

Utility-Optimal Random Access: Reduced Complexity, Fast Convergence, and Robust Performance

A. Hamed Mohsenian-Rad, *Student Member, IEEE*, Jianwei Huang, *Member, IEEE*,
Mung Chiang, *Senior Member, IEEE*, and Vincent W.S. Wong, *Senior Member, IEEE*

Abstract—In this paper, we propose two distributed contention-based medium access control (MAC) algorithms for solving a network utility maximization (NUM) problem in wireless ad hoc networks. Most of the previous NUM-based random access algorithms have one or more of the following performance bottlenecks: (1) extensive signaling among the nodes to achieve semi-distributed implementations, (2) synchronous updates of contention probabilities, (3) small update stepsizes to ensure convergence but with typically slow speed, and (4) supporting a limited range of utility functions under which the NUM is shown to be convex. Our proposed algorithms overcome the bottlenecks in all four aspects. First, only limited amount of message passing among nodes is required. Second, fully asynchronous updates of contention probabilities are allowed. Furthermore, our algorithms are robust to arbitrary large message passing delay and message loss. Third, we do not utilize any stepsize during updates, thus our algorithms can achieve faster convergence. Finally, our proposed algorithms have provable convergence, optimality, and robustness properties under a wider range of utility functions, even if the NUM problem is non-convex. Simulation results show the optimality and fast convergence of our algorithms, performance improvements compared with the subgradient-based MAC, and better efficiency-fairness tradeoff compared with the IEEE 802.11 distributed coordination function.

Index Terms—Network utility maximization, contention-based medium access control, non-convex optimization, complexity reduction, robust design, α -fair utility functions.

I. INTRODUCTION

THERE are two major types of wireless medium access control (MAC) protocols: scheduling-based (e.g., in cellular systems) and contention-based (e.g., in wireless local area

networks). In this paper, we focus on the study of contention-based MAC, where nodes *randomly* and *distributively* access the shared channel with certain transmission probabilities.

The contention-based protocols are scalable and inherently flexible, but they typically have poor performance due to insufficient feedback. For example, in IEEE 802.11 distributed coordination function (DCF) [1], a node updates its transmission probability based on the *binary* feedback of its data transmission: success (no collision) or failure. This leads to low throughput and unfair resource allocation. Such mechanism also cannot achieve a stable equilibrium [2].

In this paper, we design distributed contention-based random MAC algorithms through the framework of network utility maximization (NUM). Several related algorithms on the same NUM framework have been proposed in [3]–[7]. They have one or more of the following performance bottlenecks: (1) extensive message passing among nodes to achieve semi-distributed implementation, (2) synchronous updates of contention probabilities that require homogeneous computational capabilities and software implementations among nodes, (3) small update stepsizes to guarantee convergence with typically slow speed, and (4) supporting only a limited range of utility functions due to non-convexity.

Our proposed algorithms overcome the above performance bottlenecks in all four aspects. First, they only require limited message passing (i.e., signalling) among nodes. Based on the messages from other nodes, each node updates its persistent probabilities by solving a *local* and *myopic* optimization problem in an attempt to maximize the *total* network utility. Compared to the NUM-based random access algorithm in [3], our algorithms can reduce the total signalling overhead by a factor of ten. Second, our algorithms allow asynchronous updates of messages and contention probabilities. They can tolerate arbitrary large and finite asynchronism and message delay and are also robust to message loss. For example, even when the packet loss rate of the underlying communication channel is up to 0.5 (i.e., on average, half of the messages are lost), our algorithms can still achieve the optimal performance within a short time. Third, in our algorithms, nodes update their contention probabilities through best response updates, thus no small stepsizes are needed. This enables our algorithms to achieve a much faster convergence compared with the previous subgradient-based update methods (e.g., in [3]–[6]). Finally, our algorithms have provable convergence property under a

Manuscript received December 24, 2007; revised July 7, 2008; accepted September 7, 2008. The associate editor coordinating the review of this paper and approving it for publication was G. Xue.

A. H. Mohsenian-Rad and V.W.S. Wong are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada, (e-mail: {hamed, vincentw}@ece.ubc.ca).

J. Huang is with the Information Engineering Department, Chinese University of Hong Kong, Hong Kong, China, (e-mail: jwhuang@ie.cuhk.edu.hk).

M. Chiang is with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, (e-mail: chiangm@princeton.edu).

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Competitive Earmarked Research Grants (Project Number 412308) under the University Grant Committee of the Hong Kong Special Administrative Region, China, the Direct Grant (Project Number C001-2050398) of The Chinese University of Hong Kong, the National Key Technology R&D Program (Project Number 2007BAH17B04) established by the Ministry of Science and Technology of the People's Republic of China, as well as NSF CNS-0720570, AFOSR FA9550-06-1-0297, and U.S. ONR Grant N00014-07-1-0864. Part of this paper was presented at the MILCOM'08 Conference, San Diego, CA, November 2008.

Digital Object Identifier 10.1109/TWC.2008.071359.

wider range of utility functions, even if the NUM problem cannot be transformed into a convex optimization problem. The techniques that we use in this paper are general and may be used to tackle other non-convex optimization problems in communications and networking. Our results have been recently extended in [8] to achieve utility-optimal network performance *without* message passing.

Besides the NUM-based approach, another related thread of research focuses on the analysis of random access algorithms using game theory (e.g., [9]–[12]). The focus is on noncooperative interaction among wireless nodes, while here, we focus on *global* network performance optimization. Our results may also be extended to a mechanism design for selfish behavior of nodes so as to align them the social utility-optimality.

The rest of this paper is organized as follows. The system model is described in Section II. Our proposed distributed algorithms are presented in Section III. The convergence, optimality, and robustness of our algorithms are analytically proved in Section IV. Simulation results and performance assessments are reported in Section V. Conclusions as well as outlines for future works are given in Section VI.

II. SYSTEM MODEL

Consider a wireless ad-hoc network. Let $\mathcal{N} = \{1, \dots, N\}$ denote the set of nodes and $\mathcal{L} = \{1, \dots, L\}$ denote the set of unidirectional wireless links. For each node $n \in \mathcal{N}$, we denote the set of its *outgoing* links by $\mathcal{L}_n \subset \mathcal{L}$, with size $L_n = |\mathcal{L}_n|$. Each node n has L_n separate queues and each queue holds the packets for one of its outgoing links (see Fig. 1). Time is divided into equal-length slots. At each time slot, node n may choose to transmit on one of its outgoing links $i \in \mathcal{L}_n$ with a *persistent probability* p_i . The probabilities need to satisfy $\sum_{i \in \mathcal{L}_n} p_i \leq P_n^{\max} < 1$, where P_n^{\max} denotes the maximum *total* persistent probability on node n . Node n may remain silent in some slots. For the network in Fig. 1, node a has $L_a = 2$ outgoing links where $\mathcal{L}_a = \{1, 2\}$. In node a , those packets which are destined to node b are enqueued in queue 1. Similarly, the packets which are destined to node c are enqueued in queue 2. At each time slot, a packet from queue 1 is sent over link 1 with probability p_1 , and a packet from queue 2 is sent over link 2 with probability p_2 . Notice that links 1 and 2 will not be active at the same time.

For each node $n \in \mathcal{N}$, if the *receiver* node of link $i \in \mathcal{L}_n$ is within the *interference range* of another node $s \in \mathcal{N} \setminus \{n\}$, then any transmission by node s (i.e., transmission on any link $j \in \mathcal{L}_s$) interferes with transmission of link i . Those nodes which interfere with transmission of link i are denoted by set \mathcal{N}_i . Similar to [3], [6], we assume *infinite-backlog* in all nodes. For each node $n \in \mathcal{N}$, let r_i denote the average data rate for link $i \in \mathcal{L}_n$, which is a function of the persistent probability vector $\mathbf{p} = (p_i, \forall i \in \mathcal{L})$ of all links. We have [13]:

$$r_i(\mathbf{p}) = \gamma_i p_i \prod_{s \in \mathcal{N}_i} (1 - \sum_{j \in \mathcal{L}_s} p_j), \quad (1)$$

where γ_i denotes the peak data rate for link i (i.e., the rate achieved by link i if no node in set \mathcal{N}_i is active). Here, we assume that either the communication channel is to be fixed or it changes very slowly such that γ_i assume to be fixed for a long period of time for all $i \in \mathcal{L}$. In addition, to ensure that

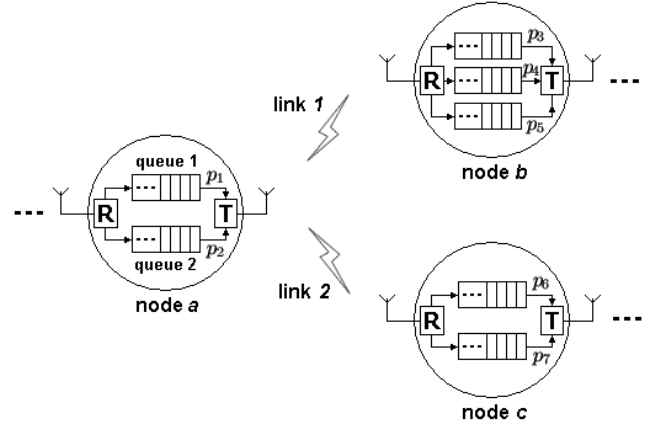


Fig. 1. A sample wireless ad-hoc network. We have, $\mathcal{L}_a = \{1, 2\}$, $\mathcal{L}_b = \{3, 4, 5\}$, and $\mathcal{L}_c = \{6, 7\}$. In node a , those packets which are assigned to be sent to node b (over link 1) are enqueued in queue 1. Similarly, those packets that are assigned to be sent to node c (over link 2) are enqueued in queue 2. R and T boxes represent receiver and transmitter units, respectively.

no link is starved, for each node n and any link $i \in \mathcal{L}_n$, we require $p_i \geq P_n^{\min} > 0$ and $L_n P_n^{\min} \leq P_n^{\max}$. Thus,

$$r_i(\mathbf{p}) \geq \gamma^{\min} P^{\min} (1 - P^{\max})^{N-1} > 0, \quad \forall i \in \mathcal{L}, \quad (2)$$

where $P^{\min} = \min_{n \in \mathcal{N}} P_n^{\min}$, $P^{\max} = \max_{n \in \mathcal{N}} P_n^{\max}$, and we have: $\gamma^{\min} = \min_{i \in \mathcal{L}} \gamma_i$.

Each link $i \in \mathcal{L}$ has a *utility function* which is an increasing and concave function of its rate r_i and indicates link i 's degree of satisfaction on its average data rate. The utility of link i is denoted by $u(r_i(\mathbf{p}))$, which is also a function of the persistent probabilities \mathbf{p} of all links. We are interested in finding the value of \mathbf{p} that solves the following *network utility maximization* (NUM) problem [14]:

$$\max_{\mathbf{p} \in \mathcal{P}} \sum_{i \in \mathcal{L}} u(r_i(\mathbf{p})), \quad (\text{NUM})$$

where the feasible persistent probability region is

$$\mathcal{P} = \left\{ \mathbf{p} : p_i \geq P_n^{\min}, \sum_{j \in \mathcal{L}_n} p_j \leq P_n^{\max}, \forall n \in \mathcal{N}, i \in \mathcal{L}_n \right\}$$

and the utility function is α -fair [15]. That is, for each link $i \in \mathcal{L}$, we have:

$$u(r_i) = \begin{cases} (1 - \alpha)^{-1} r_i^{1-\alpha}, & \text{if } \alpha \in (0, 1) \cup (1, \infty), \\ \log r_i, & \text{if } \alpha = 1. \end{cases} \quad (3)$$

Using (3), a wide range of efficient and fair allocations can be modeled. In particular, problem (NUM) reduces to throughput maximization with $\alpha \rightarrow 0$, to proportional fair allocation with $\alpha = 1$, to harmonic mean fair allocation with $\alpha = 2$, and to max-min fairness with $\alpha \rightarrow \infty$.

Although the objective function in problem (NUM) is concave in link rates $\mathbf{r} = (r_i, \forall i \in \mathcal{L})$, it is not concave in persistent probabilities \mathbf{p} due to the product form of the average rate in (1). Thus, finding the optimal solution for this *non-convex* and *coupled* optimization problem is difficult even in a centralized fashion. In this paper, we propose two algorithms which are able to find the optimal solution of problem (NUM) in a distributed fashion under verifiable sufficient conditions. In comparison with the existing algorithms, our algorithms do not require any synchronization, converge much faster, are

more robust to message delay and message loss, and support a wider range of α values in the utility function.

III. ALGORITHMS

In this section, we propose two distributed algorithms to solve optimization problem (NUM), one for *fully interfered* topologies in Section III-A and another one for *general* topologies in Section III-B. In both algorithms, each node n performs a *myopic* and local optimization, i.e., optimizing the *total* network utility by choosing the persistent probabilities of its own outgoing links, assuming others do not change theirs. Despite the complexity of the problem, we show that the solution of this local optimization problem can be obtained in closed-form, facilitated by limited message passing among nodes and a simple local sorting procedure.

A. Fully Interfered Network Topologies

We begin by considering a *fully interfered* topology, where all links interfere with each other. That is, for each $n \in \mathcal{N}$ and any $i \in \mathcal{L}_n$, the interference node set $\mathcal{N}_i = \mathcal{N} \setminus \{n\}$. This models some important practical wireless networks including *wireless personal area networks* where multiple wireless devices interact with each other over short distances, as well as *indoor wireless local area networks* where several wireless devices communicate with an access point and each other.

1) *Node n 's Local Optimization Problem*: For each wireless node n , let $\mathbf{p}_n = (p_i, \forall i \in \mathcal{L}_n)$ denote the persistent probabilities of its outgoing links. Also let $\mathbf{p}_{-n} = (p_j, \forall j \in \mathcal{L} \setminus \mathcal{L}_n)$ denote the persistent probabilities of all links *other than* the outgoing links of node n . Consider the following *local* and *myopic* optimization problem:

$$\max_{\mathbf{p}_n \in \mathcal{P}_n} \sum_{i \in \mathcal{L}} u(r_i(\mathbf{p}_n, \mathbf{p}_{-n})), \quad (\text{LOCAL-NUM})$$

where the feasible persistent probability region for node n is

$$\mathcal{P}_n = \{\mathbf{p}_n : \sum_{i \in \mathcal{L}_n} p_i \leq P_n^{\max}, p_i \geq P_n^{\min}, \forall i \in \mathcal{L}_n\}, \quad (4)$$

By solving problem (LOCAL-NUM), node n can select \mathbf{p}_n such that the *total* network utility is maximized *assuming* that \mathbf{p}_{-n} is fixed (i.e., none of the other nodes change their persistent probabilities). Clearly, nodes are *not* selfish in this case, and *cooperate* with each other. This is necessary to achieve optimal network performance in a distributed fashion.

