RAIRO Operations Research
RAIRO Oper. Res. **37** (2003) 85-97
DOI: 10.1051/ro:2003015

TRANSFORMING STOCHASTIC MATRICES FOR STOCHASTIC COMPARISON WITH THE ST-ORDER*

Tuğrul Dayar[1], Jean-Michel Fourneau[2] and Nihal Pekergin[2]

Communicated by Catherine Roucairol

Abstract. We present a transformation for stochastic matrices and analyze the effects of using it in stochastic comparison with the strong stochastic (st) order. We show that unless the given stochastic matrix is row diagonally dominant, the transformed matrix provides better st bounds on the steady state probability distribution.

Keywords. Markov processes, probability distributions, stochastic ordering, st-order.

1. Introduction

The stochastic comparison of random variables is a powerful technique in different areas of applied probability [7]. It allows the resolution of complex models involving large state spaces, and/or numerically difficult operators or distributions. There are several applications of this technique in practical problems of telecommunication engineering [9, 10] or reliability [11]. The stochastic comparison of Markov Chains (MC for short) is discussed in detail in [3, 8, 12]. The comparison of two MCs may be established by the comparison of their state probability distributions at each time instant. Obviously, if steady states exist, stochastic comparison between their steady state probability distributions is also possible.

Received July, 2001.

* *This work is supported by TÜBİTAK-CNRS joint grant.*

[1] Department of Computer Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey.

[2] PR*i*SM, Université de Versailles-St.Quentin, 45 avenue des États-Unis, 78035 France; e-mail: jmf@prism.uvsq.fr

There are different stochastic ordering relations and the most well known is the strong stochastic ordering (*i.e.*, $\leq_{\rm st}$). Intuitively speaking, two random variables X and Y which take values on a totally ordered space being comparable in the strong stochastic sense (*i.e.*, $X \leq_{\rm st} Y$) means that it is less probable for X to take larger values than Y (see [11, 12]).

Sufficient conditions for the existence of stochastic comparison of two time-homogeneous MCs are given by the stochastic monotonicity and bounding properties of their one step transition probability matrices [3, 8]. In [14], this idea is used to devise an algorithm that constructs an optimal st-monotone upper bounding MC corresponding to a given MC. Later, this algorithm is used to compute stochastic bounds on performance measures that are defined on a totally ordered and reduced state space (see [4] for a tutorial). Performance measures may be defined as reward functions of the underlying MC. In [1], states having the same reward are aggregated, so the state space size of the bounding MC is considerably reduced. St-comparison of MCs implies that both transient and steady state performance measures may be bounded. However, quite often, the transient measures are meaningless or too difficult to compute. So, we may accept to lose the transient bounds to improve the accuracy of steady state bounds.

In this note, we characterize the properties of a simple transformation on a discrete-time Markov chain (DTMC) and analyze its effects on the optimal st-monotone upper bounding matrix computed by the algorithm in [1]. This transformation keeps invariant the steady state distribution. Our motivation is to improve the accuracy of the steady state probability bounds that may be computed by stochastic comparison with the st-order. We remark that the transformation has a similar effect on the optimal st-monotone lower bounding matrix which we do not discuss here.

In this paper, we focus on the accuracy of the bounds and we do not consider the complexity issue. The matrix we obtain has the same size as the original matrix. We do not study new techniques to reduce the complexity of the resolution (see [2,6] for this topic). Indeed, the arguments developed in [14] are still valid and allow a large reduction of the state space. In our opinion, Vincent and Plateau methodology is sufficient to reduce the state space. And the accuracy of the bounds remains the major problem.

In Section 2, we provide brief background on stochastic comparison with the st-order and present an example. In Section 3, we introduce the transformation and provide a comprehensive analysis. Section 4 has concluding remarks.

2. Some preliminaries

First, we give the definition of st-ordering used in this note. For further information on the stochastic comparison method, we refer the reader to [12].

Definition 1. Let X and Y be random variables taking values on a totally ordered space. Then X is said to be less than Y in the strong stochastic sense, that is,

$X \leq_{\mathrm{st}} Y$ iff $E[f(X)] \leq E[f(Y)]$ for all nondecreasing functions f whenever the expectations exist.

