FUNDAMENTAL LIMITS OF SPECTRUM SHARING IN DEVICE-TO-DEVICE COMMUNICATION SYSTEMS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Navid Naderializadeh

January 2014

ABSTRACT

This work presents a new spectrum sharing mechanism for device-to-device communication systems. First, a K-user interference channel is considered and a sufficient condition under which using Gaussian codebooks at the sources and treating interference as noise at the destinations is information-theoretically optimal is derived. Afterwards, the notion of *information-theoretic independent sets* (in short, ITIS) is defined which denotes the subsets of users in a wireless network with n source-destination pairs inside each of which treating interference as noise is optimal. Then, the novel spectrum sharing mechanism of *information-theoretic link scheduling* (in short, ITLinQ) is proposed which at each time schedules the users that form an ITIS. Furthermore, a distributed way of implementing ITLinQ is presented and it is shown, through numerical analysis, to outperform similar state-of-the-art spectrum sharing mechanisms such as FlashLinQ. Finally, the impact of network topology is studied in wireless networks where there is no channel state information at the transmitters using more sophisticated transmission schemes.

BIOGRAPHICAL SKETCH

Navid Naderializadeh was born in 1990 in Qaemshahr, a small beautiful city in Northern Iran. He attended Kharazmi for his primary school and Shahid Beheshti Center of Education for his secondary and high school. After achieving the 1st rank in the Iranian nationwide entrance examination of Iranian universities in 2007, Navid joined Sharif University of Technology in Tehran, Iran for his undergraduate studies in electrical engineering. He selected the branch of communications and obtained his Bachelor of Science in 2011 from Sharif University. In the same year, Navid got admitted to the graduate program of the department of electrical and computer engineering (ECE) at Cornell University as a Jacobs scholar. Since then, he has been with the foundations of information engineering (FoIE) where he conducted his research in the field of network information theory. Navid completed his Master's studies in December 2013.

To My Dearest Father, Mother and Brother.

ACKNOWLEDGEMENTS

In the course of my education at Cornell University and preparing this thesis, I received help from several people who have paved the way toward obtaining my Master of Science degree here at Cornell.

First of all, I would like to thank my advisor, Prof. Salman Avestimehr, who has immensely helped me to conduct my research work in the best way. Through our weekly meetings, he always motivated me to meet high standards in my research and this was a key factor in my success over these years. I am also grateful to Prof. Lang Tong, who agreed to be on my graduate committee, for his assistance during this period. Moreover, thanks to Professors Ehsan Afshari, Aaron Wagner, James Renegar and Kevin Tang for their help and guidance while I have been at Cornell.

I had numerous friends in Ithaca who made me feel really comfortable in my first experience of studying abroad. Among them, I am grateful to my closest friends at Cornell, Hamidreza Aghasi and Sina Lashgari, for helping me in many aspects in these years. Also, thanks to my other friends, Mohammad Mahmoody, Sepehr Saroukhani, Rad Niazadeh, Hadi Hosseinzadegan, Vahid Edriss, Amirahmad Tarkeshdouz, Ali Mostajeran, Somayeh Khiyabani, Nima Taherkhani, Hedyeh Beyhaghi, Naser Nikandish, Hamid Khatibi, Vahnood Pourahmad, Javad Rostami, Elham Alipanahi, Abolhassan Vaezi, Mahya Mehrmohamadi, Yahya Tousi, Marjaneh Mottaghi, Arash Beheshtian, and Saeid Alaei. They will always bring me pleasant memories of Ithaca and Cornell.

I have enjoyed spending time in the same office with my research teammates, Ibrahim Issa, Ilan Shomorony, Alireza Vahid, David Kao and Silas Fong over these years. We had fruitful discussions on various topics and I am grateful to all of them. I have also made a lot of friends in the department of ECE at

Cornell during these years. Special thanks to Amandy Nwana, Sinem Unal, Bye-sah Gantsog, Raphael Louca, Daniel Munoz Alvarez, and Yuting Ji for their help and kindness.

Two other people in the ECE department of Cornell deserve special acknowledgement. First, I would like to thank Scott Coldren, who was of incredibly great help for all the academic issues that I had at Cornell. Also, thanks to Daniel Richter for his sincere assistance and support through these years.

Clearly, there has also been help and support from outside Ithaca that contributed to my successful performance at Cornell. Especially, I would like to thanks my close friends outside Cornell, Amirkhosro Gouran, Shervin Minaee, Morteza Hashemi, and Alireza Sheikhattar. Finally, huge thanks to my family, my parents and my brother, who helped me through every moment of my life to reach this point. This thesis is dedicated to them.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Device-to-device (D2D) communication is expected to play a fundamental role in next generation wireless systems. This type of communication bypasses the main base station of the network and allows nearby devices to communicate directly. D2D communication has been shown to provide significant improvement in resource utilization of the wireless networks which leads to an enhancement in system performance (see, e.g. [1, 2]). It has many broad applications ranging from proximity-based services (such as Internet of Things) to on-demand video caching networks and underlay to cellular networks (see, e.g. [3, 4, 5, 6, 7, 8]).

However, given the increasing traffic volume in wireless networks, the main bottleneck in these networks is the issue of interference. To date, the interference management mechanisms can be categorized into two major classes. The first type of approaches are fully-coordinated approaches relying on advanced physical layer mechanisms such as interference alignment [9, 10]. On the other hand, the second type of approaches includes fully-distributed WiFi-type mechanisms (such as CSMA/CA). Both of these types of approaches have drawbacks in the sense that the first class of approaches are hard to be implemented in practice since they need very high levels of centralization and coordination among wireless nodes, whereas the approaches of the second type experience performance degradation with increasing number of users.

This motivates an alternative interference management strategy in which in a system of multiple users (i.e., source-destination pairs), a subset of users with *sufficiently* low level of interference among them is selected and the users in that subset are scheduled to transmit at the same time. Such a subset is called an *independent set*. In fact, the *independent set scheduling* scheme is a scheme in which two users are considered to be non-interfering if the interference that they cause on each other is below a certain threshold (see, e.g. [11, 12, 13, 14, 15] and the protocol model in [16]). The problem with such a scheme is that the threshold is fixed at a certain value (often at noise level) which does not capture various parameters of the network, in particular the strength of direct desired signals and also the number of users.

To overcome these issues, a more recent spectrum sharing approach, called FlashLinQ [17], has been proposed which modifies the criteria of finding an independent set and maintains its promising performance for a large number of users while needing a low level of coordination among the wireless nodes. In a system of multiple users, FlashLinQ first orders the users according to a random priority list. Then, starting from the higher priority users, each user is scheduled to transmit if it does not receive/cause *much* interference from/at higher priority users. This way, FlashLinQ forms an independent subset of the users by considering the metric of signal-to-interference ratio (SIR). This subset is formed in a distributed way which makes the scheme suitable to be implemented in practice.

Therefore, both the regular independent set scheduling and the FlashLinQ schemes intend to find subsets of users with *sufficiently* low levels of interference. A natural question that comes to mind is: What is a *theoretically-justified criteria* to find such a subset?

In this work, we propose an answer to this question. We focus on the subsets of the users inside which *treating interference as noise* is information-theoretically optimal. This requires the derivation of a sufficient condition under which treating interference as noise (in short, TIN) is optimal. To this end, we first consider an interference channel comprising multiple source-destination pairs. Then we focus on the simple scheme of power control and using Gaussian codebooks at the transmitters (i.e., sources) and treating interference as noise at the receivers (i.e., destinations), and we derive a condition under which this scheme can achieve the whole capacity region of the network to within a constant gap.

After characterizing this condition, we call the subsets of users which satisfy this condition *information-theoretic independent subsets* (in short, ITIS). This introduces a refinement to the aforementioned regular notion of independent sets. Based on this concept, we propose our novel spectrum sharing mechanism which we call *information-theoretic link scheduling* (in short, ITLinQ). This scheme simply schedules the users inside an ITIS to transmit data while they share the same time and frequency resource. To assess the performance of this scheme, we first perform a capacity analysis of ITLinQ and show that it can achieve a certain fraction of the capacity region of the whole network (to within a gap) in a specific network model. Moreover, we present a way of implementing ITLinQ in a distributed fashion in which the nodes gain the required channel state information through a careful two-phase signaling mechanism. We will also show, through numerical analysis, that the distributed ITLinQ scheme can achieve a sum-rate gain of over 110% with respect to FlashLinQ.

Finally, as an extension to the treating interference as noise scheme, we study the impact of topology on partially-connected interference networks where there is no channel state information at the transmitters (referred to as no CSIT). We develop several linear-algebraic and graph-theoretic outer and inner bounds on the symmetric degrees-of-freedom (DoF) of these networks. We also evaluate our bounds for two classes of networks to show their tightness for most networks in these categories and also to quantify the gain of our inner bounds over benchmark interference management strategies.

CHAPTER 2

STUDYING THE OPTIMALITY OF TREATING INTERFERENCE AS NOISE IN INTERFERENCE CHANNELS

In this chapter, we consider K-user fully-connected fully-asymmetric Gaussian interference channels and present a general condition under which treating interference as noise (in short, TIN) is information-theoretically optimal. In fact, we will show that under this condition, TIN can achieve the whole generalized degrees-of-freedom (GDoF) region of the network (corresponding to the high-SNR regime) and can also achieve the entire capacity region of the network to within a constant gap (corresponding to the finite-SNR regime). As an extension, we will also characterize the entire GDoF region that TIN is able to achieve for any set of channel gain values.[1]

2.1 System Model and Preliminaries

As our starting point, consider the canonical model of a fully-asymmetric K-user wireless interference channel, with the input-output relationship

$$Y_k(t) = \sum_{i=1}^{K} h_{ki}\tilde{X}_i(t) + Z_k(t), \quad \forall k \in \{1, 2, ..., K\}, \tag{2.1}$$

where at each time index t, $\tilde{X}_i(t)$ is the transmitted symbol of transmitter i, $Y_k(t)$ is the received signal of receiver k, h_{ki} is the complex channel gain value from transmitter i to receiver k, and $Z_k(t) \sim \mathcal{CN}(0, 1)$ is the additive white Gaussian noise (AWGN) at receiver k. All the symbols are complex. Each transmitter i is subject to the power constraint $E[|\tilde{X}_i(t)|^2] \le P_i$.

[1]The remaining portion of the chapter is mainly taken from [18, 19], coauthored by the author of this thesis.

We will translate the standard channel model (2.1) into an equivalent normalized form that is more conducive for GDoF studies. We define the signal-to-noise ratio (SNR) of user i and interference-to-noise ratio (INR) of transmitter i at receiver k as follows[2].

$$\mathrm{SNR}_i \triangleq \max(1, |h_{ii}|^2 P_i), \quad \mathrm{INR}_{ki} \triangleq \max(1, |h_{ki}|^2 P_i), \quad i \neq k, \quad i, k \in \{1, 2, ..., K\}. \tag{2.2}$$

As in [20], for the GDoF metric, we preserve the ratios of different signal strengths in dB scale as all SNRs approach infinity. To this end, taking $P > 1$ as a nominal power value, we define

$$\alpha_{ii} \triangleq \frac{\log \mathrm{SNR}_i}{\log P}, \quad \alpha_{ki} \triangleq \frac{\log \mathrm{INR}_{ki}}{\log P}, \quad i \neq k, \quad i, k \in \{1, 2, ..., K\}, \tag{2.3}$$

implying that for each user i, $\mathrm{SNR}_i = P^{\alpha_{ii}}$ and for any two distinct users i, k, $\mathrm{INR}_{ki} = P^{\alpha_{ki}}$.

Now according to (2.2) and (2.3), we can represent the original channel model in (2.1) in the following form,

$$Y_k(t) = \sum_{i=1}^{K} \sqrt{P^{\alpha_{ki}}} e^{j\theta_{ki}} X_i(t) + Z_k(t), \quad \forall k \in \{1, 2, ..., K\}. \tag{2.4}$$

In this equivalent channel model, $X_i(t) = \tilde{X}_i(t)/\sqrt{P_i}$ is the transmit symbol of transmitter i, and the power constraint for each transmitter is normalized to unity; i.e., $E[|X_i(t)|^2] \leq 1$, $\forall i \in \{1, 2, ..., K\}$. The transmit power in the original channel model is absorbed in the channel coefficients, so that $\sqrt{P^{\alpha_{ki}}}$ and θ_{ki} are the magnitude and the phase, respectively, of the channel between transmitter i and receiver k, $\forall i, k \in \{1, 2, ..., K\}$. We will call the exponent α_{ki} the channel

[2] It is not difficult to verify that assigning a value of 1 to SNR's and INR's that are less than 1, or equivalently, assigning a 0 value to α_{ij} that might otherwise be negative, is only a matter of convenience, and has no impact on the GDoF or the constant gap result.

strength level of the link between transmitter i and receiver k. *In the rest of this chapter, we will only consider the equivalent channel model in (2.4).*

Since this is a K-user interference channel, transmitter i has message W_i intended for receiver i, and the messages W_i are independent, $\forall i \in \{1, 2, ..., K\}$. We denote the size of the message set of user i by $|W_i|$. For codewords spanning n channel uses, the rates $R_i = \frac{\log|W_i|}{n}, i \in \{1, 2, \cdots, K\}$, are achievable if the probability of error at all the receivers can be made arbitrarily small as n approaches infinity. The channel capacity region C is the closure of the set of all achievable rate tuples. Collecting the channel strength levels and phases in the sets

$$\alpha \triangleq \{\alpha_{ki}\}, \quad \theta \triangleq \{\theta_{ki}\}, \quad \forall i, k \in \{1, 2, ..., K\}, \tag{2.5}$$

the capacity region is a function of α, θ, P, and is denoted as $C(P, \alpha, \theta)$.

2.1.1 Generalized Degrees of Freedom

The GDoF region of the K-user interference channel as represented in (2.4) is defined as

$$\begin{aligned}\mathcal{D}(\alpha, \theta) \triangleq \Big\{(d_1, d_2, ..., d_K) :\ d_i = \lim_{P\to\infty} \frac{R_i}{\log P}, \\ \forall i \in \{1, 2, ..., K\}, (R_1, R_2, ..., R_K) \in C(P, \alpha, \theta)\Big\}.\end{aligned}$$

In general, the channel capacity (GDoF) region of complex Gaussian interference channel may depend on both the channel strength levels α, and the channel phases θ. However, the capacity (GDoF) inner and outer bounds that we present in this chapter depend *only* on the channel strength levels α. As such, our results hold regardless of whether or not the channel phase information is available to the transmitters.

2.1.2 Capacity Region within a Constant Gap

Following the same definition as in [20] and [21], an achievable region is said to be within x bits of the capacity region if for any rate tuple $(R_1, R_2, ..., R_K)$ on the boundary of the achievable region, the rate tuple $(R_1 + x, R_2 + x, ..., R_K + x)$ is outside the channel capacity region.

2.1.3 Achievable Rate Region of TIN Scheme

In the TIN scheme, transmitter i uses a transmit power of P^{r_i}, $r_i \leq 0$ and each receiver treats all the incoming interference as noise, so that the SINR at receiver i is given by

$$\mathrm{SINR}_i = \frac{P^{\alpha_{ii}} \times P^{r_i}}{1 + \sum_{j \neq i} P^{\alpha_{ij}} \times P^{r_j}}.$$

This implies that the rate achieved by user i through TIN is equal to

$$R_i = \log(1 + \mathrm{SINR}_i) = \log\left(1 + \frac{P^{\alpha_{ii}+r_i}}{1 + \sum_{j \neq i} P^{\alpha_{ij}+r_j}}\right), \quad (2.6)$$

and therefore, the generalized degrees-of-freedom (GDoF) achieved by user i equals

$$d_i = \max\{0, \alpha_{ii} + r_i - \max\{0, \max_{j:j \neq i}(\alpha_{ij} + r_j)\}\}. \quad (2.7)$$

The achievable GDoF region through TIN, which we denote by $\mathcal{P}^*$, is the set of all K-tuples $(d_1, ..., d_K)$ for which there exist r_i's, $r_i \leq 0$, $i \in \{1, ..., K\}$, such that (2.7) holds for all $i \in \{1, ..., K\}$.

2.2 Condition for Optimality of TIN

The main result of this section is the following theorem, which introduces a condition under which TIN is GDoF-optimal.

Theorem 2.1. *In a K-user interference channel, where the channel strength level from transmitter i to receiver j is equal to α_{ji}, $\forall i, j \in \{1, ..., K\}$, if the following condition is satisfied*

$$\alpha_{ii} \geq \max_{j:j\neq i}\{\alpha_{ji}\} + \max_{k:k\neq i}\{\alpha_{ik}\}, \quad \forall i, j, k \in \{1, 2, ...K\}, \tag{2.8}$$

then power control and treating interference as noise can achieve the whole GDoF region. Moreover, the GDoF region is the set of all K-tuples $(d_1, d_2, ..., d_K)$ satisfying

$$0 \leq d_i \leq \alpha_{ii}, \qquad \forall i \in \{1, ..., K\} \tag{2.9}$$

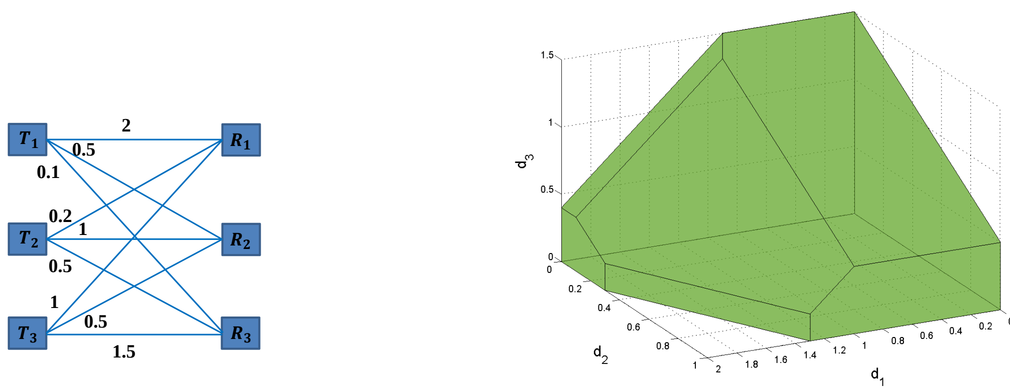
$$\sum_{j=1}^{m} d_{i_j} \leq \sum_{j=1}^{m}(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}), \ \forall (i_0, i_1, ..., i_m) \in \Pi_K, \ \ \forall m \in \{2, 3, ...K\}, \tag{2.10}$$

where Π_K is the set of all possible cyclic sequences of all subsets of $\{1, ..., K\}$, and the modulo-m arithmetic is implicitly used on the user indices, e.g., $i_m = i_0$.

Remark. Condition (2.8) can be stated in words as — *for each user the desired signal strength is no less than the sum of the strengths of the strongest interference from this user and the strongest interference to this user (all values in dB scale).* Theorem 2.1 claims that under this condition, TIN is GDoF-optimal.

Remark. Both the condition (2.8) and the GDoF region specified by (2.9)-(2.10) display a natural duality in the sense that they are both unchanged if the roles of the transmitters and receivers are switched, i.e., if all α_{ij} values are switched with α_{ji} values. In other words, for the same channel strengths, if we consider the reciprocal network (in the same sense as a multiple access channel being the reciprocal of a broadcast channel), then again under condition (2.8), TIN is GDoF-optimal, and the GDoF region is the same as in the original network. Such a duality holds also for the entire TIN region $\mathcal{P}^*$, and a similar duality relationship for the symmetric rate has been observed in [22].

Example 1. To interpret the results in Theorem 2.1, we derive and plot the GDoF region for a 3-user network in which the condition (2.8) is satisfied. Consider the 3-user network in Fig. 2.1(a). In this network, the channel strength



(a) (b)

Figure 2.1: (a) A 3-user interference channel, where the value on each link is equal to its channel strength level, and (b) The GDoF region of this network, which is a convex polyhedron and can be achieved by TIN.

level between transmitter i and receiver j, α_{ji}, is shown on the corresponding link, $\forall i, j \in \{1, 2, 3\}$. For the case of $K = 3$, $\Pi_K = \{(1, 2), (1, 3), (2, 3), (1, 2, 3), (1, 3, 2)\}$. According to Theorem 2.1, the GDoF region is the set of all (d_1, d_2, d_3) satisfying

$$0 \le d_1 \le 2$$

$$0 \le d_2 \le 1$$

$$0 \le d_3 \le 1.5$$

$$d_1 + d_2 \le 2.3$$

$$d_1 + d_3 \le 2.4$$

$$d_2 + d_3 \leq 1.5$$

$$d_1 + d_2 + d_3 \leq 3.7$$

$$d_1 + d_2 + d_3 \leq 2.5,$$

which is depicted in Fig. 2.1(b). Recall that the condition (2.8) is satisfied in the network of Fig. 2.1(a) for all users $i \in \{1, 2, 3\}$. Therefore, Theorem 2.1 implies that TIN achieves the entire GDoF region of this network. $\triangle$

We prove Theorem 2.1 through the following steps. We first show that under the condition stated in (2.8), the achievable GDoF region of TIN simplifies into a polyhedral region. We study the polyhedral TIN region in some detail to understand its structure. In particular, we show that the polyhedral TIN region can be characterized by checking the existence of a potential function for an induced fully-connected directed graph, with nodes representing the source-destination pairs in the original interference channel (with the addition of a "ground" node) and a specific assignment of lengths to the arcs of the graph. Afterwards, we derive a dual characterization of the polyhedral TIN region and use the outer bounds developed in [21] to prove the optimality of polyhedral TIN, hence TIN, whenever condition (2.8) holds.

2.2.1 Polyhedral Relaxation of TIN

In the first step toward proving Theorem 2.1, we introduce a polyhedral version of the TIN scheme. Ignoring the first $\max\{0, ...\}$ term in (2.7) changes the scheme to a relaxed version, which we call the *polyhedral TIN* scheme. With this modification, the GDoF achieved by user i will be

$$d_i = \alpha_{ii} + r_i - \max\{0, \max_{j:j\neq i}(\alpha_{ij} + r_j)\}, \tag{2.11}$$

and we denote the achievable GDoF region via polyhedral TIN by $\mathcal{P}$.

In general, comparing (2.7) and (2.11) shows that this modification can only shrink the achievable GDoF region of TIN. However, as we will show in the following, under the condition (2.8), the above relaxation incurs no loss in the GDoF region of TIN. In other words, *when the condition (2.8) is satisfied, the TIN region $\mathcal{P}^*$ is equal to the polyhedral TIN region $\mathcal{P}$.* From (2.11), the polyhedral TIN region $\mathcal{P}$ can be characterized by a number of linear inequalities, which, as we will see, significantly contributes to understanding the TIN region $\mathcal{P}^*$. In fact, $\mathcal{P}$ is the set of all K-tuples $(d_1, ..., d_K)$ for which there exist r_i's, $i \in \{1, ..., K\}$, such that

$$\begin{aligned}
r_i &\leq 0, && \forall i \in \{1, ..., K\} \\
d_i \leq \alpha_{ii} + r_i \Leftrightarrow r_i &\geq d_i - \alpha_{ii}, && \forall i \in \{1, ..., K\} \\
d_i \leq \alpha_{ii} + r_i - (\alpha_{ij} + r_j) \Leftrightarrow r_i - r_j &\geq \alpha_{ij} + (d_i - \alpha_{ii}), && \forall i, j \in \{1, ..., K\}, i \neq j.
\end{aligned}$$

As we will show, the region $\mathcal{P}$ can be fully characterized by (2.9)-(2.10). Moreover, as demonstrated in Example 1, the region $\mathcal{P}$ is a polyhedron, which is why the scheme is called polyhedral TIN. Note in general it is obvious that $\mathcal{P} \subseteq \mathcal{P}^*$, because TIN performs no worse than polyhedral TIN.

2.2.2 Dual Characterization of Polyhedral TIN Region via Potential Functions

Equipped with the aforementioned description of polyhedral TIN, we now characterize the polyhedral TIN region $\mathcal{P}$ for general channel strength levels. As mentioned earlier, $(d_1, d_2, ..., d_K) \in \mathcal{P}$ if and only if there exist r_i's, $i \in \{1, ..., K\}$, satisfying

$$r_i \leq 0, \qquad \forall i \in \{1, ..., K\} \tag{2.12}$$

$$r_i \geq d_i - \alpha_{ii}, \qquad \forall i \in \{1, ..., K\} \tag{2.13}$$

$$r_i - r_j \geq \alpha_{ij} + (d_i - \alpha_{ii}), \ \forall i, j \in \{1, ..., K\}, i \neq j. \tag{2.14}$$

Now, we define a directed graph $D = (V, A)$, as shown in Fig. 2.2(a), where

$$\begin{aligned} V &= \{v_1, ..., v_K, u\} \\ A &= A_1 \cup A_2 \cup A_3 \\ A_1 &= \{(v_i, v_j) : i, j \in \{1, ..., K\}, i \neq j\} \\ A_2 &= \{(v_i, u) : i \in \{1, ..., K\}\} \\ A_3 &= \{(u, v_i) : i \in \{1, ..., K\}\}, \end{aligned}$$

and we assign a length $l(a)$ to every arc $a \in A$ as follows.

$$\begin{aligned} l(v_i, v_j) &= \alpha_{ii} - d_i - \alpha_{ij} \\ l(v_i, u) &= \alpha_{ii} - d_i \\ l(u, v_i) &= 0. \end{aligned}$$

As an example, the corresponding directed graph D for Example 1 is drawn in Fig. 2.2(b). Evidently, this is a fully-connected directed graph, in which the

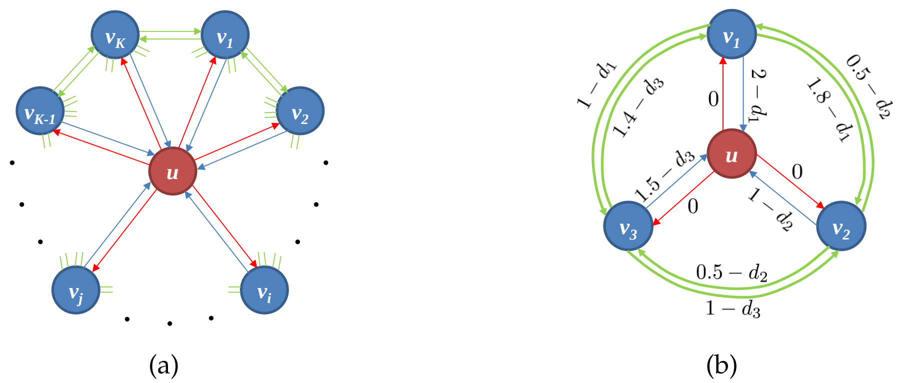


Figure 2.2: (a) The directed graph D in which the green, blue and red arcs belong to A_1, A_2 and A_3, respectively. For simplicity, only some parts of the edges are shown in this figure. (b) The corresponding directed graph D for Example 1.

length of each arc depends on the channel strength levels and the GDoFs we intend to achieve. This careful assignment of the lengths to the arcs of this graph allows us to use the following lemma.

Lemma 2.1. *If $\mathcal{P}$ denotes the polyhedral TIN region of a K-user interference channel, then $(d_1, ..., d_K) \in \mathcal{P}$ if and only if there exists a valid potential function for the graph D.*

Proof of Lemma 2.1. By definition [23], a function $p : V \to \mathbb{R}$ is called a potential if for every two nodes $a, b \in V$ such that $(a, b) \in A$, $l(a, b) \geq p(b) - p(a)$. These inequalities only depend on the *difference* between potential function values. Therefore, without loss of generality, if there exists a valid potential function for the graph, we can make one node, say node u, *ground*; i.e., $p(u) = 0$. Letting $r_i := p(v_i)$, the potential function values should satisfy the following conditions.

$$\alpha_{ii} - d_i - \alpha_{ij} \geq r_j - r_i, \qquad \forall i, j \in \{1, ..., K\}, i \neq j \tag{2.15}$$

$$\alpha_{ii} - d_i \geq -r_i, \qquad \forall i \in \{1, ..., K\} \tag{2.16}$$

$$0 \geq r_i, \qquad \forall i \in \{1, ..., K\}. \tag{2.17}$$

The above inequalities exactly match the ones in (2.12)-(2.14). This completes the proof. $\square$

Next, we invoke the potential theorem of [23], re-stated below, to complete the characterization of the polyhedral TIN region, $\mathcal{P}$.

Potential Theorem [Theorem 8.2 of [23]]*: There exists a potential function for a directed graph D if and only if each directed circuit in D has nonnegative length.*

Combining Lemma 2.1 and the potential theorem, we conclude that $(d_1, ..., d_K) \in \mathcal{P}$ if and only if each directed circuit in the graph D has nonnegative length. Therefore, it just remains to interpret the conditions of nonnegative length for the circuits.

We can categorize the circuits of D in three classes:

- Circuits in the form of (u, v_i, u). For these circuits, we have

$$\alpha_{ii} - d_i \geq 0 \Leftrightarrow d_i \leq \alpha_{ii}. \tag{2.18}$$

- Circuits in the form of $(v_{i_0}, v_{i_1}, ..., v_{i_m})$, where $i_0 = i_m$, or in other words, the circuits which do not include node u. For these circuits, the nonnegative length condition will be

$$\begin{aligned} \sum_{j=0}^{m-1}(\alpha_{i_j i_j} - d_{i_j} - \alpha_{i_j i_{j+1}}) \geq 0 &\Leftrightarrow \sum_{j=0}^{m-1} d_{i_j} \leq \sum_{j=0}^{m-1}(\alpha_{i_j i_j} - \alpha_{i_j i_{j+1}}) \\ &\overset{(a)}{\Leftrightarrow} \sum_{j=1}^{m} d_{i_j} \leq \sum_{j=1}^{m}(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}). \end{aligned} \tag{2.19}$$

where in step (a) we just reorder the terms in the right hand side and recall that $i_m = i_0$.

- Circuits in the form of $(u, v_{i_1}, ..., v_{i_m}, u)$, where $m > 1$. For these circuits, the following inequality should hold.

$$\sum_{j=1}^{m-1}(\alpha_{i_j i_j} - d_{i_j} - \alpha_{i_j i_{j+1}}) + (\alpha_{i_m i_m} - d_{i_m}) \geq 0. \tag{2.20}$$

  Since $\alpha_{i_m i_1} \geq 0$, we have $\alpha_{i_m i_m} - d_{i_m} \geq \alpha_{i_m i_m} - d_{i_m} - \alpha_{i_m i_1}$. Therefore, given the conditions (2.19), the conditions in this class of circuits are redundant.