Although problem (LOCAL-NUM) is difficult to solve, we can convert it to an equivalent and more illuminating representation. Its objective function in the case of fully interfered network topologies can be written as:

$$\begin{aligned} \sum_{i \in \mathcal{L}} u(r_i(\mathbf{p}_n, \mathbf{p}_{-n})) &= \left(\prod_{c \in \mathcal{N} \setminus \{n\}} (1 - \sum_{l \in \mathcal{L}_c} p_l) \right)^{1-\alpha} \\ & (1-\alpha) \left[\sum_{i \in \mathcal{L}_n} (\gamma_i p_i)^{1-\alpha} + (1 - \sum_{i \in \mathcal{L}_n} p_i)^{1-\alpha} \right. \\ & \left. \sum_{s \in \mathcal{N} \setminus \{n\}} \sum_{j \in \mathcal{L}_s} (\gamma_j p_j)^{1-\alpha} / (1 - \sum_{l \in \mathcal{L}_s} p_l)^{1-\alpha} \right]. \end{aligned}$$

Since the multiplicative term $(\prod_{c \in \mathcal{N} \setminus \{n\}} (1 - \sum_{l \in \mathcal{L}_c} p_l))^{1-\alpha}$ does not depend on the vector variable \mathbf{p}_n , problem (LOCAL-NUM) can be *equivalently* written as:

$$\max_{\mathbf{p}_n \in \mathcal{P}_n} \sum_{i \in \mathcal{L}_n} u(\gamma_i p_i) + v_n(\mathbf{p}_{-n}) u(1 - \sum_{i \in \mathcal{L}_n} p_i), \quad (5)$$

where

$$v_n(\mathbf{p}_{-n}) = \sum_{s \in \mathcal{N} \setminus \{n\}} (1 - \sum_{l \in \mathcal{L}_s} p_l)^{\alpha-1} \left(\sum_{j \in \mathcal{L}_s} (\gamma_j p_j)^{1-\alpha} \right). \quad (6)$$

Since $\sum_{i \in \mathcal{L}_n} u(\gamma_i p_i)$ and $u(1 - \sum_{i \in \mathcal{L}_n} p_i)$ are strictly concave functions with respect to \mathbf{p}_n , and $v_n(\mathbf{p}_{-n})$ is independent of \mathbf{p}_n , problem (5) is strictly concave in local variable vector \mathbf{p}_n . In other words, there exists a unique optimal solution of problem (5) and thus problem (LOCAL-NUM).

2) Closed-Form Solution of Problem (LOCAL-NUM):

Next, we show how to obtain a *closed-form* optimal solution for problem (5). Consider a node $n \in \mathcal{N}$ and the set of its outgoing links \mathcal{L}_n . We define a permutation, i_1, \dots, i_{L_n} , of the link indices in set \mathcal{L}_n such that for any j and l that satisfy $1 \leq j \leq l \leq L_n$, we have $\sqrt[\alpha]{\gamma_{i_j}^{\alpha-1}} \leq \sqrt[\alpha]{\gamma_{i_l}^{\alpha-1}}$. Thus, in the case of $\alpha \geq 1$, we have $\gamma_{i_1} \leq \dots \leq \gamma_{i_{L_n}}$, and in the case of $\alpha \in (0, 1)$, we have $\gamma_{i_1} \geq \dots \geq \gamma_{i_{L_n}}$. For example, let $\mathcal{L}_n = \{4, 7, 12\}$, $\gamma_4 = 18$ Mbps, $\gamma_7 = 24$ Mbps, and $\gamma_{12} = 6$ Mbps. If $\alpha \geq 1$, then we have $i_1 = 12$, $i_2 = 4$, and $i_3 = 7$. On the other hand, if $\alpha \in (0, 1)$, then $i_1 = 7$, $i_2 = 4$, and $i_3 = 12$.

Let σ denote the *smallest* number in the set $\{0, \dots, L_n - 1\}$ such that we have:

$$\frac{1}{P_n^{\min}} - L_n + \sigma \leq \sum_{l=1}^{\sigma} \sqrt[\alpha]{\left(\frac{\gamma_{i_{\sigma+1}}}{\gamma_{i_l}}\right)^{\alpha-1}} + \sqrt[\alpha]{\gamma_{i_{\sigma+1}}^{\alpha-1}} v_n(\mathbf{p}_{-n}). \quad (7)$$

We can show that (see Appendix A), if condition (7) holds for σ , then it also holds for $\sigma + 1$. We define the set $\mathcal{B}_n = \{i_{\sigma+1}, \dots, i_{L_n}\}$, with its size $B_n = |\mathcal{B}_n| = L_n - \sigma$. Notice that if condition (7) does not hold for any $\sigma \in \{0, \dots, L_n - 1\}$, then we set $\mathcal{B}_n = \{\}$ and $B_n = 0$.

Similarly, let ς denote the *smallest* number in the set $\{0, \dots, L_n - 1\}$ such that we have:

$$P_n^{\max} / P_n^{\min} - L_n + \varsigma \leq \sum_{l=1}^{\varsigma} \sqrt[\alpha]{(\gamma_{i_{\varsigma+1}} / \gamma_{i_l})^{\alpha-1}}. \quad (8)$$

Again, we can show that (see Appendix A), if condition (8) holds for ς , then it also holds for $\varsigma + 1$. We define $\mathcal{C}_n = \{i_{\varsigma+1}, \dots, i_{L_n}\}$, and its size $C_n = |\mathcal{C}_n| = L_n - \varsigma$. If condition (8) does not hold for any $\varsigma \in \{0, \dots, L_n - 1\}$, then we set $\mathcal{C}_n = \{\}$ with its size $C_n = 0$.

We now define $\mathcal{A}_n = \mathcal{B}_n \cup \mathcal{C}_n$ with its size $A_n = |\mathcal{A}_n| = L_n - \kappa$ where $\kappa = \min\{\sigma, \varsigma\}$. In fact, $\mathcal{A}_n = \{i_{\kappa+1}, \dots, i_{L_n}\}$. Depending on the value of $v_n(\mathbf{p}_{-n})$, either $\mathcal{A}_n = \mathcal{B}_n$ or $\mathcal{A}_n = \mathcal{C}_n$ (see Appendix A). Using \mathcal{A}_n , the closed-form solution of problem (5) can be obtained as follows.

Theorem 1: For each node $n \in \mathcal{N}$, the unique optimal solution of problem (5) is $\mathbf{p}_n^*(\mathbf{p}_{-n}) = \mathbf{f}_n(\mathbf{p}_{-n}) = (f_i(\mathbf{p}_{-n}), \forall i \in \mathcal{L}_n)$, where for each link $i \in \mathcal{L}_n$, the mapping $f_i(\mathbf{p}_{-n})$ is defined as:

$$f_i(\mathbf{p}_{-n}) = \begin{cases} P_n^{\min}, & \text{if } i \in \mathcal{A}_n, \\ \left[\frac{1}{\alpha \sqrt[\alpha]{\gamma_i^{\alpha-1}}} \frac{1 - A_n P_n^{\min}}{w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}} \right]_{P_n^{\min}}^{\frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n}}, & \text{if } i \in \mathcal{L}_n \setminus \mathcal{A}_n, \end{cases} \quad (9)$$

where $[x]_b^a = \max[\min[x, a], b]$, and

$$w_n = \sum_{j \in \mathcal{L}_n \setminus \mathcal{A}_n} \sqrt[\alpha]{(1/\gamma_j)^{\alpha-1}}. \quad (10)$$

The proof of Theorem 1 is available in Appendix A. The key is to show that $f_n(\mathbf{p}_{-n})$ satisfies the *necessary* and *sufficient* Karush-Kuhn-Tucker (KKT) optimality conditions [16, pp. 244]. Since problems (LOCAL-NUM) and (5) are equivalent, $\mathbf{p}_n^*(\mathbf{p}_{-n})$ is also the unique global optimal solution of problem (LOCAL-NUM). We notice that, regardless of the selected system parameters and the value of v_n , for each $i \in \mathcal{L}_n \setminus \mathcal{A}_n$, the upper bound $(P_n^{\max} - A_n P_n^{\min}) / (\sqrt[\alpha]{\gamma_i^{\alpha-1} w_n})$ in (9) is always greater than or equal to the lower bound P_n^{\min} .

Clearly, to compute $f_i(\mathbf{p}_{-n})$ in (9), node n needs to obtain $v_n(\mathbf{p}_{-n})$. If each node s announces a message m_s where

$$m_s = \left(1 - \sum_{j \in \mathcal{L}_s} p_j\right)^{\alpha-1} \left(\sum_{j \in \mathcal{L}_s} (\gamma_j p_j)^{1-\alpha}\right), \quad \forall s \in \mathcal{N} \setminus \{n\},$$

then node n can compute $v_n(\mathbf{p}_{-n}) = \sum_{s \in \mathcal{N} \setminus \{n\}} m_s$. This motivates us to propose our first algorithm¹.

3) *A Distributed MAC Algorithm*: Our proposed distributed random access algorithm is given in Algorithm 1. In this algorithm, each node $n \in \mathcal{N}$, regardless of how many outgoing links it has, announces only a *single* message m_n . All wireless nodes choose the persistent probabilities of their outgoing links based on the messages that they receive from other nodes.

Algorithm 1 - Used by each wireless node $n \in \mathcal{N}$ in a fully interfered topology.

- 1: Allocate memory for messages $\mathbf{m} = (m_1, \dots, m_N)$.
 - 2: Randomly choose $p_i \geq P_n^{\min} > 0$ for each link $i \in \mathcal{L}_n$ such that $\sum_{i \in \mathcal{L}_n} p_i \leq P_n^{\max} \leq 1$.
 - 3: Randomly choose $m_s > 0$ for all $s \in \mathcal{N}$.
 - 4: **repeat**
 - 5: Transmit on outgoing link $i \in \mathcal{L}_n$ with probability p_i .
 - 6: **if** $t \in T_{n,p}$ **then**
 - 7: Set $\mathcal{A}_n = \text{getA}(n, v_n(\mathbf{p}_{-n}), L_n, \gamma_1, \dots, \gamma_{L_n})$.
 - 8: Set $p_i = P_n^{\min}$ for all $i \in \mathcal{A}_n$.
 - 9: Set $p_i = \left[\frac{1}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})})} \right] \frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n}$
 - 10: for all $i \in \mathcal{L}_n \setminus \mathcal{A}_n$.
 - 11: **if** $t \in T_{n,qm}$ **then**
 - 12: Set $m_n = (1 - \sum_{i \in \mathcal{L}_n} p_i)^{\alpha-1} \left(\sum_{i \in \mathcal{L}_n} \left(\frac{1}{\gamma_i p_i}\right)^{\alpha-1}\right)$.
 - 13: Broadcast m_n .
 - 14: **if** a message is received **then** Update \mathbf{m} .
 - 15: **until** node n decides to leave the network.
-

The persistent probabilities and messages are *asynchronously* updated. Let $T_{n,p}$ and $T_{n,m}$ be two unbounded sets of time slots at which node n updates \mathbf{p}_n and m_n , respectively. We assume that the asynchronism of the updates is bounded; i.e., there exists a finite H (called *asynchronism measure* [17]) such that:

$$\forall t_1 \in T_{n,p}, \exists t_2 \in T_{n,p} \text{ such that } t_2 - t_1 \leq H, \quad (11)$$

$$\forall t_3 \in T_{n,m}, \exists t_4 \in T_{n,m} \text{ such that } (t_4 - t_3) + D \leq H, \quad (12)$$

¹Similar way of constructing messages based on KKT conditions has been considered in [12] for distributed power control. On the other hand, we mainly use contraction and monotone mapping to prove the properties of our proposed algorithms here, instead of the supermodular game theory used in [12].

- 1: **function** $\text{getA}(n, v_n(\mathbf{p}_{-n}), L_n, \gamma_1, \dots, \gamma_{L_n})$
- 2: Set $\mathcal{B}_n = \{\}$ and $\mathcal{C}_n = \{\}$.
- 3: **if** $\alpha \geq 1$ **then** Set i_1, \dots, i_{L_n} so that $\gamma_{i_1} \leq \dots \leq \gamma_{i_{L_n}}$.
- 4: **else** Set i_1, \dots, i_{L_n} so that $\gamma_{i_1} \geq \dots \geq \gamma_{i_{L_n}}$.
- 5: **for** $\sigma = 0, \dots, L_n - 1$ **do**
- 6: **if** $1/P_n^{\min} - L_n + \sigma \leq \sum_{l=1}^{\sigma} \sqrt[\alpha]{(\gamma_{i_{\sigma+1}}/\gamma_{i_l})^{\alpha-1}} + \sqrt[\alpha]{\gamma_{i_{\sigma+1}}^{\alpha-1} v_n(\mathbf{p}_{-n})}$ **then**
- 7: Set $\mathcal{B}_n = \{i_{\sigma+1}, \dots, i_{L_n}\}$.
- 8: Break.
- 9: **end for**
- 10: **for** $\varsigma = 0, \dots, L_n - 1$ **do**
- 11: **if** $P_n^{\max}/P_n^{\min} - L_n + \varsigma \leq \sum_{l=1}^{\varsigma} \sqrt[\alpha]{\left(\frac{\gamma_{i_{\varsigma+1}}}{\gamma_{i_l}}\right)^{\alpha-1}}$ **then**
- 12: Set $\mathcal{C}_n = \{i_{\varsigma+1}, \dots, i_{L_n}\}$.
- 13: Break.
- 14: **end for**
- 15: Set $\mathcal{A}_n = \mathcal{B}_n \cup \mathcal{C}_n$.
- 16: **return** \mathcal{A}_n .

where D denotes an upper bound on *communication delay*. From (11), each node updates the persistent probabilities of its outgoing links at least once during any time interval of length H slots. From (12), the information used by each node is outdated by at most H time slots. We notice that H can be arbitrarily large as long as it is bounded. The exact value of H is *not* important and needs *not* be known by all nodes.

Compared with the distributed MAC algorithms proposed in the literature, Algorithm 1 has several distinct features: (i) less explicit message passing is needed (e.g., in the subgradient-based algorithm proposed in [3], each node needs to announce *two* messages), (ii) asynchronous updates with arbitrarily finite delay, which minimizes the coordination overhead and allows maximum heterogeneity among nodes, and (iii) does not use any stepsizes, which avoids the slow convergence problem due to small stepsizes in the commonly used subgradient methods.

We note that if Algorithm 1 has a fixed point, then every node n achieves the optimal solution of problem (LOCAL-NUM) and each node will not change its persistent probability vector. In Section IV-A, we show that this fixed point corresponds to the global optimal solution of problem (NUM) under proper sufficient conditions, and Algorithm 1 converges to such a fixed point with fast and robust performance.

