RESTART simulation of non-Markov consecutive-k-out-of-n: F repairable systems

José Villén-Altamirano *

Departamento de Matemática Aplicada (E.U. Informática), Universidad Politécnica de Madrid, Calle Arboleda s/n, 28031 Madrid, Spain

ARTICLE INFO

Keywords:
Dependability
Consecutive-k-out-of-n: F repairable system
Rare-event simulation
RESTART

ABSTRACT

The reliability of consecutive-k-out-of-n: F repairable systems and $(k-1)$-step Markov dependence is studied. The model analyzed in this paper is more general than those of previous studies given that repair time and component lifetimes are random variables that follow a general distribution. The system has one repair service which adopts a priority repair rule based on system failure risk. Since crude simulation has proved to be inefficient for highly dependable systems, the RESTART method was used for the estimation of steady-state unavailability, MTBF and unreliability. Probabilities up to the order of 10^{-16} have been accurately estimated with little computational effort. In this method, a number of simulation retrials are performed when the process enters regions of the state space where the chance of occurrence of a rare event (e.g., a system failure) is higher. The main difficulty for the application of this method is to find a suitable function, called the importance function, to define the regions. Given the simplicity involved in changing some model assumptions with RESTART, the importance function used in this paper could be useful for dependability estimation of many systems.

1. Introduction

The reliability of the consecutive-k-out-of-n: F system (or C(k, n: F) system) has aroused great interest since it was first studied by Kontoleon in 1980 [1]. The system consists of a sequence of n ordered components along a line such that the system fails if and only if at least k consecutive components in the system have failed. A list of typical applications of C(k, n: F) systems was given by Yam et al. [2]. For an extensive review of the topic refer to the survey paper by Chao et al. [3] and the books [4,5]. A natural extension of this model is the system of consecutive failures with sparse d, introduced by Zhao et al. [6]. For $d=0$, this system is the C(k, n: F) system. Recently, other extensions of the C(k, n: F) model have appeared in the literature, see e.g., [7–9] and references therein. Another type of system closely related to the C(k, n: F) system is the C(k, n: G) system (see e.g., [10,11]). It consists of a sequence of n ordered components along a line such that the system succeeds if and only if at least k consecutive components in the system succeed.

Although in most related research work, all the components of the system are assumed to have an equal failure rate, this is not always an appropriate assumption, as can be seen in the example provided by Fu [12]. Suppose that we want to transport oil from place A to place B by an oil pipeline and that there are n pressure pumps equally spaced between A and B. Each pump can transport the oil no more than a distance of k pumps. This is obviously a C(k, n: F) system. If pumps $(i-k+1)$ to $(i-1)$ fail and the system still functions, then the pump i has to work very hard to raise the oil pressure to the fixed level. Therefore, pump i will have a higher probability of failure, and the failure rate of a pump would depend on the states of the preceding $(k-1)$ pumps. This dependence is called the $(k-1)$-step Markov dependence. Reliability properties of consecutive k-out-of-n systems with other types of dependence are studied in [13,14].

Additionally, there has been increasing interest in the study of C(k, n: F) repairable systems (see e.g., [15,16]). Lam and Zhang [17] and Lam et al. [18] studied a model for a C(k, n: F) repairable system with $(k-1)$-step Markov dependence. In this model, the lifetime of components and repair times were considered exponential random variables. A priority repair rule based on system failure risk was adopted. Some dependability measures were evaluated by a numerical method. However, a large n would complicate the use of the method. Moreover, repair time and component lifetimes usually do not follow exponential distributions. Xiao et al. [19] revised the model by assuming that repair time is a random variable following a general distribution. In this situation, the system is a non-Markov C(k, n: F) system with $(k-1)$-step Markov dependence. These authors used Monte Carlo simulation to estimate the dependability (including reliability, transient availability, MTTF and MTBF) of the new model. Since crude simulation is inefficient for highly dependable systems, fast simulation methods for rare events such as importance sampling, conditional expectation and a combination of the two methods were used. Xiao and Li [20] studied parameter sensitivity of this model for exponential lifetime distributions using conditional expectation.

In this paper, we extend the model in Ref. [19] by assuming that not only repair time but also the lifetime distribution of components are random variables following a general distribution. Moreover, we estimate a dependability measure of great interest, the steady-state availability, which was not estimated in [19]. We use the rare event simulation method RESTART for estimating all the measures. This method has a precedent, of much more limited scope [21], in the splitting method described in [22]. M. Villén-Altamirano and J. Villén-Altamirano [23] coined the name RESTART and made a theoretical analysis that yields the variance of the estimator and the gain obtained with one threshold. A detailed analysis with multiple thresholds is made in [24].

In this method a more frequent occurrence of a formerly rare event is achieved by performing a number of simulation retrials when the process enters regions of the state space where the importance is greater, i.e., regions where the chance of occurrence of the rare event is higher. These importance regions are defined by comparing the value taken by a function of the system state, called importance function, with certain thresholds. Optimal values for thresholds and the number of retrials that maximize the gain were derived in [24].

The application of this method to particular models requires the choice of a suitable importance function. An inefficiency factor related to the importance function was analyzed in [21] and guidelines for selecting heuristically such a function were provided. In [25], an importance function for estimating reliability measures that could be valid for many systems is provided. In this paper we will use such an importance function to simulate the system.

A limitation of the RESTART and splitting methodologies for simulating highly reliable systems is the difficulty in defining sufficient thresholds. For this reason, L'Ecuyer et al. [26] pointed out that this methodology is not appropriate for this type of system and Xiao et al. [19] suggested that "importance splitting is hard to be adopted for dependability estimation of non-Markov systems, because thresholds function is hard to be presented under this situation". However, as will be shown in the paper, probabilities up to the order of 10^{-16} can be accurately estimated with little computational effort.

