



Markovian bounds on functions of finite Markov chains

James Ledoux, Laurent Truffet

► To cite this version:

James Ledoux, Laurent Truffet. Markovian bounds on functions of finite Markov chains. *Advances in Applied Probability*, 2001, 33, pp.505-519. hal-00852402

HAL Id: hal-00852402

<https://hal.science/hal-00852402>

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.

Markovian Bounds on functions of Finite Markov Chains

James Ledoux* and Laurent Truffet†

8 March 2001

Abstract

In this paper, we obtain Markovian bounds on a function of a homogeneous discrete time Markov chain. For deriving such bounds, we use well known results on stochastic majorization of Markov chains and the Rogers-Pitman's lumpability criterion. The proposed method of comparison between functions of Markov chains is not equivalent to generalized coupling method of Markov chains although we obtain same kind of majorization. We derive necessary and sufficient conditions for existence of our Markovian bounds. We also discuss the choice of the geometric invariant related to the lumpability condition that we use.

KEYWORDS: weak lumpability, stochastic comparison, strong ordering.
AMS 60J10

1 Introduction

We are considering dynamic systems which could be modeled by a homogeneous discrete time Markov chain (DTMC) on a totally ordered finite state space. In the context of quick estimation of performance parameter we present a bounding method for a performance criterion which only depends on information about subsets of states (*aggregates* or *lumped states*) of the Markov chain.

More formally, let us consider a S -valued DTMC with $S = \{1, \dots, \eta\}$, $X = (X_n)_{n \in \mathbb{N}}$, also denoted by $X = (\boldsymbol{\alpha}, \mathbf{P})$, where $\boldsymbol{\alpha}$ is the probability distribution of X_0 and \mathbf{P} the $\eta \times \eta$ transition probability matrix (t.p.m.). Let us consider a performance parameter $(\theta_n)_{n \in \mathbb{N}}$ such that

$$\forall n \in \mathbb{N}, \quad \theta_n = \xi_n(\phi(X_0), \dots, \phi(X_n)), \quad (1)$$

*IRMAR UMR-CNRS 6625 & Institut National des Sciences Appliquées, 20 avenue des Buttes de Coesmes, 35708 Rennes Cedex 7, France

†IRCCyN UMR-CNRS 659 & Ecole des Mines de Nantes, Dpt. Automatique-Productique, 4 rue Alfred Kastler BP. 20722, 44307 Nantes Cedex 3, France

where $\phi : S \rightarrow \Sigma = \{1, \dots, N\}$, with $N < \eta$, and $\xi_n : \Sigma^{n+1} \rightarrow \mathbb{R}$. The main problem is that $(\phi(X_n))_{n \in \mathbb{N}}$, the *aggregated process*, is not a homogeneous Markov chain except under conditions called weak lumpability criteria [6]. Thus, under the following assumptions:

A1: ϕ is nondecreasing and $\phi(S) = \Sigma$,

A2: for every n , ξ_n is nondecreasing function in the sense of componentwise ordering,

we investigate a method to get Σ -valued Markov chains $(\underline{Y}_n)_{n \in \mathbb{N}}$ and $(\overline{Y}_n)_{n \in \mathbb{N}}$ such that:

$$\forall n \in \mathbb{N}, \quad (\underline{Y}_0, \dots, \underline{Y}_n) \leq_{st} (\phi(X_0), \dots, \phi(X_n)) \leq_{st} (\overline{Y}_0, \dots, \overline{Y}_n) \quad (2)$$

where \leq_{st} denotes the strong order of random vectors [12]. An interesting consequence of (2) is that:

$$\forall n \in \mathbb{N}, \quad \underline{\theta}_n \leq_{st} \theta_n \leq_{st} \overline{\theta}_n, \quad (3)$$

where $\underline{\theta}_n = \xi_n(\underline{Y}_0, \dots, \underline{Y}_n)$ and $\overline{\theta}_n = \xi_n(\overline{Y}_0, \dots, \overline{Y}_n)$ for any $n \in \mathbb{N}$.

In this paper we focus on the way to obtain stochastic majorization (2). Our methodology is based on well-known results on stochastic majorization [5], [7], [12] and on weak lumpability [6], [9]. It is a generalization of a first work [13] in the following ways. In [13], the author uses a weak lumpability criterion pointed out by Schweitzer [10]. Weak lumpability property has been proved to be equivalent to the existence of some invariant subspace by matrix \mathbf{P} (see [8]). Thus, every known criterion corresponds to a specific choice of this subspace. We deal here with a geometric invariant discussed in [6],[9] which is more general than Schweitzer's one. Therefore, we obtain bounds in cases which are not covered by results in [13] (e.g. see Subsection 5.3). Moreover, the invariant space in the Schweitzer's criterion is *a priori* known. This is not the case here where we have to propose a method to select suitable invariant space.

The paper is organized as follows. We present in Section 2 background material on stochastic majorization and Rogers-Pitman's lumpability criterion that will be used throughout this paper. Moreover, we explain how such results can be combined so as to derive Markovian bounds. Section 3 deals with existence of Markovian bounds on the aggregated process. It is very similar to the study of existence of exact lumpable bounds in [13]. In Section 4 we develop a policy of choice of the geometric invariant involved in the lumpability criterion. In Section 5 these results are applied to the cases of one or two aggregates. Such a context allows us to provide additional material to results of Section 4. Numerical aspects of our method are presented from an illustrative example. We conclude in Section 6.

Let us introduce some notations.

