CALCULATING TRANSIENT CHARACTERISTICS OF THE ERLANG LOSS MODEL BY NUMERICAL TRANSFORM INVERSION

J. ABATE
900 Hammond Road
Ridgewood, NJ 07450-2908

W. WHITT
AT&T Labs–Research
Room 2C-178
Murray Hill, NJ 07974-0636

Stochastic Models 3 (1998) 663–680

Key Words: Erlang loss model, transient behavior, blocking probability, first passage time, covariance function, numerical transform inversion, scaling, reflected Ornstein-Uhlenbeck diffusion process

ABSTRACT

We show how to compute the time-dependent blocking probability given an arbitrary initial state, the distribution of the time that all servers first become busy given an arbitrary initial state, the time-dependent mean number of busy servers given an arbitrary initial state, and the stationary covariance function for the number of busy servers over time in the Erlang loss model by numerically inverting the Laplace transforms of these quantities with respect to time. Algorithms for computing the transforms are available in the literature, but they do not seem to be widely known. We derive a new revealing expression for the transform of the covariance function. We show that the inversion algorithm is effective for large systems by doing examples with up to 10,000 servers. We also show that computations for very large systems (e.g., 10^6 servers) can be done with computations for moderately sized systems (e.g., 10^2-10^3 servers) and scaling associated with the heavy-traffic limit involving convergence of a normalized process to the reflected Ornstein-Uhlenbeck diffusion process. By the same reasoning, the Erlang model computations also can be used to calculate corresponding transient characteristics of the limiting reflected Ornstein-Uhlenbeck diffusion process.

1. Introduction

In this paper we consider the classical Erlang loss model, i.e., the M/M/c/0 system with Poisson arrival process, exponential service times, c servers and no extra waiting space, where blocked calls are lost. We let the individual service rate be 1 and the arrival rate (which coincides with the offered load) be a. We show how to compute several transient characteristics by numerical transform inversion. Transience arises by considering arbitrary fixed initial states.

In particular, we develop algorithms for computing: (1) the time-dependent blocking probability starting at an arbitrary initial state i, i.e., the transition probability

$$P_{ic}(t) \equiv P(N(t) = c | N(0) = i) \ , \tag{1.1}$$

where $N(t)$ is the number of busy servers at time t; (2) the complementary cumulative distribution function (ccdf) $F^c_{ic}(t)$ of the time T_{ic} all servers first become busy starting at an arbitrary initial state i; i.e.,

$$T_{ic} \equiv \inf\{t \geq 0 : N(t) = c | N(0) = i\} \tag{1.2}$$

and

$$F^c_{ic}(t) \equiv 1 - F_{ic}(t) \equiv P(T_{ic} > t) \ ; \tag{1.3}$$

(3) the time dependent mean

$$M_i(t) \equiv E(N(t) | N(0) = i) \ ; \tag{1.4}$$

and (4) the (stationary) covariance function

$$\begin{aligned} R(t) &\equiv Cov(N_s(u), N_s(u+t)) \qquad (1.5) \\ &= E(N_s(u)N_s(u+t)) - EN_s(u)EN_s(u+t) \ , \end{aligned}$$

where $\{N_s(t) : t \geq 0\}$ is a stationary version of $\{N(t) : t \geq 0\}$, i.e., where $N_s(u)$ in (1.5) is distributed according to the steady-state distribution

$$\pi_j \equiv P(N_s(u) = j) = \frac{a^j/j!}{\sum_{k=0}^{c} a^k/k!} \ . \tag{1.6}$$

We also show how to compute these quantities for very large systems by performing computations for moderately sized systems and using scaling based on the heavy-traffic limit in which $(N^{(a)}(t) - a)/\sqrt{a}$ converges to the reflected Ornstein-Uhlenbeck (ROU) process as $a \to \infty$ with $i(a) - a \sim \gamma_1\sqrt{a}$ and $c(a) - a \sim \gamma_2\sqrt{a}$, where $f(a) \sim g(a)$ means that $f(a)/g(a) \to 1$ as $a \to \infty$; see p. 177 of Borovkov [6] and Srikant and Whitt [26].

