Decentralized Optimization for Multichannel Random Access*

Guowang Miao † , *Student Member, IEEE*, Geoffrey Ye Li, *Fellow, IEEE*, and Ananthram Swami, *Fellow, IEEE*

***Abstract*—We consider schemes for decentralized cross-layer optimization of multichannel random access by exploiting local channel state and traffic information. In the network we are considering, users are not necessarily within the transmission ranges of all others; therefore, when a user is transmitting, it may only interfere with some users, which is different from most existing channel aware Aloha schemes. Besides, we also consider complicated traffic distribution, e.g. each user may choose to send packets to or receive packets from different users simultaneously. We develop *decentralized optimization for multichannel random access* (DOMRA). DOMRA consists of three steps: neighborhood information collection, transmission control of the MAC layer based on the instantaneous channel state information, and power allocation for each traffic flow on each subchannel. Simulation results demonstrated that DOMRA significantly outperforms existing channel aware Aloha schemes due to its exploitation of both multiuser diversity through cross-layer design and the inhomogeneous characteristics of traffic spatial distribution in the network. Besides, DOMRA performs closely to the globally optimum solution, which requires full network knowledge to be obtained. DOMRA can be applied to different types of wireless networks, such as wireless sensor networks and mobile *ad hoc* networks, to improve quality of service.**

***Index Terms*—cross-layer, decentralized, random access, multichannel, channel aware**

I. INTRODUCTION

In traditional networks, especially wired networks, different layers are treated separately to simplify the design. However, in wireless networks, channel states are different for different users and change with time. Network efficiency can be improved significantly if layer-wise design is substituted by cross-layer design. Therefore, cross-layer optimization is becoming a more and more important research area in wireless communications [1]–[3]. Much of the research in this area focuses on multiuser wireless communication environments, where each user is scheduled to transmit when it has favorable channel conditions so as to obtain multiuser diversity [4].

For centralized scheduling [5]–[7], the best performance can be obtained with the help of *channel state information* (CSI) from all active users. However, CSI feedback consumes a large amount of resources, especially for networks with a large number of users at high mobility. To reduce CSI feedback, decentralized cross-layer design approaches can be considered. Opportunistic random access schemes have been studied in [8]–[15] and the references therein. For opportunistic random access, each user exploits its own CSI to optimize transmission performance. In [8], each user transmits only if its channel power gain is above a pre-determined threshold that is chosen to maximize the probability of successful transmissions. For clustered *orthogonal frequency-division multiplexing* (OFDM) based wireless networks, an opportunistic multichannel Aloha has been designed in [9], in which a user transmits if it has at least some subchannels with gains greater than a threshold. In [10], some negotiations between the users and the base station are designed prior to data transmission to allow successful transmission of the user with best channel condition. A channel aware multicarrier random access scheme has been proposed in [11], where each user selects some subchannels with the best channel power gains for data transmission. Inspired by [8], it is proposed in [12] that each user in a cellular network sends request packets when the channel fading level exceeds a predetermined threshold, after which the BS processes downlink transmissions. Although the thresholds in [12] are chosen to optimize downlink throughput, the proposed scheme actually reduces uplink request collisions, and hence also deals with random access. In [13], based on decentralized CSI, a general expression for the transmission probability that may depend on the channel and the physical layer implementation is given, and the transmission probability is optimized to achieve maximum stable throughput in the *medium access control* (MAC) layer. Through slotted Aloha, a reservation-based MAC scheme is found in [14] to maximize the overall throughput. The capacity of slotted Aloha is analyzed and the optimal transmission probabilities is obtained for multi-packet reception MAC model [15]. All these works are for wireless networks where users transmit to a common receiver, e.g., the base station.

This scenarios does not fit many wireless communication environments, such as sensor networks [16], *ad hoc* networks [17], or wireless mesh networks [18]. In sensor networks [16], each user sends packets to oth-

* This work was supported by the research gift of Intel, the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-20-0011, and SABA Program of Motorola.

* Part of this work has been published at Globecom 2007

† Corresponding author. Phone: (404) 894-9111. Fax: (404) 894-7883. E-Mail: gmiao3@gatech.edu. School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250

TABLE I
NOTATIONS IN THE PAPER

K	number of subchannels	k	subchannel index
i, j	user indices	$(i, j)_k$	link from Users i to j on subchannel k
(i, j)	link from Users i to j	$\mathcal{V}$	set of active users, $\{1, 2, ..., N\}$
$\mathcal{E}$	set of links over all subchannels, $\{(i, j)_k \| i, j \in \mathcal{V}, i \neq j; k = 1, 2, ..., K\}$	$\mathcal{L}$	set of links available for communication, $\{(i, j) \| i, j \in \mathcal{V}, i \neq j$, i and j can receive packets from each other$\}$
G	directed graph denoting the network, defined as $(\mathcal{V}, \mathcal{E}, \mathcal{L})$	$\mathcal{N}_i$	interfering neighbor set of i, $\{j \| \forall j \in \mathcal{V}, s.t. (i, j) \in \mathcal{L}\}$
$\mathcal{T}_i$	set of users receiving packets from i, $\{j \| (i, j) \in \mathcal{L}$, (i, j) is backlogged $\}$	$\mathcal{S}_j$	set of users sending packets to j, $\{i \| (i, j) \in \mathcal{L}, (i, j)$ is backlogged$\}$
$\mathbf{E}$	statistical average	P_t	transmit power
P_a	average power constraint	P_m	maximum allowed transmit power
h	channel power gain	$f(h)$	probability density of h
$F(h)$	distribution of h	$\overline{H}_{(i,j)_k}$	predetermined channel power gain threshold of link $(i, j)_k$
$\|\mathcal{X}\|$	number of elements in set $\mathcal{X}$	T	throughput
p_{i_k}	transmission probability of User i on subchannel k	$p_{(i,j)_k}$	transmission probability of link $(i, j)_k$
$R_{(i,j)_k}$	average data rate on link $(i, j)_k$	$P_{(i,j)_k}(h)$	transmit power on link $(i, j)_k$ with channel power gain h
$P_{r(i,j)_k}$	received power level of link $(i, j)_k$	$\mathcal{P}_{ri}$	set of received power level of User i, defined as $\{P_{r(i,j)_k} \| j \in \mathcal{T}_i, k = 1, ..., K\}$
W	total system bandwidth		
$R(\eta)$	data rate function for a given SNR η	$\eta(h)$	received SNR given h
		A^*	the optimal solution of A
N_o	noise spectral density	$\overline{\mathcal{H}}$	set of predetermined channel power gain threshold, defined as $\{\overline{H}_{(i,j)_k} \| (i, j)_k \in \mathcal{E}, j \in \mathcal{T}_i\}$
$\mathcal{P}$	power allocation policy, defined as $\{P_{(i,j)_k}(h) \| (i, j)_k \in \mathcal{E}, j \in \mathcal{T}_i$, $h > \overline{H}_{(i,j)_k}\}$. $\mathcal{P}_i$ is the power allocation policy of User i.	$\mathcal{C}$	transmission control of the network, defined as $\{\overline{\mathcal{H}}, \mathcal{P}\}$

ers with minimized energy and only interferes with its neighboring users. Hence, different users have different sets of channel access competitors. Besides, many existing policies [8]–[13] are designed such that each user has the same transmission probability. Although this guarantees absolute fairness among all users, the network performance is not optimal when the traffic flows are not uniformly distributed or users are not necessarily within the transmission range of the others. Based on inhomogeneous characteristics of channel, interference, and traffic of different users, we propose a *decentralized optimization for multichannel random access* (DOMRA) scheme. The novel scheme consists of three steps: (1) neighborhood information collection; (2) transmission control of the MAC layer based on the instantaneous channel state information; (3) power allocation of each traffic flow on each subchannel. Simulation results show that the proposed scheme significantly outperforms existing random access schemes due to the exploitation of both multiuser diversity through cross-layer design and the inhomogeneous characteristics of traffic spatial distribution in the network. Besides, DOMRA performs closely to the globally optimum solution, which requires full network knowledge to be obtained.

The rest of this paper is organized as follows. First we introduce the physical and MAC layers of the system in Section II. In Section III, we describe the transmission policy and formulate the problem. Then in Section IV, we decompose the cross-layer optimization into two subproblems and provide suboptimal solutions. Finally, we demonstrate the performance improvement of the proposed scheme by computer simulations in Section V and conclude the paper in Section VI.

The notations in this paper are summarized in Table I.

II. SYSTEM DESCRIPTION

Consider multichannel wireless networks. The whole band is divided into K subchannels. All channels between pairs of users are assumed to be reciprocal, i.e. when no interference exists, User A can receive signal from User B if and only if User B can receive signal from User A with the same channel gain. However, the interference environments at Users A and B may be different since they are at different locations. Each user has knowledge of its own CSI and makes independent transmission control decisions, including whether to transmit given the CSI, what data rate to use and where to transmit, etc. Each user applies the same transmission control policy. In order to avoid onerous signalling burden, no communication pair has instantaneous cooperation, such as exchange of CSI, transmit power, or subchannel selections.

