The Capacity Region of Large Wireless Networks

Urs Niesen, Piyush Gupta, and Devavrat Shah

Abstract

The capacity region of a wireless network with n nodes is the set of all simultaneously achievable rates between all possible n^2 node pairs. In this paper, we consider the question of determining the *scaling*, with respect to the number of nodes n, of the capacity region when the nodes are placed uniformly at random in a square region of area n and they communicate over Gaussian fading channels. We identify this scaling of the capacity region in terms of $\Theta(n)$, out of 2^n total possible, cuts. Our results are constructive and provide optimal (in the scaling sense) communication schemes. In the case of a restricted class of traffic requirement (*permutation traffic*), we determine the precise scaling in terms of a natural generalization of the transport capacity. We illustrate the strength of these results by computing the capacity scaling in a number of scenarios with non-uniform traffic patterns for which no such results have been available before.

I. Introduction

Characterizing the capacity region of wireless networks is a long standing open problem in information theory. The exact capacity region is, in fact, not known for even simple networks like a three node relay channel or a four node interference channel. In this paper, we consider the question of approximately determining the capacity region of wireless networks by identifying its scaling in terms of the number of nodes in the network.

A. Related Work

In the last decade or so, exciting progress has been made towards approximating the capacity region of wireless networks. We shall briefly recall a small subset of work related to this paper. In [1], Gupta and Kumar proposed a simpler but insightful question. First, instead of asking for the entire n^2-dimensional capacity region of a wireless network with n nodes, attention was restricted to the scenario where each node is source and destination for exactly one communication pair (called *permutation traffic* in the following). All these source-destination pairs want to communicate at the same rate, and the interest is in finding the maximal achievable such rate. Second, instead of insisting on finding this maximal rate exactly, they focused on its asymptotic behavior as the number of nodes n grows to infinity.

This setup has indeed turned out to be more amenable to analysis. In [1], it was shown that under random placement of nodes in a given region and under certain models of communication motivated by current technology (called *combinatorial channel model* in the following), the per-node rate for random permutation traffic can scale at most as $O(n^{-1/2})$ and this can be achieved (within poly-logarithmic factor in n) by a simple scheme based on multi-hop communication. Many works since then have broadened the channel and communication models under which similar results can be proved (for example, see [2]–[12]). In particular, under the *Gaussian fading channel model* with a power-loss of $r^{-\alpha}$ for signals sent over a distance of r, [11], [12] have shown that in *extended wireless networks* (i.e., n nodes are randomly located in a region with area $\Theta(n)$) the largest per-node rate achievable by all source-destination pairs under random permutation traffic scales essentially like $\Theta\big(n^{1-\min\{3,\alpha\}/2}\big)$.

U. Niesen and D. Shah are with the Laboratory for Information and Decision Systems, Department of EECS at the Massachusetts Institute of Technology. Email: {uniesen,devavrat}@mit.edu

P. Gupta is with the Mathematics of Networks and Communications Research Department, Bell Labs, Alcatel-Lucent. Email: pgupta@research.bell-labs.com

The work of U. Niesen and D. Shah was supported by DARPA grant (ITMANET) 18870740-37362-C; the work of P. Gupta was supported in part by NSF Grants CCR-0325673 and CNS-0519535.

It can be shown that determining the scaling of the maximal achievable per-node rate under random permutation traffic as considered above is equivalent to finding the scaling of the maximal achievable per-node rate under *uniform traffic*, in which each node wants to send data at equal rate to all other nodes (see [13]). That is, analyzing random permutation traffic yields a one dimensional projection of the n^2 dimensional capacity region. Hence, the results in [1] and in [11], [12] mentioned above provide a complete characterization of the scaling of this one dimensional projection for the combinatorial and Gaussian fading channel models, respectively. It is therefore natural to ask if the scaling of the entire n^2 dimensional capacity region can be characterized. To this end, we describe two related approaches taken in recent works.

One approach, taken by Madan, Shah, and Lévêque [13], builds upon the celebrated works of Leighton and Rao [14] and Linial, London, and Rabinovich [15] on the approximate characterization of the capacity region of capacitated wireline networks. For such wireline networks, the scaling of the capacity region is determined (within poly-logarithmic factor in n) by the minimum weighted cut of the network graph. As shown in [13], this naturally extends to wireless networks under the combinatorial channel model, providing an approximation of the capacity region in this case.

Another approach, first introduced by Gupta and Kumar [1], utilizes geometric properties of the wireless network. Specifically, the notion of the *transport capacity* of a network, which is the rate-distance product summed over all source-destination pairs, was introduced in [1]. It was shown that in an extended wireless network with n nodes and under the combinatorial channel model, the transport capacity can scale at most as $\Theta(n)$. This bound on the transport capacity provides a hyper-plane which has the capacity region and origin on the same side. Through a repeated application of this transport capacity bound at different scales, together with the traditional cut-set bound, [16], [17] obtained an implicit characterization of the capacity region under the combinatorial channel model.

For the Gaussian fading channel model, asymptotic upper bounds for the transport capacity were obtained in [2], [3], and for arbitrary weighted sum-rates in [18].

B. *Our Contributions*

Despite the long list of results, the question of approximately characterizing the capacity region under Gaussian fading channel model for general power-loss remains far from being resolved. As the main result of this paper, we resolve this question successfully for extended networks under random node placement.

Our approximate characterization of the capacity region is expressed as the minimum over $\Theta(n)$, of all 2^n total possible, cuts. The upper bound (converse) follows through consideration of appropriate cut-set bounds. The lower bound (achievability) is established through a novel scheme that routes data on a virtual tree constructed using either cooperative or multi-hop communication. Information is sent along an edge towards the root of this tree by distributing it over more nodes in the network, and information is sent along an edge towards the leaves of this tree by concentrating it on fewer nodes.

As mentioned above, the approximate characterization of the capacity region is expressed as a minimization problem, and hence does not admit a succinct analytic expression. Such an expression can, however, be found in the case of general (i.e., not necessarily random) permutation traffic. To this end, we identify a generalization of the notion of transport capacity, resulting in a clean analytic formula for the scaling for this kind of traffic.

C. *Organization*

The remainder of this paper is organized as follows. Section II introduces the channel model and notations. Section III presents our main results and illustrates them with a few example scenarios. Section IV describes at a higher level the proposed communication schemes. Sections V-VIII contain proofs, and Section IX contains concluding remarks.

II. Model

Consider $A(n) \triangleq [0, \sqrt{n}]^2$ and let $V(n) \subset A(n)$ be a set of $|V(n)| = n$ nodes on $A(n)$. We use the same channel model as in [11]. Namely, if $\{x_u[t]\}_{u,t}$ are the (sampled) signals sent by the nodes in $V(n)$, then the (sampled) received signal at node v and time t is

$$y_v[t] = \sum_{u \in V(n)\setminus\{v\}} h_{u,v}[t] x_u[t] + z_v[t] \tag{1}$$

for all $v \in V(n), t \in \mathbb{N}$. Here $\{z_v[t]\}_{v,t}$ are i.i.d. circularly symmetric complex Gaussian random variables with mean 0 and variance 1, and

$$h_{u,v}[t] = r_{u,v}^{-\alpha/2} \exp(\sqrt{-1}\theta_{u,v}[t]),$$

for *path-loss exponent*[1] $\alpha > 2$, and where $r_{u,v}$ is the Euclidean distance between u and v. $\{\theta_{u,v}[t]\}_{u,v}$ is assumed to be i.i.d. with uniform distribution on $[0, 2\pi)$. We either assume that $\{\theta_{u,v}[t]\}_t$ is stationary and ergodic as a function of t which is called *fast fading* in the following, or we assume $\{\theta_{u,v}[t]\}_t$ is constant as a function of t, which is called *slow fading* in the following. In either case, we assume full channel state information is available at all nodes, i.e., each node knows all $\{h_{u,v}[t]\}_{u,v}$ at time t. We also impose an average power constraint of P on the signal $\{x_u[t]\}_t$ for every node $u \in V(n)$.

Let $\Lambda(n) \subset \mathbb{R}_+^{n\times n}$ be the capacity region of the wireless network, i.e., $\lambda \in \Lambda(n)$ if and only if every source-destination pair $(u, v) \in V^2(n)$ can reliably communicate independent messages at rate $\lambda_{u,v}$. Partition $A(n)$ into squares $\{A_{\ell,i}(n)\}_{i=1}^{4^\ell}$ of sidelength $2^{-\ell}\sqrt{n}$, and let $V_{\ell,i}(n)$ be the nodes in $A_{\ell,i}(n)$. Define

$$\begin{aligned} V_0^2(n) &\triangleq \big\{(u, v) \in V^2(n) : \sqrt{n} \leq r_{u,v} \leq 2\sqrt{n}\big\}, \\ V_\ell^2(n) &\triangleq \big\{(u, v) \in V^2(n) : 2^{-\ell}\sqrt{n} \leq r_{u,v} < 2^{-\ell+1}\sqrt{n}\big\}, \end{aligned}$$

for $\ell > 0$. Finally, let[2]

$$L(n) \triangleq \frac{1}{2} \log(n)\big(1 - \log^{-1/2}(n)\big),$$

and note that $L(n)$ is chosen such that

$$4^{-L(n)} n = n^{\log^{-1/2}(n)},$$

and hence

$$\lim_{n\to\infty} \mathbb{E}(\big|A_{L(n),i}(n)\big|) = \lim_{n\to\infty} 4^{-L(n)} n = \infty$$

while at the same time

$$\mathbb{E}(\big|A_{L(n),i}(n)\big|) = 4^{-L(n)} n \leq n^{o(1)},$$

as $n \to \infty$.

Throughout, $\{K_i\}_i$, K, $\widetilde{K}$, $\ldots$, indicate strictly positive finite constants independent of n and ℓ. To simplify notation, we assume, when necessary, that fractions are integers and omit $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$ operators. For the same reason, we also suppress dependence on n within proofs whenever this dependence is clear from the context.

[1] It is worth pointing out that recent results [19] seem to suggest that for $\alpha \in (2, 3)$ and very large values of n, the channel model becomes invalid.

[2] All logarithms are with respect to base 2.

III. Main Results

A. Permutation Traffic

Define the *generalized transport capacity* as

$$T_\alpha(n) \triangleq \sup_{\lambda\in\Lambda(n)} \sum_{(u,v)\in V^2(n)} \lambda_{u,v} f_\alpha(r_{u,v}),$$

where

$$f_\alpha(r) \triangleq \begin{cases} r^{\min\{3,\alpha\}-2} & \text{if } r \geq 1, \\ 1 & \text{if } 0 < r < 1, \\ 0 & \text{if } r = 0. \end{cases}$$

Theorem 1. *Under either fast or slow fading, for any $\alpha > 2$, $\varepsilon > 0$,*

$$T_\alpha(n) = O(n^{1+\varepsilon})$$

with probability $1 - o(1)$ as $n \to \infty$.