For example, suppose that we want to compute $P_{ic}^{(a)}(t)$ and $c(a) = \lfloor a + \sqrt{a} \rfloor$, $i(a) = \lfloor a - 2\sqrt{a} \rfloor$ and $a = 10^8$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. The scaling implies that $P_{ic}^{(a)}(t)/B(c,a)$ should be approximately independent of a, where $B(c,a) \equiv P_{ic}^{(a)}(\infty) \equiv \pi_c^{(a)}$ is the steady-state Erlang blocking probability, which is known to have the asymptotic relation

$$B(c,a) \sim \frac{1}{\sqrt{a}} \frac{\phi(\gamma)}{\Phi(-\gamma)} \quad \text{as} \quad a \to \infty \ , \tag{1.7}$$

where ϕ is the density and Φ is the cdf of a standard (mean 0, variance 1) normal distribution and γ is the limit of $(a - c)/\sqrt{a}$; see Jagerman [13], Whitt [29], and (15) of Srikant and Whitt [26]. Hence, we can compute using $a = 400$ and obtain

$$P_{ic}^{(10^8)}(t) \approx \frac{B(10^8 + 10^4, 10^8)}{B(400 + 20, 400)} P_{ic}^{(400)}(t) \approx \left(\frac{20}{10^4}\right) P_{ic}^{(400)}(t) \tag{1.8}$$

with i and c chosen appropriately in each case. We will show the effectiveness of the scaling in our numerical examples.

Our algorithms are based on computing the Laplace transforms of these quantities with respect to time. For the most part, algorithms for computing the transforms are available in the literature. In particular, an algorithm to calculate the Laplace transform of $P_{ij}(t)$ is given on pp. 81–84 of Riordan [24], but it does not seem to be widely known. Formulas for the Laplace transform of the mean and the covariance are given in Beneš [4], [5] and Jagerman [15], but the formula for the covariance transform in (15) on p. 209 of [5] and (15) on p. 136 of [4] has a sign error. We contribute to the basic theory by deriving a new revealing formula for the covariance transform. (See Theorem 2.)

Jagerman [15] evidently first had the idea of calculating these transient characteristics by numerical transform inversion, but the approach in [15] tends to produce a rough approximation rather than an accurate numerical algorithm. Also, Jagerman considered only the cases $c = 4, 8$ and 12 in his numerical examples. We wanted to know if an effective algorithm could be developed for large c. We show that it can by considering values of c up to 10^4.

Given the Laplace transforms, we apply the Fourier-series method of numerical transform inversion, as in Abate and Whitt [2], [3], but we could also apply other inversion algorithms, such as the Laguerre-series method in Abate, Choudhury and Whitt [1]. To accurately compute extremely small probabilities, we could use the transform scaling in Choudhury and Whitt [9] (which is unrelated to the scaling for large systems that we do use), but for typical cases that is not necessary, and we do not use it. Previous applications of numerical transform inversion to calculate transient characteristics of single-server queues are contained in Choudhury, Lucantoni and Whitt [7], [8] and Lucantoni, Choudhury and Whitt [21]. In contrast to the single-server algorithms, which use multi-dimensional inversion, we use only one-dimensional inversion here.

Our numerical inversion algorithm is an alternative to the spectral expansion described in Beneš [4], [5] and Riordan [24]. The spectral expansion is efficient for computing values at many time points, because the eigenvalues and eigenvectors need only be computed once. However, the inversion algorithm is also fast, and remarkably simple. The inversion program can be quite short, less than 100 lines. (See the displayed programs in [2], [3]).

Our numerical inversion algorithm is also an alternative to the numerical solution of a system of ordinary differential equations (ODEs), which has often been used; e.g., see Koopman [19], Taaffe and Ong [27], Ong and Taaffe [23], Mitra and Weiss [22], Green, Kolesar and Svoronos [12] and Davis, Massey and Whitt [11]. Numerical solution of ODEs has the advantage that it applies to time-dependent models as well as the transient behavior of stationary models with nonstationary initial conditions. However, when the numerical inversion algorithm applies, it has the advantage that it can produce calculations at any desired t without having to compute the function over a large set of time points in the interval $[0, t]$. The numerical inversion algorithm can also more easily produce high accuracy. (To illustrate with an extreme example, using the scaling in Choudhury and Whitt [9], we can compute probabilities of order 10^{-40} with relative error 10^{-8} using only standard double precision. This evidently is not possible with the system of ODEs.)

