HAL open science

A geometric invariant in weak lumpability of finite Markov chains

James Ledoux

To cite this version:

James Ledoux. A geometric invariant in weak lumpability of finite Markov chains. Journal of Applied Probability, 1997, 34 (4), pp.847-858. hal-00852311

HAL Id: hal-00852311

https://hal.science/hal-00852311v1

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.

A geometric invariant in weak lumpability of finite Markov chains

James Ledoux*

Abstract

We consider weak lumpability of finite homogeneous Markov chains, that is when a lumped Markov chain with respect to a partition of the initial state space is also a homogeneous Markov chain. We show that weak lumpability is equivalent to the existence of a direct sum of polyhedral cones which is is positively invariant by the transition probability matrix of the original chain. It allows us, in a unified way, to derive new results on lumpability of reducible Markov chains and to obtain spectral properties associated with lumpability.

AMS 1991 Subject Classification : Primary 60J10
Secondary 15A48.

Keywords: States aggregation, Positive invariance of cones.

1 Introduction

Finite Markov chains are extensively used as an analytic tool for systems modeling, in particular for dependability evaluation. Generally, we have to compute dependability measures which are only concerned with information on lumped states of the Markovian model or the transient features of the model can be only achieved by means of aggregation techniques which assume that the aggregated model is again a Markov chain. In this paper we are interested in deciding whether an aggregated chain is Markovian or not. Formally, let us consider a homogeneous Markov chain X, in discrete or continuous time, on a finite state space which is assumed to be $S = \{1, \ldots, N\}$. Let $\mathcal{P} = \{C(1), \ldots, C(M)\}$ be a fixed partition of S in $M < N$ classes. We associate with the given chain X the aggregated chain Y, over the state space $\hat{S} = \{1, \ldots, M\}$, defined by:

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

We are interested in the initial distributions of X which give an aggregated homogeneous Markov chain Y. If such a distribution exists, we say that the family of Markov chains sharing the same transition probability matrix (t.p.m.) is *weakly lumpable*. This problem has been addressed in Kemeny and Snell [5], Abdel-Moneim and Leysieffer [1], Rubino and Sericola [10] for Markovian models of reparable systems, that is when the generated chain is irreducible. When the system is non-reparable, the Markovian model has an absorbing state. If we may assume that it has an only

*INSA, 20 Avenue des Buttes de Cöesmes 35043 Rennes Cedex, FRANCE, email: ledoux@{univ-rennes1.fr}{irisa.fr}.

one irreducible transient states class, then weak lumpability for such a chain is tackled in Ledoux et al. [8]. In order to properly assess the effectiveness of systems where performance is "degradable", measures of performability over finite mission time have been proposed. Such measures call for general reducible Markov chains. In the present paper we propose new results on weak lumpability of reducible Markov chains. Note that aggregation of any finite continuous time Markov chain can be replaced in the discrete time context (see [8]) and, therefore, will not be discussed in the sequel.

A large amount of work on lumpability Markov chains is concerned in proving the unicity of the transition probability matrix of the Markovian aggregated chains. This property is vital to the computation of the set of all initial distributions leading to an aggregated Markov chain with the linear system approach developed in [1] and [10]. This paper emphasizes geometrical properties associated with the weak lumpability condition when we are interested in aggregated Markov chains sharing the same t.p.m. In particular, it can be used to give an unified view of the previous works and to derive new results for general finite Markov chains. After reviewing some preliminaries on polyhedral cones, we analyze in Section 2, for a general finite Markov chain with transition probability matrix P, the set of all initial distributions which give aggregated Markov chains sharing the same t.p.m. Pointing out the relation between lumpability and positive invariance of cones in Section 3, we show that this set is non empty if there exists a family of M polyhedral cones which are "invariant" under sub-matrices of matrix P. This result allows us to state in Section 4 that if the partition $\mathcal{P}$ is a refinement of the partition of the state space S induced by the usual "communication" equivalence relation, then we obtain an explicit formula for the transition probability matrix of any Y, which depends only on $\mathcal{P}$ and P. Throughout Section 3 and Section 4, various properties reported in Ledoux [6], Abdel-Moneim and Leysieffer [2] and Peng [9] are extended to general finite Markov chains and new spectral results are also derived.

Notation

- The set of all probability distributions on S will be denoted by $\mathcal{A}$. The support of a probability distribution α is defined as the subset of S corresponding to all positive components of α.

- By convention, vectors are row vectors. Column vectors are indicated by means of the transpose operator $(.)^{\mathsf{T}}$. The vector with all its components equal to 1 (respectively 0) is denoted merely by $\mathbf{1}$ (respectively $\mathbf{0}$). The vector e_i denotes the ith vector of the canonical basis of $\mathbb{R}^N$. We denote by I the identity matrix and by $\mathrm{diag}(v)$ (by $\mathrm{diag}(H_i)$) the (block) diagonal matrix with generic diagonal (block) entry $v(i)$ (the matrix H_i), the dimensions being defined by the context.

- The cardinality of the class $C(l)$ is denoted by $n(l)$. We assume the states of S ordered such that $C(l) = \{n(1) + \cdots + n(l-1) + 1, \ldots, n(1) + \cdots + n(l)\}$ for $1 \leq l \leq M$ (with $n(0) = 0$).

- For any subset C of S (whose cardinality is n) and $\alpha \in \mathcal{A}$, the restriction of α to C, i.e. the vector $(\alpha(i), i \in C)$, is denoted by α_C or $R_C\alpha$. On the other hand, a vector β on $[0,1]^{n(l)}$ can be viewed as the vector on $[0,1]^N$ defined by: $[R_l^{-1}\beta](i) = 0$ if $i \notin C(l)$ and $[R_l^{-1}\beta](i) = \beta(i - n(1) - \cdots - n(l-1))$ if $i \in C(l)$. If $\mathcal{C}$ is a subset of $\mathcal{A}$ (respectively of $[0,1]^{n(l)}$) then $R_l\mathcal{C}$ (respectively $R_l^{-1}\mathcal{C}$) denotes the set $\{\alpha_{C(l)} \mid \alpha \in \mathcal{C}\} \subseteq [0,1]^{n(l)}$ (respectively $\{R_l^{-1}\beta \mid \beta \in \mathcal{C}\} \subseteq \mathbb{R}^N$).

- If $C \subseteq S$ and $\alpha_C \mathbf{1}^\mathsf{T} \neq 0$, α^C is the vector of $\mathcal{A}$ defined by $\alpha^C(i) = \alpha(i)/\alpha_C \mathbf{1}^\mathsf{T}$ if $i \in C$ and by 0 if $i \notin C$.