We say that $\Pi(n) \subset V^2(n)$ is a *permutation traffic* if for every $u \in V(n)$ there is exactly one $v \in V(n) \setminus \{u\}$ such that $(u, v) \in \Pi(n)$ and exactly one $\tilde{v} \in V(n) \setminus \{u\}$ such that $(\tilde{v}, u) \in \Pi(n)$.

Theorem 2. *Under either fast or slow fading, for any $\alpha > 2$, and any sequence of permutation traffics $\{\Pi(n)\}_{n\geq 1}$,*

$$\sup_{\lambda\in\Lambda(n)} \min_{(u,v)\in\Pi(n)} \lambda_{u,v} f_\alpha(r_{u,v}) \geq n^{-o(1)}$$

with probability $1 - o(1)$ as $n \to \infty$.

Corollary 3. *Under either fast or slow fading, for any $\alpha > 2$,*

$$T_\alpha(n) \geq n^{1-o(1)}$$

with probability $1 - o(1)$ as $n \to \infty$.

Together with Theorem 1, Theorem 2 and its corollary show that

$$\lim_{n\to\infty} \frac{\log(T_\alpha(n))}{\log(n)} = 1.$$

However, Theorem 2 proves a much stronger result than just that. Indeed, it shows that the scaling of $T_\alpha(n)$ is achievable even if we are restricted to an arbitrary permutation traffic and provides a pointwise lower bound on how the rate achievable depends on the distances between the source-destination pairs. More precisely, Theorem 2 guarantees a rate of

$$n^{-o(1)} f_\alpha(r_{u,v})^{-1}$$

for any source-destination pair (u, v) of the permutation traffic. It is worth pointing out that guarantees of this sort *cannot* be made when considering the standard transport capacity

$$T(n) \triangleq \sup_{\lambda\in\Lambda(n)} \sum_{(u,v)\in V^2(n)} \lambda_{u,v} r_{u,v}.$$

Indeed, the arguments in [11], [12] show that under our channel model, the transport capacity is upper bounded by

$$T(n) = O\big(n^{(5-\min\{3,\alpha\})/2+\varepsilon}\big),$$

and that for *random* permutation traffic a transport rate of at least

$$n^{(5-\min\{3,\alpha\})/2-o(1)}$$

is achievable with probability $1 - o(1)$ as $n \to \infty$. However, if we are restricted to a permutation traffic where all source-destination pairs are at a distance[3] $n^{o(1)}$, then it is easy to show that we can at most achieve a transport rate of

$$n^{1+o(1)} \ll n^{(5-\min\{3,\alpha\})/2-o(1)}$$

for $\alpha \in (2,3)$. In other words, the choice of f_α in the definition of the generalized transport capacity $T_\alpha(n)$ is crucial to obtain a tight characterization for all (as opposed to just random) permutation traffics. For $\alpha \geq 3$, the generalized transport capacity essentially coincides with the traditional transport capacity, as defined in [1].

B. General Traffic

A *traffic matrix* is an element $\lambda \in \mathbb{R}_+^{n\times n}$. For any traffic matrix λ, let

$$\rho_\lambda^* = \sup\{b \geq 0 : b\lambda \in \Lambda(n)\}.$$

The next theorem shows how ρ_λ^* can be asymptotically computed for any $\lambda \in \mathbb{R}^{n\times n}$. By convexity of $\Lambda(n)$, this yields an asymptotic characterization of the entire capacity region $\Lambda(n)$ of the wireless network.

For traffic matrix λ, define

$$\begin{aligned}
\phi_\lambda^*(n) &\triangleq \min_{\ell\in\{1,\dots,L(n)\}\cup\{\log(n)\}} g_\alpha(4^{-\ell}n) \min_{i\in\{1,\dots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i}(n))} \\
&= \min\Bigg\{ \min_{\ell\in\{1,\dots,L(n)\}} g_\alpha(4^{-\ell}n) \min_{i\in\{1,\dots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i}(n))}, \\
&\qquad g_\alpha(n^{-1}) \min_{\substack{i\in\{1,\dots,n^2\}: \\ |V_{\log(n),i}(n)|>0}} \frac{1}{D_\lambda(V_{\log(n),i}(n))} \Bigg\},
\end{aligned} \tag{2}$$

where

$$g_\alpha(r) \triangleq \begin{cases} r^{2-\min\{3,\alpha\}/2} & \text{if } r \geq 1, \\ 1 & \text{else,} \end{cases}$$

and, for any $U \subset V(n)$,

$$D_\lambda(U) \triangleq \sum_{u\in U, v\in U^c} (\lambda_{u,v} + \lambda_{v,u}).$$

Note that the second minimization in (2) is over at most n terms, since there are at most n values of i such that $|V_{\log(n),i}(n)| > 0$. Hence $\phi_\lambda^*(n)$ can be computed as a minimum over $\Theta(n)$ terms.

Theorem 4. *Under either fast or slow fading, for any $\alpha > 2$, $\varepsilon > 0$, and any sequence $\{\lambda(n)\}_{n\geq 1}$ of traffic matrices,*

$$n^{-o(1)}\phi_{\lambda(n)}^*(n) \leq \rho_{\lambda(n)}^*(n) = O\big(n^\varepsilon \phi_{\lambda(n)}^*(n)\big),$$

with probability $1 - o(1)$ as $n \to \infty$.

Theorem 4 provides a tight scaling characterization of the entire capacity region $\Lambda(n)$ of the wireless network. Note that $\Lambda(n)$ is a n^2 dimensional set. On the other hand, noting that the minimization in the definition of $\phi_\lambda^*(n)$ for $\ell = \log(n)$ can be restricted to at most n non-empty squares $\{V_{\ell,i}\}_i$ (see (2)), the characterization of $\Lambda(n)$ is given in terms of a minimization problem of dimension $\Theta(n)$. In other words, Theorem 4 provides a $\Theta(n)$ parameter description of $\Lambda(n)$.

Since Theorem 4 characterizes the entire capacity region, it certainly subsumes the results in Theorem 1 and Theorem 2. This is, however, at the expense of a more complex description. Indeed let λ be a traffic matrix corresponding to a permutation traffic (i.e., there are only n non-zero entries in λ), and those $\lambda_{u,v}$

[3]Note that a permutation traffic of this form has, indeed, probability $o(1)$ of resulting from choosing a permutation traffic at random.

that are positive take a value that depends only on $r_{u,v}$. As we have argued in the last paragraph, to check if $\lambda \in \Lambda(n)$ (asymptotically) using Theorem 4, we have to check $\Theta(n)$ conditions. To check the same using Theorems 1 and 2, we only need to compute one inner product, i.e., only one condition needs to be checked. Thus Theorems 1 and 2 provide a one parameter description of $\Lambda(n)$ when restricted to permutation traffics of this form.

C. Example Scenarios

We next illustrate the strength of the above results by determining achievable rates of a few specific wireless network scenarios with non-uniform traffic patterns. While most of these scenarios consider permutation traffic, it is easy to show that the same results hold also if the source-destination pairing is chosen at random (possibly with non-uniform distribution). For example, if each source chooses its destination uniformly at random then the resulting pairing can be decomposed into at most $\log^2(n)$ permutation traffics with probability $1-o(1)$ as $n \to \infty$, and time sharing between those $\log^2(n)$ permutation traffics yields only an additional factor $n^{o(1)}$ loss in rate.

Example 1. *Multiple Classes of Source-Destination Pairs*

There are K classes of source-destination pairs, for some fixed K. Each source node in class i generates traffic at the same rate $\lambda_i(n)$ for a destination node that is chosen at distance $\Theta(n^{\beta_i})$, for some fixed $\beta_i \in [0, 0.5]$ and such that the resulting source-destination pairing yields a permutation traffic. Each node picks the class it belongs to in an arbitrary fashion. Then, Theorem 1 provides the following upper bound on the rates obtained by different classes:

$$\lambda_i(n) = O(n^{-\beta_i \bar{\alpha} + \varepsilon}),$$

for any $\varepsilon > 0$, and where

$$\bar{\alpha} \triangleq \min\{3, \alpha\} - 2.$$

The achievability of essentially the same order follows from Theorem 2, i.e.,

$$\lambda_i(n) \geq n^{-\beta_i \bar{\alpha} - o(1)}.$$

Hence, for a fixed number of classes K, the K dimensional projection of the capacity region considered here is rectangular (in the scaling sense), with source nodes in each class obtaining rates as a function of only the source-destination separation in that class. $\Diamond$

Example 2. *Traffic Variation with Source-Destination Separation*

Pick a permutation traffic at random, as in the traditional setting. However, instead of all sources generating traffic at the same rate, source node u generates traffic at rate that is a function of its separation from destination v, i.e., the traffic matrix is given by $\lambda_{u,v} = \psi(r_{u,v})$ for some function ψ. In particular, let us consider

$$\psi(r) \triangleq \begin{cases} r^\beta & \text{if } r \geq 1, \\ 1 & \text{else} \end{cases}$$

for some fixed $\beta \in \mathbb{R}$. The traditional setting corresponds to $\beta = 0$. Then, Theorem 1 gives the following upper bound on the capacity scaling for this traffic matrix

$$\rho^*_\lambda(n) = \begin{cases} O(n^{-(\bar{\alpha}+\beta)/2+\varepsilon}) & \text{if } \beta \geq -\bar{\alpha} - 2, \\ O(n^{1+\varepsilon}) & \text{else.} \end{cases}$$

Applying Theorem 2 shows that

$$\rho^*_\lambda(n) \geq \begin{cases} n^{-(\bar{\alpha}+\beta)/2-o(1)} & \text{if } \beta \geq -\bar{\alpha}, \\ n^{-o(1)} & \text{else.} \end{cases}$$

The two bounds coincide order wise for $\beta \geq -\bar{\alpha}$, yielding the capacity scaling in this case. For $\beta < -\bar{\alpha}$, the above upper bound is loose and we need to utilize Theorem 4 to establish the capacity scaling as

$$\rho_\lambda^*(n) = \begin{cases} \Theta(n^{-(\bar{\alpha}+\beta)/2\pm\varepsilon}) & \text{if } \beta \geq -\bar{\alpha}, \\ \Theta(n^{\pm\varepsilon}) & \text{else.} \end{cases}$$

For $\beta = 0$, and noting that $0 \geq -\bar{\alpha}$, this recovers the results from [11], [12] for random permutation traffic with uniform rate. $\Diamond$

