Statistics & Probability Letters 60 (2002) 445–457

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

A characterization for mixtures of semi-Markov processes☆

I. Epifani[a], S. Fortini[b], L. Ladelli[c,*]

[a] *Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milano, Italy*
[b] *Istituto di Metodi Quantitativi, Università "L. Bocconi", Milano, Italy*
[c] *Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy*

Received April 2001; received in revised form March 2002

Abstract

Mixtures of recurrent semi-Markov processes are characterized through a partial exchangeability condition of the array of successor states and holding times. A stronger invariance condition on the joint law of successor states and holding times leads to mixtures of Markov laws.

MSC: primary: 60G05; secondary: 60J05; 60K15; 62A15

Keywords: Holding times; Markov chains; Markov exchangeability; Partial exchangeability; Semi-Markov processes; Successor state

1. Introduction

The present paper aims at characterizing mixtures of distributions of right-continuous recurrent semi-Markov processes with countable state space. Mixtures of Markov jump processes are obtained as a special case. No results are available, to our knowledge, on mixtures of laws of general semi-Markov processes. An initial result about the problem of characterizing mixtures of distributions of continuous time Markov chains is due to Freedman (1963). More recently, Freedman (1996) gave a simpler proof. He characterized mixtures of distributions of Markov chains with a single recurrence class of stable states as those probability laws that lead to *discrete skeleton processes at the scale h* satisfying a *Markov exchangeability* condition (F), for each $h > 0$. A random sequence $(X_n)_{n \geqslant 0}$ meets (F) if the law of $(X_n)_{n \geqslant 0}$ is invariant under all finite permutations which do not

☆ Research partially supported by Progetti MURST "Modelli Statistici: basi probabilistiche e procedure per l'inferenza e le decisioni" (1998) e "Processi Stocastici, Calcolo Stocastico e Applicazioni" (1999).

* Corresponding author. Tel.: +39-02-2399-4599; fax: +39-02-2399-4568.
E-mail address: ladelli@mate.polimi.it (L. Ladelli).

PII: S0167-7152(02)00324-3

alter the initial state and the number of transitions between any two states of the chain. Condition (F) has been introduced in Freedman (1962) and taken up again in Diaconis and Freedman (1980) in characterizing mixtures of laws of discrete-time Markov chains. The problem of characterizing discrete-time Markov chains with countable state space has been addressed by many other authors. See, for example, Kallenberg (1982), Zaman (1984, 1986), Zabell (1995). In Fortini et al. (1999), a characterization is obtained in terms of partial exchangeability of the array of the *successor states*. Given a process $(X_n)_{n \geqslant 0}$ with values in a countable space I, let us consider the matrix S, whose (i,n)th entry denotes the position of the process immediately after the nth visit to the state i. Fortini et al. (1999) prove that $(X_n)_{n \geqslant 0}$ is recurrent and satisfies condition (F) if and only if S is row-wise partially exchangeable. Actually, the idea of characterizing mixtures of distributions of Markov chains through partial exchangeability of successor states dates back to de Finetti (1959). In Fortini et al. (1999) an analogous result is obtained in the case of general state space through a countable classification of the realizations of the process.

In this paper, we apply this last result to obtain the following characterization. We are interested in minimal chains $(X_t)_{t \geqslant 0}$ with countable state space I, i.e. right-continuous step processes that may have infinitely many jumps in a finite time interval and then get stuck in an extra state not belonging to I. Such processes are completely described in terms of their jump chain and holding times sequence. We consider the matrix (S,T) whose (i,n)th entry (S_{in}, T_{in}) gives the state visited by $(X_t)_{t \geqslant 0}$ after the nth visit to i, and the holding time in i at the same visit, respectively. We show that the law of $(X_t)_{t \geqslant 0}$ is a mixture of laws of recurrent minimal semi-Markov processes if and only if row-wise partial exchangeability of (S,T) holds. The unicity of the mixing measure is discussed. A characterization for mixtures of laws of minimal recurrent Markov chains is achieved by a stronger partial exchangeability condition on (S,T) together with the symmetry condition on the law of $(T_{in})_{n \geqslant 1}$ (for all $i \in I$) borrowed from Diaconis and Ylvisaker (1985) which leads to mixtures of exponential laws.

Compared to our result, Freedman (1996) does not require the smoothness conditions on the sample paths that we assume and, therefore, he comes to represent a larger class of Markov exchangeable processes. On the other hand, our result applies to mixtures of semi-Markov laws as well. Finally, we stress that our approach is different. Our characterization has some statistical implications, since it is based on the exchangeability of a suitable array of observable random variables. This fact allows to give an interpretation of the mixing law.

The paper is organized as follows. In Section 2, we recall some definitions and some results concerning semi-Markov processes with discrete state space. Moreover, we define the array of successor states and holding times and we introduce the partial exchangeability condition (PE1) for this array. Section 3 contains the main result and gives an alternative partial exchangeability hypothesis (PE2) that leads to an equivalent representation. Finally, Section 4 deals with mixtures of laws of continuous time minimal Markov chains.

2. Notation and preliminary results

We focus our attention on processes whose trajectories are right-continuous step functions with values in a countable set I. Moreover, the process may disappear after leaving some state j in I or it may make infinitely many jumps in a finite interval. In this last case, after the explosion time the process starts up again. Anyway, we are not interested in the behavior of the process after the explosion.

