J. Appl. Prob. 27, 744–756 (1989)
Printed in Israel

SOJOURN TIMES IN FINITE MARKOV PROCESSES

GERARDO RUBINO AND
BRUNO SERICOLA,* *IRISA*

Abstract

Sojourn times of Markov processes in subsets of the finite state space are considered. We give a closed form of the distribution of the nth sojourn time in a given subset of states. The asymptotic behaviour of this distribution when time goes to infinity is analyzed, in the discrete time and the continuous-time cases. We consider the usually pseudo-aggregated Markov process canonically constructed from the previous one by collapsing the states of each subset of a given partition. The relation between limits of moments of the sojourn time distributions in the original Markov process and the moments of the corresponding holding times of the pseudo-aggregated one is also studied.

AGGREGATION; TRANSIENT ANALYSIS; PERFORMANCE EVALUATION; RELIABILITY MODELING

1. Introduction

Consider a homogeneous irreducible Markov process X evolving in discrete or continuous time; in the first case we speak of a Markov chain. We can distinguish a subset of states and then consider the random variable 'time spent by process X in the given subset of the state space in its nth visit to the subset'. This kind of subject is of interest in areas such as reliability or performability [2], when a subset of the state space corresponds, for instance, to a fixed level of performance of the system [6], [1]. Since the processes considered here are irreducible, in such a context this means that we are concerned with systems that always have recovery procedures, that is, models without absorbing classes. The analysis of these random variables is the main topic of this paper.

The *sojourn* time of a process X in a subset of states will be an integer-valued random variable if X is a chain or a real-valued one in the case of a continuous-time process. The distributions of these random variables are given in Sections 2 and 3 for the discrete-time and continuous-time cases respectively. We also study the behaviour of these distributions when time goes to infinity, that is, for $n \to +\infty$ when considering the nth visit to the the given subset.

Let $E = \{1, \cdots, N\}$ be the finite state space of the given process and let $\mathscr{B} = \{B(1), \cdots, B(M)\}$ be a partition of E. Let F be the set of integers $\{1, 2, \cdots, M\}$. To

Received 15 April 1988; revision received 4 October 1988.

* Postal address for both authors: IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France.

the given process X we associate the aggregated stochastic process Y with values on F, defined by: $Y_t = m \Leftrightarrow X_t \in B(m)$ for all values of t ($t \in \mathbb{N}$ or $t \in \mathbb{R}$). We easily deduce from this definition and the irreducibility of X that the process Y obtained is also irreducible but it need not be Markov, not even homogeneous. In [3] and in [4], we can find conditions under which the aggregated chain Y (i.e. in the discrete-time case) is also Markov homogeneous. Anyway, we can construct a homogeneous Markov process Z on the state space F canonically associated to the given process X which will be called the *pseudo-aggregated* process of X with respect to the partition $\mathscr{B}$. This pseudo-aggregation is briefly presented in Section 4 and the relation between the holding times of Z and the corresponding sojourn times of X is analyzed. Section 5 presents an application to a fault-tolerant multiprocessor system [6] and the last section contains some conclusions.

2. Sojourn times in the discrete-time case

Let $X = (X_n)_{n \geq 0}$ be a homogeneous irreducible Markov chain with transition probability matrix P. Denote by α its initial probability distribution and by x its equilibrium probability distribution, that is: $x = xP$, $x > 0$ and $x1^T = 1$, where 1 denotes a row vector with all the entries equal to the scalar 1, the dimension being defined by the context. Let B be a proper subset of the state space E and denote by B^c the complementary subset $E - B$. We assume for simplicity that $B = \{1, 2, \cdots, L\}$, $1 \leqq L < N$. The partition $\{B, B^c\}$ of E induces a decomposition of P into four submatrices and a decomposition of α and x into two subvectors:

$$P = \begin{pmatrix} P_B & P_{BB^c} \\ P_{B^cB} & P_{B^c} \end{pmatrix} \quad \alpha = (\alpha_B \quad \alpha_{B^c}) \quad x = (x_B \quad x_{B^c}).$$

We shall need the two following elementary lemmas.

Lemma 2.1. *The matrix $I - P_B$ is invertible.*

Proof. Consider a Markov chain on the state space $B \cup \{a\}$ with transition probability matrix

$$P' = \begin{pmatrix} P_B & (I - P_B)1^T \\ 0 & 1 \end{pmatrix}.$$

P being irreducible, the states of B are transient and a is absorbing. Therefore

$$\lim_{k \to +\infty} P_B^k(i,j) = 0, \qquad \forall i,j \in B,$$

which implies that $I - P_B$ is invertible.

In the same way, the matrix $I - P_{B^c}$ is invertible.

