Shariatpanahi et al. EURASIP Journal on Wireless Communications and Networking 2012, 2012:290
http://jwcn.eurasipjournals.com/content/2012/1/290

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Analysis of quantization effects on the performance of cooperative MIMO techniques in wireless networks

Seyed Pooya Shariatpanahi[1], Babak Hossein Khalaj[1], Hamed Shah-Mansouri[1] and Javier Del Ser[2*]

Abstract

The most successful achievable schemes for ad hoc wireless networks are those based on establishing cooperative multiple-input and multiple-output links. In this article, we analyze one of the important design parameters of such schemes, namely the number of quantization bits. Due to the digital architecture of these schemes, the received signal at nodes should become quantized before further processing. The scheme's aggregate throughput highly depends on the resolution of the quantization process. We demonstrate that there is an optimum number of quantization bits which maximizes the network throughput. We show that the optimum number of quantization bits scales as $\beta \log_2(\text{SNR})$, for any strictly positive β independent of SNR, for the high SNR regime. Furthermore, we derive the optimum scaling of network throughput in such a regime. It is concluded that a good management of the number of quantization bits as a design parameter has a significant impact on the network performance.

Keywords: Ad hoc wireless networks, Cooperative MIMO, Throughput maximization, Quantization effects

Introduction

Many applications in the wireless communication technology involve ad hoc deployment of a large number of wireless nodes. Due to the broadcast and superposition nature of the wireless medium, we face the interference phenomenon in such networks. The throughput of the network is the end-to-end rate that all the source–destination nodes can communicate with, in the presence of interference. In order to analyze the performance of such ad hoc wireless network architectures, one can think of the wireless network as a graph. In this graph, the vertices are the wireless nodes, and the edges are wireless links. As a result, by using graph-theoretic tools, the information flow properties of the network can be analyzed [1-3].

However, as graph models cannot capture all properties of wireless networks, the graph-theory-based approach has serious limitations. One of such limitations is that a graph cannot properly model broadcast and superposition nature of wireless channels. Therefore, other works address the problem from a semi-information-theoretic perspective. Based on this approach, the first successful attempt at deriving the scaling properties of the network capacity was presented by Gupta and Kumar [4]. Their result shows aggregate throughput scaling of order $\Theta(\sqrt{n})$ for arbitrary networks of n nodes. The article by Franceschetti et al. [5] extends the above result to random networks. A number of other articles, such as [6,7], also verified the $\Theta(\sqrt{n})$ aggregate throughput limitation, which results in $\Theta(1/\sqrt{n})$ throughput per node. These results indicate that the throughput per node does not scale with the number of nodes, which seems unsatisfying. However, almost all the aforementioned articles are based on multihop routing of information. The main reason for the $\Theta(1/\sqrt{n})$ limitation is that each block of information should pass through a large number of hops before reaching its destination.

The network performance, however, is not always so disappointing if we do not bound ourselves to the multihop transmission technology. Based on this fact, the work of Aeron and Saligrama [8] and Özgür et al. [9] proposes schemes employing cooperative multiple-input and multiple-output (MIMO) techniques as their main strategy. In such schemes each source node cooperates with

*Correspondence: javier.delser@tecnalia.com
[2]OPTIMA Business Area, TECNALIA RESEARCH & INNOVATION, 48170 Zamudio, Spain
Full list of author information is available at the end of the article

the nearby nodes to form a virtual MIMO transmitter. In addition, local nodes at the receive side form a virtual MIMO receiver. Upon transmission of the MIMO signal, the received signals at the receive nodes should become quantized. After quantization and further cooperation, the data are sent to and decoded at the destination node.

The scheme proposed in [9] achieves aggregate throughput of order n, which is superior comparing with the multihop-based results. This superior performance is due to the cooperative nature of the scheme. Consequently, it has sparked a new wave of interest in the researchers to investigate the role of cooperative MIMO schemes in ad hoc networks (e.g., see [10-12]).

However, in order to make such schemes practical and usable in real-life protocols, there is still a large number of issues to be addressed. For example, there is the question whether the scaling-based analysis results (as $n \to \infty$) in [8,9] are reliable for finite size networks. In other words, how large should n be for the results to be valid? Similarly, many design parameters in [9] are still not addressed in the literature. While optimizing these parameters has a great impact on proper operation of the network, it is not clear what the optimum values for such design parameters are as functions of network characteristics.

In this article[a], we focus on one of the key design parameters of such schemes. As mentioned earlier, the nodes who act as the virtual MIMO receiver should quantize the received signal before sending these quantized data to the destination node for decoding. We focus on this quantization process, as this process is critical for the cooperative MIMO operation. The main questions we address in this article are: Does optimization of the number of quantization bits have a great effect on network throughput? How many number of bits each receive node should use for the quantization process? At high signal-to-noise ratios (SNR), what is the optimum choice for the number of quantization bits? How does the network throughput behave in high SNR regime?

In order to address the above issues, we analyze an inherent trade-off in determining the "proper" number of quantization bits. By increasing the number of quantization bits, the capacity of the virtual MIMO link is increased. However, the amount of data that the receive nodes should send to the destination node is increased as well. These two phenomena have competing effects on the network throughput. For analyzing this trade-off, we formulate the problem and find the optimum number of quantization bits maximizing the network throughput.

In order to clarify contributions of this article compared with [9,10], it should be noted that our focus in this article is on the optimization of the quantization process, rather than on proposing a new network communication scheme. In fact, the issue of optimization of number of quantization bits is not addressed in [9,10] while as will be shown, intelligent management of this issue results in significant throughput enhancement.

The structure of the article is as follows. In "Network model" section, we state our network model. In "Managing the number of quantization bits" section, we investigate the role of quantization bits and the resulting inherent trade-off. "Numerical illustrations" section contains helpful numerical illustrations. Finally, the article ends with conclusion.

At last, we review the notations used throughout the article. We use Knuth's asymptotic notation as follows [13]: $f(n) = O(g(n))$ if there exists positive constants c_0 and n_0 such that for all $n \geq n_0$ we have $0 \leq f(n) \leq c_0 g(n)$, $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$, $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. Also, matrices and vectors are indicated in boldface, $(\cdot)^H$ is the Hermitian operator, the expression $f_1(x) \sim f_2(x)$ is equivalent to $\lim_{x\to\infty} \frac{f_1(x)}{f_2(x)} = \text{constant} > 0$, and, $H(\cdot)$ is the entropy function.

Network model

In this section, we describe the network model used in this article. First, general assumptions about the network and the wireless channel model are described. Then, the model used for establishing cooperative MIMO links in the network is described.