Definition 2. Let X and Y be random variables taking values on the finite state space $\{1, 2, \ldots, n\}$. Let p and q be probability distribution vectors such that

$$p_j = Pr(X = j) \quad \text{and} \quad q_j = Pr(Y = j) \quad \text{for} \;\; j = 1, 2, \ldots, n.$$

Then X is said to be less than Y in the strong stochastic sense, that is, $X \leq_{\mathrm{st}} Y$ iff

$$\sum_{j=k}^{n} p_j \leq \sum_{j=k}^{n} q_j \quad \text{for} \;\; k = 1, 2, \ldots, n.$$

It is shown in Theorem 3.4 of [8] (p. 355) that monotonicity and comparability of the one step transition probability matrices of time-homogeneous MCs yield sufficient conditions for their stochastic comparison, which is summarized in:

Theorem 1. *Let P and Q be stochastic matrices respectively characterizing time-homogeneous MCs $X(t)$ and $Y(t)$. Then $\{X(t),\ t \in \mathcal{T}\} \leq_{\mathrm{st}} \{Y(t),\ t \in \mathcal{T}\}$ if*

- $X(0) \leq_{\mathrm{st}} Y(0)$*;*
- st-*monotonicity of at least one of the matrices holds, that is,*

$$\textit{either} \quad P_{i,*} \leq_{\mathrm{st}} P_{j,*} \quad \textit{or} \quad Q_{i,*} \leq_{\mathrm{st}} Q_{j,*} \quad \forall i, j \;\; \textit{such that} \;\; i \leq j;$$

- st-*comparability of the matrices holds, that is, $P_{i,*} \leq_{\mathrm{st}} Q_{i,*}$ $\forall i$.*

Here $P_{i,*}$ refers to row i of P.

On page 11 of [1], the following algorithm is presented to construct the optimal st-monotone upper bounding DTMC Q for a given DTMC P.

Algorithm 1. *Construction of optimal* st-*monotone upper bounding DTMC Q corresponding to DTMC P of order n:*

$q_{1,n} = p_{1,n}$;
for $i = 2, 3, \ldots, n$,
  $q_{i,n} = \max(q_{i-1,n}, p_{i,n})$;
for $l = n-1, n-2, \ldots, 1$,
  $q_{1,l} = p_{1,l}$;
  for $i = 2, 3, \ldots, n$,
    $q_{i,l} = \max(\sum_{j=l}^{n} q_{i-1,j}, \sum_{j=l}^{n} p_{i,j}) - \sum_{j=l+1}^{n} q_{i,j}$.

Let U be another st-monotone upper bounding DTMC for P. Then Q is optimal in the sense that $Q \leq_{\mathrm{st}} U$.

The following example provides the results of applying Algorithm 1 to two MCs that have the same steady state probability distribution, and shows that it may be possible to obtain different steady state st bounds.

Example 1. Consider the following (4×4) DTMC

$$P = \begin{pmatrix} 0.2 & 0 & 0.3 & 0.5 \\ 0.1 & 0 & 0.6 & 0.3 \\ 0.4 & 0.3 & 0.1 & 0.2 \\ 0.3 & 0.3 & 0.3 & 0.1 \end{pmatrix}$$

whose steady state probability distribution is given by $\pi_P = [0.2686,\ 0.1688,\ 0.2922,\ 0.2704]$. Application of Algorithm 1 to P yields the st-monotone upper bounding DTMC

$$Q = \begin{pmatrix} 0.2 & 0 & 0.3 & 0.5 \\ 0.1 & 0 & 0.4 & 0.5 \\ 0.1 & 0 & 0.4 & 0.5 \\ 0.1 & 0 & 0.4 & 0.5 \end{pmatrix}.$$

The steady state probability distribution of Q given by $\pi_Q = [0.1111,\ 0.0000,\ 0.3889,\ 0.5000]$ provides an st upper bound on π_P (*cf.* Def. 2). Note that it is possible to obtain an st-monotone upper bounding DTMC having transient states with Algorithm 1 even though the input DTMC was irreducible. Nevertheless, we will always have a single irreducible subset of states, which includes the last state, in the output DTMC with Algorithm 1 if there is a path from each state to the last state in the input DTMC [1].

Application of Algorithm 1 to

$$R = \begin{pmatrix} 0.6 & 0 & 0.15 & 0.25 \\ 0.05 & 0.5 & 0.3 & 0.15 \\ 0.2 & 0.15 & 0.55 & 0.1 \\ 0.15 & 0.15 & 0.15 & 0.55 \end{pmatrix},$$

which has the same steady state probability distribution as P, yields the st-monotone upper bounding DTMC

$$S = \begin{pmatrix} 0.6 & 0 & 0.15 & 0.25 \\ 0.05 & 0.5 & 0.2 & 0.25 \\ 0.05 & 0.3 & 0.4 & 0.25 \\ 0.05 & 0.25 & 0.15 & 0.55 \end{pmatrix}.$$

The steady state probability distribution of S is given by $\pi_S = [0.1111,\ 0.3110,\ 0.2207,\ 0.3571]$, and it is clearly a better st-upper bound on π_P than π_Q. Now, we show how R is obtained from P using a simple linear transformation.

