UNIFORMIZATION FOR λ-POSITIVE MARKOV CHAINS

Richard J. Boucherie
Department of Econometrics, Universiteit van Amsterdam
Roetersstraat 11, 1018 WB Amsterdam, The Netherlands
boucheri@butler.fee.uva.nl

Erik A. van Doorn
Faculty of Applied Mathematics, University of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands
doorn@math.utwente.nl

Abstract. We show that a variant of the uniformization method for the numerical calculation of transient state probabilities of positive recurrent Markov chains, in which the truncation error can be made arbitrarily small uniformly over time, can be extended to λ-positive Markov chains.

Keywords: uniformization, truncation error, λ-positive recurrence

1 INTRODUCTION

The *uniformization* or *randomization method*, introduced by Jensen in [10], is a technique for obtaining transient state probabilities of continuous-time Markov chains. The method relates the evolution of a continuous-time Markov chain to the evolution of a Markov jump chain in which the jump epochs are generated by a Poisson process and the jumps themselves are determined by an appropriate discrete-time Markov chain. The uniformization method has been used and extended by several authors, e.g., Gross and Miller [7] and Melamed and Yadin [13].

It was already pointed out by Jensen [10] (see also van den Hout [9, Ch. 3]) that the truncation error in the numerical evaluation of the transient state distribution via the uniformization method can be made arbitrarily small uniformly over time by appropriate use of the stationary distribution of the chain, if it exists. This note shows that a similar result can also be obtained if the chain is *not* positive recurrent, provided it is λ-positive.

In Section 2 we review some results on uniformization, and describe the role of the stationary distribution. We recall some relevant facts from Markov-chain theory in Section 3. Our variant of the uniformization method for the calculation of state probabilities of λ-positive Markov chains is presented in Section 4, and we discuss two examples in Section 5.

2 PRELIMINARIES

We consider a continuous-time Markov chain $\mathcal{X} \equiv \{X(t),\ t \geq 0\}$ taking values in a finite or countably infinite state space S. The chain is assumed to be irreducible and its q-matrix $Q \equiv (q_{ij},\ i,j \in S)$ is assumed to be stable, but not necessarily conservative. If Q is not conservative, so that $\sum_{j\in S} q_{ij} < 0$ for some $i \in S$, it is possible for the chain to leave S to an ignored absorbing state, that is, an absorbing state not included in S. When S is countably infinite, we will also suppose that the chain is *uniformizable*, that is,

$$\sup_{i\in S}\{-q_{ii}\} < \infty.$$

We write $p_i(t) \equiv \Pr\{X(t) = i\}$ and $p_{ij}(t) \equiv \Pr\{X(t) = j|\ X(0) = i\}$, and let $P(.) \equiv (p_{ij}(.),\ i,j \in S)$ be the transition matrix of the chain $\mathcal{X}$. For any sequence $(a_i,\ i \in S)$ we denote the row vector with components $a_i,\ i \in S$, by $\mathbf{a}$. In particular, $\mathbf{0}$ and $\mathbf{1}$ denote the row vectors consisting of 0's and 1's, respectively, and $\mathbf{p}(t)$ is the row vector representing the state probabilities $p_i(t),\ i \in S$, of $\mathcal{X}$ at time t. A superscript $^\top$ denotes transpose, so that $\mathbf{a}^\top$ is the column vector with elements $a_i,\ i \in S$. In what follows $\|\ .\ \|$ denotes the ℓ_∞-norm, that is, $\|A\| = \sup_i \sum_j |a_{ij}|$ for any matrix $A \equiv (a_{ij})$, and, in particular, $\|\mathbf{a}\| = \sum_{i\in S} |a_i|$.

The uniformizability of $\mathcal{X}$ implies that its transition matrix $P(.)$ is the unique solution of the Kolmogorov forward equations

$$P'(t) = P(t)Q, \quad t > 0, \tag{1}$$

see Anderson [1, Proposition 2.9]. Hence $P(t)$ is given by

$$P(t) = \mathrm{e}^{Qt} \equiv \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!}, \quad t \geq 0.$$

Alternatively, see [1, Proposition 2.10], the transition matrix can be written as

$$P(t) = \mathrm{e}^{-qt} \sum_{n=0}^{\infty} \frac{(qt)^n}{n!} R^n, \quad t \geq 0, \tag{2}$$

where $R \equiv (r_{ij},\ i,j \in S)$ is the transition probability matrix defined by

$$R \equiv I + \frac{1}{q} Q, \tag{3}$$

and q is some number satisfying

$$q \geq q_i \equiv -q_{ii}, \quad i \in S, \tag{4}$$

while it will also be convenient to let $q > q_i$ for at least one state $i \in S$. Thus for a given initial distribution $\mathbf{p}(0)$ we have

$$\mathbf{p}(t) = \mathbf{p}(0)P(t) = \mathrm{e}^{-qt} \sum_{n=0}^{\infty} \frac{(qt)^n}{n!} \mathbf{p}(0)R^n, \quad t \geq 0. \tag{5}$$

The argument leading to the representation (5) (or (2)) is known as the uniformization method, see [10], [16], [6], or [18, Sect. 8.2].

