



HAL
open science

On weak lumpability of denumerable Markov chains

James Ledoux

► **To cite this version:**

James Ledoux. On weak lumpability of denumerable Markov chains. *Statistics and Probability Letters*, 1995, 25, pp.329-339. hal-00852300

HAL Id: hal-00852300

<https://hal.science/hal-00852300>

Submitted on 20 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On Weak Lumpability of Denumerable Markov Chains

James Ledoux*

16 December 1994

Abstract

We consider weak lumpability of denumerable Markov chains evolving in discrete or continuous time. Specifically, we study the properties of the set of all initial distributions of the starting chain leading to an aggregated homogeneous Markov chain with respect to a partition of the state space.

Keywords: Weak Lumpability, Positive recurrence, R -positivity, Quasi-stationary distribution, Uniform semi-group.

1 Introduction

Let us consider a homogeneous Markov chain X , in discrete or continuous time, on a countably infinite state space denoted by E , which without loss of generality we assume to be a subset of the natural numbers \mathbb{N} (i.e. $E \subseteq \mathbb{N}$.) Let $\mathcal{B} = \{B(0), B(1), \dots\}$ be a fixed partition of E . We associate with the given chain X the aggregated chain Y , over the state space $F = \{0, 1, \dots\}$, defined by:

$$Y_t = l \iff X_t \in B(l), \text{ for any } t.$$

We are interested in the set of all initial distributions of X which give an aggregated homogeneous Markov chain Y . If this set is not empty, we say that the family of Markov chains sharing the same transition semi-group is *weakly lumpable*. Most of the literature on lumpability has been concerned with the *strong lumpability* situation, that is, when any initial distribution leads to an aggregated homogeneous Markov chain. To the best of my knowledge, the weak lumpability problem with countably infinite state space has been addressed only recently in Ball and Yeo (1993) for (irreducible positive-recurrent) continuous time Markov chains. The purpose of this note is to propose some results in discrete or continuous time, prolonging the studies reported in Rubino and Sericola (1989,1991,1993) and Ledoux and al. (1994) for a finite state space. Section 2 deals with discrete time Markov chains and mainly concerns weak lumpability for irreducible positive-recurrent or R -positive chains. In particular, we discuss the ergodic interpretation of the quasi-stationary distribution. The third section shows that lumpability for any denumerable continuous time Markov chains with an uniform transition semi-group can always be replaced in the discrete time context. The sequel of this result are also discussed for irreducible positive-recurrent or λ -positive continuous time Markov chains.

By convention, vectors are row vectors. Column vectors are indicated by means of the transpose operator $(.)^*$. The vector with all its components equal to 1 (resp. 0) is denoted merely by 1 (resp. 0). The set of all probability distributions on E will be denoted by \mathcal{A} . For any subset

*IRISA, Campus de Beaulieu 35042 Rennes Cedex

This work was partially supported by the Regional Council of Brittany under Grant 290C2010031305061.

B of E and $\alpha \in \mathcal{A}$, the restriction of α to B , i.e. the vector $(\alpha(i), i \in B)$, is denoted by α_B ; if $\alpha_B 1^* \neq 0$, α^B is the vector defined by $\alpha^B(i) = \alpha(i) / \sum_{j \in B} \alpha(j)$ if $i \in B$ and by 0 if $i \notin B$.

2 Weak lumpability in discrete time

Let $X = (X_n)_{n \geq 0}$ be a homogeneous Markov chain over state space E , given by its transition probability matrix $P = (P(i, j))_{i, j \in E}$ and its initial distribution α ; when necessary we denote it by (α, P) . Let $P(i, B)$ denote the transition probability of moving in one step from state i to the subset B of E , that is $P(i, B) = \sum_{j \in B} P(i, j)$. We denote the aggregated chain constructed from (α, P) with respect to the partition \mathcal{B} by $agg(\alpha, P, \mathcal{B})$.

Definition 2.1 *A sequence (C_0, C_1, \dots, C_j) of subsets of E is called possible for the initial distribution α iff $\mathbb{P}_\alpha(X_0 \in C_0, X_1 \in C_1, \dots, X_j \in C_j) > 0$. Given any distribution $\alpha \in \mathcal{A}$ and a possible sequence (C_0, C_1, \dots, C_j) for α , we can define the vector $f(\alpha, C_0, C_1, \dots, C_j) \in \mathcal{A}$ recursively by:*

$$\begin{aligned} f(\alpha, C) &= \alpha^C \\ f(\alpha, C_0, C_1, \dots, C_k) &= (f(\alpha, C_0, C_1, \dots, C_{k-1})P)^{C_k}. \end{aligned}$$

For any $B \in \mathcal{B}$, $\mathcal{A}(\alpha, B)$ denotes the subset of all distributions of the form $f(\alpha, C_1, \dots, C_j, B)$.

By definition, the aggregated chain $Y = agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain if and only if $\forall l, m \in F, \forall n \geq 0$ and $\forall (C_0, C_1, \dots, C_{n-1}, B(l))$ possible for α ,

$$\mathbb{P}_\alpha(X_{n+1} \in B(m) \mid X_n \in B(l), X_{n-1} \in C_{n-1}, \dots, X_0 \in C_0) = \mathbb{P}_\alpha(X_{n+1} \in B(m) \mid X_n \in B(l))$$

and the probability in the right-hand side does not depend on n ; in that case, it describes the probability of going from state l to state m in one step for the aggregated chain $agg(\alpha, P, \mathcal{B})$. The approach developed in Kemeny and Snell (1976) and in Rubino and Sericola (1989) consists in rewriting the above conditional expression as

$$\mathbb{P}_\beta(X_1 \in B(m)) \quad \text{with } \beta = f(\alpha, C_0, \dots, B(l)).$$

that is, in including the past into the initial distribution. In the same way as in Kemeny and Snell (1976), a necessary and sufficient condition for Y to be a homogeneous Markov chain can be exhibited without any particular assumption on X .

Theorem 2.2 *The chain $Y = agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain iff $\forall l, m \in F$, the probability $\mathbb{P}_\beta(X_1 \in B(m))$ is the same for every $\beta \in \mathcal{A}(\alpha, B(l))$. This common value is the transition probability for the chain Y to move from state l to state m .*

The aim of this section is to study the properties of the set of distributions

$$\mathcal{A}_M = \{\alpha \in \mathcal{A} / agg(\alpha, P, \mathcal{B}) \text{ is a homogeneous Markov chain } \}.$$

2.1 Weak lumpability for irreducible positive-recurrent Markov chains

Throughout this subsection, we assume that the considered Markov chain is irreducible positive-recurrent. Therefore, there exists a unique probability vector, denoted by π , which satisfies $\pi P = \pi$. Let g be a real function on E and m a probability measure on E ; g is m -integrable if

$$m(|g|) \triangleq \sum_{i \in E} m(i) |g(i)| = E_m[|g|] < \infty.$$

For such a Markov chain, we have the following standard corollary of the ergodic theorem.

Result 2.3 For any bounded real function g on E , we have for all $\alpha \in \mathcal{A}$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}_\alpha[g(X_k)] = \pi(g).$$

We only need the following lemma to derive Theorem 2.5 from Theorem 2.2 with similar arguments as for Theorem 3.5 in Rubino and Sericola (1989).

