International Journal of Reliability, Quality and Safety Engineering

BOUNDS FOR THE RELIABILITY OF MULTISTATE SYSTEMS WITH PARTIALLY ORDERED STATE SPACES AND STOCHASTICALLY MONOTONE MARKOV TRANSITIONS

BO HENRY LINDQVIST
Department of Mathematical Sciences,
Norwegian University of Science and Technology,
N-7491 Trondheim, Norway

Received (received date)
Revised (revised date)

Consider a multistate system with partially ordered state space E, which is divided into a set C of working states and a set D of failure states. Let $X(t)$ be the state of the system at time t and suppose $\{X(t)\}$ is a stochastically monotone Markov chain on E. Let T be the failure time, i.e. the hitting time of the set D. We derive upper and lower bounds for the reliability of the system, defined as $P_m(T > t)$ where m is the state of perfect system performance.

Keywords: Multistate system; Markov chain; partial order; stochastic monotonicity; quasi-stationary distribution; exponential distribution

1. Introduction

Consider a system consisting of n components, where the set of possible states for the kth component is a finite set which we denote by E_k $(k = 1, 2, \ldots, n)$. The set of possible states of the system is $E = E_1 \times \cdots \times E_n$. Let the system be monitored from time $t = 0$ and define $X_k(t)$ to be the state of component k at time t, defined for $t \geq 0$. The state of the system at time t is then given by $X(t) = (X_1(t), \ldots, X_n(t))$.

For systems of independent components we may construct Markov models for $\{X(t)\}$ by modelling separately each component process $\{X_k(t)\}$ as a Markov chain. In the case of dependent components, however, Markovian component processes will not in general imply a Markovian system process. Thus, in order to obtain tractable Markov models in the presence of dependence it is convenient to start by the assumption that the system state $\{X(t)\}$ forms a Markov chain on E and then define the possible transitions and transition rates between the states. In the present article we shall study such Markov chains $\{X(t)\}$ on E.

Suppose $C \subset E$ is the set of states in which the system is working, while $D \equiv E \setminus C$ is the set of failure states. The *failure time* T is the time when the set D is first hit, more precisely $T = \inf\{t \geq 0 : X(t) \in D\}$. We define the *reliability* of the system to be the probability $P_m(T > t)$, with the index m meaning that the Markov chain $\{X(t)\}$ starts at time 0 in state m, which we define to be the perfect functioning state of the system. The purpose of the article is to derive upper and lower bounds for the reliability function $P_m(T > t)$.

Bounds for $P_m(T > t)$ in the present context have earlier been derived under the assumption that $\{X(t)\}$ is so called associated in time, see Esary and Proschan[1] for the case of binary components (each E_k is the set $\{0, 1\}$) and Funnemark and Natvig[2] for the case of finite totally ordered state spaces E_k. Practice has shown, however, that bounds based on association may often be too wide to be useful. In this paper we use a slightly different approach based on properties of the so-called *quasi-stationary* distribution of the Markov chain $X(t)$ (see Darroch and Seneta[3]). Moreover, we assume that the Markov chain $\{X(t)\}$ is stochastically monotone with respect to a partial order on the state space E. In fact such an assumption is closely related to association (see Lindqvist[4]), making our approach closely related to the ones cited above.

To increase generality, we assume that the partial order on E in turn is induced by partial orders $\preceq_k$ on each set E_k. The convention is here that larger states with respect to the ordering correspond to better performance of the component. Most literature on multistate systems assume that the state spaces of the components are totally ordered. In practice, however, a totally ordered state space may not be the most natural. For example, the possible states could be "functioning", "failure of type A", "failure of type B", etc. In this case the failure states can not necessarily be ordered in a definite way. Hence we rather have a *partial* ordering, with "functioning" being better than each of the failure states, while no other pairs of states are comparable. More specifically, Caldarola[5] argued that it would be very hard to decide whether or not the state "failed closed" of a circuit breaker is more critical than the state "failed open". In addition, for a given component the ordering of the states may depend upon the system to which the component belongs. A more artificial reason for introducing partially ordered state spaces of components would be the following. Consider a binary system of binary components, where the component set can be partitioned into subsets containing stochastically dependent components, but where the collection of subsets are mutually independent. Then in order to achieve stochastically independent components, one may define "supercomponents" corresponding to each of the independent sets. A supercomponent involving m binary components would then for example have 2^m partially ordered states. This idea was considered by Caldarola[5].

