KYBERNETIKA — VOLUME 33 (1997), NUMBER 2, PAGES 171–184

ERROR BOUNDS FOR ARBITRARY APPROXIMATIONS OF "NEARLY REVERSIBLE" MARKOV CHAINS AND A COMMUNICATIONS EXAMPLE

NICO M. VAN DIJK

A condition is provided to conclude error bounds when using an arbitrary steady state approximation of a "nearly reversible" Markov chain. The error bound is of the form ΔR where

(i) Δ can be computed by the approximation in order

(ii) R can be obtained analytically by the system of interest.

The results will be illustrated for a communication system with different source characteristics. An approximation is suggested based on truncating the corresponding Möbius-function. An R-value is obtained by an inductive Markov reward equation. Numerical illustration indicates that the error bound can be useful for practical purposes.

1. INTRODUCTION

Motivation

Markov chain theory has proven to be a powerful tool for performance evaluation of computer and communication systems. Unfortunately, such systems rarely exhibit a closed form expression, like a product form, due to practical phenomena such as blocking or source interactions. As exact numerical analysis can be computationally expensive or even infeasible, approximations have been widely developed.

Approximative approaches are usually supported by extensive experimental illustration and heuristic or intuitive argumentation. Analytic a priori or on line error bounds, however, are rarely reported and seem more or less restricted to numerical or exact decomposition and aggregation procedures (cf. [4, 5, 10, 11, 15, 20, 25]).

Robust but secure error bounds are of practical interest to obtain:

- A 100 % secure order of magnitude.
- A restricted interval of possible values to which attention can be restricted, such as for simulation or optimization purposes.
- A guarantee of possible correctness or incorrectness of model assumptions, conjectures and approximation techniques.

Recently, in [20] and [22] conditions have been provided by which one can derive analytic a priori error bounds for the effect of small perturbations, system modifications or state space truncations. In practice, however, approximations may not be based on just a perturbation, modification or truncation. In contrast, approximations may involve a totally different underlying law of motion and not even be interpretable as corresponding to some modified system.

For example, approximations for queueing networks with blocking or failures are usually based on (iteratively) adapting effective service rates as if service stations can be regarded in isolation and using these in the analogue system without such features (see for example various papers in [1]). Or approximations may follow by analytic simplifications which do no longer fit a direct probabilistic or system descriptive interpretation (cf. [25]).

General result

This paper therefore aims to provide a tool by which one may also derive analytic error bounds for arbitrary approximations, not necessarily based on a system modification, but instead, which are merely based on some given approximation for the steady state distribution. These error bounds can be excepted to be *"reasonable"* when the system in *"nearly reversible"*.

Nearly reversibility

Reversibility (cf. [8]) is a most important concept in queueing network theory, as it can be shown to be an indirect characterization of so-called product form results (cf. [7]). While (strict) reversibility of communication networks is limited to special or artificial protocols which ignore practical features such as collisions, propagation delays and retransmissions, "nearly reversibility" appears to be quite common in practical communication networks. Here *"nearly reversibility"* is not a standard or well-defined concept in the literature but roughly stands for a strict reversibility up to some minor modification of one or a few of the underlying descriptions, or more precisely, up to some reasonably small discrepancy in the reversibility (balance) equations for the steady state distribution. For example, a single server system with breakdowns is strictly reversible up to the occurrence of these breakdowns, which will only take place rarely. Similarly, a communication network such as an ALOHA or a CSMA system is strictly reversible up to the occurrence of collisions, which should not take place too often.

Steps involved

Two steps are involved in order to establish an error bound of the form $\boldsymbol{\Delta R}$:

(i) The definition of an artificial Markov chain based on the steady state approximation used. This directly leads to the computation of the difference value $\boldsymbol{\Delta}$.

(ii) The estimation of so-called bias terms by a value $\boldsymbol{R}$ for the given underlying Markov reward structure. This step does not depend on the used approximation and in concrete situations can usually be established analytically.

By providing the estimate $\boldsymbol{R}$, error bounds $\boldsymbol{\Delta}\,\boldsymbol{R}$ can thus be compared for various alternative approximations.

Special application

Most of the paper will be concerned with the illustration of both steps for a particular application of practical interest: *An ALOHA-system with inhomogeneous source characteristics and collision probabilities.* To perform the first step an approximation will be given based on a so-called Möbius-expansions [24]. This approximation is chosen as it cannot just be seen as some sort of physical modification (or perturbation) of the original system. A simple bound on the essential bias-term will be derived. An explicit error bound for the system throughput is hereby obtained. Numerical support indicates that the error bound is useful in practice.

Related literature

The definition of the artificial Markov chain used seems to be new in the present setting but is related to the splitting of linear operators in a symmetric (or self-adjoint) and anti-symmetric part (see [2] and [3]).

The comparison of the artificial chain and the original model is closely related to a theorem that has recently been reported in [20] and [22] to establish perturbation or truncation results. However, it does not fit in either of these references directly. The estimation of so-called bias terms by means of an inductive proof-technique has already successfully been employed in a number of queueing situations. The current application to an ALOHA-system, however, is new and involves special technicalities as state-dependent collision probabilities that are to be dealt with. The approximation for this application is adopted from a recently developed approach in [24] based on the truncation of so-called Möbius-expansions.

