Asymptotic Analysis of Multi-Stage Cooperative Broadcast in Wireless Networks

Birsen Sirkeci Mergen, *Student Member, IEEE,* Anna Scaglione, *Member, IEEE,* and Gökhan Mergen, *Member, IEEE*

***Abstract*— Cooperative broadcast aims to deliver a source message to a locally connected network by means of collaborating nodes. In traditional architectures, node cooperation has been at the network layer. Recently, physical layer cooperative schemes have been shown to offer several advantages over the network layer approaches. This form of cooperation employs distributed transmission resources at the physical layer as a single radio with spatial diversity. In decentralized cooperation schemes, collaborating nodes make transmission decisions based on the quality of the received signal, which is the only parameter available locally. In this case, critical parameters that influence the broadcast performance include the source/relay transmission powers and the *decoding threshold* (the minimum SNR required to decode a transmission). We study the effect of these parameters on the number of nodes reached by cooperative broadcast. In particular, we show that there exists a *phase transition* in the network behavior: if the decoding threshold is below a *critical value*, the message is delivered to the whole network. Otherwise, only a fraction of the nodes is reached, which is proportional to the source transmit power. Our approach is based on the idea of *continuum approximation*, which yields closed-form expressions that are accurate when the network density is high.**

***Index Terms*— Broadcast, continuum, cooperative communication, limiting behavior of dense networks, multihop diversity, phase transition, wireless networks.**

I. Introduction

A. Motivation

In distributed ad hoc networks, most network protocols require multicast or broadcast of certain control messages. These messages generally constitute a significant portion of network traffic, and they may cause performance bottlenecks. Several authors have studied how to optimally transmit broadcast information to minimize the total number of transmissions or the energy consumption in large wireless networks (*e.g.*, see [1], [2]).

An important property of the wireless medium is that the transmitted packets are heard not only by their intended recipients but also by other neighboring nodes. While such unintended receptions are harmful when there is a single intended recipient, they may be *exploited* in broadcast mode. Moreover, "*collisions*" between different transmitting nodes, which hinder point-to-point communication, may actually be *beneficial*, when the transmitting nodes are broadcasting the same message.

In this paper, we analyze the transmission dynamics of a *cooperative broadcast protocol*, in which nodes sequentially transmit the same message in large groups to increase the received power. Here, we take the view that the group transmissions (hence, the *intentional* collisions) are actually beneficial, since they increase the received power and the transmission range. This approach is in contrast with the traditional network layer flooding that treats each link individually and attempts to eliminate collisions as much as possible. Compared to multihop broadcast, cooperative broadcast results in more rapid message propagation with fewer number of steps (Fig. 1).

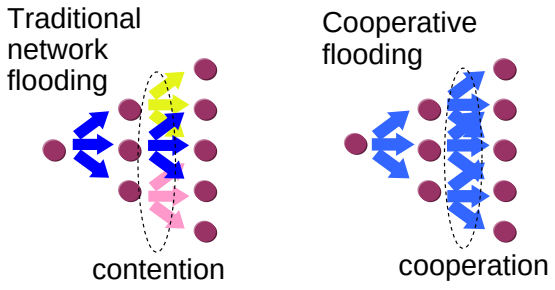

Fig. 1. Non-Cooperative versus Cooperative Broadcast

B. Problem Setup and Our Contribution

In this paper, we analyze the transmission dynamics of a simple *cooperation* protocol for broadcast over multiple stages of relays. In the considered setup, a source node initiates the broadcast by transmitting a packet. Every node who can hear the source with sufficient signal-to-noise ratio (SNR), decodes and retransmits the same packet. A training preamble in the message helps nodes to detect the packet's presence, estimate the received power and synchronize the retransmissions. The retransmissions are done *simultaneously*, even though they may not be symbol synchronized. The first group excites a second group of nodes and the retransmissions continue until every node who hears the others with sufficient SNR, retransmits once. The subsequent groups of nodes that are activated are referred to as *levels*.

The nodes use a simple *SNR threshold* criterion to decide if they are going to retransmit or not, *i.e.*, every node monitors its received SNR and decodes and retransmits if and only

Manuscript is submitted March 14, 2005; revised October 26, 2005. This work was supported by National Science Foundation under grant ITR CCR - 0428427. The material in this paper was presented in part at the 2004 IEEE Workshop on Signal Processing Advances in Wireless Communications, Lisbon, Portugal, July 2004, 2004 International Conference on Acoustics, Speech, and Signal Processing, Montreal, Quebec, Canada, May 2004 and 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing, Philadelphia, PA, May 2005.

B. S. Mergen and A. Scaglione are with Cornell University, Ithaca, NY, 14853 (e-mail: bs233@cornell.edu; anna@ece.cornell.edu).

G. Mergen is currently with Qualcomm, Inc., 675 Campbell Technology Parkway, Campbell, CA 95008 (e-mail: gmergen@ece.cornell.edu).

if its SNR exceeds a certain pre-determined threshold. In this way, the network can operate in a distributed fashion, since the nodes only use the locally available received SNR information to make transmission decisions. We assume that appropriate channel coding is used so that the decoding and retransmissions are correct as long as the received SNR is above the threshold.

In our analysis, we consider two different models for receptions. The first one, which we call the *deterministic model*, assumes that the power of simultaneously transmitted packets is equal to the sum of individual powers. This model is valid if the relays transmit in orthogonal channels, as in FDMA or CDMA, or if the relays use orthogonal space-time codes as considered in [3]. In case of orthogonal channels, a large bandwidth is required, *i.e.*, the network should operate in the *wideband regime*. In our second setup, we derive and consider a *random channel model* applicable for narrowband communication. Here, the impulse response of the channel with multiple transmitters is modelled as a Gaussian random vector. This model takes into account the effects of channel fading, time differences between simultaneous transmissions and random phases.

The objective of cooperative broadcast is to deliver the source message to the whole network. However, this goal may or may not be achieved depending on certain network parameters such as the source/relay transmission powers and the decoding threshold. In this paper, we analyze the effect of these parameters on the number of nodes reached by cooperative broadcast. In particular, we show that there exists a *phase transition* in the network behavior: if the decoding threshold is below a *critical value*, the message is delivered to the whole network. Otherwise, only a fraction of the nodes is reached proportional to the source transmit power.

The two regimes above and below the critical threshold are depicted in Fig. 2. Here, after the source transmission, other nodes in the network transmit in levels. The levels move outward as the transmissions continue. In Fig. 2(a), the number of simultaneously transmitting nodes grows at each stage, and the packet is distributed to the whole network in growing steps. On the other hand, in Fig. 2(b), the number of transmitting nodes diminishes in time, and the transmissions die out. We would like to note that the full-broadcast behavior is obtained at the cost of reduced SNR threshold, which results in a reduced communication rate in order to avoid erroneous decoding.

In our analysis, we first consider a random network in which the node locations are randomly and uniformly distributed. A *continuum model* is obtained from the random network by letting the number of nodes go to infinity while the total relay power is fixed. The continuum approach was previously used in different contexts in [4], [5]. Using the deterministic channel model with squared path-loss model, in the continuum limit, we provide a complete characterization of the broadcast levels, the total area reached by broadcast, and the critical threshold. In particular, for the deterministic model the critical threshold is shown to be equal to

$$\mathrm{SNR}_c = (\pi \ln 2) P_r \rho,$$

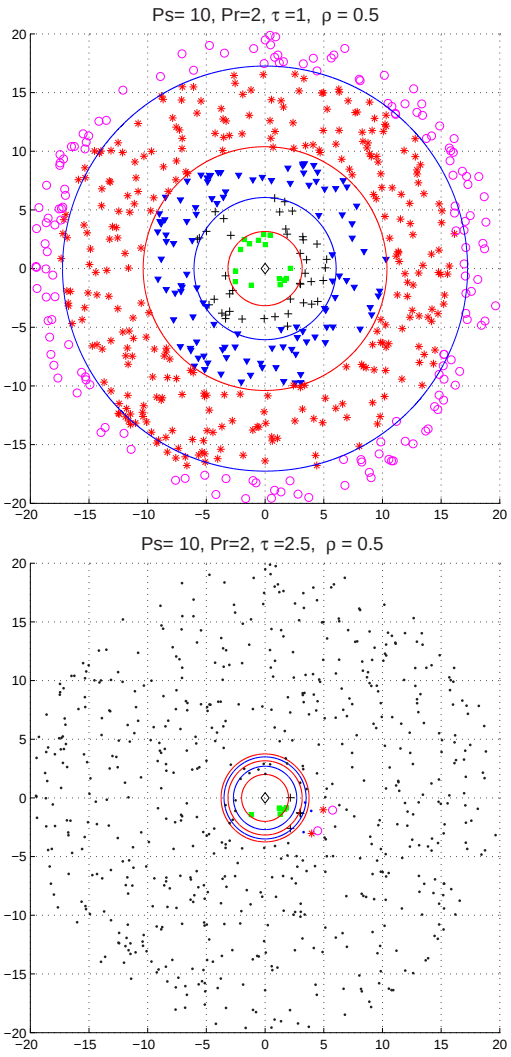

Fig. 2. (a) Transmissions propagate. (b) Transmissions die off. Nodes belonging to different levels are represented with different symbols. The nodes that didn't receive/retransmit the source message are shown with dots.

where P_r is the relay transmit power, ρ is the relay density [node/area], and the channel noise is of unit power.

We further consider the case that the relays exploit the received signal not only from their immediate neighbors (i.e. the previous level), but from m previous levels that repeated the same source message. The importance of using transmissions coming from multiple hops was recognized by [6]–[8]. Following the terminology in [6], we shall call this *multihop diversity*. In the case of m-level multihop diversity, it is shown that the phase transition occurs at the critical threshold $\mathrm{SNR}_c = [\pi \ln(m+1)] P_r \rho$.

In the second part of the paper, we derive equivalent channel models for networks with channel fading. Both orthogonal and non-orthogonal relay transmissions are considered. The network behavior is characterized in the continuum limit by the solution of a nonlinear deterministic dynamical system. Furthermore, an upper bound on the critical threshold is found for the non-orthogonal case. For both the deterministic and random channel models, it is shown by simulations that the continuum model provides reasonably accurate performance

estimates for dense random networks.

Using the continuum model for orthogonal and non-orthogonal relay channels, we analyze the speed of propagation under these two scenarios. More specifically, we fix the duration of the source message, and compare the number of hops necessary to reach a given distance from the source. Interestingly, our results indicate that the speed of propagation in the high-density network with the narrowband non-orthogonal scheme is faster than that of the wideband orthogonal scheme. Although this appears highly non-intuitive, we reason , in the first system, that there is a possibility that the signals may add up constructively. In the asymptotic regime, a fraction of the nodes will receive a signal power equal to the power that would have been obtained if the transmitters were *beamforming* towards these fortunate destinations. Obviously, the orthogonal transmission scheme does not enable beamforming gains for any of the users. The nodes that do receive with beamforming gain grow as the groups expand in size, creating a positive feedback effect that explains why the non-orthogonal scheme outperforms the orthogonal one. In the literature, there are other examples where channel randomization improves the system performance such as the opportunistic communication method proposed in [9].

C. Related Work

Most of the previous analysis of cooperation protocols ([10]–[15]) considered *single- or two-hop* communications. To the best of our knowledge, continuum analysis of cooperative *multi-hop* networks has not been done previously for any cooperative protocol. Part of the results of this paper was initially reported in [16]–[18]. The authors studied the dynamics of cooperative transmissions in a multi-hop network with *a single source-destination pair* in [19].

Cooperative protocols can be categorized according to the number of relays for which they are designed. The papers [10]–[15] investigate the spatial diversity in cooperative networks with a few number of nodes. More specifically, [13] develops low complexity cooperation protocols exploiting the spatial diversity. In [11], cooperative transmissions are considered for improving the uplink capacity. In both [13] and [11], a network with two sources and a single destination is considered. It is also assumed that the nodes are both sources and relays at the same time, and each node has an orthogonal channel assigned exclusively. It is worth noting that the channel assignment to different nodes in most proposed methods usually requires a central control unit.

Most of the protocols for multiple relays are generally extensions of designs for a few number of nodes. The extension is done by using relays in parallel (*i.e.,* multiple relays transmit simultaneously in groups—*e.g.,* [3], [20]–[22]), or in series (*i.e.,* relays transmit sequentially—*e.g.,* [6], [23]), or a combination of these two (*e.g.,* [24]). In [6], [23], four different network models are considered, which are grouped according to relay processing (amplifying or decoding) and signal reception model (from all previously transmitted terminals or from the immediate terminal). In [24], authors derive symbol error probability expressions, valid under high SNR, for networks with parallel, serial, and also combined configurations. Both [7], [25] investigate the energy efficiency of cooperative transmissions over multi-hop networks for different setups. In [26] and [27], the authors consider the channel capacity with multiple relays.

Another related scheme is *Opportunistic Large Arrays* proposed in [8], which relies on a distributed rule, referred to as the *integrate and fire model*, to synchronize the nodes. In this method, the nodes select a firing time based on the energy accumulated at the receiver. The nodes emit their decision at the firing times, which are decided in a distributed fashion. This scheme has low complexity compared to the centralized cooperative schemes, and it eliminates the problem of scheduling transmissions; however, it cannot guarantee diversity gains since the transmitted signals can overlap in time, and it requires non-negligible bandwidth overhead (see [8] for details).

In a recent work on cooperative transmission [28], the authors show that there exists a critical rate C such that the outage probability of every receiver converges to zero, for rates below C, as the number of nodes goes to infinity. The analysis is done under a sum power constraint and also independent and identically distributed channel gains. Interesting phase transitions also arise in applications of percolation theory such as the connectivity analysis of random networks [29], [30].

The organization of the paper is as follows. In Section II, the analysis of the network with the deterministic channel model is presented. In Section III, the random channel model is derived and in Section IV the continuum network behavior is characterized for the random channel models. In Section V, simulation results are presented. We conclude the paper in Section VI.

II. Network Behavior Under Deterministic Channel Model

In this section, we consider a simple *deterministic* model for modelling the received power of simultaneously transmitted signals. Suppose that every transmission with power P is received with power $P\ell(d)$ at distance d, where $\ell(.)$ is a *path-loss attenuation function*, which is assumed to be continuous and non-increasing (*e.g.,* $\ell(d) = 1/d^2$). For convenience, we use the notation $\ell(x, y)$ to denote $\ell(\sqrt{x^2 + y^2})$.

In the deterministic channel model, it is assumed that if a set of relay nodes (say, level-n nodes $= \mathcal{L}_n$) transmits simultaneously, then node j receives with power

$$Pow(j) = P \sum_{i \in \mathcal{L}_n} \ell(d_{ij}) \tag{1}$$

where d_{ij} is the distance between the i'th and j'th nodes. In practice, this received power can be achieved via *maximal ratio combining* under the following scenarios: (i) the nodes in a given level transmit in orthogonal channels (as in TDMA or FDMA); (ii) the nodes use orthogonal or pseudo-orthogonal spreading codes with desirable correlation properties; (iii) simultaneously transmitting nodes employ a distributed orthogonal space-time code [3] designed for a large number of nodes.

A. Random Network

Suppose that N nodes are uniformly and randomly distributed in a disc with radius R and a single source is located at the center of a circular region.[1] Let the source transmit with power P_s, and the relays transmit with power P_r. Let $\mathcal{S}$ denote the set of locations of relay nodes. The relay nodes decode and retransmit if and only if their SNR exceeds a certain threshold τ. Under the assumption that the channel noise is of unit power, the SNR threshold criterion is equivalent to a received power criterion, i.e.

$$Pow(j) \geq \tau.$$

At every broadcast step, the set of nodes with reception power exceeding τ, which has not transmitted so far, joins the next level. Therefore, the set of level-1 nodes is given by

$$\mathcal{S}_1 = \{(x,y) \in \mathcal{S} : P_s \ell(x,y) \geq \tau\}. \tag{2}$$

Similarly, the set of level-k nodes for $k \geq 2$ is given by

$$\mathcal{S}_k = \{(x,y) \in \mathcal{S} \setminus \bigcup_{i=1}^{k-1} \mathcal{S}_i : \sum_{(x',y') \in \mathcal{S}_{k-1}} P_r \ell(x-x', y-y') \geq \tau\}. \tag{3}$$

When the node locations are random, a random number of nodes is reached by the source in every realization of the network. In order to study the effect of P_s, P_r and τ on the broadcast behavior, we will consider the continuum model, which is introduced next.

