The Transport Capacity of Wireless Networks over Fading Channels

Feng Xue, *Student Member, IEEE,* Liang-Liang Xie, *Member, IEEE,* and P. R. Kumar, *Fellow, IEEE*

***Abstract*— We consider networks consisting of nodes with radios, and without any wired infrastructure, thus necessitating all communication to take place only over the shared wireless medium. The main focus of this paper is on the effect of fading in such wireless networks. We examine the attenuation regime where either the medium is absorptive, a situation which generally prevails, or the path loss exponent is greater than 3. We study the transport capacity, defined as the supremum over the set of feasible rate vectors of the distance weighted sum of rates.**

We consider two assumption sets. Under the first assumption set, which essentially requires only a mild time average type of bound on the fading process, we show that the transport capacity can grow no faster than $O(n)$, where n denotes the number of nodes, even when the channel state information (CSI) is available non-causally at both the transmitters and the receivers. This assumption includes common models of stationary ergodic channels; constant, frequency selective channels; flat, rapidly varying channels; and flat slowly varying channels.

In the second assumption set, which essentially features an independence, time average of expectation, and nonzeroness condition on the fading process, we constructively show how to achieve transport capacity of $\Omega(n)$ even when the CSI is unknown to both the transmitters and the receivers, provided that every node has an appropriately nearby node. This assumption set includes common models of i.i.d. channels; constant, flat channels; and constant, frequency selective channels. The transport capacity is achieved by nodes only communicating with neighbors, and only using point-to-point coding.

The thrust of these results is that the multi-hop strategy, towards which much protocol development activity is currently targeted, is appropriate for fading environments. The low attenuation regime is open.

***Index Terms*— Wireless networks, fading channels, capacity, transport capacity.**

I. Introduction

RECENT years have seen research as well as development efforts [1] focusing on wireless networks consisting of nodes with radios. Two examples of much topical interest are ad hoc networks [1], [2] and sensor networks [3]. Such networks have no wired backbone network, differentiating them for example from cellular systems, and all communication must take place only over the shared wireless medium. Due to the fact that all nodes hear a superposition of the attenuated signals transmitted by all other nodes, there are several ways, some quite complex, by which information originating at a source can reach its destination. Thus one would like to have an information theoretic basis for organizing such information transfer.

Current protocol development efforts [1] are aimed at realizing the following strategy. Packets are relayed from node to node until they reach their intended destination. At each hop a packet is fully decoded, thus digitally generated, and then retransmitted to the next node on its path. The decoding of a packet at a node is done by treating all interference from concurrent transmissions as useless noise. For brevity, we call this the "multi-hop" strategy. To realize this strategy requires a suite of protocols. A medium access control protocol is needed to avoid excessive interference at receivers, since transmitters in their vicinity need to be silenced. This is the goal of the IEEE 802.11 protocol in DCF mode, as well as proposals such as DBTMA [4] and SEEDEX [5]. A power control protocol [6] is needed not only to save battery life, but also to regulate the power of transmissions which need only traverse a short hop while avoiding creating unnecessary interference for other concurrent transmissions. This is the essence of the notion of spatial reuse of the spectrum. A routing protocol is needed [7], [8], [9], [10], [11], [12] to determine the path to be followed by packets from a source to its destination.

This multi-hop strategy however foregoes many possibilities for enhancing information transfer, and it is important to characterize how much has or has not been sacrificed. For example, it does not take advantage of multi-user estimation [13] which can enable a receiver to decode several concurrent transmissions. In fact, by subtracting the components corresponding to transmissions not of interest to it, a node could enhance the signal-to-noise ratio of the transmissions of interest to it. This is the successive interference subtraction strategy which has been shown to attain the capacity of the multiple access channel [14], [15]. Even when performing the simple operation of "relaying," there are alternatives such as "amplify and forward" rather than "decode and forward," which are known to be superior in some settings [16]. In fact relaying itself is not a simple problem - to date the capacity of the simple three node relay problem is unknown [17]. Actually, in the wireless world, much stranger forms of cooperation are possible. For example, just as in acoustic active noise cancellation [18], a node could help a second node by transmitting a signal which nulls out the transmission of a third node as perceived at the second node. Since the design space is so rich with complexities, it is

Manuscript received xxxx; revised xxxx. This material is based upon work partially supported by NSF under Contract Nos. NSF ANI 02-21357 and CCR-0325716, USARO under Contract Nos. DAAD19-00-1-0466 and DAAD19-01010-465, DARPA/AFOSR under Contract No. F49620-02-1-0325, AFOSR under Contract No. F49620-02-1-0217, and DARPA under Contact Nos. N00014-0-1-1-0576 and F33615-0-1-C-1905.

Feng Xue and PR Kumar are with the Department of Electrical and Computer Engineering, and Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA. (e-mail: {fengxue, prkumar}@uiuc.edu).

Liang-Liang Xie is with the Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China. (e-mail: xie@iss.ac.cn).

necessary to have a theoretical basis which allows us to choose from among all these possibilities. For example, we would like to know how much we lose in capacity if we sacrifice multi-user estimation, or, say, if we use point-to-point coding rather than network coding.

This problem was addressed in [19]. The usual information theoretic model was enriched by taking into account the distances between nodes located on a plane with a minimum separation distance between nodes. Each node is power limited, and a performance measure C_T, called the transport capacity, was studied, which is the supremum of the distance weighted sum of rates taken over all feasible rate vectors. The wireless medium itself was very simplistically modeled. An attenuation of the form $e^{-\gamma\rho}/\rho^{\delta}$ was presumed, as a signal traveled a distance ρ, where $\gamma \geq 0$ is the absorption constant, and δ is the path loss exponent.

However, a very important issue in practical wireless networks, and unmodeled above, is the presence of multi-path fading [20]. In a wireless network, due to the physical environment, the electromagnetic waves travel to receivers along a multitude of paths, encountering delays and suffering gains which vary with time. Depending on the frequency bandwidth used, and how fast the environment changes, the fading can be divided into four cases. If the bandwidth W of the signal is much smaller than the channel coherence bandwidth B_{coh}, i.e., $W \ll B_{coh}$, then the channel is frequency non-selective or flat fading. This means that the channel only has a multiplicative effect on the signal. If on the other hand $W \geq B_{coh}$, the receiver will get several resolvable signal components, and such a channel is called frequency selective. We can also characterize a fading channel by comparing the time duration T_s of a signal symbol with the channel coherence time T_{coh}. A channel is called slow fading if $T_s \ll T_{coh}$, or fast fading if $T_s \geq T_{coh}$. Combining these two factors, we basically have four types of fading channels: *flat slow, flat fast, frequency selective slow, and frequency selective fast fading channels* (see [20], [21], [22]). In fact, the point-to-point fading channel has been an active research field for decades, and is still the subject of much research effort [21].

In order to understand the role of multi-path fading in wireless networks, we address the following question in this paper: *In the information-theoretic sense, what is the transport capacity of wireless networks when the transmissions encounter multi-path fading, and how should information transfer be organized in such a fading environment?* We study all the four fading cases mentioned above. Each node is subject to a common power constraint, and the signals sent out encounter both power-loss due to distance as well as fading, before reaching their destination nodes. Our main contribution consists of the following two results:

a) An upper bound.: If either the path loss exponent is greater than three or the absorption constant is positive, then there is a constant $c_1 > 0$ such that the transport capacity of wireless networks with n nodes, mutually separated by a minimum positive distance ρ_{min}, is upper bounded by $c_1 n$, where n is the number of nodes in the network. This is true even if the nodes have perfect non-causal channel state information (CSI) of all the fading channels.

This result implies that though techniques such as diversity (multiuser, space, time, etc., [23]) may increase the throughput and reliability, they cannot change the order of the transport capacity [19]. We do not consider the type of diversity that multiple antennas at a node can provide, and thus it is of interest to extend our results to wireless networks where nodes have multiple antennas.

In the course of the proof of this result, we extend a useful max-flow min-cut bound to a time-varying fading environment. We also show that there is a simple and interesting connection between the flow across cut-sets and the transport capacity[1]. In fact, this result makes the transport capacity an even more natural quantity to study in wireless networks. Both the above results may be of independent interest in their own right.

To obtain the upper bound, the channel state information is allowed to be perfectly known in advance to both senders and receivers, a best case scenario. The following result addresses the opposite situation, a sort of worst case where the fading is independent from time to time. Assuming no CSI at all, we exhibit a feasible lower bound of the same order, $\Omega(n)$, for all networks where every node has a nearby node within a fixed multiple of the above minimum positive distance, when $\gamma > 0$ or $\delta > 3$. Thus, these two results together delineate the effect of fading, and show that the fundamental scaling law [19] remains the same even under fading environments.

b) A feasible lower bound.: Assume each node faces a fading process independent from time to time. If within a distance $\zeta\rho_{min}$ of every node there is another node, then a transport capacity of at least $c_2 n$ is achievable for a positive c_2. The scheduling, coding and decoding do not require any CSI at any node, and in fact require very little statistical knowledge of the fading process. In the example we construct, the signals are "peaky" and only a small fraction of the time is used in transmission – similar to the signaling strategy used in [25]. The transmissions are coordinated carefully and random phases[2] are introduced in signaling in order to avoid strong interference coming from nearby transmissions. In the scenario studied, communications are only between neighbors, and coding is only point-to-point. The thrust of this result is that the multi-hop strategy is a reasonable one for organizing the flow of information when attenuation is high and load can be balanced across the network.

Above we have only addressed the high attenuation regime, and have not said anything about the case where there is no absorption and attenuation is small, i.e., $\gamma = 0$ and $\delta \leq 3$. The reason for our inability is that in such scenarios one can exploit coherence to obtain capabilities not feasible in the relatively high attenuation regime; for example, unbounded transport capacity is feasible even when the sum of the transmission powers of the nodes is fixed; see [19]. However, in a fast fading environment, one cannot employ a strategy capitalizing on coherence. Feasibility results are constructive, and we have yet to find a scheme which works in fading environments. This

[1] A similar idea is mentioned in [24].

[2] Random fading is also intentionally induced in [26] to facilitate communication, but the purpose in [26] is to increase the channel fluctuation to improve the multiuser diversity.

remains a significant open problem. The recent breakthrough on the capacity of some relay channels with fading in [27] may possibly prove valuable in this regard.

It should be noted that fundamental studies based on first principle photonic collision models have shown that the case $\gamma > 0$ generally prevails, unless one is in free space; see [28]. Thus the results presented here, which include this case, have bearing for practical environments, though of course γ could possibly be small.

The remainder of the paper is organized as follows. In Sections II and III we formulate the model and state the main results. In Sections IV to VI we prove the upper bound, while in Sections VII to XI we prove the achievability results. We conclude with some remarks in Section XII.

II. The model

We consider a wireless communication network $\mathcal{G}_n$ consisting of a group of n nodes, $N := \{1, 2, \ldots, n\}$, located on the plane. The base-band model for the communications among them is described by the following equation:

$$Y_j(t) = \sum_{i \neq j} \frac{\beta e^{-\gamma \rho_{ij}}}{\rho_{ij}^{\delta}} \left(\sum_{l=0}^{\infty} H_{ijl}(t) \cdot X_i(t - \tau_{ij} - l) \right) + Z_j(t), \quad t \geq 1, j \in N. \tag{1}$$

We will consider two alternative assumption sets (A1) and (A2).