2 Preliminaries on cones and weak lumpability

2.1 Cone, polyhedral cone of $\mathbb{R}^n$

The basic definitions on the cones are reviewed from Berman and Plemmons [3]. Throughout this subsection $\mathcal{C}$ denotes a subset of $\mathbb{R}^n$. For any $\mathcal{C}$, Span($\mathcal{C}$) (respectively Aff($\mathcal{C}$)) refers to the linear (respectively affine) hull of $\mathcal{C}$. The set Cone($\mathcal{C}$) denotes the conical hull of C that is the set of all finite nonnegative linear combinations of the elements of $\mathcal{C}$. The elements of $\mathcal{C}$ are called the generators of Cone($\mathcal{C}$). If Cone($\mathcal{C}$) $= \mathcal{C}$ then $\mathcal{C}$ is called a *cone*. Conv($\mathcal{C}$) is the set of all finite convex linear combinations of the elements of $\mathcal{C}$. The *dimension* of a subset $\mathcal{C}$ is defined by dim($\mathcal{C}$) $=$ dim Aff($\mathcal{C}$). The interior of $\mathcal{C}$ relative to the affine space A is denoted by $\text{int}_A(\mathcal{C})$. A convex cone $\mathcal{C}$ is *pointed* if $\mathcal{C} \cap (-\mathcal{C}) = \{\mathbf{0}\}$ and *solid* if $\text{int}_{\mathbb{R}^n}(\mathcal{C}) \neq \emptyset$. Note that a convex subset $\mathcal{C}$ is such that $\text{int}_{\text{Aff}(\mathcal{C})}(\mathcal{C}) \neq \emptyset$. Finally, a closed, pointed, solid convex cone is called a *proper* cone.

Definition 2.1 *A polyhedral cone $\mathcal{C}$ of $\mathbb{R}^n$ is the solution set of a system of linear homogeneous inequalities, i.e. $\mathcal{C} = \{\ x \in \mathbb{R}^n \mid xH \geq \mathbf{0}\ \}$ where $H \in \mathbb{R}^{n\times m}$. Such a cone is a closed convex subset of $\mathbb{R}^n$ and is non-trivial if it is not reduced to $\{\mathbf{0}\}$. We recall that a bounded solution set of a system of linear inequalities is called a polytope of $\mathbb{R}^n$.*

Definition 2.2 *Let $\mathcal{C}$ be a cone of $\mathbb{R}^n$, $\mathcal{C}_1$ and $\mathcal{C}_2$ be two sub-cones of $\mathcal{C}$. The cone $\mathcal{C}$ is the direct sum of $\mathcal{C}_1$ and $\mathcal{C}_2$, that is denoted by $\mathcal{C} = \mathcal{C}_1 \oplus \mathcal{C}_2$, if* Span$(\mathcal{C}_1) \cap$ Span$(\mathcal{C}_2) = \{\mathbf{0}\}$ *and $\mathcal{C} = \mathcal{C}_1 + \mathcal{C}_2$.*

2.2 Weak lumpability of a finite Markov chain

Let $X = (X_n)_{n\geq 0}$ be a homogeneous Markov chain over state space S, 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 $P(i,C)$ denote the transition probability of moving in one step from state i to the subset C of S, that is $P(i,C) = \sum_{j\in C} P(i,j)$. Let $P_{C(l)C(m)}$ be the $n(l) \times n(m)$ sub-matrix of P given by $(P(i,j))_{i\in C(l), j\in C(m)}$. We denote the aggregated chain constructed from (α, P) with respect to the partition $\mathcal{P}$ by $agg(\alpha, P, \mathcal{P})$. A sequence $(C_0, C_1, \ldots, C_j)$ of subsets of S is called possible for the initial distribution α if $\mathbb{P}_\alpha(X_0 \in C_0, X_1 \in C_1, \ldots, X_j \in C_j) > 0$.

The approach developed in [5] and in [10] consists in rewriting the conditional expression $\mathbb{P}_\alpha(X_{n+1} \in C(m) \mid X_n \in C(l), X_{n-1} \in C_{n-1}, \ldots, X_0 \in C_0)$ (defined for any $(C_0, C_1, \ldots, C_{n-1}, C(l))$ possible for α) as $\mathbb{P}_\beta(X_1 \in C(m))$ where β is the probability distribution $(\mathbb{P}_\alpha(X_n = i \mid X_n \in C(l), \ldots, X_0 \in C_0))_{i\in S}$. Roughly speaking, it consists in including the past into the initial distribution. Consequently, $agg(\alpha, P, \mathcal{P})$ is a homogeneous Markov chain if such a probability $\mathbb{P}_\beta(X_1 \in C(m))$ depends only on l and m. If so, it is the transition probability from state l to m for the lumped Markov chain. However, the transition probability matrix of the aggregated process which may depend on the initial distribution α (see Ledoux [7]).

Let $\widehat{P}$ be a $M \times M$ stochastic matrix. We will study in the set, denoted by $\mathcal{A}_{\mathcal{M}}(\widehat{P})$, of all initial distributions α leading to an aggregated homogeneous Markov process $agg(\alpha, P, \mathcal{P})$ with transition probability matrix $\widehat{P}$:

$$\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \{\ \alpha \in \mathcal{A} \mid agg(\alpha, P, \mathcal{C}) \text{ is a homogeneous Markov chain with t.p.m. } \widehat{P}\ \}$$

To lighten the presentation, $\mathcal{A}_{\mathcal{M}}$ will also refer to $\mathcal{A}_{\mathcal{M}}(\widehat{P})$ if there is no ambiguity. The aim of this subsection is to analyze properties of this set when it is not empty.

Let us define the following matrices.