Without loss of generality, we will assume $I=\mathbb{N}$. Let us introduce a new fictitious state $\partial \notin I$ and let us put $I^*=\mathbb{N}\cup\{\partial\}$. Let Ω be the set of *generalized right-continuous step functions* (abbreviated g.s.f.) from $[0,\infty)$ into I^*, namely: a function $\omega\in\Omega$ is a right-continuous step function with values in I^* up to explosion time, with ∂ an absorbing state, and it remains constant and equal to ∂ after an explosion. Let $\sigma_1(\omega)<\sigma_2(\omega)<\cdots$ be the discontinuities of ω and set $\sigma_0(\omega):=0$. If $\sigma_n(\omega)<+\infty$ let $\xi_n(\omega)$ be the value of ω on the interval $[\sigma_n(\omega),\sigma_{n+1}(\omega))$ and let $\tau_n(\omega)=\sigma_{n+1}(\omega)-\sigma_n(\omega)$ be the holding time of ω in $\xi_n(\omega)$. If ω has only n discontinuities, set $\sigma_{n+1}(\omega)=\sigma_{n+2}(\omega)=\cdots=+\infty$, $\xi_{n+1}(\omega)=\xi_{n+2}(\omega)=\cdots=\xi_n(\omega)$ and $\tau_n(\omega)=+\infty$. Finally, $\zeta(\omega)=\sum_{n=0}^{\infty}\tau_n(\omega)$ is the explosion time of ω. Introduce now the canonical process $X=(X_t)_{t\geqslant 0}$ on Ω given by $X_t(\omega)=\omega(t)$ and endow Ω with the smallest σ-field $\mathscr{F}$ w.r.t. which all X_t are measurable. One can show that ξ_n and τ_n are $\mathscr{F}$-measurable for every $n\geqslant 0$. Actually, the sequences $\xi=(\xi_n)_{n\geqslant 0}$ and $\tau=(\tau_n)_{n\geqslant 0}$ span $\mathscr{F}$. In the following, we will deal only with probability measures P on $(\Omega,\mathscr{F})$ such that $P\{X_0=i_0\}=1$, for a fixed $i_0\in I$. Let us now recall the notion of semi-Markov process (see, for instance, Pyke, 1961). To simplify it we introduce the fictitious random variable $\tau_{-1}:=1$.

Definition 1. Let $H=(H_i)_{i\in I^*}$ be a kernel on $I^*\times(\mathscr{P}(I^*)\otimes\mathscr{B}((0,+\infty]))$ such that, under P, $(\xi_n,\tau_{n-1})_{n\geqslant 0}$ is a two-dimensional Markov process that satisfies $\xi_0=i_0$ and

$$P\{\xi_n=j,\tau_{n-1}\in C\mid \xi_0,\xi_1,\ldots,\xi_{n-1},\tau_0,\ldots,\tau_{n-2}\}=H_{\xi_{n-1}}(\{j\},C)\quad \text{a.s.-}P$$

for all $n\geqslant 1$, $j\in I^*$ and $C\in\mathscr{B}((0,+\infty])$. Then X is an I^*-valued *semi-Markov process* starting from i_0.

By construction of Ω, ∂ is an absorbing state, hence $H_\partial=\delta_{\{\partial\}\times\{+\infty\}}$. Moreover, it follows from the definition of ξ that $H_i(\{i\},(0,+\infty))=0$. For the sake of simplicity, in what follows we will write $H_i(j,\cdot)$ instead of $H_i(\{j\},\cdot)$ and we shall refer to a semi-Markov process starting from i_0 with kernel H as $S(i_0,H)$. For semi-Markov processes the following characterization holds: the process X is $S(i_0,H)$ if and only if there exist a stochastic matrix Γ on I^* and a matrix Q of probability measures on $((0,\infty],\mathscr{B}((0,\infty]))$ such that

(S–M) ξ is an I^*-valued Markov chain starting from i_0 with transition matrix Γ and for each $n\geqslant 1$, conditionally on $\xi_1,\ldots,\xi_n$, the holding times $\tau_0,\ldots,\tau_{n-1}$ are independent random variables with distributions $Q_{i_0\xi_1},\ldots,Q_{\xi_{n-1}\xi_n}$, respectively.

Moreover, H,Γ,Q satisfy

$$H_i(j,\cdot)=\Gamma_{ij}Q_{ij}(\cdot)\quad \forall i,j\in I^*, \tag{1}$$

$$Q_{ij}=\delta_{+\infty}\quad \text{if } \Gamma_{ij}=0\quad \forall (i,j)\neq(\partial,\partial). \tag{2}$$

It is immediate that

$$\Gamma_{\partial\partial}=1,\quad Q_{\partial\partial}=\delta_{+\infty},\quad Q_{ii}=\delta_{+\infty}\quad \text{if } \Gamma_{ii}>0,\quad Q_{ij}((0,\infty))=1\quad \text{if } \Gamma_{ij}>0 \text{ and } i\neq j. \tag{3}$$

Let us now define the array of successor states and holding times of X. For any i in I^* let v_{im} be the mth visiting time (w.r.t. the embedded Markov chain ξ_n) to i:

$$v_{im} = \inf\{n \geqslant 0 : v_{im-1} < n, \xi_n = i\}, \quad m = 1, 2, \ldots \ (v_{i0} := -1, \inf \emptyset = +\infty).$$

The successor state of the mth visit of X to i is

$$S_{im} := \xi_{v_{im}+1}$$

and the holding time of X in the state i at the mth visit is

$$T_{im} := \tau_{v_{im}}$$

with the convention $\xi_{+\infty} = \partial$ and $\tau_{+\infty} = +\infty$. Note that for the extra state ∂ we have $S_{\partial m} = \partial$ and $T_{\partial m} = +\infty \ \forall m \geqslant 1$. The elements of the array $(S, T) := ((S_{im}, T_{im}))_{i \in I^*, m \geqslant 1}$ are row-wise P-partially exchangeable if

$$P\left(\bigcap_{i \in \mathscr{K}} \bigcap_{m=1}^{n} \{S_{im} \in A_{im}, T_{im} \in C_{im}\}\right) = P\left(\bigcap_{i \in \mathscr{K}} \bigcap_{m=1}^{n} \{S_{im} \in A_{i\pi_i(m)}, T_{im} \in C_{i\pi_i(m)}\}\right) \tag{PE1}$$

for all $\mathscr{K} = \{\partial, 1, \ldots, k\}$, $A_{im} \in \mathscr{P}(I^*)$, $C_{im} \in \mathscr{B}((0, +\infty])$ and for each π_i (with $i \in \mathscr{K}$) varying on the permutations of $\{1, \ldots, n\}$ with $k, n \geqslant 1$. Under (PE1), the following results hold. If $\mathbb{H}$ stands for the set of all probability measures on $(I^* \times (0, +\infty], \mathscr{P}(I^*) \otimes \mathscr{B}((0, +\infty]))$, made into a topological space by the topology of weak convergence, then there is a sequence of random probability measures $\tilde{H} = (\tilde{H}_i)_{i \in I^*}$ from $(\Omega, \mathscr{F})$ into $\mathbb{H}^\infty$ such that

$$\frac{1}{n} \sum_{m=1}^{n} \delta_{(S_{im}, T_{im})} \Rightarrow \tilde{H}_i \quad \text{a.s.-}P \tag{4}$$

