Fast Estimation of the Statistics of Excessive Backlogs in Tandem Networks of Queues.

M.R. FRATER and B.D.O. ANDERSON

Department of Systems Engineering
Australian National University

The estimation of the statistics of buffer overflows in networks of queues is inherently costly, simply because of the rarity of these events. This paper presents two extensions to the known theory for finding optimal transformations for simulating such networks. The first is a direct proof of the asymptotic (as buffer size becomes large) optimality of exchanging the arrival and service rates in the simulation of an M/M/1 queue under the assumption that the arrival and service rates of the simulation system are state-independent. The second extension is a proof, using large deviations theory, that a similar technique can be used for tandem networks of queues. This transformation is also shown to be unique.

Keywords: Queueing networks, Markov chains, Importance sampling, Large deviations, Exit problem.

1 INTRODUCTION

The efficient estimation of the statistics of rare events has been of interest for a number of years [1,2]. More recently, specific application of this theory has been made to queueing networks ([3,4]), in which large deviations theory and numerical techniques have been used to find optimally fast simulation systems for networks of queues.

After a review of the use of importance sampling in simulating the M/M/1 queue (Section 2), the paper extends these ideas by providing a direct analytic proof of the optimality of exchanging the arrival and service rates for the simulation of an M/M/1 queue (Section 3), and extending the large deviations results to include an analytic minimization of variance for the simulation of a tandem network of queues (Section 4).

2 REVIEW OF M/M/1 QUEUES AND IMPORTANCE SAMPLING

Before stating the new results, a brief summary of the relevant theory relating to M/M/1 queues will be given. Consider an M/M/1 queue with arrival rate λ and service rate μ with $\mu/\lambda > 1$ (see Fig. 1). Suppose we are interested in estimating for some large integer N the value of the expectation $E[T]$, where T denotes the first time the queue length hits N, assuming it is initially zero. (The mean time between successive occasions with queue length N could also be of interest; generally, it will differ very little from $E[T]$.)

Because of the rarity of these events, simulation to determine the expected recurrence time is expensive. Importance sampling can be used to provide a more efficient simulation. The idea in importance sampling is as follows. Suppose we are interested in certain (rare) events occurring in a system S that we can simulate on a computer. Then instead of simulating S we simulate a second system $\overline{S}$, which has the property that events in S and $\overline{S}$ correspond in some way. In particular, to the rare events A in S correspond events $\overline{A}$ in $\overline{S}$. The correspondence is such that i) the events $\overline{A}$ in $\overline{S}$ are more frequent than the events A in S, and ii) the connection between S and $\overline{S}$ allows one to infer $P(A)$ if one knows $\overline{P}(\overline{A})$. ($\overline{P}(\overline{A})$ is the probability of event $\overline{A}$ in system $\overline{S}$.)

An analytic solution to the problem of finding the mean time between overflows for an M/M/1 queue is given below. Nevertheless, it also serves to illustrate the idea in importance sampling, which is also used for speeding up the simulation of networks. Take the M/M/1 queue shown in Fig. 1 as the system S. The second system $\overline{S}$ we consider is shown in Fig. 2. It is also an M/M/1 queue, but with different arrival and service rates. The fact that S and $\overline{S}$ have the same structure with different transition probabilities between states is typical of importance sampling. Notice that in fact the queue in Fig. 2 is unstable. In Fig. 1, A is the

Paper received 17 February 1989.
Final version 18 May 1989.

This work was supported by the Australian Telecommunications and Electronics Research Board (ATERB).

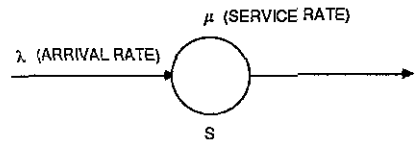


FIG. 1 — A Stable Queue ($\lambda<\mu$)

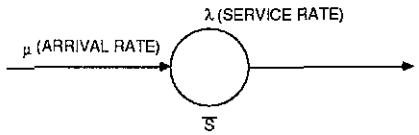


FIG. 2 — An Unstable Queue ($\lambda<\mu$)

event that the queue length, starting at 0, reaches N before returning to 0. In Fig. 2, $\overline{A}$ is the event that the queue length, starting at 0, reaches N before returning to 0. Obviously, the computation of the mean time for $\overline{A}$ to occur via simulation will be rapid. From it, one can deduce the mean time for A to occur as we now outline.

