CSMA over Time-varying Channels: Optimality, Uniqueness and Limited Backoff Rate

Se-Young Yun, Jinwoo Shin, and Yung Yi

***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 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 highly 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.**

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 [24]. 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 efficiency 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], [26] for surveys.

Recently, there has been exciting progresses 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], [13], [16], [19]. 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 [29] for more details.

However, the recent CSMA algorithms crucially rely on the assumption of static channel conditions. 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 [23] 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. A CSMA schedule at some instant may not have enough time to be close to the desired 'stationary' distribution before the channel changes. 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. 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 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), 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 5.1). Our bound depends on a combinatorial property of the underlying interference graph (*i.e.*, its chromatic number), 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 Corollary 5.1).

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 induced by it [6], [13], [16], [19]. 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 of 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. [5], [6]) and Rajagopalan et al. (cf. [19], [21]), 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.[1] ***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. Furthermore, one can design a dynamic high-throughput A-CSMA algorithm with limited backoff rates using ***C3*** (similarly as ***C1*** is used for ***C2***), but in the current paper we do not present further details due to the space limitation.

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

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. [19], [21]), 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. showed that if the CSMA parameters are changing very slowly with respect to the queue length changes, the mixing time is much faster than the queue length changes so that the realized link schedules can provably emulate Max-Weight very well. Although their key intuitions are apparently different, analytic techniques are quite similar, *i.e.*, both require to understand the long-term behavior (*i.e.* stationarity) of the Markov chains formed by CSMA.

These throughput optimality results motivate further research on design and analysis of CSMA algorithms. The work by Liu et al. [13] 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 [17]. The throughput optimality of MIMO networks under SINR model is also shown in [18]. As opposed to the continuous-time setting that 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], [22] and [9], where the authors in [9] consider imperfect sensing information. In [13], the authors studied the impact of collisions and the tradeoff between short-term fairness and efficiency. The authors in [16] considered a synchronous system consisting of the control phase, which eliminates the chances of data collisions via a simple 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], [10], [12], [14], [20], speeding up the convergence [28], and developing a practical protocol based on the CSMA theory with experimental validation [11], [15].

To the best of our knowledge, CSMA under time-varying channels has been studied only in [12] for only complete interference graphs, when the arbitrary backoff rate is allowed, and more seriously, under the time-scale separation assumption, which does not often hold in practice and extremely simplifies the analysis (no mixing time related details are needed).

II. Model and Preliminaries

A. Network Model

We consider a network consisting of a collection of n queues (or links) $\{1, \ldots, n\}$ and time is indexed by $t \in \mathbb{R}_+$. Let $Q_i(t) \in \mathbb{R}_+$ denote the amount of work in queue i at time t and let $\boldsymbol{Q}(t) = [Q_i(t)]_{1 \le i \le n}$. The system starts empty, *i.e.*, $Q_i(0) = 0$. We assume work arrives at each queue i as per an exogenous Poisson process with rate $\lambda_i > 0$, where $A_i(s,t) < \infty$ denotes the cumulative arrival to queue i in the time interval $(s, t]$. Each queue 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 finite state Markov time-varying channels [25]: each $\{\boldsymbol{c}(t) = [c_i(t)] : t \ge 0\}$ is a continuous-time, time-homogeneous, irreducible Markov process, where each link has m states channel space such that $c_i(t) \in \mathcal{H} := \{h_1, \ldots, h_m\}$ and $0 < h_1 < \cdots < h_m = 1$. We denoted 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.[2] We call $\max_{\boldsymbol{u} \in \mathcal{H}^n} \{\sum_{\boldsymbol{v} \in \mathcal{H}^n : \boldsymbol{v} \neq \boldsymbol{u}} \gamma^{\boldsymbol{u} \to \boldsymbol{v}}\}$ the *channel varying speed.* The inverse of *channel varying speed* indicates the maximum of the expected number of channel transitions during the unit-length time interval. We consider only single-hop sessions (or flows), *i.e.*, once work departs from a queue, it leaves the network.

The queues 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 queues, where the vertices $V = \{1, \ldots, n\}$ represent queues and the edges $E \subset V \times V$ represent interferences between queues: $(i,j) \in E$, if queues 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 node i and a schedule at time t, *i.e.*, whether queues transmit at time t, respectively, where $\sigma_i(t) = 1$ represents transmission of queue i at time t. Then, interference imposes the constraint that for all $t \in \mathbb{R}_+$, $\boldsymbol{\sigma}(t) \in \mathcal{I}(G)$, where

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

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

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

[2] 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.

