COMMUN. STATIST.-STOCHASTIC MODELS, 4(2), 315-359 (1988)

THE CORRELATION FUNCTIONS OF RBM AND M/M/1

Joseph Abate

AT&T Bell Laboratories
Warren, New Jersey 07060

Ward Whitt

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

This paper describes the (auto) correlation functions of regulated or reflecting Brownian motion (RBM) and several processes associated with the M/M/1 queue. For RBM and the M/M/1 continuous-time queue-length process, the correlation function of the stationary process coincides with the complementary stationary-excess cdf (cumulative distribution function) associated with a previously studied first-moment cdf. The first-moment cdf is the mean as a function of time given that the process starts at the origin, normalized by dividing by the steady-state limit. The M/M/1 first-moment cdf in turn is the stationary-excess cdf associated with the M/M/1 busy-period cdf. In fact, all the moment cdf's and correlation functions can be expressed directly in terms of the busy-period cdf. This structure provides the basis for simple approximations of the correlation functions and the moments as functions of time by hyperexponentials.

1. Introduction

The primary purpose of this paper is to provide useful descriptions of the (auto) correlation functions of stationary versions of regulated (a.k.a.

reflecting) Brownian motion (RBM) and four fundamental stochastic processes associated with the M/M/1 queue. The M/M/1 stochastic processes considered are: number in system, number in service, number waiting (excluding the one in service, if any) and workload or virtual waiting time, all in continuous time. These stochastic processes have been studied quite extensively, e.g., [5], [7], [9], [13]-[27], [31] and references cited there, but we believe that we have important contributions, which are largely based on our recent results for the transient behavior of RBM and M/M/1 [1]-[4].

We focus on the processes starting at the origin. We regard the moments (as functions of time) under this special initial condition as cdf's (cumulative distribution functions) after normalizing by the steady-state limits. As shown in [1]-[4], these moment cdf's have interesting properties; e.g., they can be expressed quite directly in terms of the busy-period cdf. Here we show that the correlation functions of the stationary processes also have very simple expressions in terms of these moment cdf's, and thus in terms of the busy-period cdf. As a consequence, we obtain nice approximations. We also show that all the correlation functions are monotone, and all but the correlation function of the M/M/1 workload are completely monotone (mixtures of exponentials).

The fact that there is a close connection between the M/M/1 transient behavior and the busy-period cdf also follows from the associated process construction used by Karlin and McGregor [13] to obtain the spectral representation and from related duality; see Chapter 3 of van Doorn [27]. However, the rich M/M/1 structure described here evidently has not been uncovered before.

By appropriate scaling of time and space, RBM appears as the limit of the M/M/1 queue-length and workload processes as $\rho \rightarrow 1$ where ρ is the traffic intensity. Thus, our results for RBM appear as a special case of our results for M/M/1. We can thus derive the RBM results in two ways:

either directly or from the M/M/1 results. RBM results have been obtained from M/M/1 results previously by Ott [22] and Cohen and Hooghiemstra [8]. The general approach is also commonly applied with the invariance principle; e.g., Section 11 of Billingsley [6].

The rest of this paper is organized as follows. After briefly discussing RBM in Section 2, we treat the M/M/1 processes representing number in system, number in service, number waiting, and workload in Sections 3-6, respectively. For the last three M/M/1 processes we also describe the moment cdf's for the first time here. In Section 7 we exhibit all derivatives at the origin of all the cdf's and thus obtain power series representations. In Section 8, we discuss approximations for the new moment cdf's and the correlation functions. In Section 9 we conclude with some bounds and inequalities.

2. RBM

RBM is Brownian motion on the positive half line with constant negative drift and constant diffusion coefficient, modified by an impenetrable reflecting barrier at the origin; see Harrison [10]. For both deriving and expressing results about RBM, it is convenient to work with *canonical RBM* which has drift -1 and diffusion coefficient $+1$. This simplification is without loss of generality because, if $R(t;\, \mu,\sigma^2,X)$ denotes RBM with drift μ, diffusion coefficient σ^2 and random initial position X, then $\{aR(bt;\, \mu,\sigma^2,X)\colon t \geqslant 0\}$ has the same distribution (as a stochastic process, i.e., the same finite-dimensional distributions) as $\{R(t;\, -1,1,aX)\colon t \geqslant 0\}$ where a and b are the scaling constants

$$a = |\mu|/\sigma^2 \quad \text{and} \quad b = \sigma^2/\mu^2. \tag{2.1}$$

Let $c(t;\, \mu,\sigma^2)$ denote the correlation function of the stationary version of $\{R(t;\, \mu,\sigma^2,X)\colon t \geqslant 0\}$, which is obtained by letting X be exponentially distributed with mean 1/2a. As a consequence,

$$c(t;\mu,\sigma^2) = 4a^2E[XR(t;\mu,\sigma^2,X)] - 1 \tag{2.2}$$
$$= 4E[aXR(t/b;-1,-1,aX)] - 1 = c(t/b;-1,1).$$

Henceforth, we only consider canonical RBM, suppressing μ and σ^2 from the notation. Our object now is to describe the correlation function $c(t)$. We obtain descriptions of $c(t)$ for RBM as a special case of the M/M/1 queue (with an appropriate time scaling) in which the traffic intensity is $\rho = 1$. Alternately, all RBM results could be obtained directly from [1], [2]. In Corollary 1 to Theorem 1 in Section 3 below, we obtain a nice explicit expression for $c(t)$ for RBM in terms of the standard normal density and cdf.

3. The M/M/1 Queue-Length Process

Let $Q(t)$ represent the number in system, i.e., the queue-length (including the customer in service, if any) at time t in the M/M/1 queueing model; e.g., [3], [4], [7]. Without loss of generality, let the service rate be 1, so that the arrival rate coincides with the traffic intensity ρ. Assume that $\rho < 1$, so that the system is stable with $Q(t)$ converging in distribution to $Q(\infty)$ as $t \to \infty$ where $P(Q(\infty) = n) = (1-\rho)\rho^n$ for $n \geq 0$.

As in (2.1) of [4], we introduce additional time scaling, working with the *transition function*

$$P_{in}(t) = P(Q(2(1-\rho)^{-2}t) = n \mid Q(0) = i),\ t \geq 0\,. \tag{3.1}$$

This time scaling (plus an additional space scaling) makes RBM results appear explicitly as the special case of M/M/1 in which $\rho = 1$. (This scaling appears in the heavy-traffic limit theorems in which $Q(t)$ appropriately normalized converges to RBM as $\rho \to 1$, e.g., [12]. Since ρ appears in the time scaling in (3.1), the limit here at $\rho = 1$ has negative drift.) In conjunction with this scaling, let $\theta = (1-\rho)/2$.

Let $m_k(t,i)$ be the associated k^{th} *moment function*, defined by

$$m_k(t,i) = \sum_{n=0}^{\infty} n^k P_{in}(t),\ t \geq 0\,, \tag{3.2}$$

and let $H_k(t)$ be the associated k^{th}*-moment cdf*

$$H_k(t) = m_k(t,0)/m_k(\infty),\ t \geq 0\,, \tag{3.3}$$

where $m_k(\infty) = \lim_{t\to\infty} m_k(t,i)$. (The normalization by $m_k(\infty)$ introduces the appropriate space scaling to obtain RBM as $\rho \to 1$.)

Let $c_q(t)$ be the correlation function of $Q(t)$ in the time scale (3.1), which is

$$\begin{aligned} c_q(t) &= \frac{E[Q(\infty)m_1(t,Q(\infty))] - [EQ(\infty)]^2}{Var[Q(\infty)]} \\ &= \frac{(1-\rho)^2}{\rho} E[Q(\infty)m_1(t,Q(\infty))] - \rho\,. \end{aligned} \tag{3.4}$$

From (3.4) we see that $c_q(0) = 1$. We now apply results about $m_1(t,i)$ in [4] to characterize $c_q(t)$. In particular, we exploit spatial homogeneity away from the barrier and time reversibility.

For any cdf $G(t)$ on $[0,\infty)$ with mean m, let $G^c(t) = 1 - G(t)$ be the associated *complementary cdf* and let $G_e(t)$ be the associated *stationary-excess cdf*, defined by

$$G_e(t) = m^{-1}\int_0^t G^c(s)\,ds,\ \ t \geq 0. \tag{3.5}$$

Thus, $H_{1e}^c(t)$ is the complementary stationary-excess cdf associated with the first-moment cdf $H_1(t)$ defined in (3.3).

Theorem 1. $c_q(t) = H_{1e}^c(t) = 1 - 2\int_0^t H_1^c(s)\,ds,\ \ t \geq 0.$

Proof. By Theorem 8.1 of [4] (the conservation law), $m_1(t,i)$ is differentiable in t with derivative

$$m_1'(t,i) = \frac{2}{(1-\rho)^2}[P_{i0}(t) - (1-\rho)],\ t \geqslant 0.$$

Since the M/M/1 queue-length process is time reversible, $P_{i0}(t) = \rho^{-i}P_{0i}(t)$; see (4.3) of [4]. Therefore, the derivative of $c_q(t)$ can be expressed as

$$\begin{aligned} c_q'(t) &= \frac{(1-\rho)^2}{\rho} E[Q(\infty)m_1'(t,Q(\infty))] \\ &= \frac{(1-\rho)^2}{\rho}\sum_{i=0}^{\infty}(1-\rho)\rho^i i\frac{2}{(1-\rho)^2}\left[(\rho^{-i}P_{0i}(t) - (1-\rho)\right] \\ &= -2[1 - H_1(t)]\,. \end{aligned}$$

This equals $-h_{1e}(t)$ where $h_{1e}(t)$ is the density of $H_{1e}(t)$ because the mean of $H_1(t)$ is 1/2; see Corollary 5.2.4 of [4]. ∎

By Corollary 5.2.2 of [4], $H_1(t)$ converges in distribution (pointwise) to nondegenerate cdf's as $\rho \to 1$ and as $\rho \to 0$ (the case $\rho = 1$ corresponding to RBM), so that the same is true for $c_q(t)$ by virtue of Theorem 1. Combining Corollaries 1.1.1(b) and 1.5.1 of [1] with Theorem 1 above, we obtain an explicit representation for the correlation function of RBM in terms of the density $\phi(t)$ and cdf $\Phi(t)$ of a standard normal random variable $N(0,1)$.

Corollary 1. For RBM ($\rho = 1$), $c(t) = 2(1-2t-t^2)[1 - \Phi(t^{1/2})] + 2t^{1/2}(1+t)\phi(t^{1/2})$.

Remarks 3.1. (a) For RBM, but not M/M/1 with $\rho < 1$, $H_{1e} = H_2$. Moreover, for RBM H_k is the k-fold convolution of H_1. (This relation holds for the factorial-moment cdfs in M/M/1; see [3], [4].) Having the stationary-excess cdf H_{1e} coincide with the two-fold convolution $H_2 = H_1 * H_1$ characterizes H_1 for RBM; see Corollary 1.5.2 of [1].

(b) The expression for $c(t)$ in Corollary 1 agrees with the integral representations in (1.20) of Ott [22] and Theorem 2 of Woodside et al. [31]. ■

We have no such explicit expression for $c_q(t)$ when $0<\rho<1$, but we can also treat the limit as $\rho \to 0$ by applying Corollary 5.2.2(b) of [4]. There is a nontrivial limit because of the time scaling.

Corollary 2. As $\rho \to 0$, $c_q(t) \to e^{-2t}$, $t \geqslant 0$.

For $0<\rho<1$, we have a representation of $c_q(t)$ in terms of Laplace transforms. (The first transform representation for $c_q(t)$, a cosine transform, was determined by Morse [19].) For any function $f(t)$, let $\hat{f}(s)$ denote its Laplace transform. Let $h_k(t)$ be the density of $H_k(t)$. (Recall that here $H_2(t)$ is the ordinary second-moment cdf, not the second-factorial-moment cdf.) From Theorem 5.1 and (2.4) of [4],

$$\begin{aligned} \hat{H}_1^c(s) &= s^{-1}[1-\hat{h}_1(s)] \\ &= (1 + (1-\theta)s + [1 + 2(1-\theta)s + \theta^2 s^2]^{1/2})^{-1} . \end{aligned} \tag{3.6}$$

Corollary 3. $\hat{c}_q(s) = s^{-1}\left[1 - 2\hat{H}_1^c(s)\right] = \theta \hat{H}_1^c(s) + (1-\theta)\hat{H}_2^c(s)$ where $\theta = (1-\rho)/2$, $\hat{H}_1^c(s)$ is in (3.6) and $\hat{h}_k(s)$ is in Theorem 5.1 and Corollary 5.2.1 of [4].