Lemma 2.4 Let β_n be the vector $(1/n) \sum_{k=1}^n \alpha P^k$. For any $l, m \in F$, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} f(\beta_n, B(l)) &= f(\pi, B(l)), \\ \lim_{n \rightarrow \infty} \mathbb{P}_{f(\beta_n, B(l))}(X_1 \in B(m)) &= \mathbb{P}_{f(\pi, B(l))}(X_1 \in B(m)). \end{aligned}$$

Proof. To obtain the first limit, it suffices to let respectively $\forall i \in E : g_j(i) = 1_{\{j\}}(i)$ with any $j \in B(l)$ and $g(i) = 1_{B(l)}(i)$ in Result 2.3. Since $f(\beta_n, B(l))(j) = \sum_{k=1}^n \mathbb{P}_\alpha(X_k = j) / \sum_{k=1}^n \mathbb{P}_\alpha(X_k \in B(l))$, the numerator and the denominator tend respectively to $\pi(g_j) = \pi(j)$ and to $\pi(g) = \pi_{B(l)} 1^*$.

The second limit is derived from the previous considerations and from Result 2.3 letting $g(i) = 1_{B(l)}(i)P(i, B(m))$, for $i \in E$. Indeed, we can write

$$\mathbb{P}_{f(\beta_n, B(l))}(X_1 \in B(m)) = \frac{1}{\sum_{i \in B(l)} \beta_n(i)} \sum_{i \in B(l)} \beta_n(i) P(i, B(m)), \quad (1)$$

and the two factors in the right-hand side of formula (1) tend respectively to $1/\pi_{B(l)} 1^*$ and $\sum_{i \in B(l)} \pi(i) P(i, B(m))$. \diamond

Finally, we have

Theorem 2.5 If $\mathcal{A}_M \neq \emptyset$, then $\pi \in \mathcal{A}_M$ and the transition probability matrix of the homogeneous Markov chain $\text{agg}(\alpha, P, \mathcal{B})$, denoted by \widehat{P} , is the same for all $\alpha \in \mathcal{A}_M$. The entries of matrix \widehat{P} are given by

$$\widehat{P}(l, m) = \sum_{i \in B(l)} \pi^{B(l)}(i) P(i, B(m)).$$

The unicity of matrix \widehat{P} for all $\alpha \in \mathcal{A}_M$ gives the convex property to the set \mathcal{A}_M . In particular, if we construct the convex envelope of the family of vectors $\{\pi^{B(l)}, l \in F\}$, $\mathcal{A}_\pi = \sum_{l \in F} \lambda_l \pi^{B(l)}$ (with $\lambda_l \geq 0$ and $\sum_{l \in F} \lambda_l = 1$) and $\mathcal{A}_M \neq \emptyset$, then we have $\mathcal{A}_\pi \subseteq \mathcal{A}_M$. With the previous result, Theorem 3.7 from Rubino and Sericola (1989) can be extended to our denumerable context. Consequently, the set \mathcal{A}_M is the (a priori) infinite intersection of a decreasing sequence of convex sets, denoted by \mathcal{A}^j ($j \geq 1$), which are the solutions to the linear systems defined as in Rubino and Sericola (1989,1991). It can be noted, as in Rubino and Sericola (1991), that the property of P -stability of \mathcal{A}^j (i.e. $\mathcal{A}^j P \subseteq \mathcal{A}^j$) allows us to identify \mathcal{A}_M as the set \mathcal{A}^j . The example of Subsection 2.4 shows that the infinite intersection of \mathcal{A}^j 's can be finite and explicitly computed.

2.2 Weak lumpability of R -positive Markov chains

We are now concerned with denumerable Markov chains with absorbing states which are assumed to be collapsed in only one class (state labeled by 0 for the aggregated process Y) of the partition \mathcal{B} . The other classes constitute a partition of the set of transient states, denoted by T , of X . It is easy to convince oneself that weak lumpability for such a Markov chain reduces to weak lumpability of the Markov chain with only one absorbing state and absorption probabilities equal

to $P(i, B(0))$ for $i \in E$. Consequently, we consider only one absorbing state denoted by a (and $B(0) = \{a\}$.) Let us denote by $\mathcal{A}_{\mathcal{M}}^T$ (resp. \mathcal{A}^T) the subset of $\mathcal{A}_{\mathcal{M}}$ (resp. \mathcal{A}) composed by the distributions α with support T , i.e. $\sum_{i \in T} \alpha(i) = 1$. We have $\mathcal{A}_{\mathcal{M}} = (1 - \lambda_T) 1^{\{a\}} + \lambda_T \mathcal{A}_{\mathcal{M}}^T$ where $1 \geq \lambda_T \geq 0$. Therefore, we restrict the analysis to the set $\mathcal{A}_{\mathcal{M}}^T$.

In discrete time, the transition probability matrix P can be decomposed as follows:

$$P = \left(\begin{array}{c|c} 1 & 0 \\ \hline (I - Q)1^* & Q \end{array} \right),$$

where matrix Q is assumed to be irreducible. In this subsection, we recall (e.g. see Seneta (1981, Chapter VI) the definitions and the main properties of the R -classification of a non-negative irreducible matrix. It can be shown that all the power series $Q_{ij}(z) = \sum_{k=0}^{\infty} Q^k(i, j)z^k$, $i, j \in E$ have a common convergence radius, denoted by R , which is usually called the *convergence parameter* of matrix Q . If E is a finite set, then R is the inverse of the spectral radius of Q . Matrix Q is said to be *R -recurrent* if and only if all the series $\sum_k Q^k(i, j)R^k$ are divergent. Furthermore, if no sequence $(Q^k(i, j)R^k)_{k \geq 0}$ tends to 0, then the matrix is said to be *R -positive*. For an R -recurrent matrix Q , there exists an unique (up to a constant) *R -invariant measure*, (resp. *R -invariant vector*) denoted by v (resp. w), that is

$$R v P = v \quad (\text{resp. } R P w^* = w^*).$$

We can now define the **stochastic** matrix \bar{P} whose entries are given by

$$\bar{P}(i, j) \triangleq R \frac{w(j)}{w(i)} Q(i, j), \quad i, j \in E.$$

Denoting the diagonal matrix with generic diagonal entry $w(i)$ by W , the previous relation becomes

$$\bar{P} = R W^{-1} Q W. \quad (2)$$

It is easy to show (as in Seneta (1981, Th 6.4)) that matrix \bar{P} is positive-recurrent if and only if matrix Q is R -positive. The stationary probability vector of \bar{P} is $\pi = (v(i)w(i))_{i \in E}$ which gives a second characterization of the positive recurrence of \bar{P} : $\sum_i v_i w_i < \infty$. It is important to note that the R -recurrence property does not allow in any way to infer the convergence of the series $\sum_k v_k$ or $\sum_k w_k$. It was shown in Ledoux and al. (1994) that using quasi-stationary distribution can be fruitful for weak lumpability of a finite absorbing Markov chain. We propose in this subsection to extent some of those ideas to a R -positive Markov chain.