B. Continuum Network

Let $\mathbb{S} \triangleq \{(x,y) : x^2 + y^2 \leq R^2\}$ denote the disc containing the network. Let $\rho = N/\text{Area}(\mathbb{S})$ be the density [node/unit area] of relays within $\mathbb{S}$. In the continuum model, we are interested in the behavior of *high density* networks with *constant sum-power*. That is, the number of relays N goes to infinity, while $P_r N$ is fixed. This implies that the *relay power per unit area*

$$\bar{P}_r \triangleq \frac{P_r N}{\text{Area}(\mathbb{S})} = P_r \rho \tag{4}$$

is also fixed.

As the number of relays goes to infinity, the level-1 nodes become dense in the set

$$\mathbb{S}_1 \triangleq \{(x,y) \in \mathbb{S} : P_s \ell(x,y) \geq \tau\}.$$

Intuitively speaking, in this regime every infinitesimal area $dxdy$ in $\mathbb{S}_1$ contains $\rho dxdy$ nodes each with power P_r. Hence, the total transmission power from each such infinitesimal area is $P_r \rho dxdy = \bar{P}_r dxdy$. Consequently, the level-2 nodes become dense in the set

$$\mathbb{S}_2 = \{(x,y) \in \mathbb{S} \setminus \mathbb{S}_1 : \iint_{\mathbb{S}_1} \bar{P}_r \ell(x-x', y-y') dx'dy' \geq \tau\}. \tag{5}$$

[1]The continuum analysis can be adapted easily for other network topologies. See [19] for the rectangular topology, where the source is located at the boundary point of the rectangle.

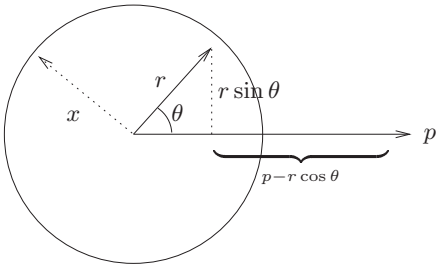

Fig. 3. Illustration of $f(x,p)$.

By recursion, it is seen that the level-k nodes, $k \geq 2$, become dense in

$$\mathbb{S}_k = \{(x,y) \in \mathbb{S} \setminus \bigcup_{i=1}^{k-1} \mathbb{S}_i : \iint_{\mathbb{S}_{k-1}} \bar{P}_r \ell(x-x', y-y') dx'dy' \geq \tau\}. \tag{6}$$

The sets $\mathbb{S}_1, \mathbb{S}_2, \cdots$ specify the continuum model.

The following theorem provides a rigorous relation between the random network and its continuum limit.

Theorem 1: Let $\bar{P}_r$ and R be fixed, and

$$P_r = \bar{P}_r/\rho, \qquad \rho = \frac{N}{\pi R^2} \tag{7}$$

be a function of N. For all $k \in \{1, 2, \cdots\}$ and open disc[2] $\mathbb{U} \subset \mathbb{R}^2$, the number of level$-k$ nodes in $\mathbb{U}$ scales as $\rho\text{Area}(\mathbb{U} \cap \mathbb{S}_k)$, i.e.,

$$\frac{|\mathbb{U} \cap \mathcal{S}_k|}{\rho\text{Area}(\mathbb{U} \cap \mathbb{S}_k)} \xrightarrow{\text{p}} 1 \quad \text{as } N \to \infty, \tag{8}$$

where $\xrightarrow{\text{p}}$ denotes convergence in probability.

Proof: See Appendix I. ∎

Theorem 1 indicates that the locations of level-k nodes are approximately uniform in the set $\mathbb{S}_k$, when the network density is high. Choosing $\mathbb{U} = \mathbb{S}$ above suggests an approximation to the number of level-k nodes:

$$|\mathcal{S}_k| \approx \rho\text{Area}(\mathbb{S}_k), \qquad \text{for large } N, \tag{9}$$

and an approximation to the total number of nodes reached by cooperative broadcast:

$$|\bigcup_{k=1}^{\infty} \mathcal{S}_k| \approx \rho\text{Area}(\bigcup_{k=1}^{\infty} \mathbb{S}_k), \qquad \text{for large } N. \tag{10}$$

We will later demonstrate the validity of (9) and (10) also by simulation.

[2]A set $\mathbb{U}$ is called an open disc if it is of the form $\{(x',y') : (x-x')^2 + (y-y')^2 < z\}$ for some $(x,y) \in \mathbb{R}^2$ and $z > 0$. Although the proof in this paper is only for disc shaped $\mathbb{U}$, the theorem actually holds for any *Jordan measurable* set $\mathbb{U}$ in $\mathbb{R}^2$. A set $\mathbb{U} \subset \mathbb{R}^2$ is called *Jordan measurable* if its area can be approximated by rectangles [31], *i.e.*, if $\sup \sum_i \text{Area}(U_i)$, where the supremum is over all disjoint rectangles (U_i's) inside $\mathbb{U}$, is equal to $\inf \sum_i \text{Area}(U_i)$, where the infimum is over all rectangles (U_i's) covering $\mathbb{U}$. The Jordan measure of a set, if exists, is the same as its Lebesgue measure.

C. Explicit Characterization of Level Sets

In this section, we will give a more explicit characterization of the level sets $\mathbb{S}_1, \mathbb{S}_2, \cdots$ for general $\ell(\cdot)$ and for $\ell(d) = 1/d^2$. For ease of presentation, we will assume an unbounded network, *i.e.,* $R = \infty$. The results for $R < \infty$ follow trivially from the results in this section.

Lemma 1: Define the function

$$f(x,p) \triangleq \int_0^x \int_0^{2\pi} \ell(p - rcos\theta, rsin\theta) r d\theta dr. \qquad (11)$$

Let $r_0, r_1, \cdots$ denote the solution of the recursive formula

$$f(r_{k-1}, r_k) - f(r_{k-2}, r_k) = \frac{\tau}{\bar{P}_r}, \qquad k = 2, 3, \cdots \qquad (12)$$

with initial conditions $r_0 = 0,\ r_1 = \ell^{-1}(\frac{\tau}{P_s})$. If the solution of (12) exists, then each $\mathbb{S}_k$ is a *disc shaped region* with inner and outer radii given by r_{k-1} and r_k, respectively.

Proof: The fact that each $\mathbb{S}_k$ is a disc follows from the assumption that the function $\ell(x,y)$ is circularly symmetric. If the solution of $P_s\ell(r_1) = \tau$ exists, then since $\ell(\cdot)$ is decreasing and continuous, $r_1 = \ell^{-1}(\frac{\tau}{P_s})$ forms the boundary of $\mathbb{S}_1$. Notice that the $\bar{P}_r f(x,p)$ is equal to the received power at a node with distance p from the source when a disc of radius x transmits (see Fig. 3). Hence, $\bar{P}_r[f(r_{k-1}, r_k) - f(r_{k-2}, r_k)]$ is the received power at a node with distance r_k from the source, when the *disc* between r_{k-2} and r_{k-1} transmits. If Eqn. (12) is satisfied, then the point r_k lies on the outer boundary of $\mathbb{S}_k$. The lemma follows. ∎

Numerical solutions of $r_0,\ r_1, \cdots$ can be obtained from (12) for general path-loss models. In the following, we explicitly solve this recursive formula for the squared-distance path-loss model by simplifying it into an equivalent homogenous linear difference equation. The relation (12) does not seem to yield closed-form expressions for other path-loss models.

Theorem 2: If $\ell(d) = \frac{1}{d^2}$, then $r_k = \sqrt{a_k}$, where the a_k depends on $\mu \triangleq \exp\left(\frac{\tau}{\bar{P}_r \pi}\right)$ as follows.

i) *Case 1* ($\mu \leq 2$): The broadcast reaches to the whole network, *i.e.,* $\lim_{k\to\infty} a_k = \infty$, where

$$a_k = \begin{cases} \frac{P_s(\mu-1)}{\tau(\mu-2)}\left(1 - \frac{1}{(\mu-1)^k}\right) & \text{if } \mu < 2 \\ \frac{P_s}{\tau}k & \text{if } \mu = 2. \end{cases} \qquad (13)$$

ii) *Case 2* ($\mu > 2$): The total area reached by the broadcast is bounded, *i.e.,*

$$\lim_{k\to\infty} a_k = \frac{P_s(\mu-1)}{\tau(\mu-2)}, \qquad (14)$$

where

$$a_k = \frac{P_s(\mu-1)}{\tau(\mu-2)}\left(1 - \frac{1}{(\mu-1)^k}\right). \qquad (15)$$

Remark: Theorem 2 implies that the network behavior goes through a *phase transition* depending on the value of μ. If

$$\mu \leq 2 \quad \Leftrightarrow \quad \tau \leq (\pi \ln 2)\bar{P}_r, \qquad (16)$$

i.e., the detection threshold is low enough with respect to the relay power per unit area, then the signal propagates to the whole network (see Fig. 4). On the other hand, if $\mu > 2$, then

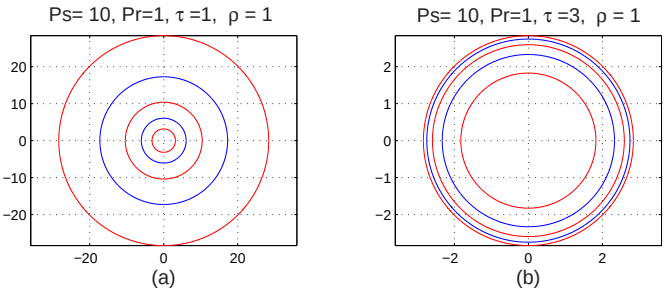

Fig. 4. (a) Transmissions propagate. (b) Transmissions die off. Notice that the scale of (a) and (b) are vastly different.

a finite portion of the network is reached, and from Theorem 1 the total number of nodes the message is delivered to is approximately equal to

$$|\bigcup_{k=1}^{\infty} \mathcal{S}_k| \approx \pi a_\infty \rho = \frac{\pi P_s \rho(\mu-1)}{\tau(\mu-2)}. \qquad (17)$$

The right hand side of Eqn. (17) implies that the number of nodes reached by the broadcast is directly proportional to the source power.

Proof: (Theorem 2) Under the squared distance path-loss model, $r_1 = \sqrt{\frac{P_s}{\tau}}$. Furthermore,

$$\begin{aligned} f(x,p) &= \int_0^x \int_0^{2\pi} \frac{r}{r^2 sin^2(\theta) + (p - rcos(\theta))^2} d\theta dr \\ &= \int_0^x \frac{2\pi r}{|p^2 - r^2|} dr \\ &= \pi \ln \frac{p^2}{|p^2 - x^2|}. \end{aligned} \qquad (18)$$

Hence, (12) is equivalent to

$$f(r_{k-1}, r_k) - f(r_{k-2}, r_k) = \pi \ln \frac{|r_k^2 - r_{k-2}^2|}{|r_k^2 - r_{k-1}^2|} = \frac{\tau}{\bar{P}_r},$$

which yields

$$r_k^2 = \frac{r_{k-1}^2 \mu - r_{k-2}^2}{\mu - 1}, \quad k = 2, 3, \ldots.$$

Defining $a_k = r_k^2$, we get a linear difference equation:

$$a_{k+1} - \frac{\mu}{\mu-1} a_k + \frac{1}{\mu-1} a_{k-1} = 0 \qquad (19)$$

with the initial conditions $a_0 = 0$ and $a_1 = \frac{P_s}{\tau}$.

Eqn. 19 can be solved as follows. Substituting $a_k = Aw^k$, we obtain

$$w^2 - \frac{\mu}{\mu-1} w + \frac{1}{\mu-1} = 0.$$

The roots are

$$w_1 = 1, \quad w_2 = \frac{1}{\mu-1}.$$

If $\mu \neq 2$, the roots are different, and the solution is

$$a_k = A_1 + A_2 w_2^k, \qquad (20)$$

where A_1 and A_2 have to be found from the initial conditions $a_0 = 0,\ a_1 = P_s/\tau$. The first condition implies $A_1 = -A_2 = A$, and the second one implies

$$A = \frac{P_s}{\tau}\left(\frac{\mu-1}{\mu-2}\right).$$

Thus, the solution is

$$a_k = \frac{P_s}{\tau}\frac{\mu-1}{\mu-2}(w_1^k - w_2^k).$$

If $\mu = 2$, then $w_1 = w_2 = 1$. In this case, (20) does not hold and the solution is of the form $a_k = A_1 + A_2 k$. Applying the boundary values, we get $A_1 = 0$ and

$$a_k = r_k^2 = \frac{P_s}{\tau}k. \tag{21}$$

The rest of the theorem follows by inspection of the solution. ∎

D. Effect of Multihop Diversity

In this section, we analyze the phase transition behavior in the case of *multihop diversity*. In multihop diversity mode, each node stores the received signals from m previous levels and combines them via maximal ratio combining. The parameter m can be interpreted as the *memory* of the receiver. With multihop diversity, after the n-th level transmission, the reception power of node j is given by

$$Pow(j) = \sum_{i \in \bigcup_{l=(n-m+1)_+}^{n} \mathcal{L}_l} \frac{P_r}{d_{ij}^2}$$

where $(x)_+ \triangleq \max(0, x)$.

Theorem 3: Consider the m-level multihop diversity network with the squared-distance path-loss model. The network exhibits two different behaviors depending on whether m is finite or infinite.

i) If $m < \infty$, then the network goes through a phase transition at

$$\mu = m+1 \quad \Leftrightarrow \quad \tau = (\pi \ln(m+1))\bar{P}_r \tag{22}$$

where $\mu = \exp\left(\frac{\tau}{\bar{P}_r \pi}\right)$ as before.

ii) If $m = \infty$, then there is no phase transition, and the message propagates to the whole network, *i.e.,* $a_k \to \infty$, regardless of the value of τ.

Proof: See Appendix II. ∎

Remark: Theorem 3 has a somewhat surprising implication: regardless of how low the $\bar{P}_r$ is, if the nodes accumulate and exploit all the transmissions before them, then the message will be delivered to the whole network. This result indicates that high-density cooperative networks with multihop diversity can indeed be very energy efficient as long as the network density is high enough.

Remark: In this section, we assume that the nodes retransmit if and only if the received SNR is above a threshold τ. We show that for the source message to propagate to the entire network, the condition $\tau \leq \tau_c$ (see Eqns. 16 and 22) should be satisfied. For a Gaussian channel, the threshold τ_c corresponds to information rate $\log(1+\tau_c)$. Hence, we can interpret $\log(1+\tau_c)$ as the maximum critical rate at which the information propagates to the entire network.

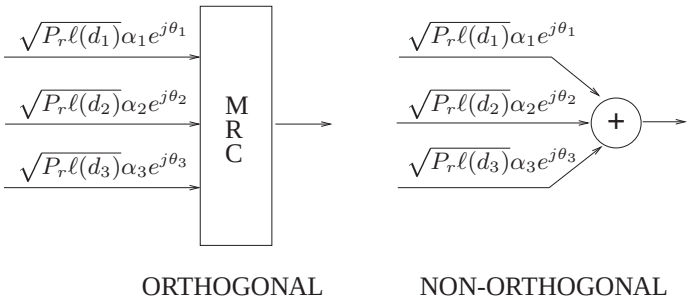

Fig. 5. The reception models for random fading corresponding to orthogonal and non-orthogonal relay transmission.

