Series Expansions for Continuous-Time Markov Processes

Bernd Heidergott[a], Arie Hordijk[b] and Nicole Leder[c]

[a] *Vrije Universiteit*, Department of Econometrics
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
email: bheidergott@feweb.vu.nl

[b] Leiden University, Mathematical Institute
P.O. Box 9512, 2300 RA Leiden, The Netherlands
email: hordijk@math.leidenuniv.nl

[c] University Hamburg, Department of Mathematics
Bundesstrasse 55, 20146 Germany
email: Nicole.Leder@studium.uni-hamburg.de

August 29, 2007

Abstract

We present exchange formulas that allow to express the stationary distribution of a continuous Markov chain with denumerable state-space having generator matrix Q^* through a continuous time Markov chain with generator matrix Q. Under suitable stability conditions, numerical approximations can be derived from the exchange formulas, and we show that the algorithms converge at a geometric rate. Applications to sensitivity analysis and bounds on perturbations are discussed as well. Numerical examples are presented to illustrate the numerical efficiency of the proposed algorithm.

1 Introduction

Let $\mathcal{X} = \{X_t, t \geq 0\}$ be a continuous-time ergodic Markov process on a denumerable state space S. Throughout this paper we will denote its transition matrix by $P(t)$, and its infinitesimal generator by Q. The infinitesimal generator is assumed to be conservative, where a matrix is called conservative if its row-sums are equal to zero. Furthermore, we assume that $\mathcal{X}$ has a unique stationary distribution, denoted by π, and we denote the associated ergodic matrix by Π, i.e., Π is a matrix with rows equal to π. Suppose that Q^* is the conservative generator matrix of another continuous time Markov chain $\mathcal{X}^* = \{X_t^*, t \geq 0\}$ defined on the same state space as $\mathcal{X}$ and denote its stationary distribution by π^* (existence assumed) and the ergodic matrix by Π^*. This paper addresses the following problem: can the difference between Π and Π^* be estimated from Q^*, Q and Π? In other words, what is the effect of changing Q to Q^* on the stationary distribution of the Markov process? For example, consider an open queueing network with arrival rate λ. What is the effect on the stationary distribution if λ is changed to λ^*?

In this paper we will present a general exchange formula that allows to express Π^* as a mapping of Q^*, Q and Π only. More precisely, we will show that

$$\Pi^* = \Pi \sum_{n \geq 0} ((Q^* - Q)D)^n,$$

where $D = (d_{ij})_{i,j \in S}$ is the deviation matrix associated with Q, which has elements

$$d_{ij} \stackrel{\text{def}}{=} \int_0^\infty (p_{ij}(t) - \pi_j) dt, \quad i, j \in S. \tag{1}$$

The deviation exists whenever all integrals in (1) are finite. Hence, provided the deviation matrix exists and the exchange formula converges, Π^* can be presented as a power series in Q^*, for given given Q, D and Π. As a first application of the exchange formula, we will derive an expression for derivatives of Π with respect to parameters of Q. Also, bounds on perturbations will be established. Moreover, we will establish sufficient conditions such that D can be obtained in closed form.

Based on this, we we derive a numerical algorithm for finite state-space Markov chains that allows to compute Π^* out of Q^*. In applications this has the nice feature that once D and Π are evaluated (either by explicit formula or numerically), our algorithm offers the opportunity of computing Π^* in a fast way. In particular, we will show that the algorithm converges exponentially fast and we will provide an efficient bound on the error made by evaluating the exchange formula only for a finite number of elements. This provides a tool for discrete optimization as once Π and D are computed, Π^* can be computed for various choices of Q^* by simple matrix multiplication, and we consider the exchange formula as a fast and efficient way of analyzing the effect of changing Q to Q^* on the stationary distribution for various possible alternatives simultaneously in a fast way.

In addition, a functional version of the algorithm exists that allows to present Π as function of θ where θ is a parameter of Q. More precisely, let Q depend on some parameter θ in a linear way. For example if Q is the generator matrix of an open Jackson network, then θ may be the arrival rate and the entries of Q_θ are linear mappings in θ. Writing Q_θ for Q at θ and denoting the ergodic matrix of Q_θ by Π_θ and the deviation matrix by D_θ, we show that

$$\Pi_{\theta+\Delta} = \Pi_\theta \sum_{n \geq 0} \Delta^n ((Q_{\theta'} - Q_\theta) D_\theta)^n,$$

for any $\theta' \neq \theta$. The above series expansion is called a *functional series expansion* as it allows to obtain $\Pi_{\theta+\Delta}$ as a polynomial in Δ. This provides a tool for continuous optimization in the following way. Suppose that one is interested in the maximal value of θ such that $\pi_\theta g \leq c$ for some cost function g and given constant c. Then, the functional series expansion can be used for obtaining $\pi_{\theta+\Delta} g$ as a polynomial in Δ. Hence, $\pi_{\theta+\Delta} g = c$ can be solved for, say Δ^*, and $\theta^* = \theta + \Delta^*$ yields an (approximate) solution to the problem $\pi_\theta g \leq c$.

The paper is organized as follows. In Section 2 basic properties of continuous-time Markov chains are presented. The exchange formula and its application to sensitivity analysis and bounding of perturbations is investigated in Section 3. Numerical algorithms are discussed in Section 4. Functional series expansions are presented in Section 5.

2 Preliminaries on Denumerable Markov Chains

2.1 Basic Properties of the Generator Matrix

We first establish basic properties of the infinitesimal generator Q and the ergodic matrix Π.

Lemma 2.1 *If it exists the ergodic matrix of a continuous-time Markov process $\mathcal{X}$ with infinitesimal generator Q satisfies*

(i) $\Pi Q = Q \Pi = 0$, *and even $A\Pi = 0$ for all conservative matrices $A \in \mathbb{R}^{S \times S}$,*

(ii) $B\Pi = \Pi$ *for all stochastic matrices $B \in \mathbb{R}^{S \times S}$.*

Proof: Due to the equal rows of Π the proof of (ii) and the second equation of (i) can be executed straightforward by extracting π_j, $j \in S$ outside the sum. Then $\pi_j \sum_{j\in S} a_{i,j} = 0$ and $\pi_j \sum_{i\in S} b_{i,j} = \pi_j$ follow from the conservativeness of A and the properties of the stochastic matrix B, respectively. For the proof of $\Pi Q = 0$ it is advisable to apply the transition rates of the related time reversed process $\mathcal{X}^{rev}$. The existence of such a process is assured by Theorem 1.3 in [16] which provides the necessary and sufficient condition that for $\mathcal{X}$ it exists at least one $\lambda_i > 0$, $i \in S$, satisfying the local balance equation

$$\lambda_i q_{i,j} = \lambda_j q_{j,i}, \quad i, j \in S.$$

Obviously, this condition holds for the ergodic distribution π so that all processes investigated in this paper are assured to be reversible. The transition rates of the reversed process are

$$q_{i,j}^{rev} = \frac{\pi_j q_{i,j}}{\pi_i}, \quad i, j \in S, \tag{2}$$

and it holds for the ergodic distributions

$$\pi^{rev} = \pi$$

(for a more detailed presentation, we refer to [16]). By applying (2) and replacing $q_{i,j}$ by the reversed rates we obtain the desired result

$$\sum_{i\in S} \pi_i \frac{\pi_j q_{j,i}^{rev}}{\pi_i} = \pi_j \sum_{i\in S} q_{j,i}^{rev} = 0, \quad i, j \in S.$$

■

The following lemma establishes key properties of the deviation matrix for time-continuous Markov chains.

Lemma 2.2 *If it exists the deviation matrix of a continuous-time Markov process with infinitesimal generator Q and ergodic matrix Π satisfies*

(i) $\Pi D = 0$,

(ii) $-QD = I - \Pi$.

Proof: Since the proof requires a splitting of the integrals in (1) with the separate integrals infinite, we replace the entries of the deviation matrix by the associated Laplace transforms. For appropriate $\alpha > 0$ and $i, j \in S$ it holds

$$\sum_{k\in S} \pi_k \hat{d}_{kj}(\alpha) = \sum_{k\in S} \pi_k \int_0^\infty e^{-\alpha t}(p_{kj}(t) - \pi_j)dt = \sum_{k\in S} \pi_k \int_0^\infty e^{-\alpha t} p_{kj}(t)dt - \sum_{k\in S} \pi_k \int_0^\infty e^{-\alpha t} \pi_j dt.$$

Justified by Fubini's theorem, we can interchange summation and integration so that we obtain from Lemma 2.1 (ii)

$$\sum_{k\in S} \pi_k \hat{d}_{kj}(\alpha) = \int_0^\infty e^{-\alpha t} \pi_j dt - \int_0^\infty e^{-\alpha t} \pi_j dt = 0.$$

By applying the same Laplace transform to (ii), we get

$$\sum_{k\in S} q_{ik} \hat{d}_{kj}(\alpha) = \sum_{k\in S} q_{ik} \int_0^\infty e^{-\alpha t}(p_{kj}(t) - \pi_j)dt = \int_0^\infty e^{-\alpha t} \left(\sum_{k\in S} q_{ik} p_{kj}(t) - \pi_j \sum_{k\in S} q_{ik} \right) dt.$$

Justified by the conservativeness of Q we have $\sum_{k\in S} q_{ik} = 0$ and from Kolmogorov's backward equation which is valid for all kinds of continuous-time Markov processes even with unbounded transition rates we get $\sum_{k\in S} q_{ik} p_{kj}(t) = p'_{ij}(t)$. Hence it holds

$$\sum_{k\in S} q_{ik}\hat{d}_{kj}(\alpha) = \int_0^\infty e^{-\alpha t} p'_{ij}(t)dt.$$

Now we let α converge to 0 and obtain for the right-hand side of the equation with the interchange of limit and integration justified by the monotone convergence theorem

$$\int_0^\infty p'_{ij}(t)dt = \lim_{t\to\infty} p_{ij}(t) - \lim_{t\to 0} p_{ij}(t) = \pi_j - \delta_{ij}$$

where δ_{ij} denotes Kronecker's delta, i.e., $\delta_{ij} = 1$ for $i = j$ and zero otherwise. Now it remains to show that the limits for $\alpha \to 0$ of $\sum_{k\in S} \pi_k \hat{d}_{kj}(\alpha)$ and $\sum_{k\in S} q_{ik}\hat{d}_{kj}(\alpha)$ equal their values at $\alpha = 0$. Therefore, we refer to the proof's of Syski's Proposition 3.6 and Corollary 3.7 in [25] which are based upon the dominated convergence theorem.

■

Definition 2.3 *A generator matrix Q is called uniformizable with rate μ if $\mu = \sup_j |q_{jj}| < \infty$.*

Any finite dimensional generator matrix is uniformizable. A classical example of a Markov chain on denumerable state-space that fails to be uniformizable is the M/M/∞ queue. Note that if Q is uniformizable with rate μ, then Q is uniformizable with rate η for any $\eta > \mu$.