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.*,

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

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 $\boldsymbol{\lambda} = [\lambda_i] \in \boldsymbol{\Lambda} = \boldsymbol{\Lambda}(\boldsymbol{\gamma}, G)$, where

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

where $\boldsymbol{\lambda} \le \boldsymbol{\lambda}'$ corresponds to the component-wise inequality, *i.e.*, if $\boldsymbol{\lambda} \notin \boldsymbol{\Lambda}$, queues should grow linearly over time 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.

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

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

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.

Lemma 2.1: A scheduling algorithm is rate-stable if

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

We omit the proof due to the space constraint.

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 queue waits for a random amount of time – also known as *backoff time*. Each queue can sense the medium perfectly and instantly, *i.e.*, knows if any other interfering queue is transmitting at a given time instance. If a queue 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 queue 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:

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

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.

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

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

Proof: We prove this by contradiction. 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 and $\frac{f_i(h_u)}{g_i(h_u)} \neq \frac{f_i(h_v)}{g_i(h_v)}$ for some link i. From the reversibility, the transition path $(\boldsymbol{\sigma}_i^0, \boldsymbol{c}_i^u) \rightarrow (\boldsymbol{\sigma}_i^0, \boldsymbol{c}_i^v) \rightarrow (\boldsymbol{\sigma}_i^1, \boldsymbol{c}_i^v)$ has to satisfy the following balance equations:

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

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

$$\begin{aligned}
\pi_{\boldsymbol{\sigma}_i^0,\boldsymbol{c}_i^u} f_i(h_u) &= \pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^u} g_i(h_u), \ \text{ and} \\
\pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^u} \gamma^{\boldsymbol{c}_i^u\rightarrow\boldsymbol{c}_i^v} &= \pi_{\boldsymbol{\sigma}_i^1,\boldsymbol{c}_i^v} \gamma^{\boldsymbol{c}_i^v\rightarrow\boldsymbol{c}_i^u}. \qquad (4)
\end{aligned}$$

From (3) and (4),

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

which contradicts the assumption $\frac{f_i(h_u)}{g_i(h_u)} \neq \frac{f_i(h_v)}{g_i(h_v)}$. This completes the proof of Theorem 2.1. ■

We note that the non-reversible property makes it hard to characterize the stationary distribution $[\pi_{\boldsymbol{\sigma},\boldsymbol{c}}]$ of the Markov process induced by A-CSMA.

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:

$$\log \frac{f_i(x)}{g_i(x)} \ = \ r_i \cdot x, \quad \text{for } x \in [0,1], \qquad (6)$$

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

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 continuous function $k : [0,1] \rightarrow \mathbb{R}_+$, we define EXP($k$)-A-CSMA as the sub-class of A-CSMA algorithms with the following form of functions:

$$\log \frac{f_i(x)}{g_i(x)} \ = \ r_i \cdot k(x), \quad \text{for } x \in [0,1], \qquad (7)$$

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

$$\text{EXP}(k)\text{-A-CSMA} \ = \ \text{EXP-A-CSMA}.$$

The proofs of Theorems 3.1 and 3.2 are given in Sections III-A and III-C, respectively. For the proof of Theorem 3.1, Section III-B describes the proof of Lemma 3.2, which is a key lemma of this work. In the following proofs (and throughout this paper), we commonly let $[\pi_{\boldsymbol{\sigma},\boldsymbol{c}}]$, $[\pi_{\boldsymbol{c}}]$ and $[\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}]$ be the stationary distributions of Markov processes $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$, $\{\boldsymbol{c}(t)\}$ and $\{\boldsymbol{\sigma}(t), \boldsymbol{c}\}$ induced by an A-CSMA algorithm, respectively.

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

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

and every EXP-A-CSMA algorithm with

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

satisfies

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

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

$$\min_{i\in V, h\in\mathcal{H}}\{f_i(h), g_i(h)\} \ge \frac{\psi \cdot m^{2^n m^n (n+1)}}{\delta_2},$$

it follows that

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

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.

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]. Due to the space constraint, we omit the proof which can be found in [27].

