1988

On the Maximum Queue Length with Applications to Data Structures: A Simple But Yet Asymptotically Exact Approach

Wojciech Szpankowski
Purdue University, spa@cs.purdue.edu

Report Number:
88-785

ON THE MAXIMUM QUEUE LENGTH WITH APPLICATIONS TO DATA STRUCTURES:
A Simple But Yet Asymptotically Exact Approach

Wojciech Szpankowski

CSD-TR-785
June 1988

ON THE MAXIMUM QUEUE LENGTH WITH APPLICATIONS TO DATA STRUCTURES: A Simple But Yet Asymptotically Exact Approach

*Wojciech Szpankowski**

Department of Computer Science
Purdue University
West Lafayette, IN 47907, USA

Abstract

A dynamic data structure called *queue* is analyzed in this paper from the viewpoint of its maximum size. By dynamic queue we understand any data structure that is built during a sequence of insertions and deletions. The maximum size of such a structure is a fundamental quantity and is directly related to many problems of resource allocations. We assume that each element of the structure (we call it further a *customer)* stays for a random time in the system and then leaves it. Furthermore, the interarrival time of customers is a generally distributed random variable. Adopting queueing theory language, we say the data structure is G|G|c queueing systems, where c (may be infinite) is the maximum number of items that can simultaneously leave the system (number of servers). We shall show that for *stable stationary* queue the maximum queue length observed by the n-th arriving customer grows asymptotically *in probability* as $\log_\alpha n$, where α is a system constant.

* This research was supported in part by NSF under grant NCR-8702115

1. MOTIVATIONS AND INTRODUCTION

Queues and other data structures [AHU] behaviors are natural models for many dynamic phenomena. They correspond to important processes in the areas of algorithms, operating systems, distributed systems, computer networks, etc. Queues are understood here as a dynamic data structure with random insertions and deletions, that is, an item may arrive at any random moment of time and may stay in the structure for a random time (service time) until it is deleted (served). We do not impose any special restrictions on the distributions of the interarrival times of items and service times. In queueing theory terminology, we consider GI|G|c systems [ASM, KLE] where c (may be infinite) stands for the number of servers, and GI/G means respectively that the interarrival and service times are generally distributed.

The dynamics of queues can be studied through the transient analysis, however, this seems to be hopeless in our general setting [cf. ASM]. Nevertheless, some important information about dynamics of the structure can be obtained by analyzing the maximum size of the queue over a period of time. Such information, without any doubt, has obvious significance to issues of resource allocation (e.g., the design of a buffer size in a distributed system). The maximum queue length was extensively studied in the seventieth by queueing theoreticians. Heyde [HEY] was the first who predicted the asymptotic growth of maximum queue length in the GI|M|1 system. His result was extended by Cohen in [COH] to M|G|1 systems, and finally Iglehart [IGL] completed these studies by providing the growth of rate for GI|G|1 queue. Unfortunately, all the works but Heyde are rather limited to queueing theory and the methodology is too complicated, with no clear extensions to other dynamic structures. In this paper, we provide a new simple methodology to study the maximum queue length in GI|G|1 and PH|PH|c [TAK] queueing systems. The bad news from this analysis (cf. [HEY, COH, IGL]) is that the maximum queue length observed by the n-th arriving customer grows asymptotically *in probability* like $\log_\alpha n$, where α is a system parameter depending upon the interarrival and service distribution functions. In addition, the proposed methods is robust in that it applies to several different models, including discrete-time priority queue [AHU], hashing with lazy deletion [MSW, WV], geometric adjacency problem that arose in the analysis of VLSI [SW], performance evaluation of digital trees [SZ1], graph optimization problem [SZ2], etc.

This paper is organized as follows. In the last section, after presenting some preliminary results form queueing theory, we state our main results. In Section 3 we apply these results to study hashing with lazy deletion.

