Stochastic Block-Monotonicity in the Approximation of the Stationary Distribution of Infinite Markov Chains

Haijun Li[1]
Department of Pure and Applied Mathematics
Washington State University
Pullman, WA 99164 U.S.A.
lih@haijun.math.wsu.edu

Yiqiang Q. Zhao[2]
Department of Mathematics and Statistics
University of Winnipeg
Winnipeg, Manitoba
Canada R3B 2E9
zhao@io.uwinnipeg.ca

April, 1997
Revised: March, 1999

[1]The research of this author is supported in part by the NSF grant DMI 9812994
[2]The research of this author is supported by a Research Grant from NSERC

Abstract

Markov chains, whose transition matrices reveal a certain type of block-structure, find many applications in various areas. Examples include Markov chains of $GI/M/1$ type and $M/G/1$ type, and more generally Markov chains of Toeplitz type. Some Markov chains without a block-repeating structure can be also included; for example, level-dependent-quasi-birth-and-death (LDQBD) processes. In analyzing this type of Markov chains, one may find that properties and/or probabilistic measures described or expressed by probability transition blocks from level to level often play a dominating role, while detailed transitions between states within the same level (block) are less important. In this paper, we introduce the concept of block-monitonicity and apply this notion to dealing with Markov chains possessing a block structure. A successful application in approximating stationary probability vectors of an infinite-state Markov chain is provided. We also hope that more applications of this concept can be exposed in the future.

AMS Subject Classification: 60J10, 60K25, 60E15
Key words and phrases: Stochastic ordering, block-monotonicity, block-augmentations, stationary distributions

1 Introduction

Markov chains, whose transition matrices possess a certain type of block structures, have been drawing people's attention for about two decades because of the increasing importance of their applications in many areas. In this area, the classical work on Markov chains of $GI/M/1$ type and $M/G/1$ type by Neuts (for example, Neuts (1981, 1989)) has been well recognized. Many other researchers have also made important contributions in this area. Updated references could be found in Neuts (1998). The scope in studying structured Markov chains has been broadened to including Markov chains of non-skip-free $GI/M/1$ and $M/G/1$ type, or $GI/M/1$ and $M/G/1$ type with multiple boundaries (Gail, Hantler and Taylor (1996, 1997)), LDQBD processes (Bright and Taylor (1996)), Markov chains of Toeplitz type (Zhao, Li and Braun (1998a, 1998b, 1999)). and block structured transition matrices with infinite blocks (Neuts (1998)). It is well-known that Markov chains of $GI/M/1$ type and $M/G/1$ type can be considered as a natural generalization of the classical $GI/M/1$ queue and $M/G/1$ queue. The transition matrix of the embbeded Markov chains for both $GI/M/1$ queue and $M/G/1$ queue is stochastic monotone. When a Markov chain has a monotone transition matrix, it often leads to new properties or/and makes analysis of the chain easier. For example, the discussion of approximating the stationary distribution of infinite Markov chains becomes easier and unified (Gibson and Seneta (1987)).

Specifically, let us consider a discrete time Markov chain with state space $\mathcal{N} = \{0, 1, 2, \ldots\}$ and transition matrix $P = (p_{i,j})$. Suppose that P is irreducible and positive recurrent. Let $\boldsymbol{\pi} = \{\pi_i\}$ be the unique stationary probability vector. In many situations, where the stationary distribution vector $\boldsymbol{\pi}$ of P is not analytically determined, numerical approximations are needed to obtain $\boldsymbol{\pi}$ (see Wolf (1980), Gibson and Seneta (1987), Heyman (1991), Zhao and Liu (1996), and the references therein). One of the ideas in computing the stationary distribution is the following: We first truncate the original transition matrix P into a finite matrix P_n by keeping all the $(n+1) \times (n+1)$ entries in the northwest corner of P. We then add appropriate numbers to entries of P_n to make P_n stochastic in some convenient way (this procedure is called *augmentation*). Denote the resulting matrix by $\tilde{P}_n$ say, and solve the finite system $\boldsymbol{\pi}^{(n)} \tilde{P}_n = \boldsymbol{\pi}^{(n)}$ and $\sum_{i=0}^{n} \boldsymbol{\pi}_i^{(n)} = 1$. Under certain conditions on the Markov chain, the stationary distribution vector $\boldsymbol{\pi}$ can be approximated by $\{\boldsymbol{\pi}^{(n)}, n \geq 1\}$ in the sense that for each fixed i,

$$\lim_{n \to \infty} \boldsymbol{\pi}_i^{(n)} = \boldsymbol{\pi}_i. \tag{1.1}$$

Seneta (1980) presented a necessary and sufficient condition for (1.1) to be true, that is, for any augmentation $\tilde{P}_n$, the sequence $\{\boldsymbol{\pi}^{(n)}\}$ is tight (see later for a definition or 37 of Billingsley (1969)) if and only if (1.1) holds. It was proved later in Gibson and Seneta (1987) that if P is stochastic monotone (that is, $\sum_{j=n}^{\infty} p_{i,j} \leq \sum_{j=n}^{\infty} p_{i+1,j}$ for any n and i), then $\{\boldsymbol{\pi}^{(n)}\}$ is tight. Therefore, (1.1) holds. This work provides us a unified treatment for the convergence of a sequence of finite-state Markov chains to an infinite-state Markov chain.

However, in many cases, P is not stochastic monotone, but it has a certain kind of block property when P is partitioned properly into submatrices (blocks). For example, the transition matrix P of a Markov chain of $GI/M/1$ type or $M/G/1$ type, including the case with multiple boundaries, a Markov chain of Toeplitz type, and some of LDQBD processes. For matrices with a block structure, it seems intuitively natural to augment the truncated matrix block-wise and analyze the systems by treating blocks as entries. In the literature, to deal with the convergence in approximations, either a scalar augmentation is used (Grassmann and Heyman (1993)), or the method is case sensitive (Latouche (1993)), or some algorithm developed from the censored process (Latouche (1993), Grassmann and Heyman (1990), Bright and Taylor (1996), Zhao, Li and Braun (1998a,b)). The motivation of writing this paper is to generalize the concept of stochastic monitonicity to stochastic block-monotonicity, that will play a similar role in analyzing at least those matrices mentioned above to that the usual stochastic monitonicity does in analyzing stochastic monotone matrices. In this paper, we first introduce the notion of block-monitonicity and provide an alternative or unified way of proving the convergence defined in (1.1) when P is block-monotone.

The rest of this paper is organized as follows. In Section 2, we introduce the notion of stochastic block-monotonicity. The main approximation result is given in Section 3. Examples to show how to apply the main result on various queueing models, including some queues with non-repeating block rows, are discussed in Section 4. Throughout this paper, the terms ‘increasing’ and ‘decreasing’ mean ‘nondecreasing’ and ‘nonincreasing’ respectively.

2 Block-Monotonicity Property of Stochastic Matrices

In this section, we first introduce a notion of stochastic block-ordering for vectors partitioned into blocks. Based on this ordering, matrices partitioned into blocks might be compared. We then introduce the concept of block-monitonicity and discuss properties of stochastic block-monotone matrices.

Stochastic block-rodering is defined based on $\mathcal{F}$-orderings (see, for example, Li and Shaked (1994)). This definition includes enough stochastic matrices for study and also leads to a mathematically tractable analysis. Throughout the paper, row vectors will be denoted by lower case bold letters and column vectors by the transpose of row vectors. For example, $\boldsymbol{a}$ is a row vector and $\boldsymbol{a}^T$ is a column vector. A real function defined on $\mathcal{N} = \{0, 1, 2, \ldots, \}$ is writen as a row vector $\boldsymbol{f} = (f(0), f(1), \ldots)$.

Definition 2.1 Let $\boldsymbol{f}_i = (f_i(1), \ldots, f_i(m))$, $i = 1, 2, \ldots$, be row vectors of size m. The function $\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots)$ is said to be block-increasing with block size m if for each $i = 1, 2, \ldots$, $\boldsymbol{f}_i \leq \boldsymbol{f}_{i+1}$ element-wise, that is, $f_i(j) \leq f_{i+1}(j)$ for $j = 1, \ldots, m$.

It is obvious from Definition 2.1 that $\boldsymbol{f}$ is a block-increasing function with block size m if and only if $(f_1(j), f_2(j), f_3(j), \ldots,)$ is an increasing function for each of $j = 1, \ldots, m$. For example, let $\boldsymbol{a} = (2, 1)$ and $\boldsymbol{b} = (2, 3)$ be two sub-vectors of size 2, then $\boldsymbol{f} = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{b}, \boldsymbol{b}, \ldots) = (2, 1, 2, 3, 2, 3, 2, 3, \ldots)$ is a block-increasing function with block size 2. Note that $\boldsymbol{f}$ is not an increasing function. It is obvious that every increasing function is block-increasing with any block size m.

Let $\mathcal{F}_m$ be the set of all block-increasing functions with block size m. Note that $\mathcal{F}_1$ is the set of all increasing functions defined on $\mathcal{N}$. For a subclass $\mathcal{F} \subseteq \mathcal{F}_m$, we can generate an ordering relation on $\mathcal{P}(\mathcal{N})$, where $\mathcal{P}(\mathcal{N})$ is the set of all probability measures on $\mathcal{N} = \{0, 1, 2, \ldots\}$, expressed as row vectors $\boldsymbol{p} = (p(0), p(1), \ldots)$.

