University of California
Los Angeles

# On Optimal Transmissions in Interference Networks

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

**Yue Zhao**

2011

The dissertation of Yue Zhao is approved.

Lieven Vandenberghe

Richard D. Wesel

Mario Gerla

Gregory J. Pottie, Committee Chair

University of California, Los Angeles

2011

# Table of Contents

**1 Introduction** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **1**

1.1 Background and Motivation . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

1.3 Outline of the Thesis . . . . . . . . . . . . . . . . . . . . . . . . . . 6

**2 Spectrum Management in Gaussian Interference Channels** . . **10**

2.1 Continuous Frequency Spectrum Management . . . . . . . . . . . 14

2.1.1 Problem Model and Two Basic Co-existence Strategies . . 14

2.1.2 The Conditions for the Optimality of FDMA . . . . . . . . 19

2.1.3 Optimal Spectrum Management in Two-user Symmetric Channels . . . . . . . . . . . . . . . . . . . . . . . . . . 24

2.1.4 Optimal Spectrum Management in the General Cases . . . 37

2.1.5 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . 41

2.2 Discrete Frequency Spectrum Management . . . . . . . . . . . . . 45

2.2.1 Channel Assignment and Power Allocation: a Vertical Decomposition . . . . . . . . . . . . . . . . . . . . . . . . . 48

2.2.2 Complexity Reduction via Dual Decomposition . . . . . . 49

2.2.3 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . 51

2.3 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 52

**3 Channel Allocation in Wireless Cellular Interference Networks** **54**

3.1 System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

3.2 Uplink Channel Allocation in One Dimensional Networks: Signal Scale Interference Alignment . . . . . . . . . . . . . . . . . . . . . . . 59

3.2.1 Revenue-Cost Separation Principle . . . . . . . . . . . . . 60

3.2.2 Optimal Solution for the Local Problem . . . . . . . . . . 61

3.2.3 Optimal Solution in Infinite One Dimensional Networks . . 66

3.2.4 Performance Evaluation: Optimal CA vs. CDMA . . . . . 68

3.3 Downlink Channel Allocation in One Dimensional Networks: Decomposition and Assignment Problem . . . . . . . . . . . . . . . . 70

3.3.1 Iterative Distributed Heuristics Based on Signal Scale Interference Alignment . . . . . . . . . . . . . . . . . . . . . 73

3.3.2 Iterative Decomposed Heuristics Based on Local Assignment Problems . . . . . . . . . . . . . . . . . . . . . . . . 76

3.4 Decomposition Framework with General Network and User Settings 86

3.4.1 Two or More Dimensional Networks with Flat Channels . 86

3.4.2 Frequency Selective Channels and Additively Separable Utility Functions . . . . . . . . . . . . . . . . . . . . . . . . 88

3.5 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89

**4 Optimal Transmissions with Successive Decoding . . . . . . . . 90**

4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . 90

4.2 Problem formulation in Gaussian channels . . . . . . . . . . . . . 93

4.2.1 Gaussian Superposition Coding and Successive Decoding: a Power and Decoding Order Optimization . . . . . . . . . 93

4.2.2 Successive Decoding of Gaussian Codewords vs. Gaussian Han-Kobayashi Schemes with Joint Decoding . . . . . . . 96

4.3 Sum-capacity in deterministic interference channels . . . . . . . . . 100

4.3.1 Channel Model and Problem Formulation . . . . . . . . . . 100

4.3.2 Symmetric Interference Channels . . . . . . . . . . . . . . . 104

4.4 Approximate Sum-capacity for Successive Decoding in Gaussian Interference Channels . . . . . . . . . . . . . . . . . . . . . . . . . 115

4.4.1 Achievable Sum-rate Motivated by the Optimal Scheme in the Deterministic Channel . . . . . . . . . . . . . . . . . . . 115

4.4.2 Upper Bounds on the Sum-capacity with Successive Decoding of Gaussian Codewords . . . . . . . . . . . . . . . . . 118

4.4.3 Performance Evaluation . . . . . . . . . . . . . . . . . . . 122

4.5 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 125

4.5.1 Complementarity Conditions and Gaussian Codewords . . 126

4.5.2 Modulo-2 Additions and Lattice Codes . . . . . . . . . . . 127

**5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 129**

**A Proof of Lemma 12 and 13 . . . . . . . . . . . . . . . . . . . . . . 132**

**B Sum-capacity of Deterministic Asymmetric Interference Channels . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 135**

B.1 Sum-Capacity without Constraint on the Number of Messages . . 135

B.1.1 $\delta_1 \geq 0$ and $\delta_2 \geq 0$ . . . . . . . . . . . . . . . . . . . . . . . 135

B.1.2 $\delta_1 \geq 0$ and $\delta_2 < 0$ . . . . . . . . . . . . . . . . . . . . . . . 138

B.1.3 $\delta_1 < 0$ and $\delta_2 \geq 0$ . . . . . . . . . . . . . . . . . . . . . . . . 141

B.2 The Case with a Limited Number of Messages . . . . . . . . . . . 142

**C Proof of Lemma 14** . . . . . . . . . . . . . . . . . . . . . . . . . . . **145**

**References** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **146**

# List of Figures


1.1 $K$-user interference channel. . . . . . . . . . . . . . . . . . . . . . . . 6

2.1 Power allocations of flat frequency sharing and flat FDMA, also an illustration of flat FDMA re-allocation. . . . . . . . . . . . . . . 18

2.2 PSD compositions at receiver 0 before and after a flat FDMA re-allocation of user 1 and user 2. . . . . . . . . . . . . . . . . . . . . . . . 21

2.3 The power region in which flat FDMA has higher sum-rate than flat frequency sharing. . . . . . . . . . . . . . . . . . . . . . . . . . 26

2.4 The maximum achievable rate as the convex hull of the rates of flat FDMA and flat frequency sharing. . . . . . . . . . . . . . . . . 29

2.5 The optimal spectrum and power allocation as a mixture of flat FDMA and flat frequency sharing. . . . . . . . . . . . . . . . . . . . . 30

3.1 Three consecutive cells in a uplink one dimensional network. Users connected by lines are suggested co-channel users. . . . . . . . . . 62

3.2 After replacing $(a_1, b_{f(1)})$ and $(a_{f^{-1}(1)}, b_1)$ with $(a_1, b_1)$ and $(a_{f^{-1}(1)}, b_{f(1)})$, the total cost decreases, (or remains unchanged.) . . . . . . . . . 65

3.3 Superiority of optimal CA over CDMA . . . . . . . . . . . . . . . . 69

3.4 Three consecutive cells in a downlink one dimensional network. . . 71

3.5 Iterative distributed two-stage algorithm for channel allocation. In each stage, the solid circles denote the cells whose CA are fixed, and the open circles denote the cells whose CA are optimized as local signal scale interference alignment. . . . . . . . . . . . . . . 75

3.6 Interference Neighborhood $\mathcal{N}(C)$, Correlated Neighborhood $\mathcal{N}'(C)$ in a one dimensional network. . . . . . . . . . . . . . . . . . . . . . . 77

3.7 Correlated Neighborhood $\mathcal{N}'(C)$, where $A \in \mathcal{N}'(C)$ but $B \notin \mathcal{N}'(C)$. Shaded areas represent the interference neighborhood. . . 78

3.8 The open circles represent a maximum independent set in a one dimensional network. Each dotted circle denotes the cells that are influenced by an open circle. . . . . . . . . . . . . . . . . . . . . . . 81

3.9 Iterative decomposed three-stage algorithm for channel allocation. In each stage, the solid circles denote the cells whose CA are fixed, and the open circles denote the cells whose CA are optimized as local assignment problems. . . . . . . . . . . . . . . . . . . . . . . . 83

3.10 Performance of iterative algorithms 3.2, 3.3 with M=20. . . . . . . 85

3.11 Cells $\{1, 1_a, 1_b, 1_c, 1_d, 1_e, 1_f\}$ belongs to the same maximum independent set of cells. $\mathcal{N}(1) = \{2, 3, 4, 5, 6, 7\}$. $\mathcal{N}'(1) = \mathcal{N}(1) \cup \{2', 2'', 3', 3'', 4', 4'', 5', 5'', 6', 6'', 7', 7''\}$. . . . . . . . . . . . . 87

4.1 Two-user Gaussian interference channel. . . . . . . . . . . . . . . . 93

4.2 Our approach to solving problem (4.3). . . . . . . . . . . . . . . . . 95

4.3 Illustrations of the difference between the achievable sum-rate with Han-Kobayashi schemes and that with successive decoding of Gaussian codewords. . . . . . . . . . . . . . . . . . . . . . . . . . . 98

4.4 Two-user deterministic interference channel. Levels A and B interfere at the $1^{st}$ receiver, and cannot be fully active simultaneously. 101

4.5 Interval representation of the two-user deterministic interference channel. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103

4.6 The symmetric capacity with successive decoding in symmetric deterministic interference channels. . . . . . . . . . . . . . . . . . . 105

4.7 Segmentation of the information levels, $\frac{2n-1}{2n} < \alpha \leq \frac{2n}{2n+1}$. . . . . . . 107

4.8 Segmentation of the information levels, $\frac{2n}{2n+1} < \alpha \leq \frac{2n+1}{2n+2}$. . . . . . . 110

4.9 The symmetric capacity with a limited number of messages. . . . 114

4.10 The optimal scheme in the symmetric deterministic interference channel. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116

4.11 Performance evaluation: achievability vs. upper bounds. . . . . . 123

4.12 Sum-capacity differences: Han-Kobayashi vs. successive decoding at $\alpha = 0.75$, and successive decoding vs. the optimal single message scheme at $\alpha = 0.66$. . . . . . . . . . . . . . . . . . . . . . . 124

A.1 $C_1$ partitioned into three parts for Lemma 12. . . . . . . . . . . . 133

A.2 $C_1$ partitioned into three parts for Lemma 13. . . . . . . . . . . . 134

B.1 $n_{11} \geq n_{21}$, $n_{22} \geq n_{12}$, and $n_{22} \geq n_{21}$. . . . . . . . . . . . . . . . . . . 136

B.2 $n_{11} \geq n_{21}$, $n_{22} \geq n_{12}$, and $n_{21} > n_{22}$. . . . . . . . . . . . . . . . . . . 138

B.3 $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, $n_{22} \geq n_{21}$, and $n_{12} > n_{11}$. . . . . . . . . . . . 139

B.4 $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, $n_{22} \geq n_{21}$, and $n_{12} \leq n_{11}$, scheme I (non-optimal). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 140

B.5 $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, $n_{22} \geq n_{21}$, and $n_{12} \leq n_{11}$, scheme II (optimal). 140

B.6 $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, and $n_{22} < n_{21}$. . . . . . . . . . . . . . . . . . . 141

B.7 $n_{11} < n_{21}$ and $n_{22} \geq n_{12}$. . . . . . . . . . . . . . . . . . . . . . . . 142

# List of Tables

Abstract of the Dissertation

# On Optimal Transmissions in Interference Networks

by

**Yue Zhao**

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2011

Professor Gregory J. Pottie, Chair

An interference network is a general model of many wireline and wireless communication networks. Depending on the problem models, the complexities of maximizing the achievable rates in interference networks vary greatly. In this thesis, two types of coding - decoding assumptions on how interference is treated are considered: i) users treat interference as noise, and ii) users apply superposition coding and successive decoding. Under the former assumption, finding the optimal transmission schemes is an NP complete non-convex problem. The latter assumption complicates the problem even further to an information theoretic level.

Assuming that interference is treated as noise, we study continuous frequency spectrum management in $K$-user interference networks. A simple pairwise channel condition for FDMA schemes to achieve all Pareto optimal points of the rate region is derived. Furthermore, a convex optimization formulation is established for this classic non-convex optimization, and the main computational complexity lies in computing convex hull functions based on channel parameters. We then study the NP hard discrete frequency spectrum management, and provide

a provably optimal decomposition of the problem into channel allocation (CA) and power allocation (PA). We show that, given the optimal CA, the globally optimal PA can be solved by a convex optimization. This suggests that finding a near optimal CA is the key problem, and its combinatorial complexity is what carries the NP hardness.

Next, we investigate approaching the optimal channel allocation in large-scale wireless cellular interference networks. We develop a very low-complexity algorithm that achieves the globally optimal uplink CA in one dimensional cellular networks with flat frequency fading. The algorithm is based on local signal scale interference alignment. For networks of much more general settings, we develop a low-complexity iterative distributed CA algorithm, based on decomposed local optimizations formulated as assignment problems. The algorithm approaches the global optimum very closely.

Finally, we study the two-user Gaussian interference channel with Gaussian superposition coding and successive decoding. We first examine an approximate deterministic formulation of the problem, and show that the constrained sum-capacity oscillates as a function of the cross link gain parameters between the information theoretic sum-capacity and the sum-capacity with interference treated as noise. Furthermore, we show that if the number of messages of either user is fewer than the minimum number required to achieve the constrained sum-capacity, the maximum achievable sum-rate drops to that with interference treated as noise. By translating the optimal schemes in the deterministic channel model to the Gaussian channel model, we show that the constrained sum-capacity in the Gaussian channels oscillates between the sum-capacity with Gaussian Han-Kobayashi schemes and that with single message schemes.

# CHAPTER 1

# Introduction

## 1.1 Background and Motivation

*Interference* is one of the defining features of communication networks. Whenever there are multiple users sharing resources, coping with mutual interference arises as a fundamental problem. In wireless systems, interference may come from nearby users' signal propagation. In wireline systems, interference may come from electromagnetic coupling between bundled lines. For users transmitting independent messages, interference among the signals is in general detrimental as opposed to helpful. Thus, designing and optimizing communication schemes and systems to mitigate the undesirable effect of interference has been a long-standing quest in communications, optimization, and information theory communities.

Along with the development of large-scale communication networks throughout the past several decades, many interference mitigation methods have emerged both in practice and in theory. The general ideas include interference avoidance, interference averaging, interference cancellation and interference alignment.

### Interference Avoidance

The idea of interference avoidance is to let physically interfering users use *orthogonal* resources. Typical ways of implementing this idea include Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA). In theory, all kinds of orthogonal resources are equivalent, and can be thought of as *degrees of freedom*, or *dimensions*. (In practice, however, different ways of implementation entail different practical constraints including synchronization requirement, peak to average power ratio, etc.) Optimal orthogonal resource allocation is in essence a *scheduling* problem, or *channel allocation* problem, which is in general *NP hard* in interference networks. In spite of the NP hardness, without optimality guarantee, many dynamic channel (frequency or time) allocation heuristics have been developed [KN96].

### Interference Averaging

The idea of interference averaging is to *pool* interference from different interferers together, and *average* over them using, e.g., spread spectrum techniques. As a result of interference averaging, the effective interference strength that different users see become similar, and hence the necessity of scheduling largely disappears. *Power control*, or *power allocation*, is thus the central task in optimizing interference averaging schemes [HT99]. Power allocation has been studied in many different scenarios, including multiple access channel, interference channel, cellular networks, ad hoc networks, etc. In interference networks (which embody all of the previously listed scenarios,) the problem of power allocation is *in general a non-convex optimization* due to interference coupling [LZ08]. In spite of the non-convexity in general, there has been a rich literature in studying specialized and simpler formulation (e.g., [KG06]), as well as the general non-convex optimization of power allocation (e.g., [LZ08]).

### Interference Cancellation

The idea of interference cancellation is to let users not only decode its own message, but also decode part of the received interference from other users, and hence remove it. In general, the implementation complexity of interference cancellation is higher than treating interference as noise, and lower than joint decoding schemes. In practice, interference cancellation has long been studied in many different forms, e.g., with CDMA transceivers, with multiple antennas, in cellular uplink and downlink transmissions [And05]. Information theoretically, it has been shown to achieve the capacity region in both *multiple access* channels and general classes of *broadcast* channels [CT91]. In *interference networks* where users transmit independent messages without transmitter or receiver cooperation, however, the maximum achievable rate region using interference cancellation with arbitrary message splitting is *unknown even for the two-user case.*

**Interference Alignment**

The idea of interference alignment is to control interference from different interferers such that they *reside in fewer signal dimensions at the desired receiver.* As a novel and general philosophy of interference mitigation, interference alignment have been recently studied in two particular forms: signal vector interference alignment, and signal scale interference alignment [CJ09]. In the literature, both forms of interference alignment have been mainly focused on achieving the maximum sum *degrees of freedom* in interference networks, which has been shown to be in general an *NP hard* problem [RSL10]. Low complexity iterative algorithms on signal vector interference alignment have been devised to approach the maximum sum degrees of freedom, in finite dimensional and finite $\mathsf{SINR}$ scenarios [GCJ08].

Combination of the above four techniques can be used in interference mitigation and increasing network capacity. However, finding *optimal* schemes with the above four ideas are in general *NP hard* in the number of users, and thus cannot be efficiently solved in large-scale networks.

In addressing the fundamental difficulties of these interference mitigation ideas, it is crucial to note that there are different *problem model assumptions* underlying the above ideas. In the literature, there have been in general two types of problem models considered:

1. The model in which each user treats the interference it receives from other users as noise.

2. The information theoretic model.

The model with interference treated as noise is a baseline model which is in practice easier to satisfy, whereas the information theoretic model is more general but in practice has a higher implementation complexity. For the model with interference treated as noise, many optimization problems (including *sum-rate maximization*) have been shown to be *NP complete* in the number of users [LZ08], which prevents us from pursuing the optimal solution for large-scale networks. Although many practical heuristics with polynomial computational complexity have been developed in the literature, they cannot provide theoretical guarantees on optimality (or approximate optimality). With the information theoretic model, even the capacity region for the *two-user* interference channel remains open, for which significant progress have recently been made on *approximate* characterizations of this capacity region [ADT11]. It has been extremely hard to analyze any networks with a size *larger than two* users (except for some networks with simplified and special structures) with the information theoretic model.

This thesis studies *optimal* transmission schemes in interference networks, and how to approach them in practice:

1. For the model with interference treated as noise, we seek to close the gap

between the theoretical optimality and practical algorithms in interference network optimization. We investigate complexity reduction methods in approaching the optimal transmission schemes in general interference networks, and low-complexity distributed algorithms that closely approach the optimal performance and apply to *large-scale* wireless cellular interference networks.

2. We study the optimal transmission schemes for the model with Gaussian superposition coding and successive decoding, which *bridges* the complexity of the model with interference treated as noise and the information theoretic model. We seek to understand in what situations and how much performance gain we can get, as compared to the implementation complexity and the associated overhead by using such techniques.

## 1.2 Model

A general network model that characterizes the interference among a set of users is the *interference channel* model. As depicted in Figure 1.1, there are $K$ users each consisting of a pair of transmitter and receiver. $x_i$ and $y_i$ denote the transmitted signal from transmitter $i$ and the received signal at receiver $i$ respectively. $h_{ij}$ represents the cross channel function through which the signal from transmitter $i$ is received at receiver $j$. $z_i$ represents the noise signal at receiver $i$, superposed on the other signals that receiver $i$ sees. Interference network, as the underlying network model for many wireless/wireline communication networks, incorporates the highly non-trivial interference channel nature into communication network optimization. In this thesis, we consider *bandlimited Additive Gaussian* interfer-

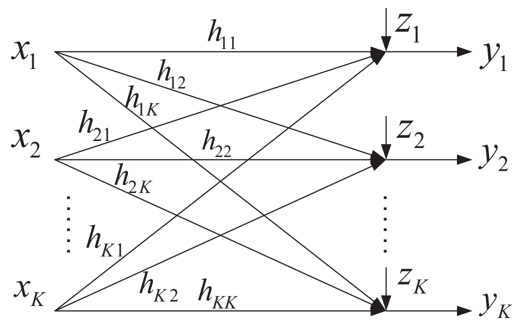



Figure 1.1: $K$-user interference channel.

ence networks,

$$y_i(f) = h_{ii}(f)x_i(f) + \sum_{j \neq i} h_{ji}(f)x_j(f) + z_i(f), \ i = 1, \ldots, K, \forall f. \tag{1.1}$$

## 1.3 Outline of the Thesis

In Chapter 2, we consider the model with interference treated as noise, and study optimal spectrum and power allocation for Gaussian interference channels (cf. Figure 1.1). The main difficulty of this problem is that it is a non-convex optimization[1] due to the interference coupling. As a result, it has been proven that most optimization problems in interference networks are NP hard [LZ08].

We first study the continuous frequency model. Intuitively, when the cross interference channel gain is sufficiently strong, it is beneficial to have users occupy orthogonal channels. We prove that, without loss of any Pareto optimality, every pair (among all) of the users can independently decide whether or not it should be orthogonalized by checking whether the cross channel gains between them satisfy a "strong coupling condition". We present the weakest, and hence the

[1]Non-convex optimization is in general not guaranteed to have efficient algorithm that solves for its globally optimal solution, whereas convex optimization in general has algorithms that run in polynomial time that solve for its globally optimal solution [BV04]

most general "strong coupling condition" for this purpose. This is an example of distributed decision making with optimality guarantees. For finding the general optimal spectrum management solution in all channel conditions, we provide a primal domain convex formulation of this classic non-convex optimization, and have a complete characterization of the optimal solution by computing the convex hull of the non-convex objective functions.

For the NP hard discrete frequency problem with $K$ users and $M$ channels, we provide a provably optimal decomposition of the problem into channel allocation (CA) and power allocation (PA). At first glance, neither of these decomposed problems is easy, as CA is a hard combinatorial optimization, and PA is again hard from its non-convexity. However, we show that given the optimal CA, an approximate optimal PA can be solved by a convex optimization, such that the achievable objective is within a constant gap from the global optimum. The problem is then "vertically" decomposed into CA and PA, and all the hardness lies only in CA. We further show that, via Lagrange dual decomposition [BV04, CYM06], the problem scale of the CA problem can be reduced to just a *single* channel, and the overall computational complexity of solving the Lagrange dual problem of optimal spectrum management is $O(M2^K)$. The vertical decomposition of CA and PA suggests that finding a near optimal CA is the central problem, and the combinatorial problem of finding the optimal CA is what carries the NP hardness.

In Chapter 3, we develop low complexity near optimal channel allocation algorithms for large-scale wireless cellular interference networks. We first show that optimizing CA is essentially a signal scale interference alignment procedure. Although this insight does not solve the NP hard combinatorial optimization of CA, by assuming *perfect* interference alignment, an efficient upper bound on the global maximum can be computed. To develop near global optimal CA algorithms, we exploit the interference as a double edged sword: while interference coupling is the defining feature of wireless networks, resulting in non-convexity and NP hardness, it also has a natural locality due to propagations losses. We establish an algorithmic framework that fully respects the effect of interference, and yet "horizontally" decomposes the optimization of CA to local *assignment problems* [BDM09]. The respect of interference ensures that the algorithm is "context-aware" (i.e., not over-simplifying the network physical layer), achieving very close to global optimal performance. The horizontal decomposition makes the algorithm distributed. It can be applied to arbitrarily large networks with low complexity.

In Chapter 4, dropping the assumption of treating interference as noise, we study the model with Gaussian superposition coding and successive decoding, and investigate the benefit from applying this more complicated scheme versus its overhead. An approximate modeling that simplifies the problem and yet captures its essence — the deterministic channel model — proves to be crucial and powerful. With the deterministic channel model, we introduce the complementarity conditions on the bit levels that capture the use of Gaussian coding and successive decoding in the Gaussian model. In the deterministic channel problem, we show that the constrained sum-capacity *oscillates* as a function of the cross link gain parameters between the information theoretic sum-capacity and the sum-capacity with interference treated as noise. Furthermore, we show that if the number of messages of either user is fewer than the minimum number required to achieve the constrained sum-capacity, the maximum achievable sum-rate drops to that with interference treated as noise. We translate the optimal schemes in the deterministic channel model to the Gaussian channel model, and also derive two upper bounds on the constrained sum-capacity. Numerical evaluations show that the constrained sum-capacity in the Gaussian channels oscillates between the sum-capacity with Gaussian Han-Kobayashi schemes [HK81] and that with single message schemes. In sum, the power of good approximate modeling is that it crystallizes the intuition of the problem, removes technical cumbersomeness as much as possible, and enables us to attack the central hardness efficiently.

Chapter 5 summarizes the main points of the thesis and suggests directions for future research.

Parts of this thesis are published in [ZP09b, ZP09a, ZP10, ZP11a, ZTA11a, ZP11b, ZTA11b]

# CHAPTER 2

# Spectrum Management in Gaussian Interference Channels

In this chapter, we consider Gaussian frequency selective interference channels (cf. Figure 4.4) in which $K(\geq 2)$ users share a common frequency bandwidth $I$. We make the assumption that interference is treated as noise at each receiver. This is in practice the simplest decoding constraint to satisfy, and is thus employed in most commercial multi-carrier communication networks. With this assumption, the two fundamental resources to allocate for each user are *power and bandwidth*. The optimization of multi-user power and bandwidth allocation in Gaussian interference channels is known in the literature as the *spectrum management* or the *spectrum balancing* problem. Depending on the constraints on power allocation, there are two different models for spectrum management problems: the *continuous* frequency model and the *discrete* frequency model. In this chapter, we mainly focus on the continuous frequency model (Section 2.1). Some key results on the discrete frequency model are provided at the end (Section 2.2) which motivate further discussions in the next chapter.

### Continuous Frequency Model

In the continuous frequency model, the channel frequency responses, the noise spectral densities, and the transmit power spectral densities (PSD) can be arbitrary *bounded piecewise continuous functions*[1] of frequency within the bandwidth of interest. Thus, in this model, there is practically no constraint on the form of the transmit PSD of each user. As a result, the transmit PSD $p_i(f), f \in I$ corresponds to an uncountably infinite number of optimization variables, for which describing the complexity of solving the optimal spectrum management in general is pointless. However, from analyzing the continuous frequency optimization problem, there are still significant insights that can be incorporated into algorithm designs in practice.

There are essentially two strategies for multiple users to co-exist: Frequency Division Multiple Access (FDMA) and frequency sharing (overlapping). As the cross coupling varies from being extremely strong to extremely weak, the preferable co-existence strategies intuitively shift from complete avoidance (FDMA) to pure frequency sharing. We start from the strong coupling scenario, and investigate the weakest interference condition under which FDMA is still guaranteed to be optimal, regardless of the power constraints. In the literature, a relatively strong pairwise coupling condition for FDMA to achieve all Pareto optimal points of the rate region is derived [EPT07]. By pairwise we mean that whether two users should be orthogonalized in frequency only depends on the interference condition between those two users. For sum-rate maximization, the required coupling strengths for FDMA to be optimal are further lowered, approaching roughly the weakest possible [HL09]. However, this condition is derived in a group-wise form, requiring the couplings between all users to be sufficiently strong.

In this chapter, the weakest possible pairwise condition for FDMA to achieve all Pareto optimal points of the rate region is proved: for any two (among all of the $K$) users, as long as the two normalized cross channel gains between them are

[1] More generally, Lebesgue integrable functions can be considered [LZ08].

both larger than or equal to 1/2, an FDMA allocation between these two users benefits every one of the $K$ users. When the cross channel gain is less than 1/2 in symmetric channels, we precisely characterize the non-empty power constraint region within which frequency sharing between two users leads to a higher rate than an FDMA allocation between them.

For the general spectrum management problem, a variety of commonly used objective functions (including weighted sum-rate maximization) lead to *non-convex optimization*, constituting the main difficulty in spectrum management. To attack this difficulty, we develop a new method that transforms the problem in the primal domain into an equivalent convex optimization with the continuous frequency model.

1. We begin with sum-rate maximization in two-user symmetric flat channels. We show that the optimal spectrum management can be solved by computing a convex hull function. As a result, the optimal spectrum management always consists of one sub-band of flat frequency sharing and one sub-band of flat FDMA. The optimal solution for the sum-rate maximization was also independently derived in [SZB08] for two-user asymmetric flat channels, and in [BHB10] for $K$-user ($K \geq 2$) symmetric flat channels.

2. We first generalize our results to two-user symmetric frequency selective channels, and show that a convex relaxation of the original non-concave objective actually leads to the same optimal value as the original problem. Next, we generalize our results to $K$-user asymmetric flat channels for arbitrary weighted sum-rate maximization, and show that the optimal solution can be found by computing a convex hull function. Finally, we combine the ideas of these generalizations, and establish the equivalent primal domain convex optimization for the spectrum management problem in its general

form, i.e., arbitrary weighted sum-rate maximization for $K$-user $(K \geq 2)$ asymmetric frequency selective channels.

### Discrete Frequency Model

In the discrete frequency model, *piecewise constant* channel frequency responses and the noise spectral densities are considered. As a result, the entire bandwidth becomes a set of *frequency flat* sub-channels $(I =) I_1 \cup I_2 \cup \ldots \cup I_M$. The key constraint that makes the discrete frequency model different from the continuous one is that the *transmit PSD must be constant in each flat sub-channel.* Accordingly, the optimization variables are $\{p_i(m), i = 1, \ldots, K, m = 1, \ldots, M\}$, where $p_i(m)$ is user $i$'s constant PSD in sub-channel $m$.

In the literature, the discrete frequency model has attracted enormous research effort in the past decade. With the commonly used non-concave objective function, the spectrum management problem with the discrete frequency model has been shown to be NP complete in the number of users even for the single carrier case [LZ08]. For the single carrier sum-rate maximization problem, two special cases have been solved: the two-user case of all channel parameters [EMK06, GGO06], and the $K$-user $(K \geq 2)$ case of fully symmetric channels [BHB10].

For the multi-carrier weighted sum-rate maximization problem, dual decomposition methods have been widely applied to decompose the problem in frequency [CHC07, CYM06, YL06]. While these methods effectively reduced the scale of the problem to solve, two remaining issues are as follows.

- While the dual master problem is a convex optimization (which can be solved by e.g. subgradient method [YL06],) the single carrier sub-problem is still an NP-complete non-convex optimization.

- The dual optimal solution does not necessarily give a primal optimal solution.

Addressing the second issue, a significant result is that the duality gap of the spectrum management problem goes to zero as the number of sub-channels goes to infinity, under mild technical conditions [YL06, LZ08].

In the literature, convexifications of the weighted sum-rate objective have been proposed so that the problem is approximated as convex optimizations [Chi05a]. However, the gap between the optimum of the convex approximation and the original problem can be unbounded. Addressing this problem, we decompose the spectrum management problem into two steps: channel allocation and power allocation. We show that provided with the optimal *channel* allocation, the solution of a convex approximation of the *power* allocation problem can be guaranteed to be within a constant gap from the global optimum of the original problem. This suggests that it is finding the optimal channel allocation that carries the NP hardness of the overall spectrum management problem.

## 2.1 Continuous Frequency Spectrum Management

### 2.1.1 Problem Model and Two Basic Co-existence Strategies

#### 2.1.1.1 Channel Model and Rate Density Function

As depicted in Figure 4.4, a $K$-user Gaussian interference channel is modeled by (1.1). WLOG, we assume that the channel is over a unit bandwidth frequency band $[0, 1]$. The results derived directly generalize to frequency bands with arbitrary bandwidths. The channel frequency selectivity is characterized by the channel gain functions $\{h_{ij}(f), 1 \leq i, j \leq K\}$ and the noise spectral density

$\{\sigma_i(f), 1 \le i \le K\}$. Denote the transmit PSD of user $i$ by $p_i(f)$, and define

$$\mathbf{p}(f) \triangleq [p_1(f), p_2(f), \ldots, p_K(f)]^T, \forall f \in [0,1]. \tag{2.1}$$

We assume that $\{h_{ij}(f), 1 \le i, j \le K\}, \{\sigma_i(f), p_i(f), 1 \le i \le K\}$ are all *piecewise bounded continuous functions* over the band $f \in [0,1]$, with a *finite* number of discontinuities.

We assume that every user uses a random Gaussian codebook, and only decodes the signal from its own transmitter, treating interference from other transmitters as noise. Employing the Shannon capacity formula for Gaussian channels, we have the following achievable rate for user $i$:

$$\begin{aligned} R_i &= \int_0^1 \log\left(1 + \frac{p_i(f)|h_{ii}(f)|^2}{\sigma_i(f) + \sum_{j\neq i} p_j(f)|h_{ji}(f)|^2}\right) df \\ &= \int_0^1 \log\left(1 + \frac{p_i(f)}{n_i(f) + \sum_{j\neq i} p_j(f)\alpha_{ji}(f)}\right) df, \end{aligned} \tag{2.2}$$

where $\alpha_{ji}(f) \triangleq |\frac{h_{ji}(f)}{h_{ii}(f)}|^2, n_i(f) \triangleq \frac{\sigma_i(f)}{|h_{ii}(f)|^2}$ are the cross channel gains and the noise power normalized by the direct channel gains. We further make a technical assumption that

$$\exists n_\epsilon > 0,\ s.t.\ \forall f \in [0,1], n_i(f) \ge n_\epsilon, \forall i = 1, \ldots, K. \tag{2.3}$$

which naturally holds in all physical channels. As we consider the continuous frequency model, we define the *rate density function* as follows:

**Definition 1.** $\forall f \in [0,1]$, *with* $\mathbf{P} = [P_1, P_2, \ldots, P_K]^T$,

$$r_i(\mathbf{P}, f) \triangleq \log\left(1 + \frac{P_i}{n_i(f) + \sum_{j\neq i} P_j \alpha_{ji}(f)}\right). \tag{2.4}$$

*The rate density function of user* $i$ *at frequency* $f$ *is*

$$r_i(\mathbf{p}(f), f) \triangleq \log\left(1 + \frac{p_i(f)}{n_i(f) + \sum_{j\neq i} p_j(f)\alpha_{ji}(f)}\right), \tag{2.5}$$

*and* $\mathbf{r}(\mathbf{p}(f), f) \triangleq [r_1(\mathbf{p}(f), f), r_2(\mathbf{p}(f), f), r_K(\mathbf{p}(f), f)]^T$.

Accordingly, the rate of user $i$ can be re-written as

$$R_i = \int_0^1 r_i(\mathbf{p}(f), f)df, i = 1, \ldots, K. \tag{2.6}$$

#### 2.1.1.2 Piecewise Continuous Functions as Limits of Piecewise Flat Functions

We consider the channel responses and power allocations as piecewise bounded continuous functions of frequency. Intuitively, one may approximate continuous functions by piecewise constant functions, by subdividing the support (frequency) to a sufficiently large number of small pieces. We make use of this idea in this chapter, and provide a technical lemma for this purpose.

**Lemma 1** (Approximation Lemma). *Given* $\{p_i(f)\}, \{\alpha_{ji}(f)\}, \{n_i(f)\}, f \in [0,1]$ *all piecewise bounded continuous, for any utility function* $\mathbf{U}(\mathbf{p}, \alpha, \mathbf{n})$ *that is a uniformly continuous function of* $\{p_i, \alpha_{ji}, n_i, 1 \leq i, j \leq K\}$, $\forall \epsilon > 0$, *there exists a set of piecewise flat power allocation functions and channel responses,*

$$\bar{\mathbf{p}}(f) = [\bar{p}_1(f), \ldots, \bar{p}_K(f)]^T, \{\bar{\alpha}_{ji}(f)\}, \{\bar{n}_i(f)\}, f \in [0,1],$$

*for which the band is divided into* $M(< \infty)$ *intervals* $I_1, \ldots, I_M, I_m = [f_{m-1}, f_m]$, *with* $f_0 = 0, f_M = 1, f_{m-1} < f_m$, *and*

$$\forall m, \forall f \in I_m, \bar{\mathbf{p}}(f) = \mathbf{P}(m), \bar{\alpha}_{ji}(f) = \alpha_{ji}(m), \bar{n}_i(f) = n_i(m), \forall i, j, \tag{2.7}$$

*where* $\mathbf{P}(m) = [P_1(m), \ldots, P_K(m)]^T, \{\alpha_{ji}(m), n_i(m)\}$ *are constants that only depend on the interval index* $m$, *such that the following three properties hold:*

*P1.* $\forall f \in [0,1], \bar{p}_i(f) \leq p_i(f), \forall i = 1, \ldots, K$;

*P2.* $\forall f \in [0,1], \bar{\alpha}_{ji}(f) \geq \alpha_{ji}(f), \forall i \neq j, \bar{n}_i(f) \geq n_i(f), \forall i$;

*P3.* $\forall f \in [0,1], |U(\bar{\mathbf{p}}(f), \bar{\alpha}(f), \bar{\mathbf{n}}(f)) - U(\mathbf{p}(f), \alpha(f), \mathbf{n}(f))| < \epsilon$.

From now on, we name the $\bar{\mathbf{p}}(f), \{\bar{\alpha}_{ji}(f)\}$ and $\{\bar{\mathbf{n}}(f)\}$ found in Lemma 1 a "piecewise flat $\epsilon$-approximation".