Let Q be uniformizable with rate μ and introduce the Markov chain P_μ as follows

$$[P_\mu]_{ij} = \begin{cases} q_{ij}/\mu & i \neq j \\ 1 + q_{ii}/\mu & i = j, \end{cases} \tag{3}$$

for $i, j \in S$, then it holds that

$$P(t) = e^{-\mu t} \sum_{n=0}^\infty \frac{(\mu t)^n}{n!} (P_\mu)^n, \quad t \geq 0. \tag{4}$$

Moreover, the stationary distribution of P_μ and $P(t)$ coincide, in formula: $\Pi_\mu = \Pi$. The Markov chains $\mathcal{X}_\mu = \{X_n^\mu : n \geq 0\}$ with transition probability P_μ is called the *subordinate* chain. The relationship between $\mathcal{X}$ and $\mathcal{X}_\mu$ can be expressed as follows. Let $N_\mu(t)$ denote a Poisson process with rate μ, then $X^\mu_{N_\mu(t)}$ and X_t are equal in distribution for all $t \geq 0$. The deviation matrix associated with P_μ is defined by

$$D_\mu = \sum_{n\geq 0} ((P_\mu)^n - \Pi_\mu) = \sum_{n\geq 0} (P_\mu - \Pi_\mu)^n + \Pi_\mu,$$

provided the sum exists. If Q is uniformizable with rate μ and the deviation matrices associated with P_μ and $P(t)$ exist, then

$$\frac{1}{\mu} D_\mu = D, \tag{5}$$

for a proof see [7]. Let Q^* be uniformizable with rate μ^* and let Q be uniformizable with rate μ. Let $\eta = \max(\mu, \mu^*)$, then (3) implies that

$$P^*_\eta - P_\eta = \frac{1}{\eta}(Q^* - Q)$$

and by (5) (with $\mu = \eta$) it follows that

$$(P^*_\eta - P_\eta) D_\eta = (Q^* - Q) D. \tag{6}$$

2.2 Geometric Ergodicity

The main tool for our analysis is the weighted supremum norm, also called v-norm, denoted by $\|\cdot\|_v$, where v is some vector, with elements $v_i \geq 1$ for all $i \in S$, and for any $w \in \mathbb{R}^{\mathcal{S}}$

$$||w||_v \stackrel{\text{def}}{=} \sup_{i \in \mathcal{S}} \frac{|w(i)|}{v(i)}. \tag{7}$$

For a matrix $A \in \mathbb{R}^{S \times S}$ the v-norm is given by

$$\|A\|_v \stackrel{\text{def}}{=} \sup_{i, ||w||_v \leq 1} \frac{\sum_{j=1}^{S} |A(i,j) w(j)|}{v(i)},$$

which implies

$$\max_{j \in S} |A|(i,j) \leq ||A||_v \, v(i) \,, \quad i \in S. \tag{8}$$

Note that v-norm convergence to 0 implies elementwise convergence to 0. With the help of the above concepts, v-geometric ergodicity (also called v-normed ergodicity) of $P(t)$ can be introduced as follows.

Definition 2.4 *The Markov chain $\mathcal{X}$ is v-geometric ergodic if $c < \infty$ and $\beta < 1$ exist such that*

$$\|P(t) - \Pi\|_v \leq c\beta^t,$$

for all $t \geq 0$.

Note that

$$||D||_v \leq \int_0^\infty ||P(t) - \Pi||_v dt$$

and it is straightforward to check that geometric v-norm ergodicity implies existence of $||D||_v$. Unfortunately, geometric v-norm ergodicity is almost impossible to check in a direct way. One of the reasons is that $P(t)$ is in general not known in explicit form. If Q is uniformizable, v-norm ergodicity of $P(t)$ can be deduced from v-norm ergodicity of P_μ in discrete time. The precise statement is in the following lemma.

Lemma 2.5 *Let Q be uniformizable with rate μ. If finite constants N, c and β, with $0 \leq \beta < 1$ exist such that $||(P_\mu)^n - \Pi||_v \leq c\beta^n$ for all n, then $P(t)$ is v-norm ergodic.*

Proof: If $\mathcal{X}$ is uniformizable with rate μ, then by (4)

$$\begin{aligned}
||P(t) - \Pi||_v &\leq e^{-\mu t} \sum_{n=0}^{\infty} \frac{(\mu t)^n}{n!} ||(P_\mu)^n - \Pi||_v \\
&\leq c e^{-\mu t} \sum_{n=0}^{\infty} \frac{(\mu t)^n}{n!} \beta^n \\
&= c e^{-\mu(1-\beta)t} \\
&\leq c \left(e^{-\mu(1-\beta)} \right)^t .
\end{aligned}$$

Noting that $e^{-\mu(1-\beta)} < 1$ proves the claim.

■

Like in discrete time case, geometric v-norm ergodicity is sufficient for the deviation matrix to exist.

Lemma 2.6 *Let Q be uniformizable with rate μ. If P_μ is v-norm ergodic, then $||D||_v$ is finite.*

Proof: If Q is uniformizable with rate μ, then (4) yields for the deviation matrix

$$D = \int_0^\infty (P(t) - \Pi) dt = \int_0^\infty e^{-\mu t} \sum_{n=0}^\infty \frac{(\mu t)^n}{n!} ((P_\mu)^n - \Pi).$$

Suppose that $||(P_\mu)^n - \Pi||_v \leq c\beta^n$, for $0 \leq \beta < 1$ and all n. This implies

$$\begin{aligned} ||D||_v &\leq \int_0^\infty e^{-\mu t} \sum_{n=0}^\infty \frac{(\mu t)^n}{n!} ||(P_\mu)^n - \Pi|| dt \\ &\leq \int_0^\infty e^{-\mu t} \sum_{n=0}^\infty \frac{(\mu t)^n}{n!} c\beta^n dt \\ &= c \int_0^\infty e^{-\mu t(1-\beta)} dt \\ &= \frac{c}{\mu(1-\beta)}. \end{aligned}$$

■

In the following lemma we derive an explicit representation of the deviation matrix.

Lemma 2.7 *Let Q be uniformizable with rate μ. If P_μ is geometrically v-norm ergodic, then D_μ and D exist and it holds that*

$$D_\mu = \sum_{n \geq 0} ((P_\mu)^n - \Pi_\mu)$$

and

$$D = \frac{1}{\mu} \sum_{n \geq 0} ((P_\mu)^n - \Pi).$$

Proof: Let P_μ is geometrically v-norm ergodic such that $||P_\mu^n - \Pi_\mu||_v \leq c\beta^n$ for some $\beta < 1$ and all n. Since

$$\left\| \sum_{n \geq 0} ((P_\mu)^n - \Pi_\mu) \right\|_v \leq \sum_{n \geq 0} ||((P_\mu)^n - \Pi_\mu)||_v \leq \frac{c}{1-\beta}$$

it follows that the sum on the above left hand side converges, which proves the first part of the lemma.

The second part of the lemma follows from the fact that $\Pi_\mu = \Pi$ and $D_\mu = \mu D$.

■

Let

$$m_{e,j} = \sum_{i \in S} \pi_i m_{i,j},$$

where $m_{i,j}$ is the mean first entrance time for state i to state j. Kemeny and Snell showed in [18] that $D = \sum_{n \geq 0} (I + Q - \Pi)^n - \Pi$ provided that $m_{e,j} < \infty$ for all $j \in S$. According to Coolen-Schrijner and van Doorn [7], the aforementioned condition on $m_{e,j}$ can be relaxed to the following: there is one $j \in S$ such that $m_{e,j} < \infty$. The precise statement is given in the lemma below.

Lemma 2.8 *If $m_{e,j} < \infty$ for at least one $j \in S$, then*

$$D = \sum_{n \geq 0} (I + Q - \Pi)^n - \Pi.$$

Note that $m_{e,j}$ is finite for finite state space S. Hence, the above lemma provides a proof for the fact that for any finite Markov chain the deviation matrix can be obtained through the inverse of $\Pi - Q$.

Example 2.9 *Consider a stable $M/M/\infty$. Then, $m_{e,0} < \infty$ and the deviation matrix exists and can be computed as described in Lemma 2.8.*

Remark 2.10 *The weighted supremum norm goes back to [20]. Normed ergodicity dates back to the early eighties, see [12] and the revised version which was published as [8]. It was originally used in analysis of Blackwell optimality; see [8], and [14] for a recent publication on this topic. Since then, it has been used in various forms under different names in many subsequent papers. In [13] it was shown for a countable Markov chain which may have one or several classes of essential states (a so-called multichained Markov chain) that normed ergodicity is equivalent to geometrical recurrence (for a similar result in Markov decision chains see [9]). Inspired by this result for a countable Markov chain a similar result was proved for a Harris chain in [22]. In this paper we use the recent results of [2], the first part of this technical report has appeared as [3].*

We will frequently use results from [11] on v-norms. While the analysis in [11] has been carried out for finite Markov chains, part of the results in [11] are obtained by purely norm-theoretic arguments applied to products of matrices and can be carried over to the denumerable state-space case without any harm. In the following we give for easy reference a summary of the results from [11] that we will use.

Let A, B, C, F be square matrices (possibly infinite dimensional), and set

$$\mathrm{H}(n) = F \sum_{k=0}^{n} ((A - B)C)^k,$$

$$\mathrm{R}(n) = F \sum_{k=n+1}^{\infty} ((A - B)C)^k,$$

and

$$G = F \sum_{k=0}^{\infty} ((A - B)C)^k,$$

provided that the sum converges. The following type of condition will be frequently used:

[A,B,C] *There exists a finite number N such that we can find $\delta_N \in (0, 1)$ which satisfies:*

$$\|((A - B)C)^N\|_v < \delta_N,$$

and we set

$$c^v_{A,B,C} \stackrel{\text{def}}{=} \frac{1}{1 - \delta_N} \left\| \sum_{k=0}^{N-1} ((A - B)C)^k \right\|_v .$$

The following result has been established in [11].

Lemma 2.11 *Suppose that $\|F\|_v$ is finite. Let*

(i) *Condition [A,B,C] holds.*

(ii) *There exists $\rho \in (0, \infty)$ and $\delta \in (0, 1)$ such that for all k*

$$||((A - B)C)^k||_v \leq \rho \delta^k.$$

(iii) *For all k it holds that*

$$||\mathrm{R}(k-1)||_v \leq c^v_{A,B,C}||F||_v \rho\delta^k,$$

with δ and ρ as in (ii) *and $c^v_{A,B,C}$ as in* (i).

