Distributed Medium Access Over Time-Varying Channels

Se-Young Yun, *Member, IEEE*, Jinwoo Shin, *Member, IEEE*, and Yung Yi, *Member, IEEE*

***Abstract*—Recent studies on MAC scheduling have shown that carrier sense multiple access (CSMA) algorithms can be throughput optimal for arbitrary wireless network topology. However, these results are highly sensitive to the underlying assumption on 'static' or 'fixed' system conditions. For example, if channel conditions are time-varying, it is unclear how each node can adjust its CSMA parameters, so-called backoff and channel holding times, using its local channel information for the desired high performance. In this paper, we study 'channel-aware' CSMA (A-CSMA) algorithms in time-varying channels, where they adjust their parameters as some function of the current channel capacity. First, we assume that backoff rates can be arbitrary large and show that the achievable rate region of A-CSMA equals to the maximum rate region if and only if the function is exponential. Furthermore, given an exponential function in A-CSMA, we design updating rules for their parameters, which achieve throughput optimality for an arbitrary wireless network topology. They are the first CSMA algorithms in the literature which are proved to be throughput optimal under time-varying channels. Moreover, we also consider the case when back-off rates of A-CSMA are restricted compared to the speed of channel variations, and characterize the throughput performance of A-CSMA in terms of the underlying wireless network topology. Our results not only guide a high-performance design on MAC scheduling under highly time-varying scenarios, but also provide new insights on the performance of CSMA algorithms in relation to their backoff rates and underlying network topologies.**

***Index Terms*—CSMA, time-varying channel, backoff, wireless ad-hoc network.**

I. Introduction

A. Motivation

HOW TO access the shared medium is a crucial issue in achieving high performance in many applications, *e.g.*, wireless networks. In spite of a surge of research papers in this area, it's the year 1992 that the seminal work by Tassiulas and Ephremides proposed a throughput optimal medium access algorithm, referred to as Max-Weight [29]. Since then, a huge array of subsequent research has been made to develop distributed medium access algorithms with high performance guarantee and low complexity. However, in many cases the tradeoff between complexity and achievable rate region has been observed, or even throughput optimal algorithms with polynomial complexity have turned out to require heavy message passing, which becomes a major hurdle to becoming practical medium access schemes, *e.g.*, see [7], [30] for surveys.

Recently, there has been exciting progress that even fully distributed medium access algorithms based on CSMA (Carrier Sense Multiple Access) with no or very little message passing can achieve optimality in both throughput and utility, *e.g.*, see [6], [17], [21], [24]. The main intuition underlying these results is that nodes dynamically adjust their CSMA parameters, *backoff* and *channel holding* times, using local information such as queue-length so that they solve a certain network-wide optimization problem for the desired high performance. We refer the readers to a survey paper [32] for more details.

However, the recent CSMA algorithms assume static channel conditions, and it is far from being clear how they perform for time-varying channels, which frequently occurs in practice. Note that it has already been shown that the Max-Weight is throughput optimal for time-varying channels [28] and joint scheduling and congestion control algorithms based on the optimization decomposition, *e.g.*, [2], are utility optimal by selecting the schedules over time, both of which essentially track the channel conditions quickly. However, a similar channel adaptation for CSMA algorithms may not be feasible for the following two reasons. First, each node in a network only knows its local channel information, and cannot track channel conditions of other nodes. Second, there exists a non-trivial coupling between CSMA's performance under time-varying channels and the speed of channel variations, since CSMA algorithms might not react quickly to the channel variations.

In this paper, we formalize and quantify this coupling, and study when and how CSMA algorithms perform depending on the network topologies and the speed of channel variations.

B. Our Contribution

In this paper, we model time-varying channels by a Markov process, and study 'channel-aware' CSMA (A-CSMA) algorithms where each link adjusts its CSMA parameters, backoff and channel holding times, as some function of its (local) channel capacity. We compare A-CSMA with 'channel-unaware' CSMA (U-CSMA) algorithms where backoff and channel holding times do not depend on the channel capacity

Manuscript received November 20, 2013; revised May 31, 2014; October 05, 2014; April 13, 2015; and September 23, 2015; accepted October 25, 2015; approved by IEEE/ACM Transactions on Networking Editor X. Lin. This work was supported by BK21 plus, the Basic Science Research Program through the National Research Foundation of Korea (NRF) under Grant NRF-2013R1A2A2A01067633 and the Ministry of Science, ICT and Future Planning under the Center for Integrated Smart Sensors as Global Frontier Project CISS-2012M3A6A6054195. This work was presented in part at ACM MobiHoc 2013.

S. Yun is with the MSR-INRIA Joint Center, Palaiseau 91120, France (e-mail: seyoung.yun@inria.fr).

J. Shin and Y. Yi are with the Department of Electrical Engineering, KAIST, Daejeon 305-701, Korea (e-mail: jinwoos@kaist.ac.kr, yiyung@kaist.edu).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2015.2503394

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

(see Section II-C for the formal definitions of A-CSMA and U-CSMA). In what follows, we first summarize our main contributions and then describe more details.

C1—*Achievable rate region of A-CSMA*. We show that the achievable rate region of A-CSMA is maximized if and only if the function for backoff and channel holding times is exponential. In particular, we prove that A-CSMA can achieve an arbitrary large fraction of the capacity region for exponential functions (see Theorem 3.1) with arbitrary large backoff rates, which turns out to be *impossible* for non-exponential functions (see Theorem 3.2).

C2—*Dynamic throughput optimal A-CSMA*. We develop two types of throughput optimal A-CSMA algorithms, where links dynamically update their CSMA parameters based on both (a) the exponential function of the channel capacity in ***C1*** and (b) the empirical local load or the local queue length, without knowledge of the speed of channel variation and the arrival statistics (such as its mean) in advance (see Theorems 4.1 and 4.2).

C3—*Achievable rate region of A-CSMA with limited backoff rates.* We provide a lower bound for the achievable rate region of A-CSMA when their backoff rates are highly limited compared to the speed of channel variations (see Theorem 5.3). Our bound depends on a combinatorial property of the underlying interference graph, *i.e.*, its chromatic number χ,[1] and is independent of backoff rates or the speed of channel variations. Moreover, it is noteworthy that the achievable rate region of A-CSMA includes the achievable rate region of channel-unaware CSMA (U-CSMA) for any limited backoff rate (see Theorem 5.2).

A typical necessary step to analyze and design a CSMA algorithm of high performance in static channels is to characterize the stationary distribution of the Markov chain of schedules induced by it [6], [17], [21], [24]. However, this task is much harder for A-CSMA in time-varying channels, since the Markov chain induced by A-CSMA is *non-reversible* (see Theorem 2.1), *i.e.*, it is unlikely that its stationary distribution has a 'clean' formula to analyze, being in sharp contrast to the CSMA analysis for static channels. To overcome this technical issue, we first show that the stationary distribution approximates to a product-form distribution when backoff rates are sufficiently large. Then, for ***C1***, we study the product-form to guarantee high throughput of A-CSMA, where the exponential functions are found. The main novelty lies in establishing the approximation scheme, using the *Markov chain tree theorem* [1], which requires counting the weights of arborescences induced by the non-reversible Markov process to understand its stationary distribution.

For ***C2***, we combine ***C1*** with existing techniques: our first and second throughput optimal algorithms are 'rate-based' and 'queue-based' ones originally studied in static channels by Jiang *et al.* (cf. [6], [5]) and Rajagopalan *et al.* (cf. [24], [26]), respectively. To extend these results to time-varying channels, our specific choice of holding times as exponential functions of the channel capacity plays a key role in establishing the desired throughput optimal performance. To our best knowledge, they are the first CSMA algorithms in the literature which are proved to be throughput optimal under general Markovian time-varying channel models.

C3 is motivated by observing that a CSMA algorithm in fast time-varying channels inevitably has to be of high backoff rates for the desired throughput performance, *i.e.*, high backoff rates are needed for tracking time-varying channel conditions fast enough. However, backoff rates are bounded in practice, which may cause degradation in the CSMA's performance. We note that CSMA algorithms with limited backoff or holding rates have been little analyzed in the literature, despite of their practical importance.[2] ***C3*** provides a lower bound for A-CSMA throughputs regardless of restrictions on their backoff rates or sensing frequencies. For example, if the interference graph is bipartite (*i.e.*, its chromatic number is two), our bound implies that A-CSMA is guaranteed to have at least 50%—throughput even with arbitrary small backoff rates.

C. *Related Work*

The research on throughput optimal CSMA has been initiated independently by Jiang *et al.* (cf. [5], [6]) and Rajagopalan *et al.* (cf. [24], [26]), where both consider the continuous time and collision free setting. Under exponential distributions on backoff and holding times, the system is modeled by a continuous time Markov chain, where the backoff rate or channel holding time at each link is adaptively controlled by the local (virtual or actual) queue lengths. Jiang *et al.* proved that the long-term link throughputs are the solution of an utility maximization problem assuming the infinite backlogged data. Rajagopalan *et al.* [24] showed that if the CSMA parameters are changing very slowly with respect to the queue length changes, the realized link schedules provably emulate Max-Weight very well. Although their key intuitions are apparently different, both require to understand the long-term behavior (*i.e.*, stationarity) of the Markov chains on schedules formed by CSMA.

These throughput optimality results motivate further research on design and analysis of CSMA algorithms. The work by Liu *et al.* [17] proves the utility optimality using a stochastic approximation technique, which has been extended to the multi-channel, multi-radio case with a simpler proof in [22]. The throughput optimality of MIMO networks under SINR model is also shown in [23]. As opposed to the continuous-time setting where carrier sensing is perfect and instantaneous (and hence no collision occurs), more practical discrete time settings that carrier sensing is imperfect or delayed (and hence collisions occur) have been also studied. The throughput optimality of CSMA algorithms in discrete time settings with collisions is established in [8], [27] and [10], where the authors in [10] consider imperfect sensing information. In [17], the authors studied the impact of collisions and the tradeoff between short-term fairness and efficiency. The authors in [21] considered a synchronous system consisting of the control phase, which eliminates the chances of data collisions via a simple

[1]Given a graph, its chromatic number χ is the smallest number of colors needed to color the vertices so that no two adjacent vertices share the same color.

[2]Even in static channels, restrictions on backoff or holding rates may degrade the throughput or delay performances of CSMA algorithms.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

message passing, and the data phase, which actually enables data transmission based on the discrete-time Glauber dynamics. There also exist several efforts on improving or analyzing delay performance [3], [4], [11], [15], [18], [25], speeding up the convergence [31], and developing a practical protocol based on the CSMA theory with experimental validation [12], [13], [19]. To the best of our knowledge, CSMA under time-varying channels has been studied only in [16] for only complete interference graphs, when the arbitrary backoff rate is allowed. Besides CSMA, distributed scheduling under time-varying channel was recently studied in [9]: the authors provide an algorithm with $1/\Delta$ throughput guarantee where Δ is the maximum degree of the interference graph. Our proposed A-CSMA algorithm provides the best throughput guarantee in time-varying channels among low complexity distributed scheduling algorithms in the literature: it achieves $1/\chi$ throughput even with limited backoff rates (see Theorem 5.3) and $\chi \le \Delta$ due to the Brook's theorem.

II. Model and Preliminaries

A. Network Model

We consider a network consisting of a collection of n links $\{1,\dots,n\}$ where each link has a queue and time is indexed by $t \in \mathbb{R}_+$. Let $Q_i(t) \in \mathbb{R}_+$ denote the amount of work in the queue of link i at time t and let $\boldsymbol{Q}(t) = [Q_i(t)]_{1\le i\le n}$. The system starts empty, *i.e.*, $Q_i(0) = 0$. The arrival process is assumed to be discrete-time with unit-sized packets arriving to queues as in [24], for convenience. Let $A_i(s,t) < \infty$ denotes the cumulative arrival to link i in the time interval $(s,t]$. For simplicity, we assume $A_i(\tau,\tau+1)$ are independent Bernoulli processes with parameter λ_i for all $\tau \in \mathbb{N}_0$. Each link i can be serviced at rate $c_i(t) \ge 0$ representing the potential departure rate of work from the queue $Q_i(t)$. We consider m discrete channel states such that $c_i(t) \in \mathcal{H} := \{h_1,\dots,h_m\}$ and $0 \le h_1 < \cdots < h_m = c_{\max}$ for all link i and time $t \ge 0$. We consider finite state Markov time-varying channels: $\{\boldsymbol{c}(t) = [c_i(t)]_{1\le i\le n} : t \ge 0\}$ is a continuous-time, time-homogeneous, and irreducible Markov process. We denote by $\gamma^{\boldsymbol{u}\to\boldsymbol{v}}$ the 'transition-rates' on the channel state for $\boldsymbol{u} \to \boldsymbol{v}, \boldsymbol{u},\boldsymbol{v} \in \mathcal{H}^n$. For the time-varying channels, we assume that each link i knows the channel state $c_i(t)$ before it transmits.[3] We call $\max_{\boldsymbol{u}\in\mathcal{H}^n}\{\sum_{\boldsymbol{v}\in\mathcal{H}^n:\boldsymbol{v}\ne\boldsymbol{u}}\gamma^{\boldsymbol{u}\to\boldsymbol{v}}\}$ the *channel varying speed*. The *channel varying speed* indicates the maximum of the expected number of channel transitions during the unit-length time interval. We let $[\pi_{\boldsymbol{c}}]$ denote the stationary distribution of Markov process $\{\boldsymbol{c}(t)\}$. We consider only single-hop sessions (or flows), *i.e.*, once work departs from a queue, it leaves the network.

