Stochastic Processes and their Applications 45 (1993) 115-125
North-Holland

A finite characterization of weak lumpable Markov processes. Part II: The continuous time case

Gerardo Rubino and Bruno Sericola
IRISA-INRIA, Campus de Beaulieu, Rennes, France

Received 19 June 1990
Revised 10 August 1990 and 3 January 1992

We analysed in the companion paper (Stochastic Process. Appl. 38, 1991), the conditions under which the aggregated process constructed from an irreducible and homogeneous discrete time Markov chain over a given partition of its state space is another homogeneous Markov chain. The obtained result is a characterization of this situation by means of a finite algorithm which computes the set of all the initial probability distributions of the starting chain such that the aggregated one is also Markov homogeneous. In this paper, we consider the same problem in continuous time. Our main result is that it is always possible to come back to the discrete time case using uniformization.

Markov processes * aggregation * weak lumpability * uniformization

1. Introduction

This paper is an extension of the work performed in [2]. Our goal is to show that the problem of markovian state aggregation in irreducible and homogeneous continuous time Markov processes can always be reduced to the same problem in discrete time [2] by means of the uniformization technique. Let us recall briefly the problem of weak lumpability in Markov processes.

Let $X=(X_t)_{t\in\mathcal{T}}$ be an irreducible and homogeneous Markov process evolving in continuous time ($\mathcal{T}=\mathbb{R}_+$) or in discrete time ($\mathcal{T}=\mathbb{N}$). The state space is assumed to be finite and is denoted by $E=\{1,2,\ldots,N\}$. The stationary distribution of X is denoted by π. Let us denote by $\mathcal{B}=\{B(1),B(2),\ldots,B(M)\}$ a partition of the state space. From here the state space E and the partition $\mathcal{B}$ are fixed.

With the given process X we associate the aggregated stochastic process Y with values on $F=\{1,2,\ldots,M\}$, defined by

$$Y_t=m \overset{\text{def}}{\iff} X_t\in B(m) \quad \text{for all } t\in\mathcal{T}.$$

It is easily checked from this definition and the irreducibility of X that the obtained process Y is also irreducible in the following sense: For any $m\in F$ and $l\in F$ such that $\mathbb{P}(Y_0=l)>0$, there exists $t\in\mathcal{T}$, $t\neq 0$, such that $\mathbb{P}(Y_t=m/Y_0=l)>0$.

Correspondence to: Drs. G. Rubino and B. Sericola, IRISA-INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France.

We analyse under which conditions the process Y is also a homogeneous Markov process. In the continuous time case, the homogeneous Markov process X is given by its transition rate matrix (infinitesimal generator) A, in which we define $A(i, i) \stackrel{\text{def}}{=} -\sum_{j \neq i} A(i, j)$. In the discrete time case, X is given by its transition probability matrix P. In order to deal simultaneously with the continuous and discrete time cases, we will use the notation Q to represent both matrices A or P depending on the context. We shall denote by (α, Q) the Markov process X when its initial distribution is α. We shall denote by $\text{agg}(X)$ (resp. by $\text{agg}(\alpha, Q)$ when its initial distribution is α) the aggregated process constructed from X (resp. from (α, Q)) over the given partition $\mathscr{B}$. Let us denote by $\mathscr{A}$ the set of all probability vectors with N entries and by $\mathscr{A}_{\mathcal{M}}(X)$ the set of all initial probability distributions of X which gives to the aggregated process the Markov homogeneous property. That is,

$$\mathscr{A}_{\mathcal{M}}(X) \stackrel{\text{def}}{=} \{\alpha \in \mathscr{A} \mid Y = \text{agg}(\alpha, Q) \text{ is Markov homogeneous}\}.$$

In [2], it is shown that when X is a discrete time Markov process, there exists a finite algorithm constructing the set $\mathscr{A}_{\mathcal{M}}(X)$. In this companion paper, it is shown how the continuous time case can be reduced to the discrete time one. The used tool is the uniformization technique.