Even if another algorithm is used, such as the numerical solution of a system of ODEs, our scaling to reduce very large systems to approximately equivalent smaller systems can play an important role. With the spectral expansion, fewer roots need to be computed. With the ODEs, fewer ODEs need to be considered.

Finally, asymptotic formulas can serve as alternatives to exact numerical algorithms in the appropriate asymptotic regimes. Such asymptotic formulas are given in Mitra and Weiss [22], Knessl [18] and Chapter 12 of Shwartz and

Weiss [25]. These asymptotic formulas are very attractive when they are both simple and sufficiently accurate. If the asymptotic formulas are not simple, then they properly should be viewed as alternatives to numerical algorithms. For example, the direct numerical inversion here would seem to be preferable to the refined asymptote in Theorem 2 of Mitra and Weiss [22], which itself involves a Laplace transform that appears to be no easier to compute than $\hat{P}_{ic}(s)$ itself (because that transform is expressed as an integral). The asymptotic expressions in Knessl [18] also seem quite complicated.

2. Time-Dependent Blocking Probabilities

As shown on pp. 81–84 of Riordan [24], the Laplace transform

$$\hat{P}_{ij}(s) \equiv \int_0^\infty e^{-st} P_{ij}(t) dt \tag{2.1}$$

is easily computed recursively, exploiting relations among the Poisson-Charlier polynomials. (Earlier work used the related sigma functions; see Kosten, Manning and Garwood [20].) Since Riordan was not developing a numerical inversion algorithm, he was not interested in a numerical algorithm for computing the transform, so it is not highlighted, but it is there. The key relation is (8) on p. 84 of [24] using the recursions (3) and (4). the determinant $|D|$ in (8) is evaluated in (6).

Remark 2.1. The Laplace transform of $P_{ij}(t)$ in the more general GI/M/s/0 model is given in Chapter 4 of Takács [28], but that transform appears to be more difficult to compute.

We will focus on $P_{ij}(t)$ only for $j = c$, but the general case can be computed as well. To express the result for $P_{ic}(t)$, let

$$d_n \equiv d_n(s, a) = (-1)^n C_n(-s, a) \ , \tag{2.2}$$

where s is a complex variable and $C_n(s,a)$ are the Poisson-Charlier polynomials; i.e.,

$$d_n = \frac{1}{a^n}\sum_{k=0}^{n}\binom{n}{k}s(s+1)\ldots(s+k-1)a^{n-k} \; ; \tag{2.3}$$

e.g.,

$$\begin{aligned} d_0 &= 1 \; , \quad d_1 = \frac{1}{a}(a+s) & (2.4)\\ d_2 &= \frac{1}{a^2}(a^2+(2a+1)s+s^2) \; . & (2.5)\end{aligned}$$

We now specify the algorithm for computing $\hat{P}_{ic}(s)$ for any desired i, c and complex s. We use the polynomials d_n, but we do not compute them via (2.3); instead we compute them recursively. Our algorithm follows from the recursive relations in Riordan [24].

Theorem 1. *The Laplace transform of the time-dependent blocking probability is*

$$\hat{P}_{ic}(s) = d_i\hat{P}_{0c}(s) \; , \tag{2.6}$$

where

$$\hat{P}_{0c}(s) = \frac{1}{a(d_{c+1}-d_c)} \; , \tag{2.7}$$

d_0 *and* d_1 *are given in (2.4) and*

$$d_{n+1} = (1+\frac{n}{a}+\frac{s}{a})d_n - \frac{n}{a}d_{n-1} \; , \quad n \geq 1 \; . \tag{2.8}$$

Since $\{N_s(t) : t \geq 0\}$ is a stationary reversible process, e.g., see p. 26 of Keilson [16], $\pi_i P_{ic}(t) = \pi_c P_{ci}(t)$. Hence, we can also calculate $P_{ci}(t)$ directly from $P_{ic}(t)$ by

$$P_{ci}(t) = (\pi_i/\pi_c)P_{ic}(t) = \frac{a^i c!}{a^c i!}P_{ic}(t) \; . \tag{2.9}$$

