# Utility-Optimal Random Access without Message Passing

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*— Random access has been studied for decades to achieve simple wireless medium access control (MAC). Some of the distributed scheduling algorithms for throughput or utility maximization also take the form of random access, although extensive message passing among the nodes is required. In this paper, we would like to answer this question: is it possible to design a MAC algorithm that can achieve the optimal network performance without message passing? We provide the first positive answer to this question in a simple Aloha-type random access wireless ad-hoc network. In particular, we propose a distributed random medium access control algorithm based on the framework of network utility maximization (NUM). We prove the convergence of our algorithm for certain sufficient conditions on the system parameters, e.g., with a large enough user population. If each wireless node is capable of decoding the source MAC address of the transmitter from the interferring signal, then our algorithm indeed converges to the global optimal solution of the NUM problem. If such decoding is inaccurate, then the algorithm still converges, although optimality may not be always guaranteed.**

***Index Terms*— Network utility maximization, random access, non-convex optimization, eliminating message passing.**

## I. Introduction

For over thirty years, researchers have studied how well simple random access protocols can work. Since 1992, some of the distributed scheduling algorithms for throughput maximization take the form of random access [1]–[3], although message passing among the nodes is required. Similarly, utility optimization with infinite backlog has been achieved with the help of message passing [4]–[6]. Very recently, in [7] a CSMA-type random access algorithm without message passing is developed that is proved to be utility-optimal. However, it still remains open whether even simpler protocols, such as Aloha-type wireless random access control without carrier sensing, can also achieve utility-optimality. In this paper, we provide the first positive answer to this question.

Manuscript received on December 24, 2007; revised on July 7, 2008; and accepted on August 30, 2008. 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 Chinese University of Hong Kong, as well as NSF CNS-0720570, ONR N00014-07-1-0864, and AFOSR FA9550-06-1-0297. The review of this paper was coordinated by Prof. Guoliang Xue.

A. H. Mohsenian-Rad and V.W.S. Wong are with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada, emails: {*hamed, vincentw*}*@ece.ubc.ca*. J. Huang is with the Information Engineering Department, Chinese University of Hong Kong, Hong Kong, China, email: *jwhuang@ie.cuhk.edu.hk*. M. Chiang is with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, email: *chiangm@princeton.edu*.

In the existing contention-based medium access control (MAC) protocols, there is usually a tradeoff between system performance (e.g., throughput and fairness) and the amount of explicit message passing required among users. One example is the IEEE 802.11 distributed coordination function (DCF), where users do not explicitly exchange any message related to their transmission probabilities[1] and adapt their transmission probabilities only based on the binary implicit feedback from the network (e.g., collision or not). This typically leads to low throughput and unfair resource allocation [8]. On the other hand, several MAC algorithms (e.g., [4]–[6]) have been designed based on the framework of network utility maximization (NUM) which lead to the optimal system performance without taking the signalling overhead into account. However, these algorithms require extensive frequent message passing among users. Considering the fact that any message transmission leads to additional contention in the network, it is of practical importance to design a MAC algorithm that can achieve the optimal performance without message passing.

In this paper, we propose an efficient NUM-based MAC algorithm without message passing based on our recent results in [9]. Compared with the previous algorithms (e.g., [4]–[6]), our NUM-based random access algorithms in [9] support a wider range of utility functions, converge faster, and allow fully asynchronous operations among users. However, message exchange is still needed in [9]. In this paper, we show that in the simple case of a fully interfered wireless network topology where all users are within the interference range of each other (e.g., as in wireless personal and local area networks), we can completely eliminate the need for message passing. We prove the convergence of our algorithm for certain sufficient conditions on the system parameters, e.g., sufficiently large user population. If each wireless node is capable of decoding the source MAC address of the transmitter from the interferring signal, then our algorithm converges to the global optimal solution of the NUM problem. If such decoding is inaccurate, the algorithm still converges, although optimality may not be guaranteed. The estimation techniques we used in this paper are related to [10], [11], although our estimation model is more elaborate and captures more information (i.e., each user's transmission probability). Simulation results show that our algorithm is robust to changes in user populations and channel conditions. These encouraging results provide important in-

[1] In this paper, we use the term "messages" to denote control signals that are explicitly related to wireless users' transmission probabilities. We notice that IEEE 802.11 DCF does not have any explicit message passing, although it has various other control signals (e.g., RTS/CTS/ACK).

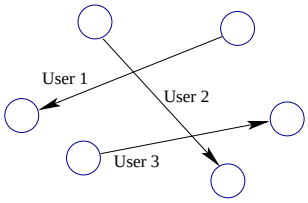


Fig. 1. A single-hop wireless ad-hoc network with $N$ = 3 users. Each user includes a wireless link and its dedicated transmitter and receiver nodes.

sights to design distributed optimal random MAC algorithms without message passing for more general topologies.

The rest of this paper is organized as follows. The system model and problem formulation are described in Section II. Our algorithm is presented in Section III. Convergence and optimality of the proposed algorithm are analytically proved in Section IV. Simulation results are shown in Section V. Conclusions and future work are discussed in Section VI.

## II. System Model

Consider a single-hop wireless ad-hoc network with $\mathcal{N} = \{1, \ldots, N\}$ as the set of wireless links. Each link, together with its dedicated transmitter and receiver nodes, is called a *user*. A sample network with 3 users is shown in Fig. 1. We assume that each user's receiver node can hear other user's transmissions. Thus, each user interferes with all other users. This models some important wireless networks including wireless personal area networks where wireless devices interact with each other (e.g., in an office) as well as indoor wireless local area networks where nodes interact with each other and an access point (e.g., in a large conference room). Time is divided into equal-length slots. At each time slot, user $i$ transmits with probability $p_i \in \mathcal{P}_i = [P_i^{\min}, P_i^{\max}]$, where $0 < P_i^{\min} < P_i^{\max} < 1$. A transmission is successful if it is the only transmission in the current time slot. Similar to [4], [6], we assume *infinite-backlog* at the transmitter nodes of the users. Let $r_i$ denote the average rate for user $i$. We have[2] [12]:

$$r_i(\boldsymbol{p}) = \gamma_i p_i \prod_{j \in \mathcal{N}\setminus\{i\}} (1 - p_j), \qquad \forall\, i \in \mathcal{N}, \tag{1}$$

where $\boldsymbol{p} = (p_i, \forall i \in \mathcal{N})$ is the vector of all users' transmission probabilities and $\gamma_i$ denotes the peak data rate for user $i$. Here, we assume that either the channel is fixed or it changes very slowly such that $\gamma_i$ can assume to be fixed for all $i \in \mathcal{N}$.

Each link $i \in \mathcal{N}$ maintains a *utility* which is an increasing and concave function of $r_i$ and indicates link $i$'s level of satisfaction on its average data rate. The utility of link $i$ is denoted by $u_i(r_i(\boldsymbol{p}))$ which is also a function of $\boldsymbol{p}$. We are interested in finding the value of $\boldsymbol{p}$ that solves the following *network utility maximization* (NUM) problem [13]:

$$\max_{\boldsymbol{p} \in \mathcal{P}} \ \sum_{i \in \mathcal{N}} \ u_i(r_i(\boldsymbol{p})), \tag{NUM}$$

where $\mathcal{P} = \{\boldsymbol{p} : p_i \in \mathcal{P}_i, \forall i \in \mathcal{N}\}$, and the utility functions are $\alpha$-fair [14]. That is, $u_i(r_i(\boldsymbol{p})) = (1-\alpha)^{-1} r_i(\boldsymbol{p})^{1-\alpha}$ if $\alpha \in (0,1) \cup (1, \infty)$, and $u_i(r_i(\boldsymbol{p})) = \log r_i(\boldsymbol{p})$, if $\alpha = 1$.