Example 3. *Source-Destination Separation Variation*

Each source generates traffic at the same rate ρ. We consider a sequence of permutation traffics $\{\Pi(n)\}_{n\geq 1}$ such that for any $\delta > 0$ and $0 < r \leq 1 - \delta$,

$$\lim_{n\to\infty} \frac{1}{n} \left|(u,v) \in \Pi(n) : r_{u,v}/\sqrt{n} \in [r, r+\delta)\right| = \int_{x=r}^{r+\delta} \psi(x)dx,$$

for some function ψ. In particular, let $\psi(r) \propto r^\beta$ for some fixed $\beta \in \mathbb{R}$. Note that the traditional setup of choosing a permutation traffic at random corresponds essentially to $\beta = 1$. Then, an upper bound on ρ is given by Theorem 1 as

$$\rho_\lambda^*(n) = \begin{cases} O(n^{-\bar{\alpha}/2+\varepsilon}) & \text{if } \beta \geq -1, \\ O(n^{-(\bar{\alpha}+\beta+1)/2+\varepsilon}) & \text{if } -1-\bar{\alpha} \leq \beta < -1, \\ O(n^{\varepsilon}) & \text{else.} \end{cases}$$

The achievability of essentially the same order follows from Theorem 4. For $\beta = 1$ this coincides again with the results from [11], [12] for random permutation traffic with uniform rate. $\Diamond$

Example 4. *Sources with Multiple Destinations*

All the example scenarios so far are concerned with permutation traffic. Here we consider more general traffic patterns. There are K classes of source nodes, for some fixed K. Each source node in class i has $\Theta(n^{\beta_i})$ destination nodes for some fixed $\beta_i \in [0, 1]$ and generates independent traffic at the same rate $\lambda_i(n)$ for each of them. Each of these destination nodes is chosen uniformly at random among the n nodes. Every node picks the class it belongs to independently and uniformly at random. Then, Theorem 4 provides the following bounds on the rates obtained by different classes:

$$n^{-\beta_i-\bar{\alpha}/2-o(1)} \leq \lambda_i(n) = O(n^{-\beta_i-\bar{\alpha}/2+\varepsilon}),$$

for any $\varepsilon > 0$ as $n \to \infty$. In other words, time sharing between all K classes and then (within each class) between all $\Theta(n^{\beta_i})$ destination nodes is order optimal in this scenario. $\Diamond$

IV. Communication Schemes

In this section, we provide a high-level description of the communication schemes used to prove achievability in Theorem 2 (see Section IV-B below) and in Theorem 4 (see Section IV-C below). We start off in Section IV-A by recalling results from prior work that will be used as building blocks in the following.

A. Hierarchical Relaying and Multi-Hop Schemes

Here we discuss (asymptotically) optimal communication schemes for permutation traffic with uniform rate on $A(n)$ in which most source-destination pairs are at a distance of $\Theta(\sqrt{n})$. Permutation traffics of this sort occur with high probability if they are generated uniformly at random. We shall use these communication schemes as building blocks in the following.

The type of optimal communication scheme depends drastically on the path loss exponent α. For $\alpha \in (2, 3]$, i.e., the path loss exponent is small, cooperative communication on a global scale is necessary to achieve optimal performance. For $\alpha > 3$, i.e., the path loss exponent is large, only local communication between neighboring nodes is necessary, and traffic is routed in a multi-hop fashion from the source to the destination. We will refer to the optimal scheme for $\alpha \in (2, 3]$ as hierarchical relaying scheme, and to the optimal scheme for $\alpha > 3$ as multi-hop scheme.

Given a permutation traffic on $V(n)$. For $\alpha \in (2, 3]$, hierarchical relaying achieves a per-node rate of $n^{1-\alpha/2-o(1)}$. For $\alpha > 3$, multi-hop communication achieves a per-node rate of $n^{-1/2}$. By choosing the appropriate scheme, we can thus achieve a per-node rate of $n^{1-\min\{3,\alpha\}/2}$. We provide a short description of the hierarchical relaying scheme in the following. The details can be found in [12].

Consider n nodes placed independently and uniformly at random on $A(n)$. Divide $A(n)$ into

$$n^{\frac{2}{\alpha}\log^{-1/3}(n)}$$

squarelets of equal size. Call a squarelet *dense*, if it contains a number of nodes proportional to its area. For each source-destination pair, choose such a dense squarelet as a *relay*, over which it will transmit information (see Figure 1).

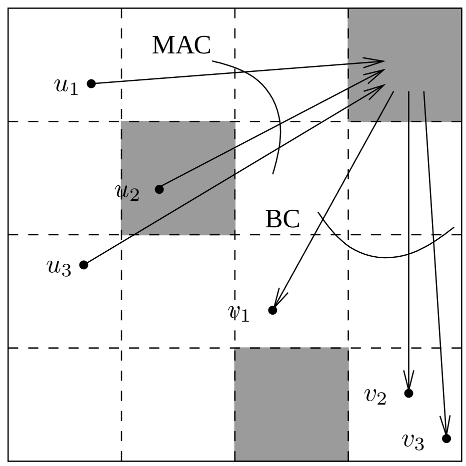

Fig. 1. Sketch of one level of the hierarchical relaying scheme. Here $\{(u_i, v_i)\}_{i=1}^3$ are three source-destination pairs. Groups of source-destination pairs relay their traffic over relay squarelets, which contain a number of nodes proportional to their area (shaded). We time share between the different relay squarelets. Within all relay squarelets the scheme is used recursively to enable joint decoding and encoding at each relay.

Consider now one such relay squarelet and the nodes that are transmitting information over it. If we assume for the moment that the nodes within the relay squarelets could cooperate then between the source nodes and the relay squarelet we would have a multiple access channel (MAC), where each of the source nodes has one transmit antenna, and the relay squarelet (acting as one node) has many receive antennas. Between the relay squarelet and the destination nodes, we would have a broadcast channel (BC), where each destination node has one receive antenna, and the relay squarelet (acting again as one node) has many transmit antennas. The cooperation gain from using this kind of scheme arises from the use of multiple antennas for this MAC and BC.

To actually enable this kind of cooperation at the relay squarelet, local communication within the relay squarelets is necessary. It can be shown that this local communication problem is actually the same as the original problem, but at a smaller scale. Indeed, we are now considering a square of size

$$n^{1-\frac{2}{\alpha}\log^{-1/3}(n)}$$

with equal number of nodes (at least order wise). Hence we can use the same scheme recursively to solve this subproblem. We terminate the recursion after

$$\log^{1/3}(n)$$

iterations, at which point we use simple TDMA to bootstrap the scheme.

Observe that at the final level of the scheme, we have divided $A(n)$ into

$$\left(n^{\frac{2}{\alpha}\log^{-1/3}(n)}\right)^{\log^{1/3}(n)} = n^{2/\alpha}$$

squarelets. A sufficient condition for the scheme to succeed is that all these squarelets are dense (i.e., contain a number of nodes proportional to their area). However much weaker conditions are sufficient as well (see [12]). The per-node rate achievable with this scheme is at least

$$n^{1-\alpha/2-o(1)},$$

and for traffic matrices where a constant fraction of source-destination pairs are at distance $\Theta(\sqrt{n})$ (as is the case with probability $1-o(1)$ as $n \to \infty$ if the source-destination traffic is chosen uniformly at random), this is asymptotically the best uniformly achievable per-node rate.

B. Permutation Traffic

As pointed out in the last section, for permutation traffic hierarchical relaying and multi-hop communication achieve a per-node rate of $n^{1-\min\{3,\alpha\}/2-o(1)}$. This rate is independent of the distance between source-destination pairs. It is shown in [11], [12] that for *random* source-destination pairing (in which most of the source-destination pairs are at a distance of order $\Theta(\sqrt{n})$) no communication scheme can uniformly over all such pairs achieve a per-node rate of more than $O(n^{1-\min\{3,\alpha\}/2+\varepsilon})$ for any $\varepsilon > 0$. In other words, for any communication scheme, there exists at least one source-destination pair whose rate is upper bounded by $O(n^{1-\min\{3,\alpha\}/2+\varepsilon})$. On the other hand, one suspects that certain source-destination pairs should be able to communicate at a rate that is considerably higher than that.

As an example, consider a situation where half of the source-destination pairs are at a distance of order $\Theta(\sqrt{n})$ and the other half are at a distance of $\Theta(1)$. By operating the network in a "long-distance" and a "short-distance" mode, one should be able to achieve the same $n^{1-\min\{3,\alpha\}/2-o(1)}$ per-node rate for those source-destination pairs at distance $\Theta(\sqrt{n})$, while being able to communicate at much higher rates between source-destination pairs at distance $\Theta(1)$ (see Figure 2). Theorem 2 shows that this is indeed the case. In fact, it shows that those source-destination pairs at distance $\Theta(1)$ can communicate at a per-node rate of $n^{-o(1)}$. This is within a $n^{o(1)}$ factor of the best scheme possible even without having to support the source-destination pairs at distance $\Theta(\sqrt{n})$.

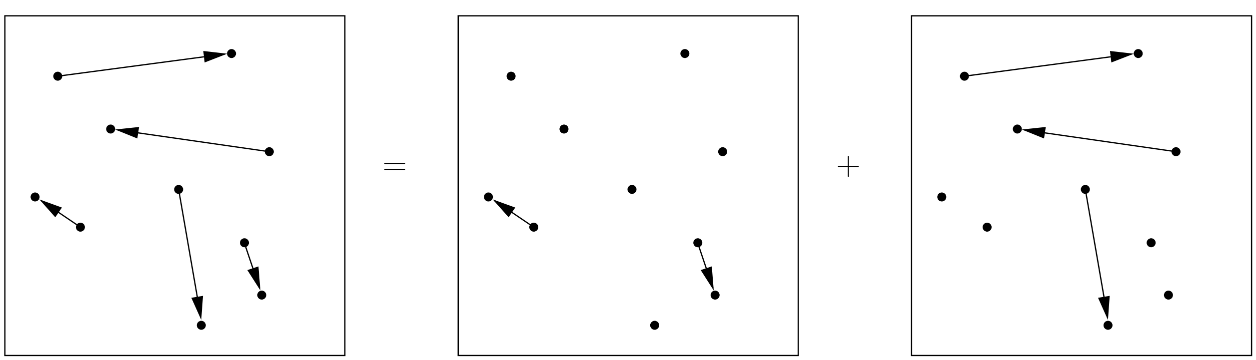

Fig. 2. Sketch of the decomposition of a permutation traffic into sub-traffics with roughly equal source-destination distances.

The proof of Theorem 2 formalizes this idea of decomposing the permutation traffic into source-destination pairs at different scales. More precisely, each permutation traffic is decomposed into subtraffics at $\Theta(\log(n))$ different distance scales, and the scheme operates by time-sharing between those subtraffics.

C. General Traffic

So far, we have only considered permutation traffic. In other words, each node is source and destination exactly once. Moreover, transmission rates were only allowed to depend on the distance between source-destination pairs. While useful, this is still a rather restrictive setup. In the most general form, we would like to answer the following question. Given a traffic matrix $\lambda \in \mathbb{R}^{n\times n}$, is it possible to simultaneously transmit independent messages between each node pair $(u,v) \in V^2(n)$ at rate $\lambda_{u,v}$? Or, in other words, is $\lambda \in \Lambda(n)$?

