Information theoretic upper bounds on the capacity of large extended ad hoc wireless networks*

Olivier Lévêque, Emre Telatar
Information Theory Laboratory
Swiss Federal Institute of Technology, Lausanne

To appear in the IEEE Transactions on Information Theory

Abstract

We derive an information theoretic upper bound on the rate per communication pair in a large ad hoc wireless network. We show that under minimal conditions on the attenuation due to the environment and for networks with a constant density of users, this rate tends to zero as the number of users gets large.

1 Introduction

The study of the feasability of large ad hoc wireless networks from an information theoretic point of view is a subject of both mathematical and practical interest. Various techniques, ranging from percolation theory to random matrix theory, can be used in order to tackle this problem. An important issue is the evaluation of the capacity of such networks. In the seminal work of P. Gupta and P. R. Kumar [3], it has been shown that under some assumptions, the transport capacity of such (planar) networks scales with $\sqrt{nA}$ where n is the number of users and A is the area occupied by the network. Even if the assumptions made in [3] are quite realistic regarding state of the art wireless communications, the question remains whether the result obtained in there, more precisely the upper bound, can be confirmed from an information theoretic point of view, that is, without any particular assumption on the way communication is established. A first confirmation of this result from an information theoretic point of view has been obtained in [6]. It was however assumed in there that signals are strongly attenuated over distance (power decay of order $\frac{1}{r^\alpha}$ with $\alpha > 6$).

The fact that the transport capacity scales with $\sqrt{nA}$ implies in particular that if there are order n pairs in the network willing to establish communication at a common rate R and if we assume that the pairs are chosen at random, without any consideration on the users' respective locations (so the average distance between paired users is of order $\sqrt{A}$), then the maximum achievable R decreases like $\frac{1}{\sqrt{n}}$ as n gets large. Our aim in the present paper is to give an information theoretic proof of the fact that in this particular scenario (and for a uniformly distributed network with a constant density of users), the maximum achievable R tends to zero under a minimal assumption on the attenuation function (power decay of order $\frac{1}{r^\alpha}$ with $\alpha > 2$).

*The work presented in this paper was supported in part by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

We would like to point out that our result does not say anything about the transport capacity of the network in general. Moreover, our upper bound is not as tight as the $\frac{1}{\sqrt{n}}$ behaviour found in [3, 6] (under stronger assumptions).

Our results apply to d-dimensional networks in general (see also [4] for an extension of the results of [3] to three dimensions). We also consider the model where an additional exponential factor is present in the attenuation function (as also considered in [6]). Section 2 is devoted to the study of uniformly distributed networks. In Section 3, we consider the particular situation of a "regular" network, where the users are placed on a grid.

2 Uniformly distributed networks

The network we consider consists of an even number n of users independently and uniformly distributed in the d-dimensional region

$$\Omega_n = [-n^{1/d}, n^{1/d}] \times [0, n^{1/d}]^{d-1}$$

of volume $|\Omega_n| = 2n$, therefore expanding with the number of users (when $d = 1$, Ω_n is the interval $[-n, n]$ and when $d = 2$, Ω_n is the rectangle $[-\sqrt{n}, \sqrt{n}] \times [0, \sqrt{n}]$) [1]. Note that because of this assumption of an "extended" network, the density of users remains constant as n increases.

Let us then divide these users into two arbitrary groups of $n/2$ users and assume that each user of the first group wishes to establish (one way) communication with a correspondent chosen at random in the second group (without any consideration on their respective locations) [2]. We assume that there is no fixed infrastructure that helps relaying communications, but we also assume no restriction on the kind of help the users can give to each other; in particular, any user may act as a relay for the communicating pairs. We moreover assume that in order to establish communication, each user has a device of power P which transmits with attenuation $g(r)$ over a distance r, where g is the function given by [3]

$$g(r) = \frac{e^{-\beta r/2}}{r^{\alpha/2}}, \quad r > 0. \tag{1}$$

Note that g describes the decay of the amplitude of the electric field and not that of the power. This model of decay is accepted as a standard one in wireless communications. The case $\alpha = 2$ and $\beta = 0$ describes the decrease of the electric field in the empty space. Because of cancelling reflections, the coefficient α is usually taken to be greater than 2, whereas a non-zero exponential factor β takes into account absorption in the air.

Let now R be the the maximum achievable rate per communication pair in the network. What we prove in the following is that R tends almost surely to zero as n gets large, under the assumption that either $\alpha > d \vee 2(d-2) = \max(d, 2(d-2))$ or $\beta > 0$. In the particular case $d = 2$, our result says more precisely that $R \le K\,\frac{n^{1/\alpha}\log n}{\sqrt{n}}$ when $\alpha > 2$ and $\beta = 0$, and that

[1]The fact that the region Ω_n is rectangular, and not square, is of little importance since we are only interested in the asymptotic behaviour of the network capacity.

[2]It could be raised here that this situation does not take place in a real network; however, the argument developed hereafter holds even if only a *constant fraction* of the users wish to establish communication without any consideration on their respective locations.

[3]One could raise again an objection here: without any constraint on the minimum distance between users, the above attenuation function may take arbitrarily large values. Because of our assumption of "extended" network however, points are likely to be sufficiently far apart form each other.

$R \leq K \frac{(\log n)^2}{\sqrt{n}}$ when $\beta > 0$ (see theorems 2.5 and 2.10). We therefore see that in the case $\beta = 0$, the bound obtained is quite distant from the $\frac{1}{\sqrt{n}}$ bound of [3], especially when α is small.

As a first step, we divide the domain Ω_n into two equal parts separated by the hyperplane $x^{(1)} = 0$, where $x^{(i)}$ denotes the i^{th} coordinate of $x \in \Omega_n$. Statistically, there are about $n/2$ users on the left-hand side of the domain; moreover, about half of these are transmitters and half of these transmitters wish to establish communication with a receiver on the right-hand side of the domain. In total, there are therefore about $n/8$ communications which need to cross the imaginary boundary from left to right, and it is easy to see that as n gets large, deviations from this idealized situation are of order much smaller than n with high probability.

