CSCanada
Progress in Applied Mathematics
Vol. 2, No. 1, 2011, pp. 01-34
DOI: 10.3968/j.pam.1925252820110201.012

ISSN 1925-251X [Print]
ISSN 1925-2528 [Online]
www.cscanada.net
www.cscanada.org

On the Finite Capacity Shortest Queue Problem

Charles Knessl[1]; Haishen Yao[2,*]

[1]Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60607-7045
[2]Department of Mathematics and Computer Science, Queensborough Community College (QCC), City University of New York, 222-05 56th Avenue, Bayside, NY 11364
Supported by (in part) NSF grants DMS 02-02815, DMS 05-03745, NSA Grant H 98230-08-1-0102, PSC-CUNY Research Awards No. 61430-0039 and No. 62472-00 40
*Correponding author.
Address: Department of Mathematics and Computer Science, Queensborough Community College (QCC), City University of New York, 222-05 56th Avenue, Bayside, NY 11364. Email: Hyao@qcc.cuny.edu

Received 10 May 2011; accepted 16 May 2011

Abstract
We consider two parallel queues. There is one server tending to each queue and the capacity of each queue is K. The network is fed by a single Poisson arrival stream of rate λ, and the two servers are identical exponential servers working at rate μ. A new arrival is routed to the queue with the smaller number of customers. If both have the same number of customers then the arrival is routed randomly, with the probability of joining either queue being $1/2$. If there are more than $2K$ customers in the system, further arrivals are turned away and lost. We let $\rho = \lambda/\mu$ and take $K \to \infty$, and consider the cases $\rho < 2$, $\rho > 2$ and $\rho - 2 = O(K^{-1})$. We shall obtain asymptotic approximations to the joint steady state distribution of finding m customers in the first queue and n in the second. The asymptotic approximations are shown to be quite accurate numerically. We shall identify precisely for what ranges of m and n can the finite capacity model be approximated by the infinite capacity one. We will also show that the marginal distribution of finding n customers in the second queue undergoes a transition when $\rho = 4$.
Key words
Shortest queue problem; Finite capacity; Poisson arrival stream; Analytical approximations

Charles Knessl, & Haishen Yao (2011). On the Finite Capacity Shortest Queue Problem. *Progress in Applied Mathematics, 2*(1), 01-34. Available from: URL: http://www.cscanada.net/index.php/pam/article/view/j.pam.1925252820110201.012
DOI: http://dx.doi.org/10.3968/j.pam.1925252820110201.012

INTRODUCTION

The shortest queue (SQ) problem is one of the classic models within queueing theory. In its simplest form, there are two parallel queues with identical exponential servers and a single Poisson arrival stream. An arrival is routed to the queue with the smaller number of customers, and goes to either queue with

probability $1/2$ if they have the same numbers of customers. We denote the arrival rate by λ, the mean service time by $1/\mu$, and let N_j denote the number of customers in the j^{th} queue, for $j = 1,\ 2$.

There are many applications of this model, which include computer network communications, packet switched data networks [1, 2], and airports with two runways [3]. The first complete solution of the symmetric, infinite capacity SQ model was apparently obtained in [4], via the use of generating functions and complex variable arguments. Previously some asymptotic results were obtained in [5], and simple bounds are discussed in [6]. A more explicit form of the steady state probability distributions, $Pr[N_1 = m, N_2 = n]$, was later obtained by Adan et. al. [7], by using a "compensation approach". The solution is an infinite mixture of geometric terms of the form ${a_j}^m {b_j}^n,\ j \geqslant 0$, with the particular mixture determined by satisfying boundary conditions along $m = n$ and $n = 0$ (or $m = 0$). This compensation approach also works for the non-symmetric model, where the two servers work at different rates μ_1 and μ_2 [8]. Variants of the SQ problem are discussed in [9–15], which include models characterized by the unfinished work rather than the number of customers, and models with jockeying, where a customer may be switched from one queue to the other.

A natural variant of the symmetric SQ problem is a finite capacity version, where N_1 and N_2 are restricted to be at most K. If an arrival occurs with $N_1 = N_2 = K$, this customer is turned away and lost. Note that a loss can occur only if both queues are at capacity, in view of the SQ routing policy. The finite capacity model was used by Conolly [16] to model toll booths on the Italian autostrada network. The author analyzes the difference equations satisfied by the joint steady state distribution of (N_1, N_2) using a spectral expansion, but the spectral coefficients must be determined numerically.

Other recent work on SQ problems with finite capacity includes Tarabia [17, 18], who also allows for jockeying and different service rates in the two queues. In [17] matrix–analytic methods are used to obtain an explicit expression for $Pr[N_1 = m, N_2 = n]$, but its evaluation involves computing a product of a possibly large number of matrices. In [18] the transient solution of this model is analyzed numerically, using both Runge-Kutta and randomization methods.

Here we consider the symmetric finite capacity model in the asymptotic limit $K \to \infty$. Defining $\rho = \lambda/\mu$ we shall consider the cases $\rho < 2$, $\rho > 2$ and $\rho - 2 = O(K^{-1})$. Note that if $\rho < 2$ and $K = \infty$, we recover the standard stable infinite capacity SQ model. We shall obtain serveral different asymptotic formulas for the steady state distribution $p(m, n) = Pr[N_1 = m, N_2 = n]$, valid in different portions of the state space $\{(m, n)\colon 0 \leqslant m, n \leqslant K\}$. In view of the symmetry of the problem we can restrict ourselves to the lattice triangle $\{(m, n)\colon 0 \leqslant n \leqslant m \leqslant K\}$. By introducing the scaled variables $x = m/K$ and $y = n/K$, which correspond to the fractions of the capacities that are filled, we shall obtain serveral different asymptotic expansions over the triangle $\mathcal{T} = \{(x, y)\colon 0 \leqslant y \leqslant x \leqslant 1\}$. We will also identify precisely when the finite capacity model may be approximated by the infinite capacity model, which will turn out to be possible only over a subtriangle of $\mathcal{T}$. We will also see that in some cases the finite capacity SQ model may be approximated by the solution of the *longer queue* model, which was analyzed by Flatto in [19]. This model may be viewed as a "dual" of the SQ model, in that there are two identical parallel queues, each fed by its own Poisson arrival stream, but there is only one server that tends to the longer of the two queues. In most ranges of $\mathcal{T}$ we shall obtain simple analytical approximations to $p(m, n)$,which we shall show to be numerically accurate even for moderate values of the capacities K.

Our analysis is based on singular perturbation expansions of the basic difference equation(s) satisfied by $p(m, n)$, after introduction of the scaled variables $(x, y) = K^{-1}(m, n)$. The analysis does make some assumptions about the forms of various expansions, and the asymptotic matching between expansions in different portions of $\mathcal{T}$.

While we do not attempt to survey all previous work on SQ models, some numerical approaches are discussed in [20–22] while asymptotic approaches appear in [23–30]. The asymptotic analyses involve either diffusion models/approximations (in one [23] or two [24, 25] dimensions), tail behavior [26, 27], or large deviations theory [28–30]. The tail behavior is often obtained by identifying the dominant singularity of a generating function, but this assumes that a reasonably explicit expression can be obtained for this function. Large deviations theory typically characterizes the decay rate of $p(m, n)$ in terms of the solution of a variational problem. However, this theory seems to estimate only the growth rate, i.e., $\log[p(m, n)]$. Here we

shall obtain the full asymptotics of $p(m,n)$, giving explicitly the leading term and indicating how to compute correction terms. We have previously done similar analyses of SQ problems with an infinite number of servers [31, 32] and many server loss models [33].

The paper is organized as follows. In section 1. we state the basic equations and summarize our asymptotic results. In section 2. we derive the expansions in two main subsets of the triangle $\mathcal{T}$. In sections 3.–6. we analyze various boundary and corner regions of $\mathcal{T}$. Finally, in section 7. we compare asymptotic and numerical results, to assess the quality of the approximations.

1. PROBLEM STATEMENT AND SUMMARY OF RESULTS

The capacity will be denoted by the positive integer K. We let $N_j(t)$ be the number of customers in the j^{th} queue at time t, for $j = 1$ and $j = 2$. We clearly have $0 \leqslant N_j(t) \leqslant K$. The arrival rate will be denoted by λ and both servers work at rate μ. We then set $\rho = \lambda/\mu$ and consider the steady state limit, where $t \to \infty$. We thus define

$$p(m,n) = \lim_{t\to\infty} Pr[N_1(t)=m, N_2(t)=n | N_1(0)=m_0, N_2(0)=n_0], \tag{1.1}$$

which is independent of the initial numbers (m_0, n_0), and exists for all values of ρ. We have not been able to solve for $p(m,n)$ exactly, so we consider the large capacity limit, where K is large but finite.

The balance equations are

$$(\rho+2)p(m,n) = \rho p(m,n-1) + p(m+1,n) + p(m,n+1); \tag{1.2}$$

$$n+1 \leqslant m \leqslant K-1, n > 0,$$

$$\begin{aligned}(\rho+2)p(n+1,n) = {}& \frac{\rho}{2}p(n,n) + \rho p(n+1,n-1) \\ & + p(n+2,n) + p(n+1,n+1),\end{aligned} \tag{1.3}$$

$$0 < n, m = n+1, m \leqslant K-1,$$

$$\begin{aligned}(\rho+2)p(n,n) &= \rho p(n,n-1) + \rho p(n-1,n) \\ &\quad + p(n+1,n) + p(n,n+1) \\ &= 2\rho p(n,n-1) + 2p(n+1,n),\end{aligned} \tag{1.4}$$

$$0 < n, m = n, m \leqslant K-1.$$

In (1.4) we used the symmetry $p(m,n) = p(n,m)$, which allows us to consider the problem for $m \geqslant n$ only. When $n = 0$, the balance equation becomes

$$(\rho+1)p(m,0) = p(m+1,0) + p(m,1), \quad 2 \leqslant m \leqslant K-1. \tag{1.5}$$

We also have the two corner conditions

$$(\rho+1)p(1,0) = \frac{\rho}{2}p(0,0) + p(2,0) + p(1,1), \tag{1.6}$$

and

$$\rho p(0,0) = p(1,0) + p(0,1).$$

When $m = K$ we have

$$(\rho + 2)p(K,n) = \rho p(K,n-1) + p(K,n+1), \quad 1 \leqslant n \leqslant K-2, \tag{1.7}$$

and

$$(\rho + 1)p(K,0) = p(K,1). \tag{1.8}$$

Finally, at the "corner " (K,K), the balance equations are

$$\begin{aligned} 2p(K,K) &= \rho p(K,K-1) + \rho p(K-1,K) \\ &= 2\rho p(K,K-1), \end{aligned} \tag{1.9}$$

$$\begin{aligned} (\rho + 2)p(K,K-1) = \frac{\rho}{2} p(K-1,K-1) \\ + \rho p(K,K-2) + p(K,K). \end{aligned} \tag{1.10}$$

The normalization requirement is

$$\sum_{m=0}^{K}\sum_{n=0}^{K} p(m,n) = \sum_{n=0}^{K} p(n,n) + 2\sum_{m=1}^{K}\sum_{n=0}^{m-1} p(m,n) = 1. \tag{1.11}$$

In Figure 1 we sketch the transitions for $(N_1(t), N_2(t))$, which may be viewed as a state-dependent continuous time random walk on the lattice square $\{0,1,2,\ldots,K\} \times \{0,1,2,\ldots,K\}$. Alternately we can restrict to $m \geqslant n$ and view it as a random walk on a triangle, with a reflection law along the diagonal $m = n$ arising from the $m \leftrightarrow n$ symmetry. Then the transitions are indicated in Figure 2.

We introduce the scaled variables (x,y) with

$$x = \frac{m}{K} = \varepsilon m, \quad y = \frac{n}{K} = \varepsilon n, \tag{1.12}$$

where

$$\varepsilon \equiv \frac{1}{K} \to 0^+. \tag{1.13}$$

Because $m, n \leqslant K$ and we restrict to $m \geqslant n$, we must analyze the problem in the triangular domain $\mathcal{T}$

$$\mathcal{T} = \{(x,y)\colon\ 0 \leqslant y \leqslant x \leqslant 1\}. \tag{1.14}$$

The probability distribution will be concentrated near $(x,y) = (0,0)$ if $\rho < 2$. If $\rho > 2$, it will be concentrated near $(1,1)$, and if $\rho \approx 2$, it will be spread out along the line segment $x = y, 0 < x < 1$. In terms of the scaled variables (x,y) we let

$$p(m,n) = P(x,y) \tag{1.15}$$

and note that p and P depend also on $\varepsilon = 1/K$ and ρ.

Below we summarize our main asymptotic results. First, we find that the form of $P(x,y)$ is different in two main parts of the triangle $\mathcal{T}$. We thus divide the interior $\mathcal{T}_0$ of $\mathcal{T}$ into the two parts $\mathcal{D}$ and $\mathcal{R}$ with

$$\mathcal{T}_0 = \mathcal{D} \cup \mathcal{R} \cup \mathcal{L}_1, \tag{1.16}$$

where

$$\mathcal{D} = \{(x,y)\colon\ \max\{0, Y(x)\} < y < x < 1\}, \tag{1.17}$$

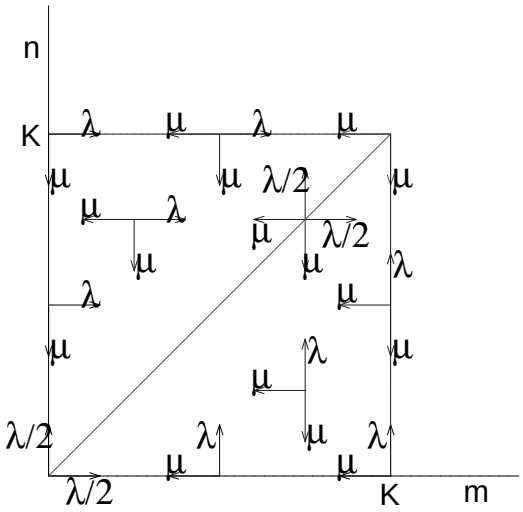

Figure 1
A Sketch of the Transition Rates for the Random Walk

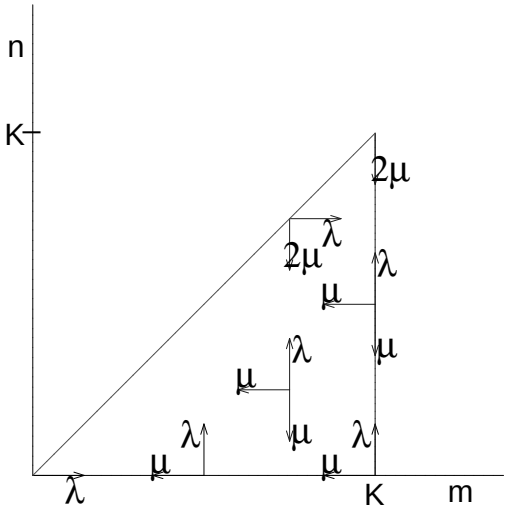

Figure 2
A Sketch of the Transition Rates for the Reflected Random Walk

$$\mathcal{R} = \{(x, y)\colon\ 0 < y < \max\{0, Y(x)\}, x < 1\}, \tag{1.18}$$

and $Y(x)$ is the line

$$Y(x) = \left(1 + \frac{4}{\rho} + \frac{16}{\rho^2}\right)x - \frac{4}{\rho} - \frac{16}{\rho^2}, \tag{1.19}$$

which lies inside the domain $\mathcal{T}$ for

$$X_0(\rho) < x < 1, \quad X_0(\rho) = \frac{16 + 4\rho}{16 + 4\rho + \rho^2}. \tag{1.20}$$

Then $\mathcal{L}_1$ is the curve separating $\mathcal{D}$ and $\mathcal{R}$, hence

$$\mathcal{L}_1 = \{(x, y)\colon y = Y(x), y > 0, x < 1\}.$$

We sketch the domains $\mathcal{D}$ and $\mathcal{R}$ in Figure 3. Note that along the separating curve we have $X_0(\rho) < x < 1$, and since $Y(X_0(\rho)) = 0$, the curve hits the boundary of $\mathcal{T}$ at the two points $(X_0(\rho), 0)$ and $(1, 1)$.