(iv) $||\mathrm{H}(n)||_v$ *converges* $||G||$ *as n tends to infinity, in particular, and* $||\mathrm{R}(n)||_v$ *converges to zero as n tends to infinity at a geometric rate.*

(v) $\mathrm{H}(n)$ *converges G as n tends to infinity.*

Then,

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(iv)} \Rightarrow \text{(v)}.$$

For finite matrices such that F has identical rows, it holds for all k that

$$||\mathrm{R}(k-1)||_v \leq c^{\mathbf{1}}_{A,B,C}||F((A-B)C)^k||_{\mathbf{1}},$$

where $\mathbf{1}$ *denotes the unit vector.*

Note that $||F||_v < \infty$ in the finite dimensional case. In [11] it has been shown that for finite matrices it holds that (i) to (v) are equivalent in Lemma 2.11.

3 Series Representation for Denumerable Markov Chains

This section presents series expansions for continuous-time Markov chains with denumerable state-space. The overall formula is derived in Section 3.1. In Section 3.2, sufficient conditions for the convergence of the series are presented. An application to sensitivity analysis is provided in Section 3.3. Eventually, we will present some new results on bounds on perturbations of Markov chains in Section 3.4.

3.1 The Exchange Formula

Let us consider now a Markov process $\mathcal{X}^*$ uniquely determined by its infinitesimal generator Q^*. To compute the corresponding ergodic matrix Π^*, we derive its series expansion based on the generator Q, the ergodic matrix Π and the deviation matrix D of a process for which these are well-known or easily calculated. By adding Q^*D to the equation put forward in Lemma 2.2 (ii), we get

$$(Q^* - Q)D = Q^*D + I - \Pi.$$

Multiplying this equation with Π^* yields

$$\Pi^*(Q^* - Q)D = \Pi^*Q^*D + \Pi^* - \Pi^*\Pi.$$

Which can be simplified by applying Lemma 2.1 (i) and (ii) so that we obtain

$$\Pi^*(Q^* - Q)D = \Pi^* - \Pi.$$

This can be written as

$$\Pi^* = \Pi + \Pi^*(Q^* - Q)D. \tag{9}$$

Inserting (9) into its right side yields

$$\Pi^* = \Pi + \Pi(Q^* - Q) + \Pi^*\big((Q^* - Q)D\big)^2. \tag{10}$$

From inserting now (9) into the right side of (10) we get

$$\Pi^* = \Pi + \Pi(Q^* - Q)D + \Pi\big((Q^* - Q)D\big)^2 + \Pi^*\big((Q^* - Q)D\big)^3.$$

By repeating this step n times we obtain

$$\Pi^* = \Pi \sum_{k=0}^{n} \big((Q^* - Q)D\big)^k + \Pi^* \big((Q^* - Q)D\big)^{n+1}. \tag{11}$$

We call the presentation in (11) the *continuous time exchange formula*. We separate (11) into the *series approximation of degree n*

$$H(n) \stackrel{\text{def}}{=} \Pi \sum_{k=0}^{n} \big((Q^* - Q)D\big)^k$$

and *remainder term*

$$R(n) \stackrel{\text{def}}{=} \Pi^* \big((Q^* - Q)D\big)^{n+1}.$$

The representation in (11) provides a scheme for approximately computing Π^*, provided that $R(n)$ tends to 0 as n tends to infinity. A sufficient condition for this is that $||((Q^* - Q)D)^n||_v$ tends to zero as n tends to infinity. Moreover, provided this condition holds, then for any f, such that $f(x) \leq cv(x)$ for all $x \in S$ and some c (in formula: $||f||_v < \infty$), it holds that $|\pi^* f - H(n)f|$ converges to zero as n tends to infinity.

A general sufficient condition for $||R(n)||_v$ to converge to zero is that D exists and that Q^* and Q are so close in v-norm sense that $||(Q^* - Q)D||_v = \rho < 1$. For example, according to Example 2.9, the deviation matrix for the M/M/∞ system exists and choosing Q^* close to Q will guarantee convergence of $||R(n)||_v$ to zero. Unfortunately, in applications the Q^* of interest typically fails to be close to Q in the above sense.

When we use the exchange formula for numerical purposes in Section 4, we will assume that a finite number N and constant $\delta_N < 1$ exist such that

$$\|((Q^* - Q)D)^N\|_v < \delta_N. \tag{12}$$

As we will show, condition (12) implies that $|\pi^* f - H(n)f|$ tends to zero as n tends to infinity at an exponential rate for any f with $||f||_v < \infty$.

Remark 3.1 *Provided that Q and Q^* are uniformizable with rate μ and μ^*, respectively, the effect of switching from Q to Q^* on the stationary distribution can alternatively be expressed via the corresponding subordinate chains. Let $\eta \geq \max(\mu, \mu^*)$. Recall that the stationary distributions of the subordinate chain and the continuous time chain coincide: $\Pi_\eta = \Pi$ and $\Pi^*_\eta = \Pi^*$. Inserting this together with (6) into (11) yields the following alternative expansion*

$$\Pi^* = \Pi \sum_{k=0}^{n} \big((P^*_\eta - P_\eta)D_\eta\big)^k + \Pi^* \big((P^*_\eta - P_\eta)D_\eta\big)^{n+1}. \tag{13}$$

We call the presentation in (13) the suboridnate exchange formula*, and we denote by*

$$H_\eta(n) \stackrel{\text{def}}{=} \Pi \sum_{k=0}^{n} \big((P^*_\eta - P_\eta)D_\eta\big)^k$$

for the subordinate series approximation of degree n *and by*

$$R_\eta(n) \stackrel{\text{def}}{=} \Pi^* \big((P^*_{\eta^*} - P_\eta)D_\eta\big)^{n+1}$$

the remainder term*. While the continuous-time exchange formula holds for continuous-time Markov chains with conservative generator matrix, the subordinate exchange formula only applies to uniformizable chains, which excludes, for example, the M/M/∞ system.*

In order to be able to apply the exchange formula in (11), one requires sufficient conditions such that $R(n)$ tends to zero as n tends to infinity. Such conditions are hard to get in the general case (i.e., for $R(n)$ without uniformization). A detailed discussion of sufficient conditions for convergence of $H(n)$ towards Π^* will be presented in Section 3.2.

3.2 Convergence of $H(n)$

For practical purposes it is important that convergence of $H(n)$ towards Π occurs at geometric rate. Since the geometric rate with which the series converges is only in special cases computable, we propose to find the convergence rate in an iterative way. The key observation is that if there exists N such that $\|((Q^* - Q)D)^N\|_v < \delta_N$ for $\delta_N < 1$, then $H(k)$ converges at exponential rate of at least δ_N. We introduce the following condition.

(C) *There exists a finite number N such that we can find $\delta_N \in (0,1)$ which satisfies:*

$$\|((Q^* - Q)D)^N\|_v < \delta_N,$$

and we set

$$c^v_{\delta_N} \stackrel{\text{def}}{=} \frac{1}{1-\delta_N} \left\| \sum_{k=0}^{N-1} ((Q^* - Q)D)^k \right\|_v .$$

Note that **(C)** is in fact condition **[A,B,C]** for $A = Q^*$, $B = Q$ and $C = D$. The factor $c^v_{\delta_N}$ in condition **(C)** allows to establish an upper bound for the remainder term that is independent of Π^*.

Denote by $T(k) = \Pi((Q^* - Q)D)^k$ the kth element of the series in (11). The following results have been established in [11] and apply to denumerable chains as well. Applying Lemma 2.11 for $A = Q^*$, $B = Q$, $C = D$, $F = \Pi$ and $G = \Pi^*$ yields the following result for the continuous time exchange formula.

Lemma 3.2 *Under* **(C)** *it holds that for any $v \geq 1$ that*

(i) $\|R(k-1)\|_v \leq c^v_{\delta_N} \|T(k)\|_v$ *for any k.*

(ii) $\lim_{k\to\infty} H(k) = \Pi \sum_{n=0}^{\infty} ((Q^* - Q)D)^n = \Pi^*$.

(iii) *There exist $\delta \in (0,1)$ and $\rho \in (0,\infty)$ exist such that for any k*

$$||T(k)||_v \leq \rho ||\Pi||_v \delta^k$$

and the upper bound for $R(k)$ in (i) *converges geometrically fast towards zero.*

Condition **(C)** is of key importance and this gives rise to the question for which class of systems **(C)** holds. It has been been shown in [11] for finite discrete-time Markov chains that (the discrete-time counterpart of condition) **(C)** is equivalent to the convergence of $H(n)$ as n tends to infinity. Unfortunately, **(C)** is a stronger condition than v-norm ergodicity; see the example of a finite discrete-time Markov chain that is v-norm ergodic but fails to satisfy **(C)** in [11]. The counterexample in [11] is, however, a rather esoteric Markov chain and in applications we have so far encountered no system that violates **(C)**. In the following we will show that v-norm ergodicity of the subordinate chain is sufficient for convergence of $H_\eta(n)$ as n tends to infinity applied to appropriate powers of the subordinate chain. This result will hold without any condition of the type of **(C)**.

Let Q^* and Q be uniformizable with rate μ^* and μ, respectively, and let $\eta = \max(\mu^*, \mu)$. In the following we will establish for uniformizable chains sufficient conditions for the convergence of the sum on the righthand side in (13) for appropriate powers of P^*_η and P_η. The key condition is the following:

(C') *There exist finite numbers k and N such that we can find $\delta_N \in (0,1)$ which satisfies:*

$$\left\| \left(((P^*_\eta)^k - (P_\eta)^k) D_{\eta,k} \right)^N \right\|_v < \delta_N,$$

where $D_{\eta,k}$ is the deviation matrix associated with $(P_\eta)^k$ an set

$$c^v_{\delta_{N,\eta}} \stackrel{\text{def}}{=} \frac{1}{1-\delta_N} \left\| \sum_{k=0}^{N-1} (((P^*_\eta)^k - P^k_\eta) D_\eta)^k \right\|_v .$$

Note that **(C')** is in fact condition [**A,B,C**] for $A = (P^*_\eta)^k$, $B = (P_\eta)^k$ and $C = D_{\eta,k}$.

As we will show in the next theorem, v-norm ergodicity of P^*_η and P_η implies that condition **(C')** is satisfied for $k \geq 2$, and thereby yields convergence of $H_\eta(n)$ as n tends to infinity.

Theorem 3.3 *Let Q and Q^* be uniformizable, such that $P^*_{\mu^*}$ and P_μ are v-norm ergodic. Then, condition* **(C')** *is satisfied for $k \geq 2$ and it holds that*

$$\Pi^* = \Pi \sum_{n \geq 0} \left(((P^*_\eta)^k - P^k) D_{\eta,k} \right)^n,$$

for any $k \geq 2$.