$(i \in I^*, n \to +\infty)$ (where $\Rightarrow$ denotes weak convergence). Moreover, if $P_{\tilde{H}}$ denotes the conditional probability on $(\Omega, \mathscr{F})$ given $\tilde{H}$, (PE1) is equivalent to

$$P_{\tilde{H}}\left(\bigcap_{i \in \mathscr{K}} \bigcap_{m=1}^{n} \{S_{im} \in A_{im}, T_{im} \in C_{im}\}\right) = \prod_{i \in \mathscr{K}} \prod_{m=1}^{n} \tilde{H}_i(A_{im}, C_{im}) \quad \text{a.s.-}P. \tag{5}$$

See de Finetti's representation for partially exchangeable arrays (de Finetti, 1938; Link, 1980).

3. Partial exchangeability of (*S, T*) and mixtures of semi-Markov processes

In this section, we characterize laws which are mixtures of recurrent semi-Markov distributions. In Theorem 1 the characterization is given in terms of the partial exchangeability by rows of the array of the bidimensional random variables (S_{in}, T_{in}) $i \in I^*$, $n \geqslant 1$, namely (PE1). The mixing measure turns out to be a probability on a class of kernels H. Proposition 1 states the equivalence between (PE1) and a partial exchangeability condition involving the holding times in i when the process next makes a transition into state j, for all i, j (see condition (PE2)). This result leads to the equivalent representation (11) for the law of X as a mixture of semi-Markov laws. Furthermore, if we consider (PE2), the mixing measure is a probability measure on the set of the couples (Γ, Q) of the jump matrix and the conditional distributions of the holding times.

Here we adopt the same definitions of the previous section. The following lemma concerns the recurrence of the initial state i_0 of X (as defined in Norris, 1997) and the consequent regularity properties of the trajectories of X entailed by (PE1).

Lemma 1. *If property* (PE1) *holds, then* i_0 *is a recurrence state for* X, *i.e.*

$$P\{\omega\in\Omega : A_{i_0}(\omega) \text{ is unbounded}\}=1, \tag{6}$$

where $A_{i_0}(\omega)=\{t: X_t(\omega)=i_0\}$. *Moreover*, (6) *implies*

$$P\{\xi\in I^\infty\}=1 \tag{7}$$

and

$$P\{\zeta<+\infty\}=0. \tag{8}$$

Proof. First, from the row-wise partial exchangeability of the elements in T, it follows that $P\{T_{i1}<+\infty,\dots,T_{im}<+\infty,T_{im+1}=+\infty\}=P\{T_{i1}=+\infty,\dots,T_{im}<+\infty,T_{im+1}<+\infty\}=0$. Hence $\{T_{i1}<+\infty\}=\bigcap_{n\geqslant 1}\{T_{in}<+\infty\}$ a.s.-P $\forall i\in I$. Moreover, $\{T_{in}<+\infty\}\subset\{v_{in}<+\infty\}$. Let λ denote the Lebesgue measure on $(0,+\infty]$. Then

$$\lambda(A_{i_0}(\omega))=\sum_{n=1}^{\infty}T_{i_0n}(\omega)\mathbf{1}(v_{i_0n}(\omega)<+\infty)$$

and

$$P\{\omega\in\Omega:\lambda(A_{i_0}(\omega))=+\infty\}=P\{T_{i_01}=+\infty\}+P\left\{\sum_{n=2}^{\infty}T_{i_0n}=+\infty,T_{i_01}<+\infty\right\}.$$

Let us now prove that

$$P\left\{\sum_{n=2}^{\infty}T_{i_0n}=+\infty,T_{i_01}<+\infty\right\}=P\{T_{i_01}<+\infty\}.$$

Suppose $P\{T_{i_01}<+\infty\}>0$. The sequence of strictly positive random variables $(T_{i_0n})_{n\geqslant 2}$ is exchangeable w.r.t. the probability measure $P^1(\cdot):=P(\cdot\mid T_{i_01}<+\infty)$. If $\mathrm{E}^1(T_{i_02})<+\infty$, by the strong law of large numbers for exchangeable sequences, there exists a σ-field $\mathscr{A}_\infty$ such that $\sum_{m=2}^{n+1}T_{i_0m}/n$ converges to the positive random variable $\mathrm{E}^1(T_{i_02}\mid\mathscr{A}_\infty)$ a.s.-P^1 (see Theorem 1.62 in Schervish, 1997). Therefore, $P^1\{\sum_{n=2}^{\infty}T_{i_0n}=+\infty\}=1$. If $\mathrm{E}^1(T_{i_02})=+\infty$ this result remains true and (6) holds. Condition (7) follows immediately from (6) because the extra state ∂ is absorbing, whereas, (8) is implied by (6) since $\zeta(\omega)=\sum_{n=1}^{\infty}\tau_n(\omega)\geqslant\lambda(A_{i_0}(\omega))$ for any $\omega\in\Omega$. □

For every kernel H in $\mathbb{H}^\infty$, let $\{\mathscr{T}_H,\mathscr{R}_H\}$ be the partition of I^* defined by

$$\mathscr{T}_H=\{i\in I^*: i\neq i_0 \text{ and } H_i=\delta_{\{\partial\}\times\{+\infty\}}\} \tag{9}$$

and let $\mathscr{H}_0$ be the measurable subset of kernels H such that $\mathscr{R}_H$ is a single indecomposable class.