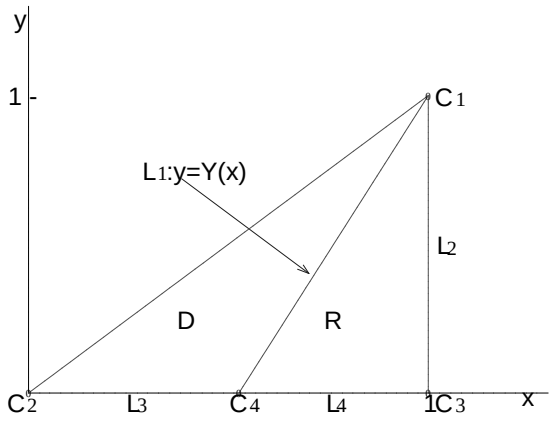

Figure 3
A Sketch of the Main Regions $\mathcal{D}$ and $\mathcal{R}$, as well as the Boundary and Corner Regions $\mathcal{L}_j$ and C_j

In region $\mathcal{D}$ we obtain the following

$$P(x, y) \sim C \exp\left[K\Phi(x, y)\right], \quad x > y, \tag{1.21}$$

where

$$\Phi(x, y) = x \log\left(\frac{\rho^2}{2(\rho + 4)}\right) + y \log\left(\frac{\rho + 4}{2}\right). \tag{1.22}$$

The constant $C = C(\rho, \varepsilon)$ has the following asymptotic forms:

$$\rho < 2\colon \quad C \sim \frac{(4 - \rho^2)(4 - \rho)}{\rho(4 + \rho)}, \tag{1.23}$$

$$\rho - 2 = \alpha\varepsilon = O(\varepsilon)\colon \quad C \sim \frac{2}{3}\frac{\alpha\varepsilon}{(e^{\alpha} - 1)}, \tag{1.24}$$

$$\rho > 2\colon \quad C \sim \left(\frac{2}{\rho}\right)^{2K} \frac{(\rho^2 - 4)}{(\rho + 4)}\frac{4}{\rho^2}. \tag{1.25}$$

By using $m = Kx$, $n = Ky$, another form of (1.21) is

$$p(m,n) \sim C\left(\frac{\rho^2}{2(\rho+4)}\right)^m \left(\frac{\rho+4}{2}\right)^n, \tag{1.26}$$

so that the distribution is approximately of product form. The expansion (1.21) applies only for $x > y$ ($m > n$) and cannot be used to compute the diagonal probabilities $p(n,n)$. To obtain the asymptotic form of $p(n,n)$, we can use (1.26) in (1.4) to get

$$p(n,n) = P(y,y) \sim C\frac{\rho}{\rho+2}\left(\frac{\rho}{2}\right)^{2n}, \quad m = n. \tag{1.27}$$

Next we consider region $\mathcal{R}$, where we obtain the approximation

$$P(x,y) \sim \widetilde{k}(\varepsilon)\widetilde{L}(x,y)\exp[K\Psi(x,y)], \tag{1.28}$$

where

$$\widetilde{k}(\varepsilon) = \frac{\sqrt{\varepsilon}}{\sqrt{2\pi}}\left(\frac{\rho}{2}\right)^{2K}\frac{\rho^2(\rho+4)}{4(\rho+2)}C, \tag{1.29}$$

$$\begin{aligned}\Psi(x,y) = &\left[\log(\rho+2-\frac{\rho}{u}-u)\right](\frac{\rho}{u}+u-\rho-2)\tau \\ &+ (\log u)(\frac{\rho}{u}-u)\tau,\end{aligned} \tag{1.30}$$

$$\widetilde{L}(x,y) = \frac{\widetilde{L}_0(u)}{\sqrt{\tau}}, \tag{1.31}$$

$$\widetilde{L}_0(u) = \frac{(u-1)(2u-\rho-2)}{\sqrt{(\rho+2)u-u^2-\rho}\,\sqrt{(\rho+2)u^2-4\rho u+\rho^2+2\rho}(2u-\rho-4)}, \tag{1.32}$$

and (u,τ) are related to (x,y) via the mapping

$$\begin{aligned}x &= (\frac{\rho}{u}+u-\rho-2)\tau+1, \\ y &= (\frac{\rho}{u}-u)\tau+1.\end{aligned} \tag{1.33}$$

When $\tau = 0$ we have $(x,y) = (1,1)$ so that all the curves in (1.33) are straight lines that emanate from the upper corner of the domain $\mathcal{T}$. Note that these curves fill $\mathcal{T}$ as u varies over the range

$$u_{\min} \leqslant u \leqslant u_{\max}, \tag{1.34}$$

where

$$u_{\min} = \frac{\rho+4}{2}, \quad u_{\max} = \frac{\rho+2+\sqrt{\rho^2+4}}{2}. \tag{1.35}$$

The curve in (1.33) with $u = u_{\min}$ is the same as $\mathcal{L}_1$ in (1.19). When u increases to $u_{\max}$ (1.33) corresponds to $x = 1$. For $u \in (u_{\min}, u_{\max})$ and

$$0 < \tau < \frac{u}{u^2-\rho}, \tag{1.36}$$

the curves in (1.33) fill the domain $\mathcal{R}$. When $\tau = u/(u^2-\rho)$ we have $y = 0$ and then (1.28) ceases to be valid, as discussed below. We can also invert explicitly the transformation in (1.33), and write (u,τ) in terms of (x,y) as

$$u = \frac{(1-y)(\rho+2)+\sqrt{(\rho+2)^2(1-y)^2+4\rho(y-x)(2-x-y)}}{2(2-x-y)}, \tag{1.37}$$

$$\tau = \frac{(1-x)(\rho+2) + \sqrt{(\rho+2)^2(1-y)^2 + 4\rho(y-x)(2-x-y)}}{\rho^2+4}. \tag{1.38}$$

Then we have the explicit expression

$$\begin{aligned}\Psi(x,y) = {} & (x-1)\log(1-x) \\ & + (y-1)\log\left[\frac{(1-y)(\rho+2) + \sqrt{(\rho+2)^2(1-y)^2 + 4\rho(y-x)(2-x-y)}}{2(2-x-y)}\right] \\ & + (1-x)\log\left[\frac{(1-x)(\rho+2) + \sqrt{(\rho+2)^2(1-y)^2 + 4\rho(y-x)(2-x-y)}}{\rho^2+4}\right]\end{aligned} \tag{1.39}$$

and $\widetilde{L}$ may be rewritten as

$$\begin{aligned}\widetilde{L}(x,y) = {} & \sqrt{\frac{\tau}{1-x}}\frac{(u-1)(2u-2-\rho)}{u(2u-4-\rho)} \\ & \times [(\rho+2)^2(1-y)^2 + 4\rho(y-x)(2-x-y)]^{-1/4}.\end{aligned} \tag{1.40}$$

From (1.39) we see that Ψ_x has a logarithmic singularity as $x \uparrow 1$, while from (1.40) it follows that $\widetilde{L}$ blows up like $(1-x)^{-1/2}$ as $x \uparrow 1$, and is also singular when $u \to u_{\min}$, since $2u_{\min} - 4 - \rho = 0$.

The expansions in (1.21) and (1.28) are valid in the interior $\mathcal{T}_0$ of $\mathcal{T}$, excluding the curve $\mathcal{L}_1$. After analyzing the singularities of these expansions, we found that different expansions must be constructed in various boundary and corner regions, which we summarize below.

The boundary regions are

$$\mathcal{L}_1 \colon y - Y(x) = O(\sqrt{\varepsilon}), \quad 0 < y < 1, \tag{1.41}$$

$$\mathcal{L}_2 \colon 1 - x = O(\varepsilon)\ (K - m = O(1)), \quad 0 < y < 1, \tag{1.42}$$

$$\mathcal{L}_3 \colon y = O(\varepsilon)\ (n = O(1)), \quad 0 < x < X_0(\rho) = \frac{16+4\rho}{16+4\rho+\rho^2}, \tag{1.43}$$

$$\mathcal{L}_4 \colon y = O(\varepsilon)\ (n = O(1)), \quad X_0(\rho) < x < 1. \tag{1.44}$$

The corner regions correspond to the scalings

$$C_1 \colon 1 - x = O(\varepsilon),\ 1 - y = O(\varepsilon)\ (K - m = O(1),\ K - n = O(1)), \tag{1.45}$$

$$C_2 \colon x = O(\varepsilon),\ y = O(\varepsilon)\ (m = O(1),\ n = O(1)), \tag{1.46}$$

$$C_3 \colon 1 - x = O(\varepsilon),\ y = O(\varepsilon)\ (K - m = O(1),\ n = O(1)), \tag{1.47}$$

$$C_4 \colon x - X_0(\rho) = O(\sqrt{\varepsilon}),\ y = O(\varepsilon)\ (n = O(1)). \tag{1.48}$$

Note that $\mathcal{L}_2$ and $\mathcal{L}_3 \cup \mathcal{L}_4$ correspond to boundaries of the triangle $\mathcal{T}$; C_1, C_2 and C_3 are the three corners of $\mathcal{T}$; and C_4 is where the transition layer $\mathcal{L}_1$ hits $y = 0$.

In the transition layer $\mathcal{L}_1$ we define

$$\eta = \frac{y - Y(x)}{\sqrt{\varepsilon}} = \sqrt{K}[y - Y(x)] = O(1). \tag{1.49}$$

Then the approximation to $p(m,n)$ has the form

$$p(m,n) \sim A^m B^n \bar{P}(x,\eta), \tag{1.50}$$

where

$$A = \frac{\rho^2}{2(\rho+4)},\ B = \frac{\rho+4}{2}, \tag{1.51}$$

$$\bar{P}(x,\eta) = \frac{C}{2\sqrt{a\pi}} \int_{-\infty}^{\eta/\sqrt{1-x}} \exp(-\frac{u^2}{4a})\, du, \tag{1.52}$$

$$a = \frac{(\rho+4)(\rho^3+6\rho^2+8\rho+32)}{\rho^4}. \tag{1.53}$$

In the boundary layer $\mathcal{L}_2$ we use the variables $l = K - m$ and $y = n\varepsilon$, with $0 < y < 1$, and obtain

$$\begin{aligned} p(m,n) \sim & \sqrt{\frac{2\pi}{\varepsilon}}\, \widetilde{k}(\varepsilon)\, \frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}\, \frac{\varepsilon^{-l}}{l!} \\ & \times \exp(\frac{F(y)}{\varepsilon})\, [D(y)]^l, \end{aligned} \tag{1.54}$$

where

$$F(y) = (y-1)\log\left(\frac{\rho+2+\sqrt{\rho^2+4}}{2}\right), \tag{1.55}$$

$$D(y) = \frac{1-y}{\sqrt{\rho^2+4}}. \tag{1.56}$$

Note also that $\exp[F(y)/\varepsilon] = [u_{\max}]^{n-K}$.

In the boundary layer $\mathcal{L}_3$ we use the original variables (m,n) and obtain

$$\begin{aligned} p(m,n) &\sim CA^m B^n \left[1 + \frac{2(4-\rho)}{(2+\rho)(4+\rho)}\left(\frac{4\rho}{(\rho+4)^2}\right)^n\right] \\ &= CA^m \left[\left(\frac{\rho+4}{2}\right)^n + \frac{2(4-\rho)}{(2+\rho)(4+\rho)}\left(\frac{2\rho}{\rho+4}\right)^n\right], \end{aligned} \tag{1.57}$$

but this applies only for $\varepsilon m < X_0(\rho)$. For $n \gg 1$ the second geometric term in (1.57) becomes negligible and (1.57) reduces to (1.26). Note also that when $\rho = 4$ the second geometric term is absent.

In the boundary layer $\mathcal{L}_4$ we use the variables x and n and obtain, for $X_0(\rho) < x < 1$,

$$\begin{aligned} p(m,n) \sim & \widetilde{k}(\varepsilon)\widetilde{L}(x,0)e^{K\Psi(x,0)} \\ & \times \left[e^{n\Psi_y(x,0)} + \frac{\rho e^{-\Psi_y(x,0)}-1}{1-e^{-\Psi_y(x,0)}}\rho^n e^{-n\Psi_y(x,0)}\right], \end{aligned} \tag{1.58}$$

where $\Psi(x,0)$ and $\Psi_y(x,0)$ can be computed from (1.39), with

$$\Psi_y(x,0) = \log\left[\frac{\rho+2+\sqrt{\rho^2+4+4\rho(1-x)^2}}{2(2-x)}\right]. \tag{1.59}$$

Next we give results for the four corner regions. The "corner" C_4 is where the expansions in (1.57) and (1.58) meet. Here we use the variables (ξ, n) where $\xi = [x - X_0(\rho)]/\sqrt{\varepsilon}$ and obtain

$$\begin{aligned} p(m,n) \sim & CA^m B^n \left[\frac{1}{\sqrt{2\pi}} \int_{\xi'}^{\infty} e^{-v^2/2}\, dv\right] \\ & \times \left[1 + \frac{2(4-\rho)}{8+6\rho+\rho^2}\left(\frac{4\rho}{(\rho+4)^2}\right)^n\right] \end{aligned} \tag{1.60}$$

where

$$\xi' = \frac{(\rho^2+4\rho+16)^{3/2}}{\sqrt{\rho^3+6\rho^2+8\rho+32}}\, \frac{1}{\rho\sqrt{2(\rho+4)}}\xi. \tag{1.61}$$

As ξ (hence ξ') $\to -\infty$ (1.60) reduces to (1.57), while as $\xi \to +\infty$ it can be shown that (1.60) asymptotically matches to (1.58), as $x \downarrow X_0(\rho)$.

The corner C_3 corresponds to $m = K - O(1)$ and $n = O(1)$, so that the first queue is at or near capacity while there are only a few customers in the second queue. Now we have

$$\begin{aligned} p(m,n) \sim \widetilde{k}(\varepsilon)\sqrt{2\pi K}\,\frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}\,\frac{K^l}{l!}\left(\frac{1}{\rho^2+4}\right)^{l/2} \\ \times\left(\frac{\rho+2+\sqrt{\rho^2+4}}{2}\right)^{n-K} \\ \times\left[1+\left(1+\frac{\rho^2}{2}-\frac{\rho}{2}\sqrt{\rho^2+4}\right)\frac{(4\rho)^n}{\left(\rho+2+\sqrt{\rho^2+4}\right)^{2n}}\right], \end{aligned} \tag{1.62}$$

where $l = K - m$ and $K = 1/\varepsilon$. It can be shown using Stirling's formula that (1.62) as $l \to \infty$ agrees with (1.58) as $x \uparrow 1$. Also, for $l = O(1)$ and $n \to \infty$ (1.62) agrees with (1.54) as $y \downarrow 0$.