A *quasi-stationary distribution* is a probability measure which makes stationary the following conditional probabilities : for all $i \in T$, $\mathbb{P}_{\alpha}(X_n = i | X_n \in T)$, that is the vector $(\mathbb{P}_{\alpha}(X_n = i | X_n \in T))_{i \in T}$ is independent of n . The existence of such a measure under milder conditions than R -recurrence is discussed in many recent papers. But it can be seen that R -recurrence is nearly a “minimal” assumption (up to Harrys Veech conditions, e.g. see Pruitt (1964).) The R -positivity property of matrix Q is also the nearly “minimal” condition to have an ergodic interpretation of such a quasi-stationary distribution with **any** probability vector as initial distribution of the Markov chain X . Moreover the results must include the finite state space ones reported in Ledoux and al. (1994). The following theorem gives an ergodic interpretation to the R -invariant measure v when it defines a probability distribution. Note that we don’t make any distinction between periodic and aperiodic cases. Throughout the remainder of this subsection, we will assume that any initial distribution $\alpha \in \mathcal{A}^T$ satisfies a constraint of the type:

$$\alpha_T \leq C_{\alpha} v \quad (3)$$

where C_α is a positive scalar. It allows to consider $\alpha_T W$ as a summable series because $0 < \alpha_T W 1^* \leq C_\alpha v W 1^* = C_\alpha \pi 1^* = C_\alpha$. Therefore the vector $(\alpha_T W)/(\alpha_T W 1^*)$ defines a probability distribution.

Lemma 2.6 *Let g be a non-negative function on E assumed to be π -integrable. For any initial distribution $\alpha \in \mathcal{A}^T$ with $\alpha_T \leq C_\alpha v$, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (\alpha_T W \bar{P}^k)(g) = \alpha_T W 1^* \pi(g).$$

Proof. From Result 2.3, we have that for any $i \in T$,

$$\lim_{n \rightarrow +\infty} \alpha_T W 1^* \frac{1}{n} \sum_{k=1}^n \left(\frac{\alpha_T W}{\alpha_T W 1^*} \bar{P}^k \right)(i) g(i) = \alpha_T W 1^* \pi(i) g(i).$$

Moreover, condition (3) required on α gives the following inequality for any $i \in T$:

$$\left(\alpha_T W \frac{1}{n} \sum_{k=1}^n \bar{P}^k \right)(i) g(i) \leq C_\alpha \left(\frac{1}{n} \sum_{k=1}^n \pi \bar{P}^k \right)(i) g(i) = C_\alpha \pi(i) g(i).$$

Since $\sum_{i \in T} \pi(i) g(i) = \pi(g) < \infty$, the dominated convergence theorem allows us to write

$$\lim_{n \rightarrow \infty} \sum_{i \in T} \left(\alpha_T W \frac{1}{n} \sum_{k=1}^n \bar{P}^k \right)(i) g(i) = \sum_{i \in T} \alpha_T W 1^* \pi(i) g(i) = \alpha_T W 1^* \pi(g).$$

◇

Theorem 2.7 *Let Q be a R -positive matrix such that its R -invariant measure v satisfies $v 1^* < \infty$. Assume that α is a probability distribution which verifies relation (3). If we define the vector*

$$p_{n,\alpha} = \frac{\sum_{k=1}^n R^k \alpha_T Q^k}{\sum_{k=1}^n R^k \alpha_T Q^k 1^*}, \quad (4)$$

then we have

$$\lim_{n \rightarrow \infty} p_{n,\alpha} = \frac{v}{v 1^*}.$$

This result can also be derived from Seneta and Vere-Jones (1966) but we use here standard arguments on regular Markov chains which give insight into the considered assumptions.

Proof. From definition (2) of matrix \bar{P} , we have

$$p_{n,\alpha} = \frac{(\alpha_T W \sum_{k=1}^n \bar{P}^k) W^{-1}}{(\alpha_T W \sum_{k=1}^n \bar{P}^k) W^{-1} 1^*}.$$

Let $g_j = W^{-1} e_j^* \geq 0$ and $g = W^{-1} 1^* \geq 0$, then we have $\pi(g_j) = v(j)$ and $\pi(g) = v 1^* < \infty$. The Lemma 2.6 allows us to write for all initial distribution such that $\alpha_T \leq C_\alpha v$:

$$\lim_{n \rightarrow \infty} p_{n,\alpha} = \lim_{n \rightarrow \infty} \frac{(\alpha_T W \sum_{k=1}^n \bar{P}^k)(g_j)}{(\alpha_T W \sum_{k=1}^n \bar{P}^k)(g)} = \frac{\alpha_T W 1^* \pi(g_j)}{\alpha_T W 1^* \pi(g)} = \frac{v(j)}{v 1^*}.$$

◇

When the state space E is finite, it is clear that relation (3) is always satisfied and the convergence in Theorem 2.7 holds for any initial distribution. Under the assumptions of Theorem 2.7, we can derive an analogous result to Theorem 2.5. The proof is obtained with similar arguments, that is, firstly establishing the lemma

Lemma 2.8 *For any distribution $\alpha \in \mathcal{A}^T$ satisfying constraint (3), let β_n be the vector*

$$\beta_n = \frac{\sum_{k=1}^n R^k \alpha_T Q^k}{\sum_{k=1}^n R^k \alpha_T Q^k \mathbf{1}^*}.$$

Then, for all $l \neq 0$ and $m \in F$, we have the same conclusions as in Lemma 2.4.

Proof. The first limit is obtained by combining transformation (2) and Lemma 2.6 with, for $i \in E$, $g_j(i) = 1_{\{j\}}(i)/w(j)$ ($j \in B(l)$ is fixed) and $g(i) = 1_{B(l)}(i)/w(i)$. Since $f((0, \beta_n), B(l))(j) = \sum_{k=1}^n R^k \alpha_T Q^k(g) / \sum_{k=1}^n R^k \alpha_T Q^k(g)$, the limit, as n goes to infinity, is the ratio $v(j) / \sum_{i \in B(l)} v(i) = f((0, v), B(l))(j)$.

With the help of the previous limits, the second convergence derives from transformation (2) and Lemma 2.6 with function g defined by: $g(i) = 1_{B(l)}(i)P(i, B(m))/w(i)$, for $i \in E$. Therefore, the two factors in the right-hand side of relation (1) (with the new definition of vector β_n) tend respectively to $v_{B(l)} \mathbf{1}^*$ and to $\sum_{i \in B(l)} v(i) P(i, B(m))$. ◇

The set $\{\alpha \in \mathcal{A}^T / \alpha_T \leq C_\alpha v \text{ and } \text{agg}(\alpha, P, \mathcal{B}) \text{ is a homogeneous Markov chain}\}$ is denoted by $\mathcal{A}_{\mathcal{M}}^T(v)$. We are in position to show the following result:

Theorem 2.9 *Let v be the quasi-stationary distribution associated with the R -positive Markov chain X . If $\mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$ then $(0, v) \in \mathcal{A}_{\mathcal{M}}^T(v)$. Moreover, if \hat{P} denotes the transition probability matrix of the homogeneous Markov chain $\text{agg}(\alpha, P, \mathcal{B})$ then this matrix is the same for all $\alpha \in \mathcal{A}_{\mathcal{M}}^T(v)$.*

Proof. Let $\alpha \in \mathcal{A}^T$ satisfying (3) such that $\text{agg}(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain with transition probability from state l to m denoted by $\hat{P}(l, m)$. Let α_k be the vector

$$\alpha_k = \left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k \mathbf{1}^*} \right).$$

For any k such that $\mathbb{P}_\alpha(X_1 \in B(l)) > 0$ ($l \neq 0$), we have:

$$\begin{aligned} \hat{P}(l, m) &= \mathbb{P}_\alpha(X_{k+1} \in B(m) \mid X_k \in B(l)) \\ &= \mathbb{P}_{(\alpha_T Q^k)_{B(l)}}(X_1 \in B(m)) \\ &= \mathbb{P}_{\alpha_k}_{B(l)}(X_1 \in B(m)). \end{aligned} \tag{5}$$

