On the power efficiency of sensory and ad-hoc wireless networks — **Source link** ⬈

A.F. Dana, Babak Hassibi

Institutions: California Institute of Technology

Related papers:

- Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks

- User cooperation diversity. Part I. System description

- Fading relay channels: performance limits and space-time signal design

- Cooperative diversity in wireless networks: Efficient protocols and outage behavior

- User cooperation diversity. Part II. Implementation aspects and performance analysis

On the Power Efficiency of Sensory and *Ad Hoc* Wireless Networks

Amir F. Dana, *Student Member, IEEE*, and Babak Hassibi

Abstract—We consider the power efficiency of a communications channel, i.e., the maximum bit rate that can be achieved per unit power (energy rate). For additive white Gaussian noise (AWGN) channels, it is well known that power efficiency is attained in the low signal-to-noise ratio (SNR) regime where capacity is proportional to the transmit power. In this paper, we first show that for a random sensory wireless network with n users (nodes) placed in a domain of fixed area, with probability converging to one as n grows, the power efficiency scales at least by a factor of $\sqrt{n}$. In other words, each user in a wireless channel with n nodes can support the same communication rate as a single-user system, but by expending only $\frac{1}{\sqrt{n}}$ times the energy. Then we look at a random *ad hoc* network with n relay nodes and r simultaneous transmitter/receiver pairs located in a domain of fixed area. We show that as long as $r \leq \sqrt{n}$, we can achieve a power efficiency that scales by a factor of $\sqrt{n}$. We also give a description of how to achieve these gains.

Index Terms—Capacity, sensor networks, wireless communication systems and networks.

I. INTRODUCTION

IN recent years, there has been great interest in the analysis of wireless networks. Most of the analyses have dealt with the capacity of different types of wireless networks especially the following two types of networks [1], [2]:

1 **Sensory networks:** A sensory network consists of $n + 1$ fixed nodes with a single receiver that collects data/information from the sensor nodes. At any given time, there can be at most one sensory transmitter. All other nodes in the network can be thought of as relay nodes. (See Fig. 1(a).)

2 *Ad hoc* **networks:** At any time, an *ad hoc* network consists of n fixed relay nodes and r fixed simultaneous transmitter/receiver pairs, where $r \leq n$. In this network, relay nodes cooperate for transmission of information from one transmit node to the corresponding receiver node. (See Fig. 1(b).)

It is shown in [1] that for a sensory network, the capacity scales as $\Theta(\log n)$.[1] For *ad hoc* networks, the problem is much

[1]The following notation will be used in this paper. For two functions f, g defined on natural numbers we have $f(n) = O(g(n))$ if $\lim_{n \to \infty} \inf f(n)/g(n) < \infty$, we have $f(n) = \Omega(g(n))$ if $\lim_{n \to \infty} \inf f(n)/g(n) > 0$ and we have $f(n) = o(g(n))$ if $\lim_{n \to \infty} \inf f(n)/g(n) = 0$. Finally, we have $f(n) = \Theta(g(n))$ if $f(n) = \Omega(g(n))$ and $f(n) = O(g(n))$.

more challenging. The groundbreaking work of [2] shows that the capacity grows at least as $\Theta(\sqrt{n})$. Using information-theoretic tools, it is shown in [3]–[5] that under some mild assumptions on the channel model, $\Theta(\sqrt{n})$ is an upper bound on the sum–capacity in the extended wireless networks, i.e., networks where the density of the nodes per area does not increase with the number of nodes. In both sensory and *ad hoc* wireless networks, these results are discouraging from a practical point of view because they suggest that for sensory and *ad hoc* wireless networks, the per-user capacity scales as $\Theta(\frac{\log n}{n})$ and $\Theta(\frac{\sqrt{n}}{n})$, respectively. This represents rewards that rapidly diminish to zero as the number of nodes (users) in the network increases.

Therefore, one interesting problem is to see whether there exists any favorable scalings in *ad hoc* and sensory wireless networks. In other words, are there any scenarios in which it is actually beneficial to form a sensory or *ad hoc* network and obtain increasing gains as the network size grows? Several researchers have looked at this problem from different points of view. In [6], the authors look at a wireless network in which users are mobile (not fixed) and they show that the total capacity of such a network scales like $\Theta(n)$. The work in [7]–[9] also considers the feasibility of wireless networks from a distributed source coding point of view. In this paper, we look at wireless networks from the power consumption point of view.

One of the main concerns in wireless networks especially in sensory networks, is power consumption [10]. Since the source of energy for each user is limited (usually a battery), users in these networks need to use power efficiently. Two major sources of power consumption at each node are the computational power and the transmit power. In this paper, we only consider the power consumption due to transmission and not due to computation. However, we should mention that it is not clear whether at low SNR (where many wireless networks usually operate at) the computation power is negligible compared to the transmit power. Incorporating computation power as well is very interesting and can be a subject of research in itself.

We will show that it is beneficial to form large networks of users in terms of power consumption. We consider sensory and *ad hoc* wireless networks where the users are placed randomly in a domain of fixed area A. We show that users in these networks can support the same rate as a single-user system, but by expending less power. Furthermore, the power that each user needs to expend decreases as we increase the size of the network where the rate of communication is kept fixed. To look at the power efficiency of these networks we will follow the same approach and concept as in [11]. The power efficiency η of a communications channel is defined as
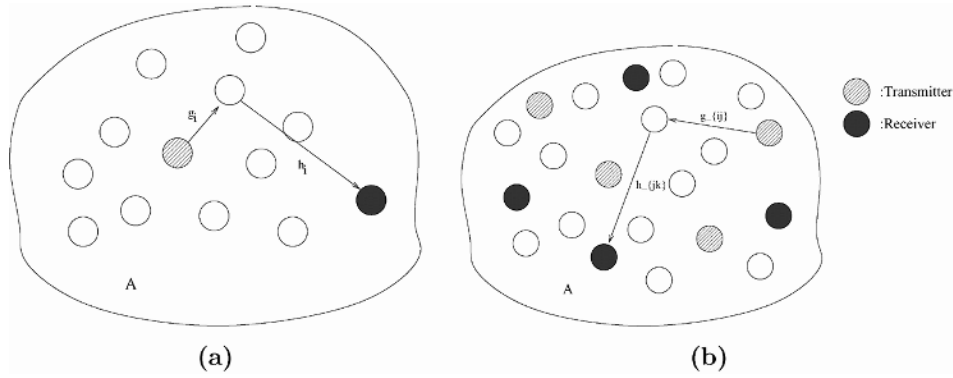
Fig. 1. Sensory and *ad hoc* wireless networks.

the ratio between the capacity (data rate) of the channel and the transmit power (energy rate). For AWGN channels this is given by

$$\eta = \frac{\log(1 + \sigma_s^2/\sigma_n^2)}{\sigma_s^2} \tag{1}$$

where σ_s^2 represents the transmit power and σ_n^2 represents the noise power. Clearly, for fixed σ_n^2 as $\sigma_s^2 \to \infty$, the value of η approaches zero, meaning that we are highly power inefficient at high SNR. On the other hand, we are power efficient at low SNR and, in fact

$$\max_{\sigma_s^2} \eta = \lim_{\sigma_s^2 \to 0} \eta = \frac{\log e}{\sigma_n^2}. \tag{2}$$

This implies that, at low SNR, capacity is proportional to the transmit power. In [11], the power efficiency (or capacity per unit cost as the author defines it) of several other communication systems is computed.

In this paper, we will find a lower bound for the power efficiency of sensory and *ad hoc* wireless networks formed in a domain of fixed area. For this, we will propose a protocol for communication among the nodes. The key idea used in the protocol is to exploit features of wireless networks and operate the network at low SNR (thereby avoiding the logarithmic scaling of the capacity). The main features that distinguish wireless networks from wireline networks are path loss, fading, and interference. Path loss has been exploited in cellular networks. Fading also is exploited in multiuser systems by scheduling transmissions when a user has favorable channel conditions [12][13]. However, most current approaches avoid interference in the network. For instance, in [2] most of the emphasis is on interference avoidance and the construction of a multihop network. In our protocol, we will exploit the interference and fading inherent in any wireless network for achieving good power efficiency. Also, the protocol proposed in this paper is a double-hop protocol. Although it is thought that the power efficiency of multihop networks is better than that of double-hop networks it can be shown that if the nodes are placed in a domain of fixed area this is not true. A similar observation is made in [14]. The authors have observed that the most energy-efficient protocol to use depends on the network topology and the radio parameters of the system.

We have shown in [15] that for sensory and *ad hoc* wireless networks for which the channel coefficients between users can be modeled by independent zero-mean, unit variance and bounded fourth-order moment random variables, the power efficiency scales at least as $\Theta(\sqrt{n})$. However, the model used for channel coefficients in this paper is more general. We will see that even with this general model we are still able to achieve a power efficiency that scales favorably as the size of the network grows. The net result is that under some mild assumptions on the channel coefficients that will be mentioned in Section II-B, with high probability the power efficiency of a random network, i.e., the data rate per energy rate, scales as $\Theta(\sqrt{n})$ for each user.

This paper is organized as follows. Section II describes the system model and assumptions and presents the statement of the problem considered in this paper. In Section III, we will compute the power efficiency of a multiple-antenna communication system for comparison. In Section IV, we consider the power efficiency of sensory wireless networks. We describe the proposed "Listen and Transmit" protocol for achieving scalable power efficiency for sensory networks. In Section V, power efficiency of *ad hoc* wireless networks is considered and analyzed. We first present a generalization of the Listen and Transmit protocol for *ad hoc* networks and then optimally allocate powers to achieve a scalable power efficiency for the network. At the end of that section, we will compare the performance of our protocol with an interference suppression scheme that requires complete knowledge of the channel. Conclusions and proposals of further work are provided in Section VI.

II. Notation and System Model

A. Notation and Definitions

Throughout this paper matrices and vectors are denoted by boldface characters. $\mathrm{tr}(\boldsymbol{A})$, $\lambda_{\max}(\boldsymbol{A})$, and $\lambda_{\min}(\boldsymbol{A})$ denote the trace, the maximum eigenvalue, and the minimum eigenvalue of a square Hermitian matrix $\boldsymbol{A}$. The superscript $*$ denotes conjugate transposition for matrices and complex conjugate for scalars. Complex conjugation for matrices is shown by using bar. Transposition is also denoted by superscript T. $\boldsymbol{A}^*$, $\boldsymbol{A}^T$, and $\bar{\boldsymbol{A}}$ are the conjugate transpose, transpose, and conjugate of the matrix $\boldsymbol{A}$, and α^* is the complex conjugate of the scalar α. $\boldsymbol{I}_r$ is the $r \times r$ identity matrix. For a matrix $\boldsymbol{A}$, $\mathrm{vec}(\boldsymbol{A})$ denotes

the vector obtained from stacking all the columns of $\boldsymbol{A}$, one on top of another. For a vector $\boldsymbol{g} = (g_1, \ldots, g_n)$, $\mathrm{diag}(\boldsymbol{g})$ denotes the $n \times n$ diagonal matrix with ith diagonal element equal to g_i. We may also write $\mathrm{diag}(\boldsymbol{g})$ as $\mathrm{diag}(g_1, \ldots, g_n)$. Finally, $\|\boldsymbol{g}\|$ denotes the Euclidean norm of vector $\boldsymbol{g}$.

In this paper, we will be studying random wireless networks. Thus, we will consider a fixed area A and will randomly select points in A to form the nodes of the network (either as transmitters, receivers, or relay nodes). Since the network is wireless, the connections between any two nodes will be subject to fading. Thus, the randomness in the network will be due to two sources: the random choice of points in A and the random fading between the connections. When we fix the position of the nodes, we denote the expectation over channel fading by $\mathrm{E}_{\mathrm{f}}[.]$. The expectation over the location of some set of the nodes, say P, in a random network is denoted by $\mathrm{E}_P[.]$ which from now on we shall call the spatial average. For instance, the spatial average of the mean value of the channel coefficient between node i and node j, h_{ij}, over the position of node i while node j is fixed, can be written as $\mathrm{E}_{\{i|j\}}[\mathrm{E}_{\mathrm{f}}[h_{ij}]]$. The expectation over the location of all the nodes is denoted by $\mathrm{E}_{\mathrm{loc}}$ and whenever $\mathrm{E}[.]$ is used without any subscript, expectation over both fading and the location of the nodes is implied. Channel coefficients are denoted by h, g, or c depending on the context. Usually c is used as a generic channel between two arbitrary points in the domain.

B. System Model and Problem Statement

Sensory Networks: As mentioned earlier, by a sensory wireless network we mean one with n relay nodes and a single transmitter/receiver pair (see Fig. 1). We assume that the nodes are placed randomly and independently according to some distribution function (not necessarily uniform) in a domain of fixed area, say A. We denote the channel coefficient from the transmitter to the relay node i by g_i and the channel coefficient from the relay node i to the receiver by h_i. We assume that, averaged over the fading, different channels are independent. Furthermore, we assume that each node i, knows only its local connections h_i, g_i, but not the other connections in the network.

Ad Hoc Networks: As mentioned earlier, for *ad hoc* networks we assume that at any time there are n relay nodes and at most r simultaneous transmit/receive pairs in the network. The nodes are placed randomly and independently according to some distribution function (not necessarily uniform) in a domain of fixed area, say A. The channel coefficient from transmitter i $(i = 1, \ldots, r)$ to relay node j $(j = 1, \ldots, n)$ is denoted by g_{ij} and from relay node j to receive node k $(k = 1, \ldots, r)$ is denoted by h_{jk}. Similar to the sensory case, we assume that, averaged over the fading, distinct channels are independent. Furthermore, if we fix the location of the transmitters and the receivers, and randomly choose relay nodes j and j', the channel coefficients $\{g_{ij}, h_{jk}\}$ and $\{g_{ij'}, h_{j'k}\}$ are independent for all i and k. As with the sensory case, we assume that all the relay nodes know their local connections, but not the remaining connections in the network. In other words, node j knows all the connections $\{g_{ij}, i = 1, \ldots, r\}$ and $\{h_{jk}, k = 1, \ldots, r\}$.

Additional Assumptions for Ad Hoc Networks: For *ad hoc* networks, we have a few more assumptions. Thus, denote the
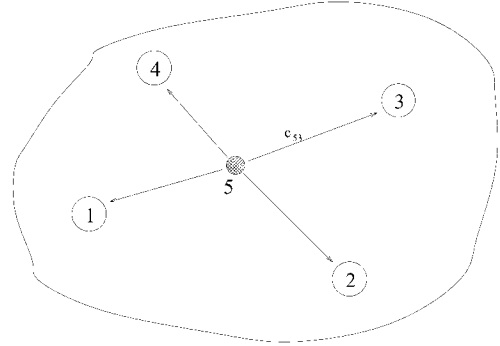

Fig. 2.   Condition on the channel coefficients.

channel coefficient between two points x_i and x_j by c_{ij}. With this notation, we have the following additional assumptions.

A. $\mathrm{E}_{\{x_5 \mid x_2\}}[\mathrm{E}_{\mathrm{f}}[c_{52}]] = 0, \qquad \forall x_2 \in A$

B. $\mathrm{E}_{\{x_5 | x_1, x_2\}}[\mathrm{E}_{\mathrm{f}}[c_{51} c_{52}^*]] = 0, \qquad \forall x_1 \neq x_2 \in A$

C. $\mathrm{E}_{\{x_5 | x_1, x_2, x_3, x_4\}}[\mathrm{E}_{\mathrm{f}}[c_{51}^* c_{52}^* c_{53} c_{54}]] = 0,$
$\qquad \forall x_1, x_2, x_3, x_4 \in A$
$\qquad$ at least 3 of the x_i are distinct.

Note that the above conditions are clearly met if the fading is zero mean. In general, however, there may be line-of-sight between different nodes in the network and the fading may be nonzero mean. The above conditions are more general and do not require zero mean fading. The first assumption says that the spatial average of the mean of a channel coefficient between a random point and a fixed point is zero. The second assumption is that the channel coefficients between one random point, x_5, and two different points, x_1, x_2 are uncorrelated when averaged over both the fading and the point placement of x_5. In other words, although the channels c_{51} and c_{52}, given that x_1 and x_2 are fixed, are not independent and may be correlated but the spatial average of the correlation between these two channels is zero. The last condition also says that the expectation of the product of the channel coefficients between one random point and four fixed points averaged over the location of the random point is zero (see Fig. 2). These assumptions appear to be reasonable, especially if we assume that the environment is rich in scattering. In this paper, we obtain two achievable bounds for the power efficiency of *ad hoc* networks. The first bound (Theorem 2) relies only on the first and second assumptions and the second bound (Theorem 3) requires the last assumption as well.

Power Assumptions: In the sensory network, we assume that the transmit power is p. For *ad hoc* networks we assume that *all* the transmitters transmit with the same power p. In both cases, we will assume that the relay nodes transmit with identical power σ_r^2. The noise introduced in every reception is an additive white circularly symmetric Gaussian noise with zero mean and variance σ_n^2 which is denoted by $\mathcal{CN}(0, \sigma_n^2)$.

Path Loss: In this paper, we will not be concerned with explicit path loss models. The main reason is that, since we consider a fixed domain A, the only characteristics of the path loss that enter our analysis are the second- and fourth-order moments of the channel. In fact, a strength of our results is that the asymptotics are not sensitive to the path-loss model. (The model and
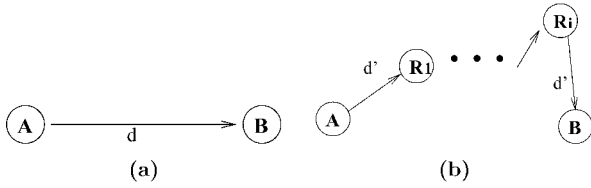
Fig. 3. Single-hop versus multihop.

the geometry of the domain A affect the constants but not the scaling behavior.) We further discuss path-loss models when comparing single-hop and multihop systems below.

Single-Hop Versus Multihop Communication: In this paper, we propose a protocol for achieving a power efficiency that scales with the number of nodes n. The communication model that we are using is a double-hop (transmit and relay) communication protocol. In this protocol, which will be explained in detail later, the communication is done in two intervals. In the first interval, the transmit nodes send their data signal. In the second interval, relay nodes send a signal based on what they have received in the first interval.

Typically, in order to increase power efficiency in wireless networks, one must move toward a multihop system so as to avoid long hops (which are subject to severe path loss) [16], [17]. While this is certainly true for networks that grow in physical size as the number of nodes increases (thereby increasing the size of the hops), it is not true for networks in which the physical domain is *fixed* while the number of nodes increases.[2] In this case, there is nothing to be gained by using multihop schemes in which the number of hops scales with the number of nodes in the network and the length of the hops becomes shorter and shorter as the number of nodes increases. To make this more explicit, we use the following qualitative argument. Suppose that n nodes are located in a domain of fixed area A. Consider two nodes (users) of distance d, which want to communicate with each other (Fig. 3(a)). Assume that the channel is AWGN and that the power loss between any two points is a decreasing function of their distance d and is denoted by $f(d)$. In this case, the relation between the transmitted signal s from A and the received signal y at B is

$$y = \sqrt{f(d)}s + v$$

The capacity is clearly $\log(1 + f(d)\frac{p}{\sigma_n^2})$ and the power efficiency achieved at low SNR is

$$\eta = \frac{f(d)}{\sigma_n^2}.$$

Assume now that we employ a multihop scheme to communicate between A and B where each node relays to its nearest neighbor (Fig. 3(b)). Since we have n nodes, the distance to a nearest neighbor will be $d' = O(\frac{1}{\sqrt{n}})d$ and the number of

hops will be of $\Omega(\sqrt{n})$. Here, each relay will communicate at rate $\log(1 + f(d')\frac{\sigma_r^2}{\sigma_n^2})$ and since the total transmit power is $\Omega(\sqrt{n})\sigma_r^2$, the power efficiency achieved at low SNR will be

$$\eta' = \frac{f(d')}{\sigma_n^2 \Omega(\sqrt{n})} = \frac{f\left(dO\left(\frac{1}{\sqrt{n}}\right)\right)}{\sigma_n^2 \Omega(\sqrt{n})}$$

For any reasonable path-loss model $\lim_{d \to 0} f(d) = \text{constant}$.[3] Therefore, the power efficiency of the multihop system scales like $O(\frac{1}{\sqrt{n}})$ as n increases. This says that for a fixed-size network increasing the number of hops in fact reduces the power efficiency.

