On the Large-Deviations Optimality of Scheduling Policies Minimizing the Drift of a Lyapunov Function

Xiaojun Lin and V. J. Venkataramanan,*
School of ECE, Purdue University, West Lafayette, IN 47907
Email: {linx,vvenkat}@ecn.purdue.edu

Abstract

We show that for a large class of scheduling algorithms, when the algorithm minimizes the drift of a Lyapunov function, the algorithm is optimal in maximizing the asymptotic decay-rate of the probability that the Lyapunov function value exceeds a large threshold. The result in this paper extends our prior results to the important and practically-useful case when the Lyapunov function is not linear in scale, in which case the evolution of the fluid-sample-paths becomes more difficult to track. We use the notion of generalized fluid-sample-paths to address this difficulty, and provide easy-to-verify conditions for checking the large-deviations optimality of scheduling algorithms. As an immediate application of the result, we show that the log-rule is optimal in maximizing the asymptotic decay-rate of the probability that the sum queue exceeds a threshold B.

1 Introduction

In this paper we are interested in link scheduling algorithms for wireless networks supporting delay-sensitive applications. In many cases, the performance objective of these applications can be mapped to a bound on the

*This work has been partially supported by NSF grants CNS-0643145, CNS-0626703, CNS-0721484, and a grant from Purdue Research Foundation.

queue-overflow probability [1–6]. Specifically, in order to meet delay constraints with high probability, we would like to ensure that the probability with which some function of the global queue-length vector exceeds an overflow threshold is below a small value. For example, such a function of the queue-length vector could be the maximum queue length among all users, or the sum of the queue length over all users. Often, a closed-form solution of the queue length distribution is not available. In that case, we could instead use large-deviations theory [7] to study the asymptotic decay-rate of the queue-overflow probability, as the overflow threshold increases to infinity [1–6]. A larger decay-rate may then be interpreted as better delay performance.

Unfortunately, due to both the radio interference and the time-varying channel conditions in wireless systems, even the large-deviation decay-rate can be difficult to characterize. Specifically, in order to minimize the queue-overflow probability, it is often necessary to use queue-length-based link scheduling algorithms, which compute the link schedule at each time based on the current queue backlog vector [1,2]. However, for such queue-length-based scheduling algorithms, computing the asymptotic decay-rate of the queue-overflow probability involves a multi-dimensional calculus-of-variations problem that is very difficult to track [1, 2]. Recently, there have been some progress in using Lyapunov functions to deal with this difficulty [4]. Through this new approach, the form of scheduling algorithms that maximize the asymptotic decay-rate of the probability of some specific form of overflow event is characterized [4, 5]. See also the related results in [3]. However, the proof techniques in these papers tend to be quite involved.

In this paper, we would like to establish a simpler and more general result of the following type.

Statement 1: *If an algorithm minimizes the drift of a Lyapunov function at every time, then such an algorithm is optimal in the sense that it maximizes the asymptotic decay-rate of the probability that the Lyapunov function value exceeds a threshold, as the threshold approaches infinity.*

Note that many queue-length-based scheduling algorithms are designed by minimizing the drift of some Lyapunov functions. For example, the max-weight algorithm minimizes the drift of the Lyapunov function $V(\vec{X}) = \sum_i X_i^2$ where X_i is the queue length of user i. Similarly, the α-algorithm in [4, 5] minimizes the drift of the Lyapunov function $V(\vec{X}) = \sum_i X_i^{\alpha+1}$. Hence, if the above result indeed holds, it will allow us to easily conclude the large-deviations optimality of a large class of link scheduling algorithms. Further, it will help us to search for the optimal scheduling algorithm by choosing the appropriate Lyapunov function.

In our prior work [6], we have established Statement 1 for Lyapunov functions that satisfy the following conditions. First, the Lyapunov function $V(\vec{X})$ must be linear in scale. In other words, if the global queue-length vector $\vec{X}$ is multiplied by a positive scalar β, then $V(\beta\vec{X}) = \beta V(\vec{X})$. Second, the Lyapunov function must be convex. These conditions are satisfied by Lyapunov functions of the form $V_\alpha(\vec{X}) = (\sum_i X_i^\alpha)^{1/\alpha}$. Note that as $\alpha \to \infty$, $V_\alpha(\vec{X}) \to \max_i X_i$. Hence, we can conclude that, as $\alpha \to \infty$, the α-algorithms asymptotically achieve the maximum asymptotic decay-rate of the probability that $\max_i X_i \geq B$. Note that this conclusion recovers the result that was first reported in [4, 5].

However, not all Lyapunov functions (and their corresponding scheduling algorithms) satisfy the afore-mentioned conditions required in [6]. A notable case is the so-called log-rule [8], which has been conjectured to maximize the asymptotic decay-rate of the overflow probability that $\sum_i X_i \geq B$. The log-rule can be viewed as minimizing the drift of the Lyapunov function $V(\vec{X}) = \sum_i (X_i + 1)\log(X_i + 1) - X_i$. This Lyapunov function is not linear in scale. Hence, we cannot use the result of [6] to study its optimality.

In this paper, we extend the result of [6] to more general forms of Lyapunov function, which include the Lyapunov function for the log-rule. We show that under suitable conditions, Statement 1 is true even when the Lyapunov function is not linear in scale. A main difficulty in establishing Statement 1 for Lyapunov functions that are not linear in scale is that the resulting fluid-sample-paths are more difficult to track. We use the recently-developed theory of generalized fluid-sample-paths [3] to address this difficulty. The result of this paper allows us to apply Statement 1 to a much larger class of Lyapunov functions and scheduling algorithms. In particular, as an immediate application, we show that the log-rule maximizes the asymptotic decay-rate of the probability that $\sum_i X_i \geq B$. This result generalizes the result of [8], which was for a log-rule-like scheduling algorithm and was for only two users.

2 The System Model

For simplicity, we focus on the downlink of a single-cell serving multiple users (although the techniques here can also be applied to more general network settings, e.g., multi-hop wireless networks). The wireless channel can be in one of $\mathcal{S}$ states. We assume that time is divided to slots with unit-length. At time-slot t, let $C(t)$ denote the channel state. We assume that the channel states are *i.i.d.* over time, and let $p_j = \mathbf{P}[C(t) = j], j = 1, 2, ..., \mathcal{S}$, denote the probability that the channel state is j at time t. The base-station serves N users. Let $A_i(t)$ denote the number of packets for user i that arrive at the base-station at time-slot t. We assume that $A_i(t)$ are *i.i.d.* over time, and are independent across users. We further assume that $A_i(t)$ is bounded for all users i and all time-slots t. Define $\lambda_i \triangleq E[A_i(t)]$ and $\vec{\lambda} = (\lambda_1, \ldots, \lambda_N)$. We assume that $\vec{\lambda}$ belongs to the interior of the capacity region [9], and hence the system can be stabilized by some scheduling policy. Due to interference, at each time-slot the base-station can only serve packets for one user. Let F_j^i denote the rate that the base-station can serve user i when the channel state is j, if the base-station chooses to serve user i. Let $U(t)$ denote the index of the user that the base-station chooses to serve at time-slot t. Then the evolution of the queue backlog for user i can be written as:

$$X_i(t+1) = \left[X_i(t) + A_i(t) - \sum_{j=1}^{\mathcal{S}} F_j^i \mathbf{1}_{\{C(t)=j, U(t)=i\}} \right]^+ .$$

Let $\vec{X} = [X_i, i = 1, ..., N]$. Let $\tilde{V}(\vec{X})$ denote a given non-negative and component-wise non-decreasing function of the global queue vector $\vec{X}$. In this paper, we are interested in the asymptotic decay-rate of the probability that $\tilde{V}(\vec{X})$ exceeds some threshold $\tilde{f}(B)$, when the scaling parameter B approaches infinity. In other words, for a particular scheduling policy π under which the system is stationary and ergodic, we are interested in the following quantity:

$$I(\pi) = -\lim_{B\to\infty} \frac{1}{B} \log \mathbf{P}_\pi[\tilde{V}(\vec{X}(0)) \geq \tilde{f}(B)], \quad (1)$$

whenever such a limit exists, where $\mathbf{P}_\pi[\cdot]$ denote the stationary distribution under the scheduling policy π. Further, let I_{opt} denote the maximum value of $I(\pi)$ over all policies. we are interested in finding the scheduling policy that can achieve the maximum decay rate I_{opt}.

Remark: The function $\tilde{f}(B)$ needs to be properly chosen so that the limit in (1) does not become trivial. We will provide more comments on the choice of $\tilde{f}(B)$ at the end of Section 3.

