Spectral Conditions and Bounds for the Rate of Convergence of countable Markov Chains

F.M. Spieksma[1)2)]

November 17, 1992

Abstract

This paper studies the convergence rate in the M-uniform ergodicity property, which is a matrix norm version of geometric ergodicity, for countable Markov chains in discrete time. By analysing the spectral properties of the transition matrix P (with respect to the associated M-norm) this convergence rate will be connected to the "second largest" eigenvalue of P similarly to the finite state space case.

Under an additional condition on the hitting times of a special state the contraction factor in Popov's criterion for geometric ergodicity is shown to provide an upperbound for this convergence rate.

The results are finally applied to a single server phase-type queue. Explicit expressions for these convergence rates are obtained for a discrete time approximation of the model, which can then be transformed to analogous bounds for the rate of convergence of the continuous time process. For the special case of the M/M/1-queue we also present numerical results on the bounding constants in the geometric and M-uniform ergodicity properties.

Keywords: Markov Chain, geometric ergodicity, spectral analysis, operator norm, rates of convergence, phase-type queues

AMS 1980 Subject Classification: 60J10, 60K25, 47B65

[1)] The research of the author has been supported by the Netherlands Organisation for Scientific Research N.W.O.

[2)] Postal address: Dept of Mathematics and Computer Science, University of Leiden, Niels Bohrweg 1, 2333CA Leiden, The Netherlands. Email: spieksma@rulwi.leidenuniv.nl

1. Introduction

Let $\Xi = \{\Xi_n,\, n \in \mathbb{Z}_+\}$ be a time-homogeneous Markov chain on a countable state space E with n-step transition matrix P^n, where

$$P^n_{ij} = \mathbb{P}\{\Xi_n = j \mid \Xi_0 = i\},\ i,j \in \mathsf{E}$$

and $P^0 = I$ for I the identity matrix. We will assume that Ξ is aperiodic and irreducible. Then the limits $\Pi_{ij} \stackrel{\triangle}{=} \lim_{n\to\infty} P^n_{ij}$ constitute a possibly defective probability distribution that is independent of the initial state of the system, and hence the matrix $\Pi = (\Pi_{ij})_{i,j\in\mathsf{E}}$ has equal rows.

In this paper we will focus on cases, where the n-step transition probabilities converge at a *geometric* rate; that is, when there are constants c_i, $i \in \mathsf{E}$, and $\alpha < 1$, such that for $i,j \in \mathsf{E}$

$$\|P^n_{i\bullet} - \Pi_{i\bullet}\| = \sum_j |P^n_{ij} - \Pi_{ij}| \le c_i\alpha^n,\ n = 0,1,\ldots, \tag{1.1}$$

where $\|\cdot\|$ is the total variation norm. Ξ is said to be *geometrically ergodic*, when (1.1) holds (cf. [10]).

In [11] Popov has shown that geometric ergodicity together with positive recurrence is equivalent to the following *contractive* drift condition: there is a vector $M \ge 1$ on E and a constant $\beta < 1$, such that

$$\sum_j P_{ij}M_j \le \beta M_i, \quad i \ne 0 \tag{1.2a}$$

$$\sum_j P_{0j}M_j < \infty. \tag{1.2b}$$

Hordijk and Spieksma refined this result in [3] and showed that a solution $M \ge 1$ of (1.2) implies *M-uniform ergodicity*[3] for the same vector M, that is

$$\|P^n - \Pi\|_M = \sup_{i\in\mathsf{E}} M_i^{-1} \sum_j |P^n_{ij} - \Pi_{ij}|M_j \le c\cdot\gamma^n,\ n = 0,1,\ldots \tag{1.3a}$$

$$\|P\|_M \le c,' \tag{1.3b}$$

for constants $\gamma < 1$, c and c', where $\|\cdot\|_M$ is the operator norm associated with the weighted supremum norm with weighting vector M. Since (1.3) trivially implies positive recurrence of Ξ, if $M \ge 1$, geometric ergodicity together with positive recurrence are therefore equivalent to M-uniform ergodicity for some $M \ge 1$.

The purpose of this paper is to analyse the convergence rates and bounding constants in (1.1) and (1.3a). Through the characterisation of geometric ergodicity in (1.3) operator theoretic tools become available, and these turn out to be quite fruitful for our analysis.

Section 2 studies the precise relation between the spectral properties of the transition matrix, M-uniform ergodicity and the recurrence conditions introduced in [3], which are also formulated in terms of M-norms. As a corollary it follows that γ is bigger, but can be chosen arbitrarily close to the second largest "eigenvalue"[4] (in absolute value) of the transition matrix P as an M-bounded operator. This extends well-known results (see [13], p.9) for finite Markov chains, although they are slightly weaker. A similar result has been proved in [4] for *uniformly ergodic* Markov chains on a

[3] In [3] this property was named M-geometric ergodicity, but renamed in [7] to establish a more consistent terminology.

[4] This will be defined more rigorously in the next section.

countable state space, but the proofs in the present paper have been complicated by the fact that in general the M-norm of P is bigger than 1, whereas the uniform ergodicity property uses supremum norms and the supremum norm of P clearly equals 1.

As a corollary of Theorem 2.2 we prove in section 2 , that geometric ergodicity is in fact equivalent to 1 being an isolated point of the spectrum of P for some M-norm. This proves a conjecture hinted at by Vere-Jones in [23], who studied the spectrum of a transformed transition matrix with respect to the ℓ_p-norm for some specific queueing examples, but did not further pursue the matrix norm approach taken here.

In section 3 we then relate β in Popov's criterion (1.2) to the second largest "eigenvalue" of P with respect to the associated M-norm. In particular, we can choose $\gamma = \beta$ in (1.3a) under the condition that the only 1-point of $F_{00}(z)$ is achieved by $z = 1$, where $F_{00}(z)$ is the generating function of the first hitting probabilities of state 0, when starting in 0. Thus this result enhances the importance of constructing solutions to Popov's criterion (1.2): not only does it yield the M-vector in the M-uniform ergodicity, but it also provides explicit bounds for corresponding rate of convergence γ, and in some special cases even computable bounds for the constant in (1.3a).

By suitably varying the solutions M to (1.2), it is possible to obtain the sharpest possible rate of convergence α for the unweighted distances of P^n and Π in (1.1) as well as the bounding constant. Our derivations resemble the arguments used by Kendall [6] and Vere-Jones [22] by a similar use of the generating functions of the transition probabilities and first hitting probabilities. However, while our manipulations typically require the same regular structures of the models to be analysed (see [23, 24] for the embedded M/M/1, E_k/M/1 and M/E_k/1-queues), using matrix norms allows far stronger results with respect to associated reward processes.

Although the condition on $F_{00}(z)$ looks fairly restrictive, the results so far suggest that it holds for homogeneous random walk type processes on $\mathbb{Z}_+^n$. Section 4 verifies this condition for a discrete time approximation of the single server phase-type queue, for which our method is ideally suited. By subsequently constructing a solution to Popov's criterion (1.2) with equality, we then obtain closed form expressions for the rates of convergence γ in (1.3) and α in (1.1). These yield the explicit convergence rates for the corresponding continuous time process by applying a solidarity result from [17]. It is not easy to provide closed form expressions for the bounding constants in (1.1) and (1.3) and therefore we will list some numerical results for the special case of the M/M/1-queue by way of illustration. The method extends to other more-dimensional queueing models like the K-competing queues and retrial queues.

All models we have studied, have the left skipfree property, and so they are not strongly ergodic, whilst they are geometrically ergodic indeed (see [16], [18]). This bars obtaining bounds for the convergence rates through δ-coefficients [4]. Due to computational complexity the graph-theoretical bounds developed for finite state Markov chains (see e.g. [8], [14]) do not seem suitable neither for the above models, especially since the bounds in these papers explicitly use the stationary probabilities. Here we leave the problem aside whether the graph-theoretical results extend to countable state space chains. It is an interesting problem how to reformulate the (merging) conductance in these papers in terms of M-weighted flows.

In [2] another approach is taken for the analysis of one-dimensional random walks. The random walk polynomials are exploited to yield a precise description of the transient behaviour of the marginal distributions of the process and obviously yields more exact results for the M/M/1-queue. However, results for the associated reward processes and in particular of the marginal moments do not seem to be included.

2. Spectral Properties of the Transition Matrix

For linear operators on a general state space we need a more refined notion of "eigenvalues". For simplicity we will adapt the general notation to our special situation. So let

$$\mathsf{V}_M \triangleq \{x \in \mathbb{C}^\infty : \|x\|_M = \sup_{i\in\mathsf{E}} \frac{|x_i|}{M_i} < \infty\}$$

be the normed complex Banach space of M-bounded vectors on E, and let similarly M_M be the space of M-bounded matrices on $\mathsf{E}\times\mathsf{E}$. Note, that the associated operator norm is the same whether V_M is a complex or real vector space.

For $A \in \mathsf{M}_M$ the *resolvent* set $\rho_M(A)$ is the set $\{\lambda \in \mathbb{C} : \lambda I - A$ has an M-bounded inverse$\}$. The inverse $(\lambda I - A)^{-1}$ is denoted by $R(\lambda, A)$, which is the *resolvent* of A at λ. Then $R(\lambda, A)$ is an *analytic* matrix function in $\lambda \in \rho_M(A)$ (to M_M). Moreover, the resolvent set is *open*.

