Delay Optimal CSMA With Linear Virtual Channels Under a General Topology

Donggyu Yun, Dongmyung Lee, Se-Young Yun, Jinwoo Shin, and Yung Yi

Abstract—In the past few years, an exciting progress has been made on CSMA (Carrier Sense Multiple Access) algorithms that achieve throughput and utility optimality for wireless networks. However, most of these algorithms are known to exhibit poor delay performance making them impractical for implementation. Recently, several papers have addressed the delay issue of CSMA and yet, most of them are limited, in the sense that they focus merely on specific network scenarios with certain conditions rather than general network topology, achieve low delay at the cost of throughput reduction, or lack rigorous provable guarantees. In this paper, we focus on the recent idea of exploiting multiple channels (actually or virtually) for delay reduction in CSMA, and prove that it is *per-link delay order-optimal*, i.e., $O(1)$-asymptotic-delay per link, if the number of virtual channels is logarithmic with respect to mixing time of the underlying CSMA Markov chain. The logarithmic number is typically small, i.e., at most linear with respect to the network size. In other words, our contribution provides not only a provable framework for the multiple-channel based CSMA, but also the required explicit number of virtual-multi-channels, which is of great importance for actual implementation. The key step of our analytic framework lies in using quadratic Lyapunov functions in conjunction with (recursively applying) Lindley equation and Azuma's inequality for obtaining an exponential decaying property in certain queueing dynamics. We believe that our technique is of broader interest in analyzing the delay performance of queueing systems with multiple periodic schedulers.

Index Terms—CSMA, delay performance, Glauber dynamics, wireless ad-hoc network.

I. INTRODUCTION

IN WIRELESS and computer networks, multiple nodes share a communication medium for transmitting their data packets. In order to achieve an efficient channel utilization to resolve any potential conflicts or interferences between competing nodes, designing a good scheduling algorithm, or medium access control (MAC) protocol is of crucial importance. In their seminal work, Tassiulas and Ephremides [1] proposed a scheduling algorithm known as 'Max-Weight' (MW) that achieves the throughput optimality. However, the MW algorithm requires to solve a NP-hard optimization problem in a centralized manner at each time slot, which is its main drawback for applying it to a large-scale network. Much efforts for designing simpler or distributed implementations of MW have been made since then, e.g., [2]–[4], but these algorithms, though distributed, require heavy message passing.

In the past few years or so, there has been a breakthrough, where the throughput/utility optimality can be achieved by just locally controlling the CSMA (Carrier Sense Multiple Access) parameters without explicit knowledge of neighboring information [5]–[9] hence providing a simple and distributed MAC with optimal performance guarantee. These algorithms are all based on one of the Markov Chain Monte Carlo (MCMC) methods called Glauber dynamics which can be used to sample the independent sets of a graph according to a product-form stationary distribution. However, it has been reported that the CSMA algorithms in general are known to suffer from poor delay performance [10]. Thus, it remains to develop or even verify the existence of such fully-distributed, yet highly delay-efficient MAC algorithms without message passing (possibly based on CSMA).

In this paper, we resolve the delay issue of CSMA algorithms, where we summarize our contributions in what follows.

A. C1—Provable Delay Order-Optimal CSMA

We show the existence of a throughput-optimal, CSMA-based (thus fully-distributed) MAC with a provable *per-link* order-optimal delay performance for *general* wireless network topologies. To the best of our knowledge, this is the first such work in the literature, where even the network-wide order-optimal CSMA delay (which is weaker than our per-link optimality) is not known (see Section II for more details).

B. C2—Connection Between Virtual Channels and Mixing Time

The CSMA algorithm studied in this paper employ virtual multi-channels, referred to as the *delayed* CSMA [11]. In particular, we prove that it has the $O(1)$-asymptotic-delay *per link*, if the number of virtual channels is in the logarithmic order of the mixing time of the underlying CSMA Markov chain. The important implication of our result is that a small number of channels suffices to achieve the optimal delay performance of the CSMA algorithm since the logarithmic number is typically small, i.e., at most linear with respect to the network size [12].

The algorithm is easy to implement and fully-distributed, i.e., no message passing is necessary. The key part of our analytic framework is the use of quadratic Lyapunov functions in conjunction with (recursively applying) Lindley equation

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

2

IEEE/ACM TRANSACTIONS ON NETWORKING

and Azuma's inequality for obtaining an exponential decaying property in certain queueing dynamics (see Section V-A for more details). We believe that our technique is of broader interest in analyzing the delay performance of queueing systems with multiple periodic schedulers.

The rest of this paper is organized as follows. We first review the related work in Section II. Section III describes the system model and introduces the notion of the delay-optimality. Section IV explains the *delayed* CSMA algorithm and presents the main theorem—the delay-optimality of the *delayed* CSMA. The correctness of the main theorem is proved in Section V. In Section VI, we present various simulation results to evaluate the performance of the *delayed* CSMA under different scenarios. Section VII concludes this paper.

II. RELATED WORK

In the literature, several papers proposed different CSMA-based MACs, all of which are provably (close to be) delay-optimal under certain restrictions. Shah and Shin [13] proposed a modified queue-based CSMA algorithm where at each time instance a very small fraction of *frozen* nodes do not execute CSMA operations. By appropriately selecting such frozen nodes, the proposed algorithm leads to the network-wide $O(n)$ queue length, where n is the number of links. Lotfinezhad and Marbach [10] prove that by periodically resetting all links to become silent and then immediately restarting the classical CSMA protocol, the algorithm leads to the $O(1)$ per-packet delay for grid or torus interference graphs. Jiang *et al.* [14] study CSMA algorithms based on parallel Glauber dynamics. The authors prove that the algorithms can achieve the network-wide $O(n \log n)$ queue length, and Subramanian and Alanyali [15] further tightened the bound. The limitations in the above papers are *(i)* [10], [13] are only applicable to a specific type of topologies, i.e., geometric or torus (inference) network topologies, and *(ii)* [14], [15] require reduced offered loads and a tradeoff between throughput and delay occurs (i.e., in both papers, a certain amount of throughput should be sacrificed for low delay). We overcome these limitations in the current paper, and more importantly, we obtain the per-link $O(1)$ queue length, which is much stronger than the network-wide (or mean) queue length bound.

Another direction of designing CSMA algorithms for better delay performance is to use (actual or virtual) 'multi-channels,' which is in essence motivated by *resolving the temporal link starvation of CSMA via 'de-correlating' the temporal accessibility* to the wireless medium. Lam *et al.* [16] considers a CSMA algorithm with multiple frequency agility, such that more than one frequency channel is available yet a link can transmit on at most one of the channels. They associate temporal starvation to the mixing time and then show that the region of fast mixing time implies that of low temporal starvation through simulations. Huang and Lin [17] proposed an algorithm called VMC-CSMA in which multiple virtual channels (defined by dividing the time line) are used to emulate a multi-channel system and address the starvation problem. The algorithm randomly selects a virtual channel, and the schedule corresponding to this chosen channel is used at each time slot. The authors show that the expected packet delay for each link equals to the inverse of its long-term average rate, and the distribution of its head-of-line (HOL) waiting time can be asymptotically bounded. The

multi-channel idea is also used in [11], where the authors propose the so-called *delayed* CSMA inducing multiple CSMA dynamics in a round robin manner. They showed that its asymptotic-delay performance can be improved by exploiting more channels (i.e., more rounds in the *delayed* CSMA). However, an explicit delay bound has yet to be studied, where especially a precise relation between the CSMA delay and the number of channels is practically important for implementation.

The question of designing MAC scheduling algorithms with low delay, not restricted to CSMA, in wireless networks has been also studied from a while ago. To name just a few, the MW algorithm empirically has a good delay performance, but its (network-wide or per-link) delay-optimality for general topology is not analytically known. Neely *et al.* [18] proved that maximal scheduling, which is suboptimal in terms of throughput, achieves $O(1)$ delay. Yi and Chiang [19] studied the 3-D tradeoff among delay, throughput, and complexity for a large class of queue-based scheduling schemes. A 'batching' policy, first considered by Neely *et al.* [20], is known to be almost delay-optimal for input-queued switch networks (a special topology of wireless networks), where its per-link delay is $O(\log n)$. Recently, Shah *et al.* [21] developed a centralized delay-optimal scheduling algorithm for wireless networks including input-queued switch networks. On the negative side, Shah *et al.* [22] showed that there exists no polynomial-time scheduler (including CSMA) with a polynomial delay for arbitrary network topologies unless $\mathbf{NP} \subseteq \mathbf{BPP}$. However, the authors consider the supremum of temporal delays over time, which does not imply that the asymptotic time-averaged delay performance is necessarily bad. Somewhat surprisingly, in the current paper we show that even a simple fully-distributed CSMA algorithm can achieve the asymptotic time-averaged $O(1)$ delay per link.

Other relevant approaches include the one recently studied in [23], which proposes a CSMA-based utility optimal MAC protocol with *the worst case delay guarantee* such that each job is either scheduled for transmission or dropped before the deadline. The main interest of [23] lies in ensuring the worst delay of the admission control for jobs, while this paper focuses on reducing the averaging queueing delay.