For any $B > 0$ and $T > 0$, define the scaled channel-state process $s_j^B(t)$, scaled arrival process $g_i^B(t)$, and scaled queue process $x_i^B(t)$ as $s_j^B(0) = g_i^B(0) = 0$, $x_i^B(0) = X_i(0)/B$,

$$s_j^B(t) = \frac{1}{B}\sum_{\tau=1}^{Bt} \mathbf{1}_{\{C(\tau)=j\}}, \qquad g_i^B(t) = \frac{1}{B}\sum_{\tau=1}^{Bt} A_i(\tau),$$
$$x_i^B(t) = \frac{1}{B}X_i(Bt),$$

for $t = \frac{m}{B}$, $m = 1, ..., BT$, and by linear interpolation otherwise. Let $\vec{s}^B(t) = [s_j^B(t), j = 1, ..., \mathcal{S}]$, $\vec{g}^B(t) = [g_i^B(t), i = 1, ..., N]$, and $\vec{x}^B(t) = [x_i^B(t), i = 1, ..., N]$. For any $\vec{\phi} = [\phi_j, j = 1, ..., \mathcal{S}] \geq 0$ and $\sum_{j=1}^{\mathcal{S}} \phi_j = 1$, define $H(\vec{\phi}||\vec{p}) = \sum_{j=1}^{\mathcal{S}} \phi_j \log \frac{\phi_j}{p_j}$. (Here we use the convention that $0 \log 0 = 0$.) Further, define

$$L_i(a) = \sup_\theta \left(\theta a - \log \mathbf{E}[\exp(\theta A_i(0))]\right).$$

For any $\vec{a} = [a_i, i = 1, ..., N]$, let $L(\vec{a}) = \sum_{i=1}^{N} L_i(a_i)$. With a suitable choice of the topological space, the sequence of processes $\vec{s}^B(\cdot)$ and $\vec{g}^B(\cdot)$ are known to satisfy sample-path large deviation principles [7, p176] with large-deviation rate-functions given by

$$\begin{aligned} I_s^T(\vec{s}(\cdot)) &= \int_0^T H(\frac{d\vec{s}(t)}{dt}||\vec{p})dt \\ I_g^T(\vec{g}(\cdot)) &= \int_0^T L(\frac{d\vec{g}(t)}{dt})dt, \end{aligned}$$

whenever the processes $\vec{s}(\cdot)$ and $\vec{g}(\cdot)$ are absolute continuous. Finally, for any $(\vec{s}(\cdot), \vec{g}(\cdot))$, define the large-deviations cost over a time interval $[t_1, t_2]$ as

$$J_{[t_1,t_2]}(\vec{s}(\cdot), \vec{g}(\cdot)) = \int_{t_1}^{t_2} H(\frac{d\vec{s}(t)}{dt}||\vec{p}) + L(\frac{d\vec{g}(t)}{dt})dt.$$

3 An Upper Bound on the Asymptotic Decay-Rate of the Queue Overflow Probability

Given any non-negative and component-wise non-decreasing function $\tilde{V}(\vec{X})$ of the global queue-length vector $\vec{X}$, define the following optimization problem for all $B > 0$, $\vec{\phi}$ and $\vec{a}$:

$$\begin{aligned}
\tilde{l}^B(\vec{\phi},\vec{a}) = \min \quad & \tilde{V}(B\vec{X}) \\
\text{subject to} \quad & X_i = [a_i - \sum_{j=1}^{\mathcal{S}} F_j^i u_j^i]^+ \\
& [u_j^i] \geq 0, \quad \sum_{i=1}^{N} u_j^i = \phi_j \\
& \text{for all channel states } j = 1, ..., \mathcal{S}.
\end{aligned}$$

The parameter u_j^i can be interpreted as the long-term fraction of time that the base-station serves user i at state j. The value $\tilde{l}^B(\vec{\phi},\vec{a})$ can then be viewed as the slowest way that $\tilde{V}(\vec{X})$ can grow when the channel-state process and the arrival process satisfy $\frac{d\vec{s}}{dt} = \vec{\phi}$ and $\frac{d\vec{g}}{dt} = \vec{a}$ at all time. For an increasing overflow threshold function $\tilde{f}(B)$, assume that the following limit exists for all $\vec{\phi}$ and $\vec{a}$:

$$\tilde{w}(\vec{\phi},\vec{a}) = \lim_{B\to\infty} \frac{1}{B}\tilde{f}^{-1}(\tilde{l}^B(\vec{\phi},\vec{a})). \quad (2)$$

Roughly speaking, $\tilde{w}(\vec{\phi},\vec{a})$ can be interpreted as the slowest speed of growth of $\tilde{f}^{-1}(\tilde{V}(\vec{X}))$ and hence $1/\tilde{w}(\vec{\phi},\vec{a})$ is the maximum-time $\tilde{V}(\vec{X})$ would take to exceed $\tilde{f}(B)$ when the channel-state process and the arrival process satisfy $\frac{d\vec{s}}{dt} = \vec{\phi}$ and $\frac{d\vec{g}}{dt} = \vec{a}$. Let

$$I_{opt} = \inf_{\tilde{w}(\vec{\phi},\vec{a})>0} \frac{H(\vec{\phi}||\vec{p}) + L(\vec{a})}{\tilde{w}(\vec{\phi},\vec{a})}.$$

We first have the following upper bound on the asymptotic decay-rate of the queue-overflow probability.

Proposition 1 *Assume that the limit in (2) exists for all $\vec{\phi}$ and $\vec{a}$. Then*

$$\liminf_{B\to\infty} \frac{1}{B} \log \mathbf{P}_\pi[\tilde{V}(\vec{X}(0)) \geq \tilde{f}(B)] \geq -I_{opt}.$$

Remark: The function $\tilde{f}(B)$ must be chosen such that the value I_{opt} is not trivial. Roughly speaking, $\tilde{f}(B)$ must be on the same order as $\tilde{V}(B\vec{X})$ when $B \to \infty$. For example, when $\tilde{V}(\vec{X}) = \sum_i X_i^2$, then $\tilde{f}(B)$ may be chosen as $\tilde{f}(B) = B^2$. If $\tilde{f}(B)$ "grows" too fast, it may happen that $\tilde{w}(\vec{\phi}, \vec{a}) = 0$ for all $\vec{\phi}$ and $\vec{a}$. In this case, $I_{opt} = +\infty$, and hence the probability $\mathbf{P}[\tilde{V}(\vec{X}(0)) \geq \tilde{f}(B)]$ decreases super-exponentially to zero as $B \to \infty$. The other extreme is when $\tilde{f}(B)$ "grows" too slowly. Specifically, if for all $\vec{X}$

$$\lim_{B\to\infty} \frac{\tilde{f}^{-1}(\tilde{V}(B\vec{X}))}{B} = +\infty,$$

then $\tilde{w}(\vec{\phi}, \vec{a}) = +\infty$ and $I_{opt} = 0$. In this case, the probability $\mathbf{P}[\tilde{V}(\vec{X}(0)) \geq \tilde{f}(B)]$ may approach a non-zero constant as $B \to \infty$. Neither of these two situations are desirable for an LDP result. To summarize, the suitable choice of $\tilde{f}(B)$ should ensure that:

(a) $\tilde{w}(\vec{\phi}, \vec{a}) > 0$ for some $\vec{\phi}$ and $\vec{a}$.

(b) For some $\vec{X}$, $\lim_{B\to\infty} \frac{\tilde{f}^{-1}(\tilde{V}(B\vec{X}))}{B} < +\infty$.

Proof of Proposition 1 : Fix a small $\epsilon > 0$. By the definition of I_{opt}, there exist $\vec{\phi}_0$ and $\vec{a}_0$ such that

$$\frac{H(\vec{\phi}_0||\vec{p}) + L(\vec{a}_0)}{\tilde{w}(\vec{\phi}_0, \vec{a}_0)} \leq I_{opt} + \epsilon.$$

Let $T_0 = \frac{1}{(1-\epsilon)\tilde{w}(\vec{\phi}_0,\vec{a}_0)}$ and $T = T_0/(1-\epsilon)$. Consider a scaled channel-state process $\vec{s}_0(\cdot)$ and a scaled arrival process $\vec{g}_0(\cdot)$ in the interval $[0, T]$ such that $\vec{s}_0(t) = t\vec{\phi}_0$ and $\vec{g}_0(t) = t\vec{a}_0$. Let $\delta > 0$ be a small number, which we will choose soon. Consider a set Γ of pairs of scaled channel-state process $\vec{s}(\cdot)$ and scaled arrival process $\vec{g}(\cdot)$ such that for each $(\vec{s}(\cdot), \vec{g}(\cdot)) \in \Gamma$, the following holds

$$\sup_{t\in[0,T]} ||\vec{s}(t) - \vec{s}_0(t)|| < \delta, \ \sup_{t\in[0,T]} ||\vec{g}(t) - \vec{g}_0(t)|| < \delta.$$

We will show that with a suitable choice of δ the queue must overflow (in the sense that $\tilde{V}(\vec{X}) > \tilde{f}(B)$) at time BT for any $(\vec{s}(\cdot), \vec{g}(\cdot)) \in \Gamma$. To see this, for any $(\vec{s}(\cdot), \vec{g}(\cdot)) \in \Gamma$, let $\vec{\phi} = (\vec{s}(T) - \vec{s}(0))/T$ and $\vec{a} = (\vec{g}(T) - \vec{g}(0))/T$.

