

CENTRAL LIMIT THEOREM FOR HITTING TIMES OF FUNCTIONALS OF MARKOV JUMP PROCESSES.

CHRISTIAN PAROISSIN¹ AND BERNARD YCART¹

Abstract. A sample of i.i.d. continuous time Markov chains being defined, the sum over each component of a real function of the state is considered. For this functional, a central limit theorem for the first hitting time of a prescribed level is proved. The proof uses a functional central limit Theorem for the sum, together with exponential bounds on the tail probabilities. The result extends the classical central limit theorem for order statistics. Various reliability models are presented as examples of applications.

Résumé. Etant donné un échantillon de processus markoviens de saut i.i.d., on considère la somme des images de chaque composant par une fonction réelle de l'état de ce composant. Pour cette fonctionnelle, on démontre un théorème central limite pour le temps d'atteinte d'un niveau donné. La démonstration utilise un théorème fonctionnel pour le processus somme, plus une borne exponentielle pour la queue de distribution. Le résultat étend le théorème central limite classique pour les statistiques d'ordre. Plusieurs modèles en fiabilité sont envisagés à titre d'exemples d'applications.

1991 Mathematics Subject Classification. 60F05, 60J25, 60K10.

The dates will be set by the publisher.

1. INTRODUCTION

Models made of n -tuples of i.i.d. Markov chains appear in different domains of applications. They can represent locations of a set of customers in a closed Jackson network [29], the current state of a parallel MCMC algorithm [10], or a description of a coherent system in reliability [3]. We shall use here the reliability interpretation as a support for intuition, although the applicability of our results is not restricted to that area. In [27], the relaxation time of samples of i.i.d. Markov chains was studied and a cutoff phenomenon was shown to occur. In [28] the hitting time of functionals of such a sample was used for detecting access to equilibrium, thus providing stopping rules for MCMC algorithms. Our main focus here is on the behavior of those functionals far from equilibrium.

Let $X = \{X(t), t \geq 0\}$ be a Markov chain in continuous time, with values in a finite state space E . The random variable $X(t)$ will be interpreted as the state at time t of a component bound to degradations and failures. Consider a system made of n independent such components. Its state will be described by a n -tuple $\tilde{X} = (X_1, \dots, X_n)$ where the X_i 's are independent copies of X . This process is again a continuous time Markov chain $\{\tilde{X}(t), t \geq 0\}$, with values in the product space E^n . Let f be a function from E to \mathbb{R} . In the reliability

Keywords and phrases: central limit Theorem, hitting time, exponential inequalities, reliability, failure time.

¹ Laboratoire MAP5, 45 rue des Saints-Pères 75270 Paris cedex 06, France; e-mail: parc@math-info.univ-paris5.fr & ycart@math-info.univ-paris5.fr

interpretation, $f(e)$ measures a level of degradation for a component in state e . The total degradation of the system in state $\eta = (\eta_i) \in E^n$ will be measured by the sum $\sum_i f(\eta_i)$. So we shall focus on the process $S_n = \{S_n(t), t \geq 0\}$, where $S_n(t)$ is the total degradation of the system at time t :

$$S_n(t) = \sum_{i=1}^n f(X_i(t)) .$$

It is natural to consider the instants at which $S_n(t)$ reaches a prescribed level of degradation. Let $k = (k(n))$ be a sequence of reals. Our main object is the *failure time* T_n , defined as:

$$T_n = \inf\{t \geq 0; S_n(t) \geq k(n)\} .$$

In the particular case where $E = \{\textit{working}, \textit{failed}\}$ (binary components), and f is the indicator of a failed component, then $S_n(t)$ simply counts the number of failed components at time t , and our system is a so-called “ k -out-of- n ” system [4].

Let $e_0 = X(0)$ be the initial state of all components (in reliability e_0 will be the *perfect state*). We shall denote by $m(t)$ (respectively: $v(t)$) the expectation (resp.: the variance) of the degradation at time t .

$$m(t) = \mathbb{E}[f(X(t)) | X(0) = e_0] , \quad v(t) = \text{Var}[f(X(t)) | X(0) = e_0] .$$

Since all components are i.i.d., the average degradation $S_n(t)/n$ converges in probability to its expectation $m(t)$. We shall assume that $m(t)$ is strictly increasing on the interval $[0, \tau]$, with $0 < \tau \leq +\infty$ (the degradation starting from the perfect state increases on average). We consider a “mean degradation level” α , such that $m(0) < \alpha < m(\tau)$. Assume the threshold $k(n)$ is such that:

$$k(n) = \alpha n + o(\sqrt{n}) .$$

The failure time T_n has to be close to the instant at which $m(t)$ crosses level α :

$$t_\alpha = \inf\{t; m(t) = \alpha\} .$$

Our main result is the following central limit Theorem.

Theorem 1.1. *Under the above hypotheses,*

$$\sqrt{n}(T_n - t_\alpha) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2) ,$$

with:

$$\sigma^2 = \frac{v(t_\alpha)}{(m'(t_\alpha))^2} .$$

The proof will be decomposed into several steps, two of which are of independent interest. We start by proving a functional CLT for the process S_n (Proposition 2.1). This result is essentially an application of general convergence Theorems by Hahn [14] and Whitt [26], related to a huge corpus of diffusion approximations (see Ethier and Kurtz [12] for a general reference). Once properly normalized, S_n converges to a Gaussian process which is also a diffusion and a Brownian martingale.