Definition 2.2 Let $\mathcal{F} \subseteq \mathcal{F}_m$ and let $\boldsymbol{p}$ and $\boldsymbol{q}$ be two probability vectors in $\mathcal{P}(\mathcal{N})$. $\boldsymbol{p}$ is said to be stochastically block-less than $\boldsymbol{q}$ with respect to $\mathcal{F}$ (denoted as $\boldsymbol{p} \leq_{\mathcal{F}} \boldsymbol{q}$) if

$$\boldsymbol{p}\boldsymbol{f}^T = \sum_{k \in \mathcal{N}} p(k) f(k) \leq \sum_{k \in \mathcal{N}} q(k) f(k) = \boldsymbol{q}\boldsymbol{f}^T, \tag{2.1}$$

for all block-increasing functions $\boldsymbol{f} \in \mathcal{F}$.

Clearly, the relation $\leq_{\mathcal{F}}$ defined by (2.1) is reflexive and transitive, but not necessarily a partial order because the following property need not hold: $\boldsymbol{p} \leq_{\mathcal{F}} \boldsymbol{q}$ and $\boldsymbol{p} \geq_{\mathcal{F}} \boldsymbol{q}$ together do not necessarily imply that $\boldsymbol{p} = \boldsymbol{q}$. The relation $\leq_{\mathcal{F}}$ defined by (2.1) becomes a partial order if $\mathcal{F}$ is a determining class (Billingsley 1968), which is not necessary to assume. We do assume, however, that $\mathcal{F}$ satisfies some closure properties in Section 3.

Recall that $\boldsymbol{p}$ is less than $\boldsymbol{q}$ in the usual stochastic order, denoted as $\boldsymbol{p} \leq_{st} \boldsymbol{q}$, if $\boldsymbol{p}\boldsymbol{f}^T \leq \boldsymbol{q}\boldsymbol{f}^T$ for all increasing functions $\boldsymbol{f}$. Since each increasing function is block-increasing, the usual stochastic comparison is a special case of the stochastic block-comparison.

Remark 2.3 In order to verify the stochastic block-comparison $\leq_{\mathcal{F}_m}$ defined in Definition 2.2, it is sufficient, as in the case of the usual stochastic order, that one only needs to verify whether or not (2.1) holds for all *nonnegative* block-increasing functions. The reason for this is the following. Suppose that $\boldsymbol{p}\boldsymbol{f}^T \leq \boldsymbol{q}\boldsymbol{f}^T$ for all nonnegative block-increasing functions of block size m. For any block-increasing function $\boldsymbol{f}$ of block size m that is bounded from below, or $f_i(k) \leq M < \infty$ for all i and k, $\boldsymbol{f} + \alpha\mathbf{1}$ is nonnegative for some positive constant α, where $\mathbf{1}$ is the vector of ones. Thus, $\boldsymbol{p}(\boldsymbol{f} + \alpha\mathbf{1})^T \leq \boldsymbol{q}(\boldsymbol{f} + \alpha\mathbf{1})^T$, which implies that $\boldsymbol{p}\boldsymbol{f}^T \leq \boldsymbol{q}\boldsymbol{f}^T$. For any block-increasing function $\boldsymbol{f}$ of block size m, we can approximate $\boldsymbol{f}$ element-wise by using a monotone sequence of block-increasing functions that are bounded from below, and hence it follows from the Lebesgue convergence theorem that (2.1) holds for any block-increasing function $\boldsymbol{f} \in \mathcal{F}_m$.

Example 2.4 Let $\mathcal{F} \subseteq \mathcal{F}_m$ be a subclass of $\mathcal{F}_m$ and let $\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \dots)$ be a probability vector partitioned into sub-vectors $\boldsymbol{p}_i$ of a common size m. $\mathbf{0}$ is the vector of zeros of size m.

1. Define $\boldsymbol{q} = (\sum_{i=1}^{k} \boldsymbol{p}_i,\, \boldsymbol{p}_{k+1}, \dots)$. For any block-increasing function $\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \dots,)$ with block size m, $(\sum_{i=1}^{k} \boldsymbol{p}_i)\boldsymbol{f}_1^T + \sum_{j \geq k+1} \boldsymbol{p}_j \boldsymbol{f}_{j-k+1}^T \leq \sum_{i=1}^{\infty} \boldsymbol{p}_i \boldsymbol{f}_i^T$. Thus, $\boldsymbol{q} \leq_{\mathcal{F}} \boldsymbol{p}$. In particular, $(\sum_{i=1}^{\infty} \boldsymbol{p}_i, \mathbf{0}, \dots) \leq_{\mathcal{F}} \boldsymbol{p}$.

2. For any block-increasing function $\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \dots,)$ with block size m, and any $k \geq 1$, we have $\boldsymbol{p}_1\boldsymbol{f}_1^T + \dots + \boldsymbol{p}_{k-1}\boldsymbol{f}_{k-1}^T + (\sum_{j \geq k} \boldsymbol{p}_j)\boldsymbol{f}_k^T \leq \sum_{j \geq 1} \boldsymbol{p}_j \boldsymbol{f}_j^T$. Therefore, for any $k \geq 1$,

$$(\boldsymbol{p}_1, \dots, \boldsymbol{p}_{k-1}, \sum_{j=k}^{\infty} \boldsymbol{p}_j, \mathbf{0}, \mathbf{0}, \dots) \leq_{\mathcal{F}} \boldsymbol{p}.$$

3. Denote $\boldsymbol{p}^{[k]} = (\mathbf{0}, \ldots, \mathbf{0}, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots,)$ where there are k vectors of zeros of size m before $\boldsymbol{p}_1$. For any block-increasing function $\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots,)$ with block size m, we have $\sum_{j\geq k+1} \boldsymbol{p}_{j-k}\boldsymbol{f}_j^T \leq \sum_{j\geq k+1} \boldsymbol{p}_{j-k}\boldsymbol{f}_{j+1}^T$, which implies that $\boldsymbol{p}^{[k]} \leq_{\mathcal{F}} \boldsymbol{p}^{[k+1]}$ for any $k \geq 0$.

We now define stochastic block-monotone matrices and introduce the notion of the stochastic block-comparison for matrices.

Definition 2.5 Let $\mathcal{F} \subseteq \mathcal{F}_m$. Suppose that $A = (A_{i,j})$ is a stochastic matrix, where all entries $A_{i,j}$ are sub-matrices of size $m \times m$. A is said to be stochastically block-monotone with respect to $\mathcal{F}$ if $\boldsymbol{p} \leq_{\mathcal{F}} \boldsymbol{q}$ implies that $\boldsymbol{p}A \leq_{\mathcal{F}} \boldsymbol{q}A$.

Corollary 2.6 *Let $\mathcal{F} \subseteq \mathcal{F}_m$. For a stochastic matrix A, if $A\boldsymbol{f}^T \in \mathcal{F}$ for all $\boldsymbol{f} \in \mathcal{F}$, then A is stochastically block-monotone with respect to $\mathcal{F}$.*

Proof: If $\boldsymbol{p} \leq_{\mathcal{F}} \boldsymbol{q}$ then for $A\boldsymbol{f}^T \in \mathcal{F}$, we have $\boldsymbol{p}A\boldsymbol{f}^T \leq \boldsymbol{q}A\boldsymbol{f}^T$. This implies that $\boldsymbol{p}A \leq_{\mathcal{F}} \boldsymbol{q}A$.

Definition 2.7 Let $\mathcal{F} \subseteq \mathcal{F}_m$, and let A and B be two stochastic matrices partitioned into blocks of size $m \times m$. A is said to be stochastically block-less than B with respect to $\mathcal{F}$, denoted by $A \leq_{\mathcal{F}} B$, if $\boldsymbol{a}_i \leq_{\mathcal{F}} \boldsymbol{b}_i$ for all i, where $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are i-th row vectors of A and B respectively.

Example 2.8 Let A_i, $i = 0, 1, \ldots$, be non-negative matrices of size $m \times m$. Assume that $\sum_{j=0}^{\infty} A_j$ is stochastic. Consider the following stochastic matrix

$$P = \begin{bmatrix} A_0 + A_1 & A_2 & A_3 & A_4 & \cdots \\ A_0 & A_1 & A_2 & A_3 & \cdots \\ & A_0 & A_1 & A_2 & \cdots \\ & & A_0 & A_1 & \cdots \\ & & & A_0 & \cdots \\ & & & & \ddots \end{bmatrix},$$

We show that P is block-monotone with respect to $\mathcal{F}$ by proving that $P\boldsymbol{f}^T \in \mathcal{F}_m$ for all $\boldsymbol{f} \in \mathcal{F}_m$. For this, let B_j denote the jth block row of P, which is a matrix of size $m \times \infty$. Using Example 2.4 (1), we have $B_1\boldsymbol{f}^T \leq B_2\boldsymbol{f}^T$ element-wise for all $\boldsymbol{f} \in \mathcal{F}_m$. From Example 2.4 (3), it follows that for any $j \geq 2$, $B_j\boldsymbol{f}^T \leq B_{j+1}\boldsymbol{f}^T$ element-wise for all $\boldsymbol{f} \in \mathcal{F}_m$. Therefore, $P\boldsymbol{f}^T$ is a block-increasing function with size m for all $\boldsymbol{f} \in \mathcal{F}_m$.