Proof of Theorem 3.1. We now complete the proof of Theorem 3.1 using Lemmas 3.1 and 3.2. For a given arrival rate $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o$, there exists $\varepsilon \in (0,1)$ such that $\boldsymbol{\lambda} \in (1-\varepsilon)\boldsymbol{\Lambda}^o$ since $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o$. 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

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

where we choose

$$f_i(c_i) = R = \eta\exp(\kappa), \quad g_i(c_i) = R\cdot\exp(-r_i\cdot c_i),$$

$$\kappa = \kappa(\delta_1, G, \boldsymbol{\gamma}) := \frac{4n^2\log|\mathcal{I}(G)|}{\delta_1^2 \min_i\{\left(\sum_{\boldsymbol{c}\in\mathcal{H}^n} c_i\pi_{\boldsymbol{c}}\right)^2\}},$$

and

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

Therefore, it follows that

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

where the last inequality is from the ergodicity of Markov process $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$. This leads to the rate-stability using Lemma 2.1, and hence completes the proof.

B. 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 $\mathcal{V}\backslash\{(\boldsymbol{\sigma}, \boldsymbol{c})\}$, 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, Markov chain tree theorem [1] implies that

$$\pi_{\boldsymbol{\sigma},\boldsymbol{c}} = \frac{w(\mathcal{A}_{\boldsymbol{\sigma},\boldsymbol{c}})}{\sum_{(\boldsymbol{\rho},\boldsymbol{d})\in\mathcal{I}(G)\times\mathcal{H}^n} w(\mathcal{A}_{\boldsymbol{\rho},\boldsymbol{d}})}. \tag{8}$$

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. Then, we have

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

where (a) is from the condition in Lemma 3.2 and for (b) 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 (8) and

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

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.

C. Proof of Theorem 3.2

Consider a star interference graph G, where 1 denotes the center vertex and $\{2,\dots,n\}$ is the set of other outer vertices. For the time-varying channel model, we set each element of channel by $h_j = \frac{j}{m}$ and assume the channel transition satisfies $\pi_{\boldsymbol{c}} = \frac{1}{m^n}$. For the arrival rate, we choose $\boldsymbol{\lambda} = \arg\max\limits_{\boldsymbol{\lambda}\in(1-\varepsilon)\boldsymbol{C}} \sum_i \lambda_i$, where $\varepsilon \in (0,1)$ will be chosen later.

Under this setup, suppose the conclusion of Theorem 3.1 holds, *i.e.*, there exists a rate-stable EXP(k)-A-CSMA. Then, from the ergodicity of Markov process $\{(\boldsymbol{\sigma}(t), \boldsymbol{c}(t))\}$, we have

$$
\lambda_i \;\leq\; \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} c_i\sigma_i\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}, \qquad \text{for all } i\in V. \tag{9}
$$

Taking the summation over $i \in V$ in both sides of the above inequality and using $\boldsymbol{\lambda} = \arg\max\limits_{\boldsymbol{\lambda}\in(1-\varepsilon)\boldsymbol{C}} \sum_i \lambda_i$, it follows that

$$
\begin{aligned}
\sum_i \lambda_i &= (1-\varepsilon) \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i \\
&\leq \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \pi_{\boldsymbol{\sigma}|\boldsymbol{c}} \sum_i c_i\sigma_i.
\end{aligned}
$$

By rearranging terms in the above inequality, we have

$$
\begin{aligned}
\frac{\varepsilon}{1-\varepsilon}\sum_i \lambda_i &\geq \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \left(\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i - \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \pi_{\boldsymbol{\sigma}|\boldsymbol{c}} \sum_i c_i\sigma_i\right) \\
&= \sum_{\boldsymbol{c}\in\mathcal{H}^n} \pi_{\boldsymbol{c}} \cdot E\left[\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i - \sum_i c_i\sigma_i\right],
\end{aligned}
$$

where the expectation is taken with respect to random variable $\boldsymbol{\sigma} = [\sigma_i]$ of which distribution is $[\pi_{\boldsymbol{\sigma}|\boldsymbol{c}}]$. Since we know $\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i - \sum_i c_i\sigma_i \geq 0$ with probability 1, we further have that for all channel state $\boldsymbol{c}$,

$$
\begin{aligned}
E\left[\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i - \sum_i c_i\sigma_i\right] &\leq \frac{\sum_i \lambda_i^{\max} \cdot \frac{\varepsilon}{1-\varepsilon}}{\pi_{\boldsymbol{c}}} \\
&\leq \frac{n\cdot\frac{\varepsilon}{1-\varepsilon}}{\pi_{\boldsymbol{c}}} = n\cdot m^n \cdot \frac{\varepsilon}{1-\varepsilon}.
\end{aligned}
$$

