A Queueing System with Queue Length Dependent Service Times, with Applications to Cell Discarding in ATM Networks

by

Doo Il Choi, Charles Knessl and Charles Tier

University of Illinois at Chicago

851 South Morgan St.

Chicago, IL 60607

e-mail: dichoi@hit.halla.ac.kr; knessl@uic.edu; tier@uic.edu

Abstract

A queueing system (M/G$_1$,G$_2$/1/K) is considered in which the service time of a customer entering service depends on whether the queue length, $N(t)$, is above or below a threshold L. The arrival process is Poisson and the general service times S_1 and S_2 depend on whether the queue length at the time service is initiated is $< L$ or $\geq L$, respectively. Balance equations are given for the stationary probabilities of the Markov process $(N(t), X(t))$ where $X(t)$ is the remaining service time of the customer currently in service. Exact solutions for the stationary probabilities are constructed for both infinite and finite capacity systems. Asymptotic approximations of the solutions are given, which yield simple formulas for performance measures such as loss rates and tail probabilities. The numerical accuracy of the asymptotic results is tested.

AMS Classification: 60K25, 24E05

Keywords: queueing systems, asymptotic methods, ATM networks

1 Introduction

1.1 Background

Queueing systems arise in a wide variety of applications such as computer systems and communication networks. A queueing system is a mathematical model to characterize the system, in which the arrivals and the service of customers (users, packets or cells) occur randomly. The customers arrive at the facility and wait in the queue (or buffer) if the server is not available. If there are many customers in the queue, they may suffer long delays which cause poor system performance. Thus, the arrival rate or the service rate may need to be controlled to reduce the delays. These systems may

This research was supported in part DOE Grant DE-FG02-96ER25168. D.I. Choi was supported by the Korea Research Foundation. Current address of D.I. Choi: Department of Mathematics, Halla Institute of Technology, San 66 HeungUp-Li HeungUp-Myon, WonJu-Shi, KangWon-Do, 220-712, South Korea be represented by queueing systems with queue-length-dependent arrival rates or service times. That is, if the queue length exceeds a threshold value, the arrival rate may be reduced (e.g. overload control), or the service rate may be increased (e.g. the cell discarding scheme [2]). Many schemes of traffic control in ATM (asynchronous transfer mode) networks have been analyzed using such threshold-based queueing systems ([2],[5],[8]-[12]).

In this paper, we analyze a queueing system with queue-length-dependent service times. Customers arrive at the queue by a Poisson process, and there is only one server. The service times of customers depend on the queue length. Concretely, we specify a threshold value L for the queue. If the queue length at service initiation of a customer is less than the threshold L (respectively, greater than or equal to the threshold L), the service time of the customer follows a distribution with probability density function $b_1(\cdot)$ (respectively, $b_2(\cdot)$). We believe that our analysis can be extended to the case of multiple thresholds. Both infinite (M/G_1,G_2/1) and finite capacity (M/G_1,G_2/1/K) queues are considered.

The analysis of this queueing system was directly motivated by the cell-discarding scheme for voice packets in ATM networks (see [2, 9, 11]). In [11], a system with deterministic service times was proposed, in which voice packets are divided into high and low priority ATM cells. The cells arrive as a concatenated pair (i.e. two cells per arrival) and the threshold L and the capacity K are measured in terms of cell pairs. The cell discarding occurs at the output of the queue and immediately prior to transmission, based upon the total number of cell pairs in the queue. If this number is less then L then the next pair of cells is transmitted (no cells are discarded). However, if the queue length is greater than or equal to L then only the high priority cell is transmitted. Thus, the low priority cell is discarded. The analysis in [11] is based on numerically solving an embedded chain formulation of the problem. The system is studied only at service time completions. Our results, when specialized to G_i = D_i (deterministic service), are directly applicable to this model.

We analyze this queueing system using the supplementary variable method. We first consider the case of an infinite capacity queue, and obtain an explicit formula for the steady-state queue length distribution p_n. When the service times have different exponential distributions, the queue length distribution has a simple, closed form. We also compute asymptotic approximations to the queue length distribution p_n for various choices of the system parameters. Next, we examine the finite capacity queue, and again obtain explicit expressions for p_n and in particular the probability p_K that the queue is full and the probability p_0 that it is empty. Also, we investigate asymptotic approximations for the queue length distribution for large values of the threshold L and the queue capacity K. We show that this queueing system has very different tail behavior (and hence loss probability) than other $M/G/1$ type models, and that the service tails can sometime determine the tail of the queue length.

There has been some previous analytical work on queueing systems with queue-length-dependent service times [1, 3, 4, 6, 7]. C. M. Harris [6] considered the $M^X/G/1$ queue with queue-length-dependent service times. In particular, if there are i customers in the queue, the service time of the customer starting service has a general distribution depending on i. By using the embedded Markov chain method, C. M. Harris [6] derived the probability generating function for the queue length at the departure epochs. However, the obtained probability generating function contains infinitely many unknown constants. A closed form was obtained only for some special service times of two types. Fakinos [4] analyzed the $G/G/1$ queue in which the service discipline is last-come first-served and the service time depends on the queue length at the arrival epoch of each customer. Abolnikov, Dshalalow and Dukhovny [1] considered queues with bulk arrivals (i.e. compound Poisson input) and state-dependent arrivals and service. Assuming that the state-dependence applies only when the queue length is below a critical level, the authors characterize the generating function (for both transient and steady-state cases) of the queue length probabilities in terms of the roots of a certain equation. Ivnitskiy [7] also considers a model with bulk, state-dependent arrivals and state-dependent service. Using the supplementary variable method, he obtains a recursion relation for the Laplace transforms of the transient queue length probabilities. For a very good recent survey of work on state-dependent queues, we refer the reader to Dshalalow [3].

The model here is a special case of that studied in [7]. However, we are able to give more explicit analytical expressions, from which we can easily obtain asymptotic expansions for tail probabilities and loss rates. These clearly show the qualitative dependence of these performance measures on the arrival rate and the service distribution(s). In particular, we show that the tail behavior of the model with threshold L is much different than the tail behavior of the standard M/G/1 and M/G/1/K models.

1.2 Statement of the Problem

We let $N(t)$ be the queue length at time t, including the customer in service, and let $X(t)$ be the remaining service time of the customer currently in service. Customers arrive according to a Poisson process with arrival rate λ and are served on a first-come first-serve basis. There is a single server and a queue with finite capacity K. An arrival that would cause $N(t)$ to exceed K is lost, without effecting future arrivals. The service time of each customer depends on the queue length at the time that customer's service begins. If the queue length at service initiation is less than L, the service time of that customer is S_1, while if the queue length is greater than or equal to L, the service time is S_2. The service times S_1 and S_2 have density functions $b_1(\cdot)$ and $b_2(\cdot)$ with means m_1 and M_1, respectively. We define $\rho_1 = \lambda m_1$ and $\rho_2 = \lambda M_1$. The process $(N(t), X(t))$ is Markov and $X(t)$ is referred to as a supplementary variable.

The stationary probabilities are denoted by

$$p_n(x)\,dx = \lim_{t\to\infty} Pr[N(t)=n,\ X(t)\in(x,x+dx)],\quad n=1,\ldots,K, \tag{1.1}$$

$$p_0 = \lim_{t\to\infty} Pr[N(t)=0]. \tag{1.2}$$

For finite capacity systems ($K<\infty$), these limits clearly exist. For infinite capacity systems ($K=\infty$), we assume the stability condition

$$\rho_2 = \lambda M_1 = \lambda\int_0^\infty t\, b_2(t)\,dt < 1. \tag{1.3}$$

The balance equations for (1.1)-(1.2) are

$$\begin{aligned}
p_1(0) &= \lambda p_0 && (1.4)\\
-\frac{dp_1}{dx} &= -\lambda p_1(x) + \lambda b_1(x)p_0 + b_1(x)p_2(0) && (1.5)\\
-\frac{dp_n}{dx} &= -\lambda p_n(x) + \lambda p_{n-1}(x) + b_1(x)p_{n+1}(0),\quad n=2,\ldots,L-1 && (1.6)\\
-\frac{dp_L}{dx} &= -\lambda p_L(x) + \lambda p_{L-1}(x) + b_2(x)p_{L+1}(0) && (1.7)\\
-\frac{dp_n}{dx} &= -\lambda p_n(x) + \lambda p_{n-1}(x) + b_2(x)p_{n+1}(0),\quad n=L+1,\ldots,K-1 && (1.8)\\
-\frac{dp_K}{dx} &= \lambda p_{K-1}(x). && (1.9)
\end{aligned}$$

The normalization condition is

$$p_0 + \sum_{n=1}^{K}\int_0^\infty p_n(x)\,dx = 1. \tag{1.10}$$

For infinite capacity systems, we omit equation (1.9) and let $K\to\infty$ in (1.8) and (1.10).

An important local balance result can be obtained by integrating the balance equations with respect to x from 0 to ∞, which leads to

$$p_{n+1}(0) = \lambda\int_0^\infty p_n(x)dx,\quad n\ge 1. \tag{1.11}$$

In the following sections, we construct exact solutions to the infinite capacity model and then the finite capacity model. As we will show, the solution of the infinite capacity model can be used to construct the solution of the finite capacity model. Then we obtain simple formulas for the performance measures by constructing asymptotic approximations to the exact solutions.

2 Infinite Capacity System ($K=\infty$)

We consider the infinite capacity model ($K=\infty$) described by equations (1.4)-(1.8) with normalization condition (1.10). For this model, equation (1.8) is valid for $n\ge L+1$.