From Corollary 3.3.1 of [3], we obtain an important shape property. (See [16] for background.)

Corollary 4. $c_q(t)$ has a completely monotone density $c_q'(t)$; i.e., $c_q'(t)$ is a mixture of exponentials.

Proof. Complete monotonicity is obviously inherited through the stationary-excess operator. (This is the same reasoning applied for Theorem 1.7 of [1] and Corollary 3.3.1 of [3].) ■

Remarks 3.2. (a) Ott [22] obtained Corollary 4 for the special case of $\rho = 1$, i.e., for RBM.

(b) The M/M/1 busy-period cdf $B(t)$ and the stationary-excess operator in (3.5) play a fundamental role. By Corollary 3.1.3 of [4], $H_1(t) = B_e(t)$, the stationary-excess distribution associated with $B(t)$. Since $c_q(t) = H_{1e}^c(t)$, $1 - c_q(t)$ is the two-fold iterated stationary-excess operator of $B(t)$; see [29]. ■

Since the k^{th} moment of the stationary-excess cdf $G_e(t)$ in (3.5) is $m_{k+1}/(k+1)m_1$ when m_k is the k^{th} moment of G, we can apply Corollary 5.2.4 of [4] to compute all moments of the *correlation cdf* $1 - c_q(t)$. These moments are useful for making approximation by moment matching.

Corollary 5. The correlation cdf $1-c_q(t)$ has first three moments $m_1 = (1+\rho)/2$, $m_2 = (1+3\rho+\rho^2)/2$ and $m_3 = 3(1+\rho)(1+5\rho+\rho^2)/4$.

We mostly discuss approximations in Section 8, but we present a sample here. As a consequence of Corollary 4, the correlation cdf $1-c_q(t)$ admits a (necessarily unique) hyperexponential (H_2, a mixture of two exponentials) fit to the first three moments in Corollary 5. A fitting scheme is described in Section 5 of [1]. It turns out that we obtain a simple explicit expression for this three-moment fit.

Corollary 6. The unique H_2 fit to the first three moments of the correlation cdf $1-c_q(t)$ yields the approximation

$$c_q(t) \approx 0.5e^{-\frac{2t}{1+\rho+\sqrt{\rho}}} + 0.5e^{-\frac{2t}{1+\rho-\sqrt{\rho}}}, \quad t \geqslant 0 . \tag{3.7}$$

For the case of RBM ($\rho = 1$), Table 5 of [1] shows that (37) is an excellent approximation, provided t is neither very small nor very large. (Recall that $H_2(t) = H_{1e}(t) = 1 - c(t)$ for RBM.) Tables 1 and 2 here show that (3.7) is also an excellent approximation for $\rho = 0.50$ and $\rho = 0.75$, again provided t is neither very small nor very large. (By

TABLE I

A comparison of hyperexponential and normal approximations of the correlation function $c_q(t) = V_2^{\xi}(t)$ with exact values obtained by Laplace transform inversion: the case $\rho = 0.50$.

time t	exact by transform inversion	three-moment H_2 fit (Corollary 6 to Theorem 1)		RBM normal approximation (3.8)
		two terms	one term	
0.01	0.9805	0.9830	-	0.9764
0.05	0.9087	0.9186	-	0.8991
0.10	0.832	0.845	-	0.821
0.15	0.765	0.779	-	0.754
0.20	0.707	0.719	-	0.696
0.25	0.655	0.665	0.399	0.645
0.50	0.461	0.459	0.318	0.456
0.75	0.335	0.329	0.253	0.333
1.00	0.249	0.242	0.202	0.249
1.25	0.187	0.182	0.161	0.188
1.50	0.143	0.140	0.128	0.144
1.75	0.109	0.108	0.102	0.111
2.00	0.0847	0.0849	0.0816	0.0866
2.50	0.0518	0.0528	0.0519	0.0533
3.00	0.0321	0.0332	0.0330	0.0333
3.50	0.0198	0.0210	0.0210	0.0211
4.00	0.0127	0.0133	0.0133	0.0135
4.50	0.0081	0.0085	0.0085	0.0088
5.00	0.0051	0.0054	0.0054	0.0057
6.00	0.00219	0.00218	0.00218	0.00245
7.00	0.00096	0.00088	0.00088	0.00107
8.00	0.00030	0.00036	0.00036	0.00047

TABLE 2

A comparison of hyperexponential and normal approximations of the correlation function $c_q(t) = V_2^c(t)$ with exact values obtained by Laplace transform inversion: the case $\rho = 0.75$. (Also see the third column of Tables 4 and 5.)

time t	exact by transform inversion	three-moment H_2 fit (Corollary 6 to Theorem 1)		RBM normal approximation (3.8)
		two terms	one term	
0.01	0.9805	0.9850	-	0.9796
0.05	0.9139	0.9278	-	0.9117
0.10	0.844	0.862	-	0.842
0.15	0.783	0.802	-	0.781
0.20	0.731	0.747	-	0.728
0.25	0.683	0.697	-	0.681
0.50	0.502	0.502	0.341	0.501
0.75	0.381	0.373	0.282	0.380
1.00	0.293	0.285	0.233	0.293
1.25	0.229	0.222	0.192	0.229
1.50	0.181	0.176	0.159	0.181
1.75	0.144	0.141	0.132	0.144
2.00	0.115	0.114	0.108	0.116
2.50	0.0748	0.0757	0.0739	0.0752
3.00	0.0495	0.0510	0.0505	0.0498
3.50	0.0330	0.0346	0.0344	0.0333
4.00	0.0222	0.0235	0.0235	0.0225
4.50	0.0153	0.0160	0.0160	0.0154
5.00	0.0104	0.0109	0.0109	0.0105
6.00	0.00474	0.00509	0.00509	0.00505
7.00	0.00242	0.00237	0.00237	0.00245
8.00	0.00106	0.00110	0.00110	0.00121
9.00	0.00067	0.00051	0.00051	0.00060

Corollary 2, (3.7) is exact for $\rho = 0$.) The exact numerical values in Tables 1 and 2 are obtained by Laplace transform inversion, as indicated in Section 4.4 of [1].

From Tables 1 and 2, it might appear that the H_2 approximations are uniformly good as $t \to \infty$, but this is not so, because any H_2 complementary cdf is asymptotically exponential as $t \to \infty$, whereas $c_q(t) \sim At^{-3/2}e^{-t/\tau}$ for $\tau = (1+\sqrt{\rho})^2/2$ as $t \to \infty$, where $f(t) \sim g(t)$ means that $f(t)/g(t) \to 1$ as $t \to \infty$; see Corollary 1.1.2 of [1] and Theorem 3.1 and Corollaries 5.2.3 and 5.2.5 of [4]. Moreover, as illustrated by Tables 3-5 of [1], the asymptotic expressions are not good approximations until t is extraordinarily large. The H_2 approximation (3.7) seems to perform well for the t of primary interest; see Sections 1.4-1.6 of [1].

Tables 1 and 2 also describe another approximation based on the correlation function $c(t)$ for RBM in Corollary 1. This *RBM normal approximation* is simply

$$c_q(t) \approx c(t/(1-\theta)), \quad t \geqslant 0 . \tag{3.8}$$

It is motivated by the Laplace transform $\hat{H}_1^c(s)$ in (3.6). If we ignore the $\theta^2 s^2$ term (since $\theta \to 0$ as $\rho \to 1$), we see that s appears multiplied by $(1-\theta)$. This suggests approximation (3.8). The first three moments of the cdf $1 - c(t/(1-\theta))$ are $m_1 = (1-\theta)$, $m_2 = 5(1-\theta)^2/2$ and $m_3 = 21(1-\theta)^3/2$, agreeing with the moments of $1 - c_q(t)$ in Corollary 5 up to terms of order θ^2.

For higher ρ such as 0.75, the RBM normal approximation in (3.8) performs better than the three moment H_2 fit, but the H_2 fit seems easier to work with. For t sufficiently large, e.g., $t \geqslant 1.5$, one exponential term dominates and thus provides a good simple approximation. For obtaining numbers, working with (3.8) is actually not easier than direct transform inversion, but (3.8) shows how $c_q(t)$ depends on ρ for larger ρ. Even with the time scaling (3.1), $c_q(t)$ tends to increase (slowly) with ρ; see Tables 1 and 2. However, for very small t, $c_q(t)$ is actually decreasing in ρ, as can be seen from (3.6). ($c_q(0) = 1$ and $c_q'(\theta) = -2$ for all ρ, but $c_q''(0)$ is a function of ρ. The cases $\rho = 0$ and $\rho = 1$ can be seen from Corollaries 1 and 2.) From Corollary 5, we see that the mean of $1-c_q(t)$ is $1-\theta = (1+\rho)/2$. Formula (3.8) contains a time scaling to produce a one-moment match.

Since we calculated the first moment of $1 - c_q(t)$, we have also calculated the integral of $c_q(t)$, here denoted by $\bar{c}_q$. (It was first found by Morse [19].)

Corollary 7. (Morse) $\bar{c}_q \equiv \int_0^\infty c_q(t)\, dt = \int_0^\infty t\, dH_{1e}(t) = (1+\rho)/2.$

Remark 3.3. The integral in Corollary 7 is one possible representation of the *relaxation time*, i.e., the rate at which the process approaches steady state; see [1]-[3], p. 161 of Keilson [16] and p. 159 of Ott [21]. In fact, Morse [19] in (11) on p. 260 obtains this result, but to develop a rough approximation for the rate of approach to steady state he works instead with the cosine transform. His relaxation time constant of ρ mentioned in Section 2.4 of [3] seems to be based on an approximation for this cosine transform. Of course, the dominant portion of the relaxation time appears in the time scaling (3.1). Morse's result and all the different versions described in Table 4 of [3] capture this first-order effect. ■

Aside from the normalization by the stationary variance, the integral in Corollary 7 yields the limiting average variance of the integral and the normalization in the central limit theorem. We review these (known) results in Corollaries 8-10 below. See Iglehart [11] for generalizations and previous work.

Corollary 8. For all initial states i,

$$\lim_{t\to\infty} t^{-1}\, Var\left[\int_0^t [Q(2(1-\rho)^{-2}s) \mid Q(0) = i]\, ds\right] = 2\, Var Q(\infty)\int_0^\infty c_q(t)\,dt$$

$$= \frac{2\rho}{(1-\rho)^2}\int_0^\infty c_q(t)\,dt = \frac{\rho(1+\rho)}{(1-\rho)^2}.$$

so that

$$\lim_{t\to\infty} t^{-1}\, Var\left[\int_0^t [Q(s)\,|\,Q(0) = i]ds\right] = \frac{2\rho(1+\rho)}{(1-\rho)^4}\,.$$

Let $R(t,x)$ denote RBM starting at x.

Corollary 9. $\lim_{t\to\infty} t^{-1}\, Var\left[\int_0^t R(s,x)\, ds\right] = 2\, VarR(\infty) \int_0^\infty c(t)dt = 1/2.$

Remark 3.4. As in Section 5 of Reynolds [24], (3.16) of Cox and Isham [9], Whitt [30] and Woodside et al. [31], Corollaries 8 and 9 are useful to determine the asymptotic efficiency of the natural estimator $t^{-1}\int_0^t Q(s)ds$ of the steady state mean. For example, we can easily determine how one sample of length t for large t compares to n i.i.d observations of the stationary queue length $Q(\infty)$. ■

We can apply Corollaries 8 and 9 together with additional structure (regenerative or stationary and ϕ-mixing) to obtain central limit theorems for M/M/1 and RBM. Let => denote weak convergence (convergence in distribution) as in Billingsley [6].

Corollary 10. As $t\to\infty$,

(a) $$t^{-1/2}\left[\int_0^t [Q(2(1-\rho)^{-2}s)\,|\,Q(0)=i)ds - \rho(1-\rho)^{-1}t\right] \Rightarrow N(0,\sigma^2)$$

for any i, where $\sigma^2 = \rho(1+\rho)/(1-\rho)^2$.