General assumptions

Consider a wireless network consisting of n nodes capable of both transmitting and receiving signals. The nodes are uniformly distributed in area A in a random manner. Each node intends to transmit information to exactly one other node, and is also the receiver of information from one other node, through a uniform random mapping. If the set of transmitting nodes at each time slot is called $\mathbb{S}$, then the signal received by the ith node at time slot m is given by

$$Y_i[m] = \sum_{k\in\mathbb{S}} H_{ik}[m]\, X_k[m] + Z_i[m]\,, \tag{1}$$

where $X_k[m]$ is the signal transmitted by node k at the time slot m, and $Z_i[m]$ denotes noise at the receiver. The wireless channel between nodes i and k at time slot m is modeled as

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

where r_{ik} is the distance between nodes i and k, α is the path loss exponent, θ_{ik} is the random phase change between nodes i and k which is uniformly distributed in the range $[0, 2\pi)$, and G denotes antenna gain. This channel model has broadly been used in the literature, e.g., see [9,10]. Network throughput is defined as the rate at which all the source nodes send their information to destination nodes and depends on the spatio-temporal transmission strategies used in the network as discussed later.

Establishing cooperative MIMO links

The scheme considered in this article is based on cooperative MIMO transmission introduced in [9] and consists of the following three phases. In the first phase, each node shares its data between nearby nodes, such that these nodes can help in transmitting the data. In the second phase, those nearby nodes form virtual MIMO transmitters and transmit MIMO signals toward their destinations. In the last phase, upon reception of the signals transmitted in the second phase, the nearby nodes at the destination side form a virtual MIMO receiver and cooperate with each other to decode the original data at the destination node. As an example of such strategy, we continue our discussion based on the framework developed in [10]. In the network, each node is a source of data for exactly one other node, and also, it is destination of data for exactly one other node. Now, we describe the transmission procedure through which all source nodes transmit their data to the corresponding destinations in three phases. According to [10], the three transmission phases can be detailed as follows:

- **Phase 1:** Partition the network area into equal clusters such that each cluster contains M nodes. Suppose that every source node has ML bits to transmit to its destination. Let us focus on one source node s. Node s distributes its ML bits among other nodes inside its cluster, L bits for each. In order to do this, node s maps a data block of length L to a Gaussian codeword of length C_0, and transmits this codeword to another node inside its cluster. This action takes C_0 time slots. Then, s repeats the same action for all other nodes in the cluster in a round robin manner to distribute all the ML bits. Accordingly, such process takes $(M-1)C_0$ time slots. Since all the nodes in this cluster are source for some other nodes in the network, each node inside the cluster should repeat the whole procedure described above, as s did, one-by-one. This results in $M(M-1)C_0$ time slots needed for this specific cluster. This intra-cluster activity is done simultaneously in one-fourth of the clusters by a 4-TDMA spatial reuse scheme. (The k-TDMA spatial reuse strategy was first proposed in [4] and is a widely known concept in the literature. For details refer to [4,9,10].) Thus, the total time needed for all the source nodes in the network (which are, in fact, all the nodes in the network) to distribute their data inside their clusters is
$$\mathbb{D}(\text{phase 1}) = 4M(M-1)C_0 = 4M(M-1)\frac{L}{R_0}, \quad (3)$$
where the rate of each transmission—i.e., the number of bits per slot—is $R_0 \triangleq L/C_0$. In this phase, the transmission from node i to node j is with transmission power Pr_{ij}^{α}. It should be noted that by increasing transmit power P, R_0 cannot be made arbitrarily large, and it reaches a saturation value due to inter-cluster interference.
- **Phase 2:** Consider the source node s. Its destination is located in some other cluster called the receive cluster. After completing phase 1, ML bits of s is distributed among M nodes inside its cluster. Thus, each node in the cluster has L bits of s. In this phase, every node maps L bits of s to a C_1-long Gaussian codeword. Then, all nodes transmit their codewords simultaneously similar to a virtual MIMO transmitter. Each node transmits with power $P\frac{r_{SD}^{\alpha}}{M}$, where r_{SD} is the distance between the centers of the transmit cluster and the receive cluster. We denote the transmission rate of each node by $R_1 \triangleq \frac{L}{C_1}$. Each source node in the network needs a MIMO transmission shot as explained above, in order to deliver its data. Since we have a total of n source nodes in the network, an order of n MIMO shots, and accordingly, n time slots is required for this phase. Thus, the total number of time slots for this phase is[b]
$$\mathbb{D}(\text{phase 2}) = 2nC_1 = 2n\frac{L}{R_1}. \quad (4)$$
- **Phase 3:** Consider a given receive cluster with M nodes. Each receive cluster, in phase 2, has acted as a virtual MIMO receiver. However, since the received signals at each virtual MIMO receiver (each cluster) are not collocated (and are distributed among the nodes in the cluster), we cannot decode the message. Thus, in phase 3 we should gather these distributed signals at the destination, which will then decode the original message.
Consider a specific destination node named d which we assume to be the destination of source node s. As described earlier, in the second phase the nodes in the cluster of s send the information belonging to s in C_1 time slots in a MIMO fashion. Thus, after the second phase, every node in the receive cluster has received C_1 real numbers. These real numbers are the MIMO observations received at the antenna of each node in the receive cluster. Since we have to gather all these analog signals at the destination nodes d for the decoding process, we first have to quantize the analog observations. Thus, every node quantizes each observation with Q bits. Therefore, each node in the receive cluster has QC_1 bits which should be passed to d. Consequently, every node passes QC_1 bits to d in a TDMA manner using codewords of length C_0 exactly as in phase 1, providing the total of

$(M-1)C_0$ codewords for destination d. Finally, from these $(M-1)C_0$ codewords (along with the analog data already received at the second phase in d), node d decodes the whole ML bits of s.
This procedure should be repeated for all destination nodes in receive cluster. Since, there are M destination nodes in the receive cluster, each one should get its $(M-1)QC_1$ bits from other nodes, one-by-one. Considering the 4-TDMA spatial reuse strategy between clusters (as in the first phase), the total time slots needed for this phase is given by

$$\mathbb{D}(\text{phase 3}) = 4M(M-1)C_0C_1Q/L = 4M(M-1)\frac{LQ}{R_0R_1}. \quad (5)$$

Consequently, the total time required for this scheme for large values of M is

$$\mathbb{D} = 4M^2\frac{L}{R_0} + 2n\frac{L}{R_1} + 4M^2\frac{LQ}{R_0R_1}. \quad (6)$$

Since the total of nML bits is delivered, the effective throughput will be

$$T = \frac{nML}{4M^2\frac{L}{R_0}\left(1+\frac{Q}{R_1}\right) + 2n\frac{L}{R_1}}. \quad (7)$$

Accordingly, by optimizing M with respect to the achievable throughput T, as [10]:

$$M^{opt} = \left(\frac{R_0}{2(R_1+Q)}\right)^{1/2}\sqrt{n}, \quad (8)$$

the optimum throughput will be

$$T^{opt} = \left(\frac{R_0R_1^2}{32(R_1+Q)}\right)^{1/2}\sqrt{n}. \quad (9)$$

It is useful to note that the receive power at each of the nodes is bounded in the third phase as [10]

$$Pa^2 \le P^{rec} \le Pb^2, \quad (10)$$

where

$$a \triangleq \left(2-\sqrt{2}\right)^{\alpha/2}, \quad b \triangleq \left(2+\sqrt{2}\right)^{\alpha/2}. \quad (11)$$

The above framework is sufficient for our purposes in this article. However, an interested reader can find more detailed explanation of the scheme by referring to [10]. In the next section, we use this simple scheme to begin analyzing the quantization effects in the second and third phases described above.