Assumption set (A1):

(A1.i) δ, γ, β, ρ_{ij} *and* τ_{ij} *are* deterministic real *variables known to all the nodes:*

- δ *and* γ *are the path-loss exponent and absorption constant of the attenuation, respectively. We assume that either* $\gamma = 0$ *with* $\delta > 3$, *or* $\gamma > 0$ *with* $\delta \geq 0$. $\beta > 0$ *is just a constant gain.*
- ρ_{ij} *is the distance between node* i *and* j. *We assume that there is a minimum distance between nodes, i.e.,* $\rho_{ij} \geq \rho_{min} > 0$.
- $\tau_{ij} := \lfloor \frac{\rho_{ij}}{\rho_0} \rfloor$ *is the propagation delay for signals from* i *to* j, *where* ρ_0 *is the distance that a signal travels in one time slot. (Later on we will see that the results do not really depend on the precise value of* τ_{ij}.*)*

(A1.ii) $\{H_{ijl}(t)\}$, $\{Z_j(t)\}$, $\{Y_j(t)\}$ *and* $\{X_i(t)\}$ *are* complex *variables.*

- $\{H_{ijl}(t) : t \geq 1, i, j \in N\}$ *is the random fading process.*
- $\{Z_j(t) : j \in N, t \geq 1\}$ *are i.i.d. complex circular Gaussian noises*[3] *independent of the fading process* $\{H_{ijl}(t)\}$, *and are not observable to the users. We suppose* $E|Z_j(t)|^2 = \sigma^2 > 0$.
- $\{X_i(t)\}$ *is the complex base-band signal sequence node* i *transmits, and* $\{Y_j(t)\}$ *is the complex base-band signal sequence node* j *receives.*

(A1.iii) *Each node is subject to an individual power constraint* P. *Since the channel has multiple paths with delays, we need to model what may have been transmitted before time 0, the time when the useful transmissions begin. We simply suppose that the signals prior to* $t = 0$ *satisfy*

$$|X_i(t)|^2 \leq \bar{P} < \infty, \qquad \forall t \leq 0, i \in N, \tag{2}$$

and are unknown to the nodes.

(A1.iv) *There exist positive constants* $\alpha \in (0, 1)$ *and* $\bar{H} > 0$, *such that for any* $i, j \in N$, *the fading process satisfies*

$$\limsup_{T} \frac{1}{T} \sum_{t=1}^{T} \sum_{l=0}^{\infty} \alpha^{-l} |H_{ijl}(t)|^2 \quad \leq \quad \bar{H}, a.s. \tag{3}$$

Assumption set (A2):

(A2.i, ii, iii) *Same as (A1.i, ii, iii).*

(A1.iv) *The fading process* $\{H_{ijl}(t)\}$ *satisfies the following:*

- $\{\mathcal{H}_j(t),\ t \geq 1\}$ *is a sequence of independent random vectors, where* $\mathcal{H}_j(t) := (H_{ijl}(t), i \in N, l \geq 0)$ *for all* $j \in N$. *That is, the fading parameters are independent from time to time.*
- *There exist* $a > 0$ *and* $p^* > 1/2$ *such that, for all* $i, j \in N$, $t \geq 1$, *Pr*$(|H_{ij0}(t)| \geq a) \geq p^*$.
- *There exist* $\alpha \in (0, 1)$ *and* $\tilde{H} > 0$ *such that, for all* $i, j \in N$,

$$\sum_{l=0}^{\infty} \alpha^{-l} E|H_{ijl}(t)|^2 \quad \leq \quad \tilde{H}. \tag{4}$$

(A2.v) *There exists* $\zeta \geq 1$ *such that for every node* j *there exists another node* $\hat{j}$ *with* $\rho_{j,\hat{j}} \leq \zeta \rho_{min}$.

Remark 2.1: In Theorem 3.1, we will prove an $O(n)$ upper bound on the transport capacity (defined below) for Assumption set (A1) even when the CSI is known non-causally and *a priori* at both the transmitters and receivers. In Theorem 3.2, we will prove a constructive $\Omega(n)$ lower bound on the transport capacity for Assumption set (A2) even when the CSI is neither known at the transmitters nor the receivers. Both results hold whenever there is any absorption, i.e., $\gamma > 0$, or if the path loss exponent is greater than 3. It has been shown in [28] that generally the absorption is positive.

Assumption set (A1) includes channel models where the fading is flat, i.e. $H_{ijl}(t) \equiv 0$ for all $l \geq 1$ and $i, j \in N$. By allowing for CSI to be known at both the transmitters and the receivers, (A1) covers what is usually meant by "slow fading" where the channel varies statistically but with a rate of variation substantially smaller than the signaling period, so that by use of a relatively short training sequence the channel state can be estimated well, and then be regarded as known for the symbol period. Thus the upper bound result under (A1) with the CSI known at the transmitter and receiver covers the case of "flat, slow" fading. However the assumption (A1) allows for a much larger class of channels, including, for example, flat or frequency selective, stationary ergodic channels, or channels varying in a deterministic time varying fashion, since basically all it requires is the condition (A1.iv).

Assumption set (A2) includes the case of frequency selective fading since it allows H_{ijl} to depend on $l \geq 1$. It also includes "fast fading" where $H_{ijl}(t)$ is an independent stochastic process rapidly varying in comparison to the symbol duration. Thus it covers the case of "frequency selective,

[3] A complex random variable Z is circular Gaussian if it can be represented as $Z = Z_1 + \nu Z_2$ where Z_1 and Z_2 are two i.i.d. (real) Gaussian random variables, and ν is the square root of -1.

rapidly varying channels.” The condition (A2.v) effectively requires that every node has a nearby node, a reasonable requirement for networks, and is needed by us to prove the feasibility result. The $\Omega(n)$ feasibility result for (A2) holds even without knowing CSI at either transmitters or receivers. Note that assumption (A2) also allows other channels, for example, flat, constant fading, or flat, independent fading channels.

Definition of feasible rate vectors

For a given wireless network $\mathcal{G}_n$, we employ the standard definition of what is meant by a *feasible information rate vector* $R = \{R_{ij}, i, j \in N, i \neq j\}$; see [29]:

1) With W_{ij} denoting the message to be sent from node i to node j, we assume that all the messages W_{ij} are independent, and uniformly distributed over their respective ranges $\{1, 2, \ldots, 2^{TR_{ij}}\}$.
2) The symbol $X_i(t)$, for $t \geq 1$, that node i sends out at time t depends on its own outgoing messages $\{W_{ij}, j \in N, j \neq i\}$, as well as the values of its past received symbols $\{Y_i(s), 0 \leq s \leq t-1\}$. An encoding scheme of block length T consists of a set of encoding and decoding functions, one for each node i, as follows:
   - *Encoders* $X_i(t; W_{i1}, W_{i2}, \ldots, W_{in}; Y_i(1), Y_i(2), \ldots, Y_i(t-1))$, $t = 1, \ldots, T$. The encoder maps the messages that node i wants to send to the other nodes, and its past received symbols, into the symbol $X_i(t)$ transmitted at time t. The average power of the symbols $X_i(t)$ is required to satisfy the constraint,
$$\frac{1}{T}\sum_{t=1}^{T}|X_i(t)|^2 \quad \leq \quad P. \tag{5}$$
   - *Decoders* $\widehat{W}_{ji}(Y_i(1), \ldots, Y_i(T); W_{i1}, \ldots, W_{in})$, $j = 1, 2, \ldots, n$. The decoder $\widehat{W}_{ji}$ at node i, for node j's message, maps the received symbols in each block, and its own information, to form an estimate of the message W_{ji} intended for it from node $j \in N$.
3) For all (i, j), the corresponding probability of error that the message sent from node i to j will not be decoded correctly, converges to zero as T goes to infinity, i.e.,
$$P_e^T(i,j) = \Pr(\widehat{W}_{ij}(Y_j(1), \ldots, Y_j(T); W_{j1}, \ldots, W_{jn}) \neq W_{ij}) \to 0, \quad \text{as } T \to \infty.$$

Now we are in a position to define the *transport capacity* of wireless networks.

Definition 2.1: For a given wireless network $\mathcal{G}_n$ with fixed node locations, *its transport capacity* $C_T(\mathcal{G}_n)$ is defined as
$$C_T(\mathcal{G}_n) \quad := \quad \sup_{R \in \mathcal{R}(\mathcal{G}_n)} \sum_{i,j} R_{ij}\rho_{ij},$$
where the optimization is over $\mathcal{R}(\mathcal{G}_n)$, the set of all feasible information rate vectors for network $\mathcal{G}_n$, i.e., without changing the node locations in $\mathcal{G}_n$.

We also define the very best transport capacity that can be delivered by any network with n nodes.

Definition 2.2: The transport capacity $C_T^{(n)}$ of the class of wireless networks with n nodes is defined as
$$C_T^{(n)} \quad = \quad \sup_{\mathcal{G}_n} C_T(\mathcal{G}_n) = \sup_{\mathcal{G}_n} \sup_{R \in \mathcal{R}(\mathcal{G}_n)} \sum_{i,j} R_{ij}\rho_{ij}.$$
That is, the optimization is over all wireless networks with n nodes.

III. Main results

Our main results are the following two theorems.

Theorem 3.1: Under assumption set (A1), even if the CSI is known non-causally to all transmitters and receivers, the transport capacity is bounded as
$$C_T^{(n)} \quad \leq \quad c_1 \cdot n, \qquad \text{for all } n,$$
where
$$c_1 := \begin{cases} \frac{8\beta\sqrt{P\bar{H}}\log e}{\sigma\sqrt{1-\alpha}(\rho_{min}/\sqrt{2})^{\delta-1}}\left(\frac{2(\delta-2)}{\delta-3} + \frac{\delta-1}{\delta-2}\right), \\ \qquad\qquad \text{if } \gamma = 0, \delta > 3; \\ \frac{48\beta\sqrt{P\bar{H}}\log e \cdot \rho_{min}^{1-\delta}}{\sigma\sqrt{1-\alpha}} \frac{e^{-(\sqrt{2}/2)\gamma\rho_{min}}}{\left(1-e^{-(\sqrt{2}/4)\gamma\rho_{min}}\right)^4}, \\ \qquad\qquad \text{if } \gamma > 0, \delta \geq 0. \end{cases}$$

Theorem 3.2: Under assumption set (A2), even if the CSI is unknown to transmitters or receivers, for any $\bar{p} \in (1/2, p^*)$ there exists a constant $c_2 > 0$, such that for any network $\mathcal{G}_n$,
$$C_T(\mathcal{G}_n) \geq c_2 \cdot n, \qquad \text{for all } n,$$
where $\quad c_2 :=$
$$\begin{cases} \min\left\{\frac{1}{2}, \frac{P\beta^2(\zeta\rho_{min})^{-2\delta}a^2\cdot(p^*-\bar{p})}{\frac{16\tilde{H}P\beta^2\rho_{min}^{-2\delta}}{1-\alpha}(1+2^{2+\delta}(\frac{2\delta-1}{\delta-1}+\frac{2\delta}{2\delta-1}))+4\delta^2}\right\} \\ \qquad \cdot(1-H(\bar{p}))\cdot\rho_{min}, \text{ if } \delta > 3, \gamma = 0; \\ \min\left\{\frac{1}{2}, \frac{P\beta^2(\zeta\rho_{min})^{-2\delta}e^{-2\gamma\zeta\rho_{min}}a^2\cdot(p^*-\bar{p})}{\frac{16\tilde{H}P\beta^2\rho_{min}^{-2\delta}}{1-\alpha}\left(e^{-2\gamma\rho_{min}}+\frac{12e^{-\sqrt{2}\gamma\rho_{min}}}{(1-e^{-\sqrt{2}\gamma\rho_{min}})^2}\right)+4\delta^2}\right\} \\ \qquad \cdot(1-H(\bar{p}))\cdot\rho_{min}, \text{ if } \delta \geq 0, \gamma > 0, \end{cases}$$
with $H(\bar{p}) := -\bar{p}\log(\bar{p}) - (1-\bar{p})\log(1-\bar{p})$. In particular,
$$C_T^{(n)} \geq c_2 \cdot n, \qquad \text{for all } n.$$

Remark 3.1: It may be noted that the above result generalizes the $\Theta(n)$ feasibility result shown in [19] for regular networks, where nodes are located at integer lattice sites in a square, to the more general class of networks satisfying property Assumption (A2.v).

IV. A max-flow min-cut lemma for time-varying media

In this section we present a useful lemma showing that the information rate one set of nodes can receive from the rest of the network is upper-bounded by a function of the power it can receive from the other nodes. It generalizes the corresponding result in [19] to fading environments.