It is well known that one can set up an embedded discrete time Markov chain $\{X_k, k = 0, 1, 2, \ldots\}$, with X_k denoting the queue length just after the k-th change of that length, i.e. just after the k-th occurrence of an arrival or departure. For convenience, one can rescale time so that $\lambda+\mu = 1$. Then, λ and μ are respectively the probabilities of upward and downward transitions, given $X_k > 0$. Where $X_k = 0$, the probability of an upward transition is 1. At this stage, one could appeal to the methods of Neuts [5] to evaluate numerically the recurrence time of the event A, using linear algebraic techniques for distributions of *Phase Type*. These latter methods have the advantage of directly calculating the distribution function, and are potentially superior to direct simulation for finding exit times. However, we believe that the methods described here, although less direct then Neuts' approach, may have computational advantages in problems of large dimension, such as queueing networks.

Now let α be the probability that with $X_0 = 0$, X_k hits N before hitting zero again. Let T denote the first time X_k reaches N and let J denote the time to hit either 0 or N for the first time after leaving 0. It is not hard to check (see also [4]) that

$$E[T] = \frac{1}{\alpha}E[J], \tag{1}$$

and that α and $E[J]$ are given by:

$$\alpha = \frac{1-\frac{\mu}{\lambda}}{1-(\frac{\mu}{\lambda})^N} \tag{2}$$

$$E[J] = \frac{1-N\alpha}{\mu-\lambda} \tag{3}$$

Our interest, however, is to understand how these values can be obtained by (efficient) simulation. Now, $E[J]$ can be easily obtained by direct simulation on S, while α, which will be small, is obtainable from $\overline{S}$, as follows. Let us call a *cycle* of the system S or $\overline{S}$ a movement from 0 to the first time *either* 0 is reached again, *or* N is reached. Define $V_k = 1_{\{X_m \text{reaches } N \text{ in cycle } k\}}$. For S, we have

$$E[V_k] = \alpha \tag{4}$$

Let L_k denote the likelihood ratio $\frac{dP}{d\overline{P}}$ during cycle k. Notice that the L_k are i.i.d. and

$$\overline{E}[L_k V_k] = E[V_k] = \alpha \tag{5}$$

Now L_k is readily definable for S and $\overline{S}$. Suppose that $V_k = 1$, and that there are ℓ departures and therefore $N+\ell$ arrivals in the cycle. (There must be N more arrivals than departures for a cycle starting at 0 to end at N.) Let ς be a trajectory starting at 0, ending at N, with $N+\ell$ arrivals and ℓ departures in the cycle. Then

$$\begin{aligned}\overline{P}(\varsigma) &= \mu^{N+\ell-1}\lambda^{\ell} \\ P(\varsigma) &= \lambda^{N+\ell-1}\mu^{\ell}\end{aligned} \tag{6}$$

and so

$$L_k = \frac{dP}{d\overline{P}} = \left(\frac{\lambda}{\mu}\right)^{N-1} \tag{7}$$

on the set $\{V_k = 1\}$.

There are frequent occurrences of the set $\{V_k = 1\}$ for the system $\overline{S}$, since it is unstable. We examine p cycles for $\overline{S}$ and estimate α by

$$\begin{aligned}\hat{\alpha} &= \frac{L_1V_1 + L_2V_2 + \ldots + L_pV_p}{p} \\ &= \left(\frac{\lambda}{\mu}\right)^{N-1} \frac{\text{Number of cycles for which } V_k=1}{p}\end{aligned} \tag{8}$$

The speed up factor in simulation time obtained by using $\overline{S}$ instead of S and requiring equal accuracy in estimating α turns out to be, see [4]:

$$\left[N\left(\frac{\lambda}{\mu}\right)^N\left(1-\frac{\lambda}{\mu}\right)\right]^{-1} \tag{9}$$

In the above analysis we have shown how knowing and simulating $\overline{S}$, it can be used to compute α and thus important statistics concerning S. But we have not described what led us to the particular $\overline{S}$ chosen, rather than something else. *A major issue in the use of importance sampling is how one should construct $\overline{S}$ from S.* To an extent, the problems of obtaining the probability of a rare event or the mean time between occurrences of a rare event (which are both problems of excessive computer time) are being replaced by another difficult problem ('How should we obtain $\overline{S}$ from S?') in importance sampling.

