Bounds on the Accuracy of the Reduced-Load Blocking Formula in Some Simple Circuit-Switched Networks

Bruce Hajek Arvind Krishna
Coordinated Science Laboratory and the
Department of Electrical and Computer Engineering
University Of Illinois at Urbana-Champaign
Urbana, IL 61801

Abstract

Circuit switched communication networks are considered in which each route is two links long, and each link can carry one call at a time. No symmetry assumption is made. Simple bounds, depending on the maximum call arrival rate and the maximum sum of rates at a link, are given on the blocking probabilities. An implication is that if the maximum per-route arrival rate converges to zero with a fixed bound on the sum of rates at links, then the well known reduced-load blocking approximation is asymptotically exact, uniformly over all network topologies.

1 Introduction

Circuit switched communication networks have long been used, most noticeably for telephone service. Recent development of high-speed wide-area networks has intensified interest in circuit switching techniques due to several factors, including the need for real-time delivery of video signals and the drive to simplify node routing functions to enable speedy implementation. In addition, techniques such as virtual circuit switching, asynchronous time-division multiplexing (ATM), virtual cut-through routing, and wormhole routing all, in some regimes, behave much like basic circuit switching. Several factors will complicate the analysis of blocking probabilities for systems likely to be developed. These factors include the fact that in the near future, fewer circuits per trunk are likely to be used than are used in the current telephone system, and a wide range of transmission rates will probably need to be accommodated.

[0]This research was supported by the National Science Foundation under contract NSF ECS 83 52030.

The set of known useful bounds on blocking probability in circuit switched networks is not yet complete. Existing approximations tend to be inaccurate when the degrees of switches are small. In many other cases, where existing approximations are typically accurate, there is still a lack of simple tight bounds. This paper provides a new technique for providing bounds on blocking probabilities. The model considered in this paper is extremely simple, but it is hoped that the technique can be extended to more general situations. The technique appears best suited for situations in which traffic routes are "diverse." The intuitive idea is that if a large number of routes goes through each link, but if the amount of traffic that goes through any two links is small, then the states of different links should be nearly independent in equilibrium. See Ziedens and Kelly [11] for more discussion of diverse routing.

Our paper is organized as follows. In this section, the network model will be introduced. It is the classical one with product form equilibrium distribution. The exact expressions for the blocking probabilities are reviewed, but no computationally feasible method is known for computing the exact blocking probabilities for large networks. In Section 2, approximations and bounds, including the reduced-load blocking approximation [10, 2], are discussed. Beginning in Section 3, attention is restricted to networks with routes covering two links each and with links that can carry one call each. A bound is provided which shows that if the maximum per-route offered traffic rate is small then the probability that two links are free is close to the product of the probabilities that each is free. Moreover, it is shown the the *actual* blocking probabilities are given by the reduced-load approximation for a certain modified vector of offered traffic rates. This leads to the consideration in Section 4 of the sensitivity of the blocking probabilities predicted by the reduced-load approximation with respect to small changes in the offered traffic rates. Combining the sensitivity analysis of Section 4 with the bounds of Section 3 yields the simple bounds on actual blocking probabilities, given in Corollary 2.

The mathematical model is as follows. A network is denoted by $(\mathcal{L}, \mathcal{R}, \mathbf{C}, \nu)$, where $\mathcal{L}$ is the set of links, $\mathcal{R}$ is the set of routes, where each route is a subset of $\mathcal{L}$, $\mathbf{C} = (C_l : l \in \mathcal{L})$ specifies how many circuits are in each link, and $\nu = (\nu_r : r \in \mathcal{R})$ specifies the rate that call requests are made for each route. Suppose that call requests arrive for route r according to a Poisson process with rate ν_r. A call requesting route r is accepted if each link in r has an available circuit. If accepted, a call holds one circuit in each link of its requested route for an amount of time that is exponentially distributed with mean one. If not accepted, a call is lost. The various arrival streams and call holding times are mutually independent. Let $\mathbf{n} = (n_r : r \in \mathcal{R})$ denote the number of calls in progress for each route. It is well known (for example see Kelly [3]) that the equilibrium distribution for the network is given by

