Received: 31 October 2022 | Revised: 30 April 2023 | Accepted: 2 May 2023

DOI: 10.1002/asmb.2771

RESEARCH ARTICLE

WILEY

Phase-type models involving restarting and instantaneous transitions, with applications to degradation and maintenance

Bo Henry Lindqvist

Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence
Bo Henry Lindqvist, Department of Mathematical Sciences, N-7491 Trondheim, Norway.
Email: bo.lindqvist@ntnu.no

[Correction added on 31 May 2023, after first online publication: The copyright line was changed.]

Abstract

A phase-type distribution is the distribution of time to absorption for an absorbing continuous-time finite state Markov chain. The paper first reviews the extension of the phase-type setting to modeling of competing risks by introducing multiple absorbing states. The main study of the paper is the further extension to introducing instantaneous transitions at certain stages of the original models. The motivation is from applications to repair and maintenance, bringing failed systems into working ones by instantaneous repair actions. Two slightly different approaches are studied. The first one is based on restarting the original Markov chain upon absorption, leading to the consideration of a Markov renewal process. The second approach involves periodically inspected systems, where maintenance actions are modeled by instantaneous transitions made at regular inspection times. For both approaches are suggested measures of reliability and maintenance based on long run properties.

KEYWORDS

competing risks, Coxian distribution, degradation, maintenance, periodic inspection, phase-type distribution

1 | INTRODUCTION

A phase-type distribution can be defined to be the distribution of the time to absorption for an absorbing continuous-time finite state Markov chain. Simple examples are mixtures and convolutions of exponential distributions, thus generalizing Erlang distributions. Neuts[1] gave a theoretical background for phase-type distributions in the general setting. Later, a series of papers by O'Cinneide, for example, References 2-4, introduced and clarified a number of theoretical issues. Nice introductions to applications in survival analysis are Aalen[5] and Slud and Suntornchost.[6] The latter paper is also an excellent source for the history and motivation for phase-type distributions and their analysis. See also Lindqvist[7] for a recent review.

The literature contains a number of articles demonstrating the use of phase-type distributions in various applications, both with respect to modeling and statistical inference. There is currently a particular interest in the use of so-called Coxian phase-type models, first considered by Cox,[8] where individuals or items go through successive stages (phases), and may transit to an absorbing state (corresponding to the event of interest) from any phase. Such models have lately been popular in health care studies. For example, Faddy et al.,[9] McGrory et al.,[10] Tang et al.,[11] and Rizk et al.[12] modeled hospital length of stay by Coxian phase-type models. Faddy et al.[9] claimed the superiority of Coxian phase-type models over common parametric models like gamma and lognormal.

The present paper is motivated by applications of phase-type models in reliability. A remarkable paper by Cumani[13] appeared in 1981, considering Markov models for failure distributions, involving acyclic transition graphs. The author showed that such models can be uniquely transferred into Coxian models with a special structure. Other early references involving reliability applications are Neuts and Meier[14] and Neuts, Perez-Ocon and Torres-Castro.[15]

Later, simple Markov chain models for systems or components subjected to degradation and maintenance were given by Hokstad and Frøvig[16] and Langseth et al..[17] Typically, critical failures here correspond to transitions to absorbing states. Pham et al.[18] considered Markov chain models for components which might not always fail fully, but may degrade with multiple stages of degradation. The model is essentially a Coxian type model, but includes the possibility of recovery from absorption in failure states. Such features will in fact play a main role in the present paper.

As indicated above, phase-type distributions have proven useful for modeling degradation and maintenance of systems or components. Typically, degradation can then be modeled through stages corresponding to the transient states of the underlying Markov chain. There is then commonly a need to consider more than one absorbing state, corresponding to different kinds of failures or shocks, or possibly maintenance actions. Lindqvist[7] gave a background for phase-type models with multiple absorbing states in a competing risks setting (see Section 2.2).

Phase-type models with multiple absorbing states have earlier been considered by, for example, McClean et al.[19] and Rizk et al.[12] in the modeling of patient pathways in hospitals. In Aalen,[5] the primary interest was in phase-type models for lifetimes, but the paper also considered the case with two absorbing states. The idea was that for some applications, not every individual will experience the event of interest. The connection was then made to so called "cure models". In a similar way, Slud et al.[6] studied phase-type models for survival data, introducing a second absorbing state for direct transition to the state of death or cure. Other applications in survival analysis involving competing risks and Markov chains, are given in, for example, Llorca et al.,[20] Abner et al.[21] and Garcia-Maya et al.[22] The latter authors considered phase-type modeling of competing risks in a semi-Markov process framework. In a recent paper, Wu and Cui[23] studied periodically inspected reliability systems involving competing risks, under environment processes modeled by absorbing Markov chains.

The purpose of the present paper is to show how phase-type models for competing risks can be extended by introducing instantaneous transitions, motivated by maintenance actions. Two slightly different approaches are studied. The first one is based on always restarting the original Markov chain upon absorption. The theoretical motivation is here a simple idea of Neuts,[1] defining a phase-type renewal process by considering instantaneous restarts of the Markov chain. In the case of multiple absorbing states, it will be seen that restarting of the process in general leads to Markov renewal processes. In reliability applications, the absorbing states will usually correspond to various types of failures or maintenance events. The main idea is here that the Markov chain is restarted in a "working state" after failure and a repair action, where the latter takes negligible time. A few associated reliability measures will be suggested.

The second approach considered in the paper adapts an idea of Lindqvist and Amundrustad,[24] who modeled periodically inspected systems by introducing maintenance by instantaneous transitions at regular points in time. Of interest in applications are, for example, whether the system is found in an absorbing state ("failure state") when inspected, and in case, the time that has been spent in such a state. Reliability measures based on such matters will be derived under consideration of stationarity of the underlying processes.

The rest of the paper is organized as follows. In Section 2, we give definitions and some main results on phase-type modeling needed in the following. The emphasis is here on the multiple absorbing state case, presented in a competing risks setting. Section 3 considers the case of restarting the underlying Markov chain upon absorption and includes a subsection on applications to maintenance, as well as a numerical example. Section 4 first reviews the main ideas for modeling of periodic inspections suggested by Lindqvist and Amundrustad[24] and then adapts them to the phase-type framework of the present paper. Some concluding remarks are given in Section 5.

2 | SOME THEORY OF PHASE-TYPE DISTRIBUTIONS

We shall let vectors and matrices be given by bold letters, respectively lowercase and uppercase. Vectors will always be assumed to be column vectors. We shall use $\mathbf{I}$ to mean the identity matrix, where the dimension will be clear from the connection. This also applies to the use of the vectors or matrices $\mathbf{0}$ which are vectors or matrices of all 0s, and the vector $\mathbf{1}$ which is a vector of all 1s. Transposes of vectors will be marked by $'$, for example, $\mathbf{p}'$. A vector of length r will for short be called an r-vector.

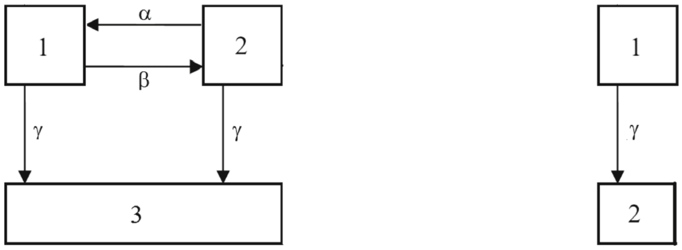

FIGURE 1 Two phase-type representations of the exponential distribution with rate γ.

2.1 Ordinary phase-type distributions

Consider a continuous-time homogeneous Markov chain $\{X(t); t \geq 0\}$, where the chain moves through some or all of m, say, transient states, or phases, before moving to a single absorbing state, $m+1$. The time of absorption, T, is then said to have a phase-type distribution.

