

# Bounding Steady-state Availability Models with Group Repair and Phase Type Repair Distributions

Juan A. Carrasco  
Departament d'Enginyeria Electrònica  
Universitat Politècnica de Catalunya  
Diagonal 647, plta. 9  
08028 Barcelona, Spain  
juan.a.carrasco@upc.edu

Except for formatting details, this version matches exactly the version published with the same title and authors in *Performance Evaluation*, vol. 35, no. 3–4, 1999, pp. 193–214

## Abstract

We propose an algorithm to obtain bounds for the steady-state availability using Markov models in which only a small portion of the state space is generated. The algorithm is applicable to models with group repair and phase type repair distributions and involves the solution of only four linear systems of the size of the generated state space, independently on the number of “return” states. Numerical examples are presented to illustrate the algorithm and compare it with a previous bounding algorithm.

**Keywords:** Fault-tolerant systems; Repairable systems; Steady-state availability; Bounds; Group repair; Phase type distributions

# 1 Introduction

A major drawback of continuous-time Markov chain (CTMC) models is that they usually have state space cardinalities which are far beyond the available computational resources. An approach which has been developed in the last few years is the use of bounding algorithms which require the generation of only a portion of the state space [2, 3, 5, 6, 8, 9, 10, 12, 17, 18]. Those algorithms perform well when, as in the case of availability models, the probability mass is concentrated in a small portion of the state space. The first of such algorithms was developed by Muntz et al. [12] using results from Courtois and Semal [5, 6] concerning bounds for conditional steady-state distributions in subsets of Markov chains. Let  $N$  be the number of components of the system. Denoting by  $C_k$ ,  $0 \leq k \leq N$ , the subset of states with exactly  $k$  failed components, by  $G$  the subset of generated states, and by  $U$  the subset of non-generated states, the basic algorithm proposed in [12] takes  $G = \cup_{0 \leq k \leq K} C_k$  and bounds the behavior in  $U$  using a submodel with states  $c_k$ ,  $K < k \leq N$  associated to the subsets  $C_k$ . This basic algorithm requires the solution of  $|C_K|$  linear systems of size  $|G| + N - K$ , which is typically very costly. In order to reduce the computational cost of the algorithm a state cloning technique is developed in [12] which introduces some looseness in the bounds but reduces the number of linear system to be solved to  $|C_F|$ , where  $0 \leq F < K$ . Lui and Muntz [8] have proposed a refinement of the algorithm for the particular case  $F = 0$  including a reuse technique which, at the price of an additional looseness in the bounds, avoids a complete reapplication of the algorithm each time  $K$  is incremented in the search for the desired accuracy. The additional looseness has been reduced in another paper from the same authors [9]. Souza e Silva and Ochoa [18] have developed state space exploration techniques in which  $G$  is generated incrementally following heuristics which try to obtain the tightest possible bounds for a given number of generated states. Semal has developed recently [17] a bounding algorithm which refines iteratively the bounds using detailed knowledge about the model in  $U$  in the proximities of  $G$ . In [2] a bounding algorithm is developed which exploits the failure distance concept to bound the behavior in  $U$  more tightly than in [12]. State space exploration techniques specifically targeted to that bounding algorithm have also been developed [3]. Finally, the algorithm described in [12] has been extended in [10] to models with infinite state spaces and subsets  $C_k$ ,  $k > K$  in which no every state has a transition to the left (subset  $C_{k-1}$ ). Performance models were considered in [10] and the bounding part of the model was found using special developments for the models under consideration.

All previous algorithms to bound the steady-state availability assume that repair actions involve just one component and assume exponential repair time distributions (the only exception being the machine repair model considered in [11], an extended version of [10], but the developments were specific for the considered model). In this paper we develop a new bounding algorithm for a larger class of models of repairable fault-tolerant systems which allow group repair (the simultaneous repair of several components) and phase type repair time distributions. The algorithm generates the subset of states  $G = \cup_{0 \leq k \leq K} C_k$  and computes the bounds without using state cloning techniques by solving only four linear systems of size  $|G|$ . The rest of the paper is organized as follows. Section 2 describes the type of models considered. Section 3 describes the bounding algorithm.

Section 4 compares the efficiency of the algorithm with that of the algorithm proposed in [12] using an example without group repair and with exponential repair time distributions and illustrates the extended range of applicability of the proposed algorithm using an example with group repair and phase type repair time distributions. Section 5 concludes the paper.

## 2 Type of models and assumed knowledge

We consider fault-tolerant systems made up of components which fail and are repaired. Failure processes have exponential distributions; repair processes have phase type distributions [13]. Components are grouped into types, the components of the same type being indistinguishable, and thus collections of components will be bags of component types (see, for instance [14] for a brief summary of bag theory). Any bag of component types which can fail simultaneously will be called *failure bag*. Let  $E$  be the set of failure bags of the model. In general, we will assume that for each component type there is a collection of failed modes in which the components of the type may fail. Failed modes differ in how the failed components are repaired. Failure bags may occur with rates which depend on the bag of failed component types and the failed modes of the failed components. We will assume known  $E$  and, for each  $e \in E$ , an upper bound  $[\lambda(e)]_{\text{ub}}$  for its rate. Repair actions can involve any bag of failed component types. We will denote by  $\beta$  the maximum cardinality of the bags of component types involved in repair actions. Each repair action  $i$  has a repair time phase type distribution  $P_i$ . Each phase type distribution  $P_i$  is defined by a transient CTMC  $Z_i = \{Z_i(t), t \geq 0\}$  with finite state space  $L_i \cup \{a\}$ , where all states in  $L_i$  are transient,  $a$  is an absorbing state and  $P[Z_i(0) \in L_i] = 1$ : the repair time is the time to absorption of  $Z_i$ . We allow repair interruption. Thus, the failure of a component of higher repair priority may preempt an undergoing repair process; the repair process may be resumed later from the point it was stopped (preemptive-resume) or retaken as it had just started (preemptive-restart). The state of the system can be completely described by giving the number of unfailed components of each type, the number of failed components of each type in each failed mode, the set of scheduled repair actions, which of them are active (in progress), and for each scheduled repair action  $i$  the state  $a \in L_i$  of the corresponding phase type repair distribution  $P_i$ .

We will denote by  $X = \{X(t), t \geq 0\}$  the resulting CTMC model and by  $\Omega$  its state space. Let  $N$  be the number of components of the system and let  $C_k$  be the subset of  $\Omega$  including the states with  $k$  failed components. As in [12] we will take  $G = \cup_{0 \leq k \leq K} C_k$  and, accordingly,  $U = \cup_{K < k \leq N'} C_k$ , where  $K < N' \leq N$ . According to the assumed type of state description, we will have  $|C_0| = 1$  and will denote by  $o$  the only state belonging to  $C_0$ . We will assume that some repair process is active in every state with failed components. Thus,  $o$  will be the only state without active repair processes and  $X$  will be irreducible.

### 3 Bounding algorithm

#### 3.1 Preliminaries

Although our bounding algorithm is mainly addressed to the computation of bounds for the steady-state availability, it can, in fact, be used to bound any steady-state reward rate measure. Let  $r_i, i \in \Omega$  be an arbitrary reward rate structure defined over  $X$ . We are interested in bounding the steady-state reward rate

$$R = \lim_{t \rightarrow \infty} E[r_{X(t)}] = \sum_{i \in \Omega} r_i p_i,$$

where  $p_i = \lim_{t \rightarrow \infty} P[X(t) = i]$ . The steady-state availability is a particular case of  $R$  in which  $r_i = 1$  for the up (operational) states and  $r_i = 0$  for the down (non-operational) states. Let  $S = \cup_{\max\{0, K+1-\beta\} \leq k \leq K} C_k$  be the subset of states in  $G$  which may have some transition from  $U$  (the so-called ‘‘return’’ subset), and for each  $s \in S$  consider the CTMC  $X_s = \{X_s(t), t \geq 0\}$  obtained from  $X$  by redirecting to  $s$  all transitions from  $U$  to  $S$ . Consider the regenerative behavior of  $X_s$  with  $X_s(0) = s$  ( $X_s$  may be in general non-irreducible) defined by the times at which  $X_s$  hits  $s$  from  $U$ . In this section we obtain lower and upper bounds for  $R_s$  expressed in terms of metrics related to  $X_s, s \in S$ . Our bounding algorithm is based on these bounds. The bounds are identical to those obtained in [12], but are expressed in a way from which our bounding algorithm follows naturally. Also, the bounds are justified using semi-regenerative and regenerative Markov process theory instead of results from Courtois and Semal [5, 6], as it was done in [12]. Let  $T_s$  and  $C_s$  be, respectively, the mean time and mean reward of  $X_s$  between regenerations. Using semi-regenerative Markov process theory [4, Section 10.6] we have:

**Theorem 1.** *There exist  $\beta_s, s \in S$  with  $\beta_s > 0, \sum_{s \in S} \beta_s = 1$  such that  $R = (\sum_{s \in S} \beta_s C_s) / (\sum_{s \in S} \beta_s T_s)$ .*

Let  $R_s$  the steady-state reward rate of  $X_s$  with  $X_s(0) = s$ , i.e.

$$R_s = \lim_{t \rightarrow \infty} E[r_{X_s(t)} | X_s(0) = s],$$

we have:

**Corollary 1.**  $\min_{s \in S} \{R_s\} \leq R \leq \max_{s \in S} \{R_s\}$ .

*Proof.* From regenerative process theory (see, for instance, [15]) we have  $R_s = C_s/T_s$ . Let  $n = |S|$  and assume the states in  $S$  numbered from 1 to  $n$ . The proof is by induction on  $n$ . The case  $n = 1$  is trivial. Consider the case  $n = 2$ . Using  $\beta_1 + \beta_2 = 1$  we have

$$R = \frac{\beta_1 C_1 + \beta_2 C_2}{\beta_1 T_1 + \beta_2 T_2} = \frac{C_2 + \beta_1 (C_1 - C_2)}{T_2 + \beta_1 (T_1 - T_2)}.$$

We have  $dR/d\beta_1 = (C_1 T_2 - C_2 T_1) / [T_2 + \beta_1 (T_1 - T_2)]^2$ . Note that  $T_2 + \beta_1 (T_1 - T_2) = \beta_1 T_1 + \beta_2 T_2 > 0$ . It follows that  $dR/d\beta_1 \geq 0$  if and only if  $C_1 T_2 - C_2 T_1 \geq 0$ . This implies that  $R$

is either monotonically increasing or monotonically decreasing on  $\beta_1$  and that either its maximum is  $R(1) = C_1/T_1$  and its minimum is  $R(0) = C_2/T_2$  or its maximum is  $R(0) = C_2/T_2$  and its minimum is  $R(1) = C_1/T_1$ . This completes the case  $n = 2$ .