The links are offered service as per the constraint imposed by interference. To model this constraint, we adopt a popular graph-based approach, where denote by $G = (V,E)$ the inference graph among n links, where the vertices $V = \{1,\dots,n\}$ represent links and the edges $E \subset V \times V$ represent interferences between links: $(i,j) \in E$, if links i and j interfere with each other. Let $\mathcal{N}(i) = \{j \in V : (i,j) \in E\}$ and $\boldsymbol{\sigma}(t) = [\sigma_i(t)] \in \{0,1\}^n$ denote the neighbors of link i and a schedule at time t, *i.e.*, whether links transmit at time t, respectively, where $\sigma_i(t) = 1$ represents transmission of link i at time t. Then, interference imposes the constraint that for all $t \in \mathbb{R}_+, \boldsymbol{\sigma}(t) \in \mathcal{I}(G)$, where

$$\mathcal{I}(G) := \{\boldsymbol{\rho} = [\rho_i] \in \{0,1\}^n : \rho_i + \rho_j \le 1, \forall (i,j) \in E\}.$$

The resulting queueing dynamics are described as follows. For $0 \le s < t$ and $1 \le i \le n$,

$$Q_i(t) = Q_i(s) - \int_s^t \sigma_i(r)c_i(r)\mathbf{1}_{\{Q_i(r)>0\}}dr + A_i(s,t),$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function. Finally, we define the cumulative actual and potential departure processes $\boldsymbol{D}(t) = [D_i(t)]$ and $\widehat{\boldsymbol{D}}(t) = [\widehat{D}_i(t)]$, respectively, where

$$D_i(t) = \int_0^t \sigma_i(r)c_i(r)\mathbf{1}_{\{Q_i(r)>0\}}dr,\ \widehat{D}_i(t) = \int_0^t \sigma_i(r)c_i(r)dr.$$

B. Scheduling, Rate Region and Metric

The main interest of this paper is to design a scheduling algorithm which decides $\boldsymbol{\sigma}(t) \in \mathcal{I}(G)$ for each time instance $t \in \mathbb{R}_+$. Intuitively, it is expected that a good scheduling algorithm will keep the queues as small as possible. To formally discuss, we define the maximum achievable rate region (also called capacity region) $\boldsymbol{C} \subset [0,1]^n$ of the network, which is the convex hull of the feasible scheduling set $\mathcal{I}(G)$, *i.e.*,

$$\boldsymbol{C} = \boldsymbol{C}(\boldsymbol{\gamma},G) = \left\{ \sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\sum_{\boldsymbol{\rho}\in\mathcal{I}(G)}\alpha_{\boldsymbol{\rho},\boldsymbol{c}}\boldsymbol{c}^T\cdot\boldsymbol{\rho} : \alpha_{\boldsymbol{\rho},\boldsymbol{c}} \ge 0 \right.$$
$$\left. \text{and} \sum_{\boldsymbol{\rho}\in\mathcal{I}(G)}\alpha_{\boldsymbol{\rho},\boldsymbol{c}} = 1 \text{ for all } \boldsymbol{c}\in\mathcal{H}^n \right\},$$

where $\boldsymbol{c}^T\cdot\boldsymbol{\rho} = [c_i\rho_i]$ and $\pi_{\boldsymbol{c}}$ denotes the stationary distribution of channel state $\boldsymbol{c}$ under the channel-varying Markov process. The intuition behind this definition comes from the facts: (a) any scheduling algorithm has to choose a schedule from $\mathcal{I}(G)$ at each time and channel state where $\alpha_{\boldsymbol{\rho},\boldsymbol{c}}$ denotes the fraction of time selecting schedule $\boldsymbol{\rho}$ for given channel state $\boldsymbol{c}$ and (b) for channel state $\boldsymbol{c} \in \mathcal{H}^n$, the fraction in the time domain where $\boldsymbol{c}(t) = [c_i(t)]$ is equal to $\boldsymbol{c}$ is $\pi_{\boldsymbol{c}}$. Hence the time average of the 'service rate' induced by any algorithm must belong to $\boldsymbol{C}$.

We call the arrival rate $\boldsymbol{\lambda}$ *admissible* if there is $\boldsymbol{\lambda}' \in \boldsymbol{C}(\boldsymbol{\gamma},G)$ such that $\boldsymbol{\lambda} \le \boldsymbol{\lambda}'$ corresponds to the component-wise inequality. Thus, the *admissible* arrival rate region is defined as $\boldsymbol{\lambda} = [\lambda_i] \in \boldsymbol{\Lambda} = \boldsymbol{\Lambda}(\boldsymbol{\gamma},G)$, where

$$\boldsymbol{\Lambda}(\boldsymbol{\gamma},G) := \left\{\boldsymbol{\lambda} \in \mathbb{R}_+^n : \boldsymbol{\lambda} \le \boldsymbol{\lambda}',\ \text{for some } \boldsymbol{\lambda}' \in \boldsymbol{C}(\boldsymbol{\gamma},G)\right\}.$$

If $\boldsymbol{\lambda} \notin \boldsymbol{\Lambda}$, queues cannot be stabilized under any scheduling algorithm. Further, $\boldsymbol{\lambda}$ is called *strictly admissible* if $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o = \boldsymbol{\Lambda}^o(\boldsymbol{\gamma},G)$ and

$$\boldsymbol{\Lambda}^o(\boldsymbol{\gamma},G) := \left\{\boldsymbol{\lambda} \in \mathbb{R}_+^n : \boldsymbol{\lambda} < \boldsymbol{\lambda}',\ \text{for some } \boldsymbol{\lambda}' \in \boldsymbol{C}(\boldsymbol{\gamma},G)\right\}.$$

We now define the performance metric.

[3]The channel information can be achieved using control messages such as RTS and CTS in IEEE 802.11, and links can adapt their transmission parameters to channel transitions for every transmission by changing coding and modulation parameters.

Definition 2.1: A scheduling algorithm is called rate-stable for a given arrival rate $\boldsymbol{\lambda}$, if

$$\lim_{t\to\infty} \frac{1}{t}\boldsymbol{D}(t) = \boldsymbol{\lambda} \text{ (with probability 1)}. \tag{1}$$

Furthermore, we say a scheduling algorithm has α-throughput if it is rate-stable for any $\boldsymbol{\lambda} \in \alpha\boldsymbol{\Lambda}^o(\boldsymbol{\gamma}, G)$. In particular, when $\alpha = 1$, it is called throughput optimal.

We note that (1) is equivalent to $\lim_{t\to\infty} \frac{1}{t}\boldsymbol{Q}(t) = 0$, since $\lim_{t\to\infty} \frac{A_i(0,t)}{t} = \lambda_i$ (because the arrival process is stationary ergodic). The following lemma implies that the potential departure process suffices to study the rate-stability [20].

Lemma 2.1: A scheduling algorithm is rate-stable if

$$\lim_{t\to\infty} \frac{1}{t}\widehat{\boldsymbol{D}}(t) > \boldsymbol{\lambda}.$$

C. Channel-Aware CSMA Algorithm: A-CSMA

The algorithm to decide $\boldsymbol{\sigma}(t)$ utilizing the local carrier sensing information can be classified as CSMA (Carrier Sense Multiple Access) algorithms. In between two transmissions, a link waits for a random amount of time—also known as *backoff time*. Each link can sense the medium perfectly and instantly, *i.e.*, knows if any other interfering link is transmitting at a given time instance. If a link that finishes waiting senses the medium to be busy, it starts waiting for another random amount of time; else, it starts transmitting for a random amount of time, called *channel holding time*. We assume that link i's backoff and channel holding times have exponential distributions with mean $1/R_i$ and $1/S_i$, respectively, where $R_i = R_i(t) > 0$ and $S_i = S_i(t) > 0$ may change over time. We define A-CSMA (channel-aware CSMA) to be the class of CSMA algorithms where $R_i(t)$ and $S_i(t)$ are decided by some functions of the current channel capacity, *i.e.*, $R_i(t) = f_i(c_i(t))$ and $S_i(t) = g_i(c_i(t))$ for some functions f_i and g_i. In the special case when $R_i(t)$ and $S_i(t)$ are decided independently of current channel information (*e.g.*, f_i's and g_i's are constant functions), we specially say a CSMA algorithm is U-CSMA (channel-unaware CSMA).

Then, given functions $[f_i]$ and $[g_i]$, it is easy to check that $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t)) : t \geq 0\}$ under A-CSMA is a continuous time Markov process, whose kernel (or transition-rates) is given by:

$$\begin{aligned}
&(\boldsymbol{\sigma}, \boldsymbol{u}) \to (\boldsymbol{\sigma}, \boldsymbol{v}) \text{ with rate } \gamma^{\boldsymbol{u}\to\boldsymbol{v}} \\
&(\boldsymbol{\sigma}_i^0, \boldsymbol{c}) \to (\boldsymbol{\sigma}_i^1, \boldsymbol{c}) \text{ with rate } f_i(c_i) \cdot \prod_{j:(i,j)\in E} (1 - \sigma_j) \\
&(\boldsymbol{\sigma}_i^1, \boldsymbol{c}) \to (\boldsymbol{\sigma}_i^0, \boldsymbol{c}) \text{ with rate } g_i(c_i) \cdot \sigma_i,
\end{aligned} \tag{2}$$

where $\boldsymbol{\sigma}_i^0$ and $\boldsymbol{\sigma}_i^1$ denote two 'almost' identical schedule vectors except i-th elements which are 0 and 1, respectively. Since $\{\boldsymbol{c}(t)\}$ is a time-homogeneous irreducible Markov process, $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$ is ergodic, *i.e.*, it has the unique stationary distribution $[\pi_{\boldsymbol{\sigma},\boldsymbol{c}}]$. For example, when functions f_i and g_i are constant (*i.e.*, U-CSMA with fixed $R_i(t) = R_i$ and $S_i(t) = S_i$),

$$\pi_{\boldsymbol{\sigma},\boldsymbol{c}} = \pi_{\boldsymbol{c}} \cdot \frac{\exp\left(\sum_i \sigma_i \log \frac{R_i}{S_i}\right)}{\sum_{\boldsymbol{\rho}=[\rho_i]\in\mathcal{I}(G)} \exp\left(\sum_i \rho_i \log \frac{R_i}{S_i}\right)},$$

and if $\{\boldsymbol{c}(t)\}$ is (time-)reversible, $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$ is as well. In general, $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$ is not reversible unless functions f_i/g_i are constant, as we state in the following theorem. The non-reversible property, unfortunately, makes it hard to characterize the stationary distribution $[\pi_{\boldsymbol{\sigma},\boldsymbol{c}}]$ of the Markov process induced by A-CSMA.

Theorem 2.1: If $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$ is reversible,

$$\frac{f_i(x)}{g_i(x)} = \frac{f_i(y)}{g_i(y)}, \quad \text{for all } x, y \in \mathcal{H}, i \in V.$$

Proof: Denote by $\boldsymbol{c}_i^u$ and $\boldsymbol{c}_i^v$ two almost identical channel state vectors except i-th elements, which are h_u and h_v, respectively. Suppose that $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$ is reversible. From the reversibility, the transition path $(\boldsymbol{\sigma}_i^0, \boldsymbol{c}_i^u) \to (\boldsymbol{\sigma}_i^0, \boldsymbol{c}_i^v) \to (\boldsymbol{\sigma}_i^1, \boldsymbol{c}_i^v)$ has to satisfy the following balance equations:

$$\begin{aligned}
\pi_{\boldsymbol{\sigma}_i^0,\boldsymbol{c}_i^u} \gamma^{\boldsymbol{c}_i^u\to\boldsymbol{c}_i^v} &= \pi_{\boldsymbol{\sigma}_i^0,\boldsymbol{c}_i^v} \gamma^{\boldsymbol{c}_i^v\to\boldsymbol{c}_i^u} \\
\pi_{\boldsymbol{\sigma}_i^0,\boldsymbol{c}_i^v} f_i(h_v) &= \pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^v} g_i(h_v),
\end{aligned} \tag{3}$$

Similarly, for the transition path $(\boldsymbol{\sigma}_i^0, \boldsymbol{c}_i^u) \to (\boldsymbol{\sigma}_i^1, \boldsymbol{c}_i^u) \to (\boldsymbol{\sigma}_i^1, \boldsymbol{c}_i^v)$,

$$\pi_{\boldsymbol{\sigma}_i^0,\boldsymbol{c}_i^u} f_i(h_u) = \pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^u} g_i(h_u),$$

and

$$\pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^u} \gamma^{\boldsymbol{c}_i^u\to\boldsymbol{c}_i^v} = \pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^v} \gamma^{\boldsymbol{c}_i^v\to\boldsymbol{c}_i^u}. \tag{4}$$

From (3) and (4),

$$\frac{\pi_{\boldsymbol{\sigma}_i^0,\boldsymbol{c}_i^u}}{\pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^v}} = \frac{\gamma^{\boldsymbol{c}_i^v\to\boldsymbol{c}_i^u} g_i(h_v)}{\gamma^{\boldsymbol{c}_i^u\to\boldsymbol{c}_i^v} f_i(h_v)} = \frac{\gamma^{\boldsymbol{c}_i^v\to\boldsymbol{c}_i^u} g_i(h_u)}{\gamma^{\boldsymbol{c}_i^u\to\boldsymbol{c}_i^v} f_i(h_u)}, \tag{5}$$

which indicates that $\frac{f_i(h_u)}{g_i(h_u)} = \frac{f_i(h_v)}{g_i(h_v)}$. This completes the proof of Theorem 2.1. ∎

III. Achievable Rate Region of A-CSMA

In this section, we study the achievable rate region of A-CSMA algorithms given (fixed) functions $[f_i]$ and $[g_i]$. We show that the achievable rate region of A-CSMA is maximized for the following choices of functions:

$$\log \frac{f_i(x)}{g_i(x)} = r_i \cdot x, \quad \text{for } x \in [0, c_{\max}], \tag{6}$$

where $r_i \in \mathbb{R}$ is some constant. Namely, the ratio $f_i(x)/g_i(x)$ is an exponential function in terms of x. We let EXP-A-CSMA denote the sub-class of A-CSMA algorithms with functions satisfying (6) for some $[r_i]$. The following theorem justifies the optimality of EXP-A-CSMA in terms of its achievable rate region.

Theorem 3.1 (Optimality): For any interference graph G, and channel transition-rate $\boldsymbol{\gamma}$, and arrival rate $\boldsymbol{\lambda} = [\lambda_i] \in \boldsymbol{\Lambda}^o(\gamma, G)$, there exists $[r_i]$, $[f_i]$ and $[g_i]$ satisfying (6) such that the corresponding EXP-A-CSMA algorithm is rate-stable.