**Remark 1.** *Property P1 ensures that the approximate piecewise flat power allocations consume less power than the original ones. Property P2 ensures that the approximate piecewise flat channel responses are "worse" than the original ones (as the cross channel gains and the noise power are all stronger, and interference is treated as noise.) Nonetheless, property P3 ensures that under these "adverse" conditions, these approximations can still achieve the original utility U arbitrarily closely.*

We note that with finite power constraints and non-degenerate channel parameters (2.3), most utility functions considered in practice (e.g. a weighted sum-rate) satisfy the uniform continuity condition of $U(\mathbf{p}, \alpha, \mathbf{n})$.

#### 2.1.1.3 Two Basic Co-existence Strategies and One Basic Transformation

There are essentially two co-existence strategies for users to reside in a common band: frequency sharing and FDMA. We introduce two basic forms of these two strategies: Flat Frequency Sharing and Flat FDMA, both defined in flat channels. We will see that these two basic strategies are the building blocks of general non-flat co-existence strategies in frequency selective channels. Consider a two-user flat channel:

$$\forall f \in [0,1], n_1(f) = n_1, n_2(f) = n_2, \alpha_{21}(f) = \alpha_{21}, \alpha_{12}(f) = \alpha_{12}. \tag{2.8}$$

**Definition 2.** *A flat frequency sharing scheme of two users is any power allocation in the form of*

$$\exists p_1, p_2, \forall f \in [0,1], p_1(f) = p_1, p_2(f) = p_2. \tag{2.9}$$

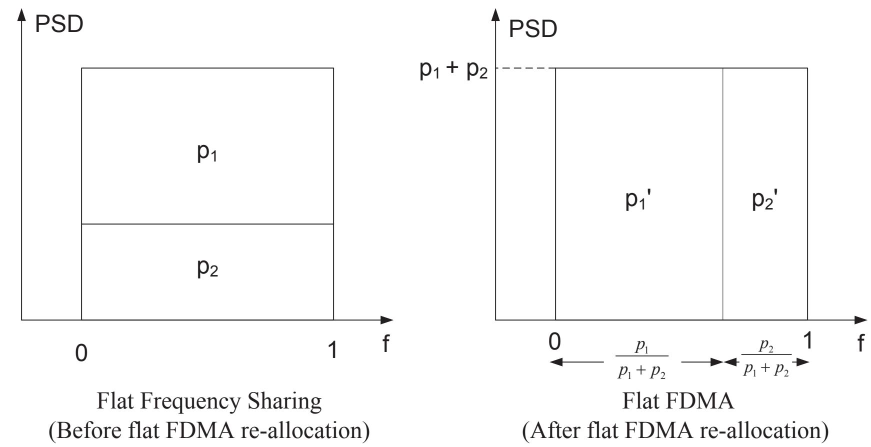



Figure 2.1: Power allocations of flat frequency sharing and flat FDMA, also an illustration of flat FDMA re-allocation.

**Definition 3.** *A flat FDMA scheme of two users is any power allocation in the form of*

$$\exists p, \forall f \in [0,1], p_1(f)p_2(f) = 0, p_1(f) + p_2(f) = p. \tag{2.10}$$

**Definition 4.** *A flat FDMA reallocation is the following power invariant transform that reallocates the power of the two users from a flat frequency sharing scheme to a flat FDMA scheme:*

*User 1 and user 2 reallocate their power within sub-bands of bandwidths* $\frac{p_1}{p_1+p_2}$ *and* $\frac{p_2}{p_1+p_2}$ *respectively, both with a flat PSD of* $p_1' = p_2' = p_1 + p_2$.

Illustrations of the power allocations of the two basic co-existence strategies before and after a flat FDMA re-allocation are depicted in Figure 2.1. Clearly, the total power of each user does not change after a re-allocation. Similarly, flat frequency sharing schemes, flat FDMA schemes, and flat FDMA re-allocation can be defined for any $k(= 1, 2, \ldots)$ users. ($k = 1$ is the degraded case in which flat frequency sharing is the same as flat FDMA.)

**Remark 2.** *A flat FDMA scheme is mathematically the same as multiple dis-*

*joint bands each seeing a flat frequency sharing of only one user. Thus, it is actually sufficient to only define flat frequency sharing schemes of any* $k(= 1, 2, \ldots)$ *users, without introducing the definition of flat FDMA schemes. This alternative approach is used later in Section 2.1.4 for the general optimization in* $K$*-user frequency selective channels. Here, flat FDMA and flat FDMA re-allocation are explicitly defined, because they offer clear intuitions for optimizing spectrum management as will be shown in the following sections.*

### 2.1.2 The Conditions for the Optimality of FDMA

We investigate the conditions under which the optimal spectrum and power allocation takes the form of FDMA. We show that our results apply to all Pareto optimal points of the achievable rate region, by proving the following simple pairwise condition:

*For any two of the* $K$ *users, as long as the normalized cross channel gains between them are both larger than or equal to* $1/2$*, every one of the* $K$ *users will benefit from an FDMA allocation between these two users.*

We arrive at this result by two steps:

1. We show the coupling condition under which FDMA achieves all Pareto optimal rate tuples within a group of strongly coupled users.

2. For the users outside a strongly coupled group, we show that they always benefit from an FDMA allocation within the strongly coupled group.

#### 2.1.2.1 The Optimality of FDMA within Strongly Coupled Users

We now give a sufficient condition for $K$-user interference channels under which FDMA among all users can achieve any Pareto optimal rate tuple. This condition requires that between every pair of users, the normalized cross channel gains must be stronger than $\frac{1}{2}$. We begin with two-user flat channels, and extend the results to $K$-user frequency selective channels.

**Theorem 1.** *Consider a two-user flat interference channel* (2.8)*. Suppose the two users co-exist in a flat frequency sharing manner* (2.9)*. If* $\alpha_{12} \geq \frac{1}{2}$ *and* $\alpha_{21} \geq \frac{1}{2}$*, then with a flat FDMA power re-allocation, both users' rates will be higher (or unchanged.)*

Theorem 1 can be generalized to the $K$-user case as follows.

**Corollary 1.** *Consider a $K$-user flat interference channel,* $\alpha_{ji}(f) = \alpha_{ji}, n_i(f) = n_i$*. Suppose the $K$ users co-exist in a flat frequency sharing manner:* $\mathbf{p}_i(f) = \mathbf{p}_i, \forall f \in [0, 1]$*. If* $\forall j \neq i, \alpha_{ji} \geq \frac{1}{2}$*, then with a flat FDMA power re-allocation, all users' rates will be higher or unchanged.*

Generalization to frequency selective channels also immediately follows.

**Corollary 2.** *Consider a $K$-user frequency selective interference channel. Suppose we have an arbitrary spectrum and power allocation scheme* $\mathbf{p}(f)$ *with some frequency sharing (overlapping) in the band. If* $\alpha_{ji}(f) \geq \frac{1}{2}, \forall j \neq i, \forall f \in [0, 1]$*, we can always find an FDMA power re-allocation scheme* $\tilde{\mathbf{p}}(f)$*, satisfying* $\int_0^1 \tilde{p}_i(f)df = \int_0^1 p_i(f)df, i = 1, \ldots, K$*, with which all users' rates are higher or unchanged.*

We summarize the above results as follows: Pick any sub-band $[f_1, f_2]$, as long as all the users are strongly coupled by $\forall j \neq i, \forall f \in [f_1, f_2], \alpha_{ji}(f) \geq \frac{1}{2}$, then for any power allocation scheme having frequency sharing used anywhere within $[f_1, f_2]$, there always exists an FDMA power re-allocation scheme (with every user's total power unchanged) that leads to a rate higher than or equal to the original sharing scheme for every user.

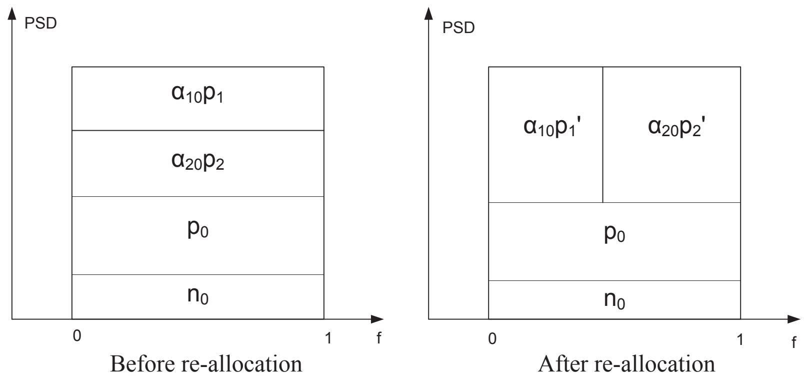



Figure 2.2: PSD compositions at receiver 0 before and after a flat FDMA re-allocation of user 1 and user 2.

#### 2.1.2.2 FDMA Within a Subset of Users Benefits All Other Users

We have seen that by properly separating a group of strongly coupled users to orthogonal channels, every user among them will have a rate higher than or equal to the rate of any frequency sharing (overlapping) scheme. In this section, we show that an FDMA allocation among a group of users also benefits all other users outside this group. This result completes the fundamental fact that, all $K$-user Pareto optimal rate tuples can be achieved by all the strongly coupled users (among all the users) separated into disjoint frequency bands. We begin with the two-interferer flat channels, and extend the results to $K$-interferer frequency selective channels.

**Lemma 2.** *Consider a three-user (one user + two interferers) flat channel: $\forall i, j, \alpha_{ji}(f) = \alpha_{ji}, n_i(f) = n_i$. Suppose the three users co-exist in a flat frequency sharing manner $p_i(f) = p_i, \forall f \in [0, 1], i = 0, 1, 2$: From user 0's perspective, a flat FDMA power re-allocation of its two interferers, namely user 1 and user 2, always leads to a rate higher than or equal to the original rate for user 0.*

*Proof.* At the receiver of user 0, the received PSDs before and after the flat

FDMA power re-allocation of its interferers are depicted in Figure 2.2. User 0's rates before and after the re-allocation are

$$R_0 = \log\left(1 + \frac{p_0}{\alpha_{10}p_1 + \alpha_{20}p_2 + n_0}\right),$$
$$R_0' = \frac{p_1}{p_1+p_2}\log\left(1 + \frac{p_0}{\alpha_{10}(p_1 + p + 2)}\right) + \frac{p_2}{p_1+p_2}\log\left(1 + \frac{p_0}{\alpha_{20}(p_1+p_2) + n_0}\right). \tag{2.11}$$

With straightforward calculations, one can verify that the function $\log(1 + \frac{P}{I+N})$ is convex in I. Therefore, By Jensen's Inequality, $R_0' > R_0, \forall p_1, p_2 \geq 0$. □

Lemma 2 can be generalized to an arbitrary number of users as in the following corollary.

**Corollary 3.** *Consider a $K+1$-user (one user + $K$ interferers) flat channel: $\forall i, j, \alpha_{ji}(f) = \alpha_{ji}, n_i(f) = n_i$. Suppose the $K+1$ users co-exist in a flat frequency sharing manner: $\forall f \in [0,1], \mathbf{p}(f) = \mathbf{p}$. From user 0's perspective, a flat FDMA power re-allocation of its $K$ interferers, namely user $1, 2, \ldots, K$, always leads to a rate higher than or equal to the original rate for user 0.*

Finally, the benefits of an FDMA within a subset of users to the other users can be generalized to frequency selective channels.

**Corollary 4.** *Consider a $K+1$-user (one user + $K$ interferers) frequency selective channel. Suppose we have an arbitrary spectrum and power allocation scheme $p_i(f), i = 0, 1, \ldots, K$, in which user $1, \ldots, K$ are not completely FDMA. Then, from user 0's perspective, there is always a corresponding FDMA power re-allocation of its $K$ interferers, namely user $1, \ldots, K$, that leads to a rate higher than or equal to the original rate for user 0.*

*Proof.* $\forall \epsilon > 0$, by Lemma 1, take a piecewise flat $\epsilon$-approximation $\bar{\mathbf{p}}(f), \{\bar{\alpha}_{ji}(f)\}$, $\{\bar{n}_i(f)\}$, such that

$$|\bar{R}_0 - R_0| < \epsilon,$$

where $\bar{R}_0$ is user 0's rate computed with $\bar{\mathbf{p}}(f), \{\bar{\alpha}_{ji}(f)\}, \{\bar{n}_i(f)\}$. If $\bar{p}_1(f), \ldots, \bar{p}_K(f)$ is not completely FDMA yet, do a flat FDMA reallocation to $\bar{p}_1(f), \ldots, \bar{p}_K(f)$ in every flat sub-channel that has a flat frequency sharing of any subset of the $K$ interferers. By Corollary 3, the resulting rate of user 0 $\bar{R}'_0$ satisfies $\bar{R}'_0 \geq \bar{R}_0 > R_0 - \epsilon$. Finally, let $\epsilon \to 0$. □

#### 2.1.2.3 Pairwise Condition for the Optimality of FDMA

Combining Theorem 1 and Lemma 2.2, we arrive at the following conclusion:

**Theorem 2.** *For any two users $i$ and $j$ (among all the $K$ users), for any frequency band $[f_1, f_2]$, if the normalized cross channel gains $\alpha_{ji}(f) \geq \frac{1}{2}, \alpha_{ij}(f) \geq \frac{1}{2}, \forall f \in [f_1, f_2]$, then no matter from which user's point of view, an FDMA of user $i$ and user $j$ within this band is always preferred.*

*Proof.* Suppose the spectrum and power allocation for user $i$ and $j$ are not FDMA, take a piecewise flat $\epsilon$-approximation $\bar{\mathbf{p}}(f), \{\bar{\alpha}_{ji}(f)\}, \{\bar{n}_i(f)\}$ such that $|\bar{R}_k - R_k| < \epsilon, k = 1, \ldots, K$. As in the proof of Corollary 2.2, with a flat FDMA reallocation of $\bar{p}_i(f)$ and $\bar{p}_j(f)$ in every flat sub-channel in $[f_1, f_2]$ that has a flat frequency sharing of user $i$ and $j$,

- Theorem 1 implies that user $i$ and $j$'s rates are increased or unchanged;
- Lemma 2.2 implies that every one of the other $K-2$ users' rate is increased or unchanged.

Finally, let $\epsilon \to 0$. □

The pairwise condition $\alpha_{ji} \geq \frac{1}{2}, \alpha_{ij} \geq \frac{1}{2}$ makes determining whether any two users should be orthogonally channelized depend only on the coupling conditions between the two of them. Furthermore, since this condition guarantees that an FDMA allocation between user $i$ and user $j$ benefits every one of the $K$ users, under this condition, all the Pareto optimal points of the rate region can be achieved with these two users having an FDMA allocation. Thus, this pairwise condition is an example of distributed decision making (on whether to orthogonalize any pair of users) with optimality guarantees.

### 2.1.3 Optimal Spectrum Management in Two-user Symmetric Channels

In this section, we start to investigate solving the complete optimal spectrum management scheme, and provide a complete solution of two-user (potentially frequency selective) *symmetric* Gaussian interference channels, defined as follows:

$$\alpha_{12}(f) = \alpha_{21}(f), n_1(f) = n_2(f), \forall f \in [0, 1]. \tag{2.12}$$

Generalizations are provided later in Section 2.1.4.

We consider the objective of sum-rate maximization. We assume an equal power constraint, or equivalently, a sum-power constraint:

$$\begin{aligned} &\max_{p_1(f), p_2(f)} R_1 + R_2 && \Leftrightarrow \quad \max_{p_1(f), p_2(f)} R_1 + R_2 \\ &s.t. \int_0^1 p_i(f) df \leq p/2, i = 1, 2 && \quad\;\; s.t. \int_0^1 (p_1(f) + p_2(f)) df \leq p \end{aligned} \tag{2.13}$$

(Equivalency is shown later in this section.)

We begin with flat channels, and solve the optimal spectrum and power allocation by computing a convex hull. Based on this result, we show that finding the spectrum and power allocation that maximizes the non-concave sum-rate objective in symmetric frequency selective channels can be equivalently transformed into a convex optimization in the primal domain.

#### 2.1.3.1 Optimal Solutions for Flat Channels

Consider a two-user symmetric flat Gaussian interference channel model:

$$\alpha_{12}(f) = \alpha_{21}(f) = \alpha < \frac{1}{2},\ n_1(f) = n_2(f) = n, \forall f \in [0,1]. \tag{2.14}$$

**Remark 3.** *For the case of $\alpha_{12}(f) = \alpha_{21}(f) \geq \frac{1}{2}, \forall f \in [0,1]$, from Theorem 1, the optimal solution is an FDMA allocation in which each user uses half of the total bandwidth and a uniform PSD therein.*

WLOG, we can normalize the power and their constraints by the noise: $p_i(f) \leftarrow \frac{p_i(f)}{n}, p \leftarrow \frac{p}{n}$, and assume $n = 1$.

**Definition 5.** *Define $r^o(p)$ to be maximum achievable sum-rate with a sum-power constraint $p$:*

$$r^o(p) \triangleq \max_{p(f)} \int_0^1 r_1(\mathbf{p}(f), f) + r_2(\mathbf{p}(f), f) df \tag{2.15}$$

$$\text{s.t.} \int_0^1 (p_1(f) + p_2(f))\, df \leq p,\ p_i(f) \geq 0, i = 1, 2, \forall f \in [0,1],$$

$$r_1(\mathbf{p}(f), f) = \log\left(1 + \frac{p_1(f)}{\alpha p_2(f)}\right), r_2(\mathbf{p}(f), f) = \log\left(1 + \frac{p_2(f)}{\alpha p_1(f)}\right). \tag{2.16}$$

Firstly, we have the following theorem on the condition under which a flat FDMA scheme is better than a flat frequency sharing scheme. Denote by $p_i$ the PSD of user $i$ in a flat frequency sharing scheme.

**Lemma 3.** *For any flat frequency sharing power allocation, a flat FDMA power re-allocation (Figure 2.1) leads to a higher or unchanged sum-rate if and only if*

$$p_1 + p_2 \geq 2\left(\frac{1}{2\alpha^2} - \frac{1}{\alpha}\right). \tag{2.17}$$

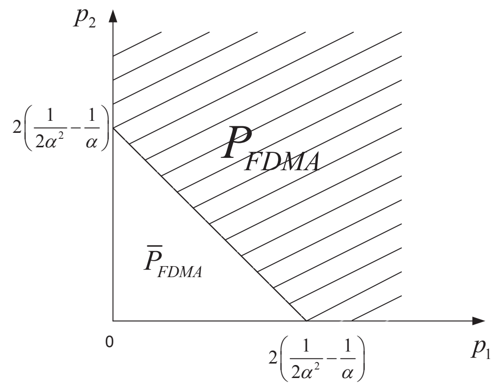



Figure 2.3: The power region in which flat FDMA has higher sum-rate than flat frequency sharing.

Given the cross channel gains $\alpha$, Lemma 3 provides us a power region $P_{FDMA}$ within which flat FDMA has a higher sum-rate than flat frequency sharing, depicted as the shaded area in Figure 2.3 (with the complement region $\bar{P}_{FMDA}$ also depicted). Clearly, if and only if $\alpha \geq \frac{1}{2}$, $P_{FDMA}$ contains the entire non-negative quadrant. This provides a "weak" converse argument on the necessity of the $\frac{1}{2}$ coupling condition derived in Section 2.1.2, for FDMA to be always optimal regardless of the power budget.

Next, we derive the optimal flat frequency sharing scheme and the optimal flat FDMA scheme.

Denote the sum-rate of a flat frequency sharing by

$$f(p_1, p_2) = \log(1 + \frac{p_1}{1 + \alpha p_2}) + \log(1 + \frac{p_2}{1 + \alpha p_1}). \tag{2.18}$$

The maximum achievable sum-rate with flat frequency sharing with a sum-power constraint, denoted by $f^*(p)$, is defined as the optimal value of the following optimization problem:

**Definition 6.**

$$f^*(p) \triangleq \max_{p_1 \geq 0, p_2 \geq 0} f(p_1, p_2)$$
$$s.t. \; p_1 + p_2 \leq p. \tag{2.19}$$

We have the following lemma on the form and the concavity of $f^*(p)$ in the region of $P_{FDMA}$.

**Lemma 4.** *When* $0 < p \leq 2\left(\frac{1}{2\alpha^2} - \frac{1}{\alpha}\right)$,

$$f^*(p) = 2 \log \left(1 + \frac{p/2}{1 + \alpha p/2}\right) \tag{2.20}$$

*is a concave function of the constraint* $p$. *The optimal flat frequency sharing scheme is* $p_1 = p_2 = \frac{p}{2}$.

In comparison, we compute the maximum achievable sum-rate with a sum-power constraint for FDMA schemes, denoted by $h^*(p)$:

**Definition 7.**

$$h^*(p) \triangleq \max_{p_1(f) \geq 0, p_2(f) \geq 0} R_1 + R_2 \tag{2.21}$$
$$s.t. \; \int_0^1 (p_1(f) + p_2(f)) df \leq p, \; p_1(f) p_2(f) = 0, \forall f \in [0, 1],$$
$$R_1 = \int_0^1 (1 + p_1(f)) df, R_2 = \int_0^1 (1 + p_2(f)) df. \tag{2.22}$$

From the FDMA and the symmetry assumption of the channel, the sum-rate of both users is equivalent to the rate of a single user with a power constraint of $p$. From the water-filling principle, $h^*(p)$ is achieved when the PSD over the whole band is flat. In other words, both users' powers are allocated mutually non-overlapped and collectively filling the whole band uniformly. Accordingly, we have the following lemma.

**Lemma 5.** *The maximum achievable sum-rate with FDMA is*

$$h^*(p) = \log(1+p). \tag{2.23}$$

Define the critical point $p_0 \triangleq 2\left(\frac{1}{2\alpha^2} - \frac{1}{\alpha}\right)$. From Lemma 3, it can be verified that $f^*(p_0) = h^*(p_0)$. As $f^*(p), h^*(p)$ are both increasing and concave, the upper envelope of $f^*(p)$ and $h^*(p)$ is given by

$$r(p) \triangleq \max f^*(p), h^*(p) = \begin{cases} f^*(p), & p \in [0, p_0] \\ h^*(p), & p \in (p_0, \infty] \end{cases}. \tag{2.24}$$

Furthermore, as $0 < \alpha < \frac{1}{2}$,

$$\frac{d}{dp} f^*(p)|_{p=p_0} = \frac{4\alpha^3}{1-\alpha} < \frac{\alpha^2}{(1-\alpha)^2} = \frac{d}{dp} h^*(p)|_{p=p_0}, \tag{2.25}$$

and the upper envelope $r(p)$ is non-concave in $[0, \infty)$.

**Definition 8.** *$r^*(p)$ is defined to be the unique convex hull of $r(p)$:*

$$r^*(p) \triangleq conv_p(r(p)). \tag{2.26}$$

A typical plot of $f^*(p), h^*(p), r^*(p)$ is given in Figure 2.4. Since $f^*(p)$ and $h^*(p)$ are themselves concave, the convex hull of the upper envelope is found by computing their common tangent line. For example, In Figure 2.4, $\alpha$ is chosen to be 0.1. $f^*(p), h^*(p)$ intersect at $p_0 = 80$. The two points of tangency are $p_f = 54.391, p_h = 115.938$.

In order to find the common tangent line of $f^*(p), h^*(p)$, the two points of tangency $p_f, p_h$ are determined by

$$\frac{d}{dp} f^*(p)|_{p=p_f} = \frac{d}{dp} h^*(p)|_{p=p_h} = \frac{h^*(p_h) - f^*(p_f)}{p_h - p_f},$$

which simplifies to finding $p_f$ by solving

$$\frac{p_f(\alpha(1+\alpha)p_f + 4\alpha - 2)}{(\alpha p_f + 2)((1+\alpha)p_f + 2)} = \log\left(\frac{(\alpha p_f + 2)^3}{4((1+\alpha)p_f + 2)}\right) \tag{2.27}$$

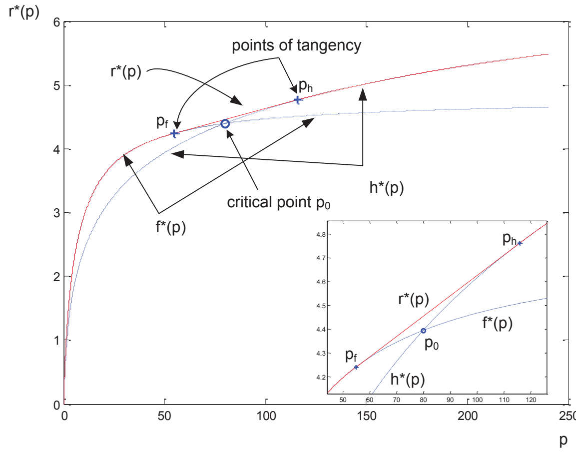



Figure 2.4: The maximum achievable rate as the convex hull of the rates of flat FDMA and flat frequency sharing.

and computing $p_h$ by

$$p_h = \frac{1}{4} p_f(\alpha(1+\alpha)p_f + 4\alpha + 2). \tag{2.28}$$

$p_f, p_h$ can be obtained by solving the closed form equation (2.27) where various numerical methods can be applied. From many numerical examples, we observed that (2.27) always has one valid fix point solution.

Next, we provide the main theorem of this section.

**Theorem 3.**

$$r^o(p) = r^*(p), \forall p \geq 0. \tag{2.29}$$

While the proof of the achievability of $r^*(p)$ is fairly straightforward, the proof of the converse follows from Jensen's inequality, as we recognize that all allocation schemes $\mathbf{p}(f)$ are pointwise either flat frequency sharing or flat FDMA.

*Proof of Theorem 3.*

i) $r^*(p) \leq r^o(p)$ (Achievability of $r^*(p)$).

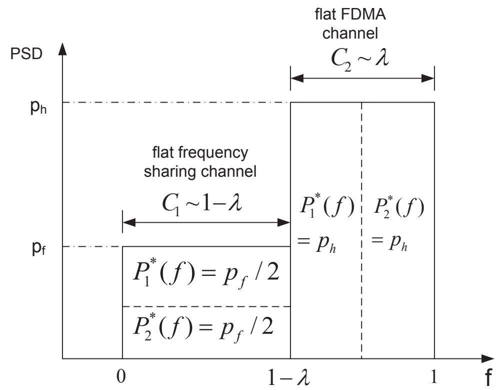



Figure 2.5: The optimal spectrum and power allocation as a mixture of flat FDMA and flat frequency sharing.

The achievability of $r^*(p)$ when $0 \leq p \leq p_f$ or $p \geq p_h$ is immediate.

When $p_f \leq p \leq p_h$,

$$r^*(p) = f^*(p_f) + \lambda(h^*(p_h) - f^*(p_f)), \tag{2.30}$$

where $\lambda = \frac{p-p_f}{p_h-p_f}$, and $r^*(p)$ is achievable by the following scheme as depicted in Figure 2.5. The whole band is split into two orthogonal channels: $C_1$ with bandwidth $1-\lambda$ and $C_2$ with bandwidth $\lambda$. In $C_1$, a flat frequency sharing with a PSD of $\frac{p_f}{2}$ for each user is applied, achieving a sum-rate of $f^*_{C_1} = (1-\lambda)f^*(p_f)$. In $C_2$, a flat FDMA with a PSD of $p_h$ for each user is applied, achieving a sum-rate of $h^*_{C_2} = \lambda h^*(p_h)$. Note that the sum-power constraint is satisfied by such a combination of flat frequency sharing and flat FDMA: $(1-\lambda)p_f + \lambda p_h = p$. Therefore, the sum-rate

$$f^*_{C_1} + h^*_{C_2} = (1-\lambda)f^*(p_f) + \lambda h^*(p_h) = r^*(p) \tag{2.31}$$

can be achieved in the original problem (2.15).

ii) $r^o(p) \leq r^*(p)$ (Converse) For any given $p$, let $\{p_1^o(f), p_2^o(f)\}$ be an optimal scheme that achieves $r^o(p)$. Define the sum-rate density

$$r_p^o(f) \triangleq \log\left(1+\frac{p_1^o(f)}{1+\alpha p_2^o(f)}\right)+\log\left(1+\frac{p_2^o(f)}{1+\alpha p_1^o(f)}\right),$$

and sum-PSD $p^o(f) = p_1^o(f) + p_2^o(f)$. Clearly, $r^o(p) = \int_0^1 r_p^o(f)df$.

From Lemmas 2.3, 4, 5,

$$\text{when } p^o(f) \le 2\left(1-\frac{1}{2\alpha(f)^2}-\frac{1}{\alpha(f)}\right), r_p^o(f) \le f^*(p^o(f)).$$
$$\text{when } p^o(f) > 2\left(1-\frac{1}{2\alpha(f)^2}-\frac{1}{\alpha(f)}\right), r_p^o(f) \le h^*(p^o(f)).$$

Thus,

$$\begin{aligned} &r_p^o(p) \le \max\left(f^*(p^o(f)), h^*(p^o(f))\right) \le r^*(p^o(f)) \\ \Rightarrow &r^o(p) = \int_0^1 r_p^o(f)df \le \int_0^1 r^*(p^o(f))df \le r^*\left(\int_0^1 p^o(f)df\right) \le r^*(p). \end{aligned} \tag{2.32}$$

The second inequality arises from the concavity of $r^*(p)$ and Jensen's inequality, and the last inequality arises from the sum-power constraint and the fact that $r^*(p)$ is an increasing function. □

The mixture of a flat frequency sharing and a flat FDMA shown in Figure 2.5 represents the general form of the optimal spectrum and power allocation achieving $r^*(p)$. The computation of the optimal spectrum management scheme is summarized in Procedure 1. Note that there always exists an optimal spectrum and power allocation with two users each using the same total power of $\frac{p}{2}$. Therefore, the above optimal solution with a sum-power constraint directly leads to the optimal solution with equal individual power constraints:

**Corollary 5.** *The maximum sum-rate with equal individual power constraints*

$$\begin{aligned} &\max_{\mathbf{p}(f)} \int_0^1 r_1(\mathbf{p}(f), f) + r_2(\mathbf{p}(f), f)df \\ &s.t. \int_0^1 p_i(f)df \le \frac{p}{2},\ p_i(f) \ge 0, i = 1, 2, \forall f \in [0, 1] \end{aligned} \tag{2.33}$$

*is* $r^*(p)$.

*Proof.* On the one hand, the equal power constraints imply the sum-power constraint. On the other hand, the optimal value with the sum-power constraint can be achieved with the equal power constraints. □

*Procedure 1: Computing the optimal spectrum management for two-user symmetric flat channels.*

---

Step 1: Solve the two points of tangency $p_f, p_h$ on $r^*(p)$.

a. Solve (2.27) numerically to find $p_f$:

b. Compute $p_h$ by (2.28).

Step 2: Compute the maximum achievable sum-rate $r^*(p)$:

If $p \le p_f,\ r^*(p) = f^*(p)$. Allocate $p_1(f) = p_2(f) = \frac{p}{2}, \forall f$.

If $p \ge p_h,\ r^*(p) = h^*(p)$. Allocate $p_1(f), p_2(f)$ such that

$p_1(f)p_2(f) = 0, p_1(f) + p_2(f) = p, \forall f$

If $p_f < p < p_h, r^*(p) = f^*(p) + \frac{h^*(p_h) - f^*(p_f)}{p_h - p_f}(p - p_f)$.

a. Compute $\lambda = \frac{p - p_f}{p_h - p_f}$.

b. Separate $[0, 1]$ into two disjoint channels:

$C_1$ with bandwidth $1 - \lambda$, $C_2$ with bandwidth $\lambda$.

c. Allocate power as follows (Figure 2.5):

$\forall f \in C_1, p_1(f) = p_2(f) = \frac{p_f}{2}$;

$\forall f \in C_2, p_1(f)p_2(f) = 0, p_1(f) + p_2(f) = p_h$.

---

#### 2.1.3.2 Generalizations to Frequency Selective Channels

We now extend the sum-rate maximization problem to the symmetric frequency selective Gaussian interference channel:

$$\alpha_{12}(f) = \alpha_{21}(f) = \alpha(f), n_1(f) = n_2(f) = n(f), \forall f \in [0, 1]. \tag{2.34}$$

With

$$r_1(\mathbf{p}(f), f) = \log\left(1 + \frac{p_1(f)}{n(f) + p_2(f)\alpha(f)}\right),$$
$$r_2(\mathbf{p}(f), f) = \log\left(1 + \frac{p_2(f)}{n(f) + p_1(f)\alpha(f)}\right),$$

define $r^o$ to be the maximum achievable sum-rate with a sum-power constraint as follows:

**Definition 9.**

$$r^o \triangleq \max_{\mathbf{p}(f)\geq 0} \int_0^1 r_1(\mathbf{p}(f), f) + r_2(\mathbf{p}(f), f) df \tag{2.35}$$
$$s.t. \int_0^1 p_1(f) + p_2(f) df \leq p.$$

Note that the objective function is *separable in* $f$. (The whole problem is, of course, not immediately separable in $f$ because of the total power constraint across the whole band.)

**Remark 4.** *Because for every fixed* $f \in [0, 1]$*,* $r_1(\mathbf{p}(f), f) + r_2(\mathbf{p}(f), f)$ *is non-concave in* $\{p_1(f), p_2(f)\}$*, the above infinite dimensional problem is a non-convex optimization.*

Next, we derive a primal domain convex relaxation of (2.35). We first normalize the PSD and the sum-PSD by $n(f)$:

$$\forall f \in [0, 1], \tilde{p}_1(f) \triangleq \frac{p_1(f)}{n(f)}, \tilde{p}_2(f) \triangleq \frac{p_2(f)}{n(f)}, \tilde{p}(f) \triangleq \tilde{p}_1(f) + \tilde{p}_2(f). \tag{2.36}$$

**Definition 10.** $\forall f \in [0,1]$, *in the same form of* (2.20) *and* (2.23) *with* $\alpha(f)$ *instead of* $\alpha$*:*

$$f^*(p,f) \triangleq 2\log\left(1+\frac{p/2}{1+\alpha(f)p/2}\right), h^*(p,f) \triangleq \log(1+p),$$
$$r^*(p,f) \triangleq conv_p\left(\max(f^*(p,f), h^*(p,f))\right). \tag{2.37}$$

Note that the convex hull operation is done along the *power* dimension for every *fixed* $f$, (not along the frequency dimension.) $\forall f \in [0,1], p_f(f), p_h(f)$, and $r^*(p,f), p \geq 0$ are computed in the same way as in Procedure 1 with $\alpha(f)$ instead of $\alpha$.

Now, in the (separable) objective function of (2.35), at every frequency $f$, we replace the non-concave $r_1(\mathbf{p}(f),f)+r_2(\mathbf{p}(f),f)$ with the concave $r^*(\tilde{p}(f),f)$ (concave in the first variable $\tilde{p}(f)$), and define $r^*$ to be the corresponding maximum achievable value as follows:

**Definition 11.**

$$r^* \triangleq \max_{\tilde{p}(f)\geq 0} \int_0^1 r^*(\tilde{p}(f),f)df \tag{2.38}$$
$$s.t. \int_0^1 \tilde{p}(f)n(f)df \leq p, \forall f \in [0,1].$$

**Remark 5.** *For every fixed* $f$*,* $r^*(\tilde{p}(f),f)$ *is concave in* $\tilde{p}(f)$*. The constraint is linear in* $\tilde{p}(f), \forall f$*. Thus, the above infinite dimensional problem is a convex optimization.*

Now, we have the following theorem:

**Theorem 4.**

$$r^o = r^*. \tag{2.39}$$

The proof of the converse is similar to that in Theorem 3. For the proof of the achievability of $r^*$, as the channel is frequency selective, we need to introduce a piecewise flat $\epsilon$-approximation, and the remaining proof exactly follows that in Theorem 3.

*Proof of Theorem 4.*

i) $r^o \leq r^*$ (Converse).

It is sufficient to prove the inequality between the integrands in (2.35) and (2.38). As in the proof of Theorem 3, from Lemmas 2.3, 4, 5,

$$r_1(\mathbf{p}(f), f) + r_2(\mathbf{p}(f), f) = \log\left(1 + \frac{\tilde{p}_1(f)}{\tilde{p}_2(f)\alpha(f)}\right) + \log\left(1 + \frac{\tilde{p}_2(f)}{\tilde{p}_1(f)\alpha(f)}\right)$$
$$\leq \max\left(f^*(\tilde{p}(f), f), h^*(\tilde{p}(f), f)\right) \leq r^*(\tilde{p}(f), f).$$

ii) $r^* \leq r^o$ (Achievability). Let sum-PSD $\tilde{p}^*(f)$ be an optimal solution of (2.38) such that $\int_0^1 r^*(\tilde{p}^*(f), f)df = r^*$. Then, $\forall \epsilon > 0$:

By Lemma 1, based on $\tilde{p}^*(f)$ and $\{\alpha_{ji}(f)\}$, take a piecewise flat $\epsilon$-approximation $\bar{p}^*(f)$ and $\{\bar{\alpha}_{ji}(f)\}$, s.t.

$$\left| \int_0^1 \bar{r}^*(\bar{p}^*(f), f)df - r^* \right| < \epsilon, \tag{2.40}$$

where $\bar{p}^*(f)$ is a piecewise flat sum-PSD, and $\bar{r}^*(\bar{p}^*(f), f)$ is computed with $\{\bar{\alpha}_{ji}(f)\}$. (Note that, since the noise PSD is already normalized to 1 as in (2.36), no further piecewise flat approximation of the noise is needed.)

