

# On the Broadcast Capacity Scaling of Large Wireless Networks at Low SNR

Serj Haddad, Olivier Lévêque

Information Theory Laboratory – School of Computer and Communication Sciences  
EPFL, 1015 Lausanne, Switzerland – {serj.haddad,olivier.leveque}@epfl.ch

## Abstract

The present paper focuses 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 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.

## Index Terms

wireless networks, broadcast capacity, low SNR communications, beamforming strategies, random matrices

## I. INTRODUCTION

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. [1], [2], [3]). 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 or to enhance cooperation among nodes belonging to the same cluster or cell. In the present paper, 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 [4], [5], [6], [7], [8]. 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 multi-hopping 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.

In the present paper, we address the low SNR regime and consider the line-of-sight (LOS) propagation model ([9], [10]) described in Section II below. In this regime, the power available does not allow for a source node to successfully transmit a message to its nearest neighbour without waiting for some amount of time in order to spare power. This raises an interesting question, as a naive cut-set argument 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 multi-hop broadcasting) can be shown to be asymptotically optimal.

This issue was first revealed in [10] in the context of one-dimensional networks, under the LOS model. For such networks, the authors proposed a hierarchical beamforming scheme to broadcast data to the network, that 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 [11], where a strategy was developed to enhance multiple-unicast communications in wireless networks under the LOS model. Taking inspiration from this paper, we propose below a new multi-stage beamforming scheme which is shown to achieve asymptotically optimal<sup>1</sup> performance for broadcasting information in a two-dimensional wireless network.

This paper was presented at ISIT 2015.

<sup>1</sup>Technically speaking, the performance of our scheme is shown below to be asymptotically optimal up to a multiplicative factor  $n^\varepsilon$ , where  $\varepsilon$  can be taken arbitrarily small, but fixed.

We give a detailed description of the scheme in Section III, as well as a proof of its optimality in Section IV. 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 IV.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 the 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.

## II. 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  $R_n = n r_n$  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

$$h_{jk} = \sqrt{G} \frac{\exp(2\pi i r_{jk}/\lambda)}{r_{jk}}, \quad (1)$$

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 [12] to be inaccurate in the setting considered in the present paper. 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 [10], 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. This could be the case e.g., in a sensor network with low battery nodes, or in a sparse network with long distances between neighboring nodes. 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 [13], [14]), 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  $\text{SNR}_s = P$ .

## III. BACK-AND-FORTH BEAMFORMING STRATEGY

First note that under the LOS model (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>2</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

<sup>2</sup>We coarsely approximate  $\log P$  by 1 here!

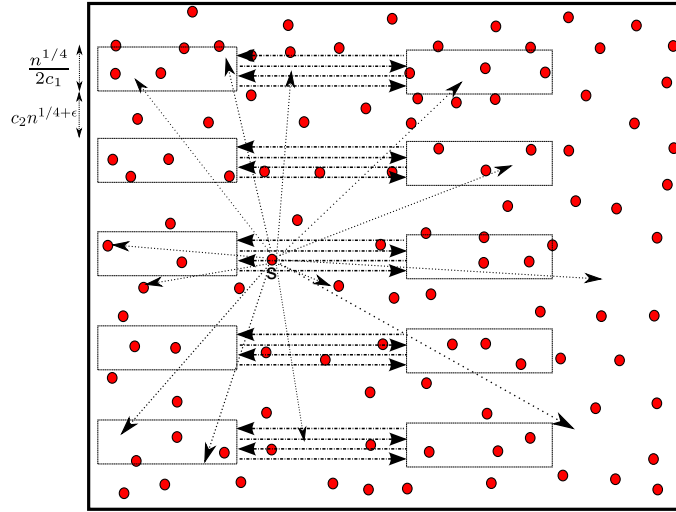


Fig. 1.  $\sqrt{n} \times \sqrt{n}$  network divided into clusters of size  $M = \frac{n^{1/4}}{2c_1} \times \frac{n^{1/2}}{4}$ . Two clusters of size  $M$  placed on the same horizontal line and separated by distance  $d = \frac{n^{1/2}}{4}$  pair up and start back-and-forth beamforming. The vertical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

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 III.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. 1. During each transmission, there are  $\Theta(n^{1/4-\epsilon})$  cluster pairs operating in parallel (see below), so  $\Theta(n^{1-\epsilon})$  nodes are communicating in total. The number of rounds needed to serve all nodes must therefore be  $\Theta(n^\epsilon)$ .

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 cluster pairs, as illustrated on Fig. 1. We show below that the broadcast rate between the operating clusters is  $\Theta(n^{1/2}P)$ . Since we only need  $\Theta(n^\epsilon)$  number of rounds to serve all clusters, phase 2 requires  $\Theta(n^{-1/2+\epsilon}P^{-1})$  time slots per message. As such, back-and-forth beamforming achieves a broadcast rate of  $\Theta(n^{1/2-\epsilon}P)$  bits per time slot.

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

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

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

is achievable with high probability<sup>3</sup> in the network. As a consequence, when  $P = \Omega(n^{-1/2})$ , 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 III.1 also says that our scheme requires asymptotically less power to achieve the same performance as plain time-

<sup>3</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.

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.

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 [15], but is also provided in the Appendix for completeness.

**Lemma III.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. 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

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

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 III.3 and Lemma III.4. The proofs are presented in the Appendix.

**Lemma III.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 cluster  $i$ ) 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 III.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 III.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

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

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 III.3 and Lemma III.4 to lower bound

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

For the sake of clarity, we can therefore approximate<sup>4</sup> the expression in (2) 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

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

where

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

Note again that  $\mathbb{E}(|Z^{(s)}|^2) = \Theta(1)$ , and  $Z_j^{(t)}$  is additive white Gaussian noise of variance  $\Theta(1)$ . Finally, note that Lemma III.4 ensures an upper bound on the beamforming gain of the noise signals, i.e.,

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

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

$$\begin{aligned} \mathbb{E} \left( \left( \left( \frac{AM}{d} \right)^t \sqrt{\text{SNR}} X \right)^2 \right) &= \left( \frac{AM}{d} \right)^{2t} \text{SNR} = \Theta(1) \\ \Rightarrow A &= \Theta \left( \frac{d}{M} \text{SNR}^{-\frac{1}{2t}} \right). \end{aligned} \quad (3)$$

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

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

<sup>4</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.