Remark: Note that if the size of the domain also increases with n, then d will also increase. In this case, since

$$\frac{\eta'}{\eta} = \frac{f\left(dO\left(\frac{1}{\sqrt{n}}\right)\right)}{f(d)\Omega(\sqrt{n})}$$

depending on the path-loss model $f(.)$ and *how* d scales with n, it may be more power efficient to use multihop. □

Channel Knowledge and Synchronicity: As mentioned earlier, we have assumed that the nodes have knowledge of their local connections. This is a much more reasonable assumption than the nodes knowing the entire network. However, it does require that the network remain relatively stationary in time, so that the local connections can be learned via the transmission of pilot symbols, etc. Furthermore, we assume a synchronous system. In other words, all the transmissions and receptions are synchronized. Later, we will argue that the system performance is not very sensitive to timing errors and lack of perfect synchronicity.

III. AN EXAMPLE: MULTIPLE-ANTENNA SYSTEMS

In order to obtain some insight into how the power efficiency of a sensory or *ad hoc* wireless network might scale, it is useful to look at the example of a multiple-antenna system. For more details see [18].

Consider an n transmit single receive multiple-antenna channel, described by the channel vector

$$H = \begin{bmatrix} h_1 & h_2 & \dots & h_n \end{bmatrix}$$

where h_i denotes the channel coefficient from the ith transmitter to the receiver. (Assume that the channel coefficients are zero-mean and unit variance and have fourth-order moment κ.) Two cases can be envisioned.

- The channel matrix is *known* to the transmitter: In this case, the optimal scheme is for beam forming. Thus, if each antenna transmits with power p the power efficiency becomes

$$\eta = \frac{\mathrm{E} \log\left(1 + \frac{p}{\sigma_n^2}(\sum_{i=1}^n |h_i|^2)^2\right)}{np}.$$

---

[2]If the area of the network increases with the number of nodes, a combination of multihop routing and the Listen and Transmit protocol described here is necessary to achieve a power efficiency that scales with the number of nodes.

[3]Note that the common power law function used in literature, $f(d) = \frac{1}{d^n}$ $n \geq 2$, does not satisfy this property since this model is only valid for far field approximation.

This is maximized when $p \rightarrow 0$, which yields

$$
\begin{aligned}
\eta &= \frac{\log e}{n\sigma_n^2} \mathrm{E} \left(\sum_{i=1}^{n} |h_i|^2 \right)^2 \\
&= \frac{\log e}{n\sigma_n^2} (n\kappa + n(n-1)) = \Theta(n). \quad (3)
\end{aligned}
$$

- The channel matrix is *unknown* to the transmitter. In this case, beam forming cannot be done. However, the capacity is known from [19] and so the power efficiency becomes

$$
\eta = \frac{\mathrm{E} \log \left(1 + \frac{p}{\sigma_n^2} \sum_{i=1}^{n} |h_i|^2 \right)}{np}.
$$

And so, at low SNR

$$
\eta = \frac{\log e}{n\sigma_n^2} \mathrm{E} \left(\sum_{i=1}^{n} |h_i|^2 \right) = \frac{\log e}{n\sigma_n^2} n = \Theta(1). \quad (4)
$$

What distinguishes an $n \times 1$ multiple-antenna system from an n-node sensory network is that the n antenna elements are allowed to cooperate, but the n nodes in a sensory network are not. What the above result says is that when the nodes are allowed to cooperate *and* the nodes know the channel coefficients the power efficiency scales as $\Theta(n)$. However, even if the nodes are allowed to cooperate, as long as they do not know the channel coefficients, the power efficiency does not improve over $\Theta(1)$. But what about a sensory network, where the nodes are not allowed to cooperate but know the local channel coefficients? Moreover, what about *ad hoc* networks? These are the questions we shall address.

IV. SENSORY NETWORKS

We begin by describing a simple protocol that achieves a power efficiency of $\Theta(\sqrt{n})$ for random sensory and, as we shall see in the next section, with some modification for *ad hoc* wireless networks. As mentioned earlier, the protocol assumes synchronous transmission and receptions, as well as local channel knowledge at the nodes.

A. Listen and Transmit Protocol

Consider a random sensory network with n relay nodes and one transmitter/receiver pair. We are interested in a probabilistic bound for the achievable power efficiency in this network, i.e., a bound that with high probability is achievable for a random network in the domain. We begin by explaining the protocol that achieves power efficiency of $\Theta(\sqrt{n})$ for sensory wireless networks. In this, so-called Listen and Transmit protocol, communication is done in two intervals:

1. **Listen interval:** In this interval, the transmitter sends the data and the relay nodes only listen. Relay node i receives

$$
r_i = g_i s + v_i, \qquad i = 1, 2, \ldots, n \quad (5)
$$

where v_i is $\mathcal{CN}(0, \sigma_n^2)$.

2. **Transmit interval:** Each node, using its knowledge of the local connections, transmits a scaled version of the signal it has received in the first interval

$$
t_i = \frac{\sigma_r g_i^* h_i^*}{|g_i||h_i|\sqrt{|g_i|^2 p + \sigma_n^2}} r_i, \qquad i = 1, 2, \ldots, n. \quad (6)
$$

The scalar is chosen so that the relay node power is σ_r^2 and so that the signal parts coherently add at the receiver.

This protocol is similar to the protocol proposed in [20]. In [20], the relay nodes transmit the exact signals they have received, scaled to meet the power constraint. In the Listen and Transmit protocol, the channel coefficients can be complex. Therefore, the relay nodes change the phase of their received signal appropriately so that the signal parts of the received signal (at the receiver) add up coherently. The received signal at the receiver is

$$
\begin{aligned}
y &= \sum_{i=1}^{n} h_i t_i + w \\
&= \underbrace{\sum_{i=1}^{n} \frac{\sigma_r |g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}}}_{\alpha} s + \underbrace{\sum_{i=1}^{n} \frac{\sigma_r |h_i| g_i^*}{|g_i|\sqrt{|g_i|^2 p + \sigma_n^2}} v_i + w}_{w'}
\end{aligned}
$$

$$(7)$$

where w is $\mathcal{CN}(0, \sigma_n^2)$. From (7) it is clear that the signal part from different relay nodes adds up coherently but the noise part does not. In this sense, the Listen and Transmit protocol can be regarded as performing *distributed beamforming*.

B. Finding a Lower Bound

We break α, defined in (7), into $\bar{\alpha}$ and $\tilde{\alpha} = \alpha - \bar{\alpha}$, where $\bar{\alpha} = \mathrm{E}_f[\alpha]$ and $\mathrm{E}_f[\tilde{\alpha}] = 0$. Now if we rewrite (7) as $y = \bar{\alpha}s + \tilde{\alpha}s + w'$, then as shown in [21], the capacity of this system can be lower-bounded by the capacity of the AWGN channel $y = \bar{\alpha}s + w''$, where w'' is a Gaussian noise with variance equal to the variance of $\tilde{\alpha}s + w'$. (In this analysis, we assume that the receiver is provided with the mean of α.) Therefore, the capacity of the system in (7) may be lower-bounded by

$$
C \geq \frac{1}{2} \log \left(1 + \frac{|\bar{\alpha}|^2 p}{\mathrm{Var}_f[w'] + \mathrm{Var}_f[\tilde{\alpha}]p} \right). \quad (8)
$$

Note that the $\frac{1}{2}$ in front of the logarithmic term comes from the fact that the transmitter transmits half of the time. By substituting $\bar{\alpha}$, $\mathrm{Var}_f[w']$, and $\mathrm{Var}_f[\tilde{\alpha}]$ in (8) with

$$
\bar{\alpha} = \mathrm{E}_f[\alpha] = \sigma_r \sum_{i=1}^{n} \mathrm{E}_f \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}}
$$

$$
\mathrm{Var}_f[\tilde{\alpha}] = \mathrm{Var}_f[\alpha] = \sigma_r^2 \sum_{i=1}^{n} \mathrm{E}_f \frac{|g_i|^2 |h_i|^2}{|g_i|^2 p + \sigma_n^2}
$$

$$
- \sigma_r^2 \sum_{i=1}^{n} \left(\mathrm{E}_f \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}} \right)^2
$$

$$
\mathrm{Var}_f[w'] = \sigma_n^2 \left(1 + \sigma_r^2 \sum_{i=1}^{n} \mathrm{E}_f \frac{|h_i|^2}{|g_i|^2 p + \sigma_n^2} \right)
$$

and rearranging the terms, we get (9) at the bottom of the page. Define $c_i = \mathrm{E}_f \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}}$ and $b_i = \mathrm{E}_f |h_i|^2$. By ignoring the negative term in the denominator we can rewrite (9) as

$$C \geq \frac{1}{2} \log \left(1 + \frac{n^2 \sigma_r^2 p \left(\frac{\sum_{i=1}^n c_i}{n} \right)^2}{\sigma_n^2 + n\sigma_r^2 \left(\frac{\sum_{i=1}^n b_i}{n} \right)} \right). \qquad (10)$$

The above lower bound holds for every fixed network. For a random network, the capacity, the power efficiency, and the above lower bound are random variables depending on the placement of the nodes in the network. Since the nodes are placed independently and according to the same distribution on the available area, the b_i's and c_i's are independent and identically distributed (i.i.d.) random variables for different relay nodes (i.e., different i's). Therefore, denoting the transmitter and receiver location by $\{t, r\}$, we have

$$\mathrm{E}_{\mathrm{loc}} \left(\sum_{i=1}^n b_i \right) = \mathrm{E}_{\{t,r\}} \sum_{i=1}^n \mathrm{E}_{\{i|t,r\}}[b_i] = n\mathrm{E}_{\mathrm{loc}}[b]$$

$$\mathrm{Var}_{\mathrm{loc}} \left(\sum_{i=1}^n b_i \right) = \mathrm{E}_{\{t,r\}} \sum_{i=1}^n \mathrm{E}_{\{i|t,r\}} \left([b_i^2] - [b_i]^2 \right)$$
$$= n\mathrm{Var}_{\mathrm{loc}}[b]$$

$$\mathrm{E}_{\mathrm{loc}} \left(\sum_{i=1}^n c_i \right) = \mathrm{E}_{\{t,r\}} \sum_{i=1}^n \mathrm{E}_{\{i|t,r\}}[c_i] = n\mathrm{E}_{\mathrm{loc}}[c]$$

$$\mathrm{Var}_{\mathrm{loc}} \left(\sum_{i=1}^n c_i \right) = \mathrm{E}_{\{t,r\}} \sum_{i=1}^n \mathrm{E}_{\{i|t,r\}} \left([c_i^2] - [c_i]^2 \right)$$
$$= n\mathrm{Var}_{\mathrm{loc}}[c],$$

where $c = \mathrm{E}_f \frac{|g||h|}{\sqrt{|g|^2 p + \sigma_n^2}}$ and $b = \mathrm{E}_f |h|^2$ are the random variables depending on the channel coefficients between one random point and two other random points. Now for any $\epsilon > 0$, using Chebyshev's inequality and the union bound on the probability of the events, we have

$$\mathrm{P} \left\{ \left| \frac{\sum_{i=1}^n c_i}{n} - \mathrm{E}_{\mathrm{loc}} c \right| \leq \epsilon, \left| \frac{\sum_{i=1}^n b_i}{n} - \mathrm{E}_{\mathrm{loc}} b \right| \leq \epsilon \right\}$$
$$\geq 1 - \frac{\mathrm{Var}_{\mathrm{loc}}[b] + \mathrm{E}_{\mathrm{loc}}[c^2]}{n\epsilon^2}. \quad (11)$$

The inequality of (11) shows that as $n \to \infty$ the quantities

$$\frac{\sum_{i=1}^n c_i}{n} \quad \text{and} \quad \frac{\sum_{i=1}^n b_i}{n}$$

behave like their spatial averages. This implies that with high probability C in (10) is bounded by

$$\frac{1}{2} \log \left(1 + \frac{n^2 \sigma_r^2 p \mathrm{E}_{\mathrm{loc}}^2[c]}{\sigma_n^2 + n\sigma_r^2 \mathrm{E}_{\mathrm{loc}}[b]} \right).$$

Remark: Note that $\mathrm{E}_{\mathrm{loc}}[b]$ and $\mathrm{E}_{\mathrm{loc}}[c]$ depend only on the domain A, on the fading characteristics, and on the distribution of the points and they do not depend on n. Thus, for fixed p and σ_r^2, as $n \to \infty$ the lower bound on capacity with high probability behaves like

$$\frac{1}{2} \log n + O(1).$$

This is the same asymptotic growth obtained for Gaussian relay channels in [1]. Thus, we conclude that the Listen and Transmit protocol (i.e., distributed beamforming) achieves the optimal asymptotic capacity growth. We, of course, are not primarily interested in capacity but rather in power efficiency. $\qquad \square$

Now we will focus on how to optimally allocate the powers (p and σ_r^2 as a function of n) to maximize the power efficiency. As mentioned earlier, $\mathrm{E}_{\mathrm{loc}}[b]$ and $\mathrm{Var}_{\mathrm{loc}}[b]$ do not depend n. Using the Taylor series expansion of $\mathrm{E}_{\mathrm{loc}}[c]$ and $\mathrm{E}_{\mathrm{loc}}[c^2]$ in p, we have

$$\mathrm{E}_{\mathrm{loc}}[c] = \frac{\mathrm{E}_{\mathrm{loc}}\mathrm{E}|h||g|}{\sigma_n} - \frac{\mathrm{E}_{\mathrm{loc}}\mathrm{E}|h||g|^3}{\sigma_n^3} p + o(p)$$

$$\mathrm{E}_{\mathrm{loc}}[c^2] = \frac{\mathrm{E}_{\mathrm{loc}}\mathrm{E}|h|^2|g|^2}{\sigma_n^2} - \frac{\mathrm{E}_{\mathrm{loc}}\mathrm{E}|h|^2|g|^4}{\sigma_n^4} p + o(p). \quad (12)$$

Note that $|h|$ and $|g|$ do not depend on n, so the only dependence of $\mathrm{E}_{\mathrm{loc}}[c]$ and $\mathrm{E}_{\mathrm{loc}}[c^2]$ on n can be through p. Since the total power consumed in the network is $\frac{1}{2}(p + n\sigma_r^2)$ ($\frac{1}{2}$ comes from the fact that each node is sending only half of the time), the power efficiency is

$$\eta = \frac{2C}{p + n\sigma_r^2}.$$

From (11) and (10), we can find a probabilistic lower bound for the power efficiency of the network. In other words, for a random placement of the nodes in the domain we have

$$\mathrm{P} \left\{ \eta = \frac{2C}{n\sigma_r^2 + p} \right.$$
$$\geq \frac{1}{p + n\sigma_r^2} \log \left(1 + \frac{n^2 \sigma_r^2 p (\mathrm{E}_{\mathrm{loc}}[c] - \epsilon)^2}{\sigma_n^2 + n\sigma_r^2 (\mathrm{E}_{\mathrm{loc}}[b] + \epsilon)} \right) \right\}$$
$$\geq 1 - \frac{\mathrm{Var}_{\mathrm{loc}}[b] + \mathrm{E}_{\mathrm{loc}}[c^2]}{n\epsilon^2}.$$

$$C \geq \frac{1}{2} \log \left(1 + \frac{\sigma_r^2 p \left(\sum_{i=1}^n \mathrm{E}_f \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}} \right)^2}{\sigma_n^2 + \sigma_r^2 \left(\sum_{i=1}^n \mathrm{E}_f |h_i|^2 - \sum_{i=1}^n p \left(\mathrm{E}_f \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}} \right)^2 \right)} \right). \qquad (9)$$

By choosing the transmit power, $p = \frac{1}{\sqrt{n}}$ and the relay node power, $\sigma_r^2 = \frac{1}{n\sqrt{n}}$ we have from (12) and the above equation that

$$P\left\{\eta \geq \frac{\sqrt{n}}{2}\log\left(1 + \frac{\left(\beta_1 + o\left(\frac{1}{\sqrt{n}}\right) - \epsilon\right)^2}{\sigma_n^2 + \frac{E_{loc}[b]+\epsilon}{\sqrt{n}}}\right)\right\}$$
$$\geq 1 - \frac{Var_{loc}[b] + \beta_2 + o\left(\frac{1}{\sqrt{n}}\right)}{n\epsilon^2}$$

where

$$\beta_1 = \frac{E_{loc}E|h||g|}{\sigma_n} \quad \text{and} \quad \beta_2 = \frac{E_{loc}E|h|^2|g|^2}{\sigma_n^2}$$

are some constants independent of n. Close inspection of the above inequality reveals that the term in the logarithm is of order one. This says that the capacity achieved with the Listen and Transmit protocol is of $\Theta(1)$. Moreover, there exist a constant $K_1 \geq 0$ such that

$$P\{\eta \geq K_1\sqrt{n}\} \geq 1 - \frac{Var_{loc}[b] + \beta_2 + o\left(\frac{1}{\sqrt{n}}\right)}{n\epsilon^2}. \quad (13)$$

From the preceding inequality, we can see that for a random placement of the nodes in A, with a high probability that approaches one as the number of nodes increases, we can achieve a power efficiency that grows like $\sqrt{n}$. Also, the rate achieved is of $\Theta(1)$. The choice of transmit and relay node power in this case is $p = n\sigma_r^2 = \frac{1}{\sqrt{n}}$.

C. Finding an Upper Bound

We can also find an upper bound on the achievable rates using the Listen and Transmit protocol. For this, we consider the case where the receiver in (7) knows $\{h_i, g_i, v_i\}$ for $i = 1, 2, \ldots, n$. In this case we have

$$C = \sup_{p(s)} \mathcal{I}(y; s) \leq \sup_{p(s)} \mathcal{I}(s; y|\{h_i, g_i, v_i\}) \quad (14)$$

where $\mathcal{I}(y; s)$ is the mutual information between y and s. Now if the receiver knows the channel coefficients and the v_i's then the system in (7) becomes an AWGN channel and therefore,

$$\sup_{p(s)} \mathcal{I}(s; y|\{h_i, g_i\}) = \frac{1}{2}E_{f\{h_i, g_i\}}\left[\log\left(1 + \frac{|\alpha|^2 p}{\sigma_n^2}\right)\right]. \quad (15)$$

The v_i's do not contribute to the noise power in the denominator of (15) since the receiver has complete knowledge of them and can cancel out their effect. Combining (14) and (15), and using the convexity of the $\log$ function, we may write

$$C \leq \frac{1}{2}\log\left(1 + \frac{E_f|\alpha|^2 p}{\sigma_n^2}\right). \quad (16)$$

Substituting the value of α from (7) in the above equation gives

$$C \leq \frac{1}{2}\log\left(1 + \frac{\sigma_r^2 p}{\sigma_n^2}E_f\left(\sum_{i=1}^{n}\frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}}\right)^2\right)$$

where again the expectation is taken over the fading of the channels for a fixed placement of the nodes. Without loss of generality we can ignore the $|g_i|^2 p$ term in denominator and rewrite the above equation as

$$C \leq \frac{1}{2}\log\left(1 + \frac{\sigma_r^2 p}{\sigma_n^4}E_f\left(\sum_{i=1}^{n}|g_i||h_i|\right)^2\right). \quad (17)$$

Since, averaged over the fading, the $|h_i||g_i|$'s are independent for different i's, we have

$$E_f\left(\sum_{i=1}^{n}|g_i||h_i|\right)^2 \leq \left(\sum_{i=1}^{n}E_f|g_i||h_i|\right)^2 + \sum_{i=1}^{n}E_f|h_i|^2|g_i|^2$$

and therefore,

$$C \leq \frac{1}{2}\log\left(1 + \frac{n^2\sigma_r^2 p}{\sigma_n^4}\left(\frac{\sum_{i=1}^{n}E_f|g_i||h_i|}{n}\right)^2\right.$$
$$\left. + \frac{n\sigma_r^2 p}{\sigma_n^4}\frac{\sum_{i=1}^{n}E_f|h_i|^2|g_i|^2}{n}\right). \quad (18)$$

Given that the location of the transmitter and the receiver is fixed, $E_f|h_i||g_i|$ and $E_f|h_i|^2|g_i|^2$ for all i depend only on the placement of the relay nodes and are i.i.d. random variables. Thus, according to the law of large numbers, their average converges to their statistical mean. More specifically, for any $\epsilon > 0$, using Chebyshev's inequality and the union bound on the probability of the events we have

$$P\left\{\left|\frac{\sum_{i=1}^{n}E_f|h_i||g_i|}{n} - \kappa\right|\right.$$
$$\left. \geq \epsilon, \left|\frac{\sum_{i=1}^{n}E_f|h_i|^2|g_i|^2}{n} - \mu\right| \geq \epsilon\right\}$$
$$\geq 1 - \frac{Var_{loc}E_f|h||g| + Var_{loc}E_f|h|^2|g|^2}{n\epsilon^2} \quad (19)$$

where $\kappa = E|h||g|$ and $\mu = E|h|^2|g|^2$. Combining (19) and (18) gives

$$P\left\{\eta \leq \frac{1}{p + n\sigma_r^2}\log\left(1 + \frac{n\sigma_r^2 p}{\sigma_n^4}\left(n(\kappa + \epsilon)^2 + \mu + \epsilon\right)\right)\right\}$$
$$\geq 1 - \frac{Var_{loc}E_f|h||g| + Var_{loc}E_f|h|^2|g|^2}{n\epsilon^2}. \quad (20)$$

It can be easily verified that for the extreme point of the above upper bound (with respect to p and $n\sigma_r^2$), we have $p = n\sigma_r^2$. Therefore,

$$\eta \leq \max_{p}\left\{\frac{1}{4p}\log\left(1 + \frac{(n(\kappa + \epsilon)^2 + \mu + \epsilon)p^2}{\sigma_n^4}\right)\right\}. \quad (21)$$

By defining $x = \frac{\sqrt{n(\kappa+\epsilon)^2+\mu+\epsilon}}{\sigma_n^2}p$, (21) may be written as

$$\eta \leq \frac{\sqrt{n(\kappa + \epsilon)^2 + \mu + \epsilon}}{4\sigma_n^2}\underbrace{\max_{x}\frac{\log(1 + x^2)}{x}}_{<2}. \quad (22)$$

Since κ, μ, ϵ, and σ_n^2 do not depend on n, it is clear from the preceding equation that $\eta \leq O(\sqrt{n})$. Also, the maximization over

x is uniquely achieved by some constant x in the interval $[0,2]$. Therefore, from the definition of x, the optimal value of p and hence $n\sigma_r^2$ is $\Theta(\frac{1}{\sqrt{n}})$. Thus, we have shown that for a random placement of nodes in the domain A, the Listen and Transmit protocol with high probability achieves a power efficiency of at most of order $\sqrt{n}$.

