

Hierarchical Beamforming for Large Wireless Networks

Alla Merzakreeva, Olivier Leveque
Swiss Federal Institute of Technology - Lausanne, Switzerland
{alla.merzakreeva, olivier.leveque}@epfl.ch

Ayfer Özgür
Stanford University, Stanford CA
aozgur@stanford.edu

Abstract—We consider a wireless network with a large number of nodes distributed over a line. Under line-of-sight propagation, this network has only one degree of freedom for communication. At high SNR, this one degree of freedom can be readily achieved by multi-hop. At low SNR, however, the performance is dominated by the power transfer in the network. We show that none of the existing architectures, neither hierarchical cooperation nor multi-hop, can achieve optimal scaling of the capacity. We develop a digital hierarchical beamforming architecture and show that it is scaling optimal. This result reveals a new regime for large wireless networks, where beamforming techniques are needed to achieve capacity.

I. INTRODUCTION

What are good architectures for communicating in wireless ad hoc networks? To infer architectural insights that can guide communication system design for future wireless networks, we follow the *scaling law* approach initiated in [1]. Motivated by the massive proliferation of wireless devices, this approach focuses on the scaling of the capacity of a random network as the number of users gets large, and seeks to identify architectures that exhibit the optimal scaling. The intensive research effort in this line [2]–[7] has led to the characterization of scaling optimal architectures for large wireless networks in many regimes (see [8] for an overview).

The operating regime of a large network is determined by three parameters:

- the average SNR between neighboring nodes in the network, which depends both on the power available at the nodes and the distances between them;
- the spatial degrees of freedom of the network, determined by the area and the carrier wavelength;
- and the power path loss exponent α of the environment, capturing how fast signal power decays with distance.

Roughly speaking, when SNR is high and there are sufficient spatial degrees of freedom in the network (this is for example the case when the pairwise channels are subject to i.i.d. fading), cooperative MIMO based architectures can provide significant capacity gains over multi-hop [5]. Multi-hop is the traditional communication architecture for wireless networks, where information is routed from source nodes to destinations via multiple intermediate nodes, just like in wired

networks. The situation is more tricky for networks at low SNR. When the SNR is low but there are still sufficient spatial degrees of freedom in the network (as with i.i.d. fading), the optimal architecture depends on α . When α is small, a hierarchical cooperation architecture based on bursty distributed MIMO transmissions is optimal. At large α , multi-hopping is the right strategy. Interestingly, none of these two strategies makes use of beamforming, which is known to be the right strategy for point-to-point MIMO channels at low SNR [9]. Under i.i.d. fading, the distributed MIMO channels available in the network are well-conditioned and beamforming (or waterfilling over the eigenvalues of the channel) provides little gain. Transmitting independent streams for each node during the distributed MIMO transmissions is optimal.

The recent work [6], however, reveals that degrees of freedom in a wireless network can be limited by physical constraints in the spatial channel. This can be thought of as the spatial channel introducing correlation between pairwise gains. When the available degrees of freedom in the network are very few, they can be readily achieved by multi-hop. This makes multi-hop scaling optimal for such networks in the high SNR regime [6]. However, at low SNR, the performance is dominated by the power transfer in the network and it is not clear whether any of the existing architectures achieves the optimal scaling of the capacity.

In this paper, we explore a new regime where the network is *both* limited in power (operating at low SNR) and in spatial degrees of freedom (operating under strongly correlated pairwise channels). We show that a new class of cooperative beamforming architectures outperform classical multi-hopping when α is small. To capture this regime in a simple setup, we focus on a one-dimensional wireless network in a line-of-sight propagation environment. This leads to the extremal case when there is only a single degree of freedom for communicating in the network. We develop a hierarchical beamforming architecture for this network, where nodes first broadcast their information to a small cluster around them. This allows nodes to beamform and distribute this information over a larger scale. Continuing in a hierarchical fashion, the information of each source node is broadcasted to the whole network, including the destination node. This architecture is digital, as opposed to amplify-and-forward based beamforming techniques considered in the literature [10].

The work of A. Özgür was supported by the ERC grant NOWIRE ERC-2009-StG-240317 when the author was with the Swiss Federal Institute of Technology - Lausanne, Switzerland.