3. The transformation and its analysis

Proposition 1. *Let P be a regular DTMC of order n. Consider the transformation*

$$R = (1-\delta)I + \delta P \quad for \quad \delta \in (0,1). \tag{1}$$

(i) *Then R is a regular DTMC of order n, where*

$$r_{i,j} = \begin{cases} 1 - \delta(1 - p_{i,i}), & i = j \\ \delta p_{i,j}, & i \neq j \end{cases} \quad \textit{for} \quad i, j = 1, 2, \ldots, n; \tag{2}$$

(ii) *R has the same steady state probability distribution as P.*

Proof. By construction, R is a DTMC of order n and its elements are given by equation (2). Furthermore, the off-diagonal part of R has the same nonzero structure as that of P because I is the identity matrix with ones on the diagonal and zeros elsewhere. Since P is regular ([13], p. 120) (*i.e.*, finite, irreducible, aperiodic), then so must be R. Existence of the steady state probability distribution of P follows from the fact that P is regular. The steady state distribution is the only stationary distribution, and it satisfies $\pi P = \pi$, $\|\pi\|_1 = 1$. Since R is regular, π is also the stationary distribution of R:

$$\pi R = \pi[(1 - \delta)I + \delta P] = (1 - \delta)\pi + \delta\pi = \pi. \qquad \square$$

Corollary 1. *If P is a DTMC of order n, then the transformation in equation (1) for $\delta \in (0, 1)$ satisfies: (i) $0 \leq \sum_{j\neq i}^{n} r_{i,j} \leq \delta$, (ii) $1 - \delta \leq r_{i,i} \leq 1$ for $i = 1, 2, \ldots, n$.*

Proof. From equation (2) and $0 \leq \sum_{j\neq i}^{n} p_{i,j} \leq 1$, we have $0 \leq \sum_{j\neq i}^{n} r_{i,j} = \delta \sum_{j\neq i}^{n} p_{i,j} \leq \delta$ for $i = 1, 2, \ldots, n$. This proves part (i). To prove part (ii), we write $r_{i,i} = 1 - \sum_{j\neq i}^{n} r_{i,j}$ and use part (i). $\square$

Definition 3. A stochastic matrix is said to be row diagonally dominant (RDD) if all of its diagonal elements are greater than or equal to 0.5.

Theorem 2. *Let P be a DTMC of order n that is not RDD. Consider the transformation in equation (1) for*[3]

$$\delta_* = \frac{0.5}{1 - \min_{1\leq i\leq n} p_{i,i}}, \tag{3}$$

and let S be the st*-monotone upper bounding DTMC for R computed by Algorithm 1. Then:*

(i) $0.5 \leq r_{i,i} = 1 - \sum_{j\neq i}^{n} r_{i,j} \leq 1$ *for $i = 1, 2, \ldots, n$ (i.e., R is RDD);*
(ii) $0 \leq \sum_{j=i+1}^{n} s_{i,j} \leq 0.5$*, for $i = 1, 2, \ldots, n - 1$;*
(iii) $0.5 \leq \sum_{j=i}^{n} s_{i,j} = \sum_{j=i}^{n} r_{i,j} \leq 1$ *for $i = 1, 2, \ldots, n$.*

Proof. We remark that δ_* is the largest positive scalar within (0,1) that makes R RDD. Part (i) follows from Corollary 1 and that $\delta_* \leq 0.5$. Now, consider the

[3] δ_* assumes the value in equation (3) throughout this note.

implications of st-monotonicity and st-comparability on S. From Algorithm 1, we have

$$\sum_{j=i+1}^{n} s_{i,j} = \max_{1\le m\le i}\left(\sum_{j=i+1}^{n} r_{m,j}\right) \quad \text{for} \quad i = 1,2,\ldots,n-1.$$

Since, $0 \le \sum_{j=i+1}^{n} r_{m,j} \le 0.5$ for $m \le i$ from part (i) of Theorem 2, part (ii) is proved. Again, consider how st-monotonicity and st-comparability are imposed on S:

$$\sum_{j=i}^{n} s_{i,j} = \max_{1\le m\le i}\left(\sum_{j=i}^{n} r_{m,j}\right) \quad \text{for} \quad i = 1,2,\ldots,n.$$

However, $\max_{1\le m\le i}(\sum_{j=i}^{n} r_{m,j}) = \sum_{j=i}^{n} r_{i,j}$ for $i = 1,2,\ldots,n$ and $0.5 \le \sum_{j=i}^{n} \times r_{i,j} \le 1$ from part (i) of Theorem 2, implying part (iii). □

When R is RDD, its diagonal serves as a barrier for the perturbation moving from the upper-triangular part to the strictly lower-triangular part in forming S. We stress that it is the concept of row diagonal dominance together with the semantics of Algorithm 1 (*i.e.*, st-monotonicity and st-comparability) and nothing more that enable us to develop the results in this note.