Based on the piecewise flat $\epsilon$-approximation, in *every flat sub-channel with a flat* $\bar{p}^*(f)$, as in the proof of Theorem 3, $\bar{r}^*(\bar{p}^*(f), f)$ can be achieved by further dividing this flat sub-channel into two sub-bands, applying a flat frequency sharing and a flat FDMA respectively (cf. Figure 2.5). Removing the normalization by multiplying by $n(f)$, denote the resulting allocation scheme by

$\mathbf{p}^o(f) = [p_1^o(f), p_1^o(f)]^T$, achieving the same sum-rate

$$\int_0^1 \bar{r}_1(\mathbf{p}^o(f), f) + \bar{r}_2(\mathbf{p}^o(f), f)df = \int_0^1 \bar{r}^*(\bar{p}^*(f), f),$$

where $\bar{r}_1(\mathbf{p}^o(f), f)$ and $\bar{r}_1(\mathbf{p}^o(f), f)$ are computed with the piecewise flat approximate channel responses $\{\bar{\alpha}_{ji}(f)\}$. Then,

$$\begin{aligned} r^o \geq \int_0^1 r_1(\mathbf{p}^o(f), f) + r_2(\mathbf{p}^o(f), f)df \geq \int_0^1 \bar{r}_1(\mathbf{p}^o(f), f) + \bar{r}_2(\mathbf{p}^o(f), f)df \\ = \int_0^1 \bar{r}^*(\bar{p}^*(f), f) > r^* - \epsilon, \end{aligned} \quad (2.41)$$

where the first inequality occurs because $\mathbf{p}^o(f)$ is a *feasible* solution of (2.35); the second inequality arises because (by $\mathcal{P}2$ from Lemma 1) $\bar{\alpha}_{ji}(f) \geq \alpha_{ji}(f), \forall i, j, \forall f \in [0, 1]$, i.e. the $\epsilon$-approximation worsens the channel responses, resulting in lower rates.

Finally, let $\epsilon \to 0$. □

Therefore, although the integrand in (2.38) is a direct convex relaxation of that in (2.35), the optimal objective value of the problem does not change, and the original non-convex optimization (2.35) is equivalently transformed to the convex optimization (2.38). Finally, for the same reasons as in section 2.1.3.1, the optimal solution with equal individual power constraints is the same as that with a corresponding sum-power constraint.

**Remark 6.** *Throughout section 2.1.3, we have worked with a sum-power constraint to gain brevity in derivations of the results for the fully symmetric cases. One may also derive the results directly with equal individual power constraints. In Section 2.1.4, as we consider potentially asymmetric channels, we will directly work with individual power constraints.*

### 2.1.4 Optimal Spectrum Management in the General Cases

In Section 2.1.3, we solved the sum-rate maximization problem in two-user symmetric frequency selective channels with equal power (or sum-power) constraints. In this section, we make the following generalizations:

- Two-user $\rightarrow$ $K$-user,
- Equal power constraints $\rightarrow$ arbitrary individual power constraints,
- Symmetric channel $\rightarrow$ arbitrary asymmetric channel,
- Sum-rate $\rightarrow$ arbitrary weighted sum-rate.

The general optimization problem is thus the following:

$$\max_{\mathbf{p}(f)\geq 0} \int_0^1 \sum_{i=1}^{K} w_i r_i(\mathbf{p}(f), f) df \tag{2.42}$$

$$s.t. \int_0^1 \mathbf{p}(f) df \leq \mathbf{p},\ r_i(\mathbf{p}(f), f) = \log\left(1 + \frac{p_i(f)}{n_i(f) + \sum_{j\neq i} p_j(f)\alpha_{ji}(f)}\right).$$

Next, we analyze this general problem in parallel with the analysis in Section 2.5, and show that the same basic ideas therein generalize here.

#### 2.1.4.1 Optimal Solutions for Flat Channels

Consider a $K$-user (potentially asymmetric) flat channel:

$$\alpha_{ji}(f) = \alpha_{ji}, n_i(f) = n_i, \forall f \in [0, 1], \forall i, j.$$

First, consider the weighted sum-rate achieved with *flat* power allocations $\mathbf{p}(f) = \mathbf{P}, \forall f \in [0, 1]$, defined as

$$R(\mathbf{P}) \triangleq \sum_{i=1}^{K} w_i \log\left(1 + \frac{P_i}{n_i + \sum_{j\neq i} P_j\alpha_{ji}}\right). \tag{2.43}$$

Denote its $K$ dimensional convex hull function by

$$R^*(\mathbf{P}) \triangleq conv_{\mathbf{P}}(R(\mathbf{P})). \tag{2.44}$$

The original problem (2.42) in flat channels can be rewritten as

**Definition 12.**

$$\begin{aligned} R^o(\mathbf{p}) \triangleq \max_{\mathbf{p}(f)\geq 0} \int_0^1 R(\mathbf{p}(f))df \\ s.t. \int_0^1 \mathbf{p}(f)df \leq \mathbf{p}. \end{aligned} \tag{2.45}$$

Now, we have the following theorem,

**Theorem 5.**

$$R^o(\mathbf{p}) = R^*(\mathbf{p}),$$

*and the optimal spectrum and power allocation* $\mathbf{p}^o(f)$ *consists of* $K+1$ *sub-bands, with* $\mathbf{p}^o(f)$ *flat in each of the sub-bands.*

*Proof.* The proof is in parallel with that of Theorem 3.

1. $R^*(\mathbf{p}) \leq R^o(\mathbf{p})$ (Achievability).

As $R^*(\mathbf{P}) \triangleq conv_{\mathbf{P}}(R(\mathbf{P}))$, by Carathéodory's theorem,

$$\exists c^{(k)} \geq 0, k=1,\ldots,K+1, \sum_{k=1}^{K+1} c^{(k)} = 1, \sum_{k=1}^{K+1} c^{(k)}\mathbf{p}^{(k)} = \mathbf{p}, \ s.t.$$
$$R^*(\mathbf{p}) = \sum_{k=1}^{K+1} c^{(k)} R(\mathbf{p}^{(k)}).$$

Accordingly, we can divide the band $[0,1]$ into $K+1$ sub-bands, each with a bandwidth of $c^{(k)}$ and uses the flat power levels of $\mathbf{p}^{(k)} = [p_1^{(k)},\ldots,p_K^{(k)}]^T$ for the $K$ users.

2. $R^o(\mathbf{p}) \leq R^*(\mathbf{p})$ (Converse).

For any feasible allocation scheme $\mathbf{p}(f), f \in [0,1]$.

$$\int_0^1 R(\mathbf{p}(f)) \leq \int_0^1 R^*(\mathbf{p}(f))df \leq R^*\left(\int_0^1 \mathbf{p}(f)df\right) \leq R^*(\mathbf{p}), \tag{2.46}$$

where the first inequality is from (2.44), the second inequality arises from Jensen's inequality, and the third inequality arises from the fact that $R^*(\mathbf{P})$ is increasing in $\mathbf{P}$. □

**Remark 7.** *In the literature, it was first shown that allocation schemes consisting of* $2K$ *sub-bands of flat allocations are sufficient to achieve any Pareto optimality [EPT07], and this sufficient number of sub-bands was later refined to* $K+1$ *[SZB08]. From Theorem 5, the sufficiency of* $K+1$ *sub-bands is also immediately implied by the fact that the optimal value and solution are obtained by nothing more than computing the convex hull* (2.44) *of a non-concave function* (2.43).

#### 2.1.4.2 Generalizations to Frequency Selective Channels

In frequency selective channels, define the weighted sum-rate density function as

$$R(\mathbf{p}, f) \triangleq \sum_{i=1}^{K} w_i r_i(\mathbf{P}, f), \tag{2.47}$$

where $r_i(\mathbf{P}, f)$ is defined in (2.4). Problem (2.42) can then be rewritten as:

**Definition 13.**

$$R^o \triangleq \max_{\mathbf{p}(f)\geq 0} \int_0^1 R(\mathbf{p}(f), f)df \tag{2.48}$$
$$s.t. \int_0^1 \mathbf{p}(f)df \leq \mathbf{p}.$$

Clearly, for every fixed $f \in [0,1]$, $R(\mathbf{p}(f), f)$ is non-concave in $\mathbf{p}(f)$, and (2.48) is an infinite dimensional non-convex optimization.

At every frequency $f \in [0,1]$, define

$$R^*(\mathbf{P}, f) \triangleq conv_{\mathbf{P}} R(\mathbf{P}, f), \tag{2.49}$$

i.e., the *convex hull of* $R^*(\mathbf{P}, f)$ *along the* $K$ *dimensions of power* $\mathbf{P}$. Note that the convex hull operation is *not* taken along the frequency dimension $f$. ($R^*(\mathbf{P}, f)$ is concave in $\mathbf{P}$ for every fixed $f$, but not necessarily jointly concave in $\mathbf{P}, f$.)

Next, we derive the following primal domain convex relaxation of (2.48): *At every frequency* $f$, we replace the non-concave $R(\mathbf{p}(f), f)$ with the concave $R^*(\mathbf{p}(f), f)$ (concave in the first variable $\mathbf{p}(f)$), and define $R^*$ to be the corresponding maximum achievable value as follows:

**Definition 14.**

$$R^* \triangleq \max_{\mathbf{p}(f)\geq 0} \int_0^1 R^*(\mathbf{p}(f), f) df \tag{2.50}$$
$$s.t. \int_0^1 \mathbf{p}(f) df \leq \mathbf{p}.$$

Clearly, (2.50) is an infinite dimensional convex optimization, because

- $\forall f \in [0,1]$, the integrand is a concave function of the variables $\mathbf{p}(f)$;
- The constraint is linear in $\{\mathbf{p}(f), f \in [0,1]\}$.

Finally, we have the following theorem

**Theorem 6.**

$$R^o = R^*. \tag{2.51}$$

Therefore, the optimal value for the non-convex optimization (2.48) equals that of its convex relaxation (2.50).

### 2.1.5 Discussion

#### 2.1.5.1 On the Complexity of Solving the General Continuous Frequency Problem

To avoid an uncountably infinite number of dimensions, consider any channel response in the form of a piecewise flat functions of frequency. Denote the corresponding flat sub-channels by $I_1, I_2, \ldots, I_M$, each with bandwidth $b_m$. (Note that the channels $\{I_m\}$ are viewed as given, and their bandwidths $\{b_m\}$ are not variables to optimize.) We now recall that the key difference between the continuous frequency model and the discrete frequency model lies in the assumption of power allocations:

- Continuous frequency: $\mathbf{p}(f)$ is piecewise bounded continuous.
- Discrete frequency: $\mathbf{p}(f)$ must be flat in every flat sub-channel $I_m$.

For example, consider a single flat band. It makes a fundamental difference whether we allow a user to subdivide this flat band and use different PSD in different sub-bands. If so, the problem model is still continuous frequency. Otherwise, the problem model is discrete frequency.

It has been proven that finding the optimal solution with the discrete frequency model is NP hard [LZ08]. This is not inconsistent with the convex formulations for the continuous frequency model in this section, because the assumptions made on power allocation are different. Next, we discuss the complexity of solving the continuous frequency optimal spectrum management in piecewise flat channels. From Theorem 6, it is sufficient to solve the convex optimization (2.50), which consists of two general steps:

- Step 1: Compute the convex hull function $R^*(\mathbf{P}, f)$ at every frequency $f \in [0, 1]$.
- Step 2: Optimize $\mathbf{p}(f)$ with the objective $\int_0^1 R^*(\mathbf{p}(f), f)df$.

In Step 1, given the channel parameters for each flat sub-channel $I_m, m = 1, \ldots, M$, a convex hull function $R_m^*(\mathbf{P}) \left(\triangleq R^*(\mathbf{P}, f), f \in I_m\right), \forall \mathbf{P} \geq 0$ is computed. Numerically and approximately computing a convex hull is itself a broad and important topic (see e.g. [Cha96]). This computational issue is not further discussed here, and is left for future work. We note that it remains unclear whether this computational issue is easier to deal with than the NP hardness in the discrete frequency model.

In Step 2, given the convex hull functions for all the flat sub-channels, as the number of sub-channels $M$ is finite, problem (2.50) becomes finite dimensional, with an increasing *concave* utility function $R_m^*(\mathbf{p}(f))$ in each sub-channel $I_m$. Now, because each $I_m$ is a *flat* channel and $R_m^*(\mathbf{P})$ is increasing concave, by Jensen's inequality, the optimal solution must satisfy that $\mathbf{p}(f)$ is flat in each sub-channel $I_m$, i.e., $\forall m = 1, \ldots, M, \exists \mathbf{p}(m) \geq 0, s.t.\ \mathbf{p}(f) = \mathbf{p}(m), \forall f \in I_m$. Problem (2.50) then becomes

$$\max_{\mathbf{p}(m)} \sum_{m=1}^{M} b_m R^*(\mathbf{p}(m)) \tag{2.52}$$
$$s.t.\ \sum_{m=1}^{M} b_m \mathbf{p}(m) \leq \mathbf{p}, \mathbf{p}(m) \geq 0, \forall m = 1, 2, \ldots, M.$$

(Recall that $b_m$ $(m = 1, \ldots, M)$ is the bandwidth of sub-channel $I_m$, and is not an optimization variable). Problem (2.52) is a convex optimization that has efficient polynomial time algorithms to solve the globally optimal solution. (For example, a dual decomposition algorithm works, see e.g. [Chi05b] among many others.)

In summary, the critical complexity in solving the general problem (2.48) based on Theorem 6 lies in computing approximate convex hull functions. While computing convex hull functions given channel parameters may be computationally costly, this two-step method does have the following advantage:

**Corollary 6.** *Once the channel parameters are given, the $M$ convex hull functions $R_m^*(\mathbf{P}), m = 1, \ldots, M$ are computed* for one time*. Then, no matter how the power constraints may vary due to problem needs, the additional complexity cost of solving the optimal solution (Step 2) is only polynomial time.*

This separation of the complexity in dealing with channel responses and power constraints does not appear in the discrete frequency model, due to the fundamental difference between the power allocation assumption of the continuous frequency model and that of the discrete frequency model. For the discrete frequency model, the constraint that a user must use a flat PSD within every (flat) sub-channel leads to the well known NP hardness. In contrast, for the continuous frequency model, the main complexity is from computing convex hull functions.

#### 2.1.5.2 On the Zero Duality Gap

It has been proved that the continuous frequency non-convex optimization (2.48) has an exact zero duality gap [LZ08]. It is pointed out that the zero duality gap comes from a time sharing condition [YL06]. It is also proved using the nonatomic property of the Lebesgue measure [LZ08]. We show that this is also immediately implied by Theorem 6.

**Definition 15.** *For problem* (2.48)*, its Lagrange dual is defined as*

$$L(\mathbf{p}(f), \lambda) \triangleq \int_0^1 R(\mathbf{P}(f), f)df - \lambda^T \left( \int_0^1 \mathbf{p}(f)df - \mathbf{p} \right).$$

*Its dual objective and dual optimal value are defined as*

$$g(\lambda) \triangleq \sup_{\mathbf{p}(f)\geq 0} L(\mathbf{p}(f), \lambda), \text{ and } D^o \triangleq \min_{\lambda \geq 0} g(\lambda).$$

*Similarly, for problem* (2.50)*, its Lagrange dual, dual objective, and dual optimal value are defined as*

$$\hat{L}(\mathbf{p}(f), \lambda) \triangleq \int_0^1 R^*(\mathbf{P}(f), f)df - \lambda^T \left( \int_0^1 \mathbf{p}(f)df - \mathbf{p} \right).$$

$$\hat{g}(\lambda) \triangleq \sup_{\mathbf{p}(f)\geq 0} \hat{L}(\mathbf{p}(f), \lambda), \text{ and } D^* \triangleq \min_{\lambda \geq 0} \hat{g}(\lambda).$$

**Corollary 7.** *The non-convex optimization* (2.48) *has a zero duality gap.*

*Proof.* Since $R^*(\mathbf{P}, f) \geq R(\mathbf{P}, f), \forall \mathbf{P}, \forall f$, we have

$$\begin{aligned} \hat{L}(\mathbf{p}(f), \lambda) \geq L(\mathbf{p}(f), \lambda) &\Rightarrow \hat{g}(\lambda) \geq g(\lambda), \lambda \geq 0 \\ &\Rightarrow D^* \geq D^o. \end{aligned}$$

Therefore,

$$R^* = D^* \geq D^o \geq R^o = R^* \Rightarrow D^o = R^o,$$

where the first equality occurs because problem (2.50) is a convex optimization and has *strong* duality [BV04]; the second inequality is from the weak duality of the non-convex optimization (2.48); the key step is the second equality $R^o = R^*$ from Theorem 6. □

Furthermore, it has been shown that, under mild technical conditions, the non-convex optimization for the discrete frequency model has an *asymptotically* zero duality gap as the number of sub-channels goes to infinity [YL06]. The result is rigorously generalized to include Lebesgue integrable PSDs in [LZ08]. Indeed, for a piecewise bounded continuous frequency channel, as it is divided into more and finer/flatter sub-channels, the difference between the power allocation assumption of discrete frequency model and that of continuous frequency model vanishes. The intuition is that we can bundle a large number of similar flat sub-channels, treat them as one combined flat channel, compute the continuous frequency power allocation, and accordingly distribute the power within these roughly identical sub-channels (as a discrete approximation of the continuous allocation.)

## 2.2 Discrete Frequency Spectrum Management

In this section, we turn our attention to the discrete frequency model with $M$ parallel channels each having a unit bandwidth. Consider a weighted sum-rate maximization problem

$$\max_{\substack{P_i^m \geq 0, 1 \leq i \leq K \\ 1 \leq m \leq M}} \sum_{i=1}^{K} \sum_{m=1}^{M} w_i \log(1 + \mathsf{SINR}_i^m) \tag{2.53}$$
$$s.t. \;\; \mathsf{SINR}_i^m = \frac{g_{ii}^m P_i^m}{\sum_{j \neq i} g_{ji}^m P_j^m + N_i^m}, \;\; \sum_{m=1}^{M} P_i^m \leq p_i, \forall i = 1, \ldots, K,$$

where $P_i^m, \mathsf{SINR}_i^m$ are the transmit power and the signal to interference plus noise ratio of user $i$ in channel $m$, $g_{ij}^m = |h_{ij}|^2$ is the channel gain from transmitter $i$ to receiver $j$ in channel $m$, $N_i^m$ is the noise power at receiver $i$ in channel $m$. It has been shown that this problem is NP complete in both the number of users $K$ and the number of channels $M$ [LZ08].

In the literature, an approximation of the objective that yields a convex optimization formulation is the following [Chi05a]:

$$\log(1 + \mathsf{SINR}_i^m) \approx \log(\mathsf{SINR}_i^m) = \log\left(\frac{g_{ii}^m P_i^m}{\sum_{j \neq i} g_{ji}^m P_j^m + N_i^m}\right), \tag{2.54}$$

The convexified maximization problem is thus

$$\max_{\substack{P_i^m \geq 0, 1\leq i \leq K \\ 1\leq m \leq M}} \sum_{i=1}^{K} \sum_{m=1}^{M} w_i \log(\mathsf{SINR}_i^m) \tag{2.55}$$
$$s.t. \sum_{m=1}^{M} P_i^m \leq p_i, \forall i = 1, \ldots, K.$$

The convexity can be seen by a change of variables: with $\tilde{P}_i^m \triangleq \log(P_i^m)$,

$$\log\left(\frac{g_{ii}^m P_i^m}{\sum_{j\neq i} g_{ji}^m P_j^m + N_i^m}\right) = \log(g_{ii}^m e^{\tilde{P}_i^m}) - \log(\sum_{j\neq i} g_{ji}^m e^{\tilde{P}_j^m} + N_i^m). \tag{2.56}$$

Note that the first term is linear in $\tilde{P}_i^m$, and the second term is concave in $\tilde{P}_i^m$ because the log of a sum of exponentials is a convex function. The power constraints are transformed to the following convex constraints

$$\sum_{m=1}^{M} e^{\tilde{P}_i^m} \leq p_i, \forall i = 1, \ldots, K. \tag{2.57}$$

The convex approximation (2.54) is a good approximation if $\mathsf{SINR}_i^m \gg 0$. To understand its implications in approaching the optimum of the original problem (2.53) by solving (2.55), we examine a sum-rate maximization problem, i.e., $w_1 = \ldots = w_K = 1$:

First, if the optimal solution of (2.53) satisfies that *all* users have *high* SINR in *all* channels, we can expect getting a close approximation of it by solving (2.55).

However, if the optimal solution of (2.53) does require some users having *low* SINR in some channels, solving (2.55) will not be able to find such a solution. This is because maximization with log(SINR) imposes extremely high penalty on low SINR. In particular, if the optimal solution of (2.53) requires user $i$ not to use channel $m$ at all, i.e., $P_i^m = 0$, it will cause $\log(\mathsf{SINR}_i^m) = -\infty$. As a result, such a solution will be avoided by solving the convex approximation (2.55).

Moreover, because $\log(1+\mathsf{SINR}) - \log(\mathsf{SINR}) \to \infty$ as $\mathsf{SINR} \to 0$ $(-\infty dB)$, for a fixed number of users and channels, the gap between the optimal value of (2.53) and that of (2.55) can be arbitrarily large.

To address the above issues of using $\log(\mathsf{SINR})$ in the objective, we consider the use of $\log(\mathsf{SINR})^+ \triangleq \max(0, \log(\mathsf{SINR}))$, and solving the following problem

$$\max_{\substack{P_i^m \geq 0, 1 \leq i \leq K \\ 1 \leq m \leq M}} \sum_{i=1}^{K} \sum_{m=1}^{M} w_i \log(\mathsf{SINR}_i^m)^+ \tag{2.58}$$
$$s.t. \sum_{m=1}^{M} P_i^m \leq p_i, \forall i = 1, \ldots, K.$$

The inclusion of the zero floor simply disallows any negative rate from computing $\log(\mathsf{SINR})$. It is immediate to verify that

$$\max_{\mathsf{SINR} \in \mathbb{R}} |\log(1+\mathsf{SINR}) - \log(\mathsf{SINR})^+| = \log 2 = 1 \ bit, \tag{2.59}$$

and the maximum discrepancy is reached at $\mathsf{SINR} = 1$ $(0dB)$. We then have the following lemma:

**Lemma 6.** *The difference between the optimal value of* (2.53) *and that of* (2.58) *is no greater than $KM$ bits.*

In other words, the per user per channel rate difference between solving (2.53) and (2.58) is at most one bit.

Although the use of $\log(\mathsf{SINR})^+$ leads to this bounded gap property, (2.58) is *no longer a convex optimization.* In what follows, we analyze the computational complexity of solving (2.58), and show that the main complexity lies in the problem of *channel allocation.*

### 2.2.1 Channel Assignment and Power Allocation: a Vertical Decomposition

We provide an alternative form of the bounded gap approximation (2.58) by characterizing the channel assignment decision explicitly:

**Definition 16.** *Define a $K \times M$ channel assignment matrix* $\mathbf{A}$*:*

$$A_{im} = \begin{cases} 1, & \textit{if } \log(\mathsf{SINR}_i^m) \geq 0 \\ 0, & \textit{otherwise.} \end{cases}, i = 1, \ldots, K, m = 1, \ldots, M. \tag{2.60}$$

In other words, if user $i$ has a *positive dB* SINR in channel $m$, $A_{im} = 1$, and we say that *channel $m$ is assigned to user $i$*. One channel can be assigned to multiple users, and one user can have multiple channels assigned to it. Clearly,

$$\log(\mathsf{SINR}_i^m)^+ = A_{im} \log(\mathsf{SINR}). \tag{2.61}$$

Define $\mathbf{A}^*$ to be the *optimal channel assignment matrix* derived from the optimal solution of (2.58). We then have the following theorem:

**Theorem 7.** *Given the optimal channel assignment matrix* $\mathbf{A}^*$*,*

$$\begin{aligned} \max_{\substack{P_i^m \geq 0, 1 \leq i \leq K \\ 1 \leq m \leq M}} & \sum_{i=1}^{K} \sum_{m=1}^{M} w_i A_{im}^* \log(\mathsf{SINR}_i^m) \\ s.t. & \sum_{m=1}^{M} P_i^m \leq p_i, \forall i = 1, \ldots, K \end{aligned} \tag{2.62}$$

*has the same optimal value as* (2.58).

*Proof.* (2.62) $\leq$ (2.58) is immediate because $A_{im}^* \log(\mathsf{SINR}_i^m) \leq \log(\mathsf{SINR}_i^m)^+$ always.

In order to show (2.58) $\leq$ (2.62), note that (2.58) is achieved by its *optimal* solution with which $\mathbf{A}^*$ is derived; by applying the same solution in (2.62) where $\mathbf{A}^*$ is applied, the same optimal value is achieved. □

Note that, with $\mathbf{A}^*$ given, (2.62) is a *convex optimization* for the same reason as in (2.55), and solving this convex optimization gives the optimal solution of (2.58). We thus have the following conclusion:

**Corollary 8** (Decomposition of Channel Assignment (CA) and Power Allocation (PA))**.** *Solving the bounded gap approximation* (2.58) *of the optimal spectrum management problem* (2.53) *can be decomposed into two steps:*

1. *Channel Assignment (a combinatorial problem): finding the optimal channel assignment matrix* $\mathbf{A} = \mathbf{A}^*$*.*

2. *Power Allocation (a convex optimization): given a channel assignment matrix* $\mathbf{A}$*, solving the convex approximation* (2.62) *to determine the transmit powers* $\{P_i^m\}$*.*

As there are in total $2^{KM}$ channel assignment matrices, a traversal of all of them (each followed by a convex optimization of power allocation) has a prohibitively high computational complexity. However, as we show next, with a dual decomposition method, the above CA-PA decomposition is used *in each channel independently*, and the overall computational complexity is reduced from $2^{KM}$ to $M2^K$.

### 2.2.2 Complexity Reduction via Dual Decomposition

We consider the Lagrange dual problem [BV04] of (2.58):

$$\min_{\lambda \geq 0} \sum_{m=1}^{M} g_m(\lambda) + \sum_{i=1}^{K} \lambda_i p_i \tag{2.63}$$

where $\lambda = [\lambda_1, \ldots, \lambda_K]^T$, and $\forall m = 1, \ \ldots, M$,

$$g_m(\lambda) = \sup_{\substack{P_i^m \geq 0, \\ i=1,\ldots,K}} \sum_{i=1}^{K} \left( w_i \log(\mathsf{SINR}_i^m)^+ - \lambda_i P_i^m \right). \tag{2.64}$$

We make the following observations:

- The dual master problem (2.63) is a convex optimization [BV04].
- In each outer iteration of updating $\lambda$ while solving the dual master problem, we need to solve $M$ sub-problems (2.64) which are still *non-convex* optimizations because of the use of $\log(\mathsf{SINR})^+$.

Note that, however, the objective function of the sub-problem (2.64) is just a *single-channel* case (with some additional linear terms) of the objective function (2.58). Therefore, Theorem 7 can be applied to each one of the sub-problems. Similarly to Definition 16, define $\mathbf{A}_m^*$ to be the optimal $K \times 1$ $(0,1)$ *channel assignment vector* of the $m^{th}$ channel: based on the optimal solution of the $m^{th}$ sub-problem (2.64), $\mathbf{A}_m^*(i) = 1$ iff user $i$ has a positive dB $\mathsf{SINR}$ in channel $m$. We then have the following corollary of Theorem 7:

**Corollary 9.** *Given the optimal channel assignment vector* $\mathbf{A}_m^*$*,*

$$\max_{P_i^m \geq 0, i=1,\dots,K} \sum_{i=1}^{K} \left( w_i A_m^*(i) \log(\mathsf{SINR}_i^m) - \lambda_i P_i^m \right) \tag{2.65}$$

*has the same optimal value as that of the sub-problem* (2.64)*.*

**Remark 8** (Decomposition in Channel)**.** *While solving the dual sub-problems* (2.64) (2.65)*, each sub-problem corresponds to one channel. The optimal solution in each channel does not depend on solutions of other channels*[2]*. The channels are only coupled through* $\lambda$ *in the dual master problem.*

For any channel $m = 1, 2, \dots, M$, the total number of possible $A_m^*$ vectors is $2^K$, which upper bounds the computational complexity of solving the $m^{th}$ sub-problem by a traversal of all these vectors. Due to the dual decomposition in

[2] *This desirable property has been observed and exploited in the literature for spectrum management (see [CYM06] among others.)*

channel, there are $M$ sub-problems *independently* solved in each outer iteration of solving the dual master problem. Thus, we arrive at the following result in complexity reduction.

**Corollary 10.** *As a function of $M, K$, the computational complexity of solving the Lagrange dual problem* (2.63) *grows no greater than $M2^K$.*

Now, the remaining issue is that solving the dual optimal solution does not in general give the primal optimal solution. This is because the bounded gap approximation (2.58) is *non-convex*, and does not guarantee strong duality. However, for the same reasons as in the previous analysis of the continuous frequency model ([LZ08, YL06], and also Section 2.1,) as the number of sub-channels goes to infinity, the duality gap of this problem (2.58) goes to zero. In other words, the dual decomposition method is asymptotically optimal.

### 2.2.3 Discussion

Combining Lemma 6, Corollary 10 and the asymptotic zero duality gap, we conclude that the computational complexity of solving the optimal discrete spectrum management (2.53) *to within an asymptotically constant gap* grows no greater than $M2^K$.

We have seen that the optimal spectrum management problem can be vertically decomposed into two steps (Corollary 8): i) finding the optimal channel assignment, and ii) given a channel assignment, finding the optimal power allocation. We have seen that step i) (CA) leads to the non-polynomial computational complexity of $2^K$, whereas step ii) (PA) can be done in polynomial time. Therefore, finding a good channel allocation is *the essential task* in approaching the optimal spectrum management. As the combinatorial problem of CA is in general NP, designing good low complexity CA heuristics is crucial for practical applications in large-scale networks.

## 2.3 Summary

This chapter considers weighted sum-rate maximization in interference networks with interference treated as noise.

For continuous frequency spectrum management, we show that for any two (among $K$) users, as long as the two normalized cross channel gains between them are both larger than or equal to $\frac{1}{2}$, an FDMA allocation between these two users benefits every one of the $K$ users. Therefore, under this pairwise condition, all Pareto optimal points of the $K$-user rate region can be achieved with this pair of users using orthogonal channels. The pairwise nature of the condition allows a completely distributed decision on whether any two users should use orthogonal channels, without loss of any Pareto optimality. Next, we show that the non-convex weighted sum-rate maximization in $K$-user asymmetric frequency selective channels can be equivalently transformed in the primal domain to a convex optimization, and the main computational complexity lies in computing convex hull functions based on channel parameters.

For discrete frequency spectrum management which is NP complete in both the number of users $K$ and the number of channels $M$, we show that an approximate formulation achieving a bounded gap to the original optimum leads to a vertical decomposition of the problem into channel allocation and power allocation. Given the optimal channel allocation, the optimal power allocation can be solved by a convex optimization. This decomposition indicates that the NP hardness of the problem lies in the combinatorial optimization of channel allocation. Via Lagrange dual decomposition, the problem scale of the optimization of channel allocation is further reduced to just a single channel, and the overall computational complexity of solving the approximate formulation is $O(M2^K)$. As it is channel allocation that carries the NP hardness of the problem, designing good channel allocation heuristics is the fundamental task in approaching the optimal spectrum management solution.

# CHAPTER 3

# Channel Allocation in Wireless Cellular Interference Networks

In this chapter, we consider the problem of *channel allocation (CA)* in discrete frequency interference networks. As pointed out in the last chapter, finding good channel allocation schemes is the essential task in approaching the optimal joint power and channel allocation solution. For generic interference networks, however, finding the *globally optimal* channel allocation is a combinatorial problem that consumes a worst-case computational complexity of $O(2^K)$ (where $K$ is the total number of users in the network).

Instead of considering arbitrary interference networks, we focus on channel allocation in *wireless cellular networks*, and design low complexity algorithms for large-scale networks. As a simple hierarchical model, the cellular model successfully exploits wireless propagation losses in coordinating users across large areas, and has been widely applied in wireless communication networks.

We show that, in the simplest case of *uplink channel allocation in one dimensional networks*, the globally optimal channel allocation can be achieved with a complexity of $O(K_{cell} M \log M)$ by a rippling of local *signal scale interference alignment*, (where $K_{cell}$ is the number of cells and $M$ is the number of channels.) Extending it to downlink CA or to two or more dimensional networks, however, signal scale interference alignment does not provide polynomial time algorithms that achieves the globally optimal solution. In spite of this, we show that a bound on the globally optimal value can be derived from assuming "perfect" signal scale interference alignment. From simulation results, the bound is shown to be reasonably tight.

To develop near globally optimal CA algorithms for general cellular networks, we exploit an important nature of wireless interference: its *locality* due to propagations losses. We establish an algorithmic framework that fully respects the effect of inter-cell interference, and yet "horizontally" decomposes the optimization of CA to local *assignment problems*, each solvable with a complexity of $O(M^3)$. The respect of interference ensures that the algorithm is "context-aware" (i.e., not over-simplifying the network physical layer), achieving very close to globally optimal performance. The horizontal decomposition makes the algorithm distributed. Hence, it can be applied to arbitrarily large networks with low complexity.

The rest of the chapter is organized as follows: In Section 3.1, we establish the system model. In Section 3.2, we develop a low complexity CA algorithm that achieves the globally optimal uplink channel allocation in one dimensional networks. In Section 3.3, we develop a low complexity decomposition framework based on local assignment problems for the *downlink* channel allocation in one dimensional networks. In Section 3.4, we generalize the decomposition framework to two or more dimensional cellular networks.

## 3.1 System Model

We make the following assumptions throughout the chapter:

1. Interference is treated as noise.

2. There are $M$ parallel channels with unit bandwidth.

3. Users within the same cell do not reuse the same channel.

4. A frequency reuse factor of *one* is applied among all the cells.

5. In every cell, all the $M$ channels are utilized, i.e., there is no vacant channel in any cell.

6. Every user occupies exactly one channel, which implies that the number of users in each cell equals $M$, since we assume no channel is vacant.

7. Each user transmits *at its own fixed* power level.

**Remark 9.**

- *From assumption 3, there is no intra-cell interference, and all co-channel interference are inter-cell interference.*

- *By assumption 5, we essentially consider a crowded network environment.*

- *For a user occupying multiple channels, an equivalent view of it is that there are multiple co-located users each occupying one channel. Thus, provided that the number of channels a user occupies is fixed, this equivalent transformation to satisfy assumption 6 does not lose generality.*

From assumption 7, users do not vary their selected power levels, and we focus on the problem of *channel allocation:* For every channel, choose the co-channel users in all the cells.

Furthermore, we characterize the inter-cell interference by introducing the following definition:

**Definition 17** (Interference Neighborhood)**.** *For any cell $A$, its interference neighborhood $\mathcal{N}(A)$ is the set of cells such that, for any cell $B \notin \mathcal{N}(A),\ B \neq A$,*

1. *The signal from cell $A$ is negligible in cell $B$ because of the propagation loss from $A$ to $B$.*

2. *The signal from cell $B$ is negligible in cell $A$ because of the propagation loss from $B$ to $A$.*

3. $A \notin \mathcal{N}(A)$.

**Remark 10.** *Given any critical interference level below which we consider the interference to be negligible, $\mathcal{N}(A)$ can be determined accordingly. Thus, by tuning $\mathcal{N}(A)$, it can characterize the actual inter-cell interference within the network arbitrarily closely.*

From the definition of interference neighborhood, we immediately observe the following:

**Remark 11.** $B \in \mathcal{N}(A)$ *if and only if* $A \in \mathcal{N}(B)$.

### Notations

- For the $i^{th}$ user $(i = 1, 2, \ldots, M)$ in cell $A$, we denote its index by $A_i$. In other words, a user is assigned with two labels: the cell that it belongs to ($A$), and its user index within this cell ($i$).