Consequently, we will end up with the conditions (2.18)-(2.19), which coincide accurately with the conditions (2.9)-(2.10), except for the non-negativity constraint on d_i's, which is needed for the generalized degrees-of-freedom to be meaningful. This directly leads us to the following theorem which characterizes the polyhedral TIN region $\mathcal{P}$ for general channel strength levels.

Theorem 2.2. *The GDoF region achieved through polyhedral TIN, denoted by $\mathcal{P}$, is the set of all K-tuples $(d_1, d_2, ..., d_K)$ satisfying*

$$0 \leq d_i \leq \alpha_{ii}, \qquad \forall i \in \{1, ..., K\} \tag{2.21}$$

$$\sum_{j=1}^{m} d_{i_j} \leq \sum_{j=1}^{m}(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}), \ \forall (i_0, i_1, ..., i_m) \in \Pi_K, \quad \forall m \in \{2, 3, ...K\}, \tag{2.22}$$

where Π_K is the set of all possible cyclic sequences of all subsets of $\{1, ..., K\}$, and the modulo-m arithmetic is implicitly used on the user indices, e.g., $i_m = i_0$.

Now, we are at a stage to complete the proof of Theorem 2.1.

2.2.3 Proof of Theorem 2.1

To prove Theorem 2.1, we show that under the condition (2.8), the polyhedral TIN region $\mathcal{P}$ coincides with the GDoF region outer bound, therefore establishing the optimality of TIN under (2.8) and proving Theorem 2.1. Note that from

Theorem 2.2, the region (2.9)-(2.10) is exactly equal to the polyhedral TIN region $\mathcal{P}$ and therefore this GDoF region can be achieved by TIN. Therefore, we only need to prove the outer bounds on the GDoF region.

In order to prove the converse, we use the outer bounds presented in [21] for cyclic Gaussian interference channels. In [21], the authors investigate an interesting K-user cyclic Gaussian interference channel, where the k-th user only interferes with the $(k-1)$-th user (mod K) as shown in Fig. 2.3. For completeness, let us re-state the key result of [21] that we need to complete the proof.

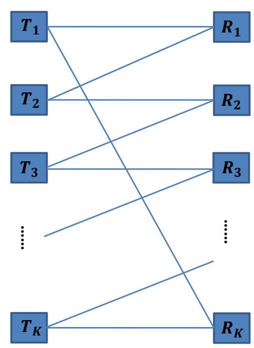


Figure 2.3: K-user cyclic interference channel.

For the K-user cyclic Gaussian interference channel, if we denote the channel gain between transmitter i and receiver j as $h_{j,i}$ and assume that each transmitter i is subject to the power constraint P_i and the additive white Gaussian noise (AWGN) at each receiver follows the distribution $\mathcal{CN}(0, \sigma^2)$, then we define the signal-to-noise and interference-to-noise ratios for each user as follows

$$\mathrm{SNR_i} = \frac{|h_{i,i}|^2 P_i}{\sigma^2}, \quad \mathrm{INR}_i = \frac{|h_{i-1,i}|^2 P_i}{\sigma^2}, \quad \forall i \in \{1, 2, ..., K\},$$

where modulo arithmetic is used on the user indices. The K-user cyclic Gaussian interference channel is in the *weak* interference regime if

$$\mathrm{INR}_i \le \mathrm{SNR}_i \;\; \forall i \in \{1, 2, ..., K\}.$$

The following theorem gives the channel capacity outer bounds for this K-user cyclic channel in the weak interference regime.

Theorem 2.3. (Theorem 2 in [21])[3] *For the K-user cyclic Gaussian interference channel in the weak interference regime, the capacity region is included in the set of rate tuples $(R_1, R_2, ..., R_K)$ such that*

$$R_i \le \lambda_i, \tag{2.23}$$

$$\sum_{j=m}^{m+l-1} R_j \le \min\Big\{\gamma_m + \sum_{j=m+1}^{m+l-2} \kappa_j + \beta_{m+l-1},\ \mu_m + \sum_{j=m}^{m+l-2} \kappa_j + \beta_{m+l-1}\Big\}, \tag{2.24}$$

$$\sum_{j=1}^{K} R_j \le \min\Big\{\sum_{j=1}^{K} \kappa_j, \rho_1, \rho_2, ..., \rho_K\Big\}, \tag{2.25}$$

$$\sum_{j=1}^{K} R_j + R_i \le \beta_i + \gamma_i + \sum_{j=1, j\ne i}^{K} \kappa_j, \tag{2.26}$$

where $i,m \in \{1, 2, ..., K\}$, $l \in \{2, 3, ..., K-1\}$, and

$$\begin{aligned}
\kappa_i &= \log\left(1 + \mathrm{INR}_{i+1} + \frac{\mathrm{SNR}_i}{1 + \mathrm{INR}_i}\right) \\
\beta_i &= \log\left(\frac{1 + \mathrm{SNR}_i}{1 + \mathrm{INR}_i}\right) \\
\gamma_i &= \log(1 + \mathrm{INR}_{i+1} + \mathrm{SNR}_i) \\
\lambda_i &= \log(1 + \mathrm{SNR}_i) \\
\mu_i &= \log(1 + \mathrm{INR}_i) \\
\rho_i &= \beta_{i-1} + \gamma_i + \sum_{j=1, j\notin\{i,i-1\}}^{K} \kappa_j.
\end{aligned}$$

Equipped with Theorem 2.3, we can now prove the converse for Theorem 2.1. The individual bounds (2.9) follow directly from the inequalities (2.23). In fact, from (2.23) we have

$$d_i = \lim_{P\to\infty} \frac{R_i}{\log P} \le \lim_{P\to\infty} \frac{\log(1 + P^{\alpha_{ii}})}{\log P} = \alpha_{ii},$$

[3] There is a minor change of notation compared to [21] in order to avoid confusion with the notations in this chapter.

for any $i \in \{1, 2, ..., K\}$. Also, the cyclic outer bounds (2.10) follow from the outer bounds (2.25) under the condition (2.8). Under this condition, for any cycle $(i_0, i_1, ..., i_m) \in \Pi_K$ we have

$$\begin{aligned}
\lim_{P\to\infty} \frac{\sum_{j=1}^{m} \kappa_{i_j}}{\log P} &= \lim_{P\to\infty} \frac{\sum_{j=1}^{m} \log\left(1 + P^{\alpha_{i_j i_{j+1}}} + \frac{P^{\alpha_{i_j i_j}}}{1+P^{\alpha_{i_{j-1} i_j}}}\right)}{\log P} \\
&= \sum_{j=1}^{m} \max\{0, \alpha_{i_j i_{j+1}}, \alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}\} \\
&= \sum_{j=1}^{m} (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}),
\end{aligned} \tag{2.27}$$

and it follows that for any $k \in \{1, 2, ..., m\}$,

$$\begin{aligned}
\lim_{P\to\infty} \frac{\rho_{i_k}}{\log P} &= \lim_{P\to\infty} \frac{\beta_{i_{k-1}} + \gamma_{i_k} + \sum_{j=1, j\notin\{k,k-1\}}^{m} \kappa_{i_j}}{\log P} \\
&= \lim_{P\to\infty} \frac{\log\left(\frac{1+P^{\alpha_{i_{k-1} i_{k-1}}}}{1+P^{\alpha_{i_{k-2} i_{k-1}}}}\right) + \log(1 + P^{\alpha_{i_k i_{k+1}}} + P^{\alpha_{i_k i_k}})}{\log P} + \sum_{j=1, j\notin\{k,k-1\}}^{m} (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \\
&= (\alpha_{i_{k-1} i_{k-1}} - \alpha_{i_{k-2} i_{k-1}}) + \alpha_{i_k i_k} + \sum_{j=1, j\notin\{k,k-1\}}^{m} (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \\
&= \alpha_{i_k i_k} + \sum_{j=1, j\notin\{k\}}^{m} (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}).
\end{aligned} \tag{2.28}$$

Therefore, comparing (2.27) and (2.28) implies that under the condition (2.8), we have

$$\lim_{P\to\infty} \frac{\sum_{j=1}^{m} \kappa_{i_j}}{\log P} \leq \lim_{P\to\infty} \frac{\rho_{i_k}}{\log P}, \quad \forall k \in \{1, 2, ..., m\},$$

and then the outer bound (2.25) implies that

$$\sum_{j=1}^{m} d_{i_j} \leq \lim_{P\to\infty} \frac{\sum_{j=1}^{m} \kappa_{i_j}}{\log P} = \sum_{j=1}^{m} (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}),$$

for any cycle $(i_0, i_1, ..., i_m) \in \Pi_K$. This completes the outer bound.

Note that as we explained before, when the condition (2.8) is satisfied, the TIN region $\mathcal{P}^*$ is equal to the polyhedral TIN region $\mathcal{P}$, which is a convex polyhedron as shown in Theorem 2.2. This means that in this regime, time-sharing *cannot* help enlarge the GDoF achievable region via TIN.

2.3 Constant Gap to Capacity

In this section, we show that when condition (2.8) holds, so that TIN is GDoF-optimal, we can apply the insight gained in the GDoF study to prove that TIN can also achieve the whole channel capacity region to within a constant gap at any finite SNR. The main result of this section is mentioned in the following theorem.

Theorem 2.4. *In a K-user interference channel, where the channel strength level between transmitter i and receiver j is α_{ji}, if condition (2.8) holds, then TIN can achieve to within $\log_2(3K)$ bits of the capacity region.*

Proof. (**Converse**) Using Theorem 2.3, we obtain the following outer bounds.

$$R_i \leq \log_2(1 + P^{\alpha_{ii}}), \qquad \forall i \in \{1, 2, ..., K\} \tag{2.29}$$

$$\sum_{j=1}^{m} R_{i_j} \leq \sum_{j=1}^{m} \log_2\left(1 + P^{\alpha_{i_j i_{j+1}}} + \frac{P^{\alpha_{i_j i_j}}}{1 + P^{\alpha_{i_{j-1} i_j}}}\right), \quad \forall (i_0, i_1, ..., i_m) \in \Pi_K, \quad \forall m \in \{2, 3, ...K\}. \tag{2.30}$$

Since $P > 1$, it follows that

$$R_i \leq \log_2(1 + P^{\alpha_{ii}}) \leq \alpha_{ii} \log_2 P + 1, \quad \forall i \in \{1, 2, ..., K\} \tag{2.31}$$

$$
\begin{aligned}
\sum_{j=1}^{m} R_{i_j} &\leq \sum_{j=1}^{m} \log_2(1 + P^{\alpha_{i_j i_{j+1}}} + \frac{P^{\alpha_{i_j i_j}}}{1 + P^{\alpha_{i_{j-1} i_j}}}) \\
&< \sum_{j=1}^{m} \log_2(1 + P^{\alpha_{i_j i_{j+1}}} + \frac{P^{\alpha_{i_j i_j}}}{P^{\alpha_{i_{j-1} i_j}}}) \\
&= \sum_{j=1}^{m} \log_2(\frac{P^{\alpha_{i_{j-1} i_j}} + P^{\alpha_{i_j i_{j+1}} + \alpha_{i_{j-1} i_j}} + P^{\alpha_{i_j i_j}}}{P^{\alpha_{i_{j-1} i_j}}}) \\
&\leq \sum_{j=1}^{m} \log_2(\frac{3P^{\alpha_{i_j i_j}}}{P^{\alpha_{i_{j-1} i_j}}}) \\
&= \sum_{j=1}^{m} [(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \log_2 P + \log_2 3],
\end{aligned}
\tag{2.32}
$$

for all cycles $(i_0, i_1, ..., i_m) \in \Pi_K$, $\forall m \in \{2, 3, ...K\}$.

(**Achievability**) Consider the power control and TIN scheme, where the power allocated to each transmitter is equal to P^{r_i} ($r_i \leq 0$, $\forall i \in \{1, 2, ..., K\}$), and the achievable rate for each user is

$$
R_{i,\text{TIN}} = \log_2(1 + \frac{P^{r_i + \alpha_{ii}}}{1 + \sum_{j \neq i} P^{r_j + \alpha_{ij}}}). \tag{2.33}
$$

From the proof of Theorem 2.1, we know that under the condition (2.8), if d_i's satisfy (2.9) and (2.10), then there exist r_i's such that

$$
r_i + \alpha_{ii} - \max_{j \neq i}\{0, r_j + \alpha_{ij}\} = d_i, \quad \forall i, j \in \{1, 2, ..., K\} \tag{2.34}
$$

$$
r_i \leq 0, \quad \forall i \in \{1, 2, ..., K\}. \tag{2.35}
$$

Therefore, we can write

$$
\begin{aligned}
R_{i,\text{TIN}} &= \log_2(1 + \frac{P^{r_i + \alpha_{ii}}}{1 + \sum_{j \neq i} P^{r_j + \alpha_{ij}}}) \\
&\geq \log_2(\frac{P^{r_i + \alpha_{ii}}}{P^0 + \sum_{j \neq i} P^{r_j + \alpha_{ij}}}) \\
&\geq \log_2(\frac{P^{r_i + \alpha_{ii}}}{KP^{r_i + \alpha_{ii} - d_i}}) \\
&= d_i \log_2 P + \log_2(\frac{1}{K}).
\end{aligned}
\tag{2.36}
$$

In other words, when d_i's satisfy (2.9) and (2.10), the rates in (2.36) are always achievable by TIN, $\forall i \in \{1, ..., K\}$. Therefore, the achievable rate region by TIN includes the rate tuples $(R_{1,\text{TIN}}, R_{2,\text{TIN}}, ..., R_{K,\text{TIN}})$ satisfying

$$R_{i,\text{TIN}} \leq \alpha_{ii} \log_2 P + \log_2(\frac{1}{K}) \quad \forall i \in \{1, 2, ..., K\} \tag{2.37}$$

$$\begin{aligned} \sum_{j=1}^{m} R_{i_j,\text{TIN}} &= \sum_{j=1}^{m} [d_{i_j} \log_2 P + \log_2(\frac{1}{K})] \\ &\leq \sum_{j=1}^{m} [(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \log_2 P + \log_2(\frac{1}{K})], \end{aligned} \tag{2.38}$$

for all cycles $(i_0, i_1, ..., i_m) \in \Pi_K$, $\forall m \in \{2, 3, ...K\}$.

Comparing (2.31)-(2.32) with (2.37)-(2.38), we can characterize the approximate channel capacity to within a constant gap, which is only dependent on the number of users K. We can show that TIN achieves to within $\log_2(3K)$ bits of the capacity region. To this end, we need to show that each of the rate constraints in (2.37) and (2.38) is within $\log_2(3K)$ bits of its corresponding outer bound in (2.31) and (2.32), i.e., the following inequalities always hold[4],

$$\begin{aligned} \sigma_{R_i} &< \log_2(3K), \qquad \forall i \in \{1, 2, ..., K\} \\ \sigma_{\sum_{j=1}^{m} R_{i_j}} &\leq m \log_2(3K), \quad \forall (i_0, i_1, ..., i_m) \in \Pi_K, \ \forall m \in \{2, 3, ..., K\}, \end{aligned} \tag{2.39}$$

where $\sigma_{(.)}$ denotes the difference between the achievable rate in (2.37) and (2.38) and its corresponding outer bound in (2.31) and (2.32). For σ_{R_i}, we have

$$\begin{aligned} \sigma_{R_i} &= [\alpha_{ii} \log_2 P + 1] - [\alpha_{ii} \log_2 P + \log_2(\frac{1}{K})] \\ &= 1 + \log_2 K < \log_2(3K), \end{aligned} \tag{2.40}$$

[4]Notice that since in the second line of (2.32) there exists a "<", "≤" is fine for the second inequality in (2.39).

and for $\sigma_{\sum_{j=1}^{m} R_{i_j}}$, we have,

$$\begin{aligned}\sigma_{\sum_{j=1}^{m} R_{i_j}} &= \sum_{j=1}^{m}[(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j})\log_2 P + \log_2 3] - \sum_{j=1}^{m}[(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j})\log_2 P + \log_2(\frac{1}{K})] \\ &= \sum_{j=1}^{m}[\log_2 3 + \log_2 K] = m\log_2(3K).\end{aligned} \tag{2.41}$$

Since (2.40) and (2.41) hold for all ranges of i and m, the proof is complete. □

2.4 The General Achievable GDoF Region of TIN

In this section, we remove the constraint (2.8) on the channel gains, and investigate the achievable GDoF region by TIN for K-user interference channels with general channel strength levels. As we show, the TIN region $\mathcal{P}^*$ is equal to the union of multiple polyhedra, each of which is in the form of the polyhedral TIN region of a subset of the users of the network. Remarkably, the TIN region is almost the same as the polyhedral TIN region in the sense that the measure of the difference of the two sets is zero in $\mathbb{R}^K$.

We have shown that when (2.8) holds, the original TIN region $\mathcal{P}^*$ is equal to the polyhedral TIN region $\mathcal{P}$. Now, the natural question to ask is what the TIN region $\mathcal{P}^*$ is for K-user interference channels with general channel strength levels. The following theorem settles this issue.

Theorem 2.5. *In a K-user interference channel, where the channel strength level from transmitter i to receiver j is equal to α_{ji}, the achievable GDoF region through power control and treating interference as noise, denoted by $\mathcal{P}^*$, is equal to*

$$\mathcal{P}^* = \bigcup_{\mathcal{S}\subseteq\{1,\ldots,K\}} \mathcal{P}_{\mathcal{S}}, \tag{2.42}$$

where $\mathcal{P}_{\mathcal{S}}$, $\mathcal{S} \subseteq \{1, ..., K\}$, is defined as

$$\mathcal{P}_{\mathcal{S}} = \{(d_1, ..., d_K) : d_i = 0, \forall i \in \mathcal{S}, 0 \leq d_j \leq \alpha_{jj}, \forall j \in \mathcal{S}^c,$$
$$\sum_{j=1}^{m} d_{i_j} \leq \sum_{j=1}^{m} (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}), \forall (i_0, i_1, i_2, ..., i_m) \in \Pi_{\mathcal{S}^c}\},$$

and $\Pi_{\mathcal{S}^c}$ is the set of all possible cyclic sequences of all subsets of $\mathcal{S}^c$.

In words, the TIN region $\mathcal{P}^*$ is the union of the polyhedral TIN regions $\mathcal{P}_{\mathcal{S}}$, each of which corresponds to the case where the users in $\mathcal{S}$ are made silent. The proof is given in Appendix A.

As Theorem 2.5 shows, the TIN region $\mathcal{P}^*$ is almost the same as the polyhedral TIN region $\mathcal{P}$ in the sense that the measure of the difference of the two sets is zero in $\mathbb{R}^K$. Furthermore, as opposed to the polyhedral TIN region, the TIN region may not be convex in general, and if time-sharing is allowed alongside with TIN, the achievable region may become substantially larger. Therefore, the above theorem also reveals that *when the sufficient condition (2.8) is violated, time-sharing may help enlarge the achievable GDoF region of TIN.*

Example 2. Consider the 3-user cyclic channel shown in Fig. 2.4. Notice that for user 3 the sufficient condition (2.8) does not hold.

First, if all the users are active, we can get the polyhedral TIN region as follows.

$$\begin{aligned}\mathcal{P}_{\emptyset} = \{(d_1, d_2, d_3) :\ & 0 \leq d_i \leq 1,\ \forall i \in \{1, 2, 3\}, \\ & d_1 + d_2 \leq 1.9,\ d_2 + d_3 \leq 1.4,\ d_1 + d_3 \leq 1.1, \\ & d_1 + d_2 + d_3 \leq 1.4\},\end{aligned} \tag{2.43}$$

which is in fact the polyhedral TIN region $\mathcal{P}$ we defined earlier.

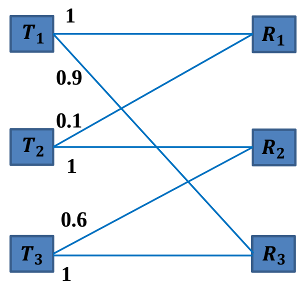


Figure 2.4: A 3-user cyclic channel, where the strength levels for each link is shown in the figure.

Then, consider the cases in which only one of the three users is made silent and hence has GDoF zero, and the other two users are active. In such cases, we only need to consider the Z-channel between the remaining two users, implying that

$$\mathcal{P}_{\{1\}} = \{(d_1, d_2, d_3) : \ d_1 = 0, \ 0 \le d_2 \le 1, \ 0 \le d_3 \le 1, \ d_2 + d_3 \le 1.4\}$$

$$\mathcal{P}_{\{2\}} = \{(d_1, d_2, d_3) : \ d_2 = 0, \ 0 \le d_1 \le 1, \ 0 \le d_3 \le 1, \ d_1 + d_3 \le 1.1\}$$

$$\mathcal{P}_{\{3\}} = \{(d_1, d_2, d_3) : \ d_3 = 0, \ 0 \le d_1 \le 1, \ 0 \le d_2 \le 1, \ d_1 + d_2 \le 1.9\}.$$

It is easy to verify that

$$\mathcal{P}_{\{1\}} \subseteq \mathcal{P}_{\emptyset}, \ \mathcal{P}_{\{2\}} \subseteq \mathcal{P}_{\emptyset},$$

but

$$\mathcal{P}_{\{3\}} \nsubseteq \mathcal{P}_{\emptyset}.$$

For instance, the GDoF tuple $(1, 0.9, 0) \in \mathcal{P}_{\{3\}}$ is not in the GDoF region $\mathcal{P}_{\emptyset}$ since it violates the cycle bound $d_1 + d_2 + d_3 \le 1.4$.

Next, consider the cases in which two users are made silent.

$$\mathcal{P}_{\{2,3\}} = \{(d_1, d_2, d_3) : \ 0 \le d_1 \le 1, \ d_2 = d_3 = 0\}$$
$$\mathcal{P}_{\{1,3\}} = \{(d_1, d_2, d_3) : \ 0 \le d_2 \le 1, \ d_1 = d_3 = 0\}$$
$$\mathcal{P}_{\{1,2\}} = \{(d_1, d_2, d_3) : \ 0 \le d_3 \le 1, \ d_1 = d_2 = 0\},$$

and it can be verified that

$$\mathcal{P}_{\{2,3\}} \subseteq \mathcal{P}_{\emptyset}, \ \mathcal{P}_{\{1,3\}} \subseteq \mathcal{P}_{\emptyset}, \ \mathcal{P}_{\{1,2\}} \subseteq \mathcal{P}_{\emptyset}.$$

Finally, we have

$$\mathcal{P}_{\{1,2,3\}} = \{(d_1, d_2, d_3) : \ d_1 = d_2 = d_3 = 0\} \subseteq \mathcal{P}_{\emptyset}.$$

Therefore, the TIN region is equal to

$$\mathcal{P}^* = \mathcal{P}_{\emptyset} \cup \mathcal{P}_{\{1\}} \cup \mathcal{P}_{\{2\}} \cup \mathcal{P}_{\{3\}} \cup \mathcal{P}_{\{1,2\}} \cup \mathcal{P}_{\{2,3\}} \cup \mathcal{P}_{\{1,3\}} \cup \mathcal{P}_{\{1,2,3\}} = \mathcal{P}_{\emptyset} \cup \mathcal{P}_{\{3\}}. \tag{2.44}$$

This region is illustrated in Fig. 2.5, where the yellow region corresponds to $\mathcal{P}_{\emptyset}$ and the blue region corresponds to $\mathcal{P}_{\{3\}}$. Note that since for user 3, the sufficient condition (2.8) is violated, the polyhedral TIN region $\mathcal{P} = \mathcal{P}_{\emptyset}$ is not the whole GDoF region for this 3-user cyclic channel. Moreover, as Fig. 2.5 shows, the region $\mathcal{P}^*$ is not convex. Therefore, time-sharing between $\mathcal{P}_{\emptyset}$ and $\mathcal{P}_{\{3\}}$ can help enlarge the achievable GDoF region via TIN. $\triangle$

2.5 Numerical Analysis

In this section, we numerically compute the probability that the sufficient condition (2.8) is satisfied in a typical wireless scenario. We consider a circular cell

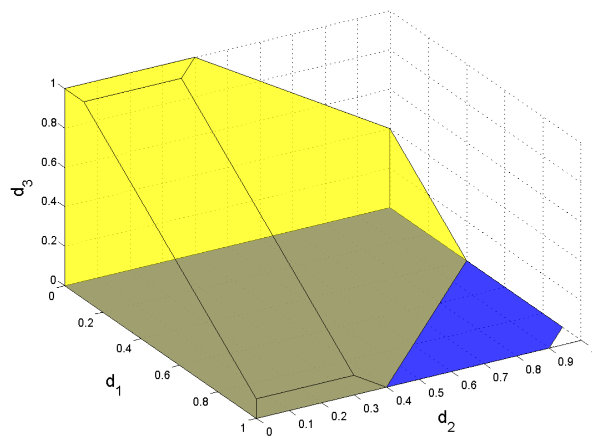
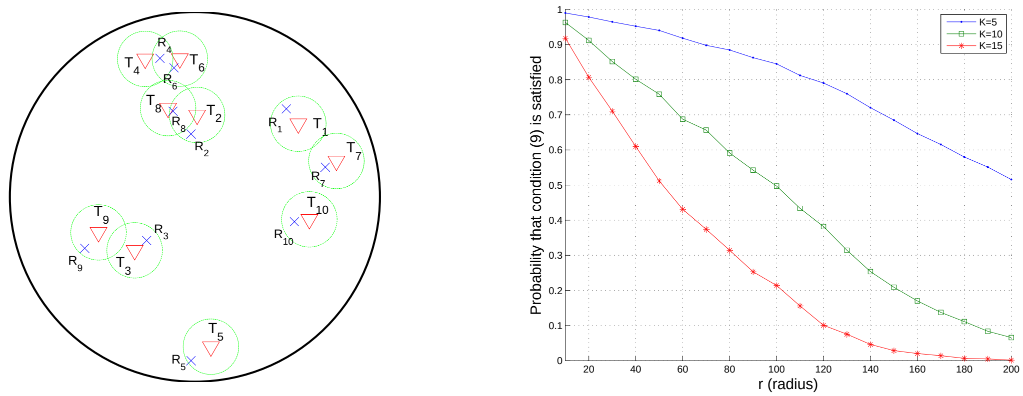


Figure 2.5: The TIN region of the network in Figure 2.4, which is the union of the yellow region ($\mathcal{P}_{\emptyset}$) and the blue region ($\mathcal{P}_{\{3\}}$).

with a radius of 1 km and place K base stations (transmitters) randomly and uniformly over the cell area. Each base station is assumed to have a coverage radius of r. In order to create a K-user interference channel with strong enough direct links, we consider K mobile receivers such that the i-th mobile receiver is located randomly and uniformly inside the coverage area of the i-th base station, $i \in \{1, 2, ..., K\}$. A realization of such a network scenario is depicted in Fig. 2.6(a).

For the channel gain values, we make use of the Erceg model [24], operating at a frequency of 2GHz and using the terrain category of hilly/light tree density. Taking the noise floor as -110 dBm, we choose the transmit power of all the base stations such that the expected value of the SNR at the boundary of their coverage area is 0 dB. Then, we randomly locate the base stations and mobile receivers according to the coverage radius r. Fig. 2.6(b) demonstrates the result of our numerical analysis.

As illustrated in this plot, the probability that the sufficient condition (2.8) for



(a) (b)

Figure 2.6: (a) A 10-user interference channel where the black circle, green circles, red triangles, and blue crosses represent the whole cell area, the coverage area of base stations, base stations (transmitters), and receivers, respectively. The coverage radius of each transmitter is taken to be $r = 100$m. (b) Effect of the coverage radius and the number of users on the probability that the sufficient condition (2.8) is satisfied.

the GDoF-optimality of TIN is satisfied decreases as the density of the network increases, either by increasing the number of users or by increasing the coverage radius of each base station. However, as a typical scenario, it is noteworthy that for the case of a 10-user interference channel with the coverage radius of 100m for each base station, the sufficient condition (2.8) is satisfied half the times. This means that with a probability of 50%, TIN is GDoF-optimal and can also achieve the whole capacity region of the network to within a constant gap. It therefore implies that the sufficient condition (2.8) can be actually satisfied in practice with a reasonably high probability, enabling optimality condition of treating interference as noise to be put into use in practice.

CHAPTER 3

SPECTRUM SHARING VIA INFORMATION-THEORETIC LINK SCHEDULING (ITLINQ)

As mentioned in Chapter 1, our basic goal is to define a *theoretically-justified* criteria to find subsets of users in a wireless network among which the level of interference is "sufficiently" low. In this chapter, we propose our criteria for finding such subsets based on the condition for the optimality of treating interference as noise developed in the previous chapter. This leads to our novel spectrum sharing mechanism of information-theoretic link scheduling (in short, ITLinQ). We will first provide a performance guarantee of ITLinQ by quantifying the fraction of the capacity region that it is able to achieve in a specific network scenario. We will also propose a distributed way of implementing ITLinQ and show, through numerical analysis, that it demonstrates a considerable sum-rate improvement over FlashLinQ, a similar state-of-the-art spectrum sharing mechanism.[1]

3.1 Description and Analysis of the Information-Theoretic Link Scheduling Scheme

In this section, we introduce our scheduling scheme, which we call "information-theoretic link scheduling" (in short, "ITLinQ"). We start by defining the notion of "information-theoretic independent set" (in short, "ITIS") and then move forward to describe the ITLinQ scheme. Afterwards, we will consider a specific network setting and in that setting, we will characterize the frac-

[1]The remaining portion of the chapter is mainly taken from [25], coauthored by the author of this thesis.

tion of capacity region that ITLinQ is able to achieve to within a gap.

3.1.1 Description of ITIS and ITLinQ

We consider a wireless network composed of n sources $\{\mathrm{S}_i\}_{i=1}^n$ and n destinations $\{\mathrm{D}_i\}_{i=1}^n$ in which each source aims to communicate a message to its corresponding destination. All the links (i.e., source-destination pairs)[2] are considered to share the same spectrum, which gives rise to interference among all the transmissions. We assume that all the nodes (i.e., all the sources and the destinations) know how many links exist in the network and they also agree on a specific ordering of the links, where by ordering we mean a labeling of the links from 1 to n. Furthermore, we assume that the nodes are synchronous; i.e., there exists a common clock among them.