Proof: By v-norm ergodicity, we have

$$||(P_\eta)^n - \Pi_\eta||_v \leq c\beta^n \,, \tag{14}$$

for $n \geq 1$, with $\beta \in (0,1)$ and $c \in (0,\infty)$. Note that P_η and $(P_\eta)^n$ have the same ergodic projector Π_η. Indeed, if $\Pi_\eta P_\eta = \Pi_\eta$, then $\Pi_\eta (P_\eta)^n = \Pi_\eta$ for any n. By computation,

$$\begin{aligned} \lim_{n\to\infty} D_{\eta,n} &= \lim_{n\to\infty} \sum_{k\geq 0} ((P_\eta)^{nk} - \Pi_\eta) \\ &= (I - \Pi) + \lim_{n\to\infty} \sum_{k\geq 1}^{\infty} ((P_\eta)^{nk} - \Pi_\eta), \end{aligned}$$

which implies that

$$\lim_{n\to\infty} ||D_{\eta,n} - (I - \Pi_\eta)||_v = \lim_{n\to\infty} \left\| \sum_{k\geq 1}^{\infty} ((P_\eta)^{nk} - \Pi_\eta) \right\|_v . \tag{15}$$

Since $(P_\eta)^n$ tends to Π_η as n tends to infinity it holds it holds that

$$\forall k : \quad \lim_{n\to\infty} ||(P_\eta)^{nk} - \Pi_\eta||_v = 0.$$

This together with (14) allows for applying Dominated Convergence to show that

$$\lim_{n\to\infty} \left\| \sum_{k\geq 1}^{\infty} ((P_\eta)^{nk} - \Pi_\eta) \right\|_v = 0,$$

and we arrive at

$$\lim_{n\to\infty} ||D_{\eta,n}||_v = ||I - \Pi_\eta||_v. \tag{16}$$

As n goes to infinity $(P^*_\eta)^n$ tends to Π^*_η in v-norm, which implies

$$\lim_{n\to\infty} ||((P^*_\eta)^n - (P_\eta)^n) D_{n,\eta}||_v = ||(\Pi^*_\eta - \Pi_\eta)(I - \Pi_\eta)||_v = ||\Pi^*_\eta - \Pi_\eta||_v.$$

Since $(\Pi^*_\eta - \Pi_\eta)^2 = 0$, the above equation implies

$$\lim_{n\to\infty} ||(((P^*_\eta)^n - (P_\eta)^n) D_{\eta,n})^k||_v = 0, \tag{17}$$

for all $k \geq 2$. Hence, applying (6) it follows that **(C')** holds for $k = 2$, which proves the first part of the theorem.

For the proof of the second part of the theorem, follows from Lemma 2.11 for $A = (P^*_\eta)^k$, $B = (P_\eta)^k$, $C = D_{\eta,k}$, $F = \Pi$ and $G = \Pi^*$. Hence, the fact that condition **(C')** holds for any $k \geq 2$ implies the desired convergence.

■

To summarize, **(C)** is a sufficient condition for $H(n)$ to converge as n tends to infinity and to guarantee that the v-norm of the remainder decreases at an geometrically rate. The following general result holds without any type **(C)** condition. Uniformizability together with v-norm ergodicity of the subordinate chains is sufficient for the subordinate series $H_\eta(n)$ to converge as n tends to infinity and for the remainder term to decrease at geometric rate, provided the series is developed for powers of the subordinate chains.

3.3 Sensitivity Analysis

Suppose that Q depends on some parameter $\theta \in (a,b) \subset \mathbb{R}$. For example, let Q be the infinitesimal generator of the continuous-time queue length process of the M/M/1 queue with arrival rate λ and service rate μ; then, Q may be interpreted as a mapping of $\theta = \mu$ with $\theta \in (\lambda, \infty)$, in writing Q_θ. In perturbation analysis of Markov chains one is typically interested in the effect of a change in θ on the stationary distribution. More formally, let π_θ denote the stationary distribution associated with Q_θ, then perturbation analysis seeks to compute $d\pi_\theta/d\theta$. The often simple structure of Q_θ motivates the following condition

$$(K) \quad \forall i,j \in S: \quad \frac{1}{\Delta}|Q_{\theta+\Delta}(i,j) - Q_\theta(i,j)| \leq K.$$

Indeed, in the M/M/1 example the entries of Q_θ are linear mappings of θ. Given D_θ let d_θ denote the vector of absolute column sums for D_θ, i.e.,

$$d_\theta(j) = \sum_i |D_\theta(i,j)|, \quad j \in S. \tag{18}$$

Theorem 3.4 *Let condition (K) be satisfied. If the vector d_θ defined in (18) is finite, then π_θ is continuous at θ.*

If, in addition, Q_θ is elementwise differentiable, then

$$\pi'_\theta = \pi_\theta Q'_\theta D_\theta.$$

Proof: We apply (11) for $n = 1$ to $Q^* = Q_{\theta+\Delta}$ and $Q = Q_\theta$, then

$$\begin{aligned} \pi_{\theta+\Delta} - \pi_\theta &= \pi_{\theta+\Delta}(Q_{\theta+\Delta} - Q_\theta)D_\theta \\ &= \pi_\theta(Q_{\theta+\Delta} - Q_\theta)D_\theta + (\pi_{\theta+\Delta} - \pi_\theta)(Q_{\theta+\Delta} - Q_\theta)D_\theta. \end{aligned} \tag{19}$$

Note that

$$|((Q_{\theta+\Delta} - Q_\theta)D_\theta)(i,j)| \leq \sum_k |(Q_{\theta+\Delta})(i,k) - (Q_\theta)(i,k)|\, |D_\theta(k,j)|, \quad i,j \in S.$$

Applying condition (K) yields

$$|((Q_{\theta+\Delta} - Q_\theta)D_\theta)(i,j)| \leq |\Delta| K \sum_k |D_\theta(k,j)| \;\leq\; |\Delta| K d_\theta(j), \quad i \in S.$$

Let $\hat{D}_\theta$ denote the matrix with rows identical to d_θ, then

$$|(Q_{\theta+\Delta} - Q_\theta)D_\theta| \leq |\Delta| K \hat{D}_\theta.$$

Since $\hat{D}_\theta$ has identical rows, it holds that $\pi_\theta \hat{D}_\theta = d_\theta$, which is finite by assumption. Applying Dominated Convergence then yields

$$\lim_{\Delta \to 0} \pi_\theta (Q_{\theta+\Delta} - Q_\theta)D_\theta = 0.$$

Following the same line of argument it follows that

$$|(\pi_{\theta+\Delta} - \pi_\theta)(Q_{\theta+\Delta} - Q_\theta)D_\theta| \quad \leq \quad |\Delta|K|(\pi_{\theta+\Delta}\hat{D}_\theta + \pi_\theta\hat{D}_\theta),$$

where the fact that $\hat{D}_\theta$ has identical rows implies that $\pi_{\theta+\Delta}\hat{D}_\theta$ and $\pi_\theta\hat{D}_\theta$ are finite. As for the limit, we obtain

$$\lim_{\Delta\to 0}(\pi_{\theta+\Delta} - \pi_\theta)(Q_{\theta+\Delta} - Q_\theta)D_\theta = 0.$$

Hence, π_θ is continuous at θ, which proves the first part of the theorem.

We turn to the proof of the second part of the theorem. Elaborating on (19), we first show that

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\pi_\theta(Q_{\theta+\Delta} - Q_\theta)D_\theta = \pi_\theta Q'_\theta D_\theta.$$

This is a direct consequence of condition (K), which allows to apply Dominated Convergence and the existence of Q'_θ. As for the second term, continuity of π_θ and differentiability of Q_θ implies that

$$\begin{aligned}\lim_{\Delta\to 0}\frac{1}{\Delta}(\pi_{\theta+\Delta} - \pi_\theta)(Q_{\theta+\Delta} - Q_\theta)D_\theta &= \lim_{\Delta\to 0}(\pi_{\theta+\Delta} - \pi_\theta)\lim_{\Delta\to 0}\frac{1}{\Delta}(Q_{\theta+\Delta} - Q_\theta)D_\theta \\ &= 0,\end{aligned}$$

which proves the second claim.

■

Finiteness of d_θ is guaranteed for finite Markov chains. In the denumerable case, a sufficient condition for d_θ to be finite is $||D_\theta^T||_v < \infty$. It is worth noting that Theorem 3.4 applies without uniformization. To see this, note that the deviation matrix of M/M/∞ satisfies the condition put forward in Theorem 3.4 although the system fails to be uniformizable. The above perturbation formula extends the result in [4] to non uniformizable chains.

3.4 Perturbation Bounds

The study of perturbation bounds for Markov chains is known as *perturbation analysis of Markov chains* (PAMC) in the literature. PAMC is a classical topic in Markov chain literature and dates back to [23]. The key task in PAMC is the following: Provide bounds on the effect of perturbing P to P^* on the stationary behavior. The above problem can be phrased as follows: Can $\|\pi^* - \pi\|_v$ be approximated or bounded in terms of $\|P^* - P\|_v$?

In the following , we provide a simple bound for denumerable continuous-time Markov chains. By (11) it holds that

$$\Pi^* = \Pi + \Pi^*(Q^* - Q)D.$$

Subtracting Π on both sides and taking v-norms yields

$$||\Pi^* - \Pi||_v = ||\Pi^*(Q^* - Q)D||_v.$$

Provided that condition **(C)** holds, it follows from Lemma 3.2 that

$$\begin{aligned}||\Pi^* - \Pi||_v &= ||\Pi^*(Q^* - Q)D||_v \\ &\leq c^v_{\delta_N}||\Pi||_v\,||Q^* - Q||_v\,||D||_v.\end{aligned}$$

The following lemma lists some bounds on perturbations.