(b) $$t^{-1/2}\left[\int_0^t R(s,x)\, ds - t/2\right] \Rightarrow N(0,\ 1/2)$$

for any x.

Remarks 3.5. Corollary 10 can be extended to functional central limit theorems in the usual way; see Theorem 20.1 of [6] and [11]. Results paralleling Corollaries 8-10 hold for the other processes in this paper, but they will not be discussed.

4. The Number in Service

Let $S(t)$ denote the number in service (which is either 1 or 0); let $m_{sk}(t, i)$ denote the k^{th} moment of the number in service at time t given that $Q(0) = i$; let $S_k(t)$ be the associated k^{th}-moment cdf defined as in (3.3), i.e., starting at the origin; let $c_s(t)$ be the correlation function; let $B(t)$ be the busy-period cdf; and let the time scale be as in (3.1). The following theorem combines elementary calculations with Corollary 4.2.3 of [4], which establishes that $P_{00}(t) = 1 - \rho B(t)$.

Theorem 2. For each $k \geqslant 1$, $i \geqslant 0$ and $t \geqslant 0$,

(a) $m_{sk}(t, i) = 1 - P_{i0}(t)$,

(b) $S_k(t) = [1-P_{00}(t)]/\rho = B(t)$,

(c) $c_s(t) = B^c(t)$,

(d) $\bar{c}_s \equiv \int_0^\infty c_s(t)dt = \theta = (1-\rho)/2$.

Corollary 1. The functions $S_k(t)$ and $1 - c_s(t)$ are bonafide cdf's with completely monotone densities.

Proof. Theorem 3.3 of [3]. ■

Remark 4.1. All moments of the busy-period cdf $B(t)$ and thus also the cdf's $S_k(t)$ and $1 - c_s(t)$ are given in Theorem 3.2 of [4] and Section 7 here.

5. The Number Waiting

Let $L(t)$ denote the number in line or the number waiting, defined by $L(t) = Q(t) - S(t)$, $t \geqslant 0$. Let $m_{lk}(t, i)$ denote the k^{th} moment at time t given that $Q(0) = i$; let $U_k(t) = m_{lk}(t, 0)/m_{lk}(\infty)$ be the associated k^{th}-moment cdf; and let $c_l(t)$ be the correlation function. As before, let the time scale be as in (3.1).

To determine the moment cdf's $U_k(t)$, we express the moments of $L(t)$ in terms of the moments of $Q(t)$.

Lemma 1. For each $k \geqslant 1$, $i \geqslant 0$, and $t \geqslant 0$,

$$m_{lk}(t, i) = \sum_{j=0}^{k-1} \binom{k}{j} (-1)^j \, m_{k-j}(t, i) + (-1)^k [1 - P_{i0}(t)] .$$

Proof. Note that $L(t)^k = [Q(t) - S(t)]^k$, $E[Q(t)^i S(t)^j] = E[Q(t)^i]$ for $i > 0$ and $E[S(t)^j] = ES(t)$ for $j > 0$. ■

Let $*$ denote convolution. As before, let $\theta = (1-\rho)/2$.

Theorem 3. The first two moment densities of $L(t)$ are

(a) $u_1(t) = [h_1(t) - 2\theta b(t)]/\rho = b(t) * h_1(t)$.

(b) $u_2(t) = [(1-\theta) h_2(t) + 2\theta^2 b(t) - 2\theta h_1(t)]/\rho(1-\theta) = b(t) * h_2(t)$.

Proof. The first parts follow easily from Lemma 1. By Corollary 5.2.1 of [4], the transforms in (a) are related by $\hat{h}_1(s) = \hat{f}_{\epsilon 0}(s) \hat{b}(s)$, so that $\hat{u}_1(s) = \hat{b}(s)[\hat{f}_{\epsilon 0}(s) - 2\theta]/\rho$. However, $[\hat{f}_{\epsilon 0}(s) - 2\theta]/\rho$ is easily seen to be the Laplace Stieltjes transform of the equilibrium time to emptiness conditional on not starting empty. By Corollary 3.1.3 of [3], this is $\hat{h}_1(s)$. A similar argument yields (b). ■

Corollary 1. The functions $U_1(t)$ and $U_2(t)$ are bonafide cdf's.

Corollary 2. The complementary moment cdf's $U_1^c(t)$ and $U_2^c(t)$ satisfy

$$U_1^c(t) = \rho^{-1} [H_1^c(t) - 2\theta B^c(t)] = 1 - \frac{t^2}{4\theta^3} + o(t) \quad \text{as } t \to 0$$

and

$$U_2^c(t) = (\rho(1-\theta))^{-1} \, [V_2^c(t) - 3\theta H_1^c(t) + 2\theta^2 B^c(t)]$$

$$= 1 - \frac{t^2}{4\theta^3 (1-\theta)} + o(t) \quad \text{as } t \to 0 .$$

so that $U_1^c(t)$ and $U_2^c(t)$ are *not* log-convex.

Proof. We give the details only for $U_1^c(t)$. Since

$$H_1^c(t) = 1 - \frac{t}{\theta} + \frac{t^2}{4\theta^3} + o(t^2) \quad \text{as } t \to 0$$

and

$$B^c(t) = 1 - \frac{t}{2\theta^2} + \frac{(1-\theta)t^2}{4\theta^4} + o(t^2) \quad \text{as } t \to 0\,,$$

e.g., see Section 7,

$$U_1^c(t) = \rho^{-1}[H_1^c(t) - 2\theta B^c(t)] = 1 - \frac{t^2}{4\theta^3} + o(t) \quad \text{as } t \to 0$$

and

$$u_1(t) = \frac{t}{2\theta^3} + o(t) \quad \text{as } t \to 0\,. \quad \blacksquare$$

Corollary 3. The first four moments of $U_1(t)$ are $m_1 = (2-\rho)/2$, $m_2 = (3-\rho)/2$, $m_3 = 3(4+2\rho-\rho^2)/4$ and $m_4 = 3(5+10\rho-\rho^3)/2$.

Remark 5.1. Odoni and Roth [20], Lee [17] and Lee and Roth [18] develop empirical approximations for the expected number waiting, and thus $U_1(t)$. Lee [17] and Lee and Roth [18] empirically discovered that $U_1(t)$ is not log-convex, but that difficulties appear only for very small ρ. From the structure established here and in [1]-[4], it seems more natural to use hyperexponential and exponential approximations for the completely monotone functions $H_1^c(t)$ and $B^c(t)$, and then use the composite approximation for $U_1(t)$ based on Corollary 2 above; see Section 8. $\blacksquare$

Lemma 1 allows us to represent $U_k(t)$ in terms of $[H_k(t), H_{k-1}(t), \ldots, H_1(t), B(t)]$ for each k. Theorem 3 leads us to make the following conjecture (which we have verified for $k = 3$, 4, and 5; see Appendix B).

Conjecture 1. For all $k \geq 1$, $u_k(t) = b(t) * h_k(t)$.

We now turn to the correlation function $c_l(t)$. Recall that $Var\, L(\infty) = \rho^2[1+2\theta-(2\theta)^2]/4\theta^2 = \rho^2(1+\rho-\rho^2)/(1-\rho)^2$. We express

$c_l(t)$ in terms of the correlation function $c_w(t)$ and the second-moment cdf $V_2(t)$ for the workload, to be treated in Section 6. We omit our proof (using Laplace transforms).

Theorem 4. The correlation function $c_l(t)$ can be expressed as

$$1 - c_l(t) = \frac{(1+2\theta)[1-c_w(t)]-4\theta^2 U_1(t)}{1+2\theta-4\theta^2} = \frac{V_2(t)-8\theta^2 H_1(t)+8\theta^3 B(t)}{1-8\theta^2(1-\theta)}$$

$$= \frac{\rho(U_1 * H_1)(t) + 4\theta\rho U_1(t)+4\theta^2 B(t)}{\rho+4\theta\rho+4\theta^2}.$$

Corollary 1. $1 - c_l(t)$ is a bonafide cdf.

Corollary 2. As $\rho \to 1$, $1 - c_l(t)$, $1 - c_w(t)$ and $V_2(t) = 1 - c_q(t)$ coincide (all approach $1 - c(t)$).

Corollary 3. $\bar{c}_l \equiv \int_0^\infty c_l(t)dt = (1+4\rho - 4\rho^2 + \rho^3)/2(1+\rho-\rho^2)$ and

$$\bar{C}_l \equiv Var\, L(\infty)\, \bar{c}_l = \rho^2(1+4\rho-4\rho^2+\rho^3)/2(1-\rho)^2.$$

Corollary 4. The first three moments of $1 - c_l(t)$ are $m_1 = [1+\rho(2-\rho)^2]/2(1+\rho-\rho^2)$, $m_2 = [1+\rho(3-\rho)^2]/2(1+\rho-\rho^2)$ and $m_3 = 3[1+\rho(16+\rho-5\rho^2+\rho^3)]/4(1+\rho-\rho^2)$.

From Theorem 4 and Corollary 2 to Theorem 3, it would appear that $c_l(t)$ might not be completely monotone, but in Section 7 we show that it actually is.

6. The Workload

Let $W(t)$ represent the workload or virtual waiting time at time t. For the M/M/1 model, $W(t)$ is distributed for each t as the random sum $v_1 + \dots + v_{Q(t)}$ where $\{v_k: k \geqslant 1\}$ is a sequence of i.i.d. service times that are independent of $Q(t)$. Hence, the k^{th} moment of $W(t)$ given i initial customers is easily computed from the moment functions for $Q(t)$, i.e.,

$m_j(t,i)$ for $1 \leqslant j \leqslant k$. However, we shall consider $m_{wk}(t, x)$, the k^{th} moment of the workload at time t given that $W(0) = x$. Of course, these moment functions coincide when $i = 0$ and $x = 0$. Let $V_k(t) = m_{wk}(t, 0)/m_{wk}(\infty)$ be the k^{th} moment cdf of $W(t)$ and let $c_w(t)$ be the correlation function. Let the time scaling be as in (3.1). Since the service time and waiting time are times, they are scaled by $2\theta^2$ too; e.g., $m_{w1}(t, 0) = 2\theta^2 m_1(t, 0)$.

As indicated above, it is not difficult to obtain the moment cdf's of $W(t)$ by applying previous results for $Q(t)$. Note that the second moment cdf $V_2(t)$ coincides with $1-c_q(t)$.

Theorem 5. $V_1(t) = H_1(t)$ and

$$V_2(t) = V_{1e}(t) = H_{1e}(t) = \theta H_1(t) + (1-\theta)H_2(t).$$

Corollary 1. $V_1(t)$ and $V_2(t)$ are bonafide cdf's with completely monotone densities.

More generally, paralleling Section 9 of [2] and Section 2 of [4], the double Laplace transform of $(W(t)|W(0) = x)$ admits a factorization. As before, one term is the special case in which the initial state is $x = 0$. Let

$$\bar{g}(\sigma,s;x) = \int_0^\infty e^{-st} E(e^{-\sigma[W(t)|W(0)=x]})dt,$$

$$\hat{G}(y,s;x) = \int_{0-}^\infty e^{-st} P(W(t) \leqslant y|W(0)=x)dt \tag{6.1}$$

and $\hat{g}(y,s;x)$ be the density of $\hat{G}$ (in the generalized distribution sense, i.e., there typically is an atom at the origin).

Theorem 6. $\bar{g}(\sigma,s;x) = \bar{g}(\sigma,s;0)\,\bar{d}(\sigma,s;x)$, where

$$\bar{g}(\sigma,s;0) = \frac{(r_1/2\theta)(1+2\theta^2\sigma)}{s(\sigma+[r_1/2\theta])} \text{ and } \bar{d}(\sigma,s;x) = \frac{\sigma e^{-(xr_2/2\theta)} - (r_2/2\theta)e^{-x\sigma}}{\sigma-(r_2/2\theta)}$$

where r_1 and r_2 are given in (2.4) of [4].

Let $\delta(y)$ be the Dirac delta function, corresponding to an atom at the origin. We obtain the time transform $\hat{g}(y,s;\,0)$ by inverting $\bar{g}(\sigma,s;\,0)$ directly by inspection.