Besides the asymptotic distribution of S_n , the proof involves control on the probability of tail events. In the reliability interpretation, this amounts to proving an exponential inequality for the probability of uninterrupted functioning up to time t , which is called the *reliability* of the system. To the best of our knowledge, Proposition 2.2 is new, even though many bounds have been proposed for reliabilities (cf. for instance [6, 8, 9, 13]). Section 2 is devoted to the precise statements and proofs of our results. It ends with the complete proof of Theorem 1.1.

In section 3 we give some examples of reliability models to which our results apply. The simplest of them is the k -out-of- n model, with binary components having exponential failure and repair times (see Bhat [5]). Our results also apply when the failure and repair times have phase type distributions [18, 19, 25], and we shall study the case where a deterministic repair time is approximated by an Erlang distribution. In the particular case of a k -out-of- n system with non-repairable components, the failure time T_n of the system is an order statistic of failure times for the components. The central limit Theorem for order statistics is well known (see for instance [23]). Here it comes out as a particular case of Theorem 1.1. A model with 3 states components, studied by Pham et al. [21] will be considered next.

2. ASYMPTOTIC RESULTS

Our main results are stated and proved in this section. The first one describes the asymptotics for the renormalized degradation process S_n (Proposition 2.1). The second one is an exponential bound for the reliability (Proposition 2.2). The proof of Theorem 1.1 will follow.

Let us first precise some notations. As explained before, the state of one component is described by a continuous time Markov chain $X = \{X(t), t \geq 0\}$. Its infinitesimal generator will be denoted by Λ , and its transition rate from state e to state e' by $\lambda_{ee'}$. It is well known (see for instance Bhat [5] section 7.7, p. 219-224), that the distribution at time t of $X(t)$, starting from e_0 , is the e_0 -th column of the matrix $\exp({}^t\Lambda t)$. It will be denoted by $p(t) = (p_e(t))$, $e \in E$. (All our vectors are columns and tA denotes matrix transpose).

We consider a n -tuple $\tilde{X} = (X_1, \dots, X_n)$ of independent copies of X . It is a continuous time Markov chain on E^n , and its infinitesimal generator is the Kronecker sum of n copies of Λ ([27] p. 92-94). We shall also need the ‘‘counting process’’ N_n , defined as the number of components of \tilde{X} in each state of E : $N_n = \{N_n(t), t \geq 0\}$, with:

$$N_n(t) = (N_{n,e}(t)), e \in E \text{ and } N_{n,e}(t) = \sum_{i=1}^n \mathbb{1}_e(X_i(t)),$$

where $\mathbb{1}_e$ denotes the indicator function of state e . Thus N appears as the sum of n independent copies of the Markov chain $\mathbb{1}(X) = (\mathbb{1}_e(X))$, $e \in E$. The expectation of $\mathbb{1}(X(t))$ is $p(t)$. For $0 \leq s \leq t$ the covariance matrix of $\mathbb{1}(X(s))$ with $\mathbb{1}(X(t))$, denoted by $K(s, t)$ can be expressed as:

$$K(s, t) = \text{diag}(p(s)) \exp(\Lambda(t-s)) - p(s) {}^t p(t). \quad (2.1)$$

The degradation function f , from E to \mathbb{R} will be identified to the vector $(f(e))$, $e \in E$. The degradation at time t of one component can be written as a linear combination of $\mathbb{1}(X(t))$:

$$f(X(t)) = {}^t f \mathbb{1}(X(t)).$$

Thus its expectation at time t is:

$$m(t) = {}^t f p(t).$$

For $0 \leq s \leq t$ the covariance function is:

$$\text{Cov}[f(X(s)), f(X(t))] = {}^t f K(s, t) f.$$

In particular for $s = t$, the variance of the degradation at time t is:

$$v(t) = \text{Var}[f(X(t)) | X(0) = e_0] = {}^t f K(t, t) f.$$

It is easy to prove that the counting process N_n is again a continuous time Markov chain (see [29] p. 296 for the Jackson network interpretation of N). Since the X_i 's are independent, the distribution of the random

vector $N_n(t)$ is multinomial with parameters n and $p(t)$. Of course, the “total degradation” process $S_n(t)$ is a linear combination of $N_n(t)$: $S_n(t) = {}^t f N_n(t)$.

Our first result is a functional CLT for S_n .

Proposition 2.1. *For all $n = 1, 2, \dots$, let $Z_n = \{Z_n(t), t \geq 0\}$ be defined by:*

$$Z_n(t) = \frac{1}{\sqrt{n}} \left(S_n(t) - n m(t) \right). \quad (2.2)$$

The sequence of processes (Z_n) , $n \in \mathbb{N}$ converges in distribution to a centered Gaussian process Z with covariance function defined for $0 \leq s \leq t$ by:

$$\text{Cov}[Z(s), Z(t)] = {}^t f K(s, t) f,$$

where the covariance matrix $K(s, t)$ is given by (2.1). Moreover, Z has continuous sample paths.

This CLT is simply an application of a result proved by Hahn ([14], Corollary 3, p. 95), using Whitt’s tools ([26], section 1, p. 68 and section 4, p. 78).

Proposition 2.1 shows that the degradation process S should remain at distance $O(\sqrt{n})$ from the deterministic function $n m$. Our assumption is that m is strictly increasing over the interval $[0, \tau]$, where $0 < \tau \leq +\infty$. For $m(0) < \alpha < m(\tau)$, there exists a unique instant $t_\alpha \in [0, \tau]$ such that $m(t_\alpha) = \alpha$. We want to study the hitting time of level $k = k(n)$, where:

$$k(n) = \alpha n + o(\sqrt{n}).$$

The hitting time T_n is defined as:

$$T_n = \inf \{ t \geq 0; S_n(t) \geq k(n) \}.$$

It is intuitively clear that T_n should be close to t_α , with an error of order $1/\sqrt{n}$.