Lemma 3.5

(i) *Provided that* **(C)** *holds, we have*

$$||\Pi^* - \Pi||_v \leq c^v_{\delta_N}||\Pi||_v\,||Q^* - Q||_v\,||D||_v.$$

(ii) *Let Q be uniformizable with rate μ. Suppose that c and β, with $0 \leq \beta < 1$ exist such that $||(P_\mu)^n - \Pi||_v \leq c\beta^n$ for all n, (in words, P_μ is v-norm ergodic). If* **(C)** *holds, then*

$$||\Pi^* - \Pi||_v \leq c^v_{\delta_N} ||\Pi||_v \, ||Q^* - Q||_v \, \frac{c}{\mu(1-\beta)}.$$

(iii) *Let Q and Q^* be uniformizable with rate η. Suppose that c and β, with $0 \leq \beta < 1$ exist such that $||(P_\mu)^n - \Pi||_v \leq c\beta^n$ for all n, (in words, P_μ is v-norm ergodic). Then constants c_k exists such that for each $k \geq 2$ it holds that*

$$||\Pi^* - \Pi||_v \leq c_k ||\Pi||_v \, ||(P^*_\eta)^k - P^k_\eta||_v \, \frac{c}{\eta(1-\beta^k)}.$$

Proof: The first part of the lemma has already been shown. For the second part of the lemma follows by bounding the v-norm of D as in Lemma 2.6. For the third part of the lemma, note that Π_η is the ergodic projector of $(P_\eta)^k$ and Π^*_η is the ergodic projector of $(P^*_\eta)^k$, which yields

$$\Pi^* - \Pi = \Pi^*((P^*_\eta)^k - (P_\eta)^k) D_{\eta,k},$$

where

$$R_{\eta,k}(0) = \Pi^*((P^*_\eta)^k - (P_\eta)^k) D_{\eta,k}.$$

Applying Lemma 2.11 for $A = (P^*_\eta)^k$, $B = (P_\eta)^k$, $C = D_{\eta,k}$, $F = \Pi$ and $G = \Pi^*$ yields

$$||R_{\eta,k}(0)||_v \leq c^v_{\delta_{N,k}} ||\Pi||_v \, ||(P^*_\eta)^k - (P_\eta)^k D_{\eta,k}||_v,$$

and we arrive at

$$||\Pi^* - \Pi||_v \quad \leq \quad c^v_{\delta_{N,k}} ||\Pi||_v \, ||(P^*_\eta)^k - (P_\eta)^k||_v \, ||D_{\eta,k}||_v. \tag{20}$$

Bounding the v-norm of $D_{\eta,k}$ as in Lemma 2.6 yields

$$||D_{\eta,k}||_v \leq \frac{c}{\eta(1-\beta^k)}.$$

Inserting this into (20) proves the claim.

■

There exists an extensive literature on PAMC for finite state Markov chains; see, for example, [19, 24, 26] for recent results and the excellent overview in [21]. Best to our knowledge, Lemma 3.5 is a first result on a perturbation bound for a denumerable continuous-time Markov chain.

The statement in Lemma 3.5 is also of interest for the study of strong stable Markov chains (for details see [15]) as it states that any uniformizable continuous time Markov with v-norm ergodic subordinate chain is strongly stable provided that condition **(C)** holds.

4 Numerical Algorithm for Finite Markov Chains

Throughout this section we assume that S is finite. In the following we will show how the continuous-time exchange formula can be made fruitful for numerical approximations. The numerical algorithm is presented in Section 4.1. The performance of the algorithm is illustrated in Section 4.2 with numerical examples.

4.1 The Algorithm

The following result is the version of Lemma 2.11 for finite state Markov chains.

Lemma 4.1 *Under* **(C)** *it holds that for any* $v \geq 1$ *that*

(i) $\|R(k-1)\|_v \leq c^{\mathbf{1}}_{\delta_N} \|T(k)\|_{\mathbf{1}}$ *for any* k*, with* $\mathbf{1}$ *the vector with all elements equal to one.*

(ii) $\lim_{k\to\infty} H(k) = \Pi \sum_{n=0}^{\infty}((Q^* - Q)D)^n = \Pi.$

(iii) $\rho \in \mathbb{R}$ *and* $\delta < 1$ *exist such that* $\|((Q^* - Q)D)^k\|_v < \rho\delta^k$ *for all* k*.*

(iv) *For all* k *it holds that* $\|T(k)\|_v < \rho\delta^k ||\Pi||_v$*, with* ρ *and* δ *as in* (iii)*.*

In addition, condition **(C)** *and (ii) are equivalent.*

With Lemma 4.1 we arrive at the following numerical approach. First we search for N such that $1 > \delta_N \overset{\text{def}}{=} ||((Q^* - Q)D_N||_v$, which implies that **(C)** holds for N and δ_N. In words, we establish the minimal power of $((Q^* - Q)D)$ that yields geometrical convergence of $H(n)$. Then, we choose a precision ϵ up to which we want to approximate Π^*. The algorithm computes the elements $T(k)$ of $H(n)$ until our upper bound for $R(k)$, given by $c^v_{\delta_N}||((Q^* - Q)D)^{k+1}||_v$, drops below ϵ.

We can now describe an algorithm that yields an approximation for π^* with ϵ precision.

Algorithm 1

Chose precision $\epsilon > 0$*. Set* $k = 1$*,* $T(1) = \Pi(Q - P)D$ *and* $H(0) = \Pi$*.*

Step 1: Find N *such that* $||((Q^* - Q)D)^N||_v < 1$*. Set* $\delta_N = ||((Q^* - Q)D)^N||_v$ *and compute*

$$c^v_{\delta_N} = \frac{1}{1-\delta_N} \left\| \sum_{k=0}^{N-1} ((Q^* - Q)D)^k \right\|_v .$$

Step 2: If

$$c^v_{\delta_N} \|T(k)\|_v < \epsilon,$$

the algorithm terminates and $H(k-1)$ *yields the desired approximation. Otherwise, go to step 3.*

Step 3: Set $H(k) = H(k-1) + T(k)$*. Set* $k := k+1$ *and* $T(k) = T(k-1)(Q^* - Q)D$*. Go to step 2.*

As shown in Lemma 4.1 the above algorithm terminates geometrically fast. In case S is finite, all norms are equivalent with respect to norm ergodicity and, without loss of generality, we take $v = \mathbf{1}$.

4.2 Example: Approximating the Retrial-Queue through Simple Queues

Objective of this section is to provide an algorithm for the ergodic distribution of a queueing process with impatient customers who recall after they hung up. The optimization of such a system was so far hindered by the fact that a representation of the stationary distribution was not available. But by applying the series approximation $\lim_{n\to\infty} H(n)$ introduced in section 3, we can approximate Π^*.

Let $\mathcal{X}^*$ be the ergodic queue-length process with states $(x_1, x_2)^t \in S \overset{\text{def}}{=} \mathbb{N}_0 \times \mathbb{N}_0$ where x_1 denotes the sum of customers in service and waiting in the queue and x_2 refers to the impatient customers intending to recall. We regard this model as an open Jackson network with two nodes. External arrivals - modeling first callers - enter the system with rate λ at the first node where they are served by c servers with each providing service at rate μ, where they wait for their service or abandon if their waiting time exceeds their exponentially-α distributed patience. Customers who hung up are considered to enter a second node - the orbit - which they leave by recalling after an exponentially-β distributed time. Therefore the first node is an M/M/c+M queue while the latter one is an M/M/∞ queue. Transition rates for $x, y \in S$ are given as follows

$$q^*_{x,y} = \begin{cases} \lambda & y = (x_1+1, x_2),\ x_1, x_2 \geq 0 \\ \min\{x_1, c\}\mu & y = (x_1-1, x_2),\ x_1 \geq 1, x_2 \geq 0 \\ \max\{x_1 - c, 0\}\alpha & y = (x_1-1, x_2+1),\ x_1 \geq 1, x_2 \geq 0 \\ x_2\beta & y = (x_1+1, x_2-1),\ x_1 \geq 0, x_2 \geq 1 \\ -(\lambda + \min\{x_1, c\}\mu + \max\{x_1 - c, 0\}\alpha + x_2\beta) & y = (x_1, x_2),\ x_1, x_2 \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

An overview of the system provides Figure 1.

Insert Figure 1 here.

Like stated before, a direct computation of the stationary distribution is not possible. Hence, we introduce a related process $\mathcal{X}$ with x_1 the number of customers waiting or being served in an M/M/c queue and x_2 the number of customers being served in an independently acting M/M/∞ queue. The arrival and service rates of the first queue remain the same as in the initial model while external customers enter the M/M/∞ queue at rate α and leave the system after being served at rate $x_2\beta$, see Figure 2.

Insert Figure 2 here.

Therefore it holds for the transition rates with $x, y \in S$

$$q_{x,y} = \begin{cases} \lambda & y = (x_1+1, x_2),\ x_1, x_2 \geq 0 \\ \min\{x_1, c\}\mu & y = (x_1-1, x_2),\ x_1 \geq 1, x_2 \geq 0 \\ \alpha & y = (x_1, x_2+1),\ x_1, x_2 \geq 0 \\ x_2\beta & y = (x_1, x_2-1),\ x_1 \geq 0, x_2 \geq 1 \\ -(\lambda + \min\{x_1, c\}\mu + \alpha + x_2\beta) & y = (x_1, x_2),\ x_1, x_2 \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

and we have the joint ergodic distribution

$$\pi_x = \left(\sum_{k=0}^{\infty} \frac{c^{\min\{c-k,0\}} \left(\frac{\lambda}{\mu}\right)^k}{\min\{k, c\}!} \right)^{-1} \frac{c^{\min\{c-x_1,0\}} \left(\frac{\lambda}{\mu}\right)^{x_1}}{\min\{x_1, c\}!} e^{-\frac{\alpha}{\beta}} \frac{\left(\frac{\alpha}{\beta}\right)^{x_2}}{x_2!}, \quad x \in S. \tag{23}$$

To receive the stationary distribution of our initial model we have to compute the remaining parts of $H(n)$ first. From (21) and (22) we get the entries of $(Q^* - Q)$ as follows

$$(Q^* - Q)_{x,y} = \begin{cases} -\alpha & y = (x_1, x_2+1),\ x_1, x_2 \geq 0 \\ \max\{x_1 - c, 0\}\alpha & y = (x_1-1, x_2+1),\ x_1 \geq 1, x_2 \geq 0 \\ -x_2\beta & y = (x_1, x_2-1),\ x_1 \geq 0, x_2 \geq 1 \\ x_2\beta & y = (x_1+1, x_2-1),\ x_1 \geq 0, x_2 \geq 1 \\ \alpha - \max\{x_1 - c, 0\}\alpha & y = (x_1, x_2),\ x_1, x_2 \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We illustrate Algorithm 1 with the following example. The arrival rate of external arrivals is $\lambda = 3$ and there are $c = 3$ servers each providing service at rate $\mu = 1$. Customers abandon with rate $\alpha = 0.5$ and customers who abandoned leave the orbit in order to recall with rate $\beta = 3$. Furthermore, we have assumed that the service center has a total capacity of $N_1 = 4$ and orbit of $N_2 = 6$. Customers that find upon arrival to the service center or the orbit no available space are lost. The performance measure we are interested in is the mean stationary number of customers at the service center, i.e., customers in orbit are not counted. More formally, let $w(x, y) = x$ for $x, y \in S$, our goal is to approximately compute $\pi^* w$. Figure 3 shows the absolute error in percentage for predicting $\pi^* w$ by $H(n)w$, i.e., the figure plots $|\pi^* w - H(n)|/\pi^* w$ as a mapping of n. The numerical value of $\pi w = 2.4967$.

Insert Figure 3 here.