- For any $l \in \hat{S}$, P_l denotes the $n(l) \times N$ sub-matrix of P: $(P(i,j))_{i \in C(l), j \in S}$.
- $\widetilde{P}$ denotes the $N \times M$ matrix defined by: $\forall i \in S, \forall m \in \hat{S}$, $\widetilde{P}(i,m) = P(i, C(m))$. For any $l \in \hat{S}$, we denote by $\widetilde{P}_l$ the $n(l) \times M$ sub-matrix of $\widetilde{P}$: $(\widetilde{P}(i,m))_{i \in C(l), m \in \hat{S}}$.
- The lth row of the stochastic matrix $\widehat{P}$ is denoted by $\widehat{P}_l$.
- For all $l \in \hat{S}$, we set $H_l = \widetilde{P}_l - \mathbf{1}^{\mathsf{T}} \widehat{P}_l$ $(n(l) \times M)$ and for any $j \geq 1$, we define the following $N \times M^{j+1}$ block diagonal matrices

$$H^{[1]} = \text{diag}(H_l), \qquad H^{[j+1]} = \text{diag}(P_l H^{[j]}). \tag{1}$$

We are in position to adopt the linear system approach from [10] and in the same manner, we have:

$$\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \bigcap_{j \geq 1} \mathcal{A}^j \quad \text{where } \mathcal{A}^j = \{\ \alpha \in \mathcal{A} \mid \alpha H^{[k]} = \mathbf{0}, \ \text{ for } k \geq 1\ \}.$$

Now, each polytope $\mathcal{A}^j$ can be seen as the trace on the set $\mathcal{A}$ of the following polyhedral cone:

$$\mathcal{C}^j \stackrel{\text{def}}{=} \{\alpha \geq \mathbf{0} \mid \alpha H^{[k]} = \mathbf{0}, \text{ for } 1 \leq k \leq j\}, \tag{2}$$

that is $\mathcal{A}^j = \mathcal{C}^j \cap \mathcal{A}$ for $j \geq 0$ (with the convention $\mathcal{C}^0 = \mathbb{R}_+^N$ and $\mathcal{A}^0 = \mathcal{A}$). Consequently, we note that $\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \mathcal{C}_{\mathcal{M}}(\widehat{P}) \cap \mathcal{A}$ where $\mathcal{C}_{\mathcal{M}}(\widehat{P}) \stackrel{\text{def}}{=} \bigcap_{j \geq 1} \mathcal{C}^j$ and we have

$$\mathcal{A}_{\mathcal{M}}(\widehat{P}) \neq \emptyset \Longleftrightarrow \mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}.$$

Next, if we note that $\mathcal{C}^{j+1}$ is deduced from $\mathcal{C}^j$ by attaching the (eventually) additional constraints $(\alpha H^{[j+1]} = \mathbf{0})$ and that $\dim(\mathcal{C}^1) \leq N$ then the following extension of Theorem 3.4 from [10] is intuitively clear:

$$\mathcal{C}_{\mathcal{M}}(\widehat{P}) = \mathcal{C}^N \tag{3}$$

where N is the number of states of the original chain.

We note from the diagonal structure of the matrices $H^{[j]}$, that, for $\alpha \geq \mathbf{0}$, $\alpha \in \mathcal{C}^j$ is equivalent to $R_l^{-1} \alpha_{C(l)} \in \mathcal{C}^j$ for all $l \in \hat{S}$. It allows us to derive part of the following lemma.

Lemma 2.3 *Let us set $\mathcal{C}_l^j = R_l \mathcal{C}^j$ for every $l \in \hat{S}$. We have, for all $j \geq 1$, $\mathcal{C}^j = \oplus_{l \in \hat{S}} R_l^{-1} \mathcal{C}_l^j$ where $R_l^{-1} \mathcal{C}_l^j \subseteq \mathcal{C}^j$ is a polyhedral cone of $\mathbb{R}^N$ ($\mathcal{C}_l^j$ is a polyhedral cone of $\mathbb{R}^{n(l)}$).*

proof. We can check from the definition of the sets $\mathcal{C}^j$ (see (1), (2)) that for $j \geq 1$,

$$\mathcal{C}_l^j = \left\{ \beta \in \mathbb{R}_+^{n(l)} \mid \beta H_l = \mathbf{0} \ \text{ and } \ \beta P_l H^{[k]} = \mathbf{0}, \ 1 \leq k \leq j-1 \right\}. \tag{4}$$

Consequently, $\mathcal{C}_l^j$ (respectively $R_l^{-1} \mathcal{C}_l^j$) is a polyhedral cone of $\mathbb{R}^{n(l)}$ (respectively $\mathbb{R}^N$). ∎

The well-known necessary and sufficient condition reported in [5] for having *strong lumpability* of $(., P)$ with an irreducible matrix P can be extended to a general stochastic matrix. The only requirement is that all the aggregated chains share the same t.p.m. $\widehat{P}$. In that case, by definition, the family $(., P)$ of Markov chains is strongly lumpable if $\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \mathcal{A}$ or $\mathcal{C}_{\mathcal{M}}(\widehat{P}) = \mathbb{R}_+^N$ for any $\alpha \in \mathcal{A}$. In fact, it is equivalent to require that $\mathcal{A}^1 = \mathcal{A}^0$ or $\mathcal{C}^1 = \mathcal{C}^0$. Now, $\mathcal{C}^1 = \mathbb{R}_+^N$ is equivalent to $H^{[1]} = \mathbf{0}$ or to $(H_l = \mathbf{0},\ \forall l \in \hat{S})$ which are precisely the conditions given by the following theorem.

Theorem 2.4 *If we require that all the aggregated chains share the same transition probability matrix, then the family $(., P)$ of Markov chains is strongly lumpable if and only if for each pair of classes $C(l)$ and $C(m)$, $P(i, C(m))$ does not depend on $i \in C(l)$.*

In particular, this result is necessary to derive some results in [2] though the characterization that they explicitly used is the Kemeny and Snell's one with the irreducibility assumption.

3 Lumpability and positive invariance

Definition 3.1 *A matrix A leaves a cone $\mathcal{C}$ of $\mathbb{R}^N$ invariant or matrix A is nonnegative on the cone $\mathcal{C}$, that will be denoted by $A \overset{\mathcal{C}}{\geq} 0$, if for every $x \in \mathcal{C}$ the vector $xA \in \mathcal{C}$ (i.e. $\mathcal{C}A \subseteq \mathcal{C}$). The cone $\mathcal{C}$ is said to be positively invariant by matrix A.*

Some spectral properties of matrices leaving a proper cone invariant are reviewed from [3].

Result 3.2 *If matrix A leaves a proper cone $\mathcal{C}$ invariant then the spectral radius $\rho(A)$ is an eigenvalue of A and $\mathcal{C}$ contains a left eigenvector of A corresponding to $\rho(A)$.*

Note that a nonnegative matrix is a matrix which leaves the proper cone $\mathbb{R}_+^N$ of $\mathbb{R}^N$ invariant. We will deal with cones which are not solid. Consequently, we have to derive a weaker result than the previous one.

