On the Queue-Overflow Probabilities of A Class of Distributed Scheduling Algorithms ☆

Can Zhao[a], Xiaojun Lin[*,a]

[a]*School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, 47906, USA*

Abstract

In this paper, we are interested in using large-deviations theory to characterize the asymptotic decay-rate of the queue-overflow probability for distributed wireless scheduling algorithms, as the overflow threshold approaches infinity. We consider ad-hoc wireless networks where each link interferes with a given set of other links, and we focus on a distributed scheduling algorithm called Q-SCHED, which is introduced by Gupta et al. First, we derive a lower bound on the asymptotic decay rate of the queue-overflow probability for Q-SCHED. We then present an upper bound on the decay rate for all possible algorithms operating on the same network. Finally, using these bounds, we are able to conclude that, subject to a given constraint on the asymptotic decay rate of the queue-overflow probability, Q-SCHED can support a provable fraction of the offered loads achievable by any algorithms.

Key words: Wireless Networks, Quality of Service, Large Deviations, Scheduling Algorithms.

---

1. Introduction

Link scheduling is an important problem for ad-hoc wireless networks. In wireless networks the transmissions at neighboring links can interfere with each other. Hence, in order to maximize the capacity of the system, it is critical to schedule only a subset of non-interfering links at each time. There

☆An earlier version of this paper has been accepted for presentation at the 48th IEEE Conference on Decision and Control, Shanghai, China, December 2009 [1].

*Corresponding author. Tel: +1(765)494-0626. Email: linx@ecn.purdue.edu.

have been many studies on designing and analyzing scheduling algorithms for wireless network. A notable result is the well-known maximum-weight scheduling algorithm, which has been shown to be throughput-optimal, i.e., it can stabilize the network at the largest set of offered loads [2]. However, this algorithm is centralized and with high computational complexity. Therefore, many researchers have proposed low-complexity and distributed scheduling algorithms (see, e.g. [3–8]). Often, the goal is to be able to stabilize the network for a provable fraction of the capacity region. For example, the low-complexity algorithm in [5] has been shown to sustain close to 1/2 of the capacity region under the node-exclusive model.

To date most studies of wireless scheduling algorithms have mainly focused on stabilities. In other words, they ensure that the queues do not grow to infinity. Although stability is an important criterion, for many real-time applications stability is far from being sufficient. For example, when watching streaming video or listening to streaming audio, the user would expect that the delay of every packet can be upper bounded with high probability. As stability only ensures that the queue-length of each link remains finite, it cannot guarantee such type of stringent quality-of-service (QoS) requirements.

In certain cases, the probability of delay violation can be mapped to the probability of queue overflow. Unfortunately, both problems have been known to be very difficult. First, the exact probability distribution is usually mathematically intractable. Hence, one often has to turn to asymptotic techniques, such as large-deviations. For wireline networks, many results have been obtained using large-deviations techniques [9], based on the assumption that the packet arrival process is known and the service rate of each link is time-invariant. However, in wireless networks, the service rate process is time-varying. Some progress has been made for the case when the scheduling decision is based only on the channel state, which means that the service rate process has known statistics [10, 11]. However, for many wireless scheduling algorithms, even the statistics of the service rate process are unknown.

Recently, the queue-overflow probability for a number of queue-length based scheduling algorithms, for which the statistics of the service rate process are unknown, have been studied in [12–16] using sample-path large-deviations. In these works, the algorithms are centralized and are for a single cell. Further, the algorithms are deterministic in the sense that the scheduling decision is a deterministic function of the system state.

In this paper, we will develop techniques to estimate and control the QoS of distributed scheduling algorithms for ad hoc networks. We will focus on a random access algorithm for ad hoc wireless networks called Q-SCHED [5]. Note that due to the distributed and random nature of Q-SCHED, the techniques in prior works [12–16] do not apply directly. This is because the sample-path large deviations techniques in these prior works require the cost of each sample-path to be known. Note that for the scheduling algorithm in these prior works, since the scheduling decision is a deterministic function of the system state, the statistics of the service rate given the system state is known. Hence, the cost of each sample-path can be written down explicitly. However, for the Q-SCHED algorithm, due to the randomness and distributive nature of the algorithm, the statistics of the service rate given the current system state is not precisely known. (In fact, only bounds on its statistics are known, as can be seen from Lemma 1.) Hence, we can not specify the cost of a sample path explicitly, and thus, we can not use the methodologies from [12–16] directly.

As in [12–16] the questions that we are interested in are: a) how to estimate the decay rate of the queue overflow probability of this algorithm, and b) given an overflow constraint, how to calculate the set of offer-load vectors that this algorithm can support. To answer these questions, we will first obtain a lower bound on the decay rate of the overflow probability for Q-SCHED. Then, based on this bound, we provide a lower bound on the set of offer-load vectors that this algorithm could support at a given queue-overflow constraint. To the best of our knowledge, this is the first work that characterizes the queue-overflow probability of distributed scheduling algorithms for ad-hoc networks in a large-deviations setting. Finally, we show that subject to a given queue-overflow constraint, the offer load supported by Q-SCHED is at least a provable fraction of the offered load supported by any other algorithms.

2. System Model

We use the model from [5]. We consider a wireless network of N nodes. Let V be the set of nodes, E be the set of directed links between nodes, and $G(V, E)$ be the directed connectivity graph of the network. Each link $l \in E$ interferes with a set of other links in E, which we denote as $\mathcal{E}_l$. We assume that if $k \in \mathcal{E}_l$ then $l \in \mathcal{E}_k$, i.e., the interference relationship is symmetric. We also let $l \in \mathcal{E}_l$, i.e., $\mathcal{E}_l = \{l\} \cup \{l' \in E : l' \text{ interferes with } l\}$. This interference set varies when different communication techniques are used. For example, for bluetooth, we use the node-exclusive interference model, also known as the primary interference model or the one-hop interference model, where $\mathcal{E}_l$ is the set of all links that are connected to either end-point of l. In IEEE 802.11 WLAN, the interference set $\mathcal{E}_l$ will be the two-hop neighbors of l, including l.

We assume a slotted system. Let $a_l(n)$ denotes the number of packets that arrive at link l in time-slot n. We assume that for each l, $a_l(1), a_l(2), ...$ are *i.i.d.* and $\lambda_l = \mathbf{E}[a_l(1)]$. Moreover, we assume that $a_l(n)$ is upper bounded by A_M for all $n > 0$ and all $l \in E$, i.e., $0 \leq a_l(n) < A_M$, which means that the number of arrival packets is finite in each time slot.

Let $d_l(n)$ denote the number of packets that can be served by link l in time-slot n. Assume that the capacity of each link is a fixed number c_l. Let $s_l(n) = 1$ indicates that link l is scheduled in time-slot n, $s_l(n) = 0$ otherwise. Clearly, $d_l(n) = c_l s_l(n)$. We assume a single-hop system, i.e., packets served at link l immediately leave the system. Let $q_l(n)$ denote the backlog of link l in slot n, and $\vec{q}(n) = \big(q_1(n), q_2(n), ..., q_{|E|}(n)\big)$. Then the evolution of each $q_l(n)$ is given by $q_l(n+1) = [q_l(n) + a_l(n) - d_l(n)]^+$, where $[\cdot]^+$ denote the projection to $[0, \infty)$. We also define:

$$A_i(n) \triangleq \sum_{l \in \mathcal{E}_i} \frac{a_l(n)}{c_l}, \qquad D_i(n) \triangleq \sum_{l \in \mathcal{E}_i} \frac{d_l(n)}{c_l}. \tag{1}$$

These two variables are the sum of normalized arrival and service in each interference set at time slot n, and will be used frequently in Section 3.

We consider the algorithm Q-SCHED that was introduced in [5]. In this algorithm, it is assumed that at the beginning of each time-slot every link l knows the queue-lengths of all links in its interference set $\mathcal{E}_l$ and also the queue-lengths of all links in the interference set $\mathcal{E}_k$ for every $k \in \mathcal{E}_l$. Each time slot is divided into two parts: a scheduling slot and a data transmission slot. Links that are chosen in the scheduling slot will transmit their packets in the data transmission slot. The scheduling slot is further divided into M mini-slots. At the beginning of each time-slot n, each link l first computes:

$$P_l(n) = \alpha \frac{\frac{q_l(n)}{c_l}}{\max_{i \in \mathcal{E}_l} \sum_{k \in \mathcal{E}_i} \frac{q_k(n)}{c_k}},$$

where $\alpha = \log(M)$. Then, each link l picks a backoff time $Y_l(n)$ from

$\{1, 2, ..., M+1\}$ according to the following probabilities:

$$\mathbf{P}(Y_l(n) = M+1) = e^{-P_l(n)};$$
$$\mathbf{P}(Y_l(n) = m) = e^{-P_l(n)\frac{m-1}{M}} - e^{-P_l(n)\frac{m}{M}}, m = 1, 2, ..., M.$$

A link that chooses backoff time $Y_l(n) = k \leq M$ will start transmission at the k-th mini-slot unless it has already heard a transmission from one of its interfering links. If a link chooses a backoff time equals to $M+1$ it will not attempt to transmit in this time slot. If two or more links that interfere with each other begin to transmit simultaneously, collision will occur and all of these transmissions will fail. Finally, any link that hears the collision will not attempt to transmit in this time slot.

We now present an important lemma proved in [5] for Q-SCHED, which will be used in our derivation. Define $V(n) = \max_{i \in E} \sum_{l \in \mathcal{E}_i} \frac{q_l(n)}{c_l}$, which denotes the largest sum of backlog in any interference neighborhood.