M/M$_1$,M$_2$/1 Queue: To illustrate the important characteristics of the solution, we first consider the case in which the service times, S_1 and S_2, are exponential with probability density functions

$$b_1(t) = \mu_1 e^{-\mu_1 t}, \quad b_2(t) = \mu_2 e^{-\mu_2 t}. \tag{2.1}$$

The solution of (1.4)-(1.8) can be constructed in a straightfoward manner as follows. For $n < L$, we assume a solution of (1.4)-(1.6) in the form $p_n(x) = e^{-\mu_1 x} P_n$ which leads to the difference equation

$$P_{n+1} - (1+\rho_1)P_n + \rho_1 P_{n-1} = 0 \tag{2.2}$$

$$P_2 = \rho_1 P_1, \quad P_1 = \lambda p_0 \tag{2.3}$$

where

$$\rho_1 = \frac{\lambda}{\mu_1}. \tag{2.4}$$

The solution of (2.2)-(2.3) is

$$P_n = \lambda p_0 \rho_1^{n-1}, \quad n = 1, \ldots, L-1 \tag{2.5}$$

so that

$$p_n(x) = \lambda p_0 \rho_1^{n-1} e^{-\mu_1 x}, \quad n = 1, \ldots, L-1. \tag{2.6}$$

We now compute $p_L(x)$ by first finding the value of $p_L(0)$ using (1.6) with $n = L-1$ and the known functions $p_{L-1}(x)$ and $p_{L-2}(x)$, which leads to

$$p_L(0) = \lambda p_0 \rho_1^{L-1}. \tag{2.7}$$

We next solve (1.7) for $p_L(x)$ in terms of $p_{L+1}(0)$ to obtain

$$p_L(x) = \frac{\lambda^2 p_0 \rho_1^{L-2}}{\lambda+\mu_1} e^{-\mu_1 x} + \frac{\mu_2}{\lambda+\mu_2} e^{-\mu_2 x} p_{L+1}(0). \tag{2.8}$$

To compute $p_{L+1}(0)$, we substitute (2.8) into the local balance result (1.11) with $n = L$. This leads to

$$p_{L+1}(0) = \lambda p_0 \rho_2 \rho_1^{L-1} \frac{\lambda+\mu_2}{\lambda+\mu_1}, \quad \rho_2 = \frac{\lambda}{\mu_2}$$

which when used in (2.8) gives

$$p_L(x) = \lambda p_0 \rho_1^{L-1} \left[\frac{\mu_1}{\lambda+\mu_1} e^{-\mu_1 x} + \frac{\lambda}{\lambda+\mu_1} e^{-\mu_2 x} \right]. \tag{2.9}$$

For $n > L$, we seek a solution of (1.8) of the form

$$p_n(x) = A_n e^{-\mu_1 x} + B_n e^{-\mu_2 x} \tag{2.10}$$

which leads to the system of difference equations

$$
\begin{aligned}
(\lambda+\mu_1)A_n &= \lambda A_{n-1} \\
(\lambda+\mu_2)B_n &= \lambda B_{n-1}+\mu_2[A_{n+1}+B_{n+1}].
\end{aligned}
$$

The solution of the above equations is

$$
A_n = \left(\tfrac{\rho_1}{1+\rho_1}\right)^{n-L} A_L
$$

$$
B_n = c\rho_2^{n-L} + \frac{\mu_2}{\mu_1}\frac{\lambda}{\mu_2-\mu_1-\lambda}\left(\frac{\rho_1}{1+\rho_1}\right)^{n-L} A_L
$$

where c and A_L are to be determined. By setting $n = L$ in (2.10), with A_n and B_n defined above, and equating the result to (2.9), we find that

$$
\begin{aligned}
A_L &= \lambda p_0 \rho_1^{L-1}\frac{\mu_1}{\lambda+\mu_1} \\
c &= \lambda p_0 \rho_1^{L-1}\frac{\lambda}{\lambda+\mu_1-\mu_2}.
\end{aligned}
$$

Thus, for $n > L$, we have

$$
p_n(x) = \lambda p_0 \rho_1^{L-1}\Bigg\{\left(\frac{\rho_1}{1+\rho_1}\right)^{n-L}\frac{1}{1+\rho_1}e^{-\mu_1 x}
+\left[\frac{\lambda}{\lambda+\mu_1-\mu_2}\rho_2^{n-L}+\frac{-\mu_2}{\lambda+\mu_1-\mu_2}\left(\frac{\rho_1}{1+\rho_1}\right)^{n-L+1}\right]e^{-\mu_2 x}\Bigg\}.
$$

The constant p_0 is determined by normalization using (1.10) and the marginals $p_n = \int_0^\infty p_n(x)dx$ are summarzied below.

Theorem 1 *M/M$_1$,M$_2$/1 Queue: Let $b_i(x) = \mu_i e^{-\mu_i x}$ and $\rho_i = \lambda/\mu_i$ for $i = 1, 2$ and assume that the stability condition $\rho_2 < 1$ is satisfied. The marginal probabilities are given by*

$$
p_n = p_0\rho_1^n, \quad n = 0, 1, \ldots, L-1, \tag{2.11}
$$

$$
p_n = p_0\rho_1^{L-1}\left\{\frac{\lambda}{\lambda+\mu_1-\mu_2}\rho_2^{n-L+1}+\frac{\mu_1-\mu_2}{\lambda+\mu_1-\mu_2}\left(\frac{\rho_1}{1+\rho_1}\right)^{n-L+1}\right\}, \quad n \geq L \tag{2.12}
$$

where

$$
p_0 = 1/\left[\frac{1-\rho_1^L}{1-\rho_1}+\frac{\rho_1^L}{1-\rho_2}\right]. \tag{2.13}
$$

An interesting aspect of the result is the tail behavior as $n - L \to \infty$. The tail probability has a different form depending on the parameters:

$$
p_n \sim p_0\rho_1^{L-1}\begin{cases} \dfrac{\lambda}{\lambda+\mu_1-\mu_2}\rho_2^{n-L+1} & \text{if } \lambda+\mu_1 > \mu_2 \\[2ex] \dfrac{\mu_2-\mu_1}{\mu_2-\lambda-\mu_1}\left(\dfrac{\rho_1}{1+\rho_1}\right)^{n-L+1} & \text{if } \lambda+\mu_1 < \mu_2. \end{cases} \tag{2.14}
$$

M/G$_1$,G$_2$/1 Queue: We now consider the system in which the service times S_1 and S_2 have general distributions. For $n < L$, we solve (1.4)-(1.6) with $L = \infty$ by introducing the generating function

$$G(z,x) = \sum_{n=1}^{\infty} p_n(x) z^n \tag{2.15}$$

into (1.5)-(1.6) to obtain

$$G_x(z,x) + \lambda(z-1)G(z,x) = b_1(x)\lambda(1-z)p_0 - \frac{1}{z} b_1(x) G(z,0) \tag{2.16}$$

where $G_x(z,x) = \partial G(z,x)/\partial x$. The solution of (2.16) is given by

$$G(z,x) = \left[\frac{G(z,0)}{z} - \lambda(1-z)p_0\right] \int_x^{\infty} e^{\lambda(z-1)(t-x)} b_1(t)dt. \tag{2.17}$$

The unknown function $G(z,0)$ is found by setting $x = 0$ in (2.17) to obtain

$$G(z,0) = \frac{\lambda p_0 z(z-1)\hat{b}_1(\lambda - \lambda z)}{z - \hat{b}_1(\lambda - \lambda z)} \tag{2.18}$$

where $\hat{b}_1(s)$ is the Laplace transform of $b_1(t)$. Combining the result (2.18) for $G(z,0)$ in (2.17) and simplifying we obtain

$$G(z,x) = \frac{\lambda p_0 z(z-1)}{z - \hat{b}_1(\lambda - \lambda z)} \int_x^{\infty} e^{\lambda(z-1)(t-x)} b_1(t)dt. \tag{2.19}$$

We invert $G(z,x)$ to find an integral representation of the stationary probabilities in the form

$$p_n(x) = \frac{\lambda p_0}{2\pi i} \int_C \frac{z-1}{z^n[z - \hat{b}_1(\lambda - \lambda z)]} \int_x^{\infty} e^{\lambda(z-1)(t-x)} b_1(t)dt\, dz. \tag{2.20}$$

The contour C is a loop in the complex z-plane which encircles the origin but excludes any other poles of the integrand. By setting $n = 1$ and $x = 0$ in (2.20), we find that (1.4) is satisfied. Based on the above calculation, we see that (2.20) is in fact a solution of (1.5)-(1.6) for $L < \infty$. Thus, we assume that for $n < L$, the solution is of the form

$$p_n(x) = \frac{\lambda k}{2\pi i} \int_C \frac{z-1}{z^n[z - \hat{b}_1(\lambda - \lambda z)]} \int_x^{\infty} e^{\lambda(z-1)(t-x)} b_1(t)dt\, dz, \quad n = 1, \ldots, L-1, \tag{2.21}$$

for some constant k, which will be determined later.

For $n \geq L$, defining a different generating function

$$H(z,x) = \sum_{n=L}^{\infty} z^{n-L} p_n(x) \tag{2.22}$$

the equations (1.7)-(1.8) are transformed into

$$H_x(z,x) + \lambda(z-1)H(z,x) = -\lambda p_{L-1}(x) + \frac{b_2(x)}{z}[p_L(0) - H(z,0)]. \tag{2.23}$$

The solution of (2.23) is

$$H(z,x) = \lambda \int_x^\infty p_{L-1}(t) e^{\lambda(z-1)(t-x)} dt - \frac{p_L(0) - H(z,0)}{z} \int_x^\infty b_2(t) e^{\lambda(z-1)(t-x)} dt. \tag{2.24}$$

As before, we determine $H(z,0)$ by setting $x=0$ in (2.24) to find that

$$H(z,0) = \frac{-p_L(0)\hat{b}_2(\lambda-\lambda z) + \lambda z \int_0^\infty p_{L-1}(t) e^{\lambda(z-1)t} dt}{[z - \hat{b}_2(\lambda - \lambda z)]}. \tag{2.25}$$

The numerator of (2.25) vanishes as $z \to 1^-$ using the local balance result (1.11) with $n = L-1$, i.e.

$$p_L(0) = \lambda \int_0^\infty p_{L-1}(t) dt. \tag{2.26}$$

We use the known solution for $p_{L-1}(x)$ given by (2.21) to compute the integral in the numerator of (2.25) as

$$\int_0^\infty p_{L-1}(t) e^{\lambda(z-1)t} dt = \frac{\lambda k}{2\pi i} \int_C \frac{w-1}{w^{L-1}[w - \hat{b}_1(\lambda - \lambda w)]} \frac{-\hat{b}_1(\lambda - \lambda w) + \hat{b}_1(\lambda - \lambda z)}{\lambda(z-w)} dw \tag{2.27}$$

and using (2.26) we find that for $L \geq 3$

$$p_L(0) = \frac{\lambda k}{2\pi i} \int_C \frac{w-1}{w^{L-1}[w - \hat{b}_1(\lambda - \lambda w)]} dw. \tag{2.28}$$

