A Fast-CSMA Algorithm for Deadline-Constrained Scheduling over Wireless Fading Channels

Bin Li and Atilla Eryilmaz
Department of Electrical and Computer Engineering
The Ohio State University, Columbus, Ohio 43210 USA.
Email: {lib, eryilmaz}@ece.osu.edu

***Abstract*—Recently, low-complexity and distributed Carrier Sense Multiple Access (CSMA)-based scheduling algorithms have attracted extensive interest due to their throughput-optimal characteristics in general network topologies. However, these algorithms are not well-suited for serving real-time traffic under time-varying channel conditions for two reasons: (1) the mixing time of the underlying CSMA Markov Chain grows with the size of the network, which, for large networks, generates unacceptable delay for deadline-constrained traffic; (2) since the dynamic CSMA parameters are influenced by the arrival and channel state processes, the underlying CSMA Markov Chain may not converge to a steady-state under strict deadline constraints and fading channel conditions.**

In this paper, we attack the problem of distributed scheduling for serving real-time traffic over time-varying channels. Specifically, we consider fully-connected topologies with independently fading channels (which can model cellular networks) in which flows with short-term deadline constraints and long-term drop rate requirements are served. To that end, we first characterize the maximal set of satisfiable arrival processes for this system and, then, propose a Fast-CSMA (FCSMA) policy that is shown to be optimal in supporting any real-time traffic that is within the maximal satisfiable set. These theoretical results are further validated through simulations to demonstrate the relative efficiency of the FCSMA policy compared to some of the existing CSMA-based algorithms.

I. Introduction

Wireless networks are expected to serve real-time traffic, such as video or voice applications, generated by a large number of users over potentially fading channels. These constraints and requirements, together with the limited shared resources, generate a strong need for distributed algorithms that can efficiently utilize the available resources while maintaining high quality-of-service for the real-time applications. Yet, the strict short-term deadline constraints and long-term drop rate requirements associated with most real-time applications complicate the development of provably good distributed solutions.

In the recent years, there has been an increasing understanding on the modeling and service of such real-time traffic in wireless networks (e.g., [4], [5], [6], [2]). However, existing works in this domain assume centralized controllers, and hence are not suitable for distributed operation in large-scale networks. In a separate line of work, it has also been shown that CSMA-based distributed scheduling (e.g., [7], [12], [3], [13]) can maximize long-term average throughput for general wireless topologies. However, these results also do not apply to strictly deadline-constrained traffic that we target, since their throughput-optimality relies: (i) on the convergence time of the underlying Markov Chain to its steady-state, which grows with the size of the network; and (ii) on relatively stationary conditions in which the CSMA parameters do not change significantly over time so that the instantaneous service rate distribution can stay close to the stationary distribution. Both of these conditions are violated in our context: (i) packets of deadline constrained traffic are likely to be dropped before the CSMA-based algorithm converges to its steady-state; and (ii) the time-varying fading creates significant variations on the CSMA parameters, in which case the instantaneous service rate distribution cannot closely track the stationary distribution.

While achieving low delay via distributed scheduling in general topologies is a difficult task (see [14]), in a related work [9] that focuses on grid topologies, the authors have designed an Unlocking CSMA (UCSMA) algorithm with both maximum throughput and order optimal average delay performance, which shows promise for distributed scheduling in special topologies. However, UCSMA also does not directly apply to deadline-constrained traffic since its measure of delay is on average. Moveover, it is not clear how existing CSMA or UCSMA implementations will perform under fading channel conditions.

With this motivation, in this work, we address the problem of distributed scheduling in fully connected networks (e.g., Cellular network, WLAN) for serving real-time traffic over independently fading channels. Our contributions are:

- In Section III-A, we characterize the maximal set of satisfiable real-time traffic characteristics as a function of their drop rate requirements and channel statistics.
- In Section III-B, we propose an FCSMA algorithm that differs from existing CSMA policies in its design principle: rather than evolving over the set of schedules to reach a favorable steady-state distribution, the FCSMA policy aims to quickly reach one of a set of favorable schedules and stick to it for a duration related to deadline constraints of the application. While the performance of the former strategy is tied to the mixing-time of a Markov Chain, the performance of our strategy is tied to the absorption time, and hence, yields significant advantage for strictly deadline-constrained flows.
- In Theorem 1, we prove that the FCSMA policy is optimal in the sense that it can satisfy the deadline and drop rate requirements for any real-time traffic within the characterized maximal satisfiable set.
- In Section IV, we compare the performance of FCSMA

with some of the existing CSMA policies under different scenarios, both to validate the theoretical claims, and to demonstrate the performance gains due to our proposed strategy.