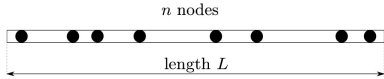


Fig. 1: One-dimensional network

II. MODEL

There are n nodes uniformly and independently distributed along a line of length L , as illustrated on Figure 1. Every node is both a source and a destination, and the sources and destinations are randomly paired up one-to-one. All source nodes wish to send a constant number of bits to their corresponding destination node at a common rate $R(n)$. The maximum achievable rate $R(n)$ is called the per-node throughput of the network. Correspondingly, the aggregate throughput of the network is defined as $T(n) = nR(n)$.¹

We assume that communication takes place over a flat channel with bandwidth W and that the received signal $Y_k[m]$ by node k at time m is given by

$$Y_k[m] = \sum_{j \in J} h_{kj} X_j[m] + Z_k[m]$$

where J is the set of transmitting nodes, $X_j[m]$ is the signal sent at time m by node j and $Z_k[m]$ is additive white circularly symmetric Gaussian noise (AWGN) of power spectral density $N_0/2$ Watts/Hz. In a line-of-sight environment, the complex baseband-equivalent channel gain h_{kj} between transmit node j and receive node k is given by

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

where G is Friis' constant, λ is the carrier wavelength, r_{kj} is the distance between node k and node j and $\alpha \geq 1$ is the power path loss exponent. Notice that the assumption $\alpha \geq 1$ in one-dimensional networks replaces the traditional assumption $\alpha \geq 2$ made in two-dimensional wireless networks. We assume a common average power budget per node P .

III. MAIN RESULT

Let us denote by SNR_s the signal-to-noise ratio over the typical nearest neighbor distance in the network². In a one-dimensional network, the typical nearest neighbor distance is $\frac{L}{n}$, therefore,

$$\text{SNR}_s = \frac{GP}{N_0 W} \left(\frac{n}{L} \right)^\alpha$$

A relatively straightforward analysis reveals that in one-dimensional networks, the multi-hop scheme described in [1] achieves with high probability³ an aggregate throughput of order

$$T(n) = \Omega \left(\frac{\text{SNR}_s}{\log n} \right)$$

¹Due to space constraints and for the sake of clarity, we will restrict ourselves in the proofs of the statements made below to *regular networks*, where nodes are equally spaced.

²where "s" stands for "short range SNR"

³meaning with probability $\geq 1 - \exp(-cn)$ for some constant $c > 0$

when $\text{SNR}_s \leq 0$ dB and $\alpha \geq 1$.

On the other hand, the best known information theoretic upper bound on the throughput scaling of such networks is given in [4]⁴:

$$T(n) = O(\log^3 n)$$

where again, $\text{SNR}_s \leq 0$ dB and $\alpha \geq 1$.

This shows that for constant SNR_s , multi-hop cooperation is order optimal and the aggregate throughput is constant, up to logarithmic factors. In the low SNR regime however (that is, when $\text{SNR}_s \ll 0$ dB), the question remains whether a more sophisticated strategy would not allow to achieve higher throughput scaling than multi-hop. We answer this question by the affirmative in the following theorem, in the case where the path loss exponent α lies between 1 and 2.

Theorem 1: Let us assume that $1 \leq \alpha < 2$ and $\text{SNR}_s \ll 0$ dB (i.e. $\text{SNR}_s = n^{-\gamma}$ for some $\gamma > 0$). Then for any $\varepsilon > 0$, there exists a communication scheme (referred to as "hierarchical beamforming" in the sequel) that achieves the following aggregate throughput with high probability as n gets large:

$$T(n) = \Omega(\min\{\text{SNR}_s n^{2-\alpha-\varepsilon}, n^{-\varepsilon}\}) \quad (2)$$

The above aggregate throughput scaling is strictly higher than that achieved by multi-hop. In particular, when $\alpha = 1$ and $\text{SNR}_s \leq 1/n$, $T(n) = \Omega(\text{SNR}_s n^{1-\varepsilon})$, which is an order n improvement over multi-hop. The hierarchical beamforming architecture achieving this performance is described in detail in the next section.

Is this strategy optimal or can we do better? Before answering this question, let us introduce the notion of a *broadcasting scheme* below.

Definition 1: A communication scheme achieving a per-node throughput $R(n)$ for n S-D pairs is said to be a broadcasting scheme if at this same rate $R(n)$, all destinations are able to decode the information sent not only by their corresponding source, but also by all the other sources.

As we will see, hierarchical beamforming enters into this category, and so does classical multi-hop in one-dimensional networks (at the price of a small adaptation of the original scheme). An interesting open question, that we do not address in the present paper, is whether *any* scaling optimal scheme is a broadcasting scheme in a one-dimensional network or not.