Thus, $H(z,0)$ is determined by (2.25) using (2.27) and (2.28) up to the constant k. We now use $H(z,0)$ and $p_L(0)$ to construct $H(z,x)$ using (2.24). After some algebra, we find that

$$\begin{aligned} H(z,x) = {} & \frac{\lambda k}{2\pi i} \int_C \frac{w-1}{w^{L-1}[w - \hat{b}_1(\lambda - \lambda w)]} \frac{\int_x^\infty (e^{-\lambda(z-1)(x-u)} - e^{-\lambda(w-1)(x-u)}) b_1(u) du}{z-w} dw \\ & + \frac{\lambda k}{2\pi i (z - \hat{b}_2(\lambda - \lambda z))} \int_C \frac{w-1}{w^{L-1}[w - \hat{b}_1(\lambda - \lambda w)]} \left[\frac{\hat{b}_1(\lambda - \lambda z) - \hat{b}_1(\lambda - \lambda w)}{z-w} - 1 \right] dw \\ & \times \int_x^\infty e^{-\lambda(z-1)(x-u)} b_2(u) du \end{aligned} \tag{2.29}$$

Thus, for $n > L$, the stationary probabilities can be found by inverting (2.22) using

$$p_n(x) = \frac{1}{2\pi i} \int_C \frac{1}{z^{n-L+1}} H(z,x) dz. \tag{2.30}$$

The quantities of interest are the marginal probabilities p_n which are obtained by removing the dependence on x. We first compute

$$H(z) = \int_0^\infty H(z,x) dx = \sum_{n \geq L} p_n z^{n-L}.$$

The following identity is useful in calculating the marginal probabilities

$$\int_0^\infty \left[\int_x^\infty e^{A(x-u)} b_j(u)du\right] dx = \frac{1-\hat{b}_j(A)}{A}.$$

We find that by integrating $H(z,x)$ and simplifying that

$$H(z) = \frac{k}{2\pi i}\frac{\hat{b}_1(\lambda-\lambda z)-z}{z-\hat{b}_2(\lambda-\lambda z)}\int_C \frac{w-1}{w^{L-1}[w-\hat{b}_1(\lambda-\lambda w)]}\frac{dw}{z-w}, \text{ on } C \;\; |w|>|z|. \tag{2.31}$$

In addition, we assume that $L \geq 3$ which is needed for the above simplification and to avoid a degenerate problem. We invert (2.31) to obtain the marginal probabilities for $n = L, L+1, \ldots$,

$$p_n = \frac{k}{(2\pi i)^2}\int\int_{|z|<|w|} \frac{\hat{b}_1(\lambda-\lambda z)-z}{z-\hat{b}_2(\lambda-\lambda z)}\frac{w-1}{w^{L-1}z^{n-L+1}}\frac{dwdz}{[w-\hat{b}_1(\lambda-\lambda w)](z-w)}. \tag{2.32}$$

The contours for the double complex integral are such that $|w|>|z|$ on the w-contour and both are small loops about the origin.

The marginals for $n<L$ are obtained by integrating (2.21) with respect to x which leads to

$$p_n = \frac{k}{2\pi i}\int_C \frac{-1+\hat{b}_1(\lambda-\lambda z)}{z-\hat{b}_1(\lambda-\lambda z)}\frac{1}{z^n}dz, \quad n = 1,\ldots,L-1, \tag{2.33}$$

and, using (1.4), we find that

$$p_0 = \frac{k}{2\pi i}\int_C \frac{(z-1)\hat{b}_1(\lambda-\lambda z)}{z-\hat{b}_1(\lambda-\lambda z)}\frac{1}{z}dz = k. \tag{2.34}$$

Again the contours are small loops inside the unit circle.

To complete the solution, we find the constant k using the normalization condition (1.10). First, we compute

$$\sum_{n=L}^\infty p_n = H(1) = \frac{k}{2\pi i}\frac{\lambda\hat{b}_1'(0)+1}{1+\lambda\hat{b}_2'(0)}\int_{C+}\frac{dw}{w^{L-1}[w-\hat{b}_1(\lambda-\lambda w)]}$$

where on the contour $C+$ we choose $|w|>1$. The only poles inside $C+$ are at at $w=0$ and $w=1$ if $\rho_1<1$. Since $-\hat{b}_1'(0) = m_1$ and $-\hat{b}_2'(0) = M_1$, we find that

$$\sum_{n=L}^\infty p_n = \frac{k}{2\pi i}\frac{1-\rho_1}{1-\rho_2}\int_{C+}\frac{dw}{w^{L-1}[w-\hat{b}_1(\lambda-\lambda w)]} \tag{2.35}$$

where on the contour $C+$ we have $|w|>1$ and $|w|<1+\delta$. In $1<|w|<1+\delta$, the function $w-\hat{b}_1(\lambda-\lambda w)$ is assumed non-zero.

For $n < L$, we must compute $p_0 + \sum_{n=1}^{L-1} p_n$. We simplify (2.33) as follows

$$\begin{aligned} p_n &= \frac{k}{2\pi i}\int_{|z|<1} \frac{\hat{b}_1(\lambda-\lambda z) - z + z - 1}{z - \hat{b}_1(\lambda-\lambda z)}\frac{1}{z^n}dz \\ &= -k\delta_{n1} + \frac{k}{2\pi i}\int_{|z|<1} \frac{z-1}{z-\hat{b}_1(\lambda-\lambda z)}\frac{1}{z^n}dz \end{aligned}$$

which leads to

$$p_0 + \sum_{n=1}^{L-1} p_n = \frac{k}{2\pi i}\int_{|z|<1} \frac{1}{z-\hat{b}_1(\lambda-\lambda z)}\Big(1 - \frac{1}{z^{L-1}}\Big)dz. \tag{2.36}$$

Combining the two sums, we find that

$$\begin{aligned} \frac{1}{p_0} = \frac{1}{k} = \frac{1}{2\pi i}\frac{1-\rho_1}{1-\rho_2}\int_{C+} \frac{dw}{w^{L-1}[w-\hat{b}_1(\lambda-\lambda w)]} \\ + \frac{1}{2\pi i}\int_{|w|<1} \frac{1}{w-\hat{b}_1(\lambda-\lambda w)}\Big(1-\frac{1}{w^{L-1}}\Big)dw. \end{aligned} \tag{2.37}$$

We simplify (2.37) by noting that if $\rho_1 < 1$ then

$$\frac{1}{2\pi i}\int_{C+} \frac{dw}{w-\hat{b}_1(\lambda-\lambda w)} = Res_{w=1}\Big[\frac{1}{w-\hat{b}_1(\lambda-\lambda w)}\Big] = \frac{1}{1-\lambda m_1}$$

so that

$$\frac{1}{p_0} = \frac{1}{k} = \frac{1}{1-\rho_1} + \left[\frac{1-\rho_1}{1-\rho_2} - 1\right]\frac{1}{2\pi i}\int_{C+} \frac{dw}{w^{L-1}[w-\hat{b}_1(\lambda-\lambda w)]}. \tag{2.38}$$

By shifting the integration contour, we can remove the restriction that $\rho_1 < 1$, and obtain the alternate form for p_0 in (2.43).

The final results for the marginals are summarized below.

Theorem 2 *(M/G_1,G_2/1) Let $b_i(x)$ for $i = 1, 2$ be the density functions for the general service times with moments defined by*

$$m_i = \int_0^\infty x^i b_1(x)dx, \quad M_i = \int_0^\infty x^i b_2(x)dx \tag{2.39}$$

and let $\rho_1 = \lambda m_1$ and $\rho_2 = \lambda M_1$. We assume the stability condition $\rho_2 = \lambda M_1 < 1$ is satisfied and that $L \geq 3$. The marginal probabilities are given by

$$p_n = \frac{k}{2\pi i}\int_C \frac{-1+\hat{b}_1(\lambda-\lambda z)}{z-\hat{b}_1(\lambda-\lambda z)}\frac{1}{z^n}dz, \quad n = 1, \ldots, L-1, \tag{2.40}$$

and for $n = L, L+1, \ldots$,

$$p_n = \frac{k}{(2\pi i)^2}\int\int_{|z|<|w|} \frac{\hat{b}_1(\lambda-\lambda z)-z}{z-\hat{b}_2(\lambda-\lambda z)}\frac{w-1}{w^{L-1}z^{n-L+1}}\frac{dwdz}{[w-\hat{b}_1(\lambda-\lambda w)](z-w)}. \tag{2.41}$$

The probability that the system is empty is

$$p_0 = k \tag{2.42}$$

where

$$\frac{1}{k} = \frac{1}{1-\rho_2} + \frac{\rho_2 - \rho_1}{1-\rho_2}\frac{1}{2\pi i}\int_C \frac{dw}{w^{L-1}[w - \hat{b}_1(\lambda - \lambda w)]}. \tag{2.43}$$

All the integration contours are small loops about the origin.