The paper is organized as follows: the model C(k, n: F) with $k-1$ step Markov dependence is introduced in Section 2. Section 3 presents the RESTART method and describes some simulation features. In Section 4 numerical examples are presented to demonstrate the method and, finally, conclusions are stated in Section 5.

2. Model assumptions

As mentioned above, the model under study is a C(k, n: F) repairable system with $(k-1)$-step Markov dependence, which has the following features:

(a) Each of the n components is either operational or failed.
(b) The system is either operational or failed.
(c) The system fails if and only if k or more consecutive components have failed.
(d) The lifetime of components has a general distribution. In the numerical examples, we will consider the exponential distribution, whose density is given by:

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0$$

and the Raleigh distribution (Weibull distribution with shape parameter equal to 2), whose density is given by

$$f(t) = \beta t \exp\left(-\frac{\beta}{2}t^2\right), \quad t > 0$$

The failure rates of the previous distributions are λ and βt, respectively.
(e) If there are h $(h < k)$ consecutive failed components that precede the component i, and component lifetimes are exponentially distributed, the residual lifetime X_i of component i will have an exponential distribution of parameter λ_h, $0 < h < k$, with $0 < \lambda < \lambda_1 < \cdots < \lambda_{k-1}$. If the lifetime of components has a Raleigh distribution, the distribution function of the residual lifetime X_i of component i is given by

$$P\{X_i \le t+x / X_i \ge t\} = 1 - \exp\left(-\frac{\beta_h}{2}(x^2 + 2xt)\right),$$

and the failure rate at time $t+x$ of the residual distribution is given by $\beta_h(t+x)$, $0 < h < k$, with: $0 < \beta < \beta_1 < \cdots < \beta_{k-1}$. This assumption represents the effect of $(k-1)$-step dependence.
(f) The repair time of a failed component has a general distribution. In the numerical examples we will consider the lognormal distribution, whose density is given by

$$g(t) = \frac{1}{\sqrt{2\pi}\sigma t}\exp\left(-\frac{(\ln t - \mu)^2}{2\sigma^2}\right), \quad t > 0$$

where μ and σ are the mean and variance of $\ln(T)$, respectively.
(g) After repair, a failed component will be "as good as new".
(h) There is one repairman only, so that no more than one component can be repaired at the same time.
(i) The system is closed when it fails, this means that if the system has failed, then no more components will fail.
(j) The following priority repair rule is adopted. A failed component has the highest repair priority if the system failure risk (defined as the probability that the system will fall into failure states on the next state transitions) after repair of that component is lower than after repair of any other failed component. This implies that the component which, when repaired, causes the highest reduction of the system failure intensity in this model has the largest repair priority. If two or more failed components have the same repair priority, then they will be repaired in the order of arrival.

Our concern is to estimate transient measures, such as system unreliability, and steady-state measures, such as steady-state unavailability and mean time between failures (MTBF).

3. RESTART simulation features in this model

Crude simulation can be used to estimate the dependability of the model given in Section 2. However, as system failure is a rare event in highly reliable models, crude Monte Carlo requires prohibitively long execution times to produce precise estimates. For this reason fast simulation methods for rare event simulations (importance sampling and conditional expectation) were used in [19], but not importance splitting or RESTART because the authors thought that "threshold function is hard to be presented under this situation". However, we use the RESTART method here because it is possible to define a good importance (threshold) function for this system and because it has some advantages over the other methods mentioned. In contrast with importance sampling or conditional expectation, for example, this method is not so dependent on particular features of the system and allows general component lifetime distributions and other generalizations of the model.

3.1. Description of RESTART

This method has been described in detail in several papers, e.g., [21,24]. Nevertheless it is described here in order to make this paper more self-contained.

Let Ω denote the state space of a process $X(t)$ and A the rare set whose probability must be estimated. A nested sequence of sets of states C_i, ($C_1 \supset C_2 \supset \cdots \supset C_M$) is defined, which determines a partition of the state space Ω into regions $C_i - C_{i+1}$; the higher the value of i, the higher the importance of the region $C_i - C_{i+1}$. These sets are defined by means of a function $\Phi : \Omega \to \Re$, called the importance function. Thresholds T_i ($1 \le i \le M$) of Φ are defined so that each set C_i is associated with $\Phi \ge T_i$.

An event at which the system is in a state of the set A is referred to as an event A. Two additional events, B_i and D_i, are defined as follows:

B_i: event at which $\Phi \ge T_i$ having been $\Phi < T_i$ at the previous event;
D_i: event at which $\Phi < T_i$ having been $\Phi \ge T_i$ at the previous event.

RESTART works as follows:

- A simulation path, called main trial, is performed in the same way as if it were a crude simulation. It lasts until it reaches a predefined "end of simulation" condition.
- Each time an event B_1 occurs in the main trial, the system state is saved, the main trial is interrupted, and $R_1 - 1$ retrials $[B_1, D_1)$ are performed. Each of these retrials is a simulation path that starts with the state saved at B_1 and finishes when an event D_1 occurs.
- After the $R_1 - 1$ retrials $[B_1, D_1)$ have been performed, the main trial continues from the state saved at B_1. Note that the total number of simulated paths $[B_1, D_1)$, including the portion $[B_1, D_1)$ of the main trial, is R_1. Each of these R_1 paths is called a trial $[B_1, D_1)$. The main trial, which continues after D_1, leads to new sets of retrials $[B_1, D_1)$ if new events B_1 occur.
- Events B_2 may occur during any trial $[B_1, D_1)$. Each time an event B_2 occurs, an analogous process is set in motion: $R_2 - 1$ retrials $[B_2, D_2)$ are performed, leading to a total number of R_2 trials $[B_2, D_2)$. The trial $[B_1, D_1)$, which continues after D_2, may lead to new sets of retrials $[B_2, D_2)$ if new events B_2 occur.
- In general, R_i trials $[B_i, D_i)$ ($1 \le i \le M$) are performed each time an event B_i occurs in a trial $[B_{i-1}, D_{i-1})$. The number R_i is constant for each value of i.
- A retrial that starts at B_i also finishes if it reaches the "end of simulation" condition before the occurrence of event D_i.