D. Main Result: Sensory Case

In previous sections, we found a lower bound on the power efficiency of sensory networks. Combining these bounds together, we have the following theorem.

Theorem 1: Consider a random sensory network with a transmitter/receiver pair and n relay nodes, where all the nodes are placed randomly and independently on a domain of fixed area A. Assume that averaged over the fading, the various channels are independent, i.e., for every two different channels c_1, c_2 we have $\mathrm{E}_\mathrm{f}[c_1 c_2] = \mathrm{E}_\mathrm{f}[c_1]\mathrm{E}_\mathrm{f}[c_2]$ and the measurement noises are all i.i.d. $\mathcal{CN}(0, \sigma_n^2)$. Furthermore, assume that the relay nodes have knowledge of their channels to and from the receiver and the transmitter and that the receiver knows the mean of α in (7). Then with high probability, the power efficiency of the network is at least $\Theta(\sqrt{n})$, i.e., there exist a scheme such that

$$\mathrm{P}\{K_2\sqrt{n} \geq \eta \geq K_1\sqrt{n}\} \geq 1 - \frac{K}{n} \qquad (23)$$

where K_1, K_2, and K are constants depending on the domain and the fading characteristics, but not on n. Moreover, the Listen and Transmit protocol achieves $\eta = \Theta(\sqrt{n})$ with the power allocation $p = n\sigma_r^2 = \Theta(\frac{1}{\sqrt{n}})$.

Remark:
- It was shown that in the Listen and Transmit protocol the rate of communication is of order constant. Therefore, we are getting the same rate of communication as the case when the transmitter and the receiver communicate in isolation. The difference is that in the former protocol, the total power consumption is of order $\Theta(\frac{1}{\sqrt{n}})$ which is $\sqrt{n}$ time less than the power consumption in the later case. Thus, we are getting a fixed rate with less power consumption.

- Implicit in the Listen and Transmit protocol there is a notion of fairness: nodes in relay mode consume n times less power that the node in transmit mode.

- Comparing the power efficiency achieved in the sensory networks with the power efficiency of multiple-antenna systems, we observe that it is better than the power efficiency of a $n \times 1$ multiple-antenna system with no channel knowledge at the transmit antennas where unlike the sensory case, cooperation between different antennas is allowed. However, as we expected it is worse than the power efficiency of an $n \times 1$ multiple-antenna system with perfect knowledge at the transmit antennas. □

E. Discussion on Synchronicity

The key idea in the Listen and Transmit protocol is to scale the received signals at the relay stage in such a way that the information-bearing signal parts add up coherently at the receiver.

Therefore, the protocol is sensitive to any error in the phase and hence to synchronicity. In this section, we try to make a qualitative analysis of the effect of asynchronicity on the Listen and Transmit protocol.

Instead of considering an asynchronous system we consider the lack of synchronicity by introducing a phase error in the channel knowledge used by the relay nodes. More precisely, we assume that instead of knowing the channel h_i perfectly, the ith relay node uses $h_i e^{j\delta_i}$ for processing its received signal, where δ_i is the phase error that models the time lag corresponding to the transmission from ith relay node to the receiver. We assume that the phase errors are i.i.d. random variables and independent from the channel coefficients. Furthermore, we assume that $\mathrm{E}[e^{j\delta_i}]$ is not zero and is equal to some constant $|\lambda| \neq 0$. In other words, we assume that by the aid of the receiver and by using a training sequence, the relay nodes have some estimate of their time lag and therefore the phase error is not distributed uniformly over the unit circle.

In this case, the received signal at the receiver is

$$y = \sum_{i=1}^{n} h_i t_i e^{j\delta_i} + w$$

where t_i, defined in (6), is the transmitted signal in the case of perfect synchronicity. By plugging in t_i from (6) and using the same approach as before, we have

$$C \geq \frac{1}{2} \log \left(1 + \frac{n^2 \sigma_r^2 p \left| \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}_\mathrm{f} \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}} e^{j\delta_i} \right|^2}{\sigma_n^2 + \sigma_r^2 \sum_{i=1}^{n} \mathrm{E}_\mathrm{f}|h_i|^2} \right).$$

$$(24)$$

Note that the lack of synchronicity appears as the phase errors $e^{j\delta_i}$ in the lower bound. Looking at the numerator of the lower bound, since the phase errors are independent of the channels, we can see that as we increase n

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}_\mathrm{f} \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}} e^{j\delta_i} \rightarrow \lambda \mathrm{E}_\mathrm{loc}[\mathrm{E}_\mathrm{f} \frac{|g_i||h_i|}{\sqrt{|g_i|^2 p + \sigma_n^2}}] \quad \text{a.s.}$$

In other words, as the number of nodes increases, for any random network, the term in the numerator of (24) with high probability is close to its average over the phase error and the location of the points. Therefore, using the same approach as in the previous section, with high probability, the power efficiency is lower-bounded by

$$\sqrt{n}\log(1 + K'|\lambda|^2)$$

where K' depends only on the geometry of the domain and the fading characteristics. From the above discussion, we have the following observations.

- As $|\lambda|$ decreases, i.e., as we become more and more uncertain about the phase of the channel, the power efficiency also decreases. For the case where $|\lambda| = 0$, i.e., the case where we have no estimate of the phase, the lower bound on the achieved power efficiency become zero.

- In terms of n, we see that, as long as $|\lambda| \neq 0$, the asymptotics of the lower bound does not change and we can still achieve a power efficiency of $\Theta(\sqrt{n})$ with the Listen and Transmit protocol.

V. *AD HOC* NETWORKS

We now turn our attention to *ad hoc* networks. The key difference, compared to the sensory networks, is that we now have r simultaneous transmitter/receiver pairs. Therefore, we are interested in the following question. Assume that in isolation, to maintain some fixed communication rate, each transmitter/receiver pair needs to operate at some power. Now if the r transmitter/receiver pairs are required to communicate simultaneously, and are members of a random wireless network with n nodes, how much can the total power consumption in the network be reduced (from the power consumption required in isolation) to maintain the same communication rate between the transmitters and receivers? We remark that since the capacity of an *ad hoc* wireless network scales as $\Theta(\sqrt{n})$ [2]–[5], to maintain a fixed rate for each transmitter/receiver pair we need to assume that $r \leq \sqrt{n}$. This will be our standing assumption throughout. To answer the question above, we will construct an extension of Listen and Transmit protocol developed for sensory networks. As in the sensory case, the main idea is to exploit interference in the network. For *ad hoc* networks the power efficiency is defined as the ratio between the sum of the mutual information of different transmitter/receiver pairs and the total power consumption of the network

$$\eta = \frac{C_{\text{total}}}{P_0}.$$

Note that both C_{total} and P_0 are random variables that depend on the placement of the points. We first consider an alternative form of power efficiency, namely, $\eta' = \frac{C_{\text{total}}}{\mathrm{E}_{\text{loc}}[P_0]}$, where the denominator is averaged over all point placements in the network. The reason is that it is easier to establish scaling laws for η'. We then show that similar scaling laws apply to η.

A. *Listen and Transmit Protocol*

As in the sensory case, the communication in the Listen and Transmit protocol is divided into two intervals.

1. **Listen interval:** Each of the r transmit users transmit the signal s_i, where $\{s_i\}$ are independent random variables. All other nodes listen. Relay node j, receives

$$r_j = \sum_{i=1}^r g_{ij} s_i + v_j \quad i = 1, 2, \dots, r, \ j = 1, 2, \dots, n.$$

$$(25)$$

2. **Transmit interval:** Each relay node j, transmits t_j, a scaled version of what it has previously received

$$t_j = \sigma_r d_j r_j, \qquad j = 1, 2, \dots, n$$

where the scalar d_j can depend only on the local knowledge of the channel coefficients at relay node j. Before describing the particular choice of $\{d_j\}$, it is instructive to consider what can be accomplished by having the relay nodes just scale their received signals. To this end, if, for a

particular choice of the $\{d_j\}$, we focus on the $r \times r$ channel matrix relating the r transmit signals to the r receive signals, it is clear that the entries of this matrix are linear combinations of the $\{d_j\}$ (see Section V-G and (64) later). Since the channel matrix has r^2 entries, if $n \geq r^2$ then we have enough free parameters in the $\{d_j\}$ to "generically" make the channel matrix diagonal. This totally suppresses the interference and yields r independent channels. Therefore, in principle, a sum–rate of order r is achievable.[4]

The problem with this approach is that it requires complete knowledge of all the channel coefficients at every node of the network (so that each node can solve the system of linear equations required to diagonalize the channel). Since this is not acceptable, we need to introduce a method that only uses local channel knowledge and so we propose the following choice for d_j:

$$d_j = \frac{\left(\sum_{i=1}^r g_{ij}^* h_{ji}^*\right)}{\sqrt{r\kappa_1 \sigma_n^2 + r((r-1)\kappa_2 + \kappa_3)p}} \qquad (26)$$

where κ_1, κ_2, and κ_3 are defined as

$$\kappa_1 = \mathrm{E}_{\{j\}}(\mathrm{E}_{\{\gamma|j\}} \mathrm{E}_{\mathrm{f}} |g_{\gamma j}|^2)^2$$
$$\kappa_2 = \mathrm{E}_{\{j\}}(\mathrm{E}_{\{\gamma|j\}} \mathrm{E}_{\mathrm{f}} |g_{\gamma j}|^2)^3$$
$$\kappa_3 = \mathrm{E}_{\{j\}}(\mathrm{E}_{\{\gamma|j\}} \mathrm{E}_{\mathrm{f}} |g_{\gamma j}|^2 \mathrm{E}_{\{\gamma|j\}} \mathrm{E}_{\mathrm{f}} |g_{\gamma j}|^4) \qquad (27)$$

where γ is a location of a random point in the domain. Note that these parameters do not depend on n and r and depend only on the geometry of the domain and the fading characteristics.

With the above choice of d_j's the operation of the relay nodes can be regarded as performing *distributed performing*. It is further shown in Appendix A that with this choice of d_j, the average transmit power for the relay nodes of the random network is σ_r^2, where the averaging is over both placement of the network as well as channel fading. Since d_j depends only on the local knowledge of the channel coefficients at relay node j, the d_j's are identical and independent random variables when the location of the relay nodes is random and the transmitters and the receivers are fixed. We will use this fact later on in our results. Finally, we remark that the above mentioned scheme may be interpreted as follows:

- Each relay node estimates each of the r transmitted signals as

$$\hat{s}_i^j = g_{ij}^* r_j, \qquad i = 1, 2, \dots, r, \ j = 1, 2, \dots, n.$$

Of course, these are very inaccurate estimates.

- Each node attempts to coherently add its estimate of signal s_i, for the ith receiver via multiplication by h_{ji}^* and normalize the sum to power σ_r^2

$$t_j = \frac{1}{\sqrt{r\kappa_1 \sigma_n^2 + r((r-1)\kappa_2 + \kappa_3)p}} \sum_{i=1}^r \hat{s}_i^j h_{ij}^*.$$

Note that in both steps of the protocol, we are exploiting interference. Since the wireless medium is a shared medium, each relay node can estimate each of the r transmitted signals. Also,

---

[4]Of course, one should worry about satisfying the power constraints. As we will see later, this scheme is not power efficient.

because of the interference, each receiver will receive a summation of the scaled versions of the signals that the relay nodes have sent. So there are n indirect paths for the signal transmitted from transmitter i to receiver i, each passing through one relay node. Each of the relay nodes has transmitted a signal that has a part that adds coherently for receiver i. Therefore, there are n signal parts that add up coherently at receiver i.

The received signal at receiver k is

$$
\begin{aligned}
y_k &= \sum_{j=1}^{n} h_{jk} t_j + w_k = \sum_{j=1}^{n} \sigma_r h_{jk} d_j r_j + w_k \\
&= \sum_{j=1}^{n} \sigma_r h_{jk} d_j \left(\sum_{i=1}^{r} g_{ij} s_i + v_j \right) + w_k \\
&= \sigma_r \sum_{i=1}^{r} \underbrace{\left(\sum_{j=1}^{n} g_{ij} h_{jk} d_j \right)}_{\alpha_i^k} s_i + \sigma_r \sum_{j=1}^{n} \underbrace{h_{jk} d_j}_{\beta_j^k} v_j + w_k.
\end{aligned}
$$
(28)

We should remark that α_i^k for all i is a sum of n independent random variables. Also notice that since the relay nodes are placed independently and the β_j^k's depend only on the channel coefficients between relay node j and the transmitters and the receivers, they are independent for different j's.

B. Finding Upper and Lower Bounds

By using the same technique as Section IV-A, we can find a lower and upper bound for the mutual information between y_k and s_k. Using the results of [21] again, the maximum value of the mutual information $C_k = \sup_{p(s_k)} \mathcal{I}(y_k; s_k)$, can be lower-bounded by the capacity of the AWGN channel with input/output equation

$$
y = \sigma_r E_f[\alpha_k^k] s_k + w_k'
$$
(29)

where w' is a zero-mean complex Gaussian noise with variance

$$
\begin{aligned}
\mathrm{Var}_f[w_k'] &= \mathrm{Var}_f \Big[\sigma_r \left(\alpha_k^k - E_f\left[\alpha_k^k \right] \right) s_k \\
&\quad + \sigma_r \sum_{i=1,i\neq k}^{r} \alpha_i^k s_i + \sigma_r \sum_{j=1}^{n} \beta_j^k v_j + w_k \Big] \\
&= \mathrm{Var}_f \left[\alpha_k^k \right] \sigma_r^2 p + \left(\sum_{i=1,i\neq k}^{r} E_f |\alpha_i^k|^2 \right) \sigma_r^2 p \\
&\quad + \left(\sigma_r^2 \sum_{j=1}^{n} E_f |\beta_j^k|^2 + 1 \right) \sigma_n^2.
\end{aligned}
$$
(30)

Therefore, we have

$$
\begin{aligned}
C_k &= \sup_{p(s_k)} \mathcal{I}(y_k; s_k) \geq \sup_{p(s_k)} \mathcal{I}(y; s_k) \\
&= \frac{1}{2} \log \left(1 + \frac{\sigma_r^2 p |E_f\left[\alpha_k^k \right]|^2}{\mathrm{Var}_f[w_k']} \right).
\end{aligned}
$$
(31)

Note that the $\frac{1}{2}$ in front of the logarithmic term comes from the fact that the transmitters transmit half of the times. We can obtain an upper bound on C_k, considering the case that the receiver

k knows $\{v_j, s_i, i = 1, \ldots, r, i \neq k, j = 1, \ldots, n\}$ and all the channel coefficients. For this case we have

$$
\begin{aligned}
C_k &= \sup_{p(s_k)} \mathcal{I}(y_k; s_k) \\
&\leq \sup_{p(s_k)} \mathcal{I}(y_k; s_k | \{v_j, s_i, i \neq k, g_{ij}, h_{jk}\}) \\
&= \frac{1}{2} E_{f\{g_{ij}, h_{jk}\}} \left[\log \left(1 + \frac{|\alpha_k^k|^2 \sigma_r^2 p}{\sigma_n^2} \right) \right].
\end{aligned}
$$
(32)

Using the convexity of the log function, we may rewrite the above equation as

$$
C_k = \sup_{p(s_k)} \mathcal{I}(y_k; s_k) \leq \frac{1}{2} \log \left(1 + \frac{E_f |\alpha_k^k|^2 \sigma_r^2 p}{\sigma_n^2} \right).
$$
(33)

In order to compute the lower and the upper bound in (31) and the above equation, first- and second-order moments of $\{\alpha_i^k\}$ and $\{\beta_j^k\}$ are required. In the following lemma, we give probabilistic bounds on $\mathrm{Var}_f[w_k']$ and $E_f[\alpha_k^k]$. The proof of this lemma is given in Appendix B.