Managing the number of quantization bits

The quantization process in the cooperative MIMO strategy is a critical operation before any further action for decoding. That is due to the fact that in order to pass the received information-carrying signal to the destination node (the decoder), these signals should first become quantized. These quantized signals constitute the information based on which the original data are recovered as will be discussed in this section.

Two stage cooperative MIMO network

Our goal is to study the effect of choosing a proper value for the number of quantization bits, on the performance of the network in terms of throughput. Thus, here we analyze a trade-off between the number of quantization bits, and the achieved throughput of the virtual MIMO channel. In fact, by decreasing Q, each node has to represent its received analog signal with more ambiguity. The distortion due to quantization, acts as an added noise and makes the MIMO decoding at the destination node more difficult. Consequently, the capacity of the virtual MIMO link is decreased. In other words, by decreasing the number of quantization bits, Q, R_1 in Equation (9) is decreased. In addition, as our goal is to achieve a low traffic load in the third phase, we have to decrease the number of quantization bits. In other words, by decreasing Q, the denominator of (9) is decreased. These are two competing effects which shape the trade-off behavior inherent in the problem.

As it is clear from Equation (9), optimizing Q will result in improvement in the pre-constant value of the throughput. Thus, the best Q can be found by solving the following optimization problem:

$$\arg\max_{Q} \frac{R_1(Q)^2}{R_1(Q)+Q}. \quad (12)$$

To this end, we need to understand how R_1 increases as Q is increased. R_1 is proportional to the MIMO capacity between the clusters in the second phase, i.e.,

$$R_1 = \frac{1}{M}C_{MIMO}, \quad (13)$$

where C_{MIMO} is the capacity of the cooperative MIMO link. The second phase signal model results in [9]

$$\mathbf{Y}_j = \mathbf{H}_{ij}\mathbf{X}_i + \mathbf{N}_j, \quad (14)$$

where $\mathbf{X}_i$ is the transmitted signal vector by the nodes in the transmit cluster, and $\mathbf{Y}_j$ is the received signal vector at the nodes in the receive cluster. Also, $\mathbf{H}_{ij}$ is the channel matrix of the cooperative MIMO setup, and $\mathbf{N}_j$ is the additive noise-plus-distortion vector added at the receive nodes. The noise-plus-distortion vector consists of two parts, namely the thermal noise ($\mathbf{Z}$), and the quantization distortion ($\mathbf{D}$)

$$\mathbf{N}_j = \mathbf{Z}_j + \mathbf{D}_j. \quad (15)$$

We consider the thermal noise of receiving nodes to be independent from each other. The thermal noise at different nodes are assumed to have power σ^2. The quantization distortion of the receive node i depends on Q, and based on the rate-distortion theory [14] is equal to

$$D_i = (P_i^{\text{rec}} + \sigma^2)2^{-Q}, \tag{16}$$

where P_i^{rec} is the power of the received signal at the node i. The received power for node i is the sum of the transmitted signals' power, diminished by the path loss effect. Given the fact that the power of each transmit node in the second phase is $P\frac{r_{SD}^{\alpha}}{M}$, we will have that [9]

$$P_i^{\text{rec}} = \sum_{k=1}^{M} \frac{P}{M}\left(\frac{r_{SD}}{r_{ki}}\right)^{\alpha}. \tag{17}$$

By considering the quantization distortion as noise, the channel model represented in (14) will be a MIMO channel with unequal noise power at the receive antennas. Obviously, this is not the best strategy since the distortion vector contains information about the transmitted signal and is not useless noise. However, with this assumption, we can note that R_1 will be

$$R_1 = \frac{1}{M}C_{MIMO} = \frac{1}{M}\mathbb{E}\left[\log_2\left(\frac{\det(\mathbf{\Sigma}'_N + \frac{\rho}{M}\mathbf{F}\mathbf{F}^H)}{\det \mathbf{\Sigma}'_N}\right)\right], \tag{18}$$

(The detailed derivation of (18) is presented in the Appendix.) where SNR is defined to be

$$\rho \triangleq P/\sigma^2, \tag{19}$$

and the expectation is with respect to the random channel phases. Also, we define the entries of the matrix $\mathbf{F}$ as $(\mathbf{F})_{ik} \triangleq \gamma_{ik}\exp(j\theta_{ik})$, where $\gamma_{ik} \triangleq (\frac{r_{SD}}{r_{ki}})^{\alpha/2}$. Also, from the geometry of the problem, we have that

$$a \leq \gamma_{ik} \leq b. \tag{20}$$

In addition, $\mathbf{\Sigma}'_N$ is the noise-plus-distortion covariance matrix given by

$$\mathbf{\Sigma}'_N = \mathbb{E}[\,\mathbf{N}\mathbf{N}^H], \tag{21}$$

where the expectation is with respect to $\mathbf{X}$ and $\mathbf{Z}$. Therefore, it is diagonal with diagonal entries

$$(\mathbf{\Sigma}'_N)_{ii} = 1 + \left(1 + \frac{\rho}{M}\sum_{k=1}^{M}\left(\frac{r_{SD}}{r_{ki}}\right)^{\alpha}\right)2^{-Q}. \tag{22}$$

It should be noted that due to the normalized values of distances in equation (22), the distance of the transmit and receive clusters is not important, and consequently the value of R_1 for all the cluster pairs is almost the same[c].

Now we are in a position to note how decreasing Q decreases R_1. By decreasing Q, the diagonal elements of the noise-plus-distortion covariance matrix stated in (22) (i.e., $(\Sigma'_N)_{ii}$) will be increased. This results in a decreased value for R_1. By understanding these two competing effects, we have evaluated the trade-off mentioned above, which suggests that there is an optimum number of quantization bits. As will be shown later, proper management of this trade-off has a great effect on enhancing the network throughput.