Corollary 1. $\hat{g}(y,s;\,0) = s^{-1}\left[\theta r_1\delta(y)+(1-\theta r_1)(r_1/2\theta)e^{-yr_1/2\theta}\right].$

Remark 6.2. In Section 10 of [4] we proposed an operational calculus for determining M/M/1 queue-length quantities from corresponding quantities for RBM. Since $W(t)$ has a continuous state space, the connection is even more direct for it: simply let the M/M/1 workload state be $w=2\theta x$ where x is the RBM state, and let the M/M/1 quantities r_1 and r_2 be the same as the RBM quantities, with the understanding that for RBM r_1 and r_2 solve one quadratic equation, while for M/M/1 $z_1=(1-\theta r_1)/\rho$ and $z_2=(1+\theta r_2)/\rho$ solve another. This prescription works for the downward first passage times, but fails to capture the atom at 0 in Corollary 1.

Paralleling (4.2) of [4], Corollary 1 shows that the time transform has the same form as the steady-state distribution, namely, an exponential density plus an atom at the origin. We obtain the steady-state distribution by an elementing limit. Let $g(y)$ be the density of $W(\infty)$, again in the generalized distribution sense.

Corollary 2. $g(y) = \lim_{s\to 0} s\hat{g}(y,s;\,0) = (1-\rho)\delta(y) + \rho\theta^{-1}e^{-y\theta^{-1}}.$

However, more important for our purposes, we obtain all moments $m_{wk}(t,0)$ and thus all moment cdf's from Corollary 1. We combine Corollary 1 and Theorems 3.1(a) and 3.3 of [4] to characterize the Laplace transform of the k^{th}-moment density $v_k(t)$. Let $\hat{f}_{e0}(s)$ be the Laplace-Stieltjes transform of the cdf of the equilibrium time to emptiness.

Corollary 3. $\hat{m}_{wk}(s,0) = k!(1-\theta r_1)\left(\dfrac{2\theta}{r_1}\right)^k s^{-1}$, so that

$$\hat{v}_k(s) = \left(\frac{1-\theta r_1}{\rho}\right)\left(\frac{2}{r_1}\right)^k = \hat{b}(s)\hat{f}_{e0}(s)^k,\ k \geqslant 1\ .$$

Remark 6.1. We can also obtain Theorem 5 directly from Corollary 3 above plus Corollary 5.2.1 and Theorem 9.1 of [4].

From Corollary 3, we can obtain any desired moment of $V_k(t)$ for any k.

Corollary 4. The mean of $V_k(t)$ is $\theta + k\rho/2$, $k \geqslant 1$.

We now obtain expressions for $V_k(t)$ for all $k \geqslant 3$. This relation generalizes Theorem 1.5 of [1] for RBM (i.e., reduces to it when $\rho = 1$).

Theorem 7.
$v_{k+1}(t) = [(k-1)\rho + 1]v_{ke}(t) - \rho[v_2(t) + ... + v_k(t)]$, $t \geqslant 0$, $k \geqslant 2$.

Proof. Applying (2.4), (2.5) and Theorem 3.3 of [4], we obtain the following relations for the Laplace transforms $\hat{v}_k \equiv \hat{v}_k(s)$:

$$
\begin{aligned}
[(k-1)\rho + 1]\,\hat{v}_{ke} &= 2s^{-1}(1-\hat{v}_k) = 2s^{-1}(1-z_1\hat{f}_{\epsilon 0}^{\,k}) \\
&= 2z_1 s^{-1}\left[z_1^{-1}-\hat{f}_{\epsilon 0}^{\,k}\right] = 2z_1 s^{-1}\left[\rho z_2-\hat{f}_{\epsilon 0}^{\,k}\right] \\
&= 2z_1 s^{-1}\left[1-\hat{f}_{\epsilon 0}^{\,k}+\theta r_2\right] \\
&= 2z_1 s^{-1}\left[1-\hat{f}_{\epsilon 0}\right]\left[1+\hat{f}_{\epsilon 0} + ... + \hat{f}_{\epsilon 0}^{\,k-1}\right] + 2\theta z_1 r_2 s^{-1} \\
&= z_1^2\rho\hat{f}_{\epsilon 0}^{\,2}\left[1+\hat{f}_{\epsilon 0} + ... + \hat{f}_{\epsilon 0}^{\,k-1}\right] + 2\theta\hat{v}_1 \\
&= \rho\hat{v}_1\,(\hat{v}_1 + ... + \hat{v}_k) + 2\theta\hat{v}_1 \\
&= (\hat{f}_{\epsilon 0}-2\theta)(\hat{v}_1 + ... + \hat{v}_k) + 2\theta\hat{v}_1 \\
&= \rho(\hat{v}_2 + ... + \hat{v}_k) + \hat{v}_{k+1}\,. \quad \blacksquare
\end{aligned}
$$

We now turn to the correlation function $c_w(t)$ and the associated covariance function $C_w(t)$. Note that $Var\ W(\infty) = (1-4\theta^2)\theta^2$ in the time scale (3.1). (The time scaling $2\theta^2$ appears twice here.)

Theorem 8. $1 - c_w(t) = U_{1e}(t) = [V_2(t)-4\theta^2 H_1(t)]/(1-4\theta^2)$.

Proof. Let $C_w(t)$ be the covariance of $W(t)$. The Laplace transform satisfies

$$s\hat{C}_w(s) = 2\theta s\hat{m}_{w1}(s,0) + \int_0^\infty s\rho\theta^{-1}e^{-x\theta^{-1}}\hat{m}_{w1}(s,x)dx - (\rho\theta)^2,$$

where

$$s\hat{m}_{w1}(s,x) = x - 2\theta s^{-1} + (2\theta/r_2)e^{-r_1x/2\theta}$$

by virtue of the conservation law (cf. Theorem 8.1 of [4] and Theorem 5 of Benes [5])

$$m'_{w1}(t,x) = P_{x0}(t) - (1-\rho) \text{ and } \hat{P}_{x0}(s) = \hat{P}_{00}(s)\hat{f}(s;x,0) = \left(\frac{2\theta}{r_2}\right)e^{-r_1x/2\theta}$$

where $f(t;x,0)$ is the first-passage time density. (The argument parallels the proof of Theorem 3.1 (a) of [4]: To be at 0 at time t starting at x you must reach 0 for a first time somewhere in the interval $[0,t]$. Alternatively, see p. 260 of Cohen [7].) After integrating and rearranging terms, we get

$$s\hat{C}_w(s) = C_w(0)\left[1-\frac{2}{2-\rho}\left(s^{-1}-\frac{4}{r_2(2+r_2)^2}\right)\right].$$

Since $z_1 = r_1/(r_2+2)$,

$$\frac{4}{r_2(2+r_2)^2} = \left(\frac{2}{r_2r_1}\right)\left(\frac{2}{r_2}\right)z_1^2 = s^{-1}\hat{h}_1(s)\hat{b}(s);$$

see (2.4), (2.5) and Theorems 3.1 and 5.1 of [4]. Hence,

$$s\hat{C}_w(s) = C_w(0)[1-SE(\hat{h}_1(s)\hat{b}(s))] = C_w(0)\,[1-SE(\hat{u}_1(s))]. \quad \blacksquare$$

Now that we have derived all the correlation functions, we summarize the expressions in Table 3.

Corollary 1. (Ott) $c_w(t)$ is monotone but not log-convex, and thus not completely monotone. In particular,

$$C_w(t) = C(0) - 2\rho\theta^2 t + \rho t^3/6\theta + o(t^3) \quad as \quad t \to 0,$$

so that $C_w'''(0) > 0$.

TABLE 3

The correlation functions. (SE represents the stationary-excess operator in (3.5).)

the process	correlation cdf	expression	result
M/M/1			
Number in System	$1 - c_q(t)$	$SE(H_1) = (SE)^2(B) = V_2$	Theorems 1 and 5
Number in Service	$1 - c_s(t)$	$B = S_1 = S_k$	Theorem 2
Number in Line	$1 - c_l(t)$	$a_1 V_2 - a_2 H_1 + a_3 B$	Theorem 4
Workload	$1 - c_w(t)$	$SE(U_1) = b_1 V_2 - b_2 H_1$	Theorem 8
RBM	$1 - c(t)$	$SE(H_1) = H_2$	Theorem 1 and Remark 3.1

Remark 6.3. From Corollary 1, we see that $C_w'(0) = -2\rho\theta^2$ and $\lim_{t \to 0} t^{-1} C_w''(t) = \rho/\theta$. Without time scaling, these values are ρ and $2\theta\rho = \rho - \rho^2$, which agree with (2.13) and (2.15) on p. 161 of Ott [21].

Corollary 2. (Ott) $\bar{c}_w \equiv \int_0^\infty c_w(t)dt = \frac{1+\theta}{1+2\theta} = \frac{3-\rho}{2(2-\rho)}$.

Proof. From Theorem 8, we see that $\bar{c}_w$ is the mean of U_{1e}, which in turn is the second moment of U_1 divided by twice the first moment. Alternatively, we can use the means of $V_2(t)$ and $H_1(t)$, namely, $(1-\theta)$ and $1/2$:

$$\int_0^\infty c_w(t)dt = \left[\int_0^\infty [1-V_2(t)]dt - 4\theta^2 \int_0^\infty [1-H_1(t)]dt\right] /(1-4\theta^2)$$

$$= [(1-\theta)-4\theta^2(1/2)]/(1-4\theta^2) = (1+\theta)/(1+2\theta) . \quad \blacksquare$$

Remarks 6.3. (a) In scaled time,

$$\int_0^\infty C_w(t)dt = Var\ W(\infty) \left[\frac{1+\theta}{1+2\theta}\right] = \rho(3-\rho)\theta^2/2 .$$

Without time scaling,

$$\int_0^\infty C_w(t)dt = [\rho(3-\rho)\theta^2/2](1/2\theta^2)^3 = \frac{\rho(3-\rho)}{(1-\rho)^4},$$

in agreement with (3.15) on p. 563 of Cox and Isham [9].

(b) Since $W(t)$ is Markov even for the more general M/G/1 model, it is natural to expect that many of the results in this section can be extended to M/G/1. This is indeed so and is intended to be the topic of a future paper. The situation for M/G/1 seems sufficiently more complex to justify a separate treatment of M/M/1.

7. Exploiting the Stationary-Excess Relations

In Sections 3-6 we have shown that all quantities of interest can be simply expressed in terms of the basic three cdf's $B(t)$, $H_1(t)$ and $V_2(t)$, where $H_1(t)$ and $V_2(t)$ in turn can be simply expressed in terms of the busy-period cdf $B(t)$. In particular, $H_1(t)$ and $V_2(t)$ are associated stationary-excess cdf's: $H_1(t) = B_e(t) = SE(B)$ and $V_2(t) = H_{1e}(t) = (SE)^2(B)$. In this section we exploit these relations together with properties of the busy-period cdf. First, we remark that these relations can be used to describe the asymptotic behavior as $t \to \infty$ of all these moment-function cdf's and correlation functions. They are all asymptotically of the form $At^{-3/2}e^{-t/\tau}$ where $\tau = (1+\sqrt{\rho})^2/2$; apply Theorem 3.1 (b) and (c) and Corollary 5.2.3 of [4].

In this section, we exploit the fact that the stationary-excess operator acts as a simple shift on the moments and derivatives. For a cdf $G(t)$ without any atoms, suppose that m_k is the k^{th} moment and d_k the k^{th} derivative of the cdf $G(t)$ at $t = 0$; e.g., $d_1 = g(0)$ where $g(t)$ is the density of $G(t)$. Then the k^{th} moment m_{ek} and k^{th} derivative at $t = 0$, d_{ek}, of the associated stationary-excess cdf $G_e(t)$ defined in (3.5) are

$$m_{ek} = m_{k+1}/(k+1)m_1, \qquad d_{e1} = 1/m_1 \qquad \text{and} \qquad d_{e(k+1)} = -d_{e1} d_k, \; k \geq 1$$

As a consequence, specifying the first i moments m_k and the first j derivatives d_k of the cdf $G(t)$ is equivalent to specifying the first $(i-1)$ moments m_{ek} and the first $(j+1)$ derivatives d_{ek} of the cdf $G_e(t)$.