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) is true if and only if  $\text{SNR} = \Omega(N_C/M) = \Omega(n^{-1/2-\epsilon})$  (check eq. (3)), 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.  $\square$

#### IV. 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 IV.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 [16, 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

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

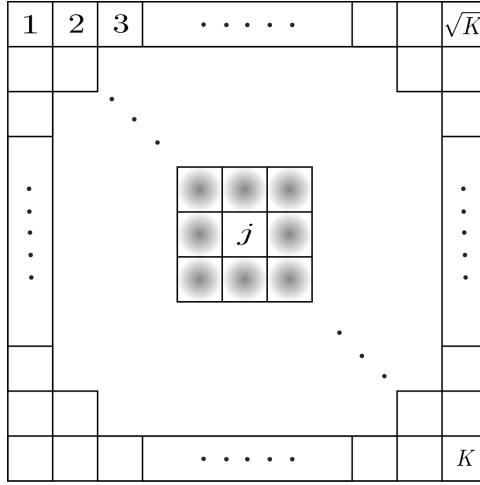


Fig. 2.  $\sqrt{n} \times \sqrt{n}$  network split into  $K$  clusters and numbered in order. As such,  $R_j = \{j - \sqrt{K} - 1, j - \sqrt{K}, j - \sqrt{K} + 1, j - 1, j, j + 1, j + \sqrt{K} - 1, j + \sqrt{K}, j + \sqrt{K} + 1\}$ , which represents the center square containing the cluster  $j$  and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

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

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

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  $\log \det(I + AB) = \log \det(I + BA)$ :

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

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 IV.1 to line-of-sight fading, where the matrix  $H$  is given by

$$h_{jk} = \begin{cases} 0 & \text{if } j = k \\ \frac{\exp(2\pi i r_{jk})}{r_{jk}} & \text{if } j \neq k \end{cases} \quad (4)$$

The rest of the section is devoted to proving the proposition below which, together with Theorem IV.1, shows the asymptotic optimality of the back-and-forth beamforming scheme presented in Section III for two-dimensional networks at low SNR and under LOS fading<sup>5</sup>.

<sup>5</sup>Note that for a one-dimensional network in LOS environment, Theorem IV.1 allows to recover the result already obtained in [10].

**Proposition IV.2.** Let  $H$  be the  $n \times n$  matrix given by (4). For every  $\varepsilon > 0$ , there exists a constant  $c > 0$  such that

$$\|H\|^2 \leq cn^{\frac{1}{2}+\varepsilon}$$

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 IV.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 the Appendix. The second building block of this proof is the following lemma, the proof of which is also given in the Appendix.

**Lemma IV.4.** Let  $\hat{H}$  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

$$\|\hat{H}\|^2 \leq c \frac{M^{1+\varepsilon}}{d}$$

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 IV.2.* The strategy for the proof is now the following: in order to bound  $\|H\|$ , we divide the matrix into smaller blocks, apply Lemma IV.3 and Lemma IV.4 in order to bound the off-diagonal terms  $\|H_{jk}\|$ . For the diagonal terms  $\|H_{jj}\|$ , we reapply Lemma IV.3 and proceed in a recursive 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 IV.3, we obtain

$$\|H\| \leq \max \left\{ \max_{1 \leq j \leq K} \sum_{k=1}^K \|H_{jk}\|, \max_{1 \leq j \leq K} \sum_{k=1}^K \|H_{kj}\| \right\} \quad (5)$$

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 IV.4, if  $d_{jk} \geq 2\sqrt{M}$ , then there exists a constant  $c > 0$  such that

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

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). By the above inequality, we obtain

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

with high probability as  $M$  gets large. Observe that as there are  $8l$  clusters or less at distance  $l\sqrt{M}$  from cluster  $j$ , so 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

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

Using the symmetry of this bound and (5), we obtain

$$\|H\| \leq 9 \max_{1 \leq j \leq K} \|H(R_j)\| + \sqrt{cn^\varepsilon} \frac{n^{3/4}}{M^{1/2}} \quad (6)$$



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 IV.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 subblock  $H_M$  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 (6), 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 IV.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 IV.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 IV.5 successively, we obtain a decreasing sequence of upper bounds on  $\|H\|$ :

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

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 IV.2. □

## V. CONCLUSION

In this work, we characterize the broadcast capacity of two-dimensional wireless networks at low SNR in line-of-sight environment, which is achieved via a back-and-forth beamforming scheme. We showed 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.

## VI. ACKNOWLEDGMENT

S. Haddad's work is supported by Swiss NSF Grant Nr. 200020-156669.

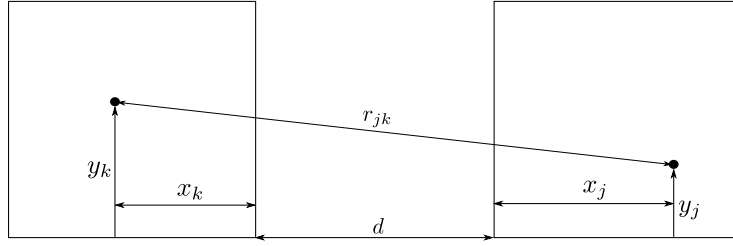


Fig. 3. Coordinate system.