In order to obtain an upper bound on R, we make a series of optimistic assumptions: we first assume that only the above $n/8+o(n)$ communications need to be established. We then introduce n additional "mirror" user that help relaying communications (where the mirror location of $x \in \Omega_n$ is $\tilde{x} = (-x^{(1)}, x^{(2)}, \ldots, x^{(n)})$). We see that there are now exactly n users on each side of the domain, which are moreover independently and uniformly distributed on each side. There is however a more important reason for introducing these "mirror" users: it brings a helpful symmetry in the problem, as we shall see below (remark 2.1). Besides, one has to keep in mind that our ultimate goal is to show an asymptotic result on the capacity in terms of the number of users. Therefore, doubling the number of users cannot harm by more than a factor 2.

Let us further assume that all the users on the left-hand side can share instantaneous information and even distribute their power resources among themselves in order to establish communication in the most efficient way with the users on the right-hand side, which in turn are able to distribute the received information instantaneously among themselves. We also assume that the user locations are known to all users. Following the argument of [2, Thm 14.10.1], we obtain the following upper bound on the sum of the rates of communications going from left to right:

$$\sum_{x \in S, y \in S^c} R_{xy} \leq I(X_S; Y_{S^c} | X_{S^c}),$$

where S denotes the set of users' locations on the left-hand side of Ω_n and S^c those on the right-hand side; $X_S = (X_x, x \in S)$ denotes the messages sent by the users in S, Y_{S^c} those received by the users in S^c and X_{S^c} those sent back by the users in S^c (which takes into account the effect induced by some eventual feedback).

In our setting, we have the following formal relation between X_S and Y_{S^c}:

$$Y_{S^c} = G\, X_S + Z, \tag{2}$$

where $G = (G_{xy}, x \in S, y \in S^c)$ is the matrix whose entries are given by $G_{xy} = g(|x-y|)$, with g given by (1), and Z is independent additive white Gaussian noise. From this relation, we deduce that

$$\begin{aligned}
I(X_S; Y_{S^c} | X_{S^c}) &= H(G\, X_S + Z \,| X_{S^c}) - H(G\, X_S + Z \,| X_S, X_{S^c}) \\
&= H(G\, X_S + Z \,| X_{S^c}) - H(Z) \\
&\leq H(G\, X_S + Z) - H(Z) \\
&= H(G\, X_S + Z) - H(G\, X_S + Z \,| X_S) = I(X_S; Y_{S^c}),
\end{aligned}$$

where we have used the fact that Z is independent from the other variables and that conditioning reduces entropy.

From now on, we will adopt the following notations (since we know that there are exactly n

users on each side):

$$S=\{x_1,\ldots,x_n\},\ S^c=\{y_1,\ldots,y_n\},\ X_S=(X_1,\ldots,X_n),\ Y_{S^c}=(Y_1,\ldots,Y_n),$$

where we assume that (x_i, y_i) forms a pair of "mirror" users for each $i \in \{1,\ldots,n\}$. With this notation, the channel model (2) becomes:

$$Y_j=\sum_{i=1}^n G_{ij}\,X_i+Z_j,\quad j=1,\ldots,n,$$

where $G_{ij}=g(|x_i-y_j|)$ and $Z=(Z_1,\ldots,Z_n)$ is a vector of independent circularly symmetric complex Gaussian random variables with unit variance. Under the power constraint

$$\sum_{i=1}^n \mathbb{E}(|X_i|^2)\le nP,$$

(arising from the fact that the users are assumed to be able to distribute their power resources among themselves), the mutual information $I(X_1,\ldots,X_n;Y_1,\ldots,Y_n)$ is maximum when $(X_1,\ldots,X_n)$ is a jointly Gaussian vector with some covariance matrix Q, so

$$\begin{aligned}\sum_{i,j=1}^n R_{ij} &\le \max_{p_X:\sum_{i=1}^n \mathbb{E}(|X_i|^2)\le nP} I(X_1,\ldots,X_n;Y_1,\ldots,Y_n)\\ &= \max_{Q\ge 0, Tr(Q)\le nP} \log\det(I+GQG^\dagger)\end{aligned}$$

By unitary transformation of the matrix G (see [9]), one finally obtains

$$\sum_{i,j=1}^n R_{ij}\le \max_{P_k\ge 0:\sum_{k=1}^n P_k\le nP}\ \sum_{k=1}^n \log(1+P_k\,\lambda_k^2)=:C_n, \tag{3}$$

where λ_k are the singular values of the matrix G (in decreasing order and repeated by multiplicity). Since about $n/8$ communications need to be established from left to right and since we wish to achieve the same common rate R on all these communications, we deduce from the above inequality that in order to prove that R tends to zero as n gets large, it is sufficient to prove that the capacity C_n defined in (3) grows *sublinearly* in n, that is, decreases to zero when divided by n.

Note finally that $y_i^{(1)}=-x_i^{(1)}$ and $y_i^{(k)}=x_i^{(k)}$ for $k\in\{2,\ldots,d\}$, so the matrix G is symmetric; it has therefore real eigenvalues μ_k and the singular values λ_k are equal to $|\mu_k|$.

2.1 No absorption case

In this section, we assume that $\beta=0$, that is, there is no absorption which creates an exponential decay of the power over the distance. It is shown in Appendix A that if

$$g(r)=\frac{1}{r^{\alpha/2}},\quad r>0,$$

and $\alpha>2(d-2)\vee 0$, then G is a non-negative (definite) matrix, so its eigenvalues μ_k are also non-negative (and equal to λ_k).

