Queueing Syst (2009) 62: 159–193
DOI 10.1007/s11134-009-9118-9

Monotonicity and error bounds for networks of Erlang loss queues

Richard J. Boucherie · Nico M. van Dijk

Received: 9 October 2007 / Revised: 16 April 2009 / Published online: 18 June 2009

Abstract Networks of Erlang loss queues naturally arise when modelling finite communication systems without delays, among which, most notably are

(i) classical circuit switch telephone networks (loss networks) and
(ii) present-day wireless mobile networks.

Performance measures of interest such as loss probabilities or throughputs can be obtained from the steady state distribution. However, while this steady state distribution has a closed product form expression in the first case (loss networks), it does not have one in the second case due to blocked (and lost) handovers. Product form approximations are therefore suggested. These approximations are obtained by a combined modification of both the state space (by a hypercubic expansion) and the transition rates (by extra redial rates). It will be shown that these product form approximations lead to

- upper bounds for loss probabilities and
- analytic error bounds for the accuracy of the approximation for various performance measures.

The proofs of these results rely upon both monotonicity results and an analytic error bound method as based on Markov reward theory. This combination and its technicalities are of interest by themselves. The technical conditions are worked out and verified for two specific applications:

R.J. Boucherie (✉)
Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: r.j.boucherie@utwente.nl

N.M. van Dijk
Department of Operations Research, Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands

- pure loss networks as under (i)
- GSM networks with fixed channel allocation as under (ii).

The results are of practical interest for computational simplifications and, particularly, to guarantee that blocking probabilities do not exceed a given threshold such as for network dimensioning.

Keywords Network of Erlang loss queues · Blocking probabilities · Error bounds

Mathematics Subject Classification (2000) Primary 90B22 · Secondary 60K25

1 Introduction

1.1 Background

The classical Erlang loss model, initially developed for a single telephone switch, is probably the most commonly known queueing model. The loss network is its generalisation to more complex circuit switched systems with multiple links, multiple switches, and multiple types of calls (see [11] for an overview). The loss network is widely used for telephone system dimensioning. The common feature of these networks is that a call arriving to the system either obtains a number of circuits from source to destination and occupies these circuits for its entire duration, or that the call is blocked and cleared because the required circuits for that call are not all available. The corresponding blocking probabilities are among the key performance measures in circuit switched telephone systems. Due to the simple structure of loss networks, their equilibrium distribution has the appealing so-called product form. This product form can be seen as a truncated multidimensional Poisson distribution, where the dimensionality is determined by the number of call types, the parameter of the Poisson distribution is determined by the load offered by all call types, and the truncation is determined by the capacity constraints of the circuits:

$$\pi_{\text{loss}}(\mathbf{n}) = G^{-1}\prod_{k=1}^{N}\frac{\nu_k^{n_k}}{n_k!}, \quad \mathbf{n}\in S,\ G=\sum_{\mathbf{n}\in S}\prod_{k=1}^{N}\frac{\nu_k^{n_k}}{n_k!},\ S=\big\{\mathbf{n}=(n_1,\dots,n_N): A\mathbf{n}\le\mathbf{s}\big\}, \tag{1}$$

where G is a normalising constant, A a $d\times N$ matrix, $\mathbf{s}=(s_1,\dots,s_d)$, with s_i the capacity constraint on circuit i, $i=1,\dots,d$, and d the number of constraints on the capacity of the circuits, $\nu_k=\lambda_k/\mu_k$, with λ_k the arrival rate and $1/\mu_k$ the mean holding time of type k calls, $k=1,\dots,N$, and N is the number of call types, see [11].

A loss network can also be seen as a network of Erlang loss queues with common capacity restrictions. An additional appealing property of the equilibrium distribution π_{loss} is that it is insensitive to the distribution of the call length or holding time apart from its mean. As blocking probabilities can readily be expressed in terms of this equilibrium distribution, the insensitivity property obviously carries over to these blocking probabilities. Although these blocking probabilities are available in closed form, numerical evaluation requires evaluation of the normalising constant G^{-1}. The size of the state space considerably complicates this evaluation. To this end, various efficient numerical evaluation and approximation schemes have been developed, including Monte Carlo summation, and Erlang fixed point methods, see [11, 20].

In mobile communications networks, a call may transfer from one cell to another while in progress. As a consequence, in addition to fresh call blocking of a newly arriving call, handover blocking for a call which attempts to route to another cell, but which finds all circuits available for this cell occupied, becomes of practical interest. In that case, the blocked handover is cleared and lost. A network of Erlang loss queues with routing and common capacity restrictions is a natural representation of this network.

The equilibrium distribution for a network of Erlang loss queues with handover blocking is, unfortunately, not available in closed form. Various approximations have therefore been suggested in the literature. The most appealing among these approximations is the redial rate approximation introduced in [4]. Under the redial rate approximation, an extra arrival rate of calls in cells surrounding a blocked cell is introduced. This redial rate mimics the behaviour of calls that are lost when transferring to the blocked cell. This approximation retains the call blocking structure of the original model. Under maximal redial rates, when all blocked calls attempt to redial, the equilibrium distribution is of product form, similar to that for the loss network. Moreover, the equilibrium distribution and blocking probabilities inherit the appealing insensitivity property. As the equilibrium distribution under the redial rate approximation also has a truncated multidimensional Poisson distribution, computational techniques developed for loss networks can be carried over to numerically evaluate fresh call and handover blocking probabilities.

1.2 Results

The redial rate approximation of blocking probabilities introduces an approximation error. However, as of yet no formal support for the accuracy of this approximation or other approximations appears to be available in the literature. For practical purposes, at least an upper bound for blocking probabilities would be of most interest as blocking probabilities are mainly used for dimensioning. In addition, an error bound on the accuracy of this bound would substantially enlarge its applicability. This paper therefore aims to establish both

- upper bounds for blocking probabilities, and
- analytical error bounds on the approximation error for specific performance measures as based on Erlang loss queue approximations.

The first result (a monotonicity result) may seem intuitively obvious, since the redial rate approximation introduces an extra arrival rate of fresh calls on circuits that are neighbours of a blocked circuit. However, as shown by an example, see Sect. 4.4, the result does not apply in general: adding extra calls on some circuits may reduce blocking probabilities in particular circuits. It is the careful combination of redial rates and state space modification that yields the monotonicity result. The monotonicity results are not only of interest to establish the bounds, but are also required for obtaining the error bounds.

The approximation error is shown to be roughly of the order of magnitude of the blocking probabilities. For dimensioning of networks with an increasing offered load this is appealing, since dimensioning based on the upper bound guarantees that blocking probabilities do not exceed a given threshold. For example, with approximate loss probabilities in the order of up to 0.5%, it would secure actual loss probabilities in the order of 1%.

As both a system and state space modification are involved, the bounds and the approximation errors need to be obtained in two steps. These steps have not been used before in the literature and appear to become rather technical. First, we will obtain a bound and an error bound due to increasing the state space to a hypercube $S_{\mathrm{hc}} = \{\mathbf{n} : 0 \le n_i \le N_i\}$, $N_i = \max\{n_i : \mathbf{n} \in S\}$, $i = 1, \ldots, N$, that contains the original state space S. The equilibrium distribution of both the original process and the process on this hypercube are not available in closed form. Next, we show that increasing the redial rates for the process on the hypercube increases blocking probabilities. In addition, an error bound is established for the accuracy under increasing redial rates. In particular, under maximal redial rates, when all calls that have lost their connection attempt to redial, the equilibrium distribution has a truncated multivariate Poissonian form, which leads to a closed form expression for the blocking probabilities.

The monotonicity and error bound results cover performance measures which are increasing in all components of the state. This includes fresh call and handover blocking probabilities as well as throughputs. With A_0 the performance measure for the original process, and $A_{\mathrm{hc},r}$ for the process on the hypercube under redial rates, the main result states that

$$A_{\mathrm{hc},r} - (\beta + \beta_{r0}) \le A_0 \le A_{\mathrm{hc},r} \le A_0 + (\beta + \beta_{r0}),$$

where the parameter β characterises the approximation error due to the state space modification from S to S_{hc}, and the parameter β_{r0} characterises the error due to the redial rate approximation on the hypercube state space. The parameters are determined by the arrival and service rates, and the equilibrium distribution on the hypercube under maximal redial rates is of product form:

$$\pi_{\mathrm{hc},r}(\mathbf{n}) = \prod_{i=1}^{N} \left[\frac{\nu_i^{n_i}}{n_i!} \Big/ \sum_{j=0}^{N_i} \frac{\nu_i^j}{j!} \right], \quad \mathbf{n} \in S_{\mathrm{hc}}.$$

The result states that the approximation $A_{\mathrm{hc},r}$ is an upper bound on A_0, and that this upper bound differs no more than $\beta + \beta_{r0}$ from A_0. In applications, $\beta + \beta_{r0}$ is often of the order of magnitude of $A_{\mathrm{hc},r}$, so that the bound is applicable for dimensioning: dimensioning the system based on a guaranteed upper bound implies that the actual system performs better than the target values.

1.2.1 Outline of proofs

The proofs are obtained in two steps. First monotonicity is demonstrated for the state space modification, where the original process is shown to be stochastically dominated by the process with the same transition structure on a larger state space, e.g. on the hypercube $S_{\mathrm{hc}} \supset S$. Then, monotonicity is demonstrated in the redial rates of the process on the hypercube. For the maximal value of the redial rates the process has a product form equilibrium distribution. Due to the hypercube state space, this enables us to obtain blocking probabilities directly from the Erlang loss formula.

For the second result (the error bound) first a general error bound result will be presented that expresses the error in the equilibrium distribution of the approximating model. Next, as a special case, a simple analytical bound is provided for the redial rate approximation on the hypercube. The proof of the error bound result requires both the monotonicity results and a Markov reward approach. In the Markov reward approach, rewards are associated with the performance measures. For example, for a blocking probability the process incurs a reward rate 1 per unit time spent in a state in which blocking would take place. Based upon the combination of the special reward and structural properties of the transition structure, monotonicity properties and error bounds for that specific performance measure can then be derived.

1.3 Literature

The results of this paper are based on monotonicity and error bounds that relate performance measures to their approximation by a product form network. The equilibrium distribution of the product form network coincides with that of an Erlang loss network. Product form approximations for networks of Erlang loss queues with routing have been discussed by various authors, see e.g. [4, 8, 18]. The redial rate approximation was introduced in [4], and generalised to networks with general call lengths in [5], which also investigates insensitivity. Performance measures for networks of Erlang loss queues with routing have been analysed in a variety of papers, see e.g. [9, 18, 19]. Performance measures and their numerical evaluation and approximation for loss networks have been addressed in a series of papers, see [11], and [20] for an overview and further references.

For the estimation of blocking probabilities, in this paper we have a twofold interest: to prove an upper bound and to establish an error bound for its accuracy. To prove bounds, the stochastic monotonicity approach by sample path comparison is widely used in the literature, see [2, 10, 12–17, 26, 28, 29]. However, while this approach is straightforward for unrestricted (or infinite) queueing systems (e.g. [2, 16, 17, 22, 28]), it is not for finite systems. For finite queueing systems a proof of stochastic monotonicity leads to complications as ‘overtaking’ might take place so that interchangeability arguments have to be used based on exponentiality assumptions [1, 26]. However, these arguments cannot be applied in mobile networks as exponential calls are no longer indistinguishable due to their location (also see [14]). In order to establish error bounds, in this paper therefore we will use a combined approach based on both monotonicity results and the Markov reward technique, see e.g. [23, 24, 27] for a survey of this technique.

1.4 Organisation

The organisation of this paper is as follows. Section 2 contains the model, the performance measures of interest, and the product form modifications. In particular, a network with unlimited capacity is used to introduce the offered load that characterises the equilibrium distribution under the redial rate approximation that is described in Sect. 2.3. Section 3 contains the main monotonicity and error bound results. The technical proofs of these results are concentrated in Sect. 5 along with additional comments. Section 4 provides two special applications which include

- a computational simplification for loss networks
- and an explicit error bound for GSM networks with fixed channel allocation.

2 Model

2.1 Markov chain

Consider a wireless communication network consisting of N cells, labelled $i = 1, 2, \ldots, N$. Calls arrive to cell i according to a Poisson process with rate λ_i (fresh calls). A successfully completed call has a negative exponentially distributed call length with mean $1/\mu$. Calls may move around in the network. A call may move from cell i to neighbouring cell k at exponential rate λ_{ik} (handover), provided the new state is feasible, $i, k = 1, \ldots, N$. A fresh call or handover leading to an infeasible state is blocked and cleared. This is referred to as fresh call blocking and handover blocking. The network can thus be represented by an exponential queueing network, with

$$\begin{array}{ll} \lambda_i & \text{arrival rate to cell } i, \\ \mu_i = \mu + \sum_k \lambda_{ik} & \text{holding time parameter in cell } i, \\ p_{ij} = \lambda_{ij}/\mu_i & \text{handover probability from cell } i \text{ to cell } j \text{, and,} \\ p_{i0} = \mu/\mu_i & \text{the successful call completion probability in cell } i. \end{array} \tag{2}$$

A state of this network is a vector $\mathbf{n} = (n_1, n_2, \ldots, n_N)$, where n_i is the number of calls in progress in cell i, $i = 1, 2, \ldots, N$. Due to interference constraints or resource sharing, the states are limited to a set of feasible states

$$S = \{\mathbf{n} : A\mathbf{n} \le \mathbf{s}\}, \tag{3}$$

where A is a $d \times N$ matrix, $\mathbf{s}$ is a d-vector, and d is the number of constraints, see [9]. A state space of this form also arises in a loss network, see [11].