Corollary 2. *Let P be a DTMC of order n that is not RDD. Consider the transformation in equation (1) for δ_*, and let S be the* st*-monotone upper bounding DTMC for R computed by Algorithm 1. Then:*

(*i*) $s_{2,1} = r_{2,1}$;
(*ii*) $s_{n,n} = r_{n,n}$;
(*iii*) $s_{i,i} \le r_{i,i}$ *for* $i = 2,3,\ldots,n-1$.

Proof. To prove part (i), we write $s_{2,1} = 1 - \sum_{j=2}^{n} s_{2,j}$ and use part (iii) of Theorem 2. To prove part (ii), recall that $s_{n,n} = \max(s_{n-1,n}, r_{n,n})$. But, $s_{n-1,n} \le 0.5$ and $0.5 \le r_{n,n} \le 1$. So, the maximum may be taken as the second argument, and we have $s_{n,n} = r_{n,n}$. To prove part (iii), note that part (iii) of Theorem 2 directly gives $s_{i,i} + \sum_{j=i+1}^{n} s_{i,j} = r_{i,i} + \sum_{j=i+1}^{n} r_{i,j}$. Since $\sum_{j=i+1}^{n} r_{i,j} \le \sum_{j=i+1}^{n} s_{i,j}$ for $i = 2,3,\ldots,n-1$ due to st-comparability from Algorithm 1, we have $s_{i,i} \le r_{i,i}$ for $i = 2,3,\ldots,n-1$. □

Theorem 3. *Let P be a DTMC of order n that is not RDD. Consider the transformation in equation (1) for two different values $\delta_1, \delta_2 \in (0, \delta_*]$ such that $\delta_1 < \delta_2$, and let $S(\delta_l)$ be the* st*-monotone upper bounding DTMC for $R(\delta_l)$, $l = 1,2$, computed by Algorithm 1. Then $s_{i,j}(\delta_1) = \rho s_{i,j}(\delta_2)$ for $i \ne j = 1,2,\ldots,n$, where $\rho = \delta_1/\delta_2 \in (0,1)$.*

Proof. Due to the form of the transformation in equation (1), we have $r_{i,j}(\delta_1) = \rho r_{i,j}(\delta_2)$ for $i \ne j = 1,2,\ldots,n$ from equation (2). Furthermore, due to Algorithm 1, the first rows of $S(\delta_l)$ and $R(\delta_l)$, $l = 1,2$, are identical. Hence, we have $s_{1,j}(\delta_1) = \rho s_{1,j}(\delta_2)$ for $j = 2,3,\ldots,n$, and the theorem holds for all off-diagonal elements in the first rows of $S(\delta_1)$ and $S(\delta_2)$. Now, consider how $s_{2,n}(\delta_l)$, $l = 1, 2$, is computed:

$$s_{2,n}(\delta_l) = \max(s_{1,n}(\delta_l), r_{2,n}(\delta_l)).$$

But, $\max(s_{1,n}(\delta_1), r_{2,n}(\delta_1)) = \rho \max(s_{1,n}(\delta_2), r_{2,n}(\delta_2)) = \rho s_{2,n}(\delta_2)$. Hence, the theorem holds for $s_{2,n}(\delta_1)$ and $s_{2,n}(\delta_2)$. Next, consider how $s_{2,k}(\delta_l)$, $l = 1, 2$, for $k = 3, 4, \ldots, n-1$ is computed starting from column $n-1$ left to column 3:

$$s_{2,k}(\delta_l) = \max\left(\sum_{j=k}^{n} s_{1,j}(\delta_l), \sum_{j=k}^{n} r_{2,j}(\delta_l)\right) - \sum_{j=k+1}^{n} s_{2,j}(\delta_l).$$

But $\sum_{j=k}^{n} s_{1,j}(\delta_1) = \rho \sum_{j=k}^{n} s_{1,j}(\delta_2)$, $\sum_{j=k}^{n} r_{2,j}(\delta_1) = \rho \sum_{j=k}^{n} r_{2,j}(\delta_2)$, and $\sum_{j=k+1}^{n} s_{2,j}(\delta_1) = \rho \sum_{j=k+1}^{n} s_{2,j}(\delta_2)$. Hence, $s_{2,k}(\delta_1) = \rho s_{2,k}(\delta_2)$ for $k = 3$, $4, \ldots, n$. Finally, from part (i) of Corollary 2, we have $s_{2,1}(\delta_l) = r_{2,1}(\delta_l)$, $l = 1, 2$, which implies $s_{2,1}(\delta_1) = \rho s_{2,1}(\delta_2)$. Hence, the theorem holds for all off-diagonal elements in the second rows of $S(\delta_1)$ and $S(\delta_2)$. This is the basis step. Now, let the induction hypothesis be $s_{i,j}(\delta_1) = \rho s_{i,j}(\delta_2)$ for $i = 3, 4, \ldots, m-1$. Then, we must show that $s_{m,j}(\delta_1) = \rho s_{m,j}(\delta_2)$ for $j \neq m$. The proof for $j > m$ is similar to that of row 2. That is, one starts with the proof for the last column and moves to the left till column $m+1$. Hence, we concentrate on the proof for columns $j = 1, 2, \ldots, m-1$. Let us consider how $s_{m,m-1}(\delta_l)$, $l = 1, 2$, is computed:

$$s_{m,m-1}(\delta_l) = \max\left(\sum_{j=m-1}^{n} s_{m-1,j}(\delta_l), \sum_{j=m-1}^{n} r_{m,j}(\delta_l)\right) - \sum_{j=m}^{n} s_{m,j}(\delta_l).$$

But, $\sum_{j=m-1}^{n} s_{m-1,j}(\delta_l) = 1 - \sum_{j=1}^{m-2} s_{m-1,j}(\delta_l)$, $\sum_{j=m-1}^{n} r_{m,j}(\delta_l) = 1 - \sum_{j=1}^{m-2} \times$ $r_{m,j}(\delta_l)$, and $\sum_{j=m}^{n} s_{m,j}(\delta_l) = \sum_{j=m}^{n} r_{m,j}(\delta_l) = 1 - \sum_{j=1}^{m-1} r_{m,j}(\delta_l)$ from part (iii) of Theorem 2. Therefore, $s_{m,m-1}(\delta_l) = \min(\sum_{j=1}^{m-2} s_{m-1,j}(\delta_l), \sum_{j=1}^{m-2} r_{m,j}(\delta_l)) + \sum_{j=1}^{m-1} r_{m,j}(\delta_l)$. The elements of $S(\delta_l)$ that contribute to $s_{m,m-1}(\delta_l)$ come from the induction hypothesis and those of $R(\delta_l)$ are in the strictly lower-triangular part of row m. Hence, $s_{m,m-1}(\delta_1) = \rho s_{m,m-1}(\delta_2)$. The proof for columns $j < m-1$ in row m is similar. $\square$

Corollary 3. *Let P be a DTMC of order n that is not RDD. Consider the transformation in equation (1) for two different values $\delta_1, \delta_2 \in (0, \delta_*]$ such that $\delta_1 < \delta_2$, and let $S(\delta_l)$ be the* st*-monotone upper bounding DTMC for $R(\delta_l)$, $l = 1, 2$, computed by Algorithm 1. Then $S(\delta_1)$ and $S(\delta_2)$ have the same steady state probability distribution.*

Proof. Since, both $S(\delta_1)$ and $S(\delta_2)$ are DTMCs by construction, from Theorem 3 we must have a transformation of the form

$$s_{i,j}(\delta_1) = \begin{cases} 1 - \rho(1 - s_{i,i}(\delta_2)), & i = j \\ \rho s_{i,j}(\delta_2), & i \neq j, \end{cases}$$

where $\rho = \delta_1/\delta_2 \in (0,1)$. But, this is a transformation as in equation (1). Since, $S(\delta_1)$ and $S(\delta_2)$ have the same nonzero structure, they will have the same steady state probability distribution whenever it exists, as we already proved in part (ii) of Proposition 1. The existence of the steady state distributions follows from the fact that $S(\delta_l)$, $l = 1, 2$, is finite, has one irreducible subset of states including state n [1], and $0.5 \leq s_{n,n}(\delta_l)$ from part (ii) of Corollary 2 implying aperiodicity. □

An important consequence of Corollary 3 is that one cannot improve the steady state probability bounds by choosing a smaller δ value to transform an already RDD DTMC.

Corollary 4. *Let P be a DTMC of order n that is RDD and Q be the corresponding* st-*monotone upper bounding DTMC computed by Algorithm 1. Consider the transformation in equation (1) for $\delta \in (0,1)$, and let S be the* st-*monotone upper bounding DTMC for R computed by Algorithm 1. Then Q and S have the same steady state probability distribution.*

Proof. Follows from Corollary 3 by noticing that $R(\delta_1)$ and $R(\delta_2)$ are both RDD. □

The discussion so far sheds light on the characteristics of the optimal st-monotone upper bounding DTMC computed by Algorithm 1 using the transformed matrix, and its steady state probability distribution. Having set the stage, we adapt a different approach to state the main result about the quality of this distribution. Specifically, our goal is to prove:

Theorem 4. *Let P be a DTMC of order n and Q be the corresponding* st-*monotone upper bounding DTMC computed by Algorithm 1. Consider the transformation in equation (1) for $\delta \in (0,1)$, and let S be the* st-*monotone upper bounding DTMC for R computed by Algorithm 1. Then $\pi_S \leq_{\mathrm{st}} \pi_Q$, where π_S and π_Q are respectively the steady state probability distributions of S and Q.*

To enhance readability, from now on we denote the (i,j)-th element of the matrix A as $A[i,j]$ rather than $a_{i,j}$.