[2]We notice that if the underlying communication channel is lossy and we know exact packet error rate, then the average rate model in (1) needs to be modified to incorporate the impact of packet loss due to channel imperfections as in (7). However, in this paper, we assume that the packet error rate is *not* known; thus, the average data rate is estimated as the current form in (1).

## III. Algorithm

*1) Local Optimization:* For each wireless user $i$, consider the following *local* optimization problem:

$$\max_{p_i \in \mathcal{P}_i} \ \sum_{j \in \mathcal{N}} \ u_j(r_j(p_i, \boldsymbol{p}_{-i})), \tag{LOCAL-NUM}$$

where $\boldsymbol{p_{-i}} = (p_j, \ \forall j \in \mathcal{N}\setminus\{i\})$ denotes the transmission probabilities of all users *other than* user $i$. To solve optimization problem (LOCAL-NUM), user $i$ will choose $p_i$ to maximize the *total* network utility, *assuming* that none of the other users change their transmission probabilities.

*Theorem 1:* For each user $i \in \mathcal{N}$, the unique global optimal solution of problem (LOCAL-NUM) is $p_i^*(\boldsymbol{p}_i) = f_i(\boldsymbol{p}_{-i})$, where the mapping function $f_i(\boldsymbol{p}_{-i})$ is defined as:

$$f_i(\boldsymbol{p}_{-i}) = \left[1/\left(1 + \sqrt[\alpha]{v_i(\boldsymbol{p}_{-i})}\ \right)\right]_{P_i^{\min}}^{P_i^{\max}}. \tag{2}$$

Here $[x]_b^a = \max\left[\min\left[x, a\right], b\right]$ and

$$v_i(\boldsymbol{p}_{-i}) = {\gamma_i}^{\alpha-1} \sum_{j \in \mathcal{N}\setminus\{i\}} (1/\gamma_j)^{\alpha-1} (1/p_j - 1)^{\alpha-1}. \tag{3}$$

The proof of Theorem 1 is similar to that of [9, Theorem 1] and is omitted for brevity. It is clear that to compute $f_i(\boldsymbol{p}_{-i})$ in (2), the only information user $i$ needs from other users is $v_i(\boldsymbol{p}_{-i})$. If each user $i$ can *estimate* the value of

$$m_j = (1/\gamma_j)^{\alpha-1} (1/p_j - 1)^{\alpha-1}, \qquad \forall\, j \in \mathcal{N}\setminus\{i\}, \tag{4}$$

then it can compute $v_i(\boldsymbol{p}_{-i}) = {\gamma_i}^{\alpha-1} \sum_{j \in \mathcal{N}\setminus\{i\}} m_j$ and set its transmission probability $p_i = f_i(\boldsymbol{p}_{-i})$. Note that for each $j \in \mathcal{N}\setminus\{i\}$, $m_j$ is bounded between $M^{\min}$ and $M^{\max}$, which are defined as follows. If $\alpha \geq 1$, then $M^{\min} = (1/\gamma^{\max})^{\alpha-1}$ $(1/P^{\max} - 1)^{\alpha-1}$ and $M^{\max} = (1/\gamma^{\min})^{\alpha-1}(1/P^{\min} - 1)^{\alpha-1}$, where $P^{\min} = \min_{i \in \mathcal{N}} P_i^{\min}$, $P^{\max} = \max_{i \in \mathcal{N}} P_i^{\max}$, $\gamma^{\min} = \min_{i \in \mathcal{N}} \gamma_i$, and $\gamma^{\max} = \max_{i \in \mathcal{N}} \gamma_i$. If $\alpha < 1$, then $M^{\min} = (1/\gamma^{\min})^{\alpha-1}(1/P^{\min} - 1)^{\alpha-1}$ and $M^{\max} = (1/\gamma^{\max})^{\alpha-1}(1/P^{\max} - 1)^{\alpha-1}$. As shown in [9, Section IV-A], if each user $i$ updates its transmission probability $p_i$ according to (2), then the system will converge to the optimal solution of problem (NUM). The question is how to obtain the values of $m_j$ for all $j \neq i$. Next, we show how this can be achieved through local observations of the shared channel.

*2) Learning from Contention History:* From (4), we can see that only the values of $\gamma_j$ and $p_j$ are required to calculate the value of message $m_j$. Notice that $\alpha$ is the same for all users. The value of the peak rate $\gamma_j$ depends on the channel gain between the transmitter and receiver of user $j$. Thus, it can only be measured by user $j$ and then announced to the whole network once user $j$ joins the network. The remaining task is to determine how to obtain the value of $p_j$.

From user $i$'s viewpoint, any time slot falls into one of the following possible states: *idle* (no user transmits), *busy* (at least one other user transmits), *success* (user $i$ transmits successfully), and *failure* (user $i$ transmits but it fails). Let $p_i^{\text{idle}}$, $p_i^{\text{busy}}$, $p_i^{\text{succ}}$, and $p_i^{\text{fail}}$ denote the probabilities of experiencing these four states, respectively. Also let $p_{i,j}^{\text{err}}$ denote the *packet error rate* of the communication channel from the transmitter

node of user $j$ to the receiver node of user $i$. We have:

$$p_i^{\text{idle}} = \prod_{j\in\mathcal{N}}(1-p_j), \tag{5}$$

$$\begin{aligned} p_i^{\text{busy}} &= (1-p_i)\left(1-\prod_{j\in\mathcal{N}\setminus\{i\}}(1-p_j)\right) \\ &= (1-p_i)-\prod_{j\in\mathcal{N}}(1-p_j) \\ &= (1-p_i)-p_i^{\text{idle}}, \end{aligned} \tag{6}$$

$$\begin{aligned} p_i^{\text{succ}} &= p_i\left(\prod_{j\in\mathcal{N}}(1-p_j)\right)\left(1-p_{i,i}^{\text{err}}\right) \\ &= (p_i/(1-p_i))\left(\prod_{j\in\mathcal{N}}(1-p_j)\right)\left(1-p_{i,i}^{\text{err}}\right) \\ &= (p_i/(1-p_i))\; p_i^{\text{idle}}\,(1-p_{i,i}^{\text{err}}), \end{aligned} \tag{7}$$

and

$$\begin{aligned} p_i^{\text{fail}} &= p_i\left(\left(1-\tfrac{1}{1-p_i}p_i^{\text{idle}}\right)+p_{i,i}^{\text{err}}-\left(1-\tfrac{1}{1-p_i}p_i^{\text{idle}}\right)p_{i,i}^{\text{err}}\right) \\ &= p_i\left(\left(1-\frac{1}{1-p_i}p_i^{\text{idle}}\right)+\left(\frac{1}{1-p_i}p_i^{\text{idle}}\right)p_{i,i}^{\text{err}}\right) \\ &= p_i\left(1-\left(\frac{1}{1-p_i}p_i^{\text{idle}}\right)\left(1-p_{i,i}^{\text{err}}\right)\right) \\ &= p_i-\left(\frac{p_i}{1-p_i}\;p_i^{\text{idle}}\right)\left(1-p_{i,i}^{\text{err}}\right). \end{aligned} \tag{8}$$

