

Exact aggregation of absorbing Markov processes using quasi-stationary distribution*

James Ledoux, Gerardo Rubino, Bruno Sericola.
IRISA – INRIA
Campus de Beaulieu 35042 Rennes, FRANCE.

12 October 1993

Abstract

We characterize the conditions under which an absorbing Markovian finite process (in discrete or continuous time) can be transformed into a new aggregated process conserving the Markovian property, whose states are elements of a given partition of the original state space. To obtain this characterization, a key tool is the quasi-stationary distribution associated with absorbing processes. It allows the absorbing case to be related to the irreducible one. We are able to calculate the set of all initial distributions of the starting process leading to an aggregated homogeneous Markov process by means of a finite algorithm. Finally, it is shown that the continuous time case can always be reduced to the discrete one using the uniformization technique.

Keywords: Weak Lumpability, Uniformization.

AMS 1991 Subject Classification: Primary 60J10
Secondary 60J27.

1 Introduction

A large amount of work has been devoted to the study of functional transformations of Markov processes. One of the main reasons for this is that such transformations often correspond to the only observable or desirable “response” of a system modeled by a Markovian evolution. The functional studied here agrees to the situation where the observer keeps in sight (or receives) information only provided by lumped states of the initial process. The suitable function is the aggregated process for which the state space consists of state classes of the original one. We are interested in keeping the power of Markov processes theory, which means that we want to state whether the aggregated process is Markovian or not. Necessary and sufficient conditions have been exhibited by Kemeny and Snell [3], Rubino and Sericola [7],[8] for irreducible homogeneous finite processes. In this paper we characterize this situation in the case of homogeneous absorbing finite processes.

The text is organized as follows. Section 2 introduces definitions and notation. Section 3 reviews the notion of quasi-stationary distribution which plays a central role in the paper. In Section 4, we consider discrete time processes and we outline the extensions to our context of the main results of [7]. We also show that the absorbing and irreducible cases can be related to each other. As a consequence, we can use the finite algorithm provided by [7] to give the

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

set, eventually empty, of initial distributions of the original process leading to a Markovian aggregated one. In Section 5, the continuous time case is reduced to the discrete time case using the uniformization technique. Section 6 describes how the algorithm of [7] can be used when dealing with absorbing processes.

2 Notation and definitions

This section introduces the main notation, compatible with [7] and [8]. We consider a homogeneous absorbing Markov process X in discrete or continuous time.

- The state space is finite and denoted by $E = \{1, 2, \dots, N\}$. Let $\mathcal{B} = \{B(0), B(1), \dots, B(M)\}$ be a partition of the state space E , where $B(0)$ contains only the absorbing states. The cardinality of the class $B(l)$ is denoted by $n(l)$.
- We assume this process having only one absorbing state (without any loss of generality [4].) The state 1 in E is absorbing and the remaining states are transient. We denote by T the set $\{2, \dots, N\}$ of transient states which is assumed to be irreducible.
- With the given process X , we associate the aggregated process Y with state space $F = \{0, 1, 2, \dots, M\}$ defined by:

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

- By convention, vectors are row vectors. Column vectors are indicated by means of the transpose operator $(\cdot)^*$. If v is a vector then the notation $v > 0$ notifies that each component is positive. The vector e_i denotes the i th vector of the canonical basis of $\mathbb{R}^N$. The vector with all components equal to 1 (resp. 0) is denoted merely by 1 (resp. 0), its dimension being defined by the context. We denote by I the identity matrix and by $\text{diag}(v)$ (by $\text{diag}(H_i)$) the (block) diagonal matrix with generic diagonal (block) entry $v(i)$ (the matrix H_i), the dimensions being defined by the context.
- We denote the set of all N -dimensional probability vectors by $\mathcal{A}$ and its subset of distributions with support T by $\mathcal{A}^T$, i.e. $\mathcal{A}^T = \{\alpha \in \mathcal{A} / \sum_{i \in T} \alpha(i) = 1\}$.
- Expressions of the form $\mathbb{P}_u(f(X))$ are to be interpreted as the probability of event $f(X)$ for the homogeneous Markov chain X with fixed transition probabilities and initial distribution $u \in \mathcal{A}$.
- For $l \in F$ and $\alpha \in \mathcal{A}$, the restriction of α to $B(l)$ denoted by $\alpha_{B(l)}$ is the vector with $n(l)$ entries $(\alpha(i), i \in B(l))$. This construction is also referred later with respect to any subset B of E .
- For $\alpha \in \mathcal{A}$ and $B \subset E$ such that $\alpha_B 1^* \neq 0$, let α^B denote the vector in $\mathcal{A}$ defined by $\alpha^B(i) = \alpha(i) / \sum_{j \in B} \alpha(j)$ if $i \in B$ and by 0 if $i \notin B$.

