THE POSTERIOR DISTRIBUTION OF THE PARAMETERS OF COMPONENT LIFETIMES BASED ON AUTOPSY DATA IN A SHOCK MODEL

JØRUND GÅSEMYR and BENT NATVIG
University of Oslo

ABSTRACT. In this paper we consider a binary, monotone system whose component states are dependent through the possible occurence of independent common shocks, i.e. shocks that destroy several components at once. The individual failure of a component is also thought of as a shock. Such systems can be used to model common cause failures in reliability analysis. The system may be a technological one, or a human being. It is observed until it fails or dies. At this instant, the set of failed components and the failure time of the system are noted. The failure times of the components are not known. These are the socalled autopsy data of the system. For the case of independent components, i.e. no common shocks, Meilijson (1981), Nowik (1990), Antoine et al. (1993) and Gåsemyr (1995) discuss the corresponding identifiability problem, i.e. whether the component life distributions can be determined from the distribution of the observed data. Assuming a model where autopsy data is known to be enough for identifiability, Meilijson (1994) goes beyond the identifiability question and into maximum likelihood estimation of the parameters of the component lifetime distributions based on empirical autopsy data from a sample of several systems. He also considers lifemonitoring of some components and conditional lifemonitoring of some other. In Gåsemyr & Natvig (1994) a corresponding Bayesian approach is presented. Due to prior information one advantage is that the identifiability problem represents no obstacle. Here, the first part of our previous paper is extended to the shock model, the motivation being that the autopsy model is of special importance when components cannot be tested separately because it is difficult to reproduce the conditions prevailing in the functioning system.

Key words: fatal set, critical set, critical shock, natural conjugate prior, mixture of products of gamma distributions

1. Introduction and basic model

In this section the basic model is introduced and motivated whereas the main results are given in Section 2. Some aspects of the computation of the distribution of the autopsy data are considered in Appendix 1. In Section 3 a parallel system of two components, also subjected to a common shock, is treated in depth. Some technical details in the deductions are left to Appendix 2.

Consider a binary, monotone system (E, ϕ), where $E = \{1, \cdots, n\}$ is the set of components and ϕ is the structure function describing the state of the system in terms of the binary states of the components. The system may be a technological one, or a human being. It is

observed until it fails or dies. At this instant, the set of failed components and the failure time of the system are noted. The failure times of the components are not known. These are the socalled autopsy data of the system.

For the case of independent components Meilijson (1981), Nowik (1990), Antoine et al. (1993) and Gåsemyr (1995) discuss the corresponding identifiability problem, i.e. whether the component life distributions can be determined from the observed data. For a very readable presentation of these efforts we recommend to start with the paper Antoine et al. (1993). They stress the importance of the autopsy model in situations in which components cannot be tested separately for instance because it is difficult to reproduce the conditions prevailing in the functioning system. The most obvious such situation is when components are actually dependent, which, however, is outside the scope of the mentioned papers. In Crowder (1994) the socalled identifiability crises in competing risks is reviewed. Its origins can be traced back to Bernoulli (1760) who attempted to disentangle the risks of dying from smallpox and other causes. Although much of this literature covers dependent components it is of less interest in reliability since competing risks just correspond to a series system.

Hence what is of real interest is to treat the autopsy model for an arbitrary system of dependent components. In the present paper we consider this problem in the case where component states are dependent through the possible occurence of independent common shocks, i.e. shocks that destroy several components at once. The individual failure of a component is also thought of as a shock. This shock model was introduced by Boyles & Samaniego (1984) and some aspects of reliability analysis in such a model are considered in Gåsemyr & Natvig (1995a). A special case is the multivariate exponential distribution introduced in Marshall & Olkin (1967). Using expert opinions in Bayesian prediction of component lifetimes in this model is treated in Gåsemyr & Natvig (1995b). Our present work is an attempt to get as much information as possible from a failing system of dependent components. This parallels and should be combined with the latter paper.

We now present the shock model in more detail. There are n individual shocks numbered $l = 1, \cdots, n$. In addition there are p common shocks numbered $n+1, \cdots, n+p$. Introduce $(l = 1, \cdots, n+p)$

$$D_l = \quad \text{the set of components destroyed by the lth shock}$$

Obviously, $D_l = \{l\}, l = 1, \cdots, n$. More generally for a set of shocks $B \subset \{1, \cdots, n+p\}$ let

$$D_B = \bigcup_{l \in B} D_l = \quad \text{the set of components destroyed by the set of shocks } B$$

Furthermore, denote the time until the lth shock by $V_l, l = 1, \cdots, n+p$. The variables V_l are assumed independent given a parameter vector $\underline{\theta}$, having absolutely continuous distributions with distribution function $F_l(t)$, survival function $\bar{F}_l(t) = 1 - F_l(t)$, p.d.f. $f_l(t)$ and failure rate $\lambda_l(t) = f_l(t)/\bar{F}_l(t), l = 1, \cdots, n+p$.

On the other hand introduce $(i = 1, \cdots, n)$

$$E_i = \{l | i \epsilon D_l\} = \quad \text{the set of shocks that destroys the ith component}$$

More generally for a set of components $A \subset E$ let

$$E_A = \bigcup_{i \epsilon A} E_i = \quad \text{the set of shocks that destroys components in the set } A$$

Now let T_i be the lifetime of the ith component and T the lifetime of the system. Then obviously $(i = 1, \cdots, n)$

$$T_i = \min\{V_l, l \epsilon E_i\} \tag{1.1}$$

From Barlow & Proschan (1975) it follows that for instance

$$T = \max_{j=1,\cdots,r} \min_{i \epsilon P_j} T_i,$$

where $P_1, \cdots, P_r$ are the minimal path sets of the system (E, ϕ). These are minimal component sets each of which ensures the functioning of the system if all components in the set function.

2. Main results

We start by giving some key definitions. Remember that a (minimal) cut set, K, of the system (E, ϕ) is a (minimal) set of components which ensures the failure of the system if all components in K have failed.

Definition 1

Let A be a cut set of the system (E, ϕ) and let $H \subset A$. H is a *critical set* for A if the functioning of all components in $H \cup A^c$ ensures the functioning of the system even if all components in $A - H$ have failed, i.e. $H \cap K \neq \emptyset$ for all minimal cut sets K of the system (E, ϕ) such that $K \subset A$.

Since A is a cut set, obviously the failure of all components in a critical set H for A leads to system failure if all components in $A - H$ have already failed. Note that a critical set H for A is not necessarily a minimal set.

Definition 2

A is a *fatal set* for the system (E, ϕ) if and only if there exists a shock $\{l\} \epsilon \{1, \cdots, n + p\}$ such that D_l is a critical set for A. In this case $\{l\}$ is a *critical shock* for A. The set of critical shocks for A is denoted by C_A.

Let

$$D = \quad \text{the set of failed components } = \{i \epsilon E | T_i \leq T\}$$

Obviously, A is a fatal set if and only if $P(D = A) > 0$. Let

$$\mathcal{A} = \{\text{fatal sets}\} = \{A \subset E | P(D = A) > 0\}$$
$$= \{A_1, \cdots, A_m\}$$

(T, D) are the autopsy data of the system. Its distribution is given by

$$G_A(t) = P(T \leq t, D = A)$$

3

with density function

$$g_A(t) = \frac{d}{dt}G_A(t)$$

The latter can be considered as a likelihood function on the space $R^+ \times \{1, 2, \cdots, m\}$ with respect to the measure

$$\mu = \text{Lebesgue measure} \times \text{counting measure}$$

For the case of no common shocks the following result, essentially given in Meilijson (1981), is straightforward

$$g_A(t) = \sum_{j \in C_A} \prod_{l \in A^c \cup \{j\}} \bar{F}_l(t) \prod_{l \in A-j} F_l(t) \lambda_j(t) \tag{2.1}$$