$\sigma_M(A) \triangleq (\rho_M(A))^c$ is the so-called spectrum of A. The notion "eigenvalue" used before, was meant to denote elements from the spectrum. The *spectral radius* $r_{\sigma,M}(A)$ is defined as $\lim_{n\to\infty} \|A^n\|_M^{1/n}$. It can be shown, that $r_{\sigma,M}(A) = \sup\{|\lambda| \mid \lambda \in \sigma_M(A)\}$. Moreover, $\sum_{n=0}^\infty \lambda^{-n}A^n$ converges in M-norm for $|\lambda| > r_{\sigma,M}(A)$ and $\|A\|_M \lambda^{-n} \to 0$, for $n \to \infty$. In this case $R(\lambda, A)$ is equal to $\lambda^{-1}\sum_{n=0}^\infty \lambda^{-n}A^n$. For $|\lambda| < r_{\sigma,M}(A)$ the summation diverges and if $A \geq 0$ then $r_{\sigma,M}(A) \in \sigma_M(A)$.

These definitions and results can be found quite extensively in [25], pp. 209-212 and in [5]. Since our derivations use generating functions, we use a slightly different notation. Write

$$A(z) \triangleq \sum_{n=0}^\infty A^n z^n$$

with i,j^{th} element $A_{ij}(z)$. Let further

$$\mathcal{D}_{x,r} \triangleq \{z \in \mathbb{C} \mid |z-x| < r\}$$

$$\mathcal{C}_{x,r} \triangleq \{z \in \mathbb{C} \mid |z-x| = r\}.$$

Then clearly $R(z^{-1}, A) = zA(z)$ and $\|A\|_M z^n \to 0$, for $n \to \infty$, for $z \in \mathcal{D}_{0,1/r_{\sigma,M}(A)}$. All other statements are similarly reformulated as statements on generating functions.

As a first step towards the characterisation of the convergence rate in the M-uniform ergodicity property we will generalise Cauchy's inequality to power series of matrices (cf. [19], p. 84).

Lemma 2.1 *Let G_n, $n = 0,1,\ldots$ be a sequence of matrices on $\mathsf{E}\times\mathsf{E}$, such that $G(z) = \sum_{n=0}^\infty G_n z^n \in \mathsf{M}_M$ for $z \in \overline{\mathcal{D}}_{0,R}$. If $U(r)$ is an upperbound of $\|G(z)\|_M$ on the circle $\mathcal{C}_{0,r}$, $r \leq R$, then $\|G_n\|_M \leq U(r)r^{-n}$.*

Proof $G_{ij}(z)$ is a power series for $z \in \mathcal{D}_{0,R}$ and consequently analytic in this region. Let $r \in (0,R)$. By virtue of Cauchy's integral theorem

$$\begin{aligned} G_{n,ij} &= \tfrac{1}{2\pi i}\oint_{\mathcal{C}_{0,r}} z^{-n-1}G_{ij}(z)dz \\ &= \tfrac{1}{2\pi i} r^{-n}\int_0^{2\pi} ie^{-ni\phi}G_{ij}(re^{i\phi})d\phi,\ n \in \mathbb{N}_0, \end{aligned}$$

where $G_{n,ij}$ denotes the i,j^{th} element of G_n. Consequently

$$\sum_{j\in\mathsf{E}} |G_{n,ij}|M_j \leq \tfrac{1}{2\pi}\sum_{j\in\mathsf{E}} r^{-n}\int_0^{2\pi} |e^{-ni\phi}|\cdot|G_{ij}(re^{i\phi})|d\phi\cdot M_j$$

$$= \frac{1}{2\pi} r^{-n} \int_0^{2\pi} \sum_{j\in\mathsf{E}} |G_{ij}(re^{i\phi})| M_j d\phi$$
$$\le r^{-n} U(r) M_i, \quad n \in \mathbb{N}_0,$$

for $i \in \mathsf{E}$, using the Fubini-Tonelli theorem (see [12], p. 307-309) to justify the last equality. The statement for $r = R$ follows directly by taking the limit $r \uparrow R$. **QED**

Before proceeding to the main result in this section, we will need some more notation. In order to allow matrix operations we define the taboo probabilities slightly different from [1], pp. 45, 46, namely

$${}_S P_{ij} \triangleq \begin{cases} P_{ij}, \, j \notin S \\ 0, \quad j \in S, \end{cases}$$

if $S \subset \mathsf{E}$ is the taboo set. Then the first hitting probabilities of state j are defined as follows:

$$f_{ij}^n \triangleq \mathbb{P}\{\Xi_n = j, \Xi_k \neq j, k = 1, \ldots, n-1 \mid \Xi_0 = i\} = \sum_k {}_j P_{ik}^{n-1} P_{kj}, \; n \ge 1,$$

where ${}_jP^n$ is the n^{th} iterate of ${}_jP$ and ${}_jP^0 = I$. For the matrix generating function $F(z)$ of these first hitting probabilities we then have

$$F_{ij}(z) \triangleq \sum_{n=1}^{\infty} f_{ij}^n z^n = z \sum_k {}_j P_{ik}(z) P_{kj}. \tag{2.1}$$

The following well-known relations between $F_{ij}(z)$ and $P_{ij}(z)$ hold for $z \in \mathcal{D}_{0,1}$ ([1], p. 55)

$$P_{jj}(z) = \frac{1}{1 - F_{jj}(z)} \tag{2.2}$$
$$P_{ij}(z) = F_{ij}(z) P_{jj}(z), \tag{2.3}$$

and each of these functions will be used to analytically extend the other outside the open unit disc.

Next we will investigate the spectral properties of P and the relation with the spectrum of $(P - \Pi)$.

THEOREM 2.2 *Let $M \ge 1$ be a vector on E and suppose that $P \in \mathsf{M}_M$.*
i) If $\mathcal{D}_{1,r} \setminus \{1\} \subset \rho_M(P)$ for some real $r > 0$, then
a) $z = 1$ is a simple pole of the resolvent $R(z, P)$ and hence Ξ is positive recurrent.
b) $\exists R < 1$, such that $\sigma_M(P) \setminus \{1\} \subset \overline{\mathcal{D}}_{0,R}$.
ii) For any $R < 1$ we have

$$\sigma_M(P - \Pi) \subset \overline{\mathcal{D}}_{0,R} \text{ and } \Pi \in \mathsf{M}_M \iff \sigma_M(P) \setminus \{1\} \subset \overline{\mathcal{D}}_{0,R}.$$

Theorem 2.2 then implies the following interesting conclusions.

COROLLARY 2.3: *If Ξ is M-uniformly ergodic, then $r_{\sigma_M(P-\Pi)} = \beta_M(P)$, where*

$$\beta_M(P) \triangleq \sup\{|z| \mid z \in \sigma_M(P), \, z \neq 1\}.$$

Consequently, any number $\gamma > \beta_M(P)$ is a feasible convergence rate in (1.3a).

Corollary 2.4: *Geometric ergodicity (1.1) together with positive recurrence is equivalent to the existence of a vector $M \geq 1$ on E, such that 1 is an isolated point of the spectrum $\sigma_M(P)$ of P with respect to the associated M-norm.*

Note, that Corollary 2.3 is a weaker statement than the finite state space version ([13], p. 9) in that $\beta_M(P)$ itself is not in general a feasible rate of convergence in (1.3a). However, in our queueing examples $\beta_M(P)$ is a feasible rate for the unweighted differences in (1.1) of the marginal and the stationary probabilities.

Finally Theorem 2.5 fully characterises the recurrence conditions developed in [3] via spectral properties. Roughly speaking, these recurrence conditions require the recurrence times to state 0 to have exponentially decaying tails in the M-norm. For bounded M they are equivalent to the Doeblin condition (see e.g. [7] for a detailed discussion). Combination of Theorems 2.2 and 2.5 then establishes the equivalence of these recurrence conditions and M-uniform ergodicity, thus providing an alternative way of proving the Key Theorem in [3].

Theorem 2.5 *Let $M \geq 1$ be a vector on E and suppose that $P \in \mathsf{M}_M$. The following are equivalent*
i) *1 is an isolated point of the spectrum $\sigma_M(P)$.*
ii) *There are constants $\delta < 1$ and c, such that $\|{}_0P^n\|_M \leq c \cdot \delta^n$, $n = 0, 1, 2, \ldots$.*
iii) *There are a vector $\widetilde{M}$ on E and a constant $\beta < 1$ solving Popov's criterion (1.2), such that $c_1 M \leq \widetilde{M} \leq c_2 M$ for some positive constants c_1, c_2.*

In [15], Chapter IV, we show that Theorems 2.2 and 2.5 (and hence the subsequent corollaries) continue to hold, if the irreducibility assumption is relaxed to allowing a finite number of closed classes and a set of transient states.

The proof of the statements will occupy the remainder of this section. In essence it consists of constructing analytic extensions of $F(z)$ or $P(z)$ as M-bounded matrix functions, as opposed to (and based on) Vere-Jones [22], whose proofs reduce to constructing *elementwise* analytic extensions.

