Forty-Seventh Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 30 - October 2, 2009

The Impact of Queue Length Information on Buffer Overflow in Parallel Queues

Krishna Jagannathan and Eytan Modiano

***Abstract*—We consider a system consisting of N parallel queues, served by one server. Time is slotted, and the server serves one of the queues in each time slot, according to some scheduling policy. In the first part of the paper, we characterize the buffer overflow exponents and the likeliest overflow trajectories under the Longest Queue First (LQF) scheduling policy. Under statistically identical arrivals to each queue, we show that the buffer overflow exponent can be simply expressed in terms of the total system occupancy exponent of m parallel queues, for some $m \leq N$. We next turn our attention to the rate of queue length information needed to operate a scheduling policy, and its relationship to the buffer overflow exponents. It is known that LQF scheduling has superior overflow exponents compared to queue blind policies such as processor sharing (PS) and random scheduling. However, we show that the overflow exponent of the LQF policy can be preserved under arbitrarily infrequent queue length information.**

I. Introduction

Scheduling is an essential component of any queueing system where the server resources need to be shared between many queues. Perhaps the most basic requirement of a scheduling algorithm is to ensure the stability of all queues in the system, whenever feasible. Much research work has been reported on "throughput optimal" scheduling algorithms that achieve stability over the entire capacity region of a network [1], [2]. While stability is an important and necessary first-order metric, most practical queueing systems have more stringent Quality of Service (QoS) requirements. For example, streaming voice and video streams are delay sensitive. Further, due to the finiteness of the buffers in practical systems, maintaining a low buffer overflow probability is an important objective.

In this paper, we consider a system consisting of N parallel queues and a single server. A scheduling policy decides which of the queues gets service in each time slot. Our goal is to better understand the relation between the buffer overflow probability and the amount of queue length information required to operate a scheduling policy. The scheduling decisions may take into account the current queue lengths in the system, in which case we will call the policy 'queue aware.' If the scheduling decisions do not depend on the current queue lengths, except to the extent of knowing whether or not a queue is empty, we will call it a 'queue blind' policy.

In the first part of this paper, we analyze the large deviation behavior of the widely studied Longest Queue First (LQF) policy. We assume that the queues are fed by statistically identical arrival processes. However, the input statistics could otherwise be very general. Under such a symmetric traffic pattern, we show that the buffer overflow exponent under LQF scheduling is expressible purely in terms of the *total system occupancy* exponent of an m queue system, where $m \leq N$ is determined by the input statistics. We also show that the likeliest overflow trajectories are straight lines.

In the second part of the paper, we turn our attention to the rate of queue length information needed to operate a scheduling policy, and its relationship to the buffer overflow exponent. Although any work conserving policy (such as LQF, processor sharing (PS) or random scheduling) will achieve the same throughput region and total system occupancy distribution, the LQF policy outperforms the queue blind policies in terms of the buffer overflow probability. Equivalently, this implies that the buffer requirements are lower under LQF scheduling than under queue blind scheduling, if we want to achieve a given overflow probability. For example, our study indicates that under Bernoulli and Poisson traffic, the buffer size required under LQF scheduling is only about 55% of that required under random scheduling, when the traffic is relatively heavy. On the other hand, with LQF scheduling, the scheduler needs queue length information in *every* time slot, which leads to a significant amount of control signalling. Hence, we identify a "hybrid" scheduling policy, which achieves the same buffer overflow exponent as the LQF policy, with arbitrarily infrequent queue length information.

A. Related Work

To our knowledge, Bertsimas et. al. [4] were among the first to analyze the large deviations behavior of parallel queues. They consider the case of *two* parallel queues, and characterize the buffer overflow exponents under two important service disciplines, namely Generalized Processor Sharing (GPS) and Generalized Longest Queue First (GLQF). We also refer to the related papers [5]–[7] where the authors analyze a system of parallel queues, with deterministic arrivals and time-varying connectivity. In [8], the authors study large deviations for the largest weighted delay first policy, and [9] deals with large deviations of max-weight scheduling for general convex rate regions. In each case, the optimal exponent and the likeliest overflow trajectory are obtainable by solving a variational control problem. Often times, the optimal solution to the variational problem can be found by solving a finite dimensional optimal control problem [4], [8].

The rest of this paper is organized as follows. In Section II, we present the system description, and some preliminaries on large deviations. Our main result on the large deviation

This work was supported by NSF grants CNS-0626781 and CNS-0915988, and by ARO Muri grant number W911NF-08-1-0238.

The authors are with the Massachusetts Institute of Technology, Cambridge, MA 02139, USA. krishnaj@mit.edu, modiano@mit.edu

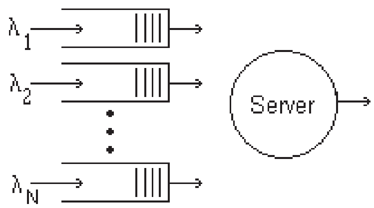


Fig. 1. N parallel queues served by one server

behavior of LQF scheduling is presented in Section III. Section IV compares LQF scheduling to queue blind scheduling in terms of the overflow probability. In Section V, we study scheduling policies with infrequent queue length information.

II. System Description and Preliminaries

Fig. 1 depicts a system consisting of N parallel queues, served by one server. We assume that time is slotted, and the server is capable of serving one packet per slot. Arrivals occur according to a random process $A_i(t), i = 1, \ldots, N$, which denotes the number of packets that arrive at queue i during slot t. The arrivals to the different queues are independent. We assume a symmetric traffic pattern, i.e., the arrival processes to each queue are statistically identical to each other. For ease of exposition, let us assume that the arrivals are independent across time slots, although our results hold under more general assumptions[1]. The average arrival rate to a queue is $\mathbb{E}\left[A_i[t]\right] = \lambda$ packets/slot for each i. For stability, we assume that the condition $\lambda < \frac{1}{N}$ is satisfied. Let us also define

$$S_i[t_1, t_2] = \sum_{\tau=t_1}^{t_2} A_i[\tau],\ t_1 \le t_2$$

as the number of arrivals to queue i between time slots t_1 to t_2.