Lemma 1. *Q-SCHED scheduling policy guarantees that for any $\epsilon_0 \geq 0$ and constants $C_1, C_2 \geq 0$, there exists a constant R such that if $V(n) \geq R$, then for any $\eta \in [0, 1]$ and for any link i such that*

$$\sum_{l \in \mathcal{E}_i} \frac{q_l(n)}{c_l} \geq \eta \left(V(n) - C_1 - C_2\epsilon_0\right),$$

the following holds,

$$\sum_{l \in \mathcal{E}_i} Pr\{\textit{Link } l \textit{ is scheduled}\} \geq \eta \left(1 - \frac{\log(M) + 1}{M} - \epsilon_0\right).$$

Note that although the original statement of Lemma 1 in [5] requires that $\epsilon_0, C_1, C_2 > 0$, the proof there also trivially holds for the case when $\epsilon_0, C_1, C_2 \geq 0$. Letting $\eta = 1$, this then implies that, when $V(n)$ is large, with high probability at least one link will be scheduled in those interference neighborhood with sum of backlog close to $V(n)$. In [5], this lemma has been used to establish the negative drift of the Lyapunov function $V(n)$ whenever the offered load satisfies that, for some $\epsilon_0 > 0$,

$$\sum_{l \in \mathcal{E}_i} \frac{\lambda_l}{c_l} \leq 1 - \frac{\log(M) + 1}{M} - \epsilon_0, \text{ for all links } i. \tag{2}$$

For the rest of the paper, we assume that (2) holds because otherwise we do not know the stability of the system.

In this paper, we are interested in queue-overflow probabilities. For example, we may want to know the probability that the maximum queue length exceeds a given threshold B. On the other hand, with the techniques developed in this paper, it is more convenient to work with

$$\mathbf{P}\left(\max_{i\in E}\sum_{l\in\mathcal{E}_i}\frac{q_l}{c_l}\geq B\right). \tag{3}$$

Note that when the arrival rates are constant, i.e. $a_i(t)=\lambda_i$ for all i,t, the above probability can be mapped to an upper bound of the probability of delay-violation. To see this, let $\tau_i(t)$ denotes the delay of the latest packet arrived at link i in time slot t. More specifically, $\tau_i(t)$ is the number of time slots for that packet remained in the system. Then for any $b>0$,

$$\begin{aligned}
&\tau_i(0)\geq b\\
\Leftrightarrow&\text{Cumulative service of link } i \text{ over}[0,b-1]<q_i(0)+\lambda_i\\
\Leftrightarrow&q_i(b)\geq b\lambda_i-(q_i(0)+\lambda_i)\\
\Leftrightarrow&q_i(b)/c_i-\lambda_i(b-1)/c_i\geq 0
\end{aligned}$$

If we consider the maximum delay in the neighborhood of link i, we will get

$$\begin{aligned}
&\max_{l\in\mathcal{E}_i}\tau_l(0)\geq b\\
\Leftrightarrow&\text{For some } l\in\mathcal{E}_l, q_l(b)/c_l-\lambda_l(b-1)/c_l\geq 0\\
\Leftrightarrow&\max_{l\in\mathcal{E}_i}\left(q_l(b)/c_l-\lambda_l(b-1)/c_l\right)\geq 0\\
\Rightarrow&\max_{l\in\mathcal{E}_i}\left(\sum_{l\in\mathcal{E}_i}\frac{q_l(b)}{c_l}-\frac{\lambda_l}{c_l}(b-1)\right)\geq 0\\
\Leftrightarrow&\sum_{l\in\mathcal{E}_i}\frac{q_l(b)}{c_l}\geq\min_{l\in\mathcal{E}_i}\frac{\lambda_l}{c_l}(b-1)
\end{aligned}$$

Hence, if the system is stationary one can conclude that [1]

$$\mathbf{P}\left(\max_{l\in\mathcal{E}_i}\tau_l(0)\geq b\right)\leq\mathbf{P}\left(\sum_{l\in\mathcal{E}_i}\frac{q_l(b)}{c_l}\geq\min_{l\in\mathcal{E}_i}\frac{\lambda_l}{c_l}(b-1)\right)$$
$$=\mathbf{P}\left(\sum_{l\in\mathcal{E}_i}\frac{q_l(0)}{c_l}\geq\min_{l\in\mathcal{E}_i}\frac{\lambda_l}{c_l}(b-1)\right)\qquad(4)$$

However, even calculating (3) is mathematically intractable. Hence, we will use large-deviations theory to estimate it. We are interested in the following limits:

$$I_0(\vec{\lambda})\triangleq-\limsup_{B\to\infty}\frac{1}{B}\log\mathbf{P}\left(\max_{i\in E}\sum_{l\in\mathcal{E}_i}\frac{q_l}{c_l}\geq B\right),$$
$$J_0(\vec{\lambda})\triangleq-\liminf_{B\to\infty}\frac{1}{B}\log\mathbf{P}\left(\max_{i\in E}\sum_{l\in\mathcal{E}_i}\frac{q_l}{c_l}\geq B\right).$$

Clearly, $I_0(\vec{\lambda})$ provides a lower bound on the decay rate of (3) and $J_0(\vec{\lambda})$ provides an upper bound.

3. The lower bound

We first develop a lower bound for $I_0(\vec{\lambda})$. For any link i in E, define the scaled queue length: $q_i^B(t)=\frac{1}{B}q_i\left(\lfloor Bt\rfloor\right)$. Note that this expression represents the standard large-deviations scaling that shrinks both time and magnitude. We also define the scaled version of the Lyapunov function: $v^B(t)=V(\vec{q}^B(t))$. The queue overflow criterion is $\left\{v^B(t)\geq 1\right\}$. For ease of exposition, we consider a system that starts at $t=0$. For a given $T>0$, we are interested in the following probability:

$$I_0^T(\vec{\lambda})\triangleq-\limsup_{B\to\infty}\frac{1}{B}\log\mathbf{P}\left(v^B(T)\geq 1\,\middle|\,v^B(0)=0\right).$$

[1] We note however that inequality (4) only provides an upper bound on the delay violation-probability. An algorithm that minimize the right hand side of (4) does not necessarily minimize the delay-violation probability. We will comment on the implication of this point in Section 7.

Intuitively, as $T \to \infty$, one would expect that $I_0^T(\vec{\lambda})$ approaches $I_0(\vec{\lambda})$, the lower bound on the decay rate of the stationary overflow probability [14, 15]. We now introduce the following main result, which provides a lower bound on $I_0^T(\vec{\lambda})$.

Theorem 2. *Assume that for some $\epsilon_0 > 0$, inequality (2) holds. For any small and positive ξ such that $0 < \xi < \epsilon_0$, let*

$$\epsilon = \frac{\log(M) + 1}{M} + \xi. \tag{5}$$

Moreover, let $\hat{D}(n)$, $n = 1, 2, ...$ be i.i.d random variables with distribution

$$\hat{D}(n) = \begin{cases} 1, & \text{with prob. } 1 - \epsilon \\ 0, & \text{with prob. } \epsilon \end{cases}$$

For any $T > 0$, the lower bound on the decay rate function satisfies

$$I_0^T(\vec{\lambda}) \geq \inf_{W \geq 0} \min_{i \in E} \inf_{d \leq (1-\epsilon)} \frac{(I_i^A(d + W) + I^{\hat{D}}(d))}{W} \triangleq L. \tag{6}$$

where

$$I^{\hat{D}}(d) = \sup_{\theta \in \mathbb{R}} \left\{ \theta d - \log \mathbf{E}\left(e^{\theta \hat{D}(1)}\right) \right\} = \sup_{\theta \in \mathbb{R}} \left\{ \theta d - \log\left(\epsilon + (1-\epsilon)e^{\theta_2}\right) \right\}.$$
$$I_i^A(a) = \sup_{\theta \in \mathbb{R}} \left\{ \theta a - \log \mathbf{E}\left(e^{\theta A_i(1)}\right) \right\}. \tag{7}$$

Note that the bound on the right hand side of (6) in Theorem 2 is independent from T. Assuming that $I_0^T(\vec{\lambda}) \to I_0(\vec{\lambda})$ we would then expect that $I_0(\vec{\lambda}) \geq L$. Such a limiting argument can be made rigorously using the Freidlin-Wentzell construction as in [14, 15].

Before we elaborate on the proof of Theorem 2, we would like to provide some intuitions behind the result. Recall from Lemma 1 that for all interference sets whose backlogs are almost the largest, i.e.

$$\sum_{l \in \mathcal{E}_i} \frac{q_l(n)}{c_l} \geq V(n) - C_1 - C_2\epsilon_0,$$

the following holds

$$\sum_{l \in \mathcal{E}_i} \Pr\{\text{Link } l \text{ is scheduled}\} \geq 1 - \frac{\log(M) + 1}{M} - \epsilon_0. \tag{8}$$

Now for simplicity, let us consider a fictitious algorithm, which guarantees that inequality (8) holds for every interference set at all time, regardless of their backlogs. Under this fictitious algorithm, the backlog of each interference set can be stochastically bounded from below by a single server queue with the same arrivals, and with the service rate given by $\hat{D}(n)$. Hence, the quantity L in (6) provides a lower bound on the decay rate of the overflow probability under this fictitious algorithm. However, the above argument does not directly apply to Q-SCHED. In Q-SCHED, inequality (8) only holds for those interference sets with the largest backlog, and such sets change from time to time, which makes it difficult for us to track the system dynamics directly by Lemma 1. To address the problem, in the following derivation we divide the entire scaled time into many small intervals. In each small interval, the interference sets that have almost the largest backlog do not change and therefore we are able to use Lemma 1 to estimate $I_0^T(\vec{\lambda})$.