Lemma 2.2. *The vector x_B satisfies $x_B = x_B U_B$ where*

$$U_B = P_{BB^c}(I - P_{B^c})^{-1} P_{B^cB}(I - P_B)^{-1}.$$

Proof. The result follows directly from the following decomposition of the system $x = xP$:

$$x_B = x_B P_B + x_{B^c} P_{B^cB}$$

$$x_{B^c} = x_B P_{BB^c} + x_{B^c} P_{B^c}$$

by replacing in the first equation the value of the vector x_{B^c} obtained from the second one.

Remark that in the same way, we have

$$x_{B^c} = x_{B^c} U_{B^c} \quad \text{where } U_{B^c} = P_{B^cB}(I - P_B)^{-1} P_{BB^c}(I - P_{B^c})^{-1}.$$

Definition 2.3. *We call 'sojourn of X in B' any sequence $X_m, X_{m+1}, \cdots, X_{m+k}$ where $k \geqq 1$, $X_m, X_{m+1}, \cdots, X_{m+k-1} \in B$, $X_{m+k} \notin B$ and if $m > 0$, $X_{m-1} \notin B$. This sojourn begins at time m and finishes at time $m + k$. It lasts k.*

Let V_n, $n \geqq 1$ be the random variable 'state of B in which the nth sojourn of X begins'. The hypothesis of irreducibility of the Markov chain X assures the existence of an infinity of sojourns of X in B with probability 1. It is immediate to verify that $(V_n)_{n \geqq 1}$ is a homogeneous Markov chain on the state space B. Let G be the $L \times L$ transition probability matrix of this chain and v_n its probability distribution vector after the nth transition: $v_n = (\mathbb{P}(V_n = 1), \cdots, \mathbb{P}(V_n = L))$ (we denote by $\mathbb{P}$ the probability measure). We have obviously $v_n = v_1 G^{n-1}$. $(V_n)_{n \geqq 1}$ is characterized by v_1 and G which are given in the following theorem.

Theorem 2.4. *Matrix G and vector v_1 are given by the following expressions*:
(i) $v_1 = \alpha_B + \alpha_{B^c}(I - P_{B^c})^{-1} P_{B^cB}$;
(ii) $G = (I - P_B)^{-1} P_{BB^c}(I - P_{B^c})^{-1} P_{B^cB} = (I - P_B)^{-1} U_B (I - P_B)$.

Proof. (i) Let $i \in B^c$ and $j \in B$. We define $H(i,j) \stackrel{\text{def}}{=} \mathbb{P}(V_1 = j \mid X_0 = i)$ and let H be the $(N - L) \times L$ matrix with entries $H(i,j)$. The matrix H is the matrix of absorption probabilities from B^c to B, that is (see [3], p. 53):

$$H = (I - P_{B^c})^{-1} P_{B^cB}.$$

Conditioning on the initial state, we then obtain, for every $j \in B$:

$$\mathbb{P}(V_1 = j) = \alpha(j) + \sum_{i \in B^c} \mathbb{P}(X_0 = i) H(i,j).$$

This gives, in matrix notation,

$$v_1 = \alpha_B + \alpha_{B^c}(I - P_{B^c})^{-1} P_{B^cB}.$$

(ii) Let V_1^c be the random variable 'state of B^c in which the first sojourn of X in B^c begins'. In a similar way as previously, for every $i \in B$ and $j \in B^c$, we define: $R(i,j) \stackrel{\text{def}}{=} \mathbb{P}(V_1^c = j \mid X_0 = i)$. The $L \times (N - L)$ matrix R is the matrix of absorption probabilities from B to B^c, that is:

$$R = (I - P_B)^{-1} P_{BB^c}.$$

So, for every $i, j \in B$ and using the Markov properties of X, we have

$$\begin{aligned} G(i,j) &= \mathbb{P}(V_2 = j \mid V_1 = i) = \mathbb{P}(V_2 = j \mid X_0 = i) \\ &= \sum_{k \in B^c} \mathbb{P}(V_1^c = k \mid X_0 = i)\mathbb{P}(V_2 = j \mid V_1^c = k, X_0 = i) \\ &= \sum_{k \in B^c} \mathbb{P}(V_1^c = k \mid X_0 = i)\mathbb{P}(V_1 = j \mid X_0 = k) \\ &= \sum_{k \in B^c} R(i,k)H(k,j). \end{aligned}$$

This gives, in matrix notation, $G = RH$.

