Euro. Jnl of Applied Mathematics (2016), *vol.* 27, *pp.* 738–755.
doi:10.1017/S0956792516000048

On a storage allocation model with finite capacity

EUNJU SOHN[1] and CHARLES KNESSL[2]

[1]*Department of Science and Mathematics, Columbia College Chicago*
600 South Michigan Avenue, Chicago, IL 60605-1996, USA
email: esohn@colum.edu
[2]*Department of Mathematics, Statistics and Computer Science*
University of Illinois at Chicago, Chicago, Illinois 60607-7045, USA
email: knessl@uic.edu

(*Received 30 July 2015; revised 18 January 2016; accepted 19 January 2016; first published online 17 February 2016*)

We consider a storage allocation model with a finite number of storage spaces. There are m primary spaces that are ranked $\{1, 2, \ldots, m\}$ and R secondary spaces ranked $\{m+1, m+2, \ldots, m+R\}$. Items arrive according to a Poisson process, occupy a space for a random exponentially distributed time, and an arriving item takes the lowest ranked available space. Letting N_1 and N_2 denote the numbers of occupied primary and secondary spaces, we study the joint distribution $\text{Prob}[N_1 = k, N_2 = r]$ in the steady state. The joint process (N_1, N_2) behaves as a random walk in a lattice rectangle. We shall obtain explicit expressions for the distribution of (N_1, N_2), as well as the marginal distribution of N_2. We also give some numerical studies to illustrate the qualitative behaviors of the distribution(s). The main contribution is to study the effects of a finite secondary capacity R, whereas previous studies had $R = \infty$.

Key words: dynamic storage allocation, finite capacity, Poisson process, random walk, steady-state joint distribution, wasted space, Erlang loss model, maximum occupied space

1 Introduction

We consider the following storage allocation model. There are a total of m primary and R secondary storage spaces. The spaces are numbered and ranked, with the primary spaces numbered $\{1, 2, \ldots, m\}$ and the secondary ones $\{m+1, m+2, \ldots, m+R\}$. Customers arrive according to a Poisson process of rate λ, and each customer occupies a storage space for an exponentially distributed amount of time, with the mean occupation time being denoted by $1/\mu$. A new arrival takes the lowest ranked available space, which will be a primary space if one of these is available. If all $m + R$ spaces are filled, then a new arrival is turned away and lost. The policy of taking the lowest ranked space is called "first-fit allocation".

We can view the storage spaces (or servers) as representing parking spaces near a restaurant, with the primary spaces being in a lot right next to the restaurant, and the secondary spaces being located somewhere further away. Higher ranked spaces will be further from the restaurant so it is natural that a customer would use the first-fit policy. Since spaces are sometimes occupied and sometimes not, models of this type are referred to as dynamic storage allocation. Design and analysis of algorithms for dynamic storage allocation is a fundamental part of computer science, as discussed in the classic book of Knuth [7]. In such applications, the stored items (or customers) correspond to records, files or lists, and the storage device is usually a set of consecutive locations or addresses, but it can also be a magnetic tape or disc. As time evolves, items are inserted and deleted, and the storage device, which is a linear array of "cells", will have regions of occupied cells alternating with interior holes. This is referred to as memory fragmentation in computers, and collapsing the holes corresponds to running a defragmentation program. Worst case studies of dynamic storage allocation date back to the 1960's (see [2,7,13]), while average case studies are more recent, from the mid-1980's.

In the language of queueing theory, the model described above with $R = \infty$, i.e., with infinitely many secondary storage spaces, can be called an $M/M/\infty$ queue with ranked servers. If R is finite the model may be called the $M/M/(m+R)/(m+R)$ queue (which is the Erlang loss model) with ranked servers.

In the memory fragmentation applications, it is certainly reasonable to assume a finite memory, but there we would not distinguish between primary and secondary spaces. But it was shown in [3] that distinguishing between primary and secondary spaces leads to a two-dimensional Markov chain whose solution allows for the computation of important quantities, such as wasted space and rank of the maximum occupied cell, associated with the fragmentation model. To be more precise, we let S be the set of occupied spaces, so if spaces 1, 3, 4, 7 and 8 are occupied we have $S = \{1, 3, 4, 7, 8\}$. Then, $\max(S)$ will be the random variable denoting the rank of the highest occupied space, and $W = \max(S) - |S|$ will be called the "wasted space" ($\max(S) = 8$, $W = 8 - 5 = 3$ for the example). These definitions apply equally to the model with finite or infinite storage capacity, but we note that if the total number of storage spaces was 10 ($= m + R$), then the unoccupied spaces ranked 9 and 10 would not be considered as wasted. In Section 2, we discuss in more detail how to calculate the distributions of $\max(S)$ and W from the present model.

We let N_1 and N_2 be the numbers of occupied primary and secondary spaces, and we will focus on the joint distribution of N_1 and N_2, in the steady state. The distributions of both N_1 and $N_1 + N_2$ are readily computed, as these processes behave as Erlang loss models, with m and $m + R$ servers, respectively. Thus, their steady state distributions are truncated Poisson distributions. However, the joint distribution and the distribution of the number N_2 of occupied secondary spaces are much more complicated.

There has been much past work on the model with an infinite (secondary) storage capacity ($R = \infty$), dating back at least to Kosten [8] in 1937. Various aspects of the solution with $R = \infty$ were also studied in [1,3,9] and [10]. In [3] generating functions and analyticity arguments were used to determine $E[z_1^{N_1} z_2^{N_2}]$, but the solution is in a complicated form, that is difficult to evaluate asymptotically, say for $\rho = \lambda/\mu \to \infty$, due to the presence of an alternating sum. The entire monograph of Newell [9] studies this particular model, and probabilistic arguments are used to analyze various limiting cases. Aldous [1] also used probabilistic arguments and in particular showed that the mean wasted space behaves as $E[W] \sim \sqrt{2\rho \log\log\rho}$, $\rho \to \infty$. Preater [10] re-examines the model and gives a new probabilistic derivation of the distribution of $\max(S)$, by reducing the problem to the solution of a random difference equation. In [12] we gave a new derivation of the joint steady state distribution of the (N_1, N_2) process, where instead of using generating functions we used a discrete version of the classic method of separation of variables, to obtain the solution as a contour integral that involves certain polynomials related to hypergeometric functions. From such representations it is possible to obtain a complete set of asymptotic results, and in [4] we studied the distribution of $\max(S)$, in [11] the distribution of the wasted space W, and in [5,6] the joint distribution $\text{Prob}[N_1 = k, N_2 = r]$. In each case, the asymptotics were for $\rho \to \infty$ with various assumptions on m, the number of primary storage spaces. Whereas in [6] the asymptotics were obtained starting from a contour integral representation of the solution, which was expanded by a combination of the saddle point method, singularity analysis and special function asymptotics, in [5] we showed how to obtain as complete a set of asymptotic results by using only the basic difference equation(s) (forward Kolmogorov equation(s)) satisfied by $\pi(k,r) = \text{Prob}[N_1 = k, N_2 = r]$. This analysis involved singular perturbation methods, such as the ray method and asymptotic matching.

The main contribution here is to study the effects of the finite secondary storage capacity ($R < \infty$). Our numerical and analytic studies show that the effects of finite R can be non-trivial and strong. The remainder of the paper is organized as follows. In Section 2, we state the basic equations and summarize our main results. The proofs and detailed calculations are given in Section 3, while Section 4 contains some numerical results. We include a brief discussion of possible future research in Section 5.