The infinitesimal transition matrix $\mathbf{A}$ of this Markov chain is an $(m+1) \times (m+1)$ matrix given in block form as

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \boldsymbol{\ell} \\ \mathbf{0}' & 0 \end{bmatrix}. \tag{1}$$

Here $\mathbf{Q}$ is the $m \times m$ matrix corresponding to transitions between the transient states, and $\boldsymbol{\ell}$ is the m-vector defining direct transition intensities from the transient states to the absorbing state. Letting $\mathbf{P}(t)$ be the matrix of transition probabilities $P_{ij}(t) = P(X(t) = j | X(0) = i)$, it can be shown that (1) implies

$$\mathbf{P}(t) = \begin{bmatrix} e^{\mathbf{Q}t} & \mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell} \\ \mathbf{0}' & 1 \end{bmatrix}. \tag{2}$$

Now let $\mathbf{p}$ define the initial distribution of the chain, that is, let $\mathbf{p}$ be an m-vector with entries $p_i = P(X(0) = i)$ for $i = 1, \ldots, m$, summing to 1. A phase type distribution is thus determined by a representation of the form $(\mathbf{p}, \mathbf{Q})$ (Here it should be noted that the transition matrix $\mathbf{A}$ in (1) is completely determined by $\mathbf{Q}$ since necessarily $\boldsymbol{\ell} = -\mathbf{Q}\mathbf{1}$).

By using (2) we obtain the cumulative distribution function of T based on the representation $(\mathbf{p}, \mathbf{Q})$,

$$F(t) = P(T \leq t) = P(X(t) = m+1) = \mathbf{p}'\mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell}.$$

Representations $(\mathbf{p}, \mathbf{Q})$ of phase-type distributions are, however, well known to be non-unique. For example, Figure 1 illustrates two representations for an exponential distribution with rate γ, with $m = 2$ and 1, respectively, where in the first case $\mathbf{p}$ is arbitrary, while it necessarily gives mass 1 to state 1 in the second case.

This motivates the definition of *order* of a phase-type distribution, which is defined to be the minimal number m of transient states of all its representations (O'Cinneide[3]). Unique representations of phase type distributions are given through Laplace transforms. The Laplace transform for the representation $(\mathbf{p}, \mathbf{Q})$ for T is

$$f^*(s) = E\left(e^{-sT}\right) = \mathbf{p}'(s\mathbf{I} - \mathbf{Q})^{-1}(-\mathbf{Q})\mathbf{1}.$$

This is a rational function of s, that is, of the form $f^*(s) = N(s)/D(s)$ for polynomials $N(s)$ and $D(s)$. The degree of the denominator polynomial, $D(s)$, after having canceled possible equal factors in the numerator and denominator, is called the *degree* of the phase-type distribution. It can be shown[3] that the order of a phase-type distribution is at least as large as its degree, but that it may be strictly larger.

Two remarkable results by O'Cinneide[3] state that, (i) a distribution on $(0, \infty)$ is a phase-type distribution if and only if it has a strictly positive continuous density and a rational Laplace transform with a unique pole of maximal real part, and (ii) if all poles are *real*, then the distribution can be represented by a Coxian model.

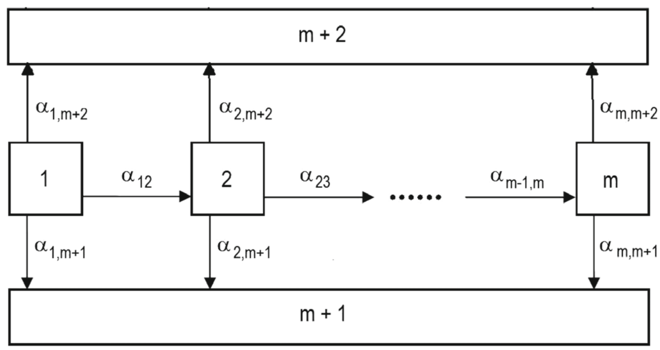

FIGURE 2 Coxian phase-type model for $K = 2$ competing risks. The figure is reproduced from Lindqvist[7] (under Creative Commons Attribution 4.0 International License).

By the non-uniqueness of representations of phase-type distributions, it is of interest to determine when two representations $(\mathbf{p}^{(a)}, \mathbf{Q}^{(a)})$ and $(\mathbf{p}^{(b)}, \mathbf{Q}^{(b)})$ lead to the same phase-type distribution in the sense of having the same Laplace transform. A key result is here a theorem given by Telek and Horvath.[25] Following,[25] we shall say that a representation $(\mathbf{p}, \mathbf{Q})$ of dimension m is *nonredundant* if the degree of the corresponding phase-type distribution is m. For intuition, we may think of nonredundancy as a way of excluding representations of phase-type distributions where there are also representations with a lower dimension, see Figure 1. As noted in Appendix A2 of Lindqvist,[7] a representation $(\mathbf{p}, \mathbf{Q})$ is nonredundant only if the minimal polynomial of $\mathbf{Q}$ equals the characteristic polynomial, where the latter condition is equivalent to $\mathbf{Q}$ being *simple* in the notion of O'Cinneide.[2] In the following we shall always consider representations $(\mathbf{p}, \mathbf{Q})$ that are nonredundant.

Theorem 1 (Telek and Horvath[25]). *Let $(\boldsymbol{p}^{(a)}, \boldsymbol{Q}^{(a)})$ and $(\boldsymbol{p}^{(b)}, \boldsymbol{Q}^{(b)})$ be nonredundant representations of phase-type distributions with the same dimension m. Let the corresponding cumulative distribution functions be $F^{(a)}(t)$ and $F^{(b)}(t)$, respectively. Then $F^{(a)}(t) = F^{(b)}(t)$ for all t if and only if there exists a nonsingular $m \times m$ matrix $\boldsymbol{B}$ with $\boldsymbol{B}\mathbf{1} = \mathbf{1}$, such that $\boldsymbol{p}^{(b)\prime} = \boldsymbol{p}^{(a)\prime}\boldsymbol{B}$ and $\boldsymbol{Q}^{(b)} = \boldsymbol{B}^{-1}\boldsymbol{Q}^{(a)}\boldsymbol{B}$.*

As noted by Telek and Horvath,[25] a general nonredundant phase-type distribution of order m can be fitted by $2m - 1$ independent parameters. This result is essentially well known and is of special interest since, apparently, a representation $(\mathbf{p}, \mathbf{Q})$ would need $m^2 + m - 1$ independent parameters.

2.2 | Phase-type models for competing risks

Consider now the extension to having $K > 1$ absorbing states in the Markov chain $\{X(t)\}$. In this setup, the Markov chain moves among m transient states before it is absorbed in one of the absorbing states, named $m + 1, m + 2, \ldots, m + K$, say. Let T be the time of absorption in any one of the absorbing states, and let C (the "cause") represent the state where absorption occurs, defining $C = j$ if $X(T) = m + j$; $j = 1, 2, \ldots, K$ (see Lindqvist[7]). Then the pair (T, C) is an observation from a classical *competing risks* model with possible causes $1, \ldots, K$. The case $K = 2$ is illustrated in Figure 2.

By extending the matrix (1) to include K absorbing states, we obtain the infinitesimal transition matrix of the modified Markov chain to be the $(m + K) \times (m + K)$ matrix given in block form as

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \mathbf{L} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{3}$$

Here $\mathbf{Q}$ as before represents transitions between the transient states, while the m-vector $\boldsymbol{\ell}$ is replaced by the $m \times K$ matrix $\mathbf{L}$ of transition intensities from the transient states to the absorbing states.

Similarly to (2), we obtain the matrix of transition probabilities $P_{ij}(t)$ given by

$$\mathbf{P}(t) = \begin{bmatrix} e^{\mathbf{Q}t} & \mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\mathbf{L} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{4}$$

Letting the m-vector $\mathbf{p}$ be the initial distribution of the Markov chain, the triple $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$ now determines a competing risks model. Such a representation will be called nonredundant if $(\mathbf{p}, \mathbf{Q})$ is nonredundant as an ordinary phase-type representation. In this case, the representation $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$ has $(K+1)m-1$ independent parameters.[7]

From (4) we obtain expressions for the *subdistribution functions* (also called cumulative incidence functions in the competing risks literature),

$$F_j(t) = P(T \le t, C = j) = P(X(t) = m + j) = \mathbf{p}'\mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell}_j$$

for $j = 1, \dots, K$, where $\mathbf{p}$ is the m-vector defining the initial distribution of the Markov chain and $\boldsymbol{\ell}_j$ is the jth column of $\mathbf{L}$.