$$\pi(\mathbf{n}) = G(\mathbf{C}) \prod_{r \in \mathcal{R}} \frac{\nu_r^{n_r}}{n_r!} \qquad \mathbf{n} \in \mathcal{S}(\mathbf{C}) \tag{1}$$

where $\mathcal{S}(\mathbf{C})$ is the set of feasible states,

$$G(\mathbf{C}) = \left(\sum_{\mathbf{n} \in \mathcal{S}(\mathbf{C})} \prod_{r \in \mathcal{R}} \frac{\nu_r^{n_r}}{n_r!} \right)^{-1}, \quad \mathcal{S}(\mathbf{C}) = \{\mathbf{n} : \sum_{r: l \in r \in \mathcal{R}} n_r \leq C_l \text{ for } l \in \mathcal{L}\}. \tag{2}$$

The equilibrium probability that there are exactly C_l calls in progress using link l is called the blocking probability at link l and it will be denoted by $\tilde{B}_l$. The notation B_l (without a tilde) will be reserved for the reduced load approximation to the actual blocking probability $\tilde{B}_l$. The blocking probability $\tilde{B}_l$ can be expressed using the function $G(\mathbf{C})$ defined in (2) as follows. Let I_A denote the indicator function of an event A, so $I_A = 1$ if A is true and $I_A = 0$ otherwise. Let $\mathbf{e}^l = (e_k^l : k \in \mathcal{L})$ be the vector with l^{th} coordinate given by $e_k^l = I_{\{k=l\}}$. Then, $\tilde{B}_l = 1 - G(\mathbf{C})/G(\mathbf{C} - \mathbf{e}^l)$. where subtraction (or addition) of vectors is performed componentwise. Also, the probability that a typical call for route r is accepted is $G(\mathbf{C})/G(\mathbf{C} - \mathbf{e}^r)$. where $\mathbf{e}^r = \sum_{l \in r} \mathbf{e}^l$.

2 Reduced-Load Approximation

Unfortunately, the normalizing constant $G(\mathbf{C})$ is difficult to compute for large networks, though some progress can be made using convolution algorithms as in the theory of closed queueing networks [8]. The work done by Le Gall [5] helps to reduce the computation time for $G(\mathbf{C})$ when the capacities are large. In general, the solution is complicated enough that simple approximations are useful. Whitt [10] and Kelly [2] investigated an approximation that Whitt termed the reduced load approximation. The approximation is obtained by focusing on how many circuits are used at different links, rather than on the coordinates of $\mathbf{n}$, and writing equations assuming that links behave independently, as explained in the next paragraph.

Consider a link l and a route r with $l \in r$. For each $j \in r$ with $j \neq l$, the stream of calls offered by route r can be considered to be thinned by a factor $1 - B_j$ at link j before presentation to link l. Pretending that the thinnings occur independently at different links yields that route r effectively offers calls to link l at rate $\nu_r \prod_{j \in r - \{i\}} (1 - B_j)$. Summing these rates over all routes using link l yields the approximate *reduced rate* that calls are offered to link l. Let $E(\nu, C)$ denote Erlang's formula for the blocking probability at a link with C circuits being offered traffic at rate ν, i.e. $E(\nu, C) = \frac{\nu^C}{C!} \left(\sum_{n=0}^{n=C} \frac{\nu^n}{n!} \right)^{-1}$. Assuming each link behaves like an isolated link with Poisson traffic offered at the reduced rate, the *reduced-load approximation* for the blocking probabilities results:

$$B_l = E\left(\sum_{r: l \in r \in \mathcal{R}} \nu_r \prod_{j \in r - \{l\}} (1 - B_j),\ C_l \right) \qquad l \in L. \tag{3}$$

For a fixed vector of offered traffic rates $\nu = (\nu_r : r \in \mathcal{R})$, view the reduced load approximation as a set of equations for the vector of blocking probabilities $\mathbf{B} = (B_l : l \in \mathcal{L})$. The existence of a solution follows from Brouwer's fixed point theorem [1] and uniqueness was proven by Kelly [2].