Extension to hierarchical cooperation schemes

The scheme explained in "Establishing cooperative MIMO links" section is a simple form of a more general scheme named Hierarchical Cooperation [9]. Consider a typical transmit cluster as explained in the previous sections. In the first phase, each node inside the cluster should distribute different data between other nodes inside the cluster. One can note that the process in phase 1 is very similar to the original problem in a smaller scale (by the original problem we mean: n nodes randomly distributed each having data for another node in the network). That is because the process in phase 1 and the original problem both consist of uniform traffic demand between the nodes of an specific area. Thus, we can divide the phase 1 into three sub-phases similar to the above-mentioned approach. The same story is true for the phase 3. Therefore, we can divide the phase 3 into three sub-phases too. This concept is illustrated in Figure 1. The big rectangles in Figure 1 illustrate the three phases at the first level (top level) of the hierarchy. The first and third phases of the first level are divided again into another three phases which are shown by small rectangles. These small rectangles constitute the second level of the hierarchy. Consequent levels of hierarchy are built in the same manner up to h levels. The hierarchical cooperation with h stages has been analyzed in [9,10] in detail. By this approach, the network throughput is [10][d]

$$T = \left(\frac{R_0}{R_1}\right)^{\frac{1}{h}} \frac{R_1}{h\left[4\left(1+\frac{Q}{R_1}\right)\right]^{(h-1)/2}}\left(\frac{n}{2}\right)^{1-\frac{1}{h}}. \tag{23}$$

As it is clear from Equation (23), optimizing Q will result in improvement in the pre-constant value of the throughput. Consequently, the optimization problem becomes

$$\arg\max_{Q}\left(\frac{R_0}{R_1(Q)}\right)^{\frac{1}{h}} \frac{R_1(Q)}{h\left[4\left(1+\frac{Q}{R_1(Q)}\right)\right]^{(h-1)/2}}. \tag{24}$$

This optimization problem reduces to (12) by putting $h = 2$.

In [11,12], it has been observed that phases 1 and 3 can be manipulated in a more effective way in terms of delay.

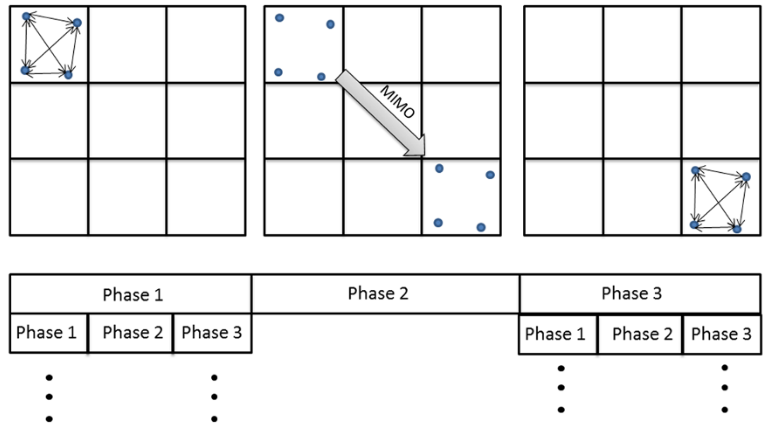

Figure 1 Illustration of the hierarchical cooperation concept. The whole operation is divided into three consecutive phases (the big rectangles). Then, in a hierarchical manner, the first and third phases are again divided into three phases (the small rectangles). This hierarchy is repeated h times.

It has been shown that this modified scheme achieves the throughput[e] given by [12]

$$T = \frac{R_0}{h\left(1+\frac{R_1}{Q}\right)^{(h-1)/h}\left(\frac{4Q}{R_1}\right)^{(h-1)/2}}\left(\frac{n}{2}\right)^{\frac{h-1}{h}}. \qquad (25)$$

For this delay-effective scheme, the optimization problem becomes

$$\arg\max_{Q} \frac{R_0}{h\left(1+\frac{R_1(Q)}{Q}\right)^{(h-1)/h}\left(\frac{4Q}{R_1(Q)}\right)^{(h-1)/2}}. \qquad (26)$$

In order to analyze the above optimization problems ((24) and (26)), we need to have more understanding of the behavior of R_1 as a function of Q. The exact expression for R_1 is the solution to the quantized MIMO channel capacity problem (Equation 18), which is very difficult to work with.

Therefore, by some simple observations, we propose a simple approximation for the behavior of R_1 as a function of Q as follows. Consider a typical MIMO transmission between two clusters. By varying the value of Q, the observed quantization noise in the diagonal elements of $\mathbf{\Sigma}'_N$ in Equation (22) varies accordingly. By substituting the covariance matrix $\mathbf{\Sigma}'_N$ in Equation (18), we can quantify the functionality of R_1 as a function of Q. Figure 2 shows a typical numerical example of R_1 as a function of Q derived by Equation (18). This figure shows that at small values of Q, the rate grows almost linearly with Q, until it gets a saturation value. In this saturation point, increasing Q does not help improving R_1. Let R_s be the saturation value for R_1. Then we have

$$R_s = \frac{1}{M}\mathbb{E}\left[\log_2 \det\left(\mathbf{I}+\frac{\rho}{M}\mathbf{F}\mathbf{F}^H\right)\right]. \qquad (27)$$

We can roughly say that this happens when the quantization distortion is negligible comparing with the thermal noise. Also, we see that until we reach this saturation point, it makes sense to approximate the curve with a linear one. Thus, we approximate $R_1(Q)$ by a piecewise linear function as depicted in Figure 3. The first parameter of this approximation is R_s which is the saturation capacity. The second parameter is the number of bits required for getting closed to this saturation capacity, namely Q_s. However, it should be noted that there are also some limitations for this kind of approximation. First, it should be noted that at low values of Q, Equation (18) is not valid and apparently the value of R_1 at $Q = 0$ is zero. Second, it should be noted that Equation (18) is not the capacity, and is just a lower bound, since we have assumed the quantization distortion to be just noise.

In order to find the optimum value of Q for the hierarchical cooperation scheme, we have to solve the problem (24). It is clear that the optimum value for Q happens before or at the saturation point (i.e., Q_s). By assuming

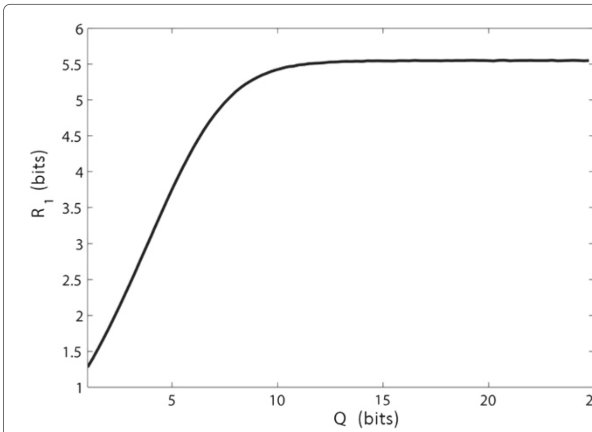

Figure 2 A numerical example to show how R_1 varies as Q varies.