To illustrate the formulas in Theorem 2, we consider the case of exponential service times in which the density functions are defined in (2.1). The Laplace transforms of the density functions are defined by

$$\hat{b}_i(s) = \frac{\mu_i}{\mu_i + s}, \quad i = 1, 2 \tag{2.44}$$

so that for $n < L$, (2.40) reduces to

$$p_n = \frac{k}{2\pi i}\int_C \frac{-dz}{z^n(z - 1/\rho_1)} = k\rho_1^n. \tag{2.45}$$

For $n \geq L$, we must compute the double contour integral. For the exponential case, the formula (2.41) becomes

$$p_n = \frac{-k}{(2\pi i)^2}\int\int_{|z|<|w|} \frac{\mu_2 + \lambda - \lambda z}{\mu_1 + \lambda - \lambda z}\frac{z - 1/\rho_1}{z - 1/\rho_2}\frac{1}{(-\lambda)}\frac{\mu_1 + \lambda - \lambda w}{w - 1/\rho_1}\frac{1}{w^{L-1}}\frac{1}{z^{n-L+1}}\frac{dzdw}{z - w}. \tag{2.46}$$

We compute the w integral first by re-writting the above integral as

$$p_n = \frac{-k}{2\pi i}\int_C \frac{\mu_2 + \lambda - \lambda z}{\mu_1 + \lambda - \lambda z}\frac{z - 1/\rho_1}{z - 1/\rho_2}\frac{1}{z^{n-L+1}}dz\frac{1}{2\pi i}\int_{|z|<|w|}\frac{1}{(-\lambda)}\frac{\mu_1 + \lambda - \lambda w}{w - 1/\rho_1}\frac{1}{w^{L-1}}\frac{dw}{z - w}. \tag{2.47}$$

The poles of the w integral are located at $0, z, 1/\rho_1 > 1$ so that this integral equals $-$(Residue at $1/\rho_1$), which gives

$$p_n = \frac{-k}{2\pi i}\rho_1^{L-1}\int_C \frac{\mu_2 + \lambda - \lambda z}{\mu_1 + \lambda - \lambda z}\frac{1}{z - 1/\rho_2}\frac{1}{z^{n-L+1}}dz. \tag{2.48}$$

The remaining integral can be evaluated by observing that the integrand has poles at $z = 0$, $1/\rho_2$, and $1 + 1/\rho_1$. Since the only pole inside of C is at $z = 0$, we have the equality: $-$(Residue at 0) = (Residue at $1/\rho_2$) + (Residue at $1 + 1/\rho_1$) so that (2.48), after some simplification, reduces to

$$p_n = k\rho_1^{L-1}\left\{\frac{\lambda}{\lambda + \mu_1 - \mu_2}\rho_2^{n-L+1} + \frac{\mu_1 - \mu_2}{\lambda + \mu_1 - \mu_2}\left(\frac{\rho_1}{1 + \rho_1}\right)^{n-L+1}\right\}, \ n \geq L. \tag{2.49}$$

The above results agree with (2.11) and (2.12) in Theorem 1. Also, the formula for p_0 in Theorem 2 reduces to (2.13).

Asymptotic Approximations: Simple formulas can be obtained by asymptotically expanding the exact integral representations for the stationary probabilities given in Theorem 2. An important limit is $L \gg 1$ and $n - L \gg 1$. The asymptotic expansion in this limit depends on the location of the zeros of $w - \hat{b}_1(\lambda - \lambda w)$, $w - \hat{b}_2(\lambda - \lambda w)$, and poles of $\hat{b}_1(w)$. Specifically, we need to locate the singularities of the integrand which are closest to the origin. Let A be the non-zero solution of

$$\lambda + A = \lambda \int_0^\infty e^{Az} b_1(z)\, dz. \tag{2.50}$$

We assume that $\hat{b}_i(\theta)$ are analytic for some θ with $\Re(\theta) < 0$, which insures the existence of a unique solution to (2.50). Clearly, $A > -\lambda$ and satisfies $A \gtreqless 0$ if $\rho_1 \lesseqgtr 1$. We let B be the solution of

$$\lambda + B = \lambda \int_0^\infty e^{Bz} b_2(z)\, dz \tag{2.51}$$

which satisfies $B > 0$ when $\rho_2 < 1$ and $B \downarrow 0$ as $\rho_2 \uparrow 1$. The Laplace transform $\hat{b}_1(\theta)$ may have singularities in the half-plane $\Re(\theta) < 0$. We assume the singularity with the largest real part is at $\theta = -C$ and that

$$\hat{b}_1(\theta) \sim \frac{D}{(\theta + C)^\nu}, \quad \theta \to -C \tag{2.52}$$

for some constants $\nu,\ D > 0$. For example, if b_1 is exponential then $\hat{b}_1(\theta) = \dfrac{\mu_1}{\mu_1 + \theta}$ so that $D = \mu_1$, $C = \mu_1$, and $\nu = 1$ in (2.52). When $b_1(z) = \dfrac{1}{\Gamma(r)} e^{-\mu_1 z r} (\mu_1 r)^r z^{r-1}$, $r > 0$, i.e. r-stage Erlang, then $\hat{b}_1(\theta) = \left(\dfrac{\mu_1 r}{\mu_1 r + \theta}\right)^r$ so that $D = (\mu_1 r)^r$, $C = \mu_1 r$ and $\nu = r$ in (2.52). If the service time is deterministic, i.e. $b_1(z) = \delta(z - m_1)$, then no singularity exists.

To derive the asymptotic formula for $n \geq L$, we again view the double contour integral as an iterated integral and approximate the w integration first. The pole of $\frac{w-1}{w - \hat{b}_1(\lambda - \lambda w)}$ that is closest to the origin is at $w = 1 + A/\lambda$ where A is the root of (2.50). Thus, if $L \gg 1$,

$$\begin{aligned}\int_{|w|=\varepsilon} \frac{w-1}{w - \hat{b}_1(\lambda - \lambda w)} \frac{1}{z - w} \frac{1}{w^{L-1}} dw &\sim -2\pi i Res_{w=1+A/\lambda} \\ &= -2\pi i \frac{(1 + \frac{A}{\lambda}) - 1}{z - (1 + \frac{A}{\lambda})} \frac{1}{1 + \lambda \hat{b}_1'(-A)} \left(\frac{\lambda}{\lambda + A}\right)^{L-1}.\end{aligned} \tag{2.53}$$

We define

$$I_l(A) = \int_0^\infty z^l e^{Az} b_1(z) dz \tag{2.54}$$

so that $I_0(A) = 1 + A/\lambda$ and

$$p_n \sim \frac{A}{\lambda} \left(\frac{\lambda}{\lambda + A}\right)^{L-1} \frac{k}{2\pi i} \frac{1}{\lambda I_1(A) - 1} \int_{|z|=\varepsilon_1} \frac{\hat{b}_1(\lambda - \lambda z) - z}{[z - (1 + A/\lambda)][z - \hat{b}_2(\lambda - \lambda z)]} \frac{dz}{z^{n-L+1}}. \tag{2.55}$$

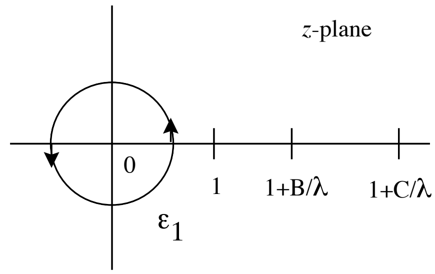

Figure 1: A sketch of the z-plane indicating the location of the poles of the integrand.

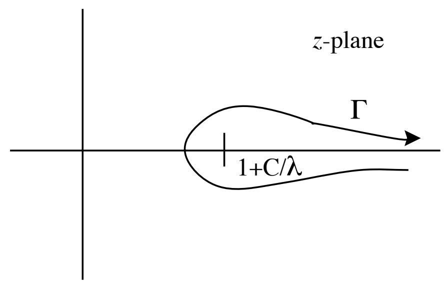

Figure 2: The contour Γ.

If $n - L = \mathcal{O}(1)$, there is no further simplification. However, if $n - L \gg 1$, then we can replace the integral in (2.55) by an asymptotic approximation. The integrand is analytic at $z = 1 + A/\lambda$ but has a pole at $z = 1 + B/\lambda > 1$ and (possibly) an algebraic singularity at $z = 1 + C/\lambda > 1$, (see Figure 1). The asymptotic approximation depends on the relative sizes of C and B.

If $C > B$, then the integral in (2.55) can be approximated as $\int_{|z|=\varepsilon_1}(\cdots)dz \sim -2\pi i\times$ [Residue at $z = \dfrac{B}{\lambda}$] so that as $n - L \to \infty$

$$p_n \sim p_0 \left(\frac{\lambda}{\lambda + A}\right)^{L-1} \left(\frac{\lambda}{\lambda + B}\right)^{n-L+1} \frac{A}{B-A} \frac{\hat{b}_1(-B) - (1 + \frac{B}{\lambda})}{1 - \lambda J_1(B)} \frac{-1}{\lambda I_1(A) - 1} \tag{2.56}$$

where

$$J_l(B) = \int_0^\infty z^l e^{Bz} b_2(z)dz, \quad J_0(B) = 1 + \frac{B}{\lambda}.$$

When $C < B$, the singularity at $1 + \frac{C}{\lambda}$ of $\hat{b}_1(\lambda - \lambda z)$, (cf. (2.52)), determines the dominant asymptotic behavior. The integral in (2.55) can be approximated as $\int_{|z|=\varepsilon_1}(\cdots)dz \sim \int_\Gamma(\cdots)dz$ where the contour Γ is drawn in Figure 2. This leads to the following approximation for (2.55) as

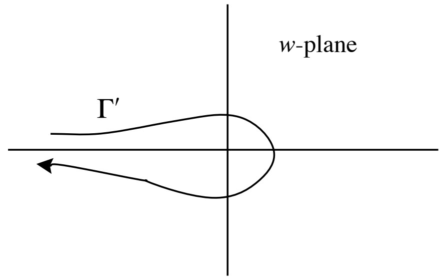

Figure 3: The contour Γ'.