The log-moment generating function of the input process[2] to each queue is assumed to exist, and is given by

$$\Lambda(\theta) = \log \mathbb{E}\left[\exp(\theta A_i[t])\right].$$

The *convex dual* of $\Lambda(\theta)$ is defined by

$$\Lambda^*(x) = \sup_{\theta}[\theta x - \Lambda(\theta)].$$

$\Lambda^*(x)$ is referred to as the *rate function* of the large deviation principle (LDP) satisfied by each input process.

We are interested in the probability of a buffer overflow, i.e., $\mathbb{P}\{\max_{i=1,\ldots,N} Q_i[0] \ge M\}$, under a given scheduling policy Π, where, $Q_i[0]$ is the queue length at time slot 0[3]. More specifically, we are interested in the exponent of the above probability under the large-buffer scaling, which is defined as

$$E_N^{\Pi} = \lim_{M \to \infty} -\frac{1}{M} \ln \mathbb{P}\{\max_{i=1,\ldots,N} Q_i[0] \ge M\}. \tag{1}$$

We emphasize that this exponent depends on the scheduling policy Π, as well as the system size N and the input statistics. We also define the exponent corresponding to the *total system occupancy* exceeding a certain limit:

$$\Theta_N = \lim_{q \to \infty} -\frac{1}{q} \ln \mathbb{P}\{\sum_{i=1}^{N} Q_i[0] \ge q\}. \tag{2}$$

As we shall see, the system occupancy exponent in (2) plays an important role in our analysis of the buffer overflow exponent (1). The following well-known lemma asserts that Θ_N is the same for all *work-conserving* scheduling policies.

Lemma 1: All work conserving policies achieve the same steady-state system occupancy distribution (and hence the same system exponent Θ_N).

In fact, the above result holds at a sample-path level, since one packet would leave the system every time slot if the system is not empty, under any work conserving policy.

We mainly analyze the Longest Queue First (LQF) scheduling policy, which, as the name states, serves the longest queue in each slot, with an arbitrary tie-breaking rule. We also consider two other work-conserving policies: *random scheduling* (RS), which serves a random occupied queue in each slot (each with equal probability), and *processor sharing* (PS), which divides the server capacity equally between all occupied queues. Note that LQF scheduling is queue-aware, while RS and PS are queue-blind.

III. Large Deviation Analysis of LQF Scheduling

In this section, we present our main results regarding the buffer overflow exponents and trajectories under LQF scheduling. We begin by characterizing the system occupancy exponent Θ_N for a work conserving policy.

Proposition 1: Under any work conserving policy, the system occupancy exponent is given by

$$\Theta_N = \inf_{a>0} \frac{1}{a} \Lambda^*(a + \frac{1}{N}) \tag{3}$$

Proof: (Outline) The result is a consequence of the fact that the total system occupancy distribution is the same as the queue length distribution of a single queue, served by the same server, but fed by the sum process $\sum_i A_i(t)$. Since the input processes to the different queues are independent and identically distributed (i.i.d), the log-moment generating function of the sum process is $N\Lambda(\theta)$. From the definition of the convex dual, the rate function of the sum process can be expressed as $N\Lambda^*(x/N)$. Once the rate function of the input process is known, the overflow exponent of a single server queue can be easily computed; see [10]. □

Let us denote by a_N^* the optimizing value of a in (3).

We now define scaled processes for the arrivals and queue lengths, which are often used to study sample path large

[1] We only need the input processes to satisfy a sample path large deviation principle (LDP), as detailed in [4].

[2] This definition applies when the inputs are independent across time. If the inputs are correlated across time slots, we define $\Lambda(\theta) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[\exp(\theta \sum_{t=1}^{n} A_i[t])\right]$.

[3] The queues are assumed to be initialized such that $Q_i[-MT] = 0, 1 \le i \le N$ for some $T > 0$. As $M \to \infty$ with T fixed, $Q_i[0]$ will approach the steady-state.

deviations in the large buffer regime. For every sample path that leads to a buffer overflow at time slot 0, there exists a time $-n \leq 0$ for which both queues are empty. Since we are interested in large M asymptotics, we let $T = -\frac{n}{M}$, and define the sequence of scaled processes[4]

$$q_i(t) = \frac{Q_i[\lfloor Mt \rfloor]}{M}, i = 1, \dots, N, \ t \in [-T, 0],$$

and

$$S_i(t) = \frac{S_i[-MT, \lfloor Mt \rfloor]}{M},$$

where $\lfloor \cdot \rfloor$ denotes the floor function. The initial condition implies that $q_i(-T) = 0, i \leq N$, and $q_i(0) = 1$ implies the overflow of queue i at time 0. We say that the input process on queue i has *empirical rate* $x_i(t) \geq 0$ in the interval $[-T, 0]$ if the following is satisfied for any $\epsilon > 0$ and large enough M :

$$\left| S_i(t) - \int_{-T}^{0} x_i(t) \mathrm{d}t \right| < \epsilon, \ \forall t \in [-T, 0]$$

Loosely speaking, the exponent corresponding to the input process on queue i exhibiting an empirical rate $x_i(t), \ t \in [-T, 0]$ is given by Mogulskii's theorem [11]:

$$\int_{-T}^{0} \Lambda^*(x_i(t)) \mathrm{d}t. \tag{4}$$

We remark that Mogulskii's theorem applies only to arrivals processes which are independent across time. If the arrivals are correlated in time, the results in this paper will still apply, if we take equation (4) as a starting point. That is, we need to *assume* that the arrival process to each queue satisfies a sample path LDP with rate function given by (4).