Lemma 1: For every domain A of fixed area and every placement of the nodes of the network, there exist constants B_i, $i = 1, \ldots, 8$, κ, κ_1, and κ' ($\kappa' \leq \kappa$) that depend only on the domain A and the fading characteristics such that for every positive ϵ_i, $i = 1, \ldots, 5$ that $\epsilon_4 < \frac{B_7}{\sqrt{r(\kappa_1 \sigma_n^2 + rp\kappa)}}$ and every positive ρ and ζ, we have the following relations:

$$
\begin{aligned}
&P\left\{ \left| E_f\left[\alpha_k^k \right] \right| > n \left(\frac{B_7}{\sqrt{r \left(\kappa_1 \sigma_n^2 + rp\kappa \right)}} - \epsilon_4 \right) \right\} \\
&\geq 1 - \frac{B_8}{n\epsilon_4^2 \left(\kappa_1 \sigma_n^2 + rp\kappa' \right)} \\
&P\left\{ E_f \left| \alpha_k^k \right|^2 < n \left(\frac{B_2}{r\kappa'p + \kappa_1 \sigma_n^2} + \epsilon_2 \right) \right. \\
&\quad \left. + n^2 \left(\frac{B_7}{\sqrt{r \left(\kappa_1 \sigma_n^2 + r\kappa p \right)}} + \epsilon_4 \right)^2 \right\} \\
&\geq 1 - \left(\frac{B_8}{n\epsilon_4^2 \left(\kappa_1 \sigma_n^2 + rp\kappa' \right)} \right. \\
&\quad \left. + \frac{3B_3}{n\epsilon_2^2 \left(\kappa'rp + \kappa_1 \sigma_n^2 \right)^2} \right) \\
&P\left\{ \mathrm{Var}[w_k'] < \sigma_r^2 p \left(n\epsilon_2 + nr\epsilon_3 + \frac{nr(B_2 + B_5)}{\kappa'rp + \kappa_1 \sigma_n^2} + n^\rho r^\zeta \epsilon_5 \right) \right. \\
&\quad \left. + \sigma_n^2 \left(1 + n\sigma_r^2 \left(\epsilon_1 + \frac{\kappa}{\kappa rp + \kappa_1 \sigma_n^2} \right) \right) \right\} \\
&\geq 1 - \frac{1}{\left(\kappa'rp + \kappa_1 \sigma_n^2 \right)^2} \\
&\quad \times \left(\frac{3B_1}{n\epsilon_1^2} + \frac{3B_3}{n\epsilon_2^2} + \frac{3(r-1)B_4}{n\epsilon_3^2} + h(n, r, \epsilon_5) \right)
\end{aligned}
$$
(34)

where

$$
h(n, r, \epsilon_5) = \begin{cases} \frac{rn(\kappa'rp + \kappa_1 \sigma_n^2)}{n^\rho r^\zeta \epsilon_5}, & \text{Using Assumptions A, B} \\ \frac{rn(r+n)B_6}{n^{2\rho} r^{2\zeta} \epsilon_5^2}, & \text{Using Assumptions A, B, C} \end{cases}
$$

Using Lemma 1, we can combine (31) and (33) to get upper and lower bounds for C_k. For this, define C_{lower} (see (35) at

the bottom of the page), then we have the following relations as probabilistic lower and upper bounds for C_k:

$$P\{C_k \geq C_{\text{lower}}\} \geq 1 - \frac{1}{(\kappa' rp + \kappa_1 \sigma_n^2)^2}$$
$$\times \left(\frac{3B_1}{n\epsilon_1^2} + \frac{3B_3}{n\epsilon_2^2} + \frac{3(r-1)B_4}{n\epsilon_3^2} \right.$$
$$\left. + \frac{B_8 (\kappa' rp + \kappa_1 \sigma_n^2)}{n\epsilon_4^2} + h(n, r, \epsilon_5) \right)$$

$$P\{C_k \leq C_{\text{upper}}\} \geq 1$$
$$- \left(\frac{B_8}{n\epsilon_4^2 (\kappa_1 \sigma_n^2 + rp\kappa')} + \frac{3B_3}{n\epsilon_2^2 (\kappa' rp + \kappa_1 \sigma_n^2)^2} \right). \quad (36)$$

In the Listen and Transmit protocol, since there are r transmitters and n relay nodes and all the nodes are transmitting half of the times, the average total power consumption is $\frac{1}{2}(rp + n\sigma_r^2)$. The total capacity of the network is $C_{\text{total}} = \sum_{k=1}^{r} C_k$. Therefore, the power efficiency of the network is

$$\eta' = \frac{2C_{\text{total}}}{rp + n\sigma_r^2} = \frac{2 \sum_{k=1}^{n} C_k}{rp + n\sigma_r^2}. \quad (37)$$

Remark: As mentioned earlier in Lemma 1, the constants B_i and κ, κ_1, and κ' do not depend on n and r. Now if we fix σ_r^2, p, and ϵ_i, $i = 1, 2, 5$ and set $\epsilon_3 = \frac{1}{r}$, $\epsilon_4 = \frac{B_7}{2\sqrt{r(\kappa_1 \sigma_n^2 + rp\kappa)}}$, and $\zeta = 0$ in (35) and (36), then the total capacity achieved by the Listen and Transmit protocol is bounded probabilistically as

$$P\left\{ C_{\text{total}} \geq K_1' r \log \left(1 + \frac{n^2}{r^2(n + n^\rho)} \right) \right\}$$
$$> 1 - \frac{rK_2'}{r^2} \left(\frac{r^3}{n} + \frac{r}{n^{2(\rho-1)}} \right)$$

where K_1' and K_2' are some constants and we have considered assumptions A, B, and C in Section II-B. Therefore, by setting $\rho = 1 + \theta$, $\theta > 0$, we have

$$P\left\{ C_{\text{total}} \geq K_1' r \log \left(1 + \frac{n^{1-\theta}}{r^2} \right) \right\} \geq 1 - K_2' \left(\frac{r^2}{n} + \frac{1}{n^{2\theta}} \right).$$

Now note that the maximum of the bound is achieved for $r = \Theta(n^{\frac{1-\theta}{2}})$ and in that case we have

$$P\left\{ C_{\text{total}} \geq K_1' n^{\frac{1-\theta}{2}} \right\} \geq 1 - \frac{K_2'}{n^\theta}, \qquad \forall \theta > 0. \quad (38)$$

From (38) we see that, with high probability, by using the Listen and Transmit protocol, we can get arbitrary close to the $\Theta(\sqrt{n})$ result of [2]. This result is interesting since we are only using the local knowledge of the channel coefficients at the relay nodes and the protocol is very simple. (It is double-hop and requires no routing.) $\square$

C. Power Allocation

We will now focus on how to optimally allocate the transmit and relay node powers (i.e., p and σ_r^2 as functions of r and n) to maximize the power efficiency. Define

$$\eta'_{\text{lower}} = \frac{2rC_{\text{lower}}}{rp + n\sigma_r^2} \quad \text{and} \quad \eta'_{\text{upper}} = \frac{2rC_{\text{upper}}}{rp + n\sigma_r^2}.$$

By using union bound on the probability of events, we get the following probabilistic lower and upper bounds for the power efficiency of the network using the Listen and Transmit protocol:

$$P\{\eta'_{\text{lower}} \leq \eta\} \geq P\{C_{\text{lower}} \leq C_k \quad k = 1, \dots, r\}$$
$$\geq 1 - rP\{C_k < C_{\text{lower}}\}$$
$$P\{\eta'_{\text{upper}} \geq \eta\} \geq P\{C_{\text{upper}} \geq C_k \quad k = 1, \dots, r\}$$
$$\geq 1 - rP\{C_k < C_{\text{upper}}\}. \quad (39)$$

We will consider the lower bound first. We try to choose the values for p, σ_r^2, ζ, and ρ so that with high probability we can achieve a power efficiency that scales with the number of nodes in the network. For this goal we take ϵ_1, ϵ_2, ϵ_3, and ϵ_5 all to be equal to a positive constant denoted by ϵ. We also choose $\epsilon_4 = \frac{B_7}{2\sqrt{r(\kappa_1 \sigma_n^2 + r\kappa p)}}$. We further consider the network operating in the low-SNR regime so that rp is at most constant (in terms of how it scales with n). Later on, when we are looking at the upper bound, we will show that the optimal operating point for this protocol is indeed when rp is of $O(1)$. Using these assumptions in (35) and adding $K_3 rp$ to the denominator, we have the following lower bound for the power efficiency η' (see (40) at the bottom of the page). $K_i, i = 1, \dots, 5$ are constants and

$$C_{\text{lower}} = \frac{1}{2} \log \left(1 + \frac{n^2 \sigma_r^2 p \left(\frac{B_7}{\sqrt{r(\kappa_1 \sigma_n^2 + rp\kappa)}} - \epsilon_4 \right)^2}{p\sigma_r^2 \left(n\epsilon_2 + nr\epsilon_3 + \frac{nr(B_2 + B_5)}{\kappa' rp + \kappa_1 \sigma_n^2} + n^\rho r^\zeta \epsilon_5 \right) + \sigma_n^2 \left(1 + n\sigma_r^2 \left(\epsilon_1 + \frac{\kappa}{\kappa rp + \kappa_1 \sigma_n^2} \right) \right)} \right)$$

$$C_{\text{upper}} = \frac{1}{2} \log \left(1 + n \frac{\sigma_r^2 p}{\sigma_n^2} \left(\frac{B_2}{r\kappa' p + \kappa_1 \sigma_n^2} + \epsilon_2 \right) + n^2 \frac{\sigma_r^2 p}{\sigma_n^2} \left(\frac{B_7}{\sqrt{r(\kappa_1 \sigma_n^2 + r\kappa p)}} + \epsilon_4 \right)^2 \right) \quad (35)$$

$$\eta'_{\text{lower}} = \frac{r}{rp + n\sigma_r^2} \log \left(1 + \frac{K_1 \frac{n^2}{r} \sigma_r^2 p}{p\sigma_r^2 (n + nrK_2 + n^\rho r^\zeta)\epsilon + \sigma_n^2 + K_3 (n\sigma_r^2 + rp)} \right)$$

$$P\{\eta'_{\text{lower}} \leq \eta\} \geq 1 - r \left(\frac{rK_4}{n\epsilon^2} + g(n, r, \epsilon) \right). \quad (40)$$

do not depend on r or n. $g(n, r, \epsilon)$ is also derived from $h(n, r, \epsilon)$ after applying the simplifications

$$g(n, r, \epsilon_5) = \begin{cases} \frac{K_5 rn}{n^\rho r^\zeta \epsilon}, & \text{Using Assumptions A, B} \\ \frac{rn(r+n)B_6}{n^{2\rho} r^{2\zeta} \epsilon^2}, & \text{Using Assumptions A, B, C} \end{cases}$$

(41)

Looking at (40), the following conditions are necessary in order to have $P\{\eta'_{\text{lower}} \leq \eta'\} \to 1$ for large n and r:

$$\frac{r^2}{n} \to 0 \qquad g(n, r, \epsilon)r \to 0. \qquad (42)$$

Therefore, from the above equation it is clear that this analysis is valid for the case where $r = O(\sqrt{n})$. ζ and ρ should be chosen so that the second condition in (42) is satisfied. These two parameters determine the rate of convergence in probability. By looking at (41), we observe that the second condition in (42) implies that $\frac{1}{n^{\rho-1} r^{\zeta-1}} \to 0$ an n grows. In this stage, we maximize the power efficiency with respect to the total transmit power rp and total relay power $n\sigma_r^2$ and subject to the constraints in (42). It can be easily verified by taking partial derivatives with respect to rp and $n\sigma_r^2$ that the expression is maximized for $rp = n\sigma_r^2 = x$. Hence, we can write the maximization problem as

$$\eta^\star_{\text{lower}} = \max_{rp, n\sigma_r^2} \eta'_{\text{lower}}$$

$$= \max_{x \geq 0} \frac{r}{2x} \log \left(1 + \frac{K_1 \frac{n}{r^2} x^2}{\sigma_n^2 + 2K_3 x + (n + nrK_2 + n^\rho r^\zeta)\epsilon x^2}\right).$$

(43)

Let $x = \Theta(n^\alpha)$, $\alpha \leq 0$ and $r = \Theta(n^\nu)$, $0 \leq \nu \leq 1/2$. Using the fact that $\alpha \leq 0$ and $\frac{1}{n^{\rho-1} r^{\zeta-1}} \to 0$ in (43), we can write

$$\eta^\star_{\text{lower}}$$

$$= \Theta \left(\max_\alpha n^{\nu-\alpha} \log \left(1 + \frac{n^{1+2\alpha}}{n^{2\nu} (\sigma_n^2 + n^{\rho+2\alpha+\nu(\zeta-1)-1})}\right)\right)$$

(44)

With the following constraints

$$\alpha \leq 0, \nu \leq \frac{1}{2}$$

$$\begin{cases} \rho - 1 + \nu(\zeta - 2) > 0, & \text{If using assumptions A, B} \\ \rho - 1 + \nu(\zeta - 1) > 0, & \text{If using assumptions A, B, C} \end{cases}$$

(45)

where the last two constraints are consequences of (42).

Consider that we use assumptions A, B, and C. Later, we analyze the performance of the protocol when only assumptions A and B are used. Set $\rho - 1 + \nu(\zeta - 1)$ fixed and equal to μ. In this case, the rate of convergence in the probability expression of (40) is $\min\{1 - 2\nu, \mu\}$. Now we are interested in the maximum achievable power efficiency for a fixed μ. We consider the following cases.

1) $\mu + 2\alpha \leq 0$: In this case, we can see that $\rho + 2\alpha + \nu(\zeta - 1) - 1 \leq 0$. Therefore, the noise power is dominant to the interference in (44) and we can simplify the expression as

$$\eta^\star_{\text{lower}} = \Theta \left(n^{\nu-\alpha} \log(1 + n^{1-2(\nu-\alpha)})\right)$$

$$= \Theta \left(n^\beta \log(1 + n^{1-2\beta})\right)$$

(46)

where β is defined as $\beta = \nu - \alpha$. Note that we have $\mu + 2\nu \leq 2\beta$. Now we consider the following cases.

• $\beta \geq \frac{1}{2}$: In this case the power of n in the $\log$ function in (46) is negative so it is of $\Theta(n^{1-2\beta})$. Therefore we have

$$\eta^\star_{\text{lower}} = \Theta(n^{1-\beta}).$$

The total rate of transmission R_{sum} is of $\Theta(n^{\nu+1-2\beta})$ in this case. The maximum order of power efficiency is achieved when β takes its smallest possible value, i.e., $\beta = \max\{\frac{1}{2}, \frac{\mu+2\nu}{2}\}$. For this case, the maximum achievable power efficiency and the total rate of transmission R_{sum} are respectively

$$\eta^\star_{\text{lower}} = \Theta \left(n^{1-\max\{\frac{1}{2}, \frac{\mu+2\nu}{2}\}}\right)$$

$$R_{\text{sum}} = \Theta(n^{\min\{\nu, 1-(\nu+\mu)\}}).$$

The transmit power and the relay node power for achieving the maximum power efficiency are

$$rp = n\sigma_r^2 = \Theta(n^\alpha) = \Theta(n^{\nu-\beta})$$

$$= \Theta(n^{\min\{\frac{2\nu-1}{2}, \frac{-\mu}{2}\}}) = \Theta \left(\min\{\frac{r}{\sqrt{n}}, \frac{1}{\sqrt{n^\mu}}\}\right).$$

(47)

Therefore, with the choice of the transmit and relay node power as above, we have

$$P\left\{\eta' = \Theta \left(n^{1-\max\{\frac{1}{2}, \frac{\mu+2\nu}{2}\}}\right)\right\}$$

$$\geq 1 - \Theta \left(\frac{1}{n^{\min\{\mu, 1-2\nu\}}}\right). \qquad (48)$$

• $\beta < \frac{1}{2}$: In this case, (46) can be rewritten as

$$\eta^\star_{\text{lower}} = \Theta(n^\beta \log n^{1-2\beta}) = (1 - 2\beta)\Theta(n^\beta \log n).$$

Now, since $\log n$ grows slower than any polynomial function in n and β is strictly less than $\frac{1}{2}$, the maximum achievable power efficiency in this case cannot be better than the previous case and thus operating the network in this region is not favorable.

2) $\mu + 2\alpha \geq 0$: For this regime, the interference will be the dominant term in (44) and therefore we have

$$\eta^\star_{\text{lower}} = \Theta \left(n^{\nu-\alpha} \log \left(1 + n^{1-(2\nu+\mu)}\right)\right).$$

As we can see from the preceding equation, the power efficiency is maximized when α takes is greatest possible value $\alpha = \frac{-\mu}{2}$. In this case, the power efficiency of the network is

$$\eta^\star_{\text{lower}} = \Theta \left(n^{\frac{2\nu+\mu}{2}} \log \left(1 + n^{1-(3\nu+\mu)}\right)\right). \qquad (49)$$

It can be easily checked that the maximum power efficiency achieved in this region is always less than or equal to the case where $\mu + 2\alpha \leq 0$.

The preceding discussion gives the probabilistic lower bound of (48) for power efficiency when assumptions A, B and C from Section II-B can be used. If only assumptions A and B can be used, then applying the same technique as above, we can easily check that the power efficiency is maximized when the network operates in the noise dominant regime (i.e., $\mu + \nu + 2\alpha \leq 0$) rather than interference dominant regime (i.e., $\mu + \nu + 2\alpha \geq 0$). Also, similar to the previous discussion $\beta \triangleq \nu - \alpha$ should be greater than equal to $\frac{1}{2}$.

In this region, $\rho + 2\alpha + \nu(\zeta - 1) - 1 \leq 0$ and we are in the noise dominant regime. Hence, we can simplify (44) to

$$\eta^{\star}_{\text{lower}} = \Theta\left(n^{\nu-\alpha}\log(1 + n^{1-2(\nu-\alpha)})\right). \qquad (50)$$

Therefore, we have $\mu + 3\nu \leq 2\beta$. Also, since $\beta \geq \frac{1}{2}$, the power of n in the $\log(.)$ function in (50) is negative, so it is of $\Theta(n^{1-2\beta})$. Therefore, we have

$$\eta^{\star}_{\text{lower}} = \Theta(n^{1-\beta})$$
$$R_{\text{sum}} = \Theta(n^{\nu+1-2\beta})$$

The best achievable power efficiency is for the case when β takes its smallest possible value. In this case

$$\beta = \max\left\{\frac{1}{2}, \frac{\mu + 3\nu}{2}\right\}$$

and the maximum achievable power efficiency and the total rate of transmission R_{sum} are respectively

$$\eta^{\star}_{\text{lower}} = \Theta\left(n^{1-\max\left\{\frac{1}{2}, \frac{\mu+3\nu}{2}\right\}}\right)$$
$$R_{\text{sum}} = \Theta(n^{\min\{\nu, 1-(2\nu+\mu)\}}).$$

The transmit power and the relay node power for achieving the maximum power efficiency are

$$rp = n\sigma_r^2 = \Theta\left(n^{\min\left\{\frac{2\nu-1}{2}, \frac{-\nu-\mu}{2}\right\}}\right)$$
$$= \Theta\left(\min\left\{\frac{r}{\sqrt{n}}, \frac{1}{\sqrt{n^{\mu}r}}\right\}\right). \qquad (51)$$

With this choice of transmit and relay node power we have the following probabilistic lower bound on the power efficiency (using assumptions A and B only):

$$P\left\{\eta = \Theta\left(n^{1-\max\left\{\frac{1}{2}, \frac{\mu+3\nu}{2}\right\}}\right)\right\} \geq 1 - \Theta\left(\frac{1}{n^{\min\{\mu, 1-2\nu\}}}\right). \qquad (52)$$

D. Main Result: Ad Hoc Case

The analysis in the previous section shows the following result.

Theorem 2: Consider an n-node random *ad hoc* network where the nodes are placed randomly and independently on a domain of fixed area where averaged over the fading, the various channels are independent, i.e., for every two different channels c_1, c_2 we have $E_f[c_1 c_2] = E_f[c_1].E_f[c_2]$. Furthermore,

assume conditions A, B, and C given in Section II-B and that at any given time there are $r = O(n^{\nu})$ $\nu \leq \frac{1}{2}$ transmit/receive pairs. Also the measurement noises are all i.i.d. $\mathcal{CN}(0, \sigma_n^2)$. If we denote the power efficiency of the network by η' (i.e, $\eta' = \frac{C_{\text{total}}}{E_{\text{loc}}[P_0]}$) then for every $\mu > 0$

$$P\left\{\eta' \geq K_1\left(n^{1-\max\left\{\frac{1}{2}, \frac{\mu+3\nu}{2}\right\}}\right)\right\} \geq 1 - \frac{K_2}{n^{\min\{1-2\nu, \mu\}}} \qquad (53)$$

where K_1 and K_2 are independent of n and r but depend on the domain and the fading characteristics. Moreover, the Listen and Transmit protocol achieves this lower bound. The transmit and the relay node powers that achieve this power efficiency are given in (47).

The following corollary is an immediate consequence of Theorem 2 by setting $\mu = 1 - 2\nu$.

Corollary 1: Consider the network model described in Theorem 2. If the number of transmitter/receiver pairs in the network is of $O(n^{\frac{1-\epsilon}{2}})$, where $\epsilon > 0$, then we have

$$P\{\eta \geq K_3\sqrt{n}\} \geq 1 - \frac{K_4}{n^{\epsilon}} \qquad (54)$$

where K_3 and K_4 are independent of n and r but depend on the domain and the fading characteristics. Moreover, by choosing the transmit and relay node powers as $rp = n\sigma_r^2 = \Theta(\frac{r}{\sqrt{n}})$, the Listen and Transmit protocol achieves this lower bound.

Remark: Note that from (54), we can see that there is a tradeoff between the number of transmitter/receiver pairs r, and the rate of convergence. As we increase r from $\Theta(1)$ to $\Theta(\sqrt{n})$, the convergence slows down. $\square$

For the case, when only assumptions A and B are used, we have the following result from (51) and (52).

Theorem 3: Consider an n-node random *ad hoc* network in a domain of fixed area where averaged over the fading, the various channels are independent. Furthermore assume conditions A, B given in Section II-B and that at any given time there are $r = O(n^{\nu})$ $\nu \leq \frac{1}{2}$ transmit/receive pairs. Also the measurement noises are all i.i.d. $\mathcal{CN}(0, \sigma_n^2)$. If we denote the power efficiency of the network by η' (i.e., $\eta' = \frac{C_{\text{total}}}{E_{\text{loc}}[P_0]}$) then for every $\mu > 0$

$$P\left\{\eta' \geq K_5 n^{1-\max\left\{\frac{1}{2}, \frac{\mu+3\nu}{2}\right\}}\right\} \geq 1 - \frac{K_6}{n^{\min\{1-2\nu, \mu\}}}$$

where K_5 and K_6 are independent of n and r and depend on the domain and fading characteristics. Moreover, the Listen and Transmit protocol achieves this lower bound. The transmit and relay node powers achieving this power efficiency are given in (51).

By considering the case where $\mu + 3\nu = 1$, we have the following corollary.