Further, let $\bar{u}_j^i$ be the corresponding fraction of time in the interval $[0, T]$ that user i is served whenever the channel state is j. Note that $\sum_{i=1}^{N} \bar{u}_j^i = 1$ for all states j. Let

$$x_i = \left[a_i - \sum_{j=1}^{\mathcal{S}} \phi_i F_j^i \bar{u}_j^i \right]^+, x_{i,0} = \left[a_{i,0} - \sum_{j=1}^{\mathcal{S}} \phi_{i,0} F_j^i \bar{u}_j^i \right]^+,$$

where $a_i, a_{i,0}, \phi_i$ and $\phi_{i,0}$ are the components of $\vec{a}, \vec{a}_0, \vec{\phi}$, and $\vec{\phi}_0$, respectively. Let $\vec{x} = [x_i]$, and $\vec{x}_0 = [x_{i,0}]$. Then by choosing δ to be sufficiently small, we can ensure that $x_i \geq (1-\epsilon) x_{i,0}$ for all $(\vec{s}(\cdot), \vec{g}(\cdot)) \in \Gamma$. Hence, we have

$$\begin{aligned} V(\vec{X}(BT)) &\geq V(\vec{X}(BT) - \vec{X}(0)) \geq V(BT\vec{x}) \\ &\geq V(BT_0 \vec{x}_0) \geq \tilde{l}^{BT_0}(\vec{\phi}_0, \vec{a}_0). \end{aligned}$$

By the definition of $\tilde{w}(\vec{\phi}, \vec{a})$, there exists B_0 such that for all $B \geq B_0$, we have

$$\tilde{f}^{-1}(\tilde{l}^{BT_0}(\vec{\phi}_0, \vec{a}_0)) \geq (1-\epsilon) BT_0 \tilde{w}(\vec{\phi}_0, \vec{a}_0) = B.$$

Hence, we have

$$\tilde{V}(\vec{X}(BT)) \geq \tilde{l}^{BT_0}(\vec{\phi}_0, \vec{a}_0) \geq \tilde{f}(B).$$

In other words, the queue must overflow at time BT for any $(\vec{s}(\cdot), \vec{g}(\cdot)) \in \Gamma$. Hence,

$$\begin{aligned} & \liminf_{B \to \infty} \frac{1}{B} \log \mathbf{P}_\pi [\tilde{V}(\vec{X}(0)) \geq \tilde{f}(B)] \\ = & \liminf_{B \to \infty} \frac{1}{B} \log \mathbf{P}_\pi [\tilde{V}(\vec{X}(BT)) \geq \tilde{f}(B)] \\ \geq & - \inf_{(\vec{s}(\cdot), \vec{g}(\cdot)) \in \Gamma^o} J_{[0,T]}(\vec{s}(\cdot), \vec{g}(\cdot)) \geq -J_{[0,T]}(\vec{s}_0(\cdot), \vec{g}_0(\cdot)) \\ = & -T(H(\vec{\phi}_0 || \vec{p}) + L(\vec{a}_0)) \geq -\frac{I_{opt} + \epsilon}{(1-\epsilon)^2}. \end{aligned}$$

Since this is true for all $\epsilon > 0$, the result of the proposition then follows by letting $\epsilon \to 0$. *Q.E.D.*

4 A Lower Bound on the Asymptotic Decay-Rate of the Queue-Overflow Probability

Next, we construct a lower bound on the asymptotic decay-rate of the queue overflow probability using a Lyapunov function.

4.1 Lyapunov functions

Often, the stability of the system under a particular scheduling policy π is established through a Lyapunov function. Let $||\vec{X}||$ be an L^p norm with $p \geq 1$. The Lyapunov function $V(\vec{X})$ for a given scheduling policy π is a function that satisfies the following conditions:

Assumption 1 *(a) $V(\vec{X})$ is a continuous function of $\vec{X}$, and $V(\vec{X}) \geq 0$.*

(b) $V(\vec{X}) \to +\infty$ as $||\vec{X}|| \to \infty$.

(c) There exists a large B such that whenever $||\vec{X}|| \geq B$,

$$\mathbf{E}[V(\vec{X}(t+1)) - V(\vec{X}(t))|\vec{X}(t) = \vec{X}] < -\xi, \tag{3}$$

for some $\xi > 0$.

The last condition implies a negative drift of the Lyapunov function. Hence, the system under policy π must be stable. Often, the negative drift is attained when the scheduling policy π chooses a schedule that minimizes the drift $V(\vec{X}(t+1)) - V(\vec{X}(t))$ at each time slot t. Note that if the Lyapunov function is differentiable, then under fairly general assumptions, the drift of the Lyapunov function may be written (for large $\vec{X}(t)$) as

$$\begin{aligned} V(\vec{X}(t+1)) - V(\vec{X}(t)) &= \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} A_i(t) \\ &- \sum_{j=1}^{\mathcal{S}} \mathbf{1}_{\{C(t)=j\}} \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} F_j^i \mathbf{1}_{\{U(t)=i\}} \\ &+ o(\nabla V(\vec{X}(t))), \end{aligned}$$

where $\nabla V(\vec{X})$ is the gradient of $V(\vec{X})$ and is given by $\nabla V(\vec{X}) = \left[\frac{\partial V}{\partial X_i}, i = 1, ..., N\right]$. Hence, ignoring the small-o term, we can define a scheduling policy that minimizes the drift of the Lyapunov function as follows.

Definition 2 *A scheduling policy π is said to minimize the drift of the Lyapunov function $V(\vec{X})$ if at any time t, when the channel state is j, the scheduling policy picks the user i that maximizes the value $\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} F_j^i$.*

With such a scheduling policy, the one-step drift of $V(\vec{X})$ can be further simplified to

$$V(\vec{X}(t+1)) - V(\vec{X}(t)) = \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} A_i(t) \tag{4}$$

$$- \sum_{j=1}^{\mathcal{S}} \mathbf{1}_{\{C(t)=j\}} \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} F_j^i + o(\nabla V(\vec{X}(t))).$$

In this section, we will establish a lower bound on the asymptotic decay-rate of the queue-overflow probability for scheduling policies of the above form, which will then help us prove Statement 1 for this class of scheduling policies. However, for this purpose we need some stronger conditions on the Lyapunov functions. These conditions essentially require that the scheduling policy not only minimizes the *one-step* drift of the Lyapunov function, it must also minimize the drift over each time-interval of length B^η for some $\eta \in (0,1)$, whenever the queue-length vector $\vec{X}(t)$ preceding this time-interval is on the order B. Such drift-minimization must hold *even when compared to another policy that knows the channel-states and arrivals in this time-interval of length B^η in advance.* For this purpose, the condition below essentially requires that the gradient of the Lyapunov function does not change much during such a time-interval of length B^η (please see part (b) of Assumption 2). Since the drift of the Lyapunov function is dependent on its gradient, under these conditions a scheduling policy that minimizes the one-step drift should also (approximately) minimize the B^η-step drift (see Proposition 4 below).

Assumption 2 *(a) $\nabla V(\vec{X}) \geq 0$ for all $\vec{X}$, and*

$$\frac{\partial V}{\partial X_i} \to +\infty \text{ as } X_i \to \infty \text{ for all } i.$$

Further, for all $M > 0$ and i, $\frac{\partial V}{\partial X_i}$ is bounded whenever $X_i \leq M$.

(b) *For any $\epsilon > 0$, $M > 0$ and $0 < v_0 < v_1$, there exists B_0 and $\eta_0 \in (0,1)$ such that for all $B \geq B_0$, $0 < \eta < \eta_0$, $||\vec{X}_0|| \in (v_0 B, v_1 B)$, and $||\Delta\vec{X}|| \leq MB^{\eta}$, the following holds*

$$||\nabla V(\vec{X}_0 + \Delta\vec{X}) - \nabla V(\vec{X}_0)|| \leq \epsilon ||\nabla V(\vec{X}_0)||.$$

(c) *For any $\epsilon > 0$, $M > 0$ and $0 < v_0 < v_1$, there exists B_0 and $\eta_0 \in (0,1)$ such that for all $B \geq B_0$, $0 < \eta < \eta_0$, $||\vec{X}_0|| \in (v_0 B, v_1 B)$, and $||\vec{X}_1|| \leq MB^{\eta}$, the following holds,*

$$||\nabla V(\vec{X}_1)|| \leq \epsilon ||\nabla V(\vec{X}_0)||, \text{ for all } i.$$

(d) *The function $f(B)$ is convex and increasing, and the following condition holds*

$$\limsup_{B\to\infty} \frac{1}{f'(B)} \sup_{\{\vec{X}:V(\vec{X})=f(B)\}} ||\nabla V(\vec{X})|| < +\infty.$$

From these conditions, we can obtain the following lemma.