2. MAIN RESULTS

In this section we establish our main results. We begin with some preliminary definitions from queueing theory, then we formulate our proposition, and finally we discuss some consequences of our main results. The proof is presented in the last two subsections, each one dealing with upper and lower bounds on the maximum queue length.

2.1 Preliminary Results

We analyze a single server queueing system with arbitrary interarrival times and service times, that is, GI|G|1 model. Let $A(t)$ and $B(t)$ respectively, represent distribution functions of the interarrival times and service times. We denote by $A^*(s)$ and $B^*(s)$ the corresponding Laplace-Stieltjes transforms of $A(t)$ and $B(t)$. The interarrival times, as well as the service times, are mutually independent, that is, both processes form a renewal process. Two quantities are of particular interest, namely the queue length Q_k and waiting time W_k at the moment of the k-th arrival of a customer. Our purpose is to estimate

$$\overline{Q}_n = \max_{1 \le k \le n} Q_k \tag{2.1a}$$

$$\overline{W}_n = \max_{1 \le k \le n} W_k \tag{2.1b}$$

as n tends to infinity. More precisely, we establish asymptotic growth of the r-th moments $E\overline{Q}_n^r$

and $E\overline{W}_n^r$ of the maximum queue length and waiting time, as well as convergence of these quantities in probability.

It turns out that the asymptotics of $\overline{Q}_n$ and $\overline{W}_n$ depend on the tail of the stationary distributions of the queue length and the waiting time. In the further part of this paper, we assume that the system is *stable*, that is, $\rho = \lambda/\mu < 1$ with λ and μ being intensities of the arrival and service processes respectively. Under this condition, a stationary distribution exists [ASM, KLE], and we deal further only with the *stationary processes* Q_k and W_k. Then, we denote by $Q(m) = Pr\{Q_k \le m\}$ and $W(x) = Pr\{W_k < x\}$, respectively, the distribution functions of the queue length Q_k and the waiting time W_k. The following well known result of Feller [FEL] describes the tail distributions of these two processes. Let θ be a unique solution, if it exists, of the following complex equation

$$A^*(\theta)B^*(-\theta) = 1 \tag{2.2}$$

Then, Feller proves that for $x \to \infty$

$$1 - W(x) = c_1 e^{-x\theta}(1 + o(1)) \tag{2.3}$$

or in other notation $1 - W(x) \sim c_1 e^{-x\theta}$ for $x \to \infty$. As a simple consequence of (2.3), we obtain the tail distribution of the queue length. Indeed, define ω as

$$\omega = A^*(\theta) < 1 \tag{2.4}$$

then for $m \to \infty$

$$1 - Q(m) = c_2 \omega^m (1 + o(1)) \tag{2.5}$$

where c_1 and c_2 are constants.

The above results have been recently extended to some c-servers queueing systems, namely PH|PH|c where PH stands for phase distribution (see [TAK, ASM, KLE]). Takahashi proved that for FIFO PH|PH|c systems, the following holds

$$1 - W(x) \sim c_3 e^{-x\theta} \tag{2.6a}$$

$$1 - Q(m) - c_4 \omega^m \tag{2.6b}$$

where θ and ω are defined respectively as

$$A^*(c\theta) B^*(-\theta) = 1 \tag{2.7a}$$

$$\omega = A^*(c\theta) \tag{2.7b}$$

Using the above asymptotics we prove our main results.

PROPOSITION. For GI|G|1 and PH|PH|c queueing systems the following holds.

(i) The r-th moments $E\overline{Q}_n^r$ and $E\overline{W}_n^r$ of the maximum queue length and waiting time become

$$E\overline{Q}_n^r = \log_\omega^r n^{-1} (1 + o(1)) \tag{2.8a}$$

$$E\overline{W}_n^r = \log^r n^{1/\theta} (1 + o(1)) \tag{2.8b}$$

for large n.