Lemma 4.1 (Max-flow min-cut bound): Suppose a wireless communication network is modeled as
$$Y_j(t) = \sum_{i \neq j}\sum_{l=0}^{\infty} A_{ijl}(t)X_i(t-\tau_{ij}-l) + Z_j(t),\ j \in N,$$

where: *(i)* The τ_{ij}'s are deterministic non-negative integers. *(ii)* $\{A_{ijl}(t)\}$ is a sequence of *known deterministic* complex numbers. *(iii)* $\{Z_j(t)\}$ is the i.i.d. circular Gaussian noise process independent of the signal process $\{X_i(t)\}$, and $E|Z_j(t)|^2 = \sigma^2$.

Then for any subset S of N, the rate vector $\{R_{ij}, i, j \in N\}$ satisfies

$$R_{SD} \le 1 + TR_{SD}P_e^{(T)} + \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in N\backslash S} \log(1 + \frac{E|\sum_{i\in S}\sum_{l=0}^{\infty} A_{ijl}(t)X_i(t-\tau_{ij}-l)|^2}{\sigma^2}),$$

where $N\backslash S$ denotes those nodes in N but not in S, $R_{SD} := \sum_{i\in S, j\in N\backslash S} R_{ij}$, and $P_e^{(T)}$ is the probability of decoding error.

Proof: The proof is similar to the proof of Lemma 4.1 in [19]. Let $D := N\backslash S$ be a set of destination nodes, and let $W_{ij} := \{1, 2, \dots, 2^{TR_{ij}}\}$ denote the message set from node i to j. We use the following notation:

$$\begin{aligned} U_j(t) &:= \sum_{i\in S}\sum_{l=0}^{\infty} A_{ijl}(t)X_{i,T}(t-\tau_{ij}-l), \quad j \in D; \\ V_j(t) &:= U_j(t) + Z_j(t); \\ W_{SD} &:= \{W_{ij} : i \in S, j \in D\}; \\ W_D &:= \{W_{ij} : i \in D, j \in N\}; \\ W_i &:= \{W_{ij} : j \in N\}. \end{aligned}$$

Let $V_D(t) := \{V_j(t) : j \in D\}$, $V_D^t := \{V_D(k) : k = 1, \dots, t\}$, and similarly for Y, U, Z. Also, let $\Gamma_S(t) := \{A_{ijl}(t)X_i(t-\tau_{ij}-l) : i \in S, j \in D, l \ge 0\}$.

First we want to show that $W_{SD} \to \{V_D^T, W_D\} \to \{Y_D^T, W_D\}$ forms a Markov chain. This can be done by showing that Y_D^T is a deterministic function of (V_D^T, W_D) and the fading coefficients $A^T := \{A_{ijl}(t), \forall i, j, l, t \ge 1\}$. Actually, for any $j \in D$, $2 \le t \le T$,

$$\begin{aligned} Y_j(t) &= V_j(t) + \sum_{i\in D, i\ne j}\sum_{l} A_{ijl}(t)X_{i,T}(t-\tau_{ij}-l) \\ &= V_j(t) + \sum_{i\in D, i\ne j}\sum_{l} A_{ijl}(t)f_{i,t}(Y_i^{t-\tau_{ij}-l-1}, W_i), \end{aligned}$$

and

$$Y_j(1) = V_j(1) + \sum_{i\in D, i\ne j}\sum_{l} A_{ijl}(1)f_{i,1}(W_i).$$

Now by Fano's Lemma and the property of a Markov chain, we have

$$H(W_{SD}|V_D^T, W_D) \le H(W_{SD}|Y_D^T, W_D) \le 1 + TR_{SD}P_e^{(T)}.$$

Thus,

$$\begin{aligned} &TR_{SD} \\ &= H(W_{SD}) = I(W_{SD}; V_D^T, W_D) + H(W_{SD}|V_D^T, W_D) \\ &\le I(W_{SD}; V_D^T, W_D) + 1 + TR_{SD}P_e^{(T)} \\ &= I(W_{SD}; W_D) + I(W_{SD}; V_D^T|W_D) + 1 + TR_{SD}P_e^{(T)} \\ &= 0 + h(V_D^T|W_D) - h(V_D^T|W_{SD}, W_D) + 1 + TR_{SD}P_e^{(T)} \\ &\le h(V_D^T) - h(V_D^T|W_{SD}, W_D) + 1 + TR_{SD}P_e^{(T)}, \end{aligned}$$

with

$$\begin{aligned} &h(V_D^T|W_{SD}, W_D) \\ &= \sum_{t=1}^{T} h(V_D(t)|V_D(1), \dots, V_D(t-1), W_{SD}, W_D) \\ &\ge \sum_{t=1}^{T} h(V_D(t)|V_D(1), \dots, V_D(t-1), \Gamma_S(t), W_{SD}, W_D) \\ &= \sum_{t=1}^{T} h(V_D(t)|\Gamma_S(t)) \ge \sum_{t=1}^{T} h(V_D(t)|U_D(t)). \end{aligned}$$

Hence,

$$\begin{aligned} &TR_{SD} \\ &\le h(V_D^T) - \sum_{t=1}^{T} h(V_D(t)|U_D(t)) + 1 + TR_{SD}P_e^{(T)} \\ &= h(V_D^T) - \sum_{t=1}^{T}\sum_{j\in D} h(Z_j(t)) + 1 + TR_{SD}P_e^{(T)} \\ &\le \sum_{t=1}^{T}\sum_{j\in D} h(V_j(t)) - \sum_{t=1}^{T}\sum_{j\in D}\log(\pi e\sigma^2) \\ &\qquad + 1 + TR_{SD}P_e^{(T)} \\ &\le \sum_{t=1}^{T}\sum_{j\in D}(\log(\pi e(E|U_j(t)|^2 + \sigma^2)) - \log(\pi e\sigma^2)) \\ &\qquad + 1 + TR_{SD}P_e^{(T)} \\ &= \sum_{t=1}^{T}\sum_{j\in D}\log(1 + \frac{E|U_j(t)|^2}{\sigma^2}) + 1 + TR_{SD}P_e^{(T)}, \end{aligned}$$

where the last inequality follows from Lemma 2 in [30]. □

Remark 4.1: Lemma 4.1 differs from Lemma 4.1 of [19] in allowing for a time-varying fading process.

With Lemma 4.1 in hand, we can easily get the following corollary.

Corollary 4.2: If a wireless network satisfies assumption set (A1), and the fading process is *known beforehand*, then for any subset S of N, and feasible rate vector $\{R_{ij}, i, j \in N\}$, $R_{SD} := \sum_{i\in S, j\in N\backslash S} R_{ij}$ satisfies

$$R_{SD} \le \log e \cdot \frac{2\beta\sqrt{P\bar{H}}}{\sigma\sqrt{1-\alpha}} \cdot \sum_{i\in S, j\in N\backslash S} \rho_{ij}^{-\delta} e^{-\gamma\rho_{ij}}. \tag{6}$$

Proof: Let $D = N\backslash S$. Since the rate vector is feasible, we know $P_e^{(T)} \to 0$ as $T \to \infty$. So by the proof of Lemma 4.1, we have

$$R_{SD} \le \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in D} \log\left(1 + \frac{\beta^2 E\left|\sum_{i\in S}\sum_{l=0}^{\infty}\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}} H_{ijl}(t)X_{i,T}(t-\tau_{ij}-l)\right|^2}{\sigma^2}\right) + o(1).$$

Furthermore,

$$
\begin{aligned}
&\frac{\beta^2}{\sigma^2} E\left|\sum_{i\in S}\sum_{l=0}^{\infty}\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}H_{ijl}(t)X_{i,T}(t-\tau_{ij}-l)\right|^2 \\
&= \frac{\beta^2}{\sigma^2} E\left|\sum_{i\in S}\sum_{l=0}^{\infty}\left(\rho_{ij}^{-\delta/2}e^{-\gamma\rho_{ij}/2}H_{ijl}(t)\alpha^{-l/2}\right)\cdot\right. \\
&\qquad\qquad \left.\left(\rho_{ij}^{-\delta/2}e^{-\gamma\rho_{ij}/2}X_{i,T}(t-\tau_{ij}-l)\alpha^{l/2}\right)\right|^2 \\
&\le \frac{\beta^2}{\sigma^2} E\left(\sum_{i\in S}\sum_{l=0}^{\infty}\left|\rho_{ij}^{-\delta/2}e^{-\gamma\rho_{ij}/2}H_{ijl}(t)\alpha^{-l/2}\right|^2\right)\cdot \\
&\qquad\qquad \left(\sum_{i\in S}\sum_{l=0}^{\infty}\left|\rho_{ij}^{-\delta/2}e^{-\gamma\rho_{ij}/2}X_{i,T}(t-\tau_{ij}-l)\alpha^{l/2}\right|^2\right) \\
&= \frac{\beta^2}{\sigma^2}\left(\sum_{i\in S}\sum_{l=0}^{\infty}\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\cdot|H_{ijl}(t)|^2\cdot\alpha^{-l}\mu\right)\cdot \\
&\qquad \left(\sum_{i\in S}\sum_{l=0}^{\infty}\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\cdot E\left|X_{i,T}(t-\tau_{ij}-l)\right|^2\cdot\alpha^{l}\mu^{-1}\right),
\end{aligned}
$$

where the inequality comes from the Cauchy-Schwarz inequality, and the last equality follows from the fact that we know $H_{ijl}(t)$ beforehand, with μ a constant to be determined later.

It is easy to verify that for any non-negative a and b, $\log(1+ab)\le(a+b)\log e$. Hence

$$
\begin{aligned}
& R_{SD}-o(1) \\
\le\;& \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in D}\log\left(1+\left(\sum_{i\in S}\sum_{l=0}^{\infty}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta\mu}{\sigma}|H_{ijl}(t)|^2\alpha^{-l}\right)\right. \\
&\quad\left.\cdot\left(\sum_{i\in S}\sum_{l=0}^{\infty}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta\mu^{-1}}{\sigma}E|X_{i,T}(t-\tau_{ij}-l)|^2\alpha^{l}\right)\right) \\
\le\;& \log e\cdot\frac{1}{T}\sum_{t=1}^{T}\sum_{j\in D}\sum_{i\in S}\sum_{l=0}^{\infty}\left(\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta\mu}{\sigma}|H_{ijl}(t)|^2\alpha^{-l}\right. \\
&\quad\left.+\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta\mu^{-1}}{\sigma}E|X_{i,T}(t-\tau_{ij}-l)|^2\alpha^{l}\right) \\
=\;& \log e\sum_{i\in S,j\in D}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta}{\sigma}\left(\mu\cdot\frac{1}{T}\sum_{t=1}^{T}\sum_{l=0}^{\infty}|H_{ijl}(t)|^2\alpha^{-l}\right. \\
&\quad\left.+\mu^{-1}\cdot\frac{1}{T}\sum_{t=1}^{T}\sum_{l=0}^{\infty}E|X_{i,T}(t-\tau_{ij}-l)|^2\alpha^{l}\right) \\
\overset{(a)}{\le}\;& \log e\sum_{i\in S,j\in D}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta}{\sigma}\Big(\mu\bar{H}+o(1) \\
&\quad+\mu^{-1}\cdot\sum_{l=0}^{\infty}E\left(\frac{1}{T}\sum_{t=1}^{T}|X_{i,T}(t-\tau_{ij}-l)|^2\right)\alpha^{l}\Big) \\
\overset{(b)}{\le}\;& \log e\sum_{i\in S,j\in D}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta}{\sigma}\Big(\mu\bar{H}+o(1) \\
&\quad+\mu^{-1}\cdot\sum_{l=0}^{\infty}\left(P+\frac{\tau_{ij}+l}{T}\bar{P}\right)\alpha^{l}\Big) \\
=\;& \log e\sum_{i\in S,j\in D}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta}{\sigma}\Big(\mu\bar{H}+o(1) \\
&\quad+\mu^{-1}\frac{P}{1-\alpha}+\mu^{-1}\frac{\tau_{ij}}{T}\frac{\bar{P}}{1-\alpha}+\mu^{-1}\frac{\bar{P}}{T}\sum_{l=0}^{\infty}l\alpha^{l}\Big) \\
=\;& \log e\sum_{i\in S,j\in D}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta}{\sigma}\Big(\mu\bar{H}+o(1) \\
&\quad+\mu^{-1}\frac{P}{1-\alpha}+\frac{\mu^{-1}}{T}\left(\frac{\tau_{ij}\bar{P}}{1-\alpha}+\bar{P}\frac{\alpha}{(1-\alpha)^2}\right)\Big),
\end{aligned}
$$

where (a) is because of (A1.iv), and (b) is because of (2) and (5).