The main task of this section is to generalize (2.1) to the common shock case. To help the reader we will all the way refer to the system in Figure 2.1.
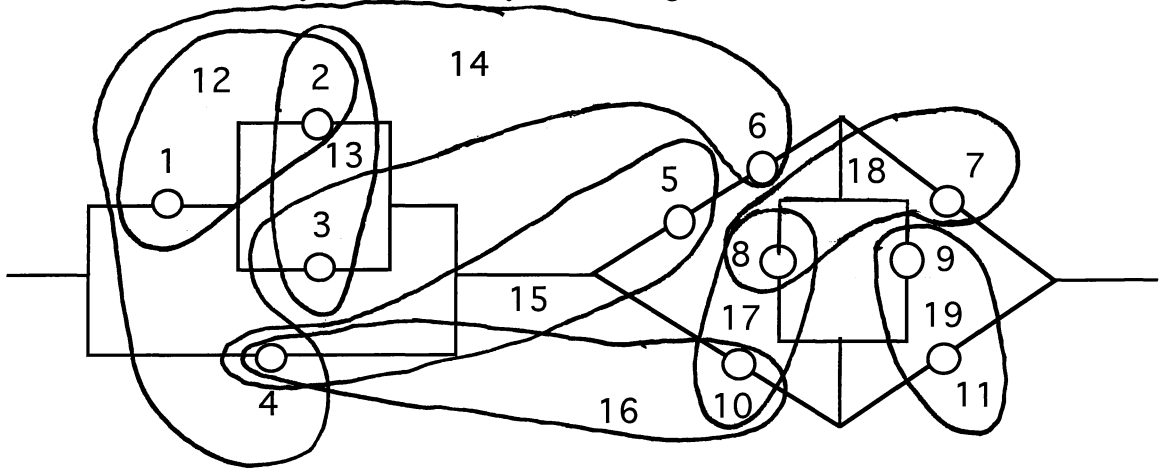

Figure 2.1 Reference system subjected to 8 common shocks.

In addition to the 11 individual shocks corresponding to the 11 components, we have 8 common shocks. These are defined by D_l being respectively $\{1, 2\}, \{2, 3\}, \{1, 2, 4, 6\}, \{4, 5\}, \{4, 10\}, \{8, 10\}, \{7, 8\}, \{9, 11\}$ for $l = 12, 13, 14, 15, 16, 17, 18, 19$. We will consider the case $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}, A^c = \{10, 11\}$. The set of critical shocks for A is $C_A = \{4, 12, 14, 15\}$.

Before proceeding further we have to explain some set manipulations. Consider a component $l \in E$ and let C be a set of shocks. Assume

$$E_l - C = \{l\}$$

This means that the lth component cannot be destroyed by common shocks outside the set C. In our reference system we for instance have $E_9 - \{19\} = \{9\}$.

Consider a set of components A and a set of shocks C. Assume

$$D_{E_A-C} = A$$

Note that $E_A - C$ is the set of shocks, except for the ones in C, that destroys components in the set A. D_{E_A-C} is the set of components destroyed by the set of shocks $E_A - C$. If

4

$D_{E_A - C} = A$, it means that the components in the set A cannot be destroyed by common shocks, also destroying components in A^c, outside the set C. In our reference system we for instance have

$$D_{E_{\{7,8,9\}} - \{17,19\}} = \{7, 8, 9\}$$

Now let

$$C_1 = E_{A^c} = \quad \text{set of shocks that cannot have occured before}$$
$$\text{time } t \text{ according to a first sorting level}$$

Let furthermore A_1 be a maximal set of components with respect to the following properties

i) $A_1 \subset A - \bigcup_{j \in C_A} D_j$

ii) $D_{E_{A_1} - C_1} = A_1$

A_1 is a maximal set of components that must fail before time t according to the first sorting level.

$$F_1 = E_{A_1} - C_1$$
$$B_1 = \{l \epsilon A_1 | E_l - C_1 = \{l\}\}$$
$$\mathcal{B}_1 = \text{Minimal sets in } \{B \subset F_1 - B_1 | D_B = A_1 - B_1\}$$

In our reference system we have

$$C_1 = \{10, 11, 16, 17, 19\}$$
$$A_1 = \{7, 8, 9\}$$
$$F_1 = \{7, 8, 9, 18\}$$
$$B_1 = \{9\}$$
$$\mathcal{B}_1 = \{\{18\}, \{7, 8\}\}$$

The contribution to $g_A(t)$ from the first sorting level is

$$\prod_{l \epsilon C_1} \bar{F}_l(t) \prod_{l \epsilon B_1} F_l(t) P \left[\bigcup_{B \epsilon \mathcal{B}_1} (V_l \leq t \; \forall l \epsilon B) \right] \tag{2.2}$$

Note that the last factor is given in a general way. In order not to derail the reader the corresponding computational aspects are treated in Appendix 1 focusing especially on the technique of recursive disjoint products inspired by Abraham (1979), Ball & Provan (1988) and Locks (1980, 1982). In our reference system this gives the factor

$$F_{18}(t) + \bar{F}_{18}(t) F_7(t) F_8(t)$$

We now turn to the second level of sorting of shocks. Let

$\mathcal{H}(A) = \{H | (H \text{ is a critical set for } A) \cap (\exists j \epsilon C_A \ni H \subset D_j\}$

$C(H) = E_H - C_1 = $ set of shocks that cannot have occured before time t according to the second sorting level

$A(H) = A - A_1 \cup H = $ set of components that must fail before time t according to the second sorting level

5

$$F(H) = E_{A(H)} - C_1 \cup C(H)$$

$$B(H) = \{l\epsilon A(H)|E_l - C_1 \cup C(H) = \{l\}\}$$

$$\mathcal{B}(H) = \text{Minimal sets in } \{B \subset F(H) - B(H)|D_B = A(H) - B(H)\}$$

In our reference system we have

$$\mathcal{H}(A) = \{\{1,2\}, \{1,4\}, \{1,2,4\}, \{1,2,6\}, \{1,4,6\}, \{1,2,4,6\},$$
$$\{2,4\}, \{2,4,6\}, \{4\}, \{4,5\}, \{4,6\}\},$$

and for instance

$$C(\{4\}) = \{4,14,15\}$$

$$A(\{4\}) = \{1,2,3,5,6\}$$

$$F(\{4\}) = \{1,2,3,5,6,12,13\}$$

$$B(\{4\}) = \{5,6\}$$

$$\mathcal{B}(\{4\}) = \{\{12,13\}, \{1,13\}, \{3,12\}, \{1,2,3\}\}$$

The contribution to $g_A(t)$ from the second sorting level is

$$\sum_{H\epsilon\mathcal{H}(A)} \prod_{l\epsilon C(H)} \bar{F}_l(t) \prod_{l\epsilon B(H)} F_l(t) P\left[\bigcup_{B\epsilon\mathcal{B}(H)} (V_l \le t \,\forall l\epsilon B)\right] \tag{2.3}$$

Note again the computational aspects associated with the last factor. In our reference system the efficient calculation of the factor corresponding to for instance $\mathcal{B}(\{4\})$ is

$$F_{12}(t)F_{13}(t) + \bar{F}_{12}(t)F_1(t)F_{13}(t) + \bar{F}_{13}(t)F_3(t)F_{12}(t) + \bar{F}_{12}(t)\bar{F}_{13}(t)F_1(t)F_2(t)F_3(t)$$

On the third and final level, where the critical set H is given, we just have to sum up the failure rates of the critical shocks that can destroy all components in H.

By combining this with (2.2) and (2.3) we end up with

$$g_A(t) = \prod_{l\epsilon C_1} \bar{F}_l(t) \prod_{l\epsilon B_1} F_l(t) \, P\left[\bigcup_{B\epsilon\mathcal{B}_1} (V_l \le t \,\forall l\epsilon B)\right]$$
$$\times \sum_{H\epsilon\mathcal{H}(A)} \prod_{l\epsilon C(H)} \bar{F}_l(t) \prod_{l\epsilon B(H)} F_l(t) \, P\left[\bigcup_{B\epsilon\mathcal{B}(H)} (V_l \le t \,\forall l\epsilon B)\right] \tag{2.4}$$
$$\times \sum_{\{j\epsilon C_A|H\subset D_j\}} \lambda_j(t)$$

(2.4) can be given in a more efficient way by introducing

$$\mathcal{L} = \{L \subset \{1,\cdots,n+p\}|\exists H\epsilon\mathcal{H}(A) \ni F(H) - B(H) = L\}$$
$$\mathcal{B}(L) = \text{Minimal sets in } \{B \subset L|D_B = D_L\}$$
$$\mathcal{H}(A,L) = \{H\epsilon\mathcal{H}(A)|F(H) - B(H) = L\}$$

We then get by interchanging the order of summation

$$g_A(t) = \prod_{l \in C_1} \bar{F}_l(t) \prod_{l \in B_1} F_l(t) P\left[\bigcup_{B \in \mathcal{B}_1} (V_l \le t \ \forall l \in B)\right]$$

$$\times \sum_{L \in \mathcal{L}} P\left[\bigcup_{B \in \mathcal{B}(L)} (V_l \le t \ \forall l \in B)\right] \tag{2.5}$$

$$\times \sum_{H \in \mathcal{H}(A,L)} \prod_{l \in C(H)} \bar{F}_l(t) \prod_{l \in B(H)} F_l(t) \sum_{\{j \in C_A | H \subset D_j\}} \lambda_j(t)$$

A further simplification of this expression is considered at the end of Appendix 1.