The paper is organized as follows. In the next section, we introduce the problem of weak lumpability in Markov processes, we fix the notation and we prove some preliminary results. Also, the principal results of [2] are recalled. Section 3 contains the main contribution of this work, namely Theorem 3.2 which shows how to reduce the continuous time case to the discrete time one. Conclusions are reported in Section 4.

2. Weak lumpable Markov processes

For every instant $u \in \mathscr{T}$, we denote by $Q(u)$ the N-dimensioned square matrix whose (i, j) entry is $\mathbb{P}(X_u = j / X_0 = i)$. That is, $Q(n) = P^n$ in discrete time, $Q(t) = e^{At}$ in continuous time.

As in [2], we denote by $n(i)$ the cardinal of $B(i)$ and we assume the states of E ordered such that

$$\begin{aligned}
B(1) &= \{1, \ldots, n(1)\},\\
&\vdots\\
B(m) &= \{n(1) + \cdots + n(m-1) + 1, \ldots, n(1) + \cdots + n(m)\},\\
&\vdots\\
B(M) &= \{n(1) + \cdots + n(M-1) + 1, \ldots, N\}.
\end{aligned}$$

Some other notation taken from [2] are the following.

- For any $l \in F$ and $\alpha \in \mathcal{A}$, we denote by $T_l.\alpha$ the vector with $n(l)$ components whose ith entry is $(T_l.\alpha)(i) = \alpha(n(1) + \cdots + n(l-1) + i)$, for $i = 1, 2, \ldots, n(l)$.
- For any real vector $\gamma \geq 0$, we denote by $\|\gamma\|$ the sum of its components.
- For any $l \in F$ and $\alpha \in \mathcal{A}$ such that $T_l.\alpha \neq 0$, we denote by $\alpha^{B(l)}$ the vector of $\mathcal{A}$ defined by $\alpha^{B(l)}(i) = \alpha(i)/\|T_l.\alpha\|$ if i belongs to $B(l)$, 0 otherwise.

The reader can see [2] for some examples. To simplify the notation, each time that in the sequel we shall write a vector of the form $\gamma^{B(m)}$ with $\gamma \in \mathcal{A}$ we implicitely mean that this vector is defined (that is, that $T_m.\gamma \neq 0$).

Let α be an element of $\mathcal{A}$, $(B(m_0), B(m_1), \ldots, B(m_j))$ be a finite sequence of elements of $\mathcal{B}$ and $0 = t_0 < t_1 < \cdots < t_j$ be a sequence of instants (in discrete or continuous time) such that $\mathbb{P}(Z_0 \in B(m_0), Z_{t_1} \in B(m_1), \ldots, Z_{t_j} \in B(m_j)) > 0$. For these two finite sequences, we define the vector $h(\alpha, (B(m_0), 0), (B(m_1), t_1), \ldots, (B(m_j), t_j))$ recursively by

$$h(\alpha, (B(m_0), 0)) = \alpha^{B(m_0)}$$

and for $1 \leq k \leq j$,

$$h(\alpha, (B(m_0), 0), \ldots, (B(m_k), t_k)) = (h(\alpha, (B(m_0), 0), \ldots, (B(m_{k-1}), t_{k-1}))Q(t_k - t_{k-1}))^{B(m_k)}.$$

A first property of the function h that will be needed in the sequel is that the mapping

$$\alpha \mapsto h(\alpha, (B(m_0), 0), \ldots, (B(m_k), t_k))$$

is continuous. The proof is as in [3, Lemma 3.2]. A second technical property of the function h is stated in the following lemma.

