Beyond Multi-hop: Optimal Cooperation in Large Wireless Networks

(Invited Paper)

Ayfer Özgür, Olivier Lévêque
EPFL, Switzerland
{ayfer.ozgur, olivier.leveque}@epfl.ch

David Tse
University of California at Berkeley
dtse@eecs.berkeley.edu

***Abstract*—Multi-hop is the traditional architecture for wireless adhoc networks. In this paper, we investigate the potential gains from more sophisticated cooperation in large wireless adhoc networks. While the capacity of multi-hop is limited to $\Theta(\sqrt{n})$ due to interference, we show that a hierarchical cooperation architecture can achieve *linear* capacity scaling in the number of users n. We also characterize how the cooperation gain is affected when the network is limited in either power or space.**

I. Introduction

Multi-hop is the communication architecture of current wireless networks such as mesh or adhoc networks. Packets are sent from each source to its destination via multiple point-to-point transmissions between relaying nodes. The origins of this approach are rooted in the practice of traditional wire-line networks. Today, the increasing need to connect a massive number of wireless devices and to support various resource intensive applications necessitates to discuss the large system performance of this architecture: Can multi-hop efficiently support communication in large wireless networks or do we need new architectures for the rapidly growing wireless networks of the future? In particular, can other architectures, tailored more carefully for wireless networks, significantly outperform multi-hop? These are important questions concerning future architectures for such networks. In this paper, we overview a theory on the capacity of large wireless networks that is able to shed some light on these questions.

The motivation for raising the above questions comes from the fact that signal interactions in wire-line and wireless networks are fundamentally different. Wire-line networks are composed of isolated point-to-point links over which signals travel independently. The signal interaction in wireless networks, however, is complex. The signal transmitted by a particular transmitter is not only heard by its intended receiver, but by all receivers in the vicinity of the transmitter. This broadcasting property can be viewed as an additional resource to be exploited or as harmful interference to be avoided. Starting with the objective to create isolated point-to-point links inside the wireless network, the multi-hop architecture is compelled to the second point of view. It designates nodes as transmitter-receiver pairs and each receiver is to decode the message from its designated transmitter, regarding signals from other transmitters as interference degrading the quality of the communication. To allow receivers to recover their messages, excessive interference between simultaneous transmissions needs to be systematically avoided. In particular, direct communication between far-away source and destination pairs in the network is not preferable, as the interference generated by long-distance transmission would preclude most of the other nodes from communicating. Instead, it is more beneficial to confine to nearest neighbor communication and maximize the number of simultaneous point-to-point transmissions inside the network. However, this means that each packet has to be retransmitted many times before getting to its final destination, bringing a relaying burden over the network.

The impact of this relaying burden over system throughput is especially detrimental at large system size, when typically there are many source-destination pairs inside the network that want to communicate with each other. This impact has been quantified by the seminal work of Gupta and Kumar [1] in 2000. They show that in a network of n nodes, where typically there are $\Theta(n)$ source-destination pairs, the multi-hop architecture provides an aggregate throughput of $\Theta(\sqrt{n})$. This implies that the rate per source-destination pair has to decrease to zero like $\Theta(1/\sqrt{n})$ with increasing number of nodes n. This limitation is essentially due to the interference between long-distance point-to-point transmissions. Therefore, a natural question is whether this interference limitation can be overcome by removing the restriction to point-to-point communication and allowing more sophisticated cooperation between the nodes. In particular, are there cooperation architectures whose performance can *scale* with system size? This is the first question we answer in this paper. We present a hierarchical cooperation architecture that achieves an aggregate throughput of $\Theta(n^{1-\epsilon})$ for any $\epsilon > 0$. A scaling arbitrarily close to linear means that there is essentially no interference limitation: The rate for each source-destination pair does not degrade significantly, even if there are more and more users entering the network. The performance boost comes from cooperative MIMO transmissions between clusters of nodes. The key ingredient is a multi-scale, hierarchical cooperation architecture without significant overhead. This hierarchical architecture allows to efficiently organize communication so that each source-destination pair in the network can benefit from high-capacity, long-distance cooperative MIMO transmissions.