Proof: The proof is given in Section III-A. ∎

We also establish that Theorem 3.1 is tight in the sense that it does not hold for other A-CSMA algorithms that have different ways of reflecting channel capacity in adjusting CSMA parameters. To state it formally, given a non-negative monotonically increasing continuous function $k : [0, c_{\max}] \to \mathbb{R}_+$, we define EXP($k$)-A-CSMA as the sub-class of A-CSMA algorithms with the following form of functions:

$$\log \frac{f_i(x)}{g_i(x)} = r_i \cdot k(x), \quad \text{for } x \in [0, c_{\max}], \tag{7}$$

where $r_i \in \mathbb{R}$ is some constant. The following theorem states that EXP-A-CSMA is the unique class of A-CSMA maximizing its achievable rate region.

Theorem 3.2 (Uniqueness): If the conclusion of Theorem 3.1 holds for EXP(k)-A-CSMA, then $k(x)$ is a linear function.

Proof: The proof is given in Sections III-D. ∎

In the following proofs (and throughout this paper), we commonly use $[\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}]$ to denote the stationary distribution of Markov process $\{\boldsymbol{\sigma}(t), \boldsymbol{c}\}$ (with fixed channel state $\boldsymbol{c}$) induced by an A-CSMA algorithm. It is noteworthy that with fixed channel state EXP-A-CSMA algorithms are the same with U-CSMA and

$$\pi_{\boldsymbol{\sigma}|\boldsymbol{c}} = \frac{\exp\left(\sum_i \sigma_i r_i c_i\right)}{\sum_{\boldsymbol{\rho}=[\rho_i]\in\mathcal{I}(G)} \exp\left(\sum_i \rho_i r_i c_i\right)}.$$

A. Proof of Theorem 3.1

To begin with, we recall that the channel varying speed ψ is defined as: $\psi = \max_{\boldsymbol{u}\in\mathcal{H}^n}\{\sum_{\boldsymbol{v}\in\mathcal{H}^n:\boldsymbol{v}\neq\boldsymbol{u}} \gamma^{\boldsymbol{u}\to\boldsymbol{v}}\}$. We first state Lemmas 3.1 and 3.2, which are the key lemmas to the proof of Theorem 3.1.

Lemma 3.1: For any $\delta_1 \in (0,1)$, arrival rate $\boldsymbol{\lambda} = [\lambda_i] \in (1-\delta_1)\boldsymbol{\Lambda}^o$, interference graph G and channel transition-rate $\boldsymbol{\gamma}$, there exists $[r_i] \in \mathbb{R}^n$ such that

$$\max_i |r_i| \le \frac{4n^2 c_{\max} \log|\mathcal{I}(G)|}{\delta_1^2 \min_i\{\left(\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i \pi_{\boldsymbol{c}}\right)^2\}},$$

and every EXP-A-CSMA algorithm with

$$\log\frac{f_i(h)}{g_i(h)} = r_i \cdot h, \quad \text{for all } i \in V, h \in \mathcal{H}$$

satisfies

$$\lambda_i \le \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1} \pi_{\boldsymbol{\sigma}|\boldsymbol{c}}, \text{ for all } i \in V.$$

Proof: This proof is presented in Section III-B. ∎

Lemma 3.2: For any $\delta_2 \in (0,1)$, interference graph G and channel transition-rate $\boldsymbol{\gamma}$ and EXP-A-CSMA algorithm with functions $\boldsymbol{f} = [f_i]$ and $\boldsymbol{g} = [g_i]$ satisfying (7) and

$$\min_{i\in V, h\in\mathcal{H}}\{f_i(h), g_i(h)\} \ge \frac{\psi m^{2^n m^n (n+1)} \exp(n c_{\max} \max_i |r_i|)}{\delta_2},$$

it follows that

$$\max_{(\boldsymbol{\sigma},\boldsymbol{c})\in\mathcal{I}(G)\times\mathcal{H}^n} \left|1 - \frac{\pi_{\boldsymbol{\sigma},\boldsymbol{c}}}{\pi_{\boldsymbol{c}}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}}\right| < \delta_2.$$

Proof: This proof is presented in Section III-C. ∎

Lemma 3.2 implies that if f_i, g_i are large enough, the stationary distribution $[\pi_{\boldsymbol{\sigma},\boldsymbol{c}}]$ approximates to a product-form distribution $[\pi_{\boldsymbol{c}}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}]$, where under EXP-A-CSMA,

$$\pi_{\boldsymbol{\sigma}|\boldsymbol{c}} \propto \exp\left(\sum_i \sigma_i r_i c_i\right),$$

due to the reversibility of Markov process $\{\boldsymbol{\sigma}(t), \boldsymbol{c}\}$. On the other hand, Lemma 3.1 implies that arrival rate $\boldsymbol{\lambda}$ is stabilized under the distribution $[\pi_{\boldsymbol{c}}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}]$. Therefore, combining two above lemmas will lead to the proof of Theorem 3.1.

Proof of Theorem 3.1: We now complete the proof of Theorem 3.1 using Lemmas 3.1 and 3.2. Consider a given arrival rate $\boldsymbol{\lambda} \in (1-\varepsilon)\boldsymbol{\Lambda}^o$ with $\varepsilon \in (0,1)$. If we apply Lemmas 3.1 and 3.2 with $(1+\varepsilon)\boldsymbol{\lambda} \in (1-\varepsilon^2)\boldsymbol{\Lambda}^o$ (*i.e.*, $\delta_1 = \varepsilon^2$ and $\delta_2 = \frac{\varepsilon}{1+\varepsilon}$), we have that there exists an EXP-A-CSMA algorithm with constant $[r_i]$ and functions $[f_i]$ and $[g_i]$ such that

$$\eta \exp(n c_{\max} \max_i |r_i|) \le \min_{i\in V, h\in\mathcal{H}}\{f_i(h), g_i(h)\}$$
$$(1+\varepsilon)\lambda_i \le \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1} \pi_{\boldsymbol{\sigma}|\boldsymbol{c}},$$

where we choose

$$f_i(x) = R = \eta\exp(\kappa), \quad g_i(x) = R\cdot\exp(-r_i\cdot x), \tag{8}$$

$$\kappa = \kappa(\delta_1, G, \boldsymbol{\gamma}) := \frac{8n^3 c_{\max}^3 \log|\mathcal{I}(G)|}{\delta_1^2 \min_i\{\left(\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right)^2\}}, \tag{9}$$

and

$$\eta = \eta(\delta_2, G, \boldsymbol{\gamma}) := \frac{\psi \cdot m^{2^n m^n (n+1)}}{\delta_2}. \tag{10}$$

Therefore, it follows that

$$\begin{aligned}\lambda_i &\le \left(1 - \frac{\varepsilon}{1+\varepsilon}\right) \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1} \pi_{\boldsymbol{\sigma}|\boldsymbol{c}} \\ &< \sum_{\boldsymbol{c}\in\mathcal{H}^n} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1} c_i \pi_{\boldsymbol{\sigma},\boldsymbol{c}} \\ &= \lim_{t\to\infty} \frac{1}{t}\widehat{D}_i(t),\end{aligned}$$

where the last equality is from the ergodicity of Markov process $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$. This leads to the rate-stability using Lemma 2.1, and hence completes the proof. Since $\liminf_{\varepsilon\to 0}(1-\varepsilon)\Lambda^o = \Lambda^o$, we can stabilize any arrival rate $\boldsymbol{\lambda} \in \Lambda^o$ with some EXP-A-CSMA algorithms.

B. Proof of Lemma 3.1

We remark that Lemma 3.1 is a non-trivial generalization of Lemma 8 in [5] (for static channels), which corresponds to a special case of Lemma 3.1 with $\pi_{\boldsymbol{c}} = 1$ for $\boldsymbol{c} = [1]$. The proof of Lemma 3.1 uses a similar strategy with that of Lemma 8 in [5].

Since $\boldsymbol{\lambda} \in (1-\delta_1)\boldsymbol{\Lambda}^o$, there exists $\boldsymbol{\lambda}^{(a)} = [\lambda_i^{(a)}] \in (1-\delta_1)\boldsymbol{C}$ such that $\boldsymbol{\lambda} \le \boldsymbol{\lambda}^{(a)}$ and there exists $\boldsymbol{\lambda}^{(b)} = [\lambda_i^{(b)}] \in \delta_1/2\boldsymbol{C}$ such that $\lambda_i^{(b)} \ge \frac{\delta_1}{2n}\cdot\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}$ for all $i \in V$. Let $\boldsymbol{\lambda}' = \boldsymbol{\lambda}^{(a)} + \boldsymbol{\lambda}^{(b)}$. Then, $\boldsymbol{\lambda}' \in (1-\delta_1/2)\boldsymbol{C}, \boldsymbol{\lambda} \le \boldsymbol{\lambda}'$, and

$$\lambda_i' \ge \frac{\delta_1}{2n}\cdot\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}} \quad \text{for all } i \in V.$$

For such a choice of $\boldsymbol{\lambda}'$, we consider the following function $F: \mathbb{R}^n \to \mathbb{R}$:

$$F(\boldsymbol{r}) = \boldsymbol{\lambda}'\cdot\boldsymbol{r} - \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \log\left(\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \exp\left(\sum_i \sigma_i c_i r_i\right)\right).$$

One can easily check that F is strictly concave and bounded above. Hence, there exists a unique maximizer $\boldsymbol{r}^* \in \mathbb{R}^n$ such that $F(\boldsymbol{r}^*) = \sup_{\boldsymbol{r}\in\mathbb{R}^n} F(\boldsymbol{r})$ and $\nabla F(\boldsymbol{r}^*) = 0$. We prove the following.

$$\max_i r_i^* \le \frac{2\log|\mathcal{I}(G)|}{\delta_1 \min_i\left\{\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right\}} \tag{11}$$

$$\min_i r_i^* \ge -\frac{4n^2 c_{\max}\log|\mathcal{I}(G)|}{\delta_1^2 \min_i\left\{\left(\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right)^2\right\}} \tag{12}$$

Proof of (11): We introduce $[\widehat{\pi}_{\boldsymbol{\sigma}|\boldsymbol{c}}]$ such that $\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \widehat{\pi}_{\boldsymbol{\sigma}|\boldsymbol{c}} = 1$,

$$\boldsymbol{\lambda}' = \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \widehat{\pi}_{\boldsymbol{\sigma}|\boldsymbol{c}}[c_i\sigma_i]_{1\le i\le n}, \quad \text{and} \quad \widehat{\pi}_{\boldsymbol{0}|\boldsymbol{c}} \ge \frac{\delta_1}{2}. \tag{13}$$

There exists such $[\widehat{\pi}_{\boldsymbol{\sigma}|\boldsymbol{c}}]$. From the definition of $\boldsymbol{C}$ there exists $[\alpha_{\boldsymbol{\sigma}|\boldsymbol{c}}]$ satisfying that $\frac{1}{1-\delta_1/2}\boldsymbol{\lambda}' = \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \alpha_{\boldsymbol{\sigma}|\boldsymbol{c}}[c_i\sigma_i]_{1\le i\le n}$ and $\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \alpha_{\boldsymbol{\sigma}|\boldsymbol{c}} = 1$ for all $\boldsymbol{c} \in \mathcal{H}^n$ since $\frac{1}{1-\delta_1/2}\boldsymbol{\lambda}' \in \boldsymbol{C}$. When we let $[\beta_{\boldsymbol{\sigma}|\boldsymbol{c}}]$ be a trivial vector such that for all $\boldsymbol{c} \in \mathcal{H}^n \beta_{\boldsymbol{0}|\boldsymbol{c}} = 1$ and $\beta_{\boldsymbol{\sigma}|\boldsymbol{c}} = 0$ for all $\boldsymbol{\sigma} \ne \boldsymbol{0}, [\widehat{\pi}_{\boldsymbol{\sigma}|\boldsymbol{c}}] = (1 - \frac{\delta_1}{2})[\alpha_{\boldsymbol{\sigma}|\boldsymbol{c}}] + \frac{\delta_1}{2}[\beta_{\boldsymbol{\sigma}|\boldsymbol{c}}]$ satisfies the conditions (13).

Suppose there exists i such that

$$r_i > \frac{2\log|\mathcal{I}(G)|}{\delta_1 \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}}.$$

Then, $\boldsymbol{r}$ cannot be a maximizer of F since

$$\begin{aligned}
F(\boldsymbol{r}) &= \boldsymbol{\lambda}'\cdot\boldsymbol{r} - \sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\log\left(\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\exp\left(\sum_i \sigma_i c_i r_i\right)\right)\\
&= \sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\sum_{\boldsymbol{\rho}\in\mathcal{I}(G)}\widehat{\pi}_{\boldsymbol{\rho}|\boldsymbol{c}}\log\frac{\exp\left(\sum_i \rho_i c_i r_i\right)}{\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\exp(\sum_i \sigma_i c_i r_i)}\\
&\le \sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\widehat{\pi}_{\boldsymbol{0}|\boldsymbol{c}}\log\frac{\exp(0)}{\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\exp(\sum_i \sigma_i c_i r_i)}\\
&\le -\sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\widehat{\pi}_{\boldsymbol{0}|\boldsymbol{c}}c_i r_i \le -\frac{\delta_1}{2}\cdot r_i\cdot\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\\
&< -\log|\mathcal{I}(G)| = F(\boldsymbol{0}) \le \sup_{\boldsymbol{r}\in\mathbb{R}^n} F(\boldsymbol{r}).
\end{aligned}$$

This completes the proof of (11).

Proof of (12): From (11) it suffices to prove that $\boldsymbol{r}$ cannot be a maximizer of F if there exists j such that, for all $k \ne j$,

$$r_j < -\frac{4n^2 c_{\max}\log|\mathcal{I}(G)|}{\delta_1^2 \min_i\left(\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right)^2}$$

and

$$r_k \le \frac{2\log|\mathcal{I}(G)|}{\delta_1 \min_i\{\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\}}.$$

The proof is completed by the following:

$$\begin{aligned}
F(\boldsymbol{r}) &= \boldsymbol{\lambda}'\cdot\boldsymbol{r} - \sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\log\left(\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\exp\left(\sum_i \sigma_i c_i r_i\right)\right)\\
&\le \boldsymbol{\lambda}'\cdot\boldsymbol{r}\\
&\le (n-1)\cdot\frac{2c_{\max}\log|\mathcal{I}(G)|}{\delta_1 \min_i\{\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\}} + \lambda'_j r_j\\
&\le \frac{(n-1)2c_{\max}\log|\mathcal{I}(G)|}{\delta_1 \min_i\{\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\}} + \frac{\delta_1}{2n}\cdot\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_j\pi_{\boldsymbol{c}}\cdot r_j\\
&< -\log|\mathcal{I}(G)| \le F(\boldsymbol{0}) \le \sup_{\boldsymbol{r}\in\mathbb{R}^n} F(\boldsymbol{r}),
\end{aligned}$$

where the first inequality stems from the empty schedule $\boldsymbol{\sigma} = \boldsymbol{0}(\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\exp\left(\sum_i \sigma_i c_i r_i\right) \ge \exp\left(\sum_i 0\cdot c_i r_i\right) = 1)$.