3.1. Local Rate Function

For a fixed t, let $\delta > 0$ be a small number. Let $\Delta v^B(\delta, t) = v^B(t+\delta) - v^B(t)$ denote the drift of the scaled Lyapunov function. Let $\mathcal{Q}$ be a closed and bounded set such that $V(\vec{q}) \geq v > 0$ for all $\vec{q} \in \mathcal{Q}$. Our first goal is to find the following limit given $\vec{q} \neq \vec{0}$ and $W > 0$,

$$-\lim_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}\left(\Delta v^B(\delta, t) \geq \delta W \left| \vec{q}^B(t) = \vec{q}\right.\right). \tag{9}$$

We call (9) the local rate function, which is the maximum asymptotic decay rate of the probability that the growth rate of v^B is no smaller than W over all $\vec{q} \in \mathcal{Q}$, conditioned on $\vec{q}^B(t) = \vec{q}$. Since the arrival and departure are both bounded, for any $i \in E$ there must exist C_i such that

$$\left| \sum_{l\in\mathcal{E}_i} \frac{q_l(n+1)}{c_l} - \sum_{l\in\mathcal{E}_i} \frac{q_l(n)}{c_l} \right| \leq C_i$$

for all n. We next define the set $\mathcal{I}(\vec{q}, \delta)$ as

$$\mathcal{I}(\vec{q}, \delta) = \left\{ i \in E \left| \sum_{l\in\mathcal{E}_i} \frac{q_l}{c_l} \geq V(\vec{q}) - \delta C_i \right.\right\}. \tag{10}$$

Intuitively, $\mathcal{I}(\vec{q}^B(t), \delta)$ is the set of links that have the close-to-largest sum of backlog $\sum_{l\in\mathcal{E}_i} \frac{q_l^B(t)}{c_l}$ in their respective interference range. Given that $W > 0$, if the event $\{\Delta v^B(\delta,t) \geq \delta W\}$ happens, we will have $v^B(t) < v^B(\delta+t)$. For any B, if $i \in E$ and $i \notin \mathcal{I}(\vec{q}^B(t),\delta)$, then $\sum_{l\in\mathcal{E}_i} \frac{q_l^B(t)}{c_l} < v^B(t) - \delta C_i$, and hence

$$\sum_{l\in\mathcal{E}_i} \frac{q_l^B(t+\delta)}{c_l} < v^B(t) - \delta C_i + \delta C_i = v^B(t) < v^B(\delta+t).$$

Therefore, for any large B, only $i \in \mathcal{I}(\vec{q}^B(t),\delta)$ could potentially maximize $\sum_{l\in\mathcal{E}_i} \frac{q_l^B}{c_l}$ at time $t+\delta$. So the change between $v^B(t)$ and $v^B(t+\delta)$ can be bounded by the maximum increment of $\sum_{l\in\mathcal{E}_i} \frac{q_l^B}{c_l}$ among those $i \in \mathcal{I}f(\vec{q}^B(t),\delta)$. More specifically, we have

$$\Delta v^B(\delta,t) \leq \max_{i\in\mathcal{I}(\vec{q}^B(t),\delta)} \sum_{n=\lfloor Bt\rfloor+1}^{\lfloor B(t+\delta)\rfloor} \frac{1}{B} \sum_{l\in\mathcal{E}_i} \frac{a_l(n)-d_l(n)}{c_l}.$$

Let $\bar{A}_i$ be the mean of $A_i(n)$. Note that by our assumption, there exists $\epsilon_0 > 0$ such that $\bar{A}_i < 1 - \frac{\log(M)+1}{M} - \epsilon_0$, for all links i.

Now consider Equation (9), since $q(t)$ is Markovian, so is $q^B(t)$. We thus have

$$\begin{aligned}&\lim_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}\left(\Delta v^B(\delta,t) \geq \delta W \,\middle|\, \vec{q}^B(t) = \vec{q}\right)\\ &= \lim_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}\left(\Delta v^B(\delta,0) \geq \delta W \,\middle|\, \vec{q}^B(0) = \vec{q}\right).\end{aligned}$$

Hence, for the following derivation, we will take $t=0$, and drop the variable t when there is no source of confusion. Moreover, for ease of exposition, let $\mathbf{P}_{\vec{q}}(\cdot)$ denote the probability distribution conditioned on $\vec{q}(0) = \vec{q}$.

We first show the following property on the service process.

Lemma 3. *Assume that $\vec{q}^B(0) = \vec{q}$ and $V(\vec{q}) \geq v > 0$. For any $i \in \mathcal{I}(\vec{q},\delta)$, $D_i(n)$ is defined in* (1). *For any $v, \xi > 0$, there exists δ_0 and B_0 such that for all $\delta \leq \delta_0$ and for all $B \geq B_0$, the following holds for $1 \leq n \leq \lfloor B\delta \rfloor$,*

$$\begin{aligned}&\mathbf{P}_{\vec{q}}(D_i(n) \geq 1 \,|\, D_i(n-1), A_i(n-1), ..., D(1), A(1)) \geq 1-\epsilon,\\ &\mathbf{P}_{\vec{q}}(D_i(n) = 0 \,|\, D_i(n-1), A_i(n-1), ..., D(1), A(1)) \leq \epsilon,\end{aligned}$$

where ϵ is defined as (5)

Proof. For any $\eta \in [0, 1)$, choose δ such that

$$\delta \leq \frac{(1-\eta)v}{4 \max_{i \in E} C_i} \triangleq \delta_0.$$

If $v^B(0) = V(\vec{q}) \geq v > 0$, because the changing rate of $v^B(t)$ is at most $\max_{i \in E} C_i$, it follows that

$$v^B(0) + \frac{(1-\eta)v}{4} \geq v^B(n/B) \geq \frac{3v}{4}, \text{for all } n \in [0, \lfloor B\delta \rfloor].$$

Hence, for $i \in \mathcal{I}(\vec{q}, \delta)$, we have

$$\begin{aligned}
\sum_{l \in \mathcal{E}_i} \frac{q_l^B(n/B)}{c_l} &\geq \sum_{l \in \mathcal{E}_i} \frac{q_l^B(0)}{c_l} - \delta C_i \geq v^B(0) - 2\delta C_i \\
&\geq v^B(n/B) - \frac{3}{4}(1-\eta)v \\
&\geq v^B(n/B) - (1-\eta)v^B(n/B) \\
&\geq \eta v^B(n/B).
\end{aligned}$$

Therefore, the unscaled backlog must satisfy

$$\sum_{l \in \mathcal{E}_i} \frac{q_l(n)}{c_l} \geq \eta V(n), \text{for all } n \in [0, \lfloor B\delta \rfloor].$$

Take the constants $\epsilon_0 = 0, C_1 = 0, C_2 = 0$ in Lemma 1. Also, since $v^B(n/B) \geq 3v/4$ for $n \in [0, \lfloor B\delta \rfloor]$, then for large enough B, $V(n) \geq R$ for $n \in [0, \lfloor B\delta \rfloor]$, where R is given by Lemma 1. Therefore, for any $i \in \mathcal{I}(\vec{q}, \delta)$, the following holds according to Lemma 1:

$$\sum_{l \in \mathcal{E}_i} \mathbf{P}_{\vec{q}}(\text{Link } l \text{ is scheduled}) \geq \eta \left(1 - \frac{\log(M)+1}{M}\right).$$

For any $\xi > 0$, we could choose η close enough to 1 so that the result of Lemma 3 holds. □

Lemma 3 implies that, when δ is small, the service rate at each neighborhood of link $i \in \mathcal{I}(\vec{q}, \delta)$ is no smaller than 1 with probability no smaller than $1 - \epsilon$.

Let $A_i^B(\delta) \triangleq \frac{1}{B}\sum_{n=1}^{[B\delta]} A_i(n)$, $D_i^B(\delta) \triangleq \frac{1}{B}\sum_{n=1}^{[B\delta]} D_i(n)$, $\hat{D}^B(\delta) \triangleq \frac{1}{B}\sum_{n=1}^{[B\delta]} \hat{D}(n)$, $Z_i(n) \triangleq A_i(n) - D_i(n)$, $Z_i^B(\delta) \triangleq A_i^B(\delta) - D_i^B(\delta)$, $\hat{Z}(n) \triangleq A_i(n) - \hat{D}_i(n)$, and $\hat{Z}_i^B(\delta) \triangleq A_i^B(\delta) - \hat{D}_i^B(\delta)$. We then have

$$\Delta v^B(\delta) \leq \max_{i \in \mathcal{I}(\vec{q}^B(0),\delta)} Z_i^B(\delta), \tag{11}$$

For each $i \in E$, let

$$H_i^{\hat{Z}} \triangleq \limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}\left(\hat{Z}_i^B(\delta) \geq \delta W\right),$$
$$H_{\max}^Z \triangleq \limsup_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}_{\vec{q}}\left(\max_{i\in\mathcal{I}(\vec{q},\delta)} Z_i^B(\delta) \geq \delta W\right).$$

From (11), we have

$$\limsup_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}\left(\Delta v^B(\delta) \geq \delta W \,\middle|\, \vec{q}^B(0) = \vec{q}\right) \leq H_{\max}^Z.$$

Unfortunately, $H_{\max}^Z$ is difficult to compute directly. On the other hand, $H_i^{\hat{Z}}$'s are fairly easy to compute. We will establish a relationship between $H_{\max}^Z$ and $H_i^{\hat{Z}}$'s and estimate $H_{\max}^Z$ by $H_i^{\hat{Z}}$'s.

Lemma 4. *Assume that $\vec{q}^B(0) = \vec{q}$ and $V(\vec{q}) \geq v > 0$. For any $\xi > 0$, define ϵ as (5). For any $v > 0$, there exists $\delta_0 > 0$ and $B_0 > 0$, such that for any $W > 0$, any $i \in \mathcal{I}(\vec{q},\delta)$ and for all $0 < \delta \leq \delta_0$ and $B \geq B_0$, we will have*

$$\mathbf{P}_{\vec{q}}(Z_i^B(\delta) \geq W) \leq \mathbf{P}(\hat{Z}_i^B(\delta) \geq W).$$

The proof of Lemma 4 uses the property in Lemma 3 that $\hat{D}(n)$ stochastically dominates $D(n)$. The details of the proof is provided in the appendix.

Using Lemma 4, we can then establish the following relationship between $H_{\max}^Z$ and $H_i^{\hat{Z}}$.