For the corner layer C_2, where $m, n = O(1)$, $p(m,n)$ can be approximated by the exact solution to the infinite capacity SQ problem, which was obtained in [4] and later in a simpler form in [5]. We write this solution as an infinite mixture of exponentials as

$$p(m,n) \sim C\left[{a_0}^m {b_0}^n + \sum_{j=1}^{\infty} c_j {a_j}^m {b_j}^n\right], \quad m > n, \tag{1.63}$$

where C is given by (1.23)–(1.25) for the three cases of ρ, $a_0 = A$ and $b_0 = B$ are as in (1.51), and

$$\begin{aligned} a_{2N} = a_{2N+1}, \quad a_{2N+1}b_{2N+1} = a_{2N+2}b_{2N+2}, \quad N \geqslant 0, \\ b_{2N} = \frac{\rho+2}{2} + \frac{\sqrt{\rho^2+4}}{2} - \sqrt{\rho^2+4}\Big/\left[1+\frac{\sqrt{\rho^2+4}+2}{\sqrt{\rho^2+4}-2}\left(\frac{\rho+2+\sqrt{\rho^2+4}}{\rho+2-\sqrt{\rho^2+4}}\right)^N\right], \\ b_{2N-1} = \frac{\rho+2}{2} - \frac{\sqrt{\rho^2+4}}{2} + \sqrt{\rho^2+4}\Big/\left[1+\frac{\sqrt{\rho^2+4}+2}{\sqrt{\rho^2+4}-2}\left(\frac{\rho+2+\sqrt{\rho^2+4}}{\rho+2-\sqrt{\rho^2+4}}\right)^N\right]. \end{aligned} \tag{1.64}$$

From (1.64) we find that $a_1 = \rho^2/[2(\rho+4)]$, $b_1 = 2\rho/(\rho+4)$ and, generally, $b_{2N}b_{2N+1} = \rho$ ($N \geqslant 0$) and $b_{2N} + b_{2N-1} = \rho + 2$ ($N \geqslant 1$). Furthermore, the constants c_j in (1.63) can be obtained recursively from

$$\frac{c_{2N+1}}{c_{2N}} = -\frac{b_{2N+1}-1}{b_{2N}-1}, \quad \frac{c_{2N}}{c_{2N-1}} = -\frac{b_{2N}-1}{b_{2N-1}-1}\,\frac{2-b_{2N+1}}{2-b_{2N-2}} \tag{1.65}$$

with $c_0 = 1$. Finally, $p(n,n)$ can be obtained by using (1.63) in (1.4).

When $\rho < 2$ we compute C from (1.23) and then (1.63) corresponds to the *exact solution* of the symmetric infinite capacity ($K = \infty$) SQ problem. However, when $\rho > 2$, C is given asymptotically by (1.25) and then (1.63) is exponentially small for $K \to \infty$, of the order $(4/\rho^2)^K$. When $\rho - 2 = O(K^{-1})$ the constant C, and hence (1.63), is of order $O(K^{-1})$. Note also that if $\rho > 2$, the right side of (1.63) grows geometrically for $m = n \to \infty$, since $a_0 b_0 = \rho^2/4$. It is possible that the series in (1.63) truncates if $\rho > 2$, which occurs for example $\rho = 4$, as then $b_0 = 4$ so that $b_1 = 0$ and $c_j = 0$ for $j \geqslant 1$. For $\rho > 2$ and $\rho \sim 2$, (1.63) represents only a local approximation to $p(m,n)$, that applies for $m, n = O(1)$.

The final region is C_1, where (x,y) is close to $(1,1)$. We use the variables (l,k) where

$$m = K - l, \quad n = K - k \tag{1.66}$$

and set $p(m,n) \sim Q(l,k)$. We also note that $m > n$ corresponds to $k > l$. Then we define $b(w)$ and $f_0(w)$ by

$$b(w) = (2+\rho)\,w - \rho - w^2, \tag{1.67}$$

$$f_0(w) = \frac{(w-1)(2w-2-\rho)}{w(2w-4-\rho)} \tag{1.68}$$

and obtain the approximation in the form of contour integrals:

$$\begin{aligned} p(m,n) &= p(K-l,K-k) \sim Q(l,k) \\ &= \frac{-C_*}{2\pi i}\int_{\Gamma(w_0)} f_0(w)[b(w)]^{-l-1}w^{l-k}\,dw, k > l \geqslant 0, \end{aligned} \tag{1.69}$$

$$\begin{aligned} p(m,m) &= p(K-l,K-l) \sim Q(l,l) \\ &= \frac{-2C_*\rho}{(\rho+2)2\pi i}\int_{\Gamma(w_0)} \frac{f_0(w)}{w}[1+\frac{b(w)}{\rho}][b(w)]^{-l-1}\,dw, \quad l \geqslant 1, \end{aligned} \tag{1.70}$$

$$p(K,K) \sim Q(0,0) = \frac{C_*}{\rho}. \tag{1.71}$$

Here

$$C_* = \sqrt{\frac{2\pi}{\varepsilon}}\widetilde{k} = \left(\frac{\rho^2}{4}\right)^{K+1}\frac{\rho+4}{\rho+2}C, \tag{1.72}$$

in view of (1.29). In (1.69) and (1.70) the integration contour $\Gamma(w_0)$ is a small counterclockwise loop about $w = w_0$, where $b(w_0) = 0$ with

$$w_0 = w_0(\rho) = 1 + \frac{\rho}{2} + \sqrt{1+\frac{\rho^2}{4}}. \tag{1.73}$$

This completes the summary of the asymptotic results for $p(m,n)$. We will give the derivations in section 3–7.

We can also interpret our asymptotic results in terms of the exact solutions to the symmetric SQ model ([4], [5]) and the longer queue problem studied by Flatto in [19]. Let us denote the exact SQ solution by $p_{SQ}(m,n;\lambda,\mu)$, where $K = \infty$ and $\rho = \lambda/\mu < 2$. The longer queue model in [19] has two parallel queues fed by independent Poisson arrival streams of rate λ', and there is a single server that works at rate μ' and tends only to the longer of the two queues. We denote the solution in [19] by $p_{LQ}(m,n;\lambda',\mu')$ and note that the stability condition is $2\lambda' < \mu'$. Our asymptotic results show that the present finite capacity SQ problem, calling its distribution $p(m,n;\lambda,\mu;K)$, can be approximated by

$$p(m,n;\lambda,\mu;K) \approx p_{SQ}(m,n;\lambda,\mu), \tag{1.74}$$

for $K^{-1}(m,n)$ lying in region $\mathcal{D}$, if $\rho < 2$. This approximation applies also along $x = y < 1$, in the corner layer C_2, and in boundary layer $\mathcal{L}_3$. It becomes invalid in region $\mathcal{R}$ and in the boundary and corner regions that border it. Furthermore, (1.74) applies in $\mathcal{D}$ for $\rho > 2$ and $\rho \sim 2$, up to the multiplicative constant C (or $\widetilde{k}$ or C_*). Thus even if $\rho \geqslant 2$ the relative probabilities $p(m \pm 1, n \pm 1)/p(m,n)$ can be inferred from (1.74).

When $\rho > 2$ ($\rho' = \lambda'/\mu' < 1/2$) we have the approximation

$$p(m,n;\lambda,\mu;K) \approx p_{LQ}(K-m,K-n;\mu,\lambda), \tag{1.75}$$

which holds in the entirety of the triangle $\mathcal{T}_0 = \mathcal{D} \cup \mathcal{R} \cup \mathcal{L}_1$ and along $x = y > 0$. The approximation fails only when $n = O(1)$, which corresponds to the boundary layers $\mathcal{L}_3$ and $\mathcal{L}_4$, and corner layers C_2, C_3 and C_4. If $\rho < 2$ and $\rho \sim 2$ then (1.75) applies in $\mathcal{T}_0$ up to a multiplicative constant, and can be used to compute the relative probabilities for $\rho \leqslant 2$.

This discussion shows that (1.74) and (1.75) provide more uniform approximations, in that they apply over larger ranges of the state space, than the approximations we derived. However, the approximations

in (1.74) and (1.75) are quite complicated (cf. (1.63) and (1.69)) whereas the formulas we gave are much simpler, except near the two corners C_1 and C_2.

Finally we consider the marginal probability $\mathcal{P}(n)$ of finding n customers in the second queue. In view of the symmetry $p(m,n) = p(n,m)$ we have

$$\mathcal{P}(n) = \sum_{m=0}^{n-1} p(n,m) + p(n,n) + \sum_{m=n+1}^{K} p(m,n). \tag{1.76}$$

Using our asymptotic results for $p(m,n)$ we can easily obtain the asymptotics for $\mathcal{P}(n)$ for $K \to \infty$. The results are different for $n = O(1), n = Ky$ with $0 < y < 1$, and $n = K - O(1)$; furthermore, the cases $\rho < 4$, $\rho > 4$ and $\rho \sim 4$ lead to different results. We summarize the different cases below.

If $\rho < 4$ and $0 < n/K < 1$ we obtain

$$\mathcal{P}(n) \sim C\frac{2(\rho+4)}{(4-\rho)(\rho+2)}\left(\frac{\rho^2}{4}\right)^n, \tag{1.77}$$

where C is given by (1.23)–(1.25), according as $\rho < 2$, $\rho \sim 2$ or $2 < \rho < 4$. For $n = O(1)$ and $\rho < 4$ we use (1.63) (and a corresponding result for $m = n$) in (1.76), to find that

$$\begin{aligned}\mathcal{P}(n) \sim & C\sum_{j=0}^{\infty} c_j {a_j}^n {b_j}^n \left[\frac{1}{b_j - 1} + \frac{a_j}{1-a_j} + \frac{2}{\rho+2}\left(\frac{\rho}{b_j} + a_j\right)\right] \\ & + C\sum_{j=0}^{\infty} \frac{c_j}{1-b_j} {a_j}^n,\ n = O(1),\ n \geqslant 1.\end{aligned} \tag{1.78}$$

Here $c_0 = 1$ and the a_j, b_j and c_j can be evaluated using (1.64) and (1.65). Actually, (1.78) contains (1.77) as a special case, since for $n \to \infty$ only the first term in the first sum in (1.78) is important asymptotically. When $n = K - k = K - O(1)$ we must use (1.69)–(1.73) to evaluate (1.76), with the result that

$$\mathcal{P}(n) = \mathcal{P}(K-k) \sim \frac{-C_*}{2\pi i}\int_{\Gamma(w_0)} f_1(w)\,[b(w)]^{-k-1}\,dw, \quad k \geqslant 1, \tag{1.79}$$

where

$$f_1(w) = \frac{2w-2-\rho}{2w-4-\rho}\left[1 - \frac{2}{\rho+2}(w-1) + \frac{1}{w-\rho}\right],$$

and

$$\mathcal{P}(K) \sim C_*\left[\frac{1}{2} + \frac{1}{\rho} + \frac{\sqrt{\rho^2+4}}{2\rho}\right]. \tag{1.80}$$

For $k \to \infty$ the behavior of the contour integral in (1.79) is determined by the pole at $w = \rho$ if $\rho > 4$, and the pole at $w = (\rho+4)/2$ if $\rho < 4$. In the latter case the expansion of (1.79) as $k = K - n \to \infty$ reduces to (1.77), once we note that C and C_* are related by (1.72).

Now we consider $\rho > 4$. The evaluation of the right hand side of (1.76) is now different. For $\rho < 4$ the dominant terms had $m = n + O(1)$ but when $\rho > 4$ the second sum in (1.76) dominates, and the major contribution comes from where $m = K[\frac{1-y}{1-\rho} + 1]$, which is a line that is contained within the range $\mathcal{R}$ (see (1.18) and Figure 3). Asymptotically evaluating the second sum using (1.28) and Laplace's method leads to

$$\begin{aligned}\mathcal{P}(n) \sim & \sqrt{2\pi K}\,\widetilde{k}(\varepsilon)\frac{\rho-2}{\rho(\rho-4)}\exp[-K(1-y)\log\rho] \\ = & C_*\frac{\rho-2}{\rho-4}\rho^{n-K-1}.\end{aligned} \tag{1.81}$$

This applies for $K \to \infty$ and either $n \to \infty$ with $0 < n/K < 1$ or for $n = O(1)$ (as long as $n \geqslant 1$). For $n = K - O(1)$ the results in (1.79) and (1.80) apply, as these do not depend on whether $\rho < 4$ or $\rho > 4$. Note also that if we expand the contour integral in (1.79) for $k \to \infty$ we also obtain the expansion in (1.81), as now the pole at $w = \rho$ dominates, and $b(\rho) = \rho$.

We note that (1.77) and (1.81) both become singular as $\rho \to 4$ and a new expansion is needed. Now $p(m, n)$ will be approximately constant for $n = Ky$ and $m/K \in [y, (y + 2)/3]$. We introduce the parameter β by

$$\rho = 4 + \frac{\beta}{K} = 4 + \beta\varepsilon, \quad \beta = O(1). \tag{1.82}$$

Then for $K \to \infty$ and β fixed we find that

$$C \sim \frac{3}{8} e^{-\beta/2} 4^{-K} \tag{1.83}$$

and then

$$\mathcal{P}(n) \sim K\beta^{-1} e^{K(y-1)\log 4} e^{-\beta/2} \left\{ \exp\left[\frac{\beta(y+1)}{4}\right] - \exp\left[\frac{\beta y}{2}\right] \right\}. \tag{1.84}$$

This applies both for $n = O(1)$ (then we can replace y by n/K in the exponent and y by 0 in the term in (1.84) in braces) and for $n = Ky$ with $0 < y < 1$. For $n = K - O(1)$ (1.79) and (1.80) still hold, but can be simplified by replacing ρ by 4 in $f_1(w)$ and $b(w)$, and expanding C_* using (1.72) and (1.83), which leads to $C_* \sim 2e^{\beta/2}$. When $\rho = 4$, $f_1(w)$ has a double pole at $w = 4$ and this determines the asymptotic behavior of (1.79) as $k \to \infty$, and then (1.79) will asymptotically match to (1.84), as $y \to 1$. We can also show that for a fixed $0 < y < 1$, (1.84) will match to (1.77) as $\beta \to -\infty$, and to (1.81) as $\beta \to +\infty$.

2. RAY EXPANSIONS

With the scaling $(m, n) = (Kx, Ky)$ in (1.12) and with (1.15) the main balance equation (1.2) becomes

$$\begin{aligned}(\rho + 2)P(x, y) &= \rho P(x, y - \varepsilon) + P(x + \varepsilon, y) + P(x, y + \varepsilon), \\ &0 < y < x < 1.\end{aligned} \tag{2.1}$$

We can use (1.4) to eliminate the diagonal probabilities $p(n, n) = P(y, y)$ and $p(n + 1, n + 1) = P(y + \varepsilon, y + \varepsilon)$ in (1.3). Then (1.4) becomes

$$P(y, y) = \frac{2\rho}{\rho + 2} P(y, y - \varepsilon) + \frac{2}{\rho + 2} P(y + \varepsilon, y), \tag{2.2}$$

and (1.3) becomes

$$\begin{aligned}&(\rho+2)P(y+\varepsilon, y) = \rho P(y+\varepsilon, y-\varepsilon) + P(y+2\varepsilon, y) \\ &+ \frac{2\rho}{\rho+2} P(y+\varepsilon, y) + \frac{2}{\rho+2} P(y+2\varepsilon, y+\varepsilon) \\ &+ \frac{\rho}{2}\left[\frac{2}{\rho+2} P(y+\varepsilon, y) + \frac{2\rho}{\rho+2} P(y, y-\varepsilon)\right].\end{aligned} \tag{2.3}$$

We seek an asymptotic expansion of the form

$$P(x, y) = \varepsilon^{\nu_1} e^{\Phi(x,y)/\varepsilon} [L(x, y) + \varepsilon L^{(1)}(x, y) + O(\varepsilon^2)] \tag{2.4}$$

where ν_1 is a constant and $\varepsilon = K^{-1}$. Then (2.1) implies that Φ satisfies the "eiconal" equation

$$\rho + 2 = \rho e^{-\Phi_y} + e^{\Phi_x} + e^{\Phi_y} \tag{2.5}$$

and $L(x,y)$ satisfies the (linear) "transport" equation

$$\rho\left(\frac{\Phi_{yy}}{2}L - L_y\right)e^{-\Phi_y} + \left(\frac{\Phi_{xx}}{2}L + L_x\right)e^{\Phi_x} + \left(\frac{\Phi_{yy}}{2}L + L_y\right)e^{\Phi_y} = 0. \tag{2.6}$$

From (2.3) we can infer a boundary condition for Φ along $x = y$, namely

$$(\rho+2)e^{\Phi_x} = \rho e^{\Phi_x}e^{-\Phi_y} + e^{2\Phi_x} + \frac{3\rho}{\rho+2}e^{\Phi_x} + \frac{2}{\rho+2}e^{2\Phi_x}e^{\Phi_y} + \frac{\rho^2}{\rho+2}e^{-\Phi_y}, \quad x = y. \tag{2.7}$$

We discuss the condition for $L(x,y)$ along $x = y$ later. We solve the PDE (2.5) by the method of characteristic; the characteristic equations are

$$\frac{dx}{dt} = -e^{\Phi_x}, \quad \frac{dy}{dt} = \rho e^{-\Phi_y} - e^{\Phi_y}, \tag{2.8}$$

$$\frac{d\Phi}{dt} = -\Phi_x e^{\Phi_x} + \Phi_y\left(\rho e^{-\Phi_y} - e^{\Phi_y}\right), \tag{2.9}$$

$$\frac{d\Phi_x}{dt} = 0, \quad \frac{d\Phi_y}{dt} = 0. \tag{2.10}$$

Here t is a parameter along a characteristic curve, or ray, and a ray must enter the domain $x > y$ for $t > 0$. We choose $t = 0$ when the ray hits the line $x = y$, so that the "initial manifold" for the problem is $(x(s), y(s)) = (s, s)$, and a ray starts from the point (s, s) with $0 < s < 1$.