We proceed with the induction step for  $R \leq \max_{s \in S} \{R_s\}$ . Without loss of generality assume  $\max_{1 \leq s \leq n} \{C_s/T_s\} = C_1/T_1$ . Assume that there exist  $\beta_1, \dots, \beta_n$  with  $\beta_s > 0$ ,  $\sum_{s=1}^n \beta_s = 1$  for which  $R > C_1/T_1$ . We can write

$$R = \frac{\beta_1 C_1 + (1 - \beta_1) \sum_{s=2}^n q_s C_s}{\beta_1 T_1 + (1 - \beta_1) \sum_{s=2}^n q_s T_s}, \quad (1)$$

with  $q_s > 0$ ,  $\sum_{s=2}^n q_s = 1$ . Using the induction hypothesis for  $n = 2$ ,  $R > C_1/T_1$  implies  $\sum_{s=2}^n q_s C_s / \sum_{s=2}^n q_s T_s > C_1/T_1$ . Using the induction hypothesis for  $n - 1$ , there must exist  $i$ ,  $2 \leq i \leq n$ , such that  $C_i/T_i \geq \sum_{s=2}^n q_s C_s / \sum_{s=2}^n q_s T_s$ , which implies  $C_i/T_i > C_1/T_1$ , a contradiction. Therefore, we have  $R \leq C_1/T_1 = \max_{1 \leq s \leq n} \{C_s/T_s\}$ .

Similarly, for the induction step for  $R \geq \min_{s \in S} \{R_s\}$ , let  $\min_{1 \leq s \leq n} \{C_s/T_s\} = C_1/T_1$  and assume that there exist  $\beta_1, \dots, \beta_n$  with  $\beta_s > 0$ ,  $\sum_{s=1}^n \beta_s = 1$  for which  $R < C_1/T_1$ . Considering again (1) with  $q_s > 0$ ,  $\sum_{s=2}^n q_s = 1$  and using the induction hypothesis for  $n = 2$ ,  $R < C_1/T_1$  implies  $\sum_{s=2}^n q_s C_s / \sum_{s=2}^n q_s T_s < C_1/T_1$ . Using the induction hypothesis for  $n - 1$ , there must exist  $i$ ,  $2 \leq i \leq n$ , such that  $C_i/T_i \leq \sum_{s=2}^n q_s C_s / \sum_{s=2}^n q_s T_s$ , which implies  $C_i/T_i < C_1/T_1$ , a contradiction. Therefore, we have  $R \geq C_1/T_1 = \min_{1 \leq s \leq n} \{C_s/T_s\}$ .  $\square$

Corollary 1 allows us to compute lower ( $[R]_{\text{lb}}$ ) and upper ( $[R]_{\text{ub}}$ ) bounds for  $R$  from lower ( $[R_s]_{\text{lb}}$ ) and upper ( $[R_s]_{\text{ub}}$ ) bounds for  $R_s$ ,  $s \in S$ :

$$[R]_{\text{lb}} = \min_{s \in S} \{[R_s]_{\text{lb}}\}, \quad (2)$$

$$[R]_{\text{ub}} = \max_{s \in S} \{[R_s]_{\text{ub}}\}. \quad (3)$$

Let  $T_{G,s}$  and  $T_{U,s}$  ( $C_{G,s}$  and  $C_{U,s}$ ) be the contributions of, respectively, the states in  $G$  and  $U$  to  $T_s$  ( $C_s$ ). We have ( $R_s = C_s/T_s$  from regenerative process theory)

$$R_s = \frac{C_{G,s} + C_{U,s}}{T_{G,s} + T_{U,s}}.$$

Assume that  $C_{G,s}$ ,  $T_{G,s}$ , an upper bound  $[T_{U,s}]_{\text{ub}}$  for  $T_{U,s}$ , and lower and upper bounds  $[r]_{\text{lb}}$  and  $[r]_{\text{ub}}$  for  $r_i$ ,  $i \in \Omega$  are known (for the steady-state availability we would take  $[r]_{\text{lb}} = 0$  and  $[r]_{\text{ub}} = 1$ ). We have:

**Theorem 2.**

$$[R_s]_{\text{lb}} = \frac{C_{G,s} + [r]_{\text{lb}}[T_{U,s}]_{\text{ub}}}{T_{G,s} + [T_{U,s}]_{\text{ub}}} \leq R_s \leq \frac{C_{G,s} + [r]_{\text{ub}}[T_{U,s}]_{\text{ub}}}{T_{G,s} + [T_{U,s}]_{\text{ub}}} = [R_s]_{\text{ub}}. \quad (4)$$

*Proof.* Consider the function  $f_1(x) = (C_{G,s} + [r]_{\text{ub}}x)/(T_{G,s} + x)$ . Since  $[r]_{\text{ub}}$  upper bounds the reward rate from any state of  $X_s$ , we have  $C_{G,s} \leq [r]_{\text{ub}}T_{G,s}$  and  $df_1/dx = ([r]_{\text{ub}}T_{G,s} - C_{G,s})/(T_{G,s} + x)^2 \geq 0$ . Also,  $C_{U,s} \leq [r]_{\text{ub}}T_{U,s}$ . Then

$$R_s \leq \frac{C_{G,s} + [r]_{\text{ub}}T_{U,s}}{T_{G,s} + T_{U,s}} = f_1(T_{U,s}) \leq f_1([T_{U,s}]_{\text{ub}}) = \frac{C_{G,s} + [r]_{\text{ub}}[T_{U,s}]_{\text{ub}}}{T_{G,s} + [T_{U,s}]_{\text{ub}}}.$$

Similarly, consider the function  $f_2(x) = (C_{G,s} + [r]_{\text{lb}}x)/(T_{G,s} + x)$ . Since  $[r]_{\text{lb}}$  lower bounds the reward rate from any state of  $X_s$ , we have  $C_{G,s} \geq [r]_{\text{lb}}T_{G,s}$  and  $df_2/dx = ([r]_{\text{lb}}T_{G,s} - C_{G,s})/(T_{G,s} + x)^2 \leq 0$ . Also,  $C_{U,s} \geq [r]_{\text{lb}}T_{U,s}$ . Then

$$R_s \geq \frac{C_{G,s} + [r]_{\text{lb}}T_{U,s}}{T_{G,s} + T_{U,s}} = f_2(T_{U,s}) \geq f_2([T_{U,s}]_{\text{ub}}) = \frac{C_{G,s} + [r]_{\text{lb}}[T_{U,s}]_{\text{ub}}}{T_{G,s} + [T_{U,s}]_{\text{ub}}}. \quad \square$$

### 3.2 Derivation of $[T_{U,s}]_{\text{ub}}$

In the rest of the paper we will denote by  $\lambda_{ij}$ ,  $i, j \in \Omega$  the transition rate from state  $i$  to state  $j$ , by  $\lambda_i = \sum_{j \in \Omega} \lambda_{ij}$ ,  $i \in \Omega$  the output rate of state  $i$ , and by  $\lambda_{iC} = \sum_{j \in C} \lambda_{ij}$ ,  $i \in \Omega$ ,  $C \subset \Omega$  the transition rate from  $i$  to subset  $C$ , all of them referred to  $X$ , unless otherwise stated. We will also consider a number of transient CTMCs  $Y$ . Each CTMC  $Y$  has a state space of the form  $B \cup \{a\}$ , where all states in  $B$  are transient and  $a$  is an absorbing state, and has a well-defined initial probability distribution with  $P[Y(0) \in B] = 1$ . We will denote by  $\tau(i, Y)$ ,  $i \in B$  the mean time spent by  $Y$  in  $i$  before absorption ( $\tau(i, Y) = \int_0^\infty P[Y(t) = i]dt$ ). We will also use the notation  $\tau(C, Y) = \sum_{i \in C} \tau(i, Y)$ . It is well-known (see, for instance, [1]) that the mean times to absorption vector  $\boldsymbol{\tau} = (\tau(i, Y))_{i \in B}$  is the solution of the linear system  $\boldsymbol{\tau}^T \mathbf{A}_B = -\mathbf{q}^T$ , where  $\mathbf{A}_B$  is the restriction of the transition rate matrix of  $Y$  to  $B$  and  $\mathbf{q} = (P[Y(0) = i])_{i \in B}$ . The expected number of times that a transition  $(i, j)$  with rate  $\lambda_{ij}$  is followed is  $\mu_{ij} = \tau(i, Y)\lambda_{ij}$ . The result follows easily:  $\mu_{ij} = \int_0^\infty P[Y(t) = i]\lambda_{ij}dt = \lambda_{ij} \int_0^\infty P[Y(t) = i]dt = \lambda_{ij}\tau(i, Y)$ . It can be similarly shown that, given a partition  $B \cup B^c$  of the state space of  $X$  and assuming  $X(0) \in B$ , the probability that  $X$  enters  $B^c$  through a state  $j \in B^c$  is  $\sum_{i \in B} \tau(i, Y_B)\lambda_{ij}$ , where  $Y_B$  is the transient CTMC tracking  $X$  till exit of  $B$  ( $Y_B$  has state space  $B' \cup \{a\}$ , where  $a$  is an absorbing state and  $B'$  is the subset of  $B$  including the states reachable before exiting  $B$  from the states with non-null initial probability, same initial probability distribution in  $B'$  and transition rates among states in  $B'$  as  $X$ , and transition rates  $\lambda'_{i,a} = \lambda_{i,B^c}$ ,  $i \in B'$ , so that  $Y_B$  enters  $a$  whenever  $X$  exits  $B$ ). Note that  $\tau(i, Y_B) > 0$  for  $i \in B'$ .

In this section we derive an upper bound for  $T_{U,s}$ ,  $[T_{U,s}]_{\text{ub}}$ , which can be obtained by solving a ‘‘bounding’’ transient CTMC  $Y$  with failure and repair transitions. The hardcore of this section is Lemma 1, which generalizes the related mean holding time lemma of Muntz et al. [12] by allowing group repair.

Let  $Y_U^m$ ,  $m \in U$  be the transient CTMC with initial state  $m$  tracking  $X$  from  $m$  till exit from  $U$  and let  $T_U^m$  be the mean time to absorption of  $Y_U^m$ . Let  $\alpha_{s,m}$  be the probability that  $X$  with initial

state  $s \in S$  will enter  $U$  through state  $m$ . We have

$$T_{U,s} = \sum_{m \in U} \alpha_{s,m} T_U^m. \quad (5)$$

Invoking Eq. (5), we can easily upper bound  $T_{U,s}$  from upper bounds for  $T_U^m$ ,  $m \in U$ . To obtain these bounds we will invoke the exact aggregation theorem for transient CTMCs and a lemma, which generalizes the mean holding time lemma proved in [12]. Exact aggregation results for irreducible CTMCs are given in [5]. These results extend easily to transient CTMCs. We have:

**Theorem 3** (Exact aggregation for transient CTMCs). *Let  $Y = \{Y(t); t \geq 0\}$  be a transient CTMC with state space  $B \cup \{a\}$ , where all states in  $B$  are transient and  $a$  is an absorbing state, transition rates  $\lambda_{ij}$ ,  $i \in B$ ,  $j \in B \cup \{a\}$ ,  $i \neq j$ , and initial probability distribution  $P[Y(0) = i] = \pi_i$ ,  $i \in B$ ,  $\sum_{i \in B} \pi_i = 1$ . Assume  $\tau(i, Y) > 0$  for all  $i \in B$ . Let  $B_1 \cup B_2 \cup \dots \cup B_n$  be a partition of  $B$ . Then, there exists a transient CTMC  $Y' = \{Y'(t); t \geq 0\}$  (the exact aggregation of  $Y$ ) with state space  $\{b_1, b_2, \dots, b_n\} \cup \{a\}$ , transition rates  $\lambda'_{b_k, b_l} = \sum_{i \in B_k} w_i^k \lambda_{i, b_l}$ ,  $1 \leq k, l \leq n$ ,  $k \neq l$  and  $\lambda'_{b_k, a} = \sum_{i \in B_k} w_i^k \lambda_{i, a}$ ,  $1 \leq k \leq n$ , with  $w_i^k > 0$ ,  $\sum_{i \in B_k} w_i^k = 1$ , and initial probability distribution  $P[Y'(0) = b_k] = \pi'_k = \sum_{i \in B_k} \pi_i$ , such that  $\tau(b_k, Y') = \tau(B_k, Y)$ .*

*Proof.* See Appendix A. □

**Note** The condition  $\tau(i, Y) > 0$ ,  $i \in B$  of Theorem 3 is verified if and only if each state  $i \in B$  is reachable from some state with non-null initial probability.

