Probability in the Engineering and Informational Sciences, **33**, 2019, 438–459.
doi:10.1017/S0269964818000177

SIGNATURE-BASED INFORMATION MEASURES OF MULTI-STATE NETWORKS

S. ZAREZADEH
Department of Statistics
Shiraz University, Shiraz 71454, Iran
E-mail: s.zarezadeh@shirazu.ac.ir

M. ASADI
Department of Statistics
University of Isfahan, Isfahan 81744, Iran
and
School of Mathematics, Institute of Research in Fundamental Sciences (IPM)
P.O Box 19395-5746, Tehran, Iran
E-mail: m.assadi@sci.ui.ac.ir

S. EFTEKHAR
Department of Statistics
Shiraz University, Shiraz 71454, Iran
E-mail: sana.eftekhar@shirazu.ac.ir

The signature matrix of an n-component three-state network (system), which depends only on the network structure, is a useful tool for comparing the reliability and stochastic properties of networks. In this paper, we consider a three-state network with states *up*, *partial performance*, and *down*. We assume that the network remains in state up, for a random time T_1 and then moves to state partial performance until it fails at time $T > T_1$. The signature-based expressions for the conditional entropy of T given T_1, the joint entropy, Kullback–Leibler (K–L) information, and mutual information of the lifetimes T and T_1 are presented. It is shown that the K–L information, and mutual information between T_1 and T depend only on the network structure (i.e., depend only to the signature matrix of the network). Some signature-based stochastic comparisons are also made to compare the K–L of the state lifetimes in two different three-state networks. Upper and lower bounds for the K–L divergence and mutual information between T_1 and T are investigated. Finally the results are extended to n-component multi-state networks. Several examples are examined graphically and numerically.

Keywords: joint entropy, Kullback–Liebler information, mutual information, order statistics, signature, shannon entropy

© Cambridge University Press 2018 0269-9648/18 $25.00

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

1. INTRODUCTION

Nowadays networks play an important role in different areas of humane life. A network is defined as a collection of *nodes* and *links* in which some particular nodes in the network are called *terminals*. Examples of nodes are road intersections, rail stations, computers, etc., and examples of links are roads, railways, communication channels, etc. We assume that a network, which is designed for some specified purpose, has elements (links or nodes) that are subject to failure over time. The failure of links (or nodes) can be occurred based on events such as electrical shocks, earthquakes, floods, or enemy attacks. The failure of a link means that the link is erased and the node failure means that all links incident to that node are erased. Assume that the performance of a network has several *states* which is defined according to existing connections between a set of specific terminals. The *state* of the network and hence its performance may change due to the failure of the links (or nodes); see, for example, [6,7]. It should be pointed out that a coherent system can be considered as a two-state network with two terminals. Throughout the paper, we assume that the links (which we call them components) are subject to failure and the nodes are absolutely reliable.

In assessing the lifetime properties and stochastic behavior of a network, the notion of *signature* is very important. The concept of signature for a single-step multi-state network is introduced in [6]. With a single-step network we mean that the failure of one component does not change the state of the network or does change it by one. We consider a single-step three-state network consisting of n binary components and assume that the network has three states: *up* ($l = 2$), *partial performance* ($l = 1$), and *down* ($l = 0$). The state $l = 2$, corresponds to the complete functioning of the network, the state $l = 1$ indicates that the network works partially, and the state $l = 0$, corresponds to the complete failure of the network. Let the lifetimes of the network components be denoted by $X_1, X_2, \ldots, X_n$ where we assume that X_i's are independent and identically distributed (i.i.d.) continuous random variables (r.v.s) with common cumulative distribution function (cdf) F. Denote also by $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ the order statistics corresponding to X_i's. Let the network be at state $l = 2$ when it starts to operate at time $t = 0$. Denote by T_1 the time duration that the network remains in state $l = 2$ and let T be the network lifetime, that is, the entrance time into state $l = 0$. Then the two-dimensional signature of the network is defined to be a probability matrix $\mathcal{S}$ with elements defined by

$$s_{i,j} = P(T_1 = X_{i:n}, T = X_{j:n}), \quad 1 \leq i < j \leq n. \tag{1.1}$$

It can be verified that $s_{i,j} = (n_{i,j}/n!)$, for $1 \leq i < j \leq n$, where $n_{i,j}$ is the number of permutations of ordinal numbers of components in which the ith and the jth component failures cause the state of the network changes from $l = 2$ to $l = 1$ and from $l = 1$ to $l = 0$, respectively. According to this, $s_{i,j}$ shows the probability that the component with lifetime $X_{i:n}$ causes transition from state $l = 2$ to state $l = 1$ and the component with lifetime $X_{j:n}, i < j$ causes transition from state $l = 1$ to state $l = 0$. Because the signature matrix $\mathcal{S}$ is free of the distribution of X_i's, an application of the law of total probability shows that, the joint survival function of (T_1, T) can be written as

$$P(T_1 > t_1, T > t) = \sum_{i,j} s_{i,j} P(X_{i:n} > t_1, X_{j:n} > t), \quad t_1 > 0, t > 0, \tag{1.2}$$

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

where

$$P(X_{i:n} > t_1, X_{j:n} > t) = \sum_{k=0}^{i-1}\sum_{l=k}^{j-1} \frac{n!}{k!(l-k)!(n-l)!} F^k(t_1)(F(t)-F(t_1))^{l-k}\bar{F}^{n-l}(t), \quad t_1 < t;$$

see [6]. Note that in representation (1.2), $s_{i,j} = 0$ for $i \geq j$. For results on signature-based properties of the lifetime of networks and coherent systems with two or more states, we refer the reader to [1,2,7,11,14,15,24]. It should be mentioned that the probability matrix with elements defined as (1.1) is equal to the bivariate signature matrix (joint signature) corresponding to two systems with shared components given in [12] (see also [22]). Hence, the reliability function (1.2) is clearly equivalent to the reliability function of the lifetimes of two coherent (or semi-coherent) systems with shared components presented in Theorem 2.1 of [12]. Zarezadeh, Mohammadi, and Balakrishnan [26] gave some pseudo-mixture representations for the joint reliability function of the lifetimes of more than two systems with shared components.

The purpose of this paper is to investigate the signature-based information properties of the state lifetimes of a multi-state network. Expressions for Shannon entropy, Rényi entropy, and Kullback–Leibler (K–L) information for the lifetime of k-out-of-n systems have been obtained in [5,25]. Recently, information properties of general coherent systems based on signature have been considered by authors. Toomaj and Doostparast [18,19] obtained signature-based expressions for the Shannon entropy and K–L information of mixed coherent systems. Some properties of the cumulative residual entropy of mixed systems were explored in [20]. Asadi et al. [3] used the Jensen–Shannon divergence to compare different coherent structures based on signatures.

The rest of the paper is organized as follows: In Section 2, we give some preliminary results which are useful in subsequent sections. These include, among others, the definitions of differential entropy, K–L divergence, conditional entropy and mutual entropy. In Section 3, we consider a three-state network described above and obtain signature-based representations for the conditional entropy of T given T_1 and the joint entropy of (T_1, T). It is shown that the K–L divergence and mutual information (as a dependency measure) between T and T_1 are free of the components lifetime and depend only on the network structure (signature). Stochastic orderings results are also given to compare the K–L of the state lifetimes in different three-state networks based on signature. Upper and lower bounds (under decreasing failure rate condition) of the K–L divergence are investigated. An upper bound is also presented for the mutual information between T and T_1. The results are examined numerically and graphically for some basic networks. In Section 4, the results are extended to the multi-state networks. General signature-based representations for information measures of lifetimes of a multi-state network are demonstrated.

2. PRELIMINARIES

Let X be a continuous r.v. with probability density function (pdf) f_X. The Shannon entropy (or differential enrtopy) of X, denoted by $H(X)$, is defined as

$$H(X) = -\int f_X(t)\log f_X(t)dt;$$

see [17]. The entropy $H(X)$ measures the amount of uncertainty or uniformity of the pdf f_X. As $H(X)$ increases, the pdf f_X tends to a uniform. Hence, the concentration of probabilities

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

decreases and it becomes more difficult to predict the values of outcomes generated from pdf f_X. The quantity $I(X) = -H(X)$ is called the information measure corresponding to X. When X and Y are continuous r.v.s with pdf's f_X and f_Y, respectively, the K–L divergence between f_Y and f_X is defined as

$$K(X:Y) = K(f_Y : f_X) = \int f_Y(t) \log \frac{f_Y(t)}{f_X(t)} dt \geq 0,$$

provided that the cdf F_Y of Y is absolutely continuous with respect to the cdf F_X of X; see [10]. The K–L divergence, which is also known as *information divergence*, *information gain* and *relative entropy*, measures the difference between two pdf's f_Y and f_X when the distribution f_X is used as the reference distribution. $K(X:Y) = 0$ if and only if $f_Y(t) = f_X(t)$ almost everywhere. For two r.v.s X and Y, the conditional entropy of Y given X, denoted by $H(Y|X)$, specifies the expected amount of information needed to predict the outcome of Y given that the value of the X is known. The conditional entropy of Y given X is defined as

$$H(Y|X) = \int f_X(x) H(Y|X = x) dx, \tag{2.1}$$

where $H(Y|X = x) = -\int f_{Y|X}(y|x) \log f_{Y|X}(y|x) dy$ is the Shannon entropy of Y under the condition that X takes a certain value x. The Shannon entropy of random vector (X, Y), with joint pdf $f_{X,Y}(x, y)$, is called joint entropy and it is defined as

$$H(X,Y) = -\iint f_{X,Y}(x,y) \log f_{X,Y}(x,y) dx dy. \tag{2.2}$$

An important criterion to measure the mutual dependency between two r.v.s X and Y, is the mutual information. More specifically, the mutual information quantifies the "amount of information" gained about r.v. Y, through the r.v. X. The mutual information between two r.v.s X and Y with pdf's f_X and f_Y is defined by

$$M(X,Y) = K(f_{X,Y} : f_X f_Y) \geq 0.$$

It can be shown that $M(X, Y) = 0$ if and only if two r.v.s X and Y are independent.