The exponentiality assumptions imply that the state of the network can be represented as a continuous-time Markov chain, $\mathbf{X} = (X(t),\ t \ge 0)$, that records the number of calls in the cells. The Markov chain has transition rates, $Q = (q(\mathbf{n}, \mathbf{n}'),\ \mathbf{n}, \mathbf{n}' \in S)$, with non-zero entries for $\mathbf{n}' \ne \mathbf{n}$ given by

$$q(\mathbf{n}, \mathbf{n}') = \begin{cases} \lambda_i 1(\mathbf{n} + \mathbf{e}_i \in S), & \mathbf{n}' = \mathbf{n} + \mathbf{e}_i, \quad \text{fresh call,} \\ n_i \mu_i p_{i0}, & \mathbf{n}' = \mathbf{n} - \mathbf{e}_i, \quad \text{call completion,} \\ n_i \mu_i p_{ik} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_k \in S), & \mathbf{n}' = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_k, \quad \text{handover,} \\ \sum_{k=1}^{N} n_i \mu_i p_{ik} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_k \notin S), & \mathbf{n}' = \mathbf{n} - \mathbf{e}_i, \quad \text{blocked handover,} \end{cases} \tag{4}$$

where $\mathbf{e}_i$ is the i-th unit vector with 1 in place i, 0 elsewhere, $1(A)$ is the indicator function of event A, that is 1 when A occurs, 0 otherwise, and the diagonal elements $q(\mathbf{n}, \mathbf{n})$ are such that the row sums equal zero. Note that the transition rates for a successful call completion or a blocked handover effectively lead to the same transition and can be combined. Nevertheless, we have listed these transition rates separately to distinguish the two events, which may have different consequences for the performance measure of interest, e.g. throughput or handover blocking. This wireless network can thus be regarded as a network of Erlang loss queues with additional state space restrictions in which customers arriving to a queue resulting in an unfeasible state are blocked and cleared from the system. For a more detailed description of a wireless network, its relation to a queueing network, and generalisations to general holding times, see [4, 5]. The equilibrium distribution, π, is the unique non-negative probability solution of the *global balance equations*

$$\pi Q = 0.$$

Remark 2.1 (Product form?) We distinguish two cases of computational interest. *Without handovers*, i.e., $p_{ij} = 0$ for all i, j, the network is called a loss network. In this case, the equilibrium distribution π is well known to have a truncated multivariate Poisson distribution as represented by (1), see [11]. This distribution is also referred to as a product form distribution. Nevertheless, due to the state space restrictions its computation can still be numerically demanding. *With handovers*, this appealing product form property will in general no longer apply due to the capacity restrictions, except for special instances such as with reversible routing. Several modifications of the transition rates have been suggested in the literature, e.g. [4, 18]. In this paper, we use the redial rate approximation introduced in [4]. This approximation is based on a truncation of a network with unlimited capacity, such that the transition rates resulting in blocked and cleared calls are preserved and compensated. The redial rate approximation will be introduced in Sect. 2.3. In this paper, we will show that this approximation leads to bounds for loss probabilities and we will derive an analytic error bound on the error in the blocking probabilities.

2.2 Performance measures

The *fresh call blocking probability*, B_i, that an additional call in cell i cannot be accepted, can be expressed as a summation of π over part of the boundary of the state space (see [3], or directly by using PASTA):

$$B_i = \frac{\sum_{\mathbf{n}\in S} \pi(\mathbf{n})\lambda_i 1(\mathbf{n}+\mathbf{e}_i \notin S)}{\sum_{\mathbf{n}\in S} \pi(\mathbf{n})\lambda_i} = \sum_{\mathbf{n}\in T_i} \pi(\mathbf{n}), \qquad T_i := \{\mathbf{n} : \mathbf{n}\in S,\ \mathbf{n}+\mathbf{e}_i \notin S\}.$$

The *handover blocking probability*, B_{ij}, that a handover from cell i to cell j is blocked, is (see [3])

$$B_{ij} = \frac{\sum_{\mathbf{n}\in S} \pi(\mathbf{n}) n_i \mu_i p_{ij} 1(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j \notin S)}{\sum_{\mathbf{n}\in S} \pi(\mathbf{n}) n_i \mu_i p_{ij}} = \frac{\sum_{\mathbf{n}\in S} \pi(\mathbf{n}) n_i 1(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j \notin S)}{\sum_{\mathbf{n}\in S} \pi(\mathbf{n}) n_i}.$$

The *call dropping probability*, D_i, that a call terminates in cell i due to an unsuccessful handover, is expressed by

$$D_i = \frac{\sum_{\mathbf{n}\in S}\sum_j \pi(\mathbf{n}) n_i \mu_i p_{ij} 1(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j \notin S)}{\sum_{\mathbf{n}\in S}\sum_j \pi(\mathbf{n}) n_i \mu_i p_{ij} 1(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j \notin S) + \sum_{\mathbf{n}\in S} \pi(\mathbf{n}) n_i \mu_i p_{i0}}.$$

The throughput or number of successful call completions, H_i, is given by

$$H_i = \sum_{\mathbf{n}\in S} \pi(\mathbf{n}) n_i \mu_i p_{i0},$$

which can be used to obtain the denominator of the handover blocking probabilities.

2.3 Product form modification

This section presents two modifications to obtain an amenable product form distribution. The first one is the system with unlimited capacity. This system has a natural interpretation of the *traffic equations* and their solution, the *offered load*, that characterise product forms. The second one is the redial rate approximation which we will use as product form approximation throughout this paper.

2.3.1 Unlimited capacity

For the system with unlimited capacity, the state space is unlimited, that is $S_\infty = \{\mathbf{n} : \mathbf{n} \geq \mathbf{0}\}$, and the equilibrium distribution also exhibits the factorising multidimensional Poisson form (1) but with $G = \prod_k e^{\nu_k}$, and $\{\nu_i\}_{i=1}^N$ the unique solution of the *traffic equations*

$$\nu_i \mu_i = \lambda_i + \sum_{j=1}^N \nu_j \mu_j p_{ji}, \quad i = 1, \dots, N. \tag{5}$$

In this case the equilibrium distribution satisfies the partial balance equations

$$\sum_{j=0}^N \big\{\pi_\infty(\mathbf{n}) q(\mathbf{n}, \mathbf{n}-\mathbf{e}_i+\mathbf{e}_j) - \pi_\infty(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j) q(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j, \mathbf{n})\big\} = 0, \quad i = 0, \dots, N,$$

where $\mathbf{e}_0 = \mathbf{0}$, the vector with each element zero.

Remark 2.2 (Traffic equations; offered load) The traffic equations (5) determine the average load of the cells in the case of infinite capacities: ν_i can be interpreted as the load offered per time unit to cell i, which consists of the arrival rate of fresh calls, λ_i, and the arrival rate, $\nu_j \mu_j p_{ji}$, due to handovers from other cells $j = 1, \dots, N$. To this end, observe that in the network with infinite capacity calls move independently among the cells of the network, so that the mean flow of calls from cell k to cell i is

$$\sum_{\mathbf{n}\geq\mathbf{0}} \pi_\infty(\mathbf{n}) n_k \mu_k p_{ki} = \nu_k \mu_k p_{ki}.$$

2.3.2 Redial rates

For networks with finite capacities, closed form solutions for the equilibrium distribution or blocking probabilities are generally not available. In [4], it is shown that the introduction of *redial rates* re-establishes a *product form* or *truncated multidimensional Poisson* equilibrium distribution. Such distributions are commonly used for studying circuit switched or wireless communications networks, most notably loss networks. Various computational methods for efficiently computing performance measures have therefore been studied, see e.g. [20] for Monte Carlo methods, and [6] for an efficient asymptotic approximation method.

Under the redial rate approximation from [4], the state space S is allowed to have the general form (3). The Markov chain $X_r = (X_r(t),\ t > 0)$ now has transition rates $Q_r = (q_r(\mathbf{n}, \mathbf{n}'),\ \mathbf{n}, \mathbf{n}' \in S)$, with non-zero entries for $\mathbf{n}' \neq \mathbf{n}$ given by

$$q_r(\mathbf{n}, \mathbf{n}') = \begin{cases} \lambda_i 1(\mathbf{n} + \mathbf{e}_i \in S), & \mathbf{n}' = \mathbf{n} + \mathbf{e}_i & \text{fresh call,} \\ n_i \mu_i p_{i0}, & \mathbf{n}' = \mathbf{n} - \mathbf{e}_i & \text{call completion,} \\ n_i \mu_i p_{ik} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_k \in S), & \mathbf{n}' = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_k & \text{handover,} \\ \sum_{k=1}^{N} n_i \mu_i p_{ik} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_k \notin S), & \mathbf{n}' = \mathbf{n} - \mathbf{e}_i & \text{blocked handover,} \\ \sum_{k=1}^{N} r_{ki} 1(\mathbf{n} + \mathbf{e}_i \in S,\ \mathbf{n} + \mathbf{e}_k \notin S), & \mathbf{n}' = \mathbf{n} + \mathbf{e}_i & \text{redial attempt,} \end{cases} \tag{6}$$

where r_{ki} is the redial rate in cell i when the neighbouring cell k is blocked, and the diagonal elements $q_r(\mathbf{n}, \mathbf{n})$ are such that the row sums equal zero. The following result is obtained in [4], where it is shown that the redial rates preserve partial balance at the boundary of the state space. The redial rates are discussed in Remark 2.5 below.

Theorem 2.3 *Let $\{\nu_i\}_{i=1}^N$ be the (unique) solution of the traffic equations (5), and assume that the redial rates are such that*

$$r_{ki} = \nu_k \mu_k p_{ki}, \quad k, i = 1, \ldots, N. \tag{7}$$

Then the equilibrium distribution π_r of X_r is a truncated multivariate Poisson distribution

$$\pi_r(\mathbf{n}) = G^{-1} \prod_{k=1}^{N} \frac{\nu_k^{n_k}}{n_k!}, \quad \mathbf{n} \in S,\ G = \sum_{\mathbf{n} \in S} \prod_{k=1}^{N} \frac{\nu_k^{n_k}}{n_k!}. \tag{8}$$

Remark 2.4 (Notation) Note that the original process is obtained by setting $r_{kj} = 0$ for all k, j. We will formulate our results for general values for the redial rates r_{kj}, with the original process as a special case.

Remark 2.5 (Interpretation of the redial rates; maximal redial rates) The redial rates r_{ki} are introduced for analytical tractability. For the values given in (7) the equilibrium distribution is a multivariate Poisson distribution.

The redial rates can be interpreted as follows. The redial rate r_{ki} represents the subscribers that have lost their connection in cell k (as fresh call, as handover, or possibly due to fading). These subscribers try to re-establish their connection in neighbouring cells when they are close to the border of cell k. Since the mean rate of subscribers with blocked calls from cell k to cell i cannot exceed the mean flow of handovers from cell k to cell i in the system with unlimited capacity, it is natural to restrict the redial rates such that

$$0 \le r_{ki} \le \nu_k \mu_k p_{ki}, \tag{9}$$

where the maximal value corresponds to the network in which *all* subscribers try to re-establish their connection. Since the redial behaviour is modelled as a Poisson arrival process, this is clearly an approximation of the actual redial behaviour that may occur in a mobile network. Intuition suggests that the redial rate approximation leads to an overestimation of blocking probabilities since the network seems to contain more calls. Due to the intricate relation between the constraints determining the state space S this can, in general, not be shown at the sample path level. Nevertheless, in Sect. 3 we show that blocking probabilities under the maximal redial rates, defined as $r_{ki} = \nu_k \mu_k p_{ki}$, do indeed overestimate the actual blocking probabilities.

