Uncertainty Bounds for a Monotone Multistate System

Helge Langseth
SINTEF Industrial Management
Department of Safety and Reliability

Bo Henry Lindqvist
Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

We consider a monotone multistate system with conditionally independent components given the component reliabilities, and random component reliabilities. Upper and lower bounds are derived for the moments of the random reliability function, extending results for binary systems. The second moment of the reliability function is given special attention, as this quantity is used to calculate the standard deviation of the system availability estimate.

KEY WORDS: MONOTONE MULTISTATE SYSTEM; RELIABILITY FUNCTION; UNCERTAINTY; ASSOCIATION

1 Introduction

We shall consider a *multistate monotone system* with n components, as defined and studied for example in Griffith [9] and Natvig [14].

For $i = 1, \ldots, n$, let X_i denote the state of the ith component, the set of possible states being $S_i = \{0, 1, \ldots, M_i\}$. We assume that the states in each S_i are linearly ordered, with 0 denoting the worst state ("complete failure") and $M_i \geq 1$ being the best state ("perfect functioning").

The state vector of the system is $\boldsymbol{X} = (X_1, \ldots, X_n)$, with possible values in $\mathcal{S} \equiv S_1 \times \cdots \times S_n$. We furnish $\mathcal{S}$ with the componentwise partial order $\leq$ defined by

$$\boldsymbol{x} = (x_1, \ldots, x_n) \leq \boldsymbol{y} = (y_1, \ldots, y_n) \Leftrightarrow x_i \leq y_i \text{ for } i = 1, \ldots, n$$

For convenience we shall later write $\boldsymbol{x} < \boldsymbol{y}$ to mean that $\boldsymbol{x} \leq \boldsymbol{y}$ *and* $x_i < y_i$ for at least one i.

In multistate reliability theory, one defines a *structure function* $\Phi : \mathcal{S} \to \{0, 1, \ldots, M\}$, with $\Phi(\boldsymbol{x})$ denoting the *state of the system* when the components are in states $\boldsymbol{x}$, see e.g. [14]. Here $M(\geq 1)$ denotes the perfect system state, and states are then ordered down to 0, which corresponds to complete system failure. The function Φ is assumed to be *increasing* with respect to the partial order on $\mathcal{S}$, i.e.

$$\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \Phi(\boldsymbol{x}) \leq \Phi(\boldsymbol{y})$$

which means that improvements of one or more component states cannot lead to a lower performance of the system. Natural requirements are also $\Phi(0, 0, \ldots, 0) = 0$ and $\Phi(M_1, M_2, \ldots, M_n) =$

M. Note that here and in the sequel we shall write *increasing* to mean the more precise notion of *non-decreasing*.

Let

$$P_{ij} = P(X_i = j) \ ; \ i = 1, \ \ldots, \ n, \ \ j = 0, \ \ldots, \ M_i$$

The marginal distribution of X_i is thus given by the vector $\boldsymbol{P}_i \equiv (P_{i0}, \ \ldots, \ P_{iM_i})$. We shall in this paper assume that the component states $X_1, \ \ldots, \ X_n$ are stochastically independent given the $\boldsymbol{P}_i$, so that the array

$$\boldsymbol{P} \equiv \begin{pmatrix} \boldsymbol{P}_1 \\ : \\ \boldsymbol{P}_n \end{pmatrix}$$

completely specifies the joint distribution of $\boldsymbol{X}$.

For later use we shall also define

$$F_i(j) = \mathrm{P}(X_i \leq j), \quad R_i(j) = \mathrm{P}(X_i \geq j)$$

for $i = 1, \ \ldots, \ n, \ j = 0, 1, \ \ldots, \ M_i$.

A main objective in multistate reliability theory is to compute probabilities of the form

$$h_j \equiv h_j(\boldsymbol{P}) = \mathrm{P}(\Phi(\boldsymbol{X}) \geq j) \ ; \ j = 1, \ \ldots, \ M \tag{1}$$

More generally one might be interested in computing expected values $\mathrm{E}\Phi(\boldsymbol{X})$. But

$$\mathrm{E}\Phi(\boldsymbol{X}) = \sum_{j=1}^{M} \mathrm{P}(\Phi(\boldsymbol{X}) \geq j) = \sum_{j=1}^{M} h_j$$

so we are back to the problem of computing the h_j in (1). (See, however, the last section regarding the computation of moments of $\mathrm{E}\Phi(\boldsymbol{X})$ when $\boldsymbol{P}$ and hence the h_j are random).

The sets $\{\boldsymbol{x} : \Phi(\boldsymbol{x}) \geq j\}$ are *increasing sets* with respect to the partial order $\leq$ on $\mathcal{S}$, (recall that a subset $\mathcal{C}$ of $\mathcal{S}$ is called *increasing* if $\boldsymbol{x} \in \mathcal{C}$ and $\boldsymbol{y} \geq \boldsymbol{x}$ implies $\boldsymbol{y} \in \mathcal{C}$). Thus one is in effect interested in computing probabilities

$$h_{\mathcal{C}} \equiv h_{\mathcal{C}}(\boldsymbol{P}) = P(\boldsymbol{X} \in \mathcal{C})$$

for increasing subsets $\mathcal{C}$ of $\mathcal{S}$. Note that to any increasing set $\mathcal{C}$ there exists a structure function Φ and a j such that $h_{\mathcal{C}}$ is of the form h_j, just let Φ be the indicator function of $\mathcal{C}$,

$$I_{\mathcal{C}}(\boldsymbol{x}) = \begin{cases} 1 \text{ if } \boldsymbol{x} \in \mathcal{C} \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

We may think of $h_{\mathcal{C}}$ as the *functioning* probability of the system.

Exact computation of the $h_{\mathcal{C}}$ for given $\boldsymbol{P}$ and $\mathcal{C}$ is in principle straightforward (see section 2), but can for large systems be a formidable (often practically impossible) task. Upper and lower bounds can be computed for example by the methods considered in Funnemark and Natvig [8] (see lemma 5 in section 5).

In this paper we assume that the array $\boldsymbol{P} = (P_{ij})$ defining the distribution of $\boldsymbol{X}$ is *random*. This in turn makes $h_{\mathcal{C}}(\boldsymbol{P})$ a random variable, which is the focus of our interest in the paper.

As there is usually no hope of finding the exact distribution of $h_{\mathcal{C}}(\boldsymbol{P})$, we will mainly be concerned with its moments, i.e.

$$\mu_m \equiv \mathrm{E}[h_{\mathcal{C}}^m(\boldsymbol{P})] \; ; \; m = 1, 2, \ldots$$

Note that here and in the following we use the notation $h^m(\cdot)$ to mean $\{h(\cdot)\}^m$.

The motivation for our study is to establish a basis for *uncertainty analysis* and *Bayesian estimation* in connection with multistate systems. The distribution of $\boldsymbol{P}$ then reflects our uncertainty about $\boldsymbol{P}$. Moreover, a prior uncertainty may be updated through observations to give a posterior distribution for $\boldsymbol{P}$.

As an example, assume that the $\boldsymbol{P}_i = (P_{i0}, \ \ldots, \ P_{iM})$ for $i = 1, \ldots, n$ are stochastically independent random vectors. A rather flexible and mathematically tractable probability model for the distribution of $\boldsymbol{P}_i$ is the *Dirichlet-distribution.* Recall that a random vector $\boldsymbol{Y} = (Y_0, Y_1, \ldots, Y_M)$ is *Dirichlet-distributed* with parameters $(\theta_0, \theta_1, \ \ldots, \ \theta_M)$ if the joint density is given by

$$\frac{\Gamma(\sum_{j=0}^M \theta_j)}{\prod_{j=0}^M \Gamma(\theta_j)} \prod_{j=0}^M y_j^{\theta_j - 1} \text{ for } y_0, \ \ldots, \ y_M \geq 0, \sum_{j=0}^M y_j = 1 \tag{3}$$

where $\Gamma(\cdot)$ is the usual gamma-function. Note that for $j = 0, \ \ldots, \ M$, the marginal distribution of Y_j is a Beta-distribution with parameters $(\theta_j, \sum_{i \neq j} \theta_i)$. The Dirichlet distribution is the conjugate family for the multinomial distribution. This means that if the prior distribution of the $\boldsymbol{P}_i$ are Dirichlet distributions, and we have multinomial data on the components, then the posterior will also be a Dirichlet distribution.