Lemma 2.1. *Let $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence of initial probability distributions (that is, a sequence of elements of $\mathcal{A}$). Let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence of non-negative real numbers whose sum is equal to 1. There exists a sequence $(\mu_n)_{n \in \mathbb{N}}$ of non-negative real numbers whose sum is equal to 1 such that*

$$h\left(\sum_{n=0}^{\infty} \lambda_n \alpha_n, (B(m_0), 0), \ldots, (B(m_k), t_k)\right) = \sum_{n=0}^{\infty} \lambda_n \mu_n h(\alpha_n, (B(m_0), 0), \ldots, (B(m_k), t_k)).$$

Proof. Check first that

$$\sum_{n=0}^{\infty} \lambda_n \alpha_n \in \mathcal{A}.$$

Let $\alpha_0, \alpha_1 \in \mathcal{A}$. Let $\lambda_0, \lambda_1 \geq 0$ such that $\lambda_0 + \lambda_1 = 1$. Since

$$(\lambda_0 \alpha_0 + \lambda_1 \alpha_1)^{B(m_0)} = \lambda_0 \mu_0 \alpha_0^{B(m_0)} + \lambda_1 \mu_1 \alpha_1^{B(m_0)},$$

where

$$\mu_r = \frac{\sum_{j\in B(m_0)} \alpha_r(j)}{\lambda_0 \sum_{j\in B(m_0)} \alpha_0(j) + \lambda_1 \sum_{j\in B(m_0)} \alpha_1(j)}, \quad r=0,1,$$

we have

$$\begin{aligned} &h(\lambda_0\alpha_0+\lambda_1\alpha_1, (B(m_0),0), \ldots, (B(m_k),t_k)) \\ &\quad = \lambda_0\mu_0 h(\alpha_0+\lambda_1\alpha_1, (B(m_0),0), \ldots, (B(m_k),t_k)) \\ &\qquad + \lambda_1\mu_1 h(\alpha_1, (B(m_0),0), \ldots, (B(m_k),t_k)). \end{aligned}$$

The result is obviously valid for every finite sum, that is

$$\begin{aligned} &h\left(\sum_{n=0}^{H} \lambda_n\alpha_n, (B(m_0),0), \ldots, (B(m_k),t_k)\right) \\ &\quad = \sum_{n=0}^{H} \lambda_n\mu_n h(\alpha_n, (B(m_0),0), \ldots, (B(m_k),t_k)). \end{aligned} \tag{1}$$

Taking limits when $H \to \infty$, the right-hand side of this equality obviously converges. In the left-hand side, it suffices to verify that the interchange of limits is valid, which follows from the continuity of the mapping $\alpha \mapsto h(\alpha, (B(m_0),0), \ldots, (B(m_k),t_k))$. □

Now, for every $m \in F$, we denote by $\mathcal{M}(\alpha, B(m))$ the set

$$\mathcal{M}(\alpha, B(m)) \stackrel{\text{def}}{=} \{\beta \in \mathcal{A} \mid \text{there exists } j \geq 0 \text{ and a sequence } (B(m_0),0), \ldots, (B(m_j),t_j) \text{ with } m_j = m \text{ such that } \beta = h(\alpha, (B(m_0),0), \ldots, (B(m_j),t_j))\}.$$

The next result is an extended version of [1, Theorem 6.4.1]. We denote here by $\mathbb{P}_\alpha(\ldots)$ the probability of any event concerning the Markov process X when its initial probability distribution is α.

Theorem 2.2. *The process $Y = \text{agg}(\alpha, Q)$ is a homogeneous Markov process if and only if $\forall m, l \in F$ and $\forall t \in \mathcal{T}$, the probability $\mathbb{P}_\beta(X_t \in B(l))$ has the same value for every $\beta \in \mathcal{M}(\alpha, B(m))$. This common value is the probability that the homogeneous Markov process Y is in state l at time t given that it starts in state m.*

