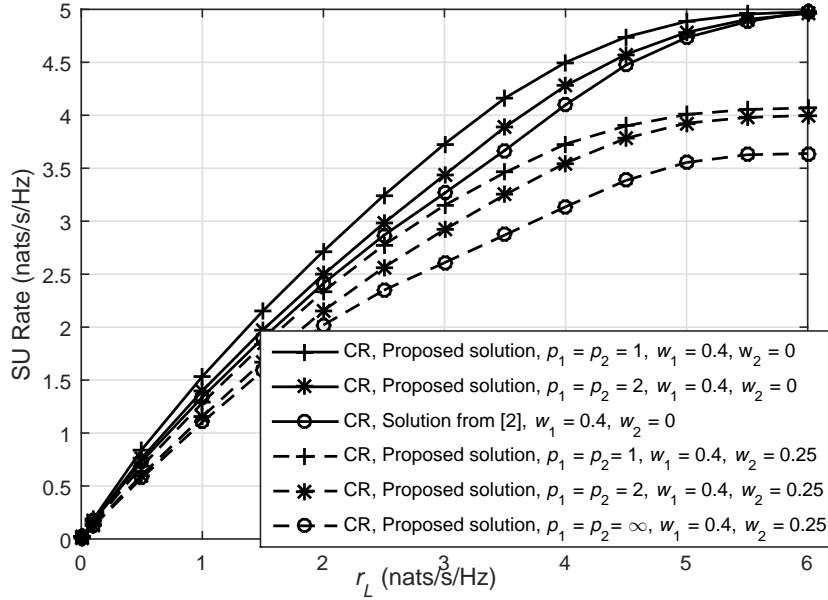


Figure 3.2: Rate for SU from the proposed CR solution and from [2] for different values of r_L (nats/s/Hz), SNR = 5dB, $M_2 = N_2 = N_1 = 4$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$ and $\hat{\mathbf{R}}_{21}$.

Figure 3.5 shows the average SU rates when the error w_2 increases with SNR = 4dB and $p_2 = 2$ or ∞ . The power and antenna settings of Figure 3.5 are the same as those in Figure 3.4. The proposed precoders show better SU rate than the beamforming and equal power designs. The uncertainty set defined with $p_2 = \infty$ has the largest shape among the Schatten norm uncertainty sets; hence the proposed solution at $p_2 = \infty$ would produce the most conservative SU rate. Similar to Figure 3.4, the precoders from AI(O)-FDPS and AI(SO)-FDPS are quite comparable in performance at different values of p_2 and w_2 .

Finally, we investigate the proposed particular solutions for the CR and CC problems in Section 3.4. Figure 3.6 evaluates the performance of the particular solution for the CR problem from Section 3.4.1 at different values of $w_1 = w_2$ as the leakage

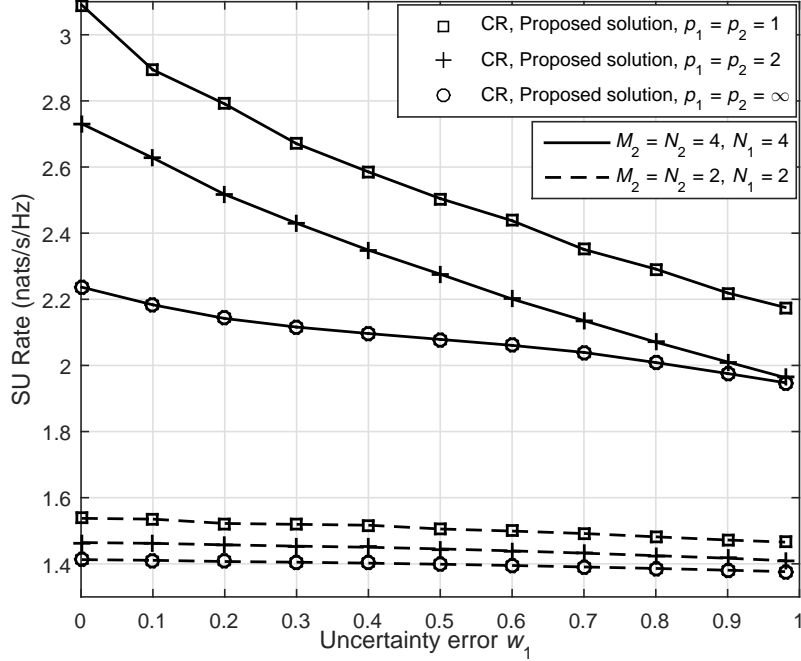


Figure 3.3: Rate for SU from the proposed CR solution versus the uncertainty error w_1 , $r_L = 3$ (nats/s/Hz), $w_2 = 0.5$, SNR = 5dB, $M_2 = N_2 = N_1 = 2$ or 4. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$ and $\hat{\mathbf{R}}_{21}$.

rate limit r_L increases, where $p_1 = p_2 = \infty$ and the power set \mathcal{Q} is relaxed. The SU rate increases with r_L as expected. The increase in the SU rate is more sensitive at smaller normalized uncertainty radii w_1 and w_2 .

Figure 3.7 shows the SU rates for the CC problem as SNR varies. The curves denoted by “ \mathbf{Q} diagonalization” and “ \mathbf{Q} , $\mathbf{\Delta}_{22}$ diagonalization” correspond to the optimal precoder that diagonalizes the available channel estimate $\hat{\mathbf{R}}_{22}$ and the sub-optimal solution in which both \mathbf{Q} and $\mathbf{\Delta}_{22}$ diagonalize $\hat{\mathbf{R}}_{22}$ for any value of p_2 . The SU rates for both precoder solutions are very comparable. The proposed precoders are applicable to a wide class of situations and covers some precoder solutions from the literature that are obtained under more restrictive scenarios. This is illustrated in

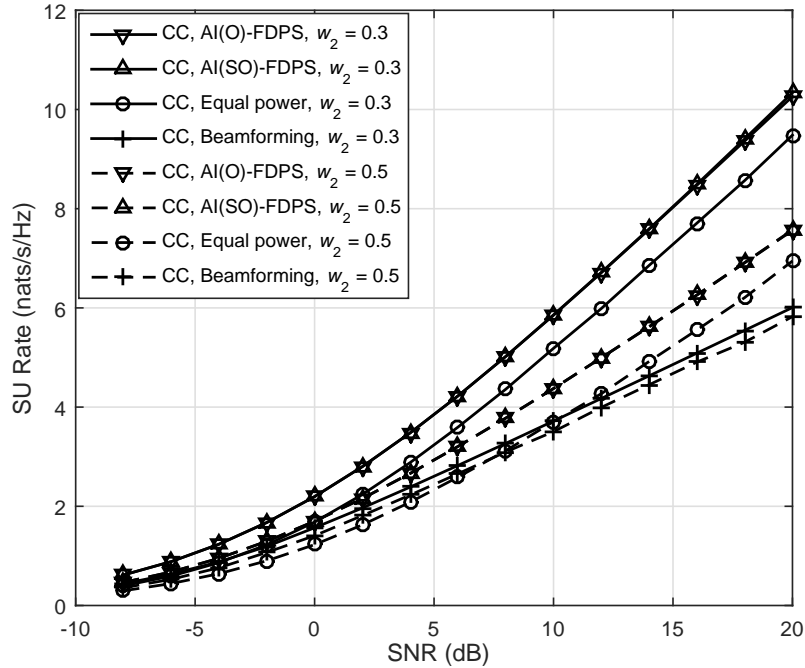


Figure 3.4: Rate for SU from the proposed CC solution, the beamforming and the equal power solutions versus SNR (dB), $M_2 = N_2 = 4$. The results shown are the average over 100 realizations of \mathbf{R}_{22} .

the figure where the proposed Q diagonalization precoder yields identical performance with the precoder from [49, 82] that is applicable to $p_2 = \infty$ only.

3.6 Conclusion

We have developed in this chapter the solution of a robust precoder for an SU in an underlay cognitive radio framework, where the available CSI of the direct link of SU and the interference link to PU have uncertainties and the QoS of PU must be maintained. In the absence of PU, the proposed solution addresses the CC problem in the literature. The CSI uncertainties are modelled by convex sets with the Schatten

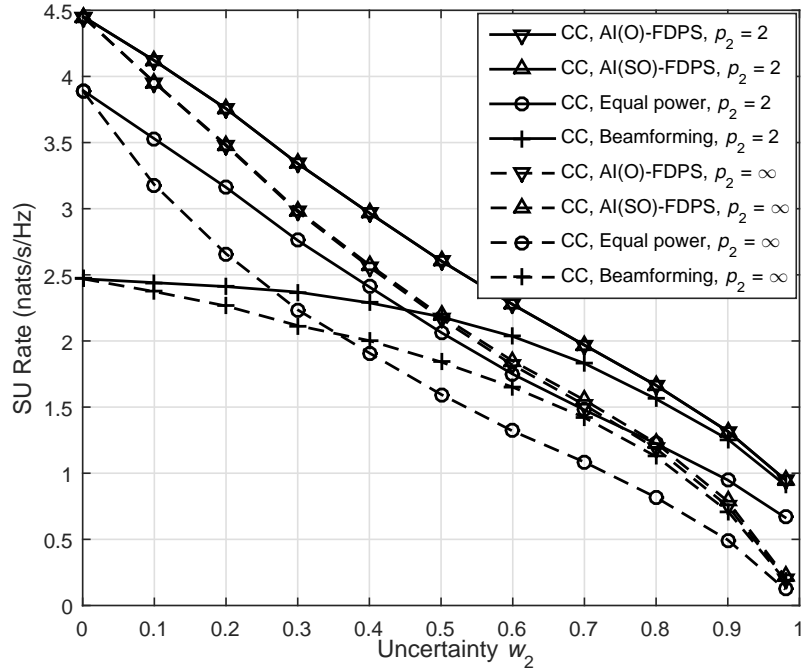


Figure 3.5: Rate for SU from the proposed CC solution, the beamforming and the equal power solutions at different values of w_2 , $M_2 = N_2 = 4$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$.

norm measure. We formulate the design as a maximin optimization problem that seeks a precoder to maximize the MI of SU under the most conservative anticipated performance for SU while ensuring the QoS of PU under all possible CSI uncertainties expected in the interference link. The maximin problem is solved iteratively, using the explicit solution for the worst case CSI deviation of the interference link, the AI technique for the worst case CSI error of the direct link, and the FDPS method for the precoder. Convergence analysis of the proposed solution is established. Simpler solutions for some specific choices of the norm measure and transmission power restrictions are also derived. Simulations validate the performance improvement of the proposed precoders for the CR and CC problems.

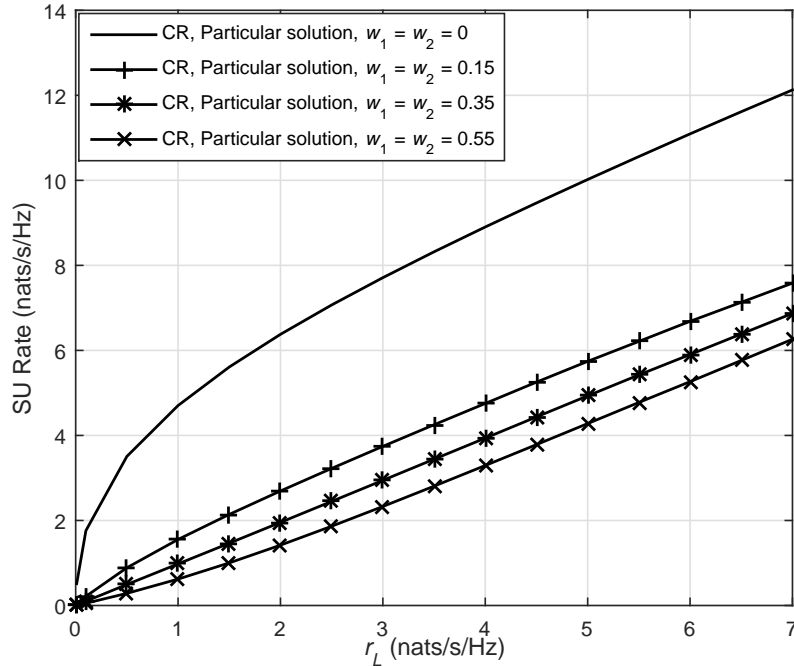


Figure 3.6: Rate for SU from the particular CR solution versus r_L , $M_2 = N_2 = N_1 = 5$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$ and $\hat{\mathbf{R}}_{21}$.