Choose n_0 large enough such that $\forall n \geq n_0$, $\sum_{k=1}^n (R^k \alpha_T Q^k)_{B(l)} \mathbf{1}^* > 0$ (T is irreducible). The transition probability $\hat{P}(l, m)$ can be rewritten in denoting by γ_k ($k = 1, \dots, n$) the scalar

$$\gamma_k = \frac{(R^k \alpha_T Q^k)_{B(l)} \mathbf{1}^*}{\sum_{k=1}^n (R^k \alpha_T Q^k)_{B(l)} \mathbf{1}^*},$$

$$\begin{aligned}\widehat{P}(l, m) &= \sum_{\substack{1 \leq k \leq n, \\ \mathbb{P}_\alpha(X_k \in B(l)) > 0}} \gamma_k \mathbb{P}_{\alpha_k^{B(l)}}(X_1 \in B(m)) \quad \text{with (5)} \\ &= \mathbb{P}_\Gamma(X_1 \in B(m))\end{aligned}$$

$$\text{where } \Gamma = \sum_{\substack{1 \leq k \leq n, \\ \mathbb{P}_\alpha(X_k \in B(l)) > 0}} \gamma_k (\alpha_k)^{B(l)} = \beta_n^{B(l)}.$$

Therefore, we obtain

$$\widehat{P}(l, m) = \mathbb{P}_{f(\beta_n, B(l))}(X_1 \in B(m)).$$

As n goes to infinity, we derive from Lemma 2.8 that the transition probabilities of the aggregated chain are (independent of α and) given by

$$\widehat{P}(l, m) = \sum_{i \in B(l)} v^{B(l)}(i) P(i, B(m)) \quad l \neq 0, m \in F \quad (6)$$

◇

The convexity of the set $\mathcal{A}_{\mathcal{M}}^T(v)$ follows from the unicity of the transition probability matrix for the aggregated chain.

Corollary 2.10 *If $\mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$ then $\mathcal{A}_{\mathcal{M}}^T(v)$ is a convex set and it necessarily includes the convex subset $\mathcal{A}_v = \sum_{l \in F} \lambda_l v^{B(l)}$ with $\lambda_l \geq 0$ and $\sum_{l \in F} \lambda_l = 1$.*

By definition, the set $\mathcal{A}_{\mathcal{M}}^T(v)$ is a subset of $\mathcal{A}_{\mathcal{M}}^T$. If $\mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$ we trivially have $\mathcal{A}_{\mathcal{M}}^T \neq \emptyset$. The converse is true at least in the following specific cases which respectively ensure that $(0, v) \in \mathcal{A}_{\mathcal{M}}^T$.

Corollary 2.11 *If the set $\mathcal{A}_{\mathcal{M}}^T$ includes a distribution with finite support or if there is a class $B(l)$ ($l \neq 0$) within the partition \mathcal{B} , which collapses a finite number of states, then $(0, v) \in \mathcal{A}_{\mathcal{M}}^T$ and we have: $\mathcal{A}_{\mathcal{M}}^T \neq \emptyset \iff \mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$.*

2.3 Quasi-stationary distribution as a distribution of reset after absorption

In this subsection, we will show that the set $\mathcal{A}_{\mathcal{M}}^T(v)$ is non empty if and only if the set $\mathcal{A}_{\mathcal{M}}$ associated with a positive-recurrent chain is not empty too. Under the condition $v1^* = 1$, where v is the R -invariant measure of the R -positive matrix Q , we can define the following transition probability matrix denoted by $P^{(v)}$:

$$P^{(v)} = \begin{pmatrix} 0 & v \\ (I - Q)1^* & Q \end{pmatrix}.$$

Lemma 2.12 *The Markov chain with transition probability matrix $P^{(v)}$ is irreducible and positive-recurrent. Its invariant probability measure is given by*

$$\pi^{(v)} = \lambda_R 1^{\{a\}} + (1 - \lambda_R)(0, v) \quad \text{with } \lambda_R = (R - 1)/(2R - 1). \quad (7)$$

Proof. The convergence parameter of the R -positive matrix Q is such that $R > 1$ (see Seneta (1981).) Let us consider the ‘‘taboo’’ probability denoted by $f_{00}^{(k)}$ and defined by the probability of going from state 0 to state 0 in k steps without revisiting state 0 in the meantime. The irreducible matrix $P^{(v)}$ will be recurrent if and only if $\sum_{k \geq 1} f_{00}^{(k)} = 1$. Since v is R -invariant, we have $\sum_{k \geq 1} f_{00}^{(k)} = \sum_{k \geq 1} vQ^{k-1}(I - Q)1^* = \sum_{k \geq 1} (1/R)^{k-1}(1 - (1/R)) = 1$. Finally, the positive

recurrence follows from checking that the invariant probability measure of matrix $P^{(v)}$ is given by formula (7). \diamond

We can now show the main result of this subsection.

Theorem 2.13 *If $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ is the set of all initial distributions α such that $\text{agg}(\alpha, P^{(v)}, \mathcal{B})$ is a homogeneous Markov chain, we have $\mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset \iff \mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$; in that case, we have $\mathcal{A}_{\mathcal{M}}^T(v) \subseteq \mathcal{A}_{\mathcal{M}}(P^{(v)}) \subseteq \mathcal{A}_{\mathcal{M}}$. The transition probability matrix $\widehat{P}^{(v)}$ of $\text{agg}(\alpha, P^{(v)}, \mathcal{B})$ is given for every $m \in F$ by $\widehat{P}^{(v)}(l, m) = \widehat{P}(l, m)$ with $l = 1, \dots, M$ and by $\widehat{P}^{(v)}(0, m) = v_{B(m)}1^*$ (matrix \widehat{P} is given by relation (6).)*

Proof. The above one to one correspondence between the respective entries of matrices \widehat{P} and $\widehat{P}^{(v)}$ is deduced from relation (7) in the previous lemma and from the definition of matrix $P^{(v)}$.

The inclusion of $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ in $\mathcal{A}_{\mathcal{M}}$ follows in the same manner as in the finite case (see Ledoux and al. (1994)) and is not reproduced here. We have only to prove that if $\mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$ then $\mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$. Indeed if $\mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$ then $\pi^{(v)} \in \mathcal{A}_{\mathcal{M}}(P^{(v)}) \subseteq \mathcal{A}_{\mathcal{M}}$ from Theorem 2.5. It easily follows that $(\pi^{(v)})^T = (0, v) \in \mathcal{A}_{\mathcal{M}}^T$ and therefore, that $\mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$.

The proposition $(\alpha \in \mathcal{A}_{\mathcal{M}}^T(v)) \implies (\alpha \in \mathcal{A}_{\mathcal{M}}^T(P^{(v)}))$ results directly from the proof of the inclusion $\mathcal{A}_{\mathcal{M}} \subseteq \mathcal{A}_{\mathcal{M}}^T(P^{(v)})$ in the finite case which can be found in Ledoux and al. (1994). \diamond

We have already noted that $\mathcal{A}_{\mathcal{M}} = \lambda 1^{\{a\}} + (1 - \lambda)\mathcal{A}_{\mathcal{M}}^T$ and that, in the finite case, $\mathcal{A}_{\mathcal{M}}^T(v) = \mathcal{A}_{\mathcal{M}}^T$. Therefore, the two sets $\mathcal{A}_{\mathcal{M}}$ and $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ are identical. We are not able to establish the same equality in the denumerable case. Another important fact is that the two sets $\lambda 1^{\{a\}} + (1 - \lambda)\mathcal{A}_{\mathcal{M}}^T(v)$ and $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ are distinct in general (this will be illustrated in the example.) The equality will hold only in the case where any distribution in $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ can be majorized by a multiple of the stationary distribution $\pi^{(v)}$ of $P^{(v)}$.