The next question we investigate in this paper is whether more sophisticated cooperation can still yield significant gain when the network is power-limited. This can be the case due to a number of reasons: 1) The transmit power available at the nodes can be small. 2) The geographical area of the network can be large. 3) The attenuation in the environment can be high. 4) The network can be operating on a large bandwidth. The objective in such wireless networks is not (only) to deal with interference but to transfer power efficiently to the receivers. We show that by turning mutually interfering signals into useful ones, cooperative MIMO based architectures can exploit the broadcasting nature of the wireless media for received power gain. This power gain, in turn, translates into significant capacity gain in power-limited wireless network.

In the last section of this paper, we investigate the impact of space on the maximal cooperation gain in wireless networks. The geographical area of the network together with the carrier wavelength determine the diversity available in the physical wireless channel. This, in turn, determines the number of spatial degrees of freedom available for communication. The multi-hop architecture is able to achieve at most $\Theta(\sqrt{n})$ degrees of freedom. We show that the spatial degrees of freedom available in the network can be significantly larger than $\Theta(\sqrt{n})$. The hierarchical cooperation architecture is able to achieve the full degrees of freedom in the network.

This paper is an initial exploration of radically new architectures for wireless networks, posing more questions than it answers. The theoretical basis is an approximate characterization of the information-theoretic capacity of wireless networks using scaling laws. There are many remaining questions related to exact performance gains (characterization of the constants preceding the scaling laws), optimization of the proposed architectures for complexity and delay, extension of the results to arbitrary topologies and traffic, the characterization of the overhead for coordinating nodes etc. We do not address these questions here, but rather seek to establish that the potential benefits from more sophisticated cooperation justify further investigation in this direction.

The current paper presents an overview mainly based on the results published in [4], [5], [6] and [7]. The reader is referred to these references for detailed proofs and a historical overview of the subject.

II. Interference

We consider the following random network model, initially proposed in [1]: There are n wireless nodes with transmitting and receiving capabilities, that are uniformly and independently distributed in a square of area A. Each node has an average power of P Watts and the network is allocated a total bandwidth of W Hertz. The signal received by node i at time-slot m is given by

$$Y_i[m] = \sum_{k \neq i} H_{ik}[m] X_k[m] + Z_i[m]$$

where $X_k[m]$ is the signal sent by node k at time m and $Z_i[m]$ is white circularly symmetric Gaussian noise of power spectral density $N_0/2$ Watts/Hz. The complex baseband-equivalent channel gain between node i and node k at time m is given by:

$$H_{ik}[m] = \sqrt{G}\; r_{ik}^{-\alpha/2}\; e^{j\,\theta_{ik}[m]} \tag{1}$$

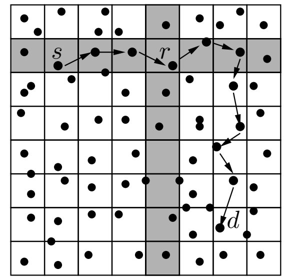

Fig. 1. The Multi-hop Architecture

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 i.i.d. across node pairs.[1] The phases can change ergodically over time but the distances r_{ik} are fixed over the duration of communication. We assume that the channel gains are known at all the nodes. The parameters G and $\alpha \geq 2$ are constants; α is called the path loss exponent of the environment.

Every node in the network is both a source and a destination for some traffic. The sources and destinations are randomly paired up one-to-one without any consideration of node locations. Each source has the same traffic rate $R(n)$ to send to its destination node. The aggregate throughput of the system is $T(n) = nR(n)$ in bits/s/Hz.We are interested in characterizing the scaling of the aggregate throughput $T(n)$ with increasing number of nodes n when the other parameters of the network A, P and W remain constant.

Gupta and Kumar showed in [1] that multi-hop achieves an aggregate throughput $\Theta(\sqrt{n})$ under this random network model w.h.p.[2] This result can be understood as follows: Let us divide the network into square cells of area A_c, containing $M = A_c n/A$ nodes on the average. In the multi-hop scheme, the messages of a source-destination pair s-d are relayed by hopping (decode-and-forward) from one cell to the next. Assume we follow a simplistic path between the source-destination pairs, first proceeding horizontally and then vertically as shown in Fig 1. Let us assign one node in each cell to do the relaying job. It is easy to observe that the relaying traffic at a particular relay node r is generated by either the source nodes located in the same horizontal slab or the destination nodes located in the same vertical slab as r. There are roughly $\sqrt{Mn}$ nodes contained in a slab of area $\sqrt{A_c A}$. This means that the relaying rate of each relay node has to be shared among $\Theta(\sqrt{Mn})$ source-destination pairs, giving $\Theta(1/\sqrt{Mn})$ throughput per source-destination pair. Reducing A_c to the nearest neighbor scale, which corresponds to $M = 1$, maximizes this rate and yields $\Theta(\sqrt{n})$ aggregate throughput.