Lemma 3.3 *If matrix A leaves a non-trivial, closed, pointed convex cone $\mathcal{C}$ invariant then there exists a nonnegative eigenvalue λ of A such that $\mathcal{C}$ contains a left eigenvector of A associated with λ.*

If a nonnegative matrix A is irreducible and leaves a non-trivial, closed, convex cone $\mathcal{C} \subseteq \mathbb{R}_+^N$ invariant then $\mathcal{C}$ contains the positive left eigenvector corresponding to the spectral radius $\rho(A)$.

proof. Matrix A represents the matrix of a linear operator f on $\mathbb{R}^N$ with respect to the canonical basis (with the convention that $f(e_i)$, for every $i \in N$, is the ith row of matrix A, that is $f(x) = xA$ for all $x \in \mathbb{R}^N$). Matrix A is nonnegative on $\mathcal{C}$ means that $f(\mathcal{C}) \subseteq \mathcal{C}$. Consequently, f leaves the linear subspace $L = \text{Span}(\mathcal{C}) \subseteq \mathbb{R}^N$ invariant and it implies that the restriction of f to the subspace L, denoted by $f_{|L}$, is a linear operator from L to L. The cone $\mathcal{C}$ is also invariant by $f_{|L}$ and is solid with respect to L. Thus, the proper cone $\mathcal{C}$ is positively invariant by the matrix $A_{|L}$ of the operator $f_{|L}$. Result 3.2 can be applied to $A_{|L}$ and conclusions are associated with the spectral radius of that matrix. However, the eigenvectors and the spectral radius of $f_{|L}$ are eigenvectors and a nonnegative eigenvalue of the initial linear operator f on domain $\mathbb{R}^N$, that gives the first part of the lemma.

If the nonnegative matrix A is irreducible, then there exists an unique (up to a constant multiple) left eigenvector of A in $\mathbb{R}_+^N$ (in fact in int$\mathbb{R}_+^N$ i.e. it is a positive left eigenvector) which corresponds

to the spectral radius of A. Now, for any closed, (pointed) convex cone $\mathcal{C} \subseteq \mathbb{R}_+^N$, if $A \overset{\mathcal{C}}{\geq} 0$ then we deduce from the first part of the proof that there exists a nonnegative left eigenvector of matrix A in $\mathcal{C} \subseteq \mathbb{R}_+^N$. Since there is only one left eigenvector of A in $\mathbb{R}_+^N$, it is positive and associated with the spectral radius of the matrix A. The second part of the lemma holds. ∎

We want emphasize that the positive invariance of polytope, used in Lemma 3.5 from [10] as a simple stop test in their incremental computation of $\mathcal{A}_{\mathcal{M}}$ from the $\mathcal{A}^j$ ones, is a central geometric invariant of the weak lumpability property as soon as we are interested in aggregated Markov chains sharing the same transition probability matrix.

Theorem 3.4 *The set $\mathcal{A}_{\mathcal{M}}(\widehat{P}) \neq \emptyset$ or $\mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$ if and only if there exists a non-trivial polyhedral cone $\mathcal{C} \subseteq \mathcal{C}^1$ such that $P \overset{\mathcal{C}}{\geq} 0$ and $\mathcal{C}$ is the direct sum $\oplus_{l \in \hat{S}} R_l^{-1}\mathcal{C}_l$ where $\mathcal{C}_l \overset{def}{=} R_l C$ for all $l \in \hat{S}$.*

proof. Suppose that $\mathcal{C}_{\mathcal{M}} \neq \{\mathbf{0}\}$. Let us verify that $\mathcal{C}_{\mathcal{M}}$ fulfills the required conditions. We have $\mathcal{C}_{\mathcal{M}} = \mathcal{C}^N$ from relation (3). Since $\mathcal{C}^N = \mathcal{C}^{N+1}$, we have for any vector $\alpha = \sum_{l \in \hat{S}} R_l^{-1}\alpha_{C(l)} \in \mathcal{C}^N$ and for any j such that $1 \leq j \leq N$

$$\begin{aligned} \alpha H^{[j+1]} = \mathbf{0} \iff & \ \forall l \in \hat{S},\ \ \alpha_{C(l)} P_l H^{[j]} = \mathbf{0} \text{ (by definition of system } H^{[j+1]}) \\ \implies & \ \alpha P H^{[j]} = \sum_{l \in \hat{S}} \alpha_{C(l)} P_l H^{[j]} = \mathbf{0}, \end{aligned}$$

that is $\alpha P \in \mathcal{C}^N$. The set $\mathcal{C}_{\mathcal{M}}$ is the direct sum of its M "projections" from relation (3) and Lemma 2.3.

Conversely, if there exists a polyhedral cone $\mathcal{C} \subseteq \mathcal{C}^1$, which is distinct from $\{\mathbf{0}\}$ and is positively invariant by P, such that $\mathcal{C} = \oplus_{l \in \hat{S}} R_l^{-1}\mathcal{C}_l$ then we show by induction that

$$\mathcal{C} \subseteq \mathcal{C}^j \ \ \forall j \geq 1.$$

The first step is obvious. Let us assume that $\mathcal{C} \subseteq \mathcal{C}^j$ with $j > 1$. For every $\alpha \in \mathcal{C}$, we have $R_l^{-1}\alpha_{C(l)} \in \mathcal{C} \subseteq \mathcal{C}^j$ for all $l \in \hat{S}$ (since $\mathcal{C}$ is a direct sum), next $[R_l^{-1}\alpha_{C(l)}]P = \alpha_{C(l)} P_l \in \mathcal{C} \subseteq \mathcal{C}^j$ for all $l \in \hat{S}$ (because $P \overset{\mathcal{C}}{\geq} 0$). We conclude that $[\forall l \in \hat{S}, \alpha_{C(l)} P_l H^{[j]} = \mathbf{0}]$ or $\alpha H^{[j+1]} = \mathbf{0}$. Thus, we have $\mathcal{C} \subseteq \mathcal{C}^{j+1}$. Finally, we obtain $\mathcal{C} \subseteq \bigcap_{j \geq 1} \mathcal{C}^j = \mathcal{C}_{\mathcal{M}}$. ∎

Using the direct sum property, Theorem 3.4 can be reformulated with "local" characteristics. That gives the main result of this section.