3 Quasi-stationary distribution

The stationary distribution associated with the irreducible case is a central notion for weak lumpability characterization. In our absorbing case, this distribution is trivial and without interest. The vector playing the equivalent role in our context is the quasi-stationary distribution [2],[9]. In this Section, we recall the definitions and the main properties of this distribution. In particular we discuss the periodic case which is not considered in [2], [9].

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

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

where matrix Q is irreducible. In the same way, the initial distribution vector can be written $\alpha = (\alpha_{\{1\}}, \alpha_T)$. For any $n \in \mathbb{N}_*$, we define the following probability vector π_n on T :

$$\text{if } Q \text{ is aperiodic: } \pi_n(j) \stackrel{\text{def}}{=} \mathbb{P}_\alpha(X_n = j | X_n \in T) = \frac{\alpha_T Q^n e_j^*}{\alpha_T Q^n 1^*}, \quad (1)$$

$$\text{if } Q \text{ is periodic: } \pi_n(j) \stackrel{\text{def}}{=} \frac{\sum_{k=1}^n \alpha_T r^k Q^k e_j^*}{\sum_{k=1}^n \alpha_T r^k Q^k 1^*}, \quad (2)$$

where $r = 1/\rho$ with ρ the spectral radius of Q and $v > 0$ is the unique probability left eigenvector of Q , associated with ρ . Let us denote by w the unique probability right eigenvector of Q associated with ρ . With $W = \text{diag}(w)$, define

$$\overline{Q} = \frac{1}{\rho} W^{-1} Q W. \quad (3)$$

Matrix $\overline{Q}$ is stochastic and irreducible. It is aperiodic (resp. periodic) if and only if matrix Q is aperiodic (resp. periodic). Taking back expression (3) in definitions (1), (2) and using standard arguments about the asymptotic behavior of Markov chains, we obtain the basic property of the *quasi-stationary distribution* v of the absorbing Markov chain X :

$$\text{for any initial distribution } \alpha \in \mathcal{A}, \quad \lim_{n \rightarrow \infty} \pi_n = v > 0.$$

4 Weak lumpability in discrete time

Let $X = (X_n)_{n \geq 0}$ be a homogeneous Markov chain with state space E , given by its transition probability matrix P and its initial distribution α ; when necessary we denote it by (α, P) . The partition $\mathcal{B}$ is fixed. Let $P(i, B)$ denote the transition probability of moving in one step from the 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 $\text{agg}(\alpha, P, \mathcal{B})$. We shall also consider the family of homogeneous Markov chains over the state space E , sharing the same transition probability matrix P , which will be denoted by $(\cdot, P)$.

4.1 Local characterization

Definition 4.1 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 $f(\alpha, C_1, \dots, C_j, B)$, that is:

$$\mathcal{A}(\alpha, B) \stackrel{\text{def}}{=} \{ \beta \in \mathcal{A} \mid \exists j \geq 0 \text{ and a possible sequence } (C_1, \dots, C_j, B) \text{ for } \alpha, \text{ reduced to } (B) \text{ if } j = 0, \text{ such that } \beta = f(\alpha, C_1, \dots, C_j, B) \}.$$

For any $\alpha \in \mathcal{A}$ such that $\alpha_T 1^* \neq 0$ and any $B \in \mathcal{B}$, it is easily checked that $\mathcal{A}(\alpha, B)$ is not empty, which follows from the irreducibility of Y on $F \setminus \{0\}$ and the fact that $q^* \neq 0$. In a similar way as in [7, Theorem 2.2], a necessary and sufficient condition for Y to be a homogeneous Markov chain can be exhibited after noting that any possible sequence for α whose last element is $B(l)$, $l \neq 0$, never contains $B(0)$ and that for any $\beta \in \mathcal{A}(\alpha, B(0))$, $\mathbb{P}_\beta(X_1 \in B(0)) = 1$ and $\mathbb{P}_\beta(X_1 \in B(m)) = 0$ for $m \neq 0$. We deduce

Result 4.2 The chain $Y = \text{agg}(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain iff $\forall l \in F \setminus \{0\}, \forall m \in F$, 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 .

