Exact Capacity Scaling
of Extended Wireless Networks

Ayfer Özgür, Olivier Lévêque
Faculté Informatique et Communications
Ecole Polytechnique Fédérale de Lausanne
1015 Lausanne, Switzerland
{ayfer.ozgur,olivier.leveque}@epfl.ch

David Tse
Department of EECS
University of California at Berkeley
Berkeley, CA 94720, USA
dtse@eecs.berkeley.edu

Abstract— n source and destination pairs randomly located in an area extending linearly with n want to communicate with each other. Signals transmitted from one user to another at distance r apart are subject to a power attenuation of $r^{-\alpha}$ and random phase changes. Classical multihop architectures that decode and forward packets can deliver a $\sqrt{n}$-scaling of the aggregate throughput, while recently proposed hierarchical cooperation achieves $n^{2-\alpha/2}$-scaling, which is superior to multi-hop for $\alpha < 3$. The study of information theoretic upper bounds has revealed the optimality of multi-hop for $\alpha > 4$, while the moderate-attenuation regime ($2 \leq \alpha \leq 4$) remains uncharacterized. We close this gap by deriving a tight upper bound on the scaling of the aggregate throughput, valid for all $\alpha \geq 2$. Our result shows that the mentioned schemes are scaling-optimal, namely that no other scheme can beat hierarchical cooperation when $\alpha < 3$, nor can it beat classical multi-hop when $\alpha \geq 3$. The key ingredient is a careful evaluation of the scaling of the cut-set bound.

I. INTRODUCTION

The seminal paper by Gupta and Kumar [1] initiated the study of scaling laws in large ad-hoc wireless networks. Their by-now familiar model considers n nodes randomly located in the unit disk, each of which wants to communicate to a random destination node at a rate $R(n)$ bits/second. They ask what is the maximally achievable scaling of the total throughput $T(n) = nR(n)$ with the system size n. They showed that classical multihop architectures with conventional single-user decoding and forwarding of packets cannot achieve a scaling better than $O(\sqrt{n})$, and that this scaling is indeed achievable by a multihop scheme that uses only nearest-neighbor communication. Although their original set-up considers a fixed area network, their results readily apply to the case where the network size scales with n to cover an increasing geographical area, keeping the density of nodes fixed.

A natural question is whether this result is a consequence of the physical-layer assumptions or whether one can do better using more sophisticated physical-layer processing. A recent work [2] shows that the answer is the latter: authors exhibit a hierarchichal cooperation scheme which yields a throughput scaling of $\Theta(n^{2-\alpha/2-\epsilon})$ bits/second for any $\epsilon > 0$, when the signals transmitted from one user to another at distance r apart are subject to a power attenuation of $r^{-\alpha}$ and to random phase changes. Therefore for $\alpha < 3$, this scheme performs strictly better than multihop strategies. However, it is not clear if one can do even better. In fact, how does the information

theoretic capacity of the network scale? Since the publication of [1], there have been several works dealing with information theoretic scaling laws for wireless adhoc networks [3], [4], [5], [6], [7], [8]. For the deterministic channel model with no random phases, [8] establishes a $O(n^{\frac{1}{\alpha+8}}\sqrt{n})$-scaling for the aggregate throughput, which is not much larger than the $\Theta(\sqrt{n})$-scaling achieved by multihop, even for small values of α. For the more general channel model with random phases considered in this paper, the best result to date [7] shows that whenever $\alpha > 4$ ($\alpha > 4.5$ in [6], $\alpha > 5$ in [4]), nearest-neighbour multihop is order-optimal. No scaling laws exist for the moderate attenuation regime when $2 \leq \alpha \leq 4$. By evaluating a cutset upper bound in the following, we identify exactly the capacity scaling law in extended adhoc wireless networks for *all* values of $\alpha \geq 2$. Our result shows that the threshold above which nearest neighbor multihop is order-optimal is indeed $\alpha = 3$. For $2 \leq \alpha \leq 3$, the hierarchical cooperation scheme is the one meeting the upper bound.