B. General Network Topologies

We now consider the general case, where each node is within the interference range of an *arbitrary subset* of the other nodes. For each node $n \in \mathcal{N}$ and any of its outgoing links $i \in \mathcal{L}_n$, the set of nodes that interfere with link i is an arbitrary subset of all nodes, i.e., $\mathcal{N}_i \subseteq \mathcal{N} \setminus \{n\}$. In this case, the objective function of problem (LOCAL-NUM), i.e., the summation $\sum_{i \in \mathcal{L}} u(r_i(\mathbf{p}_n, \mathbf{p}_{-n}))$, can be written as:

$$(1 - \alpha)^{-1} \times \left[\sum_{i \in \mathcal{L}_n} \left((\gamma_i \prod_{c \in \mathcal{N}_i} (1 - \sum_{l \in \mathcal{L}_c} p_l)) p_i \right)^{1-\alpha} + (1 - \sum_{i \in \mathcal{L}_n} p_i)^{1-\alpha} \sum_{s \in \mathcal{N} \setminus \{n\}} \sum_{j \in \mathcal{L}_s | n \in \mathcal{N}_j} (\gamma_j p_j \prod_{c \in \mathcal{N}_j \setminus \{n\}} (1 - \sum_{l \in \mathcal{L}_c} p_l))^{1-\alpha} + \sum_{s \in \mathcal{N} \setminus \{n\}} \sum_{j \in \mathcal{L}_s | n \notin \mathcal{N}_j} (\gamma_j p_j \prod_{c \in \mathcal{N}_j} (1 - \sum_{l \in \mathcal{L}_c} p_l))^{1-\alpha} \right].$$

Since the last term in the bracket does not depend on \mathbf{p}_n , problem (LOCAL-NUM) can be *equivalently* written as:

$$\max_{\mathbf{p}_n \in \mathcal{P}_n} \sum_{i \in \mathcal{L}_n} u(\gamma'_i(\mathbf{p}_{-n})p_i) + v'_n(\mathbf{p}_{-n}) u(1 - \sum_{i \in \mathcal{L}_n} p_i), \quad (13)$$

where for outgoing each link $i \in \mathcal{L}_n$, auxiliary term γ'_i is defined as $\gamma'_i(\mathbf{p}_{-n}) = \gamma_i \prod_{s \in \mathcal{N}_i} (1 - \sum_{l \in \mathcal{L}_s} p_l)$ and

$$v'_n(\mathbf{p}_{-n}) = \sum_{s \in \mathcal{N} \setminus \{n\}} \sum_{j \in \mathcal{L}_n | n \in \mathcal{N}_j} (\gamma_j p_j \prod_{s \in \mathcal{N}_j \setminus \{n\}} (1 - \sum_{l \in \mathcal{L}_s} p_l))^{1-\alpha}.$$

Notice that $\gamma'_i(\mathbf{p}_{-n})$ does not represent the peak data rate of wireless link i . We can show that problem (13) is strictly concave in vector \mathbf{p}_n and has a unique optimal solution.

The closed-form solution of problem (13) can be obtained similarly as that of problem (5) in the case of fully interfered network topologies. For the outgoing link set \mathcal{L}_n of node n , we can define a permutation of link indices in this set, i_1, \dots, i_{L_n} , such that for any j and l that satisfy $1 \leq j \leq l \leq L_n$, we have $\sqrt[\alpha]{\gamma'_{i_j}(\mathbf{p}_{-n})^{\alpha-1}} \leq \sqrt[\alpha]{\gamma'_{i_l}(\mathbf{p}_{-n})^{\alpha-1}}$. In the case of $\alpha \geq 1$, we have $\gamma'_{i_1} \leq \dots \leq \gamma'_{i_{L_n}}$. If $\alpha \in (0, 1)$, we have $\gamma'_{i_1} \geq \dots \geq \gamma'_{i_{L_n}}$. Let σ' denote the *smallest* value in $\{0, \dots, L_n - 1\}$ such that

$$\frac{1}{P_n^{\min}} - L_n + \sigma' \leq \sum_{l=1}^{\sigma'} \sqrt[\alpha]{(\gamma'_{i_{\sigma'+l}}(\mathbf{p}_{-n})/\gamma'_{i_l}(\mathbf{p}_{-n}))^{\alpha-1}} + \sqrt[\alpha]{\gamma'_{i_{\sigma'+1}}(\mathbf{p}_{-n})^{\alpha-1}} v'_n(\mathbf{p}_{-n}). \quad (14)$$

Similarly, let ς' denote the *smallest* value in $\{0, \dots, L_n - 1\}$ such that we have:

$$\frac{P_n^{\max}}{P_n^{\min}} - L_n + \varsigma' \leq \sum_{l=1}^{\varsigma'} \sqrt[\alpha]{\left(\gamma'_{i_{\varsigma'+l}}(\mathbf{p}_{-n})/\gamma'_{i_l}(\mathbf{p}_{-n})\right)^{\alpha-1}}. \quad (15)$$

We define $\mathcal{B}'_n = \{i_{\sigma'+1}, \dots, i_{L_n}\}$, with its size $B'_n = |\mathcal{B}'_n| = L_n - \sigma'$. If condition (14) does not hold for any $\sigma' \in \{0, \dots, L_n - 1\}$, then we set $\mathcal{B}'_n = \{\}$ with $B'_n = 0$. Similarly, we define $\mathcal{C}'_n = \{i_{\varsigma'+1}, \dots, i_{L_n}\}$, with its size $C'_n = |\mathcal{C}'_n| = L_n - \varsigma'$. If condition (15) does not hold for any $\varsigma' \in \{0, \dots, L_n - 1\}$, then we set $\mathcal{C}'_n = \{\}$ with $C'_n = 0$. Given \mathcal{B}'_n and \mathcal{C}'_n , we define $\mathcal{A}'_n = \mathcal{B}'_n \cup \mathcal{C}'_n$ with its size $A'_n = |\mathcal{A}'_n| = L_n - \kappa'$ where $\kappa' = \min\{\sigma', \varsigma'\}$.

Theorem 2: For each node $n \in \mathcal{N}$, the global optimal solution of problem (13) is $\mathbf{p}_n^*(\mathbf{p}_{-n}) = \mathbf{f}'_n(\mathbf{p}_{-n}) = (f'_i(\mathbf{p}_{-n}), \forall i \in \mathcal{L}_n)$, where for each $i \in \mathcal{L}_i$, function $f'_i(\mathbf{p}_{-n})$ is defined as:

$$f'_i(\mathbf{p}_{-n}) = \begin{cases} P_n^{\min}, & \text{if } i \in \mathcal{A}'_n, \\ \left[\frac{1}{\sqrt[\alpha]{\gamma'_i(\mathbf{p}_{-n})^{\alpha-1}}} \frac{1 - A'_n P_n^{\min}}{w'_n(\mathbf{p}_{-n}) + \sqrt[\alpha]{v'_n(\mathbf{p}_{-n})}} \right]_{P_n^{\min}}^{\frac{P_n^{\max} - A'_n P_n^{\min}}{\sqrt[\alpha]{\gamma'_i(\mathbf{p}_{-n})^{\alpha-1}} w'_n(\mathbf{p}_{-n})}}, & \text{otherwise,} \end{cases} \quad (16)$$

with $w'_n(\mathbf{p}_{-n}) = \sum_{j \in \mathcal{L}_n \setminus \mathcal{A}'_n} \sqrt[\alpha]{(1/\gamma'_j(\mathbf{p}_{-n}))^{\alpha-1}}$.

The proof of Theorem 2 is similar to that of Theorem 1. Eq. (16) provides the optimal solution for problem (LOCAL-NUM) for general topologies, which include the fully interfered case as a special case. We can define node s ' messages:

$$q_s = 1 - \sum_{j \in \mathcal{L}_j} p_j, \quad (17)$$

$$m_{s,n} = \sum_{j \in \mathcal{L}_s | n \in \mathcal{N}_j} 1/(\gamma_j p_j \prod_{c \in \mathcal{N}_j \setminus \{n\}} q_c)^{\alpha-1}. \quad (18)$$

Then $v'_n(\mathbf{p}_{-n}) = \sum_{s \in \mathcal{N} \setminus \{n\}} m_{s,n}$ and $\gamma'_i(\mathbf{p}_{-n}) = \gamma_i \prod_{s \in \mathcal{N}_i} q_s$ for all $i \in \mathcal{L}_n$. Message q_s simply denotes the probability that node s remains silent at a time slot. Also note that for each node $n \neq s$, if there does not exist any $j \in \mathcal{L}_s$ such that $n \in \mathcal{N}_j$, then $m_{s,n} = 0$ (i.e., node n does not cause interference to any outgoing link of node s).

Algorithm 2 - Used by each node $n \in \mathcal{N}$ in a general topology.

- 1: Allocate memory for $\mathbf{m} = (m_{1,n}, \dots, m_{N,n})$ and $\mathbf{q} = (q_1, \dots, q_N)$.
- 2: Randomly choose $p_i \geq P_n^{\min} > 0$ for all $i \in \mathcal{L}_n$ such that $\sum_{i \in \mathcal{L}_n} p_i \leq P_n^{\max} \leq 1$.
- 3: Randomly choose $m_{s,n} > 0$ and $q_s \in (0, 1)$ for all $s \in \mathcal{N}$.
- 4: **repeat**
- 5: Transmit on outgoing link $i \in \mathcal{L}_n$ with probability p_i .
- 6: **if** $t \in T_{n,p}$ **then**
- 7: Set $\mathcal{A}'_n = \text{getA}(n, v'_n(\mathbf{p}_{-n}), L_n, \gamma'_1, \dots, \gamma'_{L_n})$.
- 8: Set $p_i = P_n^{\min}$ for all $i \in \mathcal{A}'_n$.
- 9: Set $p_i = \left[\frac{1}{\sqrt[\alpha]{\gamma'_i(\mathbf{p}_{-n})^{\alpha-1}}} \frac{1 - A'_n P_n^{\min}}{w'_n(\mathbf{p}_{-n}) + \sqrt[\alpha]{v'_n(\mathbf{p}_{-n})}} \right]_{P_n^{\min}}^{\frac{P_n^{\max} - A'_n P_n^{\min}}{\sqrt[\alpha]{\gamma'_i(\mathbf{p}_{-n})^{\alpha-1}} w'_n(\mathbf{p}_{-n})}}$
- 10: for all $i \in \mathcal{L}_n \setminus \mathcal{A}'_n$.
- 11: **if** $t \in T_{n,qm}$ **then**
- 12: Set $q_n = 1 - \sum_{i \in \mathcal{L}_n} p_i$.
- 13: Set $m_{n,s} = \sum_{i \in \mathcal{L}_n : s \in \mathcal{N}_i} 1/(\gamma_i p_i \prod_{c \in \mathcal{N}_i \setminus \{s\}} q_c)^{\alpha-1}$ for any $s \neq n$.
- 14: Inform $m_{n,s}$ to all $s \in \cup_{i \in \mathcal{L}_n} \mathcal{N}_i$.
- 15: Inform q_n to all $s \in \mathcal{N} \setminus \{n\}$ if $\exists j \in \mathcal{L}_s$ and $n \in \mathcal{N}_j$.
- 16: **if** a message is received **then** Update \mathbf{m} and \mathbf{q} .
- 17: **until** node n decides to leave the network.

Our second proposed algorithm works for any general topology and is shown in Algorithm 2. In this algorithm, each node $n \in \mathcal{N}$ informs $m_{n,s}$ to all nodes s whose transmissions interfere with transmissions of at least one of the outgoing links of node n . It also informs q_n to all nodes s whose outgoing transmissions is interfered by transmissions from node n . All nodes then choose the persistent probabilities of their outgoing links based on the received messages from other nodes. In Algorithm 2, $T_{n,p}$ and $T_{n,qm}$ are two unbounded sets of time slots at which node n updates \mathbf{p}_n and announces q_n , and $m_{n,s}$ for all $s \neq n$, respectively. The assumptions on asynchronism measure are the same as those in Algorithm 1. We show in Section IV-B that for *any* topology, the fixed point of Algorithm 2 also corresponds to the global optimal solution of the *non-convex* problem (NUM) under proper conditions.

In comparison with the prior algorithms in the literature (e.g., [3]), Algorithms 1 and 2 are more *robust*, converge *faster*, and require *less signalling*. We further discuss the properties of our proposed algorithms in Sections IV and V.

IV. CONVERGENCE, OPTIMALITY, AND ROBUSTNESS

In this section, we prove the convergence, optimality and robustness properties of our algorithms by using results in distributed computation [17] and non-linear optimization [18].

A. Fully Interfered Network Topologies

Here we study Algorithm 1 which was proposed to solve problem (LOCAL-NUM) in a fully interfered topology. We first show that if Algorithm 1 has a unique fixed point, then it will globally converge to that fixed point. After that, we provide the conditions under which the uniqueness of the fixed point of Algorithm 1 is guaranteed. We also show that such unique fixed point corresponds to the unique global optimal solution of non-convex optimization problem (NUM).

Definition 1: A vector mapping $\mathbf{g}(\cdot)$ is *monotone increasing* [17, pp. 191] if for any $\tilde{\mathbf{p}}, \hat{\mathbf{p}} \in \mathcal{P}$ such that $\tilde{\mathbf{p}} \preceq \hat{\mathbf{p}}$, we have $\mathbf{g}(\tilde{\mathbf{p}}) \preceq \mathbf{g}(\hat{\mathbf{p}})$, where the inequalities are interpreted as coordinate-wise. A vector mapping $\mathbf{g}(\cdot)$ is *monotone decreasing* if for any $\tilde{\mathbf{p}}, \hat{\mathbf{p}} \in \mathcal{P}$ such that $\tilde{\mathbf{p}} \preceq \hat{\mathbf{p}}$, we have $\mathbf{g}(\tilde{\mathbf{p}}) \succeq \mathbf{g}(\hat{\mathbf{p}})$.

We define $\mathbf{f}(\mathbf{p}) = (\mathbf{f}_n(\mathbf{p}), \forall n \in \mathcal{N})$, where $\mathbf{f}_n(\mathbf{p})$ is as in Theorem 1. Recall that a fixed point of mapping $\mathbf{f}(\mathbf{p})$ is also a fixed point of Algorithm 1. We can show that²:

Proposition 1: $\mathbf{f}(\mathbf{p})$ is monotone increasing if $\alpha \geq 1$, and is monotone decreasing if $\alpha \leq 1$.

The proof of Proposition 1 is given in Appendix B. This enables us to show the following:

Theorem 3: Suppose \mathbf{f} has a unique fixed point \mathbf{p}^* . Starting from any initial point $\mathbf{p} \in \mathcal{P}$, Algorithm 1 converges to \mathbf{p}^* .

The proof of Theorem 3 is available in Appendix C. The key idea is to show that the monotone mapping $\mathbf{f}(\mathbf{p})$ satisfies *synchronous convergence* and *box* conditions; thus, *asynchronous convergence theorem* [17, pp. 431] is applicable. Theorem 3 is general and applies to any choice of system parameters. It only requires that $\mathbf{f}(\cdot)$ has a unique fixed point. Next, we will show that not only Algorithm 1 has a unique fixed point under mild technical conditions, the fixed point is indeed the global optimal solution of problem (NUM).

Let \mathcal{F} denote the set of fixed points of Algorithm 1. For each $\mathbf{p}^* \in \mathcal{F}$ and any link $i \in \mathcal{L}_n$, we have $p_i^* = f_i(\mathbf{p}_{-n}^*)$. We also let \mathcal{S} denote the set of *stationary points* [18, pp. 194] of problem (NUM). Note that all local (and global) optimal solutions of problem (NUM) belong to set \mathcal{S} .

Theorem 4: $\mathcal{F} = \mathcal{S}$.

The proof of Theorem 4 is available in Appendix D. From Theorems 3 and 4, we have:

Corollary 1: If either \mathcal{S} or \mathcal{F} is a *singleton* set (i.e., it has one element), then Algorithm 1 asynchronously converges to the unique global optimal solution of problem (NUM).