- For a cell $A$, we denote by $A(m)$ the index of *the user in cell $A$ that occupies channel $m$*. E.g., $A(m) = A_i$ means that in cell $A$, channel $m$ is occupied by the $i^{th}$ user in this cell. As each channel is occupied by one and only one user in a cell, $A(m), 1 \leq m \leq M$ is always a *permutation* function within cell $A$.

- For user $A(m)$, we denote by $p_{A(m)}$ the transmit power of this user, and $N_{A(m)}$ the noise power seen by this user in channel $m$. We denote by $g_{A(m)B}$

the channel gain from user $A(m)$ to the base station in cell $B$ in channel $m$, and $g_{BA(m)}$ the channel gain from the base station in cell $B$ to user $A(m)$ in channel $m$.

**Remark 12.**

- *$N_{A(m)}$ represents the noise level at the base station in the uplink case, and that at the mobile in the downlink case.*

- *Given $B$, $g_{A(m)B}, g_{BA(m)}, g_{A(m)A}, g_{AA(m)}$, and $N_{A(m)}$ are functions of not only the cell index $A$, the channel index $m$, but also the user index $A(m)$. (E.g., $A(m) = A_i$ and $A(m) = A_j$ may result in different $g_{A(m)B}$.)*
  *In other words, $A(m)$ is itself a function that specifies the user index, and the user index is thus implicitly included in the sub-indices of $g_{A(m)B}$, $g_{BA(m)}$, $g_{A(m)A}$, $g_{AA(m)}$, and $N_{A(m)}$.*

In comparison, one may also use more elaborate notations, e.g., $g_{A_{i,m}B}$, denoting the channel gain from user $A_i$ to the BS of $B$ *if user $A_i$ occupies channel $m$ (i.e., $A(m) = A_i$).* However, we find that using $A(m)$ in the sub-indices as a *compact* notation representing cell, channel and user indices all together leads to simpler expressions, and yet better clarity. In particular, the representation of co-channel inter-cell interference becomes simple and clear.

As a result, the rate of user $A(m)$ has the following expressions:

$$\text{Uplink: } R_{A(m)} = \log\left(1 + \frac{p_{A(m)}g_{A(m)A}}{\sum_{B\in\mathcal{N}(A)} p_{B(m)}g_{B(m)A} + N_{A(m)}}\right). \tag{3.1}$$

$$\text{Downlink: } R_{A(m)} = \log\left(1 + \frac{p_{A(m)}g_{AA(m)}}{\sum_{B\in\mathcal{N}(A)} p_{B(m)}g_{BA(m)} + N_{A(m)}}\right). \tag{3.2}$$

## 3.2 Uplink Channel Allocation in One Dimensional Networks: Signal Scale Interference Alignment

In this section, we study the problem of finding the *uplink* channel allocation that maximizes the *total throughput* in one-dimensional cellular networks, (i.e., with BS positioned on a straight line:)

$$\max_{A(m),\forall m,A} \sum_{A} \sum_{m=1}^{M} R_{A(m)} \tag{3.3}$$

We consider the case of *frequency flat* fading: for *any pair* of transmitter (mobile) and receiver (BS), it sees the same channel gain and noise level in all the $M$ channels, further described as follows.

- $g_{A(m)B}$ given that $A(m) = A_i$ is equal to $g_{A(m')B}$ given that $A(m') = A_i$, $\forall A, B, i, m \neq m'$.

  Thus, specifying the *user index* at the transmitter and the *cell indices* is sufficient to determine the channel gain, regardless of the channel index. E.g., $g_{A_iB}$ has an unambiguous value, denoting the channel gain from user $A_i$ to the BS of $B$.

- $\forall A,\ N_{A(m)} = N_A$ regardless of $A(m)$, denoting the frequency flat noise level at the BS of $A$.

We further assume that, for any cell, interference from users not from the immediate neighboring cells is ignored. In other words, $\mathcal{N}(A)$ only contains the cells that are immediate neighbors of $A$.

We exploit a revenue-cost separation principle, and equivalently transform maximizing the total throughput into minimizing the total interference cost of all cells. The key idea in obtaining the optimal channel allocation is the alignment of interference signal scale. We show that with very low complexity, a complete optimal channel allocation can be constructed by rippling along all the cells the proposed signal scale interference alignment procedure. From analyzing the consequence of the signal scale interference alignment, the superiority of the optimal channel allocation over CDMA is justified.

### 3.2.1 Revenue-Cost Separation Principle

From the last chapter, we have seen that an approximation of (3.1) is

$$R_{A(m)} \approx \log\left(\frac{p_{A(m)}g_{A(m)A}}{\sum_{B\in\mathcal{N}(A)} p_{B(m)}g_{B(m)A} + N_A}\right), \forall A, \tag{3.4}$$

and a *bounded gap* approximation of (3.1) is

$$R_{A(m)} \approx \log\left(\frac{p_{A(m)}g_{A(m)A}}{\sum_{B\in\mathcal{N}(A)} p_{B(m)}g_{B(m)A} + N_A}\right)^+, \forall A, \tag{3.5}$$

Using (3.5), *provided that* the optimal CA maximizing (3.3) satisfies that *every user has a positive dB* $\mathsf{SINR}$, we can safely drop the $^+$ in (3.5), and obtain the *same* optimal CA by maximizing (3.3) using (3.4) instead.

In this section, we assume that the optimal CA does lead to every user having a positive dB $\mathsf{SINR}$, and we use (3.4) as the rate expression in maximizing (3.3). This is a reasonable assumption because *every user only sees interference from other cells*. Accordingly, we have the revenue-cost separation principle:

$$\begin{aligned} R_{A(m)} &= \log(p_{A(m)}g_{A(m)A}) - \log(\sum_{B\in\mathcal{N}(A)} p_{B(m)}g_{B(m)A} + N_A) \\ &= \mathsf{revenue}_{A(m)} - \mathsf{cost}_{A(m)},\ \forall A, m, \end{aligned} \tag{3.6}$$

where

$$\mathsf{revenue}_{A(m)} \triangleq \log(p_{A(m)} g_{A(m)A}), \tag{3.7}$$

$$\mathsf{cost}_{A(m)} \triangleq \log(\sum_{B \in \mathcal{N}(A)} p_{B(m)} g_{B(m)A} + N_A). \tag{3.8}$$

From the definition (3.7), we immediately observe the following:

**Lemma 7** (Revenue Invariance)**.** *The revenue of any user $A_i$ is invariant to the CA of $A$.*

*Proof.* Suppose user $A_i$ occupies channel $m$ in cell $A$, i.e., $A(m) = A_i$. Then

$$\mathsf{revenue}_{A_i} = \log(p_{A_i} g_{A(m)A}). \tag{3.9}$$

Because user $A_i$ transmits at a fixed power level $p_{A_i}$, and we assume frequency flat fading, (3.9) is a constant that does not depend on the channel index $m$ as long as $A(m) = A_i$. □

Therefore, to *maximize* the network total throughput, it is equivalent to find the optimal CA that has the *minimum total* $\mathsf{cost}$ incurred:

$$\min_{A(m), \forall m, A} \sum_{A} \sum_{m=1}^{M} \mathsf{cost}_{A(m)} \tag{3.10}$$

### 3.2.2 Optimal Solution for the Local Problem

Consider any three consecutive cells on a line, we denote the center cell by $C$, the left cell by $L$, and the right cell by $R$ (Figure 3.1). By assumption we have $\mathcal{N}(C) = \{L, R\}$.

We now consider the local problem of minimizing the *total* cost that $C$ sus-

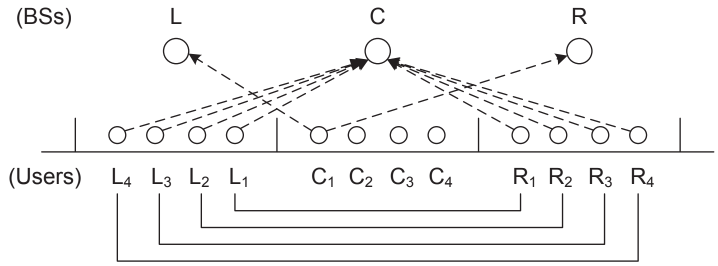



Figure 3.1: Three consecutive cells in a uplink one dimensional network. Users connected by lines are suggested co-channel users.

tains, generated from $L$ and $R$:

$$\begin{aligned}&\min_{L(m),C(m),R(m)} \sum_{m=1}^{M} \mathsf{cost}_{C(m)} \\ &= \min_{L(m),C(m),R(m)} \sum_{m=1}^{M} \log(p_{L(m)} g_{L(m)C} + p_{R(m)} g_{R(m)C} + N_C). \end{aligned} \tag{3.11}$$

Clearly, (3.11) does not depend on the CA of cell $C$ itself, but depends on the CA of $L$ and $R$. Furthermore, from the symmetry provided by the flatness of the channels, we have the following lemma:

**Lemma 8.** *The optimal solution of* (3.11) *can be achieved by first determining either $L$ or $R$'s CA arbitrarily, and then optimizing the other one's CA correspondingly.*

Now, consider the *two* cells $L$ and $C$ only. We have the following lemma on the interference received by $C$ from $L$.

**Lemma 9.** *The set of (a total of $M$) interference strengths generated from the users in $L$ to the BS of $C$ is invariant to the CA of $L$.*

Similarly to Lemma 7, this is an immediate implication of the flat channel assumption. In fact, for *any one* user $L_i$, its interference strength seen at the

BS of $C$ does not depend on which channel it occupies. According to Lemma 9, denote the set of interference strengths from $L$ to $C$ by

$$\{I_L^1, \ldots, I_L^m\},\ I_L^1 \geq I_L^2 \ldots \geq I_L^M. \tag{3.12}$$

Due to the flat channel assumption, WLOG, we *index* the $M$ channels such that $I_L^m$ is the interference from $L$ to $C$ in channel $m$. Consequently, the channel allocation function $L(m)(m = 1, \ldots, M)$ implied by such channel indexing satisfies that $I_L^m = p_{L(m)} g_{L(m)C}$.

Next, we would like to optimize the CA in $R$, i.e., $\{R(m), m = 1, \ldots, M\})$, such that the total cost generated from $\{L\ \textit{and}\ R\}$ to $C$ is minimized (3.11).

We start with an arbitrary initial CA $R(m)$, and compute

$$\{I_R^m = g_{R(m)C} p_{R(m)},\ m = 1 \ldots M\}. \tag{3.13}$$

With the channel allocation functions $L(m)$ and $R(m)$, the total cost that the users in $C$ sustain is

$$\sum_{m=1}^{M} \mathsf{cost}_{C(m)} = \sum_{m=1}^{M} \log(I_L^m + I_R^m + N_C), \tag{3.14}$$

where $I_L^m$ and $I_R^m$ $(m = 1, \ldots, M)$ are co-channel interferences.

For any *other* CA $R'(m)$, it can be represented by a permutation function $\mathcal{P}(m)(m = 1, \ldots, M)$ applied to the initial CA $R(m)$, such that

$$R'(m) = R(\mathcal{P}(m)).$$

In other words, *channel $m$ is assigned to the user who initially has channel $\mathcal{P}(m)$ assigned to it.* With $L(m)$ and $R'(m)$, the total cost that the users in $C$ sustain changes to

$$\sum_{m=1}^{M} cost'_{C(m)} = \sum_{m=1}^{M} \log(I_L^m + I_R^{\mathcal{P}(m)} + N_C), \tag{3.15}$$

where $I_L^m$ and $I_R^{\mathcal{P}(m)}$ $(m = 1, \ldots, M)$ are the new co-channel interferences.

Therefore, the optimal CA problem is formulated as follows:

$$\min_{\mathcal{P}(m)} \sum_{m=1}^{M} \log(I_L^m + I_R^{\mathcal{P}(m)} + N_C). \tag{3.16}$$

The key in finding the optimal $\mathcal{P}(m)$ (and hence the optimal CA in $R$) among the $M!$ permutations is an idea of *signal scale interference alignment*. We sort $\{I_R^m, m = 1, \ldots, M\}$ in descending order: $I_R^{m_1} \geq I_R^{m_2} \geq \ldots \geq I_R^{m_M}$, and *define* $\mathcal{P}^*(j) \triangleq m_j, j = 1, \ldots, M$. The following theorem can be shown:

**Theorem 8** (Signal Scale Interference Alignment)**.** *Among all permutation functions $\mathcal{P}(m)$, $\{\mathcal{P}^*(m), m = 1, \ldots, M\}$ yields the minimum total cost that the users in $C$ sustain.*

**Remark 13.** *From the definition of $\mathcal{P}^*(m)$, we have $I_L^1 \geq I_L^2 \ldots \geq I_L^M$ and $I_R^{\mathcal{P}^*(1)} \geq I_R^{\mathcal{P}^*(2)} \geq \ldots \geq I_R^{\mathcal{P}^*(M)}$. What has been done in the optimal CA is that we align the strongest interference each from L and R to co-exist in the same channel, and so on for the $2^{nd}$ strongest, ..., all the way to aligning the weakest interference each from L and R to co-exist in the same channel.*

*Proof of Theorem 8.* We prove an equivalent form of the theorem: If $a_1 \geq a_2 \geq \ldots \geq a_M \geq 0,\ b_1 \geq b_2 \geq \ldots \geq b_M \geq 0,\ N \geq 0$, then $\sum_{m=1}^{M} \log(a_m + b_m + N) \leq \sum_{m=1}^{M} \log(a_m + b_{f(m)} + N)$, for all permutation functions $f(m), m = 1, \ldots, M$.

We use induction on $M$ as follows.

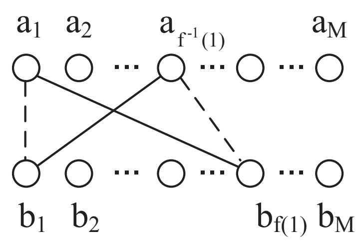



Figure 3.2: After replacing $(a_1, b_{f(1)})$ and $(a_{f^{-1}(1)}, b_1)$ with $(a_1, b_1)$ and $(a_{f^{-1}(1)}, b_{f(1)})$, the total cost decreases, (or remains unchanged.)

$M = 1$ is trivial. For $M = 2$, by Jensen's inequality,

$$\frac{a_1 - a_2}{a_1 - a_2 + b_1 - b_2} \log(a_1 + b_1 + N) + \frac{b_1 - b_2}{a_1 - a_2 + b_1 - b_2} \log(a_2 + b_2 + N) \leq \log(a_1 + b_2 + N) \tag{3.17}$$

$$\frac{b_1 - b_2}{a_1 - a_2 + b_1 - b_2} \log(a_1 + b_1 + N) + \frac{a_1 - a_2}{a_1 - a_2 + b_1 - b_2} \log(a_2 + b_2 + N) \leq \log(a_2 + b_1 + N) \tag{3.18}$$

(3.17)+(3.18) implies the theorem for $M = 2$.

Suppose the theorem holds for all $M \leq k$.

For $M = k+1$, we first represent the problem by a bipartite graph (Figure 3.2): it consists of upper $M$ points $a_1, \ldots, a_M$ and lower $M$ points $b_1, \ldots, b_M$, and there is an (undirected) edge between every $a_m$ and $b_{f(m)}$. The theorem is then equivalent to claiming that the graph with all "vertical" edges $(a_m, b_m), m = 1, \ldots, M$ yields the minimum total cost. Proof follows:

Given a graph generated from a permutation function $f(m)$:

i) If edge $(a_1, b_1)$ is in the graph, then after removing $(a_1, b_1)$, the rest of the graph degrades to an $M = k$ case, and applying the induction assumption proves this case.

ii) If edge $(a_1, b_1)$ is not in the graph, applying the induction assumption with

$M = 2$ yields,

$$\log(a_1 + b_{f(1)} + N) + \log(a_{f^{-1}(1)} + b_1 + N) \geq \log(a_1 + b_1 + N) + \log(a_{f^{-1}(1)} + b_{f(1)} + N) \tag{3.19}$$

In other words, after replacing the two edges $(a_1, b_{f(1)})$ and $(a_{f^{-1}(1)}, b_1)$ with $(a_1, b_1)$ and $(a_{f^{-1}(1)}, b_{f(1)})$, the total cost decreases (or remains unchanged), and the new bipartite graph falls into the case of i). □

### 3.2.3 Optimal Solution in Infinite One Dimensional Networks

Now, consider a two-sided infinite one dimensional cellular network. The above local alignment procedure can then be used in a rippling manner along the one dimensional network to obtain the optimal CA of all cells — indexed as $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$. We show that the globally optimal CA over the entire network can be achieved by a two-stage optimization, with each stage optimizing half of the cells in the network.

**Stage 1** We minimize the total cost sustained by cells $\{\ldots, -3, -1, 1, 3, \ldots\}$: this leads to the optimal CA in cells $\{\ldots, -2, 0, 2, \ldots\}$, specified by the following algorithm.

*Algorithm 3.1*: two-stage rippling of signal scale interference alignment

---

*Step 0.* Assign an arbitrary CA to cell 0;

*Step 1a.* Based on the CA in cell 0, assign the CA in cell 2 according to the the alignment rules in Theorem 8, and the total cost that cell 1 sustains is minimized;

Similarly,

*Step 1b.* CA in cell 0 $\Rightarrow$ CA in cell -2: the cost in cell -1 is minimized;

*Step 2a.* CA in cell 2 $\Rightarrow$ CA in cell 4: the cost in cell 3 is minimized;

*Step 2b.* CA in cell -2 $\Rightarrow$ CA in cell -4: the cost in cell -3 is minimized;

And so on.

---

Note that in Step 0, the arbitrary CA assignment of cell 0 does not lose any generality, and hence optimality (cf. Lemma 8).

**Stage 2** Similarly to Stage 1, we minimize the total cost sustained by cells $\{\ldots, -2, 0, 2, \ldots\}$ by applying the rippling procedure to cells $\{\ldots, -3, -1, 1, 3, \ldots\}$.

After the above two stages of CA, we obtain a complete optimal CA that yields the *minimum total cost* of all the cells. Because of the revenue invariance property (cf. Lemma 7), it achieves the *maximum total throughput*.

**Remark 14.** *Finding a complete optimal CA has a complexity of* $O(K_{cell} M \log M)$*, where* $K_{cell}$ *is the number of cells: The term* $K_{cell}$ *comes from the rippling procedure as above, and the term* $M \log M$ *comes from sorting the interference strengths for each cell (using e.g. Heapsort.)*

**Remark 15.** *For all the users in any one cell, the ordering of their interference strengths at the left neighboring BS could be different from that at the right neigh-*

*boring BS. For example, in Figure 3.1, assuming interference strength decreases as the transmission distance increases, user $C_1$ (among all the users in $C$) creates the strongest interference to the BS of $L$, but the weakest interference to the BS of $R$. To guarantee that a CA is globally optimal, users within each cell must align their interference strengths with those from both their $2^{nd}$ left and $2^{nd}$ right neighboring cells, (as achieved by the proposed rippling procedure.)*

### 3.2.4 Performance Evaluation: Optimal CA vs. CDMA

In this section, we give a typical example that numerically compare the *average spectral efficiency* of the optimal CA with that of CDMA schemes. We assume the simplified path loss model [Gol05] $P_r = P_t K(\frac{d_0}{d})^\gamma$ with $d_0 = 50m$ (outdoor environment), and no multipath or shadow fading. The parameter $K$ is irrelevant to the comparison between optimal CA and CDMA, and is assumed to be 1. $\gamma$ between 2.5 and 4 will be tested below. We assume a one dimensional cellular network with cell radius $= 500m$, and that there are $M$ users equally spaced in every cell.

In this example, we assume that all users transmit at the same power level. We compute a completely interference limited case, i.e., noise power is ignored. In this case, the value of the PSD is irrelevant to the comparison between optimal CA and CDMA, and is assumed to be 1. For any three consecutive cells $L$, $C$, $R$, the optimal CA derived in Section 3.2.2 yields a fully symmetric CA in $L$ and $R$ such that users in $L$ and $R$ with the same distance to the BS of $C$ co-exist in the same channel (Figure 3.1). The average cost per user is

$$\overline{cost}_{\mathsf{opt.CA}} = \frac{1}{M} \sum_{i=1}^{M} \log(2I_i) \quad \text{bits/sec/Hz} \tag{3.20}$$

where $I_i$ $(i = 1, \ldots, M)$ traverses the interference from all $M$ positions of the users in $L$ (and $R$ symmetrically.)

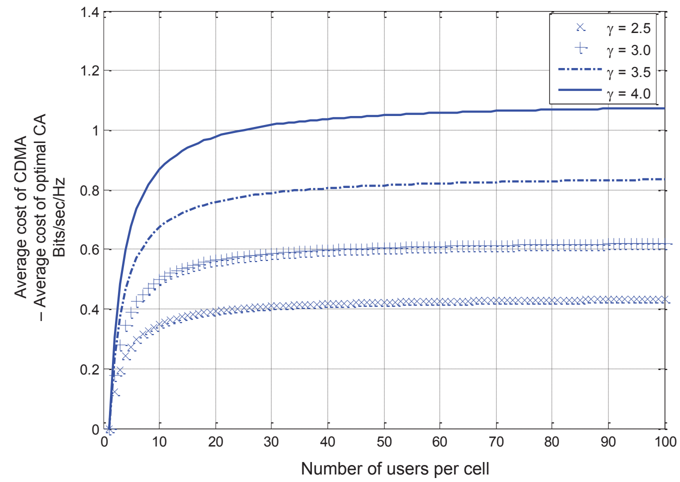



Figure 3.3: Superiority of optimal CA over CDMA

With CDMA, every user in a cell sustains the *same* amount of cost from other cells (since interference is averaged:)

$$\overline{cost}_{\mathsf{CDMA}} = \log(\frac{1}{M}\sum_{i=1}^{M} 2I_i) \quad \text{bits/sec/Hz}, \tag{3.21}$$

From Jensen's inequality, the superiority of the optimal CA over CDMA becomes evident: $\overline{cost}_{\mathsf{CDMA}} \geq \overline{cost}_{\mathsf{opt.CA}}$ always. We plot in Figure 3.3 the difference $\overline{cost}_{\mathsf{CDMA}} - \overline{cost}_{\mathsf{opt.CA}}$ as a function of $M$, parameterized by $\gamma = 2.5, 3, 3.5, 4$. Note that from revenue invariance, $-(\overline{cost}_{\mathsf{CDMA}} - \overline{cost}_{\mathsf{opt.CA}})$ equals the throughput difference in terms of the average spectral efficiency. We observe the following:

1. As the number of users increases and/or as $\gamma$ increases, the *variation* in the set of interference strengths $\{I_i\}$ increases. Thus the gap from Jensen's inequality, and hence the cost difference, increases.

2. As one numerical rule of thumb in this particular example, with $\gamma = 4$, the superiority in throughput of optimal CA over CDMA reaches above 1 bits/sec/Hz when the number of users reaches 25 per cell.

## 3.3 Downlink Channel Allocation in One Dimensional Networks: Decomposition and Assignment Problem

In this section, we turn our focus to *downlink* channel allocation. We assume the same network and channel settings as in the last section: one dimensional cellular networks with $\mathcal{N}(A)$, $\forall A$ containing the immediate neighboring cells of $A$, and frequency flat fading. We again assume that the optimal CA leads to every user having a positive dB $\mathsf{SINR}$. In addition, we also assume an *interference limited* condition: in every cell, the noise level is ignored compared to the interference.

Similarly to the uplink case, we use the following approximate rate expression with a revenue-cost separation principle in the downlink case:

$$\begin{aligned} R_{A(m)} &= \log(p_{A(m)} g_{AA(m)}) - \log(\sum_{B \in \mathcal{N}(A)} p_{B(m)} g_{BA(m)}) \\ &= \mathsf{revenue}_{A(m)} - \mathsf{cost}_{A(m)}, \ \forall A, m, \end{aligned} \quad (3.22)$$

where

$$\mathsf{revenue}_{A(m)} \triangleq \log(p_{A(m)} g_{AA(m)}), \quad (3.23)$$

$$\mathsf{cost}_{A(m)} \triangleq \log(\sum_{B \in \mathcal{N}(A)} p_{B(m)} g_{BA(m)}). \quad (3.24)$$

Similarly to the last section, we again have the revenue invariance property, i.e., Lemma 7 holds in the downlink case. Thus, the network total throughput maximization problem is equivalently transformed to network total cost minimization problem:

$$\min_{A(m), \forall m, A} \sum_{A} \sum_{m=1}^{M} \mathsf{cost}_{A(m)} \quad (3.25)$$

**Intractability of the Local Problems**

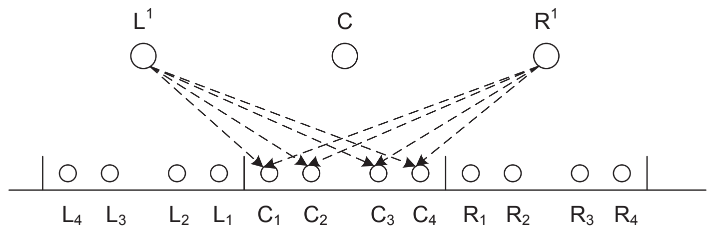



Figure 3.4: Three consecutive cells in a downlink one dimensional network.

Consider any three consecutive cells $\{L^1,\ C,\ R^1\}$[1], and the local problem of minimizing the *total* cost that $C$ sustains, generated from $L^1$ and $R^1$:

$$\begin{aligned} &\min_{L^1(m),C(m),R^1(m)} \sum_{m=1}^{M} \mathsf{cost}_{C(m)} \\ =&\min_{L^1(m),C(m),R^1(m)} \sum_{m=1}^{M} \log(p_{L^1(m)} g_{L^1C(m)} + p_{R^1(m)} g_{R^1C(m)}). \end{aligned} \tag{3.26}$$

As a key difference from the uplink case (3.11), (3.26) not only depends on the CA of $L^1$ and $R^1$, but also depends on the CA of cell $C$ itself. Nevertheless, we have the following lemma similar to Lemma 8 based on the symmetry provided by the flatness of the channels:

**Lemma 10.** *The optimal solution of* (3.26) *can be achieved by first arbitrarily determining the CA of any one of the three cells $L^1, C, R^1$, and then optimizing the other two cells's CA correspondingly.*

For example, we can first let $C(m) = C_m, m = 1, 2, \ldots, M$ WLOG, and (3.26) can then be solved by jointly optimizing the CA of $L^1$ and $R^1$ :

$$\min_{L^1(m),R^1(m)} \sum_{m=1}^{M} \log(p_{L^1(m)} g_{L^1C_m} + p_{R^1(m)} g_{R^1C_m}). \tag{3.27}$$

where $L^1(m), R^1(m)$ specifies the users in $L^1$ and $R^1$ that are co-channel with user $C_m$ in $C$. Similarly, (3.26) can also be solved by first arbitrarily determining

[1]The super-indices "1" are added to $L$ and $R$ to denote that they are the $1^{st}$ left and $1^{st}$ right neighboring cells of $C$.

the CA of either $L^1$ or $R^1$, and then jointly optimize the CA of the remaining two cells.

However, We have the following conjecture on the hardness of solving a problem like (3.27)[2]:

**Conjecture 1.** *Given the CA of any one of the three consecutive cells $L^1, C, R^1$, the local problem of minimizing the total cost in $C$ by jointly optimizing the CA of the two remaining cells (e.g.* (3.27)*) is $NP$ hard in the number of users $M$.*

Conjecture 1 motivates us to find other forms of local problems that can be solved with low complexity and also effectively applied in optimizing the entire network.

### Lower Bound on the Minimum Total Cost

Although (3.27) cannot be efficiently solved, by applying Jensen's Inequality, we obtain the following lower bound on the optimal value of (3.27):

$$\begin{aligned}
&\sum_{m=1}^{M} \log(p_{L^1(m)} g_{L^1 C_m} + p_{R^1(m)} g_{R^1 C_m}) \\
\geq & \frac{1}{2}\left( \sum_{m=1}^{M} \log(2 p_{L^1(m)} g_{L^1 C_m}) + \sum_{m=1}^{M} \log(2 p_{R^1(m)} g_{R^1 C_m}) \right) \qquad (3.28) \\
= & M + \frac{1}{2}\left( \sum_{m=1}^{M} \log p_{L^1(m)} + \sum_{m=1}^{M} \log g_{L^1 C_m} + \sum_{m=1}^{M} \log p_{R^1(m)} + \sum_{m=1}^{M} \log g_{R^1 C_m} \right) \\
= & M + \frac{1}{2}\left( \sum_{m=1}^{M} \log p_{L^1_m} + \sum_{m=1}^{M} \log g_{L^1 C_m} + \sum_{m=1}^{M} \log p_{R^1_m} + \sum_{m=1}^{M} \log g_{R^1 C_m} \right) \qquad (3.29)
\end{aligned}$$

where (3.28) comes from the concavity of the log function and Jensen's Inequality, and (3.29) occurs because the channel allocation functions $L^1(m)$ and $R^1(m)$ are permutation functions within cells $L^1$ and $R^1$ respectively. Note that (3.29) is a

[2]The proof of the conjecture remains an open problem.

constant that does not depend on $L^1(m)$ and $R^1(m)$, and thus serves as a lower bound on the optimal value of (3.27).

We note that the lower bound (3.29) is tight if and only if (3.28) is met with equality, i.e.,

$$\forall m = 1, 2, \ldots, M, \; p_{L^1(m)} g_{L^1 C_m} = p_{R^1(m)} g_{R^1 C_m}. \tag{3.30}$$

(3.3.1) implies that, in every channel, "perfect" signal scale interference alignment is achieved: the interference from the left neighboring base station $L^1$ exactly equals to the interference from the right neighboring base station $R^1$. In light of this, the intuition of minimizing the total cost is really to approach as much signal scale interference alignment as possible. However, perfect interference alignment (3.3.1) may not be achievable even with the *optimal* solution of (3.27).

### 3.3.1 Iterative Distributed Heuristics Based on Signal Scale Interference Alignment

#### Simplified Local Problems: Signal Scale Interference Alignment

For minimizing the total cost in $C$, instead of jointly optimizing the CA of two cells among $L^1, C, R^1$ while fixing the other one, we simplify the the problem to *optimizing the CA of one cell while fixing the other two*: Given $L^1(m), R^1(m)$,

$$\min_{C(m)} \sum_{m=1}^{M} \mathsf{cost}_{C(m)} = \min_{C(m)} \sum_{m=1}^{M} \log(p_{L^1(m)} g_{L^1 C(m)} + p_{R^1(m)} g_{R^1 C(m)}) \tag{3.31}$$

We make the following observation:

$$\begin{aligned}(3.31) &= \min_{C(m)} \sum_{m=1}^{M} \log \left( g_{L^1 C(m)} \left( p_{L^1(m)} + p_{R^1(m)} \frac{g_{R^1 C(m)}}{g_{L^1 C(m)}} \right) \right) \\ &= \sum_{m=1}^{M} \log(g_{L^1 C(m)}) + \min_{C(m)} \sum_{m=1}^{M} \log(p_{L^1(m)} + \tilde{p}_{C(m)}), \end{aligned} \tag{3.32}$$

where $\tilde{p}_{C(m)} \triangleq p_{R^1(m)} \frac{g_{R^1C(m)}}{g_{L^1C(m)}}$. Note that (3.32) and (3.16) (with $N_C = 0$) are essentially the *same* problem, and we have the following corollary of Theorem 8:

**Corollary 11** (Signal Scale Interference Alignment)**.** *After transforming to* (3.32), (3.31) *can solved with* $O(M \log M)$ *complexity by sorting and aligning* $\{p_{L^1(m)}\}$ *and* $\{\tilde{p}_{C(m)}\}$.

Note that, in (3.16), the physical meaning of signal scale interference alignment is clear: sorting and aligning the interference strengths (seen at the BS of the center cell) from the mobiles in the left neighboring cell and the right neighboring cell. In (3.32), however, the meaning of sorting and aligning $\{p_{L^1(m)}\}$ and $\{\tilde{p}_{C(m)}\}$ is less evident. To understand that it is indeed signal scale interference alignment, consider the case where $p_{L^1(m)} = \tilde{p}_{C(m)}$, i.e.,

$$p_{L^1(m)} = p_{R^1(m)} \frac{g_{R^1C(m)}}{g_{L^1C(m)}} \Leftrightarrow p_{L^1(m)} g_{L^1C(m)} = p_{R^1(m)} g_{R^1C(m)}. \tag{3.33}$$

(3.33) means that the interference strength from the left neighboring BS to the mobile occupying channel $m$ in the center cell is equal to that from the right neighboring BS, corresponding to the *perfect* signal scale interference alignment in reaching the lower bound (3.29). Thus, we see that the intuition of Corollary 11 is again signal scale interference alignment.

**Network Optimization:**

**Iterative Distributed Signal Scale Interference Alignment**

Now, Consider a two-sided infinite one dimensional cellular network with cell indices $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$. Using the above simplified local optimization (3.31) as building blocks, we propose an *iterative* two-stage algorithm as follows (cf. Figure 3.5):

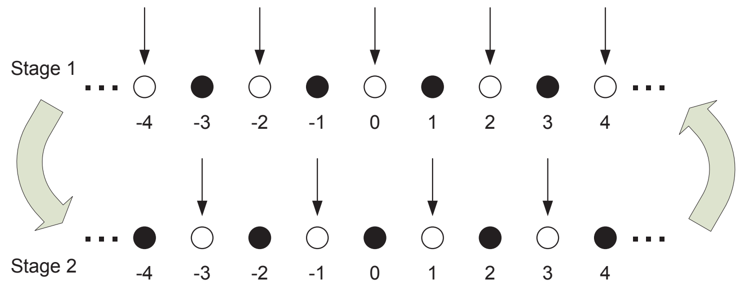



Figure 3.5: Iterative distributed two-stage algorithm for channel allocation. In each stage, the solid circles denote the cells whose CA are fixed, and the open circles denote the cells whose CA are optimized as local signal scale interference alignment.

*Algorithm 3.2*: iterative distributed signal scale interference alignment

---

Assign arbitrary initial CAs to all cells.

Repeat

Stage 1: with the CAs in cells $\ldots, -3, -1, 1, 3, \ldots$ fixed,

optimize the CAs of cells $\ldots, -2, 0, 2, \ldots$ by (3.31).

Stage 2: with the CAs in cells $\ldots, -2, 0, 2, \ldots$ fixed,

optimize the CAs of cells $\ldots, -3, -1, 1, 3, \ldots$ by (3.31).

until approximate convergence.

---

Note that, in each stage, all the local optimization are independently performed. Thus, Algorithm 3.2 is a *distributed* iterative algorithm. To perform signal scale interference alignment, however, information exchange is needed between adjacent cells. For example, in forming (3.32), although the received signal power $p_{L^1(m)}g_{L^1C(m)}$ can be measured at the mobile $C(m)$ in cell $C$, either the transmit power $p_{L^1(m)}$ or channel gain $g_{L^1C(m)}$ needs to be known at cell $C$.

Despite its distributed nature, Algorithm 3.2 does not guarantee that the network total cost is non-decreasing after each stage, and hence does not guarantee convergence. For example, optimizing the CA of cell 0 affects not only the cost in cell 0, *but also the cost in cells -1 and 1.* However, the local optimization at cell 0 does not take its influence in cells -1 and 1 into account. Therefore, Stage 1 has *no control* on the resulting total cost in cells $\ldots, -3, -1, 1, 3, \ldots$, and Stage 2 has no control on the resulting total cost in cells $\ldots, -2, 0, 2, \ldots$. As will be shown in Section 3.3.2.3, by iterating the two stages repeatedly, Algorithm 3.2 can only achieve *approximate* convergence.

We address the above issues in the next section, and develop a much more general framework of optimizing CA over the entire network, with yet better performance.

### 3.3.2 Iterative Decomposed Heuristics Based on Local Assignment Problems

#### 3.3.2.1 Local problems as Assignment Problems

To optimize the CA of $C$, in addition to the total cost in $C$ itself, we also include into the objective the total cost *in C's entire interference neighborhood* $\mathcal{N}(C)$:

$$\min_{C(m)} \sum_{A \in C \cup \mathcal{N}(C)} \sum_{m=1}^{M} \mathsf{cost}_{A(m)}, \tag{3.34}$$

$$\text{where} \quad \mathsf{cost}_{A(m)} = \log\Big( \sum_{B \in \mathcal{N}(A)} p_{B(m)} g_{BA(m)} \Big). \tag{3.35}$$

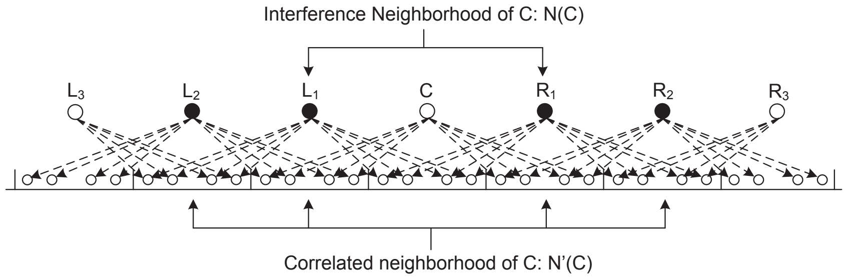



Figure 3.6: Interference Neighborhood $\mathcal{N}(C)$, Correlated Neighborhood $\mathcal{N}'(C)$ in a one dimensional network.