Definition 4. Let $\mathcal{B}$ be the set of DTMCs of order n, and let $P \in \mathcal{B}$. We define the following three operators to assist us in proving Theorem 4:

(i) t is the operator corresponding to the transformation in equation (1):

$$t(P)[i,j] = \begin{cases} 1 - \delta + \delta P[i,i], & i = j \\ \delta P[i,j], & i \neq j. \end{cases}$$

We remark that $t(P) \in \mathcal{B}$;

(ii) r is the summation operator used in the st-comparison:

$$r(P)[i,j] = \sum_{k=j}^{n} P[i,k].$$

We remark that $r(P)$ is not a stochastic matrix. Let $\mathcal{A}$ be the set of matrices defined by $r(P)$, where $P \in \mathcal{B}$;

(iii) v is the following operator which transforms $P \in \mathcal{B}$ to a matrix in $\mathcal{A}$:

$$v(P)[i,j] = \begin{cases} \sum_{k=j}^{n} P[1,k], & i = 1 \\ \max(v(P)[i-1,j], \sum_{k=j}^{n} P[i,k]), & i > 1. \end{cases}$$

Proposition 2. *Let $Z \in \mathcal{A}$. Then r^{-1}, the inverse operator of r, is given by*

$$r^{-1}(Z)[i,j] = \begin{cases} Z[i,n], & j = n \\ Z[i,j] - Z[i,j+1], & j < n. \end{cases}$$

We remark that $r^{-1}(Z) \in \mathcal{B}$.

Proposition 3. *Unrolling v yields the simpler representation*

$$v(P)[i,j] = \max_{m \le i} \left(\sum_{k \ge j} P[m,k] \right).$$

Proposition 4. *The operator corresponding to Algorithm 1 is $r^{-1}v$.*

The proofs of Propositions 2 through 4 are straightforward. From Proposition 4, we have $Q = r^{-1}v(P)$, $S = r^{-1}vt(P)$ and S is st-monotone. Our objective is to prove that

$$r^{-1}vt(P) \le_{\mathrm{st}} tr^{-1}v(P) \tag{4}$$

since it implies $\pi_S \le_{\mathrm{st}} \pi_{t(Q)} = \pi_Q$, the equality following from part (ii) of Proposition 1. To this end, we need to specify the composition of the operators in (4).

Proposition 5. *Let $P \in \mathcal{B}$. Then the composition vt is given by*

$$vt(P)[i,j] = \begin{cases} 1, & i = 1,\ j = 1 \\ \delta v(P)[i,j], & i \ge 1,\ j > i \\ \max_{m \le i}(\sum_{k \ge j}(\delta P[m,k] + (1-\delta)1_{m=k})), & i > 1,\ j \le i. \end{cases}$$

Proof. The result follows from substituting part (i) of Definition 4 in Proposition 3 and algebraic manipulations. $\square$

Proposition 6. *Let $Z \in \mathcal{A}$. Then the composition rtr^{-1} is given by*

$$rtr^{-1}(Z)[i,j] = \begin{cases} \delta Z[i,j], & i < j \\ 1 - \delta + \delta Z[i,j], & i \ge j. \end{cases}$$

Proof. From algebraic manipulations using Definition 4, Proposition 2, and substitution, we have

$$rtr^{-1}(Z)[i,j] = \begin{cases} 1-\delta+\delta Z[n,n], & i=n,\ j=n \\ \delta Z[i,n], & i<n,\ j=n \\ rtr^{-1}(Z)[i,j+1]+\delta Z[i,j]-\delta Z[i,j+1], & i<n,\ j\neq i \\ rtr^{-1}(Z)[i,i+1]+1-\delta+\delta Z[i,i]-\delta Z[i,i+1], & i<n,\ j=i. \end{cases}$$

The result follows after unrolling the recurrences in the last two lines. □

Proposition 7. *Let $P \in \mathcal{B}$. Then the composition $rtr^{-1}v$ is given by*

$$rtr^{-1}v(P)[i,j] = \begin{cases} 1, & i=1,\ j=1 \\ \delta v(P)[1,j], & i=1,\ j>1 \\ 1-\delta+\delta v(P)[i,j], & i>1,\ j\leq i \\ \delta v(P)[i,j], & i>1,\ j>i. \end{cases}$$