III. SYSTEM MODEL

A. Model

We consider a network model where the interference relationship among the wireless links can be represented by the so-called *interference graph* $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N}$ is the set of links, and $\mathcal{E}$ is the set of edges representing (symmetric) interference relationship between links. An edge $(i, j) \in \mathcal{E}$ exists between two links i and j if the corresponding wireless links interfere with each other. We denote $\mathcal{N}_v = \{w \in \mathcal{N} : (v, w) \in \mathcal{E}\}$ as the set of neighbors of vertex v. Time is divided into discrete slots, indexed by $t = 0, 1, \ldots$. Let $\boldsymbol{\sigma}(t) = (\sigma_v(t))_{v \in \mathcal{N}} \in \{0, 1\}^{|\mathcal{N}|}$ be a *schedule* that represents the set of transmitting links at time t. A link v (or vertex v in the interference graph $\mathcal{G}$) is active if it is included in the schedule, i.e., $\sigma_v(t) = 1$, and is inactive otherwise. We denote by $\Omega \subseteq \{0, 1\}^{|\mathcal{N}|}$ the set of all feasible schedules on $\mathcal{G}$, where a *feasible* schedule $\boldsymbol{\sigma}(t)$ should satisfy the independent set constraint i.e.,

$$\boldsymbol{\sigma}(t) \in \mathcal{I}(\mathcal{G}) \triangleq \{\boldsymbol{\chi} \in \{0, 1\}^{|\mathcal{N}|} : \chi_v + \chi_w \leq 1, \forall (v, w) \in \mathcal{E}\}.$$

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

YUN *et al.*: DELAY OPTIMAL CSMA WITH LINEAR VIRTUAL CHANNELS UNDER A GENERAL TOPOLOGY

3

Each link is associated with a queue that has a dedicated exogenous arrival process. Let $A_v(t)$ denote the number of packet (of unit-size) arrival of link v at time t. We assume Bernoulli arrivals, i.e., $A_v(t) \in \{0, 1\}$ with $\mathbb{P}[A_v(t) = 1] = \lambda_v$. Let $\boldsymbol{Q}(t) = (Q_v(t))_{v \in \mathcal{N}}$ denote the vector of queue sizes at time t. Then it has the following dynamics:

$$Q_v(t) = \max\{Q_v(t-1) + A_v(t) - \sigma_v(t), 0\} \qquad t \geq 1,$$

which is known as the *Lindley* equation. We denote by $A_v[t, t']$ and $D_v[t, t']$ the number of arrivals and (potential) departures in the interval $[t, t']$, respectively, i.e.,

$$A_v[t, t'] = \sum_{s=t}^{t'} A_v(s) \quad \text{and} \quad D_v[t, t'] = \sum_{s=t}^{t'} \sigma_v(s).$$

In addition, we let $\Delta_v[t, t'] = A_v[t, t'] - D_v[t, t']$. Then, by recursively applying the Lindley equations, we obtain the following:

$$\begin{aligned} Q_v(t) &= \max\{Q_v(t-1) + A_v[t, t] - D_v[t, t], 0\} \\ &= \max\{\max\{Q_v(t-2) + A_v[t-1, t-1] \\ &\quad - D_v[t-1, t-1], 0\} + A_v[t, t] - D_v[t, t], 0\} \\ &\quad \cdots \\ &= \max\Big\{Q_v(0) + A_v[1, t] - D_v[1, t], \max_{2 \leq s \leq t} \\ &\quad \{A_v[s, t] - D_v[s, t]\}, 0\Big\}. \end{aligned} \qquad (1)$$

The above equation will play a crucial role in our analysis.

We define the capacity region $\mathcal{C} \subset [0, 1]^{|\mathcal{N}|}$ of a network as the convex hull of the feasible scheduling set $\mathcal{I}(\mathcal{G})$, i.e.,

$$\mathcal{C} = \left\{ \sum_{\chi \in \mathcal{I}(\mathcal{G})} \alpha_\chi \chi : \sum_{\chi \in \mathcal{I}(\mathcal{G})} \alpha_\chi = 1, \alpha_\chi \geq 0, \forall \chi \in \mathcal{I}(\mathcal{G}) \right\}.$$

Let $\boldsymbol{\lambda} = [\lambda_v]$ and it is called *admissible* if $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, where

$$\boldsymbol{\Lambda} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^{|\mathcal{N}|} : \lambda_v \leq \gamma_v, \forall v, \text{ for some } \boldsymbol{\gamma} = [\gamma_v] \in \mathcal{C} \right\}.$$

The intuition behind this notion of $\boldsymbol{\Lambda}$ comes from the fact that any scheduling algorithm has to choose a schedule from $\mathcal{I}(\mathcal{G})$ at each time and hence the arrival rate must belong to $\boldsymbol{\Lambda}$ (otherwise, queues should grow over time). In addition, for given $\varepsilon > 0$, $\boldsymbol{\lambda}$ is called $\varepsilon - admissible$ if $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^\varepsilon$, where

$$\boldsymbol{\Lambda}^\varepsilon = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^{|\mathcal{N}|} : \lambda_v + \varepsilon < \gamma_v, \forall v, \text{ for some } \boldsymbol{\gamma} = [\gamma_v] \in \mathcal{C} \right\}.$$

Finally, $\boldsymbol{\lambda}$ is called *strictly admissible* if $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^\circ$ where $\boldsymbol{\Lambda}^\circ = \bigcup_{\varepsilon > 0} \boldsymbol{\Lambda}^\varepsilon$.

B. Performance Metric

A scheduling algorithm decides a sequence of $\boldsymbol{\sigma}(t) \in \mathcal{I}(\mathcal{G})$ over $t = 0, 1, \ldots$. In this section, we introduce the main performance metrics, which are the throughput and delay of scheduling algorithms. First, we define the throughput-optimality:

Definition 3.1 (Throughput-Optimality): A scheduling algorithm is called *throughput-optimal* if for any strictly admissible arrival rate $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^\circ$, queues remain finite with probability 1 under the algorithm, i.e.,

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{v \in \mathcal{N}} Q_v(t) < \infty, \qquad \text{with probability 1.} \quad (2)$$

A popular approach for showing the throughput-optimality is (a) defining the underlying network Markov chain induced by a scheduling algorithm and (b) proving its *positive recurrence*. We now introduce the delay optimality studied in this paper.

Definition 3.2 (Delay-Optimality): A scheduling algorithm is called *per-link delay-optimal* (or simply delay-optimal),[1] if for any ε-admissible arrival rate $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^\varepsilon$ with $\varepsilon = \Omega(1)$,

$$\limsup_{t \to \infty} \mathbb{E}[Q_v(t)] = O(1), \qquad \text{for all } v \in \mathcal{N}.$$

In the above definition, the orders $\Omega(1)$ and $O(1)$ are with respect to the network size $|\mathcal{N}|$, i.e., delay-optimality means that the *per-link* queue-size remains 'constant' as the network size grows.

IV. MAIN RESULT

Our interest is in fully-distributed CSMA scheduling algorithms, where in particular, we study the *delayed* CSMA proposed in [11]. The main idea of the *delayed* CSMA is to use multiple schedulers in a round-robin manner in order to effectively reduce the correlations between the link state process, in an attempt to alleviate the so-called *starvation problem*, i.e., once a schedule is chosen, it keeps being scheduled without any change for a large number of slots. In this section, we explain the *delayed* CSMA algorithm in more detail and then present the main result of this paper—the delay-optimality of the *delayed* CSMA algorithm. We first summarize the main contribution of our paper beyond [11]. The authors have shown that under *delayed* CSMA, the probability that a queue length is larger than some value x, is approximated by a function that is exponentially decreasing in x and T, where T is the number of schedulers. However, the result is quite asymptotic and does not imply delay-optimality, i.e., it is not clear whether the expected queue length becomes $O(1)$ as T grows. Furthermore, a more precise relation between the queue length and T is of significant importance for its actual implementation, which we characterize in this paper using the mixing time of the underlying CSMA Markov chains.

A. Delayed CSMA

The *delayed* CSMA algorithm is formally described in Algorithm 1. In the *delayed* CSMA, at each time slot, a decision schedule is chosen $\mathcal{D}(t) \in \mathcal{I}(\mathcal{G})$, which corresponds to a selection of an independent set of $I(\mathcal{G})$. The active links in the decision schedule become the candidate links which may change their state. There are various ways to choose a decision schedule $\mathcal{D}(t) \in \mathcal{I}(\mathcal{G})$ at each time slot. For example, each link simply attempts to access the medium with a fixed access probability a_v and then $v \in \mathcal{D}(t)$ with probability $a_v \prod_{u \in \mathcal{N}_v} (1 - a_u)$, or a randomized scheme with light control message exchanges can be used, as in [24]. In general, we assume that $\{\mathcal{D}(t)\}$ is a set of independent identical random variables such that $\mathbb{P}[v \in \mathcal{D}(t)] > 0$ for all v. Under this assumption on $\mathcal{D}(t)$ and given the 'fugacity' $[r_v]$, the schedule $\{\boldsymbol{\sigma}(t) : t \equiv k \pmod{T}\}$ forms a (discrete-time) irreducible and aperiodic Markov chain for $k = 0, 1, \ldots, T - 1$, e.g., the

---

[1] This per-link optimality is much stronger than the 'network-wide' optimality defined by the averaged delay over all links.