Markov's inequality implies that

$$
\Pr\left[\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i - \sum_i c_i\sigma_i \geq \frac{1}{m}\right] \;\leq\; n\cdot m^{n+1}\cdot\frac{\varepsilon}{1-\varepsilon}.
$$

If we choose $\varepsilon = \frac{1}{4n\cdot m^{n+1}}$, then

$$
\Pr\left[\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i c_i\rho_i - \sum_i c_i\sigma_i \geq \frac{1}{m}\right] \;<\; \frac{1}{2}. \tag{10}
$$

In the star graph G, $\max_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \sigma_i c_i$ is c_1 or $\sum_{j=2}^n c_j$ and under the channel model, if $c_1 \neq \sum_{j=2}^n c_j$, $\left|c_1 - \sum_{j=2}^n c_j\right| \geq \frac{1}{m}$. If $c_1 > \sum_{i=2}^n c_i$, r_i Thus, from (10), if $c_1 > \sum_{i=2}^n c_i$,$P[\sigma_1 = 1] > \frac{1}{2}$ and if $c_1 < \sum_{i=2}^n c_i$, $P[\sigma_1 = 1] < \frac{1}{2}$, which implies that $r_1 k(c_1) \geq \sum_{i=2}^n r_i k(c_i)$ and $r_1 k(c_1) \leq \sum_{i=2}^n r_i k(c_i)$, respectively. Therefore, for every channel state $\boldsymbol{c}$ with $0 < \sum_{i=2}^n c_i < 1$,

$$
r_1 k\!\left(\sum_{i=2}^n c_i + \frac{1}{m}\right) > \sum_{i=2}^n r_i k(c_i) > r_1 k\!\left(\sum_{i=2}^n c_i - \frac{1}{m}\right), \tag{11}
$$

which implies that $r_i k(x)$ is a strictly increasing function. In addition, $r_1 k(c_1) > 0$ from (10), because $\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G)} \sigma_1\pi_{\boldsymbol{\sigma}|\boldsymbol{c}} > \pi_{\boldsymbol{0}|\boldsymbol{c}}$ when $c_1 > \sum_{i=2}^n c_i$. Since $k(x)$ is non-negative and $r_i k(x)$ is strict increasing for all link i, $r_i > 0$. Thus, when we devide both sides of (11) by r_i,

$$
k\left(\sum_{i=2}^n c_i + \frac{1}{m}\right) > \sum_{i=2}^n \frac{r_i}{r_1}\cdot k(c_i) > k\left(\sum_{i=2}^n c_i - \frac{1}{m}\right). \tag{12}
$$

By choosing $x = c_2 = \cdots = c_n$ and taking $m \to \infty$ in (12), it follows that $\lim_{m\to\infty} \sum\limits_{i=2}^n \frac{r_i}{r_1}$ exists,[3] and for any $0 < x < 1/n$,

$$
k\left((n-1)x\right) = \lim_{m\to\infty} \sum_{i=2}^n \frac{r_i}{r_1}\cdot k(x),
$$

where $\lim_{m\to\infty} \sum\limits_{i=2}^n \frac{r_i}{r_1} > 1$ since $k(x)$ is strictly increasing. Hence, if we take $x \to 0$ in the above inequality, $k(0) = 0$

[3] Recall that $[r_i]$ is a function of m.

