

Product Form Approximations for Highly Linear Loss Networks with Trunk Reservation *

Mark Bebbington

Institute of Information Sciences and Technology, Massey University, New Zealand

Phil Pollett

Department of Mathematics, The University of Queensland, Brisbane, Australia

Ilze Ziedins

Department of Statistics, The University of Auckland, New Zealand

Abstract

Admission controls, such as trunk reservation, are often used in loss networks to optimize their performance. Since the numerical evaluation of performance measures is complex, much attention has been given to finding approximation methods. The Erlang Fixed Point Approximation, which is based on an independent blocking assumption, has been used for networks both with and without controls. Certain methods of accounting for dependencies in blocking behaviour have been developed for the uncontrolled setting. We explore extensions to the controlled case, in order to gain insight into the essential elements of an effective approximation. We are able to isolate the dependency factor by restricting our attention to a highly linear network.

1 Introduction

Circuit-switched networks of the kind depicted in Figure 1 consist of a set of links indexed by $j = 1, 2, \dots, K$, with C_j circuits comprising each link j , and a collection of routes \mathcal{R} . Each route $r \in \mathcal{R}$ is a set of links. Calls using route r are offered at rate ν_r from independent Poisson streams, and use $a_{jr} (\geq 0)$ circuits from link j . Calls requesting route r are blocked and lost if, on *any* link j , there are fewer than a_{jr} available circuits. Otherwise, the call is connected and simultaneously holds a_{jr} circuits on each link j for the duration of the call. Call durations are iid exponential random variables with unit mean, independent of the arrival processes. For simplicity, we shall take $a_{jr} \in \{0, 1\}$.

Although it is possible to construct an explicit expression for the blocking probability, it can not (usually) be computed in polynomial time. Thus, for networks with even moderate capacity, one is forced to use alternative methods, arguably the most important of which is the *Erlang Fixed Point Approximation* (EFP); see Kelly [6].

*This work was funded by the Australian Research Council (Grant No. A49702317).

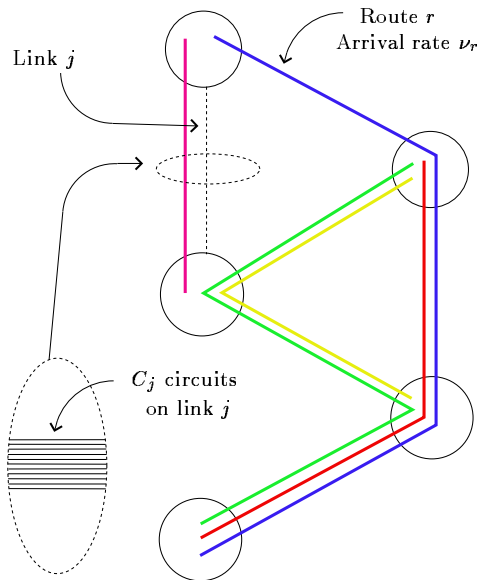


Figure 1: A typical circuit-switched network
(5 nodes, 6 links and 5 routes)

The rationale for the EFP is one of *independent blocking*. The EFP performs particularly well under two limiting regimes (Kelly [6], Ziedins and Kelly [12]), but if neither regime is operative, may not perform as well: in particular, in highly linear networks and/or networks with low capacities. Further, adding controls to the network may cause the method to perform badly under otherwise favourable regimes. A particularly useful control is *trunk reservation*, where traffic streams are assigned priorities and calls are accepted only if the occupancies of links along their route are below a given threshold, the level of which depends on the type of call. With such a control operating in a network of reasonable size, the occupancy of neighbouring links may be highly dependent and the equilibrium distribution will no longer have a product form, as it does for the corresponding uncontrolled network. Modelling dependencies in this context is thus critical.

Consider a loss network with K links forming a loop, and each link having the same capacity C . There are two types of traffic: one-link routes (type-1 traffic) and two-link routes comprising pairs of adjacent links (type-2 traffic). Type- t traffic is offered at rate ν_t on each type- t route. The result (a *symmetric ring network*) should be a simple, highly linear, network with correlations between adjacent links. This correlation reduces the effectiveness of the EFP.