The following definition is used to obtain some ordering results on network lifetimes.

DEFINITION 2.1: *For two r.v.s X and Y with survival functions $\bar{F}$ and $\bar{G}$, and pdf's f and g, respectively,*

(i) *X is said to be smaller than Y in the usual stochastic order, denoted by $X \leq_{st} Y$, if $\bar{F}(t) \leq \bar{G}(t)$ for all t;*

(ii) *X is said to be smaller than Y in the likelihood ratio order, denoted by $X \leq_{lr} Y$, if $(g(t)/f(t))$ is increasing in t.*

It is remarkable that $X \leq_{st} Y$ if and only if $E(\phi(X)) \leq E(\phi(Y))$ for all increasing function $\phi(\cdot)$. Also, $X \leq_{lr} Y$ implies that $X \leq_{st} Y$. For more details on the various concepts of stochastic orderings, we refer the reader to [16].

We also recall that r.v. X is said to have beta distribution with parameters (a, b) if its pdf is given as

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, \quad 0 < x < 1,\ a > 0, b > 0, \tag{2.3}$$

where $B(a,b) = (\Gamma(a)\Gamma(b)/\Gamma(a+b))$, $a > 0, b > 0$, is the beta function and $\Gamma(c) = \int_0^\infty x^{c-1} e^{-x} dx$, $c > 0$, is the gamma function. The ordered-Dirichlet distribution is a generalization of beta distribution (see [23], p. 182). A random vector $\mathbf{X} = (X_1, \ldots, X_m)$ is

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

said to have ordered-Dirichlet distribution with parameters $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{m+1})$, denoted by $\mathbf{X} \sim OD(\boldsymbol{\alpha})$, if the joint pdf of $\mathbf{X}$ is given by

$$f(\mathbf{x}|\boldsymbol{\alpha}) = \frac{1}{B(\boldsymbol{\alpha})}(1-x_m)^{\alpha_{m+1}-1}\prod_{i=1}^{m}(x_i - x_{i-1})^{\alpha_i - 1}, \quad 0 = x_0 < x_1 < x_2 < \cdots < x_m < 1, \tag{2.4}$$

where $B(\boldsymbol{\alpha})$ is the multivariate beta function defined, for $k = 2, 3, \ldots$, as

$$B(a_1, \ldots, a_k) = \frac{\prod_{i=1}^{k}\Gamma(a_i)}{\Gamma(\sum_{i=1}^{k} a_i)}, \quad a_1, \ldots, a_k > 0.$$

In fact, the beta distribution is obtained when $m = 1$. If $U_{k:n}$ denotes the kth order statistic associated to a random sample with size n from standard uniform distribution, then $(U_{n_1:n}, U_{n_2:n}, \ldots, U_{n_p:n})$, $1 \le n_1 < \cdots < n_p \le n$, $p = 1, \ldots, n$ has ordered-Dirichlet distribution with parameters $\alpha_i = n_i - n_{i-1}$, $i = 1, \ldots, p+1$, $n_0 = 0$, $n_{p+1} = n+1$. Throughout the paper, we show the joint pdf of $(U_{n_1:n}, U_{n_2:n}, \ldots, U_{n_p:n})$ by $h_{n_1,\ldots,n_p}(u_1, \ldots, u_p)$ and the pdf of beta distribution with parameters $(j, n-j+1)$ by h_j. We set also by convention

$$C^n_{i_1,\ldots,i_r} = \frac{1}{B(i_1, i_2 - i_1, \ldots, n - i_r + 1)}, \quad 1 \le i_1 < i_2 < \cdots < i_r \le n.$$

3. INFORMATION MEASURES OF THREE-STATE NETWORKS

Consider a network with three states: *up state*, *partial performance*, and *down state*. Let the network have n components with i.i.d. lifetimes $X_1, \ldots, X_n$ such that X_1 has pdf f, cdf F and reliability function $\bar{F} = 1 - F$. Further, assume that T_1 and T $(T_1 < T)$ denote the times of entrance into partial performance and down state, respectively. The marginal survival functions of T_1 and T are given, respectively, as

$$P(T_1 > t) = \sum_{i=1}^{n} s_i^{(1)} P(X_{i:n} > t) \quad \text{where} \quad s_i^{(1)} = P(T_1 = X_{i:n}) = \sum_{j=1}^{n} s_{i,j}, \quad i = 1, \ldots, n, \tag{3.1}$$

and

$$P(T > t) = \sum_{j=1}^{n} s_j^{(2)} P(X_{j:n} > t) \quad \text{where} \quad s_j^{(2)} = P(T = X_{j:n}) = \sum_{i=1}^{n} s_{i,j}, \quad j = 1, \ldots, n.$$

Using (1.2), the joint pdf of (T_1, T) can be written as

$$f_{T_1,T}(t_1, t) = \sum_{i,j} s_{i,j} f_{i,j:n}(t_1, t), \tag{3.2}$$

where $f_{i,j:n}$ is the joint pdf of $(X_{i:n}, X_{j:n})$ given by

$$f_{i,j:n}(t_1, t) = C^n_{i,j} f(t) f(t_1) F^{i-1}(t_1)[F(t) - F(t_1)]^{j-i-1}\bar{F}^{n-j}(t).$$

Note that (T_1, T) has an absolutely continuous distribution due to the 'single-step' property of the network (which implies $T > T_1$). From (3.2), the conditional pdf of T given T_1 is

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

written as

$$
\begin{aligned}
f_{T|T_1}(t|t_1) &= \frac{\sum_{i,j} s_{i,j} C_{i,j}^n f(t) f(t_1) F^{i-1}(t_1)[F(t) - F(t_1)]^{j-i-1} \bar{F}^{n-j}(t)}{\sum_{i=1}^n s_i^{(1)} C_i^n f(t_1) F^{i-1}(t_1) \bar{F}^{n-i}(t_1)} \\
&= \sum_{i,j} p_{i,j}(t_1) \frac{C_{i,j}^n}{C_i^n} \frac{f(t)}{(\bar{F}(t_1))^{n-i}} (\bar{F}(t_1) - \bar{F}(t))^{j-i-1} (\bar{F}(t))^{n-j} \\
&= \sum_{i,j} p_{i,j}(t_1) C_{j-i}^{n-i} \frac{f(t)}{\bar{F}(t_1)} [1 - \frac{\bar{F}(t)}{\bar{F}(t_1)}]^{j-i-1} (\frac{\bar{F}(t)}{\bar{F}(t_1)})^{n-j} \\
&= \sum_{i,j} p_{i,j}(t_1) g_{j-i:n-i}(t|t_1),
\end{aligned}
$$

where $g_{j-i:n-i}(t|t_1)$ is the pdf of $(j-i)$th order statistic from conditional r.v. $(X|X > t_1)$ with sample size $(n-i)$ and

$$
p_{i,j}(t_1) = \frac{s_{i,j} C_i^n f(t_1) F^{i-1}(t_1) \bar{F}^{n-i}(t_1)}{\sum_{i=1}^n s_i^{(1)} C_i^n f(t_1) F^{i-1}(t_1) \bar{F}^{n-i}(t_1)}. \tag{3.3}
$$

It is interesting to note that $p_{i,j}(t_1)$ is in fact the probability that the ith and the jth component failures cause to enter the network state from $l = 2$ to $l = 1$ and from $l = 1$ to $l = 0$, given that the time of entrance to partial performance is t_1; that is,

$$
p_{i,j}(t_1) = P(T_1 = X_{i:n}, T = X_{j:n} | T_1 = t_1).
$$

This is so, because we can write

$$
\begin{aligned}
s_{i,j} &= P(T_1 = X_{i:n}, T = X_{j:n}) \\
&= P(T_1 = X_{i:n}) P(T = X_{j:n} | T_1 = X_{i:n}) = s_i^{(1)} s_{j|i},
\end{aligned}
$$

where

$$
s_{j|i} := P(T = X_{j:n} | T_1 = X_{i:n}). \tag{3.4}
$$

On the other hand, from [21],

$$
\begin{aligned}
p_i^{(1)}(t_1) &:= P(T_1 = X_{i:n} | T_1 = t_1) \\
&= \frac{s_i^{(1)} C_i^n f(t_1) F^{i-1}(t_1) \bar{F}^{n-i}(t_1)}{\sum_{i=1}^n s_i^{(1)} C_i^n f(t_1) F^{i-1}(t_1) \bar{F}^{n-i}(t_1)}.
\end{aligned}
$$

Then $p_{i,j}(t_1)$ in (3.3) is rewritten as

$$
\begin{aligned}
p_{i,j}(t_1) &= s_{j|i} p_i^{(1)}(t_1) \\
&= P(T_1 = X_{i:n}, T = X_{j:n} | T_1 = t_1),
\end{aligned} \tag{3.5}
$$

where $s_{j|i}$ is defined in (3.4).