A slightly different intuitive justification of the reduced-load approximation (3) is given next. Let x_l represent the number of circuits in use at link l. The probability balance equation for the partition of the statespace $\mathcal{S}(\mathbf{C}) = \{x_l \geq i\} \cup \{x_l < i\}$ is

$$i\ P[x_l = i] = \sum_{r:l \in r \in \mathcal{R}} P[x_k < C_k, k \in r;\ x_l = i-1]\nu_r \quad l \in \mathcal{L}, 1 \leq i \leq C_l \tag{4}$$

where $P[A]$ denotes the equilibrium probability of the event A. Pretending that events at distinct links are independent,

$$P[x_k < C_k, k \in r; x_l = j] \ = \ P[x_l = j] \prod_{k \in r, k \neq l} P[x_k < C_k] \ = \ P[x_l = j] \prod_{k \in r, k \neq l} (1 - B_k).$$

Using this approximation, Equation (4) becomes

$$i\ P[x_l = i] \ = \ P[x_l = i-1] \left(\sum_{r:l \in r \in \mathcal{R}} \nu_r \prod_{k \in r, k \neq l} (1 - B_k) \right) \quad l \in \mathcal{L}, i \leq C_l,$$

These equations can be recognized as the probability balance equations for a single link with capacity C_l being offered traffic at the reduced rate discussed earlier, and Equation (3) is again suggested.

It is easy to see that $B_i \leq \bar{B}_i$ where $\bar{B}_i$ is defined to be the right-hand side of (3) when B_j is set to zero for all j. Whitt [10] established the much less obvious fact that the probability a call request for route r is accepted is greater than or equal to $\prod_{i \in r}(1 - \bar{B}_i)$. Whitt's bound is close to the actual acceptance probability when the offered traffic rates are small. An improved bound needs to be found for the case of moderate to large traffic rates.

An ideal outcome would be to obtain bounds which insure the accuracy of the reduced load approximation, which has been found to be accurate in many instances. There is some work in this direction, in the form of asymptotic analysis, which we now describe. Following Pippenger's suggestive terminology [7], the limits that have been considered for the networks $(\mathcal{L}, \mathcal{R}, \mathbf{C}, \nu)$ roughly fall into one of two classes, either *vertical* limits or *horizontal* limits. For vertical limits the network topology is fixed, and the link capacities and the traffic intensities are both increased without bound with their ratios fixed. Kelly [2] proved that in the vertical limit the reduced load approximation correctly predicts the limiting blocking probabilities. Mitra [6] provided an asymptotic expansion in the vertical limit for a special network topology.

For horizontal limits, the link capacities and per-route traffic intensities are kept at moderate levels, but the number of links and routes in the network are increased without bound. Intuitively one feels that if the routes are sufficiently diverse, so that not too much traffic need pass through any two links, then the reduced load approximation should again be asymptotically exact. So far, this result has only been established for highly symmetric networks. The most definitive results are probably those of Ziedens and Kelly [11], which extend results of Whitt [10]. While the published asymptotic results for horizontal limits give some insight, they require a high degree of symmetry in the network. The purpose of this paper is to provide bounds on the accuracy of the reduced load approximation in the case of horizontal limits without invoking a symmetry assumption. Our results to date are limited to the case of networks with all routes using only two links and all link capacities equal to one.