The ODEs in (2.10) are readily solved to give

$$\Phi_x = A_0(s), \quad \Phi_y = B_0(s), \tag{2.11}$$

where $A_0(s)$ and $B_0(s)$ are constants along a ray, and only depend on s. With (2.11) the equations in (2.8) become linear ODEs whose solutions are

$$x = -e^{A_0(s)}t + s, \quad y = (\rho e^{-B_0(s)} - e^{B_0(s)})t + s. \tag{2.12}$$

Here we imposed the conditions $(x,y) = (s,s)$ at $t = 0$. Using (2.11) and (2.12) in (2.9), we obtain

$$\frac{d\Phi}{dt} = -A_0(s)e^{A_0(s)} + B_0(s)(\rho e^{-B_0(s)} - e^{B_0(s)}), \tag{2.13}$$

whose solution is

$$\Phi(s,t) = \left[-A_0(s)e^{A_0(s)} + B_0(s)(\rho e^{-B_0(s)} - e^{B_0(s)})\right]t + C_0(s). \tag{2.14}$$

We need to determine A_0, B_0 and C_0, which are all functions of s. We have the strip condition

$$\Phi_s = \Phi_x x'(s) + \Phi_y y'(s), \quad t = 0, \tag{2.15}$$

and thus

$$C_0'(s) = A_0(s) + B_0(s). \tag{2.16}$$

Using (2.5) at $t = 0$ we obtain

$$\rho + 2 = \rho e^{-B_0(s)} + e^{A_0(s)} + e^{B_0(s)}. \tag{2.17}$$

We solve (2.17) together with (2.7) and obtain

$$A_0(s) = \log\left[\frac{\rho^2}{2(\rho+4)}\right], \quad B_0(s) = \log\left[\frac{\rho+4}{2}\right], \tag{2.18}$$

so that A_0 and B_0 are constants. Then we define A, B by $A = e^{A_0} = \rho^2/[2(\rho+4)]$, $B = e^{B_0} = (\rho+4)/2$. Solving (2.16) and letting $C_0(0) = 0$, we obtain

$$C_0(s) = (A_0 + B_0)s. \tag{2.19}$$

Note that $C_0(0)$ can be absorbed into a normalization constant. Thus,

$$\Phi(s,t) = A_0[-e^{A_0}t + s] + B_0[(\rho e^{-B_0} - e^{B_0})t + s]. \tag{2.20}$$

Using (2.12) in (2.20), we can also write Φ in terms of (x,y) as

$$\Phi = A_0 x + B_0 y, \tag{2.21}$$

so that $\exp(\varepsilon^{-1}\Phi) = e^{A_0 m}e^{B_0 n} = A^m B^n$ is of product form.

The rays in (2.12) are thus given by

$$x = -\frac{\rho^2}{2(\rho+4)}t + s, \quad y = (\frac{2\rho}{\rho+4} - \frac{\rho+4}{2})t + s. \tag{2.22}$$

For $s \in (0,1)$ and $0 < t < s(\rho+4)/(\rho^2+4\rho+16)$ the rays fill the domain $\mathcal{D}$ (cf. Figure 3).

Using (2.21), (2.18) and (2.22) to simplify (2.6) we obtain

$$L_t = 0. \tag{2.23}$$

Thus $L(x,y)$ only depends upon s, and we write

$$L(x,y) = L(s). \tag{2.24}$$

By considering the higher order terms in the expansion of (2.3) using (2.4), we can derive a boundary condition for $L(x,y)$ along $x = y$, and this will show that L is not just a constant along a ray, but globally a constant in the region $\mathcal{D}$. Summarizing our results we have shown that rays from $0 < x = y < 1$ lead to the asymptotic solution

$$P(x,y) \sim C\left(\frac{\rho^2}{2(\rho+4)}\right)^m \left(\frac{\rho+4}{2}\right)^n. \tag{2.25}$$

Since C may depend upon $\varepsilon = K^{-1}$, we can set $\nu_1 = 0$ in (2.4).

The determination of the constant C is different according to whether $\rho < 2, \rho \sim 2$ or $\rho > 2$. If $\rho < 2$ and $K \to \infty$, most of the probability mass is concentrated in the range $m, n = O(1)$. There we can approximate the solution by the solution to the stable, infinite capacity SQ problem, as given by (1.63). Then, taking into account the different form of $p(m,n)$ along $m = n$, we can determine C by normalizing (1.63), and this leads to (1.23). When $\rho \sim 2$ the probability mass is spread evenly along $0 < x < 1$, but with $x - y = O(\varepsilon)$ (i.e., $m - n = O(1)$). For $\rho - 2 = O(\varepsilon) = \alpha\varepsilon$, we can approximate (2.25) by

$$P(x,y) \sim C\left(\frac{\rho}{2}\right)^{m+n}\left(\frac{\rho}{\rho+4}\right)^{m-n} \sim Ce^{\alpha x}\left(\frac{1}{3}\right)^L, L \geqslant 1, \tag{2.26}$$

where $L = m-n = (x-y)/\varepsilon = O(1)$ and $0 < x < 1$. From (1.27) we obtain the corresponding approximation for the diagonal probabilities as

$$P(y,y) \sim \frac{1}{2}Ce^{\alpha y}, \quad 0 < y < 1. \tag{2.27}$$

By using (2.26) and (2.27) in the normalization condition (1.11) we obtain

$$\frac{1}{2}CK\int_0^1 e^{\alpha y}dy + 2\sum_{L=1}^{\infty} CK\int_0^1 e^{\alpha x}3^{-L}dx \sim 1,$$

which leads to the expression for C in (1.24). To determine C in the case $\rho > 2$ we must carefully analyze the scale $K - m, K - n = O(1)$, which is done in section 4. In this case both queues tend to be filled nearly to their capacities.

Note that the rays in (2.22) will not fill the entire domain $0 \leqslant y \leqslant x \leqslant 1$. Thus we need to consider the region $\mathcal{R}$, which is a "shadow"region for the rays from $x = y$.

To construct the asymptotic solution in region $\mathcal{R}$ we use the ansatz

$$p(m,n) = P(x,y) \sim \widetilde{k}(\varepsilon)\widetilde{L}(x,y)\exp\left[\frac{\Psi(x,y)}{\varepsilon}\right] \tag{2.28}$$

where $\widetilde{k}$ is a multiplicative constant, whose inclusion allows use to define an additive constant that arises in Ψ in a convenient form.

Using (2.28) in (2.1) we find that $\Psi(x,y)$ satisfies the same PDE as $\Phi(x,y)$ in (2.5). The characteristic equations are thus given by

$$\frac{dx}{d\tau} = -e^{\Psi_x}, \quad \frac{dy}{d\tau} = \rho e^{-\Psi_y} - e^{\Psi_y}, \tag{2.29}$$

$$\frac{d\Psi}{d\tau} = -\Psi_x e^{\Psi_x} + \Psi_y\left(\rho e^{-\Psi_y} - e^{\Psi_y}\right), \tag{2.30}$$

$$\frac{d\Psi_x}{d\tau} = 0, \quad \frac{d\Psi_y}{d\tau} = 0. \tag{2.31}$$

Here we use τ as a parameter along a ray, to distinguish the rays in $\mathcal{R}$ from those in $\mathcal{D}$. Solving the differential equations in (2.31) leads to

$$\Psi_x = A_1, \quad \Psi_y = B_1 \tag{2.32}$$

where A_1 and B_1 are constant along a ray. We now use rays that emanate from the corner point $(x,y) = (1,1)$. Thus at $\tau = 0$ we have $x = 1$ and $y = 1$. Then we use (2.32) in (2.29) and solve the latter equation(s) to obtain

$$x = -e^{A_1}\tau + 1, \quad y = (\rho e^{-B_1} - e^{B_1})\tau + 1. \tag{2.33}$$

Now, Ψ satisfies (2.5) and at $\tau = 0$ this yields, using (2.32) and (2.33),

$$\rho + 2 = \rho e^{-B_1} + e^{A_1} + e^{B_1}. \tag{2.34}$$

Let us set $e^{B_1} = u$ and use u to index the rays, so that (2.34) becomes

$$e^{A_1} = \rho + 2 - \frac{\rho}{u} - u, \tag{2.35}$$

and the rays (in parametric form) are as in (1.33).

Now consider the permissible range of the parameter u. From (2.33) and (2.35) we find that

$$\left.\frac{dy}{dx}\right|_{\tau=0} = \frac{\rho - u^2}{u^2 - (\rho+2)u + \rho}, \tag{2.36}$$

which gives the slope at which the ray indexed by u enters the corner point $(1,1)$. When $u = (\rho+2)/2$ this slope is $\dfrac{dy}{dx} = 1$. As $u \to u_{\max}$, where $u_{\max} = (\rho + 2 + \sqrt{\rho^2+4})/2$ (cf. (1.35)), $\dfrac{dy}{dx}|_{\tau=0} \to \infty$, so that the ray is locally vertical. When $u = u_{\min}$, where $u_{\min} = (\rho+4)/2$ (cf.(1.35)), the ray becomes the same as the curve $y = Y(x)$ in (1.19), that separates the regions $\mathcal{D}$ and $\mathcal{R}$.

Using (2.33), (2.35), and (2.32) in (2.30), we solve the resulting ODE and obtain (1.30). Then (1.39) gives Ψ explicitly in terms of (x,y). Here we have chosen $\Psi(1,1) = 0$, since we have yet to determine $\widetilde{k}$ in (2.28).

Now, $\widetilde{L}(x,y)$ satisfies the transport equation in (2.6) with $\Phi(x,y)$ replaced by $\Psi(x,y)$, and t replaced by τ. We first compute the Jacobian of the mapping in (1.33), to obtain

$$\mathcal{J} = \mathcal{J}(\tau,u) = \frac{\tau}{u^2}\left[(\rho+2)u^2 - 4\rho u + \rho^2 + 2\rho\right]. \tag{2.37}$$

Using (2.37) and (1.30) in (2.6), we obtain, after some simplification

$$\frac{\widetilde{L}_\tau}{\widetilde{L}} = -\frac{1}{2\tau}, \tag{2.38}$$

whose general solution is

$$\widetilde{L}(x,y) = \frac{\widetilde{L}_0(u)}{\sqrt{\tau}}, \tag{2.39}$$

with u, τ given in (1.37) and (1.38). Note that (2.39) has the same form as (1.31). We will determine $\widetilde{k}(\varepsilon)$ and $\widetilde{L}_0(u)$ in section 3., by asymptotic matching considerations.

Note that for $\frac{1}{2}(\rho+2) < u < u_{\max}$ the rays from the corner point $(x,y) = (1,1)$ fill the entire triangle $\mathcal{T}$, and for $\frac{1}{2}(\rho+2) < u < u_{\min}$ the rays fill the subtriangle $\mathcal{D}$. Thus in $\mathcal{D}$ we have two ray expansions. However, after we relate the constants C and $\widetilde{k}$ we will see that in $\mathcal{D}$ (2.28) leads to a solution that is asymptotically smaller than (2.25). Hence in $\mathcal{D}$ the product form approximation in (2.25) dominates, and the rays from the corner are only important if $u_{\min} < u < u_{\max}$, which corresponds to region $\mathcal{R}$. We also note that for each u in this range, we reach a critical value of τ where the ray hits $y = 0$ and exits the domain.

3. CORNER REGION NEAR (1, 1)

We consider the vicinity of the point $(x,y) = (1,1)$; this region we previously denoted by C_1. We use the local variables l and k, with

$$l = K - m = O(1), \quad k = K - n = O(1), \tag{3.1}$$

and note that $m \geqslant n$ implies that $k \geqslant l$. In terms of x and y we have

$$x = 1 - \varepsilon l, \quad y = 1 - \varepsilon k. \tag{3.2}$$

We then define $Q(l,k)$ by

$$p(m,n) = P(x,y;\varepsilon) = Q(l,k;\varepsilon) = Q(l,k)\left[1 + o(1)\right]. \tag{3.3}$$

Using (3.3) in (1.2) and using the symmetric property $Q(l,k) = Q(k,l)$ we obtain

$$(1+\frac{2}{\rho})Q(l,k) = Q(l,k+1) + \frac{1}{\rho}Q(l-1,k) + Q(l,k-1), \tag{3.4}$$
$$k > l+1.$$

For $k = l+1$, from (1.3) we obtain

$$\begin{aligned}(1+\frac{2}{\rho})&Q(k-1,k) \\ &= \frac{1}{2}Q(k,k) + Q(k-1,k+1) \\ &\quad + \frac{1}{\rho}[Q(k-2,k) + Q(k-1,k-1)], \quad k = l+1,\end{aligned} \tag{3.5}$$

and for $k = l$ (1.4) implies that

$$(1+\frac{2}{\rho})Q(k,k) = Q(k,k+1) + Q(k+1,k) \tag{3.6}$$
$$+\frac{1}{\rho}Q(k-1,k) + \frac{1}{\rho}Q(k,k-1)$$
$$= 2[Q(k,k+1) + \frac{1}{\rho}Q(k-1,k)], \quad k \geqslant 1.$$

We also obtain a boundary condition along $l = 0$, which follows from (1.7),

$$(1+\frac{2}{\rho})Q(0,k) = Q(0,k+1) + \frac{1}{\rho}Q(0,k-1), \quad k \geqslant 2. \tag{3.7}$$

From (1.9) and (1.10) we obtain the corner conditions

$$\frac{1}{\rho}Q(0,0) = Q(0,1), \tag{3.8}$$

and

$$(1+\frac{2}{\rho})Q(0,1) = \frac{1}{2}Q(1,1) + Q(0,2) + \frac{1}{\rho}Q(0,0). \tag{3.9}$$

These equations are the same as those for Flatto's longer queue model, studied in [19]. Here we would identify $1/\rho$ with the arrival rate and set the service rate equal to one. Then we use the results in [33] to obtain (1.69)–(1.71). Note that the longer queue model is stable only if $\rho > 2$, but (1.69)–(1.71) satisfy (3.4)–(3.9) for any value of $\rho > 0$. When $\rho > 2$, the probability mass concentrates near $(x, y) = (1, 1)$, so that we apply the normalization condition (1.11) to the solution $Q(l, k)$ that is valid in this range. To leading order this gives

$$Q(0,0) + \sum_{l=1}^{\infty} Q(l,l) + 2\sum_{l=0}^{\infty}\sum_{k=l+1}^{\infty} Q(l,k) = 1, \quad \rho > 2 \tag{3.10}$$

After some calculations, using (1.69)–(1.71) in (3.10), we obtain

$$C_* = \rho - 2, \quad \rho > 2 \tag{3.11}$$

We relate the constants C and C_*, by asymptotically matching the expansions in (1.21) and (1.70). From (1.69) we have, as $k, l \to \infty$ with $1 < k/l < 1 + \frac{4}{\rho} + \frac{16}{\rho^2}$,

$$Q(l,k) \sim C_*\frac{\rho+2}{\rho+4}\,\frac{4}{\rho^2}\left[\frac{2(\rho+4)}{\rho^2}\right]^l\left(\frac{2}{\rho+4}\right)^k. \tag{3.12}$$

For this range of k/l the asymptotics of (1.69) are determined by the pole of $f_0(w)$ in (1.68) at $w = \frac{1}{2}(\rho+4)$. Hence, (3.12) should agree with (1.21) as $x \to 1$, $y \to 1$. In fact, using (1.26) and setting $m = K - l$, $n = K - k$, (1.21) becomes

$$p(m,n) \sim C\left[\frac{\rho^2}{2(\rho+4)}\right]^{K-l}\left(\frac{\rho+4}{2}\right)^{K-k}. \tag{3.13}$$

By comparing (3.12) and (3.13), we see that the asymptotic matching is possible, provided that C and C_* are related by

$$C\left(\frac{\rho}{2}\right)^{2K} = C_*\frac{4(\rho+2)}{\rho^2(\rho+4)}, \tag{3.14}$$

which leads to (1.72). From (3.11) and (3.14), thus obtain C (cf. (1.25)) for $\rho > 2$. For $\rho < 2$ and $\rho \sim 2$ we determined C in section 3, so that C_* can be obtained from (3.14).