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

4

IEEE/ACM TRANSACTIONS ON NETWORKING

Algorithm 1 Delayed CSMA **[11]**

1: *Initialize:* for all links $v \in \mathcal{N}$, $\sigma_v(t) = 0$, $t = 0, \ldots, T-1$.
2: At each time $t \geq T$: links find a decision schedule,
3: $\mathcal{D}(t) \in \mathcal{I}(\mathcal{G})$ through a randomized procedure, and
4: **for all** links $v \in \mathcal{D}(t)$ **do**
5:    **if** $\sum_{w \in N_v} \sigma_w(t - T) = 0$ **then**
6:       $\sigma_v(t) = 1$ with probability $\frac{r_v}{1+r_v}$
7:       $\sigma_v(t) = 0$ with probability $\frac{1}{1+r_v}$
8:    **else**
9:       $\sigma_v(t) = 0$
10:    **end if**
11: **end for**
12: **for all** links $w \notin \mathcal{D}(t)$ **do**
13:    $\sigma_v(t) = \sigma_v(t - T)$
14: **end for**

---

k-th Markov chain is $\{\boldsymbol{\sigma}(uT + k) : u = 0, 1, 2, \ldots\}$.[2] The common stationary distribution $\boldsymbol{\pi} = [\pi_{\boldsymbol{\sigma}}]$ is given by

$$\pi_{\boldsymbol{\sigma}} = \frac{1}{Z} \prod_{v \in \mathcal{N}} r_v^{\sigma_v}, \tag{3}$$

where $Z = \sum_{\boldsymbol{\sigma} \in \Omega} \prod_{v \in \mathcal{N}} r_v^{\sigma_v}$ is a normalizing constant. Hence, one can think that the algorithm utilizes multiple T (almost) independent Markov chains (or schedulers). From their ergodicities, we know that for all $v \in \mathcal{N}$,

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} \sigma_v(s) = \mathbb{P}_{\pi} \sigma_v = 1.$$

The essence of recently developed throughput-optimal CSMA algorithms [7], [8] is finding an appropriate fugacity $[r_v]$ such that the long-term link throughput $\lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} \sigma_v(s)$ is (strictly) greater than the arrival rate λ_v (whenever $\boldsymbol{\lambda}$ is strictly admissible). Given such a fugacity, it is not hard to prove the desired positive recurrence of the underlying network Markov chain [8]. Several authors proposed different adaptive updating rules on fugacities, which converge to the appropriate fugacity [7]–[9], [25]. Since the main focus of this paper is to analyze the delay performance of the *delayed* CSMA, we assume that links initially start with the desired fugacity. Formally speaking, for given ε-admissible arrival rate $\boldsymbol{\lambda}$, we assume that

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} \sigma_v(s) = \mathbb{P}_{\pi} \sigma_v = 1 \geq \lambda_v + \varepsilon, \text{ for all } v \in \mathcal{N}. \tag{4}$$

This is merely for simple presentation of our proof, and one can easily extend it without the assumption, i.e., under the known adaptive fugacity updating rules in [8]. This is because our delay metric is defined in an asymptotic manner (i.e., uses 'limsup') and the exact convergence in the fugacity updating rules is not crucial in our proof strategy.[3]

---

[2]We say $t \equiv k \pmod{T}$ if $t - k$ is an integer multiple of T. It is called congruent modulo t.

[3]The 'fixed' fugacity assumptions have also been taken in other papers dealing with CSMA delay, e.g., [14].

B. Preliminaries: Mixing Time

To describe our result formally, we first introduce the necessary definitions of the *total variation distance* and the corresponding *mixing time* of the CSMA Markov chain. The total variation distance between two probability distributions $\boldsymbol{\eta} = [\eta_i]$ and $\boldsymbol{\nu} = [\nu_i]$ on state space Ω is

$$\|\boldsymbol{\eta} - \boldsymbol{\nu}\|_{TV} = \frac{1}{2} \sum_{i \in \Omega} |\eta_i - \nu_i|.$$

Using this distance metric, the *mixing time* of the k-th CSMA Markov chain $\{\boldsymbol{\sigma}(uT + k) : u = 0, 1, 2, \ldots\}$ is defined as follows:

$$M^{(k)}(\delta) = \inf \left\{ s : \max_{\boldsymbol{\mu}(k)} \|\boldsymbol{\mu}(uT + k) - \boldsymbol{\pi}\|_{TV} \leq \delta, \ \forall u \geq s \right\},$$

where $\delta > 0$ is some constant (which we will choose later) and $\boldsymbol{\mu}(t)$ denotes the probability distribution of random variable $\boldsymbol{\sigma}(t)$. The mixing time measures how long it takes for the k-th CSMA Markov chain to converge to the stationary distribution for arbitrary initial distribution $\boldsymbol{\mu}(k)$. Since we assume the fixed common fugacity across the Markov chains, the mixing time $M^{(k)}(\delta)$ is identical for $k = 0, 1, \ldots, T - 1$. Hence, for simplicity, we use $M(\delta)$ instead of $M^{(k)}(\delta)$, for any $k = 0, 1, \ldots, T - 1$.

C. Main Result: Delay-Optimality

Now we are ready to state the main result of this paper, i.e., the delay-optimality of the *delayed* CSMA algorithm.

Theorem 4.1: For any ε-admissible arrival rate $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^{\boldsymbol{\varepsilon}}$, there exists $T^* = O\left(\frac{1}{\varepsilon^3} \log M(\varepsilon/2)\right)$ such that for all $T > T^*$, the corresponding *delayed* CSMA algorithm is delay-optimal, more formally,

$$\lim_{t \to \infty} \mathbb{E}[Q_v(t)] = O\left(\frac{1}{\varepsilon^4}\right), \qquad \text{for all } v \in \mathcal{N}.$$

The above theorem states that the per-link average queue-size is bounded by a constant for sufficiently large T, the number of CSMA schedulers. The purpose of choosing large T is to effectively reduce the dependency among consecutive link states, which promotes much faster link state changes and hence alleviates the starvation problem.

We further remark that Theorem 4.1 is optimal with respect to the network size, but not with respect to ε, e.g., the best order of delay in both parameters is $O(1/\varepsilon)$. In this paper, we do not make much efforts to optimize our analysis for the better delay dependency in ε and the tighter bound of T^*. For example, in our simulation results (see Section VI), we observe that $T = 2$ is enough for the order-optimal delay for grid-like graphs. We provide the proof of Theorem 4.1 in the following section.

V. PROOFS: THEOREM 4.1 AND NECESSARY LEMMAS

A. Proof Strategy

We first describe our proof strategy at a high level, followed by the detailed proof in Section V-B. We use the popular approach using a quadratic Lyapunov function to prove the desired delay bound in Theorem 4.1. In this approach, one has to define an appropriate network Markov chain and show a certain negative drift property for the Lyapunov function (see Lemma 5.1). In particular, the network Markov chain $\{X(t)\}$ under the *delayed* CSMA algorithm is

$$X(t) = \big(\boldsymbol{Q}(t), \boldsymbol{\sigma}(t), \boldsymbol{\sigma}(t-1), \ldots, \boldsymbol{\sigma}(t-T+1)\big).$$

Two major technical challenges for deriving the desired negative drift property are:

- It is necessary that the CSMA Markov chains mix, which takes the mixing time $M = M(\varepsilon/2)$. Otherwise, the scheduling dynamics is hard to analyze.
- Even after the CSMA Markov chains mix, schedules $\{\boldsymbol{\sigma}(\cdot)\}$ are correlated (i.e., they are not i.i.d random variables), but one has to show that they still satisfy some version of law of large numbers.

For the first issue, we design a stopping time obtaining the negative drift property to be much later than the mixing time so that the negative drift in the mixing period dominates the (potential) positive drift in the initial non-mixing one. For the second issue, we first observe that schedules $\{\boldsymbol{\sigma}(\cdot)\}$ in each time interval of length T are almost i.i.d random variables (after the CSMA Markov chains mix) due to the design of the delayed CSMA algorithm. Hence, we choose T large enough so that (a) the 'variance' of the sum of schedules in each time interval of length T is small enough (b) the possible correlation across the time intervals is compensated. Specifically, for (a) we use Azuma's inequality for obtaining an exponential decaying property of queueing dynamics in each time interval of length T, and for (b) the union bound is used under Lindley recursions.

B. Proof of Theorem 4.1

We now start toward proving Theorem 4.1. First observe that the network Markov chain $\{X(t)\}$ is aperiodic and irreducible. Its ergodicity (i.e., positive recurrence) can be derived using the following key lemma whose proof is presented in Section V-C.

Lemma 5.1: For any ε-admissible arrival rate $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^{\boldsymbol{\varepsilon}}$, there exist positive numbers $T^* = O\left(\frac{1}{\varepsilon^3}\log M\right)$ such that for all $T > T^*$ and $v \in \mathcal{N}$,

$$\mathbb{E}\left[Q_v(t + 7MT/\varepsilon)^2 \,|\, X(t)\right]$$
$$\leq Q_v(t)^2 - \frac{Q_v(t)}{2} + O\left(\frac{1}{\varepsilon^4}\right). \quad (5)$$

Define the quadratic Lyapunov function $L(t) = \sum_v Q_v(t)^2$. By summing the inequality (5) in Lemma 5.1 over v, we get:

$$\mathbb{E}\left[L(t + 7MT/\varepsilon) - L(t) \,|\, X(t)\right] \leq -\sum_v \frac{Q_v(t)}{2} + O\left(\frac{|\mathcal{N}|}{\varepsilon^4}\right).$$

Then, from the Lyapunov-Foster criteria, $\{X(t)\}$ is positive recurrent.

Furthermore, taking expectations with respect to the distribution of $X(t)$ in both sides of the conclusion of Lemma 5.1, we have

$$\mathbb{E}[Q_v(t + 7MT/\varepsilon)^2] \leq \mathbb{E}[Q_v(t)^2] - \frac{\mathbb{E}[Q_v(t)]}{2} + O\left(\frac{1}{\varepsilon^4}\right).$$

Due to the the ergodicity of $\{X(t)\}$, $\lim_{t\to\infty} \mathbb{E}[Q_v(t + 7MT/\varepsilon)^2]$ and $\lim_{t\to\infty} \mathbb{E}[Q_v(t)^2]$ are finite and should be equal. Therefore, we obtain the conclusion of Theorem 4.1 : $\lim_{t\to\infty} \mathbb{E}[Q_v(t)] = O\left(\frac{1}{\varepsilon^4}\right)$.