Blocking probabilities can be obtained in closed form from the distribution (8). In particular, the fresh call, $B_{r,i}$, and handover blocking probabilities, $B_{r,ij}$, have the appealing forms (see [4])

$$B_{r,i} = \frac{\sum_{\mathbf{n}\in T_i} \prod_{k=1}^N (\nu_k^{n_k}/n_k!)}{\sum_{\mathbf{n}\in S} \prod_{k=1}^N (\nu_k^{n_k}/n_k!)}, \qquad B_{r,ij} = \frac{\sum_{\mathbf{n}\in T_{ij}} \prod_{k=1}^N (\nu_k^{n_k}/n_k!)}{\sum_{\mathbf{n}\in U_i} \prod_{k=1}^N (\nu_k^{n_k}/n_k!)}, \tag{10}$$

with

$$T_i = \{\mathbf{n} : \mathbf{n} \in S,\ \mathbf{n} + \mathbf{e}_i \notin S\}, \qquad U_i := \{\mathbf{n} : \mathbf{n} + \mathbf{e}_i \in S\} \quad \text{and}$$
$$T_{ij} := \{\mathbf{n} : \mathbf{n} + \mathbf{e}_i \in S,\ \mathbf{n} + \mathbf{e}_j \notin S\}.$$

2.4 Hypercube modification

As a special redial and state space modification, for a given original network with state space S, we define the hypercube state space

$$S_{\text{hc}} = \{\mathbf{n} : 0 \le n_i \le N_i,\ i = 1, \ldots, N\}, \quad N_i = \max\{n_i : \mathbf{n} \in S\},$$

with transition rates $Q_{\text{hc},r} = (q_{\text{hc},r}(\mathbf{n}, \mathbf{n}'), n, n' \in S_{\text{hc}})$ as defined in (6), but now with S replaced by S_{hc}, and assuming the maximal redial rates: $r_{ki} = \nu_k \mu_k p_{ki}$. It can then easily be shown that the equilibrium distribution of this hypercube process factorises over the queues:

$$\pi_{\text{hc},r}(\mathbf{n}) = \prod_{i=1}^N \left[\frac{\nu_i^{n_i}}{n_i!} \bigg/ \sum_{j=0}^{N_i} \frac{\nu_i^j}{j!} \right], \quad \mathbf{n} \in S_{\text{hc}}.$$

As a consequence, with respect to blocking probabilities, each queue behaves as an Erlang loss queue in isolation with arrival rate determined by the traffic equations. The fresh call and handover blocking probabilities thus reduce to the Erlang loss probabilities, see [4]:

$$B_{\mathrm{hc},r,i} = B_{\mathrm{hc},r,ji} = B_{\mathrm{loss}} = \frac{\nu_i^{N_i}}{N_i!} \Big/ \sum_{k=0}^{N_i} \frac{\nu_i^k}{k!}, \quad i, j = 1, \ldots, N.$$

Remark 2.6 (Other product form modifications) Other product form modifications such as a stop, recirculate, and jump-over protocol can also be used, see [25]. All these protocols lead to an equilibrium distribution that is functionally the same as obtained under the redial protocol. However, under the stop and recirculate protocols, transitions leading to call blocking are removed. This is less appropriate for analysing blocking probabilities. In addition, under a stop or recirculate protocol, approximation error bounds cannot, in general, be obtained.

3 Main results

This section provides our main practical result (Corollary 3.6). This result is based on two more technical results (Theorems 3.1, 3.4). The proofs of these results are concentrated in Sect. 5. First, we investigate monotonicity of the process in the state space and the redial rates. The second result provides an analytic error bound on the redial rate approximation. This result consists of two components: an error bound for the hypercube modification, and an error bound for the redial rate approximation of the hypercube process. Examples are included in Sect. 3.2.

3.1 General results

Consider the set of functions defined as

$$C_{\mathrm{hc}} = \left\{ f : S_{\mathrm{hc}} \to [0, \infty) \,|\, f(\mathbf{n} + \mathbf{e}_i) - f(\mathbf{n}) \geq 0, \text{ for } \mathbf{n}, \mathbf{n} + \mathbf{e}_i \in S_{\mathrm{hc}} \right\}.$$

The family of functions $f \in C_{\mathrm{hc}}$ includes, for example, fresh call blocking in cell i by $f(\mathbf{n}) = 1(\mathbf{n} \notin T_i)$.

The following theorem, which combines Lemma 5.6 and Theorem 5.7, provides our main monotonicity result. For $f \in C_{\mathrm{hc}}$ for the hypercube process, $\mathbb{E}_r f \equiv \sum_{\mathbf{n} \in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) f(\mathbf{n})$ is increasing in the redial rates. This result implies that the product form approximation that is obtained under maximal redial rates provides an upper bound for $\mathbb{E}_0 f$, the expectation of f for the original process.

Theorem 3.1 (Main monotonicity result) *When $r_{ji} \geq r'_{ji}$ for all j, i then for any $f \in C_{\mathrm{hc}}$*

$$\sum_{\mathbf{n} \in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) f(\mathbf{n}) \geq \sum_{\mathbf{n} \in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r'}(\mathbf{n}) f(\mathbf{n}),$$

and for any $f \in C_{\mathrm{hc}}$

$$\sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) f(\mathbf{n}) \geq \sum_{\mathbf{n}\in S} \pi_0(\mathbf{n}) f(\mathbf{n}).$$

Remark 3.2 (Literature) Theorem 3.1 generalises a result from [1]. In this reference, a similar result was established for fresh call blocking only, by a sample path argument. In the present, more general, setting that involves both redial rates and a state space modification, a sample path argument can no longer be given. We will use Theorem 3.1 to demonstrate Theorem 3.4. Theorem 3.1 is of theoretical interest by itself and provides monotonicity results in both the redial rates and the state space modification.

Theorem 3.4 will provide both an upper and a lower bound on the approximation error. Intuitively, it seems obvious that higher redial rates result in higher blocking probabilities. However, accepting a customer in one queue may lead to a smaller number of customers in other queues due to joint capacity constraints, which may lead to counterintuitive results (see Sect. 4.4). Nevertheless, monotonicity will appear for the hypercube process.

The theorem involves the following condition on the reward rate R, where X incurs a reward $R(\mathbf{n})$ per time unit that X spends in state $\mathbf{n}$.

Condition 3.3 Assume that for all $\mathbf{n}, \mathbf{n}+\mathbf{e}_i \in S_{\mathrm{hc}}$ the reward rate is such that on the hypercube state space S_{hc}

$$\begin{aligned}
0 &\leq R(\mathbf{n}+\mathbf{e}_i) - R(\mathbf{n}) \qquad (11)\\
&\leq \lambda_i 1(\mathbf{n}+2\mathbf{e}_i \notin S) + \sum_{j=1}^{N} n_j \mu_j p_{ji} 1(\mathbf{n}+2\mathbf{e}_i \notin S) + \mu_i p_{i0}\\
&\quad + \sum_{k=1}^{N} \mu_i p_{ik} 1(\mathbf{n}+\mathbf{e}_k \notin S). \qquad (12)
\end{aligned}$$

In Sect. 3.2, it is demonstrated that this condition is satisfied for fresh call blocking and throughput.

The following theorem, which is a combination of Theorem 5.13 and Theorem 5.16, yields our main error bound result.

Theorem 3.4 (Main error bound result) *Under Condition* 3.3

$$A_{\mathrm{hc},r} - (\beta + \beta_{r0}) \leq A_0 \leq A_{\mathrm{hc},r} \leq A_0 + (\beta + \beta_{r0}), \qquad (13)$$

where

$$\beta = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) \Phi(\mathbf{n}), \qquad \beta_{rr'} = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) \Phi_{rr'}(\mathbf{n}),$$

with

$$\Phi(\mathbf{n}) = \sum_j \lambda_j 1(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}} \backslash S) + \sum_{i,j} n_i \mu_i p_{ij} 1(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j \in S_{\mathrm{hc}} \backslash S),$$

$$\Phi_{rr'}(\mathbf{n}) = \sum_{k,j} \big(r_{kj} - r'_{kj}\big) 1(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}}, \mathbf{n}+\mathbf{e}_k \notin S_{\mathrm{hc}}) \quad (\textit{with } r_{kj} \geq r'_{kj}).$$

Remark 3.5 Condition 3.3 distinguishes two conditions that each have their specific function. The monotonicity condition (11) implies the ordering $A_0 \leq A_{\mathrm{hc},0}$ so that by Theorem 3.1 also $A_0 \leq A_{\mathrm{hc},0} \leq A_{\mathrm{hc},r}$. The bounding condition (12) will lead to the error bound $A_{\mathrm{hc},r} - A_0| \leq \beta + \beta_{r0}$.

Theorem 3.4 also provides a bound on the error in the upper bound $A_{\mathrm{hc},r}$ of A_0. Often, $\beta + \beta_{r0}$ has the order of magnitude of $A_{\mathrm{hc},r}$ so that the upper bound is roughly twice the value of A_0. For applications in wireless networks, where typical values for the blocking probabilities are 1%, this is an acceptable level of accuracy: dimensioning the system based on a guaranteed upper bound of 1% implies that the actual system performs better than the target values.

The proof of Theorem 3.4 is provided in Sect. 5, and consists of two steps that cannot be combined into a single step. The first step compares the original process X_0 on state space S with the hypercube process $X_{\mathrm{hc},0}$ on state space S_{hc}. Here the boundary of the state space S plays a crucial role. The contribution to the error bound is denoted by β. The second step compares the process $X_{\mathrm{hc},0}$ with the process $X_{\mathrm{hc},r}$. The essential step consists of a comparison of the redial rates at the boundary of S_{hc}. The contribution in the error bound is denoted by β_{r0}.

Under maximal redial rates the equilibrium distribution is of product form. The following corollary is therefore of computational interest. For practical purposes, this corollary can be regarded as the main result of this paper. The result immediately follows from Theorem 3.4 and results from Sect. 2.4.

Corollary 3.6 (Main product form error bound result) *Under Condition* 3.3, *and under maximal redial rates defined as*

$$r_{ki} = \nu_k \mu_k p_{ki}, \quad k, i = 1, \ldots, N,$$

(13) *applies with*

$$\pi_{\mathrm{hc},r}(\mathbf{n}) = \prod_{i=1}^{N} \left[\frac{\nu_i^{n_i}}{n_i!} \Big/ \sum_{j=0}^{N_i} \frac{\nu_i^j}{j!} \right], \quad \mathbf{n} \in S_{\mathrm{hc}}.$$

A disadvantage of the error bound result above, or its product form version of Theorem 3.4, is that the error bound terms β_r and $\beta_{rr'}$ require summation of the equilibrium distribution $\pi_{\mathrm{hc},r}$ over $S_{\mathrm{hc}} \setminus S$. This summation can, in general, not efficiently be evaluated in closed form. Sections 4.1 and 4.3 will therefore address an efficient estimation of these summations.

3.2 Examples

The main condition for Theorem 3.4 and Corollary 3.6 is the reward condition (Condition 3.3). This condition may seem more restrictive than it actually is. For the hypercube process, it does allow performance functions that reflect fresh call blocking, handover blocking, and throughput, as will be shown below.

3.2.1 *Fresh call blocking*

For $\mathbf{n} \in S_{\mathrm{hc}}$, and fixed j, let $R(\mathbf{n}) = \lambda_j 1(\mathbf{n}+\mathbf{e}_j \notin S_{\mathrm{hc}})$. We have $T_j = \{\mathbf{n} : \mathbf{n} \in S_{\mathrm{hc}}, \mathbf{n}+\mathbf{e}_j \notin S_{\mathrm{hc}}\}$. Then, $R \in C_{\mathrm{hc}}$, and for $\mathbf{n}+\mathbf{e}_i \in S_{\mathrm{hc}}$:

$$\begin{aligned} R(\mathbf{n}+\mathbf{e}_i) - R(\mathbf{n}) &= \lambda_j 1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \notin S_{\mathrm{hc}}) - \lambda_j 1(\mathbf{n}+\mathbf{e}_j \notin S_{\mathrm{hc}}) \\ &= \lambda_j 1(\mathbf{n}+2\mathbf{e}_i \notin S_{\mathrm{hc}}) 1(i=j). \end{aligned}$$

Thus R satisfies Condition 3.3, and as the corresponding performance measure, we obtain the fresh call blocking probability in cell j

$$A_{\mathrm{hc},r} = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) R(\mathbf{n}) = \lambda_j B_{\mathrm{hc},r,j}.$$

3.2.2 *Handover blocking and dropping*

For $\mathbf{n} \in S_{\mathrm{hc}}$, and fixed k, let $R(\mathbf{n}) = \sum_{j=1}^N n_j \mu_j p_{jk} 1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \notin S_{\mathrm{hc}})$. Then, for $\mathbf{n}+\mathbf{e}_i \in S_{\mathrm{hc}}$:

$$\begin{aligned} R(\mathbf{n}+\mathbf{e}_i) - R(\mathbf{n}) &= \sum_{j=1}^N \bigl\{ \bigl(n_j + 1(i=j)\bigr) \mu_j p_{jk} 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k \notin S_{\mathrm{hc}}) \\ &\qquad - n_j \mu_j p_{jk} 1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \notin S_{\mathrm{hc}}) \bigr\} \\ &= \sum_{j=1}^N n_j \mu_j p_{ji} 1(\mathbf{n}+2\mathbf{e}_i \notin S_{\mathrm{hc}}) 1(i=k), \end{aligned}$$

where we have used the observation that the right-hand side is non-null only for $k=i$, which also implies that $j \neq i$. Clearly, R satisfies Condition 3.3. We find

$$A_{\mathrm{hc},r} = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) R(\mathbf{n}) = \sum_{\mathbf{n}\in S} \sum_{j=1}^N \pi_{\mathrm{hc},r}(\mathbf{n}) n_j \mu_j p_{jk} 1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \notin S_{\mathrm{hc}}),$$

which represents the numerator of the call dropping probability in cell k. By analogy, for $R(\mathbf{n}) = n_j \mu_j p_{jk} 1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \notin S_{\mathrm{hc}})$ we obtain the numerator of the handover blocking probability.