The theorem below, together with Theorem 2 above, shows that: a) hierarchical beamforming is scaling optimal when $\alpha = 1$; b) among all *broadcasting schemes*, hierarchical beamforming is scaling optimal when $1 \leq \alpha < 2$, and multi-hopping is when $\alpha \geq 2$.

Theorem 2: Consider a one-dimensional network with $\alpha \geq 1$ and $\text{SNR}_s \leq 0$ dB. Then:

a) The aggregate throughput scaling of any communication scheme is upper bounded with high probability by

$$T(n) = O(\min\{\text{SNR}_s n \log^2 n, \log^3 n\})$$

⁴Notice that the fading model considered in [4] is a simpler one with no phase shifts. It turns out however that in one-dimensional networks, adding phase shifts into the picture does not change the throughput scaling.

b) The aggregate throughput scaling of any broadcasting scheme is upper bounded with high probability by

$$T(n) = \begin{cases} O(\min\{\text{SNR}_s n^{2-\alpha} \log^2 n, \log^3 n\}) & \text{if } 1 \leq \alpha < 2 \\ O(\text{SNR}_s \log^2 n) & \text{if } \alpha \geq 2 \end{cases}$$

We prove this theorem in Section V.

IV. HIERARCHICAL BEAMFORMING

Let us start by considering the situation where the SNR in the network is very low. More precisely, let us assume that

$$\text{SNR}_s \leq n^{\alpha-2} \quad (\text{with } 1 \leq \alpha < 2) \quad (3)$$

In this regime, many transmissions can take place concurrently in the network (spatial reuse) without creating interference above the noise level. Under this assumption, the lower bound in Eq. (2) reads

$$T(n) = \Omega(\text{SNR}_s n^{2-\alpha-\varepsilon}) \quad (4)$$

We first sketch the hierarchical beamforming strategy we propose and then proceed to its performance analysis which also provides a more detailed description. Consider one particular source-destination pair $s-d$ in the network. For simplicity, assume that s has one bit to communicate to d . s can communicate this one bit in two steps:

- it can first broadcast this bit to a small cluster of M neighboring nodes around itself.
- the M nodes can then simultaneously transmit this bit to the destination node d by coherently combining their signals.

The beamforming gain due to the coherent combining of the M signal leads to a better performance than simply transmitting the bit from s to d .

From the network point of view, all source-destination pairs have to eventually accomplish these two steps. Step 2 is long-range communication and only one source-destination pair can operate at a time. Steps 1 involves local communication and can be parallelized across source-destination pairs. This leads to following two phases in the operation of the network:

1. The network is divided into clusters of M nodes. Each source node distributes one bit to the M nodes in its cluster. There are M source nodes in a cluster, which can simply take turns to distribute their one bit. When the total interference from the other clusters is below the noise level, this operation can be conducted in parallel among all clusters. At the end of this phase, each node has therefore received (and decoded) one bit from every other node in its cluster.

2. In the second phase, the bits are beamformed to their actual destinations one at a time. Every cluster performs M successive transmissions, in each transmission the bit of one particular source node in the cluster is beamformed to its destination. There are a total of n successive beamforming transmissions in this phase, one for each source-destination pair in the network.

A key observation is that this two phase scheme distributes the bits of every source node to all other nodes in the network,

even if this is not what we set for. In the second phase, the beamforming transmissions done one at a time can be decoded not only by the actual destination node but simultaneously by all the nodes in the network. This is a consequence of the fact that the network has only one degree of freedom. The transmitted signals from each cluster can be arranged to coherently combine simultaneously at all the remaining nodes in the network. Therefore according to Definition 1, this two phase scheme is a broadcasting scheme.

This brings the idea of recursion. The broadcasting requirement in the first phase can be handled by further dividing each cluster into smaller clusters and use the two-phase broadcast scheme we just described. The two phase scheme is illustrated in Figure 2. The recursion is summarized in the following lemma.

Note that contrary to the classical amplify-and-forward strategy that has been shown in [10] to be optimal at low power for a single S-D pair in a relay network, the scheme presented here is based on a *digital architecture*: at each step, all the nodes decode the broadcasted information before forwarding it further. This allows in particular to avoid the burden of noise amplification experienced by amplify-and-forward schemes.