In order to prove Theorem 1.1, we shall need Proposition 2.1, together with a bound on the left tail of the hitting time T_n . Proposition 2.2 below will be used then. It is of independent interest: since the failure time T_n is close to t_α , we consider an instant t before that time and evaluate the probability of remaining below $k(n)$ between 0 and t , i.e. the *reliability* of the system.

Proposition 2.2. *For any real a , define $t_n(a) = \max\{0, t_\alpha + \frac{a}{\sqrt{n}}\}$. There exist two positive constants C_1 and C_2 such that for $n > C_1$ and $a < -C_2$, the following inequality holds:*

$$\mathbb{P}[T_n \leq t_n(a)] \leq \exp \left(-\frac{1}{2} \left(\frac{a \mu(t_\alpha)}{\delta} \right)^2 \right) + 2 \exp \left(\frac{a \sqrt{n} \mu(t_\alpha)}{16 \delta} \right), \quad (2.3)$$

with $\delta = \max\{f(e) - f(e'); e, e' \in E\}$ and $\mu(t_\alpha) = \inf\{m'(t); 0 \leq t \leq t_\alpha\}$.

Before proving (2.3), we shall briefly discuss its sharpness. By Theorem 1.1, T_n is asymptotically Gaussian. Thus for large n ’s and negative a ’s, the left tail probabilities $\mathbb{P}[T_n \leq t_n(a)]$ should be of order:

$$\exp \left(-\frac{1}{2} \frac{a^2}{\sigma^2} \right) = \exp \left(-\frac{1}{2} \frac{a^2 (m'(t_\alpha))^2}{v(t_\alpha)} \right).$$

In the right-hand side of (2.3) the dominant term is the first one, since for a negative, the second one goes to zero exponentially fast with \sqrt{n} . In the first term, $\mu(t_\alpha)$ is smaller than $m'(t_\alpha)$, and δ^2 is an upperbound for $v(t)$. As a uniform inequality, (2.3) has the correct order of magnitude $\exp(-\frac{1}{2}O(a^2))$, with a constant that leaves little room for improvement.

Proof. The left tail probabilities of T_n can be rewritten in terms of S_n :

$$\mathbb{P}[T_n \leq t_n(a)] = \mathbb{P}[\exists t \leq t_n(a), S_n(t) > k(n)].$$

The problem to bound the supremum of a process over an infinite interval arises very often. The technique we use here is classical (see [22], p. 142-146). We first bound the probability over an interval of length $1/\sqrt{n}$. Denote by A the event:

$$A = \text{“}\exists t \in [t_n(a); t_n(a+1)] \ S_n(t) > k(n)\text{”}.$$

Denote by $k_1(n)$ the following degradation level:

$$k_1(n) = n m(t_n(a)) - \frac{a}{2} \sqrt{n} \mu(t_\alpha),$$

where:

$$\mu(t_\alpha) = \inf\{m'(t); 0 \leq t \leq t_\alpha\} > 0.$$

We will condition by the event B defined as:

$$B = \text{“}S_n(t_n(a)) > k_1(n)\text{”}.$$

Let us write $\mathbb{P}[A]$ as:

$$\begin{aligned} \mathbb{P}[A] &= \mathbb{P}[A | B] \mathbb{P}[B] + \mathbb{P}[A | \bar{B}] \mathbb{P}[\bar{B}] \\ &\leq \mathbb{P}[B] + \mathbb{P}[A | \bar{B}]. \end{aligned}$$

We now bound $\mathbb{P}[B]$ and $\mathbb{P}[A | \bar{B}]$ separately. For the first one, we apply Hoeffding's inequality (e.g. [20], p. 58), using the fact that $S_n(t_n(a))$ is a sum of independent bounded random variables:

$$\mathbb{P}[S_n(t_n(a)) > k_1(n)] \leq \exp\left(-\frac{2(k_1(n) - n m(t_n(a)))^2}{n \delta^2}\right),$$

with $\delta = \max\{f(e) - f(e'); e, e' \in E\}$. Hence:

$$\mathbb{P}[B] \leq \exp\left(-\frac{1}{2} \left(\frac{a \mu(t_\alpha)}{\delta}\right)^2\right). \quad (2.4)$$

Consider now $\mathbb{P}[A | \bar{B}]$.

$$\mathbb{P}[A | \bar{B}] = \sum_{\eta \in E^n} \mathbb{P}[A | \bar{B} \text{ and } \tilde{X}(t_n(a)) = \eta] \mathbb{P}[\tilde{X}(t_n(a)) = \eta | \bar{B}].$$

Let us consider a state $\eta = (\eta_i) \in E^n$ satisfying:

$$\sum_{i=1}^n f(\eta_i) \leq k_1(n).$$

We will bound $\mathbb{P}[A | \tilde{X}(t_n) = \eta]$ independently from η , so the same bound will also hold for $\mathbb{P}[A | \bar{B}]$. The process \tilde{X} is Markovian and time homogeneous. We need to bound the probability for that process to pass in an interval of time of length $1/\sqrt{n}$, from a certain state η whose total degradation is lower than $k_1(n)$, to a total degradation larger than $k(n)$. This requires a number of jumps at least equal to $(k(n) - k_1(n))/\delta$. Let us consider a harmonized version (see e.g. [17]), for which the jumps of \tilde{X} occur according to a Poisson process. The intensity λ of that Poisson process is the maximal jump rate out of any state η , that is:

$$\lambda = n \bar{\lambda} = n \max_{e \in E} \sum_{e' \neq e} \lambda_{ee'}.$$

The number of jumps over an interval of length $1/\sqrt{n}$ has Poisson distribution with parameter $\sqrt{n\bar{\lambda}}$. Its tail probabilities can be bounded using Chernov's inequality ([7] p. 55). If π denotes the Poisson distribution with parameter ρ then for $b > \rho$:

$$\pi([b, +\infty]) \leq \left(\frac{b}{\rho}\right)^{-b} e^{b-\rho}. \quad (2.5)$$