Proof. Note that, for every $j \geq 0$, for every sequences $B(m_0), \ldots, B(m_{j+1})$ of elements of $\mathcal{B}$ and $0 = t_0 < t_1 < \cdots < t_j < t_j + t$ of instants, we have the relation

$$\mathbb{P}_\alpha(X_{t_j+t} \in B(m_{j+1}) / X_{t_j} \in B(m_j), \ldots, X_0 \in B(m_0)) = \mathbb{P}_\beta(X_t \in B(m_{j+1})),$$

where

$$\beta = h(\alpha, (B(m_0),0), (B(m_1),t_1), \ldots, (B(m_j),t_j)).$$

If the condition of the theorem is satisfied, the previous relation depends only on m_j, m_{j+1} and t. This implies that the process Y is a homogeneous Markov process.

Conversely, assume that the process Y is a homogeneous Markov process. Let β be any vector of $\mathcal{M}(\alpha, B(m))$. The vector β can then be written as

$$\beta = h(\alpha, (B(m_0), 0), (B(m_1), t_1), \ldots, (B(m), t_j)).$$

Using the previous relation, we have

$$\mathbb{P}_\beta(X_t \in B(l)) = \mathbb{P}_\alpha(X_{t_j+t} \in B(l) / X_{t_j} \in B(m)).$$

This last quantity does not depend on t_j since Y is homogeneous. Therefore $\mathbb{P}_\beta(X_t \in B(l))$ has the same value for every $\beta \in \mathcal{M}(\alpha, B(m))$. □

Following [1, 3, 2], the process X is *weakly lumpable* with respect to $\mathcal{B}$ iff $\mathcal{A}_{\mathcal{U}} \neq \emptyset$. In the particular case of $\mathcal{A}_{\mathcal{U}} = \mathcal{A}$, it is *strongly lumpable* or simply *lumpable* with respect to the given partition. As a corollary of Theorem 2.2, it is easy to verify that the process X is strongly lumpable with respect to $\mathcal{B}$ iff the following proposition holds:

$$\text{for all } l, m \in F, \text{ for all } i \in B(l), \quad \sum_{j \in B(m)} Q(i, j) \text{ does not depend on } i.$$

The proof is as in [1, Theorem 6.3.2] where only the case of discrete time processes is considered.

Corollary 2.3. *If $\mathcal{A}_{\mathcal{U}}(X) \neq \emptyset$ and if $\hat{Q}$ denotes the transition matrix of the aggregated homogeneous Markov process $Y = \text{agg}(\alpha, Q)$ then $\hat{Q}$ is the same for every α leading to an aggregated homogeneous Markov process. Moreover, $\pi \in \mathcal{A}_{\mathcal{U}}(X)$.*

Proof. In the discrete time case, a proof can be found in [3]. In the continuous time case, we have $Q = A$ and $Q(u) = e^{Au}$, $u \geq 0$. Let $\alpha \in \mathcal{A}_{\mathcal{U}}(X)$. We can write for every l, $m \in F$ and for every s, $t \geq 0$,

$$e^{\hat{A}t}(l, m) = \mathbb{P}_\alpha(X_{t+s} \in B(m) / X_s \in B(l)) = \mathbb{P}_{\alpha e^{As}}(X_t \in B(m) / X_0 \in B(l)).$$

Letting now s go to infinity, we have by a continuity argument as in [3, Lemma 3.2],

$$e^{\hat{A}t}(l, m) = \mathbb{P}_\pi(X_t \in B(m) / X_0 \in B(l)) \quad \forall t \geq 0,$$

which does not depend on α. This means that $\hat{A}$ does not depend itself on α. □

For $m \in F$, we denote by $\tilde{Q}_m$ the $n(m) \times M$ matrix with entries

$$\tilde{Q}_m(i, l) = \sum_{j \in B(l)} Q(n(1) + \cdots + n(m-1) + i, j), \quad 1 \leq i \leq n(m), \ l \in F.$$

From Corollary 2.3, we deduce easily the relation $\hat{Q}_m = (T_m.\pi^{B(m)})\tilde{Q}_m$, where $\hat{Q}_m$ is the mth row of $\hat{Q}$. In the sequel, we will always define $\hat{Q}$ in this way, even if the aggregated process Y is not a homogeneous Markov process. That is, if Y is a homogeneous Markov process then $\hat{Q}$ is its transition matrix, otherwise, $\hat{Q}_m \stackrel{\text{def}}{=} (T_m.\pi^{B(m)})\tilde{Q}_m \ \forall m \in F$.