Theorem 4 provides an asymptotic answer to this question. Its proof relies on the construction of a communication graph G. This graph is a tree, whose leaf nodes represent the nodes $V(n)$ in the wireless network. The intermediate nodes of G represent larger clusters of nodes (i.e., subsets of $V(n)$) in the wireless network (see Figure 3).

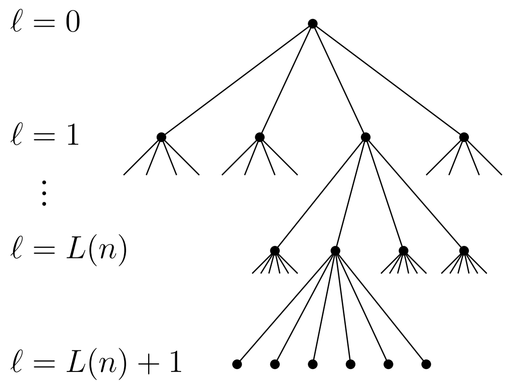

Fig. 3. Communication graph G constructed in the proof of Theorem 4. Nodes on levels $\ell \in \{0,\dots,L(n)-1\}$ have each four children, nodes on level $\ell = L(n)$ have each $\Theta\big(n^{\log^{-1/2}(n)}\big)$ children. The total number of terminal nodes is n, one representing each node in the wireless network $V(n)$. A non-terminal node in G at level $\ell \in \{0,\dots,L(n)\}$ represents the collection of nodes in $V_{\ell,i}(n)$ for some i.

Messages are sent from source to destination by routing it over G. To send information from a child node to its parent in G (i.e., towards the root node of G), the message at the cluster in $V(n)$ represented by the child node is distributed evenly among all nodes in the bigger cluster in $V(n)$ represented by the parent node. To send information from a parent node to a child node in G (i.e., away from the root node of G), the message at the cluster in $V(n)$ represented by the parent node is concentrated on the cluster in $V(n)$ represented by the child node. This distribution and concentration of messages in the wireless network is performed by either using hierarchical relaying (for $\alpha \in (2,3]$) or multi-hop communication (for $\alpha > 3$).

V. Auxiliary Lemmas

In this section, we provide auxiliary results, which will be used several times in the following. Lemmas 5 and 6 describe regularity properties exhibited with high probability by the random node placement. Lemmas 7 and 8 provide auxiliary upper bounds on the performance of any scheme in terms of cut-set bounds. Finally, Lemma 9 describes auxiliary results on the performance of hierarchical relaying and multi-hop communication as described in Section IV-A.

Lemma 5. *For any $\delta > 0$, let*

$$L_\delta(n) \triangleq \frac{1}{2}\log(n)\big(1 - \delta\log^{-1/2}(n)\big).$$

Then

$$\mathbb{P}\bigg(\bigcap_{\ell=1}^{L_\delta(n)}\bigcap_{i=1}^{4^\ell}\big\{\,|V_{\ell,i}(n)| \in [4^{-\ell-1}n, 4^{-\ell+1}n]\big\}\bigg) = 1 - o(1)$$

as $n \to \infty$*. In particular, this holds for* $L(n) = L_1(n)$*.*

Proof. Consider the j-th node and let B_j be the indicator random variable of the event that this node lies in $A_{\ell,i}$ for fixed ℓ, i. Note that

$$\sum_{j=1}^{n} B_j = |V_{\ell,i}|,$$

and that

$$\mathbb{P}(B_j = 1) = 4^{-\ell}.$$

Hence using the Chernoff bound

$$\mathbb{P}\Big(\sum_{j=1}^{n} B_j \notin [4^{-\ell-1}n, 4^{-\ell+1}n]\Big) \leq \exp(-K4^{-\ell}n),$$

for some constant $K > 0$. From this, we obtain for $\ell = L_\delta(n)$,

$$\begin{aligned}
\mathbb{P}\Big(\bigcap_{i=1}^{4^{L_\delta(n)}} \big\{ \big|V_{L_\delta(n),i}\big| \in [4^{-L_\delta(n)-1}n, 4^{-L_\delta(n)+1}n]\big\}\Big) &\geq 1 - \sum_{i=1}^{4^{L_\delta(n)}} \mathbb{P}\big(\big|V_{L_\delta(n),i}\big| \notin [4^{-L_\delta(n)-1}n, 4^{-L_\delta(n)+1}n]\big) \\
&\geq 1 - 4^{L_\delta(n)} \exp(-K4^{-L_\delta(n)}n) \\
&= 1 - o(1).
\end{aligned} \tag{3}$$

Since the $\{A_{\ell,i}\}_{\ell,i}$ are nested as a function of ℓ, we have

$$\bigcap_{\ell=1}^{L_\delta(n)} \bigcap_{i=1}^{4^\ell} \big\{ |V_{\ell,i}| \in [4^{-\ell-1}n, 4^{-\ell+1}n]\big\} = \bigcap_{i=1}^{4^{L_\delta(n)}} \big\{ \big|V_{L_\delta(n),i}\big| \in [4^{-L_\delta(n)-1}n, 4^{-L_\delta(n)+1}n]\big\},$$

which, combined with (3), proves the lemma. □

Lemma 6. *For any* $\delta > 0$*,*

$$\mathbb{P}\Big(\min_{u\in V(n), v\in V(n)\setminus\{u\}} r_{u,v} \geq n^{-1/2-\delta}\Big) = 1 - o(1),$$

as $n \to \infty$*.*

Proof. For $u, v \in V$ let

$$B_{u,v} \triangleq \{r_{u,v} < r\}.$$

Fix a node $u \in V$, then

$$\mathbb{P}(B_{u,v}|u) \leq \frac{r^2\pi}{n},$$

(the inequality being due to boundary effects). Moreover, the events $\{B_{u,v}\}_{v\in V\setminus\{u\}}$ are independent conditioned on u, and thus

$$\mathbb{P}\Big(\cap_{v\in V\setminus\{u\}} B^c_{u,v} \big| u\Big) = \prod_{v\in V\setminus\{u\}} \mathbb{P}(B^c_{u,v}|u) \geq \Big(1 - \frac{r^2\pi}{n}\Big)^n.$$

From this,

$$
\begin{aligned}
\mathbb{P}\Big(\min_{u\in V, v\in V\setminus\{u\}} r_{u,v} < r\Big) &= \mathbb{P}\Big(\cup_{u\in V, v\in V\setminus\{u\}} B_{u,v}\Big) \\
&\le \sum_{u\in V} \mathbb{P}\Big(\cup_{v\in V\setminus\{u\}} B_{u,v}\Big) \\
&= \sum_{u\in V} \Big(1-\mathbb{P}\Big(\cap_{v\in V\setminus\{u\}} B^c_{u,v}\Big)\Big) \\
&= \sum_{u\in V} \Big(1-\mathbb{E}\Big(\mathbb{P}\Big(\cap_{v\in V\setminus\{u\}} B^c_{u,v}\Big| u\Big)\Big)\Big) \\
&\le \sum_{u\in V} \Big(1-\Big(1-\frac{r^2\pi}{n}\Big)^n\Big) \\
&= n\Big(1-\Big(1-\frac{r^2\pi}{n}\Big)^n\Big),
\end{aligned}
$$

which converges to zero for $r = n^{-1/2-\delta}$. ☐

Lemma 7. *Under either fast or slow fading, for any $\alpha > 2$ and $\varepsilon > 0$, there exists $K_1 > 0$ such that for all $\lambda \in \Lambda(n)$,*

$$
\mathbb{P}\Big(\bigcap_{\ell\in\{1,\dots,L(n)\}} \bigcap_{i\in\{1,\dots,4^\ell\}} B_\ell(V_{\ell,i}(n))\Big) \ge 1-o(1), \tag{4}
$$

$$
\mathbb{P}\Big(\bigcap_{\ell\in\{1,\dots,L(n)\}} \bigcap_{i\in\{1,\dots,4^\ell\}} B_\ell(V_{\ell,i}(n)^c)\Big) \ge 1-o(1), \tag{5}
$$

as $n \to \infty$, and where for any $U \subset V(n)$

$$
B_\ell(U) \triangleq \Big\{\sum_{u\in U, v\in U^c} \lambda_{u,v} \le K_1 n^{\varepsilon}(4^{-\ell}n)^{2-\min\{3,\alpha\}/2}\Big\}.
$$

Proof. For any $U_1, U_2 \subset V$, denote by $C(U_1, U_2)$ the MIMO capacity between nodes in U_1 and nodes in U_2. The arguments of [11, Theorem 5.2] show that in the fast fading case for every $\varepsilon > 0$ there exists $K, \widetilde{K} > 0$ and a collection of node placements $\mathcal{V}$ (each of cardinality n) such that for any $V \in \mathcal{V}$, $\ell \in \{0,\dots,L(n)\}$,

$$
C(V_{\ell,i}, V^c_{\ell,i}) \le K n^{\varepsilon/2} \sum_{u\in V_{\ell,i}, v\in V^c_{\ell,i}} r^{-\alpha}_{u,v}, \tag{6}
$$

$$
C(V^c_{\ell,i}, V_{\ell,i}) \le K n^{\varepsilon/2} \sum_{u\in V^c_{\ell,i}, v\in V_{\ell,i}} r^{-\alpha}_{u,v}, \tag{7}
$$

and for adjacent squares $A_{\ell,i}, A_{\ell,j}$,

$$
\sum_{u\in V_{\ell,i}, v\in V_{\ell,j}} r^{-\alpha}_{u,v} \le \widetilde{K} n^{\varepsilon/2}(4^{-\ell}n)^{2-\min\{3,\alpha\}/2}. \tag{8}
$$

For the slow fading case, the two statements hold with probability $1-o(n^{-1})$ as $n\to\infty$. Moreover, in both cases

$$
\mathbb{P}(V\in\mathcal{V}) \ge 1-o(1) \tag{9}
$$

as $n \to \infty$. Consider now two diagonal squares $A_{\ell,i}, A_{\ell,j}$, and choose $\tilde{i}, \tilde{j}$ such that $A_{\ell,i}\cup A_{\ell,\tilde{i}}$ and $A_{\ell,j}\cup A_{\ell,\tilde{j}}$ are adjacent rectangles. Using the same arguments to these rectangles and suitably redefining $\widetilde{K}$ and $\mathcal{V}$ shows that (8) and (9) hold for diagonal squares as well. Moreover, by Lemma 5 we can assume without loss of generality that for every $V \in \mathcal{V}$ all $V_{\ell,i}$ have cardinality at most $4^{-\ell+1}n$ for each $\ell \in \{1, \dots, L(n)\}, i \in \{1, \dots, 4^{\ell}\}$.