2 Statement of the problem and summary of results

We let

$$\rho = \frac{\lambda}{\mu}, \tag{2.1}$$

denote the traffic intensity, which is a dimensionless parameter. We assume that time has been scaled so that $\mu = 1$ and thus $\rho = \lambda$. The joint process (N_1, N_2) corresponds to a continuous time random walk in a lattice rectangle, whose transition rates are indicated in Figure 1. The steady state distribution

$$\begin{aligned} \pi(k,r) &= \pi(k,r;\ m,R,\rho) \\ &= \lim_{t\to\infty} \text{Prob}\left[N_1(t)=k, N_2(t)=r \,|\, N_1(0)=k^{(0)}, N_2(0)=r^{(0)}\right], \end{aligned} \tag{2.2}$$

exists and is independent of the initial values $N_1(0)$ and $N_2(0)$. At times, it will be important to consider the dependence of $\pi(k,r)$ on the capacities m and R.

The balance equations satisfied by $\pi(\cdot,\cdot)$ can be easily obtained as

$$(\rho + k + r)\pi(k,r) = (r+1)\pi(k,r+1) + (k+1)\pi(k+1,r) + \rho\pi(k-1,r); \tag{2.3}$$

$$1 \leqslant k \leqslant m-1, \quad 0 \leqslant r \leqslant R-1$$

$$(\rho + m + r)\pi(m,r) = (r+1)\pi(m,r+1) + \rho\pi(m-1,r) + \rho\pi(m,r-1), \tag{2.4}$$

$$1 \leqslant r \leqslant R-1$$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

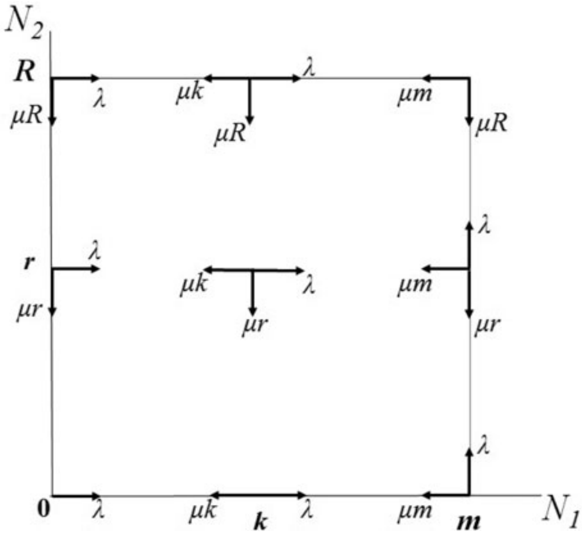

FIGURE 1. A sketch of the transition rates for the random walk.

$$(\rho + k + R)\pi(k, R) = (k + 1)\pi(k + 1, R) + \rho\pi(k - 1, R), \quad 1 \leqslant k \leqslant m - 1 \tag{2.5}$$

$$(\rho + r)\pi(0, r) = (r + 1)\pi(0, r + 1) + \pi(1, r), \quad 0 \leqslant r \leqslant R - 1 \tag{2.6}$$

$$(\rho + R)\pi(0, R) = \pi(1, R) \tag{2.7}$$

$$(\rho + m)\pi(m, 0) = \rho\pi(m - 1, 0) + \pi(m, 1) \tag{2.8}$$

$$(m + R)\pi(m, R) = \rho\pi(m, R - 1) + \rho\pi(m - 1, R). \tag{2.9}$$

The main balance equation is (2.3), which applies in the interior of the lattice rectangle, (2.4)–(2.6) correspond to boundary conditions along three of the four boundaries of the rectangle, and (2.7)–(2.9) are corner conditions. Note that (2.3) applies along $r = 0$ so the fourth boundary does not lead to a separate boundary condition. Also, (2.6) applies at $r = 0$ so the corner equation at $(0, 0)$ is $\rho\pi(0, 0) = \pi(0, 1) + \pi(1, 0)$. We also have the normalization condition

$$\sum_{k=0}^{m} \sum_{r=0}^{R} \pi(k, r) = 1. \tag{2.10}$$

If we view the process N_1 by itself it behaves precisely as the Erlang loss model, or $M/M/m/m$ queue, with m servers. This is well known to have, in the steady state, a truncated Poisson distribution, hence

$$\sum_{r=0}^{R} \pi(k, r) = \frac{\rho^k e^{-\rho}/k!}{\sum_{k=0}^{m} \rho^k e^{-\rho}/k!}, \quad 0 \leqslant k \leqslant m. \tag{2.11}$$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

The total number, $N_1 + N_2$, of occupied servers or stored items also follows a truncated Poisson distribution, with now

$$\sum_{k+r=L} \pi(k,r) = \sum_{k=\max\{0,L-R\}}^{\min\{m,L\}} \pi(k, L-k) = \frac{\rho^L e^{-\rho}/L!}{\sum_{\ell=0}^{m+R} \rho^\ell e^{-\rho}/\ell!}, \tag{2.12}$$

$$0 \leqslant L \leqslant m+R.$$

The marginal of N_2 will be denoted by

$$\mathscr{P}(r) = \sum_{k=0}^{m} \pi(k,r). \tag{2.13}$$

We next discuss the applications of this model to the memory fragmentation model that we described in the introduction. Here, we let $N = m + R$ be the total number of storage cells, no longer distinguishing primary and secondary ones, and consider the maximum occupied cell $\max(S)$ and the wasted space $W = \max(S) - |S|$. The total number $|S|$ of occupied cells follows the truncated Poisson distribution in (2.12). Writing the joint distribution as $\pi(k,r) = \pi(k,r;m,R,\rho)$ to emphasize again its dependence on the numbers m and R of storage cells, we consider $m + R \equiv N$ as fixed. Then, from π we can compute the distribution of $\max(S)$ via

$$\text{Prob}\,[\max(S) \leqslant m] = \sum_{k=0}^{m} \pi(k, 0; m, N-m, \rho), \quad 0 \leqslant m \leqslant N, \tag{2.14}$$

where $\max(S) = 0$ corresponds to the system being empty. Thus, our results can be used to evaluate $\max(S)$, but we must vary the number m of primary cells.

The wasted space W can be computed using the joint distribution by setting $r = 0$ and evaluating the sum

$$\text{Prob}[W \leqslant \mathscr{L}] = \sum_{j=0}^{N-\mathscr{L}-1} \pi(j, 0; \mathscr{L}+j, N-\mathscr{L}-j, \rho) + \sum_{j=N-\mathscr{L}}^{N} \pi(j, 0; N, 0, \rho). \tag{2.15}$$

The second sum in (2.15) is known explicitly, since

$$\pi(j, 0; N, 0, \rho) = \frac{e^{-\rho}\rho^j}{j!} \Big/ \sum_{j=0}^{N} \frac{e^{-\rho}\rho^j}{j!}, \tag{2.16}$$

as $R = 0$ implies that $r = 0$. Note that the wasted space can be at most $N - 1$, and indeed (2.15) does imply that $\text{Prob}[W = N] = 0$.

Below we summarize our main results, which will be in terms of the functions $\mathscr{F}_k^{(J)} = \mathscr{F}_k^{(J)}(\rho)$, where

$$\mathscr{F}_k^{(J)} = \frac{1}{2\pi i} \oint \frac{e^{\rho z}}{(1-z)^{J+1} z^{k+1}} dz = \sum_{i=0}^{k} \frac{\rho^i}{i!} \binom{k-i+J}{J}, \tag{2.17}$$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

and the contour integral is a small loop about $z = 0$. We thus have

$$\mathscr{F}_k^{(-1)} = \frac{\rho^k}{k!}, \quad \mathscr{F}_k^{(0)} = \sum_{i=0}^{k} \frac{\rho^i}{i!}. \tag{2.18}$$

It is possible to express $\mathscr{F}_k^{(J)}$ in terms of generalized hypergeometric functions, but this does not yield any further insight.