II. System Model

We consider a fully-connected wireless network topology where N users contend for data transmission over a single channel that is independently block fading for each user. We assume that the time scale of block fading is the same as the duration of the deadline constraint, and thus uniformly called as a *slot*. We also assume that all links start transmission at the beginning of each time slot. We capture the channel fading over link l via $C_l[t]$, which measures the maximum amount of service available in slot t, if scheduled. We assume that $\mathbf{C}[t] = (C_l[t])_{l=1}^N$ are independently distributed random variables over links and identically distributed over time. Yet, due to interference constraints, at most one link can be scheduled for service in each slot. We use a binary variable $S_l[t]$ to denote whether the link l is served at slot t, where $S_l[t] = 1$ if the link l can be served at slot t and $S_l[t] = 0$, otherwise.

Each packet has a delay bound of 1 time slot, which means that if a packet cannot be served during the slot it arrives, it will be dropped. In this context of fully-connected network, we associate each real-time flow with a link, and hence use these two terms interchangeably. Let $A_l[t]$ denote the number of packets arriving at link l in slot t that are independently distributed over links and identically distributed over time with mean λ_l, and $A_l[t] \leq A_{\max}$ for some $A_{\max} < \infty$. Each link has a maximum allowable drop rate $\rho_l \lambda_l$, where $\rho_l \in (0, 1)$ is the maximum fraction of packets that can be dropped at link l. For example, $\rho_l = 0.1$ means that at most 10% of packets can be dropped at link l on average. Under above setup, we define our stochastic control problem (SCP) as follows:

Definition 1: (SCP)

$$\underset{\{S[t]\}_{t\geq 1}}{\text{Maximize}} \quad 1 \tag{1}$$

$$\text{Subject to} \quad \overline{\lambda_l}(1-\rho_l) \leq \underline{\mu_l}, \forall l \tag{2}$$

$$\sum_l S_l[t] \leq 1 \tag{3}$$

$$S_l[t] \in \{0, 1\}, \forall l, \forall t \geq 1 \tag{4}$$

where

$$\overline{\lambda_l} = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[A_l[t]] \tag{5}$$

$$\underline{\mu_l} = \liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\min\{S_l[t]C_l[t], A_l[t]\}] \tag{6}$$

In the above maximization problem: (2) indicates that the provided average service rates satisfy the drop rate requirements of the real-time traffic; (3) indicates that at most one link is served at each slot.

Normally, it is difficult to solve SCP directly. Instead, we use the technique in [11] to introduce a virtual queue $X_l[t]$ for each link l to track the number of dropped packets at slot t. Specifically, the number of packets arriving at virtual queue l at the end of slot t is denoted as $R_l[t]$, which is equal to $A_l[t] - \min\{S_l[t]C_l[t], A_l[t]\}$. We use $I_l[t]$ to denote the service for virtual queue l at the end of the slot t with mean $\rho_l \lambda_l$, and $I_l[t] \leq I_{\max}$ for some $I_{\max} < \infty$. Further, we let $U_l[t]$ denote the unused service for queue l at the end of slot t, which is upper-bounded by $I_{\max}$. Then, the evolution of virtual queue is as follows:

$$X_l[t+1] = X_l[t] + R_l[t] - I_l[t] + U_l[t], \quad l = 1, \cdots, N.$$

In the rest of the paper, we consider the class of stationary policies $\mathcal{G}$ that select $\mathbf{S}[t]$ as a function of $(\mathbf{X}[t], \mathbf{A}[t], \mathbf{C}[t])$, which, then, forms a Markov Chain. If this Markov Chain is positive recurrent, then the average drop rate will meet the required constraint automatically (see [1]). Accordingly, we call an algorithm *optimal* if it can make this Markov Chain positive recurrent for any arrival rate vector within the maximal satisfiable region that we will characterize in the next section.

III. FCSMA Algorithm for Throughput Optimality

In this section, we first study the maximal satisfiable region given the drop rate and channel statistics. Then, we propose an optimal FCSMA algorithm.

A. Maximal Satisfiable Region

Consider the class $\mathcal{G}$ of stationary policies that base their scheduling decision on the observed vector $(\mathbf{X}[t], \mathbf{A}[t], \mathbf{C}[t])$ at slot t. The next lemma establishes a condition that is necessary for stabilizing the system.

Lemma 1: If there is a policy $G_0 \in \mathcal{G}$ that can stabilize the virtual queue $\mathbf{X}[t]$, then there exist non-negative numbers $\alpha(\mathbf{a}, \mathbf{c}; \mathbf{s})$ such that

$$\sum_{\mathbf{s}\in\mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) = 1 \tag{7}$$

$$\sum_{\mathbf{a}} P_{\mathbf{A}}(\mathbf{a}) \sum_{\mathbf{c}} P_{\mathbf{C}}(\mathbf{c}) \sum_{\mathbf{s}\in\mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \min\{\mathbf{s} \circ \mathbf{c}, \mathbf{a}\} > \lambda \circ (\mathbf{1} - \rho) \tag{8}$$

where $(\mathbf{A} \circ \mathbf{B})_i = A_i B_i$ denotes Hadamard product, $P_{\mathbf{A}}(\mathbf{a}) = P(\mathbf{A}[t] = \mathbf{a})$ and $P_{\mathbf{C}}(\mathbf{c}) = P(\mathbf{C}[t] = \mathbf{c})$.