We next present a hierarchical cooperation architecture that achieves almost linear aggregate throughput scaling under the same model.

Theorem 2.1: For any $\epsilon > 0$, there exists a constant $K_0 > 0$ independent of n such that w.h.p., an aggregate throughput

$$T(n) \geq K_0\, n^{1-\epsilon}$$

[1] The random phase model is assumed for simplicity and can be replaced by Rayleigh fading.

[2] with high probability: with probability approaching 1 as n increases.

is achievable in the network using hierarchical cooperation.

Using tools from information theory, it is easy to show that one cannot get a better capacity scaling than $O(n \log n)$, so the suggested scheme is very close to optimal [4].

The proof of Theorem 2.1 is based on the recursive application of the following proposition.

Proposition 2.1: Assume there exists a scheme that achieves an aggregate throughput

$$T(n) \geq K_1 n^b$$

in the network of n nodes w.h.p., where K_1 is a positive constant independent of n and $0 \leq b < 1$. Then one can construct another scheme for this network that achieves a *higher* aggregate throughput

$$T(n) \geq K_2 n^{\frac{1}{2-b}}$$

w.h.p., where $K_2 > 0$ is another constant independent of n.

Proof of Theorem 2.1: Proposition 2.1 is the key step to build a hierarchical architecture and prove Theorem 2.1. Since $\frac{1}{2-b} > b$ for $0 \leq b < 1$, the new scheme is always better than the old one. Therefore, as soon as we have a scheme to start with, the proposition can be applied recursively, yielding a scheme that achieves higher throughput at each step of the recursion. Starting with a simple time-division strategy between source-destination pairs that achieves $b = 0$, (alternatively we can also start with multi-hop achieving $b = 1/2$) and applying the proposition recursively h times we get a scheme that achieves $\Theta(n^{\frac{h}{h+1}})$ aggregate throughput. Choosing h large enough proves Theorem 2.1. □

We now sketch how the new scheme is constructed given the old scheme, and provide a back-of-the-envelope analysis of the scaling law it achieves.

Proof of Proposition 2.1: The scheme that proves Proposition 2.1 is based on clustering and long-range cooperative MIMO transmissions between clusters. We divide the network into clusters of M nodes. Let us focus for now on a particular source node s and its destination node d. s sends M bits to d in three steps:

(1) Node s distributes its M bits among the M nodes in its cluster, one for each node;
(2) These nodes together can then form a distributed transmit antenna array, sending the M bits *simultaneously* to the destination cluster where d lies;
(3) Each node in the destination cluster gets one observation from the cooperative MIMO transmission and it quantizes and ships the observation to d, which can then do joint MIMO processing of all the observations and decode the M transmitted bits.

From the network point of view, all source-destination pairs have to eventually accomplish these three steps. Step 2 is long-range communication and only one source-destination pair can operate at at a time. Steps 1 and 3 involve local communication and can be parallelized across source-destination pairs. Combining all this leads to three phases in the operation of the network:

Phase 1: Transmit Cooperation Clusters work in parallel. Within a cluster, each source node has to distribute its M bits among the other nodes, 1 bit for each node, such that at the end of the phase, each node has 1 bit from each of the source nodes in the same cluster. Since there are M source nodes in each cluster, this gives a total traffic of exchanging $M(M-1) \sim M^2$ bits. (Recall our assumption that each node is a source for some communication request and a destination for another.) The key observation is that this is similar to the original problem of communicating between n source and destination pairs, but on a smaller network of size M. More precisely, this traffic demand of exchanging M^2 bits can be handled by setting up M sub-phases, and assigning M source-destination pairs for each sub-phase to communicate their 1 bit. Since our channel model is scale invariant, the scheme given in the hypothesis of the proposition can be used in each sub-phase. With a scheme achieving aggregate throughput $\Theta(M^b)$, each sub-phase is completed in $\Theta(M^{1-b})$ time slots, so the whole phase takes $\Theta(M^{2-b})$ time slots.