A somewhat weaker form will actually be sufficient for our purpose:

$$\pi([b, +\infty]) \leq \exp\left(-\frac{b}{2}\left(1 - \frac{\rho}{b}\right)^2\right). \quad (2.6)$$

To go from (2.5) to (2.6), it suffices to check that for $x > 1$:

$$\log(x) - 1 + \frac{1}{x} > \frac{1}{2}\left(1 - \frac{1}{x}\right)^2.$$

Here $\rho = \sqrt{n\bar{\lambda}}$, and $b = (k(n) - k_1(n))/\delta$. The degradation level $k(n)$ is such that $k(n) = \alpha n + o(\sqrt{n})$. So there exists a constant C such that for $n > C$:

$$k(n) - \alpha n \geq -\frac{1}{2\delta}\sqrt{n}\mu(t_\alpha).$$

In the Chernov bound (2.6) above, the right-hand side is a decreasing function of b , as long as $b > \rho$. For $n > C$, a lower bound for b is as follows:

$$\begin{aligned} b &= \frac{1}{\delta}\left(k(n) - \alpha n + n m(t_\alpha) - n m(t_n(a)) + \frac{a}{2}\sqrt{n}\mu(t_\alpha)\right) \\ &\geq \frac{1}{\delta}\left(n(m(t_\alpha) - m(t_\alpha + \frac{a}{\sqrt{n}})) + \frac{a-1}{2}\sqrt{n}\mu(t_\alpha)\right) \\ &\geq \frac{1}{\delta}\left(-a\sqrt{n}\mu(t_\alpha) + \frac{a-1}{2}\sqrt{n}\mu(t_\alpha)\right) \\ &= -\frac{a+1}{2\delta}\sqrt{n}\mu(t_\alpha). \end{aligned}$$

Finally we get:

$$\mathbb{P}[A | \bar{B}] \leq \exp\left(\frac{(a+1)\sqrt{n}\mu(t_\alpha)}{4\delta}\left(1 + \frac{2\delta\bar{\lambda}}{(a+1)\mu(t_\alpha)}\right)^2\right), \quad (2.7)$$

for $-\frac{a+1}{2\delta}\sqrt{n}\mu(t_\alpha) \geq \sqrt{n\bar{\lambda}}$, i.e. $a \leq -\frac{2\bar{\lambda}\delta}{\mu(t_\alpha)} + 1$.

Now, one has to combine (2.4) and (2.7):

$$\begin{aligned}
& \mathbb{P}[\exists t \leq t_n(a), S_n(t) > k(n)] \\
& \leq \sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \mathbb{P}[\exists t \in [t_n(a-h-1); t_n(a-h)] \quad S_n(t) \geq k(n)] \\
& \leq \sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \exp\left(-\frac{1}{2} \left(\frac{(a-h-1)\mu(t_\alpha)}{\delta}\right)^2\right) \\
& \quad + \sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \exp\left(\frac{(a-h)\sqrt{n}\mu(t_\alpha)}{4\delta} \left(1 - \frac{2\delta\bar{\lambda}}{(a-h)\mu(t_\alpha)}\right)^2\right).
\end{aligned}$$

Let us consider both sums separately.

$$\begin{aligned}
\sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \exp\left(-\frac{1}{2} \left(\frac{(a-h-1)\mu(t_\alpha)}{\delta}\right)^2\right) & \leq \int_{-a}^{+\infty} \exp\left(-\frac{1}{2} \left(\frac{x\mu(t_\alpha)}{\delta}\right)^2\right) dx \\
& \leq \exp\left(-\frac{1}{2} \left(\frac{a\mu(t_\alpha)}{\delta}\right)^2\right).
\end{aligned}$$

Then, for $a \leq \min\{-\frac{2\bar{\lambda}\delta}{\mu(t_\alpha)} + 1, -\frac{4\bar{\lambda}\delta}{\mu(t_\alpha)}\}$:

$$\begin{aligned}
& \sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \exp\left(\frac{(a-h)\sqrt{n}\mu(t_\alpha)}{4\delta} \left(1 + \frac{2\delta\bar{\lambda}}{(a-h)\mu(t_\alpha)}\right)^2\right) \\
& \leq \sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \exp\left(\frac{(a-h)\sqrt{n}\mu(t_\alpha)}{4\delta} \left(1 + \frac{2\delta\bar{\lambda}}{a\mu(t_\alpha)}\right)^2\right) \\
& \leq \sum_{h=0}^{\lfloor t_\alpha \sqrt{n+a} \rfloor} \exp\left(\frac{(a-h)\sqrt{n}\mu(t_\alpha)}{16\delta}\right) \\
& \leq \exp\left(\frac{a\sqrt{n}\mu(t_\alpha)}{16\delta}\right) \left(1 - \exp\left(\frac{-\sqrt{n}\mu(t_\alpha)}{16\delta}\right)\right)^{-1} \\
& \leq 2 \exp\left(\frac{a\sqrt{n}\mu(t_\alpha)}{16\delta}\right),
\end{aligned}$$

for n large enough. □

We are now ready to complete the proof of Theorem 1.1.