Lemma 3 *Suppose that the Lyapunov function $V(\vec{X})$ satisfies Assumptions 1 and 2. For any $\epsilon > 0$, $M > 0$, and $0 < v_0 < v_1$, there exists B_0 and $\eta_0 \in (0,1)$ such that for all $B \geq B_0$, $0 < \eta < \eta_0$, $||\vec{X}_0|| \in (v_0 B, v_1 B)$, and $||\Delta\vec{X}|| \in (0, MB^{\eta})$, the following holds,*

$$\begin{aligned} V(\vec{X}_0 + \Delta\vec{X}) - V(\vec{X}_0) &\leq \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0} \Delta\vec{X}_i \\ &\quad + \epsilon ||\nabla V(\vec{X}_0)|| \cdot ||\Delta\vec{X}||. \end{aligned}$$

Proof: Let $h(t) = V(\vec{X}_0 + t\Delta\vec{X})$. Then

$$h'(t) = \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0 + t\Delta\vec{X}} \Delta X_i.$$

Hence, by the mean-value theorem, we have

$$\begin{aligned} V(\vec{X}_0 + \Delta\vec{X}) - V(\vec{X}_0) &= h(1) - h(0) \\ &= \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0 + t\Delta\vec{X}} \Delta X_i, \end{aligned}$$

for some $t \in (0,1)$. By part (b) of Assumption 2, we can find B_0 and $\eta_0 \in (0,1)$ such that for all $B \geq B_0$, $0 < \eta < \eta_0$, $||\vec{X}_0|| \in (v_0 B, v_1 B)$, and $||\Delta\vec{X}|| \leq MB^\eta$, we have,

$$\left| \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0 + t\Delta\vec{X}} - \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0} \right| \leq \epsilon ||\nabla V(\vec{X}_0)||, \text{ for all } i.$$

Hence, we have,

$$\begin{aligned} & \left| \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0 + t\Delta\vec{X}} \Delta X_i - \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}_0} \Delta X_i \right| \\ \leq \;& \epsilon N ||\nabla V(\vec{X}_0)|| \cdot ||\Delta\vec{X}||. \end{aligned}$$

The result of the lemma then follows. *Q.E.D.*

We can then obtain the following proposition, which can be viewed as a stronger version of (4). Recall that both the arrivals and the departures are assumed to be bounded.

Proposition 4 *Suppose that the scheduling policy minimizes the drift of the Lyapunov function $V(\vec{X})$, and the Lyapunov function $V(\vec{X})$ satisfies Assumptions 1 and 2. For any $\epsilon > 0$ and $0 < v_0 < v_1$, there exists B_0 and $\eta_0 \in (0,1)$ such that for all $B \geq B_0$, $0 < \eta < \eta_0$, $||\vec{X}(t_0)|| \in (v_0 B, v_1 B)$, and $t \in (0, B^\eta)$, the following holds,*

$$\begin{aligned} & V(\vec{X}(t_0 + t)) - V(\vec{X}(t_0)) \\ \leq \;& t \left[\sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} (a_i + \epsilon) \right. \\ & \left. \quad - \sum_{j=1}^{\mathcal{S}} \phi_j \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} F_j^i \right] \end{aligned} \tag{5}$$

where

$$\begin{aligned} a_i &= \frac{1}{t}\sum_{k=0}^{t-1} A_i(t_0 + k) \text{ for all users } i \\ \phi_j &= \frac{1}{t}\sum_{k=0}^{t-1} \mathbf{1}_{\{C(t_0+k)=j\}} \text{ for all states } j. \end{aligned}$$

Proof: Fix $\epsilon > 0$ and $0 < v_0 < v_1$. Since both the arrivals and the service are bounded, there exists M such that $||\vec{X}(t+1) - \vec{X}(t)|| \leq M$ for all t. Hence, using Lemma 3 and with a suitable choice of M_1, there must exist B_0 and $\eta_0 \in (0,1)$ such that for all $B \geq B_0$, $0 < \eta < \eta_0$, $||\vec{X}(t_0)|| \in (v_0 B, v_1 B)$, and $t \in (0, B^\eta)$,

$$\begin{aligned} V(\vec{X}(t_0+t)) - V(\vec{X}(t_0)) \leq & \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} (X_i(t_0+t) \\ & -X_i(t_0)) + \epsilon M_1 t \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)}. \end{aligned} \tag{6}$$

The first term on the right-hand-side can be rewritten as

$$\begin{aligned} & \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} (X_i(t_0+t) - X_i(t_0)) \\ = & \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \sum_{k=0}^{t-1} A_i(t_0+k) \\ & - \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \sum_{k=0}^{t-1} D_i(t_0+k), \end{aligned} \tag{7}$$

where $D_i(t_0+k)$ is the actual amount of service that user i receives at time t_0+k. Note that at any time-slot $t = t_0 + k$,

$$D_i(t) \leq \sum_{j=1}^{\mathcal{S}} F_j^i \mathbf{1}_{\{C(t)=j, U(t)=i\}},$$

and equality holds whenever $X_i(t) \geq M$. Further, at each time-slot t, if the channel state is j, then the scheduling policy chooses the user that maximizes $\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} F_j^i$. Hence, we must have

$$\begin{aligned} & \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} D_i(t) \\ \geq & \sum_{j=1}^{\mathcal{S}} \mathbf{1}_{\{C(t)=j\}} \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} F_j^i - M_2, \end{aligned}$$

where M_2 is a constant that bounds $\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t)} M$ for all $X_i \leq M$ and for all i. (Such a constant M_2 can be found due to part (a) of Assumption 2.) Hence, we have,

$$\begin{aligned}
&\sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} D_i(t_0+k) \\
= &\sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0+k)} D_i(t_0+k) \\
&-\sum_{i=1}^{N} \left(\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0+k)} - \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \right) D_i(t_0+k) \\
\geq &\sum_{j=1}^{\mathcal{S}} \mathbf{1}_{\{C(t_0+k)=j\}} \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0+k)} F_j^i \\
&-\sum_{i=1}^{N} \left(\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0+k)} - \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \right) D_i(t_0+k) - M_2 \\
\geq &\sum_{j=1}^{\mathcal{S}} \mathbf{1}_{\{C(t_0+k)=j\}} \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} F_j^i \\
&-\sum_{i=1}^{N} \left| \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0+k)} - \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \right| \\
&\times(\max_{j=1,\ldots,\mathcal{S}} F_j^i + D_i(t_0+k)) - M_2 \\
\geq &\sum_{j=1}^{\mathcal{S}} \mathbf{1}_{\{C(t_0+k)=j\}} \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} F_j^i \\
&-\epsilon M_3 \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} - M_2
\end{aligned}$$

for some $M_3 > 0$, where in the last step we have used part (b) of Assumption 2. Substituting into (7), we have

$$\sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} (X_i(t_0+t) - X_i(t_0))$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \sum_{k=0}^{t-1} A_i(t_0+k) \\
&\quad - \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} \sum_{k=0}^{t-1} D_i(t_0+k) \\
&\leq t \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} a_i - t \sum_{j=1}^{\mathcal{S}} \phi_j \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} F_j^i \\
&\quad + t\epsilon M_3 \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)} + tM_2.
\end{aligned}
$$

Substituting into (6), and choose a sufficient large B so that M_2 is bounded by $\epsilon \sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(t_0)}$ for any $||\vec{X}(t_0)|| \geq v_0 B$, the result of the Proposition then follows. *Q.E.D.*

4.2 The Lower Bound

Next, define the following optimization problem. For all $\vec{\phi}$ and $\vec{a}$, let

$$
\begin{aligned}
l^B(\vec{\phi}, \vec{a}) = \max \quad & \sum_{i=1}^{N} \frac{\partial V}{\partial X_i} a_i - \sum_{j=1}^{\mathcal{S}} \phi_j \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i} F_j^i \\
\text{subject to} \quad & V(\vec{X}) = f(B).
\end{aligned}
$$

Comparing with (5), the value $l^B(\vec{\phi}, \vec{a})$ can be viewed as, given $V(\vec{X}(t_0)) = f(B)$, the fastest way with which $V(\vec{X}(t))$ can grow *locally* when the channel-state process and the arrival process satisfy $\frac{d\vec{s}}{dt} = \vec{\phi}$ and $\frac{d\vec{g}}{dt} = \vec{a}$ at time t_0. Assume that the following limit exists for all $\vec{\phi}$ and $\vec{a}$:

$$
w(\vec{\phi}, \vec{a}) = \lim_{B \to \infty} \frac{1}{f'(B)} l^B(\vec{\phi}, \vec{a}).
$$

Roughly speaking, $w(\vec{\phi}, \vec{a})$ is the fastest way for $f^{-1}(V(\vec{X}))$ to grow when $V(\vec{X}) = f(B)$. Hence $\frac{1}{w(\vec{\phi}, \vec{a})}$ is the earliest time that $V(\vec{X})$ can exceed $f(B)$. Let

$$
\theta_0 = \inf_{w(\vec{\phi}, \vec{a}) > 0} \frac{H(\vec{\phi}||\vec{p}) + L(\vec{a})}{w(\vec{\phi}, \vec{a})}.
$$

Let $\mathbf{P}_0[\cdot]$ denote the distribution conditioned on $\vec{X}(0) = 0$. Then we have the following result.

Proposition 5 *Suppose that the scheduling policy minimizes the drift of the Lyapunov function $V(\vec{X})$, and the Lyapunov function $V(\vec{X})$ satisfies Assumptions 1 and 2. Assume that $\vec{X}(0) = 0$. Then for all $T \geq 0$,*

$$\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] \leq -\theta_0.$$

To prove Proposition 5, we will need the following Lemma 6 and Lemma 7 on the properties of $w(\vec{\phi}, \vec{a})$.