Lemma 5. *Let $\mathcal{Q}$ be a closed and bounded set such that $V(\vec{q}) \geq v > 0$ for all $\vec{q} \in \mathcal{Q}$. For any $\xi > 0$, define ϵ as (5). There exists $\delta_0 > 0$, such that for any $W > 0$ and for all $0 < \delta \leq \delta_0$, we have*

$$\begin{aligned} H_{max}^Z &= \limsup_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}_{\vec{q}}\left(\max_{i\in\mathcal{I}(\vec{q},\delta)} Z_i^B(\delta) \geq \delta W\right) \\ &\leq \max_{i\in E} \limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}\left(\hat{Z}_i^B(\delta) \geq \delta W\right) = \max_{i\in E} H_i^{\hat{Z}}. \end{aligned} \tag{12}$$

Proof. Choose δ_0 and B_0 as in Lemma 4. According to Lemma 4, for $0 < \delta \le \delta_0$, when $B \ge B$, we have

$$\begin{aligned}
&\mathbf{P}_{\vec{q}}\left(\max_{i\in\mathcal{I}(\vec{q},\delta)} Z_i^B(\delta) \ge \delta W\right)\\
\le&\sum_{i\in\mathcal{I}(\vec{q},\delta)} \mathbf{P}_{\vec{q}}\left(Z_i^B(\delta) \ge \delta W\right)\\
\le&\sum_{i\in\mathcal{I}(\vec{q},\delta)} \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right)\\
\le&\sum_{i\in E} \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right)\\
\le&|E|\max_{i\in E} \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right).
\end{aligned}$$

Hence,

$$\sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}_{\vec{q}}\left(\max_{i\in\mathcal{I}(\vec{q},\delta)} Z_i^B(\delta) \ge \delta W\right) \le |E|\max_{i\in E} \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right).$$

It follows that

$$\begin{aligned}
&\limsup_{B\to\infty} \frac{1}{B}\log\sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}_{\vec{q}}\left(\max_{i\in\mathcal{I}(\vec{q},\delta)} Z_i^B(\delta) \ge \delta W\right)\\
\le&\limsup_{B\to\infty} \frac{1}{B}\log|E|\max_{i\in E} \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right)\\
=&\limsup_{B\to\infty} \frac{1}{B}\log|E| + \max_{i\in E}\limsup_{B\to\infty} \frac{1}{B}\log \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right)\\
=&\max_{i\in E}\limsup_{B\to\infty} \frac{1}{B}\log \mathbf{P}\left(\hat{Z}_i^B(\delta) \ge \delta W\right).
\end{aligned}$$

□

The following main result then provides a lower bound on the local-rate function defined in (9).

Theorem 6. *Assume that for some $\epsilon_0 > 0$, inequality (2) holds. For any small and positive ξ such that $0 < \xi < \epsilon_0$, define ϵ as (5). Given t, also assume that $\vec{q}^B(t) = \vec{q} \in \mathcal{Q}$, where $\mathcal{Q}$ is a closed and bounded set such that*

$V(\vec{q}) \geq v > 0$ for all $\vec{q} \in \mathcal{Q}$. There exists $\delta_0 > 0$ such that for all $0 < \delta \leq \delta_0$ and $W > 0$,

$$-\limsup_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}\left(\Delta v^B(\delta,t) \geq \delta W \,\middle|\, \vec{q}^B(t) = \vec{q}\right)$$
$$\geq \delta \min_{i\in E} \inf_{0\leq d\leq 1-\epsilon} (I_i^A(d+W) + I^{\hat{D}}(d)). \qquad (13)$$

Proof. Since we assume finite arrivals, we have $0 \leq A_i(n) < |\mathcal{E}_i| A_M, 0 \leq \hat{D}(n) \leq 1$. From Theorem 4.5.3 of [17], since the set $\{(a,d) | W \leq a-d \leq |\mathcal{E}_i|A_M, 0 \leq d \leq 1-\epsilon\}$ is compact for any $W > 0$, there exist a rate function $I_i^{A\hat{D}}(a,d)$ such that

$$\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}\left(A_i^B(\delta) - \hat{D}^B(\delta) \geq \delta W\right) \leq -\delta \inf_{W\leq a-d\leq A_M} I_i^{A\hat{D}}(a,d), \quad (14)$$

where

$$I_i^{A\hat{D}}(a,d) = \sup_{(\theta_1,\theta_2)\in\mathbb{R}^2} \left\{\theta_1 a + \theta_2 d - \limsup_{B\to\infty} \frac{1}{B\delta} \log \mathbf{E}\left(e^{B\theta_1 A_i^B(\delta) + B\theta_2 \hat{D}^B}\right)\right\}. \quad (15)$$

Since $A_i(n)$ and $\hat{D}(n)$ are independent, we have $I_i^{A\hat{D}}(a,d) = I_i^A(a) + I^{\hat{D}}(d)$.

Choose δ_0 as in Lemma 5. According to (11), (12) and (14), we have

$$\limsup_{B\to\infty} \frac{1}{B} \log \sup_{\vec{q}\in\mathcal{Q}} \mathbf{P}\left(\Delta v^B(\delta,t) \geq \delta W \,\middle|\, \vec{q}^B(t) = \vec{q}\right)$$
$$\leq -\delta \min_{i\in E} \inf_{W\leq a-d\leq A_M} (I_i^A(a) + I^{\hat{D}}(d))$$

The mean of $\hat{D}(n)$ is $1-\epsilon$. It follows that $I^{\hat{D}}(d) \geq I^{\hat{D}}(1-\epsilon) = 0$ when $d \geq 1-\epsilon$. When $0 < \xi < \epsilon_0$, we have $W + 1 - \epsilon > \bar{A}_i$ for all links i by our assumption. If $d \geq 1-\epsilon$ and $a - d > W$, we have $a > W + d \geq W + 1 - \epsilon > \bar{A}_i$. Because $I_i^A(a)$ is increasing when $a > \bar{A}_i$, we have,

$$\inf_{\substack{W\leq a-d\leq A_M \\ d\geq 1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)) = \inf_{d\geq 1-\epsilon} (I_i^A(W+d) + I^{\hat{D}}(d)) = I_i^A(W+1-\epsilon).$$

In fact, if $d = 1-\epsilon$, we will have

$$I_i^A(W+1-\epsilon) = \inf_{\substack{W\leq a-d\leq A_M \\ d=1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)).$$

Therefore,

$$\inf_{\substack{W \le a-d \le A_M \\ d \ge 1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)) \ge \inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)),$$

and hence,

$$\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}\left(\Delta v^B(\delta,t) \ge \delta W \,\middle|\, \vec{q}^B(t) = \vec{q}\right) \le -\delta \min_{i\in E} \inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)).$$

Further, we claim that

$$\inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)) = \inf_{0 \le d \le 1-\epsilon} (I_i^A(d+W) + I^{\hat{D}}(d)). \tag{16}$$

To see this, note that

$$\begin{aligned} &\inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon}} (I_i^A(a) + I^{\hat{D}}(d)) \\ &= \min\left\{ \inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon \\ d+W \ge \bar{A}_i}} (I_i^A(a) + I^{\hat{D}}(d)) \, , \inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon \\ d+W < \bar{A}_i}} (I_i^A(a) + I^{\hat{D}}(d)) \right\}. \end{aligned}$$

For the first term, since $a \ge d+W \ge \bar{A}_i$, and $I_i^A(a)$ is increasing when $a > \bar{A}_i$, we have,

$$\inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon \\ d+W \ge \bar{A}_i}} (I_i^A(a) + I^{\hat{D}}(d)) = \inf_{\substack{a-d=W \\ d \le 1-\epsilon \\ d+W \ge \bar{A}_i}} (I_i^A(a) + I^{\hat{D}}(d)).$$

For the second term, since a could be taken to be equal to the mean, we have

$$\inf_{\substack{W \le a-d \le A_M \\ 0 \le d \le 1-\epsilon \\ d+W < \bar{A}_i}} (I_i^A(a) + I^{\hat{D}}(d)) = \inf_{\substack{0 \le d \le 1-\epsilon \\ d+W < \bar{A}_i}} I^{\hat{D}}(d).$$

However, because $\bar{A}_i - W < 1-\epsilon$ and $I^{\hat{D}}(d)$ is decreasing when $d \le 1-\epsilon$, the above quantity is actually no greater than $I^{\hat{D}}(\bar{A}_i - W)$, which is included in infimum in the first term. Hence the claim (16) holds and the local rate function can be bounded as (13). $\square$

3.2. The Lower Bound on the Decay-rate of the Overflow Probability

For fixed $T > 0$, we now derive a lower bound on $I_0^T(\vec{\lambda})$ in this subsection. Fix a small $v \in (0,1)$, choose δ_0 as in Theorem 6. Let $0 < \delta < \delta_0$ such that $\delta = T/n$ for some integer n. Since $v(0) = 0$ and the arrivals are bounded, there exists $v_{\max} > 0$, such that $v^B(t) \le v_{\max}$ for all $t \in [0,T]$. Fix $\zeta < v$. Given $v_0 = 0, v_n = 1, 0 \le v_1, ..., v_{n-1} \le v_{\max}$, define $\Gamma_k(v_k) = \left\{v_k - \zeta \le v^B(k\delta) \le v_k + \zeta\right\}$ for $k = 1, 2, ..., n-1$, $\Gamma_0(v_0) = \left\{v^B(0) = v_0\right\}$ and $\Gamma_n(v_n) = \left\{v^B(n\delta) \ge v_n\right\}$. Let m_v, $0 \le m_v < n$, be the largest integer m such that $v_{m-1} < 2v$. We first fix $v_0, ..., v_n$. For ease of exposition, we will use Γ_k to denote $\Gamma_k(v_k)$ when there is no source of confusion. Consider the follow probability:

$$\mathbf{P}\left(\Gamma_1 \cap ... \cap \Gamma_n \left| \Gamma_0 \right.\right) \tag{17}$$

Roughly speaking, this is the probability that the trajectory $v^B(t)$ follows $v_0, v_1, ..., v_n$ given that it starts at v_0. Clearly,

$$\begin{aligned}&(17) \le \\ &\mathbf{P}\left(\Gamma_n \left|\Gamma_{n-1}, ..., \Gamma_{m_v}, \Gamma_0\right.\right) \times \mathbf{P}\left(\Gamma_{n-1} \left|\Gamma_{n-2}, ..., \Gamma_{m_v}, \Gamma_0\right.\right) \times ... \times \mathbf{P}\left(\Gamma_{m_v} \left|\Gamma_0\right.\right).\end{aligned}$$

Define $\Psi_k^B(\delta,\zeta) \triangleq \left\{\Delta v^B(\delta,(k-1)\delta) \ge v_k - v_{k-1} - 2\zeta\right\}$. Let

$$\mathcal{Q}_k = \left\{ \begin{array}{l} \{\vec{q} | V(\vec{q}) \in [v_{k-1} - \zeta, v_{k-1} + \zeta]\}, k = 2, ..., n \\ \{\vec{q} | V(\vec{q}) = v_0\}, k = 1 \end{array} \right. .$$

Also define

$$\phi_k^B(\delta,\zeta) = \sup_{\vec{q} \in \mathcal{Q}_k} \mathbf{P}\left(\Psi_k^B(\delta,\zeta) \left| \vec{q}^B((k-1)\delta) = \vec{q}\right.\right).$$

We then have the following lemma.