3.7 Appendix

3.7.1 Upper Bound for MI

The MI for the transmission link from SU-Tx to SU-Rx is defined as [31]

$$\mathbb{E}[\log |\mathbf{I} + \mathbf{H}_{22} \mathbf{Q} \mathbf{H}_{22}^\dagger|] \quad (3.42)$$

where the expectation is taken over the channel \mathbf{H}_{22} . Using first the fact that [13, Corollory 2.8.5]

$$\log |\mathbf{I} + \mathbf{A} \mathbf{B}| = \log |\mathbf{I} + \mathbf{B} \mathbf{A}| \quad (3.43)$$

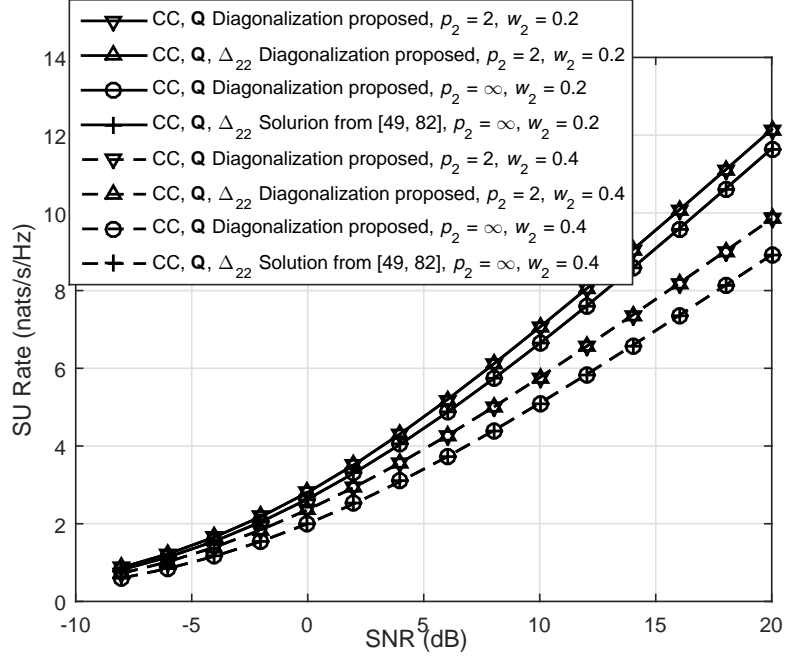


Figure 3.7: Rate for SU from the particular CC solutions versus SNR (dB), $M_2 = 6$, $N_2 = 3$. The results shown are the average over 100 realizations of $\hat{\mathbf{R}}_{22}$.

for any $\mathbf{A} \in \mathbb{C}^{n \times m}$ and $\mathbf{B} \in \mathbb{C}^{m \times n}$, and next the concavity of the $\log |\cdot|$ function over the set of PSD matrices, we have

$$\mathbb{E}[\log |\mathbf{I} + \mathbf{H}_{22} \mathbf{Q} \mathbf{H}_{22}^\dagger|] = \mathbb{E}[\log |\mathbf{I} + \mathbf{H}_{22}^\dagger \mathbf{H}_{22} \mathbf{Q}|] \quad (3.44a)$$

$$\leq \log |\mathbf{I} + \mathbb{E}[\mathbf{H}_{22}^\dagger \mathbf{H}_{22}] \mathbf{Q}|. \quad (3.44b)$$

Hence the metric $\phi(\cdot, \cdot)$ in (3.3a) is an upper bound of (3.42).

We can follow similar procedure to show that the function $g(\mathbf{Q}, \mathbf{R}_{21})$ (3.4a) is no less than $\mathbb{E}[\log |\mathbf{I} + \mathbf{H}_{21} \mathbf{Q} \mathbf{H}_{21}^\dagger|]$, where the expectation is applied to the channel \mathbf{H}_{21} .

3.7.2 Proof of Proposition 1

Consider NSD $\mathbf{\Delta}_{21} \in \mathcal{U}_1$ and PSD $\check{\mathbf{\Delta}}_{21} \in \mathcal{U}_1$. It is obvious that $-\check{\mathbf{\Delta}}_{21} \preceq -\mathbf{\Delta}_{21}$. Hence

$$\hat{\mathbf{R}}_{21} - \check{\mathbf{\Delta}}_{21} \preceq \hat{\mathbf{R}}_{21} - \mathbf{\Delta}_{21} \quad (3.45a)$$

$$\mathbf{Q}^{1/2}(\hat{\mathbf{R}}_{21} - \check{\mathbf{\Delta}}_{21})\mathbf{Q}^{1/2} \preceq \mathbf{Q}^{1/2}(\hat{\mathbf{R}}_{21} - \mathbf{\Delta}_{21})\mathbf{Q}^{1/2} \quad (3.45b)$$

$$\log |\mathbf{I} + \mathbf{Q}^{1/2}(\hat{\mathbf{R}}_{21} - \check{\mathbf{\Delta}}_{21})\mathbf{Q}^{1/2}| \leq \log |\mathbf{I} + \mathbf{Q}^{1/2}(\hat{\mathbf{R}}_{21} - \mathbf{\Delta}_{21})\mathbf{Q}^{1/2}| \quad (3.45c)$$

$$\log |\mathbf{I} + (\hat{\mathbf{R}}_{21} - \check{\mathbf{\Delta}}_{21})\mathbf{Q}| \leq \log |\mathbf{I} + (\hat{\mathbf{R}}_{21} - \mathbf{\Delta}_{21})\mathbf{Q}| \quad (3.45d)$$

where $\mathbf{Q}^{1/2}$ is the matrix square root of \mathbf{Q} . (3.45b) follows since for Hermitian matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ such that $\mathbf{A} \preceq \mathbf{B}$, $\mathbf{S}\mathbf{A}\mathbf{S}^\dagger \preceq \mathbf{S}\mathbf{B}\mathbf{S}^\dagger$ where $\mathbf{S} \in \mathbb{C}^{m \times n}$ [13, Proposition 8.1.2]. (3.45c) comes from the fact that the function $\log |\cdot|$ is strictly increasing on the set of PD Hermitian matrices [57, Ch. 16, F.2.b], and (3.45d) is by applying (3.43). It is obvious from (3.45d) that the function $g(\cdot, \cdot)$ will take a larger value if $\mathbf{\Delta}_{21}$ is NSD rather than PSD.

3.7.3 Proof of Theorem 1

The proof requires a number of lemmas. We first summarize them below,

Lemma 2: Given the Hermitian and P(S)D matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$. The matrix $\mathbf{A}(\mathbf{I} + \mathbf{B}\mathbf{A})^{-1}$ is then Hermitian and P(S)D matrix.

Proof: Let $\mathbf{A}^{1/2}$ be the matrix square root of \mathbf{A} . From direct algebraic manipula-

tion,

$$\mathbf{A}(\mathbf{I} + \mathbf{BA})^{-1} = \mathbf{A}^{1/2}\mathbf{A}^{1/2}(\mathbf{I} + \mathbf{BA}^{1/2}\mathbf{A}^{1/2})^{-1} \quad (3.46a)$$

$$= \mathbf{A}^{1/2}(\mathbf{I} + \mathbf{A}^{1/2}\mathbf{BA}^{1/2})^{-1}\mathbf{A}^{1/2}. \quad (3.46b)$$

(3.46b) is due to the fact $(\mathbf{I} + \mathbf{AB})^{-1}\mathbf{A} = \mathbf{A}(\mathbf{I} + \mathbf{BA})^{-1}$ [13, Fact 2.16.16]. The Hermitian property can be easily seen from the right side of (3.46b). Note that $\mathbf{A}^{1/2}\mathbf{BA}^{1/2}$ is P(S)D. By the definition of the P(S)D property, the right side is P(S)D and hence the left side. \blacksquare

Lemma 3: Let $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ be Hermitian with eigenvalues $\lambda_1(\mathbf{A}) \geq \dots \geq \lambda_n(\mathbf{A})$ and $\lambda_1(\mathbf{B}) \geq \dots \geq \lambda_n(\mathbf{B})$. Then, $\text{Tr}(\mathbf{AB}) \leq \sum_{i=1}^n \lambda_i(\mathbf{A})\lambda_i(\mathbf{B})$ and the equality holds if \mathbf{A} and \mathbf{B} are diagonal [13, Fact 5.12.4].

From Lemma 2 and (3.16) the matrix $(\mathbf{Q}^{(l)}\mathbf{W}_2)$ is Hermitian and PSD.

We shall begin the proof of Theorem 1 as follows. Let the Lagrange multipliers associated with (3.17b) and (3.17c) be $\gamma \geq 0$ and $\mathbf{F} \succeq 0$. We first form the Lagrangian \mathcal{L} for (3.17) and its corresponding dual function $\mathcal{D}(\cdot, \cdot)$ is

$$\mathcal{D}(\mathbf{F}, \gamma) = \max_{\Delta_{22} \succeq 0} \mathcal{L}(\Delta_{22}, \mathbf{F}, \gamma) \quad (3.47a)$$

$$= \max_{\Delta_{22} \succeq 0} \text{Tr}(\mathbf{Q}^{(l)}\mathbf{W}_2\Delta_{22}) - \gamma(\|\Delta_{22}\|_{S_{p_2}} - \epsilon_2) + \text{Tr}\left((\hat{\mathbf{R}}_{22} - \Delta_{22})\mathbf{F}\right) \quad (3.47b)$$

$$= \max_{\Delta_{22} \succeq 0} \text{Tr}\left((\mathbf{Q}^{(l)}\mathbf{W}_2 - \mathbf{F})\Delta_{22}\right) - \gamma\|\Delta_{22}\|_{S_{p_2}} + \gamma\epsilon_2 + \text{Tr}(\hat{\mathbf{R}}_{22}\mathbf{F}) \quad (3.47c)$$

$$\leq \max_{\Lambda_{\Delta_{22}} \succeq 0} \text{Tr}(\Lambda_{\mathbf{E}_2}\Lambda_{\Delta_{22}}) - \gamma\|\Lambda_{\Delta_{22}}\|_{S_{p_2}} + \gamma\epsilon_2 + \text{Tr}(\hat{\mathbf{R}}_{22}\mathbf{F}) \quad (3.47d)$$

$$\leq \max_{\Lambda_{\Delta_{22}} \succeq 0} \|\Lambda_{\mathbf{E}_2}\Lambda_{\Delta_{22}}\|_{S_1} - \gamma\|\Lambda_{\Delta_{22}}\|_{S_{p_2}} + \gamma\epsilon_2 + \text{Tr}(\hat{\mathbf{R}}_{22}\mathbf{F}) \quad (3.47e)$$

$$\leq \max_{\Lambda_{\Delta_{22}} \succeq 0} \|\Lambda_{\Delta_{22}}\|_{S_{p_2}} (\|\Lambda_{\mathbf{E}_2}\|_{S_{q_2}} - \gamma) + \gamma\epsilon_2 + \text{Tr}(\hat{\mathbf{R}}_{22}\mathbf{F}). \quad (3.47f)$$

where $\mathbf{E}_2 = (\mathbf{Q}^{(l)}\mathbf{W}_2 - \mathbf{F})$. The inequality in (3.47d) is due to Lemma 3 and equality happens when $\mathbf{U}_{\Delta_{22}} = \mathbf{U}_{\mathbf{E}_2}$. Hence, we conclude that $\mathbf{U}_{\Delta_{22}^*} = \mathbf{U}_{\mathbf{E}_2}$ in order to maximize $\mathcal{L}(\Delta_{22}, \mathbf{F}, \gamma)$.