From (5), the channel is idle if all users are silent. From (6), the channel is busy from user $i$'s viewpoint as long as user $i$ is silent (so that it can sense the channel) and at least one other user $j\neq i$ is transmitting packets. The former has probability $(1-p_i)$ while the latter has probability $(1-\prod_{j\in\mathcal{N}\setminus\{i\}}(1-p_j))$. From (7), the transmission from the transmitter node of user $i$ to the receiver node of user $i$ is successful if the transmitter node of user $i$ transmits the packet, no other user $j\neq i$ transmits any packet at the same time, and the transmitted packet is not corrupted. The latter happens with probability $(1-p_{i,i}^{\text{err}})$. Finally, from (8), user $i$ observes a failure slot if it transmits a packet and the transmitted packet either collides with some other transmission(s), or gets corrupted, or both. We notice that the probability values $p_i^{\text{idle}}$ and $p_i^{\text{busy}}$ do not depend on the value of $p_{i,i}^{\text{err}}$, while $p_i^{\text{succ}}$ and $p_i^{\text{fail}}$ do.

Since user $i$ knows the local transmission probability $p_i$, it can estimate $p_i^{\text{idle}}$ using either (5) or any of the expressions in (6)-(8). However, it is clear that user $i$ is still unable to estimate the *individual* transmission probability $p_j$ for any $j\neq i$ even if it can accurately estimate all the state probabilities $p_i^{\text{idle}}$, $p_i^{\text{busy}}$, $p_i^{\text{succ}}$, and $p_i^{\text{fail}}$. In fact, finding the values of *individual* transmission probabilities requires gathering more *individual* information from other users as we will explain next.

Recall that, at a busy time slot seen by user $i\in\mathcal{N}$, at least one other user transmits. Since users can hear each other, user $i$ may successfully *decode* the transmission of user $j\neq i$ with a probability that can be obtained as:

$$\begin{aligned} p_{i,j}^{\text{decd}} &= p_j(\textstyle\prod_{l\in\mathcal{N}\setminus\{j\}}(1-p_l))(1-p_{i,j}^{\text{err}}) \\ &= (p_j/(1-p_j))\left(\textstyle\prod_{l\in\mathcal{N}}(1-p_l)\right)(1-p_{i,j}^{\text{err}}). \end{aligned} \tag{9}$$

Let $n_{i,j}^{\text{decd}}$ denote the number of time slots between any two consecutive successful decoding of transmissions of user $j$ by user $i$. Random variable $n_{i,j}^{\text{decd}}$ has an i.i.d. geometric distribution with probability $p_{i,j}^{\text{decd}}$. Therefore, we indeed have:

$$p_{i,j}^{\text{decd}} = 1/(1+\bar{n}_{i,j}^{\text{decd}}), \tag{10}$$

where $\bar{n}_{i,j}^{\text{decd}}$ is the mean value of $n_{i,j}^{\text{decd}}$ and can be *locally* estimated by user $i$ through observation of the channel contention history. Let $T_{j,\text{decd}}^i$ denote the set of time slots at which user $i$ decodes the transmissions of user $j\neq i$. We estimate $\bar{n}_{i,j}^{\text{decd}}$ iteratively through a low-pass filter:

$$\begin{aligned} &\bar{n}_{i,j}^{\text{decd}}(t+1) = \\ &\quad(1-\varrho_{i,j}(t))\bar{n}_{i,j}^{\text{decd}}(t)+\varrho_{i,j}(t)n_{i,j}^{\text{decd}}(t)I\{t\in T_{j,\text{decd}}^i\}, \end{aligned} \tag{11}$$

where $\bar{n}_{i,j}^{\text{decd}}(t)$ and $n_{i,j}^{\text{decd}}(t)$ denote the estimation of $\bar{n}_{i,j}^{\text{decd}}$ and the measurement of $n_{i,j}^{\text{decd}}$ at time slot $t$, respectively, and $I\{\cdot\}$ is an indication function. Here $\varrho_{i,j}$ is a *diminishing* stepsize. Notice that in practice, the transmitted signal by the transmitter node of user $j$ can be decoded by the network interface of user $i$'s receiver node; however, as its destination MAC address is not the same as the one in the receiver node of user $i$, the packet is simply discarded. In our proposed scheme, the receiver node of user $i$ needs to obtain the sender's MAC address from the packet header before discarding the packet[3].

Similarly, let $n_i^{\text{idle}}$ denote the number of non-idle time slots that user $i$ observes between any two consecutive idle time slots. User $i$ can estimate $p_i^{\text{idle}}$ as follows [10]:

$$p_i^{\text{idle}} = 1/(1+\bar{n}_i^{\text{idle}}), \tag{12}$$

where $\bar{n}_i^{\text{idle}}$ is the mean value of $n_i^{\text{idle}}$. Substituting (5), (10), and (12) into (9), for each $j\in\mathcal{N}\setminus\{i\}$, we have:

$$1/p_j - 1 = \left((1+\bar{n}_{i,j}^{\text{decd}})/(1+\bar{n}_i^{\text{idle}})\right)\;(1-p_{i,j}^{\text{err}}). \tag{13}$$

Let $T_{\text{idle}}^i$ denote the set of time slots at which user $i$ observes an idle time slot. We estimate $\bar{n}_i^{\text{idle}}$ iteratively:

$$\bar{n}_i^{\text{idle}}(t+1) = (1-\rho_i(t))\bar{n}_i^{\text{idle}}(t)+\rho_i(t)n_i^{\text{idle}}(t)I\{t\in T_{\text{idle}}^i\}, \tag{14}$$

where $\bar{n}_i^{\text{idle}}(t)$ and $n_i^{\text{idle}}(t)$ denote the estimation of $\bar{n}^{\text{idle}}$ at time $t$ and the measurement of $n_i^{\text{idle}}$ at time $t$, respectively. Here $\rho_i$ is a diminishing stepsize. Based on the asynchronous stochastic approximation theory [15], we know that the error for estimating the decoding and idle probabilities decrease to zero (i.e., accurate estimations) when users do not change their transmission probabilities (i.e., the system is stationary).