Proof. As mentioned earlier, the hitting time T_n is close to t_α , with an error of order $1/\sqrt{n}$. In order to make this more precise, we need to stretch the time around t_α , by a scale factor of \sqrt{n} . Recall that for all $u \in \mathbb{R}$, $t_n(u) = \max\{0, t_\alpha + \frac{u}{\sqrt{n}}\}$. Denote by $Y_n = \{Y_n(u), u \in \mathbb{R}\}$ the process defined on the real line by :

$$Y_n(u) = Z_n(t_n(u)).$$

Since the limit process Z in Proposition 2.1 has continuous sample paths, the convergence of Y_n follows from Lemma 1.18 of Jacod [16]. Thus the sequence (Y_n) converges in distribution to a Gaussian process $Y =$

$\{Y(u), u \in \mathbb{R}\}$, with constant trajectories, where for all $u \in \mathbb{R}$,

$$Y(u) \equiv Z(t_\alpha).$$

Our goal is to compute the limit of the following probability.

$$\begin{aligned} \mathbb{P}[\sqrt{n}(T_n - t_\alpha) \leq u] &= \mathbb{P}[\exists b \leq u, S_n(t_n(b)) > k(n)] \\ &= \mathbb{P}[\exists b \leq u, Y_n(b) > \Delta_n(b)], \end{aligned}$$

with:

$$\Delta_n(b) = \frac{1}{\sqrt{n}} \left(k(n) - n m(t_n(b)) \right).$$

As n tends to infinity, $\Delta_n(b)$ tends to $\Delta(b) = -b m'(t_\alpha)$. So the process $\{X_n(b) - \Delta_n(b)\}$ converges in distribution to $\{X(b) - \Delta(b)\}$, by Lemma 1.17 of [16]. A consequence is that for all $a < u$:

$$\mathbb{P}[\forall b, a < b < u; Y_n(b) < \Delta_n(b)] \xrightarrow[n \rightarrow +\infty]{} \mathbb{P}[\forall b, a < b < u; Y(b) < \Delta(b)].$$

For $a = -\infty$ the above limit does not follow from the convergence of processes, and this is precisely where Proposition 2.2 is used.

Because $\Delta(b)$ is a decreasing function and since the process Y has constant trajectories, we have:

$$\begin{aligned} \mathbb{P}[\forall b, a < b < u; Y(b) \leq \Delta(b)] &= \mathbb{P}[Y(u) \leq \Delta(u)] \\ &= \Phi \left(-\frac{u m'(t_\alpha)}{\sqrt{v(t_\alpha)}} \right) \\ &= 1 - \Phi \left(\frac{u m'(t_\alpha)}{\sqrt{v(t_\alpha)}} \right), \end{aligned}$$

where Φ is the Gaussian distribution function.

Finally:

$$\begin{aligned} \mathbb{P}[\exists b, a < b < u; Y_n(b) > \Delta_n(b)] \\ &\leq \mathbb{P}[T_n \leq t_n(u)] \\ &\leq \mathbb{P}[\exists b, b \leq a; Y_n(b) > \Delta_n(b)] + \mathbb{P}[\exists b, a < b < u; Y_n(b) > \Delta_n(b)]. \end{aligned}$$

The first probability tends to $\Phi \left(\frac{u m'(t_\alpha)}{\sqrt{v(t_\alpha)}} \right)$. By Proposition 2.2, the first term in the last member tends to 0 as n tends to $+\infty$ and a tends to $-\infty$. The result follows. \square

As was pointed out by the referee, there exist other ways of proving Theorem 1.1 that could be more elegant and easier to extend. Indeed, he suggested to use the Skorokhod-Dudley-Wichura representation Theorem ([22] p.71), to construct a version of the limit process Z , to which Z_n converges almost surely. This somewhat simplifies the proof of tightness for the sequence $(\sqrt{n}(T_n - t_\alpha))$. As already mentioned, we believe that Proposition 2.2 is of independent interest, as a uniform bound for the reliability of a large system.

3. RELIABILITY MODELS

As explained in [21], the notion of redundancy is central to many problems in reliability. This justifies the introduction of models made of a large number of identical components. The most classical case is that of binary components: each one can be either working or failed, and the function f is 0 if the component is working, and 1 otherwise. All components are supposed to be working at time $t = 0$. The total degradation of the system is simply the number of failed components, and the breakdown occurs when the number of failed components passes a certain threshold k . These systems are called “ k -out-of- n ” (cf. for instance Barlow and Proschan [4] or Coccozza-Thivent [11]). The application of our results to k -out-of- n systems of binary components is presented in the first subsection. The second one will study the generalization to multi-state components.

3.1. Binary components

When $f(E) = \{0, 1\}$, the degradation at time t is a Bernoulli random variable. Its expectation $m(t)$ is the probability for one component to be failed at time t , and its variance $v(t)$ is $m(t)(1 - m(t))$. Recall that the instant at which the mean degradation m first passes level α is denoted by t_α . The threshold $k(n)$ is taken to be $\alpha n + o(\sqrt{n})$. Theorem 1.1 states that the asymptotic distribution of $\sqrt{n}(T_n - t_\alpha)$ is Gaussian, with variance:

$$\sigma^2 = \frac{v(t_\alpha)}{m'(t_\alpha)^2}.$$

In the particular case of binary components, $v(t_\alpha) = \alpha(1 - \alpha)$. So the knowledge of $m(t)$ suffices to determine the asymptotic distribution of the failure time T_n . In reliability theory m is called the “non-availability” of one component.