For numerical evaluation, the summation in (5) must be truncated at level N, say. That is, we approximate $\mathbf{p}(t)$ as

$$\mathbf{p}(t) \approx \mathbf{p}^{(N)}(t) \equiv \mathrm{e}^{-qt} \sum_{n=0}^{N} \frac{(qt)^n}{n!} \mathbf{p}(0)R^n, \quad t \geq 0. \tag{6}$$

When Q is conservative (and hence R stochastic), the approximation (6) has a truncation error

$$\|\mathbf{p}(t) - \mathbf{p}^{(N)}(t)\| = \mathrm{e}^{-qt} \sum_{n=N+1}^{\infty} \frac{(qt)^n}{n!} \|\mathbf{p}(0)R^n\| = \mathrm{e}^{-qt} \sum_{n=N+1}^{\infty} \frac{(qt)^n}{n!}.$$

As a consequence, for any $t \geq 0$ the approximation (6) converges to $\mathbf{p}(t)$ as $N \to \infty$, but

$$\lim_{N\to\infty} \sup_{t\geq 0} \|\mathbf{p}(t) - \mathbf{p}^{(N)}(t)\| = \lim_{N\to\infty} \sup_{t\geq 0} \; \mathrm{e}^{-qt} \sum_{n=N+1}^{\infty} \frac{(qt)^n}{n!} = 1, \tag{7}$$

so convergence is not uniform in t. The approximation performs badly for fixed N and large enough t.

A variant of the approximation (6) which does not have this drawback – but requires that $\mathcal{X}$ be positive recurrent (and hence Q conservative) – was suggested in [10] (see also [9]). To explain this variant, let $\mathbf{p}$ be the row vector representing the stationary distribution $(p_i, \; i \in S)$ of $\mathcal{X}$, so that $\mathbf{p}Q = \mathbf{0}$. By definition of R we then also have $\mathbf{p}R = \mathbf{p}$. Thus $\mathbf{p}$ constitutes the stationary distribution of the (irreducible) Markov chain with transition probability matrix R, which is aperiodic since we have assumed $q > q_i$ for some $i \in S$. As a consequence, $\mathbf{p}(0)R^n \to \mathbf{p}$ componentwise as $n \to \infty$, but, since $\|\mathbf{p}(0)R^n\| = \|\mathbf{p}\| = 1$ for all n, we actually have

$$\|\mathbf{p}(0)R^n - \mathbf{p}\| \to 0 \quad \text{as} \quad n \to \infty. \tag{8}$$

These observations lead to the representation

$$P(t) = \mathbf{1}^\top \mathbf{p} + \mathrm{e}^{-qt} \sum_{n=0}^{\infty} \frac{(qt)^n}{n!} \left(R^n - \mathbf{1}^\top \mathbf{p} \right), \quad t \geq 0, \tag{9}$$

and hence to the approximation

$$\mathbf{p}(t) \approx \tilde{\mathbf{p}}^{(N)}(t) \equiv \mathbf{p} + \mathrm{e}^{-qt} \sum_{n=0}^{N} \frac{(qt)^n}{n!} \left(\mathbf{p}(0)R^n - \mathbf{p} \right), \quad t \geq 0, \tag{10}$$

with a truncation error satisfying

$$\lim_{N\to\infty} \sup_{t\geq 0} \|\mathbf{p}(t) - \tilde{\mathbf{p}}^{(N)}(t)\| = 0, \tag{11}$$

so that the truncation level N can be chosen such that the approximation has a specified accuracy for all $t \geq 0$ simultaneously. For a detailed discussion see [9, Ch. 3].

When the Markov chain $\mathcal{X}$ is *not* positive recurrent, and hence has no stationary distribution, a similar approach is still possible, provided the chain is λ-*positive*. Before describing this approach, we must introduce, in the next section, some further concepts and results, see, e.g., Anderson [1] for more details.

3 λ-POSITIVE MARKOV CHAINS

Kingman [12] has shown that there exists a real number $\lambda \equiv \lambda_{\mathcal{X}}$, called the *decay parameter of* $\mathcal{X}$, such that

$$0 \leq \lambda \leq \inf_{i \in S} \{q_i\}, \tag{12}$$

and for each pair $i, j \in S$,

$$\lim_{t \to \infty} \frac{1}{t} \log p_{ij}(t) = -\lambda.$$

If $\lambda > 0$ the chain must be transient and is called *exponentially ergodic.* When S is finite $-\lambda$ is simply the smallest (in modulus) eigenvalue of Q, so that $\lambda = 0$ when $\mathcal{X}$ is conservative and $\lambda > 0$ otherwise.

A state $i \in S$ is said to be λ-*recurrent* if

$$\int_0^\infty \mathrm{e}^{\lambda t} p_{ii}(t) dt = \infty,$$

and λ-transient otherwise. A λ-recurrent state is said to be λ-*positive* (or λ-*positive recurrent*) if

$$\lim_{t \to \infty} \mathrm{e}^{\lambda t} p_{ii}(t) > 0.$$

Clearly, λ-(positive) recurrence and λ-transience reduce to (positive) recurrence and transience, respectively, when $\lambda = 0$.

It can be shown that λ-recurrence and λ-positive recurrence are class properties, and hence, in our setting, are properties of either all or no states in S. Accordingly, we shall call the chain $\mathcal{X}$ itself λ-(positive) recurrent, or λ-transient, whichever applies. When S is finite then either $\mathcal{X}$ is conservative and hence positive recurrent (that is, λ-positive with $\lambda = 0$), or $\mathcal{X}$ is non-conservative and hence λ-positive with $\lambda > 0$. So a finite chain is always λ-positive.

A sequence $(m_i,\ i \in S)$ of strictly positive numbers such that for some $\mu \geq 0$

$$\sum_{i \in S} m_i q_{ij} = -\mu m_j, \tag{13}$$

for all $j \in S$, is called a μ-*invariant measure (for Q).* A sequence $(v_i,\ i \in S)$ of strictly positive numbers such that for some $\mu \geq 0$

$$\sum_{j \in S} q_{ij} v_j = -\mu v_i, \tag{14}$$

for all $i \in S$, is called a μ-*invariant vector (for Q).* We shall have use for the following result, which holds true as a consequence of Kingman [12, Theorem 4] and Tweedie [19, Proposition 2]; see also Pollett [15].