The chain $(V_n)_{n \geqq 1}$ contains only one recurrent class, the set B' of the states of B directly accessible from B^c: $B' = \{j \in B \mid \exists i \in B^c,\ P(i,j) > 0\}$. Without any loss of generality, we will consider that $B' = \{1, \cdots, L'\}$ where $1 \leqq L' \leqq L$. We then denote by B'' the set $B - B'$. The partition $\{B', B''\}$ induces the following decomposition on matrices G and H:

$$G = \begin{pmatrix} G' & 0 \\ G'' & 0 \end{pmatrix} \qquad H = (H' \quad 0).$$

In the same way, the partition $\{B', B'', B^c\}$ induces the following decomposition on P:

$$P = \begin{pmatrix} P_{B'} & P_{B'B''} & P_{B'B^c} \\ P_{B''B'} & P_{B''} & P_{B''B^c} \\ P_{B^cB'} & 0 & P_{B^c} \end{pmatrix}.$$

Now, as for G, we have the following expression for G'.

Theorem 2.5.

$$G' = (I - P_{B'} - P_{B'B''}(I - P_{B''})^{-1}P_{B''B'})^{-1}(P_{B'B^c} + P_{B'B''}(I - P_{B''})^{-1}P_{B''B^c})(I - P_{B^c})^{-1}P_{B^cB'}.$$

Proof. The proof is as in (i) of Theorem 2.4. Therefore, we omit it. We give only the system satisfied by matrices G', G'' and H':

$$\begin{aligned} G' &= P_{B'}G' + P_{B'B''}G'' + P_{B'B^c}H' \\ G'' &= P_{B''B'}G' + P_{B''}G'' + P_{B''B^c}H' \\ H' &= P_{B^cB'}G' + P_{B^c}H'. \end{aligned}$$

By inferring G' from this system we obtain the given expression.

Note that the expression given in the previous theorem can considerably reduce the time necessary to compute G' when $L' < L$: instead of inverting a matrix of size L when computing G from the formula of Theorem 2.4, we have here to invert two matrices of sizes L' and $L - L'$.

We now denote by $N_{B,n}$ for $n \geqq 1$ the random variable taking values in $\mathbb{N}^*$: 'time spent by X during its nth sojourn in B'. We have the following explicit expression for the distribution of $N_{B,n}$.

Theorem 2.6.

$$\forall n \in \mathbb{N}^*, \quad \forall k \in \mathbb{N}^*, \quad \mathbb{P}(N_{B,n} = k) = v_n P_B^{k-1}(I - P_B)\, 1^T.$$

Proof. First, we derive the distribution of $N_{B,1}$. Conditioning on the state in which the sojourn in B begins, we obtain

$$\forall i \in B, \quad \mathbb{P}(N_{B,1} = 1 \mid V_1 = i) = \sum_{j \in B^c} P(i,j).$$

Let $k > 1$ and $i \in B$.

$$\mathbb{P}(N_{B,1} = k \mid V_1 = i) = \sum_{j \in B} P(i,j)\mathbb{P}(N_{B,1} = k - 1 \mid V_1 = j).$$

If we define $h_k \overset{\text{def}}{=} (\mathbb{P}(N_{B,1} = k \mid V_1 = 1), \cdots, \mathbb{P}(N_{B,1} = k \mid V_1 = L))$, we can rewrite these two relations as follows:

$$h_1^T = P_{BB^c}\, 1^T = (I - P_B)\, 1^T$$

and for $k > 1$:

$$h_k^T = P_B h_{k-1}^T$$

that is

$$h_k^T = P_B^{k-1}(I - P_B)\, 1^T, \qquad \forall k \geqq 1$$

and

$$\mathbb{P}(N_{B,1} = k) = v_1 P_B^{k-1}(I - P_B)\, 1^T, \qquad \forall k \geqq 1.$$

If we consider now the nth sojourn of X in B, we have

$$\begin{aligned}\mathbb{P}(N_{B,n} = k) &= \sum_{i \in B} \mathbb{P}(N_{B,n} = k \mid V_n = i)\mathbb{P}(V_n = i) \\ &= \sum_{i \in B} \mathbb{P}(N_{B,1} = k \mid V_1 = i)\mathbb{P}(V_n = i) \\ &= v_n h_k^T\end{aligned}$$

which is the explicit expression given in the statement.

Let us compute now the moments of the random variable $N_{B,n}$. An elementary matrix calculus gives

$$\mathbb{E}(N_{B,n}) = v_n(I - P_B)^{-1} 1^T$$

where $\mathbb{E}$ denotes expectation. For higher order moments, it is more comfortable to work with factorial moments instead of standard moments. Recall that the kth-order factorial moment of a random variable V, which will be denoted by $FM_k(V)$, is defined by:

$$FM_k(V) \stackrel{\text{def}}{=} \mathbb{E}(V(V-1)\cdots(V-k+1)).$$

See that $FM_1(V) = \mathbb{E}(V)$. The following property holds: if $\mathbb{E}(V^j) = \mathbb{E}(W^j)$ for $j = 1, 2, \cdots, k-1$ then

$$[FM_k(V) = FM_k(W) \Leftrightarrow \mathbb{E}(V^k) = \mathbb{E}(W^k)].$$

Then, by a classical matrix computation like that for the mean value of $N_{B,n}$, we have

$$(1) \qquad FM_k(N_{B,n}) = k!\, v_n P_B^{k-1}(I - P_B)^{-k} 1^T.$$

Let us now consider the asymptotic behaviour of the nth sojourn time of X in B.