Lemma 6 *If the limit in the definition of $w(\vec{\phi}, \vec{a})$ exists for all $\vec{\phi}$ and $\vec{a}$, then the function $w(\vec{\phi}, \vec{a})$ is continuous with respect to $\vec{\phi}$ and $\vec{a}$.*

Proof: Because the arrivals and the services are both bounded, we have for all B,

$$\begin{aligned} & ||l^B(\vec{\phi} + \Delta\vec{\phi}, \vec{a} + \Delta\vec{a}) - l^B(\vec{\phi}, \vec{a})|| \\ \leq \; & [\sum_{i=1}^{N} |\Delta a_i| + \sum_{j=1}^{\mathcal{S}} |\Delta \phi_j| \max_{i=1,\ldots,N} F_j^i] \\ & \times \sup_{\vec{X}: V(\vec{X}) = f(B)} ||\nabla V(\vec{X})||. \end{aligned}$$

Hence,

$$\begin{aligned} & ||w(\vec{\phi} + \Delta\vec{\phi}, \vec{a} + \Delta\vec{a}) - w(\vec{\phi}, \vec{a})|| \\ \leq \; & [\sum_{i=1}^{N} |\Delta a_i| + \sum_{j=1}^{\mathcal{S}} |\Delta \phi_j| \max_{i=1,\ldots,N} F_j^i] \\ & \times \limsup_{B\to\infty} \frac{1}{f'(B)} \sup_{\vec{X}: V(\vec{X}) = f(B)} ||\nabla V(\vec{X})||. \end{aligned}$$

The continuity of $w(\vec{\phi}, \vec{a})$ then follows from Assumption 2(d). *Q.E.D.*

Remark: If, in addition to the above result, $\vec{\phi}$ and $\vec{a}$ are constrained within a bounded set (which is true for our problem setting because the arrivals are bounded), then the function $w(\vec{\phi}, \vec{a})$ is uniformly continuous.

Lemma 7 *If the limit in the definition of $w(\vec{\phi}, \vec{a})$ exists for all $\vec{\phi}$ and $\vec{a}$ within a closed and bounded set, then the convergence of the limit is uniform for all $\vec{\phi}$ and $\vec{a}$. In other words, for any $\epsilon > 0$, there exists B_0 such that for all $B \geq B_0$ and for all $\vec{\phi}$ and $\vec{a}$, the following holds*

$$|\frac{1}{f'(B)} l^B(\vec{\phi}, \vec{a}) - w(\vec{\phi}, \vec{a})| \leq \epsilon.$$

Proof: Suppose in contrary that the convergence is not uniform. Then there must exist some $\epsilon > 0$, and a sequence of $\vec{\phi}^n$, $\vec{a}^n$ and B^n such that $B^n \to \infty$ and

$$|\frac{l^{B^n}(\vec{\phi}^n, \vec{a}^n)}{f'(B^n)} - w(\vec{\phi}^n, \vec{a}^n)| \geq \epsilon, \text{ for all } n = 1, 2, ...$$

Since $(\vec{\phi}^n, \vec{a}^n)$ are bounded, there exists a converging subsequence. Without loss of generality, denote this converging subsequence as $(\vec{\phi}^n, \vec{a}^n)$ and let $(\vec{\phi}, \vec{a})$ denote its limit. By Lemma 6, $w(\vec{\phi}^n, \vec{a}^n) \to w(\vec{\phi}, \vec{a})$. Hence, using similar bounding technique as in the proof of Lemma 6, we can find N_0 such that for all $n \geq N_0$,

$$\left| \frac{l^{B^n}(\vec{\phi}, \vec{a})}{f'(B^n)} - w(\vec{\phi}, \vec{a}) \right| \geq \frac{\epsilon}{2}.$$

This contradicts to the assumption that the limit in the definition of $w(\vec{\phi}, \vec{a})$ converges. Hence, the result of the lemma must hold. *Q.E.D.*

To prove Proposition 5, we will use the notion of a Generalized Fluid Sample Path (GFSP) introduced in [3]. Consider a sequence of scaled sample paths $(\vec{s}^B(\cdot), \vec{g}^B(\cdot), \vec{x}^B(\cdot))$ on the time-interval $[0, T]$. Define $\mu^B(t) = \frac{f^{-1}(V(B\vec{x}^B(t)))}{B}$. Fix $\eta \in (0, 1)$. For each B, divide the time interval $[0, T]$ into sub-intervals of length B^η/B, i.e., $[0, B^\eta/B]$, $[B^\eta/B, 2B^\eta/B]$, $[2B^\eta/B, 3B^\eta/B]$, and so on. For any scaled sample path $(\vec{s}^B(\cdot), \vec{g}^B(\cdot))$ (which is an element of the above sequence), linearize $\vec{s}^B$ and $\vec{g}^B$ on each such sub-interval. Let $U^B(\vec{s}^B, \vec{g}^B)$ denote such a linearized version of $\vec{s}^B$ and $\vec{g}^B$. For each t, let $\theta^B(t) = \frac{B^\eta}{B} \lfloor \frac{t}{B^\eta/B} \rfloor$. We can then define the *refined cost* of the scaled sample path $(\vec{s}^B(\cdot), \vec{g}^B(\cdot))$ on the time-interval $[0, t]$ as $\bar{J}^B(t) = J_{[0,\theta^{(B)}(t)]}(U^B(\vec{s}^B, \vec{g}^B))$.

Taking subsequence if necessary, assume that the sequence

$$(\vec{s}^B(\cdot), \vec{g}^B(\cdot), \vec{x}^B(\cdot), \mu^B(\cdot), \bar{J}^B(\cdot))$$

converges to $(\vec{s}(\cdot), \vec{g}(\cdot), \vec{x}(\cdot), \mu(\cdot), \bar{J}(\cdot))$ uniformly over the time interval $[0, T]$. This entire sequence (along with its limit) is called a Generalized Fluid Sample Path (GFSP).

The following theorem from [3] (in a slightly-varied form) establishes a lower bound on the asymptotic decay-rate of the queue overflow probability using GFSP.

Theorem 8 *Assume that $\vec{X}(0) = 0$. For any $\eta \in (0,1)$ the following holds: For any $T > 0$,*

$$\begin{aligned} & \limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] \\ \leq \; & -\inf\{\bar{J}(T)| \text{ for all GFSP's such that } \mu(0) = 0 \\ & \quad \text{ and } \mu(T) \geq 1\}. \end{aligned}$$

We are now ready to show Proposition 5.

Proof of Proposition 5 : Fix $T > 0$. From Theorem 8, we only need to show that there exists a function $\beta(\delta)$ such that $\beta(\delta) \to 0$ as $\delta \to 0$ and that for every $\delta > 0$, there exists $\eta \in (0,1)$ such that

$$\bar{J}(T) \geq \theta_0 - \beta(\delta) \tag{8}$$

for all GFSP (corresponding to η) with $\mu(0) = 0$ and $\mu(T) \geq 1$.

Fix $\delta > 0$. Let $v_0 = \delta$ and v_1 be a large number such that $||\vec{x}^B(t)|| \leq v_1$ for all $B \geq B_1$ (for some $B_1 > 0$) and $t \in (0, T)$. (Such a v_1 must exist because the arrivals are bounded.) Note that for any GFSP, the derivatives of $\vec{s}(\cdot)$ and $\vec{g}(\cdot)$ exist almost everywhere and are within a closed and bounded set. By Lemma 6, there exists $\epsilon > 0$ such that $w(\vec{\phi}, \vec{a} + \vec{\epsilon}) \leq w(\vec{\phi}, \vec{a}) + \delta$ for all $\vec{\phi}$ and $\vec{a}$ within this set, where $\vec{\epsilon}$ denote a vector whose components are all ϵ. Let M be the bound on the change of $\vec{X}$ in one time-slot. Then, according to Proposition 4, there exists $B_2 \geq B_1$ and $\eta_0 \in (0,1)$ such that the statement of Proposition 4 holds for ϵ, v_0, v_1 and M.