3 Two-Link Calls on Links with Capacity One

Assume that each route uses only two links, and then without loss of generality take $\mathcal{R}$ to be the set of all unordered pairs of elements of $\mathcal{L}$, $\mathcal{R} = \{(l,k) : l,k \in \mathcal{L}\}$. Let ν_r be written as $\nu_{l,k}$ or as $\nu_{k,l}$ when $r = (l,k)$. Also assume that $C_l = 1$ for $l \in \mathcal{L}$. Equation (4) for $i = 1$ reduces to

$$\tilde{B}_l = \sum_{k \in \mathcal{L}} P[x_k = 0, x_l = 0]\nu_{k,l} \tag{5}$$

If the probability $P[x_k = 0, x_l = 0]$ were equal to $P[x_k = 0]P[x_l = 0]$, then Equation (5) would be the same as the reduced load approximation. Thus, the following theorem provides a first step towards finding the accuracy of the reduced-load approximation.

Theorem 1 *Consider a circuit-switched network* $(\mathcal{L}, \mathcal{R}, \mathbf{C}, \nu)$*, where all calls require exactly two links and all links have capacity one. Define* ρ *by* $\rho = \max_{k,l \in \mathcal{L}} \nu_{k,l}$*. Then,*

$$\frac{1}{1+\rho} \leq \frac{P[x_k = 0, x_l = 0]}{P[x_k = 0]P[x_l = 0]} \leq 1 + \rho. \tag{6}$$

Proof. Use proof by induction on $|\mathcal{L}|$, the number of links in the network. The Theorem is trivially true if $|\mathcal{L}| = 1$, so suppose that $|\mathcal{L}| \geq 2$ and that the statement of the Theorem is true for all networks with fewer links. Set $Z_A = P[x_i = 0 \textit{ for all } i \in A]$, and in this connection write i, ij or ijk for $\{i\}$, $\{i,j\}$, or $\{i,j,k\}$, respectively.

Fix two links, k and l. Note that $x_l = 0$ if and only if exactly one of the following is true: Either $x_k = x_l = 0$, or $x_l = 0$ and a call is using route (j,k) for some other link j. Therefore, by Equation (1),

$$Z_l = Z_{kl} + \sum_{j \neq k,l} Z_{jkl}\nu_{jk} \tag{7}$$

Similarly, note that either $x_k = 0$, or there is a call on route (k,l), or there is a call on route (j,k) for some third node j, so that

$$1 = Z_k + Z_{kl}\nu_{kl} + \sum_{j \neq k,l} Z_{jk}\nu_{jk} \tag{8}$$

Multiply each side of Equation (7) by Z_k and each side of Equation (8) by Z_{kl} to obtain

$$\frac{P[x_k = 0, x_l = 0]}{P[x_k = 0]P[x_l = 0]} = \frac{Z_{kl}}{Z_k Z_l}$$

$$= \frac{(Z_k + Z_{kl}\nu_{kl})Z_{kl} + \sum_{j\neq k,l} Z_{jk}Z_{kl}\nu_{jk}}{Z_kZ_{kl} + \sum_{j\neq k,l} Z_kZ_{jkl}\nu_{jk}} \tag{9}$$

Since $Z_{kl} \leq Z_k$ and $0 \leq \nu_{kl} \leq \rho$,

$$1 \leq \frac{Z_k + Z_{kl}\nu_{kl}}{Z_k} \leq 1 + \rho$$

and by the induction hypothesis applied to the network with link k deleted,

$$\frac{1}{1+\rho} \leq \frac{Z_{jk}Z_{kl}}{Z_kZ_{jkl}} \leq 1 + \rho$$

Appling these two inequalities to Equation (9) establishes (6). The proof by induction is complete. □

Defining $\tilde{\nu}_{kl}$ to equal ν_{kl} times the quantity in the middle of (6), and applying Theorem 1, immediately yields the following corollary.