The following result in Lindqvist[7] extends Theorem 1 to the competing risks case.

Theorem 2 (Lindqvist[7]). *Let $(\boldsymbol{p}^{(a)}, \boldsymbol{Q}^{(a)}, \boldsymbol{L}^{(a)})$ and $(\boldsymbol{p}^{(b)}, \boldsymbol{Q}^{(b)}, \boldsymbol{L}^{(b)})$ be two nonredundant phase-type representations for competing risks, having subdistribution functions $F_j^{(a)}(t)$ and $F_j^{(b)}(t)$, respectively. Then $F_j^{(a)}(t) = F_j^{(b)}(t)$ for all t and j if and only if there exists a nonsingular $m \times m$ matrix $\boldsymbol{B}$ with $\boldsymbol{B}\mathbf{1} = \mathbf{1}$ such that $\boldsymbol{p}^{(b)\prime} = \boldsymbol{p}^{(a)\prime}\boldsymbol{B}$, $\boldsymbol{Q}^{(b)} = \boldsymbol{B}^{-1}\boldsymbol{Q}^{(a)}\boldsymbol{B}$ and $\boldsymbol{L}^{(b)} = \boldsymbol{B}^{-1}\boldsymbol{L}^{(a)}$.*

2.3 | Coxian competing risks models

The class of Coxian phase-type models, as briefly described in the Introduction, can in a straightforward manner be extended to the multiple absorbing state case with representations of the form $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$, with

$$\mathbf{Q} = \begin{pmatrix} -\lambda_1 & \alpha_{12} & 0 & \cdots & 0 \\ 0 & -\lambda_2 & \alpha_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda_m \end{pmatrix}, \qquad \mathbf{L} = \begin{pmatrix} \alpha_{1,m+1} & \alpha_{1,m+2} & \cdots & \alpha_{1,m+K} \\ \alpha_{2,m+1} & \alpha_{2,m+2} & \cdots & \alpha_{2,m+K} \\ \vdots & \vdots & & \vdots \\ \alpha_{m,m+1} & \alpha_{m,m+2} & \cdots & \alpha_{m,m+K} \end{pmatrix}. \tag{5}$$

Here $\lambda_i = \sum_{j=i+1}^{m+K} \alpha_{i,j}$ for $i = 1, 2, \dots, m$. For Coxian models one usually assumes, moreover, $\mathbf{p} = (1, 0, \dots, 0)'$. Figure 2 illustrates the case when $K = 2$.

An apparently more general class of models than the Coxian models above are models given by phase-type representations with *upper triangular* $\mathbf{Q}$, so that state transitions among the transient states are always to a higher numbered state. Cumani[13] showed that any phase-type distribution involving an upper triangular $\mathbf{Q}$ can be uniquely represented by a Coxian model with $\mathbf{p} = (1, 0, \dots, 0)'$ and $\mathbf{Q}$ given as in (5) with $\lambda_1 \ge \lambda_2 \ge \dots, \ge \lambda_m$. As noted by Lindqvist,[7] this result extends to the multiple absorbing states case. Thus, informally, when modeling deteriorating systems by phase-type models, one can always restrict attention to Coxian models.

Lindqvist[7] stated the following result which is a simple consequence of Theorem 2.

Theorem 3 (Lindqvist[7]). *Consider two nonredundant Coxian phase-type distributions for competing risks, given by $(\boldsymbol{p}^{(a)}, \boldsymbol{Q}^{(a)}, \boldsymbol{L}^{(a)})$ and $(\boldsymbol{p}^{(b)}, \boldsymbol{Q}^{(b)}, \boldsymbol{L}^{(b)})$, where $\boldsymbol{p}^{(a)} = \boldsymbol{p}^{(b)} = (1, 0, \dots, 0)$. Assume further that the diagonals of $\boldsymbol{Q}^{(a)}$ and $\boldsymbol{Q}^{(b)}$ are ordered in the same way. Then if $F_j^{(a)}(t) = F_j^{(b)}(t)$ for all $t > 0$ and $j = 1, 2, \dots, K$, we have $\boldsymbol{Q}^{(b)} = \boldsymbol{Q}^{(a)}$ and $\boldsymbol{L}^{(b)} = \boldsymbol{L}^{(a)}$.*

Thus, the uniqueness of Coxian models requires the same ordering of the diagonal elements of $\mathbf{Q}$. A counterexample was given in Lindqvist.[7]

3 | INSTANTANEOUS RESTART AT ABSORPTION

3.1 | Neuts' phase-type renewal process

Consider the ordinary phase-type representation $(\mathbf{p}, \mathbf{Q})$ considered in Section 2.1. Neuts,[1] p. 48, considered restarting the Markov chain $\{X(t)\}$ after absorption by introducing state $m + 1$ as an *instantaneous* state from which an immediate transition to the set of states $\{1, 2, \dots, m\}$ occurs according to the distribution $\mathbf{p}$. As noted by Neuts,[1] by considering the version of the process that is right-continuous, this leads to a Markov chain on $\{1, 2, \dots, m\}$ with infinitesimal transition matrix

$$\mathbf{A}^* = \mathbf{Q} + \boldsymbol{\ell}\mathbf{p}'. \tag{6}$$

Moreover, the successive visits to the instantaneous state form a renewal process with the underlying phase-type distribution as its interevent distribution.

Consider now instead the multiple absorbing state case of Section 2.2, represented by the triple $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$ and having the infinitesimal matrix $\mathbf{A}$ given by (3). Suppose first that, after absorption in any absorbing state, the Markov chain is restarted in the set $\{1, 2, \dots, m\}$ according to the probability vector $\mathbf{p}$. Similarly to the above, this would result in a Markov chain $\{X(t)\}$ on $\{1, 2, \dots, m\}$ with infinitesimal transition matrix

$$\mathbf{A}^* = \mathbf{Q} + \mathbf{LS}, \tag{7}$$

where $\mathbf{S}$ is a $K \times m$ matrix with each row equal to $\mathbf{p}'$. Note that $\mathbf{S} = \mathbf{1p}'$ so that from (7) we get $\mathbf{A}^* = \mathbf{Q} + \mathbf{L1p}'$ and we are back to (6) with $\boldsymbol{\ell} = \mathbf{L1}$. Thus again, times between absorption form a renewal process. A more general and more applicable situation is, however, considered below.

3.2 | A general restarted absorbing Markov chain

Let the situation be as in the previous subsection. In applications it would be more reasonable to let the restarting distribution depend on the absorbing state, thus letting the rows of $\mathbf{S}$ be possibly different probability vectors on $\{1, 2, \dots, m\}$. Still, (7) would be the infinitesimal transition matrix of the resulting (right-continuous) Markov chain.

It is seen, however, that the successive visits to the set of absorbing states in general no longer form a renewal process. Let $T_1, T_2, \dots$ denote the times of absorption and let $X(T_1), X(T_2), \dots$ be the successive states of absorption. In the following we shall sometimes find it convenient to work also with the corresponding causes $C_1, C_2, \dots$ (see Section 2.2), such that

$$C_n = j \;\Leftrightarrow\; X(T_n) = m + j \text{ for } j = 1, 2, \dots, K.$$

Now define (C_0, T_0) by $C_0 = c_0$ if the Markov chain $\{X(t)\}$ has the initial distribution given by the c_0th row of $\mathbf{S}$ and $T_0 = 0$. Then it is seen that $(C_0, T_0), (C_1, T_1), \dots$ form a Markov renewal process (see Definition 1 in the Appendix). This is because, given the full information until the nth event, the distribution of (C_{n+1}, T_{n+1}) depends only on the last absorbing state, C_n. It also follows that $\{C_n\}$ is a Markov chain on $\{1, 2, \dots, K\}$.