The physical-layer model of the network is considered to be the AWGN model in which each source S_i intends to send a message W_i to its corresponding destination D_i, and does so by encoding its message to a codeword X_i^k of length k and transmitting it within k time slots. There is a power constraint of $\mathbb{E}\left[\frac{1}{k}\|X_i^k\|^2\right] \leq P$ on the transmit vectors. The received signal vector of destination j will be equal to

$$Y_j^k = \sum_{i=1}^{n} h_{ji} X_i^k + Z_j^k,$$

where h_{ji} denotes the channel gain between source i and destination j, and Z_j^k denotes the additive white Gaussian noise vector at destination j with distribution $\mathcal{CN}(0, N\mathbf{I}_k)$, $\mathbf{I}_k$ being the $k \times k$ identity matrix. An example of such a network configuration is illustrated in Figure 3.1.

[2] In this chapter, we use the terms "user" and "link" interchangeably.

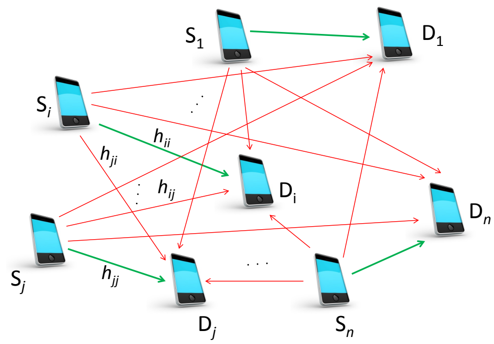


Figure 3.1: A wireless network composed of n source-destination pairs, where the green and red lines represent the direct and cross channel gains, respectively.

We assume that at each destination, all the incoming interference is treated as noise. Therefore, each source-destination pair $\mathrm{S}_i - \mathrm{D}_i$ can achieve the rate of $R_i = \log(1 + \mathrm{SINR}_i)$, where $\mathrm{SINR}_i \triangleq \frac{P|h_{ii}|^2}{\sum_{j\neq i} P|h_{ij}|^2+N}$ denotes the signal-to-interference-plus-noise ratio at destination i.

As we showed in Chapter 2, if condition (2.8) is satisfied, then treating interference as noise is information-theoretically optimal (to within a constant gap). It is easy to verify that for the aforementioned network model in this section, this condition is equivalent to the following condition

$$\mathrm{SNR}_i \geq \max_{j\neq i} \mathrm{INR}_{ij}. \max_{k\neq i} \mathrm{INR}_{ki}, \quad \forall i = 1, ..., n, \tag{3.1}$$

where $\mathrm{SNR}_i \triangleq \frac{P|h_{ii}|^2}{N}$ and $\mathrm{INR}_{ij} \triangleq \frac{P|h_{ij}|^2}{N}$ denote the signal-to-noise ratio of user i and the interference-to-noise ratio of source j at destination i, respectively.

Therefore, if we consider any subset of the source-destination pairs in a wireless network and show that condition (3.1) is satisfied in that subset, then we know that TIN is information-theoretically optimal in that subset of the users (to within a constant gap). This means that the interference is at a sufficiently low level in this subnetwork that makes it suitable to call such a subset an "information-theoretic independent subset". More formally, we have the following definition.

Definition 3.1 (ITIS)**.** In a wireless network of n users, a subset of the users $\mathcal{S} \subseteq \{1, ..., n\}$ is called an *information-theoretic independent set* (in short, ITIS) if for any user $i \in \mathcal{S}$,

$$\mathrm{SNR}_i \geq \max_{j \in \mathcal{S} \setminus \{i\}} \mathrm{INR}_{ij} . \max_{k \in \mathcal{S} \setminus \{i\}} \mathrm{INR}_{ki}. \tag{3.2}$$

As it is clear, the difference between such a concept and the regular notion of an independent set lies in the fact that in the latter case, the interference between any pair of users should be below a certain threshold (e.g., noise level), whereas in the former case, the interference between all of the users is at such a low level (determined by condition (3.1)) that makes it (to within a constant gap) information-theoretically optimal to treat all the interference as noise. Based on the concept of ITIS, we define our scheduling scheme as follows.

Definition 3.2 (ITLinQ)**.** The information-theoretic link scheduling (in short, ITLinQ) scheme is a spectrum sharing mechanism which at each time, schedules the sources in an information-theoretic independent set (ITIS) to transmit simultaneously. Moreover, all the destinations will treat their incoming interference as noise.

***Remark*.** In order to gain more intuition about the information theoretic independent sets, one can consider a simple sufficient condition for the scheduling condition in (3.2). It is easy to verify that a subset of users $\mathcal{S}$ form an ITIS if for any user $i \in \mathcal{S}$,

$$\mathrm{INR}_{ij} \leq \sqrt{\mathrm{SNR}_i}\,, \quad \mathrm{INR}_{ji} \leq \sqrt{\mathrm{SNR}_i}\,, \ \forall j \in \mathcal{S} \setminus \{i\}.$$

In fact, this condition compares the *ratio* between the INR and SNR values *in dB scale* with a fixed threshold of $\frac{1}{2}$. This is the main distinction of this condition compared to the conditions used in FlashLinQ, in which the *difference* between the INR and SNR values in dB scale is compared with a fixed threshold. We will use this sufficient condition later in this chapter for both the capacity analysis and the distributed implementation of the ITLinQ scheme.

In Section 3.2, we will show how to implement the ITLinQ scheme in a distributed way. However, for now, we will focus on characterizing the fraction of the capacity region that ITLinQ is able to achieve in a specific network setting.

3.1.2 Capacity Analysis of the ITLinQ Scheme

In this section, we analyze the fraction of the capacity region that the ITLinQ scheme can achieve to within a gap in a network with a large number of users. We consider a network in which the sources are placed uniformly and independently inside a circle of radius R, and each destination D_i is assumed to be located within a distance $r_n = r_0 n^{-\beta}$, $\beta > 0$, of its corresponding source S_i. This implies that the destination nodes get closer and closer to their corresponding source nodes as the number of users increases. Moreover, we assume that each channel gain is a deterministic function of the distance between its corresponding source and destination. In fact, we consider the path-loss model for the channel gains in which the channel gain at a distance r is deterministically equal to $h_0 r^{-\alpha}$, where α denotes the path-loss exponent.

For such a network and channel model, we have the following theorem (which will be proved later in this section) that presents a guarantee on the fraction of the capacity region that can be achieved by the ITLinQ scheme.

Theorem 3.1. *For sufficiently large number of users ($n \to \infty$) in the above model, the ITLinQ scheme can almost-surely achieve a fraction λ of the capacity region within a gap of k bits, where*

$$\begin{cases} \lambda = \frac{\sqrt{3}\pi R^2}{2\gamma^2} n^{\beta-1} \;,\; k \le \frac{\sqrt{3}\pi R^2}{2\gamma^2} \frac{\log 3n}{n^{1-\beta}} & \text{if } 0 < \beta < 1 \\ \lambda = \frac{\ln(\ln n)}{\ln n} \;,\; k \le \log(\ln n) & \text{if } \beta = 1 \\ \lambda = \frac{1}{\left\lfloor \left| \frac{1}{\beta-1} \right| + \frac{1}{2} \right\rfloor + 1} \;,\; k \le \frac{\log 3n}{\left\lfloor \left| \frac{1}{\beta-1} \right| + \frac{1}{2} \right\rfloor + 1} & \text{if } \beta > 1 \end{cases},$$

in which $\gamma = \sqrt[2\alpha]{\frac{P}{N} h_0 r_0^\alpha}$ is a constant independent of n.

Figure 3.2 illustrates the impact of the maximum source-destination distance decreasing rate on the fraction of the capacity region that can be achieved by the ITLinQ scheme. If the maximum source-destination distance is proportional to $n^{-\beta}$ such that $0 < \beta < 1$, then the ITLinQ scheme is capable of asymptotically achieving a fraction proportional to $\frac{1}{n^{1-\beta}}$ of the capacity region, within a vanishing gap. However, if the maximum source-destination distance scales as n^{-1}, then the achievable fraction of the capacity region decreases as $\frac{\ln(\ln n)}{\ln n}$ which declines much slower than the previous case. In this case, the gap increases very slowly with respect to n. Finally, in the case that the maximum distance between each source and its corresponding destination scales faster than n^{-1}, we can achieve at least a *constant* fraction of the capacity region for asymptotically large number of users which is a considerable improvement, whereas the gap is increasing with the number of users. This matches the natural intuition that the closer the destinations are located to their corresponding sources, the more the signal-to-interference-plus-noise ratio and the higher the fraction of the capacity that can be achieved by the ITLinQ scheme. Also, as a baseline, we have included the fraction of the capacity region that TDMA and independent set scheduling can achieve, which is $\frac{1}{n}$ for both schemes.[3]

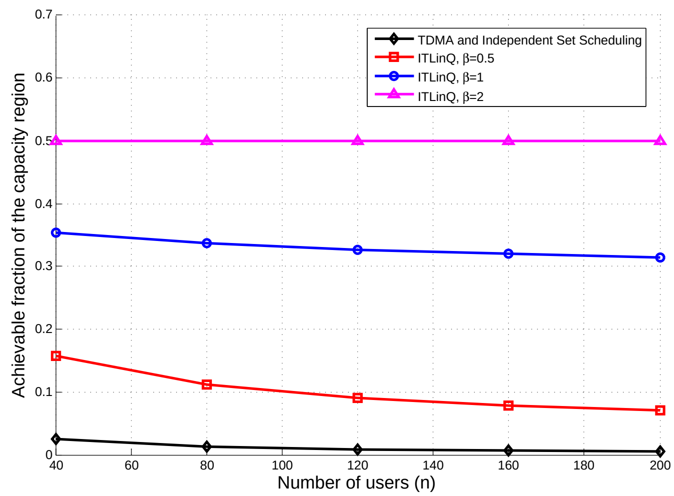


Figure 3.2: Comparison of the guaranteed achievable fraction of capacity region by the ITLinQ scheme in different regimes with TDMA and independent set scheduling.

As an immediate application of the theorem, we can consider the model in which all the n source and the n destination nodes are located uniformly and independently within a circular area of radius R, and each destination gets associated with its *closest* source. The sources and destinations are then indexed as $\{\mathrm{S}_i\}_{i=1}^n$ and $\{\mathrm{D}_i\}_{i=1}^n$, respectively. We will refer to this model as the closest access point (closest-AP) selection model. For such a model, we present the following corollary which will be proved in Appendix B.

Corollary 3.1. *For the closest-AP selection model, the ITLinQ scheme can almost-surely achieve a fraction* $\lambda = \frac{\sqrt{3}\pi R^2}{2\gamma^2} n^{\beta-1}$ *of the capacity region to within a gap of* $k \leq \frac{\sqrt{3}\pi R^2}{2\gamma^2} \frac{\log 3n}{n^{1-\beta}}$ *for any* $\beta < \frac{1}{2}$*, when* $n \to \infty$*.*

[3]The achievable fraction of the capacity region by independent set scheduling was derived through numerical analysis.

Proof of Theorem 3.1

In order to characterize the fraction of the capacity region that ITLinQ is able to achieve and prove Theorem 3.1, we seek to find the minimum number of information-theoretic independent sets which cover all the users and we will then do time-sharing among these subsets. More precisely, if we denote the set of all the information-theoretic independent subsets of a network composed of n source-destination pairs by $\mathcal{S}_n$, then we are interested in the minimum-cardinality subset of $\mathcal{S}_n$ whose members cover all the users; i.e., their union is equal to the set of all the users $\{1, ..., n\}$. Denote such a subset by $\mathcal{S}_n^*$ and let $\kappa_n = |\mathcal{S}_n^*|$. We will show that time-sharing among these κ_n information-theoretic independent sets can achieve the fractions of the capacity region mentioned in Theorem 3.1. As the first step of the proof, we characterize the achievable fraction of the capacity region by the ITLinQ scheme and its gap with respect to the random variable κ_n in the following lemma.

Lemma 3.1. *The ITLinQ scheme can achieve a fraction $\frac{1}{\kappa_n}$ of a network composed of n source-destination pairs to within a gap of $\frac{\log 3n}{\kappa_n}$.*

Proof. Consider any rate tuple $(R_1, ..., R_n)$ inside the capacity region of the network and consider any ITIS $\mathcal{U} \in \mathcal{S}_n^*$. From the result in [18], since TIN is information-theoretically optimal in $\mathcal{U}$ (to within a constant gap), the rate tuple $(\bar{R}_{1,\mathcal{U}}, ..., \bar{R}_{n,\mathcal{U}})$ is achievable in the $\frac{1}{\kappa_n}$ fraction of time which is allocated to $\mathcal{U}$, where

$$\bar{R}_{i,\mathcal{U}} = \begin{cases} R_i - \log 3|\mathcal{U}| & i \in \mathcal{U} \\ 0 & i \notin \mathcal{U} \end{cases}.$$

Therefore, the rate achieved by any user $i \in \{1, ..., n\}$ in the network through the ITLinQ scheme, denoted by $R_{i,ITLinQ}$, can be lower bounded as

$$\begin{aligned} R_{i,ITLinQ} &= \frac{1}{\kappa_n} \sum_{\mathcal{U} \in \mathcal{S}_n^*} \bar{R}_{i,\mathcal{U}} \\ &= \frac{1}{\kappa_n} \sum_{\mathcal{U} \in \mathcal{S}_n^* : i \in \mathcal{U}} (R_i - \log 3|\mathcal{U}|) \\ &\geq \frac{1}{\kappa_n} (R_i - \log 3n) \qquad (3.3) \\ &= \frac{1}{\kappa_n} R_i - \frac{\log 3n}{\kappa_n}, \end{aligned}$$

where (3.3) follows from the fact that the subsets in $\mathcal{S}_n^*$ cover all the users $\{1, ..., n\}$ and that for every $\mathcal{U} \in \mathcal{S}_n^*$, we have $|\mathcal{U}| \leq n$. This completes the proof. □

Therefore, to find an achievable fraction of the capacity region by the ITLinQ scheme, we need to find an upper bound on κ_n, that is the minimum number of information-theoretic independent subsets which cover all of the users. One way to find such an upper bound is to restrict the TIN-optimality condition in (3.1). In other words, we need to find another condition that implies condition (3.1), but is more restricted and more tractable than (3.1). Imposing such a restricted sufficient condition will reduce the number of information-theoretic independent subsets, hence leading to an upper bound on $\mathbb{E}[\kappa_n]$. To this end, we present Lemma 3.2. In the following, we denote the distance between source i and destination j by $d_{\mathrm{S}_i\mathrm{D}_j}$ and the distance between sources i and j by $d_{\mathrm{S}_i\mathrm{S}_j}$, $\forall i, j$.

Lemma 3.2. *If in a network of n source-destination pairs within the framework of the model in Section 3.1.2, the distance between* S_i *and* S_j *satisfies* $d_{\mathrm{S}_i\mathrm{S}_j} > \gamma n^{-\beta/2} + r_0 n^{-\beta}$, *then*

$$\max((\mathrm{INR}_{ji})^2, (\mathrm{INR}_{ij})^2) < \min(\mathrm{SNR}_i, \mathrm{SNR}_j).$$

Proof. Based on the model considered in Section 3.1.2, we know that $d_{\mathrm{S}_i\mathrm{D}_i} \leq r_0 n^{-\beta}$ and $d_{\mathrm{S}_i\mathrm{D}_j} \leq r_0 n^{-\beta}$. Moreover, from the triangle inequality, we will have $d_{\mathrm{S}_i\mathrm{D}_j} \geq d_{\mathrm{S}_i\mathrm{S}_j} - d_{\mathrm{S}_j\mathrm{D}_j} > \gamma n^{-\beta/2}$. Similarly, we have $d_{\mathrm{S}_j\mathrm{D}_i} > \gamma n^{-\beta/2}$. Therefore, we can get

$$\mathrm{SNR}_i = \frac{P}{N} h_0 d_{\mathrm{S}_i\mathrm{D}_i}{}^{-\alpha} \geq \frac{P}{N} h_0 \left(r_0 n^{-\beta}\right)^{-\alpha} = \frac{P}{N} h_0 r_0{}^{-\alpha} n^{\alpha\beta}, \tag{3.4}$$

and

$$\mathrm{INR}_{ji} = \frac{P}{N} h_0 d_{\mathrm{S}_i\mathrm{D}_j}{}^{-\alpha} < \frac{P}{N} h_0 \left(\gamma n^{-\beta/2}\right)^{-\alpha} = \frac{P}{N} h_0 \gamma^{-\alpha} n^{\alpha\beta/2}. \tag{3.5}$$

Combining (3.4) and (3.5), we will have

$$(\mathrm{INR}_{ji})^2 < \left(\frac{P}{N} h_0\right)^2 \gamma^{-2\alpha} n^{\alpha\beta} = \frac{P}{N} h_0 r_0{}^{-\alpha} n^{\alpha\beta} \leq \mathrm{SNR}_i, \tag{3.6}$$

and likewise, we can show that

$$(\mathrm{INR}_{ij})^2 < \mathrm{SNR}_i. \tag{3.7}$$

Combining (3.6) with (3.7) yields $\max((\mathrm{INR}_{ji})^2, (\mathrm{INR}_{ij})^2) < \mathrm{SNR}_i$. By symmetry, we will also have $\max((\mathrm{INR}_{ji})^2, (\mathrm{INR}_{ij})^2) < \mathrm{SNR}_j$. This completes the proof. $\square$

Consequently, Lemma 3.2 implies that there exists a threshold distance of $d_{th,n} = \gamma n^{-\beta/2} + r_0 n^{-\beta}$ such that if the distance between two sources is greater than this threshold, the corresponding pair of users are considered to be information-theoretically independent; i.e., the interference they cause on each other is at a sufficiently low level that it is information-theoretically optimal to treat it as noise (to within a constant gap).

Therefore, given an network of n source-destination pairs with nodes spread as mentioned in the model in the beginning of Section 3.1.2, we can build a corresponding undirected graph $G_n = (V_n, E_n)$ where $V_n = \{1, ..., n\}$ is the set of vertices and $(i, j) \in E_n$ if and only if $d_{S_iS_j} \leq d_{th,n}$; i.e., two nodes are connected together if and only if the distance between their sources is no larger than the threshold distance $d_{th,n}$. We call the resultant graph G_n the *information-theoretic conflict graph* of the original network. Clearly, this graph is a random geometric graph [26].

To return to our original problem, note that we needed to find an upper bound on κ_n. The following lemma provides such an upper bound.

Lemma 3.3. *For a large number of users (as $n \to \infty$), $\kappa_n \leq \chi(G_n)$ where $\chi(.)$ denotes the chromatic number.*

Proof. The chromatic number of G_n is the smallest number of colors that can be assigned to all of the nodes of G_n such that no two adjacent nodes have the same color. Therefore, considering the subsets of the users which receive the same color, $\chi(G_n)$ is the minimum number of subsets of the users which cover all the users and each of which consist of users whose sources have distance larger than $d_{th,n}$. From Lemma 3.2, it is easy to show that if for three distinct users i, j, k, all the pairwise source distances are larger than $d_{th,n}$, then we will have that all the squared INR's within the subnetwork consisting of users $\{i, j, k\}$ are less than all the SNR's. Extending this argument, we can see that all the independent subsets of G_n automatically satisfy the TIN-optimality condition of (3.1) as $n \to \infty$, and hence are also information-theoretic independent subsets. Therefore, κ_n, which denotes the minimum number of information-theoretic independent subsets that cover all the users, can be no more than $\chi(G_n)$, the chromatic number of G_n. □

Therefore, the final step is to characterize the asymptotic distribution of $\chi(G_n)$. This is done in the following lemma.

Lemma 3.4. *For the information-theoretic conflict graph G_n, $\chi(G_n)$ exhibits the following behavior as $n \to \infty$:*

- *If $0 < \beta < 1$, then $\frac{\chi(G_n)}{n^{1-\beta}} \xrightarrow{a.s.} \frac{2\sqrt{3}}{3\pi R^2}\gamma^2$.*
- *If $\beta = 1$, then $\frac{\chi(G_n)}{\ln n/\ln(\ln n)} \xrightarrow{a.s.} 1$.*
- *If $\beta > 1$, then $\mathbb{P}\left(\chi(G_n)\longrightarrow \left\lfloor \left|\frac{1}{\beta-1}\right| + \frac{1}{2}\right\rfloor \text{ or } \chi(G_n)\longrightarrow \left\lfloor \left|\frac{1}{\beta-1}\right| + \frac{1}{2}\right\rfloor + 1\right) = 1$.*

Proof. Since the information-theoretic conflict graph G_n is a random geometric graph with threshold distance $d_{th,n} = \gamma n^{-\beta/2} + r_0 n^{-\beta}$ and the nodes are distributed in $\mathbb{R}^2$, we can directly directly make use of the results of Theorem 1.1 in [26]. We will have the following cases:

- If $0 < \beta < 1$, then $nd_{th,n}^2 = \gamma^2 n^{1-\beta} + r_0^2 n^{1-2\beta} \gg \ln n$ (where $f(n) \ll g(n)$ is equivalent to $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$), and therefore we can use part (iv) of Theorem 1.1 in [26]. Note that the dominant term in $\gamma^2 n^{1-\beta} + r_0^2 n^{1-2\beta}$ is the first term, since $\beta > 0$. Also, as mentioned in [26], for the case of Euclidean norm in $\mathbb{R}^2$, we have $\delta = \frac{\pi}{2\sqrt{3}}$ and $\text{vol}(B) = \frac{4\pi}{3}$. Therefore, since the distribution of the nodes is uniform on a circle of radius R, we can get $\frac{\chi(G_n)}{n^{1-\beta}} \xrightarrow{\text{a.s.}} \frac{2\sqrt{3}}{3\pi R^2}\gamma^2$.
- If $\beta = 1$, then $nd_{th,n}^2 = \gamma^2 + r_0^2 n^{1-2\beta}$ which converges to a constant asymptotically, since $1 - 2\beta < 0$. This enables us to use part (ii) of Theorem 1.1 in [26], since $n^{-\epsilon} \ll \gamma^2 + r_0^2 n^{1-2\beta} \ll \ln n$ for all $\epsilon > 0$. Therefore, we have $\frac{\chi(G_n)}{\ln n/\ln(\ln n)} \xrightarrow{\text{a.s.}} 1$.

- If $\beta > 1$, then $nd_{th,n}^2 = \gamma^2 n^{-(\beta-1)} + r_0^2 n^{-(2\beta-1)}$, where $2\beta - 1 > \beta - 1 > 0$. Thus, we can make use of part (i) of Theorem 1.1 in [26] to get

$$\mathbb{P}\left(\chi(G_n) \longrightarrow \left\lfloor \left|\frac{1}{\beta-1}\right| + \frac{1}{2} \right\rfloor \text{ or } \chi(G_n) \longrightarrow \left\lfloor \left|\frac{1}{\beta-1}\right| + \frac{1}{2} \right\rfloor + 1\right) = 1.$$

□

The proof of Theorem 3.1 then follows immediately from Lemmas 3.1, 3.3 and 3.4 and also the fact that the continuous function $f(x) = \frac{1}{x}$ preserves almost-sure convergence (continuous mapping theorem [27]).

3.2 A Distributed Algorithm for ITLinQ and its comparison with FlashLinQ

In this section, we present a distributed algorithm for putting the ITLinQ scheme into practice in real-world networks. The algorithm is inspired by the FlashLinQ distributed algorithm [17] and its complexity is exactly at the same level as the FlashLinQ algorithm. However, as we will demonstrate through numerical analysis, it significantly outperforms FlashLinQ in a certain network scenario.

3.2.1 Description of the Distributed ITLinQ Algorithm

As mentioned in Section 3.1.1, we consider wireless networks consisting of n source-destination pairs. In each execution of the algorithm, to address the issue of fairness among the users, we first permute the users randomly and reindex them from 1 to n based on the realization of the random permutation, as also done in [17]. This new indexing of the users corresponds to a priority order of the users: user i has higher priority than user j if $i < j$, $\forall i, j \in \{1, ..., n\}$. Then, user 1 is always scheduled to transmit at the current time frame and for the remaining users, each user is scheduled if it does not cause and receive "too much" interference to and from the higher priority users. The conditions for defining the level of "too much" interference for user $j \in \{2, ...n\}$ are as follows, where η is a design parameter:

- At D_j, the following conditions must be satisfied:

$$\mathrm{INR}_{ji} \leq \mathrm{SNR}_j^{\eta}, \ \forall i < j, \tag{3.8}$$

  which imply that destination j does not receive too much interference from higher-priority users.

- At S_j, the following conditions must be satisfied:

$$\mathrm{INR}_{ij} \leq \mathrm{SNR}_j^{\eta}, \ \forall i < j, \tag{3.9}$$

  which imply that source j does not cause too much interference at higher-priority users.

As it is clear, there are two major differences here with respect to the FlashLinQ scheduling conditions: The first difference is that instead of considering the raw fraction $\mathrm{SIR} = \frac{\mathrm{SNR}}{\mathrm{INR}}$, here we are considering an exponent for the SNR term, which is completely inspired by the condition for the optimality of TIN (3.1). The second difference is that in condition (3.9), the outgoing interference of each user is compared to *its own* SNR rather than other users' SNR's. This is also inspired by the TIN-optimality condition (3.1).

In fact, if the parameter η is set to $\eta = 0.5$, then conditions (3.8) and (3.9) imply that the TIN-optimality condition (3.1) is satisfied at user j. This means that user j can safely be added to the information-theoretic independent subset of higher priority users and get scheduled to transmit in the current time frame. This algorithm, therefore, seeks to find the largest possible information-theoretic independent subset based on the priority ordering of the users.

However, it is clear that selecting $\eta = 0.5$ might be too pessimistic and restrictive, and may prevent some users which cause and receive low levels of interference from being scheduled. Therefore, we will leave this variable as a design parameter, and as we will see in the next section, tuning this parameter can indeed improve the achievable sum-rate by this scheduling algorithm.

The remaining question is: How can the sources and destinations check whether their pertinent conditions are satisfied? This can be done by a simple signaling mechanism which is inspired by the FlashLinQ algorithm [17] and is a two-phase process, in each of which we assume that each user uses its own frequency band and transmissions are interference-free:

- In the first phase, all the sources transmit signals at their full power P. The destinations will then receive their own desired signals and also all the interfering signals in separate frequency bands. Afterwards, the destinations estimate their received SNR's and INR's and check if their desired conditions (3.8) are satisfied. This phase is the same as that of the FlashLinQ algorithm [17].

- In the second phase, contrary to the "inverse power echo" mentioned in the FlashLinQ algorithm [17], the destinations also transmit signals at the same power level P of the sources. Similar to the first phase, in this phase all the sources can estimate the value of their desired SNR's and INR's in order to verify the validity of condition (3.9).

Remark. Clearly, power control at the transmitters may lead to an improvement in the performance of the scheme. However, due to the complication in implementing power control among the users in a distributed way, we disregard it in our scheme and use full power at all the transmitters. See e.g. [28] on power control algorithms in D2D underlaid cellular networks.

As it is obvious, the complexity of our distributed signaling mechanism is completely comparable to that of the FlashLinQ algorithm.

3.2.2 Performance Comparison of the distributed ITLinQ and FlashLinQ

In this section, we will illustrate the performance of our distributed algorithm and compare it with the FlashLinQ algorithm through numerical analysis. We drop n links randomly in a $1km \times 1km$ square. The length of each link, which is the distance between its corresponding source and destination, is taken to be a uniform random variable in the interval $[0, 40m]$. As in [17], we use the carrier frequency of 2.4 GHz and a bandwidth of 5 MHz. The noise power spectral density is considered to be -174 dBm/Hz. The transmit power is set to 20 dBm. Moreover, the channel follows the LoS model in ITU-1411 with antenna heights of 1.5m as in [17], alongside with a log-normal shadowing with standard deviation of 10 dB. The antenna gain per device is taken to be -2.5 dB and the noise figure is assumed to be 7 dB.

Figure 3.3 demonstrates the sum-rate achievable by the distributed ITLinQ scheme for different values of η and its comparison to FlashLinQ.

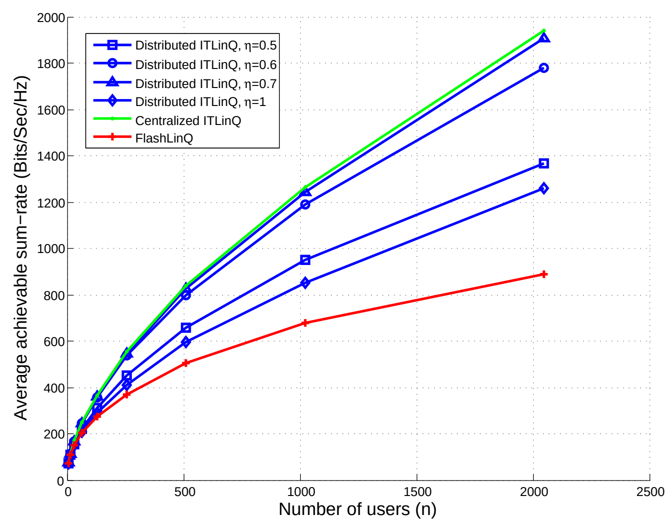


Figure 3.3: Performance comparison of distributed ITLinQ and centralized ITLinQ with FlashLinQ.

As a benchmark, we have also included a centralized version of the ITLinQ scheme in which we assume that each user i ($i = 1, ..., n$) can compute the value of $\mathrm{INR}_{ij}.\mathrm{INR}_{ki}$, $\forall j, k \neq i$. As the figure illustrates, tuning the parameter η can lead to considerable gains over FlashLinQ.[4] For the case of $\eta = 0.5$, in which conditions (3.8) and (3.9) are sufficient for the optimality of TIN (to within a constant gap), distributed ITLinQ exhibits over 50% gain over FlashLinQ for 2048 users. Interestingly, setting $\eta = 0.7$ results in more than 110% gain over FlashLinQ for 2048 users. However, as we increase η to 1, more and more users get sched-

[4]In fact, for simulation purposes, we also consider a second tuning parameter M which adds more flexibility to our scheme. With the addition of this variable, conditions (3.8) and (3.9) will change to $\mathrm{INR}_{ji} \leq M\mathrm{SNR}_j^{\eta}$, $\forall i < j$ and $\mathrm{INR}_{ij} \leq M\mathrm{SNR}_j^{\eta}$, $\forall i < j$, respectively.

uled which results in a degradation in the overall performance. Moreover, it is clear that our distributed algorithm can almost achieve the same sum-rate as the centralized ITLinQ scheme mentioned above, showing that the decentralization loss of the scheme is negligible.