As indicated in the introduction, $P_{ic}^{(a)}(t)/B(c,a)$ should be approximately independent of a provided that $i \equiv i(a) \approx a + \gamma_1\sqrt{a}$ and $c \equiv c(a) \approx a + \gamma_2\sqrt{a}$ for arbitrary constants γ_1 and γ_2 (which we think of as being in the interval

$[-5, 5]$). To calculate the Erlang blocking probability $B(c, a)$, we use the well known recurrence

$$B(c,a) = \frac{1}{1 + \frac{c}{aB(c-1,a)}} . \tag{2.10}$$

Remark 2.2. The Erlang blocking probability B is related to the polynomial d_n by $d_n(1, a) = 1/B(n, a)$. The recurrence relation (2.10) follows directly from another recurrence relation for d_n, namely,

$$d_n(s,a) = d_n(s+1,a) - \frac{n}{a} d_{n-1}(s+1,a) ; \tag{2.11}$$

see Corollary 3 on p. 549 of Jagerman [13]. The polynomials d_n are related to the sigma functions used in Beneš [5] and other early references by $\sigma_s(n) = a^n d_n(s, a)/n!$

We now illustrate the algorithm with a numerical example. We will consider five cases with five different values of a, ranging from $a = 100$ to $a = 10,000$, where $\gamma_1 = (i(a) - a)/\sqrt{a} = -3$ and $\gamma_2 = (c(a) - a)/\sqrt{a} = 2$. The five cases with steady-state performance measures are displayed in Table 1. Let M and V be the mean and variance of the steady-state number of busy servers, i.e., $M = a(1 - B)$ and

$$V = M - aB(c - M) = M - aB(c - a) - (aB)^2 . \tag{2.12}$$

The effectiveness of the scaling is shown there through the values of $\sqrt{a}B$ and V/a, which are nearly independent of a.

cases	c	a	i	B	M	V	$\sqrt{a}B$	V/a
I	120	100	70	.0056901	99.43	87.73	.056901	.877271
II	440	400	340	.0028060	398.88	352.72	.056120	.881806
III	960	900	810	.0018613	898.33	795.01	.055840	.883341
IV	2600	2500	2350	.0011122	2497.22	2211.45	.055608	.884579
V	10200	10000	9700	.0005543	9994.46	8855.13	.055430	.885513

Table 1. The five cases ($\gamma_1 = -3$ and $\gamma_2 = 2$).

Numerical values of $P_{ic}^{(a)}(t)/B(c,a)$ for nine time points are displayed in Table 2. The values of B are computed from (2.10), while the values of $P_{ic}^{(a)}(t)$ are computed by the Fourier-series method of numerical transform inversion, using Euler summation, as in [2], [3], after computing the transform values by the algorithm in Theorem 1. The inversion parameters were set so that the transform was computed at 40 values of complex s in each case. For the largest case, $a = 10^4$, the computation took about two minutes using UBASIC on a PC. As in Table 1, the effectiveness of the scaling in Table 2 is evident.

	I	II	III	IV	V
time	$(a = 100)$	$(a = 400)$	$(a = 900)$	$(a = 2,500)$	$(a = 10,000)$
1.0	.038920	.040993	.041755	.042435	.042836
1.5	.220241	.225617	.227479	.227581	.230147
2.0	.459358	.464459	.466181	.467744	.468612
2.5	.657298	.660662	.661786	.662651	.663363
3.0	.792636	.794518	.795143	.795656	.796044
4.0	.928489	.928951	.929102	.929222	.929311
5.0	.976022	.976108	.976135	.976156	.976171
7.0	.9973498	.9973442	.9973420	.9973401	.9973386
10.0	.99990311	.99990208	.99990172	.99990141	.9999011

Table 2. Values of $P_{ic}^{(a)}(t)/B(c,a)$ in the five cases.