We now specify the evolution of the scaled queue lengths $q_i(t)$ under LQF scheduling. Let $\mathcal{I}$ be any non-empty subset of $\{1, 2, \dots, N\}$. We define $\mathcal{R}_\mathcal{I}$ as the subset of $[0, 1]^N$, such that $(q_1(t), \dots, q_N(t)) \in \mathcal{R}_\mathcal{I}$ iff $q_i(t) = q_j(t) \ \forall i, j \in \mathcal{I}$, and for any $k \notin \mathcal{I}$, $j \in \mathcal{I}$, $q_k(t) < q_j(t)$. Intuitively, in region $\mathcal{R}_\mathcal{I}$, the queues in the index set $\mathcal{I}$ 'grow together', and all other queues are smaller. It is clear that the regions $\mathcal{R}_\mathcal{I}$ are convex, and constitute a partition of the set $[0, 1]^N$ as $\mathcal{I}$ ranges over all non-empty index sets. The queue evolution equation in region $\mathcal{R}_\mathcal{I}$ is given by

$$\begin{aligned} \sum_{i \in \mathcal{I}} \dot{q}_i &= \sum_{i \in \mathcal{I}} x_i(t) - 1; \\ \dot{q}_k &= x_k(t), \ \forall k \notin \mathcal{I} \end{aligned} \tag{5}$$

We now state the main result regarding the large deviations behavior of LQF scheduling.

Theorem 1: Under statistically independent and identical arrival processes to each queue, the LD behavior of LQF scheduling is given as follows

(i) The exponent is given by

$$E_N^{LQF} = \min_{k=1,\dots,N} k\Theta_k, \tag{6}$$

where Θ_k is the system occupancy exponent for k parallel queues, given by (3).

(ii) For a given λ, suppose that a unique $j \leq N$ minimizes (6), i.e.,

$$j = \arg \min_{k=1,\dots,N} k\Theta_k.$$

Then, for that λ, the likeliest overflow trajectory consists of j queues reaching overflow. More specifically, the likeliest overflow trajectory[5] (in the $(q_1(t), \dots, q_N(t))$ space) is the line segment joining the origin to the point $(q_1(0) = 1, \dots, q_j(0) = 1, q_{j+1}(0) = \frac{\lambda}{a_j^*}, \dots, q_N(0) = \frac{\lambda}{a_j^*})$, where $\frac{\lambda}{a_j^*} < 1$

The proof of the theorem follows a rather elaborate sample path large deviations argument that involves solving a variational problem. We relegate the proof to the appendix, and discuss the theorem intuitively.

The first part of the theorem states that the buffer overflow exponent under LQF scheduling is only a function of the system occupancy exponents Θ_k for $k \leq N$. The second part of the theorem assets that if E_N^{LQF} equals $j\Theta_j$ for a unique $j \leq N$, then the likeliest overflow scenario consists of j queues overflowing, and the other $N - j$ queues grow approximately to $M\frac{\lambda}{a_j^*}$, which is less than M. In particular, the queues that do not overflow are never the longest, and hence get no service at all. The service is shared equally among the j queues that overflow, and a_j^* denotes the likeliest *rate* at which the j queues overflow in spite of getting all the service. On the other hand, the queues that do not overflow get to keep all their arrivals, which occur at the average rate λ. The exponent for this case is given by $j\Theta_j$, which corresponds to all the queues in a j-queue system overflowing together. This is because the other $N - j$ queues which do not get service, get arrivals at the average rate, and hence do not contribute to the exponent.

A. Illustrative Examples with Bernoulli Traffic

In this section, we obtain the LQF exponents explicitly for a system with symmetric Bernoulli inputs to each queue. We deal with $N = 2$ and $N = 3$, since these cases are easily visualized, and elucidate the nature of the solution particularly well. We begin by making the following elementary observation regarding LQF scheduling and Bernoulli arrivals

Proposition 2: Under Bernoulli arrivals and LQF scheduling, the system evolves such that the two longest queues never differ by more than two packets.

Next, we state a well known result regarding the rate function $\Lambda^*(\cdot)$ for a Bernoulli process.

Proposition 3: For a Bernoulli process of rate λ, the rate function is given by

$$\Lambda^*(x) = D(x||\lambda) := x \log \frac{x}{\lambda} + (1 - x) \log \frac{1 - x}{1 - \lambda},$$

where $D(x||\lambda)$ is the Kullback-Liebler (KL) divergence (or the relative entropy) between x and λ.

[4]We suppress the dependence on M for simplicity of notation.

[5]The symmetry allows us to only present the case where the first j queues overflow.

The result is a consequence of Sanov's theorem for finite alphabet [10].

Let us now consider a two queue system with Bernoulli arrivals. For this simple system, it turns out that the exponent can be computed from first principles, without resorting to sample path large deviations. First, the system exponent Θ_2 under Bernoulli arrivals can be computed directly from the system occupancy Markov chain, yielding

$$\Theta_2 = 2 \ln \frac{1-\lambda}{\lambda}.$$

Proposition 4: Under LQF scheduling and Bernoulli arrivals, the following statements hold for the case $N = 2$:

(i) The likeliest overflow trajectory is along the *diagonal,* $(q_1 = q_2)$

(ii) $E_2^{LQF} = 2\Theta_2 = 4 \ln \frac{1-\lambda}{\lambda}$.

Proof: Part (i) of the result is a simple consequence of proposition 2. Specifically, suppose that one of the queues (say Q_1) overflows, so that $Q_1 \geq M$. From proposition 2, it follows that $Q_2 \geq M - 2$. Thus, when an overflow occurs in one queue, the other queue is also about to overflow, so that the only possible (and thus the likeliest) overflow trajectory is along the diagonal.