The inequality $\text{Tr}(\Lambda_{\mathbf{E}_2}\Lambda_{\Delta_{22}}) \leq \|(\Lambda_{\mathbf{E}_2}\Lambda_{\Delta_{22}})\|_{S_1}$ in (3.47e) is from [13, Fact 9.14.3]. The inequality in (3.47f) is due to Hölder inequality in Lemma 2 in chapter 2. Equality for (3.47f) can be achieved provided that the eigenvalues of the suboptimal solution Δ_{22}^* and that of \mathbf{E}_2 satisfy (2.53). Thus, for $p_2 \in (1, \infty)$,

$$\frac{|\lambda_1(\mathbf{E}_2)|^{1/p_2}}{\lambda_1^{1/q_2}(\Delta_{22}^*)} = \dots = \frac{|\lambda_i(\mathbf{E}_2)|^{1/p_2}}{\lambda_i^{1/q_2}(\Delta_{22}^*)} \quad (3.48)$$

and (3.18) in Theorem 1 follows.

From the Schatten norm definition in (2.15) and (3.18), when the constraint (3.17b) is active for reaching a solution of P-II.3, we obtain $\lambda_1(\Delta_{22}^*)$ as

$$\epsilon_2^{p_2} = \lambda_1(\Delta_{22}^*)^{p_2} + \sum_{k=2}^{M_2} \left(\frac{|\lambda_k(\mathbf{E}_2)|}{|\lambda_1(\mathbf{E}_2)|} \right)^{q_2} \lambda_1(\Delta_{22}^*)^{p_2}. \quad (3.49)$$

(3.49) leads to (3.19) directly.

For $p_2 = 1$, from (2.53a) a possible choice for $\lambda_i(\Delta_{22}^*)$ that achieves the equality in (3.47f) and fulfills (3.17b) is $\lambda_1(\Delta_{22}^*) = \epsilon_2$ and $\lambda_i(\Delta_{22}^*) = 0$ for $i \neq 1$. For $p_2 = \infty$, using (2.53c) we can select $\lambda_i(\Delta_{22}^*) = \epsilon_2$ to reach the equality in (3.47f) and satisfy (3.17b). The above two choices for $\lambda_i(\Delta_{22}^*)$ are valid regardless of the value of the matrix \mathbf{F}^* in \mathbf{E}_2 . If the eigenvalues of \mathbf{E}_2 as well as Δ_{22}^* are considered variables to be evaluated, the conditions (2.53a) or (2.53c) can lead to other choices for $\lambda_i(\Delta_{22}^*)$.

Proof of Corollary 1: If \mathbf{F}^* is the zero matrix, then the matrix \mathbf{E}_2 would be PSD and the inequality in (3.47e) becomes equality, which makes Δ_{22}^* the optimal

solution for the problem P-II.3 when $p_2 \in (1, \infty)$. For $p_2 = 1$ or $p_2 = \infty$ the choice $(\lambda_i(\mathbf{\Delta}_{22}^*) = \epsilon_2$ for $i = 1$ and zero otherwise) or $(\lambda_i(\mathbf{\Delta}_{22}^*) = \epsilon_2$ for every i) is optimal as \mathbf{E}_2 becomes a fixed matrix and the equality condition (2.53a) or (2.53c) is determined by $\mathbf{\Delta}_{22}^*$ only.

3.7.4 Evaluation of the Matrix \mathbf{F}^*

Proof of Corollary 2: From Appendix 3.7.3, an upper bound for the dual function $\mathcal{D}(\cdot, \cdot)$ of the problem (3.17) is the right side of (3.47f). $\mathcal{D}(\cdot, \cdot)$ would take the following values

$$\mathcal{D}(\mathbf{F}, \gamma) \leq \begin{cases} \infty, & \gamma < \|\mathbf{\Lambda}_{\mathbf{E}_2}\|_{S_{q_2}}, \\ \gamma\epsilon_2 + \text{Tr}(\hat{\mathbf{R}}_{22}\mathbf{F}), & \gamma \geq \|\mathbf{\Lambda}_{\mathbf{E}_2}\|_{S_{q_2}}. \end{cases} \quad (3.50)$$

We can find the variables γ and \mathbf{F} by minimizing the upper bound of $\mathcal{D}(\cdot, \cdot)$ when $\gamma \geq \|\mathbf{\Lambda}_{\mathbf{E}}\|_{S_{q_2}}$. That is,

$$\min_{\gamma \geq 0, \mathbf{F} \succeq 0} \gamma\epsilon_2 + \text{Tr}(\hat{\mathbf{R}}_{22}\mathbf{F}) \quad (3.51a)$$

$$\text{s.t. } \gamma \geq \|\mathbf{\Lambda}_{\mathbf{E}_2}\|_{S_{q_2}}. \quad (3.51b)$$

It is evident that γ should be chosen as its smallest possible value, $\gamma = \|\mathbf{\Lambda}_{\mathbf{E}_2}\|_{S_{q_2}}$. Substituting the optimal value of γ back into (3.51a) produces the problem (3.20).

Derivative of $h(\mathbf{F})$: For $q_2 \in [1, \infty)$, the derivative of the function $h(\mathbf{F})$ in (3.20)

with respect to \mathbf{F} is

$$\mathbf{D}_h(\mathbf{F}) = 2\epsilon_2 \frac{d \left[\text{Tr} \left(\left(\sqrt{\mathbf{E}_2^\dagger \mathbf{E}_2} \right)^{q_2} \right) \right]^{1/q_2}}{d\mathbf{F}} + \hat{\mathbf{R}}_{22} \quad (3.52a)$$

$$= -\epsilon_2 \|\mathbf{E}_2\|_{S_{q_2}}^{-q_2/p_2} \left(\sqrt{\mathbf{E}_2^\dagger \mathbf{E}_2} \right)^{q_2-2} \mathbf{E}_2 + \hat{\mathbf{R}}_{22}. \quad (3.52b)$$

In (3.52b), the matrix $(\mathbf{E}_2^\dagger \mathbf{E}_2)$ should be non-singular for $q_2 < 2$, otherwise we can consider its pseudo-inverse to obtain a well-defined derivative $\mathbf{D}_h(\mathbf{F})$ [13, Sec. 6.1].

We shall provide next a reformulation for (3.20) when $q_2 = \infty$. Since $\|\mathbf{E}_2\|_{S_\infty}$ is the largest eigenvalue of $\sqrt{\mathbf{E}_2^\dagger \mathbf{E}_2}$ according to (2.15), the optimization problem (3.20) can be casted as

$$\min_{t \geq 0, \mathbf{F} \succeq 0} \epsilon_2 t + \text{Tr}(\hat{\mathbf{R}}_{22} \mathbf{F}) \quad (3.53a)$$

$$\text{s.t.} \quad \begin{bmatrix} t\mathbf{I} & \mathbf{E}_2^\dagger \\ \mathbf{E}_2 & t\mathbf{I} \end{bmatrix} \succeq 0 \quad (3.53b)$$

where we have introduced t such that $\mathbf{E}_2^\dagger \mathbf{E}_2 \leq t^2 \mathbf{I}$ and used the Schur complement property of a partitioned PSD matrix [13, Definition 6.1.8]. Thus, (3.20) reduces to (3.53) when $q_2 = \infty$, which is an SDP problem of an affine objective function with a linear matrix inequality constraint.

3.7.5 Convergence Analysis

We shall establish the convergence of Algorithm 1 to a stationary point with exogenous step-size. For a given \mathbf{Q} , Δ_{21} has an optimal solution and the AI method is ensured

to yield a solution for Δ_{22} [33]. We therefore only need to show the convergence of the FDPS iteration for \mathbf{Q} . It is achieved based on the Danskins min-max Theorem [15, Proposition 4.5.1] which is repeated here for completeness.

Lemma 4: (Danskin's min-max Theorem [15, Proposition 4.5.1]) Let \mathcal{Y} be a compact set and let the function $f : \mathbb{C}^{n \times m} \times \mathcal{Y} \rightarrow \mathbb{R}$ be continuous and differentiable such that $f(\cdot, \mathbf{Y}) : \mathbb{C}^{n \times m} \rightarrow \mathbb{R}$ is convex for each $\mathbf{Y} \in \mathcal{Y}$. Then,

1. The maximum function $p(\mathbf{X}) \triangleq \max_{\mathbf{Y} \in \mathcal{Y}} f(\mathbf{X}, \mathbf{Y})$ is convex and has a subgradient $\nabla_{\mathbf{X}} f(\mathbf{X}, \bar{\mathbf{Y}})$ with respect to \mathbf{X} , where $\bar{\mathbf{Y}} \in \mathcal{Y}(\mathbf{X})$ and $\mathcal{Y}(\mathbf{X}) \triangleq \{\bar{\mathbf{Y}} : f(\mathbf{X}, \bar{\mathbf{Y}}) = \max_{\mathbf{Y} \in \mathcal{Y}} f(\mathbf{X}, \mathbf{Y})\}$ is the set of maximizing points for $p(\cdot)$.

2. If $f(\cdot, \mathbf{Y})$ is differentiable for any $\mathbf{Y} \in \mathcal{Y}$ and $\nabla_{\mathbf{X}} f(\mathbf{X}, \cdot)$ is continuous on \mathcal{Y} for each \mathbf{X} , then

$$\partial p(\mathbf{X}) = \text{conv}\{\nabla_{\mathbf{X}} f(\mathbf{X}, \bar{\mathbf{Y}}) : \bar{\mathbf{Y}} \in \mathcal{Y}(\mathbf{X})\} \quad (3.54)$$

where $\partial p(\mathbf{X})$ is the subdifferential of the convex function $p(\mathbf{X})$ at \mathbf{X} , and $\text{conv}\{\cdot\}$ denotes the convex hull.

To make use of the min-max theorem, we associate the variables to our problem as follows: $f(\cdot, \cdot)$ is $-\phi(\cdot, \cdot)$, $p(\cdot)$ is $\Psi(\cdot)$, \mathbf{X} is \mathbf{Q} , \mathbf{Y} is Δ_{22} , and $\bar{\mathbf{Y}}$ is $\Delta_{22}^{(l)}$.

The following assumption is needed. *Assumption 1:* The Hessian inverse approximation matrix $\mathbf{N}^{(l)}$, $l = 0, 1, 2 \dots$, satisfies the following condition

$$\underline{c}\|\mathbf{z}\|^2 \leq \|\mathbf{z}\|_{\mathbf{N}^{(l)}}^2 \leq \bar{c}\|\mathbf{z}\|^2 \quad (3.55)$$

where $0 < \underline{c} \leq \bar{c}$. Assumption 1 indicates that the eigenvalues of $\mathbf{N}^{(l)}$ are bounded from below and above, which can be easily checked to ensure it is valid.