(a) By convention, vectors are row vectors. Column vectors are indicated by mean of the transpose operator $(\cdot)^*$. $\mathbf{1}$ denotes vector with all its components equal to 1. Notation $\mathbf{1}_n$, $n \geq 1$, stands for n -dimensional vector $\mathbf{1}$. $\text{diag}(\mathbf{D}_i)$ denotes block diagonal matrix with generic diagonal block entry \mathbf{D}_i .

(b) $\mathcal{A}(S)$ and $\mathcal{A}(\Sigma)$ denote the sets of all probability distributions on S and Σ , respectively.

(c) Let I be an element of Σ . By convention the i th element of set $\phi^{-1}(I) = [a_I, b_I]$ will be denoted i_I , $i_I = 1, \dots, \eta_I$, where η_I is the number of elements of set $\phi^{-1}(I)$. Note that it is also the $(a_I - 1 + i)$ th element of S since ϕ is nondecreasing.

(e) Let us define the $\eta \times N$ matrix, \mathbf{V} , associated with function ϕ , by: $V(s, I) \triangleq 1_{\{\phi(s)=I\}}$, $s \in S, I \in \Sigma$. Note that: $\mathbf{V} = \text{diag}(\mathbf{1}_{\eta_I}^*)$.

(f) Let \mathbf{A} be a $\eta \times \eta$ matrix: $\mathbf{A}_{I,J} = [A(i, j)]_{i \in \phi^{-1}(I), j \in \phi^{-1}(J)}$, $\mathbf{A}_{I,\cdot} = [A(i, j)]_{i \in \phi^{-1}(I), j \in S}$. $\mathbf{A}(k, \cdot)$ and $\mathbf{A}(\cdot, j)$ denote the k th row and the j th column of matrix \mathbf{A} , respectively.

(g) With conventions (c) and (f) and by definition of the partition $(\phi^{-1}(I))_{I \in \Sigma}$ of S , scalar $A_{I,J}(l_I, k_J)$ refers to entry $A(a_I - 1 + l, a_J - 1 + k)$ of a $\eta \times \eta$ matrix \mathbf{A} .

2 Preliminaries

2.1 Stochastic majorization.

We recall well-known results on comparison of random processes in the sense of *strong order* [12] because aggregated process $X^\phi = (\phi(X_n))_{n \in \mathbb{N}}$ is a nondecreasing function of X . Componentwise ordering between any vectors $\mathbf{x} = (x(1), \dots, x(n))$ and $\mathbf{y} = (y(1), \dots, y(n))$ of \mathbb{R}^n is defined by

$$\mathbf{x} \leq \mathbf{y} \iff (\forall i = 1, \dots, n : x(i) \leq y(i)).$$

Definition 2.1 (Strong ordering) Let X, Y be two \mathbb{R}^n -valued random variables (RVs). X is smaller than Y in the sense of strong ordering, denoted by $X \leq_{st} Y$, iff for all nondecreasing real functions f from \mathbb{R}^n (in the sense of componentwise ordering on \mathbb{R}^n), we have: $E[f(X)] \leq E[f(Y)]$ provided that expectations exist.

Situation of special interest. In the case where X and Y are $\{1, \dots, m\}$ -valued RVs with respective probability distributions $\mathbf{x} = (x(1), \dots, x(m))$ and $\mathbf{y} = (y(1), \dots, y(m))$, we have:

$$X \leq_{st} Y \quad \text{iff} \quad \mathbf{x} \mathbf{U}_m \leq \mathbf{y} \mathbf{U}_m \text{ (componentwise),}$$

where

$$\mathbf{U}_m = [1_{\{i \geq j\}}]_{i,j=1,\dots,m}.$$

Since X and Y are compared with the \leq_{st} order through their probability distributions, inequality $X \leq_{st} Y$ is also denoted by $\mathbf{x} \leq_{st} \mathbf{y}$. Sometimes we shall write $\mathbf{x} \prec_{\mathbf{U}_m} \mathbf{y}$ to emphasize the order of involved stochastic vectors.

We write $\mathbf{A} \prec_U \mathbf{B}$ for two $m \times m$ stochastic matrices \mathbf{A} and \mathbf{B} if

$$\mathbf{A}(i, \cdot) \leq_{st} \mathbf{B}(i, \cdot) \quad \forall i = 1, \dots, m \quad \text{or} \quad \mathbf{A} \mathbf{U}_m \leq \mathbf{B} \mathbf{U}_m \text{ (coefficient-wise).}$$

A $m \times m$ matrix \mathbf{M} is said to be a \leq_{st} -monotone matrix if

$$\mathbf{M}(i, \cdot) \leq_{st} \mathbf{M}(i+1, \cdot) \quad \forall i = 1, \dots, m-1 \quad \text{or} \quad \mathbf{U}_m^{-1} \mathbf{M} \mathbf{U}_m \geq \mathbf{0} \text{ (coefficient-wise).}$$

Let us now recall main result on stochastic majorization. For more references on this topic see e.g. [5], [12] and [7].

Result 2.1 (Sufficient condition to compare DTMC with same state space)

Let $Z = (\lambda, \mathbf{A})$ and $W = (\nu, \mathbf{B})$ be two DTMC taking values in the same state space, with initial distributions λ and ν , t.p.m \mathbf{A} and \mathbf{B} , respectively. If

- (a) $\lambda \leq_{st} \nu$;
- (b) $\mathbf{A} \prec_U \mathbf{B}$;
- (c) there exists a \leq_{st} -monotone matrix \mathbf{M} such that

$$\mathbf{A} \prec_U \mathbf{M} \prec_U \mathbf{B},$$

then

$$\forall n \in \mathbb{N}, \quad (Z_0, \dots, Z_n) \leq_{st} (W_0, \dots, W_n). \quad (4)$$

When (4) holds, we will say that process W is an \leq_{st} -upper bound on process Z .