Theorem 1 *The distribution in* (2.2) *is given by*

$$\pi(k,r) = \sum_{J=r}^{R} (-1)^{J-r} \binom{J}{r} \mathscr{C}_J \mathscr{F}_k^{(J-1)}; \quad 0 \leqslant k \leqslant m, \quad 0 \leqslant r \leqslant R \tag{2.19}$$

where $\mathscr{F}$ *is in* (2.17) *and*

$$\mathscr{C}_J = \frac{\rho^J \mathscr{B}_J}{J!\mathscr{F}_m^{(J)} \mathscr{F}_m^{(J-1)}}, \quad 0 \leqslant J \leqslant R \tag{2.20}$$

with

$$\begin{aligned} \mathscr{B}_J &= \frac{\rho^m}{m!} - \frac{\rho^m}{(m+R)!\mathscr{F}_{m+R}^{(0)}} \left[\sum_{j=1}^{J} \binom{R}{j} \frac{j!\rho^{R+1-j}\mathscr{F}_m^{(j-1)}}{R+1-j} \right] \\ &= \frac{\rho^m R!}{(m+R)!\mathscr{F}_{m+R}^{(0)}} \left[\sum_{l=0}^{R-J} \frac{\rho^l}{l!} \mathscr{F}_m^{(R-l)} \right], \quad 0 \leqslant J \leqslant R. \end{aligned} \tag{2.21}$$

In particular, we have

$$\mathscr{B}_0 = \frac{\rho^m}{m!}, \quad \mathscr{C}_0 = \left[\sum_{l=0}^{m} \frac{\rho^l}{l!} \right]^{-1}, \tag{2.22}$$

$$\mathscr{B}_R = \frac{\rho^m R! \mathscr{F}_m^{(R)}}{(m+R)!} \left[\sum_{l=0}^{m+R} \frac{\rho^l}{l!} \right]^{-1}, \tag{2.23}$$

$$\mathscr{C}_R = \frac{\rho^{m+R}}{(m+R)!\mathscr{F}_m^{(R-1)}} \left[\sum_{l=0}^{m+R} \frac{\rho^l}{l!} \right]^{-1} \tag{2.24}$$

and

$$\pi(k,R) = \frac{\mathscr{F}_k^{(R-1)}}{\mathscr{F}_m^{(R-1)}} \frac{\rho^{m+R}}{(m+R)!} \left[\sum_{l=0}^{m+R} \frac{\rho^l}{l!} \right]^{-1}, \quad 0 \leqslant k \leqslant m. \tag{2.25}$$

We note that (2.25) gives the probabilities that all R secondary spaces are occupied. The constants $\mathscr{B}_J$ and $\mathscr{C}_J$ depend also on m, R, and ρ, while $\mathscr{F}_k^{(J)}$ depends on ρ but not on the storage capacities m and R. In the limit $R \to \infty$ the first expression in (2.21) shows

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

that $\mathscr{B}_J \to \rho^m/m!$ and then $\pi(k, r; m, R, \rho)$ tends to the limit

$$\pi(k, r; m, \infty, \rho) = \sum_{J=r}^{\infty} \binom{J}{r} \frac{(-1)^{J+r} \rho^{m+J}}{m!J!} \frac{\mathscr{F}_k^{(J-1)}}{\mathscr{F}_m^{(J)} \mathscr{F}_m^{(J-1)}}. \tag{2.26}$$

This expression was previously obtained in [6] and also follows from the generating function results in [3]. Alternate forms of $\pi(k, r; m, \infty, \rho)$, in the form(s) of contour integrals, appear in [6].

From Theorem 1, we can obtain the marginal distribution of N_2 in (2.13) as follows:

Corollary 1 *The distribution of N_2 is given by*

$$\begin{aligned} \mathscr{P}(r) &= \sum_{J=r}^{R} \frac{\rho^J}{J!} (-1)^{J-r} \binom{J}{r} \frac{\mathscr{B}_J}{\mathscr{F}_m^{(J-1)}} \\ &= \frac{\rho^m R!}{(m+R)! \mathscr{F}_{m+R}^{(0)}} \sum_{J=r}^{R} \sum_{l=0}^{R-J} \frac{\rho^{J+l}}{J!l!} (-1)^{J-r} \binom{J}{r} \frac{\mathscr{F}_m^{(R-l)}}{\mathscr{F}_m^{(J-1)}}. \end{aligned} \tag{2.27}$$

From (2.17) we have

$$\begin{aligned} \sum_{k=0}^{m} \mathscr{F}_k^{(J-1)} &= \frac{1}{2\pi i} \oint \frac{e^{\rho z}}{(1-z)^J} \left[\frac{1}{z} + \frac{1}{z^2} + \cdots + \frac{1}{z^{m+1}} \right] dz \\ &= \frac{1}{2\pi i} \oint \frac{e^{\rho z}}{(1-z)^{J+1}} \left[\frac{1}{z^{m+1}} - 1 \right] dz = \mathscr{F}_m^{(J)}, \end{aligned} \tag{2.28}$$

and thus (2.27) follows simply by summing (2.19) over k and using (2.20). The distributions of $\max(S)$ and W follow by using Theorem 1 to calculate (2.14) and (2.15), and here, we must consider the dependence of $\mathscr{B}_J$ and $\mathscr{C}_J$ on m and R, as these vary within (2.14) and (2.15). While this leads to closed-form expressions for these distributions, as triple finite sums, they are not particularly insightful and should be viewed primarily as numerical algorithms for computing the various distributions (cf. Section 4).

3 Analysis

We establish Theorem 1, by first considering the simple cases $R = 1$ and $R = 2$, where there are only one or two secondary spaces. This will indicate how to treat the case of general R.

3.1 The cases $R = 1$ and $R = 2$

First, we take $R = 1$, in which case (2.5) and (2.3) lead to

$$(\rho + k + 1)\pi(k, 1) = (k+1)\pi(k+1, 1) + \rho\pi(k-1, 1), \tag{3.1}$$

and

$$(\rho + k)\pi(k, 0) = \pi(k, 1) + (k+1)\pi(k+1, 0) + \rho\pi(k-1, 0). \tag{3.2}$$

In addition, we have four "corner equations"

$$(\rho+1)\pi(0,1) = \pi(1,1), \tag{3.3}$$

$$\rho\pi(0,0) = \pi(0,1)+\pi(1,0), \tag{3.4}$$

$$(\rho+m)\pi(m,0) = \rho\pi(m-1,0)+\pi(m,1), \tag{3.5}$$

$$(m+1)\pi(m,1) = \rho\pi(m,0)+\rho\pi(m-1,1), \tag{3.6}$$

and these correspond, respectively, to (2.7) with $R = 1$, (2.6) with $r = 0$, (2.8), and (2.9). Let us define $A_k = \pi(k,0)$ and $B_k = \pi(k,1)$, for $0 \leqslant k \leqslant m$.