Then, from (11) and (12),

$$\max_i |r_i^*| \le \frac{4n^2 c_{\max}\log|\mathcal{I}(G)|}{\delta_1^2 \min_i\left(\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right)^2}.$$

Furthermore, computing the first derivative of F gives us

$$\begin{aligned}
\frac{\partial}{\partial r_i}F(\boldsymbol{r}^*) &= \lambda'_i - \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\frac{\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\sigma_i\exp\left(\sum_i \sigma_i c_i r_i^*\right)}{\sum_{\boldsymbol{\rho}\in\mathcal{I}(G)}\exp\left(\sum_i \rho_i c_i r_i^*\right)}\\
&= \lambda'_i - \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}} = 0,
\end{aligned} \tag{14}$$

where we now choose an EXP-A-CSMA algorithm such that $\log\frac{f_i(h)}{g_i(h)} = r_i^*\cdot h$. Therefore, it follows that

$$\lambda_i \le \lambda'_i = \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}.$$

This completes the proof of Lemma 3.1.

C. *Proof of Lemma 3.2*

Let $\mathcal{G} = (\mathcal{V},\mathcal{E})$ denote a weighted directed graph induced by Markov process $\{(\boldsymbol{\sigma}(t),\boldsymbol{c}(t))\}$: $\mathcal{V} = \mathcal{I}(G)\times\mathcal{H}^n$ and $((\boldsymbol{\sigma}_1,\boldsymbol{c}_1),(\boldsymbol{\sigma}_2,\boldsymbol{c}_2)) \in \mathcal{E}$ if the transition-rate (which becomes the weight of the edge) from $(\boldsymbol{\sigma}_1,\boldsymbol{c}_1)$ to $(\boldsymbol{\sigma}_2,\boldsymbol{c}_2)$ is non-zero in Markov process $\{(\boldsymbol{\sigma}(t),\boldsymbol{c}(t))\}$. Hence, there are two types of edges:

I. $((\boldsymbol{\sigma}_1,\boldsymbol{c}_1),(\boldsymbol{\sigma}_2,\boldsymbol{c}_2)) \in \mathcal{E}$ and $\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_2$

II. $((\boldsymbol{\sigma}_1,\boldsymbol{c}_1),(\boldsymbol{\sigma}_2,\boldsymbol{c}_2)) \in \mathcal{E}$ and $\boldsymbol{c}_1 = \boldsymbol{c}_2$

A subgraph of $\mathcal{G}$ is called *arborescence* (or spanning tree) with root $(\boldsymbol{\sigma},\boldsymbol{c})$ if for any vertex in $\mathcal{V}\setminus\{(\boldsymbol{\sigma},\boldsymbol{c})\}$, there is exactly one directed path from the vertex to root $(\boldsymbol{\sigma},\boldsymbol{c})$ in the subgraph. Let $\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}}$ and $w(\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}})$ denote the set of *arborescence*s of which root is $(\boldsymbol{\sigma},\boldsymbol{c})$ and the sum of weights of *arborescence*s in $\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}}$, where the weight of an *arborescence* is the product of weight of edges. Then, since the induced Markov process is irreducible, Markov chain tree theorem [1] implies that

$$\pi_{\boldsymbol{\sigma},\boldsymbol{c}} = \frac{w(\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}})}{\sum_{(\boldsymbol{\rho},\boldsymbol{d})\in\mathcal{I}(G)\times\mathcal{H}^n} w(\mathcal{A}_{\boldsymbol{\rho},\boldsymbol{d}})}. \tag{15}$$

Now we further classify the set of *arborescence*s. We let $\mathcal{A}^{(i)}_{\boldsymbol{\sigma},\boldsymbol{c}} \subset \mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}}$ denote the set of *arborescence*s consisting of i edges of type I. Since there are at least $m^n - 1$ type I edges, $\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}} = \bigcup_{i\ge m^n-1}\mathcal{A}^{(i)}_{\boldsymbol{\sigma},\boldsymbol{c}}$. Then, we have

$$\begin{aligned}
w(\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}}) &= \sum_{i\ge m^n-1} w(\mathcal{A}^{(i)}_{\boldsymbol{\sigma},\boldsymbol{c}}) \overset{(a)}{\le} w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})\\
&\quad + \sum_{i\ge m^n}\left(\frac{\delta_2}{m^{2^n m^n(n+1)}}\right)^{i+1-m^n}\cdot|\mathcal{A}^{(i)}_{\boldsymbol{\sigma},\boldsymbol{c}}|\cdot w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})\\
&\le w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})\cdot\left(1+\sum_{i\ge m^n}\left(\frac{\delta_2}{m^{2^n m^n(n+1)}}\right)^{i+1-m^n}\cdot|\mathcal{A}^{(i)}_{\boldsymbol{\sigma},\boldsymbol{c}}|\right)\\
&\le w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})\cdot\left(1+\frac{\delta_2}{m^{2^n m^n(n+1)}}\cdot|\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}}|\right)\\
&\overset{(b)}{<} w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})\cdot(1+\delta_2).
\end{aligned}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

For $i \geq m^n$, we can transform every *arborescence* in $\mathcal{A}^{(i)}_{\boldsymbol{\sigma},\boldsymbol{c}}$ to one of *arborescence*s in $\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}}$ by removing $i+1-m^n$ type I edges, adding $i+1-m^n$ type II edges, and changing directions of at most $n(i+1-m^n)$ type II edges. From the definition of ψ and the condition in Lemma 3.2, the transformation increases the weight of *arborescence* at most $\left(\frac{\delta_2}{m^{2^n m^n(n+1)}}\right)^{i+1-m^n}$ times larger than before. Thus, in the above equation, (a) holds. For (b) of the above equation, we use the inequality $|\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}}| < (mn)^{2^n m^n}$. Therefore, using the above inequality, it follows that

$$\begin{aligned}\frac{\pi_{\boldsymbol{\sigma},\boldsymbol{c}}}{\pi_{\boldsymbol{c}}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}} &= \frac{w(\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}})}{w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})} \cdot \frac{\sum\limits_{\boldsymbol{d}\in\mathcal{H}^n} \sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\rho},\boldsymbol{d}})}{\sum\limits_{\boldsymbol{d}\in\mathcal{H}^n} \sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}_{\boldsymbol{\rho},\boldsymbol{d}})} \\ &< 1+\delta_2,\end{aligned}$$

where the first equality follows from (15) and

$$\pi_{\boldsymbol{c}}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}} = \frac{w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}})}{\sum\limits_{\boldsymbol{d}\in\mathcal{H}^n} \sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\rho},\boldsymbol{d}})}. \tag{16}$$

Similarly, one can also show that $\frac{\pi_{\boldsymbol{\sigma},\boldsymbol{c}}}{\pi_{\boldsymbol{c}}\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}} > 1-\delta_2$. This completes the proof of Lemma 3.2.

Proof of (16): We analogously define *arborescences* for Markov process $\{\boldsymbol{c}(t)\}$ and *arborescences* for Markov process $\{\boldsymbol{\sigma}(t), \boldsymbol{c}\}$, where $\mathcal{A}_{\boldsymbol{c}}$ and $\mathcal{A}_{\boldsymbol{\sigma}|\boldsymbol{c}}$ denote the sets of *arborescences* of which roots are $(\boldsymbol{c})$ and $(\boldsymbol{\sigma}, \boldsymbol{c})$,[4] respectively. Then,

$$w(\mathcal{A}^{(m^n-1)}_{\boldsymbol{\sigma},\boldsymbol{c}}) = w(\mathcal{A}_{\boldsymbol{c}})w(\mathcal{A}_{\boldsymbol{\sigma}|\boldsymbol{c}}) \prod_{\boldsymbol{d}\in\mathcal{H}^n, \boldsymbol{d}\neq\boldsymbol{c}} \sum_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}_{\boldsymbol{\rho}|\boldsymbol{d}}).$$

We can deduce (16) from the following:

$$\begin{aligned}\pi_{\boldsymbol{c}} &= \frac{w(\mathcal{A}_{\boldsymbol{c}})}{\sum\limits_{\boldsymbol{d}\in\mathcal{H}^n} w(\mathcal{A}_{\boldsymbol{d}})} \text{ and } \pi_{\boldsymbol{\sigma}|\boldsymbol{c}} = \frac{w(\mathcal{A}_{\boldsymbol{\sigma}|\boldsymbol{c}})}{\sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}_{\boldsymbol{\rho}|\boldsymbol{c}})} \\ &= \frac{w(\mathcal{A}_{\boldsymbol{\sigma}|\boldsymbol{c}}) \prod\limits_{\boldsymbol{d}\in\mathcal{H}^n, \boldsymbol{d}\neq\boldsymbol{c}} \sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}_{\boldsymbol{\rho}|\boldsymbol{d}})}{\sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}_{\boldsymbol{\rho}|\boldsymbol{c}}) \prod\limits_{\boldsymbol{d}\in\mathcal{H}^n, \boldsymbol{d}\neq\boldsymbol{c}} \sum\limits_{\boldsymbol{\rho}\in\mathcal{I}(G)} w(\mathcal{A}_{\boldsymbol{\rho}|\boldsymbol{d}})}.\end{aligned}$$

D. *Proof of Theorem 3.2*

We will show that if we suppose that $k(x)$ is not linear, EXP(k)-A-CSMA is not throughput-optimal, i.e., there exists a channel model and an interference graph where EXP(k)-A-CSMA is not rate-stable for some arrival rate vector within the maximum achievable rate region.

First, from our hypothesis that $k(x)$ is not linear, we can first find positive constants x_1, x_2, ε, and δ such that $0 < \delta \leq 1, 0 < \varepsilon \leq \min\{x_1, x_2\}$ and for all $0 < \varepsilon_1, \varepsilon_2 \leq \varepsilon$,

$$\frac{k(x_2+\varepsilon_2)-k(x_2)}{\varepsilon_2} \geq \frac{k(x_1+\varepsilon_1)-k(x_1)}{\varepsilon_1}(1+\delta). \tag{17}$$

The above equation roughly means that there exist two "line segments" (where each line segment is generated by two points over $k(x)$), whose slope difference exceed one.

Second, we now present an interference graph and a channel model of our interest, and the achievable rate vector under the assumed interference graph and the channel.

∘ *Interference graph:* We consider a complete bipartite interference graph with four links $\{1, 2, 3, 4\}$, where each odd link 1 and 3 is connected to (thus interferes with) each even link 2 and 4.

∘ *Channel model:* For x_1, x_2, ε, and δ in (17), we consider the following 5 channel states, each of which is characterized by the capacities of four links:

$$\begin{aligned}a &: && (x_1, x_1, x_2, x_2), \\ b &: && (x_1, x_2, x_2, x_1), \\ c &: && \left(x_1 + \tfrac{\varepsilon(1+\delta)}{2}, x_2 + \tfrac{\varepsilon}{2}, x_2, x_1\right), \\ d &: && (x_2, x_1, x_1, x_2), \\ e &: && \left(x_2, x_1, x_1 + \tfrac{\varepsilon(1+\delta)}{2}, x_2 + \tfrac{\varepsilon}{2}\right).\end{aligned} \tag{18}$$

Note that in the channel state c, the sum of channel capacities of odd links is constructed to be larger than that of even links by $\varepsilon\delta/2$, and analogously for the channel state e. The channel transition rates are assumed to be: $\gamma^{i\to j} = 1$ if $i \in \{a, b, d\}, j \in \{c, e\}$ or $i \in \{c, e\}, j \in \{a, b, d\}$, but $\gamma^{i\to j} = 0$ otherwise. Under these transition rates, we can easily check that the stationary channel occurrence probability is equal to 1/5 for all channels.

∘ *Arrival vector.* Consider the following arrival rate vector $\boldsymbol{\lambda} = [\lambda_i]$:

$$\begin{aligned}\lambda_1 = \lambda_3 &= \frac{1}{5}\left(x_1 + x_2 + \frac{\varepsilon(1+\delta)}{2} - \frac{\varepsilon\delta}{40}\right) \\ \lambda_2 &= \frac{1}{5}(2x_1 + x_2), \quad \text{and} \quad \lambda_4 = \frac{1}{5}(x_1 + 2x_2),\end{aligned} \tag{19}$$

which clearly leads to $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o$, because $\boldsymbol{\lambda}$ is constructed by considering the service rates from a scheduler that for the channel states $\{a, b, d\}$ even links are scheduled and for the channel states $\{c, e\}$ odd links are scheduled, from which $\varepsilon\delta/40$ is subtracted in λ_1 and λ_3.

In what follows, we denote by $\pi_{S,Z}$ the stationary probability that a schedule $S \subset \{1, 2, 3, 4\}$ holds channel under each channel state $Z \in \{a, b, c, d, e\}$.[5] To prove that with the non-linear $k(\cdot)$ it is impossible to stabilize $\boldsymbol{\lambda}$, we will take the following two steps. *Step (i):* we will first show that the following **C1** is a necessary condition to stabilize $\boldsymbol{\lambda}$:

C1. $\pi_{\{1,3\},Z} > \frac{3}{20}$ when $Z \in \{c, e\}$, and $\pi_{\{2,4\},Z} > \frac{3}{20}$ when $Z \in \{a, b, d\}$.