Moreover, in the same setting, we also study the cumulative distribution function (CDF) of the average link rate in a network of 128 users. The result is depicted in Figure 3.4.

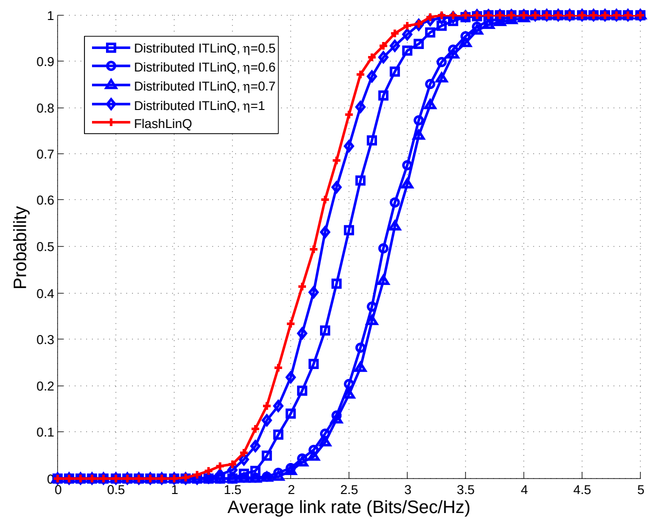


Figure 3.4: Comparison of the cumulative distribution function of the average link rate achieved by distributed ITLinQ and FlashLinQ.

Again, the same trend occurs in this plot, showing that distributed ITLinQ, especially for the value of $\eta = 0.7$, can result in considerable uniform gain over the average link rate achievable by FlashLinQ. For instance, the probability that the average link rate is greater than 2.5 bits/sec/Hz in FlashLinQ is around 0.2, while this probability is around 0.8 for the case of distributed ITLinQ.

CHAPTER 4

STUDYING THE IMPACT OF NETWORK TOPOLOGY ON INTERFERENCE NETWORKS WITH NO CSIT

In this chapter, we study the fundamental limits of interference management in partially-connected interference networks in which the transmitters have no information about the channel gain values (referred to as no-CSIT assumption). In particular, some of the channel gains in the network are allowed to be below the noise level, which makes the corresponding transmitter and receiver get disconnected. This results in an arbitrary network topology which is represented by the adjacency matrix of the network connectivity graph. In the scenario where the transmitters know the network topology but are unaware of the channel gain values, we will characterize the symmetric degrees-of-freedom (DoF) of the network. To this end, we will derive new linear-algebraic outer bounds and graph-theoretic inner bounds on this value. For the inner bounds, we present a new achievable scheme of "structured repetition coding", which goes beyond the treating interference as noise scheme that we have studied in the previous chapters. Finally, we will show that our inner and outer bounds meet each other, hence characterizing the symmetric DoF, in two distinct network scenarios through numerical analysis. We will also quantify the gain of structured repetition coding over two benchmark interference management schemes that we will introduce in this chapter.[1]

[1]The remaining portion of the chapter is mainly taken from [29, 30], coauthored by the author of this thesis.

4.1 Problem Formulation and Notations

A K-user interference network ($K \in \mathbb{N}$) is defined as a set of K transmitter nodes $\{T_i\}_{i=1}^K$ and K receiver nodes $\{D_i\}_{i=1}^K$. To model propagation path loss and interference topology, we consider a similar model to [31] in which the network is partially connected represented by the adjacency matrix $\mathbf{M} \in \{0,1\}^{K \times K}$, such that $\mathbf{M}_{ij} = 1$ if and only if transmitter T_i is connected to receiver D_j (i.e. D_j is in the coverage radius of T_i). We assume there exist direct links between each transmitter T_i and its corresponding receiver D_i (i.e. $\mathbf{M}_{ii} = 1$, $\forall i \in [1 : K]$, where we use the notation $[1 : m]$ to denote $\{1, 2, ..., m\}$ for $m \in \mathbb{N}$). We also define the set of interfering nodes to receiver D_j as $\mathcal{IF}_j := \{i : \mathbf{M}_{ij} = 1,\ i \neq j\}$.

The communication is time-slotted. At each time slot l ($l \in \mathbb{N}$), the transmit signal of transmitter T_i is denoted by $X_i[l] \in \mathbb{C}$ and the received signal of receiver D_j is denoted by $Y_j[l] \in \mathbb{C}$ given by

$$Y_j[l] = g_{jj}[l]X_j[l] + \sum_{i \in \mathcal{IF}_j} g_{ij}[l]X_i[l] + Z_j[l],$$

where $Z_j[l] \sim \mathcal{CN}(0,1)$ is the additive white Gaussian noise and $g_{ij}[l]$ is the channel gain from transmitter T_i to receiver D_j at time slot l. If transmitter T_i is not connected to receiver D_j (i.e. $\mathbf{M}_{ij} = 0$), then $g_{ij}[l]$ is assumed to be identically zero at all times. We assume that the non-zero channel gains (i.e. $g_{ij}[l]$'s s.t. $\mathbf{M}_{ij} = 1$) are independent and identically distributed (with a continuous distribution $f_G(g)$) through time and also across the users, and are also independent of the transmit symbols. The distribution $f_G(g)$ needs to satisfy three regularity conditions: $\mathbb{E}[|g|^2] < \infty$, $f_G(g) = f_G(-g), \forall g \in \mathbb{C}$, and $\exists f_{max}$ s.t. $f_{|G|}(r) \leq f_{max}, \forall r \in \mathbb{R}^+$, where $f_{|G|}(.)$ is the distribution of $|g|$. The noise terms are also assumed i.i.d. among the users and the time slots, and also independent of the transmit symbols and channel gains.

It is assumed that the transmitters $\{\mathrm{T}_i\}_{i=1}^K$ are only aware of the connectivity pattern of the network (or the network topology), represented by the adjacency matrix $\mathbf{M}$, and also the distribution f_G of the non-zero channel gains; i.e. the transmitters only know which users are interfering to each other and they also know the statistics of the channel gains, not the actual gains of the links. In this chapter, we refer to this assumption as no channel state information at the transmitters (no CSIT). As for the receivers $\{\mathrm{D}_j\}_{j=1}^K$, we assume that they are aware of the adjacency matrix $\mathbf{M}$ and the channel gain realizations of their incoming links. In other words, receiver D_j is aware of $\mathbf{M}$ and $g_{ij}[l]$, $\forall i \in \{j\} \cup \mathcal{IF}_j$, $\forall l$.

In this network, every transmitter T_i intends to deliver a message W_i to its corresponding receiver D_i. The message W_i is encoded to a vector $X_i^n = [X_i[1]\ X_i[2]\ \ldots\ X_i[n]]^T \in \mathbb{C}^n$ through an encoding function $e_i(W_i|\mathbf{M}, f_G)$; i.e. transmitters use their knowledge of network topology and the distribution of the channel gains to encode their messages. There is also a transmit power constraint $\mathbb{E}\left[\frac{1}{n}\|X_i^n\|^2\right] \le P$, $\forall i \in [1:K]$. This encoded vector is transmitted within n time slots through the wireless channel to the receivers. Each receiver D_j receives the vector $Y_j^n = [Y_j[1]\ Y_j[2]\ \ldots\ Y_j[n]]^T$ and uses a decoding function $e_j'(Y_j^n|\mathbf{M}, \mathcal{G}_j^n)$ to recover its desired message W_j. Here, $\mathcal{G}_j^n := \{g_{ij}^n : i \in [1:K]\}$ where $g_{ij}^n := [g_{ij}[1]\ g_{ij}[2]\ \ldots\ g_{ij}[n]]^T$ denotes the vector of the channel gain realizations from transmitter T_i to receiver D_j during n time slots. We also denote the set of all channel gains in all time slots by $\mathcal{G}^n = \{\mathcal{G}_1^n, ..., \mathcal{G}_K^n\}$.

The rate of transmission for user i is denoted by $R_i(P) := \frac{\log|W_i(P)|}{n}$ where $|W_i(P)|$ is the size of the message set of user i and we have explicitly shown the dependence of W_i on P. Denoting the maximum error probability at the receivers by

$\Pr_e(P) = \max_{j\in[1:K]} \Pr\left[W_j(P) \neq e'_j(Y_j^n|\mathbf{M}, \mathcal{G}_j^n)\right]$, a rate tuple $(R_1(P), ..., R_K(P))$ is said to be achievable if $\Pr_e(P)$ goes to zero as n goes to infinity.

In this chapter, the considered metric is the symmetric degrees-of-freedom (DoF) metric, which is defined as follows. If a rate tuple $(R_1(P), ..., R_K(P))$ is achievable and we let $d_i = \lim_{P\to\infty} \frac{R_i(P)}{\log(P)}$, then the DoF tuple of $(d_1, ..., d_K)$ is said to be achievable. The symmetric degrees-of-freedom d_{sym} is defined as the supremum d such that the DoF tuple $(d, ..., d)$ is achievable.

Therefore, the main problem we are going to address in this chapter is that given a K-user interference network with adjacency matrix $\mathbf{M}$ (which is known by every node in the network) and channel gains distribution f_G, what the symmetric degrees-of-freedom d_{sym} is, under no-CSIT assumption. We will start by presenting our outer bounds on d_{sym} in the next section.

4.2 Outer Bounds on d_{sym}

In this section, we will present our outer bounds for the symmetric DoF of K-user interference networks. To this end, we provide two types of outer bounds and we will motivate each outer bound through an introductory example. The main idea in both of the outer bounds is to create a set of signals by which we can sequentially decode the messages of all the users with a finite number of bits provided by a genie. This set of signals corresponds to a matrix called a *generator*. We will show systematically that for any network topology, there are some linear algebraic conditions that a matrix should satisfy to be called a generator. Therefore, our outer bounds rely highly on the topology of the network graph and the goal is to algebraically explain how these bounds are derived. The first converse generally states that the number of signals corresponding to a generator is an upper bound for the sum degrees-of-freedom of the network. However, the second converse enhances the first one, showing that there may be tighter upper bounds on the sum degrees-of-freedom due to the specific topology of the network.

4.2.1 Upper Bounds Based on the Concept of Generators

We start by presenting our first outer bound through the notion of generators. The main idea of this outer bound is presented in Example 3. Before starting the example, we need to define some notation.

- If $\mathcal{S} \subseteq [1 : K]$ is a subset of users in a K-user interference network with adjacency matrix $\mathbf{M}$, then $\mathbf{M}^{\mathcal{S}}$ denotes the adjacency matrix of the corresponding subgraph and $\mathbf{I}^{|\mathcal{S}|}$ denotes the $|\mathcal{S}| \times |\mathcal{S}|$ identity matrix.
- For a general $m \times n$ matrix $\mathbf{A}$ and $\mathcal{N} \subseteq [1 : n]$, $\mathbf{A}_{\mathcal{N}}$ denotes the submatrix of $\mathbf{A}$ composed of the columns whose indices are in $\mathcal{N}$. For the sake of brevity, if $\mathcal{N} = \{i\}$, i.e. if $\mathcal{N}$ has only one member, we use $\mathbf{A}_i$ to denote the i^{th} column of $\mathbf{A}$.
- For a general matrix $\mathbf{A}$, $c(\mathbf{A})$ denotes the number of columns of $\mathbf{A}$.

We will also need the following definition.

Definition 4.1. If $\mathbf{v} \in \{0, \pm 1\}^{n \times 1}$ and $\mathcal{V}$ is a subspace of $\mathbb{R}^n$, then $\mathbf{v} \in^{\pm} \mathcal{V}$ means that there exists a vector $\tilde{\mathbf{v}}$ in $\mathcal{V}$ which is the same as $\mathbf{v}$ up to the sign of its elements; i.e.

$$\mathbf{v} \in^{\pm} \mathcal{V} \Leftrightarrow \exists \tilde{\mathbf{v}} \in \mathcal{V} \text{ s.t. } |\tilde{\mathbf{v}}_j| = |\mathbf{v}_j|, \ \forall j \in [1 : n].$$

Moreover, if $i \in [1:n]$, then $\mathbf{v} \in_i^{\pm} \mathcal{V}$ is defined as

$$\mathbf{v} \in_i^{\pm} \mathcal{V} \Leftrightarrow \exists \tilde{\mathbf{v}} \in \mathcal{V} \text{ s.t. } |\tilde{\mathbf{v}}_i| = |\mathbf{v}_i| \text{ and } \tilde{\mathbf{v}}_j(|\tilde{\mathbf{v}}_j| - |\mathbf{v}_j|) = 0, \ \forall j \in [1:n] \setminus \{i\},$$

implying that there exists a vector $\tilde{\mathbf{v}}$ in $\mathcal{V}$ whose i^{th} element is the same as the i^{th} element of $\mathbf{v}$ up to its sign, while every other element of $\tilde{\mathbf{v}}$ either equals zero or matches the corresponding element of $\mathbf{v}$ up to its sign.

Example 3. Consider the 5-user interference network in Figure 4.1. We claim that the symmetric DoF of this network with no CSIT is upper bounded by $\frac{2}{5}$.

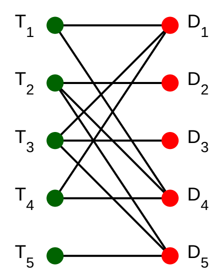


Figure 4.1: A 5-user interference network in which $d_{sym} \leq \frac{2}{5}$.

Suppose rates R_i, $i \in [1:5]$, are achievable. We define the signals

$$\begin{aligned} \tilde{Y}_1^n &= g_1^n X_1^n + g_3^n X_3^n + g_4^n X_4^n + \tilde{Z}_1^n \\ \tilde{Y}_5^n &= g_2^n X_2^n + g_3^n X_3^n + g_5^n X_5^n + \tilde{Z}_5^n, \end{aligned}$$

where $\tilde{Z}_1^n$ and $\tilde{Z}_5^n$ have the same distributions as the original noise vectors, but are independent of them and also of each other and $g_i^n = g_{ii}^n$, $i \in [1:5]$. We now show that $H(W_1, ..., W_5 | \tilde{Y}_1^n, \tilde{Y}_5^n, \mathcal{G}^n) \leq no(\log(P)) + n\epsilon_n$, which implies

$$\begin{aligned} \sum_{i=1}^{5} R_i &= \frac{1}{n} H(W_1, ..., W_5 | \mathcal{G}^n) \\ &= \frac{1}{n} \left[I(W_1, ..., W_5; \tilde{Y}_1^n, \tilde{Y}_5^n | \mathcal{G}^n) + H(W_1, ..., W_5 | \tilde{Y}_1^n, \tilde{Y}_5^n, \mathcal{G}^n) \right] \\ &\leq 2\log(P) + o(\log(P)) + \epsilon_n, \end{aligned}$$

hence $d_{sym} \leq \frac{2}{5}$. This is obtained through the following steps, which are explained intuitively here and their formal proof is discussed in the proof of Theorem 4.1 for general network topologies.

- Step 1: $H(W_1, W_5|\tilde{Y}_1^n, \tilde{Y}_5^n, \mathcal{G}^n) \leq n\epsilon_n$, due to the fact that $\tilde{Y}_1^n$ and $\tilde{Y}_5^n$ are statistically the same as Y_1^n and Y_5^n, respectively, and Fano's inequality.
- Step 2: $H(W_4|\tilde{Y}_1^n, \tilde{Y}_5^n, W_1, W_5, \mathcal{G}^n) \leq no(\log(P)) + n\epsilon_n$. This is obtained by noting that from W_5, one can create X_5^n and then by using the other terms in the conditioning, we can construct $\tilde{Y}_4^n = \tilde{Y}_1^n - \tilde{Y}_5^n + g_5^n X_5^n = g_1^n X_1^n - g_2^n X_2^n + g_4^n X_4^n + \tilde{Z}_1^n - \tilde{Z}_5^n$, which is statistically the same as Y_4^n except for a larger, but bounded, noise variance. This is because the distribution of the channel gains is symmetric around zero ($f_G(g) = f_G(-g), \forall g \in \mathbb{C}$). The desired inequality then follows, where the $n\epsilon_n$ term is due to Fano's inequality and the $no(\log(P))$ term is due to the larger noise variance, treated more formally in Lemma 4.1 which appears later.
- Step 3: $H(W_3|\tilde{Y}_1^n, \tilde{Y}_5^n, W_1, W_5, W_4, \mathcal{G}^n) \leq n\epsilon_n$, obtained by noting that from W_1 and W_4, one can create X_1^n and X_4^n and then by using the other terms in the conditioning, we can construct $\tilde{Y}_3^n = \tilde{Y}_1^n - g_1^n X_1^n - g_4^n X_4^n = g_3^n X_3^n + \tilde{Z}_1^n$, which is statistically the same as Y_3^n. The inequality then follows from Fano's inequality.
- Step 4: $H(W_2|\tilde{Y}_1^n, \tilde{Y}_5^n, W_1, W_5, W_4, W_3, \mathcal{G}^n) \leq n\epsilon_n$, obtained by noting that from W_3 and W_5, one can create X_3^n and X_5^n and then by using the other terms in the conditioning, we can construct $\tilde{Y}_2^n = \tilde{Y}_5^n - g_3^n X_3^n - g_5^n X_5^n = g_2^n X_2^n + \tilde{Z}_2^n$, which is statistically the same as Y_2^n. The inequality then follows from Fano's inequality.

Adding the above inequalities and using the chain rule for entropy yield the desired result. Therefore, starting from $\{\tilde{Y}_1^n, \tilde{Y}_5^n\}$, we created a sequence of users {1,5,4,3,2} in which we could successively generate statistically similar versions of the signals at their receivers (with a bounded difference in noise variance) by a linear combination of the signals available at each step, and at the end of the final step, we could decode the messages of all users by initially having the two signals $\{\tilde{Y}_1^n, \tilde{Y}_5^n\}$.

This process can be explained in a more systematic and linear algebraic form. Each of the signals discussed above (ignoring the noise term) can be represented as a 5×1 column vector whose i^{th} element, $i \in [1:5]$, is equal to the coefficient of $g_i^n X_i^n$ in that signal. For instance, $\tilde{Y}_1^n$ corresponds to $\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \end{bmatrix}^T$ and $\tilde{Y}_5^n$ corresponds to $\begin{bmatrix} 0 & 1 & 1 & 0 & 1 \end{bmatrix}^T$. We concatenate these two vectors so that $\{\tilde{Y}_1^n, \tilde{Y}_5^n\}$ can be represented by the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}^T. \tag{4.1}$$

Now, using the notation introduced in Definition 4.1, the successive decoding steps mentioned earlier in this example can be expressed in a linear algebraic form. In what follows, $\mathcal{S} = [1:5]$.

- Step 1 is equivalent to $\mathbf{M}_1 \in^{\pm} \mathrm{span}(\mathbf{A})$. The reason is as follows. First, note that $\mathbf{M}_1 = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \end{bmatrix}^T$ is the first column of the adjacency matrix, which corresponds to the signal received at receiver 1, namely Y_1^n (because $Y_1^n = \begin{bmatrix} g_{11}^n X_1^n & \dots & g_{51}^n X_5^n \end{bmatrix} \mathbf{M}_1 + Z_1^n$). Therefore, $\mathbf{M}_1 \in^{\pm} \mathrm{span}(\mathbf{A})$ means that by a combination of the signals $\{\tilde{Y}_1^n, \tilde{Y}_5^n\}$, we can create a statistically-similar version of Y_1^n (actually, the combination is $\tilde{Y}_1^n$ itself) . Since the distribution of the channel gains is symmetric around zero ($f_G(g) = f_G(-g), \forall g \in \mathbb{C}$), the sign of each element $g_{i1}^n X_i^n$ in Y_1^n is not important, therefore letting us use the notation developed in Definition 4.1. In the same way, we have $\mathbf{M}_5 \in^{\pm} \text{span}(\mathbf{A})$, which means that by a combination of the signals $\{\tilde{Y}_1^n, \tilde{Y}_5^n\}$, we can create a statistically-similar version of Y_5^n.

- Step 2 is equivalent to $\mathbf{M}_4 \in^{\pm} \text{span}(\mathbf{A}, \mathbf{I}_{\{1,5\}}^{|\mathcal{S}|})$. The reason is as follows. First, note that columns 1 and 5 of the identity matrix are now included since we have already decoded W_1 and W_5 in the previous step, and by having them and the channel gains, we can create the signals $g_1^n X_1^n$ and $g_5^n X_5^n$ which correspond to $\mathbf{I}_1^{|\mathcal{S}|}$ and $\mathbf{I}_5^{|\mathcal{S}|}$, respectively. Therefore, ignoring the noise terms because of their finite variance, we can create a statistically similar version of Y_4^n by having $\tilde{Y}_1^n, \tilde{Y}_5^n, W_1, W_5$ and the channel gains.

- Step 3 is equivalent to $\mathbf{M}_3 \in^{\pm} \text{span}(\mathbf{A}, \mathbf{I}_{\{1,5,4\}}^{|\mathcal{S}|})$. The reason is as follows. First, note that before this step, we have already decoded W_1, W_5 and W_4, and by having them and the channel gains, we can create the signals $g_1^n X_1^n$, $g_5^n X_5^n$ and $g_4^n X_4^n$ which correspond to $\mathbf{I}_1^{|\mathcal{S}|}$, $\mathbf{I}_5^{|\mathcal{S}|}$ and $\mathbf{I}_4^{|\mathcal{S}|}$, respectively. Therefore, we can create a statistically similar version of Y_3^n by having $\tilde{Y}_1^n, \tilde{Y}_5^n, W_1, W_5, W_4$ and the channel gains.

- Step 4 is equivalent to $\mathbf{M}_2 \in^{\pm} \text{span}(\mathbf{A}, \mathbf{I}_{\{1,5,4,3\}}^{|\mathcal{S}|})$, which means that we can create a statistically similar version of Y_2^n by having $\tilde{Y}_1^n, \tilde{Y}_5^n, W_1, W_5, W_4, W_3$ and the channel gains.

$\triangle$

Motivated by Example 3, we now formally define the notion of generators.

Definition 4.2. Consider a K-user interference network with adjacency matrix $\mathbf{M}$ and assume $\mathcal{S} \subseteq [1:K]$ is a subset of users. $\mathbf{A} \in \{\pm 1, 0\}^{|\mathcal{S}| \times r}$ ($r \in \mathbb{N}$) is called a *generator* of $\mathcal{S}$ if there exists a sequence $\Pi_{\mathcal{S}} = (i_1, ..., i_{|\mathcal{S}|})$ of the users in $\mathcal{S}$ such that

$$\mathbf{M}_{i_j}^{\mathcal{S}} \in_{i_j}^{\pm} \text{span}\,(\mathbf{A}, \mathbf{I}_{\{i_1,...,i_{j-1}\}}^{|\mathcal{S}|}),\ \forall j \in [1 : |\mathcal{S}|].$$

We use $\mathcal{J}(\mathcal{S})$ to denote the set of all generators of $\mathcal{S}$.

To gain intuition about the above definition, similar to what we mentioned in Example 3, each column of a generator $\mathbf{A}$ of $\mathcal{S}$ corresponds to a signal which is a linear combination of the transmit symbols X_i^n, $i \in \mathcal{S}$. Therefore, the number of columns of $\mathbf{A}$, denoted by $c(\mathbf{A})$, represents the number of these signals. Consequently, the spanning relationships in Definition 4.2 represent a sequence of users in which all the messages can be decoded by having $c(\mathbf{A})$ signals, as in Example 3. Also, the reason that we have used the notation $\in_{i_j}^{\pm}$ instead of $\in^{\pm}$ (which we were using in Example 3) is that intuitively, it is not necessary to generate (a statistically-similar version of) the received signal at receiver D_{i_j} *exactly*. Instead, it suffices to generate a *less-interfered* version of its received signal (by deleting some of the interference terms) and still be able to decode its message, because interference only hurts.

By having the definition of the generator in mind, we can present our first converse as follows.

Theorem 4.1. *The symmetric DoF of a K-user interference network with no CSIT is upper bounded by*

$$d_{sym} \leq \min_{\mathcal{S}\subseteq[1:K]} \min_{\mathbf{A}\in\mathcal{J}(\mathcal{S})} \frac{c(\mathbf{A})}{|\mathcal{S}|},$$

where for each $\mathcal{S} \subseteq [1 : K]$, $\mathcal{J}(\mathcal{S})$ denotes the set of all generators of $\mathcal{S}$ (Definition 4.2) and $c(\mathbf{A})$ denotes the number of columns of $\mathbf{A}$.

Before proving the theorem, we present the following lemma, which is proved in Appendix C.

Lemma 4.1. *For a discrete random variable W, continuous random vector Y^n, and two complex Gaussian noise vectors Z_1^n and Z_2^n, where each element of Z_1^n and Z_2^n are $\mathcal{CN}(0,1)$ and $\mathcal{CN}(0,N)$ random variables, respectively and all the random variables are mutually independent, if $H(W|Y^n + Z_1^n) \leq n\epsilon$, then $H(W|Y^n + Z_2^n) \leq n\epsilon + n\log(N+1)$.*

Proof of Theorem 4.1. Consider a generator of $\mathcal{S}$ denoted by $\mathbf{A}$. Without loss of generality, assume that $\mathcal{S} = [1:m]$, $c(\mathbf{A}) = m'$ ($m' \leq m$) and $\Pi_{\mathcal{S}} = (1, ..., m)$. Define $\tilde{Y}_i^n = \begin{bmatrix} g_1^n X_1^n & \dots & g_m^n X_m^n \end{bmatrix} \mathbf{A}_i + \tilde{Z}_i^n$, $i \in [1:m']$, where $g_i^n = g_{ii}^n$, $\forall i \in [1:m]$ and the noise vectors $\tilde{Z}_i^n$ have exactly the same distributions as the original noises, but are independent of them and also of each other. Suppose rates R_i, $i \in \mathcal{S}$ are achievable. Then, we will have:

$$\begin{aligned}
n\sum_{i\in\mathcal{S}} R_i &= H(W_1, ..., W_m|\mathcal{G}^n) \\
&= I(W_1, ..., W_m; \tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n|\mathcal{G}^n) + H(W_1, ..., W_m|\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n, \mathcal{G}^n) \\
&= h(\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n|\mathcal{G}^n) - h(\tilde{Z}_1^n, ..., \tilde{Z}_{m'}^n) + H(W_1, ..., W_m|\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n, \mathcal{G}^n) \\
&= h(\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n|\mathcal{G}^n) + no(\log(P)) + H(W_1, ..., W_m|\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n, \mathcal{G}^n). \qquad (4.2)
\end{aligned}$$

Now, we prove that $H(W_l|\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n, W_1, ..., W_{l-1}, \mathcal{G}^n) \leq no(\log(P)) + n\epsilon_{l,n}$ for $l \in [1:m]$. By Definition 4.2, $\mathbf{M}_l^{\mathcal{S}} \in_l^{\pm}$ span $(\mathbf{A}, \mathbf{I}_{\{1,...,l-1\}}^{|\mathcal{S}|})$, implying that there exist a vector $\tilde{\mathbf{v}} \in \mathbb{R}^{|\mathcal{S}|}$ and coefficients c_i ($i \in [1:m']$) and d_k ($k \in [1:l-1]$) such that

$$\tilde{\mathbf{v}} = \sum_{i=1}^{m'} c_i \mathbf{A}_i + \sum_{k=1}^{l-1} d_k \mathbf{I}_k^{|\mathcal{S}|} \qquad (4.3)$$

$$|\tilde{\mathbf{v}}_l| = |\mathbf{M}_{ll}^{\mathcal{S}}| = 1 \qquad (4.4)$$

$$\tilde{\mathbf{v}}_j(|\tilde{\mathbf{v}}_j| - |\mathbf{M}_{jl}^{\mathcal{S}}|) = 0, \ \forall j \in [1:m] \setminus \{l\}. \qquad (4.5)$$

Note that if $j \in \mathcal{IF}_l$, then $\mathbf{M}^{S}_{jl} = 1$ and (4.5) implies that $\tilde{\mathbf{v}}_j$ can either be equal to 0 or ± 1; i.e. $\tilde{\mathbf{v}}_j \in \{0, \pm 1\}$. On the other hand, if $j \notin \mathcal{IF}_l$, then $\mathbf{M}^{S}_{jl} = 0$ and (4.5) implies that $\tilde{\mathbf{v}}_j = 0$. Multiplying $\begin{bmatrix} g_1^n X_1^n & \dots & g_m^n X_m^n \end{bmatrix}$ by both sides of (4.3), hence, yields

$$\tilde{\mathbf{v}}_l g_l^n X_l^n + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g_j^n X_j^n = \sum_{i=1}^{m'} c_i \tilde{Y}_i^n + \sum_{k=1}^{l-1} d_k g_k^n X_k^n + \tilde{Z'}_l^n,$$

where the variance of each element of $\tilde{Z'}_l^n$ is $N_l = \sum_{i=1}^{m'} c_i^2 < \infty$. Therefore, we can write:

$$\begin{aligned} H(W_l | \sum_{i=1}^{m'} c_i \tilde{Y}_i^n + \sum_{k=1}^{l-1} d_k g_k^n X_k^n, \mathcal{G}^n) &= H(W_l | \tilde{\mathbf{v}}_l g_l^n X_l^n + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g_j^n X_j^n - \tilde{Z'}_l^n, \mathcal{G}^n) \\ &= H(W_l | \tilde{\mathbf{v}}_l g_l^n X_l^n + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g_j^n X_j^n - \tilde{Z'}_l^n, \sum_{j \in \mathcal{IF}_l} (1 - |\tilde{\mathbf{v}}_j|) g_j^n X_j^n, \mathcal{G}^n) \quad (4.6) \\ &\leq H(W_l | \tilde{\mathbf{v}}_l g_l^n X_l^n + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}'_j g_j^n X_j^n - \tilde{Z'}_l^n, \mathcal{G}^n) \quad (4.7) \\ &\leq n o(\log(P)) + n\epsilon_{l,n}, \quad (4.8) \end{aligned}$$

where (4.6) is true because, as discussed before, for all $j \in \mathcal{IF}_l$, $\tilde{\mathbf{v}}_j$ can only take the values in $\{\pm 1, 0\}$ and therefore the signals in $\sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g_j^n X_j^n$ and $\sum_{j \in \mathcal{IF}_l} (1 - |\tilde{\mathbf{v}}_j|) g_j^n X_j^n$ do not have common terms.[2] In (4.7), $\tilde{\mathbf{v}}'_j$ is defined as $\tilde{\mathbf{v}}'_j := \tilde{\mathbf{v}}_j + (1 - |\tilde{\mathbf{v}}_j|)$. Clearly $\tilde{\mathbf{v}}'_j$ can only take the values in $\{+1, -1\}$ because $\tilde{\mathbf{v}}_j \in \{\pm 1, 0\}$. Also, (4.4) implies that $\tilde{\mathbf{v}}_l \in \{+1, -1\}$. Therefore, $\tilde{\mathbf{v}}_l g_l^n X_l^n + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}'_j g_j^n X_j^n - \tilde{Z'}_l^n$ is statistically the same as Y_l^n (with a bounded difference in noise variance), because the channel gains have a symmetric distribution around zero ($f_G(g) = f_G(-g)$, $\forall g \in \mathbb{C}$). This, together with Lemma 4.1 and Fano's inequality, implies that (4.8) is correct. Hence, using the chain rule for entropy yields

$$H(W_1, ..., W_m | \tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n, \mathcal{G}^n) = \sum_{l=1}^{m} H(W_l | \tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n, W_1, ..., W_{l-1}, \mathcal{G}^n)$$

[2] If $\tilde{\mathbf{v}}_j = 0$, then $1 - |\tilde{\mathbf{v}}_j| = 1$, and if $\tilde{\mathbf{v}}_j = 1$ or $\tilde{\mathbf{v}}_j = -1$, then $1 - |\tilde{\mathbf{v}}_j| = 0$. Hence, either $\tilde{\mathbf{v}}_j$ or $1 - |\tilde{\mathbf{v}}_j|$ is non-zero, but not both.

$$\leq \sum_{l=1}^{m} no(\log(P)) + n\epsilon_{l,n}$$

$$= no(\log(P)) + n\epsilon_n,$$

which together with (4.2) implies

$$n \sum_{i \in \mathcal{S}} R_i \leq h(\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n | \mathcal{G}^n) + no(\log(P)) + n\epsilon_n \tag{4.9}$$

$$\leq nm' \log(P) + no(\log(P)) + n\epsilon_n.$$

Letting n and then P go to infinity, we will have:

$$\sum_{i \in \mathcal{S}} d_i \leq m' \Rightarrow |\mathcal{S}| d_{sym} \leq c(\mathbf{A}) \Rightarrow d_{sym} \leq \frac{c(\mathbf{A})}{|\mathcal{S}|} \Rightarrow d_{sym} \leq \min_{\mathcal{S} \subseteq [1:K]} \min_{\mathbf{A} \in \mathcal{J}(\mathcal{S})} \frac{c(\mathbf{A})}{|\mathcal{S}|}.$$

□

A simple corollary of Theorem 4.1 is the following.