Let \mathbf{Q}^* be an optimal point for P-I in (3.5). We begin with the weighted Euclidean

distance between $\mathbf{Q}^{(l+1)}$ from (3.29) and \mathbf{Q}^* . Using (3.22) and (3.24),

$$\begin{aligned} & \|\text{vec}(\mathbf{Q}^{(l+1)}) - \text{vec}(\mathbf{Q}^*)\|_{\mathbf{N}^{-1(l)}}^2 \\ & \leq \|\text{vec}(\mathbf{Q}^{(l)}) + \alpha_l \mathbf{d}^{(l)} - \text{vec}(\mathbf{Q}^*)\|_{\mathbf{N}^{-1(l)}}^2 \end{aligned} \quad (3.56a)$$

$$\leq \|\text{vec}(\mathbf{Q}^{(l)}) + \alpha_l \bar{\mathbf{d}}^{(l)} - \text{vec}(\mathbf{Q}^*)\|_{\mathbf{N}^{-1(l)}}^2 \quad (3.56b)$$

$$\begin{aligned} & = \|\text{vec}(\mathbf{Q}^{(l)}) - \text{vec}(\mathbf{Q}^*)\|_{\mathbf{N}^{-1(l)}}^2 - 2\alpha_l \text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)}))^\dagger (\text{vec}(\mathbf{Q}^{(l)}) - \text{vec}(\mathbf{Q}^*)) \\ & \quad + \alpha_l^2 \|\text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)}))\|_{\mathbf{N}^{(l)}}^2 \end{aligned} \quad (3.56c)$$

$$\leq \|\text{vec}(\mathbf{Q}^{(l)}) - \text{vec}(\mathbf{Q}^*)\|_{\mathbf{N}^{-1(l)}}^2 - 2\alpha_l (\Psi(\mathbf{Q}^{(l)}) - \Psi(\mathbf{Q}^*)) + \alpha_l^2 \bar{c} \|\text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)}))\|^2. \quad (3.56d)$$

The inequalities in (3.56a) and (3.56b) are due to the fact that $\mathbf{Q}^{(l+1)}$ from (3.29) is a projection of that from (3.22) and the descent direction $\mathbf{d}^{(l)}$ from (3.26) is a projection of $\bar{\mathbf{d}}^{(l)}$. Note that $\text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)}))$ is a subgradient for Ψ according to part 2 of Lemma 4, and $\|\bar{\mathbf{d}}^{(l)}\|_{\mathbf{N}^{-1(l)}}^2 = \|\text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)}))\|_{\mathbf{N}^{(l)}}^2$. The last two terms in (3.56d) follows from the subgradient definition [15, Sec. 4.2] and Assumption 1, respectively. The quantity $\|\text{vec}(\Psi_g^{(l)}(\mathbf{Q}^{(l)}))\|^2$ is bounded from above for any $\Delta_{22}^{(l)}$ as the feasible set \mathcal{Q} is compact.

From (3.56d), we can proceed to prove the convergence of the proposed FDPS method for exogenous step-size using the procedure presented in [60].

3.7.6 Proof of Proposition 2

Lemma 5: For Hermitian PSD matrices $\mathbf{A} \in \mathbb{C}^{n \times n}$ and $\mathbf{B} \in \mathbb{C}^{n \times n}$ whose EVDs are $\mathbf{A} = \mathbf{U}_A \Lambda_A \mathbf{U}_A^\dagger$ and $\mathbf{B} = \mathbf{U}_B \Lambda_B \mathbf{U}_B^\dagger$, where the eigenvalues are arranged in non-

increasing order. Then, we have

$$|\mathbf{I} + \mathbf{A}\mathbf{B}| \geq |\mathbf{I} + \mathbf{I}^a \mathbf{\Lambda}_A \mathbf{I}^a \mathbf{\Lambda}_B|. \quad (3.57)$$

The equality is achieved when $\mathbf{U}_A = \mathbf{U}_B \mathbf{I}^a$ [65, Lemma 1].

We can prove the precoder structure (3.36) is optimal using a similar approach to that in [113, Theorem 1].

We can show that $\mathbf{U} = \mathbf{U}_{\bar{\mathbf{R}}} \mathbf{I}^a$ optimizes (3.35) as follows. Substituting (3.36), the left side of the constraint (3.35b) becomes $\log |\mathbf{I} + \bar{\mathbf{R}} \mathbf{U} \Sigma \mathbf{U}^\dagger|$. According to (3.57), we have $\log |\mathbf{I} + \bar{\mathbf{R}} \mathbf{U} \Sigma \mathbf{U}^\dagger| \geq \log |\mathbf{I} + \mathbf{I}^a \mathbf{\Lambda}_{\bar{\mathbf{R}}} \mathbf{I}^a \Sigma|$. Clearly, when the left side of (3.35b) is replaced by the expression $\log |\mathbf{I} + \mathbf{I}^a \mathbf{\Lambda}_{\bar{\mathbf{R}}} \mathbf{I}^a \Sigma|$, the problem (3.35) would have a larger feasible set and hence $\mathbf{U} = \mathbf{U}_{\bar{\mathbf{R}}} \mathbf{I}^a$ maximizes P-V.

3.7.7 Proof of Theorem 2

Lemma 6: Let \mathcal{J} be the set of all $L = 2^n$ diagonal matrices \mathbf{J} that have either +1 or -1 as the diagonal elements. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a Hermitian matrix and \mathbf{D}_A be a diagonal matrix such that $[\mathbf{D}_A]_{ii} = [\mathbf{A}]_{ii}$. Then, $\mathbf{D}_A = \frac{1}{L} \sum_{\mathbf{J} \in \mathcal{J}} \mathbf{J} \mathbf{A} \mathbf{J}$ [57, Ch 9].

Define the matrices $\bar{\mathbf{Q}} \triangleq \mathbf{U}_{\hat{\mathbf{R}}_{22}}^\dagger \mathbf{Q} \mathbf{U}_{\hat{\mathbf{R}}_{22}}$ and $\bar{\Delta}_{22} \triangleq \mathbf{U}_{\hat{\mathbf{R}}_{22}}^\dagger \Delta_{22} \mathbf{U}_{\hat{\mathbf{R}}_{22}}$. Under the unitarily-invariant set \mathcal{Q}_u , the CC problem P-II can be expressed as

$$\min_{\bar{\mathbf{Q}} \in \mathcal{Q}_u} \left[\Psi(\bar{\mathbf{Q}}) \triangleq \max_{\bar{\Delta}_{22} \in \bar{\mathcal{U}}_2} -\log |\mathbf{I} + (\mathbf{\Lambda}_{\hat{\mathbf{R}}_{22}} - \bar{\Delta}_{22}) \bar{\mathbf{Q}}| \right] \quad (3.58)$$

where the uncertainty set $\bar{\mathcal{U}}_2 \triangleq \{ \|\bar{\Delta}_{22}\|_{S_{p_2}} \leq \epsilon_2, \mathbf{\Lambda}_{\hat{\mathbf{R}}_{22}} - \bar{\Delta}_{22} \succeq 0, \bar{\Delta}_{22} \succeq 0 \}$. It is equivalent to \mathcal{U}_2 since the Schatten norm is invariant under the transformation and

so does the PSD property [13, Proposition 8.1.2]. Let $\Delta_{22}^o = \mathbf{J}\bar{\Delta}_{22}\mathbf{J}$. From (3.58) and using (3.43), we have

$$\Psi(\mathbf{J}\bar{\mathbf{Q}}\mathbf{J}) = \max_{\bar{\Delta}_{22} \in \bar{\mathcal{U}}_2} -\log |\mathbf{I} + (\mathbf{J}\Lambda_{\hat{\mathbf{R}}_{22}}\mathbf{J} - \mathbf{J}\bar{\Delta}_{22}\mathbf{J})\bar{\mathbf{Q}}| \quad (3.59a)$$

$$= \max_{\Delta_{22}^o \in \mathcal{U}_2^o} -\log |\mathbf{I} + (\Lambda_{\hat{\mathbf{R}}_{22}} - \Delta_{22}^o)\bar{\mathbf{Q}}| \quad (3.59b)$$

$$= \Psi(\bar{\mathbf{Q}}) \quad (3.59c)$$

where (3.59b) follows as $\bar{\mathcal{U}}_2$ is equivalent to \mathcal{U}_2^o [13, Proposition 8.1.2]. The function $\Psi(\bar{\mathbf{Q}})$ is convex in $\bar{\mathbf{Q}}$, and hence

$$\Psi(\bar{\mathbf{Q}}) = \left[\frac{1}{L} \sum_{\mathbf{J} \in \mathcal{J}} \Psi(\mathbf{J}\bar{\mathbf{Q}}\mathbf{J}) \right] \geq \Psi\left(\frac{1}{L} \sum_{\mathbf{J} \in \mathcal{J}} \mathbf{J}\bar{\mathbf{Q}}\mathbf{J}\right) \quad (3.60a)$$

$$= \Psi(\mathbf{D}_{\bar{\mathbf{Q}}}) \quad (3.60b)$$

where (3.60a) is due to the Jensen's inequality and (3.60b) are obtained from Lemma 6. The inequality in (3.60a) becomes equality when $\bar{\mathbf{Q}}$ is a diagonal matrix, which can be achieved when $\mathbf{U}_{\mathbf{Q}} = \mathbf{U}_{\hat{\mathbf{R}}_{22}}$.

Chapter 4

Bayesian Multi-antenna Sensing in Cognitive Radio Networks Using Fractional Bayes Factor

4.1 Introduction

In this chapter, we propose a Bayesian detector for spectrum sensing in a multi-antenna CR network in which no CSI is available. The Bayesian approach for detection requires a prior distribution of the CSI in terms of the spatial covariance matrix, and unfortunately it is often improper and cannot be applied directly. We shall introduce the use of the FBF approach to handle improper prior, which in turn yields a well-defined Bayes factor as the test statistic for detection. A number of priors of the CSI are examined and a closed-form expression for the test statistics is derived. The developed Bayesian detector is compared with those by using the conjugate priors for both hypotheses and GLRT, and it yields considerable improvement in detection

performance.

4.1.1 Background

Opportunistic access for CR network can provide an efficient use of the limited spectrum resources and allow interweaving between heterogeneous networks [32]. Essentially, the SU in a CR network seeks opportunistic access to the spectral band of a licensed primary network when the PU is idle. Spectrum sensing in the interweave paradigm is an essential component for the design of a CR network [6]. The capabilities of SU to detect the presence of PU can be enhanced significantly through incorporating multiple antennas at the terminals of SU and PU, see [45, 50, 54, 79, 86] and [16, 41, 73, 87].

Spectrum sensing in the interweave multi-antenna CR requires a detector that generates a defined test statistic to be compared against a specific threshold value to attain a reliable belief about the activity of PU. Such a detector can be separated into the deterministic and Bayesian categories. For the deterministic category, the test statistic can be based on energy [41], multivariate cyclostationary [73], eigenvalues of the sample covariance matrix [45, 79, 86, 87], or GLRT [16, 50, 90]. The first three kinds have limitations as they assume, respectively, the noise power is accurately known, the PU signal has a format with its cyclic frequency known, or the PU signal does not have any structure. The GLRT detector does not guarantee in general optimality for the employed test statistic [40, Ch. 6]. On the other hand, the Bayesian approach avoids estimating the unknown parameters through introducing prior distributions for them and marginalizes the likelihood function [5, 21, 40, 54]. In the Bayesian framework, the Bayes factor is used as the test statistic and it can be considered as

the odds of one hypothesis to another provided by the data. In this work, we shall regard the Bayesian detection for spectrum sensing, where the unknown parameters are the spatial covariance matrix that represents the CSI from PU to SU.

Several approaches have been proposed in the statistics community for the evaluation of the Bayes factor, please refer to [11] for a good overview. Obtaining the Bayes factor is not straightforward, due to the choice of prior, the improper prior behavior and the integration for marginalization. Indeed, utilizing proper priors for hypothesis testing is crucial to ensure a well behaved Bayes factor [11]. The conjugate prior is proper but it often yields inadequate results for Bayesian detection [54] as indicated in Section 4.4. In this chapter, we shall introduce the FBF approach to define the Bayes factor for the spectrum sensing problem that can provide better performance [59,64].

FBF can work with different priors that can be improper, and it can transform improper priors into proper ones through the concept of training samples. Specifically, FBF uses a fraction of the likelihood function to make the priors become proper and the remaining for hypothesis evaluation. Consequently, it avoids the need in determining the rather difficult non-informative or objective conventional priors (CPs) [7]. Furthermore, it is computationally attractive and does not need averaging over all the possible training data as compared to the intrinsic Bayes factor [10].

Recently, the authors of [81] considered FBF and used it for associating proper priors of the unknown parameters. Previous attempts in the engineering literature are limited to the conjugate priors [54] or CPs [61,62].

4.1.2 Chapter Organization

The organization of this chapter is as follows. Section 4.2 formulates the problem and defines the Bayes factor for spectrum sensing. Section 4.3 presents the FBF technique, introduces prior distributions for the unknown CSI parameters, and determines the expression for the associated FBF. Section 4.4 compares the proposed Bayesian detector with the state-of-art detectors and Section 4.5 concludes chapter four.