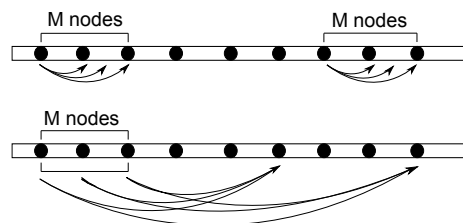


Fig. 2: Two-phase beamforming

Lemma 1: Consider $1 \leq \alpha < 2$ and a one-dimensional network with n nodes and $\text{SNR}_s \ll 0$ dB, subject to an additive external interfering source with bounded average power. If in this network, there exists a broadcasting scheme achieving with high probability an aggregate throughput

$$T(n) = \Omega(\text{SNR}_s n^\beta)$$

for some $\beta \leq 2 - \alpha$, then there exists another broadcasting scheme achieving with high probability an aggregate throughput

$$T(n) = \Omega(\text{SNR}_s n^{f(\beta)})$$

where

$$f(\beta) = 1 - \frac{\alpha(1-\beta)}{2-\beta} \quad (5)$$

Notice that $f(\beta) > \beta$ for all $1 \leq \alpha < 2$ and $\beta < 2 - \alpha$, so the performance of the new scheme is always strictly better than that of the original one. Figure 3 below illustrates the behavior of $f(\beta)$, for $\alpha = 1$ and $\alpha = 1.5$.

Proof of Lemma 1: Consider a regular network of n nodes and let us divide it into clusters of M nodes⁵, where $1 \ll M \ll n$. Based on the assumption made in the lemma, the

⁵In the random setting, these are clusters of length LM/n , containing each with high probability order M nodes.

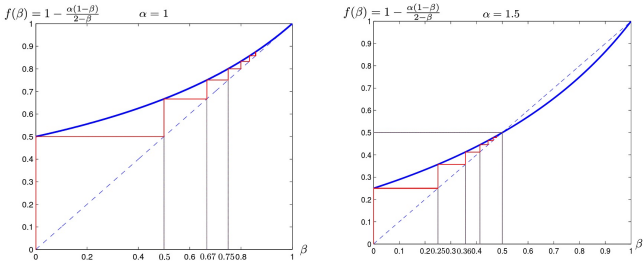


Fig. 3: Growth of the aggregate throughput exponent

following communication scheme will be shown to achieve the desired throughput scaling.

Phase 1. Source nodes broadcast information to every other node inside their cluster, using the original scheme with aggregate throughput $T(M) = \Omega(\text{SNR}_s M^\beta)^6$. This step is parallelized across clusters.

Phase 2. For each source node inside a cluster of M nodes, all the nodes inside the cluster simultaneously beamform the received bits to the rest of the network. During this second phase, only one cluster operates at a time.

Performance Analysis: In the first phase, clusters work in parallel. In order to avoid collisions between neighboring clusters, a simple time-division scheme with two rounds is used, where half of the clusters are active at a time: this only affects the throughput by a factor two and allows clear reception of the signals inside each cluster. One can indeed check that because of the assumption (3), the average power of the interference caused in one cluster by simultaneous transmissions in the other clusters is bounded.

The broadcasting rate achieved by the scheme inside each cluster is $R(M) = \Omega(\text{SNR}_s M^{\beta-1})$, so the total time taken by this first phase is upper bounded by

$$t_1 = O\left(\frac{1}{\text{SNR}_s M^{\beta-1}}\right)$$

In the second phase, M broadcast transmissions are performed sequentially from each cluster towards the rest of the network. As there are n/M clusters, the total number of transmissions is therefore equal to n (that is, one transmission takes place for each source node). The SNR of each transmission is lower bounded by

$$\text{SNR}_s \frac{n}{M} M^2 n^{-\alpha} = \text{SNR}_s M n^{1-\alpha}$$

where the above factors are explained as follows:

- the factor n/M is due to the fact that each cluster only transmits a fraction M/n of the time, so power can be spared during the rest of the time;
- the factor M^2 is the beamforming gain (notice that because of the line-of-sight channel model (1) and the assumption of a one-dimensional network, it is indeed possible to beamform the signal towards all destinations simultaneously);
- the factor $n^{-\alpha}$ is a lower bound on the power attenuation over distance.

⁶Notice that SNR_s , that only depends on the distance between neighboring nodes, remains unchanged for a cluster of size M or for the whole network.