In order to show part (ii), we first argue that $E_2^{LQF} \geq 2\Theta_2$. Indeed, when a buffer overflow occurs, the total system occupancy is at least $2M - 2$. Thus, the buffer overflow probability is upper-bounded by the probability of the total system occupancy being at least $2M - 2$:

$$\mathbb{P}\{Q_1 \geq M\} \leq \mathbb{P}\{Q_1 + Q_2 \geq 2M - 2\}.$$

We thus have,

$$\begin{aligned} E_2^{LQF} &= \lim_{M \to \infty} -\frac{1}{M} \ln \mathbb{P}\{Q_1 \geq M\} \\ &\geq \lim_{M \to \infty} -\frac{1}{M} \ln \mathbb{P}\{Q_1 + Q_2 \geq 2M - 2\} = 2\Theta_2, \end{aligned}$$

where the last equality follows from the definition of Θ_2. To show a matching upper bound, note that when the system occupancy is $2M$ or greater, at least one of the queues will necessarily overflow. We can then argue as above that $E_2^{LQF} \leq 2\Theta_2$. □

Let us now analyze a system with three queues, fed by symmetric Bernoulli traffic. In this case, although the longest two queues grow together, it is not immediately clear how the third queue behaves during overflow. Once the system exponent Θ_3 is computed from (3), we can invoke the theorem 1, and conclude that the exponent is given by $\min(2\Theta_2, 3\Theta_3)$. Note that Θ_1 is infinite in this case, since a single queue fed by Bernoulli input cannot overflow. Fig. 2 shows a plot of $2\Theta_2$ and $3\Theta_3$ as functions of the input rate λ on each queue. It is clear from the figure that for small values of λ, the exponent $2\Theta_2$ dominates the overflow behavior. In this regime, the likeliest manner of overflow involves two queues reaching overflow, while the third queue grows to approximately $M\frac{\lambda}{1/2-\lambda}$. For larger values of λ (> 0.07), the exponent is $3\Theta_3$, and all three queues overflow together.

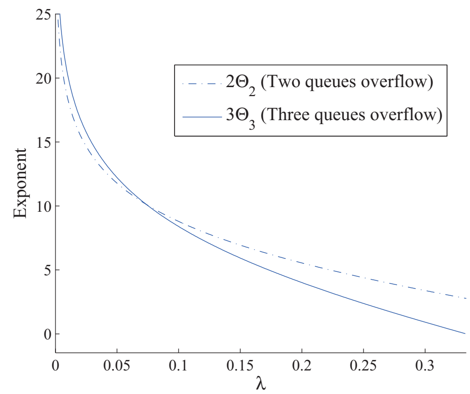


Fig. 2. Exponent behavior for $N = 3$ under Bernoulli traffic.

IV. LQF VS. QUEUE BLIND POLICIES

In this section, we compare the performance of LQF scheduling with that of queue-blind policies. We only consider a two queue system, since the large deviation behavior of PS and RS is difficult to characterize for $N > 2$. The following result for processor sharing follows from [4].

Proposition 5: The buffer overflow exponent for a two queue system under PS is given by

$$E_2^{PS} = \inf_{a>0} \frac{1}{a} \left[\Lambda^*(a + \frac{1}{2}) + \Lambda^*(\frac{1}{2}) \right]. \tag{7}$$

The likeliest manner of overflow under processor sharing is as follows. Suppose it is the first queue that overflows. The second queue receives traffic at rate $1/2$, which is also its service rate. Thus, the second queue grows to at most $o(M)$. The first queue receives service at rate $1/2$ and input traffic at rate $a_{ps}^* + 1/2$, where a_{ps}^* optimizes (7). Thus, a_{ps}^* is the rate of overflow of the first queue.

Next, we present the exponent for random scheduling.

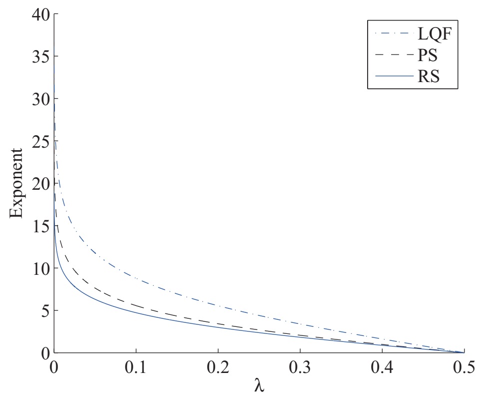
Proposition 6: The buffer overflow exponent for a two queue system under RS is given by

$$E_2^{RS} = \inf_{a>0} \frac{1}{a} \inf_{\phi \in (0,1)} \left[\Lambda^*(a + 1 - \phi) + \Lambda^*(\phi) + D(\phi || \frac{1}{2}) \right]. \tag{8}$$

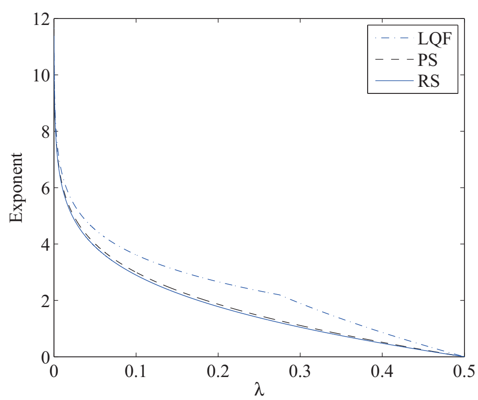
The proof is outlined in the appendix. We now describe the most likely overflow event. Suppose queue 1 overflows. The parameter ϕ that appears in the inner infimization in (8) denotes the empirical fraction of service received by queue 2. In other words, the 'fair' coin tosses that decide which queue to serve when both queues are nonempty, 'misbehave' statistically. The exponent corresponding to this event is given by $D(\phi||\frac{1}{2})$. If ϕ^* is the optimal value of ϕ in (8), the second queue receives traffic at rate ϕ^*, and therefore grows to an $o(M)$ level. The first queue receives traffic at rate $a_{rs}^* + 1 - \phi^*$, where a_{rs}^* is the optimizing value of a in (8).

Proposition 7: It holds that $E_2^{RS} \leq E_2^{PS} \leq E_2^{LQF}$.