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

3.1. Predictability of T Given T_1

In this section, we obtain the conditional entropy of T given T_1 and the joint entropy of T and T_1. First note that, based on what we have given above, we get

$$\begin{aligned} H(T|T_1 = t_1) &= -\int_0^\infty f_{T|T_1}(t|t_1) \log f_{T|T_1}(t|t_1) dt \\ &= -\int_{F(t_1)}^1 q(t_1, v) \log(f(F^{-1}(v)) q(t_1, v)) dv, \end{aligned} \tag{3.6}$$

where we used the transformation $V = F(T)$ and $q(t_1, v)$ is given as:

$$q(t_1, v) = \frac{\sum_{i,j} s_{i,j} C_{i,j}^n F^{i-1}(t_1)[v - F(t_1)]^{j-i-1}(1-v)^{n-j}}{\sum_{i=1}^n s_i^{(1)} C_i^n F^{i-1}(t_1) \bar{F}^{n-i}(t_1)}.$$

Now, we have the following theorem.

THEOREM 3.1: *Consider a three-state network with n components having i.i.d. lifetimes $X_1, \ldots, X_n$ and signature matrix $\mathcal{S} = \{s_{i,j}\}_{n \times n}$. Further, assume that T_1 and T denote the time instances of entrance into partial performance and down state, respectively. If the component lifetimes have common pdf f and cdf F, and $W_{j:n}$ has beta distribution with parameters $(j, n-j+1)$, then the conditional entropy of T given T_1 is given as*

$$\begin{aligned} H(T|T_1) &= H(V_1, V) - \sum_{j=1}^n s_j^{(2)} E[\log f(F^{-1}(W_{j:n}))] - H(V_1) \\ &= H(V|V_1) - \sum_{j=1}^n s_j^{(2)} E[\log f(F^{-1}(W_{j:n}))] \end{aligned} \tag{3.7}$$

where $V_1 = F(T_1)$ and $V = F(T)$ with joint pdf as

$$f_{V_1,V}(v_1, v) = \sum_{i,j} s_{i,j} h_{i,j}(v_1, v),$$

and hence

$$H(V_1, V) = -\sum_{i,j} s_{i,j} \int_0^1 \int_{v_1}^1 h_{i,j}(v_1, v) \log \sum_{i,j} s_{i,j} h_{i,j}(v_1, v) dv dv_1, \tag{3.8}$$

$$H(V_1) = -\sum_{i=1}^n s_i^{(1)} \int_0^1 h_i(v_1) \log \sum_{i=1}^n s_i^{(1)} h_i(v_1) dv_1, \tag{3.9}$$

in which h_i and $h_{i,j}$ denote, respectively, the pdf of beta distribution with parameters $(i, n-i+1)$ and the pdf of $OD(i, j-i, n-j+1)$ in (2.4).

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

PROOF: Since the pdf of T_1 is given as

$$f_{T_1}(t_1) = f(t_1) \sum_{i=1}^{n} s_i^{(1)} C_i^n F^{i-1}(t_1) \bar{F}^{n-i}(t_1),$$

from (2.1) and using the probability integral transformations $V_1 = F(T_1)$ and $V = F(T)$, we have

$$\begin{aligned}
H(T|T_1) &= \int_0^\infty f_{T_1}(t_1) H(T|t_1) dt_1 \\
&= -\int_0^\infty \int_{t_1}^\infty f_{T_1,T}(t_1,t) \log f_{T|T_1}(t|t_1) dt dt_1 \\
&= -\sum_{i,j} s_{i,j} \int_0^1 \int_{v_1}^1 h_{i,j}(v_1,v) \log \frac{f(F^{-1}(v)) \sum_{i,j} s_{i,j} h_{i,j}(v_1,v)}{\sum_{i=1}^n s_i^{(1)} h_i(v_1)} dv dv_1 \\
&= -\sum_{i,j} s_{i,j} \int_0^1 \int_{v_1}^1 h_{i,j}(v_1,v) \log \sum_{i,j} s_{i,j} h_{i,j}(v_1,v) dv dv_1 \\
&\quad - \sum_{i,j} s_{i,j} \int_0^1 \int_{v_1}^1 h_{i,j}(v_1,v) \log f(F^{-1}(v)) dv dv_1 \\
&\quad + \sum_{i,j} s_{i,j} \int_0^1 \int_{v_1}^1 h_{i,j}(v_1,v) \log \sum_{i=1}^n s_i^{(1)} h_i(v_1) dv dv_1 \\
&= H(V_1,V) - \sum_{j=1}^n s_j^{(2)} E[\log f(F^{-1}(W_{j:n}))] - H(V_1) \\
&= H(V|V_1) - \sum_{j=1}^n s_j^{(2)} E[\log f(F^{-1}(W_{j:n}))] \qquad \textbf{(3.10)}
\end{aligned}$$

where the last equality follows from equation $H(V|V_1) = H(V_1,V) - H(V_1)$. This completes the theorem. ■

The function $f(F^{-1}(\cdot))$, in this representation, is called density-quantile function; see [13]. This representation also shows that the difference between the conditional entropies of $H(T|T_1)$ and $H(V|V_1)$ depends to the signature matrix through the marginal signatures $s_j^{(2)} = P(T = X_{j:n})$, $j = 1, \ldots, n$.

The following theorem gives the joint entropy of T_1 and T.

THEOREM 3.2: *Under the assumptions of Theorem 3.1, the joint entropy of* (T_1, T) *is represented as*

$$H(T_1,T) = H(V_1,V) - \sum_{i=1}^n (s_i^{(1)} + s_i^{(2)}) E[\log f(F^{-1}(W_{i:n}))],$$

where $H(V_1,V)$ *is given in* (3.8).

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

PROOF: Since T_1 can be considered as the lifetime of a two-state network with signature vector $\mathbf{s}^{(1)} = (s_1^{(1)}, \ldots, s_n^{(1)})$, based on relation (6) of [18], we have

$$H(T_1) = H(V_1) - \sum_{i=1}^{n} s_i^{(1)} E[\log f(F^{-1}(W_{i:n}))]. \tag{3.11}$$

Therefore, from (3.7) and using the fact that $H(T, T_1) = H(T|T_1) + H(T_1)$, we get the required result. ■

In the following, we provide a lower bound for the joint entropy of (T_1, T). To this end, first we prove the following lemma.

LEMMA 3.3: *If $W_{i:n}$ denotes the ith order statistic associated to a random sample with size n from standard uniform distribution, then the joint entropy of $(W_{i:n}, W_{j:n})$, $1 \le i < j \le n$, can be written as*

$$\begin{aligned} H(W_{i:n}, W_{j:n}) = &- \log C_{i,j}^n - (i-1)\psi(i) - (j-i-1)\psi(j-i) \\ &- (n-j)\psi(n-j+1) + (n-2)\psi(n+1), \end{aligned}$$

where $\psi(t) = (d/dt)\log\Gamma(t)$, $t > 0$, is digamma function.

PROOF: Based on the definition of joint entropy, we have

$$\begin{aligned} H(W_{i:n}, W_{j:n}) &= - \log C_{i,j}^n - (i-1) \int_0^1 \int_0^{w_j} \log(w_i) f_{W_{i:n},W_{j:n}}(w_i, w_j) dw_i dw_j \\ &\quad - (j-i-1) \int_0^1 \int_0^{w_j} \log(w_j - w_i) f_{W_{i:n},W_{j:n}}(w_i, w_j) dw_i dw_j \\ &\quad - (n-j) \int_0^1 \int_0^{w_j} \log(1-w_j) f_{W_{i:n},W_{j:n}}(w_i, w_j) dw_i dw_j \\ &= - \log C_{i,j}^n - (i-1)E(\log W_{i:n}) - (j-i-1)E(\log(W_{j:n} - W_{i:n})) \\ &\quad - (n-j)E(\log(1 - W_{j:n})) \\ &= - \log C_{i,j}^n - (i-1)\psi(i) - (j-i-1)\psi(j-i) \\ &\quad - (n-j)\psi(n-j+1) + (n-2)\psi(n+1). \end{aligned}$$

This completes the proof of lemma. ■

Now, we prove the following theorem which gives a lower bound for joint entropy of (T_1, T).

THEOREM 3.4: *A lower bound for the joint entropy of (T_1, T) is*

$$H_L(T_1, T) = \sum_{i,j} s_{i,j} H(W_{i:n}, W_{j:n}) - \sum_{i=1}^{n} (s_i^{(1)} + s_i^{(2)}) E[\log f(F^{-1}(W_{i:n}))],$$

where $H(W_{i:n}, W_{j:n})$ is given in Lemma 3.3.