II. MODEL

There are n nodes uniformly and independently distributed in a square of area $\sqrt{n} \times \sqrt{n}$. Every node is both a source and a destination. The sources and destinations are randomly paired up one-to-one without any consideration on node locations. Each source has the same traffic rate $R(n)$ to send to its destination node and a common average transmit power budget of P Watts. The total throughput of the system is $T(n) = nR(n)$.

We assume that communication takes place over a flat channel of bandwidth W Hz around a carrier frequency of f_c, $f_c \gg W$. The complex baseband-equivalent channel gain between node i and node k at time m is given by:

$$H_{ik}[m] = \sqrt{G}\, r_{ik}^{-\alpha/2} \exp(j\theta_{ik}[m]) \qquad (1)$$

where r_{ik} is the distance between the nodes, $\theta_{ik}[m]$ is the random phase at time m, uniformly distributed in $[0, 2\pi]$ and $\{\theta_{ik}[m]; 1 \leq i \leq n, 1 \leq k \leq n\}$ is a collection of i.i.d random processes. The $\theta_{ik}[m]$'s and the r_{ik}'s are also assumed to be independent. The parameters G and $\alpha \geq 2$ are assumed to be constants; α is called the power path loss exponent.

Note that the channel is random, depending on the location of the users and the phases. The locations are assumed to
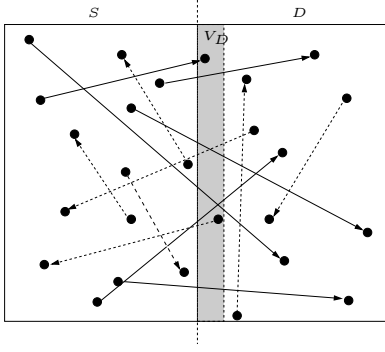
Fig. 1. The cut-set considered in the proof of Theorem 1. The communication requests that pass across the cut from left to right are depicted in bold lines.

be fixed over the duration of the communication. The phases are assumed to vary in a stationary ergodic manner (fast fading). We assume that the channel gains are known at all the nodes. The received signal at a node is the sum of the signals received from all the nodes in the network plus white circularly symmetric Gaussian noise of variance 1 per symbol.

III. CUTSET UPPER BOUND

The following theorem is the main result of this paper. The rest of the paper is devoted to its proof.

Theorem 1: For any $\epsilon > 0$, the aggregate throughput of an extended random network with n nodes is bounded above by

$$T(n) \leq \begin{cases} K\, n^{2-\alpha/2+\epsilon} & 2 \leq \alpha < 3 \\ K\, n^{1/2+\epsilon} & \alpha \geq 3 \end{cases}$$

with high probability (w.h.p.) for a constant $K > 0$ independent of n.

Proof of Theorem 1: We consider a cut dividing the $\sqrt{n} \times \sqrt{n}$ network area into two equal halves (see Figure 1). The following lemma states several properties that will be used in the sequel. We skip its proof due to space constraints.

Lemma 1: The random network with random source-destination pairing satisfies the following properties w.h.p.:

a) Let the network area be divided into n squarelets of unit area. Then, there are less than $\log n$ nodes inside all squarelets.

b) Let the network area be divided into $\frac{n}{2\log n}$ squarelets each of area $2\log n$. Then, there is at least one node inside each squarelet.

c) The number of communication requests with sources on the left-half network and destinations on the right-half network is between $((1-\delta)n/4, (1+\delta)n/4)$ for any $\delta > 0$.

We are interested in bounding above the sum of the rates of communications passing through the cut from left to right. By Part (c) of the lemma, this sum-rate is equal to $1/4$'th of the total throughput $T(n)$ w.h.p. The maximum achievable sum-rate between these S-D pairs is bounded above by the capacity of the MIMO channel between nodes S located to the left of

the cut and nodes D located to the right. Under the fast fading assumption, we have

$$\sum_{k \in S, i \in D} R_{ik} \leq \max_{\substack{Q(H) \geq 0 \\ \mathbb{E}(Q_{kk}(H)) \leq P,\, \forall k \in S}} \mathbb{E}\left(\log \det(I + HQ(H)H^*)\right) \tag{2}$$

where

$$H_{ik} = \frac{\sqrt{G}\, e^{j\,\theta_{ik}}}{r_{ik}^{\alpha/2}}, \quad k \in S, i \in D$$