If L_t is the EFP for the loss probability of type- t calls, then it is easy to show that $L_1 = B$ and $L_2 = 1 - (1 - B)^2$, where the Erlang Fixed Point B is the unique solution to $B = E(\nu_1 + 2\nu_2(1 - B), C)$.

$$E(\nu, C) = \frac{\nu^C}{C!} \left(\sum_{n=0}^C \frac{\nu^n}{n!} \right)^{-1}$$

is *Erlang's Formula* for the loss probability on a single link with C circuits and Poisson traffic offered at rate ν .

The blocking probabilities can be estimated more accurately by specifically accounting for the dependencies between adjacent links. Bebbington *et al.* [1] proposed the following improvements to the fixed point approximations for the ring network without controls, using state-dependent arrival rates in two-link subnetworks. Take links 1 and 2 as reference links and consider the subnetwork depicted in Figure 2. We identify three routes in the subnetwork: $\{1\}$, $\{2\}$ and $\{1,2\}$. If m_r denotes the number of calls on route r , then m_1 is the number of calls occupying capacity on link 1 *but not* on link 2, that is $m_1 = n_1 + n_{K1}$, m_2 is the number occupying capacity on link 2 *but not* on link 1, that is $m_2 = n_2 + n_{23}$, and, $m_{12}(= n_{12})$ is the number of calls occupying capacity on both links.

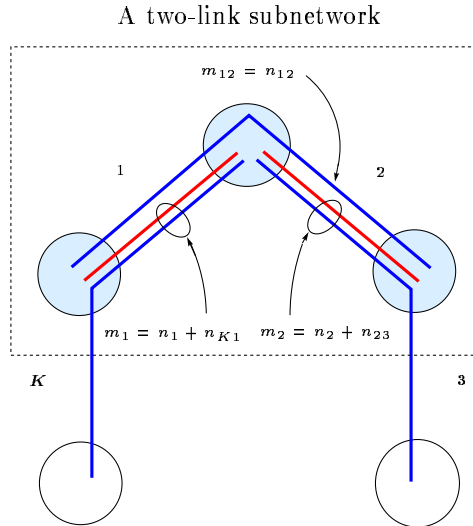


Figure 2: Definition of m_1 , m_2 and m_{12} for the symmetric ring network

The first approximation (*Approximation I*) is obtained by adapting the method of Pallant [9], where the network is decomposed into independent subnetworks and the stationary distribution is evaluated for each. For example, if we take our subnetwork to be the one depicted in Figure 2, then its state space will be $S = \{(m_1, m_2, m_{12}) : m_i + m_{12} \leq C, i = 1, 2\}$ with stationary distribution

$$\pi(\mathbf{m}) = \Phi^{-1} \frac{(\nu_1 + \nu_2(1 - B))^{m_1 + m_2} \nu_2^{m_{12}}}{m_1! m_2! m_{12!}},$$

where Φ is a normalizing constant. We then estimate B , the probability that a link adjacent to the two-link subnetwork is fully occupied, using the subnetwork itself;

$$B = \sum_{m : m_1 + m_{12} = C} \pi(m_1, m_2, m_{12}) = \sum_{m_{12}=0}^C \sum_{m_2=0}^{C-m_{12}} \pi(C - m_{12}, m_2, m_{12}). \quad (1)$$

These expressions iteratively determine a fixed point B , and we estimate $L_1 = B$ and $L_2 = 2L_1 - \sum_{m_{12}=0}^C \pi(C - m_{12}, C - m_{12}, m_{12})$.