We can therefore obtain the following successive upper bounds on C_n. Noting first that $P_k \le nP$ for all k, we obtain

$$C_n \le \sum_{k=1}^n \log(1 + nP\mu_k^2).$$

Since $\mu_k \ge 0$, we further obtain

$$C_n \le \sum_{k=1}^n \log((1 + \sqrt{nP}\,\mu_k)^2) = 2\sum_{k=1}^n \log(1 + \sqrt{nP}\,\mu_k).$$

The computation of the asymptotic behaviour of the eigenvalues μ_k is not an easy task. We are therefore going to use the following majorization argument: from [7, Thm 9.B.1, p. 218], we know that the eigenvalues μ_k majorize the diagonal elements of G, that is,

$$\sum_{k=1}^l \mu_k \ge \sum_{k=1}^l G_{kk}, \quad \forall l \in \{1, \ldots, n-1\}, \tag{4}$$

and

$$\sum_{k=1}^n \mu_k = \sum_{k=1}^n G_{kk}. \tag{5}$$

On the other hand, by [7, Prop. 3.C.1, p. 64], we know that the function

$$(x_1, \ldots, x_n) \mapsto \sum_{k=1}^n \log(1 + \sqrt{nP}\,x_k)$$

is Schur-concave (recall [7, Def. 3.A.1, p. 54]: a Schur-concave funtion is a function $f : \mathbb{R}^n \to \mathbb{R}$ such that $f(x_1, \ldots, x_n) \le f(y_1, \ldots, y_n)$ as long as $(x_1, \ldots, x_n)$ majorizes $(y_1, \ldots, y_n)$ in the sense defined above). We therefore conclude that

$$C_n \le 2\sum_{k=1}^n \log(1 + \sqrt{nP}\,G_{kk}).$$

Moreover, $G_{kk} = g(|x_k - y_k|) = g\left(2|x_k^{(1)}|\right)$, so

$$C_n \le D_n^0 := 2\sum_{k=1}^n \log\left(1 + \sqrt{nP}\,g\left(2|x_k^{(1)}|\right)\right). \tag{6}$$

Remark 2.1. Note that the above majorization argument does not work in general if we replace the eigenvalues μ_k by the singular values λ_k: this is because the singular values of the matrix G satisfy (4) but not (5). This explains why we need G to be non-negative (i.e., the assumption of mirror users).

In order to obtain an upper bound on the average behaviour of D_n^0, we need the following technical lemma.

Lemma 2.2. *For any $C, p > 0$ and $\alpha > 1$, there exists a constant $K > 0$ such that for all sufficiently large n, we have*

$$\int_0^n dx\ \log\left(1 + \frac{Cn^p}{x^\alpha}\right) \le K\, n^{\frac{p}{\alpha} \wedge 1}\, \log n,$$

where $a \wedge b$ denotes the minimum value of a and b.

Proof. Let us define $x_0 = n^{\frac{p}{\alpha}\wedge 1}$ and compute

$$\begin{aligned}
&\int_0^n dx\ \log\left(1+\frac{Cn^p}{x^\alpha}\right)\\
&= \int_0^1 dx\ \log\left(1+\frac{Cn^p}{x^\alpha}\right)+\int_1^{x_0} dx\ \log\left(1+\frac{Cn^p}{x^\alpha}\right)+\int_{x_0}^n dx\ \log\left(1+\frac{Cn^p}{x^\alpha}\right)\\
&\le \int_0^1 dx\ \log\left(\frac{1+Cn^p}{x^\alpha}\right)+\int_1^{x_0} dx\ \log\left(1+Cn^p\right)+\int_{x_0}^n dx\ \left(\frac{Cn^p}{x^\alpha}\right)\\
&= \log\left(1+Cn^p\right)+\alpha\int_0^1 dx\ \log\left(\frac{1}{x}\right)+(x_0-1)\ \log\left(1+Cn^p\right)+\frac{Cn^p}{\alpha-1}\ \left(\frac{1}{x_0^{\alpha-1}}-\frac{1}{n^{\alpha-1}}\right).
\end{aligned}$$

Replacing x_0 by its value and checking separately the two cases $\frac{p}{\alpha}<1$ and $\frac{p}{\alpha}\ge 1$ leads then to the conclusion. □

This lemma allows us to deduce the following.

Proposition 2.3. *There exists a constant $K>0$ (possibly depending on α) such that for all sufficiently large n,*

$$\mathbb{E}(D_n^0)\le K n^{\frac{d-1}{d}+(\frac{1}{\alpha}\wedge\frac{1}{d})}\ \log n,$$

so $\mathbb{E}(D_n^0)$ is sublinear in n if $\alpha>d$.

Proof. Set $m=n^{\frac{1}{d}}$ (which needs not be an integer). First note that

$$\mathbb{E}(D_n^0)=n\ \mathbb{E}\left(\log\left(1+\sqrt{Pn}g\left(2|x_1^{(1)}|\right)\right)\right)=m^{d-1}\int_0^m dx\ \log\left(1+\frac{\sqrt{P}\,m^{d/2}}{(2x)^{\alpha/2}}\right).$$

So we obtain by Lemma 2.2 that for sufficiently large n,

$$\mathbb{E}(D_n^0)\le K\ m^{d-1}\ m^{(\frac{d}{\alpha}\wedge 1)}\ \log m=\frac{K}{d}\ n^{\frac{d-1}{d}+(\frac{1}{\alpha}\wedge\frac{1}{d})}\ \log n,$$

and this concludes the proof. □

There remains to prove that the sublinear behaviour of D_n^0 in n takes place almost surely. We prove this by showing that the deviation of D_n^0 from its average is indeed almost surely sublinear in n.