Corollary 2.7. *For any $k \geqq 1$, the sequence $(\mathbb{P}(N_{B,n} = k))_{n \geqq 1}$ converges in the sense of Cesaro as $n \to \infty$ and the limit is $vP_B^{k-1}(I - P_B) 1^T$ where v is the equilibrium probability distribution of $(V_n)_{n \geqq 1}$. Vector v is given by: $v = (1/K)x_B(I - P_B)$ where K is the normalizing constant $x_B(I - P_B) 1^T$. The convergence is simple for any initial distribution of X if and only if G' is aperiodic.*

Proof. The proof is an elementary consequence of general properties of Markov chains, since the sequence depends on n only through v_n and $(V_n)_{n \geqq 1}$ is a finite homogeneous Markov chain with only one recurrent class B'. See that $v = (v' \ \ 0)$ according to the partition $\{B', B''\}$ of B. If we consider the block decomposition of G with respect to the partition $\{B', B''\}$, a simple recurrence on integer n allows us to write

$$G^n = \begin{pmatrix} G'^n & 0 \\ G''G'^{n-1} & 0 \end{pmatrix}.$$

G'^n converges in the sense of Cesaro to $1^T v'$ and in the same way, $G''G'^{n-1}$ converges in the sense of Cesaro to $1^T v'$ because $G'' 1^T = 1^T$ (recall that the dimension of vector 1 is given by the context). The expression for v is easily checked by using Lemma 2.2. The second part of the proof is an immediate corollary of the convergence properties of Markov chains.

Let us define the random variable $N_{B,\infty}$ with values in $\mathbb{N}^*$ by making its distribution equal to the previous limit:

$$\mathbb{P}(N_{B,\infty} = k) = vP_B^{k-1}(I - P_B) 1^T \quad \text{for any } k \geqq 1.$$

Then, by taking Cesaro limits in expression (1) giving the factorial moments of $N_{B,n}$, we obtain the limit given below which needs no proof.

Corollary 2.8. *The sequence of factorial moments $(FM_k(N_{B,n}))_{n \geqq 1}$ converges in the sense of Cesaro and we have*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{l=1}^{n} FM_k(N_{B,l}) = FM_k(N_{B,\infty}) = k!\, vP_B^{k-1}(I - P_B)^{-k} 1^T.$$

The convergence is simple for any initial distribution of X if and only if the matrix G' is aperiodic.

Remark. There is no relation between the periodicity of P and the periodicity of G'. That is, the four situations obtained by combination of the two properties *periodicity* and *aperiodicity* of each matrix are possible. It is immediate to construct four examples illustrating this.

3. Sojourn times in the continuous-time case

Now let $X=(X_t)_{t\geqq 0}$ be an irreducible homogeneous Markov process on the state space $E=\{1,\cdots,N\}$. Let A be the infinitesimal generator of this process, where $A(i,i)\stackrel{\text{def}}{=}-\Sigma_{j\neq i}A(i,j)$ and define $\lambda(i)\stackrel{\text{def}}{=}-A(i,i)\ \forall i\in E$. That is, $\lambda(i)$ represents the output rate of state i. Let P be the transition probability matrix of the uniformized chain [5] of X. We then have the following relation between the matrices A and P:

$$P=I+A/\lambda$$

where λ is the rate of a Poisson process $\{N(t),t\geqq 0\}$ independent of the uniformized chain and satisfying $\lambda\geqq\max(\lambda(i),i\in E)$. We will denote by α and x respectively the initial and the stationary distributions of the process X.

Remark that x is also the stationary distribution of the uniformized chain of X $(x=xP)$. As in the discrete-time case, we consider a subset B of E and we conserve the notations for B, B', B'', for the decomposition of P and the decomposition of A with respect to the partitions $\{B,B^c\}$ or $\{B',B'',B^c\}$. A sojourn of X in B is now a sequence $X_{t_m},\cdots,X_{t_{m+k}}$, $k\geqq 1$, where the t_i are the instants of transition, $X_{t_m},\cdots,X_{t_{m+k-1}}\in B$, $X_{t_{m+k}}\notin B$ and if $m>0$, $X_{t_{m-1}}\notin B$. This sojourn begins at time t_m and finishes at time t_{m+k}. It lasts $t_{m+k}-t_m$.