Corollary 2: Consider the network described in Theorem 3. If the number of transmit/receive pairs in the network is of $O(n^{\frac{1-\epsilon}{3}})$, where $\epsilon > 0$, then we have

$$P\{\eta' \geq K_7\sqrt{n}\} \geq 1 - \frac{K_8}{n^{\epsilon}}$$

where K_7 and K_8 only depend on the domain and fading characteristics. Moreover, by choosing the transmit and relay node powers as $rp = n\sigma_r^2 = \Theta(\frac{r}{\sqrt{n}})$, the Listen and Transmit protocol achieves this lower bound.

Corollary 2 implies that the maximum number of transmit/receive pairs that the network can support with power efficiency of $\Theta(\sqrt{n})$ is $\Theta(n^{\frac{1}{3}})$. On the other hand, considering the maximum power efficiency of the network with the number of transmit/receive pairs up to $n^{\frac{1}{2}}$ one can write the following corollary by setting μ equal to $1-2\nu$ in Theorem 3. We should also remark that comparing Corollaries 1 and 2, we see that the effect of assumption C is on the number of simultaneous transmitter/receiver pairs that can be in the network.

Corollary 3: Consider again the network described in Theorem 3. If the number of transmit/receive pairs in the network is of $\Theta(n^{\frac{1-\epsilon}{2}})$, where $0 < \epsilon < 1$, then we have

$$P\left\{\eta' \geq Kn^{(\frac{1+\epsilon}{4})}\right\} \geq 1 - \frac{K'}{n^\epsilon}$$

where K, K' are some constants. Therefore, in this case, if the number of transmit/receive pairs is near to $\sqrt{n}$ we can achieve a power efficiency that scales like $n^{\frac{1}{4}}$.

The following remarks are in order. The previous discussion shows the following.

- If the number of the transmitter/receiver pairs is less than $\sqrt{n}$, it was shown that a power efficiency that scales with the number of nodes, n, is achievable. The rate per transmitter/receiver pair in this case is of order constant. If we increase the number of simultaneous transmissions to more than $\sqrt{n}$, we can still achieve power efficiency of $\sqrt{n}$ using time-sharing and Listen and Transmit protocol together. In this case, at each time instant $\sqrt{n}$ of the transmitters transmit and all the others act as relay nodes. However, in this case, the rate per transmitter/receiver will not be of order constant and it will be of order $\frac{1}{\sqrt{n}}$. This is in agreement with the result of [2]–[4] in that achieving a constant rate per node in this case would require a total sum–capacity larger than $\Theta(\sqrt{n})$ which is not possible.

- There is a notion of fairness implicit in the protocol, in the sense that nodes in the relay mode consume $\frac{r}{n}$th power of the nodes in the transmit mode.

- For the case where $r = \Theta(n^{\frac{1-\epsilon}{2}})$ (or $r = \Theta(n^{\frac{1-\epsilon}{3}})$, if we do not have assumption C in Section II-B), the optimal choice of the transmit power and relay power is $p = \Theta(\frac{1}{\sqrt{n}})$, $\sigma_r^2 = \Theta(\frac{1}{n})$. The total power consumption is $\Theta(\frac{r}{\sqrt{n}})$ and the total rate is $\Theta(r)$.

- In the case of *ad hoc* networks, by using the Listen and Transmit protocol, we can see that we are keeping the rate of transmission for each transmitter/receiver fixed and of order $\Theta(1)$ but the total power consumption decreases as the number of nodes grows larger, as long as $r = O(n^{\frac{1-\epsilon}{2}})$ for some positive ϵ. If we do not have assumption C on the channel coefficients, we still have this property for $r = O(n^{\frac{1-\epsilon}{3}})$ for some positive ϵ.

We should mention that with high probability we cannot get a better power efficiency for *ad hoc* networks with this protocol. We can show this by using (35), (36), and (39) to find a probabilistic upper bound. With an argument like the one for the lower bound or the one in [15], we can show that with high probability

the maximum achievable power efficiency with this protocol is $O(\sqrt{n})$. The proof is included in Appendix D.

E. A Further Result

As mentioned earlier, the power efficiency that was considered up to now was defined as the ratio between the sum rate capacity for a specific placement of the nodes of the network and the average of the power consumption over all possible point placements of the network. In other words, for a specific placement of the nodes with sum rate capacity of C_{total} and power consumption of P_0, we defined power efficiency as $\eta' = \frac{C_{\text{total}}}{E_{\text{loc}}[P_0]}$. An alternative to this definition is to consider the ratio between the rate and power consumption for a specific network, i.e. $\eta = \frac{C_{\text{total}}}{P_0}$ as the power efficiency. In this case, the power efficiency η is a random variable depending on the placement of the nodes. However, because of the law of large numbers, as the size of the network increases, η will be close to its average, and we observe the same behavior as η' for the power efficiency. In order to state this formally we will need the following lemma the proof of which we have omitted due to similarity to the computations done in Appendix B. We should remark that in proving this lemma one only needs assumptions A and B of Section II-B. This lemma gives a bound on the power consumption at the relay stage.

Lemma 2: Consider an n-node *ad hoc* network with assumptions provided in Theorem 2; then for any specific placement of the nodes in the network, the total power consumption at relay nodes P_r can be bounded as

$$P\left\{\left|P_r - n\sigma_r^2\right| < \frac{n^{1+\gamma}\sigma_r^2}{\sqrt{r}}\left(\sigma_n^2 + rp\right)\right\}$$
$$\geq 1 - \frac{K'}{n^{2\gamma}(\kappa_1\sigma_n^2 + r\kappa p)^2} \quad (55)$$

where γ is any positive number and K', κ, κ_1 are constant independent of n and r.

Using this lemma and the fact that the network operate at low-SNR regime, i.e., $rp \ll 1$, we can bound the total power consumption of the network $P_0 = rp + P_r$ as follows:

$$P\left\{P_0 \leq \left(1 + \frac{n^\gamma}{\sqrt{r}}\right)\left(rp + n\sigma_r^2\right)\right\} \geq 1 - \frac{K}{n^{2\gamma}}.$$

Therefore, for $\eta = \frac{R_{\text{sum}}}{P_0}$ we have

$$P\left\{\left(1 + \frac{n^\gamma}{\sqrt{r}}\right)^{-1}\eta' \leq \eta\right\} \geq 1 - \frac{K}{n^{2\gamma}}. \quad (56)$$

One can combine this relation with the results on the power efficiency η' to get new bounds on η. The following theorems are immediate consequences of (56), Theorem 2, and Theorem 3.

Theorem 4: Consider an n-node random *ad hoc* network where the nodes are placed randomly and independently on a domain of fixed area where averaged over the fading, the various channels are independent, i.e., for every two different channels c_1, c_2 we have $E_f[c_1c_2] = E_f[c_1].E_f[c_2]$. Furthermore assume conditions A, B, and C given in Section II-B and that at any given time there are $r = O(n^\nu)$ $\nu \leq \frac{1}{2}$ transmit/receive

pairs. Also the measurement noises are all i.i.d. $\mathcal{CN}(0, \sigma_n^2)$. For a specific placement of the nodes of the random network, let C_{total} be the total rate of communication and P_0 be the total power consumption in the network. Then for every $\mu, \gamma > 0$, the power efficiency of the network defined as $\eta = \frac{C_{\text{total}}}{P_0}$ satisfies

$$\mathrm{P}\left\{\eta \geq K_1 \frac{\left(n^{1-\max\{\frac{1}{2}, \frac{\mu+2\nu}{2}\}}\right)}{1 + \frac{n^\gamma}{\sqrt{r}}}\right\} \geq 1 - \frac{K_2}{n^{\min\{1-2\nu, \mu, 2\gamma\}}}$$
(57)

where K_1 and K_2 are independent of n and r but depend on the domain and the fading characteristics. Moreover, the Listen and Transmit protocol achieves this lower bound.

By setting $\gamma = \frac{\nu}{2}$ and $\mu = 1 - 2\nu$ we have the following corollary.

Corollary 4: For the network described in Theorem 4, we have

$$\mathrm{P}\{\eta \geq K_1\sqrt{n}\} \geq 1 - \frac{K_2}{n^{\min\{1-2\nu, \nu\}}}.$$
(58)

Therefore, it becomes clear that by considering the power efficiency as the ratio between the sum–rate of transmission and the total power consumption for specific placement of the nodes in the network, we still have similar scaling behavior.

Remark: We should remark that the rate of convergence obtained for the probability of the event in (58) is not tight (for small r). One expects that as r, the number of pairs requesting service from the network, decreases, the rate of convergence improves. (For instance, we can see from Theorem 1 that for sensory networks in which $r = 1$, the rate of convergence is proportional to $\frac{1}{n}$.) Looking at (58), we observe that as r decreases to constant the convergence slows down. This is an artifact of our approach in bounding η. □

If we use only assumptions A and B in Section II-B then we can write the following theorem using (56) and Theorem 3.

Theorem 5: Consider an n-node random *ad hoc* network where the nodes are placed randomly and independently on a domain of fixed area where averaged over the fading, the various channels are independent, i.e., for every two different channels c_1, c_2 we have $\mathrm{E}_{\mathrm{f}}[c_1 c_2] = \mathrm{E}_{\mathrm{f}}[c_1].\mathrm{E}_{\mathrm{f}}[c_2]$. Furthermore, assume conditions A, B given in Section II-B and that at any given time there are $r = O(n^\nu) \quad \nu \leq \frac{1}{2}$ transmit/receive pairs. Also, the measurement noises are all i.i.d. $\mathcal{CN}(0, \sigma_n^2)$. For a specific placement of the nodes of the random network, let C_{total} be the total rate of communication and P_0 be the total power consumption in the network. Then for every $\mu, \gamma > 0$, the power efficiency of the network defined as $\eta = \frac{C_{\text{total}}}{P_0}$ satisfies

$$\mathrm{P}\left\{\eta \geq K_5 \frac{\left(n^{1-\max\{\frac{1}{2}, \frac{\mu+3\nu}{2}\}}\right)}{1 + \frac{n^\gamma}{\sqrt{r}}}\right\} \geq 1 - \frac{K_6}{n^{\min\{1-2\nu, \mu, 2\gamma\}}}$$
(59)

where K_5 and K_6 are independent of n and r but depend on the domain and the fading characteristics. Moreover, the Listen and Transmit protocol achieves this lower bound.

By setting $\gamma = \frac{\nu}{2}$ and $\mu = 1 - 3\nu$, we have the following corollary.

Corollary 5: Consider the network model described in Theorem 4. Then for $r = O(n^\nu) \quad \nu \leq \frac{1}{3}$ we have

$$\mathrm{P}\{\eta \geq K_1\sqrt{n}\} \geq 1 - \frac{K_2}{n^{\min\{1-3\nu, \nu\}}}.$$
(60)

F. Discussion on Synchronicity

Similar to the sensory case, the key idea of the Listen and Transmit protocol used for *ad hoc* networks, is the coherent and synchronous reception of the signals. Therefore, the protocol is sensitive to synchronicity. In this subsection, we discuss the effect of asynchronocity on our protocol.

Like sensory networks, instead of considering an asynchronous system, we consider the lack of synchronicity by introducing a phase error in the channel knowledge used by the relay nodes. More precisely, we assume that instead of knowing the channel h_{ji} perfectly, the jth relay node uses $h_{ji}e^{j\delta_{ji}}$ for processing its received signal. δ_{ji} is the phase error that models the time lag corresponding to the transmission from the jth relay node to the ith receiver. We assume that the phase errors are i.i.d. random variables and are independent from channel coefficients. Furthermore, we assume that $\mathrm{E}[e^{j\delta_{li}}]$ is not zero and is equal to some constant λ. In other words, we assume that by the aid of the receivers and by using a training sequence, the relay nodes have some estimate of their time lag and therefore the phase error is not distributed uniformly over the unit circle.

In this case, the scalar $\hat{d}_k$ used by the kth relay node is proportional to

$$\sum_{l=1}^{r} g_{lk}^* h_{kl}^* e^{-j\delta_{lk}}.$$

Using these $\hat{d}_k$'s, we can find the new α_k^i's and β_k^i's (28) in terms of δ_{li}. Following the lines of Section V-A, Section V-B and Appendix B, it can be verified that we will still have the same asymptotic behavior for power efficiency in terms of n (i.e., the asymptotic behavior of the achieved power efficiency is still like $\sqrt{n}$), but the constants appearing in the relations will now depend on λ as well. Also, we should remark that in Appendix B, we bounded the required moments of α_k^i's and β_k^i's over the fading by their average value over the geometry of the domain. In the presence of asynchronicity, we should also take into account the averaging over the δ_{lj}'s in our bounding. Therefore, the reader can verify that the bounds will still hold but the constants will change and similarly to the sensory case as $|\lambda|$ decreases, the power efficiency will also decrease and finally for the limiting case of $|\lambda| = 0$, the lower bound on the achieved power efficiency also becomes zero.

G. Complete Knowledge of the Channel

In the Listen and Transmit protocol we assumed that relay nodes have only local knowledge of the channels, i.e., they only know their connections to the transmitter and receiver nodes.

We addressed another scenario in the previous sections, where the nodes have complete knowledge of all the channel coefficients and try to diagonalize the channel matrix between the transmitters and their corresponding receivers. In this subsection, we analyze the effect of perfect knowledge of the channel on the power efficiency achieved by diagonalizing the channel matrix.

In this subsection, we make an additional assumption that the channel coefficients are independent complex random variables with zero mean and unit variance. Using (25) and (28), we can describe our protocol by the following matrix relations:

$$
\begin{aligned}
\boldsymbol{r} &= \boldsymbol{s}\boldsymbol{G} + \boldsymbol{v} \\
\boldsymbol{t} &= \boldsymbol{r}\boldsymbol{D} \\
\boldsymbol{y} &= \boldsymbol{s}\boldsymbol{G}\boldsymbol{D}\boldsymbol{H} + \boldsymbol{v}\boldsymbol{D}\boldsymbol{H} + \boldsymbol{w} \\
p\operatorname{tr}\left(\boldsymbol{D}^*\boldsymbol{G}^*\boldsymbol{G}\boldsymbol{D}\right) &+ \sigma_n^2 \operatorname{tr}\left(\boldsymbol{D}^*\boldsymbol{D}\right) = n\sigma_r^2
\end{aligned} \tag{61}
$$

where $\boldsymbol{s} \in \mathbb{C}^{1 \times r}$ is the transmitted vector, $\boldsymbol{y} \in \mathbb{C}^{1 \times r}$ is the received vector and $\boldsymbol{r}, \boldsymbol{t} \in \mathbb{C}^{1 \times n}$ are the respective received and transmitted vectors at the relay stage. $\boldsymbol{w} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}_r)$ and $\boldsymbol{v} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}_n)$ are the corresponding vectors of noise introduced at the receivers and at the relay stage respectively. $\boldsymbol{G} \in \mathbb{C}^{r \times n}$ is the channel matrix between the transmitters and the relay nodes and $\boldsymbol{H} \in \mathbb{C}^{n \times r}$ is the channel matrix between the relay nodes and the receivers. Finally, $\boldsymbol{D} = \operatorname{diag}(d_1, \ldots, d_n) \in \mathbb{C}^{n \times n}$ is a diagonal matrix with diagonal entries corresponding to the scalars chosen by the relay nodes. Notice that the d_j's depend on the channel gains. The last equation in (61) is a consequence of the power constraint for the relay nodes. We remark that the power constraint considered here is more general than what was assumed in previous sections.

From (61), the equivalent channel matrix between the transmitters and the receivers is $\boldsymbol{G}\boldsymbol{D}\boldsymbol{H}$. Therefore, diagonalizing the channel matrix amounts to finding diagonal matrix $\boldsymbol{D}$ such that $\boldsymbol{G}\boldsymbol{D}\boldsymbol{H} = \alpha\boldsymbol{I}_r$ for some complex scalar α. The number of complex equations is r^2 and the number of variables is n. Therefore, generically, this equation has a solution for $r^2 \leq n$. In this case, by looking at (61) we can write the received signal at receive node k as

$$
y_k = \alpha s_k + w_k + \sum_{j=1}^{n} h_{jk} d_j v_j.
$$

We can find an upper bound on the achievable rates using the scheme described above by considering that the receiver node has knowledge of the different noises introduced in the relay stage. Hence, we can bound the capacity of the channel between the transmit/receive pair k as

$$
C_k \leq \frac{1}{2}\mathrm{E_f}\log\left(1 + \frac{|\alpha|^2 p}{\sigma_n^2}\right) \leq \frac{1}{2}\log\left(1 + \frac{p\mathrm{E_f}|\alpha|^2}{\sigma_n^2}\right).
$$

The power efficiency can be bounded as follows:

$$
\eta' = \frac{2\sum_{i=1}^{r} C_i}{rp + n\sigma^2} \leq \frac{2r}{rp + n\sigma_r^2}\log\left(1 + \frac{\mathrm{E_f}|\alpha|^2 p}{\sigma_n^2}\right). \tag{62}
$$

Thus, we only need to find the mean of the maximum possible value of $|\alpha|^2$ subject to the following constraints:

$$
\begin{aligned}
\boldsymbol{G}\boldsymbol{D}\boldsymbol{H} &= \alpha\boldsymbol{I}_r \\
p\operatorname{tr}(\boldsymbol{D}^*\boldsymbol{G}^*\boldsymbol{G}\boldsymbol{D}) &+ \sigma_n^2\operatorname{tr}(\boldsymbol{D}^*D) = n\sigma_r^2.
\end{aligned} \tag{63}
$$

First, we try to solve the first equation in (63). Define $\boldsymbol{d} = [d_1, \ldots, d_n]^t$, $\boldsymbol{b} = \operatorname{vec}(\boldsymbol{I}_r)$. Equation (63) can be written in terms of $\boldsymbol{d}$ and $\boldsymbol{b}$ as

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{d} &= \alpha\boldsymbol{b} \\
\boldsymbol{d}^*\boldsymbol{\Lambda}\boldsymbol{d} &= n\sigma_r^2
\end{aligned} \tag{64}
$$

where it can be easily verified that $\boldsymbol{A} \in \mathbb{C}^{r^2 \times n}$ and $\boldsymbol{\Lambda} \in \mathbb{C}^{n \times n}$ are

$$
\boldsymbol{A}_{r^2 \times n} = \begin{bmatrix} \boldsymbol{H}^T & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}^T & \cdots & \boldsymbol{0} \\ \vdots & \cdots & \cdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{H}^T \end{bmatrix}_{r^2 \times rn} \cdot \begin{bmatrix} \operatorname{diag}(\boldsymbol{G}_1) \\ \operatorname{diag}(\boldsymbol{G}_2) \\ \vdots \\ \operatorname{diag}(\boldsymbol{G}_r) \end{bmatrix}_{rn \times n},
$$

$$
\boldsymbol{\Lambda}_{n \times n} = \operatorname{diag}\left(p\|\boldsymbol{g}_1\|^2 + \sigma_n^2, \ldots, p\|\boldsymbol{g}_n\|^2 + \sigma_n^2\right). \tag{65}
$$

$\boldsymbol{G}_i$, $i = 1, \ldots, r$, and $\boldsymbol{g}_j$, $j = 1, \ldots, n$, denote the ith row and jth column of $\boldsymbol{G}$, respectively. If we define $\boldsymbol{B}_{r^2 \times n} = \boldsymbol{A}\boldsymbol{\Lambda}^{-\frac{1}{2}}$, by using QR-type decomposition [22] for $\boldsymbol{B}^*$ we can write $\boldsymbol{B}^* = \boldsymbol{Q}\begin{pmatrix}\boldsymbol{R} \\ \boldsymbol{0}\end{pmatrix}$, where $\boldsymbol{Q}_{n \times n}$ is unitary matrix (i.e., $\boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{I}_n$) and $\boldsymbol{R}_{r^2 \times r^2}$ is a lower triangular matrix with diagonal elements equal to unity. By writing

$$
\boldsymbol{Q}^*\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{d} = \begin{pmatrix} \boldsymbol{d}^{(1)}_{r^2 \times 1} \\ \boldsymbol{d}^{(2)}_{(n-r^2) \times 1} \end{pmatrix}
$$

we can rewrite (64) as

$$
\begin{aligned}
\boldsymbol{R}^*\boldsymbol{d}^{(1)} &= \alpha\boldsymbol{b} \\
\left\|\boldsymbol{d}^{(1)}\right\|^2 + \left\|\boldsymbol{d}^{(2)}\right\|^2 &= n\sigma_r^2.
\end{aligned} \tag{66}
$$

