

# Throughput Analysis of Large Networks: Spatial Diversity, Beamforming Gain, and Transmission Modes

THÈSE N° 7639 (2017)

PRÉSENTÉE LE 27 MARS 2017

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS  
LABORATOIRE DE THÉORIE DE L'INFORMATION  
PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

**Serj HADDAD**

acceptée sur proposition du jury:

Dr N. Macris, président du jury  
Prof. E. Telatar, Dr O. Lévêque, directeurs de thèse  
Prof. R. Couillet, rapporteur  
Prof. A. Ozgur, rapporteuse  
Prof. M. Kapralov, rapporteur



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2017



# **Throughput Analysis of Large Networks: Spatial Diversity, Beamforming Gain, and Transmission Modes**

Serj Haddad

March 17, 2017

Thesis presented to the faculty of computer and communication sciences for  
obtaining the degree of Docteur ès Sciences

Committee members:

**Dr. N. Macris**  
President of the jury

**Prof. E. Telatar** and **Dr. O. Lévêque**  
Thesis directors

**Prof. R. Couillet**  
Expert

**Prof. M. Kapralov**  
Expert

**Prof. A. Özgür**  
Expert

Ecole Polytechnique Fédérale de Lausanne, 2017

## Acknowledgments

---

It has been a great pleasure and privilege to work under the direction of Prof. Emre Telatar and Dr. Olivier Lévêque who welcomed me in the Information Theory Laboratory.

I would like to begin the acknowledgement section by expressing my deepest gratitude to my main supervisor Dr. Olivier Lévêque for his patience and guidance during my PhD. Every time I had an idea no matter how vague it was, he always had the patience to listen and discuss with me the mathematical details I overlooked, which turned out to be the bulk of my thesis. He always supported any plan that could expand my research horizons, be it a research internship abroad or attending conferences and schools on information theory. Many thanks.

I am also truly grateful to Prof. Emre Telatar for all the interesting discussions in diverse research topics. I have learned a lot from his wisdom, modesty, and kindness. Although he was not always directly involved in the details of my research work, his advice and ideas were always interesting and highly influential.

I was very fortunate to have the opportunity to collaborate with Prof. Ayfer Özgür who welcomed me in the Information Systems Laboratory at Stanford University for an internship. Despite her tight schedule, I highly appreciate her effort to make sure my visit had a fruitful outcome. Thanks to her, I was exposed to a new research topic that enriched and diversified the current dissertation.

I would like to further thank the jury members of my thesis, Prof. Ayfer Özgür, Prof. Romain Couillet, Prof. Michael Kapralov, and the president of the jury Dr. Nicolas Macris, for their careful reading of my thesis and their valuable comments.

The information Processing Group (IPG) has been a great working environment. I would like to thank Françoise, Muriel and Damir who made sure that everything is running flawlessly in the lab. Furthermore, the lab would not have been such an amazing place were it not for the great friends and colleagues I encountered there who made my stay in Switzerland very pleasant and enjoyable. I extend my thanks to all my Lebanese friends who created a New Lebanon in Lausanne and made it feel like home.

It is difficult and strange to express my appreciation to my family, as my brother and I were taught by our parents that there is no such thing as “to apologize” or “to thank” between the four of us. We are a family. I would like to break this rule and thank my brother Jack for all his support and encouragement at all times. To my parents, there are no words that can express my gratitude. I cannot thank you enough for all your unconditional love, sacrifice, and support. As a hopeless attempt,

*“I dedicate this dissertation to you,  
my father Hratch  
and my mother Hasmig.”*

My foremost recognition and heartfelt gratitude go to my wife, Sahar, for she was with me and by my side at all times. This dissertation belongs to both of us, for it is the result of an equal, not to say more, effort and sacrifice from your side. It is a small accomplishment of ours compared to the family we are building together, especially that we are impatiently awaiting our daughter, Sofia. An infinite love and gratitude to you, Sahar.

## Abstract

---

While wired infrastructure constitutes the backbone of most wireless networks, wireless systems appeal the most to the dynamic and rapidly evolving requirements of today's communication systems because of their ease of deployment and mobility, not to mention the high cost of building a wired infrastructure. This led to an increased interest in the so called wireless *ad hoc* networks formed of a group of users, known as nodes, capable of communicating with each other through a shared wireless channel. Needless to say, these nodes are asked to use the shared wireless medium in the most efficient fashion, which is not an easy task given the absence of wired backbone. This requires a profound understanding of the wireless medium to establish a decentralized cooperation scheme, if needed, that best utilizes the resources available in the wireless channel. A significant part of this thesis focuses on the properties of the shared wireless channel, whereby we are interested in studying the *spatial diversity* and the *beamforming* capabilities in large wireless networks which are crucial in analyzing the throughput of ad hoc networks.

In this thesis, we mainly focus on the problem of broadcasting information in the most efficient manner in a large two-dimensional ad hoc wireless network at low SNR and under line-of-sight propagation. A new communication scheme, which we call multi-stage back-and-forth beamforming, is proposed, where source nodes first broadcast their data to the entire network, despite the lack of sufficient available power. The signal's power is then reinforced via successive back-and-forth beamforming transmissions between different groups of nodes in the network, so that all nodes are able to decode the transmitted information at the end. This scheme is shown to achieve asymptotically the broadcast capacity of the network, which is expressed in terms of the largest singular value of the matrix of fading coefficients between the nodes in the network. A detailed mathematical analysis is then presented to evaluate the asymptotic behavior of this largest singular value. We further characterize the maximum achievable broadcast rate under different sparsity regimes. Our result shows that this rate depends negatively on the sparsity of the network. This is to be put in contrast with the number of degrees of freedom available in the network, which have been shown previously to increase with the sparsity of the network. In this context, we further characterize the degrees of freedom

versus beamforming gain tradeoff, which reveals that high beamforming gains can only be obtained at the expense of reduced spatial degrees of freedom.

Another important factor that impacts the throughput in wireless networks is the transmit/receive capability of the transceiver at the nodes. Traditionally, wireless radios are half-duplex, i.e. they can not transmit and receive at the same time over the same frequency band. However, building on self-interference cancellation techniques, full-duplex radios have emerged as a viable paradigm over the recent years. In the last part of this thesis, we ask the fundamental question: how much can full-duplex help? Intuitively, one may expect that full-duplex radios can at most double the capacity of wireless networks, since they enable nodes to transmit and receive at the same time. However, we show that the capacity gain can indeed be larger than a factor of 2; in particular, we construct a specific instance of a wireless network where the full-duplex capacity is triple the capacity of the half-duplex network. We also propose a universal schedule for half-duplex networks of independent, memoryless, point-to-point channel capable of ensuring a fraction of  $1/4$  of the full-duplex capacity. This means that, for point-to-point networks full-duplex cannot more than quadruple the capacity of wireless networks.

**Keywords:** Wireless ad hoc networks, scaling laws, broadcast capacity, low SNR communications, beamforming strategies, sparsity, degrees of freedom versus beamforming gain tradeoff, random matrix theory, full-duplex relaying, half-duplex relaying, point-to-point channels, scheduling



## Résumé

---

Tandis qu'une infrastructure câblée est à la base de la plupart des réseaux sans fil actuels, l'évolution dynamique et rapide des nouveaux réseaux sans fil appelle à de nouvelles infrastructures elles-mêmes sans fil pour faciliter le déploiement de ces réseaux et mieux gérer la mobilité des utilisateurs, ainsi que pour minimiser les coûts d'installation de telles infrastructures. Ce fait a considérablement renforcé l'intérêt pour l'étude de réseaux sans fil dits *ad hoc*, consistant en des groupes d'utilisateurs, appelés noeuds, capables de communiquer à travers un canal sans fil partagé. Ces noeuds ont pour but de communiquer de manière optimale à travers ce canal commun, ce qui n'est pas une tâche aisée, étant donné l'absence d'infrastructure câblée pour relayer les communications. Ceci requiert une pleine compréhension du milieu sans fil pour établir des schémas de coopération décentralisés qui utilisent au mieux les ressources présentes. Une partie significative de la présente thèse s'intéresse aux propriétés de ce milieu sans fil, en particulier à la *diversité spatiale* ainsi qu'au *gain de puissance par formation de faisceaux* qu'il est possible d'atteindre dans un réseau sans fil de grande taille. Ces deux grandeurs sont en effet cruciales pour déterminer la capacité de tels systèmes.

Dans cette thèse, nous nous concentrons sur le problème de la diffusion la plus efficace d'information dans un réseau ad hoc sans fil de grande taille, lorsque le rapport signal sur bruit (SNR) est bas. Le modèle de propagation étudié est celui de la propagation des ondes électromagnétiques dans le vide. Un nouveau schéma de communication est proposé, que nous appelons schéma d'augmentation de gain de puissance à plusieurs étapes, dans lequel les noeuds transmetteurs diffusent d'abord leur données à tout le réseau, malgré le manque de puissance disponible. La puissance du signal émis est alors renforcée par des transmissions successives d'une partie du réseau à l'autre, de telle manière à ce que tous les noeuds puissent décoder l'information au bout du processus. Il est ensuite montré que ce schéma de communication atteint asymptotiquement la capacité de diffusion maximale du réseau, que l'on peut exprimer mathématiquement en termes de la plus grande valeur singulière de la matrice des coefficients d'atténuation entre les noeuds du réseau. Une analyse mathématique détaillée est ensuite présentée pour évaluer le comportement asymptotique de cette plus grande valeur singulière.

Nous caractérisons de plus le taux maximum de diffusion pour des réseaux de densité variable. Nos résultats montrent que ce taux augmente avec la densité du réseau. Ce résultat est à mettre en perspective avec les résultats obtenus précédemment sur le nombre de degrés de liberté du réseau, qui eux décroissent avec la densité du réseau. Dans ce contexte, nous caractérisons le compromis qui existe entre la diversité spatiale du milieu sans fil et le gain de puissance atteignable dans un tel milieu, et montrons que de forts gains de puissance ne peuvent être obtenus qu'au prix d'une réduction du nombre de degrés de liberté.

Un autre facteur important qui impacte la capacité de réseaux sans fil est lié aux capacités émettrices-réceptrices des noeuds. Traditionnellement, les radios sans fil travaillent en mode semi-duplex, i.e., elles ne peuvent émettre et recevoir en même temps sur la même bande de fréquence. Toutefois, en utilisant des techniques d'annulation d'auto-interférence, des radios travaillant en mode duplex intégral ont vu le jour ces dernières années. Dans la dernière partie de cette thèse, nous posons cette question fondamentale: de quel facteur le mode duplex intégral peut-il augmenter la capacité d'un réseau sans fil? De manière intuitive, on pourrait s'attendre à ce que le mode duplex intégral permette au mieux de doubler la capacité d'un réseau, du fait que ce mode permet une transmission et une réception simultanées. Nous montrons cependant que ce gain de capacité peut être plus grand qu'un facteur 2; en particulier, nous construisons une instance particulière de réseau dont la capacité en mode duplex intégral est le triple de celle en mode semi-duplex. Nous proposons aussi un schéma universel de communication pour les réseaux en mode semi-duplex, avec une unique source et une unique destination, capable d'atteindre un facteur  $1/4$  de la capacité en mode duplex intégral de même réseau. Ceci signifie pour ces réseaux que le mode duplex intégral ne peut au mieux que quadrupler la capacité du réseau.

**Mots-clés:** réseaux sans fil ad hoc, lois d'échelle, capacité de diffusion, communications à bas SNR, stratégies de gain de puissance, compromis degrés de liberté - gain de puissance, matrices aléatoires, mode duplex intégral, mode semi-duplex

# Contents

---

Acknowledgments	i
Abstract	iii
Résumé	v
Contents	vii
<b>1 Introduction</b>	<b>1</b>
1.1 Throughput Scaling in Ad Hoc Networks . . . . .	3
1.1.1 Broadcasting in Ad Hoc Networks . . . . .	4
1.1.2 Physical Limits in Wireless Networks . . . . .	5
1.2 Full-Duplex vs Half-Duplex . . . . .	7
<b>2 Broadcast Capacity of Wireless Networks</b>	<b>9</b>
2.1 Model . . . . .	9
2.2 Back-and-Forth Beamforming Strategy . . . . .	11
2.3 Optimality of the Scheme . . . . .	17
2.4 Summary . . . . .	22
2.A Appendix . . . . .	23
<b>3 Communication Tradeoffs in Wireless Networks</b>	<b>41</b>
3.1 Model . . . . .	42
3.2 Main Result . . . . .	43
3.3 Broadcast Capacity . . . . .	45
3.3.1 At High SNR . . . . .	45
3.3.2 At Low SNR . . . . .	45
3.4 Multiple-Unicast Capacity . . . . .	56
3.4.1 At High SNR . . . . .	56
3.4.2 At Low SNR . . . . .	59
3.5 Communication Tradeoffs . . . . .	62
3.A Appendix . . . . .	63
<b>4 Full-Duplex vs Half-Duplex</b>	<b>73</b>

---

4.1	Problem Formulation and Background . . . . .	74
4.2	Full-Duplex can Triple the Capacity of Wireless Networks . . .	75
4.3	Full-Duplex Cannot More Than Quadruple . . . . .	79
4.3.1	Universal Deterministic Schedule . . . . .	80
4.4	Summary . . . . .	82
<b>5</b>	<b>Conclusions and Future Directions</b>	<b>85</b>
5.1	Impact of Communication Medium on Network Throughput . .	85
5.1.1	Future Research Directions . . . . .	86
5.2	Impact of Transmission Mode on Network Throughput . . . . .	86
5.2.1	Future Research Directions . . . . .	86
	<b>Bibliography</b>	<b>89</b>
	<b>Curriculum Vitae</b>	<b>95</b>

---

## Introduction

---

# 1

We live in a world where wireless communication networks are in the heart of all kinds of industries and businesses, not to mention their fundamental role and impact in our everyday lives. This fact prompted huge and profound research work in *wireless communication theory* to better understand the fundamentals of wireless technology. In fact, a better understanding of the properties of this technology improves cellular networks and wireless LANs (local area networks), which are examples of wireless networks, in terms of speed of communication, quality of service, and design of such systems.

While most wireless networks rely on wired infrastructure leaving the wireless system as the last stage of communication, researchers put wireless systems to the forefront of communication networks. The ease of deploying wireless systems and their mobility, to name but a few, highly appeal to the requirements of communication networks. On the other hand, building wired infrastructure to serve as a high capacity backbone is difficult, expensive, and lacks the flexibility required by the dynamic and the highly evolving network systems. This discussion leads us to the so called wireless *ad hoc* networks. These networks rely merely on wireless communication without any wired infrastructure. An ad hoc network is formed of a group of users, known as nodes, capable of communicating with each other through wireless transmission. Each node has a certain power to transmit information to the rest of the nodes and is equivalently capable of receiving information from other nodes. This communication is carried out through a shared wireless medium without any infrastructure assistance.

The absence of a wired backbone connecting base stations to assist the nodes in ad hoc networks throws the burden of traffic management throughout the network on the nodes themselves. In other words, the nodes present in the network must organize themselves in a decentralized fashion and cooperate

in an efficient way to best utilize the resources available in the shared wireless medium. This requires a thorough understanding of the properties of this channel on one hand and the transmit-receive capabilities of the nodes themselves on the other hand. Together they constitute the core of this thesis.

Concerning the properties of the shared wireless channel, we are interested in studying the *spatial diversity* and the *beamforming* capabilities in large wireless networks which play a central role in analyzing the throughput of ad hoc networks. Furthermore, regarding the *transmission mode* of each user, our goal is to identify the gain, in terms of the throughput across wireless networks, of having *full-duplex* nodes that can transmit and receive at the same time instead of *half-duplex* nodes that cannot transmit and receive simultaneously.

While spatial diversity is determined by the spatial *degrees of freedom* of the channel matrix  $H$ , the beamforming gain is dictated by the channel matrix *spectral norm*  $\|H\|$ , i.e. the *largest singular value* of the channel matrix  $H$ . Consequently, a significant part of this thesis involves *random matrix theory* required to provide a rigorous mathematical analysis of unconventional random matrices, as we will see. Contrary to both spatial degrees of freedom and beamforming gain, the transmission mode is not an innate property of the channel model considered in the network. For this reason, in this thesis, we analyze the impact of the user transmission mode on the throughput across the network under simplistic point-to-point wireless transmission model, which paves the road to broader analysis with a less restrictive assumption about the channel model. This approach diverts the focus from the channel model to the duplexing ability of the nodes, which reduces the problem to a combinatorial optimization problem. Our goal is to find the maximum throughput gain one gets from having full-duplex nodes instead of half-duplex nodes, which involves discovering optimal schedules for the half-duplex scenario.

So we disclosed the motive behind studying the impact of the transmission mode of the nodes in wireless networks, which will be handled in Chapter 4. On the other hand, the analysis of spatial degrees of freedom and spectral norm of the channel matrix aims at characterizing the optimal communication schemes in large ad hoc networks. While the degrees of freedom of the channel matrix proved to have a significant impact on the multiple-unicast capacity of large wireless networks at high signal-to-noise ratio (SNR) regimes [45, 46], its spectral norm turns out to be the other main factor affecting both the broadcast and the multi-unicast capacity of large wireless networks at low-SNR regime. The current dissertation concentrates on the broadcast capacity of wireless ad hoc networks at low SNR, which was the driving force behind characterizing the beamforming gain in wireless networks. In Chapter 2, we fully characterize the broadcast capacity of *extended*<sup>1</sup> wireless networks. Generalizing this result to wireless networks with different sparsity, in terms of how the area of the network relates to the number of users, will be the focus of Chapter 3, which, as we will see, boils down to studying the tradeoff between the spatial diversity and the beamforming gain.

---

<sup>1</sup>The number of the users in the network is equal to the area of the network.

We adopt standard notation from complexity theory;  $O(\cdot)$ ,  $\Omega(\cdot)$ , and  $\Theta(\cdot)$  describe asymptotic upper, lower, and tight bounds, respectively, throughout the current dissertation.

## 1.1 Throughput Scaling in Ad Hoc Networks

In the past decades, a lot of effort was put into understanding the behavior of ad hoc networks, which requires a vivid understanding of the fundamental capacity limits of large wireless networks. Beside the virtue they have from the theoretical point of view, results on network capacity provide insight into sophisticated cooperation strategies and protocol design in wireless networks, which may alter the state-of-the-art physical-layer wireless systems that are mostly restricted to multihop protocol relying merely on point-to-point communication. While the latter simply mimics the operation of wired network and replaces the wire between two nodes by a wireless link of certain capacity, sophisticated strategies involve cooperation and coordination among multiple users in the network. Information-theoretic scaling laws, independent of the communication strategy, manifest either the optimality of simple multihop scheme or the urge to discover novel communication schemes that can attain higher throughput scaling.

The seminal work [22] by Gupta and Kumar, introduced a simple yet insightful model for wireless ad hoc networks with a large number of users. Although their model is restricted to multihop transmission with pairwise coding and decoding, it explicitly takes into account the significant properties of wireless networks; namely, the spatial distribution of nodes, the nature of the wireless medium, the signal attenuation with distance, and the traffic requirement between the users whereby all the nodes in the network are required to transmit at a common rate. The ultimate goal is to figure out the scaling limit of the achievable rate as the number of nodes in the network grows. Following this work, numerous research work was conducted to obtain scaling laws independent of the communication protocol, which led to many interesting yet sophisticated cooperative strategies such as network coding [7, 15, 16, 32], hierarchical cooperation [20, 45, 46], and space-time coding [10, 50, 53].

A shallow literature review reveals diverse research directions pertaining to the throughput analysis of large wireless networks. For instance, the authors in [27] study the broadcast capacity of multihop wireless network as the maximum rate at which a set of nodes generate broadcast packets in the network whereby all nodes are required to receive the packets successfully. Another related research direction is to consider mobile nodes, contrary to static networks where the location of the nodes are fixed and modeled according to some process, for instance Poisson point process as in [56]. Mobility-based routing method [14, 21] proves to increase the per node capacity in condition that we allow packet delay to grow arbitrarily large. This led to many interesting results on throughput scaling under different constraints on the mobility of the nodes or the delay to further characterize the throughput-delay tradeoff for mobile

networks [3, 18, 34, 40, 47]. Beside the constraints on the routing protocol or the mobility of the users in the network, other limiting factors on the broadcast capacity can be the power available at the nodes, prompting the study of broadcast capacity of power-constrained networks [26, 56].

### 1.1.1 Broadcasting in Ad Hoc Networks

The literature on the study of scaling laws in large *ad hoc* wireless networks concentrates mainly on *multiple-unicast* (one-to-one) transmissions (see e.g. [22, 43, 45]). This does not degrade by any means the importance of investigating *multicast* (one-to-many) transmissions for several reasons such as the need of many network protocols to broadcast control signals carrying channel state information or to enhance cooperation among nodes belonging to the same cluster or cell. Broadcasting control signals to the entire network is indispensable to all routing protocols for they assist not only in route discovery, but also in network monitoring and maintenance [24, 27, 51]. In other words, broadcast communication is an integral part of many multiple-unicast routing protocols, such as Dynamic Source Routing (DSR), Ad Hoc On Demand Distance Vector (AODV), Zone Routing Protocol (ZRP), and Location Aided Routing (LAR) [24]. Thus, despite the fact that multiple-unicast transmission is the main service targeted in ad hoc networks, the throughput analysis of ad hoc networks remains incomplete if we ignore the limitations imposed by network-wide broadcast communication. For this reason, we dedicate Chapters 2 and 3 to address an a priori much easier scenario than the multiple-unicast scenario. Instead of every source node willing to communicate each to a different destination node, we consider the *broadcast scenario*, where each source node wishes to send some piece of information to all the other nodes in the network.

As in the case of multiple-unicast ad hoc networks, for broadcast ad hoc networks, the ultimate goal resides in establishing information-theoretic scaling laws independent of the communication strategy. Equivalently, we are interested in studying how source nodes can broadcast their data to the whole network in the most efficient way. Previous works investigated the broadcast capacity of wireless networks under specific channel models and mainly at high SNR [25, 27, 33, 49, 51]. Of course, multiple strategies exist in this context, but from the scaling law point of view (that is, for large networks), the simplest communication strategy, where source nodes take turns broadcasting their messages to the entire network, can be shown to be asymptotically optimal (up to logarithmic factors), when the power path loss is that of free space propagation. For a stronger power path loss, still at high SNR, simple multihopping strategies also allow to achieve an asymptotically optimal broadcast capacity, so there is not much to be discussed either in this case from the scaling law point of view.

We address the low-SNR regime and consider the line-of-sight (LOS) propagation model [36, 52] described in Sections 2.1 and 3.1. In this regime, the power available does not allow for a source node to successfully transmit a



message to its nearest neighbor without waiting for some amount of time in order to spare power.

This raises an interesting question, as a naive cut-set argument upper-bounding the broadcast capacity seems to indicate that potentially much higher rates could be achievable in this case. As a consequence, none of the two strategies described above (time-division or multihop broadcasting) is asymptotically optimal.

This issue was first revealed in [36] in the context of one-dimensional networks, under the line-of-sight model. For such networks, the authors proposed a hierarchical beamforming scheme to broadcast data to the network, which was proven to achieve asymptotic optimal performance.

The key idea behind the scheme is that in a one-dimensional network and under line-of-sight fading, a group of nodes sharing some common information and using a proper precoding scheme, can beamform this information simultaneously to all the other nodes in the network, which allows to compensate for the lack of available transmit power. This idea can be used recursively to beamform information to larger and larger groups of nodes, reaching an optimal beamforming gain at the final stage.

The generalization of this idea to two-dimensional networks is not immediate. Indeed, a particular feature of one-dimensional networks is that it is always possible for a group of nodes to beamform a given signal to *all* the other nodes in the network *simultaneously*. In two dimensions, a full beamforming gain is only achievable between groups of nodes that are sufficiently far apart from each other. This was already observed in [37], where a strategy was developed to enhance multiple-unicast communications in wireless networks under the line-of-sight model. Taking inspiration from this paper, we propose a new multi-stage back-and-forth beamforming scheme which is shown to achieve asymptotically optimal<sup>2</sup> performance for broadcasting information in a two-dimensional wireless network.

### 1.1.2 Physical Limits in Wireless Networks

Chapters 2 and 3, as already mentioned, focus on the problem of broadcasting information in the most efficient manner in a large two-dimensional *ad hoc* wireless network at low SNR. We propose a novel broadcast communication scheme, where source nodes first broadcast their data to the entire network, despite the lack of sufficient available power. The signal's power is then reinforced via successive back-and-forth beamforming transmissions between different groups of nodes in the network, so that all nodes are able to decode the transmitted information at the end. This scheme is shown to achieve asymptotically the broadcast capacity of the network, which is expressed in terms of the largest singular value of the matrix of fading coefficients between the nodes in the

---

<sup>2</sup>Technically speaking, the performance of our scheme is shown below to be asymptotically optimal up to a multiplicative factor  $n^\epsilon$ , where  $\epsilon$  can be taken arbitrarily small, but fixed.

network. A detailed mathematical analysis is then presented to evaluate the asymptotic behavior of this largest singular value.

We characterize the maximum achievable broadcast rate in a wireless network at low SNR and under line-of-sight fading assumption. Our result shows that this rate depends negatively on the sparsity of the network. This is to be put in contrast with the number of degrees of freedom available in the network, which have been shown previously to increase with the sparsity of the network.

Because of the inherent broadcast nature of wireless signals, managing the *interference* between the multiple source-destination pairs is a key issue and has led to various interesting proposals [1, 2, 6, 22, 38, 39, 41, 45]. In some of these works, it appeared that the model considered for the fading environment may substantially impact the performance of the proposed communication schemes (see [17]). In particular, the *channel diversity*, both spatial and temporal, turns out to be a key parameter for the analysis of the various schemes.

In this context, the broadcast nature of the wireless medium can only help relaying communications, so that the situation seems simpler to handle, if not trivial. What we show in Chapter 3 is that even in this simpler scenario, the optimal communication performance highly depends on the nature of the wireless medium. The conclusions we draw put again channel diversity to the forefront. But whereas diversity was beneficial for establishing multiple parallel communication channels in the multiple-unicast scenario, it turns out that in the present case, diversity is on the contrary detrimental to a proper broadcasting of information. A *duality* is further established between the number of *degrees of freedom* available for multi-party communications and the *beamforming gain* of broadcast transmissions, which allows for a better dissemination of information. At one end, in a rich scattering environment, degrees of freedom are prominent, while beamforming is practically infeasible. At the other end, degrees of freedom become a scarce resource, while high beamforming gains can be achieved via collaborative transmissions.

This discussion brings us back to the limitations physics of wave propagation impose on the throughput scaling in large ad hoc networks [17, 46]. As much as it is important to have throughput scaling laws independent of any communication strategy, which was first studied in [54], it is crucial to avoid assumptions on the fading model, such as the path loss gain or the phase shifts introduced to the transmitted signals, that defy physics laws of electromagnetic propagation leading to non-realistic bounds on the throughput scaling. While the authors in [17] and [46] are interested in the spatial degrees of freedom dictated by Maxwell's physics of wave propagation, in this dissertation we are interested in the beamforming capabilities allowed by the very same Maxwell's equations.

For the multiple-unicast scenario at high SNR, having uncorrelated fading coefficients increases the throughput across the network. In other words, ideally we want the channel coefficients between the nodes to be independent in order to achieve a linear scaling (with respect to the number of nodes) multi-unicast throughput. The linear scaling can be achieved when the nodes form distributed MIMO arrays through hierarchical cooperation architecture [45].

[17] shows that this assumption is only valid if the nodes are far apart and the density of the network is less than the inverse of the number of nodes in the network. Otherwise, Maxwell's propagation model establishes correlation among the fading coefficients, which increases with the increase in the density of the network above this threshold ( $1/n$ , where  $n$  is the number of nodes). Equivalently, the spatial degrees of freedom increase as the area of the network increases<sup>3</sup>. This realization shuffled all the cards concerning optimal multi-unicast strategies available and led to sparsity oriented approach whereby networks of different sparsity may have different optimal communication strategies. This problem was put to rest in [46], where the authors show that depending on how the area of the network<sup>4</sup> relates to the number of nodes (i.e. the sparsity of the network), the spatial diversity is dictated. As a consequence, the authors prove that in the multi-unicast networks high-SNR regime, while multihopping is optimal in extended networks<sup>5</sup>, sophisticated cooperation schemes, such as hierarchical cooperation or MIMO-multihop schemes, are required in sparser networks to fully exploit the spatial degrees of freedom available. What about the broadcast scenario? How does sparsity impact the broadcast capacity? What about in the multi-unicast networks low-SNR regime?

In this thesis, we show that, contrary to the multi-unicast scenario at high SNR, correlation among fading coefficients comes in favor of broadcast capacity at low SNR. Consequently, the denser an ad hoc network becomes the more the nodes are required to cooperate in order to exploit the beamforming capabilities in the network. The duality between the spatial degrees of freedom and the beamforming gain is translated to the following relation: 1) in a rich scattering environment, while for the multi-unicast scenario at high SNR the users are asked to cooperate to exploit the spatial degrees of freedom available in the network, for the broadcast scenario at low SNR multihopping is optimal. 2) in a poor scattering environment, while for the multi-unicast scenario at high SNR multihopping is optimal, for the broadcast scenario at low SNR the users are required to cooperate to exploit the correlation among the fading coefficients. An interesting scenario, previously studied in [35], that may require the cooperation among the nodes available in the network in both rich and poor scattering environments is the multi-unicast scenario at low SNR, which is out of the scope of the current dissertation.

## 1.2 Full-Duplex vs Half-Duplex

Consider a wireless network where information is relayed with the help of many intermediate relay nodes. In common practice, these relays are half-duplex, i.e. they cannot transmit and receive at the same time over the same frequency

<sup>3</sup>Note that the users are randomly and uniformly distributed over a square area.

<sup>4</sup>For simplicity, we do not mention the carrier wavelength.

<sup>5</sup>Needless to say that multihopping is optimal not only in extended networks but in denser networks as well.

band. Over the recent years however, full-duplex relays have become increasingly viable. These full-duplex relays can transmit and receive at the same time and intuitively can be expected to double the capacity of wireless networks. In Chapter 4, we show that the capacity gain can indeed be larger than a factor of 2. We construct a specific instance of a wireless network and show that its full-duplex capacity (the capacity of the network when relays were full-duplex) is 3 times its half-duplex capacity (the capacity of the same network when relays are half-duplex). This is a single-source single-destination network consisting of a long chain of orthogonal channels.

A natural follow-up question is whether there are wireless networks where the ratio between the full-duplex and the half-duplex capacity can be even larger than 3, possibly increasing unboundedly with the size of the network. Answering this question requires one to either construct instances of wireless networks where the gain is even larger or prove universal upper bounds on the capacity gain with full-duplex operation. To this end, we show in this dissertation that for wireless networks composed of independent, memoryless, point-to-point channels the capacity gain provided by full-duplex operation cannot be larger than a factor of 4.

While various works in the literature have considered the approximation of the capacity of half-duplex networks [8, 44], the complexity of half-duplex scheduling [5, 9], and the computational complexity of computing capacity [4], we are not aware of any works that tackle the question we consider: how much can full-duplex operation help to increase the capacity of wireless networks? Chapter 4 formulates this question and provides an example which demonstrates that the correct answer is not the intuitive one. Our discussion is mostly restricted to a single-source single-destination network with orthogonal noisy (AWGN) channels, but the question we raise can be studied in much more generality for networks with multiple source-destination pairs and with channel models that incorporate wireless broadcast and superposition.

---

# Broadcast Capacity Scaling of Large Wireless Networks at Low SNR

---

# 2

In this chapter, we fully characterize the broadcast capacity of wireless networks with density 1; a network of area  $A = n$  contains  $n$  users. We address the low-SNR regime and consider the line-of-sight (LOS) propagation model ([36, 52]) described in Sections 2.1. In this regime, the power available does not allow for a source node to successfully transmit a message to its nearest neighbor without waiting for some amount of time in order to spare power.

We give a detailed description of the broadcast scheme in Section 2.2, as well as a proof of its optimality in Section 2.3. The proof of optimality is done in two steps. We first provide a general upper bound on the broadcast capacity of wireless networks (see Theorem 2.3.1), whose expression involves the matrix made of fading coefficients between the nodes in the network. We then proceed to characterize the broadcast capacity of two-dimensional wireless networks under LOS model, by obtaining an asymptotic upper bound on the largest singular value of the above mentioned matrix. This result is of interest in its own right, as such matrices have not been previously studied in the mathematical literature. In particular, there is much less randomness in such a matrix than in classically studied random matrices. We propose here a recursive method to upper bound its largest singular value.