For each user $i$ and any other user $j\neq i$, given the system parameters $\gamma_j$, $\bar{n}_{i,j}^{\text{decd}}$ and $\bar{n}_i^{\text{idle}}$, we define:

$$m_j^i(t) = (1/\gamma_j)^{\alpha-1}\left((1+\bar{n}_{i,j}^{\text{decd}}(t))/(1+\bar{n}_i^{\text{idle}}(t))\right)^{\alpha-1}, \tag{15}$$

where $m_j^i(t)$ denotes the estimation of $m_j$ made by user $i$ at time slot $t$. In general, we have:

$$m_j^i(t) = \beta_j^i(t)\; m_j(t), \tag{16}$$

where $\beta_j^i(t)>0$ is the *estimation gain*, which can represent either accurate estimation (i.e., $\beta_j^i(t)=1$), over-estimation (i.e., $\beta_j^i(t)>1$) or under-estimation (i.e., $\beta_j^i(t)<1$). From (4) and (13), if the estimations on $\bar{n}_{i,j}^{\text{decd}}$ and $\bar{n}_i^{\text{idle}}$ are accurate (i.e., they accurately represent the mean values of random variables $n_{i,j}^{\text{decd}}$ and $n_i^{\text{idle}}$, respectively) and the channel is also

[3]We notice that, in general, decoding header information of packets is more accurate than payload information. For example, in HiperLAN, header is transmitted at lower bit rate and higher reliability. Therefore, decoding MAC address from the header is less prone to error due to channel imperfections.

perfect (i.e., it has zero packet error rate), then $\beta_j^i(t) = 1$ and $m_j^i(t) = m_j(t)$ for all $j \in \mathcal{N}\backslash\{i\}$. Notice that if the value of the existing packet error rate $p_{i,j}^{\text{err}}$ is known (e.g., via measurements at the physical layer), then we can redefine:

$$m_j^i(t) = (1/\gamma_j)^{\alpha-1} \frac{\left((1+\bar{n}_{i,j}^{\text{decd}}(t))/(1+\bar{n}_i^{\text{idle}}(t))\right)^{\alpha-1}}{(1-p_{i,j}^{\text{err}})^{\alpha-1}}, \quad (17)$$

and obtain a more accurate estimation of $m_j(t)$ by canceling out the effect of wireless channel imperfections. However, in this paper, we consider the general case where the packet error rates are *not* known by the users.

For each user $i \in \mathcal{N}$ and for all $j \in \mathcal{N}\backslash\{i\}$, let $T_{j,m}^i$ denote the set of all time slots at which user $i$ updates its estimation of $m_j^i$ according to (17). We select set $T_{j,m}^i$ such that as time goes by, the minimum time difference between any two consecutive slots in the union of sets $\{T_{j,m}^i, \forall j \in \mathcal{N}\backslash\{i\}\}$ increases. As a result, for each user $j$, we update $m_j^i$ less frequently to collect more samples of $n_i^{\text{idle}}$ and $n_{i,j}^{\text{decd}}$. Thus, the estimations of mean values $\bar{n}_i^{\text{idle}}$ and $\bar{n}_{i,j}^{\text{decd}}$ improve gradually and become asymptotically accurate. We also reset the diminishing step-sizes $\rho_i$ and $\varrho_{i,j}$ to 1 after each $t \in T_{j,m}^i$ so that the errors in previous estimations do not affect new estimations. Based on these assumptions, there exists a constant $\beta_j^i > 0$ such that $\lim_{t\to\infty} \beta_j^i(t) = \beta_j^i$. From (15) and (16), we have:

$$\beta_j^i = 1/(1-p_{i,j}^{\text{err}})^{\alpha-1}, \quad \forall i,j \in \mathcal{N},\ i \neq j. \quad (18)$$

From (18), if the communication channel is perfect, then $\beta_j^i = 1$ and all estimations are *asymptotically accurate*. For a lossy channel, if $\alpha < 1$, then $\beta_j^i < 1$ and $m_j^i$ is *asymptotically under-estimated* for all $j \neq i$. On the other hand, if $\alpha > 1$, then $\beta_j^i > 1$ and $m_j^i$ is *asymptotically over-estimated*.

*3) Distributed MAC Algorithm:* Our proposed distributed random access algorithm without explicit message passing (except when each user joins or leaves the network) is shown in Algorithm 1. In this algorithm, each user $i \in \mathcal{N}$ continuously updates $\bar{n}_i^{\text{idle}}$ and $\bar{\boldsymbol{n}}_i^{\text{decd}} = (\bar{n}_{i,j}^{\text{decd}},\ \forall j \in \mathcal{N}\backslash\{i\})$ based on its local observations from the channel to estimate $\boldsymbol{m}^i = (m_j^i,\ \forall j \in \mathcal{N}\backslash\{i\})$. Then, it chooses $p_i$ according to (2) with $v_i = \gamma_i^{\alpha-1} \sum_{j\in\mathcal{N}\backslash\{i\}} m_j^i$. Here $T_{i,p}$ denotes the set of all time slots at which user $i$ updates its transmission probability $p_i$. Notice that the updates are *asynchronous* across wireless users which include synchronous updates as a special case.

**Algorithm 1** *Executed by each user $i \in \mathcal{N}$.*

1: Allocate memory for $p_i$ and $\boldsymbol{m}^i = (m_1^i, \cdots, m_N^i)$.
2: Allocate memory for $\bar{n}_i^{\text{decd}}$ and $\bar{\boldsymbol{n}}_i^{\text{decd}} = (\bar{n}_{i,1}^{\text{decd}}, \cdots, \bar{n}_{i,N}^{\text{decd}})$.
3: Randomly choose $p_i \in \left[P_i^{\min}, P_i^{\max}\right]$.
4: Randomly choose $m_j^i \in \left[M^{\min}, M^{\max}\right]$ for all $j \in \mathcal{N}$.
5: Choose $\bar{n}_i^{\text{idle}} = 1$ and $\bar{n}_{i,j}^{\text{decd}} = 1$ for all $j \in \mathcal{N}$.
6: Broadcast the fixed data rate $\gamma_i$ to all other users.
7: **repeat**
8: Transmit with probability $p_i$.
9: Update $\bar{n}_i^{\text{idle}}$ and $\bar{\boldsymbol{n}}_i^{\text{decd}}$ according to Eqs. (14) and (11).
10: **if** $t \in T_{i,p}$ **then**
11: Update $p_i = \left[1/\left(1+\sqrt[\alpha]{\gamma_i^{\alpha-1}\sum_{j\in\mathcal{N}\backslash\{i\}} m_j^i}\ \right)\right]_{P_i^{\min}}^{P_i^{\max}}$.
12: **end if**
13: **if** $t \in T_{j,m}^i$ **then**
14: Update $m_j^i$ according to Eq. (15).
15: **end if**
16: **until** the user decides to leave the network.
17: Broadcast termination message.

## IV. Convergence and Optimality

For each $i \in \mathcal{N}$, and at any time $t \in T_{i,p}$, Algorithm 1 updates

$$p_i(t+1) = f_i'(\boldsymbol{p}_{-i}, t) = \left[1/\left(1+\sqrt[\alpha]{v_i'(\boldsymbol{p}_{-i}, t)}\ \right)\right]_{P_i^{\min}}^{P_i^{\max}},$$

where $v_i'(\boldsymbol{p}_{-i}, t) = \sum_{j\in\mathcal{N}\backslash\{i\}} (\gamma_i/\gamma_j)^{\alpha-1} (1/p_j - 1)^{\alpha-1} \beta_j^i(t)$. For any $t \geq 0$, we define $f'(\boldsymbol{p}, t) = (f_i'(\boldsymbol{p}_{-i}, t), \forall\, i \in \mathcal{N})$. Here $f'(\boldsymbol{p}, t)$ is a time-varying vector mapping. Since $\beta_j^i(t)$ approaches $\beta_j^i$ as $t \to \infty$ for all $i, j \in \mathcal{N}$, the sequence of mapping $\{f'(\boldsymbol{p}, t)\}$ converges to a unique mapping $f'(\boldsymbol{p}, \infty)$ as $t \to \infty$. That is, for any $\boldsymbol{p} \in \mathcal{P}$ and any $\epsilon > 0$, there exists $t_\epsilon \geq 0$ such that $\|f'(\boldsymbol{p}, t) - f'(\boldsymbol{p}, \infty)\| < \epsilon$ for all $t \geq t_\epsilon$.