$Q(\cdot)$ is a mapping from the set of possible channel realizations H to the set of positive semi-definite transmit covariance matrices. The diagonal element $Q_{kk}(H)$ corresponds to the power allocated to the kth node at channel state H. A natural way to upperbound (2) is by relaxing the individual power constraint to a total transmit power constraint. In the present context however, this is not convenient: some nodes in S are close to the cut and some are far apart, so the impact of these nodes on the system performance is quite different. A total transmit power constraint allows the transfer of power from the nodes far apart to those close to the cut, resulting in a loose bound. Instead, we will relax the individual power constraints to a total *weighted* power constraint, where the weight assigned to a node is set to be the total *received* power on the other side of the cut per watt of transmit power from that node. However, before doing that, we need to isolate the contribution of some nodes in D that are located very close to the cut. Typically, there are few nodes on both sides of the cut that are located at a distance as small as order $\frac{1}{\sqrt{n}}$ from the cut. If included, the contribution of these few pairs to the total received power would be excessive, resulting in a loose bound in the discussion below.

Let V_D denote the set of nodes located on the $1 \times \sqrt{n}$ rectangular area immediately to the right of the cut. Note that there are no more than $\sqrt{n}\log n$ nodes in V_D by Part (a) of Lemma 1. By generalized Hadamard's inequality, we have

$$\log \det(I + HQ(H)H^*)$$
$$\leq \log \det(I + H_1Q(H)H_1^*) + \log \det(I + H_2Q(H)H_2^*)$$

where H_1 and H_2 are obtained by partitioning the original matrix H: H_1 is the rectangular matrix with entries $H_{ik}, k \in S, i \in V_D$ and H_2 is the rectangular matrix with entries $H_{ik}, k \in S, i \in D \setminus V_D$. In turn, (2) is bounded above by

$$\sum_{k \in S, i \in D} R_{ik} \leq \max_{\substack{Q(H_1) \geq 0 \\ \mathbb{E}(Q_{kk}(H_1)) \leq P,\, \forall k \in S}} \mathbb{E}\left(\log \det(I + H_1Q(H_1)H_1^*)\right)$$
$$+ \max_{\substack{Q(H_2) \geq 0 \\ \mathbb{E}(Q_{kk}(H_2)) \leq P,\, \forall k \in S}} \mathbb{E}\left(\log \det(I + H_2Q(H_2)H_2^*)\right) \tag{3}$$

The first term in (3) can be easily upperbounded by applying Hadamard's inequality once more or equivalently by considering the sum of the capacities of the individual MISO channels between nodes in S and each node in V_D. Using moreover the fact that the minimum distance between any two nodes in the network is larger than $\frac{1}{n^{1/2+\delta}}$ w.h.p for any $\delta > 0$, we get the

following upper bound for the first term

$$\max_{\substack{Q(H_1)\geq 0 \\ \mathbb{E}(Q_{kk}(H_1))\leq P,\,\forall k\in S}} \mathbb{E}\left(\log\det(I+H_1Q(H_1)H_1^*)\right)$$
$$\leq \quad K\,\sqrt{n}(\log n)^2$$

where $K>0$ is a constant independent of n.

The second term in (3) is the capacity of the MIMO channel between nodes in S and nodes in $D\setminus V_D$. This is the term that dominates in (3) and thus its scaling determines the scaling of (2). The result is given by the following lemma, which completes the proof of Theorem 1.

Lemma 2: Let $P_{tot}(n)$ be the total power received by all the nodes in $D\setminus V_D$, when nodes in S transmit *independent* signals at full power. Then for every $\epsilon>0$,

$$\max_{\substack{Q(H_2)\geq 0 \\ \mathbb{E}(Q_{kk}(H_2))\leq P,\,\forall k\in S}} \mathbb{E}\left(\log\det(I+H_2Q(H_2)H_2^*)\right)\leq n^{\epsilon}P_{tot}(n).$$

Moreover, the scaling of the total received power is upper-bounded by

$$P_{tot}(n)\leq\begin{cases} K\,n\,(\log n)^3 & \alpha=2 \\ K\,n^{2-\alpha/2}\,(\log n)^2 & 2<\alpha<3 \\ K\,\sqrt{n}\,(\log n)^3 & \alpha=3 \\ K\,\sqrt{n}\,(\log n)^2 & \alpha>3 \end{cases}$$

w.h.p. for a constant $K>0$ independent of n. $\qquad\square$

Lemma 2 identifies the total received power under independent transmissions as the fundamental quantity limiting performance. Depending on α, there is a dichotomy on how this quantity scales with the system size. This dichotomy can be interpreted as follows.