Remark 2.1 Because projectors are nondecreasing functions, Relation (4) implies that for every $n \in \mathbb{N}$, $Z_n \leq_{st} W_n$. This could be rewritten as

$$\forall n \geq 0, \quad \lambda \mathbf{A}^n \leq_{st} \nu \mathbf{B}^n. \quad (5)$$

2.2 \mathbf{C} -lumpable matrix

In general a function of a Markov chain $X^\phi = (\phi(X_n))_{n \in \mathbb{N}}$ may be not a homogeneous Markov chain. The Markov property for X^ϕ is strongly related to the initial distribution of the original chain $(X_n)_{n \in \mathbb{N}}$ (see [6]). We recall some basic facts on a lumpability criterion introduced in Kemeny-Snell's book on finite Markov chain [6] and generalized by Rogers and Pitman [9] for Markov process with continuous state space.

Let us consider a $N \times \eta$ stochastic matrix $\mathbf{C} = \text{diag}(\mathbf{c}_I)$ with \mathbf{c}_I is a η_I -dimensional probability vector. Note that $\mathbf{C}\mathbf{V}$ is the identity matrix of order $N \times N$.

Definition 2.2 Matrix \mathbf{L} is said to be a \mathbf{C} -lumpable matrix if

$$\mathbf{C}\mathbf{L} = (\mathbf{C}\mathbf{L}\mathbf{V})\mathbf{C}. \quad (6)$$

We have the following properties for such a \mathbf{C} -lumpable matrix.

1. For any convex combination $\beta\mathbf{C}$ of row vectors of matrix \mathbf{C} , $(\phi(X_n))_{n \in \mathbb{N}}$ is a version of the Markov chain $(\beta\mathbf{C}\mathbf{V}, \hat{\mathbf{L}}) = (\beta, \hat{\mathbf{L}})$ where

$$\hat{\mathbf{L}} = \mathbf{C}\mathbf{L}\mathbf{V}.$$

In particular, the process $(\phi(X_n))_{n \in \mathbb{N}}$ has the same one-dimensional distributions as Markov chain $(\beta, \hat{\mathbf{L}})$, that is

$$\forall n \geq 0, \quad \beta\mathbf{C}\mathbf{L}^n\mathbf{V} = \beta\hat{\mathbf{L}}^n.$$

Note that initial distribution giving a Markov chain $(\phi(X_n))_{n \in \mathbb{N}}$ may be not restricted to the previous convex combinations.

2. Relation (6) is equivalent to say that vectors \mathbf{c}_I ($I \in \Sigma$) satisfy

$$\forall I, J \in \Sigma, \quad \mathbf{c}_I \mathbf{L}_{I,J} = \widehat{L}(I, J) \mathbf{c}_J. \quad (7)$$

3. The following relation is a direct consequence of the equality (6):

$$\forall n \geq 0, \quad \mathbf{C} \mathbf{L}^n = (\mathbf{C} \mathbf{L} \mathbf{V})^n \mathbf{C} = \widehat{\mathbf{L}}^n \mathbf{C}. \quad (8)$$

Relation (8) states that state probability vectors for the original model $(X_n)_{n \in \mathbb{N}}$ with any initial distribution $\beta \mathbf{C}$ are obtained from a computation with a $N \times N$ matrix. Such a relation is known (e.g. [10], [1]) when vector \mathbf{c}_I is an uniform distribution over $\phi^{-1}(I)$. Introduction of such a matrix \mathbf{C} was mainly motivated by this property.

2.3 Markovian bounds

The key idea to get a Markovian bound for the aggregated process is summarized in the following proposition, which combines results on stochastic majorization and \mathbf{C} -lumpability. We only consider upper bounds (lower case being obviously deduced).

Proposition 2.1 *Let $(X_n)_{n \in \mathbb{N}} = (\alpha, \mathbf{P})$ be a S -valued Markov chain, \mathbf{L} be a \mathbf{C} -lumpable matrix into matrix $\widehat{\mathbf{L}}$. We assume that \mathbf{P} is \leq_{st} -monotone. For $\beta \in \mathcal{A}(\Sigma)$, we denote the Markov chain $(\beta \mathbf{C}, \mathbf{L})$ by $(W_n)_{n \in \mathbb{N}}$.*

If

$$\alpha \leq_{st} \beta \mathbf{C} \quad \text{and} \quad \mathbf{P} \prec_U \mathbf{L}$$

then

$$\forall n \geq 0, \quad (\phi(X_0), \dots, \phi(X_n)) \leq_{st} (\phi(W_0), \dots, \phi(W_n)). \quad (9)$$

where $(\phi(W_n))_{n \in \mathbb{N}}$ is the Markov chain $(\beta, \widehat{\mathbf{L}})$. In particular, we have for the state probability vectors

$$\forall n \geq 0, \quad \alpha \mathbf{P}^n \prec_{U_\eta} \beta \widehat{\mathbf{L}}^n \mathbf{C} \quad (10)$$

$$\alpha \mathbf{P}^n \mathbf{V} \prec_{U_N} \beta \widehat{\mathbf{L}}^n. \quad (11)$$

Proof. Result 2.1 yields $(X_0, \dots, X_n) \leq_{st} (W_0, \dots, W_n)$ for every $n \geq 0$. Since ϕ is nondecreasing, we obtain (9). We get $X_n \leq_{st} W_n$ and $\phi(X_n) \leq_{st} \phi(W_n)$ from Remark 2.1, that is

$$\alpha \mathbf{P}^n \prec_{U_\eta} \beta \mathbf{C} \mathbf{L}^n$$

and Relation (11). Since \mathbf{L} is \mathbf{C} -lumpable into $\widehat{\mathbf{L}}$, $(\phi(W_n))_{n \in \mathbb{N}}$ is the Markov chain $(\beta, \widehat{\mathbf{L}})$. The last inequality and Relation (8) give Inequality (10). \square