This problem can be posed as an optimization problem, in the following way. Let A be a rare event for S: $\alpha = P(A) \ll 1$. For a direct Monte Carlo simulation involving n independent experiments we could estimate α via

$$\hat{\alpha}_n = \frac{1}{n}\sum_{i=1}^{n} 1_A(\omega_i) \tag{10}$$

where the ω_i are the i.i.d. outcomes of the experiments. The variance of $\hat{\alpha}_n$ is easily computed as

$$E[\alpha - \hat{\alpha}_n]^2 = \frac{1}{n}(\alpha - \alpha^2) \tag{11}$$

Alternatively, consider a probability measure $\overline{P}$ associated with a system $\overline{S}$, with P absolutely continuous with respect to $\overline{P}$. Using $\overline{S}$ we can obtain a second estimate

$$\hat{\overline{\alpha}}_n = \frac{1}{n}\sum_{i=1}^{n} 1_A(\overline{\omega}_i)L(\overline{\omega}_i) \tag{12}$$

where $L = \frac{dP}{d\overline{P}}$ and the $\overline{\omega}_i$ are the outcome of n experiments using $\overline{S}$. The variance of $\hat{\overline{\alpha}}$ is different to (11), and is obtainable as

$$\frac{1}{n}\left(\int_A L^2(\omega)d\overline{P}(\omega) - \alpha^2\right) \tag{13}$$

We want this to be as accurate as possible. So we want to adjust all the transition probabilities in S to new ones in $\overline{S}$ so that

$$(\sigma^*)^2 = \int_A L^2(\omega)d\overline{P}(\omega) \tag{14}$$

is minimized.

3 OPTIMAL SIMULATION OF M/M/1 QUEUE.

In general, direct minimization of the variance of the probability estimator, as defined in the previous section, is extremely difficult and less direct arguments using large deviations theory have been used [2,4]. However, for the example of the M/M/1 queue, given the assumption that the values of arrival and service rates in $\overline{S}$ are independent of state, we have found that a direct minimization of this quantity is possible in the sense that the optimality is analytically expressible as a function of a single variable that defines all possible $L(\omega)$ choices; further, the minimization is analytically achievable in the asymptotic case of large N.

For the M/M/1 queue, suppose that $\overline{S}$ is obtained from S by varying λ, μ to λ^* and μ^*. Then, except for the step 0 to 1, the likelihood ratio (L) for a single step takes on two values, one for an upward step ($\frac{\lambda}{\lambda^*}$), the other for a downward step ($\frac{\mu}{\mu^*}$). Since the step 0 to 1 occurs with probability 1 in S and $\overline{S}$, the likelihood ratio is 1 for this step. The value of the likelihood ratio for a trajectory is found by multiplying the likelihood ratios for the steps making up the trajectory. Since this system is discrete both in time and state-space, Equation (14) can be rewritten as follows:

$$(\sigma^*)^2 = \sum_A \left(\frac{\lambda^2}{\lambda^*}\right)^{N-n+\ell-1}\left(\frac{\mu^2}{\mu^*}\right)^{\ell} \tag{15}$$

where A is the set of all trajectories that reach N before reaching 0. Let

$$C_n = \sum_{A_n} \left(\frac{\lambda^2}{\lambda^*}\right)^{N+\ell-1}\left(\frac{\mu^2}{\mu^*}\right)^{\ell} \tag{16}$$

where A_n is the set of all trajectories that start at n and reach N before reaching 0. Then we can use the first step method to find C_n, and hence $(\sigma^*)^2$. The first step equation is:

$$C_n = \frac{\lambda^2}{\lambda^*}C_{n+1} + \frac{\mu^2}{\mu^*}C_{n-1} \tag{17}$$

with certain boundary conditions. These are obtained as follows. Consideration of the case $n = 1$ shows by a first-step type argument that