(ii) The following convergences hold

$$\lim_{n \to \infty} \frac{\overline{Q}_n}{\log_\omega n^{-1}} = \lim_{n \to \infty} \frac{\overline{W}_n}{\log n^{1/\theta}} = 1 \tag{2.9}$$

in probability sense.

□

We prove the Proposition in the next two subsections, however we first note that Proposition (ii) is a simple consequence of Proposition (i) and Chebyshev's inequality. Indeed, by Proposition (i) the variance of $\overline{Q}_n$ and $\overline{W}_n$ are $var\, \overline{Q}_n = o(1) \log_\omega^2 n$ and $var\, \overline{W}_n = o(1) \log^2 n^{1/\theta}$, respectively. Then, by Chebyshev's inequality

$$Pr\left\{ \left| \frac{\overline{W}_n}{EW_n} - 1 \right| > \varepsilon \right\} \le \frac{var\, \overline{W}_n}{\varepsilon^2 (E\overline{W}_n)^2} = o(1) \tag{2.10}$$

The same is true for $\overline{Q}_n$. Hence (2.10) immediately implies (2.9), and in the next part of this section we only concentrate on proving Proposition (i). The plan for doing that is the following: in the next subsection we show how to establish an upper bound on the maximum of the waiting time , and then we prove a tight lower bound. To simplify our discussion, we shall deal mainly with the waiting time $\overline{W}_n$ knowing that the maximum queue length $\overline{Q}_n$ can be derived in a similar way.

2.2 Upper Bound

The upper bound on the maximum queue length and waiting time derive easily from the following result of Lai and Robbins [LR1, LR2], which is slightly generalized in the next lemma (see also [SZ1, BMS]).

Lemma 1. Let $Y_1, Y_2, \ldots, Y_n$ are identically distributed nonnegative random variables with distribution function $G(y)$. Denote $M_n = \max_{1 \le k \le n} Y_k$.

(i) The following bound holds for the r-th moment EM_n^r of M_n

$$EM_n^r \le a_n + n \int_{a_n}^{\infty} [1 - G(y^{1/r})]dy \tag{2.11}$$

where a_n is the smallest solution of

$$n[1 - G(a_n^{1/r})] = 1 \tag{2.12}$$

(ii) If $Y_1, Y_2, \ldots, Y_n$ are independent random variables satisfying, in addition, the following constraints

$$EY_1^p < \infty \quad \text{for some} \quad p \tag{2.13a}$$

$$G(y) < 1 \quad \text{for} \quad y < \infty \tag{2.13b}$$

$$\lim_{y \to \infty} \frac{1 - G(cy)}{1 - G(y)} = 0 \quad \text{for all} \quad c > 1 \tag{2.13c}$$

then

$$\lim_{n \to \infty} \frac{EM_n^r}{a_n^r} = 1 \tag{2.14}$$

that is, $EM_n^r = a_n^r(1 + o(1))$ for large n. In fact, a_n in (2.14) is merely a solution of (2.12) with $r=1$, that is,

$$n\,[\,1-G(a_n)\,]=1 \tag{2.12a}$$

Proof. To prove Lemma 1(i), we first note that $EM_n^r = E\max\{Y_1^r, Y_2^r, \ldots, Y_n^r\}$ and $Pr\{Y_1^r < y\} = Pr\{Y_1 < y^{1/r}\} = G(y^{1/r})$. Then, as in Lai and Robbins [LR1], one easily shows the following

$$M_n^r \le a + \sum_{k=1}^{n} [Y_k^r - a]^+ \tag{2.15}$$

where a is any number and $x^+ = \max\{0,x\}$. Taking the expectation of (2.15), we obtain

$$EM_n^r \le a + n\int_a^\infty [1 - G(y^{1/r})]dy, \tag{2.16}$$

and finally, minimizing the RHS of (2.16) with respect to a leads to a_n defined in (2.12). Part (ii) of the Lemma 1 is proved in [LR2].