Let $V^{\min}$ and $V^{\max}$ denote the lower and upper bounds on $v_i'(\boldsymbol{p}_{-i}, t)$ for each $i \in \mathcal{N}$ and at any time $t$. If $\alpha \geq 1$, then:

$$V^{\min} = (N-1) M^{\min} (\gamma^{\min})^{\alpha-1}, \quad (19)$$
$$V^{\max} = (N-1) M^{\max} (\gamma^{\max})^{\alpha-1}. \quad (20)$$

If $\alpha < 1$, then we define:

$$V^{\min} = (N-1) M^{\min} (\gamma^{\max})^{\alpha-1}, \quad (21)$$
$$V^{\max} = (N-1) M^{\max} (\gamma^{\min})^{\alpha-1}. \quad (22)$$

*Theorem 2:* Assume that there exists $t_0' \geq 0$ such that for all $t \geq t_0'$ and any $\boldsymbol{p} \in \mathcal{P}$, we have:

$$\frac{\beta^{\max}(t)}{\beta^{\min}(t)} \left(\frac{|1-\alpha|}{\alpha}\, \Psi\, \Phi(V^{\min}, V^{\max})\right)^2 \left(\frac{\gamma^{\max}}{\gamma^{\min}} \Gamma\right)^{|1-\alpha|} < 1, \quad (23)$$

where $\beta^{\min}(t) = \min_{i,j\in\mathcal{N}} \beta_j^i(t)$, $\beta^{\max}(t) = \max_{i,j\in\mathcal{N}} \beta_j^i(t)$,

$$\Psi = \max\left\{\frac{1}{P^{\min}(1-P^{\min})}, \frac{1}{P^{\max}(1-P^{\max})}\right\}, \quad (24)$$

$$\Gamma = \frac{P^{\max}(1-P^{\min})}{P^{\min}(1-P^{\max})}, \quad (25)$$

and

$$\Phi(V^{\min}, V^{\max}) = \begin{cases} \frac{(V^{\max})^{1/\alpha}}{(1+(V^{\max})^{1/\alpha})^2}, & \text{if } V^{\max} \leq 1, \\ \frac{(V^{\min})^{1/\alpha}}{(1+(V^{\min})^{1/\alpha})^2}, & \text{if } V^{\min} \geq 1, \\ 0.25, & \text{otherwise.} \end{cases} \quad (26)$$

Then, Algorithm 1 globally and asynchronously converges to the unique fixed point of mapping $f'(\boldsymbol{p}, \infty)$.

The proof of Theorem 2 is given in the Appendix. Notice that, at any time $t \geq 0$, $\beta^{\min}(t)$ and $\beta^{\max}(t)$ are bounded:

$$(M^{\min}/M^{\max}) \leq \beta^{\min}(t) \leq \beta^{\max}(t) \leq (M^{\max}/M^{\min}). \quad (27)$$

Therefore, all the terms in (23), except $\Phi$, are bounded and independent of the number of users $N$. Thus, $\Phi$ can be arbitrarily close to 0 if $N$ is large enough. Therefore,

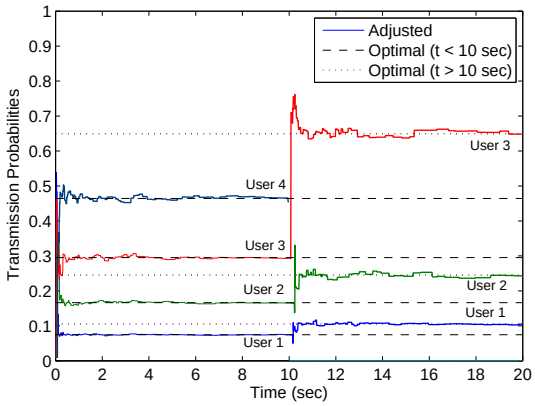


Fig. 2. Simulation results for Algorithm 1 when $\alpha$ = 0.6. The number of users and the features of the communication channel change after $t = 10s$. The optimal transmission probabilities before $t = 10s$ (i.e., dashed lines) and after $t = 10s$ (i.e., dotted lines) are obtained using [9, Algorithm 1].

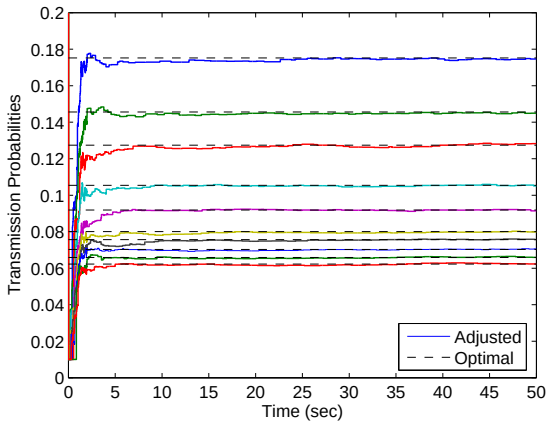


Fig. 3. Trend of the adjusted transmission probabilities when Algorithm 1 is being used, the number of users $N = 10$, and we have: $\alpha$ = 2.

*Corollary 1:* For any choice of system parameters, there exists $\hat{N} > 0$, such that Algorithm 1 converges to the unique fixed point of mapping $f'(\boldsymbol{p}, \infty)$, if the number of users $N > \hat{N}$, i.e., there are enough users competing for the channel.

As an example, assume that $P^{\min} = 0.1$, $P^{\max} = 0.9$, $\alpha = 0.9$, $\gamma^{\max} = \gamma^{\min}$, and for all $t > 1$ we have: $\beta^{\max}(t)/\beta^{\min}(t) \leq 2$. In that case, from Theorem 2, convergence of Algorithm 1 is guaranteed as long as the number of users $N \geq 4$.

We notice that Theorem 2 and Corollary 1 are general and do not depend on the exact values of the estimation errors as $t \to \infty$; however, the performance at the asymptotic fixed point still depends on the accuracy of the estimations. In fact, as long as there are enough learning opportunities for users, learning the behavior of other users can catch up with the environment changes. The following key theorem can be shown for the case of perfect communication channels.

*Theorem 3:* If the channel is perfect such that $\lim_{t\to\infty} \beta^{\min}(t) = \lim_{t\to\infty} \beta^{\max}(t) = 1$, then the fixed point of Algorithm 1 is the global optimal solution of problem (NUM).

The proof of Theorem (3) is similar to that of [9, Theorem 4]. Notice that since $\lim_{t\to\infty} \beta_j^i(t) = 1$, we have $f'(\boldsymbol{p}, \infty) = f(\boldsymbol{p}) = \big(f_i(\boldsymbol{p}_{-i}),\ \forall i \in \mathcal{N}\big)$, where $f_i(\boldsymbol{p}_{-i})$ is as in (2). In fact, when each learning step is accurate, then distributed learning becomes as effective as message passing.