2.4 Example

Let us consider the following partition $\mathcal{B} = \{B(0) = \{0\}, B(1) = \{i \geq 1\}\}$ of the state space $E = \mathbb{N}$. The transition probability matrix P is given by:

$$\left(\begin{array}{c|cc} P(0,0) = 1 & P(0,1) = 0 & P(0,n) = 0 \text{ for } n \geq 2 \\ \hline P(1,0) = 0 & P(1,n) = (1/6)(5/6)^{n-1} & \text{for } n \geq 1 \\ P(n,0) = 7/8 & P(n,1) = 1/8 & P(k,n) = 0 \\ \text{for any } n \geq 2 & \text{for } n \geq 2 & \text{for } k \geq 2, n \geq 2. \end{array} \right).$$

The submatrix Q of transition probabilities between transient states (here, $T = B(1) = \{i \geq 1\}$) is clearly irreducible. We deduce that $\sum_{k \geq 1} f_{11}^{(k)} z^k = (1/6)z + (5/48)z^2$ (with the same notation as in the proof of Lemma 2.12.) It follows from Seneta (1981, Def 6.2) that state 1 is R -positive with $R = 12/5$ and therefore that all the transient states are R -positive too. The $12/5$ -invariant probability measure v is given by

$$v(1) = \frac{1}{3}, \quad v(2) = \frac{1}{9}, \quad v(n) = \frac{1}{9} \left(\frac{5}{6} \right)^{n-1} \quad \forall n \geq 2.$$

We can directly check that $(0, v) \in \mathcal{A}_{\mathcal{M}}^T$. Indeed, we have only one transient class $B(1)$. The aggregated chain is a homogeneous Markov chain if and only if the distribution of the sojourn times in this class $B(1)$ is geometric with parameter $\widehat{P}(1,1) = 5/12$; this is immediate because vector v is precisely an $12/5$ -invariant measure associated with matrix Q .

Let us consider matrix $P^{(v)}$ which has the following vector as invariant probability measure: $\pi^{(v)} = (7/19)1^{\{a\}} + (12/19)(0, v)$ (relation (7).) We can verify (construct the convex set \mathcal{A}^1 as defined in Rubino and Sericola (1991) and check its P -stability) that

$$\mathcal{A}_{\mathcal{M}}(P^{(v)}) = \{\alpha \in \mathcal{A} / 3\alpha(1) = 1 - \alpha(0)\}.$$

We can note that $\mathcal{A}_{\mathcal{M}}^T(v) \subset \mathcal{A}_{\mathcal{M}}(P^{(v)})$. Indeed, choose $\alpha \in \mathcal{A}^T$ such that

$$\alpha(0) = 0, \quad \alpha(1) = 1/3, \quad \alpha(n) = \frac{2}{33} \left(\frac{11}{12}\right)^{n-1} \quad \forall n \geq 2.$$

Since $3\alpha(1) = 1 - \alpha(0)$, we have $\alpha \in \mathcal{A}_{\mathcal{M}}(P^{(v)})$ but the ratio $\alpha(n)/v(n) \propto (11/10)^{n-1}$ is unbounded as n goes to infinity. Therefore, vector α cannot satisfy relation (3), so $\alpha \notin \mathcal{A}_{\mathcal{M}}^T(v)$.

3 Weak lumpability in continuous time

The weak lumpability property has been recently addressed in Ball and Yeo (1993) for denumerable irreducible positive-recurrent Markov chains evolving in continuous time. Their main result (Theorem 2.3) is the counterpart of Theorem 2.5 in continuous time. Here, we propose to briefly discuss weak lumpability for denumerable Markov chains with the only assumption of having an uniform transition semi-group denoted by $(P_t)_{t \geq 0}$ (e.g. see Freedman (1983).) The generator of such a Markov chain is denoted by A and it is uniformly bounded. In particular, any finite Markov chain has an uniform transition semi-group. The Markov chain $(X_t)_{t \geq 0}$ is stochastically equivalent to the one with transition semi-group

$$\sum_{n=0}^{\infty} e^{-at} \frac{(at)^n}{n!} U^n \quad \text{where } a \geq \sup\{i : |A(i, i)|\} \text{ and } U = I + A/a, \quad (8)$$

The discrete-time Markov chain $(U_n)_{n \geq 0}$ with transition probability matrix U is usually called the “uniformized” chain associated with $(X_t)_{t \geq 0}$. In Ledoux and al. (1994), the result showing how to reduce the weak lumpability property from continuous time to discrete time is proved in the finite state space context. The proof is direct, avoiding preliminary works as in Rubino and Sericola (1993) (irreducible case.) Since the statement is only based on the definition of the Markov property and in the previous stochastic equivalence, this scheme still holds in the denumerable state space case. Therefore, we just express the result omitting the proof.

Theorem 3.1 *Let X be a Markov chain with an uniform transition semi-group and generator A . The chain $\text{agg}(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain iff $\text{agg}(\alpha, U, \mathcal{B})$ is also homogeneous Markov chain. So we have*

$$\mathcal{C}_{\mathcal{M}} \triangleq \{\alpha \in \mathcal{A} / \text{agg}(\alpha, A, \mathcal{B}) \text{ is a homogeneous Markov chain}\} = \mathcal{A}_{\mathcal{M}}(U).$$

If $\alpha \in \mathcal{C}_{\mathcal{M}}$ then the Markov chain $\text{agg}(\alpha, A, \mathcal{B})$ has a generator, denoted by \hat{A} , which is given by $\hat{A} = a(\hat{U} - I)$ where \hat{U} is the transition probability matrix of $\text{agg}(\alpha, U, \mathcal{B})$.

This result allows us to derive the unicity of the generator \hat{A} for all aggregated Markov chains under the assumptions of Theorems 2.5 or 2.9 for the (discrete time) “uniformized” chain $(U_n)_{n \geq 0}$. Specifically, if $(U_n)_{n \geq 0}$ is R -positive then the continuous time Markov chain $(X_t)_{t \geq 0}$ is λ -positive (in the terminology proposed by Kingman (1963)) with $\lambda = a(1 - 1/R)$ (see Buiculescu (1972).) Finally, we obtain

Corollary 3.2 *Let X be a Markov chain with an uniform transition semi-group and generator A .*

1. *Assume that X is irreducible positive-recurrent with invariant probability measure π . If $\text{agg}(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain then it admits the generator \widehat{A} given by*

$$\widehat{A}(l, m) = \sum_{i \in B(l)} \pi^{B(l)}(i) A(i, B(m)), \quad \forall l, m \in F.$$

2. *Let X be a Markov chain with an irreducible transient class T and all its absorbing states collapsed in the class $B(0)$ of the partition \mathcal{B} . The chain X is assumed to be λ -positive with a λ -invariant probability measure v . For any initial distribution α such that $\alpha_T \leq C_\alpha v$, where C_α is a positive real, if $\text{agg}(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain, then its generator \widehat{A} is given by*

$$\widehat{A}(l, m) = \sum_{i \in B(l)} v^{B(l)}(i) A(i, B(m)), \quad \forall l \in F \setminus \{0\}, \forall m \in F.$$