Theorem 1. *Let $I^* = \mathbb{N} \cup \{\partial\}$. The elements of the array (S, T) are P-partially exchangeable by rows if and only if there is a probability measure μ on $(\mathbb{H}^\infty, \mathscr{B}(\mathbb{H}^\infty))$ such that*

(i) $\mu(\mathscr{H}_0) = 1$,
(ii) $\mu\{H \in \mathscr{H}_0$ *such that* i_0 *is recurrent for* $S(i_0, H)\} = 1$,
(iii) *for any* $i_1, \ldots, i_n$ *in* I, $C_0, \ldots, C_{n-1}$ *in* $\mathscr{B}((0, +\infty])$ *and* $n \geqslant 1$:

$$P\{\xi_1 = i_1, \ldots, \xi_n = i_n, \tau_0 \in C_0, \ldots, \tau_{n-1} \in C_{n-1}\} = \int_{\mathbb{H}^\infty} \prod_{s=0}^{n-1} H_{i_s}(i_{s+1}, C_s)\mu(\mathrm{d}H).$$

Furthermore the mixing measure μ is uniquely determined.

Proof. The proof is obtained by applying the results in Section 5 in Fortini et al. (1999) to the chain $W = (W_n)_{n \geqslant 0} = (\xi_n, \tau_{n-1})_{n \geqslant 0}$, with state space $\mathbb{S} = I \times (0, +\infty]$. Fortini et al. (1999) give a characterization of a class of discrete-time processes with values in a Polish space, whose law is a mixture of recurrent Markov laws, in terms of partial exchangeability of a suitable array of "successor states".

Consider $\mathbb{S}^* = \mathbb{S} \cup \{(\partial, +\infty)\}$, the Borel σ-algebra $\mathscr{S}$ on $\mathbb{S}$ and the σ-algebra $\mathscr{S}^*$ generated by $\mathscr{S} \cup \{(\partial, +\infty)\}$. Let $(A_i)_{i \geqslant 0}$ be the partition of $\mathbb{S}^*$ defined by $A_0 = \{(\partial, +\infty)\}$ and $A_i = \{i\} \times (0, +\infty]$ for $i = 1, 2, \ldots$. According to Fortini et al. (1999), introduce the I^*-valued process $Y = (Y_n)_{n \geqslant 0}$

$$Y_n = \partial \mathbf{1}_{A_0}(W_n) + \sum_{i=1}^{\infty} i \mathbf{1}_{A_i}(W_n)$$

and consider the time of the nth visit of Y to i, say $\bar{v}_{in}$, and the successor state $\beta_n^{(i)} := W_{\bar{v}_{in}+1}$. Then, $\bar{v}_{jn} = v_{jn}$ (with v_{jn} defined as in Section 2) and hence

$$\beta_n^{(i)} = (S_{in}, T_{in}) \quad \forall i \in I^*, \ \forall n \geqslant 1.$$

Assume first the elements of (S, T) are P-partially exchangeable by rows. Then, by (7) in Lemma 1, $P\{W = (\xi_n, \tau_{n-1})_{n \geqslant 0} \in \mathbb{S}^\infty\} = 1$. We can apply Theorem 5.2 in Fortini et al. (1999) to the process W. Hence (i) and (iii) are proved. As far as condition (ii) is concerned, it follows from (6) in Lemma 1. The "only if" part can be obtained using Theorem 5.3 of Fortini et al. (1999) for our construction. □

Condition (iii) can be expressed by saying that there exists a random kernel $\tilde{H}$ such that

$$P_{\tilde{H}}\{\xi_1 = i_1, \ldots, \xi_n = i_n, \tau_0 \in C_0, \ldots, \tau_{n-1} \in C_{n-1}\} = \prod_{s=0}^{n-1} \tilde{H}_{i_s}(i_{s+1}, C_s) \tag{10}$$

holds a.s.-P for any $i_1, \ldots, i_n$ in I^*, $C_0, \ldots, C_{n-1}$ in $\mathscr{B}((0, +\infty])$ and $n \geqslant 1$. It follows from the proofs of Theorems 5.2 and 5.3 in Fortini et al. (1999) that $\tilde{H}$ is the a.s.-P limit of the empirical process of (S, T) given in (4). On the other hand, Eq. (10) can be rewritten as

$$P_{\tilde{H}}\{\xi_1 = i_1, \ldots, \xi_n = i_n, \tau_0 \in C_0, \ldots, \tau_{n-1} \in C_{n-1}\} = \prod_{s=0}^{n-1} \tilde{\Gamma}_{i_s i_{s+1}} \tilde{Q}_{i_s i_{s+1}}(C_s), \quad \text{a.s.-}P \tag{11}$$

whenever the stochastic matrix $\tilde{\Gamma} = (\tilde{\Gamma}_{ij})_{i,j \in I^*}$ and the family of random probability measures $\tilde{Q} = (\tilde{Q}_{ij})_{i,j \in I^*}$ on $((0,\infty], \mathscr{B}((0,+\infty]))$ are defined by

$$\tilde{\Gamma}_{ij} = \tilde{H}_i(j,(0,+\infty]), \tag{12}$$

$$\tilde{Q}_{ij}(\cdot) = \frac{\tilde{H}_i(j,\cdot)}{\tilde{\Gamma}_{ij}} \mathbf{1}(\tilde{\Gamma}_{ij} > 0) + \delta_{+\infty}(\cdot)\mathbf{1}(\tilde{\Gamma}_{ij} = 0). \tag{13}$$