C. Proof of Lemma 5.1

It will be clear that our proof technique for Lemma 5.1 can be applied to any $v \in \mathcal{N}$, for notional simplicity we henceforth omit the subscript v and use $Q(t)$ and Δ instead of $Q_v(t)$ and Δ_v. We first present the following lemma which plays a crucial role for proving Lemma 5.1.

Lemma 5.2: For any $q \geq 0$ and $\varphi \geq q - MT/2$, it holds that

$$\mathbb{P}\left[Q\left(t + 7MT/\varepsilon\right) \geq \varphi \,|\, \mathcal{L}\right] \leq g(T, \varphi),$$

where $g(T, \varphi) = \frac{140M}{\varepsilon^3}\exp\left(-\frac{\varepsilon^2 T}{80}\right) + \frac{18}{\varepsilon^2}\exp\left(-\frac{\varphi\varepsilon}{9}\right)$ and

$$\mathcal{L} = \{Q(t) = q, \boldsymbol{\sigma}(t), \ldots, \boldsymbol{\sigma}(t - T + 1)\}.$$

The proof of the above lemma is provided in Section V-D. Lemma 5.2 implies that given a network state at time t, the distribution of the (per-link) queue length at time $t + 7MT/\varepsilon$ has an exponential decaying property (i.e., light tail). We will use this property to bound the expected quadratic queue length at time $t + 7MT/\varepsilon$ for completing the proof of Lemma 5.1.

We will also use the following standard equality for the expectation of a nonegative integer-valued random variable.

Lemma 5.3: If X is a nonnegative integer-valued random variable, then

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{P}\left[X \geq i\right].$$

Since the proof of the above lemma is straightforward, we omit it here.

We proceed the proof by studying two disjoint cases: **Case (i):** $Q(t) = q \leq C$ and **Case (ii):** $Q(t) = q > C$, where $\varepsilon^2 C = O(1)$ is some large constant.

Case (i): $Q(t) = q \leq C$: In this case, we obtain

$$\mathbb{E}\left[Q(t + 7MT/\varepsilon)^2 - Q(t)^2 \,|\, \mathcal{L}\right]$$
$$\leq \mathbb{E}\left[Q(t + 7MT/\varepsilon)^2 \,|\, \mathcal{L}\right] - Q(t)$$
$$\stackrel{(a)}{\leq} \sum_{i=1}^{(C+7MT/\varepsilon)^2} \mathbb{P}\left[Q(t + 7MT/\varepsilon)^2 \geq i \,|\, \mathcal{L}\right] - Q(t)$$
$$= \sum_{i=1}^{(C+7MT/\varepsilon)^2} \mathbb{P}\left[Q(t + 7MT/\varepsilon) \geq \sqrt{i} \,|\, \mathcal{L}\right] - Q(t)$$
$$\stackrel{(b)}{\leq} \sum_{i=1}^{(C+7MT/\varepsilon)^2} g\left(T, \sqrt{i}\right) - Q(t),$$

where (a) is because

$$Q(t + 7MT/\varepsilon) \leq Q(t) + 7MT/\varepsilon \leq C + 7MT/\varepsilon,$$

with Lemma 5.3 and we apply Lemma 5.2 for (b) since $q - MT/2 < 0$, i.e., $q \leq C = O(1/\varepsilon^2)$ and $T > T^* = \Omega\left(\frac{1}{\varepsilon^2}\right)$. Observe that

$$\sum_{i=1}^{(C+7MT/\varepsilon)^2} g\left(T, \sqrt{i}\right) - Q(t)$$
$$= \sum_{i=1}^{(C+7MT/\varepsilon)^2}\left[\frac{O(M)}{\varepsilon^3}\exp\left(-\frac{\varepsilon^2 T}{80}\right)\right.$$
$$\left. + \frac{O(1)}{\varepsilon^2}\exp\left(-\frac{\varepsilon\sqrt{i}}{9}\right)\right] - Q(t)$$
$$\leq O\left(\frac{M^3 T^2}{\varepsilon^5} + \frac{CM^2 T}{\varepsilon^4} + \frac{C^2 M}{\varepsilon^3}\right)\exp\left(-\frac{\varepsilon^2 T}{80}\right)$$
$$+ \sum_{i=1}^{\infty}\frac{O(1)}{\varepsilon^2}\exp\left(-\frac{\varepsilon\sqrt{i}}{9}\right) - Q(t)$$
$$\stackrel{(c)}{=} O(1) + \frac{O(1)}{\varepsilon^2}\underbrace{\sum_{i=1}^{\infty}\exp\left(-\frac{\varepsilon\sqrt{i}}{9}\right)}_{(*)} - Q(t), \quad (6)$$

where (c) is due to large enough T with $T > T^* = \Omega\left(\frac{1}{\varepsilon^3}\log M\right)$ and $C = O(1/\varepsilon^2)$. Therefore, it suffices to bound the term $(*)$ for the proof of Lemma 5.1.

Let $k = \frac{\varepsilon}{9}$. Then, it follows that

$$
\begin{aligned}
&\sum_{i=1}^{\infty} \exp\left(-k\sqrt{i}\right) \\
&\leq \int_0^{\infty} \exp\left(-k\sqrt{i}\right) \mathrm{d}i \\
&= 2\int_0^{\infty} u \exp\left(-ku\right) \mathrm{d}u \qquad \text{(by substituting } u = \sqrt{i}) \\
&= 2\left\{-\frac{1}{k}u\exp\left(-ku\right)\Big|_0^{\infty} + \frac{1}{k}\int_0^{\infty}\exp\left(-ku\right)\mathrm{d}u\right\} \\
&= -\frac{2}{k^2}\exp\left(-ku\right)\Big|_0^{\infty} = \frac{2}{k^2} = O\left(\frac{1}{\varepsilon^2}\right).
\end{aligned} \tag{7}
$$

Combining (6) and (7) completes the proof of Lemma 5.1 for **Case (i)**.

Case (ii): $Q(t) = q > C$: In this case, we define the following (sub-)events of $\mathcal{L}$:

$$
\begin{aligned}
\mathcal{L}_- &= \{Q(t+7MT/\varepsilon) - Q(t) \leq -1\} \cap \mathcal{L}. \\
\mathcal{L}_+ &= \{Q(t+7MT/\varepsilon) - Q(t) > -1\} \cap \mathcal{L}.
\end{aligned}
$$

We also let $\delta(t) \equiv Q(t+7MT/\varepsilon) - Q(t)$. Using those notations, we have

$$
\begin{aligned}
&\mathbb{E}\left[Q(t+7MT/\varepsilon)^2 - Q(t)^2 \,|\, \mathcal{L}\right] \\
&= \mathbb{E}\left[Q(t+7MT/\varepsilon)^2 - Q(t)^2 \,|\, \mathcal{L}_-\right]\mathbb{P}\left[\mathcal{L}_- \,|\, \mathcal{L}\right] \\
&\quad + \mathbb{E}\left[Q(t+7MT/\varepsilon)^2 - Q(t)^2 \,|\, \mathcal{L}_+\right]\mathbb{P}\left[\mathcal{L}_+ \,|\, \mathcal{L}\right] \\
&= \mathbb{E}\left[\delta(t)(Q(t+7MT/\varepsilon) + Q(t)) \,|\, \mathcal{L}_-\right]\mathbb{P}\left[\mathcal{L}_- \,|\, \mathcal{L}\right] \\
&\quad + \mathbb{E}\left[\delta(t)(\delta(t) + 2Q(t)) \,|\, \mathcal{L}_+\right]\mathbb{P}\left[\mathcal{L}_+ \,|\, \mathcal{L}\right] \\
&\leq \mathbb{E}\left[-Q(t+7MT/\varepsilon) - Q(t) \,|\, \mathcal{L}_-\right]\mathbb{P}\left[\mathcal{L}_- \,|\, \mathcal{L}\right] \\
&\quad + \mathbb{E}\left[3\delta^2(t)Q(t) \,|\, \mathcal{L}_+\right] \\
&\leq -Q(t)\mathbb{P}\left[\mathcal{L}_- \,|\, \mathcal{L}\right] + 3Q(t)\mathbb{E}\left[\delta^2(t) \,|\, \mathcal{L}_+\right] \\
&\overset{(d)}{\leq} -Q(t)\Big(1 - g(T, Q(t)-1)\Big) \\
&\quad + 3Q(t)\sum_{i=1}^{49M^2T^2/\varepsilon^2}\mathbb{P}\left[Q(t+7MT/\varepsilon) - Q(t) \geq \sqrt{i} \,|\, \mathcal{L}\right] \\
&\leq \underbrace{\left(g(T,Q(t)-1) + 3\sum_{i=1}^{49M^2T^2/\varepsilon^2} g\left(T, Q(t)+\sqrt{i}\right) - 1\right)}_{(**)}Q(t),
\end{aligned}
$$

where (d) is from Lemma 5.3 and Lemma 5.2 and we also apply Lemma 5.2 for the last inequality and one can show that the term $(**)$ is smaller than $1/2$ if $Q(t) > C$ and $\varepsilon^2 C = O(1)$ is a large enough constant, similarly as we bound the term $(*)$ in **Case (i)**:

$$
g(T,Q(t)-1) + 3\sum_{i=1}^{49M^2T^2/\varepsilon^2} g\left(T, Q(t)+\sqrt{i}\right) \leq \frac{1}{2}. \tag{8}
$$

The complete proof of the above inequality is given in the Appendix. This completes the proof of Lemma 5.1 for **Case (ii)**.