$$C_1 = \frac{\lambda^2}{\lambda^*}C_2, \tag{18}$$

while consideration of $n = N - 1$ shows that

$$C_{N-1} = \frac{\lambda^2}{\lambda^*} + \frac{\mu^2}{\mu^*}C_{N-2}. \tag{19}$$

It follows, in order that (17) hold for $n = 1, \ldots, N-1$, that $C_0 = 0$ and $C_N = 1$. Solving, we find:

$$C_n = \left(2\frac{\lambda^2}{\lambda^*}\right)^{N-n} \frac{(1+\sqrt{x})^n - (1-\sqrt{x})^n}{(1+\sqrt{x})^N - (1-\sqrt{x})^N} \qquad (20)$$

where x represents the quantity $1 - 4\frac{\lambda^2}{\lambda^*}\frac{\mu^2}{\mu^*}$ and where $\lambda^* + \mu^* = 1$. Because we are simulating the embedded Markov chain, all the trajectories have as their first step a jump from 0 to 1. Hence, the quantity we desire to minimize is

$$(\sigma^*)^2 = C_1 \qquad (21)$$

When C_1 is differentiated with respect to λ^*, it is found that its derivative becomes very small for large N when $\lambda^* = 1 - \lambda$. It is clear that as $\lambda^* \to 0$ and $\mu^* \to 0$ $(\lambda^* \to 1)$, $C_1 \to \infty$. Hence, it is clear that C_1 has a minimum with respect to λ^* very close to $\lambda^* = \mu$.

Unfortunately, it does not seem to be possible to apply the methods used here to more complicated examples, such as a network of M/M/1 queues, as in these examples, analytic solutions of the first-step equations cannot be found.

4 OPTIMAL SIMULATION OF QUEUES IN TANDEM

While the mean time between overflows can be calculated analytically for a single M/M/1 queue (e.g. [4]), the first step equations for a network of queues cannot be solved analytically because the order of the characteristic equation becomes large. Therefore, simulation is necessary. Just as it may not be practical to perform direct simulation of a single M/M/1 queue, other techniques may be required for the simulation of networks.

The analysis of Jackson networks presents significant difficulties. Large Deviations Theory has been used to obtain a number of approximate results [4]. Some of these, along with our extensions to these results, are presented below.

Consider the case of two queues in tandem (Fig. 3). If, for example $\mu_1 > \mu_2$, then if the size of the buffers (N) is large, the overflow statistics ought to be dominated by the second buffer's behaviour, and μ_1 should be of little importance, i.e. from the point of view of buffer overflows, the system should behave essentially as a single buffer with arrival rate λ and service rate μ_2. Hence, it is reasonable to suspect that for large N, the optimal simulation system will behave very like a single buffer with arrival rate μ_2 and service rate λ, even if the difference between μ_1 and μ_2 is small. Let us now see how this conclusion can be rigorously established, using methods that will apply

FIG. 3 — A Pair of Queues in Tandem

to general Jackson networks and possibly to general queueing networks.

Let $\xi_1 \ldots \xi_d$ be i.i.d. random variables in $\Re^d$. Let F be the distribution function of ξ_k and m its mean. Assume that the Laplace transform of F

$$M(s) = \int_{\Re^d} \exp < s, z > dF(z) \qquad (22)$$

is finite in a neighbourhood of 0. Then the Cramér or Legendre transform is defined as [2]

$$h(y) = \sup_{s \in \Re^d} [< s, y > - \log M(s)] \qquad (23)$$

For example, the Cramér transform of an exponential distribution with parameter λ is

$$h_\lambda(u) = \begin{cases} \lambda u - \log(\lambda u) - 1 & u > 0 \\ \infty & \text{otherwise} \end{cases} \qquad (24)$$

The following properties of the Cramér transform are used in this paper:

1. $h(\cdot)$ is convex;
2. $h(\cdot)$ is non-negative;
3. $h(y) = 0$, if and only if $y = m$, where m is the mean of the distribution function;
4. $h'(m) = 0$

Properties 3 and 4 are easily verified for $h_\lambda(u)$ in (24) on setting $u = \frac{1}{\lambda}$.