All users are not necessarily within the transmission

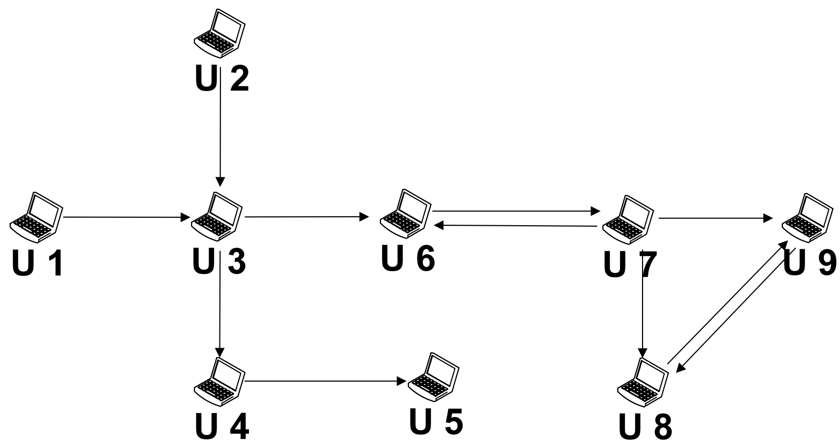

Fig. 1. Network architecture example

ranges of the others, which means that some users may not be able to receive packets from others due to weak received signal power. For simplicity, we assume those that can communicate with each other experience isotropic channels, i.e. channel power gains of different links are independent and identically distributed with probability density function, $f(h)$, and distribution function, $F(h)$. No capture is assumed for signal reception, i.e., the receiver cannot receive any signal successfully if any of its interfering neighbors, which are within the transmission range of the receiver, is transmitting simultaneously. A user can not transmit and receive simultaneously on the same subchannel; however, it may transmit on a set of subchannels and receive on a different set of subchannels at the same time. Each user may choose to send packets to or receive packets from different users on different channels, and we assume that the links that carry traffic are backlogged, i.e., they always have packets to transmit.

During transmission, each user is subject to both average and instantaneous power constraints [19]. The average power constraint is due to heat accumulation and overall power consumption, while the instantaneous power constraint comes from the limited linear range of amplifiers. Two power allocation policies will be considered. In the first one, called channel inversion, each user transmits with just sufficient power to keep the received power constant so that the signal can be reliably detected. In the second, called adaptive modulation and power allocation, each user can vary both the modulation and transmit power during each transmission time slot to maximize throughput.

III. Problem Formulation

In this section, we describe our wireless network model, and propose a channel aware multi-channel random access scheme. The characteristics of the proposed scheme are analyzed, after which a criterion for cross-layer design is provided.

Denote the wireless network as a directed graph $G(\mathcal{V}, \mathcal{E}, \mathcal{L})$, where $\mathcal{V}$, $\mathcal{E}$, and $\mathcal{L}$ are the set of active users, the set of all links over all K subchannels, and the set of links available for communication (for detailed definition, please refer to Table I. We denote $\mathcal{N}_i$ as the interfering neighbor set of User i. Each user may choose to send packets to or receive packets from several users, and $\mathcal{T}_i$ denotes the set of users receiving packets from i and $\mathcal{S}_j$ the set of users sending packets to j.

Figure 1 shows an example topology of a wireless network. The users are on a grid with unit spacing, and the transmission range is $\sqrt{2}$. The set of links available for communication is $\mathcal{L} = \{(1,3), (1,2), (1,4), (2,3), (2,6), (3,4), (3,5), (3,6), (4,6), (4,5), (5,6), (5,7), (5,8), (6,7), (6,8), (7,8), (7,9), (8,9)\}$. The arrows show the traffic flows in the network. For example, since $(4,6) \in \mathcal{L}$, any transmission by Users 4 or 6 will be received by the other though they may not have packets to send to each other. So Users 4 and 6 constitute an interfering pair and they interferer packet reception of each other. Observing User 3, it is easy to see that $\mathcal{T}_3 = \{4,6\}$, $\mathcal{S}_3 = \{1,2\}$, while $\mathcal{N}_3 = \{1,2,4,5,6\}$.

Slotted Aloha is a typical random access scheme. In slotted Aloha, the MAC layer makes transmission decisions based on the buffer occupancy and QoS requirement, and does not utilize the knowledge of the physical layer at all. Hence, when the MAC decides to transmit a frame, the channel may be in deep fade, but the physical layer still carries out the transmission, and causes a waste of power. On the other hand, the MAC layer may decide not to transmit even though the channel power gain is high, because it does not have this information from the PHY layer; this leads to wasted opportunity. With channel knowledge, the sender will transmit only when the channel power gain is above a certain threshold *. Therefore, we propose the following *decentralized optimization for multichannel random access* (DOMRA).

*Channel gains may be inferred either due to CSI feedback or via channel reciprocity.

DOMRA: *User i ($i \in \mathcal{V}$) decides to send packets to User j on subchannel k when the following conditions are satisfied: 1). User i has packets to send to j, $j \in \mathcal{T}_i$; 2). on subchannel k, link (i,j) has the best channel power gain, $h_{(i,j)_k} = \max_{l\in\mathcal{T}_i}\{h_{(i,l)_k}\}$; 3). the channel power gain is above a threshold, $h_{(i,j)_k} \geq \overline{H}_{(i,j)_k}$, where $\overline{H}_{(i,j)_k}$ is predetermined for link $(i,j)_k$. The transmission is then optimized according to $\overline{H}_{(i,j)_k}$, CSI and capability constraints.*

In DOMRA, each user transmits on the link with the best channel power gain provided that the gain is above a predetermined threshold. Proper choice of thresholds $\{\overline{H}_{(i,j)_k}|(i,j)_k \in \mathcal{E}\}$ and data transmission rates of all traffic flows, i.e. power allocation, will be determined in the following paragraphs so that overall network performance is optimized from certain perspectives.

As pointed out before, while many existing channel aware schemes such as [8]–[11], assume that each user has only one traffic flow to send and is within the transmission range of all other users, the DOMRA will provide solutions to networks in which users are not necessarily within the transmission ranges of all other users, and each user could send packets to or receive packets from different users simultaneously on different subchannels.

A. MAC Layer Analysis

According to the above transmission policy and the homogeneity assumption, the probability of a transmission on link $(i,j)_k \in \mathcal{E}$ is given by

$$p_{(i,j)_k} = \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right) \tag{1}$$

where $|\cdot|$ denotes the number of elements in the respective set. The proof of (1) is given in Appendix I.

The probability that User i transmits on subchannel k is

$$p_{i_k} = \sum_{j\in\mathcal{T}_i} p_{(i,j)_k} = \sum_{j\in\mathcal{T}_i}\frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right). \tag{2}$$

Hence, the throughput on link $(i,j)_k$ is

$$T_{(i,j)_k} = R_{(i,j)_k} p_{(i,j)_k}(1-p_{j_k})\prod_{a\in\mathcal{N}_j, a\neq i}(1-p_{a_k}), \tag{3}$$

where $R_{(i,j)_k}$ is the average data rate given that the user has decided to transmit on link $(i,j)_k$, and depends on the modulation and power allocation policy. $(1-p_{j_k})\prod_{a\in\mathcal{N}_j,a\neq i}(1-p_{a_k})$ is the probability that neither user j nor its neighboring users except user i will transmit on subchannel k, which means successful transmission on link $(i,j)_k$.

For example, in Figure 1, the transmission from User 3 to User 6 on subchannel k succeeds only when neither User 6 nor his neighbors excluding User 3, i.e., users in $\mathcal{N}_6\backslash\{3\} = \{2,4,5,7,8\}$, transmit. Hence, the throughput from User 3 to 6 on subchannel k is $T_{(3,6)_k} = p_{(3,6)_k}(1-p_{6_k})(1-p_{2_k})(1-p_{4_k})(1-p_{5_k})(1-p_{7_k})(1-p_{8_k})R_{(i,j)_k}$.

B. Physical Layer Analysis

The average transmit power on link $(i,j)_k$ is the average of transmit power over all time slots, whether or not transmission happens on this link. According to the ergodicity of the channel, it is the average of transmit power over all channel states. Hence, we have

$$\begin{aligned}
\mathbf{E}\{P_{(i,j)_k}\} = &\int_0^\infty \Pr\{H_{(i,j)_k} = h, \text{and User } i \text{ transmits} \\
&\text{on } (i,j)_k\}P_{(i,j)_k}(h)\mathrm{d}h \\
=&\int_0^\infty \Pr\{H_{(i,j)_k} = h\}\Pr\{\text{ User } i \text{ transmits} \\
&\text{on } (i,j)_k|H_{(i,j)_k} = h\}P_{(i,j)_k}(h)\mathrm{d}h \\
=&\int_{\overline{H}_{(i,j)_k}}^\infty f(h)F^{|\mathcal{T}_i|-1}(h)P_{(i,j)_k}(h)\mathrm{d}h \\
=&\frac{1}{|\mathcal{T}_i|}\int_{\overline{H}_{(i,j)_k}}^\infty P_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h),
\end{aligned} \tag{4}$$

where $\mathbf{E}\{\}$ denotes expectation, $P_{(i,j)_k}(h)$ is the transmit power on link $(i,j)_k$ when the channel has power gain h and it depends on modulation and power allocation policy. For example, in order to achieve a constant SNR at the receiver, $P_{(i,j)_k}(h)$ is allocated such that $P_{(i,j)_k}(h) = \frac{P_r}{h}$, where P_r is the received power level satisfying the SNR requirement. According to the average power constraint, we have

$$\sum_{j\in\mathcal{T}_i, k=1,\ldots,K}\mathbf{E}\{P_{(i,j)_k}\} \leq P_a, \qquad \forall i,j\in\mathcal{V}. \tag{5}$$

In existing channel access protocols, there are usually several subchannels to be selected for utilization. For example, the IEEE 802.11b physical layer [22] has 14 subchannels, 5MHz apart in frequency, all of which have the same transmission capability. However, typically there is only one single RF chain, and the peak constraint on the instantaneous transmit power has to be satisfied for the total combined transmission. We have the instantaneous power constraint

$$\sum_k\left(\max_{h,j} P_{(i,j)_k}(h)\right) \leq P_m, \qquad \forall i,j\in\mathcal{V}. \tag{6}$$