The total received power is dominated either by the power transferred between nodes near the cut (order 1 distance) or by the power transferred between nodes far away from the cut. There are relatively fewer node *pairs* near the cut than away from the cut (order $\sqrt{n}$ versus order n^2), but the channels between the nodes near the cut are considerably stronger than between the nodes far away from the cut. When the attenuation parameter α is less than 3, the received power is dominated by transfer between nodes far away from the cut. The hierarchical cooperation scheme [2], which involves at the top level of the hierarchy MIMO transmissions between clusters of size $n^{1-\epsilon}$ at distance $\sqrt{n}$ apart, achieves arbitrarily closely the required power transfer and is therefore optimal in this regime. When $\alpha\geq 3$, the received power in the cutset bound is dominated by the power transfer by the nodes near the cut. This can be achieved by nearest neighbor multihop and multihop is therefore optimal in this regime.

Proof of Lemma 2: We are interested in the scaling of the second term in (3). Let us rescale each column k of the matrix H_2 by the (square root of the) *total received power* on the right from source node k on the left. Let indeed P_k denote the total received power in $D\setminus V_D$ of the signal sent by user $k\in S$:

$$P_k=P\,G\sum_{i\in D\setminus V_D}r_{ik}^{-\alpha}:=P\,G\,d_k.$$

The second term in (3) is then equal to

$$\max_{\substack{\tilde{Q}(\tilde{H})\geq 0 \\ \mathbb{E}(\tilde{Q}_{kk}(\tilde{H}))\leq P_k,\,\forall k\in S}} \mathbb{E}\left(\log\det(I+\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*)\right)$$

where

$$\tilde{H}_{ik}=\frac{e^{j\,\theta_{ik}}}{r_{ik}^{\alpha/2}}\frac{1}{\sqrt{d_k}},\qquad k\in S\text{ and }i\in D\setminus V_D.$$

The above expression is in turn bounded above by

$$\max_{\substack{\tilde{Q}(\tilde{H})\geq 0 \\ \mathbb{E}(\mathrm{Tr}\tilde{Q}(\tilde{H}))\leq P_{tot}(n)}} \mathbb{E}\left(\log\det(I+\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*)\right)$$

where $P_{tot}(n)=\sum_{k\in S}P_k=P\,G\sum_{k\in S,\,i\in D\setminus V_D}r_{ik}^{-\alpha}$.

Let us now define, for given $n\geq 1$ and $\varepsilon>0$, the set

$$B_{n,\varepsilon}=\{\|\tilde{H}\|^2>n^{\varepsilon}\}$$

where $\|A\|$ denotes the largest singular value of the matrix A. Note that the matrix $\tilde{H}$ is better conditioned than the original channel matrix H_2: all the diagonal elements of $\tilde{H}\tilde{H}^*$ are roughly of the same order (up to a factor $\log n$), and it can be shown that there exists $K_1>0$ such that

$$\mathbb{E}(\|\tilde{H}\|^2)\leq K_1\,(\log n)^3$$

for all n. Actually, the following stronger statement holds, whose proof is relegated to the Appendix.

Lemma 3: For any $\varepsilon>0$ and $p\geq 1$, there exists $K_1>0$ and independent of n such that for all n,

$$\mathbb{P}(B_{n,\varepsilon})\leq\frac{K_1}{n^p}$$

It follows that

$$\mathbb{E}\left(\log\det(I+\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*)\right)$$
$$\leq\quad \mathbb{E}\left(\log\det(I+\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*)\,1_{B_{n,\varepsilon}}\right)$$
$$+\,\mathbb{E}\left(\mathrm{Tr}(\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*)\,1_{B_{n,\varepsilon}^c}\right)\qquad(4)$$

The first term in (4) refers to the event that the channel matrix $\tilde{H}$ is accidentally ill-conditioned. Since the probability of such an event is polynomially small by Lemma 3, the contribution of this first term is actually negligible. In the second term in (4), the matrix $\tilde{H}$ is well conditioned, and this term is actually proportional to the maximum power transfer from left to right. Details follow below.