A considerable number of papers consider uncertainty analysis for binary systems (see e.g. [12] for further references). The purpose of this paper is to show how some of the ideas from the binary case can be carried over to the multistate one. As we shall see, this extension requires some care concerning assumptions on the distribution of $\boldsymbol{P}$. As for the binary case, the systems very easily become too complex to allow exact computation of moments. Thus approximative methods are needed. Simulation is one possibility. We shall instead focus attention on easily computable upper and lower bounds for the moments. Bounds for the second moment are of particular interest in practice since they lead to bounds for the standard deviation.

The given setup can easily be modified to cover more general applications, e.g. flow networks (Doulliez-Jamoulle [7], Aven [4]) and Pert networks (see e.g. Stoyan [18]). Typically, in these cases the states of the components are modeled as discrete nonnegative random variables with state spaces $S_i = \{x_{i0}, \ \ldots, \ x_{iM_i}\}$, and the main interest is in the throughput (of a flow network) and the duration time of the project (for Pert-networks), which are often given by increasing functions Φ of the component state vector.

In section 2 we consider exact computation of the moments μ_m. The main section is section 3. Here we establish an upper bound for μ_m in the case of independent component reliability vectors $\boldsymbol{P}_i$. In section 5 we consider positively dependent (associated) component reliability vectors. The paper is concluded by an example, and some important and useful extensions of the results.

2 Exact computation of moments

As in section 1, let $\mathcal{C}$ be an increasing subset of $\mathcal{S}$, with indicator function $I_{\mathcal{C}}$ and functioning probability $h_{\mathcal{C}}$. In order to simplify notation, we shall mostly consider the increasing set $\mathcal{C}$ fixed, and we may thus drop the subscript $\mathcal{C}$.

Then for given $\boldsymbol{P}$ we have

$$h(\boldsymbol{P}) \equiv \mathrm{P}(\boldsymbol{X} \in \mathcal{C}) = \sum_{\boldsymbol{x} \in \mathcal{C}} \prod_{i=1}^{n} P_{i,x_i} \tag{4}$$

$$= \sum_{\boldsymbol{x} \in \mathcal{S}} I(\boldsymbol{x}) \prod_{i=1}^{n} P_{i,x_i} \tag{5}$$

where $I(\cdot) \equiv I_{\mathcal{C}}(\cdot)$ is given by (2). Computing the mth power and taking expectation we get

$$\mu_m = \sum_{\boldsymbol{x}^{(1)},\ \ldots,\ \boldsymbol{x}^{(m)} \in \mathcal{S}^m} \{\prod_{k=1}^{m} I(\boldsymbol{x}^{(k)})\}\mathrm{E}\{\prod_{i=1}^{n}\prod_{k=1}^{m} P_{i,x_i^{(k)}}\} \tag{6}$$

It is clear that this formula is not suited for practical computations, since the sum is over $(\prod_{i=1}^{n}(M_i+1))^m$ terms. However, the expression will prove useful for establishing an upper bound, see section 3

Note that if the random vectors $\boldsymbol{P}_1,\ \ldots,\ \boldsymbol{P}_n$ are stochastically independent, then the expectation operator E in (6) can be put inside the product sign $\prod_{i=1}^{n}$. In this case we have for $m = 1$,

$$\mu_1 \equiv \mathrm{E}h(\boldsymbol{P}) = h(\mathrm{E}\boldsymbol{P})$$

where $\mathrm{E}\boldsymbol{P}$ denotes the matrix of expected values of the elements in $\boldsymbol{P}$. Corresponding simple exact results do not exist for the higher moments, however.

For practical computation of the μ_m we need an expression for $h(\boldsymbol{P})$ with considerably fewer terms than (4)-(5). For binary systems useful expressions are obtained e.g. by algorithms for "sum of disjoint products", see the review by Agrawal and Barlow [1]. A corresponding algorithm for multistate systems has been suggested by Aven [3], who adapts an algorithm by Doulliez-Jamoulle [7] for expected throughput in flow-networks to multistate reliability systems with given minimal path and -cut vectors.

For later reference we recall the definition of minimal path vectors and minimal cut vectors. First, a vector $\boldsymbol{y} \in \mathcal{S}$ is a *path* vector to the increasing set $\mathcal{C}$ if $\boldsymbol{y} \in \mathcal{C}$. We call $\boldsymbol{y}$ a *minimal path vector* if $\boldsymbol{y}$ is a path vector and $\boldsymbol{x} < \boldsymbol{y} \Rightarrow \boldsymbol{x} \notin \mathcal{C}$. The minimal path vectors will be denoted $\boldsymbol{y}^{(1)},\ \ldots,\ \boldsymbol{y}^{(k)}$.

A vector $\boldsymbol{z} \in \mathcal{S}$ is a *cut* vector to the increasing set $\mathcal{C}$ if $\boldsymbol{z} \notin \mathcal{C}$. We call $\boldsymbol{z}$ a *minimal cut vector* if $\boldsymbol{z}$ is a cut vector and $\boldsymbol{x} > \boldsymbol{z} \Rightarrow \boldsymbol{x} \in \mathcal{C}$. The minimal cut vectors will be denoted $\boldsymbol{z}^{(1)},\ \ldots,\ \boldsymbol{z}^{(l)}$.

It is well known that an increasing set $\mathcal{C}$ can be represented by means of either the minimal path or minimal cut vectors in the following manner,

$$\mathcal{C} = \cup_{r=1}^{k}\{\boldsymbol{x} : \boldsymbol{x} \geq \boldsymbol{y}^{(r)}\} \tag{7}$$

$$\mathcal{C} = \cap_{r=1}^{l}\{\boldsymbol{x} : \boldsymbol{x} \not\leq \boldsymbol{z}^{(r)}\} \tag{8}$$

3 Upper bound for μ_m in the case of independent $\mathbf{P}_i$

Let the situation be as in the previous section. The purpose of this section is to show how the upper bounds derived in Lindqvist [11] can be generalized to the multistate case. One advantage of these bounds is that they can be computed from a single evaluation of $h(\boldsymbol{A})$, with a suitably chosen parameter matrix $\boldsymbol{A}$. Their usefulness for binary systems is studied in [13].

The point of departure of this section is the expression (6) for μ_m. The observation underlying the following lemma is fundamental for the derivation of upper bounds in this section. The corresponding result for binary systems was utilized in [11].

Lemma 1 *For each $m = 1, 2, \ldots$, the mth moment μ_m can be represented as the functioning probability of an (artificial) system with state vector*

$$\boldsymbol{y} = (\boldsymbol{x}^{(1)}, \ \ldots, \ \boldsymbol{x}^{(m)}) \in \mathcal{S}^m, \tag{9}$$

structure function

$$\Psi(\boldsymbol{y}) = \prod_{k=1}^{m} I(\boldsymbol{x}^{(k)})$$

and probability distribution

$$P_*(\boldsymbol{Y} = \boldsymbol{y}) = E\{\prod_{i=1}^{n}\prod_{k=1}^{m} P_{i,x_i^{(k)}}\} \tag{10}$$

Thus

$$\mu_m = \sum_{\boldsymbol{y}} \Psi(\boldsymbol{y}) P_*(\boldsymbol{Y} = \boldsymbol{y}) \equiv E_* \Psi(\boldsymbol{Y}) \tag{11}$$

Proof: This is immediate from (6).

Note that we use $\mathrm{P}_*, \mathrm{E}_*$ to denote probabilities and expectations with respect to the artificial system, while P, E correspond to the original distribution of $\boldsymbol{P}$.