Theorem 3.5 *The set $\mathcal{A}_{\mathcal{M}}(\widehat{P}) \neq \emptyset$ or $\mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$ if and only if there exists a family of M polyhedral cones $(\mathcal{C}_l)_{l \in \hat{S}}$, not all trivial, such that*

$$\begin{cases} \mathcal{C}_l \subseteq \mathcal{C}_l^1 \subseteq \mathbb{R}_+^{n(l)} & \forall l \in \hat{S}, \\ \mathcal{C}_l P_{C(l)C(m)} \subseteq \mathcal{C}_m & \forall l, m \in \hat{S}. \end{cases}$$

Remark 1 The polyhedral cone $\mathcal{C}_{\mathcal{M}}(\widehat{P})$, when it is distinct from $\{\mathbf{0}\}$, satisfies the conditions of Theorem 3.5. From Theorem 3.4, it follows that $\mathcal{C}_{\mathcal{M}}(\widehat{P})$ is the largest polyhedral sub-cone of $\mathcal{C}^1$ which is positively invariant by P and is the direct sum of M polyhedral cones. However, it may exist a smaller polyhedral sub-cone of $\mathcal{C}^1$ than $\mathcal{C}_{\mathcal{M}}(\widehat{P})$ which is only positively invariant (for instance the polyhedral cone $\mathcal{C}_{\mathcal{M}} P \overset{\text{def}}{=} \{\alpha P \mid \alpha \in \mathcal{C}_{\mathcal{M}}\}$ if $\mathcal{C}_{\mathcal{M}} P \subset \mathcal{C}_{\mathcal{M}}$ [7]). ◀

Theorems 3.4 and 3.5 can be associated with the Lemma 3.3 to give the following corollary.

Corollary 3.6 *If $\mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$ then it contains a nonnegative left eigenvector corresponding to a nonnegative eigenvalue of P.*

For any $l \in \hat{S}$, if the cone $R_l\mathcal{C}_{\mathcal{M}}(\widehat{P})$ of $\mathbb{R}_+^{n(l)}$ is non-trivial then it contains a nonnegative left eigenvector corresponding to the nonnegative eigenvalue $\widehat{P}(l,l)$ of $P_{C(l)C(l)}$.

proof. The first assertion is a direct consequence from positive invariance of $\mathcal{C}_{\mathcal{M}}$ and from Lemma 3.3. Since $[R_l\mathcal{C}_{\mathcal{M}}]P_{C(l)C(l)} \subseteq [R_l\mathcal{C}_{\mathcal{M}}]$ (Theorem 3.5) and $R_l\mathcal{C}_{\mathcal{M}} \neq \{\mathbf{0}\}$, there exists a nonzero left eigenvector v_l in $R_l\mathcal{C}_{\mathcal{M}}$ associated with an eigenvalue ρ_l of $P_{C(l)C(l)}$ with Lemma 3.3; and if so, we have with vector $(R_l^{-1}v_l)^{C(l)}$ as initial distribution for the original chain $\widehat{P}(l,l) = \mathbb{P}_{(R_l^{-1}v_l)^{C(l)}}(X_1 \in C(l)) = v_l P_{C(l)C(l)}\mathbf{1}^{\mathsf{T}}/v_l\mathbf{1}^{\mathsf{T}} = \rho_l$. ∎

Remark 2 The fact that $\widehat{P}(l,l)$ is necessarily an eigenvalue of $P_{C(l)C(l)}$ completely generalizes the result given in [6] for an irreducible original chain. It was based on the fact that Markovian property induces geometric sojourn times in each class $C(l)$ and on the Jordan's canonical form of a matrix. Recall that $\widehat{P}(l,l)$ may not be $\rho(P_{C(l)C(l)})$ (see [7]) ◀

From Corollary 3.6, a cone which may fulfill the sufficient condition for weak lumpability given in Theorem 3.4 is the one which can be formed from a family $\{v_l,\ l \in \hat{S}\}$ of nonnegative left eigenvectors (and nonzero vectors) associated with the family of sub-matrices $\{P_{C(l)C(l)},\ l \in \hat{S}\}$. Let us set

$$\mathcal{C}_v \stackrel{\text{def}}{=} \text{Cone}(\{R_l^{-1}v_l,\ l \in \hat{S}\}) = \bigoplus_{l\in\hat{S}} \text{Cone}(R_l^{-1}v_l). \tag{5}$$

Since $v_l \neq \mathbf{0}$ for every $l \in \hat{S}$, we choose as $\widehat{P}$ the following stochastic matrix defined by

$$\forall l \in \hat{S},\ \ \widehat{P}_l = (R_l^{-1}v_l)^{C(l)}\widetilde{P}.$$

Thus, we deduce from Theorem 3.4 that

$$P \overset{\mathcal{C}_v}{\geq} 0 \Longrightarrow \mathcal{C}_v \subseteq \mathcal{C}_{\mathcal{M}}(\widehat{P}).$$

Such a situation raises with the *exact lumpability* property described in Schweitzer (1984). Indeed, it corresponds to assume that for all $l \in \hat{S}$, $\sum_{i\in C(l)} P(i,j)$ depends only on l and m for every $j \in C(m)$. Consequently, for every $l \in \hat{S}$, the vector $v_l = \mathbf{1}_{C(l)}$ is a left eigenvector corresponding to the eigenvalue $\sum_{i\in C(l)} P(i,j)$ of nonzero matrix $P_{C(l)C(l)}$ such that

$$\mathbf{1}_{C(l)}P_{C(l)C(m)} = [\sum_{i\in C(l)} P(i,j)]\ \mathbf{1}_{C(m)} \quad \forall m \in \hat{S},$$

and we have $\mathcal{C}_v \subseteq \mathcal{C}_{\mathcal{M}}(\widehat{P})$ according to the previous discussion. The fact that exact lumpability implies weak lumpability is well known. For instance, exact lumpability is used in Buchholz [4] for exact determination of stationary and transient measures of the initial chain from an aggregated Markov chain.

The following corollary takes advantage of the identification of the sub-cone $\mathcal{C}_v$ of $\mathcal{C}_{\mathcal{M}}(\widehat{P}) = \mathcal{C}^N$ defined in (5) and of the affine independence of the M vectors $R_l^{-1}v_l$ (i.e. dim $\mathcal{C}_v = M$).

Corollary 3.7 *We have $\mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$ if and only if $\mathcal{C}_l^{N-M} \neq \{\mathbf{0}\}$ for some $l \in \hat{S}$. In that case, we have*

$$\mathcal{C}_{\mathcal{M}}(\widehat{P}) = \underset{l \in \hat{S}}{\oplus} R_l^{-1} \mathcal{C}_l^{N-M}.$$

When $C(l)$ is is an irreducible state class of S then Corollary 3.6 and Lemma 3.3 give the following additional assertions. The final part also uses the positive invariance properties of cones $R_l \mathcal{C}_{\mathcal{M}}(\widehat{P})$ ($l \in \hat{S}$) given in Theorem 3.5.