Consider the exact aggregation,  $Y_U^{m'}$  of  $Y_U^m$ ,  $m \in C_k$ ,  $K < k \leq N'$  under the partition  $\cup_{k=K+1}^{N'} C_k^m$ , where  $C_k^m$  is the subset of  $C_k$  including the states reachable from  $m$  before exiting  $U$  and  $K + 1 \leq N'_m \leq N'$ .  $Y_U^{m'}$  has a transition state diagram like the one given in Fig. 1(a) with  $N'$  substituted by  $N'_m$ . The following lemma shows how the times to absorption vector of  $Y_U^{m'}$  can be upper bounded.

**Lemma 1.** *Assume  $N' \leq N$ . Let  $Y' = \{Y'(t); t \geq 0\}$  be a transient CTMC with the state transition diagram of Fig. 1(a) and initial probability distribution  $P[Y'(0) = c_i] = \pi_i$ ,  $K + 1 \leq i \leq N'$ ,  $\sum_{i=K+1}^{N'} \pi_i = 1$ . Let  $Y = \{Y(t); t \geq 0\}$  be the transient CTMC with the state transition diagram of Fig. 1(b) and initial probability distribution  $P[Y(0) = c_i] = \pi_i$ ,  $K + 1 \leq i \leq N'$ ,  $P[Y(0) = c_i] = 0$ ,  $N' < i \leq N$ . Assume  $f_{i,j} \leq f_{i,j}^+$  and  $\sum_{j=1}^{\beta} g_{i,j} \geq g_i^- > 0$ ,  $K + 1 \leq i \leq N'$ . Then,  $\tau(c_i, Y) \geq \tau(c_i, Y')$ ,  $K + 1 \leq i \leq N'$ .*

*Proof.* For notational conciseness let  $\tau_i = \tau(c_i, Y)$ ,  $\tau'_i = \tau(c_i, Y')$ . We will use as a basic tool the balance equation for a subset of states of a transient CTMC, which establishes that the initial probability of the subset plus the expected number of entries must be equal to the final probability of the subset plus the expected number of exits. The states  $c_i$  of  $Y$  and  $Y'$  are transient and, therefore, have final probabilities equal to 0.

The proof is by induction on  $k$ ,  $k = K + 1, \dots, N'$ . The balance equation applied to  $Y'$  and the subset  $\{c_{K+1}, c_{K+2}, \dots, c_{N'}\}$  gives

$$1 = \sum_{i=K+1}^{N'} \tau'_i \sum_{j=i-K}^{\beta} g_{i,j}, \quad (6)$$

$$\tau'_{K+1} = \frac{1 - \sum_{i=K+2}^{N'} \tau'_i \sum_{j=i-K}^{\beta} g_{i,j}}{\sum_{j=1}^{\beta} g_{K+1,j}}. \quad (7)$$

The balance equation applied to  $Y$  and the subset  $\{c_{K+1}, c_{K+2}, \dots, c_N\}$  gives

$$1 = \tau_{K+1} g_{K+1}^-, \quad (8)$$

$$\tau_{K+1} = \frac{1}{g_{K+1}^-}. \quad (9)$$

Since  $g_{K+1}^- \leq \sum_{j=1}^{\beta} g_{K+1,j}$ , using (9) and (7), we have

$$\tau_{K+1} \geq \frac{1}{\sum_{j=1}^{\beta} g_{K+1,j}} \geq \tau'_{K+1},$$

showing the base case.

For the induction step, consider  $K + 1 < k \leq N'$  and assume  $\tau_i \geq \tau'_i$ ,  $K + 1 \leq i < k$ . The balance equation applied to  $Y'$  and the subset  $\{c_{K+1}, c_{K+2}, \dots, c_{k-1}\}$  gives

$$\sum_{i=K+1}^{k-1} \pi_i + \sum_{i=k}^{N'} \tau'_i \sum_{j=i-k+1}^{i-K-1} g_{i,j} = \sum_{i=K+1}^{k-1} \tau'_i \sum_{j=i-K}^{\beta} g_{i,j} + \sum_{i=K+1}^{k-1} \tau'_i \sum_{j=k-i}^{N'-i} f_{i,j},$$

which, using (6) and  $1 - \sum_{i=K+1}^{k-1} \pi_i = \sum_{i=k}^{N'} \pi_i$  gives

$$\begin{aligned} \tau'_k \sum_{i=1}^{k-K-1} g_{k,i} &= \sum_{i=k}^{N'} \pi_i + \sum_{i=K+1}^{k-1} \tau'_i \sum_{j=k-i}^{N'-i} f_{i,j} - \tau'_k \sum_{j=k-K}^{\beta} g_{k,j} - \sum_{i=k+1}^{N'} \tau'_i \sum_{j=i-K}^{\beta} g_{i,j} - \sum_{i=k+1}^{N'} \tau'_i \sum_{j=i-k+1}^{i-K-1} g_{i,j}, \\ \tau'_k &= \frac{\sum_{i=k}^{N'} \pi_i + \sum_{i=K+1}^{k-1} \tau'_i \sum_{j=k-i}^{N'-i} f_{i,j} - \sum_{i=k+1}^{N'} \tau'_i \sum_{j=i-K}^{\beta} g_{i,j} - \sum_{i=k+1}^{N'} \tau'_i \sum_{j=i-k+1}^{i-K-1} g_{i,j}}{\sum_{i=1}^{\beta} g_{k,i}}. \end{aligned} \quad (10)$$

The balance equation applied to  $Y$  and the subset  $\{c_{K+1}, c_{K+2}, \dots, c_{k-1}\}$  gives

$$\sum_{i=K+1}^{k-1} \pi_i + \tau_k g_k^- = \tau_{K+1} g_{K+1}^- + \sum_{i=K+1}^{k-1} \tau_i \sum_{j=k-i}^{N-i} f_{i,j}^+,$$



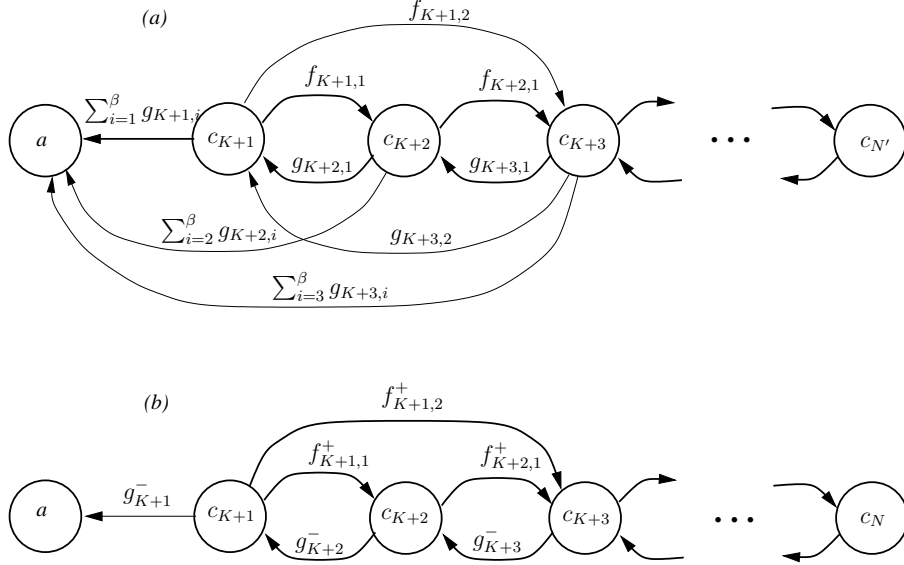


Figure 1: State transition diagrams of CTMCs of Lemma 1.

which, using (8) and  $1 - \sum_{i=K+1}^{k-1} \pi_i = \sum_{i=k}^{N'} \pi_i$  gives

$$\tau_k = \frac{\sum_{i=k}^{N'} \pi_i + \sum_{i=K+1}^{k-1} \tau_i \sum_{j=k-i}^{N-i} f_{i,j}^+}{g_k^-}. \quad (11)$$

Finally, using (11), (10),  $N \geq N'$ ,  $f_{i,j}^+ \geq f_{i,j}$ ,  $\sum_{i=1}^{\beta} g_{k,i} \geq g_k^-$ , and the induction hypothesis

$$\tau_k \geq \frac{\sum_{i=k}^{N'} \pi_i + \sum_{i=K+1}^{k-1} \tau_i \sum_{j=k-i}^{N-i} f_{i,j}^+}{\sum_{i=1}^{\beta} g_{k,i}} \geq \frac{\sum_{i=k}^{N'} \pi_i + \sum_{i=K+1}^{k-1} \tau'_i \sum_{j=k-i}^{N-i} f_{i,j}}{\sum_{i=1}^{\beta} g_{k,i}} \geq \tau'_k. \quad \square$$

Let  $\alpha_s(k) = \sum_{m \in C_k} \alpha_{s,m}$  be the probability that  $X$  with initial state  $s \in S$  will enter  $U$  through subset  $C_k$ . Let  $f_{i,j}^+$  be upper bounds for the transition rates  $f_{i,j}^{m'}$  of  $Y_U^{m'}$  from  $c_i$  to  $c_{i+j}$  associated with the failure of  $j$  components. Let  $g_i^-$  be lower bounds for  $\sum_{j=1}^{\beta} g_{i,j}^{m'}$ , where  $g_{i,j}^{m'}$  is the transition rate of  $Y_U^{m'}$  from  $c_i$  to  $c_{i-j}$  if  $j \leq i - K - 1$ ,  $\sum_{j=i-K}^{\beta} g_{i,j}^{m'}$  is the transition rate of  $Y_U^{m'}$  from  $c_i$  to  $a$ , and  $g_{i,j}^{m'}$  is associated with the repair of  $j$  components. Let  $T(k)$  be the mean time to absorption of the transient CTMC  $Y$  of Fig. 1(b) with initial state  $c_k$ . We have

**Theorem 4.**  $T_{U,s} \leq \sum_{k=K+1}^{N'} \alpha_s(k) T(k) = [T_{U,s}]_{\text{ub}}$ .

*Proof.* Let  $Y^k$  be the transient CTMC with the state transition diagram of Fig. 1(b) and initial state  $c_k$ .  $Y_U^{m'}$ ,  $m' \in C_k$  and  $Y^k$  satisfy the conditions of Lemma 1 and, therefore,  $\tau(c_i, Y_U^{m'}) \leq \tau(c_i, Y^k)$ . By Theorem 3 we have  $T_U^m = \sum_{i=K+1}^{N'_m} \tau(c_i, Y_U^{m'})$ . Then,  $T_U^m \leq \sum_{i=K+1}^{N'_m} \tau(c_i, Y^k) \leq$