Corollary 1 *Consider a circuit-switched network $(\mathcal{L}, \mathcal{R}, \mathbf{C}, \nu)$, where all calls require exactly two links and all links have capacity one. Then*

$$\frac{\tilde{B}_l}{1-\tilde{B}_l} = \sum_{k\in\mathcal{L}} \tilde{\nu}_{k,l}(1-\tilde{B}_k), \qquad l \in \mathcal{L} \tag{10}$$

for some vector of of rates $(\tilde{\nu}_{k,l})$ satisfying

$$\frac{1}{1+\rho} \leq \frac{\tilde{\nu}_{k,l}}{\nu_{k,l}} \leq 1+\rho, \qquad l, k \in \mathcal{L}. \tag{11}$$

The reduced-load equations under the conditions of this section simplify to

$$\frac{B_l}{1-B_l} - \sum_{j\in\mathcal{L}} \nu_{j,l}(1-B_j) = 0, \qquad l \in \mathcal{L}. \tag{12}$$

Note the similarity between Equations (10) and (12). Since, as noted in Section 2, $\nu = (\nu_r : r \in \mathcal{R})$ and Equation (12) uniquely determine $\mathbf{B} = (B_l : l \in \mathcal{L})$, we can write $\mathbf{B}(\nu)$ to denote $\mathbf{B}$ as a function ν. We can also write $\tilde{\mathbf{B}}(\nu)$ to denote the vector of actual blocking probabilities as a function of ν. Corollary 1 can be restated as $\tilde{\mathbf{B}}(\nu) = \mathbf{B}(\tilde{\nu})$ for some vector $\tilde{\nu}$ satisfying the conditions (11).

This result may be of some interest in its own right. However, to bound how far $\tilde{\mathbf{B}}(\nu) = \mathbf{B}(\tilde{\nu})$ is from $\mathbf{B}(\nu)$, the behavior of $\mathbf{B}$ as a function of ν must be investigated. Specifically, a bound on the sensitivity of $\mathbf{B}(\nu)$, with respect to small changes in ν, must be obtained. The program is carried out in the next section, and the result is summarized in Corollary 2 below.

4 Sensitivity Analysis

Comparison of $\mathbf{B}(\nu)$ and $\mathbf{B}(\tilde{\nu})$ is achieved in this section by bounding the partial derivative of $\mathbf{B}$ along a line segment from ν to $\tilde{\nu}$. First, a way to bound the matrix of partial derivatives $\partial \mathbf{B}/\partial \nu$ is found, and then the chain rule of calculus is applied. The matrix of partial derivatives is useful in its own right for the purpose of optimization, and it plays a key role in Kelly's study of *shadow prices* [4]. Let $F(\nu, \mathbf{B})$ denote the function given by the left-hand side of Equation (12). Then F is a continuous and continuously differentiable mapping from $\mathbf{R}_+^{\mathcal{R}} \times [0,1)^{\mathcal{L}}$ to $\mathbf{R}^{\mathcal{L}}$. The solution we seek corresponds to the set which maps onto 0. Since Equation (12) has a unique solution, the implicit function theorem[1] implies that if the derivative of the mapping F with respect to $\mathbf{B}$ is nonsingular on the set where $F = 0$, then $\mathbf{B}$ can be written as a differentiable function of ν on that set. Let F_l denote the l^{th} coordinate of F. The partial derivatives are easily calculated to be

$$\frac{\partial F_l}{\partial B_i} = \frac{1}{(1-B_l)^2} I_{\{i=l\}} + \nu_{i,l}, \qquad l, i \in \mathcal{L}. \tag{13}$$

$$\frac{\partial F_l}{\partial \nu_\alpha} = -\sum_{j \in \mathcal{L}} (1-B_j) I_{\{(j,l)=\alpha\}}, \qquad l \in \mathcal{L}, \alpha \in \mathcal{R}. \tag{14}$$

For ease of notation, we define matrices β, T, and R as follows, where $|\mathcal{L}|$ represents the cardinality of set $\mathcal{L}$, and $W_{i,j}$ is the entry in row i and column j of a matrix W.