We thus determine all moments and all derivatives of all moment cdf's and correlation functions by determining all moments and derivatives of $B(t)$. All moments of $B(t)$ were given in Theorem 3.2 of [4]. We now give all derivatives of $B(t)$ at $t = 0$ in terms of the moments. Let $f^{(k)}(0)$ be the k^{th} derivative of $f(t)$ at $t = 0$.

Theorem 9. For all $k \geq 0$, $(-1)^k b^k(0) = m_{k+2}/(k+2)!\theta^{2k+3}$, so that

$$c^2 \equiv (m_2 - m_1^2)/m_1 = (1-\theta)/\theta = (1+\rho)/(1-\rho)$$

and

$$B^c(t) = 1 + \sum_{k=0}^{\infty} (-1)^{k+1} \frac{m_{k+2}}{(k+2)!\theta^{2k+3}} \left[\frac{t^{k+1}}{(k+1)!} \right].$$

Proof. We relate two representations for the Laplace transform $\hat{b}(s)$. From (2.4) and Theorem 3.1 of [4],

$$\hat{b}(s) = z_1 = \theta\rho^{-1} \left[(1-\theta)\theta^{-1} + \theta s - [1+2(1-\theta)s+\theta^2 s^2]^{1/2} \right]$$
$$= \frac{\theta^2 s}{\rho} \left[\frac{1-\theta}{s\theta^2} + 1 - \left[1 + \frac{2(1-\theta)}{s\theta^2} + \frac{1}{s^2\theta^2} \right]^{1/2} \right]. \tag{7.2}$$

The first representation in terms of the derivatives is $\hat{b}(s) = \sum_{k=0}^{\infty} b^{(k)}(0) s^{-(k+1)}$, which comes from the Taylor series expansion $b(t) = \sum_{k=0}^{\infty} b^{(k)}(0) t^k/k!$ The second representation is the usual series involving the moments $\hat{b}(\xi) = \sum_{n=0}^{\infty} (-1)^n m_n \xi^n/n!$ (We have changed the argument from s to ξ.) The quadratic form in (7.2) enables us to easily relate these representations after setting $\xi = 1/s\theta^2$. Then

$$\frac{\theta}{\rho}\left[\frac{1-\theta}{\theta} + \frac{1}{\theta s} - \left[1 + \frac{2(1-\theta)}{s\theta^2} + \frac{1}{s^2\theta^2}\right]^{1/2}\right] = \sum_{n=0}^{\infty}(-1)^n \frac{m_n}{n!}\left[\frac{1}{s\theta^2}\right]^n$$

and

$$-\frac{s\theta^2}{\rho}\left[1 + \frac{2(1-\theta)}{s\theta^2} + \frac{1}{s^2\theta^2}\right]^{1/2} = \theta s \sum_{n=0}^{\infty}(-1)^n \frac{m_n}{n!}\left[\frac{1}{s\theta^2}\right]^n$$

$$-\frac{\theta(1-\theta)s}{\rho} - \frac{\theta}{\rho}.$$

Substituting this into (7.2), we get

$$\hat{b}(s) = \frac{1-2\theta}{\rho} - \frac{(1-2\theta)}{\rho}\theta s + \theta s \sum_{n=0}^{\infty}(-1)^n \frac{m_n}{n!}\left[\frac{1}{s\theta^2}\right]^n$$

$$= \sum_{n=2}^{\infty}(-1)^n \frac{m_n}{n!\theta^{2n-1}s^{n-1}} = \sum_{n=0}^{\infty}(-1)^n \frac{m_{n+2}}{(n+2)!\theta^{2n+3}}\left[\frac{1}{s^{n+1}}\right]. \quad \blacksquare$$

Corollary 1. In the time scale (3.1), the parameters of $B(t)$ are $m_0 = 1$, $m_1 = m_2 = \theta$, $m_3 = 3\theta(1-\theta)$, $m_4 = 3\theta[4(1-\theta)^2+\rho]$, $m_5 = 15(1-\theta)\theta[1+5\rho+\rho^2]$, $b(0) = 1/2\theta^2$, $-b^{(1)}(0) = (1-\theta)/2\theta^4$ and $b^{(2)}(0) = 3\theta(4(1-\theta)^2+\rho)/24\theta^7$.

Corollary 2. For the first-moment cdf $H_1(t)$, $(-1)^k h_1^{(k)}(0) = m_k/k!\theta^{2k+1}$, $k \geqslant 0$, where m_k is the k^{th} moment of $H_1(t)$ given in Corollary 5.2.4 of [4], e.g., $m_0 = 1$, $m_1 = 1/2$, $m_2 = (1-\theta)$, $m_3 = 3(1+3\rho+\rho^2)/4$ and $c^2 = 1 + 2\rho$.

Corollary 3. The first three derivatives of $H_1(t)$ at $t = 0$ are $h_1(0) = \theta^{-1}$, $h_1^{(1)}(0) = -(2\theta^3)^{-1}$ and $h_1^{(2)}(0) = (1-\theta)/2\theta^5$.

Corollary 4. For $V_2(t)$, $v_2(0) = 2$, $v_2^{(1)}(0) = -2/\theta$ and $(-1)^k v_2^{(k)}(0) = m_{k-2}/(k-2)!\theta^{2k-1}$, $k \geqslant 2$, where m_k is the k^{th} moment of $V_2(t)$, e.g., $m_0 = 1$, $m_1 = (1-\theta)$, $m_2 = (1+3\rho+\rho^2)/2$, $m_3 = 3(1+\rho)(1+5\rho+\rho^2)/4$ and $c^2 = 1 + \rho/(1+\rho)$.

Corollary 5. The first four derivatives of $V_2(t)$ at $t = 0$ are $\nu_2(0) = 2$, $\nu_2^{(1)}(0) = -2\theta^{-1}$, $\nu_2^{(2)}(0) = \theta^{-3}$, and $\nu_2^{(3)}(0) = -(1-\theta)\theta^{-5}$.

Corollary 6. If m_k is the k^{th} moment of the busy-period cdf $B(t)$ in (A-1) of Appendix A, then

$$H_1^c(t) = 1 + \sum_{k=0}^{\infty} (-1)^{k+1} \frac{m_{k+1}}{(k+1)!\theta^{2k+2}} \left[\frac{t^{k+1}}{(k+1)!} \right]$$

and

$$V_2^c(t) = 1 + \sum_{k=0}^{\infty} (-1)^{k+1} \frac{2m_k}{k!\theta^{2k}} \left[\frac{t^{k+1}}{(k+1)!} \right] .$$

Remarks 7.1 (a) As reflected by the c^2 values for $B(t)$, $H_1(t)$ and $V_2(t)$, these cdf's are "smoothed" by the stationary-excess operator, i.e., they become more nearly exponential: $c^2_{B(t)} \geqslant c^2_{H_1(t)} \geqslant c^2_{V_2(t)} \geqslant 1$, see [29]. (However, the stationary-excess operator does not always make c^2 decrease for completely monotone densities.) As $\rho \to 0$, $c^2 \to 1$ in each case too.

(b) The transforms of the three basic complementary cdf's have a nice exponential-like form, i.e.,

$$\hat{B}^c(s) = \frac{1}{s+(r_1/2\theta)}, \qquad \hat{H}_1^c(s) = \frac{1}{s+r_1}, \qquad \hat{V}_2^c(s) = \frac{1}{s+\left[\dfrac{2r_1}{2+\rho r_1}\right]}$$

where of course r_1 is a function of s. In each case, with $\hat{G}^c(s) = 1/(s+\alpha(s))$, $\alpha(s) \to m_1^{-1}$ as $s \to 0$ as it should.

We now combine Theorems 4 and 9 to obtain a Taylor series representation of the correlation function $c_l(t)$ for the number waiting, from which we establish that $c_l(t)$ is completely monotone.

Theorem 10. $c_l(t) = 1 + \sum_{k=0}^{\infty} (-1)^{k+1} a_k \left[\dfrac{2}{(k+2)!\theta^{2k}} \right] \left[\dfrac{t^{k+1}}{(k+1)!} \right]$ where

$$a_k = \frac{(k+2)(k+1)m_k - 4(k+2)m_{k+1} + 4m_{k+2}}{\rho(1+\rho-\rho^2)} > 0$$

for all k, with m_k being the k^{th} moment of $B(t)$, so that $c_l(t)$ is completely monotone.

Proof. From Theorem 4, $c_l(t)$ is a linear combination of $B^c(t)$, $H_1^c(t)$ and $V_2^c(t)$. The Taylor series representation thus follows from the corresponding Taylor series representations for $B^c(t)$, $H_1^c(t)$ and $V_2^c(t)$ determined by Theorem 9 and Corollary 6, together with (7.1). Nonnegativity of a_k is proved in Appendix A. ■

8. Hyperexponential Approximations

The representations in Sections 3 through 6 are ideal for developing relatively simple approximations for the M/M/1 correlation functions and moment functions. In particular, we suggest developing approximations in two steps. In the *first step* we produce hyperexponential (e.g., H_2) approximations for the three basic cdf's $B(t)$, $H_1(t) = V_1(t)$ and $V_2(t) = 1 - c_q(t)$. Since the basic three cdf's $B(t)$, $H_1(t)$ and $V_2(t)$ are all mixtures of exponentials, it is appropriate to use hyperexponential approximations. We propose doing an H_2 fit (two exponentials), but of course it is also possible to use more exponentials. Using more exponentials and more information typically leads to better approximations. As shown in [28], approximations by moment matching can also be interpreted as bounds within the full hyperexponential class (to which $B(t)$, $H_1(t)$ and $V_2(t)$ belong). The use of hyperexponential approximations follows [1]-[3] and is in the same spirit of much earlier work by Riordan [25]; e.g., pp. 106-108.

In the *second step* we obtain other approximations of interest by exploiting expressions for these quantities in terms of the basic three cdf's $B(t)$, $H_1(t)$ and $V_2(t)$. Note that all the correlation functions ($c_q(t)$, $c_s(t)$, $c_l(t)$ and $c_w(t)$) and all the moment functions ($H_k(t)$, $S_k(t)$, $U_k(t)$ and $V_k(t)$) can be expressed directly in terms of the three basic cdf's.

It is also significant that there is a very strong connection among the basic three cdf's $B(t)$, $H_1(t)$ and $V_2(t)$. In particular, $H_1(t)$ is the stationary-excess cdf of $B(t)$ and $V_2(t)$ is the stationary-excess cdf of $H_1(t)$. Moreover, the stationary-excess operator maps an H_k density

$f(t) = \sum_{i=1}^{k} p_i \lambda_i e^{-\lambda_i t}$, $t \geq 0$, into another H_k density with the same exponential parameters, i.e., $f_e(t) = \sum_{i=1}^{k} q_i \lambda_i e^{-\lambda_i t}$ where $q_i = p_i \lambda_i^{-1} / \sum_{i=1}^{k} p_i \lambda_i^{-1}$. Consequently, an H_k approximation for any one of these three basic cdf's immediately provides corresponding H_k approximations for all three basic cdf's.

We propose constructing hyperexponential approximations by matching the first i moments and the first j derivatives of the cdf at the origin. For an H_k fit, $2k - 1$ parameters need to be determined. The case of principal interest is H_2 with three parameters. As noted in Section 7, it is significant that the stationary-excess operator acts as a shift on these parameters. Since the stationary-excess cdf of an H_k distribution is again H_k with the same exponential parameters, there is a one-to-one correspondence between H_k distributions partially specified in this way. As a result of this one-to-one correspondence, any desired H_k approximation for $B(t)$, $H_1(t)$ and $V_2(t)$ can be obtained from the appropriate moments m_k and derivatives $b^{(k)}(0)$ of the busy-period cdf $B(t)$.

Given the parameters above, we can fit H_2 cdf's in the manner described in Section 5 of [1]. Because of the stationary-excess relations, we can always work with three moments. It turns out that all the H_2 fits for $B(t)$, $H_1(t)$ and $V_2(t)$ can be expressed explicitly as functions of ρ, many in remarkably simple form. The various H_2 approximations are displayed in Table 4. Associated numerical values for the case $\rho = 0.75$ are also given in Table 5. The H_2 approximations are developed for all combinations of parameter triples ranging from the first three derivatives of the cdf to the first three moments (four cases for each cdf). Also displayed for each cdf are the one-moment exponential fit and the two-moment H_2 bounding cdf with an atom at 0 (corresponding a degenerate exponential with mean 0) obtained from (6) of [28]. This cdf matches the first two

TABLE 4

Hyperexponential (H_2) approximations for the basic three moment cdf's.