Let us mention interesting consequences of the previous result. Assume that there exists Markov chains $\underline{Y} = (\underline{\beta}, \widehat{\underline{L}})$ and $\overline{Y} = (\overline{\beta}, \widehat{\overline{L}})$ which are a \leq_{st} -lower bound and a \leq_{st} -upper bound on $(\phi(X_n))_{n \in \mathbb{N}}$, respectively, and that

$$\underline{\beta} \underline{C} \leq_{st} \alpha \leq_{st} \overline{\beta} \overline{C},$$

with $\underline{C} = \text{diag}(\underline{c}_I)$ and $\overline{C} = \text{diag}(\overline{c}_I)$. First of all, it is possible to estimate error made in approximating the aggregated process by \underline{Y} or \overline{Y} :

$$\forall n \geq 0, \quad d(n) \triangleq \max_{I=1, \dots, N} |\overline{\beta} \widehat{\overline{L}}^n(I) - \underline{\beta} \widehat{\underline{L}}^n(I)|. \quad (12)$$

In a second time, we can also quickly estimate error made on original chain using only knowledge on \underline{Y} and \overline{Y} :

$$\forall n \geq 0, \quad D(n) \triangleq \max_{I=1, \dots, N; i_I=1, \dots, \eta_I} |\overline{\beta} \widehat{\overline{L}}^n(I) \overline{c}_I(i_I) - \underline{\beta} \widehat{\underline{L}}^n(I) \underline{c}_I(i_I)| \quad (13)$$

The next section focuses on the existence of \mathbf{C} -lumpable matrix \mathbf{L} such that $\mathbf{P} \prec_U \mathbf{L}$ when \mathbf{C} is fixed. Section 4 deals with the problem of the choice of matrix \mathbf{C} .

3 Existence of Markovian bounds on aggregated process

Let us recall that $X = (\alpha, \mathbf{P})$ denotes the initial DTMC. Existence and computation of an \leq_{st} -upper Markovian bound on process X^ϕ can be stated from the following theorem. This result is a slight extension of [13] where \mathbf{C} was the Schweitzer's matrix, i.e. $\mathbf{c}_I = \frac{1}{\eta_I} \mathbf{1}_{\eta_I}$, $I \in \Sigma$.

As it is shown in [13], we can assume without loss of generality that \mathbf{P} is a \leq_{st} -monotone matrix on S till the end of the paper.

Theorem 3.1 *There exist at least one \mathbf{C} -lumpable matrix, \mathbf{L} , such that:*

$$\mathbf{P} \prec_U \mathbf{L},$$

if and only if there exists a $\Sigma \times \Sigma$ stochastic matrix $\widehat{\mathbf{L}}$ whose entries are solution of system

$$\begin{aligned} I, J = 1, \dots, N : \\ \mathbf{c}_I \mathbf{P}_{I,J} \mathbf{U}_{\eta_J} + (\sum_{K=J+1}^N \mathbf{c}_I = \mathbf{P}_{I,K} \mathbf{1}_{\eta_K}^*) \mathbf{1}_{\eta_J} \leq \\ \widehat{\mathbf{L}}(I, J) \mathbf{c}_J \mathbf{U}_{\eta_J} + (\sum_{K=J+1}^N \widehat{\mathbf{L}}(I, K)) \mathbf{1}_{\eta_J} \quad (\text{componentwise}) \end{aligned} \quad (14)$$

Proof. We could give a similar proof than in [13]. In particular, the necessary condition is as in [13]. However, proof of the converse statement has been drastically

shortened and gives a clear insight into the geometry underlying the construction of the \mathbf{C} -lumpable matrix \mathbf{L} .

(if) It is easy to see that $\mathbf{P} \prec_U \mathbf{L}$ iff

$$\forall I = 1, \dots, N \quad \mathbf{P}_{I,\cdot} \mathbf{U}_\eta \leq \mathbf{L}_{I,\cdot} \mathbf{U}_\eta.$$

Left-multiplying this last inequality by the stochastic vector \mathbf{c}_I and using (7) we obtain that entries of the stochastic matrix $\widehat{\mathbf{L}}$ satisfy system (14).

(Only if) Construction of the matrix \mathbf{L} is in the same way as in [13]. For any $I \in \Sigma$, we choose a \leq -nonincreasing sequence of vectors $(\mathbf{d}_j)_{j=1}^\eta$ where \mathbf{d}_j is an element of set

$$\mathcal{D}_j = \{\mathbf{x} \in [0, 1]^{\eta_I} \mid \mathbf{x} \geq \mathbf{m}_j \text{ (componentwise)}, \mathbf{c}_I \mathbf{x}^* = r_j\},$$

where $\mathbf{m}_j^* = \sum_{k=j}^\eta \mathbf{P}_{I,\cdot}(\cdot, k)$ and $r_j = (\mathbf{c}_{\phi(j)} \mathbf{U}_{\eta_{\phi(j)}} \mathbf{e}_{j-a_{\phi(j)}+1}^* \widehat{L}(I, \phi(j)) + \sum_{K=\phi(j)+1}^N \widehat{L}(I, K)$.

For any $I \in \Sigma$, we define matrix $\mathbf{L}_{I,\cdot}$ from vectors $(\mathbf{d}_j)_{j=1}^\eta$ by

$$\forall j \in S, \quad \mathbf{L}_{I,\cdot}(\cdot, j) = \mathbf{d}_j^* - \mathbf{d}_{j+1}^* \text{ (with convention } \mathbf{d}_{\eta+1} = \mathbf{0}).$$

It is easy to check that above-defined matrix \mathbf{L} is \mathbf{C} -lumpable and such that $\mathbf{P} \prec_U \mathbf{L}$.