Corollary 4.1. *Consider a K-user interference network with adjacency matrix $\mathbf{M}$. If $\mathcal{A} \subseteq \mathcal{S} \subseteq [1 : K]$ and there exists a sequence $\Pi_{\mathcal{S} \backslash \mathcal{A}} = (i_1, ..., i_{|\mathcal{S} \backslash \mathcal{A}|})$ of the users in $\mathcal{S} \backslash \mathcal{A}$ such that:*

$$\mathbf{M}_{i_j}^{\mathcal{S}} \in_{i_j}^{\pm} \text{span}\,(\mathbf{M}_{\mathcal{A}}^{\mathcal{S}}, \mathbf{I}_{\mathcal{A} \cup \{i_1, ..., i_{j-1}\}}^{|\mathcal{S}|}),\ \forall j \in [1 : |\mathcal{S} \backslash \mathcal{A}|],$$

then $d_{sym} \leq \frac{|\mathcal{A}|}{|\mathcal{S}|}$.

Proof. If $\mathcal{A}$ satisfies the conditions in the corollary, then it is easy to show that $\mathbf{M}_{\mathcal{A}}^{\mathcal{S}}$ is a generator of $\mathcal{S}$ and hence Theorem 4.1 yields $d_{sym} \leq \frac{c(\mathbf{M}_{\mathcal{A}}^{\mathcal{S}})}{|\mathcal{S}|} = \frac{|\mathcal{A}|}{|\mathcal{S}|}$. □

Remark. Theorem 4.1 and Corollary 4.1 both depend completely on the set of interferers to the receivers or, equivalently, the adjacency matrix. Therefore, they both highlight the special role of the topology of the network on the outer bounds. Moreover, Corollary 4.1 implies that it may be sufficient to only consider as the generators the matrices which are a subset of the columns of the adjacency matrix; i.e. only considering a subset of the received signals as our initial signals. This in fact worked for the case of Example 3 where we could derive the outer bound of $\frac{2}{5}$ for the symmetric degrees-of-freedom.

It is important to notice that in the final step of the proof of Theorem 4.1, we used the trivial upper bound of $c(\mathbf{A})n\log(P)$ for the joint entropy of the signals corresponding to the generator $\mathbf{A}$. However, there may be a way to derive a tighter upper bound for this joint entropy in some network topologies, and as we see in the next section, this is in fact the case; i.e. there exist some network topologies in which the upper bound of Theorem 4.1 can be improved. Hence, we will illustrate a method to tighten the upper bound in the following section.

4.2.2 Upper Bounds Based on the Concept of Fractional Generators

We will now introduce the notion of *fractional generators* to enhance the outer bound of Theorem 4.1. The idea is that we can make use of the signal interactions and interference topology at the receivers to derive possibly tighter upper bounds for the entropy of the signals corresponding to a generator. To be precise, if a signal is composed of a subset of interferers to a receiver, there is a tighter upper bound than $n\log(P)$ for that signal. To clarify this concept, we will again go through an introductory example.

Example 4. Consider the 6-user network shown in Figure 4.2. We claim that the symmetric DoF for this network is upper bounded by $\frac{2}{7}$, while the best upper bound based on Theorem 4.1 is $\frac{2}{6}$.

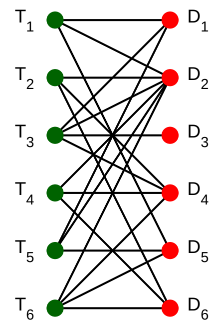


Figure 4.2: A 6-user interference network in which the upper bound of Theorem 4.1 is not tight.

The best upper bound of Theorem 4.1 for this example can be shown to be $\frac{2}{6}$, which is obtained by, for example, using $\mathbf{A} = \mathbf{M}^{\mathcal{S}}_{\{1,4\}}$ as a generator of the entire network $\mathcal{S} = [1:6]$ with $\Pi_{\mathcal{S}} = \{1,4,2,5,3,6\}$. We now show how the proof steps of Theorem 4.1 can be enhanced to obtain a tighter upper bound on d_{sym}.

Following the proof of Theorem 4.1 for the network in Figure 4.2 until equation (4.9) provides

$$n\sum_{i=1}^{6} R_i \leq h(\tilde{Y}_1^n, \tilde{Y}_4^n|\mathcal{G}^n) + no(\log(P)) + n\epsilon_n, \tag{4.10}$$

where $\tilde{Y}_1^n = g_1^n X_1^n + g_3^n X_3^n + g_5^n X_5^n + \tilde{Z}_1^n$ and $\tilde{Y}_4^n = g_2^n X_2^n + g_3^n X_3^n + g_4^n X_4^n + g_6^n X_6^n + \tilde{Z}_4^n$, $\tilde{Z}_1^n$ and $\tilde{Z}_4^n$ have the same distributions as the original noise vectors, but are independent of them and also of each other and $g_i^n = g_{ii}^n$, $i \in [1:6]$. Now, instead of simply upper bounding $h(\tilde{Y}_1^n, \tilde{Y}_4^n|\mathcal{G}^n)$ as $h(\tilde{Y}_1^n, \tilde{Y}_4^n|\mathcal{G}^n) \leq h(\tilde{Y}_1^n|\mathcal{G}^n) + h(\tilde{Y}_4^n|\mathcal{G}^n) \leq 2n\log(P) + no(\log(P))$, we show that a tighter upper bound can be found for $h(\tilde{Y}_1^n|\mathcal{G}^n)$, hence improving the upper bound on d_{sym}.

The idea is that in the network of Figure 4.2, D_1 receives signals from transmitters 1, 3 and 5. However, these transmitters are a *subset of the interferers to receiver 2*; i.e. $\{1, 3, 5\} \subseteq \mathcal{IF}_2$. This leads to a tighter upper bound of $h(\tilde{Y}_1^n|\mathcal{G}^n) \leq n(\log(P) - R_2) + no(\log(P)) + n\epsilon_n$, which can be proved as follows.

First, note that

$$H(W_2) - H(W_2|g_{22}^n X_2^n + \tilde{Y}_1^n, \mathcal{G}^n) = h(g_{22}^n X_2^n + \tilde{Y}_1^n|\mathcal{G}^n) - h(g_{22}^n X_2^n + \tilde{Y}_1^n|W_2, \mathcal{G}^n),$$

since both sides are equal to $I(g_{22}^n X_2^n + \tilde{Y}_1^n; W_2|\mathcal{G}^n)$. Therefore,

$$\begin{aligned} h(g_{22}^n X_2^n + \tilde{Y}_1^n|W_2, \mathcal{G}^n) &= H(W_2|g_{22}^n X_2^n + \tilde{Y}_1^n, \mathcal{G}^n) + h(g_{22}^n X_2^n + \tilde{Y}_1^n|\mathcal{G}^n) - H(W_2) \\ &\leq n\epsilon_n + h(g_{22}^n X_2^n + \tilde{Y}_1^n|\mathcal{G}^n) - H(W_2) \qquad (4.11) \\ &\leq n\epsilon_n + no(\log(P)) + n\log(P) - nR_2, \qquad (4.12) \end{aligned}$$

where (4.11) holds because of the same arguments as in the proof of Theorem 4.1 (less-interfered version of a received signal is sufficient to decode its corresponding symbol). On the other hand, since X_2^n is a function of W_2, we have

$$h(g_{22}^n X_2^n + \tilde{Y}_1^n|W_2, \mathcal{G}^n) = h(\tilde{Y}_1^n|\mathcal{G}^n),$$

which together with (4.12) yields $h(\tilde{Y}_1^n|\mathcal{G}^n) \leq n(\log(P) - R_2) + no(\log(P)) + n\epsilon_n$. Hence we can continue (4.10) as

$$\begin{aligned} n\sum_{i=1}^{6} R_i &\leq h(\tilde{Y}_1^n|\mathcal{G}^n) + h(\tilde{Y}_4^n|\mathcal{G}^n) + no(\log(P)) + n\epsilon_n \\ &\leq 2n\log(P) - nR_2 + no(\log(P)) + n\epsilon_n. \end{aligned}$$

Letting n and then P go to infinity and setting all the DoFs to be equal to d_{sym}, we will have:

$$6d_{sym} \leq 2 - d_{sym} \Rightarrow d_{sym} \leq \frac{2}{7},$$

which is strictly tighter than the previous outer bound of $\frac{2}{6}$ based on Theorem 4.1.

Now, we will illustrate the improvement of the outer bound in a linear algebraic form. The key part in the enhancement was that by adding $g_{22}^n X_2^n$ to $\tilde{Y}_1^n$, we could create a signal which was able to decode W_2. As we have discussed before, if $\mathcal{S} = [1:6]$, then $\tilde{Y}_1^n$ corresponds to the vector $\mathbf{M}_1^{\mathcal{S}} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}^T$. Therefore, adding $g_{22}^n X_2^n$ to $\tilde{Y}_1^n$ can be translated to adding $\mathbf{I}_2^{|\mathcal{S}|} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T$ to $\mathbf{M}_1^{\mathcal{S}}$. Moreover, the fact that W_2 can be decoded from $g_{22}^n X_2^n + \tilde{Y}_1^n$ is equivalent to $\mathbf{M}_2^{\mathcal{S}} \in_2^{\pm} \text{span}(\mathbf{M}_1^{\mathcal{S}} + \mathbf{I}_2^{|\mathcal{S}|})$. We will call $\mathbf{M}_1^{\mathcal{S}}$ a *fractional generator* of $\mathcal{S}'$ in $\mathcal{S}$ where $\mathcal{S}' = \{2\}$. This means that by expanding the signal corresponding to $\mathbf{M}_1^{\mathcal{S}}$ (through adding $\mathbf{I}_2^{|\mathcal{S}|}$ to $\mathbf{M}_1^{\mathcal{S}}$ or equivalently $g_{22}^n X_2^n$ to $\tilde{Y}_1^n$), the resulting expanded signal is able to decode W_2. This is the method that we will use to linear algebraically describe the improvement in the outer bound on d_{sym}. $\triangle$

Remark. A similar approach in [31] has been taken to derive an upper bound for the symmetric DoF of general network topologies. In particular, if for the network in Example 4, we set $h_{26} = h_{46} = h_{66} = -\sqrt{\text{SNR} \times \frac{N_0}{P}}$ and the other channel gains h_{ji} to $\sqrt{\text{SNR} \times \frac{N_0}{P}}$, then *maximum cardinality of an acyclic subset of messages*, denoted by Ψ, is equal to 3 and the *minimum internal conflict distance*, denoted by Δ, is equal to 1 for the network of Figure 4.2. Therefore, both of the bounds presented in Theorem 4.12 and Corollary 4.13 of [31] for the network of Figure 4.2 are equal to $\frac{1}{3}$, while the outer bound of $\frac{2}{7}$ that we derived in Example 4 is strictly tighter.

To generalize the improvement of the outer bound to all network topologies, we define the concept of *fractional generator*.

Definition 4.3. Consider a K-user interference network with adjacency matrix $\mathbf{M}$ and suppose $\mathcal{S}' \subseteq \mathcal{S} \subseteq [1:K]$. A vector $\mathbf{c} \in \{\pm 1, 0\}^{|\mathcal{S}|}$ is called a *fractional generator of $\mathcal{S}'$ in $\mathcal{S}$* if $\mathbf{c}_k = 0, \forall k \in \mathcal{S}'$ and there exists a sequence $\Pi_{\mathcal{S}'} = (i_1, ..., i_{|\mathcal{S}'|})$

of the users in $\mathcal{S}'$ such that:

$$\mathbf{M}^{\mathcal{S}}_{i_j} \in^{\pm}_{i_j} \text{span}\,(\mathbf{c} + \sum_{k\in\mathcal{S}'} \mathbf{I}^{|\mathcal{S}|}_k, \mathbf{I}^{|\mathcal{S}|}_{\{i_1,\ldots,i_{j-1}\}}),\ \forall j \in [1 : |\mathcal{S}'|].$$

We use the notation $\mathcal{J}_{\mathcal{S}}(\mathcal{S}')$ to denote the set of all fractional generators of $\mathcal{S}'$ in $\mathcal{S}$.

Intuitively, a fractional generator of $\mathcal{S}'$ in $\mathcal{S}$ is a column vector whose corresponding signal can decode the messages of the users in $\mathcal{S}'$ (which is a subset of the set of entire users $\mathcal{S}$) sequentially, after expansion by adding $\sum_{k\in\mathcal{S}'} \mathbf{I}^{|\mathcal{S}|}_k$ to it (or equivalently, by adding $\sum_{k\in\mathcal{S}'} g^n_{kk} X^n_k$ to its corresponding signal).

After having the definition of fractional generators, we can state the following lemma, which is proved in Appendix D.

Lemma 4.2. *Consider a subset of users $\mathcal{S} \subseteq [1 : K]$ in a K-user interference network, and suppose $\mathbf{c} \in \mathcal{J}_{\mathcal{S}}(\mathcal{S}')$. If rates R_i are achievable for all $i \in \mathcal{S}'$, then*

$$h(\sum_{j\in\mathcal{S}} \mathbf{c}_j g^n_j X^n_j + Z^n|\mathcal{G}^n) \leq n(\log(P) - \sum_{i\in\mathcal{S}'} R_i) + no(\log(P)) + n\epsilon_n,$$

where $g^n_j = g^n_{jj}$, $\forall j \in \mathcal{S}$, $\epsilon_n \to 0$ as $n \to \infty$, and each element in Z^n is a $\mathcal{CN}(0, 1)$ random variable.

We will also define $n_{\mathcal{S}}(\mathbf{c})$ as follows for a vector $\mathbf{c} \in \{\pm 1, 0\}^{|\mathcal{S}|}$.

Definition 4.4. Consider a subset of users $\mathcal{S} \subseteq [1 : K]$ in a K-user interference network. For a vector $\mathbf{c} \in \{\pm 1, 0\}^{|\mathcal{S}|}$, $n_{\mathcal{S}}(\mathbf{c})$ is defined as

$$\begin{aligned} n_{\mathcal{S}}(\mathbf{c}) := \max_{\mathcal{S}'} &\ |\mathcal{S}'| \\ \text{s.t. } &\ \mathbf{c} \in \mathcal{J}_{\mathcal{S}}(\mathcal{S}'). \end{aligned}$$

Therefore, $n_{\mathcal{S}}(\mathbf{c})$ is equal to the size of the largest subset of the users $\mathcal{S}' \subseteq \mathcal{S}$ such that $\mathbf{c}$ is a fractional generator of $\mathcal{S}'$ in $\mathcal{S}$. Note that, due to Lemma 4.2, finding $n_{\mathcal{S}}(\mathbf{c})$ leads to the tightest upper bound for the signal corresponding to $\mathbf{c}$. Therefore, we are now at a stage to state our second converse.

Theorem 4.2. *The symmetric DoF of a K-user interference network with no CSIT is upper bounded by*

$$d_{sym} \leq \min_{\mathcal{S}\subseteq[1:K]} \min_{\mathbf{A}\in\mathcal{J}(\mathcal{S})} \frac{c(\mathbf{A})}{|\mathcal{S}| + \sum_{i=1}^{c(\mathbf{A})} n_{\mathcal{S}}(\mathbf{A}_i)},$$

where for each $\mathcal{S} \subseteq [1:K]$, $\mathcal{J}(\mathcal{S})$ denotes the set of all generators of $\mathcal{S}$ (Definition 4.2), $c(\mathbf{A})$ denotes the number of columns of $\mathbf{A}$ and $n_{\mathcal{S}}(\mathbf{A}_i)$ is defined as in Definition 4.4.

Proof. Following the proof of Theorem 4.1 until equation (4.9), we know that if $\mathcal{S} = \{1, ..., m\}$, $c(\mathbf{A}) = m'$, $\Pi_{\mathcal{S}} = (1, ..., m)$, $\tilde{Y}_i^n = \begin{bmatrix} g_1^n X_1^n & \dots & g_m^n X_m^n \end{bmatrix} \mathbf{A}_i + \tilde{Z}_i^n$, $i \in [1 : m']$ and if rates R_i ($i \in \mathcal{S}$) are achievable, we will have

$$\begin{aligned} n\sum_{i\in\mathcal{S}} R_i &\leq h(\tilde{Y}_1^n, ..., \tilde{Y}_{m'}^n|\mathcal{G}^n) + no(\log(P)) + n\epsilon_n \\ &\leq \sum_{i=1}^{m'} h(\tilde{Y}_i^n|\mathcal{G}^n) + no(\log(P)) + n\epsilon_n. \end{aligned} \tag{4.13}$$

Now, if $\mathbf{A}_i \in \mathcal{J}_{\mathcal{S}}(\mathcal{S}')$, then Lemma 4.2 implies

$$\begin{aligned} n(\log(P) - \sum_{j\in\mathcal{S}'} R_j) + no(\log(P)) + n\epsilon_n &\geq h\left(\sum_{j\in\mathcal{S}} \mathbf{A}_{ji} g_j^n X_j^n + \tilde{Z}_i^n|\mathcal{G}^n\right) \\ &= h\left(\begin{bmatrix} g_1^n X_1^n & \dots & g_m^n X_m^n \end{bmatrix} \mathbf{A}_i + \tilde{Z}_i^n|\mathcal{G}^n\right) \\ &= h(\tilde{Y}_i^n|\mathcal{G}^n). \end{aligned} \tag{4.14}$$

Thus, to find the tightest upper bound on $h(\tilde{Y}_i^n|\mathcal{G}^n)$ for every $i \in [1 : c(\mathbf{A})]$, we need to find the largest subset $\mathcal{S}'$ such that $\mathbf{A}_i \in \mathcal{J}_{\mathcal{S}}(\mathcal{S}')$, which we denote by

${\mathcal{S}'}_i^*$; i.e. ${\mathcal{S}'}_i^* = \arg\max_{\mathcal{S}'} |\mathcal{S}'|$ s.t. $\mathbf{A}_i \in \mathcal{J}_{\mathcal{S}}(\mathcal{S}')$. Combining this with (4.13) and (4.14) yields

$$n \sum_{i \in \mathcal{S}} R_i \leq \sum_{i=1}^{c(\mathbf{A})} n(\log(P) - \sum_{j \in {\mathcal{S}'}_i^*} R_j) + no(\log(P)) + n\epsilon_n.$$

Letting n and then P go to infinity and setting all the DoFs to be equal to d_{sym}, we will have:

$$\begin{aligned}
|\mathcal{S}| d_{sym} &\leq c(\mathbf{A}) - \sum_{i=1}^{c(\mathbf{A})} |{\mathcal{S}'}_i^*| d_{sym} = c(\mathbf{A}) - \sum_{i=1}^{c(\mathbf{A})} n_{\mathcal{S}}(\mathbf{A}_i) d_{sym} \\
\Rightarrow d_{sym} &\leq \frac{c(\mathbf{A})}{|\mathcal{S}| + \sum_{i=1}^{c(\mathbf{A})} n_{\mathcal{S}}(\mathbf{A}_i)} \\
\Rightarrow d_{sym} &\leq \min_{\mathcal{S} \subseteq [1:K]} \min_{\mathbf{A} \in \mathcal{J}(\mathcal{S})} \frac{c(\mathbf{A})}{|\mathcal{S}| + \sum_{i=1}^{c(\mathbf{A})} n_{\mathcal{S}}(\mathbf{A}_i)}.
\end{aligned}$$

□

As it is clear from the above discussion, the outer bound of Theorem 4.2 captures the impact of network topology on upper bounding the symmetric DoF more strongly than Theorem 4.1. In fact, Theorem 4.2 tries to focus on the signal and interference interactions at the receivers through Lemma 4.2, which is the key aspect of the improvement of the bound compared to the bound suggested by Theorem 4.1.

4.3 Inner Bounds on d_{sym}

In this section, we derive inner bounds on the symmetric degrees-of-freedom. In particular, we focus on two benchmark schemes, namely *random Gaussian coding* and *interference avoidance*, and introduce a new scheme called *structured repetition coding*. The structured repetition coding scheme in general performs better than (or at least the same as) the first two schemes and as we illustrate in Section 4.4, it closes the gap between the inner and outer bounds in many networks where the first two schemes fail to do so.

4.3.1 Benchmark Schemes

We start by presenting two benchmark schemes and we will compare them with each other through examples to study their performance with respect to our outer bounds in Section 4.2.

Random Gaussian Coding and Interference Decoding

In the first scheme, we use random Gaussian coding, such that all interfering messages at each receiver are decoded. Consider a K-user interference network and look at one of the receivers, say D_j. It receives signals from T_i, $i \in \{j\} \cup \mathcal{IF}_j$. Therefore, we can see this subnetwork as a multiple access channel (MAC) to receiver j. It is well known [32] that in the fast fading settings, the capacity region of MAC with no CSIT is specified by

$$\sum_{i \in \mathcal{S}} R_i \leq \mathbb{E}\left[\log\left(1 + \sum_{i \in \mathcal{S}} |g_{ij}|^2 P\right)\right], \ \forall \mathcal{S} \subseteq \{j\} \cup \mathcal{IF}_j,$$

where the expectation is taken with respect to the channel gains. Now, since $\log(x) < \log(1+x)$ for all positive x, the rates R_i are achievable if they satisfy

$$\sum_{i \in \mathcal{S}} R_i \leq \mathbb{E}\left[\log\left(\sum_{i \in \mathcal{S}} |g_{ij}|^2 P\right)\right] = \log(P) + \mathbb{E}\left[\log\left(\sum_{i \in \mathcal{S}} |g_{ij}|^2\right)\right], \ \forall \mathcal{S} \subseteq \{j\} \cup \mathcal{IF}_j.$$

Then, because $\log(.)$ is a monotonically increasing function, the rates R_i are achievable if the following holds.

$$\sum_{i \in \mathcal{S}} R_i \leq \log(P) + \mathbb{E}\left[\log\left(|g_{i_{\mathcal{S}}j}|^2\right)\right], \ \forall \mathcal{S} \subseteq \{j\} \cup \mathcal{IF}_j,$$

where for each $\mathcal{S} \subseteq \{j\} \cup \mathcal{IF}_j$, $i_{\mathcal{S}}$ is some user in $\mathcal{S}$. From the regularity conditions on the distribution of the channel gains (mentioned in Section 4.1), it can be shown that $\mathbb{E}\left[\log\left(|g|^2\right)\right] > -\infty$ (a more general case is proved in Appendix F). Therefore, dividing the above equations by $\log(P)$ and letting P go to infinity leads to the fact that the degrees-of-freedom d_j are achievable if

$$\sum_{i \in \mathcal{S}} d_i \leq 1, \ \forall \mathcal{S} \subseteq \{j\} \cup \mathcal{IF}_j.$$

For the degrees-of-freedom to be symmetric, we will therefore have $d_{sym} \leq \frac{1}{|S|}$ which should hold for every $\mathcal{S} \subseteq \{j\} \cup \mathcal{IF}_j$. Choosing the largest subset $\mathcal{S}$ yields $d_{sym} \leq \frac{1}{1+|\mathcal{IF}_j|}$. Furthermore, all the rates (degrees-of-freedom) in this region can be achieved using random Gaussian codebooks of size $2^{nR_i} \times n$ generated for each user, in which all the elements are i.i.d. $\mathcal{CN}(0, P)$. The message W_i is the index of the row of this codebook matrix and the transmit vector will be the corresponding row of the codebook. Therefore, by applying the viewpoint of multiple access channels to all the receivers in the interference network, this theorem follows immediately.

Theorem 4.3. *Consider a K-user interference network with adjacency matrix $\mathbf{M}$. If we denote the maximum receiver degree by Δ_R (defined as $\Delta_R := 1 + \max_{j \in [1:K]} |\mathcal{IF}_j| = \max_{j \in [1:K]} \sum_{i=1}^{K} \mathbf{M}_{ij}$), then the symmetric DoF of $\frac{1}{\Delta_R}$ is achievable.*

Theorem 4.3 only considers the maximum degree among the receivers to derive an inner bound on d_{sym}. However, it fails to capture how further details of network topology can affect the achievable symmetric DoF. In other words, this theorem suggests a similar inner bound for all network topologies whose maximum receiver degrees are identical, implying its possible suboptimality for many networks. Therefore, we should seek for other schemes that exploit other structures in the network topology.

Interference Avoidance

As the name suggests, this scheme is based on avoiding the interference by all the users. Each transmitter, aware of the network topology, knows the receivers which receive interference from itself and also the transmitters who cause interference at its corresponding receiver. Therefore, it can avoid sending its symbols at the same time as those users. In other words, in this scheme, each user uses a time slot to transmit data if and only if the users who receive interference from/cause interference at that user do not use that time slot. This is tightly connected to the concept of *independent sets*.

Suppose we have a K-user interference network. $\mathcal{U} \subseteq [1:K]$ is an independent set if for all two distinct users i and j in $\mathcal{U}$, $\mathbf{M}_{ij} = \mathbf{M}_{ji} = 0$; i.e. users i and j are mutually non-interfering. Obviously, all the users in an independent set can transmit their symbols at the same time without experiencing any interference. This is the essence of the *interference avoidance* scheme. Naturally, it is best if the largest possible subset of the users send together, leading to the concept of *maximal* independent sets. $\mathcal{U}$ is a maximal independent set if it is an independent set, but for all $l \in [1:K] \backslash \mathcal{U}$, $\mathcal{U} \cup \{l\}$ is not an independent set.

After describing the above scheme, we can state our second inner bound.

Theorem 4.4. *Consider a K-user interference network with adjacency matrix $\mathbf{M}$ and* *suppose $\mathcal{U} = \{\mathcal{U}_1, ..., \mathcal{U}_m\}$ is the set of all maximal independent sets of this network. Then, the following symmetric DoF is achievable by interference avoidance.*

$$\sup_{n \in \mathbb{N}} \max_{\mathcal{U}'_1,...,\mathcal{U}'_n \in \mathcal{U}} \min_{i \in [1:K]} \frac{\sum_{j=1}^{n} \mathbf{1}(i \in \mathcal{U}'_j)}{n},$$

where for an event A, $\mathbf{1}(A) = 1$ if A occurs and $\mathbf{1}(A) = 0$ otherwise.