In [3], it has been shown that the set of stationary points \mathcal{S} is a singleton set for all $\alpha \geq 1$. They used logarithmic mapping and transformed problem (NUM) to an equivalent convex problem and showed that it has a unique stationary point. However, this transformation does not work if $\alpha \in (0, 1)$. That is the reason the algorithm proposed in [3] does not support the α -fair utility functions with $\alpha \in (0, 1)$. Here we are able to provide sufficient conditions under which the non-convex problem (NUM) has a unique optimal solution with $\alpha \in (0, 1)$. To achieve this, we need to define a different mapping concept.

Definition 2: A mapping function $\mathbf{g}(\mathbf{p})$ is a *contraction mapping* [17, pp. 181] if there exists a constant $\zeta \in (0, 1)$,

such that $\|\mathbf{g}(\tilde{\mathbf{p}}) - \mathbf{g}(\hat{\mathbf{p}})\| \leq \zeta \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|$ for all $\tilde{\mathbf{p}}, \hat{\mathbf{p}} \in \mathcal{P}$. Here $\|\cdot\|$ is some vector norm. A contraction mapping has unique fixed point [17, pp. 183].

Theorem 5: Consider the case where $\alpha \in (0, 1)$. Set \mathcal{F} is a singleton if the following holds:

$$\left(\frac{1-\alpha}{\alpha} \Psi \Phi(V^{\min}, V^{\max})\right)^2 \left(\frac{\gamma^{\max}}{\gamma^{\min}} \Gamma\right)^{1-\alpha} \left(\Omega - \frac{1}{L/L^{\min} - 1}\right) < 1, \quad (19)$$

where $L^{\min} = \min_{n \in \mathcal{N}} L_n$, $L^{\max} = \max_{n \in \mathcal{N}} L_n$, $\Gamma = (P^{\max}(1 - P^{\min})) / (P^{\min}(1 - P^{\max}))$, $\Psi = L^{\max} / (1 - P^{\max}) + 1/P^{\min}$, $\Omega = \sum_{n \in \mathcal{N}} 1/(L/L_n - 1)$, $V^{\min} = (N - 1) (\gamma^{\max}(1/P^{\min} - 1) / \gamma^{\min})^{\alpha-1}$, $V^{\max} = (N - 1) (\gamma^{\min}(1/P^{\min} - 1) / \gamma^{\max})^{\alpha-1}$, and we have:

$$\Phi(V^{\min}, V^{\max}) = \begin{cases} \frac{(V^{\max})^{1/\alpha}}{(1+(V^{\max})^{1/\alpha})^2}, & \text{if } V^{\max} \leq 1, \\ \frac{(V^{\min})^{1/\alpha}}{(1+(V^{\min})^{1/\alpha})^2}, & \text{if } V^{\min} \geq 1, \\ 0.25, & \text{otherwise.} \end{cases} \quad (20)$$

The proof of Theorem 5 is available in Appendix E. The key is to show that if (19) holds, then mapping \mathbf{f} is not only a monotone mapping, but also an l_2 -norm contraction mapping.

In general, all the terms in (19), except Φ , are independent of the number of nodes N . The value of Φ can be arbitrarily close to 0 if N is large enough. Therefore,

Corollary 2: For any $\alpha \in (0, 1)$ and any choice of other parameters, there exists a positive integer \hat{N} such that Algorithm 1 has a unique fixed point if the number of nodes $N > \hat{N}$.

Theorem 5 provides practical bounds on system parameters that guarantee the uniqueness of the fixed point. For example, consider the IEEE 802.11a standard where $\gamma^{\min} = 6$ Mbps and $\gamma^{\max} = 54$ Mbps. In Fig. 2, we plot the sufficient conditions on upper bounds of P^{\max} and lower bounds of P^{\min} for utility parameter $\alpha \in [0.1, 0.9]$ and number of nodes $N \in [2, 100]$, where each node has one outgoing link. As we can see, the difference between the lower and upper bounds increases as α or N increases, indicating the convergence condition is less restrictive. In many cases, convergence of Algorithm 1 can be obtained even when the sufficient condition (19) is not satisfied. For example, it is easy to numerically verify that for $N = 2$, problem (NUM) with $\alpha \in (0.5, 1)$ has a unique global optimal solution with any choice of system parameters.

Theorems 3 to 5 together show that Algorithm 1 asynchronously converges to the unique global optimal solution of the problem (NUM) when either $\alpha \in (0, 1)$ (under condition (19)) or $\alpha \geq 1$ (with any system parameters). In particular, Algorithm 1 works properly under delayed or even occasionally lost messages. To have a better understanding on how the system behaves with message loss, consider M consecutive messages announced by an arbitrary node n . The first $M - 1$ messages are lost (e.g., due to collisions) while the last message is properly received by all other nodes $s \in \mathcal{N} \setminus \{n\}$. In this case, all derived results will go through with another asynchronism measure. Let $\hat{H} = MH$. Since H and M are bounded, \hat{H} is also bounded. Considering \hat{H} as the new asynchronism measure, Theorems 3 to 5 can still be applied. Thus, convergence and optimality of Algorithm 1 are still guaranteed. Interestingly, this robust behavior is accompanied with fast convergence speed as shown in Section V.

²For the special case when $\alpha = 1$, vector mapping $\mathbf{f}(\mathbf{p}_{-n})$ is a *constant* mapping as its value does not depend on \mathbf{p}_{-n} . Thus, the convergence proof for Algorithm 1 is trivial if $\alpha = 1$. Same statement is true for Algorithm 2.

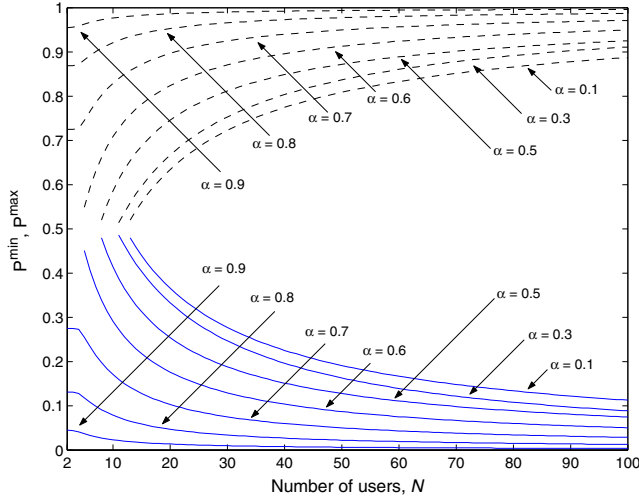


Fig. 2. Sufficient conditions on the upper bounds of P^{\max} and lower bounds of P^{\min} for $\alpha \in [0.1, 0.9]$ and $N \in [2, 100]$ when Algorithm 1 is being used and each node has one outgoing link. Solid lines represent lower bounds on P^{\min} and dashed lines represent upper bounds on P^{\max} .

B. General Network Topologies

Consider vector mapping $\mathbf{f}'(\mathbf{p}) = (\mathbf{f}'_n(\mathbf{p}_{-n}), \forall n \in \mathcal{N})$, where $\mathbf{f}'_n(\mathbf{p}_{-n})$ is defined in Theorem 2. We denote the set of fixed points of mapping $\mathbf{f}'(\mathbf{p})$ by \mathcal{F}' , which is the set of fixed points of Algorithm 2. We also denote the set of stationary points of problem (NUM) by \mathcal{S}' in this case.

Theorem 6: $\mathcal{F}' = \mathcal{S}'$.

The proof of Theorem 6 is similar to that of Theorem 4. If $\alpha \geq 1$, then from [3, Lemma 1] we know that stationary point set \mathcal{S}' is a singleton set. Together with Theorem 6, we have:

Corollary 3: If $\alpha \geq 1$ and Algorithm 2 converges, it converges to the unique global optimal solution of problem (NUM).

If $\alpha \in (0, 1)$, we can use the same idea of Theorem 5 and obtain sufficient conditions to assure that the stationary point set \mathcal{S}' is a singleton set. We first notice that since not all links interfere with each other, for each node $n \in \mathcal{N}$ and any link $i \in \mathcal{L}_n$, function f'_i may only depend on a small subset of entries in vector \mathbf{p}_{-n} . For example, consider the *chain topology* in Fig. 3, where the interferences are within *one hop*. For each node i , $\mathcal{N}_i = \{n+1, n+2\}$. In this figure, f'_i only depends on $(p_{i-5}, \dots, p_{i-1}, p_{i+2}, \dots, p_{i+4})$. Notice that γ'_i depends on $(p_{i+1}, \dots, p_{i+4})$ and γ'_{i-1} depends on $(p_{i-5}, \dots, p_{i-2})$. Thus w'_n depends on $(p_{i-5}, \dots, p_{i-1}, p_{i+2}, \dots, p_{i+4})$. In addition, $m_{n-2,n}$, $m_{n-1,n}$, $m_{n+1,n}$, and $m_{n+2,n}$ depend on $(p_{i-4}, p_{i-3}, p_{i-2})$, $(p_{i-2}, p_{i+1}, p_{i+2})$, $(p_{i-3}, p_{i-2}, p_{i+1})$, and $(p_{i+1}, p_{i+2}, p_{i+3})$, respectively. Thus, v'_n also depends on $(p_{i-5}, \dots, p_{i-1}, p_{i+2}, \dots, p_{i+4})$. We define set $\mathcal{X}_i = \{i-5, \dots, i-1, i+2, \dots, i+4\}$ as the *dependency set* for link i . Similarly, we can define \mathcal{X}_i for all $i \in \mathcal{L}$ in any arbitrary topology. That is, for any $i, j \in \mathcal{L}$, we have $j \in \mathcal{X}_i$ if and only if p_j appears in the formulation of f'_i . Let $X_i = |\mathcal{X}_i|$ denote the size of set \mathcal{X}_i . We define $X^{\max} = \max_{i \in \mathcal{L}} X_i$. As an example, for the chain topology in Fig. 3, $X^{\max} = 8$.

Theorem 7: For any general topology, the fixed point set \mathcal{F}' is a singleton if $\alpha \in (0, 1)$ and

$$\frac{1-\alpha}{\alpha} X^{\max} \Lambda \Phi(Z^{\min}, Z^{\max}) < 1, \quad (21)$$

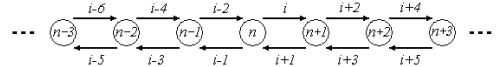


Fig. 3. A chain topology with one hop interference. Arrows indicate the direction of the corresponding unidirectional wireless links. Function f'_i only depends on $(p_{i-5}, \dots, p_{i-1}, p_{i+2}, \dots, p_{i+4})$. Notice that γ'_i , γ'_{i-1} , $m_{n-2,n}$, $m_{n-1,n}$, $m_{n+1,n}$, and $m_{n+2,n}$ depend on $(p_{i+1}, \dots, p_{i+4})$, $(p_{i-5}, \dots, p_{i-2})$, $(p_{i-4}, p_{i-3}, p_{i-2})$, $(p_{i-2}, p_{i+1}, p_{i+2})$, $(p_{i-3}, p_{i-2}, p_{i+1})$, and $(p_{i+1}, p_{i+2}, p_{i+3})$, respectively.

where Φ is as in (20), $\gamma'^{\min} = \gamma^{\min}(1 - P^{\max})^{N-1}$, $\gamma'^{\max} = \gamma^{\max}$, $\Lambda = 1/P^{\min} + 1/(1 - P^{\max})$, and we have:

$$Z^{\min} = \left(\frac{\gamma'^{\min}}{\gamma'^{\max}}\right)^{1-\alpha} (L^{\min} - 1)^\alpha, \quad (22)$$

$$Z^{\max} = \left(\frac{\gamma'^{\max}}{\gamma'^{\min}}\right)^{1-\alpha} \left((L^{\min} - 1) + \sqrt[1-\alpha]{(L - L^{\min})P^{\max}} \right)^\alpha, \quad (23)$$

Theorem 7 guarantees that Algorithm 2 has a unique fixed point which is the global optimal solution of problem (NUM). Notice that Z^{\min} and Z^{\max} are lower and upper bounds on $\gamma'^{\alpha-1}(w_n(\mathbf{p}_{-n}) + \sqrt[1-\alpha]{v_n(\mathbf{p}_{-n})})^\alpha$ for any $n \in \mathcal{N}$ and any $i \in \mathcal{L}_n$. The proof of Theorem 7 is similar to that of Theorem 5. Theorem 7 is general and applies to *any* topology. Given the particular topology of interest, we can further refine (21), e.g., as in (19). Notice that condition (21) is only sufficient as in the fully interfered case. For example, we can numerically verify that for many practical topologies (e.g., chain topologies), problem (NUM) with $\alpha \in (0.5, 1)$ has a unique local (thus global) optimal solution for *any* choice of system parameters.

Unlike mapping \mathbf{f} , mapping \mathbf{f}' may not always be monotone. Therefore, the convergence results in Theorem 3 do not apply in general. However, we can still obtain sufficient conditions under which the convergence to the unique fixed point is guaranteed. Next, we define another mapping which helps us prove the asynchronous convergence of Algorithm 2.

Definition 3: A mapping function $\mathbf{g}(\mathbf{p})$ is a *weighted maximum norm contraction mapping* [17, pp. 434] if $\mathbf{g}(\mathbf{p})$ is a contraction mapping with respect to a weighted maximum norm $\|\mathbf{g}(\mathbf{p})\|_\infty = \max_{i \in \mathcal{L}} |g_i(\mathbf{p})|/\omega_i$, where $\boldsymbol{\omega} = (\omega_1, \dots, \omega_L)$ is a vector with positive coordinates.

Similar to monotone mappings, the weighted maximum norm contraction mappings satisfy both *synchronous convergence* and *box* conditions [17, pp. 434-437]. From *asynchronous convergence theorem* [17, pp. 431], this implies that a weighted norm contraction mapping globally and asynchronously converges to its unique fixed point, based on which we can show the following key results:

Theorem 8: For any general topology, Algorithm 2 globally and asynchronously converges to the unique global optimal solution of problem (NUM) if

$$\frac{|1-\alpha|}{\alpha} \sqrt{N} X^{\max} \Lambda \Phi(Z^{\min}, Z^{\max}) < 1. \quad (24)$$

If $\alpha \in (0, 1)$, then Z^{\min} and Z^{\max} are as in (22) and (23). If $\alpha \geq 1$, $Z^{\min} = (\frac{\gamma'^{\min}}{\gamma'^{\max}})^{\alpha-1} (L^{\min} - 1)^\alpha$ and $Z^{\max} = (\frac{\gamma'^{\max}}{\gamma'^{\min}})^{\alpha-1} ((L^{\min} - 1) + \sqrt[1-\alpha]{(L - L^{\min})P^{\max}})^\alpha$.

The proof of Theorem 8 is available in Appendix F. The idea is to use the relationship between l_2 and l_∞ norms to obtain a sufficient condition under which \mathbf{f}' is a weighted

maximum norm contraction mapping with *unit* weights. Thus, Algorithm 2 asynchronously converges to its unique fixed point. From Theorem 6, the convergence is indeed towards the unique global optimal solution of problem (NUM). Notice that condition (24) is a sufficient (but not necessary) condition for asynchronous convergence. Simulation results in Section V verify that Algorithm 2 converges under a wide range of system parameters. We also notice that if (24) holds for some $\alpha \in (0, 1)$, then (21) also holds for the same α .

The exact value of asynchronism measure H is *not* important for any of the proofs. Following the same argument in Section IV-A, Algorithm 2 works properly under delayed or lost messages. In Section V, we assess the optimality, robustness, and convergence of Algorithm 2 for several randomly selected topologies and under different channel conditions.