As in [2], we define the following sets, for every $j \geq 0$:

$$\mathcal{M}^{j+1}(X) \stackrel{\text{def}}{=} \{\alpha \in \mathcal{A} \mid \text{for all } \beta = h(\alpha, (B(m_0), 0), \ldots, (B(m_k), t_k)) \text{ where } 0 \leq k \leq j,\ T_{m_k}.\beta\tilde{Q}_{m_k} = \hat{Q}_{m_k}\}.$$

The next lemma allows a recurrent construction of the sequence $(\mathcal{M}^j(X))_{j\geq 1}$.

Lemma 2.4. *For every $j \geq 1$, we have*

$$\mathcal{M}^{j+1}(X) = \{\alpha \in \mathcal{M}^j(X) \mid \forall t,\ \alpha^{B(l)}Q(t) \in \mathcal{M}^j(X) \text{ for every } l \in F\}.$$

Proof. The proof is simply based upon the following property of the function h:

$$h(\alpha^{B(l)}Q(t), (B(m_0), 0), (B(m_1), t_1), \ldots, (B(m_k), t_k))$$
$$= h(\alpha, (B(l), 0), (B(m_0), t), (B(m_1), t_1 + t), \ldots, (B(m_k), t_k + t)).$$

Let $j \geq 1$. By definition of $\mathcal{M}^{j+1}(X)$, $\alpha \in \mathcal{M}^{j+1}(X)$ is equivalent to the following proposition: for all $\beta = h(\alpha, (B(l), 0), (B(m_0), t), (B(m_1), t_1 + t), \ldots, (B(m_k), t_k + t))$, we have

$$T_{m_k}.\beta\tilde{Q}_{m_k} = \hat{Q}_{m_k}, \quad 0 \leq k \leq j - 1.$$

Using the previous property of the function h, this last equality can be written

$$T_{m_k}.h(\alpha^{B(l)}Q(t), (B(m_0), 0), (B(m_1), t_1), \ldots, (B(m_k), t_k))\tilde{Q}_{m_k} = \hat{Q}_{m_k},$$
$$0 \leq k \leq j - 1,$$

which is equivalent to $\alpha^{B(l)}Q(t) \in \mathcal{M}^j(X)$. □

Theorem 2.5.

$$\mathcal{A}_{\mathcal{M}}(X) = \bigcap_{j \geq 1} \mathcal{M}^j(X).$$

Proof. The proof is as in [3, Theorem 3.7]. It is based on the equivalence

$$\alpha \in \mathcal{A}_{\mathcal{M}}(X) \iff \text{for all } m \in F \text{ and } \beta \in \mathcal{M}(\alpha, B(m)),\ T_m.\beta\tilde{Q}_m = \hat{Q}_m,$$

which follows directly from Theorem 2.2.

Now, from the definition of $\mathcal{M}(\alpha, B(m))$, the right-hand side of the equivalence can be rewritten as follows:

$$[\text{for all } \beta = h(\alpha, (B(m_0), 0), (B(m_1), t_1), \ldots, (B(m_k), t_k)),$$
$$\text{we have } T_{m_k}.\beta\tilde{Q}_{m_k} = \hat{Q}_{m_k}],$$

which is equivalent to

$$[\text{for all } j \geqslant 0, \text{ for all } \beta = h(\alpha, (B(m_0), 0), (B(m_1), t_1), \ldots, (B(m_k), t_k))$$
$$\text{where } 0 \leqslant k \leqslant j, \text{ we have } T_{m_k}.\beta\tilde{Q}_{m_k} = \hat{Q}_{m_k}].$$