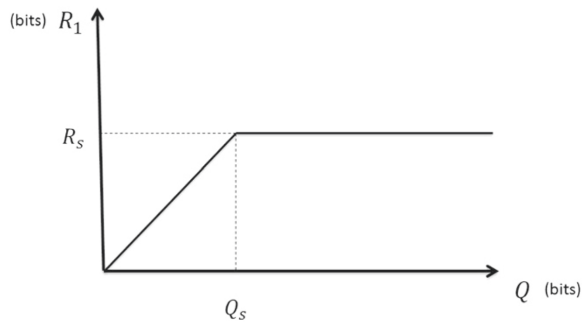

Figure 3 The proposed piecewise linear approximation for R_1, as a function of Q.

$Q \leq Q_s$ in Equation (24), we will have the following quantity to be maximized:

$$\max_{Q} \frac{1}{h} R_0^{1/h} \frac{\left(\frac{R_s}{Q_s} Q\right)^{1-1/h}}{\left[4\left(1+\frac{Q_s}{R_s}\right)\right]^{(h-1)/2}}, \tag{28}$$

which is maximized by putting $Q = Q_s$. This results in the following pre-constant value

$$\frac{1}{h} R_0^{1/h} \frac{R_s^{1-1/h}}{\left[4\left(1+\frac{Q_s}{R_s}\right)\right]^{(h-1)/2}}. \tag{29}$$

According to the above result, we should use the number of quantization bits needed to get near optimal MIMO capacity at saturation. But, further bit allocation will result in throughput loss, due to the extra load for the third phase.

In a similar way, we can analyze the optimization problem for the variant of the hierarchical cooperation scheme, proposed in [11]. By analyzing (26) we see that any $Q \leq Q_s$ maximizes the throughput pre-constant. Thus, the modified hierarchical cooperation scheme proposed in [11] is very insensitive to the choice of the number of quantization bits, as long as, we do not choose a number larger than Q_s. The corresponding pre-constant will be

$$\frac{R_0}{h\left(1+\frac{R_s}{Q_s}\right)^{(h-1)/h}\left(\frac{4Q_s}{R_s}\right)^{(h-1)/2}}. \tag{30}$$

The above discussions show that, in the hierarchical cooperation scheme and its variant, it is best to allocate enough quantization bits to have a near-optimal MIMO link in the second phase.

Analysis in the high SNR regime

One important issue is the behavior of the optimum number of quantization bits, and the network throughput pre-constant in the high SNR regime (i.e., $\rho \rightarrow \infty$). In this section, we carry out this analysis, and prove that the optimum number of quantization bits scales as $\Theta(\log_2(\rho))$. With this optimum scaling, the optimum throughput pre-constant scaling is of order $\Theta(\log_2(\rho)^{1-1/h})$.

In order to understand this issue, first, we should analyze R_1 in the high SNR regime. Equation (18) is very difficult to work with. Thus, we derive upper and lower bounds for (18) which are easier to work with. To do so, we put

$$1+\frac{(P_{\min}^{\text{rec}}+\sigma^2)2^{-Q}}{\sigma^2} \leq \left(\boldsymbol{\Sigma}'_N\right)_{ii} \leq 1+\frac{(P_{\max}^{\text{rec}}+\sigma^2)2^{-Q}}{\sigma^2} \tag{31}$$

where by using $P_{\min}^{\text{rec}}$ and $P_{\max}^{\text{rec}}$ from (10) we will get

$$1+\frac{(Pa^2+\sigma^2)2^{-Q}}{\sigma^2} \leq \left(\boldsymbol{\Sigma}'_N\right)_{ii} \leq 1+\frac{(Pb^2+\sigma^2)2^{-Q}}{\sigma^2}. \tag{32}$$

Also, we have from (20)

$$\begin{aligned} a\exp(j\theta_{ik}) \leq &(\mathbf{F})_{ik} \leq b\exp(j\theta_{ik}) \\ a(\mathbf{G})_{ik} \leq &(\mathbf{F})_{ik} \leq b(\mathbf{G})_{ik} \end{aligned} \tag{33}$$

where the entries of the matrix $\mathbf{G}$ are defined to be $(\mathbf{G})_{ik} \triangleq \exp(j\theta_{ik})$. Now that we have bounds for the elements of the matrices $\boldsymbol{\Sigma}'_N$ and $\mathbf{F}$ in (32) and (33), respectively, by inserting them into (18) the following bounds will result

$$\begin{aligned} \frac{1}{M}\mathbb{E}\left[\log_2\det\left(\mathbf{I}+\frac{1}{M}\frac{\rho a^2}{1+(\rho b^2+1)2^{-Q}}\mathbf{G}\mathbf{G}^H\right)\right] \leq R_1 \leq \frac{1}{M}\mathbb{E} \\ \left[\log_2\det\left(\mathbf{I}+\frac{1}{M}\frac{\rho b^2}{1+(\rho a^2+1)2^{-Q}}\mathbf{G}\mathbf{G}^H\right)\right]. \end{aligned} \tag{34}$$

In other words, for deriving (34) we have used Equation (18), where, we need using the bounds for the received signal powers at different nodes stated in (10) and the inequalities in (20), both originally derived in [9].

The above bounds on R_1 demonstrate an interesting fact. If we do not increase Q as SNR increases, for R_1 we will have

$$\begin{aligned} \frac{1}{M}\mathbb{E}\left[\log_2\det\left(\mathbf{I}+\frac{1}{M}\frac{a^2}{b^2}2^Q\mathbf{G}\mathbf{G}^H\right)\right] &\leq \lim_{\rho\rightarrow\infty} R_1 \\ &\leq \frac{1}{M}\mathbb{E}\left[\log_2\det\left(\mathbf{I}+\frac{1}{M}\frac{b^2}{a^2}2^Q\mathbf{G}\mathbf{G}^H\right)\right], \end{aligned} \tag{35}$$

and consequently, the pre-constant for the two-stage system (9) is bounded[f] as

$$\lim_{\rho\rightarrow\infty}\sqrt{\frac{R_0 R_1^2}{32(R_1+Q)}} < \infty. \tag{36}$$