$\sum_{i=K+1}^N \tau(c_i, Y^k) = T(k)$ . It follows (5):

$$T_{U,s} = \sum_{m \in U} \alpha_{s,m} T_U^m = \sum_{k=K+1}^{N'} \sum_{m \in C_k} \alpha_{s,m} T_U^m \leq \sum_{k=K+1}^{N'} \sum_{m \in C_k} \alpha_{s,m} T(k) = \sum_{k=K+1}^{N'} \alpha_s(k) T(k). \quad \square$$

Upper bounds  $f_{i,j}^+$  for the transition rates  $f_{i,j}^{m'}$  can be easily derived. Let  $E_j$  be the subset of  $E$  including the failure bags of cardinality  $j$ . It is clear that  $\lambda_{n,C_{i+j}}$ ,  $n \in C_i$  is upper bounded by  $\sum_{e \in E_j} [\lambda(e)]_{\text{ub}}$ . Using Theorem 3:

$$f_{i,j}^{m'} = \sum_{n \in C_i^m} w_n^i \lambda_{n,C_{i+j}},$$

with  $w_n^i > 0$ ,  $\sum_{n \in C_i^m} w_n^i = 1$ . Then, it follows

$$f_{i,j}^{m'} \leq \sum_{n \in C_i^m} w_n^i \sum_{e \in E_j} [\lambda(e)]_{\text{ub}} = \sum_{e \in E_j} [\lambda(e)]_{\text{ub}} = f_{i,j}^+. \quad (12)$$

In [12] the lowest repair rate of the model is used as lower bounds  $g_i^-$ . Unfortunately, a similar approach cannot be taken for the models considered in this paper, since depending on the characteristics of the phase type distributions the lowest rate to the left from the states of a subset  $C_i$  may be 0. A more sophisticated approach is needed. That approach is developed in the next section.

### 3.3 Computation of $g_i^-$

In this section we derive lower bounds  $g_i^- > 0$  using easy to compute characteristics of the phase type repair distributions  $P_i$  of the model. To derive the bounds  $g_i^-$  we need results from [10, 16] which are obtained for irreducible CTMCs. To establish a link with these results we define irreducible CTMCs  $X_U^m$ ,  $m \in U$  as follows:  $X_U^m$  has state space  $U^m \cup \{a\}$ , where  $U^m$  is the subset of  $U$  including the states reachable from  $m$  before exiting  $U$ , transition rates from  $U^m$  to  $U^m \cup \{a\}$  as  $Y_U^m$  and a transition rate 1 from  $a$  to  $m$ . Let  $X_U^{m'}$  be the exact aggregation of  $X_U^m$  under the partition  $(\cup_{k=K+1}^{N'_m} C_k^m) \cup \{a\}$ . Given the connection between Theorem 3 and the exact aggregation theorem for irreducible CTMCs [5] and the relationships between the involved CTMCs, it is easy to prove that the transition rates of  $X_U^{m'}$  from  $\{c_{K+1}, \dots, c_{N'_m}\}$  to  $\{c_{K+1}, \dots, c_{N'_m}, a\}$  are equal to the corresponding transition rates of  $Y_U^{m'}$ . Thus, we will consider the CTMCs  $X_U^{m'}$  instead of  $Y_U^{m'}$ .

Let  $q_{k,i}^{m,L}$ ,  $i \in C_k^m$  be the probability that  $X_U^m$  will jump from  $C_k^m$  to  $\cup_{j=K+1}^{k-1} C_j^m \cup \{a\}$  (i.e. to the left) given entry in  $C_k^m$  through state  $i$  and let  $h_{k,i}^m$ ,  $i \in C_k^m$  be the mean holding time of  $X_U^m$  in  $C_k^m$  given entry in  $C_k^m$  through state  $i$ . Let  $q_k^{m,i'}$  be the probability that  $X_U^{m'}$  will jump from  $c_k$  to  $c_{k-i}$ , let  $q_k^{m,a'}$  be the probability that  $X_U^{m'}$  will jump from  $c_k$  to  $a$ , let  $q_k^{m,L'} = \sum_{i=1}^{k-K-1} q_k^{m,i'} + q_k^{m,a'}$  be the probability that  $X_U^{m'}$  will jump from  $c_k$  to  $\cup_{i=K+1}^{k-1} c_i \cup \{a\}$  (i.e. to the left), and let  $h_k^{m'}$  be the mean holding time of  $X_U^{m'}$  in  $c_k$ . We have

$$h_k^{m'} = \frac{1}{\beta + \sum_{i=1}^{N'_m-k} g_{k,i}^{m'} + \sum_{i=1} f_{k,i}^{m'}}$$

$$q_k^{m,i'} = h_k^{m'} g_{k,i}^{m'}, \quad 1 \leq i \leq k - K - 1,$$

$$q_k^{m,a'} = h_k^{m'} \sum_{i=k-K}^{\beta} g_{k,i}^{m'}.$$

Combining them we obtain

$$\sum_{i=1}^{\beta} g_{k,i}^{m'} = \frac{\sum_{i=1}^{k-K-1} q_k^{m,i'} + q_k^{m,a'}}{h_k^{m'}} = \frac{q_k^{m,L'}}{h_k^{m'}}. \quad (13)$$

Denote by  $\mathbf{A}^m$  the transition rate matrix of  $X_U^m$ , by  $\mathbf{A}_C^m$  the restriction of  $\mathbf{A}^m$  to the subset of states  $C$ , by  $\mathbf{A}_{C,C'}^m$  the subblock of  $\mathbf{A}^m$  including the transition rates from states in  $C$  to states in  $C'$ , and let  $L_k^m = \cup_{i=K+1}^{k-1} C_i^m \cup \{a\}$ . Let  $\mathbf{v}_k^m$  denote the steady-state entry distribution of  $X_U^m$  in  $C_k^m$ , and denote by  $v_{k,i}^m$  the component of  $\mathbf{v}_k^m$  associated with state  $i \in C_k^m$ . Denote by  $\mathbf{1}$  a column vector of all ones of the appropriate dimension and by  $\mathbf{v}^T$  the transpose of vector  $\mathbf{v}$ . From [16, Corollary 4.6] we have:

**Lemma 2.**

$$h_k^{m'} = -\mathbf{v}_k^{mT} \mathbf{A}_{C_k^m}^{m-1} \mathbf{1}.$$

Lemma 2 says that  $h_k^{m'}$  is equal to the mean holding time of  $X_U^m$  in  $C_k^m$  when  $C_k^m$  is entered with probability distribution  $\mathbf{v}_k^m$ . Then, it follows that

$$h_k^{m'} = \sum_{i \in C_k^m} v_{k,i}^m h_{k,i}^m. \quad (14)$$

Also, by analogy with Lemma 2 of [10], we have:

**Lemma 3.**

$$q_k^{m,L'} = -\mathbf{v}_k^{mT} \mathbf{A}_{C_k^m}^{m-1} \mathbf{A}_{C_k, L_k^m}^m \mathbf{1}.$$

Lemma 3 says that  $q_k^{m,L'}$  is equal to the probability that  $X_U^m$  will jump from  $C_k^m$  to  $L_k^m$  when  $C_k^m$  is entered with probability distribution  $\mathbf{v}_k^m$ . Then, it follows that

$$q_k^{m,L'} = \sum_{i \in C_k^m} v_{k,i}^m q_{k,i}^{m,L}. \quad (15)$$

Combining the results obtained so far it can be proved:

**Theorem 5.**

$$\sum_{i=1}^{\beta} g_{k,i}^{m'} \geq \min_{i \in C_k^m} \frac{q_{k,i}^{m,L}}{h_{k,i}^m}.$$

*Proof.* Combining (13)–(15) we have

$$\sum_{i=1}^{\beta} g_{k,i}^{m'} = \frac{\sum_{i \in C_k^m} v_{k,i}^m q_{k,i}^{m,L}}{\sum_{i \in C_k^m} v_{k,i}^m h_{k,i}^m}.$$

Note that  $v_{k,i}^m \geq 0$  and  $\sum_{i \in C_k^m} v_{k,i}^m = 1$  and we are in a position similar to proving  $R \geq \min_{s \in S} \{R_s\}$  of Corollary 1. Then, the result follows.  $\square$

Assume that a lower bound,  $q^-$ , for  $q_{k,i}^{m,L}$ ,  $K+1 \leq k \leq N'_m$ ,  $i \in C_k^m$  and an upper bound,  $h^+$ , for  $h_{k,i}^m$ ,  $K+1 \leq k \leq N'_m$ ,  $i \in C_k^m$  are available. Using Theorem 5 we have

$$\sum_{i=1}^{\beta} g_{k,i}^{m'} \geq \frac{q^-}{h^+} = g_k^-. \quad (16)$$

In the following we show how  $q^-$  and  $h^+$  can be derived. To that end we first introduce some notation. Let a state  $i \in C_k^m$ ,  $K+1 \leq k < N'_m$ . We will denote by  $\lambda_{i,C_{>k}}$  the transition rate from  $i$  to  $\cup_{k'=k+1}^{N'_m} C_{k'}^m$  (note that  $\lambda_{j,C_{>k}}$  is the same for all states  $j$  which are visited in  $C_k^m$  from a given entry state  $i$ , since all these states have the same bag of failed component types and same failed modes of the failed components), by  $A_i$  the number of active repair processes in  $i$ , by  $a_j(i)$ ,  $1 \leq j \leq A_i$  the phase type distribution of the  $j$ th active repair process in state  $i$  ( $1 \leq a_j(i) \leq L$ ), and by  $s_j(i)$  the state of the phase type distribution  $Z_{a_j(i)}$  in state  $i$ . Let  $W_j^s$  be the random variable time to absorption of  $Z_j$  with initial state  $s$ . Let  $\lambda_{\text{ub}} = \sum_{e \in E} [\lambda(e)]_{\text{ub}}$  and let  $\text{EXP}(\lambda)$  denote an exponential random variable with parameter  $\lambda$ . Since  $\lambda_{i,C_{>k}} \leq \lambda_{\text{ub}}$  and the random variables  $W_j^s$ ,  $\text{EXP}(\lambda_{i,C_{>k}})$  and  $\text{EXP}(\lambda_{\text{ub}})$  are independent we have:

$$\begin{aligned} q_{k,i}^{m,L} &= P\left[\min_{1 \leq j \leq A_i} W_{a_j(i)}^{s_j(i)} < \text{EXP}(\lambda_{i,C_{>k}})\right] \geq P\left[\min_{1 \leq j \leq A_i} W_{a_j(i)}^{s_j(i)} < \text{EXP}(\lambda_{\text{ub}})\right] \\ &\geq \min_{1 \leq j \leq A_i} P\left[W_{a_j(i)}^{s_j(i)} < \text{EXP}(\lambda_{\text{ub}})\right] \geq \min_{1 \leq j \leq L} \min_{s \in L_j} P[W_j^s < \text{EXP}(\lambda_{\text{ub}})]. \end{aligned}$$