With our assumption on the interference neighborhood in one dimensional networks, $\mathcal{N}(C) = \{L^1, R^1\}$, and

$$(3.34) = \min_{C(m)} \sum_{m=1}^{M} (\mathsf{cost}_{L^1(m)} + \mathsf{cost}_{C(m)} + \mathsf{cost}_{R^1(m)}), \tag{3.36}$$

$$\text{where } \mathsf{cost}_{C(m)} = \log(p_{L^1(m)} g_{L^1C(m)} + p_{R^1(m)} g_{R^1C(m)}), \tag{3.37}$$

$$\mathsf{cost}_{L^1(m)} = \log(p_{L^2(m)} g_{L^2L^1(m)} + p_{C(m)} g_{CL^1(m)}), \tag{3.38}$$

$$\mathsf{cost}_{R^1(m)} = \log(p_{C(m)} g_{CR^1(m)} + p_{R^2(m)} g_{R^2R^1(m)}), \tag{3.39}$$

For notation, we use $L^k$ ($R^k$) to denote the $k^{th}$ left (right) neighboring cell of $C$. From (3.37),(3.38),(3.39), we see that to solve (3.36), we need to fix not only the CA in $L^1, R^1$, but also the CA in $L^2, R^2$. We call the set of cells $\{L^2, L^1, R^1, R^2\}$ the *Correlated Neighborhood* of $C$ (cf. Figure 3.6).

For the above local problem in its *general* form (3.34), in order to solve it, we need to fix the CA of all cells (other than $C$ itself) that appear in (3.35):

- Clearly, the CA of $\{A \mid A \in \mathcal{N}(C)\}$ need to be known and fixed. From Remark 11, this is equivalent to fixing the CA of

$$\{A \mid \mathcal{N}(A) \cap C \neq \emptyset\}. \tag{3.40}$$

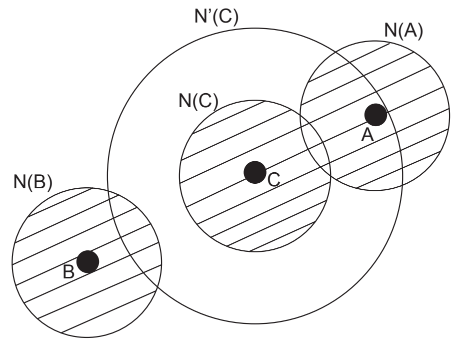



Figure 3.7: Correlated Neighborhood $\mathcal{N}'(C)$, where $A \in \mathcal{N}'(C)$ but $B \notin \mathcal{N}'(C)$. Shaded areas represent the interference neighborhood.

- Moreover, the CA of $\{B \mid \exists A \in \mathcal{N}(C), B \in \mathcal{N}(A)\}$ need to be known and fixed. From Remark 11, this is equivalent to fixing the CA of

$$\{B \mid \mathcal{N}(B) \cap \mathcal{N}(C) \neq \emptyset\}. \tag{3.41}$$

Combining (3.40) and (3.41), we see that the set of cells that need to be known and fixed while solving (3.34) is the following:

**Definition 18.** *The Correlated Neighborhood of cell $C$ is defined to be*

$$\mathcal{N}'(C) \triangleq \{A \mid \mathcal{N}(A) \cap (C \cup \mathcal{N}(C)) \neq \emptyset\}. \tag{3.42}$$

A conceptual depiction of the correlated neighborhood of $C$ is given in Figure 3.7. We summarize the general form local problem (3.34) as follows:

**Remark 16.**

1. *The optimization variables are the CA of $C$ itself.*

2. *The objective function includes the cost within the interference neighborhood of $C$, in addition to $C$ itself.*

3. *The set of cells whose CA need to be fixed consists of the correlated neighborhood of $C$.*

As a result, in (3.34), we include in the objective *all* the cells that are influenced by the CA of $C$, with necessarily more neighboring cells' CA fixed.

Now, we show that (3.34) can be solved efficiently as an *Assignment Problem* [Kuh55, BDM09] as follows:

**Theorem 9.** *Given the CA of all the cells in $C$'s correlated neighborhood $\mathcal{N}'(C)$,* (3.34) *can be solved as an assignment problem with a computational complexity of $O(M^3)$.*

*Proof.* Construct an $M \times M$ matrix $\mathbf{X}$, where $\forall m, n = 1, 2, \ldots, M$,

$$X_{mn} = \sum_{A \in C \cup \mathcal{N}(C)} \mathsf{cost}_{A(m)} \quad \textit{given that } C(m) = C_n. \tag{3.43}$$

In other words, $X_{mn}$ is the *total* cost sustained by a set of co-channel users *in case it is user $C_n$ that occupies channel $m$ in cell $C$*, and this set of co-channel users consists of all the users from $C \cup \mathcal{N}(C)$ *that occupy channel $m$*. (Clearly, this set of co-channel users automatically contains $C_n$ as the user from $C$ that occupies channel $m$.)

For example, in the above problem (3.36) of one dimensional network, $\mathcal{N}(C) = \{L^1, R^1\}$, $\mathcal{N}'(C) = \{L^2, L^1, R^1, R^2\}$. Then, $\forall m, n = 1, 2, \ldots, M$,

$$\begin{aligned} X_{mn} &= \mathsf{cost}_{C(m)} + \mathsf{cost}_{L^1(m)} + \mathsf{cost}_{R^1(m)} \quad \textit{given that } C(m) = C_n \\ &= \log(p_{L^1(m)} g_{L^1 C_n} + p_{R^1(m)} g_{R^1 C_n}) + \log(p_{L^2(m)} g_{L^2 L^1(m)} + p_{C_n} g_{C L^1(m)}) + \\ &\quad \log(p_{C_n} g_{C R^1(m)} + p_{R^2(m)} g_{R^2 R^1(m)}), \end{aligned} \tag{3.44}$$

Note that, for (3.43) (of which (3.44) is a special case), given the CA of $\mathcal{N}'(C)$ and given that $C(m) = C_n$, $X_{mn}$ has no ambiguity and can be readily computed.

With the matrix $\mathbf{X}$ constructed by (3.43), the total cost within the entire $C \cup \mathcal{N}(C)$ can be rewritten as

$$\sum_{A \in C \cup \mathcal{N}(C)} \sum_{m=1}^{M} \mathsf{cost}_{A(m)} = \sum_{m=1}^{M} X_{mC(m)}, \tag{3.45}$$

where we define $X_{mC(m)}$ to be $X_{mn}$ when $C(m) = C_n$, thus omitting the cell index "$C$" in $C_n$ without ambiguity. The local problem in its general from (3.34) can be rewritten as

$$\min_{C(m)} \sum_{m=1}^{M} X_{mC(m)}. \tag{3.46}$$

Note that $C(m), m = 1, 2, \ldots, M$ is a permutation function of $C_1, C_2, \ldots, C_M$, i.e., it is a bipartite matching of the $M$ channels to the $M$ users in $C$. Thus, (3.46) is exactly in the form of an assignment problem [BDM09], and hence can be solved by the Hungarian Algorithm [Kuh55] with a computational complexity of $O(M^3)$. □

#### 3.3.2.2 Iterative Decomposed Network Optimization

For a general wireless cellular network, we define sets of cells that can be independently and simultaneously optimized as *independent sets:*

**Definition 19.** *A set of cells $\mathcal{I}$ is an* independent set of cells, *if*

$$\forall A, B \in \mathcal{I}, A \neq B,\ A \notin \mathcal{N}'(B),\ \textit{and}\ B \notin \mathcal{N}'(A). \tag{3.47}$$

Based on the local problem (3.34), while optimizing the CA of any cell $A \in \mathcal{I}$, its correlated neighborhood $\mathcal{N}'(A)$ must be fixed. (3.47) guarantees that there is no cell in $\mathcal{I}$ that falls into the correlated neighborhood of another cell in $\mathcal{I}$.

Note that

$$(3.47)\ \Leftrightarrow\ \forall A, B \in \mathcal{I}, A \neq B,\ (A \cup \mathcal{N}(A)) \cap (B \cup \mathcal{N}(B)) = \emptyset. \tag{3.48}$$

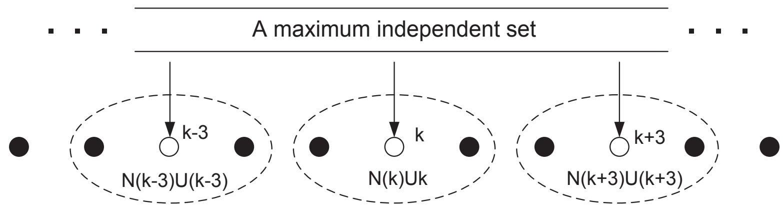



Figure 3.8: The open circles represent a maximum independent set in a one dimensional network. Each dotted circle denotes the cells that are influenced by an open circle.

In other words, the local optimization of the cells within $\mathcal{I}$ all have *disjoint* sets of cells that are influenced. Thus, the local optimization of the cells within $\mathcal{I}$ are *decomposed*. As a result, we have the following corollary:

**Corollary 12.** *For any independent set of cells $\mathcal{I}$, with the CAs of $\mathcal{I}^c$ (i.e. all cells* not *in $\mathcal{I}$) fixed, to achieve the global minimum cost over the entire network, every cell in $\mathcal{I}$ can optimize its own CA independently and simultaneously.*

Now, consider a two-sided infinite one dimensional cellular network with cell indices $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$. $\forall$ cell $k, \mathcal{N}'(k) = \{k-2, k-1, k+1, k+2\}$. We define a *maximum* independent set of cells to consist *one of every three* consecutive cells, with a uniform separation of two cells (Figure 3.8). Clearly, this is the *densest* set of independent cells in a one dimensional network. Note that there are three disjoint maximum independent set of cells whose union is the entire network. We denote them by $\mathcal{I}_0, \mathcal{I}_1, \mathcal{I}_2$, where $\mathcal{I}_k = \{k | k \equiv k \pmod 3\}$.

Based on the three maximum independent sets, we propose the following iterative *three-stage* algorithm as follows (cf. Figure 3.9):

*Algorithm 3.3:* iterative decomposed network optimization
based on local assignment problems

---

Assign arbitrary initial CAs to all cells.

Repeat

Stage 1: with the CA of cells in $\mathcal{I}_2, \mathcal{I}_3$ fixed,
optimize the CA of cells in $\mathcal{I}_1$ by (3.36).

Stage 2: with the CA of cells in $\mathcal{I}_1, \mathcal{I}_3$ fixed,
optimize the CA of cells in $\mathcal{I}_2$ by (3.36).

Stage 3: with the CA of cells in $\mathcal{I}_1, \mathcal{I}_2$ fixed,
optimize the CA of cells in $\mathcal{I}_3$ by (3.36).

until convergence.

---

**Remark 17.**

- *At every iteration, one third of the cells' CAs are optimized with an objective function covering all the cells.*

- *While optimizing a maximum independent set of cells $\mathcal{I}_i$, $\forall k \in \mathcal{I}_i$, the total cost within $k \cup \mathcal{N}(k) = \{k-1, k, k+1\}$ is minimized and hence non-increasing. From* (3.48)*, it is guaranteed that after every iteration, the total cost in the entire network is non-increasing. Therefore, the algorithm is guaranteed to converge.*

- *As each local optimization consumes a computational complexity of $O(M^3)$, the overall computational complexity of Algorithm 3.3 is $O(K_{cell}M^3)$. Moreover, the* running time *of Algorithm 3.3 is $O(M^3)$ as all the local optimiza-*

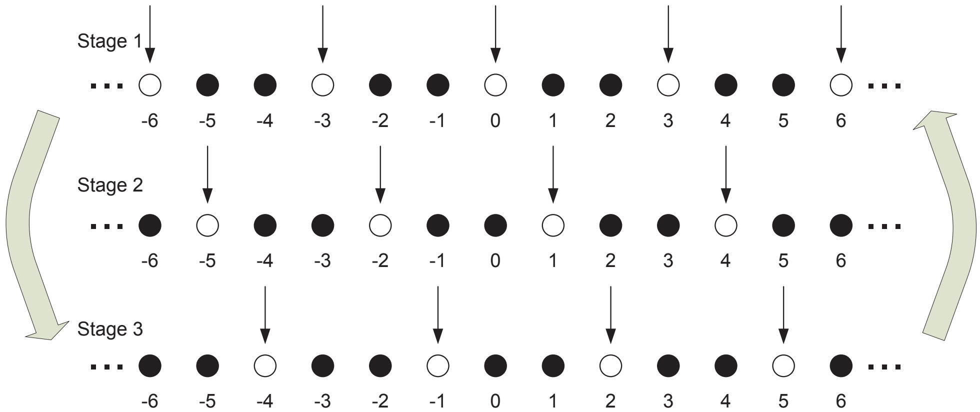



Figure 3.9: Iterative decomposed three-stage algorithm for channel allocation. In each stage, the solid circles denote the cells whose CA are fixed, and the open circles denote the cells whose CA are optimized as local assignment problems.

*tion in each stage are* simultaneously and independently *performed, i.e., distributed and decomposed.*

#### 3.3.2.3 Performance Evaluation of the Iterative Algorithms

We first examine CDMA (interference averaging) as a baseline scheme. For any cell $C$, the total cost within this cell with all the cells in the network using CDMA is given by:

$$\mathsf{cost}^{C}_{\mathsf{CDMA}} = \sum_{m=1}^{M} \log\left(\frac{\sum_{n=1}^{M} p_{L_1(n)}}{m} g_{L_1C(m)} + \frac{\sum_{n=1}^{M} p_{R_1(n)}}{m} g_{R_1C(m)}\right). \tag{3.49}$$

The average cost *per user* is

$$\overline{\mathsf{cost}}_{\mathsf{CDMA}} \triangleq \frac{1}{K_{cell}M} \sum_{C} \mathsf{cost}^{C}_{\mathsf{CDMA}}. \tag{3.50}$$

In comparison, the average cost per user with CA is

$$\overline{\mathsf{cost}}_{\mathsf{CA}} \triangleq \frac{1}{K_{cell}M} \sum_{C} \sum_{m=1}^{M} \mathsf{cost}_{C(m)}. \tag{3.51}$$

.

With Algorithms 3.2 and 3.3, we compute the *average cost difference* at every iteration: $\overline{\mathsf{cost}}_{\mathsf{CDMA}} - \overline{\mathsf{cost}}_{\mathsf{CA}}$. The higher the cost difference is, the better optimality the CA achieves. We also use the lower bound on the minimum total cost (3.29), which provides an *upper bound* on the average cost difference.

Note that the average cost difference has a physical meaning of *the difference in average spectral efficiency* (averaged over all $M$ channels and $K_{cell}$ cells) in bits/sec/Hz.

**Simulation parameters and results**

Similarly to Section 4.4.3, we assume the simplified path loss model [Gol05] $P_r = P_t K(\frac{d_0}{d})^\gamma$ with $d_0 = 50m$ and the path loss exponent $\gamma = 4$ (outdoor environment), and no multipath or shadow fading. Note that the above average cost *difference* does not depend on the parameter $K$, and we let $K = 1$. We consider a one-dimensional cellular network with cell radius equal to $500m$. We assume that the number of cells $K_{cell} = 30$, (or equivalently, a two-sided infinite periodical cellular network with a period of 30 cells.)

We simulate with $M = 20$ parallel channels, and assume that the 20 users in every cell are equally spaced. Finally, the transmit power $p_{A_n}, n = 1, \ldots, M, \forall A$ can be chosen arbitrarily: In our simulation, we choose the transmit power such that all users' *received power from their own BS* are the same, (i.e. fairness in $\mathsf{revenue}_{A(m)}$.)

We plot the average cost difference curves with Algorithms 3.2 and 3.3 as in Figure 3.10, compared with the upper bound.

**Remark 18.**

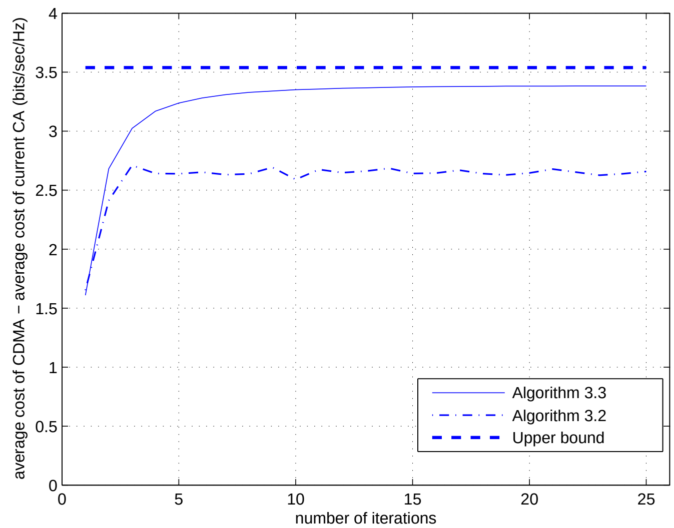



Figure 3.10: Performance of iterative algorithms 3.2, 3.3 with M=20.

- *The average cost difference is always positive, meaning that through channel allocation we achieve lower interference cost than using CDMA.*

- *Algorithm 3.3 achieves a 3.38 bits/sec/Hz better spectral efficiency than CDMA. Compared with the upper bound which is 3.53 bits/sec/Hz, it has achieved 95% of the globally optimal value. Furthermore, within* 5 iterations, *it already achieved 90% of the global optimum.*

- *As mentioned in Section 3.3.1, Algorithm 3.2 does not guarantee a monotonically increasing average cost difference. The average cost difference approximately converges (with small fluctuations) to above 2.6 bits/sec/Hz, which is about 74% of the global optimum.*

- *Both iterative channel allocation algorithms 3.2 and 3.3 achieve significantly higher spectral efficiency than CDMA.*

In general, the maximum achievable cost difference between CA and CDMA depends on channel and power parameters. The proposed iterative algorithms have exhibited a robust performance on their efficiency. In particular, Algorithm 3.3 always achieves within a few percent of the upper bound on the average cost difference.

## 3.4 Decomposition Framework with General Network and User Settings

In this section, we show that the local assignment problem (3.34), and the iterative decomposed Algorithm 3.3 apply to general wireless cellular networks, frequency selective channels, and a general set of objective functions.

### 3.4.1 Two or More Dimensional Networks with Flat Channels

In the flat channel cases, generalization to two or more dimensional networks is straightforward:

- The definition of the interference neighborhood $\mathcal{N}(A)$ applies.
- The definitions of correlated neighborhood $\mathcal{N}'(A)$ and independent set of cells $\mathcal{I}$ apply, as they are both derivatives of the definition of interference neighborhood.
- The local problem (3.34) of optimizing $C$'s CA to minimize the total cost in $C \cup \mathcal{N}(C)$, while fixing the CA in $\mathcal{N}'(C)$, applies (cf. Remark 16). The assignment problem formulation (3.46) applies.
- The iterative decomposed Algorithm 3.3 can be directly generalized: instead of having 3 maximum independent sets of cells, in the two or more

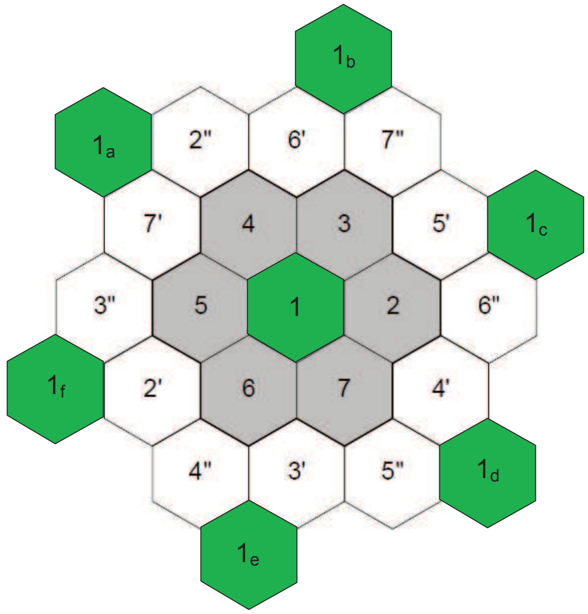



Figure 3.11: Cells $\{1, 1_a, 1_b, 1_c, 1_d, 1_e, 1_f\}$ belongs to the same maximum independent set of cells. $\mathcal{N}(1) = \{2, 3, 4, 5, 6, 7\}$. $\mathcal{N}'(1) = \mathcal{N}(1) \cup \{2', 2'', 3', 3'', 4', 4'', 5', 5'', 6', 6'', 7', 7''\}$.

dimensional case, there are more independent sets to iteratively optimize one after another. For example, in the regular hexagonal cells with the assumption that $\mathcal{N}(A)$ consists of $A$'s six immediate neighboring cells, there are 7 maximum independent sets of cells that jointly cover the entire network (cf. Figure 3.11). In each iteration, we optimize one maximum independent set, with the CA of the other six independent sets of cells fixed.

- The decomposition framework applies to both uplink and downlink channel allocation.

### 3.4.2 Frequency Selective Channels and Additively Separable Utility Functions

In frequency selective channels, we consider a general network optimization objective as follows:

$$\max \sum_{A} \sum_{m=1}^{M} U_{A(m)}(R_{A(m)}) \tag{3.52}$$

where $U_{A(m)}$ is the utility function of user $A(m)$, as a function of its achievable rate. In the form of (3.52), we have assumed that the network objective can be written as the *sum of utilities* of all users in all cells. We say that the objective has *additively separable* utility functions. We do not make any further assumption on the utility function $U_{A(m)}$.

We generalize Theorem 9 as follows:

**Corollary 13.** *For any cell $C$, given all $B(m), m = 1, 2, \ldots, M,\ B \in \mathcal{N}'(C)$,*

$$\max_{C(m)} \sum_{A \in \{C \cup \mathcal{N}(C)\}} \sum_{m=1}^{M} U_{A(m)}(R_{A(m)}) \tag{3.53}$$

*is an Assignment Problem, and can be solved with a computational complexity of $O(M^3)$.*

The proof is essentially the same as for Theorem 9, by constructing a utility matrix $\mathbf{U}$: $\forall m, n = 1, 2, \ldots M$,

$$U_{mn} = \sum_{A \in (C \cup \mathcal{N}(C))} U_{A(m)}(R_{A(m)}), \quad \textit{given that } C(m) = C_n. \tag{3.54}$$

In other words, $U_{mn}$ is the total utility in channel $m$ within $C \cup \mathcal{N}(C)$, *in case it is user $C_n$ that occupies channel $m$ in cell $C$.*

Comparing (3.53) with (3.34), the *only* difference is the utility function for the assignment problem. The problem structure is exactly the same. Therefore, the iterative decomposed Algorithm 3.3 can be directly generalized with the general assignment problem (3.53) as the building block local optimization. Combining the generalizations in network topologies discussed in section 3.4.1, the frequency selectivity, and the generalizations in the objective function (3.52), we see that the framework of *iterative decomposed network optimization based on local assignment problems* applies to all these general network and user settings.

## 3.5 Summary

In this chapter, we consider the problem of approaching the globally optimal channel allocation in large-scale wireless cellular interference networks. We show that by applying local signal scale interference alignment, the uplink channel allocation maximizing the network throughput can be achieved with a computational complexity of $O(K_{cell}M \log M)$ in one dimensional networks. With general additively separable utility functions in two or more dimensional networks, we propose a low complexity algorithmic decomposition framework on downlink and uplink channel allocation over networks with arbitrarily large sizes. In this algorithmic framework, optimization over the entire network is decomposed into local optimization that are completely decoupled due to wireless propagation losses and optimized in a distributed manner. Each local optimization is formulated as an assignment problem which can be efficiently solved. We show that an iterative algorithm that in each stage simultaneously solves decomposed local assignment problems can approach the global optimum very closely. The computational complexity of the iterative algorithm based on decomposed local assignment problems is $O(K_{cell}M^3)$.

# CHAPTER 4

# Optimal Transmissions with Successive Decoding

## 4.1 Introduction

In this chapter, we consider the sum-rate maximization problem in two-user Gaussian interference channels (cf. Figure 4.1) under the constraints of successive decoding. While the information theoretic capacity region of the Gaussian interference channel is still not known, it has been shown that a Han-Kobayashi scheme with random Gaussian codewords can achieve within 1 bit/s/Hz of the capacity region [ETW08], and hence within 2 bits/s/Hz of the sum-capacity. In this Gaussian Han-Kobayashi scheme, each user first decodes both users' common messages jointly, and then decodes its own private message. In comparison, the simplest commonly studied decoding constraint is that each user treats the interference from the other users as noise, i.e., without any decoding attempt. Using Gaussian codewords, the corresponding constrained sum-capacity problem can be formulated as a non-convex optimization of power allocation, which has an analytical solution in the two-user case [EMK06]. It has also been shown that within a certain range of channel parameters for *weak* interference channels, treating interference as noise achieves the information theoretic sum-capacity [AV09, MK09, SKC09]. For general interference channels with *more than two* users, there is so far neither a near optimal solution information theoretically, nor a polynomial time algorithm that finds a near optimal solution with interference treated as noise [LZ08] [TFL11].

We consider a decoding constraint — *successive decoding of Gaussian superposition codewords* — that bridges the complexity between joint decoding (e.g. in Han-Kobayashi schemes) and treating interference as noise. We investigate the constrained sum-capacity and its achievable schemes. Compared to treating interference as noise, allowing successive cancellation yields a much more complex problem structure. To clarify and capture the key aspects of the problem, we resort to the deterministic channel model [ADT11]. In [BT08], the information theoretic capacity region for the two-user deterministic interference channel is derived as a special case of the El Gamal-Costa deterministic model [GC82], and is shown to be achievable using Han-Kobayashi schemes.

We transmit messages using a superposition of Gaussian codebooks, and use successive decoding. To capture the use of successive decoding of Gaussian codewords, in the deterministic formulation, we introduce the *complementarity conditions* on the bit levels, which have also been characterized using a conflict graph model in [SCA10]. We develop transmission schemes on the bit-levels, which in the Gaussian model corresponds to message splitting and power allocation of the messages. We then solve the constrained sum-capacity, and show that it *oscillates* (as a function of the cross link gain parameters) between the information theoretic sum-capacity and the sum-capacity with interference treated as noise. Furthermore, the minimum number of messages needed to achieve the constrained sum-capacity is obtained. Interestingly, we show that if the number of messages is limited to even *one less* than this minimum capacity achieving number, the sum-capacity drops to that with interference treated as noise.

We then translate the optimal schemes in the deterministic channel to the Gaussian channel, using a rate constraint equalization technique. To evaluate the optimality of the translated achievable schemes, we derive and compute two upper bounds on the sum-capacity of Gaussian Han-Kobayashi schemes[1]. Since a scheme using superposition coding with Gaussian codebooks and successive decoding is a special case of Han-Kobayashi schemes, these bounds automatically apply to the sum-capacity with such successive decoding schemes as well. We select two mutually exclusive subsets of the inequality constraints that characterize the Gaussian Han-Kobayashi capacity region. Maximizing the sum-rate with each of the two subsets of inequalities leads to one of the two upper bounds. The two bounds are shown to be tight in different ranges of parameters. Numerical evaluations show that the sum-capacity with Gaussian superposition coding and successive decoding oscillates between the sum-capacity with Han-Kobayashi schemes and that with single message schemes.

The remainder of the chapter is organized as follows. Section 4.2 formulates the problem of sum-capacity with successive decoding of Gaussian superposition codewords in Gaussian interference channels, and compares it with Gaussian Han-Kobayashi schemes. Section 4.3 reformulates the problem with the deterministic channel model, and then solves the constrained sum-capacity. Section 4.4 translates the optimal schemes in the deterministic channel back to the Gaussian channel, and derives two upper bounds on the constrained sum-capacity. Numerical evaluations of the achievability against the upper bounds are provided. Section 4.5 concludes the chapter with a short discussion on generalizations of the coding-decoding assumptions and their implications.

[1] Throughout this chapter, when we refer to the Han-Kobayashi scheme, we mean the Gaussian Han-Kobayashi scheme, unless stated otherwise.

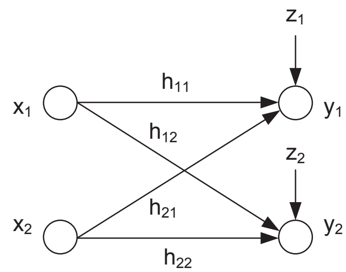



Figure 4.1: Two-user Gaussian interference channel.

## 4.2 Problem formulation in Gaussian channels

We consider the two-user Gaussian interference channel shown in Figure 4.1. The received signals of the two users are

$$y_1 = h_{11}x_1 + h_{21}x_2 + z_1,$$

$$y_2 = h_{22}x_2 + h_{12}x_1 + z_2,$$

where $\{h_{ij}\}$ are constant complex channel gains, and $z_i \sim \mathcal{CN}(0, N_i)$. Define $g_{ij} \triangleq |h_{ij}|^2, (i, j = 1, 2)$.

There is an average power constraint equal to $\bar{p}_i$ for the $i^{th}$ user ($i = 1, 2$). In the following, we first formulate the problem of finding the optimal Gaussian superposition coding and successive decoding scheme, and then provide an illustrative example to show that successive decoding schemes do not necessarily achieve the same capacity as Han-Kobayashi schemes.

### 4.2.1 Gaussian Superposition Coding and Successive Decoding: a Power and Decoding Order Optimization

Suppose the $i^{th}$ user uses a superposition of $L_i$ messages $x_i^{(\ell)} (1 \leq \ell \leq L_i)$. Denote by $r_i^{(\ell)}$ the rate of message $x_i^{(\ell)}$. For a given block length $n$, for each message $x_i^{(\ell)}$, a codebook of size $2^{nr_i^{(\ell)}}$ is generated by using IID random variables of $\mathcal{CN}(0, 1)$.

The codebooks for different messages are independently generated. For the $i^{th}$ user, the transmit signal $x_i$ is a superposition of $L_i$ Gaussian codewords, with its individual power constraint $\bar{p}_i$ satisfied, i.e.,

$$x_i = \sum_{\ell=1}^{L_i} \sqrt{p_i^{(\ell)}} x_i^{(\ell)},$$

$$\sum_{\ell=1}^{L_i} p_i^{(\ell)} \leq \bar{p}_i, \quad i = 1, 2. \tag{4.1}$$

The $i^{th}$ receiver attempts to decode all $x_i^{(\ell)}$, $\ell = 1, \ldots, L_i$, using successive decoding as follows. It chooses a decoding order $\mathcal{O}_i$ of all the $L_1 + L_2$ messages from both users. It starts decoding from the first message in this order (by treating all other messages that are not yet decoded as noise,) then peeling it off and moving to the next one, until it decodes all the messages intended for itself — $x_i^{(\ell)}$, $\ell = 1, \ldots, L_i$.

Denote the message that has order $q$ in $\mathcal{O}_i$ by $x_{t_{q,i}}^{(\ell_{q,i})}$, i.e., it is the ${\ell_{q,i}}^{th}$ message of the ${t_{q,i}}^{th}$ user. Then, the achievable rate for the successive decoding procedure to have a vanishingly small error probability as the block length $n \to \infty$ yields the following constraints on the rates of the messages:

$$r_{t_{q,i}}^{(\ell_{q,i})} \leq \log\left(1 + \frac{p_{t_{q,i}}^{(\ell_{q,i})} g_{t_{q,i}i}}{\sum_{s=q+1}^{L_1+L_2} p_{t_{s,i}}^{(\ell_{s,i})} g_{t_{s,i}i} + N_i}\right), \quad \forall 1 \leq q \leq \max_{1\leq\ell\leq L_i}\{\text{order of } x_i^{\ell} \text{ in } \mathcal{O}_i\},\ i = 1, 2. \tag{4.2}$$

Now, we can formulate the sum-rate maximization problem as:

$$\max_{\substack{\{p_i^{(\ell)}\},\mathcal{O}_i,\\ i=1,2}} \sum_{i=1}^{2} \sum_{\ell=1}^{L_i} r_i^{(\ell)} \tag{4.3}$$

$$\text{subject to: } (4.1), (4.2).$$

Note that problem (4.3) involves both a *combinatorial optimization* of the decoding orders $\{\mathcal{O}_i\}$ and a *non-convex optimization* of the transmit power $\{p_i^{(\ell)}\}$. As

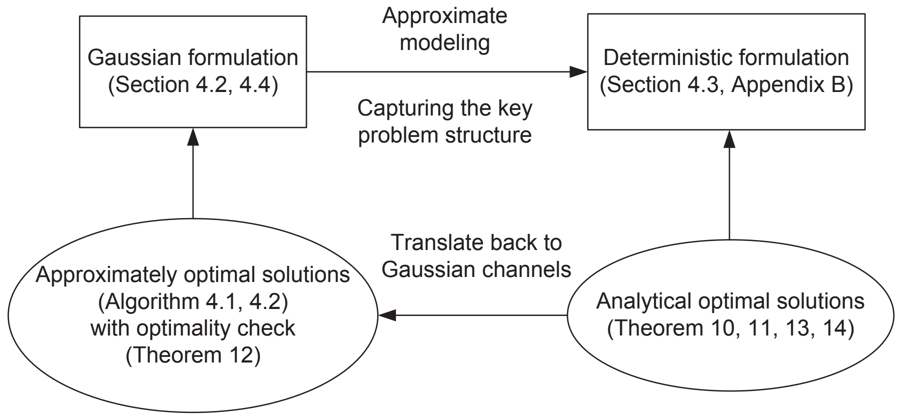



Figure 4.2: Our approach to solving problem (4.3).

a result, it is a hard problem from an optimization point of view which has not been addressed in the literature.

Interestingly, we show that an "indirect" approach can effectively and fruitfully provide approximately optimal solutions to the above problem (4.3). Instead of directly working with the Gaussian model, we approximate the problem using the recently developed deterministic channel model [ADT11]. The approximate formulation successfully captures the key structure and intuition of the original problem, for which we give a complete analytical solution that achieves the constrained sum-capacity in all channel parameters. Next, we translate this optimal solution in the deterministic formulation back to the Gaussian formulation, and show that the resulting solution is indeed close to the optimum. This indirect approach of solving (4.3) is outlined in Figure 4.2.

Next, we provide an illustration of the following point: Although the constraints for the achievable rate region with Han-Kobayashi schemes share some similarities with those for the capacity region of multiple access channels, successive decoding in interference channels does *not* always have the same achievability as Han-Kobayashi schemes, (whereas time-sharing of successive decoding schemes does achieve the capacity region of multiple access channels.)

### 4.2.2 Successive Decoding of Gaussian Codewords vs. Gaussian Han-Kobayashi Schemes with Joint Decoding

We first note that Gaussian superposition coding - successive decoding is a special case of the Han-Kobayashi scheme, using the following observations. For the $1^{st}$ user, if its message $x_1^{(\ell)}(1 \leq \ell \leq L_1)$ is *decoded* at the $2^{nd}$ receiver according to the decoding order $\mathcal{O}_2$, we categorize it into the *common* information of the $1^{st}$ user. Otherwise, $x_1^{(\ell)}$ is treated as noise at the $2^{nd}$ receiver, i.e., it appears *after* all the messages of the $2^{nd}$ user in $\mathcal{O}_2$, and we categorize it into the *private* information of the $1^{st}$ user. The same categorization is performed for the $L_2$ messages of the $2^{nd}$ user. Note that every message of the two users is either categorized as private information or common information. Thus, every successive decoding scheme is a special case of the Han-Kobayashi scheme, and hence the capacity region with successive decoding of Gaussian codewords is included in that with Han-Kobayashi schemes.

However, the inclusion in the other direction is untrue, since Han-Kobayashi schemes allow joint decoding. In the following sections, we will give a characterization of the difference between the maximum achievable sum-rate using Gaussian successive decoding schemes and that using Gaussian Han-Kobayashi schemes. This difference appears *despite* the fact that the sum-capacity of a Gaussian multiple access channel is achievable using successive decoding of Gaussian codewords. In the remainder of this section, we show an illustrative example that provides some intuition into this difference.