III. Derivation of Equivalent Channel Model for Fading Channels

A simplistic assumption we made in the previous section is that the power of simultaneously transmitted packets is equal to the sum of individual powers. This assumption, however, does not hold if simultaneous transmissions are not in orthogonal dimensions. In this section, we first derive a *random channel model* in which the impulse response of simultaneously transmitted packets is modelled as a proper complex Gaussian random vector:

$$\mathbf{h} = \mathcal{N}_c(0, \tilde{\Sigma}), \tag{23}$$

where $\tilde{\Sigma}$ is a covariance matrix that depends on the network physical layer. The Gaussian channel assumption, *i.e.*, *Rayleigh* distributed multi-path channel, is widely used in wireless communications [32]. In our setup, Gaussianity of $\mathbf{h}$ comes from the fact that there are many transmitters each with small power in the continuum asymptote. By choosing the $\tilde{\Sigma}$ appropriately, the model (23) can be used to take into account non-orthogonal narrowband transmissions, and in particular, the effects of channel fading, time differences between simultaneous transmissions and random phases. In sections III-A and III-B, we provide a rigorous derivation of (23) and find the covariance matrix $\tilde{\Sigma}$ as a function of network physical layer parameters. The network behavior with the random channel model is analyzed in Section IV.

We also consider the case that the transmitted signals go through orthogonal *fading* channels, and the receiver does maximal ratio combining of channel outputs (Fig. 5). Interestingly, in the continuum limit the equivalent channel converges to a deterministic constant, despite the existence of fading. We discuss this convergence in Section III-C. The behavior of the networks with orthogonal transmissions and non-orthogonal transmissions is compared in Section IV.

A. Transmitted and Received Signals for Non-orthogonal Transmissions

Consider a group of relay nodes $\mathcal{L}$ that transmit the same message simultaneously. Suppose that the message consists of M complex-valued samples $c[1], \cdots, c[M]$. Assume that all nodes in the network use an identical pulse-shaping filter $p(t)$. The base-band transmitted signal by the l'th node is modelled

as

$$s^{(l)}(t) = \sqrt{P_r} \sum_{m=1}^{M} c[m]p(t - t_l - mT), \qquad (24)$$

where T is the symbol period; t_l is the time the l'th node starts its transmission (*i.e.,* the relay time).

Consider a hypothetical node H at a given location (x, y). Let v_c denote the speed of light. The base-band received signal at node H is modelled as

$$r(t) = \sum_{l\in\mathcal{L}} \alpha_l e^{j\theta_l} \sqrt{\ell(d_l)} s^{(l)}(t - \frac{d_l}{v_c}) + w(t) \qquad (25)$$

where α_l is the fading coefficient between nodes l and H; θ_l is the phase difference between modulator and demodulator clocks at nodes l and H; d_l is the distance between nodes l and H; d_l/v_c is the propagation delay; $w(t)$ is additive channel noise.

We assume that the broadcast signal is narrowband, *i.e.,* the coherence bandwidth of equivalent link provided to the relays is much larger than the transmission bandwidth. Hence in the following, we neglect the propagation delays d_l/v_c in (25).

After sampling at the symbol rate, the baseband received signal becomes

$$r[n] \triangleq r(nT) = \sum_{m=0}^{M-1} c[m] \sum_{l\in\mathcal{L}} \sqrt{P_r}\beta_l p_l[n-m] + w[n] \quad (26)$$

where

$$\beta_l \triangleq \sqrt{\ell(d_l)}\alpha_l e^{j\theta_l} \qquad (27)$$

$$p_l[n-m] \triangleq p(nT - t_l - mT). \qquad (28)$$

For consistency with the previous section, we will assume that $w[\cdot]$ is white with unit power. If we define the channel coefficients as

$$h[n] \triangleq \sqrt{P_r} \sum_{l\in\mathcal{L}} \beta_l p_l[n], \qquad (29)$$

then the received signal can be compactly represented as

$$r[n] = h[n] * c[n] + w[n]. \qquad (30)$$

In the following, $h[n], n = 0, \pm 1, \pm 2, \cdots$ will be referred to as the *equivalent channel impulse response* of the cooperative channel. We will also assume that the $h[n]$ is practically non-zero only for $2D + 1$ terms, $h[-D], \cdots, h[D]$, and is zero elsewhere.

B. Asymptotic Channel Distribution

The set $\mathcal{L}$ is viewed as the set of nodes belonging to a certain level of broadcast. For tractability purposes, we make the following assumptions:

i) The locations of the nodes in $\mathcal{L}$ are independent and identically distributed (i.i.d.), and therefore, the distances d_l's are i.i.d. for different l.
ii) The starting times t_l's, $l \in \mathcal{L}$ are *zero-mean* i.i.d. distributed with a certain pdf $f(t)$. This models the situation that all the nodes in the same level transmit approximately around time zero, but there may be small variations due to differences between processing/relaying times at different nodes.
iii) The fading coefficients α_l's are i.i.d. with unit variance for different $l \in \mathcal{L}$. The phases θ_l are i.i.d. Uniform$[0, 2\pi]$ (*i.e.,* the modulator/demodulator clocks at different nodes are asynchronous).
iv) The $d_l,\ t_l,\ \alpha_l,\ \theta_l,\ l \in \mathcal{L}$ are independent.

The next theorem characterizes the asymptotic distribution of the channel.

Theorem 4: Let L denote the number of nodes in $\mathcal{L}$. Suppose that the relay power P_r varies with L, and the total relay power converges to P_T as $L\to\infty$, *i.e.,*

$$LP_r \to P_T, \quad \text{as } L\to\infty. \qquad (31)$$

Then, under assumptions i)-iv), the channel impulse response $\mathbf{h} \triangleq (h[-D], \cdots, h[D])$, satisfies

$$\mathbf{h} \xrightarrow{\mathrm{d}} \mathcal{N}_c(0, P_T\mathbb{E}\{\ell(d_l)\}\Sigma) \quad \text{as } L\to\infty, \qquad (32)$$

where $\xrightarrow{\mathrm{d}}$ denotes convergence in distribution, and Σ is a matrix with entries

$$\Sigma[n, m] = \int_{-\infty}^{\infty} f(t)p(nT - t)p^*(mT - t)dt,$$

for $n,\ m \in \{-D, \cdots, D\}$.

Proof: Rearrange $h[n]$ to get

$$h[n] = \sqrt{LP_r}\left[\frac{h[n]}{\sqrt{LP_r}}\right] = \sqrt{LP_r}\left[\frac{1}{\sqrt{L}}\sum_{l\in\mathcal{L}} \beta_l p_l[n]\right] \qquad (33)$$

Here, the first quantity converges to $\sqrt{P_T}$ as $L\to\infty$, and the expression inside the parenthesis is the addition of L i.i.d. copies of the same signal divided by $\sqrt{L}$. We will use the Multivariate Central Limit Theorem (CLT) [31] to obtain (32). In order to apply CLT, observe that $\mathbb{E}\{\beta_l p_l[n]\} = 0$, since each β_l has zero mean. Furthermore, the autocovariance of $(\beta_l p_l[n] : -D \le n \le D)$ is given by

$$\begin{aligned} R[n, m] &= \mathbb{E}\{\beta_l\beta_l^* p_l[n]p_l^*[m]\} \\ &= \mathbb{E}\{|\beta_l|^2\}\mathbb{E}\{p_l[n]p_l^*[m]\} \\ &= \mathbb{E}\{\ell(d_l)\}\Sigma[n, m]. \end{aligned} \qquad (34)$$

On the other hand, the pseudo autocovariance of $(\beta_l p_l[n] : -D \le n \le D)$ is

$$\bar{R}[n, m] = \mathbb{E}\{\beta_l^2\}\mathbb{E}\{p_l[n]p_l[m]\} = 0, \qquad (35)$$

because each θ_l is Uniform$[0, 2\pi]$. Eqn. 35 implies that $(\beta_l p_l[n] : -D \le n \le D)$ is proper complex. From Multivariate CLT, we get that

$$\frac{\mathbf{h}}{\sqrt{LP_r}} \xrightarrow{\mathrm{d}} \mathcal{N}_c(0, \mathbb{E}\{\ell(d_l)\}\Sigma).$$

The theorem follows. ■

C. Orthogonal Transmissions over Fading Channels

Another possibility is to consider the above model under the assumption that the transmissions are in orthogonal channels (*e.g.,* as in FDMA). In this scenario, for simplicity, we will assume that the receiver can perfectly recover the timing of each transmitted signal and sample it at time zero. Hence, the sampled received signal from the l'th node is

$$r_l[n] = \sum_{m=0}^{M} c[m]\sqrt{P_r}\beta_l p[n-m] + w_l[n], \qquad (36)$$

where $p[n-m] \triangleq p(nT-mT)$, and $w_l[n]$ is the noise in the l'th orthogonal channel. We assume that each $w_l[\cdot]$ is white with unit power, and is independent of one another.

If a Nyquist pulse $p(\cdot)$ is used, the signal $r_l[\cdot]$ has no intersymbol interference (*i.e.,* $p[n-m] = p(0)\delta[n-m]$). Consequently, maximal ratio combining of $r_l[\cdot],\ l \in \mathcal{L}$ gives the highest SNR:

$$\text{SNR} = P_r \sum_{l\in\mathcal{L}} |\beta_l|^2 |p(0)|^2. \qquad (37)$$

Theorem 5: Under assumptions i)-iv) and (31), the SNR of the maximal-ratio-combined received signal converges to a deterministic limit as $L\to\infty$, *i.e.,*

$$\text{SNR} \to P_T \mathbb{E}\{\ell(d_l)\}|p(0)|^2, \quad \text{as } L\to\infty, \qquad (38)$$

almost surely.

Remark: Notice that the bandwidth requirement of the orthogonal system is proportional to L if FDMA or CDMA type signaling is used. Hence, taking the limit $L\to\infty$ without sacrificing from the transmission rate requires potentially infinite bandwidth, *i.e.,* the network operates in the wideband regime. On the other hand, by using distributed orthogonal space-time codes, it may be possible to get to the above performance without sacrificing bandwidth.

Proof: Notice that

$$\text{SNR} = LP_r \left[\frac{1}{L}\sum_{l\in\mathcal{L}} |\beta_l|^2 |p(0)|^2\right].$$

Here, the first term converges to P_T and the second term converges almost surely to its mean due to the strong law of large numbers. The theorem follows. ∎

The theorem indicates that the system with orthogonal channels, despite the existence of fading and randomness in the channel, has a deterministic SNR in the limit. As we will see in the next section, this property greatly simplifies the analysis in certain cases.

IV. Network Behavior with Random Channels

In this section, we derive the continuum model for networks with random channels. Following the order of presentation in Section II, we will first describe the network with random topology, and obtain its continuum model in the limit. This will be done both for orthogonal and non-orthogonal transmission models described in Section III-A. The broadcast performance of these models will be compared in Section IV-E.

A. Random Network

Consider a network with N relay nodes located in the disc $\mathbb{S}$, where the source is located at the center. For a subset of relay nodes $\mathcal{L} \subset \{1,\cdots,N\}$, let

$$\mathbf{h}_{\mathcal{L}}(x,y) = \begin{cases} \sqrt{P_r}\sum_{l\in\mathcal{L}} \beta_l(x,y)\mathbf{p}_l, & \text{non-orthogonal} \\ \sqrt{P_r}\sqrt{\sum_{l\in\mathcal{L}} |\beta_l(x,y)|^2}\mathbf{p}, & \text{orthogonal} \end{cases} \qquad (39)$$

be the channel impulse response vector from level set $\mathcal{L}$ to a hypothetical node at (x,y) in the non-orthogonal and orthogonal channel models. In (39), we made the dependence on the location of the receiving node (x,y) explicit wherever possible. Here, $\mathbf{p}_l = (p(nT-t_l) : -\infty < n < \infty)$ refers to the pulse shaping filter delayed by t_l sampled at the Nyquist rate, and $\mathbf{p} = (p(0)\delta[n] : -\infty < n < \infty)$ is the sampled filter without delay. For the source transmission, the channel impulse response is

$$\mathbf{h}_0(x,y) = \sqrt{P_s}\beta_0(x,y)\mathbf{p} \qquad (40)$$

for both the non-orthogonal and orthogonal transmissions.

The decision criterion of when to relay packets is a subtle issue in intersymbol interference channels. In practice, the packets are coded according to a certain channel code, and CRC (Cylic Redundancy Check) bits are placed into each packet. A packet reception is considered successful if the CRC test passes after decoding the channel code. A simple way to model this phenomena is via the notion of matched-filter upper bound $||\mathbf{h}_{\mathcal{L}}(x,y)||^2$ on received SNR, *i.e.,* to consider a reception successful if $||\mathbf{h}_{\mathcal{L}}(x,y)||^2$ exceeds a certain threshold τ. A more elaborate model for receptions can be derived based on the notion of outage capacity (*i.e.,* a reception is considered successful if the instantaneous mutual information of the equivalent channel exceeds a certain threshold). However, in this paper we will focus on the simpler matched-filter upper bound approach.

Let $\mathcal{S} = \{(x_i,y_i) : i = 1,\cdots,N\}$ be the set of relay nodes that are randomly and uniformly distributed in $\mathbb{S}$. If the channel is random, then the set of level-1 nodes is given by

$$\mathcal{S}_1 = \{(x,y) \in \mathcal{S} : ||\mathbf{h}_0(x,y)||^2 \geq \tau\}.$$

The locations of the level-k nodes are given by

$$\mathcal{S}_k = \{(x,y) \in \mathcal{S} \setminus \bigcup_{i=1}^{k-1} \mathcal{S}_i :\ ||\mathbf{h}_{\mathcal{L}_{k-1}}(x,y)||^2 \geq \tau\}, \qquad (41)$$

where $\mathcal{L}_{k-1} \subset \{1,\cdots,N\}$ is the index set of level $k-1$ nodes.

Throughout this section, we will assume that t_l's are i.i.d. for different l, $\alpha_l(x,y)$'s are i.i.d. for different l and (x,y); so are $\theta_l(x,y)$'s. Moreover, t_l's are zero-mean with pdf $f(t)$, $\theta_l(x,y)$'s are Uniform$[0,2\pi]$, and $\alpha_l(x,y)$, t_l and $\theta_l(x,y)$'s are independent from one another. These assumptions extend ii)-iv) in Section III-B considering the spatial domain (x,y) as well as the node index l. We assume spatially independent fading for its simplicity. For correlated fading scenarios, refer to Section IV-G.

B. Continuum Network

In this section, we derive the continuum model for the network with random channels using the channel asymptotics derived in Section III.

In the random network, a node at location (x, y) receives the source transmission successfully with probability

$$P_1(x, y) = \Pr\{||\mathbf{h}_0(x, y)||^2 \geq \tau\}.$$

This is also the probability that a node at (x, y) joins level-1. It follows from the law of large numbers that for each measurable set $\mathbb{U} \subset \mathbb{S}$, the number of level-1 nodes in $\mathbb{U}$ scales as $\iint_{\mathbb{U}} \rho P_1(x', y')dx'dy'$, *i.e.,*

$$\frac{|\mathbb{U} \cap \mathcal{S}_1|}{\iint_{\mathbb{U}} \rho P_1(x', y')dx'dy'} \to 1 \quad \text{as } \rho \to \infty \tag{42}$$

almost surely. When $P_r\rho$ is fixed to $\bar{P}_r$, the total transmit power of level-1 nodes $P_r|\mathbb{U} \cap \mathcal{S}_1|$ converges to

$$P_T = \iint_{\mathbb{S}} \bar{P}_r P_1(x', y')dx'dy' \tag{43}$$

almost surely. Furthermore, the locations of level-1 nodes are distributed according to the density $\tilde{\rho}(x', y') \triangleq \frac{P_1(x',y')}{\iint_{\mathbb{S}} P_1(x',y')dx'dy'}$. Hence,

$$\mathbb{E}\{\ell(d_l(x, y))\} = \iint_{\mathbb{S}} \tilde{\rho}(x', y')\ell(x - x', y - y')dx'dy', \tag{44}$$

for all $l \in \mathcal{L}_1$, and the total received power at location (x, y) due to level-1 transmissions is

$$\begin{aligned}\sigma_1^2(x, y) &\triangleq P_T\mathbb{E}\{\ell(d_l(x, y))\} \\ &= \iint_{\mathbb{S}} \bar{P}_r P_1(x', y')\ell(x - x', y - y')dx'dy'. \end{aligned}\tag{45}$$

If this $\sigma_1^2(x, y)$ is substituted into Theorems 4 and 5, we see that a node at (x, y) receives the level-1 transmission successfully with probability $\Pr\{||\mathbf{h}_1(x, y)||^2 \geq \tau\}$, where

$$\begin{aligned}\mathbf{h}_1(x, y) &\sim \mathcal{N}_c(0, \sigma_1^2(x, y)\Sigma), \quad \text{non-orthogonal} \\ ||\mathbf{h}_1(x, y)||^2 &= \sigma_1^2(x, y)|p(0)|^2, \quad \text{orthogonal}\end{aligned}$$

is the equivalent channel distribution in the limit that the number of level-1 nodes goes to infinity. The probability that a node at (x, y) joins level-2 is

$$\begin{aligned}P_2(x, y) &= \Pr\{\text{receives from level-1,} \\ &\qquad \text{does not receive from the source}\} \\ &= \Pr\{||\mathbf{h}_1(x, y)||^2 \geq \tau\} \cdot \\ &\qquad [1 - \Pr\{||\mathbf{h}_0(x, y)||^2 \geq \tau\}]. \end{aligned}\tag{46}$$

Now, we can generalize what is done so far.