3. First Passage Times

Let $f_{ij}(t)$ be the probability density function (pdf) of the first passage time T_{ij} from state i to state j in the M/M/c/0 model. Clearly,

$$P_{ij}(t) = f_{ij}(t) * P_{jj}(t) \tag{3.1}$$

for all i and j, where $*$ denotes convolution. Hence, if

$$\hat{f}_{ij}(s) \equiv \int_0^\infty e^{-st} f_{ij}(t)dt \ , \tag{3.2}$$

then

$$\hat{f}_{ij}(s) = \hat{P}_{ij}(s)/\hat{P}_{jj}(s) \ . \tag{3.3}$$

Since

$$\hat{F}^c_{ij}(s) = \frac{1 - \hat{f}_{ij}(s)}{s} \tag{3.4}$$

where

$$\hat{F}^c_{ij}(s) \equiv \int_0^\infty e^{-st} F^c_{ij}(t)dt \tag{3.5}$$

and $F^c_{ij}(t)$ is the ccdf of T_{ij}, we can calculate $F^c_{ij}(t)$ by numerical inversion too. In particular, given the algorithm for calculating $\hat{P}_{ic}(s)$ in Theorem 1, we can calculate $\hat{F}^c_{ic}(s)$ and $F^c_{ic}(t)$.

In fact, as is well known, it is easy to derive a recursion for the transform $\hat{f}_{i,i+1}(s)$ directly. Considering the first transition, we have

$$\hat{f}_{i,i+1}(s) = \left(\frac{a+i}{a+i+s}\right) \left(\frac{a}{a+i} + \left(\frac{i}{a+i}\right) \hat{f}_{i-1,i}(s)\hat{f}_{i,i+1}(s)\right) \tag{3.6}$$

or, equivalently,

$$\hat{f}_{i,i+1}(s) = \frac{a}{a+i+s-i\hat{f}_{i-1,i}(s)} \ . \tag{3.7}$$

On the other hand, we can derive (3.7) from Section 2 because

$$\hat{f}_{i,i+1}(s) = \frac{\hat{f}_{i,c}(s)}{\hat{f}_{i+1,c}(s)} = \frac{\hat{P}_{ic}(s)}{\hat{P}_{i+1,c}(s)} = \frac{d_i(s,a)}{d_{i+1}(s,a)} \tag{3.8}$$

and

$$\hat{f}_{0,i}(s) = 1/d_i(s,a) \ . \tag{3.9}$$

The recursion (3.7) also follows from (2.8) and (3.8).

By the scaling for large a, the distribution of T_{ic} should be approximately independent of a when $c(a) = \lfloor a + \gamma_1\sqrt{a}\rfloor$ and $i(a) = \lfloor a + \gamma_2\sqrt{a}\rfloor$. Indeed, as $a \to \infty$ with $c(a) - a \sim \gamma_1\sqrt{a}$ and $i(a) - a \sim \gamma_2\sqrt{a}$, T_{ic} converges in distribution to the first passage time τ_{γ_2,γ_1} of the Ornstein-Uhlenbeck (OU) diffusion process from γ_2 to γ_1; see Darling and Siegert [10] and Keilson and Ross [17].

We now give a numerical example. We compute the cdf $F_{ac}(t)$ for several values of t in the five cases given in Table 1. We let the initial state here be a instead of i; i.e., $\gamma_1 = 0$ instead of $\gamma_1 = -3$. The results are shown in Table 3.

time	$I(a = 100)$	$II(a = 400)$	$III(a = 900)$	$IV(a = 2,500)$	$V(a = 10,000)$
2	.1755	.1694	.1674	.1657	.1644
4	.3318	.3230	.3199	.3175	.3156
6	.4564	.4461	.4426	.4397	.4375
8	.5576	.5467	.5429	.5398	.5375
10	.6400	.6291	.6252	.6221	.6197
20	.8715	.8638	.8611	.8588	.8571
30	.9541	.9500	.9485	.9473	.9463
40	.9836	.9817	.9809	.9803	.9798
80	.9997	.9997	.9996	.9996	.9996

Table 3. Values of the first-passage-time cdf $F_{ac}(t)$ in the five cases ($\gamma_1 = 0$ and $\gamma_2 = 2$).

It is also possible to calculate the transforms and means of the first passage times $T_{i,i+1}$ recursively; e.g., by a variant of the argument in (3.6)

$$ET_{i,i+1} = \frac{1}{a} + \frac{i}{a} ET_{i-1,i} \ . \tag{3.10}$$

Then, for $i < j$,

$$ET_{i,j} = ET_{i,i+1} + \ldots + ET_{j-1,j} \ . \tag{3.11}$$

For the five cases in Table 3, the mean first passage times ET_{ac} are 9.81, 10.11, 10.21, 10.30 and 10.36.