The simplest expressions are obtained when the failure and repair times of each component are exponentially distributed, with respective rates λ and μ . Easy computations (see for instance [5], section 7.5, p. 210-226 and section 15.6, p. 471) lead to:

$$m(t) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t}.$$

This function is strictly increasing over $[0, +\infty]$. Thus one can choose $\tau = +\infty$, and Theorem 1.1 can be applied for any α such that $0 < \alpha < \frac{\lambda}{\lambda + \mu}$. We get:

$$t_\alpha = -\frac{1}{\lambda + \mu} \log \left(1 - \frac{\alpha(\lambda + \mu)}{\lambda} \right).$$

Thus the asymptotic variance σ^2 is:

$$\sigma^2 = \frac{\alpha(1 - \alpha)}{(\lambda - \alpha(\lambda + \mu))^2}.$$

Consider now binary components having non-exponential repair or failure times. Since they are not Markovian, our results do not apply directly. However, any distribution can be approximated by a phase-type distribution. There exists an extensive literature on this subject, for which important references are Neuts' [18, 19] and Tijms' [25] books. These techniques coming from queuing theory are often called matrix-analytic methods (see for instance Asmussen's book [1] and his recent survey on the subject [2]). Phase-type distributions are laws of sojourn times in continuous time Markov chains. Their set is dense in the set of distributions with values on \mathbb{R}^+ ([18], p. 243).

Assume one has to study binary components for which the failure and repair times have phase-type distributions. For one component, one can replace each one of the two states by a set of “fictitious” states, designed so as to make the sojourn time in all fictitious working (resp.: failed) states distributed as the prescribed failure

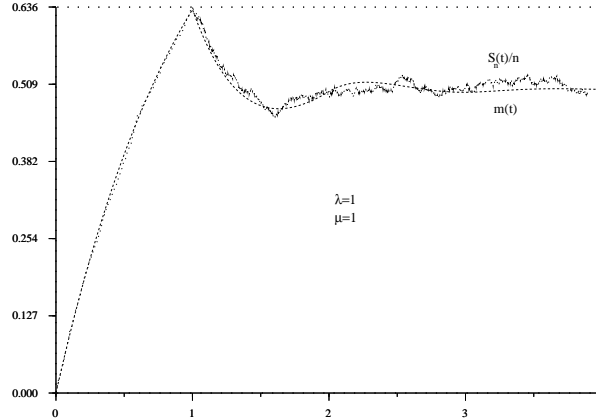


FIGURE 1. Non-availability and simulation of a system with deterministic repair times

(resp.: repair) time. Of course the value of the degradation function f will be 0 (resp.: 1) on all fictitious working (resp.: failed) states.

Replacing the initial distributions by phase-type approximations before applying Theorem 1.1, supposes that the law of the failure time T_n should not change by much. We shall not attempt to rigorously prove that intuition here. The referee suggested to use functional CLT's which are uniform with respect to the distribution of variables (e.g. [24]). Proposition 2.1 could easily be extended to general failure and repair distributions. However we do not see how to generalize our reliability bound (Proposition 2.2), which specifically relies on Markovian techniques.

We will only show that the intuition above can be supported by simulation experiments on a particular case. Suppose that the repair time of a component is deterministic, equal to $1/\mu$, whereas its failure time is exponential with parameter λ as before. The behavior of such a binary component is a so called "alternate renewal process". It can be checked (see e.g. [15], p. 311-313) that the probability of failure at time t has the following expression:

$$m(t) = 1 - \sum_{j=0}^{+\infty} \frac{\lambda^j}{j!} \left(t - \frac{j}{\mu}\right)^j e^{-\lambda(t - \frac{j}{\mu})} \mathbb{1}_{[\frac{j}{\mu}, +\infty)}(t). \quad (3.1)$$

Figure 1 represents $m(t)$ together with one trajectory of the process S_n/n , for a system with $n = 1000$ components, the failure and repair rates being $\lambda = 1$ and $\mu = 1$. Notice that m is strictly increasing over the interval $[0; \mu]$, then decreases and oscillates toward its asymptotic value $\frac{\lambda}{\lambda + \mu}$.

To approximate the Dirac mass at $1/\mu$ by a phase-type distribution, it suffices to consider an Erlang distribution with parameters r and $r\mu$, for r large enough. So we will consider a Markov chain X on the state space

$E = \{0, \dots, r\}$, having the following infinitesimal generator:

$$\Lambda_r = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -r\mu & r\mu & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & -r\mu & r\mu \\ r\mu & 0 & \dots & 0 & -r\mu \end{pmatrix}.$$

State 0 is the perfect state and states $1, \dots, r$ represent fictitious failed states. The Laplace transform of the marginal distribution $p(t)$ is easily deduced from the Chapman-Kolmogorov system. Denoting by \mathcal{L}_m the Laplace transform of the mean degradation m , one gets:

$$\mathcal{L}_m(s) = \frac{1}{s} - \frac{1}{s + \lambda - \lambda \left(\frac{r\mu}{s+r\mu} \right)^r}. \quad (3.2)$$

The Laplace transform of m' is immediately deduced from \mathcal{L}_m . However m and m' do not have tractable explicit expressions, and must be computed numerically.