Finally we have

$$\begin{aligned} \mathcal{C}_{\mathcal{M}}(v) &\triangleq \{ \alpha \in \mathcal{A} / \alpha_T \leq C_\alpha v \text{ and } \text{agg}(\alpha, A, \mathcal{B}) \text{ is a homogeneous Markov chain} \} \\ &= \{ \alpha \in \mathcal{A} / \alpha_T \leq C_\alpha v \text{ and } \text{agg}(\alpha, U, \mathcal{B}) \text{ is a homogeneous Markov chain} \}. \end{aligned}$$

We note that Corollary 2.11 can be expressed in the continuous time context. Another remark is that the first part of the previous corollary, is stated under milder conditions in Ball and Yeo (1993, Th 2.5). We end the discussion in pointing out that the equivalence between discrete time and continuous time using the uniformization technique, is also reported in Sumita and Rieders (1989) for finite ergodic Markov chains. But it is based on an erroneous characterization given in Sumita and Rieders (1989, page 66) of the weak lumpability property for discrete time Markov chains. In fact, they characterize the markovian property of the aggregated chain in terms of the Chapman-Kolmogorov equation associated with. This equivalence is false in general and has been the purpose of famous counter-examples. For instance, we can take back the irreducible transition probability matrix P considered in Rosenblatt (1971, Chap 3, Section 1):

$$P = \left(\begin{array}{cc|cc} 1/4 & 1/2 & 0 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ \hline 1/4 & 0 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right).$$

with stationary distribution $\pi = (1/4, 1/4, 1/4, 1/4)$. The partition is composed of $B(0) = \{1, 2\}$ and $B(1) = \{3, 4\}$. We can verify after some algebra that the condition from Sumita and Rieders (1989) is satisfied but the chain $\text{agg}(\pi, P, \mathcal{B})$ is not markovian since

$$\mathbb{P}_\pi(X_2 \in B(0), X_1 \in B(0), X_0 \in B(0)) = \frac{3}{16} \neq \frac{25}{128} = (\pi(1) + \pi(2))(\widehat{P}(0, 0))^2$$

with $\widehat{P}(0, 0) = 5/8$. Since matrix P is irreducible, no initial distribution can lead to an aggregated markovian chain by virtue of Theorem 2.5.

References

- [1] Ball F. and Yeo G.F (1993), Lumpability and marginalisability for continuous-time Markov chains. *J. Appl. Prob.* **19**, 518–528.

- [2] Buiculescu (1972), Quasi-stationary distributions on continuous Markov chains. *Revue Roumaine de Mathématiques Pures and Appliquées* **17**, 1013–1023.
- [3] Freedman D. (1983), *Markov chains*. (Springer-Verlag).
- [4] J.G. Kemeny and J.L Snell (1976). *Finite Markov chains*. (Springer-Verlag, New York).
- [5] Kingman J.C. (1963), The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* **13**, 337–358.
- [6] Ledoux J., Rubino G. and Sericola B. (1994), Exact aggregation of absorbing Markov processes using quasi-stationary distribution. *To appear in Journal of Applied Probability*.
- [7] Pruitt W.E. (1964), Eigenvalues of non-negative matrices. *Ann. Math. Statist.* **35**, 1797–1800.
- [8] Rosenblatt M. (1971), *Markov Processes. Structure and Asymptotic Behavior*. (Springer-Verlag, New York).
- [9] G. Rubino and B. Sericola (1989), On weak lumpability in Markov chains. *J. Appl. Prob.* **26**, 446–457.
- [10] Rubino G. and Sericola B. (1991), A finite characterization of weak lumpable Markov processes. Part I: The discrete time case. *Stoch. Proc. and Appl.* **38**, 195–204.
- [11] Rubino G. and Sericola B. (1993) A finite characterization of weak lumpable Markov processes Part II: The continuous time case. *Stoch. Proc. and Appl.* **45**, 115–125.
- [12] Seneta E. (1981) *Non-negative matrices and Markov chains*. (Springer-Verlag, New York).
- [13] Seneta E. and Vere-Jones D. (1966), On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J.Appl.Prob.* **3**, 403–434.
- [14] Sumita U. and Rieders M. (1989), Lumpability and time reversibility in the aggregation-disaggregation method for large Markov chains. *Comm. Statist.- Stochastic Models*, **5**, 63–81.

Additional material not to be published

The following sections give some details which can be helpful for Theorem 2.13 and Theorem 3.1. This material consists more or less in adaptations of proofs which will appear in Ledoux and al. (1994) for the finite space context. Therefore, there are not to be included in the paper. The other notes concern technical details about the example listed in Subsection 2.4 and a discussion on the weak lumpability condition proposed in Sumita and Rieders (1989) (Section 3.)

Proof of Theorem 2.13

Denote the function of Definition 2.1 by f_P (resp. $f_{P^{(v)}}$) when related to P (resp. $P^{(v)}$.) In a similar way, X (resp. $X^{(v)}$) denotes a Markov chain with transition probability matrix P (resp. $P^{(v)}$).

We have to show that $\mathcal{A}_{\mathcal{M}}(P^{(v)}) \subseteq \mathcal{A}_{\mathcal{M}}$. If $\alpha \in \mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$, then we have by the characterization of Theorem 2.2 that for any $l \in F$ and any $\beta' = f_{P^{(v)}}(\alpha, B_0, \dots, B(l))$,

$$\widehat{P^{(v)}}(l, m) = \mathbb{P}_{\beta'}(X_1^{(v)} \in B(m)) = \sum_{i \in B(l)} \beta'(i) P^{(v)}(i, B(m)), \quad m \in F.$$

The construction of matrix $P^{(v)}$ implies that for any $l \neq 0$ and $\beta = f_P(\alpha, B_0, \dots, B(l))$, $\beta = \beta'$ and for any $m \in F$

$$\begin{aligned} \mathbb{P}_{\beta}(X_1 \in B(m)) &= \sum_{i \in B(l)} \beta'(i) P^{(v)}(i, B(m)) \\ &= \widehat{P^{(v)}}(l, m) = \widehat{P}(l, m). \end{aligned}$$

The characterization condition of Theorem 2.2 is then satisfied.

Conversely, suppose that $\alpha \in \mathcal{A}_{\mathcal{M}}^T(v) \neq \emptyset$. If $l = 0$, any vector $\beta' = f_{P^{(v)}}(\alpha, B_0, \dots, B(l))$ reduces to $1^{\{a\}}$. Therefore, we have

$$\mathbb{P}_{\beta'}(X_1^{(v)} \in B(m)) = \begin{cases} (1^{\{a\}} P^{(v)})_{B(m)} \cdot 1^* & = v_{B(m)} \cdot 1^* & \text{if } m \neq 0, \\ 0 & & \text{if } m = 0, \end{cases}$$

and this probability, which depends only on m , is therefore $\widehat{P^{(v)}}(0, m)$.