4.2 Analysis of the initial distributions

In this subsection, we study the set of initial distributions α leading to a homogeneous Markov chain $Y = \text{agg}(\alpha, P, \mathcal{B})$. We denote this set by $\mathcal{A}_\mathcal{M}$, that is

$$\mathcal{A}_\mathcal{M} \stackrel{\text{def}}{=} \{ \alpha \in \mathcal{A} \mid Y = \text{agg}(\alpha, P, \mathcal{B}) \text{ is a homogeneous Markov chain} \}.$$

Let us denote by $\mathcal{A}_\mathcal{M}^T$ the subset of $\mathcal{A}_\mathcal{M}$ composed by the distributions with support T . We have

$$\mathcal{A}_\mathcal{M} = \lambda_1 \{e_1\} + \lambda_T \mathcal{A}_\mathcal{M}^T \text{ where } \lambda_T, \lambda_1 \geq 0 \text{ and } \lambda_T + \lambda_1 = 1. \quad (4)$$

Therefore, we restrict the analysis to the set $\mathcal{A}_\mathcal{M}^T$. When necessary, we shall specify the dependency of these sets with respect to matrix P , denoting them by $\mathcal{A}_\mathcal{M}(P)$ and $\mathcal{A}_\mathcal{M}^T(P)$.

We can obtain analogous properties as in [7] of this set. The first one is:

$$\text{If } \alpha \in \mathcal{A}_\mathcal{M}^T \text{ then for any } n \geq 1, \quad \left(0, \frac{\alpha_T Q^n}{\alpha_T Q^n 1^*} \right) \in \mathcal{A}_\mathcal{M}^T, \quad \left(0, \frac{\sum_{k=1}^n \alpha_T Q^k r^k}{\sum_{k=1}^n \alpha_T Q^k r^k 1^*} \right) \in \mathcal{A}_\mathcal{M}^T. \quad (5)$$

Using the asymptotic interpretation of the quasi-stationary distribution and the previous relations, an analogous relation of [7, (3)] is obtained : if $\mathcal{A}_\mathcal{M}^T \neq \emptyset$ then $(0, v) \in \mathcal{A}_\mathcal{M}^T$. Then it

becomes clear with Result 4.2 that the transition probability matrix of the homogeneous Markov chain $Y = \text{agg}(\alpha, P, \mathcal{B})$, denoted by $\hat{P}$, is the same for every $\alpha \in \mathcal{A}_{\mathcal{M}}^T$: for all l, m in F , $l \neq 0$,

$$\hat{P}(l, m) = \sum_{i \in B(l)} v^{B(l)}(i) P(i, B(m)). \quad (6)$$

Definition 4.3 *The family of homogeneous Markov chains $(., P)$ is weakly lumpable with respect to the partition $\mathcal{B}$ iff $\mathcal{A}_{\mathcal{M}}^T \neq \emptyset$. For any $\alpha \in \mathcal{A}_{\mathcal{M}}$, the aggregated chain $Y = \text{agg}(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain and its transition probability matrix is given by relation (6).*

An important particular case occurs when every $\alpha \in \mathcal{A}$ leads to a homogeneous and Markovian chain $Y = \text{agg}(\alpha, P, \mathcal{B})$. In this situation, the family $(., P)$ is *strongly lumpable* or simply *lumpable* with respect to the partition $\mathcal{B}$ and is characterized by the following property:

for every pair of subsets $D, B \in \mathcal{B}$, the probability $P(d, B)$ is the same for any $d \in D$.

4.3 Link with the irreducible case

From the decomposition of the transition probability matrix P in Section 3, let us define the following matrix $P^{(v)}$:

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

where v is the quasi-stationary probability vector associated with P . This matrix $P^{(v)}$ is the transition probability matrix of an irreducible Markov chain.

Result 4.4 *If $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ denotes the set of initial distributions α such that $Y = \text{agg}(\alpha, P^{(v)}, \mathcal{B})$ is Markovian homogeneous, then we have: $\mathcal{A}_{\mathcal{M}}(P) = \mathcal{A}_{\mathcal{M}}(P^{(v)})$. Moreover, the transition probability matrix $\widehat{P}^{(v)}$ of Y is given for every $m \in F$ by $\widehat{P}^{(v)}(l, m) = \hat{P}(l, m)$ with $l = 1, \dots, M$ and by $\widehat{P}^{(v)}(0, m) = v_{B(m)}1^*$ (matrix $\hat{P}$ is given by relation (6).)*

Proof. Denote the function of Definition 4.1 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. $P^{(v)}$).