Now, letting $T\to\infty$, we get

$$
R_{SD}\;\le\;\log e\sum_{i\in S,j\in D}\frac{\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}}\beta}{\sigma}\left(\mu\bar{H}+\mu^{-1}\frac{P}{1-\alpha}\right).
$$

The result holds by setting $\mu=\sqrt{\frac{P}{\bar{H}(1-\alpha)}}$. □

V. From cut-sets to distances

We now show a natural relationship between cut-sets and the distance-rate product. It allows us to easily convert results for rate vectors across cut-sets, a staple feature in network information theory, to results on the transport capacity. It also renders the transport capacity an even more appealing quantity to study in networks.

Lemma 5.1: If for a set of numbers $\{a_{ij}, i,j\in N\}$, $\sum_{i,j}a_{ij}I_{[i\in S,j\in N\setminus S]}\ge 0$ holds for every subset S of N, then $\sum_{i,j}a_{ij}\rho_{ij}\ge 0$.

Proof: By the symmetry of the condition, for any subset S of N,

$$
\sum_{i=1}^{n}\sum_{j\ne i}a_{ij}\left(I_{[i\in S,j\notin S]}+I_{[i\notin S,j\in S]}\right)\ge 0. \tag{7}
$$

Now we will construct a model to randomly select a subset S of N. Suppose the n nodes of N are located within a disk of radius r centered at the origin, as depicted in Figure 1.

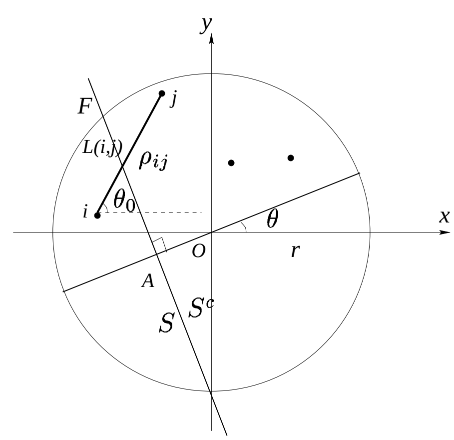

Fig. 1. The procedure for choosing a random subset S of N in Lemma 5.1

The selection proceeds as follows. First we pick an angle θ uniformly in $[0,2\pi)$. Then, on the diameter corresponding to that angle, we pick a point uniformly. Denote the point by A. Then the line F that is perpendicular to the diameter, and intersecting it at point A, will divide the n nodes into two groups. We randomly and equiprobably pick one of these to be the set S.

Suppose the line interval $L(i,j)$ connecting nodes i and j makes an angle θ_0 with the x-axis. Then,

$$\begin{aligned}
&E(I_{[i\in S, j\notin S]} + I_{[i\notin S, j\in S]}) \\
&= EI_{[\text{Line } F \text{ separates points } i \text{ and } j.]} \\
&= \text{Prob (Line } F \text{ separates points } i \text{ and } j.) \\
&= \int_0^{2\pi} \frac{\rho_{ij}|\cos(\theta-\theta_0)|}{2r} \cdot \frac{1}{2\pi} d\theta \\
&= \frac{\rho_{ij}}{4\pi r} \int_0^{2\pi} |\cos(\theta-\theta_0)| d\theta \\
&= \frac{\rho_{ij}}{4\pi r} \int_0^{2\pi} |\cos\theta| d\theta \\
&= \frac{\rho_{ij}}{4\pi r} 4 \int_0^{\pi/2} \cos\theta d\theta = \frac{\rho_{ij}}{\pi r}.
\end{aligned}$$

Now, taking the expectation on both sides of (7), we get

$$\begin{aligned}
&E \sum_{i=1}^{n} \sum_{j\neq i} a_{ij} (I_{[i\in S, j\notin S]} + I_{[i\notin S, j\in S]}) \\
&= \sum_{i=1}^{n} \sum_{j\neq i} a_{ij} E(I_{[i\in S, j\notin S]} + I_{[i\notin S, j\in S]}) \\
&= \sum_{i=1}^{n} \sum_{j\neq i} a_{ij} \frac{\rho_{ij}}{\pi r} \geq 0.
\end{aligned}$$

Hence $\sum_{i=1}^{n} \sum_{j\neq i} a_{ij}\rho_{ij} \geq 0$. □

VI. An upper bound on the transport capacity

To show the proof of the upper bound, we need the following lemma. (Lemma 6.1 (iii) is needed for Lemma 9.1).

Lemma 6.1: (i) For any $i_0 \in N$ and $\delta' > 2$,

$$\sum_{j\neq i_0} \rho_{i_0,j}^{-\delta'} \leq \frac{4}{(\rho_{min}/\sqrt{2})^{\delta'}} \left(\frac{2(\delta'-1)}{\delta'-2} + \frac{\delta'}{\delta'-1} \right).$$

(ii) For any $i_0 \in N$, $\delta \geq 0$, and $\gamma > 0$,

$$\sum_{j\neq i_0} \rho_{i_0,j}^{1-\delta} e^{-\gamma\rho_{i_0 j}} \leq \frac{24\rho_{min}^{1-\delta} e^{-(\sqrt{2}/2)\gamma\rho_{min}}}{(1-e^{-(\sqrt{2}/4)\gamma\rho_{min}})^4}.$$

(iii) For any $i_0 \in N$ and $\gamma > 0$,

$$\sum_{j\neq i_0} e^{-2\gamma\rho_{i_0 j}} \leq \frac{12e^{-\sqrt{2}\gamma\rho_{min}}}{(1-e^{-\sqrt{2}\gamma\rho_{min}})^2}.$$

Proof: Without loss of generality, suppose that $i_0 = 1$, and that it is located at the origin. Let us tessellate the plane by a square grid of size $\rho' := (\sqrt{2}/2)\rho_{min}$ with the origin being one of the corners of the grid; see Figure 2.

Since the diagonals of the small squares have lengths ρ_{min}, each of them can contain at most one node. Also, there is no node within the interior of the four squares surrounding the origin.

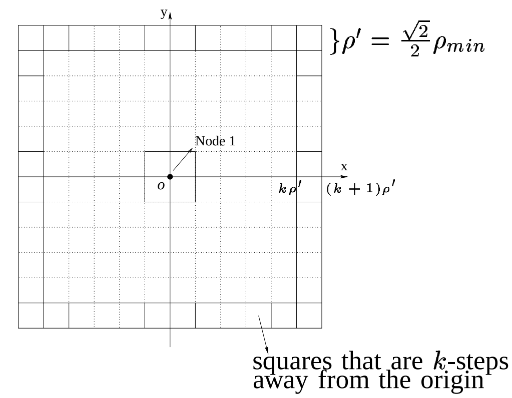

Fig. 2. Square tessellation of the plane

Since there are $4(2k+1)$ squares that are k steps away (in an l_1-sense) from the origin (see Figure 2), for case (i) we have

$$\begin{aligned}
\sum_{j\neq 1} \rho_{1j}^{-\delta'} &\leq \sum_{k=1}^{\infty} 4(2k+1)(k\rho')^{-\delta'} \\
&= \frac{4}{\rho'^{\delta'}} \sum_{k=1}^{\infty} (2k+1) \frac{1}{k^{\delta'}} \\
&= \frac{8}{\rho'^{\delta'}} \sum_{k=1}^{\infty} \frac{1}{k^{\delta'-1}} + \frac{4}{\rho'^{\delta'}} \sum_{k=1}^{\infty} \frac{1}{k^{\delta'}} \\
&\leq \frac{8}{\rho'^{\delta'}} (1 + \int_1^{\infty} \frac{1}{x^{\delta'-1}} dx) + \frac{4}{\rho'^{\delta'}} (1 + \int_1^{\infty} \frac{1}{x^{\delta'}} dx) \\
&= \frac{8}{\rho'^{\delta'}} (1 + \frac{1}{\delta'-2}) + \frac{4}{\rho'^{\delta'}} (1 + \frac{1}{\delta'-1}) \\
&= \frac{4}{\rho'^{\delta'}} \left(\frac{2(\delta'-1)}{\delta'-2} + \frac{\delta'}{\delta'-1} \right).
\end{aligned}$$

Similarly, for case (ii),

$$\begin{aligned}
&\sum_{j\neq 1} \rho_{1j}^{1-\delta} e^{-\gamma\rho_{1j}} \\
&\leq \rho_{min}^{-\delta} \sum_{j\neq 1} \rho_{1j} e^{-\gamma\rho_{1j}} \\
&\leq \rho_{min}^{-\delta} \sum_{k=1}^{\infty} 4(2k+1) \cdot (k+1)\sqrt{2}\rho' \cdot e^{-\gamma k\rho'} \\
&\leq \rho_{min}^{1-\delta} \sum_{k=1}^{\infty} 4 \cdot 3k \cdot 2k \cdot e^{-\gamma k\rho'} \\
&= 24\rho_{min}^{1-\delta} \sum_{k=1}^{\infty} k^2 (e^{-\gamma\rho'})^k \\
&\leq 24\rho_{min}^{1-\delta} \left(\sum_{k=1}^{\infty} k(e^{-\gamma\rho'/2})^k \right)^2 \\
&= 24\rho_{min}^{1-\delta} \left(e^{-\gamma\rho'/2} \sum_{k=0}^{\infty} k(e^{-\gamma\rho'/2})^{k-1} \right)^2 \\
&= 24\rho_{min}^{1-\delta} \left(\frac{e^{-\gamma\rho'/2}}{(1-e^{-\gamma\rho'/2})^2} \right)^2
\end{aligned}$$

$$= 24\rho_{min}^{1-\delta}\frac{e^{-\gamma\rho'}}{\left(1-e^{-\gamma\rho'/2}\right)^4}.$$

For case (iii), we have

$$\sum_{j\neq 1} e^{-2\gamma\rho_{1j}} \leq \sum_{k=1}^{\infty} 4(2k+1)e^{-2\gamma k\rho'} \leq \sum_{k=1}^{\infty} 12k(e^{-2\gamma\rho'})^k$$
$$= 12e^{-2\gamma\rho'}\sum_{k=1}^{\infty} k(e^{-2\gamma\rho'})^{k-1} = \frac{12e^{-2\gamma\rho'}}{(1-e^{-2\gamma\rho'})^2}.\square$$

Now we prove the upper bound of Theorem 3.1.

Proof of Theorem 3.1: By Corollary 4.2, we know that for any subset S of N, $\sum_{i,j} R_{ij} I_{[i\in S, j\in N\setminus S]} \leq \log e \cdot \frac{2\beta\sqrt{P\bar{H}}}{\sigma\sqrt{1-\alpha}} \cdot \sum_{i,j}\rho_{ij}^{-\delta}e^{-\gamma\rho_{ij}} I_{[i\in S, j\in N\setminus S]}$. Hence by Lemma 5.1,

$$\sum_{ij} R_{ij}\rho_{ij} \quad\leq\quad \log e\cdot\frac{2\beta\sqrt{P\bar{H}}}{\sigma\sqrt{1-\alpha}}\cdot\sum_{ij}\rho_{ij}^{1-\delta}e^{-\gamma\rho_{ij}}.$$

Applying Lemma 6.1 (i) and (ii) we get the desired result. $\square$

VII. Networks which achieve the lower bound under independent fading

Beginning with this section, we constructively show that the linear growth rate of the transport capacity is achievable in wireless networks. Actually, we show that it is achievable in any network that satisfies the property specified in assumption (A2.v); see Figure 3. The simplest example of a network satisfying this property is the regular network defined in [19], which is a network with nodes located at the coordinates $(k\rho_{min}, l\rho_{min})$ for $1 \leq k, l \leq \sqrt{n}$.