Take any $0 < \eta < \eta_0$. Take any GFSP corresponding to this η such that $\mu(0) = 0$ and $\mu(T) \geq 1$. Define the following for the limiting sample path $(\vec{s}(\cdot), \vec{g}(\cdot), \vec{x}(\cdot), \mu(\cdot))$. Let $T_1 = \inf\{t \geq 0 | \mu(t) \geq 1\}$ be the first time such that $\mu(t) \geq 1$. Let $T_0 = \sup\{t \leq T_1 \mid ||\vec{x}(t)|| \leq 3\delta \text{ or } \mu(t) \leq 3\delta\}$ be the last time before T_1 such that $||\vec{x}(t)|| \leq 3\delta$ or $\mu(t) \leq 3\delta$. Further, there exists B_3 such that for all $B \geq B_3$, the difference between $(\vec{s}^B(\cdot), \vec{g}^B(\cdot), \vec{x}^B(\cdot), \mu^B(\cdot), \bar{J}^B(\cdot))$

and $(\vec{s}(\cdot), \vec{g}(\cdot), \vec{x}(\cdot), \mu(\cdot), \bar{J}(\cdot))$ is less than δ. Hence, during the time interval (T_0, T_1), we must have, for all $B \geq B_3$ and $t \in [T_0, T_1]$,

$$||\vec{x}^B(t)|| \geq 2\delta, \quad 2\delta \leq \mu^B(t) \leq 1 + \delta,$$
$$\mu^B(T_0) \leq 4\delta \text{ and } \mu^B(T_1) \geq 1 - \delta.$$

Finally, according to Lemma 7, we can take another large $B_4 \geq B_3$ such that for all $B \geq B_4$ and for all $\vec{\phi}$ and $\vec{a}$,

$$|\frac{1}{f'(B\delta)} l^{B\delta}(\vec{\phi}, \vec{a}) - w(\vec{\phi}, \vec{a})| \leq \delta. \tag{9}$$

Fix some $B \geq \max\{B_2, B_4\}$ and divide the interval $[0, T]$ to sub-intervals of length B^η/B. Let $[k_0 B^\eta/B, (k_0+1)B^\eta/B]$ and $[k_1 B^\eta/B, (k_1+1)B^\eta/B]$ be the first and last sub-intervals, respectively, that are completely contained in $[T_0, T_1]$. By choosing a sufficiently large B, we can ensure that $k_0 < k_1$, and $\mu^B(k_0 B^\eta/B) \leq 5\delta$ and $\mu^B((k_1+1)B^\eta/B) \geq 1 - 2\delta$.

Consider any such sub-interval k between k_0 and k_1. Denote it by

$$[kB^\eta/B, (k+1)B^\eta/B].$$

Let $t_0 = kB^\eta/B$ and $t = B^\eta/B$. According to Proposition 4, the change of the Lyapunov function must satisfy:

$$\begin{aligned} & V(\vec{X}(B(t_0+t))) - V(\vec{X}(Bt_0)) \\ \leq \;& Bt \left[\sum_{i=1}^{N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(Bt_0)} (a_i + \epsilon) \right. \\ & \left. - \sum_{j=1}^{\mathcal{S}} \phi_j \max_{i=1,\ldots,N} \frac{\partial V}{\partial X_i}\Big|_{\vec{X}(Bt_0)} F_j^i \right] \end{aligned}$$

where

$$\begin{aligned} a_i &= \frac{1}{t}[g_i^B((t_0+t)) - g_i^B(t_0)] \text{ for all users } i \\ \phi_j &= \frac{1}{t}[s_j^B((t_0+t)) - s_j^B(t_0)] \text{ for all states } j. \end{aligned}$$

Consider the function $\mu^B(V) = \frac{1}{B} f^{-1}(V)$. Its derivative is given by $\frac{d\mu^B}{dV} = \frac{1}{Bf'(B\mu^B)}$. Further, we assume in part (d) of Assumption 2 that $f(\cdot)$ is convex and increasing. Hence, $f^{-1}(\cdot)$ is concave and increasing. Therefore, the drift of $\mu^B(t)$ must satisfy

$$\begin{aligned} & \mu^B(t_0+t)-\mu^B(t_0) \\ \le\; & \frac{t}{f'(B\mu^B(t_0))}\left[\sum_{i=1}^{N}\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(Bt_0)}(a_i+\epsilon)\right. \\ & \left.-\sum_{j=1}^{\mathcal{S}}\phi_j\max_{i=1,\ldots,N}\frac{\partial V}{\partial X_i}\Big|_{\vec{X}(Bt_0)}F_j^i\right]. \end{aligned}$$

Note that the quantity in the bracket on the right hand side is no greater than $l^{B\mu^B(t_0)}(\vec{\phi},\vec{a}+\vec{\epsilon})$. Further, by (9), the quantity on the right hand side is no greater than $t[w(\vec{\phi},\vec{a}+\vec{\epsilon})+\delta]$ since $\mu^B(t_0)\ge\delta$.

Using the above inequality and the definition of θ_0, the refined cost for such a sub-interval satisfies

$$\begin{aligned} t[H(\vec{\phi}||\vec{p})+L(\vec{a})] &\ge t\theta_0 w(\vec{\phi},\vec{a})\ge t\theta_0[w(\vec{\phi},\vec{a}+\vec{\epsilon})-\delta] \\ &\ge \theta_0[\mu^B(t_0+t)-\mu^B(t_0)-2t\delta]. \end{aligned}$$

Note that this is true for all sub-intervals k. Summing over all subintervals between k_0 and k_1, we have

$$\begin{aligned} \bar{J}^B(T) &\ge J_{[k_0B^\eta/B,(k_1+1)B^\eta/B]}(U^B[\vec{s}^B,\vec{g}^B]) \\ &\ge \theta_0[\mu^B((k_1+1)B^\eta/B)-\mu^B(k_0B^\eta/B)] \\ &\quad -2\theta_0 T\delta \\ &\ge \theta_0(1-7\delta)-2\theta_0T\delta. \end{aligned}$$

Hence,

$$\bar{J}(T)\ge\bar{J}^B(T)-\delta\ge\theta_0(1-7\delta)-2\theta_0T\delta-\delta.$$

Hence, we have shown (8). The result of the Proposition then follows. *Q.E.D.*

5 Large Deviations Optimality of Scheduling Algorithms that Minimize the Drift of a Lyapunov Function

If $\tilde{V}(\cdot) = V(\cdot)$ and $\tilde{f}(\cdot) = f(\cdot)$, the upper bound I_{opt} and the lower bound θ_0 differ only in their dependence on $\tilde{w}(\vec{\phi}, \vec{a})$ versus $w(\vec{\phi}, \vec{a})$.

Define $\Lambda(\vec{\phi})$ as the rate-region of the system (i.e., the set of all feasible offered-load vectors $\vec{\lambda}$) when the channel distribution is twisted to $\vec{\phi}$. Take $\mathbf{dist}(\vec{a}, \Lambda(\vec{\phi})) \triangleq \inf_{\vec{y} \in \Lambda(\vec{\phi})} ||\vec{a} - \vec{y}||$.

First, we note that there exists a $\hat{\delta} > 0$ such that the infimum in the definition of I_{opt} and θ_0 can be taken over the set of $\vec{\phi}$ and $\vec{a}$ such that $\mathbf{dist}(\vec{a}, \Lambda(\vec{\phi})) > \hat{\delta}$. This is because when $\vec{a}$ is very close to the set $\Lambda(\vec{\phi})$, the values $w(\vec{\phi}, \vec{a})$ and $\tilde{w}(\vec{\phi}, \vec{a})$ are close to 0 and hence it can be shown that they will not influence the infimum.

We then have the following main result.

Proposition 9 *Take $\tilde{V}(\cdot) = V(\cdot)$ and $\tilde{f}(\cdot) = f(\cdot)$. Under Assumptions 1 and 2, if $\tilde{w}(\vec{\phi}, \vec{a}) \geq w(\vec{\phi}, \vec{a})$ for all $\vec{\phi}$ and $\vec{a}$ such that $\mathbf{dist}(\vec{a}, \Lambda(\vec{\phi})) > \hat{\delta}$, then $I_{opt} = \theta_0$, and the scheduling algorithm that minimizes the drift of the Lyapunov function $V(\vec{X})$ must maximize the asymptotic decay-rate of the probability that the Lyapunov function value exceeds $f(B)$. More precisely, under such a scheduling algorithm π, there must exist $T_0 > 0$ such that for all $T \geq T_0$,*

$$\liminf_{B \to \infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)]$$
$$= \limsup_{B \to \infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] = -I_{opt}.$$

Proof: For any $\epsilon > 0$, there exists $\vec{\phi}_0$ and $\vec{a}_0$ such that

$$\frac{H(\vec{\phi}_0||\vec{p}) + L(\vec{a}_0)}{\tilde{w}(\vec{\phi}_0, \vec{a}_0)} \leq I_{opt} + \epsilon.$$

Let $T_0 > \frac{1}{\tilde{w}(\vec{\phi}_0, \vec{a}_0)}$. Using similar techniques as the proof of Proposition 1, we can show that for all $T \geq T_0$,

$$\liminf_{B \to \infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] \geq -(I_{opt} + \epsilon).$$

By Proposition 5, we must then have

$$\begin{aligned} I_{opt} + \epsilon &\geq -\liminf_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] \\ &\geq -\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] \\ &\geq \theta_0. \end{aligned}$$

If $\tilde{w}(\vec{\phi}, \vec{a}) \geq w(\vec{\phi}, \vec{a})$ for all $\vec{\phi}$ and $\vec{a}$ such that $\mathbf{dist}(\vec{a}, \Lambda(\vec{\phi})) > \hat{\delta}$, then by the definition of I_{opt} and θ_0, we must have $I_{opt} \leq \theta_0$. Since ϵ can be chosen to be arbitrarily small, We can then conclude that $I_{opt} = \theta_0$, and we can find T_0 such that for all $T \geq T_0$,

$$\lim_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] = -I_{opt}.$$

Q.E.D.