Proof of Theorem 2.2 First we show i). If $\mathcal{D}_{1,r} \backslash \{1\} \subset \rho_M(P)$, then necessarily

$$r_{\sigma,M}(P) \leq 1. \tag{2.4}$$

For suppose, that $r_{\sigma,M}(P) = c > 1$. Then $c \in \sigma_M(P)$, since P is a positive operator. $xP(x)$ is the (possibly unbounded in M-norm) inverse of $(x^{-1}I - P)$ for positive, real $x < 1$. By assumption $c^{-1}P(c^{-1}) \notin \mathsf{M}_M$, and since $xP(x)$ increases in x, necessarily $xP(x) \notin \mathsf{M}_M$ for $x \in [c^{-1}, 1)$. Consequently $x^{-1} \in \sigma_M(P)$ for these values of x. However, by assumption $x^{-1} \in \rho_M(P)$ for $x \neq 1$ sufficiently close to 1. Hence $r_{\sigma,M}(P) > 1$ is not possible.

Next we show, that the spectral radius equals 1, i.e. $1 \in \sigma_M(P)$. Suppose on the contrary, that $1 \in \rho_M(P)$. Then the resolvent $R(z^{-1}, P)$ is analytic in $z^{-1} \in \mathcal{D}_{1,r}$ as an M-bounded matrix function. (2.4) implies that $R(x^{-1}, P) = xP(x)$ for $x < 1$, and, because of the analyticity of the resolvent we find that $R(1, P) = \lim_{x\uparrow 1} xP(x) = P(1)$ by virtue of the monotone convergence theorem ([12], p. 265).

We conclude that $\sum_{n=0}^{\infty} P^n \in \mathsf{M}_M$. Because the P_{ij}^n are non-negative, $\sum_j P_{ij}^n M_j \to 0$, as n tends to infinity. Since $M \geq 1$, this contradicts the stochasticity of P and hence 1 is an isolated point of the spectrum.

By virtue of Dunford's integral theorem (cf. [25], pp. 228-230) this implies, that the Laurent expansion of $R(z^{-1}, P)$ exists in $z = 1$. As $\lim_{x\uparrow 1}(1 - x)P_{ij}(x) = \Pi_{ij}$,[6] $z = 1$ can at most be a

[6] The Abel-limit and the limit are equal, if the limit exists.

simple pole. If Ξ is not positive recurrent, then the Laurent expansion is a Taylor expansion. This contradicts the fact that $1 \in \sigma_M(P)$, and thus the proof of ia) is completed.

For the proof of ib) we first list some consequences of the fact that $z = 1$ is a simple pole of the resolvent. By Dunford's integral theorem there are $A-1$, $A_0 \in \mathsf{M}_M$, such that

$$R(z^{-1}, P) = \frac{z}{1-z} A_{-1} + \sum_{n=0}^{\infty} \Big(\frac{z-1}{z}\Big)^n A_0^{n+1}, \quad \Big|\frac{z-1}{z}\Big| < \| A_0 \|_M^{-1}, \tag{2.5}$$

where $A_0 A_{-1} = A_{-1} A_0 = 0$, $(P - I)A_0 = A_{-1} - I$ and $PA_0 = A_0 P = A_0$. By the M-boundedness of A_{-1} and A_0 we can similarly show that $A_0(P - I) = A_{-1} - I$ by multiplying (2.5) to the right by $z^{-1}I - P$ and by subsequently setting $z = 1$. Clearly $A_{-1} = \Pi$.

As a consequence $(1-z)z^{-1}R(z^{-1}, P)$ is an analytic extension of $(1-z)P(z)$ as an M-bounded matrix function in a disc, say $\mathcal{D}_{1,r'}$. r' can be chosen small enough to satisfy $(1-z)P_{00}(z) \geq \epsilon > 0$ for all $z \in \mathcal{D}_{1,r'}$, for some positive ϵ, because $(1-z)P_{00}(z) \to \Pi_{00} > 0$ as $z \to 1$. This allows the following analytic extension of $F_{i0}(z)$ in $\mathcal{D}_{1,r'}$ (cf. (2.2), (2.3))

$$F_{i0}(z) = \frac{(1-z)P_{i0}(z)}{(1-z)P_{00}(z)}, \quad i \neq 0 \tag{2.6}$$

$$F_{00}(z) = 1 - \frac{1-z}{(1-z)P_{00}(z)}. \tag{2.7}$$

Since the Taylor expansion of $F_{i0}(z)$ in 0 has non-negative coefficients, standard arguments on the analytic extension of a power series with non-negative coefficients then suffice to show, that in fact (2.1) holds for z in the larger disc $\mathcal{D}_{0,1+r'}$. For $z \in \mathcal{D}_{1,r'}$ we have $|P_{i0}(z)| \leq c_z M_i$ for some constant c_z depending on z, so that the vector $F_{\bullet 0}(z)$ with i^{th} component $F_{i0}(z)$, $i \in \mathsf{E}$, has a finite M-norm for z in this region. The non-negativity of the Taylor coefficients of $F_{i0}(z)$ finally implies that $F_{\bullet 0}(z)$ has a finite M-norm for $z \in \mathcal{D}_{0,1+r'}$.

The final step is to construct the resolvent $R(z^{-1}, P)$ for all $z \in \mathcal{D}_{0,R'} \setminus \{1\}$ and some $R' > 1$. To this end we use first-entrance decomposition ([1], p. 46) to state 0, i.e. for $n \geq 1$

$$P_{ij}^n = {}_0P_{ij}^n + \sum_{m=1}^{n} f_{i0}^m P_{0j}^{n-m}.$$

This yields the following relation for $z \in \mathcal{D}_{0,1}$

$$P_{ij}(z) = {}_0P_{ij}(z) + F_{i0}(z)P_{0j}(z), \tag{2.8}$$

with ${}_0P_{ij}(z) = \sum_{n=0}^{\infty} {}_0P_{ij}^n z^n$. Extend $P(z)$ as a meromorphic M-bounded matrix function in $\mathcal{D}_{1,r'}$ through (2.5), i.e.

$$P(z) = z^{-1}R(z^{-1}, P) = \frac{1}{1-z}\Pi + \frac{1}{z}\sum_{n=0}^{\infty}\Big(\frac{z-1}{z}\Big)^n A_0^{n+1},$$

while plugging in $A_{-1} = \Pi$. Denote the analytic part of the foregoing expression by $D(z)$, then ${}_0P_{ij}(z)$ is analytically extensible in $\mathcal{D}_{1,r'}$ as follows

$$\begin{aligned} {}_0P_{ij}(z) &= P_{ij}(z) - F_{i0}(z)P_{0j}(z) \qquad (2.9) \\ &= \frac{F_{i0}(z) - F_{i0}(1)}{z-1}\Pi_{0j} + D_{ij}(z) - F_{i0}(z)D_{0j}(z) \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} F_{i0}^{(n)}(1)(z-1)^{n-1}\Pi_{0j} + D_{ij}(z) - F_{i0}(z)D_{0j}(z), \qquad (2.10) \end{aligned}$$

where $F_{i0}^{(n)}(z)$ is the n^{th} derivative of $F_{i0}(z)$. Since ${}_0P_{ij}^n$ are all non-negative, we conlude similarly to case of $F_{i0}(z)$, that ${}_0P_{ij}(z)$ equals its power series expansion $\sum_{n=0}^{\infty} {}_0P_{ij}^n z^n$ on the whole disc $\mathcal{D}_{0,1+r'}$. Note further, that the first summation in the righthandside of (2.10) is an M-bounded vector for each $z \in \mathcal{D}_{0,r'}$, again using the non-negativity of the $F_{i0}^{(n)}(1)$ and the M-boundedness of $F_{\bullet 0}(z)$ in this region. Since $D(z) \in \mathsf{M}_M$ for $z \in \mathcal{D}_{1,r'}$ and $\Pi \in \mathsf{M}_M$, we conclude that ${}_0P(z) \in \mathsf{M}_M$ for $z \in \mathcal{D}_{1,r'}$, and, because of the non-negativity of ${}_0P$, hence for $z \in \mathcal{D}_{0,1+r'}$.

This gives us the necessary tools for constructing the resolvent of P. Because Ξ is aperiodic, $F_{00}(z) = 1$ on $\overline{\mathcal{D}}_{0,1}$ in $z = 1$ only. By continuity arguments there must be a disc $\mathcal{D}_{0,R'} \subset \mathcal{D}_{0,1+r'}$, for some $R' > 1$, such that $z = 1$ is the only 1-point of $F_{00}(z)$ in this larger disc. By applying last-exit decomposition ([1], p.46) to (2.8), it is easily seen that $O(z)$ given by

$$O_{ij}(z) \triangleq z \cdot {}_0P_{ij}(z) + z\frac{F_{i0}(z) {}_0P_{0j}(z)}{1 - F_{00}(z)} \tag{2.11}$$

is the resolvent of P at z^{-1} for $z \in \mathcal{D}_{0,1}$. By our previous argumentation $O(z) \in \mathsf{M}_M$ for $z \in \mathcal{D}_{0,R'}$ and it is easily checked that $(z^{-1}I - P)O(z) = O(z)(z^{-1}I - P) = I$. Consequently $O(z) = R(z^{-1}, P)$ for $z \in \mathcal{D}_{0,R'}$. Note that the interchange of the order of summation necessary for proving this, is allowed by the Fubini-Tonelli theorem and the proper M-boundedness of all operators involved (e.g. by using that $\|{}_0P^n\|_M \le c|x|^{-n}$ for $x > |z|$ and some constant c depending on x). The conclusion in ib) follows.