The precise definition of a stochastically monotone Markov chain on a partially ordered state space is given in the next section. Intuitively, stochastic monotonicity means an ageing property of the system in the sense that the remaining time to failure decreases stochastically as the system state is getting worse (with respect to the partial order). Although this is a reasonable assumption in many reliability applications, one may of course think of important cases where it does not hold. It is noticeable, however, that any birth and death process is stochastically monotone. Moreover, stochastic monotonicity of component processes imply stochastic monotonicity of the system process in the case of independent components. This also holds for certain cases of dependent components, under reasonable assumptions on the joint behavior of component processes[4]. Stochastically monotone Markov chains on totally ordered state spaces were considered by Keilson[6], while Lindqvist[4] has studied the finite partially ordered case.

As mentioned above, a key ingredient in our approach is the quasi-stationary distribution of the Markov chain $\{X(t)\}$. This is a distribution $\rho = (\rho_i : i \in E)$ with support on C which is defined as the limiting distribution of the state at time $t \to \infty$, conditioned on the event that the process has not yet (at time t) exited from C (see precise definition in Section 2). The basic property of a quasi-stationary distribution ρ that we will use is that

$$P_\rho(T > t) = e^{-t/E_\rho(T)},$$

which states that the failure time T is exactly exponentially distributed when the initial distribution of the chain is ρ. Quasi-stationary distributions were first considered by Darroch and Seneta[3]. A nice introduction is given by Keilson[6].

The plan of the article is as follows. In Section 2 we give precise definitions and some basic results concerning partial orders, stochastic monotonicity and quasi-stationary distributions. In Sections 3 and 4 we derive upper and lower bounds for the reliability function. A numerical comparison of the results from the present article and results of Esary and Proschan[1] is given in Section 5. Some concluding remarks are given in Section 6.

2. Precise Definitions and Basic Results

2.1. *Partial order*

Recall first that a relation $\preceq$ on a set X is a partial order if it is (i) reflexive, i.e. $x \preceq x$ for all x, (ii) antisymmetric, i.e. $x \preceq y$ and $y \preceq x$ imply $x = y$ and (iii) transitive, i.e. $x \preceq y$ and $y \preceq z$ imply $x \preceq z$.

For the systems considered in this article we assume that the component state spaces are finite partially ordered sets $(E_k, \preceq_k)$. We define the partial order $\preceq$ on their product space E to be the product order defined by

$$(i_1, i_2, \ldots, i_n) \preceq (j_1, j_2, \ldots, j_n) \text{ if and only if } i_k \preceq_k j_k \text{ for } k = 1, 2, \ldots, n.$$

We shall assume throughout the article that each set E_k contains a unique maximal element m_k such that $i_k \preceq m_k$ for all $i_k \in E_k$. The system state $m \equiv (m_1, \ldots, m_n)$ is then a unique maximal element of E. The m_k correspond to perfect functioning of component k while m corresponds to perfect functioning of the system.

Example 1. A simple nontrivial example of a partially ordered state space can be given as follows. Let $E = \{0, 1, 2, 3\}$, where $0 \preceq 1 \preceq 3$, $0 \preceq 2 \preceq 3$, but assume no relation between 1 and 2. Then 3 can be thought of as the perfect state, 0 as the complete failure state while 1 and 2 are intermediate states corresponding to non-perfect conditions which are not necessarily ordered in severity. In fact, the set E may be viewed as the state space of a supercomponent (see Section 1) resulting from two binary components. In this case state 3 corresponds to both components working, state 2 corresponds to component 1 working, component 2 failed etc.

A set $A \subseteq E$ is called *increasing (decreasing)* if $i \in A,\ j \succeq i\ (j \preceq i)$ imply $j \in A$. Note that the set A is decreasing if and only if its complement A^c is increasing. Throughout the present article we shall assume that the set C of working states is an increasing set. This is a reasonable assumption since it means that if the system is working when in a certain state, then any better state (with respect to the partial order) is also a working state. The set $D = E \setminus C$ is then necessarily a decreasing set.

2.2. *Stochastically monotone Markov chain*

Let $\alpha = (\alpha_i : i \in E)$ and $\beta = (\beta_i : i \in E)$ be probability distributions on E. We shall say that α is dominated by β, written $\alpha \preceq \beta$, if $\alpha(A) \leq \beta(A)$ for all increasing subsets $A \subseteq E$.