Now notice that $\boldsymbol{R}$ is invertible and therefore we can find $\boldsymbol{d}^{(1)}$ from (66) and by substituting its value in the second relation of (66) we get

$$
|\alpha|^2 \boldsymbol{b}^*(\boldsymbol{R}^*\boldsymbol{R})^{-1}\boldsymbol{b} + \|\boldsymbol{d}^{(2)}\|^2 = n\sigma_r^2.
$$

It can be easily verified that $\boldsymbol{R}^*\boldsymbol{R} = \boldsymbol{B}\boldsymbol{B}^* = \boldsymbol{A}\boldsymbol{\Lambda}^{-1}\boldsymbol{A}^*$. Also, the maximum of $|\alpha|^2$ is when $\boldsymbol{d}^{(2)} = \boldsymbol{0}_{n-r^2 \times 1}$. Therefore,

$$
|\alpha|^2_{\max} = \frac{n\sigma_r^2}{\boldsymbol{b}^*(\boldsymbol{A}\boldsymbol{\Lambda}^{-1}\boldsymbol{A}^*)^{-1}\boldsymbol{b}}. \tag{67}
$$

Now using the following inequality for positive definite matrix B and any vector x ([22, p. 452])

$$
(\boldsymbol{x}^*\boldsymbol{B}\boldsymbol{x})(\boldsymbol{x}^*\boldsymbol{B}^{-1}\boldsymbol{x}) \geq (\boldsymbol{x}^*\boldsymbol{x})^2
$$
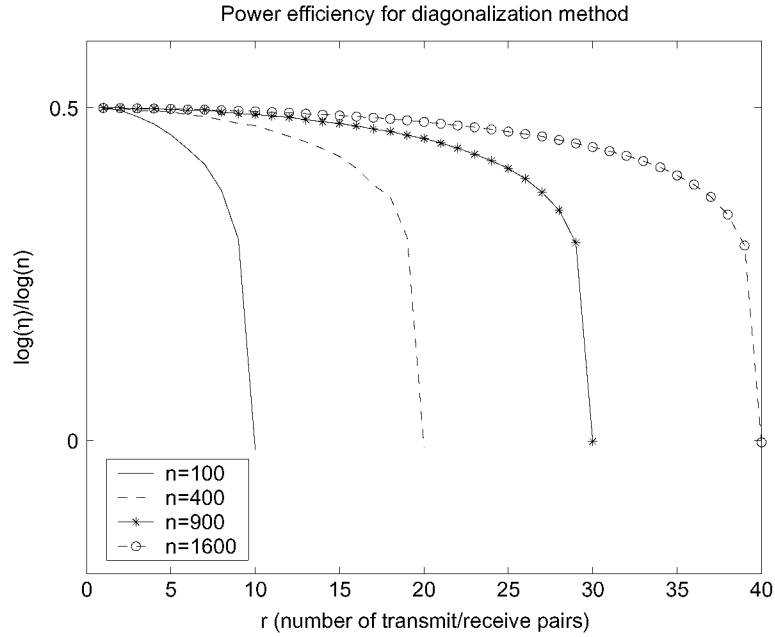
Fig. 4. Power efficiency for interference suppression case.

we have

$$|\alpha|^2_{\max} \leq \frac{n\sigma_r^2(\boldsymbol{b}^*(\boldsymbol{A}\boldsymbol{\Lambda}^{-1}\boldsymbol{A}^*)\boldsymbol{b})}{(\boldsymbol{b}^*\boldsymbol{b})^2}.$$

$\boldsymbol{b}^*\boldsymbol{b} = r$ and $\frac{1}{\sigma_{\bar{n}}^2}\boldsymbol{I}_n - \boldsymbol{\Lambda}^{-1}$ is positive semidefinite. Therefore, we have

$$E_f|\alpha|^2_{\max} \leq \frac{n\sigma_r^2}{r^2\sigma_n^2}E_f[\boldsymbol{b}^*(\boldsymbol{A}\boldsymbol{A}^*)\boldsymbol{b}]. \tag{68}$$

Using the structure of matrix $\boldsymbol{A}$ in (65), it can be verified that

$$E_f[\boldsymbol{b}^*(\boldsymbol{A}\boldsymbol{A}^*)\boldsymbol{b}] = \sum_{i,j=1}^{r} E_f[\boldsymbol{b}_i^*\boldsymbol{H}^T \text{diag}(\boldsymbol{G}_i)\text{diag}^*(\boldsymbol{G}_j)\bar{\boldsymbol{H}}\boldsymbol{b}_j]$$

where b_i is the ith column of $\boldsymbol{I}_r$. Now since the entries of $\boldsymbol{H}$ and $\boldsymbol{G}$ are independent from each other, the expectation inside the above summation is zero for $i \neq j$. Therefore, we have

$$\begin{aligned}
E_f[\boldsymbol{b}^*(\boldsymbol{A}\boldsymbol{A}^*)\boldsymbol{b}] &= \sum_{i} E_f[\boldsymbol{b}_i^*(\boldsymbol{H}^*\boldsymbol{H})^T\boldsymbol{b}_i] \\
&= E_f \text{tr}\left((\boldsymbol{H}^*\boldsymbol{H})^T \cdot \sum_{i=1}^{r} \boldsymbol{b}_i\boldsymbol{b}_i^*\right) \\
&= E_f \text{tr}((\boldsymbol{H}^*\boldsymbol{H})^T \cdot \boldsymbol{I}_r) \\
&= nr
\end{aligned} \tag{69}$$

where we have used the fact that $\text{tr}(\boldsymbol{A}\boldsymbol{B}) = \text{tr}(\boldsymbol{B}\boldsymbol{A})$. Using the above result in (68), we have

$$E_f|\alpha|^2_{\max} \leq \frac{n^2\sigma_r^2}{r\sigma_n^2}. \tag{70}$$

Combining (70) and (62) we have

$$\eta' \leq \frac{r}{rp + n\sigma_r^2} \log\left(1 + \frac{n^2\sigma_r^2 p}{r\sigma_n^2}\right). \tag{71}$$

Using an argument to previous sections (e.g., Section IV-C), we know that the maximum of the right-hand side expression is less than $\kappa\sqrt{n}$ for some constant κ dependent on σ_n^2. Therefore,

$$\eta' \leq \kappa\sqrt{n}$$

which is the same as the case when we have only local knowledge of the channel and the Listen and Transmit protocol is used. We have found the upper bound for the maximum power efficiency, η', using the actual value for $E_f|\alpha|^2$ from (67) and Matlab simulation. We have plotted the ratio $\frac{\log(\eta')}{\log n}$ for different values of $r \leq \sqrt{n}$ for $n = 100, 400, 900$, and 1600. As we can see from the plots, the upper bound suggests that we cannot do better than $\sqrt{n}$ (or equivalently $\frac{1}{2}$ in Fig. 4). Also, as the number of the simultaneous transmitter/receiver pairs r increases, the upper bound on the power efficiency of the interference cancellation method becomes smaller. This suggest that this method is not power efficient.

Based on the preceding argument, we have the following theorem.

Theorem 6: Consider a wireless *ad hoc* network with n relay nodes and r transmit/receive pair in which $r^2 \leq n$. Moreover, assume the channel coefficients can be modeled by independent zero mean unit variance complex random variables. Also assume that the relay nodes have *complete knowledge* of the channel coefficients; then if the relay nodes cancel out the interference at the receiver nodes the power efficiency scales as $O(\sqrt{n})$.

VI. CONCLUSION

In this paper, we address the power efficiency of random sensory and *ad hoc* wireless networks formed in a domain of fixed area. Under some assumptions on the moments of the channel coefficients, we show that asymptotically, as the number of nodes in the network, n, grows larger, with high probability we can achieve a power efficiency of $\Theta(\sqrt{n})$ for sensory networks. For *ad hoc* networks if the number of transmitter/receiver pairs is of $O(\sqrt{n})$, we can achieve the same result. We also described the protocol used to achieve this power efficiency. Although the best results for capacity per node in sensory and *ad hoc* wireless networks decrease as the size of networks grows larger [1], [2], we can see that it pays off to consider these networks in terms of power efficiency.

We can think of the protocol used in this paper as a simple yet powerful memoryless linear coding scheme for the relay nodes, i.e., the relay nodes simply relay a scaled version of what they have heard. One can generalize this protocol by using other coding schemes for relay nodes. Another interesting problem is to look at the spectral efficiency of sensory and *ad hoc* networks and its tradeoff with the power efficiency. Also whether a power efficiency of greater than $\Theta(\sqrt{n})$ is possible or not is still an open problem.

APPENDIX

A. Average Power of the Nodes at the Relay Stage

In this appendix, we will show that using d_j, as defined in (26), the average of the power of the signal transmitted by relay nodes over all the point placements is σ_r^2. First, notice that the transmitted signal by the jth relay node is

$$t_j = \sigma_r d_j r_j = \frac{\sigma_r \sum_{l=1}^{r} g_{lj}^* h_{jl}^* \left(\sum_{i=1}^{r} g_{ij} s_i + v_j \right)}{\sqrt{r \kappa_1 \sigma_n^2 + r((r-1)\kappa_2 + \kappa_3)p}} \quad \text{(A1)}$$

Since v_j's and s_i's are independent and zero mean, we can write the average power of t_j over the fading and location of the nodes as (A2) (shown at the bottom of the page). Thus, we only need to compute the expectations in the numerator. We have

$$\mathrm{E}\left|\sum_{l=1}^{r} g_{lj}^* h_{jl}^*\right|^2$$

$$= \sum_{l,l'}^{r} \mathrm{E}[g_{lj}^* h_{jl}^* g_{l'j} h_{jl'}]$$

$$= \sum_{l,l'}^{r} \mathrm{E}_{\{j\}} \mathrm{E}_{\{l,l'|j\}} [\mathrm{E}_{\mathrm{f}}[g_{lj} g_{l'j}^*] \mathrm{E}_{\mathrm{f}}[h_{jl'}^* h_{jl}]]$$

$$= \sum_{l,l'}^{r} \mathrm{E}_{\{j\}} (\mathrm{E}_{\{l,l'|j\}} [\mathrm{E}_{\mathrm{f}}[g_{lj} g_{l'j}^*]] \mathrm{E}_{\{l,l'|j\}} [\mathrm{E}_{\mathrm{f}}[h_{jl'}^* h_{jl}]])$$

$$\overset{(a)}{=} \sum_{l=1}^{r} \mathrm{E}_{\{j\}} [\mathrm{E}_{\{l|j\}} \mathrm{E}_{\mathrm{f}} |g_{lj}|^2]^2$$

$$= r \mathrm{E}_{\{j\}} [\mathrm{E}_{\{l|j\}} \mathrm{E}_{\mathrm{f}} |g_{lj}|^2]^2 = r \kappa_1 \quad \text{(A3)}$$

where (a) follows from the fact that for $l \neq l'$ the channels between j and l, l' are independent from each other. Therefore, we will have a term like the spatial averaging of a channel coefficient, i.e., $\mathrm{E}_{\{l|j\}} \mathrm{E}_{\mathrm{f}}[g_{lj}]$, which by assumption A is zero

$$\mathrm{E}\left|g_{ij} \sum_{l=1}^{r} g_{lj}^* h_{jl}^*\right|^2$$

$$= \sum_{l,l'}^{r} \mathrm{E}[|g_{ij}|^2 g_{lj}^* h_{jl}^* g_{l'j} h_{jl'}]$$

$$= \sum_{l,l'}^{r} \mathrm{E}_{\{j\}} \mathrm{E}_{\{l,l'|j\}} [\mathrm{E}_{\mathrm{f}}[|g_{ij}|^2 g_{lj} g_{l'j}^*] \mathrm{E}_{\mathrm{f}}[h_{jl'}^* h_{jl}]]$$

$$= \sum_{l,l'}^{r} \mathrm{E}_{\{j\}} (\mathrm{E}_{\{l,l'|j\}} [\mathrm{E}_{\mathrm{f}}[|g_{ij}|^2 g_{lj} g_{l'j}^*]] \mathrm{E}_{\{l,l'|j\}} [\mathrm{E}_{\mathrm{f}}[h_{jl'}^* h_{jl}]])$$

$$\overset{(a)}{=} \sum_{l=1}^{r} \mathrm{E}_{\{j\}} [\mathrm{E}_{\{l|j\}} \mathrm{E}_{\mathrm{f}} |g_{ij} g_{lj}|^2]^2$$

$$= (r-1) \mathrm{E}_{\{j\}} [\mathrm{E}_{\{l|j\}} \mathrm{E}_{\mathrm{f}} |g_{lj}|^2]^3$$

$$+ \mathrm{E}_{\{j\}} [\mathrm{E}_{\{i|j\}} \mathrm{E}_{\mathrm{f}} |h_{ji}|^2 \mathrm{E}_{\{i|j\}} \mathrm{E}_{\mathrm{f}} |g_{ij}|^4] = (r-1)\kappa_2 + \kappa_3$$

where again (a) follows from the fact that for $l \neq l'$, h_{jl} and $h_{jl'}$ are independent and so the term in the preceding equation can be written in terms of their spatial averaging which by assumption is zero. By substituting the values from (A3) and the above equation in (A2), we see that the average of the relay node power over all the random networks is σ_r^2.

B. Proof of Lemma 1

We intend to find probabilistic bounds for the $\mathrm{Var}_{\mathrm{f}}[w_k{}']$ and $\mathrm{E}_{\mathrm{f}}[\alpha_k^k]$ for all k. For this, we will use Markov's and Chebyshev's inequalities and also assumptions A, B, (and/or C).

By using Chebyshev's inequality we have the following inequalities:

$$\mathrm{P}\{E_k^1\} = \mathrm{P}\left\{\left|\frac{\sum_{j=1}^{n} \mathrm{E}_{\mathrm{f}} |h_{jk} d_j|^2}{n} - \mathrm{E}_{\mathrm{loc}} \mathrm{E}_{\mathrm{f}} |h_{jk} d_j|^2\right| > \epsilon_1\right\}$$

$$\leq \frac{\mathrm{Var}_{\mathrm{loc}} \mathrm{E}_{\mathrm{f}} |h_{jk} d_j|^2}{n \epsilon_1^2}$$

$$\mathrm{P}\{E_{k,k}^2\} = \mathrm{P}\left\{\left|\frac{\sum_{j=1}^{n} \mathrm{E}_{\mathrm{f}} |g_{kj} h_{jk} d_j|^2}{n}\right.\right.$$

$$\mathrm{E}[|t_j|^2] = \frac{\sigma_r^2 \left(\mathrm{E}\left|\sum_{l=1}^{r} g_{lj}^* h_{jl}^*\right|^2 \sigma_n^2 + p \sum_{i=1}^{r} \mathrm{E}\left[|g_{ij}|^2 \left|\sum_{l=1}^{r} g_{lj}^* h_{jl}^*\right|^2 \right] \right)}{r \kappa_1 \sigma_n^2 + r((r-1)\kappa_2 + \kappa_3)p}. \quad \text{(A2)}$$

$$\left. -\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|g_{kj}h_{jk}d_j|^2 \right| > \epsilon_2 \right\}$$

$$\leq \frac{\mathrm{Var}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|g_{kj}h_{jk}d_j|^2}{n\epsilon_2^2}$$

$$\mathrm{P}\{E_{i,k}^2\} = \mathrm{P}\left\{ \left| \frac{\sum_{j=1}^n \mathrm{E}_{\mathrm{f}}|g_{ij}h_{jk}d_j|^2}{n} \right.\right.$$

$$\left.\left. -\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|g_{ij}h_{jk}d_j|^2 \right| > \epsilon_3 \right\}$$

$$\leq \frac{\mathrm{Var}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|g_{ij}h_{jk}d_j|^2}{n\epsilon_3^2}$$

$$\mathrm{P}\{E_k^3\} = \mathrm{P}\left\{ \left| \frac{\sum_{j=1}^n \mathrm{E}_{\mathrm{f}}[g_{kj}h_{jk}d_j]}{n} \right.\right.$$

$$\left.\left. -\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}[g_{kj}h_{jk}d_j] \right| > \epsilon_4 \right\}$$

$$\leq \frac{\mathrm{Var}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}[g_{kj}h_{jk}d_j]}{n\epsilon_4^2} \qquad \text{(B1)}$$

Also using Markov's and Chebyshev's inequalities, we have

$$\mathrm{P}\{E_k^4\} = \mathrm{P}\left\{ \sum_{i=1, i\neq k}^r |\sum_{j=1}^n \mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]|^2 > n^\rho r^\zeta \epsilon_5 \right\}$$

$$\leq \frac{m_k}{n^\rho r^\zeta \epsilon_5} \qquad \text{(B2)}$$

$$\mathrm{P}\{E_k^5\} = \mathrm{P}\left\{ \left| \sum_{i=1, i\neq k}^r |\sum_{j=1}^n \mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]|^2 - m_k \right| > n^\rho r^\zeta \epsilon_5 \right\}$$

$$\leq \frac{V_k}{n^{2\rho} r^{2\zeta} \epsilon_5^2} \qquad \text{(B3)}$$

where m_k and V_k are the mean and the variance of

$$\sum_{i=1, i\neq k}^r |\sum_{j=1}^n \mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]|^2$$

respectively. Throughout this appendix we define

$$A = \sqrt{r(p(\kappa_3 + (r-1)\kappa_2) + \sigma_n^2)}.$$

We should remind that d_j defined in (26) of Section V is equal to

$$d_j = \frac{\sum_{i=1}^r g_{ij}^* h_{ji}^*}{\sqrt{r\kappa_1\sigma_n^2 + r((r-1)\kappa_2 + \kappa_3)p}} = \frac{\sum_{i=1}^r g_{ij}^* h_{ji}^*}{A}. \qquad \text{(B4)}$$

Note: In this appendix, we are using the same index for designating the ith receiver and the ith transmitter. Both of them are addressed by index i. In our expectations, whenever i appears as an index of a channel coefficient denoted by h (for instance, h_{ji}), we are addressing the ith receiver and when i appears as an index of a channel coefficient denoted by g (for instance, g_{il}), we are addressing the ith transmitter. For instance, in the expectation $\mathrm{E}_{\{l',l,k|j\}}\mathrm{E}_{\mathrm{f}}[h_{jl}^* h_{jl'}|h_{jk}|^2]$, we are averaging over the

locations of the l'th, lth, and kth receivers, where in $\mathrm{E}_{\{l|j\}}\mathrm{E}_{\mathrm{f}}[g_{lj}^*]$ we are averaging over the location of the lth transmitter. Therefore, the expectations will be clear by looking at the channel coefficients that appears in the formula.

For brevity, we only include the derivation of $\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$, $\mathrm{Var}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$, m_k, and V_k here. The derivation of all the other moments appearing in (B1) is very similar to the ones calculated here.