Let $H_{B,n}$ be the random variable 'time spent during the nth sojourn of X in B' and let $N_{B,n}$ be as in the previous section for the uniformized Markov chain of X. The hypothesis of irreducibility of the Markov process X assures the existence of an infinity of sojourns of X in B with probability 1. Define V_n as in the previous case for the uniformized discrete-time Markov chain.

Theorem 3.1.

$$\mathbb{P}(H_{B,n}\leqq t)=1-v_n\exp(A_Bt)\,1^T$$

where v_n is as described in Theorem 2.4.

Proof. Let T_n, $n\geqq 1$, be the nth entrance time in the subset B. We have for every $n\geqq 1$,

$$\mathbb{P}(H_{B,n}\leqq t)=\sum_{k=1}^{+\infty}\mathbb{P}(N(T_n+t)-N(T_n)=k)\mathbb{P}(H_{B,n}\leqq t\mid N(T_n+t)-N(T_n)=k).$$

Since T_n is a stopping time, we obtain

$$\mathbb{P}(H_{B,n}\leqq t)=\sum_{k=1}^{+\infty}e^{-\lambda t}\frac{(\lambda t)^k}{k!}\mathbb{P}(N_{B,n}\leqq k)=\sum_{k=1}^{+\infty}e^{-\lambda t}\frac{(\lambda t)^k}{k!}(1-v_nP_B^k\,1^T)$$

$$=1-v_n\exp(-\lambda(I-P_B)t)\,1^T=1-v_n\exp(A_Bt)\,1^T$$

which is the explicit form of the distribution of $H_{B,n}$.

Remark that v_1 and G can also be written as follows:

$$v_1 = \alpha_B - \alpha_{B^c}A_{B^c}^{-1}A_{B^cB}; \quad G = A_B^{-1}A_{BB^c}A_{B^c}^{-1}A_{B^cB}.$$

The moments of $H_{B,n}$ are then easily derived. We have

$$\mathbb{E}(H_{B,n}^k) = (-1)^k k!\, v_n A_B^{-k} 1^T \quad \text{for any } k \geqq 1.$$

Corollary 2.7 has an equivalent in the continuous-time case, with identical (omitted) proof.

Corollary 3.2. *The sequence $(\mathbb{P}(H_{B,n} \leqq t))_{n\geqq 1}$ converges in the sense of Cesaro as $n \to \infty$ and the limit is $1 - v\exp(A_B t)1^T$ where v is given in Corollary 2.7. The convergence is simple for any initial distribution of X if and only if G' is aperiodic.*

As in the discrete-time case, we define the positive real-valued random variable $H_{B,\infty}$ by: $\mathbb{P}(H_{B,\infty} \leqq t) \overset{\text{def}}{=} 1 - v\exp(A_B t)1^T$. We can then derive the Cesaro limits of the kth-order moments of $H_{B,n}$ and easily verify the following relation.

Corollary 3.3. *For any $k \geqq 1$, the sequence $(\mathbb{E}(H_{B,n}^k))_{n\geqq 1}$ converges in the sense of Cesaro and we have*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{l=1}^{n} \mathbb{E}(H_{B,l}^k) = \mathbb{E}(H_{B,\infty}^k) = (-1)^k k!\, v A_B^{-k} 1^T.$$

The convergence is simple for any initial distribution of X if and only if the matrix G' is aperiodic.

4. Pseudo-aggregation

Define the mapping $T: \mathbb{R}^N \to \mathbb{R}^M$ by: $T \cdot v = w$ where $w(m) = \sum_{i\in B(m)} v(i)$.

4.1. *The pseudo-aggregated process.* Let $X = (X_t)_{t\geqq 0}$ be a homogeneous irreducible Markov process with transition rate matrix A and equilibrium probability distribution x. We construct the pseudo-aggregated homogeneous Markov process $Z = (Z_t)_{t\geqq 0}$ from X with respect to the partition $\mathscr{B}$ by defining its transition rate matrix $\hat{A}$ as follows:

$$\forall i, j \in F, \quad \hat{A}(i,j) \overset{\text{def}}{=} \sum_{k\in B(i)} \frac{x(k)}{\sum_{h\in B(i)} x(h)} \sum_{l\in B(j)} A(k,l).$$

We have the same relation concerning the transition probability matrices of the uniformized Markov chains X^* and Z^* of X and Z respectively, denoted by P and $\hat{P}$ respectively. That is

$$\forall i, j \in F, \quad \hat{P}(i,j) \overset{\text{def}}{=} \sum_{k\in B(i)} \frac{x(k)}{\sum_{h\in B(i)} x(h)} \sum_{l\in B(j)} P(k,l).$$

If we denote by α the initial distribution of X, the initial distribution of Z is $T\cdot\alpha$ where the operator T has been defined at the beginning of this section. If the initial distribution α leads to a homogeneous Markov chain for Y, then Z^* and Y define the same homogeneous Markov chain (see [3] or [4]). The stationary distribution of Z is $T\cdot x$ as the following lemma shows.