In Example 2.8, we show that a certain stochastic matrix with repeating block rows is stochastically block-monotone with respect to $\mathcal{F}_m$. To establish stochastic monotonicity of some stochastic matrices with non-repeating block rows, the following simple fact is useful (see Section 4.2 for some applications).

Lemma 2.9 *Let $A = (A_0, A_1, A_2, \dots,)$ and $B = (B_0, B_1, B_2, \dots,)$ be two $m \times \infty$ stochastic matrices where A_i and B_i are $m \times m$ matrices, $i \geq 0$. Let C be an $m \times m$ non-negative matrix, such that for some j, every entry of $B_j - C$ is non-negative and every entry of $B_{j+1} + C$ is less than or equal to 1. If $A \leq_{\mathcal{F}} B$ then*

$$A \leq_{\mathcal{F}} (B_0, B_1, \dots, B_j - C, B_{j+1} + C, B_{j+2}, \dots,). \tag{2.2}$$

Proof: In fact, (2.2) follows from the fact that $\sum_{i=0}^{\infty} A_i \boldsymbol{f}_i^T \leq \sum_{i=0}^{\infty} B_i \boldsymbol{f}_i^T + C(\boldsymbol{f}_{j+1} - \boldsymbol{f}_j)^T$ element-wise for any $\boldsymbol{f} = (\boldsymbol{f}_0, \boldsymbol{f}_1, \boldsymbol{f}_2, \dots,) \in \mathcal{F} \subseteq \mathcal{F}_m$. In particular,

$$(A_0, A_1, A_2, \dots,) \leq_{\mathcal{F}} (A_0, A_1, \dots, A_j - C, A_{j+1} + C, A_{j+2}, \dots,),$$

when $A_j - C$ is non-negative element-wise.

From Definition 2.5, the following results are also immediate.

Proposition 2.10 Let $\mathcal{F} \subseteq \mathcal{F}_m$. If A and B are stochastically block-monotone with respect to $\mathcal{F}$, so are

1. $\alpha A + (1 - \alpha)B$, $0 \leq \alpha \leq 1$;
2. AB; and
3. A^k, $k = 0, 1, 2, \dots$.

Proposition 2.11 If A and B are stochastic matrices with $A \leq_{\mathcal{F}} B$, and either A or B is stochastically block-monotone with respect to $\mathcal{F}$, then $A^k \leq_{\mathcal{F}} B^k$ for any positive integer k.

Proof: Suppose that B is stochastically block-monotone, and $A^{k-1} \leq_{\mathcal{F}} B^{k-1}$ for $k > 1$. It follows from Definitions 2.5 and 2.7 that

$$A^{k-1}B \leq_{\mathcal{F}} B^{k-1}B = B^k.$$

On the other hand, since $A \leq_{\mathcal{F}} B$, we have $A\boldsymbol{f}^T \leq B\boldsymbol{f}^T$ for any $\boldsymbol{f} \in \mathcal{F}$. Therefore, $(\boldsymbol{p}A)\boldsymbol{f}^T \leq \boldsymbol{p}B\boldsymbol{f}^T$ for any probability vector $\boldsymbol{p}$. This implies that $A^k = A^{k-1}A \leq_{\mathcal{F}} A^{k-1}B$. The case when A is stochastically block-monotone can be proved similarly. ∎

3 Approximation of the Stationary Distribution of an Infinite Block-Monotone Stochastic Matrix

In this section and the next, we provide an useful application and interesting examples of block-monotone matrices. We show that the stationary distribution of an infinite block-monotone stochastic matrix can be approximated by the stationary distribution of a finite block-augmented matrix. This generalizes the result of Gibson and Seneta (1987) to a much wider class of Markov chains that arises naturally in stochastic modeling. The concepts of block-monotonicity, block-augmentation, and the tightness of a set of probability measures play a key role here.

To explain block-augmentations to a matrix, let us first discuss block augmentations of a probability vector. Let $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,)$ be a probability vector partitioned into blocks $\boldsymbol{a}_i$ of size m, where $\boldsymbol{a}_i = (a_i(1), \ldots, a_i(m))$. Define ${}_{(n)}\boldsymbol{a} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ to be the vector formed from the first n blocks of $\boldsymbol{a}$. A block-augmentation to ${}_{(n)}\boldsymbol{a}$ is a probability vector ${}_{(n)}\tilde{\boldsymbol{a}} = ({}_{(n)}\tilde{\boldsymbol{a}}_1, {}_{(n)}\tilde{\boldsymbol{a}}_2, \ldots, {}_{(n)}\tilde{\boldsymbol{a}}_n)$ such that

$$ {}_{(n)}\tilde{\boldsymbol{a}} \geq {}_{(n)}\boldsymbol{a} \ \text{ element-wise and } \sum_{i=1}^{n} {}_{(n)}\tilde{\boldsymbol{a}}_i(j) = \sum_{i=1}^{\infty} \boldsymbol{a}_i(j) \tag{3.1} $$

for all $j = 1, 2, \ldots, m$. Note that using Fatou's lemma, (3.1) implies that ${}_{(n)}\tilde{\boldsymbol{a}} \to \boldsymbol{a}$ element-wise as $n \to \infty$. For example, the following linear block augmentation is a special case of the above procedure, for $1 \leq i \leq n$,

$$ {}_{(n)}\tilde{\boldsymbol{a}}_i = \boldsymbol{a}_i + \alpha_i \sum_{j>n} \boldsymbol{a}_j $$

where $(\alpha_1, \ldots, \alpha_n)$ is a probability vector.

We now introduce block-augmentations to a matrix. Let $P = (P_{i,j})$ be a stochastic matrix on $\mathcal{N}$, where entries $P_{i,j}$ are sub-matrices of size $m \times m$. For each $n \geq 1$, we denote by ${}_{(n)}P$ the northwest corner of P consisting of $n \times n$ blocks. A block-augmentation ${}_{(n)}\tilde{P}$ of ${}_{(n)}P$ is a stochastic matrix such that each row of ${}_{(n)}\tilde{P}$ is a block-augmentation of the corresponding row vector of ${}_{(n)}P$. In particular, let ${}_{(n)}Q = ({}_{(n)}Q_{i,j})$ such that

$$ {}_{(n)}Q_{i,j} = P_{i,j} \text{ for } 1 \leq i \leq n, 1 \leq j \leq n-1, \text{ and } {}_{(n)}Q_{i,n} = \sum_{j=n}^{\infty} P_{i,j} \text{ for } 1 \leq i \leq n. \tag{3.2} $$

${}_{(n)}Q$ is called the last column block-augmentation.

The block-augmentation coincides the scalar one if the block size $m = 1$. When $m > 1$, any block augmentation is also a scalar augmentation, but a scalar augmentation may not be a block-augmentation.

To establish our main results, we require that $\mathcal{F} \subseteq \mathcal{F}_m$ satisfies the following tightness or/and the closure property.

Definition 3.1 Let $\mathcal{F} \subseteq \mathcal{F}_m$ be a subclass of the block-monotone functions.

1. $\mathcal{F}$ is said to satisfy the tightness condition if $\mathcal{F}$ contains a sequence of infinitely many *standard block jump functions*, where a function $\boldsymbol{f}^n = (\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots)$ is standard block jump if $\boldsymbol{f}_i = \mathbf{0}$ for $i < n$ and $\boldsymbol{f}_i = \mathbf{1}$ for $i \geq n$.

2. $\mathcal{F}$ is said to satisfy the closure property of block-monotonicity if for any n,

$$\boldsymbol{f} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n, \boldsymbol{f}_{n+1}, \ldots) \in \mathcal{F} \text{ implies } \bar{\boldsymbol{f}}_n = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n, \boldsymbol{f}_n, \ldots) \in \mathcal{F}. \tag{3.3}$$

Note that $\mathcal{F}_m$ satisfies both properties described in Definition 3.1.

Recall that a set $\{\boldsymbol{\pi}_n, n \geq 1\}$ of probability measures on a metric space S is said to be *tight* if for every positive ϵ there exists a compact subset $K \subseteq S$ such that $\boldsymbol{\pi}_n(K) > 1 - \epsilon$ for all $n \geq 1$ (Page 37 of Billingsley 1968). If S is a Polish space (complete and separable metric space), then the tightness of $\{\boldsymbol{\pi}_n, n \geq 1\}$ is equivalent to the *relative compactness*, that is, any infinite subset of $\{\boldsymbol{\pi}_n, n \geq 1\}$ possesses an infinite subsequence of probability measures converging weakly to a proper probability distribution. The tightness is a crucial condition needed to establish the results on approximating the stationary distribution of an infinite Markov chain, as showed by the following lemma.

Lemma 3.2 [Seneta 1980] Let $\boldsymbol{\pi}$ be the unique stationary distribution of the transition matrix P of an infinite, irreducible and positive recurrent Markov chain, and let ${}_{(n)}\boldsymbol{\pi}$ be a stationary distribution of a scalar augmentation of the $n \times n$ northwest corner truncation of P. Then ${}_{(n)}\boldsymbol{\pi} \to \boldsymbol{\pi}$ element-wise as $n \to \infty$ if and only if $\{{}_{(n)}\boldsymbol{\pi}, n \geq 1\}$ is tight.