V. SIMULATION RESULTS

In this section, we assess the optimality, convergence and robust performance of our algorithms. In particular, we show the advantages of Algorithms 1 and 2 compared with the previously proposed subgradient-based algorithm in [3] (also see [19]) as well as the IEEE 802.11 DCF. The simulation environment is MATLAB. For DCF, we only implemented the basic features, mostly the well-known *binary exponentially backoff* (BEB) mechanism to update persistent probabilities, without any carrier sensing mechanism. The parameters of BEB are tuned based on the results in [2, Theorem 4].

A. Convergence and Optimality

We first consider a fully interfered topology with $N = 3$ nodes and $L = 6$ links. In this network, each node has two outgoing links, one to each of the other two nodes. For all nodes $n \in \mathcal{N}$, we set the minimum persistent probability $P_n^{\min} = 0.01$ and the maximum total probability $P_n^{\max} = 0.99$. We also set asynchronism measure $H = 10$. The peak rates of 6 links are $\gamma_1 = 6$ Mbps, $\gamma_2 = 36$ Mbps, $\gamma_3 = 9$ Mbps, $\gamma_4 = 12$ Mbps, $\gamma_5 = 18$ Mbps, and $\gamma_6 = 54$ Mbps. Communication delays among nodes are up to 10 slots and the packet error rate is 0.1 (i.e., on average, 10% of the messages are lost). Fig. 4(a) shows the trajectories of adjusted persistent probabilities and their optimal values when $\alpha = 2$ (which is *greater* than 1). In this case, $\mathbf{p}^* = [0.26, 0.11, 0.21, 0.18, 0.16, 0.09]^T$. We can see that Algorithm 1 converges to the optimal solution \mathbf{p}^* within less than 300 slots, even with communication delay and message loss. Similar results when $\alpha = 0.6$ (which is *less* than 1) are shown in Fig. 4(b). In this case, $\mathbf{p}^* = [0.06, 0.21, 0.07, 0.09, 0.18, 0.38]^T$. Again, we see that Algorithm 1 converges to \mathbf{p}^* within 320 slots. Similar results can be obtained when the simulated network has a *chain* topology and Algorithm 2 is being used.

Next, we compare the convergence speed of Algorithm 2 and that of the subgradient-based algorithm [3] in a network with a randomly generated general topology, $N = 10$ nodes, and $L = 24$ wireless links. Results are shown in Fig. 5. Here we assume that there is no communication delay and no packet corruption in the underlying communication channel. We can see that Algorithm 2 converges much faster than the subgradient-based algorithm. In fact, Algorithm 2 converges

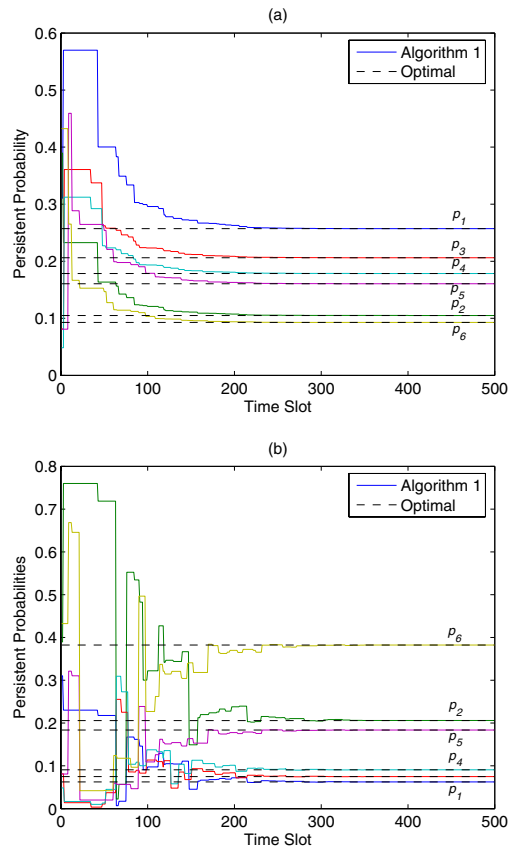


Fig. 4. Simulation results for a fully interfered topology with three nodes and six links. Algorithm 1 is used. Communications delay is 10 time slots, packet error rate is 0.1, and $\gamma_1 = 6, \gamma_2 = 36, \gamma_3 = 9, \gamma_4 = 12, \gamma_5 = 18, \gamma_6 = 54$ (all in Mbps). (a) Utility parameter $\alpha = 2$, (b) Utility parameter $\alpha = 0.6$.

to the optimal persistent probabilities after only 11 iterations. However, it takes 182 iterations for the subgradient-based algorithm to reach the exact optimal persistent probabilities.

B. Signalling Overhead

In this section, we compare the signalling overhead in our algorithms with the subgradient-based algorithm [3]. In the simulation model, nodes are *randomly* located in $100m \times 100m$ and $1km \times 1km$ fields for fully interfered and general topologies, respectively. The communication and interference ranges are $150m$ and $300m$, respectively. Each node has an outgoing link to any of its neighboring nodes within its communication range. The peak data rates (i.e., γ_i for all $i \in \mathcal{L}$) are selected *randomly* between 6 to 54 Mbps. Parameter α is set to 2 which models harmonic mean fair allocation. We assume that each *message value*, both in our proposed algorithms and the subgradient-based algorithm, requires two bytes. We notice that since Algorithm 2 and the subgradient-based algorithm [3] require transmission of *two* message values, the size of their messages is twice larger than the size of the messages of Algorithm 1. Recall that Algorithm 1 only requires transmission of a single message value. Given the message sizes, the signalling overhead for each algorithm is defined as the *total* required message exchange (in Kbytes) that the algorithm needs before it reaches the corresponding optimal solution of problem (NUM). Simulation results for

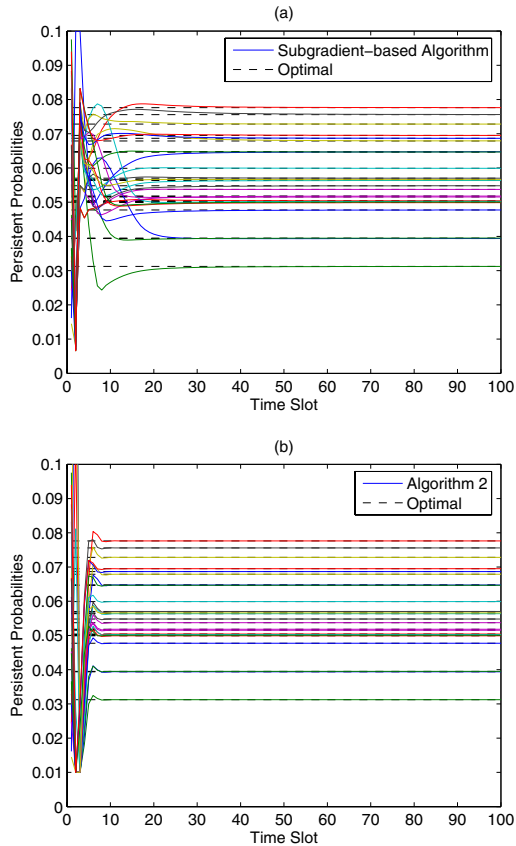


Fig. 5. Comparison between Algorithm 2 and the subgradient-based algorithm [3] in term of convergence speed in a network with $N = 10$ nodes, $L = 24$ links, and a randomly generated general topology. (a) Subgradient-based algorithm is being used. (b) Algorithm 2 is being used.

general topologies when number of nodes N varies from 5 to 30 are shown in Fig. 6. We see that increasing the number of nodes increases the signalling overhead. However, Algorithm 1 and 2 manage to reach the optimal solutions via much less signalling. Compared to the subgradient-based algorithm and when $N = 30$, Algorithms 1 and 2 reduce the signalling overhead by 1120% (from 55.2 KByte to 4.5 KByte) and 810% (from 111.3 KByte to 10.8 KByte), respectively. Notice that one reason for the superiority of our proposed algorithms is their faster convergence speed. In addition, Algorithm 1 reduces the message amount by half, which also contributes to reducing the signalling overhead.

C. Robustness

Since the underlying communication channels are not ideal in practice, transmitted messages by MAC protocols may be delayed or even be lost. In this section, we show that our proposed algorithms are robust with respect to both message delay and loss. The simulation model is the same as that in Section V-B. We only consider general topologies with $N = 30$ nodes. Results for fully interfered topologies are similar. The simulation time is set to 50,000 time slots. For each algorithm, we measure the optimality of the achieved network utility at the end of each simulation run.

First, we assume that the communication delay varies from 10 to 50 time slots. Results are shown in Fig. 7(a). We see that

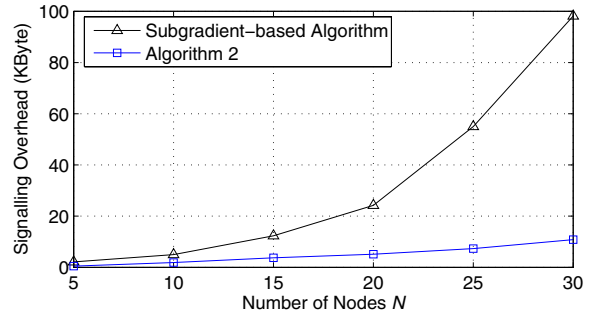


Fig. 6. Comparison between Algorithm 2 and the subgradient-based algorithm [3] in term of signalling overhead when the number of wireless nodes varies from 5 to 30. Each point represents the average results from simulating 10 random general topologies. Simulation results for the special case of fully interfered topology are similar and omitted.

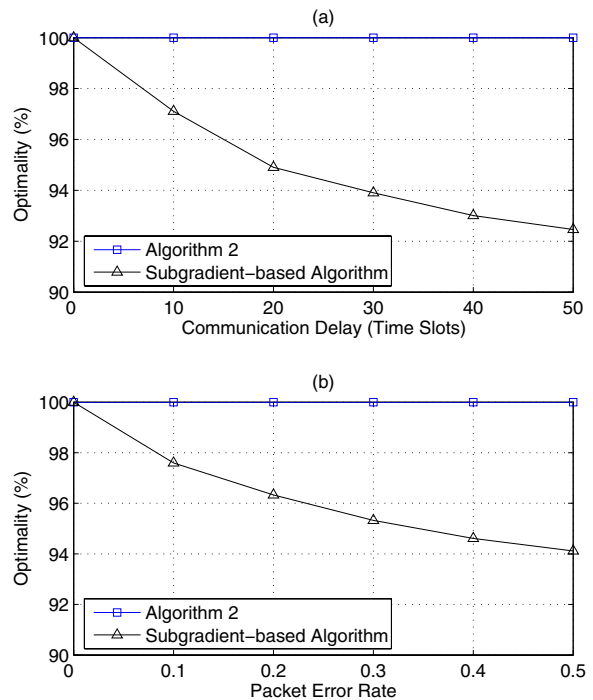


Fig. 7. Comparison between Algorithm 2 and the subgradient-based algorithm [3] in term of robustness with respect to message delay and loss. Each point represents the average results from simulating 10 random general topologies, each including 30 nodes. (a) Optimality in percentage when maximum communication delay varies from 10 to 50 time slots, (b) Optimality in percentage when packet error rate varies from 0.1 to 0.5.

by increasing delay up to 50 time slots, the subgradient-based algorithm leads to 8.4% optimality loss while Algorithm 2 can always find the exact optimal solutions. Note that although Algorithm 2 is robust to delay, higher delays can cause more iterations for the algorithm to converge. Algorithm 2 converges to the optimal persistent probabilities on average after 421, 1581, 3641, 6472, and 9923 time slots when communication delay is 10, 20, 30, 40, and 50 time slots, respectively.

Next, we assess the effect of message delay when the packet error rate varies from 0.1 (i.e., 10% of the messages are lost) to 0.5. Results are shown in Fig. 7(b). We see that Algorithm 2 is

VI. CONCLUSIONS AND FUTURE WORK

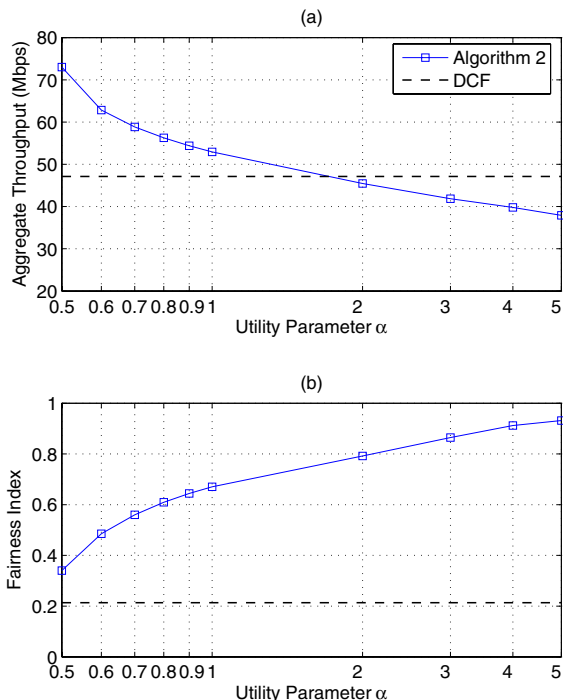


Fig. 8. Comparison between DCF and Algorithm 2 when utility parameter α varies from 0.5 to 5. Each point represents the average results from simulating 10 random general topologies, each including 30 nodes. Parameter α acts as a knob in Algorithm 2 to control the tradeoff between efficiency and fairness. As α increases, the resource allocation becomes more fair, but less efficient and vice versa. (a) Aggregate network throughput, (b) Jain's fairness index.

robust to message loss. On average, it converges to the optimal probabilities after 312, 473, 531, 629, and 727 slots when packet error rate is 0.1, 0.2, 0.3, 0.4, and 0.5, respectively.

D. Comparison with IEEE 802.11 DCF

Since our algorithms achieve the global optimal solution of problem (NUM), they establish a performance upper-bound for all MAC algorithms that are designed to solve the same problem. On the other hand, they can resolve some of the existing problems in the current 802.11 DCF, e.g., its well-known *short-term fairness* problem due to binary exponential backoff. Next, we compare Algorithm 2 with DCF in terms of system throughput and Jain's fairness index [20] when $N = 30$. Short-term fairness is obtained using a sliding window of size 200 slots. Results when utility parameter α varies from 0.5 to 5 are shown in Fig. 8. Each point represents the average results from simulating 10 random general topologies. We see that, parameter α acts as a knob in Algorithm 2 to control the tradeoff between efficiency and fairness. By increasing α we can make the system more fair but less efficient (and vice versa). If $\alpha = 0.5$, then the throughput is 54.9% higher than DCF (see Fig. 8(a)). Besides, for any $\alpha \in [0.5, 5]$, the fairness is much better than DCF (see Fig. 8(b)). We can see that, regardless of the network size, by proper selection of α (between 0.5 and 2), Algorithm 2 can improve both system throughput and fairness. Similar results are obtained when we increase the number of nodes.