follows. Similarly, by choosing $x = c_2$, $c_3 = \cdots = c_n = 1/m$ and taking $m \to \infty$ in (12), it follows that that for any $0 < x < 1$,

$$k(x) = \lim_{m\to\infty} \frac{r_2}{r_1} \cdot k(x),$$

where we use $k(0) = 0$ and $\limsup_{m\to\infty} \frac{r_i}{r_1} < \infty$ due to the existence of $\lim_{m\to\infty} \sum_{i=2}^{n} \frac{r_i}{r_1}$. Thus, $\lim_{m\to\infty} \frac{r_2}{r_1} = 1$, and more generally, $\lim_{m\to\infty} \frac{r_i}{r_1} = 1$ using same arguments. Furthermore, by choosing $x = c_2$, $y = c_3$, $c_4 = \cdots = c_n = 1/m$, and taking $m \to \infty$ in (12), we have that for any $0 < x + y < 1$,

$$k(x + y) = k(x) + k(y),$$

where we use $\lim_{m\to\infty} \frac{r_2}{r_1} = 1$. This implies that $k(x)$ is a linear function (with $k(0) = 0$), and hence the conclusion of Theorem 3.2 follows.

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 [21]).

A. Rate-based Algorithm

The first algorithm, at each queue i, updates $(f_i^{(t)}, g_i^{(t)})$ at time instances $L(j), j \in \mathbb{Z}_+$ with $L(0) = 0$, and $(f_i^{(t)}, g_i^{(t)})$ remains fixed between in the time-interval $[L(j), L(j+1))$ for all $j \in \mathbb{Z}_+$, where we define $T(j) = L(j+1) - L(j)$ for $j \geq 0$. With an abuse of notation, $f_i^{(j)}, g_i^{(j)}$ denotes 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, queue 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 queue i in $[L(j), L(j+1))$, *i.e.*,

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

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

$$\begin{aligned} g_i^{(j+1)}(x) &= \frac{j+2}{j+1} \cdot g_i^{(j)}(x) \cdot \exp\left(x \cdot \alpha(j) \cdot (\hat{s}_i(j) - \hat{\lambda}_i(j))\right) \\ f_i^{(j+1)}(x) &= j+2, \end{aligned} \tag{13}$$

with initial condition $f_i^{(0)}(x) = g_i^{(0)}(x) = 1$. It is easy to check that the A-CSMA algorithm with functions $[f_i^{(j)}]$ and $[g_i^{(j)}]$ lies in EXP-A-CSMA:

$$\log \frac{f_i^{(j)}(x)}{g_i^{(j)}(x)} = r_i(j) \cdot x,$$

where $r_i(0) = 0$ and

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

Note that, under this update rule, the algorithm at each queue 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, channel transition-rate $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o(\boldsymbol{\gamma}, G)$, the A-CSMA algorithm with updating functions as per (13) is rate-stable.

The proof of Theorem 4.1 uses the same strategy in [5] in conjunction with Lemma 3.1 and Lemma 3.2. Due to the space limitation, we present the proof in [27].

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.

$$f_i^{(t)}(x) = \left(g_i^{(t)}(x)\right)^2 \quad \text{and}$$
$$g_i^{(t)}(x) = \exp\left(x \cdot \max\left\{w(Q_i(\lfloor t \rfloor)), \sqrt{w(Q_{\max}(\lfloor t \rfloor))}\right\}\right), \tag{14}$$

where $Q_{\max}(\lfloor t \rfloor) = \max_j Q_j(\lfloor t \rfloor)$ and $w(x) = \log\log(x+e)$. One can interpret this as an EXP-A-CSMA algorithm since

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

where $r_i(t) = \max\left\{w(Q_i(\lfloor t \rfloor)), \sqrt{w(Q_{\max}(\lfloor t \rfloor))}\right\}$. The global information of $Q_{\max}(\lfloor t \rfloor)$ can be replaced by its approximate estimation that can computed through a very simple distributed algorithm (with message-passing) in [19] or a learning mechanism (without message-passing) in [22]. 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, channel transition-rate $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o(\boldsymbol{\gamma}, G)$, the A-CSMA algorithm with functions as per (14) is rate-stable.

Due to the space constraint, we omit the proof of Theorem 4.2 which can be found in [27].

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

In practice, it might be hard to have arbitrary large backoff rate because of physical constraints. From this motivation, in this section, we 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, we cannot guarantee the optimality of EXP-A-CSMA. The main result of this section is that, even with highly limited backoff rate, say at most $\delta > 0$, EXP-A-CSMA is guaranteed to have at least α-throughput, where α is *independent* of the channel varying speed and the maximum backoff rate δ. More formally, we obtain the following result.