For the first term in (4), we use Hadamard's inequality and obtain

$$\mathbb{E}\left(\log\det(I+\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*)\,1_{B_{n,\varepsilon}}\right)$$
$$\leq\quad \mathbb{E}\left(\sum_{i\in D\setminus V_D}\log(1+\tilde{H}_i\tilde{Q}(\tilde{H})\tilde{H}_i^*)\,\bigg|\,B_{n,\varepsilon}\right)\mathbb{P}(B_{n,\varepsilon})$$

where $\tilde{H}_i$ is the ith row of $\tilde{H}$. By Jensen's inequality, this expression in turn is bounded above by

$$\sum_{i \in D \setminus V_D} \log \left(1 + \mathbb{E} \left(\|\tilde{H}_i\|^2 \operatorname{Tr}\tilde{Q}(\tilde{H}) \Big| B_{n,\varepsilon} \right) \right) \mathbb{P}(B_{n,\varepsilon})$$

$$\leq \sum_{i \in D \setminus V_D} \log \left(1 + \mathbb{E} \left(\|\tilde{H}_i\|^2 \operatorname{Tr}\tilde{Q}(\tilde{H}) \right) \Big/ \mathbb{P}(B_{n,\varepsilon}) \right) \mathbb{P}(B_{n,\varepsilon})$$

$$\leq n \log \left(1 + \frac{n \, P_{tot}(n)}{\mathbb{P}(B_{n,\varepsilon})} \right) \mathbb{P}(B_{n,\varepsilon}),$$

since $\|\tilde{H}_i\|^2 = \sum_{k \in S} r_{ik}^{-\alpha} / d_k \leq \sum_{k \in S} 1 \leq n$ and $\mathbb{E}(\operatorname{Tr}\tilde{Q}(\tilde{H})) \leq P_{tot}(n)$. The fact that the minimum distance between the nodes in S and $D \setminus V_D$ is at least 1 yields $P_{tot}(n) \leq PGn^2$. Noting that $x \mapsto x \log(1 + 1/x)$ is increasing on $[0,1]$ and using Lemma 3, we finally obtain that for any $p \geq 1$, there exists $K_1 > 0$ such that

$$\max_{\substack{\tilde{Q}(\tilde{H}) \geq 0 \\ \mathbb{E}(\operatorname{Tr}\tilde{Q}(\tilde{H})) \leq P_{tot}(n)}} \mathbb{E}\left(\log \det(I + \tilde{H}\tilde{Q}((\tilde{H}))\tilde{H}^*) 1_{B_{n,\varepsilon}} \right)$$

$$\leq K_1 n^{1-p} \log \left(1 + \frac{n^{3+p}}{K_1} \right)$$

which decays polynomially to zero with arbitrary exponent as n tends to infinity.

For the second term in (4), we simply have

$$\mathbb{E}\left(\operatorname{Tr}(\tilde{H}\tilde{Q}(\tilde{H})\tilde{H}^*) 1_{B_{n,\varepsilon}^c} \right) \leq \mathbb{E}\left(\|\tilde{H}\|^2 \operatorname{Tr}\tilde{Q}(\tilde{H}) 1_{B_{n,\varepsilon}^c} \right)$$
$$\leq n^\varepsilon P_{tot}(n).$$

Thus, the last thing that needs to be checked is the scaling of

$$P_{tot}(n) = PG \sum_{k \in S} d_k = PG \sum_{k \in S, i \in D \setminus V_D} r_{ik}^{-\alpha} \qquad (5)$$

stated in Lemma 2. Note that the node locations are random and hence $P_{tot}(n)$ is a random quantity. However, the regularity property in Part (a) of Lemma 1 allows to apply a binning argument and get a high probability upper bound on $P_{tot}(n)$. Let us divide the network area into n squarelets of area 1. By Part (a) of the lemma, there are no more than $\log n$ nodes in each squarelet w.h.p. Note also that each of the (positive) terms in the summation in (5) can only increase if we move the left and right hand-side nodes towards the cut in the middle. In fact, an argument based on moving the nodes inside each squarelet onto the squarelet vertex, allows to conclude that $P_{tot}(n)$ for the random network is w.h.p. less than the same quantity computed for a regular network with $\log n$ nodes at each left-hand side vertex and $2 \log n$ nodes at each right-hand side vertex.