Given power allocation $P_{(i,j)_k}(h)$, the achieved average data rate given that a user has decided to transmit on link $(i,j)_k$ is

$$\begin{aligned}
R_{(i,j)_k} =&\mathbf{E}\Big\{R(\eta(h))|\text{User } i \text{ transmits on } (i,j)_k\Big\} \\
=&\int_{\overline{H}_{(i,j)_k}}^\infty R(\eta(h))\Pr\{H_{(i,j)_k} = h|\text{User } i \\
&\text{transmits on } (i,j)_k\}\mathrm{d}h \\
=&\int_{\overline{H}_{(i,j)_k}}^\infty R(\eta(h))\frac{A}{B}\mathrm{d}h,
\end{aligned} \tag{7}$$

where

$$
\begin{aligned}
A =& \Pr\{H_{(i,j)_k} = h, \text{User } i \text{ transmits on } (i,j)_k\} \\
=& f(h) F^{|\mathcal{T}_i|-1}(h),
\end{aligned} \tag{8}
$$

and

$$
\begin{aligned}
B =& \Pr\{\text{User } i \text{ transmits on } (i,j)_k\} \\
=& \int_{\overline{H}_{(i,j)_k}}^{\infty} \Pr\{H_{(i,j)_k} = g, \text{User } i \text{ transmits on} \\
& (i,j)_k\} \mathrm{d}g \\
=& 1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k}).
\end{aligned} \tag{9}
$$

Hence,

$$
R_{(i,j)_k} = \frac{\int_{\overline{H}_{(i,j)_k}}^{\infty} R(\eta(h)) \mathrm{d}F^{|\mathcal{T}_i|}(h)}{1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})}, \tag{10}
$$

where $\eta(h) = \frac{hP_{(i,j)_k}(h)}{N_o W/K}$ is the received SNR, N_o is noise spectral density, W is the total system bandwidth, and $R(\eta)$ is the instantaneous data rate when channel has SNR η.

If channel capacity is achieved in AWGN channels †, $R(\eta) = W \log_2(1+\eta)$. Assuming continuous-rate M-QAM modulation and given the BER requirement, $R(\eta)$ can be expressed as $R(\eta) = W \log_2(1 + \frac{3\eta}{-2\ln(5BER)})$ according to [20]. It is easy to see that in both cases, $R(\eta)$ is strictly concave in η. In general, we assume that $R(\eta)$ is continuously differentiable with first order derivative $R^{'}(\eta)$ positive and strictly decreasing in η.

C. Criterion for Cross-Layer Design

When optimizing multi-user networks, we have to take both overall network throughput and fairness into consideration. A very commonly discussed fairness criterion is max-min fairness [21]. When max-min fairness is achieved, the throughput of a certain link can not be increased without simultaneously decreasing the throughput of another link which already has smaller throughput. Usually, max-min fairness just implies to equal sharing of channel resources on each link, which compromises the overall throughput of the wireless network a lot since different links usually have different transmission conditions. Hence, we consider proportional fairness, the objective of which is to maximize the product of throughput of all links, or the geometric average [24]. As pointed out in [25], a vector of throughputs $T = (T_1, T_2, \cdots, T_n)$ is proportionally fair if it satisfies required constraints, and for any other feasible vector $\overline{T}$, the aggregate of proportional changes is non-positive, i.e. $\sum_{i=1}^{n} \frac{\overline{T}_i - T_i}{T_i} \leq 0$. Some analysis has been given in [24] from a game-theoretic standpoint and it is shown that a strategy achieving proportional fairness satisfies certain axioms of fairness and is a Nash arbitration strategy [26]. With proportional fairness, the network will be operated at a Pareto equilibrium, which corresponds to the situation where no user can improve its throughput without affecting at least one user adversely.

Denote transmission control of the whole network as $\mathcal{C} = \{\overline{\mathcal{H}}, \mathcal{P}\}$, where $\overline{\mathcal{H}}$ is the set of predetermined channel power gain thresholds and $\mathcal{P}$ is the set of power allocation policies. With the constraints in (5) and (6), the optimal configuration of the whole network, $\mathcal{C}^* = \{\overline{\mathcal{H}}^*, \mathcal{P}^*\}$, that achieves proportional fairness among all subchannels carrying traffic flows will be

$$
\mathcal{C}^* = \arg \max_{\{\overline{\mathcal{H}}, \mathcal{P}\}} \sum_{(i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i} \ln(T_{(i,j)_k}), \tag{11a}
$$

subject to

$$
\sum_{j \in \mathcal{T}_i, k=1,\ldots,K} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h) \leq P_a, \tag{11b}
$$

and

$$
\sum_k \left(\max_{h,j} P_{(i,j)_k}(h) \right) \leq P_m, \tag{11c}
$$

where throughput $T_{(i,j)_k}$ is given by (3). Denote utility $U_{(i,j)_k} = \ln(T_{(i,j)_k})$. Problem (11) aims to maximize overall network utility subject to individual power limits.

IV. Decentralized Optimization

In the previous section, we have discussed a criterion for cross-layer design. The optimization of (11) depends on the threshold configuration, $\overline{\mathcal{H}}$, power allocation, $\mathcal{P}$, and modulation policy. The global optimization of the problem is difficult and computationally expensive, and requires complete network knowledge for each user. Therefore, in this section, we find a suboptimal solution, which only needs decentralized neighborhood information.

From (11), we have

$$
\begin{aligned}
\mathcal{C}^* =& \arg \max_{\{\overline{\mathcal{H}}, \mathcal{P}\}} \sum_{(i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i} \Bigg(\ln \Big(p_{(i,j)_k} (1 - p_{j_k}) \\
& \prod_{a \in \mathcal{N}_j, a \neq i} (1 - p_{a_k}) \Big) + \ln \left(R_{(i,j)_k} \right) \Bigg).
\end{aligned} \tag{12}
$$

(12) reveals two ways to improve the overall system performance. One way is to reduce the probability of collisions in the whole network, whose effect is captured by the term $p_{(i,j)_k}(1-p_{j_k}) \prod_{a \in \mathcal{N}_j, a \neq i} (1-p_{a_k})$. The other is to allocate power properly so that the achieved data rate of each individual user can be maximized. Hence, we decompose it into two related problems, and find a suboptimal transmission control policy. The solution to find optimal MAC layer transmission control $\overline{\mathcal{H}}^*$ to resolve collisions in the whole network while guaranteeing

†In slow fading channels, channel varies slightly within each packet. With sufficiently long packet length, ideal coding can be applied to achieve channel capacity

proportional fairness can be formulated by

$$\overline{\mathcal{H}}^* = \arg\max_{\overline{\mathcal{H}}} \sum_{(i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i} \Bigg(\ln \big(p_{(i,j)_k} (1 - p_{j_k}) \prod_{a \in \mathcal{N}_j, a \neq i} (1 - p_{a_k}) \big) \Bigg). \tag{13}$$

Given MAC transmission decision, in order to maximize the mean physical layer throughput within power capability, the optimal power allocation $\mathcal{P}_i^*$ of User i is formulated by

$$\mathcal{P}_i^* = \arg\max_{\mathcal{P}_i} \sum_{j \in \mathcal{T}_i, k} R_{(i,j)_k}, \tag{14a}$$

subject to (11b)

$$\sum_{j \in \mathcal{T}_i, k=1,\ldots,K} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h) \leq P_a, \tag{14b}$$

and (11c)

$$\sum_k \left(\max_{h,j} P_{(i,j)_k}(h) \right) \leq P_m, \tag{14c}$$

where $\{\overline{H}^*_{(i,j)_k}\}$ is the solution of (13) and $R_{(i,j)_k}$ is given by (10). Although problem (11) has been decomposed into (13) and (14) to resolve network collisions and improve individual transmission capability respectively, these two problems are closely coupled through $\overline{\mathcal{H}}^*$.

A. MAC Layer Transmission Control

When optimizing the network with proportional fairness in (13), all users are assumed to transmit at the same data rate once the channel power gain is above a certain threshold. Problem (13) turns out to be similar with the problem of finding distributed access control strategy to achieve proportional fairness in traditional Aloha networks [27] and [28]. By applying techniques used in [27] and [28], the optimal transmission probability is readily achieved,

$$p^*_{(i,j)_k} = \frac{1}{|\mathcal{S}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|}. \tag{15}$$

Combining (1) and (15), Theorem 1 follows immediately, and the proof is omitted.

Theorem 1: The optimal predetermined channel power gain threshold for any link $(i,j)_k \in \mathcal{E}$ where $j \in \mathcal{T}_i$, $\overline{H}^*_{(i,j)_k}$, as defined in (13), is given by

$$\overline{H}^*_{(i,j)_k} = F^{-1} \left[(1 - \frac{|\mathcal{T}_i|}{|\mathcal{S}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|})^{\frac{1}{|\mathcal{T}_i|}} \right]. \tag{16}$$

From threshold (16), the optimal threshold of User i is independent of the receiver j but depends on the neighborhood information of User i itself, including the number of users receiving packets from User i, $|\mathcal{T}_i|$, the number of users sending packets to User i, $|\mathcal{S}_i|$, and the total number of users sending packets to the interfering neighbors of User i, $\sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|$. The first two are local information while $|\mathcal{S}_m|'s, m \in \mathcal{N}_i$, is information about interfering neighbors. The number of flows each interfering neighbor receives, i.e. $|\mathcal{S}_m|$ for all $m \in \mathcal{N}_i$, can be obtained through broadcasting of the interfering neighbor whenever this numbers changes. Since this knowledge needs to be broadcast to notify the interfering neighbors, we call it *two-hop knowledge*. The broadcasting of this two-hop knowledge incurs only trivial signalling overhead since only when either a traffic session or the network topology varies will this broadcasting be triggered. Besides, some form of two-hop knowledge is typical in many protocols, like routing information discovery in mobile *ad hoc* networks [29], [30]. Hence, it can be easily obtained.