Applying this on our reference system, we end up with the following expression which is surprisingly simple.

$$g_{\{1,2,3,4,5,6,7,8,9\}}(t) = \bar{F}_{10}(t)\bar{F}_{11}(t)\bar{F}_{16}(t)\bar{F}_{17}(t)\bar{F}_{19}(t)F_9(t)$$

$$\times [F_{18}(t) + \bar{F}_{18}(t)F_7(t)F_8(t)]$$

$$\times \{[F_{12}(t)F_{13}(t) + \bar{F}_{12}(t)F_1(t)F_{13}(t) + \bar{F}_{13}(t)F_3(t)F_{12}(t) + \bar{F}_{12}(t)\bar{F}_{13}(t)F_1(t)F_2(t)F_3(t)]$$

$$\times \bar{F}_4(t)\bar{F}_{14}(t)\bar{F}_{15}(t)[F_5(t)F_6(t)(\lambda_4(t) + \lambda_{14}(t) + \lambda_{15}(t))$$

$$+ \bar{F}_6(t)F_5(t)\lambda_{14}(t) + \bar{F}_5(t)F_6(t)\lambda_{15}(t)]$$

$$+ [F_{15}(t) + \bar{F}_{15}(t)F_4(t)F_5(t)]\bar{F}_1(t)\bar{F}_2(t)\bar{F}_{12}(t)\bar{F}_{13}(t)\bar{F}_{14}(t)F_3(t)$$

$$\times [F_6(t)(\lambda_{12}(t) + \lambda_{14}(t)) + \bar{F}_6(t)\lambda_{14}(t)]$$

$$+ [F_{13}(t) + \bar{F}_{13}(t)F_2(t)F_3(t)]\bar{F}_1(t)\bar{F}_4(t)\bar{F}_{12}(t)\bar{F}_{14}(t)\bar{F}_{15}(t)F_5(t)\lambda_{14}(t)$$

$$+ \bar{F}_2(t)\bar{F}_4(t)\bar{F}_{12}(t)\bar{F}_{13}(t)\bar{F}_{14}(t)\bar{F}_{15}(t)F_3(t)F_5(t)\lambda_{14}(t)\}$$

$$\tag{2.6}$$

Assuming a model where autopsy data is known to be enough for identifiability, Meilijson (1994) goes beyond the identifiability question and into maximum likelihood estimation of the parameter vector $\underline{\theta}$ based on empirical autopsy data from a sample of several systems. Here a corresponding Bayesian approach is indicated. Let the prior distribution of $\underline{\theta}$ be $\pi(\underline{\theta})$. Then the posterior distribution of $\underline{\theta}$ given the autopsy data $(T = t, D = A)$ is obviously

$$\pi(\underline{\theta}|T = t, D = A) \propto g_A(t)\pi(\underline{\theta}) \tag{2.7}$$

The posterior distribution of $\underline{\theta}$ gives through (1.1) the basis for Bayesian inference on component lifetimes. A specific parameter may for instance be estimated by the expectation in its posterior marginal distribution. Taking prior knowledge into account this approach is especially suitable in reliability where data often are scarce and asymptotic properties of estimators are of less help.

Now assume that V_l is exponentially distributed with failure rate $\theta_l, l = 1, \cdots, n + p$. Then $(T_1, \cdots, T_n)$ has a Marshall-Olkin multivariate exponential distribution. We have $\underline{\theta} = (\theta_1, \cdots, \theta_{n+p})$. Assume furthermore the prior distributions of θ_l to be independent and gamma with shape parameter a_l and scale parameter $b_l, l = 1, \cdots, n + p$. Denote the corresponding p.d.f. $g(\theta_l; a_l, b_l)$. For the case of no common shocks it is shown in Gåsemyr & Natvig (1994) that $\pi(\underline{\theta}|T = t, D = A)$ is a mixture of products of gamma distributions. From (2.5) and (2.7) it follows that this is true also in our general case.

7

To deal with autopsy data from a sample of several systems we indicate how the updating works when autopsy data from a single system arrives. Assume the prior distribution of $\underline{\theta}$ to be given as a mixture of products of gamma distributions; i.e.

$$\pi(\underline{\theta}) = \sum_{k=1}^{K} w_k \prod_{l=1}^{n+p} g(\theta_l; a_{kl}, b_{kl}),$$

where some of the w_k's may be negative. Then as in Gåsemyr & Natvig (1994) it follows from (2.5) and (2.7) that the posterior distribution is again a mixture of products of gamma distributions. Hence this class of distributions, when properly defined, is the natural conjugate prior for $\underline{\theta}$ with respect to our multivariate exponential autopsy model. This seems to be a completely new generalization of the fact that the gamma distribution is the natural conjugate prior for the failure rate in an exponential model.

3. A parallel system of two components subjected to a common shock

Note that with the Bayesian approach the identifiability problem represents no obstacle. To illustrate this in detail we now consider a parallel system of two components subjected to a common shock. From the references in Section 1 it is well known that the lifetime distributions of the two components are unidentifiable even under independence. This is obvious since, under the autopsy model, one in effect observes only the system failure time, which has the distribution function $F_1(t)F_2(t)$, from which it is impossible to single out $F_1(t)$ and $F_2(t)$.

We now make the same assumptions as at the end of Section 2 by restricting to the Marshall-Olkin multivariate exponential distribution. $\underline{\theta} = (\theta_1, \theta_2, \theta_3)$ and the prior distributions of θ_l are assumed independent and gamma distributed with shape parameter a_l and scale parameter $b_l, l = 1, 2, 3$. Obviously $T = \max(T_1, T_2)$ and the only fatal set is $A = \{1, 2\}$. Introduce

$$B_i(t) = \{V_i = t, V_j \le t, V_3 > t\} \quad i = 1, 2; j \ne i$$
$$B_{i+2}(t) = \{V_i > t, V_j \le t, V_3 = t\} \quad i = 1, 2; j \ne i$$
$$B_5(t) = \{V_1 > t, V_2 > t, V_3 = t\}$$
$$p_i(t) = P(B_i(t)|T = t), \ i = 1, 2, 3, 4, 5$$

Then

$$\{T = t\} = \bigcup_{i=1}^{5} B_i(t),$$

where the events $B_i(t), i = 1, 2, 3, 4, 5$ are a.s. disjoint. The posterior distribution of $\underline{\theta}$, given the autopsy data, $T = t$, can now be written as

$$\pi(\underline{\theta}|T = t) = \sum_{i=1}^{5} p_i(t)\pi(\underline{\theta}|B_i(t)) \tag{3.1}$$

From (2.4) noting that $\mathcal{H}(\{1, 2\}) = \{\{1\}, \{2\}, \{1, 2\}\}$, we get

$$g_{\{1,2\}}(t) = e^{-(\theta_1+\theta_3)t}(1 - e^{-\theta_2 t})(\theta_1 + \theta_3) + e^{-(\theta_2+\theta_3)t}(1 - e^{-\theta_1 t})(\theta_2 + \theta_3) + e^{-(\theta_1+\theta_2+\theta_3)t}\theta_3 \tag{3.2}$$