Definition Let $P_k(x, y)$ denote the probability that a node at location (x, y) joins level-k, and $\sigma_k^2(x, y)$ be the sum of signal powers from level-k at location (x, y). For $k = 1, 2, 3, \cdots$, the equations

$$\begin{aligned}P_{k+1}(x, y) &= \Pr\{||\mathbf{h}_k(x, y)||^2 \geq \tau\} \cdot \\ &\qquad \prod_{i=0}^{k-1} [1 - \Pr\{||\mathbf{h}_i(x, y)||^2 \geq \tau\}], \end{aligned}\tag{47}$$

$$\sigma_k^2(x, y) = \iint_{\mathbb{S}} \bar{P}_r P_k(x', y')\ell(x - x', y - y')dx'dy', \tag{48}$$

where

$$\begin{aligned}\mathbf{h}_k(x, y) &\sim \mathcal{N}_c(0, \sigma_k^2(x, y)\Sigma), \quad \text{non-orthogonal} \\ ||\mathbf{h}_k(x, y)||^2 &= \sigma_k^2(x, y)|p(0)|^2, \quad \text{orthogonal}\end{aligned}$$

specify the *continuum model* for networks with random channels.

The functions P_k, σ_k^2 define a non-linear dynamical system which evolves with k. Although the analytical solution of this system is hard, it can be usually evaluated numerically. Another property of the continuum model is that the $P_k(x, y)$ and $\sigma_k^2(x, y)$ are only functions of $r = \sqrt{x^2 + y^2}$. Therefore, the above dynamical system evolves only over 1-dimensional functions. For convenience, we will use the notations $P_k(x, y)$ and $P_k(r)$ interchangeably.

For our numerical evaluations in this section, we will use the following path-loss model

$$\ell(d) \triangleq \begin{cases} 1/d^2 & d_0 \leq d \\ 1/d_0^2 & 0 \leq d \leq d_0, \end{cases} \tag{49}$$

with a small $d_0 > 0$ to avoid the singularity in the integral (48). The squared-distance attenuation model $\ell(d) = 1/d^2$ comes from the free-space attenuation of electromagnetic waves, and it *does not* hold when d is very small [33]. This issue has been recognized by several researchers (*e.g.,* [34], [35]). We expect the results obtained from (49) to be more practically relevant.

C. Behavior of Continuum Network with Orthogonal Channels

The above equations for continuum network greatly simplify in case of orthogonal channels. There are two possibilities:

i) If there is no fading from the source to the relays, then $P_1(x, y)$ is binary (*i.e.,* $P_1(x, y) \in \{0, 1\}$). Furthermore, each $\sigma_k^2(x, y)$ for $k = 2, 3, \cdots$ is deterministic. Therefore, $\Pr\{||\mathbf{h}_k(x, y)||^2 \geq \tau\}$ is binary as well. If we define

$$\mathbb{S}_k \triangleq \{(x, y) \in \mathbb{S} : P_k(x, y) = 1\},$$

then

$$\sigma_k^2(x, y) = \iint_{\mathbb{S}_k} \bar{P}_r \ell(x' - x, y' - y)dx'dy'.$$

Hence, for this scenario the continuum model reduces to the continuum model for deterministic channels.

ii) If the channels from the source to relays have fading, then $P_1(x, y)$ takes continuous values in $[0, 1]$, but $\Pr\{||\mathbf{h}_k(x, y)||^2 \geq \tau\}$ is still binary. In our numerical evaluations we observed that the effect of $P_1(x, y)$ is transient, and the asymptotic behavior of the network is as in the deterministic model. In Fig. 6 and 7, we plot $P_k(r)$ in both low and high threshold regimes.

D. Behavior of Continuum Network with Non-orthogonal Channels

The following lemma simplifies the task of computing $\Pr\{||\mathbf{h}_k(x, y)||^2 \geq \tau\}$ for non-orthogonal transmissions.

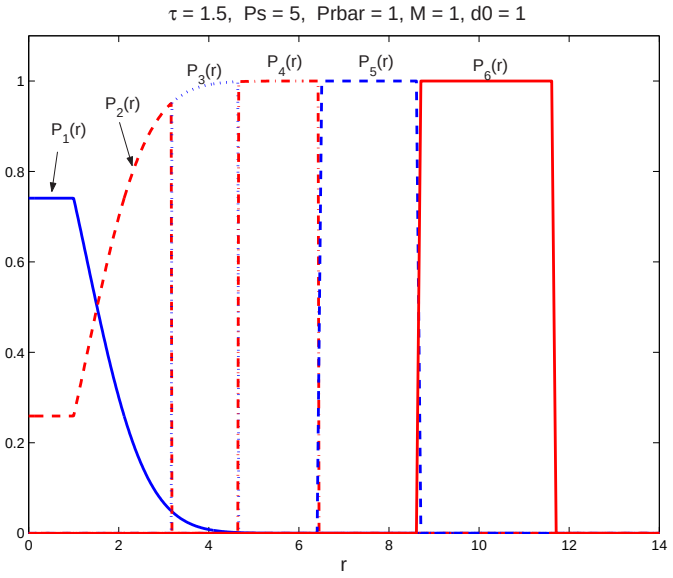

Fig. 6. The parameters are $\tau = 1.5, P_s = 5, \bar{P}_r = 1, d_0 = 1, \Sigma = 1$. The transmissions continue.

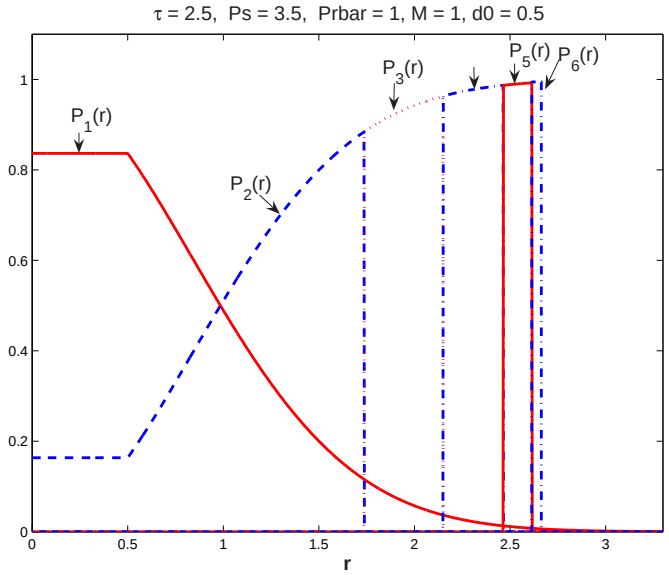

Fig. 7. The parameters are $\tau = 2.5, P_s = 3.5, \bar{P}_r = 1, d_0 = 0.5, \Sigma = 1$. The transmissions die out.

Lemma 2: Let $\lambda_1, \cdots, \lambda_M$ denote the distinct eigenvalues of Σ with multiplicities $a_1, \cdots, a_M$, respectively. The characteristic function of $||\mathbf{h}_k(x,y)||^2$, where $\mathbf{h}_k(x,y) \sim \mathcal{N}(0, \sigma_k^2(x,y)\Sigma)$, is given by

$$\Phi_k(jw) = \prod_{i=1}^{M} \frac{1}{(1 + jw\lambda_i \sigma_k^2(x,y))^{a_i}}. \quad (50)$$

Let $\Phi_k(jw)$ have the partial fraction expansion

$$\Phi_k(jw) = \sum_{i=1}^{M} \sum_{m=1}^{a_i} \frac{A_{im}}{(1 + jw\lambda_i \sigma_k^2(x,y))^m}. \quad (51)$$

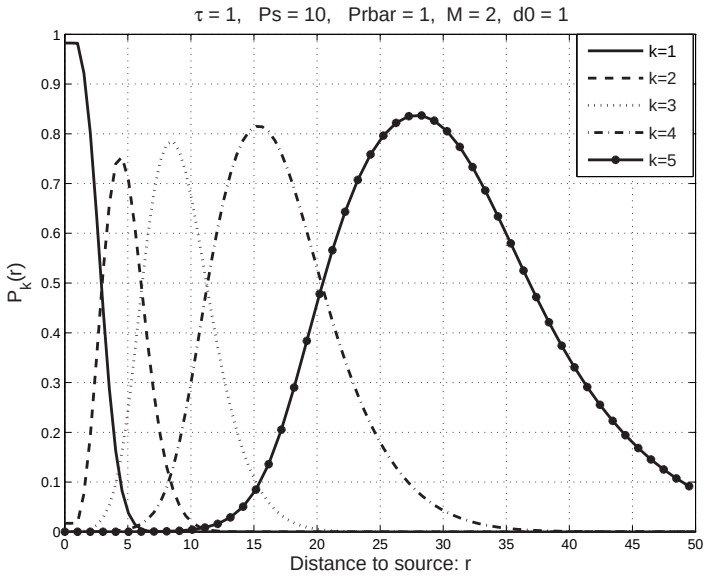

Fig. 8. Transmissions continue. The parameters are $\tau = 1, P_s = 10, \bar{P}_r = 1, M = 2, d_0 = 1, \Sigma = \frac{1}{M}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

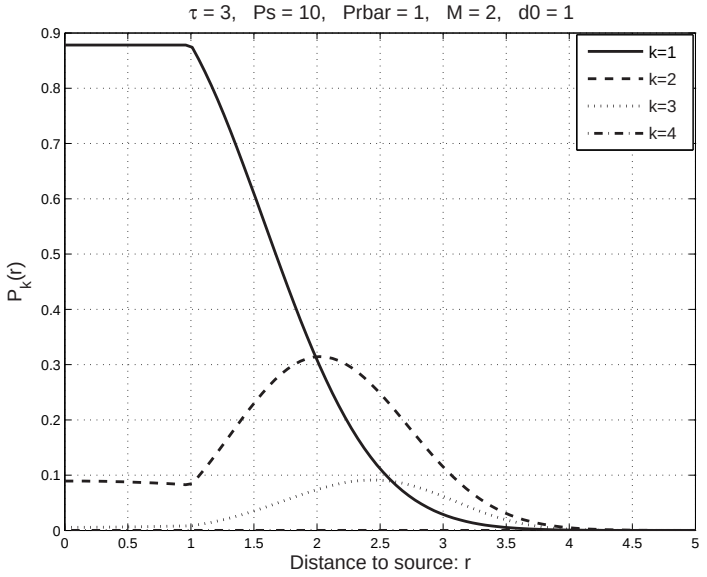

Fig. 9. Transmissions die out. The parameters are $\tau = 3, P_s = 10, \bar{P}_r = 1, M = 2, d_0 = 1, \Sigma = \frac{1}{M}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Then,

$$\Pr\{||\mathbf{h}_k(x,y)||^2 \geq \tau\} = \sum_{i=1}^{M} \sum_{n=1}^{a_i} \frac{A_{in}}{(n-1)!} \Gamma(n, \frac{\tau}{\lambda_i \sigma_k^2(x,y)}), \quad (52)$$

where $\Gamma(a,x) = \int_x^\infty e^{-t} t^{a-1} dt$. If the eigenvalues of Σ are distinct, then the above expression simplifies to

$$\Pr\{||\mathbf{h}_k(x,y)||^2 \geq \tau\} = \sum_{i=1}^{M} A_{i1} e^{-\tau/\sigma_k^2(x,y)\lambda_i}. \quad (53)$$

Remark: Computation of the right hand side of (52)-(53) can be done in two steps. First, apply an eigenvalue decomposition

to Σ to obtain $\lambda_1, \cdots, \lambda_M$. Then, apply partial fraction expansion to (50) to get A_{im}'s.

Proof: The proof can be easily done utilizing well-known techniques [18]. ■

The analytical solution of the continuum network in the case of non-orthogonal channels appears to be a non-trivial problem. In order to gain intuition, we evaluated (47) and (48) numerically for large R. Similar to the case of deterministic channels, it is observed that there exists a critical threshold τ_c. For $\tau > \tau_c$, the transmissions eventually die out, *i.e.,*

$$\sup_{(x,y)\in\mathbb{R}^2} P_k(x,y) \to 0 \quad \text{as} \quad k \to \infty.$$

Otherwise, the transmissions propagate to the whole network, while the level curves, $P_k(r), r \in \mathbb{R}$, become wider as k increases. See Figs. 8 and 9 for these two regimes.

In general, we do not have an explicit characterization of τ_c. However, the following Theorem gives a sufficient condition for the transmissions to die out.

Theorem 6: Consider an infinite disc (*i.e.,* $\mathbb{S} = \mathbb{R}^2$). If the path-loss model satisfies

$$\iint_{\mathbb{R}^2} l(x',y')dx'dy' < \infty, \tag{54}$$

then there exists $\tau_c < \infty$ such that if $\tau > \tau_c$, then the transmissions eventually die out, *i.e.,*

$$\sup_{(x,y)\in\mathbb{R}^2} P_k(x,y) \to 0, \quad \text{as} \quad k \to \infty. \tag{55}$$

Remark: Condition (54) does not hold for (49). However, it is satisfied for all $\ell(d) = \min\{1/d_0^u, 1/d^u\}, \quad u > 2, \ d_0 > 0$.

Proof: See Appendix III. ■

E. Comparison between non-orthogonal and orthogonal cooperative broadcast

In this section, we compare the message propagation behavior in random orthogonal and random nonorthogonal channels. Fig. 10 shows the $P_k(r)$ for both models. These plots are obtained for the parameters $\tau = 1, P_s = 5, \bar{P}_r = 1, M = 1, d_0 = 1$.

For both orthogonal and non-orthogonal models, $P_1(\cdot)$ is the same; however, the level curves $P_k(\cdot)$ differ significantly for large k. In particular, the level curves in the non-orthogonal case move faster. This is a rather counter-intuitive result, because the orthogonal system (with FDMA/CDMA) uses much more bandwidth than the non-orthogonal one, and the use of orthogonal channels is more reliable in the sense that the receiver eliminates/reduces the effects of fading via maximal ratio combining. Our result implies that the system *with fading* provides better broadcast behavior than the one without.

We believe that this fact can be explained as follows. The maximal ratio combining method reduces the probability that the combined signal experiences a deep fade at the cost of reducing the probability that the signal experiences a favorable fade. In a dense network, favorable fading realizations are very valuable, because when the node density is high, although there is a small probability of having a good fading realization, there is always a fraction of nodes that experiences them.