Now fix one of the absorbing states, for example $m + c$, that is, cause $C = c$. It is clear that for $C_0 = c$, the succeeding visits to this state, at times $U_1, U_2, \dots$, say, form a renewal process. The times $U_1, U_2, \dots$ are likewise regeneration points of the process $\{X(t); t \geq 0\}$, which is hence a regenerative process (Definition 2 in the Appendix).

Following Cocozza-Thivent,[26,27] $\{X(t)\}$ is also a *semi*-regenerative process with respect to the Markov renewal process $\{(C_n, T_n)\}$ (Definition 3 in the Appendix). This follows since conditioning on $(C_0, T_0), (C_1, T_1), \dots, (C_n, T_n)$ with $C_n = c$, $\{X(T_n + t)\ ;\ t \geq 0\}$ has the same distribution as $\{X(t)\ ;\ t \geq 0\}$ given $C_0 = c$.

3.3 | Long run properties of reward processes connected to absorption

Cocozza-Thivent[27] showed how to apply the above facts to reward processes. In our case, suppose for example that the cost of absorption in state $m + j$ is a_j per visit. A relevant question is then what is the long run expected cost per time unit for the process. Since $U_1, U_2, \dots$ is a renewal process, the solution is well known to be the ratio of the total expected cost per renewal cycle divided by the expected length of a renewal cycle.[27,28] This is, however, not necessarily straightforward to calculate. Instead we shall use the result by Cocozza-Thivent,[27] given as Theorem 4 in the Appendix.

Consider first the fixed cause c and let

$$\Phi^{(c)}(t) = \sum_{n:T_n \leq t} I(C_n = c).$$

In words, this can be interpreted as the number of absorptions into state $m+c$ in the time interval $(0, t]$. The process $\{\Phi^{(c)}(t)\}$ now satisfies the condition of Theorem 4 in the Appendix, and hence the long run expected relative number of visits to $m+c$ is

$$\lim_{t\to\infty}\frac{E(\Phi^{(c)}(t))}{t} = \frac{E_{\boldsymbol{\pi}}(\Phi^{(c)}(T_1))}{E_{\boldsymbol{\pi}}(T_1)}, \tag{8}$$

where $\boldsymbol{\pi}$ is the stationary distribution of the Markov chain $\{C_n; n \geq 0\}$ (Here $E_{\boldsymbol{\pi}}$ means expected value with C_0 being distributed as $\boldsymbol{\pi}$).

In order to derive $\boldsymbol{\pi}$, consider the limit as t tends to infinity in (4). The upper right block of (4) then in the limit gives the probabilities of absorption in each absorbing state, for each starting state. Since necessarily $e^{\mathbf{Q}t} \to \mathbf{0}$ as $t \to \infty$, it follows that the transition matrix of $\{C_n\}$ is given by the $K \times K$ matrix

$$-\mathbf{S}\mathbf{Q}^{-1}\mathbf{L}. \tag{9}$$

Assuming this transition matrix is irreducible, we can find uniquely the stationary distribution $\boldsymbol{\pi}$ by solving

$$-\boldsymbol{\pi}'\mathbf{S}\mathbf{Q}^{-1}\mathbf{L} = \boldsymbol{\pi}',$$
$$\boldsymbol{\pi}'\mathbf{1} = 1.$$

Having solved for $\boldsymbol{\pi}$, we next consider calculation of the right hand side of (8). It is clear by stationarity that

$$E_{\boldsymbol{\pi}}(\Phi^{(c)}(T_1)) = \pi_c \tag{10}$$

with $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$.

To derive an expression for the denominator of (8), we will first calculate, for the phase-type model of Section 2.2, the expected times μ_i to absorption (in any absorbing state), given start in state $i \in \{1, 2, \ldots, m\}$. Considering (4), and recalling the general formula $E(T) = \int_0^\infty P(T > t)dt$ for a general lifetime T, we get

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_m)' = \int_0^\infty e^{\mathbf{Q}t}\mathbf{1}dt = -\mathbf{Q}^{-1}\mathbf{1}. \tag{11}$$

It follows that the expected time to absorption starting with an instantaneous transition from absorbing state $m+j$ is $-\mathbf{S}_j\boldsymbol{\mu} = -\mathbf{S}_j\mathbf{Q}^{-1}\mathbf{1}$, where $\mathbf{S}_j$ is the jth row of $\mathbf{S}$. Hence the denominator of (8) is

$$E_{\boldsymbol{\pi}}(T_1) = -\boldsymbol{\pi}'\mathbf{S}\mathbf{Q}^{-1}\mathbf{1}. \tag{12}$$

We can thus complete the calculation of the right hand side of (8) to get

$$\lim_{t\to\infty}\frac{E(\Phi^{(c)}(t))}{t} = \frac{\pi_c}{-\boldsymbol{\pi}'\mathbf{S}\mathbf{Q}^{-1}\mathbf{1}}. \tag{13}$$

Considering instead the renewal process $\{U_n\}$, it follows from the theory of renewal reward processes that the right hand side of (13) equals $1/r_c$, where r_c is the expected time between absorptions in state $m+c$. Thus (10)–(12) imply that the expected recurrence time of absorption in $m+c$ is

$$r_c = \frac{E_{\boldsymbol{\pi}}(T_1)}{\pi_c} = \frac{-\boldsymbol{\pi}'\mathbf{S}\mathbf{Q}^{-1}\mathbf{1}}{\pi_c}.$$

Going back to the more general case where the cost of absorption in state $m+j$ is a_j per visit, it is readily verified from the above that the long run expected cost per time unit is

$$\frac{\boldsymbol{\pi}'\mathbf{a}}{-\boldsymbol{\pi}'\mathbf{S}\mathbf{Q}^{-1}\mathbf{1}}, \tag{14}$$

where $\mathbf{a} = (a_1, \ldots, a_K)$.

3.3.1 | Invariance of underlying phase-type failure model

It is the purpose of the present subsection to show that even if the representations of competing risks as considered in Section 2.2 are not unique (Theorem 2), the main results (13)-(14) of the previous subsection are invariant with respect to equivalent representations of the underlying phase-type distribution. Indeed, it can be argued from Theorem 2 that the process (C_n, T_n) will have a distribution independent of the underlying phase-type models.

To be more precise, recall the construction which involves a competing risks representation $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$ extended by a matrix $\mathbf{S}$ with rows defining the starting distributions from the respective absorbing states at the restarts. Suppose now that we have two representations of this type, $(\mathbf{p}^{(a)}, \mathbf{Q}^{(a)}, \mathbf{L}^{(a)}, \mathbf{S}^{(a)})$ and $(\mathbf{p}^{(b)}, \mathbf{Q}^{(b)}, \mathbf{L}^{(b)}, \mathbf{S}^{(b)})$. Theorem 2 gives necessary and sufficient conditions that the two underlying competing risks representations are equivalent. Suppose these conditions hold. By the relation $\mathbf{p}^{(b)\prime} = \mathbf{p}^{(a)\prime}\mathbf{B}$ between initial distributions we find it natural to assume in addition,

$$\mathbf{S}^{(b)} = \mathbf{S}^{(a)}\mathbf{B}.$$

Consider now the transition matrix (9) of $\{C_n\}$. From Theorem 2 and the above follows

$$-\mathbf{S}^{(b)}\mathbf{Q}^{(b)-1}\mathbf{L}^{(b)} = -\mathbf{S}^{(a)}\mathbf{B}\mathbf{B}^{-1}\mathbf{Q}^{(a)-1}\mathbf{B}\mathbf{B}^{-1}\mathbf{L}^{(a)} = -\mathbf{S}^{(a)}\mathbf{Q}^{(a)-1}\mathbf{L}^{(a)}.$$

We can hence in particular conclude that the stationary distribution $\boldsymbol{\pi}$ of the process $\{C_n\}$ is invariant under equivalent representations of the competing risks model. Likewise, the right hand side of (12) as well as (14) are invariant, since

$$-\boldsymbol{\pi}^{(b)}\mathbf{S}^{(b)}\mathbf{Q}^{(b)-1}\mathbf{1} = -\boldsymbol{\pi}^{(a)}\mathbf{S}^{(a)}\mathbf{B}\mathbf{B}^{-1}\mathbf{Q}^{(a)-1}\mathbf{B}\mathbf{1} = -\boldsymbol{\pi}^{(a)}\mathbf{S}^{(a)}\mathbf{Q}^{(a)-1}\mathbf{1}$$

using that $\mathbf{B}\mathbf{1} = \mathbf{1}$ and that the stationary distribution $\boldsymbol{\pi}$ is invariant.