However, if we set $Q = \beta \log_2(\rho)$, for values of $0 < \beta < 1$, the upper bound in the high SNR regime will be given by

$$\begin{aligned} R_{up} &\triangleq \frac{1}{M}\mathbb{E}\left[\log_2 \det\left(\mathbf{I} + \frac{1}{M}\frac{\rho}{1+(\rho a^2+1)2^{-Q}} b^2 \mathbf{G}\mathbf{G}^H\right)\right] \\ &\sim \frac{1}{M}\mathbb{E}\left[\log_2 \det\left(\mathbf{I} + \frac{1}{M}\frac{b^2}{a^2}\rho^{\beta} \mathbf{G}\mathbf{G}^H\right)\right] \\ &= \frac{1}{M}\mathbb{E}\left[\sum_i \log_2\left(1 + \frac{1}{M}\frac{b^2}{a^2}\rho^{\beta}\lambda_i^2\right)\right] \\ &\sim \log_2\left(\frac{b^2}{a^2}\rho^{\beta}\right) + C = \log_2\left(\frac{b^2}{a^2}\right) + \beta\log_2(\rho) + C, \end{aligned} \tag{37}$$

where

$$C \triangleq \frac{1}{M}\sum_i \mathbb{E}\log_2\left(\frac{1}{M}\lambda_i^2\right). \tag{38}$$

In (37) and (38), λ_i's are singular values of $\mathbf{G}$ which are independent of ρ. Also, in the case of $\beta > 1$, by similar calculations, we will have

$$R_{up} \sim \log_2(b^2) + \log_2(\rho) + C. \tag{39}$$

Finally, for $\beta = 1$:

$$R_{up} \sim \log_2\left(\frac{b^2}{1+a^2}\right) + \log_2(\rho) + C. \tag{40}$$

A similar analysis holds for the lower bound. Thus, by choosing $Q = \beta\log_2(\rho)$, for any $\beta > 0$,

$$\log_2(\rho) + C_1' \le R_1 \le \log_2(\rho) + C_2', \tag{41}$$

where C_1' and C_2' are two constant values independent of SNR. The inequalities in (41) indicate that R_1 is of the order $\log_2(\rho)$, and accordingly the pre-constant for the two-stage scheme (9) will increase as

$$\sqrt{\frac{R_0 R_1^2}{32(R_1+Q)}} \sim \sqrt{\log_2(\rho)}. \tag{42}$$

Also, by a similar approach, we see that for the h-stage hierarchical cooperation scheme, with the throughput indicated in (23), the pre-constant scaling is of order

$$\left(\frac{R_0}{R_1}\right)^{\frac{1}{h}} \frac{R_1}{h\left[4\left(1+\frac{Q}{R_1}\right)\right]^{(h-1)/2}} \sim \log_2(\rho)^{1-\frac{1}{h}}. \tag{43}$$

In addition, for the modified h-stage scheme, with the throughput stated in (25), the pre-constant will be independent of ρ in the scaling sense.

Finally, we prove that the scaling choice of $Q = \Theta(\log_2(\rho))$ is order-optimal. First, we show that no scaling order larger than this results in higher throughput. We have observed that by the scaling of $Q = \Theta(\log_2(\rho))$, the scaling of R_1 is as $R_1 = \Theta(\log_2(\rho))$. Choosing a larger scaling than $Q = \Theta(\log_2(\rho))$ for Q, does not help improving the scaling order of R_1. That is because $R_1 = \Theta(\log_2(\rho))$ is the best achievable scaling even in the case of $Q = \infty$. Thus, increasing the scaling order of Q does not help improving the scaling order of R_1. On the other hand, this action will deteriorate the pre-constant factor by increasing the third phase load. Consequently, no scaling larger than $\Theta(\log_2(\rho))$ is the optimum scaling order.

Now, we prove that no scaling smaller than $\Theta(\log_2(\rho))$ is the optimum choice. Suppose we scale Q as $Q = \Theta(f(\rho))$, in which $\lim_{\rho\to\infty} \frac{f(\rho)}{\log_2(\rho)} = 0$. Thus, we have

$$\begin{aligned} g(\rho) &\triangleq \frac{\rho}{1+(\rho b^2+1)2^{-f(\rho)}} \\ &= \frac{\rho}{1+(\rho b^2+1)2^{-\log_2(\rho)\frac{f(\rho)}{\log_2(\rho)}}} \\ &= \frac{\rho}{1+(\rho b^2+1)\rho^{-\frac{f(\rho)}{\log_2(\rho)}}} \\ &\sim \frac{1}{b^2}\rho^{\frac{f(\rho)}{\log_2(\rho)}}. \end{aligned} \tag{44}$$

Therefore, by similar calculations as in (37), and using (44), for the lower bound in (34) we will have

$$\begin{aligned} &\frac{1}{M}\mathbb{E}\left[\log_2 \det\left(\mathbf{I} + \frac{1}{M}\frac{a^2}{b^2}\rho^{\frac{f(\rho)}{\log_2(\rho)}}\mathbf{G}\mathbf{G}^H\right)\right] \\ &\qquad \sim \log_2\left(\frac{a^2}{b^2}\rho^{\frac{f(\rho)}{\log_2(\rho)}}\right) \sim f(\rho). \end{aligned} \tag{45}$$

Similar analysis holds for the upper bound in (34), and thus, we will have $R_1 = \Theta(f(\rho))$. Accordingly, the pre-constant scaling for the hierarchical cooperation scheme is

$$\left(\frac{R_0}{R_1}\right)^{\frac{1}{h}} \frac{R_1}{h\left[4\left(1+\frac{Q}{R_1}\right)\right]^{(h-1)/2}} \sim f(\rho)^{1-\frac{1}{h}}, \tag{46}$$

which has become worse when compared with (43). Thus, the optimal scaling for Q is

$$Q_{opt} = \Theta(\log_2(\rho)). \tag{47}$$

Numerical illustrations

In this section, with the help of numerical illustrations, we investigate the above-discussed results and facts. The system considered in this section is the cooperative MIMO scheme explained in "Establishing cooperative MIMO links" section.

Consider Figure 4. This figure shows the pre-constant factor of the throughput (i.e., $\left(\frac{R_0 R_1^2}{32(R_1+Q)}\right)^{1/2}$ from Equation (9)) as function of Q in $\rho = 25$ dB. The figure contains three curves. The middle curve is the pre-constant based on the relationship (18) used for R_1, which is the exact capacity of a MIMO channel with different noise powers at the receivers. The upper curve is based on the lower bound on R_1 in (34). The lower curve is

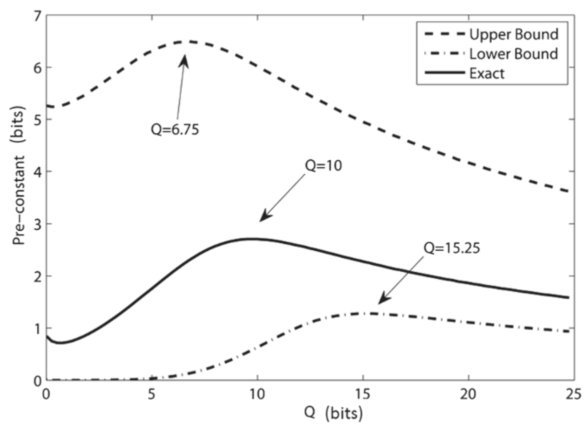

Figure 4 The throughput pre-constant (in bits) as a function of the number of quantization bits *Q*. The upper and lower curves indicate the pre-constant when the upper and lower bound relations for R_1 (in Equation 34) is used. The middle curve is based on the exact capacity in (18).