1) Computing $\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$ and $\mathrm{Var}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$: For the mean term we have

$$\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$$

$$= \frac{1}{A^2}\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}\left| \sum_{l=1}^r g_{lj}^* h_{jl}^* h_{jk} \right|^2$$

$$= \frac{1}{A^2} \sum_{l,l'} \mathrm{E}[g_{lj}^* h_{jl}^* g_{l'j} h_{jl'}|h_{jk}|^2]$$

$$= \frac{1}{A^2}\left(\sum_{l=1}^r \mathrm{E}[|g_{lj}h_{jl}h_{jk}|^2] + \sum_{l\neq l'}\mathrm{E}[g_{lj}^* h_{jl}^* g_{l'j} h_{jl'}|h_{jk}|^2] \right)$$

$$= \frac{1}{A^2}\left(\sum_{l=1}^r \mathrm{E}_{\{j\}}\left[(\mathrm{E}_{\{l,k|j\}}\mathrm{E}_{\mathrm{f}}[|g_{lj}|^2]) \cdot (\mathrm{E}_{\{l,k|j\}}\mathrm{E}_{\mathrm{f}}[|h_{jl}||h_{jk}|^2]) \right] \right.$$

$$+ \sum_{l\neq l'}\mathrm{E}_{\{j\}}\left[(\mathrm{E}_{\{l|j\}}\mathrm{E}_{\mathrm{f}}[g_{lj}^*]) \right.$$

$$\left.\left. \cdot (\mathrm{E}_{\{l'|j\}}\mathrm{E}_{\mathrm{f}}[g_{l'j}]) \cdot (\mathrm{E}_{\{l',l,k|j\}}\mathrm{E}_{\mathrm{f}}[h_{jl}^* h_{jl'}|h_{jk}|^2]) \right] \right) \qquad \text{(B5)}$$

where we have substituted d_j by its value from (B4). Now notice that by assumption A, $\mathrm{E}_{\{l|j\}}\mathrm{E}_{\mathrm{f}}[g_{lj}^*] = 0$, so the second summation is zero. Thus, we can write

$$\mathrm{E}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$$

$$= \frac{1}{A^2}((r-1)\mathrm{E}_{\{j\}}[\mathrm{E}_{\{l|j\}}\mathrm{E}_{\mathrm{f}}|g_{lj}|^2]^3$$

$$+ \mathrm{E}_{\{j\}}[\mathrm{E}_{\{k|j\}}\mathrm{E}_{\mathrm{f}}|g_{kj}|^2\mathrm{E}_{\{k|j\}}\mathrm{E}_{\mathrm{f}}|h_{jk}|^4)$$

$$= \frac{(r-1)\kappa_2 + \kappa_3}{rp((r-1)\kappa_2 + \kappa_3) + r\kappa_1\sigma_n^2} \overset{(a)}{<} \frac{\kappa}{r\kappa p + \kappa_1\sigma_n^2} \qquad \text{(B6)}$$

where $\kappa = \max\{\kappa_2, \kappa_3\}$. Also, (a) follows from the fact that $f(x) = \frac{x}{L_1^2 x + L_2^2}$ is increasing in x. For the variance we have

$$\mathrm{Var}_{\mathrm{loc}}\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2$$

$$\leq \mathrm{E}_{\mathrm{loc}}[\mathrm{E}_{\mathrm{f}}|h_{jk}d_j|^2]^2$$

$$= \frac{1}{A^4}\mathrm{E}_{\mathrm{loc}}\left(\sum_{l,l'}\mathrm{E}_{\mathrm{f}}[|h_{jk}|^2 g_{lj}^* g_{l'j} h_{jl}^* h_{jl'}] \right)^2$$

$$= \frac{1}{A^4}\mathrm{E}_{\{j\}}\sum_{l_i, i=1\ldots4}\mathrm{E}_{\{l_i,k|j\}}\left[(\mathrm{E}_{\mathrm{f}}[|h_{jk}|^2 g_{l_1 j}^* g_{l_2 j} h_{jl_1}^* h_{jl_2}]) \right.$$

$$\left. \cdot (\mathrm{E}_{\mathrm{f}}[|h_{jk}|^2 g_{l_3 j} g_{l_4 j}^* h_{jl_3} h_{jl_4}^*]) \right]$$

$$\overset{(b)}{=} \frac{1}{A^4}\mathrm{E}_{\{j\}}\sum_{l_i, i=1\ldots4}(\mathrm{E}_{\{l_i,k|j\}}[\mathrm{E}_{\mathrm{f}}[|h_{jk}|^2 h_{jl_1}^* h_{jl_2}]$$

$$\cdot \mathrm{E}_{\mathrm{f}}[|h_{jk}|^2 h_{jl_3} h_{jl_4}^*]]$$

$$\cdot \underbrace{\mathrm{E}_{\{l_i|j\}}[\mathrm{E}_{\mathrm{f}}[g_{l_1 j}^* g_{l_2 j}]\mathrm{E}_{\mathrm{f}}[g_{l_3 j} g_{l_4 j}^*]])}_{(c)}$$

where we have plugged in the value of d_j from (B4). Equality (b) follows from the fact that once the relay node is fixed, the coefficients to the transmitters and coefficient to the receivers are independent. Now if at least one l_t is different from the others then because of assumption A in Section II-B the (c) term will be zero,. The only case when the (c) term in the above summation is nonzero is for $l_{i_1} = l_{i_2}, l_{i_3} = l_{i_4}$ where (i_1, i_2, i_3, i_4) is a permutation of $(1, 2, 3, 4)$. Thus, there are at most $r + 3r(r - 1)$ nonzero terms in the summation. Note that each of the terms in the summation above depends only on the channel coefficients between at most six random points and thus depends only on the geometry of the domain and not on n or r. If we consider B_1 to be an upper bound for these terms then the variance can be bounded by

$$\text{Var}_{\text{loc}} \text{E}_{\text{f}} |h_{jk} d_j|^2 \leq \frac{(r + 3r(r - 1))B_1}{r^2((p(r - 1)\kappa_2 + \kappa_3) + \kappa_1 \sigma_n^2)^2}$$
$$\leq \frac{3B_1}{(r\kappa' p + \kappa_1 \sigma_n^2)^2} \quad \text{(B7)}$$

where $\kappa' = \min\{\kappa_2, \kappa_3\}$.

2) Computing m_k and V_k: In this appendix, we will bound m_k and V_k. The derivation in this part is rather long but the idea is essentially like that of the previous parts. We start with m_k (as defined in the equation at the top of the following page), where the value of d_j is substituted from (B4). Since $j \neq j'$ and $i \neq k$, using assumptions A and B, it can be easily verified that each of the $\star$ terms is zero. Therefore, the first summation does not contribute to m_k and we can simplify the above equation to

$$m_k = \sum_{i=1, i \neq k}^{r} \sum_{j=1}^{n} \text{E}_{\text{loc}} |\text{E}_{\text{f}}[g_{ij} h_{jk} d_j]|^2 \quad \text{(B8)}$$

By substituting d_j in the above equation with its value from (B4) we have the second equation at the top of the following page, where in (a) we have used assumption A of Section II-B and the fact that since $i \neq k$ the only possibility for l, l' that results in a nonzero term in the summation is when they are equal. Inequality (b) is also a consequence of the fact that the expectations in the last equality do not depend on r and n. Hence, they can be upper-bounded by a constant B_5.

For V_k we have

$$V_k + m_k^2 = \text{E}_{\text{loc}} \left[\sum_{i=1, i \neq k}^{r} \left| \sum_{j=1}^{n} \text{E}_{\text{f}}[g_{ij} h_{jk} d_j] \right|^2 \right]^2$$

$$= \text{E}_{\text{loc}} \left[\sum_{i=1, i \neq k}^{r} \sum_{j, j'} \text{E}_{\text{f}}[g_{ij} h_{jk} d_j] \text{E}_{\text{f}}[g_{ij'}^* h_{j'k}^* d_{j'}^*] \right]^2$$

$$= \frac{1}{A^4} \text{E}_{\text{loc}} \left[\sum_{i=1, i \neq k}^{r} \sum_{j, j'=1}^{n} \sum_{l, l'=1}^{r} \text{E}_{\text{f}} \right.$$
$$\left. \times [g_{ij}^* h_{jk}^* h_{jl} g_{lj}] \text{E}_{\text{f}}[g_{ij'} h_{j'k} h_{j'l'}^* g_{l'j'}^*] \right]^2$$

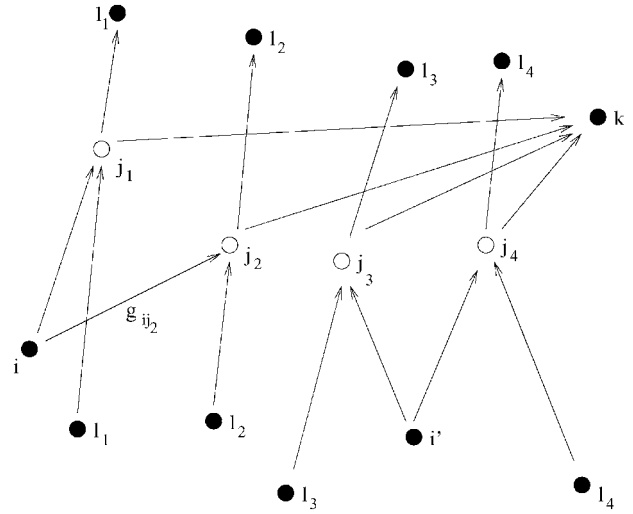Fig. 5. A generic setting of the points for computing V_k.

$$= \frac{1}{A^4} \text{E}_{\text{loc}} \left[\sum_{i, i'=1, i, i' \neq k}^{r} \sum_{j_1, j_2, j_3, j_4=1}^{n} \sum_{l_1, l_2, l_3, l_4=1}^{r} \text{E}_{\text{f}} \right.$$
$$\times [g_{ij_1}^* h_{j_1 k}^* h_{j_1 l_1} g_{l_1 j_1}] \text{E}_{\text{f}}[g_{ij_2} h_{j_2 k} h_{j_2 l_2}^* g_{l_2 j_2}^*]$$
$$\left. \times \text{E}_{\text{f}}[g_{i'j_3}^* h_{j_3 k}^* h_{j_3 l_3} g_{l_3 j_3}] \text{E}_{\text{f}}[g_{i'j_4} h_{j_4 k} h_{j_4 l_4}^* g_{l_4 j_4}^*] \right]$$

$$= \frac{1}{A^4} \sum_{\mathcal{I}} \text{E}_{\text{loc}} \left[\text{E}_{\text{f}}[g_{ij_1}^* g_{l_1 j_1}] \text{E}_{\text{f}}[h_{j_1 k}^* h_{j_1 l_1}] \right.$$
$$\times \text{E}_{\text{f}}[g_{ij_2} g_{l_2 j_2}^*] \text{E}_{\text{f}}[h_{j_2 k} h_{j_2 l_2}^*]$$
$$\times \text{E}_{\text{f}}[g_{i'j_3}^* g_{l_3 j_3}] \text{E}_{\text{f}}[h_{j_3 k}^* h_{j_3 l_3}]$$
$$\left. \times \text{E}_{\text{f}}[g_{i'j_4} g_{l_4 j_4}^*] \text{E}_{\text{f}}[h_{j_4 k} h_{j_4 l_4}^*] \right] \quad \text{(B9)}$$

where $\mathcal{I} = \{i, i', j_1, j_2, j_3, j_4, l_1, l_2, l_3, l_4\}$ is the set of indices representing transmitters, relay nodes, and receivers, respectively. The summation is over all possible choices of $\mathcal{I}$ conditioned that $i, i' \neq k$. In the following paragraphs, we will look at the cases when the expectation terms in the above summation is nonzero. A generic configuration of the points is depicted in Fig. 5.

First, if at least one j_t is distinct from the others then the expectation in the sum will be zero. This can be verified by fixing the location of the four nodes connected to j_t and taking the expectation over the placement of j_t. Since $i, i' \neq k$ the expectation will be of the form given in assumption C of Section II-B which is zero. Therefore, we should have $j_{t_1} = j_{t_2}, j_{t_3} = j_{t_4}$, where (t_1, t_2, t_3, t_4) is a permutation of $(1, 2, 3, 4)$. Using a similar argument, it can be shown that the expectation in the above summation is nonzero, iff $l_{t'_1} = l_{t'_2}, l_{t'_3} = l_{t'_4}$ for some permutation of indices in set $\{1, 2, 3, 4\}$.

Due to the symmetry in the connections between the relay nodes and the transmitters and receivers (Fig. 5), we only consider two cases among the three possible combinations of $j_{t_1} = j_{t_2}, j_{t_3} = j_{t_4}$. In fact, the cases where (t_1, t_2, t_3, t_4) is $(1, 3, 2, 4)$ or $(1, 4, 2, 3)$ result in the same expression. Therefore, we only need to compute the expression for

$$m_k = \mathrm{E}_{\mathrm{loc}}\left[\sum_{i=1,i\neq k}^{r}\left|\sum_{j=1}^{n}\mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]\right|^2\right]$$

$$= \sum_{i=1,i\neq k}^{r}\sum_{j,j'}\mathrm{E}_{\mathrm{loc}}[\mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]\mathrm{E}_{\mathrm{f}}[g_{ij'}^{*}h_{j'k}^{*}d_{j'}^{*}]]$$

$$= \sum_{i=1,i\neq k}^{r}\sum_{j=1}^{n}\mathrm{E}_{\mathrm{loc}}|\mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]|^2 + \frac{1}{A^2}\sum_{i=1,i\neq k}^{r}\sum_{j\neq j'}^{n}\sum_{l,l'}^{r}\mathrm{E}_{\mathrm{loc}}[\mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}h_{jl}^{*}g_{lj}^{*}]\cdot\mathrm{E}_{\mathrm{f}}[g_{ij'}^{*}h_{j'k}^{*}h_{j'l'}g_{l'j'}]]$$

$$= \frac{1}{A^2}\sum_{i=1,i\neq k}^{r}\sum_{j\neq j'}^{n}\sum_{l,l'=1}^{r}\mathrm{E}_{\{j,j'\}}\underbrace{\mathrm{E}_{\{l',l,k,i|j,j'\}}[\mathrm{E}_{\mathrm{f}}[g_{ij}g_{lj}^{*}]\cdot\mathrm{E}_{\mathrm{f}}[h_{jk}h_{jl}^{*}]\cdot\mathrm{E}_{\mathrm{f}}[g_{ij'}^{*}g_{l'j'}]\cdot\mathrm{E}_{\mathrm{f}}[h_{j'k}^{*}h_{j'l'}]]}_{\star}$$

$$+ \sum_{i=1,i\neq k}^{r}\sum_{j=1}^{n}\mathrm{E}_{\mathrm{loc}}|\mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]|^2$$

$$m_k = \sum_{i=1,\neq k}^{r}\sum_{j=1}^{n}\mathrm{E}_{\mathrm{loc}}|\mathrm{E}_{\mathrm{f}}[g_{ij}h_{jk}d_j]|^2$$

$$= \frac{1}{A^2}\sum_{i=1,i\neq k}^{r}\sum_{j=1}^{n}\sum_{l,l'}^{r}\mathrm{E}_{\{j\}}\mathrm{E}_{\{l,l',i,k|j\}}[\mathrm{E}_{\mathrm{f}}[g_{ij}g_{lj}^{*}]\mathrm{E}_{\mathrm{f}}[h_{jk}h_{jl}^{*}]\cdot\mathrm{E}_{\mathrm{f}}[g_{ij}^{*}g_{l'j}]\cdot\mathrm{E}_{\mathrm{f}}[h_{jk}^{*}h_{jl'}]]$$

$$\overset{(a)}{=} \frac{1}{A^2}\sum_{i=1,i\neq k}^{r}\sum_{j=1}^{n}\sum_{l=1}^{r}\mathrm{E}_{\{j\}}\left[\mathrm{E}_{\{l,i,k|j\}}[|\mathrm{E}_{\mathrm{f}}[g_{ij}g_{lj}^{*}]|^2]]\cdot\mathrm{E}_{\{l,i,k|j\}}[\mathrm{E}_{\mathrm{f}}[h_{jk}h_{jl}^{*}]|^2]\right]$$

$$= \frac{(r-1)n\left((r-2)\mathrm{E}_{\{j\}}(\mathrm{E}_{\{i|j\}}[|\mathrm{E}_{\mathrm{f}}[g_{ij}]|^2])^4 + 2\mathrm{E}_{\{j\}}[(\mathrm{E}_{\{k|j\}}|\mathrm{E}_{\mathrm{f}}[h_{jk}]|^2)^2\cdot(\mathrm{E}_{\{i|j\}}[\mathrm{E}_{\mathrm{f}}|g_{ij}|^2]^2)]\right)}{r(\kappa_1\sigma_n^2 + p(\kappa_3 + (r-1)\kappa_2))}$$

$$\overset{(b)}{\leq} \frac{(r-1)nB_5}{\kappa_1\sigma_n^2 + p(\kappa_3 + (r-1)\kappa_2)} \leq \frac{rnB_5}{\kappa_1\sigma_n^2 + pr\kappa'}$$

$(t_1, t_2, t_3, t_4) = (1, 2, 3, 4)$ and $(t_1, t_2, t_3, t_4) = (1, 3, 2, 4)$. We consider the following cases.

- $j_1 = j_3 \neq j_2 = j_4$ and $l_1 = l_4 \neq l_2 = l_3$. In this case, considering the connections at receiver side, since $l_1 \neq l_2$, terms like the cross-correlation of two channels between one random point and two fixed point will appear which we know by assumption B is zero. Thus, all the terms in this case result in zero (see Fig. 6(a)).

- $j_1 = j_3 \neq j_2 = j_4$ and $l_1 = l_3 \neq l_2 = l_4$. In this case, if $i \neq i'$, the expectation will be zero by assumption B. Therefore, the total number of terms in this case is $n(n-1)r(r-1)^2$ (see Fig. 6(b)).

- $j_1 = j_3 \neq j_2 = j_4$ and $l_1 = l_2 \neq l_3 = l_4$. Comparing Fig. 6(c) to Fig. 6(a), we see that all the terms in this case are zero as well.

- $j_1 = j_3 \neq j_2 = j_4$ and $l_1 = l_3 = l_2 = l_4$. We can easily see that the only possible way to get a nonzero term is when $i = i'$. The total number of terms in this case is $n(n-1)r(r-1)$ (see Fig. 6(d)).

- $j_1 = j_2 \neq j_3 = j_4$ and $l_1 = l_2 \neq l_3 = l_4$. In this case, we will always get a nonzero result for the expectation. Hence, there are $n(n-1)r(r-1)^3$ nonzero terms. Of all the terms related to this case, there are $n(n-1)\prod_{q=1}^{4}(r-q)$ terms for which we have the additional property that l_1, l_3, i, i', and k are all distinct. Notices that all these terms appear in m_k^2. To observe this, note that because of the distinctness of all the nodes the expectation term in (B9) can be written as

the product of two terms: the first one is the product of the channel coefficients between transmitters i, l_1, relay nodes j_1, and receivers k, l_1. The second one is the product of the channel coefficients between transmitters i', l_3, relay node j_3, and receivers k, l_3. But each of these terms is of the form that appears in (B9). Therefore, the summation of the expectations terms, corresponding to this case, will be canceled out by subtracting m_k^2 from $V_k + m_k^2$ in (B9). Thus, only

$$n(n-1)(r(r-1)^3 - \prod_{q=1}^{4}(r-q)) \leq 12n(n-1)r^3$$

terms appear in the variance V_k (see Fig. 6(e)).

- $j_1 = j_2 \neq j_3 = j_4$ and $l_1 = l_3 \neq l_2 = l_4$. Since $l_1 \neq l_2$, we will have a term like the cross-correlation of two channels between a random point and two fixed points, which we know by assumption B is zero. Hence, all the terms in this case are zero (see Fig. 6(f)).

- $j_1 = j_2 \neq j_3 = j_4$ and $l_1 = l_4 \neq l_2 = l_3$. For this case we always get a zero term (see Fig. 6(g)).

- $j_1 = j_2 \neq j_3 = j_4$ and $l_1 = l_2 = l_3 = l_4$. We always get a nonzero term in this case. There are $n(n-1)r(r-1)^2$ nonzero terms in this case (see Fig. 6(h)).

- $j_1 = j_2 = j_3 = j_4$ and $l_{t_1} = l_{t_2}, l_{t_3} = l_{t_4}$ for (t_1, t_2, t_3, t_4) is a permutation of $(1, 2, 3, 4)$. For this case, we will always get a nonzero expression. The number of nonzero terms in this case is $n(r-1)^2(3r^2 - 2r)$ (see Fig. 6(i)).
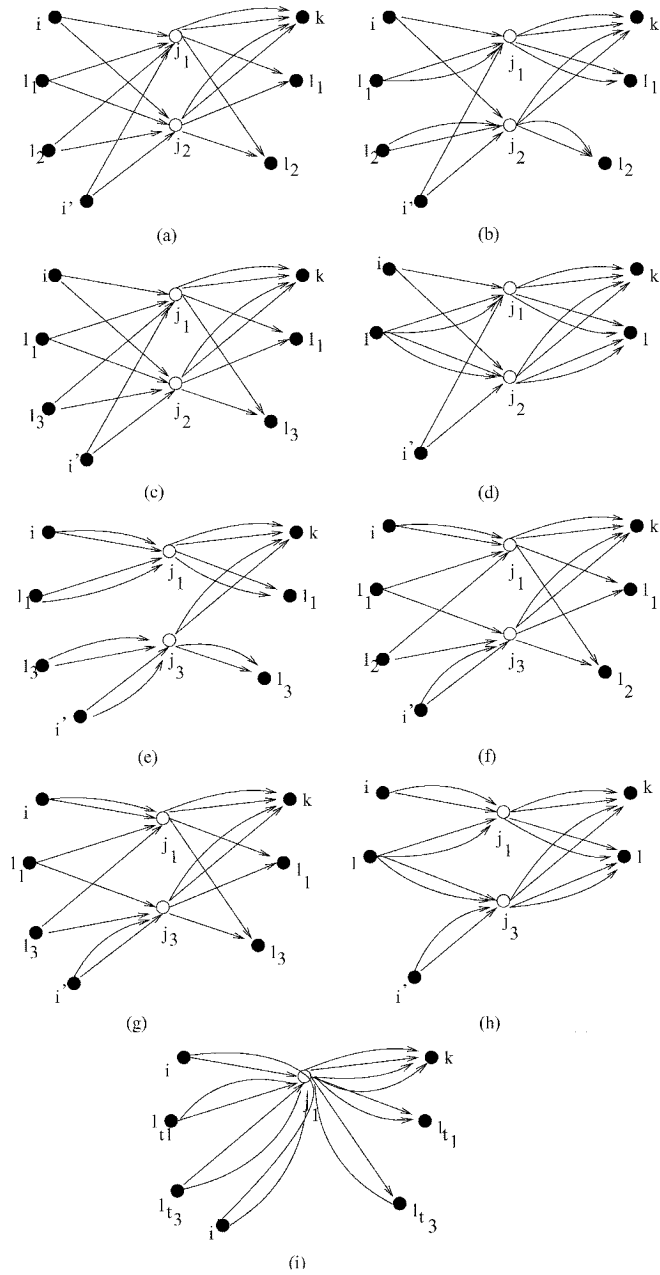
Fig. 6. Possible configurations appearing in computing V_k. In all the cases, each directed link indicates that the corresponding channel coefficient appears in the expression for computing V_k.