A more accurate approximation (*Approximation II*) uses additional knowledge of the state of a given link in estimating the probability that the adjacent link is

full. We use *state-dependent* arrival rates, $\rho_n = \nu_1 + \nu_2(1 - b_n)$, $n \in \{0, 1, \dots, C-1\}$, where b_n is the probability that link K is fully occupied, conditional on $m_1 = n$ (b_n is also the probability that link 3 is fully occupied, conditional on $m_2 = n$), so that

$$\pi(\mathbf{m}) = \Phi^{-1} \frac{\nu_2^{m_{12}} \left(\prod_{n=0}^{m_1-1} \rho_n \right) \left(\prod_{n=0}^{m_2-1} \rho_n \right)}{m_1! m_2! m_{12}!}.$$

Once b_n is estimated and π determined, we estimate L_1 and L_2 as for Approximation I. An estimate of b_n is found by assuming that b_n does not depend on m_{12} . For $n = 0, \dots, C-1$, we set

$$b_n = \frac{\sum_{m=0}^n p(n-m, C-m, m)}{\sum_{m=0}^n \sum_{r=0}^{C-m} p(n-m, r, m)},$$

where

$$p(n_1, m_K, n_{K1}) = \frac{\nu_1^{n_1} \nu_2^{n_{K1}} \left(\prod_{s=0}^{m_K-1} \rho_s \right)}{n_1! n_{K1}! m_K!}.$$

The dependence of b_n on m_{12} is due to the cyclic nature of the network, but is expected to be slight for large networks.

While Approximation I gives some improvement in accuracy over the EFP, the improvement obtained using Approximation II is considerable, with relative error of the order of 10^{-5} of the EFP. See Bebbington *et al.* [1] for details. Zachary and Ziedins [11] derive an approximation for more general networks, which corresponds to Approximation II in the case of a ring network, although expressed differently.

2 Trunk Reservation

In a *trunk reservation* policy, a call of type- l is accepted on a link with capacity C only if there will be at least t_l circuits free on that link after the call is accepted. Such a policy is usually applied independently for each link involved in a multilink route. If each type ($l = 1, 2$) of call produces a reward w_l when accepted, then such a policy is optimal for a single link (Lippman [8]), although not necessarily for a multilink system.

In the trunk reservation case, provided all call types have the same capacity requirements and holding time distributions, a one-dimensional description of the occupancy of each link suffices, and the EFP can be extended in a straight-forward manner, although the equilibrium distribution of the network no longer has product form. If this is not the case, then a higher dimensional description of the occupancy of each link is required. Coyle *et al.* [5] discuss two ways of dealing with the latter case, with the object of achieving a product form approximation. This is then combined, in their situation, with a modularization technique developed for Stochastic Petri Net models by Ciardo and Trivedi [3, 4]. We observe that in the two link subnetwork of Figure 2, product form is lost immediately on the imposition of a control, such as trunk reservation.

The first method approximates the state of each link by a birth and death process,

giving an approximate distribution $\{p_n\}$ for the number of occupied circuits as

$$p_n \propto \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i}, \quad (2)$$

for $n = 0, 1, \dots, C$, where λ_i and μ_i are the (circuit) arrival and clear-down rates when i circuits are occupied. Obviously it will be necessary to approximate, as in the EFP, arrivals due to two-link calls by assuming that the neighbouring link blocks independently. The second method simply ignores the physical trunk reservation, instead *thinning* the unprotected stream by the amount which would be blocked due to trunk reservation. Thus we have the *reduced load* approximation on the two-dimensional state space, which then gives a product-form equilibrium distribution on that two-dimensional state space. (Note that Coyle *et al.* [5] recommend truncating the state space “appropriately”.) In our case, we are interested in a situation with differing traffic types, which take the same number of circuits per link, rather than a different number of circuits on the same number of links, so some adaptation is necessary. We will have the reduced arrival rates

$$\nu_l^* = \nu_l \left(\frac{\sum_{i=0}^{C-t_l-1} p_i}{\sum_{i=0}^{C-1} p_i} \right)^l, \quad (3)$$

$l = 1, 2$, where the exponent l is due to the fact that the two-link traffic can run afoul of trunk reservation on both links. This will be slightly modified for two-link traffic when we consider conditional approximations. The state probabilities $\{p_i\}$ are estimated by means of (2).