□

To prove upper bounds on $\overline{W}_n = \max_{1\le k\le n} W_k$ and $\overline{Q}_n = \max_{1\le k\le n} Q_k$, we use Lemma 1(i) and our estimates (2.3) and (2.5) on the tail of the queue length and the waiting time distributions. Let us focus on the waiting time. For stable systems, i.e., $\rho < 1$, a stationary distribution of W_k exists, hence $\{W_k\}_{k=1}^n$ are identically distributed, hence Lemma 1(i) can be applied, and without loosing generality we can assume $r=1$. To compute a_n from (2.12a) and our tail estimate (2.3), we note that there exists $\varepsilon > 0$ tending to zero for $n \to \infty$ such that [BMS]

$$a_n = \frac{\log nc_1(1+\varepsilon)}{\theta} \tag{2.17a}$$

In the case of queue length a_n becomes

$$a_n = \frac{\log nc_2(1+\varepsilon)}{\log \omega^{-1}} \tag{2.17b}$$

Applying the above to (2.16), one immediately proves that the second part (i.e., the integral part) of the RHS of (2.16) is O(1). Therefore,

$$E\overline{W}_n^r \le a_n^r + O(1) \tag{2.18a}$$

$$E\overline{Q}_n^r \le a_n^r + O(1) \tag{2.18b}$$

where a_n is given in (2.17a) for waiting time, and in (2.17b) for the queue length.

2.3 Lower Bound

Here is the strategy to attack our problem. Let L_n denote the number of busy periods completed just prior to the arrival of the n-th customer. We shall prove that $L_n \sim n\alpha$ *almost surely* (a.s.) for some constant α. Let a random customer arriving *during the k-th busy period* sees U_k customers in the queue. Busy periods in GI|G|1 are i.i.d. random variables, hence $\{U_n\}_{k=1}^{L_n}$ is i.i.d. sequence of random variables. Naturally

$$\overline{Q}_n \ge \max_{1 \le k \le L_n} U_k \tag{2.19}$$

where '$\ge$' means "stochastically greater" [STO]. Let $M_{L_n} = \max_{1 \le k \le L_n} \{U_k\}$. In particular, (2.19) implies that $E\overline{Q}_n^r \ge EM_{L_n}^r$ for any $r > 0$. To compute EM_{L_n} we first prove $L_n \sim n\alpha$ *a.s.* which directly implies that $EM_{L_n} \sim EM_{n\alpha}$. Since U_k are i.i.d. and they admit the same tail behavior as Q_k (see (2.5)), we shall use Lemma 1(ii) to show that $EM_{n\alpha} \sim \log_\omega(n\alpha)^{-1}$. This, and our upper bound (2.18) derived above, will imply Proposition 1(i) formula (2.8a), and simple manipulation will lead to (2.8b).

The plan just described needs to be accompanied with a proof of some technicalities which follow. We first show that

Lemma 2. For large n and $\rho < 1$, $L_n \sim n\alpha$ *almost surely,* where α^{-1} is the average number of customers served in a busy periods.

Proof. The proof follows, with some slight changes, from the proof of Lemma 1 in Heyde [HEY]. For the completeness, we provide a sketch of the proof. Let D_n be the number of customer servers in the first n busy periods. Then $D_n = \sum_{i-1}^{n} V_i$ where V_i is the number of customers served in the i-th busy period. Since V_i are i.i.d. [ASM], and letting $\alpha^{-1} = EV_i > 0$, then by strong law of large numbers, $W_n/n \sim \alpha^{-1}$ *a.s.* But, with details found in [GAL, BAN], we note that $D_{L_n}/L_n \sim n/L_n \sim \alpha$ *a.s.*, hence $L_n \sim n\alpha$ *a.s.*