- β: a $|\mathcal{L}| \times |\mathcal{L}|$ diagonal matrix, with $\beta_{i,i} = (1 - B_i)$, $i \in \mathcal{L}$.
- T: a $|\mathcal{L}| \times |\mathcal{R}|$ matrix, with $T_{i,j} = I_{[i \in j]} \prod_{k \in j} (1 - B_k)$, $i \in \mathcal{L}$, $j \in \mathcal{R}$.
- R: a $|\mathcal{L}| \times |\mathcal{L}|$ matrix, with $R_{i,j} = \nu_{i,j}$, $i, j \in \mathcal{L}$.

Equation (13) becomes $\partial F/\partial \mathbf{B} = \beta^{-2} + R$ and Equation (14) becomes $\partial F/\partial \nu = -\beta^{-1} T$. The implicit function theorem implies that if the required matrices are invertible, $\partial \mathbf{B}/\partial \nu = (\beta^{-2} + R)^{-1} \beta^{-1} T = \beta \Lambda T$, where $\Lambda = (I + \beta R \beta)^{-1}$. This expression for $\partial \mathbf{B}/\partial \nu$ is essentially a special case of an equation derived by Kelly[4]. Kelly's paper implies that the partial derivatives are well-defined (i.e. the matrix Λ is well defined), but a bound on Λ is also needed. A simple bound exists for the network model considered in this section:

Lemma 1 *The matrix Λ is well-defined and*

$$\sum_{j \in \mathcal{L}} |\Lambda_{i,j}| \leq \frac{1}{1 - B_*} \qquad i \in \mathcal{L}$$

where $B_* = \max_{l \in \mathcal{L}} B_l$.

Proof. Let $M = \beta R \beta$, so the diagonal elements $M_{i,i}$ are all equal to zero and the off-diagonal elements are given by $M_{i,j} = (1 - B_i)(1 - B_j)\nu_{i,j}$. The elements of M are all nonnegative and Equation (12) may be written as

$$\sum_{j\in\mathcal{L}} M_{i,j} = (1-B_i)\sum_{j\in\mathcal{L}} \nu_{i,j}(1-B_j) = B_i.$$

Thus, if $\mathbf{e}$ denotes the vector with all coordinates equal to 1, $M\mathbf{e} \le B_*\mathbf{e}$, where inequality of vectors is understood coordinate-by-coordinate. By induction, $M^n\mathbf{e} \le B_*^n\mathbf{e}$ for $n \ge 0$, so that Λ is given by the absolutely convergent series $\Lambda = \sum_{n=0}^{\infty}(-1)^n M^n$. Moreover, if $|\Lambda|$ is defined by $|\Lambda|_{i,j} = |\Lambda_{i,j}|$, then

$$|\Lambda|\mathbf{e} \le \sum_{n=0}^{\infty} M^n\mathbf{e} \le \sum_{n=0}^{\infty} B_*^n\mathbf{e} = \frac{1}{1-B_*}\mathbf{e},$$

and the Lemma is proved. □

Recall that the main goal of this section is to compare $\mathbf{B}(\nu)$ and $\mathbf{B}(\tilde{\nu})$ where $\tilde{\nu}$ was introduced in Corollary 1. The bounds (11) imply that $\tilde{\nu}_\alpha = \nu_\alpha(1+\delta_\alpha\rho)$, for some vector $\delta = (\delta_\alpha : \alpha \in \mathcal{R})$ such that $-1/(1+\rho) \le \delta_\alpha \le 1$ and ρ is defined in Theorem 1. As s varies from 0 to ρ, the vector $\hat{\nu}$ defined by $\hat{\nu}_\alpha = \nu_\alpha(1+\delta_\alpha s)$ varies from ν to $\tilde{\nu}$. We now alter the notation and write $\mathbf{B}(s)$ for $\mathbf{B}(\hat{\nu})$. Thus, $\mathbf{B}(0)$ is the vector of approximate blocking probabilities determined by ν and $\mathbf{B}(\rho) = \tilde{\mathbf{B}}(\nu)$ is the vector of actual blocking probabilities determined by ν.