#### APPENDIX

*Proof of Lemma III.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 III.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). 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:

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

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

*Proof of Lemma III.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 these). 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

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

where  $\mathcal{T}'_l$  denotes the  $l$ -th interfering transmit cluster that is at a vertical distance of  $l \left( \frac{n^{1/4}}{2c_1} + c_2 n^{1/4+\epsilon} \right)$  from the desired receiving cluster  $\mathcal{R}_i$ . 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):

$$\begin{aligned} \left| \sum_{k \in \mathcal{T}'_l} \frac{\cos(2\pi(r_{jk} - x_k))}{r_{jk}} \right| &= \left| \sum_{k \in \mathcal{T}'_l} X_k^{(l)} \right| \\ &= \left| \sum_{k \in \mathcal{T}'_l} (X_k - \mathbb{E}(X_k^{(l)})) + \sum_{k \in \mathcal{T}'_l} \mathbb{E}(X_k^{(l)}) \right| \\ &\stackrel{(a)}{\leq} \left| \sum_{k \in \mathcal{T}'_l} (X_k^{(l)} - \mathbb{E}(X_k^{(l)})) \right| + \left| \sum_{k \in \mathcal{T}'_l} \mathbb{E}(X_k^{(l)}) \right| \\ &\stackrel{(b)}{=} M \left| \frac{1}{M} \sum_{k \in \mathcal{T}'_l} (X_k^{(l)} - \mathbb{E}(X_k^{(l)})) \right| + M \left| \mathbb{E}(X_1^{(l)}) \right| \end{aligned} \quad (7)$$

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 bound the second term of (7):  $\forall k \in \mathcal{T}'_l$ , we have

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

is a  $C^2$  function and

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

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

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

which in turn yields the upper bound

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

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

$$\begin{aligned}
\left| \mathbb{E} \left( X_k^{(l)} \right) \right| &= \left| \frac{4}{n^{1/2}} \int_0^{n^{1/4}} dx_k \frac{1}{|y_{k_1} - y_{k_0}|} \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi r_{jk})}{r_{jk}} \right| \\
&\leq \frac{4}{n^{1/2} |y_{k_1} - y_{k_0}|} \int_0^{n^{1/4}} dx_k \left| \int_{y_{k_0}}^{y_{k_1}} dy_k \frac{\cos(2\pi r_{jk})}{r_{jk}} \right| \\
&\leq \frac{9/(2\pi)}{|y_{k_1} - y_{k_0}| l c_2 n^{1/4+\epsilon}} \leq \frac{9c_1}{\pi c_2 l n^{1/2+\epsilon}} = \frac{9c_1}{\pi c_2 l d n^\epsilon}. \tag{8}
\end{aligned}$$

We further upper bound the first term in (7) by using the Hoeffding's inequality [17]. 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) is true if  $t = \frac{1}{d} \sqrt{\frac{2n^\epsilon}{M}}$ . Therefore, we have

$$\left| \frac{1}{M} \sum_{k \in \mathcal{T}'_l} \left( X_k^{(l)} - \mathbb{E} \left( X_k^{(l)} \right) \right) \right| \leq \frac{1}{d} \sqrt{\frac{2n^\epsilon}{M}} \tag{9}$$

with probability  $\geq 1 - 2 \exp(-n^\epsilon)$ . Combining (8) and (9), we can upper bound (7) as follows

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

Finally, we have

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

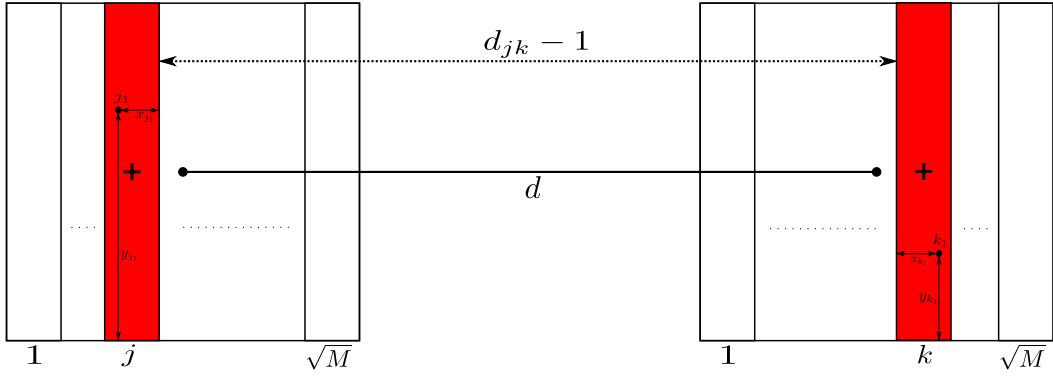


Fig. 4. Two square clusters that have a center-to-center distance  $d$ , with each cluster decomposed into  $\sqrt{M}$  vertical  $\sqrt{M} \times 1$  rectangles.  $d_{jk}$  is distance between the centers (marked with cross) of the two rectangles  $j$  and  $k$ . Moreover, we have the points  $j_1(x_{j_1}, y_{j_1})$  and  $k_1(x_{k_1}, y_{k_1})$  in the rectangles  $j$  and  $k$ , respectively.

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

*Proof of Lemma IV.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

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

by the assumption made above. As  $u \neq 0$ ,  $\|u_j\| > 0$ , so 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 IV.4.* As in the case of  $\|H\|$ , analyzing directly the asymptotic behavior of  $\|\widehat{H}\|$  reveals itself difficult. We therefore decompose our proof into simpler subproblems. The strategy is essentially the following: in order to bound  $\|\widehat{H}\|$ , we divide the matrix into smaller blocks, bound the smaller blocks  $\|\widehat{H}_{jk}\|$ , and apply Lemma IV.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. 4).