The most convenient way to index the node positions in the resulting regular network is to use double indices. The left-hand side nodes are located at positions $(-k_x + 1, k_y)$ and those on the right at positions (i_x, i_y) where $k_x, k_y, i_x, i_y = 1, \ldots, \sqrt{n}$, so that

$$d_k = d_{k_x,k_y} = \sum_{i_x,i_y=1}^{\sqrt{n}} \frac{1}{((i_x + k_x - 1)^2 + (i_y - k_y)^2)^{\alpha/2}} \quad (6)$$

and the following upper bound on $P_{tot}(n)$ holds for the random network:

$$P_{tot}(n) \leq 2(\log n)^2 PG \sum_{k_x,k_y=1}^{\sqrt{n}} d_{k_x,k_y}. \qquad (7)$$

The following lemma establishes the scaling of d_{k_x,k_y}.

Lemma 4: There exist constants $K_2, K_3 > 0$ independent of k_x, k_y and n such that

$$d_{k_x,k_y} \leq \begin{cases} K_2 \log n & \text{if } \alpha = 2, \\ K_2 k_x^{2-\alpha} & \text{if } \alpha > 2, \end{cases}$$

and

$$d_{k_x,k_y} \geq K_3 k_x^{2-\alpha} \qquad \text{for} \qquad \alpha \geq 2.$$

Due to space constraints, we do not give the proof of the above lemma, but let us mention the following heuristic argument, that can be obtained through Laplace's principle. The summation in d_{k_x,k_y} scales the same as the maximum term in the sum times the number of terms which have roughly this maximum value. The maximum term is of the order of $1/k_x^\alpha$. The terms that take on roughly this value are those for which i_x runs from 1 to the order of k_x and i_y runs from k_y to k_y plus or minus the order of k_x. There are roughly k_x^2 such terms. Hence $d_{k_x,k_y} \approx k_x^2 \cdot 1/k_x^\alpha = k_x^{2-\alpha}$.

The upper bound given in the above lemma now yields:

$$\sum_{k_x,k_y=1}^{\sqrt{n}} d_{k_x,k_y} \leq \begin{cases} K_4 \, n \log n & \text{if } \alpha = 2, \\ K_4 \, n^{2-\alpha/2} & \text{if } 2 < \alpha \leq 3, \\ K_4 \, \sqrt{n} \log n & \text{if } \alpha = 3, \\ K_4 \, \sqrt{n} & \text{if } \alpha > 3 \end{cases}$$

for another constant $K_4 > 0$ independent of n. This upper bound combined with (7) completes the proof of Lemma 2.$\square$

IV. CONCLUSION

In this paper, we have identified the capacity scaling law of extended networks for *all* values of the power path loss exponent $\alpha \geq 2$. We have focused on the 2D setting, where nodes are located on a planar area, but our results generalize naturally to d-dimensional networks.

ACKNOWLEDGMENT

V. APPENDIX

Proof of Lemma 3: Let us start by considering the $2m^{th}$ moment of the spectral norm of $\tilde{H}$ given by (see [9, Ch. 5])

$$\|\tilde{H}\|^{2m} = \rho(\tilde{H}^* \tilde{H})^m = \lim_{l \to \infty} \{\operatorname{Tr}((\tilde{H}^* \tilde{H})^l)\}^{m/l}$$

By the dominated convergence theorem and Jensen's inequality, we have

$$\mathbb{E}(\|\tilde{H}\|^{2m}) \leq \lim_{l \to \infty} \{\mathbb{E}(\operatorname{Tr}((\tilde{H}^* \tilde{H})^l))\}^{m/l}$$