Suppose the $i^{th}$ user $(i = 1, 2)$ uses *two* messages: a common message $x_i^c$ and a private message $x_i^p$. We consider a power allocation to the messages, and denote the power of $x_i^c$ and $x_i^p$ by $q_i^c$ and $q_i^p$, $(i = 1, 2.)$ Denote the achievable rates of $x_i^c$ and $x_i^p$ by $r_i^c$ and $r_i^p$. In a Han-Kobayashi scheme, at each receiver, the common messages and the intended private message are *jointly* decoded, treating the unintended private message as noise. This gives rise to the achievable rate region with any given power allocation as follows:

$$r_1^c + r_1^p + r_2^c \leq \log(1 + \frac{q_1^c + q_1^p + g_{21}q_2^c}{g_{21}q_2^p + N_1}),\ r_2^c + r_2^p + r_1^c \leq \log(1 + \frac{q_2^c + q_2^p + g_{12}q_1^c}{g_{12}q_1^p + N_2}), \tag{4.4}$$

$$r_1^c + r_2^c \leq \log(1 + \frac{q_1^c + g_{21}q_2^c}{g_{21}q_2^p + N_1}),\ r_2^c + r_1^c \leq \log(1 + \frac{q_2^c + g_{12}q_1^c}{g_{12}q_1^p + N_2}), \tag{4.5}$$

$$r_1^c + r_1^p \leq \log(1 + \frac{q_1^c + q_1^p}{g_{21}q_2^p + N_1}),\ r_2^c + r_2^p \leq \log(1 + \frac{q_2^c + q_2^p}{g_{12}q_1^p + N_2}), \tag{4.6}$$

$$r_1^p + r_2^c \leq \log(1 + \frac{q_1^p + g_{21}q_2^c}{g_{21}q_2^p + N_1}),\ r_2^p + r_1^c \leq \log(1 + \frac{q_2^p + g_{12}q_1^c}{g_{12}q_1^p + N_2}), \tag{4.7}$$

$$r_1^c \leq \log(1 + \frac{q_1^c}{g_{21}q_2^p + N_1}),\ r_2^c \leq \log(1 + \frac{q_2^c}{g_{12}q_1^p + N_2}), \tag{4.8}$$

$$r_2^c \leq \log(1 + \frac{g_{21}q_2^c}{g_{21}q_2^p + N_1}),\ r_1^c \leq \log(1 + \frac{g_{12}q_1^c}{g_{12}q_1^p + N_2}), \tag{4.9}$$

$$r_1^p \leq \log(1 + \frac{q_1^p}{g_{21}q_2^p + N_1}),\ r_2^p \leq \log(1 + \frac{q_2^p}{g_{12}q_1^p + N_2}). \tag{4.10}$$

In a successive decoding scheme, depending on the different decoding orders applied, the achievable rate regions have different expressions. In the following, we provide and analyze the achievable rate region with the decoding orders at receiver 1 and 2 being $(x_1^c \to x_2^c \to x_1^p)$ and $(x_2^c \to x_1^c \to x_2^p)$ respectively. The intuition obtained with these decoding orders holds similarly for other decoding orders. With any given power allocation, we have

$$r_1^c \leq \min\left(\log(1 + \frac{q_1^c}{q_1^p + g_{21}(q_2^c + q_2^p) + N_1}), \log(1 + \frac{g_{12}q_1^c}{q_2^p + g_{12}q_1^p + N_2})\right), \tag{4.11}$$

$$r_2^c \leq \min\left(\log(1 + \frac{q_2^c}{q_2^p + g_{12}(q_1^c + q_1^p) + N_2}), \log(1 + \frac{g_{21}q_2^c}{q_1^p + g_{21}q_2^p + N_1})\right), \tag{4.12}$$

$$r_1^p \leq \log(1 + \frac{q_1^p}{g_{21}q_2^p + N_1}), r_2^p \leq \log(1 + \frac{q_2^p}{g_{12}q_1^p + N_2}). \tag{4.13}$$

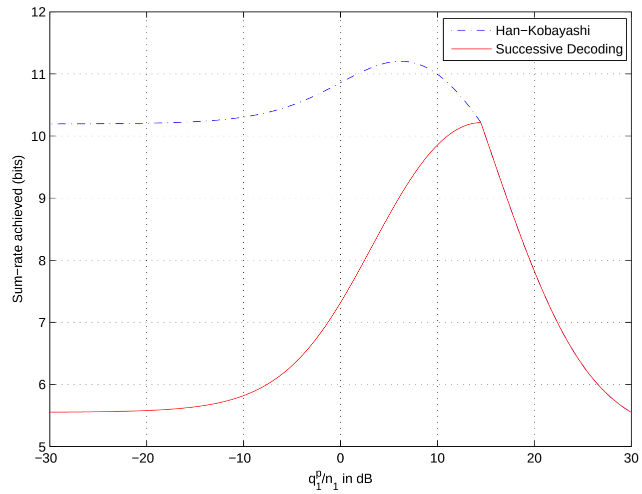



Figure 4.3: Illustrations of the difference between the achievable sum-rate with Han-Kobayashi schemes and that with successive decoding of Gaussian codewords.

It is immediate to check that (4.11) $\sim$ (4.13) $\Rightarrow$ (4.4) $\sim$ (4.10), but not vice versa.

To observe the difference between the constrained sum-capacity with (4.4) $\sim$ (4.10) and that with (4.11) $\sim$ (4.13), we examine the following symmetric channel,

$$g_{11} = g_{22} = 1, g_{12} = g_{21} = 0.17, N_1 = N_2 = 1, \tag{4.14}$$

in which we apply symmetric power allocation schemes with $q_1^c = q_2^c$ and $q_1^p = q_2^p$, and a power constraint of $\bar{p} = \bar{p}_i = q_i^p + q_i^p = 1000, i = 1, 2$.

**Remark 19.** *Note that* $\mathsf{SNR} = \frac{g_{11}\bar{p}}{N_i} = 1000 \sim 30dB, \mathsf{INR} = \frac{g_{21}\bar{p}}{N_j} = 170 \sim 22.5dB \Rightarrow \frac{\log \mathsf{INR}}{\log \mathsf{SNR}} \approx \frac{3}{4}$. *As indicated in Figure 19 of [BT08], under this parameter setting, simply using successive decoding of Gaussian codewords can have an arbitrarily large sum-capacity loss compared to joint decoding schemes, as* $\mathsf{SNR} \to \infty$.

We plot the sum-rates with the private message power $q_i^p$ sweeping from nearly zero (-30dB) to the maximum (30dB) as in Figure 4.3. As observed, the difference between the two schemes is evident when the private message power $q_i^p$ is sufficiently smaller than the common message power $q_i^c$ (with $q_i^p + q_i^c = 1000$.) The intuition of why successive decoding of Gaussian codewords is not equivalent to the Han-Kobayashi schemes is best reflected in the case of $q_i^p = 0$. In the above parameter setting, with $q_i^p = 0$, (4.4) $\sim$ (4.10) translate to

$$r_1^c + r_2^c \leq \log(1+\frac{q_1^c + g_{21}q_2^c}{N_1}) = 10.19\ bits, \tag{4.15}$$

$$r_1^c \leq \log(1+\frac{g_{12}q_1^c}{N_2}) = 7.42\ bits,\ \ r_2^c \leq \log(1+\frac{g_{21}q_2^c}{N_1}) = 7.42\ bits, \tag{4.16}$$

whereas (4.11) $\sim$ (4.13) translate to

$$r_1^c \leq \min\{\log(1+\frac{q_1^c}{g_{21}q_2^c + N_1}), \log(1+\frac{g_{12}q_1^c}{N_2})\} = \min\{2.78, 7.42\} = 2.78\ bits, \tag{4.17}$$

$$r_2^c \leq \min\{\log(1+\frac{q_2^c}{g_{12}q_1^c + N_2}), \log(1+\frac{g_{21}q_2^c}{N_1})\} = \min\{2.78, 7.42\} = 2.78\ bits. \tag{4.18}$$

As a result, the maximum achievable sum-rates with the Han-Kobayashi scheme and that with the successive decoding scheme are 10.19 bits and 5.56 bits respectively. Here, the key intuition is as follows: for a common message, its individual rate constraints at the two receivers in a successive decoding scheme (4.11), (4.12) are tighter than those in a joint decoding scheme (4.8), (4.9). In the following sections, we will see that the constraints (4.11), (4.12) lead to a non-smooth behavior of the sum-capacity using successive decoding of Gaussian codewords. Finally, we connect the results shown in Figure 4.3 to the results shown later in Figure 4.11 of Section 4.4.3:

**Remark 20.** *In Figure 4.3, the optimal symmetric power allocation for a Han-Kobayashi scheme and that for a successive decoding scheme are* $q_1^p/N_1 = 6.2dB$

*and 14.5dB respectively, leading to sum-rates of 11.2 bits and 10.2 bits. This result corresponds to the performance evaluation at* $\alpha = \frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})} = 0.75$ *in Figure 4.11.*

## 4.3 Sum-capacity in deterministic interference channels

### 4.3.1 Channel Model and Problem Formulation

In this section, we apply the deterministic channel model [ADT11] as an approximation of the Gaussian model on the two-user interference channel. We define

$$n_{11} \triangleq \log(\mathsf{SNR}_1) = \log(\frac{g_{11}\bar{p}_1}{N_1}) = \log(\tilde{g}_{11}\bar{p}_1), \tag{4.19}$$

$$n_{22} \triangleq \log(\mathsf{SNR}_2) = \log(\frac{g_{22}\bar{p}_2}{N_2}) = \log(\tilde{g}_{22}\bar{p}_2), \tag{4.20}$$

$$n_{12} \triangleq \log(\mathsf{INR}_1) = \log(\frac{g_{21}\bar{p}_2}{N_1}) = \log(\tilde{g}_{21}\bar{p}_2), \tag{4.21}$$

$$n_{21} \triangleq \log(\mathsf{INR}_2) = \log(\frac{g_{12}\bar{p}_1}{N_2}) = \log(\tilde{g}_{12}\bar{p}_1), \tag{4.22}$$

where $\tilde{g}_{ij} \triangleq g_{ij}/N_j$ are the channel gains normalized by the noise power. Without loss of generality (WLOG), we assume that $n_{11} \geq n_{22}$. Now, $n_{ji}$ counts the bit levels of the signal sent from the $i^{th}$ transmitter that are above the noise level at the $j^{th}$ receiver. Further, we define

$$\delta_1 \triangleq n_{11} - n_{21} = -\log(\frac{\tilde{g}_{12}}{\tilde{g}_{11}}),\ \ \delta_2 \triangleq n_{22} - n_{12} = -\log(\frac{\tilde{g}_{21}}{\tilde{g}_{22}}), \tag{4.23}$$

which represent the cross channel gains relative to the direct channel gains, in terms of the number of bit-level shifts. To formulate the optimization problem, we consider $\{n_{ji}\}$ to be *real* numbers. (As will be shown later in Remark 23, with *integer* bit-level channel parameters, our derivations automatically give integer bit-level optimal solutions.)

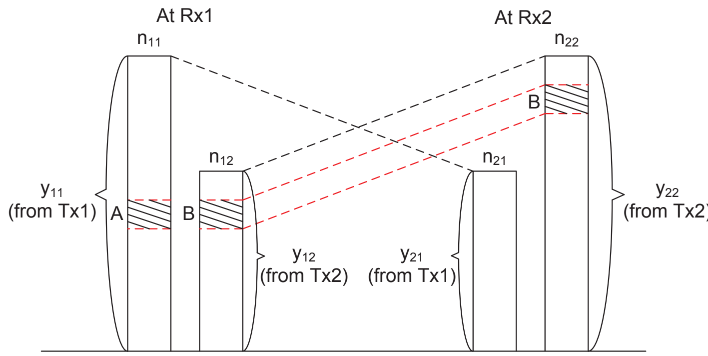



Figure 4.4: Two-user deterministic interference channel. Levels A and B interfere at the $1^{st}$ receiver, and cannot be fully active simultaneously.

In Figure 4.4, the desired signal and the interference signal at both receivers are depicted. $y_{11}$ and $y_{12}$ are the sets of received *information levels* at receiver 1 that are above the noise level, from users 1 and 2 respectively. $y_{21}$ and $y_{22}$ are the sets of received information levels at receiver 2. A more concise representation is provided in Figure 4.5:

- The sets of information levels of the *desired* signals at receivers 1 and 2 are represented by the continuous intervals $I_1 = [0, n_{11}]$ and $I_2 = [n_{11} - n_{22}, n_{11}]$ on two parallel lines, where the leftmost points correspond to the most significant (i.e., highest) information levels, and the points at $n_{11}$ correspond to the positions of the noise levels at both receivers.

- The positions of the information levels of the *interfering* signals are indicated by the dashed lines crossing between the two parallel lines.

Note that an information level (or simply termed *"level"*) is a real *point* on a line, and the measure of a set of levels (e.g. the length of an interval) equals the amount of information that this set can carry. The design variables are *whether each level of a user's received desired signal carries information for this user,* characterized by the following definition:

**Definition 20.** *$f_i(x)$ is the indicator function on whether the levels inside $I_i$ carry information for the $i^{th}$ user.*

$$f_i(x) = \begin{cases} 1, & \textit{if } x \in I_i, \textit{ and level } x \textit{ carries information for the } i^{th} \textit{ user,} \\ 0, & \textit{otherwise.} \end{cases} \quad (i = 1, 2.) \tag{4.24}$$

As a result, the rates of the two users are $R_1 = \int_0^{n_{11}} f_1(x)\mathrm{d}x, R_2 = \int_0^{n_{11}} f_2(x)\mathrm{d}x$. For an information level $x$ s.t. $f_i(x) = 1$, we call it an *active* level for the $i^{th}$ user, and otherwise an *inactive* level.

The constraints from superposition of Gaussian codewords with successive decoding (4.11) $\sim$ (4.13) translate to the following *Complementarity Conditions* in the deterministic formulation.

$$f_1(x)f_2(x + \delta_1) = 0, \forall - \infty < x < \infty, \tag{4.25}$$

$$f_2(x)f_1(x + \delta_2) = 0, \forall - \infty < x < \infty, \tag{4.26}$$

where $\delta_1$ and $\delta_2$ are defined in (4.23). The interpretation of (4.25) and (4.26) are as follows: for any two levels each from one of the two users, if they interfere with each other at any of the two receivers, they cannot be simultaneously active. For example, in Figure 4.4, information levels $A$ from the $1^{st}$ user and $B$ from the $2^{nd}$ user interfere at the $1^{st}$ receiver, and hence cannot be fully active simultaneously. These complementarity conditions have also been characterized using a conflict graph model in [SCA10].

**Remark 21.** *For any given function $f_i(x), x \in I_i$, every disjoint segment within $I_i$ with $f_i(x) = 1$ on it corresponds to a distinct message. Adjacent segments that can be so combined as a super-segment having $f_i(x) = 1$ on it, are viewed*

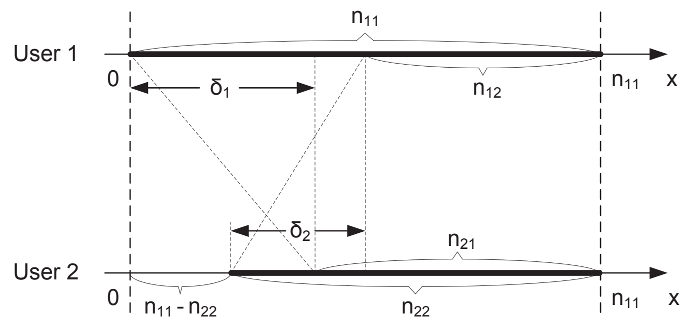



Figure 4.5: Interval representation of the two-user deterministic interference channel.

*as* one *segment, i.e., the combined super-segment. Thus, for two segments* $s_1 = [a, b] \in I_i$ *and* $s_2 = [c, d] \in I_i, (b < c,)$ *satisfying* $f_i(x) = 1, \forall x \in s_1 \cup s_2$*, if* $\exists x_0 \in (b, c), f(x_0) = 0$*, then* $s_1, s_2$ *separated by the point* $x_0$ *have to correspond to two distinct messages.*

Finally, we note that

$$(4.25) \Leftrightarrow f_2(x)f_1(x - \delta_1) = 0, \forall -\infty < x < \infty,$$
$$\text{and} \quad (4.26) \Leftrightarrow f_1(x)f_2(x - \delta_2) = 0, \forall -\infty < x < \infty.$$

Thus, we have the following result:

**Lemma 11.** *The parameter settings* $\begin{cases} \delta_1 = a \\ \delta_2 = b \end{cases}$ *and* $\begin{cases} \delta_1 = -b \\ \delta_2 = -a \end{cases}$ *correspond to the same set of complementarity conditions.*

We consider the problem of maximizing the sum-rate $R^{sum} \triangleq R_1 + R_2$ of the two users employing successive decoding, formulated as the following continuous support (infinite dimensional) optimization problem:

$$\max_{f_1(x), f_2(x)} (R^{sum} =) \int_0^{n_{11}} f_1(x) + f_2(x) \mathrm{d}x \tag{4.27}$$
$$\text{subject to } (4.24), (4.25), (4.26).$$

Problem (4.27) does not include upper bounds on the number of messages $L_1, L_2$. Such upper bounds can be added based on Remark 21. We will analyze the cases without and with upper bounds on the number of messages. We first derive the constrained sum-capacity in *symmetric* interference channels in the remainder of this section. Results are then generalized using similar approaches to *general (asymmetric)* interference channels in Appendix B.

### 4.3.2 Symmetric Interference Channels

In this section, we consider the case where $n_{11} = n_{22}, n_{12} = n_{21}$. Define $\alpha \triangleq \frac{n_{12}}{n_{11}}, \beta \triangleq 1 - \alpha$. WLOG, we normalize the amount of information levels by $n_{11}$, and consider $n_{11} = n_{22} = 1$, and $n_{12} = n_{21} = \alpha$. Note that in symmetric channels, $\beta = \delta_1 = \delta_2$.

Now, (4.25) (4.26) becomes

$$f_1(x)f_2(x+\beta) = 0, \forall -\infty < x < \infty, \tag{4.28}$$

$$f_2(x)f_1(x+\beta) = 0, \forall -\infty < x < \infty. \tag{4.29}$$

Problem (4.27) becomes

$$\max_{f_1(x),f_2(x)} (R^{sum} =) \int_0^1 f_1(x) + f_2(x)\mathrm{d}x \tag{4.30}$$
$$\text{subject to } (4.24), (4.28), (4.29).$$

From Lemma 11, *it is sufficient to only consider the case with $\beta \geq 0$, i.e. $\alpha \leq 1$.*

We next derive the constrained sum-capacity using successive decoding for $\alpha \in [0, 1]$, first without upper bounds on the number of messages, then with upper bounds. We will see that in symmetric channels, the constrained sum-capacity $R^{sum*}$ is achievable with $R_1 = R_2$. Thus, we also use the maximum

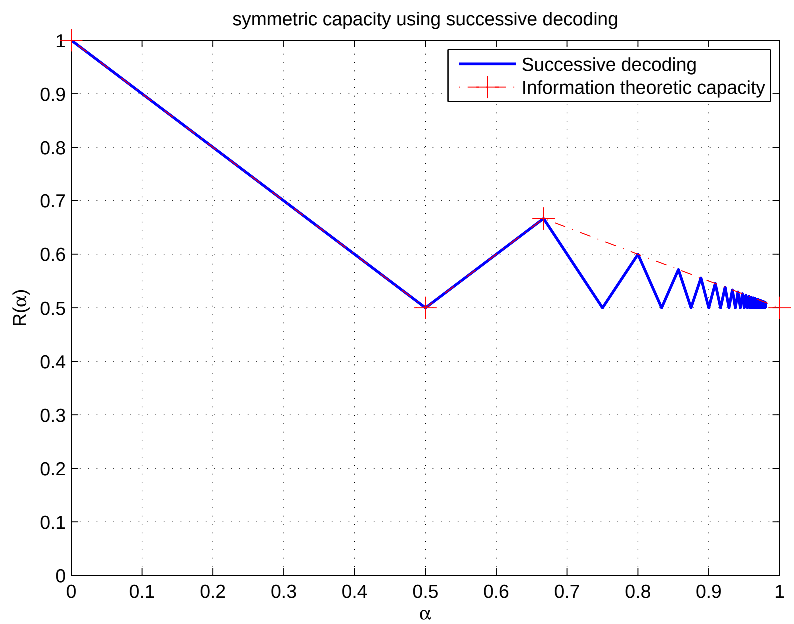



Figure 4.6: The symmetric capacity with successive decoding in symmetric deterministic interference channels.

achievable *symmetric* rate, denoted by $R(\alpha)$ as a function of $\alpha$, as an equivalent performance measure. $R(\alpha)$ is thus one half of the optimal value of (4.30).

#### 4.3.2.1 Symmetric Capacity without Constraint on the Number of Messages

**Theorem 10.** *The maximum achievable symmetric rate using successive decoding, (i.e., having constraints* (4.28), (4.29)*),* $R(\alpha)$ $(\alpha \in [0,1])$*, is characterized by*

- $R(\alpha) = 1 - \frac{\alpha}{2}$*, when* $\alpha = \frac{2n}{2n+1}, n = 0, 1, 2, \ldots$.
- $R(\alpha) = \frac{1}{2}$*, when* $\alpha = \frac{2n-1}{2n}, n = 1, 2, 3, \ldots$.

- *In every interval* $[\frac{2n}{2n+1}, \frac{2n+1}{2n+2}], n = 0, 1, 2, \ldots,$ $R(\alpha)$ *is a decreasing linear function.*

- *In every interval* $[\frac{2n-1}{2n}, \frac{2n}{2n+1}], n = 1, 2, 3, \ldots,$ $R(\alpha)$ *is an increasing linear function.*

- $R(1) = \frac{1}{2}$.

**Remark 22.** *We plot* $R(\alpha)$ *in Figure 4.6, compared with the information theoretic capacity [BT08].*

The key ideas in deriving the constrained sum-capacity are to *decompose* the effects of the complementarity conditions, such that the resulting sub-problems become easier to solve.

*Proof of Theorem 10.* i) When $\frac{2n-1}{2n} < \alpha \leq \frac{2n}{2n+1}, n = 1, 2, 3, \ldots, \frac{1}{2n+1} \leq \beta < \frac{1}{2n}$. We divide the interval $[0, 1]$ into $2n + 1$ segments $\{s_1, \ldots, s_{2n+1}\}$, where the first $2n$ segments have length $\beta$, and the last segment has length $1 - 2n\beta \in (0, \frac{1}{2n+1}]$ (cf. Figure 4.7.) With these, the complementarity conditions (4.28) (4.29) are equivalent to the following:

$$\begin{cases} \forall x \in s_1 (\Leftrightarrow x + \beta \in s_2), f_1(x) f_2(x + \beta) = 0, \\ \forall x \in s_2 (\Leftrightarrow x + \beta \in s_3), f_2(x) f_1(x + \beta) = 0, \\ \qquad \cdots \\ \forall x \in s_{2n-1} (\Leftrightarrow x + \beta \in s_{2n}), f_1(x) f_2(x + \beta) = 0, \end{cases} \tag{4.31}$$

$$\text{and} \qquad \forall x + \beta \in s_{2n+1}, \ f_2(x) f_1(x + \beta) = 0, \tag{4.32}$$

(Relations (4.31) and (4.32) correspond to the shaded strips in Figure 4.7.)

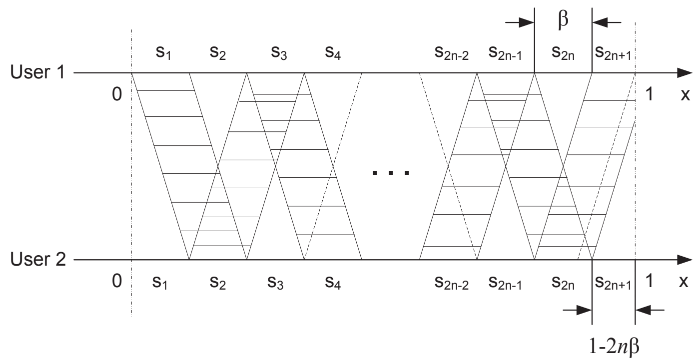



Figure 4.7: Segmentation of the information levels, $\frac{2n-1}{2n} < \alpha \leq \frac{2n}{2n+1}$.

Similarly,

$$\begin{cases} \forall x \in s_1 (\Leftrightarrow x + \beta \in s_2), f_2(x) f_1(x+\beta) = 0, \\ \forall x \in s_2 (\Leftrightarrow x + \beta \in s_3), f_1(x) f_2(x+\beta) = 0, \\ \qquad \cdots \\ \forall x \in s_{2n-1} (\Leftrightarrow x + \beta \in s_{2n}), f_2(x) f_1(x+\beta) = 0, \end{cases} \tag{4.33}$$

$$\text{and} \qquad \forall x + \beta \in s_{2n+1}, \; f_1(x) f_2(x+\beta) = 0. \tag{4.34}$$

We partition the set of all segments into two groups:

$$\mathcal{G}_1 = s_1 \cup s_3 \cup \ldots \cup s_{2n+1} \text{ and } \mathcal{G}_2 = s_2 \cup s_4 \cup \ldots \cup s_{2n}.$$

Note that

- (4.31) and (4.32) are constraints on $f_1(x)$ with support in $\mathcal{G}_1$, and on $f_2(x)$ with support in $\mathcal{G}_2$.
- (4.33) and (4.34) are constraints on $f_1(x)$ with support in $\mathcal{G}_2$, and on $f_2(x)$ with support in $\mathcal{G}_1$.

Consequently, instead of viewing the (infinite number of) optimization variables as $f_1(x)|_{[0,1]}$ and $f_2(x)|_{[0,1]}$, it is more convenient to view them as

$$C_1 \triangleq \{f_1(x)|_{\mathcal{G}_1}, f_2(x)|_{\mathcal{G}_2}\} \text{ and } C_2 \triangleq \{f_1(x)|_{\mathcal{G}_2}, f_2(x)|_{\mathcal{G}_1}\}, \tag{4.35}$$

because there is *no constraint between $C_1$ and $C_2$* from the complementarity conditions. In other words, $C_1$ and $C_2$ can be optimized *independently* of each other. Define

$$R^{sum}_{C_1} \triangleq \int_{\mathcal{G}_1} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_2} f_2(x)\mathrm{d}x,$$
$$R^{sum}_{C_2} \triangleq \int_{\mathcal{G}_2} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_1} f_2(x)\mathrm{d}x.$$

Clearly, $R^{sum} = R^{sum}_{C_1} + R^{sum}_{C_2}$. Hence (4.30) can be solved by separately solving the following two sub-problems:

$$\max_{f_1(x)|_{\mathcal{G}_1}, f_2(x)|_{\mathcal{G}_2}} (R^{sum}_{C_1} =) \int_{\mathcal{G}_1} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_2} f_2(x)\mathrm{d}x \tag{4.36}$$
$$\text{subject to } (4.24), (4.31), (4.32),$$

and

$$\max_{f_1(x)|_{\mathcal{G}_2}, f_2(x)|_{\mathcal{G}_1}} (R^{sum}_{C_2} =) \int_{\mathcal{G}_2} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_1} f_2(x)\mathrm{d}x \tag{4.37}$$
$$\text{subject to } (4.24), (4.33), (4.34).$$

We now prove that the optimal value of (4.36) is $R^{sum*}_{C_1} = 1 - n\beta$:

- (Achievability:) $1 - n\beta$ is achievable with $f_1(x) = 1, \forall x \in \mathcal{G}_1$, and $f_2(x) = 0, \forall x \in \mathcal{G}_2$.

- (Converse:) (4.31) $\Rightarrow \forall i \in \{1, 2, \ldots, n\}$, $\int_{s_{2i-1}} f_1(x)\mathrm{d}x + \int_{s_{2i}} f_2(x)\mathrm{d}x \leq \beta$

$$\Rightarrow \int_{\mathcal{G}_1} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_2} f_2(x)\mathrm{d}x = \sum_{i=1}^{n} \Big( \int_{s_{2i-1}} f_1(x)\mathrm{d}x + \int_{s_{2i}} f_2(x)\mathrm{d}x \Big) + \int_{s_{2i+1}} f_1(x)\mathrm{d}x$$
$$\leq \beta \cdot n + (1 - 2n\beta) = 1 - n\beta. \tag{4.38}$$

By symmetry, the solution of (4.37) can be obtained similarly, and the optimal value is $R_{C_2}^{sum*} = 1 - n\beta$ as well. Therefore, the optimal value of (4.30) is $R^{sum*} = 2(1 - n\beta)$.

As the above maximum achievable scheme is symmetric, i.e.,

$$f_1(x) = f_2(x) = \begin{cases} 1, & \forall x \in \mathcal{G}_1 \\ 0, & \forall x \in \mathcal{G}_2 \end{cases}, \tag{4.39}$$

the symmetric capacity is

$$R(\alpha) = 1 - n\beta = n\alpha + 1 - n. \tag{4.40}$$

Clearly, $R(\alpha)$ is an *increasing linear* function of $\alpha$ in every interval $(\frac{2n-1}{2n}, \frac{2n}{2n+1}], n = 1, 2, 3, \ldots$. It can be verified that $R(\alpha)|_{\frac{2n-1}{2n}} = \frac{1}{2}$, and $R(\alpha)|_{\frac{2n}{2n+1}} = 1 - \frac{\alpha}{2}$.

ii) When $\frac{2n}{2n+1} < \alpha \leq \frac{2n+1}{2n+2}, n = 0, 1, 2, \ldots, \frac{1}{2n+2} \leq \beta < \frac{1}{2n+1}$. Similarly to i), we divide the interval $[0, 1]$ into $2n + 2$ segments $\{s_1, \ldots, s_{2n+2}\}$, where the first $2n + 1$ segments have length $\beta$, and the last segment has length $1 - (2n + 1)\beta \in (0, \frac{1}{2n+2}]$ (cf. Figure 4.8). Then, the complementarity conditions (4.28), (4.29) are equivalent to the following:

$$(4.31), (4.32) \text{ and } f_1(x)f_2(x + \beta) = 0, \forall x + \beta \in s_{2n+2}, \tag{4.41}$$

$$\text{and } \; (4.33), (4.34) \text{ and } f_2(x)f_1(x + \beta) = 0, \forall x + \beta \in s_{2n+2}. \tag{4.42}$$

Similarly to i), with $\mathcal{G}_1 = s_1 \cup s_3 \cup \ldots \cup s_{2n+1}$ and $\mathcal{G}_2 = s_2 \cup s_4 \cup \ldots \cup s_{2n+2}$, (4.30) can be solved by separately solving the following two sub-problems:

$$\max_{f_1(x)|_{\mathcal{G}_1}, f_2(x)|_{\mathcal{G}_2}} (R_{C_1}^{sum} =) \int_{\mathcal{G}_1} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_2} f_2(x)\mathrm{d}x \tag{4.43}$$

$$\text{subject to } (4.24), (4.31), (4.32), (4.41),$$

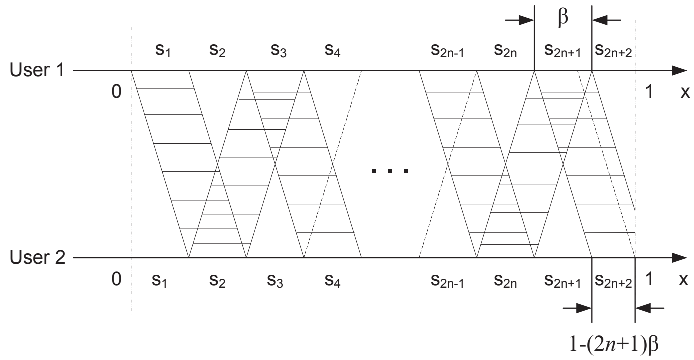



Figure 4.8: Segmentation of the information levels, $\frac{2n}{2n+1} < \alpha \leq \frac{2n+1}{2n+2}$.

and

$$\max_{f_1(x)|_{\mathcal{G}_2}, f_2(x)|_{\mathcal{G}_1}} (R_{C_2}^{sum} =) \int_{\mathcal{G}_2} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_1} f_2(x)\mathrm{d}x \tag{4.44}$$

$$\text{subject to } (4.24), (4.33), (4.34), (4.42).$$

We now prove that the optimal value of (4.43) is $(n+1)\beta$:

- (Achievability:) $(n+1)\beta$ is achievable with $f_1(x) = 1, \forall x \in \mathcal{G}_1$, and $f_2(x) = 0, \forall x \in \mathcal{G}_2$.
- (Converse:) $(4.31), (4.32), (4.41) \Rightarrow \forall i \in \{1, 2, \ldots, n+1\}$, $\int_{s_{2i-1}} f_1(x)\mathrm{d}x + \int_{s_{2i}} f_2(x)\mathrm{d}x \leq \beta$

$$\begin{aligned}\Rightarrow \int_{\mathcal{G}_1} f_1(x)\mathrm{d}x + \int_{\mathcal{G}_2} f_2(x)\mathrm{d}x =& \sum_{i=1}^{n+1} \Big( \int_{s_{2i-1}} f_1(x)\mathrm{d}x + \int_{s_{2i}} f_2(x)\mathrm{d}x \Big) \\ \leq& (n+1)\beta. \end{aligned} \tag{4.45}$$

By symmetry, the solution of (4.44) can be obtained similarly. Thus, the optimal value of (4.30) is $2(n+1)\beta$. The maximum achievable scheme is also characterized by (4.39), and the symmetric rate is

$$R(\alpha) = (n+1)\beta = -(n+1)\alpha + n + 1. \tag{4.46}$$

Clearly, $R(\alpha)$ is a *decreasing linear* function of $\alpha$ in every interval $(\frac{2n}{2n+1}, \frac{2n+1}{2n+2}], n = 0, 1, 2, \ldots$. It can be verified that $R(\alpha)|_{\frac{2n}{2n+1}} = 1 - \frac{\alpha}{2}$, and $R(\alpha)|_{\frac{2n+1}{2n+2}} = \frac{1}{2}$.

iii) It is clear that $R(0) = 1$, which is achievable with $f_1(x) = f_2(x) = 1, \forall x \in (0, 1)$, and $R(1) = \frac{1}{2}$, which is achievable by time sharing $\begin{cases} f_1(x) = 1, & x \in [0,1] \\ f_2(x) = 0, & x \in [0,1] \end{cases}$ and $\begin{cases} f_1(x) = 0, & x \in [0,1] \\ f_2(x) = 1, & x \in [0,1] \end{cases}$. □

We summarize the optimal scheme that achieves the constrained symmetric capacity as follows:

**Corollary 14.** *When $\alpha \in (0, 1)$, the constrained symmetric capacity is achievable with*

$$f_1(x) = f_2(x) = \begin{cases} 1, & \forall x \in \mathcal{G}_1 \\ 0, & \forall x \in \mathcal{G}_2 \end{cases}, \tag{4.47}$$

*where $\mathcal{G}_1 = \bigcup_{i=1,2,\ldots} s_{2i-1}$ and $\mathcal{G}_2 = \bigcup_{i=1,2,\ldots} s_{2i}$.*

In the special cases when $\alpha = \frac{2n-1}{2n}, (n = 1, 2, 3, \ldots,)$ and $\alpha = 1$, the constrained symmetric capacity drops to $\frac{1}{2}$ which is also achievable by time sharing $\begin{cases} f_1(x) = 1, & x \in [0,1] \\ f_2(x) = 0, & x \in [0,1] \end{cases}$ and $\begin{cases} f_1(x) = 0, & x \in [0,1] \\ f_2(x) = 1, & x \in [0,1] \end{cases}$.

We observe that the *numbers of messages used* by the two users — $L_1, L_2$ — in the above optimal schemes are as follows:

**Corollary 15.**

- *when $\alpha \in (\frac{2n-1}{2n}, \frac{2n+1}{2n+2}), (n = 1, 2, 3, \ldots)$, $L_1 = L_2 = n + 1$;*
- *when $\alpha \in [0, \frac{1}{2}]$, $\alpha = \frac{2n-1}{2n}, (n = 1, 2, 3, \ldots)$, or $\alpha = 1$, $L_1 = L_2 = 1$.*

**Remark 23.** *In the original formulation of the deterministic channel model [ADT11],* $\{n_{ij}\}$ *are considered to be* integers, *and the achievable scheme must also have integer bit-levels. In this case,* $\alpha = \frac{n_{12}}{n_{11}}$ *is a* rational *number. As a result, the optimal scheme* (4.47) *will consist of active segments* $\mathcal{G}_1$ *that have rational boundaries with the same denominator* $n_{11}$*. This indeed corresponds to an integer bit-level solution.*

From Theorem 10 (cf. Figure 4.6), it is interesting to see that the constrained symmetric capacity oscillates as a function of $\alpha$ between the information theoretic capacity and the baseline of 1/2. This phenomenon is a consequence of the complementarity conditions. In Section 4.5, we further discuss the connections of this result to other coding-decoding constraints.