In [16] a perturbation theorem is provided for estimating the difference between the steady state distribution of a finite Markov chain with that of a perturbation of that chain. This theorem is also related to earlier theorems such as in [10] and [15] and provides error bounds which explicitly depend on the finite number of states. Such a condition is not required in this paper. Furthermore, the error bound is expressed in L_1-norms, while the bias-term approach adopted herein from [20] allows to obtain explicit error bounds in state dependent terms. In particular, this allows one to use state dependent modifications and even truncations, provided their likelihood is sufficiently small or provided a weighting function can be given, as illustrated for example in [20]. Furthermore, this set-up also provides comparison or monotonicity results at the same time (see Corollary 2.3). For the present applications in finite cases, the error bounds from [20] and [22] could lead to a similar bound as in [16]. Similarly, in [6] an approach has been developed by estimating arbitrary Markov chains by reversible upper and lower bound modifications. These references though merely provide rough performance bounds and not (small) error bounds for (accurate) given approximations as in this paper.

2. MODEL AND RESULT

Consider a continuous-time Markov chain with state space $N = \{1, 2, \ldots\}$ and transition rates $q(i,j)$ for a transition from state i into state j. Without restriction of generality assume that this chain is irreducible at some set S with unique stationary distribution $\{\pi(i), i \in S\}$.

Let $\{\overline{\pi}(i), i \in \overline{S}\}$ be some given *arbitrary approximating probability distribution* at some subset $\overline{S} \subset S$ and *define* transition rates at $\overline{S}$ by:

$$\boxed{\overline{q}(i,j) = \tfrac{1}{2}\left[q(i,j) + q(j,i)\,\tfrac{\overline{\pi}(j)}{\overline{\pi}(i)}\right]} \tag{2.1}$$

for all $i, j \in \overline{S}$ while $\overline{q}(i,j)$ is defined to be equal to 0 otherwise. Then for all i, j one directly verifies the reversibility property:

$$\boxed{\overline{\pi}'(i)\,\overline{q}(i,j) = \overline{\pi}(j)\,\overline{q}(j,i)} \tag{2.2}$$

Without restriction of generality, also assume that the approximate Markov chain with transition rates $\overline{q}(i,j)$ as per (2.1) is irreducible at $\overline{S}$, so that its unique stationary distribution is given by $\{\overline{\pi}(\cdot), i \in \overline{S}\}$. (Note that the global balance equations are directly verified by summing (2.2) over all j.)

Let $r : S \to \mathbb{R}$ be some given reward function, to be interpreted as a reward per step, and consider the stationary (or average) reward measures:

$$\begin{cases} G = \sum_{i \in S} \pi(i)\, r(i) \\ \overline{G} = \sum_{i \in \overline{S}} \overline{\pi}(i)\, \overline{r}(i) \end{cases} \tag{2.3}$$

where we assume that these are well-defined. We wish to evaluate the difference

$$|G - \overline{G}|.$$

First, let us make some notational conventions. An upper bar "–" symbol indicates the approximate model. Without further mentioning we only give definitions for the original chain and we directly adopt these definitions for the approximate model by adding the upper bar "–" symbol. With reference to Remark 2.4 below, for convenience assume that for some finite scalar $Q < \infty$:

$$Q \geq \sup_i \sum_j q(i,j). \tag{2.4}$$

Then by the standard uniformization (or randomization) technique (cf. [18], p. 110) the distribution π is equal to the unique stationary distribution of the discrete time Markov chain with one step transition probabilities $p(i,j)$ defined by:

$$\begin{cases} p(i,j) = q(i,j)/Q \quad (j \neq i) \\ p(i,i) = \left[1 - \sum_{j \neq i} q(i,j)/Q\right]. \end{cases} \tag{2.5}$$

Now let $\mathbb{C} = \{f | f : S \to \mathbb{R}\}$ be the space of real-valued functions on S and define transition operators $\{\boldsymbol{T}_t | t = 0, 1, 2, \ldots\} : \mathbb{C} \to \mathbb{C}$ by $\boldsymbol{T}_0 f = f$ and for $t \geq 0$:

$$(\boldsymbol{T}_{t+1} f)(i) = \sum_j \boldsymbol{p}(i, j)(\boldsymbol{T}_t f)(j). \tag{2.6}$$

In words that is, $\boldsymbol{T}_t f(i)$ represents the expected value of function f at time t under one-step transition probabilities $\boldsymbol{p}(\cdot, \cdot)$ and starting in state i at time 0. Now define functions $\{\boldsymbol{V}_n | n = 0, 1, 2 \ldots\} : S \to \mathbb{R}$ by:

$$\boldsymbol{V}_n = Q^{-1} \sum_{t=0}^{n-1} \boldsymbol{T}_t r. \tag{2.7}$$

In words that is, $\boldsymbol{V}_n(i)$ represents the expected total reward received over n periods under one-step transition probabilities $\boldsymbol{p}(i, j)$ when starting in state i at time 0 and incurring a one-step reward $\boldsymbol{r}(j)$ whenever the system visits state j. By virtue of the uniformization technique (cf. [18], p. 110) and the irreducibility assumptions, for any $\ell \in S$, we then have:

$$\boldsymbol{G} = \lim_{N \to \infty} \frac{Q}{N} \boldsymbol{V}_N(\ell) \tag{2.8}$$

provided this limit exists. This value $\boldsymbol{G}$ represents the expected average reward per unit of time. We note here that the factors Q^{-1} and Q in (2.7) and (2.8) could be omitted but are included for convenience later on. We can now present a general comparison result to compare the original model with an arbitrary approximation. A more practical corollary will be concluded directly.