## 2.1 Model

There are  $n$  nodes uniformly and independently distributed in a square of area  $A = n$ , so that the node density remains constant as  $n$  increases. Every node wants to broadcast a different message to the whole network, and all nodes want to communicate at a common *per user* data rate  $r_n$  bits/s/Hz. We denote by  $R_n = nr_n$  the resulting *aggregate* data rate and will often refer to it simply as “broadcast rate” in the sequel. The broadcast capacity of the network,

denoted as  $C_n$ , is defined as the maximum achievable aggregate data rate  $R_n$ . We assume that communication takes place over a flat channel with bandwidth  $W$  and that the signal  $Y_j[m]$  received by the  $j$ -th node at time  $m$  is given by

$$Y_j[m] = \sum_{k \in \mathcal{T}} h_{jk} X_k[m] + Z_j[m],$$

where  $\mathcal{T}$  is the set of transmitting nodes,  $X_k[m]$  is the signal sent at time  $m$  by node  $k$  and  $Z_j[m]$  is additive white circularly symmetric Gaussian noise (AWGN) of power spectral density  $N_0/2$  Watts/Hz. We also assume a common average power budget per node of  $P$  Watts, which implies that the signal  $X_k$  sent by node  $k$  is subject to an average power constraint  $\mathbb{E}(|X_k|^2) \leq P$ . In line-of-sight environment, the complex baseband-equivalent channel gain  $h_{jk}$  between transmit node  $k$  and receive node  $j$  is given by

$$h_{jk} = \sqrt{G} \frac{\exp(2\pi i r_{jk}/\lambda)}{r_{jk}}, \quad (2.1)$$

where  $G$  is Friis' constant,  $\lambda$  is the carrier wavelength, and  $r_{jk}$  is the distance between node  $k$  and node  $j$ . Let us finally define

$$\text{SNR}_s = \frac{GP}{N_0W},$$

which is the SNR available for a communication between two nodes at distance 1 in the network.

It should be noticed that the above line-of-sight model departs from the traditional assumption of i.i.d. phase shifts in wireless networks. The latter assumption is usually justified by the fact that inter-node distances are in practice much larger than the carrier wavelength, implying that the numbers  $2\pi r_{jk}/\lambda$  can be roughly considered as i.i.d. This approximation was however shown in [17] to be inaccurate in the setting considered in our work. A second remark is that no multipath fading is considered here, which would probably reduce in practice the efficiency of the strategy proposed in the following paragraph.

We focus in the following on the low-SNR regime, by which we mean, as in [36], that  $\text{SNR}_s = n^{-\gamma}$  for some constant  $\gamma > 0$ . This means that the power available at each node does not allow for a constant rate direct communication with a neighbor. It is important to note here that making the assumption that the SNR decays as a inverse power of the number of nodes is key to uncover the fact that plain time-division fails to be optimal at low SNR. This type of assumption was already made in previous contributions regarding the multiple unicast problem (see in particular [42, 48]), leading to similarly interesting conclusions.

In order to simplify notation, we choose new measurement units such that  $\lambda = 1$  and  $G/(N_0W) = 1$  in these units. This allows us to write in particular that  $\text{SNR}_s = P$ .

## 2.2 Back-and-Forth Beamforming Strategy

First note that under LOS model (2.1) and the assumptions made in the previous section, the time division scheme described in the introduction achieves a broadcast (aggregate) rate  $R_n$  of order  $\min(P, 1)$ . Indeed, a rate of order 1 is obviously achieved at high SNR<sup>1</sup>. At low SNR (i.e. when  $P \sim n^{-\gamma}$  for some  $\gamma > 0$ ), each node can spare power while the others are transmitting, so as to compensate for the path loss of order  $1/n$  between the source node and other nodes located at distance at most  $\sqrt{2n}$ , leading to a broadcast rate of order  $R_n \sim \log(1 + nP/n) \sim P$ . As we will see, this broadcast rate is not optimal at low SNR.

In the following, we propose a new broadcasting scheme that will prove to be order-optimal. In this new scheme, source nodes still take turns broadcasting their messages, but each transmission is followed by a series of network-wide back-and-forth transmissions that reinforce the strength of the signal, so that at the end, every node is able to decode the message sent from the source.

The reason why back-and-forth transmissions are useful here is that in line-of-sight environment, nodes are able to (partly) align the transmitted signals so as to create a significant beamforming gain for each transmission (whereas this would not be the case in high scattering environment with i.i.d. fading coefficients).

In what follows, we describe the scheme used to broadcast the message of a given source node to the entire network. In other words, the scheme described below is repeated  $n$  times to ensure the broadcast of the message of each and every node to the entire network.

### Scheme Description.

The scheme is split into two phases:

**Phase 1. Broadcast Transmission.** The source node broadcasts its message to the whole network. All the nodes receive a noisy version of the signal in this phase, which remains undecoded. This phase only requires one time slot.

**Phase 2. Back-and-Forth Beamforming with Time Division.** Let us first present here an idealized version of this second phase: upon receiving the signal from the broadcasting node, nodes start multiple back-and-forth beamforming transmissions between the two halves of the network, in order to enhance the strength of the signal. Although this simple scheme probably achieves the optimal performance claimed in Theorem 2.2.1 below, we lack the analytical tools to prove it. We therefore propose a time-division strategy, where clusters of size  $M = \frac{n^{1/4}}{2c_1} \times \frac{n^{1/2}}{4}$  and separated by horizontal distance  $d = \frac{n^{1/2}}{4}$  pair up for the back-and-forth transmissions, as illustrated on Fig. 2.1. During each transmission, there are  $\Theta(n^{1/4-\epsilon})$  cluster pairs operating in parallel (see (2.2)), so  $\Theta(n^{1-\epsilon})$  nodes are communicating in total. The number of rounds needed to serve all nodes must therefore be  $\Theta(n^\epsilon)$ .

<sup>1</sup>We coarsely approximate  $\log P$  by 1 here!

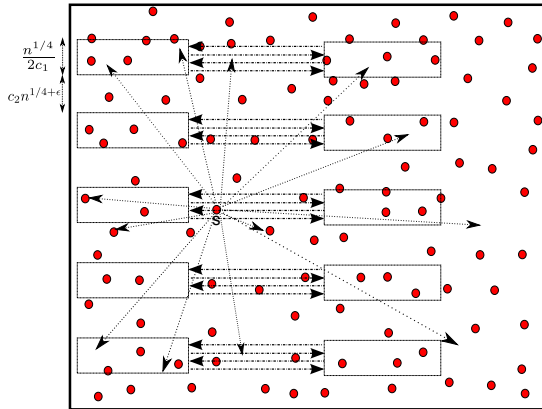


Figure 2.1 –  $\sqrt{n} \times \sqrt{n}$  network divided into clusters of size  $M = \frac{n^{1/4}}{2c_1} \times \frac{n^{1/2}}{4}$ . Two clusters of size  $M$  placed on the same horizontal line and separated by distance  $d = \frac{n^{1/2}}{4}$  pair up and start back-and-forth beamforming. The vertical separation between adjacent cluster pairs is  $c_2 n^{1/4+\epsilon}$ .

After each transmission, the signal received by a node in a given cluster is the sum of the signals coming from the facing cluster, of those coming from other clusters, and of the noise. We assume a sufficiently large vertical distance  $c_2 n^{1/4+\epsilon}$  separating any two adjacent cluster pairs, as illustrated on Fig. 2.1. We show below that the broadcast rate between the operating clusters is  $\Theta(n^{\frac{1}{2}}P)$ . Since we only need  $\Theta(n^\epsilon)$  number of rounds to serve all clusters, phase 2 requires  $\Theta(n^{-\frac{1}{2}+\epsilon}P^{-1})$  time slots per bit. As such, back-and-forth beamforming achieves a broadcast rate of  $\Theta(n^{\frac{1}{2}-\epsilon}P)$  bits per time slot.

In view of the described scheme, we are able to state the following result.

**Theorem 2.2.1.** *For any  $\epsilon > 0$  and  $P = O(n^{-\frac{1}{2}})$ , the following broadcast rate*

$$R_n = \Omega\left(n^{\frac{1}{2}-\epsilon}P\right)$$

*is achievable with high probability<sup>2</sup> in the network. As a consequence, when  $P = \Omega(n^{-\frac{1}{2}})$ , a broadcast rate  $R_n = \Omega(n^{-\epsilon})$  is achievable with high probability.*

The broadcast rate achieved by our scheme outperforms therefore plain time-division in a large network. Interestingly, Theorem 2.2.1 also says that our scheme requires asymptotically less power to achieve the same performance as plain time-division. In a large network, this could allow for example to send control signals or channel state information at low cost in the network, without hurting other transmissions.

<sup>2</sup>that is, with probability at least  $1 - O\left(\frac{1}{n^p}\right)$  as  $n \rightarrow \infty$ , where the exponent  $p$  is as large as we want.



Before proceeding with the proof of the theorem, the following lemma provides an upper bound on the probability that the number of nodes inside each cluster deviates from its mean by a large factor. Its proof can be found in [35], but is also provided in Appendix 2.A for completeness.

**Lemma 2.2.2.** *Let us consider a cluster of area  $M$  with  $M = n^\beta$  for some  $0 < \beta < 1$ . The number of nodes inside each cluster is then between  $((1 - \delta)M, (1 + \delta)M)$  with probability larger than  $1 - \frac{n}{M} \exp(-\Delta(\delta)M)$  where  $\Delta(\delta)$  is independent of  $n$  and satisfies  $\Delta(\delta) > 0$  for  $\delta > 0$ .*

As shown in Fig. 2.1, two clusters of size  $M = \frac{n^{1/4}}{2c_1} \times \frac{n^{1/2}}{4}$  placed on the same horizontal line and separated by distance  $d = \frac{n^{1/2}}{4}$  form a cluster pair. During the back-and-forth beamforming phase, there are many cluster pairs operating simultaneously. Given that the cluster width is  $\frac{n^{1/4}}{2c_1}$  and the vertical separation between adjacent cluster pairs is  $c_2 n^{1/4+\epsilon}$ , there are

$$N_C = \frac{n^{1/2}}{\frac{n^{1/4}}{2c_1} + c_2 n^{1/4+\epsilon}} = \Theta\left(n^{1/4-\epsilon}\right) \quad (2.2)$$

cluster pairs operating at the same time. Let  $\mathcal{R}_i$  and  $\mathcal{T}_i$  denote the receiving and the transmitting clusters of the  $i$ -th cluster pair, respectively.

Two key ingredients for analyzing the multi-stage back-and-forth beamforming scheme are given in Lemma 2.2.3 and Lemma 2.2.4. The proofs are presented in Appendix 2.A.

**Lemma 2.2.3.** *The maximum beamforming gain between the two clusters of the  $i$ -th cluster pair can be achieved by using a compensation of the phase shifts at the transmit side which is proportional to the horizontal positions of the nodes. More precisely, there exist a constant  $c_1 > 0$  (remember that  $c_1$  is inversely proportional to the width of  $i$ -th cluster) and a constant  $K_1 > 0$  such that the magnitude of the received signal at node  $j \in \mathcal{R}_i$  is lower bounded with high probability by*

$$\left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \geq K_1 \frac{M}{d},$$

where  $x_k$  denotes the horizontal position of node  $k$ .

**Lemma 2.2.4.** *For every constant  $K_2 > 0$ , there exists a sufficiently large separating constant  $c_2 > 0$  such that the magnitude of interfering signals from the simultaneously operating cluster pairs at node  $j \in \mathcal{R}_i$  is upper bounded with high probability by*

$$\left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \leq K_2 \frac{M}{d n^\epsilon} \log n.$$

*Proof of Theorem 2.2.1.* The first phase of the scheme results in noisy observations of the message  $X$  at all nodes, which are given by

$$Y_k^{(0)} = \sqrt{\text{SNR}_k} X + Z_k^{(0)},$$

where  $\mathbb{E}(|X|^2) = \mathbb{E}(|Z_k^{(0)}|^2) = 1$  and  $\text{SNR}_k$  is the signal-to-noise ratio of the signal  $Y_k^{(0)}$  received at the  $k$ -th node before the back-and-forth beamforming starts (time (0) of the back-and-forth beamforming, denoted by the superscript (0) in the variables  $Y_k^{(0)}$  and  $Z_k^{(0)}$ ). In what follows, we drop the index  $k$  from  $\text{SNR}_k$  and only write  $\text{SNR} = \min_k \{\text{SNR}_k\}$ . Note that it does not make a difference at which side of the cluster pairs the back-and-forth beamforming starts or ends. Hence, assume the left-hand side clusters ignite the scheme by amplifying and forwarding the noisy observations of  $X$  to the right-hand side clusters. The signal received at node  $j \in \mathcal{R}_i$  (denoted by the subscript) after the 1st cluster-to-cluster transmission (denoted by the superscript (1)) is given by

$$Y_j^{(1)} = \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} A Y_k^{(0)} + Z_j^{(1)} \quad (2.3)$$

where  $A$  is the amplification factor (to be calculated later) and  $Z_j^{(1)}$  is additive white Gaussian noise of variance  $\Theta(1)$ . We start by applying Lemma 2.2.3 and Lemma 2.2.4 to lower bound

$$\begin{aligned} \left| \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &\geq \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\quad - \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\geq \left( K_1 - K_2 \frac{\log n}{n^\epsilon} \right) \frac{M}{d} = \Theta \left( \frac{M}{d} \right). \end{aligned}$$

For the sake of clarity, we can therefore approximate<sup>3</sup> the expression in (2.3)

<sup>3</sup>We make this approximation to lighten the notation and make the exposition clear, but needless to say, the whole analysis goes through without the approximation; it just becomes barely readable.

as follows

$$\begin{aligned}
Y_j^{(1)} &= \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} A \sqrt{\text{SNR}_k} X \\
&+ \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} A Z_k^{(0)} + Z_j^{(1)} \\
&\simeq \frac{AM}{d} \sqrt{\text{SNR}} X + \frac{A\sqrt{N_C M}}{d} Z^{(0)} + Z_j^{(1)} \\
&= \frac{AM}{d} \sqrt{\text{SNR}} X + \frac{AM}{d} \sqrt{\frac{N_C}{M}} Z^{(0)} + Z_j^{(1)},
\end{aligned}$$

where

$$Z^{(0)} = \frac{d}{\sqrt{N_C M}} \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} Z_k^{(0)}.$$

Note that  $\mathbb{E}(|Z^{(0)}|^2) = \Theta(1)$ . Repeating the same process  $t$  times (denoted by the superscript  $(t)$ ) in a back-and-forth manner results in a final signal at node  $j \in \mathcal{R}_i$  in the left or the right cluster (depending on whether  $t$  is odd or even) that is given by

$$\begin{aligned}
Y_j^{(t)} &= \left(\frac{AM}{d}\right)^t \sqrt{\text{SNR}} X + \left(\frac{AM}{d}\right)^t \sqrt{\frac{N_C}{M}} Z^{(0)} \\
&+ \dots + \left(\frac{AM}{d}\right)^{t-s} \sqrt{\frac{N_C}{M}} Z^{(s)} + \dots + Z_j^{(t)},
\end{aligned}$$

where for  $0 \leq s \leq t-1$ ,

$$Z^{(s)} = \frac{d}{\sqrt{N_C M}} \sum_{b=1}^{N_C} \sum_{k \in \mathcal{T}_b} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} Z_k^{(s)}.$$

Note again that for  $0 \leq s \leq t-1$ ,  $\mathbb{E}(|Z^{(s)}|^2) = \Theta(1)$ , and  $Z_j^{(t)}$  is additive white Gaussian noise of variance  $\Theta(1)$ . Finally, note that the reason behind the amplification of the noise at each beamforming cycle by a factor  $A\frac{M}{d}$  is Lemma 2.2.4 which ensures an upper bound on the beamforming gain of the noise signals, i.e.,

$$\begin{aligned}
\left| \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &\leq \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\
&+ \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\
&\leq \left(1 + K_2 \frac{\log n}{n^\epsilon}\right) \frac{M}{d}.
\end{aligned}$$

(notice indeed that the first term in the middle expression

$$\left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \leq \frac{M}{d},$$

because it contains  $M$  terms, all less than  $1/d$ ). Now, we want the power of the signal to be of order 1, that is:

$$\begin{aligned} \mathbb{E} \left( \left( \left( \frac{AM}{d} \right)^t \sqrt{\text{SNR} X} \right)^2 \right) &= \left( \frac{AM}{d} \right)^{2t} \text{SNR} = \Theta(1) \quad (2.4) \\ \Rightarrow A &= \Theta \left( \frac{d}{M} \text{SNR}^{-\frac{1}{2t}} \right). \end{aligned}$$

Since at each round of TDMA cycle there are  $\Theta(N_C M)$  nodes transmitting, then every node will be active  $\Theta\left(\frac{N_C M}{n}\right)$  fraction of the time. As such, the amplification factor is given by

$$A = \Theta \left( \sqrt{\frac{n}{N_C M} \tau P} \right),$$

where  $\tau$  is the number of time slots between two consecutive transmissions, i.e. every  $\tau$  time slots we have one transmission. Therefore, we have

$$\begin{aligned} A &= \Theta \left( \frac{d}{M} \text{SNR}^{-\frac{1}{2t}} \right) = \Theta \left( \sqrt{\frac{n}{N_C M} \tau P} \right) \\ \Rightarrow \tau &= \Theta \left( \frac{N_C d^2}{n M P} \text{SNR}^{-\frac{1}{t}} \right). \end{aligned}$$

We can pick the number of back-and-forth transmissions  $t$  sufficiently large to ensure that  $\text{SNR}^{-\frac{1}{t}} = O(n^\epsilon)$ , which results in

$$\tau = O \left( \frac{1}{n^{1/2} P} \right).$$

Moreover, the noise power is given by

$$\begin{aligned} &\sum_{s=0}^{t-1} \mathbb{E} \left( \left( \left( \frac{AM}{d} \right)^{t-s} \sqrt{\frac{N_C}{M}} Z^{(s)} \right)^2 \right) + \mathbb{E} \left( \left( Z_j^{(t)} \right)^2 \right) \\ &\leq t \mathbb{E} \left( \left( \left( \frac{AM}{d} \right)^t \sqrt{\frac{N_C}{M}} Z^{(0)} \right)^2 \right) + 1 \\ &\leq t \left( \frac{AM}{d} \right)^{2t} \frac{N_C}{M} + 1 \\ &\stackrel{(a)}{\leq} t + 1 = \Theta(1), \end{aligned}$$

where (a) holds if and only if  $\text{SNR} = \Omega(N_C/M) = \Omega(n^{-1/2-\epsilon})$  (check eq. (2.4)), which is true: Distance separating any two nodes in the network is at most  $\sqrt{2n}$ , which implies that the SNR of the received signal at all the nodes in the network is  $\Omega(n^{-1/2})$ .

Given that the required  $\tau = O\left(\frac{1}{n^{1/2}P}\right)$ , we can see that for  $P = O(n^{-1/2})$  the broadcast rate between simultaneously operating clusters is  $\Omega(n^{1/2}P)$ . Finally, applying TDMA of  $\frac{n}{N_C M} = \Theta(n^\epsilon)$  steps ensures that  $X$  is successfully decoded at all nodes and the broadcast rate  $R_n = \Omega(n^{1/2-\epsilon}P)$ .

As a last remark, let us mention that the consequence stated in the theorem for the regime where more power is available at the transmitters is an obvious one: by simply reducing the amount of power used at each node to exactly  $n^{-1/2} \leq P$ , one achieves the following broadcast rate, using the first part of the theorem:

$$R_n = \Omega\left(n^{\frac{1}{2}-\epsilon} n^{-\frac{1}{2}}\right) = \Omega(n^{-\epsilon}).$$

This completes the proof of the theorem.  $\square$

### 2.3 Optimality of the Scheme

In this section, we first establish a general upper bound on the broadcast capacity of wireless networks at low SNR, which applies to a general fading matrix  $H$  (with proper measurement units such that again,  $\text{SNR}_s = P$  in these units).

**Theorem 2.3.1.** *Let us consider a network of  $n$  nodes and let  $H$  be the  $n \times n$  matrix with  $h_{jj} = 0$  on the diagonal and  $h_{jk}$  = the fading coefficient between node  $j$  and node  $k$  in the network. The broadcast capacity of such a network with  $n$  nodes is then upper bounded by*

$$C_n \leq P \|H\|^2$$

where  $P$  is the power available per node and  $\|H\|$  is the spectral norm (i.e. the largest singular value) of  $H$ .

*Proof.* Using the classical cut-set bound [11, Theorem 15.10.1], the following upper bound on the broadcast capacity  $C_n$  is obtained:

$$C_n \leq \max_{\substack{P_X: \\ \mathbb{E}(|X_k|^2) \leq P, \forall 1 \leq k \leq n}} \min_{1 \leq j \leq n} I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j | X_j).$$

Moreover, we have

$$\begin{aligned} I(X_{\{1, \dots, n\} \setminus \{j\}}, X_j; Y_j) &= I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j) + I(X_j; Y_j | X_{\{1, \dots, n\} \setminus \{j\}}) \\ &\stackrel{(a)}{=} I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j) \\ &= I(X_j; Y_j) + I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j | X_j) \\ &\stackrel{(b)}{\geq} I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j | X_j), \end{aligned}$$

where (a) follows from the fact that  $X_j - X_{\{1, \dots, n\} \setminus \{j\}} - Y_j$  forms a Markov chain, which means that  $I(X_j; Y_j | X_{\{1, \dots, n\} \setminus \{j\}}) = 0$ , and (b) follows from the fact that  $I(X_j; Y_j) \geq 0$ . Therefore, we get

$$\begin{aligned} C_n &\leq \max_{\substack{p_X: \\ \mathbb{E}(|X_k|^2) \leq P, \forall 1 \leq k \leq n}} \min_{1 \leq j \leq n} I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j | X_j) \\ &\leq \max_{\substack{p_X: \\ \mathbb{E}(|X_k|^2) \leq P, \forall 1 \leq k \leq n}} \min_{1 \leq j \leq n} I(X_{\{1, \dots, n\} \setminus \{j\}}; Y_j) \\ &\leq \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \min_{1 \leq j \leq n} \log(1 + h_j Q_X h_j^\dagger) \end{aligned}$$

where  $h_j = (h_{j1}, \dots, h_{j,j-1}, 0, h_{j,j+1}, \dots, h_{jn})$ , as the joint distribution  $p_X$  maximizing the above expression is clearly Gaussian. Using then the fact that the minimum of a set of numbers is less than its average, the above expression can be further bounded by

$$\begin{aligned} C_n &\leq \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \frac{1}{n} \sum_{j=1}^n \log(1 + h_j Q_X h_j^\dagger) \\ &= \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \frac{1}{n} \sum_{j=1}^n \log \det(I_n + h_j^\dagger h_j Q_X) \\ &\leq \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \log \det \left( I_n + \frac{1}{n} \sum_{j=1}^n h_j^\dagger h_j Q_X \right) \end{aligned}$$

using successively the property that  $\log \det(I + AB) = \log \det(I + BA)$  and the fact that  $\log \det(\cdot)$  is concave. Observing now that the  $n \times n$  matrix  $H$  whose entries are given by  $h_{jk} = (h_j)_k$  is the one in the theorem statement and that  $\sum_{j=1}^n h_j^\dagger h_j = H^\dagger H$ , we can rewrite, using again  $\log \det(I + AB) = \log \det(I + BA)$ :

$$\begin{aligned} C_n &\leq \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \log \det \left( I_n + \frac{1}{n} H Q_X H^\dagger \right) \\ &\leq \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \frac{1}{n} \text{Tr}(H Q_X H^\dagger) \\ &\leq \max_{\substack{Q_X \geq 0 \\ (Q_X)_{kk} \leq P, \forall 1 \leq k \leq n}} \frac{1}{n} \text{Tr}(Q_X) \|H\|^2 = P \|H\|^2 \end{aligned}$$

where the last inequality follows from the fact that  $\text{Tr}(BAB^\dagger) \leq \|B\|^2 \text{Tr}(A)$ , for any matrix  $B$  and  $A \geq 0$ . This completes the proof.  $\square$

We now aim to specialize Theorem 2.3.1 to line-of-sight fading, where the

matrix  $H$  is given by

$$h_{jk} = \begin{cases} 0 & \text{if } j = k \\ \frac{\exp(2\pi i r_{jk})}{r_{jk}} & \text{if } j \neq k. \end{cases} \quad (2.5)$$

The rest of the section is devoted to proving the proposition below which, together with Theorem 2.3.1, shows the asymptotic optimality of the back-and-forth beamforming scheme presented in Section 2.2 for two-dimensional networks at low SNR and under LOS fading. It is also worth mentioning that for a one-dimensional network in LOS environment, Theorem 2.3.1 allows to recover the result already obtained in [36].

**Proposition 2.3.2.** *Let  $H$  be the  $n \times n$  matrix given by (2.5). For every  $\epsilon > 0$ , there exists a constant  $c > 0$  such that*

$$\|H\|^2 \leq c n^{\frac{1}{2} + \epsilon}$$

with high probability as  $n$  gets large.

Analyzing directly the asymptotic behavior of  $\|H\|$  reveals itself difficult. We therefore decompose our proof into simpler subproblems. The first building block of the proof is the following Lemma, which can be viewed as a generalization of the classical Geršgorin discs' inequality.

**Lemma 2.3.3.** *Let  $B$  be an  $n \times n$  matrix decomposed into blocks  $B_{jk}$ ,  $j, k = 1, \dots, K$ , each of size  $M \times M$ , with  $n = KM$ . Then*

$$\|B\| \leq \max \left\{ \max_{1 \leq j \leq K} \sum_{k=1}^K \|B_{jk}\|, \max_{1 \leq j \leq K} \sum_{k=1}^K \|B_{kj}\| \right\}$$

The proof of this Lemma is relegated to Appendix 2.A. The second building block of this proof is the following lemma, the proof of which is also given in Appendix 2.A.

**Lemma 2.3.4.** *Let  $\widehat{H}$  be the  $M \times M$  channel matrix between two square clusters of  $M$  nodes distributed uniformly at random, each of area  $A = M$ . Then there exists a constant  $c > 0$  such that*

$$\|\widehat{H}\|^2 \leq c \frac{M^{1+\epsilon}}{d}$$

with high probability as  $M$  gets large, where  $2\sqrt{M} \leq d \leq M$  denotes the distance between the centers of the two clusters.

*Proof of Proposition 2.3.2.* The strategy for the proof is now the following: in order to bound  $\|H\|$ , we divide the matrix into smaller blocks, apply Lemma 2.3.3 and Lemma 2.3.4 in order to bound the off-diagonal terms  $\|H_{jk}\|$ . For the diagonal terms  $\|H_{jj}\|$ , we reapply Lemma 2.3.3 and proceed in a recursive

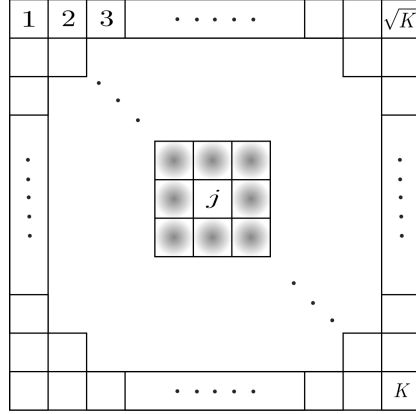


Figure 2.2 –  $\sqrt{n} \times \sqrt{n}$  network split into  $K$  clusters and numbered in order. As such,  $R_j = \{j - \sqrt{K} - 1, j - \sqrt{K}, j - \sqrt{K} + 1, j - 1, j, j + 1, j + \sqrt{K} - 1, j + \sqrt{K}, j + \sqrt{K} + 1\}$ , which represents the center square containing the cluster  $j$  and its 8 neighbors (marked in shades).

manner, until we reach small size blocks for which a loose estimate is sufficient to conclude. Let us therefore decompose the network into  $K$  clusters of  $M$  nodes each, with  $n = KM$ . By Lemma 2.3.3, we obtain

$$\|H\| \leq \max \left\{ \max_{1 \leq j \leq K} \sum_{k=1}^K \|H_{jk}\|, \max_{1 \leq j \leq K} \sum_{k=1}^K \|H_{kj}\| \right\}, \quad (2.6)$$

where the  $n \times n$  matrix  $H$  is decomposed into blocks  $H_{jk}$ ,  $j, k = 1, \dots, K$ , with  $H_{jk}$  denoting the  $M \times M$  channel matrix between cluster number  $j$  and cluster number  $k$  in the network. Let us also denote by  $d_{jk}$  the corresponding inter-cluster distance, measured from the centers of these clusters. According to Lemma 2.3.4, if  $d_{jk} \geq 2\sqrt{M}$ , then there exists a constant  $c > 0$  such that

$$\|H_{jk}\|^2 \leq c \frac{M^{1+\epsilon}}{d_{jk}} \leq cn^\epsilon \frac{M}{d_{jk}}$$

with high probability as  $M \rightarrow \infty$ .

Let us now fix  $j \in \{1, \dots, K\}$  and define  $R_j = \{1 \leq k \leq K : d_{jk} < 2\sqrt{M}\}$  and  $S_j = \{1 \leq k \leq K : d_{jk} \geq 2\sqrt{M}\}$  (see Fig. 2.2). By the above inequality, we obtain

$$\sum_{k=1}^K \|H_{jk}\| \leq \sum_{k \in R_j} \|H_{jk}\| + \sqrt{cn^\epsilon} \sum_{k \in S_j} \sqrt{\frac{M}{d_{jk}}}$$

with high probability as  $M$  gets large. Observe that as there are  $8l$  clusters or



less at distance  $l\sqrt{M}$  from cluster  $j$ , we obtain

$$\sum_{k \in S_j} \sqrt{\frac{M}{d_{jk}}} \leq \sum_{l=2}^{\sqrt{K}} 8l \sqrt{\frac{M}{l\sqrt{M}}} = O\left(M^{1/4}K^{3/4}\right) = O\left(\frac{n^{3/4}}{M^{1/2}}\right)$$

as  $K = n/M$ . There remains to upper bound the sum over  $R_j$ . Observe that this sum contains at most 9 terms: namely the term  $k = j$  and the 8 terms corresponding to the 8 neighboring clusters of cluster  $j$ . It should then be observed that for each  $k \in R_j$ ,  $\|H_{jk}\| \leq \|H(R_j)\|$ , where  $H(R_j)$  is the  $9M \times 9M$  matrix made of the  $9 \times 9$  blocks  $H_{j_1, j_2}$  such that  $j_1, j_2 \in R_j$ . Finally, this leads to

$$\sum_{k=1}^K \|H_{jk}\| \leq 9\|H(R_j)\| + \sqrt{cn^\epsilon} \frac{n^{3/4}}{M^{1/2}}$$

Using the symmetry of this bound and (2.6), we obtain

$$\|H\| \leq 9 \max_{1 \leq j \leq K} \|H(R_j)\| + \sqrt{cn^\epsilon} \frac{n^{3/4}}{M^{1/2}}. \quad (2.7)$$

A key observation is now the following: the  $9M \times 9M$  matrix  $H(R_j)$  has exactly the same structure as the original matrix  $H$ . So in order to bound its norm  $\|H(R_j)\|$ , the same technique may be reused! This leads to the following recursive Lemma.