Seneta (1980) proved Lemma 3.2 for the column augmentation (this is, augmenting some specific column) of the $n \times n$ northwest corner truncation of P. It is easy to see that his result holds for any scalar augmentation (also see Gibson and Seneta 1987).

It is well known, and easy to verify that for a sequence of probability vectors $\{\boldsymbol{p}_n, n \geq 0\}$ and a probability vector $\boldsymbol{p}$, if $\boldsymbol{p}_n \leq_{st} \boldsymbol{p}$ for any n, then $\{\boldsymbol{p}_n, n \geq 0\}$ is tight. For stochastic block-comparisons, we have,

Lemma 3.3 Let $\mathcal{F} \subseteq \mathcal{F}_m$ be a subclass of the block-monotone functions, and let $\{\boldsymbol{p}_n, n \geq 0\}$ be a sequence of probability vectors. If $\mathcal{F}$ satisfies the tightness condition described in Definition 3.1 (1), and for any n, $\boldsymbol{p}_n \leq_{\mathcal{F}} \boldsymbol{p}$ where $\boldsymbol{p}$ is a probability vector, then $\{\boldsymbol{p}_n, n \geq 0\}$ is tight.

Proof. Since $\boldsymbol{p}_n \leq_{\mathcal{F}} \boldsymbol{p}$ for any n, we have $\boldsymbol{p}_n \boldsymbol{f}^T \leq \boldsymbol{p}\boldsymbol{f}^T$ for any $\boldsymbol{f} \in \mathcal{F}$. Since $\mathcal{F}$ satisfies the property of the tightness, $\mathcal{F}$ contains an infinite sequence of standard block jump functions. Let $\{n_m, m \geq 1\}$ be such a sequence. Then, for all $n \geq 1$, and all $m \geq 1$,

$$\sum_{j=1}^{n_m} \boldsymbol{p}_n(j) \geq \sum_{j=1}^{n_m} \boldsymbol{p}(j).$$

For any small positive ϵ, select a finite n_k such that $\sum_{j=1}^{n_k} \boldsymbol{p}(j) > 1 - \epsilon$. Therefore, $\sum_{j=1}^{n_k} \boldsymbol{p}_n(j) > 1 - \epsilon$ for all n, which means that $\{\boldsymbol{p}_n, n \geq 0\}$ is tight. ■

Theorem 3.4 Let $\mathcal{F} \subseteq \mathcal{F}_m$. Suppose that P is positive recurrent and stochastically block-monotone with respect to $\mathcal{F}$ and that ${}_{(n)}\tilde{P} \geq {}_{(n)}P$ is any stochastic block-augmentation of ${}_{(n)}P$ with $n \times n$ blocks of size m. Let ${}_{(n)}\boldsymbol{\pi}$ be a stationary probability vector of ${}_{(n)}\tilde{P}$. Then ${}_{(n)}\boldsymbol{\pi} \to \boldsymbol{\pi}$ in the sense of (1.1) where $\boldsymbol{\pi}$ is the unique stationary probability vector of P.

Proof: Let ${}_{(n)}\tilde{P}_{i,j} = ({}_{(n)}\tilde{P}_{i,j})$, where ${}_{(n)}\tilde{P}_{i,j}$ is a sub-matrix of size $m \times m$ for $i, j = 1, \ldots, n$). We extend the definition of ${}_{(n)}\tilde{P}$ to a stochastic matrix on $\mathcal{N}$ by putting

$$ {}_{(n)}\tilde{P}_{i,1} = \sum_{j=1}^{\infty} P_{i,j}, \text{ for } i > n, \tag{3.4}$$

and zeros for all other block-enties not having been defined. The stationary probability vector ${}_{(n)}\boldsymbol{\pi}$ is extended to $\mathcal{N}$ by putting zeros for the elements beyond the n-th block.

The key is to show that ${}_{(n)}\tilde{P} \leq_{\mathcal{F}} P$. Let $\alpha_k = ({}_{(n)}\tilde{\boldsymbol{a}}_1, {}_{(n)}\tilde{\boldsymbol{a}}_2, \ldots, {}_{(n)}\tilde{\boldsymbol{a}}_n, \boldsymbol{0}, \ldots)$ be the k-th row of ${}_{(n)}\tilde{P}$, and $\beta_k = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n, \boldsymbol{a}_{n+1}, \ldots)$ be the k-th row of P, where ${}_{(n)}\tilde{\boldsymbol{a}}_j$, and $\boldsymbol{a}_j$ are the row vectors of size m. For $k > n$, we have $\alpha_k \leq_{\mathcal{F}} \beta_k$ from Example 2.3 (1). Now consider $k \leq n$. Let $\boldsymbol{b}_i = {}_{(n)}\tilde{\boldsymbol{a}}_i - \boldsymbol{a}_i$ for $i \leq n$. Since ${}_{(n)}\tilde{\boldsymbol{a}}_i \geq \boldsymbol{a}_i$ element-wise and

$\sum_{i=1}^{n} {}_{(n)}\tilde{\boldsymbol{a}}_i(j) = \sum_{i=1}^{\infty} \boldsymbol{a}_i(j)$, we have $\sum_{i=1}^{n} \boldsymbol{b}_i(j) = \sum_{i=n+1}^{\infty} \boldsymbol{a}_i(j)$. For any $\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots) \in \mathcal{F}$, we have

$$\sum_{i=1}^{n} \boldsymbol{b}_i \boldsymbol{f}_i^T \leq \sum_{i=1}^{n} \boldsymbol{b}_i \boldsymbol{f}_n^T = \sum_{i=n+1}^{\infty} \boldsymbol{a}_i \boldsymbol{f}_n^T \leq \sum_{i=n+1}^{\infty} \boldsymbol{a}_i \boldsymbol{f}_i^T,$$

which implies that $\alpha_k \leq_{\mathcal{F}} \beta_k$. Therefore, ${}_{(n)}\tilde{P} \leq_{\mathcal{F}} P$ with P being stochastically block-monotone. Thus, by Proposition 2.11, ${}_{(n)}\tilde{P}^k \leq_{\mathcal{F}} P^k$ for any $k \geq 1$. Hence we have ${}_{(n)}\boldsymbol{\pi} \leq_{\mathcal{F}} \boldsymbol{\pi}$ for any n. From Lemma 3.3, the sequence $\{{}_{(n)}\boldsymbol{\pi}\}$ is tight. Therefore, ${}_{(n)}\boldsymbol{\pi} \longrightarrow \boldsymbol{\pi}$ element-wise follows from Lemma 3.2. ■

If we in Theorem 3.4 only assume that P is a stochastic matrix and irreducible and $P \leq_{\mathcal{F}} R$ where R is positive recurrent and stochastically block-monotone with respect to $\mathcal{F}$, then the conclusion still holds. This is similar to the situation of the usual stochastic monotonicity (Gibson and Seneta 1987).

For the scalar case, Gibson and Seneta (1987) showed that the best, in the stochastic sense, approximation ${}_{(n)}\boldsymbol{\pi}$ to $\boldsymbol{\pi}$ is obtained by the last column augmentation. To obtain a similar result for stochastic block-monotonicity, we need the following lemma.

Lemma 3.5 Let $\mathcal{F} \subseteq \mathcal{F}_m$, and let $\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots)$ and $\boldsymbol{q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots)$ be probability vectors with blocks of size m. If $\boldsymbol{p} \leq_{\mathcal{F}} \boldsymbol{q}$ with $\mathcal{F}$ satisfying the closure property of block-monotonicity (3.3), then for any n, we have,

$$(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_{n-1}, \sum_{i=n}^{\infty} \boldsymbol{p}_i, \mathbf{0}, \mathbf{0}, \ldots) \leq_{\mathcal{F}} (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{n-1}, \sum_{i=n}^{\infty} \boldsymbol{q}_i, \mathbf{0}, \mathbf{0}, \ldots). \tag{3.5}$$

Theorem 3.6 Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots)$ be the stationary distribution of P, and let ${}_{(n)}\boldsymbol{\pi} = ({}_{(n)}\boldsymbol{\pi}_1, {}_{(n)}\boldsymbol{\pi}_2, \ldots)$ be a stationary distribution of the block augmentation ${}_{(n)}\tilde{P}$ of ${}_{(n)}P$, where Both $\boldsymbol{\pi}$ and ${}_{(n)}\boldsymbol{\pi}$ are partitioned into blocks of size m. Suppose that $\mathcal{F} \subseteq \mathcal{F}_m$ satisfies the tightness condition and the closure property of block-monotonicity. Let $P = (P_{i,j})$ be positive recurrent and stochastically block-monotone with respect to $\mathcal{F}$, and let ${}_{(n)}Q$ be the stochastic matrix formed from ${}_{(n)}P$ by augmenting the n-th block column (see (3.2)). If ${}_{(n)}Q$ has a unique stationary distribution vector ${}_{(n)}\boldsymbol{\nu} = ({}_{(n)}\boldsymbol{\nu}_1, {}_{(n)}\boldsymbol{\nu}_2, \ldots)$, then there exists an infinite sequence $1 \leq n_1 \leq n_2 \leq \ldots$ such that

$$0 \leq \sum_{i=1}^{n_k} ({}_{(n)}\boldsymbol{\nu}_i - \boldsymbol{\pi}_i)\mathbf{1}^T \leq \sum_{i=1}^{n_k} ({}_{(n)}\boldsymbol{\pi}_i - \boldsymbol{\pi}_i)\mathbf{1}^T \tag{3.6}$$

for every k, where $\mathbf{1}$ is the row vector of ones of size m. Therefore, ${}_{(n)}\boldsymbol{\nu}$ is at least as good an approximation to $\boldsymbol{\pi}$ as any ${}_{(n)}\boldsymbol{\pi}$ in the sense of (3.6).