Let $\{X(t)\}$ be a time-homogeneous Markov chain on E with intensity matrix $Q = (q_{ij})$. If α is a distribution on E, then $P_\alpha(\cdot)$ denotes probabilities computed when the distribution of $X(0)$ is α. In order to stress the dependence on Q, we shall sometimes write $P_{\alpha,Q}(\cdot)$ for $P_\alpha(\cdot)$. For short, we write $P_i(\cdot)$ when α is the measure concentrated in state i. The Markov chain $\{X(t)\}$ is *stochastically monotone* if for any increasing set $A \subseteq E$ and any $t > 0$, $P_i(X(t) \in A)$ is a non-increasing function of i on E (with respect to the partial order $\preceq$). The following equivalent definition in terms of the intensity matrix Q is given in Lindqvist[4]:

The intensity matrix Q is stochastically monotone if for all $i \preceq j$ in E we have

$$Q_i(A) \leq Q_j(A) \text{ for all increasing } A \subseteq E \text{ with } i, j \notin A \tag{1}$$

$$Q_i(A) \geq Q_j(A) \text{ for all decreasing } A \subseteq E \text{ with } i, j \notin A \tag{2}$$

Here $Q_i(A) = \sum_{j \in A} q_{ij}$.

By a trivial generalization of the proof of Theorem 9.3F in Keilson[6] (to allow *partial* order) it follows that if $\{X(t)\}$ is stochastically monotone, and E contains a unique maximal element m, then $P_m(X(t) \in A)$ is a non-increasing function of t for any increasing set A. In the setting of the present article this means that when the system state is defined by a stochastically monotone Markov chain, starting in its “best” state, then the system deteriorates (stochastically) with time.

To describe additional properties of stochastically monotone Markov chains we introduce the space of *sample paths* of the chain $\{X(t)\}$. Let $u > 0$ be a fixed time point, and consider the process $\{X(t)\}$ on the closed time interval $[0, u]$. It is well known that the sample paths of $\{X(t)\}$ for $t \in [0, u]$ can be chosen as members of the space E^u of functions from $[0, u]$ to E which are right continuous and have left limits at all $t \in [0, u]$. Moreover, E^u becomes a Polish space when furnished with the Skorohod metric (see e.g. Kamae et al.[7]). Now the $P_\alpha(\cdot)$ can be viewed as measures on the measurable space $(E^u, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-field in E^u. We shall in the following tacitly assume that any considered subset of E^u is a member of $\mathcal{B}$. A natural partial order on E^u is the pointwise order, i.e. for $x(\cdot), y(\cdot) \in E^u$ we have $x \leq y$ iff $x(t) \preceq y(t)$ for all $t \in [0, u]$. This makes it possible to consider increasing (decreasing) subsets of E^u.

Let now R and Q be two intensity matrices for the Markov chains $\{X(t)\}$ and $\{Y(t)\}$, respectively, both defined on E, but not necessarily stochastically monotone. We shall say that R is *dominated* by Q, written $R \leq Q$, if for all $i \preceq j$ in E we have

$$R_i(A) \leq Q_j(A) \text{ for all increasing } A \subseteq E \text{ with } i, j \notin A \tag{3}$$

$$R_i(A) \geq Q_j(A) \text{ for all decreasing } A \subseteq E \text{ with } i, j \notin A \tag{4}$$

If at least one of R and Q is stochastically monotone, then it is enough to consider $i = j$ in the above definition. If $R \leq Q$, then it follows from Theorem 5 of Kamae et al.[7] that for initial distributions $\alpha \preceq \beta$ and any increasing set $C \subseteq E^u$ we have

$$P_{\alpha,R}(C) \leq P_{\beta,Q}(C) \tag{5}$$

Let α and β be probability distributions on E such that $\alpha \preceq \beta$, and let $\{X(t)\}$ be stochastically monotone with intensity matrix Q. Then by (5) applied to the case $R = Q$ it follows that for any increasing set $C \subseteq E^u$ we have

$$P_\alpha(C) \leq P_\beta(C) \tag{6}$$

Now, let D be a decreasing subset of E and let T denote the hitting time of the set D, as defined in Section 1. Then $\{T > t\}$ defines an increasing set in E^u when $u > t$. Hence by (6), for stochastically monotone $\{X(t)\}$, we have

$$P_\alpha(T > t) \leq P_\beta(T > t) \text{ whenever } \alpha \preceq \mu \tag{7}$$

The relation (7) is intuitively reasonable in our setting, as it may be interpreted to say that with a better initial state, the probability of no failure before time t increases.