It is seen that $\Psi(\boldsymbol{y})$ is the structure function (0-1 valued) of a series system of m modules each with structure function $\boldsymbol{I}$.

Assume now that the rows $\boldsymbol{P}_i$ of $\boldsymbol{P}$ are stochastically P-independent. With respect to the artificial system it is then seen that the n vectors $\boldsymbol{Y}_i \equiv (X_i^{(1)}, \ \ldots, \ X_i^{(m)})$, $i = 1, \ \ldots, \ n$, are stochastically P_*-independent, with probability functions

$$\mathrm{P}_*(\boldsymbol{Y}_i = \boldsymbol{y}_i) = \mathrm{E}[\prod_{k=1}^{m} P_{i,x_i^{(k)}}]$$

but that the components within each $\boldsymbol{Y}_i$ are dependent. Note that we now represent $\boldsymbol{y}$ as $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ where $\boldsymbol{y}_i = (x_i^{(1)}, \ldots, x_i^{(m)})$ (compare this to (9)).

The idea of our upper bound is to construct a set of P_*-independent random vectors

$$\boldsymbol{Y}_i^+ = (X_i^{+(1)}, \ \ldots, \ X_i^{+(m)}), \ ; \ i = 1, \ \ldots, \ n$$

which are P_*-stochastically larger than the $\boldsymbol{Y}_i$, i.e.

$$\boldsymbol{Y}_i \overset{\mathrm{st}}{\leq} \boldsymbol{Y}_i^+ \tag{12}$$

such that for each fixed $i = 1, \ \ldots, \ n$,

$$X_i^{+(1)}, \ \ldots, \ X_i^{+(m)}$$

are independent and identically distributed with, say,

$$\mathrm{P}_*(X_i^{+(k)} = j) = \alpha_{ij} \ ; \ j = 0, \ \ldots, \ M_i, \ 1 \leq k \leq m \tag{13}$$

(Recall that a random vector $\boldsymbol{V}$ in R^m is *stochastically larger* than a random vector $\boldsymbol{V}$ in R^m, written $\boldsymbol{U} \overset{\text{st}}{\leq} \boldsymbol{V}$, if $Eg(\boldsymbol{U}) \leq Eg(\boldsymbol{V})$ for all increasing functions g on R^m).

We can now state the following key result:

Theorem 1 *Suppose the $\boldsymbol{P}_i$ comprising $\boldsymbol{P}$ are stochastically independent. Let*

$$\boldsymbol{A} = (\alpha_{ij} \ : \ i = 1, \ \ldots, \ n; \ j = 0, \ \ldots, \ M_i)$$

be given by the preceding construction, i.e. such that (12) and (13) hold for $i = 1, \ldots, n$. Then

$$\mu_m \leq h^m(\boldsymbol{A}) \tag{14}$$

Proof: By lemma 1 and the fact that Ψ is increasing,

$$\begin{aligned}
\mu_m &= \mathrm{E}_*\Psi(\boldsymbol{Y}) \\
&\leq \mathrm{E}_*\Psi(\boldsymbol{Y}^+) \\
&= \prod_{k=1}^{m} \mathrm{E}_* I(\boldsymbol{X}^{+(k)}) \\
&= \prod_{k=1}^{m} h(\boldsymbol{A}) \\
&= h^m(\boldsymbol{A})
\end{aligned}$$

The question in practice is how to construct the X_i^+-variables. Of course, $X_i^+ = M_i$ a.s. for all i, k would do, but would only give the trivial upper bound 1 for the μ_m. The goal must be to choose the X_i^+ as small as possible, in a stochastic sense. This problem may for a fixed component with $M_i = M$ and a fixed moment m be formalized as follows:

Let $\boldsymbol{Z} = (Z_1, \ \ldots, \ Z_m)$ be random variables with values in $\{0, 1, \ \ldots, \ M\}$. Suppose there is a random vector $\boldsymbol{T} = (T_0, \ \ldots, \ T_M)$ with $T_j \geq 0$ a.s. for all j and $\sum_{j=0}^{M} T_j = 1$ a.s., such that $Z_1, \ \ldots, \ Z_m$ given $\boldsymbol{T}$ are conditionally independent and identically distributed with

$$P(Z_k = j | \boldsymbol{T}) = T_j \ ; \ j = 0, \ \ldots, \ M$$

We want to construct a random vector

$$\boldsymbol{W} = (W_1, \ \ldots, \ W_m)$$

of i.i.d. components W_k with values in $\{0, 1, \ldots, M\}$ and

$$P(W_k = j) = \alpha_j, \ \ \alpha_j \geq 0, \ \ \sum_{j=0}^{M} \alpha_j = 1$$

such that

$$\boldsymbol{Z} \overset{\text{st}}{\leq} \boldsymbol{W} \tag{15}$$

In practice, $\boldsymbol{Z}$ corresponds to $\boldsymbol{Y}_i$, $\boldsymbol{W}$ corresponds to $\boldsymbol{Y}_i^+$, $\boldsymbol{T}$ corresponds to $\boldsymbol{P}_i$ while α_j corresponds to α_{ij} in (13). The array $\boldsymbol{A}$ in theorem 1 can in this way be constructed row by row by letting i run through $\{1, \ldots, n\}$.

A set of necessary conditions on the α_j for (15) to hold is given by the following:

Theorem 2 *Let $\boldsymbol{Z}$ and $\boldsymbol{W}$ have the same meaning as above. For (15) to hold it is necessary that*

$$(\sum_{j=l}^{M} \alpha_j)^m \geq E(\sum_{j=l}^{M} T_j)^m \tag{16}$$

for $l = 1, \ \ldots, \ M$.

Proof: If (15) holds, then for each $l = 1, \ \ldots, \ M$ we have

$$P(\cap_{k=1}^{m}(W_k \geq l)) \geq P(\cap_{k=1}^{m}(Z_k \geq l))$$

which is the same as (16).

The necessary condition of theorem 2 becomes particularly simple for binary components (i.e. $M = 1$). As shown in [11], the condition (16) is then also sufficient. The following example shows, however, that the conditions (16) are *not* in general sufficient for $\boldsymbol{Z} \overset{\text{st}}{\leq} \boldsymbol{W}$ when $M > 1$.

Example

Let $M = 2$ and $m = 2$. Further let U be uniformly distributed on the interval $[0, 1]$ and define

$$T_0 = U/2, \quad T_1 = 1 - U, \quad T_2 = U/2$$

If the conditions (16) were sufficient, then we could have chosen α_1, α_2 such that

$$\alpha_2 = \sqrt{\mathrm{E}T_2^2} = \sqrt{1/12}, \quad \alpha_1 + \alpha_2 = \sqrt{\mathrm{E}(T_1 + T_2)^2} = \sqrt{7/12}$$

Consider, however, the increasing set

$$A = \{\boldsymbol{z} = (z_1, z_2) : \boldsymbol{z} = (1, 1) \text{ or } z_1 = 2 \text{ or } z_2 = 2\}$$

Then

$$\begin{aligned} P(\boldsymbol{Z} \in A) &= \mathrm{E}[T_1^2 + 2T_2 - T_2^2] \\ &= \mathrm{E}[1 - U + \frac{3}{4}U^2] = 0.75 \end{aligned}$$

while

$$P(\boldsymbol{W} \in A) = \alpha_1^2 + 2\alpha_2 - \alpha_2^2 = 0.7197$$

Thus $\boldsymbol{W}$ is *not* stochastically larger than $\boldsymbol{Z}$, although the conditions (16) are satisfied.

Before we go on to study in detail how to construct the array $\boldsymbol{A}$ in the case $M = 2, m = 2$, it is instructive to note that we can obtain equality in (14) when Φ is a kind of generalized series system (Corollary 1). More generally we can formulate the following necessary and sufficient condition for equality:

Theorem 3 *Suppose the $\boldsymbol{P}_i$ comprising $\boldsymbol{P}$ are stochastically independent, and that $\boldsymbol{P}$ with probability 1 gives positive mass to each possible state in $\mathcal{S}$.*

Let $\mathcal{C}$ be a given subset of $\mathcal{S}$ and for the ith component ($i = 1, \ldots, n$), let

$$D_i = \{d \in \mathcal{S}_i : d \textit{ is the ith coordinate of some minimal path vector to } \mathcal{C}\}$$

Then equality holds in (14) of theorem 1 if and only if the following conditions are both satisfied for each component $i = 1, \ldots, n$.