By Lemma IV.3, we obtain

$$\|\widehat{H}\| \leq \max \left\{ \max_{1 \leq j \leq \sqrt{M}} \sum_{k=1}^{\sqrt{M}} \|\widehat{H}_{jk}\|, \max_{1 \leq j \leq \sqrt{M}} \sum_{k=1}^{\sqrt{M}} \|\widehat{H}_{kj}\| \right\} \quad (10)$$

where the  $M \times M$  matrix  $\widehat{H}$  is decomposed into blocks  $\widehat{H}_{jk}$ ,  $j, k = 1, \dots, \sqrt{M}$ , with  $\widehat{H}_{jk}$  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. 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

$$\|\widehat{H}_{jk}\|^2 \leq c' \frac{M^\epsilon}{d_{jk}} \leq c \frac{M^\epsilon}{d} \quad (11)$$

with high probability as  $M \rightarrow \infty$ . Applying (10) and (11), we get

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

Therefore, what remains to be proven is inequality (11). The strategy we propose in order to upper bound  $\|\widehat{H}_{jk}\|^2$  is to use the moments' method, relying on the following inequality:

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

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

$$\mathbb{E}(\text{Tr}((\widehat{H}_{jk}\widehat{H}_{jk}^\dagger)^\ell)) \leq \frac{M(c \log M)^{\ell-1}}{d_{jk}^{\ell+1}} \quad (12)$$

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

$$\mathbb{E}(\text{Tr}(FF^\dagger)) = \sum_{j_1, k_1=1}^{\sqrt{M}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^*) = \sum_{j_1, k_1=1}^{\sqrt{M}} \mathbb{E}(|f_{j_1 k_1}|^2) = \sum_{j_1, k_1=1}^{\sqrt{M}} \frac{1}{r_{j_1 k_1}^2} \leq \frac{M}{d_{jk}^2} \quad (13)$$

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

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

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  $S_2$ .

$$S_2 = |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^*)| \\ = \left| \frac{1}{M^2} \int_0^1 dx_{j_1} \int_0^{\sqrt{M}} dy_{j_1} \int_0^1 dx_{j_2} \int_0^{\sqrt{M}} dy_{j_2} \int_0^1 dx_{k_1} \int_0^{\sqrt{M}} dy_{k_1} \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \frac{e^{2\pi i(g_{j_1 j_2}(k_1) + g_{j_2 j_1}(k_2))}}{\rho_{j_1 j_2}(k_1) * \rho_{j_2 j_1}(k_2)} \right|, \quad (14)$$

where

$$g_{j_1 j_2}(k_1) = r_{j_1 k_1} - r_{j_2 k_1} = -g_{j_2 j_1}(k_1) \\ = \sqrt{(d_{jk} - 1 + x_{j_1} + x_{k_1})^2 + (y_{j_1} - y_{k_1})^2} - \sqrt{(d_{jk} - 1 + x_{j_2} + x_{k_1})^2 + (y_{j_2} - y_{k_1})^2} \quad (15)$$

and

$$\rho_{j_1 j_2}(k_1) = r_{j_1 k_1} \cdot r_{j_2 k_1} = \rho_{j_2 j_1}(k_1) \geq d_{jk}^2, \quad (16)$$

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. 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 (15) and (16), we can rewrite (14) as follows

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

where

$$B_{2,1} = \left| \int dk_2 \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| = \left| \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| \\ \leq \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \left| \frac{e^{2\pi i g_{j_2 j_1}(k_2)}}{\rho_{j_2 j_1}(k_2)} \right| \\ = \int_0^1 dx_{k_2} \int_0^{\sqrt{M}} dy_{k_2} \frac{1}{\rho_{j_2 j_1}(k_2)} \leq \frac{\sqrt{M}}{d_{jk}^2} \quad (17)$$

We therefore obtain

$$S_2 \leq \frac{1}{M^{3/2} d_{jk}^2} \int dj_1 \cdot A_{1,2} \quad (18)$$

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 (18), we present the following lemma, taken from [18] and adapted to the present situation.

**Lemma 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

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

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 (18) as follows

$$\begin{aligned} A_{1,2} &= \int dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\ &= \int_{|y_{j_2} - y_{j_1}| < \epsilon\sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| + \int_{|y_{j_2} - y_{j_1}| \geq \epsilon\sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\ &\leq \int_{|y_{j_2} - y_{j_1}| < \epsilon\sqrt{M}} dj_2 \int dk_1 \frac{1}{\rho_{j_1 j_2}(k_1)} + \int_{|y_{j_2} - y_{j_1}| \geq \epsilon\sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\ &\leq \frac{\epsilon M}{d_{jk}^2} + \int_{|y_{j_2} - y_{j_1}| \geq \epsilon\sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \end{aligned} \tag{19}$$

Furthermore, note that

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

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

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



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

$$\begin{aligned} \left| \frac{\partial g_{j_1 j_2}(k_1)}{\partial y_{k_1}} \right| &\geq c_3 \frac{|y_{j_2} - y_{j_1}|}{d_{jk}} - \frac{|y_{k_1} - y_{j_1}| \cdot |x_{j_2} - x_{j_1}|}{d_{jk}^2} \\ &\geq c_3 \frac{|y_{j_2} - y_{j_1}|}{d_{jk}} - \frac{\sqrt{M}}{d_{jk}^2} \\ &\stackrel{(a)}{\geq} \frac{c_3 |y_{j_2} - y_{j_1}| - 1}{d_{jk}}, \end{aligned} \quad (20)$$

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 (16) and (20), we can apply lemma A.1 and upper bound the second term in (19) as follows

$$\begin{aligned} \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \left| \int dk_1 \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| &\leq \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dj_2 \int_0^1 dx_{k_1} \left| \int_0^{\sqrt{M}} dy_{k_1} \frac{e^{2\pi i g_{j_1 j_2}(k_1)}}{\rho_{j_1 j_2}(k_1)} \right| \\ &\leq \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} dy_{j_2} \frac{7}{\pi \frac{c_3 |y_{j_2} - y_{j_1}| - 1}{d_{jk}} d_{jk}^2} \\ &\leq \frac{7}{\pi c_3 d_{jk}} \int_{|y_{j_2} - y_{j_1}| \geq \epsilon \sqrt{M}} \frac{1}{|y_{j_2} - y_{j_1}| - 1/c_3} dy_{j_2} \\ &\leq \frac{7}{\pi c_3 d_{jk}} \log\left(\frac{1}{\epsilon}\right) \end{aligned} \quad (21)$$

which gives the following upper bound on (19)

$$A_{1,2} \leq \frac{\epsilon M}{d_{jk}^2} + \frac{7}{\pi c_3 d_{jk}} \log\left(\frac{1}{\epsilon}\right) \stackrel{(a)}{=} O\left(\frac{\log M}{d_{jk}}\right), \quad (22)$$