Example 2. Suppose E is given as in Example 1 and let Q be the intensity matrix of a Markov chain on E. The increasing subsets of E are $\{3\}$, $\{1,3\}$, $\{2,3\}$, $\{1,2,3\}$, while the decreasing ones are $\{0\}$, $\{0,1\}$, $\{0,2\}$, $\{0,1,2\}$. Thus by going through all possible cases of (1) and (2) in the definition of stochastic monotonicity we conclude that Q is stochastically monotone if and only if the following eight relations all hold:

$$\begin{array}{rccc}
\text{(i)} & q_{10} + q_{12} & \geq & q_{30} + q_{32} \\
\text{(ii)} & q_{20} + q_{21} & \geq & q_{30} + q_{31} \\
\text{(iii)} & q_{12} + q_{13} & \geq & q_{02} + q_{03} \\
\text{(iv)} & q_{21} + q_{23} & \geq & q_{01} + q_{03} \\
\text{(v)} & q_{10} & \geq & q_{30}
\end{array}$$

(vi)	q_{20}	$\geq$	q_{30}
(vii)	q_{13}	$\geq$	q_{03}
(viii)	q_{23}	$\geq$	q_{03}

If transitions between non-related states are not allowed (i.e. Q is *up-down*, to be defined below), then $q_{12} = q_{21} = 0$ and it is seen that the inequalities (v)-(viii) are implied by (i)-(iv). If we further assume that direct transitions between the extreme states 0 and 3 are not possible, then Q is stochastically monotone if and only if $q_{10} \geq q_{32},\ q_{20} \geq q_{31},\ q_{13} \geq q_{02},\ q_{23} \geq q_{01}$. In the case of a supercomponent (see Example 1), these inequalities state that the failure rate of each (sub)component increases and the repair rate decreases, when the other component goes from working to failed.

2.3. *Stationary and quasi-stationary distribution*

Next we introduce the concept of an *ML-matrix*. A square finite matrix $B = (b_{ij})$ is called an ML-matrix if $b_{ij} \geq 0$ for all $i \neq j$. If B is an ML-matrix, then for sufficiently large a, $S = aI + B$ is a nonnegative matrix, where I is the identity matrix. We call B irreducible if S is irreducible, i.e. if for any pair i, j there is a positive integer n with $s_{ij}^{(n)} > 0$. If B is irreducible, then by Chapter 2 in Seneta[8] there exists an eigenvalue τ (which we shall call the Perron-Frobenius eigenvalue of B) such that $\tau > \mathrm{Re}(\nu)$ for any other eigenvalue ν of B (where $\mathrm{Re}(\nu)$ is the real part of the possibly complex number ν). Moreover, to τ correspond unique, up to constant multiples, strictly positive left and right eigenvectors.

Any intensity matrix Q of a Markov chain $\{X(t)\}$ is an ML-matrix. The property that Q is irreducible as an ML-matrix then corresponds to the irreducibility of the Markov chain $\{X(t)\}$ in the ordinary terminology, i.e. that any state can be reached from any other state. Moreover, for an intensity matrix Q we have $\tau = 0$, while the left and right eigenvectors of Q at 0 are, respectively, the stationary distribution π (if it exists) and a column of all $1s$.

The intensity matrix Q is called *up-down* if $q_{ij} > 0$ implies $i \preceq j$ or $i \succeq j$ for all $i, j \in E,\ i \neq j$. This means that any change of state of the system is to a state which is either better or worse with respect to the partial order.

In the article we will use the notation $Q_C = (q_{ij} : i, j \in C)$. Thus Q_C is the restriction of Q to C. We will also use the corresponding notation for vectors, e.g. $\rho_C = (\rho_i : i \in C)$.

It is well known that Q irreducible implies the existence of a unique stationary distribution $\pi = (\pi_i : i \in E)$. Moreover, Q_C irreducible implies the theorem below.

Theorem 1 (Darroch and Seneta[3]) *If Q_C is irreducible, then for any initial distribution α with $\alpha(C) > 0$, the limits*

$$\rho_j = \lim_{t\to\infty} P_\alpha(X(t) = j \mid T > t) \ ;\ j \in C \tag{8}$$

exist, with the ρ_j being strictly positive and not depending on α. The vector ρ_C is the unique normalized (to norm 1) left eigenvector of Q_C for the Perron-Frobenius *eigenvalue* $\tau \equiv -a$ (< 0). *Moreover, define* ρ *to be the probability distribution on all of* E *with probability mass* ρ_j *for* $j \in C$ *and mass* 0 *for* $j \in D$. *Then*

$$P_\rho(T > t) = e^{-at} \quad \textit{for all } t > 0 \tag{9}$$

i.e. the distribution of T *is exactly exponential when* ρ *is the initial distribution.*

Following Darroch and Seneta[3] we shall call the distribution ρ the *quasi-stationary distribution* of $\{X(t)\}$. Note the dependence of ρ on the set C. As indicated by (8), the quasi-stationary distribution can be interpreted as the conditional distribution of the state of the system, $X(t)$, at a large time t, conditioned on the event that the system has not yet failed (i.e. $X(t)$ has not yet left the set C). Moreover, (9) can be interpreted to say that under this condition, the remaining time to failure is exactly exponentially distributed, with expectation $E_\rho(T) = 1/a$.