Phase 2: Cooperative MIMO We perform successive long-distance cooperative MIMO transmissions between source-destination pairs, one at a time. In each one of the MIMO transmissions, say one between s and d, the M bits of s are simultaneously transmitted by the M nodes in its cluster to the M nodes in the cluster of d. The long-distance MIMO transmissions are repeated for each source-destination pair in the network, hence we need $\Theta(n)$ time-slots to complete the phase.

Phase 3: Receive Cooperation Clusters work in parallel. Since there are M destination nodes inside each cluster, each cluster has received M MIMO transmissions in phase 2. Each MIMO transmission is intended for a different destination node. Thus, each node in the cluster has M received observations, one from each of the MIMO transmissions, and each observation is to be conveyed to a different destination node in its cluster. Nodes quantize each observation into fixed Q bits (independent of M and n), so there are now a total of QM^2 bits to exchange inside each cluster. Using exactly the same strategy as in Phase 1, we conclude the phase in $\Theta(QM^{2-b})$ time slots.

Assuming that each destination node is able to decode the transmitted bits from its source node from the M quantized signals it gathers by the end of Phase 3, we can calculate the rate of the scheme as follows. Each source node is able to transmit M bits to its destination node, hence nM bits in total are delivered to their destinations in $\Theta(M^{2-b} + n + QM^{2-b})$ time slots, yielding an aggregate throughput of the order of

$$\frac{nM}{M^{2-b} + n + QM^{2-b}} \quad \text{bits per time slot.}$$

Maximizing this throughput by choosing $M = n^{\frac{1}{2-b}}$ yields $T(n) = \Theta(n^{\frac{1}{2-b}})$ for the aggregate throughput, which is the result in Proposition 2.1. □

Proving Theorem 2.1 by recursively applying Proposition 2.1, we have built a hierarchical architecture to achieve the desired throughput. At the lowest level of the hierarchy,

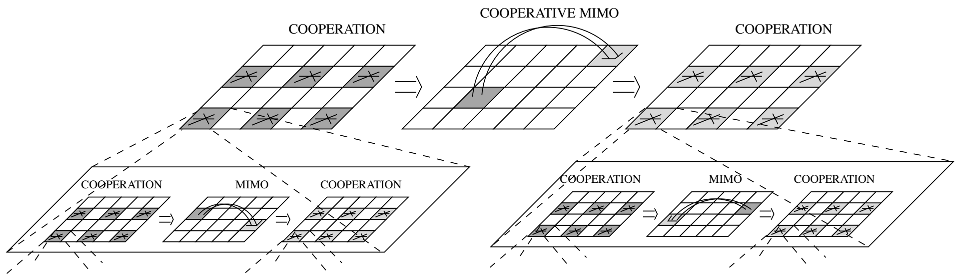

Fig. 2. The figure illustrates the salient features of the hierarchical cooperation architecture.

we use the simple time-division scheme to exchange bits for cooperation among small clusters. Combining this with longer range MIMO transmissions, we get a higher throughput scheme for cooperation among nodes in larger clusters at the next level of the hierarchy. Finally, at the top level of the hierarchy, the cooperation clusters are of the order of $\Theta(n^{1-\epsilon})$, almost the size of the network, and the MIMO transmissions are over the global scale, $\Theta(\sqrt{A})$ to meet the desired traffic demands. Figure 2 shows the resulting hierarchical scheme with a focus on the top two levels.

It is important to understand the aspects of the channel model which the scheme made use of in achieving the linear capacity scaling:

- The path attenuation decay law $1/r^\alpha$ ($\alpha \geq 2$) ensures that the *aggregate* signals from far away nodes are much weaker than signals from close-by nodes. This enables (a constant fraction of) the clusters to operate simultaneously in the first and the third phases. (This is similar to the spatial reuse in multi-hop.)
- Since received SNR between every pair of nodes in the network is lower bounded by a constant $\frac{GP}{N_0 W(\sqrt{2A})^\alpha}$ for every n, the scheme does not suffer any power-limitation.
- The random i.i.d. channel phases enable full spatial multiplexing gain for the long-range MIMO transmissions.