Suppose that the expression of β' does not contain the class $B(0)$. In this case, $\beta' = f_P(\alpha, B_0, \dots, B(l))$ because only matrix Q appears. We deduce $\mathbb{P}_{\beta'}(X_1^{(v)} \in B(m)) = \widehat{P}(l, m)$. Finally, suppose that $l \neq 0$ and that in the definition of vector β' the set $B(0)$ appears at least once. This vector can be written as follows:

$$\beta' = f_{P^{(v)}}(\alpha, B_0, \dots, B_{j-1}, B(0), B_{j+1}, \dots, B_n, B(l))$$

where j is the largest integer between 0 and n such that the sequence $B_{j+1}, \dots, B_n, B(l)$ does not contain $B(0)$ (with the convention that if $j = n$ then the sequence is reduced to $B(l)$.) Using the recursive definition of f , β' can also be expressed as

$$f_{P^{(v)}}\left(f_{P^{(v)}}(\alpha, \dots, B_{j-1}, B(0)) P^{(v)}, B_{j+1}, \dots, B_n, B(l)\right) = f_{P^{(v)}}((0, v), B_{j+1}, \dots, B_n, B(l)).$$

Therefore, we return to the previous situation and the proof is ended.

Some details on the example

We compute the transition probability matrix \widehat{P} associated with the aggregated chain $agg(\alpha, P, \mathcal{B})$ from relation (6)

$$\widehat{P} = \begin{pmatrix} 1 & 0 \\ 7/12 & 5/12 \end{pmatrix}.$$

If we form the irreducible recurrent positive matrix $P^{(v)}$ as in Subsection 2.3, then $agg(\alpha, P^{(v)}, \mathcal{B})$ has transition probability matrix $\widehat{P}^{(v)}$ (with Theorem 2.13)

$$\widehat{P}^{(v)} = \begin{pmatrix} 0 & 1 \\ 7/12 & 5/12 \end{pmatrix}.$$

According to Rubino and Sericola (1991), let us define matrices $\widetilde{P}_a^{(v)}$ and $\widetilde{P}_1^{(v)}$ by

$$\begin{aligned} \widetilde{P}_a^{(v)} &= \begin{pmatrix} P^{(v)}(a, B(0)) & P^{(v)}(a, B(1)) \end{pmatrix} = (1 \ 0) \\ \text{and } \widetilde{P}_1^{(v)} &= \begin{pmatrix} P^{(v)}(1, B(0)) & P^{(v)}(1, B(1)) \\ P^{(v)}(2, B(0)) & P^{(v)}(2, B(1)) \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 7/8 & 1/8 \\ \vdots & \vdots \end{pmatrix}. \end{aligned}$$

Furthermore, let us form the block diagonal $H = \text{diag}(H_l)$ with $H_0 = \widetilde{P}_a^{(v)} - 1^* \widehat{P}_a^{(v)} = 0$ and

$$H_1 = \widetilde{P}_1^{(v)} - 1^* \widehat{P}_1^{(v)} = \begin{pmatrix} -7/12 & 7/12 \\ 7/24 & -7/24 \\ \vdots & \vdots \end{pmatrix}.$$

The convex polyhedron \mathcal{A}^1 is defined by (as in Rubino and Sericola (1991))

$$\mathcal{A}^1 = \{\alpha \in \mathcal{A} / \alpha H = 0\}.$$

The linear system reduces to

$$\begin{aligned} \mathcal{A}^1 &= \{\alpha \in \mathcal{A} / \alpha_{B(1)} H_1 = 0\} \\ &= \{\alpha \in \mathcal{A} / -2\alpha(1) + \sum_{i \geq 2} \alpha(i) = 0\} \\ &= \{\alpha \in \mathcal{A} / 3\alpha(1) = 1 - \alpha(0)\}. \end{aligned}$$

We check now the $P^{(v)}$ -stability of polyhedron \mathcal{A}^1 , i.e. $\mathcal{A}^1 P^{(v)} \subseteq \mathcal{A}^1$. If $\alpha \in \mathcal{A}^1$ then we can write

$$(\alpha P^{(v)})(k) = \begin{cases} (7/8)(\sum_{i \geq 2} \alpha(i)) & \text{for } k = 0, \\ (1/3)\alpha(0) + (1/6)\alpha(1) + (1/8)(\sum_{i \geq 2} \alpha(i)) & \text{for } k = 1, \\ (1/9)(5/6)^{k-2}\alpha(0) + (1/6)(5/6)^{k-1}\alpha(1) & \text{for } k \geq 2. \end{cases}$$

From relation $3\alpha(1) = 1 - \alpha(0)$, it follows that

$$(\alpha P^{(v)})(k) = \begin{cases} (7/4)\alpha(1) & \text{for } k = 0, \\ 1/3 - (7/12)\alpha(1) & \text{for } k = 1. \end{cases}$$

Finally $1 - (\alpha P^{(v)})(0) = 3(\alpha P^{(v)})(1)$ and consequently $\alpha P^{(v)} \in \mathcal{A}^1$. The polyhedron \mathcal{A}^1 is stable by $P^{(v)}$ and we have $\mathcal{A}_{\mathcal{M}}(P^{(v)}) = \mathcal{A}^1$ as noted in Subsection 2.1.

Sketch of proof for Theorem 3.1

Consider a Poisson process $N = (N_t)_{t \geq 0}$ with rate a , such that $a \geq \sup(-A(i, i), i \in E)$. Assume that the uniformized Markov chain $(U_n)_{n \geq 0}$ is independent of N . The process $(U_{N_t})_{t \geq 0}$ has a transition semi-group given at the beginning of Section 3 and is stochastically equivalent to $X = (X_t)_{t \geq 0}$. Using this property, we prove Theorem 3.1 as follows. For all $k \in \mathbb{N}$, $B_0, \dots, B_k \in \mathcal{B}$, $0 < t_1 < \dots < t_k$ and $0 < n_1 < \dots < n_k$, we define, to simplify the notation,

$$F_X(k) = \mathbb{P}_\alpha(X_{t_k} \in B_k, \dots, X_{t_1} \in B_1, X_0 \in B_0),$$

$$F_U(k) = \mathbb{P}_\alpha(U_{n_k} \in B_k, \dots, U_{n_1} \in B_1, U_0 \in B_0),$$

$$F_N(k) = \mathbb{P}(N_{t_k} = n_k, \dots, N_{t_1} = n_1).$$

Since N is a Poisson process with rate a , we have $F_N(k) > 0 \forall k \in \mathbb{N}$ and

$$F_N(k) = F_N(k-1) \mathbb{P}(N_{t_k - t_{k-1}} = n_k - n_{k-1}). \quad (9)$$

Probability $F_X(k)$ can be expressed in terms of the equivalent stochastic process U_{N_t} as:

$$\begin{aligned} F_X(k) &= \mathbb{P}_\alpha(U_{N_{t_k}} \in B_k, \dots, U_{N_{t_1}} \in B_1, U_{N_{t_0}} \in B_0) \quad \forall k \geq 1 \\ &= \sum_{n_1 \geq 0} \sum_{n_2 \geq n_1} \dots \sum_{n_{k-1} \geq n_{k-2}} \sum_{n_k \geq n_{k-1}} F_U(k) F_N(k) \\ &\quad (\text{from the independence of } U \text{ and of } N). \end{aligned} \quad (10)$$

Assume that $\text{agg}(\alpha, U, \mathcal{B})$ is Markov homogeneous. This implies

$$F_U(k) = F_U(k-1) \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}). \quad (11)$$

We have to show that

$$F_X(k) = F_X(k-1) \mathbb{P}_\alpha(X_{t_k - t_{k-1}} \in B_k \mid X_0 \in B_{k-1}).$$