Lemma 7. *For any k, $m_v < k \le n$, the following holds*

$$\mathbf{P}\left(\Gamma_k \left|\Gamma_{k-1}, ..., \Gamma_{m_v}, \Gamma_0\right.\right) \le \phi_k^B(\delta,\zeta). \tag{18}$$

Proof.

$$\begin{aligned}\mathbf{P}\left(\Gamma_k \left|\Gamma_{k-1}, ..., \Gamma_{m_v}, \Gamma_0\right.\right) \le& \mathbf{P}\left(\Psi_k^B(\delta,\zeta) \left|\Gamma_{k-1}, ..., \Gamma_{m_v}, \Gamma_0\right.\right) &(19)\\ =& \frac{\mathbf{P}\left(\left(\bigcap_{i=m_v}^{k-1} \Gamma_i\right) \bigcap \Psi_k^B(\delta,\zeta) \left|\Gamma_0\right.\right)}{\mathbf{P}\left(\bigcap_{i=m_v}^{k-1} \Gamma_i \left|\Gamma_0\right.\right)}. &(20)\end{aligned}$$

The numerator of (20) can be written as

$$
\begin{aligned}
&\mathbf{E}\Big\{\mathbf{1}_{\left\{\left(\bigcap_{i=m_v}^{k-1}\Gamma_i\right)\bigcap\Psi_k^B(\delta,\zeta)\right\}}\,|\Gamma_0\Big\}\\
=&\mathbf{E}\Big\{\mathbf{E}\Big\{\mathbf{1}_{\left\{\left(\bigcap_{i=m_v}^{k-1}\Gamma_i\right)\bigcap\Psi_k^B(\delta,\zeta)\right\}}\,\Big|\vec{q}^B((k-1)\delta),...,\vec{q}^B(m_v\delta)\Big\}\Big|\Gamma_0\Big\}\\
=&\mathbf{E}\Big\{\mathbf{E}\Big\{\mathbf{1}_{\left\{\Psi_k^B(\delta,\zeta)\right\}}\,\Big|\vec{q}^B((k-1)\delta),...,\vec{q}^B(m_v\delta)\Big\}\cdot\mathbf{1}_{\left\{\bigcap_{i=m_v}^{k-1}\Gamma_i\right\}}\Big|\Gamma_0\Big\} \qquad (21)
\end{aligned}
$$

Since $v^B(t)$ satisfies the Markov property, conditioned on $\vec{q}^B((k-1)\delta)$, $\{\vec{q}^B(t), t > (k-1)\delta\}$ is independent from $\{\vec{q}^B(t), t < (k-1)\delta\}$. Define,

$$
\begin{aligned}
f_k(\vec{q}_{k-1},...,\vec{q}_{m_v}) \triangleq&\mathbf{E}\left\{\mathbf{1}_{\left\{\Psi_k^B(\delta,\zeta)\right\}}\,\Big|\vec{q}^B((k-1)\delta)=\vec{q}_{k-1},...,\vec{q}^B(m_v\delta)=\vec{q}_{m_v}\right\}\\
=&\mathbf{P}\left(\Psi_k^B(\delta,\zeta)\,\Big|\vec{q}^B((k-1)\delta)=\vec{q}_{k-1}\right).
\end{aligned}
$$

By the definition of $\phi_k^B(\delta,\zeta)$, we have

$$
f_k(\vec{q}_{k-1},...,\vec{q}_{m_v}) \le \phi_k^B(\delta,\zeta),
$$

when $\vec{q}_{k-1} \in \mathcal{Q}_k$. Then

$$
\begin{aligned}
(21) =&\mathbf{E}\left\{f_k(\vec{q}^B((k-1)\delta),...,\vec{q}^B(\delta))\cdot\mathbf{1}_{\left\{\bigcap_{i=m_v}^{k-1}\Gamma_i\right\}}\Big|\Gamma_0\right\}\\
\le&\phi_k^B(\delta,\zeta)\cdot\mathbf{E}\left\{\mathbf{1}_{\left\{\bigcap_{i=m_v}^{k-1}\Gamma_i\right\}}\Big|\Gamma_0\right\}.
\end{aligned}
$$

Noting that

$$
\mathbf{E}\left\{\mathbf{1}_{\left\{\bigcap_{i=m_v}^{k-1}\Gamma_i\right\}}\Big|\Gamma_0\right\}=\mathbf{P}\left(\bigcap_{i=m_v}^{k-1}\Gamma_i\,|\Gamma_0\right),
$$

we therefore have (20) $\le \phi_k^B(\delta,\zeta)$. □

We now can prove Theorem 2.

Proof of Theorem 2. For any $\zeta > 0$, there exists a finite set $\mathcal{V}$ of vectors $(v_n,...,v_0)$, $v_0 = 0, v_n = 1, 0 \le v_1,...,v_{n-1} \le v_{\max}$, such that

$$
\begin{aligned}
&\bigcup_{(v_n,...,v_0)\in\mathcal{V}}\Gamma_{n-1}(v_{n-1})\times...\times\Gamma_1(v_1) \supseteq\\
&\Big\{\left(v^B((n-1)\delta),...,v^B(\delta)\right)\Big|0\le v^B((n-1)\delta)\le v_{\max},...,0\le v^B(\delta)\le v_{\max}\Big\},
\end{aligned}
$$

where $\times$ denotes the Cartesian product. Taking the advantage of Lemma 7, one can show that

$$\mathbf{P}\left(v^B(T) \geq 1 \,\middle|\, v^B(0) = v_0\right) \leq \sum_{(v_n,\ldots,v_0)\in\mathcal{V}} \prod_{k=m_v+1}^{n} \phi_k^B(\delta,\zeta).$$

If we take the log and let B go to infinity, we will have

$$\begin{aligned}&\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}\left(v^B(T) \geq 1 \,\middle|\, v^B(0) = v_0\right)\\ \leq &\max_{(v_n,\ldots,v_0)\in\mathcal{V}} \limsup_{B\to\infty} \frac{1}{B} \sum_{k=m_v+1}^{n} \log \phi_k^B(\delta,\zeta).\end{aligned}$$

To estimate the limit in the above inequality, we will use the local rate function derived in Section III.A. From the definition of $\phi_k^B(\delta,\zeta)$ and Theorem 6, if $\zeta < 0$, since $v_k \geq 2v, m_v < k \leq n$, we will have

$$\limsup_{B\to\infty} \frac{1}{B} \log \phi_k^B(\delta,\zeta) \leq -(v_k - v_{k-1} - 2\zeta)L$$

if $v_k - v_{k-1} - 2\zeta \geq 0$. On the other hand, if $v_k - v_{k-1} - 2\zeta < 0$, the above inequality still holds because the left hand side is always less than 0. Therefore, taking the sum from $k = m_v + 1$ to n, one will get

$$\begin{aligned}&\max_{(v_n,\ldots,v_0)\in\mathcal{V}} \limsup_{B\to\infty} \frac{1}{B} \sum_{k=m_v+1}^{n} \log \phi_k^B(\delta,\zeta)\\ \leq &- \min_{(v_n,\ldots,v_0)\in\mathcal{V}} \left(1 - 2v - 2(n - m_v)\zeta\right) L\\ \leq &- \left(1 - 2v - 2n\zeta\right) L\end{aligned}$$

Let $v \to 0$ and $\zeta \to 0$, we have

$$I_0^T(\vec{\lambda}) = -\limsup_{B\to\infty} \frac{1}{B} \log \mathbf{P}\left(v^B(T) \geq 1 \,\middle|\, v^B(0) = 0\right) \geq L.$$

□

4. An Upper Bound

In this section, we will develop an upper bound for the decay-rate function $J_0(\vec{\lambda})$. We use the notion of "interference degree" that is first introduced in

[7]. The interference degree K_l of a link l is the maximum number of links in its interference set $\mathcal{E}_l$ that can transmit simultaneously. The interference degree $K(V,E)$ of a network $G(V,E)$ is the maximum interference degree over all links in E, i.e., $K(V,E) = \max_{l \in E} K_l$. For a given network topology, the interference degree can be obtained based on the interference model we use. For some classes of interference models, $K(V,E)$ can be bounded for all networks. For instance, if we consider node-exclusive interference model, we will have $K(V,E) \leq 2$. For any given network $G(V,E)$, let its interference degree be K. Consider an fictitious algorithm such that

$$\mathbf{P}\left(\sum_{l \in \mathcal{E}_i} \frac{d_l(n)}{c_l} = K\right) = 1, \text{for all } n.$$

Clearly, this fictitious algorithm will provide a lower bound on the overflow probability over all possible algorithms. We denote such algorithm as OPTIMAL. Note that OPTIMAL may not exist, and it is only used to derive an upper bound on $J_0(\vec{\lambda})$. We now consider (3) under algorithm OPTIMAL. Let $Z_i^B(t), A_i^B(t), D_i^B(t)$ have the same meaning as before. Now, the derivation are much easier since $Z_i^B(t)$ are i.i.d. across t. According to Theorem 6.6 in [18], the overflow rate function of OPTIMAL for each link i is given by

$$\lim_{B \to \infty} \frac{1}{B} \log \mathbf{P}\left(\sum_{l \in \mathcal{E}_i} \frac{q_l^B(0)}{c_l} > 1\right) = -\inf_{x > 0} \frac{I_{i,\text{opt}}^Z(x)}{x}$$

where $I_{i,\text{opt}}^Z(x)$ is the rate function of Z_i^B under OPTIMAL algorithm. It is trivial to show that $I_{i,\text{opt}}^Z(x) = I_i^A(x+K)$. Hence, the decay-rate of the queue-overflow probability of this fictitious system is given by

$$\lim_{B \to \infty} \frac{1}{B} \log \mathbf{P}\left(v^B(0) > 1\right) = -\min_{i \in E} \inf_{a > 0} \frac{I_i^A(a+K)}{a}. \tag{22}$$