#### 4.3.2.2 The Case with a Limited Number of Messages

In this subsection, we find the maximum achievable sum/symmetric rate using successive decoding when there are constraints on the maximum number of messages for the two users respectively. Clearly, the constrained symmetric capacity achieved with $\alpha \in [0, 1]$ will be lower than $R(\alpha)$. We start with the following two lemmas, whose proofs are relegated to Appendix A:

**Lemma 12.** *If there exists a segment with an even index* $s_{2i}$ $(i \geq 1)$ *and* $s_{2i}$ *does not end at 1, such that*

$$f_1(x) = 1, \forall x \in s_{2i}, \text{ or } f_2(x) = 1, \forall x \in s_{2i},$$

*(with* $f_i(x)$ *defined as in* (4.24)*,) then* $R^{sum} \leq 1$.

**Lemma 13.** *If there exists a segment with an odd index* $s_{2i-1}$ $(i \geq 1)$*, such that*

$$f_1(x) = 0, \forall x \in s_{2i-1}, \text{ or } f_2(x) = 0, \forall x \in s_{2i-1},$$

*then* $R^{sum} \leq 1$.

Recall that the optimal scheme (4.47) requires that, for both users, *all* segments in $\mathcal{G}_2$ are fully inactive, and *all* segments in $\mathcal{G}_1$ are fully active. The above two lemmas show the cost of violating (4.47): if one of the segments in $\mathcal{G}_2$ becomes fully active for either user (cf. Lemma 12), or one of the segments in $\mathcal{G}_1$ becomes fully inactive for either user (cf. Lemma 13), the resulting sum-rate cannot be greater than 1. We now establish the following theorem:

**Theorem 11.** *Denote by* $L_i (i = 1, 2)$ *the number of messages used by the* $i^{th}$ *user. When* $\alpha \in (\frac{2n-1}{2n}, \frac{2n+1}{2n+2}), (n = 1, 2, \ldots,)$ *if* $L_1 \leq n$ or $L_2 \leq n$*, the maximum achievable sum-rate is 1.*

*Proof.* WLOG, assume that there is a constraint of $L_1 \leq n$.

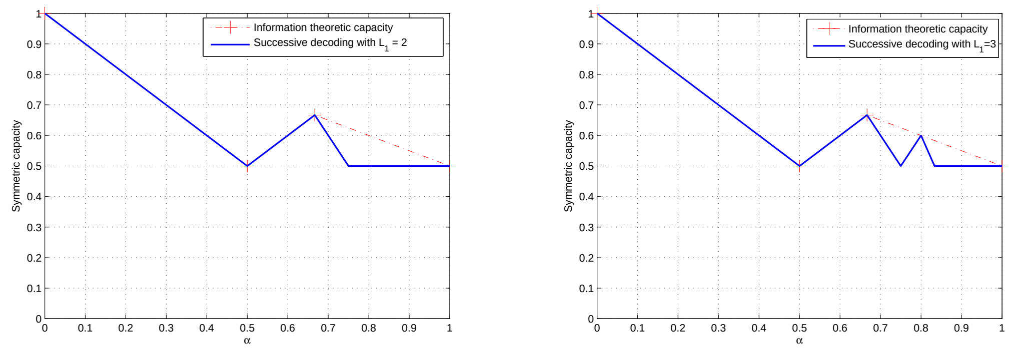
i) First, the sum-rate of 1 is always achievable with

$$f_1(x) = 1, f_2(x) = 0, \forall x \in [0, 1].$$

ii) If there exists $s_{2i}, 1 \leq i \leq n$, such that *either* $f_1(x) = 1, \forall x \in s_{2i}$, *or* $f_2(x) = 1, \forall x \in s_{2i}$, then from Lemma 12, the achieved sum-rate is no greater than 1.

iii) If *for all* $s_{2i}, 1 \leq i \leq n$, there exists $x_i$ in the interior of $s_{2i}$ such that $f_1(x_i) = 0$:

Note that $x_i$ *separates* the two segments $s_{2i-1}, s_{2i+1}$ for the $1^{st}$ user. From Remark 21, $s_{2i-1}$ and $s_{2i+1}$ have to be *two distinct messages* provided that both of them are (at least partly) active for the $1^{st}$ user. On the other hand, there are




(a) Maximum achievable symmetric rate with $L_1 \leq 2$.

(b) Maximum achievable symmetric rate with $L_1 \leq 3$.

Figure 4.9: The symmetric capacity with a limited number of messages.

$n+1$ such segments $\mathcal{G}_1 = \{s_1, s_3, \ldots, s_{2n+1}\}$ (cf. Figures 4.7 and 4.8), whereas the number of messages of the $1^{st}$ user is upper bounded by $L_1 \leq n$. Consequently, $\exists 1 \leq i_1 \leq n+1$, such that $f_1(x) = 0, \forall x \in s_{2i_1-1}$. In other words, there must be a segment in $\mathcal{G}_1$ that is fully inactive for the $1^{st}$ user. By Lemma 13, in this case, the achieved sum-rate is no greater than 1. $\square$

Comparing Theorem 11 with Corollary 15, we conclude that if the number of messages used for *either* of the two users is fewer than the number used in the optimal scheme (4.47) (as in Corollary 15), the maximum achievable symmetric rate drops to $\frac{1}{2}$. This is illustrated in Figure 4.9(a) with $L_1 \leq 2$ (or $L_2 \leq 2$), and in Figure 4.9(b) with $L_1 \leq 3$ (or $L_2 \leq 3$).

Complete solutions (without and with constraints on the number of messages) in *asymmetric* channels follow similar ideas, albeit more tediously. Detailed discussions are relegated to Appendix B.

## 4.4 Approximate Sum-capacity for Successive Decoding in Gaussian Interference Channels

In this section, we turn our focus back to the two-user Gaussian interference channel, and consider the sum-rate maximization problem (4.3). Based on the relation between the deterministic channel model and the Gaussian channel model, we *translate* the optimal solution of the deterministic channel into the Gaussian channel. We then derive upper bounds on the optimal value of (4.3), and evaluate the achievability of our translation against these upper bounds.

### 4.4.1 Achievable Sum-rate Motivated by the Optimal Scheme in the Deterministic Channel

As the deterministic channel model can be viewed as an approximation to the Gaussian channel model, optimal schemes of the former suggest approximately optimal schemes of the latter. In this subsection, we show the translation of the optimal scheme of the deterministic channel to that of the Gaussian channel. We show in detail *two forms* (simple and fine) of the translation for symmetric interference channels:

$$g_{11} = g_{22}, g_{12} = g_{21}, N_1 = N_2, \bar{p}_1 = \bar{p}_2 = \bar{p}.$$

The translation for asymmetric channels can be derived similarly, albeit more tediously.

#### 4.4.1.1 A simple translation of power allocation for the messages

Recall the optimal scheme for symmetric deterministic interference channels

(Corollary 14,) as plotted in Figure 4.10. $x_i^{(\ell)}, \ell = 1, \ldots, L$ represent the segments (or *messages* as translated to the Gaussian channel) that are active for the $i^{th}$ user. Recall that

$$-\beta = -(1-\alpha) = n_{21} - n_{11} = \log(\frac{g_{12}}{g_{11}}). \tag{4.48}$$

Thus, a shift of $\beta$ to the right (i.e. lower information levels) in the deterministic channel approximately corresponds to a power scaling factor of $\frac{g_{12}}{g_{11}}$ in the Gaussian channel. Accordingly, a simple translation of the symmetric optimal scheme (cf. Figure 4.10) into the Gaussian channel is given as follows:

*Algorithm 4.1: A simple translation by direct power scaling.*

---

Step 1: Determine the number of messages $L_1 = L_2 = L$ for each user as the same number used in the optimal deterministic channel scheme.

Step 2: Let $\frac{p^{(2)}}{p^{(1)}} = \frac{p^{(3)}}{p^{(2)}} = \ldots = \frac{p^{(L)}}{p^{(L-1)}} = \left(\frac{g_{12}}{g_{11}}\right)^2$, and normalize the power by $\sum_{\ell=1}^{L} p^{(\ell)} = \bar{p}$.

---

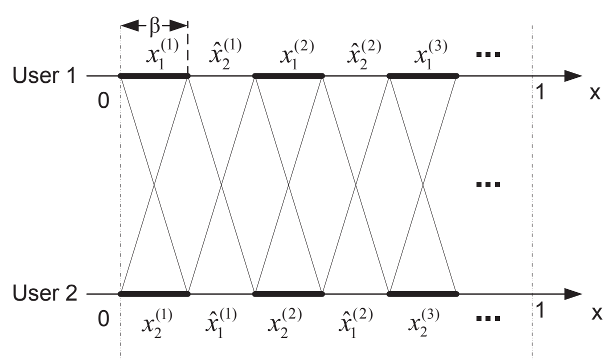



Figure 4.10: The optimal scheme in the symmetric deterministic interference channel.

#### 4.4.1.2 A finer translation of power allocation for the messages

In this part, for notational simplicity, we assume WLOG that the noise power $N_1 = N_2 = 1$ and $g_{11} = 1$. We consider the case where the cross channel gain is no greater than the direct channel gain: $0 \leq g_{12} \leq g_{11}$.

In the optimal deterministic scheme, the key property that ensures optimality is the following:

**Corollary 16.** *A message $x_i^{(\ell)}$ that is decoded at* both *receivers is subject to the* same *achievable rate constraint at both receivers.*

For example, In the optimal deterministic scheme (cf. Figure 4.10), message $x_1^{(1)}$ is subject to an achievable rate constraint of $|x_1^{(1)}|$ at the $1^{st}$ receiver, and that of $|\hat{x}_1^{(1)}|$ at the $2^{nd}$ receiver, with $|x_1^{(1)}| = |\hat{x}_1^{(1)}| = \beta$. In general, $x_1^{(1)}, \ldots, x_2^{(L-1)}$ and $x_2^{(1)}, \ldots, x_2^{(L-1)}$ are the messages that are *decoded at both receivers*, whereas $x_1^{(L)}, x_2^{(L)}$ are decoded only at their intended receiver (and treated as noise at the other receiver.)

According to Corollary 16, we show that a finer translation of the power allocation for the messages is achieved by *equalizing the two rate constraints* for every common message ($x_i^{(1)}, \ldots, x_i^{(L-1)}, i = 1, 2$). (However, rates of different common messages are not necessarily the same.)

As the $1^{st}$ step of determining the power allocations, we give the following lemma on the power allocation of $x_1^{(1)}$ (with the proof found in Appendix C):

**Lemma 14.**

*1) If $\bar{p} \leq \frac{1-g_{12}}{g_{12}^2}$, then $L = 1$, and $x_1^{(1)}(x_2^{(1)})$ is treated as noise at the $2^{nd}(1^{st})$ receiver, with $p^{(1)} = \bar{p}$. In this case, there is only one message for each user (as its private message.) rate constraint equalization is* not *needed.*

*2) If $\bar{p} > \frac{1-g_{12}}{g_{12}^2}$, then $L \geq 2$, and $x_1^{(1)}(x_2^{(1)})$ are decoded at both receivers. To* equalize *its rate constraints at both receivers, we must have*

$$p^{(1)} = 1 - g_{12} + (1 - g_{12}^2)\bar{p} \quad (< \bar{p}). \tag{4.49}$$

Next, we observe that after decoding $x_1^{(1)}, x_2^{(1)}$ at both receivers, determining $p^{(2)}$ for $x_1^{(2)}, x_2^{(2)}$ can be transformed to an equivalent $1^{st}$ step problem with $\bar{p} \leftarrow \bar{p} - p^{(1)}$: solving the new $p^{(1)}$ of the transformed problem gives the correct equalizing solution for $p^{(2)}$ of the original problem. In general, we have the following recursive algorithm in determining $L$ and $p^{(1)}, \ldots, p^{(L)}$.

*Algorithm 4.2, A finer translation by adapting L and the powers using rate constraint equalization.*

---

Initialize $L = 1$.

Step 1: If $\bar{p} \leq \frac{1-g_{12}}{g_{12}^2}$, then $p^{(L)} \leftarrow \bar{p}$ and terminate.

Step 2: $p^{(L)} \leftarrow 1 - g_{12} + (1 - g_{12}^2)\bar{p}$. $L \leftarrow L + 1$. $\bar{p} \leftarrow \bar{p} - p^{(1)}$. Go to Step 1.

---

Numerical evaluations of the above simple and finer translations of the optimal scheme of the deterministic channel into that of the Gaussian channel are provided later in Figure 4.11.

### 4.4.2 Upper Bounds on the Sum-capacity with Successive Decoding of Gaussian Codewords

In this subsection, we provide two upper bounds on the optimal solution of (4.3) for general (asymmetric) channels. More specifically, the bounds are derived for the sum-capacity with Han-Kobayashi schemes, which automatically upper bound the sum-capacity with successive decoding of Gaussian codewords (as shown in Section 4.2.2.) We will observe that the two bounds have complementary efficiencies, i.e., each being tight in a different regime of parameters.

Similarly to Section 4.2.2, we denote by $x_i^p$ the private message of the $i^{th}$ user, and $x_i^c$ the common message ($i = 1, 2$.) We denote $q_i$ to be the power allocated to each *private* message $x_i^p$, $i = 1, 2$. Then, the power of the common message $x_i^c$ equals $\bar{p}_i - q_i$. WLOG, we normalize the channel parameters such that $g_{11} = g_{22} = 1$. Denote the rates of $x_i^p$ and $x_i^c$ by $r_i^p$ and $r_i^c$. The sum-capacity of Gaussian Han-Kobayashi schemes is thus the following:

$$\begin{aligned} \max_{q_1,q_2}\ & r_1^c + r_1^p + r_2^c + r_2^p \\ & s.t.\ (4.4) \sim (4.10). \end{aligned} \tag{4.50}$$

To bound (4.50), we select two mutually exclusive *subsets* of the constraints: $\{(4.4), (4.10)\}$ and $\{(4.7)\}$. Then, with each subset of the constraints, a relaxed sum-rate maximization problems can be solved, leading to an *upper bound* to the original constrained sum-capacity (4.50).

The first upper bound on the constrained sum-capacity is as follows (whose proof is immediate from (4.4) and (4.10)):

**Lemma 15.** *The sum-capacity using Han-Kobayashi schemes is upper bounded by*

$$\begin{aligned} opt_1 \triangleq \max_{q_1,q_2}\ \min\Big\{ & \log(1+\frac{\bar{p}_1 + g_{21}(\bar{p}_2 - q_2)}{g_{21}q_2 + N_1}) + \log(1+\frac{q_2}{g_{12}q_1 + N_2}), \\ & \log(1+\frac{\bar{p}_2 + g_{12}(\bar{p}_1 - q_1)}{g_{12}q_1 + N_2}) + \log(1+\frac{q_1}{g_{21}q_2 + N_1})\ \Big\}. \end{aligned} \tag{4.51}$$

*Computation of the upper bound* (4.51). Note that

$$\log(1+\frac{\bar{p}_1+g_{21}(\bar{p}_2-q_2)}{g_{21}q_2+N_1})+\log(1+\frac{q_2}{g_{12}q_1+N_2})$$
$$=\log(c_1)-\log(g_{21}q_2+N_1)-\log(g_{12}q_1+N_2)+\log(g_{12}q_1+q_2+N_2), \tag{4.52}$$

$$\text{and}\quad \log(1+\frac{\bar{p}_2+g_{12}(\bar{p}_1-q_1)}{g_{12}q_1+N_2})+\log(1+\frac{q_1}{g_{21}q_2+N_1})$$
$$=\log(c_2)-\log(g_{12}q_1+N_2)-\log(g_{21}q_2+N_1)+\log(g_{21}q_2+q_1+N_1), \tag{4.53}$$

where $c_1 \triangleq N_1+\bar{p}_1+g_{21}\bar{p}_2, c_2 \triangleq N_2+\bar{p}_2+g_{12}\bar{p}_1$. Clearly, the minimum of (4.52) and (4.53)) is

$$-\log(g_{21}q_2+N_1)-\log(g_{12}q_1+N_2) \tag{4.54}$$
$$+\log\big(\min\{c_1(g_{12}q_1+q_2+N_2), c_2(g_{21}q_2+q_1+N_1)\}\big).$$

Now, consider the halfspace $(q_1, q_2) \in \mathcal{H}$ defined by the linear constraint

$$c_1(g_{12}q_1+q_2+N_2) \le c_2(g_{21}q_2+q_1+N_1) \Leftrightarrow (c_1g_{12}-c_2)q_1 \le (c_2g_{21}-c_1)q_2+c_2N_1-c_1N_2. \tag{4.55}$$

In $\mathcal{H}$,

$$(4.54)=\log(c_1)-\log(g_{21}q_2+N_1)-\log(g_{12}q_1+N_2)+\log(g_{12}q_1+q_2+N_2) \triangleq f(q_1,q_2). \tag{4.56}$$

Note that $\frac{\partial f(q_1,q_2)}{\partial q_1} < 0, \forall q_1 \ge 0$. Thus, depending on the sign of $c_1g_{12}-c_2$, we have the following two cases:

*Case 1:* $c_1g_{12}-c_2 \ge 0$. Then, (4.55) gives an *upper* bound on $q_1$. Consequently, to maximize (4.56), the optimal solution is achieved with $q_1 = 0$. Thus, maximizing (4.56) is equivalent to

$$\max_{q_2} \; -\log(g_{21}q_2 + N_1) + \log(q_2 + N_2) \tag{4.57}$$

$$\text{s.t.} \quad 0 \le q_2 \le \bar{p}_2, \tag{4.58}$$

in which the objective (4.57) is *monotonic*, and the solution is either $q_2 = 0$ or $q_2 = \bar{p}_2$.

*Case 2:* $c_1 g_{12} - c_2 < 0$. Then, (4.55) gives a *lower* bound on $q_1$,

$$q_1 \ge \frac{(c_1 - c_2 g_{21})q_2 + c_1 N_2 - c_2 N_1}{c_2 - c_1 g_{12}}. \tag{4.59}$$

Consequently, to maximize (4.56), the optimal solution is achieved with $q_1 = \frac{(c_1 - c_2 g_{21})q_2 + c_1 N_2 - c_2 N_1}{c_2 - c_1 g_{12}}$, which is a linear function of $q_2$. Substituting this into (4.56), we need to solve the following problem:

$$\max_{q_2} \; -\log(a_1 q_2 + b_1) - \log(a_2 q_2 + b_2) + \log(a_3 q_2 + b_3) \tag{4.60}$$

$$\text{s.t.} \quad 0 \le q_2 \le \bar{p}_2,$$

where $a_i, b_i, (i = 1, 2, 3)$ are constants determined by $c_1, c_2, g_{12}, g_{21}, N_1, N_2$. Now, (4.60) can be solved by taking the first derivative w.r.t. $q_2$, and checking the two stationary points and the two boundary points.

In the other halfspace $\mathcal{H}^c$, the same procedure as above can be applied, and the maximizer of (4.54) within $\mathcal{H}^c$ can be found. Comparing the two maximizers within $\mathcal{H}$ and $\mathcal{H}^c$ respectively, we get the global maximizer of (4.51). □

The second upper bound on the constrained sum-capacity is as follows (whose proof is immediate from (4.7)):

**Lemma 16.** *The sum-capacity using Han-Kobayashi schemes is upper bounded by*

$$opt_2 \triangleq \max_{q_1, q_2} \; \log\left(1 + \frac{q_1 + g_{21}(\bar{p}_2 - q_2)}{g_{21}q_2 + N_1}\right) + \log\left(1 + \frac{q_2 + g_{12}(\bar{p}_1 - q_1)}{g_{12}q_1 + N_2}\right). \tag{4.61}$$

*Computation of the upper bound* (4.61). Note that

$$\log\left(1+\frac{q_1+g_{21}(\bar{p}_2-q_2)}{g_{21}q_2+N_1}\right)+\log\left(1+\frac{q_2+g_{12}(\bar{p}_1-q_1)}{g_{12}q_1+N_2}\right)$$

$$= \log(q_1+g_{21}\bar{p}_2+N_1)-\log(g_{12}q_1+N_2) \quad (4.62)$$

$$+\log(q_2+g_{12}\bar{p}_1+N_2)-\log(g_{21}q_2+N_1), \quad (4.63)$$

where (4.62) is a function only of $q_1$, and (4.63) is a function only of $q_2$. Clearly, max (4.62), $s.t.\ 0 \le q_1 \le \bar{p}_1$ and max (4.63), $s.t.\ 0 \le q_2 \le \bar{p}_2$ can each be solved by taking the first order derivatives, and checking the stationary points and the boundary points. □

We combine the two upper bounds (4.51) and (4.61) as the following theorem:

**Theorem 12.** *The sum-capacity using Gaussian superposition coding-successive decoding is upper bounded by* $\min(opt_1, opt_2)$.

### 4.4.3 Performance Evaluation

We numerically evaluate our results in a symmetric Gaussian interference channel. The $\mathsf{SNR}$ is set to be $30dB$. To evaluate the performance of successive decoding, we sweep the parameter range of $\alpha = \frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})} \in [0.5, 1]$, as when $\alpha \in [0, 0.5]$, the approximate optimal transmission scheme is simply treating interference as noise without successive decoding.

In Figure 4.11, the simple translation by Algorithm 4.1 and the finer translation by Algorithm 4.2 are evaluated, and the two upper bounds derived above (4.51), (4.61) are computed. The maximum achievable sum-rate with a single message for each user ($L_1 = L_2 = 1$) is also computed, and is used as a baseline scheme for comparison.

We make the following observations:

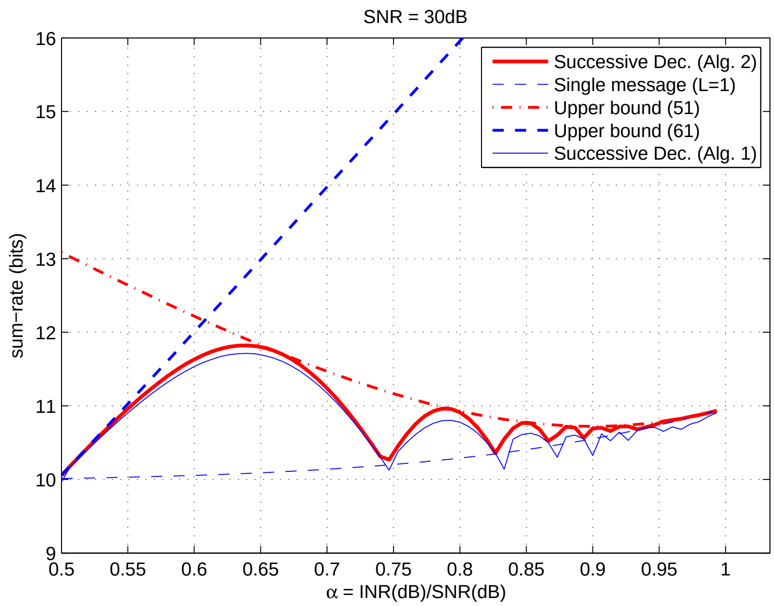



Figure 4.11: Performance evaluation: achievability vs. upper bounds.

- The finer translation of the optimal deterministic scheme by Algorithm 4.2 is strictly better than the simple translation by Algorithm 4.1, and is also strictly better than the optimal single message scheme.

- The first upper bound (4.51) is tighter for higher $\mathsf{INR}$ ($\alpha \geq 0.608$ in this example), while the second upper bound (4.61) is tighter for lower $\mathsf{INR}$ ($\alpha < 0.608$ in this example).

- A phenomenon similar to that in the deterministic channels appears: the sum-capacity with successive decoding of Gaussian codewords oscillates between the sum-capacity with Han-Kobayashi schemes and that with single message schemes.

- The largest difference between the sum-capacity of successive decoding and that of single message schemes appears at around $\frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})} = 0.64$, which is about 1.8 bits.

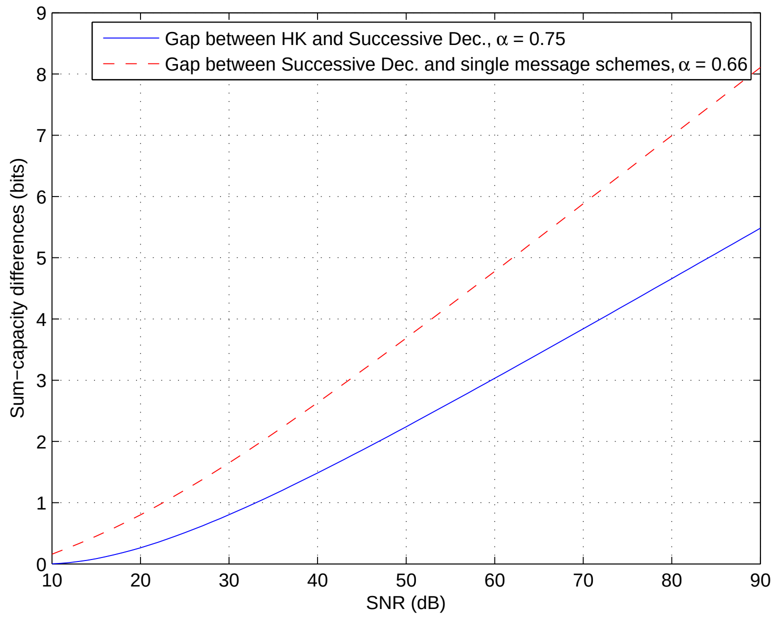



Figure 4.12: Sum-capacity differences: Han-Kobayashi vs. successive decoding at $\alpha = 0.75$, and successive decoding vs. the optimal single message scheme at $\alpha = 0.66$.

- The largest difference between the sum-capacity of successive decoding and that of joint decoding (Han-Kobayashi schemes) appears at around $\frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})} = 0.74$. This corresponds to the same parameter setting as discussed in Section 4.2.2 (cf. Figure 4.3). We see that with 30dB $\mathsf{SNR}$, this largest sum-capacity difference is about 1.0 bits.

For this particular case with $\mathsf{SNR} = 30dB$, the observed sum-capacity differences (1.8 bits and 1.0 bits) may not seem very large. However, the capacity curves shown with the deterministic channel model (cf. Figure 4.6) indicate that these differences can go to infinity as $\mathsf{SNR} \to \infty$. This is because a rate point $d_{sym}(\alpha)$ on the symmetric capacity curve in the deterministic channel has the following interpretation of *generalized degrees of freedom* in the Gaussian channel

[ETW08, BT08].

$$d_{sym}(\alpha) = \lim_{\mathsf{SNR},\mathsf{INR}\to\infty, \frac{\log \mathsf{INR}}{\log \mathsf{SNR}}=\alpha} \frac{C_{sym}(\mathsf{INR},\mathsf{SNR})}{C_{awgn}(\mathsf{SNR})}, \tag{4.64}$$

where $C_{awgn}(\mathsf{SNR}) = \log(1+\mathsf{SNR})$, and $C_{sym}(\mathsf{INR},\mathsf{SNR})$ is the symmetric capacity in the two-user symmetric Gaussian channel as a function of $\mathsf{INR}$ and $\mathsf{SNR}$.

Since $C_{awgn}(\mathsf{SNR}) \to \infty$ as $\mathsf{SNR} \to \infty$, for a fixed $\alpha$, any finite gap of the achievable rates in the deterministic channel indicates a rate gap that goes to infinity as $\mathsf{SNR} \to \infty$ in the Gaussian channel. To illustrate this, we plot the following sum-capacity differences in the Gaussian channel, with $\mathsf{SNR}$ growing from $10dB$ to $90dB$:

- The sum-capacity gap between Gaussian superposition coding - successive decoding schemes and single message schemes, with $\alpha = \frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})} = 0.66$.
- The sum-capacity gap between Han-Kobayashi schemes and Gaussian superposition coding - successive decoding schemes, with $\alpha = \frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})} = 0.75$.

As observed, the sum-capacity gaps increase asymptotically linearly with $\log \mathsf{SNR}$, and will go to infinity as $\mathsf{SNR} \to \infty$.

## 4.5 Summary

In this chapter, we studied the problem of sum-rate maximization with Gaussian superposition coding and successive decoding in two-user interference channels. This is a hard problem that involves both a combinatorial optimization of decoding orders and a non-convex optimization of power allocation. To approach this problem, we used the deterministic channel model as an educated approximation of the Gaussian channel model, and introduced the complementarity conditions that capture the use of successive decoding of Gaussian codewords. We solved the sum-capacity of the deterministic interference channel with the complementarity conditions, and obtained the capacity achieving schemes with the minimum number of messages. We showed that the constrained sum-capacity oscillates as a function of the cross link gain parameters between the information theoretic sum-capacity and the sum-capacity with interference treated as noise. Furthermore, we showed that if the number of messages used by either of the two users is fewer than its minimum capacity achieving number, the maximum achievable sum-rate drops to that with interference treated as noise. Next, we translated the optimal schemes in the deterministic channel to the Gaussian channel using a rate constraint equalization technique, and provided two upper bounds on the sum-capacity with Gaussian superposition coding and successive decoding. Numerical evaluations of the translation and the upper bounds showed that the constrained sum-capacity oscillates between the sum-capacity with Han-Kobayashi schemes and that with single message schemes.

Next, we discuss some intuitions and generalizations of the coding-decoding assumptions.

### 4.5.1 Complementarity Conditions and Gaussian Codewords

The complementarity conditions (4.25), (4.26) in the deterministic channel model has played a central role that leads to the discovered oscillating constrained sum-capacity (cf. Theorem 10). The intuition behind the complementarity conditions is as follows: At any receiver, if two active levels from different users interfere with each other, then *no* information can be recovered at this level. In other words, the *sum of interfering codewords provides nothing helpful*.

This is exactly the case when random Gaussian codewords are used in Gaussian channels with successive decoding, because the sum of two codewords from random Gaussian codebooks cannot be decoded as a valid codeword. This is the reason why the usage of Gaussian codewords with successive decoding is translated to complementarity conditions in the deterministic channels. (Note that the preceding discussions do not apply to *joint* decoding of Gaussian codewords as in Han-Kobayashi schemes.)

### 4.5.2 Modulo-2 Additions and Lattice Codes

In the deterministic channel, a relaxation on the complementarity conditions is that the *sum* of two interfering active levels can be decoded as their *modulo-2 sum*. As a result, the aggregate of two interfering codewords still provides something valuable that can be exploited to achieve higher capacity. This assumption is part of the original formulation of the deterministic channel model [ADT11], with which the information theoretic capacity of the two-user interference channel (cf. Figure 4.6 for the symmetric case) can be achieved with Han-Kobayashi schemes [BT08].

In Gaussian channels, to achieve an effect similar to decoding the modulo-2 sum with successive decoding, *lattice* codes are natural candidates of the coding schemes. This is because lattice codebooks have the group property such that the *sum* of two lattice codewords can still be decoded as a valid codeword. Such intermediate information can be decoded first and exploited later during a successive decoding procedure, in order to increase the achievable rate. For this to succeed in interference channels, *alignment of the signal scales* becomes essential [MDF11]. However, our preliminary results have shown that the ability to decode the sum of the lattice codewords does *not* provide sum-capacity increase for low and medium $\mathsf{SNR}$s. In the above setting of $\mathsf{SNR} = 30dB$ (which is typically considered as a high $\mathsf{SNR}$ in practice,) numerical computations show that the sum-capacity using successive decoding of lattice codewords with alignment of signal scales is *lower* than the previously shown achievable sum-rate using successive decoding of Gaussian codewords (cf. Figure 4.11), for the entire range of $\alpha = \frac{\log \mathsf{INR}}{\log \mathsf{SNR}} \in [0.5, 1]$. The reason is that the cost of alignment of the signal scales turns out to be higher than the benefit from it, if $\mathsf{SNR}$ is not sufficiently high. In summary, no matter using Gaussian codewords or lattice codewords, the gap between the achievable rate using successive decoding and that using joint decoding can be significant for typical $\mathsf{SNR}$s in practice.

# CHAPTER 5

# Conclusion

This thesis studies optimal transmission schemes in interference networks under a set of different problem models. Each model entails a different level of problem complexity, for which methods and algorithms are developed that effectively reduce the computational complexities in approaching the optimal solution with performance guarantees.

**The model with interference treated as noise and general Gaussian interference networks**

With the continuous frequency model, for any two (among $K$) users, if the two normalized cross channel gains between them are both larger than or equal to $\frac{1}{2}$, an FDMA allocation between them benefits every one of the $K$ users. For the classic non-convex optimization of power and spectrum management, an equivalent primal domain convex formulation is established. For piecewise flat channels, we showed that the main computational complexity lies in computing convex hull functions.

With the discrete frequency model, a provably optimal "vertical" decomposition of the spectrum management problem into channel allocation and power allocation was developed. With this vertical decomposition, the global optimum can be achieved to within a constant number of bits. If the channel allocation is globally optimal, the globally optimal power allocation can be solved by a convex optimization. Applying dual decomposition methods further reduced the computational complexity of finding the globally optimal spectrum management. This result suggests that approaching the optimal channel allocation is the essential problem, and its combinatorial complexity is what carries the NP hardness.

### The model with interference treated as noise and wireless cellular interference networks

Algorithms that approach the optimal channel allocation for arbitrarily large wireless cellular networks were proposed. For one-dimensional uplink cellular networks with flat fading channels, a two-stage channel allocation algorithm with $O(K_{cell} M \log M)$ complexity that maximizes the network throughput was found. The key idea is local signal scale interference alignment. Unfortunately, this interference alignment approach does not generalize in a low-complexity manner to more general cases. Instead, we developed a local optimization which can be formulated as an assignment problem to be solved efficiently. Using it as a building block, an iterative decomposed network optimization algorithm with $O(K_{cell} M^3)$ was developed, and was shown to very closely approach the globally maximum network throughput. This decomposition framework based on local assignment problems is applicable to very general optimization objectives and network settings. An interesting future research direction is to combine the low complexity channel allocation algorithms with power allocation algorithms (e.g., in an alternating manner,) and seek to approach the globally optimal solution for the classic non-convex joint spectrum and power optimization in large-scale wireless cellular networks.

### The model with Gaussian superposition coding - successive decoding and two-user Gaussian interference channel

We used the deterministic channel model as an educated approximation of the Gaussian channel model, and introduced the complementarity conditions that capture the use of successive decoding of Gaussian codewords. We showed that the constrained sum-capacity in the deterministic interference channel oscillates as a function of the cross link gain parameters between the information theoretic sum-capacity and the sum-capacity with interference treated as noise. Furthermore, if the number of messages used by either of the two users is fewer than the minimum number required to achieve the constrained sum-capacity, the maximum achievable sum-rate drops to that with interference treated as noise. Translating the optimal schemes in the deterministic channel back to the Gaussian channel, we showed that the constrained sum-capacity with successive decoding oscillates between the sum-capacity with Han-Kobayashi schemes and that with single message schemes. Based on our results on successive decoding capacity in the two-user interference channel, it is an interesting future research direction to apply the insight into networks with larger sizes and special structures (e.g. cellular networks,) and to further evaluate how much gain from message splitting and successive decoding can be obtained in practical scenarios.

In general, optimization in interference networks is fundamentally hard. In this thesis, we illustrated a trade-off between approaching the network optimality and pursuing practical solutions with low implementation complexity. Having constructed a variety of levels of abstraction in the problem model, we developed capacity approaching schemes of reasonable complexity with each level of abstraction.

# APPENDIX A

## Proof of Lemma 12 and 13

*Proof of Lemma 12.* By symmetry, it is sufficient to prove for the case $f_2(x) = 1, \forall x \in s_{2i}$, for some $s_{2i}$ that does not end at 1.

Now, consider the sum-rate achieved within $C_1$ (4.35). As shown in Figure A.1, $C_1$ can be partitioned into three parts: $C_{11} = \{f_1(x)|_{s_1,s_3,\ldots,s_{2i-3}}, f_2(x)|_{s_2,s_4,\ldots,s_{2i-2}}\}$, $C_{12} = \{f_1(x)|_{s_{2i-1},s_{2i+1}}, f_2(x)|_{s_{2i}}\}$, and $C_{13} = \{f_1(x)|_{s_{2i+3},\ldots}, f_2(x)|_{s_{2i+2},\ldots}\}$, ($C_{11}, C_{12}, C_{13}$ can be degenerate.) Note that

- From the achievable schemes in the proof of Theorem 10, the maximum achievable sum-rate within $C_{11} \cup C_{13}$ can be achieved with $f_2(x) = 1, \forall x \in s_2 \cup s_4 \cup \ldots \cup s_{2i-2} \cup s_{2i+2} \cup \ldots$, and $f_1(x) = 0, \forall x \in s_1 \cup s_3 \cup \ldots \cup s_{2i-3} \cup s_{2i+3} \cup \ldots$.
- By the assumed condition, $f_2(x) = 1, \forall x \in s_{2i} \Rightarrow f_1(x) = 0, \forall x \in s_{2i-1} \cup s_{2i+1}$.