3.2.3 Throughput

For $\mathbf{n} \in S_{\mathrm{hc}}$, let $R(\mathbf{n}) = n_j \mu_j p_{j0}$. Then, for $\mathbf{n} + \mathbf{e}_i \in S_{\mathrm{hc}}$:

$$R(\mathbf{n} + \mathbf{e}_i) - R(\mathbf{n}) = \mu_i p_{i0} 1(i = j),$$

so that R satisfies Condition 3.3. This leads to the throughput of cell j:

$$A_{\mathrm{hc},r} = \sum_{\mathbf{n} \in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) R(\mathbf{n}).$$

4 Applications

In this section, we will provide a separate example to illustrate the error due to

- the state space modification from S to S_{hc} (Sect. 4.1),
- the redial rate approximation (Sect. 4.2).

In Sect. 4.4 we provide a counterexample to indicate that the monotonicity result of Theorem 3.1 is not generally valid.

Section 4.1 considers the classical loss network for circuit switched communications systems. As the equilibrium distribution in this case is multivariate Poisson, the effect of the state space modification can be illustrated nicely. Section 4.2 considers a GSM network with fixed channel allocation. This is the key application which motivated our research.

4.1 Loss networks

This example considers the error due to the state space modification, where the process on the original state space S is approximated by the process on the hypercube state space S_{hc}. For a loss network the equilibrium distribution on both state spaces can, in principle, be evaluated in closed form, so that it provides a good test case for the accuracy of the state space modification. Furthermore, it is of interest to note that the easily computable Erlang loss probabilities bound for the hypercube process indeed bounds the blocking probabilities of the original process.

When handovers do not occur, i.e., $p_{ij} = 0$ for all i, j, the network is a loss network. The equilibrium distribution $\pi_0 = \pi_{\mathrm{loss}}$ is given in (1). Interesting performance measures are the blocking probability B_i, and the throughput $H_i = \lambda_i(1 - B_i)$. Although the blocking probability B_i is available in closed form, this form is not amenable for computation. Often, Monte Carlo summation is used to evaluate the sum [4, 20]. When the state space S is close to the hypercube state space S_{hc}, blocking probabilities can be rapidly evaluated using the convolution algorithm of [7].

The reward rate $R(\mathbf{n}) = \lambda_i 1(\mathbf{n} + \mathbf{e}_i \notin S)$ yields the blocking probability via $A_0 = \lambda_i B_i$. We have an explicit product form distribution on both S and S_{hc}. To this end, note that $\pi_{\mathrm{hc},0}(\mathbf{n}) = G_{\mathrm{hc}}^{-1} \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!}$, $\mathbf{n} \in S_{\mathrm{hc}}$, where $G_{\mathrm{hc}} = \prod_{i=1}^{N} [\sum_{j=0}^{N_i} \frac{\nu_i^j}{j!}]$ so that the normalising constant G_{hc} is readily evaluated. As a consequence,

$$A_0 = \sum_{\mathbf{n}\in S} R(\mathbf{n})\pi_0(\mathbf{n}) = \lambda_i \sum_{\mathbf{n}\in T_i} G^{-1} \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!} = \lambda_i B_i,$$

and

$$A_{\mathrm{hc},0} = \sum_{\mathbf{n}\in S} R(\mathbf{n})\pi_{\mathrm{hc},0}(\mathbf{n}) = \lambda_i \sum_{\mathbf{n}\in T_i\cup(S_{\mathrm{hc}}\setminus S)} G_{\mathrm{hc}}^{-1} \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!}.$$

Evaluation of $A_{\mathrm{hc},0}$ requires summation of $\pi_{\mathrm{hc},0}(\mathbf{n})$ over the set $T_i \cup (S_{\mathrm{hc}} \setminus S)$. When this set is small, i.e., when S does not deviate too much from a hypercube, evaluation of $A_{\mathrm{hc},0}$ is much faster than evaluation of A_0 that requires evaluation of the normalising constant G, which involves a summation of $\prod_k \frac{\nu_k^{n_k}}{n_k!}$. Below we also provide a readily computable bound on $A_{\mathrm{hc},0} - A_0$.

The error due to the state space modification is expressed by β as

$$\beta = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc}}(\mathbf{n})\Phi(\mathbf{n}) = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \sum_{j=1}^{N} \lambda_j 1(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}} \setminus S) G_{\mathrm{hc}}^{-1} \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!}.$$

Especially when some of the λ_j for $j \neq i$ are large, we have $A_{\mathrm{hc}} - \beta < 0$ so that the lower bound is not of practical value. An upper bound is of great practical interest. This can be obtained as follows.

Let $\mathbf{M} = (M_1, \ldots, M_N)$ be an upper corner of the hypercube that is completely contained in S, let

$$S_{\mathrm{hc}}^{M} = \{\mathbf{n} : 0 \le n_i \le M_i,\ i = 1, \ldots, N\} \subset S,$$

and let

$$\beta_M = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc}}(\mathbf{n}) \sum_{j=1}^{N} \lambda_j 1\big(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}} \setminus S_{\mathrm{hc}}^{M}\big).$$

As $\beta \le \beta_M$ and taking into account the explicit expression for the equilibrium distribution π_{hc}, we obtain

$$|A_{\mathrm{hc}} - A_0| \le \beta_M \le \left(\sum_{j=1}^{N} \lambda_j\right) \sum_{\ell=1}^{N} \sum_{n_\ell=M_\ell}^{N_\ell} \left[\frac{\nu_\ell^{n_\ell}}{n_\ell!} \Big/ \sum_{j=0}^{N_\ell} \frac{\nu_i^{j}}{j!}\right].$$

This result may be sharpened by carefully taking into account the state space summations involved in the definition of β_M. In addition, note that the selection of $\mathbf{M}$ need not be unique, which allows flexibility for minimisation of the upper bound. We have thus obtained an explicit upper bound on the error in the blocking probabilities due to state space modification.

4.2 Fixed channel allocation: a hypercube space process

In a GSM network operating under fixed channel allocation, each cell is assigned a fixed number of channels that can be used by calls in that cell only. As a consequence, the state space is a hypercube $S_{\mathrm{hc}} = \{\mathbf{n} : 0 \le n_i \le N_i\}$, where N_i is the number of channels assigned to cell i. Under maximal redial rates $r_{kj} = \nu_k \mu_k p_{kj}$

$$\begin{aligned}\beta_{r0} &= \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) \sum_{k,j=1}^{N} r_{kj} 1(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}}, \mathbf{n}+\mathbf{e}_k \notin S_{\mathrm{hc}}) \\ &= \sum_{k,j=1}^{N} \nu_k \mu_k p_{kj} B_{\mathrm{hc},r,k}(1 - B_{\mathrm{hc},r,j}),\end{aligned}$$

where we have used the fact that the state space is a hypercube. We thus obtain

$$\begin{aligned}&B_{\mathrm{hc},r,j} - \sum_{k,\ell=1}^{N} \frac{\nu_k \mu_k p_{k\ell}}{\lambda_j} B_{\mathrm{hc},r,k}(1 - B_{\mathrm{hc},r,\ell}) \\ &\quad \le B_{\mathrm{hc},0,j} \le B_{\mathrm{hc},r,j} \le B_{\mathrm{hc},0,j} + \sum_{k,\ell=1}^{N} \frac{\nu_k \mu_k p_{k\ell}}{\lambda_j} B_{\mathrm{hc},r,k}(1 - B_{\mathrm{hc},r,\ell}),\end{aligned}$$

where

$$B_{\mathrm{hc},r,j} = \frac{\nu_j^{N_j}}{N_j!}\left[\sum_{t=0}^{N_j} \frac{\nu_j^t}{t!}\right]^{-1},$$

the Erlang loss probability. From the expressions for blocking probabilities obtained in [4], for maximal redial rates $B_{\mathrm{hc},r,jk} = B_{\mathrm{hc},r,k}$.

The term $\sum_{k,\ell=1}^{N} \frac{\nu_k \mu_k p_{k\ell}}{\lambda_j}$ may be small, especially when $p_{k0} \approx 1$. This is in accordance with intuition, as in this regime handovers are rare, and redial rates are small, so that the redial rate approximation is likely to be accurate.

Notice that the lower bound may actually be below zero. In applications, the upper bound is often of more importance than the lower bound. Observe that

$$\begin{aligned}B_{\mathrm{hc},0,j} + \sum_{k,\ell=1}^{N} \frac{\nu_k \mu_k p_{k\ell}}{\lambda_j} B_{\mathrm{hc},r,k}(1 - B_{\mathrm{hc},r,\ell}) &\le B_{\mathrm{hc},0,j} + \sum_{k,\ell=1}^{N} \frac{\nu_k \mu_k p_{k\ell}}{\lambda_j} B_{\mathrm{hc},r,k} \\ &= B_{\mathrm{hc},0,j} + \sum_{k=1}^{N} \frac{\nu_k \mu_k (1 - p_{k0})}{\lambda_j} B_{\mathrm{hc},r,k}.\end{aligned}$$

When the upper bound $B_{\mathrm{hc},r,j} < 1\%$, the error in the blocking probability of the actual fresh call blocking probabilities $B_{\mathrm{hc},0,j}$ is of that order of magnitude, too. Thus, it is sufficient to dimension the system with maximal redial rates to guarantee a Quality of Service limit of 1% of the blocking probabilities, in which case the actual blocking probabilities will be in the range 0.5%–1%.

4.3 General result including routing

The approach for a loss network without routing as in Sect. 4.1 can readily be extended to networks with routing. Note that in this case the equilibrium distribution of the original chain is not known. However, the bounds are expressed in terms of the equilibrium distribution of the hypercube process with redial rates. Under maximal redial rates the resulting truncated Poisson equilibrium distribution is explicitly known and amenable for computation since its normalising constant is known in closed form.

The bound consists of two parts: β and β_{r0}. Under maximal redial rates:

$$\begin{aligned}
\beta + \beta_{r0} = \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) \Bigg\{ & \sum_{j=1}^{N} \lambda_j 1\big(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}} \backslash S_{\mathrm{hc}}^{M}\big) \\
&+ \sum_{i,j=1}^{N} n_i \mu_i p_{ij} 1\big(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j \in S_{\mathrm{hc}} \backslash S_{\mathrm{hc}}^{M}\big) \\
&+ \sum_{k,j=1}^{N} r_{kj} 1(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}}, \mathbf{n}+\mathbf{e}_k \notin S_{\mathrm{hc}}) \Bigg\} \\
= \sum_{\mathbf{n}\in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) \Bigg\{ & \sum_{j=1}^{N} \lambda_j 1\big(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}} \backslash S_{\mathrm{hc}}^{M}\big) \\
&+ \sum_{i,j=1}^{N} \nu_i \mu_i p_{ij} 1\big(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}} \backslash S_{\mathrm{hc}}^{M}\big) \\
&+ \sum_{i,j=1}^{N} (\nu_i \mu_i p_{ij}) 1(\mathbf{n}+\mathbf{e}_j \in S_{\mathrm{hc}}, \mathbf{n}+\mathbf{e}_i \notin S_{\mathrm{hc}}) \Bigg\}.
\end{aligned}$$

Following the steps as in Sect. 4.1, we readily obtain

$$\begin{aligned}
\beta_M \le & \sum_{\ell=1}^{N} \sum_{n_\ell=M_\ell}^{N_\ell} \left[\frac{\nu_\ell^{n_\ell}}{n_\ell!} \Big/ \sum_{j=0}^{N_\ell} \frac{\nu_i^{j}}{j!} \right] \left\{ \sum_{j=1}^{N} \lambda_j + \sum_{i=1}^{N} \nu_i \mu_i (1-p_{i0}) \right\} \\
& + \sum_{i=1}^{N} \sum_{i,j=1,\ j\neq i}^{N} (\nu_i \mu_i p_{ij}) \frac{\nu_i^{N_i}}{N_i!} \Big/ \sum_{k=0}^{N_k} \frac{\nu_i^{k}}{k!}.
\end{aligned}$$

Remark 4.1 (Complete sharing) Under complete sharing of capacity, all cells share the common capacity s. The state space is

$$S_s = \left\{ \mathbf{n} : \sum_{i=1}^{N} n_i \le s \right\},$$

and handovers cannot be blocked. The PASTA property implies that

$$B_j = B_{r,j} = \frac{(\sum_{j=1}^{N} \nu_j)^s}{s!} \left[\sum_{t=0}^{s} \frac{(\sum_{j=1}^{N} \nu_j)^t}{t!} \right]^{-1},$$

and

$$B_{ij} = B_{r,ij}.$$

When the state space S is close to that of complete sharing, we may use S_s instead of S_{hc}^M to approximate the error bound.