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

$$\mathbb{E}(\text{Tr}((FF^\dagger)^2)) \leq 2 \frac{M}{d_{jk}^3} + M^2 S_2 = O\left(M \frac{\log M}{d_{jk}^3}\right). \quad (23)$$

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

**Lemma A.2.** For  $\ell \geq 1$ , let

$$S_\ell = \left| \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* \cdots f_{j_\ell k_\ell} f_{j_1 k_\ell}^*) \right|, \quad (24)$$

with  $j_1 \neq \dots \neq j_\ell$  and  $k_1 \neq \dots \neq k_\ell$ . Note that  $S_\ell$  does not depend on the particular choice of  $j_1 \neq \dots \neq j_\ell$  and  $k_1 \neq \dots \neq k_\ell$ . We can upper bound  $S_\ell$  given in (24) as follows

$$S_\ell \leq \frac{1}{M^{\ell-1} d_{jk}^2} \left( c \frac{\log M}{d_{jk}} \right)^{\ell-1}.$$

*Proof of Lemma A.2.* Note that the lemma is true for  $\ell = 1, 2$ , since  $S_1 \leq \frac{1}{d_{jk}^2}$  and  $S_2 \leq \frac{1}{M d_{jk}^3} \log M$  (as we showed above). In general, for any  $\ell \geq 1$ , we have

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

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 (22) and (17), respectively, we can upper bound  $A_{t,t+1}$  (for  $2 \leq t \leq \ell$ ) and  $B_{\ell,1}$ . Therefore, we get

$$\begin{aligned} S_\ell &\leq \frac{1}{M^\ell} \int dj_1 A_{1,2} \cdot A_{2,3} \cdots A_{\ell-1,\ell} \cdot B_{\ell,1} \\ &\leq \frac{1}{M^\ell} \int dj_1 \left( c \frac{\log M}{d_{jk}} \right)^{\ell-1} \frac{\sqrt{M}}{d_{jk}^2} = \frac{1}{M^{\ell-1} d_{jk}^2} \left( c \frac{\log M}{d_{jk}} \right)^{\ell-1}. \end{aligned} \quad (25)$$

□

As one can notice from (25), the upper bound on  $S_\ell$  is independent of the choice of the node  $j_1$ . As such, the same upper bound applies on the expectation conditioned on any value of  $j_1$  as follows

$$\tilde{S}_\ell \stackrel{\text{def}}{=} |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* \cdots f_{j_\ell k_\ell} f_{j_1 k_\ell}^* | j_1)| \leq \frac{1}{M^{\ell-1} d_{jk}^2} \left( c \frac{\log M}{d_{jk}} \right)^{\ell-1}. \quad (26)$$

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 proved it for  $\ell = 1, 2$  and will further prove it for  $\ell = 3$  and  $\ell = 4$ . After that, the mechanics of the computation is the same for all  $\ell > 4$ .

In what follows, when we say that there exists an “equality”, we mean an index is equal to another index. For example, when we say there are 2 “equality”s, we mean that there are either two disjoint pairs of equal indices (for example,  $j_1 = j_2$  and  $k_2 = k_4$ ) or 3 equal indices (for example,  $j_1 = j_2 = j_3$ ). Let us start with  $\ell = 3$ ,

$$\begin{aligned} \mathbb{E}(\text{Tr}((FF^\dagger)^3)) &= \sum_{\substack{j_1, j_2, j_3=1, \\ k_1, k_2, k_3=1}}^{\sqrt{M}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) \\ &= \sum_{\substack{j_1 \neq j_2 \neq j_3, \\ k_1 \neq k_2 \neq k_3}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) \\ &\quad + \sum_{\substack{j_1, j_2, j_3, \\ k_1, k_2, k_3: \\ \exists \text{ exactly 1 "equality"}}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) \\ &\quad + \sum_{\substack{j_1, j_2, j_3, \\ k_1, k_2, k_3: \\ \exists \geq 2 \text{ "equality"s}}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*). \end{aligned}$$

Note that if there exists exactly 1 “equality”, then we have 6 different cases; namely,  $j_1 = j_2$  or  $j_2 = j_3$  or  $j_1 = j_3$  or  $k_1 = k_2$  or  $k_2 = k_3$  or  $k_1 = k_3$ . Since in all the 6 cases we end up with adjacent indices<sup>6</sup> being equal, then these 6 cases are identical

<sup>6</sup>Two indices are said to be adjacent if they result in a term and its conjugate when they are equal. In other words, if two  $j$  indices are connected to the same  $k$  index but in opposite direction, then they are said to be adjacent indices.

and it makes no difference which case we consider (we consider below the case where  $j_1 = j_2$ ). Therefore, we have

$$\begin{aligned}
\sum_{\exists \text{ exactly 1 "equality"}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) &= 6 \sum_{\substack{(j_1=j_2) \neq j_3, \\ k_1 \neq k_2 \neq k_3}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) \\
&\leq 6 \sum_{j_1} \mathbb{E} \left( \sum_{k_1} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* | j_1) \sum_{\substack{j_3 \neq j_1, \\ k_2 \neq k_3}} \mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^* | j_1) \right) \\
&\leq 6 \sum_{j_1} \mathbb{E} \left( \sum_{k_1} |\mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* | j_1)| \sum_{\substack{j_3 \neq j_1, \\ k_2 \neq k_3}} |\mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^* | j_1)| \right) \\
&= 6(\sqrt{M})^5 \tilde{S}_1 \tilde{S}_2 = 6M^{5/2} \tilde{S}_1 \tilde{S}_2.
\end{aligned}$$