Condition **C1** means that even (resp. odd) link schedule should be played more than odd (resp. even) ones, when channel state $Z \in \{a, b, d\}$ (resp. $Z \in \{c, e\}$). For instance, $\pi_{\{2,4\},a} > \sum_{S:S\neq\{2,4\}} \pi_{S,a}$ since $\pi_c = \sum_S \pi_{S,c} = \frac{1}{5}$ and $\pi_{\{2,4\},a} \geq \frac{3}{20}$. Next, in *Step (ii)*, we will show that $k(x)$ satisfying (17) cannot satisfy **C1.**

[4]Since Markov process $\{\boldsymbol{\sigma}(t), \boldsymbol{c}\}$ changes only $\boldsymbol{\sigma}$, the state space is the scheduling set.

[5]For ease of presentation, we temporarily use this notation $\pi_{S,Z}$ only in this proof to mean that S and Z is a schedule (i.e., a subset of non-interfering links) and a channel state, respectively, different from $\pi_{\boldsymbol{\sigma},\boldsymbol{c}}$ elsewhere.

Step (i). We start by noting the following, which provides the bound on the "distance" between the rate region boundary and the given arrival rate (19):

$$\max_{\boldsymbol{\rho}\in\boldsymbol{\Lambda}} \sum_{i=1}^{4}(\rho_i - \lambda_i) = 2 \times \frac{1}{5} \times \frac{\varepsilon\delta}{40} = \frac{\varepsilon\delta}{100}. \tag{20}$$

A necessary condition for $\boldsymbol{\lambda}$-stabilization is that the sum of the potential service rates over even links $\{2,4\}$ exceeds the sum of arrival rates over those links, i.e., $\lambda_2 + \lambda_4 = \frac{3}{5}(x_1 + x_2)$, and

$$\begin{aligned}\frac{3}{5}(x_1+x_2) \le& \sum_{S:|S\cap\{2,4\}|\ge 1} \sum_{Z\in\{a,b,d\}} (x_1+x_2)\pi_{S,Z} \\ &+ \sum_{S:|S\cap\{2,4\}|\ge 1} \sum_{Z\in\{c,e\}} (x_1+x_2+\frac{\varepsilon}{2})\pi_{S,Z} \\ \le& \sum_{S:|S\cap\{2,4\}|\ge 1} \sum_{Z\in\{a,b,d\}} (x_1+x_2)\pi_{S,Z} + \frac{x_1+x_2}{40},\end{aligned} \tag{21}$$

where the last inequality is due to $\varepsilon \le \min\{x_1, x_2\}$ (thus $\frac{\varepsilon}{2} \le \frac{x_1+x_2}{4}$) and the following upper bound on the even links' potential service rates for the $\{c,e\}$ channels:

$$\sum_{S:|S\cap\{2,4\}|\ge 1} \sum_{Z\in\{c,e\}} \pi_{S,Z} \le \frac{\varepsilon\delta}{100}\frac{2}{\varepsilon\delta} \le \frac{1}{50}, \tag{22}$$

which holds because from (20) and the sum of channel capacities of odd links exceeds that of even links by $\frac{\varepsilon\delta}{2}$ for the channel state $\{c,e\}$.

Note that the capacity of each link for every channel state is at least $\min\{x_1,x_2\}$ from our channel construction in (18). Then, since an idle link makes at least $\min\{x_1,x_2\}$ rate loss and $\varepsilon\delta \le \min\{x_1,x_2\}$, (20) requires to satisfy:

$$\sum_{S:|S|=2} \sum_{Z} \pi_{S,Z} \ge 1 - \frac{\varepsilon\delta}{100\min\{x_1,x_2\}} \ge 1 - \frac{1}{100}. \tag{23}$$

By dividing LHS and RHS of (21) by (x_1+x_2), we deduce another bound

$$\sum_{S:|S\cap\{2,4\}|\ge 1} \sum_{Z\in\{a,b,d\}} \pi_{S,Z} \ge \frac{3}{5} - \frac{1}{40}. \tag{24}$$

Recall that $\pi_Z = 1/5$ for all channel states Z. Then, from (22) and (23), $\pi_{\{1,3\},c} + \pi_{\{1,3\},e} \ge \frac{2}{5} - \frac{1}{50} - \frac{1}{100}$ and from (23) and (24), $\pi_{\{2,4\},a} + \pi_{\{2,4\},b} + \pi_{\{2,4\},d} \ge \frac{3}{5} - \frac{1}{40} - \frac{1}{100}$, which implies **C1.**

Step (ii). We will prove that under **C1**, non-linear $k(x)$ leads to a contradiction. We first let c_i^Z denote the capacity of link i at channel state Z. Then, we first present the key lemma to this proof, which tells us the relation of the links' intensities under **C1** for the channel states $\{a,b,d\}$ and $\{c,e\}$.

Lemma 3.3: Under Condition **C1**, we should have:

$$\begin{aligned}&r_1k(c_1^Z) + r_3k(c_3^Z) < r_2k(c_2^Z) + r_4k(c_4^Z) \text{ if } Z \in \{a,b,d\},\\ &r_1k(c_1^Z) + r_3k(c_3^Z) > r_2k(c_2^Z) + r_4k(c_4^Z) \text{ if } Z \in \{c,e\}.\end{aligned}$$

Proof: This proof is presented in the supplementary material. ■

To intuitively understand why Lemma 3.3 holds, let's consider the case when the backoff rates $[f_i(c_i^Z)]$ and the holding rates $[g_i(c_i^Z)]$ significantly exceed the channel varying speed, so that the condition of Lemma 3.2 holds for sufficiently small δ_2. Then, Lemma 3.2 and Condition **C1** leads to:

$$\pi_{\{2,4\}|a} \ge \frac{1}{1-\delta_2}\frac{\pi_{\{2,4\},a}}{\pi_a} \ge \frac{1}{2}, \tag{25}$$

where $\pi_{\{2,4\}|a}$ is the stationary probability of the schedule $\{2,4\}$ given the channel state a. In order for (25) to hold, the sum of transmission intensities of even links should be larger than that of odd links (*i.e.*, $r_2k(c_2^a) + r_4k(c_4^a) > r_1k(c_1^a) + r_3k(c_3^a)$). Lemma 3.3 formalizes this intuition by considering the general $[f_i(c_i^Z)]$ and $[g_i(c_i^Z)]$.

By using the actual capacities in (18) for each channel state a, b, c, d, and e, under **C1** Lemma 3.3 gives us the following:

$$r_1k(x_1) + r_3k(x_2) < r_2k(x_1) + r_4k(x_2), \tag{26}$$

$$r_1k(x_1) + r_3k(x_2) < r_2k(x_2) + r_4k(x_1), \tag{27}$$

$$r_1k\left(x_1+\frac{\varepsilon(1+\delta)}{2}\right) + r_3k(x_2) > r_2k\left(x_2+\frac{\varepsilon}{2}\right) + r_4k(x_1), \tag{28}$$

$$r_1k(x_2) + r_3k(x_1) < r_2k(x_1) + r_4k(x_2), \tag{29}$$

$$r_1k(x_2) + r_3k\left(x_1+\frac{\varepsilon(1+\delta)}{2}\right) > r_2k\left(x_1+\frac{\varepsilon}{2}\right) + r_4k(x_2). \tag{30}$$

Suppose that (27), (28), (29), and (30) holds. Then, from (27) and (28), we can derive that

$$\begin{aligned}r_1 &> r_2\frac{k(x_2+\frac{\varepsilon}{2}) - k(x_2)}{k(x_1+\frac{\varepsilon(1+\delta)}{2}) - k(x_1)}\\ &= r_2\frac{(k(x_2+\frac{\varepsilon}{2}) - k(x_2))/\varepsilon}{(k(x_1+\frac{\varepsilon(1+\delta)}{2}) - k(x_1))/(\varepsilon(1+\delta))}\frac{1}{1+\delta} \ge r_2,\end{aligned}$$

where for the last inequality we use (17) from our hypothesis of the non-linearity of k. Analogously, from (29) and (30), we can conclude $r_3 > r_4$. Since $r_1 > r_2, r_3 > r_4$, and $k(x)$ is a positive function, we must have that

$$(r_2 - r_1)k(x_1) + (r_4 - r_3)k(x_2) < 0,$$

which, however, contradicts with (26), completing the proof of **Step (ii).** Then, the proof of Theorem 3.2 is completed.

IV. Dynamic Throughput Optimal A-CSMA

In the previous section, it is shown that, for any feasible arrival rate, there exists an EXP-A-CSMA algorithm stabilizing the arrivals. In this section, we describe EXP-A-CSMA algorithms which dynamically update its parameters so as to stabilize the network without knowledge of the arrival statistics. More precisely, the CSMA scheduling algorithm uses $f_i^{(t)}$ and $g_i^{(t)}$ to compute the value of parameters $R_i(t) = f_i^{(t)}(c_i(t))$ and $S_i(t) = g_i^{(t)}(c_i(t))$ at time t, respectively, and update them adaptively over time. We present two algorithms to decide $f_i^{(t)}$ and $g_i^{(t)}$. They are building upon prior algorithms in conjunction with the properties of EXP-A-CSMA established in the previous section, referred to as a rate-based (extension of [5]) and queue-based algorithm (extension of [26]).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A. Rate-Based Algorithm

The first algorithm, at each link i, updates $(f_i^{(t)}, g_i^{(t)})$ at time instances $L(j), j \in \mathbb{Z}_+$ with $L(0) = 0$. Thus, $(f_i^{(t)}, g_i^{(t)})$ is fixed for each time-interval $[L(j), L(j+1))$. Let $T(j) = L(j+1) - L(j)$ for $j \geq 0$ and, with an abuse of notation, $f_i^{(j)}, g_i^{(j)}$ denote the value of $f_i^{(t)}, g_i^{(t)}$ for $t \in [L(j), L(j+1))$, respectively. To begin with, the algorithm sets $f_i^{(0)}(x) = g_i^{(0)}(x) = 1$ (i.e., $R_i(0) = S_i(0) = 1$) for all i and all $x \in [0,1]$.

Now we describe how to choose a varying update interval $T(j)$. We select $T(j) = \exp\left(\sqrt{j}\right)$, for $j \geq 1$, and choose a step-size $\alpha(j)$ of the algorithm as $\alpha(j) = \frac{1}{j}$, for $j \geq 1$. Given this, link i updates f_i and g_i as follows. Let $\hat{\lambda}_i(j), \hat{s}_i(j)$ be empirical arrival and service observed at link i in $[L(j), L(j+1))$, *i.e.*,

$$\hat{\lambda}_i(j) = \frac{1}{T(j)} A_i(L(j), L(j+1))$$

and

$$\hat{s}_i(t) = \frac{1}{T(j)} \left[\int_{L(j)}^{L(j+1)} \sigma_i(t) c_i(t)\, dt\right].$$

Then, the update rule is defined by, for $x \in [0,1]$

$$\begin{aligned} g_i^{(j+1)}(x) &= g_i^{(j)}(x) \exp\left(2nc_{\max}^2 \alpha(j)\right) \\ f_i^{(j+1)}(x) &= g_i^{(j+1)}(x) \exp\left(r_i(j+1)\right), \end{aligned} \tag{31}$$

where

$$r_i(j+1) = r_i(j) + \alpha(j) \cdot (\hat{\lambda}_i(j) - \hat{s}_i(j))$$

with initial condition $r_i(1) = 0$.

Note that, under this update rule, the algorithm at each link i uses only its local history. Despite this, we establish that this algorithm is rate-stable, as formally stated as follows:

Theorem 4.1: For any given graph G and channel transition-rate $\boldsymbol{\gamma}$, the A-CSMA algorithm with updating functions as per (31) is throughput optimal.

Proof: This proof is presented in the supplementary material. ■

B. Queue-Based Algorithm

Now we describe the second algorithm which chooses $(f_i^{(t)}, g_i^{(t)})$ as a simple function of queue-sizes as follows.

$$f_i^{(t)}(x) = \exp\left(2nc_{\max} w(Q_{\max}(\lfloor t \rfloor)) + x r_i(t)\right)$$

and

$$g_i^{(t)}(x) = \exp\left(2nc_{\max} w(Q_{\max}(\lfloor t \rfloor))\right), \tag{32}$$

where $r_i(t) = \max\left\{w(Q_i(\lfloor t \rfloor)), \sqrt{w(Q_{\max}(\lfloor t \rfloor))}\right\}$ with weight function $w(x) = \log\log(x+e)$ and $Q_{\max}(\lfloor t \rfloor) = \max_j Q_j(\lfloor t \rfloor)$. One can interpret this as an EXP-A-CSMA algorithm since

$$\log \frac{f_i^{(t)}(x)}{g_i^{(t)}(x)} = r_i(t) \cdot x.$$

The global information of $Q_{\max}(\lfloor t \rfloor)$ can be replaced by its approximate estimation that can be computed through a very simple distributed algorithm (with message-passing) in [24] or a learning mechanism (without message-passing) in [27]. This does not alter the rate-stability of the algorithm that is stated in the following theorem.

Theorem 4.2: For any given graph G and channel transition-rate $\boldsymbol{\gamma}$, the A-CSMA algorithm with functions as per (32) is throughput optimal.

Proof: This proof is presented in the supplementary material. ■

Note that we design $f_i^{(t)}(x)$ and $g_i^{(t)}(x)$ such that they grow as the maximum queue length $Q_{\max}(t)$ increases, *i.e.*, backoff and holding rates of nodes become large when $Q_{\max}(t)$ does. Then, from Lemma 3.2, the stationary distribution $\pi_{\boldsymbol{\sigma},\boldsymbol{c}}$ induced by the queue-based algorithm is approximated well by a product-form (channel and scheduling) distribution when $Q_{\max}(t)$ is large. Once we have such a property, we can use identical proof arguments in [26] for establishing the desired throughput optimality. Irrespective of such a technical reason, increasing backoff and holding rates with respect to the maximum queue length is quite natural since it is necessary for chasing arbitrarily fast channel variations.

V. Achievable Rate Region of A-CSMA With Limited Backoff Rate

In practice, it might be hard to have arbitrary large backoff rate due to physical constraints. This motivates us to investigate the achievable rate region of A-CSMA algorithms with limited backoff rate. Note that, in the proof of Theorem 3.1, we choose the backoff rates $[f_i]$ to be proportional to the channel varying speed. Thus, when the backoff rate is limited and the channel varying speed grows up, it is clear that we cannot guarantee the optimality of EXP-A-CSMA. It can be simply shown that, for given channel transition rate $\boldsymbol{\gamma}$, the rate region of EXP-A-CSMA grows as the backoff limit decreases, where the backoff limit (denoted by ϕ) is a bound of backoff functions $[f_i]$ such that $\max_{i \in V, x \in [0, c_{\max}]} f_i(x) \leq \phi$. It is because the decreasing backoff limit removes some possible choices of $[f_i]$ in EXP-A-CSMA. In this section, we will give lower bounds of achievable rate region for such limited backoff rate.