The approximate probabilities that a type-1 or type-2 call is blocked will be denoted by L_1, L_2 respectively. We will use $B^{(1)}$ and $B^{(2)}$ for the probabilities that a link has no circuit available for type-1 and type-2 calls, respectively. Let us first suppose that type-2 traffic is only accepted when the occupancy is below a threshold $C - t_2$, thus reserving the last t_2 circuits on each link for type-1 traffic. Then

$$\lambda_i = \begin{cases} \nu_1 + 2\nu_2(1 - B^{(2)}) & i < C - t_2 \\ \nu_1 & C - t_2 \leq i < C \\ 0 & i \geq C \end{cases}$$

and $\mu_i = i$. The first approximation, the standard extension of the EFP (see, for example, Kelly [7]) and hence denoted *Approximation Ea1*, is to let $\pi(n) = \Phi^{-1}(n!)^{-1}(\nu_1 + 2\nu_2(1 - B^{(2)}))^{\min(n, C-t_2)} \nu_1^{[n-C+t_2]_+}$, where $[\cdot]_+$ denotes the positive part, $B^{(1)} = \pi(C)$, and the birth and death formula (2) is used to approximate

$$B^{(2)} = f(B^{(1)}) = B^{(1)} \sum_{i=0}^{t_2} \frac{C!}{\nu_1^i (C-i)!}. \quad (4)$$

We then have $L_1 = B^{(1)}$, $L_2 = 1 - (1 - B^{(2)})^2$. In the case of the reduced load approach, we obtain *Approximation Eb1* with $\pi(n) = \Phi^{-1}(n!)^{-1}(\nu_1 + 2\nu_2 \left(\frac{1-B^*}{1-B}\right)^2 (1-B))^n$, where $B = \pi(C)$ and $B^* = f(B)$ is obtained from (4). The estimates for the one-link and two-link traffic blocking probabilities are

$$L_1 = B, \quad L_2 = 1 - (1 - B^*)^2. \quad (5)$$

Analogous approximations *Ea2* and *Eb2* can be constructed (see Bebbington *et al.* [2] for details) when the last t_1 circuits on each link are kept for type-2 traffic.

Now let us consider how we might attempt to improve on these approximations using the ideas of Section 1. Consider a two link subnetwork, with state description (m_1, m_2, m_{12}) as in Figure 2, and with trunk reservation parameter $t_2 > 0$ on each link. Then the arrival rate for type-2 calls occupying both links is $\nu_2 I_{\{m_1+m_{12} < C-t_2\}} I_{\{m_2+m_{12} < C-t_2\}}$, and we approximate the arrival rate on link i , $i = 1, 2$, from other sources by $\nu_1 + \nu_2(1 - B^{(2)}) I_{\{m_i+m_{12} < C-t_2\}}$, provided the link is not full. Let \tilde{B} be the probability that a link is at the trunk reservation threshold, that is, there are exactly $C - t_2$ circuits in use. Then

$$\tilde{B} = \sum_{m_{12}=0}^{C-t_2} \sum_{m_2=0}^{C-m_{12}} \pi(C - t_2 - m_{12}, m_2, m_{12}). \quad (6)$$

Using the birth and death approximation (2), we can then obtain

$$B^{(1)} = \tilde{B} \frac{\nu_1^{t_2} (C - t_2)!}{C!}, \quad (7)$$

and

$$B^{(2)} = \tilde{B} \sum_{i=0}^{t_2} \frac{\nu_1^i (C - t_2)!}{(C - t_2 + i)!}. \quad (8)$$

We can now produce *Approximation Ia1*, a product form approximation for the equilibrium distribution (closely related to Approximation I of Section 1) with

$$\pi(m_1, m_2, m_{12}) = \Phi^{-1} \frac{(\nu_1 + \nu_2(1 - B^{(2)}))^{\sum_{l=1,2} \min(m_l, C-t_2)} \nu_1^{\sum_{l=1,2} [m_l - C + t_2]_+} \nu_2^{m_{12}}}{m_1! m_2! m_{12!}},$$

for $m_{12} \leq C - t_2$, where $B^{(1)}$ and $B^{(2)}$ are given by (6), (7) and (8). The estimated blocking probabilities are then

$$L_1 = B^{(1)}, \quad L_2 = 1 - \sum_{m_{12}=0}^{C-t_2-1} \sum_{m_1=0}^{C-t_2-1-m_{12}} \sum_{m_2=0}^{C-t_2-1-m_{12}} \pi(m_1, m_2, m_{12}). \quad (9)$$