Also

$$\begin{aligned} h_{k,i}^m &= E\left[\min\left\{\text{EXP}(\lambda_{i,C_{>k}}), \min_{1 \leq j \leq A_i} W_{a_j(i)}^{s_j(i)}\right\}\right] < E\left[\min_{1 \leq j \leq A_i} W_{a_j(i)}^{s_j(i)}\right] \\ &\leq \max_{1 \leq j \leq A_i} E\left[W_{a_j(i)}^{s_j(i)}\right] \leq \max_{1 \leq j \leq L} \max_{s \in L_j} E[W_j^s]. \end{aligned}$$

Let us denote  $P[W_j^s < \text{EXP}(\lambda_{\text{ub}})]$  by  $Q_j^s$  and  $E[W_j^s]$  by  $H_j^s$ . We use

$$q^- = \min_{1 \leq j \leq L} \min_{s \in L_j} Q_j^s, \quad (17)$$

$$h^+ = \max_{1 \leq j \leq L} \max_{s \in L_j} H_j^s. \quad (18)$$

Let  $\mathbf{B}_j$  be the transition rate matrix of  $Z_j$  restricted to the transient states  $L_j$  and let  $\mathbf{b}_j$  be the vector whose entries are the transition rates of  $Z_j$  from  $L_j$  to the absorbing state  $a$ . Let  $\mathbf{Q}_j$  and  $\mathbf{H}_j$  be the vectors with entries  $Q_j^s$  and  $H_j^s$ , respectively,  $s \in L_j$ .  $\mathbf{Q}_j$  and  $\mathbf{H}_j$  can be obtained as:

$$\mathbf{Q}_j = -(\mathbf{B}_j - \lambda_{\text{ub}} \mathbf{I})^{-1} \mathbf{b}_j, \quad (19)$$

$$\mathbf{H}_j = -\mathbf{B}_j^{-1}\mathbf{1}. \quad (20)$$

Eq. (20) is trivial since the component in row  $s$  and column  $i$  of  $-\mathbf{B}_j^{-1}$  is the mean time to absorption spent by  $Z_j$  in state  $i$  given that the initial state is  $s$ . Eq. (19) follows considering transient CTMCs  $Z'_j$  with state space  $L_j \cup \{a, b\}$ , where  $a$  and  $b$  are absorbing states, same transition rates from  $L_j$  to  $a$  as  $Z_j$  and transition rates  $\lambda_{s,b} = \lambda_{\text{ub}}$ ,  $s \in L_j$ . The transition rate matrix of  $Z'_j$  restricted to  $L_j$  is  $\mathbf{B}_j - \lambda_{\text{ub}}\mathbf{I}$  and  $Q_j^s$  is the probability of being absorbed in state  $a$  given that the initial state is  $s$ . These comments justify the equation.

### 3.4 Computation of $T(k)$

Let  $M$  be the set of indices  $k$  associated to the subsets  $C_k$ ,  $K < k \leq N$  with  $\lambda_{i,C_k} \neq 0$  for some  $i \in G$ . Remember that  $T(k)$  is the mean time to absorption of the transient CTMC  $Y$  of Fig. 1(b) with initial state  $c_k$ . In order to obtain the bounds  $[T_{U,s}]_{\text{ub}}$  given by Theorem 4 we have to compute  $T(k)$ ,  $k \in M$ . A direct computation of each  $T(k)$  solving  $Y$  with initial state  $c_k$  would require the solution of  $|M|$  linear systems. In this section, we derive a more efficient procedure, specially for large  $|M|$ . The procedure is based on the following equations, where  $\phi_k = g_k^- + \sum_i f_{k,i}^+$  denotes the output rate of  $Y$  from  $c_k$  (see Fig. 1(b)):

$$T(k) = \frac{1}{\phi_k} + \frac{g_k^-}{\phi_k} T(k-1) + \sum_i \frac{f_{k,i}^+}{\phi_k} T(k+i), \quad K+2 \leq k < N, \quad (21)$$

$$T(N) = \frac{1}{g_N^-} + T(N-1). \quad (22)$$

These equations are obtained as follows. First, consider (21).  $T(k)$ , mean time to absorption of  $Y$  with initial state  $c_k$ , is equal to the mean sojourn time in  $c_k$ ,  $1/\phi_k$ , plus the mean time to absorption from the next visited state, which is  $c_{k-1}$  with probability  $g_k^-/\phi_k$  and  $c_{k+i}$  with probability  $f_{k,i}^+/\phi_k$ . Eq. (22) is obtained similarly; in this case,  $\phi_N = g_N^-$  and state  $c_{N-1}$  is the next visited state with probability 1. Eqs. (21) and (22) can be solved recursively in terms of  $T(N)$ , yielding

$$T(N-1) = T(N) - \frac{1}{g_N^-}, \quad (23)$$

$$T(k) = \frac{1}{g_{k+1}^-} \left[ \phi_{k+1} T(k+1) - 1 - \sum_i f_{k+1,i}^+ T(k+1+i) \right], \quad k = N-2, \dots, K+1. \quad (24)$$

It remains to discuss the computation of  $T(N)$ . Let  $\tau_i^N$  denote the mean time to absorption in state  $c_i$  of  $Y$  with initial state  $c_N$ . Then

$$T(N) = \sum_{i=K+1}^N \tau_i^N. \quad (25)$$

The row vector  $\boldsymbol{\tau}^{NT} = (\tau_{K+1}^N \dots \tau_N^N)$  is the solution of the linear system

$$\boldsymbol{\tau}^{NT} \mathbf{A} = -(0 \dots 01), \quad (26)$$

where  $\mathbf{A}$  is the restriction of the transition rate matrix of  $Y$  to the transient states. A direct solution of (26) is possible exploiting the upper Hessenberg structure of  $\mathbf{A}$  and the fact that all components except the last one of the right-hand vector of (26) are null. Defining  $\nu_i = \tau_i^N / \tau_{K+1}^N$  ( $\nu_{K+1} = 1$ ), all the equations except the last one give a triangular linear system on  $\nu_i$ ,  $K+2 \leq i \leq N$  which can be solved easily. Substituting then  $\tau_i^N$  by  $\nu_i \tau_{K+1}^N$ ,  $K+2 \leq i \leq N$ , in the last equation of (26) and using the solution for  $\nu_i$ ,  $K+2 \leq i \leq N$  found in the previous step gives an equation on  $\tau_{K+1}^N$ . Solving that equation and using  $\tau_i^N = \nu_i \tau_{K+1}^N$ ,  $K+2 \leq i \leq N$  we obtain  $\tau_i^N$ ,  $K+2 \leq i \leq N$ . The solution procedure can be described as follows:

$$\begin{aligned} \nu_{K+1} &= 1, \\ \nu_i &= \frac{1}{g_i} \left[ \phi_{i-1} \nu_{i-1} - \sum_{j=K+1}^{i-2} f_{j,i-j-1}^+ \nu_j \right], \quad i = K+2, \dots, N, \end{aligned} \quad (27)$$

$$\begin{aligned} \tau_{K+1}^N &= \frac{1}{\phi_N \nu_N - \sum_{i=K+1}^{N-1} f_{i,N-i}^+ \nu_i}, \\ \tau_i^N &= \nu_i \tau_{K+1}^N, \quad i = K+2, \dots, N. \end{aligned} \quad (28)$$

### 3.5 Computation of the bounds

$T_{G,s}$ ,  $C_{G,s}$  and  $\alpha_s(k)$ ,  $k \in M$  could be computed from the mean times to absorption vector of the transient CTMCs  $Y_G^s$  tracking  $X$  from state  $s$  till exit from  $G$ . This however would involve the solution of  $|S|$  linear systems of size  $|G|$ , which is very expensive. In this section we develop a computational procedure which obtains the bounds  $[R]_{\text{lb}}$  and  $[R]_{\text{ub}}$  solving only four linear systems.

Let:

$$\begin{aligned} T'_s &= T_{G,s} + [T_{U,s}]_{\text{ub}}, \\ C'_s &= C_{G,s} + [r]_{\text{lb}} [T_{U,s}]_{\text{ub}}, \\ C''_s &= C_{G,s} + [r]_{\text{ub}} [T_{U,s}]_{\text{ub}}. \end{aligned}$$

Using (4), the bounds (2), (3) for  $R$  can be expressed in terms of  $T'_s$ ,  $C'_s$  and  $C''_s$ ,  $s \in S$  as

$$[R]_{\text{lb}} = \min_{s \in S} \left\{ \frac{C'_s}{T'_s} \right\}, \quad (29)$$

$$[R]_{\text{ub}} = \max_{s \in S} \left\{ \frac{C''_s}{T'_s} \right\}. \quad (30)$$

The key of the new computational procedure is the derivation of forward equations for  $T'_i$ ,  $C'_i$  and  $C''_i$ ,  $i \in G$ . To that end we first write these variables in terms of  $\alpha_i(k)$  and  $T(k)$  using Theorem 4:

$$\begin{aligned} T'_i &= T_{G,i} + \sum_{k \in M} \alpha_i(k) T(k), \quad i \in G, \\ C'_i &= C_{G,i} + [r]_{\text{lb}} \sum_{k \in M} \alpha_i(k) T(k), \quad i \in G, \end{aligned}$$

$$C_i'' = C_{G,i} + [r]_{\text{ub}} \sum_{k \in M} \alpha_i(k) T(k), \quad \text{hspace} * 1 \text{emi} \in G.$$

Each of these variables can be expressed as the sum of a contribution associated with the visit to state  $i$  plus the corresponding variable for the next visited state in  $G$ . This gives the forward equations (note that  $1/\lambda_i$  is the mean time in state  $i$ ,  $\lambda_{i,C_k}/\lambda_i$  is the jump probability from state  $i$  to subset  $C_k$ , and  $\lambda_{ij}/\lambda_i$  is the jump probability from state  $i$  to state  $j$ ):

$$T_i' = \frac{1}{\lambda_i} + \sum_{k \in M} \frac{\lambda_{i,C_k}}{\lambda_i} T(k) + \sum_{\substack{j \in G \\ j \neq i}} \frac{\lambda_{ij}}{\lambda_i} T_j', \quad i \in G, \quad (31)$$

$$C_i' = \frac{r_i}{\lambda_i} + \sum_{k \in M} \frac{\lambda_{i,C_k}}{\lambda_i} [r]_{\text{lb}} T(k) + \sum_{\substack{j \in G \\ j \neq i}} \frac{\lambda_{ij}}{\lambda_i} C_j', \quad i \in G, \quad (32)$$

$$C_i'' = \frac{r_i}{\lambda_i} + \sum_{k \in M} \frac{\lambda_{i,C_k}}{\lambda_i} [r]_{\text{ub}} T(k) + \sum_{\substack{j \in G \\ j \neq i}} \frac{\lambda_{ij}}{\lambda_i} C_j'', \quad i \in G. \quad (33)$$

Let  $\rho_{ij} = \lambda_{ij}/\lambda_i$ . The sets of equations (31)–(33) can be formulated as linear systems introducing the matrix  $\mathbf{B} = \mathbf{I} - (\rho_{ij})_{i,j \in G, i \neq j}$  and the vectors  $\mathbf{T}' = (T_i')_{i \in G}$ ,  $\mathbf{C}' = (C_i')_{i \in G}$ ,  $\mathbf{C}'' = (C_i'')_{i \in G}$ ,  $\boldsymbol{\mu}' = ((1/\lambda_i) + \sum_{k \in M} (\lambda_{i,C_k}/\lambda_i) T(k))_{i \in G}$ ,  $\mathbf{c}' = ((r_i/\lambda_i) + \sum_{k \in M} (\lambda_{i,C_k}/\lambda_i) [r]_{\text{lb}} T(k))_{i \in G}$ , and  $\mathbf{c}'' = ((r_i/\lambda_i) + \sum_{k \in M} (\lambda_{i,C_k}/\lambda_i) [r]_{\text{ub}} T(k))_{i \in G}$ :

$$\mathbf{B}\mathbf{T}' = \boldsymbol{\mu}', \quad (34)$$

$$\mathbf{B}\mathbf{C}' = \mathbf{c}', \quad (35)$$

$$\mathbf{B}\mathbf{C}'' = \mathbf{c}''. \quad (36)$$

Matrix  $\mathbf{B}$  can be large and iterative methods should be used to solve the linear systems (34)–(36). From the properties of  $\mathbf{B}$  it is easy to prove [19] that Gauss-Seidel will converge. We found though that the convergence under Gauss-Seidel was typically extremely slow. However, a decomposition technique can be used to speed up the convergence. The price is to solve one more linear system, but we have found that then Gauss-Seidel converges very fast. See [7] for an analysis of the convergence properties of the linear systems obtained with the decomposition technique.