Remark: For each B, as $T \to \infty$, the probability $\mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)]$ approaches $\mathbf{P}_\pi[V(\vec{X}(0)) \geq f(B)]$. Hence, we could infer from Proposition 9 that the scheduling algorithm π should also maximize the asymptotic decay-rate of $\mathbf{P}_\pi[V(\vec{X}(0)) \geq f(B)]$. This argument could be made rigorous using the Freidlin-Wentzell construction [3, 6].

It could be non-trivial to check $\tilde{w}(\vec{\phi}, \vec{a}) \geq w(\vec{\phi}, \vec{a})$ for all $\vec{\phi}$ and $\vec{a}$. Next, we provide a sufficient condition that is easier to check. For each $B > 0$ and $\vec{y} = [y_i, i = 1, \ldots, N] \geq 0$, define

$$\begin{aligned} \bar{l}^B(\vec{y}) = \max \quad & \sum_{i=1}^{N} \frac{\partial V}{\partial X_i} y_i \\ \text{subject to} \quad & V(\vec{X}) = f(B). \end{aligned}$$

Note that $\bar{l}^B(\vec{y})$ can be viewed as the fastest possible way that, given $V(\vec{X}(t_0)) = f(B)$, $V(\vec{X}(t))$ can grow *locally* in a particular direction $\vec{y}$. Similarly, we can interpret $V(B\vec{y})$ as the value of $V(\vec{X}(t))$ when the growth direction of $\vec{X}(t)$ is *consistently* equal to $\vec{y}$.

Proposition 10 *Suppose that for any $\eta > 0$, $\epsilon > 0$ and $M > 0$, there exists a B_0 such that for all $B \geq B_0$ and all $\eta < ||\vec{y}|| \leq M$, the following holds*

$$\frac{1}{f'(B)} \bar{l}^B(\vec{y}) \leq \frac{1}{B} f^{-1}(V(B\vec{y})) + \epsilon.$$

Then under Assumptions 1 and 2, the scheduling algorithm that minimizes the drift of the Lyapunov function $V(\vec{X})$ must maximize the asymptotic decay-rate of the probability that the Lyapunov function value exceeds $f(B)$.

Proof: Take any $\vec{\phi}$ and $\vec{a}$ such that $\mathbf{dist}(\vec{a}, \Lambda(\vec{\phi})) > \hat{\delta}$. Fix a small $\epsilon > 0$. Let B_0 be chosen according to the assumption in the proposition with $M = N \max_{i=1,\dots,N} a_i$, and $\eta = \hat{\delta}$. For any $[u_j^i] \geq 0$ such that $\sum_{i=1}^{N} u_j^i = \phi_j$ for all j, we have

$$
\begin{aligned}
&\sum_{i=1}^{N} \frac{\partial V}{\partial X_i} a_i - \sum_{j=1}^{\mathcal{S}} \phi_j \max_{i=1,\dots,N} \frac{\partial V}{\partial X_i} F_j^i \\
&\leq \sum_{i=1}^{N} \frac{\partial V}{\partial X_i} a_i - \sum_{j=1}^{\mathcal{S}} \sum_{i=1}^{N} u_j^i \frac{\partial V}{\partial X_i} F_j^i \\
&= \sum_{i=1}^{N} \frac{\partial V}{\partial X_i} (a_i - \sum_{j=1}^{\mathcal{S}} u_j^i F_j^i) \leq \sum_{i=1}^{N} \frac{\partial V}{\partial X_i} [a_i - \sum_{j=1}^{\mathcal{S}} u_j^i F_j^i]^+.
\end{aligned}
$$

Now, let $y_i = [a_i - \sum_{j=1}^{\mathcal{S}} u_j^i F_j^i]^+$ for all i. y_i can be interpreted as the distance between $\vec{a}$ and a vector in the set $\Lambda(\vec{\phi})$. Since $\mathbf{dist}(\vec{a}, \Lambda(\vec{\phi})) > \hat{\delta}$, we must have $||\vec{y}|| \geq \hat{\delta} = \eta$. Then, by the assumption of the proposition, we have, for all $B \geq B_0$,

$$
\frac{l^B(\vec{\phi}, \vec{a})}{f'(B)} \leq \frac{\bar{l}^B(\vec{y})}{f'(B)} \leq \frac{f^{-1}(V(B\vec{y}))}{B} + \epsilon.
$$

Since this is true for all $[u_j^i]$, we must have

$$
\frac{l^B(\vec{\phi}, \vec{a})}{f'(B)} \leq \frac{f^{-1}(\tilde{l}^B(\vec{\phi}, \vec{a}))}{B} + \epsilon.
$$

Taking limit as $B \to \infty$, we obtain $w(\vec{\phi}, \vec{a}) \leq \tilde{w}(\vec{\phi}, \vec{a}) + \epsilon$. Since this is true for all $\epsilon > 0$, we must have $w(\vec{\phi}, \vec{a}) \leq \tilde{w}(\vec{\phi}, \vec{a})$. The result then follows from Proposition 9. *Q.E.D.*

5.1 The Optimality of the Log-rule

Next, we will use the above result to show that the log-rule is optimal in maximizing the asymptotic decay-rate of the probability that the sum-queue exceeds a threshold B. Take the Lyapunov function $V(\vec{X}) = \sum_{i=1}^{N}(X_i+1)\log(X_i+1) - X_i$, and that $f(B) = (B+1)\log(B+1) - B$. The policy that minimizes the drift of the Lyapunov function is the log-rule: at each time, the base-station should choose the user i with the largest value of

$$\frac{\partial V}{\partial X_i}F_j^i = \log(X_i+1)F_j^i.$$

We first obtain the following proposition.

Proposition 11 *The log-rule maximizes the asymptotic decay-rate of the probability that $V(\vec{X}) \geq f(B)$.*

Proof: We can verify that the Lyapunov function $V(\vec{X})$ satisfies Assumptions 1 and 2. For example, to verify part (b) and part (c) of Assumption 2, note that for all B and $v_0 > 0$ such that $MB^\eta < v_0 B$, and for all $||\vec{X}|| \geq v_0 B$, and $||\Delta X|| \leq MB^\eta$, we have,

$$\begin{aligned} &||\nabla V(\vec{X} + \Delta X) - \nabla V(\vec{X})|| \leq ||\nabla V(|\Delta X|)|| \\ &\leq \sum_{i=1}^{N} \log(|\Delta X_i| + 1) \leq N \log(MB^\eta + 1), \end{aligned}$$

and

$$||\nabla V(\vec{X})|| \geq \frac{1}{N}\sum_{i=1}^{N}\log(X_i+1) \geq \frac{1}{N}\log(\frac{v_0 B}{N} + 1).$$

Hence, for any $\epsilon > 0$, by choosing $\eta < \frac{\epsilon}{2N^2}$, we must have

$$||\nabla V(\vec{X} + \Delta X) - \nabla V(\vec{X})|| \leq ||\nabla V(|\Delta X|)|| \leq \epsilon ||\nabla V(\vec{X})||,$$

for all sufficiently large B.

It remains to check the condition in Proposition 10. We will first show that

$$\limsup_{B\to\infty} \frac{\bar{l}^B(\vec{y})}{f'(B)} \le \sum_{i=1}^N y_i \tag{10}$$

$$\liminf_{B\to\infty} \frac{f^{-1}(V(B\vec{y}))}{B} \ge \sum_{i=1}^N y_i. \tag{11}$$

To show (10), consider the following optimization problem

$$\begin{aligned} \max \quad & \sum_{i=1}^N \log(X_i+1) y_i \\ \text{subject to} \quad & \sum_{i=1}^N (X_i+1)\log(X_i+1) - X_i = f(B). \end{aligned}$$

Let $z_i = \log(X_i+1)$. It is easy to see that the optimal value of the above problem is the same as the optimal value of the following problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^N z_i y_i \\ \text{subject to} \quad & \sum_{i=1}^N e^{z_i}(z_i-1)+1 \le f(B). \end{aligned}$$

Clearly, if $y_i = 0$, then the corresponding z_i should be equal to 0. Let $\mathcal{I} = \{i | y_i > 0\}$. Introduce a Lagrange multiplier λ for the constraint, we can show that the optimal value z_i must satisfy

$$\begin{aligned} & y_i - \lambda e^{z_i} z_i = 0 \text{ for all } i \in \mathcal{I}. \\ & \sum_{i\in\mathcal{I}} e^{z_i}(z_i-1)+1 = f(B). \end{aligned}$$

This implies that for any two $i, i' \in \mathcal{I}$, we must have

$$\frac{e^{z_i} z_i}{e^{z_{i'}} z_{i'}} = \frac{y_i}{y_{i'}}.$$

Hence, as $B \to \infty$, the optimal values z_i must increase to infinity for all $i \in \mathcal{I}$. Take any small $\epsilon > 0$. We can then find a B_0 such that for $B \ge B_0$, the optimal values z_i must satisfy $z_i \geq 1$ and $z_i \leq (1+\epsilon)(z_i - 1)$ for all $i \in \mathcal{I}$. We then have,

$$\frac{1}{\lambda}\sum_{i\in\mathcal{I}} y_i = \sum_{i\in\mathcal{I}} e^{z_i} z_i \leq (1+\epsilon)\sum_{i\in\mathcal{I}} e^{z_i}(z_i - 1)$$
$$= (1+\epsilon)(f(B)-1).$$