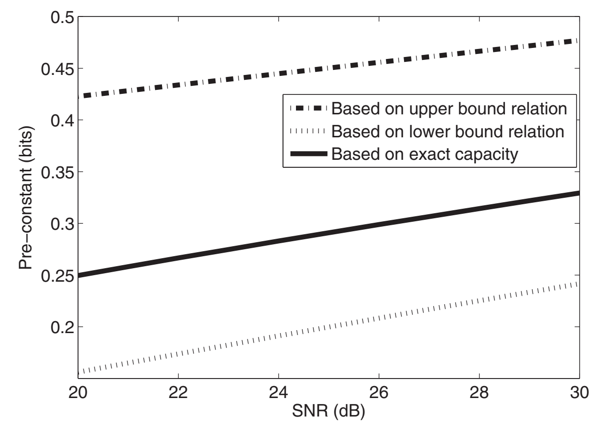

Figure 6 The optimum pre-constant value as a function of SNR level. The upper and lower curves indicate the pre-constant when the upper and lower bound relations for R_1 (in Equation 34) is used. The middle curve is based on the exact capacity relation in (18).

based on the upper bound in (34). This figure shows some interesting facts. The first fact is that there is an optimum choice of Q which maximizes the pre-constant value of the throughput. This Q is the optimal choice for handling the trade-off discussed in "Managing the number of quantization bits" section. The second fact is that the evaluation based on the lower and upper bound relations also shows the existence of an optimal value, but with a slightly different value for the optimum Q. The optimum Q derived based on these curves are 5.75, 10, and 15.25 bits at this specific SNR value.

The next issue is about the behavior of the optimum value of Q at different SNR levels. This issue is investigated in Figure 5. The three curves in this figure are derived in the same way as in Figure 4, with the difference that the upper curve is associated with the lower bound for R_1 and the lower curve is associated with the upper bound. In this figure, we note that all the three curves have almost equal slopes. This provides the answer to the question that why the bounds stated in (34) are very useful to work with, when considering the high SNR regime analysis.

Finally, Figure 6 shows the pre-constant of the network throughput when the optimum number of quantization bits derived in Figure 5 is used. Again, in this figure we note the same slope for the three curves. This fact, again, illustrates the reason for the advantage of using bounds in the high SNR regime analysis.

It should be noted that the analysis in "Analysis in the high SNR regime" section is tailored to the high SNR case. That is because the model we have considered is a dense

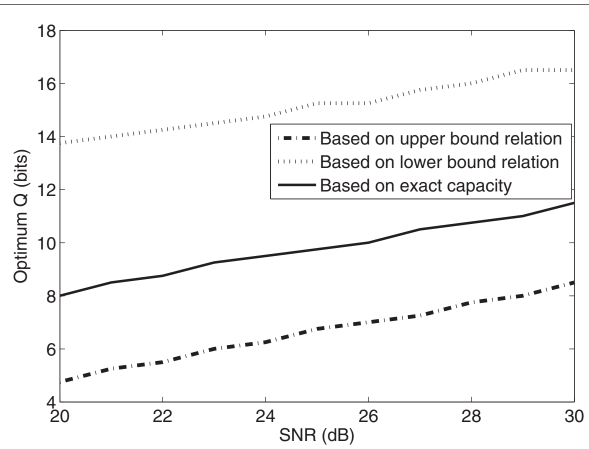

Figure 5 The optimum number of quantization bits as a function of SNR. The upper and lower curves indicate the optimum Q, when the upper and lower bound relations for R_1 (in Equation 34) is used. The middle curve is based on the exact capacity in (18).

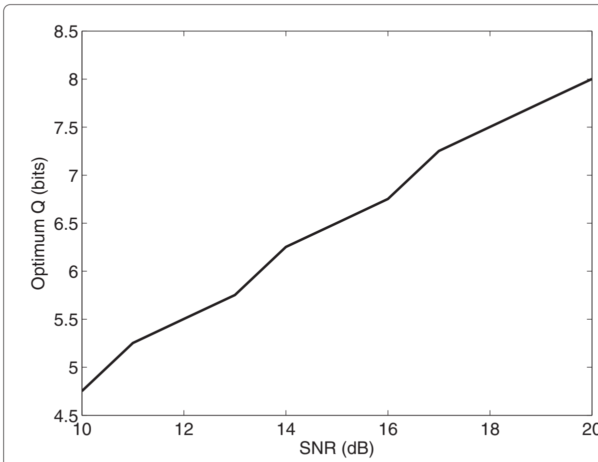

Figure 7 The optimum number of quantization bits as a function of *SNR*.

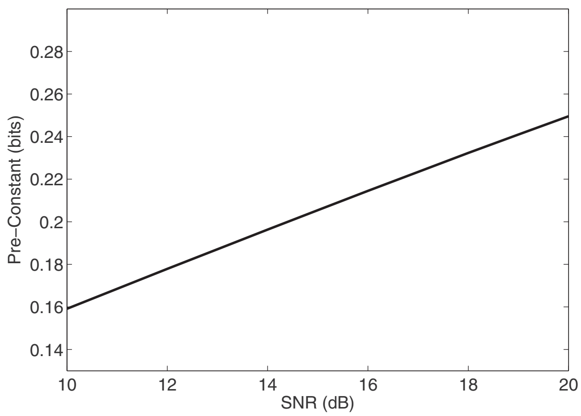

Figure 8 The optimum pre-constant value as a function of *SNR* level.

wireless network which is interference-limited in nature and not power-limited. However, extending the analytical results of this article to non-high SNR situations is an interesting problem for future work. However, we can examine the problem in that situation from the numerical viewpoint which are shown in Figures 7 and 8.

The numerical results presented in this section are very helpful to understand the inherent trade-off in the optimization problem.

Conclusion

In this article, we have analyzed one of the key practical design parameters of cooperative MIMO schemes in ad hoc networks. We have considered the role of the number of quantization bits used at the nodes acting as the virtual MIMO receiver. By increasing this number, we will have a better virtual MIMO link, at the expense of a higher traffic load for cooperation at the receive nodes. We have analyzed this trade-off and have shown there is an optimum number of quantization bits. Also, we have shown that if we do not increase this number as SNR increases, the throughput reaches a saturation value. We have shown that the optimum scaling of the number of quantization bits with SNR is as $\log_2(SNR)$. By this optimal scaling, the pre-constant of the hierarchical cooperation method increases by order $(\log_2(SNR))^{1-1/h}$, in which h is the number of hierarchy stages. The discussions are supported with numerical illustrations.