D. Proof of Lemma 5.2

From the Lindley recursion (1) and the union bound, we have

$$
\begin{aligned}
&\mathbb{P}\left[Q(t+7MT/\varepsilon) \geq \varphi \,|\, \mathcal{L}\right] \\
&\leq \underbrace{\mathbb{P}\left[Q(t+MT) + \Delta[t+MT+1, t+7MT/\varepsilon] \geq \varphi \,|\, \mathcal{L}\right]}_{(\dagger)} \\
&\quad + \underbrace{\mathbb{P}\left[\max_{MT+2 \leq s \leq 7MT/\varepsilon} \Delta[t+s, t+7MT/\varepsilon] \geq \varphi \,|\, \mathcal{L}\right]}_{(\ddagger)}.
\end{aligned}
$$

Hence, it suffices to bound both terms $(\dagger)$ and $(\ddagger)$ for the proof of Lemma 5.2, which we present in what follows.

Bound for $(\dagger)$: We will use Azuma's inequality and the union bound for obtaining the bound of $(\dagger)$.

To this end, we define the appropriate supermartingale under the condition $\mathcal{L}$ as follows: for a given integer $i \in \{M+1, M+2, \ldots, 7M/\varepsilon\}$ and each integer $s \in [t+(i-1)T+1, t+iT]$, let

$$
Y_s^i = \sum_{u=t+(i-1)T+1}^{s} (\Delta[u,u] + \varepsilon/2),
$$

and initially $Y_{t+(i-1)T}^i = 0$. Then, using the definition of the mixing time $M = M(\varepsilon/2)$, the conditional independence of $\{\Delta[s,s] : s \in [t+(i-1)T, t+iT]\}$, and the fact that $\mathbb{E}[\Delta[s,s] + \varepsilon/2] \leq -\varepsilon/2 + \varepsilon/2 = 0$, we can easily check that the following random variables form a supermartingale (whose formal proof is presented in Appendix):

$$
\{Y_s^i : s \in [t+(i-1)T, t+iT]\}. \tag{9}
$$

Using this notation, we have

$$
\begin{aligned}
&\mathbb{P}\left[Q(t+MT) + \Delta[t+MT+1, t+7MT/\varepsilon] \geq \varphi \,|\, \mathcal{L}\right] \\
&\leq \mathbb{P}\left[Q(t) + MT + \Delta[t+MT+1, t+7MT/\varepsilon] \geq \varphi \,|\, \mathcal{L}\right] \\
&= \mathbb{P}\left[\Delta[t+MT+1, t+7MT/\varepsilon] \geq \varphi - Q(t) - MT \,|\, \mathcal{L}\right] \\
&= \mathbb{P}\left[\sum_{i=M+1}^{7M/\varepsilon}\left(Y_{t+iT}^i - Y_{t+(i-1)T}^i - \frac{\varepsilon T}{2}\right) \geq \varphi - q - MT \,|\, \mathcal{L}\right] \\
&\leq \mathbb{P}\left[\bigcup_{i=M+1}^{7M/\varepsilon}\left\{Y_{t+iT}^i - Y_{t+(i-1)T}^i - \frac{\varepsilon T}{2} \geq \frac{\varphi - q - MT}{M(7/\varepsilon - 1)}\right\} \,|\, \mathcal{L}\right] \\
&\overset{(a)}{\leq} \sum_{i=M+1}^{7M/\varepsilon} \mathbb{P}\left[Y_{t+iT}^i - Y_{t+(i-1)T}^i \geq \frac{\varepsilon T}{2} - \frac{\varepsilon T}{4} \,\Big|\, \mathcal{L}\right] \\
&\overset{(b)}{\leq} \sum_{i=M+1}^{7M/\varepsilon} \exp\left(\frac{-(\frac{\varepsilon T}{4})^2}{2\sum_{r=1}^{T}(1+\varepsilon/2)^2}\right) \\
&\leq \frac{7M}{\varepsilon}\exp\left(\frac{-(\frac{\varepsilon T}{4})^2}{5T}\right) = \frac{7M}{\varepsilon}\exp\left(-\frac{\varepsilon^2 T}{80}\right),
\end{aligned} \tag{10}
$$

where (a) is due the the union bound and

$$
\frac{\varphi - q - MT}{M(7/\varepsilon - 1)} \geq \frac{-3MT/2}{M(7/\varepsilon - 1)} \geq \frac{-3MT/2}{6M/\varepsilon} = -\frac{\varepsilon T}{4},
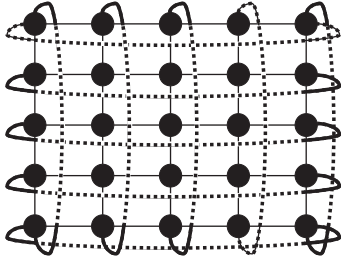$$

Fig. 1. Example: $n \times n$ torus interference graph, where $n = 5$. We scale n to see how the delay scales with n.

and (b) is due to Azuma's inequality on the supermartingale $\{Y^i_s : s \in [t + (i - 1)T, t + iT]\}$. Note that the bound (10) is identical for all $i = M + 1, M + 2, \ldots$.

Bound for (‡): One can bound (‡) similarly as we did for (†). In particular, we obtain

$$\mathbb{P}\left[\max_{MT+2 \leq s \leq 7MT/\varepsilon} \Delta[t + s, t + 7MT/\varepsilon] \geq \varphi \,|\, \mathcal{L}\right]$$
$$\leq \frac{133M}{\varepsilon^3} \exp\left(-\frac{\varepsilon^2 T}{80}\right) + \frac{18}{\varepsilon^2} \exp\left(-\frac{\varphi\varepsilon}{9}\right). \quad (11)$$

We provide the complete proof of (11) in the Appendix. Combining (10) and (11) leads to the desired conclusion of Lemma 5.2.

Remark: In Lemma 5.2, we consider $7MT/\varepsilon$ time spacing. Our main interest is to show that *asymptotic* queue lengths become a constant order with respect to the network size. In this sense, the important part is to space "enough time" itself. We does not optimize the time spacing and chooses a proper value, which is $7MT/\epsilon$, inducing simpler coefficients of many terms in our proofs.

VI. SIMULATION RESULTS

In this section, we provide simulation results to verify our analytical findings. We consider two interference graph topologies, Torus and Random, in the following two subsections.

A. Torus Topology

Setup: As depicted in Fig. 1, we first study grid-like graphs, which have been popularly used as representative interference graphs. We consider $n \times n$ torus interference graphs composed of total number of n^2 links ($n \geq 2$), where every link has exactly four interfering neighbors.

We vary the scale size n of the torus graph to see how the delay performance (measured by the queue lengths) behaves. We assume that the arrival rate of every link is the same, thus we simply denote the common arrival rate by λ. Instead of finding appropriate fugacity values in advance, we adaptively adjust fugacities by running the algorithm in [8], which utilizes both of empirical average arrival rates and service rates. In all plots except for Fig. 2(d) (which shows the delay performance for varying loads), we use a reasonably high load $\rho = 80\%$ i.e., $\lambda = 0.5 \times 0.8 = 0.4$, where note that $\lambda = 0.5$ corresponds to the boundary of the capacity region. We set the access probability a_v to 0.2 for decision schedules in the *delayed* CSMA, which maximizes the expected number of nodes in a decision schedule $\mathcal{D}(t)$ under the torus topology. We run 10^6 slots for all simulations.

Compared Algorithm: U-CSMA [10]: We also compare the *delayed* CSMA for various T with U-CSMA [10] that is provably delay-optimal in torus graphs. We comment that the original U-CSMA has been developed under a continuous time framework, which means that a procedure of determining decision schedules is not required. Recall that the key idea of U-CSMA lies in resetting the underlying CSMA Markov chain with a given period, say P. A naive candidate discrete-time version of U-CSMA is just the *delayed* CSMA with $T = 1$, that restarts itself every P time slots. However, our simulation experience of such a version of U-CSMA showed that it performs significantly poorly, because the step of finding decision schedules makes U-CSMA start with a very small number of active links at each reset period, which leads to scheduling only a small number of links during each period. This problem can be relaxed by enlarging P, but then it weakens the effect of resetting. Thus, for a fair comparison, we employ an almost "ideal" version of U-CSMA which magically has a good decision schedule, i.e., its number of links in a decision schedule is close to that of a maximum independent set.

Results: We first verify the delay optimality proved in Theorem 4.1. Fig. 2(a) shows the average queue-size vs. the scale size n in the torus. Indeed, we observe that while queue size of the "conventional CSMA" (i.e., $T = 1$ in the *delayed* CSMA) linearly increases with n, that of the *delayed* CSMA for $T = 2$ does not increase with n, i.e., $O(1)$ delay. We comment that at least in the tested torus topology, just $T = 2$ suffices to achieve very low delays, even for highly large scales, e.g., $n = 20$ (thus 400 links). This is highly valuable in practice, because a small T significantly simplifies practical implementation. Fig. 2(b) shows the average queue size traces over time for various values of $T = 1, 2, 4, 6$, where we observe significant difference between $T = 1$ and other values of T's, whereas marginal difference among $T > 2$ is observed.