Then, we have the following upper bound:

$$J_0(\vec{\lambda}) \leq I_{\text{opt}} \triangleq \min_{i \in E} \inf_{a > 0} \frac{I_i^A(a+K)}{a}.$$

We now pose a constraint on this decay rate function. Suppose that we want to guarantee that $I_{\text{opt}} \geq \theta_0$. Let Λ_i^A be the cumulant generating function of the arrival process, $\Lambda_i^A(\theta) = \log \mathbf{E}\left(e^{\theta A_i(1)}\right)$. Note that I_i^A is the Legendre transform of Λ_i^A and Λ_i^A is also the Legendre transform of I_i^A. Therefore, the following holds

$$\begin{aligned}I_{opt} \geq \theta_0 \Leftrightarrow &\inf_{a>0} \frac{I_i^A(a+K)}{a} \geq \theta_0, \forall i \in E \\ &\Leftrightarrow \sup_a \left\{\theta_0 a - I_i^A(a)\right\} - K\theta_0 \leq 0, \forall i \in E \\ &\Leftrightarrow \Lambda_i^A(\theta_0) - K\theta_0 \leq 0, \forall i \in E \Leftrightarrow \max_{i\in E} \frac{\Lambda_i^A(\theta_0)}{\theta_0} \leq K. \end{aligned} \tag{23}$$

The quantity $\frac{\Lambda_i^A(\theta_0)}{\theta_0}$ is often called the effective bandwidth of the arrival process. The inequality (23) implies that the maximum possible effective capacity region of the system under any algorithm is such that the effective bandwidth in every interference range must be no greater than K.

5. Comparisons

We now compare our lower bound in Section III with the upper bound in Section IV. Note that when the value of ϵ_0 in (2) satisfies $0 < \xi < \epsilon_0$ and $W > 0$, we have $W + 1 - \epsilon > \bar{A}_i$ for all links i by our assumption. If $d \geq 1 - \epsilon$, we have $W + d \geq W + 1 - \epsilon > \bar{A}_i$, which means that $I_i^A(d+W)$ and $I^{\hat{D}}(d)$ are increasing with respect to d. Therefore,

$$\inf_{d\leq 1-\epsilon} I_i^A(d+W) + I^{\hat{D}}(d) = \inf_d I_i^A(d+W) + I^{\hat{D}}(d).$$

Then, we have

$$\begin{aligned}L \geq \theta_0 \Leftrightarrow &\inf_{W>0} \min_{i\in E} \inf_{d\leq(1-\epsilon)} \{I_i^A(d+W) + I^{\hat{D}}(d) - \theta W\} \geq 0 \\ &\Leftrightarrow \inf_{W>0,d} \{I_i^A(d+W) + I^{\hat{D}}(d) - \theta W\} \geq 0, \forall i \in E.\end{aligned}$$

Further, note that $I_i^A(d+W) + I^{\hat{D}}(d) - \theta W \geq 0$ if $W \leq 0$. Hence, we have, for all $i \in E$

$$\begin{aligned}&\inf_{W>0,d} \{I_i^A(d+W) + I^{\hat{D}}(d) - \theta W\} \geq 0 \\ \Leftrightarrow &\inf_{W,d} \{I_i^A(d+W) + I^{\hat{D}}(d) - \theta W\} \geq 0 \\ \Leftrightarrow &\sup_{W+d} \{(d+W)\theta - I_i^A(d+W)\} + \sup d\{-d\theta - I^{\hat{D}}(d)\} \leq 0 \\ \Leftrightarrow &\Lambda_i^A(\theta) + \Lambda^{\hat{D}}(-\theta) \leq 0.\end{aligned}$$

Hence

$$L \geq \theta_0 \Leftrightarrow \Lambda_i^A(\theta) + \Lambda^{\hat{D}}(-\theta) \leq 0, \forall i \in E$$
$$\Leftrightarrow \max_{i \in E} \frac{\Lambda_i^A(\theta_0)}{\theta_0} \leq \frac{\Lambda^{\hat{D}}(-\theta_0)}{-\theta_0}.$$

Therefore, the effective capacity region of Q-SCHED is such that the sum of effective bandwidth in each interference range is no greater than $\frac{\Lambda^{\hat{D}}(-\theta_0)}{(-\theta_0)}$. Note that $\Lambda^{\hat{D}}(-\theta_0) = \log\left(\epsilon + (1-\epsilon)e^{-\theta_0}\right)$. Thus, the effective capacity region of Q-SCHED at a given constraint is at least $\frac{\log\left(\epsilon+(1-\epsilon)e^{-\theta_0}\right)}{-K\theta_0}$ of that of any other algorithms. We then consider two special cases.

5.1. Deterministic Arrivals

We first consider the case when the arrivals are deterministic, i.e., $a_i(n) = \lambda_i$ for all $i \in E$ and $n = 1, 2, ...$ Recall that $\bar{A}_i = \sum_{l \in \mathcal{E}_i} \frac{\lambda_l}{c_l}$ is the mean of $A_i(n)$, and the rate function of A_i will be

$$I_i^A(a) = \begin{cases} 0, a = \bar{A}_i \\ \infty, \text{otherwise} \end{cases}.$$

Under the assumption of deterministic arrival, we have the effective bandwidth of arrival process equals to its mean, i.e., $\Lambda_i^A(\theta_0)/\theta_0 = \bar{A}_i$. Hence,

$$L \geq \theta_0 \Leftrightarrow \Lambda^{\hat{D}}(-\theta_0)/(-\theta_0) \geq \bar{A}_i, \forall i \in E.$$

Thus, under the assumption of deterministic arrivals the effective capacity region of Q-SCHED at a given constraint is still at least $\frac{\log\left(\epsilon+(1-\epsilon)e^{-\theta_0}\right)}{-K\theta_0}$ of that of any other algorithms.

5.2. Infinite Number of Mini-slots

In this case, we assume that the number of mini-slots M is infinite. We also assume that the variable ξ in Lemma 3 is 1, which implies that ϵ in Lemma 3 equals to 0. It follows that the rate function of $\hat{D}$ will be:

$$I^{\hat{D}}(d) = \begin{cases} 0, d = 1 \\ \infty, \text{otherwise} \end{cases}.$$

Consequently, the right-hand-side of (6) can be simplified

$$L = \inf_{W \geq 0} \min_{i \in E} \frac{I_i^A(W+1)}{W}. \tag{24}$$

Using the same method as in the section IV, if we pose a constraint on the decay rate, e.g. if we want to guarantee that $L \geq \theta_0$, then we could get the "effective bandwidth" in this case: $\max_{i \in E} \frac{\Lambda_i^A(\theta_0)}{\theta_0} \leq 1$. This result implies that, at a given θ_0, the effective capacity region of the Q-SCHED algorithm is such that the sum of the effective bandwidth in each interference range is no greater than 1. Comparing with (23), we note that the effective capacity region of Q-SCHED is at least $1/K$ of that of any other algorithms.

6. Simulation Result

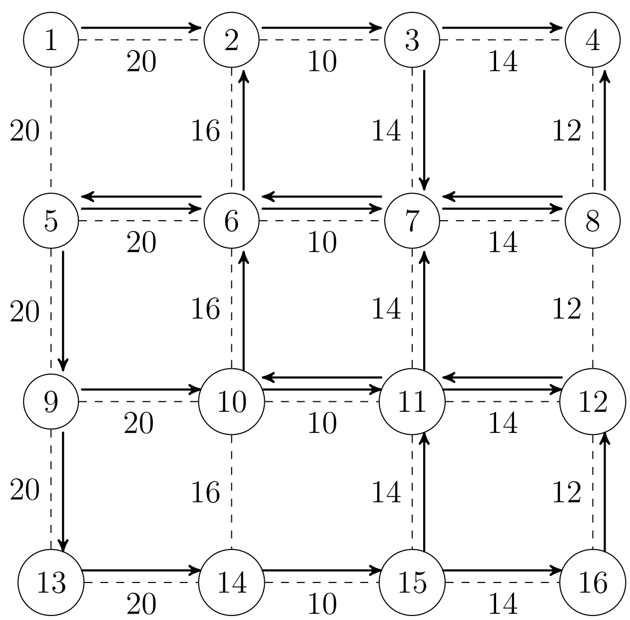

Figure 1: Network Topology

In this section, we provide the simulation results to verify our earlier results. We simulate a single-hop network with topology shown in Figure 1, which has 16 nodes. The deashed lines represent links. The number next to each link denotes its capacity. Tha arrows represent the flows. The arrival rate of all flows are the same. We run the Q-SCHED algorithm for many slots. In each time slot, we measure the quantity $\max_{i \in E} \sum_{l \in \mathcal{E}_i} \frac{q_l}{c_l}$ and compare it with a given threshold B to see if overflow occurs. The ratio between the number of times that overflow occurs and the total number of time slots will give the probability of queue-overflow. In the following, we consider two interference models: the one-hop interference model and the two-hop interference model.

6.1. One-Hop Interference Model

In this subsection we provide the simulation results using the one-hop interference model. First we assume that packets of each flow arrive at constant rates and let the arrival rate of each flow to be 1. The result is shown in Figure 2. We plot the overflow probability $\mathbf{P}\left(v^B(0) > 1\right)$ against the overflow threshold B with different numbers of minislots for Q-SCHED. We have also plotted two corresponding lines whose slope are calculated by the lower bound given in Theorem 2 with different numbers of minislots. We see that the actual overflow probability decays faster than the lower bound given by our analytical results in Theorem 2. We also observe in Figure 2 that as the number of minislots increases, the overflow probability decreases which is coincident with the nature of Q-SCHED.

We then set the volume of data that arrive in each time slot to be a uniformly distributed random variable over $[0, 2]$. Note that the mean arrival rate of each flow is still 1. The result is shown as Figure 3. We plot the overflow probability $\mathbf{P}\left(v^B(0) > 1\right)$ against the overflow threshold B with different numbers of minislots for Q-SCHED. We can also confirm that the actual overflow probability decays faster than the lower bound in Theorem 2. We also draw a line with slope equal to (24), which is the lower bound on the decay rate for the case when we have infinite minislots. The decay rate is larger if the number of minislots is larger and the decay rate is the largest if we allow infinite minislots. By comparing Figure 2 and Figure 3, we observe that under the same overflow threshold B, the overflow probability under random arrivals is slightly larger that the overflow probability under deterministic arrivals, i.e., if we add more randomness in the arrival, the overflow probability will increase.