For the proof of ii) first assume that $\sigma_M(P - \Pi) \subset \overline{\mathcal{D}}_{0,R}$ for some $R < 1$ and $\Pi \in \mathsf{M}_M$. We have to show, that $z \in \mathcal{D}_{0,R^{-1}}$ implies $z^{-1} \in \rho_M(P)$, for $z \neq 0, 1$.

By assumption $R(z^{-1}, P - \Pi) = z\sum_{n=0}^{\infty}(P - \Pi)^n z^n$ for $z \in \mathcal{D}_{0,R^{-1}}$. Using the identity

$$P(z) = \sum_{n=0}^{\infty}(P^n - \Pi)z^n + \frac{\Pi}{1 - z}, \tag{2.12}$$

for $z \in \mathcal{D}_{0,1}$, we define for $z \in \mathcal{D}_{0,R^{-1}}$ and $z \neq 1$,

$$\widehat{O}(z) \triangleq z\sum_{n=0}^{\infty}(P - \Pi)^n z^n + \frac{z^2}{1 - z}\Pi.$$

Clearly $\widehat{O}(z) \in \mathsf{M}_M$ and it is straightforward to verify, that $\widehat{O}(z)$ is the resolvent of P at z^{-1}.

For the reverse statement, first note that $r_{\sigma,M}(P)$ must be 1 by a similar reasoning as for the proof of ia), whence $z = 1$ is a simple pole of the resolvent and $\Pi \in \mathsf{M}_M$. For $z = 1$ we have to use (2.5) to construct the required resolvent. In this case

$$\widetilde{O}(1) \triangleq \Pi + A_0$$

is easily verified to be the resolvent $R(1, P - \Pi)$. The assumptions further imply, that $(1 - z)z^{-1}R(z^{-1}, P)$ is an analytic extension of $(1 - z)P(z)$ as an M-bounded matrix function in $\mathcal{D}_{0,R^{-1}}$. This extension, which we will denote as $(1 - z)P(z)$, therefore has a Taylor expansion in $\mathcal{D}_{0,R^{-1}}$ and since the (elementwise) coefficients of the Taylor expansion are unique, we must have $(1 - z)P(z) = I + \sum_{n=1}^{\infty}(P^n - P^{n-1})z^n$ for all z in the disc $\mathcal{D}_{0,R^{-1}}$. It is standard by now to check, that for $z \in \mathcal{D}_{0,R^{-1}}$, $z \neq 1$, the matrix $\widetilde{O}(z) \in \mathsf{M}_M$ defined by

$$\widetilde{O}(z) \triangleq \frac{z}{1 - z}I + \frac{z}{1 - z}\sum_{n=1}^{\infty}(P^n - P^{n-1})z^n - \frac{z^2}{1 - z}\Pi$$

is the resolvent of $P - \Pi$ at z^{-1}. By the fact that the resolvent is analytic as an M-bounded matrix function, Lemma 2.1 yields, that $\|P^n - P^{n-1}\|_M \leq cx^{-n}$ for some $x > |z|$ and some constant c depending on x. Together with the M-boundedness of the relevant operators the necessary interchange arguments are justified, similarly as before. This completes the proof of Theorem 2.2. **QED**

Corollary 2.3 then follows immediately from Theorem 2.2, by observing that the spectrum is closed, so that $\beta_M(P) \in \sigma_M(P)$. For Corollary 2.4 note, that geometric ergodicity is equivalent to M-uniform ergodicity for some $M \geq 1$. **QED**

Proof of Theorem 2.5 From the proof of that theorem follows, that i) implies $\sum_{n=0}^{\infty} {}_0P^n \in \mathsf{M}_M$. We may now apply Lemma 3.1 in [3] to conclude ii). Moreover, the equivalence of ii) and iii) follows analogously to this same Lemma 3.1. By a direct construction of the Laurent series expansion (2.5) it is possible to show that ii) implies i). This construction reduces to first proving that $\Pi \in \mathsf{M}_M$ (cf. [3], Eq. (3.9)) and then (see [15], Chapter IV.3) that the matrix A_0 given by

$$A_0 \triangleq (I - \Pi) \sum_{n=0}^{\infty} {}_0P^n (I - \Pi)$$

is the unique M-bounded matrix with $\Pi \cdot A_0 = A_0 \cdot \Pi = 0$ and $(I-P)A_0 = A_0(I-P) = I - \Pi$. Setting $A_{-1} = \Pi$ and using A_0 defined in this way, it is straightforward to show that the Laurent expansion given in (2.5) is indeed the resolvent of P at z^{-1} for $z^{-1} \in \mathcal{D}_{1,r} \setminus \{1\}$ and some $r > 0$. **QED**

3. The Connection between the second largest Eigenvalue and Popov's Criterion

In this section an important part is played by the generating functions of the first hitting times, similarly to the derivations in the previous section. In particular, by iteration it can be verified that the vector M given by $M_i = F_{i0}(\beta^{-1})$, $i \neq 0$, and $M_0 = 1$, is the *minimal solution* to Popov's criterion (1.2) with components greater or equal to 1, for the contraction factor β. Moreover, these functions solve (1.2a) with *equality* (cf. [21]).

For our results to hold, we need the following assumption. Let $\mathcal{D}$ denote the disc of convergence of the Taylor expansion of $F_{00}(z)$ in $z = 0$; $\mathcal{D}$ equals $\mathcal{D}_{0,R}$ or $\overline{\mathcal{D}}_{0,R}$, for R denoting the radius of convergence, depending on whether $F_{00}(R)$ diverges or not.

Assumption A $F_{00}(z)$ achieves the value 1 on $\mathcal{D}$ in $z = 1$ only.

Theorem 3.1 *Suppose that Assumption A holds. Let further the vector M on E and the constant $\beta < 1$ solve Popov's criterion (1.2) and write $\tilde{c} = \max\{1, \sum_j {}_0P_{0j} M_j\}$. Then*

i) $\beta_M(P) \leq \beta$.

ii) Ξ *is M-uniformly ergodic with rate of convergence $\gamma = \beta + \delta$ and bounding constant $c = c_{\beta,\delta}$, where*

$$c_{\beta,\delta} = \frac{1}{1-\gamma} M_0 + \tilde{c}\gamma \left[\frac{1}{\delta} + \frac{1}{(1-\gamma)(1-\beta)}\right] + \sup\left\{\left|\frac{1}{1 - F_{00}(z)}\right| : z \in \mathcal{C}_{0,\gamma^{-1}}\right\} \left[M_0 + \frac{\gamma \tilde{c}}{\delta}\right],$$

for any $\delta < 1 - \beta$.

iii) If $F_{00}(R)$ is finite, i.e. $\mathcal{D} = \overline{\mathcal{D}}_{0,R}$, then Ξ is geometrically ergodic (1.1) with rate of convergence $\alpha = R^{-1}$ and this is also the sharpest possible convergence rate. As the bounding constant we *may take*

$$c_i = \Big[1 + \sup\{\big|\frac{1}{1-F_{00}(z)}\big| : z \in \mathcal{C}_{0,R}\}\big(F_{00}(R)-1\big)\Big]\frac{F_{i0}(R)}{R-1}.$$

Obviously one can't avoid having the impression, that the bounds c_δ and c_i have a mere academic value. Even though this is true in general, they can be calculated for some of our queueing examples (see section 4).

Many of our queueing examples satisfy a so-called (see [7]) M-left skipfree property, i.e. for all i, $j \in \mathsf{E}$

$$P_{ij} > 0 \Rightarrow M_j \geq M_i - m$$

for some integer $m \geq 1$. This property allows to refine the result in Theorem 3.1 considerably.

THEOREM 3.2 *Suppose that Ξ is a Markov chain satisfying Assumption A. Let the vector M on E and the constant $\beta < 1$ solve Popov's criterion (1.2) with equality, such that M is unbounded off finite sets. In addition to the conclusions of Theorem 3.1, $\beta = \beta_M(P)$ if Ξ satisfies the M-left skipfree property. Hence Popov's criterion gives the sharpest possible convergence rate in the M-uniform ergodicity property. Moreover, $\beta_M(P) \in \sigma_M(P)$, i.e. the second largest eigenvalue is positive, real.*

Proof of Theorem 3.1 For the proof of i) it is sufficient to show that $(1-z)P(z)$ can be extended as an M-bounded operator in the disc $\mathcal{D}_{0,\beta^{-1}}$. Indeed, this will give us the resolvent of P at z^{-1} for z in this region, and i) then will follow from Theorem 2.2ii).