(i) *For each $d \in D_i$ we have*

$$\left(\sum_{j=d}^{M_i} \alpha_{ij}\right)^m = E\left(\sum_{j=d}^{M_i} P_{ij}\right)^m \equiv E(R_i(d))^m$$

(ii) *There exists a random variable U_i and constants $\{\gamma_{id} : d \in D_i\}$ such that*

$$R_i(d) = \gamma_{id} U_i \ ; \ d \in D_i$$

Proof: To prove the "if" part, suppose conditions (i) and (ii) hold for all components i. From (7) and the inclusion-exclusion formula for the probability of a union of events it follows that for given $\boldsymbol{P}$, $h(\boldsymbol{P})$ is a linear combination of terms of the form

$$P(X_1 \geq d_1, X_2 \geq d_2, \ldots, X_n \geq d_n)$$

where $d_i \in D_i$ for all i. From condition (ii) thus follows that, for some constant c,

$$h(\boldsymbol{P}) = c \prod_{i=1}^{n} U_i$$

Now the independence of the $\boldsymbol{P}_i$ implies the independence of the U_i, so

$$\mu_m = c^m \prod_{i=1}^{n} \mathrm{E}(U_i^m) \tag{17}$$

From condition (i) we readily get

$$\sum_{j=d}^{M_i} \alpha_{ij} = \gamma_{id} \left(\mathrm{E}(U_i^m)\right)^{1/m}$$

Thus condition (ii) holds also for the rows of the nonrandom $\boldsymbol{A}$, with U_i replaced by the constant $(\mathrm{E}(U_i^m))^{1/m}$. Therefore $h^m(\boldsymbol{A})$ equals (17) so equality holds in (14).

To prove "only if", suppose equality holds in (14) for a given value m. Fix a component, which without loss of generality can be taken to be component 1. Suppose D_1 has r elements $a_1, \ldots, a_r$ corresponding to the r different minimal path vectors to $\mathcal{C}$ given by

$$(a_l, b_{l2}, \ldots, b_{ln}) \ ; \ l = 1, \ldots, r$$

The assumption of equality in (14) implies that

$$\mathrm{E}_*\Psi(\boldsymbol{Y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n) = \mathrm{E}_*\Psi(\boldsymbol{Y}_1^+, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n) \tag{18}$$

for each fixed set of values for $\boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$. Suppose now for simplicity that $r = m$. (This is unreasonable, but is done for simplifying the exposition. We shall later see how we can avoid the assumption). Then for $i = 2, \ldots, n$ we can take

$$\boldsymbol{y}_i = (b_{1i}, b_{2i}, \ldots, b_{mi}) \tag{19}$$

The clue of our proof is that for this choice of $\boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$ we have $\Psi(\boldsymbol{v}, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n) = 1$ if and only if $v_k \geq a_k$ for $k = 1, \ldots, m$. Thus

$$\begin{aligned}
\mathrm{E}_*\Psi(\boldsymbol{Y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n) &= \mathrm{P}_*(X_1^{(1)} \geq a_1, \cdots, X_1^{(m)} \geq a_m) \\
&= \mathrm{E}[\prod_{k=1}^{m} R_1(a_k)] \\
&\leq \text{(Hölder's inequality)} \prod_{k=1}^{m} [\mathrm{E}(R_1(a_k)^m)]^{1/m} \\
&\leq \text{(Theorem 2)} \prod_{k=1}^{m} \left[\left(\sum_{j=a_k}^{M_1} \alpha_{1j} \right)^m \right]^{1/m} \\
&= \prod_{k=1}^{m} \left(\sum_{j=a_k}^{M_1} \alpha_{1j} \right) = \mathrm{E}_*\Psi(\boldsymbol{Y}_1^+, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n)
\end{aligned}$$

It follows from the last inequality above that equality in (18) implies condition (i). Moreover, the condition for equality in Hölder's inequality is that the

$$R_1(a_k)$$

are proportional (with probability 1). This is exactly condition (ii) of the theorem. Thus we are done if $r = m$.

If $r < m$, then in (19) we can fill the entries $r + 1, \ldots, m$ consistently by repeating some of the r given minimal path vectors. If $r > m$, then we can use any set of m minimal path vectors to construct (19) and every time we will get the proportionality condition of (ii) for a subset of D_1. Since any subset can be used we may conclude that condition (ii) holds for the entire D_1.

Remark: The conditions (i) and (ii) are satisfied for any component i for which $\boldsymbol{P}_i$ is nonrandom. Condition (ii) is, furthermore, trivially satisfied if D_i has cardinality 1. This leads to the following corollary:

Corollary 1 *Let the assumptions be as in theorem 3, but suppose $\mathcal{C}$ has a single minimal path vector $(a_1, \ldots, a_n)$ (we may call the corresponding system a* generalized series system, *since it works provided each component exceeds a certain level). If the α_{ij} can be chosen to satisfy (i) of theorem 3 with $d = a_i$, then equality holds in (14).*

4 Upper bound for μ_m when $M = 2, m = 2$

We study in detail this case which is interesting for applications, since it will enable us to compute upper bounds for the second moment and hence the standard deviation of $h(\boldsymbol{P})$ for systems with 3-state components.

Let $\boldsymbol{Z}$ and $\boldsymbol{W}$ be as described in section 3. We now consider vectors $\boldsymbol{Z} = (Z_1, Z_2)$ with $Z_k \in \{0, 1, 2\}$ (and similarly for $\boldsymbol{W}$). Since the distributions of $\boldsymbol{Z}$ and $\boldsymbol{W}$ are permutation symmetric by our construction (i.e. (Z_1, Z_2) and (Z_2, Z_1) have the same distribution, and similarly for $\boldsymbol{W}$), it turns out that the study of the relation

$$\boldsymbol{Z} \overset{\text{st}}{\leq} \boldsymbol{W}$$

can be reduced as follows.

Define first, for $j = 1, 2$ and $\boldsymbol{z} \in \{0, 1, 2\}^2$,

$$g_j(\boldsymbol{z}) = g_j(z_1, z_2) = \text{ number of } z_i \text{ that are } \geq j$$

and let

$$g(\boldsymbol{z}) = (g_1(\boldsymbol{z}), g_2(\boldsymbol{z})) \tag{20}$$

(For example, $g(0, 1) = (1, 0)$ and $g(2, 0) = (1, 1)$).

We claim that

$$\boldsymbol{Z} \overset{\text{st}}{\leq} \boldsymbol{W} \text{ if and only if } g(\boldsymbol{Z}) \overset{\text{st}}{\leq} g(\boldsymbol{W})$$

To see this, note that "only if" is clear since g is increasing in each variable. To prove the "if"-part, observe first that since $\boldsymbol{Z}$ and $\boldsymbol{W}$ are permutation symmetric, we have $\boldsymbol{Z} \overset{\text{st}}{\leq} \boldsymbol{W}$ if and only if $\mathrm{E}\Psi(Z_1, Z_2) \leq \mathrm{E}\Psi(W_1, W_2)$ for all increasing functions Ψ which are symmetric in the two arguments (see Scarsini and Shaked [17]). But such functions can be written as increasing functions of g, and the "if"-part of the statement follows.

Note that while $\boldsymbol{z}$ itself has $3^2 = 9$ different values, $g(\boldsymbol{z})$ has only 6 possible values

$$K = \{(2, 2), (2, 1), (1, 1), (2, 0), (1, 0), (0, 0)\}$$

For example, $g(\boldsymbol{Z}) = (1, 1)$ corresponds to the two vectors $\boldsymbol{Z} = (2, 0), (0, 2)$.