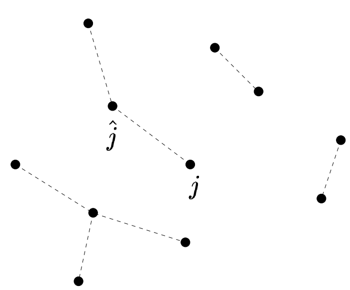

Fig. 3. A network satisfying property (A2.v). All dashed lines have lengths less than a multiple of the minimum distance.

Suppose that the destination node for the information originating at each node j is the corresponding node $\hat{j}$. Furthermore, without loss of generality, we assume that the delay $\tau_{j\hat{j}} = 0$ for all j.

In the following sections, we show how to generate the codebooks, then how to coordinate the transmissions by a schedule generated by a randomization scheme, and then how to decode the signals by thresholding. Finally, we analyze the probability of decoding error.

We begin by fixing a number $\bar{p} \in (1/2, p^*)$. For $\epsilon > 0$ sufficiently small, we introduce the following quantities for brevity:

$$P_\epsilon \quad := \quad P - \epsilon,$$

$$\lambda \quad := \quad \begin{cases} \frac{4}{\left(\rho_{min}/\sqrt{2}\right)^{2\delta}}\left(\frac{2\delta-1}{\delta-1}+\frac{2\delta}{2\delta-1}\right), \text{if } \delta > 3, \gamma = 0; \\ \frac{12\rho_{min}^{-2\delta}e^{-\sqrt{2}\gamma\rho_{min}}}{(1-e^{-\sqrt{2}\gamma\rho_{min}})^2}, \text{if } \delta \geq 0, \gamma > 0, \end{cases} \tag{8}$$

$$\theta \quad := \quad \min\{1/2, \tag{9}$$
$$\frac{(P-\epsilon)\beta^2(\zeta\rho_{min})^{-2\delta}e^{-2\gamma\zeta\rho_{min}}a^2\cdot(p^*-\bar{p})}{\frac{16\tilde{H}(P-\epsilon)\beta^2}{1-\alpha}\left(\rho_{min}^{-2\delta}e^{-2\gamma\rho_{min}}+\lambda\right)+4\delta^2}\},$$

$$\theta_\epsilon \quad := \quad \theta - \epsilon.$$

For any ϵ making the above quantities positive, and $\epsilon_1 \in (0, \bar{p}-1/2)$, we are going to show that the information rate $\theta_\epsilon(1-H(\bar{p}-\epsilon_1)-\epsilon_1)$ is achievable for every node pair j and $\hat{j}$, simultaneously. This certainly suffices to prove Theorem 3.2.

From now on, $\bar{p}$, ϵ and ϵ_1 are fixed.

VIII. Random coding

Each node is given $\theta_\epsilon T$ slots to transmit during a communication horizon T. The scheduling details will be given in the next section. In this section, we apply a random coding method to generate the codebook, similar to classical information theory. The only additional feature is that we generate the codebooks for different T's in a coupled fashion, to facilitate analysis.

For a given rate $R = 1 - H(\bar{p}-\epsilon_1) - \epsilon_1$, the n nodes generate their codebooks individually, independently of each other. Node j generates a $2^{\theta_\epsilon TR} \times \theta_\epsilon T$ random matrix with entries being i.i.d. binary valued r.v.'s with distribution $p(x)$ such that $\Pr(X = 0) = 1/2$, and $\Pr(X = \sqrt{P_\epsilon/\theta}) = 1/2$. The w^{th} codeword is the w^{th} row of this matrix. The codebook of node j is denoted as $C_j := \{X_w^j = (X_{w,1}^j, X_{w,2}^j, \ldots, X_{w,\theta_\epsilon T}^j) : w = 1, 2, \ldots, 2^{\theta_\epsilon TR}\}$, and it is revealed to the intended receiver node $\hat{j}$. We denote the codebook by $C := \{C_1, \ldots, C_n\}$.

Notice that the codebook actually depends on T. We construct the random codebooks for different T's in a coupled way that such that the codebook for a larger T is an extension of the codebooks for smaller ones. That is, for $T < T'$, $X_{w,k}^j(T) = X_{w,k}^j(T')$ whenever $k \leq \theta_\epsilon T$ and $1 \leq w \leq 2^{\theta_\epsilon TR} < 2^{\theta_\epsilon T'R}$, for all $j \in N$.

IX. Scheduling transmissions without excessive interference

If we allow all the nodes to transmit in the same time slots, then each receiver will face strong interference from nearby nodes. So we let nodes transmit in a timeshared fashion. Specifically, for any given large $T > 0$, each node $j \in N$ only transmits at a set of pre-selected increasing time-slots t_k^j, $k = 1, 2, \ldots, \theta_\epsilon T$. We call this set the *duty slots of node* j. The corresponding (intended) receiver $\hat{j}$ will decode based only on the signals it receives at time slots t_k^j, $k = 1, \ldots, \theta_\epsilon T$. (Note that here we use the assumption that $\tau_{j\hat{j}} = 0$ for all $j \in N$).

Let the indicator function $b_i(t)$, $i \in N$, be defined as follows:

$$b_i(t) \quad = \quad \begin{cases} 1, & \text{if slot } t \text{ is in node } i\text{'s duty slot;} \\ 0, & \text{if otherwise.} \end{cases} \tag{10}$$

Then the following result guarantees the existence of a "good" time-sharing schedule.

Lemma 9.1 (Bounded interference): For all T sufficiently large, there exists a set of natural numbers $\{t_k^j, k = 1, \ldots, \theta_\epsilon T; j \in N\}$ such that, if we let node j's duty slots be this set, then

$$\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} b_j(t_k^j - l)$$
$$+ \sum_{i \neq j, \hat{j}} \rho_{i\hat{j}}^{-2\delta} e^{-2\gamma\rho_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} b_i(t_k^j - \tau_{i\hat{j}} - l)$$
$$\leq \frac{4P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda), \forall j \in N, k = 1, \ldots, \theta_\epsilon T,$$

where λ is defined in (8).

Proof: We exhibit the existence of such a schedule by a probabilistic argument.

Every node $j \in N$ independently generates a sequence of i.i.d. Bernoulli r.v.'s $B_j(t)$, $t \geq 1$, with Pr $\{B_j(t) = 0\} = 1 - 2\theta$, and $\Pr\{B_j(t) = 1\} = 2\theta$. Define $B_j(t) := 0$ for $t \leq 0$. Let $\hat{t}_j(k)$ be the time slot that node j gets the k^{th} 1 in its sequence.

By the strong law of large numbers (SLLN), $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} I_{[B_j(t)=1]} = 2\theta > 2\theta_{\epsilon/2}, \forall j \in N$, a.s. Hence, for all $j \in N$, with high probability, there are at least $2\theta_{\epsilon/2}T$ 1's in $\{B_j(t), t \leq T\}$ when T is large.

Again, by the SLLN, for all $j \in N$, $\lim_T \frac{1}{2\theta_{\epsilon/2}T} \sum_{k=1}^{2\theta_{\epsilon/2}T} \left(B_j(\hat{t}_k^j - l)\alpha^l\right) = 2\theta\alpha^l$, $a.s.$, for all $l \geq 1$; $\lim_T \frac{1}{2\theta_{\epsilon/2}T} \sum_{k=1}^{2\theta_{\epsilon/2}T} \left(B_i(\hat{t}_k^j - l - \tau)\alpha^l\right) = 2\theta\alpha^l$, $a.s.$, for all $l \geq 0, \tau \geq 0$ and $i \neq j$. Since $\left|\frac{1}{2\theta_{\epsilon/2}T}\sum_{k=1}^{2\theta_{\epsilon/2}T} B_i(\hat{t}_k^j - l - \tau)\alpha^l\right| \leq \alpha^l$, $\forall i$, by the dominated convergence theorem,

$$\lim_T \frac{1}{2\theta_{\epsilon/2}T} \sum_{k=1}^{2\theta_{\epsilon/2}T} \left(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_j(\hat{t}_k^j - l)\right.$$
$$\left. + \sum_{i \neq j, \hat{j}} \rho_{i\hat{j}}^{-2\delta} e^{-2\gamma\rho_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_i(\hat{t}_k^j - \tau_{i\hat{j}} - l)\right)$$
$$= \rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} 2\theta$$
$$+ \sum_{i \neq j, \hat{j}} \rho_{i\hat{j}}^{-2\delta} e^{-2\gamma\rho_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} 2\theta$$
$$\leq \rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \frac{2P_\epsilon}{1-\alpha} + \sum_{i \neq j, \hat{j}} e^{-2\gamma\rho_{i\hat{j}}} \rho_{i\hat{j}}^{-2\delta} \beta^2 \frac{2P_\epsilon}{1-\alpha}$$

$$(\text{if } \gamma = 0, \delta > 3) \quad \leq \frac{2P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda), \quad a.s.,$$
$$(\text{if } \gamma > 0, \delta \geq 0) \quad \leq \frac{2P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \rho_{min}^{-2\delta} \sum_{i \neq j, \hat{j}} e^{-2\gamma\rho_{i\hat{j}}})$$
$$\leq \frac{2P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda), \quad a.s.,$$

where the last two inequalities follow from Lemma 6.1 (i) for $\gamma = 0$ with $\delta > 3$, and Lemma 6.1 (iii) for $\gamma > 0$ with $\delta \geq 0$.

Since at least half of the (nonnegative) elements are smaller than or equal to twice the average, for large T we claim that there are $\theta_\epsilon T$ time slots $\{t_k^j, k = 1, \ldots, \theta_\epsilon T\}$ in $\{\hat{t}_k^j, k = 1, 2, \ldots, 2\theta_{\epsilon/2}T\}$ such that

$$\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_j(t_k^j - l)$$
$$+ \sum_{i \neq j, \hat{j}} \rho_{i\hat{j}}^{-2\delta} e^{-2\gamma\rho_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_i(t_k^j - \tau_{i\hat{j}} - l)$$
$$\leq \frac{4P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda).$$

Otherwise,

$$\sum_{k=1}^{2\theta_{\epsilon/2}T} \left(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_j(\hat{t}_k^j - l)\right.$$
$$\left. + \sum_{i \neq j, \hat{j}} \rho_{i\hat{j}}^{-2\delta} e^{-2\gamma\rho_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_i(\hat{t}_k^j - \tau_{i\hat{j}} - l)\right)$$
$$> \frac{4P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda) \cdot (2\theta_{\epsilon/2} - \theta_\epsilon)T,$$

which implies

$$\frac{1}{2\theta_{\epsilon/2}T} \sum_{k=1}^{2\theta_{\epsilon/2}T} \left(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} \beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_j(\hat{t}_k^j - l)\right.$$
$$\left. + \sum_{i \neq j, \hat{j}} \rho_{i\hat{j}}^{-2\delta} e^{-2\gamma\rho_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} B_i(\hat{t}_k^j - \tau_{i\hat{j}} - l)\right)$$
$$> \frac{2P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda)\frac{\theta}{\theta - \epsilon/2}.$$

This is a contradiction. □

Based on this lemma, we first pick a $\bar{T}$ large enough, and then let node j's duty slots in the time interval $\{1, 2, \ldots, \bar{T}\}$ be the corresponding $\{t_k^j, k = 1, \ldots, \theta_\epsilon \bar{T}\}$. Now for a T that is an integral multiple of $\bar{T}$, we periodically repeat the duty slot sequence. That is, for $T = m\bar{T}$, the duty slots are $\{st_k^j : k = 1, \ldots, \theta_\epsilon \bar{T}$ and $s = 1, 2, \ldots, m\}$. In the sequel we need to, and will, only consider T's which are integer multiples of $\bar{T}$, and relabel the duty slots as $\{t_k^j, k = 1, \ldots, \theta_\epsilon T\}$.

The transmission schedule and random phases: Each node $j \in N$ chooses a message W_j uniformly from $\{1, 2, \ldots, 2^{\theta_\epsilon TR}\}$. During j's duty slot t_k^j $(k = 1, 2, \ldots, \theta_\epsilon T)$, it first generates a random phase $\exp(\nu\phi_j(t_k^j))$, where ν is the square root of -1, and $\phi_j(t_k^j) \sim U[0, 2\pi)$. Then it transmits $X_{w_j,k}^j \cdot \exp(\nu\phi_j(t_k^j))$. The random phases $\exp(\nu\phi_j(t))$'s are introduced just to help the decoding by eliminating the possible correlations among signals and fading processes. The receivers need not know the exact values.