One can get the capacity region for Q-SCHED once the arrival processes are given. We choose the arrival rate of each flow to be the same. We can use simulation to obtain the maximum possible arrival rate for the

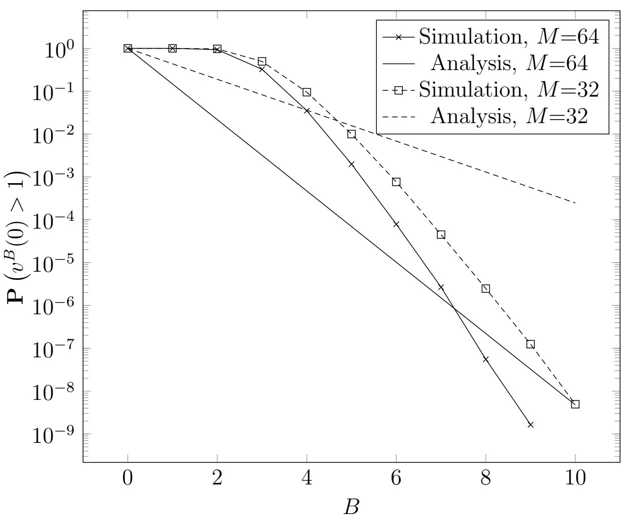

Figure 2: Deterministic Arrivals under One-Hop Interference Model: The Overflow Probability versus Overflow Threshold

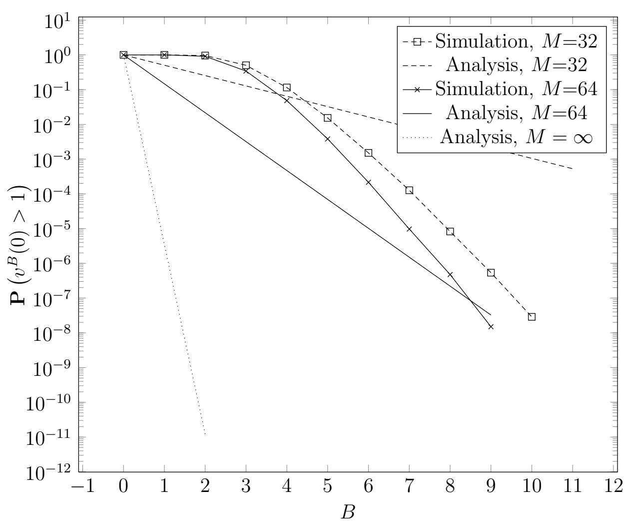

Figure 3: Random Arrivals under One-Hop Interference Model: The Overflow Probability versus Overflow Threshold

network that satisfies a given constraint on the overflow probability that $\mathbf{P}\left(v^B(0) > 1\right) < p_0$. Correspondingly, we can convert the constraint on the overflow probability p_0 to a constraint on the decay rate θ_0 with the approximation that the overflow probability is equal to $e^{B\theta_0}$. We can then use Theorem 2 to calculate the maximum arrival rate subject to the constraint that the lower bound is greater than θ_0. This quantity then provides a lower

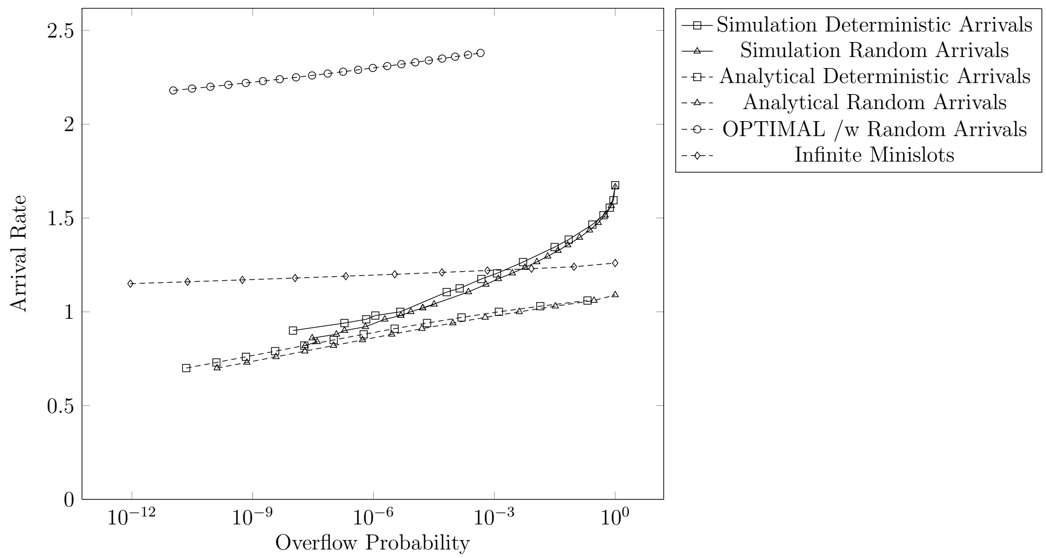

Figure 4: Capacity Region versus Overflow Probability, One-Hop Interference Model

bound on the effective capacity of the system. We plot the maximum possible arrival rate under computed from simulation and from the analytical lower bound. The result is shown in Figure 4. We set the number of minislots to be 32 under all cases expect the case infinite minislots and the case for OPTIMAL algorithm. We also set the overflow threshold $B = 8$ in all cases. We see that the capacity region for our analytical lower bound is only slightly smaller than the one obtained by simulation. We also see that, when we have infinite number of minislots, the lower bound "capacity region" is more than 1/2 of the capacity region for OPTIMAL. One can also verify that when we have deterministic arrivals, the lower bound of the capacity region under a fix overflow probability p_0, is larger than a fraction $\frac{\log\left(\epsilon+(1-\epsilon)e^{-\theta_0}\right)}{-2\theta_0}$ of the capacity region for OPTIMAL, where $\theta_0 = -\frac{1}{B}\log p_0$ is the corresponding

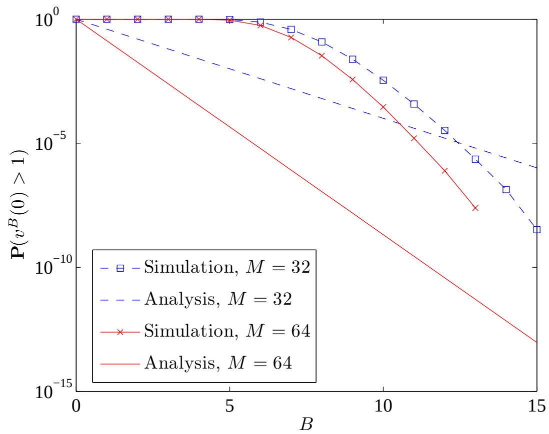

Figure 5: Deterministic Arrivals under Two-Hop Interference Model: The Overflow Probability versus Overflow Threshold

decay rate.

6.2. Two-Hop Interference Model

We next change the interference model to the two-hop interference model on the same network in Figure 1. We first simulate the case with deterministic arrivals. Data arrive to each flow at a constant rate of 1/2. In Figure 5, we plot the overflow probability $\mathbf{P}\left(v^B(0) > 1\right)$ against the overflow threshold B with two different choices on the number of minislots for Q-SCHED. Once again, we confirm that the actual overflow probability decays faster than the lower bound given by our analytical results in Theorem 2. Note that due the larger interference set, the overflow probability under the two-hop interference model (Figure 5) is larger than the overflow probability under one-hop interference model (Figure 2), even with smaller arrival rate.

We then set the volume of the data that arrive in each time slot to be a random variable with uniform distribution over [0,1]. The result is shown as Figure 6. We also see that the actual overflow probability decays faster than the lower bound in Theorem 2. By comparing Figure 5 and Figure 6, the same conclusion could be made as we did when we compare Figure 2 and Figure 3, i.e., if we add randomness in the arrival, the overflow probability

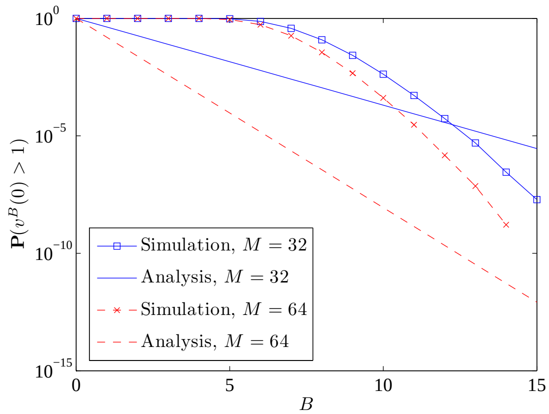

Figure 6: Random Arrivals under Two-Hop Interference Model: The Overflow Probability versus Overflow Threshold

will increase.

7. Conclusion and Discussion

In this paper, we developed a lower bound on the decay rate of the overflow probability for scheduling algorithm for ad-hoc wireless networks called Q-SCHED. We also show that the effective capacity that Q-SCHED could support is a provable fraction of the maximum possible effective capacity over all other algorithms, subject to a given constraint on the decay-rate of the queue-overflow probability. For future work, we will extend the approach of this paper to other wireless scheduling algorithms and other types of performance guarantees. We believe that the techniques developed in this paper can be used to study other distributed scheduling algorithms such as maximal matching and greedy maximal matching [6–8]. Essentially, what we need is some knowledge of the statistics of service rate process like that in Lemma 1. Then, using the technique in Theorem 2 and Section 4 we can compute the lower and upper bounds on the decay rate of the queue-overflow probability.

The lower bound on the decay-rate of the overflow probability can be mapped to a corresponding lower bound on the decay rate of the delay-violation probability as in (4), which can then be used to bound the delay-violation probability. We note however that such bounds on delay-violation probability are probably not the tightest, since we cannot derive a matching upper bound on the decay-rate of the delay-violation probability similar to (4). Hence, we believe that Q-SCHED cannot guarantee a constant fraction of the effective system capacity subject to a given constraint on the delay-violation probability. For future work, we are interested in algorithms that can achieve a constant fractions of the effective system capacity even subject to delay constrains. We conjecture that such algorithms will likely use delay, rather than queue-length, to make the scheduling decisions.