To describe the decomposition technique let us consider the generic problem of computing for  $i \in G$  the expected accumulated reward up to absorption  $V_i$  of the transient CTMC  $Y_G^i$  with initial state  $i$  tracking  $X$  till exit from  $G$  for the generic reward rate structure  $v_j$ ,  $j \in G$ . Note that  $T_i'$ ,  $C_i'$  and  $C_i''$  can be formulated as  $V_i$  with  $v_j$  equal to, respectively,  $1 + \sum_{k \in M} \lambda_{j,C_k} T(k)$ ,  $r_j + \sum_{k \in M} \lambda_{j,C_k} [r]_{\text{lb}} T(k)$  and  $r_j + \sum_{k \in M} \lambda_{j,C_k} [r]_{\text{ub}} T(k)$ . Let the vectors  $\mathbf{V} = (V_i)_{i \in G}$  and  $\mathbf{b} = (v_i/\lambda_i)_{i \in G}$ . Then,  $\mathbf{V}$  is the solution of the linear system

$$\mathbf{B}\mathbf{V} = \mathbf{b}.$$

Without loss of generality let us assume that the state  $o$  in which all components are up has index 1. Let  $\tilde{V}_i$  denote the expected accumulated reward to absorption or hit of state 1. Let  $\gamma_i$  denote

the probability that  $Y_G^i$  will exit  $G$  without hitting state 1. Decomposing  $V_i$  in its two contributions delimited by the time at which  $Y_G^i$  gets absorbed or hits state 1, we obtain

$$V_i = \tilde{V}_i + (1 - \gamma_i)V_1, \quad i \in G, \quad (37)$$

The set of equations (37) can be solved in  $V_i, i \in G$ , yielding:

$$V_i = \tilde{V}_i + \frac{1 - \gamma_i}{\gamma_1} \tilde{V}_1, \quad i \in G. \quad (38)$$

Note that  $\tilde{V}_i$  is the expected accumulated reward to absorption of the transient CTMC  $\tilde{Y}_G^i$  obtained from  $Y_G^i$  by directing to the absorbing state the entries in state 1. Then,  $\tilde{V}_i, i \in G$  can be computed as  $V_i, i \in G$ , using the matrix  $\tilde{\mathbf{B}}$ :

$$\tilde{\mathbf{B}} = \begin{pmatrix} 1 & -\rho_{12} & \cdots & -\rho_{1,|G|} \\ 0 & 1 & \cdots & -\rho_{2,|G|} \\ & & \ddots & \\ 0 & -\rho_{|G|,2} & \cdots & 1 \end{pmatrix}$$

instead of  $\mathbf{B}$ . Let the vectors  $\tilde{\mathbf{T}}' = (\tilde{T}'_i)_{i \in G}$ ,  $\tilde{\mathbf{C}}' = (\tilde{C}'_i)_{i \in G}$ ,  $\tilde{\mathbf{C}}'' = (\tilde{C}''_i)_{i \in G}$ . Applying the previous result we have that these vectors are the solutions of the linear systems

$$\tilde{\mathbf{B}}\tilde{\mathbf{T}}' = \boldsymbol{\mu}', \quad (39)$$

$$\tilde{\mathbf{B}}\tilde{\mathbf{C}}' = \mathbf{c}', \quad (40)$$

$$\tilde{\mathbf{B}}\tilde{\mathbf{C}}'' = \mathbf{c}''. \quad (41)$$

The probabilities  $\gamma_i$  can be formulated as the expected accumulated reward up to absorption of  $\tilde{Y}_G^i$  with reward rate  $\lambda_{i,U}$ . Then, letting the vectors  $\boldsymbol{\gamma} = (\gamma_i)_{i \in G}$  and  $\boldsymbol{\omega} = (\lambda_{i,U}/\lambda_i)_{i \in G}$ ,  $\boldsymbol{\gamma}$  is the solution of the linear system

$$\tilde{\mathbf{B}}\boldsymbol{\gamma} = \boldsymbol{\omega}. \quad (42)$$

Finally, using (38)  $T'_s, C'_s$  and  $C''_s, s \in S$  can be computed from  $\tilde{T}'_s, \tilde{C}'_s$  and  $\tilde{C}''_s, s \in S$  using

$$T'_s = \tilde{T}'_s + \frac{1 - \gamma_s}{\gamma_1} \tilde{T}'_1, \quad (43)$$

$$C'_s = \tilde{C}'_s + \frac{1 - \gamma_s}{\gamma_1} \tilde{C}'_1, \quad (44)$$

$$C''_s = \tilde{C}''_s + \frac{1 - \gamma_s}{\gamma_1} \tilde{C}''_1. \quad (45)$$

The complete algorithm to compute the bounds can be summarized as follows:



## Algorithm

1. Compute  $f_{i,j}^+$  using (12) and  $g_i^-$  solving (19) and (20), and using Eqs. (17), (18) and (16).
2. Compute  $T(N)$  using (27), (28) and (25).
3. Compute  $T(k)$ ,  $K + 1 \leq k < N$  using (23), (24).
4. Solve the linear systems (39)–(42).
5. Compute  $T'_s$ ,  $C'_s$  and  $C''_s$ ,  $s \in S$  using (43)–(45).
6. Compute  $[R]_{\text{lb}}$ ,  $[R]_{\text{ub}}$  using (29), (30).

## 4 Numerical results

In this section we illustrate the bounding algorithm and compare it with that proposed in [12] using two examples. The first example is an availability model without group repair and with exponentially distributed repair times. The model falls within the scope of application of the algorithm proposed in [12]. We use this first example to compare our algorithm with the algorithm proposed in [12]. For our algorithm we take as  $g_i^-$  a lower bound for the repair rate from any state with failed components, as done in the algorithm described in [12]. The example is a variant of the large example used in [12]. Fig. 2 shows the block diagram of the system. The system is operational (up) if at least one processor PA or PB is unfailed, at least one controller of each set (C1, C2) is unfailed, and at least three disks of each disk cluster (set of disks D1, D2, D3, D4, D5, and D6) are unfailed. Only one processor of each set is active. Non-active processors do not fail. A fault in the active processor PA is propagated to the active processor PB with probability 0.1. Active processors and controllers C1 fail with rate  $1/2000 \text{ h}^{-1}$ . Controllers C2 fail with rate  $1/4000 \text{ h}^{-1}$ . Disks fail with a different rate for each cluster. The disk failure rates are  $1/6000 \text{ h}^{-1}$  for disks D1,  $1/8000 \text{ h}^{-1}$  for disks D2,  $1/10\,000 \text{ h}^{-1}$  for disks D3,  $1/12\,000 \text{ h}^{-1}$  for disks D4,  $1/14\,000 \text{ h}^{-1}$  for disks D5 and  $1/16\,000 \text{ h}^{-1}$  for disks D6. Components can fail in two modes with equal probabilities. There is only one repairman which selects the component to be repaired at random from the set of failed components. The repair rate of a component depends on the failed mode of the component and on the operational/down state of the system. When the system is operational repair rates are  $0.1 \text{ h}^{-1}$  in one failed mode and  $0.05 \text{ h}^{-1}$  in the other failed mode. The repair rates are 10 times larger when the system is down. The difference in repair rates between the operational and down states of the system can be due to more careful repair procedures in the operational state to avoid system crashes as a consequence of erroneous maintenance operations. The system has a moderate complexity (36 components of 10 different types) but a very large state space: of the order of  $10^{10}$  states. The size of the state space precludes an exact numerical solution of the Markov model. Thus, the example illustrates the type of models for which bounding algorithms are an attractive approach.

Table 1 gives the failure bags of the example and for each failure bag  $e$  the corresponding upper bound for its rate  $[\lambda(e)]_{\text{ub}}$ . We use the notation  $c[n]$  to indicate  $n$  instances of component type  $c$ . The upper bounds  $f_{i,j}^+$  are  $f_{i,1}^+ = 4.93571 \times 10^{-3} \text{ h}^{-1}$  and  $f_{i,2}^+ = 5 \times 10^{-5} \text{ h}^{-1}$ . For  $g_i^-$  we take

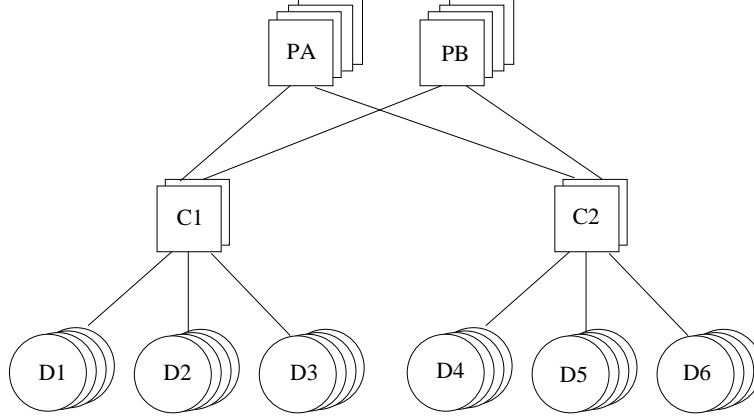


Figure 2: Block diagram of the first example.

Table 1: Failure bags  $e$  and  $[\lambda(e)]_{\text{ub}}$  in  $\text{h}^{-1}$  for the first example.

$e$	$[\lambda(e)]_{\text{ub}}$
PA[1]	$5 \times 10^{-4}$
PA[1] PB[1]	$5 \times 10^{-5}$
PB[1]	$5 \times 10^{-4}$
C1[1]	$10^{-3}$
C2[1]	$5 \times 10^{-4}$
D1[1]	$6.66667 \times 10^{-4}$
D2[1]	$5 \times 10^{-4}$
D3[1]	$4 \times 10^{-4}$
D4[1]	$3.33333 \times 10^{-4}$
D5[1]	$2.85714 \times 10^{-4}$
D6[1]	$2.5 \times 10^{-4}$

$0.05 \text{ h}^{-1}$ . The measure of interest is the steady-state unavailability, which can be formulated as  $R$  with  $r_i = 1$  for down states and  $r_i = 0$  for up states. We then have  $[r]_{\text{lb}} = 0$ ,  $[r]_{\text{ub}} = 1$ .