Theorem 2.1. Assume that for some function Φ, an initial state $\ell \in \overline{S}$, all $i \in \overline{S}$ and $t \geq 0$:

$$\left| (\boldsymbol{r} - \overline{\boldsymbol{r}})(i) + \frac{1}{2} \sum_j \left[\boldsymbol{q}(i, j) - \boldsymbol{q}(j, i) \frac{\overline{\pi}(j)}{\overline{\pi}(i)} \right] [\boldsymbol{V}_t(j) - \boldsymbol{V}_t(i)] \right| \leq \alpha\, \Phi(j) \tag{2.9}$$

and

$$\overline{\boldsymbol{T}}_t\, \Phi(\ell) \leq \beta. \tag{2.10}$$

Then

$$\boxed{|\boldsymbol{G} - \overline{\boldsymbol{G}}| \leq \alpha\, \beta.} \tag{2.11}$$

Proof. By virtue of (2.6) and (2.7) we have:

$$\begin{cases} \boldsymbol{V}_{t+1} = \boldsymbol{r}\, Q^{-1} + \boldsymbol{T}\, \boldsymbol{V}_t \\ \overline{\boldsymbol{V}}_{t+1} = \overline{\boldsymbol{r}}\, Q^{-1} + \overline{\boldsymbol{T}}\, \overline{\boldsymbol{V}}_t. \end{cases} \tag{2.12}$$

Hence, for any n and arbitrary $\ell \in \overline{S}$:

$$\begin{aligned}(\overline{\boldsymbol{V}}_n - \boldsymbol{V}_n)(\ell) &= (\overline{\boldsymbol{r}} - \boldsymbol{r})(\ell)\, Q^{-1} + (\overline{\boldsymbol{T}}\,\overline{\boldsymbol{V}}_{n-1} - \boldsymbol{T}\,\boldsymbol{V}_{n-1})(\ell) \\ &= (\overline{\boldsymbol{r}} - \boldsymbol{r})(\ell)\, Q^{-1} + (\overline{\boldsymbol{T}} - \boldsymbol{T})\,\boldsymbol{V}_{n-1}(\ell) + \overline{\boldsymbol{T}}(\overline{\boldsymbol{V}}_{n-1} - \boldsymbol{V}_{n-1})(\ell).\end{aligned}$$

By repeating this relation for $n = N, N-1, \dots, 1$ we find:

$$\begin{aligned}&(\overline{\boldsymbol{V}}_N - \boldsymbol{V}_n)(\ell) \\ &\quad = \sum_{t=0}^{N-1} \overline{\boldsymbol{T}}_t \left([\overline{\boldsymbol{r}} - \boldsymbol{r}]\, Q^{-1} + [(\overline{\boldsymbol{T}} - \boldsymbol{T})\,\boldsymbol{V}_{n-t-1}]\right)(\ell) + \overline{\boldsymbol{T}}_N(\overline{\boldsymbol{V}}_0 - \boldsymbol{V}_0)(\ell).\end{aligned} \tag{2.13}$$

Further, by (2.6), (2.5) and (2.1) we obtain for any i:

$$\begin{aligned}(\overline{\boldsymbol{T}} - \boldsymbol{T})\,\boldsymbol{V}_s(i) &= \sum_j [\overline{\boldsymbol{p}}(i,j) - \boldsymbol{p}(i,j)]\,\boldsymbol{V}_s(j) \\ &= \sum_j [\overline{\boldsymbol{p}}(i,j) - \boldsymbol{p}(i,j)]\,[\boldsymbol{V}_s(j) - \boldsymbol{V}_s(i)] \\ &= -\frac{1}{2} Q^{-1} \sum_j \left[\boldsymbol{q}(i,j) - \boldsymbol{q}(j,i)\,\frac{\overline{\pi}(j)}{\overline{\pi}(i)}\right][\boldsymbol{V}_s(j) - \boldsymbol{V}_s(i)]\,.\end{aligned} \tag{2.14}$$

By substituting (2.14) in (2.13), nothing that $\overline{\boldsymbol{V}}_0(\cdot) = \boldsymbol{V}_0(\cdot)$, taking absolute values and using that $\overline{\boldsymbol{T}}_t$ is a monotone operator, i.e. $\overline{\boldsymbol{T}}_t\, f(i) \le \overline{\boldsymbol{T}}_t\, g(i)$ if $f(j) \le g(j)$ for all j, we obtain from (2.9), (2.10), (2.13) and (2.14):

$$\left|(\overline{\boldsymbol{V}}_n - \boldsymbol{V}_n)(\ell)\right| \le \alpha\, Q^{-1} \sum_{t=0}^{N-1} \overline{\boldsymbol{T}}_t\, \Phi(\ell) \le \alpha\, \beta\, N\, Q^{-1}. \tag{2.15}$$