Theorem 1 *If the Markov chain $\mathcal{X}$ is λ-recurrent then there exist, up to constant multiples, a unique λ-invariant measure $(m_i,\ i \in S)$ and a unique λ-invariant vector $(v_i,\ i \in S)$; the chain is λ-positive if and only if $\sum_{i\in S} m_i v_i < \infty$.*

4 UNIFORMIZATION FOR λ-POSITIVE CHAINS

We return to the context of Section 2, but we will now assume that the Markov chain $\mathcal{X}$ is λ-positive. As a consequence there are, up to constant multiples, a unique λ-invariant measure $(m_i,\ i \in S)$ and a unique λ-invariant vector $(v_i,\ i \in S)$, which, when represented by row vectors, are denoted by $\mathbf{m}$ and $\mathbf{v}$, respectively. Moreover, whether S is finite or countably infinite, $\mathbf{m}\mathbf{v}^\top = \sum_{i\in S} m_i v_i < \infty$, by Theorem 1.

Following Pollett [15] we define $\bar{Q} \equiv (\bar{q}_{ij})$ by

$$\bar{Q} = \lambda I + V^{-1} Q V, \tag{15}$$

where V is the diagonal matrix with diagonal element v_i in row i, $i \in S$. The matrix $\bar{Q}$ is called the λ-*dual of* Q *with respect to* $\mathbf{v}$, and is actually a uniformizable and conservative (and hence regular) q-matrix. Indeed, by (12) and (15) we have

$$-\bar{q}_{ii} = q_i - \lambda \geq 0, \quad i \in S,$$

while

$$\sum_{j\in S} \bar{q}_{ij} = \lambda + v_i^{-1} \sum_{j\in S} q_{ij} v_j = 0, \quad i \in S.$$

Moreover, $\bar{Q}$ being regular, it is the q-matrix of a positive recurrent Markov chain, since the measure $(u_i,\ i \in S)$ defined by

$$u_i \equiv \frac{m_i v_i}{\mathbf{m}\mathbf{v}^\top}, \quad i \in S,$$

constitutes a probability distribution over S, satisfying

$$\sum_{i\in S} u_i \bar{q}_{ij} = \frac{1}{\mathbf{m}\mathbf{v}^\top} \left(\lambda m_j v_j + v_j \sum_{i\in S} m_i q_{ij} \right) = 0, \quad j \in S.$$

Choosing some positive number $\bar{q}$ such that

$$\bar{q} \geq \bar{q}_i \equiv -\bar{q}_{ii} = q_i - \lambda, \quad i \in S, \tag{16}$$

with inequality for at least one state $i \in S$, and defining

$$\bar{R} \equiv I + \frac{1}{\bar{q}} \bar{Q} = \left(1 + \frac{\lambda}{\bar{q}}\right) I + \frac{1}{\bar{q}} V^{-1} Q V,$$

we can employ the uniformization method to represent $\bar{P}(.)$, the transition matrix of the Markov chain corresponding to $\bar{Q}$, as

$$\bar{P}(t) = \mathrm{e}^{-\bar{q}t} \sum_{n=0}^{\infty} \frac{(\bar{q}t)^n}{n!} \bar{R}^n, \quad t \geq 0. \tag{17}$$

From [15] or [1] we know that the transition matrices $P(.)$ and $\bar{P}(.)$ are related through

$$P(t) = \mathrm{e}^{-\lambda t} V \bar{P}(t) V^{-1}, \quad t \geq 0. \tag{18}$$

Hence, writing

$$T \equiv V \bar{R} V^{-1} = \left(1 + \frac{\lambda}{\bar{q}}\right) I + \frac{1}{\bar{q}} Q, \tag{19}$$

we get

$$P(t) = \mathrm{e}^{-(\lambda+\bar{q})t} \sum_{n=0}^{\infty} \frac{(\bar{q}t)^n}{n!} T^n, \quad t \geq 0;$$

see also van Dijk and Sladky [4], who obtain a similar representation in a more general setting. Thus, for a given initial distribution $\mathbf{p}(0)$, we have

$$\mathbf{p}(t) = \mathbf{p}(0) P(t) = \mathrm{e}^{-(\lambda+\bar{q})t} \sum_{n=0}^{\infty} \frac{(\bar{q}t)^n}{n!} \mathbf{p}(0) T^n, \quad t \geq 0,$$

which leads to the approximation

$$\mathbf{p}(t) \approx \mathbf{p}^{(N)}(t) \equiv \mathrm{e}^{-(\lambda+\bar{q})t} \sum_{n=0}^{N} \frac{(\bar{q}t)^n}{n!} \mathbf{p}(0) T^n, \quad t \geq 0. \tag{20}$$

The approximation (20) is essentially equivalent to the approximation (6) in the sense that for each $\bar{q}$ satisfying (16) there is a q satisfying (4), such that the right-hand side of (20) equals the right-hand side of (6), and vice versa. Indeed, if we let $q = \bar{q} + \lambda$ (so that $q \geq q_i$ for all $i \in S$), and define R as in (3), then $\bar{q}T = qR$, which, upon substitution in (20), gives us (6). The converse is proven similarly.