4.4 Counterexample

This section provides an example to illustrate that the introduction of redial rates does not necessarily increase fresh call blocking probabilities at all cells. Consider a network of 5 cells, cell $1, \ldots, 5$, with common capacity constraints

$$n_1 + n_2 \leq 1, \qquad n_2 + n_3 \leq 1, \qquad n_3 + n_4 \leq 1, \qquad n_4 + n_5 \leq 1.$$

Handovers are allowed only from cell 2 to cell 3, say with probability p. The traffic equations (5) have the unique solution

$$\nu_i = \lambda_i / \mu_i, \quad i = 1, 2, 4, 5, \qquad \nu_3 = (\lambda_3 + \lambda_2 p)/\mu_3.$$

Fresh call blocking probabilities for the process without redial rates, and with maximal redial rates, then become:

$$B = \left(\frac{22\,531\,289}{129\,964\,237} \frac{17\,307\,792}{129\,964\,237} \frac{25\,390\,649}{129\,964\,237} \frac{17\,428\,912}{129\,964\,237} \frac{22\,507\,065}{129\,964\,237} \right),$$

$$B_r = \left(\frac{21}{121} \frac{16}{121} \frac{25}{121} \frac{16}{121} \frac{21}{121} \right),$$

and

$$B_i < B_{r,i}, \quad i = 1, 3, 5, \qquad B_i > B_{r,i}, \quad i = 2, 4.$$

This illustrates that for a general state space there is a knock-on effect due to the redial rates: extra calls in one cell may decrease the load in neighbouring cells, resulting in lower blocking probabilities in cells sharing a capacity constraint with that neighbouring cell.

5 Proof of the main results

This section provides the proofs of our main results and some related arguments. Some of the results are duplicated to enhance the readability of the section. Section 5.1 first establishes preliminary results on Markov reward structures and uniformisation. Next, Sect. 5.2 develops the monotonicity results, and Sect. 5.3 proves the error bound result.

5.1 Preliminaries

We will compare performance measures for the system under different conditions by means of expected rewards. To this end, let a reward $R(\mathbf{n})$ be incurred per unit time whenever the system is in state $\mathbf{n}$, and define

$$A = \sum_{\mathbf{n}\in S} \pi(\mathbf{n}) R(\mathbf{n}) = \lim_{t\to\infty} \frac{1}{t} \mathbb{E} \int_0^t R\bigl(X(u)\bigr)\, du,$$

with $\pi(\mathbf{n})$ the equilibrium distribution of the Markov chain $X(t)$. First, in order to use inductive arguments, we transfer the continuous-time setting to a discrete-time formulation by means of uniformisation. To this end, let Λ be some arbitrarily large number such that

$$\Lambda \geq \sum_{j=1}^{N} \lambda_j + \sum_{j=1}^{N}\sum_{k=0}^{N} N_j \mu_j p_{jk} + \sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj} = \sum_{j=1}^{N} \lambda_j + \sum_{j=1}^{N} N_j \mu_j + \sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}.$$

The continuous-time Markov chain $\mathbf{X}$ can then be studied via the discrete-time Markov chain with one-step transition probabilities (uniformisation), see e.g. [21, p. 110]:

$$P(\mathbf{n}, \mathbf{n}') = \begin{cases} q(\mathbf{n}, \mathbf{n}')/\Lambda, & \text{if } \mathbf{n}' \neq \mathbf{n}, \\ 1 - \sum_{n''\neq n} q(\mathbf{n}, \mathbf{n}'')/\Lambda, & \text{if } \mathbf{n}' = \mathbf{n}. \end{cases}$$

Furthermore, let the functions $V^k(\mathbf{n})$ represent the expected cumulative reward over k steps when starting in state $\mathbf{n}$ at time 0 and incurring a reward $R(\mathbf{n})/\Lambda$ per step for the corresponding discrete-time Markov chain, i.e.,

$$V^K(\mathbf{n}) = \frac{1}{\Lambda} \sum_{k=0}^{K-1} \sum_{\mathbf{n}'\in S} P^k(\mathbf{n}, \mathbf{n}') R(\mathbf{n}'), \quad \mathbf{n} \in S,\ K = 0, 1, 2, \ldots, \qquad V^0(\mathbf{n}) = 0,$$

where, by convention, $P^0(\mathbf{n}, \mathbf{n}') = 1(\mathbf{n} = \mathbf{n}')$. These functions can recursively be determined as

$$V^{K+1}(\mathbf{n}) = \frac{R(\mathbf{n})}{\Lambda} + \sum_{\mathbf{n}'\in S} P(\mathbf{n}, \mathbf{n}') V^K(\mathbf{n}'), \quad \mathbf{n} \in S,\ K = 0, 1, 2, \ldots, \qquad V^0(\mathbf{n}) = 0,$$

and by virtue of the uniformisation:

$$A = \lim_{K\to\infty} \frac{\Lambda}{K} V^K(\mathbf{n}).$$

Similarly, with the same uniformisation parameter Λ, for the modified processes with redial rates r_{kj} and the state space transformed to the hypercube S_{hc}, we can determine A_r and $A_{\text{hc},r}$ by defining the one-step matrices P_r and $P_{\text{hc},r}$ and cumulative rewards V_r^k and $V_{\text{hc},r}^k$ with q_r and $q_{\text{hc},r}$ replacing q.

5.2 Monotonicity

This section provides proofs for a variety of monotonicity results. These monotonicity results have a twofold function. First, Theorems 5.3, 5.4, and 5.7 will be essential for the proof of the error bound Theorem 3.4 as will appear in Sect. 5.3. Second, these theorems will also lead to upper bounds of practical interest by themselves. In particular, the main monotonicity result (Theorem 5.7) states that rewards for the hypercube process with arbitrary redial rates exceed those of the original process.

First, we show that rewards for the hypercube process are monotone and increasing in the number of steps of the Markov chain. Next, it is shown that the cumulative expected rewards for the hypercube process exceed those for the original process. Our main monotonicity result states that rewards for the hypercube process with arbitrary redial rates exceed those of the original process. In particular, this result allows us to select maximal redial rates under which the equilibrium distribution is truncated multivariate Poisson. The proof of this result consists of a number of steps. This section provides these steps as well as additional comments on the results.

Consider the set of functions defined as

$$C_{\mathrm{hc}} = \big\{ f : S_{\mathrm{hc}} \to [0,\infty) \,|\, f(\mathbf{n}+\mathbf{e}_i) - f(\mathbf{n}) \geq 0,\ \text{for } \mathbf{n}, \mathbf{n}+\mathbf{e}_i \in S_{\mathrm{hc}} \big\}.$$

Lemma 5.1 *C_{hc} is closed under $P_{\mathrm{hc},r}$, that is $(P_{\mathrm{hc},r} f) \in C_{\mathrm{hc}}$ for all $f \in C_{\mathrm{hc}}$.*

Proof It is sufficient to show that $(P_{\mathrm{hc},r} f)(\mathbf{n}+\mathbf{e}_i) - (P_{\mathrm{hc},r} f)(\mathbf{n}) \geq 0$ for $\mathbf{n}, \mathbf{n}+\mathbf{e}_i \in S_{\mathrm{hc}}$ for all $f \in C_{\mathrm{hc}}$. We will first establish results for the process on arbitrary state space S, and only when required in the derivation restrict ourselves to S_{hc}. For notational convenience, we omit the subscript in the transitions rates. Straightforward calculations yield, for $\mathbf{n},\ \mathbf{n}+\mathbf{e}_i \in S$,

$$\begin{aligned}
&\Lambda\big[(P_{\mathrm{hc},r} f)(\mathbf{n}+\mathbf{e}_i) - (P_{\mathrm{hc},r} f)(\mathbf{n})\big] \\
&\quad = \sum_{\mathbf{n}' \in S} q(\mathbf{n}+\mathbf{e}_i, \mathbf{n}') f(\mathbf{n}') + \Lambda f(\mathbf{n}+\mathbf{e}_i) - \sum_{\mathbf{n}' \in S} q(\mathbf{n}+\mathbf{e}_i, \mathbf{n}') f(\mathbf{n}+\mathbf{e}_i) \\
&\qquad - \sum_{\mathbf{n}' \in S} q(\mathbf{n}, \mathbf{n}') f(\mathbf{n}') - \Lambda f(\mathbf{n}) + \sum_{\mathbf{n}' \in S} q(\mathbf{n}, \mathbf{n}') f(\mathbf{n}) \\
&\quad = \sum_{j=1}^{N} \lambda_j f(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j) 1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \in S) - \sum_{j=1}^{N} \lambda_j f(\mathbf{n}+\mathbf{e}_j) 1(\mathbf{n}+\mathbf{e}_j \in S) \\
&\qquad + \sum_{j=1}^{N} \lambda_j f(\mathbf{n}+\mathbf{e}_i) 1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \notin S) - \sum_{j=1}^{N} \lambda_j f(\mathbf{n}) 1(\mathbf{n}+\mathbf{e}_j \notin S) \\
&\qquad + \sum_{j=1}^{N} \sum_{k=0}^{N} (n_j + \delta_{ij}) \mu_j p_{jk} f(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k) 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k \in S) \\
&\qquad - \sum_{j=1}^{N} \sum_{k=0}^{N} n_j \mu_j p_{jk} f(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k) 1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \in S)
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=1}^{N}\sum_{k=1}^{N}(n_j+\delta_{ij})\mu_j p_{jk} f(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\notin S)\\
&-\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk} f(\mathbf{n}-\mathbf{e}_j)1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\notin S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj} f(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j)1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)\\
&-\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj} f(\mathbf{n}+\mathbf{e}_j)1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S)\\
&+\Lambda\big[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n})\big]\\
&-\sum_{j=1}^{N}\lambda_j f(\mathbf{n}+\mathbf{e}_i)+\sum_{j=1}^{N}\lambda_j f(\mathbf{n})\\
&-\sum_{j=1}^{N}\sum_{k=0}^{N}(n_j+\delta_{ij})\mu_j p_{jk} f(\mathbf{n}+\mathbf{e}_i)+\sum_{j=1}^{N}\sum_{k=0}^{N} n_j\mu_j p_{jk} f(\mathbf{n})\\
&-\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj} f(\mathbf{n}+\mathbf{e}_i)1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj} f(\mathbf{n})1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S),
\end{aligned}
$$

so that

$$
\begin{aligned}
&\Lambda\big[(P_{\mathrm{hc},r}f)(\mathbf{n}+\mathbf{e}_i)-(P_{\mathrm{hc},r}f)(\mathbf{n})\big]\\
&\quad=\sum_{j=1}^{N}\lambda_j\big[f(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j)-f(\mathbf{n}+\mathbf{e}_j)\big]1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S)\\
&\qquad+\sum_{j=1}^{N}\lambda_j\big[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n})\big]1(\mathbf{n}+\mathbf{e}_j\notin S)\\
&\qquad+\sum_{j=1}^{N}\sum_{k=0}^{N} n_j\mu_j p_{jk}\big[f(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k)-f(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\big]\\
&\qquad\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\in S)\\
&\qquad+\sum_{k=0}^{N}\mu_i p_{ik}\big[f(\mathbf{n}+\mathbf{e}_k)-f(\mathbf{n})\big]1(\mathbf{n}+\mathbf{e}_k\in S)
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}\big[f(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-f(\mathbf{n}-\mathbf{e}_j)\big]1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\notin S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[f(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j)-f(\mathbf{n}+\mathbf{e}_j)\big]1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)\\
&+\Bigg\{\Lambda_r-\sum_{j=1}^{N}\lambda_j-\sum_{j=1}^{N}\sum_{k=0}^{N}(n_j+\delta_{ij})\mu_j p_{jk}\\
&-\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S)\Bigg\}\big[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n})\big]\\
&+\sum_{j=1}^{N}\lambda_j\big[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n}+\mathbf{e}_j)\big]1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\notin S,\ \mathbf{n}+\mathbf{e}_j\in S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}\big[f(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-f(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\big]\\
&\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\in S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[f(\mathbf{n}+\mathbf{e}_j)-f(\mathbf{n}+\mathbf{e}_i)\big]\\
&\times\big[1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)-1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S)\big].
\end{aligned}
$$

Now restrict attention to the hypercube process $X_{\mathrm{hc},r}$ with state space S_{hc}, and transition probabilities $P_{\mathrm{hc},r}$. For this process, all terms except the last three are positive due to the definition of Λ and the assumption that $f\in C_{\mathrm{hc}}$. On the hypercube state space, the last three terms are zero since for $\mathbf{n}+\mathbf{e}_i\in S_{\mathrm{hc}}$ it must be that $\mathbf{n}+\mathbf{e}_j\in S_{\mathrm{hc}}$ implies that $\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S_{\mathrm{hc}}$ unless $i=j$. However, for $i=j$ we have $[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n}+\mathbf{e}_j)]=0$. A similar argument applies to the other terms. $\square$

Remark 5.2 (C_{hc} closed under P?) The hypercube state space is essential for the proof of Lemma 5.1. In particular, in addition to the assumption that $f\in C_{\mathrm{hc}}$, for the proof to be completed the following terms must cancel:

$$
\begin{aligned}
&\sum_{j=1}^{N}\lambda_j\big[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n}+\mathbf{e}_j)\big]1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\notin S,\ \mathbf{n}+\mathbf{e}_j\in S)\\
&\quad+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}\big[f(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-f(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\big]\\
&\quad\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\in S)
\end{aligned}
$$

$$+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[f(\mathbf{n}+\mathbf{e}_j)-f(\mathbf{n}+\mathbf{e}_i)\big]$$

$$\times\big[1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)-1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S)\big].$$

To this end, recall from the proof that, on the hypercube state space, these terms are zero since for $\mathbf{n}+\mathbf{e}_i\in S$ it must be that $\mathbf{n}+\mathbf{e}_j\in S$ implies that also $\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S$ unless $i=j$. However, for $i=j$ we have $[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n}+\mathbf{e}_j)]=0$. Similarly, in the second term the indicator is non-zero only when $k=i$, but then the term in square brackets cancels. For non hypercube state spaces the contribution of $[f(\mathbf{n}+\mathbf{e}_i)-f(\mathbf{n}+\mathbf{e}_j)]$ may be arbitrary, and, in general, C_{hc} is *not* closed under P_r.