Table 2 gives the results obtained for the first example with the bounding algorithm described in [12] and the algorithm proposed in this paper. We give bounds, CPU times (measured in a 167 MHz UltraSPARC 1 workstation) and total number of Gauss-Seidel iterations (asking a relative tolerance in the solution of  $10^{-8}$ ) for the algorithm described in [12] for several pairs  $K, F$  and the CPU time and total number of Gauss-Seidel iterations consumed by our algorithm for several values of  $K$ . We also show the number of generated states ( $|G|$ ) and the number of “return” states ( $|S|$ ). Our algorithm obtains the same bounds as the algorithm proposed in [12] for  $F = K$ . The CPU times of the algorithm of [12] for  $F = K$  are large because of the high number of linear systems ( $|S| = |C_K|$ ) which have to be solved. The CPU times of the algorithm can be made reasonable selecting a small value of  $F$  (for instance,  $F = 1$ ) but then the bounds become looser and, as the example shows, they can be significantly looser than the bounds obtained with our algorithm. The

Table 2: Steady-state unavailability bounds, CPU times in seconds and number of Gauss-Seidel iterations ( $n_I$ ) for the algorithm described in [12] and the proposed algorithm for the first example.

$K$	$F$	$ G $	$ S $			CPU time ( $n_I$ )	
				lower bound	upper bound	[12]	proposed
2	0	265	1	$2.9965 \times 10^{-5}$	$1.7579 \times 10^{-3}$	0.165 (79)	
	1		55	$2.9972 \times 10^{-5}$	$1.1930 \times 10^{-3}$	0.595 (496)	
	2		209	$2.9972 \times 10^{-5}$	$6.7213 \times 10^{-4}$	4.73 (5001)	0.118 (41)
3	0	1796	1	$3.5522 \times 10^{-5}$	$1.9900 \times 10^{-4}$	1.28 (77)	
	1		55	$3.5524 \times 10^{-5}$	$1.5832 \times 10^{-4}$	4.43 (439)	
	2		209	$3.5526 \times 10^{-5}$	$1.2123 \times 10^{-4}$	40.2 (4485)	
	3		1531	$3.5526 \times 10^{-5}$	$8.4473 \times 10^{-5}$	296 (35276)	0.920 (54)
4	0	10496	1	$3.6233 \times 10^{-5}$	$4.9960 \times 10^{-5}$	10.3 (76)	
	1		55	$3.6233 \times 10^{-5}$	$4.7222 \times 10^{-5}$	35.1 (379)	
	2		209	$3.6233 \times 10^{-5}$	$4.4729 \times 10^{-5}$	308 (3876)	
	3		1531	$3.6233 \times 10^{-5}$	$4.2237 \times 10^{-5}$	2478 (31042)	
	4		8700	$3.6233 \times 10^{-5}$	$3.9768 \times 10^{-5}$	> 10000	7.26 (62)
5	0	51391	1	$3.6306 \times 10^{-5}$	$3.7359 \times 10^{-5}$	57.8 (75)	
	1		55	$3.6306 \times 10^{-5}$	$3.7184 \times 10^{-5}$	167 (305)	
	2		209	$3.6306 \times 10^{-5}$	$3.7024 \times 10^{-5}$	1492 (3148)	
	5		40895	$3.6306 \times 10^{-5}$	$3.6542 \times 10^{-5}$	> 10000	44.7 (67)

decomposition technique used by our algorithm is extremely efficient and makes the total number of Gauss-Seidel iterations required to solve the four linear systems even smaller than the number of iterations required to solve the single linear system of the algorithm proposed in [12] for  $F = 0$ . The CPU times of our algorithm are accordingly smaller than the CPU times of the algorithm described in [12] for  $F = 0$ . In summary, our algorithm compares favorably with the algorithm proposed in [12], when the latter is applicable.

The second example illustrates the broader applicability of our bounding algorithm. It has both group repair and phase type repair time distributions. The block diagram of the example is given in Fig. 3. The system is made up of two processing subsystems, each including one processor P and two memories M, two sets of controllers C1 and C2, each with two controllers, and four sets of disks D1, D2, D3 and D4, each with three disks. The system is up if at least one processor and one memory connected to it are operational, one controller of each set is operational, and two disks of each set are operational. Processors fail with rate  $10^{-5} \text{ h}^{-1}$ ; a processor failure is soft with probability 0.8 and hard with probability 0.2. In addition, either being soft or hard, a processor failure contaminates (fails) the operational memories to which it is connected with probability 0.05. Memories fail with rate  $5 \times 10^{-5} \text{ h}^{-1}$ , controllers fail with rate  $2 \times 10^{-5} \text{ h}^{-1}$ . Disks D1 fail with rate  $10^{-6} \text{ h}^{-1}$ , disks D2 fail with rate  $1.5 \times 10^{-6} \text{ h}^{-1}$ , disks D3 fail with rate  $2 \times 10^{-6} \text{ h}^{-1}$ , and disks

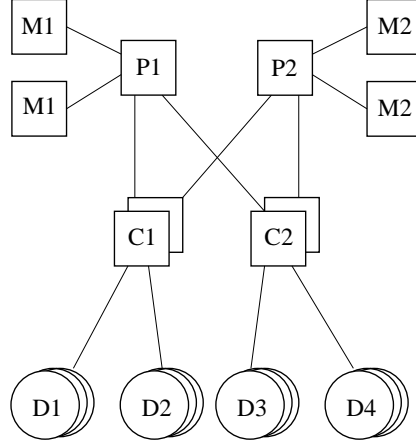


Figure 3: Block diagram of the second example.

D4 fail with rate  $3 \times 10^{-6} \text{ h}^{-1}$ . There are two repairmen. One performs restarts of processors in soft failure and the other performs all the other maintenance actions with first priority given to disks, next to controllers, next to processors, and last to memories. Failed memories of the same processing subsystem are repaired simultaneously (in a single repair action); thus the model has group repair. Components with the same repair priority are chosen at random. The policy is preemptive-resume. Fig. 4 gives the phase type distributions for all repair actions, with the initial probabilities shown inside the circles denoting the states of  $Z_i$  and all transition rates in  $\text{h}^{-1}$ . The state of the system can be described by giving the number of components of each type which are operational and for each component type, failed mode pair the number of failed components of the type in the failed mode in each state of the phase type distribution associated to the component type, failed mode pair. The complete model has about  $4.9 \times 10^9$  states, clearly outside of current computing capabilities.

The second example has 10 component types and  $N = 22$  components. Table 3 gives the failure bags of the model and for each failure bag  $e$  the upper bound  $[\lambda(e)]_{\text{ub}}$  for its rate. The upper bounds  $f_{i,j}^+$  are  $f_{i,1}^+ = 5.225 \times 10^{-4} \text{ h}^{-1}$ ,  $f_{i,2}^+ = 10^{-6} \text{ h}^{-1}$  and  $f_{i,3}^+ = 10^{-6} \text{ h}^{-1}$ . The upper bound  $\lambda_{\text{ub}}$  is  $\lambda_{\text{ub}} = 5.245 \times 10^{-4} \text{ h}^{-1}$ . We also have  $h^+ = 5 \text{ h}$ ,  $q^- = 0.997384$  and  $g_i^- = 0.199477 \text{ h}^{-1}$ .

Table 4 gives the bounds for the steady-state unavailability obtained for  $K = 2, 3, 4$  and 5. We also give the number of generated states ( $|G|$ ). By profiling the code we have found out that about 50% of the CPU time is devoted to the generation of the models, while the solution of the four linear systems accounts for the remaining 50%. The CPU time for  $K = 5$  was about 4 minutes in a 167 MHz UltraSPARC 1 workstation.

The 4-Erlang phase type distribution used for processor restarts can be imagined as an approximation to a deterministic restart time of value  $1 \text{ h}^{-1}$ . The goodness of the approximation improves with the number of exponential stages  $k$ . We explored that issue and obtained results with increasing  $k$  for  $K = 5$ . Table 5 gives the results. We can note that the steady-state unavailability is quite insensitive to the shape of the restart time distribution and a small value of  $k$  is enough to obtain an accurate approximation.

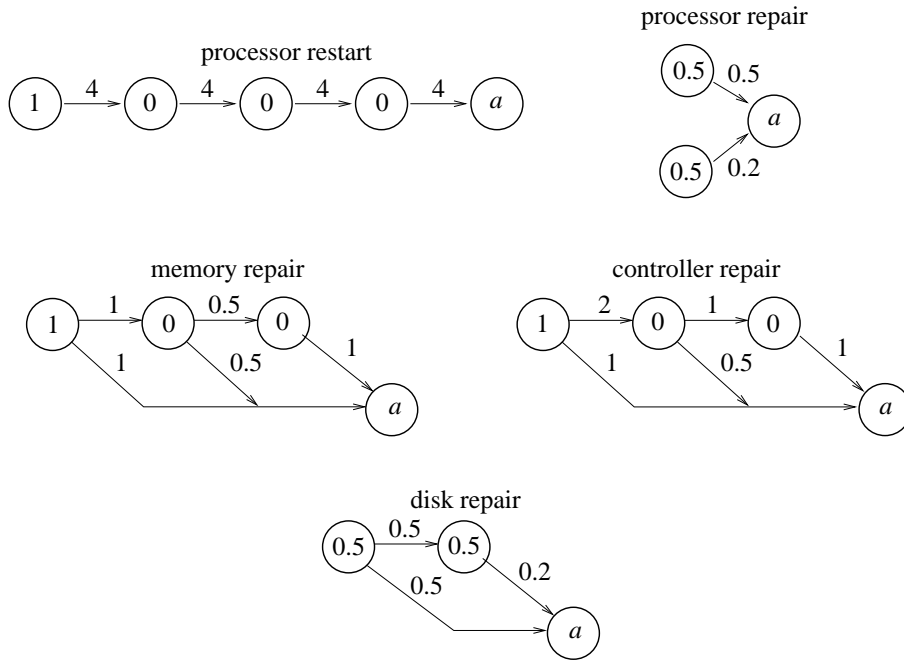


Figure 4: Phase type repair distributions for the repair actions of the second example.

Table 3: Failure bags  $e$  and  $[\lambda(e)]_{\text{ub}}$  in  $\text{h}^{-1}$  for the second example.

$e$	$[\lambda(e)]_{\text{ub}}$
P1[1]	$10^{-5}$
P1[1] M1[1]	$5 \times 10^{-7}$
P1[1] M1[2]	$5 \times 10^{-7}$
M1[1]	$10^{-4}$
P2[1]	$10^{-5}$
P2[1] M2[1]	$5 \times 10^{-7}$
P2[1] M1[2]	$5 \times 10^{-7}$
M2[1]	$10^{-4}$
C1[1]	$4 \times 10^{-5}$
C2[1]	$4 \times 10^{-5}$
D1[1]	$3 \times 10^{-6}$
D2[1]	$4.5 \times 10^{-6}$
D3[1]	$6 \times 10^{-6}$
D4[1]	$9 \times 10^{-6}$

Table 4: Results for the second example and increasing  $K$ .