Fig. 1 illustrates a RESTART simulation with $M=3$, $R_1=R_2=4$, $R_3=3$, in which the chosen importance function Φ also defines set A as $\Phi \ge L$. Bold, thin, dashed and dotted lines are used to distinguish the main trial and the retrials $[B_1, D_1)$, $[B_2, D_2)$, and $[B_3, D_3)$, respectively.

Note that for the statistics referring to all the trials, the weight assigned to a trial when it is in the region $C_i - C_{i+1}$ (C_M if $i=M$) must be the inverse of the cumulative number of trials, $1/r_i$

$$r_i = \prod_{j=1}^{i} R_j \quad (1 \le i \le M).$$

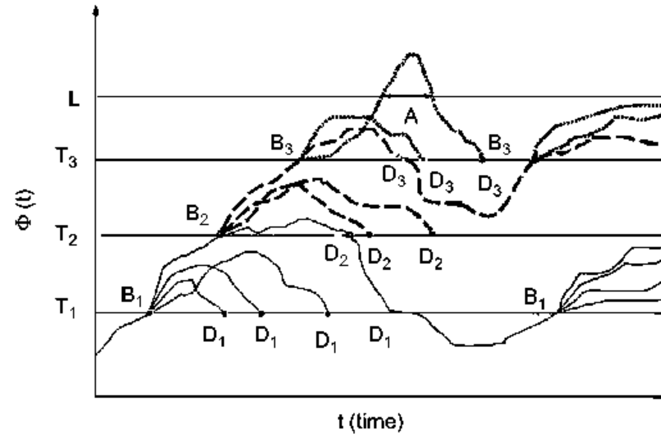

Fig. 1. Simulation with RESTART.

Some more notations

- $P = \Pr\{A\}$; $C_{M+1} = A$; $C_0 = \Omega$;
- $P_{h/i}$($0 \le i \le h \le M+1$): probability of visiting set C_h in a trajectory, given that the system is in a state of the set C_i. For $h \le M$, since $C_h \subset C_i$, $P_{h/i} = \Pr\{C_h\}/\Pr\{C_i\}$;
- N_{Aj}: total number of events A that occur in the j sample of the simulation (in the main trial or in any retrial);
- X_i ($1 \le i \le M$): random variable indicating the state of the system at an event B_i;
- Ω_i ($1 \le i \le M$): set of possible system states at an event B_i;
- P^*_{A/X_i} ($1 \le i \le M$): importance of state X_i, defined as the expected number of events A in a trial $[B_i, D_i)$ when the system state at B_i is X_i. Note that P^*_{A/X_i} is also a random variable since it depends on the initial state X_i, which is random;
- $P^*_{A/i}$ ($1 \le i \le M$): expected importance when the process enters set C_i. If $F(x_i)$ is the distribution function of X_i, then:

$$P^*_{A/i} = E[P^*_{A/X_i}] = \int_{\Omega_i} P^*_{A/x_i}\, dF(x_i).$$

- $V(P^*_{A/X_i})$ ($1 \le i \le M$): variance of the importance when the process enters set C_i:

$$V(P^*_{A/X_i}) = \int_{\Omega_i} (P^*_{A/x_i})^2\, dF(x_i) - (P^*_{A/i})^2.$$

3.2. RESTART estimators

The unreliability of a system at instant t, is defined as the probability of a system failure in the interval (0, t). For estimating the unreliability at instant t, $1-R(t)$, we simulate l times the interval (0, t). In each of the l samples, the trials finish when mission time t is reached or the system fails. The estimator of the unreliability is

$$1 - \hat{R}(t) = \sum_{j=1}^{l} \frac{N_{Aj}}{l r_M}$$

As the event A is a system failure, N_{Aj} is the number of system failures that occur in the jth sample of the simulation. Let us observe that, in contrast with crude simulation, more than one system failure can occur in the same sample. Transient unavailability, defined as the probability of the system being failed at instant t, could be calculated in a similar manner, but without finishing a trial when the system fails.

The steady-state unavailability of a system, $1-A$, is defined as the proportion of time that the system is failed in the long run. The MTBF is defined as the expected amount of time between two consecutive failures of the system. To estimate the steady-state unavailability and MTBF we simulate l long periods of time, each with a short transient period. That is, we use the independent replication method (the batch means method could also be used). The stopping condition in each sample is that the end of the long period is reached by all the survival trials (trials for which event D_i has not occurred). The estimators are

$$1-\hat{A}=\sum_{j=1}^{l}\frac{T_{Fj}}{lTr_M},\quad MTB\hat{F}=\sum_{j=1}^{l}\frac{lTr_M}{N_{Aj}}$$

where T_{Fj} is the total time that the system has failed in the jth sample, and T is the period of time simulated in each sample. We need the results of each sample for constructing the confidence intervals.

3.3. Theoretical speedup

A measure of the efficiency for computing $\hat{P}$ is given by $CV(\hat{P})$, where C is the computer cost and $V(\hat{P})$ is the variance of the estimator. The gain G (also called speedup) obtained can be defined as the ratio of $CV(\hat{P})$ obtained with crude simulation to the same product obtained with RESTART. In [24] it is shown that G is given by

$$G=\frac{1}{f_V f_O f_R f_T}\,\frac{1}{P(-\ln P+1)^2} \tag{1}$$

The term $1/(P(-\ln P+1)^2)$ can be considered the ideal gain. Factors f_V, f_O, f_R and f_T, can be considered inefficiency factors that reduce the actual gain with respect to the ideal gain. Each factor reflects

- f_R: inefficiency due to the non-optimal choice of the number of trials.
- f_T: inefficiency due to the non-optimal choice of the thresholds.
- f_V: inefficiency due to the non-optimal choice of the importance function.
- f_O: inefficiency due to the computer overhead produced by the implementation of RESTART.