Using this, we now compute the summation in (6). Consider "rings" of squares around $A_{\ell,i}$. The first such "ring" contains the (at most) 8 squares neighboring $A_{\ell,i}$. The next "ring" contains at most 16 squares. In general, "ring" k contains at most $8k$ squares. Let

$$\{A_{\ell,i}\}_{i \in I_k}$$

be the squares in "ring" k. Then

$$\sum_{u \in V_{\ell,i}, v \in V_{\ell,i}^c} r_{u,v}^{-\alpha} = \sum_{k \geq 1} \sum_{j \in I_k} \sum_{u \in V_{\ell,i}, v \in V_{\ell,j}} r_{u,v}^{-\alpha}. \tag{10}$$

By (8) and the arguments in the last paragraph,

$$\sum_{j \in I_1} \sum_{u \in V_{\ell,i}, v \in V_{\ell,j}} r_{u,v}^{-\alpha} \leq 8\widetilde{K} n^{\varepsilon/2} (4^{-\ell} n)^{2-\min\{3,\alpha\}/2}. \tag{11}$$

Now note that for $k > 1$ and $j \in I_k$, nodes in $V_{\ell,i}$ and $V_{\ell,j}$ are at least at distance $(k-1)(2^{-\ell}\sqrt{n})$. Moreover, since $V \in \mathcal{V}$, each $\{V_{\ell,j}\}_{\ell,j}$ has cardinality at most $4^{-\ell+1}n$. Thus

$$\begin{aligned}
\sum_{k > 1} \sum_{j \in I_k} \sum_{u \in V_{\ell,i}, v \in V_{\ell,j}} r_{u,v}^{-\alpha} &\leq \sum_{k > 1} 8k \big(4^{-\ell+1} n\big)^2 \big((k-1)(2^{-\ell}\sqrt{n})\big)^{-\alpha} \\
&= 128 \big(4^{-\ell} n\big)^{2-\alpha/2} \sum_{k > 1} k(k-1)^{-\alpha} \\
&= K' \big(4^{-\ell} n\big)^{2-\alpha/2},
\end{aligned} \tag{12}$$

for some $K' > 0$, and where we have used that $\alpha > 2$. Substituting (11) and (12) into (10) yields

$$\sum_{u \in V_{\ell,i}, v \in V_{\ell,i}^c} r_{u,v}^{-\alpha} \leq 8\widetilde{K} n^{\varepsilon/2} (4^{-\ell} n)^{2-\min\{3,\alpha\}/2} + K' \big(4^{-\ell} n\big)^{2-\alpha/2}, \tag{13}$$

for $V \in \mathcal{V}$.

Combining (13) with (6) and using the cut-set bound shows that

$$\mathbb{P}\big(B_\ell(V_{\ell,i}(n)) \big| V \in \mathcal{V}\big) \geq 1 - o(n^{-1}),$$

for every $\ell \in \{1, \dots, L(n)\}, i \in \{1, \dots, 4^{\ell}\}$, and under either fast or slow fading (the probability is, in fact, equal to 1 in the fast fading case). Hence, using $L(n)4^{L(n)} \leq n$ for n large enough,

$$\begin{aligned}
&\mathbb{P}\Big(\bigcap_{\ell \in \{1,\dots,L(n)\}} \bigcap_{i \in \{1,\dots,4^\ell\}} B_\ell(V_{\ell,i}(n)) \Big) \\
&\qquad \geq \mathbb{P}\Big(\bigcap_{\ell \in \{1,\dots,L(n)\}} \bigcap_{i \in \{1,\dots,4^\ell\}} B_\ell(V_{\ell,i}(n)) \Big| V \in \mathcal{V} \Big) \mathbb{P}(V \in \mathcal{V}) \\
&\qquad \geq \Big(1 - \sum_{\ell \in \{1,\dots,L(n)\}} \sum_{i \in \{1,\dots,4^\ell\}} \Big(1 - \mathbb{P}\big(B_\ell(V_{\ell,i}(n)) \big| V \in \mathcal{V} \big) \Big) \Big) \mathbb{P}(V \in \mathcal{V}) \\
&\qquad \geq \big(1 - n o(n^{-1})\big)\big(1 - o(1)\big) \\
&\qquad \geq 1 - o(1).
\end{aligned}$$

This shows (4); (5) follows from a similar argument and (7). ☐

Lemma 8. *Under either fast or slow fading, for any $\alpha > 2$, there exists $K_2 > 0$ such that for all $\lambda \in \Lambda(n)$, $u \in V(n)$,*

$$\mathbb{P}\Big(\sum_{v \in V(n)\setminus\{u\}} \lambda_{u,v} \leq K_2 \log(n) \ \forall u \in V(n) \Big) \geq 1 - o(1), \tag{14}$$

$$\mathbb{P}\Big(\sum_{v \in V(n)\setminus\{u\}} \lambda_{v,u} \leq K_2 \log(n) \ \forall u \in V(n) \Big) \geq 1 - o(1), \tag{15}$$

as $n \to \infty$.

Proof. The argument follows the one in [11, Theorem 3.1]. As before, denote by $C(U_1, U_2)$ the MIMO capacity between nodes in U_1 and nodes in U_2, for $U_1, U_2 \subset V$. Consider first (14). By the cut-set bound,

$$\sum_{v \neq u} \lambda_{u,v} \leq C(\{u\}, \{u\}^c).$$

Here $C(\{u\}, \{u\}^c)$ is the SIMO capacity between u and the nodes in $\{u\}^c$, i.e.,

$$\begin{aligned} C(\{u\}, \{u\}^c) &= \log\Big(1 + P \textstyle\sum_{v \neq u} |h_{u,v}|^2\Big) \\ &\leq \log(1 + P(n-1)n^{\alpha}), \end{aligned}$$

where for the first inequality we have used Lemma 6, which asserts that $r_{u,v} \geq n^{-1/2-\delta} \geq n^{-1}$ for all $u, v \in V$ with probability $1 - o(1)$ as $n \to \infty$.

Similarly, for (15),

$$\sum_{v \neq u} \lambda_{v,u} \leq C(\{u\}^c, \{u\}).$$

Here $C(\{u\}, \{u\}^c)$ is the MISO capacity between the nodes in $\{u\}^c$ and u, i.e.,

$$\begin{aligned} C(\{u\}^c, \{u\}) &\leq \log\Big(1 + (n-1) P \textstyle\sum_{v \neq u} |h_{v,u}|^2\Big) \\ &\leq \log(1 + P(n-1)^2 n^{\alpha}). \end{aligned}$$

□

Lemma 9. *For each $n \in \mathbb{N}$, let $\ell(n) \in \{0, \ldots, L(n)\}$ and let $\Pi_i(n)$ be any permutation traffic on $V_{\ell(n),i}(n)$ for every $i \in \{1, \ldots, 4^{\ell(n)}\}$. Then under either fast or slow fading, and for any $\alpha > 2$,*

$$\sup_{\lambda \in \Lambda(n)} \min_{i \in \{1,\ldots,4^{\ell(n)}\}} \min_{(u,v) \in \Pi_i(n)} \lambda_{u,v} \geq n^{-o(1)} (4^{-\ell(n)} n)^{1 - \min\{3,\alpha\}/2},$$

with probability $1 - o(1)$ as $n \to \infty$.

Proof. We shall use either hierarchical relaying (for $\alpha \in (2, 3]$) or multi-hop (for $\alpha > 3$) to communicate within each square $V_{\ell,i}$. See Section IV-A for a description of these communication schemes. We operate every fourth of the $V_{\ell,i}$ simultaneously, and show that the added interference due to this spatial reuse results only in a constant factor loss in rate.

Consider first $\alpha \in (2, 3]$. The squares $A_{\ell,i}$ at level ℓ have an area of $4^{-\ell} n$. In order to be able to use hierarchical relaying within each of the $\{A_{\ell,i}\}_i$, it is sufficient to show that we can partition each $A_{\ell,i}$ into $(4^{-\ell} n)^{2/\alpha}$ squarelets, each of which contains a number of nodes proportional to the area. In other words, we partition A into squarelets of size

$$\begin{aligned} \big(4^{-\ell} n\big)^{1-2/\alpha} &\geq \big(4^{-L(n)} n\big)^{1-2/\alpha} \\ &= 2^{(1-2/\alpha) \log^{1/2}(n)} \\ &= 4^{-L_{\delta}(n)} n, \end{aligned}$$

where

$$\delta \triangleq 1 - 2/\alpha > 0.$$

Hence Lemma 5 shows that with probability $1 - o(1)$ as $n \to \infty$, all

$$\{A_{i,\ell}\}_{\ell \in \{0,\dots,L(n)\}, i \in \{1,\dots,4^\ell\}}$$

are simultaneously regular enough for hierarchical relaying to be successful. This achieves a per-node rate of

$$\lambda_{u,v} \geq n^{-o(1)} (4^{-\ell} n)^{1-\alpha/2} \tag{16}$$

for any $(u,v) \in \Pi_i(n)$

That the additional interference from spatial reuse results only in a constant loss in rate, follows from the same arguments as in the proof of [12, Theorem 1] (with the appropriate modifications for slow fading as described there). Intuitively, this is the case since the interference from a square at distance r is attenuated by a factor $r^{-\alpha}$, which, since $\alpha > 2$, is summable. Hence the combined interference has power on the order of the receiver noise, resulting in only a constant factor loss in rate.