4.2 Problem Formulation

4.2.1 Problem Setup

We shall consider spectrum sensing for an interweave CR network in which SU is able to exploit the spectrum resources of the primary network whenever PU is not active. Figure 4.1 depicts the considered interweave CR model. We assume PU has l transmit and SU has p receive antennas. The transmitted signal from PU is $\mathbf{x}_k \in \mathbb{R}^{l \times 1}$, $k = 1, \dots, N$, where N is the number of samples available for detection. We shall follow [16, 21, 50, 54, 90] and consider \mathbf{x}_k is Gaussian distributed with $\mathbb{E}[\mathbf{x}\mathbf{x}^t] = \mathbf{I}$. The signal propagates through the channel represented by the matrix $\mathbf{H} \in \mathbb{R}^{p \times l}$ that is assumed static during the N sample period and reaches SU. The observed signal at SU is \mathbf{y}_k , $k = 1, \dots, N$. The collections of the transmitted and received samples form the matrices $\mathbf{X} \triangleq [\mathbf{x}_1 \cdots \mathbf{x}_N]$ and $\mathbf{Y} \triangleq [\mathbf{y}_1 \cdots \mathbf{y}_N]$.

Spectrum sensing in CR can be casted as a detection problem that intends to

distinguish between the following two hypotheses:

$$H_0 : \mathbf{Y} = \mathbf{W}, \quad (4.1a)$$

$$H_1 : \mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}, \quad (4.1b)$$

where $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_N]$, $\mathbf{w}_k \in \mathbb{R}^{p \times 1}$, is the zero-mean additive white Gaussian noise matrix. The hypotheses H_0 and H_1 in (4.1) correspond to the null (noise) model and to the data model.

It is reasonable to consider the received signal \mathbf{y}_k is IID. Thus the detection problem becomes one of choosing between two multivariate normal distributions from the observations [6, 16, 21, 54].

$$H_0 : \mathbf{y}_k \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{D}), \quad k = 1, \dots, N \quad (4.2a)$$

$$H_1 : \mathbf{y}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad k = 1, \dots, N \quad (4.2b)$$

where $\boldsymbol{\mu} \in \mathbb{R}^{p \times 1}$ is the mean vectors, $\mathbf{D} = \mathbb{E}[\mathbf{w}\mathbf{w}^t]$ is a diagonal matrix of positive diagonal elements. We do not restrict the diagonal elements to be identical to account for uncalibrated multi-antenna receivers [50, 54]. $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ is a positive definite matrix that is equal to $\boldsymbol{\Sigma} = \mathbf{H}\mathbf{H}^t + \mathbf{D}$.

The CSI for the detection problem (4.2) is not available and $\boldsymbol{\mu}$, \mathbf{D} , \mathbf{H} , and $\boldsymbol{\Sigma}$ are unknown parameters. In the development follows, we shall denote the density function for the hypothesis H_0 as $f_0(\mathbf{Y}/\boldsymbol{\mu}, \mathbf{D})$ and that for H_1 as $f_1(\mathbf{Y}/\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

4.2.2 Bayesian Detection

Let $\pi_0(\boldsymbol{\mu}, \mathbf{D})$ and $\pi_1(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ be prior distributions for the unknown parameters $(\boldsymbol{\mu}, \mathbf{D})$ and $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The marginal density functions for the hypotheses are

$$m_0(\mathbf{Y}) = \int f_0(\mathbf{Y}/\boldsymbol{\mu}, \mathbf{D}) \pi_0(\boldsymbol{\mu}, \mathbf{D}) d\boldsymbol{\mu} d\mathbf{D}, \quad (4.3a)$$

$$m_1(\mathbf{Y}) = \int f_1(\mathbf{Y}/\boldsymbol{\mu}, \boldsymbol{\Sigma}) \pi_1(\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\boldsymbol{\mu} d\boldsymbol{\Sigma}. \quad (4.3b)$$

Under the Bayesian framework, the test statistic to discriminate between the two hypotheses is the Bayes factor defined as

$$B_{10} = \frac{m_1(\mathbf{Y})}{m_0(\mathbf{Y})}. \quad (4.4)$$

Given a threshold value γ , the probability of detection and the probability of false alarm in favor of H_1 are

$$P_D = \mathbb{P}(B_{10} \geq \gamma/H_1) \quad (4.5)$$

$$P_{\text{FA}} = \mathbb{P}(B_{10} \geq \gamma/H_0), \quad (4.6)$$

where $\mathbb{P}(A)$ denotes the probability of the event A .

The Bayes factor (4.4) needs to be evaluated without any unspecified constants [11] so that the integrations in both (4.3a) and (4.3b) yield certain values. This requires that the prior $\pi_i(\cdot, \cdot)$, $i = 0, 1$, must be proper, i.e. $\int d\pi_i(\cdot, \cdot) < \infty$.

Determining proper CPs can be a challenging task [7]. Furthermore, a rich set of priors is available from Bayesian inferences and predictions and nevertheless they are improper. Rather than evaluating (4.4), we shall replace it by the FBF for detection.

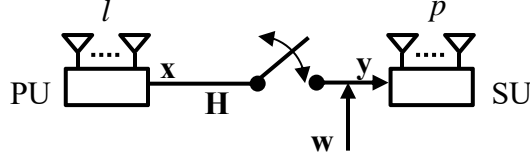


Figure 4.1: An interweave CR model between PU and SU, where each user has multiple antennas and the channel \mathbf{H} is not known.

One direct benefit is the impropriety of the priors can be easily taken care of.

The FBF for spectrum sensing that we are going to develop is for unknown parameters that are real. Extension to the complex parameters is for further study.

4.3 Proposed FBF Detector

4.3.1 FBF

FBF was first introduced by O'Hagan [64] in 1995. It applies a fraction b , $0 < b < 1$, of the likelihood $f_i(\cdot)$, $i = 0, 1$, to handle the improper behaviors of the priors and then uses the remaining $1 - b$ fraction of the likelihood to decide between the hypotheses. The FBF, B_{10}^F , is defined as [11, 64]

$$B_{10}^F = B_{10} \frac{\int f_0^b(\mathbf{Y}/\boldsymbol{\mu}, \mathbf{D})\pi_0(\boldsymbol{\mu}, \mathbf{D})d\boldsymbol{\mu}d\mathbf{D}}{\int f_1^b(\mathbf{Y}/\boldsymbol{\mu}, \boldsymbol{\Sigma})\pi_1(\boldsymbol{\mu}, \boldsymbol{\Sigma})d\boldsymbol{\mu}d\boldsymbol{\Sigma}} \quad (4.7)$$

where B_{10} is given in (4.4). The ratio term on the right of B_{10} has the purpose of cancelling out any unspecified constants in B_{10} . Thus, the FBF B_{10}^F becomes a well defined numeric value.

In (4.7), one common choice of b is m/N , where m is the minimal training sample

size [10]. Other possible values for b are $\max(m, \sqrt{N})/N$ and $\max(m, \log N)/N$ [64]. At the end of Section 4.3.4, we shall provide a requirement of b for the proposed CR detector.

Moreno [59] has shown that the FBF method can produce fractional priors, meaning that we can consider FBF as a method of generating suitable CP for hypotheses or models discrimination. The sufficient conditions for the existence of the prior are that $b = m/N$ and the models are nested.

We shall introduce next some priors for $\boldsymbol{\mu}$, \mathbf{D} and $\boldsymbol{\Sigma}$ and use them to determine B_{10}^F for the hypothesis testing problem (4.1).

4.3.2 Prior Distribution under H_0

For calculation tractability, we shall use the following conjugate proper prior for $(\boldsymbol{\mu}, \mathbf{D})$

$$\pi_0(\boldsymbol{\mu}, \mathbf{D}) \stackrel{\text{indep}}{\sim} \prod_{j=1}^p IG(h/2, \delta_{jj}/2) \quad (4.8a)$$

$$= \prod_{j=1}^p \frac{(\delta_{jj}/2)^{h/2}}{\Gamma(h/2)} d_{jj}^{-(\frac{h}{2}+1)} \exp\left(-\frac{\delta_{jj}}{2d_{jj}}\right) \quad (4.8b)$$

$$= \frac{|\boldsymbol{\Delta}|^{h/2}}{2^{hp/2} \Gamma^p(h/2)} |\mathbf{D}|^{-(\frac{h}{2}+1)} \text{etr}\left(-\frac{1}{2} \boldsymbol{\Delta} \mathbf{D}^{-1}\right) \quad (4.8c)$$

where $IG(\cdot, \cdot)$ stands for the inverse Gamma distribution, $h/2$ and $\delta_{jj}/2$, $j = 1, \dots, p$ are the shape and scale parameters of $IG(\cdot, \cdot)$ and $\Gamma(\cdot)$ denotes the gamma function. The matrix $\boldsymbol{\Delta}$ is diagonal formed by having $\delta_{jj} > 0$, $j = 1, \dots, p$, as the diagonal elements. Note that in (4.8c) the prior for $\boldsymbol{\mu}$ is uniform and independent of the prior for \mathbf{D} .

In the absence of prior information about δ_{jj} , it is customary to make the prior in (4.8) having non-influential effect by setting the hyperparameters (h, δ_{jj}) to small values such as (0.001, 0.001). We should note however that the prior $IG(\epsilon, \epsilon)$ becomes improper as $\epsilon \rightarrow 0$ [30].

Define $\tilde{h} = h + N - 1$ and the sample covariance matrix \mathbf{S} as

$$\mathbf{S} = \sum_{k=1}^N (\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^t \quad (4.9)$$

where $\bar{\mathbf{y}} = (1/N) \sum_{k=1}^N \mathbf{y}_k$ is the sample mean. It can be shown that the marginal density $m_0(\mathbf{Y})$ defined in (4.3a) under the prior distribution (4.8c) has the following closed form expression [5]

$$m_0(\mathbf{Y}) = \frac{1}{\pi^{(N-1)p/2} N^{p/2}} \frac{\Gamma^p(\tilde{h}/2)}{\Gamma^p(h/2)} \frac{|\Delta|^{h/2}}{|\Delta + \text{Diag}(\mathbf{S})|^{\tilde{h}/2}}. \quad (4.10)$$

4.3.3 Prior Distributions under H_1

We shall introduce a collection of possible improper priors for the covariance matrix Σ . It would be convenient to express Σ^{-1} in a unique decomposition form through the Cholesky factorization

$$\Sigma^{-1} = \Psi \Psi^t \quad (4.11)$$

where $\Psi \in \mathbb{R}^{p \times p}$ is an upper triangular matrix with positive diagonal elements. The off-diagonal elements are denoted by ψ_{jk} , $k > j$.

Rather than using priors on the whole Σ as in [21, 54], we shall follow [20, 83] and apply priors on the elements of Ψ . The general class of priors for $(\boldsymbol{\mu}, \Psi)$ that we

propose is

$$\pi_{1,\mathbf{a}}(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \prod_{j=1}^p \frac{1}{\psi_{jj}^{a_j}} \quad (4.12)$$

where $\mathbf{a} = [a_1, \dots, a_j, \dots, a_p]^t$. (4.12) is a common form that encompasses several priors by changing the value of \mathbf{a} . Choosing $a_j = p - j$ gives the Jeffreys' prior π_J and $a_j = p - j + 1$ the independence-Jeffreys' (left-Haar measure) prior π_{LJ} [37]. If $a_j = 2 - j$, (4.12) becomes the Geisser and Cornfield's prior π_{GC} that yields the exact frequentist matching inference for all the means and variances [29]. The right-Haar measure prior π_{RH} for exact frequentist inference is also represented by (4.12) when setting $a_j = j$. Finally, using $a_j = 1$ reduces (4.12) to the reference prior π_R [83] that is defined in [12].