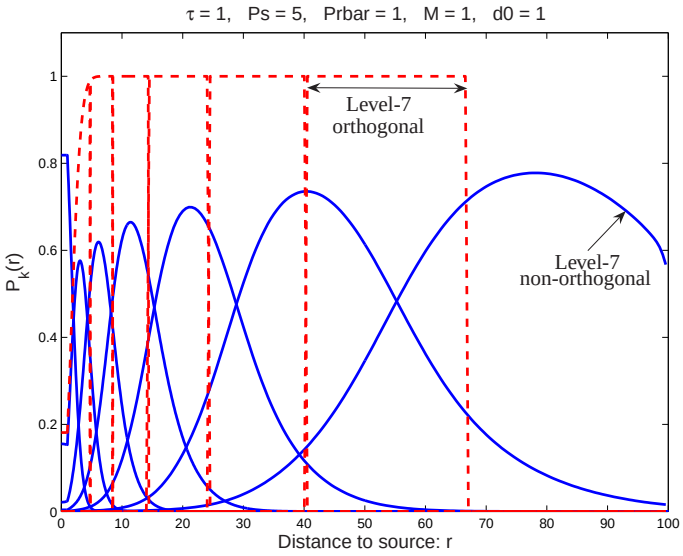

Fig. 10. Wideband orthogonal vs. narrowband non-orthogonal: the upper, rectangular shaped levels are the orthogonal, the lower wave-like levels are non-orthogonal. The parameters are $\tau = 1, P_s = 5, \bar{P}_r = 1, M = 1, d_0 = 1, \Sigma = 1$.

Once these lucky nodes receive and retransmit, the nodes neighboring them see a boost of signal power because of the properties of $\ell(d)$. In conclusion, we believe that the nodes that enable faster level movement are the ones at the forefront of each level.

In the narrowband system, favorable fading realizations occur, when the phases of two or more simultaneously transmitting nodes happen to add coherently, or when one of the transmitting nodes experiences a very good channel with the receiver. Considering that non-orthogonal transmissions do not require infinite bandwidth, we conclude that the non-orthogonal scheme is more advantageous also in terms of end-to-end delay.

F. Multihop diversity under fading channels

If the nodes listen to m previous levels, the approach presented in the previous sections using the matched filter bound can be generalized by modifying (47) as

$$P_{k+1}(x,y) = \Pr\{\sum_{j=(k-m+1)_+}^{k} ||\mathbf{h}_j(x,y)||^2 \geq \tau\} \cdot \prod_{i=0}^{(k-1)} [1 - \Pr\{\sum_{j=(i-m+1)_+}^{i} ||\mathbf{h}_j(x,y)||^2 \geq \tau\}], \tag{56}$$

where $(x)_+ \triangleq \max(0,x)$. In conjunction with (48), Eqn. 56 determines the network behavior.

In case of narrowband non-orthogonal transmissions, the multihop diversity allows the signal to flow much faster (see Fig. 11). Furthermore, we expect to observe the threshold behavior. In case of wideband orthogonal transmissions, the behavior converges to the deterministic channel behavior in Section II-D.

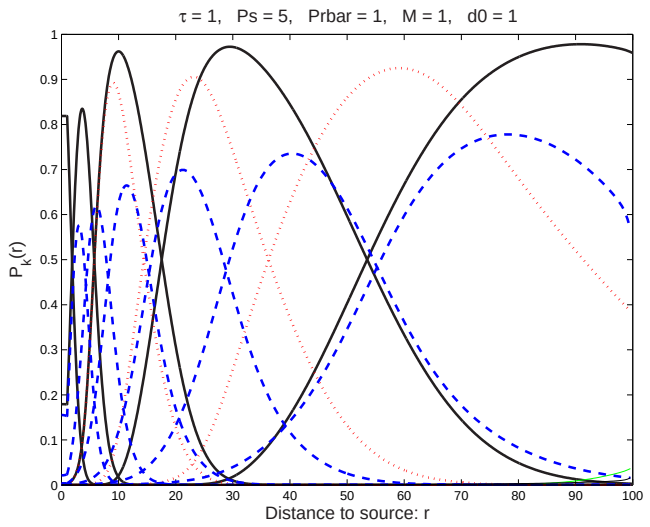

Fig. 11. Narrowband non-orthogonal transmissions: $\tau = 1, P_s = 5, \bar{P}_r = 1, M = 1, d_0 = 1, \Sigma = 1$. Three different scenarios: (i) straight line $m = 3$, (ii) dotted line $m = 2$ (iii) dashed line $m = 1$. Note that first level curve $P_1(r)$ is the same for $m = 1, 2, 3$ and the second level curve $P_2(r)$ is the same for $m = 2, 3$.

G. Extensions to correlated fading

In the previous sections, in order to derive our results, we assumed that the small-scale fading is spatially independent. The spatial independence is not actually needed in order to derive the results. The underlying assumption needed for the results in the paper is that the channel coefficients between any pair of nodes are independent from one another. It is well-known in multiple-input multiple-out (MIMO) literature, it is reasonable to model that the channel gains at two receiving antennas separated by more than half the wavelength are independent. Hence, in a rich scattering environment (such as urban areas), the independent fading assumption is expected to be accurate.

The main difference of our problem from a MIMO transmitter is that we have different transmitters, hence the clocks at different transmitters are naturally assumed to be asynchronous. In addition to the rich scattering environments, considering an environment where *there is no small scale fading* ($\alpha(x, y) =$ constant $\forall(x, y)$) such that the channel gains depend only on the independent random phases and pathloss gain, our analysis for the independent fading applies as it is.

There are also scenarios where the independent fading assumption is not valid. In the case of spatially dependent fading, the analysis is intractable in general. For example, correlated fading implies that certain regions of the network can reliably receive the source message and the rest can not. Because of this reason, in general the set of level-1 nodes looks like "swiss cheese" (see Fig. 12), containing dark regions that are in deep fade and can not receive source message. Hence the analysis becomes quite intractable for the higher levels.

Next, we provide two new methods to deal with the spatially correlated scenarios: (i) a conditional continuum model based on the instantaneous realization of the fading process α; (ii) a model for spatially dependent fading such that the CLT for dependent random variables holds.

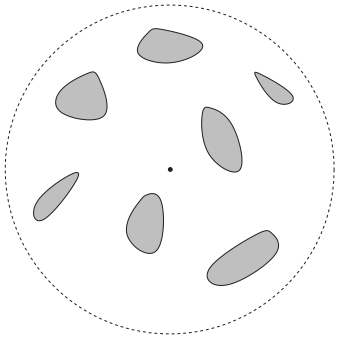

Fig. 12. First level nodes for the correlated fading model

In the following, we assume that the value of small scale fading is a function of the locations of the transmitter and the receiver. That is, assume the transmitter is located at (x', y') and the receiver is located at (x, y), then the small scale fading coefficient is denoted by $\alpha(x, y, x', y')$. Also, we assume that the phase shifts of different transmitters are independent. Next we describe the network behavior.

i) Consider a specific realization of $\alpha(x, y, x', y')$, which is assumed to be continuous with respect to (x, y), (x', y'). Given $\alpha(x, y, x', y')$ and the locations (x, y), (x', y'), define $\beta(x, y, x', y') \triangleq \sqrt{l(x - x', y - y')}\alpha(x, y, x', y')e^{\theta(x', y')}$. Note that $\beta(x, y, x', y')$'s are conditionally independent for all (x', y') since $\theta(x', y')$'s are independent. The Eqns. 33-41 are still valid in this scenario. The equations in Section IV-B can be updated as:

$$\begin{aligned} P_1(x, y|\alpha) &= \Pr\{||\mathbf{h}_0(x, y|\alpha)||^2 \geq \tau\} \\ \sigma_1^2(x, y|\alpha) &= \iint_{\mathbb{S}} \bar{P}_r |\alpha(x, y, x', y')|^2 \cdot \\ &\qquad \ell(x - x', y - y')P_1(x', y')dx'dy' \end{aligned}$$

where $\mathbf{h}_0(x, y|\alpha) = \sqrt{P_s}\beta_0(x, y, x_s, y_s)\mathbf{p}$ and the location of the source (x_s, y_s) is known and deterministic. Furthermore, for any level-k,

$$\begin{aligned} P_{k+1}(x, y|\alpha) &= \Pr\{||\mathbf{h}_k(x, y|\alpha)||^2 \geq \tau\} \cdot \\ &\qquad \prod_{i=0}^{k-1} [1 - \Pr\{||\mathbf{h}_i(x, y|\alpha)||^2 \geq \tau\}], \\ \sigma_k^2(x, y|\alpha) &= \iint_{\mathbb{S}} \bar{P}_r |\alpha(x, y, x', y')|^2 \cdot \\ &\qquad \ell(x - x', y - y')P_k(x', y')dx'dy' \end{aligned}$$

where

$$\begin{aligned} \mathbf{h}_k(x, y|\alpha) &\sim \mathcal{N}_c(0, \sigma_k^2(x, y|\alpha)\Sigma), \quad \text{non-orthogonal} \\ ||\mathbf{h}_k(x, y|\alpha)||^2 &= \sigma_k^2(x, y|\alpha)|p(0)|^2, \quad \text{orthogonal.} \end{aligned}$$

The above equations specify a conditional continuum model given the instantaneous realizations of α (*i.e.*, we treat α as a given deterministic function). One drawback of the conditional model (in general, correlated fading),

is the network evolution is strictly a function of the instantaneous fading realization. Therefore, the continuum model exhibits different behavior for different realizations of the fading process, and it is hard to come up with conclusions.

ii) If there exists a fading model for $\alpha(x,y,x',y')$ such that the CLT for dependent random variables holds, then similar to the previous case, a continuum model can be derived. We know that certain generalizations of the cental limit theorem exist for dependent random variables [31]. Unfortunately, it is hard to directly apply these results in the literature to our problem. All we can claim at the moment is that such conditions are in general looser than the condition of independence.

For the fading models where the CLT for dependent random variables holds, the Eqns. 33-41 are still valid. The equations in Section IV-B can be updated as:

$$P_1(x,y) = \Pr\{||\mathbf{h}_0(x,y)||^2 \geq \tau\}.$$

Given the location (x,y), the signal power received from level-1 at (x,y) is

$$\begin{aligned}\sigma_1^2(x,y) &= P_T \mathbb{E}_\ell\{|\beta_\ell(x,y,x_\ell,y_\ell)|^2\} \\ &= \iint_{\mathbb{S}} \mathbb{E}\{|\alpha(x,y,x',y')|^2\}\bar{P}_r \cdot \\ &\qquad P_1(x',y')\ell(x-x',y-y')dx'dy'\end{aligned}$$

Furthermore, for any level-k,

$$\begin{aligned}P_{k+1}(x,y) &= \Pr\{||\mathbf{h}_k(x,y)||^2 \geq \tau\} \cdot \\ &\qquad \prod_{i=0}^{k-1}[1-\Pr\{||\mathbf{h}_i(x,y)||^2 \geq \tau\}] \\ \sigma_k^2(x,y) &= \iint_{\mathbb{S}} \mathbb{E}\{|\alpha(x,y,x',y')|^2\}\bar{P}_r \cdot \\ &\qquad P_k(x',y')\ell(x-x',y-y')dx'dy'\end{aligned}$$

where

$$\begin{aligned}\mathbf{h}_k(x,y) &\sim \mathcal{N}_c(0,\sigma_k^2(x,y)\Sigma), \quad \text{non-orthogonal} \\ ||\mathbf{h}_k(x,y)||^2 &= \sigma_k^2(x,y)|p(0)|^2, \quad \text{orthogonal.}\end{aligned}$$

V. Simulation Results

In this section, we check the accuracy of continuum approximations in predicting the behavior of the random network.

A. Deterministic Channel Model

First, in Table 1, we compare the approximate number of nodes that transmit in each level, calculated through continuum analysis, which is equal to $\pi\rho(a_k - a_{k-1})$ [see Eqn. 9], and the average number of nodes that belong to a certain level obtained from simulating 100 random networks. Specifically, for $\tau = 1$, $\bar{P}_r = 1$, $P_s = 1$, Table 1 shows the ratio, $\frac{|S_k|}{\pi\rho(a_k - a_{k-1})}$, for $k = 1,\ldots,4$, and for $\rho = 1, 10, 100$. Similarly, in Table 2, we compare the approximate radii calculated through the continuum analysis, Eqn. 13, and the averaged radii obtained from simulating 100 random networks. Specifically, for $\tau = 1$, $\bar{P}_r = 1$, $P_s = 1$, Table 2 shows the ratio, $\frac{R_k}{\sqrt{a_k}}$, for $k = 1,\ldots,4$, and for $\rho = 1, 10, 100$, where R_k is the maximum distance between the source and nodes in $\mathcal{S}_k$. It can be clearly seen that in both tables the ratios between the asymptotic value and the numerical average tend to 1 as the node density increases.

k / ρ	1	10	100
1	0.7815	0.9651	0.9968
2	0.7428	0.9586	0.9926
3	0.6916	0.9534	0.9900
4	0.6615	0.9491	0.9895

TABLE I

The ratio of expected number of nodes in each level to the approximate number of nodes, $\frac{|S_k|}{\pi\rho(a_k - a_{k-1})}$

k / ρ	1	10	100
1	0.8992	0.9853	0.9968
2	0.9921	1.0307	1.0209
3	1.0214	1.0393	1.0221
4	1.0165	1.0327	1.0176

TABLE II

The ratio of expected radius of each level disc to approximate radius, $\frac{R_k}{\sqrt{a_k}}$

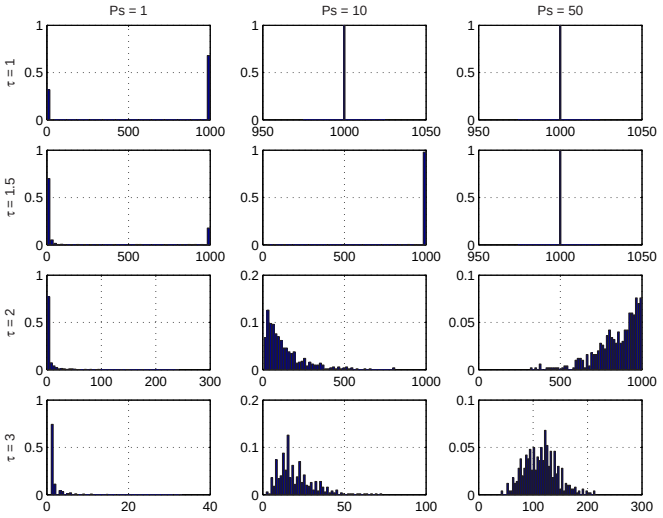

Fig. 13. Probability of number of nodes that transmit in a 1000 node network, $\rho = 1$, $P_r = 1$. Note that the scales for both horizontal and vertical axis are different.

Figure 13 shows the histogram of the number of nodes reached by broadcast as a function of N, τ, and P_s. To obtain the histograms, we simulated random networks with 1000 nodes (500 trials, $\rho = 1$, $P_r = 1$). The values are plotted for different thresholds, τ, and source powers, P_s. The theoretical critical threshold is at $(\pi \ln 2)P_r\rho \approx 2.18$ [Eqn. (16)]. With the continuum model, for $\tau < 2.18$, the predicted number of nodes with SNR $\geq \tau$ corresponds to all nodes (1000 in this case) since the $a_k \to \infty$. Instead for $\tau > 2.18$, the number of nodes

with SNR $\geq \tau$ can be calculated from (17). For $P_s = 10, 50$ and $\tau = 1, 1.5, 3$, the continuum model correctly predicts the network behavior; all nodes in the network are reached for the values of the threshold $\tau = 1, 1.5$, and only a fraction is reached for $\tau = 3$. For $\tau = 2$, which is close to the critical threshold ≈ 2.18, and for $P_s = 50$ the signal reaches to the whole network; but this is no longer the case for $P_s = 1, 10$ in contrast to what the asymptotic analysis predicts.