If $\alpha \in \mathcal{A}_{\mathcal{M}}(P^{(v)})$, then we have by definition of weak lumpability [7] that for any $l = 0, \dots, M$ and any $\beta' = f_{P^{(v)}}(\alpha, C_0, \dots, B(l))$,

$$\widehat{P}^{(v)}(l, m) = \mathbb{P}_{\beta'}(X_1^{(v)} \in B(m)) = \sum_{i \in B(l)} \beta'(i) P^{(v)}(i, B(m)), \quad m = 0, \dots, M. \quad (7)$$

We must show that $\alpha \in \mathcal{A}_{\mathcal{M}}(P)$. The construction of matrix $P^{(v)}$ implies that for any $l \neq 0$ and $\beta = f_P(\alpha, C_0, \dots, B(l))$, $\beta = \beta'$ and for any $m = 0, \dots, M$

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

The characterization condition of Result 4.2 is then satisfied.

Conversely, suppose that $\alpha \in \mathcal{A}_{\mathcal{M}}(P)$. If $l = 0$, any vector $\beta' = f_{P^{(v)}}(\alpha, C_0, \dots, B(l))$ reduces to e_1 . Therefore, we have

$$\mathbb{P}_{\beta'}(X_1^{(v)} \in B(m)) = \begin{cases} (e_1 P^{(v)})_{B(m)} \cdot 1^* & = v_{B(m)} \cdot 1^* \text{ if } m \neq 0, \\ 0 & \text{if } m = 0, \end{cases}$$

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, $\beta' = f_P(\alpha, C_0, \dots, B(l))$ because only the 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:

$$\beta' = f_{P^{(v)}}(\alpha, C_0, \dots, C_{j-1}, B(0), C_{j+1}, \dots, C_n, B(l))$$

where j is the largest integer between 0 and n such that the sequence $C_{j+1}, \dots, C_n, B(l)$ 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

$$f_{P^{(v)}}(f_{P^{(v)}}(\alpha, \dots, C_{j-1}, B(0))P^{(v)}, C_{j+1}, \dots, C_n, B(l)) = f_{P^{(v)}}((0, v), C_{j+1}, \dots, C_n, B(l)).$$

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

This last theorem allows us to use the finite algorithm described in [7] for irreducible Markov chains as we will see in Section 6.

5 Weak lumpability in continuous time

Let $X = (X_t)_{t \geq 0}$ be a homogeneous Markov process with state space E . We show in this section that the continuous time case can be reduced to the discrete time one using the uniformization technique (see for instance [6]). This can be done as in [8] where some interesting intermediate steps are given. We give a more concise proof of this result which holds also in the irreducible case. The process X is given by its transition rate matrix A in which we set $A(i, i) = -\sum_{j \neq i} A(i, j)$.

Consider a Poisson process $(N_t)_{t \geq 0}$ with rate λ , such that $\lambda \geq \max(-A(i, i), i \in E)$. Let $U = (U_n)_{n \geq 0}$ a discrete time homogeneous Markov chain independent of (N_t) on the state space E and with transition probability matrix P given by $P = I + A/\lambda$. It can then be shown that the two processes $(X_t)_{t \geq 0}$ and $(U_{N_t})_{t \geq 0}$ are stochastically equivalent. Using this property, we prove the following result.

Result 5.1 *agg($\alpha, P, \mathcal{B}$) is a homogeneous Markov chain iff agg($\alpha, A, \mathcal{B}$) is a homogeneous Markov process. So, we have*

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

Proof. 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 λ ,

$$F_N(k) = F_N(k-1)\mathbb{P}(N_{t_k-t_{k-1}} = n_k - n_{k-1}). \quad (8)$$

From the independence of U and of N ,

$$F_X(k) = \sum_{n_1=0}^{+\infty} \sum_{n_2=n_1}^{+\infty} \cdots \sum_{n_{k-1}=n_{k-2}}^{+\infty} \sum_{n_k=n_{k-1}}^{+\infty} F_U(k)F_N(k). \quad (9)$$

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

$$F_U(k) = F_U(k-1)\mathbb{P}_\alpha(U_{n_k-n_{k-1}} \in B_k | U_0 \in B_{k-1}). \quad (10)$$

We have to show that

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

Replacing $F_U(k)$ and $F_N(k)$ in (9) by the respective relations (8) and (10), we obtain

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

that is,

$$F_X(k) = F_X(k-1)\mathbb{P}_\alpha(X_{t_k-t_{k-1}} \in B_k | X_0 \in B_{k-1}). \quad (11)$$

Conversely, if relation (11) holds then we can rewrite it using (9),

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

From the positivity of the terms $F_N(k)$, we deduce that for all $n_k > \cdots > n_1 > 0$,

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

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

As stated before, in [8] (irreducible case) a different scheme is followed. For instance, before obtaining the main result, it is shown that for any α such that $agg(\alpha, A, \mathcal{B})$ is Markov homogeneous, the transition matrix $\hat{A}$ of $agg(\alpha, A, \mathcal{B})$ is the same. This also holds here and it can be directly deduced from Result 5.1, dealing to $\hat{P} = I + \hat{A}/\lambda$.