In the sequel, we are going to prove that the following upper bound holds with high probability:

$$\mathbb{E}(\mathrm{Tr}((\tilde{H}^*\tilde{H})^l)) \le t_l\, n\, (K_1 \log n)^{3l} \qquad (8)$$

where $t_l = \frac{(2l)!}{l!(l+1)!}$ are the Catalan numbers and $K_1 > 0$ is a constant independent of n. By Chebyshev's inequality, this allows to conclude that for any m,

$$
\begin{aligned}
\mathbb{P}(B_{n,\varepsilon}) &\le \frac{\mathbb{E}(\|\tilde{H}\|^{2m})}{n^{m\varepsilon}} \le \frac{1}{n^{m\varepsilon}} \lim_{l\to\infty}(t_l\, n\,(K_1 \log n)^{3l})^{m/l} \\
&\le \frac{\left(4(K_1 \log n)^3\right)^m}{n^{m\varepsilon}}
\end{aligned}
$$

since $\lim_{l\to\infty} t_l^{1/l} = 4$. For any $\varepsilon > 0$, choosing m sufficiently large shows therefore that $\mathbb{P}(B_{n,\varepsilon})$ decays polynomially with arbitrary exponent as $n \to \infty$, which is the result stated in Lemma 3.

There remains to prove the upperbound (8). Expanding the expression gives

$$
\begin{aligned}
&\mathbb{E}(\mathrm{Tr}((\tilde{H}^*\tilde{H})^l)) \\
&= \sum_{\substack{i_1,\dots,i_l \in D\setminus V_D \\ k_1,\dots,k_l \in S}} \mathbb{E}\left(\overline{\tilde{H}_{i_1 k_1}}\,\tilde{H}_{i_1 k_2}\,\overline{\tilde{H}_{i_2 k_2}}\,\tilde{H}_{i_2 k_3}\dots\overline{\tilde{H}_{i_l k_l}}\,\tilde{H}_{i_l k_1}\right) (9)
\end{aligned}
$$

Recall that the random variables $\tilde{H}_{ik}$ are independent and zero-mean, so the expectation is only non-zero when the terms in the product form conjugate pairs. From now on, we will focus on the $l = 2$ case, in order to give some intuition as to why estimate (8) holds; for further details and a complete proof, the reader is referred to [10]. We have

$$\mathbb{E}(\mathrm{Tr}((\tilde{H}^*\tilde{H})^2)) = \sum_{\substack{i_1,i_2 \in D\setminus V_D \\ k_1,k_2 \in S}} \mathbb{E}\left(\overline{\tilde{H}_{i_1 k_1}}\,\tilde{H}_{i_1 k_2}\,\overline{\tilde{H}_{i_2 k_2}}\,\tilde{H}_{i_2 k_1}\right) (10)$$

$$= \sum_{\substack{i_1,i_2 \in D\setminus V_D \\ k\in S}} |\tilde{H}_{i_1 k}|^2 |\tilde{H}_{i_2 k}|^2 + \sum_{\substack{i\in D\setminus V_D \\ k_1 \ne k_2 \in S}} |\tilde{H}_{i k_1}|^2 |\tilde{H}_{i k_2}|^2 (11)$$

$$\le \sum_{\substack{i_1,i_2 \in D\setminus V_D \\ k\in S}} |\tilde{H}_{i_1 k}|^2 |\tilde{H}_{i_2 k}|^2 + \sum_{\substack{i\in D\setminus V_D \\ k_1,k_2 \in S}} |\tilde{H}_{i k_1}|^2 |\tilde{H}_{i k_2}|^2 (12)$$

where (11) follows from the fact that the expectation in (10) is non-zero only when either $k_1 = k_2 = k$ or $i_1 = i_2 = i$. Note that we have removed the expectations since $|\tilde{H}_{ik}|^2$ is a deterministic quantity in our case. The upper bound (12) is obtained by doublecounting the terms with $i_1 = i_2 = i$ and $k_1 = k_2 = k$, i.e. the terms of the form $|\tilde{H}_{ik}|^4$.