Finally, we remark that the recursions (3.10) and (2.10) are closely related. They can be connected using the relation

$$ET_{i,i+1} = \frac{1}{aB(i,a)} \ . \tag{3.12}$$

Since $aB(i,a)$ is the overflow rate in an i-server system, $1/aB(i,a)$ is the mean time between overflows from the i-server system, which is easily seen to be $ET_{i,i+1}$.

4. The Time-Dependent Mean

The time-dependent mean in (1.4) has Laplace transform

$$\hat{M}_i(s) \equiv \int_0^\infty e^{-st} M_i(t)dt = \frac{i}{1+s} + \frac{a}{1+s}\left(\frac{1}{s} - \hat{P}_{ic}(s)\right) \ ; \qquad (4.1)$$

see p. 215 of Beneš [5]. Clearly $\hat{M}_i(s)$ is easily computed once we have $\hat{P}_{ic}(s)$.

Since $(N(t)-a)/\sqrt{a}$ converges to the ROU process as $a\to\infty$ with $i(a)-a \sim \gamma_1\sqrt{a}$ and $c(a) - a \sim \gamma_2\sqrt{a}$, we should have

$$m_i^{(a)}(t) \equiv \frac{M_i^{(a)}(t) - a}{\sqrt{a}} \to m_i(t) \quad \text{as} \quad a\to\infty \qquad (4.2)$$

provided that $i(a)$ and $c(a)$ are defined as above. We confirm the effectiveness of this scaling by computing the scaled mean $m_i^{(a)}(t)$ in (4.2) for several different values of a. In particular, values of $-m_i^{(a)}(t)$ are displayed in Table 4 for the same five cases as in Table 2. Now we let $\gamma_1 = -3$ again, as in Table 2.

time	*I* ($a = 100$)	*II* ($a = 400$)	*III* ($a = 900$)	*IV* ($a = 2,500$)	*V* ($a = 10,000$)
0.1	2.714512	2.714512	2.714512	2.714512	2.714512
0.5	1.819592	1.819592	1.819592	1.819592	1.819592
1.0	1.103903	1.103920	1.103925	1.103930	1.103638
1.5	.672385	.672445	.672466	.672483	.669390
2.0	.415669	.415718	.415733	.415743	.415751
3.0	.177146	.176943	.176865	.176800	.176748
5.0	.070190	.069547	.069316	.069124	.068976
7.0	.058365	.057607	.057335	.057111	.056938
10.0	.056954	.056174	.055895	.055664	.055486

Table 4. Values of the normalized mean $[a - M_i^{(a)}(t)]/\sqrt{a}$ in the five cases ($\gamma_1 = -3$).

5. The Covariance Function

We now give two new expressions for the Laplace transform of the covariance function in (1.5).

Theorem 2. *The covariance function $R(t)$ has Laplace transform*

$$
\begin{aligned}
\hat{R}(s) &\equiv \int_0^\infty e^{-st} R(t) dt \\
&= \frac{V}{1+s} - \frac{(M-V)}{(1+s)^2} + \frac{(aB)^2}{(1+s)^2}\left(\frac{\hat{P}_{cc}(s)}{B} - \frac{1}{s}\right) && (5.1) \\
&= \frac{V}{1+s} - \frac{(a-M)(\hat{M}_c(s) - (M/s))}{1+s}\ , && (5.2)
\end{aligned}
$$

where $B \equiv B(c,a) \equiv \pi_c$ in (1.6), $M \equiv M_i(\infty) = a(1-B)$ and $V \equiv R(0)$ is given in (2.12).