Corollary 3.8 *Let us assume that $R_l \mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$. If $P_{C(l)C(l)}$ is an irreducible matrix then $R_l \mathcal{C}_{\mathcal{M}}(\widehat{P})$ contains only one left eigenvector v_l of $P_{C(l)C(l)}$ and this vector is positive. Moreover, $\widehat{P}(l,l)$ is the spectral radius of $P_{C(l)C(l)}$. Thus, we necessarily have $\widehat{P}_l = (R_l^{-1} v_l)^{C(l)} \widetilde{P}$.*
Moreover, for any state class $C(m)$ which can be accessed from a state of $C(l)$ (i.e. there exists a possible sequence $(C(l), C(i_1), \ldots, C(i_k), C(m))$ for some e_i with $i \in C(l)$), we have $v_l P_{C(l)C(i_1)} \cdots P_{C(i_k)C(m)} \neq \mathbf{0}$ is in $R_m \mathcal{C}_{\mathcal{M}}(\widehat{P})$ and $\widehat{P}_m = (R_m^{-1}[v_l P_{C(l)C(i_1)} \cdots P_{C(i_k)C(m)}]^{C(m)} \widetilde{P}$.

4 Lumpability of reducible Markov chains

The previous results can be applied to the aggregation of Markov chain with respect to a partition $\mathcal{P}$ which is a refinement of the partition of S corresponding to the usual communication equivalence relation. This partition is denoted by $\mathcal{I} = (I_k)_{k \in J}$ throughout this section. The elements of $\mathcal{I}$ are called the communication classes or the irreducibility classes and $|J|$ denotes the cardinality of $\mathcal{I}$. Such a state class I_k induces an irreducible sub-matrix $P_{I_k I_k}$ of P. Consequently, we can associated with each state class I_k, the unique stochastic left eigenvector v_k of $P_{I_k I_k}$ corresponding to the spectral radius of $P_{I_k I_k}$. Throughout this section, we assume that the states of S are ordered such that P is a lower block-triangular matrix

$$P = \begin{pmatrix} P_{I_1 I_1} & 0 & \cdots & 0 \\ * & P_{I_2 I_2} & \ddots & \vdots \\ * & * & \ddots & 0 \\ * & * & * & P_{I_{|J|} I_{|J|}} \end{pmatrix}.$$

Partition $\mathcal{P}$ is a *refinement* of the partition $\mathcal{I}$ if $\forall l \in \hat{S}$, $\exists! k \in J$ such that $C(l) \subseteq I_k$. For each $k \in J$, there exists $L_k \subseteq \hat{S}$ such that $I_k = \uplus_{l \in L_k} C(l)$. Any nonnegative vector α on I_k can be seen as an element of $\otimes_{l \in L_k} \mathbb{R}_+^{n(l)}$. Consequently, we denote the vector on S, $\sum_{l \in L_k} R_l^{-1} \alpha_{C(l)}$, by $R_{I_k}^{-1} \alpha$.

Example 1 Let us consider the following partition $\mathcal{P} = \{C(1) = \{1\}, C(2) = \{2,3\}, C(3) = \{4\}, C(4) = \{5,6,7\}\}$ of the state space $S = \{1,2,3,4,5,6,7\}$. The reducible transition probability matrix P is given by:

$$P = \left(\begin{array}{c|cc|c|ccc} 1/4 & 1/4 & 1/2 & 0 & 0 & 0 & 0 \\ \hline 0 & 1/6 & 5/6 & 0 & 0 & 0 & 0 \\ 7/8 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ \hline 1/7 & 0 & 0 & 3/14 & 3/14 & 3/14 & 3/14 \\ \hline 1/8 & 1/24 & 0 & 1/6 & 1/6 & 1/6 & 1/3 \\ 1/12 & 0 & 0 & 1/8 & 3/8 & 1/4 & 1/6 \\ 0 & 0 & 1/12 & 3/8 & 1/8 & 1/4 & 1/6 \end{array}\right).$$

The partition in communication classes is $\mathcal{I} = \{I_1 = \{1,2,3\}\}, I_2 = \{4,5,6,7\}\}$ with $I_1 = C(1) \uplus C(2)$ and $I_2 = C(3) \uplus C(4)$. The stochastic left eigenvectors corresponding to spectral radius of respective matrices $P_{I_1 I_1}$ and $P_{I_2 I_2}$ are $v_1 = (7/16, 3/16, 6/16)$, $v_2 = (1/4, 1/4, 1/4, 1/4)$. Vectors $R_{I_1}^{-1} v_1$, $R_{I_2}^{-1} v_2$ are respectively $(7/16, 3/16, 6/16, 0, 0, 0, 0)$ and $(0, 0, 0, 1/4, 1/4, 1/4, 1/4)$. $\triangle$

A family of communication classes $(I_{i_0}, \cdots, I_{i_n})$ is called a *path* if each class $I_{i_{k-1}}$ has an access to the class I_{i_k} for $k = 1, \ldots, n$ (that is there exists a state in $I_{i_{k-1}}$ which communicates with a state of I_{i_k}). We call I_{i_0} the starting point and I_{i_n} the end point of the path.

Theorem 4.1 *Let us assume that partition $\mathcal{P}$ is a refinement of the partition $\mathcal{I} = (I_k)_{k \in J}$ of S. We have the family of vectors $(v_k)_{k \in J}$, v_k being the stochastic left eigenvector associated with the spectral radius of matrix $P_{I_k I_k}$. If $\alpha \in \mathcal{A}$ is such that $\alpha_{I_k} \neq \mathbf{0}$ and $agg(\alpha, P, \mathcal{P})$ is a homogeneous Markov chain, then, for any m such that I_m belongs to a path with starting point I_k, we have $agg(R_{I_m}^{-1} v_m, P, \mathcal{P})$ is a homogeneous Markov chain and for all $l \in \hat{S}$ such that $C(l) \subseteq I_m$:*

$$\widehat{P}_l = (R_{I_m}^{-1} v_m)^{C(l)} \widetilde{P}; \tag{6}$$

moreover the family $\mathcal{F}_m$ composed of vectors $R_l^{-1}(v_m)_{C(l)}$ is such that $\mathrm{Cone}(\mathcal{F}_m) \subseteq \mathcal{C}_{\mathcal{M}}(\widehat{P})$.