We need now an upper bound on the scaling of $\sum_{i\in D\setminus V_D} |\tilde{H}_{ik}|^2$ and $\sum_{k\in S} |\tilde{H}_{ik}|^2$. Note that these are random quantities due to the random positions of the nodes. However, a binning argument using parts (a) and (b) of Lemma 1 provides again a connection to regular networks. Skipping the binning argument itself, let us directly concentrate on the regular case when the matrix elements of $\tilde{H}$ are given by

$$\tilde{H}_{ik} = \frac{e^{j\,\theta_{ik}}}{((i_x + k_x - 1)^2 + (i_y - k_y)^2)^{\alpha/4}}\,\frac{1}{\sqrt{d_{k_x,k_y}}}$$

and d_{k_x,k_y} is given in (6). By Lemma 4, we have $d_{k_x,k_y} \ge K_3\, k_x^{2-\alpha}$ for a constant $K_3 > 0$ independent of n. This, in turn, yields the upper bound

$$
\begin{aligned}
|\tilde{H}_{ik}|^2 &\le \frac{1}{K_3}\frac{k_x^{\alpha-2}}{((i_x + k_x - 1)^2 + (i_y - k_y)^2)^{\alpha/2}} \\
&\le \frac{1}{K_3}\frac{1}{(i_x + k_x - 1)^2 + (i_y - k_y)^2} \qquad (13)
\end{aligned}
$$

Summing over either i or k, and using the upper bound in Lemma 4 with $\alpha = 2$ yields

$$\sum_{i=1}^{n} |\tilde{H}_{ik}|^2,\ \sum_{k=1}^{n} |\tilde{H}_{ik}|^2 \le K_1 \log n$$

where $K_1 = \frac{K_2}{K_3}$ with K_2 and K_3 being the constants appearing in the lemma. For the random network, we have an extra factor $(\log n)^2$ arising from the binning argument:

$$\sum_{i\in D\setminus V_D} |\tilde{H}_{ik}|^2,\ \sum_{k\in S} |\tilde{H}_{ik}|^2 \le K_1 (\log n)^3.$$

Therefore, we obtain from (12) that

$$\mathbb{E}(\mathrm{Tr}((\tilde{H}^*\tilde{H})^2)) \le 2n(K_1 \log n)^6$$

which confirms estimate (8) in the case $l = 2$.

For $l \ge 2$, the non-vanishing terms in (9) can be associated to rooted planted planar trees that have l branches. This analogy gives t_l terms in the upper bound corresponding to (12), with each term corresponding to one such tree. It can be shown that each of these terms is upperbounded by $n(K_1 \log n)^{3l}$ by generalizing the approach for the case $l = 2$. This in turn yields the upper bound (8).

REFERENCES

[1] P. Gupta and P. R. Kumar, *The Capacity of Wireless Networks*, IEEE Trans. on Information Theory 42 (2), 2000, 388-404.

[2] A. Özgür, O. Lévêque, D. Tse, *How does the information capacity of ad hoc networks scale?*, Proc. of the 44th Annual Allerton Conf. on Communication, Control and Computing, Sept. 2006.

[3] L. -L. Xie and P. R. Kumar, *A Network Information Theory for Wireless Communications: Scaling Laws and Optimal Operation*, IEEE Trans. on Information Theory 50 (5), 2004, 748-767.

[4] A. Jovicic, P. Viswanath and S. R. Kulkarni, *Upper Bounds to Transport Capacity of Wireless Networks*, IEEE Trans. on Information Theory 50 (11), 2004, 2555-2565.

[5] O. Lévêque and E. Telatar, *Information Theoretic Upper Bounds on the Capacity of Large, Extended Ad-Hoc Wireless Networks*, IEEE Trans. on Information Theory 51 (3), March 2005.

[6] S. Ahmad, A. Jovicic and P. Viswanath, *Outer Bounds to the Capacity Region of Wireless Networks*, IEEE Trans. Inform. Theory, 52 (6), 2006, 2770-2776.

[7] L. -L. Xie, P. R. Kumar, *On the Path-Loss Attenuation Regime for Positive Cost and Linear Scaling of Transport Capacity in Wireless Networks*, IEEE Trans. on Information Theory, 52 (6), 2006, 2313-2328.

[8] A. Özgür, O. Lévêque, E. Preissmann, *Scaling laws for one and two-dimensional random wireless networks in the low attenuation regime*, submitted to IEEE Trans. on Information Theory, 2006.

[9] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[10] A. Özgür, O. Lévêque, D. Tse, *Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks*, submitted to IEEE Trans. on Information Theory, 2006.