From Theorems 2 and 3, if the channel is perfect and sufficient condition (23) holds, Algorithm 1 asynchronously converges to the unique global optimal solution of non-convex problem (NUM). If the channel is not perfect, although the algorithm still converges, optimality is not always guaranteed.

## V. Simulation Results

To evaluate the performance of our proposed distributed algorithm, we develop a discrete-event simulator that implements Algorithms 1 and the IEEE 802.11 DCF access method. We first consider a network with $N = 4$, $P^{\min} = 0.01$, and $P^{\max} = 0.99$. We set $\gamma_1$ = 6, $\gamma_2$ = 18, $\gamma_3$ = 36, and $\gamma_4$ = 54, all in Mbps. Utility parameter $\alpha = 0.5 < 1$. Notice that none of the previous NUM-based MAC algorithms (e.g, [4]–[6]) support $\alpha$-fair utility functions with $\alpha \in (0, 1)$ because of non-convexity (see [9, Sections II and IV-A]). Each slot is 20 $\mu s$ (as in 802.11a) and the simulation time is $20s$. We assume that from time $t = 0$ to $t = 10s$, the channel is perfect and $N = 4$. Then, from $t = 10s$ to $t = 20s$, the channel is lossy and $N = 3$ (i.e., user 4 leaves the network). Packet error rates are randomly selected between 0 and 0.01 at $t = 10s$ and then become fixed until $t = 20s$. The diminishing stepsizes are selected in the form of $1/t$ while $t$ is the number of time slots. Recall from Section III-.2 that for any $i, j \in \mathcal{N}$, the stepsizes are reset after each $t \in T_{j,m}^i$. Results are shown in Fig. 2. We see that Algorithm 1 converges to a small neighborhood of the optimal values fast. It is robust to the change of user population and channel conditions. Similar results can be obtained for $\alpha \geq 1$ and $N > 4$. For example, the results when $\alpha = 2$ and there are $N = 10$ users are shown in Fig. 3.

Next, we compare the signalling overhead in our algorithm with Algorithm 1 in [9]. We assume that each message value requires two bytes. Simulation results are shown in Fig. 4. The total amount of messages needed depend on the number of iterations for convergence. We refer to Algorithm 1 as the algorithm *without* message passing. It only requires one time message exchange when a user joins the network. We can see that increasing the number of users increases the signalling overhead in both algorithms. However, our algorithm requires significantly less signalling compared to Algorithm 1 in [9]. When the number of users $N = 50$, our algorithm reduces the signalling overhead by more than a factor of 40.

It is well-known that 802.11 DCF has a *short-term fairness* problem, due to binary exponential backoff. Next, we compare 802.11 DCF with Algorithm 1 in terms of both throughput and Jain's fairness index [16]. The short-term fairness is obtained using sliding windows with size of 200 slots. There are $N = 10$ users in the network and their fixed peak rates are randomly selected between 6 and 54 Mbps. Simulation time is $100s$. The results when $\alpha$ varies between 0.5 to 5 are shown in Fig. 5. We see that, parameter $\alpha$ acts as a *knob* to control the tradeoff between efficiency and fairness. By increasing $\alpha$ we can make the system more fair but less efficient (and vice versa). If $\alpha = 0.5$, then the throughput is 29.7% higher than DCF (see Fig. 5(a)). Besides, for any choice of $\alpha \in [0.5, 5]$,

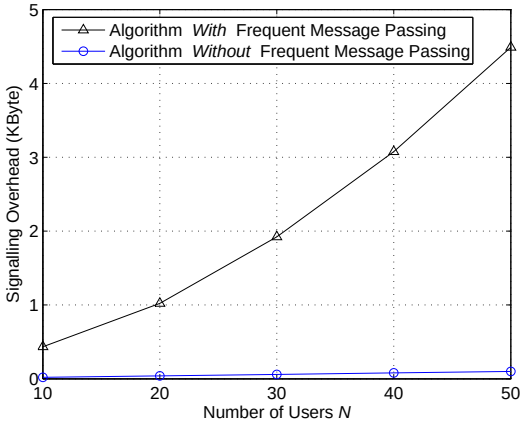


Fig. 4. Comparison between Algorithm 1 in [9] (i.e., the algorithm *with* explicit message passing) and our proposed Algorithm 1 (i.e., the algorithm *without* explicit message passing) in term of the signalling overhead when the number of users varies from 10 to 50.

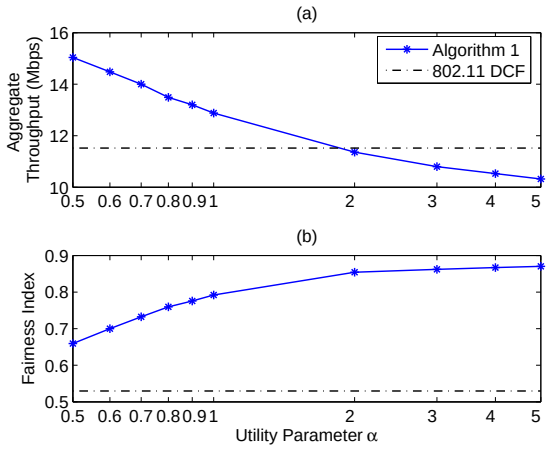


Fig. 5. Comparison between Algorithm 1 and 802.11 DCF when $N = 10$.

the fairness is much better than DCF (Fig. 5(b)).

## VI. Conclusion and Future Work

In this paper, we designed a distributed contention-based MAC algorithm to solve a NUM problem at the link-layer without explicit message passing among users. Our algorithm is asynchronous, enjoys fast convergence, and supports a wider range of utility functions compared to previously proposed NUM-based MAC algorithms. Simulation results show that our algorithm achieves a better efficiency-fairness trade-off compared with the IEEE 802.11 DCF. It is also robust to the changes of user population and channel conditions.

This work represents a first step towards building practical and optimal random access protocols without message passing. Results can be extended in several directions. There are some recent efforts to model more realistic scenarios, where the queues at each node have *finite-backlog* [17], [18]. In [18], a *queue back-pressure random access* algorithm is proposed which aims to adjust the transmission probabilities to achieve optimal network utility and queue stability. The proposed algorithm requires frequent message exchange among the nodes. In this regard, our estimation techniques in Section III can be similarly used to eliminate the need for message passing; thus, significantly reducing the signalling overhead.

We may also relax the need for equal-length time slots, following the techniques in [19, Section IV], where a utility-optimal random access algorithm is proposed for logarithmic utilities in pure (un-slotted) ALOHA systems. It is also worth mentioning that since we limit our study to a topology scenario, where each user can hear transmissions from every other user, the hidden/exposed terminal is not a problem.

Last but not least, an important extension of the current paper is studying the existence of distributed utility-optimal random MAC algorithms without signalling for *general* network topologies, where each user may hear the transmissions from only a *subset* of other users. In this case, the proof of convergence in Theorem 2 will still be valid after slight modifications. However, the performance may not be optimal in general. In fact, it has been recently shown in [20] that in a network with only three users, where the two side users cannot hear each other's transmissions and the utility parameter $\alpha = 0$, the wireless users *cannot* adjust their transmission probabilities to achieve exact optimal network performance without message passing. Therefore, we conjecture that there indeed does *not* exist any algorithm that converges to the optimal network performance without any explicit information exchange in a general topology scenario, unless all users can hear each other's transmissions as studied in this paper.