We now compare the *delayed* CSMA with U-CSMA in Fig. 2(c), where we have plotted the queue size traces for $T = 1, 2$. For U-CSMA, as indicated in [10], we use $P = 100$ for the reset period. First, we observe that U-CSMA outperforms the *delayed* CSMA with $T = 1$, which is because resetting reduces the correlations among schedules over time. However, we also see that the *delayed* CSMA with $T = 2$ is significantly better than U-CSMA (recall that y-axis is log-scaled). This demonstrates that an approach of weakening temporal correlations by running multiple Markov chains leads to much higher delay performance gain than that of re-starting new Markov chains periodically. The simulation result for different values of loads in Fig. 2(d) also shows a significant decrease of delay by the *delayed* CSMA with $T = 2$. We further report 'snapshots' of active 'even' and 'odd' links under the delayed CSMA with $T = 1, 2$ in Fig. 3, where the torus graph is bipartite with edges between even and odd nodes (or links). They also explain why $T = 2$ is better than $T = 1$ for the delay performance of the *delayed* CSMA: when $T = 1$ and $T = 2$, the sets of active links change little and much, respectively, between two consecutive times.

B. Random Topology

Setup: To simulate a scenario closer to practical ones, we construct a random geometric graph, as shown in Fig. 4(a). To do that, we set up a grid square (1000×1000 m^2), and put 30

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

8   IEEE/ACM TRANSACTIONS ON NETWORKING
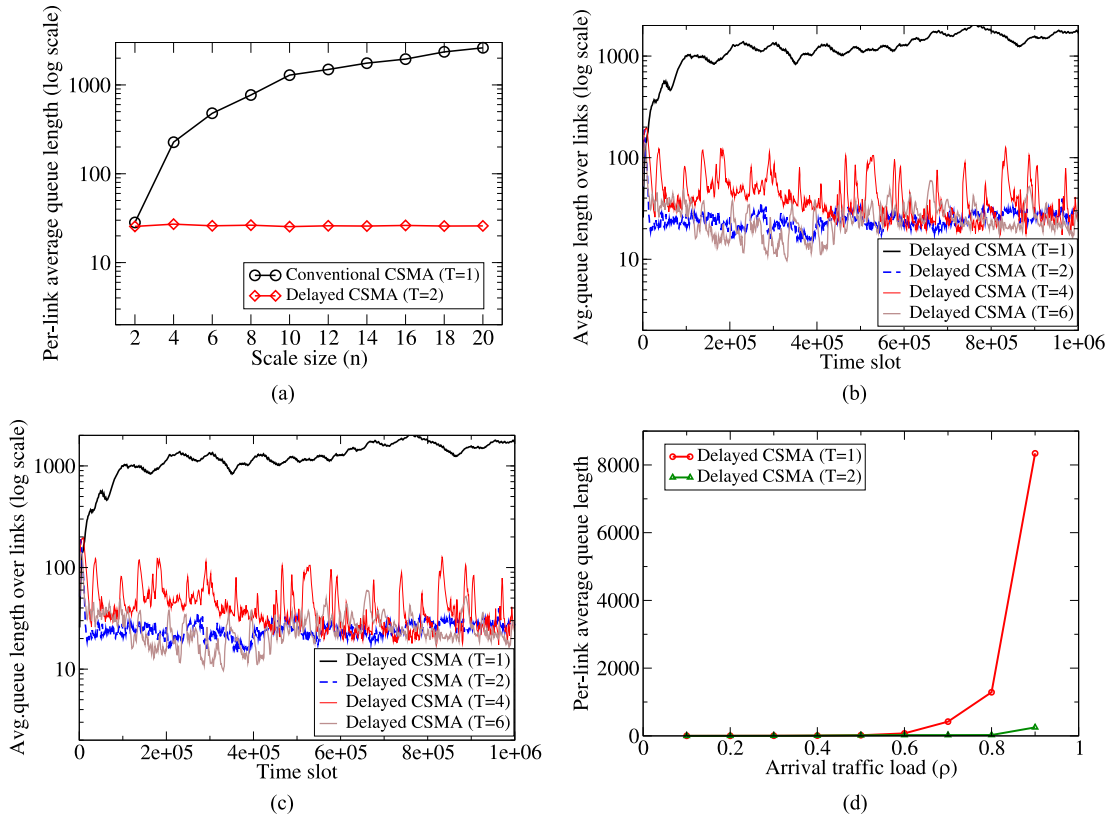
Fig. 2.   Simulation results for the torus interference graphs. (a) Average queue length (over both time slots and links) vs. scale size n in $n \times n$ torus graphs: y-axis is log-scaled. (b) Average queue length traces (over links) for the *delayed* CSMAs with different values of T ($n = 10$). (c) Average queue length traces (over links) for the *delayed* CSMA with $T = 1, 2$ and U-CSMA ($n = 10$). (d) Average queue length (over both time slots and links) for different values of loads ($n = 10$).
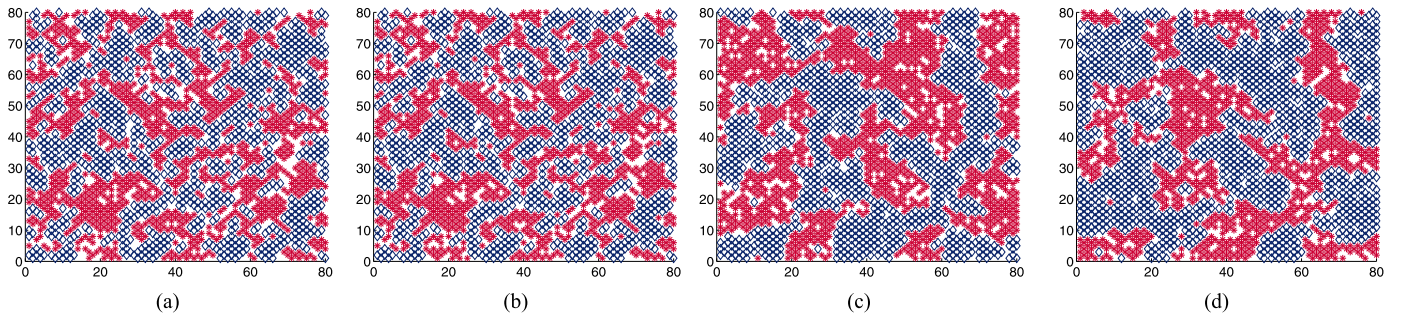

Fig. 3.   Snapshots of the schedules of the *delayed* CSMA with $T = 1$ ((a) and (b)) and $T = 2$ ((c) and (d)) for 80×80 torus. (A magenta star represents an even active link and an indigo diamond represents an odd active link.). (a) Time $t = 1999$ for $T = 1$. (b) Time $t = 2000$ for $T = 1$. (c) Time $t = 1999$ for $T = 2$. (d) Time $t = 2000$ for $T = 2$.

wireless nodes uniformly at random in the square. The transmission range of each node is set to 250 m which is a typical range of an outdoor WiFi node. Two nodes are connected if they are placed within the transmission range. Through this procedure, for the random nodes in Fig. 4(a), 57 wireless links are created. Regarding the interference relationship, we adopt an one-hop interference model so that any two links which share a node interfere with each other. Fig. 4(b) shows the corresponding interference graph of the topology Fig. 4(a), having 57 vertices (i.e., links in the original graph) and 188 edges. The simulation setting is similar to that of the torus topology, except that we use $a_v = 0.15$ in choosing decision schedules. Note that the number of nodes selected in the decision schedule, which is controlled by a_v in Section IV-A, should be large enough to reduce the mixing time of the service process. If all nodes have the same

degree (e.g., torus topology in the previous section), the value of a_v maximizing the number of selected nodes, is the inverse of the degree. However, the degree of every node is different in a general topology like the one here. Thus, we set a_v to be just the inverse of the average degree of vertices, which is 0.15. We remark that other values of a_v show similar trends to what is presented in this subsection, unless its value is not too extreme.

Results: We show the simulation results for the effect of different values of T for varying arrival rates. As expected, our simulation results reveal that large T leads to less delay. First, Fig. 5(a) displays the average queue size traces over time for $T = 1, 2, 10, 50, 500$ when the arrival rate is 0.05. If we use just $T = 2$, we can reduce the average delay by about 30% ~ 50% compared to the conventional CSMA (notice that Y-axis of Fig. 5(a) is log-scale), and we have the diminishing results in

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

YUN *et al.*: DELAY OPTIMAL CSMA WITH LINEAR VIRTUAL CHANNELS UNDER A GENERAL TOPOLOGY 9
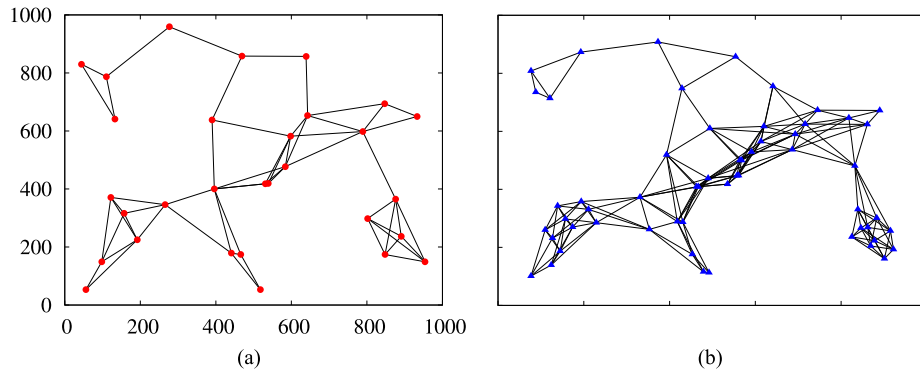