We shall next obtain $m_1(\mathbf{Y})$ described in (4.3b) for the prior (4.12). Let the Cholesky factorization of the sample covariance matrix given in (4.9) be $\mathbf{S} = \mathbf{V}^t \mathbf{V}$, where $\mathbf{V} \in \mathbb{R}^{p \times p}$ is an upper triangular matrix with diagonal elements $v_{jj} > 0$, $j = 1, \dots, p$. Define the upper triangular matrix $\mathbf{T} \triangleq \mathbf{V} \boldsymbol{\Psi}$ whose diagonal elements are $t_{jj} > 0$, $j = 1, \dots, p$. Through the Jacobian of the transformation from $\boldsymbol{\Sigma}$ to $\boldsymbol{\Psi}$ and from $\boldsymbol{\Psi}$ to \mathbf{T} , we have $d\boldsymbol{\Sigma}/d\boldsymbol{\Psi} = 2^p \prod_{j=1}^p \psi_{jj}^{-2(p+1)+j}$ and $d\boldsymbol{\Psi}/d\mathbf{T} = \prod_{j=1}^p v_{jj}^{-(p-j+1)}$. As a result, the density function $f_1(\mathbf{Y}/\boldsymbol{\mu}, \boldsymbol{\Sigma})$ becomes $f_1(\mathbf{V}/\boldsymbol{\mu}, \mathbf{T})$. Since $\psi_{jj} = t_{jj}/v_{jj}$, it can be shown that $f_1(\mathbf{V}/\boldsymbol{\mu}, \mathbf{T})$ has the following expression [5],

$$f_1(\mathbf{V}/\boldsymbol{\mu}, \mathbf{T}) = \frac{2^p N^{-p/2}}{(2\pi)^{\frac{(N-1)p}{2}}} \prod_{j=1}^p \frac{v_{jj}^c}{t_{jj}^{-d_j}} \text{etr} \left(-\frac{1}{2} \mathbf{T} \mathbf{T}^t \right) \mathcal{N} \left(\boldsymbol{\mu}, \frac{\boldsymbol{\Sigma}}{N} \right) \quad (4.13)$$

where $c = p - N + 2$ and $d_j = -2(p + 1) + j + N - 1$.

We have shown in [5] that the off-diagonal elements of \mathbf{T} follows the unit normal distribution and the diagonal elements the chi-square distribution with $d_j + 1$ degrees

of freedom (dof).

Consequently, $m_1(\mathbf{Y})$ can be evaluated explicitly as

$$m_1(\mathbf{Y}) = \frac{(2\pi)^{\frac{-(N-1)p+O}{2}}}{N^{p/2}} \prod_{j=1}^p 2^{\frac{\tilde{k}_j}{2}-1} \Gamma\left(\frac{\tilde{k}_j}{2}\right) v_{jj}^{\tilde{c}_j}, \quad (4.14)$$

where $O = p(p-1)/2$, $\tilde{d}_j = (N-1) - a_j$, $\tilde{k}_j = \tilde{d}_j + 1$, $c_j = -(N-1) - (p-j+1)$, and $\tilde{c}_j = c_j + a_j$. For further details, please refer to [5].

4.3.4 Bayes Factor Evaluation

The prior for \mathbf{D} is improper as $h, \delta \rightarrow 0$, so does the class of priors (4.12) for the matrix $\mathbf{\Sigma}$. The following proposition provides a closed form expression for the test statistic B_{10}^F .

Proposition 1 [5]: The FBF B_{10}^F can be evaluated according to (4.7) as

$$B_{10}^F = K \frac{\prod_{j=1}^p 2^{\frac{\tilde{k}_j}{2}-1} \left(\frac{b}{2}\right)^{\frac{\tilde{d}_{b,j}+1}{2}} \Gamma\left(\frac{\tilde{k}_j}{2}\right) v_{jj}^{\tilde{c}_j - \tilde{c}_{b,j}}}{\Gamma\left(\frac{\tilde{d}_{b,j}+1}{2}\right)} \frac{|\mathbf{\Delta} + \text{Diag}(\mathbf{S})|^{\frac{\tilde{h}}{2}}}{|\mathbf{\Delta} + b\text{Diag}(\mathbf{S})|^{\frac{\tilde{h}_b}{2}}} \quad (4.15)$$

where the parameters $\tilde{c}_{b,j} = bc_j + a_j - (b-1)$, $\tilde{d}_{b,j} = bN - a_j - 1$, $\tilde{h}_b = h + bN - 1$ and $K = \frac{2^pb^{O/2}\Gamma^p(\tilde{h}_b/2)}{2^{(1+(b-1)N)p/2}\Gamma^p(\tilde{h}/2)}$.

Proof: We should first substitute (4.8c) and use $\pi_{1,\mathbf{a}}(\boldsymbol{\mu}, \mathbf{\Sigma})$ from (4.12) for $\pi_1(\boldsymbol{\mu}, \mathbf{\Sigma})$ in (4.7). Then we obtain from (4.10) and (4.14) the factor B_{10} defined in (4.4). Simplifying yields the FBF expression (4.15). Please refer to [5] for the details. ■

Since the gamma function $\Gamma((\tilde{d}_{b,j}+1)/2)$ in (4.15) has to be a positive value, the

following condition must be satisfied for the fraction b ,

$$b > \frac{p+1}{N}, \quad (4.16)$$

where the number of samples should be $N > p + 1$.

4.4 Numerical Results

This section presents numerical results for the hypothesis testing problem of spectrum sensing in interweave CR. We use 10^5 realizations to generate data in each hypothesis to evaluate the probability of detection and the probability of false alarm according to (4.5) and (4.6). The detection threshold γ in (4.6) for a given P_{FA} value is determined experimentally as the probability P_{FA} does not admit a closed-form expression. The number of samples to obtain the test statistics for detection is $N = 50$.

We compare the performance of the proposed FBF detector (4.15) with the ones that are based on the conjugate prior π_{C} in both hypotheses [54] and on GLRT [87]. Their test statistics are provided in Appendices 4.6.1 and 4.6.2.

Using $l = 4$ PU transmit antennas and $p = 4$ SU receive antennas, Figure 4.2 shows the average probability of detection as SNR at SU increases while keeping a fixed P_{FA} at 10^{-3} . The fraction b for the proposed FBF is set to $b = 0.1$. We examine all members of the class of improper priors: π_{J} , π_{IJ} , π_{GC} , π_{RH} , and π_{R} for FBF. They behave similarly and the difference occurs at very low SNR where the π_{RH} prior shows the lowest P_{D} while the π_{R} prior has the best detection performance. Any of the FBF outperforms the conjugate prior [54] and the GLRT test [87] detector considerably. For example to reach a $P_{\text{D}} = 0.5$, the FBF detectors require an SNR of -2.5 dB while

the other two detectors need 1 dB and 6 dB, respectively.

Figure 4.3 shows the average probability of detection versus the probability of false alarm. The SNR is -4 dB and the numbers of antennas are $l = 5, p = 5$. The fraction for FBF is $b = 0.12$. We observe again that the behaviors of the different priors defined in (4.12) for the CSI in H_1 are similar. They all provide superior performance to the detectors using the conjugate prior π_C and GLRT test. At $P_{FA} = 10^{-2}$, the corresponding probability of detection P_D for the FBF detector with π_{RH} prior, and the π_C prior and GLRT based detectors are 0.67, 0.16 and 1.1×10^{-2} .

4.5 Conclusion

In this chapter, we have developed a Bayesian detector for spectrum sensing in interweave CR networks. The proposed detector employs FBF to handle the problem of improper prior distributions in generating the Bayes factor for detection. We have introduced a new class of improper priors for the covariance matrix that represents the CSI between PU and SU. Numerical results show that the proposed FBF Bayesian detector has superior performance to that with the conjugate prior and to the GLRT test.

4.6 Appendix

4.6.1 Bayes Factor for the Conjugate Prior

The conjugate priors for the distributions (4.2a) and (4.2b) are (4.8c) and the Wishart distributions, respectively. Let u and $\mathbf{\Upsilon}$ be the dof and the scale matrix of the Wishart distribution for (4.2b). It can be verified that the Bayes factor (4.4) corresponding to the conjugate priors B_{10}^C has the following form

$$B_{10}^C = \frac{\Gamma(\tilde{u}/2)\Gamma^p(h/2)}{\Gamma(u/2)\Gamma^p(\tilde{h}/2)} \frac{|\mathbf{\Upsilon}|^{u/2}}{|\mathbf{\Upsilon} + \mathbf{S}|^{\tilde{u}/2}} \frac{|\mathbf{\Delta} + \text{Diag}(\mathbf{S})|^{\tilde{h}/2}}{|\mathbf{\Delta}|^{h/2}}, \quad (4.17)$$

where $\tilde{u} = u + N - 1$. We set $\mathbf{\Upsilon} = \mathbf{I}$ to indicate that there is no CSI about the covariance matrix $\mathbf{\Sigma}$.

4.6.2 GLRT Formula

The GLRT test statistic from [87] is

$$T_{\text{GLRT}} = (C(1 - \kappa/p)^{p-1})^{-N} \quad (4.18)$$

where $C = (1 - 1/p)^{p-1}$ and $\kappa = \lambda_1 / \sum_{i=1}^p \lambda_i$ with λ_i being the eigenvalues of the covariance matrix $\frac{1}{N}\mathbf{S}$ arranged in non increasing order.

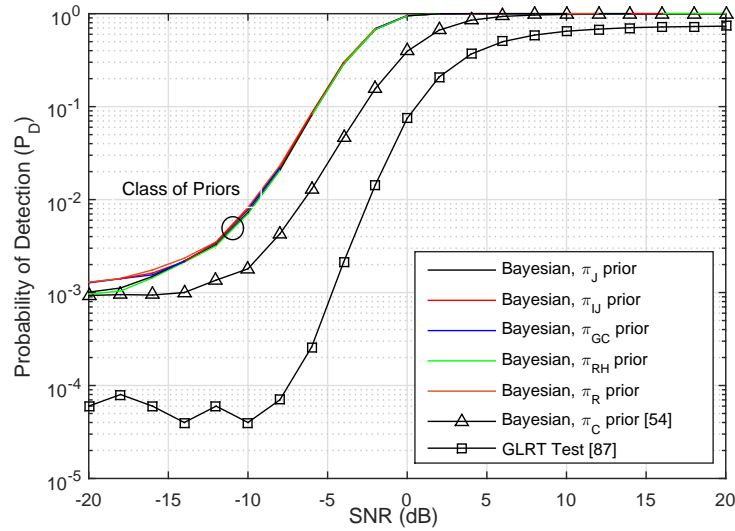


Figure 4.2: Performance of the proposed FBF, the conjugate prior [54] and the GLRT test [87] detector for PU active; the settings are $l = 4, p = 4, P_{FA} = 10^{-3}$ and $b = 0.1$.

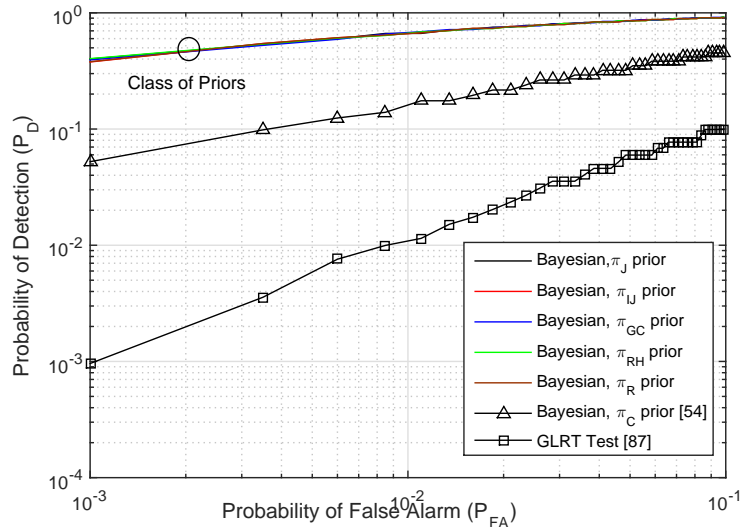


Figure 4.3: Comparison of ROC curves from different detectors for detecting PU active; the settings are $l = 5, p = 5, \text{SNR} = -4 \text{ dB}$ and $b = 0.12$.

Chapter 5

Summary and Future Work

This chapter provides a summary for our developed techniques in this thesis. It also describes the CR problems we intend to conduct in the future.

5.1 Research Summary

Chapter 2 considers a CR problem such that CSI of the link from SU-Tx to PU-Rx is uncertain or unknown while that of the link from SU-Tx to SU-Rx is perfectly known. We impose the IT and LR interference constraints to maintain the QoS of PU.