Proposition 2.4. *Fix $\alpha>0$. Then for any $\varepsilon>0$, we have*

$$\lim_{n\to\infty}\frac{|D_n^0-\mathbb{E}(D_n^0)|}{n^{\frac{1}{2}+\varepsilon}}=0,\quad \textit{almost surely.}$$

Proof. What we are going to use here is Hoeffding's inequality (see [5]). We first note that D_n^0 is the sum of n independent random variables

$$d_k=\log\left(1+\sqrt{nP}\,g\left(2|x_k^{(1)}|\right)\right).$$

However, each of these random variables is unbounded, since $|x_k^{(1)}|$ can be arbitrarily close to zero. We are going to show that with a certain scaling factor, they are all bounded away from zero with high probability as n goes to infinity, and that under the condition that they are effectively bounded away from zero, D_n^0 concentrates around its mean with a deviation of order less than $n^{\frac{1}{2}+\varepsilon}$ for any $\varepsilon > 0$. Let us then fix $\eta > 0$ and compute the probability that any of the $|x_k^{(1)}|$ is smaller than $n^{-\frac{d-1}{d}-\eta}$. Denoting by $x_{\min}$ the vector x_k whose first component is minimal (in absolute value), we obtain by the union bound that

$$\mathbb{P}\left(|x_{\min}^{(1)}| < n^{-\frac{d-1}{d}-\eta}\right) \leq n\,\mathbb{P}\left(|x_1^{(1)}| < n^{-\frac{d-1}{d}-\eta}\right) = n\,\frac{n^{-\frac{d-1}{d}-\eta}}{n^{\frac{1}{d}}} = n^{-\eta},$$

and this probability is arbitrarily small for any $\eta > 0$. On the other hand, under the condition that $|x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}$, the $|x_k^{(1)}|$ remain i.i.d. random variables, as the following calculation shows. Let $t_1, \ldots, t_n \in [n^{-\frac{d-1}{d}-\eta}, n]$; we then have

$$\begin{aligned}
&\mathbb{P}\left(|x_1^{(1)}| \geq t_1, \ldots, |x_n^{(1)}| \geq t_n \;\middle|\; |x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}\right) \\
&= \frac{\mathbb{P}(|x_1^{(1)}| \geq t_1, \ldots, |x_n^{(1)}| \geq t_n)}{\mathbb{P}(|x_1^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}, \ldots, |x_n^{(1)}| \geq n^{-\frac{d-1}{d}-\eta})} = \prod_{k=1}^n \frac{\mathbb{P}(|x_k^{(1)}| \geq t_k)}{\mathbb{P}(|x_k^{(1)}| \geq n^{-\frac{d-1}{d}-\eta})} \\
&= \prod_{k=1}^n \mathbb{P}\left(|x_k^{(1)}| \geq t_k \;\middle|\; |x_k^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}\right) = \prod_{k=1}^n \mathbb{P}\left(|x_k^{(1)}| \geq t_k \;\middle|\; |x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}\right),
\end{aligned}$$

since the $x_k^{(1)}$ are independent. Moreover, under the condition that $|x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}$, the random variables d_k are bounded:

$$|d_k| \leq \log\left(1 + \sqrt{nP}\,g\left(2n^{-\frac{d-1}{d}-\eta}\right)\right) \leq \log\left(1 + \sqrt{P}\,n^{\frac{1}{2}+\frac{\alpha}{2}(\frac{d-1}{d}+\eta)}\right) \leq C_1(\eta) + C_2(\eta)\log n.$$

We can therefore apply Hoeffding's inequality [5] which states that for all $\lambda > 0$ and sufficiently large n,

$$\mathbb{P}\left(|D_n^0 - \mathbb{E}(D_n^0)| \geq \lambda\sqrt{n} \;\middle|\; |x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}\right) \leq 2\,\exp\left(-\frac{2\,\lambda^2}{(C_1(\eta) + C_2(\eta)\log n)^2}\right),$$

and replacing λ by λn^{ε} in the preceding inequality gives

$$\mathbb{P}\left(|D_n^0 - \mathbb{E}(D_n^0)| \geq \lambda\,n^{\frac{1}{2}+\varepsilon} \;\middle|\; |x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}\right) \leq 2\,\exp\left(-\frac{2\,\lambda^2\,n^{2\varepsilon}}{(C_1(\eta) + C_2(\eta)\log n)^2}\right).$$

In conclusion, we have

$$\begin{aligned}
&\mathbb{P}\left(|D_n^0 - \mathbb{E}(D_n^0)| \geq \lambda\,n^{\frac{1}{2}+\varepsilon}\right) \\
&\leq \mathbb{P}\left(|x_{\min}^{(1)}| < n^{-\frac{d-1}{d}-\eta}\right) + \mathbb{P}\left(|D_n^0 - \mathbb{E}(D_n^0)| \geq \lambda\,n^{\frac{1}{2}+\varepsilon} \;\middle|\; |x_{\min}^{(1)}| \geq n^{-\frac{d-1}{d}-\eta}\right) \\
&\leq n^{-\eta} + 2\,\exp\left(-\frac{2\,\lambda^2\,n^{2\varepsilon}}{(C_1(\eta) + C_2(\eta)\log n)^2}\right).
\end{aligned}$$

Choosing $\eta > 1$, we therefore obtain that

$$\sum_{n\geq 1} \mathbb{P}\left(|D_n^0 - \mathbb{E}(D_n^0)| \geq \lambda\,n^{\frac{1}{2}+\varepsilon}\right) < \infty,$$

so by the Borel-Cantelli lemma, we have for any $\varepsilon > 0$ and $\lambda > 0$,

$$\mathbb{P}\left(\frac{|D_n^0 - \mathbb{E}(D_n^0)|}{n^{\frac{1}{2}+\varepsilon}} \geq \lambda \quad \text{infinitely often}\right) = 0,$$

which implies the result. $\square$

Let us summarize the results obtained in the following theorem.

Theorem 2.5. *If $\alpha > d \vee 2(d-2)$ (that is, $\alpha > d$ when $d \leq 4$) and there is no absorption (that is, $\beta = 0$), then the maximum achievable rate R per communication pair in a large uniformly distributed network decreases almost surely to zero as the number of users gets large. More precisely, under the above assumption, there exist a constant $K > 0$ such that for sufficiently large n,*

$$R \leq K \frac{\log n}{n^{\frac{1}{d}-\frac{1}{\alpha}}} \quad a.s.$$

Note that the last estimate comes from the fact that $R \leq \frac{D_n^0}{n/8}$ and propositions 2.3 and 2.4.