Proof. If we take n maximal independent sets $\mathcal{U}'_1, ..., \mathcal{U}'_n$ and allow all the users in $\mathcal{U}'_j$ to transmit simultaneously in time slot j, $j \in [1:n]$, then for every user i, $i \in [1:K]$, there will be $\sum_{j=1}^{n} \mathbf{1}(i \in \mathcal{U}'_j)$ clean, interference-free, channels between T_i and D_i. Hence, each user i achieves $\frac{\sum_{j=1}^{n} \mathbf{1}(i \in \mathcal{U}'_j)}{n}$ degrees-of-freedom. Since we are interested in the achievable symmetric degrees-of-freedom, the maximum DoF that all the users can simultaneously achieve with a specific choice of n and $\mathcal{U}'_1, ..., \mathcal{U}'_n$ is $\min_{i \in [1:K]} \frac{\sum_{j=1}^{n} \mathbf{1}(i \in \mathcal{U}'_j)}{n}$. Optimizing over n and $\mathcal{U}'_1, ..., \mathcal{U}'_n$, the best symmetric DoF achievable under interference avoidance is $\sup_{n \in \mathbb{N}} \max_{\mathcal{U}'_1,...,\mathcal{U}'_n \in \mathcal{U}} \min_{i \in [1:K]} \frac{\sum_{j=1}^{n} \mathbf{1}(i \in \mathcal{U}'_j)}{n}$. □

Remark. The aforementioned ideas of independent sets are very closely related to fractional coloring and fractional chromatic numbers of graphs in graph theory [33]. To relate the two problems, we define the *conflict graph* of a K-user interference network with adjacency matrix $\mathbf{M}$ as an undirected graph $G = (\mathcal{V}, \mathcal{E})$ with the set of vertices $\mathcal{V} = [1:K]$ and the set of edges $\mathcal{E}$ where for all $i \neq j$, $e_{ij} \in \mathcal{E}$ if $\mathbf{M}_{ij} = 1$ or $\mathbf{M}_{ji} = 1$ in the original interference network. Now, the assignment of time slots to different users based on independent sets corresponds to *coloring* the conflict graph G. An *n-coloring* of a graph $G = (\mathcal{V}, \mathcal{E})$ is an assignment of a single color out of a set of n colors to each of the vertices in $\mathcal{V}$ such that if $e_{ij} \in \mathcal{E}$, different colors are assigned to vertices i and j. The smallest n for which an n-coloring is possible for G is called the *chromatic number* of G, denoted by $\chi(G)$.

Moreover, an *m-fold coloring* (known as *fractional coloring*) of a graph G is an assignment of sets of m colors to each vertex in $\mathcal{V}$ such that if $e_{ij} \in \mathcal{E}$, the sets of colors assigned to vertices i and j are disjoint. Also, G is said to be $n : m$-*colorable* if there exists an m-fold coloring of G such that all the colors used in the coloring are drawn from a set of n distinct colors. The smallest n for which G is $n : m$-colorable is called the m-fold chromatic number of G, denoted by $\chi_m(G)$. The maximum symmetric DoF achievable by interference avoidance is $\sup_{m\in\mathbb{N}} \frac{m}{\chi_m(G)}$ which is exactly the value presented in Theorem 4.4. [3] However, the *fractional chromatic number* of G is defined as $\chi_f(G) = \inf_{m\in\mathbb{N}} \frac{\chi_m(G)}{m}$, which can also be shown to equal $\lim_{m\to\infty} \frac{\chi_m(G)}{m}$ [33]. Therefore, the best symmetric DoF achievable by interference avoidance is in fact $\frac{1}{\chi_f(G)}$.

The two schemes we presented so far, incorporate two different aspects of network topology, namely maximum receiver degree and fractional chromatic number, to improve spectral efficiency. A natural question that comes to mind is: How do these two schemes compare to each other? Is one of them superior than the other one for all network graphs? The answer is negative. We will present two examples to clarify how the schemes work and also to compare them. In the first example, random Gaussian coding performs better, while in

[3]For every $m \in \mathbb{N}$, interference avoidance can achieve the symmetric DoF of $\frac{m}{\chi_m(G)}$, because for every m-fold chromatic number $\chi_m(G)$, m is the largest $\bar{m}$ such that an $\bar{m}$-fold coloring exists for G, where the colors are selected out of a palette of $\chi_m(G)$ colors. Each color out of the total of $\chi_m(G)$ colors corresponds to an independent set. Hence, m is the maximum $\bar{m}$ such that each node appears $\bar{m}$ times in the independent sets corresponding to $\chi_m(G)$ colors. In other words, if $\mathcal{U}$ is the set of all maximal independent sets of the interference network, then

$$m = \max_{\mathcal{U}'_1,\ldots,\mathcal{U}'_{\chi_m(G)} \in \mathcal{U}} \min_{i\in[1:K]} \sum_{j=1}^{\chi_m(G)} \mathbf{1}(i \in \mathcal{U}'_j),$$

because each user appears at least $\min_{i\in[1:K]} \sum_{j=1}^{\chi_m(G)} \mathbf{1}(i \in \mathcal{U}'_j)$ times among the independent sets $\mathcal{U}'_1, \ldots, \mathcal{U}'_{\chi_m(G)}$ and m is the maximum value of this quantity where the maximization is over the selection of independent sets corresponding to $\chi_m(G)$ colors. Optimizing over m yields the inner bound in Theorem 4.4.

the second one, interference avoidance outperforms the first scheme.

Example 5. Consider the 4-user network in Figure 4.3(a). Suppose we want to

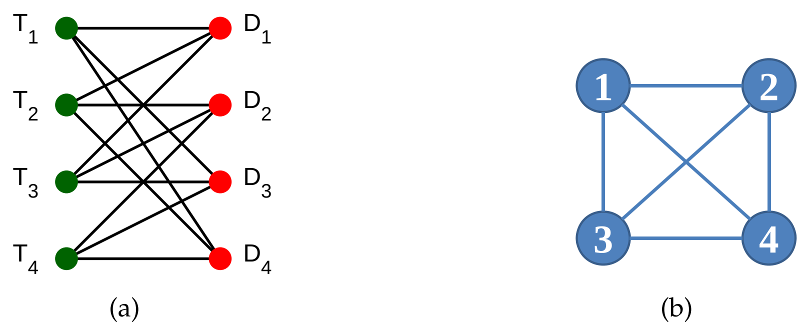


Figure 4.3: (a) A 4-user interference network in which random Gaussian coding is optimal, and (b) its corresponding conflict graph.

apply interference avoidance to this network. We should identify the independent sets in this network. Clearly, all the users are mutually interfering in this network. This can also be seen in the fully connected conflict graph of Figure 4.3(b), whose maximal independent sets are $\{1\}, \{2\}, \{3\}$ and $\{4\}$, implying that the best symmetric DoF achievable under interference avoidance is $\frac{1}{4}$. However, the maximum receiver degree in this network is $\Delta_R = 3$ and therefore, Theorem 4.3 implies that random Gaussian coding and interference decoding can achieve the symmetric DoF of $\frac{1}{3}$ which is higher than the value achieved by interference avoidance.

To show that the symmetric DoF of $\frac{1}{3}$ is optimal, it is necessary to mention the outer bound, too. If you consider the subnetwork consisting of the users $\mathcal{S} = \{1, 2, 3\}$, then clearly $[1\ 1\ 1]^T$ is a generator of $\mathcal{S}$. Therefore, using Theorem 4.1, $d_{sym} \leq \frac{1}{3}$ implying the optimality of random Gaussian coding and interference decoding in this network, whereas interference avoidance performs suboptimally in this case. $\triangle$

Example 6. As our next example, we return to the network we considered in Example 3, which is repeated in Figure 4.4(a) for convenience.

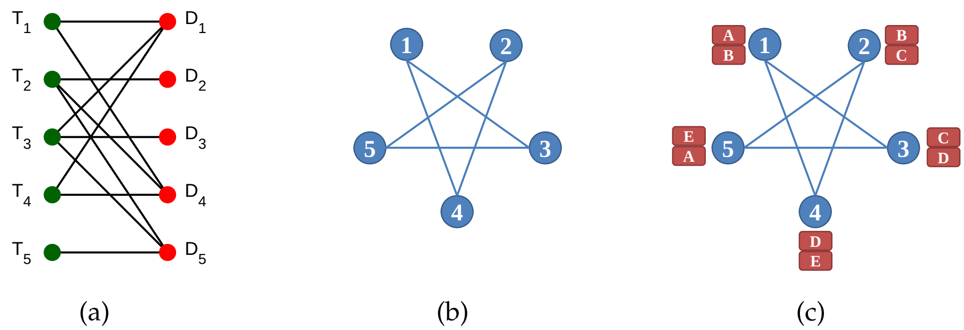


Figure 4.4: (a) A 5-user interference network in which interference avoidance is optimal, (b) the corresponding conflict graph, and (c) a 5:2-coloring.

As shown before, for this network $d_{sym} \leq \frac{2}{5}$. However, the maximum receiver degree in this network is $\Delta_R = 3$, hence random Gaussian coding and interference decoding can only achieve the symmetric DoF of $\frac{1}{3}$ which is less than the outer bound.

On the other hand, it is obvious that the maximal independent sets of the conflict graph of this network, shown in Figure 4.4(b), are $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$, $\{4, 5\}$ and $\{5, 1\}$. By assigning one time slot to each of these sets, we can achieve the symmetric DoF of $\frac{2}{5}$ because each user is repeated twice in these sets, therefore meeting the outer bound of $\frac{2}{5}$ mentioned earlier. A corresponding 5:2-coloring is also shown in Figure 4.4(c). Hence, in this example, interference avoidance outperforms random Gaussian coding and interference decoding. $\triangle$

Taking a closer look at the two schemes presented in this section, they can be viewed as two extremes of a spectrum. Random Gaussian coding and interference decoding tries to decode all the interference at all the receivers by adopting a random code which does not make efficient use of the topology of the network. On the other side, interference avoidance tries to prevent the mutually interfering nodes from transmitting at the same time, which causes no interference to occur at the receivers. Therefore, one may think of using a scheme that is naturally between these two extremes; i.e. using some kind of structured code that makes best use of the topology of the network and does not necessarily try to avoid the interference at the receivers, but at the same time enables the receivers to decode their desired messages. This leads to a new scheme which will be introduced in the following section.

4.3.2 Structured Repetition Coding

We now present a scheme based on structured repetition codes at the transmitters so that we can better exploit structure of network topology. This scheme unifies the two schemes presented in Section 4.3.1 in the way that it not only enables the receivers to decode their intended symbols without necessarily decoding all the interference, but it also allows mutually interfering users to possibly send data at the same time, implying that the scheme can potentially outperform both benchmark schemes presented in Section 4.3.1. We will motivate the idea of structured repetition coding through the following example. Before starting the example, we need the following definition.

Definition 4.5. For a graph $G = (\mathcal{V}, \mathcal{E})$, a *matching* is a subset of edges no two of which share a common vertex. The *matching number* of G, denoted by $\mu(G)$, is the size of a maximum matching of G (a matching of G containing the largest possible number of edges).

Example 7. Consider the 6-user network in Figure 4.5. We claim that in this network, the symmetric DoF of $\frac{1}{3}$ is achievable, while the benchmark schemes discussed in the previous section can at most achieve a symmetric DoF of $\frac{1}{4}$.

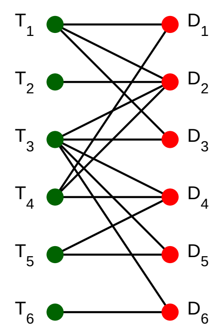


Figure 4.5: A 6-user interference network in which random Gaussian coding and interference avoidance are suboptimal.

In this network, the outer bound for the symmetric DoF is $d_{sym} \leq \frac{1}{3}$ by Corollary 4.1, because the sets $\mathcal{A} = \{2\}$ and $\mathcal{S} = \{1, 2, 3\}$ satisfy the conditions of the corollary; i.e. in the subnetwork $\mathcal{S}$, we can generate statistically similar versions of the signals at receivers 3 and 1 by having the received signal at receiver 2. Therefore, $d_{sym} \leq \frac{|\mathcal{A}|}{|\mathcal{S}|} = \frac{1}{3}$.

However, in terms of the achievable schemes, Theorem 4.3 indicates that the best symmetric DoF achievable by random Gaussian coding and interference decoding is $\frac{1}{\Delta_R} = \frac{1}{4}$. Also, the maximal independent sets of this network are $\{1, 5, 6\}$, $\{2, 5, 6\}$, $\{3\}$ and $\{4, 6\}$. Therefore, Theorem 4.4 states that the maximum symmetric DoF which interference avoidance can achieve is $\frac{1}{4}$. Thus, our two previous schemes both achieve the same symmetric DoF of $\frac{1}{4}$ which is strictly lower than the outer bound of $\frac{1}{3}$. Now, let us see if the achievable symmetric DoF can be improved.

Targeting the symmetric DoF of $\frac{1}{3}$, we can think of an achievable scheme in which each transmitter has one symbol to be sent within three time slots such that all the receivers can decode their desired messages. To this end, we create a *transmission matrix* $\mathbf{T} \in \{0,1\}^{6\times 3}$ where $\mathbf{T}_{ik} = 1$ if transmitter i sends its single symbol X_i in time slot k and $\mathbf{T}_{ik} = 0$ if transmitter i is silent in time slot k. Consider the following matrix.

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}^T . \tag{4.15}$$

As mentioned above, the first row of (4.15) means that transmitter 1 sends its only symbol X_1 in time slot 1 and remains silent otherwise, the second row means that transmitter 2 sends its symbol X_2 in time slot 2, the third row implies that transmitter 3 repeats its symbol X_3 in time slots 1 and 3, etc. We will now show that with this transmission matrix, all the receivers can create interference-free versions of their desired symbols for almost all values of channel gains. As an example, let us focus on receiver 4. The signals that D_4 receives in three time slots are as follows.

$$\begin{aligned} Y_4[1] &= g_{34}[1]X_3 + g_{44}[1]X_4 + g_{54}[1]X_5 + Z_4[1] \\ Y_4[2] &= g_{54}[2]X_5 + Z_4[2] \\ Y_4[3] &= g_{34}[3]X_3 + g_{44}[3]X_4 + Z_4[3]. \end{aligned}$$

Since D_4 is aware of the channel gains of all the links connected to it at all times, it can create the following signal.

$$Y_4'[1,2] := Y_4[1] - \frac{g_{54}[1]}{g_{54}[2]} Y_4[2] = g_{34}[1]X_3 + g_{44}[1]X_4 + Z_4'[1,2],$$

where $Z_4'[1,2]$ is a noise term with bounded variance. Now, it is clear that from $Y_4[3]$ and $Y_4'[1,2]$, D_4 can create an interference-free version of X_4 as follows.

$$\frac{g_{34}[1]Y_4[3] - g_{34}[3]Y_4'[1,2]}{g_{34}[1]g_{44}[3] - g_{34}[3]g_{44}[1]} = X_4 + \tilde{Z}_4,$$

where $\tilde{Z}_4$ has a bounded variance. The above combination of the signals is possible if $g_{34}[1]g_{44}[3] - g_{34}[3]g_{44}[1] \neq 0$ which holds for almost all values of channel gains, because the channel gains are i.i.d. and drawn from continuous distributions.

The fact that for almost all values of the channel gains, there exists a linear combination of the received signals at receiver 4 which is an interference-free version of X_4 can also be viewed in terms of the matching number of a bipartite graph. The idea is to first create an "effective" transmission matrix $\bar{\mathbf{T}}^4$ for receiver 4, which is defined as a 6×3 matrix, where $\bar{\mathbf{T}}^4_{ik} = \mathbf{M}_{i4}\mathbf{T}_{ik}, \ \forall i \in [1:6], k \in [1:3]$, as shown in (4.16). In words, $\bar{\mathbf{T}}^4$ is the same as $\mathbf{T}$ with the distinction that the rows corresponding to the transmitters which are not connected to D_4 are set to zero.

$$\bar{\mathbf{T}}^4 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}^T . \tag{4.16}$$

This matrix corresponds to a bipartite graph $\bar{G}^4$, shown in Figure 4.6(a), with the set of vertices $\{v_1, ..., v_6\} \cup \{v'_1, v'_2, v'_3\}$, where v_i is connected to v'_k if and only if $\bar{\mathbf{T}}^1_{ik} = 1, \ \forall i \in [1:6], k \in [1:3]$.

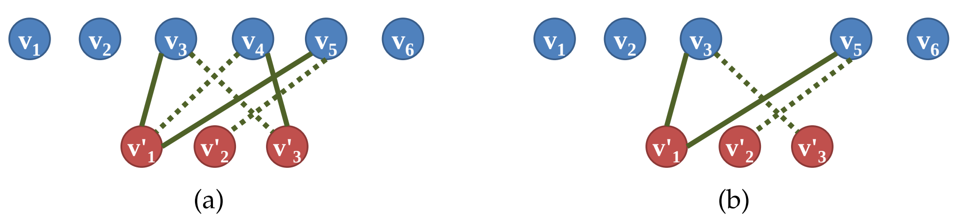


Figure 4.6: (a) The bipartite graph $\bar{G}^4$ corresponding to the matrix $\bar{\mathbf{T}}^4$ in (4.16), and (b) the graph $\bar{G}^4 \setminus 4$, which is the same as $\bar{G}^4$ after removing v_4 and its corresponding edges. In both graphs, the dashed edges correspond to a maximum matching.

Note that the matching number of $\bar{G}^4$, denoted by $\mu(\bar{G}^4)$, is equal to 3 and a maximum matching of $\bar{G}^4$ is shown in Figure 4.6(a). However, as shown in Figure 4.6(b), upon removal of v_4 and its corresponding edges from $\bar{G}^4$, the matching number reduces to 2. As we show in Lemma 4.3, this reduction in the matching number is equivalent to the fact that for almost all values of the channel gains, there exists a linear combination of the signals at receiver 4 which is an interference-free version of X_4. Theorem 4.5 shows that this procedure reduces the problem of checking whether the transmission matrix $\mathbf{T}$ is successful or not to a bipartite matching problem.

Therefore, user 4 can achieve $\frac{1}{3}$ degrees-of-freedom. Arguments similar to the one above can show that all the other receivers can create interference-free versions of their desired symbols either, by linearly combining their received signals in three time slots. In particular, D_1 needs to combine its received signals at time slots 1 and 3, whereas D_2, D_3, D_5 and D_6 only need their received signals at time slots 2, 3, 2 and 2, respectively. Therefore, this scheme, which we will call *structured repetition coding*, can achieve the symmetric DoF of $\frac{1}{3}$. This inner bound meets the outer bound, indicating that structured repetition coding is optimal in the network of Figure 4.5, contrary to the two benchmark schemes which perform suboptimally in this example. $\triangle$

Motivated by Example 7, we now formally define structured repetition coding. In what follows, for a general matrix $\mathbf{T}$, we use $\mathbf{T}_{l,*}$ to denote the l^{th} row of $\mathbf{T}$.

Definition 4.6. Consider a K-user interference network with adjacency matrix

$\mathbf{M}$. Also, consider a matrix $\mathbf{T} \in \{0,1\}^{mK \times n}$, for some $m, n \in \mathbb{N}$, satisfying

$$\sum_{l=(i-1)m+1}^{im} \mathbf{T}_{lk} \leq 1, \ \forall k \in [1:n], \ \forall i \in [1:K], \tag{4.17}$$

which, in words, means that there exists at most a single 1 in positions $(i-1)m+1$ to im of each column k, for all $i \in [1:K]$ and for all $k \in [1:n]$. Then, *structured repetition coding with transmission matrix* $\mathbf{T}$ is defined as a scheme, in which transmitter T_i ($i \in [1:K]$) intends to deliver m symbols, denoted by $\left\{\tilde{X}_l\right\}_{l=(i-1)m+1}^{im}$, to receiver D_i in n time slots, using the following encoding and decoding procedure.

- Transmitter T_i ($i \in [1:K]$) creates its transmit vector, denoted by X_i^n, as follows.

$$X_i^n = \sum_{l=(i-1)m+1}^{im} \mathbf{T}_{l,*}^T \tilde{X}_l.$$

  In words, this means that at each time slot k, transmitter i ($i \in [1:K]$) looks for index $l \in [(i-1)m+1 : im]$ such that $\mathbf{T}_{lk} = 1$ (note that due to (4.17) there is at most one such l) and transmits $\tilde{X}_l$ in that time slot (if such an index cannot be found, the transmitter will remain silent).

- At the end of the transmission, receiver D_j ($j \in [1:K]$) receives

$$\begin{aligned} Y_j^n &= \sum_{i=1}^{K} \mathbf{M}_{ij} g_{ij}^n X_i^n + Z_j^n \\ &= \sum_{i=1}^{K} \mathbf{M}_{ij} g_{ij}^n \left(\sum_{l=(i-1)m+1}^{im} \mathbf{T}_{l,*}^T \tilde{X}_l \right) + Z_j^n \\ &= \sum_{l=1}^{mK} \mathbf{M}_{\lceil \frac{l}{m} \rceil j} g_{\lceil \frac{l}{m} \rceil j}^n \mathbf{T}_{l,*}^T \tilde{X}_l + Z_j^n. \end{aligned}$$

  Then, D_j looks for vectors $\mathbf{u}_l \in \mathbb{C}^n, \ l \in [(j-1)m+1 : jm]$ such that

$$\mathbf{G}^j \mathbf{u}_l = \mathbf{I}_l^{mK}, \ \forall l \in [(j-1)m+1 : jm], \tag{4.18}$$

where $\mathbf{G}^j$ is the $mK \times n$ matrix whose lk^{th} element is defined as $\mathbf{G}^j_{lk} = \mathbf{M}_{\lceil \frac{l}{m} \rceil j} g_{\lceil \frac{l}{m} \rceil j}[k] \mathbf{T}_{lk}$, and if it can find such $\mathbf{u}_l$'s, it will reconstruct a noisy, but interference-free, version of each symbol $\tilde{X}_l$ by projecting Y^n_j along the direction of $\mathbf{u}_l$, i.e.

$$(Y^n_j)^T \mathbf{u}_l = \tilde{X}_l + (Z^n_j)^T \mathbf{u}_l, \ \forall l \in [(j-1)m + 1 : jm].$$

Remark. If the conditions in (4.18) are satisfied, then by using an outer code for each of the symbols $\tilde{X}_l$, $l \in [1 : mK]$, a rate of $C_l = \mathbb{E}\left[\log\left(1 + \frac{P}{\|\mathbf{u}_l\|_2^2}\right)\right] \geq \log(P) - \mathbb{E}\left[\log\left(\|\mathbf{u}_l\|_2^2\right)\right]$ over each symbol can be achieved, where the expectation is taken with respect to the channel gain values. Since $\mathbb{E}\left[\log\left(\|\mathbf{u}_l\|_2^2\right)\right]$ does not scale with the transmit power P, and as shown in Appendix F, its value is finite, the scheme guarantees 1 DoF per symbol.

In the remainder of this section, we will address the conditions that the transmission matrix $\mathbf{T}$ needs to satisfy in order to guarantee the existence of $\mathbf{u}_l$'s satisfying (4.18), hence being able to neutralize the interference at all the receivers. We will then use these conditions to characterize the symmetric DoF that is achievable by structured repetition coding.

Definition 4.7. Consider a K-user interference network with adjacency matrix $\mathbf{M}$ and structured repetition coding with transmission matrix $\mathbf{T} \in \{0, 1\}^{mK \times n}$. For each $j \in [1 : K]$, $\bar{\mathbf{T}}^j$ is an $mK \times n$ matrix whose lk^{th} element is defined as

$$\bar{\mathbf{T}}^j_{lk} = \mathbf{T}_{lk} \mathbf{M}_{\lceil \frac{l}{m} \rceil j}, \ \forall l \in [1 : mK], \ \forall k \in [1 : n].$$

Moreover, $\bar{G}^j$ is defined as the bipartite graph with the set of vertices $\mathcal{V} = \{v_1, ..., v_{mK}\} \cup \{v'_1, ..., v'_n\}$ whose adjacency matrix is $\bar{\mathbf{T}}^j$; i.e. v_l is connected to v'_k if and only if $\bar{\mathbf{T}}^j_{lk} = 1$, $\forall l \in [1 : mK]$, $\forall k \in [1 : n]$. Also, for all $l \in [1 : mK]$, we use the notation $\bar{G}^j \setminus l$ to denote the subgraph of $\bar{G}^j$ with node v_l and its incident edges removed.

The above definitions make us ready to state our theorem about the graph theoretic conditions that a transmission matrix $\mathbf{T}$ needs to satisfy to achieve a symmetric DoF of $\frac{m}{n}$.

Theorem 4.5. *Consider a K-user interference network with adjacency matrix $\mathbf{M}$. If a transmission matrix $\mathbf{T} \in \{0,1\}^{mK \times n}$ satisfies the following conditions*

$$\mu(\bar{G}^j) - \mu(\bar{G}^j \setminus l) = 1, \ \forall l \in [(j-1)m+1 : jm], \ \forall j \in [1:K],$$

where $\bar{G}^j$ and $\bar{G}^j \setminus l$ are defined in Definition 4.7, then for almost all values of channel gains, there exist vectors $\{\mathbf{u}_l\}_{l=1}^{mK}$ satisfying

$$(Y_j^n)^T \mathbf{u}_l = \tilde{X}_l + (Z_j^n)^T \mathbf{u}_l, \ \forall l \in [(j-1)m+1 : jm], \ \forall j \in [1:K],$$

where Y_j^n and $\tilde{X}_l$ are defined in Definition 4.6. Hence, structured repetition coding with transmission matrix $\mathbf{T}$ achieves the symmetric DoF of $\frac{m}{n}$.

Theorem 4.5 immediately leads to the following corollary.

Corollary 4.2. *Consider a K-user interference network with adjacency matrix $\mathbf{M}$. Then, the following symmetric DoF is achievable by structured repetition coding (see Definition 4.6).*

$$\begin{aligned} &\sup_{n \in \mathbb{N}} \max_{m \in [1:n]} \frac{m}{n} \\ &\qquad \text{s.t. } \exists \mathbf{T} \in \{0,1\}^{mK \times n} : \mu(\bar{G}^j) - \mu(\bar{G}^j \setminus l) = 1, \\ &\qquad\qquad\qquad\qquad \forall l \in [(j-1)m+1 : jm], \ \forall j \in [1:K], \end{aligned}$$

where $\bar{G}^j$ and $\bar{G}^j \setminus l$ are defined in Definition 4.7.

Remark. While in the optimization problem of Corollary 4.2, the value of $\max_{m \in [1:n]} \frac{m}{n}$ is optimized over $n \in \mathbb{N}$, we will limit the range space for n to be bounded as $n \in [1 : K + 1]$ in order to numerically evaluate the inner bounds in Section 4.4, and as we will see, the inner bounds derived after this reduction match the outer bounds in most of the topologies. This reduces the optimization problem in Corollary 4.2 to a combinatorial optimization problem that can be solved for relatively small networks. Finding efficient algorithms to solve it for general networks is an interesting open problem.

Remark. The structured repetition coding scheme illustrates the fact that even in the case where the channel gains change i.i.d. over time (i.e. coherence time of 1 time slot), it is possible to exploit network topology in order to design a carefully-chosen repetition pattern at the transmitters which enables the receivers to neutralize all the interference. However, as the coherence time of the channel increases, there would be other opportunities that can be utilized, such as aligning the interference, as in [31, 34].

The existence of a vector $\mathbf{u}_l$ satisfying the conditions in Theorem 4.5 is equivalent to the existence of a vector $\mathbf{u}_l$ satisfying the conditions in (4.18), i.e. $\mathbf{G}^j\mathbf{u}_l = \mathbf{I}_l^{mK}$, where $\mathbf{G}^j$ is an $mK \times n$ matrix whose entries are either zero or i.i.d. random variables (corresponding to the channel gains g_{ij}). This enables us to use the following lemma, proved in Appendix C, which addresses the existence of $\mathbf{u}_l$'s satisfying $\mathbf{G}^j\mathbf{u}_l = \mathbf{I}_l^{mK}$ for such structured random matrices $\mathbf{G}^j$.

Lemma 4.3. *Consider a bipartite graph $G = (\{v_1, ..., v_m\} \cup \{v'_1, ..., v'_n\}, \mathcal{E})$ with a corresponding $m \times n$ adjacency matrix $\mathbf{T}$ where $\mathbf{T}_{ij} = 1$ if v_i is connected to v'_j and $\mathbf{T}_{ij} = 0$ otherwise. Also, define $\tilde{\mathbf{T}}$ to be an $m \times n$ matrix for which $\tilde{\mathbf{T}}_{ij} = g_{ij}\mathbf{T}_{ij}$, where g_{ij}'s are i.i.d. random variables drawn from a continuous distribution. If for some $l \in [1 : m]$,*

$\mu(G) - \mu(G \setminus l) = 1$ (where $G \setminus l$ denotes the subgraph of G with node v_l and its incident edges removed), then for almost all values of g_{ij}'s, there exists a vector $\mathbf{u} \in \mathbb{C}^n$ such that

$$\tilde{\mathbf{T}}\mathbf{u} = \mathbf{I}_l^m,$$

where $\mathbf{I}_l^m$ is the l^{th} column of the $m \times m$ identity matrix. Moreover, $\|\mathbf{u}\|_2 = \left\|(\tilde{\mathbf{T}}^l)^{-1}\mathbf{I}_l^{\mu(G)}\right\|_2$, where $\tilde{\mathbf{T}}^l$ is a $\mu(G) \times \mu(G)$ submatrix of $\tilde{\mathbf{T}}$ corresponding to a maximum matching in G.

Proof of Theorem 4.5. Following Definitions 4.6 and 4.7, the received vector of receiver j ($j \in [1 : K]$) can be written as

$$\begin{aligned} Y_j^n &= \sum_{l=1}^{mK} \mathbf{M}_{\lceil \frac{l}{m} \rceil j} g^n_{\lceil \frac{l}{m} \rceil j} \mathbf{T}^T_{l,*} \tilde{X}_l + Z_j^n \\ &= \sum_{l=1}^{mK} g^n_{\lceil \frac{l}{m} \rceil j} (\bar{\mathbf{T}}^j_{l,*})^T \tilde{X}_l + Z_j^n, \end{aligned}$$

and it needs vectors $\{\mathbf{u}_l\}_{l=(j-1)m+1}^{jm}$ such that

$$(Y_j^n)^T \mathbf{u}_l = \tilde{X}_l + (Z_j^n)^T \mathbf{u}_l, \ \forall l \in [(j-1)m + 1 : jm]. \tag{4.19}$$

This means that for almost all values of the channel gains, there must exist vectors $\{\mathbf{u}_l\}_{l=(j-1)m+1}^{jm}$ satisfying

$$\mathbf{G}^j \mathbf{u}_l = \mathbf{I}_l^{mK}, \ \forall l \in [(j-1)m + 1 : jm],$$

where $\mathbf{I}_l^{mK}$ is the l^{th} column of the $mK \times mK$ identity matrix, and $\mathbf{G}^j$ is an $mK \times n$ matrix whose lk^{th} element is defined as $\mathbf{G}^j_{lk} = g_{\lceil \frac{l}{m} \rceil j}[k] \bar{\mathbf{T}}^j_{lk}$. Due to the specific structure of the transmission matrix $\mathbf{T}$ described in (4.17), $\mathbf{G}^j$ has i.i.d. random entries and zeros wherever $\bar{\mathbf{T}}^j$ has ones and zeros, respectively. This enables us to make use of Lemma 4.3, therefore proving the existence of vectors $\mathbf{u}_l$, $\forall l \in [(j-1)m + 1 : jm]$, $\forall j \in [1 : K]$.