Proof. Let us first extend the definition of ${}_{(n)}\tilde{P} = ({}_{(n)}\tilde{P}_{i,j})$, $i, j = 1, \ldots, n$, to a stochastic matrix on $\mathcal{N}$ as did in (3.4). We also extend the definition of ${}_{(n)}Q = ({}_{(n)}Q_{i,j})$, $i, j = 1, \ldots, n$, to a stochastic matrix on $\mathcal{N}$ by defining

$$
{}_{(n)}Q_{i,j} = \begin{cases} P_{i,j}, & \text{for } i > n, j = 1, 2, \ldots, n-1 \\ \sum_{k=n}^{\infty} P_{i,k}, & \text{for } i > n, j = n \\ 0, & \text{for blocks not having been defined.} \end{cases}
$$

Thus, ${}_{(n)}Q\boldsymbol{f}^T = P\bar{\boldsymbol{f}}_n^T$ where $\bar{\boldsymbol{f}}_n$ is defined by (3.3). The stationary probability vectors ${}_{(n)}\boldsymbol{\pi}$ and ${}_{(n)}\boldsymbol{\nu}$ are extended to $\mathcal{N}$ by making their entries beyond the n-th block zero. Since P is stochastically block-monotone and $\mathcal{F}$ satisfies (3.3), ${}_{(n)}Q$ is also stochastically block-monotone. From Theorem 3.4, ${}_{(n)}\tilde{P} \leq_{\mathcal{F}} P$, which implies that for any $\boldsymbol{f} \in \mathcal{F}$,

$$
{}_{(n)}\tilde{P}\boldsymbol{f}^T =_{(n)} \tilde{P}\bar{\boldsymbol{f}}_n^T \leq P\bar{\boldsymbol{f}}_n^T =_{(n)} Q\boldsymbol{f}^T
$$

element-wise, where $\bar{\boldsymbol{f}}_n$ is defined by (3.3). Thus, ${}_{(n)}\tilde{P} \leq_{\mathcal{F}} {}_{(n)}Q$. Also it follows from Example 2.4 (2) that ${}_{(n)}Q \leq_{\mathcal{F}} P$. Since ${}_{(n)}Q$ and P are both stochastically block-monotone, it follows from Proposition 2.11 that ${}_{(n)}\tilde{P}^k \leq_{\mathcal{F}} {}_{(n)}Q^k$ and ${}_{(n)}Q^k \leq_{\mathcal{F}} P^k$ for any $k \geq 1$. Hence we have

$$
{}_{(n)}\boldsymbol{\pi} \;\leq_{\mathcal{F}}\; {}_{(n)}\boldsymbol{\nu} \;\leq_{\mathcal{F}}\; \boldsymbol{\pi}. \tag{3.7}
$$

Since $\mathcal{F}$ contains an infinite sequence of standard block jump functions, there exists an infinite sequence $1 \leq n_1 \leq n_2 \leq \ldots$ such that

$$
\sum_{i=1}^{n_k} {}_{(n)}\boldsymbol{\pi}_i \mathbf{1}^T \geq \sum_{i=1}^{n_k} {}_{(n)}\boldsymbol{\nu}_i \mathbf{1}^T \geq \sum_{i=1}^{n_k} \boldsymbol{\pi}_i \mathbf{1}^T,
$$

which is equivalent to (3.6). ■

Note that if $\mathcal{F}$ contains all standard block jump functions of size m, then (3.6) becomes

$$
0 \leq \sum_{i=1}^{k} ({}_{(n)}\boldsymbol{\nu}_i - \boldsymbol{\pi}_i)\mathbf{1}^T \leq \sum_{i=1}^{k} ({}_{(n)}\boldsymbol{\pi}_i - \boldsymbol{\pi}_i)\mathbf{1}^T \tag{3.8}
$$

for each k.

There are two special classes of block-monotone functions that are worth discussing in more detail. First, consider the class $\mathcal{F}_m$ consisting of all block-increasing functions with block size m. The relationship between the stochastic block-order $\leq_{\mathcal{F}_m}$ and the usual stochastic order $\leq_{st}$ can be best described in the following proposition. For this, consider two probability vectors $\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots)$ and $\boldsymbol{q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots)$ partitioned into blocks with $\boldsymbol{p}_i = (p_i(1), p_i(2), \ldots, p_i(m))$ and $\boldsymbol{q}_i = (q_i(1), q_i(2), \ldots, q_i(m))$ for $i = 1, 2, \ldots$. Define, for each $j = 1, 2, \ldots, m$, the jth phase probability vectors for $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively, as

$$\boldsymbol{\alpha}_j = \left(1 - \sum_{k=1}^{\infty} p_k(j), p_1(j), p_2(j), \ldots\right)$$

and

$$\boldsymbol{\beta}_j = \left(1 - \sum_{k=1}^{\infty} q_k(j), q_1(j), q_2(j), \ldots\right).$$

It is obvious that if $\boldsymbol{p} \leq_{\mathcal{F}_m} \boldsymbol{q}$, then $\sum_{k=1}^{\infty} p_k(j) = \sum_{k=1}^{\infty} q_k(j)$, where $j = 1, 2, \ldots, m$. The following result is easy to verify.

Proposition 3.7 $\boldsymbol{p} \leq_{\mathcal{F}_m} \boldsymbol{q}$ if and only if $\boldsymbol{\alpha}_j \leq_{st} \boldsymbol{\beta}_j$ for all $j = 1, 2, \ldots, m$.

Therefore, the order $\leq_{\mathcal{F}_m}$ means a (vector-wise) usual stochastic comparison of two blocks of phase probability vectors.

To verify that a stochastic matrix A is a stochastically block monotone with respect to $\mathcal{F}_m$, the following characterization is useful.

Theorem 3.8 Let $A = (A_{i,j})$ be a stochastic matrix partitioned into sub-matrices $A_{i,j}$ of size $m \times m$. The following statements are equivalent.

1. A is a stochastically block-monotone matrix with respect to $\mathcal{F}_m$;
2. $A\boldsymbol{f}$ is block-increasing for any block-increasing function $\boldsymbol{f}$.

Proof: (2) trivially implies (1). To prove that (1) implies (2), consider, for each fixed k, the probability vector $\boldsymbol{p}_j^{(k)}$ partitioned into blocks of size m, in which the $j-$th element of the $k-$th block is 1 $(1 \leq j \leq m)$. Obviously, $\boldsymbol{p}_j^{(k)} \leq_{\mathcal{F}_m} \boldsymbol{p}_j^{(k+1)}$ for any $1 \leq j \leq m$, so from (1) we have $\boldsymbol{p}_j^{(k)} A \leq_{\mathcal{F}_m} \boldsymbol{p}_j^{(k+1)} A$ for all $1 \leq j \leq m$. Therefore, for any k,

$$(A_{k,1}, A_{k,2}, \ldots,) \leq_{\mathcal{F}_m} (A_{k+1,1}, A_{k+1,2}, \ldots,),$$

which means $A\boldsymbol{f}$ is block-increasing for any block-increasing function $\boldsymbol{f}$. ■

Since clearly $\mathcal{F}_m$ satisfies the tightness condition and the closure property of block-monotonicity, the stationary distribution vector of a stochastically $\mathcal{F}_m$-block monotone transition matrix can be approximated as described in Theorems 3.4 and 3.6.

The other interesting case involves the following class of block-monotone functions: $\mathcal{F}_+ = \{\boldsymbol{f} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots,) \in \mathcal{F}_m : \boldsymbol{a}_i = (a_i, a_i, \dots, a_i), i \geq 1\}$; that is, within each block of size m, the values of the function are same. Obviously $\mathcal{F}_+$ satisfies the tightness condition and the closure property of block-monotonicity, so the stationary distribution vector of a stochastically $\mathcal{F}_+$-block-monotone transition matrix can also be approximated as described in Theorems 3.4 and 3.6. It is interesting to notice that the stochastic $\mathcal{F}_+$-comparison of $\boldsymbol{p}$ and $\boldsymbol{q}$ is equivalent to the usual stochastic comparison of two probability vectors formed by adding up all the values in each block of $\boldsymbol{p}$ and $\boldsymbol{q}$ respectively.

Remark 3.9 We have so far discussed stochastic block-monotonicity for a stochastic matrix partitioned into equaled size blocks. Using the similar idea, it is possible to study block-monotone properties for certain transition matrix having different block sizes. For example, consider, similar to $\mathcal{F}_+$, the class of block increasing functions with different block sizes such that each block has the same function value. It is not difficult to see that by theorem-to-theorem modification of the results in Sections 2 and 3, Theorem 3.4 and Theorem 3.6 still hold for this class.