In this paper, we designed two distributed contention-based MAC algorithms to solve the NUM problem at the link-layer in wireless ad hoc networks. Both algorithms globally converge to the unique global optimal solution of the NUM problem under mild conditions on system parameters. In particular, it is true for α -fair utility functions with $\alpha \in (0, 1)$, in which case the NUM problem cannot be transformed into a convex optimization problem using logarithmic change of variables as in the previous literature, and thus, it is already very challenging to even prove the uniqueness of the optimal solution before designing any algorithm.

Besides supporting a wider range of utility functions, our proposed algorithms have several other advantages over previously proposed NUM-based random MAC algorithms, including much less message passing, fully asynchronous updates, robustness to message delays and message losses, and very fast convergence speed. Simulation results confirm that our algorithms converge faster than the recently proposed algorithm in [3] with significantly (e.g., a magnitude order of 10 times) less signaling overhead. They are robust to arbitrary delays in control message exchanges among wireless nodes, and even control message losses due to channel errors in . Moreover, our proposed algorithms achieve better efficiency-fairness trade-off compared to the IEEE 802.11 DCF.

Results in this paper can be extended in several directions. First, we can extend optimization problem (NUM) and formulate the network utility maximization problem at the transport-layer to design an optimal, fast, and robust joint congestion control and random MAC scheme. After decomposition of the joint problem (cf. [21] and [19]), the random MAC sub-problem can be solved in closed-form similar to (9) and (16) for fully interfered and general network topologies, respectively. We shall also extend our model to more realistic networking scenarios, where the queues at each node have finite-backlog [22], [23]. In this regard, the data rate model in (1) needs to be modified. Most of the techniques used in this paper, such as myopic adjustment of the persistent probabilities using best-response, are expected to be still applicable for networks with finite-backlogged queues to reduce the signaling overhead and achieve more robust performance. We may also relax the need for the existence of equal-length time slots, following the techniques used in [24, Section IV], where a utility-optimal random access algorithm is proposed for logarithmic utilities (i.e., $\alpha = 1$) in pure (un-slotted) ALOHA systems. It is also important to analyze how our algorithms perform in a highly dynamic network, where the nodes locations and wireless channels are quickly changing. Finally, we may extend our random access model by including carrier sensing. Carrier sensing can help to avoid collision among those transmissions which their transmitter nodes are within the carrier sensing range of each other. This can be modeled by redefining the interference set \mathcal{N}_i for each link $i \in \mathcal{L}$ and modify the data rate model in (1), accordingly.

APPENDIX

A. Proof of Theorem 1

Lemma 1: For each node $n \in \mathcal{N}$, we have:

$$\frac{1}{P_n^{\min}} - B_n \leq \sqrt[\alpha]{\gamma_i^{\alpha-1} (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})})}, \quad \forall i \in \mathcal{B}_n, \quad (25)$$

$$\frac{P_n^{\max}}{P_n^{\min}} - C_n \leq \sqrt[\alpha]{\gamma_i^{\alpha-1} w_n}, \quad \forall i \in \mathcal{C}_n. \quad (26)$$

$$\frac{1}{P_n^{\min}} - B_n > \sqrt[\alpha]{\gamma_i^{\alpha-1} (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})})}, \quad \forall i \in \mathcal{L}_n \setminus \mathcal{B}_n, \quad (27)$$

$$\frac{P_n^{\max}}{P_n^{\min}} - C_n > \sqrt[\alpha]{\gamma_i^{\alpha-1} w_n}, \quad \forall i \in \mathcal{L}_n \setminus \mathcal{C}_n. \quad (28)$$

Proof: From (7) and knowing that $\mathcal{L}_n \setminus \mathcal{A}_n = \{i_1, \dots, i_\kappa\}$, for each link $i \in \mathcal{B}_n$ we have:

$$\begin{aligned} \frac{1}{P_n^{\min}} - B_n &\leq \sum_{j \in \mathcal{L}_n \setminus \mathcal{A}_n} \sqrt[\alpha]{\left(\frac{\gamma_{i_{\kappa+1}}}{\gamma_j}\right)^{\alpha-1}} + \sqrt[\alpha]{\gamma_{i_{\kappa+1}}^{\alpha-1} v_n(\mathbf{p}_{-n})} \\ &= \sqrt[\alpha]{\frac{\gamma_{i_{\kappa+1}}^{\alpha-1}}{\gamma_i^{\alpha-1}}} \left(\sum_{j \in \mathcal{L}_n \setminus \mathcal{A}_n} \sqrt[\alpha]{\left(\frac{\gamma_j}{\gamma_i}\right)^{\alpha-1}} + \sqrt[\alpha]{\gamma_i^{\alpha-1} v_n(\mathbf{p}_{-n})} \right) \\ &\leq \sqrt[\alpha]{\gamma_i^{\alpha-1}} (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}), \end{aligned}$$

where the last inequality comes from the fact that $\sqrt[\alpha]{\gamma_{i_{\kappa+1}}^{\alpha-1}} \leq \sqrt[\alpha]{\gamma_{i_{\sigma+1}}^{\alpha-1}} \leq \sqrt[\alpha]{\gamma_i^{\alpha-1}}$ for all $i \in \mathcal{B}_n$. Recall that $\kappa = \min\{\sigma, \varsigma\}$ and $i_{\kappa+1} \leq i_{\sigma+1}$. Proof of (26) is similar. We prove (27) by *contradiction*. Assume that there exists $i_\varepsilon \in \mathcal{L}_n \setminus \mathcal{A}_n$ such that (27) does not hold. It is clear that $\varepsilon \leq \sigma$. We have

$$\frac{1}{P_n^{\min}} - L_n + \sigma \leq \sqrt[\alpha]{\gamma_{i_\varepsilon}} \left(\sum_{l=1}^{\sigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_l}\right)^{\alpha-1}} + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})} \right) \quad (29)$$

$$\frac{1}{P_n^{\min}} - L_n + \varepsilon - 1 \leq \sum_{l=1}^{\varepsilon-1} \sqrt[\alpha]{\left(\frac{\gamma_{i_\varepsilon}}{\gamma_l}\right)^{\alpha-1}} + \sqrt[\alpha]{\gamma_{i_\varepsilon} v_n(\mathbf{p}_{-n})}, \quad (30)$$

where the last inequality results from the fact that $\frac{\gamma_{i_\varepsilon}}{\gamma_l} \leq 1$ for all $l = \varepsilon, \dots, \delta$ which implies $\sum_{l=\varepsilon}^{\sigma} \sqrt[\alpha]{(\gamma_{i_\varepsilon}/\gamma_l)^{\alpha-1}} \leq \sigma - \varepsilon + 1$. Comparing with (7), inequality (30) implies that $i_\varepsilon \in \mathcal{A}_n$ which contradicts the assumption that $i_\varepsilon \in \mathcal{L}_n \setminus \mathcal{A}_n$. Proof of (28) is similar and is omitted for brevity. ■

Since (LOCAL-NUM) is strictly concave in \mathbf{p}_n , its optimal solution is *unique* [16, pp. 137] and satisfies the following *necessary* and *sufficient* KKT conditions [16, pp. 244]:

$$p_i \geq P_n^{\min}, \quad \forall i \in \mathcal{L}_n, \quad (31)$$

$$\sum_{i \in \mathcal{L}_n} p_i \leq P_n^{\max}, \quad (32)$$

$$\frac{1}{\gamma_i^{\alpha-1} p_i^\alpha} - \frac{v_n(\mathbf{p}_{-n})}{(1 - \sum_{i \in \mathcal{L}_n} p_i)^\alpha} = \lambda_n - \delta_i, \quad \forall i \in \mathcal{L}_n, \quad (33)$$

$$\lambda_n (\sum_{i \in \mathcal{L}_n} p_i - P_n^{\max}) = 0, \quad (34)$$

$$\delta_i (P_n^{\min} - p_i) = 0, \quad \forall i \in \mathcal{L}_n, \quad (35)$$

$$\lambda_n \geq 0, \quad \delta_i \geq 0, \quad \forall i \in \mathcal{L}_n, \quad (36)$$

where λ_n denotes the Lagrangian multiplier corresponding to constraint $\sum_{i \in \mathcal{L}_n} p_i \leq P_n^{\max}$ and δ_i denotes the Lagrangian multiplier corresponding to constraint $p_i \geq P_n^{\min}$ for each $i \in \mathcal{L}_n$. Now we need to show that (9) leads to (31) to (36).

Condition (31) directly results from (9). We also have:

$$\begin{aligned} &\sum_{i \in \mathcal{L}_n} f_i(\mathbf{p}_{-n}) \\ &\leq A_n P_n^{\min} + \sum_{i \in \mathcal{L}_n \setminus \mathcal{A}_n} \frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1} w_n}} \\ &= A_n P_n^{\min} + \frac{P_n^{\max} - A_n P_n^{\min}}{w_n} \left(\sum_{i \in \mathcal{L}_n \setminus \mathcal{A}_n} \sqrt[\alpha]{\left(\frac{1}{\gamma_i}\right)^{\alpha-1}} \right) \\ &\stackrel{\text{by (10)}}{=} P_n^{\max}. \end{aligned}$$

Thus, condition (32) also holds. Two cases are possible:

Case I: $\sum_{i \in \mathcal{L}_n} f_i(\mathbf{p}_{-n}) < P_n^{\max}$, which only happens if $\sum_{i \in \mathcal{L}_n \setminus \mathcal{A}_n} f_i(\mathbf{p}_{-n}) < P_n^{\max} - A_n P_n^{\min}$. Thus, there exists $l \in \mathcal{L}_n \setminus \mathcal{A}_n$ such that we have:

$$\begin{aligned} \frac{1}{\sqrt[\alpha]{\gamma_l^{\alpha-1} w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}}} &< \frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_l^{\alpha-1} w_n}}, \\ \frac{1 - A_n P_n^{\min}}{w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}} &< \frac{P_n^{\max} - A_n P_n^{\min}}{w_n}. \quad (37) \end{aligned}$$

Multiplying both sides of this inequality by $\sqrt[\alpha]{(1/\gamma_i)^{\alpha-1}}$, for any link $i \in \mathcal{L}_n \setminus \mathcal{A}_n$, we have:

$$\frac{1}{\sqrt[\alpha]{\gamma_i^{\alpha-1} w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}}} < \frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1} w_n}}.$$

Hence, for each $i \in \mathcal{L}_n \setminus \mathcal{A}_n$, we have:

$$f_i(\mathbf{p}_{-n}) = \max \left(P_n^{\min}, \frac{1}{\sqrt[\alpha]{\gamma_i^{\alpha-1} w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}}} \frac{1 - A_n P_n^{\min}}{w_n} \right). \quad (38)$$

In this case, we have $\sigma \leq \varsigma$ and $\kappa = \min\{\sigma, \varsigma\} = \sigma$. In other words, $\mathcal{A}_n = \mathcal{B}_n$ and $A_n = B_n$. We prove this by *contradiction*. Assume that $\sigma > \varsigma$ and $\kappa = \varsigma$. We have:

$$\begin{aligned} \frac{1 - (L_n - \varsigma) P_n^{\min}}{\sum_{l=1}^{\varsigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_{i_l}}\right)^{\alpha-1}} + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}} &\stackrel{\text{by (37)}}{<} \frac{P_n^{\max} - (L_n - \varsigma) P_n^{\min}}{\sum_{l=1}^{\varsigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_{i_l}}\right)^{\alpha-1}}} \\ &\stackrel{\text{by (8)}}{\leq} \sqrt[\alpha]{\gamma_{i_{\varsigma+1}}^{\alpha-1}} P_n^{\min}. \end{aligned}$$

Thus,

$$\frac{1}{P_n^{\min}} - L_n + \varsigma < \sqrt[\alpha]{\gamma_{i_{\varsigma+1}}^{\alpha-1}} \left(\sum_{l=1}^{\sigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_{i_l}}\right)^{\alpha-1}} + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})} \right). \quad (39)$$

Comparing (39) and (7), we have $i_{\varsigma+1} \in \mathcal{B}_n$. This implies that $\sigma \leq \varsigma$ which contradicts our counter assumption $\sigma > \varsigma$. Thus, $\kappa = \sigma$ and we have:

$$\mathcal{A}_n = \mathcal{B}_n, \quad A_n = B_n. \quad (40)$$

We can verify that conditions (33)-(35) hold for $\mathbf{p}_n = \mathbf{f}_n(\mathbf{p}_{-n})$ if $\lambda_n = 0$ and we set $\delta_i = 0$ if $i \in \mathcal{L}_n \setminus \mathcal{A}_n$ and set:

$$\delta_i = \frac{v_n(\mathbf{p}_{-n})}{\left(1 - A_n P_n^{\min} - \sum_{j \in \mathcal{L}_n \setminus \mathcal{A}_n} f_j(\mathbf{p}_{-n})\right)^\alpha} - \left(\frac{1}{\sqrt[\alpha]{\gamma_i^{\alpha-1} P_n^{\min}}} \right)^\alpha \quad (41)$$

if $i \in \mathcal{A}_n$. Next, we need to show that (36) holds. For each node $i \in \mathcal{A}_n$, we have:

$$\begin{aligned} & \frac{v_n(\mathbf{p}_{-n})}{\left(1 - A_n P_n^{\min} - \sum_{j \in \mathcal{L}_n \setminus \mathcal{A}_n} f_j(\mathbf{p}_{-n})\right)^\alpha} \\ & \stackrel{\text{by (38)}}{\geq} \frac{v_n(\mathbf{p}_{-n})}{\left(1 - A_n P_n^{\min}\right) \left(1 - \sum_{j \in \mathcal{L}_n \setminus \mathcal{A}_n} \frac{\sqrt[\alpha]{1/\gamma_j^{\alpha-1}}}{w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}}\right)^\alpha} \\ & \stackrel{\text{by (10)}}{=} \left(\frac{\sqrt[\alpha]{\gamma_i^{\alpha-1}} (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})})}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} (1 - A_n P_n^{\min})}\right)^\alpha \\ & \stackrel{\text{by (25)}}{\geq} \left(\frac{1/P_n^{\min} - A_n}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} (1 - A_n P_n^{\min})}\right)^\alpha = \left(\frac{1}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} P_n^{\min}}\right)^\alpha, \\ & \stackrel{\text{and (40)}}{\geq} \end{aligned}$$

Therefore, $\delta_i \geq 0$ for all $i \in \mathcal{L}_n$ and condition (36) also holds. Finally, for each link $i \in \mathcal{L}_n \setminus \mathcal{A}_n$ we have:

$$\frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n} \stackrel{\text{by (37)}}{>} \frac{1 - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})})} \stackrel{\text{and (27)}}{\geq} P_n^{\min}, \quad \stackrel{\text{by (40)}}{\geq} \quad (42)$$

which guarantees that the upper bound in (9) is indeed greater than the lower bound. This completes the proof for Case I.