The proof is almost the same as [15] and hence is omitted here. Note that the left hand side of inequality (8) is the total average service provided for each link during one time slot; while $\lambda \circ (1 - \rho)$ is the total average amount of data packets at each link that need to be served. Thus, to the meet the constraint of drop rate, (8) should be satisfied. We define maximal satisfiable region $\Lambda(\rho)$ as follows:

$$\Lambda(\rho) = \{\mathbf{A} : \exists \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \geq 0, \text{such that both (7) and (8)satisfy}\}$$

B. FCSMA algorithm

Before we present and analyze our proposed FCSMA algorithm, we define a set of functions (also see [8]) that allows flexibility in the design and implementation of the algorithm.

$\mathcal{F}$:= set of non-negative, nondecreasing and differentiable functions $f(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{x\to\infty} f(x) = \infty$.

$\mathcal{B} := \{f \in \mathcal{F}\colon \lim_{x\to\infty} \frac{f(x+a)}{f(x)} = 1, \text{ for any } a \in \mathbb{R}\}.$

The examples of functions that are in class $\mathcal{B}$ are $f(x) = \log x$, $f(x) = x$ or $f(x) = e^{\sqrt{x}}$. $f(x) = e^x$ is not in class $\mathcal{B}$.

Definition 2 (FCSMA Algorithm): At the beginning of each time slot t, each link l independently generates an exponentially distributed random variable with mean $f(X_l[t])^{-\min\{C_l[t],A_l[t]\}}$, and starts transmitting after this random duration unless it senses another transmission before. The link that grabs the channel transmits its packets until the end of the slot. If there are no packets awaiting in the link l, it transmits dummy packets to occupy the channel.

Remarks: (1) The absorption time of FCMSA algorithm at slot t is exponentially distributed with mean $\frac{1}{\sum_{j=1}^{N} f(X_j[t])^{\min\{C_j[t],A_j[t]\}}}$, which quickly becomes negligibly small as we demonstrate in next section.

(2) The parameter of FCSMA policy quickly adapts to arrival and channel state processes. Due to its fast absorption time, FCSMA policy yields significant advantages over existing CSMA policies evolving slowly to the steady-state. In FCSMA, the probability of serving link l in slot t will be:

$$\pi_l = \frac{f(X_l[t])^{\min\{C_l[t],A_l[t]\}}}{Z}(1 - \frac{1}{Z}) \tag{9}$$

where $Z = \sum_{j=1}^{N} f(X_j[t])^{\min\{C_j[t],A_j[t]\}}$. In equation (9), $\frac{f(X_l[t])^{\min\{C_l[t],A_l[t]\}}}{Z}$ is the probability that link l successfully grabs the channel; while $1 - \frac{1}{Z}$ is the average remaining time for serving the packet at slot t given that link l grabs the channel. Let $W^*[t] = \max_l \log f(X_l[t]) \min\{C_l[t], A_l[t]\}$. The following lemma establishes the fact that FCSMA policy picks a link with the weight close to maximum weight with high probability when the maximum weight $W^*[t]$ is large enough.

Lemma 2: Given $\epsilon > 0$ and $\zeta > 0$, $\exists \overline{W} < \infty$, such that if $W^*[t] > \overline{W}$, then FCSMA policy picks a link k satisfying

$$P\{W_k[t] \geq (1-\epsilon)W^*[t]\} \geq 1 - \zeta$$

which also implies

$$\begin{aligned}&\mathbb{E}[W_k[t]1_{\{W^*[t]\geq W\}}|\mathbf{A}[t], \mathbf{C}[t], \mathbf{X}[t]] \\ &\geq (1-\epsilon)(1-\zeta)W^*[t]1_{\{W^*[t]\geq W\}}\end{aligned} \tag{10}$$

where $W_k[t] = \log f(X_k[t]) \min\{C_k[t], A_k[t]\}$.