$n-L \to \infty$:

$$p_n \sim \frac{p_0}{2\pi i}\frac{A}{\lambda}\frac{(\frac{\lambda}{\lambda+A})^{L-1}}{\lambda I_1(A)-1}\frac{\lambda}{C-A}\frac{1}{1+\frac{C}{\lambda}-\hat{b}_2(-C)}\int_\Gamma \frac{D}{(\lambda-\lambda z+C)^\nu}\left(\frac{1}{z}\right)^{n-L+1}dz. \qquad (2.57)$$

The integral can be further simplified by letting $z = 1+\frac{C}{\lambda}-\frac{1}{\lambda}\frac{w}{n-L+1}$ to obtain

$$\begin{aligned}\int_\Gamma \frac{D}{(\lambda-\lambda z+C)^\nu}\left(\frac{1}{z}\right)^{n-L+1}dz &= -\int_{\Gamma'}\frac{1}{\lambda}\frac{D}{w^\nu}\left(1+\frac{C}{\lambda}\right)^{-n+L-1}(n-L+1)^{\nu-1} \\ &\quad \times\left[1-\frac{w}{n-L+1}\frac{1}{C+\lambda}\right]^{-n+L-1}dw \\ &\sim -\frac{D}{\lambda}\left(1+\frac{C}{\lambda}\right)^{-n+L-1}(n-L+1)^{\nu-1}\int_{\Gamma'}\frac{e^{w/(C+\lambda)}}{w^\nu}dw\end{aligned}$$

where Γ' is shown in Figure 3. Using the substitution $w = \xi(C+\lambda)$ in the integral, we find that it can be expressed in terms of the Gamma function, so that for $C < B$,

$$\begin{aligned}p_n &\sim p_0\frac{A}{C-A}\frac{1}{\lambda I_1(A)-1}\frac{1}{1+\frac{C}{\lambda}-\hat{b}_2(-C)}\frac{D}{\lambda}\frac{1}{(C+\lambda)^{\nu-1}} \\ &\times\frac{(n-L+1)^{\nu-1}}{\Gamma(\nu)}\left(\frac{\lambda}{\lambda+A}\right)^{L-1}\left(\frac{\lambda}{\lambda+C}\right)^{n-L+1}, \quad n-L\gg 1.\end{aligned}$$

We note that if $\hat{b}_1(\theta)$ has a simple pole at $\theta = -C$ then $\nu = 1$ and the algebraic factor $(n-L+1)^{\nu-1}$ disappears.

We derive an asymptotic formula for $L \gg 1$, $n \gg 1$ and $n < L$ by using (2.40) and noting that

$$\begin{aligned}p_n = \frac{p_0}{2\pi i}\int_C \frac{-1+\hat{b}_1(\lambda-\lambda z)}{z-\hat{b}_1(\lambda-\lambda z)}\frac{1}{z^n}dz &\sim \frac{p_0}{2\pi i}\int_C\frac{z-1}{z-\hat{b}_1(\lambda-\lambda z)}\frac{1}{z^n}dz \\ &\sim \frac{-p_0}{2\pi i}Res_{z=1+A/\lambda}[\cdots] \\ &= \frac{p_0}{\lambda I_1(A)-1}\frac{A}{\lambda}\left(\frac{\lambda}{\lambda+A}\right)^n.\end{aligned}$$

We now compute asymptotic approximations to p_0 for $L \gg 1$. The constant p_0 is defined in Theorem 2 as

$$\frac{1}{p_0} = \frac{1}{1-\rho_2} + \frac{\rho_2 - \rho_1}{1-\rho_2} \frac{1}{2\pi i} \int_C \frac{dw}{w^{L-1}[w - \hat{b}_1(\lambda - \lambda w)]}. \tag{2.58}$$

The poles of the integrand closest to $w = 0$ are at $w = 1$ and $w = 1 + A/\lambda$. We approximate $1/p_0$ when $L \gg 1$ as

$$\begin{aligned} \frac{1}{p_0} &\sim \frac{1}{1-\rho_2} + \frac{\rho_2 - \rho_1}{1-\rho_2} \left[-\frac{1}{1-\rho_1} - \left(\frac{\lambda}{\lambda + A} \right)^{L-1} \frac{1}{1 - \lambda I_1(A)} \right] \\ &= \frac{1}{1-\rho_1} + \frac{\rho_1 - \rho_2}{1-\rho_2} \frac{1}{1 - \lambda I_1(A)} \left(\frac{\lambda}{\lambda + A} \right)^{L-1}. \end{aligned} \tag{2.59}$$

Now if $\rho_1 = \lambda\, m_1 < 1$ then $A > 0$ and

$$p_0 \sim 1 - \rho_1. \tag{2.60}$$

When $\rho_1 > 1$, we find that $-\lambda < A < 0$ so that

$$p_0 \sim \frac{1 - \lambda I_1(A)}{\rho_1 - \rho_2} (1 - \rho_2) \left(1 + \frac{A}{\lambda} \right)^{L-1}. \tag{2.61}$$

The final case is when $\rho_1 \approx 1$. We set $\rho_1 = 1 + \frac{a}{L}$ and find that (2.50) gives

$$A \sim -\frac{2a}{\lambda m_2 L}$$

so that $A \to 0$ as $L \to \infty$. In addition, we find that

$$1 - \lambda I_1(A) = 1 - \lambda \int_0^\infty e^{Az} z b_1(z) dz \sim \frac{a}{L}, \quad A \to 0.$$

Using the above, we find that if $\rho_1 - 1 = a/L = \mathcal{O}(L^{-1})$ then

$$p_0 = \frac{a}{L} \Big/ \left[\exp\left(\frac{2a}{\lambda^2 m_2} \right) - 1 \right]. \tag{2.62}$$

The results are summarized below.

Theorem 3 *Asymptotic approximations for M/G_1,G_2/1 queue: For $L \gg 1$, and A, B, and C defined by (2.50)-(2.52), the stationary probabilities have the following asymptotic approximations:*
$n - L \gg 1$ *and* $C > B$*:*

$$p_n \sim p_0 \left(\frac{\lambda}{\lambda + A} \right)^{L-1} \left(\frac{\lambda}{\lambda + B} \right)^{n-L+1} \frac{A}{B - A} \frac{\hat{b}_1(-B) - (1 + \frac{B}{\lambda})}{1 - \lambda J_1(B)} \frac{-1}{\lambda I_1(A) - 1} \tag{2.63}$$

where

$$J_l(B) = \int_0^\infty z^l e^{Bz} b_2(z)dz, \quad J_0(B) = 1 + \frac{B}{\lambda}.$$

$n - L \gg 1$ *and* $C < B$*:*

$$p_n \sim p_0 \frac{A}{C - A} \frac{1}{\lambda I_1(A) - 1} \frac{1}{1 + \frac{C}{\lambda} - \hat{b}_2(-C)} \frac{D}{\lambda} \frac{1}{(C + \lambda)^{\nu - 1}} \times \frac{(n - L + 1)^{\nu - 1}}{\Gamma(\nu)} \left(\frac{\lambda}{\lambda + A}\right)^{L-1} \left(\frac{\lambda}{\lambda + C}\right)^{n-L+1}, \; n - L \gg 1. \tag{2.64}$$

For $n \gg 1$ *and* $n < L$*,*

$$p_n \sim \frac{p_0}{\lambda I_1(A) - 1} \frac{A}{\lambda} \left(\frac{\lambda}{\lambda + A}\right)^n \tag{2.65}$$

and

$$p_0 \sim \begin{cases} 1 - \rho_1, & \rho_1 < 1 \\ \dfrac{1 - \lambda I_1(A)}{\rho_1 - \rho_2}(1 - \rho_2)\left(1 + \dfrac{A}{\lambda}\right)^{L-1}, & \rho_1 > 1 \\ \dfrac{a}{L} \Big/ \left[\exp\left(\dfrac{2a}{\lambda^2 m_2}\right) - 1\right], & \rho_1 - 1 = a/L = \mathcal{O}(L^{-1}). \end{cases}$$

For $L \gg 1$ and $n - L$ fixed (just above the threshold), the asymptotic result is given by (2.55) with $k = p_0$ given above.

We specialize the above formulas to the exponential service case with $b_i(t) = \mu_i e^{-\mu_i t}$ for $i = 1, 2$. For this case, we find that $m_1 = 1/\mu_1$, $m_2 = 2/\mu_1^2$, $M_1 = 1/\mu_2$, and $M_2 = 2/\mu_2^2$. We explicitly solve (2.50) and (2.51) to obtain

$$A = \mu_1 - \lambda, \quad B = \mu_2 - \lambda.$$

In addition, from (2.52), we find that

$$C = \mu_1, \quad D = \mu_1, \quad \nu = 1.$$

Using the above results, we obtain

$$I_1(A) = \mu_1/\lambda^2, \quad J_1(B) = \mu_2/\lambda^2$$

so that for $n - L \gg 1$, (2.64) simplifies to

$$p_n \sim \frac{\mu_2 - \mu_1}{\mu_2 - \lambda - \mu_1} \left(\frac{\rho_1}{1 + \rho_1}\right)^{n-L+1}$$

which agrees with the result in (2.14). We note that $C < B$ is equivalent to $\mu_1 < \mu_2 - \lambda$. A similar reduction occurs for the case when $C > B$.

Finally, we examine the case when $C \approx B$, more precisely $B - C = \mathcal{O}(L^{-1})$. For $L \gg 1$, $n \gg 1$ and $n < L$, we have the approximation (2.65). We let $C = B + \eta/L$ and note that the integral in (2.55) now has singularities (at $1 + B/\lambda$ and $1 + C/\lambda$) that are close to each other. We consider the integral

$$\text{Integral } = \frac{1}{2\pi i}\int_{|z|=\varepsilon_1} \frac{\hat{b}_1(\lambda - \lambda z) - z}{[z - (1 + A/\lambda)][z - \hat{b}_2(\lambda - \lambda z)]}\frac{dz}{z^{n-L+1}}.$$

We make the change of variables

$$z = \left(1 + \frac{B}{\lambda}\right)\left(1 + \frac{u}{L}\right)$$

and use the approximations

$$\frac{1}{z - \hat{b}_2(\lambda - \lambda z)} \sim \frac{1}{1 - \lambda J_1(B)}\frac{1}{z - 1 - B/\lambda}$$

$$\hat{b}_1(\lambda - \lambda z) - z \sim \hat{b}_1(\lambda - \lambda z) \sim \frac{D}{(C + \lambda - \lambda z)^{\nu}} = DL^{\nu}[\eta - (B + \lambda)u]^{-\nu}$$

to obtain

$$\text{Integral } \sim \frac{\lambda}{B - A}\frac{-DL^{\nu}}{1 - \lambda J_1(B)}\left(\frac{\lambda}{B + \lambda}\right)^{n+1-L}\frac{1}{2\pi i}\int_{-\Gamma'}\frac{e^{\zeta(n/L-1)}}{\zeta[\eta + (B + \lambda)\zeta]^{\nu}}d\zeta.$$

Here the contour Γ' is shown in Figure 3 and it encircles both of the singularities of the integrand. The final approximation to (2.55) is, for $n - L \to \infty$,

$$p_n \sim p_0\frac{A}{B - A}\frac{DL^{\nu}}{[\lambda I_1(A) - 1][1 - \lambda J_1(B)]}\left(\frac{\lambda}{\lambda + A}\right)^{L-1}\left(\frac{\lambda}{B + \lambda}\right)^{n+1-L}\mathcal{F} \tag{2.66}$$

where

$$\mathcal{F} = \frac{1}{2\pi i}\int_{-\Gamma'}\frac{e^{\zeta(\frac{n}{L}-1)}}{\zeta[\eta + (B + \lambda)\zeta]^{\nu}}d\zeta.$$