Alternatively, using the reduced load approach we obtain *Approximation Ib1*, with

$$\pi(m_1, m_2, m_{12}) = \Phi^{-1} \frac{\left(\nu_1 + \nu_2 \left(\frac{1-B^*}{1-B} \right)^2 (1-B) \right)^{m_1+m_2} \left(\nu_2 \left(\frac{1-B^*}{1-B} \right)^2 \right)^{m_{12}}}{m_1! m_2! m_{12!}},$$

where

$$B = \sum_{m_{12}=0}^C \sum_{m_2=0}^{C-m_{12}} \pi(C - m_{12}, m_2, m_{12}) \quad (10)$$

is the blocking probability, and $B^* = f(B)$ is obtained from (4). The estimates for the type-1 and type-2 traffic blocking probabilities are again given by (5). Again,

see Bebbington *et al.* [2] for the details of similar approximations *Ia2* and *Ib2* for the case when t_1 circuits on each link are reserved for type-2 traffic.

While Approximation II of Section 1 does not lend itself to exploitation via the birth-death or reduced-load approaches, we can construct a weakened version by replacing the estimate for B in (1) with an estimate of the probability that a link is full, conditional on the adjacent link having at least one circuit free:

$$\hat{B} = \frac{\sum_{m_{12}=0}^{C-1} \sum_{m_2=0}^{C-1-m_{12}} \pi(C - m_{12}, m_2, m_{12})}{\sum_{m_{12}=0}^{C-1} \sum_{m_1=0}^{C-m_{12}} \sum_{m_2=0}^{C-1-m_{12}} \pi(m_1, m_2, m_{12})}. \quad (11)$$

As a matter of interest, in the uncontrolled system, this is more accurate than Approximation I when the arrival rates are high (relative to the capacities). In the birth and death approximation, the equations are almost identical to the Approximation Ia cases, but in order to obtain *Approximation IIa1* (type-1 traffic protected), we substitute for (8) with

$$B^{(2)} = \frac{\sum_{m_{12}=0}^{C-t_2-1} \sum_{m_2=0}^{C-t_2-1-m_{12}} \pi(C - t_2 - m_{12}, m_2, m_{12})}{\sum_{m_{12}=0}^{C-t_2-1} \sum_{m_1=0}^{C-m_{12}} \sum_{m_2=0}^{C-t_2-1-m_{12}} \pi(m_1, m_2, m_{12})} \sum_{i=0}^{t_2} \frac{\nu_1^i (C - t_2)!}{(C - t_2 + i)!},$$

where the term before the summation is the probability that a link is at the trunk reservation threshold conditional on the adjacent link being below the threshold. The blocking estimates are again given by (9). *Approximation IIa2* (type-2 traffic protected) can be obtained similarly. In the reduced load approach we must replace the independent probabilities in (3) by the two link conditional probability that a two-link call is rejected on at least one of the links. If type-1 traffic is protected we thus obtain *Approximation IIb1*, with

$$\pi(m_1, m_2, m_{12}) = \Phi^{-1} \frac{\left(\nu_1 + \nu_2 \frac{(1-B^*)(1-\hat{B}^*)}{(1-B)(1-\hat{B})} (1 - \hat{B}) \right)^{m_1+m_2} \left(\nu_2 \frac{(1-B^*)(1-\hat{B}^*)}{(1-B)(1-\hat{B})} \right)^{m_{12}}}{m_1! m_2! m_{12!}},$$

where B is found from (10), $B^* = f(B)$ from (4), \hat{B} from (11), and $\hat{B}^* = f(\hat{B})$ from (4). The estimates for the one-link and two-link traffic blocking probabilities are $L_1 = B$, $L_2 = 1 - (1 - B^*)(1 - \hat{B}^*)$. If we protect type-2 traffic, we obtain *Approximation IIb2* similarly.

Coyle *et al.* [5] use approximations Ea, Eb and Ib for a four link ring network, resulting in relative errors of 4%–10% and upwards against simulated values.