As an example, we chose again $\lambda = \mu = 1$. For $r = 10, 20, \dots, 100$, we computed the maximal difference between the exact value of m (formula (3.1)), and its phase-type approximation, by numerically solving the Chapman-Kolmogorov system. The results are the following.

r	10	20	30	40	50	60	70	80	90	100
Max. diff.	0.091	0.069	0.059	0.053	0.048	0.044	0.041	0.039	0.037	0.035

TABLE 1. Maximal difference between the exact value of m and a phase-type approximation, for a binary system with deterministic repair times.

However, the maximal difference is located at the sharp peak of $m(t)$ and the approximation is notably better for smaller values of t . From then on, we fixed $r = 100$ and $n = 500$, and considered the levels $k(n) = \lfloor \alpha n \rfloor$, for $\alpha = 0.05, 0.1, \dots, 0.5$. For those α 's, the exact values of t_α agreed with those of the phase-type approximation up to the fourth decimal. For each value of α , the mean hitting time t_α and the variance σ^2 was numerically computed for the phase-type approximation. Then we simulated 100 independent trajectories of the system with deterministic repair times, and recorded for each one the hitting time T_n of the different levels $k(n)$. For each value of α , the empirical distribution of T_n was compared with the asymptotic normal distribution given in Theorem 1.1, by computing the p-value of the Kolmogorov-Smirnov test. The results are given in Table 2. With any usual risk, the goodness-of-fit hypothesis is accepted in all cases. Notice that since hitting times were computed along the same trajectories, their values for different α 's are not independent.

Let us now consider binary, but non repairable components. Let G denote the distribution function of the failure times, and g their density. In order to apply Theorem 1.1, one can assume that this distribution is phase-type. The degradation function is the probability of failure at time t , i.e. $m(t) = G(t)$. It is strictly increasing, from 0 to 1. For each α such that $0 < \alpha < 1$, the mean hitting time t_α is the α -th quantile of the failure time: $t_\alpha = G^{-1}(\alpha)$. The asymptotic variance is now:

$$\sigma^2 = \frac{\alpha(1-\alpha)}{g(G^{-1}(\alpha))^2}.$$

α	t_α	σ^2	K.-S. stat.	p -value
0.05	0.051	0.053	0.920	0.366
0.10	0.105	0.111	1.004	0.266
0.15	0.163	0.176	0.829	0.497
0.20	0.223	0.25	0.827	0.501
0.25	0.288	0.333	0.773	0.589
0.30	0.357	0.428	0.792	0.557
0.35	0.431	0.538	1.003	0.267
0.40	0.511	0.666	0.681	0.743
0.45	0.598	0.818	0.957	0.318
0.50	0.693	1.	0.703	0.706

TABLE 2. Comparison of empirical distributions of hitting times, with asymptotic distributions on a phase type approximation.

But if the components are not repairable, the time at which the k -th of them fails is the k -th order statistic of a sample of i.i.d. failure times, with distribution function G . So in this particular case, our result amounts to the well known central limit Theorem for order statistics (e.g. Reiss [23] section 4.1, p. 108-109).

3.2. Markovian multi-state components

Our results also apply to multi-state components. The different states of one component correspond to performance levels and describe its progressive degradation. As an example, we will consider an extension of a model studied by Pham et al. in [21] where the cardinality of the state space is 3. The three states are 0 (perfect), 1 (deteriorated) and 2 (failed). A perfect component can either deteriorate or directly fail. A deteriorated component can either fail or be repaired. A failed component is restored to perfect state. The transition diagram is that of figure 2. The infinitesimal generator is the following.

$$\Lambda = \begin{pmatrix} -(\lambda_1 + \lambda_2) & \lambda_1 & \lambda_2 \\ \mu_1 & -(\mu_1 + \lambda_3) & \lambda_3 \\ \mu_2 & 0 & -\mu_2 \end{pmatrix}.$$

The matrix $\exp(\Lambda t)$ can be explicitly computed, but the expressions are somewhat lengthy. Simpler formulas are obtained for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. For the probabilities of the different states starting from 0 at time $t = 0$, one gets:

$$\begin{cases} p_0(t) = \frac{\mu_2(\lambda + \mu_1)}{(2\lambda + \mu_1)(\lambda + \mu_2)} + \frac{\lambda(2\lambda + \mu_1 - \mu_2)}{(2\lambda + \mu_1)(\lambda + \mu_1 - \mu_2)} e^{-(2\lambda + \mu_1)t} + \frac{\lambda(\mu_1 - \mu_2)}{(\lambda + \mu_2)(\lambda + \mu_1 - \mu_2)} e^{-(\lambda + \mu_2)t} \\ p_1(t) = \frac{\lambda\mu_2}{(2\lambda + \mu_1)(\lambda + \mu_2)} - \frac{\lambda(2\lambda + \mu_1 - \mu_2)}{(2\lambda + \mu_1)(\lambda + \mu_1 - \mu_2)} e^{-(2\lambda + \mu_1)t} + \frac{\lambda^2}{(\lambda + \mu_2)(\lambda + \mu_1 - \mu_2)} e^{-(\lambda + \mu_2)t} \\ p_2(t) = \frac{\lambda}{\lambda + \mu_2} - \frac{\lambda}{\lambda + \mu_2} e^{-(\lambda + \mu_2)t}. \end{cases}$$