Proof: Define

$$\mathcal{X} = \{l : \log f(X_l[t]) \min\{C_l[t], Q_l[t]\} < (1-\epsilon)W^*[t]\}$$

Then,

$$\begin{aligned}\pi(\mathcal{X}) &:= \sum_{l\in\mathcal{X}} \pi_l \\ &\leq \sum_{l\in\mathcal{X}} \frac{\exp(\log f(X_l[t]) \min\{C_l[t], Q_l[t]\})}{\sum_{j=1}^{n} \exp(\log f(X_j[t]) \min\{C_j[t], Q_j[t]\})} \\ &< \frac{|\mathcal{X}| \exp((1-\epsilon)W^*[t])}{\sum_{j=1}^{n} \exp(\log f(X_j[t]) \min\{C_j[t], Q_j[t]\})} \\ &\leq \frac{N \exp((1-\epsilon)W^*[t])}{\exp(W^*[t])} = \frac{N}{\exp(\epsilon W^*[t])}\end{aligned} \tag{11}$$

The first inequality in (11) follows the fact that $1 - \frac{1}{Z} \leq 1$. Thus, $\exists \overline{W} < \infty$ such that $W^*[t] > \overline{W}$ implies $\pi(\mathcal{X}) < \zeta$. ■

Under certain conditions for the function f, we can establish the optimality of FCSMA algorithm.

Theorem 1: FCSMA is optimal if $\log f \in \mathcal{B}$ and $f(0) \geq 1$.

Proof: See the Appendix for the proof. ■

Remarks: The optimality of FCSMA is preserved even when the slope of function f is low, which is easier to be implemented in practice.

IV. Simulation Results

In this section, we perform simulations to validate the optimality of the proposed FCSMA policy with deadline constraint 1 time slot in both fading and non-fading channels. In the simulation, there are $N = 10$ links. All links require that the maximum fraction of dropping packets cannot exceed $\rho = 0.2$. The number of arrivals in each slot follows Bernoulli distribution. For the simulations of a fading channel, all links suffer from the ON-OFF channel fading independently with probability $p = 0.9$ that the channel is available in each time slot. Under this setup, we can use the same technique in paper [16] to get the maximal satisfiable region: $\Gamma = \{\lambda : N(1-\rho)\lambda < 1 - (1-p\lambda)^N\}$. Through numerical calculation, we can get $\lambda < 0.051$ in non-fading channel and $\lambda < 0.03$ in fading channel. We compare our proposed FCSMA policy with $f(x) = e^x$ with QCSMA algorithm [12] with the weight $X_l[t] \min\{C_l[t], A_l[t]\}$ (In our setup, QCSMA algorithm with the weight $\log\log(X_l[t] \min\{C_l[t], A_l[t]\} + e)$ has much worse performance than that with $X_l[t] \min\{C_l[t], A_l[t]\}$). To that end, we divide each time slot into M mini-slots. In FCSMA policy, if the link contends for the channel successfully, it will occupy that channel in the rest of time slot; while in QCSMA policy, each link contends for the channel and transmits the data in 1 mini-slot. Here, we don't consider the overhead that the QCSMA policy needs to contend for the channel, which will greatly degrade its performance.

From Figure 1 and 2, we can observe that the average virtual queue length grows very fast under the QCSMA policy with $M = 1$ while the average queue length of FCSMA always stays at a low level. The reason for the poor performance of QCSMA scheme in deadline-constrained application is that the underlying Markov chain is controlled by the arrival and channel state processes. If the running time of QCSMA policy has the same time scale with the deadline of the packet, this Markov chain cannot converge to the steady-state. However, FCMSA policy can quickly lock into one state and exhibits good performance, which is shown in Theorem 1 to be optimal if we carefully choose the parameters. In addition, as M increases, the performance of QCSMA improves. The reason is that the underlying Markov chain has enough time to converge to the steady-state and thus yields better performance. Recall that FCSMA policy waits for random duration before accessing the channel, this random duration can be arbitrarily small when the number of links increases and the virtual queue length is high. We can see from simulations that FCSMA policy has almost the same performance as that in steady state.