## Appendix

### A. Proof of Theorem 2

For any $\boldsymbol{p} \in \mathcal{P}$ and $t \geq t_0'$, the Jacobian $J(\boldsymbol{p}, t)$ is defined as an $N \times N$ matrix whose entry in row $i$ and column $j$ is $\partial f_i(\boldsymbol{p}, t)/\partial p_j$. We can show that:

$$\|J'(\boldsymbol{p}, t)\|_\infty \leq \frac{|1-\alpha|}{\alpha}\, \Psi\, \Phi, \tag{28}$$

and

$$\|J'(\boldsymbol{p}, t)\|_1 \leq \frac{|1-\alpha|}{\alpha}\, \frac{\beta^{\max}(t)}{\beta^{\min}(t)}\, \Psi\, \Phi \left( \left(\frac{\gamma^{\max}}{\gamma^{\min}}\right) \Gamma \right)^{1-\alpha}. \tag{29}$$

Let $\tilde{\boldsymbol{p}}, \hat{\boldsymbol{p}} \in \mathcal{P}$. From (23), (28), (29), and by Cauchy-Schwarz inequality we have [21, pp. 635]:

$$\begin{aligned} &\|f'(\tilde{\boldsymbol{p}}, t) - f'(\hat{\boldsymbol{p}}, t)\|_2 \leq \|J'(\boldsymbol{p}, t)\|_2\, \|\tilde{\boldsymbol{p}} - \hat{\boldsymbol{p}}\|_2 \\ &\leq \sqrt{\|J'(\boldsymbol{p}, t)\|_\infty \|J'(\boldsymbol{p}, t)\|_1}\, \|\tilde{\boldsymbol{p}} - \hat{\boldsymbol{p}}\|_2 < \|\tilde{\boldsymbol{p}} - \hat{\boldsymbol{p}}\|_2, \end{aligned}$$

where $\boldsymbol{p}$ is any convex combination of $\tilde{\boldsymbol{p}}$ and $\hat{\boldsymbol{p}}$. Thus, for any $t \geq t_0'$, vector function $f'(\boldsymbol{p}, t)$ is a contraction mapping and has a unique fixed point [21, pp. 183], denoted by $\boldsymbol{p}_t^*$. We also denote the unique fixed point of $f'(\boldsymbol{p}, \infty)$ by $\boldsymbol{p}_\infty^*$. Thus,

$$\|f'(\boldsymbol{p}, t) - \boldsymbol{p}_t^*\|_2 \leq \eta_t\, \|\boldsymbol{p} - \boldsymbol{p}_t^*\|_2 \leq \eta \xi, \tag{30}$$

where $\eta_t = \|J'(\boldsymbol{p}, t)\|$, $\eta = \max_{t > t_0'} \eta_t$, and $\xi = \|\boldsymbol{p} - \boldsymbol{p}_t^*\|_2$. Note that $\eta < 1$, and $\xi$ is bounded. Since $f'(\boldsymbol{p}, t)$ is continuous at $\boldsymbol{p}_t^*$ and $\lim_{t\to\infty} f'(\boldsymbol{p}, t) = f'(\boldsymbol{p}, \infty)$, we indeed have $\lim_{t\to\infty} \boldsymbol{p}_t^* = \boldsymbol{p}_\infty^*$. In other words, for any choice of $\epsilon > 0$, there exists a $t_0 \geq t_0'$, such that:

$$\|\boldsymbol{p}_t^* - \boldsymbol{p}_\infty^*\|_2 \leq \epsilon, \qquad \forall t \geq t_0. \tag{31}$$

Together with (30), we have:

$$\|f'(\boldsymbol{p},t)-\boldsymbol{p}^*_\infty\|_2 \le \|f'(\boldsymbol{p},t)-\boldsymbol{p}^*_t\|_2+\|\boldsymbol{p}^*_t-\boldsymbol{p}^*_\infty\|_2 \le \eta\xi+\epsilon.$$

Similarly, we can show that:

$$\begin{aligned}\|f'(f'(\boldsymbol{p},t),t+1)-\boldsymbol{p}^*_\infty\|_2 \le& \\ &\left\|f'(f'(\boldsymbol{p},t),t+1)-\boldsymbol{p}^*_{t+1}\right\|_2+\left\|\boldsymbol{p}^*_{t+1}-\boldsymbol{p}^*_\infty\right\|_2 \\ &\le \eta\left(\|f'(\boldsymbol{p},t)-\boldsymbol{p}^*_\infty\|_2+\left\|\boldsymbol{p}^*_{t+1}-\boldsymbol{p}^*_\infty\right\|_2\right)+\epsilon \\ &\le \eta(\eta\xi+\epsilon+\epsilon)+\epsilon=\eta(\eta\xi+2\epsilon)+\epsilon.\end{aligned}$$

For any $k \ge 0$, we recursively define $f'^k(\boldsymbol{p},t) = f'(f'^{k-1}(\boldsymbol{p},t),t+k-1)$ where $f'^0=\boldsymbol{p}$. From (32), and by mathematical induction, we can show that for any $k\ge 0$,

$$\begin{aligned}\left\|f'^k(\boldsymbol{p},t)-\boldsymbol{p}^*_\infty\right\|_2 &\le \eta^k\xi+\frac{2\left(1-\eta^k\right)}{1-\eta}\epsilon-\epsilon \\ &< \eta^k\xi+\frac{1+\eta}{1-\eta}\epsilon.\end{aligned} \tag{32}$$

For any $\varepsilon>0$, there exist $k_\varepsilon$ such that if $k\ge k_\varepsilon$, then $\eta^k\xi \le \frac{\varepsilon}{2}$. By choosing:

$$\epsilon=\frac{1-\eta}{1+\eta}\frac{\varepsilon}{2}, \tag{33}$$

we have:

$$\left\|f'^k(\boldsymbol{p},t)-\boldsymbol{p}^*_\infty\right\|_\infty \le \left\|f'^k(\boldsymbol{p},t)-\boldsymbol{p}^*_\infty\right\|_2 < \tfrac{\varepsilon}{2}+\tfrac{\varepsilon}{2}=\varepsilon, \tag{34}$$

where the first inequality comes from the fact that $l_\infty$ norm is always less than or equal to $l_2$ norm. This implies that, starting from any initial point $\boldsymbol{p}\in\mathcal{P}$ at time $t\ge t_0$, the transmission probability of each user $i\in\mathcal{N}$ would be as close as desired to the $i^{\text{th}}$ entry of $\boldsymbol{p}^*_\infty$ after some finite number of slots. For all time instances $t\ge t_0$, we define:

$$\epsilon'_t=\max_{k\ge 0}\left\|\boldsymbol{p}^*_{k+t}-\boldsymbol{p}^*_\infty\right\|_\infty, \tag{35}$$

$$\varepsilon'_t=\max_{k\ge 0,\boldsymbol{p}'\in\mathcal{P}}\|f'^{k+t-t_0}(\boldsymbol{p}',t_0)-\boldsymbol{p}^*_\infty\|_\infty, \tag{36}$$