□

The next step is to show that $EM_{L_n} \sim EM_{n\alpha}$ for large n. Indeed without care of details, we can easily estimate

$$Pr\{M_{L_n} > x\} = \sum_{\ell=1}^{\infty} Pr\{U_1 > x \quad \text{or} \quad U_2 > x \text{ or}, \ldots, \text{ or} \quad U_\ell > x\} \; Pr\,\{L_n = \ell\}$$

$$= \sum_{\ell=1}^{n\alpha-\varepsilon} P(\cdot) + \sum_{\ell=n\alpha-\varepsilon}^{n\alpha+\varepsilon} P(\cdot) + \sum_{\ell=n\alpha+\varepsilon}^{\infty} P(\cdot)$$

where $P(\cdot)$ is the summand from the first line of the above. Since $L_n \sim n\alpha$ for $n \to \infty$, hence the first and the third sums tend to zero for $n \to \infty$. Then, noting $EM_{L_n} = \sum_{x=0}^{\infty} Pr\{M_{L_n} > x\}$, one proves immediately $EM_{L_n} \sim EM_{n\alpha}$ (for more detailed proof see [GAL, Sec. 6.2]).

To complete the proof, we note that for some K, $M_K = \max_{1 \le k \le K} U_k$ can be estimated by Lemma 1(ii). Indeed, U_k are i.i.d. with the tail distribution admitting the same behavior as Q_n, since $Pr\{U_k = \ell\} = Pr\{Q_k = \ell \mid \text{the queue is not empty}\} = P(Q_k = \ell)/(1 - P(Q_k = 0))$. This and (2.5) imply that conditions (2.13) hold, so applying Lemma 1(ii) one immediately proves that $M_K \sim \log K/\log \omega^{-1}$. Putting everything together, we have shown

$$EM_{L_n} \sim EM_{n\alpha} \sim \frac{\log n\,\alpha}{\log \omega^{-1}} \tag{2.20}$$

So, by (2.20), (2.19) and the upper bound (2.18), we finally proved our Proposition 1(i). At this point, it is worth to mention that our method of establishing the lower bound is quite different than the one used by Heyde [HEY], however they resemble some similarities. In [HEY] U_k

defined in (2.19) denotes the *maximum* queue length in the k-th busy period (not just merely the queue length as in our analysis), and this distribution is not easy to compute [cf. COH]. Therefore, Heyde evaluated the maximum queue length only in G|M|1 system.

Finally, we note that the results of our Proposition can be a little strengthened by showing *almost sure* convergence.

Theorem. The following holds for large n

$$\limsup_{n \to \infty} \frac{\overline{Q}_n}{\log_\omega n^{-1}} \le 1 \quad a.s. \tag{2.21}$$

$$\limsup_{n \to \infty} \frac{\overline{W}_n}{\log n^{1/\theta}} \le 1 \quad a.s. \tag{2.22}$$

Proof: We concentrate on the queue length and prove only (2.21); the other derivation is almost the same. By definition of $\overline{Q}_n = \max_{i \le k \le 1} \{Q_k\}$, the estimate (2.5) of the tail of the queue length distribution, and stationarity of the system, one finds

$$\begin{aligned} Pr\{\overline{Q}_n > r\} &= Pr\{Q_1 > r \text{ or } Q_2 > r \text{ or } \cdots \text{ or } Q_n > r\} \\ &\le n\, Pr\{Q_1 > r\} = nc_2\omega^n(1+o(1)) \end{aligned} \tag{2.23}$$

Let us set $r = (1+\varepsilon)\log_\omega n^{-1}$ for $\varepsilon > 0$, then (2.23) implies

$$Pr\{\overline{Q}_n > (1+\varepsilon)\log_\omega n^{-1}\} \le n^{-\varepsilon} c_2(1+o(1)) = o(1)$$

so (2.21) follows.

□

Roughly speaking, the Theorem points out that the maximum queue length and waiting time *with probability one* cannot exceed $\log_\omega n^{-1}$ and $\log n^{1/\theta}$ respectively. Interestingly enough, the proof of the theorem suggests the following generalization of Lemma 1.