First, we note that (3.1) does not involve $A_k = \pi(k,0)$, and is a simple second order difference equation for B_k, with two independent solutions. The boundary condition (3.3) eliminates one of these solutions and we thus conclude that

$$B_k = \frac{B_0}{2\pi i}\oint \frac{e^{\rho z}}{1-z}\frac{1}{z^{k+1}}dz = B_0\sum_{l=0}^{k}\frac{\rho^l}{l!}, \tag{3.7}$$

and then $(1+\rho)B_0 = B_1$, so (3.3) holds. With (3.7), (3.2) becomes the inhomogeneous difference equation

$$(\rho+k)A_k-(k+1)A_{k+1}-\rho A_{k-1} = B_k, \quad 1 \leqslant k \leqslant m-1, \tag{3.8}$$

while (3.4) leads to $\rho A_0 - A_1 = B_0$. In view of (3.1) with $\pi(k,1) = B_k$ we see that a particular solution to (3.8) is $A_k^P = -B_k$ and this particular solution will also satisfy (3.4). To this particular solution we must add a homogeneous solution, which will satisfy $(\rho+k)A_k^H = (k+1)A_k^H + \rho A_{k-1}^H$ and $\rho A_0^H = A_1^H$. But then A_k^H must be proportional to the Poisson distribution $\rho^k/k!$, and hence the most general solution to (3.8) and (3.4) is

$$A_k = -B_k + \frac{\rho^k}{k!}(A_0+B_0). \tag{3.9}$$

It remains only to determine A_0 and B_0. From (3.5) and (3.6) we have

$$(\rho+m)A_m = \rho A_{m-1}+B_m \quad \text{and} \quad (m+1)B_m = \rho A_m + \rho B_{m-1}. \tag{3.10}$$

The first equation in (3.10), along with (3.7) and (3.9), leads to

$$\frac{\rho^{m+1}}{m!}(A_0+B_0) = (m+1)B_0\sum_{l=0}^{m+1}\frac{\rho^l}{l!}, \tag{3.11}$$

and then the second equation in (3.10) holds automatically. Then, the normalization condition $\sum_{k=0}^{m}(A_k+B_k) = 1$ leads to $(A_0+B_0) = \left[\sum_{k=0}^{m}\rho^k/k!\right]^{-1}$. Thus, when $R = 1$ the

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

final results are

$$\pi(k,0) = \frac{\rho^k}{k!}\left[\sum_{l=0}^{m}\frac{\rho^l}{l!}\right]^{-1} - \frac{\rho^{m+1}}{(m+1)!}\sum_{l=0}^{k}\frac{\rho^l}{l!}\left[\sum_{l=0}^{m}\frac{\rho^l}{l!}\right]^{-1}\left[\sum_{l=0}^{m+1}\frac{\rho^l}{l!}\right]^{-1} \tag{3.12}$$

$$\pi(k,1) = \frac{\rho^{m+1}}{(m+1)!}\sum_{l=0}^{k}\frac{\rho^l}{l!}\left[\sum_{l=0}^{m}\frac{\rho^l}{l!}\right]^{-1}\left[\sum_{l=0}^{m+1}\frac{\rho^l}{l!}\right]^{-1}. \tag{3.13}$$

From (3.12) and (3.13) we can easily verify the truncated Poisson distribution for N_1 in (2.11), and, with a little more work, also that for $N_1 + N_2$ in (2.12). The latter requires separating the cases $L = 0$, $1 \leqslant L \leqslant m$, and $L = m + 1$ $(= m + R)$.

When $R = 2$ for convenience, we set $A_k = \pi(k, 0)$ and $B_k = \pi(k, 1)$ as before, and also let $D_k = \pi(k, 2)$. We first consider $r = 2$ and then decrement r by 1. Now (2.5) leads to

$$(\rho + k + 2)D_k = (k + 1)D_{k+1} + \rho D_{k-1}, \tag{3.14}$$

while (2.7) becomes $(\rho + 2)D_0 = D_1$. This is again a simple difference equation and boundary condition that may be easily solved using, e.g., generating functions, to give

$$\pi(k,2) = D_k = \mathscr{C}_2\sum_{l=0}^{k}\frac{\rho^l}{l!}(k+1-l) = \mathscr{C}_2\mathscr{F}_k^{(1)}, \tag{3.15}$$

where $\mathscr{C}_2 = D_0$ is a constant, which will depend only on m and ρ. Using $r = 1$ in (2.3) leads to

$$(\rho + k + 1)B_k - \rho B_{k-1} - (k+1)B_{k+1} = 2\mathscr{C}_2\mathscr{F}_k^{(1)}, \quad 1 \leqslant k \leqslant m-1, \tag{3.16}$$

while (2.6) with $r = 1$ gives $(\rho + 1)B_0 = 2D_0 + B_1 = 2\mathscr{C}_2\mathscr{F}_0^{(1)} + B_1$. Thus, a particular solution to (3.16) is $B_k^P = -2\mathscr{C}_2\mathscr{F}_k^{(1)}$ and the general solution to (3.16) and (2.6) (with $r = 1$) is

$$\pi(k,1) = B_k = \mathscr{C}_1\mathscr{F}_k^{(0)} - 2\mathscr{C}_2\mathscr{F}_k^{(1)}, \tag{3.17}$$

where $\mathscr{C}_1$ is another constant. Then, by examining (2.3) with $r = 0$ and imposing the boundary condition $\rho A_0 = A_1 + B_0$ we ultimately obtain

$$\pi(k,0) = A_k = \mathscr{C}_0\frac{\rho^k}{k!} - \mathscr{C}_1\mathscr{F}_k^{(0)} + \mathscr{C}_2\mathscr{F}_k^{(1)}. \tag{3.18}$$

It remains to fix the three constants $\mathscr{C}_0$, $\mathscr{C}_1$, and $\mathscr{C}_2$. From (3.15), (3.17), and (3.18) we have $\pi(k,0) + \pi(k,1) + \pi(k,2) = \mathscr{C}_0\rho^k/k!$, and thus by normalization (2.10),

$$\mathscr{C}_0 = \frac{1}{\mathscr{F}_m^{(0)}} = \left[\sum_{l=0}^{m}\frac{\rho^l}{l!}\right]^{-1}. \tag{3.19}$$

The remaining two constants are obtained by applying the boundary condition along $k = m$ in (2.4) with $r = 1$, along with the corner conditions in (2.9) $(r = 2 = R)$ and (2.8)

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

($r = 0$). This yields three equations but if any two hold the third is automatically satisfied. The final results for $\mathscr{C}_1$ and $\mathscr{C}_2$ are

$$\mathscr{C}_1 = \frac{\rho^{m+2}}{(m+2)!\mathscr{F}^{(0)}_{m+2}\mathscr{F}^{(1)}_{m}} + \left[\frac{\rho^{m+2}}{(m+2)!} + \frac{\rho^{m+1}}{(m+1)!}\right] \frac{l}{\mathscr{F}^{(0)}_{m}\mathscr{F}^{(0)}_{m+2}} \tag{3.20}$$

$$\mathscr{C}_2 = \frac{\rho^{m+2}}{(m+2)!\mathscr{F}^{(0)}_{m+2}\mathscr{F}^{(1)}_{m}}. \tag{3.21}$$

Thus, (3.15) and (3.17)–(3.21) summarize $\pi(k,r)$ for $0 \leqslant k \leqslant m$ and $0 \leqslant r \leqslant 2$. We can easily show that $\mathscr{C}_1$ in (3.20) agrees with (2.20) (with (2.21)) when $R = 2$. Also, with some calculation we can verify that (2.12) holds for $R = 2$ and all $0 \leqslant L \leqslant m+2$.