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

PROOF: Since the entropy is a concave function of the underlying pdf, from (3.2) we can write

$$H(T_1, T) \geq \sum_{i,j} s_{i,j} H(X_{i:n}, X_{j:n}), \tag{3.12}$$

where

$$H(X_{i:n}, X_{j:n}) = -\int_0^\infty \int_0^{x_j} f_{i,j:n}(x_i, x_j) \log f_{i,j:n}(x_i, x_j) dx_i dx_j.$$

Based on the probability integral transformations $W_{i:n} = F(X_{i:n})$ and $W_{j:n} = F(X_{j:n})$ we have

$$H(X_{i:n}, X_{j:n}) = -\int_0^1 \int_0^{w_j} h_{i,j}(w_i, w_j) \log(f(F^{-1}(w_i)) f(F^{-1}(w_j)) h_{i,j}(w_i, w_j)) dw_i dw_j.$$

Since $W_{i:n}$ and $W_{j:n}$ are the ith and the jth order statistics associated to a random sample from standard uniform distribution, then the joint pdf of $(W_{i:n}, W_{j:n})$ is as $h_{i,j}(w_i, w_j)$ and hence we conclude that

$$H(X_{i:n}, X_{j:n}) = H(W_{i:n}, W_{j:n}) - E[\log f(F^{-1}(W_{i:n}))] - E[\log f(F^{-1}(W_{j:n}))]. \tag{3.13}$$

By substituting (3.13) in (3.12), the proof is complete. ■

3.2. The K–L and Mutual Information between T and T_1

In this section, we obtain two further information measures regarding the r.v.s T_1 and T. We first obtain K–L divergence between T_1 and T in the case when the distribution of T_1 is considered as the reference distribution. It can be shown, using the transformation $V = F(T)$, that the K–L information between T_1 and T is given as

$$K(T : T_1) = \sum_{j=1}^n s_j^{(2)} \int_0^1 h_j(v) \log \frac{\sum_{j=1}^n s_j^{(2)} h_j(v)}{\sum_{i=1}^n s_i^{(1)} h_i(v)} dv, \tag{3.14}$$

where h_j is the beta distribution with parameters $(j, n-j+1)$. This shows that $K(T : T_1)$ depends only on the network structure and is free of the distribution of component lifetimes.

THEOREM 3.5: *Let $W_{i:n}$ denote the ith order statistic of a random sample with size n from standard uniform distribution. For a three-state network with state lifetimes (T_1, T) and signature matrix $\mathcal{S} = \{s_{i,j}\}_{n\times n}$, an upper bound for $K(T : T_1)$ is obtained as*

$$K_U(T : T_1) = \sum_{i=1}^n \sum_{j=1}^n s_i^{(1)} s_j^{(2)} K(W_{j:n} : W_{i:n})$$

where, for $i = 1, \ldots, n$, $j = 1, \ldots, n$,

$$K(W_{j:n} : W_{i:n}) = \log\left(\frac{\Gamma(i)\Gamma(n-i+1)}{\Gamma(j)\Gamma(n-j+1)}\right) + (j-i)[\psi(j) - \psi(n-j+1)], \tag{3.15}$$

in which $\psi(a)$, $a > 0$, is the digamma function.

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

PROOF: From (3.14), we have

$$K(T:T_1) = K(V:V_1) = \int f_V(v) \log \frac{f_V(v)}{f_{V_1}(v)} dv$$
$$= -H(V) - \int f_V(v) \log f_{V_1}(v) dv,$$

where $H(V)$ is the entropy of V presented as

$$H(V) = -\sum_{i=1}^{n} \int_0^1 s_i^{(2)} h_i(v) \log \left(\sum_{i=1}^{n} s_i^{(2)} h_i(v) \right) dv. \tag{3.16}$$

From [18], a lower bound for the Shannon entropy of V is given as

$$H(V) \geq \sum_{j=1}^{n} s_j^{(2)} H(W_{j:n}). \tag{3.17}$$

On the other hand, we have

$$\int f_V(v) \log f_{V_1}(v) dv = \sum_{j=1}^{n} s_j^{(2)} \int h_j(v) \log \left(\sum_{i=1}^{n} s_i^{(1)} h_i(v) \right) dv \tag{3.18}$$
$$\geq \sum_{i=1}^{n} \sum_{j=1}^{n} s_i^{(1)} s_j^{(2)} \int h_j(v) \log h_i(v) dv \tag{3.19}$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} s_i^{(1)} s_j^{(2)} (-H(W_{j:n}) - K(W_{j:n} : W_{i:n}))$$
$$= -\sum_{j=1}^{n} s_j^{(2)} H(W_{j:n}) - \sum_{i=1}^{n} \sum_{j=1}^{n} s_i^{(1)} s_j^{(2)} K(W_{j:n} : W_{i:n}) \tag{3.20}$$

where the inequality follows from Jensen inequality. Therefore from (3.17) and (3.20), the theorem is concluded. ■

In the following, under the condition that V_1 has decreasing failure rate, a lower bound is given for the K–L between T and T_1. We recall that the failure rate of a continuous r.v. with pdf f and survival function $\bar{F}$ is defined as $\lambda(x) = f(x)/\bar{F}(x)$.

THEOREM 3.6: *Consider a three-state network with state lifetimes (T_1, T) and signature matrix $\mathcal{S} = \{s_{i,j}\}_{n\times n}$. If V_1 has decreasing failure rate, then a lower bound for $K(T:T_1)$ is as*

$$K_L(T:T_1) = \max(0, H(V_1) - H(V)), \tag{3.21}$$

where $H(V_1)$ and $H(V)$ are given in (3.9) and (3.16), respectively.

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

PROOF: Let f_{V_1} and f_V denote the pdf's of V_1 and V, respectively. From (3.14), we have

$$\begin{aligned}
K(T:T_1) = K(V:V_1) &= \int f_V(v)\log\frac{f_V(v)}{f_{V_1}(v)}dv \\
&= \int f_V(v)\log f_V(v)dt - \int f_V(v)\log f_{V_1}(v)dv \\
&= -H(V) - \int f_V(v)\log\lambda_{V_1}(v)dv - \int f_V(v)\log\bar{F}_{V_1}(v)dv \\
&\geq -H(V) - \int f_{V_1}(v)\log\lambda_{V_1}(v)dv - \int f_{V_1}(v)\log\bar{F}_{V_1}(v)dv \\
&= -H(V) + H(V_1),
\end{aligned}$$

where the inequality follows from the fact that V_1 has decreasing failure rate and $P(V_1 < V) = 1$ which implies $V_1 \leq_{st} V$. ■

Remark 3.7: In the lower bound (3.21), a sufficient condition under which $H(V_1) \geq H(V)$ is that, under the assumptions of the theorem, V has an increasing pdf; see [4].

THEOREM 3.8: *Consider two three-state networks with state lifetimes* (T_1, T) *and* (T_1^*, T^*). *Let the components of each network have i.i.d. lifetimes with common absolutely continuous cdf* F. *Further, assume that the corresponding signatures matrices of the networks are* $\mathcal{S} = \{s_{i,j}\}_{n\times n}$ *and* $\mathcal{S}^* = \{s^*_{ij}\}_{n\times n}$, *respectively, such that* $\boldsymbol{s}^{(1)} = \boldsymbol{s}^{*(1)}$.

(a) *If* $\boldsymbol{s}^{(1)} \leq_{lr} (\geq_{lr})\boldsymbol{s}^{(2)}$ *and* $\boldsymbol{s}^{(2)} \leq_{st} (\geq_{st})\boldsymbol{s}^{*(2)}$, *then* $K(T:T_1) \leq K(T^*:T_1^*)$.

(b) *If* $\boldsymbol{s}^{*(1)} \geq_{lr} (\leq_{lr})\boldsymbol{s}^{*(2)}$ *and* $\boldsymbol{s}^{(2)} \leq_{st} (\geq_{st})\boldsymbol{s}^{*(2)}$, *then* $K(T:T_1) \geq K(T^*:T_1^*)$.