Remark 2.6. As a by-product, the above analysis also gives an upper bound on the maximal amount of information that can be carried from one part of the network to the other. This amount is bounded above by C_n, which in turn is bounded above by $K\, n^{(d-1)/d} \log n$, almost surely for sufficiently large n.

2.2 Case with absorption

In the following, we assume that $\beta > 0$. Starting back from expression (3), we follow the lines of the preceding section: using the fact that $P_k \leq nP$, we have

$$C_n \leq \sum_{k=1}^{n} \log(1 + nP\lambda_k^2).$$

Now, since λ_k^2 are the eigenvalues of the matrix $GG^\dagger$, repeating the majorization argument of Section 2.1 gives

$$C_n \leq \sum_{k=1}^{n} \log(1 + nP(GG^\dagger)_{kk})$$

Moreover, since g is decreasing, we have

$$(GG^\dagger)_{kk} = \sum_{l=1}^{n} |G_{kl}|^2 = \sum_{l=1}^{n} g(|x_k - y_l|)^2 \leq n\, g(|x_k^{(1)}|)^2,$$

So we finally obtain:

$$C_n \leq D_n := \sum_{k=1}^{n} \log\left(1 + Pn^2 g(|x_k^{(1)}|)^2\right). \tag{7}$$

We need now the following technical lemma, similar to Lemma 2.2.

Lemma 2.7. *For any $C, p, \alpha, \beta > 0$, there exists a constant $K > 0$ such that for all sufficiently large n, we have*

$$\int_0^n dx\ \log\left(1 + \frac{Cn^p e^{-\beta x}}{x^\alpha}\right) \leq K\ (\log n)^2.$$

Proof. The proof follows the lines of that of Lemma 2.2; set $x_0 = \frac{p}{\beta}\log n$ (which is smaller than

n for sufficiently large n) and compute

$$\begin{aligned}
&\int_0^n dx\ \log\left(1+\frac{Cn^p e^{-\beta x}}{x^\alpha}\right)\\
&= \int_0^1 dx\ \log\left(1+\frac{Cn^p e^{-\beta x}}{x^\alpha}\right) + \int_1^{x_0} dx\ \log\left(1+\frac{Cn^p e^{-\beta x}}{x^\alpha}\right) + \int_{x_0}^n dx\ \log\left(1+\frac{Cn^p e^{-\beta x}}{x^\alpha}\right)\\
&\le \int_0^1 dx\ \log\left(\frac{1+Cn^p}{x^\alpha}\right) + \int_1^{x_0} dx\ \log\left(1+Cn^p\right) + \int_{x_0}^n dx\ Cn^p\ e^{-\beta x}\\
&= \log\left(1+Cn^p\right) + \alpha\int_0^1 dx\ \log\left(\frac{1}{x}\right) + (x_0-1)\ \log\left(1+Cn^p\right) + \frac{Cn^p}{\beta}\ \left(e^{-\beta x_0}-e^{-\beta n}\right).
\end{aligned}$$

Replacing x_0 by its value proves then the lemma. □

From this, we deduce the following upper bound on the average behaviour of D_n.

Proposition 2.8. *There exists a constant $K > 0$ (possibly depending on α or β) such that for all sufficiently large n,*

$$\mathbb{E}(D_n) \le Kn^{\frac{d-1}{d}}\ (\log n)^2,$$

so $\mathbb{E}(D_n)$ is sublinear in n.

Proof. Set $m = n^{\frac{1}{d}}$. First note that

$$\mathbb{E}(D_n) = n\ \mathbb{E}\left(\log\left(1+Pn^2 g(|x_1^{(1)}|)^2\right)\right) = m^{d-1}\int_0^m dx\ \log\left(1+\frac{Pm^{2d}e^{-\beta x}}{x^\alpha}\right).$$

So we obtain by Lemma 2.7 that for sufficiently large n,

$$\mathbb{E}(D_n) \le K\ m^{d-1}\ (\log m)^2 = \frac{K}{d^2}\ n^{\frac{d-1}{d}}\ (\log n)^2,$$

and this concludes the proof. □

As before, there remains to prove that the sublinear behaviour of D_n in n takes place almost surely. We prove this using the following concentration result.

Proposition 2.9. *Fix $\alpha > 0$, $\beta > 0$. Then for any $\varepsilon > 0$, we have*

$$\lim_{n\to\infty} \frac{|D_n - \mathbb{E}(D_n)|}{n^{\frac{1}{2}+\varepsilon}} = 0, \quad \textit{almost surely.}$$

Proof. The proof is identical to that of Proposition 2.4, so we do not reproduce it here. □

Let us summarize the result obtained in the following theorem.

Theorem 2.10. *If there is absorption (that is, $\beta > 0$), then the maximum achievable rate R per communication pair in a large uniformly distributed network decreases almost surely to zero as the number of users gets large. More precisely, under the above assumption, there exist a constant $K > 0$ such that for sufficiently large n,*

$$R \le K\frac{(\log n)^2}{n^{\frac{1}{d}}} \quad \textit{a.s.}$$

Note that the last estimate comes from the fact that $R \le \frac{D_n}{n/8}$ and propositions 2.8 and 2.9.

3 Regular networks

Let us now consider the case where the network is a regular network in the sense that the users are placed on a regular grid inside Ω_n. For simplicity, we will assume that there are n users on each side and that $n = m^d$ for some integer m. The positions of the users on the left and the right-hand side of the region Ω_n are therefore given by

$$x_k = (-k_1 + 0.5, k_2 - 0.5, \ldots, k_d - 0.5) \quad \text{and} \quad y_l = (l_1 - 0.5, l_2 - 0.5, \ldots, l_d - 0.5),$$

where $k = (k_1, \ldots, k_d)$ and $l = (l_1, \ldots, l_d)$ denote from now on multi-indices ranging from $(1, \ldots, 1)$ to $(m, \ldots, m)$. For notational simplicity, we will go on writing $k = 1, \ldots, n$ for an enumeration of all the multi-indices.