Applying (2.8) completes the proof. □

Though condition (2.9) does in principle allows one to combine the approximate values $\overline{\pi}(\cdot)$ with the bias-terms $\boldsymbol{V}_t(j) - \boldsymbol{V}_t(i)$, it is more realistic and convenient that these will be analyzed separately. This is expressed by the following practical corollary, where for simplicity we assume $\overline{\boldsymbol{r}} = \boldsymbol{r}$ and $\Phi(\cdot) = 1$.

Corollary 2.2. Let $\overline{\boldsymbol{r}} = \boldsymbol{r}$ and assume that for some $\boldsymbol{\Delta}$ and all $i \in \overline{S}$:

$$\left|\frac{1}{2} \sum_j \left[\boldsymbol{q}(i,j) - \boldsymbol{q}(j,i)\,\frac{\overline{\pi}(j)}{\overline{\pi}(i)}\right]\right| \le \boldsymbol{\Delta} \tag{2.16}$$

and that for some $\boldsymbol{R}$, for all $i, j \in S$ with $\boldsymbol{q}(i,j) > 0$, and $t \ge 0$:

$$|\boldsymbol{V}_t(j) - \boldsymbol{V}_t(i)| \le \boldsymbol{R}. \tag{2.17}$$

Then

$$|\boldsymbol{G} - \overline{\boldsymbol{G}}| \le \boldsymbol{\Delta}\,\boldsymbol{R}. \tag{2.18}$$

It can also be of interest to investigate whether the proposed approximation provides an upper or lower bound of some performance measure. To this end, a more relaxed form of (2.9) is given in the following corollary.

Corollary 2.3 (Comparison result) Suppose that for all $i \in \overline{S}$ and $t \geq 0$:

$$(r - \overline{r})(i) + \frac{1}{2} \sum_j \left[q(i,j) - q(j,i) \frac{\overline{\pi}(j)}{\overline{\pi}(i)} \right] [V_t(j) - V_t(i)] \geq (\leq)\ 0. \tag{2.19}$$

Then

$$G \geq (\leq)\ \overline{G}.$$

Proof. This follows directly by substituting (2.14) and (2.19) in (2.13), recalling that the operators $\overline{T}_t$ are monotone and applying (2.8). □

Remark 2.4. The uniformization condition (2.4) may seem strong. However, it is always satisfied for finite systems. For infinite systems it can be violated, such as for an infinite server queue. However, in such practical cases also an "approximate uniformization" can be applied by choosing an arbitrarily large Q. As based on convergence results in [21] and noting that the value Q does not arises in (2.11), the validity of (2.11) can therefore also be concluded in such cases.

3. APPLICATION: A COMMUNICATION SYSTEM WITH INHOMOGENEOUS CHARACTERISTICS

This section deals with a special application in order to illustrate the results of Section 2, most notably condition (2.17). Herein, for presentational convenience and clarity we restrict to an ALOHA-type communication system with only four sources. The extension to any number of sources, however, merely involves more complexity and is essentially the same. The approximation used is based on a recently developed approach in [24] by truncating Möbius expansions. We choose this approximation for two reasons:

(i) To illustrate the results for an approximation that cannot be seen as just a simple modification of the system protocols or law of motion.

(ii) To advocate this new approximation approach.

3.1. Model

Consider a communication system with M sources (transmitters or processors) of which each can be in an idle (non-transmitting) or busy (transmitting) mode as follows. When idle a source h will schedule a transmission request after an exponential time with parameter γ_h. Its transmission has an exponential duration with parameter μ_h. However, with $H = \{h_1, \ldots, h_n\}$ denoting that sources $h_1, \ldots, h_n$, say in increasing order of number, are currently busy, a transmission by source $h \notin H$ is:

$$\begin{cases} \text{accepted an initiated with probability:} & \beta(h|H) \\ \text{rejected and lost with probability:} & 1 - \beta(h|H). \end{cases}$$

Here we make the natural assumption that these acceptation probabilities can only become smaller if more sources are busy, i.e. for all h, s and H:

$$\beta(h|H+s) \leq \beta(h|H). \tag{3.1}$$

When lost the source remains idle to schedule a new transmission request after another exponential time. By this latter blocking probability we can model different aspects of which we give two examples.