3.4 | Long run availability for Coxian models

In Section 3.3 we considered reward processes connected to absorption, that is, concerning the "failure" states of the considered process. Considering now the transient ("working states"), we shall be concerned with measuring the availability of the system.

Consider a Coxian model where we think of each state (stage) having a specific interpretation, ordered according to performance with 1 being the best. Then one might be interested in the long run expected relative amount of time the process spends in state r or lower for some $1 \le r < m$.

Thus, for some $r < m$, define

$$\Phi_r(t) = \text{time spent in states } \{1, 2, \dots, r\} \text{ in } [0, t]$$

and consider

$$\lim_{t\to\infty} \frac{E(\Phi_r(t))}{t}.$$

By Theorem 4 in the Appendix this can be calculated by considering the process up to the first semi-regenerative time, T_1, under the initial distribution $\boldsymbol{\pi}$.

Observe first that, under stationarity defined by $\boldsymbol{\pi}$, the probability of restarting the process in state $i \in \{1, 2, \dots, m\}$ is $\boldsymbol{\pi}'\mathbf{S}^i$, where $\mathbf{S}^i$ is the ith column of $\mathbf{S}$. Let now $\mathbf{S}^{[r]}$ be the $K \times r$ matrix with columns $\mathbf{S}^1, \mathbf{S}^2, \dots, S^r$. Then $\boldsymbol{\pi}'\mathbf{S}^{[r]}$ is the vector of the restarting probabilities for the r first states.

Next, given that restart is in state $i \in \{1, 2, \dots, r\}$, we need to calculate the expected time spent in $\{1, 2, \dots, r\}$. Since we now consider a Coxian model, where no transit to lower numbered states is allowed, we may use the result (11) to the modification $\mathbf{Q}^{[r]}$ of the $\mathbf{Q}$ in (5), given by letting $m = r$, or in other words, by cutting the rows and columns numbered from $r + 1$ to m in the $\mathbf{Q}$ of (5).

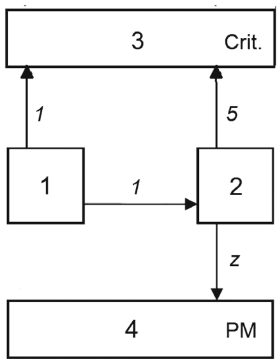

FIGURE 3 Simple phase-type model for degraded failure (state 2), critical shock failure (state 3) and PM (state 4).

Combining the above, and noting the result (11), we conclude that

$$E_{\boldsymbol{\pi}}(\Phi_r(T_1)) = -\boldsymbol{\pi}'\mathbf{S}^{[r]}\mathbf{Q}^{[r]-1}\mathbf{1}.$$

Hence, by invoking (12) and Theorem 4, we get the limiting availability result

$$\lim_{t\to\infty}\frac{E(\Phi_r(t))}{t} = \frac{\boldsymbol{\pi}'\mathbf{S}^{[r]}\mathbf{Q}^{[r]-1}\mathbf{1}}{\boldsymbol{\pi}'\mathbf{S}\mathbf{Q}^{-1}\mathbf{1}}. \tag{15}$$

Looking at (15), it is seen that the numerator in fact is the version of (12) obtained by the appropriate "pruning" of the matrices **Q** and **S** (which might have been concluded directly). It is seen, moreover, that the result in (15) trivially would equal 1 if we put $r = m$.

3.5 | Application to degradation and maintenance. A numerical example

In an application to maintenance optimization based on data from a reliability database, Langseth et al.[17] considered a Coxian model with $m = 2$ and $K = 4$. The absorbing states corresponded to, respectively, critical failure due to shock; critical failure due to deterioration; stop due to fortuitous detection of failure; preventive maintenance (PM). The authors[17] used the model to compare the effect of different degrees of PM using estimated parameters from empirical data.

In applications, the choice of the matrix **S** of restarting distributions would be an additional and important feature of the present approach. The entries of **S** must of course be depending on the application, and of the actual choice of the underlying Markov chain $\{X(t)\}$ and the interpretation of its states. The absorbing states may for example each be associated with a certain type of system failure, for which there may be a given set of repair strategies that define the corresponding row of **S**. As an example, Bedford and Lindqvist[29] considered the choice between perfect repair, minimal repair and partial repair of a system after system failure. Doyen and Gaudoin[30] considered various imperfect maintenance strategies in a competing risks setting.

Motivated by the above cited paper by Langseth et al.,[17] we consider below a simple example as illustrated in Figure 3. The figure shows a competing risks phase-type model with $m = 2$ and $K = 2$. The absorbing states 3 and 4 correspond to, respectively, critical failure and preventive maintenance (PM). State 1 is the state of perfect performance, while state 2 is a working state with degraded performance. Transition rates between the states are given in italics in the figure. As can be seen, PM is performed only from state 2. The purpose of the example is to illustrate some of the general definitions and results obtained in the previous subsections, and in particular demonstrate how the rate of PM, z, affects the performance.

The matrices **Q** and **L** given below correspond to the figure, while **S** has been chosen in the way that from state 4 (PM) there is an instant transition to state 1, while from the failure state 3 there is done a possibly partial repair, leading to a restart with probability 1/4 in state 1 and 3/4 in state 2 (somewhat arbitrarily chosen in order to give mass to both states 1 and 2). Thus,

$$\mathbf{Q}=\begin{pmatrix}-2 & 1\\ 0 & -(5+z)\end{pmatrix},\quad \mathbf{L}=\begin{pmatrix}1 & 0\\ 5 & z\end{pmatrix},\quad \mathbf{S}=\begin{pmatrix}1/4 & 3/4\\ 1 & 0\end{pmatrix}.$$

A straightforward calculation gives that the transition matrix of the Markov chain $\{C_n\}$ is (see (9)),

$$-\mathbf{SQ}^{-1}\mathbf{L}=\frac{1}{40+8z}\begin{pmatrix}40+z & 7z\\ 40+4z & 4z\end{pmatrix}$$

giving a stationary distribution $\boldsymbol{\pi}=(\pi_3,\pi_4)'$ on the absorbing states 3 and 4 as follows,

$$\pi_3=\frac{40+4z}{40+11z}\searrow,\quad \pi_4=\frac{7z}{40+11z}\nearrow.$$

Here the arrows $\nearrow$ and $\searrow$ mean that the result is respectively increasing and decreasing in z. (Note that in this example we prefer to use indices 3 and 4 for the absorbing states, instead of the *cause* numbers 1 and 2).

Next we calculate the (stationary) expected time between semi-regenerative time points,

$$E_{\boldsymbol{\pi}}(T_1)=-\boldsymbol{\pi}'\mathbf{SQ}^{-1}\mathbf{1}=\frac{12+4z}{40+11z}\nearrow.$$

Further, the expected time between critical failures (i.e., visits to state 3) is

$$r_3=\frac{E_{\boldsymbol{\pi}}(T_1)}{\pi_3}=\frac{3+z}{10+z}\nearrow$$

and the expected time between PM is

$$r_4=\frac{E_{\boldsymbol{\pi}}(T_1)}{\pi_4}=\frac{12+4z}{7z}\searrow.$$

Suppose now that cost of absorption in state 3 (critical failure) equals $w>0$ while cost of absorption in 4 (PM) equals 1. Thus, in (14) we put $\mathbf{a}=(w,1)'$ and get the long run expected cost per time unit to be

$$\frac{(40+4z)w+7z}{12+4z},\tag{16}$$

which is decreasing in z if $w>3/4$.