**Lemma 2.3.5.** *Assume there exist constants  $c > 0$  and  $b \in [1/4, 1/2]$  such that*

$$\|H\| \leq \sqrt{cn^\epsilon} n^b$$

*with high probability as  $n$  gets large. Then there exists a constant  $c' > 0$  such that*

$$\|H\| \leq \sqrt{c'n^\epsilon} n^{f(b)}$$

*with high probability as  $n$  gets large, where  $f(b) = \frac{3b}{4b+2} < b$ .*

*Proof.* The assumption made implies that there exist  $c > 0$  and  $b \in [1/4, 1/2]$  such that for every  $M \times M$  diagonal sub-block  $H_M$  of the matrix  $H$ ,

$$\|H_M\| \leq \sqrt{cM^\epsilon} M^b \leq \sqrt{cn^\epsilon} M^b$$

with high probability as  $M$  gets large. Together with (2.7), this implies that

$$\begin{aligned} \|H\| &\leq 9\sqrt{cn^\epsilon} M^b + \sqrt{cn^\epsilon} \frac{n^{3/4}}{M^{1/2}} \\ &= 10\sqrt{cn^\epsilon} \left( M^b + \frac{n^{3/4}}{M^{1/2}} \right) \end{aligned}$$

Choosing  $M = \lfloor n^{3/(4b+2)} \rfloor$ , we obtain

$$\|H\| \leq \sqrt{c'n^\epsilon} n^{3b/(4b+2)}.$$

□

Besides, it is easy to check that the assumption of Lemma 2.3.5 holds with  $b = 1/2$ . Apply for this the slightly modified version of the classical Geršgorin inequality (which is nothing but the statement of Lemma 2.3.3 applied to the case  $M = 1$ ):

$$\|H\| \leq \max \left\{ \max_{1 \leq j \leq n} \sum_{k=1}^n |h_{jk}|, \max_{1 \leq j \leq n} \sum_{k=1}^n |h_{kj}| \right\} = \max_{1 \leq j \leq n} \sum_{\substack{k=1 \\ k \neq j}}^n \frac{1}{r_{jk}}$$

For any  $1 \leq j \leq n$ , it holds with high probability that for  $c$  large enough,

$$\sum_{\substack{k=1 \\ k \neq j}}^n \frac{1}{r_{jk}} \leq \sum_{l=1}^{\sqrt{n}} (cl \log n) \frac{1}{l} = O(\sqrt{n} \log n)$$

which implies that  $\|H\| = O(\sqrt{n^{1+\epsilon}})$  for any  $\epsilon > 0$ .

By applying Lemma 2.3.5 successively, we obtain a decreasing sequence of upper bounds on  $\|H\|$ :

$$\|H\| \leq \sqrt{cn^\epsilon} n^{b_0}, \quad \leq \sqrt{cn^\epsilon} n^{b_1}, \quad \leq \sqrt{cn^\epsilon} n^{b_2}$$

where the sequence  $b_0 = 1/2$ ,  $b_1 = f(b_0) = 3b_0/(4b_0 + 2) = 3/8$ ,  $b_2 = f(b_1) = 3b_1/(4b_1 + 2) = 9/28$  converges to the fixed point  $b^* = f(b^*) = 1/4$  (as  $f$  is strictly increasing on  $[\frac{1}{4}, \frac{1}{2}]$  and  $f(b) < b$  for every  $\frac{1}{4} < b \leq \frac{1}{2}$ ). This finally proves Proposition 2.3.2.  $\square$

## 2.4 Summary

In this chapter, we characterize the broadcast capacity of two-dimensional wireless networks, specifically extended networks of area  $A = n$ , at low SNR in line-of-sight environment, which is achieved via a back-and-forth beamforming scheme. We show that the broadcast capacity is upper bounded by the total power transfer in the network, which in turn is equal to  $P \|H\|^2$ . We present a detailed analysis of the largest singular value of the fading matrix  $H$ . We further present a practical broadcasting scheme that guarantees the total power transfer throughout the network. This scheme relies on back-and-forth beamforming among clusters through multiple stage time division channel accesses. In other words, this chapter can be summarized, combining Theorem 2.2.1, Theorem 2.3.1, and Proposition 2.3.2, in the following Theorem.

**Theorem 2.4.1.** *Under line-of-sight model assumption, the aggregate broadcast capacity in a network of  $n$  nodes, area  $A = n$ , and carrier wavelength  $\lambda = 1$ , scales as*

$$C_n \sim \min\{\sqrt{n} P, 1\}$$

*up to a multiplicative factor  $n^\epsilon$ , for any fixed  $\epsilon > 0$ .*

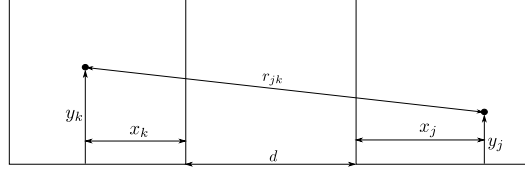


Figure 2.3 – Coordinate system.

## 2.A Appendix

*Proof of Lemma 2.2.2.* The number of nodes in a given cluster is the sum of  $n$  independently and identically distributed Bernoulli random variables  $B_i$ , with  $\mathcal{P}(B_i = 1) = M/n$ . Hence

$$\begin{aligned}
& \mathbb{P}\left(\sum_{i=1}^n B_i \geq (1 + \delta)M\right) \\
&= \mathbb{P}\left(\exp\left(s \sum_{i=1}^n B_i\right) \geq \exp(s(1 + \delta)M)\right) \\
&\leq \mathbb{E}^n(\exp(sB_1)) \exp(-s(1 + \delta)M) \\
&= \left(\frac{M}{n} \exp(s) + 1 - \frac{M}{n}\right)^n \exp(-s(1 + \delta)M) \\
&\leq \exp(-M(s(1 + \delta) - \exp(s) + 1)) = \exp(-M\Delta_+(\delta))
\end{aligned}$$

where  $\Delta_+(\delta) = (1 + \delta) \log(1 + \delta) - \delta$  by choosing  $s = \log(1 + \delta)$ . The proof of the lower bound follows similarly by considering the random variables  $-B_i$ . The conclusion follows from the union bound.  $\square$

*Proof of Lemma 2.2.3.* We present lower and upper bounds on the distance  $r_{jk}$  separating a receiving node  $j \in \mathcal{R}_i$  and a transmitting node  $k \in \mathcal{T}_i$ . Denote by  $x_j, x_k, y_j$ , and  $y_k$  the horizontal and the vertical positions of nodes  $j$  and  $k$ , respectively (as shown in Fig. 2.3). An easy lower bound on  $r_{jk}$  is

$$r_{jk} \geq x_k + x_j + d$$

On the other hand, using the inequality  $\sqrt{1 + x} \leq 1 + \frac{x}{2}$ , we obtain

$$\begin{aligned}
r_{jk} &= \sqrt{(x_k + x_j + d)^2 + (y_j - y_k)^2} \\
&= (x_k + x_j + d) \sqrt{1 + \frac{(y_j - y_k)^2}{(x_k + x_j + d)^2}} \\
&\leq x_k + x_j + d + \frac{(y_j - y_k)^2}{2d} \leq x_k + x_j + d + \frac{1}{2c_1^2}.
\end{aligned}$$

Therefore,

$$0 \leq r_{jk} - x_k - x_j - d \leq \frac{1}{2c_1^2}.$$

After bounding  $r_{jk}$ , we can proceed to the proof of the lemma as follows:

$$\begin{aligned} \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &= \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k - x_j - d))}{r_{jk}} \right| \\ &\geq \Re \left( \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k - x_j - d))}{r_{jk}} \right) \\ &\geq \sum_{k \in \mathcal{T}_i} \frac{\cos\left(\frac{\pi}{c_1^2}\right)}{r_{jk}} \geq K_1 \frac{M}{d}, \end{aligned}$$

where the constant  $c_1$  is chosen sufficiently large so that  $\cos\left(\frac{\pi}{c_1^2}\right) > 0$ .  $\square$

*Proof of Lemma 2.2.4.* There are  $N_C$  clusters transmitting simultaneously. Except for the horizontally adjacent cluster of a given cluster pair ( $i$ -th cluster pair), all the rest of the transmitting clusters are considered as interfering clusters (there are  $N_C - 1$  of them). With high probability, each cluster contains  $\Theta(M)$  nodes.

For the sake of clarity, we assume here that every cluster contains exactly  $M$  nodes, but the argument holds in the general case. In this lemma, we upper bound the magnitude of interfering signals from the simultaneously interfering clusters at node  $j \in \mathcal{R}_i$  as follows

$$\begin{aligned} &\left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \left| \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \left| \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| + \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \left| \sum_{k \in \mathcal{T}_l} \frac{\sin(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq 2 \sum_{l=1}^{N_C} \left| \sum_{k \in \mathcal{T}'_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| + 2 \sum_{l=1}^{N_C} \left| \sum_{k \in \mathcal{T}'_l} \frac{\sin(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \end{aligned}$$

where  $\mathcal{T}'_l$  denotes the  $l$ -th interfering transmit cluster that is at a vertical distance of  $(l-1)\frac{n^{1/4}}{2c_1} + lc_2n^{1/4+\epsilon}$  from the desired receiving cluster  $\mathcal{R}_i$ . Note that there are at most two clusters that are at a vertical distance of  $(l-1)\frac{n^{1/4}}{2c_1} +$

$lc_2n^{1/4+\epsilon}$  from the receiving cluster. We further upper bound the first term (cosine terms) in the equation above as follows (notice that we can upper bound the second term (sine terms) in exactly the same fashion):

$$\begin{aligned}
\left| \sum_{k \in \mathcal{T}'_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| &= \left| \sum_{k \in \mathcal{T}'_l} X_k^{(l)} \right| \\
&= \left| \sum_{k \in \mathcal{T}'_l} (X_k^{(l)} - \mathbb{E}(X_k^{(l)})) + \sum_{k \in \mathcal{T}'_l} \mathbb{E}(X_k^{(l)}) \right| \\
&\stackrel{(a)}{\leq} \left| \sum_{k \in \mathcal{T}'_l} (X_k^{(l)} - \mathbb{E}(X_k^{(l)})) \right| + \left| \sum_{k \in \mathcal{T}'_l} \mathbb{E}(X_k^{(l)}) \right| \\
&\stackrel{(b)}{=} M \left| \frac{1}{M} \sum_{k \in \mathcal{T}'_l} (X_k^{(l)} - \mathbb{E}(X_k^{(l)})) \right| + M |\mathbb{E}(X_1^{(l)})|
\end{aligned} \tag{2.8}$$

where (a) follows from the triangle inequality and (b) results from the fact that the  $X_k^{(l)}$ 's (note that  $X_k^{(l)} = (\cos(2\pi(r_{jk} - x_k)))/r_{jk} \forall k \in \mathcal{T}'_l$ ) are independent and identically distributed. Let us first upper bound the second term in (2.8):  $\forall k \in \mathcal{T}'_l$ , we have

$$|r_{jk} - x_k| = \sqrt{(x_k + x_j + d)^2 + (y_j - y_k)^2} - x_k \geq d = \frac{n^{1/2}}{4}$$

is a  $C^2$  function and

$$\begin{aligned}
\left| \frac{\partial(r_{jk} - x_k)}{\partial y_k} \right| &= \left| \frac{\partial r_{jk}}{\partial y_k} \right| = \frac{|y_k - y_j|}{r_{jk}} \\
&\geq \frac{lc_2 n^{1/4+\epsilon} + (l-1) \frac{n^{1/4}}{2c_1}}{n^{1/2}} \\
&\geq lc_2 n^{-1/4+\epsilon}
\end{aligned}$$

Moreover,  $r''_{jk}$  changes sign at most twice. By the integration by parts formula, we obtain

$$\begin{aligned}
&\int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \\
&= \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{2\pi r'_{jk}}{2\pi r'_{jk} r_{jk}} \cos(2\pi(r_{jk} - x_k)) \\
&= \frac{-\sin(2\pi(r_{jk} - x_k))}{2\pi r'_{jk} r_{jk}} \Big|_{y_{k_0}}^{y_{k_1}} + \frac{1}{2\pi} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{r_{jk} r''_{jk} + (r'_{jk})^2}{(r'_{jk} r_{jk})^2} \sin(2\pi(r_{jk} - x_k))
\end{aligned}$$

which in turn yields the upper bound

$$\begin{aligned}
& \left| \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\
& \leq \frac{1}{2\pi} \left( \frac{2}{\min_{y_k} \{|r'_{jk}| |r_{jk}|\}} + \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{|r''_{jk}|}{(r'_{jk})^2 |r_{jk}|} + \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{1}{r_{jk}^2} \right) \\
& \leq \frac{1}{2\pi} \left( \frac{4}{l c_2 n^{1/4+\epsilon}} + \frac{1}{\min_{y_k} \{|r_{jk}|\}} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{|r''_{jk}|}{(r'_{jk})^2} + \frac{|y_{k_1} - y_{k_0}|}{\min_{y_k} \{r_{jk}^2\}} \right) \\
& \leq \frac{1}{2\pi} \left( \frac{4}{l c_2 n^{1/4+\epsilon}} + \frac{4}{l c_2 n^{1/4+\epsilon}} + \frac{2}{n^{3/4}} \right) \leq \frac{9/(2\pi)}{l c_2 n^{1/4+\epsilon}}.
\end{aligned}$$

Therefore, for any  $k \in \mathcal{T}'_l$ ,

$$\begin{aligned}
\left| \mathbb{E} \left( X_k^{(l)} \right) \right| &= \left| \frac{4}{n^{1/2}} \int_0^{n^{1/4}} dx_k \frac{1}{|y_{k_1} - y_{k_0}|} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\
&\leq \frac{4}{n^{1/2} |y_{k_1} - y_{k_0}|} \int_0^{n^{1/4}} dx_k \left| \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\
&\leq \frac{9/(2\pi)}{|y_{k_1} - y_{k_0}| l c_2 n^{1/4+\epsilon}} \leq \frac{9c_1}{\pi c_2} \frac{1}{l n^{1/2+\epsilon}} \leq \frac{9c_1}{\pi c_2} \frac{1}{l d n^\epsilon}. \quad (2.9)
\end{aligned}$$

We further upper bound the first term in (2.8) by using Hoeffding's inequality [23]. Note that the  $X_k^{(l)}$ 's are i.i.d. and integrable random variables such that for any  $1 \leq l \leq N_C$  and  $\forall k \in \mathcal{T}'_l$ , we have  $X_k^{(l)} \in [-1/d, 1/d]$ . As such, Hoeffding's inequality yields

$$\begin{aligned}
\mathbb{P} \left( \left| \frac{1}{M} \sum_{k \in \mathcal{T}'_l} \left( X_k^{(l)} - \mathbb{E} \left( X_k^{(l)} \right) \right) \right| > t \right) &\leq 2 \exp \left( -\frac{M t^2}{2/d^2} \right) \\
&= 2 \exp \left( -\frac{1}{2} M d^2 t^2 \right) \\
&\stackrel{(a)}{=} 2 \exp(-n^\epsilon),
\end{aligned}$$

where (a) holds if  $t = \frac{1}{d} \sqrt{\frac{2n^\epsilon}{M}}$ . Therefore, we have

$$\left| \frac{1}{M} \sum_{k \in \mathcal{T}'_l} \left( X_k^{(l)} - \mathbb{E} \left( X_k^{(l)} \right) \right) \right| \leq \frac{1}{d} \sqrt{\frac{2n^\epsilon}{M}} \quad (2.10)$$

with probability  $\geq 1 - 2 \exp(-n^\epsilon)$ . Combining (2.9) and (2.10), we can upper bound (2.8) as follows

$$\begin{aligned} \left| \sum_{k \in \mathcal{T}'_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| &\leq M \left| \frac{1}{M} \sum_{k \in \mathcal{T}'_l} (X_k^{(l)} - \mathbb{E}(X_k^{(l)})) \right| + M |\mathbb{E}(X_1^{(l)})| \\ &\leq \frac{M}{d} \sqrt{\frac{2n^\epsilon}{M}} + \frac{9c_1}{\pi c_2} \frac{M}{d n^\epsilon}. \end{aligned}$$

Finally, we have

$$\begin{aligned} &\left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq 2 \sum_{l=1}^{N_C} \left| \sum_{k \in \mathcal{T}'_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| + 2 \sum_{l=1}^{N_C} \left| \sum_{k \in \mathcal{T}'_l} \frac{\sin(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\ &\stackrel{(a)}{\leq} 4 \sum_{l=1}^{N_C} \left( \frac{M}{d} \sqrt{\frac{2n^\epsilon}{M}} + \frac{9c_1}{\pi c_2} \frac{M}{d n^\epsilon} \right) \\ &\leq 4\sqrt{2} \frac{N_C \sqrt{n^\epsilon} M}{d} + \frac{36c_1}{\pi c_2} \frac{M}{d n^\epsilon} \log n \\ &\leq \left( 4\sqrt{2} \frac{N_C n^{3\epsilon/2}}{\sqrt{M} \log n} + \frac{36c_1}{\pi c_2} \right) \frac{M}{d n^\epsilon} \log n \\ &= \left( \Theta \left( \frac{n^{1/4-\epsilon} n^{3\epsilon/2}}{n^{3/8} \log n} \right) + \Theta(1) \right) \frac{M}{d n^\epsilon} \log n = \Theta \left( \frac{M}{d n^\epsilon} \log n \right), \end{aligned}$$

where (a) holds with high probability (more precisely, with probability  $\geq 1 - 4 N_C \exp(-n^\epsilon)$ ), which concludes the proof.  $\square$

*Proof of Lemma 2.3.3.* - Let us first consider the case where  $B$  is a Hermitian and positive semi-definite matrix. Then  $\|B\| = \lambda_{\max}(B)$ , the largest eigenvalue of  $B$ . Let now  $\lambda$  be an eigenvalue of  $B$  and  $u$  be its corresponding eigenvector, so that  $\lambda u = Bu$ . Using the block representation of the matrix  $B$ , we have

$$\lambda u_j = \sum_{k=1}^K B_{jk} u_k, \quad \forall 1 \leq j \leq K$$

where  $u_j$  is the  $j^{\text{th}}$  block of the vector  $u$ . Let now  $j$  be such that  $\|u_j\| = \max_{1 \leq k \leq K} \|u_k\|$ . Taking norms and using the triangle inequality, we obtain

$$\begin{aligned} |\lambda| \|u_j\| &= \left\| \sum_{k=1}^K B_{jk} u_k \right\| \leq \sum_{k=1}^K \|B_{jk} u_k\| \\ &\leq \sum_{k=1}^K \|B_{jk}\| \|u_k\| \leq \sum_{k=1}^K \|B_{jk}\| \|u_j\| \end{aligned}$$

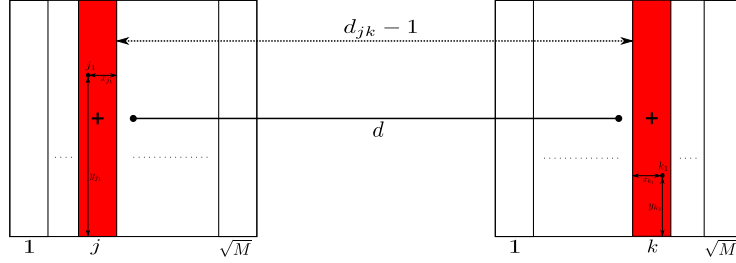


Figure 2.4 – Two square clusters that have a center-to-center distance  $d$ , with each cluster decomposed into  $\sqrt{M}$  vertical  $\sqrt{M} \times 1$  rectangles.  $d_{jk}$  is distance between the centers (marked with cross) of the two rectangles  $j$  and  $k$ . Moreover, we have the points  $j_1(x_{j_1}, y_{j_1})$  and  $k_1(x_{k_1}, y_{k_1})$  in the rectangles  $j$  and  $k$ , respectively.

by the assumption made above. As  $u \neq 0$ ,  $\|u_j\| > 0$ , we obtain

$$|\lambda| \leq \max_{1 \leq j \leq K} \sum_{k=1}^K \|B_{jk}\|$$

As this inequality applies to any eigenvalue  $\lambda$  of  $B$  and  $\|B\| = \lambda_{\max}(B)$ , the claim is proved in this case.

- In the general case, observe first that  $\|B\|^2 = \lambda_{\max}(BB^\dagger)$ , where  $BB^\dagger$  is Hermitian and positive semi-definite. So by what was just proved above,

$$\|B\|^2 = \lambda_{\max}(BB^\dagger) \leq \max_{1 \leq j \leq K} \sum_{k=1}^K \|(BB^\dagger)_{jk}\|$$

Now,  $(BB^\dagger)_{jk} = \sum_{l=1}^K B_{jl}B_{kl}^\dagger$  so

$$\begin{aligned} \sum_{k=1}^K \|(BB^\dagger)_{jk}\| &= \sum_{k=1}^K \left\| \sum_{l=1}^K B_{jl}B_{kl}^\dagger \right\| \\ &\leq \sum_{k=1}^K \sum_{l=1}^K \|B_{jl}\| \|B_{kl}\| \leq \sum_{l=1}^K \|B_{jl}\| \max_{1 \leq j \leq K} \sum_{k=1}^K \|B_{kj}\| \end{aligned}$$

and we finally obtain

$$\|B\|^2 \leq \left( \max_{1 \leq j \leq K} \sum_{l=1}^K \|B_{jl}\| \right) \left( \max_{1 \leq j \leq K} \sum_{k=1}^K \|B_{kj}\| \right)$$

which implies the result, as  $ab \leq \max\{a, b\}^2$  for any two positive numbers  $a, b$ .  $\square$



*Proof of Lemma 2.3.4.* As in the case of  $\|H\|$ , analyzing directly the asymptotic behavior of  $\|\hat{H}\|$  reveals itself difficult. We therefore decompose our proof into simpler subproblems. The strategy is essentially the following: in order to bound  $\|\hat{H}\|$ , we divide the matrix into smaller blocks, bound the smaller blocks  $\|\hat{H}_{jk}\|$ , and apply Lemma 2.3.3. Let us therefore decompose each of the two square clusters into  $\sqrt{M}$  vertical  $\sqrt{M} \times 1$  rectangles of  $\sqrt{M}$  nodes each (See Fig. 2.4).

By Lemma 2.3.3, we obtain

$$\|\hat{H}\| \leq \max \left\{ \max_{1 \leq j \leq \sqrt{M}} \sum_{k=1}^{\sqrt{M}} \|\hat{H}_{jk}\|, \max_{1 \leq j \leq \sqrt{M}} \sum_{k=1}^{\sqrt{M}} \|\hat{H}_{kj}\| \right\} \quad (2.11)$$

where the  $M \times M$  matrix  $\hat{H}$  is decomposed into blocks  $\hat{H}_{jk}$ ,  $j, k = 1, \dots, \sqrt{M}$ , with  $\hat{H}_{jk}$  denoting the  $\sqrt{M} \times \sqrt{M}$  channel matrix between  $k$ -th rectangle of the transmitting cluster and the  $j$ -th rectangle of the receiving cluster. As shown in Fig. 2.4, let us also denote by  $d_{jk}$  the corresponding inter-rectangle distance, measured from the centers of the two rectangles. We want to show that for  $2\sqrt{M} \leq d \leq M$ , where  $d$  is the distance between the centers of the two clusters, there exist constants  $c, c' > 0$  such that

$$\|\hat{H}_{jk}\|^2 \leq c' \frac{M^\epsilon}{d_{jk}} \leq c \frac{M^\epsilon}{d} \quad (2.12)$$

with high probability as  $M \rightarrow \infty$ . Applying (2.11) and (2.12), we get

$$\|\hat{H}\| \leq \max \left\{ \max_{1 \leq j \leq \sqrt{M}} \sum_{k=1}^{\sqrt{M}} \|\hat{H}_{jk}\|, \max_{1 \leq j \leq \sqrt{M}} \sum_{k=1}^{\sqrt{M}} \|\hat{H}_{kj}\| \right\} \leq \left( c \frac{M^{1+\epsilon}}{d} \right)^{1/2}$$

Therefore, what remains to be proven is inequality (2.12). The strategy we propose in order to upper bound  $\|\hat{H}_{jk}\|^2$  is to use the moments' method, relying on the following inequality:

$$\begin{aligned} \|\hat{H}_{jk}\|^2 &= \lambda_{\max}(\hat{H}_{jk} \hat{H}_{jk}^\dagger) \leq \left( \sum_{k=1}^M (\lambda_k(\hat{H}_{jk} \hat{H}_{jk}^\dagger))^\ell \right)^{1/\ell} \\ &= \left( \text{Tr} \left( (\hat{H}_{jk} \hat{H}_{jk}^\dagger)^\ell \right) \right)^{1/\ell} \end{aligned}$$

valid for any  $\ell \geq 1$ . So by Jensen's inequality, we obtain that  $\mathbb{E}(\|\hat{H}_{jk}\|^2) \leq \left( \mathbb{E}(\text{Tr}((\hat{H}_{jk} \hat{H}_{jk}^\dagger)^\ell)) \right)^{1/\ell}$ . In what follows, we show that taking  $\ell \rightarrow \infty$  leads to  $\mathbb{E}(\|\hat{H}_{jk}\|^2) \leq c \frac{\log M}{d_{jk}}$ . More precisely, we show that

$$\mathbb{E}(\text{Tr}((\hat{H}_{jk} \hat{H}_{jk}^\dagger)^\ell)) \leq \frac{M(c \log M)^{\ell-1}}{d_{jk}^{\ell+1}} \quad (2.13)$$

which implies

$$\left(\mathbb{E}(\text{Tr}((\widehat{H}_{jk}\widehat{H}_{jk}^\dagger)^\ell))\right)^{1/\ell} \leq \frac{M^{1/\ell}(c \log M)^{1-1/\ell}}{d_{jk}^{1+1/\ell}} \xrightarrow{\ell \rightarrow \infty} c \frac{\log M}{d_{jk}}.$$

We first prove (2.13) for  $\ell = \{1, 2\}$ , then generalize it to any  $\ell$ . To simplify the notation, let  $F = \widehat{H}_{jk}$ . For  $\ell = 1$ , we obtain

$$\mathbb{E}(\text{Tr}(FF^\dagger)) = \sum_{j_1, k_1=1}^{\sqrt{M}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^*) = \sum_{j_1, k_1=1}^{\sqrt{M}} \mathbb{E}(|f_{j_1 k_1}|^2) = \sum_{j_1, k_1=1}^{\sqrt{M}} \frac{1}{r_{j_1 k_1}^2} \leq \frac{M}{d_{jk}^2} \quad (2.14)$$

Note here that given the definition of  $d_{jk}$ , it only holds that  $r_{j_1 k_1} \geq d_{jk} - 1$  and not  $d_{jk}$ . However, given our assumption that  $d_{jk} \geq \sqrt{M}$ , this simplification does not matter asymptotically and also allows to lighten the notation. We will make this simplification constantly in the following. For  $\ell = 2$ , we obtain

$$\begin{aligned} \mathbb{E}(\text{Tr}((FF^\dagger)^2)) &= \mathbb{E}(\text{Tr}(FF^\dagger FF^\dagger)) \\ &= \sum_{j_1, j_2, k_1, k_2=1}^{\sqrt{M}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*) \\ &\leq \sum_{\substack{j_1=j_2 \\ k_1, k_2}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*) + \sum_{\substack{j_1, j_2 \\ k_1=k_2}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*) \\ &\quad + \sum_{\substack{j_1 \neq j_2 \\ k_1 \neq k_2}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*) \\ &\leq 2 \frac{M^{3/2}}{d_{jk}^4} + M^2 S_2 \stackrel{(a)}{\leq} 2 \frac{M}{d_{jk}^3} + M^2 S_2 \end{aligned}$$

where  $S_2 = |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*)|$  with  $j_1 \neq j_2$  and  $k_1 \neq k_2$  does not depend on the specific choice of  $j_1 \neq j_2$  and  $k_1 \neq k_2$ , and (a) results from fact that  $d_{jk} \geq \sqrt{M}$ . In what follows, we upper bound  $S_2$ .

$$\begin{aligned} S_2 &= |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*)| \\ &= \left| \frac{1}{M^2} \int_0^1 dx_{j_1} \int_0^{\sqrt{M}} dy_{j_1} \int_0^1 dx_{j_2} \int_0^{\sqrt{M}} dy_{j_2} \int_0^1 dx_{k_1} \int_0^{\sqrt{M}} dy_{k_1} \right. \\ &\quad \left. \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \frac{e^{2\pi i(g_{j_1 j_2}(k_1) + g_{j_2 j_1}(k_2))}}{\rho_{j_1 j_2}(k_1) \cdot \rho_{j_2 j_1}(k_2)} \right|, \quad (2.15) \end{aligned}$$

where

$$\begin{aligned} g_{j_1 j_2}(k_1) &= r_{j_1 k_1} - r_{j_2 k_1} = -g_{j_2 j_1}(k_1) \\ &= \sqrt{(d_{jk} - 1 + x_{j_1} + x_{k_1})^2 + (y_{j_1} - y_{k_1})^2} \\ &\quad - \sqrt{(d_{jk} - 1 + x_{j_2} + x_{k_1})^2 + (y_{j_2} - y_{k_1})^2} \quad (2.16) \end{aligned}$$

and

$$\rho_{j_1 j_2}(k_1) = r_{j_1 k_1} \cdot r_{j_2 k_1} = \rho_{j_2 j_1}(k_1) \geq d_{jk}^2, \quad (2.17)$$

where  $0 \leq x_{j_1}, x_{j_2}, x_{k_1}, x_{k_2} \leq 1$  and  $0 \leq y_{j_1}, y_{j_2}, y_{k_1}, y_{k_2} \leq \sqrt{M}$  are the horizontal and the vertical positions, respectively (see Fig. 2.4).

From now on, let us use the short-hand notation

$$\int dj \quad \text{for} \quad \int_0^1 dx_j \int_0^{\sqrt{M}} dy_j$$

Using this short-hand notation as well as equations (2.16) and (2.17), we can rewrite (2.15) as follows

$$\begin{aligned} S_2 &= \left| \frac{1}{M^2} \int dj_1 \int dj_2 \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \int dk_2 \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| \\ &\leq \frac{1}{M^2} \int dj_1 \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \cdot \left| \int dk_2 \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| \\ &= \frac{1}{M^2} \int dj_1 \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \cdot B_{2,1} \end{aligned}$$

where

$$\begin{aligned} B_{2,1} &= \left| \int dk_2 \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| = \left| \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| \\ &\leq \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \left| \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| \\ &= \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \frac{1}{\rho_{j_2 j_1}(k_2)} \leq \frac{\sqrt{M}}{d_{jk}^2} = \tilde{B}. \end{aligned} \quad (2.18)$$

We therefore obtain

$$S_2 \leq \frac{1}{M^{3/2} d_{jk}^2} \int dj_1 \cdot A_{1,2} \quad (2.19)$$

where

$$A_{1,2} = \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right|$$

Before further upper bounding (2.19), we present the following lemma, taken from [46] and adapted to the present situation.