In [24] criteria for minimizing factors f_R, f_T and f_O were given. A value of the factor f_R very close to 1 can be achieved if the number of trials is chosen according to:

$$R_i=\frac{1}{\sqrt{P_{i+1/i-1}}},\quad i=1,\ldots,M \tag{2}$$

Factor f_T is minimized by choosing very close thresholds, i.e., $P_{i+1/i}$ close to one. In reliability simulations it is difficult to choose close thresholds and f_T may be significantly greater than one. It is evaluated by

$$f_T=\frac{\left(\sum_{i=0}^{M}\left((1-P_{i+1/i})/\sqrt{P_{i+1/i}}\right)+1\right)^2}{(-\ln P+1)^2} \tag{3}$$

Factor f_O is due to the overheads involved in the implementation of RESTART: (1) for each event, an overhead mainly due to the need to evaluate the importance function and to compare it with the threshold values, and (2) for each retrial, an overhead mainly due to the restoration of the system state and to the rescheduling of the scheduled events. In Section 3.5.2 is explained how to reduce this factor by an efficient rescheduling.

In [21] guidelines for choosing the importance function that minimizes factor f_V were provided. One of the guidelines for reducing f_V is to reschedule the scheduled events at the beginning of each trial. The following upper bound of f_V was defined: $f_V\le Max_{1\le i\le M+1}(s_i)$, and also the exact formula of f_V and of factors s_i. From their analysis it was concluded that s_i may be approximated by

$1+V(P^*_{A/X_i})/(P^*_{A/i})^2$

The value of f_V can be reduced by an appropriate choice of the importance function.

3.4. Importance function

The importance function Φ is a function of the variables which defines the system state $X(t)$ and thus depends on the system under study. According to the definition of the factor f_V, to be suitable an importance function should lead to a small value for the variance $V(P^*_{A/X_i})$. This can be achieved if the importance of all the system states x_i at events B_i is of the same order of magnitude. Roughly speaking, it means that the probability of system failure in a trial $[B_i, D_i)$ must be similar for all the entry states of each set C_i.

In a general system there are minimal cutsets with different cardinalities. In a balanced system, where all the components are equally probable to fail, it is more probable that system failure will be due to the failure of all the components of a minimal cutset with the lowest cardinality. The "distance" to the system failure is related with the number of components that remain operational in the cutset with the lowest number of operational components. For this reason, the importance function (at an instant t) is defined as

$$\Phi(t)=cl-oc(t) \tag{4}$$

where cl is the cardinality of the minimal cutset with lowest cardinality and $oc(t)$ is the number of components that are operational at time t in the cutset with lowest number of operational components. Thresholds T_i are $1,2,\ldots,cl-1$ for estimating transient measures and $2,\ldots,cl-1$ for estimating steady-state measures, because in the last case the next event when all the component are operational is a failure, and so the first threshold would frequently be reached if the threshold were defined at $T=1$. However, for estimating the reliability at an instant t, most of the simulations of the interval $(0, t)$ will finish without any component failing, and so it is worth defining a threshold at $T=1$. This importance function was introduced in [25] and is valid for simulating many highly dependable systems.

In our model we have $(n-k+1)$ minimal cutsets. As all of them have the same cardinality (k), the definition of the importance function can be expressed as: "the number of components that are down at time t in the cutset with greatest number of failed components". The main differences between the importance, P^*_{A/X_i}, of the system states x_i at events B_i are: (i) whether the failed components of the cutset are consecutive or not and (ii) the total number of components in the systems that are down when the process enters each set C_i. The greater is that number, the greater is the importance of the system state. The importance is also greater if the failed components of the cutset (with greatest number of failed components) are consecutive. It seems that the difference between the importance of these states could be relatively small, and could be of the same order of magnitude in most cases. Thus, the variance $V(P^*_{A/X_i})$ could be small. Simulation results will corroborate this conjecture.

3.5. Simulation process

3.5.1. Scheduling lifetimes with exponential distribution

If lifetime distribution of all the components is exponential, the time for the next failure is determined by $\Delta t=-(1/\Lambda)\ln(u)$ where u is a random number of the $U(0,1)$ distribution and $\Lambda=\sum_{i\in O}\lambda_i$, where O is the set of operational components. The next component that fails is chosen at random with probability λ_i/Λ, $i \in O$. This rule cannot be applied to other lifetime distributions.

3.5.2. Rescheduling

As mentioned above, rescheduling the scheduled events at the beginning of each retrial increases the efficiency of RESTART. Due to the memory-less property of the exponential distribution, it is straightforward to reschedule the residual lifetime (or service time) of components that are exponentially distributed. Since only the next component failure has to be rescheduled, factor f_O is lower with exponential distribution. For the Weibull distribution, it is also possible to generate the residual lifetime (or service time) by means of a formula since the inverse of the distribution function of the residual lifetime has a closed form.

For other distributions that have no explicit formula of the distribution function, e.g. lognormal distribution, we can use two procedures:

(i) Obtaining a random value of the whole lifetime (or service time) of a component. If that value is greater than the lifetime at the current time, the residual lifetime is obtained as the difference between the two amounts. Otherwise a new random value is obtained and so on. This procedure had been used in previous studies, e.g., [25].
(ii) Calculating the approximate distribution function, $F(c)$, of the current lifetime c of a component by looking for the closest value in a table of the distribution function obtained numerically. Then obtaining the distribution function of the whole lifetime by generating a uniform random number y in the interval $(F(c), 1)$. Finally, the residual lifetime is obtained as $F^{-1}(y)-c$. In our experiments, it has been possible to reduce the computational time (and thus, factor f_O) up to 3.7 times by applying this procedure for generating residual lognormal distributed repair times.