Proof: To see the first inequality $E_2^{RS} \leq E_2^{PS}$, note that substituting $\phi = 1/2$ in the RS exponent (8) yields the PS

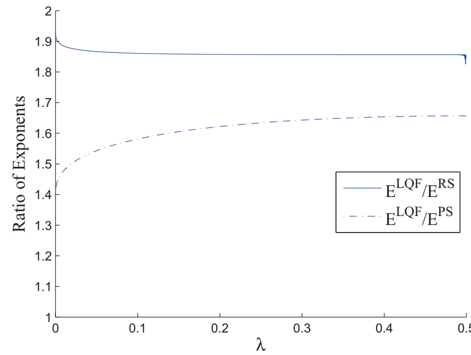


(a)

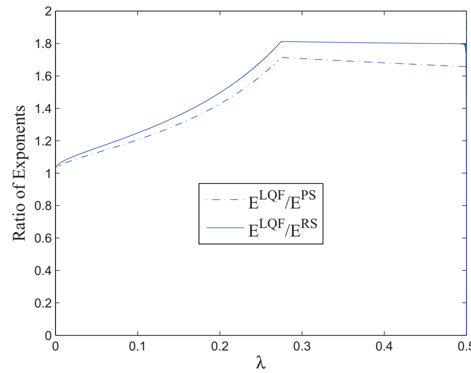


(b)

Fig. 3. Comparison of LQF, PS and RS exponents for a two queue system, under (a) Bernoulli arrivals (b) Poisson arrivals



(a)



(b)

Fig. 4. Ratio of LQF exponent to PS and RS exponents for (a) Bernoulli arrivals (b) Poisson arrivals.

exponent. To prove the second inequality, it suffices to show that $E_2^{PS} \leq \Theta_1$ and $E_2^{PS} \leq 2\Theta_2$. First note that for all $a \geq 0$, we have $\Lambda^*(a+1/2) \geq \Lambda^*(1/2)$ since the input rate λ is less than $1/2$. Thus, for all $a \geq 0$,

$$\frac{2}{a}\Lambda^*(a+1/2) \geq \frac{1}{a}[\Lambda^*(a+1/2) + \Lambda^*(1/2)].$$

Taking inf on both sides, we have $E_2^{PS} \leq 2\Theta_2$. Similarly, for all $a > 0$, it can be shown that $\Lambda^*(a+1) \geq \Lambda^*(a+1/2) + \Lambda^*(1/2)$, using the fact that $\Lambda^*(\dot{})$ is an increasing convex function, for arguments greater than λ. Dividing the preceding inequality by a and taking infimum, it follows that $E_2^{PS} \leq \Theta_1$. $\square$

In Fig. 3, we plot the exponents corresponding to LQF, PS and random scheduling for a two queue system, as a function of the arrivals rate λ. Fig. 3(a) corresponds to having Bernoulli arrivals in each time slot, while in Fig. 3(b), the number of arrivals in each slot is a Poisson random variable. The first observation we make from Fig. 3 is that, for a given arrival rate, the exponent values for a given policy are generally larger under Bernoulli traffic. This is because Poisson arrivals have a larger potential for being more bursty, and hence the overflow probability is larger (and the exponent smaller) for a given average rate. Next, notice that the LQF exponent under Poisson traffic (Fig. 3(b)) exhibits a cusp at $\lambda \approx 0.27$. This is because under Poisson traffic, we have two competing exponents Θ_1 and $2\Theta_2$, corresponding respectively to one queue and both the queues overflowing. For λ below the cusp, Θ_1 dominates, and vice-versa. On the other hand, under Bernoulli traffic, Θ_1 is infinite. Thus, the LQF exponent is given by $2\Theta_2$, which is a smooth curve as shown in Fig. 3(a).

In Fig. 4, we plot the ratio of the LQF exponent to the PS and RS exponents. This ratio is directly related to the savings in the buffer size that results from using LQF scheduling, as opposed to using one of the queue blind policies. For example, consider the ratio of the LQF exponent to the RS exponent, when the traffic is relatively heavy (say $\lambda > 0.3$). This is the regime where overflows are most likely to occur. We see that under both Bernoulli and Poisson traffic, the LQF exponent is roughly 1.8 times the RS exponent. This implies that in order to achieve a certain overflow probability, the LQF policy requires only 55% of the buffer size required under random scheduling in heavy traffic. A similar comparison can also be made between the LQF and PS exponents.

V. Scheduling with Infrequent Queue Length Information

We have seen that the LQF policy has a superior queue overflow performance compared to queue blind policies. This is because it can discern and mitigate large queue build-up on one of the queues. On the other hand, the scheduler needs to know queue length information in every slot in order to perform LQF scheduling. In this section, we will show that the buffer overflow performance of LQF scheduling can be maintained even if we allow for arbitrarily infrequent queue length information to be conveyed to the scheduler.

The basic idea is that it is sufficient to serve the longest queue only when the queues are large. When the queue lengths are all small, we can save on the queue length information by adopting a work conserving, but queue-blind scheduling strategy. To achieve this, we suggest the following scheduling policy which is a 'hybridized' version of the queue-blind RS, and the LQF policy.

Hybrid Scheduling: Let $K < M$ be a given queue length threshold. In each slot, if all queues are smaller than K, then serve any random occupied queue. If at least one queue exceeds K, serve the longest queue in that slot.

The following theorem asserts that the hybrid policy asymptotically achieves the same buffer overflow exponent as LQF scheduling, while requiring queue length information in a vanishingly small fraction of slots.

Theorem 2: Suppose K increases sub-linearly in the buffer size M (i.e., $K(M) = o(M)$). Then,

(i) The buffer overflow exponent of hybrid scheduling is equal to E_N^{LQF}, and
(ii) The fraction of slots in which queue length information is required approaches zero if $K(M) \to \infty$ as $M \to \infty$.

We provide a heuristic explanation of the result due to space constraints. Observe that queue length information is required only in time slots when the longest queue in the system is longer than K. Since RS is a stabilizing policy, the steady state probability that the longest queue exceeds K approaches zero as K becomes large. (In fact, this probability goes to zero *exponentially* in K.) Therefore, the fraction of slots in which queue length information is required can be made arbitrarily small. On the other hand, the overflow exponent remains the same as in the LQF case. This is because when we consider the scaled queue lengths as M becomes large, the hybrid policy differs from LQF scheduling only in an infinitesimal neighborhood around the origin.

VI. Acknowledgements

The authors would like to thank Prof. John Tsitsiklis for his valuable comments and suggestions.