For $\alpha > 3$, the argument is similar — instead of hierarchical relaying we now use multi-hop communication. This achieves a per-node rate of

$$\lambda_{u,v} \geq n^{-o(1)} (4^{-\ell} n)^{-1/2} \tag{17}$$

for any $(u,v) \in \Pi_i(n)$. Combining (16) and (17) yields the desired result. $\square$

VI. Proof of Theorem 1

Note that for any $u, v \in V$ with $u \neq v$,

$$(u,v) \in \bigcup_{\ell \geq 0} V_\ell^2$$

Thus we can upper bound $T_\alpha(n)$ as

$$\begin{aligned} T_\alpha(n) &= \sup_{\lambda \in \Lambda} \sum_{\ell \geq 0} \sum_{(u,v) \in V_\ell^2} \lambda_{u,v} f_\alpha(r_{u,v}) \\ &\leq \sup_{\lambda \in \Lambda} \sum_{\ell \geq 0} f_\alpha(2^{-\ell+1}\sqrt{n}) \sum_{(u,v) \in V_\ell^2} \lambda_{u,v} \\ &\leq \sum_{0 \leq \ell \leq L(n)} f_\alpha(2^{-\ell+1}\sqrt{n}) \sup_{\lambda \in \Lambda} \sum_{(u,v) \in V_\ell^2} \lambda_{u,v} + f_\alpha(2^{-L(n)}\sqrt{n}) \sup_{\lambda \in \Lambda} \sum_{\ell > L(n)} \sum_{(u,v) \in V_\ell^2} \lambda_{u,v}. \end{aligned} \tag{18}$$

We now upper bound each of the terms

$$\sup_{\lambda \in \Lambda} \sum_{(u,v) \in V_\ell^2} \lambda_{u,v}$$

for $0 \leq \ell \leq L(n)$. Let $(u,v) \in V_\ell^2$; by definition this implies $r_{u,v} \geq 2^{-\ell}\sqrt{n}$. Now pick $i \in \{1, \dots, 4^{\ell+1}\}$ such that $u \in V_{\ell+1,i}$. Observe that for any $w \in V_{\ell+1,i}$, we have $r_{u,w} < 2^{-\ell}\sqrt{n}$. Therefore $v \in V_{\ell+1,i}^c$, and hence

$$V_\ell^2 \subset \bigcup_{i=1}^{4^{\ell+1}} \big(V_{\ell+1,i} \times V_{\ell+1,i}^c\big).$$

Using this observation and Lemma 7

$$\begin{aligned}\sup_{\lambda\in\Lambda}\sum_{(u,v)\in V_\ell^2}\lambda_{u,v} &\leq \sup_{\lambda\in\Lambda}\sum_{i=1}^{4^{\ell+1}}\sum_{u\in V_{\ell+1,i},v\in V_{\ell+1,i}^c}\lambda_{u,v}\\ &\leq \sum_{i=1}^{4^{\ell+1}}K_1 n^{\varepsilon/2}(4^{-\ell-1}n)^{2-\tilde{\alpha}/2}\\ &\leq K_1 4^{\ell(\tilde{\alpha}/2-1)+1}n^{2-\tilde{\alpha}/2+\varepsilon/2}, \end{aligned}\tag{19}$$

for $0\leq\ell\leq L(n)$ and with probability $1-o(1)$ as $n\to\infty$, and where we have defined

$$\tilde{\alpha}\triangleq\min\{3,\alpha\}.$$

Consider now $\ell>L(n)$. By Lemma 8,

$$\sup_{\lambda\in\Lambda}\sum_{v\in V\setminus\{u\}}\lambda_{u,v}\leq K_2\log(n)$$

with probability $1-o(1)$ as $n\to\infty$. Hence

$$\begin{aligned}\sup_{\lambda\in\Lambda}\sum_{\ell>L(n)}\sum_{(u,v)\in V_\ell^2}\lambda_{u,v} &\leq \sum_{u\in V}\sup_{\lambda\in\Lambda}\sum_{v\in V\setminus\{u\}}\lambda_{u,v}\\ &\leq K_2 n\log(n). \end{aligned}\tag{20}$$

Substituting (19) and (20) into (18), we obtain

$$\begin{aligned}T_\alpha(n) &\leq \sum_{\ell=0}^{L(n)}f_\alpha(2^{-\ell+1}\sqrt{n})\sup_{\lambda\in\Lambda}\sum_{(u,v)\in V_\ell^2}\lambda_{u,v}+f_\alpha(2^{-L(n)}\sqrt{n})\sup_{\lambda\in\Lambda}\sum_{\ell>L(n)}\sum_{(u,v)\in V_\ell^2}\lambda_{u,v}\\ &\leq 8K_1\sum_{\ell=0}^{L(n)}4^{-\ell(\tilde{\alpha}/2-1)}n^{\tilde{\alpha}/2-1}4^{\ell(\tilde{\alpha}/2-1)}n^{2-\tilde{\alpha}/2+\varepsilon/2}+K_2\big(2^{-L(n)}n\big)^{\alpha/2-1}n\log(n)\\ &= L(n)O\big(n^{1+\varepsilon/2}\big)+O\big(n^{1+\varepsilon}\big)\\ &= O\big(n^{1+\varepsilon}\big), \end{aligned}$$

with probability $1-o(1)$ as $n\to\infty$.

VII. Proof of Theorem 2

We construct a communication scheme and lower bound the rate $\rho_{u,v}(n)$ that it achieves for any $(u,v)\in\Pi$. Define

$$\Pi_\ell\triangleq\Pi\cap V_\ell^2.$$

Consider the partition $\{A_{\ell-1,i}\}_{i=1}^{4^{\ell-1}}$ of A. For each $\ell\geq 3$, construct bigger squares by joining 4 squares in $\{A_{\ell-1,i}\}_{i=1}^{4^{\ell-1}}$. There are four different ways of doing this, yielding "shifted" versions of bigger squares (see Figure 4). Call $\{\widetilde{A}_{\ell-1,i}^j\}_i$ for $j\in\{1,\dots,4\}$ the resulting bigger squares.

Define

$$\Pi_{\ell,i}^j\triangleq\Pi_\ell\cap(\widetilde{A}_{\ell-1,i}^j\times\widetilde{A}_{\ell-1,i}^j).$$

Let $(u,v)\in\Pi_{\ell,i}^j$, then $r_{u,v}<2^{-\ell+1}\sqrt{n}$. Thus if $u\in A_{\ell-1,i}$ then v is in either $A_{\ell-1,i}$ or one of its neighbors. Hence there exists i' and j' such that $(u,v)\in\widetilde{A}_{\ell-1,i'}^{j'}$. Therefore

$$\Pi_\ell=\bigcup_{i,j}\Pi_{\ell,i}^j.$$

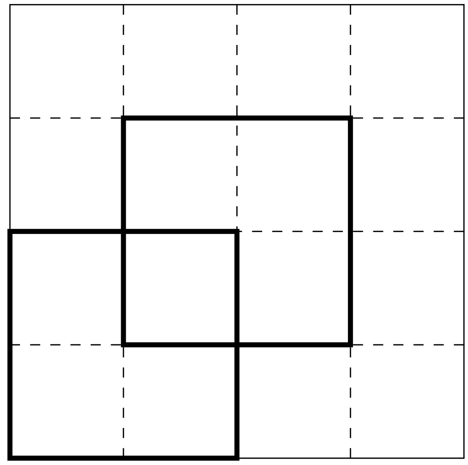

Fig. 4. Two of the four ways of defining the bigger squares $\{\widetilde{A}^j_{\ell-1,i}\}_i$ (bold lines) for $j \in \{1, \ldots, 4\}$ from the smaller squares $\{A_{\ell-1,i}\}_{i=1}^{4^{\ell-1}}$ (dashed lines).

For $\ell = 2$ the number of ways to define bigger squares is smaller, and the same approach used in the following for general ℓ will work as well. For $\ell \in \{0, 1\}$, it is not necessary to define bigger squares, essentially the same approach as for the general case will again work in this case.

We time share between $\ell \in \{0, \ldots, L(n)+1\}$ and between $j \in \{1, \ldots, 4\}$. For each $\ell \in \{0, \ldots, L(n)\}$ and $j \in \{1, \ldots, 4\}$, we communicate between the source-destination pairs within each of the squares $\{\widetilde{A}^j_{\ell-1,i}\}_i$ (with the exception of $\ell \in \{0, 1\}$ where construction of bigger squares is not necessary). A sketch of this traffic decomposition is shown in Figure 2 in Section IV-B. This time-sharing results in a total rate loss of a factor

$$\frac{1}{4(L(n)+2)} \leq n^{-o(1)}.$$

Using hierarchical relaying (for $\alpha \in (2, 3]$) or multi-hop (for $\alpha > 3$) within each square $\widetilde{A}^j_{\ell-1,i}$ according to the traffic matrix given by $\Pi^j_{\ell,i}$ yields by Lemma 9 a per-node rate of

$$\rho_{u,v}(n) \geq n^{-o(1)} \big(4^{-\ell} n\big)^{1-\tilde{\alpha}/2}, \tag{21}$$

simultaneously for all

$$(u, v) \in \bigcup_i \Pi^j_{\ell,i}.$$

Now note that each source-destination pair $(u, v) \in \Pi^j_{\ell,i}$ is at least at distance $2^{-\ell}\sqrt{n}$, and hence

$$\lambda_{u,v} f_\alpha(r_{u,v}) \geq \rho_{u,v}(n) f_\alpha(2^{-\ell}\sqrt{n}) \geq n^{-o(1)} \tag{22}$$

for any $\ell \in \{0, \ldots, L(n)\}$.

For $\ell > L(n)$, we use the multi-hop scheme to communicate between source-destination pairs in

$$\bigcup_{\ell > L(n)} \bigcup_i \Pi^j_{\ell,i},$$

time sharing between $j \in \{1, \ldots, 4\}$ and the four ways of defining bigger squares at level $\ell = L(n)$. By Lemma 9, this achieves a per-node rate of

$$\rho_{u,v}(n) \geq n^{-o(1)} (4^{-L(n)} n)^{-1/2-o(1)} \geq n^{-o(1)},$$

simultaneously for all

$$(u, v) \in \bigcup_{\ell > L(n)} \bigcup_i \Pi^j_{\ell,i}.$$

From this and using $f(r) \geq 1$ for $r > 0$, we obtain

$$\lambda_{u,v} f_\alpha(r_{u,v}) \geq \rho_{u,v}(n) \geq n^{-o(1)} \tag{23}$$

for any $\ell > L(n)$ and $(u,v) \in \Pi_\ell$.

Combining (22) and (23), and using that

$$\Pi = \bigcup_{\ell \geq 0} \Pi_\ell,$$

shows the desired result.

VIII. Proof of Theorem 4

We start with some definitions and auxiliary results concerning routing over capacitated tree graphs. Let $G = (V_G, E_G)$ be a tree graph, with edge capacities $c(e)$ in both directions for $e \in E_G$. For a non-terminal node u in G, let $\mathcal{D}(u)$ be the descendants of u in G (including u itself, i.e., $u \in \mathcal{D}(u)$), and let $\mathcal{L}(u)$ denote the leaf nodes in $\mathcal{D}(u)$. Let $e = (u,v) \in E_G$, and assume that v is the parent of u in G; with slight abuse of notation, define $\mathcal{D}(e) \triangleq \mathcal{D}(u)$.