By applying (2.7) and (3.2) this gives after rearranging

$$\pi(\underline{\theta}|T = t) \propto [\theta_1 e^{-(\theta_1+\theta_3)t}(1 - e^{-\theta_2 t})$$
$$+ \theta_2 e^{-(\theta_2+\theta_3)t}(1 - e^{-\theta_1 t}) + \theta_3 e^{-(\theta_1+\theta_3)t}(1 - e^{-\theta_2 t})$$
$$+ \theta_3 e^{-(\theta_2+\theta_3)t}(1 - e^{-\theta_1 t}) + \theta_3 e^{-(\theta_1+\theta_2+\theta_3)t}]$$
$$\times \prod_{i=1}^{3} b_i^{a_i} \theta_i^{a_i-1} e^{-b_i\theta_i}/\Gamma(a_i)$$
$$= (a_1/(b_1 + t))(b_1/(b_1 + t))^{a_1}(b_3/(b_3 + t))^{a_3}g(\theta_1; a_1 + 1, b_1 + t)g(\theta_3; a_3, b_3 + t)$$
$$\times [g(\theta_2; a_2, b_2) - (b_2/(b_2 + t))^{a_2}g(\theta_2; a_2, b_2 + t)]$$
$$+ (a_2/(b_2 + t))(b_2/(b_2 + t))^{a_2}(b_3/(b_3 + t))^{a_3}g(\theta_2; a_2 + 1, b_2 + t)g(\theta_3; a_3, b_3 + t)$$
$$\times [g(\theta_1; a_1, b_1) - (b_1/(b_1 + t))^{a_1}g(\theta_1; a_1, b_1 + t)]$$
$$+ (a_3/(b_3 + t))(b_3/(b_3 + t))^{a_3}(b_1/(b_1 + t))^{a_1}g(\theta_3; a_3 + 1, b_3 + t)g(\theta_1; a_1, b_1 + t)$$
$$\times [g(\theta_2; a_2, b_2) - (b_2/(b_2 + t))^{a_2}g(\theta_2; a_2, b_2 + t)]$$
$$+ (a_3/(b_3 + t))(b_3/(b_3 + t))^{a_3}(b_2/(b_2 + t))^{a_2}g(\theta_3; a_3 + 1, b_3 + t)g(\theta_2; a_2, b_2 + t)$$
$$\times [g(\theta_1; a_1, b_1) - (b_1/(b_1 + t))^{a_1}g(\theta_1; a_1, b_1 + t)]$$
$$+ (a_3/(b_3 + t))(b_3/(b_3 + t))^{a_3}(b_1/(b_1 + t))^{a_1}(b_2/(b_2 + t))^{a_2}$$
$$\times g(\theta_3; a_3 + 1, b_3 + t)g(\theta_1; a_1, b_1 + t)g(\theta_2; a_2, b_2 + t)$$

From (3.1) we then get $(i = 1, 2; j \neq i)$

$$\pi(\underline{\theta}|B_i(t)) = g(\theta_i; a_i + 1, b_i + t)g(\theta_3; a_3, b_3 + t)$$
$$\times [g(\theta_j; a_j, b_j) - (b_j/(b_j + t))^{a_j}g(\theta_j; a_j, b_j + t)]/[1 - (b_j/(b_j + t))^{a_j}]$$

$$\pi(\underline{\theta}|B_{i+2}(t)) = g(\theta_i; a_i, b_i + t)g(\theta_3; a_3 + 1, b_3 + t)$$
$$\times [g(\theta_j; a_j, b_j) - (b_j/(b_j + t))^{a_j}g(\theta_j; a_j, b_j + t)]/[1 - (b_j/(b_j + t))^{a_j}] \qquad (3.3)$$

$$\pi(\underline{\theta}|B_5(t)) = g(\theta_1; a_1, b_1 + t)g(\theta_2; a_2, b_2 + t)g(\theta_3; a_3 + 1, b_3 + t)$$
$$p_i(t) = \alpha_i(t)/\sum_{k=1}^{5} \alpha_k(t),$$

where $(i = 1, 2; j \neq i)$

$$\alpha_i(t) = (a_i/b_i)(b_i/(b_i + t))^{a_i+1}(b_3/(b_3 + t))^{a_3}[1 - (b_j/(b_j + t))^{a_j}]$$
$$\alpha_{i+2}(t) = (a_3/b_3)(b_i/(b_i + t))^{a_i}(b_3/(b_3 + t))^{a_3+1}[1 - (b_j/(b_j + t))^{a_j}]$$
$$\alpha_5(t) = (a_3/b_3)(b_1/(b_1 + t))^{a_1}(b_2/(b_2 + t))^{a_2}(b_3/(b_3 + t))^{a_3+1}$$

Now

$$\lim_{t \to 0} \alpha_i(t) = 0 \ , i = 1, 2, 3, 4 \quad \lim_{t \to 0} \alpha_5(t) = a_3/b_3$$

Hence

$$\lim_{t \to 0} p_i(t) = 0 \ , i = 1, 2, 3, 4 \quad \lim_{t \to 0} p_5(t) = 1,$$

and accordingly

$$\lim_{t \to 0} \pi(\underline{\theta}|T = t) = g(\theta_1; a_1, b_1)g(\theta_2; a_2, b_2)g(\theta_3; a_3 + 1, b_3)$$

This is intuitively obvious since the shape parameter a_3 is added by 1 corresponding to a common shock, whereas all scale parameters are unchanged due to zero time at test.

It easily follows from (3.3) that $(i = 1, 2, 3, 4)$

$$\lim_{t \to \infty} p_5(t)/p_i(t) = 0$$

Assume without loss of generality that $a_2 > a_1$. As in Gåsemyr & Natvig (1994) we get $(i = 2, 4)$

$$\lim_{t \to \infty} p_1(t)/p_i(t) = \lim_{t \to \infty} p_3(t)/p_i(t) = \infty$$

Now since

$$\lim_{t \to \infty} p_1(t)/p_3(t) = a_1/a_3,$$

it follows that for $a_2 > a_1$

$$\lim_{t \to \infty} p_1(t) = a_1/(a_1 + a_3) \quad \lim_{t \to \infty} p_3(t) = a_3/(a_1 + a_3)$$

By (3.1) and the expressions for $\pi(\underline{\theta}|B_i(t)), i = 1, 3$ in (3.3) the probability measure corresponding to $\pi(\theta_1, \theta_2, \theta_3|T = t)$ converges weakly to the product measure of the Dirac measures at 0, $\delta_0(\theta_1)$ and $\delta_0(\theta_3)$, and the measure corresponding to $g(\theta_2; a_2, b_2)$. For the second individual shock this is intuitively obvious since we just know that the time until it occurs is less than infinity $(\lim_{t \to \infty}(p_2(t) + p_4(t) + p_5(t)) = 0)$ and hence our prior assessment is unchanged.

Now let $a_2 > a_1, b_2 < b_1$. Here the prior mean, a_1/b_1, of θ_1 is less than the prior mean, a_2/b_2, of θ_2. Consider a vector, $\underline{T}$, of independent system lifetimes. We shall show that

$$\pi(\theta_1 \geq \theta_0|\underline{T}) < \pi(\theta_2 \geq \theta_0|\underline{T});$$

i.e. θ_2 is posterior stochastically larger than θ_1. As a special case this is also true a priori and the stochastic order is preserved. The argument is completely parallel to the one in Gåsemyr & Natvig (1994).

Denote the likelihood function by $L(\underline{\theta}; \underline{T})$. This is obviously symmetric in θ_1 and θ_2 since this is the case for $g_{\{1,2\}}(t)$ given by (3.2) and since system lifetimes are independent. Define $(i = 1, 2)$

$$I(\theta_i) = I(\theta_i \geq \theta_0)$$

We now have

$$\pi(\theta_1 \geq \theta_0|\underline{T}) - \pi(\theta_2 \geq \theta_0|\underline{T}) \propto$$

$$\int_{0 < \theta_2 < \theta_1 < \infty} L(\underline{\theta}; \underline{T})[I(\theta_1) - I(\theta_2)]\pi(\theta_1, \theta_2, \theta_3)d\theta_1 d\theta_2 d\theta_3$$

$$+ \int_{0 < \theta_1 < \theta_2 < \infty} L(\underline{\theta}; \underline{T})[I(\theta_1) - I(\theta_2)]\pi(\theta_1, \theta_2, \theta_3)d\theta_1 d\theta_2 d\theta_3$$

10

By interchanging θ_1 and θ_2 in the last integral using the symmetry of $L(\underline{\theta};\underline{T})$ this equals

$$= \int\limits_{0<\theta_2<\theta_1<\infty} L(\underline{\theta};\underline{T})[I(\theta_1) - I(\theta_2)][\pi(\theta_1,\theta_2,\theta_3) - \pi(\theta_2,\theta_1,\theta_3)]d\theta_1 d\theta_2 d\theta_3$$

$$\propto \int\limits_{0<\theta_2<\theta_1<\infty} L(\underline{\theta};\underline{T})[I(\theta_1) - I(\theta_2)](\theta_1\theta_2)^{a_1-1}e^{-(\theta_1+\theta_2)b_2}$$

$$\times [\theta_2^{a_2-a_1}e^{-(b_1-b_2)\theta_1} - \theta_1^{a_2-a_1}e^{-(b_1-b_2)\theta_2}]\theta_3^{a_3-1}e^{-b_3\theta_3}d\theta_1 d\theta_2 d\theta_3 < 0$$

Note that the argument does not depend on the values of the parameters a_3 and b_3.