PROOF: We prove part (a). The proof of part (b) is similar. From Theorem 5 of [9], the assumption of $\mathbf{s}^{(1)} \leq_{lr} (\geq_{lr})\mathbf{s}^{(2)}$ implies $V_1 \leq_{lr} (\geq_{lr})V$ and hence $(\sum_{j=1}^n s_j^{(2)}h_j(v)/\sum_{i=1}^n s_i^{(1)}h_i(v))$ is increasing (decreasing) in v. On the other hand, from $\mathbf{s}^{(2)} \leq_{st} (\geq_{st})\mathbf{s}^{*(2)}$ and Theorem 3 of [9], we have $V \leq_{st} (\geq_{st})V^*$ and hence for any increasing function $\phi(\cdot)$, $E(\phi(V)) \leq (\geq)E(\phi(V_1))$. Then

$$\begin{aligned}
K(T:T_1) &= \int_0^1 \sum_{j=1}^n s_j^{(2)}h_j(v)\log\frac{\sum_{j=1}^n s_j^{(2)}h_j(v)}{\sum_{i=1}^n s_i^{(1)}h_i(v)}dv, \\
&\leq \int_0^1 \sum_{j=1}^n s_j^{*(2)}h_j(v)\log\frac{\sum_{j=1}^n s_j^{(2)}h_j(v)}{\sum_{i=1}^n s_i^{(1)}h_i(v)}dv \\
&= \int_0^1 \sum_{j=1}^n s_j^{*(2)}h_j(v)\log\sum_{j=1}^n s_j^{(2)}h_j(v)dv - \int_0^1 \sum_{j=1}^n s_j^{*(2)}h_j(v)\log\sum_{i=1}^n s_i^{(1)}h_i(v)dv \\
&\leq \int_0^1 \sum_{j=1}^n s_j^{*(2)}h_j(v)\log\sum_{j=1}^n s_j^{*(2)}h_j(v)dv - \int_0^1 \sum_{j=1}^n s_j^{*(2)}h_j(v)\log\sum_{i=1}^n s_i^{(1)}h_i(v)dv \\
&= \int_0^1 \sum_{j=1}^n s_j^{*(2)}h_j(v)\log\frac{\sum_{j=1}^n s_j^{*(2)}h_j(v)}{\sum_{i=1}^n s_i^{(1)}h_i(v)}dv = K(T^*:T_1^*)
\end{aligned}$$

where the second inequality follows from the non-negativity property of K–L divergence. ■

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

Remark 3.9: We have been able to get the result of Theorem 3.8 under the restrictive condition that for the two networks $\mathbf{s}^{(1)} = \mathbf{s}^{*(1)}$. An interesting problem is to investigate whether the result of the theorem remains valid if one drops this condition or rather impose weaker condition.

Using the same approach as used to get (3.14), it can be seen that the mutual information between T_1 and T is as follows:

$$M(T,T_1) = \sum_{i,j} s_{i,j} \int_0^1 \int_0^v h_{i,j}(v_1,v) \log \frac{\sum_{i,j} s_{i,j} h_{i,j}(v_1,v)}{\sum_{j=1}^n s_j^{(2)} h_j(v) \sum_{i=1}^n s_i^{(1)} h_i(v_1)} dv_1 dv. \quad \textbf{(3.22)}$$

This indicates that the mutual information $M(T,T_1)$ depends only on the signature matrix and hence its values are determined only based on the network structure. That is, the reduction in the uncertainty of network lifetime, T, due to the knowledge of the lifetime T_1 depends only on the network structure and is free of the lifetime distributions of T_1 and T_2.

The following theorem gives an upper bound for the mutual information between the entrance times of a three-state network into partial performance and down state.

THEOREM 3.10: *Under the assumptions of Theorem 3.5, we have the following upper bound for $M(T,T_1)$:*

$$M_U(T,T_1) = \sum_{i,j} s_{i,j} M(W_{i:n}, W_{j:n}) + \sum_{i=1}^n \sum_{j=1}^n (s_i^{(2)} s_j^{(2)} + s_i^{(1)} s_j^{(1)}) K(W_{j:n} : W_{i:n}),$$

where, for $1 \le i < j \le n$,

$$M(W_{i:n}, W_{j:n}) = \log \frac{\Gamma(n-i+1)\Gamma(j)}{\Gamma(n+1)\Gamma(j-i)} - (j-1)\psi(j) + (j-i-1)\psi(j-i) - (n-i)\psi(n-i+1) + n\psi(n+1)$$

and $K(W_{j:n} : W_{i:n})$ is obtained from (3.15).

PROOF: From (3.22), we have $M(T_1,T) = M(V_1,V) = H(V) + H(V_1) - H(V_1,V)$. Then, using the concavity of entropy, we get

$$\begin{aligned} M(T_1,T) &\le H(V) + H(V_1) - \sum_{i=1}^n \sum_{j=1}^n s_{i,j} H(W_{i:n}, W_{j:n}) \\ &= \sum_{i=1}^n \sum_{j=1}^n s_{i,j} M(W_{i:n}, W_{j:n}) + (H(V_1) - \sum_{i=1}^n s_i^{(1)} H(W_{i:n})) \\ &\quad + (H(V) - \sum_{i=1}^n s_j^{(2)} H(W_{j:n})). \end{aligned} \quad \textbf{(3.23)}$$

On the other hand, by the Jensen inequality, we have

$$H(V) \le -\sum_{i=1}^n \sum_{j=1}^n s_i^{(2)} s_j^{(2)} \int_0^1 h_i(u) \log h_j(u) du$$

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

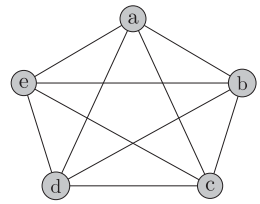


FIGURE **1.** A network with ten links, five nodes, and five terminals.

and hence

$$H(V) - \sum_{j=1}^{n} s_j^{(2)} H(W_{j:n}) \le \sum_{i=1}^{n}\sum_{j=1}^{n} s_i^{(2)} s_j^{(2)} K(W_{j:n} : W_{i:n}). \tag{3.24}$$

The same argument shows that

$$H(V_1) - \sum_{i=1}^{n} s_i^{(1)} H(W_{i:n}) \le \sum_{i=1}^{n}\sum_{j=1}^{n} s_i^{(1)} s_j^{(1)} K(W_{j:n} : W_{i:n}). \tag{3.25}$$

Therefore the required result follows from (3.23)–(3.25). ■

3.3. Examples

Let us look at the following examples:

Example 3.11: Consider a network with 10 links (components) and 5 nodes depicted in Figure 1. The network is defined to be in up state if all nodes are connected, partial performance if it is isolated into two nonconnected sets of nodes, and down state if it is isolated into more than two nonconnected sets of nodes. Let the links be subject to failure and the nodes be absolutely reliable.

In [7], the nonzero elements of the signature matrix $\mathcal{S}$ are computed as

$$s_{4,7} = 0.0047, \quad s_{4,8} = 0.0194, \quad s_{5,7} = 0.0191, \quad s_{5,8} = 0.0751,$$
$$s_{6,7} = 0.0596, \quad s_{6,8} = 0.2270, \quad s_{7,8} = 0.5951.$$

Using these values and formulas given above, one can show numerically, by computational software such as MAPLE, that $H(V_1) = -0.4504381$, $H(V) = -0.6494429$, $H(V_1, V) = -1.703062$, and $\sum_{i,j} s_{i,j} H(W_{i:n}, W_{j:n}) = -1.906097$.

Now assume that the link lifetimes are i.i.d. with either exponential distribution (Exp(λ)), or Weibull distribution (W(α,1)) with pdf's

$$g_E(t) = \frac{1}{\lambda} \exp\left(-\frac{t}{\lambda}\right), \quad t > 0,\ \lambda > 0,$$
$$g_W(t) = \alpha t^{\alpha-1} \exp(-t^{\alpha}), \quad t > 0,\ \alpha > 0,$$

respectively. Figure 2 shows the Shannon entropy of T, conditional entropy of T given T_1, the joint entropy of (T_1, T), and the lower bound obtained in Theorem 3.4, for these two distributions as a function of parameters.

It is evident from the left panel of Figure 2 that when the scale parameter (mean) of the exponential distribution increases, each of the information of the network lifetime and

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

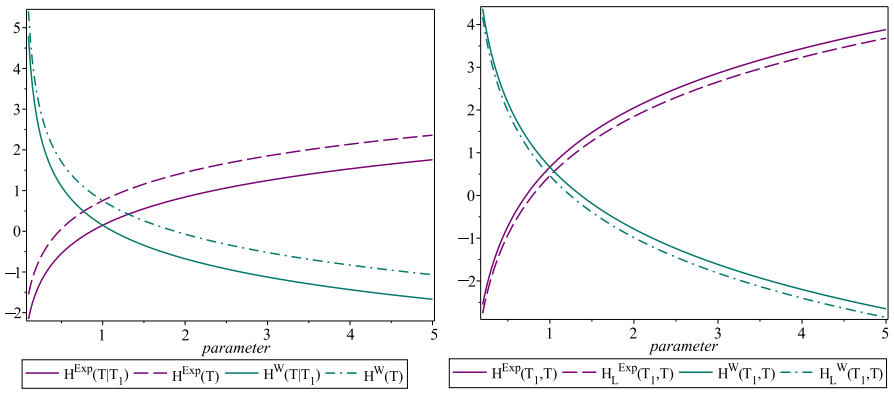


FIGURE **2.** Different information measures for the network in Example 3.11.

the information of T given T_1 increases. But in the case where the lifetime of links are distributed as Weibull, these measures decrease when the shape parameter α increases.

The right panel of Figure 2 shows the joint entropies and the corresponding lower bounds for both distributions as a function of parameters. It is seen that the values of lower bound and the joint entropy of network are relatively close to each other.

Example 3.12: Consider again Example 3.11. For this network it can be shown that, regardless of the distribution function of components lifetime, $K(T : T_1) = 0.4119464$ and $M(T, T_1) = 0.6039588$.