We just have to justify existence of vectors $(\mathbf{d}_j)_{j=1}^\eta$ for any $I \in \Sigma$. In fact system (14) gives that each vector \mathbf{m}_j satisfies

$$\mathbf{c}_I \mathbf{m}_j^* \leq r_j. \tag{15}$$

Note that the sequence $(r_j)_{j=1}^\eta$ and $(\mathbf{m}_j)_{j=1}^\eta$ are nonincreasing.

Now existence of the nonincreasing sequence $(\mathbf{d}_j)_{j=1}^\eta$ is proved by induction. We have $\mathbf{d}_1 = \mathbf{1}$ since $r_1 = 1$ and $\mathbf{m}_1 = \mathbf{1}$. Let us assume that we have obtained $\mathbf{1} \geq \mathbf{d}_2 \geq \dots \geq \mathbf{d}_j \geq \mathbf{0}$ for some $j > 1$. We compute \mathbf{d}_{j+1} as follows. We have $\mathbf{d}_j \geq \mathbf{m}_j \geq \mathbf{m}_{j+1}$ and $\mathbf{c}_I \mathbf{d}_j^* = r_j \geq r_{j+1} \geq \mathbf{c}_I \mathbf{m}_{j+1}^*$. The last inequality follows from (15). We get that \mathbf{d}_j is "over" the affine hyperplane $\{\mathbf{x} \in \mathbb{R}^{\eta_I} \mid \mathbf{c}_I \mathbf{x}^* = r_{j+1}\}$ and \mathbf{m}_{j+1} is "under" this hyperplane. Therefore, the segment of line $[\mathbf{m}_{j+1}, \mathbf{d}_j]$ joining \mathbf{m}_{j+1} to \mathbf{d}_j cuts $\{\mathbf{x} \in \mathbb{R}_+^{\eta_I} \mid \mathbf{c}_I \mathbf{x}^* = r_{j+1}\}$ in one point. Let \mathbf{d}_{j+1} be this point. Then \mathbf{d}_{j+1} is such that $\mathbf{0} \leq \mathbf{m}_{j+1} \leq \mathbf{d}_{j+1} \leq \mathbf{d}_j$ since any point of the segment $[\mathbf{m}_{j+1}, \mathbf{d}_j]$ has this property. Finally, we have $\mathbf{c}_I \mathbf{d}_{j+1}^* = r_{j+1}$ by construction. \square

Corollary 3.1 *Let us assume that there exists a stochastic matrix $\widehat{\mathbf{L}}$ whose entries satisfy system (14). The entries of any stochastic matrix $\mathbf{\Lambda}$ such that $\widehat{\mathbf{L}} \prec_U \mathbf{\Lambda}$ are still solution to system (14).*

Proof.

Assume that $\widehat{\mathbf{L}} \prec_U \mathbf{\Lambda}$ with $\mathbf{\Lambda} = [\lambda(I, J)]_{I, J \in \Sigma}$. We first show that

$$\forall J \in \Sigma, \forall \beta \in [0, 1], \quad \beta \widehat{L}(I, J) + \sum_{K=J+1}^N \widehat{L}(I, J) \leq \beta \lambda(I, J) + \sum_{K=J+1}^N \lambda(I, K). \tag{16}$$

For a fixed J , let us define $g(\beta) = \beta = (\lambda(I, J) - \widehat{L}(I, J)) + \sum_{K=J+1}^N (\lambda(I, K) - \widehat{L}(I, K))$. g is an affine function. Since $\widehat{\mathbf{L}}(I, \cdot) \leq_{st} \mathbf{A}(I, \cdot)$, g is such that $g(0) \geq 0$ and $g(1) \geq 0$. Therefore, $g(\beta) \geq 0$ for all $\beta \in [0, 1]$.

Thus, for each l in $\{0, \dots, \eta_J - 1\}$, we have inequality (16) with $\beta = \sum_{k_J=\eta_J-l}^{\eta_J} \mathbf{c}_J(k_J)$. This shows that scalars $(\lambda(I, J))_{I, J \in \Sigma}$ satisfy system (14). \square

We deduce from Theorem 3.1 and Proposition 2.1, an existence condition and the computation of an \leq_{st} -upper Markovian bound on process X^ϕ .

Corollary 3.2 *If the two following conditions are fulfilled*

- (a) *there exists a stochastic matrix $\widehat{\mathbf{L}}$ whose entries satisfy system (14)*
- (b) *there exists a probability vector $\beta \in \mathcal{A}(\Sigma)$ such that*

$$\alpha \leq_{st} \beta \mathbf{C}, \quad (17)$$

then the Σ -valued Markov chain $Y = (\beta, \widehat{\mathbf{L}})$ is an \leq_{st} -upper bound on X^ϕ , i.e.

$$\forall n \in \mathbb{N}, \quad (\phi(X_0), \dots, \phi(X_n)) \leq_{st} (Y_0, \dots, Y_n).$$

To end this section, let us mention that it is also possible to study lower bounding problem. Same kind of results could be derived by reversing inequalities in (14), Theorem 3.1 and in (17), Corollary 3.2.

4 Policy of choice of matrix \mathbf{C}

In this section we deal with the selection of the stochastic matrix $\mathbf{C} = \text{diag}(\mathbf{c}_I)$. The choice is based on the following condition of existence of a \mathbf{C} -lumpable upper bound.