6 The algorithm

We briefly discuss the computation of the sets $\mathcal{A}_{\mathcal{M}}$ or $\mathcal{C}_{\mathcal{M}}$ in Sections 4 and 5 for absorbing processes. In the discrete time case, the main result of Section 4 shows that we can use an irreducible auxiliary chain which will be the input of the algorithm of [7]. The conclusion of Section 5 is that the computation of $\mathcal{C}_{\mathcal{M}}$ for a continuous time absorbing Markov process can be reduced to the computation of $\mathcal{A}_{\mathcal{M}}$ on the associated uniformized discrete time chain. We can resume the situation by the following informal algorithm:

1. If X is a continuous time process then compute the uniformized matrix $P = I + A/\lambda$.
2. Compute the quasi-stationary distribution v which is the solution to $vQ = \rho v$ and $v1^* = 1$, where ρ is the spectral radius of the submatrix Q of P (see Section 3.)
3. Form the matrix $P^{(v)}$ of Subsection 4.3.
4. For each $l \in F$, $\widehat{P}_l^{(v)}$ is the l th row of the matrix $\widehat{P}^{(v)}$ defined in Result 4.4, $\widetilde{P}_l^{(v)}$ and $P_l^{(v)}$ denote the respective matrices

$$\left(P^{(v)}(i, B(k)) \right)_{i \in B(l), k \in F} \quad \text{and} \quad \left(P^{(v)}(i, j) \right)_{i \in B(l), j \in E}.$$

Compute the $n(l) \times (M + 1)$ -matrix $H_l = \widetilde{P}_l^{(v)} - 1^* \widehat{P}_l^{(v)}$, for each $l \in F$. Let us define the block diagonal matrices

$$\begin{aligned} H^{[1]} &= \text{diag}(H_l), \\ H^{[j+1]} &= \text{diag}(P_l^{(v)} H^{[j]}), \quad j \geq 1. \end{aligned}$$

Consider the following sequence of polytopes of $\mathbb{R}^N$:

$$\begin{aligned} \mathcal{A}^1 &= \{ \alpha \in \mathcal{A} / \alpha H^{[1]} = 0 \}, \\ \mathcal{A}^{j+1} &= \{ \alpha \in \mathcal{A}^j / \alpha H^{[j+1]} = 0 \}. \end{aligned}$$

The main result of [7, Theorem 3.4] identifies the polytop $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ as $\mathcal{A}^N$ and Result 4.4 gives

$$\mathcal{A}_{\mathcal{M}} = \mathcal{A}^N.$$

If we are interested in the minimal representation of the polytop $\mathcal{A}_{\mathcal{M}}$, i.e. the set of its vertices, then the incremental Chernikova's algorithm [5] is well-suited to compute these vertices with the above recursive definition of $\mathcal{A}^N$.

Acknowledgements

The authors would like to thank the referee for his helpful comments concerning Sections 5 and 6.

References

- [1] Buiculescu. Quasi-stationary distributions on continuous Markov chains. *Revue Roumaine de Mathématiques Pures et Appliquées*, 18(7):1013–1023, 1972.
- [2] J.N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Prob.*, 2:88–100, 1965.
- [3] J.G. Kemeny and J.L. Snell. *Finite Markov chains*. Springer-Verlag, New York Heidelberg Berlin, 1976.
- [4] J. Ledoux, G. Rubino and B. Sericola. Agrégation faible des processus de Markov absorbants. Tech. Report No 1736, INRIA, 1992. Campus de Beaulieu, 35042 Rennes Cedex, France.
- [5] T.H. Matheiss and D.S. Rubin. A survey and comparaison of methods for finding all vertices of convex polyhedral sets. *Mathematics of Operations Research*, 5(2):167–185, 1980.
- [6] S.M. Ross. *Stochastic Processes*. John Wiley and Sons, 1983.
- [7] G. Rubino and B. Sericola. A finite characterization of weak lumpable Markov processes. Part I: The discrete time case. *Stoch. Proc. and Appl.*, 38:195–204, 1991.
- [8] G. Rubino and B. Sericola. A finite characterization of weak lumpable Markov processes Part II: The continuous time case. *Stoch. Proc. and Appl.*, 45:115–125, 1993.
- [9] E. Seneta. *Non-negative matrices and Markov chains*. Springer-Verlag, 1981.