From the preceding argument, we have an upper bound for the number of nonzero terms in V_k. By looking at (B9), we see that each of the terms can be written as the ratio between an expectation of the product of at most 16 channel coefficients in the network and A^2, Therefore, the only dependency on n and r is through A^2. If we denote the maximum of the expectation terms appearing in the numerator of these ratios by $B_6 > 0$, we can bound the variance V_k by

$$V_k \leq (3n(n-1)r(r-1)^2 + 2n(n-1)r(r-1)$$
$$+ n(r-1)^2(3r^2 - 2r) + n(n-1)(12r^3))\frac{B_6}{A^2}$$
$$\leq \frac{(n^2r + nr^2)B_6}{(\kappa_1\sigma_n^2 + r\kappa'p)^2}$$

where we have taken into account all the nonzero terms produced by all the arrangements of j_t's and l_t's in the above equation and B_6 is a positive constant independent of n and r.

3) Bounds on Other Moments: Using a very similar approach to that the previous sections, we get the following bounds for the required moments appearing in (B1)

$$\mathrm{E_{loc}E_f}|g_{ij}h_{jk}d_j|^2 \leq \frac{B_2}{r\kappa'p + \kappa_1\sigma_n^2}$$
$$\mathrm{Var_{loc}E_f}|g_{kj}h_{jk}d_j|^2 \leq \frac{3B_3}{(r\kappa'p + \kappa_1\sigma_n^2)^2}$$
$$\mathrm{Var_{loc}E_f}|g_{ij}h_{jk}d_j|^2 \leq \frac{3B_4}{(r\kappa'p + \kappa_1\sigma_n^2)^2}, \qquad i \neq k$$
$$\mathrm{E_{loc}E_f}[g_{kj}h_{jk}d_j] \geq \frac{B_7}{\sqrt{r\left(\kappa_1\sigma_n^2 + rp\kappa\right)}}$$
$$\mathrm{Var_{loc}E_f}[g_{kj}h_{jk}d_j] \leq \frac{B_8}{\kappa_1\sigma_n^2 + pr\kappa'} \qquad \text{(B10)}$$

where B_i's and $\kappa, \kappa', \kappa_1$ are constants that do not depend on n and r and are only a functions of the domain and the fading characteristics. We should mention that to derive the above bounds we only need assumptions A and B of Section II-B. We use assumption C only in computation of V_k.

4) Proof of Lemma: Having found all the required expectations we now compute required bounds for $\mathrm{Var_f}[w'_k]$ and $\mathrm{E_f}[\alpha_k^k]$. From (30) we have

$$\mathrm{Var_f}[w'_k] = \mathrm{Var_f}\left[\alpha_k^k\right]\sigma_r^2 p + \left(\sum_{i=1, i\neq k}^r \mathrm{E_f}\left|\alpha_i^k\right|^2\right)\sigma_r^2 p$$
$$+ \left(\sigma_r^2\sum_{j=1}^n \mathrm{E_f}|\beta_j^k|^2 + 1\right)\sigma_n^2 \quad \text{(B11)}$$

where $\alpha_i^k = \sum_{j=1}^n g_{ij}h_{jk}d_j$ and $\beta_j^k = h_{jk}d_j$. Using the independence of fading for different channels, we have

$$\mathrm{E_f}\left|\alpha_i^k\right|^2 - \left|\mathrm{E_f}\left[\alpha_i^k\right]\right|^2$$
$$= \sum_{j=1}^n \mathrm{E_f}|g_{ij}h_{jk}d_j|^2 - \sum_{j=1}^n |\mathrm{E_f}[g_{ij}h_{jk}d_j]|^2 \quad \text{(B.12)}$$

Therefore, from (B12) and (B11) we have

$$\mathrm{Var_f}[w'_k] \leq \left(\sum_{i=1}^r \sum_{j=1}^n \mathrm{E_f}|g_{ij}h_{jk}d_j|^2\right)\sigma_r^2 p$$
$$+ \left(\sum_{i=1, i\neq k}^r \left|\sum_{j=1}^n \mathrm{E_f}[g_{ij}h_{jk}d_j]\right|^2\right)\sigma_r^2 p$$
$$+ \left(\sigma_r^2\sum_{j=1}^n \mathrm{E_f}|h_{jk}d_j|^2 + 1\right)\sigma_n^2.$$

By using union bound on the probability of events E_k^1, E_k^4 and $E_{i,k}^2$, $i = 1, \ldots, r$ from (B1) and (B2), we get

$$
P\left\{ \mathrm{Var_f}[w_k'] \right.
$$
$$
\leq \sigma_r^2 p \left(n \sum_{i=1}^{r} \mathrm{E_{loc}E_f}|g_{ij}h_{jk}d_j|^2 + nr\epsilon_3 + n\epsilon_2 + n^\rho r^\varsigma \epsilon_5 \right)
$$
$$
\left. + \sigma_n^2 \left(1 + n\sigma_r^2(\epsilon_1 + \mathrm{E_{loc}E_f}|h_{jk}d_j|^2) \right) \right\}
$$
$$
\geq 1 - \left(\frac{\mathrm{Var_{loc}E_f}|h_{jk}d_j|^2}{n\epsilon_1^2} + \frac{\mathrm{Var_{loc}E_f}|g_{kj}h_{jk}d_j|^2}{n\epsilon_2^2} \right.
$$
$$
\left. + \frac{(r-1)\mathrm{Var_{loc}E_f}|g_{ij}h_{jk}d_j|^2}{n\epsilon_3^2} + \frac{m_k}{n^\rho r^\varsigma \epsilon_5^2} \right).
$$

By substituting the required moments in the above equation with their corresponding bounds calculated in preceding paragraphs, after some simplifications, it can be shown that

$$
P\left\{ \mathrm{Var_f}[w_k'] \right.
$$
$$
< \sigma_r^2 p \left(n\epsilon_2 + nr\epsilon_3 + \frac{nrB_2}{\kappa' rp + \kappa_1 \sigma_n^2} + n^\rho r^\varsigma \epsilon_5 \right)
$$
$$
\left. + \sigma_n^2 \left(1 + n\sigma_r^2 \left(\epsilon_1 + \frac{\kappa}{\kappa rp + \kappa_1 \sigma_n^2} \right) \right) \right\}
$$
$$
\geq 1 - \frac{1}{(\kappa' rp + \kappa_1 \sigma_n^2)^2} \left(\frac{3B_1}{n\epsilon_1^2} + \frac{3B_3}{n\epsilon_2^2} + \frac{3(r-1)B_4}{n\epsilon_3^2} \right.
$$
$$
\left. + \frac{rn(\kappa' rp + \kappa_1 \sigma_n^2)}{n^\rho r^\varsigma \epsilon_5} \right). \quad \text{(B13)}
$$

Note that in this bound we have only used assumptions A, B from Section II-B on channel coefficients. By using assumption C together with assumptions A and B we can get a stronger probability bound for variance of the equivalent noise. Using (B10), (B3), the second bound can be written as

$$
P\left\{ \mathrm{Var_f}[w_k'] \right.
$$
$$
< \sigma_r^2 p \left(n\epsilon_2 + nr\epsilon_3 + \frac{nr(B_2 + B_5)}{\kappa' rp + \kappa_1 \sigma_n^2} + n^\rho r^\varsigma \epsilon_5 \right)
$$
$$
\left. + \sigma_n^2 \left(1 + n\sigma_r^2 \left(\epsilon_1 + \frac{\kappa}{\kappa rp + \kappa_1 \sigma_n^2} \right) \right) \right\}
$$
$$
\geq 1 - \frac{1}{(\kappa' rp + \kappa_1 \sigma_n^2)^2}
$$
$$
\times \left(\frac{3B_1}{n\epsilon_1^2} + \frac{3B_3}{n\epsilon_2^2} + \frac{3(r-1)B_4}{n\epsilon_3^2} + \frac{rn(n+r)B_6}{n^{2\rho} r^{2\varsigma} \epsilon_5^2} \right).
$$
$$
\quad \text{(B14)}
$$

Combining (B13) and (B14), we have the following bound for the variance of the equivalent noise:

$$
P\left\{ \mathrm{Var_f}[w_k'] < \sigma_r^2 p \left(n\epsilon_2 + nr\epsilon_3 + \frac{nr(B_2 + B_5)}{\kappa' rp + \kappa_1 \sigma_n^2} + n^\rho r^\varsigma \epsilon_5 \right) \right.
$$
$$
\left. + \sigma_n^2 \left(1 + n\sigma_r^2 \left(\epsilon_1 + \frac{\kappa}{\kappa rp + \kappa_1 \sigma_n^2} \right) \right) \right\}
$$
$$
\geq 1 - \frac{1}{(\kappa' rp + \kappa_1 \sigma_n^2)^2}
$$
$$
\times \left(\frac{3B_1}{n\epsilon_1^2} + \frac{3B_3}{n\epsilon_2^2} + \frac{3(r-1)B_4}{n\epsilon_3^2} + h(n, r, \epsilon_5) \right)
$$

where

$$
h(n, r, \epsilon_5) = \begin{cases} \frac{rn(\kappa' rp + \kappa_1 \sigma_n^2)}{n^\rho r^\varsigma \epsilon_5}, & \text{Using Assumptions A, B} \\ \frac{rn(r+n)B_6}{n^{2\rho} r^{2\varsigma} \epsilon_5^2}, & \text{Using Assumptions A, B, C.} \end{cases}
$$

This is the same bound that is given in Lemma 2 for the variance of the equivalent noise $\mathrm{Var_f}[w_k']$. For

$$
\mathrm{E_f}[\alpha_k^k] = \mathrm{E_f}\left[\sum_{i=1}^{n} g_{kj}h_{jk}d_j \right]
$$

using (B1) and the bounds on the mean and the variance of $\mathrm{E_f}[g_{kj}h_{jk}d_j]$ computed in previous sections, we get

$$
P\left\{ |\mathrm{E_f}[\alpha_k^k]| > n \left(\frac{B_7}{\sqrt{r(\kappa_1 \sigma_n^2 + rp\kappa)}} - \epsilon_4 \right) \right\}
$$
$$
\geq 1 - \frac{B_8}{n\epsilon_4^2(\kappa_1 \sigma_n^2 + rp\kappa')} \quad \text{(B15)}
$$

which is the same as the bound given in Lemma 2. Using

$$
P\{E_{k,k}^2 \cup E_k^3\} \leq P\{E_{k,k}^2\} + P\{E_k^3\}
$$

and (B12) we can write

$$
P\left\{ \mathrm{E_f}|\alpha_k^k|^2 < n(\mathrm{E_{loc}E_f}|g_{kj}h_{jk}d_j|^2 + \epsilon_2) \right.
$$
$$
\left. + n^2(|\mathrm{E_{loc}E_f}[g_{kj}h_{jk}d_j]| + \epsilon_1)^2 \right\}
$$
$$
\geq 1 - \left(\frac{\mathrm{Var_{loc}E_f}[g_{kj}h_{jk}d_j]}{n\epsilon_4^2} + \frac{\mathrm{Var_{loc}E_f}|g_{kj}h_{jk}d_j|^2}{n\epsilon_2^2} \right).
$$
$$
\quad \text{(B16)}
$$

By substituting the required moments computed in previous sections, we get

$$
P\left\{ \mathrm{E_f}|\alpha_k^k|^2 < n \left(\frac{B_2}{r\kappa' p + \kappa_1 \sigma_n^2} + \epsilon_2 \right) \right.
$$
$$
\left. + n^2 \left(\frac{B_7}{\sqrt{r(\kappa_1 \sigma_n^2 + r\kappa p)}} + \epsilon_4 \right)^2 \right\}
$$
$$
\geq 1 - \left(\frac{B_8}{n\epsilon_4^2(\kappa_1 \sigma_n^2 + rp\kappa')} + \frac{3B_3}{n\epsilon_2^2(\kappa' rp + \kappa_1 \sigma_n^2)^2} \right).
$$
$$
\quad \text{(B17)}
$$

This is the same bound given in Lemma 2 for $\mathrm{E_f}|\alpha_k^k|^2$.

C. Maximum Achievable Power Efficiency Using the Listen and Transmit Protocol

In this appendix, we prove that with high probability we cannot achieve a power efficiency of greater than $O(\sqrt{n})$ with the Listen and Transmit protocol.

Using (35), (36), and (39) we can write the upper bound for power efficiency as

$$\eta_{\text{upper}} = \frac{r}{rp + n\sigma_r^2} \log\left(1 + n\frac{\sigma_r^2 p}{\sigma_n^2}\left(\frac{B_2}{r\kappa'p + \kappa_1\sigma_n^2} + \epsilon_2\right)\right.$$
$$\left. + n^2 \frac{\sigma_r^2 p}{\sigma_n^2}\left(\frac{B_7}{\sqrt{r\left(\kappa_1\sigma_n^2 + r\kappa p\right)}} + \epsilon_4\right)^2\right)$$
$$\text{P}\{\eta_{\text{upper}} \geq \eta\} \geq 1 - r\left(\frac{B_8}{n\epsilon_4^2\left(\kappa_1\sigma_n^2 + rp\kappa'\right)}\right.$$
$$\left. + \frac{3B_3}{n\epsilon_2^2\left(\kappa'rp + \kappa_1\sigma_n^2\right)^2}\right) \qquad \text{(C1)}$$

where these inequalities hold for every $\epsilon_2 > 0$ and $\epsilon_4 < \frac{B_7}{\sqrt{r(\kappa_1\sigma_n^2 + rp\kappa)}}$. By substituting $\epsilon_4 = \frac{B_7}{2\sqrt{r(\kappa_1\sigma_n^2 + r\kappa p)}}$, we have

$$\eta_{\text{upper}} = \frac{r}{rp + n\sigma_r^2} \log\left(1 + n\frac{\sigma_r^2 p}{\sigma_n^2}\left(\frac{B_2}{r\kappa'p + \kappa_1\sigma_n^2} + \epsilon_2\right)\right.$$
$$\left. + n^2 \frac{9\sigma_r^2 p}{4\sigma_n^2}\frac{B_7^2}{r\left(\kappa_1\sigma_n^2 + r\kappa p\right)}\right)$$
$$\text{P}\{\eta_{\text{upper}} \geq \eta\} \geq 1 - r\left(\frac{4rB_8\left(\kappa_1\sigma_n^2 + r\kappa p\right)}{9nB_7^2\left(\kappa_1\sigma_n^2 + rp\kappa'\right)}\right.$$
$$\left. + \frac{3B_3}{n\epsilon_2^2\left(\kappa'rp + \kappa_1\sigma_n^2\right)^2}\right). \qquad \text{(C2)}$$

The above equation can be simplified to

$$\eta_{\text{upper}} = \frac{r}{rp + n\sigma_r^2} \log\left(1 + \frac{\sigma_r^2 p}{\sigma_n^2\left(\kappa_1\sigma_n^2 + rp\kappa'\right)}\right.$$
$$\left. \times \left(K_1 n + K_2 \frac{n^2}{r}\right)\right)$$
$$\text{P}\{\eta_{\text{upper}} \geq \eta\} \geq 1 - \frac{K_3 r^2}{n} \qquad \text{(C3)}$$

where K_1, K_2, and K_3 are constants independent of n and r. By defining $x = rp$ and $y = n\sigma_r^2$ and finding the partial derivatives of η_{upper} in (C3) with respect to x and y, we can show that values that maximize the upper bound satisfy the following relation:

$$y = \frac{x\left(\kappa'x + \kappa_1\sigma_n^2\right)}{\kappa_1\sigma_n^2}.$$

Substituting this value of y in the expression yields

$$\eta_{\text{upper}}^\star = \max_x \frac{r}{2\left(x + \frac{x(\kappa'x + \kappa_1\sigma_n^2)}{\kappa_1\sigma_n^2}\right)}$$
$$\times \log\left(1 + \frac{x^2\left(\frac{K_1}{r} + \frac{K_2 n}{r^2}\right)}{\kappa_1\sigma_n^4}\right)$$
$$\text{P}\{\eta_{\text{upper}}^\star \geq \eta\} \geq 1 - \frac{K_3 r^2}{n}. \qquad \text{(C4)}$$

If we define $z = \frac{\sqrt{n}}{r}x$ then by using the fact that $r = O(\sqrt{n})$ we will have

$$\eta_{\text{upper}}^\star = O\left(\max_z \sqrt{n}\frac{\log(1 + A'z^2)}{z(1 + \frac{r}{\sqrt{n}}z)}\right)$$
$$= \max_z \frac{\log(1 + A'z^2)}{z}O(\sqrt{n}) = O(\sqrt{n})$$
$$\text{P}\{\eta_{\text{upper}}^\star \geq \eta\} \geq 1 - \frac{K_3 r^2}{n} \qquad \text{(C5)}$$

where A' is some constant and the second equality comes from the fact that $r = O(\sqrt{n})$. Therefore, with high probability the maximum achievable power efficiency with the Listen and Transmit protocol is $O(\sqrt{n})$.

REFERENCES

[1] M. Gastpar and M. Vetterli, "On the asymptotic capacity of Gaussian relay networks," in *Proc. Int. Symp. Information Theory*, Lausanne, Switzerland, 2002, p. 195.

[2] P. Gupta and P.R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[3] ——, "Toward an information theory of large networks: an achievable rate region," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1877–1894, Aug. 2003.

[4] L-L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[5] O. Lévêque and İ.E. Telatar, "Information theoretic upper bounds on the capacity of large extended ad hoc wireless networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 858–865, Mar. 2005.

[6] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.

[7] S. D. Sevetto, "On the feasibility of large-scale wireless sensor networks," in *Proc. 40th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2002.

[8] A. Scaglione and S. D. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proc. 8th ACM Int. Conf. Mobile Computing and Networking (MobiCom)*, Atlanta, GA, Sep. 2002.

[9] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," in *The 2nd Int. Workshop on Information Processing in Sensor Networks (IPSN '03)*, Palo Alto, CA, Apr. 2003.

[10] A. Ephremides, "Energy concerns in wireless networks," *IEEE Wireless Commun.*, pp. 48–59, Aug. 2002.

[11] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[12] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Conf. Communications*, Seattle, WA, Jun. 1995.

[13] D. N. C. Tse and S. V. Hanly, "Multi-access fading channels–Part I: Polymatroid structure, optimal resource allocation, and throughput capacity," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

[14] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Hawaii Int. Conf. System Sciences*, Hawaii, Jan. 2000, pp. 3005–3014.

[15] B. Hassibi and A. Dana, "On the power efficiency of sensory and ad hoc wireless networks," in *Proc. Asilomar Conf. Signals, Systems, and Computing*, Pacific Grove, CA, Nov. 2002.

[16] V. Rodoplu and T.H. Meng, "Minimum energy mobile wireless networks," *IEEE J. Sel. Area Commun.*, vol. 17, no. 8, pp. 1333–1344, Aug. 1999.

[17] J. Gomez, A.T. Campbell, M. Naghshineh, and C. Bisdikian, "Power-aware routing in wireless packet networks," in *Proc. IEEE Int. Workshop on Mobile Mult. Communications*, Nov. 1999, pp. 380–383.

[18] A. Lozano, A. Tulino, and S. Verdú, "Multi-antenna capacity in low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[19] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[20] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. IEEE Infocom 2002*, New York, Jun. 2002.

[21] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.