3. Bounds for the Reliability Function

In the present section we show how to derive simple upper and lower bounds for the reliability function $P_m(T > t)$ in terms of the quasi-stationary distribution ρ and the associated eigenvalue $-a$.

From (7) follows directly that if Q is stochastically monotone and if a quasi-stationary distribution ρ exists (e.g. if Q_C is irreducible), then

$$0 \le P_m(T > t) - P_\rho(T > t) \le (1 - \rho_m) P_m(T > t)$$

From this we obtain

$$e^{-at} \le P_m(T > t) \le \rho_m^{-1} e^{-at} \tag{10}$$

The upper bound in (10) appears, however, to be of little value when ρ_m is not close to 1. There is thus a need for improvement of this bound. A possible approach is as follows. First, write

$$P_\rho(T > t) = \sum_{i \in C} \rho_i P_i(T > t) = \rho_m P_m(T > t) + \sum_{i \neq m} \rho_i P_i(T > t)$$

so that

$$P_m(T > t) = \rho_m^{-1} \left[P_\rho(T > t) - \sum_{i \neq m} \rho_i P_i(T > t) \right]$$

From this we get directly:

Theorem 2 *Assume that* Q *is stochastically monotone and* Q_C *is irreducible. Suppose lower bounds* $b_i(t)$ *of* $P_i(T > t)$ *are available for all* $i \in J$, *where* J *is some subset of* $C \setminus \{m\}$. *Then*

$$e^{-at} \le P_m(T > t) \le \rho_m^{-1} \left[e^{-at} - \sum_{i \in J} \rho_i b_i(t) \right] \tag{11}$$

This inequality generalizes the upper bound of (10), as the latter corresponds to $J = \emptyset$ (the empty set). It seems worthwhile to try to derive nontrivial bounds $b_i(t)$ for at least the i corresponding to the largest ρ_i, $i \neq m$. The question remains to derive functions $b_i(t)$ that are relatively simple to compute and otherwise do fairly well. Our idea here is to define new Markov chains from the old one and then apply the left inequality in Theorem 2 to the derived chains.

In the following, fix one element in J and call it m'. Define the set E' by

$$E' = \{i \in E : i \preceq m'\}$$

Then E' is a finite partially ordered set, with partial order inherited from E, and with a unique maximal element m'. We shall consider a Markov chain $\{X'(t)\}$ on E' with intensity matrix $Q' = (q'_{ij} : i, j \in E')$, where $q'_{ij} = q_{ij}$ when $i, j \in E'$, $i \neq j$, and diagonal elements q'_{ii} defined so that the row sums equal 0. (Note that $Q' \neq Q_{E'}$). Define now $D' = D \cap E'$. Then D' is a decreasing set with respect to the partially ordered set $(E', \preceq)$. Let furthermore T' denote the hitting time of D' for the Markov chain $\{X'(t)\}$. Then we have the following result.

Lemma 1 *Suppose Q is stochastically monotone and up-down. Then*

$$P_{m',Q}(T > t) \geq P_{m',Q'}(T' > t)$$

Proof. We shall define another Markov chain $\{Y(t)\}$ on E, given by the intensity matrix $R = (r_{ij})$ with off-diagonal elements given by

$$r_{ij} = \begin{cases} 0 & \text{if } i \in E',\, j \in E \setminus E' \\ q_{ij} & \text{otherwise} \end{cases}$$

and r_{ii} defined so that the row sums of R equal 0.

As there are no positive transition rates from E' to $E \setminus E'$, it is seen that for an initial state in E', the chain $\{Y(t)\}$ behaves exactly as $\{X'(t)\}$. Thus

$$P_{m',R}(T > t) = P_{m',Q'}(T' > t)$$

By (5) we are therefore done if we can show that $R \leq Q$.

Since Q is stochastically monotone, we must prove that, for any i, $R_i(A) \leq (\geq) Q_i(A)$ for all increasing (decreasing) sets $A \subseteq E$ with $i \notin A$.

Let therefore $i \in E$ be given. First, let A be an increasing set with $i \notin A$. Then $R_i(A) \leq Q_i(A)$ by the definition of R.

Next, let A be a decreasing set with $i \notin A$. We split up into the following two cases:

$i \in E \setminus E'$: Then we have $R_i(A) = Q_i(A)$ by the definition of R.