Consider User 7 in Figure 1. It is easy to see that $|\mathcal{T}_7| = 3$, $|\mathcal{S}_7| = 1$, and the two-hop knowledge of interfering neighbors $|\mathcal{S}_5| = 1$, $|\mathcal{S}_6| = 2$, $|\mathcal{S}_8| = 2$, and $|\mathcal{S}_9| = 2$. Hence, for all $j \in \mathcal{T}_7$ and $k = 1, \cdots, K$, $\overline{H}^*_{(7,j)_k} = F^{-1}\left[(1 - \frac{3}{1+7})^{1/3}\right] = F^{-1}(0.855)$. If the channel is Rayleigh with average power gain h_a, $\overline{H}^*_{(7,j)_k} = 1.931 h_a$. Hence, since there are many traffic flows in the neighborhood of User 7, it transmits only when the channel has very good condition.

As we can see above, the optimal threshold can be obtained through two-hop knowledge. In the following, we consider two special applications.

1) Transmission Control with One-Hop Knowledge
To avoid signalling broadcast, assume no user has two-hop knowledge, and it needs to be estimated to get approximation of the optimal thresholds. Since the transmission of each interfering neighbor $j \in \mathcal{N}_i$ can be detected by User i, $|\mathcal{T}_j|$ is available. User i can approximate $|\mathcal{S}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|$, the total number of received traffic flows within the interfering range of User i, to be $|\mathcal{T}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{T}_m|$, the total number of transmitted traffic flows User i can detect. Hence, instead of (16), the transmission threshold with one-hop knowledge, i.e. local knowledge, is

$$\overline{H}^*_{(i,j)_k} = F^{-1} \left[(1 - \frac{|\mathcal{T}_i|}{|\mathcal{T}_i| + \sum_{j \in \mathcal{N}_i} |\mathcal{T}_j|})^{\frac{1}{|\mathcal{T}_i|}} \right]. \tag{17}$$

Since the approximation in (17) is not always accurate, there might be some performance degradation. Approximation error happens when there exists undetectable traffic flows that are sent either into or out of the interfering range of User i.

2) Transmission Control for One-Hop Networks
Assume that all users are within the transmission range of each other, i.e., this is a one-hop network. A simple example is the uplink transmissions of different users to the access point in wireless lan, and at most one traffic flow within the network can succeed in transmission in one transmission slot on

one subchannel. Denote $n = |\mathcal{S}_i| + \sum_{m\in\mathcal{N}_i} |\mathcal{S}_m|$ for any User i, then n is the same for all users and represents the total number of traffic flows in the network. During any time slot on each subchannel, at most one traffic flow within the network can send data successfully. The transmission threshold is given by

$$\overline{H}^*_{(i,j)_k} = F^{-1}\left[(1-\frac{|\mathcal{T}_i|}{n})^{\frac{1}{|\mathcal{T}_i|}}\right]. \qquad (18)$$

If each user has only one traffic flow to send, i.e. $|\mathcal{T}_i| = 1$, the transmission threshold is

$$\overline{H}^*_{(i,j)_k} = F^{-1}\left[(1-\frac{1}{n})\right], \qquad (19)$$

which is the same as the transmission control in [8]. [8] has demonstrated that the total throughput for such a system achieves a fraction, $(1-\frac{1}{n})^{n-1}$, of its counterpart's throughput with an optimum centralized scheduler. The throughput reduction is due to the inherent contention in random access.

B. Physical Layer Optimization with Channel Inversion

Consider a simple transmitter adaptation technique, channel inversion [31], which maintains a constant received power level so that the signals can be reliably received during each traffic session. Once the MAC decides to transmit with channel power gain h, the transmit power is directly given by $P_t = P_r/h$, where P_r is the received power level. Different traffic flows may have different received power levels, P_r, according to the power allocation strategy. The reliable transmission data rate is given by $R(P_r)$. According to the assumption in Section III-B, $R(P_r)$ is strictly concave in P_r since the average noise power is constant on each subchannel.

From (4), the average transmit power on link $(i,j)_k$ is

$$\mathbf{E}\{P_{(i,j)_k}\} = \frac{1}{|\mathcal{T}_i|}\int_{\overline{H}_{(i,j)_k}}^{\infty} \frac{P_{r(i,j)_k}}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h). \qquad (20)$$

Hence, the instantaneous received power is

$$P_{r(i,j)_k} = |\mathcal{T}_i|\mathbf{E}\{P_{(i,j)_k}\}\left(\int_{\overline{H}_{(i,j)_k}}^{\infty}\frac{\mathrm{d}F^{|\mathcal{T}_i|}(h)}{h}\right)^{-1}. \qquad (21)$$

Denote by $\mathcal{P}_{ri} = \{P_{r(i,j)_k} | (i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i\}$ the set of the received power configuration of User i. According to (10), the average data rate is $R_{(i,j)_k} = \mathbf{E}\{R(\eta(h))\} = R(P_{r(i,j)_k})$. The problem in (14) is equivalent to

$$\mathcal{P}^*_{ri} = \arg\max_{\mathcal{P}_{ri}} \sum_{j\in\mathcal{T}_i,k} R(P_{r(i,j)_k}), \qquad (22a)$$

subject to

$$\sum_{j\in\mathcal{T}_i,k}\frac{1}{|\mathcal{T}_i|}\int_{\overline{H}^*_{(i,j)_k}}^{\infty}\frac{P_{r(i,j)_k}}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h) \le P_a, \qquad (22b)$$

and

$$\sum_k\left(\max_j \frac{P_{r(i,j)_k}}{\overline{H}^*_{(i,j)_k}}\right) \le P_m. \qquad (22c)$$

The above power allocation problem is solved by Theorem 2, which is proved in Appendix II.

Theorem 2: Assuming the strict concavity of the data rate function $R(P_r)$, (22) has unique globally optimal reception power levels $P^*_{r(i,j)_k}$ on any link $(i,j)_k \in \mathcal{E}$ where $j \in \mathcal{T}_i$

$$P^*_{r(i,j)_k} = \min\left(\frac{P_a}{K}\Big(\int_{\overline{H}^*_{(i,j)_k}}^{\infty}\frac{1}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h)\Big)^{-1}, \frac{P_m\overline{H}^*_{(i,j)_k}}{K}\right), \qquad (23)$$

in which $\overline{H}^*_{(i,j)_k}$ is determined by Theorem 1.

Whenever MAC decides to transmit, the physical layer always execute the transmission. However, when $\overline{H}^*_{(i,j)_k}$ is very small, (23) turns out to be very small and the physical layer has extremely low throughput due to the penalty of allowing transmission on deeply faded channels. Hence, $\overline{H}^*_{(i,j)_k}$ should be further modified by the physical layer to avoid transmitting on deeply faded channels. Observing (15), $p^*_{(i,j)_k}$ can be $1, \frac{1}{2}, \frac{1}{3}$, etc.. Assuming Rayleigh channel with average power gain h_a and one traffic flow is carried, the corresponding thresholds are $0, 0.69h_a, 1.10h_a$, etc.. Hence, transmission on deeply faded channels is possible only when $p^*_{(i,j)_k} = 1$. Thus, define $\overline{H}_o$ as

$$\overline{H}_o = \arg\max_{\overline{H}} R\Big(\frac{P_a}{K\int_{\overline{H}}^{\infty}\frac{1}{h}f(h)\mathrm{d}h}\Big)(1-F(\overline{H})), \qquad (24)$$

which leads to the maximum physical layer throughput when the physical layer is required to transmit under any channel conditions. If $\overline{H}^*_{(i,j)_k}$ determined by Theorem 1 is less than $\overline{H}_o$, then substitute it with $\overline{H}_o$. This lowers $p_{(i,j)_k}$ slightly since the channel is not deeply faded most of the time. The revision effectively improves link performance but impacts trivially overall network performance, and we do not need to further improve the thresholds of other users to adapt to this change for the sake of optimality in (13), which otherwise incurs additional signalling overhead.