Two examples

(i) (Slotted ALOHA) *In slotted ALOHA, transmissions are time-slotted in time slots of length δ and take place along a single transmission switch. As this switch can handle only one task (acceptation or completion) at a time, a transmission can be accepted only if none of the ongoing transmissions is completed during that same time-slot. Hence, with $w_s = (1 - e^{-\delta\gamma_s})$ the probability that source h completes its transmission during a time-slot of length δ is given by:*

$$\beta(h|H) = \prod_{s \in H} (1 - w_s) = \exp(-\delta \sum_{s \in G} \gamma_s). \tag{3.2}$$

(ii) (Common memory utilization) *As another example, a busy source may communicate with (retrieve data from or store data at) some memory device M, say during a fraction w_s of its busy time. However, to start a transmission, a source must first retrieve some information from this memory. As this memory can communicate with only one source at a time, none of the busy sources may thus communicate with this memory upon transmission request. This is modeled by:*

$$\beta(h|H) = \prod_{s \in H} (1 - w_s). \tag{3.3}$$

Both examples can be shown (cf. [23], [24]) to be reversible (cf. [8] for a definition) *if and only if* the source characteristics w_s are the same for all sources, i.e. $w_s = w$ for some w and all s. In that case, the steady state distribution is given by (cf. [23], [24]):

$$\pi(H) = c(1-w)^{[(M-[H])]!-1} \prod_{h \in H} [\gamma_h / \mu_h] \tag{3.4}$$

where c is normalizing constant and $[H]$ the cardinality of H. For the case with unequal values w_s, no simple explicit expression seems to be available. As the numerical computation becomes computationally expensive for reasonably large systems, an approximation that can reduce the computation is thus of interest.

3.2. Approximation

Below, we merely present the idea of a general approximation procedure which is developed in [24] and apply it to a system with four sources. The insight and underlying details are based on so-called Möbius expansions as investigated in more detail in this reference as also related to [14].

General approach

Consider a continuous-time Markov process with state description H and transition rate $\boldsymbol{q}(H, H')$ such that $\boldsymbol{q}(H, H') = 0$ if $||H'| - |H|| \geq 2$. That is, only one source can change its status at a time. When the process is reversible the stationary distribution $\pi(\cdot)$ can be expressed as (cf. [21]):

$$\begin{cases} \pi(H) \;=\; \exp \boldsymbol{W}(H) \quad \text{with} \\ \boldsymbol{W}(H) = \sum_{B \subseteq H,\, |B| \leq K} \boldsymbol{W}(B) \sum_{j=0}^{K-|B|} (-1)^j \begin{pmatrix} |H| - |B| \\ j \end{pmatrix}, \end{cases} \tag{3.5}$$

where K is some specific integer that follows from the transition structure. (Roughly speaking, K is the value such that states which differ in more than K sources do not influence each other in terms of a potential interpretation similar to physical interactions.) As a result, in the reversible case, the stationary probabilities of all states can be expressed in states with cardinality less or equal than K. Here one must typically think of K to be small, say $K = 1$, 2 or 3. For example, with $K = 2$ and $M = 4$ we would obtain

$$\begin{cases} \pi(i,j,k) \;=\; \pi(\phi)\, \frac{\pi(i,j)\,\pi(i,k)\,\pi(j,k)}{\pi(i)\,\pi(j)\,\pi(k)} \\ \pi(1,2,3,4) \;=\; [\pi(\phi)]^3\, \frac{\pi(1,2)\,\pi(1,3)\,\pi(1,4)\,\pi(2,3)\,\pi(2,4)\,\pi(3,4)}{[\pi(1)\,\pi(2)\,\pi(3)\,\pi(4)]^2}. \end{cases} \tag{3.6}$$

In the non-reversible case, these latter relations fail but still do seem *reasonable as approximations*. Roughly speaking, the approximation then comes down to ignoring interactions of states which differ in more than K sources.

Remark 3.1. Clearly, the accuracy of such an approximation will decrease the larger K, that is the more different states interact while having up to K different components, and the larger M, as this allows more different components.

Application to model of Section 3.1

Now reconsider the model of Section 3.1, where one can think for instance of either example (i) or (ii) and for presentational convenience assume $M = 4$. In the case of unequal values w_s a reasonable approximation $\overline{\pi}(\cdot)$ thus seems to be suggested by the reduced set of global balance equations:

$$\begin{cases} \overline{\pi}(H) \sum_{H'} \boldsymbol{q}(H, H') = \sum_{H'} \overline{\pi}(H')\, \boldsymbol{q}(H', H) & \text{for } H \text{ with } |H| \leq 2 \\ (3.6) \text{ with } \pi(\cdot) \text{ replaced by } \overline{\pi}(\cdot), & \text{for } H \text{ with } |H| > 2. \end{cases} \tag{3.7}$$

Here

$$\boldsymbol{q}(H, H') = \begin{cases} \gamma_h\, \beta(h|H) & \text{for } H' = H' \cup \{h\} \quad (h \notin H) \\ \mu_h & \text{for } H' = H \setminus \{h\} \quad (h \in H) \end{cases} \tag{3.8}$$

and $\boldsymbol{q}(H, H') = 0$ otherwise. Numerical computations (e.g. see Section 3.3) have shown that this approximation is quite reasonable for the total distribution. However, no theoretical and guaranteed error bounds for the accuracy of these approximations have been reported. To this end, we aim to investigate the conditions of Corollary 2.2.