Now, in order to give an interpretation of $(\tilde{\Gamma},\tilde{Q})$ in terms of empirical processes, we introduce the array $(T_{ijm})_{i,j \in I^*, m \geqslant 1}$ of the holding times in i when X next makes a jump to j, i.e. $T_{ijm} = \tau_{v_{ijm}}$, where v_{ijm} is the mth visit of X to the string (i,j):

$$v_{ijm} = \inf\{n \geqslant 0 : v_{ijm-1} < n, \xi_n = i, \xi_{n+1} = j\} \quad m \geqslant 1 \;\; (v_{ij0} := -1, \inf \emptyset = +\infty).$$

We will now prove that (PE1) is equivalent to the partial exchangeability condition given by

$$P\left(\bigcap_{i,j \in \mathscr{K}} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{ijm} \in C_{ijm}\}\right) = P\left(\bigcap_{i,j \in \mathscr{K}} \bigcap_{m=1}^{n} \{S_{im} = x_{i\pi_i(m)}, T_{ijm} \in C_{ij\rho_{ij}(m)}\}\right) \tag{PE2}$$

for all $\mathscr{K} = \{\partial, 1, \ldots, k\}$, $x_{im} \in I^*$, $C_{ijm} \in \mathscr{B}((0,+\infty])$ and for each π_i and ρ_{ij} (with $i,j \in \mathscr{K}$) varying on the permutations of $\{1,\ldots,n\}$.

Proposition 1. *Conditions* (PE1) *and* (PE2) *are equivalent.*

Proof. First, note that (PE2) entails $T_{i\partial 1} = +\infty \;\; \forall i \in I$. Then, (7) holds both under (PE1) and (PE2) so that it is enough to verify the equivalence between (PE1) and (PE2) only for $i \neq j$, $i,j \in \{1,\ldots,k\}$, x_{in} in I and $C_{in} = (0,t_{in}]$, $C_{ijn} = (0,t_{ijn}]$ with $t_{in}, t_{ijn} \in (0,+\infty)$.

Let $1 = s_{i1}, s_{i2}, \ldots, s_{in}$ be the unique integers such that

$$\bigcap_{i=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{im} \leqslant t_{im}\} = \bigcap_{i=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{ix_{im}s_{im}} \leqslant t_{im}\}.$$

Then, for every permutation π_i of $1,\ldots,n$,

$$\bigcap_{i=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{i\pi_i(m)}, T_{ix_{im}s_{im}} \leqslant t_{im}\} = \bigcap_{i=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{i\pi_i(m)}, T_{i\pi_i(m)} \leqslant t_{im}\}.$$

It follows that (PE2) implies (PE1). Conversely, assume (PE1) and verify that a.s.-P

$$P_{\tilde{\Gamma},\tilde{Q}}\left(\bigcap_{i \neq j = 1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{ijm} \leqslant t_{ijm}\}\right) = \prod_{i \neq j = 1}^{k} \prod_{m=1}^{n} \tilde{\Gamma}_{ix_{im}} \tilde{Q}_{ij}((0,t_{ijm}]) \tag{14}$$

with $\tilde{\Gamma},\tilde{Q}$ defined as in (12) and (13). To this end, note that if (PE1) is in force then

$$P_{\tilde{H}}\left(\bigcap_{i=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{im} \leqslant t_{im}\}\right) = \prod_{i=1}^{k} \prod_{m=1}^{n} \tilde{\Gamma}_{ix_{im}} \tilde{Q}_{ix_{im}}((0,t_{im}]) \quad \text{a.s.-}P. \tag{15}$$

Moreover, under (PE1), the following two conditions of recurrence of the accessible states hold:

$$\{v_{ij1} < +\infty\} = \bigcap_{m \geqslant 1} \{v_{ijm} < +\infty\} \quad \text{a.s.-}P \ \forall i, j \in I \tag{16}$$

and

$$\{v_{i1} < +\infty\} = \bigcap_{m \geqslant 1} \{v_{im} < +\infty\} \quad \text{a.s.-}P \ \forall i \in I. \tag{17}$$

In fact, if $P\{v_{ij1} < +\infty\} > 0$, we get, for every $m \geqslant 1$,

$$\begin{aligned} &P\{v_{ij1} < +\infty, \dots, v_{ijm} < +\infty, v_{ijm+1} = +\infty\} \\ &\quad = \sum_{k_1 < \cdots < k_m} P\{S_{ik_1} = \cdots = S_{ik_m} = j, v_{ijm+1} = +\infty\} \\ &\quad = \sum_{k_1 < \cdots < k_m} P\{S_{i1} = \cdots = S_{im} = j, S_{in} \neq j \ \forall n \geqslant m+1\} \quad \text{[by the partial exchangeability of } S\text{]} \\ &\quad = 0 \end{aligned}$$

so that (16) is satisfied. On the other hand, by (7) and (16) we have $\{v_{i1} < +\infty\} = \bigcup_{j \in I} \{v_{ij1} < +\infty\} = \bigcup_{j \in I} \bigcap_{m \geqslant 1} \{v_{ijm} < +\infty\} \subseteq \bigcap_{m \geqslant 1} \{v_{im} < +\infty\}$ a.s.-P.

Therefore, for fixed $k \geqslant 1$, the following equalities hold a.s.-P:

$$\begin{aligned} &P_{\tilde{H}}\left(\bigcap_{i \neq j=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{ijm} \leqslant t_{ijm}\} \right) \\ &\quad = P_{\tilde{H}}\left(\bigcap_{i \neq j=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{ijm} \leqslant t_{ijm}, v_{ijm} < +\infty\} \right) \\ &\quad = \sum_{\{m \leqslant k_{ijm} < k_{ijm+1}\}} P_{\tilde{H}}\left(\bigcap_{i \neq j=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, v_{ijm} = v_{ik_{ijm}}, S_{ik_{ijm}} = j, T_{ik_{ijm}} \leqslant t_{ijm}\} \right) \\ &\quad = \sum_{\{m \leqslant k_{ijm} < k_{ijm+1}\}} P_{\tilde{H}}\left(\bigcap_{i \neq j=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, v_{ijm} = v_{ik_{ijm}}, S_{ik_{ijm}} = j\} \right) \prod_{i \neq j=1}^{k} \prod_{m=1}^{n} \tilde{Q}_{ij}((0, t_{ijm}]) \end{aligned}$$