Moreover, if there exist 2 or more “equality”s, then we will be left with 4 indices at most. Therefore, we have

$$\begin{aligned}
\sum_{\exists \geq 2 \text{ "equality"s}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) &\leq \sum_{\exists \geq 2 \text{ "equality"s}} \mathbb{E}(|f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*|) \\
&\leq \frac{(\sqrt{M})^4}{d_{jk}^6} = \frac{M^2}{d_{jk}^6}.
\end{aligned}$$

As such, we get

$$\begin{aligned}
\mathbb{E}(\text{Tr}((FF^\dagger)^3)) &\leq \sum_{\substack{j_1 \neq j_2 \neq j_3, \\ k_1 \neq k_2 \neq k_3}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_1 k_3}^*) + 6M^{5/2} \tilde{S}_1 \tilde{S}_2 + \frac{M^2}{d_{jk}^6} \\
&\leq M^3 S_3 + 6M^{5/2} \tilde{S}_1 \tilde{S}_2 + \frac{M^2}{d_{jk}^6} \\
&\stackrel{(a)}{\leq} \frac{M(c \log M)^2}{d_{jk}^4} + \frac{M\sqrt{M}(c \log M)}{d_{jk}^5} + \frac{M^2}{d_{jk}^6} \stackrel{(b)}{=} O\left(\frac{M(c \log M)^2}{d_{jk}^4}\right),
\end{aligned}$$

where (a) follows from (25) and Lemma A.2, and (b) follows from the fact that  $d_{jk} \geq \sqrt{M}$ .

We use the same approach for  $\ell = 4$ . We have

$$\begin{aligned}
\mathbb{E}(\text{Tr}((FF^\dagger)^4)) &= \sum_{\substack{j_1, j_2, j_3, j_4=1, \\ k_1, k_2, k_3, k_4=1}}^{\sqrt{M}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&= \sum_{\substack{j_1 \neq j_2 \neq j_3 \neq j_4, \\ k_1 \neq k_2 \neq k_3 \neq k_4}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&+ \sum_{\substack{j_1, j_2, j_3, j_4, \\ k_1, k_2, k_3, k_4: \\ \exists \text{ exactly 1 "equality"}}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&+ \sum_{\substack{j_1, j_2, j_3, j_4, \\ k_1, k_2, k_3, k_4: \\ \exists \text{ exactly 2 "equality"s}}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&+ \sum_{\substack{j_1, j_2, j_3, j_4, \\ k_1, k_2, k_3, k_4: \\ \exists \geq 3 \text{ "equality"s}}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*).
\end{aligned}$$

Note that if there exists exactly 1 “equality”, then we have two different scenarios (See Fig. 5) : 1) two adjacent indices are equal; namely,  $j_1 = j_2$  or  $j_2 = j_3$  or  $j_3 = j_4$  or  $j_1 = j_4$  or  $k_1 = k_2$  or  $k_2 = k_3$  or  $k_3 = k_4$  or  $k_1 = k_4$ . 2) two non-adjacent indices are equal; namely,  $j_1 = j_3$  or  $j_2 = j_4$  or  $k_1 = k_3$  or  $k_2 = k_4$ . Since all the cases under the same scenario are identical, we consider one case for each scenario (the case  $j_1 = j_2$  for the first scenario and the case  $j_1 = j_3$  for the second scenario).

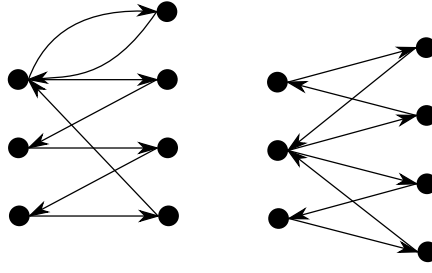


Fig. 5. There exists exactly 1 “equality”: illustrating scenario 1) to the left and scenario 2) to the right.

Therefore, we have

$$\begin{aligned}
& \sum_{\exists \text{ exactly 1 "equality"}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&= 8 \sum_{\substack{(j_1=j_2) \neq j_3 \neq j_4, \\ k_1 \neq k_2 \neq k_3 \neq k_4}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&+ 4 \sum_{\substack{j_2 \neq (j_1=j_3) \neq j_4, \\ k_1 \neq k_2 \neq k_3 \neq k_4}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^* f_{j_1 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&= 8 \sum_{j_1} \mathbb{E} \left( \sum_{k_1} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* | j_1) \sum_{\substack{j_3 \neq j_4 \neq j_1, \\ k_2 \neq k_3 \neq k_4}} \mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1) \right) \\
&+ 4 \sum_{j_1} \mathbb{E} \left( \sum_{\substack{j_2 \neq j_1, \\ k_1 \neq k_2}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^* | j_1) \sum_{\substack{j_4 \neq j_1, \\ k_3 \neq k_4}} \mathbb{E}(f_{j_1 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1) \right)
\end{aligned}$$

which can be upper bounded as such

$$\begin{aligned}
& \sum_{\exists \text{ exactly 1 "equality"}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&\leq 8 \sum_{j_1} \mathbb{E} \left( \sum_{k_1} |\mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* | j_1)| \sum_{\substack{j_3 \neq j_4 \neq j_1, \\ k_2 \neq k_3 \neq k_4}} |\mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1)| \right) \\
&+ 4 \sum_{j_1} \mathbb{E} \left( \sum_{\substack{j_2 \neq j_1, \\ k_1 \neq k_2}} |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^* | j_1)| \sum_{\substack{j_4 \neq j_1, \\ k_3 \neq k_4}} |\mathbb{E}(f_{j_1 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1)| \right) \\
&= 8(\sqrt{M})^7 \tilde{S}_1 \tilde{S}_3 + 4(\sqrt{M})^7 \tilde{S}_2 \tilde{S}_2 = 8M^{7/2} \tilde{S}_1 \tilde{S}_3 + 4M^{7/2} \tilde{S}_2 \tilde{S}_2. \tag{27}
\end{aligned}$$

If there exist exactly 2 “equality”s, we analyze the different possible cases in the following way. The first “equality” can be through either scenario 1) or scenario 2) illustrated in Fig. 5.