X. DECODING BY THRESHOLDING

Upon receiving the complex baseband signal sequence $\{Y_{\hat{j}}(t_k^j), 1 \leq k \leq \theta_\epsilon T\}$, node $\hat{j}$ first passes the sequence

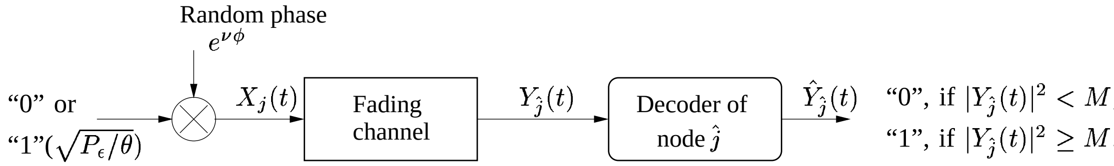

Fig. 4. The communication system

through a simple thresholding filter, as follows:

$$Y_k^{\hat{j}} = \begin{cases} 1, & \text{if } \left|Y_{\hat{j}}(t_k^j)\right|^2 \geq M; \\ 0, & \text{otherwise,} \end{cases} \quad k = 1, 2, \ldots, \theta_\epsilon T,$$

where

$$M := \left(\frac{4\tilde{H}P_\epsilon\beta^2}{1-\alpha}(\rho_{min}^{-2\delta}e^{-2\gamma\rho_{min}} + \lambda) + \sigma^2 \right) / (p^* - \bar{p}). \quad (11)$$

This is shown in Figure 4.

Then, node $\hat{j}$ declares that the index $\hat{W}_j$ was sent if

$$\frac{1}{\theta_\epsilon T}\sum_{k=1}^{\theta_\epsilon T} I_{[X^j_{\hat{W}_j,k}/\sqrt{P_\epsilon/\theta}=Y_k^{\hat{j}}]} \geq \bar{p} - \epsilon_1, \quad (12)$$

and there is no other codeword X_w^j satisfies this.

If no such $\hat{W}_j$ exists or if there is more than one such, then an error is declared. Or even if the energy in the codeword exceeds the prescribed P, an error is declared.

XI. Analysis of the probability of decoding error

Instead of calculating the probability of error for a specific codebook generated according to the procedure in Section VIII, we calculate the average over all such codes, as is standard.

Let $\mathcal{E}_j := \{\hat{W}_j \neq W_j\}$, and $\mathcal{E} := \cup_{j=1}^n \mathcal{E}_j$. Then the *average probability of error over all codebooks* satisfies

$$\Pr(\mathcal{E}) = \sum_C \Pr(C)P_e^{(n)}(C) \leq \sum_C \Pr(C)(\sum_{j=1}^n \Pr(\mathcal{E}_j|C)).$$

If we can show the average probability of decoding error from any node j to its destination $\hat{j}$ goes to zero, then $\Pr(\mathcal{E})$ does so too. To fix notation, say $j = 1$ and its destination is node 2. Furthermore, we can assume that, without loss of generality due to symmetry of codebook construction, the message $W_1 = 1$ was sent from node 1 to node 2. So we only consider $\Pr(\mathcal{E}_1|W_1 = 1)$ in the sequel.

To ease the burden of complex notation, without loss of generality, we assume that the delay $\tau_{12} = 0$, and denote node 1's duty slots as $\{t_k : k = 1, \ldots, \theta_\epsilon T\}$. Now the reception at node 2 can be described as follows:

$$\begin{aligned} Y_2(t_k) = {} & \beta\rho_{12}^{-\delta}e^{-\gamma\rho_{12}}H_{1,2,0}(t_k)X_1(t_k) \\ & +\beta\rho_{12}^{-\delta}e^{-\gamma\rho_{12}}\sum_{l=1}^{\infty}H_{1,2,l}(t_k)X_1(t_k - l) \\ & +\sum_{i\geq 3}\beta\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\sum_{l=0}^{\infty}H_{i,2,l}(t_k)X_i(t_k - \tau_{i2} - l) \\ & +Z_2(t_k), \qquad k = 1, 2, \ldots, \theta_\epsilon T. \end{aligned} \quad (13)$$

Define $A_0 := \{1/T\sum_{k=1}^{\theta_\epsilon T}|X_1(t_k)|^2 > P\}$, and $A_w :=$ {codeword X_w^1 satisfies (12)}, for $w \in \{1, 2, \ldots, 2^{\theta_\epsilon TR}\}$. Then

$$\begin{aligned} \Pr(\mathcal{E}_1|W_1 = 1) &= \Pr(A_0 \cup A_1^c \cup A_2 \cup A_3 \cup \ldots \cup A_{2^{\theta_\epsilon TR}}) \\ &\leq \Pr(A_0) + \Pr(A_1^c) + \sum_{w=2}^{2^{\theta_\epsilon TR}}\Pr(A_w). \end{aligned}$$

By the law of large numbers, we know $\Pr(A_0) \to 0$, as $T \to \infty$.

The following lemma shows that $\Pr(A_1^c) \to 0$, as $T \to \infty$.

Lemma 11.1: $\Pr(A_1^c) \to 0$, as $T \to \infty$.

Proof: It suffices to show that

$$\liminf_T \frac{1}{\theta_\epsilon T}\sum_{k=1}^{\theta_\epsilon T} I_{[X^1_{1,k}/\sqrt{P_\epsilon/\theta}=Y_k^2]} \geq \bar{p} - \epsilon_1/2, \text{a.s.} \quad (14)$$

According to (13), the reception at node 2 can be decomposed as follows:

$$\begin{aligned} & Y_2(t_k) \\ & = Z_2(t_k) + \beta\rho_{12}^{-\delta}e^{-\gamma\rho_{12}}H_{1,2,0}(t_k)X_1(t_k) \\ & \quad +\beta\rho_{12}^{-\delta}e^{-\gamma\rho_{12}}\sum_{l=1}^{t_k-1}H_{1,2,l}(t_k)X_1(t_k - l) \\ & \quad +\beta\rho_{12}^{-\delta}e^{-\gamma\rho_{12}}\sum_{l=t_k}^{\infty}H_{1,2,l}(t_k)X_1(t_k - l) \\ & \quad +\sum_{i\geq 3}\beta\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\sum_{l=0}^{t_k-\tau_{i2}-1}H_{i,2,l}(t_k)X_i(t_k - \tau_{i2} - l) \\ & \quad +\sum_{i\geq 3}\beta\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\sum_{l=t_k-\tau_{i2}}^{\infty}H_{i,2,l}(t_k)X_i(t_k - \tau_{i2} - l) \\ & := Z_2(t_k) + \mu(t_k)X_1(t_k) + \gamma_1'(t_k) + \gamma_1''(t_k) \\ & \qquad +\gamma_2'(t_k) + \gamma_2''(t_k), \quad k = 1, 2, \ldots, \theta_\epsilon T. \end{aligned} \quad (15)$$

Define $\mathcal{F}_k := \sigma\{X_{1,k}^1, X_{1,k-1}^1, \ldots, X_{1,1}^1\}$ for $k \geq 1$, as the σ-algebra generated by the first k digits of the codeword, and $\mathcal{F}_0 := \{\phi, \Omega\}$, and $Q_k := I_{[X^1_{1,k}/\sqrt{P_\epsilon/\theta}=Y_k^2]}$.

We will now show that there exists an integer $K > 0$ such that when $k \geq K$,

$$\Pr(|Y_2(t_k)|^2 < M \,|\, \mathcal{F}_{k-1}, \text{ and } |X_1(t_k)| = 0) \geq \bar{p} - \epsilon_1/2, \quad (16)$$

and

$$\begin{aligned} & \Pr(|Y_2(t_k)|^2 \geq M \,|\, \mathcal{F}_{k-1}, \text{ and } |X_1(t_k)| = \sqrt{P_\epsilon/\theta}) \\ & \qquad \geq \bar{p} - \epsilon_1/2, \end{aligned} \quad (17)$$

where M is defined in (11).

Proof of (16): We have

$$\begin{aligned}
&\Pr\left(|Y_2(t_k)|^2 \geq M | X_1(t_k)=0, \mathcal{F}_{k-1}\right)\\
&\leq \frac{1}{M} E(|Y_2(t_k)|^2 | X_1(t_k)=0, \mathcal{F}_{k-1})\\
&= \frac{1}{M} E(|\gamma_1'(t_k)+\gamma_1''(t_k)+\gamma_2'(t_k)\\
&\qquad +\gamma_2''(t_k)+Z_2(t_k)|^2 \mid \mathcal{F}_{k-1}) \quad \text{(by (15))}\\
&= \frac{1}{M} E(|\gamma_1'(t_k)|^2+|\gamma_2'(t_k)|^2+|\gamma_1''(t_k)\\
&\qquad +\gamma_2''(t_k)|^2+\sigma^2 \mid \mathcal{F}_{k-1}), \qquad (18)
\end{aligned}$$

where the inequality is because of Chebyshev's inequality, and the last equality comes from the fact that $\gamma_1'(t_k), \gamma_2'(t_k), \gamma_1''(t_k)+\gamma_2''(t_k)$ and $Z_2(t_k)$ are uncorrelated by the random phases introduced in the signaling.

Again, by the decorrelation, and recalling the definition of the indicator function $b_i(t)$ in (10),

$$\begin{aligned}
&E(|\gamma_1'(t_k)|^2+|\gamma_2'(t_k)|^2|\mathcal{F}_{k-1})\\
&\leq \beta^2\rho_{12}^{-2\delta}e^{-2\gamma\rho_{12}}\sum_{l=1}^{t_k-1}E|H_{1,2,l}(t_k)|^2\frac{P_\epsilon}{\theta}b_1(t_k-l)\\
&\quad +\sum_{i\geq 3}\beta^2\rho_{i,2}^{-2\delta}e^{-2\gamma\rho_{i2}}\cdot\\
&\qquad \sum_{l=0}^{t_k-\tau_{i2}-1}E|H_{i,2,l}(t_k)|^2\frac{P_\epsilon}{\theta}b_i(t_k-\tau_{i2}-l)\\
&\leq \beta^2\rho_{min}^{-2\delta}e^{-2\gamma\rho_{min}}\sum_{l=1}^{t_k-1}E|H_{1,2,l}(t_k)|^2\alpha^{-l}\alpha^l\frac{P_\epsilon}{\theta}b_1(t_k-l)\\
&\quad +\sum_{i\geq 3}\beta^2\rho_{i,2}^{-2\delta}e^{-2\gamma\rho_{i2}}\cdot\\
&\qquad \sum_{l=0}^{t_k-\tau_{i2}-1}E|H_{i,2,l}(t_k)|^2\alpha^{-l}\alpha^l\frac{P_\epsilon}{\theta}b_i(t_k-\tau_{i2}-l)\\
&\overset{(a)}{\leq} \beta^2\rho_{min}^{-2\delta}e^{-2\gamma\rho_{min}}\sum_{l=1}^{\infty}\tilde{H}\alpha^l\frac{P_\epsilon}{\theta}b_1(t_k-l)\\
&\quad +\sum_{i\geq 3}\beta^2\rho_{i,2}^{-2\delta}e^{-2\gamma\rho_{i2}}\sum_{l=0}^{\infty}\tilde{H}\alpha^l\frac{P_\epsilon}{\theta}b_i(t_k-\tau_{i2}-l)\\
&\leq \frac{4P_\epsilon\beta^2\tilde{H}}{1-\alpha}(\rho_{min}^{-2\delta}e^{-2\gamma\rho_{min}}+\lambda), \qquad (19)
\end{aligned}$$

where (a) holds because $E|H_{ijl}(t_k)|^2\alpha^{-l} \leq \tilde{H}$ by (4), and the last inequality comes from Lemma 9.1.