Remark 3 The previous theorem can be interpreted as follows: if a state of a class I_k is allowed to be an initial state of our Markovian model then all the rows of matrix $\widehat{P}$ corresponding to the state classes of the $\mathcal{P}$ included in I_k or in the element of a path with starting point I_k, are necessarily given by formula (6) and depend only on $\mathcal{I}$ and P. ◀

proof. We have $I_k = \uplus_{l \in L_k} C(l)$ for some $L_k \subseteq \hat{S}$. We deduce from Theorem 3.5 that if $agg(\alpha, P, \mathcal{P})$ is a homogeneous Markov chain then there exists a pointed polyhedral cone, defined by $\mathcal{C}_{I_k} = \oplus_{l \in L_k} R_l^{-1} \mathcal{C}_l$, such that cone $R_{I_k} \mathcal{C}_{I_k}$ is positively invariant by the irreducible matrix $P_{I_k I_k}$. Lemma 3.3 states that this last cone contains the stochastic left eigenvector v_k corresponding to the spectral radius of $P_{I_k I_k}$ when it is non-trivial. Since all the distributions of cone $\mathcal{C}_{I_k}$ lead to an aggregated Markov chain with the same t.p.m. $\widehat{P}$, we deduce that $agg(R_{I_k}^{-1} v_k, P, \mathcal{P})$ is a homogeneous Markov chain and that $\widehat{P}_l = (R_{I_k}^{-1} v_k)^{C(l)} \widetilde{P}$ for every $l \in \hat{S}$ such that $C(l) \subseteq I_k$.

Let us now consider a path with starting point I_k and assume that there exists a distribution α such that $\alpha_{I_k} \neq \mathbf{0}$ and $agg(\alpha, P, \mathcal{P})$ is Markov. The chain $agg(R_{I_k}^{-1} v_k, P, \mathcal{P})$ is also a homogeneous Markov chain from the first part of the proof. Since $\mathcal{C}_{\mathcal{M}}$ is positively invariant by P, we have that for any $n \geq 0$, $R_{I_k}^{-1} v_k P^n \in \mathcal{C}_{\mathcal{M}}$. The class I_k communicate with any element I_i of the path. Consequently, letting i, there exists $n_i > 0$ such that for $w_i = R_{I_k}^{-1} v_k P^{n_i}$, $R_{I_i} w_i \neq \mathbf{0}$ and $agg(w_i, P, \mathcal{P})$ is a homogeneous Markov chain. The rows of matrix $\widehat{P}$ corresponding to the classes of $\mathcal{P}$ included in I_i are necessarily given by $\widehat{P}_l = (R_{I_i}^{-1} v_i)^{C(l)} \widetilde{P}$, from the first part of the proof.

The last part of the theorem follows from the fact that each $R_l v_m$ is in $R_l \mathcal{C}_{\mathcal{M}}$ (since $R_{I_m}^{-1} v_m \in \mathcal{C}_{\mathcal{M}}$) and from the conical property of $\mathcal{C}_{\mathcal{M}}$. ■

Remark 4 If we wish that all initial distributions on $\mathcal{A}$ lead to an aggregated homogeneous Markov chain (strong lumpability property) then, for all $k \in \hat{S}$, there must exist such a distribution whose support contains states from class I_k. Thus, Theorem 4.1 allows us to conclude that all the aggregated chains share the same transition probability matrix. Consequently, the unicity condition on this matrix required in Theorem 2.4 can be dropped. ◀

Example 1 (continued) From Theorem 4.1, the 4×4 matrix $\widehat{P}$ is necessarily given by

$$\widehat{P}_1 = (1/4, 3/4, 0, 0), \qquad \widehat{P}_2 = (7/12, 5/12, 0, 0),$$
$$\widehat{P}_3 = (1/7, 0, 3/14, 9/14), \qquad \widehat{P}_4 = (5/12, 1/24, 2/9, 2/3).$$

Let us form the matrices H_1, H_2, H_3, H_4:

$$H_1 = H_3 = \mathbf{0},\ H_2 = \begin{pmatrix} -7/12 & 7/12 & 0 & 0 \\ 7/24 & -7/24 & 0 & 0 \end{pmatrix},\ H_4 = \begin{pmatrix} 1/18 & 0 & -1/18 & 0 \\ 1/72 & -1/24 & -7/72 & 1/8 \\ -5/72 & 1/24 & 11/72 & -1/8 \end{pmatrix}.$$

The nonnegative solutions to the homogeneous system associated with each previous matrix define the four following polyhedral cones $\mathcal{C}_1^1, \mathcal{C}_2^1, \mathcal{C}_3^1, \mathcal{C}_4^1$ (see formula (4)):

$$\mathcal{C}_1^1 = \mathcal{C}_3^1 = \mathbb{R}_+;\ \ \mathcal{C}_2^1 = \text{Cone}(v_2');\ \ \mathcal{C}_4^1 = \text{Cone}(v_4');$$

with $v_2' = (1, 2) = R_2 v_1 / R_2 v_1 \mathbf{1}^\mathsf{T}$ and $v_4' = (1, 1, 1) = R_4 v_2 / R_4 v_2 \mathbf{1}^\mathsf{T}$. Note that v_2' (respectively v_4') is the positive left eigenvector (unique to a constant multiple) corresponding to the spectral radius $\widehat{P}(2, 2) = 5/12$ (respectively $\widehat{P}(4, 4) = 2/3$) of the irreducible matrix $P_{C(2)C(2)}$ (respectively $P_{C(4)C(4)}$). It is easy to check that the conditions of the Theorem 3.5 are met and thus $\mathcal{C}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$. If we construct the cone $\mathcal{C}_{v'} = \text{Cone}(\{e_1, (R_2^{-1} v_2'), e_3, (R_4^{-1} v_4')\})$, then we observe that $\mathcal{C}^1 = \mathcal{C}_{v'}$. It follows that $\mathcal{C}_{\mathcal{M}}(\widehat{P}) = \mathcal{C}_{v'}$. $\triangle$

Let us define the following positive vector on S, $v \stackrel{\text{def}}{=} \sum_{k \in J} R_{I_k}^{-1} v_k$, the convex subsets of $\mathbb{R}^N$

$$\mathcal{C}_v = \bigoplus_{l \in \hat{S}} \text{Cone}(R_l^{-1} v_{C(l)}), \qquad \mathcal{A}_v = \mathcal{C}_v \cap \mathcal{A}$$

and matrix $\widehat{P}$ by $\widehat{P}_l = v^{C(l)} \widetilde{P}$ for all $l \in \hat{S}$. In the previous example, we found that $\mathcal{C}_{\mathcal{M}}(\widehat{P}) = \mathcal{C}_v$ or $\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \mathcal{A}_v$. We can verify (with Theorem 3.5) that

$$\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \mathcal{A}_v \Longrightarrow \forall l, m \in \hat{S},\ (v^{C(l)} P)^{C(m)} = v^{C(m)}. \tag{7}$$