Lemma 4.1. *If* $z \stackrel{\text{def}}{=} T\cdot x$ *then we have* $z\hat{P} = z$ (*i.e.* $z\hat{A} = 0$), $z > 0$ *and* $z 1^T = 1$.

Proof. For $1 \leqq m \leqq M$, the mth entry of $z\hat{P}$ is equal to

$$\sum_{l=1}^{M} z(l)\hat{P}(l,m) = \sum_{l=1}^{M}\left(\sum_{i\in B(l)} x(i) \sum_{j\in B(m)} P(i,j)\right) = \sum_{i=1}^{N} x(i) \sum_{j\in B(m)} P(i,j)$$

$$= \sum_{j\in B(m)}\left(\sum_{i=1}^{N} x(i)P(i,j)\right) = \sum_{j\in B(m)} x(j) = z(m);$$

the remainder of the proof is obvious.

In the discrete-time and continuous-time cases, the following property holds.

Lemma 4.2. *The pseudo-aggregated process constructed from X with respect to the partition* $\mathcal{B}$ *and the pseudo-aggregated process obtained after M successive aggregations of X with respect to each* $B(i)$, $i \in F$, *in any order, are equivalent.*

The proof is a direct consequence of the construction of the pseudo-aggregated processes. It is for this reason that in what follows we consider only the situation where $\mathcal{B}$ contains only one subset having more than one state. That is, we assume as in the previous sections that $B = \{1,\cdots,L\}$ where $1 < L < N$ and that $\mathcal{B} = \{B, \{L+1\},\cdots,\{N\}\}$ with $N \geqq 3$. The state space of the pseudo-aggregated process Z will be denoted by $F = \{b, L+1,\cdots,N\}$. We shall denote also by B^c the complementary subset $\{L+1,\cdots,N\}$.

4.2. *Pseudo-aggregation and sojourn times.* We consider the pseudo-aggregated homogeneous Markov process Z constructed from X with respect to the partition $\mathcal{B} = \{B, \{L+1\},\cdots,\{N\}\}$ of E. Although the stationary distribution of X over the sets of $\mathcal{B}$ is equal to the state stationary distribution of Z (Lemma 4.1), it is not the same with the distribution of the nth sojourn time of X^* in B and the corresponding distribution of $N_{b,n}$, the nth holding time of Z^* in b, which is independent of n and will be then denoted by N_b. This last (geometric) distribution is given by

$$\mathbb{P}(N_b = k) = \frac{x_B(I-P_B)1^T}{x_B 1^T}\left(\frac{x_B P_B 1^T}{x_B 1^T}\right)^{k-1}, \qquad k \geqq 1$$

which is to be compared with the given expression of $\mathbb{P}(N_{B,n} = k)$.

Observe that if there is no interval transition between different states inside B and if the geometric holding time distributions of X in the individual states of B have the same parameter, we obviously have identity between the different distributions of $N_{B,n}$, $N_{B,\infty}$ and N_b. That is, if $P_B = \beta I$ with $0 \leqq \beta < 1$, we have

$$\mathbb{P}(N_{B,n}=k)=\mathbb{P}(N_{B,\infty}=k)=\mathbb{P}(N_b=k)=(1-\beta)\beta^{k-1}, \qquad k \geqq 1.$$

More generally, if X^* is a lumpable Markov chain [3] over the partition $\mathscr{B}$ (i.e. the aggregated chain Y is also Markov homogeneous for any initial distribution of X^*) we have: $P_B 1^T = \beta 1^T$, $(I-P_B)1^T=(1-\beta)1^T$, where $\beta \in [0, 1[$. In this case, $P_B^{k-1}(I-P_B)1^T=(1-\beta)\beta^{k-1}$, $k \geqq 1$, and the distributions of $N_{B,n}$, $N_{B,\infty}$ and N_b are identical.

We now investigate the relation between moments (factorial moments) of $N_{B,n}$ and $N_{b,n}$. We give an algebraic necessary and sufficient condition for the equality between the Cesaro limit of the kth-order moments of the random variable $N_{B,n}$ and the kth-order moments of N_b.