Theorem 2 *Let $\mathbf{B}(s)$ be the solution to the reduced-load Equation (12) when ν_α is replaced by $\hat{\nu}_\alpha = \nu_\alpha(1+\delta_\alpha s)$, and $-1/(1+\rho) \le \delta_\alpha \le 1$. Let $B_*(s) = \max_{l\in\mathcal{L}} B_l(s)$. Then, for $0 \le s \le \rho$,*

$$B_*(0)e^{-s} \le B_*(s) \le B_*(0)e^{s}, \tag{15}$$

and

$$(1-B_l(0))e^{-\tau(s+s^2/2)} \le 1-B_l(s) \le (1-B_l(0))e^{\tau(s+s^2/2)} \tag{16}$$

where $\tau = \max_{i\in\mathcal{L}}(\sum_{j\in\mathcal{L}}\nu_{i,j})$.

Proof. Taking s as the variable, $\frac{dB_l}{ds}$ can be computed using the chain rule and the fact that $\partial\mathbf{B}/\partial\nu = \beta\Lambda T$.

$$\left|\frac{dB_l}{ds}\right| = \left|\sum_{\alpha\in\mathcal{R}} \frac{dB_l}{d\hat{\nu}_\alpha}\nu_\alpha\delta_\alpha\right| \le \sum_{\alpha\in\mathcal{R}}\left|\frac{dB_l}{d\hat{\nu}_\alpha}\nu_\alpha\delta_\alpha\right| \le \sum_{\alpha\in\mathcal{R}} |\nu_\alpha\delta_\alpha|(1-B_l)\sum_{k\in\mathcal{L}}|\Lambda_{l,k}|T_{k,\alpha}$$

Observe that $T_{k,\alpha}$ is nonzero only when $k \in \alpha$. Recall that $\mathcal{R}$ is the set of *unordered* pairs of elements of $\mathcal{L}$, $\mathcal{R} = \{(i,j) : i,j \in \mathcal{L}\}$. With these observations, and substituting for T, the above equation simplifies to

$$\begin{aligned}
\left|\frac{dB_l}{ds}\right| &\le (1-B_l)\sum_{(i,j)\in\mathcal{R}} \nu_{i,j}|\delta_{i,j}|(1-B_i)(1-B_j)(|\Lambda_{l,i}|+|\Lambda_{l,j}|) \\
&= (1-B_l)\sum_{i,j\in\mathcal{L}} \nu_{i,j}|\delta_{i,j}|(1-B_i)(1-B_j)|\Lambda_{l,i}| \\
&= (1-B_l)\sum_{i\in\mathcal{L}}|\Lambda_{l,i}|(1-B_i)\sum_{j\in\mathcal{L}}\left|\frac{\delta_{i,j}}{1+s\delta_{i,j}}\right|\hat{\nu}_{i,j}(1-B_j) \qquad (17)
\end{aligned}$$

where the first equality follows from the fact that the equation is symmetric in i and j and $\mathcal{R} = \{(i,j) : i < j, i, j \in \mathcal{L}\}$. It is easy to obtain the following bound on the $\delta_{i,j}$ term in the above equation,

$$0 \le s \le \rho, \quad \frac{-1}{1+\rho} \le \delta_{i,j} \le 1 \Longrightarrow -1 \le \frac{\delta_{i,j}}{1+s\delta_{i,j}} \le 1$$

Using the above bound and the fact that $B_j(s)$ is the solution to Equation (12) when $\hat{\nu}$ has replaced ν, the second summation in Equation (17) is bounded by $B_i/(1-B_i)$. Thus, we get

$$\left|\frac{dB_l}{ds}\right| \le (1-B_l)\sum_{i\in\mathcal{L}} |\Lambda_{l,i}|B_i \quad \le \quad (1-B_l)\frac{B_*}{1-B_*} \tag{18}$$

where the last inequality follows from Lemma 1 and $B_i \le B_*$.