Hence,

$$\lambda \geq \frac{\sum_{i\in\mathcal{I}} y_i}{(1+\epsilon)(f(B)-1)}.$$

Further, from $y_i = \lambda e^{z_i} z_i$ and $z_i \geq 1$, we have $y_i \geq \lambda e^{z_i}$ for $i \in \mathcal{I}$. Hence,

$$z_i \leq \log \frac{y_i}{\lambda} \leq \log \frac{(1+\epsilon) y_i (f(B)-1)}{\sum_{i\in\mathcal{I}} y_i},$$

and

$$\sum_{i\in\mathcal{I}} z_i y_i \leq \sum_{i\in\mathcal{I}} y_i \log \frac{(1+\epsilon) y_i}{\sum_{i\in\mathcal{I}} y_i} + \sum_{i\in\mathcal{I}} y_i \log(f(B)-1).$$

As $B \to \infty$, we then have

$$\begin{aligned}
&\limsup_{B\to\infty} \frac{1}{f'(B)} \sum_{i=1}^{N} z_i y_i \\
\leq\ &\limsup_{B\to\infty} \frac{\sum_{i\in\mathcal{I}} y_i \log[(B+1)\log(B+1) - B - 1]}{\log(B+1)} \\
=\ &\sum_{i\in\mathcal{I}} y_i = \sum_{i=1}^{N} y_i.
\end{aligned}$$

This then proves (10).

To show (11), note that for any $\vec{y} = [y_i] \geq 0$ and $\vec{y} \neq 0$,

$$\begin{aligned}
V(B\vec{y}) &= \sum_{i=1}^{N} [(By_i + 1)\log(By_i + 1) - By_i] \\
&\geq \sum_{i=1}^{N} y_i (B \log B - B) + B \sum_{i\in\mathcal{I}} y_i \log y_i,
\end{aligned}$$

and

$$
\begin{aligned}
& f(B\sum_{i=1}^N y_i) \\
= \ & (B\sum_{i=1}^N y_i + 1)\log(B\sum_{i=1}^N y_i + 1) - B\sum_{i=1}^N y_i \\
= \ & \sum_{i=1}^N y_i(B\log B - B) + B\sum_{i=1}^N y_i \log\frac{B\sum_{i=1}^N y_i + 1}{B} \\
& + \log(B\sum_{i=1}^N y_i + 1),
\end{aligned}
$$

where $\mathcal{I}$ again denotes the set of the indices i such that $y_i > 0$. For any $\epsilon > 0$, there exists B_1 such that for all $B \geq B_1$,

$$
B|\sum_{i\in\mathcal{I}} y_i \log y_i| \leq \epsilon \sum_{i=1}^N y_i(B\log B - B),
$$

and

$$
\begin{aligned}
& B\sum_{i=1}^N y_i \log\frac{B\sum_{i=1}^N y_i + 1}{B} + \log(B\sum_{i=1}^N y_i + 1) \\
& \leq \epsilon \sum_{i=1}^N y_i(B\log B - B).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
V(B\vec{y}) & \geq (1-\epsilon)\sum_{i=1}^N y_i(B\log B - B) \\
& \geq \frac{1-\epsilon}{1+\epsilon} f(B\sum_{i=1}^N y_i) \geq f(\frac{1-\epsilon}{1+\epsilon}B\sum_{i=1}^N y_i).
\end{aligned}
$$

We then have

$$
\liminf_{B\to\infty} \frac{1}{B} f^{-1}(V(B\vec{y})) \geq \frac{1-\epsilon}{1+\epsilon}\sum_{i=1}^N y_i.
$$

Since this is true for all $\epsilon > 0$, we then obtain (11). Further, using similar techniques as in the proof of Lemma 6 and Lemma 7, we can show that given any $0 < \eta < M$, the convergence in (10) and (11) is uniform over all $\vec{y}$ such that $\eta \leq ||\vec{y}|| \leq M$. The condition of Proposition 10 thus must hold.

By Proposition 10, we can then conclude that the log-rule maximizes the asymptotical decay-rate of the probability that $V(\vec{X}) \geq f(B)$. *Q.E.D.*

We now show the following result.

Proposition 12 *The log-rule maximizes the asymptotic decay-rate of the probability that* $\sum_{i=1}^{N} X_i \geq B$.

Proof: Using similar techniques as in the proof of the limit (11), we can show that $\lim_{B\to\infty} \frac{1}{B} f^{-1}(V(B\vec{y})) = \sum_{i=1}^{N} y_i$ and that, given any $0 < \eta < M$, the convergence is uniform over all $\vec{y}$ such that $\eta \leq ||\vec{y}|| \leq M$. Hence, we can show that the corresponding asymptotic decay-rate of the probability that $V(\vec{X}) \geq f(B)$ is given by

$$I_{opt} = \inf_{\tilde{w}(\vec{\phi},\vec{a})>0} \frac{H(\vec{\phi}||\vec{p}) + L(\vec{a})}{\tilde{w}(\vec{\phi},\vec{a})}, \tag{12}$$

where

$$\begin{aligned} \tilde{w}(\vec{\phi},\vec{a}) = \min \quad & \sum_{i=1}^{N} y_i \\ \text{subject to} \quad & y_i = [a_i - \sum_{j=1}^{\mathcal{S}} F_j^i u_j^i]^+ \\ & \sum_{i=1}^{N} u_j^i = \phi_j \text{ for all channel states } j. \end{aligned}$$

To show that the log-rule also maximizes the asymptotic decay-rate of the probability that $\sum_{i=1}^{N} X_i \geq B$, we consider the following minimization problem

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} (X_i + 1)\log(X_i + 1) - X_i \\ \text{subject to} \quad & \sum_{i=1}^{N} X_i \geq B. \end{aligned}$$

It is easy to check that its solution is given by $X_i = B/N$ where N is the total number of users, and the minimum value is given by

$$(B+N)\log(B+N) - (B+N)\log N - B.$$

Note that when $B \to \infty$, the ratio between this quantity and $f(B)$ converges to 1. Hence, for any $\epsilon > 0$, there exists a B_1 such that for all $B \geq B_1$, $\sum_{i=1}^{N} X_i \geq (1+\epsilon)B$ implies $V(\vec{X}) \geq f(B)$. We then have

$$\begin{aligned} &\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[\sum_{i=1}^{N} X_i(BT) \geq (1+\epsilon)B] \\ \leq\ &\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[V(\vec{X}(BT)) \geq f(B)] = -I_{opt}. \end{aligned}$$

By a change of variable, we can infer that $\limsup \frac{1}{B} \log \mathbf{P}_0[\sum_{i=1}^{N} X_i(BT) \geq B] \leq -\frac{I_{opt}}{1+\epsilon}$. Since this is true for all ϵ, we have $\limsup \frac{1}{B} \log \mathbf{P}_0[\sum_{i=1}^{N} X_i(BT) \geq B] \leq -I_{opt}$.

Finally, using $\tilde{V}(\vec{X}) = \sum X_i$ and $\tilde{f}(B) = B$, we can derive a lower bound on the overflow probability for $\sum_{i=1}^{N} X_i \geq B$ over all scheduling policies, i.e., there exists $T_0 > 0$ such that for all $T \geq T_0$,

$$\liminf_{B\to\infty} \frac{1}{B} \log \mathbf{P}_0[\sum_{i=1}^{N} X_i(BT) \geq B] \quad \geq \quad -I_{opt},$$

where I_{opt} is also given by (12). The result then follows. *Q.E.D.*

References

[1] S. Shakkottai, "Effective Capacity and QoS for Wireless Scheduling," *IEEE Transactions on Automatic Control*, vol. 53, no. 3, April 2008.

[2] L. Ying, R. Srikant, A. Eryilmaz, and G. E. Dullerud, "A Large Deviations Analysis of Scheduling in Wireless Networks," *IEEE Transactions on Information Theory*, vol. 52, no. 11, November 2006.

[3] A. L. Stolyar, "Large Deviations of Queues Sharing a Randomly Time-varying Server," *Queueing Systems*, vol. 59, pp. 1–35, 2008.

[4] V. J. Venkataramanan and X. Lin, "Structural Properties of LDP for Queue-Length Based Wireless Scheduling Algorithms," in *45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2007.

[5] ——, "On Wireless Scheduling Algorithms for Minimizing the Queue-Overflow Probability," *IEEE/ACM Transactions on Networking*, accepted for publication, also available at http://min.ecn.purdue.edu/∼linx/publications.html.

[6] ——, "On the Queue-Overflow Probabilty of Wireless Systems: A New Approach Combining Large Deviations with Lyapunov Functions," *IEEE Transactions on Information Theory*, submitted, also available at http://min.ecn.purdue.edu/∼lin/papers.html.

[7] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer-Verlag, 1998.

[8] B. Sadiq and G. de Veciana, "Optimality and large deviations of queues under the pseudo-Log rule opportunistic scheduling," in *46th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2008.

[9] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, December 1992.