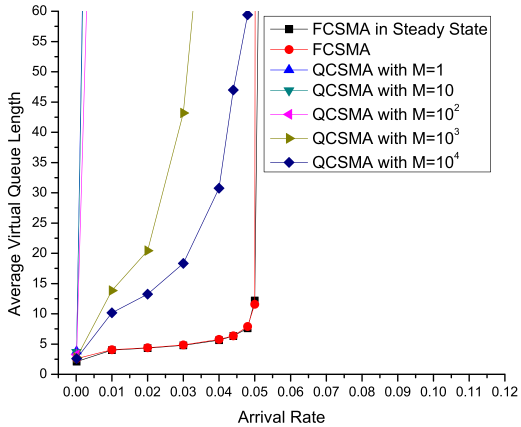

Fig. 1. Performance of FCSMA and QCSMA over non-fading channel

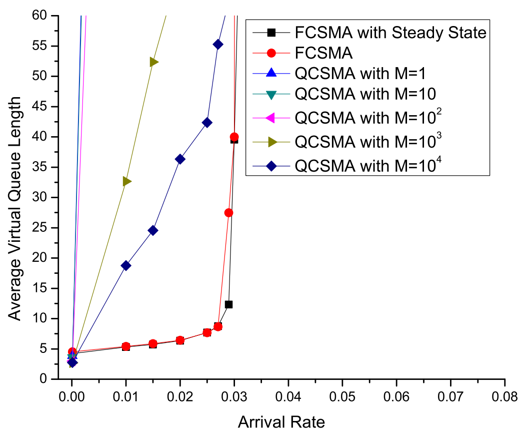

Fig. 2. Performance of FCSMA and QCSMA over fading channel

V. Conclusions

In this paper, we first characterized the maximal satisfiable set of arrival processes given the drop rate and channel statistics and then proposed a provably optimal distributed FCSMA policy for scheduling deadline-constrained traffic over fading channel. We validated the performance of FCSMA policy by comparing it with existing CSMA policies through simulations. We assumed that the time scale of channel fading is the same as the duration of the deadline constraint, which is not always the case in practical wireless networks. We will relax this assumption in our future work. Also, we will try to explore scheduling algorithms for real-time traffic over fading channel in multi-hop network topologies.

VI. Acknowledgement

This work was supported in part by DTRA Grant HDTRA 1-08-1-0016, and NSF Awards: CAREER-CNS-0953515 and CCF-0916664.

VII. Appendix
Proof of Theorem 1

Proof: Let $g(x) = \log f(x)$. Consider the Lyapunov function $V(\mathbf{X}) := \sum_{l=1}^{N} h(X_l)$, where $h'(x) = g(x)$. Then

$$\begin{aligned}\Delta V :&= \mathbb{E}\left[V(\mathbf{X}[t+1]) - V(\mathbf{X}[t])|\mathbf{X}[t]=\mathbf{X}\right]\\ &= \sum_{l=1}^{N}\mathbb{E}\left[(h(X_l[t+1]) - h(X_l[t]))|\mathbf{X}[t]=\mathbf{X}\right]\end{aligned}$$

By the mean-value theorem, we have $h(X_l[t+1]) - h(X_l[t]) = g(X_l')(X_l[t+1] - X_l[t]) = g(X_l')(R_l[t] - I_l[t] + U_l[t])$, where X_l' lies between $X_l[t]$ and $X_l[t+1]$. Hence, we get

$$\begin{aligned}\Delta V &= \sum_{l=1}^{N}\mathbb{E}\left[g(X_l')(R_l[t] - I_l[t] + U_l[t])|\mathbf{X}[t]=\mathbf{X}\right]\\ &= \underbrace{\sum_{l=1}^{N}\mathbb{E}\left[g(X_l')U_l[t]|\mathbf{X}[t]\right]}_{=:\Delta V_1} + \underbrace{\sum_{l=1}^{N}\mathbb{E}\left[g(X_l')(R_l[t] - I_l[t])|\mathbf{X}[t]\right]}_{=:\Delta V_2}\end{aligned}$$

For ΔV_1, if $X_l[t] = X_l \geq I_{\max}$, then $U_l[t] = 0$. If $X_l[t] = X_l < I_{\max}$, then $U_l[t] \leq I_{\max}$. But in this case, $X_l[t+1] \leq (I_{\max} + A_{\max})$. Hence, $g(X_l') \leq g(I_{\max} + A_{\max}) < \infty$. Thus,

$$\begin{aligned}\Delta V_1 &= \sum_{l=1}^{N}\mathbb{E}\left[g(R_l)U_l[t]\mathbf{1}_{\{X_l < I_{\max}\}}|\mathbf{X}[t]=\mathbf{X}\right]\\ &\leq N I_{\max} g(I_{\max} + A_{\max})\end{aligned} \tag{12}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function.

Next, let's focus on ΔV_2. We know that $g(X_l') = g(X_l[t] + a_l)$ $(|a_l| \leq A_{\max})$. According to the definition of function $g \in \mathcal{B}$, given $\beta > 0$, there exists $M > 0$, such that for any $X_l[t] = X_l > M$, we have $\left|\frac{g(X_l')}{g(X_l)} - 1\right| < \beta$, that is,

$$(1-\beta)g(X_l) < g(X_l') < (1+\beta)g(X_l) \tag{13}$$

Thus, we have

$$\begin{aligned}&g(X_l')(R_l[t] - I_l[t])\\ =&g(X_l')\left[(R_l[t] - I_l[t])_+ - (R_l[t] - I_l[t])_-\right]\\ <&(1+\beta)g(X_l)(R_l[t] - I_l[t])_+\\ &- (1-\beta)g(X_l)(R_l[t] - I_l[t])_-\\ =&g(X_l)(R_l[t] - I_l[t]) + \beta g(X_l)\left|R_l[t] - I_l[t]\right|\\ \leq&g(X_l)(R_l[t] - I_l[t]) + \beta A_{\max} g(X_l)\end{aligned} \tag{14}$$