The set $\mathcal{L}$ has a finite cardinality. Hence, by a standard argument involving Dini derivatives [9, Appendix I], it follows from Equation (18) that B_* satisfies the inequality, $|\frac{dB_*}{ds}| \le B_*$, when $0 \le s \le \rho$. This inequality implies (15).

It is clear from Equation (12) and the definition of $\hat{\nu}$ that $B_*/(1-B_*) < \tau(1+s)$. Then, Equation (18) implies that $|\frac{dB_l}{ds}| \le (1-B_l)\tau(1+s)$, when $0 \le s \le \rho$. The bounds (16) follow. □

Theorem 2 with $s = \rho$ and Corollary 1 together directly imply the main result of this paper, summarized in the following corollary.

Corollary 2 *The actual blocking probabilities $\tilde{B}_l$ satisfy the inequalities, $B_* e^{-\rho} \le \tilde{B}_* \le B_* e^{\rho}$ and*

$$(1-B_l)e^{-\tau(\rho+\rho^2/2)} \le 1 - \tilde{B}_l \le (1-B_l)e^{\tau(\rho+\rho^2/2)},$$

where B_l is given by the solution to Equation (12), $\rho = \max_{k,l\in\mathcal{L}} \ \nu_{k,l}$, $\quad B_* = \max_{l\in\mathcal{L}} \ B_l$, and $\tau = \max_{i\in\mathcal{L}} \left(\sum_{j\in\mathcal{L}} \nu_{i,j}\right)$.

5 Conclusion

Bounds on the accuracy of the reduced-load approximation are given in Corollary 2. The bounds show that as $\rho \to 0$ for τ fixed, the error of the reduced-load approximation tends to zero, uniformly over all networks with the restriction that routes use only two links each and link capacities are limited to one. Hopefully the bounds can be extended to more general network topologies in order to prove, as suggested by Ziedins and Kelly [11], that the reduced-load approximation is accurate for networks with diverse routing.

References

[1] K. Deimling. *Nonlinear Functional Analysis.* Springer-Verlag, Berlin, 1985.

[2] F. P. Kelly. Blocking probabilities in large circuit switched networks. *Advances in Applied Probability*, 18:473–505, 1986.

[3] F. P. Kelly. *Reversibility and Stochastic Networks*. *Wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons, Chichester, 1979.

[4] F. P. Kelly. Routing in circuit-switched networks: optimization, shadow prices and decentralization. *Advances in Applied Probability*, 20:112–144, 1988.

[5] F. LeGall. About loss probabilities for general routing policies in circuit-switched networks. *IEEE Transactions on Communications*, 37(1):57–59, 1989.

[6] D. Mitra. Some results from an asymptotic analysis of a class of simple, circuit-switched networks. In Tijms Boxma, Cohen, editor, *Teletraffic Analysis and Computer Performance Evaluation*, pages 47–62, Elsevier, New York, 1986.

[7] N. Pippenger. 1989. In keynote talk, 27th Annual Allerton Conference on Communications, Control and Computing.

[8] M. Reiser and H. Kobayashi. Queueing networks with multiple closed chains: theory and computational algorithms. *IBM Journal of Research and Development*, 19:283–294, 1975.

[9] N. Rouche, P. Habets, and M. Laloy. *Stability Theory by Liapunov's Direct Method*. Springer-Verlag, New York, Heidelberg, Berlin, 1977.

[10] W. Whitt. Blocking when service is required from several service facilities. *AT&T Tehnical Journal*, 64:1807–1856, 1985.

[11] I. B. Ziedins and F. P. Kelly. Limit theorems for loss networks with diverse routing. *Advances in Applied Probability*, 21:804–830, 1989.