Appendix

A. Proof of Theorem 1

The proof can be divided into two parts. The first part involves showing that the queue length process under LQF scheduling satisfies an LDP, whose rate function is given by the solution to a variational problem. The second step involves solving the variational problem in the case of symmetric arrivals, and proving that the optimal solution to the variational problem takes a simple form, as given by the theorem.

The existence of an LDP for the queue length was shown in [8] for longest weighted waiting time as well as longest weighted queue length scheduling. Assuming without loss of generality that the first queue overflows, the exponent is given by the following variational problem

$$\min \int_{-T}^{0} \left[\sum_{i=1}^{N} \Lambda^*(x_i(t)) \right] \mathrm{d}t \tag{9}$$

subject to

$$q_i(-T) = 0, \forall i$$

$$q_1(0) = 1,$$

$$T: \text{ free},$$

$$q_j(0): \text{ free for } j > 1,$$

and the queue length trajectories $q_i(t)$ evolve according to (5).

Our emphasis is on solving the above variational problem under the symmetric traffic scenario. In (9), the empirical rates $x_i(t)$ are the control variables, and the cost function is the exponent corresponding to the control variables, as given by Mogulskii's theorem. In words, the variational problem is to find the set of empirical rates which leads to the smallest exponent, and results in the overflow of at least one queue. Note that the above is a free time problem, i.e., the time T over which overflow occurs is not constrained. Also, it is possible for queues other than the first queue to reach overflow.

An important property which helps us solve (9) is given by the following lemma, which states that when the scaled queue lengths are within one of the regions $\mathcal{R}_{\mathcal{I}}$, the empirical rates $x_i(t)$ can be taken as constants, without loss of optimality.

Lemma 2: Fix a time interval $[-T_1, -T_2]$ and consider a control trajectory $x_i(t),\ i = 1, \ldots, N,\ t \in [-T_1, -T_2]$, such that the scaled queue lengths $q_i(t),\ i = 1, \ldots, N,\ t \in [-T_1, -T_2]$ stay within a particular region $\mathcal{R}_{\mathcal{I}}$. Define the average control trajectory $\bar{x}_i$ in the interval $[-T_1, -T_2]$ as

$$\bar{x}_i(\tau) = \frac{1}{T_1 - T_2} \int_{-T_1}^{-T_2} x_i(t) \mathrm{d}t$$

for $i = 1, \ldots, N$ and $\tau \in [-T_1, -T_2]$. Then, the queue lengths under the average control trajectory $\bar{x}_i(t)$ lie entirely within $\mathcal{R}_{\mathcal{I}}$, and satisfy the same initial and final conditions at $t = -T_1$ and $t = -T_2$ respectively. Furthermore, the cost achieved under the (constant) control trajectory $\bar{x}_i(t)$ is not larger than the cost achieved under $x_i(t)$.

The above result follows from the convexity of $\Lambda^*(\cdot)$ and of the sets $\mathcal{R}_{\mathcal{I}}$, and the proof is akin to the two dimensional case treated in [4, Lemma 5.1]. Using Lemma 2, we next compute the exponents corresponding to overflow trajectories that stay entirely within a particular region $\mathcal{R}_{\mathcal{I}}$. Later, we will show that overflow trajectories that traverse more than one region cannot have a strictly smaller exponent than trajectories that stay within exactly one of the regions. This will give us the result we want.

Consider an overflow trajectory that lies entirely within $\mathcal{R}_{\mathcal{I}_j}$, where $|\mathcal{I}_j| = j$ for some $1 \leq j < N$. In this case, the j queues in the index set $\mathcal{I}_j$ reach overflow, while the other $N - j$ queues are strictly smaller, and hence receive no service. Due to the symmetry of arrivals, we can compute the exponent assuming that $\mathcal{I}_j = \{1, \ldots, j\}$, i.e., the first j queues overflow. Lemma 2 implies that the optimal empirical rates can be restricted to constant values[6] $x_i,\ i = 1, \ldots, N$ for this particular overflow event. Let $a = 1/T$ denote the rate at which the first j queues overflow. Since each queue $k \in \{1, \ldots, j\}$ overflows at rate a, the empirical input rate x_k must be of the form $x_k = a + \phi_k$, where $\phi_k \geq 0$ can be thought of as the rate at which queue k receives service in the overflow interval. Since the first j queues receive all the service, we have $\sum_{k=1}^{j} \phi_k = 1$. Next, for $l > j$, we need $x_l \leq a$, since these queues are never the longest, and hence get no service.

The optimization in (9) takes the following form when the first j queues reach overflow.

$$\inf_{a>0} \frac{1}{a} \inf_{\substack{\phi_k \geq 0; \sum_{k=1}^{j} \phi_k = 1 \\ x_l \leq a,\ \forall\ l>j}} \sum_{k=1}^{j} \Lambda^*(a + \phi_k) + \sum_{l=j+1}^{N} \Lambda^*(x_k) \quad (10)$$

Let us now perform the inner minimization in (10). It is obvious that the minimization over $\phi_k,\ k \leq j$ and $x_l,\ l > j$ can be performed independently. Due to convexity of the rate function, we have

$$\frac{1}{j} \sum_{k=1}^{j} \Lambda^*(a + \phi_k) \geq \Lambda^*(\frac{1}{j} \sum_{k=1}^{j} (a + \phi_k)) = \Lambda^*(a + \frac{1}{j}).$$

Therefore, the optimal value of the ϕ_ks is given by $\phi_k = 1/j,\ k \leq j$. Next, consider optimizing over x_l for $l > j$. We distinguish two cases:

(i) $a > \lambda$: In this case, it is optimal to choose $x_l = \lambda$ for each $l > j$, since $\Lambda^*(\lambda) = 0$.

(ii) $a \leq \lambda$: In this case, the constraint $x_l \leq a$ has to be active, since for $x < \lambda$, $\Lambda^*(x)$ is decreasing in x. Thus, we have $x_l = a$.