4 Applications

In this section, we present some examples of transition probability matrices, which are stochastically block-monotone, and some applications in stochastic modeling to illustrate the usefulness of the results developed in previous sections. Notice that all the matrices in these examples are not stochastically monotone in the usual sense, but they demonstrate some natural block-monotone structures. For all examples discussed here, the convergence of the block-augmented matrices has been established in the previous section. The block-augmentations include some methods often used by other researchers in the literature. For example, the last block-column augmentation and the block linear augmentation are two, which are easy to implement. The last block-column augmentation provides the best approximation in the sense of block-monitonicity. We also mention that in general, the last block-column augmentation does not provide the best approximation in the sense of l_1 norm; the censored Markov chain does, Zhao and Liu (1996). Although the concept of censoring is very useful in developing efficient algorithms (for example, the censored chain in (9) of Bright and Taylor, 1996, and Grassmann and Heyman, 1990), the censored chain itselt cannot be directly used as an approximation since the implementation is often difficult.

Therefore, algorithms are often developed based on the concept of censoring; for example, (17) and (18) in Bright and Taylor, 1996. (17) is the last block-column augmentation, and (18) is one of the linear block-augmentation.

4.1 Quasi-Birth-and-Death Processes

Consider a quasi-birth-and-death (QBD) process on the state space $\mathcal{E} = \{(i,j), i \geq 0, 1 \leq j \leq m\}$. The embedded Markov chain has the following transition matrix:

$$P = \begin{pmatrix} B_0 & A_0 & & & & \\ B_1 & A_1 & A_0 & & & \\ & A_2 & A_1 & A_0 & & \\ & & A_2 & A_1 & A_0 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where $B_0 + A_0$, $B_1 + A_1 + A_0$ and $A_2 + A_1 + A_0$ are stochastic matrices.

Example 4.1 Consider the $GI/M/c$ queue in which the arrivals form a PH-renewal process with interarrival time distribution of phase type with representation $(\boldsymbol{\alpha}, T)$ of order m (Page 88, Neuts (1981)). The service rate for each of c servers is denoted by μ. The $GI/M/c$ queue is a QBD process on the state space $\mathcal{E} = \{(i,j), i \geq 0, 1 \leq j \leq m\}$ where i denotes the number of customers in the system and j denotes the phase of the arrival process. The generator of the Markov chain is given by

$$Q = \begin{pmatrix} T & T^0A^0 & & & & & & \\ \mu I & T-\mu I & T^0A^0 & & & & & \\ & 2\mu I & T-2\mu I & T^0A^0 & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & c\mu I & T-c\mu I & T^0A^0 & & \\ & & & & c\mu I & T-c\mu I & T^0A^0 & \\ & & & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where I is the $m \times m$ identity matrix. By using uniformization $P = I + Q/c_{\max}$ with $c_{\max}$ equal or greater than the biggest absolute value of all diagonal elements of Q, one can easily convert the generator into the transition matrix of a discrete time Markov chain. Note that

P still has the same form as that of Q,

$$P = \begin{pmatrix} B & B^0A^0 & & & & & \\ \lambda I & B-\lambda I & B^0A^0 & & & & \\ & 2\lambda I & B-2\lambda I & B^0A^0 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & c\lambda I & B-c\lambda I & B^0A^0 & \\ & & & & c\lambda I & B-c\lambda I & B^0A^0 \\ & & & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where $B = I + T/c_{\max}$, $\lambda = \mu/c_{\max}$ and $B^0 = T^0/c_{\max}$. To verify that P is stochastically block-monotone with respect to $\mathcal{F}_m$, let D_i denote the $m \times \infty$ sub-matrix formed by m rows between the $((i-1)m+1)$st row and the (im)th row, $i \geq 1$. It follows from Example 2.4 (3) that $D_i \leq_{\mathcal{F}_m} D_{i+1}$ for $i \geq c+1$. To show that $D_i \leq_{\mathcal{F}_m} D_{i+1}$ for $2 \leq i \leq c$, let $\boldsymbol{p}^{i,k}$ be the kth row vector of D_i, $1 \leq k \leq m$, and $\alpha_j^{i,k}$ be the jth phase probability vector of $\boldsymbol{p}^{i,k}$, $1 \leq j \leq m$. We have,

$$\alpha_j^{i,k} = (1 - b_{k,j} - c_{k,j}, 0, \ldots, 0, 0, b_{k,j}, c_{k,j}, 0, \ldots,), j \neq k$$

where there are $i-1$ zeros between $1 - b_{k,j} - c_{k,j}$ and $b_{k,j}$, and

$$\alpha_k^{i,k} = (1 - b_{k,k} - c_{k,k}, 0, \ldots, 0, (i-1)\lambda, b_{k,k} - (i-1)\lambda, c_{k,k}, 0, \ldots,),$$

where there are $i-2$ zeros between $1 - b_{k,k} - c_{k,k}$ and $(i-1)\lambda$. Here $b_{k,j}$ is the (k,j)th entry of B and $c_{k,j}$ is the (k,j)th entry of B^0A^0. It is now a simple matter to verify that $\alpha_j^{i,k} \leq_{st} \alpha_j^{i+1,k}$ for all $j \neq k$ and $\alpha_k^{i,k} \leq_{st} \alpha_k^{i+1,k}$. By Proposition 3.7, $\boldsymbol{p}^{i,k} \leq_{\mathcal{F}_m} \boldsymbol{p}^{i+1,k}$, $1 \leq k \leq m$. Thus, $D_i \leq_{\mathcal{F}_m} D_{i+1}$ for $2 \leq i \leq c$. It can be also similarly verify that $D_1 \leq_{\mathcal{F}_m} D_2$.

Therefore, $D_i \leq_{\mathcal{F}_m} D_{i+1}$ for all i, which implies, from Theorem 3.8, that P is stochastically block-monotone with respect to $\mathcal{F}_m$. Hence, the stationary probability vector can be obtained by approximation from a stochastic block augmentation of P, and the last block-column augmentation provides the best approximation in the sense (3.6).

Example 4.2 We now consider a queueing model with one Poisson arrival stream with rate λ, c (≥ 2) identical exponential servers with service rate μ for each server and c

separate queues. All necessary independent conditions are assumed. Jockeying is allowed among queues. The jockeying rule is that the last customer in the longest queue will instantaneously move to the shortest queue as soon as the difference of customer numbers between the longest queue and the shortest queue is greater than 2. Consider the numbers of customers in all queues including those in service if any. The state space of the Markov chain is $\mathcal{E} = \{(i_1, i_2, \ldots, i_c), i_k \geq 0 \text{ for } 1 \leq k \leq c, \text{ and } |i_k - i_l| \leq 1 \text{ for } 1 \leq k, l \leq c\}$. States are partitioned into blocks according to the sum of numbers of customers in all queues. Without loss of generality, we assume that $\lambda + c\mu = 1$ (see Zhao (1990) for details). To illustrate the idea of finding the stationary probability vector of the Markov chain, let us consider the following transition matrix for $c = 3$ (the general case is similar),

$$P = \begin{pmatrix} A_{0,0} & Q_{0,1} & & & & & & & \\ A_{1,0} & A_{1,1} & Q_{1,2} & & & & & & \\ & A_{2,1} & A_{2,2} & Q_{2,3} & & & & & \\ & & Q_{0,-1} & Q_{0,0} & Q_{0,1} & & & & \\ & & & Q_{1,0} & Q_{1,1} & Q_{1,2} & & & \\ & & & & Q_{2,1} & Q_{2,2} & Q_{2,3} & & \\ & & & & & Q_{0,-1} & Q_{0,0} & Q_{0,1} & \\ & & & & & & \ddots & \ddots & \ddots \end{pmatrix}$$

where $A_{i,i}$ and $Q_{i,i}$ for $i = 0, 1, \ldots, c-1$ are matrices of size $\binom{c}{i} \times \binom{c}{i}$. All the sub-matrices are now described below. $A_{0,0} = (1 - \lambda)$, $Q_{0,1} = (\lambda/3, \lambda/3, \lambda/3)$, $Q_{0,-1} = (\mu, \mu, \mu)$, $Q_{0,0} = (0)$, $Q_{1,1} = Q_{2,2} = 0$, $A_{1,1} = diag(1 - (\lambda + \mu), 1 - (\lambda + \mu), 1 - (\lambda + \mu))$, $A_{2,2} = diag(1 - (\lambda + 2\mu), 1 - (\lambda + 2\mu), 1 - (\lambda + 2\mu))$, where $diag(a_1, a_2, a_3)$ denotes a 3×3 matrix with diagonal elements being a_1, a_2 and a_3, and zero otherwise.

$$A_{1,0} = \begin{pmatrix} \mu \\ \mu \\ \mu \end{pmatrix}, \quad A_{2,1} = \begin{pmatrix} \mu & \mu & 0 \\ \mu & 0 & \mu \\ 0 & \mu & \mu \end{pmatrix}, \quad Q_{1,0} = \begin{pmatrix} 3\mu \\ 3\mu \\ 3\mu \end{pmatrix},$$

$$Q_{2,1} = \begin{pmatrix} 1.5\mu & 1.5\mu & 0 \\ 1.5\mu & 0 & 1.5\mu \\ 0 & 1.5\mu & 1.5\mu \end{pmatrix}, \quad Q_{1,2} = \begin{pmatrix} \lambda/2 & \lambda/2 & 0 \\ \lambda/2 & 0 & \lambda/2 \\ 0 & \lambda/2 & \lambda/2 \end{pmatrix}, \quad Q_{2,3} = \begin{pmatrix} \lambda \\ \lambda \\ \lambda \end{pmatrix}.$$

The transition probabilities have the following property: from any state in a given block, the sum of transition probabilities over any block is a constant. Equivalently, the Markov chain is lumpable. Note also that P is not stochastically monotone in the usual sense (for example, the fourth row probability vector is not stochastically less than the fifth row probability vector).