With channel inversion, the instantaneous transmit power allocation $\mathcal{P}^*$ is:

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P^*_{r(i,j)_k}}{h} & h \ge \overline{H}^*_{(i,j)_k} \\ 0 & otherwise \end{cases}. \qquad (25)$$

C. Physical Layer Optimization with Adaptive Modulation and Power Allocation

Consider ideal physical layer transmissions. Each user can vary both the transmit power and rate to achieve the best transmission performance. According to (10) and

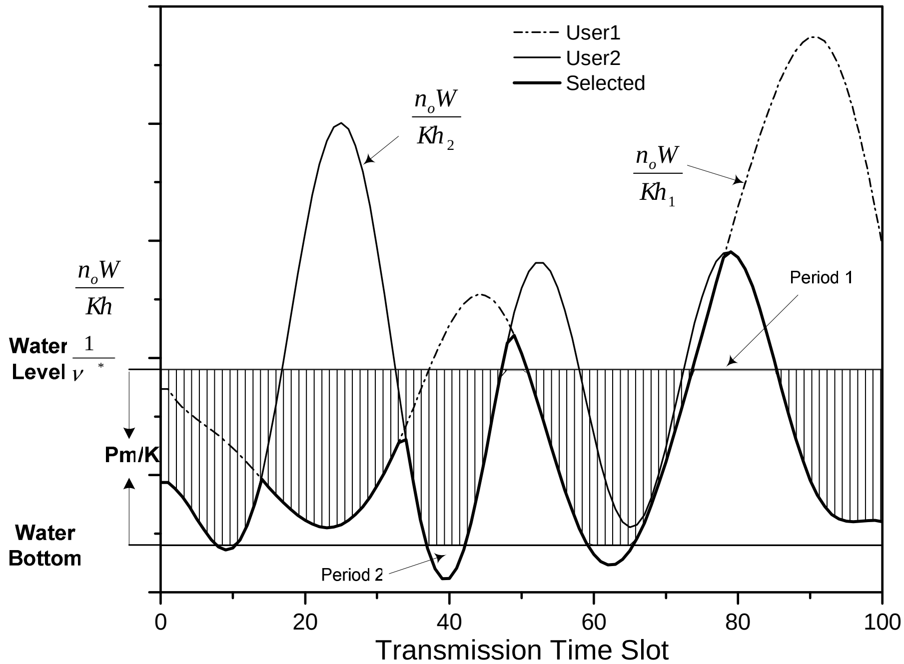

Fig. 2. Capability limited water-filling over time

(14), the power allocation strategy can be formulated by

$$\mathcal{P}_i^* = \arg\max_{\mathcal{P}_i} \sum_{j\in\mathcal{T}_i,k} \frac{\int_{\overline{H}^*_{(i,j)_k}}^{\infty} R(\eta(h))\mathrm{d}F^{\mathcal{T}_i}(h)}{(1-F^{\mathcal{T}_i}(\overline{H}^*_{(i,j)_k}))} \tag{26a}$$

subject to (11b)

$$\sum_{j\in\mathcal{T}_i,k} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) \le P_a, \tag{26b}$$

and (11c)

$$\sum_k \left(\max_{h,j} P_{(i,j)_k}(h)\right) \le P_m. \tag{26c}$$

The optimal solution of (26) is given in Theorem 3, which is proved in Appendix III.

Theorem 3: Assume the data rate function $R(\eta)$ to be continuously differentiable and the first order derivative $R^{'}(\eta)$ is positive and strictly decreasing. For any link $(i,j)_k \in \mathcal{E}$ where $j \in \mathcal{T}_i$, (26) has a unique globally optimal power allocation given by: if $P_m < \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$, $P^*_{(i,j)_k}(h) = \frac{P_m}{K}$ for $h \ge \overline{H}^*_{(i,j)_k}$; otherwise,

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P_m}{K} & \nu^* < R^{'}\left(\frac{hP_m}{n_oW}\right)\frac{hK}{n_oW}, \\ 0 & \nu^* \ge R^{'}(0)\frac{hK}{n_oW}, \\ R^{'-1}\left(\frac{\nu^* n_oW}{hK}\right)\frac{n_oW}{Kh} & \text{otherwise}, \end{cases} \tag{27}$$

for $h \ge \overline{H}^*_{(i,j)_k}$. $R^{'-1}()$ is the inverse function of $R^{'}()$. $\nu^* \ge 0$ is uniquely determined by

$$\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = \frac{P_a}{K}, \tag{28}$$

where $\overline{H}^*_{(i,j)_k}$ is given by Theorem 1.

Observing (27), when $\nu^* \ge R^{'}(0)\frac{hK}{n_oW}$, the channel is deeply faded and although the MAC layer decides to transmit, the physical layer further optimizes the transmission and decides not to transmit.

For example, assume the data rate function to be $R(\eta) = \frac{W}{K}\ln(1+\eta)$. The power allocation when $P_m \ge \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$ is given by

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P_m}{K} & \frac{1}{\nu^*} - \frac{n_oW}{Kh} > \frac{P_m}{K} \\ 0 & \frac{1}{h} \ge \frac{K}{\nu^* n_oW} \\ \frac{1}{\nu^*} - \frac{n_oW}{Kh} & \text{otherwise} \end{cases} \tag{29}$$

for $h \ge \overline{H}^*_{(i,j)_k}$, which is similar to the well-known water-filling power allocation scheme [32]–[34]. Since the proposed power allocation scheme has maximum instantaneous power constraint, we call it capability-limited water filling.

According to (27), power will be optimally distributed over both time and all subchannels. Figure 2 illustrates the capability-limited power allocation of a user that is transmitting data to User 1 and 2 on a subchannel by using (29), and the striped parts in the figure represent the amount of power allocated. The power allocation during 100 transmission time slots is shown. We assume that $\frac{\nu^* n_oW}{K} > \overline{H}^*_{(i,j)_k}$ here. According to the transmission policy, the user always selects the destination with better channel power gains. As indicated by "Period 1" in Figure

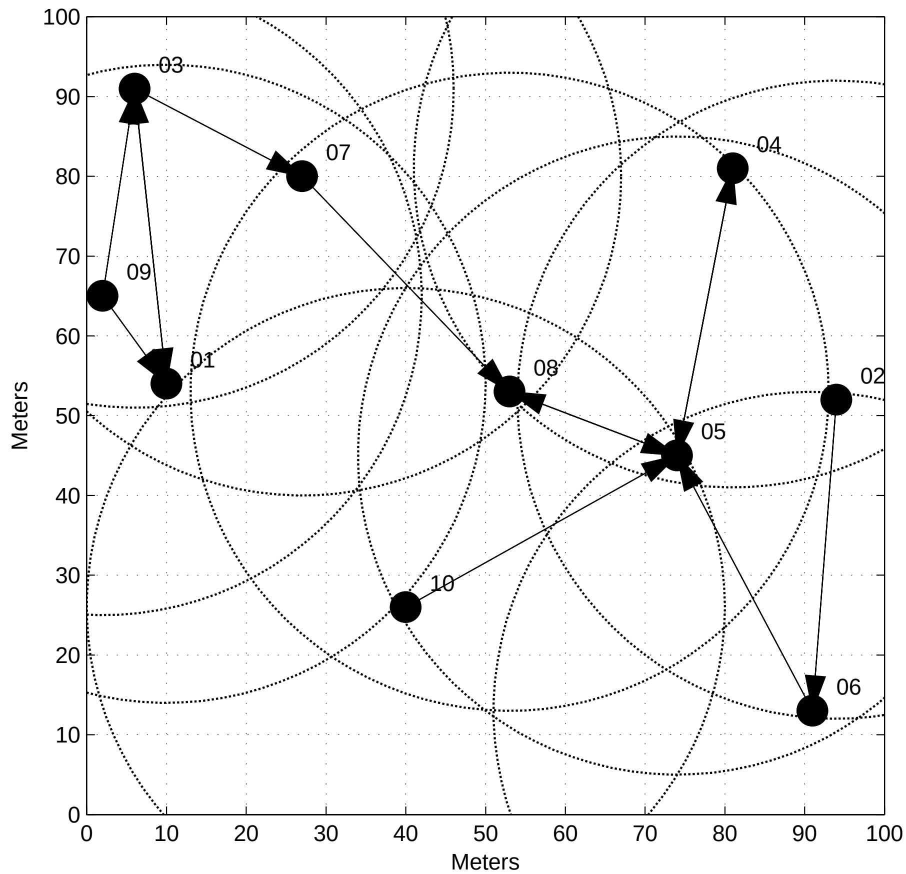

Fig. 3. Random network topology.

2 there are no transmissions when subchannels of both User 1 and 2 are deeply faded. In "Period 2" in Figure 2, although the channel conditions are so good that higher data rates can be achieved, the actual data rate is limited by the instantaneous transmission capability P_m.

When $\frac{\nu^* n_o W}{K} < \overline{H}^*_{(i,j)_k}$, since the MAC decides to transmit only when $h > \overline{H}^*_{(i,j)_k}$, the physical layer will always transmit when the MAC wants to transmit according to (29). Assuming large P_m, the power allocation is always $\frac{1}{\nu^*} - \frac{n_o W}{Kh}$. Then according to (28), the water level is

$$\frac{1}{\nu^*} = \frac{\frac{n_o W}{K} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} \frac{1}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h) + \frac{P_a}{K}}{1 - F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}. \quad (30)$$

We can always use (30) to approximate the water level since with large probability, most transmissions will fall within the normal working ranges of the transmitter.

V. Simulation Results

In this section, we first demonstrate DOMRA performance in a network with random topologies. Then we further show how closely DOMRA performs to the globally optimum solution.

A. Network Performance Improvement

Consider a network with random topologies and compare the average performance of all simulation trials. In each simulation trial, users are randomly dropped and uniformly distributed in a square area with side length one hundred meters. Each user has a transmission range of forty meters and selects neighboring users randomly for data transmission. Figure 3 illustrates a network topology in one trial, where arrows indicate traffic flows and circles transmission ranges of different users. Different schemes will be implemented to provide detailed performance comparisons.