If $\eta = 0$ (i.e. $B = C$) then $\mathcal{F}$ is explicitly evaluated as

$$\mathcal{F} = (B + \lambda)^{-\nu}\left(\frac{n}{L} - 1\right)^{\nu}/\Gamma(\nu + 1).$$

3 Finite Capacity System ($K < \infty$)

When the queue length has finite capacity, the stationary probabilities are solutions of the system (1.4)-(1.10). The result for the infinite capacity system (2.20) is still valid for $1 \leq n < L$ and the result (2.30) is valid for $L \leq n < K$ except that $k = p_0$ must be recomputed taking into account that now $0 \leq n \leq K$. In addition, the probability $p_K(x)$ must be computed be solving (1.9), i.e.

$$p_K(x) = \lambda\int_x^{\infty} p_{K-1}(s)ds. \tag{3.1}$$

We set $n = K - 1$ in (2.30) and use the result in (3.1) to obtain

$$p_K(x) = \frac{\lambda p_0}{(2\pi i)^2} \iint\limits_{|z|<|w|} \frac{1}{z^{K-L}} \frac{1}{w^{L-1}} \frac{w-1}{w - \hat{b}_1(\lambda - \lambda w)} \frac{1}{z-w} \tag{3.2}$$

$$\times \Bigg\{ \frac{\int_x^\infty b_1(u)du - \int_x^\infty e^{\lambda(1-z)(x-u)} b_1(u)du}{1-z} - \frac{\int_x^\infty b_1(u)du - \int_x^\infty e^{\lambda(1-w)(x-u)} b_1(u)du}{1-w}$$

$$+ \frac{\hat{b}_1(\lambda - \lambda z) - \hat{b}_1(\lambda - \lambda w) - z + w}{z - \hat{b}_2(\lambda - \lambda z)} \frac{\int_x^\infty b_2(u)du - \int_x^\infty e^{\lambda(1-z)(x-u)} b_2(u)du}{1-z} \Bigg\} dzdw.$$

Here we have used the fact that

$$\lambda \int_x^\infty \int_s^\infty e^{\lambda(1-z)(s-u)} b_1(u)duds = \tag{3.3}$$

$$\frac{1}{1-z} \left\{ \int_x^\infty b_1(u)du - \int_x^\infty e^{\lambda(1-z)(s-u)} b_1(u)du \right\}.$$

The main quantities of interest are again the marginal probabilities. The marginals for p_n, $1 \le n < L$ and $L \le n < K$ are given by (2.40) and (2.41) in Theorem 2, respectively. Again, the constant $k = p_0$ must be re-calculated. The marginal probability p_K must be computed using

$$p_K = \int_0^\infty p_K(x)dx.$$

We use the identities

$$\int_0^\infty \int_x^\infty b_1(u)dudx = m_1$$

and

$$\int_0^\infty \int_x^\infty e^{\lambda(1-z)(x-u)} b_1(u)dudx = \frac{1 - \hat{b}_1(\lambda - \lambda z)}{\lambda(1-z)}$$

to obtain

$$p_K = \frac{\lambda p_0}{(2\pi i)^2} \iint\limits_{|z|<|w|} \frac{1}{z^{K-L}} \frac{w-1}{w^{L-1}(w - \hat{b}_2(\lambda - \lambda w))} \frac{1}{z-w} \tag{3.4}$$

$$\times \Bigg\{ \left(\tfrac{1}{1-z} - \tfrac{1}{1-w} \right) m_1 + \tfrac{\hat{b}_1(\lambda - \lambda z) - 1}{\lambda(z-1)^2} - \tfrac{\hat{b}_1(\lambda - \lambda w) - 1}{\lambda(w-1)^2}$$

$$+ \tfrac{1}{z - \hat{b}_2(\lambda - \lambda z)} [\hat{b}_1(\lambda - \lambda z) - \hat{b}_1(\lambda - \lambda w) - z + w] \left[\tfrac{M_1 + \frac{\hat{b}_2(\lambda - \lambda z) - 1}{\lambda(z-1)}}{1-z} \right] \Bigg\}.$$

This result can be further simplified by noting that

$$\int_C \frac{1}{z-w} \frac{w-1}{w^{L-1}} dw = 0, \quad |w| > |z|, \quad L \ge 3,$$

$$\left[M_1 + \frac{1}{\lambda} \frac{\hat{b}_2(\lambda - \lambda z) - 1}{1-z} \right] \frac{1}{z - \hat{b}_2(\lambda - \lambda z)} \frac{1}{1-z} = \frac{-1}{\lambda(1-z)^2} + \frac{1 - \rho_2}{\lambda(z-1)(z - \hat{b}_2(\lambda - \lambda z))},$$

and

$$\frac{\lambda m_1 - 1}{(2\pi i)^2} \iint\limits_{|z|<|w|} \left[\frac{1}{w-1} - \frac{1}{z-1}\right] \frac{w-1}{z^{K-L} w^{L-1}} \frac{1}{z-w} \frac{1}{w - \hat{b}_1(\lambda - \lambda w)} dz dw$$
$$= \frac{1-\rho_1}{2\pi i} \int_C \frac{1}{w^{L-1}} \frac{1}{w - \hat{b}_1(\lambda - \lambda w)} dw.$$

The final result is

$$p_K = p_0 \Bigg\{ 1 + \frac{1-\rho_1}{2\pi i} \int_C \frac{1}{w^{L-1}} \frac{1}{w - \hat{b}_1(\lambda - \lambda w)} dw \tag{3.5}$$
$$+ \frac{1-\rho_2}{(2\pi i)^2} \iint\limits_{|z|<|w|} \frac{1}{z^{K-L} w^{L-1}} \frac{w-1}{z-1} \frac{1}{z-w} \frac{\hat{b}_1(\lambda - \lambda z) - z}{[w - \hat{b}_1(\lambda - \lambda w)][z - \hat{b}_2(\lambda - \lambda z)]} dz dw \Bigg\}.$$

The normalization constant p_0 is determined by (1.10), which we write as

$$\frac{1}{p_0} = 1 + \sum_{j=1}^{L-1} \left(\frac{p_j}{p_0}\right) + \sum_{j=L}^{K-1} \left(\frac{p_j}{p_0}\right) + \frac{p_K}{p_0}. \tag{3.6}$$

Substituting (2.40), (2.41) (with k replaced by p_0), and (3.5) into (3.6), we find after simplifying that

$$\frac{1}{p_0} = 1 + \frac{\rho_1}{2\pi i} \int_C \frac{1}{w^{L-1}} \frac{1}{\hat{b}_1(\lambda - \lambda w) - w} dw \tag{3.7}$$
$$+ \frac{\rho_2}{(2\pi i)^2} \iint\limits_{|z|<|w|} \frac{1}{z^{K-L} w^{L-1}} \frac{w-1}{z-1} \frac{1}{z-w} \frac{\hat{b}_1(\lambda - \lambda z) - z}{[w - \hat{b}_1(\lambda - \lambda w)][\hat{b}_2(\lambda - \lambda z) - z]} dz dw. \tag{3.8}$$

The results are summarized below.

Theorem 4 *(M/G$_1$,G$_2$/1/K) Let $b_i(x)$ for $i = 1, 2$ be the density functions for the general service times as defined in Theorem 2 and let $L \geq 3$. The marginal probabilities are given by*

$$p_n = \frac{p_0}{2\pi i} \int_C \frac{-1 + \hat{b}_1(\lambda - \lambda z)}{z - \hat{b}_1(\lambda - \lambda z)} \frac{1}{z^n} dz, \quad n = 1, \ldots, L-1, \tag{3.9}$$

and for $n = L, \ldots,$ K-1,

$$p_n = \frac{p_0}{(2\pi i)^2} \iint\limits_{|z|<|w|} \frac{\hat{b}_1(\lambda - \lambda z) - z}{z - \hat{b}_2(\lambda - \lambda z)} \frac{w-1}{w^{L-1} z^{n-L+1}} \frac{dw dz}{[w - \hat{b}_1(\lambda - \lambda w)](z-w)}. \tag{3.10}$$

The probability that the system is full is

$$p_K = p_0 \Bigg\{ 1 + \frac{1-\rho_1}{2\pi i} \int_C \frac{1}{w^{L-1}} \frac{1}{w - \hat{b}_1(\lambda - \lambda w)} dw \tag{3.11}$$
$$+ \frac{1-\rho_2}{(2\pi i)^2} \iint\limits_{|z|<|w|} \frac{1}{z^{K-L} w^{L-1}} \frac{w-1}{z-1} \frac{1}{z-w} \frac{\hat{b}_1(\lambda - \lambda z) - z}{[w - \hat{b}_1(\lambda - \lambda w)][z - \hat{b}_2(\lambda - \lambda z)]} dz dw \Bigg\}.$$

The normalization constant p_0, i.e. the probability that the system is empty, is

$$\frac{1}{p_0} = 1 + \frac{\rho_1}{2\pi i}\int_C \frac{1}{w^{L-1}}\frac{1}{\hat{b}_1(\lambda-\lambda w)-w}dw \tag{3.12}$$
$$+\frac{\rho_2}{(2\pi i)^2}\int\int_{|z|<|w|}\frac{1}{z^{K-L}w^{L-1}}\frac{w-1}{z-1}\frac{1}{z-w}\frac{\hat{b}_1(\lambda-\lambda z)-z}{[w-\hat{b}_1(\lambda-\lambda w)][\hat{b}_2(\lambda-\lambda z)-z]}dzdw.$$

On all the contours it is assumed that the origin is the only singularity within the contour.

Asymptotic Approximations: We compute asymptotic approximations for the finite capacity system as $K,\ L\to\infty$ and for various values of ρ_1 and ρ_2. The asymptotic expansions of p_n, except for p_0 and p_K, are the same as in the infinite-capacity systems and are given in Theorem 3 for the two cases $B \lesseqgtr C$.