**Lemma 2.A.1.** *Let  $g : [0, \sqrt{M}] \rightarrow \mathbb{R}$  be a  $C^2$  function such that  $|g'(y)| \geq c_1 > 0$  for all  $z \in [0, \sqrt{M}]$  and  $g''$  changes sign at most twice on  $[0, \sqrt{M}]$  (say e.g.  $g''(y) \geq 0$  in  $[y_-, y_+]$  and  $g''(y) \leq 0$  outside). Let also  $\rho : [0, \sqrt{M}] \rightarrow \mathbb{R}$  be*

a  $C^1$  function such that  $|\rho(y)| \geq c_2 > 0$  and  $\rho'(y)$  changes sign at most twice on  $[0, \sqrt{M}]$ . Then

$$\left| \int_0^{\sqrt{M}} dy \frac{e^{2\pi i g(y)}}{\rho(y)} \right| \leq \frac{7}{\pi c_1 c_2}.$$

*Proof.* By the integration by parts formula, we obtain

$$\begin{aligned} \int_0^{\sqrt{M}} dy \frac{e^{2\pi i g(y)}}{\rho(y)} &= \int_0^{\sqrt{M}} dy \frac{2\pi i g'(y)}{2\pi i g'(y)\rho(y)} e^{2\pi i g(y)} \\ &= \frac{e^{2\pi i g(y)}}{2\pi i g'(y)\rho(y)} \Big|_0^{\sqrt{M}} - \int_0^{\sqrt{M}} dy \frac{g''(y)\rho(y) + g'(y)\rho'(y)}{2\pi i (g'(y)\rho(y))^2} e^{2\pi i g(y)} \end{aligned}$$

which in turn yields the upper bound

$$\begin{aligned} \left| \int_0^{\sqrt{M}} dy \frac{e^{2\pi i g(y)}}{\rho(y)} \right| &\leq \frac{1}{2\pi} \left( \frac{1}{|g'(\sqrt{M})||\rho(\sqrt{M})|} + \frac{1}{|g'(0)||\rho(0)|} \right. \\ &\quad \left. + \int_0^{\sqrt{M}} dy \frac{|g''(y)|}{(g'(y))^2 |\rho(y)|} + \int_0^{\sqrt{M}} dy \frac{|\rho'(y)|}{g'(y)(\rho(y))^2} \right) \end{aligned}$$

By the assumptions made in the lemma, we have

$$\begin{aligned} \int_0^{\sqrt{M}} dy \frac{|g''(y)|}{(g'(y))^2 |\rho(y)|} &\leq \frac{1}{c_2} \int_0^{\sqrt{M}} dy \frac{|g''(y)|}{(g'(y))^2} \\ &= \frac{1}{c_2} \left( - \int_0^{y_-} dy \frac{g''(y)}{(g'(y))^2} + \int_{y_-}^{y_+} dy \frac{g''(y)}{(g'(y))^2} - \int_{y_+}^{\sqrt{M}} dy \frac{g''(y)}{(g'(y))^2} \right) \\ &= \frac{1}{c_2} \left( \frac{1}{g'(\sqrt{M})} - \frac{1}{g'(0)} + \frac{2}{g'(y_-)} - \frac{2}{g'(y_+)} \right) \end{aligned}$$

So

$$\int_0^{\sqrt{M}} dy \frac{|g''(y)|}{(g'(y))^2 |\rho(y)|} \leq \frac{7}{c_1 c_2}.$$

We obtain in a similar manner that

$$\int_0^{\sqrt{M}} dy \frac{|\rho'(y)|}{g'(y)(\rho(y))^2} \leq \frac{7}{c_1 c_2}$$

Combining all the bounds, we finally get

$$\left| \int_0^{\sqrt{M}} dy \frac{e^{2\pi i g(y)}}{\rho(y)} \right| \leq \frac{7}{\pi c_1 c_2}$$

□

For any  $\epsilon > 0$ , we can upper bound  $A_{1,2}$  in equation (2.19) as follows

$$\begin{aligned}
A_{1,2} &= \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\
&= \int_{|y_{j_2} - y_{j_1}| < \epsilon \sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\
&\quad + \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\
&\leq \int_{|y_{j_2} - y_{j_1}| < \epsilon \sqrt{M}} dj_2 \int dk_1 \frac{1}{\rho_{j_1 j_2}(k_1)} \\
&\quad + \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\
&\leq \frac{\epsilon M}{d_{jk}^2} + \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \tag{2.20}
\end{aligned}$$

Furthermore, note that

$$\begin{aligned}
g_{j_1 j_2}(k_1) &= r_{j_1 k_1} - r_{j_2 k_1} \\
&= - \int_{x_{j_1}}^{x_{j_2}} \frac{d_{jk} - 1 + x + x_{k_1}}{\sqrt{(d_{jk} - 1 + x + x_{k_1})^2 + (y_{j_1} - y_{k_1})^2}} dx \\
&\quad + \int_{y_{j_1}}^{y_{j_2}} \frac{y_{k_1} - y}{\sqrt{(d_{jk} - 1 + x_{j_2} + x_{k_1})^2 + (y - y_{k_1})^2}} dy
\end{aligned}$$

Therefore, the first order partial derivative of  $g_{j_1 j_2}(k_1)$  with respect to  $y_{k_1}$  is given by

$$\begin{aligned}
\frac{\partial g_{j_1 j_2}(k_1)}{\partial y_{k_1}} &= \int_{x_{j_1}}^{x_{j_2}} \frac{(y_{k_1} - y_{j_1})(d_{jk} - 1 + x + x_{k_1})}{((d_{jk} - 1 + x + x_{k_1})^2 + (y_{j_1} - y_{k_1})^2)^{3/2}} dx \\
&\quad + \int_{y_{j_1}}^{y_{j_2}} \frac{(d_{jk} - 1 + x_{j_2} + x_{k_1})^2}{((d_{jk} - 1 + x_{j_2} + x_{k_1})^2 + (y - y_{k_1})^2)^{3/2}} dy
\end{aligned}$$

From this expression, we deduce that for a constant  $c_3 > 0$

$$\begin{aligned}
\left| \frac{\partial g_{j_1 j_2}(k_1)}{\partial y_{k_1}} \right| &\geq c_3 \frac{|y_{j_2} - y_{j_1}|}{d_{jk}} - \frac{|y_{k_1} - y_{j_1}| \cdot |x_{j_2} - x_{j_1}|}{d_{jk}^2} \\
&\geq c_3 \frac{|y_{j_2} - y_{j_1}|}{d_{jk}} - \frac{\sqrt{M}}{d_{jk}^2} \\
&\stackrel{(a)}{\geq} \frac{c_3 |y_{j_2} - y_{j_1}| - 1}{d_{jk}}, \tag{2.21}
\end{aligned}$$

where (a) follows from the fact that  $d_{jk} \geq \sqrt{M}$ . For  $c_3 |y_{j_2} - y_{j_1}| - 1 > 0$  (we will tune  $\epsilon$  accordingly, as we will see), using (2.17) and (2.21), we can apply

lemma 2.A.1 and upper bound the second term in (2.20) as follows

$$\begin{aligned}
& \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\
& \leq \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \int_0^1 dx_{k_1} \left| \int_0^{\sqrt{M}} dy_{k_1} \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\
& \leq \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dy_{j_2} \frac{7}{\pi \frac{c_3 |y_{j_2} - y_{j_1}| - 1}{d_{jk}} d_{jk}^2} \\
& \leq \frac{7}{\pi c_3 d_{jk}} \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} \frac{1}{|y_{j_2} - y_{j_1}| - 1/c_3} dy_{j_2} \\
& \leq \frac{7}{\pi c_3 d_{jk}} \log \left( \frac{1}{\epsilon} \right) \tag{2.22}
\end{aligned}$$

which gives the following upper bound on (2.20)

$$A_{1,2} \leq \frac{\epsilon M}{d_{jk}^2} + \frac{7}{\pi c_3 d_{jk}} \log \left( \frac{1}{\epsilon} \right) \stackrel{(a)}{=} O \left( \frac{\log M}{d_{jk}} \right) = \tilde{A}, \tag{2.23}$$

where (a) results from choosing  $\epsilon = \frac{c_4}{\sqrt{M}}$  with sufficiently large  $c_4 > 0$ , which also ensures that  $c_3 |y_{j_2} - y_{j_1}| - 1 > 0$ . For the chosen value of  $\epsilon$ , we get  $S_2 = O \left( \frac{1}{\sqrt{M} d_{jk}^4} \right) + O \left( \frac{1}{M d_{jk}^3} \log M \right) = O \left( \frac{1}{M d_{jk}^3} \log M \right)$ . As a result, we get

$$\mathbb{E}(\text{Tr}((FF^\dagger)^2)) \leq 2 \frac{M}{d_{jk}^3} + M^2 S_2 = O \left( M \frac{\log M}{d_{jk}^3} \right). \tag{2.24}$$

Now, we generalize our result to any moment  $\ell > 2$ . We start with the following lemma.

**Lemma 2.A.2.** For  $\ell \geq 1$  and  $0 \leq i \leq 2\ell - 2$ , let

$$S_\ell^{(i)} = \left| \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* \cdots f_{j_\ell k_\ell} f_{j_1 k_\ell}^*) \right|, \tag{2.25}$$

with  $i$  ‘‘equality’’s, where by ‘‘equality’’ we mean an index is equal to another index. For example, if  $i = 0$ , then  $j_1 \neq \dots \neq j_\ell$  and  $k_1 \neq \dots \neq k_\ell$ . For all  $0 \leq i \leq 2\ell - 2$  and any  $S_\ell^{(i)}$  out of the  $\sum_{i_1=0}^i \binom{\ell}{i_1} \binom{\ell}{i-i_1}$  possible ones, we have

$$\sum_{i=0}^{2\ell-2} (\sqrt{M})^{2\ell-i} S_\ell^{(i)} = O \left( \sum_{i=0}^{\ell-1} \sqrt{M} \tilde{A}^{\ell-i-1} \tilde{B}^{i+1} + \frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}} \right),$$

where  $\tilde{A}$  and  $\tilde{B}$  are defined as in (2.23) and (2.18), respectively.

*Proof of Lemma 2.A.2.* In general, for any  $\ell \geq 1$ , we have

$$\begin{aligned}
S_\ell^{(0)} &= \left| \frac{1}{M^\ell} \int dj_1 \int dj_2 \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \int dj_3 \int dk_2 \frac{e^{2\pi i g_{j_2 j_3}(k_2)}}{\rho_{j_2 j_3}(k_2)} \cdots \right. \\
&\quad \left. \int dj_\ell \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1} j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1} j_\ell}(k_{\ell-1})} \int dk_\ell \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right| \\
&\leq \frac{1}{M^\ell} \int dj_1 \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \int dj_3 \left| \int dk_2 \frac{e^{2\pi i g_{j_2 j_3}(k_2)}}{\rho_{j_2 j_3}(k_2)} \right| \cdots \\
&\quad \int dj_\ell \left| \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1} j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1} j_\ell}(k_{\ell-1})} \right| \cdot \left| \int dk_\ell \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right| \\
&= \frac{1}{M^\ell} \int dj_1 A_{1,2} \cdot A_{2,3} \cdots A_{\ell-1,\ell} \cdot B_{\ell,1}
\end{aligned}$$

where (just as we defined  $A_{1,2}$  and  $B_{2,1}$ )

$$A_{t-1,t} = \int dj_t \left| \int dk_{t-1} \frac{e^{2\pi i g_{j_{t-1} j_t}(k_{t-1})}}{\rho_{j_{t-1} j_t}(k_{t-1})} \right| \quad \text{for } 2 \leq t \leq \ell$$

and

$$B_{\ell,1} = \left| \int dk_\ell \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right|.$$

Similarly to how we proceeded with  $A_{1,2}$  and  $B_{2,1}$  in (2.23) and (2.18), respectively, we can upper bound  $A_{t,t+1} \leq \tilde{A}$  (for  $2 \leq t \leq \ell$ ) and  $B_{\ell,1} \leq \tilde{B}$ . Therefore, we get

$$\begin{aligned}
S_\ell^{(0)} &\leq \frac{1}{M^\ell} \int dj_1 A_{1,2} \cdot A_{2,3} \cdots A_{\ell-1,\ell} \cdot B_{\ell,1} \\
&\leq \frac{1}{M^\ell} \int dj_1 \tilde{A}^{\ell-1} \tilde{B} = \frac{\tilde{A}^{\ell-1} \tilde{B}}{M^{\ell-1/2}}. \tag{2.26}
\end{aligned}$$

To generalize this result to any  $0 \leq i \leq \ell - 1$ , we use the following observation. Assume the first “equality” is given by  $k_m = k_p$ , where  $1 \leq m < p \leq \ell - 1$ . This means instead of having the term

$$A_{p,p+1} = \int dj_{p+1} \left| \int dk_p \frac{e^{2\pi i g_{j_p j_{p+1}}(k_p)}}{\rho_{j_p j_{p+1}}(k_p)} \right| \leq \tilde{A}$$

we have

$$\int dj_{p+1} \left| \frac{e^{2\pi i g_{j_p j_{p+1}}(k_m)}}{\rho_{j_p j_{p+1}}(k_m)} \right| \leq \frac{\sqrt{M}}{d_{jk}^2} = \tilde{B}.$$

Therefore, for  $1 \leq m < p \leq \ell - 1$ , we have

$$\begin{aligned}
S_\ell(k_m = k_p) &\leq \frac{1}{M^{\ell-1/2}} \int dj_1 A_{1,2} \cdots A_{p-1,p} \cdot \tilde{B} \cdot A_{p+1,p+2} \cdots A_{\ell-1,\ell} \cdot B_{\ell,1} \\
&\leq \frac{1}{M^{\ell-1/2}} \int dj_1 \tilde{A}^{\ell-2} \tilde{B}^2 = \frac{\tilde{A}^{\ell-2} \tilde{B}^2}{M^{\ell-1}}.
\end{aligned}$$

The only case remaining for the first “equality” is  $k_m = k_\ell$ , where  $1 \leq m < \ell$ . In this case, the term

$$A_{\ell-1,\ell} \cdot B_{\ell,1} = \int dj_\ell \left| \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1},j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1},j_\ell}(k_{\ell-1})} \right| \cdot \left| \int dk_\ell \frac{e^{2\pi i g_{j_\ell,j_1}(k_\ell)}}{\rho_{j_\ell,j_1}(k_\ell)} \right| \leq \tilde{A}\tilde{B}$$

is replaced by

$$\begin{aligned} & \int dj_\ell \left| \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1},j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1},j_\ell}(k_{\ell-1})} \right| \cdot \left| \frac{e^{2\pi i g_{j_\ell,j_1}(k_m)}}{\rho_{j_\ell,j_1}(k_m)} \right| \\ & \leq \int dj_\ell \int dk_{\ell-1} \left| \frac{e^{2\pi i g_{j_{\ell-1},j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1},j_\ell}(k_{\ell-1})} \right| \cdot \left| \frac{e^{2\pi i g_{j_\ell,j_1}(k_m)}}{\rho_{j_\ell,j_1}(k_m)} \right| \\ & \leq \frac{M}{d_{jk}^4} = \tilde{B}^2, \end{aligned}$$

which results in the same upper bound on  $S_\ell(k_m = k_\ell)$  as before. As such,

$$S_\ell^{(1)} = O\left(\frac{\tilde{A}^{\ell-2}\tilde{B}^2}{M^{\ell-1}}\right). \quad (2.27)$$

For the second “equality”, without loss of generality, assume  $j_m = j_p$ , where  $1 \leq m < p \leq \ell$ . If index  $k_{p-1}$  still exists (did not vanish due to the first “equality”), then instead of having the term

$$A_{p-1,p} = \int dj_p \left| \int dk_{p-1} \frac{e^{2\pi i g_{j_{p-1},j_p}(k_{p-1})}}{\rho_{j_{p-1},j_p}(k_{p-1})} \right| \leq \tilde{A}$$

we have

$$\left| \int dk_{p-1} \frac{e^{2\pi i g_{j_{p-1},j_m}(k_{p-1})}}{\rho_{j_{p-1},j_m}(k_{p-1})} \right| \leq \int dk_{p-1} \left| \frac{e^{2\pi i g_{j_{p-1},j_m}(k_{p-1})}}{\rho_{j_{p-1},j_m}(k_{p-1})} \right| \leq \frac{\sqrt{M}}{d_{jk}^2} = \tilde{B}.$$

Therefore,

$$S_\ell(j_m = j_p, k_u = k_v, p \neq v+1) \leq \frac{1}{M^{\ell-1}} \int dj_1 \tilde{A}^{\ell-3} \tilde{B}^3 = \frac{\tilde{A}^{\ell-3} \tilde{B}^3}{M^{\ell-3/2}}.$$

Note that if the index  $k_{p-1}$  vanished due to the first “equality”, then having  $j_m = j_p$  as the second “equality” results in

$$S_\ell(j_m = j_p, k_u = k_{p-1}) \leq \frac{\tilde{A}^{\ell-2} \tilde{B}^2}{M^{\ell-1}}.$$

As such, we get

$$S_\ell^{(2)} = O\left(\frac{\tilde{A}^{\ell-3} \tilde{B}^3}{M^{\ell-3/2}} + \frac{\tilde{A}^{\ell-2} \tilde{B}^2}{M^{\ell-1}}\right). \quad (2.28)$$



Note that the second term in (3.24) can be ignored, since we know from the upper bound on  $S_\ell^{(1)}$  that

$$(\sqrt{M})^{2\ell-2} \frac{\tilde{A}^{\ell-2} \tilde{B}^2}{M^{\ell-1}} \leq (\sqrt{M})^{2\ell-1} \frac{\tilde{A}^{\ell-2} \tilde{B}^2}{M^{\ell-1}}.$$

Combining (2.26), (3.23) and (3.24), we have

$$\sum_{i=0}^2 (\sqrt{M})^{2\ell-i} S_\ell^{(i)} = O\left(\sum_{i=0}^2 \sqrt{M} \tilde{A}^{\ell-i-1} \tilde{B}^{i+1}\right).$$

Note that every time we add a new “equality”, we obtain exactly one *new* term that results from replacing one  $\tilde{A}$  term by one  $\tilde{B}$  term. As such, covering all the possible less than  $\ell$  number of “equality”s gives

$$\sum_{i=0}^{\ell-1} (\sqrt{M})^{2\ell-i} S_\ell^{(i)} = O\left(\sum_{i=0}^{\ell-1} \sqrt{M} \tilde{A}^{\ell-i-1} \tilde{B}^{i+1}\right).$$

For  $\ell \leq i \leq 2\ell - 2$ , we have the following trivial bound on  $S_\ell^{(i)}$  (with any  $i$  “equality”s),

$$\begin{aligned} S_\ell^{(i)} &= |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* \cdots f_{j_\ell k_\ell} f_{j_1 k_\ell}^*)| \\ &\leq \mathbb{E}(|f_{j_1 k_1} f_{j_2 k_1}^* \cdots f_{j_\ell k_\ell} f_{j_1 k_\ell}^*|) \leq \frac{1}{d_{jk}^{2\ell}}. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \sum_{i=\ell}^{2\ell-2} (\sqrt{M})^{2\ell-i} S_\ell^{(i)} &\leq \sum_{i=\ell}^{2\ell-2} \frac{(\sqrt{M})^{2\ell-i}}{d_{jk}^{2\ell}} \\ &\leq (\ell-1) \frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}} = O\left(\frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}}\right), \end{aligned}$$

which concludes the proof of the Lemma.  $\square$

We are now set out to prove:

$$\mathbb{E}(\text{Tr}((FF^\dagger)^\ell)) = O\left(\frac{M(\log M)^{\ell-1}}{d_{jk}^{\ell+1}}\right).$$

We can write  $\mathbb{E}(\text{Tr}((FF^\dagger)^\ell))$  as follows

$$\begin{aligned} \mathbb{E}(\text{Tr}((FF^\dagger)^\ell)) &\leq \sum_{i=0}^{2\ell-2} \sum_{i_1=0}^i \binom{\ell}{i_1} \binom{\ell}{i-i_1} (\sqrt{M})^{2\ell-i} \max_{\binom{\ell}{i_1} \binom{\ell}{i-i_1} \text{ cases}} \{S_\ell^{(i)}\} \\ &\stackrel{(a)}{=} O\left(\sum_{i=0}^{\ell-1} \sqrt{M} \tilde{A}^{\ell-i-1} \tilde{B}^{i+1} + \frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}}\right) \\ &= O\left(\sqrt{M} \tilde{A}^\ell \sum_{i=0}^{\ell-1} \left(\frac{\tilde{B}}{\tilde{A}}\right)^{i+1} + \frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}}\right), \end{aligned}$$

where (a) follows from Lemma 2.A.2. Further note that since  $d_{jk} \geq \sqrt{M}$ ,

$$\tilde{A} = \frac{\log M}{d_{jk}} \geq \tilde{B} = \frac{\sqrt{M}}{d_{jk}^2}.$$

Therefore,  $\frac{\tilde{B}}{\tilde{A}} \leq 1$ , which means

$$\begin{aligned} \mathbb{E}(\text{Tr}((FF^\dagger)^\ell)) &= O\left(\sqrt{M} \tilde{A}^\ell \left(\frac{\tilde{B}}{\tilde{A}}\right) + \frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}}\right) \\ &= O\left(\frac{M(\log M)^{\ell-1}}{d_{jk}^{\ell+1}} + \frac{(\sqrt{M})^\ell}{d_{jk}^{2\ell}}\right) \\ &\stackrel{(a)}{=} O\left(\frac{M(\log M)^{\ell-1}}{d_{jk}^{\ell+1}}\right), \end{aligned}$$

where (a) follows from the fact that  $d_{jk} \geq \sqrt{M}$ . The last step, which concludes the proof, includes applying Markov's inequality to get

$$\begin{aligned} \mathbb{P}\left(\lambda_{\max}(\hat{H}_{jk} \hat{H}_{jk}^\dagger) \geq c' \frac{M^\epsilon}{d_{jk}}\right) &\leq \frac{\mathbb{E}((\lambda_{\max}(\hat{H}_{jk} \hat{H}_{jk}^\dagger))^\ell)}{(c' M^\epsilon / d_{jk})^\ell} \\ &\leq \frac{\mathbb{E}(\text{Tr}((FF^\dagger)^\ell))}{(c' M^\epsilon / d_{jk})^\ell} \\ &\leq \frac{M (c \log M)^{\ell-1} / d_{jk}^{\ell+1}}{(c' M^\epsilon / d_{jk})^\ell} \\ &\leq \frac{M (\log M)^{\ell-1}}{d_{jk} M^{\epsilon\ell}} \end{aligned}$$

which, for any fixed  $\epsilon > 0$ , can be made arbitrarily small by taking  $\ell$  sufficiently large. Proving (2.12) concludes the proof of the Lemma.

A last remark is that we proved Lemma 2.3.4 for aligned clusters. However, the proof can be easily generalized to tilted clusters, as shown in Fig. 2.5. We can always draw a larger cluster containing the original cluster and having the

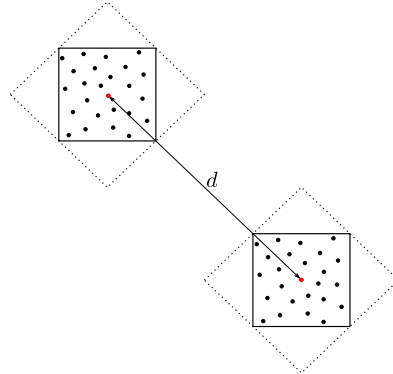


Figure 2.5 – Two tilted square clusters that have a center-to-center distance  $d$ . We can draw larger squares (drawn in dotted line) containing the original clusters with the same centers that are aligned.

same center. The larger cluster can at most contain twice as many nodes as the original cluster. The large clusters are now aligned. Moreover, the distance  $d$  from the centers of the two newly created large clusters still satisfies the required condition ( $2\sqrt{M} \leq d \leq M$ ).  $\square$



---

## Communication Tradeoffs in Wireless Networks Under Different Regimes

---

# 3

The first contribution in this chapter can be seen as a generalization of Theorem 2.4.1. While in Chapter 2 we assume a constant density independent of the number of nodes, in this chapter we study *all* sparsity regimes; ranging from very dense networks (density of  $n$ ) to very sparse networks (density  $\leq 1/n$ ). Obviously the constant density regime falls as a special case in the spectrum of regimes we study in this chapter. However, with all the technical details involved in characterizing the broadcast capacity under different sparsity assumptions, formally stated in Theorem 3.2.1 below, it was rather important, for the ease of exposition, to start by characterizing the broadcast capacity of networks with constant density. Consequently, the tools we use in this chapter are similar to those used in the previous chapter.

While the ingredients to prove Theorem 3.2.1 remain the same as those used to prove Theorem 2.4.1, new significant challenges are posed in the mathematical analysis of the maximum achievable broadcast rate and the study of the spectral norm of the channel matrix  $H$  when we assume different sparsity regimes. For instance, the recursive approach used to upper bound the norm of the channel matrix  $H$  in the case of density 1 breaks down, as we will see, when we assume the case where the density is proportional or inversely proportional to the number of nodes available in the network.

Furthermore, it is worth discussing, or at least justifying, the significance of analyzing networks with different sparsity, as it directly serves the main purpose of this chapter, as its title suggests, to study and analyze the communication tradeoffs in wireless networks. Wireless networks consist of users willing to communicate data to other users. While the purpose of communicating information remains the same, the properties of the wireless networks differ from one setting to another. While some wireless networks can be highly loaded others can be highly sparse. For example, cellular networks in urban ar-

eas versus cellular networks in rural areas, campus wireless local area network (LAN) during working days versus campus wireless LAN during vacations.

We know from [17] and [46] that under LOS model the spatial diversity, equivalently the number of degrees of freedom, of a given wireless network is proportional to the square root of its area and inversely proportional to the carrier wavelength. Consequently, in this chapter we show that for a given number of nodes  $n$  (resp. for a given area  $A$ ) decreasing the network area  $A$  (resp. increasing the number of nodes  $n$ ) increases the correlation among the channels which improves the beamforming capabilities in the network thus increasing the broadcast capacity. On the other hand, as shown in [45] and [46], multi-unicast communication is bounded by the spatial degrees of freedom present in the network. In other words, while degrees of freedom in the network are proportional to  $\sqrt{A}/\lambda$ , the beamforming gain, represented by the spectral norm of the channel matrix, is inversely proportional to  $\sqrt{A}/\lambda$ . This conflict of interest in terms of the network density between broadcast capacity and multi-unicast capacity constitutes the main contribution of this chapter. Finally, it is worth mentioning that if the cut-set bound on the multi-unicast capacity is dictated by the power transfer across the network, which corresponds to the low-SNR regime, rather than spatial degrees of freedom, then the communication scheme must exploit the beamforming capabilities together with the degrees of freedom available in the network, if possible.

As in Chapter 2, our analysis relies on the simplistic line-of-sight fading model for signal attenuation over distance, where signal amplitude attenuation is inversely proportional to distance and phase shifts are also proportional to distance. Yet, this model, along with another parameter characterizing the *sparsity* of the network, allows to capture the different regimes mentioned above and to characterize the performance tradeoffs. In addition, we would like to highlight here that despite the simplicity of the model, the mathematical analysis needed to establish the result on the maximum achievable broadcast rate in the network requires a precise and careful study of the spectral norm of unconventional random matrices, rarely studied in the mathematical literature.

### 3.1 Model

The system model is the same as that in Section 2.1, except for the area of the network. There are  $n$  nodes uniformly and independently distributed in a square network of area  $A = n^\nu$ ,  $\nu > 0$ . Each node has a different message to broadcast to the whole network, and all nodes want to communicate at a common *per user* data rate  $r_n$  bits/s/Hz. As in Chapter 2, we denote by  $R_n = nr_n$  the resulting *aggregate* data rate and refer to it as “broadcast rate”. The broadcast capacity of the network, denoted as  $C_n$ , is defined as the maximum achievable aggregate data rate  $R_n$ . The communication takes place over a flat fading channel with bandwidth  $W$  and that the signal  $Y_j[m]$  received by the

$j$ -th node at time  $m$  is given by

$$Y_j[m] = \sum_{k \in \mathcal{T}} h_{jk} X_k[m] + Z_j[m],$$

where  $\mathcal{T}$  is the set of transmitting nodes,  $X_k[m]$  is the signal sent at time  $m$  by node  $k$  and  $Z_j[m]$  is additive white circularly symmetric Gaussian noise (AWGN) of power spectral density  $N_0/2$  Watts/Hz. We also assume a common average power budget per node of  $P$  Watts, which implies that the signal  $X_k$  sent by node  $k$  is subject to an average power constraint  $\mathbb{E}(|X_k|^2) \leq P$ . In line-of-sight environment, the complex baseband-equivalent channel gain  $h_{jk}$  between transmit node  $k$  and receive node  $j$  is given by

$$h_{jk} = \sqrt{G} \frac{\exp(2\pi i r_{jk}/\lambda)}{r_{jk}}, \quad (3.1)$$

where  $G$  is Friis' constant,  $\lambda$  is the carrier wavelength, and  $r_{jk}$  is the distance between node  $k$  and node  $j$ . Let us finally define

$$\text{SNR}_s = \frac{GP}{N_0 W} n^{1-\nu},$$

which is the SNR available for a communication between two nodes at distance  $n^{\frac{\nu-1}{2}}$ ; the typical distance between two neighboring nodes in a network of area  $A = n^\nu$ .

In order to simplify notation, we choose new measurement units such that  $\lambda = 1$  and  $G/(N_0 W) = 1$  in these units. This allows us to write in particular that  $\text{SNR}_s = n^{1-\nu} P$ .

### 3.2 Main Result

We first present a known result on the multiple-unicast scenario [46]. In this case, the aggregated network throughput scales as<sup>1</sup>

$$T_n \sim \begin{cases} n & \text{if } A/\lambda^2 \geq n^2 \\ \sqrt{A}/\lambda & \text{if } n \leq A/\lambda^2 \leq n^2 \\ \sqrt{n} & \text{if } 1 \leq A/\lambda^2 \leq n \end{cases}$$

at high-SNR regime. We give the aggregate multi-unicast throughput at **high SNR**, because it is dictated by the spatial **degrees of freedom** available in the network. Such an aggregate throughput is achieved by a hierarchical cooperative strategy involving network-wide distributed MIMO transmissions in the first two cases, while a simple multihopping strategy achieves the performance claimed in the third regime.

A totally different scenario awaits us in the broadcast case. Our main result can be summarized in the following Theorem.

---

<sup>1</sup>up to logarithmic factors

**Theorem 3.2.1 (Claim  $(\star)$ ).** *The aggregate broadcast capacity scales as*

$$C_n \sim \begin{cases} \min\{\text{SNR}_s, 1\} & \text{if } (A/\lambda^2) \geq n^2 \\ \min\left\{\left(\frac{n}{\sqrt{A}/\lambda}\right) \text{SNR}_s, 1\right\} & \text{if } 1 \leq (A/\lambda^2) \leq n^2 \end{cases}$$

*up to a multiplicative factor  $n^\epsilon$ , for any fixed  $\epsilon > 0$ .*

$(\star)$  The above theorem relies on Claim 3.A.2 in Lemma 3.A.1. Since other lemmas/proposition (Specifically, Proposition 3.3.5, Lemma 3.3.6, and Lemma 3.A.1) presented in this chapter rely on the very same claim, we will denote them by  $(\star)$  in what follows.

The broadcast capacity characterized in the theorem above is achieved<sup>2</sup> by multihopping or simple time division based broadcast transmission in the first case and by a multi-stage back-and-forth beamforming strategy in the second case. The performance is further capped at 1, which means that such **beamforming gains** can only be obtained at **low SNR**.

We see here that no particular beamforming gain can be obtained for a sparse network of density  $O(1/n)$  (regime where  $A \sim n^2$ ). The beamforming gain starts appearing as the density goes above  $1/n$  and it continues increasing as the network gets denser. In the previous chapter, we already proved the result when the network is of constant density ( $A \sim n$ ).

A final observation shows the duality of the two previous results: in the regime where  $A/\lambda^2 \geq n$  (that is, for networks of constant density or sparser), we have  $\mathbf{DoF} = \min\{n, \sqrt{A}/\lambda\}$  spatial degrees of freedom, while the beamforming gain is  $\mathbf{BG} = \max\left\{1, \frac{n}{\sqrt{A}/\lambda}\right\}$ . This means that, for  $A/\lambda^2 \geq n$ , we have

$$\mathbf{DoF} \times \mathbf{BG} = n, \tag{3.2}$$

which represents the tradeoff between the degrees of freedom available and the beamforming gain achievable in the network. The relation we have in (3.2) captures the fact that high beamforming gains can only be obtained at the expense of a reduced number of degrees of freedom (or reciprocally).

Another interesting observation is that at low SNR, for any  $\epsilon > 0$ , the multiple-unicast capacity is upper bounded by  $n^{1+\epsilon}\text{SNR}_s$ , which represents the total power transfer across the network. In other words, at low SNR, the multi-unicast capacity is equal to  $(\mathbf{DoF} \times \mathbf{BG}) \times \text{SNR}_s$  up to a multiplicative factor  $n^\epsilon$ , for any  $\epsilon > 0$ . As such, to achieve the multi-unicast capacity in the low-SNR regime; equivalently, to ensure an optimal power transfer in the network, the wireless communication scheme deployed in the network must exploit (if possible) both the beamforming capabilities and the spatial degrees of freedom present in the network. This is an interesting open problem worth the investigation.

<sup>2</sup>up to a multiplicative factor  $n^\epsilon$ , for any fixed  $\epsilon > 0$



### 3.3 Broadcast Capacity

#### 3.3.1 At High SNR

Under LOS model (3.1) and the assumptions made in the Section 3.1, a simple time division scheme achieves a broadcast (aggregate) rate  $R_n$  of order  $\min(\text{SNR}_s, 1)$ . Indeed, a rate of order 1 is obviously achieved at high SNR<sup>3</sup>.

At high SNR, when  $\text{SNR}_s = \Omega(1)$ , each node is capable to broadcast to the entire network every  $n$  time slots. In other words, since the nodes take turns, then each node will remain silent for  $n$  time slots during which it can accumulate a power of  $\Omega(nP)$ . Moreover, since any two nodes are at most separated by a distance of  $\sqrt{2n^\nu}$ , then the received power at every node is

$$\Omega\left(\frac{nP}{n^\nu}\right) = \Omega(\text{SNR}_s) = \Omega(1).$$

#### 3.3.2 At Low SNR

As in Chapter 2, our focus is on low-SNR regime,  $\text{SNR}_s = n^{-\gamma}$  for some constant  $\gamma > 0$ , which means that the power available at each node is not enough to establish a constant rate direct communication with a neighbor. This could be the case e.g., in a sensor network with low battery nodes, or in a sparse network (large  $\nu$ ) with long distances between neighboring nodes.