The condition $g(\boldsymbol{Z}) \overset{\text{st}}{\leq} g(\boldsymbol{W})$ is equivalent to having

$$P(g(\boldsymbol{Z}) \in A) \leq P(g(\boldsymbol{W}) \in A) \tag{21}$$

for all increasing subsets A of K, $\emptyset \neq A \neq K$. These sets are

$$\begin{aligned}
A_1 &= \{(2, 2)\} \\
A_2 &= \{(2, 2), (2, 1)\} \\
A_3 &= \{(2, 2), (2, 1), (1, 1)\} \\
A_4 &= \{(2, 2), (2, 1), (2, 0)\} \\
A_5 &= \{(2, 2), (2, 1), (1, 1), (2, 0)\} \\
A_6 &= \{(2, 2), (2, 1), (1, 1), (2, 0), (1, 0)\}
\end{aligned}$$

Assume now that α_1, α_2 satisfy the necessary conditions (16), which in this case means

$$\begin{aligned} \alpha_2^2 &\geq \mathrm{E}T_2^2 \\ (\alpha_1+\alpha_2)^2 &\geq \mathrm{E}(T_1+T_2)^2 \end{aligned}$$

We first show that these conditions imply (21) for all sets A_j above, except A_5. Then we shall find an additional condition to assure that (21) holds for A_5. We shall also show that the additional condition is unnecessary in case $\boldsymbol{T} = (T_0, T_1, T_2)$ is Dirichlet-distributed.

Note first that (21) for $A = A_1$ and $A = A_4$ correspond to the necessary conditions, and are thus satisfied by assumption. For the remaining sets A_j we get,

A_2: Here

$$\begin{aligned} P(g(\boldsymbol{Z}) \in A_2) &= \mathrm{E}[T_2^2 + 2T_1T_2] \\ &= \mathrm{E}[2(T_1+T_2)T_2 - T_2^2] \\ &\leq 2(\mathrm{E}(T_1+T_2)^2)^{1/2}(\mathrm{E}T_2^2)^{1/2} - \mathrm{E}(T_2^2) \\ &\leq 2(\alpha_1+\alpha_2)\alpha_2 - \alpha_2^2 \\ &= P(g(\boldsymbol{W}) \in A_2) \end{aligned}$$

The two inequalities above follow from, respectively, Cauchy-Schwarz inequality and the fact that $2a^{1/2}b^{1/2} - a$ is increasing in both a and b (for $0 \leq a \leq b \leq 1$).

A_3: Here

$$\begin{aligned} P(g(\boldsymbol{z}) \in A_3) &= \mathrm{E}[2T_2 - T_2^2] \\ &\leq 2(\mathrm{E}(T_2)^2)^{1/2} - \mathrm{E}(T_2^2) \\ &\leq 2\alpha_2 - \alpha_2^2 \\ &= P(g(\boldsymbol{W}) \in A_3) \end{aligned}$$

since $2a^{1/2} - a$ is increasing in a for $0 \leq a \leq 1$

A_6: In the same manner,

$$\begin{aligned} P(g(\boldsymbol{Z}) \in A_6) &= P(g(\boldsymbol{Z}) \neq (0,0)) \\ &= 2\mathrm{E}(T_1+T_2) - \mathrm{E}(T_1+T_2)^2 \\ &\leq 2(\mathrm{E}(T_1+T_2)^2)^{1/2} - \mathrm{E}(T_1+T_2)^2 \\ &\leq 2(\alpha_1+\alpha_2) - (\alpha_1+\alpha_2)^2 \\ &= P(g(\boldsymbol{W}) \in A_6) \end{aligned}$$

It remains to consider A_5. We have

$$P(g(\boldsymbol{Z}) \in A_5) = \mathrm{E}[T_1^2 + 2T_2 - T_2^2]$$

so we need to have

$$\mathrm{E}[T_1^2 + 2T_2 - T_2^2] \leq \alpha_1^2 + 2\alpha_2 - \alpha_2^2$$

As we saw in the example following theorem 2, this condition is not implied by the sufficient conditions. Thus we state the following result, which is best possible in the sense that none of the inequalities can be removed in general:

Theorem 4 *Let $M = m = 2$. Then $\boldsymbol{Z} \overset{st}{\leq} \boldsymbol{W}$ if and only if the following inequalities are all satisfied:*

$$
\begin{aligned}
E[T_2^2] &\leq \alpha_2^2 && (22)\\
E[(T_1+T_2)^2] &\leq (\alpha_1+\alpha_2)^2 && (23)\\
E[T_1^2+2T_2-T_2^2] &\leq \alpha_1^2+2\alpha_2-\alpha_2^2 && (24)
\end{aligned}
$$

If α_2^2 is chosen to obtain equality in (22), we can replace (24) by a simpler condition:

Corollary 2 *Let $M = m = 2$. A sufficient condition for $\boldsymbol{Z} \overset{st}{\leq} \boldsymbol{W}$ is that*

$$
\begin{aligned}
E[T_2^2] &= \alpha_2^2 && (25)\\
E[(T_1+T_2)^2] &\leq (\alpha_1+\alpha_2)^2 && (26)\\
E[T_2-T_1T_2] &\leq \alpha_2-\alpha_1\alpha_2 && (27)
\end{aligned}
$$

Proof: It is enough to check that (24) is implied by (25-27). But the left hand side of (24) is

$$\mathrm{E}[(T_1+T_2)^2+2(T_2-T_1T_2)-2T_2^2]$$

which by the assumptions is

$$\leq (\alpha_1+\alpha_2)^2+2(\alpha_2-\alpha_1\alpha_2)-2\alpha_2^2=\alpha_1^2+2\alpha_2-\alpha_2^2$$

which is the right hand side of (24).

It is interesting to note that condition (24) in theorem 4 is unnecessary when $\boldsymbol{T}$ is Dirichlet-distributed (see definition in section 1).

Theorem 5 *Let $M = m = 2$ and assume that $\boldsymbol{T} = (T_0, T_1, T_2)$ is Dirichlet-distributed. Then $\boldsymbol{Z} \overset{st}{\leq} \boldsymbol{W}$ if (22) and (23) are both satisfied.*

Thus in this case the conditions (16) are also sufficient.

Proof: Suppose $\boldsymbol{T}$ is Dirichlet-distributed with parameters $(\theta_0, \theta_1, \theta_2)$. Let α_1, α_2 be such that (22) and (23) hold. Further, let $\tilde{\alpha}_1, \tilde{\alpha}_2$ satisfy $\tilde{\alpha}_2^2 = \mathrm{E}(T_2^2)$ and $(\tilde{\alpha}_1+\tilde{\alpha}_2)^2 = \mathrm{E}[(T_1+T_2)^2]$. Then if $\tilde{\boldsymbol{W}}$ is the "$\boldsymbol{W}$" corresponding to the $\tilde{\alpha}_1, \tilde{\alpha}_2$, we have $\tilde{\boldsymbol{W}} \overset{\mathrm{st}}{\leq} \boldsymbol{W}$ and hence we are done if we can prove that $\boldsymbol{Z} \overset{\mathrm{st}}{\leq} \tilde{\boldsymbol{W}}$. In essence this means that, by Corollary 2, we are done if we prove that α_1, α_2 chosen by equality in (22-23), satisfies (27).