We consider that these $2n$ users wish to form n communication pairs, choosing their correspondent at random. An argument similar to that developed in the previous section shows that there will be about $n/4$ communications needing to cross the imaginary boundary $x^{(1)} = 0$ from left to right. Repeating then the argument of the previous section leads to the following upper bound on the maximum achievable rate per communication pair in the ntework:

$$R \le \frac{C_n}{n/4 + o(n)}, \quad \text{where } C_n := \max_{P_k \ge 0: \sum_{k=1}^n P_k \le nP} \sum_{k=1}^n \log(1 + P_k \lambda_k^2),$$

Here, λ_k are the singular values of the matrix G, whose entries $G_{ij} = g(|x_i - y_j|)$ are determinisitic in the present context.

3.1 No absorption case

Let us assume that $\beta = 0$. We will show in the following a better result than that of Section 2.1, in the sense that we do not need anymore the assumption that $\alpha > 2(d-2)$, but this requires to be a little more careful in the majoration procedure. Let us first note that for any fixed vector $(P_1, \ldots, P_n)$, we have

$$\sum_{k=1}^n \log(1 + P_k \lambda_k^2) \le \sum_{k=1}^n \log(1 + P_k (GG^\dagger)_{kk}),$$

by the same majorization argument as that of Section 2.2. Moreover, since g is decreasing, we have

$$(GG^\dagger)_{kk} = \sum_{l=1}^n g(|x_k - y_l|)^2 \le \sum_{l=1}^n g(|x_k^{(1)} - y_l^{(1)}|)^2 = m^{d-1} \sum_{l_1=1}^m g(k_1 + l_1 - 1)^2$$

(recall that $x_k^{(1)} = -k_1 + 0.5$, $y_l^{(1)} = l_1 - 0.5$ and $n = m^d$ where m is an integer). Let us compute

$$\begin{aligned} \sum_{l_1=1}^m g(k_1 + l_1 - 1)^2 &= \sum_{l_1=1}^m \frac{1}{(k_1 + l_1 - 1)^\alpha} \le \frac{1}{k_1^\alpha} + \int_1^m dy \, \frac{1}{(k_1 + y - 1)^\alpha} \\ &\le \frac{1}{k_1^\alpha} + \frac{1}{\alpha - 1} \, \frac{1}{k_1^{\alpha-1}} \le \frac{\alpha}{\alpha - 1} \, \frac{1}{k_1^{\alpha-1}}. \end{aligned}$$

This leads to the following upper bound:

$$C_n \le \max \left\{ \sum_{k=1}^n \log\left(1 + \frac{\alpha}{\alpha - 1} \, \frac{P_k m^{d-1}}{k_1^{\alpha-1}}\right) \,\middle|\, P_k \ge 0, \ \sum_{k=1}^n P_k \le nP \right\},$$

which can be rewritten as

$$C_n \leq \max\left\{ \sum_{k=1}^{n} \log\left(1 + \frac{\tilde{P}_k m^{d-1}}{k_1^{\alpha-1}}\right) \,\middle|\, \tilde{P}_k \geq 0,\ \sum_{k=1}^{n} \tilde{P}_k \leq \frac{\alpha}{\alpha-1}\, nP \right\}.$$

This maximization problem has the well known "water-filling" solution:

$$C_n = \sum_{k=1}^{n} \log\left(\frac{\nu m^{d-1}}{k_1^{\alpha-1}}\right)^+,$$

where ν satisfies the constraint

$$\sum_{k=1}^{n} \left(\nu - \frac{k_1^{\alpha-1}}{m^{d-1}}\right)^+ = \frac{\alpha}{\alpha-1}\, nP, \tag{8}$$

and a^+ denotes the positive part of $a \in \mathbb{R}$. We need now the following two technical lemmas.

Lemma 3.1. *Let $\alpha > d$ and ν satisfy (8). There exists then a constant $K > 0$ (possibly depending on α) such that for all sufficiently large n, we have*

$$\nu \leq K n^{\frac{1}{d}(1-\frac{d}{\alpha})}.$$

Proof. Equation (8) implies that

$$\begin{aligned}
\frac{\alpha}{\alpha-1}\, mP &= \sum_{k_1=1}^{m} \left(\nu - \frac{k_1^{\alpha-1}}{m^{d-1}}\right)^+ \\
&\geq \int_1^m dx\, \left(\nu - \frac{x^{\alpha-1}}{m^{d-1}}\right)^+ \\
&= \int_1^{x_0} dx\, \left(\nu - \frac{x^{\alpha-1}}{m^{d-1}}\right),
\end{aligned}$$

where $x_0 = (\nu m^{d-1})^{\frac{1}{\alpha-1}}$. Computing this last expression gives

$$\begin{aligned}
\frac{\alpha}{\alpha-1}\, mP &\geq \nu(x_0 - 1) - \frac{x_0^\alpha - 1}{\alpha m^{d-1}} \\
&\geq \nu^{\frac{\alpha}{\alpha-1}}\, m^{\frac{d-1}{\alpha-1}}\, \frac{\alpha-1}{\alpha} - \nu := F(\nu).
\end{aligned}$$