Algorithm 1 terminates whenever $c^v_{\delta_N} \, \|T(n)\|_v$ drops below ϵ. For example, taking $\epsilon = 0.01$, the algorithm will compute $\pi^* w$ up to a precision of ± 0.01. The number of elements of $H(n)$ required for achieving this precision is illustrated in Figure 4, where the dotted line represents ϵ.

Insert Figure 4 here.

As can be seen from Figure 4, $H(13)$ yields the desired precision and increasing the precision to, say 0.005, would lead to $H(15)$. For the sake of completeness we state the values required in Algorithm 1. For the above example, we obtained $N = 11$, $\delta_N = 0.9179$ and $c^{\mathbf{1}}_{\delta_N} = 201.2311$.

As we have mentioned earlier, for finite state Markov chains, Algorithm 1 can also be formulated for the subordinate chain using the subordinate exchange formula. The algorithmic complexity of both algorithms will be identical, since for both algorithms the main numerical work is in computing Π and D.

5 Functional Series Expansion

In this section we present a functional version of the series expansion that allows to obtain Π as function of θ, where θ is a parameter of Q. The functional model is introduced in Section 5.1. The application to performance evaluation of finite state Markov chains is presented in Section 5.2. Eventually, we illustrate the performance of the algorithm in Section 5.3 with numerical examples.

5.1 The Basic Model

As has already been noted in Section 3.3, the elements of Q are typically linear mappings of the rates of the Markov process. Consider two generator matrices Q^* and Q and define

$$Q_\theta = \theta Q^* + (1 - \theta) Q, \quad \theta \in [0, 1].$$

Denote the stationary distribution associated with Q_θ by Π_θ and note that $\Pi^* = \Pi_1$ and $\Pi = \Pi_0$. The basic model implies that

$$Q_{\theta+\Delta} - Q_\theta = \Delta(Q^* - Q).$$

Inserting the above representation for $Q_{\theta+\Delta} - Q_\theta$ into (11) yields

$$\Pi_{\theta+\Delta} = \Pi_\theta \sum_{k=0}^{n} \Delta^k \big((Q^* - Q) D_\theta\big)^k \; + \; \Delta^{k+1} \Pi_{\theta+\Delta} \big((Q^* - Q) D_\theta\big)^{n+1},$$

with D_θ the deviation matrix associated with Q_θ. Provided that $R(n)$ tends to zero as n tends to infinity, one obtains

$$\Pi_{\Delta+\theta} = \Pi_\theta \sum_{k=0}^{\infty} \Delta^k \big((Q^* - Q) D_\theta\big)^k, \tag{24}$$

which is noticeably the Taylor series expansion of Π_θ developed at point θ. Taylor series expansions of Markov chains are topic of active research. A Taylor series expansion for irreducible discrete-time Markov chains on denumerable state space can be found in [5]. This result has been extended to v-norm ergodic discrete-time Markov chains on general state space in [10]. For discrete-time Markov chains on denumerable state space with several ergodic classes, a Taylor series expansion can be found in [1].

Developing the above equation at $\theta = 0$, yields

$$\Pi_\Delta = \Pi \sum_{k=0}^{n} \Delta^k \big((Q^* - Q)D\big)^k \; + \; \Delta^{k+1} \Pi_\Delta \big((Q^* - Q)D\big)^{n+1}. \tag{25}$$

We call the presentation in (11) the *continuous-time exchange formula*. We separate (11) into the *series approximation of degree n*

$$H_\Delta(n) \stackrel{\text{def}}{=} \Pi \sum_{k=0}^{n} \Delta^k \big((Q^* - Q)D\big)^k$$

and *remainder term*

$$R_\Delta(n) \stackrel{\text{def}}{=} \Pi^* \Delta^{n+1} \big((Q^* - Q)D\big)^{n+1}.$$

5.2 The Algorithm

For this section we assume that the state space S is finite. Following the line of thought put forward in Section 4, we obtain the following algorithm for approximating π_Δ with ϵ precision, where $T_\Delta(k) = \Delta^k \Pi((Q^* - Q)D)^k$ denotes the kth element in $H_\Delta(n)$.

Algorithm 2

Chose precision $\epsilon > 0$. Set $k = 1$, $T_\Delta(1) = \Delta\Pi(Q^* - Q)D$ and $H_\Delta(0) = \Pi$.

Step 1: Find N such that $||((Q^* - Q)D)^N||_{\mathbf{1}} < 1$. Set $\delta_N = ||((Q^* - Q)D)^N||_{\mathbf{1}}$ and compute

$$c_{\delta_N}^{\mathbf{1}} = \frac{1}{1 - \delta_N} \left\| \sum_{k=0}^{N-1} ((Q^* - Q)D)^k \right\|_{\mathbf{1}} .$$

Step 2: If

$$c_{\delta_N}^{\mathbf{1}} \, \|T_\Delta(k)\|_v < \epsilon,$$

the algorithm terminates and $H_\Delta(k-1)$ yields the desired approximation. Otherwise, go to step 3.

Step 3: Set

$$T_\Delta(k+1) = T_\Delta(k)(Q^* - Q)D$$

and $H_\Delta(k) = H_\Delta(k-1) + T_\Delta(k)$. Let $k := k+1$ and go to step 2.

The above algorithm is not guaranteed to yield the desired approximation, which stems from the fact that $|\Delta|$ may lay outside the radius of convergence of the series in (25). However, $1/\delta_N$ as computed by the above algorithm yields a lower bound for radius of convergence of $H_\Delta(n)$ to the true function and thus an indication of the maximal value of $|\Delta|$.

5.3 Example: Functional Dependence of the Retrial Queue on Impatience Rate

We illustrate Algorithm 2 with the following example. The arrival rate of external arrivals is $\lambda = 3$ and there are $c = 3$ servers with each providing service at rate $\mu = 1$. Customers who abandoned leave the orbit in order to recall with rate $\beta = 3$. Furthermore, we have assumed that the service center has a total capacity of $N_1 = 4$ and orbit of $N_2 = 6$. Customers that find upon arrival to the service center or the orbit no available space are lost. The performance measure we are interested in is the mean stationary number of customers at the service center, i.e., customers in orbit are not counted. More formally, let $w(x, y) = x$ for $x, y \in S$, our goal is to approximately compute $\pi^* w$. Let Q^* be the generator matrix of the finite state space version of the M/M/c+M system with abandonment and retrial with customers abandon with rate $\alpha^* = 0.9$ and let Q be the generator matrix of the same system except for the abandon rate which is set to $\alpha = 0.2$.

Insert Figure 5 here.

Figure 5 shows the absolute error in percentage for predicting $\pi^* w$ by $H_\Delta(n)w$, i.e., the figure plots $|\pi^* w - H_\Delta(n)|/\pi^* w$ as a mapping of $\Delta \in [0, 1.4]$ for $n = 4$. Note that the radius of convergence of $H_\Delta(n)$ is given by $\alpha + (\alpha^* - \alpha)\delta_N$. The values of auxiliary variables in Algorithm 1 are $N = 4$, $\delta_N = 0.7469$ and $c^{\mathbf{1}}_{\delta_N} = 7.2953$. Hence, $H_\Delta(n)w$ converges as n tends to ∞ to the true mean stationary queue length for $\Delta < 1/\delta_N = 1.3389$. From $\alpha = 0.2$, we conclude that $\pi_{\alpha'} w$ can be approximated by $H_\Delta(n)$ for all $\alpha' < 1/\delta_N + \alpha = 1.5389$. This range is indicated by the vertical dotted line in Figure 5.

Suppose that one is interested in the maximal abandon rate that will result in a stationary mean queue length of, say, 3.72. Then solving $3.72 = H_\Delta w$ yields $\Delta = 0.523$ and the maximal value for α is given by $\alpha' = 0.723$, see the arrows in Figure 5. Hence, α' solves the optimization problem $\max_\alpha \pi_\alpha w \leq 3.72$.

Figure 6 to Figure 8 show the decay in our bound on the remainder term as a mapping of n for various values of Δ. The horizontal line indicates the precision value $\epsilon = 0.01$. The minimal number n for which $H_\Delta(n)$ has to be evaluated in order to guarantee $|\pi_{\alpha+\Delta} w - H_\Delta(n)w| < \epsilon$ can be found to be the first integer such that the bound on the remainder drops below the horizontal line. As the figures illustrate, for $\Delta = 0.4$ we have $n = 5$, for $\Delta = 0.8$ we have $n = 7$ and for $\Delta = 1.2$ we have $n = 11$.

Insert Figure 6 to Figure 8 here.

A lower bound for the radius of convergence of $H_\Delta(n)$ is given by δ_N. Hence, decreasing δ_N increases the range for Δ for which we can be sure that $H_\Delta(n)$ converges to the correct value. This can be achieved by adjusting the computation of δ_N in Algorithm 2. For example, if a radius of convergence of at least, say, $r > 0$ is required, then Step 1 in Algorithm 1 has to be modified as follows:

Step 1(b) *Find N such that $||((Q^* - Q)D)^N||_v < 1/r$. Set $\delta_N = ||((Q^* - Q)D)^N||_v$ and compute*

$$c^v_{\delta_N} = \frac{1}{1 - \delta_N} \left\| \sum_{k=0}^{N-1} ((Q^* - Q)D)^k \right\|_v .$$

By (**??**), Step 1(b) will terminate in finite time for any $r > 0$. In other words, $H_\Delta(n)$ converges to the true function for any $\Delta \in \mathbb{R}$. Obviously Step 1(b) will require a longer time then Step 1 as the initial phase of $||((Q^* - Q)D)^n||_v$ before geometric decay occurs might be rather long. On the other hand, simply taking α^* larger, increases $||Q^* - Q||_v$ and will also decrease the performance of the algorithm. The analysis of the trade-off between increased radius of convergence and longer initial phase in Step 1(b) of the algorithm is topic of further research.

References

[1] E. Altman, K.E. Avrachenkov, and Núñez-Queija: Perturbation analysis for denumerable markov chains with application to queueing models. *Advances in Applied Probability*, 36:839–853, 2004.

[2] A.A. Borovkov and A. Hordijk: *On normed ergodicity of Markov chains*, Technical Report MI 2000-40 Leiden University, 2000.

[3] A.A. Borovkov and A. Hordijk: Characterization and sufficient conditions for normed ergodicity of Markov chains, *Advances in Applied Probability*, 36:227-242, 2004.

[4] Xi-Ren Cao, Han-Fu Chen: Perturbation realization, potentials, sensitivity analysis of Markov processes. *IEEE Transactions on Automatic Control*, 42:1382-1393, 1997.

[5] Xi-Ren Cao: The Maclaurin Series for performance unctions of Markov chains. *Advances in Applied Probability*, 30:676-692, 1998.

[6] G.E. Cho and C.D. Meyer: Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335:137–150, 2001.