The total time taken by this second phase is therefore upper bounded by

$$t_2 = O\left(\frac{n}{\text{SNR}_s M n^{1-\alpha}}\right) = O\left(\frac{1}{\text{SNR}_s M n^{-\alpha}}\right)$$

Optimal cluster size. In order to optimize the throughput of the new scheme, the optimal cluster size M^* should be chosen such that the durations of the two phases are equal, i.e. $t_1 = t_2$, which leads to

$$(M^*)^{\beta-1} = M^* n^{-\alpha} \quad \text{i.e.} \quad M^* = n^{\alpha/(2-\beta)} \quad (6)$$

(Notice that $\alpha/(2-\beta) \leq 1$, as $\beta \leq 2-\alpha$ by assumption).

Resulting aggregate throughput: With this cluster size, it is worthwhile noticing that the broadcasting rate of the new scheme is the *same* as the one achieved in each cluster with the original scheme. However, as more nodes participate to the transmission, the aggregate throughput increases as follows:

$$\begin{aligned} T(n) &= n R(M^*) = \Omega(n \text{SNR}_s (M^*)^{\beta-1}) \\ &= \Omega(\text{SNR}_s n^{f(\beta)}) \end{aligned}$$

where $f(\beta)$ is given in (5). This completes the proof. \blacksquare

Let us now explain how applying recursively Lemma 1 allows to obtain the lower bound (4) on the aggregate throughput scaling.

Let us first use multi-hop for broadcasting information at the lowest level of the hierarchy, that is, inside small clusters of M_1 nodes. Note that multi-hop can be easily transformed into a broadcasting scheme in the one-dimensional case without changing its aggregate throughput scaling; since information is routed over a line, each destination already observes the information sent by order n nodes on average. The aggregate throughput achieved inside each cluster is therefore

$$T(M_1) = \Theta\left(\frac{\text{SNR}_s}{\log M_1}\right) = \Omega(\text{SNR}_s M_1^\beta) \quad \forall \beta < 0$$

Using then the two-phase scheme described in the proof of Lemma 1, we reach for larger clusters of size M_2 (to be specified below) an aggregate throughput

$$T(M_2) = \Omega(\text{SNR}_s M_2^{f(\beta)}) \quad \forall \beta < 0$$

Iterating this procedure $h-1$ times, until the large cluster size M_h reaches the network size n , we obtain the following aggregate throughput

$$T(n) = \Omega(\text{SNR}_s n^{f^{(h-1)}(\beta)}) \quad \forall \beta < 0$$

As illustrated on Figure 3, the sequence $f^{(h-1)}(\beta)$ converges to the minimal solution of the equation

$$\beta^* = f(\beta^*)$$

which is given by $\beta^* = 2-\alpha$ for $1 \leq \alpha < 2$. For a fixed number of hierarchical levels h , the achieved aggregate throughput scaling is therefore $T(n) = \Omega(\text{SNR}_s n^{2-\alpha-\varepsilon})$, and $\varepsilon > 0$ can be made arbitrarily small by increasing the number h . \blacksquare

In addition, let us describe how to compute the optimal cluster sizes M_1, \dots, M_h in this process. From Eq. (6) in the proof of Lemma 1, we deduce that at level $1 \leq k < h$,

$$M_k = (M_{k+1})^{\alpha/(2-\beta(k))}$$

where $\beta(k)$ is the aggregate throughput exponent achieved at level k . This allows to compute recursively the cluster sizes, starting from $M_h = n$. From this analysis, it turns out that as h gets large, the optimal cluster size M_1 at the lowest level of the hierarchy converges to

$$M_1 = n^{\alpha-1}$$

So when $\alpha = 1$, the hierarchical beamforming scheme starts directly from tiny clusters, whereas when $1 < \alpha < 2$, the optimal communication strategy is first to perform multi-hop inside clusters of size $n^{\alpha-1}$, and then to use hierarchical beamforming. We therefore see that in the latter case, because of the higher value of the path loss exponent α , beamforming only helps when sufficiently many nodes participate to the transmission.

Finally, let us mention what happens at moderately low SNR, i.e. when $n^{\alpha-2} \leq \text{SNR}_s \leq 1$. In this case, the interference felt from the simultaneously transmitting clusters might hurt the transmissions inside a cluster. A simple solution to this problem is to reduce the power used by each node, so as to meet the equality $\text{SNR}_s = n^{\alpha-2}$. In this case, the aggregate throughput of the scheme is arbitrarily close to a constant, which proves the claim made in Theorem 1.

V. UPPER BOUNDS ON THE THROUGHPUT SCALING

In this section, we prove Theorem 2. Notice that in both parts a) and b), the stated $O(\log^3 n)$ bound comes from [4].