Example 3.13: Consider the bridge network with graph as depicted in Figure 3. This network has 4 nodes and 5 links (components). Let the links of the network be subject to failure and the nodes be absolutely reliable. We define the states of the network as three following cases:

(a) Suppose that the capacity of each link is one. The states of the network are described based on the maximal flow which is delivered between two nodes a and d. Then, the network is in up state if all links function or only the link between b and c fails. If there is no path between a and d, the network is in down state. In the other cases, the network has partial performance. In this case, Ashrafi and Asadi [1] obtained the non-zero elements of the signature matrix of the network as follows

$$s_{1,2} = \frac{1}{5}, \quad s_{1,3} = \frac{7}{15}, \quad s_{1,4} = \frac{2}{15}, \quad s_{2,3} = \frac{2}{15}, \quad s_{2,4} = \frac{1}{15}.$$

(b) The network is in up state if the nodes in $\mathcal{T} = \{a, c, d\}$ are connected, is in partial performance if two nodes in $\mathcal{T}$ are connected and is in down state if all nodes in $\mathcal{T}$ are disconnected. In [1], the non-zero elements of signature matrix are obtained as

$$s_{2,4} = \frac{2}{15}, \quad s_{2,5} = \frac{1}{15}, \quad s_{3,4} = \frac{7}{15}, \quad s_{3,5} = \frac{7}{30}, \quad s_{4,5} = \frac{1}{10}.$$

(c) The network functions completely when all nodes are connected to each other. It has partial performance if the nodes are separated into two non-connected sets, and is failed completely if the nodes are separated into more than two non-connected sets of nodes. In [8], the non-zero elements of signature matrix are given as

$$s_{2,4} = \frac{1}{5}, \quad s_{3,4} = \frac{4}{5}.$$

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

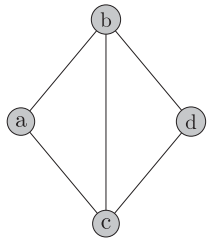


FIGURE **3.** Bridge network.

TABLE 1. $K(T_1, T)$ and $M(T_1, T)$ in Example 3.13

Model	$K(T : T_1)$	$K_U(T : T_1)$	$M(T_1, T)$	$M_U(T_1, T)$
(a)	1.22627	1.68565	0.23437	0.81179
(b)	0.89532	1.46166	0.35994	0.94780
(c)	0.55057	0.67563	0.49576	0.66714

Table 1 shows the K–L information and mutual information between the entrance times of network into partial performance T_1, and the down state T. The values of upper bounds are also presented for the K–L divergence ($K_U(T : T_1)$) and mutual information ($M_U(T_1, T)$) based on Theorems 3.5 and 3.10, respectively.

4. EXTENSION TO MULTI-STATE NETWORKS

In this section, we generalize the results of Section 3 to multi-state networks. Suppose that the network has n components with i.i.d. lifetimes $X_1, \ldots, X_n$ and r.v.s $X_{1:n}, \ldots, X_{n:n}$ denote the ordered component lifetimes. Assume further that the network has $m+1$ states $l = 0, 1, \ldots, m$ such that $l = m$ corresponds to the highest performance level of network. The network starts working at the time $t = 0$; it enters into state $m - i$ at random time T_i, $i = 1, \ldots, m$, and then moves to the next state. With $\mathbf{i} = (i_1, \ldots, i_m)$ and $\mathcal{I} = \{\mathbf{i} | 1 \le i_1 < \cdots < i_m \le n\}$, the joint survival function of $(T_1, \ldots, T_m)$ is shown to be represented as

$$P(T_1 > t_1, \ldots, T_m > t_m) = \sum_{\mathbf{i} \in \mathcal{I}} s_{i_1, \ldots, i_m} P(X_{i_1:n} > t_1, \ldots, X_{i_m:n} > t_m), \tag{4.1}$$

where $t_1 > 0, \ldots, t_m > 0$ and $s_{i_1, \ldots, i_m}$ is the $(i_1, \ldots, i_m)$th element of the m-dimensional signature matrix $\mathcal{S} = \left(s_{i_1, \ldots, i_m}\right)_{\mathbf{i} \in \mathcal{I}}$ in which the non-zero elements of $\mathcal{S}$ are defined as

$$s_{i_1, \ldots, i_m} = P(T_1 = X_{i_1:n}, \ldots, T_m = X_{i_m:n}), \quad \mathbf{i} \in \mathcal{I};$$

see [7].

The marginal survival function of T_j, $j = 1, \ldots, m$, is of the form (3.1) with signature vector $\mathbf{s}^{(j)} = (s_1^{(j)}, \ldots, s_n^{(j)})$ where

$$s_i^{(j)} = P(T_j = X_{i:n}) = \sum_{\mathbf{i}^{(j)} \in \mathcal{I}^{(j)}} s_{i_1, \ldots, i_m}, \quad j = 1, \ldots, m, \ i = 1, \ldots, n, \tag{4.2}$$

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

such that $\mathbf{i}^{(j)} = (i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_m)$ and

$$\mathcal{I}^{(j)} = \{\mathbf{i}^{(j)} | 1 \le i_1 < \cdots < i_{j-1} < i_{j+1} < \cdots < i_m \le n\}.$$

Using (4.1), the joint pdf of $(T_1, \ldots, T_m)$ is given as follows

$$f_{T_1,\ldots,T_m}(t_1, \ldots, t_m) = \sum_{\mathbf{i}\in\mathcal{I}} s_{i_1,\ldots,i_m} f_{i_1,\ldots,i_m:n}(t_1, \ldots, t_m), \tag{4.3}$$

where $f_{i_1,\ldots,i_m:n}$ is the joint pdf of $(X_{i_1:n}, \ldots, X_{i_m:n})$.

The marginal pdf of any subset of $\{T_1, \ldots, T_m\}$ can be calculated by integrating over the non-existing variables. Since $\mathcal{S} = (s_{i_1,\ldots,i_m})_{\mathbf{i}\in\mathcal{I}}$ is a probability matrix, this yields to a mixture representation similar to (4.3) with mixing probability matrix obtained by summing on the non-existing indices. For example, the joint pdf of $(T_1, \ldots, T_k)$, $k = 1, \ldots, m-1$, is given as

$$f_{T_1,\ldots,T_k}(t_1, \ldots, t_k) = \sum_{\mathbf{i}_k\in\mathcal{I}_k} s_{i_1,\ldots,i_k} f_{i_1,\ldots,i_k:n}(t_1, \ldots, t_k),$$

where $f_{i_1,\ldots,i_k:n}$ is the joint pdf of $(X_{i_1:n}, \ldots, X_{i_k:n})$, $\mathbf{i}_k = (i_1, \ldots, i_k)$, $\mathcal{I}_k = \{\mathbf{i}_k | 1 \le i_1 < \cdots < i_k \le n\}$ and

$$s_{i_1,\ldots,i_k} = P(T_1 = X_{i_1:n}, \ldots, T_k = X_{i_k:n}) = \sum_{\bar{\mathbf{i}}_k\in\bar{\mathcal{I}}_k} s_{i_1,\ldots,i_m}, \quad 1 \le i_1 < \cdots < i_k \le n,$$

such that $\bar{\mathbf{i}}_k = (i_{k+1}, \ldots, i_m)$, and $\bar{\mathcal{I}}_k = \{\bar{\mathbf{i}}_k | 1 \le i_{k+1} < \cdots < i_m \le n\}$. The marginal pdf of $(T_{k+1}, \ldots, T_m)$ is similarly followed as

$$f_{T_{k+1},\ldots,T_m}(t_{k+1}, \ldots, t_m) = \sum_{\bar{\mathbf{i}}_k\in\bar{\mathcal{I}}_k} s_{i_{k+1},\ldots,i_m} f_{i_{k+1},\ldots,i_m:n}(t_{k+1}, \ldots, t_m),$$

where $f_{i_{k+1},\ldots,i_m:n}$ denotes the joint pdf of $(X_{i_{k+1}:n}, \ldots, X_{i_m:n})$ and

$$\begin{aligned} s_{i_{k+1},\ldots,i_m} &= P(T_{k+1} = X_{i_{k+1}:n}, \ldots, T_m = X_{i_m:n}) \\ &= \sum_{\mathbf{i}_k\in\mathcal{I}_k} s_{i_1,\ldots,i_m}, \quad 1 \le i_{k+1} < \cdots < i_m \le n. \end{aligned}$$

Now some information measures for the network lifetimes, $T_1, \ldots, T_m$, are calculated.

For two random vectors $\mathbf{X}$ and $\mathbf{Y}$ with joint pdf's $f_{\mathbf{X}}$ and $f_{\mathbf{Y}}$ in R^m and R^n, respectively, the conditional entropy of $\mathbf{Y}$ given $\mathbf{X}$, denoted by $H(\mathbf{Y}|\mathbf{X})$, is defined as

$$H(\mathbf{Y}|\mathbf{X}) = \int_{R^m} \cdots \int f_{\mathbf{X}}(\mathbf{x}) H(\mathbf{Y}|\mathbf{X} = \mathbf{x}) d\mathbf{x}, \tag{4.4}$$

where $H(\mathbf{Y}|\mathbf{X} = \mathbf{x}) = -\int_{R^n} \cdots \int f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \log f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) d\mathbf{y}$.