3 Numerical Results

In order to investigate the accuracy (or otherwise) of the approximations, they were computed for the system with varying C , ν_1 , ν_2 , t_1 , and t_2 . The resulting values were compared with the simulated (Singer [10]) proportion of calls blocked in a system with $K = 24$, after a presample period long enough to discount transient effects. 95% confidence intervals for the simulated blocking probabilities were calculated and converted to confidence intervals for the relative errors. In some cases, notably Eb2, IIb2 and all the b1 approximations, the algorithms failed to converge in a straightforward manner, with blocking probabilities either exceeding unity, or displaying

two-point oscillation. This was remedied by applying a bisection procedure, with the next iterate being the average of the previous two. Iterates exceeding unity were reset to 1 before bisecting. The plots are presented in pairs, with the relative error in L_1 (L_2) on the left (right). We observe that approximations Eb, Ib and IIb give very similar results (denoted “b approx” in the figures), while approximations Ia and IIa are likewise similar (and labelled “a approx”).

If the load on the network is increased with ν_1 and ν_2 in constant proportion, Figure 3 shows that the “a” approximations seem to perform no worse than Ea when approximating L_1 for a range of moderate rates, but are markedly worse at higher rates. The “b” approximations for L_1 are much worse than Ea. Both “a” and “b” approximations are markedly worse than Ea for estimating L_2 , although the “b” estimates for L_2 are comparable with Ea at very high arrival rates for $t_1 = 3$.

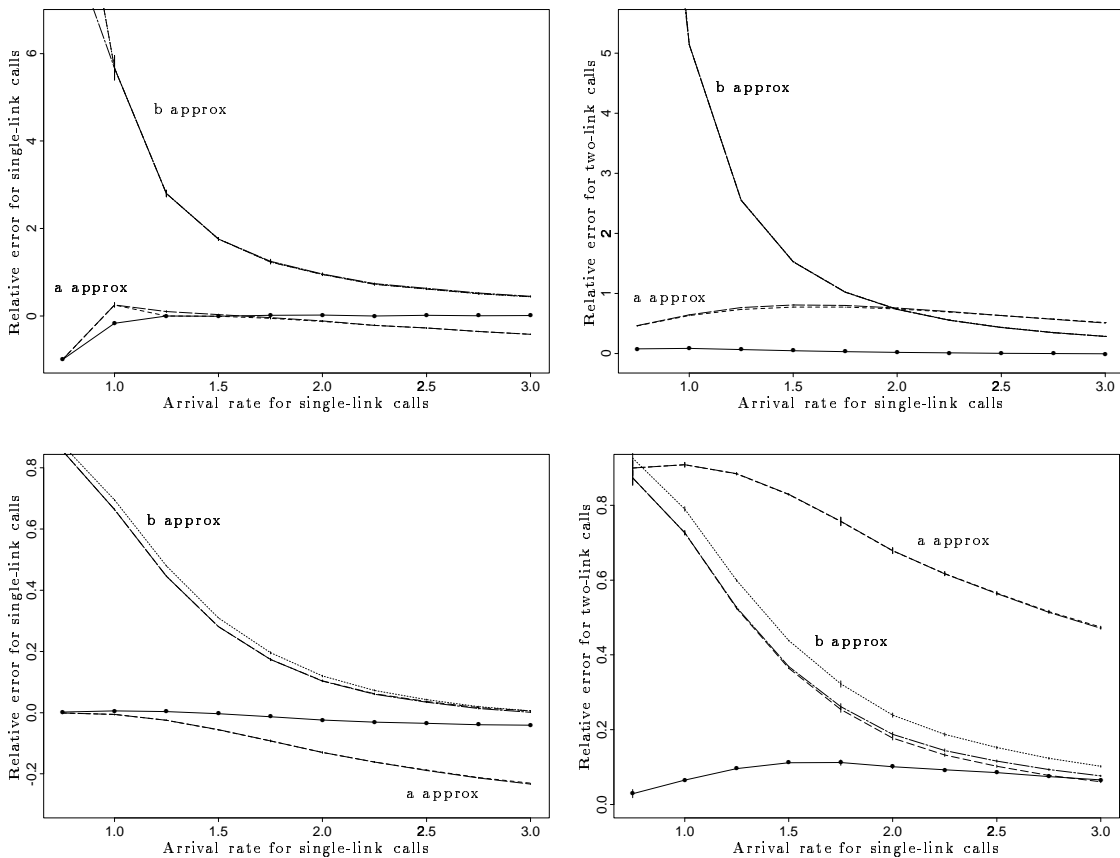


Figure 3: Variation with load. $C = 12, \nu_2 = 2\nu_1$ with ν_1 varying. The upper (lower) plots are for $t_2 = 3$ ($t_1 = 3$). Ea ————— Ia - - - - - IIa - - - - -
Eb Ib - - - - - IIb - . - . - .