Lemma 3. Let $Y_1, \cdots, Y_n$ are identically distributed random variables with distribution function $G(y)$ satisfying conditions (2.13). Let also a_n be the smallest root of the equation (2.12a), that is, $n[1 - G(a_n) = 1$. Then

$$\lim_{n \to \infty} \sup \frac{M_n}{a_n} \leq 1 \qquad a.s. \tag{2.24}$$

where $M_n = \max\{M_1, Y_n\}$.

Proof: As in (2.23), we first note the

$$Pr\{M_n > n\} \leq n\,Pr\{Y_1 > r\} \tag{2.25}$$

But, (2.13) implies that for $c = 1 + \varepsilon$ and $r = (1 + \varepsilon)a_n$ the probability $Pr\{Y_1 > (1 + \varepsilon)a_n\}$ = $o(1)Pr\{Y_1 > a_n\}$. This, and (2.12a) (i.e., $Pr\{Y_1 > a_n\} = n^{-1}$) give us

$$Pr\{M_n > (1 + \varepsilon)a_n\} \leq o(1)$$

which leads to (2.24).

□

The interesting fact about Lemma 3 is that we are dealing with *dependent* identically distributed random variables. Moreover, the *identical* distribution restriction can be dropped out in the last lemma.

3. APPLICATIONS

This section discusses in a more detailed fashion one possible application of our main result from Section 2, namely to *hashing with lazy deletion*, which was introduced by Van Wyk and Vitter in [WV], and carefully analyzed by Morrison et al. [MSW], (see also [AHU]).

Here is a sketch of the model description of the hashing with lazy deletion. For more details, the reader should consult [MSW, WV]. We quote from [MSW]: "A sequence of items is given; each item includes a search key, a starting time and an expiration time. The items arrive in the order of their starting times and each item must be kept in a dynamic dictionary (available for searching), until the arrival of an item whose starting time is later than the items expiration time". Two quantities are of interest in such a model, namely, N_t being the number of items that start at or before time t and expire at or after time t, and the actual number of items kept in the dictionary. If H is the number of buckets in the hashing, then by $U_{t,H}$, we denote the actual number of items present at time t. Naturally, $N_t \le U_{t,H}$ for all H.

The model just described can be rephrased in the queueing terminology as follows. We consider $GI|G|\infty$ queueing system. Under FIFO discipline, N_t represents the number of items in $GI|G|\infty$, while $U_{t,H}$ can be interpreted as the number of customers in a system in which only an arriving customer can free the expired (served) customers. In a more descriptive way, the hashing with lazy deletion can be viewed as a queueing system with a gate (a door), which is open *only* by arriving customers. Customers who have completed their services must wait for an arriving customer before they can leave the system.

Our purpose is to analyze $\max_{1 \le t \le n} N_t$ and $\max_{1 \le t \le n} U_{t,H}$. To compare our results with those obtained in [MSW], we adopt the same assumptions, that is, we restrict our analysis to $M|M|\infty$ queueing systems with the arrival rate λ and the service rate μ. At the beginning, we assume $H = 1$, since for any H the following holds [MSW]

$$U_{t,H} = \sum_{i=1}^{H} U_{t,1}^{(i)}(\lambda/H) \tag{3.1}$$

where $U_{t,1}^{(i)}(\lambda/H)$ is the number of customers in a $M|M|\infty$ queueing system with $H = 1$ and with Poisson arrival rate λ/H. We return to (3.1) later, and we shall use notation U_t rather than $U_{t,1}$ as long as it does not cause a confusion.