By definition of $\mathcal{M}^{j+1}(X)$, this last proposition is in turn equivalent to

$$\text{for all } j \geqslant 0, \quad \alpha \in \mathcal{M}^{j+1}(X),$$

and the result follows. □

3. Uniformization and relations between the discrete time case and the continuous time case

Let us consider the continuous time process $X = (X_t)_{t \geqslant 0}$ with transition rate matrix A. The uniformization technique consists of constructing an auxiliary homogeneous discrete time Markov chain $U = (U_n)_{n \in \mathbb{N}}$ on the same state space E, with transition probability matrix P, and a Poisson process $(N(t))_{t \geqslant 0}$ with rate λ, independent of U, such that the two processes (X_t) and $(U_{N(t)})$ are equivalent. The construction is as follows:

- we choose $\lambda \in \mathbb{R}$ such that $\lambda \geqslant \max\{-A(i, i), i = 1, \ldots, N\}$;
- we define

$$P(U_{n+1} = j / U_n = i) = P(i, j) = \begin{cases} 1 + A(i, i)/\lambda & \text{if } i = j, \\ A(i, j)/\lambda & \text{otherwise.} \end{cases}$$

Between the matrices A and P the following relation holds.

$$P = I + \frac{1}{\lambda} A, \tag{2}$$

where I is the $N \times N$ identity matrix. We shall denote also $U = \text{unif}(X)$. Observe that the two processes X and U have the same stationary distribution denoted by π. We denote by V the aggregated process of the Markov chain U with respect to the partition $\mathcal{B}$. The process V with values on F is defined by

$$V_n = m \overset{\text{def}}{\Longleftrightarrow} U_n \in B(m) \quad \text{for all } n \geqslant 0.$$

All the results described in the previous section are applicable to the processes U, V, the matrix $Q(u)$ being replaced by P^u for every $u \in \mathbb{N}$.

Observe that we can always consider the following scheme. Given the family of homogeneous and irreducible Markov processes X sharing the same transition rate matrix A, we construct the family of homogeneous and irreducible Markov chains U with common transition probability matrix P defined by (2). We also define the family Y of aggregated processes constructed from X with respect to $\mathscr{B}$ and the family V of aggregated processes constructed from U with respect to the same partition. Each $\alpha \in \mathscr{A}$ fixes an element of each family of processes X, Y, U, V.

$$\begin{array}{ccc} X & \xrightarrow{\text{aggregation}} & Y \\ {\scriptstyle\text{uniformization}}\downarrow & & \\ U & \xrightarrow[\text{aggregation}]{} & V \end{array}$$

In the next theorem, we show that if Y is Markov homogeneous, then V is also Markov homogeneous. Moreover, in this case, V is the uniformized process constructed from Y with respect to the same uniformization rate λ.

As before, we define the matrices $\hat{A}$ and $\hat{P}$, independently of the fact that Y and V are Markov or not. Let us also define for every $m \in F$, the $n(m) \times M$ matrix $\tilde{I}_m$ by

$$\tilde{I}_m(i, l) = \begin{cases} 1 & \text{if } l = m, \\ 0 & \text{otherwise,} \end{cases}$$

and the M dimensional row vector e_m by

$$e_m(l) = \begin{cases} 1 & \text{if } l = m, \\ 0 & \text{otherwise.} \end{cases}$$

Let us denote by $\hat{I}$ the $M \times M$ identity matrix. We can immediately check the following relation between matrices $\hat{P}$ and $\hat{A}$ which means, in words, that if Y and V are Markov homogeneous, then $V = \text{unif}(Y)$.