On the other hand, the Markov chain with q-matrix $\bar{Q}$ is positive recurrent with stationary probabilities $u_i,\ i \in S$, so we can use the analogue of (9) and write, instead of (17),

$$\bar{P}(t) = \mathbf{1}^\top \mathbf{u} + \mathrm{e}^{-\bar{q}t} \sum_{n=0}^{\infty} \frac{(\bar{q}t)^n}{n!} \left(\bar{R}^n - \mathbf{1}^\top \mathbf{u}\right), \quad t \geq 0.$$

Noting that

$$V \mathbf{1}^\top \mathbf{u} V^{-1} = \frac{\mathbf{v}^\top \mathbf{m}}{\mathbf{m} \mathbf{v}^\top}, \tag{21}$$

it subsequently follows from (18) that $P(t)$ can be represented alternatively as

$$P(t) = \mathrm{e}^{-\lambda t}\left(\frac{\mathbf{v}^\top \mathbf{m}}{\mathbf{m}\mathbf{v}^\top} + \mathrm{e}^{-\bar{q}t}\sum_{n=0}^{\infty}\frac{(\bar{q}t)^n}{n!}\left(T^n - \frac{\mathbf{v}^\top \mathbf{m}}{\mathbf{m}\mathbf{v}^\top}\right)\right), \quad t \geq 0.$$

Hence, assuming that the initial distribution $\mathbf{p}(0)$ is such that $\mathbf{p}(0)\mathbf{v}^\top < \infty$ (e.g., by having finite support) and writing

$$\alpha \equiv \frac{\mathbf{p}(0)\mathbf{v}^\top}{\mathbf{m}\mathbf{v}^\top}, \tag{22}$$

we get

$$\mathbf{p}(t) = \alpha\mathrm{e}^{-\lambda t}\mathbf{m} + \mathrm{e}^{-(\lambda+\bar{q})t}\sum_{n=0}^{\infty}\frac{(\bar{q}t)^n}{n!}\left(\mathbf{p}(0)T^n - \alpha\mathbf{m}\right), \quad t \geq 0,$$

which leads to the approximation

$$\mathbf{p}(t) \approx \tilde{\mathbf{p}}^{(N)}(t) \equiv \alpha\mathrm{e}^{-\lambda t}\mathbf{m} + \mathrm{e}^{-(\lambda+\bar{q})t}\sum_{n=0}^{N}\frac{(\bar{q}t)^n}{n!}\left(\mathbf{p}(0)T^n - \alpha\mathbf{m}\right), \quad t \geq 0. \tag{23}$$

Remark. Rather than starting the argument leading to the approximation (23) with the λ-dual of Q with respect to $\mathbf{v}$, one can also choose as a starting point the *λ-reverse of Q with respect to* $\mathbf{m}$ (see [15] for a definition and developments).

The main point we wish to make in this paper is that *when* λ, $\mathbf{m}$ and $\mathbf{v}$ are available, then the approximation (23) is better in some sense than the approximation (20) (or (6)), a statement which is made more precise in the next theorem. To be able to distinguish between the two approximations (20) and (23) we now use the norm $\|\ .\ \|_V$, defined in terms of the usual ℓ_∞-norm by $\|A\|_V \equiv \|AV\|$, so that

$$\|\mathbf{a}\|_V \equiv \|\mathbf{a}V\| = \sum_{i\in S}|a_i v_i|,$$

for any row vector $\mathbf{a} \equiv (a_i,\ i \in S)$.

Theorem 2 *If* $\mathbf{p}(0)\mathbf{v}^\top < \infty$*, then the approximation* (20) *has the property*

$$\lim_{N\to\infty}\sup_{t\geq 0}\ \mathrm{e}^{\lambda t}\|\mathbf{p}(t) - \mathbf{p}^{(N)}(t)\|_V = \mathbf{p}(0)\mathbf{v}^\top > 0, \tag{24}$$

while the approximation (23) *satisfies*

$$\lim_{N\to\infty}\sup_{t\geq 0}\ \mathrm{e}^{\lambda t}\|\mathbf{p}(t) - \tilde{\mathbf{p}}^{(N)}(t)\|_V = 0. \tag{25}$$

Proof. We let $\mathbf{u}(0) \equiv \mathbf{p}(0)V/\mathbf{p}(0)\mathbf{v}^\top$ and note that $\mathbf{u}(0)$ represents an honest distribution over S, so that $\|\mathbf{u}(0)\bar{R}^n\| = 1$ for all n. Hence,

$$\begin{aligned}
e^{\lambda t}\|\mathbf{p}(t) - \mathbf{p}^{(N)}(t)\|_V &= e^{-\bar{q}t}\|\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\mathbf{p}(0)T^n\|_V \\
&= e^{-\bar{q}t}\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\|\mathbf{p}(0)V\bar{R}^nV^{-1}\|_V \\
&= \mathbf{p}(0)\mathbf{v}^\top e^{-\bar{q}t}\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\|\mathbf{u}(0)\bar{R}^n\| \\
&= \mathbf{p}(0)\mathbf{v}^\top e^{-\bar{q}t}\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!},
\end{aligned}$$

from which (24) immediately follows.