1) Single-channel network: Assume that the network operates with one channel. For simplicity, assume Rayleigh fading channel and $R(P) = W \ln(1 + \frac{hP}{WN_0})$. We will compare the performance of the proposed cross-layer transmission policy with the channel-aware Aloha in [8], and the optimal traditional Aloha in [27], which does not consider cross-layer optimizations. For traditional Aloha transmissions, in order to make the comparison meaningful, the same average power constraint and instantaneous power constraint are enforced. Since there is no cooperation between MAC and the physical layer, the physical layer assumes that it keeps on transmitting except when the channel is deeply faded. In order to

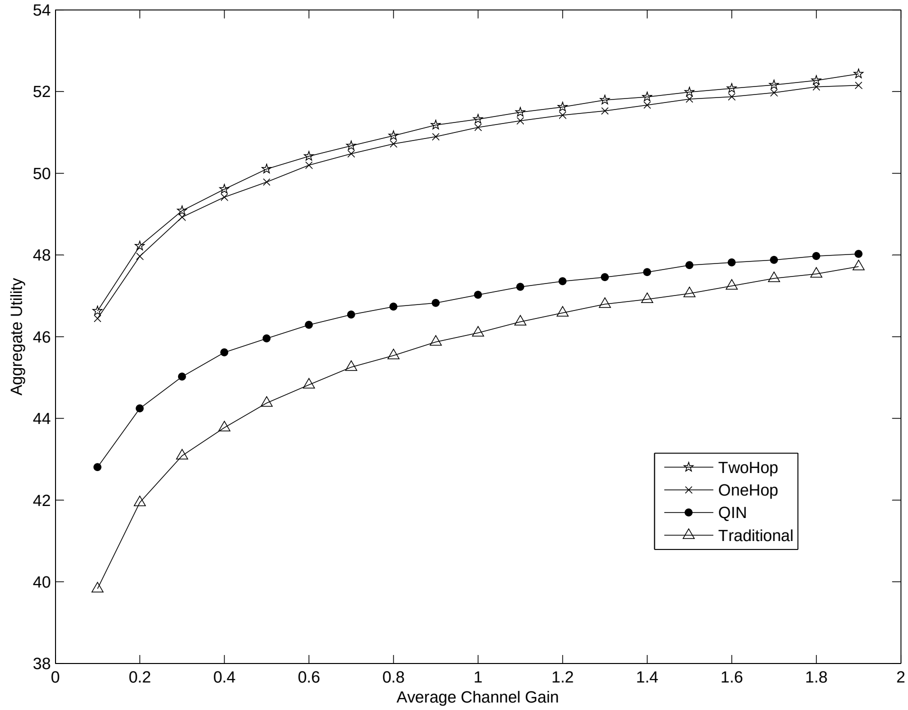

Fig. 4. Network aggregate utility comparison. $P_m = 50\text{dBm}, P_a = 43\text{dBm}$, $W = 100\text{Hz}$, and $N_o = 0.001\text{W/Hz}$.

satisfy power constraints, the transmission threshold is chosen so that the average data rate is maximized, i.e. $\overline{H} = \arg\max_H (1 - F(H))R(P_r)$ subject to the instantaneous power constraint (6), and P_r is given by (21). The threshold is found through linear search.

Figure 4 shows the aggregate utility comparison of the whole network when the channel has different average channel gains. The "TwoHop" curve represents the result of DOMRA when each user has two-hop information of the neighboring users while the "OneHop" curve represents the result when each user has only one-hop information. As we can see, with only one-hop knowledge, the system has slight performance degradation as compared with the transmissions when two-hop knowledge is available. Curve "QIN" shows the performance of [8], which assumes that each user has the knowledge of how many users there are in the whole network. Curve "Traditional" shows the result using the traditional optimal Aloha. As shown in Figure 4, with the advantage of cross layer design, the proposed scheme outperforms traditional optimal Aloha greatly. In addition, by exploiting the neighborhood information of each user, the proposed method also outperforms the existing channel-aware Aloha in [8]. This is due to the consideration of the inhomogeneous traffic spatial distribution in the proposed scheme and the channels are better utilized.

2) Multichannel network: Consider the same wireless network configurations as those in the single-channel network scenario except that there are five subchannels. Besides implementing schemes in the single-channel network scenario for multichannel environment, we also run the CAMCRA proposed in [11]. During each each transmission slot, CAMCRA chooses c subchannels with the c most significant gains, where $c = \max\left(1, \lfloor \frac{\text{subchannel number}}{\text{user number}} \rfloor\right)$. Then the method in ([8]) is applied on each subchannel given that each user knows how many users are using the subchannel. Since the number of users in each subchannel is a random variable, it is proposed in [11] to use $\max\left(1, \frac{\text{user number}}{\text{subchannel number}}\right)$ as an estimate. As shown in Figure 5, the CAMCRA in [11] has slight performance improvement as compared with the channel-aware Aloha in [8] because of exploitation of multichannel diversity. However, these two schemes do not perform good when the network has arbitrary spatial traffic distribution. Our proposed DOMRA with either two-hop or one-hop information significantly outperforms these existing schemes due to exploitation of multiuser diversity and proper adaptive transmission settings and power allocations according to inhomogeneous traffic spatial distribution in the network.

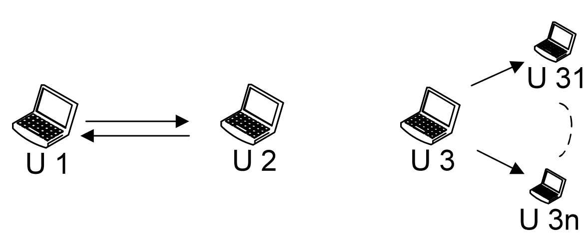

Fig. 6. A simple network topology

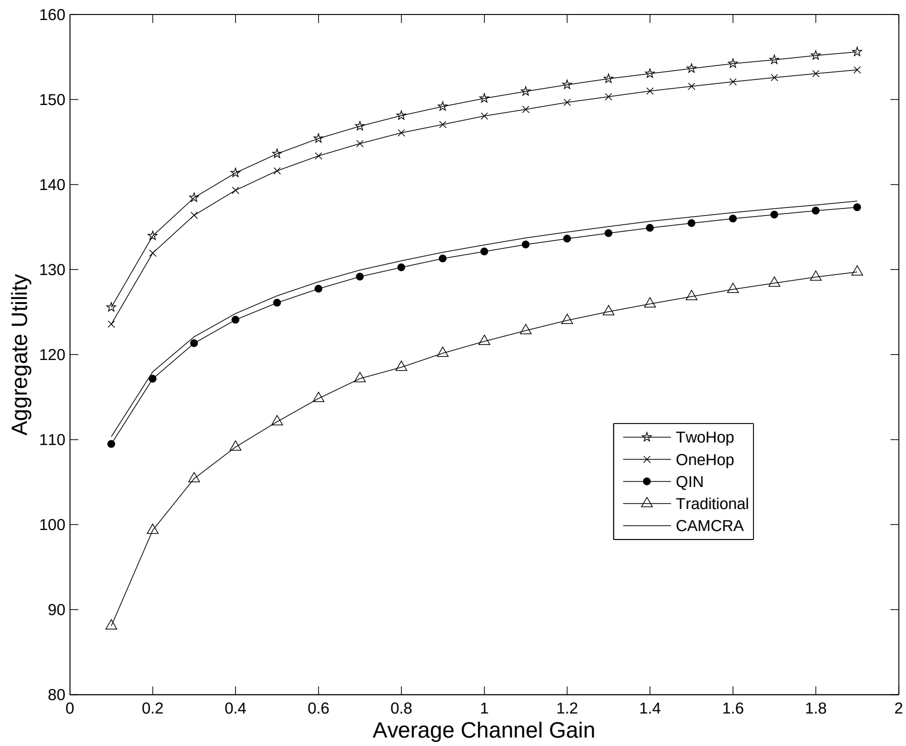

Fig. 5. Five channel network aggregate utility comparison. $P_m = 50\text{dBm}$, $P_a = 43\text{dBm}$, $W = 100\text{Hz}$, $N_o = 0.001\text{W/Hz}$.

B. *Suboptimality Gap*

Problem (11) is decomposed into subproblems (13) and (14) to obtain feasible suboptimal control policy. In order to show the suboptimality gap, we exhaustively search for the global optimum in (11), and run a simple network topology to reduce search complexity. As shown in Figure 6, arrows indicate traffic flows. User 3 is sending traffic to n receivers, who are all out of the transmission ranges of Users 1 and 2. User 1 can communicate with 2, but not 3, while User 2 can communicate with both. When n is zero, the traffic distribution is symmetric in the network. The larger the number n, the more asymmetric the traffic distribution is. We call n *traffic asymmetry*, and vary it from 0 to 8. Figure 7 compares network aggregate utility and shows the suboptimality gap. While the global optimum can only be obtained through floods of broadcast of complete network knowledge, our decomposition technique yields a feasible suboptimal decentralized solution, which requires limited (two-hop knowledge case), or no (one-hop knowledge case) signalling overhead. Besides, the proposed scheme performs closely to the global optimum, and even reaches the global optimum when the traffic is symmetric.

VI. Conclusions and Future Research

We have proposed a joint physical-MAC layer optimization policy for multichannel Aloha random access in wireless networks in which all users are not necessarily within the transmission range of each other and each user may have packets to send to or receiver from different users. The joint physical-MAC layer optimization policy exploits decentralized CSI, and achieves multi-user diversity through cross-layer design. System performance is optimized while proportional fairness is obtained with the consideration of the inhomogeneous characteristics of the traffic spatial distribution. Simulation results show that the proposed scheme significantly outperforms existing channel aware Aloha schemes. The generality of the design in this paper will allow its applications in different types of wireless networks to fully exploit the system capacity. The scheme presented here is simple but gives guidelines for decentralized cross-layer optimization in practical wireless networks. The methodology provided can be easily adapted to improve the performance of different wireless networks. For example, in networks based on 802.11 standards, besides using the backoff window technology, the transmission of RTS to compete for channel access can also be designed according to the proposed DOMRA to further decrease the collision probability and allow larger successful probability of users with better channel power gain. In the future we will consider extensions that take into accounts specific features of 802.11 and 802.16 based networks. More practical channel models will be incorporated. Implementation issues regarding modulation and coding policy will also be considered.