Theorem 5.1: 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

$$\alpha = \max\left\{\min_{i \in V} \sum_{\boldsymbol{c} \in \mathcal{H}^n} c_i \pi_{\boldsymbol{c}}, \frac{1}{\chi(G)}\right\}. \tag{15}$$

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

Theorem 5.1 implies that even with arbitrary small backoff rates, A-CSMA is guaranteed to achieve a partial fraction of the capacity region. For example, for a bipartite interference graph, at least 50%-throughput can be achieved since its chromatic number is two. The proof strategy is as follows: (i) We first find the achievable rate region of U-CSMA, and (ii) we then show that for any U-CSMA parameters, there exists a EXP-A-CSMA algorithm satisfying the backoff constraint and achieving ε-close departure rate with that by the U-CSMA (we formally state this in Corollary 5.1).

Corollary 5.1: For any $\phi > 0$, 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 $\max_{i \in V, x \in [0,1]} f_i(x) \le \phi$ and

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

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.

A. Proof of Theorem 5.1

The main strategy for the proof of Theorem 5.1 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, *i.e.*,

$$P_I(G) = \left\{\boldsymbol{x} \in [0,1]^n : \boldsymbol{x} = \sum_{\boldsymbol{\rho} \in \mathcal{I}(G)} \alpha_{\boldsymbol{\rho}} \boldsymbol{\rho}, \sum_{\boldsymbol{\rho} \in \mathcal{I}(G)} \alpha_{\boldsymbol{\rho}} = 1, \alpha_{\boldsymbol{0}} > 0\right\}. \tag{16}$$

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. ■

Lemma 5.2: 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 $\boldsymbol{S} = [S_i]$ such that

$$\max_i R_i \le \phi \quad \text{and} \quad \max_i S_i \le \phi,$$

where α is defined in (15).

Proof: 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

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

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 $\delta > 0$,

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

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 $\left[\frac{1}{\chi(G)} - \delta\right] \in P_I(G)$. 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.

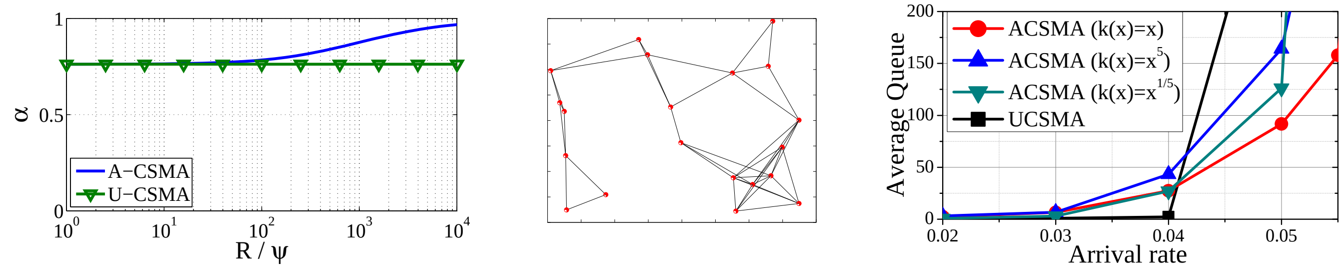
Now the second case $\boldsymbol{\lambda} \in \min_{i \in V} \sum_{\boldsymbol{c} \in \mathcal{H}^n} c_i \pi_{\boldsymbol{c}} \cdot \boldsymbol{\Lambda}^o$ can be proved in an similar manner, where we have to prove that there exists a U-CSMA algorithm satisfying

$$\lim_{t\to\infty} \mathbb{E}[\sigma_i(t)] > \rho_i \qquad \text{for all } i \in V,$$

where we define $\boldsymbol{\rho} = [\rho_i]$ as

$$\boldsymbol{\rho} = \frac{1}{\min_{i \in V} \sum_{\boldsymbol{c} \in \mathcal{H}^n} c_i \pi_{\boldsymbol{c}}} \cdot \boldsymbol{\lambda} \in \boldsymbol{\Lambda}^o \subset P_I(G).$$

This follows from Lemma 5.1 and $\boldsymbol{\rho} \in P_I(G)$. This completes the proof of Lemma 5.2. ■



(a) 5-link complete graph

(b) Random topology

(c) Random: A-CSMA vs U-CSMA and Uniqueness

Fig. 1. Numerical Results

Lemma 5.2 implies that for any arrival rate $\boldsymbol{\lambda} = [\lambda_i] \in \alpha\boldsymbol{\Lambda}^o$, there exist $\varepsilon > 0$ and a rate-stable 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 5.1.