In the following two sections, we will investigate optimal cooperation in wireless networks for which the last two conditions fail to hold.

III. Power

In the previous section, we assumed that the parameters A, W, P of the network remain constant as the number of users n becomes large. This ensures that the received SNR between the most far-away pairs in the network is lower bounded by a constant even if n increases. This model allowed us to concentrate solely on interference in wireless networks. We now want to address networks which can potentially suffer power-limitation. We model this situation by allowing the received SNR between some pairs of nodes in the network, most notably between far away pairs, to decrease to zero with increasing n. We are still interested in the best scaling achievable for the aggregate throughput $T(n)$.

Let us start by discussing the performance of the earlier two schemes, multi-hop and hierarchical cooperation, when the network is power-limited. The multi-hop scheme employs only nearest-neighbor communication. Therefore it suffers power limitation only when the nearest neighbor transmissions inside the network are power-limited. More precisely, let us define the received SNR in a point-to-point transmission over the typical nearest neighbor distance to be,

$$\mathrm{SNR}_s := \frac{GP}{N_0 W(\sqrt{A/n})^\alpha}. \tag{2}$$

where $\sqrt{A/n}$ is the typical nearest neighbor distance inside the network. We refer to this quantity as the short-distance SNR of the network. The relaying rate at each hop is given by $\log(1+\mathrm{SNR}_s)$. Therefore, the performance of multi-hop is order-wise equal to

$$\begin{aligned} T_{MH} &= \sqrt{n}\log(1+\mathrm{SNR}_s) \\ &= \begin{cases} \Theta(\sqrt{n}) & \text{if} \quad \mathrm{SNR}_s \gg 0\,\mathrm{dB} \\ \Theta(\sqrt{n}\,\mathrm{SNR}_s) & \text{if} \quad \mathrm{SNR}_s \ll 0\,\mathrm{dB}, \end{cases} \end{aligned} \tag{3}$$

where $\mathrm{SNR}_s \ll 0\,\mathrm{dB}$ stands for SNR_s decreasing to 0 with increasing n. Similarly, $\mathrm{SNR}_s \gg 0\,\mathrm{dB}$ indicates that SNR_s is constant or increases with n. (We discard possible logarithmic terms.)

On the other hand, the backbone of the hierarchical cooperation architecture is long-distance cooperative MIMO transmission. The performance of the architecture is order-wise equal to the capacity of the MIMO transmissions at the highest level of the hierarchy, between clusters of size $\Theta(n^{1-\epsilon})$ separated by a distance $\Theta(\sqrt{A})$. The capacity of these MIMO transmissions is given by

$$\begin{aligned} T_{HC} &= n^{1-\epsilon}\log(1+\mathrm{SNR}_l) \\ &= \begin{cases} \Theta(n^{1-\epsilon}) & \text{if} \quad \mathrm{SNR}_l \gg 0\,\mathrm{dB} \\ \Theta(n^{1-\epsilon}\,\mathrm{SNR}_l) & \text{if} \quad \mathrm{SNR}_l \ll 0\,\mathrm{dB}, \end{cases} \end{aligned} \tag{4}$$

where SNR_l is the long-distance SNR in the network defined as

$$\mathrm{SNR}_l := n\frac{GP}{N_0\, W\,(\sqrt{A})^\alpha}. \tag{5}$$

Note that SNR_l is not simply the received SNR in a point-to-point transmission over the diameter $\sqrt{A}$ of the network, but it is n *times* this quantity. This leads to the following interesting fact: Even if the received SNR between far away pairs inside the network, of the order of $\frac{GP}{N_0\, W\,(\sqrt{A})^\alpha}$, decreases to zero as $1/n$ with increasing number of nodes n, the network does not suffer any power-limitation since SNR_l is constant in this case. We can still achieve linear capacity scaling with hierarchical cooperation. This is due to the fact that cooperative MIMO transmission not only allows to deal with interference but converting mutually interfering signals into useful ones, it also provides a factor of n power gain. This is also the case in the classical MIMO setup [2], [3]. Note that to achieve linear scaling in the number of antennas M in a MIMO channel with M Tx and M Rx antennas, we do not require the available power at the transmitter to scale linearly with M. A constant total transmit power suffices for linear capacity scaling of MIMO.