For the tandem queues, call a cycle a piece of a trajectory starting at the zero state and terminating on the first occasion when either the the sum of the contents of the two buffers equals N or again equals zero. If α is the probability that a cycle reaches N, then for a pair of buffers in tandem, it can be argued (see [4]) that the parameters λ' etc. for the system $\overline{S}$ can be found as the arguments achieving minimization in the following large-deviations approximation for α:

$$\alpha = \exp -N \min_{\substack{\lambda'>0, \mu_1', \mu_2' \geq 0, \\ \lambda' > \mu_1' \text{ or } \lambda' > \mu_2' \\ \lambda' + \mu_1' + \mu_2' = 1}} R\left[\lambda' h_\lambda\left(\frac{1}{\lambda'}\right) + \mu_1' h_{\mu_1}\left(\frac{1}{\mu_1'}\right) + \mu_2' h_{\mu_2}\left(\frac{1}{\mu_2'}\right)\right] \qquad (25)$$

where

$$R = \begin{cases} \frac{1}{\lambda' - \mu_1'} & \lambda' > \mu_1' \text{ and } \mu_1' < \mu_2' \\ \frac{1}{\lambda' - \mu_2'} & \text{otherwise} \end{cases} \quad (26)$$

This approximation for α does not give good numerical results, but should provide a good asymptotic approximation for the parameters of $\overline{S}$ for large N because the large deviations approximation considers only that component of α that behaves exponentially in the limit of large exit time. This exponential part should dominate over other multiplicative components of α in the determination of this minimum. The quantity R is the average time for the sum of the contents of the two buffers in the system $\overline{S}$ to increase by 1.

The quantity in the exponent of Equation 25 is to be minimized subject to the constraint $\lambda' + \mu_1' + \mu_2' = 1$. This minimization has previously been carried out numerically, and the results presented here demonstrated for certain specific cases. Here, an alternative analytic minimization that makes possible an extension to tandem networks of arbitrary length is shown. This can be done using the method of Lagrange multipliers:

$$\begin{aligned} \mathcal{L} &= R\left[\lambda' h_\lambda(\frac{1}{\lambda'}) + \mu_1' h_{\mu_1}(\frac{1}{\mu_1'}) + \mu_2' h_{\mu_2}(\frac{1}{\mu_2'})\right] \\ &\quad + g(\lambda' + \mu_1' + \mu_2' - 1) \end{aligned} \quad (27)$$

Because of the symmetry of the problem with respect to μ_1 and μ_2, we assume, without loss of generality, that $\lambda' > \mu_1'$ and $\mu_1' < \mu_2'$.

Then we have:

$$\frac{\partial \mathcal{L}}{\partial \lambda'} = -\frac{K}{(\lambda' - \mu_1')^2} + \frac{1}{\lambda' - \mu_1'} \log \frac{\lambda}{\lambda'} + g = 0 \quad (28a)$$

$$\frac{\partial \mathcal{L}}{\partial \mu_1'} = \frac{K}{(\lambda' - \mu_1')^2} + \frac{1}{\lambda' - \mu_1'} \log \frac{\mu_1}{\mu_1'} + g = 0 \quad (28b)$$

$$\frac{\partial \mathcal{L}}{\partial \mu_2'} = \frac{1}{\lambda' - \mu_1'} \log \frac{\mu_2}{\mu_2'} + g = 0 \quad (28c)$$

where

$$K = \lambda' h_\lambda(\frac{1}{\lambda'}) + \mu_1' h_{\mu_1}(\frac{1}{\mu_1'}) + \mu_2' h_{\mu_2}(\frac{1}{\mu_2'})$$

Since g is the rate of change of $\mathcal{L}$ with the sum of the arrival and service rates, and we do not expect the probability that a cycle exits rather than returns to zero to depend on the scaling of time, we require $g = 0$. Hence, it can be seen that there is a unique minimum of $\mathcal{L}$ where $\lambda' = \mu_1'$, $\mu_1' = \lambda$ and $\mu_2' = \mu_2$, i.e. λ is exchanged with the smaller of μ_1 and μ_2. (See Appendix for details.)