When the backoff limit is large, EXP-A-CSMA algorithm can achieve ε-close the capacity. In Theorem 5.1, we establish a sufficient condition to achieve $(1-\varepsilon)$ fraction of capacity region, where the sufficient condition can be tightened more for complete graphs.

Theorem 5.1: For any given $0 < \varepsilon < 1$, interference graph G, channel transition-rate $\boldsymbol{\gamma}$ and arrival rate $\boldsymbol{\lambda} \in (1-\varepsilon)\boldsymbol{\Lambda}^o$, there exists a rate-stable EXP-A-CSMA algorithm with functions $[f_i]$ and $[g_i]$ such that $\max_{i \in V, x \in [0, c_{\max}]} f_i(x) \leq \phi$, if

$$\phi \geq \eta\left(\frac{\varepsilon}{1+\varepsilon}, G, \boldsymbol{\gamma}\right) \exp(\kappa(\varepsilon^2, G, \boldsymbol{\gamma})),$$

where κ and η are defined in (9) and (10), respectively. If G is a complete graph, the above condition on ϕ can be tightened further as:

$$\phi \geq \psi \cdot \frac{\exp\left(2c_{\max} \log(n+1) / (\varepsilon^2 \min_i \{\sum_{\boldsymbol{c} \in \mathcal{H}^n} c_i \pi_{\boldsymbol{c}}\})\right)}{\sqrt{1+\varepsilon} - 1},$$

where we recall the channel varying speed is defined as $\psi = \max_{\boldsymbol{u}\in\mathcal{H}^n}\{\sum_{\boldsymbol{v}\in\mathcal{H}^n:\boldsymbol{v}\neq\boldsymbol{u}}\gamma^{\boldsymbol{u}\to\boldsymbol{v}}\}$.

Proof: This proof is presented in Section V-A. ■

The above theorem implies that backoff rates might be necessarily huge for throughput optimality of EXP-A-CSMA. On the other hand, in the following theorems, we show that EXP-A-CSMA even with limited backoff rates is no worse than U-CSMA in throughput and achieves at least a constant fraction α of the maximum throughput region, where α only depends on the stationary distribution $[\pi_{\boldsymbol{c}}]$ and the interference graph G.

Theorem 5.2: For any $\varepsilon > 0, \phi > 0$, interference graph G, channel transition-rate $\boldsymbol{\gamma}$, and U-CSMA parameters, there exists a EXP-A-CSMA algorithm with functions $[f_i]$ and $[g_i]$ such that $\max_{i\in V, x\in[0,c_{\max}]} f_i(x) \le \phi$ and

$$\limsup_{t\to\infty}\left|1-\frac{\widehat{D}_i^A(t)}{\widehat{D}_i^U(t)}\right| < \varepsilon, \quad \text{for all } i\in V,$$

where $\widehat{D}_i^A$ and $\widehat{D}_i^U$ denote the cumulative potential departure processes of the EXP-A-CSMA and the U-CSMA, respectively.

Proof: This proof is presented in Section V-B. ■

Theorem 5.3: For any $\phi > 0$, interference graph G, channel transition-rate $\boldsymbol{\gamma}$ and arrival rate $\boldsymbol{\lambda} \in \alpha\boldsymbol{\Lambda}^o$, there exists a rate-stable EXP-A-CSMA algorithm with functions $[f_i]$ and $[g_i]$ such that

$$\max_{i\in V, x\in[0,1]} f_i(x) \le \phi,$$

where

$$\alpha = \max\left\{\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}\frac{c_i\pi_{\boldsymbol{c}}}{c_{\max}}, \frac{1}{\chi(G)}\right\}. \tag{33}$$

In above, $\chi(G)$ is the chromatic number of G.

Proof: This proof is presented in Section V-C. ■

The above theorem implies that for a bipartite interference graph, at least 50%-throughput can be achieved under EXP-A-CSMA regardless of constraints on backoff rates since its chromatic number is two.

A. *Proof of Theorem 5.1*

1) General Graphs: From the proof of Theorem 3.1, we naturally get this bound. We have shown that, in (8), any arrival rate $\boldsymbol{\lambda} \in (1-\varepsilon)\boldsymbol{\Lambda}^o$ can be stabilized by an EXP-A-CSMA algorithm with

$$f_i(x) = \eta\left(\frac{\varepsilon}{1+\varepsilon}, G, \gamma\right)\exp(\kappa(\varepsilon^2, G, \gamma))$$

for all link i and channel state x, since $\delta_1 = \varepsilon^2$ and $\delta_2 = \frac{\varepsilon}{1+\varepsilon}$ in the proof of Theorem 3.1.

2) Complete Graphs: Let us consider a EXP-A-CSMA with $g_i(x) = S$ and $f_i(x) = S\exp(r_i x)$. For notational convenience, let $\pi_{i,\boldsymbol{c}}$ be the stationary distribution induced by EXP-A-CSMA with the above $g_i(x)$ and $f_i(x)$ that the link i holds the channel and channel state is $\boldsymbol{c}$.

From (11), for arrival rate $\boldsymbol{\lambda}^* \in (1-\varepsilon^2)\boldsymbol{\Lambda}^o$, there exists $\boldsymbol{r}^*$ such that $\lambda_i^* = \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\sigma_i\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}$ and

$$\max_i r_i^* \le \frac{2\log(n+1)}{\varepsilon^2\min_i\left\{\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right\}}.$$

We will later show the following lower bound on $\pi_{i,\boldsymbol{c}}$: For each link i,

$$\pi_{i,\boldsymbol{c}} \ge \pi_{\boldsymbol{c}}\frac{S}{S+\psi_{\boldsymbol{c}}}\frac{S\exp(c_i r_i)}{S+\psi_{\boldsymbol{c}}+\sum_\ell S\exp(c_\ell r_\ell)}, \tag{34}$$

where $\psi_{\boldsymbol{c}} := \sum_{\boldsymbol{v}\in\mathcal{H}^n:\boldsymbol{v}\neq\boldsymbol{c}}\gamma^{\boldsymbol{c}\to\boldsymbol{v}}$. Therefore, for $g_i(x) = S = \frac{\psi}{\sqrt{1+\varepsilon}-1}$, and $f_i(x) = S\exp(r_i^* x)$, EXP-A-CSMA satisfies $\max_{i\in V, x\in[0,1]} f_i(x) \le \phi$ and stabilizes an arrival vector $\frac{1}{1+\varepsilon}\boldsymbol{\lambda}^*$ such that $\boldsymbol{\lambda}^* \in (1-\varepsilon^2)\boldsymbol{\Lambda}^o$, since

$$\begin{aligned}
\lim_{t\to\infty}\frac{1}{t}\widehat{D}_i(t) &= \sum_{\boldsymbol{c}\in\mathcal{H}^n}\sum_i c_i\pi_{i,\boldsymbol{c}} \\
&> \frac{1}{1+\psi/S}\sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\frac{c_i\exp(c_i r_i^*)}{1+\psi/S+\sum_\ell\exp(c_\ell r_\ell^*)} \\
&\ge \frac{1}{(1+\psi/S)^2}\sum_{\boldsymbol{c}\in\mathcal{H}^n}\pi_{\boldsymbol{c}}\frac{c_i\exp(c_i r_i^*)}{1+\sum_\ell\exp(c_\ell r_\ell^*)} \\
&= \frac{1}{1+\varepsilon}\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)}\sigma_i\pi_{\boldsymbol{\sigma}|\boldsymbol{c}} \\
&= \frac{1}{1+\varepsilon}\lambda_i^*,
\end{aligned}$$

which implies that for any arrival rate $\boldsymbol{\lambda} \in (1-\varepsilon)\boldsymbol{\Lambda}^o$, there exist an EXP-A-CSMA that stabilizes $\boldsymbol{\lambda}$.

Proof of (34): To complete the proof, we now show (34). Note that since G is a complete graph, only one link can be active at any time. We divide the entire time into three phases *P1, P2* and *P3*:

- ∘ Whenever any link becomes active, *P1* starts.
- ∘ *P1* moves to *P2* when the channel of any link changes.
- ∘ *P1* or *P2* moves to *P3* if no link is active.

Based on these time phases, we define two random processes $a(t)$ and $b(t)$ as follows:

- ∘ $a(t) = 1$ and $b(t) = 0$, if t is in *P1*.
- ∘ $a(t) = 0$ and $b(t) = 1$, if t is in *P2*.
- ∘ $a(t) = 0$ and $b(t) = 0$, if t is in *P3*.

One can observe that the media is sensed busy at time t if and only if $a(t) + b(t) = 1$.

Using $b(t)$, consider a random process $\{X_{\boldsymbol{c}}(t) \in \{0, 1, \ldots, n\} : t \ge 0\}$ such that

$$X_{\boldsymbol{c}}(T(t)) = \begin{cases} i & \text{if } \sigma_i(t) = 1 \text{ for some } i\in\{1, 2\ldots, n\} \\ 0 & \text{otherwise}\end{cases}$$

where

$$T(t) = \int_{\tau=0}^{t}(1-b(\tau))\mathbf{1}_{\boldsymbol{c}}(\tau)d\tau,$$

and $\mathbf{1}_{\boldsymbol{c}}(t) = 1$ when the channel state at time t is $\boldsymbol{c}$, and 0 otherwise. In other words, $\{X_{\boldsymbol{c}}(t)\}$ is a (truncated) random process on the schedule under the complete graph, where it considers only the phases *P1, P3* and when the channel state is $\boldsymbol{c}$. Then,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

one can easily check that $\{X_{\boldsymbol{c}}(t)\}$ is a Markov process, whose transition rates are given by: for all $1 \le i, j \le n$ and $i \neq j$,

$$\begin{aligned} 0 \to i &\quad \text{with rate} \quad S\exp(c_i r_i),\\ i \to 0 &\quad \text{with rate} \quad S + \psi_{\boldsymbol{c}},\\ j \to i &\quad \text{with rate} \quad 0. \end{aligned}$$

Note that the transition is not allowed between any two scheduled links i and j, because $\{X_{\boldsymbol{c}}(t)\}$ removes *P2* and takes into account only when the channel state is $\boldsymbol{c}$. Let $[\mu_i]_{0 \le i \le n}$ be the stationary distribution of $\{X_{\boldsymbol{c}}(t)\}$, which has the following form:

$$\mu_i = \frac{S\exp(c_i r_i)}{S + \psi_{\boldsymbol{c}} + \sum_\ell S\exp(c_\ell r_\ell)}. \tag{35}$$

We now introduce another random process $b'(t)$ which is equal to $b(t)$ in the phase *P2* (*i.e.*, $b'(t) = 1$ if $b(t) = 1$). The random process $b'(t)$, however, is possibly different from $b(t)$ in the phases *P1* and *P3*. At every channel transition in the phase *P3*, we set $b'(t) = 1$ for a random time period while $b(t) = 0$. More precisely, when a channel transition occurs at time t_1 in the phase *P3*, set $b'(t) = 1$ for $t_1 \le t \le \min\{t_2, t_1 + t_s\}$ where t_s is an exponential random variable with mean $1/S$ and t_2 is the next channel transition time. Note that if there is no active link at t_2 (i.e., t_2 is in the phase *P3*), we analogously set $b'(t) = 1$ for a random time period; otherwise, the phase *P2* begins and thus $b'(t) = b(t) = 1$. Therefore, every channel transition makes $b'(t) = 1$ and from the memoryless property of the exponential distribution, $b'(t)$ keeps 1 for an exponential random time with mean $1/S$. We provide an exapmle to show the difference between $b(t)$ and $b'(t)$ in the supplementary material. Using $b'(t)$, we define a random process $\{X'_{\boldsymbol{c}}(t) \in \{0,1\} : t \ge 0\}$:

$$X'_{\boldsymbol{c}}(T'(t)) = b'(t) \quad \text{and} \quad T'(t) = \int_{\tau=0}^{t} \mathbf{1}_{\boldsymbol{c}}(\tau)d\tau.$$

Then, from the definition of $b'(t)$, $\{X'_{\boldsymbol{c}}(t)\}$ is a Markov process where the transition rate from 0 to 1 is the channel transition rate $\psi_{\boldsymbol{c}}$ since every channel transition makes $b'(t) = 1$, and the transition rate from state 1 to state 0 becomes S from the memoryless property of the exponential distribution. We plot the state transition diagram of $X'_{\boldsymbol{c}}$ in Fig. 1. Therefore, we have

$$\lim_{t\to\infty} \mathbb{P}\{X'_{\boldsymbol{c}}(t) = 1\} = \frac{\psi_{\boldsymbol{c}}}{S + \psi_{\boldsymbol{c}}}. \tag{36}$$

We denote by $p_{i,\boldsymbol{c},a}$ (resp. $p_{i,\boldsymbol{c},b}$ and $p_{i,\boldsymbol{c},b'}$) the long-term time fraction that link i is active, channel state is $\boldsymbol{c}$, and $a(t) = 1$ (resp. $b(t) = 1$ and $b'(t) = 1$), *i.e.*,

$$\begin{aligned} p_{i,\boldsymbol{c},a} &= \lim_{T\to\infty} \frac{1}{T}\int_{t=0}^{T} \sigma_i(t)\mathbf{1}_{\boldsymbol{c}}(t)a(t)dt.\\ p_{i,\boldsymbol{c},b} &= \lim_{T\to\infty} \frac{1}{T}\int_{t=0}^{T} \sigma_i(t)\mathbf{1}_{\boldsymbol{c}}(t)b(t)dt,\\ p_{i,\boldsymbol{c},b'} &= \lim_{T\to\infty} \frac{1}{T}\int_{t=0}^{T} \sigma_i(t)\mathbf{1}_{\boldsymbol{c}}(t)b'(t)dt. \end{aligned}$$

Then, for each link i,

$$\pi_{i,\boldsymbol{c}} = p_{i,\boldsymbol{c},a} + p_{i,\boldsymbol{c},b} \tag{37}$$