Note that the short and the long distance SNR's in the network are related as $\text{SNR}_s = n^{\alpha/2-1}\text{SNR}_l$. For $\alpha > 2$, SNR_s is always larger than SNR_l. The network starts to experience power limitation when $\text{SNR}_l \ll 0$ dB. Considering the performances in (3) and (4) together with the relation $\text{SNR}_s = n^{\alpha/2-1}\text{SNR}_l$, we observe that hierarchical cooperation performs better than multi-hop when $2 \leq \alpha \leq 3$. Signal power decays slowly with distance in this case and hierarchical cooperation yields maximal received power by collecting the received signals of a large number of nodes. When $\alpha > 3$ signal power decays fast with distance and long-distance communication is not preferable. Multi-hop performs better than hierarchical cooperation in this case.

Is this the best performance we can get in power-limited wireless networks with $\text{SNR}_l \ll 0$ dB? We next show that a hybrid architecture combining hierarchical cooperation with multi-hop performs significantly better than either of the schemes when $\alpha > 3$ and $\text{SNR}_s \gg 0$ dB. Note that since $\text{SNR}_s = n^{\alpha/2-1}\text{SNR}_l$, there is a wide range of parameters where $\text{SNR}_s \gg 0$ dB while $\text{SNR}_l \ll 0$ dB. This corresponds to the heterogeneous case where the short-range links in the network are strong (of high SNR) and the long-range links are weak (of low SNR).

Theorem 3.1: Let $\alpha > 2$, $\text{SNR}_s \gg 0$ dB and $\text{SNR}_l \ll 0$ dB. For any $\epsilon > 0$, there exists a constant $K_3 > 0$ independent of n such that w.h.p., an aggregate throughput

$$T_{HC+MH} \geq K_3 \sqrt{n}\, \text{SNR}_s^{\frac{1}{\alpha-2}-\epsilon}.$$

is achievable in the network using a hybrid architecture combining hierarchical cooperation with multi-hop.

Combining the performances achieved by hierarchical cooperation and multi-hop together with Theorem 3.1, we obtain the following approximation formula for the capacity of large wireless networks:

$$C \propto \begin{cases} n & \text{SNR}_l \gg 0 \text{ dB} \\ n\ \text{SNR}_l & \text{SNR}_l \ll 0 \text{ dB and } 2 \leq \alpha \leq 3 \\ \sqrt{n}\ \text{SNR}_s & \text{SNR}_s \ll 0 \text{ dB and } \alpha > 3 \\ \sqrt{n}\ \text{SNR}_s^{\frac{1}{\alpha-2}} & \text{SNR}_l \ll 0 \text{ dB, } \text{SNR}_s \gg 0 \text{ dB} \\ & \quad \text{and } \alpha > 3. \end{cases} \tag{6}$$

The schemes achieving the above performances are summarized in Fig. 3. The optimality of these schemes has been established by information theoretic arguments in [5].

Proof of Theorem 3.1: We next sketch how the hybrid architecture operates. On the global scale, the hybrid scheme is similar to multi-hop. The network is divided into cells and the packets of each source-destination pair are transferred by hopping from one cell to the next. At each hop, the associated relay node decodes the packets from the previous cell and forwards them to the next. The architecture differs from multi-hop

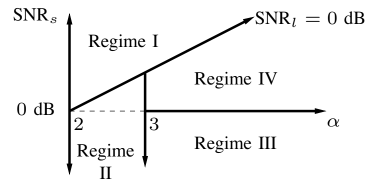

Fig. 3. The four operating regimes of large wireless networks in (6). The corresponding optimal schemes are I-II: Hierarchical Cooperation, III-Multi-hop, IV- Hybrid Multi-hop + Hierarchical Cooperation.

by the fact that each hop is performed via cooperative MIMO transmission assisted by hierarchical cooperation. Let us divide the network into cells of area A_c, containing $M = A_c n/A$ nodes on the average. As in the case of multi-hop, the relaying burden imposed on a given cell is due to the $\Theta(\sqrt{Mn})$ source-destination paths that pass through this cell.