Let $H = 1$ (single bucket) and consider two $M|M|\infty$ systems: one with FIFO discipline and the other describing the hashing system. Then, for the first system under stationary assumption [KLE]

$$p_j = Pr\{N_t = j\} = \frac{\rho^j}{j!} e^{-\rho} \tag{3.2}$$

where $\rho = \lambda/\mu$. In [MSW] it was proved that the appropriate distribution for U_t is given by

$$P_j = Pr\{U_t = j\} = p_{j-1} \quad j \geq 1 \tag{3.3}$$

We want to estimate $\overline{U}_n = \max_{1 \leq t \leq n} U_t$ and $\overline{N}_n = \max_{1 \leq t \leq n} N_t$. We cannot directly apply our Proposition, since in this model an infinite number of servers is considered. However, the methodology from Section 2 applies without any significant changes. In particular, we conclude that both $\overline{N}_n$ and $\overline{U}_n$ are asymptotically equal to the root a_n of the equation (2.12a) in Lemma 1.

Let us first focus on the maximum queue length $\overline{N}_n$ in the M|M|∞ system with FIFO discipline. The complement of the distribution function for N_t can be computed as follows [REN]

$$1 - F(x) = Pr\{N_t > x\} = \sum_{j=x}^{\infty} \frac{\rho^j}{j!} e^{-\rho} = \frac{\gamma(x,\rho)}{\Gamma(x)} \tag{3.4}$$

where $\gamma(x,\rho)$ is *the incomplete gamma* function defined as [AS, GAU]

$$\gamma(x,\rho) = \int_0^{\rho} t^{x-1} e^{-t} dt$$

and $\Gamma(x)$ is the gamma function defined as $\gamma(x,\infty)$ [AS]. As in Section 2, we need only the tail of the function $1 - F(x)$. But, [GAU]

$$\frac{\gamma(x,\rho)}{\Gamma(x)} \sim \frac{e^{-\rho}\rho^x}{x+1} \qquad x \to \infty \tag{3.5}$$

and ρ is bounded. Then by Lemma 1 and the spirit of Section 2 (which is not presented here to avoid repetitiveness), we recognize that $E\overline{N}_n \sim a_n$ where a_n is the smallest solution of the equation (2.12a) of Lemma 1 (ii), that is, thr following equation

$$n \ \frac{\gamma(a_n \ , \rho)}{\Gamma(a_n)} = 1 \tag{3.6}$$

We use (3.5) to simplify (3.6) to

$$n \ \frac{\rho^{a_n} e^{-\rho}}{a_n + 1} = 1 \tag{3.7}$$

A simple algebra reveals that [SZ2]

$$a_n \sim \frac{\log ne^{-\rho}}{\log \rho^{-1}} - \frac{\log \log ne^{-\rho}}{\log \rho^{-1}}$$

So we finally obtain

$$E\overline{N}_n^r = \log_\rho^r n^{-1}(1 + o(1)) \tag{3.8a}$$

$$\lim_{n \to \infty} \frac{\overline{N}_n}{\log_\rho n^{-1}} = 1 \quad in \ probability \tag{3.9b}$$

Also, in the light of (3.3), one immediately shows that $\overline{U}_n/\overline{N}_n \sim 1$, hence $\overline{U}_n/\log_\rho n^{-1} \sim 1$ in probability.

Finally, we investigate the hashing with $H > 1$ buckets. Let $\rho_H = \frac{\lambda}{H\mu}$. Then, by (3.1)–(3.3), we note that $U_{t,H}$ is a sum of H independent (truncated) Poisson processes, each with parameter ρ_H. This implies that $U_{t,H}$ is Poisson distributed, too, with parameter $H\rho_H = \rho = \lambda/\mu$ [REN]. Therefore, $Pr\{U_{t,H} > x\} = \gamma(x, H\rho_H)/\Gamma(x) \sim e^{-\rho}\rho^x/(x+1)$ for $x \to \infty$, and the same arguments as above lead to

$$E\overline{U}_{n,H} = \log_\rho^r n^{-1}(1 + o(1)) \tag{3.10a}$$

$$\lim_{n \to \infty} \frac{\overline{U}_{n,H}}{\log_\rho n^{-1}} = 1 \quad in \ probability \tag{3.10b}$$

These results are consistent with the extensive numerical calculations presented in [MSW]. They show the same rate of growth. However, since we restrict our interest to the leading factor in the asymptotics of $\max U_t$ and $\max N_t$, we cannot estimate $E\{\max_{1 \le t \le n} U_{t,H} - \max_{1 \le t \le n} N_{t,H}\}$. For this we need exact asymptotics up to the second leading term. From our analysis, however, we know that $E\{\max_{1 \le t \le n} U_{t,H} - \max_{1 \le t \le n} N_{t,H}\} = o(\log n)$. Numerical results reveal that this difference if $O(H)$, and a more careful analysis of the same sort as in our paper, *may* lead to that result.