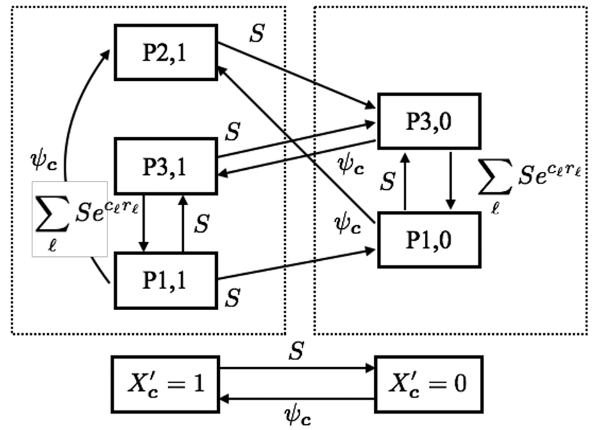


Fig. 1. State transition diagram of $X'_{\boldsymbol{c}}$, where (Px,y) refers to the phase *Px* with $b' = y$.

$$\begin{aligned} &\ge \left(\pi_{\boldsymbol{c}} - \sum_\ell p_{\ell,\boldsymbol{c},b}\right)\frac{p_{i,\boldsymbol{c},a}}{\pi_{\boldsymbol{c}} - \sum_\ell p_{\ell,\boldsymbol{c},b}}\\ &= \left(\pi_{\boldsymbol{c}} - \sum_\ell p_{\ell,\boldsymbol{c},b}\right)\mu_i && (38)\\ &\ge \pi_{\boldsymbol{c}}\left(1 - \frac{\sum_\ell p_{\ell,\boldsymbol{c},b'}}{\pi_{\boldsymbol{c}}}\right)\mu_i && (39)\\ &= \pi_{\boldsymbol{c}}\left(1 - \lim_{t\to\infty}\mathbb{P}\{X'_{\boldsymbol{c}}(t) = 1\}\right)\mu_i && (40)\\ &= \pi_{\boldsymbol{c}}\frac{S}{S + \psi_{\boldsymbol{c}}}\frac{S\exp(c_i r_i)}{S + \psi_{\boldsymbol{c}} + \sum_\ell S\exp(c_\ell r_\ell)}, && (41) \end{aligned}$$

where (37) is due to the fact that $\sigma_i(t) = \sigma_i(t)(a(t)+b(t))$. (38) holds, since $\lim_{t\to\infty}\mathbb{P}\{\sigma_i(t) = 1|\boldsymbol{c}, b(t) = 0\} = \frac{p_{i,\boldsymbol{c},a}}{\pi_{\boldsymbol{c}} - \sum_\ell p_{\ell,\boldsymbol{c},b}}$ and $\{X_{\boldsymbol{c}}(t)\}$ removes *P2*. (39) is induced from $b'(t) \ge b(t)$, and (40) is obtained from $\lim_{t\to\infty}\mathbb{P}\{X'_{\boldsymbol{c}}(t) = 1\} = \lim_{t\to\infty}\mathbb{P}\{b'(t) = 1|\boldsymbol{c}\} = \frac{\sum_\ell p_{\ell,\boldsymbol{c},b'}}{\pi_{\boldsymbol{c}}}$. We finally get (41) from (35) and (36). This completes the proof of (34).

B. *Proof of Theorem 5.2*

Let $\boldsymbol{\Lambda}^U$ denote the arrival rate region stabilized by U-CSMA. For any $\varepsilon > 0$ and arrival rate $\boldsymbol{\lambda} = [\lambda_i] \in (1-\varepsilon)\boldsymbol{\Lambda}^U$, there exist an U-CSMA algorithm with arbitrary small parameters $[R_i]$ and $[S_i]$, which stabilize arrival rate $(1+\varepsilon)\boldsymbol{\lambda}$, *i.e.*,

$$(1+\varepsilon)\lambda_i \le \sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1} \pi^*_{\boldsymbol{\sigma}},$$

where $[\pi^*_{\boldsymbol{\sigma}}]$ is the stationary distribution of Markov process $\{\boldsymbol{\sigma}(t)\}$ induced by the U-CSMA algorithm. In particular, given $\phi > 0$, one can assume $\max_i R_i \le \phi$. For the choice of $[R_i]$ and $[S_i]$, we consider an EXP-A-CSMA algorithm with functions

$$f_i(x) = R_i \quad \text{and} \quad g_i(x) = R_i\exp(-r_i x),$$

where we choose r_i to satisfy $S_i = \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} R_i \exp(-r_i \cdot c_i)$. Note that r_i satisfying the above equality always exists for given S_i, and $\max_{i\in V, x\in[0,1]} f_i(x) = \max_i R_i \le \phi$. Furthermore, one can observe that the maximum value of $f_i(x)$ and $g_i(x)$ for $x \in [0,1]$ can be made arbitrarily small due to arbitrarily small R_i, S_i. Using this observation and the Markov chain tree theorem (as we did for the proof of Lemma 3.2), one can show that

$$\max_{(\boldsymbol{\sigma},\boldsymbol{c})\in\mathcal{I}(G)\times\mathcal{H}^n}\left|1-\frac{\pi_{\boldsymbol{\sigma},\boldsymbol{c}}}{\pi_{\boldsymbol{c}}\pi^*_{\boldsymbol{\sigma}}}\right|<\varepsilon,$$

where $[\pi_{\boldsymbol{\sigma},\boldsymbol{c}}]$ denotes the stationary distribution of Markov process $\{(\boldsymbol{\sigma}(t),\boldsymbol{c}(t))\}$ by the EXP-A-CSMA algorithm. Therefore, it follows that

$$\begin{aligned}\lambda_i &\le \left(1-\frac{\varepsilon}{1+\varepsilon}\right)\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1}\pi^*_{\boldsymbol{\sigma}}\\ &< \sum_{\boldsymbol{c}\in\mathcal{H}^n}\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1} c_i\pi_{\boldsymbol{\sigma},\boldsymbol{c}}\\ &= \lim_{t\to\infty}\frac{1}{t}\widehat{D}_i(t),\end{aligned}$$

where the last inequality is from the ergodicity of Markov process $\{(\boldsymbol{\sigma}(t),\boldsymbol{c}(t))\}$. Due to Lemma 2.1, this means that the EXP-A-CSMA algorithm is rate-stable for the arrival rate $\boldsymbol{\lambda}$. This completes the proof of Theorem 5.3.

C. Proof of Theorem 5.3

The main strategy for the proof of Theorem 5.3 is that we study U-CSMA (channel-unaware CSMA) to achieve the performance guarantee of A-CSMA. We start by stating the following key lemmas about U-CSMA.

Lemma 5.1: Let $P_I(G)$ be the independent-set polytope,

$$P_I(G)=\left\{\boldsymbol{x}\in[0,1]^n:\boldsymbol{x}<\sum_{\boldsymbol{\rho}\in\mathcal{I}(G)}\alpha_{\boldsymbol{\rho}}\boldsymbol{\rho},\right.\\ \left.\sum_{\boldsymbol{\rho}\in\mathcal{I}(G)}\alpha_{\boldsymbol{\rho}}=1,\alpha_{\boldsymbol{\rho}}\ge 0 \text{ for all } \boldsymbol{\rho}\in\mathcal{I}(G)\right\} \quad (42)$$

Then, for $\boldsymbol{\lambda}\in P_I(G)$, there exists a U-CSMA algorithm with parameters $\boldsymbol{R}=[R_i]$ and $\boldsymbol{S}=[S_i]$ such that

$$\lim_{t\to\infty}\mathbb{E}[\boldsymbol{\sigma}(t)]>\boldsymbol{\lambda}.$$

Proof: The proof of Lemma 8 in [5] goes through for the proof of Lemma 5.1 in an identical manner. We omit further details.

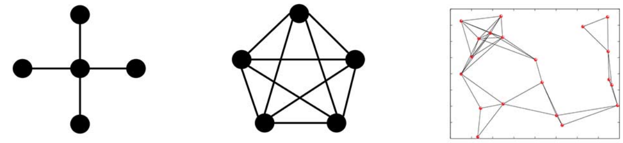
In what follows, we show that, for any $\phi>0$, interference graph G, channel transition-rate $\boldsymbol{\gamma}$ and arrival rate $\boldsymbol{\lambda}\in\alpha\boldsymbol{\Lambda}^o$, there exists a rate-stable U-CSMA algorithm with parameters $\boldsymbol{R}=[R_i]$ and $\boldsymbol{S}=[S_i]$ where α is defined in (33). Then, by Theorem 5.2, we can conclude this theorem.

To this end, it suffices to show that there exists a U-CSMA algorithm stabilizing any arrival rate $\boldsymbol{\lambda}$ such that

$$\boldsymbol{\lambda}\in\frac{1}{\chi(G)}\cdot\boldsymbol{\Lambda}^o \quad\text{or}\quad \boldsymbol{\lambda}\in\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}\cdot\boldsymbol{\Lambda}^o.$$

First, consider $\boldsymbol{\lambda}\in\frac{1}{\chi(G)}\cdot\boldsymbol{\Lambda}^o$. From Lemma 2.1 and the ergodicity of Markov process $\{\boldsymbol{\sigma}(t)\}$ and $\{\boldsymbol{c}(t)\}$ under U-CSMA, it suffices to prove that there exists a U-CSMA algorithm satisfying

$$\lim_{t\to\infty}\mathbb{E}[\sigma_i(t)c_i(t)]>\lambda_i \quad \text{for all } i\in V.$$



(a) (b) (c)

Fig. 2. Topologies. (a) Star, (b) Complete, (c) Random.

Since $\{\boldsymbol{\sigma}(t)\}$ and $\{\boldsymbol{c}(t)\}$ are independent, the above condition is equivalent to

$$\lim_{t\to\infty}\mathbb{E}[\sigma_i(t)]>\lim_{t\to\infty}\frac{\lambda_i}{\mathbb{E}[c_i(t)]} \quad\text{for all } i\in V.$$

Since $\chi(G)\cdot\lambda_i<\lim_{t\to\infty}\mathbb{E}[c_i(t)]=\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}$ (otherwise, $\chi(G)\boldsymbol{\lambda}\notin\boldsymbol{\Lambda}^o$), it is enough to prove that for an appropriately defined $\delta>0$,

$$\lim_{t\to\infty}\mathbb{E}[\sigma_i(t)]>\frac{1}{\chi(G)}-\delta \quad\text{for all } i\in V.$$

There exists a U-CSMA algorithm with parameter $\boldsymbol{R}=[R_i]$ and $\boldsymbol{S}=[S_i]$ satisfying the above inequality from Lemma 5.1 and $\left[\frac{1}{\chi(G)}-\delta\right]\in P_I(G)$, since from the definition of chromatic number there exists subset $\mathcal{I}^{(c)}(G)\subset\mathcal{I}(G)$ such that $|\mathcal{I}^{(c)}(G)|=\chi(G)$ and $\sum_{\boldsymbol{\sigma}\in\mathcal{I}^{(c)}(G)}\sigma_i=1$ for all i. Furthermore, we can make R_i and S_i arbitrarily small since $\lim_{t\to\infty}\mathbb{E}[\sigma_i(t)]$ under U-CSMA is invariant as long as ratios R_i/S_i remain same.

Now the second case $\boldsymbol{\lambda}\in\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}\frac{c_i\pi_{\boldsymbol{c}}}{c_{\max}}\cdot\boldsymbol{\Lambda}^o$ can be proved in an similar manner, where we have to prove that there exists a U-CSMA algorithm satisfying

$$\lim_{t\to\infty}\mathbb{E}[\sigma_i(t)]>\frac{\lambda_i}{\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}} \quad\text{for all } i\in V,$$

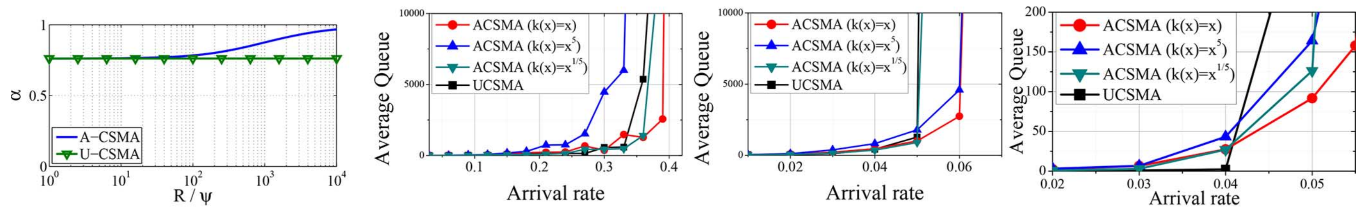
since $\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}\le\lim_{t\to\infty}\mathbb{E}[c_i(t)]$ for all $i\in V$. By Lemma 5.1, it is sufficient to show that $\frac{\boldsymbol{\lambda}}{\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}}\in P_I(G)$ for all $\frac{\boldsymbol{\lambda}}{\min_{i\in V}\sum_{\boldsymbol{c}\in\mathcal{H}^n}c_i\pi_{\boldsymbol{c}}}\in\frac{1}{c_{\max}}\cdot\boldsymbol{\Lambda}^o$, which is trivial since $\boldsymbol{\Lambda}^o\subset c_{\max}\cdot P_I(G)$ from the definitions of $\boldsymbol{\Lambda}^o$ and $P_I(G)$. This is the end of this proof.

VI. Simulation Results

In this section, we provide simulation results to demonstrate our analytical findings.