and

$$\varepsilon_t=\begin{cases}\max\left[\varepsilon'_t,\frac{2(1+\eta)}{1-\eta}\epsilon'_t\right], & \text{if } t<t_0+C,\\ \max\left[\varepsilon'_t,\frac{2(1+\eta)}{1-\eta}\epsilon'_t,\chi(C)\varepsilon_{t-C}\right], & \text{otherwise,}\end{cases} \tag{37}$$

where the function $\chi(C)=\frac{1}{2}(\frac{3+\eta}{1+\eta}\eta^C+1)$, integer constant $C=\lceil\log(\frac{1+\eta}{3+\eta})/\log(\eta)\rceil+1$, and $\lceil\cdot\rceil$ denotes the ceiling function. From (31) and (34), $\{\epsilon'_t\}$ and $\{\varepsilon_t\}$ are infinite decreasing sequences and converge to zero as $t\to\infty$. Construct a new *time sequence* $\{\bar{t}_l\}$ where $\bar{t}_l=t_0+lC$ for any integer $l\ge 0$. Since $\chi(C)<1$, the sequence $\{\varepsilon_t\}$ is decreasing and in particular, we have $\lim_{l\to\infty}\varepsilon_{\bar{t}_l}=0$. For each $l\ge 0$, define $\mathcal{P}_{\bar{t}_l}=\{\boldsymbol{p}:\|\boldsymbol{p}-\boldsymbol{p}^*_\infty\|_\infty\le\varepsilon_{\bar{t}_l}\}$. It is clear that $\boldsymbol{p}^*_\infty\in\mathcal{P}_{\bar{t}_l}$ and $\mathcal{P}_{\bar{t}_{l+1}}\subseteq\mathcal{P}_{\bar{t}_l}$ for all $l\ge 0$. Furthermore, $\mathcal{P}_{\bar{t}_{l+l'}}\subset\mathcal{P}_{\bar{t}_l}$ for some finite $l'$. For any $\boldsymbol{p}\in\mathcal{P}_{\bar{t}_l}$,

$$\left\|\boldsymbol{p}-\boldsymbol{p}^*_{\bar{t}_l}\right\|_\infty \le \|\boldsymbol{p}-\boldsymbol{p}^*_\infty\|_\infty+\left\|\boldsymbol{p}^*_{\bar{t}_l}-\boldsymbol{p}^*_\infty\right\|_\infty \le \varepsilon_{\bar{t}_l}+\epsilon'_t.$$

From (34), we know that:

$$\left\|f'^k(\boldsymbol{p},\bar{t}_l)-\boldsymbol{p}^*_\infty\right\|_\infty < \eta^k\left(\varepsilon_{\bar{t}_l}+\epsilon'_{\bar{t}_l}\right)+\frac{1+\eta}{1-\eta}\epsilon'_{\bar{t}_l}.$$

If $\varepsilon_{\bar{t}_l}=\varepsilon'_{\bar{t}_l}$, then $\epsilon'_{\bar{t}_l}\le\frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2}$. On the other hand, if $\varepsilon_{\bar{t}_l}=\frac{2(1+\eta)}{1-\eta}\epsilon'_{\bar{t}_l}$, or if $\varepsilon_{\bar{t}_l}=\chi(C)\varepsilon_{\bar{t}_l-C}$, then $\varepsilon'_{\bar{t}_l}\le\varepsilon_{\bar{t}_l}$ and $\epsilon'_{\bar{t}_l}\le\frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2}$. Thus, for all three possibilities in (37), we have

$$\begin{aligned}\|f'^C(\boldsymbol{p},\bar{t}_l)-\boldsymbol{p}^*_\infty\|_\infty &< \eta^C\left(\varepsilon_{\bar{t}_l}+\frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2}\right)+\frac{1+\eta}{1-\eta}\frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2} \\ &= \chi(C)\,\varepsilon_{\bar{t}_l}\le\varepsilon_{\bar{t}_{l+1}}.\end{aligned}$$

Thus, for any choice of $\boldsymbol{p}\in\mathcal{P}_{\bar{t}_l}$, the mapping $f'^C(\boldsymbol{p},\bar{t}_l)\in\mathcal{P}_{\bar{t}_{l+1}}$. Since both *synchronous convergence* and *box* conditions hold, Algorithm 1 globally and asynchronously converges to the unique fixed point $\boldsymbol{p}^*_\infty$ [21, pp. 431].

## REFERENCES


[1] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling for maximum throughput in multi-hop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1936–1949, Dec. 1992.

[2] P. Gupta and A. Stolyar, "Optimal throughput allocation in general random-access networks," in *Proc. of CISS*, Princeton, NJ, Mar. 2006.

[3] C. Joo and N. B. Shroff, "Performance of random access scheduling schemes in multi-hop wireless networks," in *Proc. of IEEE INFOCOM*, Anchorage, Alaska, May 2007.

[4] J. Lee, M. Chiang, and R. Calderbank, "Utility-optimal random-access control," *IEEE Trans. on Wireless Communications*, vol. 25, pp. 1135–1147, Aug. 2007.

[5] X. Wang and K. Kar, "Cross-layer rate control for end-to-end proportional fairness in wireless networks with random access," *IEEE J. on Selected Areas in Communications*, vol. 24, pp. 1548–1559, Aug. 2006.

[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] J. Liu, Y. Yi, A. Proutiere, M. Chiang, and H. V. Poor, "Maximizing utility via random access without message passing," *Microsoft Research Technical Report, 2008-128*, 2008.

[8] J. Lee, A. Tang, J. Huang, M. Chiang, and A. Calderbank, "Reverse engineering MAC: A game-theoretic model," *IEEE J. on Selected Areas in Communications*, vol. 6, pp. 2741–2751, Jul. 2007.

[9] A. H. Mohsenian-Rad, J. Huang, M. Chiang, and V. W. S. Wong, "Utility-optimal random access: Reduced complexity, fast convergence, and robust performance," *to appear in IEEE Trans. Wireless Communications*, 2008.

[10] L. Chen, S. H. Low, and J. C. Doyle, "Random access game and medium access control design," Submitted to *IEEE/ACM Trans. on Networking*, Dec. 2006.

[11] F. Calì, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Trans. on Networking*, vol. 8, no. 6, pp. 785–799, 2000.

[12] D. P. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.

[13] F. Kelly, "Charging and rate control for elastic traffic," *European Trans. on Telecommunication*, vol. 8, pp. 33–37, 1997.

[14] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Net.*, vol. 8, pp. 556–567, Oct. 2000.

[15] V. S. Borkar, "Asynchronous Stochastic Approximation," *SIAM J. on Control and Optim.*, vol. 36, pp. 840–851, May 1998.

[16] 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.

[17] C. Bordenave and D. M. A. Proutiere, "Performance of random medium access control an asymptotic approach," in *Proc. of ACM Sigmetrics*, Annapolis, MD, June 2008.

[18] J. Liu, A. Stolyar, M. Chiang, and H. V. Poor, "Queue back-pressure random access in multi-hop wireless networks: Optimality and stability," *IEEE Transactions on Information Theory (submitted)*, May 2008.

[19] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportionally fair rates using local information in aloha networks," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1858 – 1862, Oct. 2004.

[20] A. Proutiere, Y. Yi, and M. Chiang, "Throughput of random access without message passing," in *Proc. of CISS*, Princeton, NJ, Mar. 2008.

[21] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.