(here we use the measurability of the sets $\{v_{ijn} = v_{ik_{ijn}}\}$ w.r.t. S, (17) which ensures $v_{ik_{ijm}} < +\infty$ $\forall k_{ijm}$ on the set $\{S_{im} = x_{im}\}$ with $x_{im} \in I$, and (15))

$$= P_{\tilde{H}}\left(\bigcap_{i \neq j=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, v_{ijm} < +\infty\} \right) \prod_{i \neq j=1}^{k} \prod_{m=1}^{n} \tilde{Q}_{ij}((0, t_{ijm}])$$

$$= P_{\tilde{H}} \left(\bigcap_{i \neq j=1}^{k} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, v_{ij1} < +\infty\} \right) \prod_{i \neq j=1}^{k} \prod_{m=1}^{n} \tilde{Q}_{ij}((0, t_{ijm}]) \quad \text{[by (16)]}$$

$$= \prod_{i \neq j=1}^{k} \prod_{m=1}^{n} \tilde{\Gamma}_{ix_{im}} (1 - \delta_{\tilde{\Gamma}_{ij}}(0)) \tilde{Q}_{ij}((0, t_{ijm}])$$

(since $P_{\tilde{H}}(\{v_{ij1} = +\infty\}) = P_{\tilde{H}}\{S_{in} \neq j \ \forall n \geqslant 1\} = \lim_{n \to +\infty} (1 - \tilde{\Gamma}_{ij})^n = \delta_{\tilde{\Gamma}_{ij}}(0)$ a.s.-P)

$$= \prod_{i \neq j=1}^{k} \prod_{m=1}^{n} \tilde{\Gamma}_{ix_{im}} \tilde{Q}_{ij}(0, t_{ijm}]). \qquad \square$$

From the equivalence between (PE1) and (PE2) we deduce that the representation of the law of $(\xi_n, \tau_{n-1})_{n \geqslant 0}$ in (11) holds also under (PE2). Moreover, from (14) we obtain that for every $i, j \in I^*$, a.s.-P

$$\frac{1}{n} \sum_{m=1}^{n} \delta_{S_{im}} \Rightarrow \tilde{\Gamma}_{ij} \quad (n \to +\infty),$$

$$\frac{1}{n} \sum_{m=1}^{n} \delta_{T_{ijm}} \Rightarrow \tilde{Q}_{ij} \quad (n \to +\infty),$$

which explain in terms of the process X the meaning of the random matrix $\tilde{\Gamma}, \tilde{Q}$ in (11).

4. About mixtures of Markov processes

In this section, we focus our attention on measures on $(\Omega, \mathscr{F})$ that are mixtures of laws of Markov chains. Roughly speaking, a semi-Markov process X is Markovian if the “waiting” time in a state i depends only on the current state i and it is exponentially distributed. More precisely, let $\mathscr{E}(q)$ be the *enlarged exponential measure* with parameter $q \geqslant 0$: $\mathscr{E}(0)$ identifies the Dirac measure concentrated at $+\infty$. For any fixed (Γ, Q) satisfying (2) and (3), suppose there exists a sequence $q = (q_i)_{i \in I^*}$ of nonnegative numbers such that for any $i, j \in I^*$

$$Q_{ij} = \mathscr{E}(q_i) \mathbf{1}(\Gamma_{ij} > 0) + \mathscr{E}(0) \mathbf{1}(\Gamma_{ij} = 0). \tag{18}$$

Then, by (2), (3) and (18) we easily obtain

$$\Gamma_{\partial\partial} = 1 \quad \text{and} \quad q_\partial = 0$$

and, for every i in I,

$$\Gamma_{ii} = 1 \quad \text{if } q_i = 0 \quad \text{and} \quad \Gamma_{ii} = 0 \quad \text{if } q_i > 0.$$

The pair (Γ, q) defines the I^*-valued Markov processes whose sample paths are g.s.f.’s through

$$P\{\xi_n = j, \tau_{n-1} > t \mid \xi_0, \xi_1, \ldots, \xi_{n-1}, \tau_0, \ldots, \tau_{n-2}\} = \Gamma_{\xi_{n-1} j} e^{-q_{\xi_{n-1}} t} \quad \text{a.s.-}P$$

(see Freedman, 1971, Proposition 48, p. 170, or Norris, 1997, Theorem 2.8.4). If, in addition, the process is minimal then it will be called $M(i_0, \Gamma, q)$.

Let us now introduce a stronger partial exchangeability condition:

(PE3) the elements of the array

$$\begin{bmatrix} S \\ T \end{bmatrix}$$

are *row-wise P-partially exchangeable* if the law of

$$\begin{bmatrix} S \\ T \end{bmatrix}$$

is invariant w.r.t. any finite permutation acting on each row.

Recall that if $\mathbb{P}$ stands for the set of all probability measures on $(I^*, \mathscr{P}(I^*))$ and $\mathbb{Q}$ stands for the set of all probability measures on $((0, +\infty], \mathscr{B}((0, +\infty]))$, made into topological spaces by the topology of weak convergence, then, by de Finetti's Theorem, condition (PE3) holds if and only if there exist two sequences of random probability measures $\tilde{\Gamma} = (\tilde{\Gamma}_i)_{i \in I^*} : \Omega \mapsto \mathbb{P}^\infty$ and $\tilde{\boldsymbol{Q}} = (\tilde{Q}_i)_{i \in I^*} : \Omega \mapsto \mathbb{Q}^\infty$ such that

$$\frac{1}{n} \sum_{m=1}^{n} \delta_{S_{im}} \Rightarrow \tilde{\Gamma}_i, \quad \frac{1}{n} \sum_{m=1}^{n} \delta_{T_{im}} \Rightarrow \tilde{Q}_i \quad (n \to +\infty) \text{ a.s.-}P \tag{19}$$

and

$$P_{\tilde{\Gamma}, \tilde{\boldsymbol{Q}}}\left(\bigcap_{i \in \mathscr{K}} \bigcap_{m=1}^{n} \{S_{im} = x_{im}, T_{im} \in C_{im}\} \right) = \prod_{i \in \mathscr{K}} \prod_{m=1}^{n} \tilde{\Gamma}_i(x_{im}) \tilde{Q}_i(C_{im}),$$

where $P_{\tilde{\Gamma}, \tilde{\boldsymbol{Q}}}$ denotes the conditional probability on $(\Omega, \mathscr{F})$ given $\tilde{\Gamma}, \tilde{\boldsymbol{Q}}$.