When CSI is not known, we assume that the PU network exchanges control or side information that can be used to extract some useful indicators about PU's QoS. We develop an iterative algorithm that optimizes the SU rate and satisfies QoS requirements through making use of the side information. The proposed algorithm is optimal as the number of antennas of PU-Rx goes without limit. Also, the performance of the iterative algorithm is examined through simulations under different values of Doppler

frequency.

When CSI is inaccurate, we model the uncertainty error using a generic norm called the Schatten norm. We use the worst case formulation in the interference constraint to find the most conservative solution for the precoder matrix. Particularly for the IT metric, an equivalent tractable formulation for the interference constraint at the worst case channel uncertainty matrix is proposed. For the LR metric, a local approximation through linearization is developed to obtain a tractable expression for the interference constraint. This approximation serves as an upper bound and the final solution is achieved iteratively.

The study in this chapter is completed by comparing the performance of precoder solutions from the unknown CSI and the partially known CSI. It is observed that the SU rate of the former could be better than that of the latter, and we derive the conditions for this occurrence.

In chapter 3, CSI of the links from SU-Tx to SU-Rx and from SU-Tx to PU-Rx are inaccurate. The linear precoding matrix belongs to a general convex power set and the uncertainty error elements to a convex set defined by the Schatten norm. The PU's QoS is measured using the (LR) metric. We seek the precoder matrix that maximizes MI at the worst uncertainty elements in the uncertainty regions associated with both links. If the interference constraint is fully satisfied, we encounter the compound capacity problem which has its own interest. We provide equivalent and more tractable reformulations for the mutual information function, and propose two iterative algorithms to obtain the best precoder. These algorithms exhibit different convergence behavior and computational complexity. We reduce the computational time for each iterative step through deriving analytical forms for the precoder, which

may produce a lower bound for the SU rate.

For a fixed precoder matrix, the uncertainty element of the SU's direct link is obtained using the AI approach, while that of the SU's interference link is estimated analytically. The precoder is updated iteratively using the FDPS technique until a stopping criterion is met.

We finally provide simpler and attractive solutions for the CR and CC problems. For the CR problem, we find the optimal structure when the uncertainty sets are defined by the spectral norm and the transmit power requirement is always satisfied. For the CC problem, we show that the precoder matrix has the same eigenvectors as that of the nominal channel matrix when the uncertainty set is unitarily-invariant.

Chapter 4 addresses an interweave CR network in which CSI of the covariance matrices under the null and alternative hypotheses are not known. We follow the Bayesian approach that imposes priors on the unknown parameters and use the Bayes factor as a test statistic. We invoke the FBF approach to obtain a well-defined Bayes factor when the proposed priors are improper. We further use a new class of non-informative priors in the alternative hypothesis. The numerical results show that the proposed FBF detector outperforms the state-of-the art detectors.

5.2 Future Work

We are planning to extend our work towards the following directions

5.2.1 Stochastic Modeling

We shall adopt the stochastic modeling for channel uncertainties in the forthcoming work. It is however worth noting that dealing with this modeling is challenging because of the intractability of the closed form of the probability expressions or even the lack of such expressions. Indeed, several works have focused on solving safe tractable approximations to circumvent this difficulty in the context of multiuser beamforming. Nevertheless, those approximations suffer from two folds: First, the robust solutions are conservative and considerably not close to the optimum of the original problem. Second, the approximated problem may not be feasible even though the original problem is feasible.

We can summarize our objectives for this path as follows

- We shall formulate the precoder design problem by optimizing the ergodic capacity of SU under transmit-power, number of spatial data streams, and a set of QoS-outage based interference constraints. The general structure of the optimum precoder will be derived for single and multiple PUs;
- We shall make the solutions developed in the above item are applicable to a wide range of objective functions (performance measures) under the aforementioned constraints. Our aim is to make the mathematical formulation of the proposed problem covers other communication systems in addition to the CR one;
- We shall analyze the performance of the proposed techniques in the asymptotic regime as the number of antennas goes to infinity, and devise some asymptotic algorithms to obtain the precoder.

5.2.2 Bayesian Model Selection

We shall extend the FBF approach introduced in chapter 4 for Bayesian model selection. We shall consider as well the intrinsic Bayes factor (IBF) and conjugate priors.

5.2.3 Game Theory Based CR Design

We shall consider a CR network with multiple MIMO PUs and SUs. Our objectives in this path are to address the following aspects.

- We shall consider a global interference constraint from SUs to PUs rather than the individual one. This design can provide better information rate for SU than that with conservative individuals constraints. The suggested design will be accomplished using game theory to control signaling among users;
- We shall devise a distributed method to update the precoder matrices for SUs as an alternative to the centralized fashion.

Bibliography

- [1] M. H. Al-Ali and K. C. Ho, "Robust transmit precoding for underlay MIMO cognitive radio with interference leakage rate limit," in *Proc. IEEE ICASSP*, Shanghai, China, Mar. 2016, pp. 3001-3005.
- [2] M. H. Al-Ali and K. C. Ho, "Transmit precoding in underlay MIMO cognitive radio with unavailable or imperfect knowledge of primary interference channel," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5143-5155, Aug. 2016.
- [3] M. H. Al-Ali and K. C. Ho, "Bayesian multi-antenna sensing in cognitive radio networks using fractional Bayes factor," in *Proc. IEEE ICASSP*, New Orleans, Mar. 2017.
- [4] M. H. Al-Ali and K. C. Ho, "Precoding for MIMO channels in cognitive radio networks with CSI uncertainties and for compound capacity," under 2nd review, *IEEE Trans. Signal Process.*
- [5] M. H. Al-Ali and K. C. Ho, "A new class of priors for Bayesian model selection and hypothesis testing: MIMO cognitive radio application," under preparation.

- [6] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101-116, May 2012.
- [7] M. J. Bayarri, J. O. Berger, A. Forte, and G. Garcia-Donato, "Criteria for Bayesian model choice with application to variable selection," *Ann. Stat.*, vol. 40, no. 3, pp. 1550-1577, 2012.
- [8] M. Beko, "Efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4108-4117, Nov. 2012.
- [9] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," chapter 18 in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. Boca Raton, FL, USA: CRC Press, Aug. 2001.
- [10] J. O. Berger and L. R. Pericchi, "The intrinsic Bayes factor for model selection and prediction," *J. Amer. Statist. Assoc.*, vol. 91, no. 433, pp. 109-122, Mar. 1996.
- [11] J. O. Berger and L. R. Pericchi, "Objective Bayesian methods for model selection: Introduction and comparison," *Inst. Math. Stat. Lecture Notes-Monograph Series*, vol. 38, pp. 135-207, 2001.
- [12] J. O. Berger, J. M. Bernardo, and D. Sun, "The formal definition of reference priors," *Ann. Stat.*, vol. 37, no. 2, pp. 905-938, 2009.
- [13] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, Formulas*, 2nd ed. Princeton, NJ, USA: Princeton Univ. Press, 2009.
- [14] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 2nd ed., 1999.

- [15] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.
- [16] P. Bianchi, M. Debbah, M. Maida, and J. Najim, “Performance of statistical tests for single-source detection using random matrix theory,” *IEEE Trans. Info. Theory*, vol. 57, pp. 2400-2419, Apr. 2011.
- [17] E. G. Birgin, J. M. Martínez, and M. Raydan, “Nonmonotone spectral projected gradient methods of convex sets,” *SIAM J. Optimization*, vol. 10, no.4, pp. 1196-1211, Jun. 2000.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [19] S.-M. Cai and Y. Gong, “Cognitive beamforming for throughput maximization with statistical cross channel state information,” *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 2031-2034, Aug. 2014.
- [20] G. Consonni and L. L. Rocca, “Objective Bayes covariate-adjusted sparse graphical model selection,” Arxiv preprint arXiv:1510.02245, Oct. 2011.
- [21] R. Couillet and M. Debbah, “A Bayesian framework for collaborative multi-source signal sensing,” *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5186-5195, Oct. 2010.
- [22] K. Cumanan, R. Krishna, V. Sharma, and S. Lambetharan, “Robust interference control techniques for multiuser cognitive radios using worst-case performance optimization,” in *Proc. 42nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 2008, pp. 378-382.

- [23] K. Cumanan, R. Zhang, and S. Lambotharan, "A new design paradigm for MIMO cognitive radio with primary user rate constraint," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 706-709, May 2012.
- [24] E. Dall'Anese, S.-J. Kim, G. B. Giannakis, and S. Pupolin, "Power control for cognitive radio networks under channel uncertainty," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3541-3551, Oct. 2011.
- [25] 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.
- [26] H. Du and T. Ratnarajah, "Robust joint signal and interference alignment for MIMO cognitive radio network," in *Proc. IEEE WCNC*, Paris, France, Apr. 2012, pp. 448-452.
- [27] M. Fazel, H. Hindi, and S. P. Boyd, "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices," in *Proc. Amer. Control Conf.*, Denver, CO, Jun. 2003, pp. 2156-2162.
- [28] F. Gao, R. Zhang, Y.-C. Liang, and X. Wang, "Design of learning-based MIMO cognitive radio systems," *IEEE Veh. Technol.*, vol. 59, no. 4, pp. 1707-1720, Feb. 2010.
- [29] S. Geisser and J. Cornfield, "Posterior distributions for multivariate normal parameters," *J. Roy. Stat. Soc. Series B*, vol. 25, no. 2, pp. 368-376, 1963.
- [30] A. Gelman, "Prior distributions for variance parameters in hierarchical models," *Bayesian Analysis*, vol. 1, no. 3, pp. 515-533, 2006.

- [31] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.
- [32] A. J. Goldsmith, S. A. Jafar, I. Marić, and S. Srinivasa, “Breaking spectrum gridlock with cognitive radios: An information theoretic perspective,” *Proc. IEEE*, vol. 97, no. 5, pp. 894-914, May 2009.
- [33] J. Gorski, F. Pfeuffer, and K. Klamroth, “Biconvex sets and optimization with biconvex functions - A survey and extensions,” *Math. Methods of Operation Research*, vol. 66, no. 3, pp. 373-407, May 2007.
- [34] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming,” Mar. 2014 [Online]. Available: <http://cvxr.com/cvx>
- [35] Y. Guo and B. C. Levy, “Robust MSE equalizer design for MIMO communication systems in the presence of model uncertainties,” *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1840-1852, May 2006.
- [36] Y. W. Huang, Q. Li, W.-K. Ma, and S. Z. Zhang, “Robust multicast beamforming for spectrum sharing-based cognitive radios,” *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 527-533, Jan. 2012.
- [37] H. Jeffreys, *Theory of Probability*. Oxford, NY: Oxford Univ. Press, 3rd edition, 1961.
- [38] X. Kang, H. K. Garg, Y.-C. Liang, and R. Zhang, “Optimal power allocation for OFDM-based cognitive radio with new primary transmission protection criteria,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2066-2075, Jun. 2010.

- [39] X. Kang, R. Zhang, Y.-C. Liang, and H. K. Garg, "Optimal power allocation strategies for fading cognitive radio channels with primary user outage constraint," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 374-383, Feb. 2011.
- [40] S. M. Kay, *Fundamentals of Statistical Signal Processing Detection Theory*, Upper Saddle River, NJ: Prentice-Hall, 1998.
- [41] S. Kim, J. Lee, H. Wang, and D. Hong, "Sensing performance of energy detector with correlated multiple antennas," *IEEE Signal Process. Lett.*, vol. 16, no. 8, pp. 671-674, Aug. 2009.
- [42] J. B. Lasserre, "A trace inequality for matrix product," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1500-1501, Aug. 1995.
- [43] K. L. Law, I. Wajid, and M. Pesavento, "Robust downlink beamforming in multi-group multicasting using trace bounds on the covariance mismatches," in *Proc. IEEE ICASSP*, Kyoto, Japan, Mar. 2012, pp. 3229-3232.
- [44] B. C. Levy and R. Nikoukhah, "Robust least-squares estimation with a relative entropy constraint," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 89-104, Jan. 2004.
- [45] Z. Li, D. Wang, P. Qi, and B. Hao, "Maximum eigenvalue based sensing and power recognition for multi-antenna cognitive radio system," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, 2015.
- [46] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mähönen, "Cognitive radio networking and communications: An overview," *IEEE Trans. veh. Technol.*, vol. 60, no. 7, pp. 3386-3407, Sep. 2011.