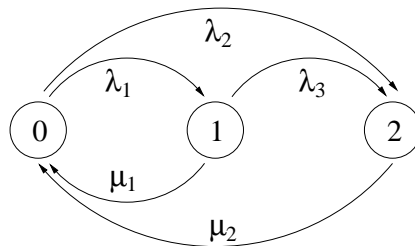


FIGURE 2. Transition diagram of a three-state component

In Pham et al. [21], a failed component is not repaired ($\mu_2 = 0$), so state 2 is absorbing. Their system is supposed to be failed as soon as k out of n components are in state 2. With our notations, this corresponds to $f(0) = f(1) = 0$, and $f(2) = 1$. The mean degradation is $m(t) = p_2(t)$. The mean hitting time t_α and the asymptotic variance σ^2 can also be explicitly computed. The expressions are similar to those of the binary case. For $\alpha < \frac{\lambda}{\lambda + \mu_2}$:

$$t_\alpha = -\frac{1}{\lambda + \mu_2} \log \left(1 - \frac{\alpha(\lambda + \mu_2)}{\lambda} \right),$$

and:

$$\sigma^2 = \frac{\alpha(1 - \alpha)}{(\lambda - \alpha(\lambda + \mu_2))^2}.$$

Consider now a variant of this model which takes into account the difference between a perfect and a degraded component. We can choose function f as follows: $f(0) = 0$, $f(1) = \frac{1}{2}$, and $f(2) = 1$. The mean degradation becomes $m(t) = \frac{1}{2}p_1(t) + p_2(t)$ which may no longer be increasing for all t , depending on the value of μ_2 . For suitable values of α , t_α can be computed at least numerically, and Theorem 1.1 can be applied.

We are pleased to thank J. Jacod and J. Rousseau for their hints. We are also indebted to the anonymous referee, whose thorough report was very helpful.

REFERENCES

- [1] S. Asmussen. *Applied probability and queues*. Wiley, New-York, 1987.
- [2] S. Asmussen. Matrix-analytic models and their analysis. *Scand. J. Statist.*, 27:193–226, 2000.
- [3] T. Aven and U. Jensen. *Stochastic models in reliability*. Springer, New-York, 1999.
- [4] R.E. Barlow and F. Proschan. *Mathematical theory of reliability*. SIAM, Philadelphia, 1996.
- [5] U.N. Bhat. *Elements of applied stochastic processes*. Wiley, New-York, 1984.
- [6] J.L. Bon. *Méthodes mathématiques de la fiabilité*. Masson, Paris, 1995.
- [7] N. Bouleau and D. Lépingle. *Numerical methods for stochastic processes*. Wiley, New York, 1994.
- [8] M.V. Boutsikas and M.V. Kontras. Generalized reliability bounds for coherent structures. *J. Appl. Probab.*, 37(3):778–794, 2000.
- [9] J.A. Carrasco. Computationally Efficient and Numerically Stable Reliability Bounds for Repairable Fault-Tolerant Systems. *IEEE Trans. on Computers*, 51(3):254–268, 2002.
- [10] D. Chauveau and J. Diébolt. An automated stopping rule for MCMC convergence assessment. *Comput. Statist.*, 14(3):419–442, 1999.
- [11] C. Coccozza-Thivent. *Processus stochastiques et fiabilité des systèmes*. Springer, Paris, 1997.
- [12] S.N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. Wiley, New-York, 1986.
- [13] R. Gupta and A. Chaudhary. Analysis of reliability bounds for series, parallel and k -out-of- n system configuration. *Microelectron. Reliab*, 34(1):183–185, 1994.
- [14] M.G. Hahn. Central limit theorem in $D[0, 1]$. *Z. Wahrsch. Verw. Geb*, 44:89–101, 1978.
- [15] A. Hølyand and M. Rausand. *System reliability theory: models and statistical methods*. Wiley, New-York, 1994.

- [16] J. Jacod. Théorèmes limite pour les processus. In P.L. Hennequin, editor, *École d'Été de Probabilités de Saint-Flour XIII (1983)*, L.N. in Math. 1117, pages 299–409. Springer, Berlin, 1985.
- [17] J. Keilson. *Markov chain models - rarity and exponentiality*. Springer, New-York, 1979.
- [18] M.F. Neuts. *Structured stochastic matrices of M/G/1 type and their applications*. Dekker, New-York, 1989.
- [19] M.F. Neuts. *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Dover, New-York, 1994.
- [20] V.V. Petrov. *Sums of independent random variables*. Springer, Berlin, 1975.
- [21] H. Pham, A. Suprasad, and R.B. Misra. Reliability analysis of k -out-of- n systems with partially repairable multi-state components. *Microelectron. Reliab*, 36(10):1407–1415, 1996.
- [22] D. Pollard. *Convergence of stochastic processes*. Springer, New-York, 1984.
- [23] R.-D. Reiss. *Approximate distributions of order statistics, with application to non-parametric statistics*. Springer, New-York, 1989.
- [24] A. Sheehy and J.A. Wellner. Uniform Donsker classes of functions. *Ann. Probab.*, 20(4):1983-2030, 1992.
- [25] H.C. Tijms. *Stochastic models: an algorithmic approach*. Wiley, Chichester, 1994.
- [26] W. Whitt. Some useful functions for functional limit theorems. *Math. Oper. Res.*, 5(1):67–85, 1980.
- [27] B. Ycart. Cutoff for samples of Markov chains. *ESAIM Probability-Statistics* 3:89–107, 1999.
- [28] B. Ycart. Stopping tests for Monte-Carlo Markov chain methods. *Meth. and Comp. in Applied Probab.*, 2(1):23-36, 2000.
- [29] B. Ycart. Cutoff for Markov chains: some examples and applications. In E. Goles and S. Martínez, editors, *Complex Systems*, pages 261–300. Kluwer, Dordrecht, 2001.