Putting the two cases together, we get from (10) the exponent E_j corresponding to exactly j queues overflowing, while the trajectory stays inside $\mathcal{R}_{\mathcal{I}_j}$.

$$E_j = \min(\chi_j, \xi_j) \quad (11)$$

with

$$\chi_j = \inf_{0<a\leq\lambda} \frac{1}{a} \left[j\Lambda^*(a + \frac{1}{j}) + (N - j)\Lambda^*(a) \right], \text{ and}$$

$$\xi_j = \inf_{a>\lambda} \frac{j}{a} \Lambda^*(a + \frac{1}{j}). \quad (12)$$

[6]For simplicity of notation, we henceforth use x_i in place of $\bar{x}_i$.

The above expression holds for $1 \leq j < N$. The exponent for all the N queues overflowing is simpler to obtain; it is given by

$$E_N = \inf_{a>0} \frac{N}{a} \Lambda^*(a + \frac{1}{N}) = N\Theta_N, \quad (13)$$

where the last equality follows by recalling (3). The optimal exponent considering the set of all overflow trajectories that stay inside *any one* of the regions $\mathcal{R}_{\mathcal{I}},\ \mathcal{I} \subset \{1, \ldots, N\}$ is obtained by minimizing E_j over $j = 1, \ldots, N$.

At this point, we are two steps away from obtaining the result. The first step involves showing that there is nothing further to be gained by considering paths that traverse more than one of the partitioning regions. This would imply that the optimal exponent is given by $\min_{1\leq j\leq N} E_j$. The second step involves showing that $\min_{1\leq j\leq N} E_j = \min_{1\leq j\leq N} j\Theta_j$, where Θ_j is the system occupancy exponent of j parallel queues, defined in (3). The following two lemmas establish what is needed.

Lemma 3: For every queue overflow trajectory that traverses more than one of the regions $\mathcal{R}_{\mathcal{I}},\ \mathcal{I} \subset \{1, \ldots, N\}$, there exists an overflow trajectory that lies entirely within one of the regions, while achieving an exponent that is no larger.

Proof: We only rule out overflow trajectories that traverse two regions; similar arguments can be used for trajectories that visit more than two regions. Consider a queue trajectory that starts out in a region $\mathcal{R}_{\mathcal{I}}$ but reaches overflow in region $\mathcal{R}_{\mathcal{J}}$, while staying in one of the two regions at every instant in between. Note that the region $\mathcal{R}_{\mathcal{I}}$ is a convex set of dimension $N - |\mathcal{I}| + 1$. That is, regions that involve a larger number of queues growing together, have a smaller dimension and vice-versa.

We will consider two cases, $\mathcal{I} \supset \mathcal{J}$ and $\mathcal{I} \subset \mathcal{J}$. Brief reflection should make it clear that if one of the above two containments is not satisfied, the trajectory has to necessarily traverse more than two regions. The arguments that follow are easier to understand if visualized in two dimensions.

Suppose $\mathcal{I} \subset \mathcal{J}$. Consider a trajectory that starts out at the origin at $t = -T$, and stays inside $\mathcal{R}_{\mathcal{I}}$ until time $t = -T_1$, when it enters $\mathcal{R}_{\mathcal{J}}$. The trajectory stays in $\mathcal{R}_{\mathcal{J}}$ until overflow at $t = 0$. Intuitively, the queues $q_i,\ i \in \mathcal{I}$ start out growing together. At time $-T_1$, the queues $q_i,\ i \in \mathcal{J} - \mathcal{I}$ 'catch up', and overflow occurs in all the queues in the index set $\mathcal{J}$. Since constant empirical input rates are optimal inside each partition region (Lemma 2), the arbitrary trajectory in $\mathcal{R}_{\mathcal{I}}$ can be replaced at no further cost by a straight segment that has the same initial and final values ($q_i(-T) = 0$, and $q_i(-T_1) \in \mathcal{R}_{\mathcal{J}}$ for each i). This segment lies entirely in $\mathcal{R}_{\mathcal{I}}$, but is arbitrarily close to the region $\mathcal{R}_{\mathcal{J}}$. (Note that $\mathcal{R}_{\mathcal{J}}$ forms one of the 'boundaries' of $\mathcal{R}_{\mathcal{I}}$). However, the cost of this replaced segment is clearly not lower than the optimal trajectory in $\mathcal{R}_{\mathcal{J}}$ with the same initial and final conditions. The part of the trajectory from $t = -T_1$ until over flow at $t = 0$, can again be replaced by the optimal trajectory in $\mathcal{R}_{\mathcal{J}}$ with the corresponding end points. Thus, overall, the cost of the original trajectory is greater than or equal to that of the optimal trajectory in $\mathcal{R}_{\mathcal{J}}$.

Now consider the case $\mathcal{I} \supset \mathcal{J}$. Intuitively, this case corresponds to the queues $q_i,\ i \in \mathcal{I}$ starting to grow together. At some time instant, the queues $q_i,\ i \in \mathcal{I} - \mathcal{J}$ start 'losing out', and overflow occurs within $\mathcal{R}_{\mathcal{J}}$. The arbitrary trajectories in each of the regions can be replaced with an optimal segment in each of the regions, with the same boundary conditions at no added cost. The cost of this replaced trajectory, is a convex combination of the optimal overflow trajectories in regions $\mathcal{R}_{\mathcal{J}}$ and $\mathcal{R}_{\mathcal{I}}$, and hence cannot be smaller than the smaller of the two costs. Thus, a strictly smaller cost cannot be obtained by a trajectory that traverses two regions. □

Lemma 4: $\min_{1\leq j\leq N} E_j = \min_{1\leq j\leq N} j\Theta_j$.