Slightly better results were obtained in our test cases applying procedure (i) a maximum of 3–4 times, and when the residual lifetime was not obtained, by applying procedure (ii).

For distributions that have a closed-form formula of the distribution function but not of the inverse of the distribution function of the residual lifetime, procedure (ii) is recommended but the exact values of $F(c)$ and of $F^{-1}(y)$ should be obtained with the formula.

3.5.3. Relative error

The interval width is evaluated using the independent replication method with a non-fixed number of replicas (samples). After each sample, the half width of the 95% t-Student confidence interval divided by the estimate (relative error) is calculated. The number of samples is not fixed beforehand and the simulation finishes when the relative error is smaller than 0.1. Steady-state unavailability and MTBF are estimated at the same time, and the stopping condition is applied for unavailability. The relative error in the estimation of MTBF is also calculated in the simulation, being slightly greater than 0.1 in most cases while slightly lower than 0.1 in others.

4. Test cases

We simulated models with component lifetimes exponentially and Weibull distributed. The repair time of a failed component has a lognormal distribution with parameters μ=1.21 and σ=0.8 in all the models. The simulations were run on a Pentium(R) D CPU at 3.01 GHz.

4.1. C(4, 60: F) repairable system

First, the same example as studied in [19] was simulated. The example is a C(4, 60: F) repairable system with three-step Markov dependence modelled as in Section 2. If there are l (l=0, 1, 2, 3) consecutive failed components that precede the component i, the residual lifetime X_i of component i has an exponential distribution with failure rates λ_i=0.001, 0.0012, 0.0017 and 0.0021, respectively (called model EL A in Table 1). Simulation was also made assuming that component lifetimes follow a Raleigh distribution (Weibull distribution with shape parameter equal to 2) with scale parameters β_i=0.00000157, 0.00000226, 0.00000454 and 0.00000693, respectively (model WL A in Table 1). The mean lifetime of the components was the same in both models. In model WL A' the components were not new, but were 640 units of time old at the beginning of the simulation. Unreliability was estimated for different small intervals $(0, t_e)$. The importance function is given by Eq. (4) with cl=4. The number of retrials were chosen according to Eq. (2), with the values of $P_{i+1/i-1}$ estimated in pilot runs (one or two for each case). We set three thresholds and the rare set can be reached in trials $[B_3, D_3)$. Results are listed in Table 1 with computing times, speedups and factor estimates.

Accurate results were obtained with short computational times (less than 12 min) for estimating probabilities up to the order of 10^{-13}. We can observe that much higher reliability is obtained with WL A model than with EL A even though the mean lifetime of the components is the same in both models. This is due to the fact that the Raleigh distribution has an increasing failure rate and, for t=1, the constant failure rate of the exponential distribution is 640 times the failure rate of the Raleigh distribution. Reliability of the same order of magnitude is obtained in model WL A' because at the beginning of the simulation the failure rate is the same in both models. The strong increase in reliability obtained using new instead of old components is of particular note.

To evaluate the gain in time with respect to a crude simulation, the computational time needed to achieve a relative error of 10% with crude simulations was measured for the EL A model with t_e=25 and with t_e=5, and for the WL A' model with t_e=25. The measured values were extrapolated to the other cases. For example, for the EL A model with t_e=5 the computational time with crude simulation was 12.8 h. As the number of samples for

Table 1
Unreliability estimates for C(4, 60: F) system. Relative error=0.1.

Model	Interval $(0, t_e)$	$1-\hat{R}(t)$	Run-time (s)	Gain in time	Factor f_T	Factors $f_V \times f_O$
EL A	(0, 25)	3.4×10^{-6}	3	4.8×10^{2}	2.6	1.3
EL A	(0, 5)	3.8×10^{-8}	6	7.7×10^{3}	4.9	2.1
EL A	(0, 1)	1.2×10^{-10}	24	4.9×10^{5}	12.5	2.4
WL A	(0, 25)	8.6×10^{-13}	697	2.1×10^{8}	28.6	2.4
WL A'	(0, 25)	8.3×10^{-6}	11	1.6×10^{1}	2.4	2.0
WL A'	(0, 1)	2.0×10^{-10}	149	3.8×10^{5}	11.0	2.2

achieving a given relative error is inversely proportional to the probability being estimated, the number of samples for the EL A model with t_e=1 is 317 times greater. The measured computational time of each sample simulated with t_e=1 is 1.24 times lower than that with t_e=5. Thus, the extrapolated computational time for simulating the EL A model with t_e=1 is 3272 h. The gain in time shown in Table 1 is obtained by dividing this time by 24 s. This extrapolation mode is checked by comparing the measured time for t_e=5 with the time obtained by extrapolating the measured time for t_e=25.

It is interesting to compare the measured gain with the theoretical gain derived from Eq. (1). As the value of R_i given by Eq. (2) is taken, then f_R=1. Factor f_T is given by Eq. (3), with the values of $P_{i/(i-1)}$ estimated in the simulation. If we assume f_V=1 and f_O=1 in Eq. (1), we obtain a theoretical gain equal to 612 for the EL A model with $t_e = 25$. We see that the theoretical gain (for f_V=1, f_O=1) is 1.3 times the actual gain in time. The ratio 1.3 can be taken as an estimate of $f_V \times f_O$ for this case. The low values of $f_V \times f_O$ in all the cases show that the choice of $\Phi(t)$=$cl - oc(t)$ is appropriate and that the application is efficient for the tested models, at least for the cases tested. The low values of factor f_V observed indicates that all the system states at the entry state of the same set C_i have a similar importance, or at least must be of the same order of magnitude. For example if set C_i is reached when two components of the same cutset fail, the importance is greater if the failed components are consecutive and/or there are many components in the repair queue. However, if the two failed components are not consecutive the importance is not much lower, and the number of components in the repair queue is not so important due to the way of selecting the component that will be repaired.