Note. In the three-moment fit to $H_1^c(t)$ and the $[g(0), m_1, m_2]$-fit to $V_2^c(t)$ below, $\lambda_1 = (2+\rho+[5-(1-\rho)(5+\rho)]^{1/2})/4$; see (2.2) and (2.7) of [3].

	the three basic complementary cdf's		
Type of Approximation	$B^c(t)$	$H_1^c(t)$	$V_2^c(t) = c_q(t)$
One-Moment Exponential Fit	$e^{-t/\theta}$ $m_1 = \theta$	e^{-2t} $m_1 = 1/2$	$e^{-t/(1-\theta)}$ $m_1 = 1-\theta$
Two-Moment Bounding H_2 (Atom at 0 plus an exponential)	$2\theta e^{-2t}$ $c^2 = (1-\theta)/\theta$ $\lambda_2^{-1} = m_1(1+c^2)/2$ $p_2 = 2/(1+c^2)$	$(1/2(1-\theta))e^{-t/(1-\theta)}$ $c^2 = 1+2\rho$ $\lambda_2^{-1} = m_1(1+c^2)/2$ $p_2 = 2/(1+c^2)$	$p_2 e^{-\lambda_2 t}$ $c^2 = 1+(\rho/2(1-\theta)^2)$ $\lambda_2^{-1} = (1-\theta)\eta$ $\eta = [1+(\rho/4(1-\theta)^2)]$ $p_2 = (1-\theta)\lambda_2$
H_2 fit based on $g''(0)$, $g'(0)$, $g(0)$	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_1^{-1} = 2\theta^2 p/(1-p)$ $\lambda_2^{-1} = 2\theta^2(1-p)/p$ $p = [1-\sqrt{\rho/(\rho+4)}]/2$	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_1^{-1} = \theta(1-\sqrt{\rho})$ $\lambda_2^{-1} = \theta(1+\sqrt{\rho})$ $p = (1-\sqrt{\rho})/2$	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_1^{-1} = \theta q/(1-q)$ $\lambda_2^{-1} = \theta(1-q)/q$ $p = q^3$ $q = [1-\sqrt{\rho(4-3\rho)}]/2$
H_2 fit based on $g'(0)$, $g(0)$, m_1	$0.5e^{-\lambda_1 t}+0.5e^{-\lambda_2 t}$ $\lambda_1^{-1} = \theta(1-\sqrt{\rho})$ $\lambda_2^{-1} = \theta(1+\sqrt{\rho})$	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_1^{-1} = \theta p/(1-p)$ $\lambda_2 = \theta(1-p)/p$ $p = [1-\sqrt{\rho/(4-3\rho)}]/2$	$\lambda_1^{-1}e^{-\lambda_1 t}+\lambda_2^{-1}e^{-\lambda_2 t}$ $\lambda_1^{-1} = (1-\sqrt{\rho})/2$ $\lambda_2^{-1} = (1+\sqrt{\rho})/2$
H_2 fit based on $g(0)$, m_1, m_2	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_1^{-1} = \theta(1-p)/p$ $\lambda_2^{-1} = \theta p/(1-p)$ $p = [1+\sqrt{\rho/(4-3\rho)}]/2$	$0.5e^{-\lambda_1 t}+0.5e^{-\lambda_2 t}$ $\lambda_1^{-1} = (1-\sqrt{\rho})/2$ $\lambda_2^{-1} = (1+\sqrt{\rho})/2$	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_2 = 4\lambda_1^{-1}$ $p = 2/(2+\lambda_1)$ see note above.
H_2 fit based on m_1, m_2, m_3	$\lambda_2^{-1}e^{-\lambda_1 t}+\lambda_1^{-1}e^{-\lambda_2 t}$ $\lambda_1^{-1} = (1-\sqrt{\rho})/2$ $\lambda_2^{-1} = (1+\sqrt{\rho})/2$	$pe^{-\lambda_1 t}+(1-p)e^{-\lambda_2 t}$ $\lambda_2 = 4/\lambda_1$ $p = \lambda_1/(2+\lambda_1)$ see note above.	$0.5e^{-\lambda_1 t}+0.5e^{-\lambda_2 t}$ $\lambda_1^{-1} = (1+\rho-\sqrt{\rho})/2$ $\lambda_2^{-1} = (1+\rho+\sqrt{\rho})/2$

TABLE 5

The hyperexponential (H_2) approximations for the basic three moment cdf's in the specific case $\rho = 0.75$.

	the three basic complementary cdf's		
Type of Approximation	$B^c(t)$	$H_1^c(t)$	$V_2^c(t) = c_q(t)$
One-Moment Exponential Fit	e^{-8t}	e^{-2t}	$e^{-1.143t}$
Two-Moment Bounding H_2 (Atom at 0 plus an exponential)	$0.25e^{-2t}$	$0.571e^{-1.143t}$	$0.8033e^{-0.918t}$
H_2 fit based on $g''(0)$, $g'(0)$, $g(0)$	$p = 0.3013$ $\lambda_1 = 74.2$ $\lambda_2 = 3.46$	$p = 0.0670$ $\lambda_1 = 59.71$ $\lambda_2 = 4.287$	$p = 0.00515$ $\lambda_1 = 38.32$ $\lambda_2 = 1.670$
H_2 fit based on $g'(0)$, $g(0)$, m_1	$p = 0.5000$ $\lambda_1 = 59.71$ $\lambda_2 = 4.287$	$p = 0.1727$ $\lambda_1 = 38.32$ $\lambda_2 = 1.670$	$p = 0.0670$ $\lambda_1 = 14.93$ $\lambda_2 = 1.072$
H_2 fit based on $g(0)$, m_1, m_2	$p = 0.8273$ $\lambda_1 = 38.32$ $\lambda_2 = 1.670$	$p = 0.500$ $\lambda_1 = 14.93$ $\lambda_2 = 1.072$	$p = 0.301$ $\lambda_1 = 4.64$ $\lambda_2 = 0.863$
H_2 fit based on m_1, m_2, m_3	$p = 0.933$ $\lambda_1 = 14.93$ $\lambda_2 = 1.072$	$p = 0.699$ $\lambda_1 = 4.64$ $\lambda_2 = 0.836$	$p = 0.500$ $\lambda_1 = 2.26$ $\lambda_2 = 0.765$
Parameters			
$g''(0)$	124,928.	14,336.0	512.0
$g'(0)$	1792.	256.0	16.00
$g(0)$	32.00	8.00	2.00
m_1	0.125	0.500	0.875
m_2	0.125	0.875	1.906
m_3	0.328	2.859	6.973

moments and yields an upper bound on the Laplace-Stieltjes transform. It is interesting that the stationary-excess cdf associated with this bounding cdf coincides with the one-moment exponential fit.

It should be clear that the three-parameter H_2 approximations improve near the origin (for larger t) if we use more derivatives (moments) in the fitting. As in [1]-[4], we have been primarily interested in larger t, so that we think of the three-moment fit as the standard one, but the others can be considered too. It is natural to ask how the approximations based on one region perform in the other region. As illustrations, note that the H_2 fit for $B^c(t)$ based on the three derivatives $b^{(2)}(0)$, $b^{(1)}(0)$ and $b(0)$ has mean $\theta(1-\rho^2)$ while the actual mean is θ. Thus the approximation for the mean is pretty good for small ρ, but not for large ρ. On the other hand, the H_2 fit for $B^c(t)$ based on the first three moments m_1, m_2 and m_3 has a density at the origin of $2(1+\rho)/(1-\rho)$ while the actual value is $b(0) = 2/(1-\rho)^2$. Again, the approximation is good for small ρ, but not for large ρ. Having a good fit for $b(0)$ is obviously a pretty strong requirement, but these tests reveal important limitations of the H_2 approximations. Even though $B(t)$ is completely monotone, it is quite different from a simple mixture of two exponentials. (Better approximations for $B(t)$ are developed in a forthcoming paper.) However, as $\rho \to 0$, $B(t)$, $H_1(t)$ and $V_2(t)$ approach a simple exponential (e.g., see Corollary 2 to Theorem 1), so that we should expect the H_2 approximations to be uniformly good for small ρ.

Remark 8.1. Riordan [25], p. 107, claims that a three-moment H_2 fit for $b(t)$ actually matches the first four moments. However, while this is nearly true, it is not actually correct because $m_4 = 3\theta[1+3\rho]$ in the three-moment H_2 fit, whereas the exact value is $m_4 = 3\theta[1+3\rho+\rho^2]$. The difference $3\theta\rho^2$ corresponds to a 20 percent relative error as $\rho \to 1$, but is negligible for small ρ. ∎

The resulting approximations for $c_q(t) = V_2^c(t)$, $c_l(t)$ and $c_w(t)$ based on a three-moment fit to each of the basic cdf's $B(t)$, $H_1(t)$ and

$V_2(t)$ are compared to exact values obtained by Laplace transform inversion (Section 4.4 of [1]) in the cases $\rho = 0.50$ and $\rho = 0.75$ in Tables 1-2 and 6-9. Note that these composite approximations involve four exponentials for $c_w(t)$ and six exponentials for $c_l(t)$ because the exponential parameters for $B(t)$, $H_1(t)$ and $V_2(t)$ based on three-moment fits are different.

For $c_l(t)$ and $c_w(t)$, we actually consider two different approximation procedures. In addition to the composite approximations based on three-moment H_2 fits to $B(t)$, $H_1(t)$ and $V_2(t)$, we also consider direct three-moment fits for the cases $\rho = 0.50$ and 0.75 in Tables 6-9. Since $c_l(t)$ is completely monotone (Theorem 10), such a three-moment H_2 fit is always possible. For $c_w(t)$, a direct three-moment H_2 fit is possible for $\rho \geqslant 0.2$; i.e., the first three moments of $c_w(t)$ are $m_1 = (3-\rho)/2(2-\rho)$, $m_2 = (4+2\rho-\rho^2)/2(2-\rho)$ and $m_3 = 3(5+10\rho-\rho^3)/4(2-\rho)$, so that $c^2 > 1$ for all ρ and $m_1 m_3/m_2^2 = 3(5+10\rho-\rho^3)(3-\rho)/2(4+2\rho-\rho^2) \geqslant 1.5$ for $\rho \geqslant 0.2$.

Finally, there is the first-moment cdf for the number waiting $U_1(t)$. (Since $V_1(t) = H_1(t)$ and $S_k(t) = B(t)$, we have already treated all other first-moment cdf's.) From Corollary 2 to Theorem 3, we know that $U_1^c(t)$ is not completely monotone. However, the first three moments of $U_1(t)$ are $m_1 = (2-\rho)/2$, $m_2 = (3-\rho)/2$ and $m_3 = 3(4+2\rho-\rho^2)/4$, so that

$$c^2 = (m_2-m_1^2)/m_1^2 = 1 + 2(3\rho-1-\rho^2)/(2-\rho)^2 \geqslant 1$$

and $m_3 m_1/m_2^2 \geqslant 1.5$ for $\rho \geqslant (1+\sqrt{5})/4 = 0.382$. Consequently, a direct three-moment H_2 fit is possible for $\rho \geqslant 0.382$. This direct three-moment H_2 fit is compared to the composite approximation involving four exponentials based on Corollary 2 to Theorem 3 and the exact values based on Laplace transform inversion for the cases $\rho = 0.50$ and 0.75 in Tables 10 and 11.

In conclusion, all the approximations perform quite well, with the exception of the H_2 approximations of the busy-period cdf $B(t)$ (which is

TABLE 6

A comparison of two approximations for the correlation function $c_I(t)$ with exact values obtained by Laplace transform inversion: the case of $\rho = 0.50$.