Proof. Formula (4.1) can be obtained by taking the Laplace transform of (15) on p. 209 of Beneš [5], after correcting a sign error ($C \to -C$ in the last term). Formula (4.1) can also be deduced from (29) on p. 217 of [5]. However, a more direct derivation is to combine equations (193) and (200) in Jagerman [15]. This yields

$$
\begin{aligned}
\hat{R}(s) = {} & \frac{V+M^2}{1+s} + \frac{aM}{s(1+s)} - \frac{M^2}{s} - \frac{acB}{(1+s)^2} \\
& - \frac{a^2B(1-B)}{s(1+s)^2} + \frac{(aB)^2}{(1+s)^2}\left[\frac{\hat{P}_{cc}(s)}{B} - \frac{1}{s}\right] .
\end{aligned} \tag{5.3}
$$

Comparing (5.3) to (4.1), we see that it suffices to show that

$$
-\frac{(M-V)}{(1+s)^2} = \frac{M^2}{1+s} + \frac{aM}{s(1+s)} - \frac{M^2}{s} - \frac{acB}{(1+s)^2} - \frac{aBM}{s(1+s)^2} . \tag{5.4}
$$

However, the first three terms on the right in (5.4) equal

$$
\frac{M(a-M)}{s(1+s)} = \frac{aBM}{s(1+s)}\ ,
$$

while the last term on the right in (5.4) equals

$$
\frac{aBM}{(1+s)}\left(\frac{1}{s} - \frac{1}{1+s}\right) = \frac{aBM}{s(1+s)} - \frac{aBM}{(1+s)^2} .
$$

Hence, the right side of (5.4) equals

$$
-\frac{(acB - aBM)}{(1+s)^2} = \frac{-aB(c-M)}{(1+s)^2} = \frac{-(M-V)}{(1+s)^2} .
$$

Finally, equation (4.2) follows from (4.1), exploiting the formula for $\hat{M}_c(s)$ in (4.1), obtained by setting $i = c$. $\square$

We can apply (4.2) to obtain useful expressions in the time domain.

Corollary. *The covariance can be expressed as*

$$R(t) = Ve^{-t} - (a - M)\int_0^t e^{-(t-u)}[M_c(u) - M)]du \leq Ve^{-t} . \tag{5.5}$$

Proof. The product of transforms in (4.2) corresponds to convolution. the inequality follows because $a > M$ and $M_c(t) \downarrow M$, the latter because the process $\{N(t) : t \geq 0\}$, being a birth-and-death process is stochastically monotone. $\square$

The Corollary to Theorem 2 yields a bound which is approached as $c \to \infty$; i.e., it is known that $R(t) = Ve^{-t}$ in the M/M/∞ model. Beneš proposes a simple approximation

$$R(t) \approx Ve^{-Mt/V} , \quad t \geq 0 ,$$

which is easy to compute and reasonably accurate; see p. 188 of [5].

Since

$$Cov\left(\frac{N_s(u) - a}{\sqrt{a}} , \frac{N_s(u+t) - a}{\sqrt{a}}\right) = \frac{Cov(N_s(u), N_s(u+t))}{a}$$

we conclude that $C^{(a)}(t)/a$ should be approximately independent of a provided that $c(a) = a + \gamma\sqrt{a}$. We confirm this scaling in our numerical example below. In particular, values of the normalized covariance function $R(t)/a$ are displayed in Table 5. We use the same five cases (values of a) and same nine time points as in Table 4. From the evident convergence, it is clear that the values can be used to approximate the covariance function of the limiting ROU diffusion process as well.

time	I $(a = 100)$	II $(a = 400)$	III $(a = 900)$	IV $(a = 2,500)$	V $(a = 10,000)$
0.1	.784019	.788345	.789814	.791000	.791895
0.5	.502346	.505750	.506913	.507853	.508564
1.0	.288786	.291173	.291990	.292652	.293153
1.5	.166203	.167816	.168370	.168819	.169159
2.0	.095700	.096765	.097132	.097429	.097655
3.0	.031748	.032192	.032345	.032469	.032564
5.0	.003496	.003219	.003589	.003608	.003623
7.0	.0003850	.0003948	.0003982	.0004010	.0004032
10.0	.00001407	.00001455	.00001472	.00001486	.00001496

Table 5. Values of the normalized covariance function $R(t)/a$.

References

[1] Abate, J., Choudhury, G. L. and Whitt, W. (1996) On the Laguerre method for numerically inverting Laplace transforms, *INFORMS J. Computing*, 8, 413–427.

[2] Abate, J. and Whitt, W. (1992) The Fourier-series method for inverting transforms of probability distributions, *Queueing Systems* 10, 5–88.