Therefore, under the assumed condition, the maximum achievable sum-rate within $C_1$ is achievable with $\{f_2(x) = 1, \forall x \in \mathcal{G}_2$, and $f_1(x) = 0, \forall x \in \mathcal{G}_1\}$.

Furthermore, from the proof of Theorem 10, we know that the maximum achievable sum-rate within $C_2$ is achievable with $\{f_2(x) = 1, \forall x \in \mathcal{G}_1$, and $f_1(x) = 0, \forall x \in \mathcal{G}_2\}$. Combining the maximum achievable schemes within $C_1$ and $C_2$, by letting $\{f_2(x) = 1, \forall x \in [0,1]$, and $f_1(x) = 0, \forall x \in [0,1]\}$, a sum-rate of 1

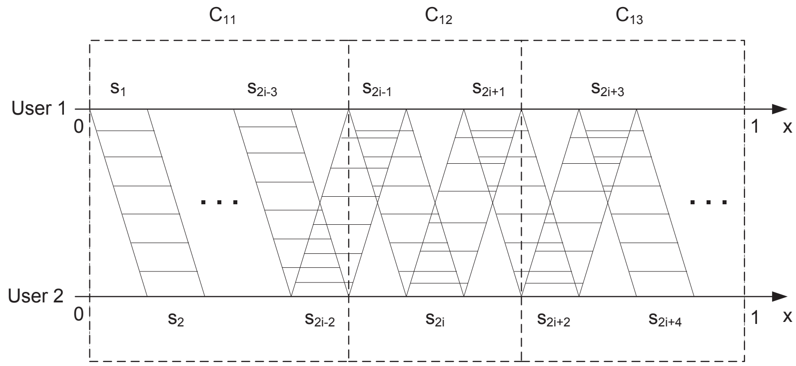



Figure A.1: $C_1$ partitioned into three parts for Lemma 12.

is achieved, and this is the maximum achievable sum-rate given the assumed condition. □

*Proof of Lemma 13.* By symmetry, it is sufficient to prove for the case $f_1(x) = 0, \forall x \in s_{2i-1}$, for some $s_{2i-1}$.

Now, consider the sum-rate achieved within $C_1$. As shown in Figure A.2, $C_1$ can be partitioned into three parts: $C_{11} = \{f_1(x)|_{s_1,s_3,\ldots,s_{2i-3}}, f_2(x)|_{s_2,s_4,\ldots,s_{2i-2}}\}$, $C_{12} = f_1(x)|_{s_{2i-1}}$, and $C_{13} = \{f_1(x)|_{s_{2i+1},s_{2i+3},\ldots}, f_2(x)|_{s_{2i},s_{2i+2},\ldots}\}$, ($C_{11}, C_{12}, C_{13}$ can be degenerate.) Note that:

- From the achievable schemes in the proof of Theorem 10, the maximum achievable sum-rate within $C_{11} \cup C_{13}$ can be achieved with $f_2(x) = 1, \forall x \in s_2 \cup s_4 \cup \ldots \cup s_{2i-2} \cup s_{2i} \cup \ldots$, and $f_1(x) = 0, \forall x \in s_1 \cup s_3 \cup \ldots \cup s_{2i-3} \cup s_{2i+1} \cup \ldots$.

- By the assumed condition, $f_1(x) = 0, \forall x \in s_{2i-1}$.

Therefore, under the assumed condition, the maximum achievable sum-rate within $C_1$ is achievable with $\{f_2(x) = 1, \forall x \in \mathcal{G}_2$, and $f_1(x) = 0, \forall x \in \mathcal{G}_1\}$.

Furthermore, from the proof of Theorem 10, we know that the maximum achievable sum-rate within $C_2$ is achievable with $\{f_2(x) = 1, \forall x \in \mathcal{G}_1$, and $f_1(x) =$

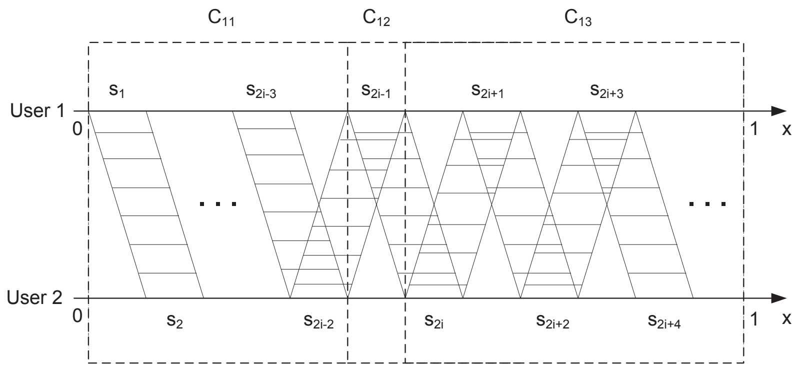



Figure A.2: $C_1$ partitioned into three parts for Lemma 13.

$0, \forall x \in \mathcal{G}_2\}$. Combining the maximum achievable schemes within $C_1$ and $C_2$, by letting $\{f_2(x) = 1, \forall x \in [0,1]$, and $f_1(x) = 0, \forall x \in [0,1]\}$, a sum-rate 1 is achieved, and this is the maximum achievable sum-rate given the assumed condition. □

# APPENDIX B

# Sum-capacity of Deterministic Asymmetric Interference Channels

We consider the general two-user interference channel where the parameters $n_{11}, n_{22}, n_{12}, n_{21}$ can be arbitrary. Still, WLOG, we make the assumptions that $n_{11} \geq n_{22}$ and $n_{11} = 1$. We will see that our approaches in the symmetric channel can be similarly extended to solving the constrained sum-capacity in asymmetric channels, without and with constraints on the number of messages.

From Lemma 11, it is sufficient to consider the following three cases:

$$\text{i) } \delta_1 \geq 0 \text{ and } \delta_2 \geq 0; \quad \text{ii) } \delta_1 \geq 0 \text{ and } \delta_2 < 0; \quad \text{iii) } \delta_1 < 0 \text{ and } \delta_2 \geq 0. \qquad \text{(B.1)}$$

## B.1 Sum-Capacity without Constraint on the Number of Messages

We provide the optimal scheme that achieves the constrained sum-capacity in each of the three cases in (B.1), respectively.

### B.1.1 $\delta_1 \geq 0$ and $\delta_2 \geq 0$

This is by definition (4.23) equivalent to $n_{21} \leq 1$ and $n_{22} \geq n_{12}$.

*Case 1*, $n_{22} \geq n_{21}$:

Define $\beta_1 \triangleq 1 - n_{12}, \beta_2 \triangleq n_{22} - n_{21}$. As depicted in Figure B.1, interval $I_1 (= [0, 1])$ is partitioned into segments $\{s_1, s_2, s_3, \ldots\}$, with $|s_1| = |s_3| = \ldots = \beta_1$ and $|s_2| = |s_4| = \ldots = \beta_2$; the last segment ending at 1 has the length of the proper residual. Interval $I_2 (= [1 - n_{22}, 1])$ is partitioned into segments $\{s'_1, s'_2, s'_3 \ldots\}$, with $|s'_1| = |s'_3| = \ldots = \beta_2$ and $|s'_2| = |s'_4| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual.

Similarly to (4.35) as in the previous analysis for the symmetric channels, we partition the optimization variables $f_1(x)|_{[0,1]}$ and $f_2(x)|_{[1-n_{22},1]}$ into

$$C_1 \triangleq \{f_1(x)|_{s_1,s_3,\ldots}, f_2(x)|_{s'_2,s'_4,\ldots}\} \text{ and } C_2 \triangleq \{f_1(x)|_{s_2,s_4,\ldots}, f_2(x)|_{s'_1,s'_3,\ldots}\}. \quad \text{(B.2)}$$

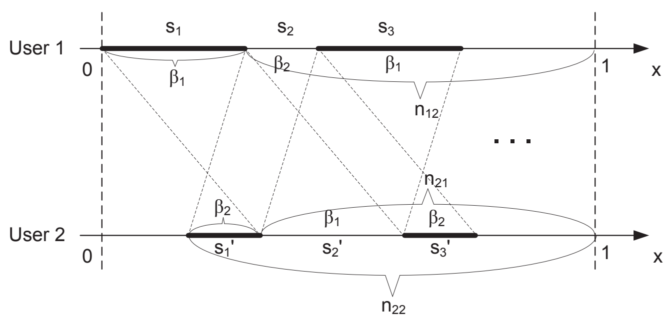



Figure B.1: $n_{11} \geq n_{21}$, $n_{22} \geq n_{12}$, and $n_{22} \geq n_{21}$.

As there is *no constraint between* $C_1$ *and* $C_2$ from the complementarity conditions (4.25) and (4.26), similarly to (4.36) and (4.37), the sum-rate maximization (4.27) is decomposed into two separate problems:

$$\max_{f_1(x)|_{s_1,s_3,\ldots}, f_2(x)|_{s'_2,s'_4,\ldots}} (R^{sum}_{C_1} =) \int_{s_1,s_3,\ldots} f_1(x)\mathrm{d}x + \int_{s'_2,s'_4,\ldots} f_2(x)\mathrm{d}x \quad \text{(B.3)}$$

$$\text{subject to } (4.24), (4.25), (4.26),$$

$$\max_{f_1(x)|_{s_2,s_4,\ldots}, f_2(x)|_{s'_1,s'_3,\ldots}} (R^{sum}_{C_2} =) \int_{s_2,s_4,\ldots} f_1(x)\mathrm{d}x + \int_{s'_1,s'_3,\ldots} f_2(x)\mathrm{d}x \quad \text{(B.4)}$$

$$\text{subject to } (4.24), (4.25), (4.26).$$

By the same argument as in the proof of Theorem 10, the optimal solution of (B.3) is given by

$$f_1(x) = 1, \forall x \in s_1 \cup s_3 \cup \ldots, \text{ and } f_2(x) = 0, \forall x \in s'_2 \cup s'_4 \cup \ldots. \tag{B.5}$$

Also, the optimal solution of (B.4) is given by

$$f_1(x) = 0, \forall x \in s_2 \cup s_4 \cup \ldots, \text{ and } f_2(x) = 1, \forall x \in s'_1 \cup s'_3 \cup \ldots. \tag{B.6}$$

Consequently, we have the following theorem:

**Theorem 13.** *A constrained sum-capacity achieving scheme is given by*

$$f_1(x) = \begin{cases} 1, & \forall x \in s_1 \cup s_3 \cup \ldots \\ 0, & otherwise \end{cases}, \text{ and } f_2(x) = \begin{cases} 1, & \forall x \in s'_1 \cup s'_3 \cup \ldots \\ 0, & otherwise \end{cases}, \tag{B.7}$$

*and the maximum achievable sum-rate is readily computable based on* (B.7).

*Case 2*, $n_{21} > n_{22}$:

Define $\beta_1 \triangleq 1 - n_{12} - (n_{21} - n_{22})$. As depicted in Figure B.2, interval $I_1(= [0, 1])$ is partitioned into segments $\{s_0, s_1, s_3, s_5, \ldots\}$, with $|s_0| = n_{21} - n_{22}$, and $|s_1| = |s_3| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual. Interval $I_2(= [1 - n_{22}, 1])$ is partitioned into segments $\{s'_2, s'_4 \ldots\}$, with $|s'_2| = |s'_4| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual. (The indexing is not consecutive as we consider $\{s_{2i}\}$ and $\{s'_{2i-1}\}$ $(i \geq 1)$ as degenerating to empty sets.)

Clearly, $s_0$ of $I_1$ does not conflict with any levels of $I_2$, and thus we let $f_1(x) = 1, \forall x \in s_0$. On all the other segments, the sum-rate maximization problem is

$$\max_{f_1(x)|_{s_1, s_3, \ldots}, f_2(x)|_{s'_2, s'_4, \ldots}} \int_{s_1, s_3, \ldots} f_1(x) \mathrm{d}x + \int_{s'_2, s'_4, \ldots} f_2(x) \mathrm{d}x \tag{B.8}$$

$$\text{subject to } (4.24), (4.25), (4.26).$$

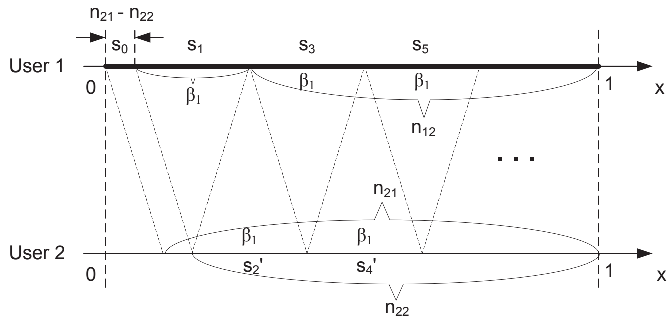



Figure B.2: $n_{11} \geq n_{21}$, $n_{22} \geq n_{12}$, and $n_{21} > n_{22}$.

By the same argument as in the proof of Theorem 10, the optimal solution of (B.8) is given by

$$f_1(x) = 1, \forall x \in s_1 \cup s_3 \cup \ldots, \text{ and } f_2(x) = 0, \forall x \in s_2' \cup s_4' \cup \ldots.$$

Thus, a sum-capacity achieving scheme is simply $f_1(x) = 1, \forall x \in I_1$, and $f_2(x) = 0, \forall x \in I_2$.

### B.1.2 $\delta_1 \geq 0$ and $\delta_2 < 0$

This is by definition (4.23) equivalent to $n_{21} \leq 1$ and $n_{22} < n_{12}$. Note that by Lemma 11, it is sufficient to only consider the case where $|\delta_1| \geq |\delta_2|$, (because in case $|\delta_1| < |\delta_2|$, we have $|-\delta_2| > |-\delta_1|$.)

*Case 1*, $n_{22} \geq n_{21}$, and $n_{12} > 1$:

Define $\beta_1 \triangleq n_{22} - n_{21} - (n_{12} - 1)$. As depicted in Figure B.3, interval $I_1(= [0,1])$ is partitioned into segments $\{s_1, s_3, \ldots\}$, with $|s_1| = |s_3| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual. Interval $I_2(= [1 - n_{22}, 1])$ is partitioned into segments $\{s_0', s_2', s_4' \ldots\}$, with $|s_0'| = n_{12} - 1$ and $|s_2'| = |s_4'| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual.

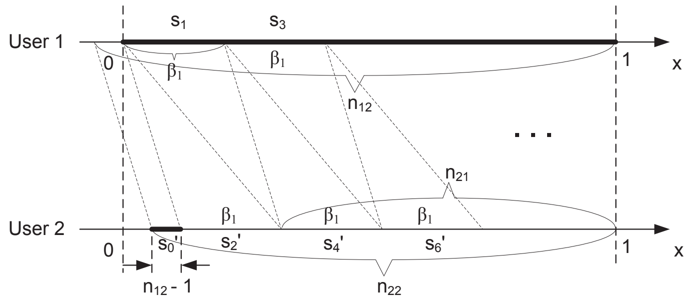



Figure B.3: $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, $n_{22} \geq n_{21}$, and $n_{12} > n_{11}$.

Clearly, $s'_0$ of $I_2$ does not conflict with any levels of $I_1$, and thus we let $f_2(x) = 1, \forall x \in s'_0$. On all the other segments, the sum-rate maximization problem is again (B.8), and the optimal solution is given by

$$f_1(x) = 1, \forall x \in s_1 \cup s_3 \cup \ldots, \text{ and } f_2(x) = 0, \forall x \in s'_2 \cup s'_4 \cup \ldots.$$

Thus, a sum-capacity achieving scheme is $f_1(x) = 1, \forall x \in I_1$, and $f_2(x) = \begin{cases} 1, & \forall x \in s'_0 \\ 0, & otherwise \end{cases}$.

*Case 2*, $n_{22} \geq n_{21}$, and $n_{12} \leq 1$:

Define $\beta_1 \triangleq 1 - n_{12}, \beta_2 \triangleq n_{22} - n_{21}$. As depicted in Figure B.4, interval $I_1(= [0, 1])$ is partitioned into segments $\{s_1, s_2, s_3, \ldots\}$, with $|s_1| = |s_3| = \ldots = \beta_1$ and $|s_2| = |s_4| = \ldots = \beta_2$; the last segment ending at 1 has the length of the proper residual. Interval $I_2(= [1 - n_{22}, 1])$ is partitioned into segments $\{s'_1, s'_2, s'_3 \ldots\}$, with $|s'_1| = |s'_3| = \ldots = \beta_2$ and $|s'_2| = |s'_4| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual.

Compare with Case 1 of Section B.1.1 and note the similarities between Figure B.4 and Figure B.1: we apply the same partition of the optimization variables (B.2), and the sum-rate maximization (4.27) is decomposed in the same way into two separate problems (B.3) and (B.4). However, while the optimal solution of

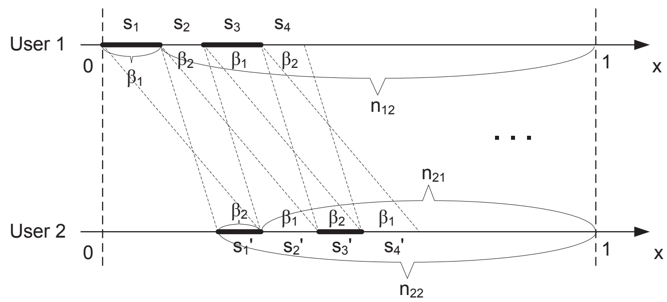



Figure B.4: $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, $n_{22} \geq n_{21}$, and $n_{12} \leq n_{11}$, scheme I (non-optimal).

(B.3) is still given by (B.5), *the optimal solution of* (B.4) *is no longer given by* (B.6). Instead, as $\delta_2 < 0$, the optimal solution of (B.4) is given by

$$f_1(x) = 1, \forall x \in s_2 \cup s_4 \cup \ldots, \text{ and } f_2(x) = 0, \forall x \in s'_1 \cup s'_3 \cup \ldots.$$

Thus, a sum-capacity achieving scheme is given by $f_1(x) = 1, \forall x \in I_1$, and $f_2(x) = 0, \forall x \in I_2$, depicted as in Figure B.5.

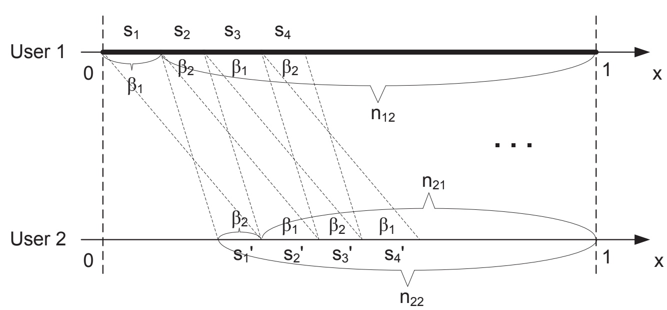



Figure B.5: $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, $n_{22} \geq n_{21}$, and $n_{12} \leq n_{11}$, scheme II (optimal).

*Case 3*, $n_{22} < n_{21}$:

Compare with Case 2 of B.1.1 (cf. Figure B.2), with the same definition of $\beta_1$ and the same partition of $I_1$ and $I_2$, the segmentation is depicted in Figure B.6.

Noting the similarities between Figure B.2 and Figure B.6, we see that the

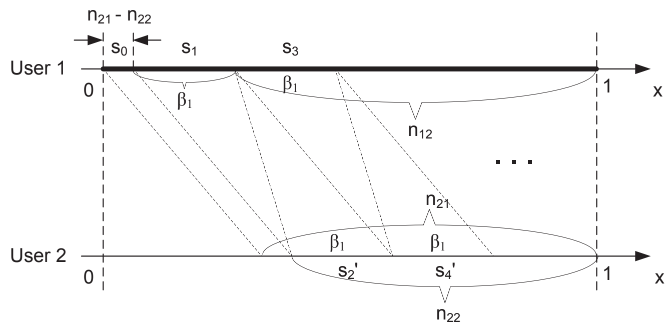



Figure B.6: $n_{11} \geq n_{21}$, $n_{22} < n_{12}$, and $n_{22} < n_{21}$.

optimal solution of the two cases are the same: $f_1(x) = 1, \forall x \in I_1$, and $f_2(x) = 0, \forall x \in I_2$.

### B.1.3 $\delta_1 < 0$ and $\delta_2 \geq 0$

This is by definition (4.23) equivalent to $n_{21} > 1$ and $n_{22} \geq n_{12}$. Note that by Lemma 11, it is sufficient to only consider the case where $|\delta_1| \leq |\delta_2|$, (because in case $|\delta_1| > |\delta_2|$, we have $|-\delta_2| \leq |-\delta_1|$.)

Define $\beta_1 \triangleq 1 - n_{12} - (n_{21} - n_{22})$. As depicted in Figure B.7, interval $I_1(= [0,1])$ is partitioned into segments $\{s_0, s_1, s_3, s_5, \ldots\}$, with $|s_0| = n_{21} - n_{22}$ and $|s_1| = |s_3| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual. Interval $I_2(= [1 - n_{22}, 1])$ is partitioned into segments $\{s_2', s_4' \ldots\}$, with $|s_2'| = |s_4'| = \ldots = \beta_1$; the last segment ending at 1 has the length of the proper residual.

Clearly, $s_0$ of $I_1$ does not conflict with any levels of $I_2$, and thus we let $f_1(x) = 1, \forall x \in s_0$. On all the other segments, the sum-rate maximization problem is again (B.8). As $\delta_1 < 0$, the optimal solution is given by

$$f_1(x) = 0, \forall x \in s_1 \cup s_3 \cup \ldots, \text{ and } f_2(x) = 1, \forall x \in s_2' \cup s_4' \cup \ldots.$$

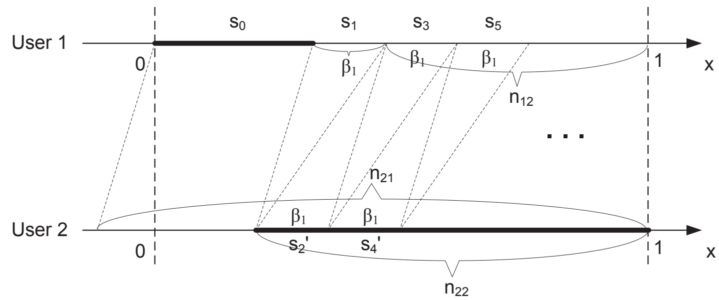



Figure B.7: $n_{11} < n_{21}$ and $n_{22} \geq n_{12}$.

Thus, a sum-capacity achieving scheme is $f_1(x) = \begin{cases} 1, & \forall x \in s_0 \\ 0, & otherwise \end{cases}$, and $f_2(x) = 1, \forall x \in I_2$.

Summarizing the discussions of the six parameter settings (cf. Figures B.1 - B.3 and B.5 - B.7) in this section, we observe:

**Remark 24.** *Except for Case 1 of Section B.1.1, the optimal schemes for the other cases all have the property that* only one message is used for each user.

## B.2 The Case with a Limited Number of Messages

In this section, we extend the sum-capacity results in Section 4.3.2.2 to the asymmetric channels when there are upper bounds on the number of messages $L_1, L_2$ for the two users respectively. From Remark 24, we only need to discuss Case 1 of Section B.1.1 (cf. Figure B.1,) with its corresponding notations.

Similarly to the symmetric channels, we generalize Lemma 12 and 13 to the following two lemmas for the general (asymmetric) channels, whose proofs are exact parallels to those of Lemma 12 and 13:

**Lemma 17.**

*1. If* $\exists s_{2i}$*,* $s_{2i}$ *does not end at 1, such that* $f_1(x) = 1, \forall x \in s_{2i}$*, then* $R^{sum} \leq 1$.

*2. If* $\exists s'_{2i}$*,* $s'_{2i}$ *does not end at 1, such that* $f_2(x) = 1, \forall x \in s'_{2i}$*, then* $R^{sum} \leq n_{22}$.

**Lemma 18.**

*1. If* $\exists s_{2i-1}$*, such that* $f_1(x) = 0, \forall x \in s_{2i-1}$*, then* $R^{sum} \leq n_{22}$.

*2. If* $\exists s'_{2i-1}$*, such that* $f_2(x) = 0, \forall x \in s'_{2i-1}$*, then* $R^{sum} \leq 1$.

We then have the following generalization of Theorem 11 to the general (asymmetric) channels:

**Theorem 14.** *Denote by* $L_i$ *the number of messages used by the* $i^{th}$ *user in any scheme, and denote by* $n_i$ *the dictated number of messages used by the* $i^{th}$ *user in the constrained sum-capacity achieving scheme* (B.7)*. Then, if* $L_1 \leq n_1 - 1$ or $L_2 \leq n_2 - 1$*, we have* $R^{sum} \leq 1$.

*Proof.* Consider $L_2 \leq n_2 - 1$. (The case of $L_1 \leq n_1 - 1$ can be proved similarly.)

i) The sum-rate of 1 is always achievable with

$$f_1(x) = 1, \forall x \in I_1, f_2(x) = 0, \forall x \in I_2.$$

ii) If there exists $s'_{2i}, (i \geq 1)$ and $s'_{2i}$ does not end at 1, such that $f_2(x) = 1, \forall x \in s'_{2i}$, then from Lemma 17, $R^{sum} \leq n_{22} \leq 1$.

iii) If *for every* $s'_{2i}, i \geq 1$ and $s'_{2i}$ does not end at 1, there exists $x_i$ in the interior of $s'_{2i}$ such that $f_2(x_i) = 0$:

For every $x_i$, since $s'_{2i}$ does not end at 1, $s'_{2i+1}$ exists. Note that $x_i$ *separates* the two segments $s'_{2i-1}, s'_{2i+1}$ for the $2^{nd}$ user. From Remark 21, $s'_{2i-1}$ and $s'_{2i+1}$ have to be *two distinct messages* provided that both of them are (at least partly) active for the $2^{nd}$ user. On the other hand, there are $n_2$ such segments $\{s'_1, s'_3, \ldots, s'_{2n_2-1}\}$, whereas the number of messages is upper bounded by $L_2 \leq n_2 - 1$. Consequently, $\exists 1 \leq i_2 \leq n_2$, such that $f_2(x) = 0, \forall x \in s_{2i_2-1}$. In other words, for the $2^{nd}$ user, there must be a segment with an odd index that is *fully inactive*. By Lemma 18, in this case, $R^{sum} \leq 1$. □

Similarly to the symmetric case, we conclude that if the number of messages used for *either* user is fewer than the number used in the optimal scheme (B.7), the maximum achievable sum-rate drops to 1.

# APPENDIX C

# Proof of Lemma 14

At the $1^{st}$ receiver, the message $x_1^{(1)}$ is decoded by treating all other messages $(x_1^{(2)}, \ldots, x_1^{(L)}, x_2^{(1)}, \ldots, x_2^{(L)})$ as noise, and has an $\mathsf{SNR}_1$ of $\frac{p^{(1)}}{(\bar{p}-p^{(1)})+g_{21}\bar{p}+1}$.

At the $2^{nd}$ receiver, $x_2^{(1)}$ is first decoded and peeled off. Suppose $x_1^{(1)}$ is also decoded at the $2^{nd}$ receiver (by treating $x_1^{(2)}, \ldots, x_1^{(L)}$, $x_2^{(2)}, \ldots, x_2^{(L)}$ as noise,) it has an $\mathsf{SNR}_2$ of $\frac{g_{12}p^{(1)}}{g_{12}(\bar{p}-p^{(1)})+(\bar{p}-p^{(1)})+1}$. To equalize the rate constraints for $x_1^{(1)}$ at both receivers, we need

$$\mathsf{SNR}_1 = \mathsf{SNR}_2 \Rightarrow p^{(1)} = 1 - g_{12} + (1 - g_{12}^2)\bar{p}.$$

Note that $p^{(1)} < \bar{p}$ requires that $\bar{p} > \frac{1-g_{12}}{g_{12}^2}$. Otherwise, $\bar{p} \leq \frac{1-g_{12}}{g_{12}^2}$, and the above $1 - g_{12} + (1 - g_{12}^2)\bar{p} \geq \bar{p}$. It implies that we should not decode $x_1^{(1)}$ at the $2^{nd}$ receiver, i.e., $x_i^{(1)} (i = 1, 2)$ is the only message $(L = 1)$ of the $i^{th}$ user, which is treated as noise at the other receiver.

# References

[ADT11] S. Avestimehr, S. Diggavi, and D.N.C. Tse. "Wireless network information flow: a deterministic approach." *To appear in IEEE Transactions on Information Theory*, April 2011.

[And05] J. G. Andrews. "Interference cancellation for cellular systems: a contemporary overview." *IEEE Wireless Communications*, **vol.12, no.2**:19–29, April 2005.

[AV09] V.S. Annapureddy and V.V. Veeravalli. "Gaussian interference networks: sum capacity in the low-interference regime and new outer bounds on the capacity region." *IEEE Transactions on Information Theory*, **vol.55, no.7**:3032–3050, 2009.

[BDM09] R. E. Burkard, M. Dell'Amico, and S. Martello. *Assignment Problems*. SIAM, Philadelphia (PA.), 2009.

[BHB10] S.R. Bhaskaran, S.V. Hanly, N. Badruddin, and J.S. Evans. "Maximizing the Sum Rate in Symmetric Networks of Interfering Links." *IEEE Transactions on Information Theory*, **vol.56, no.9** : 4471–4487, September 2010.

[BT08] G. Bresler and D.N.C. Tse. "The two-user Gaussian interference channel: a deterministic view." *European Transactions in Telecommunications*, **vol. 19**:333–354, 2008.

[BV04] S.P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[Cha96] T.M. Chan. "Optimal output-sensitive convex hull algorithms in two and three dimensions." *Discrete and Computational Geometry*, **vol.16** : 361–368, 1996.

[CHC07] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen. "Autonomous Spectrum Balancing for Digital Subscriber Lines." *IEEE Transactions on Signal Processing*, **vol.55, no.8** : 4241–4257, August 2007.

[Chi05a] M. Chiang. "Balancing transport and physical Layers in wireless multihop networks: jointly optimal congestion control and power control." *IEEE Transactions on Selected Areas in Communications*, **vol.23, no.1** : 104–116, January 2005.

[Chi05b] M. Chiang. "Geometric programming for communication systems." *Foundations and Trends in Communications and Information Theory*, **vol.2, no. 1/2** : 1–156, August 2005.

[CJ09] V. R. Cadambe and S. Jafar. "Refections on interference alignment and the degrees of freedom of the K user interference channel." *IEEE Information Theory Society Newsletter*, **vol. 59**:5–9, December 2009.

[CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., New York, 1991.

[CYM06] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen. "Optimal multiuser spectrum balancing for digital subscriber lines." *IEEE Transactions on Communications*, **vol.54, no.5** : 922–933, May 2006.

[EMK06] M. Ebrahimi, M.A. Maddah-Ali, and A.K. Khandani. "Power allocation and asymptotic achievable sum-rates in single-hop wireless networks." *Proceedings of Conference on Information Sciences and Systems*, pp. 498–503, March 2006. .

[EPT07] R. Etkin, A. Parekh, and D.N.C. Tse. "Spectrum sharing for unlicensed bands." *IEEE Journal on Selected Areas in Communications*, **vol.25, no.3** : 517–528, April 2007.

[ETW08] R H. Etkin, D.N.C. Tse, and H. Wang. "Gaussian interference channel capacity to within one bit." *IEEE Transactions on Information Theory*, **vol.54, no.12**:5534–5562, 2008.

[GC82] A.E. Gamal and M. Costa. "The capacity region of a class of deterministic interference channels." *IEEE Transactions on Information Theory*, **vol.28, no.2**:343 – 346, March 1982.

[GCJ08] K. Gomadam, V. R. Cadambe, and S. A. Jafar. "Approaching the Capacity of Wireless Networks through Distributed Interference Alignment." *Proceedings of IEEE Global Telecommunications Conference*, pp. 1–6, Dec 2008. .

[GGO06] A. Gjendemsjo, D. Gesbert, G.E. Oien, and S.G. Kiani. "Optimal Power Allocation and Scheduling for Two-Cell Capacity Maximization." *4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp. 1–6, April 2006.

[Gol05] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[HK81] T. Han and K. Kobayashi. "A new achievable rate region for the interference channel." *IEEE Transactions on Information Theory*, **vol.27, no.1**: 49–60, January 1981.

[HL09] S. Hayashi and Z.Q. Luo. "Spectrum Management for Interference-Limited Multiuser Communication Systems." *IEEE Transactions on Information Theory*, **vol.55, no.3** : 1153–1175, 2009.

[HT99] S. V. Hanly and D. N. C. Tse. "Power control and capacity of spread spectrum wireless networks." *Automatica*, **Volume 35, Issue 12**:1987–2012, December 1999.

[KG06] S. Koskie and Z. Gajic. "Signal-to-interference-based power control in wireless networks: A Survey 1992 – 2005." *Dynamics of Continuous, Discrete and Impulsive Systems: Series B, Applications and Algorithms, invited paper*, **vol. 13**:187–220, 2006. .

[KN96] I. Katzela and M. Naghshineh. "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey." *IEEE Personal Communications*, **vol.3, no.3**:10–31, June 1996.

[Kuh55] H. W. Kuhn. "The Hungarian Method for the assignment problem." *Naval Research Logistics Quarterly*, **2**: 83–97, 1955.

[LZ08] Z. Luo and S. Zhang. "Dynamic spectrum management: complexity and duality." *IEEE Journal of Selected Topics in Signal Processing*, **vol.2, no.1**:57 –73, 2008.

[MDF11] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. N. C. Tse. "Approximate capacity of Gaussian interference-relay networks with weak cross links." *To appear in IEEE Transactions on Information Theory*, May 2011. See also CoRR, abs/1005.0404.

[MK09] A.S. Motahari and A.K. Khandani. "Capacity bounds for the Gaussian interference channel." *IEEE Transactions on Information Theory*, **vol.55, no.2**:620–643, February 2009.

[RSL10] M. Razaviyayn, M. Sanjabi, and Z. Q. Luo. "Linear transceiver design for interference alignment: Complexity and computation." *CoRR*, **abs/1009.3481**, 2010.

[SCA10] Z. Shao, M. Chen, S. Avestimehr, and S.-Y.R. Li. "Cross-layer optimization for wireless networks with deterministic channel models." *Proceedings of IEEE INFOCOM (mini-conference)*, 2010.

[SKC09] X. Shang, G. Kramer, and B. Chen. "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels." *IEEE Transactions on Information Theory*, **vol.55, no.2**:689–699, Feb. 2009.

[SZB08] H. Shen, H. Zhou, R.A. Berry, and M.L. Honig. "Optimal spectrum allocation in Gaussian interference networks." *42nd Asilomar Conference on Signals, Systems and Computers*, pp. 2142–2146, October 2008.

[TFL11] C. W. Tan, S. Friedland, and S. H. Low. "Spectrum management in multiuser cognitive wireless networks: optimality and algorithm." *IEEE Journal on Selected Areas in Communications*, 2011.

[YL06] W. Yu and R. Lui. "Dual methods for nonconvex spectrum optimization of multicarrier systems." *IEEE Transactions on Communications*, **vol.54, no.7** : 1310–1322, 2006.

[ZP09a] Y. Zhao and G. J. Pottie. "Optimal Spectrum Management in Multiuser Interference Channels." *Proceedings of IEEE International Symposium on Information Theory*, pp. 2266–2270, June 2009.

[ZP09b] Y. Zhao and G. J. Pottie. "Optimal spectrum management in two-user symmetric interference channels." *Proceedings of Information Theory and Applications Workshop*, pp. 256–263, February 2009.

[ZP10] Y. Zhao and G. J. Pottie. "Optimization of Power and Channel Allocation Using the Deterministic Channel Model." *Proceedings of Information Theory and Applications Workshop*, pp. 1–8, February 2010.

[ZP11a] Y. Zhao and G. J. Pottie. "Interference Strength Alignment and Uplink Channel Allocation in Linear Cellular Networks." *to appear in Proceedings of IEEE International Conference on Communications*, June 2011.

[ZP11b] Y. Zhao and G. J. Pottie. "Optimal Spectrum Management in Multiuser Interference Channels." *CoRR*, **abs/1102.3758**, 2011.

[ZTA11a] Y. Zhao, C. W. Tan, A. S. Avestimehr, S. N. Diggavi, and G. J. Pottie. "On the sum-capacity with successive decoding in interference channels." *to appear in Proceedings of IEEE International Symposium on Information Theory*, July 2011.

[ZTA11b] Y. Zhao, C. W. Tan, A. S. Avestimehr, S. N. Diggavi, and G. J. Pottie. "On the Sum-Capacity with Successive Decoding in Interference Channels." *CoRR*, **abs/1103.0038**, 2011.