To prove (25) we note that $\mathbf{u}R = \mathbf{u}$ and

$$\|\mathbf{u}(0)\bar{R}^n - \mathbf{u}\| \to 0 \quad \text{as} \quad n \to \infty, \tag{26}$$

by analogy with (8). By (19), (21) and (22) we subsequently get

$$\begin{aligned}
e^{\lambda t}\|\mathbf{p}(t) - \tilde{\mathbf{p}}^{(N)}(t)\|_V &= e^{-\bar{q}t}\|\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\left(\mathbf{p}(0)T^n - \alpha\mathbf{m}\right)\|_V \\
&= e^{-\bar{q}t}\|\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\mathbf{p}(0)V\left(\bar{R}^n - \mathbf{1}^\top\mathbf{u}\right)V^{-1}\|_V \\
&= \mathbf{p}(0)\mathbf{v}^\top e^{-\bar{q}t}\|\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\mathbf{u}(0)\left(\bar{R}^n - \mathbf{1}^\top\mathbf{u}\right)\| \\
&\leq \mathbf{p}(0)\mathbf{v}^\top e^{-\bar{q}t}\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\|\mathbf{u}(0)\bar{R}^n - \mathbf{u}\| \\
&= \mathbf{p}(0)\mathbf{v}^\top e^{-\bar{q}t}\sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!}\|\left(\mathbf{u}(0)\bar{R}^N - \mathbf{u}\right)\bar{R}^{n-N}\| \\
&\leq \mathbf{p}(0)\mathbf{v}^\top\|\mathbf{u}(0)\bar{R}^N - \mathbf{u}\|,
\end{aligned}$$

since $\bar{R}$ is a stochastic matrix. The statement now follows with (26). □

The transformation (19) and the resulting approximation (23) are natural generalizations of the transformation (3) and the approximation (10), respectively. Indeed, when $\lambda = 0$ we can choose the parameter $\bar{q}$ in (19) equal to the parameter q in (3) and obtain $T = R$, while (23) is easily seen to reduce to (10) since, after proper normalization, $\mathbf{v} = \mathbf{1}$ and $\mathbf{m} = \mathbf{p}$. Note also that $R\mathbf{1}^\top = \mathbf{1}^\top$ and $\mathbf{p}R = \mathbf{p}$ generalize to $T\mathbf{v}^\top = \mathbf{v}^\top$ and $\mathbf{m}T = \mathbf{m}$, respectively.

When $\lambda = 0$, the relations (24) and (25) reduce to (7) and (11), respectively, since then the V-norm is just our usual ℓ_∞-norm. In the general case $\lambda \geq 0$, we must use the V-norm rather than the ℓ_∞-norm, since $\|\mathbf{m}\|$ may be infinite. Then, however, the v_i's are sufficiently small for the V-norm to be effective. Indeed, $\|\mathbf{m}\|_V = \mathbf{m}\mathbf{v}^\top < \infty$, by Theorem 1.

The preceding suggests that when $\|\mathbf{m}\| < \infty$ one might be able to distinguish between the approximations (20) and (23) in terms of the ℓ_∞-norm. Before settling this in the affirmative, we observe that if $\|\mathbf{m}\| < \infty$ and $\lambda > 0$, then Q is necessarily non-conservative and escape from S must be certain. Indeed, assuming that Q is conservative and summing (13) with $\mu \equiv \lambda > 0$ over j, leads to a contradiction. Moreover, we know from [19, Proposition 2] that

$$\sum_{i\in S} m_i p_{ij}(t) = \mathrm{e}^{-\lambda t} m_j, \quad j \in S.$$

Summing over j and letting $t \to \infty$, it subsequently follows that $\sum_{j\in S} p_{ij}(t) \to 0$ as $t \to \infty$ for any state i, and hence $\|\mathbf{p}(t)\| \to 0$ as $t \to \infty$ for any initial distribution, that is, escape from S is certain.

Theorem 3 *The the approximation* (20) *has the property*

$$\lim_{N\to\infty} \sup_{t\geq 0} \|\mathbf{p}(t) - \mathbf{p}^{(N)}(t)\| = 1, \tag{27}$$

while, if $\mathbf{p}(0)\mathbf{v}^\top < \infty$ *and* $\|\mathbf{m}\| < \infty$*, the approximation* (23) *satisfies*

$$\lim_{N\to\infty} \sup_{t\geq 0} \|\mathbf{p}(t) - \tilde{\mathbf{p}}^{(N)}(t)\| = 0. \tag{28}$$

Proof. Since the approximations (20) and (6) are essentially equivalent, (27) is a restatement of (7).

When $\lambda = 0$ the relation (28) reduces to (25), so in the remainder of this proof we will assume $\lambda > 0$. It will be convenient to define

$$r \equiv \frac{\bar{q}}{\lambda + \bar{q}} < 1, \tag{29}$$

and

$$\Pi \equiv I + \frac{1}{\lambda + \bar{q}} Q,$$

so that Π is a (strictly) substochastic transition probability matrix. Since our assumptions $\|\mathbf{m}\| < \infty$ and $\lambda > 0$ imply that the chain $\mathcal{X}$ will escape from S with probability 1, the same must be true for the discrete-time chain generated by Π, that is,

$$\|\mathbf{p}(0)\Pi^n\| \to 0 \ \text{ as } \ n \to \infty, \tag{30}$$

for any initial distribution $\mathbf{p}(0)$. Subsequently recalling that $\mathbf{m}T = \mathbf{m}$ and noting that $\Pi = rT$, we get

$$\begin{aligned}
\|\mathbf{p}(t) - \tilde{\mathbf{p}}^{(N)}(t)\| &= \mathrm{e}^{-(\lambda+\bar{q})t} \| \sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!} \left(\mathbf{p}(0)T^n - \alpha\mathbf{m}\right) \| \\
&= \mathrm{e}^{-(\lambda+\bar{q})t} \sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!} \| \left(\mathbf{p}(0)T^N - \alpha\mathbf{m}\right) T^{n-N} \| \\
&\leq \|\mathbf{p}(0)T^N - \alpha\mathbf{m}\| \mathrm{e}^{-(\lambda+\bar{q})t} \sum_{n=N+1}^{\infty} \frac{(\bar{q}t)^n}{n!} r^{N-n} \\
&= \|\mathbf{p}(0)\Pi^N - \alpha r^N \mathbf{m}\| \mathrm{e}^{-(\lambda+\bar{q})t} \sum_{n=N+1}^{\infty} \frac{((\lambda+\bar{q})t)^n}{n!} \\
&\leq \|\mathbf{p}(0)\Pi^N\| + \alpha r^N \|\mathbf{m}\|,
\end{aligned}$$

so that the result follows by (29), (30) and the fact that $\|\mathbf{m}\|$ is finite. □