Case II: $\sum_{i \in \mathcal{L}_n} f_i(\mathbf{p}_{-n}) = P_n^{\max}$, which happens only if

$$\sum_{i \in \mathcal{L}_n \setminus \mathcal{A}_n} f_i(\mathbf{p}_{-n}) = P_n^{\max} - A_n P_n^{\min}. \quad (43)$$

We already know that

$$\sum_{i \in \mathcal{L}_n \setminus \mathcal{A}_n} \frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n} \stackrel{\text{by (10)}}{=} P_n^{\max} - A_n P_n^{\min}. \quad (44)$$

From (43) and (44) and since $f_i(\mathbf{p}_{-n}) \leq (P_n^{\max} - A_n P_n^{\min}) / (\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n)$ for all $i \in \mathcal{L}_n \setminus \mathcal{A}_n$, we have:

$$f_i(\mathbf{p}_{-n}) = \frac{P_n^{\max} - A_n P_n^{\min}}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n}, \quad \forall i \in \mathcal{L}_n \setminus \mathcal{A}_n. \quad (45)$$

From (9) and (45) we have:

$$\frac{1 - A_n P_n^{\min}}{w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}} \geq \frac{P_n^{\max} - A_n P_n^{\min}}{w_n}. \quad (46)$$

In this case, we have $\sigma \geq \varsigma$ and $\kappa = \min\{\sigma, \varsigma\} = \varsigma$. In other words, $\mathcal{A}_n = \mathcal{C}_n$, and $A_n = C_n$. We prove this by *contradiction*. Assume that $\sigma < \varsigma$ and $\kappa = \sigma$. We have:

$$\begin{aligned} \frac{P_n^{\max} - (L_n - \sigma) P_n^{\min}}{\sum_{l=1}^{\sigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_{i_l}}\right)^{\alpha-1}}} & \stackrel{\text{by (46)}}{\leq} \frac{1 - (L_n - \sigma) P_n^{\min}}{\sum_{l=1}^{\sigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_{i_l}}\right)^{\alpha-1}} + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})}} \\ & \stackrel{\text{by (7)}}{\leq} \sqrt[\alpha]{\gamma_{i_{\sigma+1}}^{\alpha-1}} P_n^{\min}. \end{aligned}$$

Thus, we have:

$$\frac{P_n^{\max}}{P_n^{\min}} - L_n + \sigma \leq \sqrt[\alpha]{\gamma_{i_{\sigma+1}}^{\alpha-1}} \left(\sum_{l=1}^{\sigma} \sqrt[\alpha]{\left(\frac{1}{\gamma_{i_l}}\right)^{\alpha-1}} \right). \quad (47)$$

Comparing (47) and (8), we have $i_{\sigma+1} \in \mathcal{C}_n$. This implies that $\sigma \geq \varsigma$ which contradicts our counter assumption $\sigma < \varsigma$. Thus, $\kappa = \varsigma$ and we have:

$$\mathcal{A}_n = \mathcal{C}_n, \quad A_n = C_n. \quad (48)$$

We can verify that conditions (33)-(35) hold for each $\mathbf{p}_n = \mathbf{f}_n(\mathbf{p}_{-n})$ if we set

$$\lambda_n = \left(\frac{w_n}{P_n^{\max} - A_n P_n^{\min}} \right)^\alpha - \left(\frac{\sqrt[\alpha]{v_n(\mathbf{p}_{-n})}}{1 - P_n^{\max}} \right)^\alpha, \quad (49)$$

and

$$\delta_i = \begin{cases} \left(\frac{w_n}{P_n^{\max} - A_n P_n^{\min}} \right)^\alpha - \left(\frac{1}{\sqrt[\alpha]{\gamma_i^{\alpha-1}} P_n^{\min}} \right)^\alpha, & \text{if } i \in \mathcal{A}_n, \\ 0, & \text{if } i \in \mathcal{L}_n \setminus \mathcal{A}_n. \end{cases} \quad (50)$$

Next, we show that (36) holds. From (46) and by reordering,

$$\begin{aligned} w_n(1 - A_n P_n^{\min}) & \geq (w_n + \sqrt[\alpha]{v_n(\mathbf{p}_{-n})})(P_n^{\max} - A_n P_n^{\min}) \\ \frac{w_n}{P_n^{\max} - A_n P_n^{\min}} & \geq \frac{\sqrt[\alpha]{v_n(\mathbf{p}_{-n})}}{1 - P_n^{\max}}. \end{aligned} \quad (51)$$

On the other hand, from (26) and (48) and by reordering,

$$w_n / (P_n^{\max} - A_n P_n^{\min}) \geq 1 / (\sqrt[\alpha]{\gamma_i^{\alpha-1}} P_n^{\min}). \quad (52)$$

Replacing (51) in (49), we have $\lambda_n \geq 0$. On the other hand, replacing (52) in (50), we have $\delta_i \geq 0$ for all $i \in \mathcal{L}_n$. Thus, condition (36) holds. Finally, for each $i \in \mathcal{L}_n \setminus \mathcal{A}_n$ we have:

$$(P_n^{\max} - A_n P_n^{\min}) / (\sqrt[\alpha]{\gamma_i^{\alpha-1}} w_n) \stackrel{\text{and (28)}}{\geq} P_n^{\min}, \quad \stackrel{\text{by (48)}}{\geq} \quad (53)$$

which guarantees that the upper bound in (9) is greater than or equal to the lower bound. ■

B. Proof of Proposition 1

If $\alpha \geq 1$, then for any $\tilde{\mathbf{p}}, \hat{\mathbf{p}} \in \mathcal{P}$ such that $\tilde{\mathbf{p}} \preceq \hat{\mathbf{p}}$, we have: $\tilde{p}_j \leq \hat{p}_j$, for all $j \in \mathcal{L}$. Thus,

$$1 - \sum_{j \in \mathcal{L}_s} \tilde{p}_j \geq 1 - \sum_{j \in \mathcal{L}_s} \hat{p}_j, \quad \forall s \in \mathcal{N}, \quad (54)$$

$$1/\tilde{p}_j \geq 1/\hat{p}_j, \quad \forall j \in \mathcal{L}. \quad (55)$$

From (54) and (55),

$$m_s(\tilde{\mathbf{p}}_{-n}) \geq m_s(\hat{\mathbf{p}}_{-n}), \quad \forall s \in \mathcal{N}, \quad (56)$$

$$v_i(\tilde{\mathbf{p}}_{-n}) \geq v_i(\hat{\mathbf{p}}_{-n}), \quad \forall i \in \mathcal{L}, \quad (57)$$

$$w_i + \sqrt[\alpha]{v_i(\tilde{\mathbf{p}}_{-n})} \geq w_i + \sqrt[\alpha]{v_i(\hat{\mathbf{p}}_{-n})} \quad \forall i \in \mathcal{L}, \quad (58)$$

$$f_i(\tilde{\mathbf{p}}_{-n}) \leq f_i(\hat{\mathbf{p}}_{-n}), \quad \forall i \in \mathcal{L}. \quad (59)$$

Thus, the vector function $\mathbf{f}(\mathbf{p})$ is a monotone increasing mapping. If $\alpha \leq 1$, then the sign of the inequalities in (56)-(59) is reversed and $\mathbf{f}(\mathbf{p})$ becomes monotone decreasing. ■

C. Proof of Theorem 3

Since H is bounded, the local memory of each node $n \in \mathcal{N}$ is updated infinitely often as $t \rightarrow \infty$. Thus, the *total asynchronism* assumption [17, pp. 430] holds. First suppose with $\alpha \geq 1$. Let \mathbf{p}^{\min} and \mathbf{p}^{\max} denote two $N \times 1$ vectors with all entries equal to P_n^{\min} and P_n^{\max} , respectively. Since $\mathbf{f}(\mathbf{p}^*) = \mathbf{p}^*$, $\mathbf{p}^{\min} \preceq \mathbf{p}^* \preceq \mathbf{p}^{\max}$, and $\mathbf{f}(\cdot)$ is monotone increasing, we have:

$$\mathbf{p}^{\min} \preceq \mathbf{f}(\mathbf{p}^{\min}) \preceq \mathbf{p}^* \preceq \mathbf{f}(\mathbf{p}^{\max}) \preceq \mathbf{p}^{\max}. \quad (60)$$

Note that, for each $n \in \mathcal{N}$ and $i \in \mathcal{L}_n$, $f_i(\mathbf{p}^{\min}) \geq P_n^{\min} \geq P_n^{\min}$ and $f_i(\mathbf{p}^{\max}) \leq P_n^{\max} \leq P_n^{\max}$. Let $\mathbf{f}^k(\mathbf{p})$ denote the composition of \mathbf{f} with itself k times, where $\mathbf{f}^0(\mathbf{p}) = \mathbf{p}$. Starting from (60) and applying \mathbf{f} for $k \geq 0$ times, we have:

$$\mathbf{f}^k(\mathbf{p}^{\min}) \preceq \mathbf{f}^{k+1}(\mathbf{p}^{\min}) \preceq \mathbf{p}^* \preceq \mathbf{f}^{k+1}(\mathbf{p}^{\max}) \preceq \mathbf{f}^k(\mathbf{p}^{\max}). \quad (61)$$

We also define: $\mathcal{P}^k = \{\mathbf{p} | \mathbf{f}^k(\mathbf{p}^{\min}) \preceq \mathbf{p} \preceq \mathbf{f}^k(\mathbf{p}^{\max})\}$. If $\mathbf{p} \in \mathcal{P}^k$, then $\mathbf{f}(\mathbf{p}) \in \mathcal{P}^{k+1}$. From (61), $\mathcal{P}^{k+1} \subseteq \mathcal{P}^k$ for all $k \geq 0$. We can also show (by contradiction) that if $\mathcal{P}^k \neq \{\mathbf{p}^*\}$, then $\mathcal{P}^{k+2} \neq \mathcal{P}^k$. That is, $\mathcal{P}^{k+2} \subset \mathcal{P}^k$. Thus, $\lim_{k \rightarrow \infty} \mathbf{f}^k(\mathbf{p}^{\min}) = \lim_{k \rightarrow \infty} \mathbf{f}^k(\mathbf{p}^{\max}) = \mathbf{p}^*$. Besides, if $\mathbf{p}, \hat{\mathbf{p}} \in \mathcal{P}^k$, then $(p_1, \dots, p_{i-1}, \hat{p}_i, p_{i+1}, \dots, p_L) \in \mathcal{P}^k$ for any $k \geq 0$. Therefore, both *synchronous convergence* and *box conditions* [17, pp. 431] hold. From *asynchronous convergence theorem* [17, pp. 431], starting from any point in \mathcal{P}^0 , Algorithm 1 will converge to \mathbf{p}^* . Since $\mathcal{P} \subseteq \mathcal{P}^0$, the proof for $\alpha \geq 1$ is complete. Now consider the case with $\alpha \leq 1$, from Proposition 1, the mapping $\mathbf{f}(\cdot)$ is monotone decreasing. We have:

$$\mathbf{p}^{\min} \preceq \mathbf{f}(\mathbf{p}^{\max}) \preceq \mathbf{p}^* \preceq \mathbf{f}(\mathbf{p}^{\min}) \preceq \mathbf{p}^{\max}. \quad (62)$$

Comparing with (60), the orders of $\mathbf{f}(\mathbf{p}^{\min})$ and $\mathbf{f}(\mathbf{p}^{\max})$ is exchanged in (62). Applying $\mathbf{f}(\cdot)$ once more, we have:

$$\mathbf{f}(\mathbf{p}^{\max}) \preceq \mathbf{f}^2(\mathbf{p}^{\min}) \preceq \mathbf{p}^* \preceq \mathbf{f}^2(\mathbf{p}^{\max}) \preceq \mathbf{f}(\mathbf{p}^{\min}).$$

By mathematical induction, we can show that if k is *odd*, then $\mathbf{f}^k(\mathbf{p}^{\max}) \preceq \mathbf{f}^{k+1}(\mathbf{p}^{\min}) \preceq \mathbf{p}^* \preceq \mathbf{f}^{k+1}(\mathbf{p}^{\max}) \preceq \mathbf{f}^k(\mathbf{p}^{\min})$, and if k is *even*, then:

$$\mathbf{f}^k(\mathbf{p}^{\min}) \preceq \mathbf{f}^{k+1}(\mathbf{p}^{\max}) \preceq \mathbf{p}^* \preceq \mathbf{f}^{k+1}(\mathbf{p}^{\min}) \preceq \mathbf{f}^k(\mathbf{p}^{\max}),$$

For each (either odd or even) $k \geq 0$, we redefine set $\mathcal{P}^k = \{\mathbf{p} | \min[\mathbf{f}^k(\mathbf{p}^{\min}), \mathbf{f}^k(\mathbf{p}^{\max})] \preceq \mathbf{p} \preceq \max[\mathbf{f}^k(\mathbf{p}^{\min}), \mathbf{f}^k(\mathbf{p}^{\max})]\}$. The rest of the proof follows the case of $\alpha \geq 1$ and is omitted for brevity. ■

D. Proof of Theorem 4

For each $\hat{\mathbf{p}}^* \in \mathcal{F}$ and each link $i \in \mathcal{N}$, Theorem 1 states that $\hat{\mathbf{p}}_i^*$ is the unique global optimal solution of convex Problem (LOCAL-NUM), and satisfies the KKT conditions in (31)-(36). We denote the corresponding Lagrangian multipliers by $\hat{\lambda}_n^*$ and $\hat{\delta}_i^*$ for all $i \in \mathcal{L}_n$. We also define $\tilde{\lambda}^* = (\tilde{\lambda}_n^*, \forall n \in \mathcal{N})$ and $\tilde{\delta}^* = (\tilde{\delta}_i^*, \forall i \in \mathcal{L})$. On the other hand, since any local optimum $\hat{\mathbf{p}} \in \mathcal{S}$ is a *regular point* [18, pp. 315] of non-convex Problem (NUM), and so must satisfy the KKT *necessary conditions* [18, Proposition 3.3.1], for any $i \in \mathcal{N}$,

$$\tilde{p}_i^* \geq P_n^{\min}, \quad \forall n \in \mathcal{N}, i \in \mathcal{L}_n, \quad (63)$$

$$\sum_{i \in \mathcal{L}_n} \tilde{p}_i^* \leq P_n^{\max}, \quad \forall n \in \mathcal{N}, \quad (64)$$

$$\frac{1}{y_n(\tilde{\mathbf{p}}_{-n}^*)} \left(\frac{1}{\gamma_i^{\alpha-1} p_i^\alpha} - \frac{v_n(\mathbf{p}_{-n}^*)}{(1 - \sum_{i \in \mathcal{L}_n} p_i)^\alpha} \right) = \tilde{\lambda}_n^* - \tilde{\delta}_i^*, \quad \forall n \in \mathcal{N}, i \in \mathcal{L}_n, \quad (65)$$

$$\tilde{\lambda}_n^* (\sum_{i \in \mathcal{L}_n} p_i - P_n^{\max}) = 0, \quad \forall n \in \mathcal{N}, \quad (66)$$

$$\tilde{\delta}_i^* (P_n^{\min} - p_i) = 0, \quad \forall n \in \mathcal{N}, i \in \mathcal{L}_n, \quad (67)$$

$$\tilde{\lambda}_n^* \geq 0, \quad \tilde{\delta}_i^* \geq 0, \quad \forall i \in \mathcal{L}_n, \quad (68)$$

where $y_n(\mathbf{p}_{-n}) = (\prod_{s \in \mathcal{N} \setminus \{n\}} (1 - \sum_{j \in \mathcal{L}_s} p_j))^{1-\alpha} > 0$. We define, $\tilde{\lambda}^* = (\tilde{\lambda}_n^*, \forall n \in \mathcal{N})$ and $\tilde{\delta}^* = (\tilde{\delta}_i^*, \forall i \in \mathcal{L})$. If $(\hat{\mathbf{p}}^*, \hat{\lambda}^*, \hat{\delta}^*)$ satisfies conditions (31)-(36) for all $n \in \mathcal{N}$, then $(\hat{\mathbf{p}}^*, (\hat{\lambda}_n^*/y_n(\hat{\mathbf{p}}_{-n}^*), \forall n \in \mathcal{N}), (\hat{\delta}_i^*/y_n(\hat{\mathbf{p}}_{-n}^*), \forall n \in \mathcal{N}, i \in \mathcal{L}_n))$ satisfies conditions (63)-(68). On the other hand, if $(\tilde{\mathbf{p}}^*, \tilde{\lambda}^*, \tilde{\delta}^*)$ satisfies conditions (63)-(68), then $(\tilde{\mathbf{p}}^*, (\tilde{\lambda}_n^* y_n(\tilde{\mathbf{p}}_{-n}^*), \forall n \in \mathcal{N}), (\tilde{\delta}_i^* y_n(\tilde{\mathbf{p}}_{-n}^*), \forall n \in \mathcal{N}, i \in \mathcal{L}_n))$ satisfies conditions (31)-(36) for all $n \in \mathcal{N}$. The former implies that $\mathcal{F} \subseteq \mathcal{S}$ while the later implies $\mathcal{S} \subseteq \mathcal{F}$. Thus, $\mathcal{F} = \mathcal{S}$. ■