Since F is increasing on the domain where it is positive, we then obtain that $\nu \leq \nu_0$, where ν_0 satisfies the equation $F(\nu_0) = \frac{\alpha}{\alpha-1} mP$. This equation in turn implies that

$$\nu_0^{\frac{\alpha}{\alpha-1}}\, m^{\frac{d-1}{\alpha-1}}\, \frac{\alpha-1}{\alpha} \geq \frac{\alpha}{\alpha-1}\, mP,$$

so

$$\nu_0 \geq \left(P\left(\frac{\alpha}{\alpha-1}\right)^2\right)^{\frac{\alpha-1}{\alpha}} m^{(1-\frac{d-1}{\alpha-1})\frac{\alpha-1}{\alpha}} := K_\alpha\, m^{1-\frac{d}{\alpha}}$$

and

$$\nu_0^{-\frac{1}{\alpha-1}}\, m^{-\frac{d-1}{\alpha-1}} \leq K_\alpha^{-\frac{1}{\alpha-1}}\, m^{-\frac{1}{\alpha-1}(1-\frac{d}{\alpha})-\frac{d-1}{\alpha-1}} = K_\alpha^{-\frac{1}{\alpha-1}}\, m^{-\frac{d}{\alpha}} \underset{m\to\infty}{\longrightarrow} 0,$$

since $\alpha > d$. Now, since

$$F(\nu_0) = \nu_0^{\frac{\alpha}{\alpha-1}} m^{\frac{d-1}{\alpha-1}} \left(\frac{\alpha - 1}{\alpha} - \nu_0^{-\frac{1}{\alpha-1}} m^{-\frac{d-1}{\alpha-1}} \right),$$

we obtain that for sufficently large m, there exists $K > 0$ (depending on α) such that

$$\frac{\alpha}{\alpha - 1}\, mP = F(\nu_0) \geq K\, \nu_0^{\frac{\alpha}{\alpha-1}}\, m^{\frac{d-1}{\alpha-1}}.$$

This implies finally that

$$\nu \leq \nu_0 \leq K\, m^{\left(1-\frac{d-1}{\alpha-1}\right)\frac{\alpha-1}{\alpha}} = K\, m^{1-\frac{d}{\alpha}},$$

which concludes the proof. $\square$

Lemma 3.2. *For any $\eta > 0$ and $k_0 \geq 1$, we have*

$$\sum_{k_1=1}^{k_0} \log \left(\frac{k_0^\eta}{k_1^\eta} \right) \leq \eta\, k_0.$$

Proof. Let us simply compute

$$\begin{aligned} \sum_{k_1=1}^{k_0} \log \left(\frac{k_0^\eta}{k_1^\eta} \right) &= \eta \left(k_0\, \log k_0 - \sum_{k_1=1}^{k_0} \log k_1 \right) \\ &\leq \eta \left(k_0\, \log k_0 - \int_1^{k_0} dx\, \log x \right) \\ &= \eta\, (k_0\, \log k_0 - k_0\, \log k_0 + k_0 - 1) \leq \eta\, k_0. \end{aligned}$$

$\square$

This allows us to establish the following proposition.

Proposition 3.3. *Let $\alpha > d$. There exists then a constant $K > 0$ (possibly depending on α) such that for all sufficiently large n, we have*

$$C_n \leq K n^{\frac{d-1}{d}+\frac{1}{\alpha}}.$$

So C_n is sublinear in n when $\alpha > d$.

Proof. From the above analysis, we have the following upper bound on C_n:

$$C_n \leq m^{d-1} \sum_{k_1=1}^{m} \log \left(\frac{\nu m^{d-1}}{k_1^{\alpha-1}} \right)^+,$$

where ν satisfies the constraint (8). Let k_0 denote the smallest integer such that

$$(\nu m^{d-1})^{\frac{1}{\alpha-1}} \leq k_0.$$

We then obtain:

$$C_n \leq m^{d-1} \sum_{k=1}^{k_0} \log \left(\frac{k_0^{\alpha-1}}{k_1^{\alpha-1}} \right) \leq m^{d-1}\, \alpha\, k_0,$$

by Lemma 3.2 with $\eta = \alpha - 1$. On the other hand, by Lemma 3.1, there exists $K > 0$ such that

$$k_0 \leq K\, m^{\left(1-\frac{d}{\alpha}+d-1\right)\frac{1}{\alpha-1}} + 1 = K\, m^{\frac{d}{\alpha}} + 1,$$

for sufficiently large m, which in turn implies

$$C_n \leq K\, m^{d-1+\frac{d}{\alpha}},$$

and this completes the proof. □

This proposition leads directly to the following theorem.

Theorem 3.4. *If $\alpha > d$ and there is no absorption (that is, $\beta = 0$), then the maximum achievable rate R per communication pair in a large regular network decreases almost surely to zero as the number of users gets large.*

We obtain here the same result as the one obtained for uniformly distributed networks, without the assumption of non-negativity of the matrix G, that is, without the assumption that $\alpha > 2(d-2)$.

3.2 Case with absorption

In the case where there is absorption (that is, $\beta > 0$), we follow the lines of Section 2.2 and obtain easily:

$$\begin{aligned}
C_n &\leq \sum_{k=1}^{n} \log(1 + nP\lambda_k^2) \leq \sum_{k=1}^{n} \log(1 + nP(GG^\dagger)_{kk}) \\
&\leq m^{d-1} \sum_{k_1=1}^{m} \log\left(1 + Pm^{2d} g(k_1)^2\right) \\
&\leq m^{d-1} \sum_{k_1=1}^{m} \log\left(1 + Pm^{2d} e^{-\beta k_1}\right) \\
&\leq m^{d-1} \int_0^m dx\, \log\left(1 + Pm^{2d} e^{-\beta x}\right) \\
&\leq \frac{K}{d^2}\, n^{\frac{d-1}{d}}\, (\log n)^2,
\end{aligned}$$

by Lemma 2.7, so this proves the following theorem.

Theorem 3.5. *If there is absorption (that is, $\beta > 0$), then the maximum achievable rate R per communication pair in a large regular network decreases almost surely to zero as the number of users gets large.*

4 Conclusion and perspectives

We have proved that under minimal assumptions (that is, with a power decay of order $1/r^\alpha$ with $\alpha > d \vee 2(d-2)$ or in the presence of absorption), the maximum achievable rate per communication pair in a large extended ad hoc network has to decrease to zero as the number of users gets large. However, we have seen that our scaling law is not as tight as the one obtained in [3, 6]. In order to get a better result, a precise study of the behaviour of the singular values λ_k is necessary. This is work under progress.