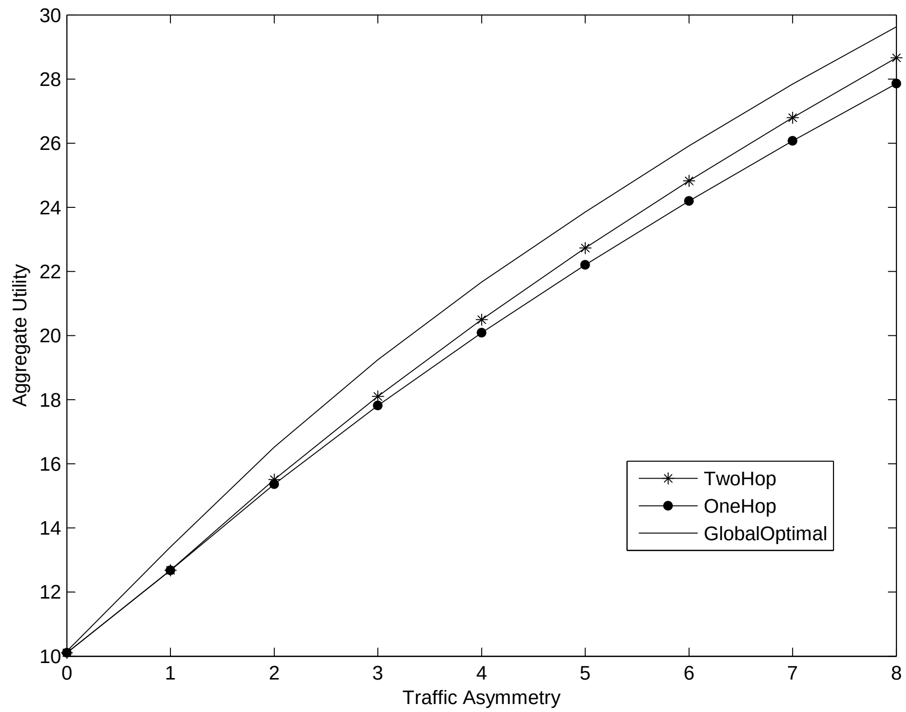

Fig. 7. Aggregate utility gap to the global optimum. $P_m = 50\mathrm{dBm}, P_a = 43\mathrm{dBm}$, $W = 100\mathrm{Hz}$, $N_o = 0.001\mathrm{W/Hz}$.

APPENDIX I
PROOF OF TRANSMISSION PROBABILITY

Since in each transmission time slot, User i sends packets to User j on subchannel k only when this subchannel has the best channel gain among all users in $\mathcal{T}_i$, and the subchannel power gain is above $\overline{H}_{(i,j)_k}$, we get the following:

$$\begin{aligned}
p_{(i,j)_k} &= \Pr\left\{h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k}), h_{(i,j)_k} \geq \overline{H}_{(i,j)_k}\right\}\\
&= \Pr\left\{h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k})\right\}\cdot\\
&\quad \Pr\left\{h_{(i,j)_k} \geq \overline{H}_{(i,j)_k}\,\middle|\, h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k})\right\}\\
&= \frac{1}{|\mathcal{T}_i|}\ \Pr\left\{\max_{a\in\mathcal{T}_i}(h_{(i,a)_k}) \geq \overline{H}_{(i,j)_k}\right\}\\
&= \frac{1}{|\mathcal{T}_i|}\left(1 - \prod_{a\in\mathcal{T}_i}\Pr\left(h_{(i,a)_k} < \overline{H}_{(i,j)_k}\right)\right)\\
&= \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right)
\end{aligned}$$

APPENDIX II
PROOF OF THEOREM 2

According to (16), we can see that $\overline{H}^*_{(i,j)_k}$ is independent of j and k. Hence,the first constraint of (22) is $\sum_{j\in\mathcal{T}_i,k=1,...,K} P_{r(i,j)_k} \leq P_a|\mathcal{T}_i|\big(\int_{\overline{H}^*_{(i,j)_k}}^{\infty}\frac{1}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h)\big)^{-1}$. Since data rate function $R()$ is assumed to be a strictly concave function, $\sum_{j\in\mathcal{T}_i,k} R(P_{r(i,j)_k}) \leq |\mathcal{T}_i|KR(\frac{\sum_{j\in\mathcal{T}_i,k}P_{r(i,j)_k}}{K|\mathcal{T}_i|})$. The equation holds if and only if $P_{r(i,j)_k}$ is the same value for all $j \in \mathcal{T}_i$ and k. Hence, for optimal solution, $\max_j \frac{P^*_{r(i,j)_k}}{\overline{H}^*_{(i,j)_k}}$ is the same for all $k = 1, ..., K$, and the second constraint of (22) is equivalent to $P^*_{r(i,j)_k} \leq \frac{P_m\overline{H}^*_{(i,j)_k}}{K}$. Then, it is easy to see that when the first constraint in (22) takes effect, the optimal solution is the first term in (23), while when the second constraint takes effect, the optimal solution is the second term in (23). Hence, (23) satisfies both constraints, and the objective value will be maximized when one constraint takes effect while satisfying the other constraint.

APPENDIX III
PROOF OF THEOREM 3

According to the symmetry of all subchannels, we see that $\max_{h,j} P_{(i,j)_k}$ is the same for all subchannels. Hence, constraint (26c) equals $\max_{h,j} P_{(i,j)_k}(h) \leq \frac{P_m}{K}$, which is the same as $P_{(i,j)_k}(h) \leq \frac{P_m}{K}$. According to (16), $\overline{H}^*_{(i,j)_k}$ is independent of j and k. Problem (26) is equivalent to:

$$\mathcal{P}_i^* = \arg\min_{\mathcal{P}_i} -\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} R\left(\frac{hP_{(i,j)_k}(h)}{n_oW/K}\right)\mathrm{d}F^{\mathcal{T}_i}(h)$$

subject to

$$\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) - |\mathcal{T}_i|P_a \leq 0, \tag{III.31a}$$

and

$$P_{(i,j)_k}(h) - \frac{P_m}{K} \le 0. \qquad \text{(III.31b)}$$

Introducing Lagrange multipliers $\lambda_{(i,j)_k}(h)$, $\gamma_{(i,j)_k}(h)$ and $\nu \ge 0$ for the three inequalities respectively, the Lagrange function associated with problem (III.31) is:

$$\begin{aligned}
&L(\mathcal{P}_i, \lambda_{(i,j)_k}, \gamma_{(i,j)_k}, \nu) \\
&= -\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} R\left(\frac{\widehat{h}P_{(i,j)_k}(\widehat{h})}{n_oW/K}\right)\mathrm{d}F^{|\mathcal{T}_i|}(\widehat{h}) + \\
&\sum_{j\in\mathcal{T}_i,k}\lambda_{(i,j)_k}(h)(-P_{(i,j)_k}(h)) + \\
&\sum_{j\in\mathcal{T}_i,k}\gamma_{(i,j)_k}(h)(P_{(i,j)_k}(h) - \frac{P_m}{K}) + \\
&\nu\left(\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(\widehat{h})\mathrm{d}F^{|\mathcal{T}_i|}(\widehat{h}) - |\mathcal{T}_i|P_a\right) \\
&= \sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty}\overline{L}_{(i,j)_k}(\widehat{h})\mathrm{d}F^{|\mathcal{T}_i|}(\widehat{h}) - \sum_{j\in\mathcal{T}_i,k}\gamma_{(i,j)_k}(h)\frac{P_m}{K} \\
&- \nu|\mathcal{T}_i|P_a,
\end{aligned}$$

where $\overline{L}_{(i,j)_k}(\widehat{h}) = -R\left(\frac{\widehat{h}P_{(i,j)_k}(\widehat{h})}{n_oW/K}\right) + \nu P_{(i,j)_k}(\widehat{h}) - \lambda_{(i,j)_k}(h)P_{(i,j)_k}(h)/C + \gamma_{(i,j)_k}(h)(P_{(i,j)_k}(h))/C$, in which $C = 1 - F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})$. According to [35], we obtain the following Karush-Kuhn-Tucker (KKT) conditions for optimal power allocation when $h \ge \overline{H}^*_{(i,j)_k}$:

$$\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) - |\mathcal{T}_i|P_a \le 0, \qquad \text{(III.32)}$$

$$P^*_{(i,j)_k}(h) - \frac{P_m}{K} \le 0, \qquad \text{(III.33)}$$

$$\lambda^*_{(i,j)_k}(h) \ge 0, \gamma^*_{(i,j)_k}(h) \ge 0, \text{ and } \nu^* \ge 0, \qquad \text{(III.34)}$$

$$\lambda^*_{(i,j)_k}(h)P^*_{(i,j)_k}(h) = 0, \qquad \text{(III.35)}$$

$$\gamma^*_{(i,j)_k}(h)\left(P^*_{(i,j)_k}(h) - \frac{P_m}{K}\right) = 0, \qquad \text{(III.36)}$$

$$\nu^*\left(\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) - |\mathcal{T}_i|P_a\right) = 0, \qquad \text{(III.37)}$$

and

$$\begin{aligned}
&\left.\frac{\partial\overline{L}_{(i,j)_k}(h)}{\partial P_{(i,j)_k}(h)}\right|_{P^*_{(i,j)_k}(h)} \\
&= -R'\left(\frac{hP_{(i,j)_k}(h)}{n_oW/K}\right)\frac{hK}{n_oW} + \nu^* - \lambda^*_{(i,j)_k}(h)/C + \\
&\quad \gamma^*_{(i,j)_k}(h)/C = 0. \qquad \text{(III.38)}
\end{aligned}$$

1°. When

$$\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) < |\mathcal{T}_i|P_a, \qquad \text{(III.39)}$$

according to (III.37), $\nu^* = 0$. From (III.38), $\gamma^*_{(i,j)_k}(h) > 0$. Hence, $P^*_{(i,j)_k}(h) = \frac{P_m}{K}$ from (III.36). (III.39) equals $P_m < \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$.