where $(x)_+ = \max\{x, 0\}$, $(x)_- = -\min\{x, 0\}$ and $|R_l[t] - I_l[t]| \leq |A_l[t]| \leq A_{\max}$. Thus, we divide ΔV_2 into two parts:

$$\begin{aligned}\Delta V_2 = &\underbrace{\sum_{l=1}^{N}\mathbb{E}\left[g(X_l')(R_l[t] - I_l[t])\mathbf{1}_{\{X_l > M\}}|\mathbf{X}[k]=\mathbf{X}\right]}_{=:\Delta V_3}\\ &+ \underbrace{\sum_{l=1}^{N}\mathbb{E}\left[g(X_l')(R_l[t] - I_l[t])\mathbf{1}_{\{X_l \leq M\}}|\mathbf{X}[t]=\mathbf{X}\right]}_{=:\Delta V_4}\end{aligned}$$

For ΔV_3, by using (14), we have

$$\Delta V_3 \leq \sum_{l=1}^{N} \mathbb{E}\left[g(X_l)(R_l[t] - I_l[t])\mathbf{1}_{\{X_l > M\}}|\mathbf{X}[t] = \mathbf{X}\right] + \sum_{l=1}^{N} \beta A_{\max} g(X_l)\mathbf{1}_{\{X_l > M\}}$$
$$= \underbrace{\sum_{l=1}^{N} \mathbb{E}[g(X_l)(A_l[t] - I_l[t])\mathbf{1}_{\{X_l > M\}}|\mathbf{X}[t] = \mathbf{X}]}_{=:L_1} \underbrace{- \mathbb{E}[\sum_{l=1}^{N} W_l^F[t]1_{\{X_l > M\}}|\mathbf{X}[t]]}_{=:L_2} + \sum_{l=1}^{N} \beta A_{\max} g(X_l)\mathbf{1}_{\{X_l > M\}}$$

where $W_l^F[t] = \log f(X_l[t]) \min\{C_l[t]S_l^F[t], A_l[t]\}$ and $\mathbf{S}^F[t]$ denotes the schedule chosen by FCSMA with $S_k^F[t] = 1$. Next, we will explore the upper bound of L_1 by using Lemma 1 and give the lower bound of L_2 by the Lemma 2.

First, let's focus on L_1. By Lemma 1, there exist non-negative numbers $\alpha(a, c; s)$ satisfying (7) and for a $\delta > 0$ small enough, we have

$$\sum_{\mathbf{a}} P_{\mathbf{A}}(\mathbf{a}) \sum_{\mathbf{c}} P_{\mathbf{C}}(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \min\{s_l c_l, a_l\} \geq \lambda_l(1 - \rho_l) + \delta \quad (15)$$

Let $W_l = g(X_l) \min\{s_l c_l, a_l\}$. In the following proof, we can also write the maximum weight $W^*[t] = \sum_{l=1}^{N} W_l^*[t]$, where $W_l^*[t] = \log f(X_l[t]) \min\{C_l[t]S_l^*[t], A_l[t]\}$ and optimal schedule $\mathbf{S}^*[t] = \arg\max_{\mathbf{S} \in \mathcal{S}} \sum_{l=1}^{N} W_l^*[t]$. By using (15), we have

$$L_1 = \sum_{l=1}^{N} g(X_l)\lambda_l(1 - \rho_l)1_{\{X_l > M\}}$$
$$\leq \sum_{\mathbf{a}} P_{\mathbf{A}}(\mathbf{a}) \sum_{\mathbf{c}} P_{\mathbf{C}}(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l 1_{\{X_l > M\}} - \delta \sum_{l=1}^{N} g(X_l)1_{\{X_l > M\}}$$
$$= \sum_{\mathbf{a}} P_{\mathbf{A}}(\mathbf{a}) \sum_{\mathbf{c}} P_{\mathbf{C}}(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l 1_{\{X_l > M, W^*[t] > \overline{W}\}}$$
$$+ \sum_{\mathbf{a}} P_{\mathbf{A}}(\mathbf{a}) \sum_{\mathbf{c}} P_{\mathbf{C}}(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l 1_{\{X_l > M, W^*[t] \leq \overline{W}\}}$$
$$- \delta \sum_{l=1}^{N} g(X_l)1_{\{X_l > M\}}$$
$$\leq \sum_{\mathbf{a}} P_{\mathbf{A}}(\mathbf{a}) \sum_{\mathbf{c}} P_{\mathbf{C}}(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l 1_{\{X_l > M, W^*[t] > \overline{W}\}} + \overline{W} - \delta \sum_{l=1}^{N} g(X_l)1_{\{X_l > M\}} \quad (16)$$