REFERENCES

[AHU] Aho, A., Hopcroft, J. and Ullman, J., *Data Structures and Algorithms*, Addison-Wesley, Reading, MA (1983).

[AS] Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions*, Dover, New York (1964).

[ASM] Asmussen, S., *Applied Probability and queues*, John Wiley & Sons, Chichester (1987).

[BMS] Baccelli, F., Makowsky, A. and Shwartz, A., The fork-join queue and related systems with synchronization constraints: Stochastic ordering, approximations and computable bounds, preprint (1987).

[BAN] Barndorff-Nielsen, O., On the limit distribution of the maximum of a random number of independent random variables, *Acta Math. Acad. Sci. Hungrar.*, 15, 399–403 (1964).

[COH] Cohen, J.W., Extreme value distribution for the M|G|1 and the GI|M|1 queueing systems, *Ann. Inst. H. Poincare Sect. B.*, 4, 83–98 (1968).

[FEL] Feller, W., *An Introduction to Probability Theory and its Applications*, Vol. II, John Wiley & Sons, New York (1971).

[GAL] Galambos, J., *The Asymptotic Theory of Extreme Order Statistics*, John Wiley & Sons, New York (1978).

[GAU] Gautschi, W., A computational procedure for incomplete gamma functions, *ACM Trans. on Mathematical Software*, 5, 466–481 (1979).

[HEY] Heyde, C.C., On the growth of the maximum queue length in a stable queue, *Operations Res.*, 44, 423–452 (1971).

[IGL] Iglehart, D., Extreme values in the GI|G|1 queue, *The Ann. Math. Statist.*, 43, 627–635 (1972).

[KLE] Kleinrock, L., *Queueing Systems, Vol. 1*, John Wiley & Sons (1976).

[LR1] Lai, T. and Robbins, M., Maximally dependent random variables, *Proc. Nat. Acad. Sci., USA*, 73, 286–288 (1986).

[LR2] Lai, T. and Robbins, H., A class of dependent random variables and their maxima, *Z. Wahrscheinhich*, 42, 89– 111 (1978).

[MSW] Morrison, J., Shepp, L. and Van Wyk, C., A queueing analysis of hashing with lazy deletion, *SIAM J. Comput.*, 16, 1155–1164 (1987).

[REN] Rényi, A., *Probability Theory*, North-Holland, Amsterdam (1970).

[STO] Stoyan, D., *Comparison Methods for Queues and Other Stochastic Models*, John Wiley & Sons, Chichester (1983).

[SW] Szymanski, T.G. and Van Wyk, C., Space efficient algorithms for VLSI artwork analysis, *Proc. 20th Design Automation Conference*, 734–739 (1983).

[SZ1] Szpankowski, W., On the analysis of the average height of a digital tree: Another approach, Purdue University CSD TR-646 (submitted to a journal) (1986).

[SZ2] Szpankowski, W., (Probably) optimal solution to some problem not only on graphs, Purdue University CSD TR-780, 1988 (submitted to a conference) .

[TAK] Takahashi, Y., Asymptotic exponentiality of the tail of the waiting-time distribution in a PH|PH|c queue, *Adv. Appl. Probab.*, 13, 619–630 (1981).

[WV] Van Wyk, C. and Vitter, J.S., The complexity of hashing with lazy deletion, *Algorithmica*, 1, 17–29 (1986).