From (4.3), the conditional pdf of $(T_{l+1}, \ldots, T_m)$ given $(T_1, \ldots, T_l)$, $l = 1, \ldots, m-1$, is written as

$$\begin{aligned} &f_{T_{l+1},\ldots,T_m|T_1,\ldots,T_l}(t_{l+1}, \ldots, t_m|t_1, \ldots, t_l) \\ &= \frac{\prod_{j=l+1}^{m} f(t_j) \sum_{\mathbf{i}\in\mathcal{I}} s_{i_1,\ldots,i_m} C^n_{i_1,\ldots,i_m} F^{i_1-1}(t_1) \bar{F}^{n-i_m}(t_m) \prod_{j=1}^{m-1} [F(t_{j+1}) - F(t_j)]^{i_{j+1}-i_j-1}}{\sum_{\mathbf{i}_l\in\mathcal{I}_l} s_{i_1,\ldots,i_l} C^n_{i_1,\ldots,i_l} F^{i_1-1}(t_1) \bar{F}^{n-i_l}(t_l) \prod_{j=1}^{l-1} [F(t_{j+1}) - F(t_j)]^{i_{j+1}-i_j-1}}. \end{aligned}$$

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

Therefore, using the integral transformations $V_j = F(T_j)$, $j = l+1, \ldots, m$, we get

$$\begin{aligned} &H(T_{l+1}, \ldots, T_m | t_1, \ldots, t_l) \\ &\quad = -\int_{F(t_l)}^{v_{l+2}} \cdots \int_{v_{m-1}}^{1} q(t_1, \ldots, t_l, v_{l+1}, \ldots, v_m) \\ &\qquad \times \log \left(\prod_{j=l+1}^{m} f(F^{-1}(v_j)) q(t_1, \ldots, t_l, v_{l+1}, \ldots, v_m) \right) dv_m \ldots dv_{l+1}, \end{aligned}$$

where

$$\begin{aligned} &q(t_1, \ldots, t_l, v_{l+1}, \ldots, v_m) \\ &\quad = \frac{\sum_{\mathbf{i} \in \mathcal{I}} s_{i_1, \ldots, i_m} C^n_{i_1, \ldots, i_m} F^{i_1 - 1}(t_1) a_{i_l, \ldots, i_m}(v_{l+1}, \ldots, v_m | t_l) \prod_{j=1}^{l-1} [F(t_{j+1}) - F(t_j)]^{i_{j+1} - i_j - 1}}{\sum_{\mathbf{i}_l \in \mathcal{I}_l} s_{i_1, \ldots, i_l} C^n_{i_1, \ldots, i_l} F^{i_1 - 1}(t_1) \bar{F}^{n - i_l}(t_l) \prod_{j=1}^{l-1} [F(t_{j+1}) - F(t_j)]^{i_{j+1} - i_j - 1}}, \end{aligned}$$

and

$$a_{i_l, \ldots, i_m}(v_{l+1}, \ldots, v_m | t_l) = [v_{l+1} - F(t_l)]^{i_{l+1} - i_l - 1} (1 - v_m)^{n - i_m} \prod_{j=l+1}^{m-1} (v_{j+1} - v_j)^{i_{j+1} - i_j - 1}.$$

Now, using the same argument as used in Theorem 3.1, we can show that:

$$\begin{aligned} H(T_{l+1}, \ldots, T_m | T_1, \ldots, T_l) &= H(V_1, \ldots, V_m) - H(V_1, \ldots, V_l) \\ &\quad - \sum_{j=l+1}^{m} \left\{ \sum_{i_j=1}^{n} s_{i_j}^{(j)} E[\log f(F^{-1}(W_{i_j:n}))] \right\} \\ &= H(V_1, \ldots, V_m | V_1, \ldots, V_l) \\ &\quad - \sum_{j=l+1}^{m} \left\{ \sum_{i_j=1}^{n} s_{i_j}^{(j)} E[\log f(F^{-1}(W_{i_j:n}))] \right\}, \end{aligned} \tag{4.5}$$

where $W_{i_j:n}$ has beta distribution with parameters $(i_j, n - i_j + 1)$ and

$$\begin{aligned} H(V_1, \ldots, V_r) = &- \sum_{\mathbf{i}_r \in \mathcal{I}_r} s_{i_1, \ldots, i_r} \int_0^1 \cdots \int_{v_{r-1}}^1 h_{i_1, \ldots, i_r}(v_1, \ldots, v_r) \\ &\times \log \sum_{\mathbf{i}_r \in \mathcal{I}_r} s_{i_1, \ldots, i_r} h_{i_1, \ldots, i_r}(v_1, \ldots, v_r) dv_r \ldots dv_1. \quad r = 2, \ldots, m. \end{aligned} \tag{4.6}$$

Remark 4.1: From (4.5), when the time that the network enters into the first state is known, the expected amount of information for predicting $(T_2, \ldots, T_m)$ can be measured by

$$H(T_2, \ldots, T_m | T_1) = H(V_1, \ldots, V_m) - H(V_1) - \sum_{j=2}^{m} \sum_{i_j=1}^{n} s_{i_j}^{(j)} E\big[\log f(F^{-1}(W_{i_j:n}))\big]. \tag{4.7}$$

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

Also, when $(T_1, \ldots, T_{m-1})$ is known, the expected amount of information needed to predict the failure time of network is calculated by

$$H(T_m|T_1, \ldots, T_{m-1}) = H(V_1, \ldots, V_m) - H(V_1, \ldots, V_{m-1}) - \sum_{i_m=1}^{n} s_{i_m}^{(m)} E\big[\log f(F^{-1}(W_{i_m:n}))\big].$$

The Shannon entropy (differential entropy) of joint random vector $\mathbf{X} = (X_1, \ldots, X_n)$ is defined as

$$H(\mathbf{X}) = -\int_{R^n} \cdots \int f_{\mathbf{X}}(\mathbf{x}) \log f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \tag{4.8}$$

where $f_{\mathbf{X}}$ is the joint pdf of $\mathbf{X}$. Hence, similar to Theorem 3.2, the joint entropy of $(T_1, \ldots, T_m)$ is obtained as

$$H(T_1, \ldots, T_m) = H(V_1, \ldots, V_m) - \sum_{j=1}^{m} \sum_{i_j=1}^{n} s_{i_j}^{(j)} E\big[\log f(F^{-1}(W_{i_j:n}))\big],$$

where $H(V_1, \ldots, V_m)$ is given in relation (4.6).

A lower bound for the joint entropy of $(T_1, \ldots, T_m)$ can be written as

$$H_L(T_1, \ldots, T_m) = \sum_{\mathbf{i} \in \mathcal{I}} s_{i_1, \ldots, i_m} H(W_{i_1:n}, \ldots, W_{i_m:n}) - \sum_{j=1}^{m} \sum_{i_j=1}^{n} s_{i_j}^{(j)} E\big[\log f(F^{-1}(W_{i_j:n}))\big],$$

where, using the fact that the joint pdf of $(W_{i_1:n}, \ldots, W_{i_m:n})$ is $h_{i_1, \ldots, i_m}(w_{i_1}, \ldots, w_{i_m})$, $H(W_{i_1:n}, \ldots, W_{i_m:n})$ is given as follows:

$$\begin{aligned} H(W_{i_1:n}, \ldots, W_{i_m:n}) = & -\log C^n_{i_1, \ldots, i_m} - (i_1 - 1)\psi(i_1) - (n - i_m)\psi(n - i_m + 1) \\ & - \sum_{j=1}^{m-1} (i_{j+1} - i_j - 1)\psi(i_{j+1} - i_j) \\ & + \left[n - m + i_1 - i_m + \sum_{j=1}^{m-1} (i_{j+1} - i_j) \right] \psi(n+1). \end{aligned}$$

The K-L of two absolutely continuous random vectors $\mathbf{X}$ and $\mathbf{Y}$ in R^n, respectively, with joint pdf's $f_{\mathbf{X}}$ and $f_{\mathbf{Y}}$ is defined as

$$K(\mathbf{X} : \mathbf{Y}) = \int_{R^n} f_{\mathbf{X}}(\mathbf{x}) \log \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{Y}}(\mathbf{y})} d\mathbf{x} \geq 0. \tag{4.9}$$

Suppose that m is an even number and $m^* = m/2$. We obtain the K-L information between the two vectors $(T_1, \ldots, T_{m^*})$ and $(T_{m^*+1}, \ldots, T_m)$, to determine the distance between the densities of these vectors. To this, the joint pdf's of $(T_1, \ldots, T_{m^*})$ and