At low SNR, each node can spare power while the others are transmitting, so as to compensate for the path loss of order  $1/n^\nu$  between the source node and other nodes located at distance at most  $\sqrt{2n^\nu}$ , leading to a broadcast rate of order  $R_n \sim \log(1 + nP/n^\nu) \sim n^{1-\nu}P = \text{SNR}_s$ .

In the following, we will see that, at low SNR, while the described simple TDMA based broadcast scheme is order-optimal<sup>4</sup> for networks of area  $A \geq n^2$ , it is not optimal for networks with area  $A < n^2$  ( $\nu < 2$ ) (for simplicity, as stated in Section 3.1, we take  $\lambda = 1$ ). On the other hand, the back-and-forth beamforming scheme, presented in the previous chapter, proves to be order-optimal<sup>4</sup> for  $A < n^2$ .

#### Broadcasting Schemes

As described in Chapter 2, the back-and-forth beamforming scheme involves source nodes taking turns to broadcast their messages. Each transmission is followed by a series of network-wide back-and-forth transmissions that reinforce the strength of the signal, so that at the end, every node is able to decode the message sent from the source.

The reason why back-and-forth transmissions are useful for **small area networks/dense networks** is that in line-of-sight environment, nodes are able to (partly) align the transmitted signals so as to create a significant beamforming gain for each transmission (whereas this would not be the case in **high**

<sup>3</sup>We coarsely approximate  $\log P$  by 1 here!

<sup>4</sup>up to a multiplicative factor  $n^\epsilon$ , for any  $\epsilon > 0$

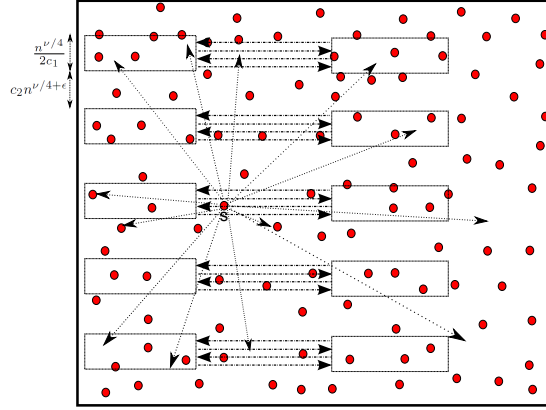


Figure 3.1 –  $\sqrt{n} \times \sqrt{n}$  network divided into clusters of size  $M = \frac{n^{\nu/4}}{2c_1} \times \frac{n^{\nu/2}}{4}$ . Two clusters of size  $M$  placed on the same horizontal line and separated by distance  $d = \frac{n^{\nu/2}}{4}$  pair up and start back-and-forth beamforming. The vertical separation between adjacent cluster pairs is  $c_2 n^{\nu/4 + \epsilon}$ .

**scattering environment/sparse networks** with i.i.d. fading coefficients). In short, the back-and-forth beamforming scheme is split into two phases:

**Phase 1. Broadcast Transmission.** The source node broadcasts its message to the whole network. A noisy version of the signal, which remains undecoded, is received at the nodes. Only one time slot is required for this phase.

**Phase 2. Back-and-Forth Beamforming with Time Division.** Upon receiving the signal from the broadcasting node, nodes start multiple back-and-forth beamforming transmissions between the two halves of the network. Clusters of size  $M = \frac{n^{\nu/4}}{2c_1} \times \frac{n^{\nu/2}}{4}$  and separated by horizontal distance  $d = \frac{n^{\nu/2}}{4}$  pair up for the back-and-forth transmissions. During each transmission, there are  $\Theta(n^{\nu/4 - \epsilon})$  cluster pairs operating simultaneously, so in total there are  $\Theta(n^{1 - \epsilon})$  nodes communicating. As such, to serve all nodes,  $\Theta(n^\epsilon)$  time rounds are required.

Upon each transmission cycle, a node in a given cluster receives a signal that is the sum of the signals coming from the facing cluster, of those coming from simultaneously operating clusters, and of the noise. For this reason, we introduce a sufficiently large vertical distance  $c_2 n^{\nu/4 + \epsilon}$  separating any two adjacent cluster pairs. We show below that the broadcast rate between the operating clusters is  $\Theta(n^{2 - \frac{3\nu}{2}} P) = \Theta(n^{1 - \frac{\nu}{2}} \text{SNR}_s)$ . Since we only need  $\Theta(n^\epsilon)$  number of rounds to serve all clusters, phase 2 requires  $\Theta(n^{-2 + \frac{3\nu}{2} + \epsilon} P^{-1})$  time slots. As such, back-and-forth beamforming achieves a broadcast rate of  $\Theta(n^{2 - \frac{3\nu}{2} - \epsilon} P) = \Theta(n^{1 - \frac{\nu}{2} - \epsilon} \text{SNR}_s)$  bits per time slot. Now that we have described the broadcasting scheme we use for networks of different sparsity, which is identical to the scheme described in the previous chapter for extended networks, we are able to state the following Theorem.

**Theorem 3.3.1.** For any  $\epsilon > 0$ ,  $0 < \nu < 2$ , and  $P = O(n^{-2+\frac{3\nu}{2}})$ , the following broadcast rate

$$R_n = \Omega\left(n^{2-\frac{3\nu}{2}-\epsilon}P\right) = \Omega\left(\frac{n}{\sqrt{A}}n^{-\epsilon}\text{SNR}_s\right)$$

is achievable with high probability<sup>5</sup> in the network. As a consequence, when  $P = \Omega(n^{-2+\frac{3\nu}{2}})$ , a broadcast rate  $R_n = \Omega(n^{-\epsilon})$  is achievable with high probability.

The proof of the above theorem proceeds in exactly the same fashion as that of Theorem 2.2.1. A direct extension of lemma 2.2.2 is given in the following lemma, which provides an upper bound on the probability that the number of nodes inside each cluster deviates from its mean by a large factor. For the sake of completeness, the proof is provided in the Appendix 3.A.

**Lemma 3.3.2.** Let us consider a cluster of area  $M$  with  $M = n^\beta$  for some  $\nu - 1 < \beta < \nu$ . The number of nodes inside each cluster is then between  $((1 - \delta)Mn^{1-\nu}, (1 + \delta)Mn^{1-\nu})$  with probability larger than  $1 - \frac{n^\nu}{M} \exp(-\Delta(\delta)Mn^{1-\nu})$  where  $\Delta(\delta)$  is independent of  $n$  and satisfies  $\Delta(\delta) > 0$  for  $\delta > 0$ .

As shown in Fig. 3.1, two clusters of size  $M = \frac{n^{\nu/4}}{2c_1} \times \frac{n^{\nu/2}}{4}$  placed on the same horizontal line and separated by distance  $d = \frac{n^{\nu/2}}{4}$  form a cluster pair. During the back-and-forth beamforming phase, there are many cluster pairs operating simultaneously. Given that the cluster width is  $\frac{n^{\nu/4}}{2c_1}$  and the vertical separation between adjacent cluster pairs is  $c_2n^{\nu/4+\epsilon}$ , there are

$$N_C = \frac{n^{\nu/2}}{\frac{n^{\nu/4}}{2c_1} + c_2n^{\nu/4+\epsilon}} = \Theta\left(n^{\nu/4-\epsilon}\right)$$

cluster pairs operating at the same time. Let  $\mathcal{R}_i$  and  $\mathcal{T}_i$  denote the receiving and the transmitting clusters of the  $i$ -th cluster pair, respectively.

Lemma 3.3.3 and Lemma 3.3.4 are the two main ingredients for analyzing the multi-stage back-and-forth beamforming scheme. While the proof of Lemma 2.2.3 can be trivially extended to prove Lemma 3.3.3, the proof of Lemma 2.2.4 requires additional effort to tailor it for Lemma 3.3.4. The proofs are presented in the Appendix 3.A.

**Lemma 3.3.3.** The maximum beamforming gain between the two clusters of the  $i$ -th cluster pair can be achieved by using a compensation of the phase shifts at the transmit side which is proportional to the horizontal positions of the nodes. More precisely, there exist a constant  $c_1 > 0$  (remember that  $c_1$  is inversely proportional to the width of  $i$ -th cluster) and a constant  $K_1 > 0$  such that the magnitude of the received signal at node  $j \in \mathcal{R}_i$  is lower bounded with high probability by

$$\left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \geq K_1 \frac{Mn^{1-\nu}}{d},$$

<sup>5</sup>that is, with probability at least  $1 - O\left(\frac{1}{n^p}\right)$  as  $n \rightarrow \infty$ , where the exponent  $p$  is arbitrary.

where  $x_k$  denotes the horizontal position of node  $k$ .

**Lemma 3.3.4.** *For every constant  $K_2 > 0$ , there exists a sufficiently large separating constant  $c_2 > 0$  such that the magnitude of interfering signals from the simultaneously operating cluster pairs at node  $j \in \mathcal{R}_i$  is upper bounded with high probability by*

$$\left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \leq K_2 \frac{Mn^{1-\nu}}{dn^\epsilon} \log n.$$

*Proof of Theorem 3.3.1.* As you will notice, here we follow exactly the same steps used to prove Theorem 2.2.1.

The first phase of the scheme results in noisy observations of the message  $X$  at all nodes, which are given by

$$Y_k^{(0)} = \sqrt{\text{SNR}_k} X + Z_k^{(0)},$$

where  $\mathbb{E}(|X|^2) = \mathbb{E}(|Z_k^{(0)}|^2) = 1$  and  $\text{SNR}_k$  is the signal-to-noise ratio of the signal  $Y_k^{(0)}$  received at the  $k$ -th node. In what follows, we drop the index  $k$  from  $\text{SNR}_k$  and only write  $\text{SNR} = \min_k \{\text{SNR}_k\}$ . Note that it does not make a difference at which side of the cluster pairs the back-and-forth beamforming starts or ends. Hence, assume the left-hand side clusters ignite the scheme by amplifying and forwarding the noisy observations of  $X$  to the right-hand side clusters. The signal received at node  $j \in \mathcal{R}_i$  is given by

$$Y_j^{(1)} = \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} A Y_k^{(0)} + Z_j^{(1)} \quad (3.3)$$

where  $A$  is the amplification factor (to be calculated later) and  $Z_j^{(1)}$  is additive white Gaussian noise of variance  $\Theta(1)$ . We start by applying Lemma 3.3.3 and Lemma 3.3.4 to lower bound

$$\begin{aligned} \left| \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &\geq \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\quad - \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\ &\geq \left( K_1 - K_2 \frac{\log n}{n^\epsilon} \right) \frac{Mn^{1-\nu}}{d} = \Theta \left( \frac{Mn^{1-\nu}}{d} \right). \end{aligned}$$

For the sake of clarity, we can therefore approximate<sup>6</sup> the expression in (3.3) as follows

$$\begin{aligned}
Y_j^{(1)} &= \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} A \sqrt{\text{SNR}_k} X \\
&+ \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} A Z_k^{(0)} + Z_j^{(1)} \\
&\simeq \frac{A M n^{1-\nu}}{d} \sqrt{\text{SNR}} X + \frac{A \sqrt{N_C M n^{1-\nu}}}{d} Z^{(0)} + Z_j^{(1)} \\
&= \frac{A M n^{1-\nu}}{d} \sqrt{\text{SNR}} X + \frac{A M n^{1-\nu}}{d} \sqrt{\frac{N_C}{M n^{1-\nu}}} Z^{(0)} + Z_j^{(1)},
\end{aligned}$$

where

$$Z^{(0)} = \frac{d}{\sqrt{N_C M n^{1-\nu}}} \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} Z_k^{(0)}.$$

Note that  $\mathbb{E}(|Z^{(0)}|^2) = \Theta(1)$ . Repeating the same process  $t$  times in a back-and-forth manner results in a final signal at node  $j \in \mathcal{R}_i$  in the left or the right cluster (depending on whether  $t$  is odd or even) that is given by

$$\begin{aligned}
Y_j^{(t)} &= \left( \frac{A M n^{1-\nu}}{d} \right)^t \sqrt{\text{SNR}} X + \left( \frac{A M n^{1-\nu}}{d} \right)^t \sqrt{\frac{N_C}{M n^{1-\nu}}} Z^{(0)} \\
&+ \dots + \left( \frac{A M n^{1-\nu}}{d} \right)^{t-s} \sqrt{\frac{N_C}{M n^{1-\nu}}} Z^{(s)} + \dots + Z_j^{(t)},
\end{aligned}$$

where for  $0 \leq s \leq t-1$ ,

$$Z^{(s)} = \frac{d}{\sqrt{N_C M n^{1-\nu}}} \sum_{b=1}^{N_C} \sum_{k \in \mathcal{T}_b} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} Z_k^{(s)}.$$

Furthermore, for  $0 \leq s \leq t-1$ ,  $\mathbb{E}(|Z^{(s)}|^2) = \Theta(1)$ , and  $Z_j^{(t)}$  is additive white Gaussian noise of variance  $\Theta(1)$ . Finally, note that Lemma 3.3.4 ensures an upper bound on the beamforming gain of the noise signals, i.e.,

$$\begin{aligned}
\left| \sum_{l=1}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &\leq \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\
&+ \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| \\
&\leq \left( 1 + K_2 \frac{\log n}{n^\epsilon} \right) \frac{M n^{1-\nu}}{d}.
\end{aligned}$$

<sup>6</sup>As in Chapter 2, we make this approximation to lighten the notation, but the whole analysis goes through without the approximation.

(notice indeed that the first term in the middle expression is trivially upper bounded by  $Mn^{1-\nu}/d$ , as it contains  $Mn^{1-\nu}$  terms, all less than  $1/d$ ). Now, we want the power of the signal to be of order 1, that is:

$$\mathbb{E} \left( \left( \left( \frac{AMn^{1-\nu}}{d} \right)^t \sqrt{\text{SNR}} X \right)^2 \right) = \left( \frac{AMn^{1-\nu}}{d} \right)^{2t} \text{SNR} = \Theta(1) \quad (3.4)$$

$$\Rightarrow A = \Theta \left( \frac{d}{Mn^{1-\nu}} \text{SNR}^{-\frac{1}{2t}} \right). \quad (3.5)$$

Since at each round of TDMA cycle there are  $\Theta(N_C Mn^{1-\nu}) = \Theta(n^{1-\epsilon})$  nodes transmitting, then every node will be active a fraction

$$f = \Theta \left( \frac{N_C Mn^{1-\nu}}{n} \right) = \Theta \left( \frac{N_C M}{n^\nu} \right)$$

of the time. Moreover, let  $\tau$  denote the number of time slots between two consecutive transmissions, i.e. every  $\tau$  time slots we have one transmission. This results in an amplification factor of

$$A = \sqrt{f \times \tau \times P} = \Theta \left( \sqrt{\frac{n^\nu}{N_C M} \tau P} \right). \quad (3.6)$$

Combining (3.5) and (3.6), we get

$$\begin{aligned} A &= \Theta \left( \frac{d}{Mn^{1-\nu}} \text{SNR}^{-\frac{1}{2t}} \right) = \Theta \left( \sqrt{\frac{n^\nu}{N_C M} \tau P} \right) \\ \Rightarrow \tau &= O \left( \frac{1}{P} \left( \frac{d}{Mn^{1-\nu}} \right)^2 n^{-\epsilon} \text{SNR}^{-1/t} \right). \end{aligned}$$

We can pick the number of back-and-forth transmissions  $t$  sufficiently large to ensure that  $\text{SNR}^{-\frac{1}{t}} = O(n^\epsilon)$ , which results in

$$\tau = O \left( \frac{1}{P} \left( \frac{d}{Mn^{1-\nu}} \right)^2 \right) = O \left( \frac{1}{n^{2-\frac{3\nu}{2}} P} \right).$$

Moreover, the noise power is given by

$$\begin{aligned} &\sum_{s=0}^{t-1} \mathbb{E} \left( \left( \left( \frac{AMn^{1-\nu}}{d} \right)^{t-s} \sqrt{\frac{N_C}{Mn^{1-\nu}}} Z^{(s)} \right)^2 \right) + \mathbb{E} \left( \left( Z_j^{(t)} \right)^2 \right) \\ &\leq t \mathbb{E} \left( \left( \left( \frac{AMn^{1-\nu}}{d} \right)^t \sqrt{\frac{N_C}{Mn^{1-\nu}}} Z^{(0)} \right)^2 \right) + 1 \\ &\leq t \left( \frac{AMn^{1-\nu}}{d} \right)^{2t} \frac{N_C}{Mn^{1-\nu}} + 1 \\ &\stackrel{(a)}{\leq} t + 1 = \Theta(1), \end{aligned}$$

where (a) holds if and only if  $\text{SNR} = \Omega\left(\frac{N_C}{Mn^{1-\nu}}\right) = \Omega(n^{\nu/2-1-\epsilon})$  (check eq. (3.4)), which is true: Distance separating any two nodes in the network is as most  $\sqrt{2n^\nu}$ , which implies that the SNR of the received signal at all the nodes in the network is  $\Omega(n\tau P/n^\nu) = \Omega(n^{\nu/2-1})$ .

Given that the required  $\tau = O\left(\frac{1}{n^{2-3\nu/2}P}\right)$ , we can see that for  $P = O(n^{3\nu/2-2})$  the broadcast rate between simultaneously operating clusters is  $\Omega(n^{2-3\nu/2}P)$ . Finally, applying TDMA of  $\frac{n}{N_C M n^{1-\nu}} = \Theta(n^\epsilon)$  steps ensures that  $X$  is successfully decoded at all nodes and the broadcast rate  $R_n = \Omega(n^{2-3\nu/2-\epsilon}P)$ . This completes the proof of the theorem.  $\square$

### Optimality of the Scheme

We start with the general upper bound already established in Theorem 2.3.1 on the broadcast capacity of wireless networks at low SNR, which applies to a general fading matrix  $H$ .

We now aim to specialize Theorem 2.3.1 to line-of-sight fading, where the matrix  $H$  is given by

$$h_{jk} = \begin{cases} 0 & \text{if } j = k \\ \frac{\exp(2\pi i r_{jk})}{r_{jk}} & \text{if } j \neq k. \end{cases} \quad (3.7)$$

The rest of the section is devoted to proving the proposition below which, together with Theorem 2.3.1, shows the asymptotic optimality of the back-and-forth beamforming scheme for **small area networks/dense networks** ( $0 < \nu < 2$ ) and the asymptotic optimality of the simple TDMA based broadcast scheme for **high scattering environment/sparse networks** ( $\nu \geq 2$ ) at low SNR and under LOS fading. The spectral norm  $\|H\|$  represents the beamforming capabilities in the network; the smaller the network gets, the denser it becomes, thus increasing the correlation between the entries of the matrix  $H$ , which results in a larger spectral norm.

**Proposition 3.3.5** ( $\star$ ). *Let  $H$  be the  $n \times n$  matrix given by (3.7). For every  $\epsilon > 0$ , there exists a constant  $c > 0$  such that*

$$\|H\|^2 \leq \begin{cases} c n^{2-\frac{3\nu}{2}+\epsilon} & \text{if } 0 < \nu < 2 \\ c n^{1-\nu+\epsilon} & \text{if } \nu \geq 2 \end{cases}$$

with high probability as  $n$  gets large.

Consequently, Theorem 2.3.1 gives the following upper bound on the broadcast capacity

$$C_n \leq P \|H\|^2 \leq \begin{cases} c n^{2-\frac{3\nu}{2}+\epsilon} P = c \frac{n}{\sqrt{A}} n^\epsilon \text{SNR}_s & \text{if } 0 < \nu < 2 \\ c n^{1-\nu+\epsilon} P = c n^\epsilon \text{SNR}_s & \text{if } \nu \geq 2, \end{cases}$$

which, together with Theorem 3.3.1, concludes the proof of Theorem 3.2.1.

For the proof of Proposition 3.3.5, analyzing directly the asymptotic behavior of  $\|H\|$  reveals itself difficult. We therefore decompose our proof into simpler subproblems. The first building block of the proof is Lemma 2.3.3, which is a generalization of the classical Geršgorin discs' inequality. The second building block of this proof is the following lemma, the proof of which is given in the Appendix 3.A.

**Lemma 3.3.6** ( $\star$ ). *Let  $\widehat{H}$  be the  $m \times m$  channel matrix between two square clusters of  $m$  nodes distributed uniformly at random, each of area  $A = m^\nu$ ,  $\nu > 0$ . Then*

$$\begin{aligned} \|\widehat{H}\|^2 &\leq \max \left\{ \frac{m^{2+\epsilon}}{Ad}, \frac{m^{1+\epsilon}}{d^2} \right\} \\ &\leq \begin{cases} \frac{m^{2+\epsilon}}{Ad} & \text{if } 0 < \nu < 2 \\ \max \left\{ \frac{m^{2+\epsilon}}{Ad}, \frac{m^{1+\epsilon}}{d^2} \right\} & \text{if } \nu \geq 2 \end{cases} \end{aligned}$$

with high probability as  $m$  gets large, where  $2\sqrt{A} \leq d \leq A$  denotes the distance between the centers of the two clusters.

*Proof of Proposition 3.3.5.* First we consider the case where  $\nu \geq 2$ . The strategy for the proof is now the following: in order to bound  $\|H\|$ , we divide the matrix into smaller blocks, apply Lemma 2.3.3 and Lemma 3.3.6 in order to bound the off-diagonal terms  $\|H_{jk}\|$ . For the diagonal terms  $\|H_{jj}\|$ , we reapply Lemma 2.3.3 and proceed in a recursive manner, until we reach small size blocks for which a loose estimate is sufficient to conclude. However, we should note that the parameter  $\nu$ , used to relate the area to the number of nodes available, increases with every new layer of the recursion. In other words, consider a cluster of area  $A_m$  that contains  $m$  nodes. We know that the density of the entire network is equal to  $n^{1-\nu}$ . Therefore, we get

$$\frac{m}{A_m} = \frac{m}{m^{\nu'}} = \frac{n}{n^\nu} \stackrel{(a)}{\leq} \frac{m}{m^\nu} \quad (3.8)$$

where (a) follows from the fact that  $\nu \geq 2 > 1$  and  $\nu'$  is the new parameter relating area  $A_m$  of the network to the number of nodes  $m$  available in the network. As a result, we get  $\nu' \geq \nu$ .

As we mentioned, a network with area  $A_0 = n^\nu$  has a density of  $n^{1-\nu}$ . This means that a cluster of area  $A_1 = m_1 n^{\nu-1}$  contains  $m_1$  nodes with high probability. Let us therefore decompose the network into  $K_1$  square clusters of area  $m_1 n^{\nu-1}$  with  $m_1$  nodes each. Without loss of generality, we assume each cluster has exactly  $m_1$  nodes and  $K_1 = n/m_1 = A_0/A_1$ . By Lemma 2.3.3, we obtain

$$\|H\| \leq \max \left\{ \max_{1 \leq j \leq K_1} \sum_{k=1}^{K_1} \|H_{jk}\|, \max_{1 \leq j \leq K_1} \sum_{k=1}^{K_1} \|H_{kj}\| \right\} \quad (3.9)$$

where the  $n \times n$  matrix  $H$  is decomposed into blocks  $H_{jk}$ ,  $j, k = 1, \dots, K_1$ , with  $H_{jk}$  denoting the  $m_1 \times m_1$  channel matrix between cluster number  $j$  and



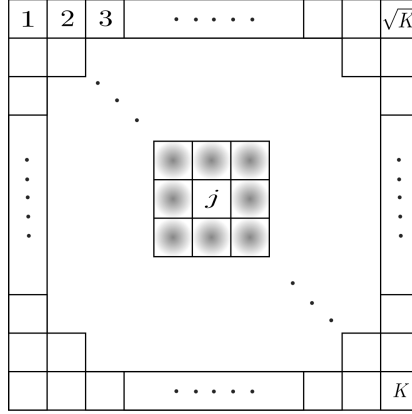


Figure 3.2 –  $n^{\nu/2} \times n^{\nu/2}$  network split into  $K$  clusters and numbered in order. As such,  $R_j = \{j - \sqrt{K} - 1, j - \sqrt{K}, j - \sqrt{K} + 1, j - 1, j, j + 1, j + \sqrt{K} - 1, j + \sqrt{K}, j + \sqrt{K} + 1\}$ , which represents the center square containing the cluster  $j$  and its 8 neighbors (marked in shades).

cluster number  $k$  in the network. Let us also denote by  $d_{jk}$  the corresponding inter-cluster distance, measured from the centers of these clusters. Based on Lemma 3.3.6, if  $d_{jk} \geq 2\sqrt{A_1}$ , we obtain

$$\|H_{jk}\|^2 \leq \max \left\{ \frac{m_1^{2+\epsilon}}{A_1 d_{jk}}, \frac{m_1^{1+\epsilon}}{d_{jk}^2} \right\} \stackrel{(a)}{=} \frac{m_1^{1+\epsilon}}{d_{jk}^2}$$

with high probability as  $m_1 \rightarrow \infty$ , where (a) follows from the fact that  $A_1/m_1 = n^{\nu-1} \geq n^{\nu/2} \geq d_{jk}$ , since  $\nu \geq 2$  (equivalently,  $\frac{m_1}{A_1} \leq \frac{1}{d_{jk}}$ ).

Let us now fix  $j \in \{1, \dots, K_1\}$  and define  $R_j = \{1 \leq k \leq K_1 : d_{jk} < 2\sqrt{A_1}\}$  and  $S_j = \{1 \leq k \leq K_1 : d_{jk} \geq 2\sqrt{A_1}\}$  (see Fig. 3.2). By the above inequality, we obtain

$$\sum_{k=1}^{K_1} \|H_{jk}\| \leq \sum_{k \in R_j} \|H_{jk}\| + \sqrt{n^\epsilon} \sum_{k \in S_j} \frac{\sqrt{m_1}}{d_{jk}}$$

with high probability as  $m_1$  gets large. Observe that as there are  $8t$  clusters or less at distance  $t\sqrt{A_1}$  from cluster  $j$ , so we obtain

$$\sum_{k \in S_j} \frac{\sqrt{m_1}}{d_{jk}} \leq \sum_{t=2}^{\sqrt{K_1}} 8t \frac{\sqrt{m_1}}{t\sqrt{A_1}} = O \left( \sqrt{\frac{K_1 m_1}{A_1}} \right).$$

There remains to upper bound the sum over  $R_j$ . Observe that this sum contains at most 9 terms: namely the term  $k = j$  and the 8 terms corresponding to the 8 neighboring clusters of cluster  $j$ . It should then be observed that for each  $k \in R_j$ ,  $\|H_{jk}\| \leq \|H(R_j)\|$ , where  $H(R_j)$  is the  $9m_1 \times 9m_1$  matrix made of the

$9 \times 9$  blocks  $H_{j_1, j_2}$  such that  $j_1, j_2 \in R_j$ . Finally, this leads to

$$\sum_{k=1}^{K_1} \|H_{jk}\| \leq 9\|H(R_j)\| + 8\sqrt{n^\epsilon} \sqrt{\frac{K_1 m_1}{A_1}}$$

Using the symmetry of this bound and (3.9), we obtain

$$\|H\| \leq 9 \max_{1 \leq j \leq K_1} \|H(R_j)\| + 8\sqrt{n^\epsilon} \sqrt{\frac{K_1 m_1}{A_1}}. \quad (3.10)$$

A key observation is now the following: For all  $1 \leq j \leq K_1$ , the  $9M \times 9M$  matrix  $H(R_j)$  has exactly the same structure as the original matrix  $H$ . Therefore, without loss of generality<sup>7</sup>, let us assume  $\|H_1\| = \max_{1 \leq j \leq K_1} \|H(R_j)\| = \|H(R_1)\|$ . Finally, to bound  $\|H_1\|$ , the same technique may be reused. Based on (3.8), we know that the parameter relating the area of the network to the number of nodes increases with the higher layers of the recursion. In other words, this parameter always remains greater than 2, which means that, at any layer of the recursion if the clusters area is  $A_m$  and it contains  $m$  nodes, then

$$\max \left\{ \frac{m^{2+\epsilon}}{A_m d_{jk}}, \frac{m^{1+\epsilon}}{d_{jk}^2} \right\} = \frac{m^{1+\epsilon}}{d_{jk}^2}.$$

This leads to the following recursive solution.

$$\begin{aligned} \|H\| &= O \left( \|H_1\| + \sqrt{n^\epsilon} \sqrt{\frac{K_1 m_1}{A_1}} \right) \\ &= O \left( \|H_2\| + \sqrt{n^\epsilon} \sqrt{\frac{K_2 m_2}{A_2}} + \sqrt{n^\epsilon} \sqrt{\frac{K_1 m_1}{A_1}} \right) \\ &= O \left( \|H_l\| + \sqrt{n^\epsilon} \sum_{t=1}^l \sqrt{\frac{K_t m_t}{A_t}} \right) \\ &= O \left( \|H_l\| + \sqrt{n^\epsilon} \sqrt{n^{1-\nu}} \sum_{t=1}^l \sqrt{K_t} \right), \end{aligned}$$

where  $m_t$  denotes the number of nodes in a square cluster of area  $A_t$ . Moreover,  $K_t = A_{t-1}/A_t = m_{t-1}/m_t$  denotes the number of square clusters of area  $A_t$  and  $m_t$  nodes in a square network of area  $A_{t-1}$  containing  $m_{t-1}$  nodes (note that  $A_0 = A = n^\nu$  and  $m_0 = n$ ). Finally,  $\|H_t\|$  denotes the norm of the channel matrix of the network with square area  $A_t$  and  $m_t$  nodes.

Note that we have a trivial bound on  $\|H_l\|$ . Apply for this the slightly modified version of the classical Geršgorin inequality (which is nothing but the statement of Lemma 2.3.3 applied to the case  $M = 1$ ):

$$\|H_l\| \leq \max \left\{ \max_{1 \leq j \leq m_l} \sum_{k=1}^{m_l} |(H_l)_{jk}|, \max_{1 \leq j \leq m_l} \sum_{k=1}^{m_l} |(H_l)_{kj}| \right\} = \max_{1 \leq j \leq m_l} \sum_{\substack{k=1 \\ k \neq j}}^{m_l} r_{jk}$$

<sup>7</sup>We present a uniform upper bound on  $\|H(R_j)\|$ .

For any  $1 \leq j \leq m_l$ , it holds with high probability that for  $c$  large enough,

$$\sum_{\substack{k=1 \\ k \neq j}}^{m_l} \frac{1}{r_{jk}} \leq \sum_{t=1}^{\sqrt{m_l}} \frac{ct \log n}{t n^{\frac{\nu-1}{2}}} = O\left(\sqrt{m_l} n^{\frac{1-\nu}{2}} \log n\right) = O\left(n^{1-\nu} \sqrt{A_l} \log n\right),$$

where the first inequality comes from the fact that at a distance  $t n^{\frac{\nu-1}{2}}$  there are at most  $ct$  clusters of area  $n^{\nu-1}$  with at most  $\log n$  nodes each. This implies that  $\|H_l\| = O\left(\sqrt{n^\epsilon} n^{1-\nu} \sqrt{A_l}\right)$  for any  $\epsilon > 0$ . Therefore, we have

$$\|H\| = O\left(\sqrt{n^\epsilon} n^{1-\nu} \sqrt{A_l} + \sqrt{n^\epsilon} \sqrt{n^{1-\nu}} \sum_{t=1}^l \sqrt{\frac{A_{t-1}}{A_t}}\right).$$

Upon optimizing over the  $A_t$ 's, we get  $A_t = n^{\nu - \frac{t}{l+1}}$ . Note that  $A_t$  is a decreasing function of  $t$  and  $A_0 = n^\nu$ . Thus, for  $\nu \geq 2$ , we get the desired result

$$\|H\| = O\left(n^{\frac{1-\nu}{2}} n^{\frac{\epsilon}{2} + \frac{1}{2(l+1)}}\right),$$

where for any  $\epsilon' > \epsilon/2$ , we can pick  $l$  large enough so that  $\frac{\epsilon}{2} + \frac{1}{2(l+1)} < \epsilon'$  (notice that  $\epsilon$  and  $\epsilon'$  can be as small as we want).