Second, let's consider L_2. Since

$$(1 - \epsilon)(1 - \zeta)\mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t] = \mathbf{X}]$$
$$\leq (1 - \epsilon)(1 - \zeta)\mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{W^*[t] > \overline{W}\}}|\mathbf{X}[t] = \mathbf{X}]$$
$$\leq \mathbb{E}[\sum_{l=1}^{N} W_l^F[t]1_{\{W^*[t] > \overline{W}\}}|\mathbf{X}[t] = \mathbf{X}] \text{(By Lemma 2)}$$
$$= \mathbb{E}[\sum_{l=1}^{N} W_l^F[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t] = \mathbf{X}] + \mathbb{E}[\sum_{l=1}^{N} W_l^F[t]1_{\{X_l \leq M, W^*[t] > \overline{W}\}}|\mathbf{X}[t] = \mathbf{X}]$$
$$\leq \mathbb{E}[\sum_{l=1}^{N} W_l^F[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t] = \mathbf{X}] + NA_{\max}g(M)$$

L_2 becomes

$$L_2 \geq \mathbb{E}[\sum_{l=1}^{N} W_l^F[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t]]$$
$$\geq (1 - \epsilon)(1 - \zeta)\mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t]] - NA_{\max}g(M) \quad (17)$$

Thus, by using (16) and (17), ΔV_3 becomes

$$\Delta V_3 \leq \sum_{\mathbf{a}} P_A(\mathbf{a}) \sum_{\mathbf{c}} P_C(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l 1_{\{X_l > M, W^*[t] > \overline{W}\}}$$
$$- \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t]]$$
$$+ (\epsilon + \zeta - \epsilon\zeta)\mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l > M, W^*[t] > \overline{W}\}}|\mathbf{X}[t]]$$
$$+ \overline{W} + NA_{\max}g(M) - \delta \sum_{l=1}^{N} g(X_l)1_{\{X_l > M\}}$$
$$+ \beta A_{\max} \sum_{l=1}^{N} g(X_l)1_{\{X_l > M\}} \quad (18)$$

Since

$$\sum_{\mathbf{a}} P_A(\mathbf{a}) \sum_{\mathbf{c}} P_C(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l - \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]|\mathbf{X}[t]]$$
$$= \sum_{\mathbf{a}} P_A(\mathbf{a}) \sum_{\mathbf{c}} P_C(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l$$
$$- \sum_{\mathbf{a}} P_A(\mathbf{a}) \sum_{\mathbf{c}} P_C(\mathbf{c}) \sum_{\mathbf{s} \in \mathcal{S}} \alpha(\mathbf{a}, \mathbf{c}; \mathbf{s}) \sum_{l=1}^{N} W_l^*$$
$$\leq 0 \quad (19)$$

Thus, we have

$$\sum_{\mathbf{a}} P_A(\mathbf{a}) \sum_{\mathbf{c}} P_C(\mathbf{c}) \sum_{\mathbf{s}\in\mathcal{S}} \alpha(\mathbf{a},\mathbf{c};\mathbf{s}) \sum_{l=1}^{N} W_l 1_{\{X_l>M, W^*[t]>\overline{W}\}}$$
$$\leq \sum_{\mathbf{a}} P_A(\mathbf{a}) \sum_{\mathbf{c}} P_C(\mathbf{c}) \sum_{\mathbf{s}\in\mathcal{S}} \alpha(\mathbf{a},\mathbf{c};\mathbf{s}) \sum_{l=1}^{N} W_l$$
$$\leq \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]|\mathbf{X}[t]=\mathbf{X}] \text{(By using (19))}$$
$$= \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l>M\}}|\mathbf{X}[t]] + \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l\leq M\}}|\mathbf{X}[t]]$$
$$\leq \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l>M, W^*[t]>\overline{W}\}}|\mathbf{X}[t]=\mathbf{X}]$$
$$+ \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l>M, W^*[t]\leq\overline{W}\}}|\mathbf{X}[t]=\mathbf{X}] + NA_{\max}g(M)$$
$$\leq \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l>M, W^*[t]>\overline{W}\}}|\mathbf{X}[t]] + \overline{W} + NA_{\max}g(M) \tag{20}$$

In addition, we have

$$\mathbb{E}[\sum_{l=1}^{N} W_l^*[t]1_{\{X_l>M, W^*[t]>\overline{W}\}}|X[t]=X]$$
$$\leq \mathbb{E}[\sum_{l=1}^{N} W_l^*[t]|X[t]=X]$$
$$\leq A_{\max}\sum_{l=1}^{N} g(X_l)$$
$$= A_{\max}\sum_{l=1}^{N} g(X_l)1_{\{X_l>M\}} + A_{\max}\sum_{l=1}^{N} g(X_l)1_{\{X_l\leq M\}}$$
$$\leq A_{\max}\sum_{l=1}^{N} g(X_l)1_{\{X_l>M\}} + NA_{\max}g(M) \tag{21}$$

then, by using (20) and (21), we have

$$\Delta V_3 \leq -\gamma \sum_{l=1}^{N} g(X_l)1_{\{X_l>M\}} + D_1 \tag{22}$$

where $D_1 = 2\overline{W} + (2+\epsilon+\zeta-\epsilon\zeta)NA_{\max}g(M)$ and $\gamma = \delta - \beta A_{\max} - A_{\max}(\epsilon+\zeta-\epsilon\zeta)$. We can choose β, ϵ, ζ small enough such that $\gamma > 0$.