From the moment-formulas of the Dirichlet distribution (see e.g. DeGroot [6, p. 49]) we get

$$\mathrm{E}T_2-\mathrm{E}T_1T_2=\frac{\theta_2(\theta_0+\theta_2+1)}{(\theta_0+\theta_1+\theta_2)(\theta_0+\theta_1+\theta_2+1)}$$

so that we need to show that

$$\frac{\theta_2(\theta_0+\theta_2+1)}{(\theta_0+\theta_1+\theta_2)(\theta_0+\theta_1+\theta_2+1)} \leq$$

$$\frac{\sqrt{\theta_2(\theta_2+1)}\left(\sqrt{(\theta_0+\theta_1+\theta_2)(\theta_0+\theta_1+\theta_2+1)}-\sqrt{(\theta_1+\theta_2)(\theta_1+\theta_2+1)}\right)+\theta_2(\theta_2+1)}{(\theta_0+\theta_1+\theta_2)(\theta_0+\theta_1+\theta_2+1)}$$

or, equivalently,

$$\sqrt{(\theta_0+\theta_1+\theta_2)(\theta_0+\theta_1+\theta_2+1)}-\sqrt{(\theta_1+\theta_2)(\theta_1+\theta_2+1)} \geq \theta_0\sqrt{\frac{\theta_2}{\theta_2+1}} \tag{28}$$

Now (28) follows from the next lemma by choosing $x = \theta_1+\theta_2$ and $\delta = \theta_0$. Thus we are done.

Lemma 2 *Let* $f(x) = \sqrt{x(x+1)}$, $x \geq 0$. *Then* $f(x+\delta) - f(x) \geq \delta$ *for all* $\delta > 0$.

Proof: $f'(x) = \frac{2x+1}{2\sqrt{x(x+1)}}$. Hence $f'(x) \geq 1$ for all $x \geq 0$. Further,

$$f(x+\delta) - f(x) = \int_x^{x+\delta} f'(u)\,du \geq \int_x^{x+\delta} du = \delta$$

We summarize the final findings of this section in a theorem, where the condition for equality follows from theorem 3:

Theorem 6 *Let* $M_i = 2$ *for* $i = 1, \ldots, n$ *and let* $\mathcal{C}$ *be an increasing set in* $\mathcal{S}$. *Further let* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_n$ *be independent with nondegenerate Dirichlet distributions. Then for* $h(\boldsymbol{P}) = P(\boldsymbol{X} \in \mathcal{C})$ *we have*

$$Eh^2(\boldsymbol{P}) \leq h^2(\boldsymbol{A}) \text{ where } \boldsymbol{A} = \begin{bmatrix} \sqrt{E\{[P_{11}+P_{12}]^2\}} - \sqrt{E\{[P_{12}]^2\}} & \sqrt{E\{[P_{12}]^2\}} \\ \sqrt{E\{[P_{21}+P_{22}]^2\}} - \sqrt{E\{[P_{22}]^2\}} & \sqrt{E\{[P_{22}]^2\}} \\ \vdots & \vdots \\ \sqrt{E\{[P_{n1}+P_{n2}]^2\}} - \sqrt{E\{[P_{n2}]^2\}} & \sqrt{E\{[P_{n2}]^2\}} \end{bmatrix}$$

with equality if and only if the system is a generalized series system, i.e. has a single minimal path vector.

Arbitrary values for M and m

We considered above the special situation $M = m = 2$, leading to the general result theorem 4 and the special case theorem 6. The same line of action can, at least in principle, be used for arbitrary choices of m and M. Start by defining a function $g(\cdot) : \{0, \ldots, M\}^m \to \{0, \ldots, m\}^M$ similar to (20). A result like (21) will still apply, but the number of increasing subsets A will grow rapidly as M and m are increased. The increasing sets should then be checked one by one, and the corresponding restriction on $\boldsymbol{\alpha}$ noted. Finally, a set of inequalities like (22)-(24) will result.

It may seem reasonable to *guess* that if $\boldsymbol{T}$ is Dirichlet distributed, then theorem 6 can be generalized to arbitrary values for M and m. We state the following conjecture which if true would lead to a straightforward extension of theorem 6.

Conjecture 1 *Let $\boldsymbol{T}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ have the same meaning as in section 3, but assume now that $\boldsymbol{T}$ is Dirichlet-distributed. Then $\boldsymbol{Z} \overset{st}{\leq} \boldsymbol{W}$ if and only if*

$$(\sum_{j=l}^{M} \alpha_j)^m \geq E(\sum_{j=l}^{M} T_j)^m \ \ \textit{for}\ l = 1,\ \ldots,\ M$$

5 Bounds based on minimal path or -cut vectors

This section deals with generalization of Lindqvist [12] and Natvig & Eide [16] to a multi-state setting. The concept of *associated random variables* will play an important role in the following.

Recall (see e.g. [5], [10]) that a set $\boldsymbol{Y} = (Y_1,\ \ldots,\ Y_r)$ of random variables is called *associated* if

$$\mathrm{E}[f(\boldsymbol{Y})g(\boldsymbol{Y})] \geq \mathrm{E}f(\boldsymbol{Y})\ \mathrm{E}g(\boldsymbol{Y})$$

for all bounded functions $f, g : R^r \to R$ which are increasing in each argument. It is well known (see references above for proofs) that any set of independent random variables is associated, and that if $\boldsymbol{Y}$ is associated, then any set of all non-decreasing or all non-increasing functions $f_1(\boldsymbol{Y}),\ \ldots,\ f_s(\boldsymbol{Y})$ is associated. These properties are used repeatedly below without necessarily explicit mention.

It follows easily from the definition that if $\boldsymbol{Y}$ is associated and $Y_j \geq 0$ for all j, then

$$\mathrm{E}[\prod_{j=1}^{r} Y_r] \geq \prod_{j=1}^{r} \mathrm{E}Y_r \tag{29}$$

Defining $\coprod_{j=1}^{r} a_j \equiv 1 - \prod_{j=1}^{r}(1 - a_j)$, it follows from (29) that if $\boldsymbol{Y}$ is associated and $Y_j \leq 1$ a.s. for all j, then

$$\mathrm{E}[\coprod_{j=1}^{r} Y_r] \leq \coprod_{j=1}^{r} \mathrm{E}Y_r \tag{30}$$

More generally, if $\boldsymbol{Y}$ is associated with $0 \leq Y_j \leq 1$ for all j, then [11, Lemma 1]

$$E\left[\left(\coprod_{r=1}^{n} Y_r\right)^m\right] \leq \coprod_{1 \leq r_k \leq n, k=1,\ldots,m} E\prod_{\ell=1}^{m} Y_{r_\ell} \tag{31}$$

where the rightmost $\coprod$ is taken over all vectors $(r_1,\ \ldots,\ r_m)$ with each $r_k \in \{1, 2,\ \ldots,\ n\}$.

This is in turn a special case of the following result, which is proved in essentially the same way as (31), and which is needed in the final section.

Lemma 3 *Suppose the set $(Y_{r,k} : k = 1,\ \ldots,\ m;\ r = 1,\ \ldots,\ n_k)$ of random variables with values in $[0, 1]$ is associated. Then*

$$E\left[\prod_{k=1}^{m} \coprod_{r=1}^{n_k} Y_{r,k}\right] \leq \coprod_{1 \leq r_k \leq n_k;\ k=1,\ \ldots,\ m} E\prod_{\ell=1}^{m} Y_{r_\ell,\ell} \tag{32}$$

We shall, furthermore, need the following result, which may be of independent interest.

Lemma 4 *Let $\boldsymbol{Y} = (Y_0, \ldots, Y_M)$ be Dirichlet-distributed as defined in (3), and put $X_k = \sum_{j=0}^{k} Y_j$ for $k = 0, \ \ldots, \ M-1$. Then*

$$X_0, X_1, \ \ldots, \ X_{M-1}$$

are associated random variables.

Remark: Note that the definition of the Dirichlet-distribution is symmetric in $Y_0, \ \ldots, \ Y_M$. Hence the lemma implies that any collection of nested sums of the Y_js is associated.

Proof: Define

$$Z_j = \frac{Y_j}{1 - X_j} \text{ for } j = 0, \ \ldots, \ M-1$$

It is shown in [2, p. 61] that the Z_j are independent. We shall show by induction that

$$X_k = \coprod_{j=0}^{k} W_j \tag{33}$$

for $k = 0, 1, , \ \ldots, \ M-1$, where $W_j = \frac{Z_j}{1+Z_j}$.