Proof of Theorem 2.a): For a regular network, the proof follows from the following simple observation: the per-node throughput $R(n)$ is upper bounded by the maximum mutual information from a given source node and the rest of the network, as illustrated on Figure 4.

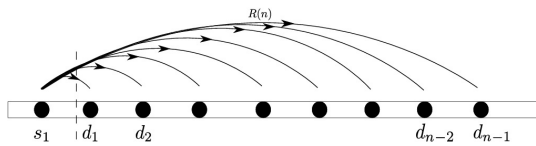


Fig. 4: Cut around a source node

In particular, using the fact that $\log(1+x) \leq x$, we obtain

$$\begin{aligned} R(n) &\leq \log \left(1 + \frac{P}{N_0 W} \sum_{k=2}^n |h_{k1}|^2 \right) \\ &\leq \text{SNR}_s \sum_{k=2}^n \frac{1}{(k-1)^\alpha} = O(\text{SNR}_s \log n) \end{aligned}$$

for any $\alpha \geq 1$, which implies the above upper bound on the aggregate throughput. An extra $\log n$ factor appears when considering random node positions. ■

Proof of Theorem 2.b): Considering again a regular network, let us examine the cut illustrated on Figure 5.

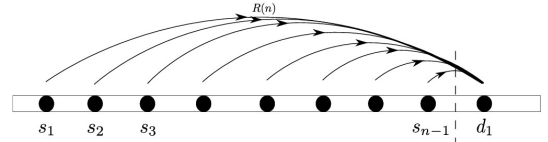


Fig. 5: Cut around a destination node

As we are assuming that each destination node must decode the information from all other nodes in the network (with each node sending a different message), the maximum mutual information between all the nodes on the left and the destination node on the right is an upper bound on the sum rate of communications going from left to right, i.e. $(n-1)R(n)$. This upper bound therefore reads

$$(n-1)R(n) \leq \sup_{Q \geq 0: \text{Tr} Q \leq P} \log \left(1 + \frac{h^\dagger Q h}{N_0 W} \right)$$

where h is the $(n-1) \times 1$ vector of fading coefficients and Q is the $(n-1) \times (n-1)$ input covariance matrix. This expression is in turn upper bounded by

$$\begin{aligned} (n-1)R(n) &\leq \log \left(1 + \frac{P}{N_0 W} \left(\sum_{j=1}^{n-1} |h_{nj}| \right)^2 \right) \\ &\leq \text{SNR}_s \left(\sum_{j=1}^{n-1} \frac{1}{(n-j)^{\alpha/2}} \right)^2 \\ &= \begin{cases} O(\text{SNR}_s n^{2-\alpha} \log n) & \text{if } 1 \leq \alpha < 2 \\ O(\text{SNR}_s \log n) & \text{if } \alpha \geq 2 \end{cases} \end{aligned}$$

As $T(n)$ is clearly of the same order as $(n-1)R(n)$, this settles the proof in the case of a regular network. The extra $\log n$ factor appears again when considering random node positions. ■

REFERENCES

- [1] P. Gupta, P. R. Kumar, *The Capacity of Wireless Networks*, IEEE Trans. on Information Theory 42 (2), 2000, 2313–2328.
- [2] 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.
- [3] 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.
- [4] A. Özgür, O. Lévêque, E. Preissmann, *Scaling Laws for One and Two-Dimensional Random Wireless Networks in the Low Attenuation Regime*, IEEE Trans. on Information Theory 53 (10), 2007, 3549–3572.
- [5] A. Özgür, O. Lévêque, D. Tse, *Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad-Hoc Networks*, IEEE Trans. on Information Theory 53 (10), 2007, 3549–3572.
- [6] M. Franceschetti, M.D. Migliore, P. Minero, *The Capacity of Wireless Networks: Information-Theoretic and Physical Limits*, IEEE Trans. on Information Theory 55 (8), 2009, 3413–3424.
- [7] U. Niesen, P. Gupta, D. Shah, *On Capacity Scaling in Arbitrary Wireless Networks*, IEEE Trans. on Information Theory 55 (9), 3959–3982, September 2009.
- [8] A. Özgür, O. Lévêque, D. Tse, *Operating Regimes of Large Wireless Networks*, Foundations and Trends in Networking, Now Publishers, 2011.
- [9] D. Tse, P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [10] U. Niesen and S. Diggavi, *The Approximate Capacity of the Gaussian N-Relay Diamond Network*, submitted to IEEE Trans. on Information Theory, August 2010.