We now regroup the blocks of P into a block matrix of block size $(2^c-1)\times(2^c-1)$ as follows

$$P = \begin{pmatrix} B_0 & C & & & \\ D & B_1 & C & & \\ & D & B_1 & C & \\ & & D & B_1 & C \\ & & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{4.1}$$

where

$$B_0 = \begin{pmatrix} A_{0,0} & Q_{0,1} & 0 \\ A_{1,0} & A_{1,1} & Q_{1,2} \\ 0 & A_{2,1} & A_{2,2} \end{pmatrix}, \quad B_1 = \begin{pmatrix} Q_{0,0} & Q_{0,1} & 0 \\ Q_{1,0} & Q_{1,1} & Q_{1,2} \\ 0 & Q_{2,1} & Q_{2,2} \end{pmatrix},$$

and

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ Q_{2,3} & 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & Q_{0,-1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

To show that P is stochastically block monotone in certain sense, we introduce the following class of block monotone functions. Define $\mathcal{F}_* = \{\boldsymbol{f} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots) : \boldsymbol{b}_i = (b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}, b_{i6}, b_{i7}), i \geq 1,$ such that $b_{i1} \leq b_{i2} = b_{i3} = b_{i4} \leq b_{i5} = b_{i6} = b_{i7}\} \subseteq \mathcal{F}_1$; that is, $\mathcal{F}_*$ consists of increasing functions of block size 2^c-1 such that each block is partitioned into sub-blocks with sizes $\binom{c}{0}$, $\binom{c}{1}$, and $\binom{c}{c-1}$ respectively, and the function takes on a constant value over each sub-block. Clearly $\mathcal{F}_*$ satisfies the tightness condition (does not satisfy the closure property of block increasingness, though).

Using the lumpable property, we now show that $P\boldsymbol{f} \in \mathcal{F}_*$ for any $\boldsymbol{f} \in \mathcal{F}_*$. For this, let $\boldsymbol{f} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots) \in \mathcal{F}_*$ where $\boldsymbol{b}_i = (b_{i1}, b_{i2}, b_{i2}, b_{i2}, b_{i3}, b_{i3}, b_{i3}), b_{i1} \leq b_{i2} \leq b_{i3}, i \geq 1$, and let $\boldsymbol{p}_j$ be the jth row vector of P. Noticing that $3\mu + \lambda = 1$, we have

$$\boldsymbol{p}_1\boldsymbol{f}^T = (1-\lambda)b_{11} + \lambda b_{12} \leq \boldsymbol{p}_2\boldsymbol{f}^T = \boldsymbol{p}_3\boldsymbol{f}^T = \boldsymbol{p}_4\boldsymbol{f}^T = \mu b_{11} + (1-\lambda-\mu)b_{12} + \lambda b_{13}$$

$$\leq \boldsymbol{p}_5\boldsymbol{f}^T = \boldsymbol{p}_6\boldsymbol{f}^T = \boldsymbol{p}_7\boldsymbol{f}^T = 2\mu b_{12} + (1-\lambda-2\mu)b_{13} + \lambda b_{21} \leq$$

$$\boldsymbol{p}_8\boldsymbol{f}^T = 3\mu b_{13} + \lambda b_{22} \leq \boldsymbol{p}_9\boldsymbol{f}^T = \boldsymbol{p}_{10}\boldsymbol{f}^T = \boldsymbol{p}_{11}\boldsymbol{f}^T = 3\mu b_{21} + \lambda b_{23}$$

$$\leq \boldsymbol{p}_{12}\boldsymbol{f}^T = \boldsymbol{p}_{13}\boldsymbol{f}^T = \boldsymbol{p}_{14}\boldsymbol{f}^T = 3\mu b_{22} + \lambda b_{31}.$$

Also for any $k \geq 1$,

$$\boldsymbol{p}_{k7+1}\boldsymbol{f}^T = 3\mu b_{k3} + \lambda b_{(k+1)2} \leq \boldsymbol{p}_{k7+2}\boldsymbol{f}^T = \boldsymbol{p}_{k7+3}\boldsymbol{f}^T = \boldsymbol{p}_{k7+4}\boldsymbol{f}^T = 3\mu b_{(k+1)1} + \lambda b_{(k+1)3}$$

$$\leq \boldsymbol{p}_{k7+5}\boldsymbol{f}^T = \boldsymbol{p}_{k7+6}\boldsymbol{f}^T = \boldsymbol{p}_{k7+7}\boldsymbol{f}^T = 3\mu b_{(k+1)2} + \lambda b_{(k+2)1}.$$

Therefore, $P\boldsymbol{f} \in \mathcal{F}_*$. Thus, P is stochastically block-monotone with respect to $\mathcal{F}_*$. So the stationary probability vector can be obtained by approximation from a stochastic block augmentation of P.

As illustrated in this example, we sometimes need to carefully design a class of block monotone functions in order to establish the block monotone property for some Markov chain.

It is also possible to establish the block monotone property for the transition matrix of the embedded Markov chain of a generalized model of the above example (Zhao and Grassmann (1995)), where the interarrival time has a general distribution function. The transition matrix is of $GI/M/1$ type.

4.2 Queues with non-repeating block rows

Our method can be also applied to queues whose transition matrices are of the forms with non-repeating block rows. Some queueing systems with state-dependent parameters have this kind of natural block monotone structures, as the following example shows.

Example 4.3 Consider a queueing system with a single server, a finite internal source of N customers and an independent non-stationary Poisson arrival stream of external customers (for a simplified version of this model, see, for example, page 299 of Neuts (1981)). Here Poisson arrival rate λ_i, depending on the number of customers in the system upon arrival, is increasing in i and bounded above by λ. Each internal customer who is not in service issues a request for service after an exponential distributed length of time of mean $1/\lambda'$. The service times of both types of customers are exponentially distributed, respectively with parameters μ for external and μ' for internal customers.

The internal customers have preemptive-resume priority over external customers. The classical independence assumptions on inter-arrival and service times are imposed. The model leads to a Markov chain on the state space $\mathcal{E} = \{(i, j), i \geq 0, 0 \leq j \leq N\}$. The index i denotes the number of external customers in the system and the index j represents the number of internal customers requesting service. When the second index assumes the value j, the arrival rate of internal customers is given by $(N - j)\lambda'$.

The generator of this Markov chain is given by

$$Q = \begin{pmatrix} A + B_0 & C_0 & & & & \\ A & B_1 & C_1 & & & \\ & A & B_2 & C_2 & & \\ & & A & B_3 & C_3 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where A, B_i's and C_i's are matrices of order $N + 1$. The matrix A is zero, except for the entry μ in the upper left-hand corner. The matrix B_i is a Jacobi matrix with diagonal elements

$$-\lambda_i - N\lambda' - \mu, \quad -\lambda_i - (N-1)\lambda' - \mu', \quad -\lambda_i - (N-2)\lambda' - \mu', \quad \ldots, \quad -\lambda_i - \lambda' - \mu', \quad -\lambda_i - \mu'.$$

The elements on the superior diagonal are given by $N\lambda', (N - 1)\lambda', \ldots, \lambda'$, and all elements on the inferior diagonal are equal to μ'. The matrix C_i is equal to $\lambda_i I$. Now let c be the largest absolute value of the diagonal entries in Q, and consider the transition matrix $P = I + Q/c$ of the embedded discrete-time Markov chain, which has a similar form as Q.

$$P = \begin{pmatrix} A' + B'_0 & C'_0 & & & & \\ A' & B'_1 & C'_1 & & & \\ & A' & B'_2 & C'_2 & & \\ & & A' & B'_3 & C'_3 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where $A' = A/c$, $B'_i = I + B_i/c$, and $C'_i = C_i/c$, $i = 0, 1, \ldots$.

Using Example 2.8, we obtain that

$$\begin{pmatrix} A' + B'_0 & C'_0 & & & & \\ A' & B'_0 & C'_0 & & & \\ & A' & B'_0 & C'_0 & & \\ & & A' & B'_0 & C'_0 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

is stochastically block-monotone with respect to $\mathcal{F}_m$. Observe that $B'_{i+1} = B'_i - (\lambda_{i+1} - \lambda_i)I$ and $C'_{i+1} = C'_i + (\lambda_{i+1} - \lambda_i)I$ for $i \geq 0$. Using Lemma 2.9, we obtain that P is stochastically block-monotone with respect to $\mathcal{F}_m$, so the stationary probability vector can be obtained by approximation from a stochastic block augmentation of P, and the last block-column augmentation provides the best approximation in the sense (3.6).

4.3 Queues of $M/G/1$ Type

In this subsection, we discuss an example in which the transition probability matrix is of $M/G/1$ type.