3.2 General R

First, consider the boundary equation (2.5) and the corner condition in (2.7). The most general solution to these is $\pi(k,R) = \mathscr{C}_R\mathscr{F}^{(R-1)}_k$, where $\mathscr{C}_R$ is a constant. Then by examining (2.3) with $r = R-1$ we conclude that $\pi(k,R-1) = \mathscr{C}_{R-1}\mathscr{F}^{(R-2)}_k - R\mathscr{C}_R\mathscr{F}^{(R-1)}_k$, where $\mathscr{C}_{R-1}$ is another constant. Proceeding by decreasing r we are led to

$$\pi(k,r) = \sum_{J=r}^{R}(-1)^{J-r}\binom{J}{r}\mathscr{C}_J\mathscr{F}^{(J-1)}_k. \tag{3.22}$$

The expression in (3.22) satisfies the main interior equation in (2.3), since every term in the sum does. Also, (3.22) holds along $r = R$ as well as $r = 0$, and the boundary condition in (2.6) holds, since $\mathscr{F}^{(J-1)}_0 = 1$, $\mathscr{F}^{(J-1)}_1 = J+\rho$, and

$$(\rho+r)\sum_{J=r}^{R}(-1)^{J-r}\binom{J}{r}\mathscr{C}_J = (r+1)\sum_{J=r+1}^{R}(-1)^{J-r-1}\binom{J}{r+1}\mathscr{C}_J + \sum_{J=r}^{R}(-1)^{J-r}(J+\rho)\binom{J}{r}\mathscr{C}_J.$$

The last equality follows from $(r-J)\binom{J}{r} = -(r+1)\binom{J}{r+1}$. Thus, (3.22) satisfies (2.3) and (2.5)–(2.7), and it remains only to satisfy the boundary equation in (2.4) along $k = m$, the two corner conditions in (2.8) and (2.9), and the normalization condition (2.10). There are $R+1$ unknown constants, $\mathscr{C}_J$ for $0 \leqslant J \leqslant R$, and (2.4) with (2.8) and (2.9) yields $R-1+2 = R+1$ homogeneous equations, but one of these must be redundant, since one constant will be determined by normalization.

Noting that, for $0 \leqslant J \leqslant R$,

$$\sum_{r=0}^{R}(-1)^{J-r}\binom{J}{r} = \sum_{r=0}^{J}(-1)^{J-r}\binom{J}{r} = \delta_{0J} = \begin{cases}1, & J=0\\ 0, & J\neq 0\end{cases}, \tag{3.23}$$

we obtain from (3.22)

$$\sum_{r=0}^{R}\pi(k,r) = \mathscr{C}_0\mathscr{F}^{(-1)}_k = \mathscr{C}_0\frac{\rho^k}{k!}. \tag{3.24}$$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

This verifies the Poisson marginal distribution of N_1 and determines $\mathscr{C}_0$ as $1/\mathscr{F}_m^{(0)}$, as in (2.22).

To determine $\mathscr{C}_J$ for $1 \leqslant J \leqslant R$ we apply (2.4). The corner condition (2.9) and (3.22) imply that

$$\left[(m+R)\mathscr{F}_m^{(R-1)} + \rho R\mathscr{F}_m^{(R-1)} - \rho\mathscr{F}_{m-1}^{(R-1)}\right]\mathscr{C}_R = \rho\mathscr{F}_m^{(R-2)}\mathscr{C}_{R-1}. \tag{3.25}$$

From (2.17) we can easily establish the following recurrence relations

$$(k+1)\mathscr{F}_{k+1}^{(J-1)} = \rho\mathscr{F}_k^{(J-1)} + J\mathscr{F}_k^{(J)} \tag{3.26}$$
$$(J+k)\mathscr{F}_k^{(J-1)} = J\mathscr{F}_k^{(J)} + \rho\mathscr{F}_{k-1}^{(J-1)} \tag{3.27}$$
$$\mathscr{F}_k^{(J-1)} = \mathscr{F}_k^{(J)} - \mathscr{F}_{k-1}^{(J)}. \tag{3.28}$$

For example, (3.26) follows from

$$\begin{aligned}(k+1)\mathscr{F}_{k+1}^{(J-1)} - \rho\mathscr{F}_k^{(J-1)} &= \frac{1}{2\pi i}\oint e^{\rho z}\left[\frac{k+1}{z^{k+2}(1-z)^J} - \frac{\rho}{z^{k+1}(1-z)^J}\right]dz \\ &= -\frac{1}{2\pi i}\oint \frac{1}{(1-z)^J}\frac{d}{dz}\left(\frac{e^{\rho z}}{z^{k+1}}\right)dz \\ &= \frac{1}{2\pi i}\oint \frac{Je^{\rho z}}{(1-z)^{J+1}z^{k+1}}dz = J\mathscr{F}_k^{(J)},\end{aligned}$$

where we integrated by parts. We have already used a summed version of (3.28) (cf. (2.28)) in calculating the marginal of N_2.

For $1 \leqslant r \leqslant R-1$, (2.4) may be replaced by the "artificial" boundary equation

$$(m+1)\pi(m+1,r) = \rho\pi(m,r-1), \quad 1 \leqslant r \leqslant R-1, \tag{3.29}$$

where $\pi(m+1,\cdot)$ is defined from (2.3), by assuming that (2.3) holds also at $k=m$. Using (3.29) instead of (2.4) will simplify some of the calculations, but it is important to note that (3.29) does not hold when $r=R$. Using (3.22) and (3.29) leads to

$$(m+1)\sum_{J=r}^{R}(-1)^{J-r}\binom{J}{r}\mathscr{C}_J\mathscr{F}_{m+1}^{(J-1)} = \rho\sum_{J=r-1}^{R}(-1)^{J-r+1}\binom{J}{r-1}\mathscr{C}_J\mathscr{F}_m^{(J-1)}. \tag{3.30}$$

Shifting the summation index in the last sum from $J \to J-1$, using the binomial identity

$$\binom{J}{r-1} = \binom{J+1}{r} - \binom{J}{r},$$

and some rearranging of (3.30) lead to

$$\begin{aligned}&\sum_{J=r}^{R}(-1)^{J-r}\binom{J}{r}\left[(m+1)\mathscr{C}_J\mathscr{F}_{m+1}^{(J-1)} - \rho\mathscr{C}_J\mathscr{F}_m^{(J-1)} - \rho\mathscr{C}_{J-1}\mathscr{F}_m^{(J-2)}\right] \\ &\quad = \rho(-1)^{R+1-r}\binom{R+1}{r}\mathscr{C}_R\mathscr{F}_m^{(R-1)},\end{aligned} \tag{3.31}$$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

and this equation holds for $r = 1, 2, \dots, R-1$. By using (3.27) with $J = R$ and $k = m$, (3.25) is equivalent to

$$\left[R\mathscr{F}_m^{(R)} + \rho R \mathscr{F}_m^{(R-1)}\right]\mathscr{C}_R = \rho \mathscr{F}_m^{(R-2)}\mathscr{C}_{R-1}. \tag{3.32}$$

If we introduce $\mathscr{B}_J$ as in (2.20) then (3.31) becomes

$$\sum_{J=r}^{R}(-1)^{J-r}\binom{J}{r}\frac{\rho^J}{(J-1)!}\frac{\mathscr{B}_J - \mathscr{B}_{J-1}}{\mathscr{F}_m^{(J-1)}} = \frac{\rho^{R+1}}{R!}(-1)^{R+1-r}\binom{R+1}{r}\frac{\mathscr{B}_R}{\mathscr{F}_m^{(R)}}. \tag{3.33}$$

Now suppose the sequences $\{f_J\}$ and $\{g_J\}$ are related by

$$\sum_{J=r}^{R}(-1)^{J-r}\binom{J}{r}f_J = g_r, \quad 1 \leqslant r \leqslant R-1. \tag{3.34}$$

Then inverting (3.34) leads to

$$f_J = \binom{R}{R-J}f_R + \sum_{l=J}^{R-1}\binom{l}{J}g_l, \quad 1 \leqslant J \leqslant R. \tag{3.35}$$