Next we asymptotically match (1.28) and (1.69), i.e., we will match the expansions in region $\mathcal{R}$ and in the corner layer C_1. This will determine $\widetilde{k}$ and $\widetilde{L}_0(u)$, and thus complete the determination of the ray expansion in region $\mathcal{R}$.

Note that the slope of the line $y = Y(x)$ that separates the regions $\mathcal{D}$ and $\mathcal{R}$ is $1 + \dfrac{4}{\rho} + \dfrac{16}{\rho^2}$, so that we must consider the limit $k, l \to \infty$ with $1 + \dfrac{4}{\rho} + \dfrac{16}{\rho^2} < \dfrac{k}{l} < \infty$. From (1.69) we find that

$$Q(l,k) \sim \frac{C_*}{\sqrt{2\pi}} w_s^{l-k} [b(w_s)]^{-l} \mathcal{M}(k,l), \tag{3.15}$$

where

$$w_s = w_s\left(\frac{k}{l}\right) = \frac{(\rho+2)k + \sqrt{(\rho^2+4)k^2 + 4\rho l^2}}{2(k+l)}, \tag{3.16}$$

$$b(w_s) = \frac{l}{2(k+l)^2}\Big[(\rho^2+4)k - 4\rho l + (\rho+2)\sqrt{(\rho^2+4)k^2 + 4\rho l^2}\Big], \tag{3.17}$$

$$\mathcal{M}(k,l) = \frac{1}{[(\rho^2+4)k^2 + 4\rho l^2]^{1/4}} \frac{(w_s - 1)\left(w_s - 1 - \dfrac{\rho}{2}\right)}{\sqrt{w_s b(w_s)}\left(w_s - 2 - \dfrac{\rho}{2}\right)}. \tag{3.18}$$

For this range the asymptotics of (1.69) are determined by a saddle point at $w = w_s(k/l)$, which is obtained by solving

$$\frac{d}{dw}\left[-l\log[b(w)] + (l-k)\log w\right] = -l\frac{b'(w)}{b(w)} + \frac{l-k}{w} = 0.$$

We then need to show that

$$\widetilde{k}\,\widetilde{L}(x,y) \exp\left[\frac{\Psi(x,y)}{\varepsilon}\right]\Big|_{(x,y)\to(1,1)} \sim Q(l,k)|_{l,k\to\infty}, \tag{3.19}$$

for $\dfrac{k}{l} = \dfrac{1-y}{1-x} > 1 + \dfrac{4}{\rho} + \dfrac{16}{\rho^2}$, which means that we are within region $\mathcal{R}$.

To evaluate the left side of (3.19) we recall that as $(x,y) \to (1,1)$ we have $\tau \to 0$ and by (1.30) and (1.33) $\exp[K\Psi(x,y)]$ becomes

$$\begin{aligned}\exp[K\Psi(x,y)] &= \exp\left[K\left((x-1)\log(\rho+2-\frac{\rho}{u}-u) + (y-1)\log(u)\right)\right] \\ &= \exp\left[(m-K)\log(\rho+2-\frac{\rho}{u}-u) + (n-K)\log u\right] \\ &= \left(\frac{1}{u}\right)^{k-l}\Big[(\rho+2)u - \rho - u^2\Big]^{-l}.\end{aligned} \tag{3.20}$$

Using (1.37) to express u in terms of x and y and then comparing (3.20) to (1.39) (or (1.30)) we find that

$$K\Psi(x,y) = (l-k)\log(w_s) - l\log[b(w_s)],$$

where we also used (3.16) and (3.17). We also note that $(\rho+2)u - \rho - u^2 = b(w_s)$ and $u = w_s(k/l)$. Hence the exponentially varying parts of (3.19) agree precisely.

To complete the matching verification we must show that

$$\widetilde{k}\,\widetilde{L}(x,y)\Big|_{(x,y)\to(1,1)} \sim \frac{C_*}{\sqrt{2\pi}}\mathcal{M}(k,l). \tag{3.21}$$

From (1.33) we obtain

$$\tau = \frac{1-x}{\rho+2-\frac{\rho}{u}-u} = \frac{\varepsilon l}{\rho+2-\frac{\rho}{u}-u}, \tag{3.22}$$

and

$$\frac{k}{l} = \frac{1-y}{1-x} = \frac{\rho-u^2}{\rho+u^2-(\rho+2)u}. \tag{3.23}$$

Using (1.31), (2.39), (3.18), (3.22) and (3.23), after some computations we obtain

$$\widetilde{k}(\varepsilon)\widetilde{L}_0(u) = \frac{C_*}{\sqrt{2\pi}} \frac{\sqrt{\varepsilon l}}{\sqrt{\rho+2-\frac{\rho}{u}-u}} \frac{1}{\left[(\rho^2+4)k^2+4\rho l^2\right]^{1/4}} \tag{3.24}$$

$$\times \frac{(w_s-1)}{\sqrt{w_s b(w_s)}} \frac{w_s-1-\frac{\rho}{2}}{w_s-2-\frac{\rho}{2}}.$$

We have thus determined $\widetilde{L}_0(u)$ and $\widetilde{k}$ as

$$\widetilde{L}_0(u) = \frac{\sqrt{u}}{\sqrt{(\rho+2)u-\rho-u^2}} \left[(\rho^2+4)v^2+4\rho\right]^{-1/4} \tag{3.25}$$

$$\times \frac{(u-1)}{\sqrt{u[(\rho+2)u-\rho-u^2]}} \frac{u-1-\frac{\rho}{2}}{u-2-\frac{\rho}{2}},$$

where

$$v = \frac{\rho-u^2}{\rho+u^2-(\rho+2)u} = \frac{k}{l}, \tag{3.26}$$

and

$$\widetilde{k}(\varepsilon) = \frac{C_*\sqrt{\varepsilon}}{\sqrt{2\pi}}. \tag{3.27}$$

After some simplification (3.25) becomes the same as (1.32). Finally, $\widetilde{k}$, C and C_* are related in (1.29) and (1.72).

We comment that in imposing the matching condition in (3.19), we expanded $Q(l,k)$ for $l,k \to \infty$ by the saddle point method, but expanding the left side as $(x,y) \to (1,1)$ amounted to doing nothing. This is because Ψ and $\widetilde{L}$ have the forms $\Psi = (1-x)\Psi_*\left(\frac{1-y}{1-x}\right)$ and $\widetilde{L} = (1-x)^{-1/2}L_*\left(\frac{1-y}{1-x}\right)$. This shows that the corner approximation $p(m,n) \sim Q(l,k)$ is actually uniformly valid in the entire triangle $\mathcal{T}$, with the exception of $y \approx 0$ which we discuss later. But, the ray approximations in the regions $\mathcal{D}$ and $\mathcal{R}$ lead to much simpler results.

4. TRANSITION LAYER EXPANSION

The expansions in region $\mathcal{D}$ (cf. (1.21)) and region $\mathcal{R}$ (cf. (1.28)) do not agree along the common ray $y = Y(x)$ (cf. (1.19)). We can show that

$$\Phi(x,Y(x)) - 2\log(\frac{\rho}{2}) = \Psi(x,Y(x)),$$

so that the exponential orders of magnitude of $Ce^{K\Phi}$ do agree with those of $C_*e^{K\Psi}$ along the common ray. But, the remaining factors in (1.21) and (1.28) do not agree, and in fact $\widetilde{L}(x,y)$ becomes singular along

$y = Y(x)$ (then $u \to \frac{1}{2}(\rho+2) = u_{\min}$ in (1.32)). Thus, there is a transition layer between $\mathcal{D}$ and $\mathcal{R}$. We consider the vicinity of the ray that separates $\mathcal{D}$ and $\mathcal{R}$, a region we called $\mathcal{L}_1$.

We introduce the scaling

$$\eta = \frac{y - Y(x)}{\sqrt{\varepsilon}} = \sqrt{K}(y - Y(x)) = O(1) \tag{4.1}$$

and set

$$p(m,n) = A^m B^n \overline{P}(x,\eta), \tag{4.2}$$

where A, B are as in (1.51).

We then use (4.1) and (4.2) in (1.2) to obtain

$$\rho B^{-1}(-\varepsilon\partial_y + \frac{\varepsilon^2}{2}\partial_y{}^2 + \cdots)\overline{P} + A(\varepsilon\partial_x + \frac{\varepsilon^2}{2}\partial_x{}^2 + \cdots)\overline{P} + B(\varepsilon\partial_y + \frac{\varepsilon^2}{2}\partial_y{}^2 + \cdots)\overline{P} = 0. \tag{4.3}$$

Here we used $\rho + 2 = \rho/B + B + A$ and expanded $P(x+\varepsilon, y)$ as $P + \varepsilon P_x + \frac{1}{2}\varepsilon^2 P_{xx} + O(\varepsilon^3)$, etc. We change variables to (x,η) noting that

$$\partial_x \to \partial_x - \frac{1}{\sqrt{\varepsilon}}Y'(x)\partial_\eta, \quad \partial_y \to \frac{1}{\sqrt{\varepsilon}}\partial_\eta. \tag{4.4}$$

After rewriting (4.3) using (4.4), at the $O(\sqrt{\varepsilon})$ order we obtain

$$-\rho B^{-1} - A \cdot Y'(x) + B = 0, \tag{4.5}$$

but this holds automatically in view of (1.19) and (1.51). At the next order ($O(\varepsilon)$) we obtain the parabolic PDE

$$A\overline{P}_x + \frac{1}{2}\left[A(Y'(x))^2 + \rho B^{-1} + B\right]\overline{P}_{\eta\eta} = 0. \tag{4.6}$$

Using the values of A, B and $Y'(x)$ we rewrite (4.6) as

$$\overline{P}_x + a\overline{P}_{\eta\eta} = 0, \tag{4.7}$$

where

$$a = \frac{(\rho+4)(\rho^3 + 6\rho^2 + 8\rho + 32)}{\rho^4}.$$

When $\eta \to \infty$, (4.2) should match to (1.21), which implies that $\overline{P} \to C$ as $\eta \to \infty$. When $\eta \to -\infty$, (4.2) should match to (1.28). Based on these observations we seek a solution of (4.7) using a similarity variable ($= \eta/\sqrt{1-x}$) and obtain

$$\overline{P}(x,\eta) = \frac{C}{2\sqrt{a\pi}}\int_{-\infty}^{\eta/\sqrt{1-x}} \exp\left(-\frac{u^2}{4a}\right) du, \tag{4.8}$$

where C is as in (1.23)–(1.25). From (4.8) we clearly satisfy the matching condition as $\eta \to -\infty$. As $\eta \to -\infty$, from (4.8) we have

$$\overline{P}(x,\eta) \sim \frac{C(\rho+4)^{1/2}(\rho^3+6\rho^2+8\rho+32)^{1/2}}{\sqrt{\pi}\rho^2}\frac{\sqrt{1-x}}{-\eta} \times \exp\left(-\frac{1}{4a}\frac{\eta^2}{1-x}\right). \tag{4.9}$$

We proceed to expand the ray expansion in region $\mathcal{R}$, as $y \to Y(x)$, and thus verify the matching between regions $\mathcal{R}$ and $\mathcal{L}_1$. With (1.33) and $m = Kx$, $n = Ky$, we write the product form $A^m B^n$ as

$$A^m B^n = e^{K\Phi(x,y)}, \tag{4.10}$$

where $\Phi(x,y)$ is in (1.22) (with (1.33)). We expand Φ and Ψ about $y = Y(x)$, using

$$\Phi(x,y) = \Phi(x,Y(x)) + \Phi_y(x,Y(x))\,\sqrt{\varepsilon}\eta + \frac{1}{2}\Phi_{yy}(x,Y(x))\varepsilon\eta^2 + O(\varepsilon^{3/2}) \tag{4.11}$$

and

$$\Psi(x,y) = \Psi(x,Y(x)) + \Psi_y(x,Y(x))\,\sqrt{\varepsilon}\eta + \frac{1}{2}\Psi_{yy}(x,Y(x))\varepsilon\eta^2 + O(\varepsilon^{3/2}), \tag{4.12}$$

where $y = Y(x) + \sqrt{\varepsilon}\eta$.

From (1.22) and (1.33) we find that

$$\Phi(x,Y(x)) - \Psi(x,Y(x)) = \log(\frac{\rho^2}{4}), \tag{4.13}$$

$$\Phi_y(x,Y(x)) = \Psi_y(x,Y(x)) = \log(\frac{\rho+4}{2}), \tag{4.14}$$

$$\Phi_{yy}(x,Y(x)) = 0, \tag{4.15}$$

and

$$\Psi_{yy}(x,Y(x)) = \frac{\rho^4}{2(\rho+4)(\rho^3+6\rho^2+8\rho+32)(x-1)} = \frac{1}{2a(x-1)}. \tag{4.16}$$

From (4.16), we see that $\frac{1}{2}\Psi_{yy}(x,Y(x))\eta^2$ is the same as $-\frac{1}{4a}\frac{\eta^2}{1-x}$. Thus the exponential parts of (4.2) and (1.28) match.