U-CSMA vs. A-CSMA. We first consider a 5-link complete interference graph, *i.e.*, all 5 links interfere with each other. All links are homogeneous and independent in terms of time-varying channels, where we assume that the channel space is simply $\{0.5,1\}$. With an abuse of notation, let $\gamma_i^{a\to b}$ denote the transition-rate for $c_i(t)$ which translates from state a to state b. We consider the transition-rate $\gamma=\gamma_i^{0.5\to 1}=\gamma_i^{1\to 0.5}$. We compare A-CSMA and U-CSMA, with the following setups:

$$\begin{aligned}&\text{A}-\text{CSMA}: f_i(x)=R,\ g_i(x)=R\cdot 10^{-4x}\\ &\text{U}-\text{CSMA}: f_i(x)=R,\ g_i(x)=R\cdot 10^{-4},\end{aligned}$$



(a) (b) (c) (d)

Fig. 3. Numerical results on A-CSMA. (a) 5-link complete graph, (b) Star: A-CSMA vs U-CSMA, (c) Complete: A-CSMA vs U-CSMA, (d) Random: A-CSMA vs U-CSMA.

so that $\log(f_i/g_i) = 4x\log(10)$ for A-CSMA and $4\log(10)$ for U-CSMA, respectively. Throughputs of A-CSMA and U-CSMA are evaluated by estimating the average rate in the potential departure process, *i.e.*, $\lim_{t\to\infty}\frac{1}{t}\widehat{\boldsymbol{D}}(t)$. Fig. 3(a) shows the results, where in x-axis, we vary the ratio of backoff rate R to the channel varying speed ψ (determined by γ) and y-axis represents the fraction of achievable rate region α (note that in a complete interference graph, the rate region is symmetric). Since the maximum rate point in $\boldsymbol{\Lambda}$ is $\lambda_i = \frac{1}{5}((1-(1/2)^5)+(1/2)^6)$ for all $i \in V$, in Fig. 3(a), $\alpha = \lim_{t\to\infty}\frac{5\widehat{\boldsymbol{D}}(t)}{((1-(1/2)^5)+(1/2)^6)t}$. We observe that (i) by reflecting the channel capacity in the CSMA parameters as an exponential function, A-CSMA has α-throughput where α approaches 100% (see Theorem 3.1), and (ii) U-CSMA has 76%-throughput. Note that $\alpha \geq 76\%$ even with limited backoff rates (*i.e.*, small R/ψ), and this matches Corollary 5.2 which states that A-CSMA's throughput is at least U-CSMA's throughput.

Linear vs. non-linear adaptation. Here, we test *dynamic* A-CSMA and U-CSMA algorithms. Let A-CSMA$(k(x))$ denote the EXP(k)-A-CSMA with queue-based algorithms. We test $k(x)=x$, $k(x)=x^5$ and $k(x)=x^{1/5}$ to compare concave and convex rate adaptation with the linear adaptive function. Each link updates the functions $[f_i]$ and $[g_i]$ according to its own queue length as stated in Section IV-B except the channel adaptation function $k(\cdot)$. In terms of channels, each link has independent and identical channels, where $\mathcal{H}=\{\frac{u}{10}: 1\leq u\leq 10\}$. For each link i, $\gamma^{u/10\to(u+1)/10} = \gamma^{u/10\to(u-1)/10} = 0.01$, and 0 otherwise. We increase the arrival rates homogeneously across all links, and plot the average queue lengths to study which arrival rates start to make the system unstable across all the tested algorithms. The average queue length will blow up when the algorithm cannot stabilize the given arrival rate.

Star graph. Fig. 3(b) shows the results of the 5-link star graph, where the average queue blows up at 0.42 (A-CSMA(x)), 0.36 (A-CSMA(x^5)), and 0.39 (A-CSMA$(x^{1/5})$ and U-CSMA). Note that for star graphs, we should schedule the inner and outer links alternatively, depending on the channel conditions to stabilize the system. Due to the channel adaptivity of x and $x^{1/5}$, A-CSMA(x) and A-CSMA $(x^{1/5})$ outperform U-CSMA. A-CSMA(x^5), however, is even worse than U-CSMA. This is because x^5 is a convex adaptive function, and thus the sum of intensities of the outer links is smaller than that of the inner link when the channel capacity of the inner link is the same with the sum of outer links' capacities. Therefore, outer links tend to loose transmission chances even for good channel conditions.

Complete graph. Fig. 3(c) shows the results of the 10-link complete graph, where the average queue blows up at 0.07 (A-CSMA(x), A-CSMA(x^5)), and 0.06 (A-CSMA$(x^{1/5})$ and U-CSMA). To understand this trend, note that for complete graphs, to achieve throughput optimality, we have to schedule the link having the largest channel capacity at each channel state. A-CSMA(x) and A-CSMA(x^5) outperforms U-CSMA because the transmission chance increases as the channel capacity increases. However, A-CSMA$(x^{1/5})$ performs similarly with U-CSMA, because with the concave function x^5, the transmission intensity of a link changes less aggressively, so small difference from U-CSMA's performance. The results for star and complete graphs validates the uniqueness of our analysis on how the system should adapt to channel variations, irrespective of interference graphs (see Theorem 3.2).

From Theorem 4.2, we know that A-CSMA(x) is throughput optimal. The results, however, are not tight with the maximum rate point in $\boldsymbol{\Lambda}$, which is

$$\lambda_i = \frac{1}{10}\sum_{k=1}^{10}\frac{k}{10}\left((\frac{k}{10})^{10}-(\frac{k-1}{10})^{10}\right) = 0.095,\ \forall i \in V,$$

whereas the average queue blows up at 0.07. This stems from the backoff rate and channel-varying speed. In Theorem 5.1, the backoff rate should exponentially grow up as the arrival rate closes to the maximum rate point. Thus, we need large enough $Q_{\max}$ in (32) to stabilize the arrivals with nearly maximum rate.

Random graph. We now study A-CSMA and U-CSMA for a random topology by uniformly locating 20 nodes in a square area and a link between two nodes are established by a given transmission range, as depicted in Fig. 2(c). To model interference, we assume the two-hop interference model (*i.e.*, any two links within two hops interfere) as in 802.11. Fig. 3(d) shows that again A-CSMA(x) outperforms others.

VII. Conclusion

Recently, it is shown that CSMA algorithms can achieve throughput (or utility) optimality where 'static' channel is assumed. However, in practice, the channel capacities are typically time-varying. In this paper, we study a generic criteria of throughput-optimal CSMA algorithms in the time-varying scenarios, and propose the A-CSMA algorithm by exploring certain sufficient conditions, *e.g.*, high back-off rates. The tight analysis of throughputs of CSMA algorithms with general backoff rates is an interesting open question in the future research.

References

[1] V. Anantharam and P. Tsoucas, "A proof of the Markov chain tree theorem," *Statist. Probab. Lett.*, vol. 8, no. 2, pp. 189–192, 1989.

[2] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Netw.*, vol. 1, no. 1, pp. 1–149, 2006.

[3] T. P. Hayes and A. Sinclair, "A general lower bound for mixing of single-site dynamics on graphs," *Ann. Appl. Probab.*, vol. 17, no. 3, pp. 931–952, 2007.

[4] L. Jiang, M. Leconte, J. Ni, R. Srikant, and J. Walrand, "Fast mixing of parallel Glauber dynamics and low-delay CSMA scheduling," in *Proc. IEEE INFOCOM*, 2011, pp. 371–375.

[5] L. Jiang, D. Shah, J. Shin, and J. Walrand, "Distributed random access algorithm: Scheduling and congestion control," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6182–6207, Dec. 2010.

[6] L. Jiang and J. Walrand, "A distributed CSMA algorithm for throughput and utility maximization in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 960–972, Jun. 2010.

[7] L. Jiang and J. Walrand*, Scheduling and Congestion Control for Wireless and Processing Networks*. San Francisco, CA, USA: Morgan & Claypool, 2010.

[8] L. Jiang and J. Walrand, "Approaching throughput-optimality in distributed CSMA scheduling algorithms with collisions," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 816–829, Jun. 2011.

[9] C. Joo, X. Lin, J. Ryu, and N. B. Shroff, "Distributed greedy approximation to maximum weighted independent set for scheduling with fading channels," in *Proc. ACM MobiHoc*, 2013, pp. 89–98.

[10] T. H. Kim, J. Ni, R. Srikant, and N. Vaidya, "On the achievable throughput of CSMA under imperfect carrier sensing," in *Proc. IEEE INFOCOM*, 2011, pp. 1674–1682.

[11] C.-H. Lee, D. Y. Eun, S.-Y. Yun, and Y. Yi, "From Glauber dynamics to Metropolis algorithm: Smaller delay in optimal CSMA," in *Proc. IEEE ISIT*, Jun. 2012, pp. 2681–2685.

[12] H. Lee, S. Moon, and Y. Yi, "A-DCF: Design and implementation of delay and queue length based wireless MAC," in *Proc. IEEE INFOCOM*, 2015, pp. 2173–2181.

[13] J. Lee, H. Lee, Y. Yi, S. Chong, B. Nardelli, and M. Chiang, "Making 802.11 DCF near-optimal: Design, implementation, and evaluation," in *Proc. IEEE SECON*, 2013, pp. 86–94.

[14] D. A. Levin, Y. Peres, and E. L. Wilmer*, Markov Chains and Mixing Times*. Providence, RI, USA: AMS, 2009.

[15] B. Li and A. Eryilmaz, "A Fast-CSMA algorithm for deadline-constrained scheduling over wireless fading channels," ArXiv e-prints, Mar. 2012.

[16] B. Li and A. Eryilmaz, "Optimal distributed scheduling under time-varying conditions: A Fast-CSMA algorithm with applications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3278–3288, Jul. 2013.

[17] J. Liu, Y. Yi, A. Proutiere, M. Chiang, and H. V. Poor, "Towards utility-optimal random access without message passing," *J. Wireless Commun. Mobile Comput.*, vol. 10, no. 1, pp. 115–128, Jan. 2010.

[18] M. Lotfinezhad and P. Marbach, "Throughput-optimal random access with order-optimal delay," in *Proc. IEEE INFOCOM*, 2011, pp. 2867–2875.

[19] B. Nardelli *et al.*, "Experimental evaluation of optimal CSMA," in *Proc. IEEE INFOCOM*, 2011, pp. 1188–1196.

[20] M. J. Neely*, Stochastic Network Optimization With Application to Communication and Queueing Systems*. San Francisco, CA, USA: Morgan & Claypool, 2010.

[21] J. Ni, B. Tan, and R. Srikant, "Q-CSMA: Queue-length based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–5.

[22] A. Proutiere, Y. Yi, T. Lan, and M. Chiang, "Resource allocation over network dynamics without timescale separation," in *Proc. IEEE INFOCOM*, 2010, pp. 1–5.

[23] D. Qian, D. Zheng, J. Zhang, and N. Shroff, "CSMA-based distributed scheduling in multi-hop MIMO networks under SINR model," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.

[24] S. Rajagopalan, D. Shah, and J. Shin, "Network adiabatic theorem: An efficient randomized protocol for contention resolution," in *Proc. ACM SIGMETRICS*, 2009, pp. 133–144.

[25] D. Shah and J. Shin, "Delay optimal queue-based CSMA," in *Proc. ACM SIGMETRICS*, 2010, pp. 373–374.

[26] D. Shah and J. Shin, "Randomized scheduling algorithm for queueing networks," *Ann. Appl. Probab.*, vol. 22, pp. 128–171, 2012.

[27] D. Shah, J. Shin, and P. Tetali, "Medium access using queues," in *Proc. IEEE FOCS*, 2011, pp. 698–707.

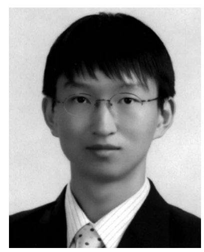
[28] L. Tassiulas, "Scheduling and performance limits of networks with constantly changing topology," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 1067–1073, May 1997.

[29] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936–1949, Dec. 1992.

[30] Y. Yi and M. Chiang", "Chapter 9: Stochastic network utility maximization and wireless scheduling," in *Next-Generation Internet Architectures and Protocols*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

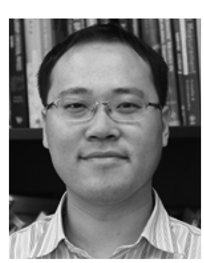
[31] S.-Y. Yun, J. Shin, and Y. Yi, "CSMA using the Bethe approximation for utility maximization," in *Proc. IEEE ISIT*, 2013, pp. 206–210.

[32] S.-Y. Yun, Y. Yi, J. Shin, and D. Y. Eun, "Optimal CSMA: A survey," in *Proc. IEEE ICCS*, 2012, pp. 199–204.



Se-Young Yun is currently a Post-Doc Researcher with the MSR-INRIA Joint Center since April 2014. Before joining the MSR-INRIA Joint Center, he spent one year with KTH, Stockholm, Sweden, as a Post-Doc Researcher. He received the B.S. and Ph.D. degrees in electrical engineering from the KAIST, Daejeon, Korea, in 2006 and 2012, respectively. From 2012 to 2013, he was a Post-Doc Researcher with the same department. His research interests include social networks, data mining, network economics, optimality on CSMA, future Internet network, and network science. He received the Best Paper Award at ACM MobiHoc 2013.

Jinwoo Shin is currently an Assistant Professor with the Department of Electrical Engineering, KAIST, Korea. He obtained the B.S. degrees in computer science and mathematics from Seoul National University in 2001 and the Ph.D. degree in mathematics from the Massachusetts Institute of Technology in 2010. After spending two years with the Algorithms and Randomness Center, Georgia Institute of Technology, and one year (2012–2013) with the Business Analytics and Mathematical Sciences Department and IBM T. J. Watson Research, he joined the KAIST department in Fall 2013. He received the Best Student Paper Award at ACM SIGMETRICS 2009, the Best MIT CS Doctoral Thesis (George M. Sprowls) Award 2010, the Best Paper Award at ACM MobiHoc 2013, the Best Publication Award from INFORMS Applied Probability Society 2013, the Bloomberg L. P. research gift award 2015, and the ACM SIGMETRICS Rising Star Award 2015.



Yung Yi received the B.S. and the M.S. degrees from the School of Computer Science and Engineering, Seoul National University, Korea, in 1997 and 1999, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Texas at Austin, in 2006. From 2006 to 2008, he was a Post-Doctoral Research Associate with the Department of Electrical Engineering, Princeton University. Now, he is an Associate Professor with the Department of Electrical Engineering, KAIST, Korea. His current research interests include the design and analysis of computer networking and wireless communication systems, especially congestion control, scheduling, and interference management, with applications in wireless ad hoc networks, broadband access networks, economic aspects of communication networks, and green networking systems. He received the best paper awards at IEEE SECON 2013 and ACM MobiHoc 2013. He is now an Associate Editor of the IEEE/ACM Transactions on Networking.