VI. Numerical Results

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

Complete interference graph. We first consider a 5-link complete interference graph, *i.e.*, all 5 links interfere with each other. All queues are homogeneous in terms of time-varying channels, where we assume that the channel space is simply $\{0.5, 1\}$ and the transition-rate $\gamma = \gamma^{0.5\to1} = \gamma^{1\to0.5}$. We compare A-CSMA and U-CSMA, with the following setups:

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

so that $\log(f_i/g_i) = 4x$ for A-CSMA and 4 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.*, $\lim_{t\to\infty}\frac{1}{t}\widehat{\boldsymbol{D}}(t)$. Figure 1(a) shows the results, where in x-axis, we vary the ratio of backoff rate R to 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). We observer that (i) by reflecting the channel capacity in the CSMA parameters as an exponential function, A-CSMA has α-throughput where α approaches 100% (this can be explained by Theorem 3.1), and (ii) U-CSMA has 76%-throughput. Note that $\alpha \ge 76\%$ even with limited backoff rates (*i.e.*, small R/ψ), and this matches Corollary 5.1 which states that A-CSMA's throughput is at least U-CSMA's throughput.

Random topology. We now study *dynamic* 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 Figure 1(b). To model interference, we assume the two-hop interference model (*i.e.*, any two links within two hops interfere) as in 802.11. Here, each link has independent and identical channels, where $\mathcal{H} = \{\frac{u}{10} : 1 \le u \le 10\}$. For all link i, $\gamma^{u/10\to(u+1)/10} = \gamma^{u/10\to(u-1)/10} = 0.01$, and 0 otherwise.

In Figure 1(c), we increase the arrival rates homogeneously for all links, and plot the average queue lengths to see which arrival rates make the system stable or unstable across the tested algorithms. The average queue length blows up when the algorithm cannot stabilize the given arrival rate. We test *dynamic* A-CSMA and U-CSMA algorithms: the queue-based

A-CSMA(x), A-CSMA(x^5), A-CSMA($x^{1/5}$), and U-CSMA, where for given function $k(x)$, A-CSMA($k(x)$) denotes the A-CSMA algorithms satisfying (7). Note that if $k(x) = 1$, A-CSMA($k(x)$) is equal to U-CSMA. The functions $[f_i]$ and $[g_i]$ are defined as stated in Section IV-B except the channel adaptation function $k(\cdot)$. Figure 1(c) shows that *(a)* A-CSMA(x) stabilize more arrival rates than A-CSMA(x^5) and A-CSMA($x^{1/5}$), which coincides with our uniqueness result (see Theorem 3.2) and *(b)* dynamic A-CSMA algorithms outperform dynamic U-CSMA when the arrival rate exceeds 0.04, which means that the achievable rate region of A-CSMA includes the achievable rate region of U-CSMA. In the low arrival-rate regime, U-CSMA could be better than A-CSMA in view of delay because the transmission intensity of U-CSMA is always high when the queue is large, while, under A-CSMA algorithm, each link waits until its channel condition being good although the queue length is large.

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 propose A-CSMA which behaves adaptively to channel variations. We first show that the achievable rate region of A-CSMA equals to the maximum rate region under a novel design of the CSMA parameter updating rules as some particular function of instantaneous link capacity. From this result, we also design throughput optimal A-CSMA algorithms. Finally, we also consider a practical case of limited backoff rates. We proved that with any backoff-rate limitation, A-CSMA has the worst-case throughput guarantee without dependence on network topologies, characterized by the achievable throughput of the channel-unaware CSMA which does not adapt to channel variations.

References

[1] V. Anantharam and P. Tsoucas. A proof of the Markov chain tree theorem. *Statistics & Probability Letters*, 8(2):189 – 192, 1989.

[2] L. Georgiadis, M. J. Neely, and L. Tassiulas. Resource allocation and cross-layer control in wireless networks. *Foundations and Trends in Networking*, 1(1):1–149, 2006.

[3] T. P. Hayes and A. Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *Annals of Applied Probability*, 17(3):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 *Proceedings of Infocom*, 2011.