By applying (3.34) and (3.35) to (3.33) we obtain the difference $\mathscr{B}_J - \mathscr{B}_{J-1}$ explicitly in terms of $\mathscr{B}_R$, as

$$\begin{aligned}\frac{\rho^J}{(J-1)!}&\frac{\mathscr{B}_J - \mathscr{B}_{J-1}}{\mathscr{F}_m^{(J-1)}} \\ &= \binom{R}{R-J}\frac{\rho^R}{(R-1)!}\frac{\mathscr{B}_R - \mathscr{B}_{R-1}}{\mathscr{F}_m^{(R-1)}} + \frac{\rho^{R+1}}{R!}\frac{\mathscr{B}_R}{\mathscr{F}_m^{(R)}}\sum_{l=J}^{R-1}(-1)^{R+1-l}\binom{l}{J}\binom{R+1}{l}.\end{aligned} \tag{3.36}$$

The sum in (3.36) can be explicitly evaluated as

$$\sum_{l=J}^{R-1}(-1)^{R+1-l}\binom{l}{J}\binom{R+1}{l} = \frac{(R+1)(R-J)}{R+1-J}\binom{R}{J}. \tag{3.37}$$

Using (3.32) and (2.20) (with $J = R$ and $J = R-1$) we find that

$$\frac{\mathscr{B}_{R-1}}{\mathscr{B}_R} = 1 + \rho\frac{\mathscr{F}_m^{(R-1)}}{\mathscr{F}_m^{(R)}}. \tag{3.38}$$

Using (3.36)–(3.38) gives

$$\mathscr{B}_J - \mathscr{B}_{J-1} = -\frac{\rho^{R+1-J}}{R+1-J}\frac{J!}{R!}\binom{R}{J}\frac{\mathscr{F}_m^{(J-1)}}{\mathscr{F}_m^{(R)}}\mathscr{B}_R. \tag{3.39}$$

Since $\mathscr{C}_0 = 1/\mathscr{F}_m^{(0)}$ we have $\mathscr{B}_0 = \rho^m/m!$ by (2.20), and then summing (3.39) leads to

$$\mathscr{B}_J = \frac{\rho^m}{m!} - \left[\sum_{j=1}^{J}\frac{\rho^{R+1-j}}{(R+1-j)!}\mathscr{F}_m^{(j-1)}\right]\frac{\mathscr{B}_R}{\mathscr{F}_m^{(R)}}. \tag{3.40}$$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

Below (3.43) we shall show that

$$\binom{m+R}{m}\mathscr{F}^{(0)}_{m+R} = \sum_{j=1}^{R+1} \frac{\rho^{R+1-j}\mathscr{F}^{(j-1)}_m}{(R+1-j)!}. \tag{3.41}$$

Then (3.40) with $J = R$ becomes

$$\left[\sum_{j=1}^{R+1} \frac{\rho^{R+1-j}\mathscr{F}^{(j-1)}_m}{(R+1-j)!}\right] \frac{\mathscr{B}_R}{\mathscr{F}^{(R)}_m} = \frac{\rho^m}{m!}, \tag{3.42}$$

and using (3.41) to evaluate the sum in (3.42) leads to the expression for $\mathscr{B}_R$ in (2.23). From (2.23) and (2.20) we conclude that

$$\pi(m,R) = \frac{\rho^{m+R}}{(m+R)!\mathscr{F}^{(0)}_{m+R}}, \tag{3.43}$$

which gives the probability that all spaces are occupied, and this agrees with (2.12) if $L = m + R$.

To prove (3.41) we use (2.17) with $J = 0$ and $k = m + R$. Then, we have

$$\begin{aligned}
\mathscr{F}^{(0)}_{m+R} &= \frac{1}{2\pi i}\oint \frac{e^{\rho z}}{(1-z)z^{m+R+1}}dz \\
&= \frac{m!}{(m+R)!}\frac{(-1)^R}{2\pi i}\oint \frac{e^{\rho z}}{1-z}\frac{d^R}{dz^R}\left(\frac{1}{z^{m+1}}\right)dz \\
&= \frac{m!}{(m+R)!}\frac{1}{2\pi i}\oint \frac{1}{z^{m+1}}\frac{d^R}{dz^R}\left(\frac{e^{\rho z}}{1-z}\right)dz \\
&= \frac{m!}{(m+R)!}\frac{1}{2\pi i}\oint \left[\sum_{j=0}^{R}\frac{d^{R-j}}{dz^{R-j}}(e^{\rho z})\frac{d^j}{dz^j}\left(\frac{1}{1-z}\right)\binom{R}{j}\right]\frac{dz}{z^{m+1}} \\
&= \frac{m!}{(m+R)!}\sum_{j=0}^{R}\binom{R}{j}\rho^{R-j}\left[\frac{1}{2\pi i}\oint \frac{j!\,e^{\rho z}}{(1-z)^{j+1}z^{m+1}}dz\right] \\
&= \frac{m!R!}{(m+R)!}\sum_{j=0}^{R}\frac{\rho^{R-j}\mathscr{F}^{(j)}_m}{(R-j)!}.
\end{aligned} \tag{3.44}$$

Here, we integrated by parts R times to get the third equality in (3.44), and used the Leibniz rule for computing the R–fold derivative of the product of $e^{\rho z}$ and $(1-z)^{-1}$. Shifting the index in (3.41) from $j \to j+1$ we conclude that (3.41) is equivalent to (3.44).

We comment that we have also verified analytically after a lengthy calculation, that $\pi(0,0) = 1/\mathscr{F}^{(0)}_{m+R}$, so that (2.12) holds if $L = 0$. We have also done extensive numerical verifications of (2.12) for $1 \leqslant L \leqslant m + R - 1$.

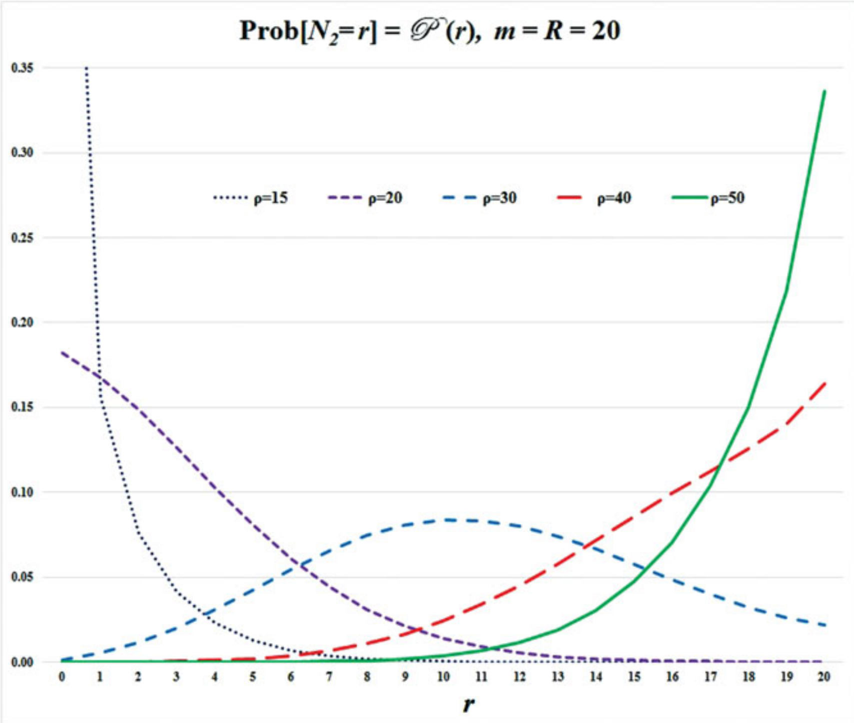

FIGURE 2. A sketch of the distribution of N_2.