**1) First “equality” through the first scenario:** Without loss of generality, consider the case  $j_1 = j_2$ . We want to have a second “equality”. Note that we have two disjoint cycles that meet at the node  $j_1$ . The second “equality” can be between two indices belonging to the same cycle or between two indices belonging to different cycles. There are 6 cases where the “equality” is between two indices belonging to the same cycle; namely,  $j_1 = j_3$  or  $j_3 = j_4$  or  $j_1 = j_4$  or  $k_2 = k_3$  or  $k_3 = k_4$  or  $k_2 = k_4$ . Without loss of generality, consider the case  $k_2 = k_3$ :

$$\begin{aligned}
& \sum_{\substack{(j_1=j_2) \neq j_3 \neq j_4, \\ k_1 \neq (k_2=k_3) \neq k_4}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_2} f_{j_4 k_2}^* f_{j_4 k_2} f_{j_4 k_4} f_{j_1 k_4}^*) \\
& \sum_{\substack{(j_1=j_2) \neq j_3 \neq j_4, \\ k_1 \neq (k_2=k_3) \neq k_4}} \mathbb{E} \left( \frac{1}{r_{j_1 k_1}^2} \frac{1}{r_{j_3 k_2}^2} f_{j_1 k_2} f_{j_4 k_2}^* f_{j_4 k_4} f_{j_1 k_4}^* \right) \\
& = \sum_{j_1, k_2} \mathbb{E} \left( \sum_{k_1} \frac{1}{r_{j_1 k_1}^2} \sum_{j_3} \frac{1}{r_{j_3 k_2}^2} \sum_{j_4 \neq j_1, k_4 \neq k_2} \mathbb{E}(f_{j_1 k_2} f_{j_4 k_2}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1, k_2) \right) \\
& \leq \frac{M}{d_{jk}^4} \sum_{\substack{j_1 \neq j_4, \\ k_2 \neq k_4}} |\mathbb{E}(f_{j_1 k_2} f_{j_4 k_2}^* f_{j_4 k_4} f_{j_1 k_4}^*)| = \frac{M^3}{d_{jk}^4} S_2.
\end{aligned}$$

On the other hand, there are 3 cases where the second “equality” is between two indices belonging to different cycles; namely,  $k_1 = k_2$  or  $k_1 = k_3$  or  $k_1 = k_4$ . Without loss of generality, consider the case  $k_1 = k_3$ :

$$\begin{aligned}
& \sum_{\substack{(j_1=j_2) \neq j_3 \neq j_4, \\ k_2 \neq (k_1=k_3) \neq k_4}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_2} f_{j_3 k_1} f_{j_4 k_1}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
& = \sum_{j_1, k_1} \mathbb{E} \left( f_{j_1 k_1} f_{j_1 k_1}^* \sum_{\substack{j_3 \neq j_4 \neq j_1, \\ k_2 \neq k_4 \neq k_1}} \mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_1} f_{j_4 k_1}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1, k_1) \right) \\
& = \sum_{j_1, k_1} \mathbb{E} \left( \frac{1}{r_{j_1 k_1}^2} \sum_{\substack{j_3 \neq j_4 \neq j_1, \\ k_2 \neq k_4 \neq k_1}} \mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_1} f_{j_4 k_1}^* f_{j_4 k_4} f_{j_1 k_4}^* | j_1, k_1) \right) \\
& \leq \frac{1}{d_{jk}^2} \sum_{\substack{j_1 \neq j_3 \neq j_4, \\ k_1 \neq k_2 \neq k_4}} |\mathbb{E}(f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_1} f_{j_4 k_1}^* f_{j_4 k_4} f_{j_1 k_4}^*)| = \frac{M^3}{d_{jk}^2} S_3.
\end{aligned}$$

Therefore, when the first “equality” happens through the first scenario, we have the following upper bound

$$\begin{aligned}
& \sum_{\substack{(j_1=j_2), j_3, j_4, \\ k_1, k_2, k_3, k_4: \\ \exists \text{ exactly 1 more "equality" }}} \mathbb{E}(f_{j_1 k_1} f_{j_1 k_1}^* f_{j_1 k_2} f_{j_3 k_2}^* f_{j_3 k_2} f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \leq 6 \frac{M^3}{d_{jk}^4} S_2 + 3 \frac{M^3}{d_{jk}^2} S_3 \stackrel{(a)}{=} O \left( \frac{M(c \log M)^4}{d_{jk}^5} \right),
\end{aligned}$$

where (a) follows from Lemma A.2 and from the fact that  $d_{jk} \geq \sqrt{M}$ .