Denote by $\Lambda_G \subset \mathbb{R}_+^{|V_G| \times |V_G|}$ the set of feasible rate matrices for G (i.e., the set of flow rates that can be routed through G). For a matrix of flow rates $\lambda \in \mathbb{R}_+^{|V_G| \times |V_G|}$, define

$$\gamma_\lambda^* \triangleq \sup\{b \geq 0 : b\lambda \in \Lambda_G\}.$$

For $e \in E_G$, let

$$d_\lambda(e) \triangleq \sum_{\substack{u \in \mathcal{D}(e) \\ v \notin \mathcal{D}(e)}} (\lambda_{u,v} + \lambda_{v,u}).$$

Lemma 10. *Let $G = (V_G, E_G)$ be an undirected capacitated tree graph, and $\lambda \in \mathbb{R}_+^{|V_G| \times |V_G|}$ a matrix of flow rates. Then*

$$\gamma_\lambda^* = \min_{e \in E_G} \frac{c(e)}{d_\lambda(e)}.$$

Proof. Let $e \in E_G$, and consider $u \in \mathcal{D}(e)$, $v \notin \mathcal{D}(e)$. Since G is a tree, the only way to route data from u to v and from v to u is through e. Therefore

$$\gamma_\lambda^* d_\lambda(e) \leq c(e)$$

for all $e \in E_G$, and thus

$$\gamma_\lambda^* \leq \min_{e \in E_G} \frac{c(e)}{d_\lambda(e)}.$$

Conversely, suppose

$$\gamma_\lambda^* < \min_{e \in E_G} \frac{c(e)}{d_\lambda(e)}.$$

Then the load over every edge $e \in E_G$ is strictly less than $c(e)$, and hence it is possible to increase the flow for each (u,v) pair by a strictly positive amount. This contradicts the definition of γ_λ^*, and hence shows that

$$\gamma_\lambda^* \geq \min_{e \in E_G} \frac{c(e)}{d_\lambda(e)}.$$

$\square$

We are now ready to embark on the proof of Theorem 4. Set

$$\lambda^* \triangleq \rho_\lambda^*(n)\lambda,$$

and note that[4] $\lambda^* \in \Lambda$. Thus by Lemma 7, we have for any $\ell \in \{1, \dots, L(n)\}$ and $i \in \{1, \dots, 4^\ell\}$,

$$\max\left\{\sum_{u\in V_{\ell,i}, v\in V_{\ell,i}^c} \lambda^*_{u,v}, \sum_{u\in V_{\ell,i}^c, v\in V_{\ell,i}} \lambda^*_{u,v}\right\} \le K_1 n^\varepsilon (4^{-\ell} n)^{2-\min\{3,\alpha\}/2}.$$

Noting that

$$\max\left\{\sum_{u\in V_{\ell,i}, v\in V_{\ell,i}^c} \lambda_{u,v}, \sum_{u\in V_{\ell,i}, v\in V_{\ell,i}^c} \lambda_{v,u}\right\} \ge \frac{1}{2} D_\lambda(V_{\ell,i}), \tag{24}$$

yields

$$\begin{aligned}
\rho^*_\lambda(n) &\le \min_{\ell\in\{1,\dots,L(n)\}} K n^\varepsilon (4^{-\ell} n)^{2-\min\{3,\alpha\}/2} \min_{i\in\{1,\dots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \\
&= \min_{\ell\in\{1,\dots,L(n)\}} K n^\varepsilon g_\alpha(4^{-\ell} n) \min_{i\in\{1,\dots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})}, \qquad (25)
\end{aligned}$$

for some constant $K > 0$.

For $\ell > L(n)$ and any $i \in \{1, \dots, 4^\ell\}$, Lemma 8 shows that

$$\begin{aligned}
\sum_{u\in V_{\ell,i}, v\in V_{\ell,i}^c} \lambda^*_{u,v} &\le \sum_{u\in V_{\ell,i}} \sum_{v\neq u} \lambda^*_{u,v} \\
&\le K_2 \left|V_{\ell,i}\right| \log(n) \\
&\le K_2 \left|V_{L(n),i}\right| \log(n) \\
&\le n^{o(1)},
\end{aligned}$$

where we have assumed without loss of generality that $V_{\ell,i} \subset V_{L(n),i}$. Therefore

$$\rho^*_\lambda(n) \le \frac{n^{o(1)}}{\sum_{u\in V_{\ell,i}, v\in V_{\ell,i}^c} \lambda_{u,v}},$$

and by a similar argument

$$\rho^*_\lambda(n) \le \frac{n^{o(1)}}{\sum_{u\in V_{\ell,i}^c, v\in V_{\ell,i}} \lambda_{u,v}}.$$

Using (24),

$$\begin{aligned}
\rho^*_\lambda(n) &\le \frac{n^{o(1)}}{D_\lambda(V_{\ell,i})} \\
&\le \frac{n^{o(1)} g_\alpha(4^{-\ell} n)}{D_\lambda(V_{\ell,i})},
\end{aligned}$$

since $g_\alpha \ge 1$. Hence

$$\begin{aligned}
\rho^*_\lambda(n) &\le \min_{\ell > L(n)} n^{o(1)} g_\alpha(4^{-\ell} n) \min_{i\in\{1,\dots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \\
&\le \min_{\ell = \log(n)} n^{o(1)} g_\alpha(4^{-\ell} n) \min_{i\in\{1,\dots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})}.
\end{aligned}$$

Combining this with (25) yields

$$\rho^*_\lambda(n) = O\big(n^\varepsilon \phi^*_\lambda(n)\big). \tag{26}$$

Now, construct a graph $G = (V_G, E_G)$ as follows. G is a full tree (i.e., all its leaf nodes are on the same level). G has n leaves, each of them representing an element of V. To simplify notation, we assume that $V \subset V_G$, so that the leaves of G are exactly the elements of $V \subset V_G$. Whenever the distinction is

[4]To be precise, $(1-\delta)\lambda^* \in \Lambda$ for any $\delta \in (0,1)$. However, since we are only interested in the scaling behavior of Λ, we can safely ignore the additional $(1-\delta)$ factor.

relevant, we use u, v for nodes in $V \subset V_G$ and μ, ν for nodes in $V_G \setminus V$ in the following. The non-terminal nodes of G correspond to $V_{\ell,i}$ for all $\ell \in \{0, \dots, L(n)\}$, $i \in \{1, \dots, 4^\ell\}$, with hierarchy induced by the one on $A(n)$ (see Figure 3 in Section IV-C). Thus, nodes in V_G at level $\ell < L(n)$ have 4 children, nodes at level $\ell = L(n)$ have between $4^{-L(n)-1}n$ and $4^{-L(n)+1}n$ children (with high probability, by Lemma 5), and nodes at level $\ell = L(n) + 1$ are the leaves of the tree. To understand the relation between V_G and V, we define the *representative* $\mathcal{R} : V_G \to 2^V$ as follows. For a leaf node $u \in V \subset V_G$ of G, let

$$\mathcal{R}(u) \triangleq \{u\}.$$

For $\mu \in V_G$ at level $L(n)$, choose $\mathcal{R}(\mu) \subset \mathcal{L}(\mu) \subset V$ such that

$$|\mathcal{R}(\mu)| = 4^{-L(n)-1}n.$$

This is possible with probability $1 - o(1)$ as $n \to \infty$ by Lemma 5. Finally, for $\mu \in V_G$ at level $\ell < L(n)$, and with children $\{\nu_i\}_{j=1}^4$, let

$$\mathcal{R}(\mu) \triangleq \bigcup_{j=1}^{4} \mathcal{R}(\nu_j).$$

We now define an edge capacity $c(\mu, \nu)$ for each edge $(\mu, \nu) \in E_G$. If μ is a leaf of G and ν its parent, set

$$c(\mu, \nu) = c(\mu, \nu) \triangleq 1. \tag{27}$$

If μ is a non-terminal node at level ℓ in G and ν its parent, then set

$$c(\mu, \nu) = c(\nu, \mu) \triangleq (4^{-\ell}n)^{2-\min\{3,\alpha\}/2}. \tag{28}$$

We argue now that if a multicommodity flow λ corresponding to a traffic matrix of the wireless network (i.e., only the leaf nodes of G are sources or destinations) can be routed through G, then $n^{o(-1)}\lambda \in \Lambda$ (i.e., almost the same flow can be reliably transmitted over the wireless network). The idea (that will be made precise) is the following. To transmit information from a non-terminal node $\mu \in V_G$ to its parent node ν, we split the message at each node in $\mathcal{R}(\mu)$ into four parts and send one part to each node in $\mathcal{R}(\nu)$. In other words, we distribute the message by a factor four over the wireless network. To transmit information from a node $\mu \in V_G$ with non-terminal children $\{\nu_j\}_{j=1}^4$ to one of them, say ν_1, we send the message parts from each $\{\mathcal{R}(\nu_j)\}_{j=2}^4$ to a corresponding node in $\mathcal{R}(\nu_1)$ and combine them there. In other words, we concentrate the message by a factor four over the wireless network. The scheme is bootstrapped at the leaves $V \subset V_G$ of G (where, by our definition of λ, all traffic originates and ends) as follows. To send a message from a leaf node $u \in V \subset V_G$ to its parent ν in G, the message is split at u into $|\mathcal{R}(\nu)|$ equal pieces, and one piece is sent to each node in $\mathcal{R}(\nu)$ over the wireless network. In other words, we distribute the message again over the wireless network, but this time by a factor of $|\mathcal{R}(\nu)|$. To send a message to a leaf node $u \in V \subset V_G$ from its parent ν in G, each node in $\mathcal{R}(\nu)$ sends its piece of the message to u over the wireless network. Thus, again we concentrate the message over the network, but this time by a factor of $|\mathcal{R}(\nu)|$.

We now analyze this scheme in more detail. Note first that by time sharing between the $L(n) + 1$ non-terminal levels of the tree, and by appropriate spatial reuse within each level, we only loose a factor of at most

$$\widetilde{K} \frac{1}{4(L(n) + 1)} \leq n^{-o(1)}$$

for some constant $\widetilde{K} > 0$ in rate. Hence it is sufficient to consider communication between a non-terminal node of G and its children.

We first consider communication up the tree (i.e., towards the root). Let $u \in V \subset V_G$ be a leaf node of G and ν be its parent. To send traffic at rate $c(u, \nu)$ from u to ν, node u splits its traffic into