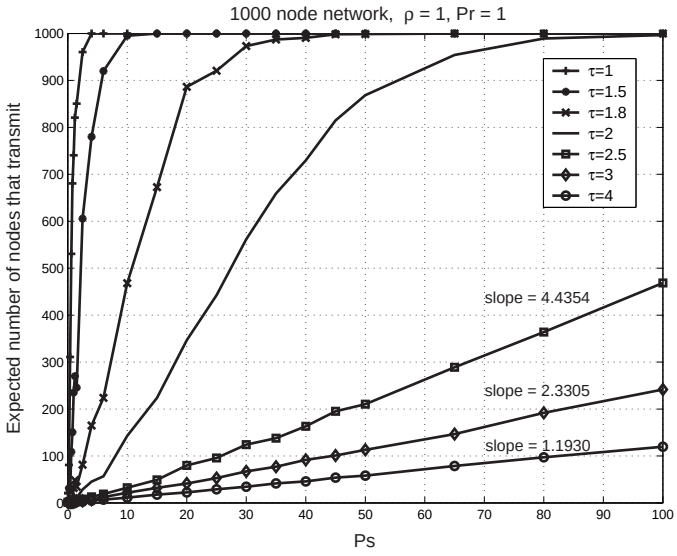

Fig. 14. Expected number of nodes that transmits vs P_s

Here, we see that the continuum approximation is not accurate in all cases ($\tau = 1$, $P_s = 1$) because we are testing the network exactly at the asymptotic threshold. This behavior is caused by the fact that the source power is too low, and for certain random network configurations, retransmission never starts. In random networks this possibility can only be avoided by boosting sufficiently the source power, P_s to initiate the broadcast with sufficient power.

Fig. 14 shows the expected number of nodes reached by the broadcast as a function of P_s and τ. Clearly, the smaller the τ, the smaller P_s is needed. Two regimes can be identified in the figure; for $\tau < 2.18$, the whole network is reached after some P_s. However, for $\tau > 2.18$, the expected number of nodes grows linearly with P_s. This is in accordance with (17). The slope s predicted by the continuum analysis is $s = \frac{\rho\pi(\mu-1)}{\tau(\mu-2)}$. Hence, for $\tau = 3$, $s = 2.7969$, and for $\tau = 4$, $s = 1.2849$. Note that these slopes are reasonably close to the predicted values. On the other hand, for $\tau = 2.5$, $s = 7.0702$, the expected slope is quite different than the simulated value, 4.4354. This is due to the fact that $\tau = 2.5$ is close to the critical threshold 2.18.

Fig. 15 gives a typical realization of a 1000-node network. Here, the dotted lines show the level radii estimated from the continuum approximation. Also, the nodes belonging to S_k, $k = 1, \ldots, 4$ are shown. The asymptotic analysis accurately predicts the locations of level sets.

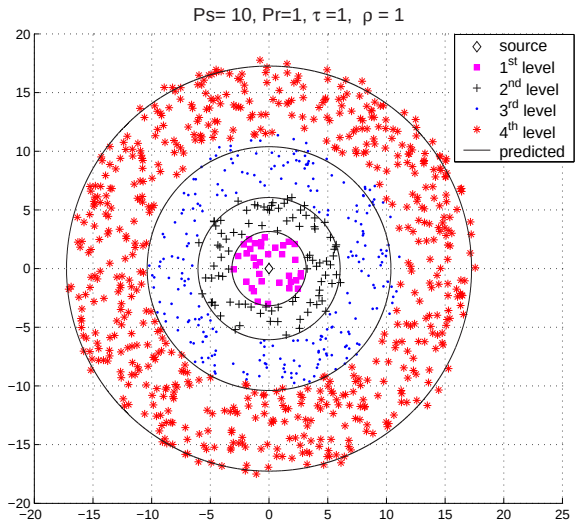

Fig. 15. Random Network Realization vs Continuum Approx.

B. Random Channel Model

1) Non-orthogonal transmission: We simulated the random network, and obtained the empirical density of nodes at level-k with respect to the distance to the source. This, together with the continuum result $P_k(r)$ is shown for $\tau = 1$ and $\tau = 3$ in Figs. 16 and 17 respectively. The parameters are $P_s = 5$, $\bar{P}_r = 1$, $\rho = 30$, $M = 3$, $\Sigma = \frac{\mathbf{I}}{M}$. The distribution of relay times, t_l's are assumed Uniform$\{-1, 0, 1\}$. The fading coefficients α_l's are $\mathcal{N}_c(0, 1)$. In Fig. 16, the moving wave behavior is observed, as expected from the continuum model. On the other hand in Fig. 17, the transmissions die out. Notice that the continuum result $P_k(r)$ is the smooth curve.

2) Orthogonal Transmission: Similar to the non-orthogonal scenario, we simulated the random network, and obtained the empirical density of nodes at level-k under wideband orthogonal channels. Figures 18 and 19 show the behavior respectively in low and high threshold regimes. The parameters are $P_s = 3.5$, $\bar{P}_r = 1$, $M = 1$, $\Sigma = \mathbf{I}$. The fading coefficients α_l are $\mathcal{N}_c(0, 1)$. The curves are obtained for a high node density, $\rho = 100$. Fig. 18 shows the behavior for the low threshold regime. The effect of fading is transient, it *i.e.*, dies out after the first two levels, and the third level curve $P_3(r)$ almost takes binary values. In Fig. 19, the behavior for the high threshold region is shown. As expected, the effect of initial fading is transient and, the curves get narrower as the signal propagates. Notice that the continuum result $P_k(r)$ is the smooth curve.

VI. Conclusion

In this paper, we analyzed the behavior of a wireless network with cooperative broadcasting. The analysis is based on the idea of continuum approximation, which models networks with high node density. We believe that the techniques used in this paper can be useful in the analysis of other cooperative protocols. The accuracy of the continuum approximation is verified by simulations. The interesting conclusion drawn from the analysis is that there exists a phase transition in the

propagation of the message, which is a function of the node powers and the reception threshold.

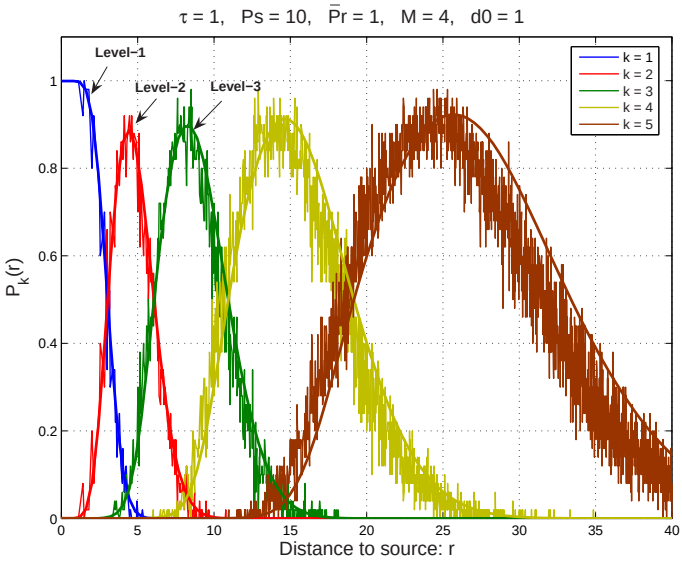

Fig. 16. Transmissions continue

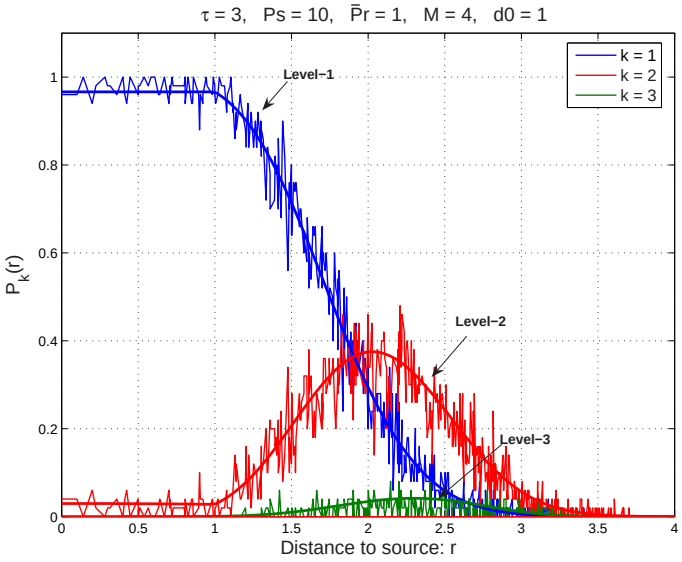

Fig. 17. Transmissions die out

Appendix I
Convergence of Random Network to the Continuum

Let

$$\mathbb{D}(r, r') = \{(x, y) : r \leq x^2 + y^2 \leq r'\}$$

denote the ring with inner and outer radii r and r', respectively. The following theorem, which is an instance of the so-called *uniform law of large numbers*, is a major step in the proof of Theorem 1.

Theorem 7: Let $X_i \in \mathbb{R}^2$, $i = 1, 2, \cdots, N$ denote the location of nodes in the random network. Let $\mathcal{U}$ denote the

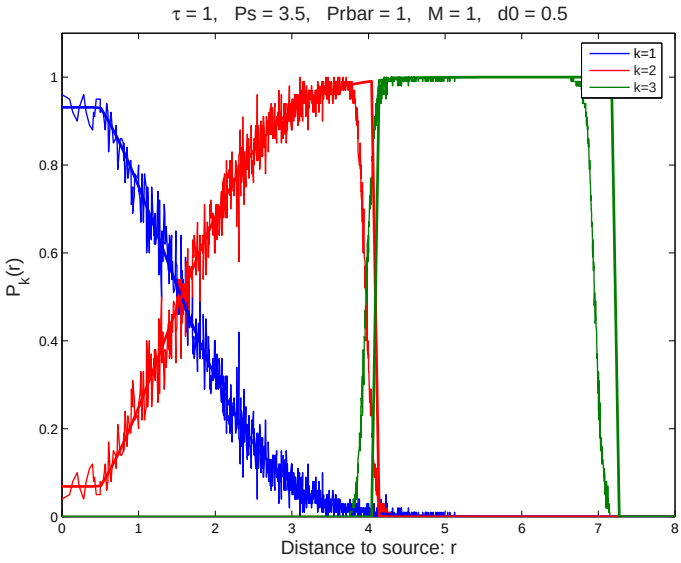

Fig. 18. Transmissions continue

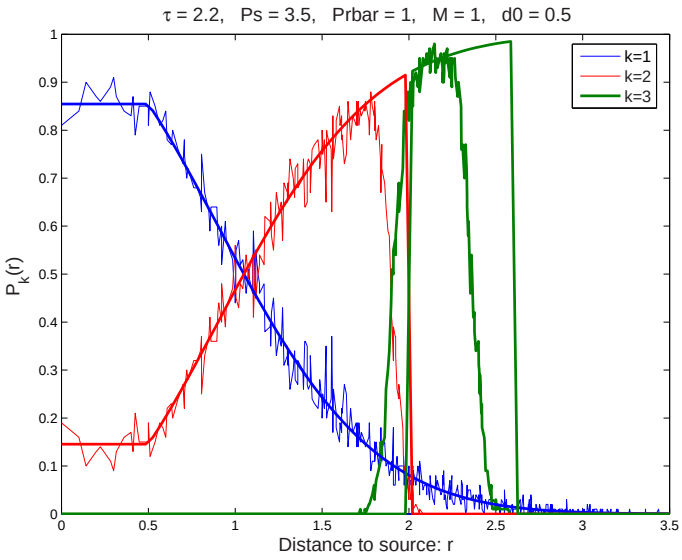

Fig. 19. Transmissions die out

set of all axis-aligned rectangles in $\mathbb{R}^2$.[3] There exists a real-valued sequence $\epsilon_N \to 0$ such that as $N \to \infty$,

$$\Pr\left\{\sup_{U\in\mathcal{U}}\left|\frac{1}{N}\sum_{i=1}^{N}1(X_i \in U) - \frac{\text{Area}(U \cap \mathbb{S})}{\pi R^2}\right| \leq \epsilon_N\right\} \to 1. \quad (57)$$

Proof: It is a classic example in statistical learning theory that the VC dimension of the set $\mathcal{U}$ is finite (*e.g.*, see [36]). The result (57) directly follows from the Vapnik-Chervonenkis Theorem [37]. ■

For convenience, we let

$$F_N(\epsilon_N) = \left\{\sup_{U\in\mathcal{U}}\left|\frac{1}{N}\sum_{i=1}^{N}1(X_i \in U) - \frac{\text{Area}(U \cap \mathbb{S})}{\pi R^2}\right| \leq \epsilon_N\right\}$$

[3]Each set in $\mathcal{U}$ is of the form $\{(x, y) : x_{\min} \leq x < x_{\max},\ y_{\min} \leq y < y_{\max}\}$ for some $x_{\min}, x_{\max}, y_{\min}, y_{\max} \in \mathbb{R}$.

denote the event in (57). Let a_N be an integer-valued sequence such that

$$\begin{aligned} &\text{(i)} && a_N \text{ is monotone increasing.} \\ &\text{(ii)} && a_N \to \infty \text{ as } N \to \infty. \\ &\text{(iii)} && \epsilon_N / 2^{-a_N} \to 0 \text{ as } N \to \infty. \end{aligned} \tag{58}$$

Later, the motivation for such a sequence will be clearer. We would like to partition the Euclidian plane with squares of size $2^{-a_N} \times 2^{-a_N}$. In particular, we use the notation

$$\begin{aligned} \Delta_N(i,j) \triangleq \{(x,y) : i2^{-a_N} \le x < (i+1)2^{-a_N}, \\ j2^{-a_N} \le y < (j+1)2^{-a_N}\} \end{aligned} \tag{59}$$

to denote these squares. The following lemma gives upper and lower bounds on the number of nodes in each $\Delta_N(i,j)$, given that $F_N(\epsilon_N)$ happened.

Lemma 3: If the event $F_N(\epsilon_N)$ happened, then for every $\Delta_N(i,j) \subset \mathbb{S}$

$$\begin{aligned} \rho 2^{-a_N+1}\left[1 - (\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right] \le |\mathcal{S} \cap \Delta_N(i,j)| \\ \le \rho 2^{-a_N+1}\left[1 + (\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right]. \end{aligned} \tag{60}$$

The upper bound is valid even for $\Delta_N(i,j)$ not in $\mathbb{S}$.

Proof: Substituting $U = \Delta_N(i,j)$ into the definition of $F_N(\epsilon_N)$ gives

$$\left| \frac{1}{N}|\mathcal{S} \cap \Delta_N(i,j)| - \frac{\text{Area}(\Delta_N(i,j) \cap \mathbb{S})}{\pi R^2} \right| \le \epsilon_N.$$

Eqn (60) is obtained easily after some algebraic manipulation. ∎

Let $E_{k,N}(\delta)$ denote the event that first k level nodes are contained in the disk $\mathbb{D}(0, r_k + \delta)$ and the $\mathcal{S}_i,\ i = 1, \cdots, k$ contains all nodes in $\mathbb{D}(r_{i-1} + \delta, r_i - \delta)$, *i.e.,*

$$\begin{aligned} E_{k,N}(\delta) = \{\cup_{i=1}^{k} \mathcal{S}_i \subset \mathbb{D}(0, r_k + \delta), \\ \mathcal{S} \cap \mathbb{D}(r_{i-1} + \delta, r_i - \delta) \subset \mathcal{S}_i, \quad \forall i \in \{1, \cdots, k\}\}. \end{aligned}$$

Theorem 8: For all $k \in \{1, \cdots, k_{\max}\}$, there exists a non-increasing sequence $\delta_N \to 0$ such that $\Pr\{E_{k,N}(\delta_N)\} \to 1$ as $N \to \infty$.

Proof: We will prove the theorem by induction, *i.e.,* we will show that

i) There exists $\delta_N \to 0$ such that $\Pr\{E_{1,N}(\delta_N)\} \to 1$.
ii) If $\Pr\{E_{k,N}(\delta_N)\} \to 1$ for some non-increasing $\delta_N \to 0$, $k \in \{1, \cdots, k_{\max} - 1\}$, then there exists non-increasing $\delta'_N \to 0$ such that $\Pr\{E_{k+1,N}(\delta'_N)\} \to 1$.