Let us finally consider availability of perfect performance in state 1, which is given by (15) with $r=1$. Now

$$\mathbf{S}^{[1]}=\begin{pmatrix}1/4\\ 1\end{pmatrix},\quad \mathbf{Q}^{[1]}=\begin{pmatrix}-1\end{pmatrix}$$

so

$$-\boldsymbol{\pi}'\mathbf{S}^{[1]}\mathbf{Q}^{[1]-1}\mathbf{1}=\frac{5+4z}{40+11z}$$

and the availability of state 1 is hence by (15),

$$\frac{5+4z}{12+4z}\nearrow.$$

Looking at the arrows after the expressions above, it is generally seen that increased rate of PM, z, is beneficial for the considered measures. For example, (16) shows the positive effect of PM when cost of critical failure is larger than cost of PM.

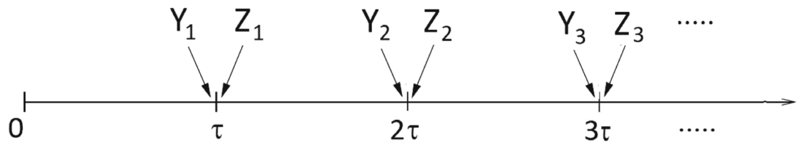

FIGURE 4 The Markov chains $\{Y_n\}$ and $\{Z_n\}$.

4 | PERIODIC INSPECTION AND MAINTENANCE

In the present section we study a slightly different extension of the competing risks phase-type models $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$ of Section 2.2. As in the previous section we introduce instantaneous transitions, but now at regular time intervals and hence not necessarily at absorption times. The motivation is from the conference paper Lindqvist and Amundrustad.[24]

4.1 | The general periodic inspection model

Following Lindqvist and Amundrustad[24] we consider periodically tested systems or components, with inspections at time $\tau, 2\tau, 3\tau, \ldots$ for a given and fixed $\tau > 0$, where maintenance is done only at these times, and with negligible repair times. Below we consider the competing risks model $(\mathbf{p}, \mathbf{Q}, \mathbf{L})$ of Section 2.2. Let now $\{X(t)\}$ be the underlying Markov chain and let it start in the usual manner from the starting state governed by the m-vector $\mathbf{p}$. At time τ the system is inspected. The probability distribution of the state at this time is $(\mathbf{p}', \mathbf{0})\ \mathbf{P}(\tau)$ with $\mathbf{P}(\cdot)$ given by (4). This state may be either a "working" state in $\{1, 2, \ldots, m\}$ or a "failure state" in $\{m+1, \ldots, m+K\}$. Upon inspection, let the state of $\{X(t)\}$ be instantaneously transformed to a new state by what we shall call a maintenance matrix $\mathbf{R}$ of dimension $(m+K) \times (m+K)$. Here $\mathbf{R}$ can in principle be any nonnegative matrix with row sums 1 (but see the special case of the next subsection). The initial distribution of the restarted process from time τ to 2τ is then $(\mathbf{p}', \mathbf{0}')\mathbf{P}(\tau)\mathbf{R}$, while the state at time 2τ, immediately before the second inspection, has distribution

$$(\mathbf{p}', \mathbf{0}')\mathbf{P}(\tau)\mathbf{R}\mathbf{P}(\tau).$$

By continuing the process in the same manner at times $3\tau, 4\tau, \ldots$, the following two Markov chains will be of particular interest:

- The Markov chain $\{Y_n\}$ with state space $\{1, 2, \ldots, m+K\}$ and transition matrix $\mathbf{R}\mathbf{P}(\tau)$, which defines the state at each inspection, that is, immediately before the instantaneous maintenance action.
- The Markov chain $\{Z_n\}$ with state space $\{1, 2, \ldots, m+K\}$ and transition matrix $\mathbf{P}(\tau)\mathbf{R}$, which defines the state immediately after the maintenance, that is, at the start of a new cycle of length τ.

Figure 4 illustrates the Markov chains $\{Y_n\}$ and $\{Z_n\}$.

In the following we shall assume that the Markov chain $\{Y_n\}$ is irreducible and aperiodic, which implies the existence of a stationary distribution $\boldsymbol{\gamma} = (\gamma_j)$ with

$$\gamma_j = \lim_{n\to\infty} P(Y_n = j) > 0 \text{ for } j = 1, 2, \ldots, m+K.$$

By the definition of $\{Z_n\}$ it is clear that also the limits

$$\rho_j = \lim_{n\to\infty} P(Z_n = j) \text{ for } j = 1, 2, \ldots, m+K$$

exist, where $\boldsymbol{\rho} = (\rho_j)$ and $\boldsymbol{\gamma}$ satisfy the two relations

$$\begin{aligned} \boldsymbol{\rho}' &= \boldsymbol{\gamma}'\ \mathbf{R}, \\ \boldsymbol{\gamma}' &= \boldsymbol{\rho}'\ \mathbf{P}(\tau). \end{aligned} \tag{17}$$

4.2 | Special case: No maintenance performed for a working system

A case of special interest is where the maintenance matrix **R** is given by

$$\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{V} & \mathbf{0} \end{bmatrix}. \tag{18}$$

Here **I** is the $m \times m$ identity matrix, while **V** is a $K \times m$ matrix with entries v_{ji} being the probability of instantaneous maintenance performed in absorbing state $m + j$ which brings the state of the system back to the working state $i \in \{1, 2, \ldots, m\}$. The identity matrix **I** means that if the state at an inspection is in one of the "working" states $\{1, 2, \ldots, m\}$, then the process is restarted in that particular state. Thus, in effect, no maintenance is performed at inspections for a working system.

Consider first the transition matrix of $\{Z_n\}$, that is, $\mathbf{P}(\tau)\mathbf{R}$, which now can be written

$$\mathbf{P}(\tau)\mathbf{R} = \begin{bmatrix} e^{\mathbf{Q}\tau} + \mathbf{Q}^{-1}\,(e^{\mathbf{Q}\tau} - \mathbf{I})\,\mathbf{LV} & \mathbf{0} \\ \mathbf{V} & \mathbf{0} \end{bmatrix}.$$

Since the columns corresponding to the states $m + 1, \ldots, m + K$ are all **0**, it is clear that the limiting probability vector $\boldsymbol{\rho}$ defined in Section 4.1 is of the form $\boldsymbol{\rho}' = (\mathbf{g}', \mathbf{0}')$, where $\mathbf{g} = (g_1, g_2, \ldots, g_m)'$ is an m-vector which is the stationary distribution corresponding to the transition matrix

$$e^{\mathbf{Q}\tau} + \mathbf{Q}^{-1}\,(e^{\mathbf{Q}\tau} - \mathbf{I})\,\mathbf{LV}.$$

The g_i are hence obtained by solving the equations

$$\mathbf{g}' = \mathbf{g}' \left(e^{\mathbf{Q}\tau} + \mathbf{Q}^{-1}\,(e^{\mathbf{Q}\tau} - \mathbf{I})\,\mathbf{LV} \right) \tag{19}$$

$$\mathbf{g}'\mathbf{1} = 1. \tag{20}$$

Consider next the stationary distribution $\boldsymbol{\gamma}$ of $\{Y_n\}$. Now write this as

$$\boldsymbol{\gamma}' = (\boldsymbol{\eta}', \boldsymbol{\nu}'),$$

where $\boldsymbol{\eta}$ is the m-vector corresponding to the states $\{1, 2, \ldots, m\}$. Then by (17), $(\boldsymbol{\eta}', \boldsymbol{\nu}') = (\mathbf{g}', \mathbf{0}')\,\mathbf{P}(\tau)$, so by (4) we get

$$\begin{aligned} \boldsymbol{\eta}' &= \mathbf{g}' e^{\mathbf{Q}\tau}, \\ \boldsymbol{\nu}' &= \mathbf{g}'\mathbf{Q}^{-1}(e^{\mathbf{Q}\tau} - \mathbf{I})\,\mathbf{L} \end{aligned} \tag{21}$$

which are hence obtained from **g**.