$K$	$ G $	lower bound	upper bound
2	513	$4.44759 \times 10^{-9}$	$1.00362 \times 10^{-5}$
3	5079	$4.47533 \times 10^{-9}$	$1.33280 \times 10^{-8}$
4	36385	$4.53283 \times 10^{-9}$	$4.56180 \times 10^{-9}$
5	202333	$4.53383 \times 10^{-9}$	$4.54889 \times 10^{-9}$

Table 5: Results for the second example with  $K = 5$  and increasing number of stages of the  $k$ -Erlang distribution of processor restart time.

$k$	$ G $	lower bound	upper bound
1	105658	$4.53468 \times 10^{-9}$	$4.54888 \times 10^{-9}$
2	134637	$4.53413 \times 10^{-9}$	$4.54886 \times 10^{-9}$
3	166862	$4.53393 \times 10^{-9}$	$4.54888 \times 10^{-9}$
4	202333	$4.53383 \times 10^{-9}$	$4.54889 \times 10^{-9}$

## 5 Conclusions

An algorithm to bound the steady-state availability applicable to models with group repair and phase type repair distributions has been developed. Previous bounding algorithms assumed that repair actions involved a single component and assumed exponential repair distributions. In addition, previous bounding algorithms either had to solve many linear systems to obtain the tightest possible bounds or introduced looseness if state cloning techniques were used to reduce the number of linear systems to be solved, whereas our algorithm does not clone states and requires the solution of only four linear systems of the size of the generated state space, being the time devoted to the solution of these linear systems comparable with the time to generate the model. Our algorithm per se is not confined to compute bounds for the steady-state availability: it can be used to compute bounds for the steady-state reward rate of models exhibiting similar structures.

## Acknowledgements

This work was supported by the ‘‘Comisin Interministerial de Ciencia y Tecnologia’’ (CICYT) under the research grant TIC95-0707-C02-02.

## Appendix A. Proof of Theorem 3

Without loss of generality, assume that the transient states of  $Y$  are sorted following the subset ordering  $B_1, B_2, \dots, B_n$ . For notational conciseness let  $\tau_i = \tau(i, Y)$  and  $\tau'_k = \tau(B_k, Y)$ . Note that  $\tau'_k > 0$ . Let the vectors  $\boldsymbol{\tau} = (\tau_i)_{i \in B}$ ,  $\boldsymbol{\pi} = (\pi_i)_{i \in B}$  and let  $\mathbf{A}$  be the transition rate matrix of  $Y$  restricted to  $B$ .  $\boldsymbol{\tau}$  satisfies the linear system

$$\boldsymbol{\tau}^T \mathbf{A} = -\boldsymbol{\pi}^T. \quad (46)$$

Let  $w_i^k = \tau_i / \tau'_k$ ,  $i \in B_k$ ,  $1 \leq k \leq n$ . Note that  $w_i^k > 0$  and  $\sum_{i \in B_k} w_i^k = 1$ . Defining the column vectors  $\mathbf{w}(k) = (w_i^k)_{i \in B_k}$ ,  $\boldsymbol{\pi}(k) = (\pi_i)_{i \in B_k}$ , we can rewrite (46) as

$$\left( \tau'_1 \mathbf{w}(1)^T \cdots \tau'_n \mathbf{w}(n)^T \right) \begin{pmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ & \ddots & \\ \mathbf{A}_{n1} & \cdots & \mathbf{A}_{nn} \end{pmatrix} = -\left( \boldsymbol{\pi}(1)^T \cdots \boldsymbol{\pi}(n)^T \right),$$

where  $\mathbf{A}_{kl}$  are the blocks of  $\mathbf{A}$  induced by the partition of  $B$ . This block decomposition gives the set of equations

$$\sum_{k=1}^n \tau'_k \mathbf{w}(k)^T \mathbf{A}_{kl} = -\boldsymbol{\pi}(l)^T, \quad 1 \leq l \leq n.$$

Postmultiplying by  $\mathbf{1}$ , a column vector of all ones with appropriate dimension

$$\sum_{k=1}^n \tau'_k \mathbf{w}(k)^T \mathbf{A}_{kl} \mathbf{1} = -\boldsymbol{\pi}(l)^T \mathbf{1}, \quad 1 \leq l \leq n.$$

Defining  $\pi'_k = \boldsymbol{\pi}(k)^T \mathbf{1} = \sum_{i \in B_k} \pi_i$ ,  $\lambda'_{b_k, b_l} = \mathbf{w}(k)^T \mathbf{A}_{kl} \mathbf{1} = \sum_{i \in B_k} w_i^k \lambda_{i, B_l}$ ,  $k \neq l$ , and  $\lambda'_{b_k} = -\mathbf{w}(k)^T \mathbf{A}_{kk} \mathbf{1}$ , we get

$$\sum_{\substack{k=1 \\ k \neq l}}^n \tau'_k \lambda'_{b_k, b_l} - \tau'_l \lambda'_{b_l} = -\pi'_l, \quad 1 \leq l \leq n.$$

Thus,  $\boldsymbol{\tau}' = (\tau'_k)_{1 \leq k \leq n}$  satisfies the linear system

$$\boldsymbol{\tau}'^T \mathbf{A}' = -\boldsymbol{\pi}'^T,$$

with  $\boldsymbol{\pi}' = (\pi'_k)_{1 \leq k \leq n}$  and

$$\mathbf{A}' = \begin{pmatrix} -\lambda'_{b_1} & \lambda'_{b_1, b_2} & \cdots & \lambda'_{b_1, b_n} \\ \lambda'_{b_2, b_1} & -\lambda'_{b_2} & \cdots & \lambda'_{b_2, b_n} \\ & & \ddots & \\ \lambda'_{b_n, b_1} & \lambda'_{b_n, b_2} & \cdots & -\lambda'_{b_n} \end{pmatrix}. \quad (47)$$

In summary, under the condition  $\lambda'_{b_k, a} = \lambda'_{b_k} - \sum_{\substack{l=1 \\ l \neq k}}^n \lambda'_{b_k, b_l} \geq 0$ ,  $1 \leq k \leq N$ ,  $\tau'_k = \tau(B_k, Y) (< \infty$  since all states in  $B$  of  $Y$  are transient) is the mean time to absorption in state  $b_k$  of the transient CTMC  $Y'$  with state space  $\{b_1, b_2, \dots, b_N\} \cup \{a\}$ , transition rate matrix (47), and initial probability distribution  $P[Y'(0) = b_k] = \pi'_k$ ,  $1 \leq k \leq N$ . The transition rates  $\lambda'_{b_k, b_l}$  satisfy the conditions of

the theorem. It remains to show that the transition rates to the absorbing state  $\lambda'_{b_k,a}$  also satisfy those conditions and are  $\geq 0$ . First, note that the output rates of  $Y'$  can be written as

$$\lambda'_{b_k} = -\mathbf{w}(k)^T \mathbf{A}_{kk} \mathbf{1} = \sum_{i \in B_k} w_i^k \lambda_i - \sum_{i \in B_k} w_i^k \lambda_{i,B_k}.$$

Then, using  $\lambda'_{b_k,a} = \lambda'_{b_k} - \sum_{\substack{l=1 \\ l \neq k}}^n \lambda'_{b_k,b_l}$  and  $\lambda_{ia} = \lambda_i - \sum_{l=1}^n \lambda_{i,B_l}$ :

$$\begin{aligned} \lambda'_{b_k,a} &= \lambda'_{b_k} - \sum_{\substack{l=1 \\ l \neq k}}^n \lambda'_{b_k,b_l} = \sum_{i \in B_k} w_i^k \lambda_i - \sum_{i \in B_k} w_i^k \lambda_{i,B_k} - \sum_{\substack{l=1 \\ l \neq k}}^n \sum_{i \in B_k} w_i^k \lambda_{i,B_l} \\ &= \sum_{i \in B_k} w_i^k \left( \lambda_i - \sum_{l=1}^n \lambda_{i,B_l} \right) = \sum_{i \in B_k} w_i^k \lambda_{ia} \geq 0. \quad \square \end{aligned}$$

## References

- [1] U. N. Bhat, *Elements of Applied Stochastic Processes*, 2nd edition, John Wiley and Sons, 1984.
- [2] J. A. Carrasco, “Tight Steady-state Availability Bounds using the Failure Distance Concept,” *Performance Evaluation*, 34, September 1998, pp. 27–64.
- [3] J.A. Carrasco, J. Escribá and A. Calderón, “Efficient Exploration of Availability Models Guided by Failure Distances”, *Performance Evaluation Review*, vol. 24, no. 1, May 1996.
- [4] E. Çinlar, *Introduction to Stochastic Processes*, Prentice-Hall, 1975.
- [5] P.J. Courtois and P. Semal, “Bounds for the positive eigenvectors of nonnegative matrices and for their approximations,” *Journal of the ACM*, vol. 31, no. 4, pp. 804–825, October 1984.
- [6] P.J. Courtois and P. Semal, “Computable bounds for conditional steady-state probabilities in large Markov chains and queueing models,” *IEEE J. Select. Areas Commun.*, vol. SAC-4, no. 6, pp. 926–937, September 1986.
- [7] P. Heidelberger, J. K. Muppala and K. S. Trivedi, “Accelerating Mean Time to Failure Computations,” *Performance Evaluation*, vol. 27–28, October 1996, pp. 627–645.
- [8] J.C.S. Lui and R. Muntz, “Evaluating Bounds on Steady-State Availability of Repairable Systems from Markov Models,” in *Numerical Solution of Markov chains*, Marcel Dekker, New York, pp. 435–454, 1991.
- [9] J.C.S. Lui and R.R. Muntz, “Computing Bounds on Steady State Availability of Repairable Computer Systems,” *Journal of the ACM*, vol. 41, no. 4, July 1994, pp. 676–707.
- [10] S. Mahévas and G. Rubino, “Bounding asymptotic dependability and performance measures,” in *Proc. 2nd IEEE Int. Performance and Dependability Symp.*, Urbana-Champaign, USA, September 1996, pp. 176–186.
- [11] S. Mahévas and G. Rubino, “Bound Computation of Dependability and Performance Measures,” INRIA Technical Report no. 3135, March 1997, available from ftp.inria.fr.



- [12] R.R. Muntz, E. de Souza e Silva and A. Goyal, "Bounding Availability of Repairable Computer Systems," *IEEE Trans. on Computers*, vol. 38, no. 12, pp. 1714–1723, December 1989.
- [13] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach*, Dover Publications Inc., New York, 1994, chapter 2.
- [14] J. L. Peterson, *Petri Net Theory and the Modeling of Systems*, Appendix, Prentice-Hall, 1981.
- [15] S.M. Ross, *Stochastic Processes*, John Wiley & Sons, New York, 1983.
- [16] G. Rubino and B. Sericola, "Sojourn Times in Finite Markov Processes," *J. Appl. Prob.*, vol. 27, 1989, pp. 744–756.
- [17] P. Semal, "Refinable Bounds for Large Markov Chains," *IEEE Trans. on Computers*, vol. 44, no. 10, October 1995, pp. 1216–1222.
- [18] E. de Souza e Silva and P.M. Ochoa, "State Space Exploration in Markov Models," *Performance Evaluation Review*, vol. 20, no. 1, June 1992, pp. 152–166.
- [19] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, Princeton, 1994.