Replacing $F_U(k)$ and $F_N(k)$ in (10) by the respective relations (9) and (11), we obtain

$$\begin{aligned} F_X(k) &= \sum_{n_1 \geq 0} \dots \sum_{n_k \geq n_{k-1}} F_U(k-1) \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}) \\ &\quad \times F_N(k-1) \mathbb{P}(N_{t_k - t_{k-1}} = n_k - n_{k-1}) \\ &= \sum_{n_1 \geq 0} \dots \sum_{n_{k-1} \geq n_{k-2}} \sum_{l=0}^{+\infty} F_U(k-1) F_N(k-1) \\ &\quad \times \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1}) \mathbb{P}(N_{t_k - t_{k-1}} = l) \\ &= \left(\sum_{n_1 \geq 0} \dots \sum_{n_{k-1} \geq n_{k-2}} F_U(k-1) F_N(k-1) \right) \\ &\quad \times \sum_{l \geq 0} \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1}) \mathbb{P}(N_{t_k - t_{k-1}} = l) \\ &= F_X(k-1) \sum_{l=0}^{+\infty} \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1}) \mathbb{P}(N_{t_k - t_{k-1}} = l), \end{aligned} \quad (12)$$

that is,

$$F_X(k) = F_X(k-1) \mathbb{P}_\alpha(X_{t_k - t_{k-1}} \in B_k \mid X_0 \in B_{k-1}). \quad (13)$$

Conversely assume that $agg(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain. Relation (13) holds, so relation (12) holds too (since only formula (9) is invoked between these two expressions.) We can write

$$\begin{aligned}
F_X(k) &= \sum_{n_1 \geq 0} \cdots \sum_{n_{k-1} \geq n_{k-2}} \sum_{l \geq 0} F_U(k-1) F_N(k-1) \\
&\quad \times \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1}) \mathbb{P}(N_{t_k - t_{k-1}} = l) \\
&= \sum_{n_1 \geq 0} \cdots \sum_{n_k \geq n_{k-1}} F_U(k-1) \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}) \\
&\quad \times F_N(k-1) \mathbb{P}(N_{t_k - t_{k-1}} = n_k - n_{k-1}) \\
&= \sum_{n_1 \geq 0} \cdots \sum_{n_k \geq n_{k-1}} F_U(k-1) \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}) F_N(k) \quad \text{with (9)}.
\end{aligned}$$

Using (10), we obtain the following relation:

$$\sum_{n_1=0}^{+\infty} \cdots \sum_{n_k=n_{k-1}}^{+\infty} F_N(k) \{F_U(k) - F_U(k-1) \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1})\} = 0.$$

Therefore, we deduce that for all $n_k > \cdots > n_1 > 0$,

$$F_U(k) = F_U(k-1) \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}),$$

and so $agg(\alpha, U, \mathcal{B})$ is a homogeneous Markov chain.

Weak lumpability condition given in Sumita and Rieders (1989)

The weak lumpability characterization proposed in Sumita and Rieders (1989, page 66, eq. (2.5)) for finite ergodic discrete time Markov chain is given by

The lumped chain $agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain iff there exists a stochastic matrix $\hat{P} = (\hat{P}(l, m))_{l, m \in F}$ such that $\forall l, m \in F$:

$$\frac{\alpha_{B(l)}}{\alpha_{B(l)} 1^*} (P^n)_{B(l)B(m)} 1^* = \hat{P}^n(l, m) \quad \forall n \geq 1. \quad (14)$$

Noting that the left hand side represents the probability $\mathbb{P}_\alpha(X_n \in B(m) \mid X_0 \in B(l))$, we see in fact that we require the Chapman-Kolmogorov condition on the transition probability matrix \hat{P} of the aggregated chain $agg(\alpha, P, \mathcal{B})$. This is generally false. Let us consider the finite aperiodic irreducible matrix P given at the end of Section 3 (see Rosenblatt (1971, Chap 3, Section 1)):

$$P = \left(\begin{array}{cc|cc} 1/4 & 1/2 & 0 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ \hline 1/4 & 0 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right).$$

with stationary distribution $\pi = (1/4, 1/4, 1/4, 1/4)$. The partition is composed of $B(0) = \{1, 2\}$ and $B(1) = \{3, 4\}$ and the transition probability matrix associated with chain $agg(\alpha, P, \mathcal{B})$ given by Theorem 2.5 is

$$\hat{P} = \begin{pmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{pmatrix}.$$

For all $n \geq 1$, we have

$$\widehat{P}^n = \begin{pmatrix} 1/2 + (1/2)(1/4)^n & 1/2 - (1/2)(1/4)^n \\ 1/2 - (1/2)(1/4)^n & 1/2 + (1/2)(1/4)^n \end{pmatrix}.$$

and we can verify that

$$P^n = \left(\begin{array}{cc|cc} 1/4 & 1/4 + (1/4)^n & 1/4 - (1/4)^n & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ \hline 1/4 & 1/4 - (1/4)^n & 1/4 + (1/4)^n & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right).$$

The following quantities are respectively $\widehat{P}^n(0,0)$, $\widehat{P}^n(0,1)$, $\widehat{P}^n(1,0)$, $\widehat{P}^n(1,1)$:

$$\begin{aligned} \mathbb{P}_\pi(X_n \in B(0) \mid X_0 \in B(0)) &= \frac{\pi_{B(0)}}{\pi_{B(0)}1^*} (P^n)_{B(0)B(0)}1^* = \frac{1}{2} [1 + (1/4)^n], \\ \mathbb{P}_\pi(X_n \in B(1) \mid X_0 \in B(0)) &= \frac{\pi_{B(0)}}{\pi_{B(0)}1^*} (P^n)_{B(0)B(1)}1^* = \frac{1}{2} [1 - (1/4)^n], \\ \mathbb{P}_\pi(X_n \in B(0) \mid X_0 \in B(1)) &= \frac{\pi_{B(1)}}{\pi_{B(1)}1^*} (P^n)_{B(1)B(0)}1^* = \frac{1}{2} [1 - (1/4)^n], \\ \mathbb{P}_\pi(X_n \in B(1) \mid X_0 \in B(1)) &= \frac{\pi_{B(1)}}{\pi_{B(1)}1^*} (P^n)_{B(1)B(1)}1^* = \frac{1}{2} [1 + (1/4)^n]. \end{aligned}$$

The condition (14) is satisfied for the initial distribution π .

We compute now

$$\begin{aligned} \mathbb{P}_\pi(X_2 \in B(0), X_1 \in B(0), X_0 \in B(0)) &= \pi_{B(0)}(P_{B(0)B(0)})^2 1^* \\ &= \frac{1}{4}(1,1) \begin{pmatrix} 1/4 & 1/2 \\ 1/4 & 1/4 \end{pmatrix} (3/4 \ 1/2)^* \\ &= 3/16. \end{aligned}$$

If $agg(\pi, P, \mathcal{B})$ was a homogeneous Markov chain according to condition (14) then the probability $\mathbb{P}_\pi(X_2 \in B(0), X_1 \in B(0), X_0 \in B(0))$ will be equal to

$$(\pi_{B(0)}1^*)(\widehat{P}(0,0))^2 = \frac{25}{128}.$$

Consequently, the aggregated chain $agg(\pi, P, \mathcal{B})$ can not be markovian and we deduce from Theorem 2.5 that no initial distribution can lead to an aggregated markovian Markov chain.