Part (i) immediately follows from the definitions of $\mathcal{S}_1$, $\mathbb{S}_1$ (choose $\delta_N = 0,\ \forall N$). To prove (ii), assume that

$$\Pr\{E_{k,N}(\delta_N)\} \to 1 \tag{61}$$

for some non-increasing $\delta_N \to 0,\ k \in \{1, \cdots, k_{\max} - 1\}$. For any sequence δ'_N, the inequality

$$\begin{aligned} \Pr\{E_{k+1,N}(\delta'_N)\} \ge \Pr\{E_{k,N}(\delta_N),\ F_N(\epsilon_N)\} \cdot \\ \Pr\{E_{k+1,N}(\delta'_N) \mid E_{k,N}(\delta_N),\ F_N(\epsilon_N)\} \end{aligned}$$

holds. Because of Theorem 7 and (61),

$$\Pr\{E_{k,N}(\delta_N),\ F_N(\epsilon_N)\} \to 1.$$

Thus, we are done if we can show the existence of $\delta'_N \to 0$ such that

$$\Pr\{E_{k+1,N}(\delta'_N) \mid E_{k,N}(\delta_N),\ F_N(\epsilon_N)\} \to 1. \tag{62}$$

In order to show (62) we will upper and lower bound the summation in (3) with relevant integrals. Our lower bound depends on a function $\underline{\ell}_N(x,y,x',y')$, which is constructed next. Consider the disc $\mathbb{D}(r_{k-1} + \delta_N, r_k - \delta_N)$ and a point (x', y') in $\mathbb{R}^2$.

- If a square $\Delta_N(i,j)$ is contained inside $\mathbb{D}(r_{k-1}+\delta_N, r_k - \delta_N)$, let the function $\underline{\ell}_N(x,y,x',y')$ take the value $\min_{(x'',y'') \in \Delta_N(i,j)} \ell(x'' - x', y'' - y')$ for all $(x,y) \in \Delta_N(i,j)$.
- If $(x,y) \in \Delta_N(i,j)$, but $\Delta_N(i,j)$ is not fully contained in $\mathbb{D}(r_{k-1} + \delta_N, r_k - \delta_N)$, let $\underline{\ell}_N(x,y,x',y')$ be zero.

More succinctly, the function $\underline{\ell}_N(x,y,x',y')$ is the *lower envelope* of $\ell(x - x', y - y')$ inside the region $\mathbb{D}(r_{k-1} + \delta_N, r_k - \delta_N)$, and is zero otherwise. The cross-sectional view of $\underline{\ell}_N(x,y,x',y')$ along the x-axis looks like Fig. 20.

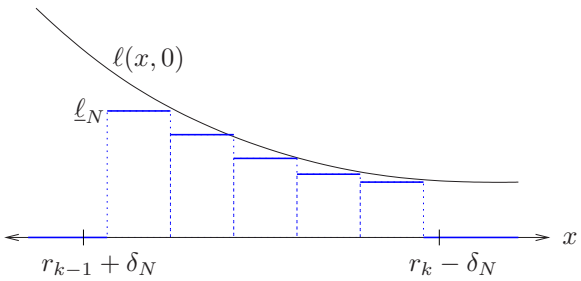

Fig. 20. Illustration of $\underline{\ell}_N(x,0,0,0)$ as a function of x.

An important property of this function is that $\underline{\ell}_N(x,y,x',y')$ is non-decreasing with N (for this, we need δ_N to be non-decreasing). Moreover, for all $(x,y), (x',y')$,

$$\underline{\ell}_N(x,y,x',y') \to \ell(x - x', y - y') \quad \text{as } N \to \infty, \tag{63}$$

since ℓ is continuous. Now, we want to examine why these properties are important. Under the assumption that the event $\{E_{k,N}(\delta_N),\ F_N(\epsilon_N)\}$ happened, for any $x, y \in \mathbb{R}$, we have

$$\begin{aligned} \sum_{(x',y') \in \mathcal{S}_{k-1}} P_r \ell(x - x', y - y') \ge \\ \sum_{(x',y') \in \mathcal{S} \cap \mathbb{D}(r_{k-1}+\delta_N, r_k - \delta_N)} P_r \ell(x - x', y - y') \ge \\ \sum_{i,j=-\infty}^{\infty} P_r |\Delta_N(i,j) \cap \mathcal{S} \cap \mathbb{D}(r_{k-1} + \delta_N, r_k - \delta_N)| \cdot \\ \min_{(x',y') \in \Delta_N(i,j)} \ell(x - x', y - y'). \end{aligned}$$

Recall that $P_r = \bar{P}_r / \rho$, and observe that the final expression is lower bounded by

$$\begin{aligned} \ge \left[1 - (\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right] \cdot \\ \int_{\mathbb{D}(r_{k-1}, r_k)} \bar{P}_r \underline{\ell}_N(x,y,x',y') dx' dy' \triangleq \underline{\mathrm{I}}_N(x,y). \end{aligned}$$

The final lower bound has several important properties. First, the term $\left[1-(\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right]$ converges to 1 due to Eqn. (58). By the Monotone Convergence Theorem, for all (x,y)

$$\int_{\mathbb{D}(r_{k-1},r_k)} \bar{P}_r \underline{\ell}_N(x,y,x',y')dx'dy' \rightarrow \int_{\mathbb{S}_k} \bar{P}_r \ell(x-x',y-y')dx'dy', \quad \text{as } N\rightarrow\infty. \quad (64)$$

Furthermore, by using the properties of the function ℓ, it can be shown that the convergence in (64) is *uniform* over $(x,y) \in \mathbb{D}(r_k+\gamma, r_{k+1})$ for all $\gamma > 0$. Also, since ℓ is non-increasing, the function $\underline{\mathrm{I}}_N(x,y)$ is non-increasing along the outward radial direction for all $(x,y) \in \mathbb{D}(r_k,\infty)$. All these properties lead to the fact that there exists non-increasing $\delta'_N \rightarrow 0$ such that

$$\underline{\mathrm{I}}_N(x,y) > \tau \quad \text{for all} \quad (x,y) \in \mathbb{D}(r_k, r_{k+1}-\delta'_N). \quad (65)$$

If the event $E_{k,N}(\delta_N)$ happened, then $\cup_{i=1}^k \mathcal{S}_i \subset \mathbb{D}(0, r_k+\delta_N)$, which means that the nodes in $\mathbb{D}(r_k+\delta_N,\infty)$ join level $k+1$ if their received power exceeds τ. Therefore,

$$\Pr\{\mathcal{S}\cap\mathbb{D}(r_k+\delta''_N, r_{k+1}-\delta''_N) \subset \mathcal{S}_{k+1} \mid E_{k,N}(\delta_N),\ F_N(\epsilon_N)\} \rightarrow 1, \quad (66)$$

as $N\rightarrow\infty$, where $\delta''_N \triangleq \max\{\delta_N, \delta'_N\}$. ■

We are done if we can show the existence of a sequence $\delta'''_N \rightarrow 0$ such that as $N\rightarrow\infty$,

$$\Pr\{\cup_{i=1}^{k+1}\mathcal{S}_i \subset \mathbb{D}(0, r_{k+1}+\delta'''_N) \mid E_{k,N}(\delta_N),\ F_N(\epsilon_N)\} \rightarrow 1. \quad (67)$$

(notice that (66) and (67) imply (62)). Since the rest of the proof is quite similar, we will only give its outline. Under the condition that the event $\{E_{k,N}(\delta_N),\ F_N(\epsilon_N)\}$ happened, if we can show that there exists $\delta'''_N \rightarrow 0$ such that

$$\underline{\mathrm{I}}_N(x,y) < \tau \quad \text{for all} \quad (x,y) \in \mathbb{D}(r_{k+1}+\delta'''_N, \infty), \quad (68)$$

we are done. First, notice that is if $E_{k,N}(\delta_N)$ happened, then

$$\mathcal{S}_k \subset [\cup_{i=1}^{k-1}\mathbb{D}(r_i-\delta_N, r_i+\delta_N) \cup \mathbb{D}(r_{k-1}+\delta_N, r_k+\delta_N)] \triangleq \tilde{\mathbb{S}}_{k,N}.$$

This is because all $\mathcal{S}_i,\ i=1,\cdots,k$ are in $\mathbb{D}(0, r_k+\delta_N)$, yet the nodes in $\mathbb{D}(r_{i-1}+\delta_N, r_i-\delta_N)$ belong to level i, for $i=1,\cdots,k-1$. Consequently, the following upper bound holds:

$$\sum_{(x',y')\in\mathcal{S}_{k-1}} P_r\ell(x-x',y-y') \leq \sum_{(x',y')\in\mathcal{S}\cap\tilde{\mathbb{S}}_{k,N}} P_r\ell(x-x',y-y').$$

$$\leq \left[1+(\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right]\int_{\mathbb{R}^2} \bar{P}_r\bar{\ell}_N(x,y,x',y')dx'dy' \triangleq \bar{I}_N(x,y), \quad (69)$$

where the function $\bar{\ell}_N(x,y,x',y')$ is the *upper-envelope* of ℓ over the region $\tilde{\mathbb{S}}_{k,N}$, *i.e.,*

$$\bar{\ell}_N(x,y,x',y') = \begin{cases} \max_{(x'',y'')\in\Delta_N(i,j)} \ell(x''-x', y''-y'), \\ \quad \text{if } (x,y)\in\Delta_N(i,j) \text{ and } \tilde{\mathbb{S}}_{k,N}\cap\Delta_N(i,j)\neq\phi, \\ 0, \qquad \text{otherwise.} \end{cases}$$

Monotone Convergence Theorem implies that $\bar{\ell}_N(x,y,x',y') \rightarrow \ell(x-x',y-y')$ for all $(x,y),\ (x',y')$. Using arguments similar to above, it is seen that (69) implies (68). The theorem follows. ■

Next, we will prove Theorem 1 using Theorem 7 and Theorem 8. In the following, we will give a proof of Theorem 1 under the assumption that $k_{\max} < \infty$. The same proof can be used for $k_{\max}=\infty$ after minor modifications.

Proof: (Theorem 1) Let $\mathbb{V} \triangleq \mathbb{S}\cap\mathbb{U}$. Notice that

$$\frac{|\mathbb{U}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{U}\cap\mathbb{S}_k)} = \frac{|\mathbb{V}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}.$$

In order to prove Theorem 1, we need to show that

$$\Pr\left\{\left|\frac{|\mathbb{V}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}-1\right| > s\right\} \rightarrow 0 \quad \text{as } N\rightarrow\infty, \quad \forall s>0,$$

However, notice that this probability is upper bounded by

$$\leq \Pr\{[F_N(\epsilon_N),\ E_{N,k_{\max}}(\delta_N)]^C\} + \Pr\left\{\left|\frac{|\mathbb{V}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}-1\right| > s | F_N(\epsilon_N), E_{N,k_{\max}}(\delta_N)\right\} \quad (70)$$

Because of Theorem 7 and Theorem 8, the term on the left hand side converges to zero. Therefore, to prove Theorem 1, it suffices to show that

$$\Pr\left\{\left|\frac{|\mathbb{V}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}-1\right| > s \mid F_N(\epsilon_N), E_{N,k_{\max}}(\delta_N)\right\} \rightarrow 0 \quad \text{as } N\rightarrow\infty, \quad \forall s>0. \quad (71)$$

Under the assumption that $E_{N,k_{\max}}(\delta_N)$ happened, we have the relation

$$\mathcal{S}\cap\mathbb{D}(r_{k-1}+\delta_N, r_k-\delta_N) \subset \mathcal{S}_k \subset \tilde{\mathbb{S}}_{k_{\max},N}, \quad (72)$$

where

$$\tilde{\mathbb{S}}_{k,N} \triangleq [\cup_{i=1}^{k_{\max}}\mathbb{D}(r_i-\delta_N, r_i+\delta_N) \cup \mathbb{D}(r_{k-1}+\delta_N, r_k+\delta_N)],$$

as before. To see (71), we shall use the "partition into rectangles" trick, used above. Consider the partitioning $\Delta_N(i,j)$ introduced above. Define the sets

$$\bar{\mathbb{V}}_{k,N} = \bigcup_{i,j}\{\Delta_N(i,j) : \Delta_N(i,j)\cap[\mathbb{V}\cap\tilde{\mathbb{S}}_{k,N}]\neq\phi\},$$

$$\underline{\mathbb{V}}_{k,N} = \bigcup_{i,j}\{\Delta_N(i,j) : \Delta_N(i,j) \subset \mathbb{V}\cap\mathbb{D}(r_{k-1}+\delta_N, r_k-\delta_N)\},$$

as outer and inner approximations of $\mathbb{V}$ with rectangles. Since $\mathbb{U}$ is a disc shaped region[4]

$$\lim_{N\rightarrow\infty}\text{Area}(\underline{\mathbb{V}}_{k,N}) = \lim_{N\rightarrow\infty}\text{Area}(\bar{\mathbb{V}}_{k,N}) = \text{Area}(\mathbb{V}\cap\mathbb{S}_k). \quad (73)$$

Now, we examine why (73) is useful.

$$\frac{|\mathbb{V}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)} \geq \frac{|\underline{\mathbb{V}}_{k,N}\cap\mathcal{S}|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)} \geq \left[\frac{1}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}\right]\frac{\text{Area}(\underline{\mathbb{V}}_{k,N})}{2^{-a_N+1}}(\rho 2^{-a_N+1})\cdot\left[1-(\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right].$$

[4]The property (73), which is fairly obvious for disc shaped $\mathbb{U}$, is also true for all Jordan measurable $\mathbb{U}$. For conciseness, we will not prove it in this paper, though.

The lower bound converges to 1. Similarly,

$$\begin{aligned}\frac{|\mathbb{V}\cap\mathcal{S}_k|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)} &\leq \frac{|\bar{\mathbb{V}}_{k,N}\cap\mathcal{S}|}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}\\ &\leq \left[\frac{1}{\rho\text{Area}(\mathbb{V}\cap\mathbb{S}_k)}\right]\frac{\text{Area}(\bar{\mathbb{V}}_{k,N})}{2^{-a_N+1}}(\rho 2^{-a_N+1})\cdot\\ &\qquad\left[1+(\pi R^2)\frac{\epsilon_N}{2^{-a_N+1}}\right],\end{aligned}$$

which converges to 1 as well. The theorem follows. ■

APPENDIX II
PROOF OF THEOREM 3

Similar to the case of $m=1$, we can derive the difference equation for $k\geq m$ as

$$a_{k+1}-\frac{\mu}{\mu-1}a_k+\frac{1}{\mu-1}a_{k-m}=0. \tag{74}$$

The initial conditions of this system are given by the following lemma.

Lemma 4: The first initial condition is $a_1=\frac{P_s}{\tau}$. Then k'th initial condition of the difference equation (74) is the unique solution of the equation

$$g(p)\triangleq\frac{P_s}{p}+\pi\bar{P}_r\ln\frac{p}{|p-a_{k-1}|}=\tau,\quad 2\leq k\leq m. \tag{75}$$

The initial conditions $a_2,\ \cdots,\ a_m$ can be obtained by solving (75) recursively.

Proof: If we assume that the nodes have memory length of m, then for $k\leq m$, k'th level nodes consider the transmission from all levels transmitted previously including the source node. Then finding a_k simplifies to solving $g(p)=\tau$ for p. The claim is that the solution of the equation $g(p)=\tau$ exists and unique. This can be seen as follows. By construction $a_k\geq a_{k-1}$. By inspecting the derivative of $g(p)$, it is seen that $g(p)$ is a decreasing continuous function for $p\geq a_{k-1}$. Note that $g(a_{k-1})=\infty$ and $g(\infty)=0$. Hence the equation $g(p)=\tau$ has a unique solution for $p\geq a_{k-1}$. ■

The characteristic function associated with the difference equation (74) is

$$f(r)=r^{m+1}-\frac{\mu}{\mu-1}r^m+\frac{1}{\mu-1}=0. \tag{76}$$

The polynomial $f(r)$ is $(m+1)$-th order and has $(m+1)$ roots $r_i,\ i=1\ldots(m+1)$. The solution of the difference equation is determined by the roots of the polynomial $f(r)$ and the initial conditions (75). If there exists a root $|r_i|>1$, then as $n\to\infty$, $a_n\to\infty$. On the other hand, if $|r_i|\leq 1,\forall i$ then as $n\to\infty$, $a_n\to a_\infty$, where a_∞ is a finite number. Hence, we would like to count the number of roots of $f(r)$ inside the unit circle and strictly outside the unit circle. The following lemma serves this purpose.