Theorem 5.3 *For any $f\in C_{\mathrm{hc}}$ and $k\geq 0$ we have with $\mathbf{0}=(0,\dots,0)$*

$$P_{\mathrm{hc},r}^{k}f(\mathbf{0})\leq P_{\mathrm{hc},r}^{k+1}f(\mathbf{0})\leq\sum_{\mathbf{n}\in S_{\mathrm{hc}}}\pi_{\mathrm{hc},r}(\mathbf{n})f(\mathbf{n}).$$

Proof We will prove the first inequality by induction in k. For $k=0$ it applies since

$$\Lambda P_{\mathrm{hc},r}f(\mathbf{0})=\Lambda f(\mathbf{0})+\sum_{j=1}^{N}\lambda_j\big(f(\mathbf{0}+\mathbf{e}_j)-f(\mathbf{0})\big)\geq\Lambda f(\mathbf{0}),$$

where we have used that $\mathbf{e}_j\in S_{\mathrm{hc}}$ for all j. Suppose that the inequality holds for $k\leq t$. Then it also holds for $k=t+1$, since

$$P_{\mathrm{hc},r}^{t+1}f(\mathbf{0})-P_{\mathrm{hc},r}^{t+2}f(\mathbf{0})=P_{\mathrm{hc},r}^{t}(P_{\mathrm{hc},r}f)(\mathbf{0})-P_{\mathrm{hc},r}^{t+1}(P_{\mathrm{hc},r}f)(\mathbf{0})\leq 0,$$

where the last inequality is obtained since $P_{\mathrm{hc},r}f\in C_{\mathrm{hc}}$ by Lemma 5.1.

The second inequality is a direct consequence of the first inequality and irreducibility of the Markov chain which implies that $\lim_{k\to\infty}P_{\mathrm{hc},r}^{k}f(\mathbf{0})=\sum_{\mathbf{n}\in S_{\mathrm{hc}}}\pi_{\mathrm{hc},r}(\mathbf{n})f(\mathbf{n})$. □

Monotonicity between the original Markov chain and the hypercube process can only be obtained for redial rates equal to zero. As we will see in Lemma 5.6, the hypercube process is monotone in the redial rates. We are now ready to state a main monotonicity result which will be used in the proof of Theorem 3.4.

Theorem 5.4 *For any $f\in C_{\mathrm{hc}}$ and $k\geq 0$ we have*

$$P_0{}^{k}f(\mathbf{0})\leq P_{\mathrm{hc},0}^{k}f(\mathbf{0}).\tag{14}$$

Moreover,

$$\sum_{\mathbf{n}\in S}\pi_0(\mathbf{n})f(\mathbf{n})\leq\sum_{\mathbf{n}\in S_{\mathrm{hc}}}\pi_{\mathrm{hc},0}(\mathbf{n})f(\mathbf{n}).\tag{15}$$

Proof For notational convenience, we introduce the Markov chain $\bar{X}_r$ as the extension of X_r to state space S_{hc}, that has transition rates $\bar{Q}_r = (\bar{q}_r(\mathbf{n}, \mathbf{n}'), \mathbf{n}, \mathbf{n}' \in S_{\text{hc}})$ for $\mathbf{n}' \neq \mathbf{n}$ defined as

$$\bar{q}_r(\mathbf{n}, \mathbf{n}') = \begin{cases} q_r(\mathbf{n}, \mathbf{n}'), & \text{if } \mathbf{n}, \mathbf{n}' \in S, \\ q_{\text{hc},r}(\mathbf{n}, \mathbf{n}'), & \text{if } \mathbf{n} \in S_{\text{hc}} \backslash S, \mathbf{n}' \in S_{\text{hc}}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that $q_r(\mathbf{n}, \mathbf{n}') = q_{\text{hc},r}(\mathbf{n}, \mathbf{n}')$ if $\mathbf{n}, \mathbf{n}' \in S_{\text{hc}} \backslash \{\{\bigcup_i T_i\} \cup \{\bigcup_{i,j} T_{ij}\}\}$, and that the states $S_{\text{hc}} \backslash S$ are transient states for $\bar{X}_r$. The chain $\bar{X}_r$ is uniformisable with transition matrix

$$\bar{P}_r(\mathbf{n}, \mathbf{n}') = \begin{cases} q_r(\mathbf{n}, \mathbf{n}')/\Lambda, & \text{if } \mathbf{n}' \neq \mathbf{n}, n, \mathbf{n}' \in S, \\ q_{\text{hc},r}(\mathbf{n}, \mathbf{n}')/\Lambda, & \text{if } \mathbf{n} \in S_{\text{hc}} \backslash S, \mathbf{n}' \in S_{\text{hc}}, \\ 1 - \sum_{n'' \neq n} q_r(\mathbf{n}, \mathbf{n}'')/\Lambda, & \text{if } \mathbf{n}' = \mathbf{n} \in S, \\ 1 - \sum_{n'' \neq n} q_{\text{hc},r}(\mathbf{n}, \mathbf{n}'')/\Lambda, & \text{if } \mathbf{n}' = \mathbf{n} \in S_{\text{hc}} \backslash S. \end{cases}$$

Note that for the process starting at S, e.g. starting empty (in state $\mathbf{0} = (0, \dots, 0)$, the evolution of the process $\bar{X}_r$ coincides with that of X_r, so that

$$\bar{P}_r^k f(\mathbf{0}) = P_r^k f(\mathbf{0}).$$

The entries of $\bar{P}_0$ and $P_{\text{hc},0}$ differ only at the boundary of S. We readily find that, for $f \in C_{\text{hc}}$, and $\mathbf{n} \in S_{\text{hc}}$

$$\begin{aligned} &(P_{\text{hc},0} - \bar{P}_0) f(\mathbf{n}) \\ &\quad = \sum_{j=1}^{N} \lambda_j 1(\mathbf{n} + \mathbf{e}_j \in S_{\text{hc}} \backslash S)\big(f(\mathbf{n} + \mathbf{e}_j) - f(\mathbf{n})\big) \\ &\qquad + \sum_{i=1}^{N} \sum_{j=1}^{N} n_i \mu_i p_{ij} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j \in S_{\text{hc}} \backslash S)\big(f(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j) - f(\mathbf{n} - \mathbf{e}_i)\big) \geq 0. \end{aligned} \tag{16}$$

Observe that

$$\begin{aligned} &\big(P_{\text{hc},0}^k - \bar{P}_0^k\big) f(\mathbf{0}) \\ &\quad = \bar{P}_0\big[\big(P_{\text{hc},0}^{k-1} - \bar{P}_0^{k-1}\big) f\big](\mathbf{0}) + (P_{\text{hc},0} - \bar{P}_0)\big(P_{\text{hc},0}^{k-1} f\big)(\mathbf{0}) \\ &\quad = \cdots = \bar{P}_0^k\big(P_{\text{hc},0}^0 f - \bar{P}_0^0 f\big)(\mathbf{0}) + \sum_{t=0}^{k-1} \bar{P}_0^t (P_{\text{hc},0} - \bar{P}_0)\big(P_{\text{hc},0}^{k-t-1} f\big)(\mathbf{0}). \end{aligned}$$

Note that $P_{\text{hc},0}^0 f = \bar{P}_0^0 f = f$ by definition. By Lemma 5.1, observe that $P_{\text{hc},0}^{k-t-1} f \in C_{\text{hc}}$ for $f \in C_{\text{hc}}$, so that by (16) $(P_{\text{hc},0} - \bar{P}_0)(P_{\text{hc},0}^{k-t-1} f)(\mathbf{0}) \geq 0$ for all $t =$

$0, \ldots, k-1$. Furthermore, since $\bar{P}_0$ is a stochastic matrix, we can use the fact that $\bar{P}_0^t g \geq 0$ if $g \geq 0$ componentwise. The proof of (14) is hereby completed. From Theorem 5.3 we obtain from (14) for $r=0$: ${P_0}^k f(\mathbf{0}) \leq \sum_{\mathbf{n}\in S_{\text{hc}}} \pi_{\text{hc},0}(\mathbf{n}) f(\mathbf{n})$ for all k. Equation (15) now follows noting that S is an irreducible class for X, so that for all $\mathbf{m} \in S$

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} P_0^k f(\mathbf{m}) = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} P_0^k f(\mathbf{0}) = \sum_{\mathbf{n}\in S} \pi_0(\mathbf{n}) f(\mathbf{n}). \qquad \square \tag{17}$$

Remark 5.5 (General redial rates) The assumption of null redial rates is used in (16). For non-null redial rates an additional negative term involving the redial rates at the boundary of S would appear.

Now we will show that $P_{\text{hc},r}^k f(\mathbf{0})$ for $f \in C_{\text{hc}}$ is strictly increasing in the redial rates, which implies that the rewards (blocking probabilities) are increasing in the redial rates. This result will enable us to provide a computable bound on the blocking probabilities for the original process (without redial rates). The main step is the following lemma.

Lemma 5.6 *Consider the processes $X_{\text{hc},r}$ and $X_{\text{hc},r'}$ on state space S_{hc} with $r_{ji} \geq r'_{ji}$ for all j, i. For $f \in C_{\text{hc}}$*

$$P_{\text{hc},r}^k f(\mathbf{0}) \geq P_{\text{hc},r'}^k f(\mathbf{0}),$$

and

$$\sum_{\mathbf{n}\in S_{\text{hc}}} \pi_{\text{hc},r}(\mathbf{n}) f(\mathbf{n}) \geq \sum_{\mathbf{n}\in S_{\text{hc}}} \pi_{\text{hc},r'}(\mathbf{n}) f(\mathbf{n}).$$

Proof Note that

$$\begin{aligned}&(P_{\text{hc},r} - P_{\text{hc},r'}) f(\mathbf{n})\\ &\quad = \sum_{k,i} (r_{ki} - r'_{ki})\big(f(\mathbf{n}+\mathbf{e}_i) - f(\mathbf{n})\big) 1(\mathbf{n}+\mathbf{e}_k \notin S_{\text{hc}},\ \mathbf{n}+\mathbf{e}_i \in S) \geq 0.\end{aligned}$$

Furthermore, C_{hc} is closed under $P_{\text{hc},r}$. The remainder of the proof can be shown along the lines of that of Theorem 5.4. $\square$

Our main monotonicity result now follows directly as a consequence of Theorem 5.4, and Lemma 5.6 for $r' = 0$.

Theorem 5.7 (Main monotonicity result) *For any $f \in C_{\text{hc}}$, $r_{ji} \geq 0$ for all j, i, and $k \geq 0$*

$$P_0^k f(\mathbf{0}) \leq P_{\text{hc},r}^k f(\mathbf{0}).$$

Moreover,

$$\sum_{\mathbf{n}\in S} \pi_0(\mathbf{n}) f(\mathbf{n}) \leq \sum_{\mathbf{n}\in S_{\text{hc}}} \pi_{\text{hc},r}(\mathbf{n}) f(\mathbf{n}).$$

Remark 5.8 (Bound by maximal redial rates) Under the conditions of Theorem 5.7, i.e., for $r_{ji} = \nu_j \mu_j p_{ji}$, $j, i = 1, \ldots, N$, an upper bound can readily be computed by

$$\pi_{\mathrm{hc},r}(\mathbf{n}) = \prod_{k=1}^{N} \left(\frac{\nu_k^{n_k}}{n_k!} \Big/ \sum_{j=1}^{N_k} \frac{\nu_j^{n_j}}{n_j!} \right).$$

Remark 5.9 (Other product form modifications) Various modifications resulting in a product form or truncated multivariate Poisson equilibrium distribution have been introduced in the literature. For these modifications, the result of Lemma 5.1 that is crucial for our main monotonicity result Theorem 5.7 cannot be obtained, since the transition rates in the modification do not lead to higher states (transitions from $\mathbf{n}$ to $\mathbf{n} + \mathbf{e}_i$ for some i).

Remark 5.10 A sample path proof for Lemma 5.6 is provided in [1] for fresh call blocking probabilities. In the present paper we have provided a direct proof for general $f \in C_{\mathrm{hc}}$.