Proof: We first prove that $\chi_j \geq E_N$ for all $j < N$. First, using convexity, we can write

$$\begin{aligned}\frac{j}{N}\Lambda^*(a+\frac{1}{j}) + \frac{N-j}{N}\Lambda^*(a) \quad &\geq \quad \Lambda^*\left(\frac{j}{N}(a+\frac{1}{j}) + \frac{N-j}{N}a\right)\\ &= \quad \Lambda^*(a+\frac{1}{N}). \end{aligned} \tag{14}$$

We now have

$$\begin{aligned}\chi_j \quad &= \quad \inf_{0<a\leq\lambda} \frac{1}{a}\left[j\Lambda^*(a+\frac{1}{j}) + (N-j)\Lambda^*(a)\right]\\ &\geq \quad \inf_{a>0} \frac{1}{a}\left[j\Lambda^*(a+\frac{1}{j}) + (N-j)\Lambda^*(a)\right]\\ &\overset{(a)}{\geq} \quad \inf_{a>0} \frac{N}{a}\Lambda^*(a+\frac{1}{N}) = E_N.\end{aligned}$$

The inequality (a) follows from (14). It is now clear that the χ_js are irrelevant, as they are always dominated by $E_N = N\Theta_N$. We next write the following series of equalities that imply the lemma.

$$\begin{aligned}\min_{1\leq j\leq N} E_j \quad &= \quad \min(\xi_1, \ldots, \xi_{N-1}, N\Theta_N)\\ &= \quad \min_{1\leq j<N} \min(\xi_j, N\Theta_N)\\ &\overset{(b)}{=} \quad \min_{1\leq j<N} \min(j\Theta_j, N\Theta_N)\\ &= \quad \min_{1\leq j\leq N} j\Theta_j\end{aligned}$$

In the above, equality (b) is shown as follows. Consider $\min(\xi_j, N\Theta_N)$. The definition of ξ_j (12) involves the infimum of a convex function of a over $a > \lambda$. If the convex function attains its *global* minimum for $0 < a < \lambda$, then the infimum in (12) will be obtained at $a = \lambda$. In this case, it is easy to show that $N\Theta_N \leq \xi_j$. Thus, if ξ_j has to be smaller than $N\Theta_N$, the infimum in (12) must be obtained at the global minimum, which lies at $a > \lambda$.[7] Thus, whenever $\min(\xi_j, N\Theta_N) = \xi_j$, we necessarily have

$$\xi_j = \inf_{a>\lambda} \frac{j}{a}\Lambda^*(a+\frac{1}{j}) = \inf_{a>0} \frac{j}{a}\Lambda^*(a+\frac{1}{j}) = j\Theta_j,$$

so that equality (b) follows, and we are done. □

[7]It follows that $\lambda/a_j^* < 1$, which proves the claim made in part (ii) of theorem 1.

B. Proof Outline of Proposition 6

Let $B_i[t] \in \{0, 1\}$ denote the i.i.d fair 'coin tosses' that decide which queue to serve when both the queues are occupied. If $B_i[t] = 1$, then the second queue is served if occupied in slot t; if $B_i[t] = 0$, the first queue is served if occupied. If one of the queues is not occupied in slot t, the occupied queue is served, and $B_i[t]$ becomes irrelevant. Let $\phi(t)$ be the empirical fraction of coin tosses in favor of the second queue, defined analogously to the empirical input rates in Section III. The dynamics of the scaled queue length processes under RS is given by

$$\begin{aligned}\dot{q}_1(t) \quad &= \quad x_1(t) - (1-\phi(t))\\ \dot{q}_2(t) \quad &= \quad x_2(t) - \phi(t),\end{aligned}$$

whenever $q_1(t)$ and $q_2(t)$ are non-zero. If either $q_1(t) = 0$ or $q_2(t) = 0$, then

$$\dot{q}_1(t) + \dot{q}_2(t) = x_1(t) + x_2(t) - 1.$$

Here, $x_1(t)$ and $x_2(t)$ are the empirical rates of the input processes.

Using a result analogous to Lemma 2, we can prove that constant empirical rates for the inputs as well as the coin tosses is optimal, within each of the regions (i) $q_1(t) > 0,\ q_2(t) > 0$ (ii) $q_1(t) > 0,\ q_2(t) = 0$, and (iii) $q_1(t) = 0,\ q_2(t) > 0$. The problem can now be mapped to an instance of generalized processor sharing with variable service rate, as treated in [4]. The result follows by applying the GPS exponent results to our symmetric case, and noting that the rate function corresponding to the fair coin tosses is given by $D(\cdot||1/2)$.

References

[1] L. Tassiulas, A. Ephremides, *Dynamic server allocation to parallel queues with randomly varying connectivity.* IEEE Transactions on Information Theory, Vol. 39, No. 2, pp. 466-478, March 1993.

[2] M. J. Neely, E. Modiano, C. E. Rohrs, *Power and server allocation in a multi-beam satellite with time varying channels,* IEEE Infocom, New York, NY, June 2002.

[3] M. J. Neely, *Delay analysis for max-weight opportunistic scheduling in wireless systems,* 46th Allerton Conference, Monticello, IL 2008.

[4] D. Bertsimas, I. C. Paschalidis, and J. N. Tsitsiklis, *Asymptotic Buffer Overflow Probabilities in Multiclass Multiplexers: An Optimal Control Approach*, IEEE Transactions on Automatic Control. Vol. 43, No. 3, March 1998, pp. 315-335.

[5] S. Shakottai, *Effective Capacity and QoS for Wireless Scheduling,* IEEE Transactions on Automatic Control, Volume 53, Issue 3, April 2008 pp.749 - 761.

[6] L. Ying, R. Srikant, A. Eryilmaz and G. Dullerud. *A Large Deviations Analysis of Scheduling in Wireless Networks.* IEEE Transactions on Information Theory, Vol. 52, No. 11, November 2006.

[7] V. J. Venkataramanan and X. Lin, *On Wireless Scheduling Algorithms for Minimizing the Queue-Overflow Probability,* submitted to IEEE/ACM Transactions on Networking, 2008.

[8] A. L. Stolyar and K. Ramanan *Largest Weighted Delay First Scheduling: Large Deviations and Optimality,* The Annals of Applied Probability 2001, Vol. 11, No. 1, pp. 148.

[9] V. G. Subramanian, *Large Deviations Of Max-Weight Scheduling On Convex Rate Regions* Proceedings of ITA 2008, UCSD, La Jolla, CA.

[10] A. Ganesh, N. O'Connel, D. Wischik, *Big Queues*, Springer-Verlag, 2004.

[11] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Second Ed., Springer Verlag, 1998, pp. 176.