Remark 3.2. Related to Remark 3.1, the accuracy for this application would even be larger if instead of $|H| \leq 2$ we would approximate at $|H| \leq 3$. This would require more approximate computations. Conversely, the accuracy would become unacceptable if we would just have approximated at $|H| \leq 1$, as if each source can be treated in isolation. Also for larger M we could still restrict to $|H| \leq 2$, but of course the accuracy will somewhat decrease in M.

3.3. Estimation of bias–terms: R

We need to verify the essential condition (2.17). Here we note that we can simply identify a state i with a state H, so that all results of Section 2 can be adopted directly for this multi-dimensional application. As performance measure of interest we aim to evaluate the throughput of the system by:

$$\boldsymbol{r}(H) = \sum_{s \notin H} \gamma_s \, \beta(s|H).$$

Further, we choose Q by:

$$Q = \sum_h [\gamma_h + \mu_h]. \tag{3.9}$$

Let $\boldsymbol{V}_t(\cdot)$ be defined by (2.5), (2.6) and (2.7) with the transition rates (3.8) substituted. As a result, we only need to verify (2.17) for states of the form $i = H$ and $j = H \setminus \{h\}$ and $j = H \cup \{h\}$. To this end, for convenience write $H - h = H \setminus \{h\}$ and $H + h = H \cup \{h\}$.

Lemma 3.1. Assume that $\mu_h \geq \sum_s \gamma_s$ for all h, then for all H, $H+h$ and $t \geq 0$:

$$\boxed{0 \leq \boldsymbol{V}_t(H) - \boldsymbol{V}_t(H+h) \leq 1.} \tag{3.10}$$

P r o o f. This will be established by induction in t. As $\boldsymbol{V}_0(\cdot) = 0$, (3.10) holds for $t = 0$. Suppose that (3.10) holds for $t \leq m$ and all H, $H+h$. The following relations then follow by expressing $\boldsymbol{V}_{m+1}(H)$ as according to (2.12). Herein, we note in advance that some terms are rewritten or artificially added and subtracted (e. g. the term $\gamma_h \, Q^{-1} \, \beta(h|H) \, \boldsymbol{V}_m(H+h)$ in the first relation) or split (e. g. $\beta(s|H) = \beta(s|H+h) + [\beta(s|H) - \beta(s|H+h)]$ also in the first relation) in order to obtain terms with equal coefficients that can be compared pairwise. Further, we recall (3.1).

Then by virtue of (2.12):

$$
\begin{aligned}
&\boldsymbol{V}_{m+1}(H) \\
&= \sum_{s \notin H} \gamma_s \, Q^{-1} \, \beta(s|H) + \sum_{s \in H} \mu_s \, Q^{-1} \, \boldsymbol{V}_m(H-s) \\
&+ \sum_{s \notin H} \gamma_s \, Q^{-1} \, \beta(s|H) \, \boldsymbol{V}_m(H+s) + \left[1 - \sum_{s \in H} \mu_s \, Q^{-1} - \sum_{s \notin H} \gamma_s \, Q^{-1} \right] \boldsymbol{V}_m(H) \\
&= \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H) + \gamma_h \, Q^{-1} \, \beta(h|H) + \sum_{s \in H} \mu_s \, Q^{-1} \, \boldsymbol{V}_m(H-s) + \mu_h \, Q^{-1} \, \boldsymbol{V}_m(H) \\
&+ \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H+h) \, \boldsymbol{V}_m(H+h) \\
&+ \sum_{s \notin H+h} \gamma_s \, Q^{-1} \left[\beta(s|H) - \beta(s|H+h) \right] \boldsymbol{V}_m(H+s) + \gamma_h \, Q^{-1} \, \beta(h|H) \, \boldsymbol{V}_m(H+h) \\
&+ \left[1 - \sum_{s \in H+h} \mu_s \, Q^{-1} - \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H) - \gamma_h \, Q^{-1} \, \beta(h|H) \right] \boldsymbol{V}_m(H).
\end{aligned}
$$