Since $\mathcal{F}_{k-1}$ is independent of the past signals $X(t)$, $t \leq 0$, we have

$$\begin{aligned}
&E(|\gamma_1''(t_k)+\gamma_2''(t_k)|^2|\mathcal{F}_{k-1})\\
&\quad = E|\gamma_1''(t_k)+\gamma_2''(t_k)|^2 \leq 2E|\gamma_1''(t_k)|^2+2E|\gamma_2''(t_k)|^2.
\end{aligned}$$

Now we show that both $\gamma_1''(t_k)$ and $\gamma_2''(t_k)$ diminish to 0 as $k \to \infty$. Actually,

$$\begin{aligned}
E|\gamma_1''(t_k)|^2 &\leq \beta^2\rho_{min}^{-2\delta}e^{-2\gamma\rho_{min}}\cdot E|\sum_{l=t_k}^{\infty}H_{1,2,l}(t_k)X_1(t_k-l)|^2\\
&\leq \beta^2\rho_{min}^{-2\delta}\cdot E|\sum_{l=t_k}^{\infty}H_{1,2,l}(t_k)\alpha^{-l/2}\cdot\alpha^{l/2}X_1(t_k-l)|^2\\
&\leq \beta^2\rho_{min}^{-2\delta}\cdot\\
&\qquad E\left(\sum_{l=t_k}^{\infty}|H_{1,2,l}(t_k)|^2\alpha^{-l}\right)\left(\sum_{l=t_k}^{\infty}\alpha^l|X_1(t_k-l)|^2\right)\\
&\overset{((2))}{=} \beta^2\rho_{min}^{-2\delta}\cdot\left(\sum_{l=t_k}^{\infty}E|H_{1,2,l}(t_k)|^2\alpha^{-l}\right)\left(\sum_{l=t_k}^{\infty}\alpha^l\bar{P}\right)\\
&\overset{((4))}{\leq} \beta^2\rho_{min}^{-2\delta}\cdot\tilde{H}\cdot\frac{\bar{P}\alpha^{t_k}}{1-\alpha}\\
&= \frac{\tilde{H}\beta^2\rho_{min}^{-2\delta}\bar{P}}{1-\alpha}\cdot\alpha^{t_k}=o(1), \quad \text{as } k\to\infty. \qquad (20)
\end{aligned}$$

For $E|\gamma_2''(t_k)|^2$ we have

$$\begin{aligned}
&E|\gamma_2''(t_k)|^2\\
&= E\left|\sum_{i\geq 3}\sum_{l=t_k-\tau_{i2}}^{\infty}\beta\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}H_{i,2,l}(t_k)X_i(t_k-\tau_{i2}-l)\right|^2\\
&= \beta^2 E\left|\sum_{i\geq 3}\sum_{l=t_k-\tau_{i2}}^{\infty}(\rho_{i,2}^{-\delta/2}e^{-\gamma\rho_{i2}/2}H_{i,2,l}(t_k)\alpha^{-l/2})\cdot\right.\\
&\qquad\qquad \left.(\rho_{i,2}^{-\delta/2}e^{-\gamma\rho_{i2}/2}X_i(t_k-\tau_{i2}-l)\alpha^{l/2})\right|^2\\
&\leq \beta^2 E\left(\sum_{i\geq 3}\sum_{l=t_k-\tau_{i2}}^{\infty}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}|H_{i,2,l}(t_k)|^2\alpha^{-l}\right)\cdot\\
&\qquad \left(\sum_{i\geq 3}\sum_{l=t_k-\tau_{i2}}^{\infty}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}|X_i(t_k-\tau_{i2}-l)|^2\alpha^l\right)\\
&\overset{((2))}{\leq} \beta^2\left(\sum_{i\geq 3}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}E\sum_{l=t_k-\tau_{i2}}^{\infty}|H_{i,2,l}(t_k)|^2\alpha^{-l}\right)\cdot\\
&\qquad \left(\sum_{i\geq 3}\sum_{l=t_k-\tau_{i2}}^{\infty}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\bar{P}\alpha^l\right)\\
&\leq \beta^2\left(\sum_{i\geq 3}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\tilde{H}\right)\left(\sum_{i\geq 3}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\bar{P}\frac{\alpha^{t_k-\tau_{i2}}}{1-\alpha}\right)\\
&\leq \frac{\beta^2\tilde{H}\bar{P}}{1-\alpha}\left(\sum_{i\geq 3}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}}\right)^2\cdot\alpha^{t_k-\max_i\tau_{i2}}. \qquad (21)
\end{aligned}$$

Since $\sum_{i\geq 3}\rho_{i,2}^{-\delta}e^{-\gamma\rho_{i2}} < \infty$, we know $E|\gamma_2''(t_k)|^2 \to 0$, as $k \to \infty$.

Combining (18)–(21), and recalling the definition of M in (11), we obtain

$$\begin{aligned}
&\Pr(|Y_2(t_k)|^2 \geq M \mid X_1(t_k)=0, \mathcal{F}_{k-1})\\
&\qquad \leq p^*-\bar{p}+\epsilon_1/2 \leq 1-\bar{p}+\epsilon_1/2,
\end{aligned}$$

when k is appropriately large, depending on $\epsilon_1/2$. This proves (16).

Proof of (17): Recall that what we need to show is that there exists a $K > 0$ such that when $k \geq K$, $\Pr(|Y_2(t_k)|^2 \geq$

$M \,|\, \mathcal{F}_{k-1}$, and $|X_1(t_k)| = \sqrt{P_\epsilon/\theta}) \geq \bar{p} - \epsilon_1/2$, where M is defined in (11).

Let event D_k be the event $\{|H_{1,2,0}(t_k)| \geq a, |\gamma_1''(t_k) + \gamma_2''(t_k)| \leq \sqrt{M}\epsilon'\}$ for some $\epsilon' > 0$. First we show that

$$\Pr(D_k|\mathcal{F}_{k-1}) = \Pr(D_k) \geq p^* + o(1), \text{ as } k \to \infty, \quad (22)$$

where p^* is as stated in Theorem 3.2.

By assumption, $\Pr(|H_{1,2,0}| \geq a) \geq p^*$. Furthermore, from (20) and (21), as $k \to \infty$,

$$\begin{aligned}
&\Pr(|\gamma_1''(t_k) + \gamma_2''(t_k)| > \sqrt{M}\epsilon') \\
&\leq \frac{1}{M\epsilon'^2} E|\gamma_1''(t_k) + \gamma_2''(t_k)|^2 \\
&\leq \frac{2}{M\epsilon'^2} \left(E|\gamma_1''(t_k)|^2 + E|\gamma_2''(t_k)|^2\right) = o(1).
\end{aligned}$$

Hence (22) follows.

Now,

$$\begin{aligned}
&\Pr(|Y_2(t_k)|^2 \geq M \mid |X_1(t_k)| = \sqrt{P_\epsilon/\theta}, \mathcal{F}_{k-1}) \\
\geq\ & \Pr(|Y_2(t_k)|^2 \geq M, \text{and } D_k \,|\, |X_1(t_k)| = \sqrt{P_\epsilon/\theta}, \mathcal{F}_{k-1}) \\
=\ & \Pr\Big(|\beta\rho_{12}^{-\delta} e^{-\gamma\rho_{12}} H_{1,2,0}(t_k)\sqrt{P_\epsilon/\theta}\exp(\nu\phi_1(t_k)) + \\
& \quad \gamma_1'(t_k) + \gamma_1''(t_k) + \gamma_2'(t_k) + \gamma_2''(t_k) + Z_2(t_k)| \\
& \quad \geq \sqrt{M}, \text{ and } D_k \,\Big|\, \mathcal{F}_{k-1}\Big) \\
\geq\ & \Pr(|\beta(\zeta\rho_{min})^{-\delta} e^{-\gamma\zeta\rho_{min}} H_{1,2,0}(t_k)\sqrt{P_\epsilon/\theta}| \geq 2\sqrt{M}, \\
& \quad |\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)| \leq \sqrt{M}(1-\epsilon'), \\
& \quad \text{and } D_k \mid \mathcal{F}_{k-1}) \\
=\ & \Pr(D_k|\mathcal{F}_{k-1}) \cdot \\
& \Pr(|\beta(\zeta\rho_{min})^{-\delta} e^{-\gamma\zeta\rho_{min}} H_{1,2,0}(t_k)\sqrt{P_\epsilon/\theta}|^2 \geq 4M, \\
& |\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)| \leq \sqrt{M}(1-\epsilon') \,|\, D_k, \mathcal{F}_{k-1}) \\
\overset{(a)}{=}\ & \Pr(D_k) \cdot \Pr(|\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)| \\
& \quad \leq \sqrt{M}(1-\epsilon') \,|\, D_k, \mathcal{F}_{k-1}) \\
\overset{(b)}{\geq}\ & \Pr(D_k) \cdot (1 - \frac{1}{M(1-\epsilon')^2} \cdot \\
& \quad E(|\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)|^2 \,|\, D_k, \mathcal{F}_{k-1})\ ) \\
\geq\ & \Pr(D_k) \cdot \Bigg(1 - \frac{1}{M(1-\epsilon')^2} \cdot \\
& \quad \frac{E(|\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)|^2 \,|\, \mathcal{F}_{k-1})}{\Pr(D_k)}\Bigg) \\
=\ & \Pr(D_k) - \frac{E(|\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)|^2 \,|\, \mathcal{F}_{k-1})}{M(1-\epsilon')^2},
\end{aligned}$$

where (a) is because $|H_{1,2,0}(t_k)| \geq a$, and since the definitions of θ and M ((9), (11)) imply that

$$\begin{aligned}
&\left|\beta(\zeta\rho_{min})^{-\delta} e^{-\gamma\zeta\rho_{min}} H_{1,2,0}(t_k)\sqrt{P_\epsilon/\theta}\right|^2 \\
&\geq \beta^2(\zeta\rho_{min})^{-2\delta} e^{-2\gamma\zeta\rho_{min}} a^2 P_\epsilon/\theta \\
&\geq \beta^2(\zeta\rho_{min})^{-2\delta} e^{-2\gamma\zeta\rho_{min}} a^2 P_\epsilon \cdot \\
&\quad \frac{1}{P_\epsilon\beta^2(\zeta\rho_{min})^{-2\delta} e^{-2\gamma\zeta\rho_{min}} a^2/4M} = 4M,
\end{aligned}$$

while (b) comes from Chebyshev's inequality.

From (18) and (19),

$$\begin{aligned}
&E(|\gamma_1'(t_k) + \gamma_2'(t_k) + Z_2(t_k)|^2|\mathcal{F}_{k-1}) \\
&\leq \frac{4P_\epsilon\beta^2\tilde{H}}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda) + \sigma^2.
\end{aligned}$$

So,

$$\begin{aligned}
&\Pr(|Y_2(t_k)|^2 \geq M||X_1(t_k)| = \sqrt{P_\epsilon/\theta}, \mathcal{F}_{k-1}) \\
\geq\ & \Pr(D_k) - \frac{\frac{4P_\epsilon\beta^2\tilde{H}}{1-\alpha}(\rho_{min}^{-2\delta} e^{-2\gamma\rho_{min}} + \lambda) + \sigma^2}{M(1-\epsilon')^2} \\
=\ & p^* + o(1) - \frac{M(p^* - \bar{p})}{M(1-\epsilon')^2} \\
=\ & p^* + o(1) - \frac{1}{(1-\epsilon')^2} \cdot (p^* - \bar{p}), \text{ as } k \to \infty.
\end{aligned}$$

By selecting ϵ' sufficiently small,

$$\begin{aligned}
&\Pr(|Y_2(t_k)|^2 \geq M \,|\, |X_1(t_k)| = \sqrt{P_\epsilon/\theta}, \mathcal{F}_{k-1}) \\
&\geq \bar{p} - \epsilon_1/2,
\end{aligned}$$

for k large enough. This proves (17).