[5] L. Jiang, D. Shah, J. Shin, and J. Walrand. Distributed random access algorithm: Scheduling and congestion control. *IEEE Transactions on Information Theory*, 56(12):6182 –6207, dec. 2010.

[6] L. Jiang and J. Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. *IEEE/ACM Transactions on Networking*, 18(3):960 –972, June 2010.

[7] L. Jiang and J. Walrand. *Scheduling and Congestion Control for Wireless and Processing Networks*. Morgan-Claypool, 2010.

[8] L. Jiang and J. Walrand. Approaching throughput-optimality in distributed CSMA scheduling algorithms with collisions. *IEEE/ACM Transactions on Networking*, 19(3):816–829, June 2011.

[9] T. H. Kim, J. Ni, R. Srikant, and N. Vaidya. On the achievable throughput of CSMA under imperfect carrier sensing. In *Proceedings of Infocom*, 2011.

[10] C.-H. Lee, D. Y. Eun, S.-Y. Yun, and Y. Yi. From Glauber dynamics to Metropolis algorithm: Smaller delay in optimal CSMA. In *Proceedings of ISIT*, June 2012.

[11] J. Lee, H. Lee, Y. Yi, S. Chong, B. Nardelli, and M. Chiang. Making 802.11 dcf near-optimal: Design, implementation, and evaluation. In *Proceedings of IEEE Secon*, 2013.

[12] B. Li and A. Eryilmaz. A Fast-CSMA Algorithm for Deadline-Constrained Scheduling over Wireless Fading Channels. *ArXiv e-prints*, Mar. 2012.

[13] J. Liu, Y. Yi, A. Proutiere, M. Chiang, and H. V. Poor. Towards utility-optimal random access without message passing. *Wiley Journal of Wireless Communications and Mobile Computing*, 10(1):115–128, Jan. 2010.

[14] M. Lotfinezhad and P. Marbach. Throughput-optimal random access with order-optimal delay. In *Proceedings of Infocom*, 2011.

[15] B. Nardelli, J. Lee, K. Lee, Y. Yi, S. Chong, E. Knightly, and M. Chiang. Experimental evaluation of optimal CSMA. In *Proc. IEEE INFOCOM*, 2011.

[16] 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 *Proceedings of Infocom*, 2010.

[17] A. Proutiere, Y. Yi, T. Lan, and M. Chiang. Resource allocation over network dynamics without timescale separation. In *Proceedings of Infocom*, 2010.

[18] D. Qian, D. Zheng, J. Zhang, and N. Shroff. CSMA-based distributed scheduling in multi-hop MIMO networks under SINR model. In *Proceedings of Infocom*, 2010.

[19] S. Rajagopalan, D. Shah, and J. Shin. Network adiabatic theorem: An efficient randomized protocol for contention resolution. In *Proceedings of ACM Sigmetrics*, 2009.

[20] D. Shah and J. Shin. Delay optimal queue-based CSMA. In *Proceedings of ACM Sigmetrics*, 2010.

[21] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *Annals of Applied Probability*, 22:128–171, 2012.

[22] D. Shah, J. Shin, and P. Tetali. Medium access using queues. In *Proceedings of IEEE FOCS*, 2011.

[23] L. Tassiulas. Scheduling and performance limits of networks with constantly changing topology. *IEEE Transactions on Information Theory*, 43(3):1067 –1073, may 1997.

[24] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12):1936–1949, December 1992.

[25] H. S. Wang and N. Moayeri. Finite-state markov channel-a useful model for radio communication channels. *IEEE Transactions on Vehicular Technology*, 44(1):163 –171, feb 1995.

[26] Y. Yi and M. Chiang. *Next-Generation Internet Architectures and Protocols*. Cambridge University Press, 2011. Chapter 9: Stochastic network utility maximization and wireless scheduling.

[27] S.-Y. Yun, J. Shin, and Y. Yi. CSMA over time-varying channels: Optimality, uniqueness and limited backoff rate. *ArXiv e-prints*, 2013.

[28] S.-Y. Yun, J. Shin, and Y. Yi. CSMA using the bethe approximation for utility maximization. In *Proceedings of ISIT*, 2013.

[29] S.-Y. Yun, Y. Yi, J. Shin, and D. Y. Eun. Optimal CSMA: A survey. In *Proceedings of ICCS*, 2012.