**2) First “equality” through the second scenario:** Without loss of generality, consider the case  $j_1 = j_3$ . We want to pick up a second “equality”. Using similar techniques as used in the previous case, one can show that

$$\begin{aligned}
& \sum_{\substack{(j_1=j_3), j_2, j_4, \\ k_1, k_2, k_3, k_4: \\ \exists \text{ exactly 1 more "equality" }}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_1 k_2}^* f_{j_1 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) = O \left( \frac{M(c \log M)^4}{d_{jk}^5} \right).
\end{aligned}$$

Therefore, combining the upper bounds obtained for the two scenarios, we get

$$\begin{aligned}
& \sum_{\substack{j_1, j_2, j_3, j_4, \\ k_1, k_2, k_3, k_4: \\ \exists \text{ exactly 2 "equality"s }}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) = O \left( \frac{M(c \log M)^4}{d_{jk}^5} \right). \tag{28}
\end{aligned}$$

Finally, if there exist 3 or more “equality”s, then we will be left with at most 5 indices. Therefore, we have

$$\begin{aligned}
& \sum_{\exists \geq 3 \text{ "equality"s}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_2 k_2}^* f_{j_3 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_3 k_3}^* f_{j_4 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
& \leq \sum_{\exists \geq 3 \text{ "equality"s}} \mathbb{E}(|f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_2 k_2}^* f_{j_3 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_3 k_3}^* f_{j_4 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*|) \\
& \leq \frac{(\sqrt{M})^5}{d_{jk}^8} = \frac{M^2 \sqrt{M}}{d_{jk}^8} \stackrel{(a)}{=} O\left(\frac{M}{d_{jk}^5}\right),
\end{aligned} \tag{29}$$

where (a) follows from the fact that  $d_{jk} \geq \sqrt{M}$ .

Combining (27), (28), and (29), we get

$$\begin{aligned}
\mathbb{E}(\text{Tr}((FF^\dagger)^4)) &= \sum_{\substack{j_1, j_2, j_3, j_4=1, \\ k_1, k_2, k_3, k_4=1}}^{\sqrt{M}} \mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_2 k_2}^* f_{j_3 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_3 k_3}^* f_{j_4 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*) \\
&\leq \sum_{\substack{j_1 \neq j_2 \neq j_3 \neq j_4, \\ k_1 \neq k_2 \neq k_3 \neq k_4}} |\mathbb{E}(f_{j_1 k_1} f_{j_2 k_1}^* f_{j_2 k_2} f_{j_2 k_2}^* f_{j_3 k_2} f_{j_3 k_2}^* f_{j_3 k_3} f_{j_3 k_3}^* f_{j_4 k_3} f_{j_4 k_3}^* f_{j_4 k_4} f_{j_1 k_4}^*)| \\
&\quad + 8M^{7/2} \tilde{S}_1 \tilde{S}_3 + 4M^{7/2} \tilde{S}_2 \tilde{S}_2 + O\left(\frac{M(c \log M)^4}{d_{jk}^5}\right) \\
&= M^4 S_4 + 8M^{7/2} \tilde{S}_1 \tilde{S}_3 + 4M^{7/2} \tilde{S}_2 \tilde{S}_2 + O\left(\frac{M(c \log M)^4}{d_{jk}^5}\right) \stackrel{(a)}{=} O\left(\frac{M(c \log M)^4}{d_{jk}^5}\right),
\end{aligned}$$

where (a) follows from (26), Lemma A.2, and from the fact that  $d_{jk} \geq \sqrt{M}$ .

For  $\ell \geq 5$ , using techniques similar to the ones used for  $\ell = 3, 4$ , one can show that

$$\mathbb{E}(\text{Tr}((FF^\dagger)^\ell)) = O\left(M^{\ell-1/2} \sum_{i=1}^{\lfloor \ell/2 \rfloor} \tilde{S}_i \tilde{S}_{\ell-i}\right) + M^\ell S_\ell + O\left(\frac{M(c \log M)^{\ell-1}}{d_{jk}^{\ell+1}}\right) \stackrel{(a)}{=} O\left(\frac{M(c \log M)^{\ell-1}}{d_{jk}^{\ell+1}}\right),$$

where (a) follows from Lemma A.2 and 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.

A last remark is that we proved lemma IV.4 for aligned clusters. However, the proof can be easily generalized to tilted clusters, as shown in Fig. 6. We can always draw a larger cluster containing the original cluster and having the 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 ( $2\sqrt{M} \leq d \leq M$ ).  $\square$

## REFERENCES

- [1] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, March 2000.
- [2] 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.

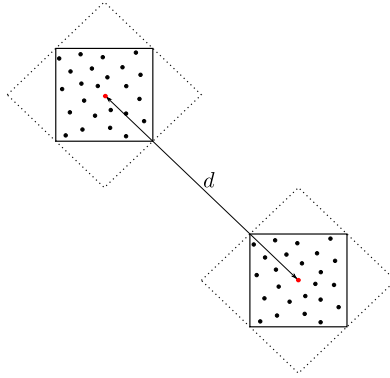


Fig. 6. Two tilted square clusters that have a center-to-center distance  $d$ . We can draw larger squares (drawn in dotted line) containing the original clusters with the same centers that are aligned.

- [3] A. Özgür, O. Lévêque, and E. Preissmann, "Scaling laws for one- and two-dimensional random wireless networks in the low-attenuation regime," *IEEE Trans. on Inform. Theory*, vol. 53, no. 10, pp. 3573–3585, October 2007.
- [4] 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.
- [5] 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.
- [6] B. Tavli, "Broadcast capacity of wireless networks," *IEEE Communications Letters*, vol. 10, no. 2, pp. 68–69, February 2006.
- [7] 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.
- [8] 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.
- [9] 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," vol. 2B, July 2005, pp. 707–710.
- [10] 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.
- [11] —, "Telescopic beamforming for large wireless networks," *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2771–2775, July 2013.
- [12] 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.
- [13] A. Özgür, R. Johari, D. Tse, and O. Lévêque, "Information theoretic operating regimes of large wireless networks," *IEEE Trans. on Inform. Theory*, vol. 56, no. 1, pp. 427–437, January 2010.
- [14] 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. on Inform. Theory*, vol. 57, no. 8, pp. 5088–5102, August 2011.
- [15] A. Merzakreeva, "Cooperation in space-limited wireless networks at low snr," Ph.D. dissertation, EPFL, Switzerland, 2014.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd edition*. Wiley, 2006.
- [17] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, p. 1330, 1963.
- [18] A. Özgür, O. Lévêque, and D. N. C. Tse, "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.