[7] P. Coolen-Schrijner, E. van Doorn: The deviation matrix of a continuous-time Markov chain. *Probability in the Engineering and Informational Sciences*, 16:351-366, 2002.

[8] R. Dekker and A. Hordijk: Average, sensitive and Blackwell optimal policies in denumerable Markov decision chains with unbounded rewards. *Mathematics of Operations Research*, 13:395-421, 1988.

[9] R. Dekker, A.Hordijk and F.M. Spieksma: On the relation between recurrence and ergodicity properties in denumerable Markov decision chains. *Mathematics of Operations Research*, 19:539-559, 1994.

[10] B. Heidergott and A. Hordijk: Taylor series expansions for stationary Markov chains. *Advances in Applied Probability*, 35:1046-1070, 2003.

[11] B. Heidergott, A. Hordijk, M. van Uitert: Series Expansions for Finite-State Markov Chains. *Probability in the Engineering and Informational Sciences*, to appear, 2007.

[12] A. Hordijk and R. Dekker, *Average, sensitive and Blackwell optimal policies in denumerable Markov decision chains with unbounded rewards*, Report No. 83-36, Institue of Applied Mathematics and Computing Science, Leiden University, 1983.

[13] A. Hordijk and F.M. Spieksma: On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Advances in Applied Probability*, 24:343-376, 1992.

[14] A. Hordijk and A.A. Yushkevich: Blackwell optimality in the class of all policies in Markov decision chains with a Borel state space and unbounded rewards. *Mathematical Methods of Operations Research*, 50:421-448, 1999.

[15] N. Kartashov: *Strong Stable Markov Chains*. VSP, Untrecht, 1996.

[16] F.P. Kelly: *Reversibility and Stochastic Networks.* Wiley, Chichester, 1979.

[17] J.G. Kemeny, J.L. Snell: *Finite Markov Chains.* Princeton, New York, 1960.

[18] J.G. Kemeny, J.L. Snell: Potentials for denumerable Markov Chains. *Journal of Mathematical Analysis and Applications*, 3:196-260, 1960.

[19] S. Kirkland: Conditioning properties of the stationary distribution for a Markov chain. *Electronic Journal of Linear Algebra*, 10:1-15, 2003.

[20] S. Lipman: On dynamic programming with unbounded rewards. *Mangement Science*, 21:1225-1233, 1974/75

[21] G. Cho and C. Meyer: Comparision of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335:137-150, 2001.

[22] S.P. Meyn and R.L. Tweedie: *Markov Chains and Stochastic Stability*, Springer, London, 1993.

[23] E. Schweitzer: Perturbation theory and finite Markov chains. *Journal of Applied Probability*, 5:401-413, 1968.

[24] E. Solan, N. Vieille: Perturbed Markov chains. *Journal of Applied Probability*, 40:107-122, 2003.

[25] R. Syski: Ergodic potential. *Stochastic Processes and Their Application*, 7:311-336, 1978.

[26] A. Mitrophnov: Sensitivity and convergence of uniformly ergodic Markov chains. *Journal of Applied Probability*, 42:1003-1014, 2005.