Note that any two neighboring cells in the network can be viewed as a small wireless network of $2M$ nodes randomly and uniformly distributed on a rectangular area $2\sqrt{A_c} \times \sqrt{A_c}$. (Consider for example the two cells highlighted in Fig. 4.) Assume we pair the M nodes in one cell with the M nodes in the other cell into M distinct source-destination pars. Assume the long distance SNR in this small network $\text{SNR}_l(A_c) \gg 0$ dB where,

$$\text{SNR}_l(A_c) = M \frac{GP}{N_0\, W\, (\sqrt{A_c})^\alpha}.$$

From (4), using hierarchical cooperation to establish the M communications between the two cells, we can get an aggregate throughput $\Theta(M^{1-\epsilon})$. This corresponds to the total outbound relaying rate of a cell. It has to be divided among the $\Theta(\sqrt{Mn})$ source-destination pairs for which the cell is responsible. This yields a rate $\Theta(\sqrt{M}M^{-\epsilon}/\sqrt{n})$ per source-destination pair. Equivalently, the aggregate throughput achieved by the hybrid architecture is given by

$$T_{HC+MH} = \Theta(\sqrt{n}\, M^{1/2-\epsilon}). \tag{7}$$

Note that combining multi-hop with hierarchical cooperation provides a-$\sqrt{M}$-fold-gain in the aggregate throughput as compared to pure multi-hop, which indeed corresponds to $M = 1$ in the above discussion. Choosing larger M yields a larger aggregate throughput as it increases the hop capacity. Indeed, if we could choose $M = n$, we could get linear scaling in which case the scheme reduces to pure hierarchical cooperation. However since $\text{SNR}_l \ll 0$ dB, the condition $\text{SNR}_l(A_c) \gg 0$ dB is not satisfied for $A_c = A$, so M can not be as large as n. The largest cluster area that satisfies the condition $\text{SNR}(A_c) \gg 0$ dB is given by

$$A_c = (A/n)\, \text{SNR}_s^{1/(\alpha/2-1)}, \qquad M = \text{SNR}_s^{1/(\alpha/2-1)}. \tag{8}$$

This is the largest geographical scale in the network over which the power-limitation is not felt. Any larger cluster size increases the relaying burden without increasing the hop capacity. Combining (8) and (7) gives the result in Theorem 3.1.

IV. Space

When discussing the hierarchical cooperation scheme in Section II, we noted that the i.i.d. random phase model

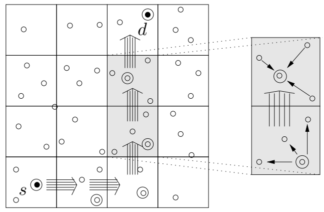

Fig. 4. Cooperate locally, multi-hop globally: A generic optimal architecture for wireless networks. The two extremes of this architecture are precisely traditional multi-hop, where the cluster size is 1 and the number of hops is $\Theta(\sqrt{n})$, and hierarchical cooperation where the cluster size is $\Theta(n)$ and the number of hops is 1.

enables full spatial multiplexing gain for the cooperative MIMO transmissions. This corresponds to having n degrees of freedom for communication inside the wireless network. When the geographical area of the network is small, the diversity in the physical channel is limited. Using principles from electromagnetics, Franceschetti et al. showed in [8] that the total degrees of freedom in a wireless network are upper bounded by

$$\Theta\left(\max\left(\sqrt{n}, \min(n, \sqrt{A}/\lambda)\right)\right),$$

where λ is the carrier wavelength. This implies that the available degrees of freedom in the network are n only if $\sqrt{A}/\lambda > n$. For networks with $\sqrt{A}/\lambda < n$, this number puts a limitation on the maximum possible cooperation gain. In particular, if $\sqrt{A}/\lambda < \sqrt{n}$, there are only $\sqrt{n}$ degrees of freedom in the network which can be achieved by multi-hop.

When $\sqrt{A}/\lambda > \sqrt{n}$, we want to investigate whether the full degrees of freedom in the network $\min(n, \sqrt{A}/\lambda)$ are achievable. For this purpose, we refine our phase model in Section II to the following physical channel model

$$H_{ik} = \sqrt{G}\,\frac{e^{j2\pi r_{ik}/\lambda}}{r_{ik}}. \tag{9}$$

This model corresponds to free-space propagation ($\alpha = 2$) in a line-of-sight type environment, a case in which spatial limitation is expected to be most severe.The following theorem is refinement of Theorem 2.1 under the above physical model.