Let us now return to the case of a general prior distribution $\pi(\underline{\theta})$. In Gåsemyr & Natvig (1994) for the case with no common shock, assuming θ_1 and θ_2 to be prior independent, it is shown that they are posterior negatively correlated. It is demonstrated here that this is not necessarily the case when a common shock is added, illustrating the consequences of dependence in reliability.

From (3.1) assuming $\theta_1, \theta_2, \theta_3$ to be prior independent

$$\mathrm{Cov}(\theta_1,\theta_2|T=t) = \sum_{\substack{i=1 \\ j\neq i}}^{2} p_i(t) \int_0^\infty \int_0^\infty \int_0^\infty [\theta_1 - E(\theta_1|T=t)][\theta_2 - E(\theta_2|T=t)]$$

$$\times \pi(\theta_i|V_i=t)\pi(\theta_j|V_j\leq t)\pi(\theta_3|V_3>t)d\theta_1 d\theta_2 d\theta_3$$

$$+ \sum_{\substack{i=1 \\ j\neq i}}^{2} p_{i+2}(t) \int_0^\infty \int_0^\infty \int_0^\infty [\theta_1 - E(\theta_1|T=t)][\theta_2 - E(\theta_2|T=t)]$$

$$\times \pi(\theta_i|V_i>t)\pi(\theta_j|V_j\leq t)\pi(\theta_3|V_3=t)d\theta_1 d\theta_2 d\theta_3$$

$$+ p_5(t) \int_0^\infty \int_0^\infty \int_0^\infty [\theta_1 - E(\theta_1|T=t)][\theta_2 - E(\theta_2|T=t)] \qquad (3.4)$$

$$\times \pi(\theta_1|V_1>t)\pi(\theta_2|V_2>t)\pi(\theta_3|V_3=t)d\theta_1 d\theta_2 d\theta_3$$

$$= \sum_{\substack{i=1 \\ j\neq i}}^{2} p_i(t)[E(\theta_i|V_i=t) - E(\theta_i|T=t)][E(\theta_j|V_j\leq t) - E(\theta_j|T=t)]$$

$$+ \sum_{\substack{i=1 \\ j\neq i}}^{2} p_{i+2}(t)[E(\theta_i|V_i>t) - E(\theta_i|T=t)][E(\theta_j|V_j\leq t) - E(\theta_j|T=t)]$$

$$+ p_5(t)[E(\theta_1|V_1>t) - E(\theta_1|T=t)][E(\theta_2|V_2>t) - E(\theta_2|T=t)]$$

By noting that $(i=1,2; j\neq i)$

$$E(\theta_i|T=t) = p_i(t)E(\theta_i|V_i=t) + (p_j(t) + p_{j+2}(t))E(\theta_i|V_i\leq t)$$
$$+ (p_{i+2}(t) + p_5(t))E(\theta_i|V_i>t) \qquad (3.5)$$

$$\sum_{i=1}^{5} p_i(t) = 1, \qquad (3.6)$$

11

it is shown in Appendix 2 that

$$
\begin{aligned}
\mathrm{Cov}(\theta_1, \theta_2 | T = t) = & \\
- & [E(\theta_1 | V_1 \le t) - E(\theta_1 | V_1 = t)][E(\theta_2 | V_2 \le t) - E(\theta_2 | V_2 = t)]p_1(t)p_2(t) \\
+ & [E(\theta_1 | V_1 = t) - E(\theta_1 | V_1 > t)][E(\theta_2 | V_2 \le t) - E(\theta_2 | V_2 > t)]p_1(t)(p_4(t) + p_5(t)) \\
+ & [E(\theta_1 | V_1 \le t) - E(\theta_1 | V_1 > t)][E(\theta_2 | V_2 = t) - E(\theta_2 | V_2 > t)]p_2(t)(p_3(t) + p_5(t)) \\
- & [E(\theta_1 | V_1 \le t) - E(\theta_1 | V_1 > t)][E(\theta_2 | V_2 \le t) - E(\theta_2 | V_2 > t)](p_1(t)p_4(t) \\
+ & p_2(t)p_3(t) + p_3(t)p_4(t))
\end{aligned}
\tag{3.7}
$$

Note that from Barlow & Proschan (1985) we have "under mild regularity conditions" that $E(\theta_i | V_i)$ is nonincreasing in $V_i, i = 1, 2, 3$.

Now specialize $\pi(\theta_i) = g(\theta_i; a_i, b_i), i = 1, 2, 3$. It is then shown in Appendix 2 that (3.7) gives

$$
\begin{aligned}
\mathrm{Cov}(\theta_1, \theta_2 | T = t) \propto & (b_1/(b_1 + t))^{a_1+1}(b_2/(b_2 + t))^{a_2+1}(b_3/(b_3 + t))^{2a_3}(a_1 a_2)/(b_1^2 b_2^2) \\
\times & \{-(b_1/(b_1 + t))(b_2/(b_2 + t))[(a_1/b_1)t - 1 + (b_1/(b_1 + t))^{a_1}][(a_2/b_2)t - 1 + (b_2/(b_2 + t))^{a_2}] \\
+ & (a_3 t/(b_3 + t))[b_1/(b_1 + t) + b_2/(b_2 + t) - t(a_1/(b_1 + t) + a_2/(b_2 + t) + a_3/(b_3 + t))]\}
\end{aligned}
\tag{3.8}
$$

By choosing $t = 1, a_1 = a_2 = 1, a_3 = 6, b_1 = b_2 = b_3 = 5$, we get as promised

$$
\mathrm{Cov}(\theta_1, \theta_2 | T = t) > 0
$$

Note that in our choice of parameters $a_3/b_3 = 6/5, a_i/b_i = 1/5 \ i = 1, 2$ giving a much higher prior mean of θ_3 than of θ_1 and θ_2.

Acknowledgement

References

Abraham, J.A. (1979). An improved algorithm for network reliability. IEEE Trans. Reliability **28**, 58-61.

Antoine, R., Doss, H. & Hollander, M. (1993). On identifiability in the autopsy model of reliability theory. J. Appl. Prob. **30**, 913-930.

Ball, M.O. & Provan, J.S. (1988). Disjoint products and efficient computation of reliability. Operations Research **36**, 703-715.

Barlow, R.E. & Proschan, F. (1975). Statistical theory of reliability and life testing. Probability models. Holt, Rinehart and Winston, New York.

Barlow, R.E. & Proschan, F. (1985). Inference for the exponential life distribution. Theory of Reliability. Soc. Italiana di Fisica, Bologna 143-164.

Bernoulli, D. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite Vérole, et des avantages de l'Inoculation pour la prévenir. Mém de l'Académie Royale de Science, 1760, 1-45.

Boyles, R.A. & Samaniego, F.J. (1984). Modelling and inference for multivariate binary data with positive dependence. J. Amer. Statist. Ass. **79**, 188-193.

Crowder, M. (1994). Identifiability crises in competing risks. Int. Statist. Rev. **62**, 379-391.

Gåsemyr, J. (1995). On identifiability in the autopsy model of reliability systems of lifemonitored and conditionally lifemonitored components. Tech. rep., Department of Mathematics, Univ. of Oslo.

Gåsemyr, J. & Natvig, B. (1994). Bayesian estimation of component lifetimes based on autopsy data. Tech. rep., Department of Mathematics, Univ. of Oslo.

Gåsemyr, J. & Natvig, B. (1995a). Some aspects of reliability analysis in shock models. Scand. J. Statist. **22**, to appear.

Gåsemyr, J. & Natvig, B. (1995b). Using expert opinions in Bayesian prediction of component lifetimes in a shock model. Maths. Operations Research **20**, 227-242.