By applying combined first-entrance and last-exit decomposition to state 0, we derive similarly to (2.8) and (2.11), that

$$P_{ij}(z) = {}_0P_{ij}(z) + \frac{F_{i0}(z)\,{}_0P_{0j}(z)}{1-F_{00}(z)}, \tag{3.1}$$

for $z \in \mathcal{D}_{0,1}$. By assumption

$$\sum_j {}_0P_{ij}M_j \leq \beta M_i,\ i \neq 0.$$

Iterating this expression gives

$$\sum_j {}_0P^n_{ij}M_j \leq \beta^n M_i,\ i \neq 0, \tag{3.2}$$

so that for $z \in \mathcal{D}_{0,\beta^{-1}}$

$$\sum_j \Big|\sum_{n=0}^{\infty} {}_0P^n_{ij}z^n\Big| M_j \leq \sum_{n=0}^{\infty}\beta^n|z|^n M_i = \frac{1}{1-\beta|z|}M_i,\quad i \neq 0. \tag{3.3}$$

For initial state $i = 0$,

$$\begin{aligned}\sum_j \Big|\sum_{n=0}^{\infty} {}_0P^n_{0j}z^n\Big| M_j &\leq M_0 + \sum_k {}_0P_{0k}\Big|\sum_j\sum_{n=0}^{\infty} {}_0P^n_{kj}z^n\Big| M_j \leq M_0 + \frac{1}{1-\beta|z|}\sum_k {}_0P_{0k}M_j \\ &\leq M_0 + \frac{\tilde{c}}{1-\beta|z|},\end{aligned} \tag{3.4}$$

by using (3.3) in the second inequality. Combining (3.3) and (3.4) shows that ${}_0P(z) \in \mathsf{M}_M$ for $z \in \mathcal{D}_{0,\beta^{-1}}$, and hence $z^{-1} \in \rho_M({}_0P)$ for z in this region.

Since $F_{i0}(\beta^{-1})$ is the minimal solution to (1.2) and the coefficients in the Taylor expansion of $F_{i0}(z)$ in $z = 0$ are non-negative, we find that $F_{i0}(z) \le M_i$ for $z \in \mathcal{D}_{0,\beta^{-1}}$ for $i \neq 0$. Together with Assumption A we finally conclude, that the matrix function with i, j^{th} element

$$(1-z)_0P_{ij}(z) + \frac{1-z}{1-F_{00}(z)} F_{i0}(z)_0P_{0j}(z), \tag{3.5}$$

is an M-bounded matrix extension of $(1-z)P(z)$ for $z \in \mathcal{D}_{0,\beta^{-1}}$. Mark, that $z = 1$ is a simple pole, because $F_{00}^{(1)}(1) \ge F_{00}(1) = 1 > 0$. In particular, using (2.12) the matrix function with i, j^{th} element $\sum_{n=0}^{\infty}(P_{ij}^n - \Pi_{ij})z^n$ can be analytically extended in this region, through

$$\sum_{n=0}^{\infty}(P_{ij}^n - \Pi_{ij})z^n \triangleq {}_0P_{ij}(z) + \frac{1}{1-F_{00}(z)} F_{i0}(z)_0P_{0j}(z) - \frac{\Pi_{ij}}{1-z}, \quad i, j \in \mathsf{E}. \tag{3.6}$$

Observe that by virtue of Theorem 2.2ii) the lefthandside of (3.6) is analytic as an M-bounded operator in $z = 1$. This proves i). The bounding constant $c_{\beta,\delta}$ in ii) follows directly from (3.6) and Lemma 2.1 by using the bounds (3.3) and (3.4) for $z \in \mathcal{C}_{0,(\beta+\delta)^{-1}}$, by M-bounding Π as follows

$$\|\Pi\|_{M} \le \sum_j \Pi_{0j} M_j = \Pi_{00} \sum_{n=0}^{\infty} \sum_j {}_0P_{0j}^n M_j \le M_0 + \frac{\tilde{c}}{1-\beta},$$

where in the last equality we use (3.4) for $z = 1$, and finally by observing that

$$\sup\{\left|\frac{1}{1-z}\right| : z \in \mathcal{C}_{0,(\beta+\delta)^{-1}}\} = \frac{1}{(\beta+\delta)^{-1} - 1} = \frac{\beta+\delta}{1-\beta-\delta}.$$

Next we will prove iii). Because Ξ is irreducible, finiteness of $F_{00}(R)$ implies, that $F_{i0}(R)$ and ${}_0P_{0j}(R)$ are finite. Consequently, the righthandside of (3.6) is finite for $z \in \mathcal{C}_{0,R}$ and all $i, j \in \mathsf{E}$. This remains true of we take the supremum of the absolute value of (3.6) over the set $\mathcal{C}_{0,R}$. Then applying Cauchy's integral theorem (of which Lemma 2.1 is the matrix version) yields

$$|P_{ij}^n - \Pi_{ij}| \le R^{-n}\Big[{}_0P_{ij}(R) + \sup\{\left|\frac{1}{1-F_{00}(z)}\right| : z \in \mathcal{C}_{0,R}\} \cdot F_{i0}(R)_0P_{0j}(R) + \frac{\Pi_{ij}}{R-1}\Big]. \tag{3.7}$$

As

$$\begin{aligned}
\sum_j {}_0P_{ij}(R) &= \sum_{j\in\mathsf{E}} \sum_{n=0}^{\infty} {}_0P_{ij}^n R^n = \delta_{i0} + \sum_{n=1}^{\infty} \mathbb{P}\{\Xi_1, \ldots, \Xi_n \neq 0 \mid \Xi_0 = i\} R^n \\
&= \delta_{i0} + \sum_{n=1}^{\infty} R^n \sum_{m=n+1}^{\infty} f_{i0}^m \\
&= \frac{F_{i0}(R) - 1}{R-1},
\end{aligned} \tag{3.8}$$

with δ denoting the Kronecker delta, we find in conjuction with (3.7) that $\alpha = R^{-1}$ is a feasible rate of convergence for the geometric ergodicity property (1.1) and the bounding constant c_i is given as in the statement of the theorem.

To show that R^{-1} is the sharpest possible convergence rate, we assume on the contrary, that the convergence radius of $\sum_n \sum_j |P_{ij}^n - \Pi_{ij}| z^n$ is bigger than R, say $\widetilde{R}$, for all states i. Through (2.12) this implies in particular, that $(1-z)P_{00}(z)$ has an analytic extension within the larger disc $\mathcal{D}_{0,\widetilde{R}}$. From (3.5) follows that

$$(1-z)P_{00}(R) = \frac{1-R}{1-F_{00}(R)},$$

which is strictly positive. By continuity arguments there is some disc $\mathcal{D}_{R,r} \subset \mathcal{D}_{0,\widetilde{R}}$, such that $(1-z)P_{00}(z) \neq 0$ for all z in this disc. However, then we can use (2.7) to analytically extend $F_{00}(z)$ in $\mathcal{D}_{R,r}$ and by now familiar arguments then show, that the Taylor expansion of $F_{00}(z)$ must converge in the larger disc $\mathcal{D}_{0,R+r}$. This contradicts our assumption that the convergence radius of $F_{00}(z)$ equals R and thus proves iii). **QED**

Proof of Theorem 3.2 Under the conditions of the theorem there is for each n a state, say $i_n \neq 0$, such that

$$\sum_j {}_0P^n_{i_n j} M_j = \beta^n M_{i_n}. \tag{3.9}$$

Let ${}_0\tilde{P}$ denote the taboo probability matrix ${}_0P$ with the 0^{th} row replaced by 0's. (3.9) implies that $\|{}_0\tilde{P}^n\|_M = \beta^n$ and so the spectral radius $r_{\sigma,M}({}_0\tilde{P}) = \lim_{n\to\infty} \|{}_0\tilde{P}^n\|_M^{1/n}$ equals β. Hence $\beta \in \sigma_M({}_0\tilde{P})$, because ${}_0\tilde{P}$ is a positive operator.

If $\beta > \beta_M(P)$, then $(1-z)P(z)$ can be extended as an M-bounded operator in a closed disc, say $\overline{\mathcal{D}}_{\beta^{-1},r}$, centered at $z = \beta^{-1}$ and not containing the point $z = 1$. Let

$$s \stackrel{\triangle}{=} \sup\{\|P(z)\|_M : z \in \overline{\mathcal{D}}_{\beta^{-1},r}\}.$$

s is finite by the analyticity of the resolvent as an M-bounded matrix function and by the fact that $\overline{\mathcal{D}}_{\beta^{-1},r}$ does not contain the point 0. Moreover, $|F_{i0}(z)| \leq M_i$, $i \neq 0$, for $|z| \leq \beta^{-1}$, because of the minimality of $F_{i0}(\beta^{-1})$ as a solution to Popov's criterion for the contraction factor β. Using (2.9) to analytically extend ${}_0\tilde{P}(z)$ in $\mathcal{D}_{\beta^{-1},r}$, we obtain for real $x \in (\beta^{-1}-r, \beta^{-1})$

$$\|{}_0\tilde{P}(x)\|_M \leq s(1+M_0).$$

As ${}_0\tilde{P}(x)$ is monotone increasing in x, we find by virtue of the monotone convergence theorem, that

$$\|{}_0\tilde{P}(\beta^{-1})\|_M \leq s(1+M_0),$$

which contradicts the fact that $\beta \in \sigma_M({}_0\tilde{P})$. Consequently $\beta = \beta_M(P)$ and so $\beta_M(P) \in \sigma_M({}_0\tilde{P})$. Then combination with (3.1) yields $\beta_M(P) \in \sigma_M(P)$. **QED**

T. Ott provided the following simple, but effective counterexample, exemplifying that in general Assumption A need not hold.