$i \in E'$: Then $R_i(A) = Q_i(A \cap E')$. We are done if we can prove that $Q_i(A \cap E') = Q_i(A)$. To do this, suppose for contradiction that $q_{ik} > 0$ for some $k \in A \setminus E'$. Then, by the assumed up-down property of Q, either $i \preceq k$ or $i \succeq k$. The former case is impossible, as it would imply $i \in A$ because $k \in A$ and A is decreasing. The latter case is, however, also impossible, as it would imply $k \in E'$ because $i \in E'$ and E' is a decreasing set in E. Thus $q_{ik} = 0$ for all $k \in A \setminus E'$ and we are done.

The lemma suggests that lower bounds $b_{m'}(t)$ may be based on $P_{m'}(T' > t)$, either by direct computation of these probabilities, or by further bounding them from below. An example is given in Section 5. If $\{X'(t)\}$ is stochastically monotone (which is not guaranteed from the monotonicity of $\{X(t)\}$) and $E' \cap C$ is irreducible (with respect to $\{X'(t)\}$), then we can use the lower bound of Theorem 2. If $\{X'(t)\}$ is not monotone, then we may construct a monotone R' with $R' \leq Q'$ and use (5) to compute a bound.

The bounds $b_i(t)$ of the preceding paragraph have the property that $b_i(0) = 1$, $b_i(\infty) = 0$. Using them in (11) thus yields an upper bound $h(t)$ for $P_m(T > t)$ with $h(0) = (1 - \sum_{i \in J} \rho_i)\rho_m^{-1}$ and $h(\infty) = 0$. Thus we can have $h(0)$ as close to 1 as we wish, by increasing J. With $J = C \setminus \{m\}$ we get $h(0) = 1$.

4. Alternative Lower Bound for the Reliability Function

Recall from Theorem 2 that under the given conditions we have $P_m(T > t) \geq e^{-at}$, where $-a$ is the Perron-Frobenius eigenvalue of the restriction Q_C of Q to the set C. In practice one may not want to compute the exact value of a, so one might be interested in more easily available upper bounds c, say, for a. Of course, if $a \leq c$, then $P_m(T > t) \geq e^{-ct}$.

The result below is a consequence of a more general result of Lindqvist[9] on the Perron-Frobenius eigenvalue of ML-matrices. The possible advantage of the result is that instead of doing computations related to the quasi-stationary distribution one needs only compute an ordinary stationary distribution. We also note that the inequality (12) does not require stochastic monotonicity of Q.

Let, as before, Q_C be the restriction of Q to C. Define now Q_C° so that Q_C and Q_C° coincide outside the main diagonal, and let the diagonal elements of Q_C° be defined so that each row of Q_C° sum up to 0. Then Q_C° is the intensity matrix of a Markov chain on C, and hence the Perron-Frobenius eigenvalue of Q_C° is 0. Note, furthermore, that

$$Q_{C,ii} - Q_{C,ii}^\circ = -Q_i(D)$$

Let π° be the stationary distribution of the Markov chain on C with intensity matrix Q_C°. Then we have:

Theorem 3 (Lindqvist[9]) *If Q_C is irreducible, then*

$$a \leq \sum_{i \in C} \pi_i^\circ Q_i(D) \equiv c, \tag{12}$$

and hence

$$P_m(T > t) \geq e^{-ct} \tag{13}$$

Example 3. Suppose $E = \{0, 1, 2\}$ with total order $0 \preceq 1 \preceq 2$ and transition rates given by $q_{21} = q_{12} = q_{02} = 100$, $q_{10} = 1$ and $q_{20} = q_{01} = 0$. The resulting Markov chain is stochastically monotone, and with $C = \{1, 2\}$ we have Q_C irreducible. Now Theorem 3 gives $a \leq 0.5$, whereas the exact value of a is 0.4988.

Remark: Time-reversible chains.

Recall that the Markov chain $\{X(t)\}$ by definition is *time-reversible* if there exist positive constants $(z_i : i \in E)$ such that

$$z_i q_{ij} = z_j q_{ji} \tag{14}$$

for all $i, j \in E$, $i \neq j$. The stationary distribution $\pi = (\pi_i : i \in E)$ of the Markov chain $\{X(t)\}$ on E is found by norming the z_i to have sum 1. The relations (14) obviously also carry over to the Markov chain with infinitesimal intensity matrix Q_C° encountered in Theorem 3. It therefore follows that in the time-reversible case we have

$$\pi_i^\circ = \frac{\pi_i}{\pi(C)} \equiv \pi_i^*$$

where π^* is the so-called *ergodic exit distribution* considered by Keilson[6].