Lemma 3.1. *Between the matrices $\hat{A}$ and $\hat{P}$ the following relation holds:*

$$\hat{P} = \hat{I} + \hat{A}/\lambda.$$

Proof. Let $m \in F$. Check first that, from (2),

$$\tilde{P}_m = \tilde{I}_m + \tilde{A}_m/\lambda.$$

Then,

$$\hat{A}_m = T_m.\pi^{B(m)}\tilde{A}_m = T_m.\pi^{B(m)}\lambda(\tilde{P}_m - \tilde{I}_m) = \lambda(\hat{P}_m - T_m.\pi^{B(m)}\tilde{I}_m),$$

and from the definition of $\tilde{I}_m$ we have

$$T_m.\pi^{B(m)}\tilde{I}_m = e_m,$$

which gives $\hat{A}_m = \lambda(\hat{P}_m - e_m)$, that is $\hat{A} = -\lambda(\hat{I} - \hat{P})$. $\square$

We are now able to prove the main result of the paper.

Theorem 3.2. *The process X is weakly lumpable with respect to $\mathcal{B}$ iff the uniformized process* unif(X) *is weakly lumpable with respect to $\mathcal{B}$. In this case, we have*

- $\mathcal{A}_{\mathcal{M}}(X) = \mathcal{A}_{\mathcal{M}}(\mathrm{unif}(X))$.
- $V = \mathrm{unif}(Y)$, *that is,* $\mathrm{agg}(\mathrm{unif}(X)) = \mathrm{unif}(\mathrm{agg}(X))$.

Proof. We prove the more precise following result:

$$\text{for all } j \geq 1, \quad \mathcal{M}^j(X) = \mathcal{M}^j(\mathrm{unif}(X)). \tag{3}$$

From Theorem 2.5, relation (3) implies that $\mathcal{A}_{\mathcal{M}}(X) = \mathcal{A}_{\mathcal{M}}(\mathrm{unif}(X))$ which in turn proves the enounced equivalence. The proof of (3) is by induction. For $j = 1$,

$$\mathcal{M}^1(X) = \{\alpha \in \mathcal{A} \mid \text{for all } \beta = \alpha^{B(l)},\ T_l.\beta\tilde{A}_l = \hat{A}_l\}$$

and

$$\mathcal{M}^1(\mathrm{unif}(X)) = \{\alpha \in \mathcal{A} \mid \text{for all } \beta = \alpha^{B(l)},\ T_l.\beta\tilde{P}_l = \hat{P}_l\}.$$

Now,

$$\begin{aligned}
\alpha \in \mathcal{M}^1(\mathrm{unif}(X)) \iff & \forall l \in F,\ T_l.\alpha^{B(l)}\tilde{P}_l = \hat{P}_l \\
\iff & \forall l \in F,\ T_l.\alpha^{B(l)}(\tilde{I}_l + \tilde{A}_l/\lambda) = e_l + \hat{A}_l/\lambda \\
& \text{(Lemma 3.1)} \\
\iff & \forall l \in F,\ T_l.\alpha^{B(l)}\tilde{A}_l = \hat{A}_l \\
\iff & \alpha \in \mathcal{M}^1(X).
\end{aligned}$$

Assume that $\mathcal{M}^k(X) = \mathcal{M}^k(\mathrm{unif}(X))$ for all $k \leq j$. We have

$$\begin{aligned}
\mathcal{M}^{j+1}(X) &= \{\alpha \in \mathcal{M}^j(X) \mid \forall t \geq 0,\ \alpha^{B(l)}\, \mathrm{e}^{At} \in \mathcal{M}^j(X) \text{ for every } l \in F\} \\
&= \{\alpha \in \mathcal{M}^j(\mathrm{unif}(X)) \mid \forall t \geq 0,\ \alpha^{B(l)}\, \mathrm{e}^{At} \in \mathcal{M}^j(\mathrm{unif}(X)) \\
&\qquad \text{for every } l \in F\}.
\end{aligned} \tag{4}$$