Endnotes

[a] An earlier version of this article was accepted at the 36th IEEE Conference on Local Computer Networks (LCN 2011) as a poster presentation.
[b] In fact, just the scaling is important which is linear with n. However, to understand why we have put $2nC_1$, and not nC_1, refer to [9,10].
[c] The reason is that in the second phase, the transmit power of the MIMO transmission is proportional to r_{SD}^{α} where r_{SD} is the distance between the centers of the source and destination clusters. On the other hand, the path loss between any two nodes in these two clusters can be approximated by the distance between the centers of the clusters. Thus, these two terms cancel the difference between different cluster pairs and we can treat them equally. Therefore, for the sake of analytical traceability, we have used this approximation and concluded that R_1 is almost the same for all the cluster pairs.
[d] The calculations are the same as [10], with the difference that we have not put $R_1 = R_0$.
[e] The calculations are the same as [12], with the difference that we have not put $R_1 = R_0$.
[f] It should be noted that there is an intrinsic saturation for R_0 due to interference in the 4-TDMA spatial reuse scheme as mentioned earlier (i.e., $\lim_{\rho\to\infty} R_0 < \infty$).

Appendix

In this appendix, we derive Equation (18) of the capacity of quantized MIMO channel with different quantization noise at the receivers. Based on [9], Equations (14) and (15) characterize the quantized MIMO channel model. If we do not use the information in $\mathbf{D}$ for recovering $\mathbf{X}$, and consider the quantization distortion as an extra gaussian noise independent of $\mathbf{X}$, we can assume $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{N})$. Then, by following the conventional capacity calculation of MIMO channels we will have

$$C_{MIMO} = \mathbb{E}\left[\log_2\left(\frac{\det(\mathbf{\Sigma}_N + \mathbf{H}\mathbf{\Sigma}_X\mathbf{H}^H)}{\det \mathbf{\Sigma}_N}\right)\right], \qquad (48)$$

where $\mathbf{\Sigma}_X$ and $\mathbf{\Sigma}_N$ are covariance matrices of $\mathbf{X}$ and $\mathbf{N}$, respectively. Therefore, they are diagonal with the entries $(\mathbf{\Sigma}_X)_{ii} = P\frac{r_{SD}^{\alpha}}{M}$, and $(\mathbf{\Sigma}_N)_{ii} = \sigma^2 + D_i$. From (48) we derive the following

$$C_{MIMO} = \mathbb{E}\left[\log_2\left(\frac{\det(\mathbf{\Sigma}_N + \frac{P}{M}\mathbf{F}\mathbf{F}^H)}{\det \mathbf{\Sigma}_N}\right)\right], \qquad (49)$$

where the entries of the matrix $\mathbf{F}$ are $(\mathbf{F})_{ik} \triangleq \gamma_{ik}\exp(j\theta_{ik})$, and $\gamma_{ik} \triangleq (\frac{r_{SD}}{r_{ki}})^{\alpha/2}$. Finally, by normalizing to σ^2 and defining $\rho \triangleq \frac{P}{\sigma^2}$ we will have the relation stated in (18).

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This study was supported in part by the Iran National Science Foundation under Grant 87041174 and in part by the Iran Telecommunications Research Center (ITRC). Also, this study was supported in part by the EMAITEK funds from the Basque Government.

Author details

[1] Department of Electrical Engineering and Advanced Communication Research Institute (ACRI) Sharif University of Technology, Azadi Ave, Tehran, Iran. [2] OPTIMA Business Area, TECNALIA RESEARCH & INNOVATION, 48170 Zamudio, Spain.

Received: 13 April 2012 Accepted: 10 August 2012
Published: 18 September 2012

References

1. K Jain, J Padhye, V Padmanabhan, L Qiu, in *Paper presented at the Annual International Conference on Mobile Computing and Networking (MobiCom) (ACM)*. Impact of interference on multi-hop wireless network performance (San Diego, California, 2003), pp. 66–80
2. R Madan, D Shah, O Leveque, Product multicommodity flow in wireless networks. IEEE Trans. Inf. Theory. **54**(4), 1460–1476 (2008)
3. C Peraki, SD Servetto, in *Paper presented at the IEEE Information Theory Workshop (ITW)*. Capacity stabilityflows in large-scale random networks (San Antonio, Texas, 2004), pp. 159–164
4. P Gupta, PR Kumar, The capacity of wireless networks. IEEE Trans. Inf. Theory. **42**(2), 388–404 (2000)
5. M Franceschetti, O Dousse, DNC Tse, Closing the gap in the capacity of wireless networks via percolation theory. IEEE Trans. Inf. Theory. **53**(3), 1009–1018 (2007)
6. SR Kulkarni, P Viswanath, A deterministic approach to throughput scaling in wireless networks. IEEE Trans. Inf. Theory. **50**(6), 1041–1049 (2004)
7. A El Gamal, J Mammen, B Prabhakar, D Shah, Optimal throughput-delay scaling in wireless networks—Part: I the fluid model. IEEE Trans. Inf. Theory. **52**(6), 2568–2592 (2006)
8. S Aeron, V Saligrama, Wireless ad hoc networks: strategies and scaling laws for the fixed SNR regime. IEEE Trans. Inf. Theory. **53**(6), 2044–2059 (2007)
9. A Özgür, O Lévêque, D Tse, Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks. IEEE Trans. Inf. Theory. **53**(10), 3549–3572 (2007)
10. J Ghaderi, L-L Xie, X Shen, Hierarchical cooperation in ad hoc networks: optimal clustering and achievable throughput. IEEE Trans. Inf. Theory. **55**(8), 3425–3436 (2009)
11. A Özgür, O Lévêque, Throughput-delay tradeoff for hierarchical cooperation in ad hoc wireless networks. IEEE Trans. Inf. Theory. **56**(3), 1369–1377 (2010)
12. L-L Xie, On information-theoretic scaling laws for wireless networks. IEEE Trans. Inf. Theory (submitted) (2008) (Revised, July 2009). available at http://arxiv.org/abs/0809.1205
13. DE Knuth, Big omicron and big omega and big theta. ACM SIGACT News. **8**, 18–24 (1976)
14. T Cover, JA Thomas, *Elements of Information Theory*. (Wiley Series in Telecommunications and Signal Processing, New York, NY,2006)

doi:10.1186/1687-1499-2012-290
Cite this article as: Shariatpanahi *et al.*: Analysis of quantization effects on the performance of cooperative MIMO techniques in wireless networks. *EURASIP Journal on Wireless Communications and Networking* 2012 **2012**:290.