Locks, M.O. (1980). Recursive disjoint products, inclusion-exclusion, and min-cut approximations. IEEE. Trans. Reliability **29**, 368-371.

Locks, M.O. (1982). Recursive disjoint products: a review of three algorithms. IEEE. Trans. Reliability **31**, 33-35.

Marshall, A.W. & Olkin, I. (1967). A multivariate exponential distribution. J. Amer. Statist. Ass. **62**, 30-44.

Meilijson, I. (1981). Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine. J. Appl. Prob. **18**, 829-838.

Meilijson, I. (1994). Competing risks on coherent reliability systems: Estimation in the parametric case. J. Amer. Statist. Ass. **89**, 1459-1464.

Nowik, S. (1990). Identifiability problems in coherent systems. J. Appl. Prob. **28**, 862-872.

Shier, D.R. & Whited, D.E. (1985). Algorithms for generating minimal cut sets by inversion. IEEE Trans. Reliability **34**, 314-319.

Appendix 1

The first part of this appendix is concerned with the computation of the set A_1. Also we consider partitioning of $A_1 - B_1$, leading to a corresponding factorization of the expression for $g_A(t)$. Similar factorizations result from treating the sets $A(H) - B(H)$ and D_L in the same way. It turns out that both questions can be handled by focusing on minimality rather than maximality.

It is convenient to introduce some new notation. For arbitrary $A' \subset E, C \subset \{1, \cdots, n+p\}$ define

$$M_{A',C} = \text{the smallest set of components } M \text{ containing } A' \text{ such that } D_{E_M - C} = M \quad (A.1)$$

Note that this is well defined since, if $A' \subset M_i$ and $D_{E_{M_i} - C} = M_i, i = 1, 2$, then also $D_{E_M - C} = M$ for $M = M_1 \cap M_2$. $M_{A',C}$ can easily be computed as follows: Put $M_1 = A'$ and recursively $M_{i+1} = D_{E_{M_i} - C}, i = 1, 2, \cdots$. Then $M_{A',C} = M_i$ for the first i such that $M_{i+1} = M_i$. Now it is easily seen that

$$A_1 = A - M_{\bigcup_{j \epsilon C_A} D_j, C_1} \quad (A.2)$$

If A', C introduced before (A.1) satisfy $D_{E_{A'} - C} = A'$, define

$$\mathcal{M}_{A',C} = \text{Minimal sets in } \{M \subset A' | D_{E_M - C} = M\} \quad (A.3)$$

$\mathcal{M}_{A',C}$ can easily be found by starting with an arbitrary component $l_1 \epsilon A'$ and forming $M_1 = M_{\{l_1\},C}$, then choosing $l_2 \epsilon A' - M_1$ etc.

For $C \subset \{1, \cdots, n+p\}$ and M as in (A.3), define

$$\mathcal{B}_{M,C} = \text{Minimal sets in } \{B \subset \{1, \cdots, n+p\} - C | D_B = M\} \quad (A.4)$$

$\mathcal{B}_{M,C}$ can be found either by inspection or by for instance the minimal path-to-minimal cut inversion algorithm of Shier & Whited (1985), with the minimal path sets $\{E_l - C | l \epsilon M\}$ as input.

Now (2.2) can be factorized as follows

$$P[\bigcup_{B \in \mathcal{B}_1} (V_l \leq t \; \forall l \epsilon B)] = \prod_{M \epsilon \mathcal{M}_{A_1 - B_1, C_1}} P[\bigcup_{B \in \mathcal{B}_{M,C_1}} (V_l \leq t \; \forall l \epsilon B)] \quad (A.5)$$

Similar factorizations of (2.3) and (2.5) corresponding to $\mathcal{B}(H)$ and $\mathcal{B}(L)$ can be done according to the sets in $\mathcal{M}_{A(H) - B(H), C_1 \cup C(H)}$ and $\mathcal{M}_{D_L, \{1, \cdots, n+p\} - L}$ respectively.

Next, we will consider the problem of computing an expression of the form

$$P[\bigcup_{B \in \mathcal{B}} (V_l \leq t \; \forall l \epsilon B)] \quad (A.6)$$

For instance, in our reference example we need to compute (A.6) with $\mathcal{B} = \mathcal{B}_1 = \{\{18\}, \{7,8\}\}$. We want to obtain an expression that is as structurally simple as possible, keeping the number of summands low. Questions concerning algorithmic implementations and computational complexity will not be considered. The computation is essentially an adaptation of the disjoint product algorithm described in Abraham (1979) and Locks (1980, 1982). We denote by $F_B(t)$ the event that $V_l \leq t$ for all $l \in B$, and let $B_1, \cdots, B_r$ be the sets in $\mathcal{B}$. The relevant event can then be written as

$$\bigcup_{B \in \mathcal{B}} F_B(t) = \bigcup_{k=1}^{r} [F_{B_k}(t) \bigcap (\bigcap_{1 \leq i < k} F_{B_i}(t)^c)], \quad (A.7)$$

14

where the union on the right hand side is disjoint (see e.g. Locks (1982), theorem 1). We want to write each set in this union, and hence the entire event, as a disjoint union of simple events, i.e. in the form

$$W = \bigcup_{i=1}^{\nu} [(\bigcap_{l \in P_i} (V_l > t)) \bigcap (\bigcap_{l \in Q_i} (V_l \leq t))] \tag{A.8}$$

where the sets P_i, Q_i satisfy $(P_i \cap Q_j) \cup (P_j \cap Q_i) \neq \emptyset, i \neq j$ to ensure disjointness. We associate with (A.8) the polynomial

$$f(\mathbf{y}) = \sum_{i=1}^{\nu} (\prod_{l \in P_i} y_l)(\prod_{l \in Q_i} \bar{y}_l) \tag{A.9}$$

defined for $(\mathbf{y} \in \{0, 1\}^{n+p}(\bar{y} = 1 - y$ by definition). Then multiplication of polynomials corresponds to intersection of events, and for disjoint events, addition corresponds to union. Putting $Y_l = I(V_l > t)$ we obtain

$$P(W) = Ef(\mathbf{Y}) = f(\bar{F}_1(t), \cdots, \bar{F}_{n+p}(t)) \tag{A.10}$$

(extending the definition of f to $[0, 1]^{n+p}$). To write (A.6) in the form (A.10) note first that $F_{B_k}(t)$ is represented by the polynomial $\prod_{l \in B_k} \bar{y}_l$, whereas $F_{B_i}(t)^c$ is represented by

$$\sum_{j \in B_i} y_j (\prod_{l \in B_i, l < j} \bar{y}_l) \tag{A.11}$$

(see Locks (1982), theorem 2). Using these building blocks, a polynomial corresponding to (A7) may be constructed. The resulting expression may, however, be considerably simplified. First, it can be seen that in the term corresponding to $F_{B_k}(t) \cap (\cap_{1 \leq i < k} F_{B_i}(t)^c)$ (A.11) may be replaced by

$$\sum_{j \in B_i - B_k} y_j (\prod_{l \in B_i - B_k, l < j} \bar{y}_l) \tag{A.12}$$

(confer Abraham (1979), theorem 1, b - ii). Moreover, if for $i_1, i_2 < k$ we have $B_{i_1} - B_k \subset B_{i_2} - B_k$, then the term corresponding to i_2 may be dropped, since $\{V_l > t\}$ for at least one $l \in B_{i_1} - B_k$ implies $\{V_l > t\}$ for at least one $l \in B_{i_2} - B_k$ (confer the argument preceding (4) in Ball & Provan (1985)). Hence, we only need the minimal sets among the sets $B_i - B_k$. Denote the corresponding set of integers $i < k$ by M_k. Note also that the ordering of the shocks in $B_i - B_k$ can be changed arbitrarily without changing the corresponding event, but with a possible influence on the resulting polynomial. For $i \in M_k$ define an ordering $\pi_{i,k}$ on $B_i - B_k$, i.e. a one-to-one function from $B_i - B_k$ onto $\{1, 2, \cdots, |B_i - B_k|\}$. We then have that (A.7) is represented by the polynomial

$$f_B(\mathbf{y}) = \sum_{k=1}^{r} (\prod_{l \in B_k} \bar{y}_l)(\prod_{i \in M_k} (\sum_{j \in B_i - B_k} y_j (\prod_{l \in B_i - B_k, \pi_{i,k}(l) < \pi_{i,k}(j)} \bar{y}_l))) \tag{A.13}$$

Combining (A.10) and (A.13) finally gives the desired expression.