As

$$\forall t \geq 0, \quad \mathrm{e}^{At} = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!} P^n,$$

we have

$$\forall t \geq 0,\ \forall l \in F, \quad \alpha^{B(l)}\, \mathrm{e}^{At} = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!} \alpha^{B(l)} P^n.$$

We can then write

$$\alpha^{B(l)} e^{At} \in \mathcal{M}^j(\text{unif}(X))$$

$$\iff \forall 0 \leq k \leq j-1,\ \forall (B(m_0), 0), \ldots, (B(m_k), n_k),$$

$$T_{m_k}.h(\alpha^{B(l)} e^{At}, (B(m_0), 0), \ldots, (B(m_k), n_k)) \tilde{P}_{m_k} = \hat{P}_{m_k}.$$

But, from Lemma 2.1,

$$h(\alpha^{B(l)} e^{At}, (B(m_0), 0), \ldots, (B(m_k), n_k))$$

$$= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mu_n(t) h(\alpha^{B(l)} P^n, (B(m_0), 0), \ldots, (B(m_k), n_k)).$$

Finally, we obtain

$$\alpha^{B(l)} e^{At} \in \mathcal{M}^j(\text{unif}(X))$$

$$\iff \forall 0 \leq k \leq j-1,\ \forall (B(m_0), 0), \ldots, (B(m_k), n_k),$$

$$\sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mu_n(t)$$

$$\times T_{m_k}.h(\alpha^{B(l)} P^n, (B(m_0), 0), \ldots, (B(m_k), n_k)) \tilde{P}_{m_k} = \hat{P}_{m_k}$$

$$\iff \forall 0 \leq k \leq j-1,\ \forall (B(m_0), 0), \ldots, (B(m_k), n_k),$$

$$\sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mu_n(t)$$

$$\times (T_{m_k}.h(\alpha^{B(l)} P^n, (B(m_0), 0), \ldots, (B(m_k), n_k)) \tilde{P}_{m_k} - \hat{P}_{m_k}) = 0$$

$$\left(\text{since } \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mu_n(t) = 1\right).$$

This last equality must be true for every $t \geq 0$, so

$$\alpha^{B(l)} e^{At} \in \mathcal{M}^j(\text{unif}(X))$$

$$\iff \forall 0 \leq k \leq j-1,\ \forall (B(m_0), 0), \ldots, (B(m_k), n_k),\ \forall n \geq 0,$$

$$T_{m_k}.h(\alpha^{B(l)} P^n, (B(m_0), 0), \ldots, (B(m_k), n_k)) \tilde{P}_{m_k} = \hat{P}_{m_k}$$

$$\iff \text{for all } n \geq 0,$$

$$\alpha^{B(l)} P^n \in \mathcal{M}^j(\text{unif}(X))$$

$$\iff \alpha \in \mathcal{M}^{j+1}(\text{unif}(X)).$$

Together with (4), this last equivalence shows that $\mathcal{M}^{j+1}(\text{unif}(X)) = \mathcal{M}^{j+1}(X)$. The second part of the theorem is just Lemma 3.1. □

4. Conclusions

In this paper we analyse the set of all initial probability distributions of an irreducible and homogeneous Markov process which lead to a homogeneous aggregated Markov process, given the transition rate matrix and a partition of the state space. The main result is the reduction to the discrete time case, which is analysed in [2], by means of the uniformization technique. As stated in the companion paper [2], the first possible direction to extend these results seems to be the case of absorbing Markov processes.

References

[1] J.G. Kemeny and J.L. Snell, Finite Markov Chains (Springer, New York, 1976).

[2] G. Rubino and B. Sericola, A finite characterization of weak lumpable Markov processes. Part I: the discrete time case, Stochastic Process. Appl. 38 (1991) 195-204.

[3] G. Rubino and B. Sericola, On weak lumpability in Markov chains, J. Appl. Probab. 26 (1989) 446-457.