E. Proof of Theorem 5

For any $\mathbf{p} \in \mathcal{P}$, Jacobian $J(\mathbf{p})$ is defined as an $L \times L$ matrix whose entry in row i and column j is $\partial f_i / \partial p_j$. Consider node $n \in \mathcal{N}$ and each $i \in \mathcal{L}_n$, if $j \in \mathcal{L}_n$ then we have $J_{i,j}(\mathbf{p}) = 0$. On the other hand, if $j \notin \mathcal{L}_n$, where $j \in \mathcal{L}_s$, then we have:

$$J_{i,j}(\mathbf{p}) = \frac{\frac{\alpha-1}{\alpha} \frac{(1-A_n P_n^{\min})}{1 - \sum_{k \in \mathcal{L}_s} p_k} \left[m_s + \left(\frac{1 - \sum_{k \in \mathcal{L}_s} p_k}{P_j \gamma_j} \right)^{\alpha-1} \left(\frac{1 - \sum_{k \in \mathcal{L}_s} p_k}{P_j} \right) \right]}{\alpha \sqrt{\gamma_i^{\alpha-1} v_n(\mathbf{p}_{-n})}^{1-1/\alpha} (w_n + v_n(\mathbf{p}_{-n}))^{1/\alpha}}. \quad (69)$$

This is under the assumption that $f_i \in (P_n^{\min}, (P_n^{\max} - A_n P_n^{\min}) / \alpha \sqrt{\gamma_i^{\alpha-1} w_n})$. If f_i is chosen to be one of the boundary points, then the corresponding entry is 0. Thus, (69) provides upper-bounds (in absolute values) on the entries of the Jacobian. We can show that:

$$\begin{aligned} \|J\|_\infty &\leq \frac{1-\alpha}{\alpha} \max_{i \in \mathcal{L}_n, n \in \mathcal{N}} \left\{ \frac{(\gamma_i v_n(\mathbf{p}_{-n}))^{1/\alpha}}{(1 + (\gamma_i v_n(\mathbf{p}_{-n}))^{1/\alpha})^2} \right\} \times \\ &\quad \left(\frac{L^{\max}}{1 - P^{\max}} + \frac{1}{P^{\min}} \right) \\ &\leq \frac{1-\alpha}{\alpha} \Psi \Phi(V^{\min}, V^{\max}). \end{aligned} \quad (70)$$

Note that $1 - A_n P_n^{\min} \leq 1$ and $\alpha \sqrt{\gamma_i^{\alpha-1} w_n} \geq 1$ for all $i \in \mathcal{L}_n \setminus \mathcal{A}_n$. Function $x^{1/\alpha} / (1 + x^{1/\alpha})^2$ is always non-negative and less than 0.25. It has a unique maximum at $x = 1$. The function is monotonically increasing for $0 \leq x < 1$ and monotonically decreasing for $x > 1$. Its value approaches zero as either $x \rightarrow 0$ or $x \rightarrow \infty$. We can similarly show that:

$$\begin{aligned} \|J(\mathbf{p})\|_1 &\leq \frac{1-\alpha}{\alpha} \Psi \Phi(V^{\min}, V^{\max}) \left(\frac{\gamma_{\max}^{\max}}{\gamma_{\min}^{\min}} \Gamma \right)^{1-\alpha} \times \\ &\quad \left(\Omega - \frac{1}{L/L^{\min} - 1} \right). \end{aligned} \quad (71)$$

From (19), (70), and (71), we have [17, pp. 635]:

$$\|J(\mathbf{p})\|_2 \leq \sqrt{\|J(\mathbf{p})\|_\infty \|J(\mathbf{p})\|_1} < 1. \quad (72)$$

Let $\tilde{\mathbf{p}}, \hat{\mathbf{p}} \in \mathcal{P}$. From (72) and by Cauchy Schwarz inequality,

$$\|\mathbf{f}(\tilde{\mathbf{p}}) - \mathbf{f}(\hat{\mathbf{p}})\|_2 \leq \|J(\mathbf{p})\|_2 \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|_2 < \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|_2,$$

where \mathbf{p} is any convex combination of $\tilde{\mathbf{p}}$ and $\hat{\mathbf{p}}$. Thus, \mathbf{f} is an l_2 -norm contraction mapping. ■

F. Proof of Theorem 8

Following the same argument as in Appendix E, condition (21) implies that for each $\mathbf{p} \in \mathcal{P}$, we have: $\sqrt{N} \|J'(\mathbf{p})\|_2 < 1$ where $J'(\mathbf{p})$ denotes the Jacobian matrix of $\mathbf{f}'(\mathbf{p})$. From linear algebra, we also know that for each $N \times 1$ vector \mathbf{a} we have $\|\mathbf{a}\|_\infty \leq \|\mathbf{a}\|_2 \leq \sqrt{N} \|\mathbf{a}\|_\infty$. Let $\tilde{\mathbf{p}}, \hat{\mathbf{p}} \in \mathcal{P}$. Using the Cauchy-Schwarz inequality, we have:

$$\begin{aligned} \|\mathbf{f}'(\tilde{\mathbf{p}}) - \mathbf{f}'(\hat{\mathbf{p}})\|_\infty &\leq \|\mathbf{f}'(\tilde{\mathbf{p}}) - \mathbf{f}'(\hat{\mathbf{p}})\|_2 \\ &\leq \|J'(\mathbf{p})\|_2 \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|_2 \\ &< \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|_2 / \sqrt{N} \\ &\leq \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|_\infty. \end{aligned}$$

Thus, condition (21) guarantees that mapping function $\mathbf{f}'(\cdot)$ is indeed an l_∞ norm contraction mapping. Thus, it asynchronously converges to its *unique* fixed point. From Theorem

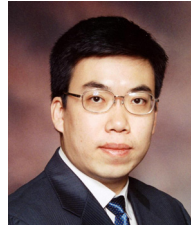
6, this further implies that Algorithm 2 converges to the unique global optimal solution of problem (NUM). ■

REFERENCES

- [1] "IEEE 802.11a," [Online] Available: <http://standards.ieee.org/getieee802/download/802.11-1999.pdf>.
- [2] J. Lee, A. Tang, J. Huang, M. Chiang, and A. Calderbank, "Reverse engineering MAC: A game-theoretic model," *IEEE J. Select. Areas in Commun.*, vol. 6, pp. 2741-2751, July 2007.
- [3] J. Lee, M. Chiang, and R. Calderbank, "Utility-optimal random-access control," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 1135-1147, Aug. 2007.
- [4] X. Wang and K. Kar, "Cross-layer rate control for end-to-end proportional fairness in wireless networks with random access," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 1548-1559, Aug. 2006.
- [5] —, "Distributed algorithms for max-min fair rate allocation in aloha networks," in *Proc. 42nd Annual Allerton Conf.*, Urbana-Champaign, IL, Oct. 2004.
- [6] L. Chen, S. Low, and J. Doyle, "Joint congestion control and media access control design for ad hoc wireless networks," in *Proc. of IEEE INFOCOM*, Miami, FL, Mar. 2005.
- [7] P. Gupta and A. Stolyar, "Optimal throughput allocation in general random-access networks," in *Proc. CISS*, Princeton, NJ, Mar. 2006.
- [8] A. H. Mohsenian Rad, J. Huang, M. Chiang, and V. W. S. Wong, "Utility-optimal random access without message passing," *IEEE Trans. Wireless Commun.* (to appear), 2009.
- [9] L. Chen, S. H. Low, and J. C. Doyle, "Contention control: A game theoretic approach," in *Proc. IEEE Conf. Decision Control*, New Orleans, LA, Dec. 2007.
- [10] A. B. MacKenzie and S. B. Wicker, "Stability of multipacket slotted aloha with selfish users and perfect information," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar. 2003.
- [11] Y. Sagduyu and A. Ephremides, "Power control and rate adaptation as stochastic games for random access," in *Proc. IEEE Conf. Decision Control*, Maui, HI, Dec. 2003.
- [12] J. Huang, R. Berry, and M. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 1074-1084, May 2006.
- [13] D. P. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.
- [14] F. Kelly, "Charging and rate control for elastic traffic," *European Trans. Telecommun.*, vol. 8, pp. 33-37, 1997.
- [15] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, pp. 556-567, Oct. 2000.
- [16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [17] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.
- [18] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Sci., 2004.
- [19] J. Zhang and D. Zheng, "A stochastic primal-dual algorithm for joint flow control and mac design in multi-hop wireless networks," in *Proc. CISS*, Princeton, NJ, Mar. 2006.
- [20] R. Jain, W. Hawe, and D. Chiu, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *Tech. Rep. DEC-TR-301*, Sept. 1984.
- [21] J. Lee, M. Chiang, and R. Calderbank, "Jointly optimal congestion and contention control in wireless ad hoc networks," *IEEE Commun. Lett.*, vol. 10, no. 3, pp. 216-218, Mar. 2006.
- [22] C. Bordenave and D. M. A. Proutiere, "Performance of random medium access control an asymptotic approach," in *Proc. ACM Sigmetrics*, Annapolis, MD, June 2008.
- [23] J. Liu, A. Stolyar, M. Chiang, and H. V. Poor, "Queue back-pressure random access in multi-hop wireless networks: Optimality and stability," *IEEE Trans. Inform. Theory* (submitted), May 2008.
- [24] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportionally fair rates using local information in aloha networks," *IEEE Trans. Automatic Control*, vol. 49, no. 10, pp. 1858-1862, Oct. 2004.



Amir-Hamed Mohsenian-Rad (S'04) received his B.Sc. degree from Amir-Kabir University of Technology (Tehran, Iran) in 2002, M.Sc. degree from Sharif University of Technology (Tehran, Iran) in 2004, and Ph.D. degree from The University of British Columbia (Vancouver, Canada) in 2008, all in electrical engineering. From March to July 2007, he was also a visiting scholar at Princeton University (Princeton, NJ). Currently, Dr. Mohsenian-Rad is a post-doctoral research and teaching fellow at the University of British Columbia. As a graduate student, he granted the UBC Graduate Fellowship as well as the Pacific Century Graduate Scholarship from the British Columbia Provincial Government. He currently serves as technical program committee (TPC) member for the IEEE Global Telecommunications Conference (Globecom'09), the IEEE International Conference on Communications (ICC'09), and the IEEE Consumer Communications and Networking Conference (CCNC'09). His research interests are in the area of optimization as well as game theory and their applications in computer communications and wireless networking.



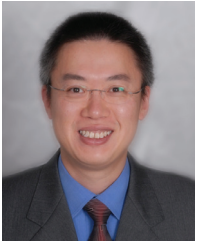
Jianwei Huang (S'01-M'06) received the B.S. degree in electrical engineering from Southeast University, Nanjing, China, in 2000, and the M.S. and Ph.D. degrees in electrical and computer engineering from Northwestern University, Evanston, IL, in 2003 and 2005, respectively. He is currently an Assistant Professor in the Department of Information Engineering at the Chinese University of Hong Kong, China. He was a Postdoctoral Research Associate at Princeton University, Princeton, NJ, during 2005-2007. He also had summer internships in Motorola,

Arlington Heights, IL, in 2004 and 2005. He conducts research in the area of nonlinear optimization and game theoretical analysis of communication networks, with current focus on network economics, cognitive radio networks, broadband communication networks, and multimedia over wireless. Dr. Huang has served as an Associate Editor of Journal of Computer and Electrical Engineering, a Lead Guest Editor of IEEE Journal of Selected Areas in Communications, a Lead Guest Editor of Journal of Advances in Multimedia, and a TPC Co-Chair of International Conference on Game Theory for Networks (GameNets'09). He was the recipient of a 2001 Walter P. Murphy Fellowship at Northwestern University and a 1999 Chinese National Excellent Student Award.



Mung Chiang (S'00-M'03-SM'08) is an Associate Professor of Electrical Engineering, and an Affiliated Faculty of Applied and Computational Mathematics and of Computer Science, at Princeton University. He received the B.S. (Honors) in Electrical Engineering and Mathematics, M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 1999, 2000, and 2003, respectively, and was an Assistant Professor at Princeton University 2003-2008. His research areas include optimization, distributed control, and stochastic analysis of communication networks, with applications to the Internet, wireless networks, broadband access networks, and content distribution.

His awards include Presidential Early Career Award for Scientists and Engineers 2008 from the White House, Young Investigator Award 2007 from ONR, CAREER Award 2005 from NSF, TR35 Young Innovator Award 2007 from Technology Review, Young Researcher Award Runner-up 2004-2007 from Mathematical Programming Society, Frontiers of Engineering Symposium participant 2008 from NAE, Wentz Junior Faculty Award 2005 and SEAS Teaching Commendation 2007 from Princeton University, and Hertz Foundation Fellowship. His paper awards include ISI citation Fast Breaking Paper in Computer Science, IEEE INFOCOM Best Paper Finalist, and IEEE GLOBECOM Best Student Paper. He also has 16 patents filed and given plenary talks at IEEE workshops RAWNET and SAM. His guest and associate editorial services include IEEE/ACM Trans. Netw., IEEE Trans. Inform. Theory, IEEE J. Sel. Area Comm., IEEE Trans. Comm., IEEE Trans. Wireless Comm., and J. Optimization and Engineering.



Vincent W.S. Wong (SM'07) received the B.Sc. degree from the University of Manitoba, Winnipeg, MB, Canada, in 1994, the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada, in 1996, and the Ph.D. degree from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2000. From 2000 to 2001, he worked as a systems engineer at PMC-Sierra Inc. He joined the Department of Electrical and Computer Engineering at UBC in 2002 and is currently an Associate Professor. His research interests are in resource and

mobility management for wireless mesh networks, wireless sensor networks, and heterogeneous wireless networks. Dr. Wong is an Associate Editor of the IEEE Transactions on Vehicular Technology and an Editor of KICS/IEEE Journal of Communications and Networks. He serves as TPC member in various conferences, including IEEE Infocom'09, ICC'09, and Globecom'09. He is a senior member of the IEEE and a member of the ACM.