The number of summands in (A.13) is potentially as large as $\sum_{k=1}^{r} \prod_{i \in M_k} |B_i - B_k|$. However, in practise the number of summands may be considerably smaller due to the orthogonality

relation $y_l \bar{y}_l = 0$. For instance, $(y_1 + \bar{y}_1 y_2)(y_1 + \bar{y}_1 y_3) = y_1 + \bar{y}_1 y_2 y_3$. Changing the ordering, we see that the same event is represented by $(y_2 + \bar{y}_2 y_1)(y_3 + \bar{y}_3 y_1) = y_2 y_3 + y_1 y_2 \bar{y}_3 + y_1 \bar{y}_2$ which illustrates the usefulness of being able to change the ordering of $B_i - B_j$. It is also important to choose a convenient ordering of $\mathcal{B}$ in the outset. In particular, it may be possible to order the sets in $\mathcal{B}$ in such a way that $|B_i - B_k| = 1$ for all $i \in M_k, k = 1, 2, \cdots, r$. This corresponds to a shelling in the sense of Ball & Provan (1985) (and to shellability of $\mathcal{B}$), and leads to $r = |\mathcal{B}|$ summands in (A.13). In the reference example, with $\mathcal{B} = \{\{18\}, \{7, 8\}\}$, we obtain a shelling by putting $B_1 = \{18\}, B_2 = \{7, 8\}$, resulting in the term $F_{18}(t) + \bar{F}_{18}(t) F_7(t) F_8(t)$. In general, it pays to start with cut sets with few shocks, i.e. containing many large shocks. A case by case inspection shows that if $\mathcal{B}$ is of the form $\mathcal{B}_{M,C}$ (confer (A.4)) and if M has at most three components, then $\mathcal{B}$ is shellable. However, with $|M| = 4$, this is no longer true.

Alternatively, one may use a disjoint product procedure based on minimal path sets rather than minimal cut sets (this is the procedure applied in the references). Examples indicate that shellability is then obtained more rarely. On the other hand, the number of minimal path sets is usually smaller, and the number of summands in the expression corresponding to (A.13) will therefore often be no larger. Another advantage is that it is easier to find the minimal path sets. In fact, the minimal path sets corresponding to the minimal cut sets in $\mathcal{B}_{M,C}$ are simply $\{E_l - C | l \epsilon M\}$.

Finally, we show how different contributions to (2.5) can be grouped together to obtain a further simplification of the expression for $g_A(t)$. Before going into details, it is worthwhile to give a few remarks concerning the set $\mathcal{L}$. Note first that for each $L \in \mathcal{L}$ if $D_l \subset D_L$, then $l \epsilon L$. Furthermore, due to the subtraction of $B(H)$ in the definition of $\mathcal{L}$ if $i \epsilon D_L$, there exists a common shock $l \epsilon L$ such that $i \epsilon D_l$. It follows that L is uniquely characterized by $L \cap \{n + 1, \cdots, n + p\}$.

The following simplification of (2.5) is obtained by grouping together critical sets belonging to the same $L \in \mathcal{L}$ and the same critical shock. For $L \in \mathcal{L}$ define

$$\mathcal{J}(L) = \{j \epsilon C_A | D_j - D_L \text{ is a critical set } \cap F(D_j - D_L) - B(D_j - D_L) = L\} \quad (A.14)$$

Furthermore, for j, L such that $j \epsilon \mathcal{J}(L)$ define

$$\mathcal{H}'(j, L) = \{H \subset D_j - D_L | H \text{ is a critical set } \cap F(H) - B(H) = L\} \quad (A.15)$$

$$\mathcal{H}(j, L) = \text{ Minimal sets in } \mathcal{H}'(j, L) \quad (A.16)$$

Clearly, $H' \in \mathcal{H}'(j, L)$ if and only if $H' \subset D_j - D_L$ and there exists $H \in \mathcal{H}(j, L)$ such that $H \subset H'$. Hence

$$\bigcup_{H \in \mathcal{H}(j,L)} \bigcap_{l \epsilon H} (V_l > t) = \bigcup_{H \in \mathcal{H}'(j,L)} \bigcap_{l \epsilon H} (V_l > t)$$

$$= \bigcup_{H \in \mathcal{H}'(j,L)} [[\bigcap_{l \epsilon H} (V_l > t)] \bigcap [\bigcap_{l \epsilon D_j - D_L - H} (V_l \leq t)]], \quad (A.17)$$

the last union being a disjoint representation.

Note also that since all sets $H \in \mathcal{H}'(j, L)$ by (A.15) give rise to the same $F(H) \cap \{n + 1, \cdots, n + p\}$, due to the relation between $F(H)$ and $C(H)$ they also must give rise to the

same $C(H) \cap \{n+1, \cdots, n+p\}$. Moreover, this set only depends on L; i.e. is independent of j. Denote this common value by $C(L)$. Furthermore, define $B(j,L) = A - A_1 \cup D_L \cup D_j$. Now the last part of (2.5) can be expressed in terms of the *minimal* sets $\mathcal{H}(j,L)$ as follows

$$
\sum_{H \in \mathcal{H}(A,L)} \prod_{l \in C(H)} \bar{F}_l(t) \prod_{l \in B(H)} F_l(t) \sum_{\{j \in C_A | H \subset D_j\}} \lambda_j(t)
$$

$$
= \prod_{l \in C(L)} \bar{F}_l(t) \sum_{j \in J(L)} \lambda_j(t) \prod_{l \in B(j,L)} F_l(t) \sum_{H \in \mathcal{H}'(j,L)} \prod_{l \in H} \bar{F}_l(t) \prod_{l \in D_j - D_L - H} F_l(t) \qquad (A.18)
$$

$$
= \prod_{l \in C(L)} \bar{F}_l(t) \sum_{j \in J(L)} \lambda_j(t) \prod_{l \in B(j,L)} F_l(t) P[\bigcup_{H \in \mathcal{H}(j,L)} \bigcap_{l \in H}(V_l > t)],
$$

having applied (A.17). The factor in brackets can be computed by a disjoint product procedure based on minimal path sets.

In our reference system we have for instance for $L = \emptyset, J(\emptyset) = 14, \mathcal{H}'(14, \emptyset) = \{\{1,2,4\}$,

$\{1,2,4,6\}, \{2,4\}, \{2,4,6\}\}, \mathcal{H}(14, \emptyset) = \{2,4\}, C(\emptyset) = \{12, 13, 14, 15\}, B(14, \emptyset) = \{3, 5\}$. This accounts for the last expression in (2.6). In this system $\mathcal{H}(j, L)$ always consists of a single element, the other possibilities being $\mathcal{H}(14, \{2, 3, 13\}) = \{1, 4\}, \mathcal{H}(j, \{1, 2, 3, 12, 13\}) = \{4\}$ for $j = 4, 14, 15, \mathcal{H}(j, \{4, 5, 15\}) = \{1, 2\}$ for $j = 12, 14$.

Appendix 2

Here we give the details first in the deduction of (3.7) from (3.4), (3.5) and (3.6) and then (3.8) from (3.7).