And similarly

$$
\begin{aligned}
&\boldsymbol{V}_{m+1}(H+h) \\
&= \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H+h) + \sum_{s \in H} \mu_s \, Q^{-1} \, \boldsymbol{V}_m(H+h-s) + \mu_h \, Q^{-1} \, \boldsymbol{V}_m(H) \\
&+ \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H+h) \, \boldsymbol{V}_m(H+h+s) \\
&+ \left[1 - \sum_{s \in H+h} \mu_s \, Q^{-1} - \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H+h) \right] \boldsymbol{V}_m(H+h) \\
&= \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H+h) + \mu_h \, Q^{-1} \, \boldsymbol{V}_m(H) + \gamma_h \, Q^{-1} \, \beta(h|H) \, \boldsymbol{V}_m(H+h) \\
&+ \sum_{s \in H} \mu_s \, Q^{-1} \, \boldsymbol{V}_m(H+h-s) + \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H+h) \, \boldsymbol{V}_m(H+h+s) \\
&+ \sum_{s \notin H+h} \mu_s \, Q^{-1} \left[\beta(s|H) - \beta(s|H+h) \right] \boldsymbol{V}_m(H+h) \\
&+ \left[1 - \sum_{s \in H+h} \mu_s \, Q^{-1} - \sum_{s \notin H+h} \gamma_s \, Q^{-1} \, \beta(s|H) - \gamma_s \, Q^{-1} \, \beta(h|H) \right] \boldsymbol{V}_m(H+h).
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\boldsymbol{V}_{m+1}(H) - \boldsymbol{V}_{m+1}(H+h) \\
&= \gamma_h\, Q^{-1}\, \beta(h|H) + \sum_{s \notin H+h} \gamma_s\, Q^{-1}[\beta(s|H) - \beta(s|H+h)] \\
&+\gamma_h\, Q^{-1}\, \beta(h|H)\, [\boldsymbol{V}_m(H+h) - \boldsymbol{V}_m(H+h)] + \mu_h\, Q^{-1}[\boldsymbol{V}_m(H) - \boldsymbol{V}_m(H)] \\
&+ \sum_{s \in H} \mu_s\, Q^{-1}[\boldsymbol{V}_m(H-s) - \boldsymbol{V}_m(H+h-s)] \\
&+ \sum_{s \notin H+h} \gamma_s\, Q^{-1}\, \beta(s|H+h)\, [\boldsymbol{V}_m(H+s) - \boldsymbol{V}_m(H+h+s)] \\
&+ \sum_{s \notin H+h} \gamma_s\, Q^{-1}[\beta(s|H) - \beta(s|H+h)]\, [\boldsymbol{V}_m(H+s) - \boldsymbol{V}_m(H+h)] \\
&+ \left[1 - \sum_{s \in H+h} \mu_s\, Q^{-1} - \sum_{s \notin H} \gamma_s\, Q^{-1}\, \beta(s|H)\right] [\boldsymbol{V}_m(H) - \boldsymbol{V}_m(H+h)] \qquad (3.11)
\end{aligned}
$$

where actually the third and fourth term are equal to 0 but kept in for clarity and the use of an argument below. First note that we can write

$$
\boldsymbol{V}_m(H+s) - \boldsymbol{V}_m(H+h) = [\boldsymbol{V}_m(H+s) - \boldsymbol{V}_m(H)] + [\boldsymbol{V}_m(H) - \boldsymbol{V}_m(H+h)]. \qquad (3.12)
$$

The first term of the right hand side of (3.12) is non-positive but estimated from below by -1 as per induction assumption (3.10) for $t = m$. In relation (3.11) this is compensated by the term:

$$
\sum_{s \notin H+h} \gamma_s\, Q^{-1}[\beta(s|H) - \beta(s|H+h)]
$$

which equals the coefficient of (3.12) in (3.11). By substituting the lower estimate 0 from (3.10) for $t = m$ in all other terms, the right hand side of (3.11) is thus estimated from below by 0. That is, we have proven the lower estimate 0 in (3.10) also for $t = m + 1$.

To conclude the upper estimate 1, now recall that the third and fourth term in the right hand side of (3.11) are equal to 0, while also $\mu_h \geq \sum_s \gamma_s$. As a consequence, these 0-terms can compensate for the first two additional positive terms. More precisely, by estimating the right hand side of (3.12) from above by $\boldsymbol{V}_m(H) - \boldsymbol{V}_m(H+h)$, as justified by the hypothesis (3.10) for $t = m$, substituting the upper estimate 1 from (3.10) in all terms and nothing that all coefficients sum up to 1, the right hand side of (3.11) is estimated from above by 1. That is, we have also proven the upper estimate 1 in (3.10) for $t = m + 1$.

The induction in m now completes the proof. □

3.4. Numerical examples

We can now apply Corollary 2.2 with $\boldsymbol{R} = 1$. The value $\boldsymbol{\Delta}$ is thus to be computed by substituting the approximations as per (3.7). Numerical illustration is provided below. Here we note that realistically for the applications as described in Section 3.1, most notably the slotted ALOHA-model, the w-value should be thought of as being small, in which case the results are quite reasonable (in the order of 1 %). But also larger *less realistic* w values are included for the purpose of *testing*. Roughly, the results indicate that the error bound as based on (2.16) will be quite rough compared to the exact error bound but yet quite reasonable, in order of a few percent, as a 100 % secure bound on the imprecision involved. The exact performance value $\boldsymbol{G}$ is included for comparison and is obtained by numerically solving the system.

Remark 3.3. Similar results can be expected with larger numbers of sources M. Of course, with the approximation kept at $|H| \leq 2$, the accuracy would somewhat decrease. The accuracy, however, will predominantly be determined by the truncation level $L : |H| \leq L$ and by the differences on the w_s-values.