2°. When $\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = |\mathcal{T}_i|P_a$, i.e. $P_m \ge \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$,

a). if $\gamma^*_{(i,j)_k}(h) > 0$, $P^*_{(i,j)_k}(h) = \frac{P_m}{K}$ from (III.36), and $\lambda^*_{(i,j)_k}(h) = 0$ from (III.35). According to (III.38), $\nu^* < R'\left(\frac{hP_m}{n_oW}\right)\frac{hK}{n_oW}$

b). if $\gamma^*_{(i,j)_k}(h) = 0$, from (III.38), $\nu^* = R'\left(\frac{hP^*_{(i,j)_k}(h)}{n_oW/K}\right)\frac{hK}{n_oW} + \lambda^*_{(i,j)_k}(h)/C \ge R'\left(\frac{hP^*_{(i,j)_k}(h)}{n_oW/K}\right)\frac{hK}{n_oW}$. According to the strictly decreasing assumption of the first order derivative of $R(\eta)$, $R'\left(\frac{hP^*_{(i,j)_k}(h)}{n_oW/K}\right) \le R'(0)$, i.e. $\nu^* \le R'(0)\frac{hK}{n_oW} + \lambda^*_{(i,j)_k}(h)/C$, and equality holds only when $P^*_{(i,j)_k}(h) = 0$. Hence, with (III.34) and (III.35), if $\nu^* \ge R'(0)\frac{hK}{n_oW}$, it is easy to see that $P^*_{(i,j)_k}(h) = 0$. If $\nu^* < R'(0)\frac{hK}{n_oW}$, $P^*_{(i,j)_k}(h) > 0$, $\lambda^*_{(i,j)_k}(h)$=0, and $R'\left(\frac{hP^*_{(i,j)_k}(h)}{n_oW/K}\right)\frac{hK}{n_oW} = \nu^*$. Then $P^*_{(i,j)_k}(h) = R'^{-1}\left(\frac{\nu^*n_oW}{hK}\right)\frac{n_oW}{Kh}$. Hence, when $h \ge \overline{H}^*_{(i,j)_k}$, we have

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P_m}{K} & \nu^* < R'\left(\frac{hP_m}{n_oW}\right)\frac{hK}{n_oW}, \\ 0 & \nu^* \ge R'(0)\frac{hK}{n_oW}, \\ R'^{-1}\left(\frac{\nu^*n_oW}{hK}\right)\frac{n_oW}{Kh} & \text{otherwise}, \end{cases} \qquad \text{(III.40)}$$

Observing that both $\overline{H}^*_{(i,j)_k}$ and $P^*_{(i,j)_k}(h)$ are independent of j and k, substituting (III.40) into the condition $\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = |\mathcal{T}_i|P_a$, we get

$$\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = \frac{P_a}{K}. \qquad \text{(III.41)}$$

Since $P^*_{(i,j)_k}(h)$ is a piecewise-linear decreasing function of ν^* with breakpoints at $R'(0)\frac{hK}{n_oW}$ and $R'\left(\frac{hP_m}{n_oW}\right)\frac{hK}{n_oW}$, (III.41) has a unique solution of ν^*.

Theorem 3 is readily obtained from both 1° and 2°. The solution is globally optimal since for convex optimizations, KKT conditions are both necessary and sufficient for a local minimum to be a global minimum. If, in addition, the objective function is strictly convex, the globally optimal solution is unique. It is easy to see that in (III.31), given that the first order derivative $R'(\eta)$ is positive and strictly decreasing, the constraints are convex, and the objective function is strictly convex of $P_{(i,j)_k}(h), j \in \mathcal{T}_i, k = 1, \cdots, K$, and the unique global optimality follows.

References

[1] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "Cross-layer design for power efficiency and QoS provisioning in multi-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 3306-3315, Nov. 2006.

[2] P. P. Pham, S. Perreau, and A. Jayasuriya, "New cross-Layer design approach to ad hoc networks under Rayleigh fading," *IEEE Trans. J. Select. Areas Commun.*, vol. 23, no. 1, pp. 28-39, Jan. 2005.

[3] M. A. Haleem and R. Chandramouli, "Adaptive downlink scheduling and rate selection: a cross-layer design," *IEEE J. Select. Areas Commun.*, vol. 23, no. 6, pp. 1287-1297, Jan. 2005.

[4] R. Knopp and P. Humblet, "Information capacity and power control in single cell multiuser communications," in *Proc. IEEE Conf. Commun. (ICC' 95)*, Seattle, WA, Jun. 1995, pp. 331-335.

[5] P. Viswanath, D. N. C. Tse, and R. Laroia "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277-1294, June 2002.

[6] G. Song and Y. (G.) Li, "Cross-layer optimization for OFDM wireless networks-part I: theoretical framework," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 614-624, Mar. 2005.

[7] G. Song and Y. (G.) Li, "Cross-layer optimization for OFDM wireless networks-part II: algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625-634, Mar. 2005.

[8] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *Proc. IEEE INFOCOM 2003*, San Francisco, CA, Apr. 2003, pp. 1084-1094.

[9] K. Bai and J. Zhang, "Opportunistic multichannel Aloha: distributed multiaccess control scheme for OFDMA wireless networks," *IEEE Trans. Veh. Tech.*, vol. 55 no. 3, pp. 848-855, Mar. 2006.

[10] X. Qin and R. Berry, "Opportunistic splitting algorithms for wireless networks," in *Proc. IEEE INFOCOM 2004*, Hong Kong, P.R. China, Apr. 2004, pp. 1662-1672.

[11] G. Ganesan, G. Song, and Y. (G.) Li, "Asymptotic throughput analysis of distributed multichannel random access schemes," in *Proc. IEEE Conf. Commun. (ICC' 2005)*, Seoul, South Korea, May 2005, pp. 3637-3641.

[12] Y. Xue, T. Kaiser, and A. B. Gershman, "Channel-Aware Aloha-Based OFDM Subcarrier Assignment in Single-Cell Wireless Communications," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 953-962, May 2007.

[13] S. Adireddy, and L. Tong, "Exploiting decentralized channel state information for random access,"*IEEE Trans. Inf. Theory*, vol 51, no. 2, pp. 537-561, Feb. 2005.

[14] A. Maharshi, L. Tong, and A. Swami, "Cross-layer designs of multichannel reservation MAC under Rayleigh fading," *IEEE Trans. Signal Processing*, vol 51, no. 8, pp. 2054-2067, Aug. 2003.

[15] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted Aloha with multipacket reception," *IEEE Trans. Inf. Theory*, vol 51, no. 7, pp. 2636-2656, July 2005.

[16] I. F. Akyildiz, Su Weilian, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol 40, no. 8, pp. 102-114, Aug. 2002.

[17] I. Chlamtac, M. Conti, and J. J. N. Liu, "Mobile ad hoc networking: imperatives and challenges," *Ad Hoc Networks*, vol 1, no. 1, pp. 13-64, July 2003.

[18] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: commodity multihop ad hoc networks," *IEEE Commun. Mag.*, vol 43, no. 3, pp. 123-131, Mar. 2005.

[19] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol 45, no. 5, pp. 1468-1489, July 1999.

[20] A.J.Goldsmith and S. G. Chua "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[21] E. L. Hahne, "Round-robin scheduling for max-min fairness in data networks," *IEEE J. Select. Areas Commun.*, vol 9, no. 7, pp. 1024-1039, Sep. 1991.

[22] IEEE 802.11 Working Group, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," 1997.

[23] IEEE 802.11 Working Group, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," 1997.

[24] R. Mazumdar, L. G. Mason, and C. Douligieris, "Fairness in network optimal flow control: optimality of product forms," *IEEE Trans. Commun.*, vol 39, no. 5, pp. 775-782, May 1991.

[25] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Operat. Res. Soc.*, vol 49, pp. 237-252, 1998.

[26] P. D. Straffin, "Game theory and strategy," *Mathematical Association of America* , 1993.

[27] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportional fairness using local information in Aloha networks," *IEEE Trans. Autom. Control*, vol. 49, no. 10, pp. 1858-1863, Oct. 2004.

[28] P. Gupta and A. L. Stolyar, "Optimal throughput allocation in general random-access networks," in *Proc. Conf. Information Science and Systems (CISS)*, Princeton, NJ, 2006, pp. 1254-1259.

[29] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint relaying: an efficient technique for flooding in mobile wireless networks," in *Proc. of 35th Annual Hawaii Int. Conf. on Sys. Sci.*, 2001.

[30] J. Lipman, P. Boustead, and J. Judge, "Neighbor aware adaptive power flooding in mobile ad hoc networks," *Int. J. Foundations of Comp. Sci.*, vol. 14, no. 2, pp. 237-252, Apr., 2003.

[31] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol 43, no. 6, pp. 1986-1992, July 1997.

[32] I. E. Telatar, "Capacity of multi-antenna Gaussian Channels," *Europ. Trans. Telec.*, vol 10, no. 6, pp. 585-595, Nov./Dec. 1999.

[33] J. Jang, and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Select. Areas Commun.*, vol 21, no. 2, pp. 171-178, Feb. 2003.

[34] G. Miao, and Z. Niu, "Practical Feedback Design based OFDM Link Adaptive Communications over Frequency Selective Channels," in *Proc. IEEE Conf. Commun. (ICC' 2006)*, Istanbul, Turkey, June 2006, pp. 4624-4629.

[35] S. Boyd and L. Vandenberghe, "Convex Optimization," *Cambridge University Press*, 2004.