When Q is non-conservative it is often of interest to know the state distribution for $\mathcal{X}$ conditional on the chain being in S. That is, we are interested in $\boldsymbol{\pi}(t) \equiv (\pi_i(t),\ i \in S)$ with

$$\pi_i(t) \equiv \frac{p_i(t)}{\sum_{j \in S} p_j(t)}, \quad i \in S,\ t \geq 0. \tag{31}$$

It may be shown on the basis of Theorem 4.2, that when $\|\mathbf{m}\| < \infty$, and escape from S is certain, approximating both numerator and denominator of (31) by using (23), and truncating at level N, results in an approximation $\tilde{\boldsymbol{\pi}}^N(t)$ which satisfies

$$\lim_{N\to\infty} \sup_{t \geq 0} \mathrm{e}^{-\lambda t} \|\boldsymbol{\pi}(t) - \tilde{\boldsymbol{\pi}}^{(N)}(t)\| = 0.$$

As an aside we remark that when $\boldsymbol{\pi} \equiv (\pi_i,\ i \in S)$ is an honest distribution over S such that

$$\pi_i = \pi_i(t), \quad i \in S,\ t \geq 0,$$

then $\boldsymbol{\pi}$ is called a *quasi-stationary distribution*. Actually, when $\mathbf{m}$ is a summable λ-invariant measure, then $\mathbf{m}/\|\mathbf{m}\|$ *is* a quasi-stationary distribution (see Nair and Pollett [14]), while it is usually also the limit as $t \to \infty$ of the conditional state distribution $\boldsymbol{\pi}(t)$.

Obviously, the applicability of the approximation (23) is considerably restricted by the fact that one needs to know the decay parameter λ, and the λ-invariant measure $\mathbf{m}$ and λ-invariant vector $\mathbf{v}$. Some examples of Markov chains for which this knowledge is available are given in the next section.

An approximation of $\tilde{\mathbf{p}}^{(N)}(t)$, and hence of $\mathbf{p}(t)$, which requires knowledge of λ, but *not* of $\mathbf{m}$ and $\mathbf{v}$, may be obtained by generalising the *steady-state-detection method*; see [9] and the references mentioned there. First note that $\bar{R}^n \to \mathbf{1}^\top \mathbf{u}$, and hence $T^n \to \mathbf{v}^\top \mathbf{m}/\mathbf{m}\mathbf{v}^\top$ componentwise as $n \to \infty$. As a consequence we have

$$\mathbf{p}(0)T^n \to \frac{\mathbf{p}(0)\mathbf{v}^\top \mathbf{m}}{\mathbf{m}\mathbf{v}^\top} = \alpha \mathbf{m} \quad \text{as} \quad n \to \infty, \tag{32}$$

componentwise, since we have assumed $\mathbf{p}(0)\mathbf{v}^\top < \infty$. The approximation now amounts to choosing N so large that $\mathbf{p}(0)T^N$ (or at least a relevant part of it) has converged in sufficient degree according to a suitable criterion which does not involve its limit $\alpha\mathbf{m}$), and replacing $\alpha\mathbf{m}$ in (23) by $\mathbf{p}(0)T^N$. A disadvantage of this method is that no exact error bounds are available, even in the ergodic case.

5 EXAMPLES

5.1 Example 1: The birth-death process

We let $\mathcal{X}$ be a birth-death process on $S = \mathbb{N}_0 \equiv \{0, 1, \ldots\}$ with birth and death rates b_i and d_i, respectively. All rates are assumed to be positive, except $d_0 \geq 0$; if $d_0 > 0$, there is an ignored absorbing state -1, say, which can only be reached via state 0. We also assume uniformizability, that is, $\sup_{i \in S}\{b_i + d_i\} < \infty$.

With λ denoting the decay parameter of $\mathcal{X}$, it is not difficult to see that the λ-invariant measure $(m_i,\ i \in S)$ and λ-invariant vector $(v_i,\ i \in S)$ for $\mathcal{X}$ may be represented as

$$m_i = \pi_i Q_i(\lambda), \quad v_i = Q_i(\lambda), \quad i \in S, \tag{33}$$

where $Q_i,\ i \in S$, are polynomials defined by the recurrence relation

$$\begin{aligned} &b_i Q_{i+1}(x) = (b_i + d_i - x)Q_i(x) - d_i Q_{i-1}(x), \quad i > 0, \\ &b_0 Q_1(x) = b_0 + d_0 - x, \quad Q_0(x) = 1, \end{aligned} \tag{34}$$

and $\pi_i,\ i \in S$, are constants given by

$$\pi_0 \equiv 1, \quad \pi_i \equiv \frac{b_0 b_1 \cdots b_{i-1}}{d_1 d_2 \cdots d_i}, \quad i > 0,$$

see, e.g., [11].