4.2.1 | Efficiency of maintenance.

The elements of $\boldsymbol{\nu}$ in (21) can be interpreted as the long run expected relative number of inspections where the process is found to be in the corresponding "failure" state. They will typically depend on the maintenance interval τ, and are hence useful as measures of the efficiency of maintenance.

4.2.2 | Expected downtime between inspections.

Of interest in applications is also the expected amount of time in an inspection interval of length τ that the system spends in a given absorbing state. Following Lindqvist and Amundrustad,[24] the expected amount of time in the interval $(n\tau, (n +$

$1)\tau]$ spent in state $m+j$, is

$$\int_{n\tau}^{(n+1)\tau} P(X(t)=m+j)\;dt = \int_0^{\tau} \sum_{i=1}^{m} P(Z_n=i)\;P_{i,m+j}(t)\;dt$$

which converges as $n\to\infty$ to

$$\int_0^{\tau} \sum_{i=1}^{m} \mathbf{g}_i\;P_{i,m+j}(t)\;dt. \tag{22}$$

Here, $P_{i,m+j}(t)$ can be found from (4). It follows that (22) is the jth entry of the K-vector

$$\mathbf{g}' \int_0^{\tau} \mathbf{Q}^{-1}(e^{\mathbf{Q}t}-\mathbf{I})\mathbf{L}dt = \mathbf{g}'[\mathbf{Q}^{-2}(e^{\mathbf{Q}\tau}-\mathbf{I})-\tau\mathbf{I}]\;\mathbf{L}, \tag{23}$$

which hence measures the expected length of the system's downtime in an inspection interval, for each of the absorbing states $m+1,\;\dots\;,m+K$.

4.2.3 | Uniqueness of maintenance measures

The situation considered above can be characterized by the quadruple $(\mathbf{p},\mathbf{Q},\mathbf{L},\mathbf{V})$, with $\mathbf{V}$ given by (18).

Suppose now that we start out by two equivalent competing risks representations $(\mathbf{p}^{(a)},\mathbf{Q}^{(a)},\mathbf{L}^{(a)})$ and $(\mathbf{p}^{(b)},\mathbf{Q}^{(b)},\mathbf{L}^{(b)})$ with relations between the $\mathbf{p}^{(\cdot)},\mathbf{Q}^{(\cdot)}$ and $\mathbf{L}^{(\cdot)}$ through a matrix $\mathbf{B}$ as in Theorem 2. Consider then the corresponding representations $(\mathbf{p}^{(a)},\mathbf{Q}^{(a)},\mathbf{L}^{(a)},\mathbf{V}^{(a)})$ and $(\mathbf{p}^{(b)},\mathbf{Q}^{(b)},\mathbf{L}^{(b)},\mathbf{V}^{(b)})$, and assume

$$\mathbf{V}^{(b)} = \mathbf{V}^{(a)}\mathbf{B}, \tag{24}$$

which is motivated by the relation between $\mathbf{p}^{(a)}$ and $\mathbf{p}^{(b)}$ in Theorem 2.

In the same manner as what we did in Section 3.3.1 we now show that the key results (21) and (23) are not affected by the choice of representation $(\mathbf{p},\mathbf{Q},\mathbf{L})$ of the initial competing risks model.

Let $\mathbf{g}^{(a)}$ and $\mathbf{g}^{(b)}$ be the (uniquely given) versions of $\mathbf{g}$ as defined by (19)-(20) corresponding to the two representations. A key result is here is that

$$\mathbf{g}^{(a)\prime} = \mathbf{g}^{(b)\prime}\mathbf{B}^{-1}. \tag{25}$$

To prove this, reconsider the equations (19)-(20) for $\mathbf{g}^{(b)}$,

$$\begin{aligned} \mathbf{g}^{(b)\prime} &= \mathbf{g}^{(b)\prime}\left(e^{\mathbf{Q}^{(b)}\tau} + \mathbf{Q}^{(b)-1}\;(e^{\mathbf{Q}^{(b)}\tau}-\mathbf{I})\;\mathbf{L}^{(b)}\mathbf{V}^{(b)}\right) \\ \mathbf{g}^{(b)\prime}\mathbf{1} &= 1 \end{aligned} \tag{26}$$

Multiplying equation (26) from the right by $\mathbf{B}^{-1}$ and using the relations from Theorem 2 as well as (24), it is seen that equation (26) is transformed to

$$\mathbf{g}^{(b)\prime}\mathbf{B}^{-1} = \mathbf{g}^{(b)\prime}\mathbf{B}^{-1}\left(e^{\mathbf{Q}^{(a)}\tau} + \mathbf{Q}^{(a)-1}\;(e^{\mathbf{Q}^{(a)}\tau}-\mathbf{I})\;\mathbf{L}^{(a)}\mathbf{V}^{(a)}\right) \tag{27}$$

It is also clear that $\mathbf{g}^{(b)\prime}\mathbf{B}^{-1}\mathbf{1}=1$ since $\mathbf{B1}=\mathbf{1}$ implies $\mathbf{B}^{-1}\mathbf{1}=\mathbf{1}$. Now (25) follows from (27) by uniqueness of the solution for $\mathbf{g}$ of (19)-(20).

By using (25) and the relations from Theorem 2 it is straightforward to show that the key results (21) and (23) are invariant with respect to the original competing risks representations.

4.3 | Modifying the maintenance matrix R

Lindqvist and Amundrustad[24] considered certain variations of the above approach, consisting in modifications of the maintenance matrix **R** in (18). One such modification would be to replace the identity matrix **I** in the upper left corner in order to allow repairs or other interventions in the system when it is in working condition at inspections. Suppose for example that state 1 means a perfectly working system. Then if a perfect repair of the system is performed at each inspection both for a failed and a working system, the matrix **I** as well as **V** in (18) would be replaced by matrices with first column consisting of all 1s and 0s otherwise. Such an assumption on perfect maintenance is essentially made in the earlier cited paper by Wu and Cui.[23]

Lindqvist and Amundrustad also considered the possibility of imperfect repair from the absorbing states. This means that the matrix **0** in the lower right corner of (18) is replaced by a matrix with some entries being positive.[24] We will not pursue this here but refer to examples in the conference paper.[24]

5 | CONCLUDING REMARKS

We have considered phase-type models for competing risks with extensions obtained by introducing instantaneous transitions, (i) in the form of restarting the Markov chain after absorption, or (ii) including periodic instantaneous interventions. The motivation has been from modeling of deteriorating and maintained systems, but other applications may of course be thought of.

Since phase-type models for lifetime distributions or more generally for competing risks distributions are non-unique,[7] an interesting conclusion in the present paper is that certain natural measures pertaining to the occurrence or cost of failures, or dormant failures of inspected systems, are invariant with respect to representations of the underlying failure distributions.

The main purpose of the paper has been to suggest and sketch a theoretical framework which involves an extension of classical phase-type modeling. There is hence room for much further research in order to consider specific applications and new adaptations of the approach.