Next, we expand p_0 given by (3.12). The asymptotic approximation for the integrals depends on the location of the singularities of the integrand that are closest to the origin and follows closely the results for the infinite capacity system in the previous section (cf. the derviation of (2.59)). For the single integral, these poles are at $w=1$ and $w=1+A/\lambda$, where $A>0$ if $\rho_1<1$. The dominant poles of the double integral are located at $z=1$, $z=1+B/\lambda$ and $z=1+C/\lambda$ where $C>0$ and $B \gtrless 0$ if $\rho_2 \lesseqgtr 1$. Thus, for $K,\ L\gg 1$, $K-L\gg 1$, and $B\neq C$, we find that

$$\frac{1}{p_0}\sim\frac{1}{1-\rho_1}+\frac{\rho_1-\rho_2}{1-\rho_2}\frac{1}{1-\lambda I_1(A)}\left(\frac{\lambda}{\lambda+A}\right)^{L-1} \tag{3.13}$$
$$+\left(\frac{\lambda}{\lambda+A}\right)^{L-1}\left(\frac{\lambda}{\lambda+B}\right)^{K-L}\frac{\rho_2 A}{B(B-A)}\frac{\lambda(I_0(B)-J_0(B))}{(1-\lambda I_1(A))(\lambda J_1(B)-1)}$$
$$+\left(\frac{\lambda}{\lambda+A}\right)^{L-1}\left(\frac{\lambda}{\lambda+C}\right)^{K-L}\frac{\rho_2 A}{C(C-A)}\frac{1}{1-\lambda I_1(A)}\frac{D}{1+C/\lambda-J_0(C)}\left(\frac{K-L}{\lambda+C}\right)^{\nu-1}\frac{1}{\Gamma(\nu)}$$

uniformily in ρ_1 and ρ_2.

We can simplify (3.13) for different values of the parameters. For example, since $C>0$, the last term in (3.13) is always negligible and may be dropped. The simplifications of (3.13) are summarized below.

1. $\rho_1<1$ and $\rho_2<1$

$$p_0\sim 1-\rho_1\equiv\nu_- \tag{3.14}$$

2. $\rho_1<1$ and $\rho_2\approx 1$

$$p_0\sim 1-\rho_1 \tag{3.15}$$

3. $\rho_1\approx 1$ and $\rho_2<1$

$$p_0\sim\frac{a}{L}\left[\frac{1}{\exp(2a/(\lambda^2 m_2))-1}\right],\quad \rho_1-1=\frac{a}{L} \tag{3.16}$$

4. $\rho_1 > 1$ and $\rho_2 < 1$

$$p_0 \sim \frac{1-\rho_2}{\rho_1-\rho_2}[1-\lambda I_1(A)]\left(1+\frac{A}{\lambda}\right)^{L-1} \tag{3.17}$$

5. $\rho_1 \approx 1$ and $\rho_2 \approx 1$ (heavy traffic limit)

$$p_0 \sim \frac{1}{L}\left[\frac{e^{2a/(\lambda^2 m_2)}-1}{a}+e^{2a/(\lambda^2 m_2)}\frac{e^{2b\beta/(\lambda^2 M_2)}-1}{b}\right]^{-1}, \tag{3.18}$$
$$\rho_1 - 1 = \tfrac{a}{L},\quad \rho_2 - 1 = \tfrac{b}{L},\quad \beta = \tfrac{K-L}{L}$$

6. $\rho_1 > 1$ and $\rho_2 > 1$

$$p_0 \sim \left(1+\frac{A}{\lambda}\right)^{L-1}\left(1+\frac{B}{\lambda}\right)^{K-L}\frac{B(B-A)[1-\lambda I_1(A)][\lambda J_1(B)-1]}{\rho_2 A\lambda[I_0(B)-J_0(B)]} \equiv \nu_+ \tag{3.19}$$

7. $\rho_1 \approx 1$ and $\rho_2 > 1$

$$p_0 \sim \nu_+ \tag{3.20}$$

8. $\rho_1 > 1$ and $\rho_2 \approx 1$

$$p_0 \sim \left(1+\frac{A}{\lambda}\right)^{L-1}[1-\lambda I_1(A)]\frac{1}{\rho_1-1}\frac{b}{L}\frac{1}{\exp(2b\beta/(\lambda^2 M_2))-1} \tag{3.21}$$

9. $\rho_1 < 1$ and $\rho_2 > 1$

   (a) $\frac{\lambda}{\lambda+A}\left(\frac{\lambda}{\lambda+B}\right)^{\beta} < 1: \quad p_0 \sim 1-\rho_1 = \nu_-$

   (b) $\frac{\lambda}{\lambda+A}\left(\frac{\lambda}{\lambda+B}\right)^{\beta} > 1: \quad p_0 \sim \nu_+$

   (c) $\frac{\lambda}{\lambda+A}\left(\frac{\lambda}{\lambda+B}\right)^{\beta} = 1+\mathcal{O}(L^{-1}): \quad p_0 \sim \dfrac{\nu_-\nu_+}{\nu_+ + \nu_-}.$

Note that in the last case p_n has a bimodal behavior as $K,\ L \to \infty$, with peaks near $n = 0$ and $n = K$.

The asymptotic expansion of p_K as $K,\ L \to \infty$ is computed by using (3.5) for p_K/p_0 and the expansions of the integrals which were derived for the infinite capacity case in Section 2. The results are summarized below.

Theorem 5 *Asymptotic expansions for the M/G$_1$,G$_2$/1/K queue: For $K,\ L \gg 1$ and A, B, and C defined by (2.50)-(2.52), the stationary probabilities p_n have the asymptotic expansion (2.65) for $n < L$ and (2.63)- (2.64) for $n - L \gg 1$. The asymptotic expansion for p_K is*

$$\frac{p_K}{p_0} \sim \left(\frac{\lambda}{\lambda+A}\right)^{L-1}\frac{A(\rho_2-1)}{1-\lambda I_1(A)}\begin{cases}\frac{\lambda[I_0(B)-J_0(B)]}{B(B-A)[\lambda J_1(B)-1]}\left(\frac{\lambda}{\lambda+B}\right)^{K-L}, & B<C\\ \left(\frac{K-L}{C+\lambda}\right)^{\nu-1}\frac{D}{C(C-A)[1+C/\lambda-J_0(C)]}\frac{1}{\Gamma(\nu)}\left(\frac{\lambda}{\lambda+C}\right)^{K-L}, & B>C\end{cases} \tag{3.22}$$

where the asymptotic expansions of p_0 are given by items 1-9 above for various values of ρ_1 and ρ_2.

4 Discussion and Numerical Results

We now demonstrate the usefulness of our results for a finite capacity system, namely the M/D$_1$,D$_2$/1/K queue. For this model, the density functions of the service times are

$$b_1(x) = \delta(x - m_1), \quad b_2(x) = \delta(x - M_1). \tag{4.1}$$

This model corresponds to cell-discarding model analyzed in [11] if $M_1 = m_1/2$.

The exact solution is given in Theorem 4 with (4.1) in place of the general density functions. We illustrate how to numerically compute the exact solution. This calculation is only feasible for moderate L and K. For large values of L and K, we constructed the asymptotic approximations in Theorem 5. We demonstrate the accuracy of our asymptotic results by comparison with the exact solution. As we will see below, our asymptotic results are quite useful for moderate values of n, L, and K and are extremely accurate when n, L, and K are large.

To evaluate the exact solution for p_n in Theorem 4, we must compute the complex integrals in (3.9)-(3.12). The simplest approach is to evaluate the integrals by using the method of residues. The residue at $z = 0$ in (3.9) is difficult to compute if n is large. This is a key motivation for the development of asymptotic expansions. For moderate values of n, we compute the residues using the symbolic computation program *Maple*.

The double complex integrals in (3.10)-(3.12) can be evaluated by re-writting the term $1/(z - w)$ (which couples the two integrals) as

$$\frac{1}{z-w} = \frac{-1}{w}\frac{1}{1 - z/w} = \frac{-1}{w}\sum_{j=0}^{\infty}\left(\frac{z}{w}\right)^j.$$

Using this result in the integrand in (3.10) allows the double integral to be separated into

$$\sum_{j=0}^{n-L}\left(\frac{1}{2\pi i}\int_C z^{j+L-n-1}\cdots dz\right)\left(\frac{1}{2\pi i}\int_C \frac{-1}{w^{j+L}}\cdots dw\right),$$

where $\cdots$ can be identified from (3.10). The sum truncates at $j = n - L$ since for $j > n - L$ the z integral vanishes. Finally, we must evaluate both integrals by computing the residues at $z = 0$ and $w = 0$ for each value of j. Again the calculation is only feasible for L and $n - L$ moderate in size. As before, we use *Maple* to perform the calculation.

The calculation of the asymptotic approximation requires computing the constants A and B by solving (2.50) and (2.51), respectively. For deterministic service times, there is no singularity C so that $C = \infty > B$. Given the constants A and B, we evaluate the formulas (2.65) and (2.63) for p_n if $n < K$ and (3.22) for p_K. The constant p_0 is given by (3.14)-(3.21), depending on the values of ρ_1 and ρ_2.

In Table 1, we consider a queue with the threshold $L = 5$ and capacity $K = 10$. The arrival rate is $\lambda = 1$ and the service times are $m_1 = 1/2$ and $M_1 = 1/4$ so that $\rho_1 = 1/2$ and $\rho_2 = 1/4$. Thus, when the queue length exceeds the threshold, the service time of the jobs entering service is half of the original service time. For these values of ρ_i, the asymptotic value of p_0 is given by (3.14). The solutions of (2.50) and (2.51) are $A \approx 2.512$ and $B \approx 9.346$, respectively.