Proof. The result follows from direct substitution using Proposition 6. □

Thus, we have to compare the two systems of recurrence equations in Propositions 5 and 7 which are based on $v(P)$. We remark that both systems are linear systems on the (max,+) semi-ring. Furthermore, the two systems have the same values in the strictly upper triangular part (*i.e.*, when $j>i$) and at the point $(1,1)$. This suggests an element-wise comparison as specified in the next proposition the proof of which follows:

Proposition 8. *$r^{-1}vt(P) \leq_{\mathrm{st}} tr^{-1}v(P)$ is equivalent to $vt(P) \leq rtr^{-1}v(P)$, where the latter comparison is element-wise.*

It is easier to use element-wise comparison (*i.e.* $\leq$) because we have to compare elements defined by recurrence relations. We do not want to unroll the recurrence relations of the operator v. So, let us proceed with the comparison of the lower triangular elements.

Lemma 1. *For all i and j such that $i \geq j$, we have $vt(P)[i,j] \leq rtr^{-1}v(P)[i,j]$.*

Proof. Recall the value of $vt(P)$ for $i \geq j$:

$$vt(P)[i,j] = \max_{m\leq i}\left(\sum_{k\geq j}(\delta P[m,k]+(1-\delta)1_{m=k})\right).$$

Using the fact that the maximum of a summation is less than or equal to the sum of the maxima, we obtain

$$vt(P)[i,j] \leq \max_{m\leq i}\left(\sum_{k\geq j}\delta P[m,k]\right) + \max_{m\leq i}\left(\sum_{k\geq j}(1-\delta)1_{m=k}\right).$$

Since $i \geq j$, the summation of the indicator function (*i.e.*, the second term) equals 0 or 1, and the value 1 is reached for some m. Thus,

$$\max_{m \leq i} \left(\sum_{k \geq j} (1-\delta) 1_{m=k} \right) = (1-\delta).$$

As the multiplication by δ is linear for both of the operators max and $+$, we identify $v(P)[i,j]$ to complete the proof:

$$vt(P)[i,j] \leq \delta v(P)[i,j] + (1-\delta) = rtr^{-1} v(P)[i,j]. \qquad \square$$

Hence, we have proved Proposition 8, which in turn completes the proof of Theorem 4, and we have $\pi_S \leq_{\mathrm{st}} \pi_Q$.

Theorem 5. *Let P be a DTMC of order n that is not RDD and Q be the corresponding* st*-monotone upper bounding DTMC computed by Algorithm 1. Consider the transformation in equation (1) for two different values $\delta_1, \delta_2 \in [\delta_*, 1)$ such that $\delta_1 < \delta_2$, and let $S(\delta_l)$ be the* st*-monotone upper bounding DTMC for $R(\delta_l)$, $l = 1, 2$, computed by Algorithm 1. Then $\pi_{S(\delta_1)} \leq_{\mathrm{st}} \pi_{S(\delta_2)} \leq_{\mathrm{st}} \pi_Q$, where $\pi_{S(\delta_l)}$ and π_Q are respectively the steady state probability distributions of $S(\delta_l)$, $l = 1, 2$, and Q. Furthermore, if $P[n,n] < \max_{1 \leq i \leq n}(P[i,n])$, then $\pi_{S(\delta_2)} \neq \pi_Q$.*

Proof. The general result follows from Theorem 4 together with Corollary 4. As for the latter part, observe that $\pi_Q = \pi_{t(Q)}$ from part (i) of Definition 4 and part (ii) of Proposition 1. Now, assume that $\pi_{S(\delta_2)} = \pi_{t(Q)}$, where t uses δ_2. We will prove by contradiction that this is not possible. By construction, we have $S(\delta_2)[i,n] = t(Q)[i,n] = \delta_2 \max_{1 \leq m \leq i}(P[m,n])$ for $i = 1, 2, \ldots, n-1$, $S(\delta_2)[n,n] = 1 - \delta_2 + \delta_2 P[n,n]$, and $t(Q)[n,n] = 1 - \delta_2 + \max_{1 \leq i \leq n}(P[i,n])$. Then $P[n,n] < \max_{1 \leq i \leq n}(P[i,n])$ implies $S(\delta_2)[n,n] < t(Q)[n,n]$. Now, notice that

$$\begin{aligned} \pi_{S(\delta_2)}[n] &= \sum_{i=1}^{n} \pi_{S(\delta_2)}[i]\ S(\delta_2)[i,n] \\ &= \pi_{S(\delta_2)}[n]\ S(\delta_2)[n,n] + \sum_{i=1}^{n-1} \pi_{S(\delta_2)}[i]\ S(\delta_2)[i,n], \\ \pi_{t(Q)}[n] &= \sum_{i=1}^{n} \pi_{t(Q)}[i]\ t(Q)[i,n] = \pi_{t(Q)}[n]\ t(Q)[n,n] + \sum_{i=1}^{n-1} \pi_{t(Q)}[i]\ t(Q)[i,n]. \end{aligned}$$