Theorem 4.3. *For any* $k \geqq 1$:

$$\left[\lim_{n\to+\infty}\frac{1}{n}\sum_{l=1}^{n} FM_k(N_{B,l}) = FM_k(N_b)\right]$$

$$\Leftrightarrow \left[x_B(P_B(I-P_B)^{-1})^{k-1}1^T = x_B 1^T\left(\frac{x_B P_B 1^T}{x_B(I-P_B)1^T}\right)^{k-1}\right].$$

Proof. For the geometric distribution of N_B, we have

$$FM_k(N_b)=\frac{k!(x_B P_B 1^T)^{k-1}x_B 1^T}{(x_B(I-P_B)1^T)^k}.$$

Let us fix $k \geqq 1$. Since from Corollary 2.7,

$$v=\frac{x_B(I-P_B)}{x_B(I-P_B)1^T}$$

we have

$$\left[\lim_{n\to+\infty}\frac{1}{n}\sum_{l=1}^{n} FM_k(N_{B,l}) = FM_k(N_b)\right]$$

$$\Leftrightarrow \left[vP_B^{k-1}(I-P_B)^{-k}1^T = (x_B 1^T)\frac{(x_B P_B 1^T)^{k-1}}{(x_B(I-P_B)1^T)^k}\right]$$

$$\Leftrightarrow \left[x_B(P_B(I-P_B)^{-1})^{k-1}1^T = (x_B 1^T)\left(\frac{x_B P_B 1^T}{x_B(I-P_B)1^T}\right)^{k-1}\right],$$

which is the given condition.

In the important case $k=1$, the above condition is always satisfied. The following corollary states this result. There is no need for a proof since it is trivial to check the condition for $k=1$.

Corollary 4.4.

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{E}(N_{B,l}) = \mathbb{E}(N_b).$$

In continuous time we have analogous results about the relation between properties of sojourn times of the given process X and the corresponding holding times of the pseudo-aggregated process Z. Denote by H_b the holding time random variable for process Z:

$$\mathbb{P}(H_b \leq t) = 1 - e^{-\mu t} \quad \text{where } \mu = -\frac{x_B A_B 1^T}{x_B 1^T}.$$

The continuous-time version of Theorem 4.3 is then the following.

Theorem 4.5. *For any* $k \geq 1$:

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{E}(H_{B,l}^k) = \mathbb{E}(H_b^k) \Leftrightarrow [(x_B(I - P_B)1^T)^{k-1} x_B(I - P_B)^{-k+1} 1^T = (x_B 1^T)^k].$$

Proof. The kth-order moment of the holding time H_b of Z in b can be written

$$\mathbb{E}(H_b^k) = \frac{k!}{\mu^k} = k!\left(-\frac{x_B 1^T}{x_B A_B 1^T}\right)^k.$$

Then, we have

$$\left[\lim_{n\to+\infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{E}(H_{B,l}^k) = \mathbb{E}(H_b^k)\right] \Leftrightarrow \left[v(I - P_B)^{-k} 1^T = \left(\frac{x_B 1^T}{x_B(I - P_B)1^T}\right)^k\right]$$

$$\Leftrightarrow [(x_B(I - P_B)1^T)^{k-1} x_B(I - P_B)^{-k+1} 1^T = (x_B 1^T)^k]$$

thanks to the expression of v given in Corollary 2.7.

As in the discrete-time case, the given condition is always satisfied for first-order moments. The following result corresponds to Corollary 4.4 with identical trivial verification.

Corollary 4.6.

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{E}(H_{B,l}) = \mathbb{E}(H_{B,\infty}) = \mathbb{E}(H_b).$$

5. Application

In [6], the authors consider a fault-tolerant multiprocessor system with k buffer stages. Processors fail independently at rate σ and are repaired singly with rate δ. Buffer stages fail independently at rate γ and are repaired singly with rate τ. Processor failure causes a graceful degradation of the system (the number of processors is decreased by one). The system is in a failed state when all processors have failed or any buffer stage has failed. No additional processor failures are assumed to occur when the system is in a failed state. The model is represented by a homogeneous Markov process with the state transition diagram shown in Figure 1. At any time the state of the system is (i,j) where $0 \leqq i \leqq 2$ is the number of non-failed processors, and j is 0 if any of the buffer stages is failed, otherwise it is 1.