Table 1

$L = 5,\ K = 10,\ \rho_1 = 1/2,\ \rho_2 = 1/4$			
n	exact	asympt.	rel. err.
0	0.50072	0.50000	.00144
1	0.32483	0.47283	.45562
2	0.12277	0.13460	.09629
3	0.03784	0.03831	.01251
4	0.01092	0.01090	.00156
5	.24230e-2	.94023e-2	2.8803
6	.41153e-3	.90872e-3	1.2081
7	.56957e-4	.87828e-4	.542002
8	.68217e-5	.84885e-5	.24434
9	.74209e-6	.82041e-6	.10555
10	.63341e-7	.65832e-7	.03933

Table 2

$L = 5,\ K = 15,\ \rho_1 = 1/2,\ \rho_2 = 1/4$			
n	exact	asympt.	rel. err.
0	0.50072	0.50000	0.00144
1	0.32483	0.47283	0.45562
2	0.12277	0.13460	0.09629
3	0.03784	0.03831	0.01251
4	0.01092	0.01090	0.00156
5	.24230e-2	.94023e-2	2.8803
6	.41153e-3	.90872e-3	1.2081
7	.56957e-4	.87828e-4	.542002
8	.68217e-5	.848857e-5	.24434
9	.74209e-6	.82041e-6	.10555
10	.76088e-7	.79293e-7	.042115
11	.75498e-8	.76636e-8	.015072
12	.73723e-9	.740688e-9	.004686
13	.71503e-10	.71587e-10	.001169
14	.69179e-11	.69188e-11	.000129
15	.55527e-12	.55518e-12	.000148

As we see from Table 1, the asymptotic expansion is quite accurate for $n = 0$, $n = 2 - 4$ and $n > 8$. It is not accurate for $n = 1$ and $n = 5 - 8$. This is consistent with our results in Theorem 5 since the asymptotic result (2.65) is valid for $n \gg 1$ and (2.63) is valid for $n - L \gg 1$. It is remarkable that our results are this accurate since n and $n - L$ are only moderate in size. We cannot expect the asymptotic solution to be accurate for $n = L$ since it assumes that $n - L \to \infty$. We have chosen values for L and K that are quite small. In reality, we would expect them to be of the order 10^2 (see [11]). For such large values of L and K, calculation of the exact solution would prove difficult while the evaluation of the asymptotic solution is straightforward.

In Table 2, we consider the same queue as in Table 1 but now with $K = 15$. The values for A and B are the same as for Table 1 since these constants are independent of L and K. We see that the exact and the asymptotic results are numerically close to those in the previous example for $n < 10$, as expected. For $n \geq 10$, the relative error starts at about 4% and rapidly decreases to under 1%.

The example in Table 3 is a queue with threshold $L = 8$ and capacity $K = 15$. The arrival rate is $\lambda = 1$ and the service times are $m_1 = 2$ and $M_1 = 1/2$ so that $\rho_1 = 2$ and $\rho_2 = 1/2$. In this example, $\rho_1 > 1$ so that in the absence of the threshold, the queue length distribution would be peaked near the capacity K. Now $A \approx -0.7968$ and $B \approx 2.512$. The asymptotic results are quite accurate when $n < L$. However, when $n \geq L$ the results are not accurate.

Table 3

$L = 8,\ K = 15,\ \rho_1 = 2,\ \rho_2 = 1/2$

n	exact	asympt.	rel. err.
0	.28305e-5	.28292e-5	.00044
1	.18084e-4	.18690e-4	.03351
2	.91796e-4	.91986e-4	.00206
3	.00045286	.00045271	.00033
4	.0022290	.0022280	.00044
5	.010970	.010965	.00044
6	.053991	.053967	.00044
7	.26572	.26560	.00044
8	.29180	4.4929	14.397
9	.19878	1.2789	5.4341
10	.10408	.36408	2.4980
11	.045665	.10364	1.2696
12	.017598	.02950	.67654
13	.0061526	.0083988	.36508
14	.0020014	.0023908	.19455
15	.00043941	.00047573	.0826589

Table 4

$L = 8,\ K = 25,\ \rho_1 = 2,\ \rho_2 = 1/2$

n	exact	asympt.	rel. err.
0	.28292e-5	.28292e-5	0.000003
1	.18076e-4	.18690e-4	0.033972
2	.91756e-4	.91986e-4	0.002507
3	0.000453	0.000453	0.000104
4	0.002228	0.002228	0.000002
5	0.010966	0.010966	0.000003
6	0.053968	0.053968	0.000003
7	0.265605	0.265604	0.000003
8	0.291673	4.492908	14.403916
9	0.198693	1.278988	5.436998
10	0.104037	0.364087	2.499576
11	0.045645	0.103644	1.270637
12	0.01759	0.029504	0.677286
13	0.00615	0.008399	0.365690
14	0.002001	0.002391	0.195078
15	0.000618	0.000681	0.100900
16	0.000185	0.000194	0.049898
17	.53894e-4	.55154e-4	0.023365
18	.15540e-4	.15700e-4	0.010298
19	.44505e-5	.44694e-5	0.004260
20	.12702e-5	.12723e-5	0.001652
21	.36196e-6	.36218e-6	0.000601
22	.10308e-6	.10310e-6	0.000205
23	.29348e-7	.29350e-7	0.000064
24	.83549e-8	.83550e-8	0.000017
25	.16624e-8	.16624e-8	0.000001

In Table 4 we retain $L = 8$ and increase K to 25. Now the asymptotic results are accurate to within 5% for $16 \leq n \leq 25$. In Tables 3 and 4, the results are not accurate for $L = 8 \leq n \leq 14$.

Apparently, $n - L = 6$ is too small a value for (2.63) to be useful, for these particular parameter values.

In Table 5, we consider the same case as in Table 3 except that now $M_1 = 2/3$ and hence $\rho_2 = 2/3$. Now we have $A \approx -0.7968$ and $B \approx 1.144$. The results are to within 7% if $n - L \geq 4$ (i.e. $12 \leq n \leq 15$). In Table 6 we increase K to 25. The error is at most 7% for all $n - L \geq 4$ (i.e. $12 \leq n \leq 25$).

Table 5

$L = 8,\ K = 15,\ \rho_1 = 2,\ \rho_2 = 2/3$

n	exact	asympt.	rel. err.
0	.21329e-5	.21219e-5	.0051
1	.13627e-4	.14017e-4	.02865
2	.69173e-4	.68989e-4	.00265
3	.00034125	.00033953	.00504
4	.0016796	.0016710	.00514
5	.0082667	.0082241	.00514
6	.040685	.040475	.00514
7	.20023	.19920	.00514
8	.25976	.85983	2.3100
9	.21133	.40103	.8976
10	.13486	.18704	.3869
11	.074723	.087239	.1675
12	.038095	.040690	.0680
13	.018535	.018978	.0239
14	.008803	.0088516	.0054
15	.0025867	.0025790	.00295

Table 6

$L = 8,\ K = 25,\ \rho_1 = 2,\ \rho_2 = 2/3$

n	exact	asympt.	rel. err.
0	.21219e-5	.21219e-5	0.000005
1	.13557e-4	.14017e-4	0.033970
2	.68817e-4	.68989e-4	0.002505
3	0.00034	0.00034	0.000103
4	0.001671	0.001671	0.000004
5	0.008224	0.008224	0.000005
6	0.040476	0.040476	0.000005
7	0.199204	0.199203	0.000005
8	0.258429	0.859833	2.327155
9	0.210249	0.401035	0.907433
10	0.134169	0.187047	0.394119
11	0.074339	0.087241	0.173549
12	0.0379	0.04069	0.073622
13	0.01844	0.018978	0.029203
14	0.008758	0.008852	0.010655
15	0.004114	0.004129	0.003551
16	0.001924	0.001926	0.001078
17	0.000897	0.000898	0.000297
18	0.000419	0.000419	0.000073
19	0.000195	0.000195	0.000014
20	.91124e-4	.91124e-4	0.000001
21	.42501e-4	.42501e-4	-
22	.19823e-4	.19823e-4	-
23	.92457e-5	.92457e-5	-
24	.43123e-5	.43123e-5	-
25	.12564e-5	.12564e-5	-

Our results should be very accurate for $K > 25$, but it then becomes difficult to evaluate the exact solution. These numerical comparisons show that the asymptotic approximations are quite robust, and are accurate even for relatively small values of L and K. Tables 4 and 6 show that when $n = 25 = K$, the two results agree to five decimal places. Loss rates can thus be calculated to a very high precision using our asymptotic formulas.

References

[1] L. ABOLNIKOV, J.H. DSHALALOW AND A. DUKHOVNY, On stochastic processes in a multilevel control bulk queueing system, *Stoch. Anal. and Appl.* 10, 1992, 155-179.

[2] B. D. CHOI AND D. I. CHOI, Queueing system with queue length dependent service times and its application to cell discarding scheme in ATM networks *IEE Proc.-Commun.* 143, Feb. 1996, 5-11.

[3] J.H. DSHALALOW, Queueing systems with state-dependent parameters, in: *Advances in Queueing* (ed. by J.H. Dshalalow), 61-116, CRC Press, Boca Raton, FL, 1995.

[4] D. FAKINOS, The G/G/1(LCFS/P) queue with service depending on queue size, *Eur. J. Oper. Res.* 59, 1992, 303-307.

[5] W. B. GONG, A. YAN AND C. G. CASSANDRAS, The M/G/1 queue with queue-length dependent arrival rate, *Stochastic Models* 8, 1992, 733-741.

[6] C.M. HARRIS, Some results for bulk-arrival queue with state-dependent service times, *Manage. Sci.* 16, 1970, 313-326.

[7] V.A. IVNITSKIY, A stationary regime of a queueing system with parameters dependent on the queue length and with nonordinary flow, *Eng. Cyb.* ,1974, 85-90.

[8] S. Q. LI, Overload control in a finite message storage buffer, *IEEE Trans. Commun.* 37, 1989, 1330-1338.

[9] K. SRIRAM, Methodologies for bandwidth allocation, transmission scheduling and congestion avoidance in broadband ATM networks, *Comput. Netw. ISDN Syst.* 26, 1993, 43-59.

[10] K. SRIRAM AND D. M. LUCANTONI, Traffic smoothing of bit dropping in a packet voice multiplexer, *IEEE Trans. Commun.* 37, 1989, 703-712.

[11] K. SRIRAM, R. S. MCKINNEY AND M. H. SHERIF, Voice packetisation and compression in broadband ATM networks, *IEEE J. Sel. Areas Commun.* 9(3), 1991, 294-304.

[12] H. TAKAGI, Analysis of a finite capacity M/G/1 queue with resume level, *Performance Evaluation* 5, 1985, 197-203.