Example 4.4 Zhao and Alfa (1995) studied the impact of the presence of impatient customers on a telephone switch. An impatient customer (incomplete call), either due to a dial tone delay or abundance from the system, still consumes about 30% to 80% of the real time of processing a complete call. They formulated the model into a discrete time Markov chain with finite capacity. Since the buffer size is usually large, we here relax the condition to infinite capacity. Specifically, we assume that the arrival process is Poisson with rate λ. The service mechanism is last-come-first-serve and always the patient customers get served first. Let T_0 be the waiting time threshold. Upon the arrival, any customer is patient. If the waiting time a customer has endured in the buffer before entering the service is at least T_0, then this customer becomes impatient and will remain impatient forever. The service times of patient and impatient customers are constants and equal to a and b, respectively, with $a > b$. Consider the time epoch t_n immediately after the n-th service completion. Let I_n and J_n be, respectively, the numbers of the impatient and patient customers in the system at time t_n. Because a patient customer may become impatient later, neither of them is Markovian. To study this system, Zhao and Alfa (1995) proposed an approximate Markovian model essentially refreshing all patient customers at each service completion epoch by forgetting their waiting time history. For convenience, let the capacity m $(= T_0/a)$ of the buffer for storing patient customers be an integer. Now, it is easy to see that $\{(I_n, J_n), n \geq 0\}$ is a Markov chain on the state space $\{(i, j), i \geq 0, j = 0, 1, \ldots, m\}$. It can be shown that the transition matrix is given by

$$P = \begin{bmatrix} B_0 & B_1 & B_2 & B_3 & \cdots \\ A_0 & A_1 & A_2 & A_3 & \cdots \\ & A_0 & A_1 & A_2 & \cdots \\ & & A_0 & A_1 & \cdots \\ & & & A_0 & \cdots \\ & & & & \ddots \end{bmatrix},$$

where

$$B_0 = \begin{bmatrix} b_0 & b_1 & b_2 & \cdots & b_{m-1} & b_m \\ a_0 & a_1 & a_2 & \cdots & a_{m-1} & a_m \\ 0 & a_0 & a_1 & \cdots & a_{m-2} & a_{m-1} \\ 0 & 0 & a_0 & \cdots & a_{m-3} & a_{m-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_0 & a_1 \end{bmatrix}, \quad B_j = \begin{bmatrix} 0 & 0 & \cdots & 0 & b_{m+j} \\ 0 & 0 & \cdots & 0 & a_{m+j} \\ 0 & 0 & \cdots & 0 & a_{m-1+j} \\ 0 & 0 & \cdots & 0 & a_{m-2+j} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{1+j} \end{bmatrix}$$

for $j = 1, 2, \ldots$, and

$$A_0 = \begin{bmatrix} b_0 & b_1 & b_2 & \cdots & b_{m-1} & b_m \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & b_{m+1} \\ a_0 & a_1 & a_2 & \cdots & a_{m-1} & a_m \\ 0 & a_0 & a_1 & \cdots & a_{m-2} & a_{m-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_0 & a_1 \end{bmatrix},$$

and

$$A_j = \begin{bmatrix} 0 & 0 & \cdots & 0 & b_{m+j} \\ 0 & 0 & \cdots & 0 & a_{m-1+j} \\ 0 & 0 & \cdots & 0 & a_{m-2+j} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_j \end{bmatrix}$$

for $j = 2, 3, \ldots$. In the above matrices, for $k = 0, 1, 2, \ldots$,

$$a_k = \frac{a^k}{k!} e^{-a} \quad \text{and} \quad b_k = \frac{b^k}{k!} e^{-b}.$$

To verify that P is stochastically block-monotone with respect to $\mathcal{F}_m$, let C_i be the $m \times \infty$ sub-matrix of P formed by m rows between the $(i-1)m+1$st row and the imth row, $i \geq 1$.

Using Example 2.4 (3), we have $C_i \leq_{\mathcal{F}_m} C_{i+1}$ for $i \geq 2$. To verify that $C_1 \leq_{\mathcal{F}_m} C_2$, let $\boldsymbol{p}_j$ and $\boldsymbol{q}_j$ be the jth row vector of C_1 and C_2 respectively, $1 \leq j \leq m$. Clearly $\boldsymbol{p}_1 = \boldsymbol{q}_1$. Also from Example 2.4 (3), we have $\boldsymbol{p}_j \leq_{\mathcal{F}_m} \boldsymbol{q}_j$ for $j \geq 2$. Thus $C_1 \leq_{\mathcal{F}_m} C_2$.

Therefore, $C_i \leq_{\mathcal{F}_m} C_{i+1}$ for any i, and it follows from Theorem 3.8 that P is stochastically block-monotone with respect to $\mathcal{F}_m$. So the stationary probability vector can be obtained by approximation from a stochastic block-augmentation of P, and the last block-column augmentation provides the best approximation in the sense (3.6).

Finally, we mention that the convergence of scaler augmentations for block-structured matrices was studied by Grassmann and Heyman (1993), where the missing probabilities are added to the last column of the northwest corner. The block augmentation can also be used to deal with the convergence of some numerical algorithms, for example, some of algorithms studied in Latouche (1993).

Acknowledgements

The authors thank two anonymous refereees for their many valuable comments and Dr. W.J. Braun for reading an early draft of this paper, which improved the quality of the paper.

References

[1] Billingsley, P. (1969), *Convergence of Probability Measures*, John Wiley & Sons, New York.

[2] Bright, L. and P.G. Taylor (1996), Equilibrium distributions for level-dependent-quasi-birth-and-death processes, in *Matrix-Analytic Methods in Stochastic Models*, Chakravarthy, S. and Alfa, A.S. eds, Marcel Dekker, New York, 359–373.

[3] Gail, H.R., S.L. Hantler and B.A. Taylor (1996), Spectral analysis of $M/G/1$ and $GI/M/1$ type Markov chains, *Adv. Appl. Prob.*, **28**, 114–165.

[4] Gail, H.R., S.L. Hantler and B.A. Taylor (1997), $M/G/1$ and $GI/M/1$ type Markov chains with multiple boundary levels, *Adv. Appl. Prob.*, **29**, 773–758.

[5] Grassmann, W.K. and D.P. Heyman (1990), Equilibrium distribution of block-structured Markov chains with repeating rows, *J. Appl. Prob.*, **27**, 557–576.

[6] Grassmann, W.K. and D.P. Heyman (1993), Computation of steady-state probabilities for infinite-state Markov chains with repeating rows, *ORSA Journal on Computing*, **5**, 292–303.

[7] Gibson, D. and E. Seneta (1987), Monotone infinite stochastic matrices and their augmented truncations, *Stochastic Processes and Their Applications*, **24**, 287–292.

[8] Heyman, D. (1991), Approximating the stationary distribution of an infinite stochastic matrix, *Journal of Applied Probability*, **28**, 96–103.

[9] Latouche, G. (1993), Algorithms for infinite Markov chains with repeating columns, in *Linear Algebra, Markov Chains and Queueing Models*, Meyer, C.D. and Plemmons, R.J. Eds, Springer Verlag, New York, 231–265.

[10] Li, H. and M. Shaked (1994), Stochastic convexity and concavity of Markov processes, *Mathematics of Operations Research* **19**, 477–493.

[11] Neuts, M.F. (1981), *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, The Johns Hopkins University Press, Baltimore.

[12] Neuts, M.F. (1989), *Structured Stochastic Matrices of $M/G/1$ Type and Their Applications*, Marcel Decker Inc., New York.

[13] Neuts, M.F. (1998), Some promising directions in algorithmic probability, in *Advances in Matrix Analysis Methods for Stochastic Models*, Alfa, A.S. and Chakravarthy, S.R. Eds, Notable Publications, New Jersey, 429–443.

[14] Seneta, E. (1980), Computing the stationary distribution for infinite Markov chains, *Linear Algebra and Its Applications*, **34**, 259–267.

[15] Wolf, D. (1980), Approximation of the invariant probability measure of an infinite stochastic matrix, *Advance in Applied Probability*, **12**, 710–726.

[16] Zhao, Y.Q. (1990), *Shortest Queue Models*, Ph.D thesis, University of Saskatchewan, Canada.

[17] Zhao, Y.Q. and W.K. Grassmann (1995), Queueing analysis of a jockeying model, *Operations Research*, **43**, 520–530.

[18] Zhao, Y.Q. and A.S. Alfa (1995), Performance analysis of a telephone system with both patient and impatient customers, *Telecommunication Systems*, **4**, 201–215.

[19] Zhao, Y.Q. and D. Liu (1996), The censored Markov chain and the best augmentation, *Journal of Applied Probability*, **33**, 623–629.

[20] Zhao, Y.Q., W. Li and W.J. Braun (1998a), Infinite block-structured transition matrices and their properties, *Adv. Appl. Prob.*, **30**, 365–384.

[21] Zhao, Y.Q., W. Li and W.J. Braun (1998b), Factorization, spectral analysis and fundamental matrix for transition matrices with block-repeating entries, submitted to *Appl. Prob.*.

[22] Zhao, Y.Q., W. Li and A.S. Alfa (1999), Duality results for block-structured transition matrices, Infinite block-structured transition matrices and their properties, to appear in *J. Appl. Prob.*, **36**.