Lemma 4.1 (Existence of upper bounds) *There exists at least one \mathbf{C} -lumpable matrix \mathbf{L} such that $\mathbf{P} \prec_U \mathbf{L}$ iff*

$$(\Gamma) \quad \forall I \in \Sigma, \quad \mathbf{c}_I \mathbf{P}_{I,N} \mathbf{U}_{\eta_N} \leq \mathbf{c}_N \mathbf{U}_{\eta_N}.$$

Proof. If there exists a \mathbf{C} -lumpable upper bound, then the scalars $\widehat{L}(I, J)$, $I, J \in \Sigma$, satisfy (14). Since $\widehat{L}(I, N) \leq 1$ for every $I \in \Sigma$, we deduce from inequalities (14) with $J = N$, that $\mathbf{c}_I \mathbf{P}_{I,N} \mathbf{U}_{\eta_N} \leq \mathbf{c}_N \mathbf{U}_{\eta_N}$ for every $I \in \Sigma$.

Conversely, let us assume that condition (Γ) is fulfilled. It is easy to check that the scalars

$$\widehat{L}(I, J) = 1_{\{J=N\}} \quad I, J \in \Sigma,$$

satisfy constraints (14) and define a stochastic matrix $\widehat{\mathbf{L}}$. \square

Remark 4.1 *For any matrix \mathbf{C} , lower bound always exists.*

Theorem 4.1 *Solution of (Γ) is obtained as follows.*

1. *We choose vector \mathbf{c}_N satisfying*

$$\mathbf{c}_N \mathbf{P}_{N,N} \mathbf{U}_{\eta_N} \leq \mathbf{c}_N \mathbf{U}_{\eta_N}. \quad (18)$$

In particular, a stochastic left-eigenvector \mathbf{v}_N corresponding to the spectral radius ρ_N of matrix $\mathbf{P}_{N,N}$ always satisfies (18).

2. *Since \mathbf{P} is \leq_{st} -monotone, every stochastic vector \mathbf{c}_I satisfies $\mathbf{c}_I \mathbf{P}_{I,N} \mathbf{U}_{\eta_N} \leq \mathbf{c}_N \mathbf{U}_{\eta_N}$ ($I = 1, \dots, N-1$).*

Proof. For the first assertion, we just have to quote that the spectral radius of a substochastic matrix is always smaller than 1.

Since \mathbf{P} is \leq_{st} -monotone, we have

$$\forall I = 1, \dots, N-1, \quad \mathbf{P}_{I,N} \mathbf{U}_{\eta_N} \leq \mathbf{P}_{N,N} \mathbf{U}_{\eta_N} \text{ (componentwise).}$$

Then we have for the following convex combination of previous inequalities

$$\mathbf{c}_I \mathbf{P}_{I,N} \mathbf{U}_{\eta_N} \leq \mathbf{c}_N \mathbf{P}_{N,N} \mathbf{U}_{\eta_N}$$

which gives the second assertion.

Therefore, if \mathbf{c}_N satisfies (18) then (Γ) is also satisfied. \square

Remark 4.2 *Condition (Γ) is always satisfied with $\mathbf{C} = \text{diag}(\mathbf{v}_I)$, where \mathbf{v}_I is a stochastic left-eigenvector associated with the spectral radius ρ_I of $\mathbf{P}_{I,I}$. Moreover, if \mathbf{P} is \mathbf{C} -lumpable, this choice for \mathbf{C} ensures that the upper bound \mathbf{L} coincides with \mathbf{P} .*

Remark 4.3 *The existence of a \mathbf{C} -lumpable upper bound does not imply existence of \mathbf{C}' -lumpable upper bound for matrix \mathbf{C}' such that $\mathbf{C}' \prec_U \mathbf{C}$ (see Subsection 5.3).*

5 Applications

We go into further details for deriving bounds on aggregated process in the two following contexts. We consider lumping in one or two classes. Such partitions are basically involved in reliability theory when each state of a system is either a up-state or a down-state. Down-states of the Markov model are lumped in one aggregate. The other ones may define a second aggregate. The illustrative example will give another instance of framework (queueing theory) where such partitions are useful. Let us recall that \mathbf{P} is assumed to be a \leq_{st} -monotone matrix on S .

5.1 One aggregate

We lump only one subset of states of S . It means that the $N - 1$ former "classes" contain only one state of S and the last class lumps the others. Thus we have $\eta_1 = \dots = \eta_{N-1} = 1$ and $\eta_N = \eta - N + 1$. In such a case, we define matrix \mathbf{C} by $\mathbf{c}_I = (1)$ for $I = 1, \dots, N - 1$ and $\mathbf{c}_N = \mathbf{v}_N$ where \mathbf{v}_N is a stochastic left eigenvector of matrix $\mathbf{P}_{N,N}$ corresponding to its spectral radius ρ_N . It follows from Remark 4.2 that we always have a stochastic matrix $\widehat{\mathbf{L}}$ solution of system (14). In fact, we can choose matrix $\widehat{\mathbf{L}}$ as follows.