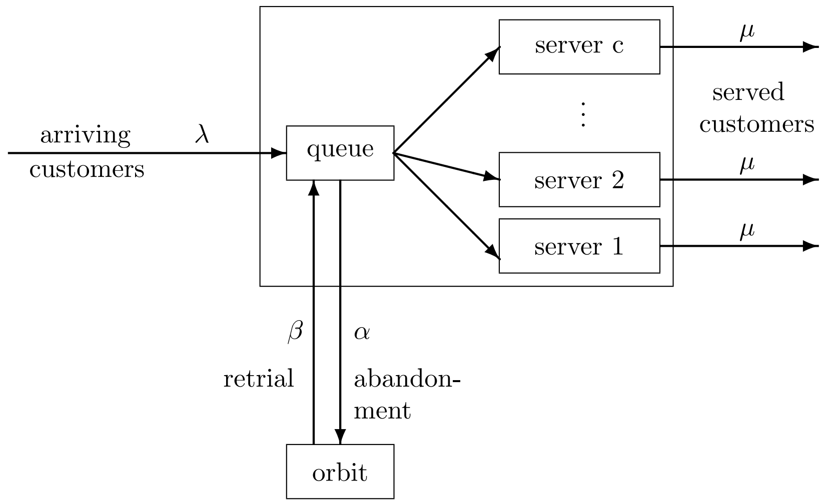

Figure 1: Structure of an M/M/c+M queueing system with abandonment and retrial

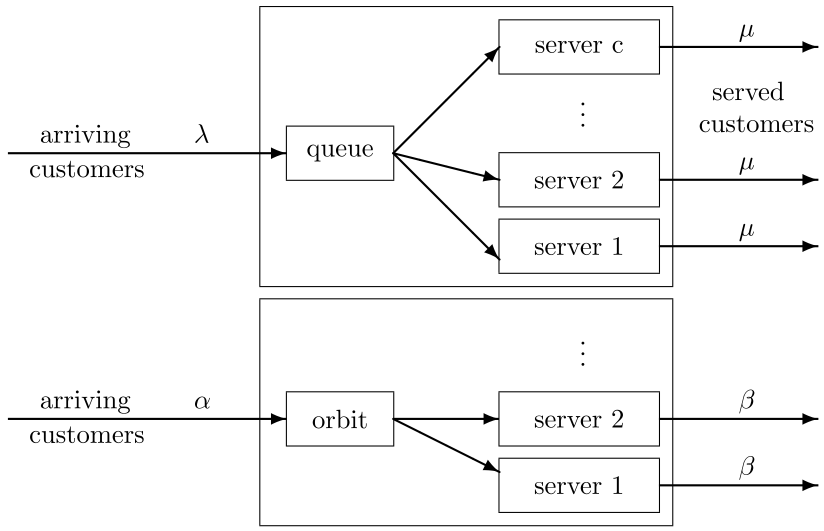

Figure 2: Structure of an M/M/c and M/M/∞ queueing system

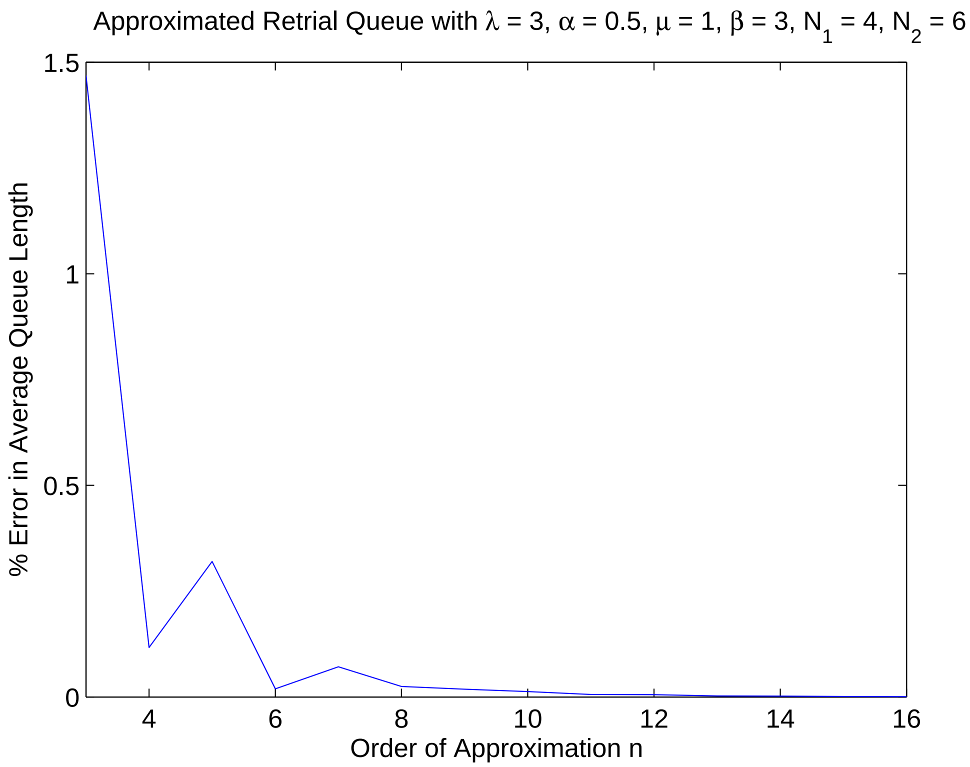

Figure 3: Error Percentage for Predicting the Stationary Queue Length via $H(n)$

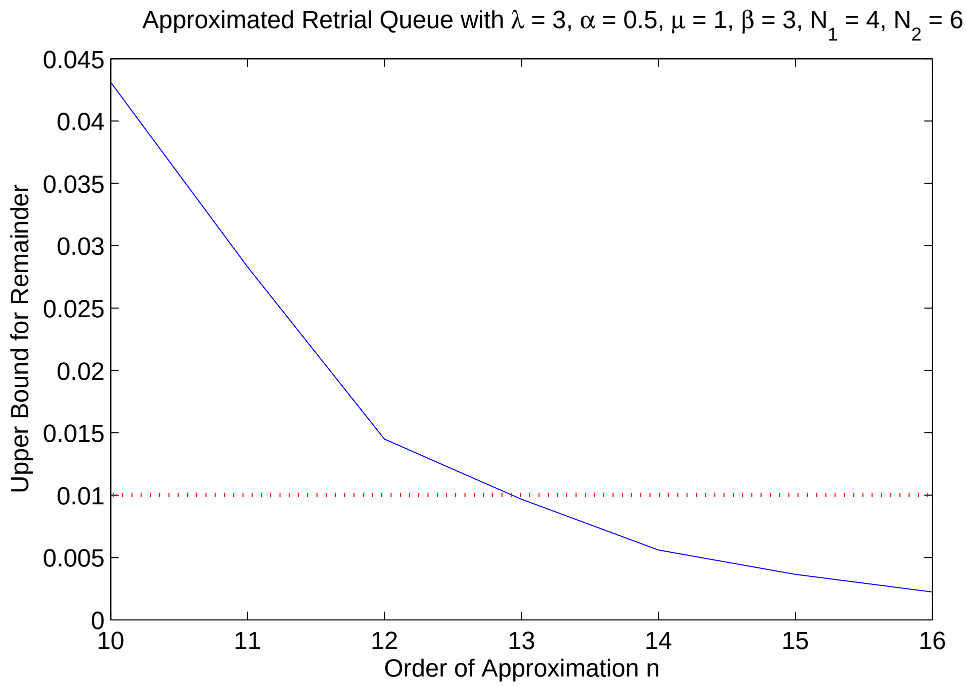

Figure 4: Bound on Remainder Term as Function of n

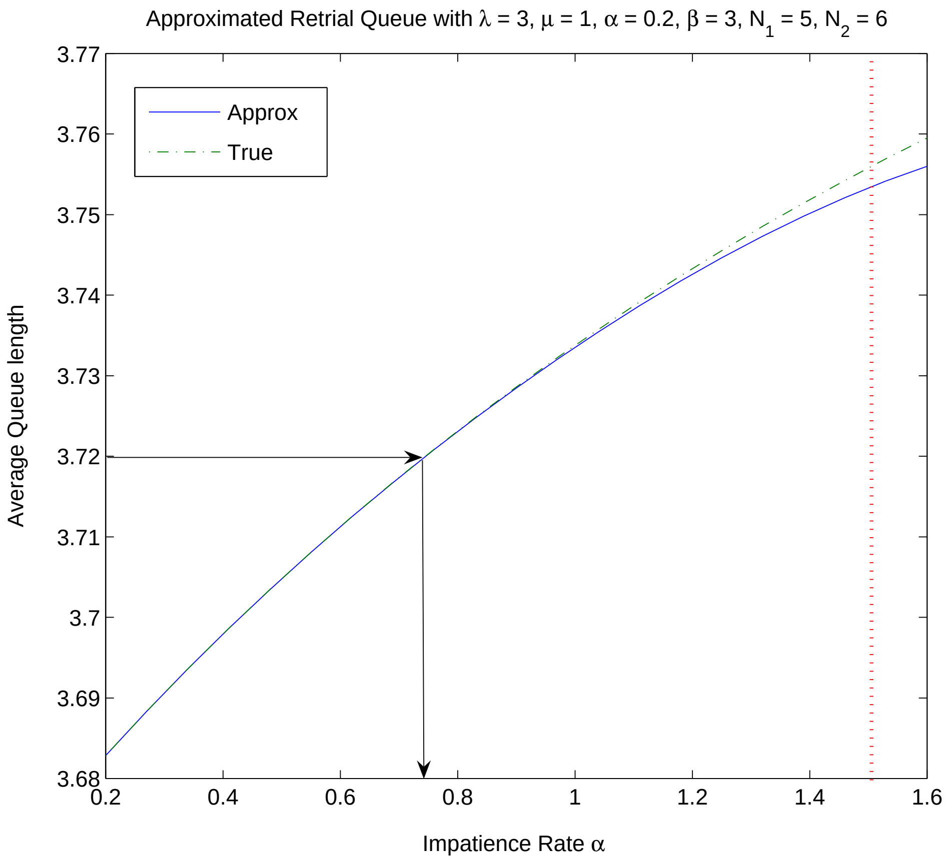

Figure 5: Error Percentage for Predicting the Stationary Queue Length via $H_\Delta(4)$

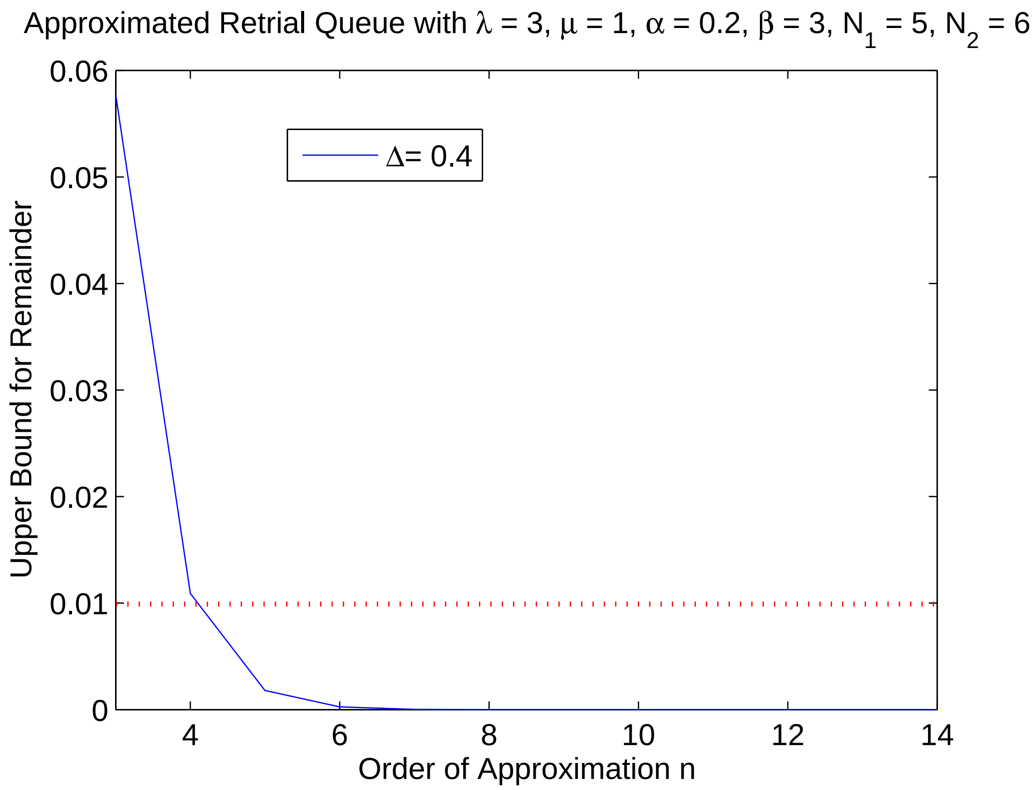

Figure 6: Bound on Remainder Term as Function of n

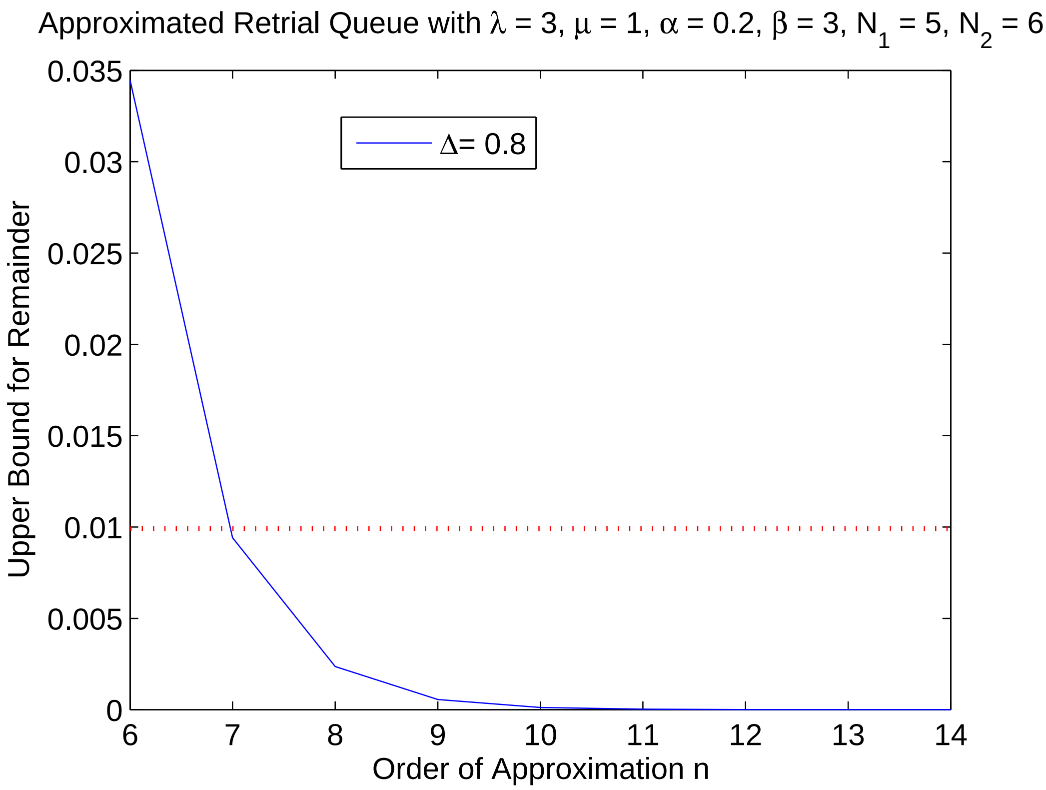

Figure 7: Bound on Remainder Term as Function of n

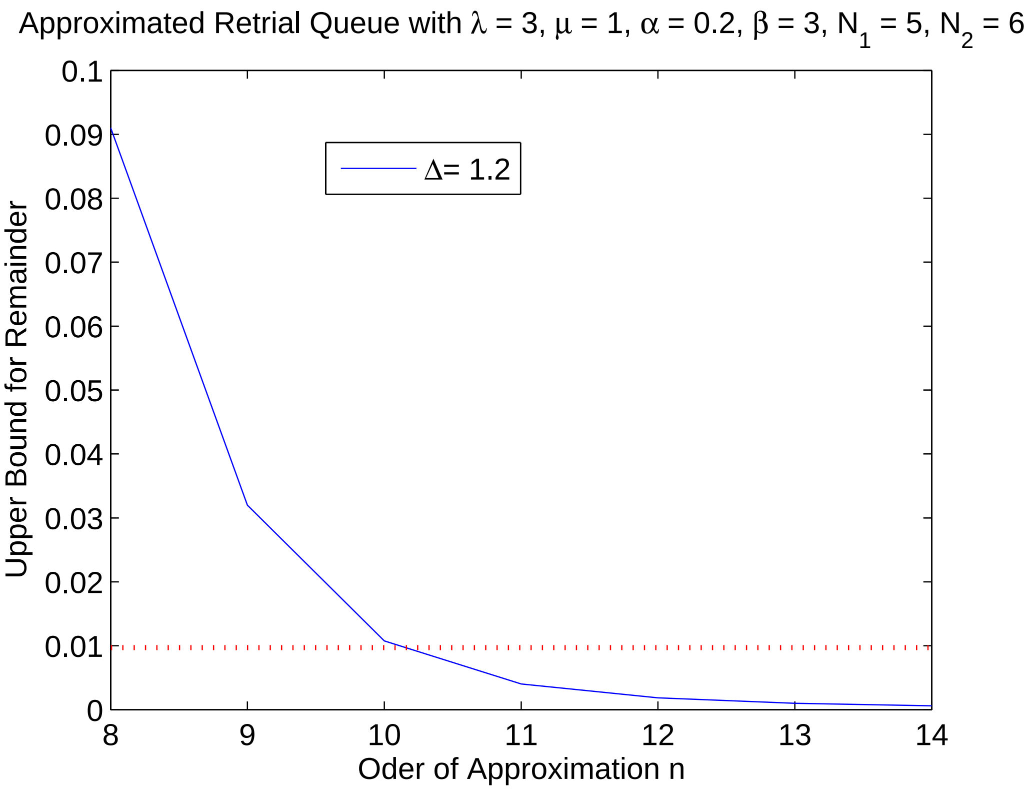

Figure 8: Bound on Remainder Term as Function of n