As in Diaconis and Ylvisaker (1985) let us consider the following symmetry condition on the law of the holding times T:

$$P\{(T_{i1}, \ldots, T_{in}) \in C\} = P\{(T_{i1}, \ldots, T_{in}) \in C + s\} \tag{20}$$

for all i in I, $n \geqslant 1$, C in $\mathscr{B}((0, +\infty]^n)$, $s = (s_1, \ldots, s_n) \in \mathbb{R}^n$ satisfying $\sum_{j=1}^{n} s_j = 0$ and $C + s \subset (0, +\infty]^n$.

The following holds:

Proposition 2. *Let the elements of the array T be row-wise P-partially exchangeable and let $\tilde{\boldsymbol{Q}}$ be as in (19). Then (20) holds if and only if there exists a random sequence $\tilde{q} = (\tilde{q}_i)_{i \in I}$, with $\tilde{q}_i : \Omega \mapsto [0, \infty)$, such that*

$$\tilde{Q}_i = \mathscr{E}(\tilde{q}_i) \quad \text{a.s.-}P. \tag{21}$$

Proof. If $P(\bigcap_{n=1}^{\infty} \{T_{in} < +\infty\}) = 0$, then $P(\bigcap_{n=1}^{\infty} \{T_{in} = +\infty\}) = 1$ and $\tilde{Q}_i = \delta_{+\infty}$ a.s.-P. Hence, (21) holds with $\tilde{q}_i = 0$ a.s.-P. Suppose now $P(\bigcap_{n=1}^{\infty} \{T_{in} < +\infty\}) > 0$. Let P^* be the probability measure on $(\Omega, \mathscr{F})$ defined by

$$P^*(A) = P\left(A \,\middle|\, \bigcap_{j=1}^{\infty} \{T_{in} < +\infty\}\right).$$

If P satisfies (20), the same condition is also true for P^*. Since under P^* the random variables T_{in}'s are a.s. finite, we can apply the result in Diaconis and Ylvisaker (1985, Theorem 7) and obtain that for every $i \in I$ the sequence $(T_{in})_{n \geqslant 1}$ is P^*-exchangeable and there exists a nonnegative random variable q_i^* on $(\Omega, \mathscr{F})$ such that

$$P^*\{T_{in} > t \,|\, Q_i^*\} = \mathrm{e}^{-q_i^* t}, \quad \forall t \geqslant 0,\ \forall n \geqslant 1 \text{ a.s.-}P^*,$$

where Q_i^* is given by

$$\frac{1}{n}\sum_{m=1}^{n} \delta_{T_{im}} \Rightarrow Q_i^* \quad \text{a.s.-}P^*.$$

On the other hand, (19) holds. Then, for every $i \in I$, we can determine a set $\mathscr{V}_i \subseteq \bigcap_{n=1}^{\infty} \{T_{in} < +\infty\}$, such that $\mathscr{V}_i = \bigcap_{n=1}^{\infty} \{T_{in} < +\infty\}$ a.s.-P and (21) holds with $\tilde{q}_i = q_i^* \mathbf{1}_{\mathscr{V}_i}$. The converse is immediate. □

We are now able to deduce from Section 3 a representation for the law of jump processes satisfying (PE3) and (20). Let $\mathbb{G}$ be the set of stochastic matrices on I^* with the topology of coordinate convergence and let $\mathbb{D} = \mathbb{G} \times [0, +\infty)^{\infty}$. Let us denote by $\mathscr{D}_0$ the measurable subset of $\mathbb{D}$ formed by the pairs (Γ, q), for which the partition $\{\mathscr{R}_{\Gamma,q}, \mathscr{T}_{\Gamma,q}\}$ of I^* defined by

$$\mathscr{T}_{\Gamma,q} = \{i \in I^* : i \neq i_0 \text{ and } q_i = 0\} \tag{22}$$

satisfies the following properties:

$\partial \in \mathscr{T}_{\Gamma,q}$;
if $i \in \mathscr{T}_{\Gamma,q}$, then $\Gamma_{ij} = 0$ for every $j \in I$;
if $\mathscr{R}_{\Gamma,q} = \{i_0\}$, then $\Gamma_{i_0 i_0} = 1$ and $q_{i_0} = 0$;
if $\mathscr{R}_{\Gamma,q} \supsetneq \{i_0\}$, then $q_{i_0} > 0$ and $\Gamma_{ii} = \Gamma_{ij} = 0$ for every $i \in \mathscr{R}_{\Gamma,q}$ and $j \in \mathscr{T}_{\Gamma,q}$.

Theorem 2. *Let X be the canonical process defined in Section 2 and $I^* = \mathbb{N} \cup \{\partial\}$. Then,* (PE3) *holds and the elements of the array T satisfy condition* (20) *if and only if there is a probability measure $\bar{\mu}$ on $(\mathbb{D}, \mathscr{B}(\mathbb{D}))$ such that*

(j) $\bar{\mu}(\mathscr{D}_0) = 1$;
(jj) $\bar{\mu}\{(\Gamma, q) \in \mathscr{D}_0$ *such that i_0 is recurrent for a $M(i_0, \Gamma, q)$ process*$\} = 1$;
(jjj) *for any F in $\mathscr{F}$*

$$P\{X \in F\} = \int_{\mathbb{D}} \mathscr{M}_{i_0,\Gamma,q}(F)\, \bar{\mu}(\mathrm{d}\Gamma\, \mathrm{d}q)$$

where, for (Γ, q) in $\mathscr{D}_0$, $\mathscr{M}_{i_0,\Gamma,q}$ is the probability measure on $(\Omega, \mathscr{F})$ under which X is a $M(i_0, \Gamma, q)$ process.