Lemma 5: If $1<\mu<m+1$, then $f(r)$ has a root in the interval $(1,\infty)$. On the other hand, if $\mu>m+1$, then $f(r)$ has m roots inside unit circle and a root on the unit circle.

Proof: The polynomial $f(r)$ is $(m+1)$-th order and has $(m+1)$ roots $r_1,\cdots,\ r_{m+1}$. We can easily rewrite

$$f(r)=\frac{(r-1)}{\mu-1}\tilde{f}(r)$$

where

$$\tilde{f}(r)=(\mu-1)r^m-r^{m-1}-r^{m-2}\ldots-r-1.$$

Hence, $r_1=1$ is a root of $f(r)$ and the rest of the roots satisfy $\tilde{f}(r)=0$. We will use this $\tilde{f}(r)$ and Rouché's Theorem [38] in the following.

- *Case 1:* For $1<\mu<m+1$, $f(1)=\mu-1-m<0$. Also we can easily see that $\lim_{r\to\infty}\tilde{f}(r)=\infty$. Since $\tilde{f}(r)$ is a polynomial, it is continuous. Hence, $\tilde{f}(r)$ has a root in $(1,\infty)$.
- *Case 2:* If $\mu>m+1$, then $\tilde{f}(r)$ have m roots inside unit circle. We can rewrite $\tilde{f}(r)$ as $\tilde{f}(r)=h(r)+g(r)$ where

$$\begin{aligned}h(r) &= (\mu-1)r^m\\ g(r) &= -(r^{m-1}+r^{m-2}+\ldots+1)\end{aligned}$$

We show that $|h(r)|>|g(r)|$ on the unit circle $\mathcal{C}$.

$$\begin{aligned}|h(r)| &=^a |(\mu-1)r^m|=|(\mu-1)|\\ &>^b m\\ &=^c |1|+|r|+|r^2|+\ldots+|r^m|\\ &>^d |1+r+r^2+\ldots+r^m|\\ &=^e |g(r)|\end{aligned} \tag{77}$$

where (a) is by definition and also $r\in\mathcal{C}$; (b) is assumed by the lemma; (c) follows from the fact that $r\in\mathcal{C}$; (d) is from the triangle inequality; (e) is by definition. Then, by using Rouché's theorem $h(r)$ and $\tilde{f}(r)=h(r)+g(r)$ have the same number of zeros inside the unit circle. The function $h(r)$ has m multiple zeros at $r=0$, then $\tilde{f}(r)$ has m roots inside unit circle.

■

Proof of Theorem 3 i) follows from Lemmas 4 and 5. Part ii), can be seen as follows. If $m=\infty$, then a_k can be found by solving the following equation recursively

$$g(p)\triangleq\frac{P_s}{p}+\pi\bar{P}_r\ln\frac{p}{|p-a_{k-1}|}=\tau,\quad k\geq 1,$$

where $a_0=0$. The existence of unique solution to this equation is discussed previously in Lemma 4. We will use proof by contraction in the following. We assume that $a_k\to B$, such that $0<B<\infty$. Note that as $k\to\infty$, $a_{k-1}\to B$. Then

$$\lim_{k\to\infty}g(a_k)=g(B)=\frac{P_s}{B}+\pi\bar{P}_r\ln\frac{B}{|B-B|}=\infty$$

which is a contraction since we know that $g(B)=\tau$ is finite.

APPENDIX III
PROOF OF THEOREM 6

Assume that $\iint_{\mathbb{S}}l(x-u,y-v)dudv=C<\infty$. We will first upper bound the $\sigma_k^2(x,y)$. Using (48), we get

$$\begin{aligned}\sigma_k^2(x,y) &\leq \bar{P}_r\left(\sup_{(u,v)\in\mathbb{S}}P_k(u,v)\right)\iint_{\mathbb{S}}l(x-u,y-v)dudv\\ &\leq C\bar{P}_r\sup_{(u,v)\in\mathbb{S}}P_k(u,v)\end{aligned} \tag{78}$$

Also, by using (47) we can derive the following upper bound on $P_{k+1}(x,y)$ in terms of $\sigma_k^2(x,y)$. For simplicity, we assume

that Σ has distinct eigenvalues (the following proof can be adapted easily for the other case). Apply (53) to obtain

$$\begin{aligned} P_{k+1}(x,y) &\leq \sum_{i=1}^{M} A_{i1} e^{-\tau/\sigma_k^2(x,y)\lambda_i} \\ &\leq \sum_{i=1}^{M} A_{i1} e^{-\tau/\lambda_i \sup_{(x,y)\in\mathbb{S}} \sigma_k^2(x,y)}, \\ &\leq \gamma e^{-\tau/\lambda_m \sup_{(x,y)\in\mathbb{S}} \sigma_k^2(x,y)} \end{aligned} \tag{79}$$

where the second inequality follows from the fact that $e^{-\tau/x}$ is an increasing function of x and in the third inequality $\gamma = M \max_i A_{i1}$ and $\lambda_m = \max_i \lambda_i$. Note that $\gamma > 0$, since it is a bound to a probability.

Let $M_k = \sup_{(x,y)\in\mathbb{S}} P_k(x,y)$. Combining (78) and (79), we have the relation

$$M_{k+1} \leq \gamma e^{-\beta/M_k}, \qquad k = 1, 2, \cdots, \tag{80}$$

where $\beta = \frac{\tau}{\lambda_m C P_r}$. The initial condition is that

$$\begin{aligned} M_1 &= \sup_{(x,y)\in\mathbb{S}} P_1(x,y) \\ &= \sup_{(x,y)\in\mathbb{S}} \sum_{i=1}^{M} A_{i1} e^{-\tau/\sigma_0^2(x,y)\lambda_i} \\ &= \sum_{i=1}^{M} A_{i1} e^{-\tau/\sup_{(x,y)\in\mathbb{S}} \sigma_0^2(x,y)\lambda_i} \\ &= \sum_{i=1}^{M} A_{i1} e^{-\tau/P_s\lambda_i \sup_{(x,y)\in\mathbb{S}} \ell(x,y)}, \end{aligned}$$

where we only used the definitions of $P_1(x,y), \sigma_0^2(x,y)$.

Next, we will show that any sequence $M_1, M_2, \cdots$ satisfying (80) converges to zero. To this end, consider a sequence $L_1, L_2, \cdots$ satisfying $L_{k+1} = \gamma e^{-\beta/L_k}$ with the initial condition $L_1 = M_1$. We will first argue that $M_k \leq L_k, \quad \forall k$. This can be easily proved by induction:
i) $M_1 \leq L_1$ by the definition of L_1.
ii) Assume that $M_k \leq L_k$ for some k. Then,

$$M_{k+1} \leq \gamma e^{-\beta/M_k} \leq \gamma e^{-\beta/L_k} = L_{k+1},$$

because the function $e^{-\beta/x}$ is increasing in x.

Second, we will observe that $L_k \to 0$ as $k \to \infty$ if

$$1 > \frac{\gamma}{\beta e} \quad \Leftrightarrow \quad \tau > \left(\gamma \lambda_m C e^{-1}\right) \bar{P}_r.$$

Observe that

$$\frac{L_{k+1}}{L_k} \leq \gamma \sup_{x\geq 0} \frac{e^{-\beta/x}}{x} = \frac{\gamma}{\beta e}; \tag{81}$$

by differentiation it can be seen that the function $\frac{e^{-\beta/x}}{x}$ is maximized at $x = \beta$. Eqn. (81) shows that $L_{k+1} \leq \frac{L_1}{(\beta e/\gamma)^k} \to 0$. Let $\tau_c = \gamma \lambda_m C e^{-1}$, then the theorem follows.

REFERENCES

[1] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "Energy-efficient broadcast and multicast trees in wireless networks," *Mob. Netw. Appl.*, vol. 7, no. 6, pp. 481–492, 2002.

[2] B. Williams and T. Camp, "Comparison of broadcasting techniques for mobile ad hoc networks," in *ACM Proc. on MOBIHOC*, 2002.

[3] J. N. Laneman and G. W. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, Oct. 2003.

[4] S. Shakkottai, R. Srikant, and N. Shroff, "Unreliable sensor grids: Coverage, connectivity and diameter," *AdHoc Networks*, 2004.

[5] P. Jacquet, "Geometry of information propagation in massively dense ad hoc networks," in *ACM MobiHoc'04*, May 2004, pp. 157–162.

[6] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Trans. Commun.*, Oct. 2004.

[7] I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, Aug. 2004.

[8] A. Scaglione and Y.-W.Hong, "Opportunistic large arrays:cooperative transmission in wireless multihop adhoc networks to reach far distances," *IEEE Trans. Signal Processing*, no. 8, Aug. 2003.

[9] P. Viswanath, D. Tse, , and R. Laroia, "Opportunistic beamforming using dump antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, June 2002.

[10] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," in *Proc. of IEEE Inter. Symp. on Inform. Theory (ISIT)*, 1998, p. 156.

[11] ——, "User cooperation - part 1: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, Nov. 2003.

[12] ——, "User cooperation - part 2: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, Nov. 2003.

[13] J. N. Laneman, D. Tse, , and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, Dec. 2004.

[14] S. Barbarossa and G. Scutari, "Distributed space-time coding strategies for wideband multihop networks: regenerative vs. non-regenerative relays," in *Proc. of IEEE Inter. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, 2004.

[15] R. U. Nabar, F. W. Kneubuhler, and H. Bölcskei, "Performance limits of amplify-and-forward based fading relay channel," in *Proc. of IEEE Inter. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, 2004.

[16] B. Sirkeci-Mergen and A. Scaglione, "Signal acquisition for cooperative transmissions in multi-hop ad-hoc networks," in *Proc. of IEEE Inter. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, 2004.

[17] ——, "Coverage analysis of cooperative broadcast in wireless networks," in *Proc. of IEEE Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, July 2004.

[18] ——, "Message propagation in a cooperative network with asynchronous receptions," in *Proc. of IEEE Inter. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, 2005.

[19] ——, "A continuum approach to dense wireless networks with cooperation," in *Proc. of Annual Joint Conf. of the IEEE Computer and Commun. Societies (Infocom)*, 2005.

[20] P. A. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a rayleigh-fading environment," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, Sep. 2004.

[21] Y. Chang and Y. Hua, "Application of space-time linear block codes to parallel wireless relays in mobile ad hoc networks," in *Signals, Systems and Computers, 2003 The Thrity-Seventh Asilomar Conference*, vol. 1, Nov. 2003, pp. 1002 – 1006.

[22] Y. Hua, Y. Mei, and Y. Chang, "Parallel wireless mobile relays with space-time modulations," in *Statistical Signal Processing, 2003 IEEE Workshop*, 28 Sept.- 1 Oct. 2003, pp. 375 – 378.

[23] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "A theoretical characterization of the multihop wireless communications channel with diversity," in *Global Telecommunications Conference*, vol. 2, 2001, pp. 841 – 845.

[24] A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," *to appear in IEEE Trans. on Wireless Comm.*, 2004.

[25] Y.-W. Hong and A. Scaglione, "Energy-efficient broadcasting with cooperative transmission in wireless sensory ad hoc networks," in *Proc. of Allerton Conf. on Commun., Contr. and Comput. (ALLERTON)*, Oct. 2003.

[26] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. of Annual Joint Conf. of the IEEE Computer and Commun. Societies (Infocom 2002)*, 2002.

[27] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," IEEE Trans. Inform. Theory, vol. 51, no. 9, Sept. 2005.
[28] A. Khisti, U. Erez, and G. Wornell, "A capacity theorem for cooperative multicasting in large wireless networks," in *Proc. of Allerton Conf. on Commun., Contr. and Comput. (ALLERTON)*, 2004.
[29] O. Dousse and P. Thiran, "Connectivity vs. capacity in dense ad hoc networks," in *Proc. of Annual Joint Conf. of the IEEE Computer and Commun. Societies (Infocom)*, 2004.
[30] Y.-C. Cheng and T. Robertazzi, "Critical connectivity phenomena in multihop radio models," *IEEE Trans. on Communications*, vol. 37, no. 7, July 1989.
[31] P. Billingsley, *Probability and Measure*, 3rd ed. Wiley Series in Probability and Mathematical Statistics, 1995.
[32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, C. U. Press, Ed.
[33] T. S. Rappaport, *Wireless Communications Principles and Practice*, 2nd ed. Prentice Hall.
[34] O. Dousse and P. Thiran, "Connectivity vs. capacity in dense ad hoc networks," in *Proceedings of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Infocom 2004)*, 2004.
[35] A. Jovicic, P. Viswanath, and S. R. Kulkarni, "Upper bounds to transport capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2555–2565, 2004.
[36] V. N. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, Ed., 1998.
[37] V. N. Vapnik and A. Y. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory Probab. Appl.*, vol. 16, pp. 264–280, 1971.
[38] J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, 6th ed. McGraw-Hill International Editions, 1996.

PLACE PHOTO HERE

Birsen Sirkeci Mergen received her B.Sc. degrees in Electrical and Electronics Engineering and Mathematics in 1998 from Middle East Technical University (METU), Ankara, Turkey. She received her M.Sc. degree in Electrical and Computer Engineering from Northeastern University, Boston, MA, in 2001. She worked as a DSP engineer at Aware Inc, Bedford, MA during 2000-2002. She is currently pursing her PhD degree in Electrical and Computer Engineering at Cornell University, Ithaca, NY. She received the Fred Ellersick Award for the best unclassified paper in MILCOM 2005 with her co-authors.

Her research interests lie mainly in the field of digital signal processing and communications. Currently, she is interested in cooperative transmission in wireless ad-hoc networks and distributed protocols.

PLACE PHOTO HERE

Anna Scaglione received the M.Sc. and the Ph.D. degrees in Electrical Engineering from the University of Rome "La Sapienza", Rome, Italy, in 1995 and 1999, respectively. She was a postdoctoral researcher at University of Minnesota, Minneapolis, MN, in 1999-2000. She is currently Assistant Professor in Electrical Engineering at Cornell University, Ithaca, NY, since 2001; prior to this she was assistant professor during the academic year 2000-2001, at the University of New Mexico, Albuquerque, NM. She received the 2000 IEEE Signal Processing Transactions Best Paper Award and the NSF Career Award in 2002. She also received the Fred Ellersick Award for the best unclassified paper in MILCOM 2005. She is an Associate Editor for the IEEE Transactions on Wireless Communications and has been Co-Guest Editor of the Communication Magazine Special Issue on Power Line Communications "Broadband is Power: Internet Access through the Power Line Networks" , May 2003. She was recently nominated as a member of the IEEE Signal Processing for Communication Technical Committee. Her research is in the broad area of signal processing for communication systems. Her current research focuses on optimal transceiver design for MIMO-broadband systems and cooperative communications systems for large scale sensor networks.

PLACE PHOTO HERE

Gökhan Mergen was born in 1978 in Ankara, Turkey. He received his B.S.degrees in Electrical and Electronics Engineering and Mathematics in 2000 from Middle East Technical University (METU), Turkey. He received his M.S. and Ph.D. degrees in Electrical and Computer Engineering from Cornell University (the latter one in Jan. 2005). His industry experience includes work at the Bell Labs, Lucent Technologies, during the summer of 2003. He is currently employed with Qualcomm, Campbell CA. His research interests lie in the areas of communications theory, information theory, data networks and statistical signal processing.