When $(x,y) \to (x,Y(x))$, $u \to u_{\min} = (\rho+4)/2$, and thus (1.31) becomes

$$\begin{aligned}\widetilde{L}(x,y) &= \frac{\widetilde{L}_0(u)}{\sqrt{\tau}} \\ &\sim \frac{\sqrt{2}\,\sqrt{\rho+4}(\rho+2)}{\sqrt{\rho^3+6\rho^2+8\rho+32}(\rho+4)}\,\frac{1}{\sqrt{1-x}}\,\frac{1}{u-u_{\min}}.\end{aligned} \tag{4.17}$$

Using (1.37) we find that

$$u - u_{\min} \sim u_y(x,Y(x))(y-Y(x)) = u_y(x,Y(x))\,\sqrt{\varepsilon}\eta, \tag{4.18}$$

where

$$u_y(x,Y(x)) = -\frac{\rho^4}{4(\rho^3+6\rho^2+8\rho+32)}\,\frac{1}{1-x}. \tag{4.19}$$

With (1.29), (4.17), (4.18), (4.19), (4.13) and (4.12) we see that $\widetilde{kL}(x,y)$, as $y \to Y(x)$, agrees with the pre-exponential factors in (4.18). We have thus verified that the local expansion in region $\mathcal{L}_1$, as given by (1.50) (with (1.52)) or (4.2) (with (4.8)), asymptotically matches to the ray expansions in regions $\mathcal{D}$ and $\mathcal{R}$.

We can also verify that (4.2) (with (4.8)) matches to the corner layer C_1. Note that as we approach the corner $(x,y) = (1,1)$,

$$\frac{\eta}{\sqrt{1-x}} = \frac{y-Y(x)}{\sqrt{\varepsilon}\,\sqrt{1-x}} = \frac{y-1-Y'(1)(x-1)}{\sqrt{\varepsilon}\,\sqrt{1-x}} = \frac{1}{\sqrt{l}}\left[(1+\frac{4}{\rho}+\frac{16}{\rho^2})l - k\right].$$

But since $Y(x)$ is a linear function of x, the above is an exact identity. We can show that when (1.69) is expanded for $k,l \to \infty$ with $k/l \sim 1 + \frac{4}{\rho} + \frac{16}{\rho^2}$, we obtain precisely (4.2) (with (4.8)). In this limit the saddle point w_s is close to the pole at $w = (\rho+2)/2$.

5. BOUNDARY LAYER EXPANSION NEAR $x = 1$

We construct an approximation in the boundary layer region $\mathcal{L}_2$, which was defined in (1.42). In this region $1 - x = O(\varepsilon)$ (hence $m = K - l = K - O(1)$) and $0 < y < 1$. On this scale we set

$$p(m,n) = P(x,y) = \overline{Q}_l(y), \tag{5.1}$$

(1.2) becomes

$$(\rho+2)\overline{Q}_l(y) = \rho\overline{Q}_l(y-\varepsilon) + \overline{Q}_{l-1}(y) + \overline{Q}_l(y+\varepsilon), \quad l \geqslant 1, \tag{5.2}$$

and (1.7) becomes

$$(\rho+2)\overline{Q}_0(y) = \rho\overline{Q}_0(y-\varepsilon) + \overline{Q}_0(y+\varepsilon), \tag{5.3}$$

where $p(K,n) = \overline{Q}_0(y)$.

We first analyze (5.3) and assume that $\overline{Q}_0$ has an expansion in the WKB form

$$\overline{Q}_0(y) = \exp\left[KF(y)\right]\left[G(y) + \frac{1}{K}G^{(1)}(y) + O(\frac{1}{K^2})\right]. \tag{5.4}$$

Then we use (5.4) in (5.3), and after expanding for $\varepsilon = K^{-1} \to 0$ we obtain the following ODEs for F and G:

$$\rho + 2 = \rho e^{-F'(y)} + e^{F'(y)}, \tag{5.5}$$

$$0 = \rho e^{-F'}\left(\frac{1}{2}F''G - G'\right) + e^{F'}\left(\frac{1}{2}F''G + G'\right). \tag{5.6}$$

Solving (5.5) yields

$$F'(y) = \log\left(\frac{\rho + 2 + \sqrt{\rho^2+4}}{2}\right). \tag{5.7}$$

The other root of (5.5) has $F' < 0$ and must be discarded in view of later asymptotic matching considerations. Integrating (5.7) we obtain

$$F(y) = \log\left(\frac{\rho + 2 + \sqrt{\rho^2+4}}{2}\right)(y-1). \tag{5.8}$$

Then $F'' = 0$ and using (5.7) in (5.6) we find that $G'(y) = 0$, so that $G(y) \equiv G_0$ is a constant.

Now consider $l \geqslant 1$ in (5.2). Setting

$$\overline{Q}_l(y) \sim \exp\left[KF(y)\right]\varepsilon^{-l}G_l(y) \tag{5.9}$$

in (5.2) and expanding the resulting equation for small ε we obtain to leading order

$$\left(e^{F'} - \rho e^{-F'}\right)G_l'(y) + G_{l-1}(y) = 0. \tag{5.10}$$

The general solution of (5.10) is

$$G_l(y) = \frac{1}{l!}[D(y)]^l G_0, \tag{5.11}$$

where

$$D'(y) = -\left[e^{F'} - \rho e^{-F'}\right]^{-1} = -\frac{1}{\sqrt{\rho^2+4}}. \tag{5.12}$$

Asymptotic matching to the corner layer $\mathcal{L}_1$ will require that $D(1) = 0$ so that the solution of (5.12) is

$$D(y) = \frac{1}{\sqrt{\rho^2+4}}(1-y). \tag{5.13}$$

From (5.4), (5.8) and (5.13) we have

$$p(m,n) \sim \exp\left[\frac{F(y)}{\varepsilon}\right]\frac{\varepsilon^{-l}}{l!}\left(\frac{1-y}{\sqrt{\rho^2+4}}\right)^l G_0, \tag{5.14}$$

where G_0 is a constant, to be fixed by asymptotic matching.

We match (5.14) with the solution in corner layer C_1. From (1.69), we obtain in the limit $l = O(1)$, $k \to \infty$,

$$Q(l,k) \sim C_* {w_0}^{-k}\frac{k^l}{l!}(\rho^2+4)^{-l/2}\mathcal{M}_1(\rho), \tag{5.15}$$

where

$$w_0 = \frac{\rho+2+\sqrt{\rho^2+4}}{2},$$

$$\mathcal{M}_1(\rho) = \frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}.$$

Note that $F'(y) = \log w_0$ by (5.7), thus $\exp[F(y)/\varepsilon] = w_0^{-k}$. In view of (5.13), we have

$$\varepsilon^{-l}[D(y)]^l = \left(\frac{k}{\sqrt{\rho^2+4}}\right)^l. \tag{5.16}$$

With (5.16) and (5.15) we see that (5.14) with (1.69) match if

$$G_0 = C_*\frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}. \tag{5.17}$$

With (5.17), (5.14), (1.72), and (5.8) we have established (1.54).

Finally, we verify the matching between (5.14) (with (5.17)) and region $\mathcal{R}$ result in (1.28). We need to expand the ray expansion in $\mathcal{R}$ as $x \uparrow 1$, for a fixed $y \in (0,1)$. In terms of the ray variables this corresponds to letting $u \to u_{\max} = \frac{1}{2}(\rho+2+\sqrt{\rho^2+4})$ for a fixed τ.

By Stirling's formula, for $l \to \infty$,

$$\frac{\varepsilon^{-l}}{l!}[D(y)]^l \sim \frac{1}{\sqrt{2\pi l}}\left(\frac{e}{l\varepsilon}\right)^l[D(y)]^l. \tag{5.18}$$

Using $1-x = \varepsilon l$ we rewrite (5.18) as

$$\frac{\varepsilon^{-l}}{l!}[D(y)]^l \sim \sqrt{\frac{\varepsilon}{2\pi}}\frac{1}{\sqrt{1-x}}\exp\left\{\frac{1}{\varepsilon}\left[-(1-x)\log(1-x)+1-x+(1-x)\log(D(y))\right]\right\}. \tag{5.19}$$

Using (1.33) we have

$$\tau = \frac{y-1}{\frac{\rho}{u}-u}, \tag{5.20}$$

and since $\rho/u_{\max} - u_{\max} = -\sqrt{\rho^2+4}$, we have, as $u \to u_{\max}$

$$\tau \to \tilde{\tau}(y) = -\frac{1}{\sqrt{\rho^2+4}}(y-1) = D(y). \tag{5.21}$$

From (1.30) and (1.33) we obtain

$$\Psi_x(x,y) = \log\left(\rho + 2 - \frac{\rho}{u} - u\right). \tag{5.22}$$

Since $\rho + 2 = \rho/u_{\max} + u_{\max}$, Ψ_x has a logarithmic singularity at $u = u_{\max}$. Using (1.33) we can also write τ in terms of x and u as

$$\tau = \frac{x-1}{\dfrac{\rho}{u} + u - \rho - 2}. \tag{5.23}$$

Using (5.23) and (5.21) in (5.22) we obtain

$$\Psi_x = \log\left[\frac{1-x}{\tau}\right] \sim \log(1-x) - \log[D(y)], \quad u \to u_{\max}. \tag{5.24}$$

By integrating (5.24) with respect to x and noting that $\Psi(1,y) = F(y)$, we obtain

$$\begin{aligned}\Psi(x,y) =& F(y) - (1-x)\log(1-x) + (1-x) \\ &+ (1-x)\log(D(y)) + o(1-x), \quad x \to 1.\end{aligned} \tag{5.25}$$

Thus, as $x \to 1$, $\exp[K\psi(x,y)]$ agrees with the exponential factor in (5.19).

To verify the matching of the algebraic factors, we must show that

$$\widetilde{k}(\varepsilon)\widetilde{L}(x,y) \sim C_* \frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}\frac{\sqrt{\varepsilon}}{\sqrt{2\pi}}\frac{1}{\sqrt{1-x}}, \quad x \to 1. \tag{5.26}$$

We use (5.23) to eliminate τ in (1.31), with which the left side of (5.26) (with (1.32)) becomes

$$\widetilde{k}(\varepsilon)\frac{(u-1)(2u-2-\rho)}{\sqrt{u}\,\sqrt{(\rho+2)u^2 - 4\rho u + \rho^2 + 2\rho}(2u-4-\rho)\sqrt{1-x}}. \tag{5.27}$$

The factors that depend on u in (5.27) are not singular as $u \to u_{\max}$ ($x \to 1$), and simply replacing u by $u_{\max}$ we can easily show that (5.27) becomes the same as the right side of (5.26) (note also that C_* and $\widetilde{k}$ are related by (1.72)). This verifies the matching between boundary layer $\mathcal{L}_1$ and the ray expansion in region $\mathcal{R}$.

6. BOUNDARY LAYER EXPANSIONS NEAR $y = 0$

In this section we consider the scale $n = O(1)$ ($y = O(\varepsilon)$) so there are only a few customers in the second queue. The structure of $p(m,n)$ is different for the following 5 ranges of m: $m = O(1)$ (corner layer C_2), $0 < m/K < X_0(\rho) = (16+4\rho)/(16+4\rho+\rho^2)$ (boundary layer $\mathcal{L}_3$), $m/K = X_0(\rho) + O(K^{-1/2})$ (corner layer C_4), $X_0(\rho) < m/K < 1$ (boundary layer $\mathcal{L}_4$), and $m = K - O(1)$ (corner layer C_3). We proceed to analyze (1.2)–(1.11) for these ranges of (m,n).

If m,n are both $O(1)$ and $K \to \infty$ we solve (1.2)–(1.6) and omit the boundary conditions (1.7)–(1.10). Using the compensation approach of Adan [9] leads to the approximation in (1.63) for $m > n$, and (1.4) can be used to compute $p(n,n)$. The value of the constant C in (1.63) follows from (1.11) if $\rho < 2$, and by asymptotic matching to the region $\mathcal{D}$ ray expansion if $\rho > 2$ or $\rho - 2 = O(\varepsilon)$.

Now consider boundary layer $\mathcal{L}_3$, where $n = O(1)$ and $0 < x < X_0(\rho)$. On this scale only (1.2) and (1.5) apply. We use the variables x and n and define $\mathcal{P}$ by

$$p(m,n) = CA^m B^n \mathcal{P}_n(x;\varepsilon). \tag{6.1}$$

Then (1.2) becomes

$$(\rho+2)\mathcal{P}_n(x;\varepsilon) = \frac{\rho}{B}\mathcal{P}_{n-1}(x;\varepsilon) + A\mathcal{P}_n(x+\varepsilon;\varepsilon) + B\mathcal{P}_{n+1}(x;\varepsilon) \tag{6.2}$$

while (1.5) yields $(\rho+1)\mathcal{P}_0(x;\varepsilon) = A\mathcal{P}_1(x+\varepsilon;\varepsilon) + B\mathcal{P}_1(x;\varepsilon)$, which is in view of (6.2) may be replaced by the "artificial" boundary condition

$$\mathcal{P}_0(x;\varepsilon) = \frac{\rho}{B}\mathcal{P}_{-1}(x;\varepsilon) = \frac{2\rho}{\rho+4}\mathcal{P}_{-1}(x;\varepsilon). \tag{6.3}$$

If (6.1) is to asymptotically match to the region $\mathcal{D}$ expansion in (1.21) we must have $\mathcal{P}_n(x;\varepsilon) \to 1$ as $n \to \infty$. Assuming that $\mathcal{P}_n(x;\varepsilon) \sim \mathcal{P}_n(x)$ as $\varepsilon \to 0$, (6.2) becomes $(\rho+2)\mathcal{P}_n(x) = (\rho/B)\mathcal{P}_{n-1}(x) + A\mathcal{P}_n(x) + B\mathcal{P}_{n+1}(x)$, which is a simple difference equation that admits solutions of the form β^n, where $\beta = 1$ or $\beta = 4\rho/(\rho+4)^2 (< 1)$. By the matching condition we must have $\mathcal{P}_n(x) = 1 + c\beta^n$, and the "reflection coefficient" c is determined by the boundary condition in (6.3). We have thus obtained the approximation in (1.57). Note that $c = 0$ if $\rho = 4$, in which case the region $\mathcal{D}$ approximation remains valid for $n = O(1)$.

Next we consider the boundary layer $\mathcal{L}_4$, where $X_0(\rho) < x < 1$. We again use the variables x and n, but now set

$$p(m,n) = \widetilde{k}(\varepsilon)e^{K\Psi(x,0)}\mathcal{R}_n(x;\varepsilon), \tag{6.4}$$

with $\mathcal{R}_n(x;\varepsilon) \to \mathcal{R}_n(x)$ as $\varepsilon = K^{-1} \to 0$. We will require that as $n \to \infty$ the expansion in (6.4) matches to the approximation in (1.28) as $y \to 0$, which necessitates the exponential factor in (6.4). This matching condition implies that the leading term $\mathcal{R}_n(x)$ satisfies

$$\mathcal{R}_n(x) \sim \widetilde{L}(x,0)e^{n\Psi_y(x,0)}, n \to \infty, \tag{6.5}$$

which we obtained by expanding (1.28) as $y \to 0$.

Using (6.4) in (1.2) and expanding for $\varepsilon \to 0$, using $K\Psi(x+\varepsilon,0) = K\Psi(x,0) + \Psi_x(x,0) + O(\varepsilon)$, leads to the limiting equation

$$(\rho+2)\mathcal{R}_n(x) = \rho\mathcal{R}_{n-1}(x) + e^{\Psi_x(x,0)}\mathcal{R}_n(x) + \mathcal{R}_{n+1}(x), \tag{6.6}$$

while (1.5) is equivalent to the boundary condition

$$\mathcal{R}_0(x) = \rho\mathcal{R}_{-1}(x). \tag{6.7}$$

Equation (6.6) is again a simple difference equation in n, but now the coefficients depend upon x. Seeking geometric solutions of the form $\mathcal{R}_n(x) = [\beta_*(x)]^n$ we see that

$${\beta_*}^2 + \rho = [\rho + 2 - e^{\Psi_x(x,0)}]\beta_*. \tag{6.8}$$

Since $\Psi(x,y)$ satisfies (2.5), we set $y = 0$ in (2.5) with Φ replaced by Ψ, and compare the result to (6.8), to conclude that $\beta_*(x) = e^{\Psi_y(x,0)}$ is one solution of (6.8). The second solution of the quadratic equation is then $\beta_* = \rho e^{-\Psi_y(x,0)}$, and we recall that $\Psi_y(x,0)$ is given explicitly in terms of x in (1.59). In view of the matching condition of (6.5) we thus write

$$\mathcal{R}_n(x) = \widetilde{L}(x,0)\left[e^{n\Psi_y(x,0)} + c_*(x)\rho^n e^{-n\Psi_y(x,0)}\right]. \tag{6.9}$$

Using (6.7) we determine $c_*(x)$ as $c_* = [\rho e^{-\Psi_y(x,0)} - 1]/[1 - e^{\Psi_y(x,0)}]$, and we thus obtain the approximation in (1.58). Note that from (1.59) we can easily show that $2\Psi_y(x,0) > \log(\rho)$, so that as $n \to \infty$ the second exponential in (6.9) becomes negligible.