This idea can be generalized to networks of M queues in tandem. In this case, the Lagrangian is differentiated with respect to λ' and the μ_i', and we obtain a similar set of equations to that obtained for the pair of tandem queues above:

$$\frac{\partial \mathcal{L}}{\partial \lambda'} = -\frac{K}{(\lambda' - \mu_i')^2} + \frac{1}{\lambda' - \mu_i'} \log \frac{\lambda}{\lambda'} + g = 0 \quad (29a)$$

$$\frac{\partial \mathcal{L}}{\partial \mu_i'} = \frac{K}{(\lambda' - \mu_i')^2} + \frac{1}{\lambda' - \mu_i'} \log \frac{\mu_i}{\mu_i'} + g = 0 \quad (29b)$$

$$\frac{\partial \mathcal{L}}{\partial \mu_j'} = \frac{1}{\lambda' - \mu_i'} \log \frac{\mu_j}{\mu_j'} + g = 0 \quad (29c)$$

where

$$K = \lambda' h_\lambda(\frac{1}{\lambda'}) + \sum_{j=1}^{M} \mu_j' h_{\mu_j}(\frac{1}{\mu_j'})$$

(29c) is repeated for all the μ_j' except for that one which dominates the overflow statistics (i). This set of equations has a unique solution that corresponds to swapping λ with the smallest of the μ_i, and leaving all the other μ_i unchanged. The argument is virtually the same as that for the case of a pair of tandem queues.

By definition, the subscript μ of $h_\mu(\cdot)$ is the expected number of services per unit time. Hence, the mean time between services is $\frac{1}{\mu}$. Therefore, for buffers whose service rates are the same in both the original and optimal-simulation systems, (i.e. $\mu_j = \mu_j'$) we have

$$h_{\mu_j}(\frac{1}{\mu_j'}) = h_{\mu_j}(\frac{1}{\mu_j}) = 0 \quad (30)$$

since the Cramér transform has value zero at the mean of its associated distribution function.

It is guaranteed that $h(\cdot)$ is zero at exactly one point, that point being the mean of the associated distribution. At that point, its derivative is also zero [2]. Therefore, for buffers other than that which dominates the overflow statistics, (29c) can be written:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \mu_j'} \frac{1}{\lambda' - \mu_i'} h_{\mu_j}(\frac{1}{\mu_j'}) \\ &= \frac{1}{\lambda' - \mu_i'} \frac{\partial}{\partial \mu_j'} h_{\mu_j}(\frac{1}{\mu_j'}) \end{aligned} \quad (31)$$

where queue i is the one that dominates the overflow statistics, and $j \neq i$. Hence, it is clear that for queues other than that which dominates the overflow statistics, $\mu_j' = \mu_j$, *regardless of the distribution associated with these service rates.*

Hence, we see that for sufficiently large N, the optimal simulation system depends only on the statistics of the service rate of one queue (that of the least serviced buffer) and the arrival process, assuming that no two service rates are actually equal, and does not depend in any way on the statistics of the service rates of buffers other than the one dominating the overflow statistics.

5 CONCLUSION

This paper has presented an asymptotic analytic solution to the problem of finding the importance sampling transformation for the simulation of a tandem network of queues that is optimal in the sense of variance. This analysis needs to be extended in two ways. The first is to allow the optimal simulation of arbitrary queueing networks, and the second is to remove the restrictions on the statistics of the arrival and service rates imposed in the results presented here. Investigation of the relationship between the results presented here and the connection between large deviations and reverse-time modeling is also required in order to establish whether similar connections exist to those known for diffusions ([6]).

APPENDIX I

SOLUTION OF EQUATIONS (28)

In (28), as argued in the body of the text, $g = 0$. Then (28c) implies:

$$\mu_2' = \mu_2 \tag{32}$$

while (28a) and (28b) yield on addition:

$$\lambda'\mu_1' = \lambda\mu_1 \tag{33}$$

Now $\lambda' + \mu_1' + \mu_2' = 1 = \lambda + \mu_1 + \mu_2$, so

$$\lambda' + \mu_1' = \lambda + \mu_1 \tag{34}$$

Together, (33) and (34) yield

$$(\lambda' - \mu_1') = \pm(\lambda - \mu_1) \tag{35}$$

and then (34) and (35) imply

$$\lambda' = \lambda \text{ or } \lambda' = \mu_1 \tag{36}$$

whence either

$$\lambda' = \lambda, \quad \mu_1' = \mu_1, \quad \mu_2' = \mu_2 \tag{37a}$$

$$\lambda' = \mu_1, \quad \mu_1' = \lambda, \quad \mu_2' = \mu_2 \tag{37b}$$

The triples in (37a) and (37b) are the only two possibilities for a stationary point of $\mathcal{L}$. However, the triple (37a) fails to satisfy the constraint $\lambda' > \mu_1'$ or $\lambda' > \mu_2'$. Hence, the triple of (37b) is the unique stationary point. That it defines a minimum is easily established.