From (3.4) and (3.5) we get

$\text{Cov}(\theta_1, \theta_2 | T = t)$
$= p_1(t)[(1 - p_1(t))E(\theta_1 | V_1 = t) - (p_2(t) + p_4(t))E(\theta_1 | V \le t) - (p_3(t) + p_5(t))E(\theta_1 | V_1 > t)]$
$\times [(1 - p_1(t) - p_3(t))E(\theta_2 | V_2 \le t) - p_2(t)E(\theta_2 | V_2 = t) - (p_4(t) + p_5(t))E(\theta_2 | V_2 > t)]$
$+ p_2(t)[(1 - p_2(t))E(\theta_2 | V_2 = t) - ((p_1(t) + p_3(t))E(\theta_2 | V_2 \le t) - (p_4(t) + p_5(t))E(\theta_2 | V_2 > t)]$
$\times [(1 - p_2(t) - p_4(t))E(\theta_1 | V_1 \le t) - p_1(t)E(\theta_1 | V_1 = t) - (p_3(t) + p_5(t))E(\theta_1 | V_1 > t)]$
$+ p_3(t)[1 - p_3(t) - p_5(t))E(\theta_1 | V_1 > t) - p_1(t)E(\theta_1 | V_1 = t) - (p_2(t) + p_4(t))E(\theta_1 | V_1 \le t)]$
$\times [(1 - p_1(t) - p_3(t))E(\theta_2 | V_2 \le t) - p_2(t)E(\theta_2 | V_2 = t) - (p_4(t) + p_5(t))E(\theta_2 | V_2 > t)]$
$+ p_4(t)[(1 - p_4(t) - p_5(t))E(\theta_2 | V_2 > t) - p_2(t)E(\theta_2 | V_2 = t) - (p_1(t) + p_3(t))E(\theta_2 | V_2 \le t)]$
$\times [(1 - p_2(t) - p_4(t))E(\theta_1 | V_1 \le t) - p_1(t)E(\theta_1 | V_1 = t) - (p_3(t) + p_5(t))E(\theta_1 | V_1 > t)]$
$+ p_5(t)[(1 - p_3(t) - p_5(t))E(\theta_1 | V_1 > t) - p_1(t)E(\theta_1 | V_1 = t) - (p_2(t) + p_4(t))E(\theta_1 | V_1 \le t)]$
$\times [(1 - p_4(t) - p_5(t))E(\theta_2 | V_2 > t) - p_2(t)E(\theta_2 | V_2 = t) - (p_1(t) + p_3(t))E(\theta_2 | V_2 \le t)]$

17

After some tedious but straightforward algebra applying (3.6) this can be rewritten as

$$
\begin{aligned}
& E(\theta_1|V_1 = t)E(\theta_2|V_2 \le t)[p_1(t)(p_2(t) + p_4(t) + p_5(t))] \\
& + E(\theta_1|V_1 = t)E(\theta_2|V_2 = t)[-p_1(t)p_2(t)] \\
& + E(\theta_1|V_1 = t)E(\theta_2|V_2 > t)[-p_1(t)(p_4(t) + p_5(t))] \\
& + E(\theta_1|V_1 \le t)E(\theta_2|V_2 \le t)[-(p_1(t) + p_3(t))(p_2(t) + p_4(t))] \\
& + E(\theta_1|V_1 \le t)E(\theta_2|V_2 = t)[p_2(t)(p_1(t) + p_3(t) + p_5(t))] \\
& + E(\theta_1|V_1 \le t)E(\theta_2|V_2 > t)[p_4(t)(p_1(t) + p_3(t)) - p_2(t)p_5(t)] \\
& + E(\theta_1|V_1 > t)E(\theta_2|V_2 \le t)[p_3(t)(p_2(t) + p_4(t)) - p_1(t)p_5(t)] \\
& + E(\theta_1|V_1 > t)E(\theta_2|V_2 = t)[-p_2(t)(p_3(t) + p_5(t))] \\
& + E(\theta_1|V_1 > t)E(\theta_2|V_2 > t)[p_5(t)(p_1(t) + p_2(t)) - p_3(t)p_4(t)],
\end{aligned}
$$

which leads to (3.7).

Now specialize $\pi(\theta_i) = g(\theta_i; a_i, b_i)$, $i = 1, 2, 3$. Since

$$
\begin{aligned}
\pi(\theta_i|V_i \le t) &\propto (1 - e^{-\theta_i t})g(\theta_i; a_i, b_i) \\
\pi(\theta_i|V_i = t) &\propto \theta_i e^{-\theta_i t}g(\theta_i; a_i, b_i) \\
\pi(\theta_i|V_i > t) &\propto e^{-\theta_i t}g(\theta_i; a_i, b_i),
\end{aligned}
$$

we get

$$
\begin{aligned}
\pi(\theta_i|V_i \le t) &= [g(\theta_i; a_i, b_i) - (b_i/(b_i + t))^{a_i}g(\theta_i; a_i, b_i + t)] \\
&\quad /[1 - (b_i/(b_i + t))^{a_i}] \\
\pi(\theta_i|V_i = t) &= g(\theta_i; a_i + 1, b_i + t) \\
\pi(\theta_i|V_i > t) &= g(\theta_i; a_i, b_i + t).
\end{aligned}
$$

Hence

$$
\begin{aligned}
E(\theta_i|V_i \le t) &= (a_i/b_i)[1 - (b_i/(b_i + t))^{a_i+1}]/[1 - (b_i/(b_i + t))^{a_i}] \\
E(\theta_i|V_i = t) &= (a_i + 1)/(b_i + t) \\
E(\theta_i|V_i > t) &= a_i/(b_i + t).
\end{aligned}
$$

Inserting this into (3.7) applying (3.3) we get

$$
\begin{aligned}
\mathrm{Cov}(\theta_1, \theta_2|T = t) = & (\sum_{i=1}^{5} \alpha_i)^{-2} \\
& \{-(a_1 a_2/(b_1 b_2))(b_1/(b_1 + t))^{a_1+1}(b_2/(b_2 + t))^{a_2+1}(b_3/(b_3 + t))^{2a_3} \\
& \times [(a_1/b_1)(1 - (b_1/(b_1 + t))^{a_1+1}) - ((a_1 + 1)/(b_1 + t))(1 - (b_1/(b_1 + t))^{a_1})] \\
& \times [(a_2/b_2)(1 - (b_2/(b_2 + t))^{a_2+1}) - ((a_2 + 1)/(b_2 + t))(1 - (b_2/(b_2 + t))^{a_2})] \\
& + (a_1 a_3/(b_1 b_3))(b_1/(b_1 + t))^{a_1+1}(b_2/(b_2 + t))^{a_2}(b_3/(b_3 + t))^{2a_3+1} \\
& \times [(a_2/b_2)(1 - (b_2/(b_2 + t))^{a_2+1}) - (a_2/(b_2 + t))(1 - (b_2/(b_2 + t))^{a_2})]/(b_1 + t) \\
& + (a_2 a_3/(b_2 b_3))(b_1/(b_1 + t))^{a_1}(b_2/(b_2 + t))^{a_2+1}(b_3/(b_3 + t))^{2a_3+1}
\end{aligned}
$$

$$\times \; [(a_1/b_1)(1 - (b_1/(b_1 + t))^{a_1+1}) - (a_1/(b_1 + t))(1 - (b_1/(b_1 + t))^{a_1})]/(b_2 + t)$$

$$- (a_3/b_3)(b_1/(b_1 + t))^{a_1}(b_2/(b_2 + t))^{a_2}(b_3/(b_3 + t))^{2a_3+1}[a_1/(b_1 + t) + a_2/(b_2 + t)$$

$$+ a_3/(b_3 + t)] \times [(a_1/b_1)(1 - (b_1/(b_1 + t))^{a_1+1}) - (a_1/(b_1 + t))(1 - (b_1/(b_1 + t))^{a_1})]$$

$$\times \; [(a_2/b_2)(1 - (b_2/(b_2 + t))^{a_2+1}) - (a_2/(b_2 + t))(1 - (b_2/(b_2 + t))^{a_2})]\}$$

$$= (\sum_{i=1}^{5} \alpha_i)^{-2}(b_1/(b_1 + t))^{a_1}(b_2/(b_2 + t))^{a_2}(b_3/(b_3 + t))^{2a_3}$$

$$\times \; \{-(a_1a_2/(b_1b_2))(b_1/(b_1 + t))(b_2/(b_2 + t))[(a_1t - b_1)/(b_1(b_1 + t)) + (b_1/(b_1 + t))^{a_1+1}/b_1]$$

$$\times \; [(a_2t - b_2)/(b_2(b_2 + t)) + (b_2/(b_2 + t))^{a_2+1}/b_2]$$

$$+ (a_1a_3/(b_1b_3))(b_1/(b_1 + t))(b_3/(b_3 + t))a_2t/((b_1 + t)b_2(b_2 + t))$$

$$+ (a_2a_3/(b_2b_3))(b_2/(b_2 + t))(b_3/(b_3 + t))a_1t/((b_1 + t)b_1(b_2 + t))$$

$$- (a_3/(b_3 + t))[a_1/(b_1 + t) + a_2/(b_2 + t) + a_3/(b_3 + t)](a_1a_2t^2/(b_1b_2(b_1 + t)(b_2 + t)))\}$$

This readily reduces to (3.8).