The only remaining issue to address is the noise variance in (4.19). The capacity of the channel in (4.19) is equal to

$$C_l = \mathbb{E}\left[\log\left(1 + \frac{P}{\|\mathbf{u}_l\|_2^2}\right)\right], \tag{4.20}$$

where the expectation is taken with respect to the channel gains. Lemma 4.3 implies that $\|\mathbf{u}_l\|_2 = \left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2$, where $\tilde{\mathbf{G}}^{j,l}$ is a $\mu(\bar{G}^j) \times \mu(\bar{G}^j)$ submatrix of $\mathbf{G}^j$ corresponding to a maximum matching in $\bar{G}^j$. Combining this with (4.20), we can write

$$\begin{aligned} C_l &\geq \mathbb{E}\left[\log\left(\frac{P}{\|\mathbf{u}_l\|_2^2}\right)\right] \\ &= \log(P) - \mathbb{E}\left[\log\left(\|\mathbf{u}_l\|_2^2\right)\right] \\ &= \log(P) - \mathbb{E}\left[\log\left(\left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2^2\right)\right]. \end{aligned} \tag{4.21}$$

Now, note that $\mathbb{E}\left[\log\left(\left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2^2\right)\right]$ does not scale with the transmit power P and as we show in Appendix F, its value is finite. Therefore, the outer code on each of the symbols $\tilde{X}_l$ guarantees 1 degree-of-freedom to be achieved for that symbol.

Hence, if all the conditions of the theorem are satisfied, then all the receivers can create interference-free versions of their m desired symbols, implying that structured repetition coding with transmission matrix $\mathbf{T}$ can achieve the symmetric DoF of $\frac{m}{n}$. □

Theorem 4.5, therefore, implies that for any given network topology, it suffices to carefully choose a well-structured transmission matrix $\mathbf{T} \in \{0,1\}^{mK\times n}$ which satisfies the graph theoretic conditions mentioned in the theorem. This makes the symmetric DoF of $\frac{m}{n}$ achievable through structured repetition coding.

4.4 Numerical Analysis

In this section, we will evaluate our inner and outer bounds for two diverse classes of network topologies. We will examine the possible network instances in two scenarios of 6-user networks with 6 square cells and 6-user networks with 1 central and 5 surrounding base stations. The goal is to study the tightness of our inner and outer bounds, compare the performance of the achievable schemes presented in Section 4.3, and study the effect of network density on the fraction of topologies in which structured repetition coding outperforms benchmark schemes. Note that for the structured repetition coding scheme, we search over all transmission matrices $\mathbf{T} \in \{0, 1\}^{mK \times n}$ for which $n \leq K + 1$, due to computational tractability. We seek to answer the following questions.

- Do there exist any network topologies in which our inner and outer bounds on the symmetric DoF do not meet? On the other hand, among the networks in which the bounds are tight, what are the possible values of the symmetric DoF and how are these values distributed?
- Focusing on the topologies in which the inner and outer bounds meet, what is the impact of the sparsity or density of the network graph on the gains that can be obtained beyond the benchmark schemes by using only the knowledge about network topology?
- What is the fraction of the topologies in which structured repetition coding can outperform the benchmark schemes? How much can the sole knowledge about network topology help to go beyond random Gaussian coding and interference avoidance?

We will address these questions in the following classes of networks.

4.4.1 6-User Networks with 6 Square Cells

The networks that we consider in this section are represented by 6 square cells, each one having a base station BS_i in the center, $i \in [1:6]$, with a mobile user inside the cell. An example can be seen in Figure 4.7.

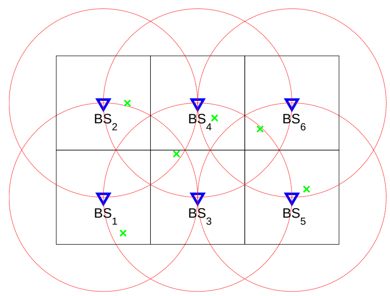


Figure 4.7: A 6-cell network realization where the blue triangles, green crosses, black squares and red circles represent base stations, mobile users, cell boundaries and coverage area of base stations, respectively.

In this figure, the blue triangles represent base stations, the green crosses represent mobile users, the black squares represent the cells and the red circles depict the coverage area of each base station. It is obvious that any placement of the mobile users corresponds to a partially-connected 6-user interference network.

In what follows, we will generalize this model to all possible topologies in which a mobile user in a cell can receive interference from any *nonempty subset* of its three adjacent BS's, together with the signal from its own BS. For instance in Figure 4.7, user 2 can receive interference from any nonempty subset of $\{\mathrm{BS}_1, \mathrm{BS}_3, \mathrm{BS}_4\}$ and user 4 can receive interference from any nonempty subset of $\{\mathrm{BS}_1, \mathrm{BS}_2, \mathrm{BS}_3\}$ or $\{\mathrm{BS}_3, \mathrm{BS}_5, \mathrm{BS}_6\}$ (corresponding to left and right halves of the cell, respectively). This implies that the degree of each receiver is no less than

2 and no more than 4. Ignoring isomorphic topologies, there are in total 22,336 unique topologies in this class. For each of these topologies, we evaluated our inner and outer bounds to draw the following conclusions.

1. We note that quite interestingly, our bounds are tight for all cases, except for 16 distinct topologies. For the remaining networks, which we will hereby focus on, the gap is zero, implying that our bounds determine the symmetric DoF for most networks in this class. In these networks, the symmetric DoF only takes 4 distinct values in $\{\frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}\}$ with the distribution shown in Figure 4.8.

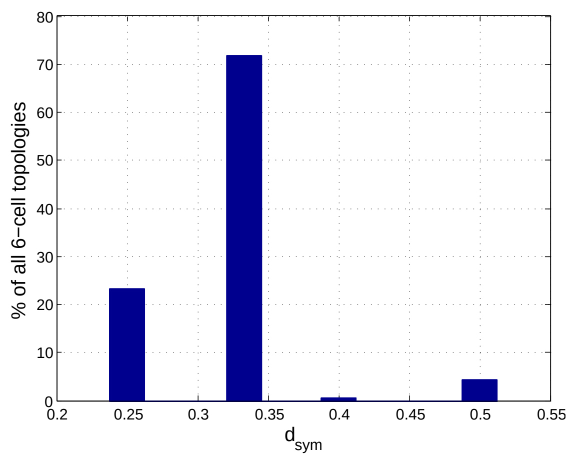


Figure 4.8: Distribution of d_{sym} among 6-cell networks in which our bounds are tight.

2. Figure 4.9 illustrates the impact of the number of interfering links on the performance of structured repetition coding compared to benchmark schemes. As it is clear, the gain is not much when the network is too dense. However, if the density of the network, characterized by the number of cross links in the network, is at a moderate level, then the gain of structured repetition coding over the benchmark schemes can be significant. It

is worth mentioning that there are totally around 50 percent and 10 percent of the networks in which structured repetition coding outperforms random Gaussian coding and interference avoidance, respectively. Moreover, structured repetition coding outperforms *both* benchmark schemes in 1167 network topologies, which constitute more than 5 percent of all the networks. This means that even with a sole knowledge of network topology, one can perform better than both of the benchmark schemes.

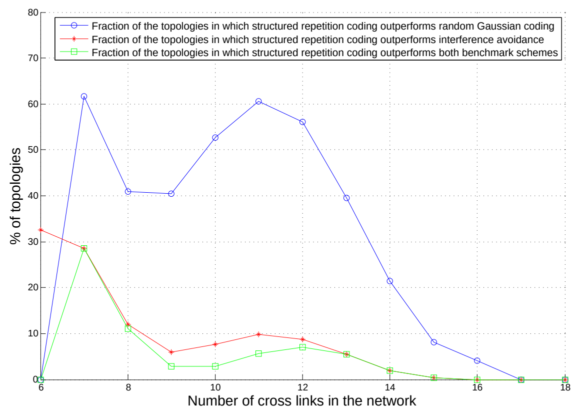


Figure 4.9: Effect of network density on the fraction of networks in which structured repetition coding outperforms benchmark schemes in 6-user cellular networks.

3. Turning our focus to the networks where structured repetition coding outperforms the benchmark schemes, it is interesting to know the value of the gains obtained over them. Among the networks in which structured repetition coding outperforms random Gaussian coding, the gain of the former scheme over the latter takes 5 distinct values in $\{\frac{6}{5}, \frac{4}{3}, \frac{3}{2}, \frac{8}{5}, 2\}$, distributed as shown in Figure 4.10(a). Also, Figure 4.10(b) illustrates the distribution of the gain of structured repetition coding over interference avoidance among the networks in which this gain is greater than unity. This gain can take 4 distinct values in $\{\frac{6}{5}, \frac{5}{4}, \frac{4}{3}, \frac{3}{2}\}$.

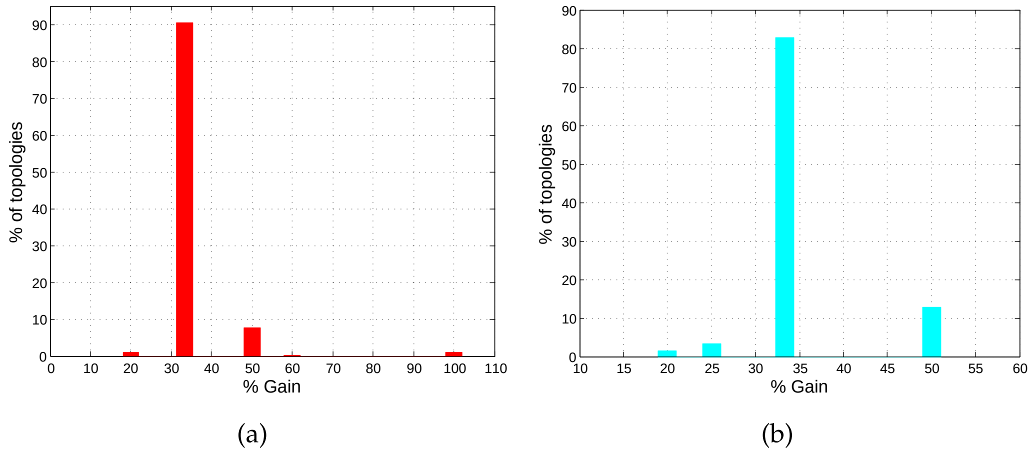


Figure 4.10: Comparison of achievable schemes in 6-user cellular networks: (a) Distribution of the gain of structured repetition coding over $\frac{1}{\Delta_R}$ (random Gaussian coding), and (b) Distribution of the gain of structured repetition coding over interference avoidance.

4. Among all the network topologies, there are 14 topologies which yield the highest gains over both random Gaussian coding and interference avoidance. As an example, one of these networks is depicted in Figure 4.11.

In the network of Figure 4.11 (and all the other 13 networks which yield the highest gains), d_{sym} is equal to $\frac{1}{2}$, which can be achieved by structured repetition coding. However, the best symmetric DoF achieved by random Gaussian coding is $\frac{1}{4}$, hence a gain of 2 can be obtained over this scheme. This implies that for all these 14 networks, there exists a receiver whose degree is 4 (receiver D_6 in Figure 4.11(a)). Moreover, another pattern that is common among these 14 "high-yield" topologies is that the three users which are interfering to the receiver with degree 4 are mutually non-interfering, hence constituting an independent set (users {3,4,5} in Figure 4.11(a)). The third common feature of all these topologies is con-

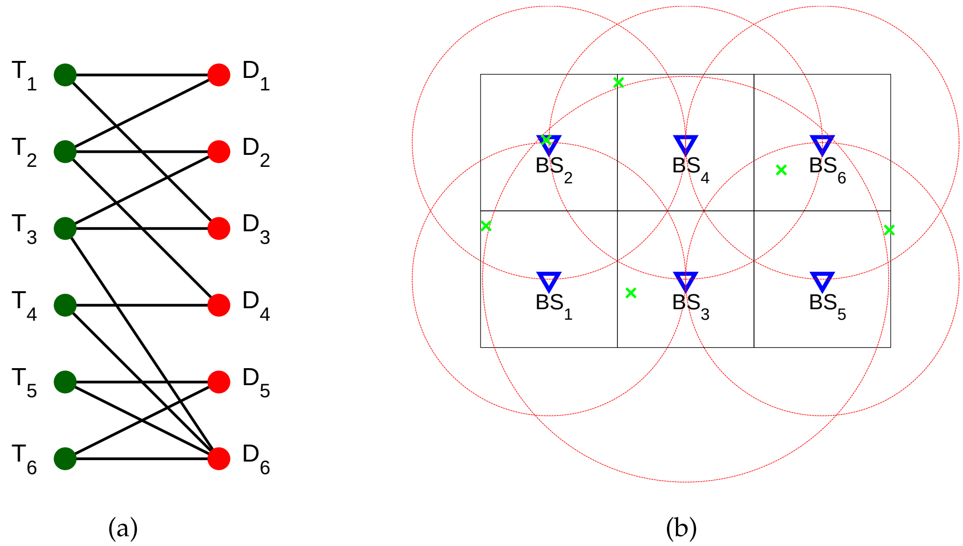


Figure 4.11: (a) A 6-user interference network in which $d_{sym} = \frac{1}{2}$ and the gain of structured repetition coding over random Gaussian coding and interference avoidance is 2 and $\frac{3}{2}$, respectively, and (b) a corresponding 6-cell realization.

taining a 3-user cyclic chain (a 3-user network with users i,j and k where T_i is connected to D_j, T_j is connected to D_k, and T_k is connected to D_i). The subgraph consisting of users $\{1,2,3\}$ in Figure 4.11(a) is a 3-user cyclic chain. This is the main reason that interference avoidance can achieve no better than the symmetric DoF of $\frac{1}{3}$ in these networks, allowing structured repetition coding to have a gain of $\frac{3}{2}$ over it.

4.4.2 6-User Networks with 1 Central and 5 Surrounding Base Stations

In this section, we explore another class of 6-user networks, consisting of 1 base station (BS) located in the center of a circle with radius 1, and 5 other base stations located uniformly on the boundary of the circle. Each base station has a coverage radius of $r < 1$, with a mobile client randomly located in its coverage area. A realization of such a network scenario is illustrated in Figure 4.12.

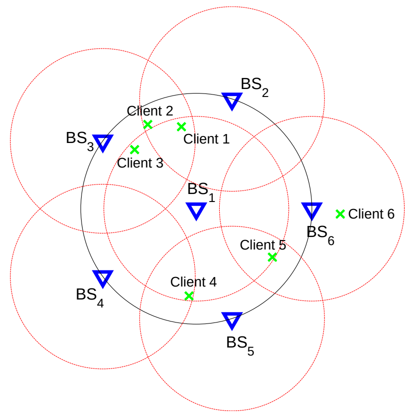


Figure 4.12: A 6-user network realization with 1 BS in the middle and 5 BS's surrounding it, where the blue triangles, green crosses, black circle and red circles represent base stations, mobile clients, unit circle and coverage area of base stations, respectively. In this figure, $r = 0.8$.

Again, as we had in Figure 4.7, the blue triangles represent base stations, the green crosses represent mobile clients, the black circle represents the unit circle and the red circles depict the coverage area of each BS. Obviously, any placement of the mobile clients corresponds to a partially-connected 6-user interference network.

To analyze our bounds for this class of networks, we generated 12000 network instances by randomly locating the mobile clients for the case of $r = 0.8$. Upon removing isomorphic graphs, we ended up with 1507 distinct topologies and evaluated our inner and outer bounds for these topologies, leading to the following conclusions.

1. We find out interestingly, that our bounds are tight in all the generated net-

work topologies, and Figure 4.13 illustrates the distribution of d_{sym} among these topologies. We note that d_{sym} takes 4 distinct values in $\{\frac{1}{3}, \frac{2}{5}, \frac{1}{2}, 1\}$. The most frequent value that d_{sym} takes is $\frac{1}{3}$, followed by $\frac{1}{2}$.

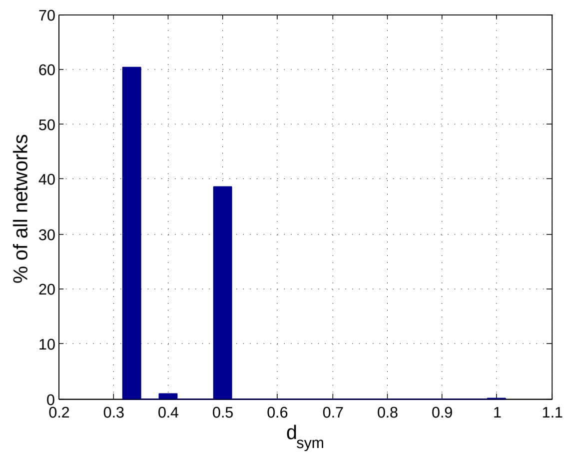


Figure 4.13: Distribution of d_{sym} among 6-user networks with 1 central and 5 surrounding BS's, where each BS has a coverage radius of $r = 0.8$.

2. Figure 4.14 illustrates the effect of the number of cross links in the network, which is a measure of density of the network graph, on the fraction of topologies which yield gains over benchmark schemes. The trend is similar to that of Figure 4.9, showing that if the network graph is too sparse (few number of cross links) or too dense (high number of cross links), there is not much gain beyond the benchmark schemes. However, if the network graph is moderately dense, then structured repetition coding can attain gain over the benchmark schemes in a larger fraction of networks. Moreover, the figure implies that, on average, interference avoidance yields higher inner bounds on d_{sym} than random Gaussian coding, in this class of networks.

3. Figure 4.15(a) illustrates the distribution of the gain of structured repetition coding over random Gaussian coding among the topologies in which

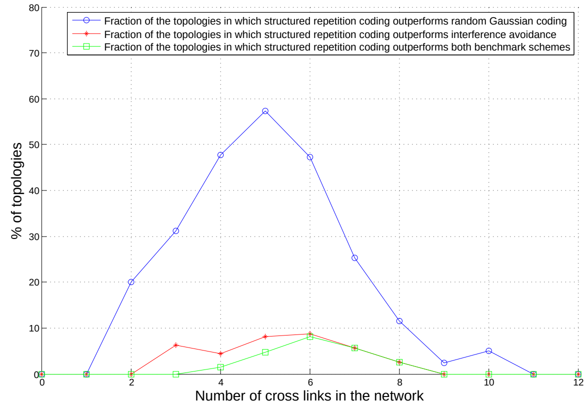
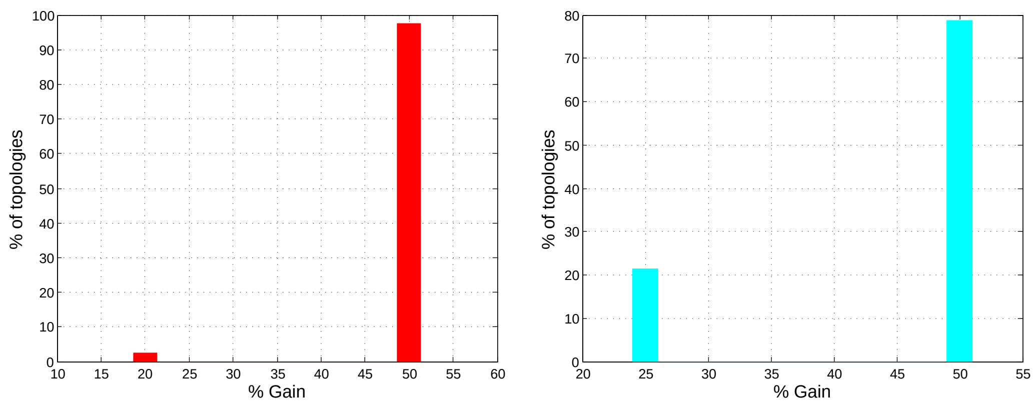


Figure 4.14: Effect of network density on the fraction of networks in which structured repetition coding outperforms benchmark schemes in 6-user networks with 1 central and 5 surrounding BS's, where each BS has a coverage radius of $r = 0.8$.

this gain is greater than unity. This gain can take 2 distinct values in $\{\frac{6}{5}, \frac{3}{2}\}$. Moreover, among the networks in which there is a gain over interference avoidance, this gain can take 2 distinct values in $\{\frac{5}{4}, \frac{3}{2}\}$, with the distribution shown in Figure 4.15(b). The most frequent value of both of the gains is $\frac{3}{2}$, which indicates a 50% improvement in the inner bound on d_{sym}.



(a) (b)

Figure 4.15: Comparison of achievable schemes in 6-user networks with 1 central and 5 surrounding BS's, where each BS has a coverage radius of $r = 0.8$: (a) Distribution of the gain of structured repetition coding over $\frac{1}{\Delta_R}$ (random Gaussian coding), and (b) Distribution of the gain of structured repetition coding over interference avoidance.

APPENDIX A

PROOF OF THEOREM 2.5

We prove the theorem in two steps.

- Step 1: $\bigcup_{\mathcal{S}\subseteq\{1,...,K\}} \mathcal{P}_{\mathcal{S}} \subseteq \mathcal{P}^*$. It suffices to show that for all $\mathcal{S} \subset \{1, ..., K\}$, $\mathcal{P}_{\mathcal{S}} \subseteq \mathcal{P}^*$; i.e., the region $\mathcal{P}_{\mathcal{S}}$ can be achieved through TIN. Note that if $\mathcal{S} = \emptyset$, then $\mathcal{P}_{\mathcal{S}} = \mathcal{P}_{\emptyset} = \mathcal{P} \subseteq \mathcal{P}^*$.

  Now, if $\mathcal{S} \neq \emptyset$, then to make the users in $\mathcal{S}$ silent, we set $r_i = -\infty$, $\forall i \in \mathcal{S}$. This forces $d_i = 0$, $\forall i \in \mathcal{S}$. Then, for the remaining users, i.e., the users in $\mathcal{S}^c$, we use polyhedral TIN. Therefore, the polyhedral TIN region where all the users in $\mathcal{S}$ are removed from the network, can be achieved. This region is in fact $\mathcal{P}_{\mathcal{S}}$, and hence, $\mathcal{P}_{\mathcal{S}} \subseteq \mathcal{P}^*$.

- Step 2: $\mathcal{P}^* \subseteq \bigcup_{\mathcal{S}\subseteq\{1,...,K\}} \mathcal{P}_{\mathcal{S}}$. To prove this, we first define the sets $\tilde{\mathcal{P}}_{\mathcal{S}}$ as $\mathcal{P}_{\mathcal{S}}$ restricted to strictly positive GDoF's for users in $\mathcal{S}^c$; i.e.,

  $$\tilde{\mathcal{P}}_{\mathcal{S}} = \{(d_1, ..., d_K) \in \mathcal{P}_{\mathcal{S}} : d_i > 0, \forall i \in \mathcal{S}^c\},$$

  for any $\mathcal{S} \subseteq \{1, 2, ..., K\}$. It is obvious that $\tilde{\mathcal{P}}_{\mathcal{S}} \subseteq \mathcal{P}_{\mathcal{S}}$ and therefore,

  $$\bigcup_{\mathcal{S}\subseteq\{1,...,K\}} \tilde{\mathcal{P}}_{\mathcal{S}} \subseteq \bigcup_{\mathcal{S}\subseteq\{1,...,K\}} \mathcal{P}_{\mathcal{S}}. \tag{A.1}$$

  Now, we show that any GDoF point $(d_1, ..., d_K)$ lying outside all of the sets $\tilde{\mathcal{P}}_{\mathcal{S}}$ should not be achievable by TIN. Such a point should satisfy at least one of the following conditions:

  - $d_i < 0$ or $d_i > \alpha_{ii}$ for some user $i \in \{1, ..., K\}$. In this case, it is trivial that the GDoF point is not achievable by TIN.

– $\sum_{j=1}^{m} d_{i_j} > \sum_{j=1}^{m}(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j})$ for some cyclic sequence $(i_0, i_1, ..., i_m) \in \Pi_K$ such that $d_{i_j} > 0, \forall j \in \{1, ..., m\}$. In this case, letting $i_{m+1} = i_1$, we have

$$\sum_{j=1}^{m} r_{i_j} + \alpha_{i_j i_j} - \max\{0, \max_{i_k \neq i_j}(r_{i_k} + \alpha_{i_j i_k})\} > \sum_{j=1}^{m}(\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j})$$
$$\Rightarrow \sum_{j=1}^{m} r_{i_j} + \alpha_{i_{j-1} i_j} - \underbrace{\max\{0, \max_{i_k \neq i_j}(r_{i_k} + \alpha_{i_j i_k})\}}_{\geq r_{i_{j+1}} + \alpha_{i_j i_{j+1}}} > 0,$$

which is a contradiction. Therefore in this case, the GDoF point is not achievable by TIN, too.

This implies that $\mathcal{P}^* \subseteq \bigcup_{\mathcal{S} \subseteq \{1,...,K\}} \tilde{\mathcal{P}}_{\mathcal{S}}$, and combining this with (A.1) yields $\mathcal{P}^* \subseteq \bigcup_{\mathcal{S} \subseteq \{1,...,K\}} \mathcal{P}_{\mathcal{S}}$.

Combining steps 1 and steps 2 leads to (2.42), therefore completing the proof.

APPENDIX B

PROOF OF COROLLARY 3.1

In the closest-AP selection model, n sources and n destinations are uniformly and independently located within a circle of radius R on the plane and then each destination is associated with its *closest* source. If there are n points (sources) located uniformly within a circle of radius R, then the probability that the minimum distance of a new point (destination) in the circle to the closest source, denoted by r_{min}, is greater than a threshold d is equal to

$$\begin{aligned}\mathbb{P}(r_{min} > d) &= \mathbb{P}(\text{no BS within distance } d \text{ of the destination})\\ &= \left(\frac{\pi R^2 - \pi d^2}{\pi R^2}\right)^n\\ &= \left(1 - \left(\frac{d}{R}\right)^2\right)^n.\end{aligned}$$

Denote the distance of destination i to its closest source as r_i. Then the probability that all the destinations are within a distance $d = Rn^{-\beta}$ of their corresponding sources can be lower bounded as

$$\begin{aligned}\mathbb{P}[\max_i r_i \le d] &= \mathbb{P}[r_1 \le d \;\&\; r_2 \le d \;\&\; \ldots \;\&\; r_n \le d]\\ &= 1 - \mathbb{P}[r_1 > d \text{ or } r_2 > d \text{ or } \ldots \text{ or } r_n > d]\\ &\ge 1 - n\,\mathbb{P}[r_1 > d] &\text{(B.1)}\\ &= 1 - n\left(1 - \left(\frac{d}{R}\right)^2\right)^n\\ &= 1 - n\left(1 - \frac{1}{n^{2\beta}}\right)^n, &\text{(B.2)}\end{aligned}$$

where in (B.1), we have used the union bound and the fact that r_i's are identically distributed. If $\beta < \frac{1}{2}$, it is easy to show that the expression in (B.2) goes to 1 as n goes to infinity because of the following lower bound:

$$1 - \frac{n}{e^{n^{1-2\beta}}} \leq 1 - n\left(1 - \frac{1}{n^{2\beta}}\right)^n,$$

which goes to 1 as n goes to infinity if $2\beta < 1$. Hence, the closest-AP selection model is almost-surely a special class of the model mentioned in Section 3.1.2 for any $\beta < \frac{1}{2}$. Moreover, in this case the first part of Theorem 3.1 shows that ITLinQ can almost-surely achieve a fraction of the capacity region proportional to $n^{\beta-1}$. Therefore, in the closest-AP selection model, ITLinQ is able to almost-surely achieve a fraction $\lambda = \frac{\sqrt{3}\pi R^2}{2\gamma^2} n^{\beta-1}$ of the capacity region to within a gap of $k \leq \frac{\sqrt{3}\pi R^2}{2\gamma^2} \frac{\log 3n}{n^{1-\beta}}$ when $n \to \infty$, for any $\beta < \frac{1}{2}$.