The accuracy of the simulation results was validated in two ways. First, by checking that similar results are obtained with crude and RESTART simulations, in those cases where crude simulation is feasible. Second, by simulating a C(4, 5: F) non-reparable system without Markov dependence (that is, λ_1=λ_2=λ_3=λ_4) and checking that the analytical results of this system, both for exponential and Weibull lifetimes, are within the confidence intervals obtained by simulation.

System steady-state unavailability and MTBF were estimated for model EL A and also for exponential lifetime distribution with the following failure rates: λ_i=0.0003, 0.00036, 0.00051, and 0.00063 (model EL B) and λ_i=0.0001, 0.00012, 0.00017, and 0.00021 (model EL C). In models WL A, WL B and WL C, component lifetimes follow a Raleigh distribution with the same mean lifetime as in the EL A, EL B and EL C models, respectively. In all the models the repair time of a failed component has a lognormal distribution with parameters μ=1.21 and σ=0.8. Results are listed in Table 2 with computing times, speedups and factor estimates.

The values of the gain and of the factors was obtained in the same way as those of Table 1. It can be observed that more computational time than in the previous case is needed for estimating probabilities of the same order of magnitude. Factor f_T is greater because, as mentioned in Section 3.4, it is only possible to set two intermediate thresholds given that the first threshold is reached when two components of the same cutset are failed.

As expected, factor f_O and thus the product $f_V \times f_O$ is greater than in Table 1. However, it is low enough to allow probabilities up to the order of 10^{-10} to be estimated accurately with short computational times. The results also corroborate that the importance function $\Phi(t)$=$cl - oc(t)$ is appropriate and that the application is efficient for the tested models. Factor f_O is lower with exponential models because the rescheduling of the lifetimes of the components at each retrial is faster with exponential distribution and also because of the way the next failed component is rescheduled, as commented in Section 3.5.2.

To estimate probabilities lower than, say, 10^{-20} would require much more computational time because factor f_T increases as the probability to be estimated decreases since it is not possible to set more than three intermediate thresholds. Also the product $f_V \times f_O$ increases slightly with lower probabilities. The example applications are not asymptotically efficient for the same reason: it is not possible to set a sufficient number of thresholds when the probability of the rare set tends to zero [21]. Nevertheless, very low probabilities, low enough for most applications, can be estimated with reasonable computational effort.

The accuracy of the simulation results was validated by checking that similar results are obtained with crude simulations, as well as by simulating a C(4, 4: F) repairable system with exponential repair time and without Markov dependence. With these assumptions, the repair queue can be modelled as an M/M/1 queue. The analytical result given by this model was seen to be within the confidence interval obtained by RESTART simulation.

4.2. *C(6, 60: F) repairable system*

We also simulated a C(6, 60: F) repairable system with 5-step Markov dependence. In cases where l ($l < 6$) consecutive failed components precede the component i, the residual lifetime X_i of component i has an exponential distribution with failure rates (λ_i) of 0.001, 0.0012, 0.0017, 0.0021, 0.0025 and 0.0029, respectively (called model EL D in Table 3). Simulation was also made assuming that component lifetime distributions were Raleigh with scale parameters β_i=0.0000157, 0.0000226, 0.0000454, 0.0000693, 0.0000982 and 0.0001149, respectively (model WL D in Table 3). Although the mean lifetime of the Raleigh components were $\sqrt{10}$ times lower than that of exponential components, the reliability was greater because it was estimated for small intervals (0, t_e) for which the failure rate is smaller with Raleigh lifetimes, as commented in Section 4.1.

System steady-state unavailability and MTBF were estimated for an exponential lifetime distribution with failure rates obtained as 0.3 and 0.1 times the failure rates of model EL D (models EL E and EL F, respectively). In WL E and WL F models, component

Table 2
Unavailability and MTBF estimates for C(4, 60: F) system. Relative error=0.1.

Model	$1-\hat{A}$	MTBF	Run-time (s)	Factor f_T	Gain in time	Factors $f_V \times f_O$
EL A	1.6×10^{-5}	5.8×10^5	4	2.6	3.8×10^1	4.4
EL B	4.6×10^{-8}	2.8×10^8	12	9.0	1.3×10^3	5.9
EL C	4.5×10^{-10}	1.7×10^{10}	35	34.9	1.4×10^4	8.6
WL A	2.2×10^{-5}	8.9×10^5	12	2.5	1.4×10^1	9.4
WL B	8.8×10^{-8}	1.4×10^8	45	8.2	2.6×10^2	5.9
WL C	5.8×10^{-10}	1.4×10^{10}	194	32.8	3.0×10^3	34.8

Table 3
Unreliability estimates for C(6, 60: F) system. Relative error=0.1.

Model	Interval (0, t_e)	$1-\hat{R}(t)$	Run-time (s)	Gain in time	Factor f_T	Factors $f_V \times f_O$
EL D	(0, 25)	1.8×10^{-9}	17	2.0×10^{5}	2.8	2.2
EL D	(0, 5)	1.6×10^{-12}	32	2.8×10^{7}	5.6	4.9
WL D	(0, 25)	3.4×10^{-13}	218	1.2×10^{8}	5.9	4.8
WL D	(0, 12)	2.1×10^{-16}	791	5.0×10^{10}	13.3	5.2

Table 4
Unavailability and MTBF estimates for C(6, 60: F) system. Relative error=0.1.