On the other hand, property in the right hand side implies that $\mathcal{A}_v \subseteq \mathcal{A}_{\mathcal{M}}(\widehat{P})$ with Theorem 3.4. Thus, it is a sufficient condition for weak lumpability with matrix $\widehat{P}$, as noted in [5] for irreducible matrix P. We also note that the right hand side in (7) gives for all $l \in \hat{S}$, $v_{C(l)} P_{C(l)C(l)} = \widehat{P}(l, l) v_{C(l)}$. Thus, for all $l \in \hat{S}$, $v_{C(l)}$ is a positive left eigenvector of matrix $P_{C(l)C(l)}$ corresponding to the eigenvalue $\widehat{P}(l, l)$. It can be useful to know when the converse implication of (7) holds. It is shown to be valid in [9] under the irreducible assumption for the initial matrix P and the additional condition (Γ):

$$(\Gamma)\text{: the column vectors of matrices } P^k V\ (k \geq 0) \text{ span } \mathbb{R}^N$$

where V is the $N \times M$ matrix defined by $V(i, l) = 1$ if $i \in C(l)$ and 0 otherwise. The previous comments specify some relation between the various equivalent conditions given in Theorem 3.1 from [9]. Since this theorem is based only on the condition (Γ) and the unicity of the t.p.m. associated with any aggregated chain from $\mathcal{A}_{\mathcal{M}}(\widehat{P})$, it can be directly extended to our context omitting the proof. Note that all Peng's results hold in the context of Section 2. Under (Γ), the following theorem states also that, for the aggregated chain, the Markovian property is equivalent to satisfy the Chapman-Kolmogorov equations.

Theorem 4.2 *Let us assume that partition $\mathcal{P}$ is a refinement of the partition $\mathcal{I}$ of S. Under the condition (Γ), the following are equivalent:*

$agg(v/v\mathbf{1}^{\mathsf{T}}, P, \mathcal{P})$ *satisfies to the Chapman-Kolmogorov equations;*

$(v^{C(l)}P)^{C(m)} = v^{C(m)}$ *for all* $l, m \in \hat{S}$*;*

$\mathcal{A}_{\mathcal{M}}(\widehat{P}) = \mathcal{A}_v$.

Theorem 4.1 of this section can be applied to derive the two main published results on weak lumpability. The first one deals with irreducible matrix P in [10], that is $\mathcal{I}$ reduces to only one class. If $agg(\alpha, P, \mathcal{P})$ is a homogeneous Markov chain then $agg(\pi, P, \mathcal{P})$ is also a homogeneous Markov chain where π is the stochastic vector solution to $\pi P = \pi$. The t.p.m. $\widehat{P}$ is the same for any aggregated homogeneous Markov chain and is given by $\widehat{P}_l = \pi^{C(l)}\widetilde{P},\ \ l \in \hat{S}$.

A second family of Markov chains can also be treated with Theorem 4.1. As in [8], let us consider a transition probability matrix P such that the partition of S induced by the communication equivalence relation is $\mathcal{I} = \{I_1, I_2\}$: where I_1 contains one absorbing state and I_2 all the transient ones. If there exists $\alpha \in \mathcal{A}$ such that $\alpha_{I_2} \neq \mathbf{0}$ and $agg(\alpha, P, \mathcal{P})$ is a homogeneous Markov chain then $agg((0, v), P, \mathcal{P})$ is also a homogeneous Markov chain with v is the stochastic vector solution to $vP_{I_2I_2} = \rho v$, where ρ is the spectral radius of the matrix $P_{I_2I_2}$. We recall that v is called the quasi-stationary distribution associated with the family $(., P)$. The t.p.m. $\widehat{P}$ is the same for any homogeneous Markov chain $agg(\alpha, P, \mathcal{P})$ with an initial distribution α whose support contains transient states. It is given by $\widehat{P}_1 = e_1$ and $\widehat{P}_l = (0, v)^{C(l)}\widetilde{P},\ \ l \in \hat{S} \setminus \{1\}$.

Conclusion

This paper extends to general finite Markov chains the linear system approach used in [1],[10] for the weak lumpability problem. In adopting here the viewpoint of positive invariance of polyhedral cones, we propose new results on weak/strong lumpability of a finite Markov chain. Most of our results are expressed with local characteristics of the chain, that is to the level of the state classes of the partition. This allows us to derive (or extend) spectral properties associated with exact aggregation. In a general manner, our work specifies some geometrical invariants corresponding to the lumpability requirement which are promising for studying related problems: investigate formally the weak lumpability of strongly structured Markovian models and analyze sensitivity to the "data" of the exact aggregation feasibility. We do not go into further details here.

References

[1] A.M. Abdel-Moneim and F.W. Leysieffer. Weak lumpability in finite Markov chains. *J. Appl. Probab.*, 19:685–691, 1982.

[2] A.M. Abdel-Moneim and F.W. Leysieffer. Lumpability for non-irreducible finite Markov chains. *J. Appl. Probab.*, 21:567–574, 1984.

[3] A. Berman and R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, 1979.

[4] P. Buchholz. Exact and ordinary lumpability in finite Markov chains. *J. Appl. Probab.*, 31:59–75, 1994.

[5] J.G. Kemeny and J.L. Snell. *Finite Markov chains*. Springer-Verlag, 1976.

[6] J. Ledoux. A necessary condition for weak lumpability. *Oper. Res. Letters*, 13:165–168, 1993.

[7] J. Ledoux. Weak lumpability of finite Markov chains and positive invariance of cones. Technical Report 2801, INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France, 1996.

[8] J. Ledoux, G. Rubino, and B. Sericola. Exact aggregation of absorbing Markov processes using quasi-stationary distribution. *J. Appl. Probab.*, 31:626–634, 1994.

[9] N. Peng. On weak lumpability of a finite Markov chain. *Statist. Probab. Letters*, 27:313–318, 1995.

[10] G. Rubino and B. Sericola. A finite characterization of weak lumpable Markov processes. Part I:The discrete time case. *Stochastic Process. Appl.*, 38:195–204, 1991.