We next analyze the corner layer C_4, where we use the variables n and ξ where $\xi = [x - X_0(\rho)]\sqrt{K} = [x - X_0(\rho)]/\sqrt{\varepsilon} = O(1)$. We set

$$p(m,n) = CA^m B^n \mathcal{F}_n(\xi;\varepsilon) \tag{6.10}$$

and $\mathcal{F}_n(\xi;\varepsilon) \sim \mathcal{F}_n(\xi)$ will be the leading term as $\varepsilon \to 0$. From (1.2) we again find that $\mathcal{F}_n$ will satisfy the difference equation $(\rho+2)\mathcal{F}_n = (\rho/B)\mathcal{F}_{n-1} + A\mathcal{F}_n + B\mathcal{F}_{n+1}$, and (1.5) leads to $\mathcal{F}_0 = (\rho/B)\mathcal{F}_{-1}$. We thus write

$$\mathcal{F}_n(\xi) = \left\{1 + \frac{2(4-\rho)}{8+6\rho+\rho^2}\left[\frac{4\rho}{(\rho+4)^2}\right]^n\right\}\mathcal{F}_*(\xi). \tag{6.11}$$

To determine the dependence on ξ we use asymptotic matching. As $n \to \infty$ in region C_4 we approach the transition line $\mathcal{L}_1$, where (1.50)–(1.53) apply. From (6.11) we have $F_n(\xi) \to F_*(\xi)$ as $n \to \infty$, while the similarity variable $\eta/\sqrt{1-x}$ in (1.52), as $(x,y) \to (X_0(\rho),0)$,becomes

$$\begin{aligned}\frac{\eta}{\sqrt{1-x}} &= \frac{y-Y(x)}{\sqrt{\varepsilon}\sqrt{1-x}} \sim -\frac{Y'(X_0)(x-X_0)}{\sqrt{\varepsilon}\sqrt{1-X_0}} \\ &= -\xi\frac{Y'(X_0)}{\sqrt{1-X_0}} = -\xi\frac{(16+4\rho+\rho^2)^{3/2}}{\rho^3}.\end{aligned} \tag{6.12}$$

By matching we thus have

$$\mathcal{F}_*(\xi) = \frac{1}{2\sqrt{a\pi}}\int_{-\infty}^{-\widetilde{\xi}} \exp\left(-\frac{u^2}{4a}\right)du \tag{6.13}$$

where $\widetilde{\xi} = \xi(16+4\rho+\rho^2)^{3/2}\rho^{-3}$. With (6.13) and (6.11), the leading term for (6.10) becomes the same as (1.63), after setting $u = -\sqrt{2a}v$ in the integral in (6.13). Note that $\xi'\sqrt{2a} = \widetilde{\xi}$, where a is as in (1.53).

We can also verify that the approximation in (1.60) for $x \approx X_0(\rho)$ asymptotically matches to those for $x < X_0(\rho)$ and $x > X_0(\rho)$ (keeping $n = O(1)$). In fact, for $\xi \to -\infty$ (1.60) reduces to (1.57), as then $\int_{\xi'}^{\infty} e^{-v^2/2}dv \to \sqrt{2\pi}$. The matching between (1.60) and (1.58) is more difficult to verify. We need to show that as ξ (hence ξ') $\to \infty$ (1.60) agrees with (1.58) as $x \downarrow X_0(\rho)$. As $x \to X_0(\rho)$ (1.59) shows that $\Psi_y(x,0) \to \Psi_y(X_0(\rho),0) = \log\left(\frac{\rho+4}{2}\right)$ so that the factor inside the brackets in (1.58) becomes B^n times the factor inside the braces in (6.11). Then letting $\xi \to \infty$ in (1.60) the matching condition is equivalent to

$$\widetilde{k}e^{K\Psi(x,0)}\widetilde{L}(x,0)\Big|_{x\downarrow X_0} \sim CA^m\frac{1}{\sqrt{2\pi}\xi'}e^{-(\xi')^2/2} \tag{6.14}$$

From (1.39) we find that

$$\Psi(X_0(\rho),0) = X_0\log A - \log\left(\frac{\rho^2}{4}\right),$$

$$\Psi_x(X_0(\rho),0) = \log A = \log\left[\frac{\rho^2}{2(\rho+4)}\right],$$

and

$$\Psi_{xx}(X_0(\rho),0) = -\frac{(16+4\rho+\rho^2)^3}{2\rho^2(\rho+4)(\rho^3+6\rho^2+8\rho+32)} = -\left(\frac{\xi'}{\xi}\right)^2.$$

Then expanding $\Psi(x,0)$ in Taylor series about $x = X_0$ and noting the relation between $\widetilde{k}$ and C (which holds for all ρ) in (1.72), we cancel the exponentially varying factors in (6.14) to obtain

$$\sqrt{\frac{\varepsilon}{2\pi}}\frac{\rho^2(\rho+4)}{4(\rho+4)}\widetilde{L}(x,0)\Big|_{x\downarrow X_0} \sim \frac{1}{\sqrt{2\pi}\xi'} = \frac{\sqrt{\varepsilon}}{2\pi}\frac{1}{x-X_0}\frac{\xi}{\xi'}. \tag{6.15}$$

As $y \to 0$ we have $(u,\tau) \to (u_*(x),\tau_*(x))$ in (1.37) and (1.38), and furthermore setting $x = X_0$ we find that

$$u_*(X_0) = B = \frac{\rho+4}{2}, \quad \tau_*(X_0) = \frac{2(\rho+4)}{\rho^2+4\rho+16} \tag{6.16}$$

and

$$\frac{u_*'(X_0)}{u_*(X_0)} = \frac{(16+4\rho+\rho^2)^2}{2(\rho+4)(\rho^3+6\rho^2+8\rho+32)}. \tag{6.17}$$

Setting $y = 0$ and letting $x \to X_0$ in (1.31) and (1.32) we find that

$$\widetilde{L}(x,0) \sim \frac{1}{\sqrt{\tau_*(X_0)}}\frac{1}{2u_*'(X_0)}\frac{1}{x-X_0}K_*, \tag{6.18}$$

$$K_* = \frac{(u_*-1)(2u_*-2-\rho)}{\sqrt{(\rho+2)u_*-{u_*}^2-\rho}\,\sqrt{(\rho+2){u_*}^2-4\rho u_*+\rho^2+2\rho}},$$

where u_* is also evaluated at $x = X_0$. To obtain (6.18) we also used $2u(x,0)-\rho-4 = 2u_*(x)-\rho-4 \sim 2u_*'(X_0)(x-X_0)$. Expression (6.18) shows that $\widetilde{L}(x,0)$ has the correct singularity as $x \to X_0$, and using (1.61) and (6.16)–(6.18) we can easily verify that the multiplicative constants in the left and right sides of (6.15) also agree. This verifies the matching between (1.58) and (1.60).

Finally we consider the corner layer C_3, where $m = K - O(1)$ and $n = O(1)$. We set $l = K - m$ and

$$p(m,n) = \mathcal{G}(l,n;\varepsilon). \tag{6.19}$$

In this corner range only the balance equations (1.2), (1.5), (1.7) and (1.8) apply, and these yield

$$(\rho+2)\mathcal{G}(l,n) = \rho\mathcal{G}(l,n-1) + \mathcal{G}(l-1,n) + \mathcal{G}(l,n+1); \quad l,n \geqslant 1 \tag{6.20}$$

$$(\rho+1)\mathcal{G}(l,0) = \mathcal{G}(l-1,0) + \mathcal{G}(l,n), \quad l \geqslant 1 \tag{6.21}$$

$$(\rho+2)\mathcal{G}(0,n) = \rho\mathcal{G}(0,n-1) + \mathcal{G}(0,n+1), \quad n \geqslant 1 \tag{6.22}$$

and

$$(\rho+1)\mathcal{G}(0,0) = \mathcal{G}(0,1). \tag{6.23}$$

The local approximation $\mathcal{G}$ must satisfy several asymptotic matching conditions. As $n \to \infty$ with $l = O(1)$ it should agree with (1.54) as $y \to 0$; as $l \to \infty$ with $n = O(1)$ it should agree with (1.58) as $x \uparrow 1$; and for $n, l \to \infty$ with $n/l \in (0,\infty)$ it should agree with the region $\mathcal{D}$ ray expansion, as the latter is expanded for $(x,y) \to (1,0)$. In order to have a chance of matching to (1.54), $\mathcal{G}$ must contain the factor $\varepsilon^{-l} = K^l$ and the factor $\exp[KF(0)] = \left[\dfrac{2}{\rho+2+\sqrt{\rho^2+4}}\right]^K$.

Let us thus set

$$\mathcal{G}(l,n) \sim K^l e^{KF(0)}\sqrt{2\pi K}\,\widetilde{k}\,\mathcal{G}_*(l,n). \tag{6.24}$$

The matching between (1.54) and (6.19) will hold if

$$\mathcal{G}_*(l,n) \sim \frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}\frac{1}{l!}\left(\frac{1}{\sqrt{\rho^2+4}}\right)^l\left(\frac{\rho+2+\sqrt{\rho^2+4}}{2}\right)^n, n \to \infty. \tag{6.25}$$

Using (6.24) in (6.20) leads to

$$(\rho+2)\mathcal{G}_*(l,n) = \rho\mathcal{G}_*(l,n-1) + \mathcal{G}_*(l,n+1); \quad l,n \geqslant 1 \tag{6.26}$$

and (6.21) yields

$$(\rho+1)\mathcal{G}_*(l,0) = \mathcal{G}_*(l,1), l \geqslant 1. \tag{6.27}$$

Equations (6.22) and (6.23) then show that (6.26) and (6.27) hold also at $l = 0$. The general solution to (6.26) and (6.27) is

$$\mathcal{G}_*(l,n) = \widetilde{\mathcal{G}}(l)\left[\left(\frac{\rho+2+\sqrt{\rho^2+4}}{2}\right)^n + \widetilde{c}\left(\frac{2\rho}{\rho+2+\sqrt{\rho^2+4}}\right)^n\right] \tag{6.28}$$

where

$$\widetilde{c} = 1 + \frac{\rho^2}{2} - \frac{\rho}{2}\sqrt{\rho^2+4}.$$

Then using the matching condition (6.25) we determine $\widetilde{\mathcal{G}}(l)$ as

$$\widetilde{\mathcal{G}}(l) = \frac{\rho+2+\sqrt{\rho^2+4}}{2\rho^2}\frac{1}{l!}\left(\frac{1}{\sqrt{\rho^2+4}}\right)^l. \tag{6.29}$$

We have thus established (1.62).

We verify that as $l \to \infty$, (1.62) matches to (1.58) as $x \uparrow 1$. From (1.59) we see that

$$e^{\Psi_y(1,0)} = \frac{\rho+2+\sqrt{\rho^2+4}}{2}, \quad \frac{\rho e^{-\Psi_y(1,0)}-1}{1-e^{\Psi_y(1,0)}} = \widetilde{c}$$

and thus the bracketed factor in (1.58) approaches, as $x \to 1$, that in (6.28). It thus remains to show that

$$K^l e^{KF(0)}\sqrt{2\pi K}\widetilde{\mathcal{G}}(l)\Big|_{l\to\infty} \sim \widetilde{L}(x,0)e^{K\Psi(x,0)}\Big|_{x\to 1}. \tag{6.30}$$

From (1.39) and (1.61) we have $\Psi(1,0) = F(0)$. For $x \to 1$ we also obtain from (1.39)

$$\Psi(x,0) = \Psi(1,0) + (x-1)\log(1-x) + 1 - x + \frac{1}{2}(x-1)\log(\rho^2+4) + o(1-x).$$

Since $1 - x = \varepsilon l$ it follows that

$$e^{K\Psi(x,0)} \sim e^{KF(0)}(\varepsilon l)^{-l}(\rho^2+4)^{-l/2}e^l, \quad x \uparrow 1. \tag{6.31}$$

Using (6.31) in (6.30) and expanding $l!$ in (6.29) by Stirling's formula we have $\sqrt{2\pi K}/l! \sim l^{-l}e^l\sqrt{K/l}$, $\sqrt{K/l} = 1/\sqrt{1-x}$, so the matching in (6.30) is possible if

$$\widetilde{L}(x,0) \sim \frac{1}{\sqrt{1-x}}\frac{\sqrt{\rho^2+4}+\rho+2}{2\rho^2}, \quad x \to 1. \tag{6.32}$$

Setting $y = 0$ and letting $x \to 1$ in (1.40) leads to

$$\lim_{x\to 1}[\sqrt{1-x}\widetilde{L}(x,0)] = \frac{\sqrt{\tau_*(1)}[u_*(1)-1][2u_*(1)-2-\rho]}{u_*(1)[2u_*(1)-4-\rho][\rho^2+4]^{1/4}} \tag{6.33}$$

But, $\tau_*(1) = \tau(1,0) = (\rho^2+4)^{-1/2}$ and $u_*(1) = \frac{1}{2}[\rho+2+\sqrt{\rho^2+4}]$ by (1.37) and (1.38), and then (6.33) agrees precisely with (6.32). This completes the matching verification between the regions $\mathcal{L}_4$ and C_3.

We comment that for $n = O(1)$ and for all 5 ranges of x, the local expansions amount simply to adding a second exponential in n, to satisfy the boundary condotion in (1.5). This shows that the boundary $n = 0$ affects the solution only locally, on the scale $n = O(1)$, for $K \to \infty$. This can also be said about the boundary condition along $m = K$ in (1.7). However, the BC along $m = n$ in (1.3) and (1.4), and the corner condition(s) at $(m,n) = (K,K)$, do affect the solution at points far from these regions.

7. NUMERICAL STUDIES

We test the accuracy of some of our asymptotic results. First we consider the diagonal probabilities, with $m = n$. Our asymptotic analysis led to different approximations for the scales $n = O(1)$ (corresponding to corner region C_2), $0 < y = n/K < 1$ (where (1.27), with (1.23)–(1.25), applies), and $n = K - l = K - O(1)$ (corresponding to C_1, where (1.70) or (1.71) applies). In Table 1 we take $\rho = 2$ and $K = 15$, and compare the exact $p(n, n)$ to the three asymptotic results. For $n = O(1)$ we used (1.63) in (1.4) to get

$$p(n,n) \sim \frac{2C}{\rho+2} \sum_{j=0}^{\infty} c_j (a_j b_j)^n \left(\frac{\rho}{b_j} + a_j \right), \quad n \geqslant 1 \tag{7.1}$$

and then $\rho p(0,0) = 2p(1,0)$ yields

$$p(0,0) \sim \frac{2C}{\rho} \sum_{j=0}^{\infty} c_j a_j. \tag{7.2}$$

Here $c_0 = 1$.