In Theorem 6.9C of Keilson[6] is shown that in the case of time-reversibility (not assuming stochastic monotonicity) we have

$$P_{\pi^*}(T > t) \leq P_\rho(T > t) = e^{-at}. \tag{15}$$

The left hand side here can be interpreted as the probability that a presently working system which has been running for a long time, will continue to work for at least a time t. The right hand side, on the other hand, is the corresponding probability for the case when the system is working and has not yet visited the failure states.

The next example shows that (15) does not necessarily hold for non-reversible stochastically monotone chains.

Example 3 (continued). The Markov chain of the example is not time-reversible since $q_{20} = 0$, while $q_{02} > 0$. Moreover, we compute $\pi^* = (0, 0.4975, 0.5025)$, so $\pi^* \succ \rho$ and hence $P_{\pi^*}(T > t) > P_\rho(T > t)$.

Remark: Bounding the quasi-stationary distribution.

The following lemma is a straightforward consequence of the fact that $-a$ is an eigenvalue of Q_C.

Lemma 2 *If Q_C is irreducible, then*

$$a = \sum_{i \in C} \rho_i Q_i(D) \tag{16}$$

Proof: The relation $\rho Q_C = \tau\rho$ is equivalent to $\sum_{i \in C} \rho_i q_{ij} = \tau\rho_j$ for all $j \in C$. Summing over j we get $\sum_{i \in C} \rho_i \sum_{j \in C} q_{ij} = \tau$ from which the lemma follows since $\sum_{j \in C} q_{ij} = -Q_i(D)$ by the fact that Q is an intensity matrix.

By Lemma 2 and Theorem 3 we have

$$\sum_{i \in C} \rho_i Q_i(D) \leq \sum_{i \in C} \pi_i^\circ Q_i(D) \tag{17}$$

Since $Q_i(D)$ is a decreasing function of $i \in C$ it might be conjectured from (17) that

$$\pi^\circ \preceq \rho_C \tag{18}$$

Indeed, (18) holds in Example 3 since we there have $\pi^\circ = (0.5, 0.5)$. However, (18) does not hold in general as will follow from the next example.

Example 4. Let $(E, \preceq)$ be as in Example 1. Let $q_{01} = 1,\ q_{10} = 1,\ q_{02} = 1,\ q_{20} = 1410,\ q_{13} = 5,\ q_{31} = 2,\ q_{23} = 90,\ q_{32} = 1$. Then Q is stochastically monotone. Let $C = \{1, 2, 3\}$. A computation shows that

$$\begin{aligned} \rho_C &= (0.2838, 0.0005, 0.7157) \\ \pi^\circ &= (0.2835, 0.0079, 0.7087) \end{aligned}$$

so $\pi^\circ \not\preceq \rho_C$ since π° gives the largest mass to the increasing set $\{2, 3\}$.

5. Comparison of Results: a Binary 2-out-of-three System

Consider a system with three binary components, so that $E_1 = E_2 = E_3 = \{0, 1\}$. The components are assumed to be independent of each other, each with a failure rate λ (transition rate from state 1 to state 0) and repair rate μ (transition rate from 0 to 1). We assume that the system is a 2-out-of-3 system, which means that the system works if and only if at least two components are in state 1.

It is convenient to denote the system states by the ordered vector of the corresponding component states, 101 etc. Then the set of working states is $C = \{111, 110, 101, 011\}$, with $m = 111$ corresponding to the perfect state. This example was considered by Esary and Proschan[1], who computed lower bounds for $P_m(T > t)$ by using properties of associated stochastic processes. The purpose of the present example is to compare our bounds with theirs, and also to exemplify the new results.

First we ascertain that the corresponding Markov chain is indeed stochastically monotone. This follows by inspection of the intensity matrix Q, but is also a simple consequence of the fact that the component processes are independent birth and death processes. Next, with C given as above we have

$$Q_C = \begin{bmatrix} -3\lambda & \lambda & \lambda & \lambda \\ \mu & -(2\lambda+\mu) & 0 & 0 \\ \mu & 0 & -(2\lambda+\mu) & 0 \\ \mu & 0 & 0 & -(2\lambda+\mu) \end{bmatrix}.$$

The quasi-stationary distribution on C is found as the unique probability vector ρ_C on C satisfying $\rho_C Q_C = \tau \rho_C$ where τ is the largest eigenvalue of Q_C. The solution is

$$\rho_C = (\rho_m, (1-\rho_m)/3, (1-\rho_m)/3, (1-\rho_m)/3),$$

where $\rho_m = (\sqrt{\delta^2+10\delta+1}-\delta-1)/4$ and $\delta = \mu/\lambda$. Moreover, $a \equiv -\tau = 2\lambda(1-\rho_m)$.