Finally, our analysis in this paper is restricted to single-hop networks. For future works, we will study how to extend the techniques to multihop systems, possibly using the techniques in [19].

Appendix: Proof of Lemma 4

The inequality we need to show is equivalent to the following:

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{\lfloor B\delta \rfloor} Z_i(n) \geq BW\right) \leq \mathbf{P}\left(\sum_{n=1}^{\lfloor B\delta \rfloor} \hat{Z}_i(n) \geq BW\right).$$

According to Lemma 3, we could choose δ_0 and B_0 such that for any $i \in \mathcal{I}(\vec{q}, \delta)$, when $B \geq B_0$ and $\delta \leq \delta_0$ the following holds for $n \leq B\delta$

$$\mathbf{P}_{\vec{q}}(D_i(n) \geq 1 \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) \geq 1 - \epsilon,$$
$$\mathbf{P}_{\vec{q}}(D_i(n) = 0 \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) \leq \epsilon.$$

Note that, δ_0 exists because $\mathcal{I}(\vec{q}, \delta)$ is a finite set. We use induction to prove that for any N, $1 \leq N \leq \lfloor B\delta_0 \rfloor$

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N} Z_i(n) \geq BW\right) \leq \mathbf{P}\left(\sum_{n=1}^{N} \hat{Z}_i(n) \geq BW\right), \qquad (25)$$

and then the result of Lemma 4 will holds.

We first show that the induction hypothesis (25) holds for $N = 1$. Note that, for $i \in \mathcal{I}(\vec{q}, \delta)$,

$$\mathbf{P}_{\vec{q}}(D_i(n) \geq 0 \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) = 1 = \mathbf{P}(\hat{D}(n) \geq 0),$$
$$\mathbf{P}_{\vec{q}}(D_i(n) \geq 1 \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) \geq 1 - \epsilon = \mathbf{P}(\hat{D}(n) \geq 1)$$

Moreover, for any $k \geq 2$,

$$\mathbf{P}_{\vec{q}}(D_i(n) \geq k \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) \geq 0 = \mathbf{P}(\hat{D}(n) \geq k).$$

Hence, for any $k = 0, 1, 2, ...$

$$\mathbf{P}_{\vec{q}}(D_i(n) \geq k \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) \geq \mathbf{P}(\hat{D}(n) \geq k),$$
$$\mathbf{P}_{\vec{q}}(D_i(n) \leq k \,|D_i(n-1), A_i(n-1), ..., D(1), A(1)\,) \leq \mathbf{P}(\hat{D}(n) \leq k).$$

Since $A_i(1)$, $D_i(1)$ and $\hat{D}(1)$ are independent, we have for any k,

$$\begin{aligned}\mathbf{P}_{\vec{q}}(D_i(1) \leq k \,|A_i(1)\,) &= \mathbf{P}_{\vec{q}}(D_i(1) \leq k) \leq \mathbf{P}(\hat{D}(1) \leq k) \\ &= \mathbf{P}(\hat{D}(1) \leq k \,|A_i(1)\,).\end{aligned}$$

Hence,

$$\mathbf{P}_{\vec{q}}(D_i(1) \leq A_i(1) - BW \,|A_i(1)\,) \leq \mathbf{P}(\hat{D}(1) \leq A_i(1) - BW \,|A_i(1)\,).$$

Using total probability equation, we have

$$\mathbf{P}_{\vec{q}}(D_i(1) \leq A_i(1) - BW) \leq \mathbf{P}(\hat{D}(1) \leq A_i(1) - BW),$$

which means that the induction hypothesis (25) holds for $N = 1$, i.e.,

$$\mathbf{P}_{\vec{q}}(Z_i(1) \geq BW) \leq \mathbf{P}(\hat{Z}(1)) \geq BW).$$

Now assume that the induction hypothesis (25) holds for $N-1$, i.e., for any $i \in \mathcal{I}(\vec{q}, \delta)$,

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N-1} Z_i(n) \geq BW\right) \leq \mathbf{P}\left(\sum_{n=1}^{N-1} \hat{Z}_i(n) \geq BW\right).$$

Then

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N} Z_i(n) \geq BW\right) = \mathbf{P}_{\vec{q}}\left(D_i(n) \leq \sum_{n=1}^{N-1} Z_i(n) + A_i(n) - BW\right) \quad (26)$$

Since $A_i(N)$, $D_i(N)$ and $\hat{D}(N)$ are independent, using similar method as we did previously, we can show that

$$\begin{aligned}&\mathbf{P}_{\vec{q}}\left(D_i(N) \leq \sum_{n=1}^{N-1} Z_i(n) + A_i(n) - BW \,\middle|\, A_i(N), D_i(N-1), ..., A(1)\right) \\ &\leq \mathbf{P}_{\vec{q}}\left(\hat{D}(N) \leq \sum_{n=1}^{N-1} Z_i(n) + A_i(n) - BW \,\middle|\, A_i(N), D_i(N-1), ..., A(1)\right).\end{aligned}$$

Therefore, from total probability equation, we have

$$\mathbf{P}_{\vec{q}}\left(D_i(N) \leq \sum_{n=1}^{N-1} Z_i(n) + A_i(N) - BW\right) \tag{27}$$

$$\leq \mathbf{P}_{\vec{q}}\left(\hat{D}(N) \leq \sum_{n=1}^{N-1} Z_i(n) + A_i(N) - BW\right)$$

$$\leq \mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N-1} Z_i(n) \geq \hat{D}(N) - A_i(N) + BW\right). \tag{28}$$

However, $\sum_{n=1}^{N-1} Z_i(n)$ is independent from $\hat{D}(N)$ and $A_i(N)$. From the induction hypothesis for $N-1$, we then have

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N-1} Z_i(n) \geq \hat{D}(N) - A_i(N) + BW \,\middle|\, A_i(N), \hat{D}(N)\right)$$
$$\leq \mathbf{P}\left(\sum_{n=1}^{N-1} \hat{Z}(n) \geq \hat{D}(N) - A_i(N) + BW \,\middle|\, A_i(N), \hat{D}(N)\right)$$

Once again, using total probability equation, we get that

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N-1} Z_i(n) \geq \hat{D}(N) - A_i(N) + BW\right) \tag{29}$$

$$\leq \mathbf{P}\left(\sum_{n=1}^{N-1} \hat{Z}(n) \geq \hat{D}(N) - A_i(N) + BW\right) \tag{30}$$

From (26), (28) and (30), we can conclude that

$$\mathbf{P}_{\vec{q}}\left(\sum_{n=1}^{N} Z_i(n) \geq BW\right) \leq \mathbf{P}\left(\sum_{n=1}^{N} \hat{Z}_i(n) \geq BW\right).$$

This proves the induction hypothesis (25) for N. The result of the Lemma then follows.

References

[1] C. Zhao, X. Lin, On the queue-overflow probabilities of distributed scheduling algorithms, in: IEEE Conference on Decision

and Control, 2009, accepted for presentation, also available at http://cobweb.ecn.purdue.edu/∼linx/publications.html.

[2] L. Tassiulas, A. Ephremides, Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks, IEEE Transactions on Automatic Control 37 (12) (1992) 1936–1948.

[3] X. Lin, N. B. Shroff, The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks, in: Proceedings of IEEE INFOCOM, Miami, Florida, 2005, pp. 1794–1803.

[4] X. Lin, S. B. Rasool, Constant-time distributed scheduling policies for ad hoc wireless networks, in: Proceedings of IEEE Conference on Decision and Control, San Diego, 2006.

[5] A. Gupta, X. Lin, R. Srikant, Low-complexity distributed scheduling algorithms for wireless networks, in: IEEE INFOCOM, 2007.

[6] X. Wu, R. Srikant, J. R. Perkins, Queue-length stability of maximal greedy schedules in wireless networks, in: Information Theory and Applications Inaugural Workshop, University of California, San Diego, 2006.

[7] P. Chaporkar, K. Kar, S. Sarkar, Throughput guarantees through maximal scheduling in wireless networks, in: 43d Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, 2005.

[8] X. Lin, N. B. Shroff, The impact of imperfect scheduling on crosslayer congestion control in wireless networks, IEEE/ACM Transactions on Networking 14 (2) (2006) 302–315.

[9] G. Kesidis, J. Walrand, C.-S. Chang, Effective bandwidths for multicalss markov fluids and other atm sources, IEEE/ACM Transactions on Networking 1 (4) (1993) 424–428.

[10] D. Wu, R. Negi, Effective capacity: A wireless link model for support of quality of service, IEEE Transactions on Wireless Communications 2 (4) (2003) 630–643.

[11] A. Eryilmaz, R. Srikant, Scheduling with quality of service constraints over rayleigh fading channels, in: Proceedings of the IEEE Conference on Decision and Control, 2004.

[12] L. Ying, R. Srikant, A. Eryilmaz, G. E. Dullerud, A large deviations analysis of scheduling in wireless networks, IEEE Transactions on Information Theory 52 (11) (2006) 5088–5098.

[13] S. Shakkottai, Effective capacity and qos for wireless scheduling, IEEE Transactions on Automatic Control 53 (3) (2008) 749–761.

[14] V. J. Venkataramanan, X. Lin, On wireless scheduling algorithms for minimizing the queue-overflow probability, IEEE/ACM Transactions on Networking, to appear, also available at http://cobweb.ecn.purdue.edu/∼linx/publications.html.

[15] A. L. Stolyar, Large deviations of queues sharing a randomly time-varying server, Queueing Systems: Theory and Applications 59 (1) (2008) 1–35.

[16] X. Lin, On characterizing the delay performance of wireless scheduling algorithms, in: 44th Annual Allerton Conference on Communication, Control, and Compting, Monticello, IL, 2006.

[17] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, 2nd Edition, Springer, 1998.

[18] A. Ganesh, N. O'Connell, D. Wischik, Big Queues, Springer, 2004.

[19] V. Venkataramanan, X. Lin, L. Ying, S. Shakkottai, On scheduling for minimizing end-to-end buffer usage over multihop wireless networks, in: Proceedings of IEEE INFOCOM, San Diego, California, 2010.