$|\mathcal{R}(\nu)| = 4^{-L(n)-1}n$ equal parts and sends each part to one node in $\mathcal{R}(\nu)$. Recall $\mathcal{R}(\nu) \subset V_{L(n),i}$ for some i. Since

$$r_{u,v} \leq 2(4^{-L(n)}n)^{1/2}$$

for any $u, v \in V_{L(n),i}$, communicating between u and v, we incur a power loss of

$$r_{u,v}^{-\alpha} \geq 2^{-\alpha}(4^{-L(n)}n)^{-\alpha/2} \geq 2^{-\alpha}n^{-\log^{-1/2}(n)\alpha/2},$$

and hence we can communicate between u and v at a rate of at least

$$\log\left(1 + P2^{-\alpha}n^{-\log^{-1/2}(n)\alpha/2}\right) \geq n^{-o(1)}.$$

By time sharing between all the destination nodes in $\mathcal{R}(\nu)$, and since all message parts are only $|\mathcal{R}(\nu)|^{-1}$ of the size of the original message, all message parts of u can be transmitted from u to $\mathcal{R}(\nu)$ at this rate. By Lemma 5,

$$|\mathcal{L}(\nu)| \leq 4^{-L(n)+1}n = 4n^{\log^{-1/2}(n)},$$

and hence further time sharing between all source nodes in $\mathcal{L}(\nu)$, we can communicate simultaneously from all leaf nodes $u \in \mathcal{L}(\nu)$ to $\mathcal{R}(\nu)$ at a rate at least

$$4^{-1}n^{-\log^{-1/2}(n)}n^{-o(1)} \geq n^{-o(1)}c(u, \nu).$$

Let now $\nu \in V_G$ be a node in level $\ell < L(n)$ in G and let $\{\mu_j\}_{j=1}^4$ be its children. Since

$$|\mathcal{R}(\mu_j)| = 4^{-\ell(n)-2}n$$

for all $j \in \{1, \ldots, 4\}$, we can find a one-to-one correspondence between $\mathcal{R}(\mu_j)$ and $\mathcal{R}(\mu_k)$. Choose an arbitrary such correspondence for each $j, k \in \{1, \ldots, 4\}, j \neq k$. Now, since the multicommodity flow corresponds to a traffic matrix λ for the wireless network, we know that the traffic to be sent from μ_j to ν originates at one or several nodes in $\mathcal{L}(\mu_j)$. Thus by construction of the previous stages, all nodes in $\mathcal{R}(\mu_j)$ possess an equal part of the total message to be transmitted from μ_j to ν. Split each such message part further into four equal parts and consider one particular node $u \in \mathcal{R}(\mu_k)$. The first part of the message stays at u. The other three parts are to be transmitted to the corresponding nodes in $\{\mathcal{R}(\mu_j)\}_{j \neq k}$. Time sharing between all 12 possible (j, k) pairs, we only incur a constant loss. Hence we can focus on communication between a particular (j, k) pair. Note that we are now in a situation with $4^{-\ell(n)-2}n$ nodes in $\mathcal{R}(\mu_j)$, each with a message of equal size for the corresponding node in $\mathcal{R}(\mu_k)$. This is a permutation traffic, and by Lemma 9 we can therefore use hierarchical relaying (for $\alpha \in (2, 3]$) or multi-hop for ($\alpha > 3$) to transmit at a total rate up to

$$n^{-o(1)}(4^{-\ell}n)^{2-\min\{3,\alpha\}/2} = n^{-o(1)}c(\mu, \nu).$$

Consider now communication down the tree (i.e., away from the root). Communication between ν and μ works in the same fashion by concentrating the messages. The same arguments as in the previous two paragraphs show that any rate up to (27) or (28) are achievable up to a factor $n^{-o(1)}$. Time sharing between the two directions, yields an additional rate loss of a factor $1/2$. Together, this shows that if a multicommodity flow λ corresponding to a traffic matrix for the wireless network can be routed through G then $n^{-o(1)}\lambda$ can be communicated over the wireless network. In other words,

$$\lambda \in \Lambda_G \;\Rightarrow\; n^{-o(1)}\lambda \in \Lambda,$$

and we therefore have

$$\begin{aligned}
\rho_\lambda^*(n) &= \sup\{b \geq 0 : b\lambda \in \Lambda\} \\
&\geq \sup\{b \geq 0 : bn^{o(1)}\lambda \in \Lambda_G\} \\
&= \sup\{n^{-o(1)}b \geq 0 : b\lambda \in \Lambda_G\} \\
&= n^{-o(1)}\gamma_\lambda^*(n).
\end{aligned} \tag{29}$$

Now, by Lemma 10, we know that by optimally routing over G (which, since G is a tree, is trivial), we can achieve

$$\gamma_\lambda^*(n) = \min_{e \in E_G} \frac{c(e)}{d_\lambda(e)}. \tag{30}$$

Consider now an edge $e = (\mu, \nu) \in E_G$, and assume that ν is the node closer to the root of G. Let ℓ be the level of μ in the tree. Then, by construction, $c(e)$ is only a function of ℓ and given by either (27) or (28). Moreover, $d_\lambda(e)$ is either equal to $D_\lambda(\{u\})$ for some $u \in V$ if $\ell = L(n)+1$, or equal to $D_\lambda(V_{\ell,i})$ for some i if $\ell \le L(n)$. Therefore

$$\min_{e \in E_G} \frac{c(e)}{d_\lambda(e)} = \min\left\{ \min_{u \in V} \frac{1}{D_\lambda(\{u\})}, \min_{\ell \in \{1,\ldots,L(n)\}} (4^{-\ell} n)^{2-\min\{3,\alpha\}/2} \min_{i \in \{1,\ldots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \right\}. \tag{31}$$

Now, by Lemma 6, $r_{u,v} \ge n^{-1/2-\delta} > n^{-1}$ for all $u, v \in V$ with probability $1 - o(1)$, and hence we have for $\ell = \log(n)$ that

$$\begin{aligned}
\max_{u \in V} D_\lambda(\{u\}) &= \max_{i \in \{1,\ldots,4^\ell\}} \max_{u \in V_{\ell,i}} D_\lambda(\{u\}) \\
&= \max_{i \in \{1,\ldots,4^\ell\}} D_\lambda(V_{\ell,i}),
\end{aligned}$$

and thus for $\ell = \log(n) > L(n)$,

$$\begin{aligned}
\min_{u \in V} \frac{1}{D_\lambda(\{u\})} &= \min_{i \in \{1,\ldots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \\
&\ge \frac{1}{g_\alpha(4^{-L(n)} n)} g_\alpha(4^{-\ell} n) \min_{i \in \{1,\ldots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \\
&\ge n^{-o(1)} g_\alpha(4^{-\ell} n) \min_{i \in \{1,\ldots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})}.
\end{aligned}$$

Together with (31), this yields

$$\begin{aligned}
\min_{e \in E_G} \frac{c(e)}{d_\lambda(e)} &= \min\left\{ \min_{u \in V} \frac{1}{D_\lambda(\{u\})}, \min_{\ell \in \{1,\ldots,L(n)\}} g_\alpha(4^{-\ell} n) \min_{i \in \{1,\ldots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \right\} \\
&\ge n^{-o(1)} \min_{\ell \in \{1,\ldots,L(n)\} \cup \{\log(n)\}} g_\alpha(4^{-\ell} n) \min_{i \in \{1,\ldots,4^\ell\}} \frac{1}{D_\lambda(V_{\ell,i})} \\
&= n^{-o(1)} \phi_\lambda^*(n).
\end{aligned} \tag{32}$$

Combining (26), (29), (30), and (32), shows that

$$\begin{aligned}
n^{-o(1)} \phi_\lambda^*(n) &\le n^{-o(1)} \gamma_\lambda^*(n) \\
&\le \rho_\lambda^*(n) \\
&= O\big(n^\varepsilon \phi_\lambda^*(n)\big),
\end{aligned}$$

completing the proof.

IX. Conclusions

We considered general traffic patterns in wireless networks with n nodes uniformly distributed on $[0, \sqrt{n}]^2$ and with a Gaussian fading channel model. We first focused attention on permutation traffic (i.e., every node in the network is source and destination exactly once) and showed that for every source-destination pair at distance r, we can guarantee a rate of

$$n^{-o(1)} f_\alpha(r)^{-1} = n^{-o(1)} \begin{cases} r^{2-\min\{3,\alpha\}} & \text{if } r \ge 1, \\ 1 & \text{if } 0 < r < 1, \\ \infty & \text{if } r = 0. \end{cases}$$

Moreover, this is essentially the best guarantee possible, since the sum of all rates, weighted by $f_\alpha(r)$ for source-destination pairs at distance r, is upper bounded by $O(n^{1+\varepsilon})$ as $n \to \infty$ for any $\varepsilon > 0$.

We then focused on completely general traffic. More precisely, for $\lambda \in \mathbb{R}^{n\times n}$, we asked if λ is an element of the n^2 dimensional capacity region $\Lambda(n)$ of the wireless network. We provided an asymptotic answer to this question in terms of $\Theta(n)$ easily computable parameters of the network topology and traffic demands. This resulted in a complete asymptotic characterization of the capacity region $\Lambda(n)$ of the wireless network.

X. Acknowledgments

The authors would like to thank David Tse and Greg Wornell for helpful discussions.

References

[1] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[2] L.-L. Xie and P. R. Kumar. A network information theory for wireless communication: Scaling laws and optimal operation. *IEEE Transactions on Information Theory*, 50(5):748–767, May 2004.

[3] A. Jovičić, P. Viswanath, and S. R. Kulkarni. Upper bounds to transport capacity of wireless networks. *IEEE Transactions on Information Theory*, 50(11):2555– 2565, November 2004.

[4] F. Xue, L.-L. Xie, and P. R. Kumar. The transport capacity of wireless networks over fading channels. *IEEE Transactions on Information Theory*, 51(3):834–847, March 2005.

[5] L.-L. Xie and P. R. Kumar. On the path-loss attenuation regime for positive cost and linear scaling of transport capacity in wireless networks. *IEEE Transactions on Information Theory*, 52(6):2313–2328, June 2006.

[6] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran. Closing the gap in the capacity of wireless networks via percolation theory. *IEEE Transactions on Information Theory*, 53(3):1009–1018, March 2007.

[7] P. Gupta and P. R. Kumar. Towards an information theory of large networks: An achievable rate region. *IEEE Transactions on Information Theory*, 49(8):1877–1894, August 2003.

[8] L.-L Xie and P. R. Kumar. An achievable rate for the multiple-level relay channel. *IEEE Transactions on Information Theory*, 51(4):1348–1358, April 2005.

[9] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, September 2005.

[10] S. Aeron and V. Saligrama. Wireless ad hoc networks: Strategies and scaling laws for the fixed SNR regime. *IEEE Transactions on Information Theory*, 53(6):2044–2059, June 2007.

[11] A. Özgür, O. Lévêque, and D. N. C. Tse. Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks. *IEEE Transactions on Information Theory*, 53(10):3549–3572, October 2007.

[12] U. Niesen, P. Gupta, and D. Shah. On capacity scaling in arbitrary wireless networks. *submitted to IEEE Transactions on Information Theory*, November 2007. Available online at http://arxiv.org/abs/0711.2745.

[13] R. Madan, D. Shah, and O. Lévêque. Product multicommodity flow in wireless networks. *IEEE Transactions on Information Theory*, 54(4):1460–1476, April 2008.

[14] F. T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, 46(6):787–832, November 1999.

[15] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, June 1995.

[16] S. Subramanian, S. Shakkottai, and P. Gupta. On optimal geographic routing in wireless networks with holes and non-uniform traffic. In *Proceedings of IEEE Infocom*, May 2007.

[17] S. Subramanian, S. Shakkottai, and P. Gupta. Optimal geographic routing for wireless networks with near-arbitrary holes and traffic. In *Proceedings of IEEE Infocom*, April 2008.

[18] S. H. A. Ahmad, A. Jovičić, and P. Viswanath. On outer bounds to the capacity region of wireless networks. *IEEE Transactions on Information Theory*, 52(6):2770–2776, June 2006.

[19] M. Franceschetti, M. D. Migliore, and P. Minero. The capacity of wireless networks; information-theoretic and physical limits. In *Allerton Conference on Communication, Control, and Computing*, September 2007.