COUNTEREXAMPLE This is a Markov chain on two states, 0 and 1, with transition matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

Ξ is aperiodic and irreducible, if $p \neq 0, 1$. For this chain the second largest eigenvalue of P is $-(1-p)$. Moreover,

$$F_{00}(z) = z^2(1-p)\frac{1}{1-pz}$$

and so there are 2 points, 1 and $-\frac{1}{1-p}$, where $F_{00}(z)$ takes the value 1 within its disc of convergence iff $p < 1-p$. In this case the point $-\frac{1}{1-p}$ determines the convergence radius of $D_{ij}(z) \triangleq \sum_n (P_{ij}^n - \Pi_{ij})z^n$.

If $p > 1-p$, then the convergence radius of $F_{00}(z)$ is *smaller* than the convergence radius of $D_{ij}(z)$, so that the $F_{i0}(z)$ do not suffice to analytically extend these functions, Consequently either the assumption that $F_{00}(z)$ also converges on the closure of its disc of convergence, or the left skipfree property of the transition matrix in conjunction with *infinite countability* of the state space, are quite imperious for our results on the sharpest convergence rates to hold.

In constructing solutions to Popov's criterion (1.2) it is important to connect these to the generating functions $F_{i0}(z)$. The next result deals with this problem.

LEMMA 3.3. *Let the vector $M \geq 1$ on E and the contraction factor $\beta < 1$ solve Popov's criterion (1.2). Then for $z \in \mathcal{D}_{0,\beta^{-1}}$, the vector V on E with*

$$\begin{aligned} V_0 &\triangleq 1 \\ V_i &\triangleq F_{i0}(z), \quad i \geq 1 \end{aligned}$$

is the unique M-bounded vector with 0^{th} component equal to 1 solving (1.2) with equality for the "contraction" factor z^{-1}.

As we shall see in section 4 the V need not be the unique solution! In particular M and V may be different vectors for $z^{-1} = \beta$.

Proof Consider the transformed taboo probability matrix ${}_0\tilde{P}$ defined in the proof of Theorem 3.2. Clearly $z^{-1} \in \rho_M({}_0\tilde{P})$ for $z \in \mathcal{D}_{0,\beta^{-1}}$ and so the inverse $R(z^{-1}, {}_0\tilde{P}) = z\sum_{n=0}^{\infty} {}_0\tilde{P}^n z^n$ exists as an M-bounded operator. By rewriting (1.2a), any solution W with 0^{th} component equal to 1 satisfies

$$\big((z^{-1}I - {}_0\tilde{P})W\big)_i = \begin{cases} z^{-1}, & i = 0 \\ P_{i0}, & i \neq 0, \end{cases}$$

and multiplying both sides with $R(z^{-1}, {}_0\tilde{P})$ yields $W = V$. **QED**

4. A single Server phase-type Queue

We will illustrate the results in the previous sections by analysing the M/PH/1-queue. This is a single service facility with an infinite capacity buffer, where jobs arrive according to a Poisson process with rate λ. With probability p_k any job receives k exponential phases of service each with mean ν^{-1}. Then the mean service time μ^{-1} equals $\sum_k p_k k/\nu$, and it is well-known that $\lambda < \mu$ is necessary and sufficient for ergodicity.

To make the stochastic process associated with the number of jobs in the system at each time instant a Markov process, it is necessary to incorporate the number of phases of service the job in service still is to get, in the state variable. This results in a Markov process on the state space

$$\mathsf{E} = \{i = (i_1, i_2) \mid i_1, i_2 \in \mathbb{Z}_+\}$$

with i_1 denoting the number of jobs in the queue and i_2 the remaining number of phases of service of the job being served (cf. [9] for the finite phase case and [20] for the more general case).

Without loss of generality we can study the h-approximation of this Markov process (see [17], Theorem 3.3). This is a discrete time Markov chain, the transition probabilities of which are first order approximations of the continuous time probabilities. The non-zero transition probabilities of the h-approximation of the M/PH/1-queue are:

$$\begin{aligned}
P_{0\,0+le_2} &= \lambda p_l h \\
P_{i\,i+e_1} &= \lambda h, && i_1, i_2 > 0 \\
P_{i\,i-e_2} &= \nu h, && i_1 > 0, i_2 > 1 \\
P_{i\,i-e_1+le_2} &= \nu p_l h, && i_1 > 0, i_2 = 1 \\
P_{i\,i} &= 1 - \sum_{j \neq i} P_{i\,j},
\end{aligned}$$

where e_k is the k^{th} unit vector and $h < (\lambda + \nu)^{-1}$.

Let us say that $F \in \mathcal{G}^+(\epsilon)$ for F the distribution function of a random variable on $[0, \infty)$, if F has a convergent Laplace-Stieltjes transform in $(-\epsilon, 0)$, i.e.

$$\int_0^\infty e^{-st} dF(t) < \infty, \quad s \in (-\epsilon, 0).$$

In [18] we have shown that this h-approximation is geometrically ergodic provided that the phase-distribution of the service times is in $\mathcal{G}^+(\epsilon)$ for some $\epsilon > 0$. More specifically, the h-approximation is M-uniformly ergodic for some vector M that increases exponentially quickly in both components of the state variable. The present paper will focus on analysing the properties of the generating functions $F_{i0}(z)$ of the first hitting times of state 0 with the purpose of verifying Assumption A, so as to determine the actual rates of convergence in the M-uniform ergodicity property.

Lemma 4.1 *There is a function $a(z)$, such that*

$$F_{i0}(z) = \big(a^*(z)\big)^{i_1} \big(a(z)\big)^{i_2} \tag{ 4.1a}$$

with

$$a^*(z) = \sum_{l \geq 1} p_l \big(a(z)\big)^l. \tag{4.1b}$$

Proof By stochastic comparison it is easy to see that

$$f^m_{(i_1+1, i_2)(i_1+1,1)} = f^m_{(i_1,i_2)(i_1,1)}$$

for $i_1 \geq 0$ and $i_2 \geq 1$. Similarly

$$f^m_{(i_1+1,1)(i_1,1)} = f^m_{(i_1,1)(i_1-1,1)}$$

for $i_1 \geq 1$. Consequently

$$a_{i_2}(z) \triangleq F_{(0,i_2)(0,1)}(z) = F_{(i_1,i_2)(i_1,1)}(z)$$

and

$$F_{(i_1,1)(0,1)}(z) = \big(a^*(z)\big)^{i_1}$$

for all $i_1 \geq 1$, where

$$a^*(z) \triangleq F_{(1,1)(0,1)}(z).$$

Each path from a state (i_1, i_2) has to pass through $(i_1, 1)$ and subsequently through $(0, 1)$ before reaching 0, and therefore

$$F_{(i_1,i_2),0}(z) = a_{i_2}(z)\big(a^*(z)\big)^{i_1} F_{(0,1),0}(z). \tag{4.2}$$

For $i_2 \geq 2$

$$F_{(i_1,i_2),0}(z) = \lambda hz F_{(i_1+1,i_2),0}(z) + z(1-\lambda h-\nu h)F_{(i_1,i_2),0}(z) + z\nu h F_{(i_1,i_2-1),0}(z).$$

Combining this with (4.2), whilst dividing by $F_{(i_1,i_2),0}(z)$, this implies

$$1 = z\lambda h a^*(z) + z(1-\lambda h-\nu h) + z\nu h \frac{a_{i_2-1}(z)}{a_{i_2}(z)}, \tag{4.3}$$

for all $i_2 \geq 2$. Hence the quotient $a_{i_2-1}(z)/a_{i_2}(z)$ is independent of $i_2 \geq 2$, and we must have

$$a_{i_2}(z) = \big(a(z)\big)^{i_2-1} \tag{4.4}$$

for

$$a(z) \triangleq F_{(0,2),(0,1)}(z).$$

Using (4.3) we can then express $a^*(z)$ as a function of $a(z)$

$$z\lambda h a^*(z) = 1 - z(1-\lambda h-\nu h) - \frac{z\nu h}{a(z)}. \tag{4.5}$$

To further determine $a^*(z)$ observe that

$$F_{(i_1,1),0}(z) = z\lambda h F_{(i_1+1,1),0}(z) + z(1-\lambda h-\nu h)F_{(i_1,1),0}(z) + z\nu h \sum_l p_l F_{(i_1-1,l),0}(z)$$

and dividing by $F_{(i_1,1),0}(z)$ gives

$$1 = z\lambda h a^*(z) + z(1-\lambda h-\nu h) + z\nu h \sum_l p_l \big(a(z)\big)^{l-1} \frac{1}{a^*(z)}. \tag{4.6}$$

Together with (4.5) we find (4.1b). Similarly rewriting $F_{(0,1),0}(z)$ yields

$$1 = z\lambda h a^*(z) + z(1-\lambda h-\nu h) + \frac{z\nu h}{F_{(0,1),0}(z)}, \tag{4.7}$$

and using (4.5) shows that $F_{(0,1),0}(z)$ is equal to $a(z)$. Note that $a_{i_2}(1) = a(1) = a^*(1) = 1$ and so these functions are not identically 0 in their domain of analyticity. As a consequence they have only got isolated zeros and hence by continuity arguments the validity of (4.3), (4.6) and (4.7) is guaranteed for all z in the domain of analyticity of the respective functions. Combination with (4.2) and (4.4) finally proves the lemma. **QED**

By the productform structure of the $F_{i0}(z)$ it is very easy now to verify Assumption A.