Based on (16) and (17), for $k \geq K$, we have

$$\begin{aligned}
&E[Q_k \,|\, \mathcal{F}_{k-1}] \\
=\ & E\left[I_{[X_{i,k}^1/\sqrt{P_\epsilon/\theta} = Y_k^2]} \,|\, \mathcal{F}_{k-1}\right] \\
=\ & E\left[I_{[|X_1(t_k)|=0]} I_{[X_{i,k}^1/\sqrt{P_\epsilon/\theta} = Y_k^2]} \,|\, \mathcal{F}_{k-1}\right] \\
& + E\left[I_{\left[|X_1(t_k)| = \sqrt{P_\epsilon/\theta}\right]} I_{\left[X_{i,k}^1/\sqrt{P_\epsilon/\theta} = Y_k^2\right]} \,|\, \mathcal{F}_{k-1}\right] \\
=\ & \frac{1}{2}\Pr(|Y_2(t_k)|^2 < M \,|\, \mathcal{F}_{k-1}, |X_1(t_k)| = 0) \\
& + \frac{1}{2}\Pr(|Y_2(t_k)|^2 \geq M \,|\, \mathcal{F}_{k-1}, |X_1(t_k)| = \sqrt{P_\epsilon/\theta}) \\
\geq\ & \bar{p} - \epsilon_1/2, \quad \text{a.s.}
\end{aligned}$$

Note here that until now we have considered the situation for fixed T, i.e., all the variables X, Y, Q, etc., depend on T. Recall now the coupled generation of the codebooks for different T's in Section VIII, by which the codebooks for a greater T are the extensions of those for smaller ones. Based on this observation, we know $\mathcal{F}_{k-1}$ and Q_k are well defined for $k \geq 1$.

Define $\bar{Q}_k := Q_k - E[Q_k|\mathcal{F}_{k-1}]$ and $S_k := \sum_{l=1}^k \bar{Q}_l$. Since $E[\bar{Q}_k|\mathcal{F}_{k-1}] = 0$, $\{\bar{Q}_k\}$ are martingale differences. Furthermore, $\{S_k\}$ is an L^2 martingale, i.e., $E|S_k|^2 < \infty$.

Since $\sum_{k=1}^\infty \frac{E|\bar{Q}_k|^2}{k^2} \leq \sum_{k=1}^\infty \frac{4}{k^2} < \infty$, by martingale theory (p. 397 [31]), we know that $S_{\theta_\epsilon T}/(\theta_\epsilon T) \to 0$, a.s. This implies

$$\liminf_T \frac{1}{\theta_\epsilon T}\sum_{k=1}^{\theta_\epsilon T} Q_k \geq \bar{p} - \epsilon_1/2,$$

which is (14).

If, furthermore, we can show that for any $w \geq 2$,

$$\Pr(A_w) \quad \leq \quad 2^{-\theta_\epsilon T(1-H(\bar{p}-\epsilon_1))}, \quad (23)$$

then $\Pr(\mathcal{E}_1|W_1 = 1) \leq o(1) + 2^{\theta_\epsilon TR - \theta_\epsilon T(1-H(\bar{p}-\epsilon_1))} = o(1) + 2^{-\theta_\epsilon T\epsilon_1} = o(1)$, as $T \to \infty$.

By symmetry we only need to show (23) for the case $w = 2$.

For any deterministic 0, 1 sequence $y = (y_1, \ldots, y_{\theta_\epsilon T})$, we know from the Chernoff bound (p.11, [32]) that $\Pr(\sum_{k=1}^{\theta_\epsilon T} I_{[X^1_{2,k}/\sqrt{P_\epsilon/\theta} = y_k]} \geq \theta_\epsilon T(\bar{p} - \epsilon_1)) \leq 2^{-\theta_\epsilon T(1-H(\bar{p}-\epsilon_1))}$. So, letting $Y := (Y_1^2, Y_2^2, \ldots, Y^2_{\theta_\epsilon T})$, we have

$$\begin{aligned}
\Pr(A_1) &= \Pr\left(\sum_{k=1}^{\theta_\epsilon T} I_{[X^1_{2,k}/\sqrt{P_\epsilon/\theta} = Y_k^2]} \geq \theta_\epsilon T(\bar{p} - \epsilon_1)\right) \\
&= \sum_y \Pr(Y = y) \cdot \\
&\qquad \Pr\left(\sum_{k=1}^{\theta_\epsilon T} I_{[X^1_{2,k}/\sqrt{P_\epsilon/\theta} = y_k]} \geq \theta_\epsilon T(\bar{p} - \epsilon_1)\right) \\
&\leq \sum_y \Pr(Y = y) 2^{-\theta_\epsilon T(1-H(\bar{p}-\epsilon_1))} \\
&= 2^{-\theta_\epsilon T(1-H(\bar{p}-\epsilon_1))}.
\end{aligned}$$

This proves Theorem 3.2.

XII. Concluding remarks

In this paper we have examined the effect of fading on wireless networks, studying in particular how the transport capacity grows with the number of nodes, in networks where nodes are separated by a minimum positive distance. We have restricted attention to the case where there is absorption, the generally prevalent case, or the path-loss exponent δ is larger than 3. When the nodes are subject to power constraints, we have shown that the transport capacity can grow no faster than linearly in the size of the network even if the fading process is known non-causally. Thus, the upper bound holds no matter what the fading environment is, and is thus a best case result. On the other hand, if we consider the opposite scenario where the fading is independent from time to time, a sort of worst case scenario, then one achieves linear growth in any network where every node has another node located within a fixed multiple of the positive minimum distance. In the constructions, communications are only between neighboring nodes, and only point-to-point coding is used. This supports the case for the use of the multi-hop strategy in wireless ad hoc networks, a strategy which is currently the target of much protocol development activity. When $\gamma = 0$ and $\delta \leq 3$, the behavior of wireless networks can be quite different, as shown in [19], through the exploitation of coherence. Whether such behavior can hold under fading, where coherence is not achievable, is an open question. Also of interest is how to exploit multiple antennas at nodes, when such are present in wireless networks.

Acknowledgments

The authors are grateful to Professor Venu Veeravalli for discussions on fading channels.

References

[1] "Mobile Ad-hoc Networks (manet). IETF Secretariat," October 10 2003. http://www.ietf.org/html.charters/manet-charter.html.

[2] "The Fourth ACM International Symposium on Mobile Ad Hoc Networking and Computing Mobihoc 2003," June 1–3 2003. Annapolis, MD, USA. http://www.sigmobile.org/mobihoc/2003/.

[3] "ACM SenSys '03," Nov. 5-7 2003. Los Angeles, CA, USA. http://www.cens.ucla.edu/sensys03/.

[4] Z. J. Haas and J. Deng, "Dual busy tone multiple access (DBMTA) – performance evaluation," in *Proceedings of the IEEE VTC '99*, (Houston, Texas), May 17–21 1999.

[5] R. Rozovsky and P. R. Kumar, "SEEDEX: A MAC protocol for ad hoc networks," in *Proceedings of the 2001 ACM International Symposium on Mobile Ad Hoc Networking and Computing*, (Long Beach, CA), pp. 67–75, Oct 4-5 2001.

[6] S. Narayanaswamy, V. Kawadia, R. S. Sreenivas, and P. R. Kumar, "Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol," in *European Wireless Conference – Next Generation Wireless Networks: Technologies, Protocols, Services and Applications*, (Florence, Italy), pp. 156–162, Feb 25–28 2002.

[7] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *Computer Communications Review*, vol. 24, no. 4, pp. 234–244, 1994. (ACM SIGCOMM '94).

[8] C. Perkins, E. Royer, and S. Das, "Ad hoc on demand distance vector (AODV) routing." http://www.ietf.org/internet-drafts/draft-ietf-manet-aodv-03.txt. IETF Internet Draft (work in progress), June 1999.

[9] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing* (T. Imielinski and H. Korth, eds.), Kluwer Academic Publishers, 1996.

[10] T. Clausen, P. Jacquet, A. Laouiti, P. Muhlethaler, and A. Qayyum et L. Viennot, "Optimized link state routing protocol." IEEE INMIC, 2001. Pakistan.

[11] R. Ogier and M. Lewis, "Topology broadcast based on reverse-path forwarding (TBRPF)." SRI, 2003. http://www.erg.sri.com/projects/tbrpf/.

[12] P. Gupta and P. R. Kumar, "A system and traffic dependent adaptive routing algorithm for ad hoc networks," in *Proceedings of the 36th IEEE Conference on Decision and Control*, (San Diego, CA), pp. 2375–2380, December 1997.

[13] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge University Press, 1998.

[14] R. Ahlswede, "Multi-way communication channels," in *Proceedings of the 2nd Int. Symp. Inform. Theory (Tsahkadsor, Armenian S.S.R.)*, (Prague), pp. 23–52, Publishing House of the Hungarian Academy of Sciences, 1971.

[15] H. Liao, *Multiple access channels*. PhD thesis, University of Hawaii, Honolulu, HA, 1972. Department of Electrical Engineering.

[16] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proceedings of the 2000 IEEE Int. Symp. Inform. Theory*, (Sorrento, Italy), p. 22, 2000.

[17] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, 1979.

[18] W. Ren and P. R. Kumar, "Stochastic parallel model adaptation: Theory and applications to active noise cancelling, feedforward control, IIR filtering and identification," *IEEE Transactions on Automatic Control*, vol. AC-37, pp. 269–307, 1992.

[19] L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation." *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, 2004.

[20] J. G. Proakis, *Digital Communications*. McGraw-Hill, fourth ed., 2001.

[21] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[22] V. V. Veeravalli, *Wideband Multi-Antenna Wireless Channels: Statistical Modeling, Analysis, and Simulation*. IEEE VTC 2001 Tutorial, October 7 2001. Tutorial #7.

[23] R. Knopp and P. Humblet, "Information capacity and power control in single cell multiuser communications," in *IEEE International Communication Conference*, (Seattle, WA), June 1995.

[24] S.R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," in *Proc. 2002 IEEE International Symposium on Information Theory*, Lausanne, Switzerland, June 30–July 5, 2002, pp. 351.

[25] I. E. Telatar and D. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1384–1400, 2002.

[26] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, pp. 1277–1294, June 2002.

[27] G. Kramer, M. Gastpar, and P. Gupta, "Capacity theorems for wireless relay channels," 41^{st} *Allerton Conference on Communication, Control and Computing.* October 1-3, 2003, Monticello, IL.

[28] M. Franceschetti, J. Bruck, and L.J. Schulman, "A random walk model of wave propagation," *IEEE Transactions on Antennas Propagation*, 52(5), pp. 1304 - 1317, May 2004

[29] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley and Sons, 1991.

[30] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov.-Dec. 1999.

[31] Y. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*. New York, NY: Springer-Verlag, 1988.

[32] B. Bollobas, *Random Graphs*. Cambridge University Press, second ed., 2001.

Feng Xue (S'02) received the B.S. degree from the Department of Mathematics, Shandong University, Jinan, China, in 1997, and the M.S. degree from the Institute of Systems Science, the Chinese Academy of Sciences, Beijing, China, in 2000.

He is currently working toward the Ph.D. degree in the Department of Electical and Computer Engineering, University of Illinois Urbana-Champaign. His research interests are in the fields of wireless networks, sensor networks, and information theory.

Liang-Liang Xie (M'03) was born in Hunan, China, in 1974. He received the B.S. degree in mathematics from Shandong University, and the Ph.D. degree in control theory from the Chinese Academy of Sciences, in 1995 and 1999, respectively.

He joined the faculty of Institute of Systems Science, Chinese Academy of Sciences in 1999. He was a Guest Researcher in the Automatic Control Group, Linkping University, Sweden during 1999-2000, and a Postdoctoral Research Associate in the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, USA during 2000-2002. His research interests include network information theory, wireless networks, adaptive control and system identification.

P R Kumar (F' 88), born 1952, obtained his B. Tech. degree from IIT Madras in 1973, and M.S. and D.Sc. degrees from Washington University, St. Louis in 1975 and 1977, respectively. From 1977-84 he was with the Department of Mathematics at the University of Maryland Baltimore County, and since 1985 he has been at the University of Illinois Urbana-Champaign, where is currently Franklin Woeltge Professor of Electrical and Computer Engineering. He was the recipient of the Donald P. Eckman Award of the American Automatic Control Council in 1985. His current research interests are in the areas of wireless networks, sensor networks, and the convergence of control with communication and computing. Within this he is working on problems in network information theory, architecture, middleware, protocol development, performance evaluation, software development, implementation, experimentation, and application development.