5 Acknowledgements

We would like to thank Shashibhushan Borade, Michael Gastpar, Piyush Gupta, Shan-Yuan Ho, P. R. Kumar, David Tse, Rüdiger Urbanke and Feng Xue for many stimulating discussions.

A Non-negativity of the matrix G in the no absorption case

Let us first consider the one-dimensional case (with $\alpha > 0$). In this case, since $y_j = -x_j$, the entries of the matrix G are given by

$$\begin{aligned} G_{ij} &= g(|x_i - y_j|) = g(x_i + x_j) = \frac{1}{(x_i + x_j)^{\alpha/2}} \\ &= \int_0^\infty dt \, \frac{t^{\alpha/2-1}}{\Gamma(\alpha/2)} \, e^{-(x_i+x_j)t}, \end{aligned}$$

where Γ is the Euler Gamma function. This implies that G is non-negative, since

$$\sum_{i,j=1}^n G_{ij} \, c_i \, c_j = \int_0^\infty dt \, \frac{t^{\alpha/2-1}}{\Gamma(\alpha/2)} \left(\sum_{i=1}^n e^{-x_i t} \, c_i \right)^2 \geq 0.$$

Let us now consider the higher-dimensional case together with the assumption that $\alpha > 2(d-2)$. We have the following expression for the entries of the matrix G:

$$G_{ij} = \frac{1}{\left((x_i^{(1)} + x_j^{(1)})^2 + |x_i^{(2)} - x_j^{(2)}|^2 + \ldots + |x_i^{(d)} - x_j^{(d)}|^2 \right)^{\alpha/4}},$$

so using the fact that the Fourier transform of

$$f_a(x_2, \ldots, x_d) = \frac{1}{(a^2 + |x_2|^2 + \ldots + |x_d|^2)^{\alpha/4}}$$

is given by (see [8, formulas I.2.7 and I.18.29])

$$\hat{f}_a(\xi_2, \ldots, \xi_d) = C_{\alpha,d} \left(\frac{\sqrt{|\xi_2|^2 + \ldots + |\xi_d|^2}}{a} \right)^{\frac{\alpha}{4} - \frac{d-1}{2}} \cdot K_{\frac{\alpha}{4} - \frac{d-1}{2}} (a \sqrt{|\xi_2|^2 + \ldots + |\xi_d|^2}),$$

where K_ν is the modified Bessel function of second kind and of order ν, we obtain that

$$\begin{aligned} G_{ij} &= \frac{C_{\alpha,d}}{(2\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} d\xi_2 \cdots d\xi_d \left(\frac{\sqrt{|\xi_2|^2 + \ldots + |\xi_d|^2}}{x_i^{(1)} + x_j^{(1)}} \right)^{\frac{\alpha}{4} - \frac{d-1}{2}} \cdot \\ & \quad \cdot K_{\frac{\alpha}{4} - \frac{d-1}{2}} \left((x_i^{(1)} + x_j^{(1)}) \sqrt{|\xi_2|^2 + \ldots + |\xi_d|^2} \right) e^{i\xi_2(x_i^{(2)} - x_j^{(2)}) + \cdots + i\xi_d(x_i^{(d)} - x_j^{(d)})}. \end{aligned}$$

Since by formula 9.6.23 in [1], we have

$$\frac{1}{r^\nu} K_\nu(r) = \frac{\sqrt{\pi}}{2^\nu \, \Gamma(\nu + \frac{1}{2})} \int_0^\infty dt \, e^{-r \cosh(t)} \sinh^{2\nu}(t),$$

for $\nu > -\frac{1}{2}$, we obtain that the matrix whose entries are given by

$$\left(\frac{\sqrt{|\xi_2|^2 + \ldots + |\xi_d|^2}}{x_i^{(1)} + x_j^{(1)}} \right)^{\frac{\alpha}{4} - \frac{d-1}{2}} \cdot K_{\frac{\alpha}{4} - \frac{d-1}{2}} \left((x_i^{(1)} + x_j^{(1)}) \sqrt{|\xi_2|^2 + \ldots + |\xi_d|^2} \right)$$

is non-negative if

$$\frac{\alpha}{4} - \frac{d-1}{2} > -\frac{1}{2}, \quad \text{that is,} \quad \alpha > 2(d-2).$$

So under the same assumption, G is a convex combination of products of non-negative matrices, it is therefore itself non-negative.

Remark. Numerical simulations indicate that G is non-negative if and only if $\alpha \geq 2(d-2) \vee 0$.

References

[1] Abramowitz M., Stegun I. A., "Handbook of mathematical functions", 1964, National Bureau of Standards.

[2] Cover T. M., Thomas J. A., "Elements of information theory", John Wiley & Sons, 1991, New York.

[3] Gupta P., Kumar P. R., "The capacity of wireless networks", IEEE Trans. on Information Theory, Vol. 46(2), 2000, pp. 388-404.

[4] Gupta P., Kumar P. R., "Internets in the sky: The capacity of three dimensional wireless networks", Communications in Information and Systems, Vol. 1(1), 2001, pp. 33-50.

[5] Hoeffding W., "Probability inequalities for sums of bounded random variables", J. of Amer. Stat. Assoc., Vol. 58, 1963, pp. 13-30.

[6] Kumar P. R., Xie L.-L., "A network information theory for wireless communications: scaling laws and optimal operation", IEEE Trans. on Information Theory, Vol. 50, Nr. 5, 2004, pp. 748-767.

[7] Marshall A. W., Olkin I., "Inequalities: theory of majorization and its applications", 1979, Academic Press.

[8] Oberhettinger F., "Tables of Fourier transforms and Fourier transforms of distributions", 1990, Springer Verlag.

[9] Telatar E., "Capacity of Multi-antenna Gaussian Channels", European Trans. on Telecommunications, Vol. 10, Nr. 6, 1999, pp. 585–595.