APPENDIX C

PROOF OF LEMMA 4.1

$$
\begin{aligned}
H(W|Y^n + Z_2^n) &= H(W|Y^n + Z_2^n, Z_1^n - Z_2^n) + I(W; Z_1^n - Z_2^n|Y^n + Z_2^n) \\
&\leq H(W|Y^n + Z_1^n) + h(Z_1^n - Z_2^n|Y^n + Z_2^n) - h(Z_1^n - Z_2^n|Y^n + Z_2^n, W) \\
&\leq n\epsilon + h(Z_1^n - Z_2^n) - h(Z_1^n - Z_2^n|Y^n + Z_2^n, W, Z_2^n) \\
&= n\epsilon + h(Z_1^n - Z_2^n) - h(Z_1^n) \\
&= n\epsilon + n\log(\pi e(N+1)) - n\log(\pi e) \\
&= n\epsilon + n\log(N+1).
\end{aligned}
$$

APPENDIX D

PROOF OF LEMMA 4.2

Without loss of generality, let $\mathcal{S} = [1 : m]$ and $\mathcal{S}' = [1 : m']$ ($m' \leq m$). Also, with respect to $\mathbf{c}$ being a fractional generator of $\mathcal{S}'$ in $\mathcal{S}$, suppose (without loss of generality) that $\Pi_{\mathcal{S}'} = (1, ..., m')$. First, note that

$$h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n | \mathcal{G}^n) - h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n | W_1, ..., W_{m'}, \mathcal{G}^n)$$
$$= H(W_1, ..., W_{m'}) - H(W_1, ..., W_{m'} | \sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n, \mathcal{G}^n),$$

since both sides are equal to $I(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n; W_1, ..., W_{m'} | \mathcal{G}^n)$. Therefore, we can write

$$\begin{aligned}
&h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n | W_1, ..., W_{m'}, \mathcal{G}^n) \\
&= H(W_1, ..., W_{m'} | \sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n, \mathcal{G}^n) \\
&\quad + h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n | \mathcal{G}^n) - H(W_1, ..., W_{m'}) \\
&\leq H(W_1, ..., W_{m'} | \sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n, \mathcal{G}^n) \\
&\quad + n(\log(P) - \sum_{i \in \mathcal{S}'} R_i) + no(\log(P)).
\end{aligned} \tag{D.1}$$

Now, we prove that $H(W_l | \sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n, W_1, ..., W_{l-1}, \mathcal{G}^n) \leq no(\log(P)) + n\epsilon_{l,n}$ for $l \in [1 : m']$. By Definition 4.3, $\mathbf{M}_l^{\mathcal{S}} \in_l^{\pm} \text{span}\,(\mathbf{c} + \sum_{k=1}^{m'} \mathbf{I}_k^{|\mathcal{S}|}, \mathbf{I}_{\{1,...,l-1\}}^{|\mathcal{S}|})$, implying that there exist a vector $\tilde{\mathbf{v}} \in \mathbb{R}^{|\mathcal{S}|}$ and coefficients α and d_k ($k \in [1 : l-1]$) such that

$$\tilde{\mathbf{v}} = \alpha \left(\mathbf{c} + \sum_{k=1}^{m'} \mathbf{I}_k^{|\mathcal{S}|} \right) + \sum_{k=1}^{l-1} d_k \mathbf{I}_k^{|\mathcal{S}|} \tag{D.2}$$

$$|\tilde{\mathbf{v}}_l| = |\mathbf{M}^S_{ll}| = 1 \tag{D.3}$$

$$\tilde{\mathbf{v}}_j(|\tilde{\mathbf{v}}_j| - |\mathbf{M}^S_{jl}|) = 0,\ \forall j \in [1:m] \setminus \{l\}. \tag{D.4}$$

Note that if $j \in \mathcal{IF}_l$, then $\mathbf{M}^S_{jl} = 1$ and (D.4) implies that $\tilde{\mathbf{v}}_j$ can either be equal to 0 or ±1; i.e. $\tilde{\mathbf{v}}_j \in \{0, \pm1\}$. On the other hand, if $j \notin \mathcal{IF}_l$, then $\mathbf{M}^S_{jl} = 0$ and (D.4) implies that $\tilde{\mathbf{v}}_j = 0$. Multiplying $\begin{bmatrix} g^n_1 X^n_1 & \dots & g^n_m X^n_m \end{bmatrix}$ by both sides of (D.2), hence, yields

$$\tilde{\mathbf{v}}_l g^n_l X^n_l + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g^n_j X^n_j = \alpha\left(\sum_{j=1}^{m} \mathbf{c}_j g^n_j X^n_j + \sum_{k=1}^{m'} g^n_k X^n_k \right) + \sum_{k=1}^{l-1} d_k g^n_k X^n_k.$$

Therefore, we can write:

$$\begin{aligned}
&H\left(W_l \middle| \alpha\left(\sum_{j=1}^{m} \mathbf{c}_j g^n_j X^n_j + \sum_{k=1}^{m'} g^n_k X^n_k + Z^n \right) + \sum_{k=1}^{l-1} d_k g^n_k X^n_k, \mathcal{G}^n \right) \\
&\quad = H(W_l | \tilde{\mathbf{v}}_l g^n_l X^n_l + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g^n_j X^n_j + \alpha Z^n, \mathcal{G}^n) \\
&\quad = H(W_l | \tilde{\mathbf{v}}_l g^n_l X^n_l + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g^n_j X^n_j + \alpha Z^n, \sum_{j \in \mathcal{IF}_l} (1 - |\tilde{\mathbf{v}}_j|) g^n_j X^n_j, \mathcal{G}^n) \qquad \text{(D.5)} \\
&\quad \le H(W_l | \tilde{\mathbf{v}}_l g^n_l X^n_l + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}'_j g^n_j X^n_j + \alpha Z^n, \mathcal{G}^n) \qquad \text{(D.6)} \\
&\quad \le no(\log(P)) + n\epsilon_{l,n}, \qquad \text{(D.7)}
\end{aligned}$$

where (D.5) is true because, as discussed before, for all $j \in \mathcal{IF}_l$, $\tilde{\mathbf{v}}_j$ can only take the values in $\{\pm1, 0\}$ and therefore the signals in $\sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}_j g^n_j X^n_j$ and $\sum_{j \in \mathcal{IF}_l} (1 - |\tilde{\mathbf{v}}_j|) g^n_j X^n_j$ do not have common terms.[1] In (D.6), $\tilde{\mathbf{v}}'_j$ is defined as $\tilde{\mathbf{v}}'_j := \tilde{\mathbf{v}}_j + (1 - |\tilde{\mathbf{v}}_j|)$. Clearly $\tilde{\mathbf{v}}'_j$ can only take the values in $\{+1, -1\}$ because $\tilde{\mathbf{v}}_j \in \{\pm1, 0\}$. Also, (D.3) implies that $\tilde{\mathbf{v}}_l \in \{+1, -1\}$. Therefore, $\tilde{\mathbf{v}}_l g^n_l X^n_l + \sum_{j \in \mathcal{IF}_l} \tilde{\mathbf{v}}'_j g^n_j X^n_j + \alpha Z^n$ is statistically the same as Y^n_l (with a bounded difference in noise variance), because the channel

[1] If $\tilde{\mathbf{v}}_j = 0$, then $1 - |\tilde{\mathbf{v}}_j| = 1$, and if $\tilde{\mathbf{v}}_j = 1$ or $\tilde{\mathbf{v}}_j = -1$, then $1 - |\tilde{\mathbf{v}}_j| = 0$. Hence, either $\tilde{\mathbf{v}}_j$ or $1 - |\tilde{\mathbf{v}}_j|$ is non-zero, but not both.

gains have a symmetric distribution around zero ($f_G(g) = f_G(-g), \forall g \in \mathbb{C}$). This, together with Lemma 4.1 and Fano's inequality, implies that (D.7) is correct. Hence, using the chain rule for differential entropy yields

$$\begin{aligned} H(W_1, ..., W_{m'} | \sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n, \mathcal{G}^n) \\ &= \sum_{l=1}^{m'} H(W_l | \sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n, W_1, ..., W_{l-1}, \mathcal{G}^n) \\ &\leq \sum_{l=1}^{m'} no(\log(P)) + n\epsilon_{l,n} \\ &= no(\log(P)) + n\epsilon_n. \end{aligned}$$

Therefore, (D.1) can be written as

$$h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n | W_1, ..., W_{m'}, \mathcal{G}^n) \leq n(\log(P) - \sum_{i \in \mathcal{S}'} R_i) + no(\log(P)) + n\epsilon_n. \tag{D.8}$$

But note that

$$h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + \sum_{k=1}^{m'} g_k^n X_k^n + Z^n | W_1, ..., W_{m'}, \mathcal{G}^n) = h(\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + Z^n | \mathcal{G}^n), \tag{D.9}$$

because by Definition 4.3, $\mathbf{c}_j = 0, \forall j \in \mathcal{S}'$ and therefore, $\sum_{j=1}^{m} \mathbf{c}_j g_j^n X_j^n + Z^n$ is independent of $W_1, ..., W_{m'}$. The lemma then follows from (D.8) and (D.9).

APPENDIX E

PROOF OF LEMMA 4.3

The fact that $\mu(G) - \mu(G \setminus l) = 1$ means that there exists a maximum matching in G which covers node v_l; i.e. there is one edge in the matching incident on v_l. This matching covers $\mu(G)$ vertices out of $\{v_1, ..., v_m\}$, including v_l for sure, and $\mu(G)$ vertices out of $\{v'_1, ..., v'_n\}$ (note that for the bipartite graph G, $\mu(G) \le \min\{m, n\}$). Therefore, it corresponds to a $\mu(G) \times \mu(G)$ submatrix of the entire adjacency matrix $\mathbf{T}$, which we will denote by $\mathbf{T}^l$, and we know that $\mathbf{T}^l$ includes (a subset of) the l^{th} row of $\mathbf{T}$. Without loss of generality, we assume that $l \in [1 : \mu(G)]$ and $\mathbf{T}^l$ consists of the first $\mu(G)$ rows and columns of $\mathbf{T}$. We will also denote the corresponding random matrix by $\tilde{\mathbf{T}}^l$; i.e. $\tilde{\mathbf{T}}^l_{ij} = g_{ij}\mathbf{T}^l_{ij}$, $\forall i \in [1 : \mu(G)]$, $\forall j \in [1 : \mu(G)]$.

Now, we show that $\det(\tilde{\mathbf{T}}^l) \neq 0$ for almost all values of g_{ij}'s. This is because

$$\det(\tilde{\mathbf{T}}^l) = \sum_{\sigma \in \Pi_{\mu(G)}} \operatorname{sgn}(\sigma) \prod_{i=1}^{\mu(G)} \tilde{\mathbf{T}}^l_{i\sigma_i}, \tag{E.1}$$

where $\Pi_{\mu(G)}$ is the set of all permutations of $[1 : \mu(G)]$ and $\operatorname{sgn}(\sigma) = 1$ if σ can be derived from $[1 : \mu(G)]$ by doing an even number of switches, and $\operatorname{sgn}(\sigma) = -1$ otherwise. Note that $\det(\tilde{\mathbf{T}}^l)$ is a multivariate polynomial of distinct i.i.d. channel gains (drawn from a continuous distribution), which is not identically zero. The reason that the polynomial is not identically zero is because the matching corresponds to a set of nonzero entries of $\mathbf{T}^l$ (and hence $\tilde{\mathbf{T}}^l$) which do not share common rows/columns, hence constituting a non-zero term in (E.1). Therefore, the Schwartz-Zippel Lemma [35, 36] states that the value of this polynomial is not zero for almost all values of g_{ij}'s.

Therefore, $\tilde{\mathbf{T}}^l$ is invertible with probability 1, implying that there exists a vector $\mathbf{u}' \in \mathbb{C}^{\mu(G)}$ such that $\tilde{\mathbf{T}}^l\mathbf{u}' = \mathbf{I}_l^{\mu(G)}$. In fact, $\mathbf{u}' = (\tilde{\mathbf{T}}^l)^{-1}\mathbf{I}_l^{\mu(G)}$. Now, let $\mathbf{u} \in \mathbb{C}^n$ be the vector such that

$$\mathbf{u}_j = \begin{cases} \mathbf{u}'_j & \text{if } j \in [1:\mu(G)] \\ 0 & \text{if } j \in [\mu(G)+1:n] \end{cases}. \tag{E.2}$$

Now, we claim that $\tilde{\mathbf{T}}\mathbf{u} = \mathbf{I}_l^m$. This is true because of the following.

- $\tilde{\mathbf{T}}_{l,*}\mathbf{u} = \tilde{\mathbf{T}}^l_{l,*}\mathbf{u}' = 1$.
- $\tilde{\mathbf{T}}_{i,*}\mathbf{u} = \tilde{\mathbf{T}}^l_{i,*}\mathbf{u}' = 0, \forall i \in [1:\mu(G)] \setminus \{l\}$.
- Also, each row $\tilde{\mathbf{T}}_{j,*}$ ($j \in [\mu(G)+1:n]$) is linearly dependent on the rows $\tilde{\mathbf{T}}_{i,*}$, $i \in [1:\mu(G)] \setminus \{l\}$, because otherwise, we would have at least $\mu(G)$ independent rows in $\tilde{\mathbf{T}}_{\setminus l,*}$ (the same matrix as $\tilde{\mathbf{T}}$ with the l^{th} row removed) and this corresponds to a matching with a size of at least $\mu(G)$ in $G \setminus l$, contradicting $\mu(G)-1$ being the size of the maximum matching in $G \setminus l$. Therefore, for all $j \in [\mu(G)+1:n]$, there exist coefficients α_{ij} ($i \in [1:\mu(G)] \setminus \{l\}$) such that $\tilde{\mathbf{T}}_{j,*} = \sum_{i\in[1:\mu(G)]\setminus\{l\}} \alpha_{ij}\tilde{\mathbf{T}}_{i,*}$, implying that $\tilde{\mathbf{T}}_{j,*}\mathbf{u} = \sum_{i\in[1:\mu(G)]\setminus\{l\}} \alpha_{ij}\tilde{\mathbf{T}}_{i,*}\mathbf{u} = 0$.

To complete the proof, note that $\|\mathbf{u}\|_2 = \|\mathbf{u}'\|_2 = \left\|(\tilde{\mathbf{T}}^l)^{-1}\mathbf{I}_l^{\mu(G)}\right\|_2$, because of the definition of $\mathbf{u}$ in (E.2).

APPENDIX F

PROOF OF FINITENESS OF NOISE VARIANCE IN (4.21)

In this appendix, we intend to show that in (4.21), $\mathbb{E}\left[\log\left(\left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2^2\right)\right] < \infty$. We need the following key lemma to prove this inequality.

Lemma F.1. [1] *Assume $p(X_1,...,X_n) = \sum_{i=1}^m a_i \prod_{j=1}^n X_j^{d_{ji}}$ ($1 \leq m \leq 2^n$) is a multivariate polynomial of complex i.i.d. random variables $X_1,...,X_n$ with a continuous distribution, where for all $i \in [1:m]$, a_i is a constant coefficient in $\mathbb{C}$ satisfying $|a_i| \geq 1$, all the monomials are assumed to be distinct, and the degree of X_j in the i^{th} monomial, denoted by d_{ji}, satisfies $d_{ji} \in \{0,1\}, \forall j \in [1:n], \forall i \in [1:m]$. If there exists $f_{max} < \infty$ such that $f_{|X|}(r) \leq f_{max}, \forall r \in \mathbb{R}^+$, where $f_{|X|}(r)$ is the distribution of $|X_i|$, $\forall i \in [1:n]$, then for all $\epsilon \in [0,1]$,*

$$\Pr\left[|p(X_1,...,X_n)| \leq \epsilon\right] \leq 2^{n+1} f_{max} \sqrt[2^{n-1}]{\epsilon}.$$

Before proving the lemma, we show how this lemma implies that $\mathbb{E}\left[\log\left(\left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2^2\right)\right] < \infty$ in (4.21). Note that $\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}$ is the l^{th} column of $\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}$, the inverse of $\tilde{\mathbf{G}}^{j,l}$. Therefore,

$$\begin{aligned}\left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}\mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2^2 &= \sum_{i=1}^{\mu(\bar{G}^j)} \left|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1}_{il}\right|^2 \\ &= \frac{1}{\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2} \sum_{i=1}^{\mu(\bar{G}^j)} |M_{li}|^2,\end{aligned}$$

where M_{li} is the determinant of $\tilde{\mathbf{G}}^{j,l}$ after removing its l^{th} row and i^{th} column. Due to the definition of determinants (see (E.1), for instance), each of the terms M_{li} and also $\det\left(\tilde{\mathbf{G}}^{j,l}\right)$ is a multivariate polynomial of i.i.d. channel gains, in which

[1] This lemma has connections to estimating the size of lemniscates of multivariate polynomials, studied in [37, 38]. However, here we prove a different form of upper bound which suits our framework to prove the finiteness of the noise variance in (4.21).

each of the random variables appear with the degree of 0 or 1 in each monomial. In other words, if we rename the i.i.d. channel gains inside $\tilde{\mathbf{G}}^{j,l}$ as $g_1, ..., g_n$, then each M_{li} can be written as

$$M_{li} = \sum_{k=1}^{m_{li}} a_{k,li} \prod_{h=1}^{n} g_h^{d_{k,h,li}}, \tag{F.1}$$

and $\det\left(\tilde{\mathbf{G}}^{j,l}\right)$ can be written as

$$\det\left(\tilde{\mathbf{G}}^{j,l}\right) = \sum_{k=1}^{m} a_k \prod_{h=1}^{n} g_h^{d_{k,h}}, \tag{F.2}$$

where $|a_{k,li}| = |a_k| = 1$ and $d_{k,h,li}, d_{k,h} \in \{0, 1\}$, for all h, k, i. Hence, we can write

$$\begin{aligned}
\mathbb{E}\left[\log\left(\left\|\left(\tilde{\mathbf{G}}^{j,l}\right)^{-1} \mathbf{I}_l^{\mu(\bar{G}^j)}\right\|_2^2\right)\right] &= \mathbb{E}\left[\log\left(\sum_{i=1}^{\mu(\bar{G}^j)} |M_{li}|^2\right)\right] - \mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right] \\
&\leq \log\left(\sum_{i=1}^{\mu(\bar{G}^j)} \mathbb{E}\left[|M_{li}|^2\right]\right) - \mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right] \qquad \text{(F.3)} \\
&\leq \log\left(\sum_{i=1}^{\mu(\bar{G}^j)} m_{li} \sum_{k=1}^{m_{li}} \mathbb{E}\left[\left|\prod_{h=1}^{n} g_h^{d_{k,h,li}}\right|^2\right]\right) - \mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right] \qquad \text{(F.4)} \\
&= \log\left(\sum_{i=1}^{\mu(\bar{G}^j)} m_{li} \sum_{k=1}^{m_{li}} \mathbb{E}\left[|g|^2\right]^{\sum_{h=1}^{n} d_{k,h,li}}\right) - \mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right], \qquad \text{(F.5)}
\end{aligned}$$

where (F.3) follows from Jensen's inequality, (F.4) follows from the Cauchy-Schwarz inequality, and (F.5) follows from g_h's being i.i.d. Hence, if $\mathbb{E}\left[|g|^2\right] < \infty$, then the first term in (F.5) is bounded. Therefore, it remains to show that $\mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right] > -\infty$. We can write

$$\mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right] \geq \mathbb{E}\left[\log\left(\min\left\{1, \left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right\}\right)\right] = -\mathbb{E}[Y], \tag{F.6}$$

where $Y = -\log\left(\min\left\{1, \left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right\}\right)$ is a non-negative random variable. Hence, we have

$$\mathbb{E}[Y] = \int_0^{\infty} \Pr[Y \geq y] dy$$

$$
\begin{aligned}
&= \int_0^\infty \Pr\left[\min\left\{1, \left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right\} \le 2^{-y}\right] dy \\
&= \int_0^\infty \Pr\left[\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2 \le 2^{-y}\right] dy \\
&= \frac{2}{\ln 2}\int_0^1 \Pr\left[\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right| \le u\right] \frac{du}{u} \qquad \text{(F.7)} \\
&\le \frac{2}{\ln 2}\int_0^1 \frac{2^{n+1} f_{max} \sqrt[2^{n-1}]{u}}{u} du \qquad \text{(F.8)} \\
&= \frac{2^{n+2} f_{max}}{\ln 2}\int_0^1 u^{\frac{1}{2^{n-1}}-1} du \\
&= \frac{2^{2n+1} f_{max}}{\ln 2} \\
&< \infty,
\end{aligned}
$$

where in (F.7), we have used the change of variables $u = 2^{-\frac{y}{2}}$ and (F.8) follows from (F.2) and Lemma F.1. This, together with (F.6) implies that $\mathbb{E}\left[\log\left|\det\left(\tilde{\mathbf{G}}^{j,l}\right)\right|^2\right] > -\infty$, hence finishing the proof.

Now, we focus on proving Lemma F.1.

Proof of Lemma F.1. We will use induction on the number of variables (n) to prove the desired inequality.

Base case: We need to prove that for all $\epsilon \le 1$, $\Pr\left[|p(X_1)| \le \epsilon\right] \le 4 f_{max}\epsilon$. In general, $p(X_1) = aX_1 + b$, where $a, b \in \mathbb{C}$ and $|a| \ge 1$ and $|b| \ge 1$. Therefore, we can write

$$
\begin{aligned}
\Pr\left[|p(X_1)| \le \epsilon\right] &= \Pr\left[|aX_1 + b| \le \epsilon\right] \\
&\le \Pr\left[||aX_1| - |b|| \le \epsilon\right] \\
&= \Pr\left[\frac{|b| - \epsilon}{|a|} \le |X_1| \le \frac{|b| + \epsilon}{|a|}\right] \\
&\le 2 f_{max}\epsilon \qquad \text{(F.9)} \\
&< 4 f_{max}\epsilon.
\end{aligned}
$$

Inductive step: Assume for all $\epsilon \in [0,1]$, $\Pr\left[|p(X_1,...,X_{k+1})| \leq \epsilon\right] \leq 2^{k+1} f_{max} \sqrt[2^{k-1}]{\epsilon}$. Now, consider the polynomial $p(X_1,...,X_k,X_{k+1}) = \sum_{i=1}^{m} a_i \prod_{j=1}^{k+1} X_j^{d_{ji}}$. Without loss of generality, we can write this polynomial as

$$p(X_1,...,X_k,X_{k+1}) = \sum_{i=1}^{m'} (a_i X_{k+1} + b_i) \prod_{j=1}^{k} X_j^{d_{ji}} + \sum_{i=m'+1}^{m} a_i \prod_{j=1}^{k} X_j^{d_{ji}}, \tag{F.10}$$

where we first factored out the monomials which include X_{k+1}, and afterwards, we lumped together the monomials that were indistinct in terms of $X_1,...,X_k$.

Now, we can write

$$\begin{aligned}
&\Pr\left[|p(X_1,...,X_k,X_{k+1})| \leq \epsilon\right] \\
&= \Pr\left[|p(X_1,...,X_k,X_{k+1})| \leq \epsilon \middle| \min_{i\in[1:m']} |a_i X_{k+1} + b_i| \leq \sqrt{\epsilon}\right] \Pr\left[\min_{i\in[1:m']} |a_i X_{k+1} + b_i| \leq \sqrt{\epsilon}\right] \\
&\quad + \Pr\left[|p(X_1,...,X_k,X_{k+1})| \leq \epsilon \middle| \min_{i\in[1:m']} |a_i X_{k+1} + b_i| > \sqrt{\epsilon}\right] \Pr\left[\min_{i\in[1:m']} |a_i X_{k+1} + b_i| > \sqrt{\epsilon}\right] \\
&\leq \Pr\left[\min_{i\in[1:m']} |a_i X_{k+1} + b_i| \leq \sqrt{\epsilon}\right] + \iint_A \Pr\left[|p(X_1,...,X_k, re^{j\phi})| \leq \epsilon\right] f_{|X|,\angle X}(r,\phi) d\phi dr
\end{aligned} \tag{F.11}$$

$$\begin{aligned}
&\leq \sum_{i=1}^{m'} \Pr[|a_i X_{k+1} + b_i| \leq \sqrt{\epsilon}] \\
&\quad + \iint_A \Pr\left[\left|\sum_{i=1}^{m'} \frac{a_i re^{j\phi} + b_i}{\sqrt{\epsilon}} \prod_{j=1}^{k} X_j^{d_{ji}} + \sum_{i=m'+1}^{m} \frac{a_i}{\sqrt{\epsilon}} \prod_{j=1}^{k} X_j^{d_{ji}}\right| \leq \sqrt{\epsilon}\right] f_{|X|,\angle X}(r,\phi) d\phi dr
\end{aligned} \tag{F.12}$$

$$\leq 2^k (2 f_{max} \sqrt{\epsilon}) + \iint_A \left(2^{k+1} f_{max} \sqrt[2^{k-1}]{\sqrt{\epsilon}}\right) f_{|X|,\angle X}(r,\phi) d\phi dr \tag{F.13}$$

$$\begin{aligned}
&\leq 2^{k+1} f_{max} \sqrt[2^k]{\epsilon} + 2^{k+1} f_{max} \sqrt[2^k]{\epsilon} \\
&= 2^{k+2} f_{max} \sqrt[2^k]{\epsilon},
\end{aligned}$$

where in (F.11-F.13), the integration is over $A = \{(r,\phi) : \min_{i\in[1:m']} |a_i re^{j\phi} + b_i| > \sqrt{\epsilon}\}$, and in (F.12), we have used the union bound. Also, (F.13) is true because of the upper bound in (F.9), the fact that $m' \leq 2^k$, and also because in (F.12), we have $\left|\frac{a_i re^{j\phi} + b_i}{\sqrt{\epsilon}}\right| >$

$1, \forall i \in [1 : m']$ and $\left|\frac{a_i}{\sqrt{\epsilon}}\right| \geq |a_i| \geq 1, \forall i \in [m'+1 : m]$, which enables us to use the inductive assumption by noting that $\sum_{i=1}^{m'} \frac{a_i r e^{j\phi} + b_i}{\sqrt{\epsilon}} \prod_{j=1}^{k} X_j^{d_{ji}} + \sum_{i=m'+1}^{m} \frac{a_i}{\sqrt{\epsilon}} \prod_{j=1}^{k} X_j^{d_{ji}}$ is a polynomial in $X_1, ..., X_k$, satisfying the conditions in the lemma. This completes the proof. $\square$

BIBLIOGRAPHY

[1] "5G Radio Access - Research and Vision," Ericsson white paper, June 2013.

[2] X. Lin, J. Andrews, and A. Ghosh, "A Comprehensive Framework for Device-to-Device Communications in Cellular Networks," e-print arXiv:1305.4219.

[3] M. Ji, G. Caire, and A. F. Molisch, "Wireless Device-to-Device Caching Networks: Basic Principles and System Performance," e-print arXiv:1305.5216.

[4] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Communications*, vol. 19, no. 3, pp. 96-104, June 2012.

[5] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-Device Communication as an Underlay to LTE-Advanced Networks, *IEEE Communications Magazine*, vol. 47, no. 12, pp. 42-49, Dec. 2009.

[6] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design Aspects of Network Assisted Device-to-Device Communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170-177, Mar. 2012.

[7] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient Resource Allocation for Device-to-Device Communication Underlaying LTE Network," in proceedings of *IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 368-375, Oct. 2010.

[8] Q. Duong and O. Shin, "Distance-Based Interference Coordination for Device-to-Device Communications in Cellular Networks," in proceedings of *Fifth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 776-779, July 2013.

[9] V. R. Cadambe and S. A. Jafar, "Interference Alignment and Degrees of Freedom of the K-User Interference Channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425-3441, Aug. 2008.

[10] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication Over MIMO X Channels: Interference Alignment, Decomposition, and Performance Analysis," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3457-3470, Aug. 2008.

[11] L. Tassiulas and A. Ephremides, "Jointly Optimal Routing and Scheduling in Packet Radio Networks," *IEEE Transactions on Information Theory*, vol. 38, no. 1, pp. 165-168, Jan. 1992.

[12] P. Chaporkar, K. Kar, and S. Sarkar, "Throughput Guarantees Through Maximal Scheduling in Wireless Networks," in proceedings of *43rd Annual Allerton Conference on Communications, Control, and Computing*, 2005.

[13] X. Lin, and N. B. Shroff, "The Impact of Imperfect Scheduling on Cross-Layer Rate Control in Multihop Wireless Networks," in proceedings of *24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2005.

[14] G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Maximum Weighted Matching with Interference Constraints," in proceedings of *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, (PerCom)*, 2006.

[15] G. Sharma, R. R. Mazumdar, and N. B. Shroff, "On the Complexity of Scheduling in Wireless Networks," in proceedings of *The Twelfth Annual ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2006.

[16] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388-404, Mar. 2000.

[17] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "FlashLinQ: A Synchronous Distributed Scheduler for Peer-to-Peer Ad Hoc Networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1215-1228, Aug. 2013.

[18] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. Jafar, "On the Optimality of Treating Interference as Noise," e-print arXiv:1305.4610.

[19] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. Jafar, "On the Optimality of Treating Interference as Noise," in proceedings of *51st Annual Allerton Conference on Communications, Control, and Computing*, 2013.

[20] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534-5562, Dec. 2008.

[21] L. Zhou, and W. Yu, "On the capacity of the K-user cyclic Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 154-165, Jan. 2013.

[22] J. Zander and M. Frodigh, "Comment on " Performance of optimum transmitter power control in cellular radio systems"", *IEEE Transactions on Vehicular Technology*, vol. 43, no. 3, Aug. 1994.

[23] A. Schrijver, *Combinatorial Optimization*, Springer, 2003.

[24] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 7, pp. 1205-1211, July 1999.

[25] N. Naderializadeh and A. S. Avestimehr, "ITLinQ: A New Approach for Spectrum Sharing in Device-to-Device Communication Systems," e-print arXiv:1311.5527.

[26] C. Mcdiarmid and T. Müller, "On the chromatic number of random geometric graphs," *Combinatorica*, vol. 31, no. 4, pp. 423-488, Nov. 2011.

[27] H. B. Mann and A. Wald, "On the Statistical Treatment of Linear Stochastic Difference Equations," *Econometrica* , vol. 11, no. 3/4, pp. 173-220, 1943.

[28] N. Lee, X. Lin, J. G. Andrews, and R. W. Heath Jr, "Power Control for D2D Underlaid Cellular Networks: Modeling, Algorithms and Analysis," e-print arXiv:1305.6161.

[29] N. Naderializadeh and A. S. Avestimehr, "Interference Networks with No CSIT: Impact of Topology," e-print arXiv:1302.0296.

[30] N. Naderializadeh and A. S. Avestimehr, "Impact of Topology on Interference Networks with No CSIT," in Proceedings of *IEEE International Symposium on Information Theory*, Istanbul, Turkey, 2013.

[31] S. A. Jafar, "Topological Interference Management through Index Coding," In *arXiv:1301.3106*.

[32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[33] E. R. Scheinerman and D. H. Ullman, *Fractional Graph Theory: A Rational Approach to the Theory of Graphs*, John Wiley & Sons, 2008.

[34] S. A. Jafar, "Elements of Cellular Blind Interference Alignment — Aligned Frequency Reuse, Wireless Index Coding and Interference Diversity," In *arXiv:1203.2384*.

[35] J. T. Schwartz, "Fast Probabilistic Algorithms for Verification of Polynomial Identities," *Journal of the ACM*, 27(4):701-717, 1980.

[36] R. Zippel, "Probabilistic Algorithms for Sparse Polynomials," in *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, pp. 216-226, 1979.

[37] A. Cuyt, K. Driver and D. S. Lubinsky, "On the Size of Lemniscates of Polynomials in One and Several Variables," in *Proceedings of the American Mathematical Society*, vol. 124, no. 7, pp. 2123-2136, July 1996.

[38] D. S. Lubinsky, "Small Values of Polynomials: Cartan, Pólya and Others," *Journal of Inequalities and Applications*, vol. 1, pp. 199-222, 1997.