For  $0 < \nu < 2$ , we will take the following approach: We notice that a dense network can be seen as a superposition of sparse networks. In other words, we will look at a network with  $n$  nodes uniformly and independently distributed over the area  $n^\nu$ , as the superposition of  $n^{1-\nu/2}$  networks with  $m = n^{\nu/2}$  nodes uniformly and independently distributed over the area  $n^\nu = m^2$ . Again, by Lemma 2.3.3, we obtain

$$\|H\| \leq \max \left\{ \max_{1 \leq j \leq n^{1-\nu/2}} \sum_{k=1}^{n^{1-\nu/2}} \|H_{jk}\|, \max_{1 \leq j \leq n^{1-\nu/2}} \sum_{k=1}^{n^{1-\nu/2}} \|H_{kj}\| \right\}$$

where the  $n \times n$  matrix  $H$  is decomposed into blocks  $H_{jk}$ ,  $j, k = 1, \dots, n^{1-\nu/2}$ , with  $H_{jk}$  denoting the  $m \times m$  channel matrix between sparse network number  $j$  and sparse network number  $k$ . Since each of these sparse networks has area  $m^2$  with  $m$  nodes, we can apply the upper bound we got for  $\nu = 2$ , and  $\forall j, k = 1, \dots, n^{1-\nu/2}$ , obtain

$$\|H_{jk}\| = O\left(m^{-\frac{1}{2} + \epsilon}\right) = O\left(n^{-\frac{\nu}{4} + \frac{\epsilon}{2}}\right),$$

which results in

$$\|H\| = O\left(n^{1 - \frac{3\nu}{4} + \frac{\epsilon}{2}}\right).$$

This finally proves Proposition 3.3.5.  $\square$

### 3.4 Multiple-Unicast Capacity

#### 3.4.1 At High SNR

##### Multiple-Unicast Schemes

In [31] and [46], the authors show that for two clusters of area  $A$  separated by distance  $d$ , the following spatial degrees of freedom are achievable

$$\begin{cases} \min \left\{ n, \sqrt{A}/\lambda \right\}, & \text{if } 1 \leq d \leq \sqrt{A} \\ \min \left\{ n, \frac{A}{\lambda d} \right\}, & \text{if } \sqrt{A} \leq d \leq A/\lambda. \end{cases} \quad (3.11)$$

A follow up result to the achievable degrees of freedom is the multi-unicast capacity scaling in large ad hoc networks.

Given that we are considering the high-SNR regime, for the multiple-unicast communications in a wireless network, we are interested in exploiting the spatial degrees of freedom available in the network. In [46], the authors show that under line-of-sight model, the distributed MIMO based hierarchical cooperation architecture in [45] achieves a capacity scaling as

$$\max \left\{ \sqrt{n}, \min \left\{ n, \frac{\sqrt{A}}{\lambda} \right\} \right\}$$

in a network of  $n$  source-destination pairs uniformly distributed over an area  $A$  and communicating around a carrier wavelength  $\lambda$ . The authors in [46], show that the scaling of the capacity depends on how  $n$  compares to  $\sqrt{A}/\lambda$ . They further uncover the missing thread when it comes to the spatial degrees of freedom available in large networks that can be divided into three categories covering all possible regimes.

1)  $\sqrt{A}/\lambda \geq n$  : Capacity scales linearly in  $n$ . There are sufficient spatial degrees of freedom for all the users. Distributed MIMO communication can fully exploit all the available degrees of freedom.

2)  $\sqrt{A}/\lambda \leq \sqrt{n}$  : Capacity scales as  $\sqrt{n}$ . The spatial degrees of freedom available in the network are highly limited. Multihop communication is optimal.

3)  $\sqrt{n} \leq \sqrt{A}/\lambda \leq n$  : The spatial limitation is present in this regime, since we do not have full  $n$  spatial degrees of freedom. However, one can have more than what simple multihopping achieves, since we have more than  $\sqrt{n}$  degrees of freedom. The authors in [46] show that either a modification of the hierarchical cooperation scheme in [45] or a version of the MIMO-multihop scheme in [42] can achieve the  $\sqrt{A}/\lambda$  available degrees of freedom and therefore the optimal scaling of the capacity in this regime.

### Optimality of the Scheme

For the multiple-unicast scenario, we still need to know if the aggregated network throughput indeed scales as

$$T_n \sim \begin{cases} n & \text{if } (A/\lambda^2) \geq n^2 \\ (\sqrt{A}/\lambda) & \text{if } n \leq (A/\lambda^2) \leq n^2 \\ \sqrt{n} & \text{if } 1 \leq A/\lambda^2 \leq n. \end{cases} \quad (3.12)$$

The authors in [46] showed that such an aggregate throughput is achieved by a hierarchical cooperative strategy involving network-wide distributed MIMO transmissions in the first two cases, while a simple multihopping strategy achieves the performance claimed in the third regime. We therefore see that the wider the area is, the more degrees of freedom are available for communication in the network. The regime where  $A \sim n^2$  (corresponding to a sparse network of density  $O(1/n)$ ) models the case when the phase shifts are large enough to ensure sufficient channel diversity and full degrees of freedom of MIMO transmissions. On the contrary, in the regime where  $A \sim n$  (corresponding to a network of constant density), and even though this may seem surprising at first sight, phase shifts do not allow for efficient MIMO transmissions, so that multihopping becomes the best way to transfer information across the network.

In [17] the authors show that in real wireless networks, under the physical line-of-sight propagation model, the network area dictates the spatial degrees of freedom available in the network, thus significantly affecting the multi-unicast capacity. They show that physics laws of electromagnetic propagation restrict the spatial degrees of freedom of long-range distributed MIMO transmissions in ad hoc networks to  $\sqrt{A}/\lambda$ . In other words, when the network area  $A$  is relatively small and  $\sqrt{A}/\lambda \leq \sqrt{n}$ , then the multi-unicast capacity is bounded by  $\sqrt{n}$ . On the other hand, if  $\sqrt{n} < \sqrt{A}/\lambda \leq n$  then the capacity is equal to the number of degrees of freedom available in the network, which is indeed equal to  $\sqrt{A}/\lambda$ . Finally, for relatively large networks,  $\sqrt{A}/\lambda > n$ , we have full degrees of freedom enabling a capacity of the order of number of nodes  $n$ . This means that, at high SNR the multi-unicast capacity is upper bounded by

$$\max \left\{ \sqrt{n}, \min \left\{ n, \frac{\sqrt{A}}{\lambda} \right\} \right\}.$$

Therefore, based on [46] and [17], we know that the aggregated multi-unicast network throughput scales as in (3.12) up to logarithmic factors.

Further note that for the case of two clusters of area  $A$  separated by distance  $d$ , the authors in [12] and [13] try to find a matching upper bound to the achievable spatial degrees of freedom in (3.11). For this reason, they approximate the original matrix with a new matrix  $G$  whose entries are given by

$$g_{jk} = \exp \left( -\frac{2\pi i y_j y_k}{\lambda d} \right), \quad (3.13)$$

where  $0 \leq y_j, y_k \leq \sqrt{A}$  are the vertical components of nodes  $j$  and  $k$  placed in two squares of area  $A$  each and separated by a distance  $d$ . In other words, the authors in [12] use a Taylor approximation to quadratic order around 0 in the variable  $(y_j - y_k)$ , to get

$$\begin{aligned} r_{jk} &= \sqrt{(x_j + d + x_k)^2 + (y_j - y_k)^2} \\ &\approx (x_j + d + x_k) + \frac{(y_j - y_k)^2}{2d} \\ &= (x_j + d + x_k) + \frac{y_j^2}{2d} + \frac{y_k^2}{2d} - \frac{y_j y_k}{d}. \end{aligned} \quad (3.14)$$

Therefore,

$$\begin{aligned} h_{jk} &:= \frac{\exp(2\pi i r_{jk}/\lambda)}{r_{jk}} \\ &\approx \tilde{h}_{jk} := \frac{\exp(2\pi i (u_j + v_k - y_j y_k/d)/\lambda)}{r_{jk}}, \end{aligned}$$

where  $u_j = d/2 + x_j + \frac{y_j^2}{2d}$  and  $v_k = d/2 + x_k + \frac{y_k^2}{2d}$ . The authors further note that the eigenvalues of  $\tilde{H}\tilde{H}^*$  do not depend on the particular values of the  $u_j$ 's and  $v_k$ 's, which means that they are the same as the eigenvalues of  $GG^*$  ( $G$  as defined in (3.13)).

The maximum number of bits per second per Hertz that can be transferred reliably from the transmit cluster to the receive cluster given by

$$\begin{aligned} C_n &= \max_{Q \geq 0: Q_{kk} \leq P, \forall k} \log \det(I + HQH^*) \\ &\leq \log \det(I + nPHH^*), \end{aligned}$$

where  $Q$  is the covariance matrix of the input signal vector. The authors claim (through numerical evidence) that the following approximation holds:

$$\log \det(I + nPHH^*) = \log \det(I + GG^*)(1 + o(1))$$

with high probability as  $n$  gets large. Finally, they present the following theorem which suggests the tightness of the lower bound found in [46] (given by (3.11)).

**Theorem 3.4.1.** *If  $\frac{A}{\lambda d} \gg \sqrt{n}$ , then there exists a constant  $K_2 > 0$  such that*

$$\log \det(I + GG^*) \leq K_2 \min \left\{ n, \frac{A}{\lambda d} \right\} \log n,$$

*with high probability, as  $n$  gets large.*

The theorem above states that if  $\frac{A}{\lambda d} \gg \sqrt{n}$ , then the number of spatial degrees of freedom of a MIMO transmission between two clusters of area  $A$  separated by distance  $d$  is of order  $\frac{A}{\lambda d}$ , up to logarithmic factors.

Therefore, for  $\lambda = 1$  and inter-cluster distance  $2\sqrt{A} \leq d \leq A$ , we get

$$\text{DoF} \sim \min \left\{ n, \frac{A}{d} \right\}. \quad (3.15)$$

Moreover, for the same channel model (LOS model) with  $\lambda = 1$ , we know from Lemma 3.3.6 and the tightness of the upper bound on the broadcast capacity that if  $\widehat{H}$  is the  $n \times n$  channel matrix between two square clusters of  $n$  nodes distributed uniformly at random, each of area  $A$ , then

$$\|\widehat{H}\|^2 \sim \max \left\{ \frac{n^2}{Ad}, \frac{n}{d^2} \right\}, \quad (3.16)$$

where  $2\sqrt{A} \leq d \leq A$  denotes the distance between the centers of the two clusters. Combining (3.15) and (3.16), we obtain

$$\text{DoF} \times \|\widehat{H}\|^2 \sim \frac{n^2}{d^2}$$

which is equivalent to the result in (3.2) ( $\text{DoF} \times \text{BG} = n$ ), since

$$\|\widehat{H}\|^2 P = \text{BG} \times \frac{nP}{d^2} = \text{BG} \times \text{SNR}_l,$$

where the  $\text{SNR}_l = \frac{nP}{d^2}$  denotes the long-range SNR at the nodes in the destination cluster, because there are  $n$  nodes transmitting with power  $P$  and the signal attenuation over the inter-cluster distance  $d$  is  $1/d^2$ . Furthermore, we consider the power path loss in the ad hoc network to be that of free space propagation (path loss gain is  $\alpha = 2$ ). Therefore, we get

$$\text{SNR}_s = \text{SNR}_l,$$

which results from the fact that

$$\text{SNR}_s = \frac{P}{\left(n^{\frac{\nu-1}{2}}\right)^\alpha} = n^{1-\nu} P$$

and

$$\text{SNR}_l = \frac{nP}{(n^{\nu/2})^\alpha} = n^{1-\nu} P,$$

where  $n^{\nu/2}$  is the average distance between two nodes in the network.

### 3.4.2 At Low SNR

#### Trivial Upper Bound

If we assume the complex baseband-equivalent channel gain between node  $j$  and node  $k$  at time  $t$  is given by

$$h_{jk}[t] = \sqrt{G} \frac{\exp(i\theta_{jk}[t])}{r_{jk}},$$

where  $\theta_{jk}[t]$  is the random phase at time  $t$ , uniformly distributed in  $[0, 2\pi]$  and  $\{\theta_{jk}[t]; j \neq k\}$  is a collection of i.i.d random process, we know from [42] that the upper bound on the multi-unicast capacity, controlled by the power transfer across the network at low-SNR regime, is of order

$$n(\log n)^3 \text{SNR}_s.$$

Furthermore, for the line-of-sight propagation model (3.1), we know from [35] that for networks of area  $A = n$  (equivalently, density 1) and at low SNR, for any fixed  $\epsilon > 0$ , the multi-unicast capacity is upper bounded by  $n^{1+\epsilon} \text{SNR}_s$ . Here we generalize this result to networks with different sparsity. Indeed, we show that, at low SNR, a trivial upper bound on the multi-unicast capacity is given by the following theorem.

**Theorem 3.4.2.** *For any  $\epsilon > 0$ , in a wireless ad hoc network with  $n$  nodes and area  $A = n^\nu$ ,  $\nu > 0$ , and under low-SNR regime, the multi-unicast capacity can be upper bounded as*

$$T(n) \leq n^{2-\nu+\epsilon} P = n^{1+\epsilon} \text{SNR}_s.$$

Before proving the theorem above, we have the following lemma.

**Lemma 3.4.3.** *In a wireless ad hoc network with  $n$  nodes and area  $A = n^\nu$ ,  $\nu > 0$ , the following properties hold with high probability as  $n$  gets large:*

1. *For any  $\delta > 0$ , there exists a constant  $c(\delta) > 0$  such that the minimum distance between a randomly chosen node and all the other nodes in the network is greater than  $\frac{c(\delta) n^{\frac{\nu-1}{2}}}{n^\delta}$ .*
2. *For any  $\epsilon > 0$ , if the network is divided into  $n$  squares of area  $n^{\nu-1}$  each, then there are less than  $n^\epsilon$  nodes in each cell.*

*Proof.* Consider randomly chosen node in the network which is at distance larger than  $\frac{n^{\frac{\nu-1}{2}}}{n^\delta}$  to all other nodes in the network. This is equivalent to saying that there are no other nodes inside a circle of area  $\pi \frac{n^{\nu-1}}{n^{2\delta}}$  around this node. The probability of such an event is

$$\begin{aligned} \mathbb{P} \left( \text{minimum distance to a node} \geq \frac{n^{\frac{\nu-1}{2}}}{n^\delta} \right) &= \left( 1 - \frac{\pi n^{\nu-1}}{n^{2\delta}} \right)^{n-1} \\ &\geq 1 - \frac{c(\delta)}{n^{2\delta}}, \end{aligned}$$

where, for some constant  $c(\delta) > 0$ ,  $c(\delta)/n^{2\delta}$  goes to 0 with increasing  $n$ .

For the second part of the lemma, the proof is similar to that of Lemma 3.3.2.  $\square$



*Proof of Theorem 3.4.2.* Consider a source-destination pair  $s - d$  in the network. The transmission rate  $R(n)$  from source node  $s$  to destination  $d$  is bounded by the capacity of the single-input multiple-output (SIMO) channel between the source node  $s$  and the rest of the network:

$$\begin{aligned} R(n) &= \max_{p_X} I(X_s; Y_{\{1, \dots, n\} \setminus \{s\}}) \\ &= \log \left( 1 + P \sum_{j \neq s} |h_{js}|^2 \right) \\ &\leq \sum_{j \neq s} P |h_{js}|^2 = \sum_{j \neq s} \frac{P}{r_{js}^2} \end{aligned} \quad (3.17)$$

We further divide the network into square cells of area  $n^{\nu-1}$  each. By the aforementioned lemma, there are at most  $n^\epsilon$  nodes in each cell. Moreover, the cell  $S$  containing the source node  $s$  has at most 8 cells in the first layer (which represent the set of adjacent cells denoted by  $S_{\text{adj}}$ ), at most 16 cells in the second layer, and at most  $8\ell$  cells in the  $\ell$ -th layer (denoted by  $S_\ell$ ), where  $2 \leq \ell \leq \sqrt{n} - 1$ . As such, we can further upper bound (3.17) as follows

$$\begin{aligned} R(n) &\leq \sum_{j \neq s} \frac{P}{r_{js}^2} \\ &\leq \sum_{j \in S \cup S_{\text{adj}}} \sum_{j \neq s} \frac{P}{r_{js}^2} + \sum_{\ell=2}^{\sqrt{n}-1} \sum_{j \in S_\ell} \frac{P}{r_{js}^2} \\ &\stackrel{(a)}{\leq} \frac{9Pn^\epsilon}{\left(\frac{c(\delta)n^{\frac{\nu-1}{2}}}{n^\delta}\right)^2} + \sum_{\ell=2}^{\sqrt{n}-1} \sum_{j \in S_\ell} \frac{P}{r_{js}^2} \\ &\stackrel{(b)}{\leq} n^{\epsilon_1} n^{1-\nu} P + \sum_{\ell=2}^{\sqrt{n}-1} \frac{8\ell P n^\epsilon}{\left(\ell n^{\frac{\nu-1}{2}}\right)^2} \\ &= O(n^{\epsilon_1} n^{1-\nu} P + n^\epsilon n^{1-\nu} \log n P) = O(n^{\epsilon_1} \text{SNR}_s), \end{aligned}$$

where  $n^{\epsilon_1} = 9n^{\epsilon+2\delta}/c(\delta)^2$ , (a) results from Lemma 3.4.3 and (b) follows from Lemma 3.4.3 and the fact that nodes in layer  $\ell$  are at a distance of order  $\ell n^{(\nu-1)/2}$ .

Finally, we conclude the proof by upper bounding the multi-unicast capacity as follows

$$T(n) \leq nR(n) \leq n^{1+\epsilon_1} \text{SNR}_s$$

with high probability.  $\square$

### Available Achievable Schemes

Contrary to the case of high SNR, where there are communication schemes that achieve the upper bound on the multi-unicast capacity in ad hoc networks for

all sparsity regimes, in the case of low SNR, we do not have achievable schemes that match the upper bound in Theorem 3.4.2 except for very dense networks (area  $A = 1$ ) and very sparse networks (area  $A \geq n^2$ ). In other words, we know that if  $A = 1$  ( $\nu = 0$ ), the back-and-forth beamforming scheme (which we originally proposed as a broadcasting scheme) achieves a rate of  $n \text{SNR}_s$  and if  $A \geq n^2$  ( $\nu \geq 2$ ), the hierarchical cooperation scheme also achieves a rate of  $n \text{SNR}_s$ . The optimality of the existing schemes in the regimes mentioned at low SNR results from the following observations:

1. In dense networks,  $A = 1$ , the phases of the signals transmitted are highly correlated, which makes the back-and-forth scheme a perfect candidate to ensure the power transfer in the network. For this reason, if the area of the network is 1, back-and-forth beamforming scheme is capable of achieving a capacity of  $n \text{SNR}_s$ .
2. In sparse networks,  $A \geq n^2$ , the phases of the signals transmitted are highly uncorrelated and converge to the identically and independently distributed case. For this reason, the hierarchical cooperation scheme is a perfect candidate to fully exploit all the spatial degrees of freedom available in the network and achieve a capacity of  $n \text{SNR}_s$ .

Based on the discussion above, at low SNR, for networks of area  $1 < A < n^2$  ( $0 < \nu < 2$ ), the communication scheme should both exploit the correlation among the phases of the transmitted signals and the degrees of freedom available in the network to ensure an optimal power transfer across the network thus achieving a capacity of  $n \text{SNR}_s$ <sup>8</sup>. In other words, one should think of a hybrid communication scheme capable of beamforming over all the degrees of freedom optimally and this remains an interesting open problem worth the investigation<sup>9</sup>.

### 3.5 Communication Tradeoffs

In this work, we characterize the broadcast capacity of a wireless network at low SNR in line-of-sight environment and under various assumptions regarding the network density. The result exhibits a dichotomy between sparse networks, where node collaboration can hardly help enhancing communication rates, and dense networks, where significant gains can be obtained via collaborative beamforming. Moreover, we discuss the multi-unicast scenario presented in previous works to highlight the importance of spatial degrees of freedom.

We realize that the increase in the number of degrees of freedom comes at the expense of the beamforming capabilities in the network. In other words, dense networks provide higher correlation among the phases of the transmitted

<sup>8</sup>Assuming the trivial upper bound given in Theorem 3.4.2 is indeed achievable up to a multiplicative factor  $n^\epsilon$ .

<sup>9</sup>For  $A = n$ , it is shown in [35] that a throughput capacity of  $n^{6/7} \text{SNR}_s$  is achievable at low SNR.

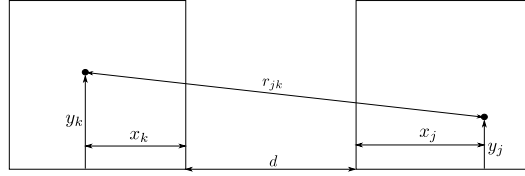


Figure 3.3 – Coordinate system.

signals which results in larger beamforming capabilities and smaller degrees of freedom, while sparse networks provide higher degrees of freedom and smaller beamforming capabilities. While the beamforming capabilities play a central role in the broadcast capacity at low SNR and the spatial degrees of freedom play a central role in multi-unicast capacity at high SNR, they both appear to be crucial when studying the multi-unicast capacity at low SNR, which remains an open problem.

### 3.A Appendix

*Proof of Lemma 3.3.2.* The number of nodes in a given cluster is the sum of  $n$  independently and identically distributed Bernoulli random variables  $B_i$ , with  $\mathcal{P}(B_i = 1) = M/n^\nu$ . Hence

$$\begin{aligned}
& \mathbb{P} \left( \sum_{i=1}^n B_i \geq (1 + \delta) M n^{1-\nu} \right) \\
&= \mathbb{P} \left( \exp \left( s \sum_{i=1}^n B_i \right) \geq \exp(s(1 + \delta) M n^{1-\nu}) \right) \\
&\leq \mathbb{E}^n (\exp(s B_1)) \exp(-s(1 + \delta) M n^{1-\nu}) \\
&= \left( \frac{M}{n^\nu} \exp(s) + 1 - \frac{M}{n^\nu} \right)^n \exp(-s(1 + \delta) M n^{1-\nu}) \\
&\leq \exp(-M n^{1-\nu} (s(1 + \delta) - \exp(s) + 1)) = \exp(-M n^{1-\nu} \Delta_+(\delta))
\end{aligned}$$

where  $\Delta_+(\delta) = (1 + \delta) \log(1 + \delta) - \delta$  by choosing  $s = \log(1 + \delta)$ . The proof of the lower bound follows similarly by considering the random variables  $-B_i$ . The conclusion follows from the union bound.  $\square$

*Proof of Lemma 3.3.3.* We present lower and upper bounds on the distance  $r_{jk}$  separating a receiving node  $j \in \mathcal{R}_i$  and a transmitting node  $k \in \mathcal{T}_i$ . Denote by  $x_j, x_k, y_j$ , and  $y_k$  the horizontal and the vertical positions of nodes  $j$  and  $k$ , respectively (as shown in Fig. 3.3). An easy lower bound on  $r_{jk}$  is

$$r_{jk} \geq x_k + x_j + d$$

On the other hand, using the inequality  $\sqrt{1+x} \leq 1 + \frac{x}{2}$ , we obtain

$$\begin{aligned} r_{jk} &= \sqrt{(x_k + x_j + d)^2 + (y_j - y_k)^2} \\ &= (x_k + x_j + d) \sqrt{1 + \frac{(y_j - y_k)^2}{(x_k + x_j + d)^2}} \\ &\leq x_k + x_j + d + \frac{(y_j - y_k)^2}{2d} \leq x_k + x_j + d + \frac{1}{2c_1^2}. \end{aligned}$$

Therefore,

$$0 \leq r_{jk} - x_k - x_j - d \leq \frac{1}{2c_1^2}.$$

After bounding  $r_{jk}$ , we can proceed to the proof of the lemma as follows:

$$\begin{aligned} \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &= \left| \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k - x_j - d))}{r_{jk}} \right| \\ &\geq \Re \left( \sum_{k \in \mathcal{T}_i} \frac{\exp(2\pi i(r_{jk} - x_k - x_j - d))}{r_{jk}} \right) \\ &\geq \sum_{k \in \mathcal{T}_i} \frac{\cos\left(\frac{\pi}{c_1^2}\right)}{r_{jk}} \geq K_1 \frac{Mn^{1-\nu}}{d}, \end{aligned}$$

where the constant  $c_1$  is chosen sufficiently large so that  $\cos\left(\frac{\pi}{c_1^2}\right) > 0$ .  $\square$

*Proof of Lemma 3.3.4.* There are  $N_C$  clusters transmitting simultaneously. Except for the horizontally adjacent cluster of a given cluster pair ( $i$ -th cluster pair), all the rest of the transmitting clusters are considered as interfering clusters (there are  $N_C - 1$  of them). With high probability, each cluster contains  $\Theta(Mn^{1-\nu})$  nodes.

For the sake of clarity, we assume here that every cluster contains exactly  $Mn^{1-\nu}$  nodes, but the argument holds in the general case. In this lemma, we upper bound the magnitude of interfering signals from the simultaneously interfering clusters at node  $j \in \mathcal{R}_i$  as follows

$$\begin{aligned} \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| &\leq \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\ &\quad + \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\sin(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \end{aligned}$$

We only upper bound the first term (cosine terms) in the equation above as follows (we can upper bound the second term (sine terms) in exactly the same fashion):

$$\begin{aligned}
& \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\
& \leq \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} - \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| \\
& \quad + \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| \\
& \leq \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} - \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| \\
& \quad + \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \left| \sum_{k \in \mathcal{T}_l} \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| \\
& \leq \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} - \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| \\
& \quad + 2 \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \left| \sum_{k \in \mathcal{T}'_l} \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| \tag{3.18}
\end{aligned}$$

where  $\mathcal{T}'_l$  denotes the  $l$ -th interfering transmit cluster that is at a vertical distance of  $(l-1)\frac{n^{\nu/4}}{2c_1} + lc_2n^{\nu/4+\epsilon}$  from the desired receiving cluster  $\mathcal{R}_i$ . Note the difference between the first terms of the equations (2.8) and (3.18). Although we apply Hoeffding's inequality to upper bound the first term in (3.18) as before, we will see that it requires a slightly different approach than the one used to upper bound the first term in (2.8). Let us first upper bound the second term in (3.18). Denote by  $X_k^{(l)} = (\cos(2\pi(r_{jk} - x_k)))/(r_{jk}) \forall k \in \mathcal{T}'_l$ . Note that  $X_k^{(l)}$ 's are independent and identically distributed. For any  $k \in \mathcal{T}'_l$ , we have

$$|r_{jk} - x_k| = \sqrt{(x_k + x_j + d)^2 + (y_j - y_k)^2} - x_k \geq d = \frac{n^{\nu/2}}{4}$$

is a  $C^2$  function and

$$\begin{aligned} \left| \frac{\partial(r_{jk} - x_k)}{\partial y_k} \right| &= \left| \frac{\partial r_{jk}}{\partial y_k} \right| = \frac{|y_k - y_j|}{r_{jk}} \\ &\geq \frac{l c_2 n^{\nu/4+\epsilon} + (l-1) \frac{n^{\nu/4}}{2c_1}}{n^{\nu/2}} \\ &\geq l c_2 n^{-\nu/4+\epsilon} \end{aligned}$$

Moreover,  $r_{jk}''$  changes sign at most twice. By the integration by parts formula, we obtain

$$\begin{aligned} &\int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \\ &= \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{2\pi r_{jk}'}{2\pi r_{jk}' r_{jk}} \cos(2\pi(r_{jk} - x_k)) \\ &= \frac{-\sin(2\pi(r_{jk} - x_k))}{2\pi r_{jk}' r_{jk}} \Big|_{y_{k_0}}^{y_{k_1}} + \frac{1}{2\pi} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{r_{jk} r_{jk}'' + (r_{jk}')^2}{(r_{jk}' r_{jk})^2} \sin(2\pi(r_{jk} - x_k)) \end{aligned}$$

which in turn yields the upper bound

$$\begin{aligned} &\left| \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq \frac{1}{2\pi} \left( \frac{2}{\min_{y_k} \{|r_{jk}'| |r_{jk}|\}} + \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{|r_{jk}''|}{(r_{jk}')^2 |r_{jk}|} + \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{1}{r_{jk}^2} \right) \\ &\leq \frac{1}{2\pi} \left( \frac{4}{l c_2 n^{\nu/4+\epsilon}} + \frac{1}{\min_{y_k} \{|r_{jk}|\}} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{|r_{jk}''|}{(r_{jk}')^2} + \frac{|y_{k_1} - y_{k_0}|}{\min_{y_k} \{r_{jk}^2\}} \right) \\ &\leq \frac{1}{2\pi} \left( \frac{4}{l c_2 n^{\nu/4+\epsilon}} + \frac{4}{l c_2 n^{\nu/4+\epsilon}} + \frac{2}{n^{3\nu/4}} \right) \leq \frac{9/(2\pi)}{l c_2 n^{\nu/4+\epsilon}}. \quad (\text{take } \nu > 2\epsilon) \end{aligned}$$

Therefore, for any  $k \in \mathcal{T}'_l$ ,

$$\begin{aligned} \left| \mathbb{E} \left( X_k^{(l)} \right) \right| &= \left| \frac{4}{n^{\nu/2}} \int_0^{\frac{n^{\nu/2}}{4}} dx_k \frac{1}{|y_{k_1} - y_{k_0}|} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq \frac{4}{n^{\nu/2} |y_{k_1} - y_{k_0}|} \int_0^{\frac{n^{\nu/2}}{4}} dx_k \left| \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| \\ &\leq \frac{9/(2\pi)}{|y_{k_1} - y_{k_0}| l c_2 n^{\nu/4+\epsilon}} \leq \frac{9c_1}{\pi c_2} \frac{1}{l n^{\nu/2+\epsilon}} \leq \frac{9c_1}{\pi c_2} \frac{1}{l d n^\epsilon}. \quad (3.19) \end{aligned}$$

We further upper bound the first term in (3.18) by using Hoeffding's inequality [23]. Denote by  $X_k = \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}}$ , where  $1 \leq k \leq N_C M n^{1-\nu} = \Theta(n^{1-\epsilon})$ .

Note that  $X_k$ 's are i.i.d. and integrable random variables that represent *all* nodes in *all* the interfering clusters. In other words, we have

$$\left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} - \mathbb{E} \left( \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right) \right| = \left| \sum_{k=1}^{n^{1-\epsilon}} (X_k - \mathbb{E}(X_k)) \right|.$$

We have  $X_k \in [-1/d, 1/d]$ . As such, Hoeffding's inequality yields

$$\begin{aligned} \mathbb{P} \left( \frac{1}{n^{1-\epsilon}} \left| \sum_{k=1}^{n^{1-\epsilon}} (X_k - \mathbb{E}(X_k)) \right| > t \right) &\leq 2 \exp \left( -\frac{n^{1-\epsilon} t^2}{2/d^2} \right) \\ &= 2 \exp \left( -\frac{1}{2} n^{1-\epsilon} d^2 t^2 \right) \\ &\stackrel{(a)}{=} 2 \exp(-n^{\epsilon_1}), \end{aligned}$$

where (a) holds if  $t = \frac{\sqrt{2n^{\epsilon+\epsilon_1-1}}}{d}$ . Therefore, we have

$$\left| \sum_{k=1}^{n^{1-\epsilon}} (X_k - \mathbb{E}(X_k)) \right| \leq n^{1-\epsilon} t = \frac{\sqrt{2n^{1-\epsilon+\epsilon_1}}}{d} \quad (3.20)$$

with probability  $\geq 1 - 2 \exp(-n^{\epsilon_1})$ . Combining (3.19) and (3.20), we can upper bound (3.18) as follows

$$\begin{aligned} \left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| &\leq \frac{\sqrt{2n^{1-\epsilon+\epsilon_1}}}{d} + \sum_{l=1}^{N_C} \frac{9c_1}{\pi c_2} \frac{Mn^{1-\nu}}{l d n^\epsilon} \\ &\leq \frac{\sqrt{2n^{1-\epsilon+\epsilon_1}}}{d} + \frac{9c_1}{\pi c_2} \frac{Mn^{1-\nu}}{d n^\epsilon} \log n. \end{aligned}$$

Note that for  $M = \Theta(n^{3\nu/4})$  and  $\nu \leq 2 - 2(\epsilon + \epsilon_1)$ , we have

$$\frac{\sqrt{2n^{1-\epsilon+\epsilon_1}}}{d} \leq \frac{9c_1}{\pi c_2} \frac{Mn^{1-\nu}}{d n^\epsilon} \log n$$

Finally, upper bounding the sine terms in the same fashion, we obtain

$$\left| \sum_{\substack{l=1 \\ l \neq i}}^{N_C} \sum_{k \in \mathcal{T}_l} \frac{\exp(2\pi i(r_{jk} - x_k))}{r_{jk}} \right| = O \left( \frac{Mn^{1-\nu}}{d n^\epsilon} \log n \right)$$

with high probability (more precisely, with probability  $\geq 1 - 4 \exp(-n^{\epsilon_1})$ ), which concludes the proof.  $\square$

*Proof of Lemma 3.3.6.* We try to extend the proof of the particular case of  $A = m$  ( $\nu = 1$ ) presented in Chapter 2.