ACKNOWLEDGMENT

The author would like to thank the Associate Editor and two referees for their constructive comments and suggestions.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Bo Henry Lindqvist https://orcid.org/0000-0001-8952-9311

REFERENCES

1. Neuts MF. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press; 1981.
2. O'Cinneide CA. On non-uniqueness of representations of phase-type distributions. *Stoch Model*. 1989;5(2):247-259.
3. O'Cinneide CA. Characterization of phase-type distributions. *Stoch Model*. 1990;6(1):1-57.
4. O'Cinneide CA. Triangular order of triangular phase-type distributions. *Stoch Model*. 1993;9(4):507-529.
5. Aalen OO. Phase type distributions in survival analysis. *Scand J Stat*. 1995;22(4):447-463.
6. Slud EEV, Suntornchost J. Parametric survival densities from phase-type models. *Lifetime Data Anal*. 2014;20(3):459-480.
7. Lindqvist BH. Phase-type models for competing risks, with emphasis on identifiability issues. *Lifetime Data Anal*. 2023;29:318-341.
8. Cox DR. A use of complex probabilities in the theory of stochastic processes. *Math Proc Camb Philos Soc*. 1955;51(2):313-319.
9. Faddy M, Graves N, Pettitt A. Modeling length of stay in hospital and other right skewed data: comparison of phase-type, gamma and log-normal distributions. *Value Health*. 2009;12(2):309-314.
10. McGrory CA, Pettitt AN, Faddy MJ. A fully Bayesian approach to inference for Coxian phase-type distributions with covariate dependent mean. *Comput Stat Data Analy*. 2009;53(12):4311-4321.
11. Tang X, Luo Z, Gardiner JC. Modeling hospital length of stay by Coxian phase-type regression with heterogeneity. *Stat Med*. 2012;31(14):1502-1516.
12. Rizk J, Walsh W, Burke K. An alternative formulation of Coxian phase-type distributions with covariates: application to emergency department length of stay. *Stat Med*. 2021;40(6):1574-1592.

13. Cumani A. On the canonical representation of homogeneous Markov processes modelling failure-time distributions. *Microelectron Reliab*. 1982;22(3):583-602.
14. Neuts MF, Meier KS. On the use of phase type distributions in reliability modelling of systems with two components. *Operat Res Spekt*. 1981;2(4):227-234.
15. Neuts MF, Pérez-Ocón R, Torres-Castro I. Repairable models with operating and repair times governed by phase type distributions. *Adv Appl Probab*. 2000;32(2):468-479.
16. Hokstad P, Frøvig AT. The modelling of degraded and critical failures for components with dormant failures. *Reliab Eng Syst Saf*. 1996;51(2):189-199.
17. Langseth H, Haugen K, Sandtorv H. Analysis of OREDA data for maintenance optimisation. *Reliab Eng Syst Saf*. 1998;60(2):103-110.
18. Pham H, Suprasad A, Misra RB. Availability and mean life time prediction of multistage degraded system with partial repairs. *Reliab Eng Syst Saf*. 1997;56(2):169-173.
19. McClean S, Garg L, Barton M, Fullerton K. Using mixed phase-type distributions to model patient pathways. *23rd International Symposium on Computer-Based Medical Systems (CBMS)*. Los Alamitos, CA: IEEE Computer Society; 2010:172-177.
20. Llorca J, Delgado-Rodríguez M. Competing risks analysis using Markov chains: impact of cerebrovascular and ischaemic heart disease in cancer mortality. *Int J Epidemiol*. 2001;30(1):99-101.
21. Abner EL, Charnigo RJ, Kryscio RJ. Markov chains and semi-Markov models in time-to-event analysis. *J Biomet Biostat*. 2013;S1:e001.
22. Garcia-Maya B, Limnios N, Lindqvist BH. Competing risks modeling by extended phase-type semi-Markov distributions. *Methodol Comput Appl Probab*. 2022;24:309-319.
23. Wu B, Cui L. Reliability analysis of periodically inspected systems with competing risks under Markovian environments. *Comput Ind Eng*. 2021;158:107415.
24. Lindqvist BH, Amundrustad H. Markov models for priodically tested components. *Safety and Reliability. Proceedings of the European Conference on Safety and Reliability-ESREL*. A.A. Balkema; 1998:191-197.
25. Telek M, Horváth G. A minimal representation of Markov arrival processes and a moments matching method. *Perform Eval*. 2007;64(9-12):1153-1168.
26. Cocozza-Thivent C. *Markov Renewal and Piecewise Deterministic Processes*. Springer Nature; 2021.
27. Cocozza-Thivent C. Convergence de fonctionnelles de processus semirégénératifs (in French), Prépublications de l'Université de Marne la Vallée 02/2000. 2000:2.
28. Ross SM. *Stochastic Processes*. Second ed. Wiley; 1995.
29. Bedford T, Lindqvist BH. The identifiability problem for repairable systems subject to competing risks. *Adv Appl Probab*. 2004;36(3):774-790.
30. Doyen L, Gaudoin O. Imperfect maintenance in a generalized competing risks framework. *J Appl Probab*. 2006;43(3):825-839.
31. Asmussen S. *Applied Probability and Queues*. 2nd ed. Springer; 2003.
32. Grall A, Dieulle L, Bérenguer C, Roussignol M. Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE Trans Reliab*. 2002;51(2):141-150.

How to cite this article: Lindqvist BH. Phase-type models involving restarting and instantaneous transitions, with applications to degradation and maintenance. *Appl Stochastic Models Bus Ind*. 2023;1-16. doi: 10.1002/asmb.2771

APPENDIX

The definitions and the theorem below are taken from Cocozza[26,27] and Asmussen.[31] See also the application in Grall.[32]

Definition 1. A stochastic process $\{(Y_n, T_n); n \geq 0\}$ with state space $\mathcal{Y} \times [0, \infty)$, with $T_0 = 0$ is called a *Markov renewal process* if for any $\mathcal{A} \subset \mathcal{Y}$ and $t > 0$,

$$\begin{aligned} &P(Y_{n+1} \in A, T_{n+1} - T_n \leq t \mid Y_0, \ldots, Y_n = y, T_0, \ldots, T_n) \\ &\quad = P(Y_{n+1} \in A, T_{n+1} - T_n \leq t \mid Y_n = y). \end{aligned} \tag{A1}$$

Moreover, the process is said to be a *homogeneous* Markov renewal process if equation (A1) is independent of n.

Definition 2. A stochastic process $\{X(t);\ t \geq 0\}$ is said to be a *regenerative* process if there exists a renewal process $\{U_n; n \geq 0\}$ with $U_0 = 0$ and $\sup_n U_n = \infty$ such that the process $\{X(U_n + t); t \geq 0\}$ is independent of $U_0, U_1, \ldots, U_n$ and has the same distribution as the process $\{X(t); t \geq 0\}$. The process $\{U_n; n \geq 0\}$ is called the *embedded* renewal process associated with $\{X(t); t \geq 0\}$.

Definition 3. A stochastic process $\{X(t); t \geq 0\}$ is said to be a *semi-regenerative* process if there exists a Markov renewal process $\{(Y_n, T_n); n \geq 0\}$ with $T_0 = 0$ and $\sup_n T_n = \infty$, such that the process $\{X(T_n + t); t \geq 0\}$ conditioning on $T_0, \ldots, T_n, Y_0, \ldots, Y_n = y)$ has the same distribution as the process $\{X(t); t \geq 0\}$ given $Y_0 = y$. The process $\{(Y_n, T_n); n \geq 0\}$ is called the *embedded* Markov renewal process associated with $\{X(t); t \geq 0\}$.

Theorem 4 (Cocozza-Thivent[26,27]). *Let $\{X(t); t \geq 0\}$ be a stochastic process with state space $\mathcal{X}$ which is both semi-regenerative with semi-regeneration times $\{T_n, n \geq 0\}$ and regenerative with regeneration times $\{U_n, n \geq 0\}$. Let $\{Y_n; n \geq 0\} = \{X(T_n); n \geq 0\}$ be the embedded Markov process with stationary distribution $\boldsymbol{\pi}$. Furthermore define $\Phi = \{\Phi(t); t \geq 0\}$, with $\Phi(0) = 0$, as a positive and increasing stochastic process with the properties that*

- $\Phi(t) = \Psi_t(X(u); 0 \leq u \leq t)$
- $\Phi(t) - \Phi(s) = \Psi_{t-s}(X(u); s \leq u \leq t), \quad 0 \leq s \leq t$

for some nonnegative function Ψ_t. If for any $t > 0$,

$$E(\Phi(t)) < \infty, \quad E(\Phi(U_1)) < \infty,$$

then

$$\lim_{t\to\infty} \frac{E(\Phi(t))}{t} = \frac{E_{\boldsymbol{\pi}}(\Phi(T_1))}{E_{\boldsymbol{\pi}}(T_1)}.$$