For $\rho = 2$, (1.24) with $\alpha = 0$ yields $C \sim 2/(3K)$. On the scale $n = O(1)$ we used (7.1) and (7.2), while for $0 < y < 1$ we used (1.27), which is this case gives $p(n,n) \sim C/2 \sim 1/(3K) = 1/45$, and this is independent of n. Table 1 shows that the $n = O(1)$ result agrees with the exact to 4 decimal places for $0 \leqslant n \leqslant 9$, while the $n = K - O(1)$ result agrees with the exact to 4 decimal places for $8 \leqslant n \leqslant 15$. The table also shows that the uniform approximation $p(n,n) \approx 1/45$ is correct to 2 decimal places for $2 \leqslant n \leqslant 12$, but breaks down near $n = 0$ and $n = K = 15$. Near these boundaries it becomes necessary to use the corner layers C_2 and C_1, as our asymptotic analysis predicts.

Table 1
$K = 15$, $\rho = 2.0$

n	Num Result	$n = O(1)$	$0 < y < 1$	$n = K - O(1)$
0	.01254	.01254	.02222	.02222
1	.01968	.01968	.02222	.02222
2	.02164	.02164	.02222	.02222
3	.02209	.02209	.02222	.02222
4	.02219	.02219	.02222	.02222
5	.02221	.02221	.02222	.02222
6	.02222	.02222	.02222	.02222
7	.02222	.02222	.02222	.02222
8	.02222	.02222	.02222	.02222
9	.02222	.02222	.02222	.02222
10	.02223	.02222	.02222	.02223
11	.02225	.02222	.02222	.02225
12	.02232	.02222	.02222	.02232
13	.02255	.02222	.02222	.02255
14	.02357	.02222	.02222	.02357
15	.03333	.02222	.02222	.03333

In Table 2 we take $\rho = 3$ and $K = 15$. Now we approximate C by (1.25), and again use (7.1) or (7.2) for $n = O(1)$, (1.27) for $0 < y < 1$, and (1.70) or (1.71) for $n = K - O(1)$. Now Table 2 shows that $p(n,n)$ increases with n, which can be expected since $\rho < 2$. The data again shows the necessity of treating separately the three scales of n, and the agreement with the exact numerical results is very good.

We next test the accuracy of the region $\mathcal{R}$ ray expansion, i.e. (1.28), for $x > y > 0$ and $y < Y(x)$. In Table 3 we take $\rho = 8$ and $K = 30$. For $\rho = 8$ the transition curve that separates regions $\mathcal{R}$ and $\mathcal{D}$ is given by $y = Y(x) = \frac{1}{4}(7x - 3)$, for $x \in (\frac{3}{7}, 1)$. Table 3 has $24 \leqslant m \leqslant 26$ and $12 \leqslant n \leqslant 15$, which corresponds to (x, y) lying in region $\mathcal{R}$. Note that by either decreasing m or increasing n we move toward the transition curve $\mathcal{L}_1$, where (1.28) ceases to be valid. The errors between exact and asymptotic results in Table 3 are typically about 20%, with the worst error being about 36%, when $(m, n) = (24, 15)$, and the least error being about 8%, when $(m, n) = (26, 12)$. Note that $(m, n) = (24, 15)$ is closest to the transition line $\mathcal{L}_1$, i.e., $y = Y(x)$.

In Table 4 we retain $\rho = 8$ but increase K to 50. We want to examine a similar window in the triangle $\mathcal{T}$ as in Table 3, so we scale the values of (m, n) in Table 3 by a factor of about 5/3, and consider in Table 4 the lattice points $\{(m, n)\colon 40 \leqslant m \leqslant 43,\ 20 \leqslant n \leqslant 25\}$. Now the worst error is about 28%, when $(m, n) = (40, 25)$, and the least error is about 5%, when $(m, n) = (43, 20)$. Typical errors are $10 - 15\%$. The improved accuracy shown in Table 4 (over Table 3) is consistent with (1.28) having an error factor of $1 + O(K^{-1})$, and this is indeed predicted by our asymptotic analysis.

Finally we test the accuracy of (1.26) (or (1.21)), the ray expansion in region $\mathcal{D}$. We now take a small value of ρ, $\rho = 0.5$, with which the transition curve $\mathcal{L}_1$ becomes $y = Y(x) = 73x - 72$, $x \in (\frac{72}{73}, 1)$. The slope of $Y(x)$ is nearly vertical and thus region $\mathcal{D}$ fills nearly the entire triangle $\mathcal{T}$. Now C is computed from (1.23) and Table 5 has $K = 30$, $13 \leqslant m \leqslant 15$ and $5 \leqslant n \leqslant 9$. For each data point the exact and asymptotic results agree to 4 decimal places. This is also consistent with the asymptotics, as when $\rho < 2$ (1.26) fails to be exact due only to the effects of the finite capacities, but these are exponentially small for K large and $(x, y) \in \mathcal{D}$.

In Tables 2–5 we use the notation $1.00(-5)$ for 1.00×10^{-5}. The numerical comparisons show the necessity of treating the separate regions of the state space, including the boundary and corner regions. The agreement between the asymptotic and numerical results is fairly good for region $\mathcal{R}$ and excellent for region $\mathcal{D}$, and this is easily explained in terms of the error factors in the expansions.

Table 2
$K = 15$, $\rho = 3.0$

n	Num Result	$n = O(1)$	$0 < y < 1$	$n = K - O(1)$
0	5.961 (−7)	5.960 (−7)	9.933 (−7)	9.933 (−7)
1	1.992 (−6)	1.992 (−6)	2.235 (−6)	2.235 (−6)
2	4.893 (−6)	4.892 (−6)	5.028 (−6)	5.028 (−6)
3	1.124 (−5)	1.123 (−5)	1.131 (−5)	1.131 (−5)
4	2.542 (−5)	2.541 (−5)	2.545 (−5)	2.546 (−5)
5	5.728 (−5)	5.725 (−5)	5.728 (−5)	5.730 (−5)
6	1.289 (−4)	1.288 (−4)	1.288 (−4)	1.289 (−4)
7	2.903 (−4)	2.899 (−4)	2.899 (−4)	2.903 (−4)
8	6.537 (−4)	6.524 (−4)	6.524 (−4)	6.537 (−4)
9	1.473 (−3)	1.468 (−3)	1.468 (−3)	1.473 (−3)
10	3.323 (−3)	3.303 (−3)	3.303 (−3)	3.323 (−3)
11	7.519 (−3)	7.432 (−3)	7.432 (−3)	7.519 (−3)
12	1.711 (−2)	1.672 (−2)	1.672 (−2)	1.716 (−2)
13	3.956 (−2)	3.762 (−2)	3.762 (−2)	3.956 (−2)
14	9.650 (−2)	8.465 (−2)	8.465 (−2)	9.650 (−2)
15	3.333 (−1)	1.904 (−1)	1.904 (−1)	3.333 (−1)

Table 3
$K = 30, \quad \rho = 8.0$

n \ m	24	25	26
12	3.155 (−18)	5.718 (−18)	9.185 (−18)
	3.715 (−18)	6.368 (−18)	9.906 (−18)
13	2.230 (−17)	4.179 (−17)	6.967 (−17)
	2.715 (−17)	4.737 (−17)	7.584 (−17)
14	1.559 (−16)	3.025 (−16)	5.239 (−16)
	1.989 (−16)	3.510 (−16)	5.772 (−16)
15	1.078 (−15)	2.166 (−15)	3.903 (−15)
	1.470 (−15)	2.595 (−15)	4.367 (−15)

Table 4
$K = 50, \quad \rho = 8.0$

n \ m	40	41	42	43
20	1.543 (−29)	2.839 (−29)	4.875 (−29)	7.718 (−29)
	1.742 (−29)	3.111 (−29)	5.228 (−29)	8.154 (−29)
21	1.085 (−28)	2.037 (−28)	3.574 (−28)	5.790 (−29)
	1.244 (−28)	2.254 (−28)	3.859 (−28)	6.145 (−28)
22	7.580 (−28)	1.452 (−27)	2.605 (−27)	4.319 (−27)
	8.856 (−28)	1.627 (−27)	2.836 (−27)	4.610 (−27)
23	5.259 (−27)	1.029 (−26)	1.887 (−26)	3.204 (−26)
	6.289 (−27)	1.170 (−26)	2.075 (−26)	3.441 (−26)
24	3.624 (−26)	7.242 (−26)	1.358 (−25)	2.362 (−25)
	4.462 (−26)	8.391 (−26)	1.511 (−25)	2.557 (−25)
25	2.479 (−25)	5.060 (−25)	9.711 (−25)	1.731 (−24)
	3.167 (−25)	5.998 (−25)	1.095 (−24)	1.890 (−24)

Table 5
$K = 30, \quad \rho = 0.5$

n \ m	13	14	15
5	1.971 (−18)	5.477 (−20)	1.521 (−21)
	1.971 (−18)	5.477 (−20)	1.521 (−21)
6	4.436 (−18)	1.232 (−19)	3.423 (−21)
	4.436 (−18)	1.232 (−19)	3.423 (−21)
7	9.983 (−18)	2.773 (−19)	7.702 (−21)
	9.983 (−18)	2.773 (−19)	7.702 (−21)
8	2.246 (−17)	6.239 (−19)	1.733 (−20)
	2.246 (−17)	6.239 (−19)	1.733 (−20)
9	5.053 (−17)	1.403 (−18)	3.899 (−20)
	5.053 (−17)	1.403 (−18)	3.899 (−20)

REFERENCES

[1] Foschini, G. (1977). On Heavy Traffic Diffusion Analysis and Dynamic Routing in Packet Switched Networks. In K. Chandy, & J. Reiser (Eds.), *Computer Performance*, (pp. 499–513). New York: North-Holland.

[2] Foschini, G., & Salz, J. (1978). A Basic Dynamic Routing Problem and Diffusion. IEEE (Trans.) *Commun., 26*(3), 320–327.

[3] Gertsbakh, I. (1984). The Shorter Queue Problem: A Numerical Study Using the Matrix-geometric Solution. *Eur. J. Oper. Res., 15* (3), 374–381.

[4] Flatto, L., & McKean, H. (1977). Two Queues in Parallel. *Commun. Pure Appl. Math., 30*(2), 255–263.

[5] Kingman, J. F. C. (1961). Two Similar Queues in Parallel. *Annals Math. Statist., 32*(4), 1314–1323.

[6] Halfin, S. (1985). The Shortest Queue Problem. *J. Appl. Probab., 22*(4), 865–878.

[7] Adan, I. J. B. F., Wessels, J., & Zijm, W. H. M. (1990). Analysis of the Symmetric Shortest Queue Problem. *Comm. Statist. Stochastic Models, 6*(4), 691–713.

[8] Adan, I. J. B. F.,Wessels, J., & Zijm, W. H. M. (1991). Analysis of the Asymmetric Shortest Queue Problem. *Queueing Systems Theory Appl., 8*(1), 1–58.

[9] Adan, I. J. B. F., J. van Houtum, G., & Van der Wal, J. (1994). Upper and Lower Bounds for the Waiting Time in the Symmetric Shortest Queue Problem. *Ann. Oper. Res., 48*(2), 197–217.

[10] Wang, P. P. (2000). Workload Distribution of Discrete Time Parallel Queues with Two Servers. *Naval Res. Logist., 47*(5), 440–454.

[11] Wu, P., & Posner, M. J. M. (1997). A Level-crossing Approach to the Solution of the Shortest Queue Problem. *Oper. Res. Lett., 21*(4), 181–189.

[12] J. van Houtum, G., Adan, I. J. B. F., Wessels, J., & Zijm, W. H. M. (2001). Performance Analysis of Parallel Identical Machines with a Generalized Shortest Queue Arrival Mechanism. *OR Spektrum, 23*(3), 411–427.

[13] Zhao, Y. Q., & Grassman, W. K. (1990). The Shortest Queue Model with Jockeying. *Naval Res. Logist., 37*(5), 773–787.

[14] Adan, I. J. B. F., Wessels, J., & Zijm, W. H. M. (1991). Analysis of the Symmetric Shortest Queue Problem with Threshold Jockeying. *Comm. Statist. Stochastic Models, 7*(4), 615–627.

[15] Adan, I. J. B. F., Wessels, J., & Zijm, W. H. M. (1993). Matrix-geometric Analysis of the Shortest Queue Problem with Threshold Jockeying. *Oper. Res. Lett., 13*(2), 107–112.

[16] Conolly, B. W. (1984). The Autostrada Queueing Problems. *J. Appl. Prob., 21*(2), 394–403.

[17] Tarabia, A. M. K. (2008). Analysis of Two Queues in Parallel with Jockeying and Restricted Capacities. *Appl. Math. Model., 32*(5), 802–810.

[18] Tarabia, A. M. K. (2009). Transient Analysis of Two Queues in Parallel with Jockeying. *Stochastics: An International Journal of Probability and Stochastic Processes, 81*(2), 129–145.

[19] Flatto, L. (1989). The Longer Queue Problem. *Probab. Eng. Inform. Sci., 3*(4), 537–559.

[20] Blanc, J. P. C. (1992). The Power-series Algorithm Applied to the Shortest-queue Model. *Oper. Res., 40*(1), 157–167.

[21] Grassman, W. K. (1980). Transient and Steady State Results for Two Parallel Queues. *Omega Int. J. Manag. Sci., 8*(1), 105–112.

[22] Hooghiemstra, G., Keane, M., & Van de Rhee, S. (1988). Power-series for Stationary Distributions of Coupled Processor Models. *SIAM J. Appl. Math., 48*(5), 1159–1166.

[23] Rieman, M. (1984). Some Diffusion Approximations with State Space Collapse. In A. V. Balakreshnan

& M. Thomas (Eds.), *Modelling and Performance Evaluation Methodology*, (pp. 209–240). Berlin: Springer-Verlag.

[24] Fleming, P. J., & Simon, B. (1999). Heavy Traffic Approximations for a System of Infinite Servers with Load Balancing. *Probab. Eng. Inform. Sci., 13*(3), 251–273.

[25] Turner, S. R. E. (2000). A Join the Shorter Queue in Heavy Traffic. *J. Appl. Probab., 37*(1), 212–223.

[26] Sakuma, Y, Miyazawa, M., & Zhao, Y. Q. (2006). Decay Rates for $PH/M/2$ Queue with Shortest Queue Discipline. *Queueing Systems, 53*(4), 189–201.

[27] Kurkova, I. A., & Suhov, Y. M. (2003). Malyshev's Theory and JS-queues: Asymptotics of Stationary Probabilities. *Ann. Appl. Probab., 13*(4), 1313–1354.

[28] McDonald, D. (1996). Overloading Parallel Servers when Arrivals Join the Shortest Queue. In P. Glasserman, K. Sigman & D. Yao (Eds.), *Stochastic Networks: Stability and Rare Events, Lecture Notes in Statistics 117*, (pp. 169–196). New York: Springer Verlag.

[29] Alanyali, M., & Hajek, B. (1998). On Large Deviations in Load Sharing Networks. *Ann. Appl. Probab., 8*(1), 67–97.

[30] Shwartz, A., & Weiss, A. (1995). *Large Deviations for Performance Analysis, Queues, Communications, and Computing, Stochastic Modeling Series.* London: Chapman and Hall.

[31] Yao, H., & Knessl, C. (2005). On the Infinite Server Shortest Queue Problem: Symmetric Case. *Stochastic Models, 21*(1), 101–132.

[32] Yao, H., & Knessl, C. (2006). On the Infinite Server Shortest Queue Problem: Non-symmetric Case. *Queueing Systems, 52*(2), 157–177.

[33] Yao, H., & Knessl, C. (2008). On the Shortest Queue Version of the Erlang Loss Model. *Studies in Appl. Math., 120*(2), 129–212.