Assuming that the chain is λ-positive, that is,

$$\gamma^{-1} \equiv \sum_{i \in S} m_i v_i = \sum_{i \in S} \pi_i Q_i^2(\lambda) < \infty,$$

the transition probabilities $p_{ij}(t),\ i,j \in S,\ t \geq 0$, can now be represented as

$$\begin{aligned} p_{ij}(t) &= \gamma\, \mathrm{e}^{-\lambda t} \pi_j Q_i(\lambda) Q_j(\lambda) \\ &\quad + \mathrm{e}^{-(\bar{q}+\lambda)t} \sum_{n=0}^{\infty} \frac{(\bar{q}t)^n}{n!} \left((T^n)_{ij} - \gamma \pi_j Q_i(\lambda) Q_j(\lambda)\right), \end{aligned} \tag{35}$$

where $\bar{q}$ is any number satisfying $\bar{q} \geq b_i + d_i - \lambda,\ i \in S$, and $T \equiv (t_{ij},\ i,j \in S)$ is the matrix with elements

$$t_{ij} = \begin{cases} \dfrac{b_i}{\bar{q}} & j = i+1,\ i \geq 0 \\ 1 + \dfrac{\lambda - b_i - d_i}{\bar{q}} & j = i,\ i \geq 0 \\ \dfrac{d_i}{\bar{q}} & j = i-1,\ i \geq 1, \end{cases}$$

and 0 otherwise. An approximation which is arbitrarily close to $p_{ij}(t)$ uniformly over t may now be obtained by truncating the infinite sum.

For many specific birth-death processes the decay parameter λ is explicitly known. An example is the process with constant rates

$$b_i = b, \quad d_{i+1} = d, \quad i = 0, 1, \ldots,$$

but $d_0 = \delta < d$. It may be shown that this process is λ-positive if and only if

$$0 \leq \delta < \sqrt{d}(\sqrt{d} - \sqrt{b}),$$

in which case

$$\lambda = \delta \left(1 - \frac{b}{d - \delta}\right),$$

cf. [5], where the discrete-time variant of this process is analysed.

5.2 Example 2: A closed Jackson network

Consider a closed Jackson network modelling a manufacturing system consisting of K stations. The system might fail due to defective routing. That is, a job might be damaged while being transported from one station to another, in which case the system has to be stopped to remove the damaged job.

Let $\mathbf{i} \equiv (i_1, i_2, \ldots, i_K)$ denote the state of the network, i_k being the number of jobs present at station k. When the network is in state $\mathbf{i}$, the service rate in station k, $k = 1, 2, \ldots, K$, is $\mu_k(\mathbf{i})$. As usual in the literature on product form distributions, cf. Boucherie and van Dijk [3] and Henderson and

Taylor [8], we assume that functions $\psi : \mathbb{N}_0^N \to [0,\infty)$, $\phi : \mathbb{N}_0^N \to (0,\infty)$ and $\theta : \{1,2,\dots,K\} \to [0,\infty)$ exist such that

$$\mu_k(\mathbf{i}) = \frac{\psi(\mathbf{i}-\mathbf{e}_k)\theta(k)}{\phi(\mathbf{i})},$$

where $\mathbf{e}_k$ denotes the k-th unit vector. Following Boucherie [2], we will additionally assume that the total service intensity of the network is fixed and normalized to 1, that is, for all states $\mathbf{i}$

$$\sum_{k=1}^{K} \mu_k(\mathbf{i}) = 1.$$

For a network consisting of single-server stations, with (basic) service rate ν_k at station k, in which surplus capacity of idle servers is proportionally shared by the busy servers, these functions are $\theta(k) = \nu_k$, $\psi(\mathbf{i}-\mathbf{e}_k) = \mathbb{1}_{\{\mathbf{i}-\mathbf{e}_k \geq \mathbf{0}\}}$, and $\phi(\mathbf{i}) = \sum_{k=1}^{K} \nu_k \mathbb{1}_{\{\mathbf{i}-\mathbf{e}_k \geq \mathbf{0}\}}$, where $\mathbb{1}_A$ denotes the indicator function of the event A.

Upon leaving station k after completion of service a customer is routed to station $\ell \neq k$ with probability $p_{k\ell}$, and is damaged with probability $1 - \sum_{\ell=1}^{K} p_{k\ell}$, the latter event resulting in termination of network operation. The substochastic matrix $P \equiv (p_{k\ell},\ k,\ell = 1,2,\dots,K)$ is assumed to be primitive. Let β be the Perron-Frobenius eigenvalue of P, and $\mathbf{c}$ and $\mathbf{y}$ the corresponding left and right eigenvectors,

$$\mathbf{c}P = \beta\mathbf{c} \quad \text{and} \quad P\mathbf{y} = \beta\mathbf{y},$$

respectively, normalized such that

$$\sum_{k=1}^{K} c_k y_k/\theta(k) = 1.$$

When the network contains M jobs, the state space of the Markov chain $\mathcal{X}$ recording the state of the network is $S = \{\mathbf{i} \equiv (i_1, i_2, \dots, i_K)|\ \sum_{k=1}^{K} i_k = M\}$, while its transition rates are given by

$$\begin{aligned} q_{\mathbf{i},\mathbf{i}-\mathbf{e}_k+\mathbf{e}_\ell} &= \mu_k(\mathbf{i})p_{k\ell}, \quad k,\ell = 1,2,\dots,K,\ k \neq \ell, \\ q_{\mathbf{i},\mathbf{i}} &= -1. \end{aligned}$$

for $\mathbf{i} \in S$. The failure rate in state $\mathbf{i}$ is $\sum_{k=1}^{K} \mu_k(\mathbf{i})(1 - \sum_{\ell=1}^{K} p_{k\ell})$.