Model	$1-\hat{A}$	MTBF	Run-time (s)	Factor f_T	Gain in time	Factors $f_V \times f_O$
EL E	4.5×10^{-11}	2.3×10^{11}	30	5.2	9.7×10^{5}	7.1
EL F	4.6×10^{-14}	2.4×10^{14}	84	15.4	1.1×10^{8}	12.5
WL E	4.8×10^{-11}	3.3×10^{11}	195	5.4	1.4×10^{5}	44.0
WL F	3.4×10^{-14}	3.5×10^{14}	715	17.7	1.8×10^{7}	91.2

lifetimes follow a Raleigh distribution with the same mean lifetime as in the EL E and EL F models, respectively. Results are listed in Table 4 with computing times, speedups and factor estimates.

The importance function is given by Eq. (4) with cl=6. In these models it is possible to set four intermediate thresholds to study steady-state unavailability and MTBF and five intermediate thresholds to study unreliability. As a consequence, factor f_T (and, thus, computer time) is significantly reduced. Therefore, the gain is greater than in the C(4, 60: F) system for estimating probabilities of the same order of magnitude. This fact also allows lower probabilities (up to the order of 10^{-16}) to be estimated with moderate computational effort. Unlike with importance sampling, RESTART works better with higher redundancies because it allows more thresholds to be set.

4.3. *Other similar systems*

A great advantage of RESTART compared with importance sampling or conditional expectation is the simplicity involved in changing some model assumptions, given that no analytical study is required in most cases. For example, for simulating the Weibull model with this method it is simply necessary to change some lines of the computer program used for the exponential model. Spare components, other types of dependence or even failure propagation (e.g., the failure of a component causes some other components to fail with given probabilities) can be easily considered in the model, in a similar manner as was done in example 2 of [25]. Another possible extension is to consider several repairmen instead of one, as in example 1 of [25]. To implement this, we would first have to write one future event for each repairman in the list of future events of the simulation program (as in crude simulation), and then reschedule several repair times instead of one at the beginning of each retrial. So, it would not be difficult to implement this extension.

It is a straightforward task to study circular systems like those considered in [2,27,28]. This system for short consists of a sequence of n ordered components along a circle (instead of along a line) so that the system fails if, and only if, at least k consecutive components in the system fail. To simulate this system with RESTART it is merely necessary to add $k-1$ cutsets to the list of cutsets of the C(k, n: F) system. The changes in the simulation program are just the same as we had to make with crude simulation. For the same parameters as in the EL E model of Table 4, we obtained an unavailability of 4.9×10^{-11} with 29 s of computational time and for the EL F model an unavailability of 5.0×10^{-14} with 81 s of computational time.

The same importance function could also be applied for studying other models such as the consecutive k-out-of-r-from-n system (see e.g., [29,30]). This system has n ordered elements and fails if less than k out of r consecutive elements are in working condition. In this model there are more cutsets, all of them with cardinality $r-k+1$. To simulate this system with RESTART, we would have to make the same changes in the list of cutsets of the C(k, n: F) system as with crude simulation. Another similar system is the k-out-of-r-from-n system with multiple failure criteria, proposed by Levitin [31], although in this model the cutsets show different cardinality. For example, the (1, 2)-out-of-(3, 5)-from-12 system with multiple failure criteria works if at least one of each three consecutive components works and 2 of each 5 consecutive components. The minimal cutsets are: (1,2,3), (2,3,4), …, (10,11,12), (corresponding to the first criterion) and (1,2,4,5), (2,3,5,6), …, (8,9,11,12) (corresponding to the second criterion). The same importance function with cl=3 could be applied. The first threshold, for example, is reached if either a component of a cutset with cardinality 3 or two components of a cutset with cardinality 4 fail. However, the importance function used in this paper is not valid for some models, e.g., the generalized k-out-of-n system proposed by Cui and Xie [32] or the linear multi-state sliding window system proposed by Levitin [33]. Further investigation is required to find a better importance function for those models.

The splitting technique described in [22] could be applied for estimating the unreliability but the run-time would be slightly greater than with RESTART. However, splitting is useless for estimating the steady-state unavailability because it is only valid for simulations made by means of short replicas, e.g., short transient simulation or regenerative simulation of simple systems [21]. The two variants of splitting that use resplits, as described in [26], might also be appropriate for this problem, given that those variants are closer to the RESTART method than to splitting.

5. Conclusions

The paper studies dependability estimation for a consecutive-k-out-of-n: F repairable system with non-exponential component lifetimes and repair time distributions and $(k-1)$-step Markov dependence. Different measures, including unreliability, MTBF and steady-state unavailability of highly dependable systems, are estimated by RESTART, a general method for rare event simulation. Probabilities up to the order of 10^{-16} have been accurately estimated with short computational effort.

The advantage of RESTART over other methods, such as importance sampling or conditional expectation, is that it is not so dependent on particular features of the system, e.g., component lifetime distributions. Only the importance function depends on the system being simulated. However, the same importance function as used in this paper can be applied to different models without additional analytical effort, regardless of the level of redundancy or of whether the model is Markovian in nature or not. This feature could extend the use of our method for dependability estimation to many other systems. In this paper, we have extended the model introduced in [19] to the case of non-exponential component lifetimes, estimating the steady-state unavailability that was not estimated therein. We have also studied higher redundancy cases and circular systems. Other possible extensions suggested in [19], such as considering r repairmen, are straightforward with RESTART. The same importance function could be applied to other models like consecutive k-out-of-r-from-n system with multiple failure criteria proposed by Levitin [31]. However, extension to some models like those considered in [32,33] would require further investigation to improve the importance function used in this paper.

Acknowledgements

The author thanks Mr. José Ángel Delgado Valderas for his assistance in writing the simulation programs which were used to obtain the numerical results presented in Section 4. The author is also grateful for the helpful comments of M. Villén-Altamirano that have improved the paper.