5.3 Error bounds

We are now also able to establish error bounds on performance measures such as the fresh call blocking probabilities and throughputs by studying cumulative reward structures of the Markov reward chains. The following lemma establishes a lower and upper bound for the different terms of the cumulative rewards for the system with redial rates r_{ij}. To make our result and the role of the state space more explicit, we formulate the results for a general state space. As a corollary we provide the result for the hypercube state space.

Lemma 5.11 *Consider the process X_r with state space S, transition rates q_r and reward rate R. A sufficient condition for*

$$0 \le \left[V_r^{K+1}(\mathbf{n} + \mathbf{e}_i) - V_r^{K+1}(\mathbf{n})\right] \le 1, \quad \mathbf{n}, \mathbf{n} + \mathbf{e}_i \in S,$$

is that

$$0 \le \left[V_r^{K}(\mathbf{n} + \mathbf{e}_i) - V_r^{K}(\mathbf{n})\right] \le 1, \quad \mathbf{n}, \mathbf{n} + \mathbf{e}_i \in S,$$

and

$$\begin{aligned}
0 \le{}& R(\mathbf{n} + \mathbf{e}_i) - R(\mathbf{n}) \\
&+ \sum_{j=1}^{N} \lambda_j \big(V_r^K(\mathbf{n} + \mathbf{e}_i) - V_r^K(\mathbf{n} + \mathbf{e}_j)\big) 1(\mathbf{n} + \mathbf{e}_j \in S,\ \mathbf{n} + \mathbf{e}_i + \mathbf{e}_j \notin S) \\
&+ \sum_{j=1}^{N} \sum_{k=1}^{N} n_j \mu_j p_{jk} \big(V_r^K(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j) - V_r^K(\mathbf{n} - \mathbf{e}_j + \mathbf{e}_k)\big) \\
&\times 1(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j + \mathbf{e}_k \notin S,\ \mathbf{n} - \mathbf{e}_j + \mathbf{e}_k \in S)
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big(V_r^K(\mathbf{n}+\mathbf{e}_i)-V_r^K(\mathbf{n}+\mathbf{e}_j)\big)\\
&\times\big[1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S)-1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)\big]\\
\le&\sum_{j=1}^{N}\lambda_j 1(\mathbf{n}+\mathbf{e}_j\in S,\ \mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\notin S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\in S)\\
&+\mu_i p_{i0}+\sum_{k=1}^{N}\mu_i p_{ik}1(\mathbf{n}+\mathbf{e}_k\notin S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[1(\mathbf{n}+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_k\notin S)-1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S,\mathbf{n}+\mathbf{e}_i+\mathbf{e}_k\notin S)\big].
\end{aligned}
\tag{18}
$$

Proof For $K+1$, a derivation similar to that in the proof of Lemma 5.1 yields, for $\mathbf{n}$, $\mathbf{n}+\mathbf{e}_i\in S$,

$$
\begin{aligned}
&\Lambda\big[V^{K+1}(\mathbf{n}+\mathbf{e}_i)-V^{K+1}(\mathbf{n})\big]\\
&\quad=R(\mathbf{n}+\mathbf{e}_i)-R(\mathbf{n})\\
&\qquad+\sum_{j=1}^{N}\lambda_j\big(V^K(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_j)\big)1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\in S)\\
&\qquad+\sum_{j=1}^{N}\lambda_j\big(V^K(\mathbf{n}+\mathbf{e}_i)-V^K(\mathbf{n})\big)1(\mathbf{n}+\mathbf{e}_j\notin S)\\
&\qquad-\sum_{j=1}^{N}\lambda_j\big(V^K(\mathbf{n}+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_i)\big)1(\mathbf{n}+\mathbf{e}_j\in S,\ \mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\notin S)\\
&\qquad+\sum_{j=1}^{N}\sum_{k=0}^{N} n_j\mu_j p_{jk}\big(V^K(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k)-V^K(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\big)\\
&\qquad\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\in S)\\
&\qquad+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}\big(V^K(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-V^K(\mathbf{n}-\mathbf{e}_j)\big)1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\notin S)\\
&\qquad+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}\big(V^K(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-V^K(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\big)
\end{aligned}
$$

$$
\begin{aligned}
&\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k \notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \in S)\\
&+\sum_{k=0}^{N}\mu_i p_{ik}\big(V^K(\mathbf{n}+\mathbf{e}_k)-V^K(\mathbf{n})\big)1(\mathbf{n}+\mathbf{e}_k \in S)\\
&+\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[V^K(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_j)\big]\\
&\times 1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \in S, \mathbf{n}+\mathbf{e}_i+\mathbf{e}_k \notin S)\\
&-\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[V^K(\mathbf{n}+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_i)\big]\\
&\times\big[1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \in S, \mathbf{n}+\mathbf{e}_i+\mathbf{e}_k \notin S)-1(\mathbf{n}+\mathbf{e}_j \in S, \mathbf{n}+\mathbf{e}_k \notin S)\big]\\
&+\Bigg(\Lambda-\sum_{j=1}^{N}\lambda_j-\sum_{j=1}^{N}\sum_{k=0}^{N}(n_j+\delta_{ij})\mu_j p_{jk}\\
&-\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}1(\mathbf{n}+\mathbf{e}_j \in S, \mathbf{n}+\mathbf{e}_k \notin S)\Bigg)\big[V^K(\mathbf{n}+\mathbf{e}_i)-V^K(\mathbf{n})\big].
\end{aligned}
$$

First consider the lower bound. Observe that

$$
1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \in S)+1(\mathbf{n}+\mathbf{e}_j \notin S)+1(\mathbf{n}+\mathbf{e}_j \in S,\ \mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \notin S)=1,
$$

and a similar relation holds for the handover and redial terms. As $0 \le [V^K(\mathbf{n}+\mathbf{e}_i)-V^K(\mathbf{n})]$, $\mathbf{n}, \mathbf{n}+\mathbf{e}_i \in S$, all terms are guaranteed positive (use definition of Λ), except the three terms

$$
\begin{aligned}
&R(\mathbf{n}+\mathbf{e}_i)-R(\mathbf{n})\\
&\quad-\sum_{j=1}^{N}\lambda_j\big(V^K(\mathbf{n}+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_i)\big)1(\mathbf{n}+\mathbf{e}_j \in S,\ \mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \notin S)\\
&\quad+\sum_{j=1}^{N}\sum_{k=1}^{N} n_j\mu_j p_{jk}\big(V^K(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-V^K(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\big)\\
&\quad\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k \notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \in S)\\
&\quad-\sum_{j=1}^{N}\sum_{k=1}^{N} r_{kj}\big[V^K(\mathbf{n}+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_i)\big]\\
&\quad\times\big[1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \in S, \mathbf{n}+\mathbf{e}_i+\mathbf{e}_k \notin S)-1(\mathbf{n}+\mathbf{e}_j \in S, \mathbf{n}+\mathbf{e}_k \notin S)\big],
\end{aligned}
$$

but this expression is non-negative by the assumption of the lemma.

Now consider the upper bound. In the expression

$$\sum_{k=0}^{N} \mu_i p_{ik}\big(V^K(\mathbf{n}+\mathbf{e}_k) - V^K(\mathbf{n})\big)1(\mathbf{n}+\mathbf{e}_k \in S)$$

the $k=0$ term cancels. This absorbs the term $\mu_i p_{i0}$ in the bound on $R(\mathbf{n}+\mathbf{e}_i) - R(\mathbf{n})$. As $[V^K(\mathbf{n}+\mathbf{e}_i) - V^K(\mathbf{n})] \le 1$, $\mathbf{n}, \mathbf{n}+\mathbf{e}_i \in S$, we obtain by insertion of the upper bound, and noting that under the upper bound all terms involving the redial rates cancel,

$$\begin{aligned}
&\Lambda\big[V^{K+1}(\mathbf{n}+\mathbf{e}_i) - V^{K+1}(\mathbf{n})\big] \\
&\quad \le \sum_{j=1}^{N} \lambda_j 1(\mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \in S) \\
&\qquad + \sum_{j=1}^{N} \lambda_j 1(\mathbf{n}+\mathbf{e}_j \notin S) \\
&\qquad + \sum_{j=1}^{N}\sum_{k=0}^{N} n_j \mu_j p_{jk} 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k \in S) \\
&\qquad + \sum_{j=1}^{N}\sum_{k=1}^{N} n_j \mu_j p_{jk} 1(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \notin S) \\
&\qquad + \sum_{k=1}^{N} \mu_i p_{ik} 1(\mathbf{n}+\mathbf{e}_k \in S) \\
&\qquad + \left(\Lambda - \sum_{j=1}^{N}\lambda_j - \sum_{j=1}^{N}\sum_{k=0}^{N}(n_j+\delta_{ij})\mu_j p_{jk}\right) \\
&\qquad + \sum_{j=1}^{N} \lambda_j 1(\mathbf{n}+\mathbf{e}_j \in S,\ \mathbf{n}+\mathbf{e}_i+\mathbf{e}_j \notin S) \\
&\qquad + \sum_{j=1}^{N}\sum_{k=1}^{N} n_j \mu_j p_{jk} 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k \notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k \in S) \\
&\qquad + \mu_i p_{i0} + \sum_{k=1}^{N} \mu_i p_{ik} 1(\mathbf{n}+\mathbf{e}_k \notin S) \\
&\quad = \sum_{j=1}^{N} \lambda_j
\end{aligned}$$

$$+\sum_{j=1}^{N}\sum_{k=0}^{N}(n_j+\delta_{ij})\mu_j p_{jk}$$

$$+\left(\Lambda-\sum_{j=1}^{N}\lambda_j-\sum_{j=1}^{N}\sum_{k=0}^{N}(n_j+\delta_{ij})\mu_j p_{jk}\right)$$

which completes the proof. □

Corollary 5.12 *Consider the hypercube process $X_{\mathrm{hc},r}$. A sufficient condition for*

$$0\le\left[V_{\mathrm{hc},r}^{K}(\mathbf{n}+\mathbf{e}_i)-V_{\mathrm{hc},r}^{K}(\mathbf{n})\right]\le 1,\qquad \mathbf{n},\mathbf{n}+\mathbf{e}_i\in S$$

is that

$$0\le R(\mathbf{n}+\mathbf{e}_i)-R(\mathbf{n}) \tag{19}$$

$$\le\lambda_i 1(\mathbf{n}+2\mathbf{e}_i\notin S)+\sum_{j=1}^{N}n_j\mu_j p_{ji}1(\mathbf{n}+2\mathbf{e}_i\notin S)+\mu_i p_{i0}$$

$$+\sum_{k=1}^{N}\mu_i p_{ik}1(\mathbf{n}+\mathbf{e}_k\notin S). \tag{20}$$

Proof We use expression (18) for which it can readily be seen that all indicator terms are equal to zero. Hence,

$$\sum_{j=1}^{N}\lambda_j\left[V^K(\mathbf{n}+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_i)\right]1(\mathbf{n}+\mathbf{e}_j\in S,\ \mathbf{n}+\mathbf{e}_i+\mathbf{e}_j\notin S)$$

$$=\sum_{j=1}^{N}\lambda_j\left[V^K(\mathbf{n}+\mathbf{e}_j)-V^K(\mathbf{n}+\mathbf{e}_i)\right]1(j=i,\ \mathbf{n}+\mathbf{e}_i\in S,\ \mathbf{n}+2\mathbf{e}_i\notin S)=0,$$

and

$$\sum_{j=1}^{N}\sum_{k=1}^{N}n_j\mu_j p_{jk}\left[V^K(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-V^K(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\right]$$

$$\times 1(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j+\mathbf{e}_k\notin S,\ \mathbf{n}-\mathbf{e}_j+\mathbf{e}_k\in S)$$

$$=\sum_{j=1}^{N}\sum_{k=1}^{N}n_j\mu_j p_{jk}\left[V^K(\mathbf{n}+\mathbf{e}_i-\mathbf{e}_j)-V^K(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_k)\right]$$

$$\times 1(k=i,\ \mathbf{n}+\mathbf{e}_i\in S,\ \mathbf{n}+2\mathbf{e}_i\notin S)=0.$$

By analogy, the redial rates term cancel, which completes the proof. □

The following result now transforms the comparison of the original and the hypercube process with null redial rates into a condition on the bias terms for only one process, the hypercube process.

Theorem 5.13 *Suppose that for some non-negative function* $\Phi \in C_{\mathrm{hc}}$, *for all* $\mathbf{n} \in S$ *and* $k = 0, 1, 2, \ldots$

$$0 \le \sum_{\mathbf{n}' \in S_{\mathrm{hc}}} \big(q_{\mathrm{hc},0}(\mathbf{n}, \mathbf{n}') - q_0(\mathbf{n}, \mathbf{n}')\big)\big(V_{\mathrm{hc},0}^k(\mathbf{n}') - V_{\mathrm{hc},0}^k(\mathbf{n})\big) < \Lambda \Phi(\mathbf{n}).$$

Then

$$A_{\mathrm{hc},0} - \beta \le A_0 \le A_{\mathrm{hc},0},$$

where

$$\beta = \sum_{\mathbf{n} \in S_{\mathrm{hc}}} \pi_{\mathrm{hc},0}(\mathbf{n}) \Phi(\mathbf{n}).$$

Proof Recall the definition of $\bar{P}$ provided in the proof of Theorem 5.4. By iteration, and by analogy with the proof of Theorem 5.4, we get

$$\big(V_{\mathrm{hc},0}^k - V_0{}^k\big)(\mathbf{0}) = \sum_{t=0}^{k-1} \bar{P}_0^t (P_{\mathrm{hc},0} - \bar{P}_0) V_{\mathrm{hc},0}^{k-t-1}(\mathbf{0}).$$

For notational convenience, we will omit the index 0.