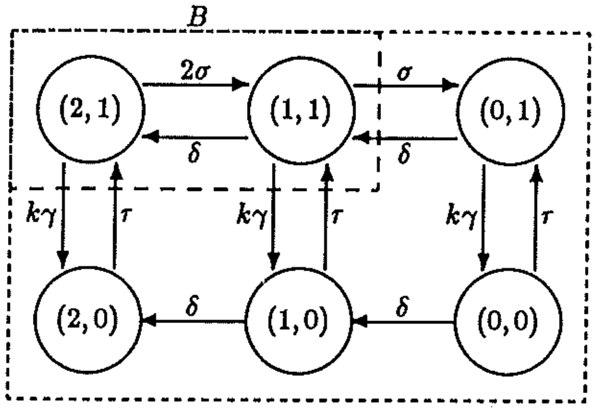

Figure 1. Model of a fault-tolerant multiprocessor system (two-processor case)

We assume the system starting with all its processors and buffers operational, i.e., the initial state is the state (2, 1) with probability 1. We are interested in the states of the set $B = \{(2, 1), (1, 1)\}$, in which the system is operational. We take the following values of the different rates: $\sigma = 6 \times 10^{-5}$, $\delta = \mu = 1/6$, $\gamma = 10^{-2}$; time is in hours [6]. Suppose the user wishes the different sojourns in B to be, in the mean, as 'long' as possible and that the only free parameter is k, the memory size. It is clear that the sequence $\mathbb{E}(H_{B,n})$ is decreasing when k increases, so, the best value for k is $k = 1$. Since, on the other side, the system needs memory enough to perform, we will look for the highest value of k such that the first mean sojourn times be, say, at least 24 hours. Assume also that k must be a power of two. If we compute the expectations $\mathbb{E}(H_{B,1}), \mathbb{E}(H_{B,2}), \cdots$, for $k = 4$, we observe that they are all between 24.9 hours and 25 hours. The reader can verify that $G' = G$ is aperiodic, so, the convergence of $\mathbb{E}(H_{B,n})$ is simple. Moreover, $v_1 \approx v$ ($\| v - v_1 \|_1 = 253431/392266256 \approx 0.000646$); that is why the sequence is almost stationary. If the same computations are carried out for $k = 8$, we get: $\mathbb{E}(H_{B,n}) \approx 12.5$ hours for each n. The distribution of $H_{B,n}$, when $k = 4$, is

$$\mathbb{P}(H_{B,n} \leqq t) = 1 - q_n(e^{ut} - e^{vt}) + \tfrac{1}{2}(e^{ut} + e^{vt})$$

where u, v (the two eigenvalues of A_B) and q_n are given by

$$u = \frac{-9259 + 50\sqrt{15670}}{75000} \qquad v = \frac{-9259 - 50\sqrt{15670}}{75000}$$

$$q_n = \frac{2280879\sqrt{15670}}{1229362446304000}\left(\frac{10559625}{108626189}\right)^{n-1} + \frac{4910386711729\sqrt{15670}}{1229362446304000}.$$

It can be also verified that the condition of Theorem 4.5 is not satisfied but, for instance, the relative deviation between the two respective second-order moments is about 1.8×10^{-5} per cent. All the computations have been done using the MACSYMA package.

6. Conclusions

In this paper we investigate the sojourn time of a homogeneous finite Markov process in a given subset of the state space in both the discrete- and continuous-time cases. In particular, analytical expression of the distributions of the nth sojourn or visit of the process to the subset are derived and their asymptotic behaviour is analyzed.

When the system analyst is interested only in steady state behaviour, it is usual to replace the original model by a 'pseudo-aggregation' where the chosen subset is collapsed into a single state. This is done for instance when the state space has too many elements. This pseudo-aggregation is constructed such that the Markovian property is conserved even if the aggregation of the original process is not Markov. The interest of this procedure is that the steady state probability that the original process will be in the given subset is equal to the steady state probability that the pseudo-aggregated one will be in the corresponding individual state. It is natural then to look at the relations between sojourn times in the first process and the corresponding holding times in the pseudo-aggregation. We show that we have equality between the limit in the sense of Cesaro of successive sojourn time expectations and the corresponding mean holding time in the associated pseudo-aggregated process, and that the equality is no longer valid when greater-order moments are considered.

References

[1] DE SOUZA E SILVA, E. AND GAIL, H. R. (1986) Calculating cumulative operational time distributions of repairable computer systems. *IEEE Trans. Computers* 35, 322–332.

[2] MEYER, J. F. (1980) On evaluating the performability of degradable computing systems. *IEEE Trans. Computers* 29, 720–731.

[3] KEMENY, J. G. AND SNELL, J. L. (1976) *Finite Markov Chains*. Springer-Verlag, New York.

[4] RUBINO, G. AND SERICOLA, B. (1989) On weak lumpability in Markov chains. *J. Appl. Prob.* 26(3).

[5] ROSS, S. M. (1983) *Stochastic Processes*. Wiley, New York.

[6] KULKARNI, V. G., NICOLA, V. F., SMITH, R. M. AND TRIVEDI, K. S. (1986) Numerical evaluation of performability and job completion time in repairable fault-tolerant systems. In *Proc. IEEE 16th Fault-Tolerant Computing Symposium (Vienna, Austria)*.