As (t_1, t_2) varies, Figure 4 shows that the “a” and “b” approximations do better than Ea when $t_1 = t_2 = 0$, as shown in Bebbington *et al.* [1]. Otherwise, both are much worse than Ea over much of their range. The exceptions are the “a” approximations for L_1 , which are comparable with Ea when $t_2 = 1$, and variants Ib and IIb, as better estimates of L_2 than Ea when $t_1 = 6$.

From Figures 3 and 4 it is obvious that the greatest systematic deficiencies are when two-link traffic is protected, and in predicting the blocking probability

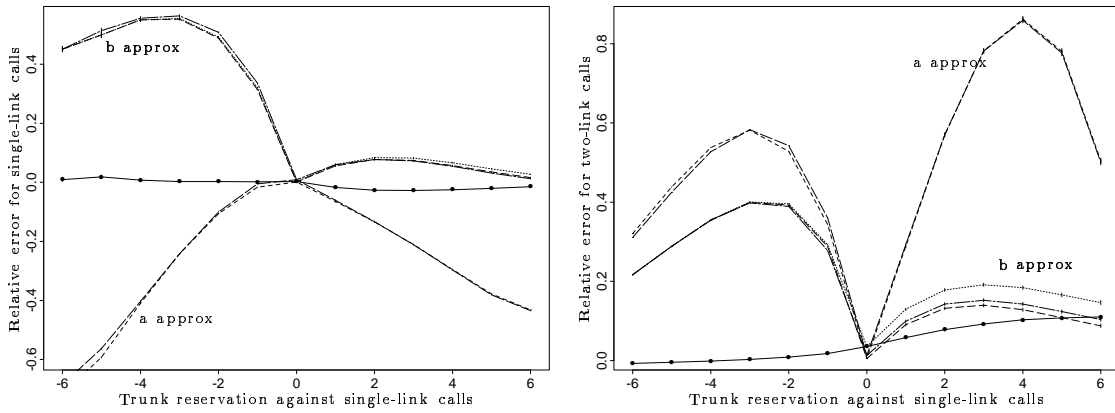


Figure 4: Variation with trunk reservation parameter. $C = 12, \nu_1 = 3, \nu_2 = 4.5$. The horizontal axis is the trunk reservation *against* type-1 traffic ($= t_1 - t_2$). Ea _____ Ia - - - - -
 - - - - - IIa - - - - - Eb Ib - - - - - IIb - . - . - .

for two-link traffic. The latter is understandable, due to the additional level of approximation in calculating L_2 from L_1 . This leads us to consider a constant loading $\nu_1 + 2\nu_2$ with $t_1 = 3$, and varying the traffic mix by means of ν_1 . We see in Figure 5 that the “b” approximations for single-link calls are markedly better than Ea for very low ν_1 , while being comparable for two-link calls, but rapidly worsen.