Fig. 4. A wireless network topology and its interference graph. (a) A random wireless network topology. We generate 30 nodes, and connect two nodes u and v when u's transmission range includes v, where 57 links are created. (b) The interference graph for the random topology, where 188 edges (i.e., pairwise interference relationships) are generated.
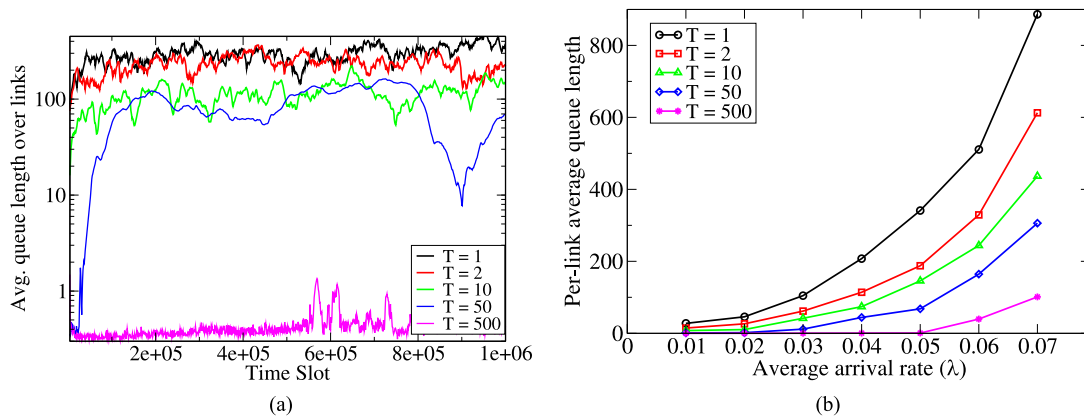

Fig. 5. Simulation results for the interference graph (Fig. 4(b)). (a) Average queue length traces (over links) for the *delayed* CSMAs with different values of T ($\lambda = 0.05$). (b) Average queue length (over both time slots and links) for different values of arrival rates λ.

terms of delay reduction, as T grows. This result is consistent with that of the Torus topology, where due to its regular pattern, only $T = 2$ was enough to achieve significant delay reduction. Fig. 5(b) shows the average queue length with varying arrival rates λ. Those curves start to rise steeply after some threshold value of λ, and such threshold λ becomes larger if we utilize more T in the *delayed* CSMA.

VII. CONCLUSION

In this paper, we have addressed the open question for designing a CSMA algorithm that is both throughput and delay optimal for general wireless network topology. We proved that one of the throughput-optimal CSMA algorithms based on the notion of virtual channels proposed in the literature has the per-link $O(1)$-asymptotic-delay for general wireless network topology if the number of virtual channels has the logarithmic order of the mixing time of the underlying CSMA Markov chain, which is at most linear order of the network size. The significance of our result lies in the proof of the existence of a scheduling policy that achieves optimality in both throughput and delay, operating in a fully distributed manner.

APPENDIX

Proof of (8): The first term in (8) is bounded by:

$$g(T, Q(t) - 1) = \frac{140M}{\varepsilon^3} \exp\left(-\frac{\varepsilon^2 T}{80}\right)$$
$$+ \frac{18}{\varepsilon^2} \exp\left(-\frac{(Q(t) - 1)\varepsilon}{9}\right) \leq 1/4 \quad (12)$$

where the last inequality is because T^*, $Q(t)$ are large enough so that $T > T^* = \Omega\left(\frac{1}{\varepsilon^3} \log M\right)$ and $Q(t) > C = \Omega(1/\varepsilon^2)$. On the other hand, for the second term in (8), we observe that

$$3 \sum_{i=1}^{49M^2T^2/\varepsilon^2} g\left(T, Q(t) + \sqrt{i}\right)$$

$$\leq q\, 3 \sum_{i=1}^{49M^2T^2/\varepsilon^2} \frac{140M}{\varepsilon^3} \exp\left(-\frac{\varepsilon^2 T}{80}\right)$$
$$+ 3 \sum_{i=1}^{\infty} \frac{18}{\varepsilon^2} \exp\left(-\frac{(Q(t) + \sqrt{i})\varepsilon}{9}\right)$$

where one can check that the first and second summations in the above inequality can be made arbitrarily small by choosing large T^*, $Q(t)$ with $T > T^* = \Omega\left(\frac{1}{\varepsilon^3} \log M\right)$ and $Q(t) > C = \Omega(1/\varepsilon^2)$, respectively, i.e., $\sum_{i=1}^{49M^2T^2/\varepsilon^2} g\left(T, Q(t) + \sqrt{i}\right) \leq \frac{1}{4}$. Combining the above inequality with (12) leads to the proof of (8).

Proof of (9)'s Supermartingale: For given i and all $s \in [t + (i - 1)T + 1, t + iT]$,

$$\mathbb{E}[Y_s^i | Y_{s-1}^i, \ldots, Y_{t+(i-1)T}^i]$$
$$= \mathbb{E}[Y_s^i | \underbrace{\Delta[s - 1, s - 1], \ldots, \Delta[t + (i - 1)T, t + (i - 1)T]}_{\delta}]$$
$$= Y_{s-1}^i + \varepsilon/2 + \mathbb{E}[\Delta[s, s] | \delta] \leq Y_{s-1}^i + \varepsilon/2 - \varepsilon/2$$
$$= Y_{s-1}^i.$$

The last inequality can be explained as follows. Note that s is the point of time at which time longer than the mixing time has elapsed since time t. Hence, $\Delta[s,s], \ldots, \Delta[t+(i-1)T, t+(i-1)T]$ become independent when $\mathcal{L}$ is given. Further, since we take the mixing time as $M = M(\varepsilon/2)$, $\mathbb{E}[D[s,s]] \geq (\lambda + \varepsilon) - \varepsilon/2 = \lambda + \varepsilon/2$, which in turn implies that $\mathbb{E}[\Delta[s,s]] \leq \lambda - (\lambda + \varepsilon/2) = -\varepsilon/2$. Therefore, $\{Y^i_{t+(i-1)T}, \ldots Y^i_{t+iT}\}$ forms a supermartingale.

Proof of (11): We first characterize two inequalities that are used to prove (11). For $\zeta > -\varepsilon T/2$ and $i = M+1, M+2, \ldots$, it follows that

$$\mathbb{P}\left[\Delta[t+(i-1)T+1, t+iT] > \zeta \mid \mathcal{L}\right]$$
$$\leq \mathbb{P}\left[Y^i_{t+iT} - Y^i_{t+(i-1)T} \geq \frac{\varepsilon T}{2} + \zeta \mid \mathcal{L}\right]$$
$$\leq \exp\left(\frac{-(\varepsilon T/2 + \zeta)^2}{2\sum_{i=1}^{T}(1+\varepsilon/2)^2}\right)$$
$$\leq \exp\left(\frac{-(\varepsilon T/2 + \zeta)^2}{5T}\right), \tag{13}$$

where we use Azuma's inequality on the supermartingale $\{Y^i_{t+(i-1)T+1}, \ldots Y^i_{t+iT}\}$. Similarly, for $\zeta > 0$, we have that

$$\mathbb{P}\left[\max_{(i-1)T+1 < s \leq iT} \Delta[t+s, t+iT] \geq \zeta \mid \mathcal{L}\right]$$
$$\leq \sum_{s=t+(i-1)T+1}^{t+iT} \mathbb{P}\left[\Delta[t+s, t+iT] \geq \zeta \mid \mathcal{L}\right]$$
$$= \sum_{s=t+(i-1)T+1}^{t+iT} \mathbb{P}\left[Y^i_{t+iT} - Y^i_{s-1} \geq \frac{(t+iT-s+1)\varepsilon}{2} + \zeta \mid \mathcal{L}\right]$$
$$\leq \sum_{s=t+(i-1)T+1}^{t+iT} \exp\left(\frac{-\{(t+iT-s+1)\varepsilon/2 + \zeta\}^2}{2(1+\varepsilon/2)^2(t+iT-s+1)}\right)$$
$$= \sum_{k=1}^{T} \exp\left(\frac{-\{\frac{\varepsilon k}{2} + \zeta\}^2}{2(1+\frac{\varepsilon}{2})^2 k}\right) \leq \sum_{k=1}^{\infty} \exp\left(\frac{-\{\frac{\varepsilon k}{2} + \zeta\}^2}{\frac{4(1+\frac{\varepsilon}{2})^2}{\varepsilon}\{\frac{\varepsilon k}{2} + \zeta\}}\right)$$
$$= \sum_{k=1}^{\infty} \exp\left(\frac{-\frac{\varepsilon^2 k}{2} - \zeta\varepsilon}{4(1+\frac{\varepsilon}{2})^2}\right) \leq \sum_{k=1}^{\infty} \exp\left(\frac{-\frac{\varepsilon^2 k}{2} - \zeta\varepsilon}{9}\right)$$
$$= \exp\left(-\frac{\zeta\varepsilon}{9}\right) \cdot \frac{1}{e^{\frac{\varepsilon^2}{18}} - 1} \leq \frac{18}{\varepsilon^2} \exp\left(-\frac{\zeta\varepsilon}{9}\right). \tag{14}$$