Theorem 4.1: Under the channel model in (9) and when $\sqrt{A}/\lambda > \sqrt{n}$, the total throughput achieved by hierarchical cooperation is lower bounded by,

$$T \geq K_4\left(\min(n, \sqrt{A}/\lambda)\right)^{1-\epsilon}$$

w.h.p., for any $\epsilon > 0$ and a constant $K_4 > 0$ independent of the system parameters n, A and λ.

Accordingly, the optimal operation of the network falls into three cases:

1) $\sqrt{A}/\lambda \leq \sqrt{n}$: The number of spatial degrees of freedom is too small, sophisticated cooperation is useless and nearest neighbor multi-hopping is optimal.
2) $\sqrt{A}/\lambda > n$: The number of spatial degrees of freedom is n, cooperation is very useful, and the optimal performance can be achieved by the hierarchical cooperation scheme. Spatial degree of freedom limitation does not come into play and the performance is *as though* the phases are i.i.d. uniform across the nodes.
3) $\sqrt{n} \leq \sqrt{A}/\lambda \leq n$: The number of degrees of freedom is smaller than n, so the spatial limitation is felt, but larger than what can be achieved by simple multi-hopping. A modification of the hierarchical cooperation scheme achieves optimal scaling in this regime.

V. Conclusion

Can more sophisticated cooperation techniques provide significant capacity gains over the conventional multi-hop architecture in large wireless networks? We showed that the precise answer to this question depends on the operating regime where a particular network lies. Each operating regime corresponds to a subset of the parameter space where the optimal cooperation architecture is different. In many regimes, architectures better tailored for wireless networks, such as hierarchical cooperation, can provide substantial capacity gains over conventional multi-hop. In certain regimes, most notably when the network is severely limited in either power or space, multi-hop is fundamentally optimal. The operating regime of a given wireless network can be determined by computing or directly measuring certain engineering quantities that we identified in our discussion, such as short-range SNR, long-range SNR, area, power path loss exponent etc.

To conclude, let us apply the insights summarized in this paper to a numerical example: Consider a large network serving $n = 10,000$ users on a campus of $A = 1$ km^2, operating at 3 GHz ($\lambda = 0.1$ m). Note that although the network is quite dense, $\sqrt{A}/\lambda = 10000$ so we do not expect to observe any space limitation. Under free-space propagation and assuming unit transmit and receive antenna gains, the attenuation G in (1) and (9), given by the Friis formula $G = \frac{G_{Tx}\cdot G_{Rx}\cdot\lambda_c^2}{16\pi^2}$, is 10^{-6}. Assuming transmit power P of 1 mW per node, thermal noise N_0 at -174 dBm, a bandwidth W of 10 MHz and noise figure NF= 10 dB, the SNR between farthest pairs is 34 dB and SNR$_l$ = 84 dB, very much in the high SNR regime. So while multi-hop can achieve a total throughput of the order of 100 bits/s/Hz, hierarchical cooperation promises an aggregate throughput of the order of 10000 bits/s/Hz.

References

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

[2] G. J. Foschini, *Layered Space-Time Architecture For Wireless Communication in a Fading Environment when Using Multi-Element Antennas*, AT&T Bell Labs. Tech. J. 1 (2), 41–59, 1996.

[3] E. Telatar, *Capacity of Multi-Antenna Gaussian Channels*, European Trans. on Telecommunications 10 (6), 585–596, November 1999.

[4] A. Özgür, O. Lévêque, D. Tse, *Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad-Hoc Networks*, IEEE Trans. on Information Theory 53 (10), pp.3549-3572, 2007.

[5] A. Özgür, R. Johari, O. Lévêque, D. Tse, *Information Theoretic Operating Regimes of Large Wireless Networks*, IEEE Trans. on Information Theory 56 (1), pp.427-437, 2010.

[6] A. Özgür, O. Lévêque and D. Tse, *Linear Capacity Scaling in Wireless Networks: Beyond Physical Limits?*, Proc. IEEE Information Theory and Applications Workshop, San Diego, Feb 2010.

[7] A. Özgür Aydin. Fundamental limits and optimal operation in large wireless networks. PhD thesis, Lausanne, 2009.

[8] M. Franceschetti, M.D. Migliore, P. Minero, *The capacity of wireless networks: information-theoretic and physical limits*, preprint, 2007.