[3] Abate, J. and Whitt, W. (1995) Numerical inversion of Laplace transforms of probability distributions, *ORSA J. Computing* 7, 36–43.

[4] Beneš, V. E. (1961) The covariance function of a simple trunk group with applications to traffic measurements. *Bell System Tech. J.* 40, 117–148.

[5] Beneš, V. E. (1965) *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York.

[6] Borovkov, A. A. (1984) *Asymptotic Methods in Queueing Theory*, Wiley, New York.

[7] Choudhury, G. L., Lucantoni, D. M. and Whitt, W. (1994) Multi-dimensional transform inversion with application to the transient M/G/1 queue, *Ann. Appl. Prob.* 4, 719–740.

[8] Choudhury, G. L., Lucantoni, D. M., and Whitt, W. (1997) Numerical Solution of $M_t/G_t/1$ queues, *Oper. Res.*, 45, to appear.

[9] Choudhury, G. L. and Whitt, W. (1996) Probabilistic scaling for the numerical inversion of non-probability transforms, *INFORMS J. Computing*, submitted.

[10] Darling, D. A. and Siegert, J. F. (1953) The first passage problem for a continuous Markov process, *Ann. Math. Statist.* 24, 624–639.

[11] Davis, J. L., Massey, W. A. and Whitt, W. (1995) Sensitivity to the service-time distribution in the nonstationary Erlang loss model, *Management Sci.* 41, 1107–1116.

[12] Green, L., Kolesar, P. and Svoronos, A. (1991) Some effects of non-stationarity on multiserver Markovian queueing systems, *Oper. Res.* 39, 502–511.

[13] Jagerman, D. L. (1974) Some properties of the Erlang loss function. *Bell System Tech. J.* 53, 525–551.

[14] Jagerman, D. L. (1975) Nonstationary blocking in telephone traffic, *Bell System Tech. J.* 54, 625–661.

[15] Jagerman, D. L. (1978) An inversion technique for the Laplace transform with applications, *Bell System Tech. J.* 57, 669–710.

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

[17] Keilson, J. and Ross, H. F. (1975) Passage Time distributions for Gaussian Markov (Ornstein-Uhlenbeck) Statistical Processes, *Selected Tables in Mathematical Statistics*, 3, 233–327.

[18] Knessl, C. (1990) On the transient behavior of the M/M/m/m Loss model. *Stochastic Models* 6, 749–776.

[19] Koopman, B. O. (1972) Air terminal queues under time-dependent conditions, *Oper. Res.* 20, 1089–1114.

[20] Kosten, L., Manning, J. R. and Garwood, F. (1949) On the accuracy of measurements of probabilities of loss in telephone systems, *J. Roy. Statist. Soc., Ser. B* 11, 54–67.

[21] Lucantoni, D. M., Choudhury, G. L. and Whitt, W. (1994) The transient BMAP/G/1 queue, *Stochastic Models* 10, 145–182.

[22] Mitra, D. and Weiss, A. (1989) The transient behavior in Erlang's model for large trunk groups and various traffic conditions, *Teletraffic Science for new cost-Effective Systems, Networks and Services, ITC-12*, M. Bonatti (ed.), Elsevier-Science Publishers, Amsterdam 1367–1374.

[23] Ong, K. L. and Taaffe, M. R. (1989) Nonstationary queues with interrupted Poisson arrivals and unreliable/repairable servers, *Queueing Systems* 4, 27–46.

[24] Riordan, J. (1962) *Stochastic Service Systems*, Wiley, New York.

[25] Shwartz, A. and Weiss, A. (1995) *Large Deviations for Performance Analysis*, Chapman and Hall, London.

[26] Srikant, R. and Whitt, W. (1996) Simulation run lengths to estimate blocking probabilities *ACM J. TOMACS*, 6, 7–52.

[27] Taaffe, M. R. and Ong, K. L. (1987) Approximating Ph(t)/M(t)/S/C queueing systems, *Annals Oper. Res.* 8, 103–116.

[28] Takács, L. (1962) *Introduction to the Theory of Queues*, Oxford University Press, New York.

[29] Whitt, W. (1984) Heavy traffic approximations for service systems with blocking, *AT&T Bell Lab. Tech. J.* 63, 689–708.