LEMMA 4.2 *The h-approximation of the M/PH/1-queue satisfies Assumption A.*

Proof Suppose that $F_{0,0}(\tilde{z}) = 1$ for some $\tilde{z} \neq 1$. Then

$$1 = F_{0,0}(z) = \tilde{z}(1 - \lambda h) + \tilde{z}\lambda h \sum_l p_l F_{(0,l),0}(z)$$
$$= \tilde{z}(1 - \lambda h) + \tilde{z}\lambda h \alpha^*(\tilde{z})$$

so that

$$a^*(\tilde{z}) = \frac{1 - \tilde{z} + \tilde{z}\lambda h}{\tilde{z}\lambda h}.$$

Plug this into (4.7), then

$$a(\tilde{z}) = F_{(0,1),0}(\tilde{z}) = 1$$

so that $a^*(\tilde{z}) = 1$ from (4.1b). Hence

$$1 = F_{0,0}(\tilde{z}) = \tilde{z}(1 - \lambda h) + \tilde{z}\lambda h = \tilde{z},$$

thus proving that $z = 1$ is the only 1-point of $F_{0,0}(z)$. **QED**

Lemma 4.2 allows to conclude from Theorem 3.1 that the contraction factor in Popov's criterion (1.2) is the second largest eigenvalue with respect to the matrix norm induced by a vector M solving (1.2a) with equality. As a candidate for this M we use $M_i = F_{i0}(\beta^{-1})$, $i \neq 0$, and we will use the structure given in Lemma 4.1 to further determine the range of β's for which the $F_{i0}(\beta^{-1})$ converge.

We can write $F_{(0,1),0}(\beta^{-1}) = 1 + x$, for $\beta < 1$ and some $x > 0$ depending on β. By virtue of Lemma 4.1 $F_{i,0}(\beta^{-1})$ then equals

$$M_i(x) \triangleq \Big(\sum_l p_l (1 + x)^l\Big)^{i_1} (1 + x)^{i_2},$$

and we will further analyse the relation between β and x. Clearly for any positive x, $M_i(x)$ solves (1.2a) with equality for the "contraction factor" $\beta(x)$ satisfying

$$\beta(x) = 1 + \lambda h \sum_l p_l \big[(1 + x)^l - 1\big] - \nu h \frac{x}{1 + x}. \tag{4.8}$$

As we shall see below, this $\beta(x)$ is indeed a contraction factor, i.e. $\beta(x) < 1$, for some $x > 0$, provided that the service time distribution lies in $\mathcal{G}^+(\epsilon)$ for some $\epsilon > 0$. This condition is not only sufficient but also necessary for geometric ergodicity: see [18], Theorem 6.2.

Under this condition on the service time distribution $M_i(x) < \infty$, $\beta(x) < \infty$ and

$$\frac{d}{dx}\beta(x) = \lambda h \sum_l l p_l (1 + x)^{l-1} - \nu h \frac{1}{(1 + x)^2},$$

for all sufficiently small $x > 0$, so that

$$\frac{d}{dx}\beta(x)\big|_{x=0} = \lambda h \sum_l l p_l - \nu h$$
$$= \nu h \Big\{\frac{\lambda}{\mu} - 1\Big\} < 0$$

in view of the M/PH/1-queue being ergodic. As $\beta(0) = 1$, this implies $\beta(x) < 1$ for small enough $x > 0$. In particular $\beta(x)$ decreases for $x \leq \tilde{x}$, where $\tilde{x}$ is given by

$$\tilde{x} \triangleq \min\{x : x \geq 0, \sum_l l p_l (1+x)^{l+1} = \frac{\nu}{\lambda}\},$$

and $\beta(x)$ increases at least for all $x > \tilde{x}$ sufficiently close to $\tilde{x}$, whilst $M_i(x)$ increases for all positive x. By virtue of Lemma 3.3 we must have for $i \neq 0$ and $x \leq \tilde{x}$

$$F_{i0}(\beta^{-1}(x)) = M_i(x),$$

with $\beta^{-1}(x) = 1/\beta(x)$. Moreover, by the continuity and montonicity of $F_{i0}(r)$ as a function of r, the convergence radius of $F_{i0}(r)$ must be equal to $\beta^{-1}(\tilde{x})$. This leads to the following conclusion.

THEOREM 4.3 *For the h-approximation of the M/PH/1-queue the disc of convergence $\mathcal{D}$ of $F_{00}(z)$ is precisely $\overline{\mathcal{D}}_{0,\beta^{-1}(\tilde{x})}$ provided the service time distribution is in $\mathcal{G}^+(\epsilon)$ for some $\epsilon > 0$. For positive x with $\beta(x) < 1$ we have that*

$$\beta_{M(x)}(P) = \beta(x) = 1 + \lambda h \sum_l p_l \big[(1+x)^l - 1\big] - \nu h \frac{x}{1+x},$$

and the conclusions of Theorem 3.1 apply, i.e. for $\delta > 0$

$$\|P^n - \Pi\|_{M(x)} \leq c_{\beta(x),\delta}(\beta(x) + \delta)^n$$

with $c_{\beta(x),\delta}$ given by Theorem 3.1. Moreover

$$\|P^n_{i\bullet} - \Pi_{i\bullet}\| \leq c_i \big[\beta(\tilde{x})\big]^n,$$

with c_i also given by Theorem 3.1.

Proof That $\beta(x)$ is the second largest eigenvalue follows from Theorem 3.2, since the h-approximation of the *M/PH/*1-queue has the $M(x)$-left skipfree property and $M(x)$ is unbounded off finite sets. The other conclusions follow from Lemma 4.2, the foregoing arguments and Theorem 3.1. **QED**

We may now apply Theorem 3.3 from [17] to obtain a similar result for the original Markov process associated with the single server phase type queue.

COROLLARY 4.4 *For the M/PH/1-queue let $P_{ij}(t) = \mathbb{P}\{\Xi_t = j \mid \Xi_0 = i\}$, $t \in \mathbb{R}_+$. Then*

$$\|P(t) - \Pi\|_{M(x)} \leq c_{\beta(x),\delta} \exp\{(\delta h^{-1} - \tilde{\beta}(x))t\}$$

with

$$\tilde{\beta}(x) = \frac{1 - \beta(x)}{h} = \nu \frac{x}{1+x} - \lambda \sum_l p_l \big[(1+x)^l - 1\big]$$

and $c_{\beta(x),\delta}$ given by Theorem 3.1. Similarly

$$\|P_{i\bullet}(t) - \Pi_{i\bullet}\| \leq c_i \exp\{-\tilde{\beta}(\tilde{x})t\};$$

$\tilde{\beta}(\tilde{x})$ is the sharpest possible contraction rate and c_i is given by Theorem 3.1.

Proof Since $\beta(x)$ depends linearly on h, the first statement follows from Theorem 3.3 in [17]. The second statement follows in a similar way. **QED**

For computing c_δ and c_i for specific parameters λ, ν and p_l, $l \geq 1$, we need to analyse the function $F_{00}(z)$. By Lemma 4.1 and Lemma 3.3 $F_{(0,1)}(z)$ is the smallest root in absolute value of the equation

$$a(z) = z\lambda h \sum_l p_l\big(a(z)\big)^{l+1} + z(1-\lambda h - \nu h)a(z) + z\nu h$$

and then $F_{00}(z) = z\lambda h \sum_l p_l(a(z))^l + z(1-\lambda)h$.

As an example we will treat the case that $p_1 = 1$ and $p_l = 0$, $l \neq 1$, which is the h-approximation of the M/M/1-queue. As state $(i_1, 1)$ unambiguously corresponds to $i_1 + 1$ jobs being present in the system, we will now write i for state $(i-1, 1)$, for $i \in \mathbb{Z}_+$. Then

$$M_i(x) = (1+x)^i,$$
$$\beta(x) = 1 + xh\big(\lambda - \frac{\nu}{1+x}\big)$$

and

$$\tilde{x} = \sqrt{\nu/\lambda} - 1,$$

so that the minimal contraction factor

$$\beta(\tilde{x}) = 1 - h\big(\sqrt{\nu} - \sqrt{\lambda}\big)^2$$

is associated with the vector $M_i(\tilde{x}) = \big(\sqrt{\nu/\lambda}\big)^i$. For fixed $\beta \in (\beta(\tilde{x}), 1)$ there are two solutions x_k, $k = 1, 2$, for which the corresponding vector $M_i(x_k)$, $i \in \mathsf{E}$, solves Popov's criterion with equality for the contraction factor β:

$$x_1 = \frac{\beta-1}{2h\lambda} + \frac{\nu-\lambda}{2\lambda} - \frac{1}{2\lambda}\sqrt{\big(\frac{\beta-1}{h} + \nu - \lambda\big)^2 - 4\lambda\frac{1-\beta}{h}}$$
$$x_2 = \frac{\beta-1}{2h\lambda} + \frac{\nu-\lambda}{2\lambda} + \frac{1}{2\lambda}\sqrt{\big(\frac{\beta-1}{h} + \nu - \lambda\big)^2 - 4\lambda\frac{1-\beta}{h}}.$$

Then $F_{i0}(\beta^{-1}) = M_i(x_1)$, $i \geq 1$, and $M_i(x_2) > F_{i0}(\beta^{-1}(\tilde{x}))$.