Using (13) and (14), we have:

$$\mathbb{P}\left[\max_{MT+2 \leq s \leq 7MT/\varepsilon} \Delta[t+s, t+7MT/\varepsilon] \geq \varphi \mid \mathcal{L}\right]$$
$$= 1 - \mathbb{P}\left[\underbrace{\max_{MT+2 \leq s \leq 7MT/\varepsilon} \Delta[t+s, t+7MT/\varepsilon] < \varphi}_{(\star)} \mid \mathcal{L}\right]$$
$$\overset{(a)}{\leq} 1 - \mathbb{P}\left[\underbrace{\max_{(i-1)T < s \leq iT} \Delta[t+s, t+iT] < \varphi + \frac{\varepsilon}{4}T}_{(\star\star)}\right]$$

for $i = M+1, \ldots, 7M/\varepsilon - 1$

and $\underbrace{\max_{(7M/\varepsilon-1)T < s \leq 7MT/\varepsilon} \Delta[t+s, t+7MT/\varepsilon] < \varphi}_{(\star\star\star)}$

and $\underbrace{\Delta[t+(i-1)T+1, t+iT] \leq -\frac{\varepsilon}{4}T}_{(\star\star\star\star)}$

for $i = M+1, \ldots, 7M/\varepsilon \mid \mathcal{L}\Big]$

$$\leq \sum_{i=M+1}^{7M/\varepsilon-1} \mathbb{P}\left[\max_{(i-1)T < s \leq iT} \Delta[t+s, t+iT] \geq \varphi + \frac{\varepsilon}{4}T \mid \mathcal{L}\right]$$
$$+ \mathbb{P}\left[\max_{(7M/\varepsilon-1)T < s \leq 7MT/\varepsilon} \Delta[t+s, t+7MT/\varepsilon] \geq \varphi \mid \mathcal{L}\right]$$
$$+ \sum_{i=M+1}^{7M/\varepsilon} \mathbb{P}\left[\Delta[t+(i-1)T+1, t+iT] > -\frac{\varepsilon}{4}T \mid \mathcal{L}\right]$$
$$\leq \left(\frac{7}{\varepsilon}M - M - 1\right)\frac{18}{\varepsilon^2}\exp\left(\frac{-(\varphi + \frac{\varepsilon}{4}T)\varepsilon}{9}\right)$$
$$+ \frac{18}{\varepsilon^2}\exp\left(\frac{-\varphi\varepsilon}{9}\right) + M\left(\frac{7}{\varepsilon} - 1\right)\exp\left(\frac{-(\frac{\varepsilon T}{2} - \frac{\varepsilon}{4}T)^2}{5T}\right)$$
$$\leq \frac{126M}{\varepsilon^3}\exp\left(\frac{-\varepsilon^2 T}{36}\right) + \frac{18}{\varepsilon^2}\exp\left(\frac{-\varphi\varepsilon}{9}\right)$$
$$+ \frac{7}{\varepsilon}M\exp\left(\frac{-\varepsilon^2 T}{80}\right)$$
$$\leq \frac{133M}{\varepsilon^3}\exp\left(-\frac{\varepsilon^2 T}{80}\right) + \frac{18}{\varepsilon^2}\exp\left(-\frac{\varphi\varepsilon}{9}\right).$$

In (a), we consider a time interval $[t+MT, t+7MT/\varepsilon]$ that are partitioned into two disjoint sub-intervals $I_1 = [t+MT, t+(7M/\varepsilon - 1)T]$ and $I_2 = [t+(7M/\varepsilon - 1)T + 1, t+7MT/\varepsilon]$, and we let s^* be the maximization point in $(\star)$. Then, if $s^* \in I_1$, $(\star\star)$ and $(\star\star\star\star)$ imply $(\star)$, and if $s^* \in I_2$, $(\star\star\star)$ is equivalent to $(\star)$. This completes the proof of (11).

REFERENCES

[1] 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.

[2] L. Tassiulas, "Linear complexity algorithms for maximum throughput in radio networks and input queued switches," in *Proc. IEEE INFOCOM*, 1998, pp. 533–539.

[3] P. Giaccone, B. Prabhakar, and D. Shah, "Randomized scheduling algorithms for high-aggregate bandwidth switches," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 4, pp. 546–559, May 2003.

[4] E. Modiano, D. Shah, and G. Zussman, "Maximizing throughput in wireless networks via gossiping," in *Proc. ACM SIGMETRICS*, New York, NY, USA, 2006, pp. 27–38.

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

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

[7] 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.

[8] 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.

[9] 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.

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

YUN *et al.*: DELAY OPTIMAL CSMA WITH LINEAR VIRTUAL CHANNELS UNDER A GENERAL TOPOLOGY

11

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

[11] J. Kwak, C.-H. Lee, and D. Y. Eun, "A high-order Markov chain based scheduling algorithm for low delay in CSMA networks a general lower bound for mixing of single-site dynamics on graphs," in *Proc. IEEE INFOCOM*, 2014, pp. 1662–1670.

[12] N. Berger, C. Kenyon, E. Mossel, and Y. Peres, "Glauber dynamics on trees and hyperbolic graphs," *Probab. Theory Related Fields*, vol. 131, no. 3, pp. 311–340, 2005.

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

[14] 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.

[15] V. Subramanian and M. Alanyali, "Delay performance of CSMA in networks with bounded degree conflict graphs," in *Proc. IEEE ISIT*, Aug. 2011, pp. 2373–2377.

[16] K.-K. Lam, C.-K. Chau, M. Chen, and S.-C. Liew, "Mixing time and temporal starvation of general CSMA networks with multiple frequency agility," in *Proc. IEEE ISIT*, 2012, pp. 2676–2680.

[17] P.-K. Huang and X. Lin, "Improving the delay performance of CSMA algorithms: A virtual multi-channel approach," in *Proc. IEEE INFOCOM*, 2013, pp. 2598–2606.

[18] M. J. Neely, "Delay analysis for maximal scheduling in wireless networks with bursty traffic," in *Proc. IEEE INFOCOM*, 2008, pp. 385–393.

[19] Y. Yi, A. Proutière, and M. Chiang, "Complexity in wireless scheduling: Impact and tradeoffs," in *Proc. ACM MobiHoc*, 2008, pp. 33–42.

[20] M. J. Neely, E. Modiano, and Y. S. Cheng, "Logarithmic delay for N × N packet switches under the crossbar constraint," *IEEE/ACM Trans. Netw.*, vol. 15, no. 3, pp. 657–668, Jun. 2007.

[21] D. Shah, N. Walton, and Y. Zhong, "Optimal queue-size scaling in switched networks," *Perform. Eval. Rev.*, vol. 40, no. 1, pp. 17–28, 2012.

[22] D. Shah, D. N. C. Tse, and J. N. Tsitsiklis, "Hardness of low delay network scheduling," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 7810–7817, Dec. 2011.

[23] H. Li and N. Vaidya, "Optimal CSMA-based wireless communication with worst-case delay and non-uniform sizes," in *Proc. IEEE INFOCOM*, 2014, pp. 2526–2534.

[24] 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.

[25] N. Hegde and A. Proutiere, "Simulation-based optimization algorithms with applications to dynamic spectrum access," in *Proc. CISS*, 2012, pp. 1–6.

Dongmyung Lee received his B.S. degree in electrical and electronic engineering from Yonsei University, Korea and Ph.D. degree in industrial engineering at Seoul National University, Korea. From 2013 to 2014, he was a Post-Doctoral Researcher in the Department of Electrical Engineering at KAIST. He is currently with Mobile Communications Division, Samsung Electronics. His research interests include network economics, performance analysis, and management strategy.
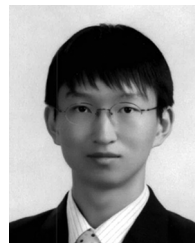
Se-Young Yun is currently a Visiting Researcher at Microsoft Research, Cambridge from June 2015. Before joining Microsoft Research, Cambridge, he spent one year at MSR-INRIA joint center as a Post-Doc Researcher. He received the B.S. and Ph.D. in electrical engineering from the KAIST, Daejeon, Republic of Korea, in 2006 and 2012. From 2012 to 2013, he was a Postdoc Researcher at the same department. He also served as a Postdoc Researcher at KTH, Stockholm, Sweden during one year (2013–2014). 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 in ACM MobiHoc 2013.

Jinswoo Shin is currently an Assistant Professor at the Department of Electrical Engineering at KAIST, Korea. His major research interest is to develop advanced algorithms for large-scale data and network analytics. He obtained the Ph.D. degree from Massachusetts Institute of Technology in 2010 and B.S. degrees (in Math and CS) from Seoul National University in 2001. After spending two years (2010–2012) at Algorithms & Randomness Center, Georgia Institute of Technology, one year (2012–2013) at Business Analytics and Mathematical Sciences Department, IBM T. J. Watson Research, he started teaching at KAIST in Fall 2013. He received the Kennneth C. Sevcik Award at SIGMETRICS 2009, George M. Sprowls Award 2010, Best Paper Award at MobiHoc 2013, Best Publication Award from INFORMS Applied Probability Society 2013, Bloomberg Scientific Research Award 2015 and ACM SIGMETRICS Rising Star Award 2015.

Donggyu Yun received the B.S. and the M.S. in electrical engineering from the KAIST, Daejeon, South Korea, in 2009 and 2011. He is currently working toward the Ph.D degree in electrical engineering at KAIST. His research interests include future internet, wireless networks, applying machine learning theory to networks.
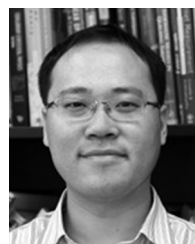
Yung Yi received his B.S. and the M.S. in the School of Computer Science and Engineering from Seoul National University, South Korea in 1997 and 1999, respectively, and his Ph.D. in the Department of Electrical and Computer Engineering at the University of Texas at Austin in 2006. From 2006 to 2008, he was a Post-Doctoral Research Associate in the Department of Electrical Engineering at Princeton University. Now, he is an Associate Professor at the Department of Electrical Engineering at KAIST, South 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 IEEE/ACM TRANSACTIONS ON NETWORKING, *Journal of Communication Networks*, and Elsevier *Computer Communications* Journal.