- [47] D. J. Love, R. W. Heath Jr, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An Overview of limited feedback in wireless communication systems," *IEEE J. Sel. areas Commun.*, vol. 26, no. 8, pp. 1341-1365, Oct. 2008.
- [48] S. Loyka and C. D. Charalambous, "On the compound capacity of a class of MIMO channels subject to normed uncertainty," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2048-2063, Apr. 2012.
- [49] S. Loyka and C. D. Charalambous, "Novel matrix singular value inequalities and their applications to uncertain MIMO channels," *IEEE Trans. Info. Theory*, vol. 61, no. 12, pp. 6623-6634, Dec. 2015.
- [50] R. Lopez-Valcarce, G. Vazquez-Vilar, and J. Sala, "Multiantenna spectrum sensing for cognitive radio: Overcoming noise uncertainty," in *Proc. 2nd Int. Workshop Cognit. Inf. Process. (CIP)*, Elba, Italy, Jun. 2010, pp. 310-315.
- [51] T. Luan, X.-D. Zhang, F. Gao, J. C. F. Li, and M. Lei, "Robust interference mitigation against channel uncertainty for MIMO cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 970-975, Feb. 2014.
- [52] B. Maham, P. Popovski, X. Zhou, and A. Hjørungnes, "Cognitive multiple access network with outage margin in the primary system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3343-3353, Oct. 2011.
- [53] B. Maham and P. Popovski, "Cognitive multiple-antenna network with outage and rate margins at the primary system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2409-2423, Jun. 2015.

- [54] J. Manco-Vásquez, M. Lázaro-Gredilla, D. Ramírez, J. Vía, and I. Santamaría, “A Bayesian approach for adaptive multiantenna sensing in cognitive radio networks,” *Signal Process., Elsevier*, vol. 96, part B, pp. 228-240, 2014.
- [55] A. Manolakos, Y. Noam, and A. J. Goldsmith, “Interference due to null space mismatch in cooperative multipoint MIMO cellular networks,” in *Proc. IEEE ICC*, Sydney, NSW, Jun. 2014, pp. 5148-5153.
- [56] J. Mao, J. Gao, Y. Liu, G. Xie, and X. Li, “Power allocation over fading cognitive MIMO channels: An ergodic capacity perspective,” *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1162-1173, Jan. 2012.
- [57] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer, 2nd ed., 2009.
- [58] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.
- [59] E. Moreno, “Bayes factor for intrinsic and fractional priors in nested models. Bayesian robustness,” *Inst. Math. Stat. Lecture Notes-Monograph Series*, vol. 31, pp. 257-270, 1997.
- [60] A. Nedić and D. P. Bertsekas, “Incremental subgradient methods for nondifferentiable optimization,” *SIAM J. Optimization*, vol. 12, no. 1, pp. 109-138, Jul. 2001.
- [61] J. K. Nielsen, M. G. Christensen, and S. H. Jensen, “Default Bayesian estimation of the fundamental frequency,” *IEEE Trans. Audio, Speech, and Language Process.*, vol. 21, no. 3, pp. 598-610, Mar. 2013.

- [62] J. K. Nielsen, M. G. Christensen, A. T. Cemgil, and S. H. Jensen, "Bayesian model comparison with g-prior," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 225-238, Jan. 2014.
- [63] Y. Noam, A. Manolakos, and A. J. Goldsmith, "Null space learning with interference feedback for spatial division multiple access," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5699-5715, Jul. 2014.
- [64] T. O'Hagan, "Fractional Bayes factor for model comparison (with discussion)," *J. R. Stat. Soc. Series B*, vol. 57, no.1, pp. 99-138, Jan. 1995.
- [65] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Uniform power allocation in MIMO channels: A game-theoretic approach," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1707-1727, Jul. 2003.
- [66] D. P. Palomar, "Unified framework for linear MIMO transceivers with shaping constraints," *IEEE Commun. Lett.*, vol. 8, no. 12, pp. 697-699, Dec. 2004.
- [67] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Found. Trends Commun. Inform. Theory*, vol. 3, no. 4-5, pp. 331-551, 2006.
- [68] D. S. Papailiopoulos and A. G. Dimakis, "Interference alignment as a rank constrained rank minimization," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4278-4288, Aug. 2012.
- [69] S. Parsaeefard and A. R. Sharafat, "Robust worst-case interference control in underlay cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 8, pp. 3731-3745, Oct. 2012.

- [70] A. Pascual-Iserte, D. P. Palomar, A. I. Prez-Neira, and M.A.Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346-360, Jan. 2006.
- [71] M. Pei, A. Mukherjee, A. L. Swindlehurst, and J. Wei, "Rank minimization designs for underlay MIMO cognitive radio networks with completely unknown primary CSI," in *Proc. IEEE GLOBECOM*, Anaheim, CA, Dec. 2012, pp. 1150-1155.
- [72] K. T. Phan, S. A. Vorobyov, N. D. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2323-2335, Jun. 2009.
- [73] D. Ramírez, P. J. Schreier, J. Vía, I. Santamaria, and L. L. Scharf, "Detection of multivariate cyclostationarity," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5395-5408, Oct. 2015.
- [74] B. Rustem and M. Howe, *Algorithms for Worst-case Design and Applications to Risk Management*. Princeton, NJ, USA: Princeton Univ. Press, 2002.
- [75] B. Rustem, S. Žaković, and P. Parpas, "An interior point algorithm for continuous minimax: Implementation and computation," *Optimization Methods Software*, vol. 23, no. 6, pp. 911-928, Dec. 2008.
- [76] L. Sboui, Z. Rezki, and M.-S. Alouini, "Achievable rate of spectrum sharing cognitive radio systems over fading channels at low-power regime," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6461-6473, Nov. 2014.

- [77] G. Scutari, D. P. Palomar, S. Barbarossa, “Cognitive MIMO radio,” *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 46-59, Nov. 2008.
- [78] N. Shariati, J. Wang, M. Bengtsson, “Robust training sequence design for correlated MIMO channel estimation,” *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp.107-120, Jan. 2014.
- [79] S. K. Sharma, S. Chatzinotas, and B. Ottersten, “Maximum eigenvalue detection for spectrum sensing under correlated noise,” in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 7268-7272.
- [80] M. Shuai and S. Dechun, “Chance constrained robust beamforming in cognitive radio networks,” *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 67-70, Jan. 2013.
- [81] I. Smith and F. Ferrari, “Bayesian detection with the posterior distribution of the likelihood ratio,” in *Proc.18th European Signal Process. Conf. (EUSIPCO)*, Aalborg, Denmark, Aug. 2010, pp. 144-148.
- [82] Y. Sun, C. E. Koksal, and N. B. Shroff, “Capacity of compound of MIMO Gaussian channels with additive uncertainty,” *IEEE Trans. Info. Theory*, vol. 58, no. 4, pp. 2048-2063, Apr. 2012.
- [83] D. Sun and J. O. Berger, “Objective Bayesian analysis for the multivariate normal model,” in *Bayesian Statistics*, Valencia, Spain, Jun. 2006, pp. 1-38.
- [84] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, “Multiple antenna spectrum sensing in cognitive radios,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814-823, Feb. 2010.

- [85] *3GPP Technical Specification Group Radio Access Network Physical Layer Procedures (FDD) (Release 5)*, 3rd Generation Partnership Project Std. S25.214 V5.11.0, 2005.
- [86] C. G. Tsinos and K. Berberidis, “Decentralized adaptive eigenvalue-based spectrum sensing for multiantenna cognitive radio systems,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1703-1715, Mar. 2015.
- [87] Y. Zeng and Y.-C. Liang, “Eigenvalue-based spectrum sensing algorithms for cognitive radio,” *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784-1793, 2009.
- [88] 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.
- [89] 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.
- [90] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, “Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach,” *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84-88, Jan. 2010.
- [91] R. Zhang, “On active learning and supervised transmission of spectrum sharing based cognitive radios by exploiting hidden primary radio feedback,” *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2960-2970, Oct. 2010.

- [92] Y. J. Zhang and A. M. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 362-373, Feb. 2011.
- [93] 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.
- [94] 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.
- [95] 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, Jul. 2009.
- [96] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Applicat.*, vol. 19, no. 2, pp. 499-533, 1998.
- [97] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313-324, Feb. 2003.
- [98] I. Wajid, M. Pesavento, Y. C. Eldar, and D. Ciochina, "Robust downlink beamforming with partial channel state information for conventional and cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3656-3670, Jul. 2013.

- [99] J. Wang and D. P. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3086-3100, Aug. 2009.
- [100] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5-23, Nov. 2010.
- [101] K.-Y. Wang, N. Jacklin, Z. Ding, and C.-Y. Chi, "Robust MISO transmit optimization under outage-based QoS constraints in two-tier heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1883-1897, Apr. 2013.
- [102] J. Wang, M. Bengtsson, B. Ottersten, and D. P. Palomar, "Robust MIMO precoding for several classes of channel uncertainty," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3056-3070, Jun. 2013.
- [103] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690-5707, Nov. 2014.
- [104] A. Wiesel, Y. C. Elder, and S. Shamai, "Optimization of the MIMO compound capacity," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1094-1101, Mar. 2007.
- [105] U. L. Wijewardhana, M. Codreanu, and M. Latva-aho, "Robust beamformer design for underlay cognitive radio network using worst case optimization," in *Proc. Symp. WiOpt*, Tsukuba Science City, Japan, May 2013, pp. 404-411.

- [106] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, ANSI/IEEE Std. 802.11b-1999 (R2003), 1999.
- [107] C. Xiao, Y. R. Zheng, and N.C. Beaulieu, “Novel sum-of-sinusoids simulation models for Rayleigh and Rician fading channels,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3667-3679, Dec. 2006.
- [108] C. Xiao and Y. R. Zheng, “On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs,” in *Proc. IEEE GLOBECOM*, New Orleans, LA, 2008, pp. 1-5.
- [109] B. Xie, Y. Li, H. Minn, and A. Nosratinia, “Adaptive interference alignment with CSI uncertainty,” *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 792-801, Feb. 2013.
- [110] Y. Xu, X. Zhao, and Y.-C. Liang, “Robust power control and beamforming in cognitive radio networks: A Survey,” *IEEE Commun. Survey Tutorials*, vol. 17, no. 4, pp. 1834-1857, 2015.
- [111] J. Yang, T. Q. Duong, M. ElKashlan, X. lei, and X. Gao, “Multiuser scheduling for cognitive MIMO with channel estimation errors and feedback delay,” in *IEEE ICC*, London, UK, Jun. 2015, pp. 7737-7742.
- [112] H. Yi, “Nullspace-based secondary joint transceiver scheme for cognitive radio MIMO networks using second-order statistics,” in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010, pp. 1-5.

- [113] 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.
- [114] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Comm. Surveys Tutorials*, vol. 11, no. 1, pp. 116-130, 2009.

VITA

Mohannad Hamed Al-Ali was born on November 14 1981 in Basrah Province, Iraq. He received the B.Sc. and M.Sc. degrees in Electrical Engineering from Basrah University, Basrah, Iraq, in 2003 and 2007, respectively. He joined the Electrical and Computer Engineering department at the University of Missouri-Columbia on August 2012 and since then he has begun working with Prof. Dominic Ho on his research. He expects to receive the Doctor of Philosophy degree on May 2017.

Al-Ali is married to Mrs Amal Jameel Al-Fadhel with four kids: Mohammad, Zahraa, Hasan and Lujain. After completing the PhD requirements, he will return back to his country and serve as an Assistant Professor in the Computer Engineering Department, Basrah University, Iraq.