For the first moment, we have

$$\begin{aligned} \mathbb{E}\left(\text{Tr}\left(\widehat{H}\widehat{H}^\dagger\right)\right) &= \sum_{j_1, k_1=1}^m \mathbb{E}(\hat{h}_{j_1 k_1} \hat{h}_{j_1 k_1}^*) \\ &= \sum_{j_1, k_1=1}^m \mathbb{E}(|\hat{h}_{j_1 k_1}|^2) = \sum_{j_1, k_1=1}^m \frac{1}{r_{j_1 k_1}^2} = O\left(\frac{m^2}{d^2}\right). \end{aligned}$$

For  $\ell \geq 1$ , the  $\ell$ -th moment is given by

$$\begin{aligned} \mathbb{E}(\text{Tr}((\widehat{H}\widehat{H}^\dagger)^\ell)) &= \sum_{\substack{j_1, \dots, j_\ell=1 \\ k_1, \dots, k_\ell=1}}^m \mathbb{E}(\hat{h}_{j_1 k_1} \hat{h}_{j_2 k_1}^* \dots \hat{h}_{j_\ell k_\ell} \hat{h}_{j_1 k_\ell}^*) \\ &\leq \sum_{i=0}^{2\ell-2} \sum_{i_1=0}^i \binom{\ell}{i_1} \binom{\ell}{i-i_1} m^{2\ell-i} \max_{\binom{\ell}{i_1} \binom{\ell}{i-i_1} \text{ cases}} \{\widehat{S}_\ell^{(i)}\}, \end{aligned}$$

where  $\widehat{S}_\ell^{(i)} = |\mathbb{E}(\hat{h}_{j_1 k_1} \hat{h}_{j_2 k_1}^* \dots \hat{h}_{j_\ell k_\ell} \hat{h}_{j_1 k_\ell}^*)|$  with  $i$  “equality”s (as in Chapter 2, by “equality” we mean an index is equal to another index). Note that the nodes corresponding to the indices are randomly and uniformly positioned on  $\sqrt{A}$ -by- $\sqrt{A}$  square clusters. In other words,  $\widehat{S}_\ell^{(i)}$  assumes that the points corresponding to  $j$ 's and  $k$ 's are randomly chosen in two squares of area  $A$ , respectively. We start with the following lemma.

**Lemma 3.A.1** ( $\star$ ). *For  $\ell \geq 1$  and  $0 \leq i \leq 2\ell - 2$ , let*

$$\widehat{S}_\ell^{(i)} = |\mathbb{E}(\hat{h}_{j_1 k_1} \hat{h}_{j_2 k_1}^* \dots \hat{h}_{j_\ell k_\ell} \hat{h}_{j_1 k_\ell}^*)|, \quad (3.21)$$

*with  $i$  “equality”s, where by “equality” we mean an index is equal to another index. For example, if  $i = 0$ , then  $j_1 \neq \dots \neq j_\ell$  and  $k_1 \neq \dots \neq k_\ell$ . For all  $0 \leq i \leq 2\ell - 2$  and any  $\widehat{S}_\ell^{(i)}$  out of the  $\sum_{i_1=0}^i \binom{\ell}{i_1} \binom{\ell}{i-i_1}$  possible ones, we have*

$$\sum_{i=0}^{2\ell-2} m^{2\ell-i} \widehat{S}_\ell^{(i)} = O\left(\sum_{i=0}^{\ell-1} m^{2\ell-i} \frac{\widetilde{A}^{\ell-i-1} \widetilde{B}^{i+1}}{(\sqrt{A})^{2\ell-i-1}} + \frac{m^\ell}{d_{jk}^{2\ell}}\right),$$

where

$$\widetilde{A} = \frac{\log A}{d_{jk}} \quad \text{and} \quad \widetilde{B} = \frac{\sqrt{A}}{d_{jk}^2}.$$



*Proof of Lemma 3.A.1.* In general, for any  $\ell \geq 1$ , we have

$$\begin{aligned} \widehat{S}_\ell^{(0)} &= \left| \frac{1}{A^{2\ell}} \int dj_1 \int dj_2 \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \int dj_3 \int dk_2 \frac{e^{2\pi i g_{j_2 j_3}(k_2)}}{\rho_{j_2 j_3}(k_2)} \cdots \right. \\ &\quad \left. \int dj_\ell \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1} j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1} j_\ell}(k_{\ell-1})} \int dk_\ell \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right| \\ &\leq \frac{1}{A^{2\ell}} \int dj_1 \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \int dj_3 \left| \int dk_2 \frac{e^{2\pi i g_{j_2 j_3}(k_2)}}{\rho_{j_2 j_3}(k_2)} \right| \cdots \\ &\quad \int dj_\ell \left| \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1} j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1} j_\ell}(k_{\ell-1})} \right| \left| \int dk_\ell \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right| \\ &= \frac{1}{A^{\ell+1/2}} \int dj_1 \widehat{A}_{1,2} \cdot \widehat{A}_{2,3} \cdots \widehat{A}_{\ell-1,\ell} \cdot \widehat{B}_{\ell,1} \end{aligned}$$

where

$$\widehat{A}_{t-1,t} = \frac{1}{\sqrt{A}} \int dj_t \left| \frac{1}{\sqrt{A}} \int dk_{t-1} \frac{e^{2\pi i g_{j_{t-1} j_t}(k_{t-1})}}{\rho_{j_{t-1} j_t}(k_{t-1})} \right| \text{ for } 2 \leq t \leq \ell$$

and

$$\widehat{B}_{\ell,1} = \left| \frac{1}{\sqrt{A}} \int dk_\ell \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right| \leq \frac{1}{\sqrt{A}} \int dk_\ell \left| \frac{e^{2\pi i g_{j_\ell j_1}(k_\ell)}}{\rho_{j_\ell j_1}(k_\ell)} \right| \leq \frac{\sqrt{A}}{d_{jk}^2} = \widetilde{B}.$$

Note that each integral in  $\widehat{A}_{t-1,t}$  and  $\widehat{B}_{\ell,1}$  is normalized by  $\sqrt{A}$  to account for the uniform distribution of the horizontal component of each point over 0 and  $\sqrt{A}$ . Despite the fact that the points corresponding to  $j$  and  $k$  indices are randomly distributed over square clusters instead of rectangular clusters, we have the following claim.

**Claim 3.A.2.** *Similar to the case of rectangular clusters studied in Chapter 2 (see (2.23)), for the case of square clusters, the upper bound on  $\widehat{A}_{t,t+1}$  is given by*

$$\widehat{A}_{t,t+1} \leq \frac{\log A}{d_{jk}} = \widetilde{A}.$$

Therefore, we have

$$\begin{aligned} \widehat{S}_\ell^{(0)} &\leq \frac{1}{A^{\ell+1/2}} \int dj_1 \widehat{A}_{1,2} \cdot \widehat{A}_{2,3} \cdots \widehat{A}_{\ell-1,\ell} \cdot \widehat{B}_{\ell,1} \\ &\leq \frac{1}{A^{\ell+1/2}} \int dj_1 \widetilde{A}^{\ell-1} \widetilde{B} = \frac{\widetilde{A}^{\ell-1} \widetilde{B}}{A^{\ell-1/2}}. \end{aligned} \quad (3.22)$$

To generalize this result to any  $0 \leq i \leq \ell - 1$ , we use the following observation. Assume the first “equality” is given by  $k_m = k_p$ , where  $1 \leq m < p \leq \ell - 1$ . This means instead of having the term

$$\widehat{A}_{p,p+1} = \frac{1}{\sqrt{A}} \int dj_{p+1} \left| \frac{1}{\sqrt{A}} \int dk_p \frac{e^{2\pi i g_{j_p j_{p+1}}(k_p)}}{\rho_{j_p j_{p+1}}(k_p)} \right| \leq \widetilde{A}$$

we have

$$\frac{1}{\sqrt{A}} \int dj_{p+1} \left| \frac{e^{2\pi i g_{j_p, j_{p+1}}(k_m)}}{\rho_{j_p, j_{p+1}}(k_m)} \right| \leq \frac{\sqrt{A}}{d_{jk}^2} = \tilde{B}.$$

Therefore, for  $1 \leq m < p \leq \ell - 1$ , we have

$$\begin{aligned} \widehat{S}_\ell(k_m = k_p) &\leq \frac{1}{A^\ell} \int dj_1 \widehat{A}_{1,2} \cdots \widehat{A}_{p-1,p} \cdot \tilde{B} \cdot \widehat{A}_{p+1,p+2} \cdots \widehat{A}_{\ell-1,\ell} \cdot \widehat{B}_{\ell,1} \\ &\leq \frac{1}{A^\ell} \int dj_1 \tilde{A}^{\ell-2} \tilde{B}^2 = \frac{\tilde{A}^{\ell-2} \tilde{B}^2}{A^{\ell-1}}. \end{aligned}$$

The only case remaining for the first ‘‘equality’’ is  $k_m = k_\ell$ , where  $1 \leq m < \ell$ . In this case, the term

$$\begin{aligned} &\widehat{A}_{\ell-1,\ell} \cdot \widehat{B}_{\ell,1} \\ &= \frac{1}{\sqrt{A}} \int dj_\ell \left| \frac{1}{\sqrt{A}} \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1}, j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1}, j_\ell}(k_{\ell-1})} \right| \cdot \left| \frac{1}{\sqrt{A}} \int dk_\ell \frac{e^{2\pi i g_{j_\ell, j_1}(k_\ell)}}{\rho_{j_\ell, j_1}(k_\ell)} \right| \\ &\leq \tilde{A} \tilde{B} \end{aligned}$$

is replaced by

$$\begin{aligned} &\frac{1}{\sqrt{A}} \int dj_\ell \left| \frac{1}{\sqrt{A}} \int dk_{\ell-1} \frac{e^{2\pi i g_{j_{\ell-1}, j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1}, j_\ell}(k_{\ell-1})} \right| \cdot \left| \frac{e^{2\pi i g_{j_\ell, j_1}(k_m)}}{\rho_{j_\ell, j_1}(k_m)} \right| \\ &\leq \frac{1}{\sqrt{A}} \int dj_\ell \frac{1}{\sqrt{A}} \int dk_{\ell-1} \left| \frac{e^{2\pi i g_{j_{\ell-1}, j_\ell}(k_{\ell-1})}}{\rho_{j_{\ell-1}, j_\ell}(k_{\ell-1})} \right| \cdot \left| \frac{e^{2\pi i g_{j_\ell, j_1}(k_m)}}{\rho_{j_\ell, j_1}(k_m)} \right| \\ &\leq \frac{A}{d_{jk}^4} = \tilde{B}^2, \end{aligned}$$

which results in the same upper bound on  $S_\ell(k_m = k_\ell)$  as before. As such,

$$\widehat{S}_\ell^{(1)} = O\left(\frac{\tilde{A}^{\ell-2} \tilde{B}^2}{A^{\ell-1}}\right). \quad (3.23)$$

For the second ‘‘equality’’, without loss of generality, assume  $j_m = j_p$ , where  $1 \leq m < p \leq \ell$ . If index  $k_{p-1}$  still exists (did not vanish due to the first ‘‘equality’’), then instead of having the term

$$\widehat{A}_{p-1,p} = \frac{1}{\sqrt{A}} \int dj_p \left| \frac{1}{\sqrt{A}} \int dk_{p-1} \frac{e^{2\pi i g_{j_{p-1}, j_p}(k_{p-1})}}{\rho_{j_{p-1}, j_p}(k_{p-1})} \right| \leq \tilde{A},$$

we have

$$\begin{aligned} &\left| \frac{1}{\sqrt{A}} \int dk_{p-1} \frac{e^{2\pi i g_{j_{p-1}, j_m}(k_{p-1})}}{\rho_{j_{p-1}, j_m}(k_{p-1})} \right| \\ &\leq \frac{1}{\sqrt{A}} \int dk_{p-1} \left| \frac{e^{2\pi i g_{j_{p-1}, j_m}(k_{p-1})}}{\rho_{j_{p-1}, j_m}(k_{p-1})} \right| \leq \frac{\sqrt{A}}{d_{jk}^2} = \tilde{B}. \end{aligned}$$

Therefore,

$$\widehat{S}_\ell(j_m = j_p, k_u = k_v, p \neq v + 1) \leq \frac{1}{A^{\ell-1/2}} \int dj_1 \widetilde{A}^{\ell-3} \widetilde{B}^3 = \frac{\widetilde{A}^{\ell-3} \widetilde{B}^3}{A^{\ell-3/2}}.$$

Note that if the index  $k_{p-1}$  vanished due to the first “equality”, then having  $j_m = j_p$  as the second “equality” results in

$$\widehat{S}_\ell(j_m = j_p, k_u = k_{p-1}) \leq \frac{\widetilde{A}^{\ell-2} \widetilde{B}^2}{A^{\ell-1}}.$$

As such, we get

$$\widehat{S}_\ell^{(2)} = O\left(\frac{\widetilde{A}^{\ell-3} \widetilde{B}^3}{A^{\ell-3/2}} + \frac{\widetilde{A}^{\ell-2} \widetilde{B}^2}{A^{\ell-1}}\right). \quad (3.24)$$

Note that the second term in (3.24) can be ignored, since we know from the upper bound on  $\widehat{S}_\ell^{(1)}$  that

$$m^{2\ell-2} \frac{\widetilde{A}^{\ell-2} \widetilde{B}^2}{A^{\ell-1}} \leq m^{2\ell-1} \frac{\widetilde{A}^{\ell-2} \widetilde{B}^2}{A^{\ell-1}}.$$

Combining (2.26), (3.23) and (3.24), we have

$$\sum_{i=0}^2 m^{2\ell-i} \widehat{S}_\ell^{(i)} = O\left(\sum_{i=0}^2 m^{2\ell-i} \frac{\widetilde{A}^{\ell-i-1} \widetilde{B}^{i+1}}{(\sqrt{A})^{2\ell-i-1}}\right).$$

Note that every time we add a new “equality”, we obtain exactly one *new* term that results from replacing one  $\widetilde{A}$  term by one  $\widetilde{B}$  term. As such, covering all the possible less than  $\ell$  number of “equality”s gives

$$\sum_{i=0}^{\ell-1} m^{2\ell-i} \widehat{S}_\ell^{(i)} = O\left(\sum_{i=0}^{\ell-1} m^{2\ell-i} \frac{\widetilde{A}^{\ell-i-1} \widetilde{B}^{i+1}}{(\sqrt{A})^{2\ell-i-1}}\right).$$

For  $\ell \leq i \leq 2\ell - 2$ , we have the following trivial bound on  $\widehat{S}_\ell^{(i)}$  (with any  $i$  “equality”s),

$$\begin{aligned} \widehat{S}_\ell^{(i)} &= |\mathbb{E}(\hat{h}_{j_1 k_1} \hat{h}_{j_2 k_1}^* \cdots \hat{h}_{j_\ell k_\ell} \hat{h}_{j_1 k_\ell}^*)| \leq \mathbb{E}\left(|\hat{h}_{j_1 k_1} \hat{h}_{j_2 k_1}^* \cdots \hat{h}_{j_\ell k_\ell} \hat{h}_{j_1 k_\ell}^*|\right) \\ &\leq \frac{1}{d_{jk}^{2\ell}}. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \sum_{i=\ell}^{2\ell-2} m^{2\ell-i} \widehat{S}_\ell^{(i)} &\leq \sum_{i=\ell}^{2\ell-2} \frac{m^{2\ell-i}}{d_{jk}^{2\ell}} \\ &\leq (\ell-1) \frac{m^\ell}{d_{jk}^{2\ell}} = O\left(\frac{m^\ell}{d_{jk}^{2\ell}}\right), \end{aligned}$$

which concludes the proof of the Lemma.  $\square$

Using Lemma 3.A.1, we can further proceed and upper bound the  $\ell$ -th moment as such

$$\begin{aligned}
\mathbb{E}(\text{Tr}((\widehat{H}\widehat{H}^\dagger)^\ell)) &= \sum_{\substack{j_1, \dots, j_\ell=1 \\ k_1, \dots, k_\ell=1}}^m \mathbb{E}(\widehat{h}_{j_1 k_1} \widehat{h}_{j_2 k_1}^* \dots \widehat{h}_{j_\ell k_\ell} \widehat{h}_{j_1 k_\ell}^*) \\
&\leq \sum_{i=0}^{2\ell-2} \sum_{i_1=0}^i \binom{\ell}{i_1} \binom{\ell}{i-i_1} m^{2\ell-i} \max_{\binom{\ell}{i_1} \binom{\ell}{i-i_1} \text{ cases}} \{\widehat{S}_\ell^{(i)}\} \\
&= O\left(\sum_{i=0}^{2\ell-2} m^{2\ell-i} \frac{\widetilde{A}^{\ell-i-1} \widetilde{B}^{i+1}}{(\sqrt{A})^{2\ell-i-1}} + \frac{m^\ell}{d_{jk}^{2\ell}}\right) \\
&= O\left(\sum_{i=0}^{2\ell-2} m^{2\ell-i} \frac{(\log A/d_{jk})^{\ell-i-1} (\sqrt{A}/d_{jk}^2)^{i+1}}{(\sqrt{A})^{2\ell-i-1}} + \frac{m^\ell}{d_{jk}^{2\ell}}\right) \\
&= O\left(\sum_{i=0}^{2\ell-2} \left(\frac{m^2 \log A}{d_{jk} A}\right)^\ell \left(\frac{A}{m d_{jk} \log A}\right)^i \frac{A}{d_{jk} \log A} + \frac{m^\ell}{d_{jk}^{2\ell}}\right) \\
&= O\left(\max\left\{\left(\frac{m^2 \log A}{d_{jk} A}\right)^\ell, \left(\frac{m}{d_{jk}^2}\right)^\ell\right\}\right).
\end{aligned}$$

As in Appendix 2.A, applying the Markov's inequality, concludes the proof of lemma 3.3.6 for aligned clusters, which can be easily generalized to tilted clusters, as shown in Fig. 2.5.  $\square$

---

# 4

## Full-Duplex vs Half-Duplex

---

Usually, wireless radios are half-duplex, i.e. they can not transmit and receive at the same time over the same frequency band. However, building on self-interference cancellation techniques, full-duplex radios have emerged as a viable paradigm over the recent years. In this chapter, we ask the following question: how much can full-duplex increase the capacity of wireless networks? Intuitively, one may expect that full-duplex radios can at most double the capacity of wireless networks, since they enable nodes to transmit and receive at the same time. However, we show that the capacity gain can indeed be larger than a factor of 2; in particular, we construct a specific instance of a wireless relay network where the capacity with full-duplex radios is triple the capacity of the network when the relays are half-duplex. We also propose a universal schedule for half-duplex networks composed of independent, memoryless, point-to-point channels which achieves at least a fraction of  $1/4$  of the corresponding full-duplex capacity. This means that for wireless networks composed of point-to-point channels full-duplex capability at the relays cannot more than quadruple the capacity of network.

The chapter is structured as follows. Section 4.1 presents some preliminaries and gives a brief description of the type of network models we want to study. In Section 4.2, we prove that there exists wireless network where the full-duplex capacity can be as large as 3 times the capacity of the corresponding half-duplex network. Finally, Section 4.3 shows that full-duplex operation can at most quadruple the capacity of wireless half-duplex networks of independent, memoryless, point-to-point AWGN channels.

## 4.1 Problem Formulation and Background

Consider a single-source single-destination network with orthogonal noisy (for example AWGN) channels and  $N$  relay nodes denoted by node 1 through node  $N$  assisting the communication between the source node 0 and the destination node  $N + 1$  (see Fig. 4.3). For the relay nodes, we consider either full-duplex operation, i.e. the relays are capable of receiving and transmitting simultaneously, or half-duplex operation, i.e. they are limited to either receiving or transmitting over a given frequency at any given time. Note that when there is broadcast and superposition, i.e. the constituent channels are not independent of each other, the information theoretic capacity of relay networks remains unknown to date, both when the relays are half-duplex and full-duplex.

When the network consists of independent, memoryless, noisy channels and the relays are full-duplex, the information theoretic capacity of the network is simple to obtain. In particular, due to network equivalence [28], the channel between any two nodes can be equivalently considered as a noiseless bit pipe with capacity equal to the capacity of the corresponding noisy channel. In this case, the highest possible rate of communication from the source to the destination with full-duplex relays is given by the max-flow min-cut theorem [19]

$$C^{(\text{FD})} = \min_{\mathcal{A} \subseteq [1:N]} \sum_{u \in \mathcal{A}, v \in \bar{\mathcal{A}}} c_{u,v}, \quad (4.1)$$

where  $c_{u,v}$  denotes the point-to-point capacity of the noisy channel from node  $u$  to node  $v$ . Note that  $c_{u,v} = 0$  if  $u$  and  $v$  are not connected or the data flow is in the opposite direction (from  $v$  to  $u$ ).

When the relays are half-duplex, the capacity remains open even when the network consists of orthogonal noisy AWGN channels. What complicates the problem in this case is that now each relay needs to develop a strategy of when to listen and when to transmit. Fixed scheduling strategies are those where the listen-talk states of the relays are established prior to the start of communication (but perhaps depending on global channel/network knowledge). However, random scheduling strategies, which allow the schedules to change during runtime, can be also used so as to convey additional information with the transmit and receive states of the relays. Such random schedules can increase capacity [30, 29, 55]. Moreover, the transmit power of the relays can be optimized across different configurations of the network while still satisfying an average power constraint at the nodes. Note that in a network of  $N$  relays where each relay can be in either transmit or receive state, there are  $2^N$  different possible configurations for the network. While the capacity of such networks remains unknown in general, inner and upper bounds on the capacity are available which differ by a constant gap that depends only on  $N$  and is independent of the channel configurations and the network topology. For example, [44] develops inner and upper bounds on the capacity of any half-duplex Gaussian network (under a general model which enables broadcast and superposition of signals, and use of

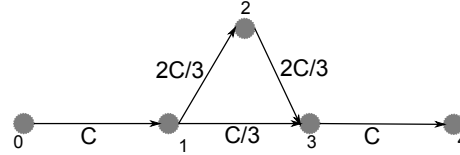


Figure 4.1 – Single-source single-destination network with 3 relay nodes.

multiple antennas at each node), where the gap is linear in the total number of nodes (or the total number of antennas in case of nodes with multiple antennas). When the upper bound in [44, Theorem 2.3] is specialized to case where the channels in the network are independent AWGN channels, it implies the highest possible rate of communication from the source to the destination with half-duplex relays can be upper bounded by the following cutset like bound

$$C^{(\text{HD})} \leq \max_{\mathbb{P}_{S_{[1:N]}}} \min_{\mathcal{A} \subseteq [1:N]} \sum_{u \in \mathcal{A}, v \in \bar{\mathcal{A}}} c_{u,v} \times \mathbb{P}(S_u = 1, S_v = 0) + 3N, \quad (4.2)$$

where  $c_{u,v}$  is the corresponding point-to-point capacity of the channel from  $u$  to  $v$  as before;  $S_{[1:N]}$  denotes the vector  $[S_1, \dots, S_N]$ . For any  $i \in [1 : N]$ ,  $S_i \in [0 : 1]$  denotes the *state* random variable that takes value 0 if node  $i$  is in receive-mode and value 1 if node  $i$  is in transmit-mode.  $\mathbb{P}_{S_{[1:N]}}$  corresponds to a half-duplex schedule specifying how much time the network spends in each state  $S_{[1:N]}$ . Note that the first term in the expression corresponds to the capacity of the half-duplex network under the best possible fixed schedule. The result (50) of [44] essentially bounds the potential gain due to random scheduling and power control across different schedules within the additive term  $3N$ . We will make use of this upper bound in the sequel. Note that a lower bound on the capacity of the half-duplex network can be obtained by computing the capacity under any fixed schedule. In this chapter, we are interested in understanding how large the ratio  $\frac{C^{(\text{FD})}}{C^{(\text{HD})}}$  can be for a wireless network.

## 4.2 Full-Duplex can Triple the Capacity of Wireless Networks

We start with a simple example which illustrates why full-duplex operation can more than double the capacity of the corresponding half-duplex network.

Consider the single-source single-destination point-to-point network with 3 relaying nodes shown in Fig. 4.1. Each link in the given network is marked by its corresponding capacity. Using (4.1), the full-duplex capacity

$$C^{(\text{FD})} = C.$$

On the other hand, to evaluate the upper bound in (4.2) on the half-duplex capacity, we need to find the best transmit/receive schedule for the relay nodes.

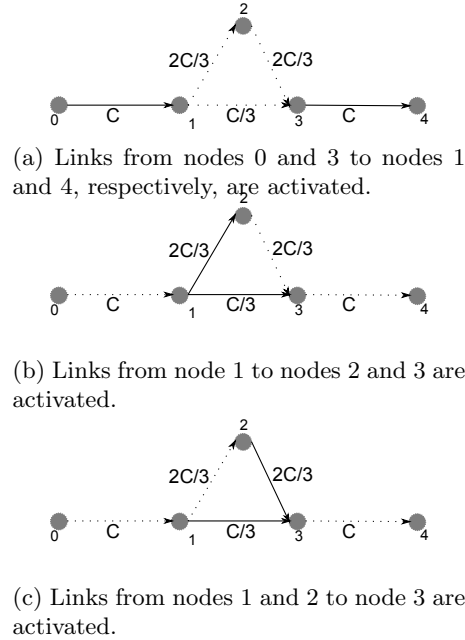


Figure 4.2 – Schedules for the network given in Fig. 4.1.

One can show that in the optimal schedule, the network in Fig. 4.1 will be in only one of the three states shown in Fig. 4.2 with states (a), (b), and (c) operating 0.4, 0.3, and 0.3 fraction of the time, respectively. As such, the upper bound in (4.2) gives

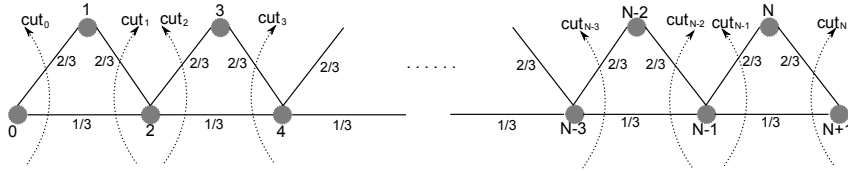
$$C^{(\text{HD})} \leq \frac{2}{5}C + 9.$$

Therefore, for  $C \gg 9$ , we have

$$\frac{C^{(\text{HD})}}{C^{(\text{FD})}} \leq \frac{\frac{2}{5}C + 9}{C} \approx \frac{2}{5} < \frac{1}{2},$$

which means that full-duplex operation indeed more than doubles the capacity of the half-duplex network given in Fig. 4.1. Note that in order to achieve half the capacity of the full-duplex network in the half-duplex network, we would need all the links to be active at least half of the time. Assume we ensure this for all links except for the link from node 1 to node 3, i.e. the remaining 4 links are active at least half of the time. In other words, consider the line network with 4 hops from node 0 to node 4. To activate these 4 links half of the time, we must activate the links in the 1st and the 3rd hop simultaneously half of the time, and activate the links in the 2nd and the 4th hop simultaneously the other half of the time. One can see that such a scheduling deactivates the link from node 1 to node 3 all the time, because nodes 1 and 2 are in transmit/receive mode simultaneously all the time. This conflict in scheduling



Figure 4.3 – Triangle chain network with  $N \geq 1$  relay nodes.

precludes the half-duplex network to achieve half of the full-duplex capacity. Based on this observation, we build a long triangle chain network as shown in Fig. 4.3, which leads to the following proposition.

**Proposition 4.2.1.** *Full-duplex radios can triple the capacity of wireless networks.*

*Proof.* We prove the proposition by showing that there exists an instance of wireless networks where the full-duplex capacity is indeed at least three times the capacity of the corresponding half-duplex network.

Consider a triangle chain network of independent, memoryless, point-to-point channels, where we have  $N \geq 1$  ( $N$  being an odd number) relaying nodes, as shown in Fig. 4.3. The links denoted by  $2/3$  have a capacity of  $(2/3)C$  and those denoted by  $1/3$  have a capacity of  $(1/3)C$ . Based on (4.1), the full-duplex capacity of this network is equal to  $C$ . We want to prove that the half-duplex capacity of the given network is indeed less than or equal to  $1/3$  of the full-duplex capacity.

Let us deactivate all the base links of the triangular subnetworks in the triangle chain network (shown in Fig. 4.3) and only consider the line network with  $N + 1$  hops from node 0 to node  $N + 1$  each of capacity  $(2/3)C$ . By simply activating the odd hops half of the time and the even hops the other half of the time, we can lower bound the half-duplex capacity by

$$C^{\text{HD}} \geq C/3.$$

We will show in what follows that the described relay scheduling is indeed optimal for large values of  $N$ , hence proving that full-duplex operation triples the capacity of this triangle chain network.

For  $1 \leq i \leq (N + 1)/2$ , let  $A_i$  denote the event that the link from node  $2i - 2$  to node  $2i$  is active,  $B_i$  denote the event that the link from node  $2i - 2$  to node  $2i - 1$  is active, and  $C_i$  denote the event that the link from node  $2i - 1$  to node  $2i$  is active. Note that events  $B_i$  and  $C_i$  are disjoint because of the half-duplex constraint at the relay node  $2i - 1$ . Furthermore, events  $A_i \cup C_i$  and  $A_{i+1} \cup B_{i+1}$  are disjoint because of the half-duplex constraint at the relay node  $2i$ . This means that

$$\mathbb{P}(A_i \cap B_i) + \mathbb{P}(A_i \cap C_i) \leq \mathbb{P}(A_i) \quad (4.3)$$

and

$$\mathbb{P}(A_i \cup C_i) + \mathbb{P}(A_{i+1} \cup B_{i+1}) \leq 1. \quad (4.4)$$

Note that the relevant cuts to upper bound the first term in (4.2) are given by the  $N + 1$  cuts shown in Fig. 4.3. To evaluate each cut, we sum over the capacity of each traversed link multiplied with the fraction of time it is active (equivalently, the probability that the link is active), which precisely corresponds to the first term in (4.2). We further normalize the value of each cut by  $C$  to avoid repetition. Thus, we can upper bound these cuts as follows:

$$\begin{aligned} \text{cut}_0 &: \frac{1}{3}\mathbb{P}(A_1) + \frac{2}{3}\mathbb{P}(B_1) \\ &= \frac{2}{3}(\mathbb{P}(A_1) + \mathbb{P}(B_1)) - \frac{1}{3}\mathbb{P}(A_1) \\ &= \frac{2}{3}(\mathbb{P}(A_1 \cup B_1) + \mathbb{P}(A_1 \cap B_1)) - \frac{1}{3}\mathbb{P}(A_1) = t_0 \\ \text{cut}_1 &: \frac{2}{3}(\mathbb{P}(A_1 \cup C_1) + \mathbb{P}(A_1 \cap C_1)) - \frac{1}{3}\mathbb{P}(A_1) = t_1 \\ &\quad \vdots \\ \text{cut}_{N-1} &: \frac{2}{3}(\mathbb{P}(A_{\frac{N+1}{2}} \cup B_{\frac{N+1}{2}}) + \mathbb{P}(A_{\frac{N+1}{2}} \cap B_{\frac{N+1}{2}})) \\ &\quad - \frac{1}{3}\mathbb{P}(A_{\frac{N+1}{2}}) = t_{N-1} \\ \text{cut}_N &: \frac{2}{3}(\mathbb{P}(A_{\frac{N+1}{2}} \cup C_{\frac{N+1}{2}}) + \mathbb{P}(A_{\frac{N+1}{2}} \cap C_{\frac{N+1}{2}})) \\ &\quad - \frac{1}{3}\mathbb{P}(A_{\frac{N+1}{2}}) = t_N. \end{aligned}$$

Therefore, using (4.2) we can upper bound the half-duplex capacity as follows

$$C^{(\text{HD})} \leq C \times \min_{i \in \{0, \dots, N\}} \{t_i\} + 3N. \quad (4.5)$$

We can further upper bound (4.5) by the average over all the cuts as follows

$$\begin{aligned} C^{(\text{HD})} &\leq \frac{C}{N+1} \sum_{i=0}^N t_i + 3N \\ &= \frac{C}{N+1} \left( \frac{2}{3} \left( \mathbb{P}(A_1 \cup B_1) + \mathbb{P}(A_{\frac{N+1}{2}} \cup C_{\frac{N+1}{2}}) \right) \right. \\ &\quad \left. + \frac{2}{3} \sum_{i=1}^{\frac{N-1}{2}} (\mathbb{P}(A_i \cup C_i) + \mathbb{P}(A_{i+1} \cup B_{i+1})) \right. \\ &\quad \left. + \frac{2}{3} \sum_{i=1}^{\frac{N+1}{2}} (\mathbb{P}(A_i \cap B_i) + \mathbb{P}(A_i \cap C_i)) - \frac{2}{3} \sum_{i=1}^{\frac{N+1}{2}} \mathbb{P}(A_i) \right) + 3N. \end{aligned}$$

Using the inequalities in (4.3) and (4.4), we get

$$\begin{aligned} C^{(\text{HD})} &\leq \frac{C}{N+1} \left( \frac{4}{3} + \frac{2}{3} \frac{N-1}{2} + \frac{2}{3} \sum_{i=1}^{\frac{N+1}{2}} \mathbb{P}(A_i) - \frac{2}{3} \sum_{i=1}^{\frac{N+1}{2}} \mathbb{P}(A_i) \right) + 3N \\ &= \frac{C}{N+1} \left( \frac{2}{3} + \frac{N+1}{3} \right) + 3N \\ &= \frac{C}{3} + \frac{2}{3} \frac{C}{N+1} + 3N. \end{aligned}$$

Therefore, as  $N \rightarrow \infty$ , the upper bound on the half-duplex capacity is dominated by  $C/3$  if we choose  $C \gg N$  (for example, choose  $C = N^2$ ). This proves that for the triangle chain network given in Fig. 4.3, full-duplex can at least triple the capacity, since

$$\lim_{N \rightarrow \infty} \frac{C^{(\text{HD})}}{C^{(\text{FD})}} \leq \lim_{N \rightarrow \infty} \frac{1}{3} + \frac{2}{3} \frac{1}{N+1} + \frac{3N}{C} = \frac{1}{3},$$

which concludes the proof.  $\square$

### 4.3 Full-Duplex Operation Cannot More Than Quadruple the Capacity of a Half-Duplex Network

The goal of this section is to prove the following proposition.