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

$(T_{m^*+1},\dots,T_m)$ can be written, respectively, as follow

$$f_{T_1,\dots,T_{m^*}}(t_1,\dots,t_{m^*}) = \sum_{\mathbf{i}_{m^*}\in\mathcal{I}_{m^*}} s_{i_1,\dots,i_{m^*}} C^n_{i_1,\dots,i_{m^*}} F^{i_1-1}(t_1)\bar{F}^{n-i_{m^*}}(t_{m^*}) \prod_{j=1}^{m^*} f(t_j) \times \prod_{j=1}^{m^*-1} [F(t_{j+1}) - F(t_j)]^{i_{j+1}-i_j-1}, \tag{4.10}$$

and

$$f_{T_{m^*+1},\dots,T_m}(t_{m^*+1},\dots,t_m) = \sum_{\bar{\mathbf{i}}_{m^*}\in\bar{\mathcal{I}}_{m^*}} s_{i_{m^*+1},\dots,i_m} C^n_{i_{m^*+1},\dots,i_m} F^{i_{m^*+1}-1}(t_{i_{m^*+1}})\bar{F}^{n-i_m}(t_m) \times \prod_{j=m^*+1}^{m} f(t_j) \prod_{j=m^*+1}^{m-1} [F(t_{j+1}) - F(t_j)]^{i_{j+1}-i_j-1}. \tag{4.11}$$

Therefore, using transformations $V_j = F(T_j)$, $j = 1,\dots,m^*$, and relations (4.10) and (4.11) we have

$$K(T_1,\dots,T_{m^*} : T_{m^*+1},\dots,T_m) = \sum_{\mathbf{i}_{m^*}\in\mathcal{I}_{m^*}} s_{i_1,\dots,i_{m^*}} \int_0^1 \dots \int_0^{v_2} h_{i_1,\dots,i_{m^*}}(v_1,\dots,v_{m^*}) \times \log \frac{\sum_{\mathbf{i}_{m^*}\in\mathcal{I}_{m^*}} s_{i_1,\dots,i_{m^*}} h_{i_1,\dots,i_{m^*}}(v_1,\dots,v_{m^*})}{\sum_{\bar{\mathbf{i}}_{m^*}\in\bar{\mathcal{I}}_{m^*}} s_{i_{m^*+1},\dots,i_m} h_{i_{m^*+1},\dots,i_m}(v_1,\dots,v_{m^*})} dv_1\dots dv_{m^*}. \tag{4.12}$$

In particular, the K–L information between two arbitrary lifetimes T_r and T_k is given as

$$K(T_r : T_k) = \int_0^\infty f_{T_r}(t_r) \log \frac{f_{T_r}(t_r)}{f_{T_k}(t_r)} dt_r = \sum_{i_r=1}^n s^{(r)}_{i_r} \int_0^1 h_{i_r}(v_r) \log \frac{\sum_{i_r=1}^n s^{(r)}_{i_r} h_{i_r}(v_r)}{\sum_{i_k=1}^n s^{(k)}_{i_k} h_{i_k}(v_r)} dv_r, \quad r\neq k, \quad r,k = 1,\dots,m. \tag{4.13}$$

It is seen that, relations (4.12) and (4.13) depend only on the network structure and are free of the distribution of components lifetime.

Note that, using (4.1) the joint reliability function of (T_r, T_k), $r \neq k$, $r,k = 1,\dots,m$ can be written as

$$P(T_r > t_r, T_k > t_k) = \sum_{i_r,i_k} s^{(rk)}_{i_r,i_k} P(X_{i_r:n} > t_r, X_{i_k:n} > t_k), \quad t_r > 0, t_k > 0, \tag{4.14}$$

where $s^{(rk)}_{i_r,i_k} = P(T_r = X_{i_r:n}, T_k = X_{i_k:n})$ is the (i_r,i_k)th element of a two-dimensional matrix $\mathcal{S}^{(rk)} = \{s^{(rk)}_{i_r,i_k}\}_{n\times n}$. Then $s^{(rk)}_{i_r,i_k} = 0$ for $i_r \geq i_k$. Based on relation (4.14) and probability integral transformation, the mutual information between two lifetimes T_r and T_k,

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

$r \neq k$, $r, k = 1, \ldots, m$, can be written as

$$M(T_r, T_k) = \int_0^\infty \int_0^{t_k} f_{T_r,T_k}(t_r, t_k) \log \frac{f_{T_r,T_k}(t_r, t_k)}{f_{T_r}(t_r) f_{T_k}(t_k)} dt_r dt_k$$
$$= \sum_{i_r,i_k} s^{(rk)}_{i_r,i_k} \int_0^1 \int_0^{v_k} h_{i_r,i_k}(v_r, v_k) \log \frac{\sum_{i_r,i_k} s^{(rk)}_{i_r,i_k} h_{i_r,i_k}(v_r, v_k)}{\sum_{i_k=1}^n s^{(k)}_{i_k} h_{i_k}(v_k) \sum_{i_r=1}^n s^{(r)}_{i_r} h_{i_r}(v_r)} dv_r dv_k.$$

The mutual information $M(T_r, T_k)$ depends only on the network structure and is free of the stochastic mechanism of components failure.

Acknowledgments

The authors would like to thank two reviewers for their constructive comments which led to improving the exposition of this paper. Asadi's research work was carried out in IPM Isfahan branch and was in part supported by a grant from IPM (No. 96620411).

References

1. Ashrafi, S. & Asadi, M. (2014). Dynamic reliability modeling of three-state networks. *Journal of Applied Probability* 51(4): 999–1020.
2. Ashrafi, S. & Asadi, M. (2015). On the stochastic and dependence properties of the three-state systems. *Metrika* 78(3): 261–281.
3. Asadi, M., Ebrahimi, N., Soofi, E.S., & Zohrevand, Y. (2016). Jensen-Shannon information of the coherent system lifetime. *Reliability Engineering & System Safety* 156: 244–255.
4. Ebrahimi, N., Soofi, E.S., & Soyer, R. (2013). When are observed failures more informative than observed survivals? *Naval Research Logistics* 60(2): 102–110.
5. Ebrahimi, N., Soofi, E.S., & Zahedi, H. (2004). Information properties of order statistics and spacings. *IEEE Transactions on Information Theory* 50: 177–183.
6. Gertsbakh, I. & Shpungin, Y. (2011). *Network Reliability and Resilience. Springer Briefs in Electrical and Computer Engineering*. Berlin, Germany: Springer-Verlag.
7. Gertsbakh, I. & Shpungin, Y. (2012). Stochastic models of network survivability. *Quality Technology & Quantitative Management* 9(1): 45–58.
8. Gertsbakh, I., Shpungin, Y., & Spizzichino, F. (2012). Two-dimensional signatures. *Journal of Applied Probability* 49(2): 416–429.
9. Kochar, S., Mukerjee, H., & Samaniego, F.J. (1999). The 'signature' of a coherent system and its application to comparison among systems. *Naval Research Logistics* 46: 507–523.
10. Kullback, S. & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics* 22(1): 79–86.
11. Navarro, J., Balakrishnan, N., & Samaniego, F.J. (2008). Mixture reptesentations of residual lifetimes of used systems. *Journal of Applied Probability* 45: 1097–1112.
12. Navarro, J., Samaniego, F.J., & Balakrishnan, N. (2013). Mixture representations for the joint distribution of two coherent systems with shared components. *Advances in Applied Probability* 45(4): 1011–1027.
13. Parzen, E. (1979). Nonparametric statistical data modelling. *Journal of American Statistical Association* 74: 105–122.
14. Samaniego, F.J. (1985). On closure of the IFR class under formation of coherent systems. *IEEE Transactions on Reliability* 34: 69–72.
15. Samaniego, F.J. (2007). *System signatures & their applications in reliability engineering*. New York, Berlin: Springer.
16. Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic orders*. New York: Springer.
17. Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423.
18. Toomaj, A. & Doostparast, M. (2014). A note on signature-based expressions for the entropy of mixed r-out-of-n systems. *Naval Research Logistics* 61(3): 202–206.
19. Toomaj, A. & Doostparast, M. (2016). On the Kullback Leibler information for mixed systems. *International Journal of Systems Science* 47(10): 2458–2465.

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press

20. Toomaj, A., Sunoj, S.M., & Navarro, J. (2017). Some properties of the cumulative residual entropy of coherent and mixed systems. *Journal of Applied Probability* 54(2): 379–393.
21. Mahmoudi, M. & Asadi, M. (2011). The dynamic signature of coherent systems. *IEEE Transactions on Reliability* 60(4): 817–822.
22. Marichal, J.L., Mathonet, P., Navarro, J., & Paroissin, C. (2017). Joint signature of two or more systems with applications to multistate systems made up of two-state components. *European Journal of Operational Research* 263: 559–570.
23. Wilks, S.S. (1962). *Mathematical statistics*. New York: Wiley.
24. Zarezadeh, S., Ashrafi, S., & Asadi, M. (2016). A shock model based approach to network reliability. *IEEE Transactions on Reliability* 65(2): 992–1000.
25. Zarezadeh, S. & Asadi, M. (2010). Results on residual Rényi entropy of order statistics and record values. *Information Sciences* 180(21): 4195–4206.
26. Zarezadeh, S., Mohammadi, L., & Balakrishnan, N. (2018). On the joint signature of several coherent systems with some shared components. *European Journal of Operational Research* 264(3): 1092–1100.

https://doi.org/10.1017/S0269964818000177 Published online by Cambridge University Press