We have $W_1 = X_1$, which proves the case $k = 0$. Suppose next that (33) holds for $k-1$. Then

$$\begin{aligned} X_k &= X_{k-1} + Y_k \\ &= X_{k-1} + W_k(1 - X_{k-1}) \\ &= W_k \coprod X_{k-1} = \textstyle\coprod_{j=0}^{k} W_j \end{aligned}$$

and (33) is proven.

The W_j are independent, because the Z_j are, and hence associated. Hence by (33), the X_j are associated.

The next lemma is fundamental for the derivation of bounds for μ_m in this section. The inequalities are straightforward extensions of similar results for binary systems [5] and are also given in [8].

Lemma 5 *With assumptions as in section 1, and with minimal path and -cut vectors given as in section 2, we have for given $\boldsymbol{P}$,*

$$\prod_{r=1}^{l} P(\boldsymbol{X} \not\leq \boldsymbol{z}^{(r)}) \leq h(\boldsymbol{P}) \leq \coprod_{r=1}^{k} P(\boldsymbol{X} \geq \boldsymbol{y}^{(r)})$$

Proof: By (8), the independence of $X_1, \ldots, X_n$ given $\boldsymbol{P}$, and the fact that the sets $\{\boldsymbol{x} : \boldsymbol{x} \not\leq \boldsymbol{z}^{(l)}\}$ are increasing sets in $\mathcal{S}$, we have by (29),

$$\begin{aligned} h(\boldsymbol{P}) &= P(\cap_{r=1}^{l} (\boldsymbol{X} \not\leq \boldsymbol{z}^{(r)})) \\ &\geq \prod_{r=1}^{l} P(\boldsymbol{X} \not\leq \boldsymbol{z}^{(r)})) \end{aligned}$$

The same line of proof gives the upper bound.

Assume now that $\boldsymbol{P}$ is random. Then we can define random variables

$$\begin{aligned} U_r &= P(\boldsymbol{X} \geq \boldsymbol{y}^{(r)}) = \prod_{i=1}^{n} R_i(y_i^{(r)}) \; ; r = 1, \ldots, k \\ V_s &= P(\boldsymbol{X} \not\leq \boldsymbol{z}^{(s)}) = 1 - \prod_{i=1}^{n} F_i(z_i^{(s)}) \; ; s = 1, \ldots, l \end{aligned}$$

for which

$$\prod_{s=1}^{l} V_s \leq h(\boldsymbol{P}) \leq \coprod_{r=1}^{k} U_r$$

Taking the mth power throughout the inequality and then taking expectations, we get the following result.

Theorem 7 *Suppose that each of the two sets $\{U_1, \ldots, U_k\}$ and $\{V_1, \ldots, V_l\}$ are associated. Then*

$$\prod_{s=1}^{l} E[V_s^m] \leq \mu_m \leq \coprod_{1 \leq r_1, \ldots, r_m \leq k} E \prod_{j=1}^{m} U_{r_j}$$

A sufficient condition for association of the two sets is that the set $\{R_i(j) : i = 1, \ldots, n; \; j = 1, \ldots, M_i\}$ is associated. In particular this holds if the vectors $\boldsymbol{P}_i$ are stochastically independent with each $\boldsymbol{P}_i$ being Dirichlet-distributed.

Proof: The first part follows from (29) and (31). The sufficient condition follows from the fact that both the U_r and the V_s are (by their definition) increasing functions of the $R_i(j)$. Finally, lemma 4 implies the result on the Dirichlet case.

6 Numerical example

We now turn to an example of a multistate system considered by Natvig et al. [15], concerning an electric system on an oil rig. Natvig et al. [15] studies the time-dependent behavior of the system, computing bounds for the system availability in fixed time intervals using a Markovian approach. Thus our application here will be quite different.

The system consists of two generators and a control unit. Each of the generators can (if working properly) supply the rig with a sufficient amount of power, and hence only one generator is normally working while the other is in cold standby. If the operating generator experiences trouble, then the control unit is responsible for starting the spare generator.

Each of the two generators can be in any of three states, which we name $\{0, 1, 2\}$. The control unit can also be in any of three states, which are 2: Operating properly, 1: Fails to start the spare generator, 0: Halts the operating generator without starting the spare.

We let X_1 represent the control unit, while X_2 and X_3 are the states of generator 1 and 2, respectively.

Assume now that the rig gets sufficient power as long as at least one of the generators are supplying power at level 1, i.e., if and only if $\boldsymbol{X} \in \mathcal{C}$, where

$$\mathcal{C} = \{\boldsymbol{X} : (X_1 \geq 1 \cap X_2 \geq 1) \cup (X_1 = 2 \cap X_3 \geq 1)\}$$

The minimal path and -cut vectors of this increasing set are:

Min path vectors		
X_1	X_2	X_3
1	1	0
2	0	1

Min cut vectors		
X_1	X_2	X_3
2	0	0
1	0	2
0	2	2

Somewhat arbitrarily we model our uncertainties about the probability distributions of each of X_1, X_2, X_3 by independent Dirichlet-distributions with parameters given by the following table.

Parameters in the distribution				
i	θ_{i0}	θ_{i1}	θ_{i2}	$\sum_{j=0}^{2} \theta_{ij}$
1	2	7	3	12
2	1	16	3	20
3	1	16	3	20

With these parameter choices, we find $Eh(\boldsymbol{P}) = h(\mathrm{E}\boldsymbol{P}) = 0.8035$, which serves as our estimate for $P(\boldsymbol{X} \in \mathcal{C})$. However, suppose we are also interested in quantifying the uncertainty of this estimate, represented by, e.g., $SD(h(\boldsymbol{P}))$. Thus we calculate bounds for $Eh^2(\boldsymbol{P})$ and utilize these to obtain bounds for the standard deviation:

Upper bounds		
Inequality	Theorem 6	Theorem 7
Result	0.65871	0.77950

Lower bounds		
Inequality	Theorem 7	Trivial
Result	0.65092	0.64568

(The "trivial" bound is simply $(\mathrm{E}h(\boldsymbol{P}))^2$). The results are very informative, in fact we find $0.651 \leq Eh^2(\boldsymbol{P}) \leq 0.659$, or equivalently, $0.0724 \leq SD[h(\boldsymbol{P})] \leq 0.114$.

7 Extensions

New bounds obtained by duality

From the example in section 6 we see that the upper bound of μ_2 given by theorem 7 is too large to be of interest. Motivated by results for binary systems ([11],[12],[13]) we try in this section to apply inequalities for μ_2 to the *dual* of the original system, and then to transform results back to the original one. In fact the idea will also work for higher moments than 2, but seems to be most successful for that case.

We can of course write

$$h(\boldsymbol{P}) = 1 - P(\boldsymbol{X} \in \mathcal{C}^c)$$

where $\mathcal{C}^c$, the complementary set of $\mathcal{C}$ in $\mathcal{S}$, is a *decreasing* set in $\mathcal{S}$. However, this set can be made increasing by reversing the order of the states within each component state space S_i. Thus relabel the states of the ith component so that the new state $\tilde{j}$ corresponds to $M_i - j$,$(i = 1, \ldots, n;\ j = 0, 1, \ldots, M_i)$. Denote by $\tilde{\mathcal{C}}$ the set $\mathcal{C}^c$ in this new setting. Then $\tilde{\mathcal{C}}$ is increasing. Moreover, letting $\tilde{X}_i = M_i - X_i$ and $\tilde{P}_{ij} P(\tilde{X}_i = \tilde{j}) P_{i,M_i-j}$, we get

$$\tilde{h}(\tilde{\boldsymbol{P}}) \equiv P(\tilde{\boldsymbol{X}} \in \tilde{\mathcal{C}}) = 1 - h(\boldsymbol{P}) \tag{34}$$

If $I_{\mathcal{C}}(\cdot)(I_{\tilde{\mathcal{C}}}(\cdot))$ is the indicator function of $\mathcal{C}$ $(\tilde{\mathcal{C}})$, then $I_{\tilde{\mathcal{C}}}(\cdot)$ is called the *dual* of $I_{\mathcal{C}}(\cdot)$.