Note. For the direct H_2 fit, the three moments are $m_1 = 0.85$, $m_2 = 1.65$ and $m_3 = 5.213$. The H_2 parameters are $p_1 = 0.449$, $\lambda_1^{-1} = 0.495$, $p_2 = 0.551$ and $\lambda_2^{-1} = 1.139$.

time t	exact by transform inversion	composite H_2 fit six exponentials (Theorem 4)	direct three-moment H_2 fit two exponentials
0.01	0.9847	0.9848	0.9862
0.05	0.927	0.928	0.933
0.10	0.864	0.866	0.872
0.15	0.806	0.810	0.815
0.20	0.755	0.759	0.762
0.25	0.707	0.713	0.713
0.50	0.519	0.523	0.519
0.75	0.387	0.387	0.384
1.00	0.293	0.289	0.289
1.25	0.224	0.219	0.220
1.50	0.172	0.168	0.169
1.75	0.133	0.130	0.132
2.00	0.103	0.102	0.103
2.50	0.0636	0.0641	0.0643
3.00	0.0399	0.0408	0.0406
3.50	0.0252	0.0261	0.0259
4.00	0.0158	0.0167	0.0166
4.50	0.0103	0.0108	0.0107
5.00	0.0069	0.0069	0.0069
6.00	0.00228	0.00285	0.00284
7.00	0.00185	0.00117	0.00118

TABLE 7

A comparison of two approximations for the correlation function $c_l(t)$ with exact values obtained by Laplace transform inversion: the case of $\rho = 0.75$.

Note. For the direct H_2 fit, the three moments are $m_1 = 0.914$, $m_2 = 2.020$ and $m_3 = 7.43$. The H_2 parameters are $p_1 = 0.493$, $\lambda_1^{-1} = 0.492$, $p_2 = 0.507$ and $\lambda_2^{-1} = 1.325$.

time t	exact by transform inversion	composite H_2 fit six exponentials (Theorem 4)	direct three-moment H_2 fit two exponentials
0.01	0.9839	0.9857	0.9863
0.05	0.925	0.932	0.934
0.10	0.861	0.872	0.872
0.15	0.803	0.817	0.816
0.20	0.753	0.766	0.764
0.25	0.705	0.719	0.716
0.50	0.524	0.528	0.526
0.75	0.399	0.395	0.395
1.00	0.309	0.301	0.303
1.25	0.242	0.235	0.236
1.50	0.192	0.186	0.187
1.75	0.153	0.149	0.149
2.00	0.122	0.120	0.121
2.50	0.0798	0.0802	0.0799
3.00	0.0526	0.0541	0.0538
3.50	0.0353	0.0368	0.0366
4.00	0.0237	0.0251	0.0249
4.50	0.0162	0.0172	0.0171
5.00	0.0111	0.0117	0.0117
6.00	0.0049	0.0055	0.0055
7.00	0.00262	0.00256	0.00258

TABLE 8

A comparison of two approximations for the correlation function $c_w(t)$ with exact values obtained by Laplace transform inversion: the case of $\rho = 0.50$.

Note. For the direct H_2 fit, the three moments are $m_1 = 0.833$, $m_2 = 1.583$ and $m_3 = 4.938$. The H_2 parameters are $p_1 = 0.5095$, $\lambda_1^{-1} = 0.527$, $p_2 = 0.4905$ and $\lambda_2^{-1} = 1.151$.

time t	exact by transform inversion	composite H_2 fit four exponentials (Theorem 8)	direct three-moment H_2 fit two exponentials
0.01	0.9868	0.9872	0.9862
0.05	0.934	0.937	0.933
0.10	0.871	0.877	0.871
0.15	0.813	0.821	0.814
0.20	0.759	0.768	0.761
0.25	0.709	0.718	0.712
0.50	0.512	0.515	0.515
0.75	0.379	0.375	0.379
1.00	0.284	0.278	0.282
1.25	0.216	0.210	0.213
1.50	0.165	0.161	0.163
1.75	0.127	0.125	0.126
2.00	0.098	0.098	0.098
2.50	0.0605	0.0613	0.0603
3.00	0.0375	0.0388	0.0379
3.50	0.0236	0.0247	0.0241
4.00	0.0152	0.0158	0.0154
4.50	0.0095	0.0101	0.0099
5.00	0.00639	0.00643	0.00641
6.00	0.00215	0.00263	0.00268
7.00	0.00139	0.00107	0.00112

TABLE 9

A comparison of two approximations for the correlation function $c_w(t)$ with exact values obtained by Laplace transform inversion: the case of $\rho = 0.75$.

Note. For the direct H_2 fit, the three moments are $m_1 = 0.90$, $m_2 = 1.975$ and $m_3 = 7.25$. The H_2 parameters are $p_1 = 0.5014$, $\lambda_1^{-1} = 1.320$, $p_2 = 0.4985$ and $\lambda_2^{-1} = 0.477$.

time t	exact by transform inversion	composite H_2 fit four exponentials (Theorem 8)	direct three-moment H_2 fit two exponentials
0.01	0.9843	0.9863	0.9859
0.05	0.923	0.933	0.932
0.10	0.857	0.872	0.869
0.15	0.798	0.815	0.812
0.20	0.745	0.762	0.759
0.25	0.698	0.713	0.710
0.50	0.516	0.518	0.518
0.75	0.392	0.386	0.388
1.00	0.303	0.295	0.296
1.25	0.237	0.230	0.231
1.50	0.187	0.182	0.183
1.75	0.149	0.146	0.146
2.00	0.120	0.118	0.118
2.50	0.078	0.078	0.078
3.00	0.00517	0.0529	0.0526
3.50	0.0341	0.0359	0.0357
4.00	0.0232	0.0245	0.0243
4.50	0.0156	0.0167	0.0166
5.00	0.0107	0.0114	0.0114
6.00	0.0048	0.0053	0.0053
7.00	0.0025	0.0025	0.0025

TABLE 10

A comparison of two approximations for the first-moment cdf $U_1(t)$ with exact values obtained from Laplace transform inversion: the case of $\rho = 0.50$.

Note. For the direct H_2 fit, the three moments are $m_1 = 0.75$, $m_2 = 1.25$ and $m_3 = 3.563$. The H_2 parameters are $p_1 = 0.854$, $\lambda_1^{-1} = 0.646$, $p_2 = 0.146$ and $\lambda_2^{-1} = 1.354$.

time t	exact by transform inversion	composite H_2 fit four exponentials (Corollary 2 to Theorem 3)	direct three-moment H_2 fit two exponentials
0.01	0.9983	0.99913	0.9858
0.05	0.970	0.981	0.931
0.10	0.911	0.936	0.867
0.15	0.844	0.876	0.808
0.20	0.778	0.811	0.753
0.25	0.716	0.746	0.702
0.50	0.481	0.475	0.495
0.75	0.335	0.315	0.352
1.00	0.239	0.223	0.252
1.25	0.175	0.166	0.182
1.50	0.130	0.127	0.132
1.75	0.0977	0.0982	0.0972
2.00	0.0743	0.0766	0.0721
2.50	0.0436	0.0470	0.0409
3.00	0.0260	0.0288	0.0242
3.50	0.0161	0.0177	0.0148
4.00	0.00996	0.0109	0.00938
4.50	0.00624	0.00665	0.00608
5.00	0.00376	0.00407	0.00402
6.00	0.00179	0.00152	0.00182
7.00	0.00039	0.00057	0.00085

TABLE 11

A comparison of two approximations for the first-moment cdf $U_1(t)$ with exact values obtained from Laplace transform inversion: the case of $\rho = 0.75$.

Note. For the direct H_2 fit, the three moments are $m_1 = 0.625$, $m_2 = 1.125$ and $m_3 = 3.703$. The H_2 parameters are $p_1 = 0.6678$, $\lambda_1^{-1} = 0.333$, $p_2 = 0.3322$ and $\lambda_2^{-1} = 1.213$.

time t	exact by transform inversion	composite H_2 fit four exponentials (Corollary 2 to Theorem 3)	direct three-moment H_2 fit two exponentials
0.01	0.9903	0.9977	0.9775
0.05	0.892	0.955	0.893
0.10	0.780	0.864	0.800
0.15	0.692	0.765	0.719
0.20	0.621	0.673	0.648
0.25	0.562	0.592	0.585
0.50	0.369	0.339	0.368
0.75	0.260	0.229	0.249
1.00	0.190	0.171	0.179
1.25	0.142	0.134	0.134
1.50	0.109	0.107	0.104
1.75	0.0837	0.0857	0.0819
2.00	0.0655	0.0690	0.0655
2.50	0.0410	0.0450	0.0426
3.00	0.0261	0.0293	0.0281
3.50	0.0171	0.0191	0.0186
4.00	0.0113	0.0124	0.0123
4.50	0.00749	0.00811	0.00813
5.00	0.00517	0.00528	0.00538
6.00	0.00242	0.00224	0.00236
7.00	0.00110	0.00095	0.00103

studied in a forthcoming paper). From Tables 6-11, it is apparent that the direct H_2 approximations perform somewhat better than the composite approximations. Since the direct H_2 approximations have a simpler form (two exponentials as opposed to four or six), they usually are to be preferred.

9. Bounds and Inequalities

One positive density $f_1(t)$ on $[0,\infty)$ is said to be less than or equal to another $f_2(t)$ in the likelihood-ratio sense, here denoted by $f_1 \leqslant_r f_2$, if $f_2(t)/f_1(t)$ is increasing in t for all $t \geqslant 0$. Since the cdf's $B(t)$, $H_1(t)$ and $V_2(t)$ have completely monotone densities $b(t)$, $h_1(t)$ and $v_2(t)$ which are connected by the stationary-excess operator, we can apply Theorem 3.1(i) of [29] to order these densities. Moreover, from the spectral representation, pp. 98-103 of [13], we know that the mixing distribution in the mixture-of-exponential representation of $b(t)$ has support on the interval $[2(1+\sqrt{\rho})^{-2}, 2(1-\sqrt{\rho})^{-2}]$ in the time scale (3.1). We thus obtain the following comparison result. Let e_λ denote an exponential density with mean λ^{-1}.

Theorem 11. For $\lambda_1 = 2(1+\sqrt{\rho})^{-2}$ and $\lambda_2 = 2(1-\sqrt{\rho})^{-2}$, $e_{\lambda 1} \leqslant_r b \leqslant_r h_1 \leqslant_r v_2 \leqslant_r e_{\lambda 2}$.

Corollary 1. For all $t \geqslant 0$, $e^{-\lambda_1 t} \leqslant B^c(t) \leqslant H_1^c(t) \leqslant V_2^c(t) \leqslant e^{-\lambda_2 t}$ for λ_1 and λ_2 in Theorem 11.

Corollary 2. When $\rho = 0$, $B^c(t) = H_1^c(t) = V_2^c(t) = e^{-2t}$, $t \geqslant 0$.

Corollary 3. When $\rho = 0$, $c_q(t) = c_l(t) = e^{-2t}$, $U_1^c(t) = (e^{-2t})^{*2} = (1+2t)e^{-2t}$ and $c_w(t) = U_{1e}^c(t) = (1+t)e^{-2t}$.

Corollary 4. For RBM, $e_{1/2} \leqslant_r h_1 \leqslant_r c' = h_2 \leqslant_r h_3$.

As a consequence of Corollary 1 above, we can also easily establish some inequalities among the correlation functions.

Corollary 5. For all t and ρ, $c_q(t) \leqslant c_l(t)$ and $c_q(t) \leqslant c_w(t)$.

Proof. Apply Theorems 4 and 8 together with Corollary 1 to Theorem 11. ∎

From the numerical results we see that $c_w(t) \leqslant c_l(t)$ for sufficiently large t and ρ, but the inequality is reversed for sufficiently small t and ρ. For $\rho \geqslant 0.50$ and $t \geqslant 1$, the three correlation functions are quite close from a practical numerical view. Using only the dominant exponential in the direct H_2 fit, we obtain

$$c_q(t) \approx 0.500 e^{\frac{-t}{1.308}}, \quad c_w(t) \approx 0.501 e^{\frac{-t}{1.320}} \quad \text{and} \quad c_l(t) \approx 0.507 e^{\frac{-t}{1.325}}$$

for the case $\rho = 0.75$.

REFERENCES

[1] J. Abate and W. Whitt, Transient behavior of regulated Brownian motion, I: starting at the origin. *Adv. Appl. Prob.* 19 (1987), 560-598.