1. $\widehat{L}(N, N) = \rho_N$, $\widehat{L}(N, J) = \mathbf{v}_N \mathbf{P}_{N,J}$ for $J = 1, \dots, N - 1$;
2. For every $I = 1, \dots, N - 1$, compute the smallest vector $\widehat{\mathbf{L}}(I, \cdot)$ with respect to the partial order \leq_{st} such that

$$\begin{aligned} J = 1, \dots, N - 1 \quad \mathbf{P}_{I,J} \mathbf{U}_{\eta_J} + \sum_{K=J+1}^N \mathbf{P}_{I,K} \mathbf{1}_{\eta_K}^* &\leq \sum_{K=J}^N \widehat{L}(I, K) \\ \mathbf{P}_{I,N} \mathbf{U}_{\eta_N} &\leq \widehat{L}(I, N) \mathbf{v}_N \mathbf{U}_{\eta_N}. \end{aligned}$$

It is not difficult to justify that such a definition of scalars $(\widehat{L}(I, J))_{I,J \in \Sigma}$ gives a stochastic matrix $\widehat{\mathbf{L}}$ whose entries are solution of system (14). Note that matrix $\widehat{\mathbf{L}}$ whose entries are defined by

$$\begin{aligned} \widehat{L}(N, N) = \rho_N, \quad \widehat{L}(N, J) = \mathbf{v}_N \mathbf{P}_{N,J} \quad \text{for } J = 1, \dots, N - 1; \\ \text{and } \widehat{L}(I, J) = 1_{\{J=N\}} \quad \text{for } I = 1, \dots, N, \end{aligned}$$

is always a solution of system (14).

5.2 two aggregates

When we have two aggregates, i.e. a partition of S in two subsets, we choose vectors $\{\mathbf{c}_1, \mathbf{c}_2\}$ and matrix $\widehat{\mathbf{L}}$ as follows.

1. $\widehat{L}(2, 2)$ is the spectral radius ρ_2 of the nonnegative matrix $\mathbf{P}_{2,2}$ and \mathbf{c}_2 is a stochastic left-eigenvector \mathbf{v}_2 of $\mathbf{P}_{2,2}$ corresponding to ρ_2 .
2. $\mathbf{c}_1 = \mathbf{v}_2 (\mathbf{I} - \mathbf{P}_{2,2})^{-1} \mathbf{P}_{2,1} = \frac{1}{1-\rho_2} \mathbf{v}_2 \mathbf{P}_{2,1}$.
3. $\mathbf{c}_1, \mathbf{c}_2, \widehat{L}(2, 2)$ are fixed. Set $\widehat{L}(1, 2) = \max(C_1, C_2)$ with

$$\begin{aligned} C_1 = \max \left(\max_{i=1, \dots, \eta_1, \mathbf{c}_1 \mathbf{U}_{\eta_1}(i) \neq 1} \frac{(\mathbf{c}_1 \mathbf{P}_{1,1} \mathbf{U}_{\eta_1} + \mathbf{c}_1 \mathbf{P}_{1,2} \mathbf{1}_{\eta_2}^* \mathbf{1}_{\eta_1} - \mathbf{c}_1 \mathbf{U}_{\eta_1})(i)}{1 - (\mathbf{c}_1 \mathbf{U}_{\eta_1})(i)}, 0 \right) \\ C_2 = \max_{i=1, \dots, \eta_2, \mathbf{c}_2 \mathbf{U}_{\eta_2}(i) \neq 0} \frac{(\mathbf{c}_1 \mathbf{P}_{1,2} \mathbf{U}_{\eta_2})(i)}{(\mathbf{c}_2 \mathbf{U}_{\eta_2})(i)} \end{aligned}$$

This choice of vectors $\{\mathbf{c}_1, \mathbf{c}_2\}$ is valid from Theorem 4.1. Now, entries of matrix $\hat{\mathbf{L}}$ are solution of system (14) with vectors $\{\mathbf{c}_1, \mathbf{c}_2\}$. We show that this choice is partly optimal in the following sense.

With only two aggregates, system (14) becomes

$$\begin{cases} \mathbf{c}_1 \mathbf{P}_{1,2} \mathbf{U}_{\eta_2} \leq \hat{L}(1, 2) \mathbf{c}_2 \mathbf{U}_{\eta_2} \\ \mathbf{c}_1 \mathbf{P}_{1,1} \mathbf{U}_{\eta_1} + \mathbf{c}_1 \mathbf{P}_{1,2} \mathbf{1}_{\eta_2}^* \mathbf{1}_{\eta_1} \leq (1 - \hat{L}(1, 2)) \mathbf{c}_1 \mathbf{U}_{\eta_1} + \hat{L}(1, 2) \mathbf{1}_{\eta_1} \end{cases} \quad (20a)$$

$$\begin{cases} \mathbf{c}_2 \mathbf{P}_{2,2} \mathbf{U}_{\eta_2} \leq \hat{L}(2, 2) \mathbf{c}_2 \mathbf{U}_{\eta_2} \\ \mathbf{c}_2 \mathbf{P}_{2,1} \mathbf{U}_{\eta_1} + \mathbf{c}_2 \mathbf{P}_{2,2} \mathbf{1}_{\eta_2}^* \mathbf{1}_{\eta_1} \leq (1 - \hat{L}(2, 2)) \mathbf{c}_1 \mathbf{U}_{\eta_1} + \hat{L}(2, 2) \mathbf{1}_{\eta_1}. \end{cases} \quad (20b)$$

Note that $\hat{\mathbf{L}} \prec_U \tilde{\mathbf{L}}$ is equivalent to

$$\hat{L}(1, 2) \leq \tilde{L}(1, 2), \quad \hat{L}(2, 2) \leq \tilde{L}(2, 2).$$