4 Numerical studies

We evaluate the distributions in (2.19), (2.27) and (2.14)–(2.16), in order to better understand the effects of the finiteness of the number of storage spaces. We take $m = R = 20$ so that the total number of storage spaces is $N = m+R = 40$, and consider various values of ρ. At times we shall discuss the numerical data in terms of the asymptotic results for the infinite capacity model (where $R = N = \infty$), for $\rho \to \infty$, which appear in [4, 6, 11], as for sufficiently large values of N (relative to ρ) the effects of finite capacity will be quite small.

In Figure 2, we plot the distribution $\mathscr{P}(r)$ of the number N_2 of secondary spaces, for $0 \leqslant r \leqslant R = 20$ and for the values $\rho = 15, 20, 30, 40, 50$. Note that for large ρ and m we would expect that $E[N_1] \approx \min\{\rho, m\}$ and then, since $E[N_1 + N_2] \approx \min\{\rho, m + R\}$ by (2.12), we have $E[N_2] \approx 0$ for $0 < \rho < m$, $E[N_2] \approx \rho - m$ for $m < \rho < m + R$ and $E[N_2] \approx R$ for $\rho > m + R$. The three cases of $E[N_2]$ correspond to not needing the secondary storage spaces, needing all primary and some secondary ones, and needing all primary and secondary spaces. Figure 2 shows that for $\rho = 15$ and $\rho = 20$ the distribution $\mathscr{P}(r)$ is maximal at $r = 0$, becoming less concentrated for the larger value of ρ. When $\rho = 30$ the distribution has a peak at $r = 10$, while for $\rho = 40$ and $\rho = 50$ the distribution is peaked at $r = R$, becoming more concentrated with increasing ρ. For the model with $R = \infty$ we obtained in [6] various limit laws for $\mathscr{P}(r)$ as $\rho \to \infty$. In particular, we showed that for m and ρ large with $\rho/m < 1$, we have $\mathscr{P}(r) \to \delta_{0r}$, where $\delta_{0r} = 0$ if $r \geqslant 1$ and $\delta_{00} = 1$, for $\rho/m = 1 + O(\rho^{-1/2})$ we get a complicated limiting distribution that involves

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

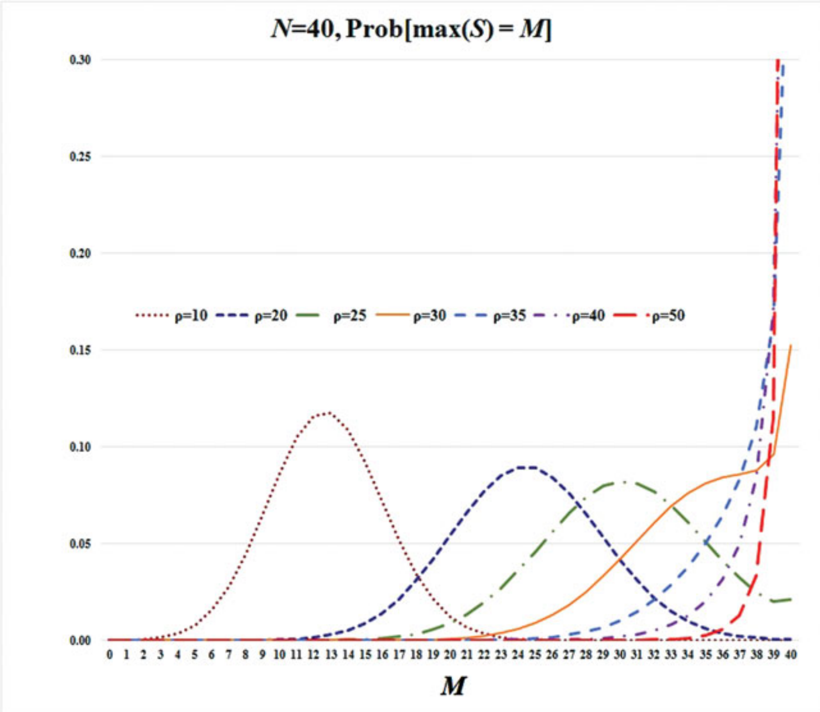

FIGURE 3. A sketch of the distribution of max(S).

the parabolic cylinder functions, while for $\rho/m > 1$ we get a Gaussian limit law, with now $\mathscr{P}(r)$ concentrated on the scale $r = \rho - m + O(\sqrt{\rho})$. Figure 2 suggests that the limit laws for $R = \infty$ may still provide good approximations for $\rho = 15, 20, 30$ but not for $\rho = 40, 50$. When ρ exceeds $m + R = 40$ the effects of the finite capacity become significant, and then presumably new asymptotic limit laws would apply. The graph with $\rho = 50$ suggests that for $(m + R)/\rho < 1$ the mass in $\mathscr{P}(r)$ may concentrate near $r = R$, on the discrete scale $r = R - O(1)$.

In Figure 3, we plot the distribution of max(S), which we computed using (2.19) with $r = 0$ to evaluate numerically the sum in (2.14). This distribution, Prob[max(S) = M], has support over $0 \leqslant M \leqslant N = 40$ and we consider $\rho = 10, 20, 25, 30, 35, 40, 50$. Figure 3 shows that when $\rho = 10$ and $\rho = 20$, this distribution has a single peak (local maximum) in the interior of $[0, 40]$, but when $\rho = 25$ it becomes bimodal, with a secondary peak forming at $M = 40$. When ρ increases to 30 the interior peak has disappeared, though the distribution still has two inflection points. Further studies show that for $N = 40$ and integer values of ρ the bimodal behavior occurs for $25 \leqslant \rho \leqslant 29$. As ρ increases past 30 the distribution is again unimodal with a peak at $M = N$, and becomes more concentrated with increasing ρ. For the model with $R = \infty$ we showed in [4] that Prob[max(S) = M] is concentrated roughly near $M = \rho + O(\sqrt{\rho})$ and with $\beta \equiv (M - \rho)/\sqrt{\rho}$ we obtained an approximation to the distribution in terms of the maximal root $z_0(\beta)$ of the parabolic cylinder function $D_z(-\beta)$. In [1] and [4] it was also shown that $E[\max(S)] = \rho + \sqrt{2\rho \log\log\rho}\,[1 + o(1)]$. The present numerical studies suggest that the large ρ asymptotics for the model with $R = \infty$

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

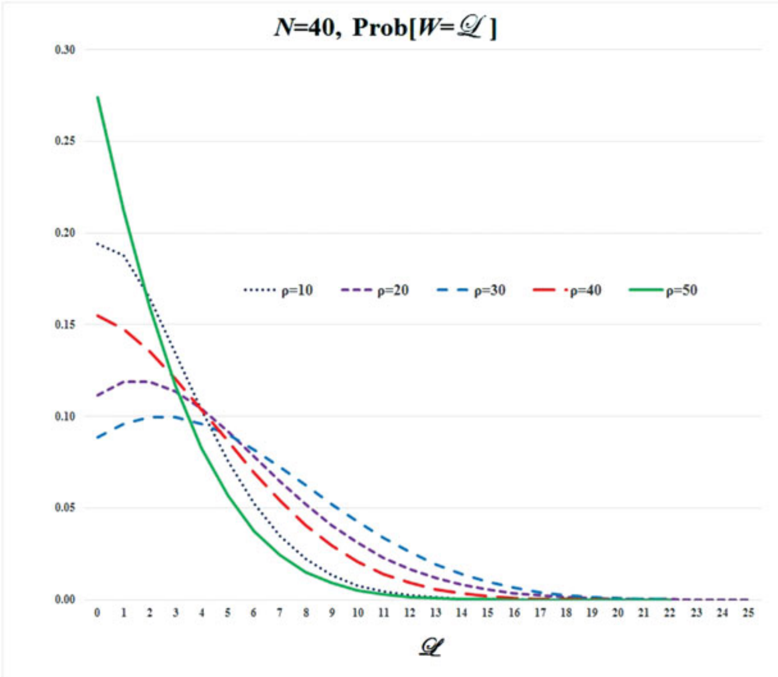

FIGURE 4. A sketch of the distribution of the wasted space W.

may still provide an adequate description when $\rho = 10, 20$ but not for $\rho \geqslant 25$. To describe the evolution into bimodality and the ultimate concentration near $M = N$ would require entirely different asymptotic analyses. The curve with $\rho = 30$ suggests that comparable mass may be accumulating on two different spatial scales, say $M = N - O(\sqrt{\rho})$ and $M = N - O(1)$, while for sufficiently large ρ/N (say $\rho/N > 1$) the latter scale dominates. Certainly considering the limits of ρ and N simultaneously large with either $\rho/N \sim 1$ or $\rho/N > 1$ should lead to some interesting asymptotic results and limit laws for Prob[max(S) = M], which will be quite different from those in [4].