Numerical results

$(\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 1)$
$(\mu_1 = \mu_2 = \mu_3 = \mu_4 = 1)$
$\boldsymbol{G}$: expected number of idle sources
$\overline{\boldsymbol{G}}$: approximate value based on $\overline{\pi}$ as per (3.7)
ε: error bound for $|\boldsymbol{G} - \overline{\boldsymbol{G}}|$ as based on (2.16) and (2.17), $(\boldsymbol{R} = 1)$

Ex.	(w_1, w_2, w_3, w_4)	$\boldsymbol{G}$	$\overline{\boldsymbol{G}}$	$\boldsymbol{\Delta}$ (by (2.16))	$\varepsilon = \boldsymbol{\Delta R}$
1	(.5, .5, .5, .45)	2.628953	2.628985	.114	4.3 %
2	(.2, .2, .2, .25)	2.299938	2.299955	.063	2.7 %
3	(.2, .25, .2, .25)	2.316404	2.316425	.083	3.6 %
4	(.1, .15, .2, .25)	2.252328	2.252416	.149	6.6 %
5	(.14, .16, .18, .20)	2.846528	2.846578	.071	2.5 %
6	(.04, .06, .08, .10)	2.103492	2.103497	.047	2.3 %
7	(.01, .02, .03, .04)	2.037369	2.037371	.021	1.0 %
8	(.01, .02, .01, .02)	2.0224361	2.0224364	.047	2.3 %

(Received April 7, 1995.)

REFERENCES

[1] T. Altiok and H. G. Perros: Queueing Networks with Blocking. North Holland, Amsterdam 1989.
[2] V. E. Beneš: A thermodynamic theory of traffic in connecting network. Bell System Technical Journal *VLII* (1963), 3, 567–607.
[3] V. E. Beneš: Mathematical Theory of Connecting Networks and Telephone Traffic. Academic Press, New York 1965.
[4] P. J. Curtois: Decomposability: Queueing and Computer System Application. Academic Press, New York 1977.

[5] P. J. Curtois and P. Semal: Bounds and conditional steady-state distribution in large Markovian and queueing models. In: Teletraffic Analalysis and Computer Performance Evaluation (O. J. Boxma, J.W. Cohen and H. C. Tijms, eds.), North Holland, Amsterdam 1986.
[6] A. Hordijk and A. Ridder: Insensitive bounds for the stationary distribution on non-reversible Markov chains. J. Appl. Probab. *25* (1988), 9–20.
[7] A. Hordijk and N. M. van Dijk: Networks of queues. Part I: Job-local balance and the adjoint process. Part II: General routing and service characteristics. Lecture Notes in Control and Inform. Sci. *60* (1983), 158–205.
[8] F. P. Kelly: Reversibility and Stochastic Networks. Wiley, New York 1979.
[9] J. G. Kemeny, J. L. Snell and A. W. Knapp: Denumerable Markov Chains. Van Nostrand, Princeton N. J. 1966.
[10] C. D. Meyer: The condition of finite Markov chain and perturbation bounds for the limiting probabilities. SIAM J. Algebraic Discrete Methods *1* (1980), 273–283.
[11] R. R. Muntz, E. De Souza e Silva and A. Goyal: Bounding availability of repairable computer systems. Performance Evaluation *17* (1989), 29–38.
[12] J. K. Muppala and K. S. Trivedi: Numerical Transient Solution of Finite Markovian Systems. Research Report, Duke University 1990.
[13] R. Nelson and L. Kleinrock: Rude-CSMA: A multihop channel access protocol. IEEE Trans. Comm. *COM 33* (1985), 8, 785–791.
[14] E. Pinsky and Y. Yemini: A statistical mechanics of some interconnection networks. In: Performance'84, Elsevier, North Holland, Amsterdam 1984.
[15] P. J. Schweitzer: Perturbation theory and undiscounted Markov renewal programming. Oper. Res. *17* (1969), 716–727.
[16] E. Seneta: Sensitivity to perturbation of the stationary distribution: Some refinements. Linear Algebra Appl. *108* (1988), 121–126.
[17] W. J. Stewart (ed.): Numerical Solutions of Markov Chains. Marcel Dekker, New York 1990.
[18] H. C. Tijms: Stochastic Modelling and Analysis. Wiley, New York 1986.
[19] J. van der Wal and P. J. Schweitzer: Iterative bounds on the equilibrium distribution of a finite Markov chain. Probab. in Engng. Inform. Sci. *1* (1987), 117–131.
[20] N. M. van Dijk: On the importance of bias terms for error bounds and comparison results. In: Numerical Solutions of Markov Chains (W. J. Stewart, ed.), Marcel Dekker, New York 1991, pp. 618–647.
[21] N. M. van Dijk: Approximate uniformization for continuous-time Markov chains with an application to performability analysis. Stochastic Process. Appl. *40* (1992), 339–357.
[22] N. M. van Dijk and M. L. Puterman: Perturbation theory for Markov reward processes with applications to queueing. Adv. in Appl. Probab. *20* (1988), 78–99.
[23] N. M. van Dijk and J. P. Veltkamp: Product form for stochastic interference systems. Probab. in Engng. Inform. Sci. *2* (1988), 355–376.
[24] J. P. Veltkamp: Cost Functions for Interference Systems and VLSI High-level Synthesis. Ph. D. Thesis. Twente University 1988.
[25] B. S. Yoon and B. S. Shanthikumar: Bounds and approximations for the transient behaviour of continuous-time Markov chains. Probab. in Engng. Inform. Sci. *3* (1989), 175–109.

Prof. Dr. Nico M. van Dijk, Faculty of Economical Sciences and Econometrics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam. The Netherlands.