By analogy with Boucherie [2], it is easy to see that $\mathcal{X}$ has a $(1-\beta)$-invariant measure

$$m_{\mathbf{i}} = \phi(\mathbf{i}) \prod_{k=1}^{K} \left(c_k/\theta(k)\right)^{i_k}, \quad \mathbf{i} \in S,$$

and a $(1-\beta)$-invariant vector

$$v_{\mathbf{i}} = \prod_{k=1}^{K} y_k^{i_k}, \quad \mathbf{i} \in S.$$

Obviously,

$$\gamma^{-1} \equiv \sum_{\mathbf{i}\in S} m_{\mathbf{i}} v_{\mathbf{i}} < \infty,$$

so from the continuous-time analogue of [17, Theorem 6.4] we obtain that $\mathcal{X}$ has decay parameter $\lambda = 1-\beta$ and λ-invariant measure $\mathbf{m} = (m_{\mathbf{i}},\ \mathbf{i} \in S)$ and vector $\mathbf{v} = (v_{\mathbf{i}},\ \mathbf{i} \in S)$.

For a system starting off empty, i.e., $p_{\mathbf{i}}(0) = 1$ for $\mathbf{i} = \mathbf{0}$, the state probabilities $p_{\mathbf{i}}(t)$ can now be represented as

$$p_{\mathbf{i}}(t) = \gamma\, \mathrm{e}^{-(1-\beta)t} m_{\mathbf{i}} + \mathrm{e}^{-(1-\beta+\bar{q})t} \sum_{n=0}^{\infty} \frac{(\bar{q}t)^n}{n!} \left((T^n)_{\mathbf{0},\mathbf{i}} - \gamma m_{\mathbf{i}}\right), \quad \mathbf{i} \in S,$$

where $\bar{q}$ is any number satisfying $\bar{q} \geq \beta$, and $T \equiv (t_{\mathbf{i},\mathbf{j}},\ \mathbf{i},\mathbf{j} \in S)$ is the matrix with elements

$$t_{\mathbf{i},\mathbf{j}} = \begin{cases} \dfrac{\mu_k(\mathbf{i}) p_{k\ell}}{\bar{q}} & \mathbf{j} = \mathbf{i} - \mathbf{e}_k + \mathbf{e}_\ell, \quad k,\ell = 1,2,\ldots,K,\ k \neq \ell, \\ 1 - \dfrac{\beta}{\bar{q}} & \mathbf{j} = \mathbf{i}, \end{cases}$$

and 0 otherwise. An approximation which is arbitrarily close to $p_{\mathbf{i}}(t)$ uniformly over t may now be obtained by truncating the infinite sum.

Acknowledgements. The research of R.J. Boucherie has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. We thank the referees and an editor for their valuable comments, and one referee in particular for pointing out an error in an earlier version of this paper.

References

[1] W.J. Anderson (1991) *Continuous-Time Markov Chains: An Applications Oriented Approach*, Springer-Verlag, New York.

[2] R.J. Boucherie (1997) On the quasi-stationary distribution for queueing networks with defective routing. *J. Austral. Math. Soc. (B)* **38**, 454-463.

[3] R.J. Boucherie and N.M. van Dijk (1991) Product forms for queueing networks with state dependent multiple job transitions. *Adv. Appl. Probab.* **23**, 152-187.

[4] N.M. van Dijk and K. Sladky (1996) A note on uniformization for dynamic nonnegative systems. Manuscript.

[5] E.A. van Doorn and P. Schrijner (1995) Ratio limits and limiting conditional distributions for discrete-time birth-death processes. *J. Math. Anal. Appl.* **190**, 263-284.

[6] W.K. Grassman (1991) Finding transient solutions in Markovian event systems through randomization. pp. 357-371 in: *Numerical Solution of Markov Chains*, W.J. Stewart, ed., Marcel Dekker, New York.

[7] D. Gross and D.R. Miller (1984) The randomization technique as a modeling tool and solution procedure for transient Markov processes. *Oper. Res.* **32**, 343-361.

[8] W. Henderson and P.G. Taylor (1990) Product forms in networks of queues with batch arrival and batch services. *Queueing Systems* **6**, 71-88.

[9] W.B. van den Hout (1996) *The Power-Series Algorithm. A Numerical Approach to Markov Processes.* CentER for Economic Research, Tilburg University, The Netherlands.

[10] A. Jensen (1953) Markoff chains as an aid in the study of Markoff processes. *Skandinavisk Aktuarietidskrift* **36**, 87-91.

[11] M. Kijima, M.G. Nair, P.K. Pollett and E.A. van Doorn (1997) Limiting conditional distributions for birth-death processes. *Adv. Appl. Probab.* **29**, 185-204.

[12] J.F.C. Kingman (1963) The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* (3rd Series) **13**, 337-358.

[13] B. Melamed and M. Yadin (1984) Randomization procedures in the computation of cumulative-time distributions over discrete state Markov processes. *Oper. Res.* **32**, 926-944.

[14] M.G. Nair and P.K. Pollett (1993) On the relationship between μ-invariant measures and quasi-stationary distributions for continuous-time Markov chains. *Adv. Appl. Probab.* **25**, 82-102.

[15] P.K. Pollett (1988) Reversibility, invariance and μ-invariance. *Adv. Appl. Probab.* **20**, 600-621.

[16] A. Reibman and K.S. Trivedi (1988) Numerical transient analysis of Markov models. *Comput. Oper. Res.* **15**, 19-36.

[17] E. Seneta (1981) *Non-negative Matrices and Markov Chains*, 2nd ed. Springer-Verlag, New York.

[18] W.J. Stewart. (1994) *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, Princeton, NJ.

[19] R.L. Tweedie (1974) Some ergodic properties of the Feller minimal process. *Quart. J. Math. Oxford* **25**, 485-495.