REFERENCES

[1] Zabczyk, J., "Exit Problem and Control Theory," Systems and Control Letters, Vol. 6, August 1985, pp. 165–172.

[2] Cottrell, M., Fort, J.C. and Malgouyres, G., "Large Deviations and Rare Events in the Study of Stochastic Algorithms," I.E.E.E. Trans. Automatic Control, Vol. AC-28, September 1983, pp. 907–918.

[3] Walrand, J., *An Introduction to Queueing Networks*. Englewood Cliffs, New Jersey: Prentice-Hall, 1988.

[4] Parekh, S. and Walrand, J., "A Quick Simulation of Excessive Backlogs in Networks of Queues," I.E.E.E. Trans. Automatic Control, Vol. 34, January 1989, pp. 54–66.

[5] Neuts, M.F., *Matrix-Geometric Solutions in Stochastic Models*. Baltimore, Maryland: Johns Hopkins University Press, 1981.

[6] Frater, M.R., Kennedy, R.A. and Anderson, B.D.O., "Reverse-time Modelling, Optimal Control and Large Deviations," Systems and Control Letters. To appear.

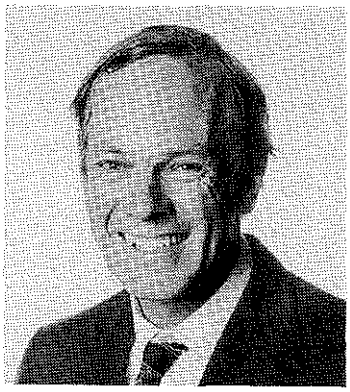
[7] Frater, M.R. and Anderson, B.D.O., "Estimation of Rare Event Statistics in Data Communications Networks," in 3rd Australian Teletraffic Research Seminar, November 1988.

[8] Mitrani, I., *Modelling of Computer and Communication Systems*, Cambridge University Press, 1987.

[9] Bertsekas, D.P. and Gallager, R.G., *Data Networks*, Englewood Cliffs, New Jersey: Prentice-Hall, 1987.

[10] Schwartz, M., *Telecommunication Networks: Protocols, Modelling and Analysis*, Reading, Massachusetts: Addison-Wesley, 1987.

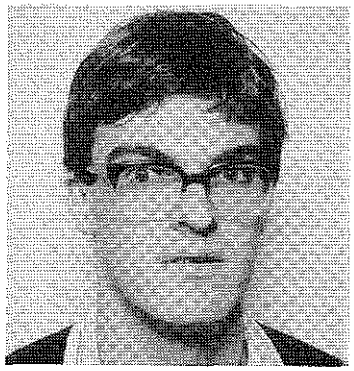
BIOGRAPHIES



Brian D.O. ANDERSON received the B.Sc. degree in mathematics and the B.E. degree in electrical engineering from the University of Sydney, and the Ph.D. degree from Stanford University in 1966.

He is currently Professor and Head of the Department of Systems Engineering at the Australian National University; from 1967 through 1981 he was Professor of Electrical Engineering at the University of Newcastle. He is co-author of several books and his research interests are in control and signal processing.

Dr Anderson is a Fellow of the Royal Society, the Australian Academy of Technological Sciences, the Australian Academy of Science and an Honorary Fellow of the Institution of Engineers, Australia. Previously, he was editor of *Automatica*, and is currently President-Elect of IFAC.



Michael R. FRATER received the B.Sc. degree in mathematics and physics (1986) and the B.E. degree in electrical engineering (1988) from the University of Sydney.

He is currently a graduate student in the Department of Systems Engineering at the Australian National University, working toward the Ph.D. degree. His research interests include aspects of communications, signal processing and control.