Any upper or lower bound derived for the moments of $h(\boldsymbol{P})$ can now be applied to $\tilde{h}(\tilde{\boldsymbol{P}})$, and through (34) it can give rise to new bounds for the moments of $h(\boldsymbol{P})$. In particular we have for the second moment,

$$\mathrm{E}[h^2(\boldsymbol{P})] = \mathrm{E}[\tilde{h}^2(\tilde{\boldsymbol{P}})] + 2\mathrm{E}h(\boldsymbol{P}) - 1$$

from which we obtain upper (lower) bounds for $\mathrm{E}[h^2(\boldsymbol{P})]$ from upper (lower) bounds of $\mathrm{E}[\tilde{h}^2(\tilde{\boldsymbol{P}})]$ and $\mathrm{E}h(\boldsymbol{P})$ (where the latter may in applications often be computed exactly).

The following remarks are important for practical application of the duality approach.

Remark 1. There is a close connection between minimal path and cut vectors corresponding to the original set $\mathcal{C}$ and the derived set $\tilde{\mathcal{C}}$. In fact, if $\boldsymbol{a} = (a_1, \ldots, a_n)$ is a minimal path (cut) vector of $\mathcal{C}$, then $(M_1 - a_1, \ \ldots, \ M_n - a_n)$ is a minimal cut (path) vector of $\tilde{\mathcal{C}}$. As we have seen earlier, equality in theorem 1 can be obtained for generalized series system. The dual of such a system is a system with a single minimal *cut* vector. This is a system which fails whenever all components fall below prespecified levels, and may be called a *generalized parallel system.* Thus, computing bounds for μ_2 by the methods of section 3 using both the system itself and its dual, we may expect to get close to the exact value for μ_2 if the system is nearly a generalized series system or is nearly a generalized parallel system.

Remark 2. If the $(R_i(j))$ corresponding to $\boldsymbol{P}$ are associated random variables, then this also holds for $(\tilde{R}_i(j))$ corresponding to $\tilde{\boldsymbol{P}}$. This follows since decreasing functions of associated random variables are associated. Thus the inequalities of section 5 can be applied to the dual if they can be applied to the original system.

Remark 3. If the random vectors $\boldsymbol{P}_i$ are Dirichlet-distributed, then this also holds for the $\tilde{\boldsymbol{P}}_i$. This is immediate from the definition of the Dirichlet-distribution. Thus theorems 6 and 7 can be applied to the dual system if they can be applied to the original system.

Numerical example (continued): The results of the example of section 6, when using the approach of this section, were found to be:

Upper bounds		
Inequality	Theorem 6	Theorem 7
Dual result	0.66231	0.66160

Lower bounds		
Inequality	Theorem 7	Trivial
Dual result	0.63233	0.64568

It is remarkable that the upper bound of theorem 7 is now very close to the best upper bound of section 6, while the same theorem applied to the *original* system gave a much too large value. However, the lower bound from theorem 7 is now much worse than before, and is in fact useless since it falls below the trivial bound $(\mathrm{E}h(\boldsymbol{P}))^2$.

The expected system state

Suppose that $\Phi(\boldsymbol{x})$ takes values in $\{0, 1, \ \ldots, \ M\}$, in which case

$$\mathrm{E}\Phi(\boldsymbol{X}) = \sum_{j=1}^{M} P(\Phi(\boldsymbol{X}) \geq j) = \sum_{j=1}^{M} h_j \equiv \gamma(\boldsymbol{P})$$

Then for $m = 1, 2, \ldots$ we get by computing the mth power and taking expectation,

$$\mathrm{E}[\gamma^m(\boldsymbol{P})] = \sum_{1 \leq j_1,\ \ldots,\ j_m \leq M} \mathrm{E}[\prod_{r=1}^{m} h_{j_r}(\boldsymbol{P})]$$

Thus, in order to compute exactly or derive bounds the mth moment of $\gamma(\boldsymbol{P})$ we need to be able to compute or bound terms of the form

$$\mu_{j_1,\ \ldots,\ j_m} \equiv \mathrm{E}[\prod_{r=1}^{m} h_{j_r}(\boldsymbol{P})] \tag{35}$$

This problem is, however, only slightly more complicated than the problem of computing or bounding the moments $\mu_m = \mathrm{E}[h^m(\boldsymbol{P})]$ considered so far in this paper.

To be specific, consider first the formula (6) for μ_m. The corresponding formula for (35) is simply

$$\mu_{j_1,\ \ldots,\ j_m} \quad \sum_{\boldsymbol{x}^{(1)},\ \ldots,\ \boldsymbol{x}^{(m)}} \{\prod_{k=1}^{m} \Phi_{j_k}(\boldsymbol{x}^{(k)})\} \mathrm{E}\{\prod_{i=1}^{n} \prod_{k=1}^{m} P_{i,x_i}\}$$

where $\Phi_j(\cdot)$ is the indicator function of the set $\{\boldsymbol{x} : \Phi(\boldsymbol{x}) \geq j\}$.

The ideas leading to an upper bound for μ_m in section 3 are easily generalized to $\mu_{j_1,\ \ldots,\ j_m}$, simply by redefining $\Psi(\boldsymbol{y})$ to $\prod_{k=1}^{m} \Phi_{j_k}(\boldsymbol{x}^{(k)})$. Similarly, the bounds in section 5 are generalized in a rather straightforward manner, using the general inequalities (29) and (32).

References

[1] Agrawal, A.& Barlow, R.E. (1984) A survey of network reliability and domination theory. *Operations Research*, **32**, 478-492.

[2] Aitchison, J. (1986) *The statistical analysis of compositional data.* Chapman and Hall. London.

[3] Aven, T. (1985) Reliability evaluation of multistate systems with multistate components. *IEEE Trans. Reliability*, **R-34**, 473-479.

[4] Aven, T. (1987) Availability evaluation of oil/gas production and transportation systems. *Reliability Engineering*, **14**, 35-44.

[5] Barlow, R.E & Proschan, F. (1975) *Statistical Theory of Reliability and Life Testing. Probability Models.* To Begin With, Silver Spring, MD.

[6] DeGroot, M.H (1970) *Optimal statistical decisions.* McGraw-Hill Inc., N.Y.

[7] Doulliez, P. & Jamoulle, J. (1972) Transportation networks with random arc capacities. *Recherche Operationelle*, **3**, 45-60.

[8] Funnemark E. & Natvig, B. (1985) Bounds for the availabilities in a fixed time interval for multistate monotone systems. *Adv. Appl. Probab.,* **17** 638–665.

[9] Griffith, W.S. (1980) Multistate reliability models. *J. Appl. Prob.,* **17** 735–744.

[10] LINDQVIST, B.H. (1988) Association of probability measures on partially ordered spaces. *J. Multivariate Analysis,* **26**, 111–132.

[11] LINDQVIST, B.H. (1992) Moment Inequalities for the reliability function. *Statistics & Probabilty Letters,* **13**, 85–88.

[12] LINDQVIST, B.H. (1994) The minimal path upper bound for the moments of a reliability function. *Scand. J. Statist.,* **21**, 83–90.

[13] LINDQVIST, B.H. (1994) On the uncertainty of a system reliability estimate. *Reliability Engineering and System Safety,* **46** , 263–269.

[14] NATVIG, B. (1980) Two suggestions on how to define a multistate coherent system. *Adv. Appl. Probab,* **14** 434–455.

[15] NATVIG, B., SØRMO, S., HOLEN, A.T. & HØGÅSEN, G. (1986) Multistate reliability theory – a case study. *Adv. Appl. Prob.* **18**, 577-579.

[16] NATVIG, B. & EIDE, H. (1987) Bayesian estimation of system reliability. *Scand. J. Statist.,* **14** 319–327.

[17] SCARSINI, M. & SHAKED, M. (1990) Stochastic ordering for permutation symmetric distributions, *Statistics & Probability Letters*, **9**, 217-222.

[18] STOYAN, D. (1983) *Comparison methods for queues and other stochastic models.* Wiley, New York.