The second term involving the summation on the right-hand side is the same in both equations since $S(\delta_2)[i,n] = t(Q)[i,n]$ for $i = 1, 2, \ldots, n-1$, and we assumed $\pi_{S(\delta_2)} = \pi_{t(Q)}$. However, the first terms are different, contradicting the assumption that $\pi_{S(\delta_2)}[n] = \pi_{t(Q)}[n]$. Hence, it must be that $\pi_{S(\delta_2)} \neq \pi_Q$. $\square$

Proposition 9. *Remember that $0.5 \leq \delta_*$. Thus, one matrix which gives the best* st*-bound for this family of matrices can be computed through $P/2 + I/2$.*

The st-monotone upper bounding matrix construction algorithm for continuous-time Markov chains (CTMCs) (see [14]) employed in [15] uses the diagonal of $(P - I)$ as a barrier for the perturbation that is moving from the upper-triangular part to the strictly lower-triangular part in forming the continuous-time st-monotone upper bounding matrix. In other words, the algorithm in [14] essentially achieves the same effect as the transformation in equation (1) for $\delta \in (0, \delta_*]$ on a stochastic matrix that is not RDD. However, to the best of our knowledge a discussion of its characteristics and an analysis of its effects on the bounding matrix do not exist.

4. Conclusion

We have presented a transformation for stochastic matrices that may be used in stochastic comparison with the strong stochastic order (see [5] for a tool on st-bounds which implements this result). We have shown that if the given stochastic matrix is not row diagonally dominant, then the steady state probability distribution of the optimal st-monotone upper bounding matrix corresponding to the row diagonally dominant transformed matrix is better in the strong stochastic sense than the one corresponding to the original matrix. And we have established that the transformation $P/2 + I/2$ provides the best bound for the family of transformation we have considered here.

References

[1] O. Abu–Amsha and J.-M. Vincent, An algorithm to bound functionals of Markov chains with large state space, in *4th INFORMS Conference on Telecommunications*. Boca Raton, Florida (1998). Available as Rapport de recherche MAI No. 25. IMAG, Grenoble, France (1996).

[2] M. Benmammoun, J.M. Fourneau, N. Pekergin and A. Troubnikoff, *An algorithmic and numerical approach to bound the performance of high speed networks*, IEEE MASCOTS 2002. Fort Worth, USA (2002) 375-382.

[3] J. Keilson and A. Kester, Monotone matrices and monotone Markov processes. *Stochastic Process. Appl.* **5** (1977) 231-241.

[4] J.M. Fourneau and N. Pekergin, *An algorithmic approach to stochastic bounds, Performance evaluation of complex systems: Techniques and Tools.* Springer, *Lecture Notes in Comput. Sci.* **2459** (2002) 64-88.

[5] J.M. Fourneau, M. Le Coz, N. Pekergin and F. Quessette, *An open tool to compute stochastic bounds on steady-state distributions and rewards*, IEEE Mascots 03. USA (2003).

[6] J.M. Fourneau, M. Le Coz and F. Quessette, *Algorithms for an irreducible and lumpable strong stochastic bound, Numerical Solution of Markov Chains.* USA (2003).

[7] M. Kijima, *Markov Processes for stochastic modeling.* Chapman & Hall (1997).

[8] W.A. Massey, Stochastic orderings for Markov processes on partially ordered spaces. *Math. Oper. Res.* **12** (1987) 350-367.

[9] N. Pekergin, Stochastic delay bounds on fair queueing algorithms, in *Proc. of INFOCOM'99.* New York (1999) 1212-1220.

[10] N. Pekergin, Stochastic performance bounds by state reduction. *Performance Evaluation* **36-37** (1999) 1-17.

[11] M. Shaked and J.G. Shantikumar, *Stochastic Orders and their Applications.* Academic Press, California (1994).

[12] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models*. John Wiley & Sons, Berlin, Germany (1983).
[13] H.M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*. Academic Press, Florida (1984).
[14] M. Tremolieres, J.-M. Vincent and B. Plateau, *Determination of the optimal upper bound of a Markovian generator*, Technical Report 106. LGI-IMAG, Grenoble, France (1992).
[15] L. Truffet, Near complete decomposability: bounding the error by stochastic comparison method. *Adv. Appl. Probab.* **29** (1997) 830-855.

To access this journal online:
www.edpsciences.org