**Proposition 4.3.1.** *Full-duplex radios cannot more than quadruple the capacity of single-input single-output wireless networks of independent, noisy (for example AWGN) point-to-point channels.*

*Proof.* We start our proof by showing that for any single-source single-destination network there exists a fixed deterministic schedule that ensures that all links are active  $\frac{1}{4}$  of the time. We prove the existence of this schedule by the probabilistic method.

Consider communication over  $T$  time-slots and construct a scheduling table  $S$  with  $N$  columns and  $T$  rows, where the rows are enumerated by the time-slots  $1 \leq i \leq T$  and the columns are enumerated by the nodes in the network  $1 \leq j \leq N$ . Each entry  $S(i, j)$  is a binary value that represents the transmission mode of the  $j$ -th node in the  $i$ -th time slot; 1 if the node is in the transmit-mode and 0 if it is in the receive-mode.

We construct a schedule randomly by filling each entry with a 0 or 1 by fair coin flips. The edge  $e_{u \rightarrow v}$  going from node  $u$  to node  $v$  is active at time slot  $i$ , if node  $u$  is in transmit-mode and node  $v$  is in receive-mode, i.e. if  $S(i, u) = 1$  and  $S(i, v) = 0$ . Denote by  $N_{e_{u \rightarrow v}}$  the number of times the edge  $e_{u \rightarrow v}$  is active over the  $T$  time-slots. Our random schedule construction for half-duplex networks ensures, for any  $\epsilon > 0$ ,

$$\mathbb{P} \left( N_{e_{u \rightarrow v}} \geq (1 - \epsilon) \frac{T}{4} \right) \rightarrow 1$$

as  $T$  gets large by the law of large numbers. We choose  $T$  large enough such that

$$\mathbb{P}\left(N_{e_{u \rightarrow v}} \geq (1 - \epsilon)\frac{T}{4}\right) > 1 - \frac{1}{M}, \quad (4.6)$$

where  $M$  denotes the number of edges in the network. Assume the edges are numbered from 1 to  $M$ . For  $1 \leq m \leq M$ , let  $e_m$ ,  $E_m$ , and  $E_m^c$  denote the  $m$ -th edge, the event  $\{N_{e_m} \geq (1 - \epsilon)T/4\}$ , and its complement event, respectively. Using (4.6) and the union bound, we get

$$\begin{aligned} \mathbb{P}\left(\bigcap_{m=1}^M E_m\right) &\geq 1 - \sum_{m=1}^M \mathbb{P}\left(E_m^c\right) = 1 - \sum_{m=1}^M (1 - \mathbb{P}(E_m)) \\ &> 1 - M\left(1 - \left(1 - \frac{1}{M}\right)\right) = 0. \end{aligned}$$

Therefore, there exists a *deterministic* schedule with  $\hat{T} \leq T$  time slots for which all the edges of the network are active for more than  $(1 - \epsilon)/4$  fraction of the time, where  $\epsilon$  here can be taken arbitrarily small.

Applying this deterministic schedule to the network over  $\hat{T}$  time-slots, we obtain a network where each link in the network is active only in certain time slots but in a total of at least  $(1 - \epsilon)\hat{T}/4$  time-slots. From a capacity perspective, such a network is equivalent to a network where links are composed of  $\hat{T}$  frequency bands but each link has access to only a subset of these frequencies, a total of at least  $(1 - \epsilon)\hat{T}/4$ . This is because from a capacity perspective we can interchange the independent channels over time with independent channels over different frequency bands provided that we normalize the capacity of the second network by  $\hat{T}$ . Since we know that every link has access to at least  $(1 - \epsilon)\hat{T}/4$  independent frequency bands, the capacity of this network, when normalized by  $\hat{T}$  is lower bounded by a fraction of  $(1 - \epsilon)/4$  of the capacity of the original full-duplex network. This implies that the deterministic schedule activating each link at least a fraction of  $(1 - \epsilon)/4$  of the time can achieve a rate in the half-duplex network which is at least as large as  $(1 - \epsilon)/4$  fraction of the full-duplex capacity. Noting that  $\epsilon$  can be made arbitrarily small concludes the proof of the proposition.  $\square$

### 4.3.1 Universal Deterministic Schedule

We next construct an explicit deterministic schedule which given any network topology with  $N$  nodes ensures that all the links in the network are active exactly  $1/4$  fraction of the time. Moreover, the number of time slots required by this schedule increases linearly with the number of nodes in the network. This is equivalent to saying that the number of active states in this schedule is linear in  $N$ . Note that it has been recently shown in [5, 9], the optimal half-duplex schedule in any Gaussian network has only  $N + 1$  states, however no efficient algorithms are known for finding this schedule.

Our schedule requires  $T = 2^{\lceil \log_2(2N) \rceil}$  time slots, which means that  $N < T \leq 2N$ . It is constructed as follows: for a positive integer  $m$  (we will see how we choose the value of  $m$  later), construct a  $2^m$ -by- $m$  binary matrix by listing all possible binary sequences of length  $m$  as its  $2^m$  rows. Call the columns of this matrix  $X_1, \dots, X_m$ . Further, construct a  $2^m$ -by- $(2^m - 1)$  matrix  $S$ , the columns of which are indexed by all non-empty subsets of  $\{1, \dots, m\}$ . The column indexed by the subset  $Q$  is equal to  $\sum_{i \in Q} X_i \pmod{2}$ .

**Lemma 4.3.2.** *Such a construction ensures that any two columns of the matrix  $S$  have the property that the combinations  $(0, 0)$ ,  $(0, 1)$ ,  $(1, 0)$ , and  $(1, 1)$  occur equal number of times, i.e.  $2^{m-1}/4$  times.*

*Proof.* In fact, the columns of matrix  $S$  constitute all the codewords, except for the all-zero codeword, of a Hadamard code with block length  $2^m$ . Every non-zero codeword in this codebook has a Hamming weight of  $2^{m-1}$ . Equivalently, every column in matrix  $S$  has equal number  $(2^{m-1})$  of zeros and ones. Consider any two columns, denoted by  $Y_1$  and  $Y_2$ , in matrix  $S$ . We want to prove that

$$\begin{aligned} & \#\{i : Y_1(i) = 0, Y_2(i) = 0\} \\ &= \#\{i : Y_1(i) = 0, Y_2(i) = 1\} \\ &= \#\{i : Y_1(i) = 1, Y_2(i) = 0\} \\ &= \#\{i : Y_1(i) = 1, Y_2(i) = 1\} = 2^{m-1}/4. \end{aligned} \quad (4.7)$$

Let  $T = 2^m$  and  $Y_3 = Y_1 \oplus Y_2$ . Note that  $Y_3$  must also be a column in  $S$ . Since every column in  $S$  has equal number of zeros and ones, we have

$$\begin{aligned} \#\{i : Y_1(i) = 0\} &= \frac{T}{2} \\ &= \#\{i : Y_1(i) = 0, Y_2(i) = 0\} \\ &\quad + \#\{i : Y_1(i) = 0, Y_2(i) = 1\}, \end{aligned} \quad (4.8)$$

$$\begin{aligned} \#\{i : Y_1(i) = 1\} &= \frac{T}{2} \\ &= \#\{i : Y_1(i) = 1, Y_2(i) = 0\} \\ &\quad + \#\{i : Y_1(i) = 1, Y_2(i) = 1\}, \end{aligned} \quad (4.9)$$

$$\begin{aligned} \#\{i : Y_3(i) = 0\} &= \frac{T}{2} \\ &= \#\{i : Y_1(i) = 0, Y_2(i) = 0\} \\ &\quad + \#\{i : Y_1(i) = 1, Y_2(i) = 1\}, \end{aligned} \quad (4.10)$$

$$\begin{aligned} \#\{i : Y_3(i) = 1\} &= \frac{T}{2} \\ &= \#\{i : Y_1(i) = 1, Y_2(i) = 0\} \\ &\quad + \#\{i : Y_1(i) = 0, Y_2(i) = 1\}. \end{aligned} \quad (4.11)$$

Solving (4.8), (4.9), (4.10), and (4.11), leads to the result in (4.7).  $\square$

t	{1}	{2}	{3}	{1,2}	{1,3}	{2,3}	{1,2,3}
<b>1</b>	0	0	0	0	0	0	0
<b>2</b>	0	0	1	0	1	1	1
<b>3</b>	0	1	0	1	0	1	1
<b>4</b>	0	1	1	1	1	0	0
<b>5</b>	1	0	0	1	1	0	1
<b>6</b>	1	0	1	1	0	1	0
<b>7</b>	1	1	0	0	1	1	0
<b>8</b>	1	1	1	0	0	0	1

Table 4.1 – Universal deterministic schedule with 8 time slots for networks of  $N \in \{4, 5, 6, 7\}$  nodes.

As before, the entry  $S(i, j)$  represents the transmission mode of the  $j$ -th node in the  $i$ -th time slot; 1 if the node is in transmit-mode and 0 if it is in receive-mode. Therefore, the matrix  $S$  corresponds to a good schedule with  $T = 2^m$  for networks of at most  $2^m - 1$  nodes. Given the number of nodes  $N$ , we choose  $m$  just large enough to ensure  $N \leq 2^m - 1$ . Equivalently, we pick  $m$ , such that  $2^{m-1} \leq N < 2^m$ , i.e.  $m - 1 = \lfloor \log_2 N \rfloor$ . Finally, we pick the submatrix with the first  $N$  columns of  $S$  (note that we can pick any  $N$  columns of the matrix  $S$ ) and  $2^m$  rows as the schedule for the  $N$  half-duplex nodes. Therefore, based on Lemma 4.3.2, the universal deterministic schedule proposed ensures that all the links in any network are “used properly” exactly  $1/4$  fraction of the time.

Let us discuss the example of universal deterministic schedule given in Table 4.1 for networks of  $N \in \{4, 5, 6, 7\}$  nodes. Assume we have a network with  $N = 5$  nodes. This means that  $m = \lfloor \log_2 N \rfloor + 1 = 3$ . We construct a  $2^3$ -by-3 binary matrix by listing the possible combinations of 3 bits as the  $2^3$  rows. Call the columns of this matrix  $X_1$ ,  $X_2$ , and  $X_3$ . Afterwards, we construct a  $2^3$ -by- $(2^3 - 1)$  matrix  $S$ , the columns of which are indexed by all non-empty subsets of  $\{1, 2, 3\}$ . The column corresponding to the subset  $Q$  is equal to  $\sum_{i \in Q} X_i \bmod 2$ , as shown in Table 4.1. Note that any two columns of the matrix  $S$  contain the combinations  $(0, 0)$ ,  $(0, 1)$ ,  $(1, 0)$ , and  $(1, 1)$  exactly 2 times. Since we have only 5 nodes in the network, we pick the submatrix with the first  $N = 5$  columns and the  $T = 8$  rows as the schedule for our half-duplex network.

#### 4.4 Summary

In this chapter, we show that full-duplex can more than double the capacity of wireless networks. Indeed, we construct an instance of wireless network where the ratio between the full-duplex and the half-duplex capacity is 3. The follow-up question that remains open is whether there exists an instance of wireless network where full-duplex can more than triple the capacity of the half-duplex network. This might require a better understanding of the optimal

---

scheduling in half-duplex multi-relay networks, which can give an insight about how to construct networks that exploit the weak points of half-duplex nodes to minimize the half-duplex capacity. Moreover, we propose a universal schedule for half-duplex networks that ensures all links are active  $1/4$  of the time. As a consequence, we show that full-duplex radios cannot more than quadruple the capacity of wireless networks of independent, memoryless, point-to-point channels.





---

## Conclusions and Future Directions

---

# 5

In this chapter, we summarize the main contributions of this dissertation and briefly discuss some open problems and possible future research directions.

### 5.1 Impact of Communication Medium on Network Throughput

The first main contribution of the thesis, presented in Chapters 2 and 3, characterizes the broadcast capacity of a wireless network at low SNR in line-of-sight environment and under different assumptions regarding the network density. Our analysis proves that to achieve the broadcast capacity at low SNR, the nodes available in the network are required to collaborate and efficiently exploit the beamforming capabilities available in the network when the network is dense. On the other hand, sparse networks lead to uncorrelated channel coefficients which results in the absence of beamforming capabilities in the network, which in turn means that no cooperation is required by the nodes to achieve the broadcast capacity.

Another contribution related to the impact of the communication medium on the network throughput resides in the following observation: an increase in the number of degrees of freedom comes at the expense of the beamforming capabilities in the network. In other words, sparse networks diminish the correlation among the channel coefficients which on one hand improves the spatial diversity, thus, increasing the number of degrees of freedom available, and on the other hand reduces the beamforming gain of the channel matrix. Equivalently, dense networks provide higher correlation among the phases of the transmitted signals which results in larger beamforming capabilities and smaller degrees of freedom.

### 5.1.1 Future Research Directions

In this dissertation, we show the impact of the beamforming gain on the broadcast capacity of ad hoc networks at low-SNR regime. We further provide an achievable scheme, namely, multi-stage back-and-forth beamforming, which exploits the beamforming gain in an optimal fashion to achieve the broadcast capacity. Furthermore, we present previous works on the multi-unicast capacity in ad hoc networks at high-SNR regime. In these works, authors propose sophisticated schemes, based on hierarchical cooperation, that can fully benefit from the spatial degrees of freedom available to achieve the multi-unicast capacity.

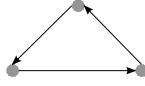
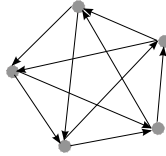
A scenario that may require a communication scheme that exploits both the beamforming gain and the degrees of freedom is the multi-unicast capacity at low SNR, which still remains an open problem. This is because at low SNR the multi-unicast capacity is given by the power transfer across the network. This may require an optimal power transfer (equivalently an optimal beamforming scheme) along each dimension of the communication medium. In other words, we suspect that the optimal communication scheme should be capable of combining strategies used by the hierarchical cooperation (MIMO transmission) with strategies used by the back-and-forth beamforming scheme, if it were to achieve the multi-unicast capacity at low-SNR regime. Such a hybrid scheme is still unknown and subject to future investigation.

## 5.2 Impact of Transmission Mode on Network Throughput

The second part of the thesis, presented in Chapter 4, emphasizes the impact of the transceiving capabilities of the nodes, rather than the communication channel, on the network throughput. We show that full-duplex can more than double the capacity of wireless networks. In fact, we construct an instance of wireless network where the full-duplex capacity is three times the half-duplex capacity. We further propose a universal schedule for half-duplex networks of independent, memoryless, point-to-point channels that is capable of achieving a fraction of  $1/4$  of the full-duplex capacity. In other words, we show that full-duplex radios cannot more than quadruple the capacity of wireless networks of independent, memoryless, point-to-point channels.

### 5.2.1 Future Research Directions

Note that in this work, we consider single-source single-destination networks. Consequently, we consider acyclic networks. However, the schedule we propose in 4.3.1 guarantees that all links are active  $1/4$  fraction of the time in both cyclic and acyclic networks. This may suggest that we may have better schedules for acyclic networks.

Figure 5.1 – Cyclic network with  $N = 3$  nodes.Figure 5.2 – Cyclic network with  $N = 5$  nodes.

### Cyclic vs Acyclic Networks

Consider the triangular cyclic network shown in Fig. 5.1. If any of the three links is activated, then none of the other two links can be activated simultaneously. As a consequence, each link can be activated only  $1/3$  fraction of the time to ensure equal data transfer through all the links. We can extend the triangular cyclic network to a pentagonal cyclic network as shown in Fig. 5.2. In this network, at most 3 links can be activated simultaneously. As a result, to ensure equal data transfer through all the 10 links, each link should be activated only  $3/10$  fraction of the time.

One can generalize the topology of cyclic networks shown in Fig.5.1 and Fig.5.2 as follows. Consider a network with  $N$  (assume  $N \gg 1$  is odd) nodes placed on a circle. Each node has  $\frac{N-1}{2}$  outgoing links connected respectively to the  $\frac{N-1}{2}$  neighboring nodes that follow it on the circle. In this cyclic network, to ensure equal data transfer through all the  $\frac{N(N-1)}{2}$  links, each link should be activated only  $1/4$  fraction of the time for sufficiently large  $N$ . *Proof Approach:* One can show that, in order to simultaneously activate the largest number of links in the given network,  $\frac{N-1}{2}$  nodes should be in transmit-mode and  $\frac{N+1}{2}$  of the remaining nodes should be in receive-mode (or vice versa). Moreover, we can randomly pick the nodes that are in transmit-mode. Therefore, without loss of generality, assume the first  $\frac{N-1}{2}$  nodes are in transmit-mode and the rest of the nodes are in receive-mode. Note that for the first transmitting node, only one link is activated. For the second transmitting node, two links are activated. This trend continues until the last transmitting node which has all its  $\frac{N-1}{2}$  outgoing links activated. This results in a total of

$$1 + 2 + \dots + \frac{N-1}{2} = \frac{N^2 - 1}{8}$$

links activated out of  $\frac{N(N-1)}{2}$  links. Therefore, for large values of  $N$ , at any

instance, we can activate

$$\frac{\frac{N^2-1}{8}}{\frac{N(N-1)}{2}} = \frac{1}{4} \frac{N+1}{N} \xrightarrow{N \rightarrow \infty} 1/4$$

fraction of the links in the network simultaneously. This means that, to ensure equal data transfer through all the links, each link should be activated 1/4 fraction of the time. This may suggest that the maximum capacity gain that full-duplex radios can achieve is indeed 4. However, this may require to come up with with a communication scenario<sup>1</sup>, whereby all the links in the above described topology are indeed readily required to be active for an equal fraction of the time.

On the other hand, the maximum gain of full-duplex networks over half-duplex networks may vary between cyclic and acyclic networks. In other words, the maximum gain might be equal to 3 for acyclic or single-source single-destination networks, and equal to 4 for cyclic or multiple-source multiple-destination networks. All these claims and suggestions are plausible and worth the investigation.

---

<sup>1</sup>For example, we may have to consider multiple-source multiple-destination networks instead of single-source single-destination networks.

## Bibliography

---

- [1] R. Ahlswede, N. Cai, S. Li, and R. Yeung, “Network information flow,” *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.
- [2] A. Avestimehr, S. Diggavi, and D. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Trans. Inform. Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.
- [3] N. Bansal and Z. Liu, “Capacity, delay and mobility in wireless ad-hoc networks,” *IEEE Infocom, 22nd International Conference on Computer Communications*, pp. 1553–1563, March 2003.
- [4] S. Brahma, C. Fragouli, and A. Özgür, “Computing half-duplex schedules in gaussian relay networks via min-cut approximations,” *IEEE Trans. Inform. Theory*, vol. 60, no. 11, pp. 7204–7220, November 2014.
- [5] —, “On the complexity of scheduling in half-duplex diamond networks,” *IEEE Trans. Inform. Theory*, vol. 62, no. 5, pp. 2557–2572, May 2016.
- [6] V. Cadambe and S. Jafar, “Interference alignment and degrees of freedom of the k-user interference channel,” *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3425–3441, August 2008.
- [7] C. Campolo, C. Casetti, C. F. Chiasserini, and S. Tarapiah, “Performance of network coding for ad hoc networks in realistic simulation scenarios,” *International Conference on Telecommunications*, pp. 31–36, May 2009.
- [8] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, “The symmetric sum-capacity of the gaussian half-duplex causal cognitive interference channel to within a constant gap,” *Proceedings of the IEEE International Symposium on Information Theory*, pp. 754–758, July 2013.
- [9] —, “On the optimality of simple schedules for networks with multiple half-duplex relays,” *IEEE Trans. Inform. Theory*, vol. 62, no. 7, pp. 4120–4134, July 2016.
- [10] M. M. Carvalho and J. J. Garcia-Luna-Aceves, “Analytical modeling of ad hoc networks that utilize space-time coding,” *4th International Symposium*

- on *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp. 1–11, April 2006.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd edition*. Wiley, 2006.
- [12] M. Desgroseilliers, O. Lévêque, and E. Preissmann, “Spatial degrees of freedom of mimo systems in line-of-sight environment,” *Proceedings of the IEEE International Symposium on Information Theory*, pp. 834–838, July 2013.
- [13] —, “Partially random matrices in line-of-sight wireless networks,” *IEEE Asilomar Conference on Signals, Systems and Computers*, pp. 1026–1030, November 2013.
- [14] S. N. Diggavi, M. Grossglauser, and D. N. C. Tse, “Even one-dimensional mobility increases the capacity of wireless networks,” *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3947–3954, November 2005.
- [15] C. Fragouli, J. Widmer, and J. Y. L. Boudec, “Efficient broadcasting using network coding,” *IEEE/ACM Transactions on Networking*, vol. 16, no. 2, pp. 450–463, April 2008.
- [16] C. Fragouli, J.-Y. Le Boudec, and J. Widmer, “Network coding: An instant primer,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 63–68, January 2006.
- [17] M. Franceschetti, M. Migliore, and P. Minero, “The capacity of wireless networks: Information-theoretic and physical limits,” *IEEE Trans. Inform. Theory*, vol. 55, no. 8, pp. 3413–3424, August 2009.
- [18] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, “Optimal throughput-delay scaling in wireless networks - part i: the fluid model,” *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2568–2592, June 2006.
- [19] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.
- [20] J. Ghaderi, L. L. Xie, and X. Shen, “Hierarchical cooperation in ad hoc networks: Optimal clustering and achievable throughput,” *IEEE Trans. Inform. Theory*, vol. 55, no. 8, pp. 3425–3436, August 2009.
- [21] M. Grossglauser and D. N. C. Tse, “Mobility increases the capacity of ad hoc wireless networks,” *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, August 2002.
- [22] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, March 2000.

- [23] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.
- [24] X. Hong, K. Xu, and M. Gerla, "Scalable routing protocols for mobile ad hoc networks," *IEEE Network*, vol. 16, no. 4, pp. 11–21, July 2002.
- [25] S. W. Jeon and S. Y. Chung, "Two-phase opportunistic broadcasting in large wireless networks," *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2771–2775, June 2007.
- [26] A. Keshavarz-Haddad and R. Riedi, "On the broadcast capacity of multihop wireless networks: Interplay of power, density and interference," *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 314–323, June 2007.
- [27] A. Keshavarz-Haddad, V. Ribeiro, and R. Riedi, "Broadcast capacity in multihop wireless networks," *Proceedings of the 12th Annual International Conference on Mobile computing and networking, MobiCom'06*, pp. 239–350, September 2006.
- [28] R. Koetter, M. Effros, and M. Medard, "A theory of network equivalence – part i: Point-to-point channels," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 972–995, Feb 2011.
- [29] G. Kramer, "Capacity of deterministic, half-duplex, line networks with two sources," pp. 1–2, April 2015.
- [30] —, "Models and theory for relay channels with receive constraints," *Proc. 42th Annu. Allerton Conf. Commun., Control, Comp.*, pp. 1312–1321, September 2004.
- [31] S. H. Lee and S. Y. Chung, "Capacity scaling of wireless ad hoc networks: Shannon meets maxwell," *IEEE Trans. Inform. Theory*, vol. 58, no. 3, pp. 1702–1715, March 2012.
- [32] L. Li, R. Ramjee, M. Buddhikot, and S. Miller, "Network coding-based broadcast in mobile ad-hoc networks," *IEEE Infocom, 26th International Conference on Computer Communications*, pp. 1739–1747, May 2007.
- [33] X. Y. Li, J. Zhao, Y.-W. Wu, S. J. Tang, X.-H. Xu, and X.-F. Mao, "Broadcast capacity for wireless ad hoc networks," *IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, pp. 114–123, September 2008.
- [34] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with brownian mobility," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2777–2784, June 2006.

- [35] A. Merzakreeva, "Cooperation in space-limited wireless networks at low snr," Ph.D. dissertation, EPFL, Switzerland, 2014.
- [36] A. Merzakreeva, O. Lévêque, and A. Özgür, "Hierarchical beamforming for large one-dimensional wireless networks," *Proceedings of the IEEE International Symposium on Information Theory*, pp. 1533–1537, July 2012.
- [37] —, "Telescopic beamforming for large wireless networks," *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2771–2775, July 2013.
- [38] A. Motahari, S. Oveis-Gharan, M.-A. Maddah-Ali, and A. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Trans. Inform. Theory*, vol. 60, no. 8, pp. 4799–4810, August 2014.
- [39] B. Nazer, M. Gastpar, S. Jafar, and S. Vishwanath, "Ergodic interference alignment," *IEEE Trans. Inform. Theory*, vol. 58, no. 10, pp. 6355–6371, October 2012.
- [40] M. J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad hoc mobile networks," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1917–1937, June 2005.
- [41] U. Niesen, P. Gupta, and D. Shah, "On capacity scaling in arbitrary wireless networks," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 3959–3982, September 2009.
- [42] A. Özgür, R. Johari, D. Tse, and O. Lévêque, "Information theoretic operating regimes of large wireless networks," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 427–437, January 2010.
- [43] A. Özgür, O. Lévêque, and E. Preissmann, "Scaling laws for one- and two-dimensional random wireless networks in the low-attenuation regime," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3573–3585, October 2007.
- [44] A. Özgür and S. N. Diggavi, "Approximately achieving gaussian relay network capacity with lattice-based qmf codes," *IEEE Trans. Inform. Theory*, vol. 59, no. 12, pp. 8275–8294, December 2013.
- [45] A. Özgür, O. Lévêque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3549–3572, October 2007.
- [46] —, "Spatial degrees of freedom of large distributed mimo systems and wireless ad hoc networks," *IEEE Journ, on Selected Areas in Communications*, vol. 31, no. 2, pp. 202–214, February 2013.
- [47] G. Sharma, R. Mazumdar, and N. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 981–992, October 2007.



- [48] W. Y. Shin, S. W. Jeon, N. Devroye, M. H. Vu, S. Y. Chung, Y. H. Lee, and V. Tarokh, "Improved capacity scaling in wireless networks with infrastructure," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 5088–5102, August 2011.
- [49] B. Sirkeci-Mergen and M. Gaspar, "On the broadcast capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 56, no. 8, pp. 3847–3861, August 2010.
- [50] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, March 1998.
- [51] B. Tavli, "Broadcast capacity of wireless networks," *IEEE Communications Letters*, vol. 10, no. 2, pp. 68–69, February 2006.
- [52] P. Uthansakul, M. E. Bialkowski, S. Durrani, K. Bialkowski, and A. Postula, "Effect of line of sight propagation on capacity of an indoor mimo system," *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, vol. 2B, pp. 707–710, July 2005.
- [53] B. Vucetic and J. Yuan, *Space-Time Coding*. Wiley, 2003.
- [54] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.
- [55] M. Yuksel and E. Erkip, "Multiple-antenna cooperative wireless systems: A diversity–multiplexing tradeoff perspective," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3371–3393, October 2007.
- [56] R. Zheng, "Information dissemination in power-constrained wireless networks," *IEEE Infocom, 25th International Conference on Computer Communications*, pp. 1–10, April 2006.



# Curriculum Vitae

---

## Serj Haddad

*Date of Birth:* September 1, 1988

*Nationality:* Lebanese

*Email:* serj.haddad@epfl.ch

### Education

- Sept. 2012–Mar. 2017:  
**PhD in Information Theory**, École Polytechnique Fédérale de Lausanne (EPFL), *Advisor:* Prof. Emre Telatar, *Co-advisor:* Dr. Olivier Lévêque
- Sept. 2010–Jun. 2012:  
**Master's Degree of Science in Computer Engineering**, Lebanese American University, Cumulative GPA: 4.0/4.0
- Sept. 2006–Jun. 2010:  
**Bachelor's Degree in Computer Engineering**, Lebanese American University, Cumulative GPA: 3.96/4.0

### Professional Experience

- Sept. 2012–Mar. 2017:  
**Doctoral Research Assistant**, Information Theory Laboratory (LTHI), EPFL, Lausanne
- Sept. 2012–Mar. 2017:  
**Teaching Assistant**, School of Communications and Computer Sciences, EPFL, Lausanne
- Feb. 2016–Jun. 2016:  
**Visiting Student Researcher**, Information Systems Laboratory, *Advisor:* Prof. Ayfer Özgür, Stanford University, California

- Sept. 2010–Jun. 2012:  
**M.Sc. Thesis**, Department of Electrical and Computer Engineering,  
*Title: "Cooperative FSO Systems"*, LAU, Byblos
- Sept. 2010–Jun. 2012:  
**Part-time Faculty**, Department of Electrical and Computer Engineering,  
LAU, Byblos

## Publications

- S. Haddad, A. Özgür, and E. Telatar, *Can full-duplex more than double the capacity of wireless networks?*, submitted to IEEE International Symposium on Information Theory.
- S. Haddad and O. Lévêque, *On the broadcast capacity scaling of large wireless networks at low snr*, to appear in IEEE Transactions on Information Theory.
- S. Haddad and O. Lévêque, *On the impact of sparsity on the broadcast capacity of wireless networks*, Proceedings of the IEEE International Symposium on Information Theory, pages 2634-2638, July 2016.
- S. Haddad and O. Lévêque, *Communication tradeoffs in wireless networks*, [online] Available: <http://arxiv.org/pdf/1601.06405.pdf>.
- S. Haddad and O. Lévêque, *On the broadcast capacity of large wireless networks at low snr*, Proceedings of the IEEE International Symposium on Information Theory, pages 171-175, June 2015.
- S. Haddad and O. Lévêque, *Diversity analysis of free-space optical networks with multihop transmissions*, Proceedings of the IEEE International Conference on Communications, pages 3406-3411, June 2014.
- C. Abou-Rjeily and S. Haddad, *Inter-relay cooperation: A new paradigm for enhanced relay-assisted fso communications*, IEEE Transactions on Communications, 62(6):1970-1982, June 2014.
- C. Abou-Rjeily and S. Haddad, *Novel cooperation strategies for free-space optical communication systems in the absence and presence of feedback*, Proceedings Part II of DICTAP, Dijon, France, pages 543-556, June 2011.
- C. Abou-Rjeily and S. Haddad, *Cooperative fso systems: Performance analysis and optimal power allocation*, IEEE Journal of Lightwave Technology, 29(7):1058-1065, April 2011.