[2] J. Abate and W. Whitt, Transient behavior of regulated Brownian motion, II: non-zero initial conditions. *Adv. Appl. Prob.* 19 (1987), 599-631.

[3] J. Abate and W. Whitt, Transient behavior of the M/M/1 queue starting at the origin. *Queueing Systems* 2 (1987), 41-65.

[4] J. Abate and W. Whitt, Transient behavior of the M/M/1 queue via Laplace transforms. *Adv. Appl. Prob.* 20 (1988), to appear.

[5] V. E. Benes, On queues with Poisson arrivals. *Ann. Math. Statist.* 28 (1957), 670-677.

[6] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[7] J. W. Cohen, *The Single Server Queue*, Second Edition, North-Holland, Amsterdam, 1982.

[8] J. W. Cohen and G. Hooghiemstra, Brownian excursion, the M/M/1 queue and their occupation times. *Math. Oper. Res.* 6 (1981), 608-629.

[9] D. R. Cox and V. Isham, The virtual waiting-time and related processes. *Adv. Appl. Prob.* 18 (1986), 558-573.

[10] J. M. Harrison, *Brownian Motion and Stochastic Flow Systems*, Wiley, New York, 1985.

[11] D. L. Iglehart, Functional limit theorems for the queue GI/G/1 in light traffic. *Adv. Appl. Prob.* 3, 269-281.

[12] D. L. Iglehart and W. Whitt, Multiple channel queues in heavy traffic, II: sequences, networks, and batches. *Adv. Appl. Prob.* 2 (1970), 355-369.

[13] S. Karlin and J. McGregor, Many server queueing processes with Poisson input and exponential service times. *Pacific J. Math.* 8 (1958), 87-118.

[14] J. Keilson, A review of transient behavior in regular diffusion and birth-death processes. *J. Appl. Prob.* 1 (1964), 247-266.

[15] J. Keilson, A review of transient behavior in regular diffusion and birth-death processes, Part. II. *J. Appl. Prob.* 2 (1965), 405-428.

[16] J. Keilson, *Markov Chain Models - Rarity and Exponentiality*, Springer-Verlag, New York, 1979.

[17] I. Lee, *Stationary Markovian Queueing Systems: An Approximation for the Transient Expected Queue Length*, M. S. dissertation, Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, 1985.

[18] I. Lee and E. Roth, Stationary Markovian queueing systems: an approximation for the transient expected queue length, 1986, submitted for publication.

[19] P. M. Morse, Stochastic properties of waiting lines. *Operations Res.* 3 (1955), 255-261.

[20] A. R. Odoni and E. Roth, An empirical investigation of the transient behavior of stationary queueing systems. *Operations Res.* 31 (1983), 432-455.

[21] T. J. Ott, The covariance function of the virtual waiting-time process in an M/G/1 queue. *Adv. Appl. Prob.* 9 (1977a), 158-168.

[22] T. J. Ott, The stable M/G/1 queue in heavy traffic and its covariance function. *Adv. Appl. Prob.* 9 (1977b), 169-186.

[23] A. G. Pakes, The correlation coefficients of the queue lengths of some stationary single server queues. *J. Australian Math. Soc.* 12 (1971), 35-46.

[24] J. F. Reynolds, The covariance structure of queues and related processes — a survey of recent work. *Adv. Appl. Prob.* 7 (1975), 383-415.

[25] J. Riordan, *Stochastic Service Systems*, Wiley, New York, 1962.

[26] L. Takács, Investigation of waiting-time problems by reduction to Markov processes. *Acta. Math. Acad. Sci. Hungary* 6 (1955), 101-129.

[27] E. van Doorn, *Stochastic Monotonicity and Queueing Applications of Birth-Death Processes*, Lecture Notes in Statistics 4, Springer-Verlag, New York, 1980.

[28] W. Whitt, On approximations of queues, III: mixtures of exponential distributions. *AT&T Bell Lab. Tech. J.* 63 (1984), 163-175.

[29] W. Whitt, The renewal-process stationary-excess operator. *J. Appl. Prob.* 22 (1985), 156-167.

[30] W. Whitt, Planning queueing simulations. *Management Sci.*, to appear.

[31] C. M. Woodside, B. Pagurek and G. F. Newell, A diffusion approximation for correlation in queues. *J. Appl. Prob.* 17 (1980), 1033-1047.

APPENDIX A

Proof of Nonnegativity in Theorem 10

In this appendix we prove the nonnegativity of a_k in Theorem 10. First note that $(1+\rho-\rho^2)a_k = 2,\ 6\theta,\ 24\theta(1-\theta),\ 30\theta(4(1-\theta)^2+2\rho)$ for $k = 0, 1, 2$ and 3, respectively. For general k, we apply the basic recursion for the busy-period moments due to Riordan (Theorem 3.2 of [4]), namely,

$$4m_{k+2} = (4k+2)(1+\rho)m_{k+1} - (k^2-1)(1-\rho)^2 m_k \qquad \text{(A-1)}$$

for $m_0 = 1$ and $m_1 = (1-\rho)/2$. From (A-1), we get

$$\begin{aligned} b_k \equiv \rho(1+\rho-\rho^2)a_k &= [(4k+2)(1+\rho) - (4k+8)]m_{k+1} \\ &\quad + [(k^2+3k+2) - (k^2-1)(1-\rho)^2]m_k \\ &= [(4k+2)\rho-6]m_{k+1} + [(k^2-1)(2\rho-\rho^2) + 3(k+1)]m_k\ . \qquad \text{(A-2)} \end{aligned}$$

To prove that $b_k \geqslant 0$ for all k, we use a lemma about the busy-period moments, which is of some independent interest. Let $x_k = 2m_{k+1}/m_k(k+1)$. Note that $x_k = 1$ for all k when $\{m_k\}$ is the sequence of moments of an exponential with mean 1/2 (the case $\rho=0$).

Lemma A1. For $k = 0$, 1 and 2, $x_k = 1 + (k-1)\rho$. For $k \geqslant 3$, $1 \leqslant x_k \leqslant 1 + (k-1)\rho$.

Proof. Dividing through (A-1) by $2(k+2)m_{k+1}$ yields the following recursion for x_k:

$$x_{k+1} = \frac{(2k+1)(1+\rho)}{k+2} - \frac{(k-1)(1-\rho)^2}{x_k(k+2)} . \tag{A-3}$$

A bound for x_k in (A-3) yields a corresponding bound for x_{k+1}. If $x_k \geqslant 1$, then

$$x_{k+1} \geqslant \frac{(2k+1)(1+\rho)}{k+2} - \frac{(k-1)(1-2\rho+\rho^2)}{k+2}$$

$$= 1 + \frac{(4k-1)\rho-(k-1)\rho^2}{k+2} \geqslant 1 .$$

If $x_k \leqslant 1 + [k-1]\rho$, then

$$x_{k+1} \leqslant \frac{(2k+1)(1+\rho)}{k+2} - \frac{(k-1)(1-2\rho+\rho^2)}{(k+2)(1+[k-1]\rho)}$$

$$= 1 + k\rho - \frac{(k-1)k^2\rho^2}{(k+2)(1+[k-1]\rho)} \leqslant 1 + k\rho . \quad \blacksquare$$

Remark. It is also easy to see that $1 < x_k < 1 + (k-1)\rho$ for all $k \geqslant 3$ and $\rho > 0$. For example, $x_3 = 1 + 2\rho - \rho^2/(1+\rho)$. $\blacksquare$

From (A-2), we see that

$$b_k = 2m_{k+1}([2k+1]\rho-3) + (k+1)m_k(3+(k-1)(2-\rho)\rho)$$

so that $b_k \geqslant 0$ if and only if either (i) $\rho \geqslant 3/(2k+1)$ or (ii) $\rho < 3/(2k+1)$ and

$$x_k \equiv \frac{2m_{k+1}}{m_k(k+1)} \leqslant \frac{3+(k-1)(2-\rho)\rho}{3-[2k+1]\rho} .$$

However, if $\rho < 3/(2k+1)$, then

$$\frac{3+(k-1)(2-\rho)\rho}{3-[2k+1]\rho} = 1 + \frac{(4k-1)\rho}{3} + \frac{k(8k+1)\rho^2}{3-[2k+1]\rho} \geqslant 1 + (k-1)\rho \geqslant x_k$$

by Lemma A-1. Hence, the proof is complete. ■

APPENDIX B

Relations Among the Moment Functions Supporting Conjecture 1

Here we present relations among the moment functions which enable us to establish Conjecture 1 in Section 5 for $k = 3$, 4 and 5, and which might be the basis for a proof for all k.

We always consider the system starting out empty here. Let m_k and $m_{(k)}$ refer to the ordinary moment function $m_k(t, 0)$ and the associated factorial moment function (see [3] and [4]), respectively. These moment functions are related by $m_k = \sum_{j=1}^{k} a_{kj} m_{(j)}$ where a_{kj} are the Stirling numbers of the second kind, see Johnson and Kotz [12a], (105) on p. 19; e.g.,

$$\begin{aligned} m_1 &= m_{(1)}, \qquad m_2 = m_{(2)} + m_{(1)} \\ m_3 &= m_{(3)} + 3m_{(2)} + m_{(1)} \\ m_4 &= m_{(4)} + 6m_{(3)} + 7m_{(2)} + m_{(1)} \\ m_5 &= m_{(5)} + 10m_{(4)} + 25m_{(3)} + 15m_{(2)} + m_{(1)} . \end{aligned} \tag{B-1}$$

Let $\dot{m}_k$ and $\dot{m}_{(k)}$ denote the derivatives with respect to time; let $\delta(t)$ be the Dirac delta function. From Theorem 5.2 of [4], we know that

$$\dot{m}_{(k)} = k! \left(\frac{\rho}{2\theta} \right)^k [h_1(t)]^{*k} \tag{B-2}$$

where $h_1(t) = b(t) * f_{e0}(t)$ with $f_{e0}(t) = \rho h_1(t) + 2\theta\delta(t)$. Consequently, we see that the factorial moments satisfy the recursion relation

$$\dot{m}_{(k)} = \rho b(t) * [\dot{m}_{(k)} + k\dot{m}_{(k-1)}] \tag{B-3}$$

where $m_{(0)} = \delta(t)$.

We can thus apply Lemma 1 in Section 5 to relate $\dot{m}_{lk}$ to $\dot{m}_{(j)}$. In particular,

$$\begin{aligned}
\dot{m}_{l1} &= \dot{m}_{(1)} - \rho b(t) \\
\dot{m}_{l2} &= \dot{m}_{(2)} - \dot{m}_{(1)} + \rho b(t) \\
\dot{m}_{l3} &= \dot{m}_{(3)} + \dot{m}_{(1)} - \rho b(t) \\
\dot{m}_{l4} &= \dot{m}_{(4)} + 2\dot{m}_{(3)} + \dot{m}_{(2)} - \dot{m}_{(1)} + \rho b(t) \\
\dot{m}_{l5} &= \dot{m}_{(5)} + 5\dot{m}_{(4)} + 5\dot{m}_{(3)} + \dot{m}_{(1)} - \rho b(t) .
\end{aligned} \tag{B-4}$$

We now can establish Conjecture 1 by combining (B-3) and (B-4); e.g., for $k = 4$,

$$\begin{aligned}
\dot{m}_{l4} &= \rho b(t) * [\dot{m}_{(4)} + 4\dot{m}_{(3)}] + 2\rho b(t) * [\dot{m}_{(3)} + 3\dot{m}_{(2)}] \\
&\quad + \rho b(t) * [\dot{m}_{(2)} + 2\dot{m}_{(1)}] - \rho b(t) * [\dot{m}_{(1)} + \delta(t)] + \rho b(t) \\
&= \rho b(t) * [\dot{m}_{(4)} + 6\dot{m}_{(3)} + 7\dot{m}_{(2)} + \dot{m}_{(1)}] \qquad \text{(B-5)} \\
&= \rho b(t) * m_4 .
\end{aligned}$$

This establishes Conjecture 1 for the case $k = 4$, since clearly $m_{lk}(\infty) = \rho m_k(\infty)$. ■

Received: 1/19/1987

Revised: 2/15/1988

Accepted: 3/20/1988

Recommended by V. Ramaswami, Editor