In Figure 4, we plot the distribution of the wasted space W, Prob[$W = \mathscr{L}$], for $N = 40$ and various ρ. We consider only $0 \leqslant \mathscr{L} \leqslant 25$ $(= 5N/8)$, since for $\mathscr{L} > 25$ we are always well into the (right) tail of the distribution. Figure 4 shows that for $\rho = 10$ the distribution is maximal at $\mathscr{L} = 0$, for $\rho = 20$ and $\rho = 30$ there is an interior peak, and for $\rho = 40$ and $\rho = 50$ the distribution is again maximal at $\mathscr{L} = 0$, becoming more concentrated with further increasing ρ. In [11] we obtained for $R = \infty$ a Gaussian limit law for Prob[$W = \mathscr{L}$], for $\rho \to \infty$ with the scaling $\mathscr{L} = \sqrt{2\rho \log\log\rho} + O(\sqrt{\rho})$. This may still provide a reasonable approximation to the finite R model if $\rho = 20$ and $\rho = 30$, and perhaps $\rho = 10$ is too small to see the asymptotic limit law, but it again seems that for $\rho/N \sim 1$ and $\rho/N > 1$ entirely different asymptotics are needed. The numerical studies suggest that for $\rho/N > 1$ the mass concentrates on the scale $\mathscr{L} = O(1)$. The results also suggest that for a fixed large N and increasing ρ (starting with a reasonably large ρ so the Gaussian peak is evident) the Gaussian peak at $\mathscr{L} = \sqrt{2\rho \log\log\rho}$ should reach some maximum value of $\mathscr{L}$, and then further increases of ρ will cause the peak to travel to

https://doi.org/10.1017/S0956792516000048 Published online by Cambridge University Press

the left, eventually disappearing into $\mathscr{L} = 0$. We would again guess that this maximum excursion value of the peak, or that of $E[\mathscr{L}] = E[\mathscr{L}](\rho, N)$, can be characterized using asymptotic analysis.

The numerical studies show that the finiteness of R and $N = m + R$ can significantly affect the distributions, and they also point the way toward possible new asymptotic limit laws, which would hold for $\rho, N \to \infty$ and either $\rho/N \sim 1$ or $\rho/N > 1$.

5 Discussion

To summarize we have obtained an explicit, albeit complicated, solution for a storage allocation model with two sets of storage spaces. Distinguishing between primary and secondary spaces allows for the calculation of the wasted space in the corresponding memory fragmentation model. The numerical studies in Section 4 show that with increasing load ρ, the finiteness of the number R of secondary spaces (or $m + R$ of total spaces) significantly affects the distributions of N_2, $\max(S)$ and W. For larger values of ρ, N_2 becomes concentrated near $N_2 = R$, $\max(S)$ concentrates near $N = m + R$, while the wasted space W concentrates more near $W = 0$, as all the available spaces will tend to be occupied.

In view of the complexity of (2.19)–(2.21) and (2.27) it would also be useful to examine some limiting cases. Here, we only considered the limit $R \to \infty$ (with fixed m and ρ). But an interesting asymptotic limit would have $\rho \to \infty$, with the storage capacities scaled to be commensurately large, with m and R of the order $O(\rho)$. For the model with $R = \infty$, we had previously obtained detailed asymptotic results for the joint distribution of (N_1, N_2) and marginal of N_2 [5,6], the distribution of $\max(S)$ [4], and that of the wasted space W [11], all in the limit of $\rho \to \infty$. Once the asymptotics of (N_1, N_2) are fully understood, those of $\max(S)$ and W may be obtained from (2.14) and (2.15), by evaluating asymptotically the sums that appear. When $R = \infty$ we obtained the asymptotics of $\pi(k, r) = \pi(k, r; m, \infty, \rho)$ by either the saddle point method [6] or singular perturbation techniques [5]. It is difficult to asymptotically evaluate (2.26) directly, in view of the alternating sum, as this causes a lot of cancellation. In [6] we represented the sum as an integral, using a Watson transformation, and then the integral could be evaluated by the saddle point method and/or singularity analysis. The model with $R < \infty$ is more complicated in that the solution in (2.19) (with (2.20) and (2.21)) involves a double sum. Thus, in order to examine the limit $\rho \to \infty$ we would likely need at least a double contour integral, and these are notoriously difficult to expand asymptotically. Alternately, we could try analyzing this limit by approximating the basic balance equations (2.3)–(2.9), using singular perturbation methods, as we did in [5] for the model with $R = \infty$. We are presently examining the large ρ asymptotics of this model, which should ultimately provide some simpler and more insightful expressions for $\pi(k, r)$ and $\mathscr{P}(r)$.

References

[1] Aldous, D. (1986) Some interesting processes arising as heavy traffic limits in an $M/M/\infty$ storage process. *Stoch. Process. Appl.* **22**, 291–313.

[2] Coffman, Jr., E. G. (1983) An introduction to combinatorial models of dynamic storage allocation. *SIAM Rev*. **25**, 311–325.

[3] Coffman, Jr., E. G., Kadota, T. T. & Shepp, L. A. (1986) A stochastic model of fragmentation in dynamic storage allocations. *SIAM J. Comput*. **14**, 416–425.

[4] Knessl, C. (2000) Asymptotic expansions for a stochastic model of queue storage. *Ann. Appl. Probab*. **10**, 592–615.

[5] Knessl, C. (2003) Geometrical optics and models of computer memory fragmentation. *Stud. Appl. Math*. **111**, 185–238.

[6] Knessl, C. (2004) Some asymptotic results for the $M/M/\infty$ queue with ranked servers. *Queueing Syst*. **47**, 201–250.

[7] Knuth, D. E. (1997) *Fundamental Algorithms*, 3rd ed., vol. 1, Addison-Wesley, Reading, MA.

[8] Kosten, L. (1937) Uber Sperrungswahrscheinlichkeiten bei Staffelschltungen. *Electra Nachr.-Tech*. **14**, 5–12.

[9] Newell, G. F. (1984) *The $M/M/\infty$ Service System with Ranked Servers in Heavy Traffic*, Springer, New York.

[10] Preater, J. (1997) A perpetuity and the $M/M/\infty$ ranked server system. *J. Appl. Probab*. **34**, 508–513.

[11] Sohn, E. & Knessl, C. (2008) The distribution of wasted spaces in the $M/M/\infty$ queue with ranked servers. *Adv. Appl. Probab*. **40**, 835–855.

[12] Sohn, E. & Knessl, C. (2008) A simple direct solution to a storage allocation model. *Appl. Math. Lett*. **21**, 172–175.

[13] Standish, T. A. (1980) *Data Structure Techniques*, Addison-Wesley, Reading, MA.