References

[1] Kontoleon JM. Reliability determination of r-successive-out-of-n: F system. IEEE Trans Reliab 1980;29(5):600–2.
[2] Yam R, Zuo MJ, Zhang YL. A method for evaluation of reliability indices for repairable circular consecutive-k-out-of-n: F systems. Reliab Eng Syst Saf 2003;79(1):1–9.
[3] Chao MT, Fu JC, Koutras MV. Survey of reliability studies of consecutive k-out-of-n: F and related systems. IEEE Trans Reliab 1995;44(1):120–7.
[4] Chang GJ, Cui LR, Hwang FK. Reliabilities of consecutive-k systems. Boston: Kluwer; 2000.
[5] Kuo W, Zuo MJ. Optimal reliability modelling: principles and applications. New York: Wiley; 2003.
[6] Zhao X, Cui L, Kuo W. Reliability for sparsely connected consecutive-k systems. IEEE Trans Reliab 2007;56(3):516–24.
[7] Demir S. Reliability of combined k-out-of-n and consecutive k_c-out-of-n systems of Markov dependent components. IEEE Trans Reliab, doi:10.1109/TR.2009.2026816.
[8] Mohan P, Agarwal M, Sen K. Combined m-consecutive-k-out-of-n: F & consecutive k_c-out-of-n: systems. IEEE Trans Reliab 2009;58(2):328–36.
[9] Eryilmaz S, Kan C, Akici F. Consecutive k-within-m-out-of-n: F system with exchangeable components. Naval Res Logist 2009;56(6):503–10.
[10] Gera AE. A consecutive k-out-of-n: G system with dependent elements–a matrix formulation and solution. Reliab Eng Syst Saf 2000;68(1):61–7.
[11] Wu Y, Guan J. Repairable consecutive-k-out-of-n: G systems with r repairmen. IEEE Trans Reliab 2005;54(2):328–37.
[12] Fu JC. Reliability of a large consecutive-k-out-of-n: F systems with $k-1$-step Markov dependence. IEEE Trans Reliab 1987;36(1):75–7.
[13] Cui L, Li H. Analytical method for reliability and MTTF assessment of coherent systems with dependent components. Reliab Eng Syst Saf 2007;92(3):300–7.
[14] Eryilmaz S. Reliability properties of consecutive k-out-of-n systems of arbitrarily dependent components. Reliab Eng Syst Saf 2009;94(2):350–6.
[15] Cheng K, Zhang YL. Analysis for a consecutive-k-out-of-n: F repairable system with priority in repair. Int J Syst Sci 2001;32(5):591–8.
[16] Guan J, Wu Y. Repairable consecutive-k-out-of-n: F system with fuzzy states. Fuzzy Sets Syst 2006;157(1):121–42.
[17] Lam Y, Zhang YL. Repairable consecutive-k-out-of-n: F system with Markov dependence. Nav Res Logist 2000;47(1):18–39.
[18] Lam Y, Keung H, Ng T. A general model for consecutive-k-out-of-n: F repairable system with exponential distribution and $(k-1)$-step Markov dependence. Eur J Oper Res 2001;129(3):663–82.
[19] Xiao G, Li Z, Li T. Dependability estimation for non-Markov consecutive-k-out-of-n: F repairable systems by fast simulation. Reliab Eng Syst Saf 2007;92(3):293–9.
[20] Xiao G, Li Z. Estimation of dependability measures and parameter sensitivities of a consecutive-k-out-of-n: F repairable system with $(k-1)$-step Markov dependence by simulation. IEEE Trans Reliab 2008;57(1):71–83.
[21] Villén-Altamirano M, Villén-Altamirano J. On the efficiency of RESTART for multidimensional state systems. ACM Trans Model Comput Sim 2006;16(3):251–79.
[22] Kahn H, Harris TE. Estimation of particle transmission by random sampling. Natl Bur Stand Appl Math Ser 1951;12:27–30.
[23] Villén-Altamirano M, Villén-Altamirano J. RESTART: a method for accelerating rare event simulations. In: Proceedings of the 13th International Teletraffic Congress, Copenhagen, 1991. p. 71–6.
[24] Villén-Altamirano M, Villén-Altamirano J. Analysis of RESTART simulation: theoretical basis and sensitivity study. Eur Trans Telecom 2002;13(4):373–86.
[25] Villén-Altamirano J. Importance functions for RESTART simulation of highly-dependable systems. Simulation 2007;83(12):821–8.
[26] L'Ecuyer P, Demers V, Tuffin B. Rare events, splitting, and quasi-Monte Carlo. ACM Trans Model Comput Sim 2007;17(2) article 9.
[27] Zhang YL, Zuo MJ, Yam RCM. Reliability analysis for a circular consecutive-2-out-of-n: F repairable system with priority in repair. Reliab Eng Syst Saf 2000;68(2):113–20.
[28] Yun WY, Kim GR, Yamamoto H. Economic design of a circular consecutive-k-out-of-n: F system with $(k-1)$-step Markov dependence. Reliab Eng Syst Saf 2007;92(4):464–78.
[29] Habib A, Szántai T. New bounds on the reliability of consecutive k-out-of-r-from-n: F system. Reliab Eng Syst Saf 2000;68(2):97–104.
[30] Dutuit Y, Rauzy A. New insight into the assessment of k-out-of-n and related systems. Reliab Eng Syst Saf 2001;72(3):303–14.
[31] Levitin G. Consecutive k-out-of-r-from-n system with multiple failure criteria. IEEE Trans Reliab 2004;53(3):394–400.
[32] Cui LR, Xie M. On a generalized consecutive k-out-of-n system and its reliability. Int J Syst Sci 2005;36(5):267–74.
[33] Levitin G. Common supply failures in linear multi-state sliding window system. Reliab Eng Syst Saf 2003;82(1):55–62.