Furthermore the mixing measure $\bar{\mu}$ is uniquely determined.

Proof. Let us assume (PE3) and (20). Then, (19) holds with $\tilde{Q}_i = \mathscr{E}(\tilde{q}_i)$ a.s.-P and the probability distribution—say $\bar{\mu}$—of $(\tilde{\Gamma}, \tilde{q})$ concentrates all its mass on $\mathscr{D}_0$. Hence, (j) is true. Moreover, (5) is satisfied with $\tilde{H}_i(j,\cdot) = \tilde{\Gamma}_{ij}\tilde{Q}_i(\cdot)$ a.s.-P for every i, j in I^*. By means of the "only if" part of Theorem 1, (jj) and (jjj) follow. Conversely, assume (j)–(jjj) and let $(\tilde{\Gamma}, \tilde{q})$ be a random pair with law $\bar{\mu}$. For every (Γ, q) in $\mathbb{D}$, let $Q(\Gamma, q)$ be defined as in (18), let $H(\Gamma, Q)$ be the kernel given in (1) and

$$\mu(A) = \bar{\mu}\{(\Gamma, q) : H(\Gamma, q) \in A\} \quad \forall A \in \mathscr{B}(\mathbb{H}^\infty).$$

It is easy to verify that for any $(\Gamma, q) \in \mathscr{D}_0$, $\mathscr{T}_{\Gamma,q} = \mathscr{T}_{H(\Gamma,q)}$ with $\mathscr{T}_H$ and $\mathscr{T}_{\Gamma,q}$ defined in (9) and (22), respectively. It follows that $\mathscr{D}_0 = \{(\Gamma, q) : H(\Gamma, q) \in \mathscr{H}_0\}$. Therefore,

(i) $\mu(\mathscr{H}_0) = \bar{\mu}(\mathscr{D}_0) = 1$,
(ii) $\mu\{H \in \mathscr{H}_0$ such that i_0 is recurrent for a $S(i_0, H)$ process$\}$ = $\bar{\mu}\{(\Gamma, q) \in \mathscr{D}_0$ such that i_0 is recurrent for a $M(i_0, \Gamma, q)$ process$\} = 1$,

whereas (iii) in Theorem 1 comes from (jjj) with a change of variables. Now from the "if" part of Theorem 1, (PE1) follows. Moreover (5) holds with $\tilde{H}_i(j,\cdot) = \tilde{\Gamma}_{ij}\mathscr{E}(\tilde{q}_i,\cdot)$ a.s.-P $\forall i, j \in I^*$. Thus (PE3) is achieved. Finally, condition (20) follows from Proposition 2. □

Acknowledgements

We would like to thank Piercesare Secchi for some useful discussions.

References

Diaconis, P., Freedman, D., 1980. De Finetti's theorem for Markov chains. Ann. Probab. 8, 115–130.

Diaconis, P., Ylvisaker, D., 1985. Quantifying prior opinion. With discussion and a reply by Diaconis. In: Bernardo, J.M., DeGroot, M.H., Lindley, D.V., Smith, A.F.M. (Eds.), Bayesian Statistics, Vol. 2 (Valencia, 1983). North-Holland, Amsterdam, pp. 133–156.

de Finetti, B., 1938. Sur la condition d'equivalence partielle. In: Actualités Scientifiques et Industrielles, Vol. 739. Hermann, Paris, pp. 5–18.

de Finetti, B., 1959. La probabilità e la statistica nei rapporti con l'induzione secondo i diversi punti di vista. In: Atti Corso CIME su Induzione e Statistica, Cremonese, Roma, pp. 1–115. [English translation 1972. In: de Finetti, B. (Ed.), Probability, Induction and Statistics. Wiley, New York, pp. 147–227].

Fortini, S., Ladelli, L., Petris, G., Regazzini, E., 1999. On mixtures of distributions of Markov chains. Technical Report, 99.9, CNR-IAMI, Milano.

Freedman, D.A., 1962. Invariants under mixing which generalize de Finetti's theorem. Ann. Math. Statist. 33, 916–923.

Freedman, D.A., 1963. Invariants under mixing which generalize de Finetti's theorem: continuous parameter case. Ann. Math. Statist. 34, 1194–1216.

Freedman, D., 1971. Markov Chains. Springer, Berlin, New York.

Freedman, D.A., 1996. De Finetti's theorem in continuous time. In: Ferguson, T.S., Sharpley, L.S., MacQueen, J.B. (Eds.), Statistics, Probability and Game Theory, IMS Lecture Notes-Monograph Series, Vol. 30. Inst. Math. Statist, Hayward, CA, pp. 83–98.

Kallenberg, O., 1982. Characterizations and embedding properties in exchangeability. Z. Wahrsch. 60, 249–281.

Link, G., 1980. Representations theorems of de Finetti type for (partially) symmetric probability measures. In: Jeffrey, R.C. (Ed.), Studies in Inductive Logic and Probability, Vol. 3. University of California Press, Berkeley, CA, pp. 207–231.

Norris, J.R., 1997. Markov Chains. Cambridge University Press, Cambridge.
Pyke, R., 1961. Markov renewal processes: definitions and preliminary properties. Ann. Math. Statist. 32, 1231–1242.
Schervish, M.J., 1997. Theory of Statistics. Springer, Berlin, New York.
Zabell, S., 1995. Characterizing Markov exchangeable sequences. J. Theoret. Probab. 8, 175–178.
Zaman, A., 1984. Urn models for Markov exchangeability. Ann. Probab. 11, 223–229.
Zaman, A., 1986. A finite form of de Finetti's theorem for stationary Markov exchangeability. Ann. Probab. 14, 1418–1427.