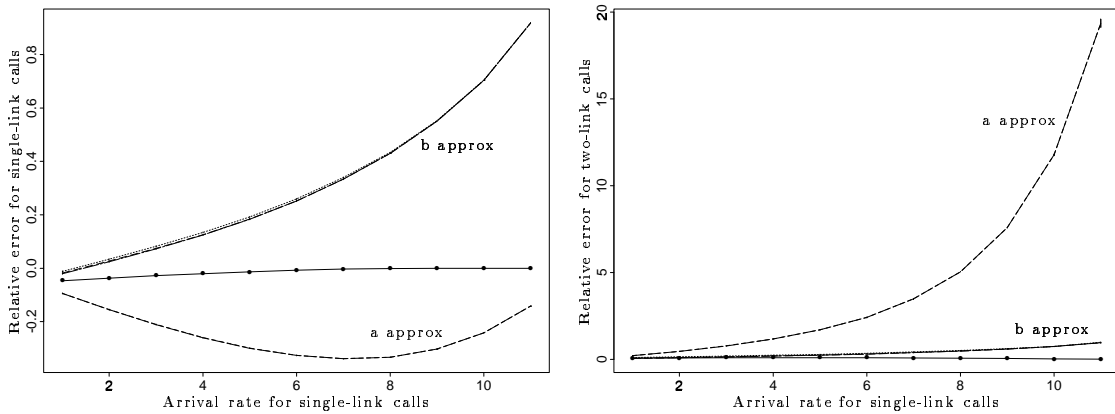


Figure 5: Variation with traffic mix (ν_1). $C = 12, t_1 = 3, 2\nu_1 + \nu_2 = 12$. Ea _____
 Ia - - - - - IIa - - - - - Eb Ib - - - - - IIb - . - . - .

We also considered increasing C with load proportional, and found that both the “a” and “b” approximations became appreciably worse, with the exception of the “a” approximation of L_1 , which however remained much poorer than Ea.

In general, the “a” approximations tend to underestimate L_1 while overestimating L_2 . The “b” approximations tend to overestimate both loss probabilities. While the “b” approximations seem to improve with higher loading or trunk reservation, the “a” approximations improve little if at all. It seemed that none of the proposed approximations performed as well in general as Ea, apart from a *very* limited range of values. Interestingly the “a” approximations are at their best when estimating L_1 , whereas the “b” approximations do better for estimating L_2 .

We have examined variations on the EFP for specialized networks with trunk reservation, which incorporate information about dependencies between neighbouring links. While our results are not as encouraging as we might have wished, they do point to the possibility of a variant of the full Approximation II, hopefully without the additional level of approximation superimposed by the birth and death product form expressions for the subnetwork. We plan to address this, and the formulation of an approximation for general networks, in future work.

References

- [1] Bebbington, M.S., Pollett, P.K., and Ziedins, I. (1998a) Two-link approximation schemes for linear loss networks without controls, *J. Korean Math. Soc.*, **35**, 539–557.
- [2] Bebbington, M.S., Pollett, P.K., and Ziedins, I. (1998b) Product form approximations for highly linear loss networks with trunk reservation, *CADSMAP Report Series*, 98-02, Mathematics Department, University of Queensland.
- [3] Ciardo, G., and Trivedi, K.S. (1990) Solution of large GSPN Models, *Proc. First Int. Conf. on the Numerical Solution of Markov Chains*. Raleigh, NC.
- [4] Ciardo, G., and Trivedi, K.S. (1991) A decomposition approach for stochastic petri net models, *Proc. of Petri Nets and Performance Models*, Melbourne, 74-85.
- [5] Coyle, A.J., Henderson, W., and Taylor P.G. (1995) *Decomposition Methods for Loss Networks with Circuit Reservation*, Teletraffic Research Centre, University of Adelaide.
- [6] Kelly, F.P. (1986) Blocking probabilities in large circuit-switched networks, *Adv. Appl. Probab.*, **18**, 473-505.
- [7] Kelly, F.P. (1990) Routing and capacity allocation in networks with trunk reservation, *Math. Oper. Res.*, **15**, 771-793.
- [8] Lippman, S.A. (1975) Applying a new device in the optimization of exponential queueing systems, *Oper. Res.*, **23**, 687-710.
- [9] Pallant, D. (1992) A reduced load approximation for cellular mobile networks including handovers. *Austral. Telecom. Res.*, **26**, 21-30.
- [10] Singer, D. (1993) *A guide to RINGNET*, Mathematics Department, University of Queensland.
- [11] Zachary, S. and Ziedins, I. (1999) Loss networks and Markov random fields. *J. Appl. Prob.*, to appear.
- [12] Ziedins, I.B. and Kelly, F.P. (1989) Limit theorems for loss networks with diverse routing. *Adv. Appl. Prob.*, **21**, 804-830.