For this example it is very easy to compute the bounds c_δ and c_i, since $F_{00}(z) = z\lambda h F_{10}(z) + z(1-\lambda_h)$ with $F_{10}(z)$ the smallest root in absolute value of

$$a(z) = z\lambda h\big(a(z)\big)^2 + z(1-\lambda h - \nu h)a(z) + z\nu h.$$

The following table shows some numerical results, where $f_{\beta+\delta} \stackrel{\triangle}{=} \sup\{|(1-F_{00}(z))^{-1}| : z \in \mathcal{C}_{0,(\beta+\delta)^{-1}}\}$. Further write $\tilde{c} \stackrel{\triangle}{=} c_i/F_{i0}(\beta^{-1}\tilde{()})$; the $c_{\beta,\delta}$ and c in the table are taken to be the smallest integers upperbounding the corresponding constants as defined through Theorem 3.1.

λh	νh	δ	$\beta(\tilde{x})$	$M_i(\tilde{x})$	$f_{\beta(\tilde{x})+\delta}$	$\gamma = \beta + \delta$	$c_{\beta,\delta}$	$\tilde{c}$
0.1	0.2	0	0.9828	$2^{i/2} = (1.4142)^i$	21.8969	0.9828		135
		0.01			58.3703	0.9928	14179	
		0.015			127.2449	0.9978	36048	
		0.017			156.9105	0.9998	379919	
0.1	0.5	0	0.8472	$5^{i/2} = (2.2361)^i$	3.6850	0.8472		13
		0.01			4.0023	0.8572	452	
		0.05			6.2836	0.8972	204	
		0.1			13.7601	0.9472	290	
		0.15			188.3280	0.9972	814	
0.1	0.8	0	0.6657	$8^{i/2} = (2.8284)^i$	1.4879	0.6657		5
		0.01			1.5587	0.6757	184	
		0.1			2.6247	0.7657	45	
		0.2			5.4381	0.8657	61	
		0.3			24.3406	0.9757	220	
		0.334			273.3084	0.9997	13736	

As a consequence we have e.g. for $\lambda = 0.1$, $\nu = 0.5$ and $h = 1$ that

$$\sum_j |P_{ij}^n - \Pi_{ij}| 5^{j/2} \le 204 \cdot (0.8972)^n 5^{i/2},$$

whilst choosing $\delta = .05$, so that for any reward vector r on E with $c \triangleq \sup\{|r_i 5^{-i/2}| : i \in \mathsf{E}\} < \infty$

$$\big|\mathbb{E}(r(\Xi_n) \,|\, \Xi_0 = i) - \sum_j \Pi_{ij} r_j\big| \le c \cdot 204 \cdot (0.8972)^n 5^{i/2}$$

and we thus have geometric convergence of all moments. Moreover for the unweighted differences there is a far better bound:

$$\sum_j |P_{ij}^n - \Pi_{ij}| \le 13 \cdot (0.8472)^n 5^{i/2}.$$

For the associated Markov process in continuous time we have a similar result:

$$\big|\mathbb{E}(r(\Xi_t) \,|\, \Xi_0 = i) - \sum_j \Pi_{ij} r_j\big| \le c \cdot 204 \cdot \exp\{-0.1028t\} 5^{i/2}$$

and for the unweighted differences

$$\sum_j |P_{ij}(t) - \Pi_{ij}| \le 13 \cdot \exp\{-0.1528t\} 5^{i/2}.$$

The vast difference between $c_{\beta,\delta}$ and $c_i/F_{i0}(\beta^{-1})$ for small δ is accounted for by the fact that in this case $c_{\beta,\delta}$ is closely related to the first derivative $F_{i0}^1(\beta^{-1})$ and in general this expression is not bounded by M, in particular not if we take $M_i = F_{i0}(\beta^{-1})$, $i \neq 0$.

Note also, that increasing $\beta(x)$, while fixing $\gamma = \beta(x) + \delta$, implies that $M_i(x)$ increases, so that geometric convergence results for a wider class of reward functionals are implied. Meanwhile $c_{\beta,\delta}$

increases and thus we get worse bounds for the marginal distributions and marginal rewards associated with polynomial immediate reward functionals.

Acknowledgement

I would like to thank A. Hordijk, T. Ott and R. Tweedie for stimulating discussions and helpful comments.

References

[1] Chung, K.L. *Markov Chains with stationary transition probabilities* (2d edition). Springer Verlag, Berlin, 1967.

[2] Doorn, E.v. Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process. *Adv. Appl. Prob.* **17**, 514-530, 1985.

[3] Hordijk, A. and Spieksma, F.M. On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. Appl. Prob.*, **24**, 343-376, 1992.

[4] Isaacson, I. and Luecke, G.R. Strongly ergodic Markov Chains and rates of convergence using spectral conditions. *Stoch. Proc. Appl.* **7**, 113-121, 1978.

[5] Karlin, S. Positive operators, *J. Math. Mech.*, **8**, 907-937, 1959.

[6] Kendall, D.G. Geometric Ergodicity and the Theory of Queues. In: *Mathematical Methods in the Social Sciences 1959* (eds. K.J. Arrow, S. Karlin, P. Suppes). Stanford University Press, Stanford, 176-195, 1960.

[7] Meyn, S. and Tweedie, R.L. *Markov Chains and Stochastic Stability.* Control and Communication in Engineering, Springer Verlag, 1992.

[8] Mihail, M. Conductance and Convergence of Markov Chains-A Combinatorial Treatment of Expanders. *30th Annual Symp. on Found. Comp. Sci.*, 526-531, 1989.

[9] Neuts, M.F. *Matrix Geometric Solutions for Stochastic Models.* John Hopkins University Press, 1981.

[10] Nummelin, E. and Tweedie Geometric ergodicity and R-positivity for general Markov chains. *Ann. Appl. Prob.* **6**, 404-420, 1978.

[11] Popov, N.N. Conditions for geometric ergodicity of countable Markov chains. *Soviet Math. Dokl.* **18**, 676-679, 1977.

[12] Royden, H.L. *Real Analysis* (3d edition). Macmillan Publishing Company, New York, 1988.

[13] Seneta, E. *Non-negative Matrices and Markov Chains.* Springer-Verlag, Berlin, 1981, 2d edition.

[14] Sinclair, A. and Jerrum, M. Approximate counting, uniform generation and rapidly mixing Markov Chains. *Inform. Comp.* **82**, 93-133, 1989.

[15] Spieksma, F.M. *Geometrically ergodic Markov Chains and the optimal Control of Queues*, Ph. D. thesis, University of Leiden, 1990 (available on request from the author).

[16] Spieksma, F.M. The existence of sensitive optimal policies in two multi-dimensional queueing models. *Ann. Operat. Res.* **28**, 273-296, 1991.

[17] Spieksma, F.M. A recurrence type characterisation of M-uniform ergodicity for Markov processes. Technical report no TW-92-03, University of Leiden, 1992, *submitted.*

[18] Spieksma, F.M. and Tweedie, R.L. Strengthening ergodicity to geometric ergodicity of Markov chains. Technical report no TW-92-05, Univ. of Leiden, 1992, *accepted for publication in Stoch.*

Models.
[19] TITCHMARSH, E.C. *The Theory of Functions* (2d edition). Oxford University Press, Oxford, 1939.
[20] TWEEDIE R.L. Operator-geometric stationary distributions for Markov chains, with applications to queueing models. *Adv. Appl. Prob.* **14**, 368-391, 1982.
[21] TWEEDIE R.L. Criteria for rates of convergence of Markov chains, with applications to queueing and storage theory. In: *Papers in Probability, Statistics and Analysis* (eds. J.F.C. Kingman, G.E.H. Reuter), 260-276, 1983.
[22] VERE-JONES, D. Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford (2)*, **13**, 7-28, 1962.
[23] VERE-JONES, D. On the Spectra of some Linear Operators associated with Queueing Systems. Z. Wahrscheinlichkeitstheorie u. verw. Geb. **2**, 12-21, 1963.
[24] VERE-JONES, D. A rate of convergence problem in the Theory of Queues. *Theory Prob. Appl.* **9**, 96-103, 1964.
[25] YOSIDA, K. *Functional Analysis* (2d edition). Springer Verlag, Berlin, 1980.