To compute c defined in Theorem 3 we first note that Q_C° is obtained by changing the three lower diagonal elements of Q_C to $-\mu$. Then from (12) we get $c = 6\lambda/(\delta+3)$.

Table 1 presents the bounds of Theorems 2 and 3, together with the exact values and the bounds of Esary and Proschan[1] (EP), computed for selected values of δ and $\gamma \equiv \lambda t$.

For computation of the upper bound of Theorem 2 we have used the somewhat crude bounds $P_i(T > t) \geq b_i(t) = e^{-2\lambda t}$ for $i = 110, 101, 011$. (These correspond to ignoring the possibility of transition to 111 before leaving C).

Table 1. Bounds for the reliability of a 2-out-of-3 system.

γ	δ	Exact	Thm 2	Thm 3	EP	Thm 2 (upper)
1	2	.44	.37	.30	.36	.60
1	10	.682	.663	.630	.650	.799
1	100	.9449	.9444	.9434	.9439	.9682
10	20	.0895	.0886	.074	.074	.101
10	100	.565	.565	.558	.559	.581

It is seen that the simple lower bound e^{-at} of Theorem 2 using the Perron-Frobenius eigenvalue of Q_C is better than the EP bound. On the other hand, the more easily computed bound e^{-ct} from Theorem 3 is beaten by the EP bound, but otherwise seems to behave satisfactorily at least for highly reliable systems. The upper bounds using Theorem 2 are further away from the exact values than are the lower bounds. However, they may still be useful especially for highly reliable systems.

6. Concluding Remarks

The motivation for the article is a study of bounds for the reliability of multistate systems. However, it is clear that the obtained results are valid for any Markov chain with finite partially ordered state space and stochastically monotone transitions. This of course extends the application area of the results.

The main idea of the approach is that, under the assumption of stochastic monotonicity, one obtains simple and useful bounds for the reliability function $P_m(T > t)$ by computing the quasi-stationary distribution and the corresponding eigenvalue. These bounds are given in Theorem 2, while Theorem 3 is a further simplification involving a stationary distribution. It is believed, moreover, that the results of the article give reasonable approximations also in the case of non-monotonicity.

The computations reported in the previous section indicate that the bounds are improving as systems become more reliable. Preliminary investigations for larger systems seem to confirm this. Theoretically, this fact is closely connected to limit results (e.g. Keilson[6]) which state that life times of highly reliable systems tend to be exponentially distributed. The advantage of giving bounds, as in this article, is that they are not merely limit results but applicable in any "non-asymptotic" case.

An interesting question is, however, whether it is worthwhile to consider just upper and lower bounds, given that modern computers are able to solve numerically the exact equations even for quite large systems. In many cases one would of course do the complete computations. However, it is clear that the numerical problem of just computing the largest eigenvalue and the corresponding eigenvector of a matrix, is considerably less intensive, and thus enabling one to consider more complex problems. Also, in many instances rough bounds of reliability are all that is needed, particularly when taking into account the uncertainty already inherent in the input data. A further motivation for our study is of course that the bounds are of theoretical interest and may lead to new insight into the properties of Markov chains.

References

1. Esary, J.D., Proschan, F. (1970). A reliability bound for systems of maintained, interdependent components. *Journal of the American Statistical Association*, **65**, 329-338.
2. Funnemark, E. and Natvig, B. (1986). Bounds for the availability in a fixed time interval for a multistate monotone system of multistate, maintained, interdependent components. *Advances in Applied Probability*, **17**, 638-665.
3. Darroch, J.N. and Seneta, E. (1965). On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, bf 2, 88-100.
4. Lindqvist, B. H. (1987). Monotone and associated Markov chains, with applications to reliability theory. *Journal of Applied Probability*, **24**, 679-695.
5. Caldarola, L. (1979) Coherent systems with multistate components, *Nuclear Engineering and Design*, **58**, 127-139.
6. Keilson, J. (1979). *Markov Chain Models - Rarity and Exponentiality.* Springer, New York.
7. Kamae, T., Krengel, U. and O'Brien, G.L. (1977). Stochastic inequalities on partially ordered spaces. *Annals of Probability* **5**, 899-912.
8. Seneta, E.(1973). *Non-Negative Matrices*. Allen & Unwin, London.
9. Lindqvist, B. H. (2002). On comparison of the Perron-Frobenius eigenvalues of two ML-matrices. *Linear Algebra and its Applications* **353**, 257-266.