Since $\sum_{\mathbf{n}' \in S_{\mathrm{hc}}} \bar{p}(\mathbf{n}, \mathbf{n}') = 1 = \sum_{\mathbf{n}' \in S_{\mathrm{hc}}} p_{\mathrm{hc},0}(\mathbf{n}, \mathbf{n}')$ we have

$$\begin{aligned}(P_{\mathrm{hc}} - \bar{P}) V_{\mathrm{hc}}^k(\mathbf{n}) &= \sum_{\mathbf{n}' \in S_{\mathrm{hc}}} \big(p_{\mathrm{hc},0}(\mathbf{n}, \mathbf{n}') - \bar{p}(\mathbf{n}, \mathbf{n}')\big) V_{\mathrm{hc}}^k(\mathbf{n}') \\ &= \sum_{\mathbf{n}' \neq \mathbf{n}} \big(p_{\mathrm{hc},0}(\mathbf{n}, \mathbf{n}') - \bar{p}(\mathbf{n}, \mathbf{n}')\big)\big(V_{\mathrm{hc},0}^k(\mathbf{n}') - V_{\mathrm{hc},0}^k(\mathbf{n})\big).\end{aligned}$$

Combination of this result with the hypothesis of the theorem gives

$$\big(V_{\mathrm{hc}}^k - V^k\big)(\mathbf{0}) \le \sum_{t=0}^{k-1} \bar{P}^t \Phi(\mathbf{0}) \le \sum_{t=0}^{k-1} P_{\mathrm{hc},0}^t \Phi(\mathbf{0}) \le k \sum_{\mathbf{n}} \pi_{\mathrm{hc},0}(\mathbf{n}) \Phi(\mathbf{n}), \tag{21}$$

where the second inequality follows from Theorem 5.4. Recall (17). Application of Theorem 5.3 completes the proof. □

Remark 5.14 Note that the condition of the theorem is for $\mathbf{n} \in S$. Further note that $q_0(\mathbf{n}, \mathbf{n}') = 0$ for $\mathbf{n} \in S$, $\mathbf{n}' \notin S$.

Remark 5.15 Note that the theorem can also be formulated with the roles of $X_{\mathrm{hc},0}$ and X_0 reversed. However, this requires an upper bound on $(V_0^k(\mathbf{n}') - V_0^k(\mathbf{n}))$ that usually cannot be obtained.

As a second comparison result, by analogy with the monotonicity result for the transition matrices, the cumulative rewards of the hypercube process also appear to be monotone in the redial rates.

Theorem 5.16 *Consider the processes $X_{\mathrm{hc},r}$ and $X_{\mathrm{hc},r'}$ on state space S_{hc}. Let $r_{ji} \geq r'_{ji}$ for all j, i. Suppose that for some non-negative function $\Phi_{rr'} \in C_{\mathrm{hc}}$, for all $\mathbf{n} \in S_{\mathrm{hc}}$ and $k = 0, 1, 2, \dots$*

$$0 \leq \sum_{\mathbf{n}' \in S_{\mathrm{hc}}} \big(q_{\mathrm{hc},r}(\mathbf{n},\mathbf{n}') - q_{\mathrm{hc},r'}(\mathbf{n},\mathbf{n}')\big)\big(V^k_{\mathrm{hc},r}(\mathbf{n}') - V^k_{\mathrm{hc},r}(\mathbf{n})\big) < \Lambda \Phi_{rr'}(\mathbf{n}).$$

Then

$$A_{\mathrm{hc},r} - \beta_{rr'} \leq A_{\mathrm{hc},r'} \leq A_{\mathrm{hc},r},$$

where

$$\beta_{rr'} = \sum_{\mathbf{n} \in S_{\mathrm{hc}}} \pi_{\mathrm{hc},r}(\mathbf{n}) \Phi_{rr'}(\mathbf{n}).$$

Proof The proof is obtained by analogy with that of Theorem 5.13 but now invoking Lemma 5.6.

For $\mathbf{n} \in S$, under the conditions of Corollary 5.12,

$$\begin{aligned}
&\sum_{\mathbf{n}' \in S_{\mathrm{hc}}} \big(q_{\mathrm{hc},0}(\mathbf{n},\mathbf{n}') - q_0(\mathbf{n},\mathbf{n}')\big)\big(V^k_{\mathrm{hc},0}(\mathbf{n}') - V^k_{\mathrm{hc},0}(\mathbf{n})\big) \\
&\quad = \sum_j \lambda_j 1(\mathbf{n} + \mathbf{e}_j \in S_{\mathrm{hc}} \backslash S)\big(V^k_{\mathrm{hc},0}(\mathbf{n} + \mathbf{e}_j) - V^k_{\mathrm{hc},0}(\mathbf{n})\big) \\
&\qquad + \sum_{i,j} n_i \mu_i p_{ij} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j \in S_{\mathrm{hc}} \backslash S)\big(V^k_{\mathrm{hc},0}(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j) - V^k_{\mathrm{hc},0}(\mathbf{n} - \mathbf{e}_i)\big) \\
&\quad \leq \sum_j \lambda_j 1(\mathbf{n} + \mathbf{e}_j \in S_{\mathrm{hc}} \backslash S) + \sum_{i,j} n_i \mu_i p_{ij} 1(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j \in S_{\mathrm{hc}} \backslash S) = \Phi(\mathbf{n}),
\end{aligned}$$

and $\Phi \in C_{\mathrm{hc}}$. For $\mathbf{n} \in S_{\mathrm{hc}}$, under the conditions of Corollary 5.12, and assuming that $r_{k,j} \geq r'_{k,j}$, for all k, j,

$$\begin{aligned}
&\sum_{\mathbf{n}' \in S_{\mathrm{hc}}} \big(q_{\mathrm{hc},r}(\mathbf{n},\mathbf{n}') - q_{\mathrm{hc},r'}(\mathbf{n},\mathbf{n}')\big)\big(V^k_{\mathrm{hc},r}(\mathbf{n}') - V^k_{\mathrm{hc},r}(\mathbf{n})\big) \\
&\quad = \sum_{k,j} (r_{k,j} - r'_{k,j}) 1(\mathbf{n} + \mathbf{e}_j \in S_{\mathrm{hc}}, \mathbf{n} + \mathbf{e}_k \notin S_{\mathrm{hc}})\big(V^k_{\mathrm{hc},r}(\mathbf{n} + \mathbf{e}_j) - V^k_{\mathrm{hc},r}(\mathbf{n})\big) \\
&\quad \leq \sum_{k,j} (r_{k,j} - r'_{k,j}) 1(\mathbf{n} + \mathbf{e}_j \in S_{\mathrm{hc}}, \mathbf{n} + \mathbf{e}_k \notin S_{\mathrm{hc}}) = \Phi_{rr'}(\mathbf{n}).
\end{aligned}$$

□

A combination of Theorem 5.13 and Theorem 5.16 yields our main error bound result of Theorem 3.4.

6 Concluding remarks

This paper has investigated analytical results for performance measures in networks of Erlang loss queues with common capacity constraints that naturally arise when modelling finite circuit switched communication systems. For such networks, the equilibrium distribution is, in general, not available in closed form. Via a state space modification, and a redial rate approximation, monotonicity results and bounds have been obtained for performance measures including blocking probabilities and throughputs. Both the approximating results for these performance measures, and bounds on the accuracy of the approximation have been obtained in closed form via the product form equilibrium distribution that is available for a network with suitably chosen redial rates.

Results for the upper bound on the approximating performance measures are amenable for dimensioning in practical systems, since the error in these bounds is roughly of the order of magnitude of the performance measure. The lower bounds have been argued to be rather loose. Further research includes improvement of the accuracy of the lower bounds. Furthermore, extension of the bounds to systems with time-dependent arrival rates is among our aims for further research.

Acknowledgements The authors would like to thank Jan van der Wal for a stimulating discussion which led to the counterexample of Sect. 4.4. The research of R.J. Boucherie is partially supported by grant IS043014EW-NL. The authors are grateful to the referees for their careful reading of the manuscript and helpful suggestions.

References

1. Abdalla, N., Boucherie, R.J.: Blocking probabilities in mobile communications networks with time-varying rates and redialing subscribers. Ann. Oper. Res. **112**, 15–34 (2002)
2. Adan, I., van der Wal, J.: Monotonicity of the throughput of a closed queueing network in the number of jobs. Oper. Res. **37**, 935–957 (1989)
3. Baccelli, F., Brémaud, P.: Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences. Springer, Berlin (1994)
4. Boucherie, R.J., Mandjes, M.: Estimation of performance measures for product form cellular mobile communications networks. Telecommun. Syst. **10**, 321–354 (1998)
5. Boucherie, R.J., van Dijk, N.M.: On a queueing network model for cellular mobile communications networks. Oper. Res. **48**, 38–49 (2000)
6. Boucherie, R.J., Mandjes, M., Verwijmeren, S.: Asymptotic evaluation of blocking probabilities in a hierarchical cellular mobile network. Probab. Eng. Inf. Sci. **14**, 81–99 (2000)
7. Coleman, J.L., Henderson, W., Taylor, P.G.: A convolution algorithm for resource allocation problems with moderate user interference. IEEE Trans. Commun. **42**, 1106–1111 (1994)
8. Everitt, D.: Product form solutions in cellular mobile communication systems. Fourth Australian Teletraffic Research Seminar, Paper No. 3.1 (1989)
9. Everitt, D., Macfadyen, N.W.: Analysis of multi-cellular mobile radiotelephone systems with loss. Br. Telecom Tech. J. **1**, 218–222 (1983)
10. Keilson, L.J., Kester, A.: Monotone matrices and monotone Markov processes. Stoch. Process. Appl. **5**, 231–245 (1977)
11. Kelly, F.P.: Loss networks. Ann. Appl. Probab. **1**, 319–378 (1991)

12. Massey, W.A.: Strong orderings for Markov processes on partially ordered spaces. Math. Oper. Res. **12**, 350–367 (1987)
13. Shaked, M., Shanthikumar, J.G.: Stochastic Orders and Their Applications. Academic Press, San Diego (1994)
14. Sonderman, D.: Comparing multi-server queues with finite waiting rooms, I: Same number of servers. Adv. Appl. Probab. **11**, 439–447 (1979)
15. Sonderman, D.: Comparing multi-server queues with finite waiting rooms, II: Different number of servers. Adv. Appl. Probab. **11**, 448–455 (1979)
16. Stoyan, D.: Bounds and approximations in queueing through monotonicity and continuity. Oper. Res. **25**, 851–863 (1977)
17. Stoyan, D.: Comparison Method for Queues and Other Stochastic Models. Wiley, Chichester (1983)
18. Pallant, D.L., Taylor, P.G.: Approximations of performance measures in cellular mobile networks with dynamic channel allocation. Telecommun. Syst. **3**, 129–163 (1994)
19. Pallant, D.L., Taylor, P.G.: Modeling handovers in cellular mobile networks with dynamical channel allocation. Oper. Res. **43**, 33–42 (1995)
20. Ross, K.W.: Multiservice Loss Models for Broadband Telecommunication Systems. Springer, London (1995)
21. Tijms, H.C.: Stochastic Models: An Algorithmic Approach. Wiley, New York (1994)
22. Tsoucas, P., Walrand, J.: Monotonicity of throughput in non-Markovian networks. J. Appl. Probab. **26**, 134–141 (1989)
23. van Dijk, N.M.: Perturbation theory for unbounded Markov reward processes with applications to queueing. Adv. Appl. Probab. **20**, 99–111 (1988)
24. van Dijk, N.M.: The importance of bias terms for error bounds and comparison results. In: Stewart, W.J. (ed.) Numerical Solution of Markov Chains. Dekker, New York (1991)
25. van Dijk, N.M.: Queueing Networks and Product Forms: A Systems Approach. Wiley, Chichester (1993)
26. van Dijk, N.M., Miyazawa, M.: A note on bounds and error bounds for non- exponential batch arrival systems. Probab. Eng. Inf. Sci. **11**, 189–201 (1997)
27. van Dijk, N.M., Putterman, M.L.: Perturbation theory for Markov reward processes with applications to queueing systems. Adv. Appl. Probab. **20**, 79–98 (1988)
28. van Dijk, N.M., Tsoucas, P., Walrand, J.: Simple bounds and monotonicity of the call congestion of infinite multiserver delay systems. Probab. Eng. Inf. Sci. **2**, 129–138 (1988)
29. Whitt, W.: Comparing counting processes and queues. Adv. Appl. Probab. **13**, 207–220 (1981)