For ΔV_4, we have

$$\Delta V_4 \leq \sum_{l=1}^{N} \mathbb{E}\left[g(X_l')R_l[t]|\mathbf{X}[t]=\mathbf{X}\right]\mathbf{1}_{\{X_l\leq M\}}$$
$$\leq \sum_{l=1}^{N} \mathbb{E}\left[g(X_l')A_l[t]|\mathbf{X}[t]=\mathbf{X}\right]\mathbf{1}_{\{X_l\leq M\}}$$
$$\leq NA_{\max}g(M+A_{\max})$$

Thus, we get

$$\Delta V < -\gamma \sum_{l=1}^{N} g(X_l)\mathbf{1}_{\{X_l>M\}} + D$$
$$\leq -\gamma \sum_{l=1}^{N} g(X_l) + E \tag{23}$$

where $D := NI_{\max}g(I_{\max}+A_{\max}) + D_1 + NA_{\max}g(M+A_{\max}) < \infty$ and $E := D + N\gamma g(M)$. Hence, by the Lyapunov Drift theorem [11], we have $\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{N}\mathbb{E}[g(X_l[t])] \leq \frac{E}{\gamma} < \infty$, which implies stability-in-the mean and thus the Markov Chain is positive recurrent [10]. ■

References

[1] J. G. Dai. On the positive harris recurrence for multiclass queueing networks: A unified approach via fluid limit models. *Annals of Applied Probability*, pages 49–77, 1995.

[2] H. Gangammanavar and A. Eryilmaz. Dynamic coding and rate-control for serving deadline-constrained traffic over fading channels. In *Proc. IEEE International Symposium on Information Theory. (ISIT)*, Austin, TX, June 2010.

[3] J. Ghaderi and R. Srikant. On the design of efficient csma algorithms for wireless networks. In *Proc. IEEE International Conference on Decision and Control. (CDC)*, Atlanta, GA, December 2010.

[4] I. Hou, V. Borkar, and P. R. Kumar. A theory of qos for wireless. In *Proc. IEEE International Conference on Computer Communications. (INFOCOM)*, Rio de Janeiro, Brazil, April 2009.

[5] I. Hou and P. R. Kumar. Scheduling heterogeneous real-time traffic over fading wireless channels. In *Proc. IEEE International Conference on Computer Communications. (INFOCOM)*, San Diego, CA, March 2010.

[6] J. Jaramillo and R. Srikant. Optimal scheduling for fair resource allocation in ad hoc networks with elastic and inelastic traffic. In *Proc. IEEE International Conference on Computer Communications. (INFOCOM)*, San Diego, CA, March 2010.

[7] L. Jiang and J. Walrand. A csma distributed algorithm for throughput and utility maximization in wireless networks. In *Proc. Allerton Conference on Communication, Control, and Computing (Allerton'2008)*, Monticello, Illinois, September 2008.

[8] B. Li and A. Eryilmaz. On the limitation of randomization for queue-length-based scheduling in wireless networks. In *Proc. IEEE International Conference on Computer Communications. (INFOCOM)*, Shanghai, China, April 2011.

[9] M. Lotfinezhad and P. Marbach. Throughput-optimal random access with order-optimal delay. In *Proc. IEEE International Conference on Computer Communications. (INFOCOM)*, Shanghai, China, April 2011.

[10] S. Meyn and R. Tweedie. Criteria for stability of markovian processes i: Discrete time chains. *Advances in Applied Probability*, 24:542–574, 1992.

[11] M. Neely. *Stochastic Network Optimization with Aapplication to Communication and Queueing Systems*. Morgan & Claypool, 2010.

[12] 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 International Conference on Computer Communications. (INFOCOM)*, mini-Conference, San Diego, CA, March 2010.

[13] S. Rajagopalan, D. Shah, and J. Shin. Network adiabatic theorem: an efficient randomized protocol for contention resolution. In *Proc. IEEE International Joint Conference on Measurement and Modeling of Computer Systems. (SIGMETRICS)*, Seattle, WA, June 2009.

[14] D. Shah, D. Tse, and J. Tsitsiklis. Hardness of low delay network scheduling. *Submitted to IEEE Transactions on Information Theory*, 2009.

[15] L. Tassiulas. Scheduling and performance limits of networks with constantly varying topology. *IEEE Transactions on Information Theory*, 43:1067–1073, May 1997.

[16] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Transactions on Information Theory*, 39(2):466–478, 1993.