For fixed vectors $\{\mathbf{c}_1, \mathbf{c}_2\}$, if entries of matrix $\hat{\mathbf{L}}$ satisfy (20a,20b) then it follows from Corollary 3.1 that entries of any stochastic matrix $\tilde{\mathbf{L}}$ such that $\hat{\mathbf{L}} \prec_U \tilde{\mathbf{L}}$ also satisfy (20a,20b). We would like to obtain the smallest matrix $\hat{\mathbf{L}}$ with respect to the partial order \leq_{st} whose entries satisfy (20a,20b). So, we have to choose $\hat{L}(2, 2)$ as follows

$$\hat{L}(2, 2) = \max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2))$$

with

$$\begin{aligned} G_1(\mathbf{c}_2) &= \max_{i=1, \dots, \eta_2; \mathbf{x}_2 = \mathbf{U}_{\eta_2}(i) \neq 0} \frac{(\mathbf{x}_2 \mathbf{P}_{2,2} \mathbf{U}_{\eta_2})(i)}{(\mathbf{x}_2 \mathbf{U}_{\eta_2})(i)} \\ G_2(\mathbf{c}_1, \mathbf{c}_2) &= \max_{i=1, \dots, \eta_1, \mathbf{x}_1 \mathbf{U}_{\eta_1}(i) \neq 1} \frac{(\mathbf{x}_2 \mathbf{P}_{2,1} \mathbf{U}_{\eta_1} + \mathbf{x}_2 \mathbf{P}_{2,2} \mathbf{1}_{\eta_2}^* \mathbf{1}_{\eta_1} - \mathbf{x}_1 \mathbf{U}_{\eta_1})(i)}{1 - (\mathbf{x}_1 \mathbf{U}_{\eta_1})(i)}. \end{aligned}$$

The optimal choice for $\hat{L}(2, 2)$, \mathbf{c}_1 and \mathbf{c}_2 would be

$$\hat{L}(2, 2) = \max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2)) = \inf_{\mathbf{c}_1 \geq \mathbf{0}, \mathbf{c}_2 \in \mathcal{D}_2} (\max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2))) \quad (21)$$

where $\mathcal{D}_2 = \{\mathbf{c}_2 \geq \mathbf{0} \mid \forall i, (\mathbf{c}_2 \mathbf{U}_{\eta_2}(i) = 0 \implies \mathbf{c}_2 \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}(i) = 0)\}$.

Since $\mathbf{v}_2 \in \mathcal{D}_2$ and $\frac{1}{1-\rho_2} \mathbf{v}_2 \mathbf{P}_{2,1}$ is a stochastic vector, we always have

$$\inf_{\mathbf{c}_1 \geq \mathbf{0}, \mathbf{c}_2 \in \mathcal{D}_2} (\max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2))) \leq \rho_2 = \max(G_1(\mathbf{v}_2), G_2(\frac{1}{1-\rho_2} \mathbf{v}_2 \mathbf{P}_{2,1}, \mathbf{v}_2)).$$

Thus, the choice $\hat{L}(2, 2) = \rho_2$, $\mathbf{c}_2 = \mathbf{v}_2$ and $\mathbf{c}_1 = \frac{1}{1-\rho_2} \mathbf{v}_2 \mathbf{P}_{2,1}$ is not optimal in general with respect to Problem (21). But, this will be the case when matrix $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}$ or $\mathbf{P}_{2,2}$ is irreducible. Note that $\mathbf{P}_{2,2}$ is a monotone matrix so that $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}$ is a nonnegative matrix.

Proposition 5.1 *If matrix $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}$ or $\mathbf{P}_{2,2}$ is irreducible then the positive vector \mathbf{v}_2 and ρ_2 are such that*

$$\rho_2 = \inf_{\mathbf{c}_1 \geq \mathbf{0}, \mathbf{c}_2 \in \mathcal{D}_2} (\max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2))) = G_1(\mathbf{v}_2) > 0$$

Proof. Proof is in two steps.

First step: show that $\rho_2 = G_1(\mathbf{v}_2) = \inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2)$;

Second step: show that $\inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2) = \inf_{\mathbf{c}_1 \geq \mathbf{0}, \mathbf{c}_2 \in \mathcal{D}_2} (\max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2)))$.

If matrix $\mathbf{P}_{2,2}$ is irreducible then the spectral radius ρ_2 is positive and there exists a unique positive stochastic left-eigenvector \mathbf{v}_2 corresponding to ρ_2 . Note that ρ_2 is also the spectral radius of matrix $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}$ since $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}$ is matrix $\mathbf{P}_{2,2}$ up to a basis change.

We always have

$$\inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2) \leq \rho_2 = G_1(\mathbf{v}_2). \quad (22)$$

Let us assume that $\inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2) = \lambda < \rho_2$. Therefore, there exists $\varepsilon > 0$ and a stochastic vector $\mathbf{w}_2 \in \mathcal{D}_2$ such that

$$G_1(\mathbf{w}_2) \leq \lambda + \varepsilon < \rho_2$$

which is equivalent to

$$\mathbf{w}_2 \mathbf{P}_{2,2} \mathbf{U}_{\eta_2} \leq (\lambda + \varepsilon) \mathbf{w}_2 \mathbf{U}_{\eta_2}. \quad (23)$$

Since $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2} \geq 0$, we easily see that

$$\forall n \geq 1, \quad \mathbf{U}_{\eta_2}^{-1} \frac{1}{n} \sum_{k=1}^n \left(\frac{\mathbf{P}_{2,2}}{\rho_2} \right)^k \mathbf{U}_{\eta_2} \geq 0. \quad (24)$$

We get from Relation (23) that

$$\mathbf{0} \leq \mathbf{w}_2 \frac{1}{n} \sum_{k=1}^n \left(\frac{\mathbf{P}_{2,2}}{\rho_2} \right)^k \mathbf{U}_{\eta_2} \leq \frac{1}{n} \sum_{k=1}^n \left(\frac{\lambda + \varepsilon}{\rho_2} \right)^k \mathbf{w}_2 \mathbf{U}_{\eta_2}. \quad (25)$$

Note that series $(\frac{1}{n} \sum_{k=1}^n \left(\frac{\mathbf{P}_{2,2}}{\rho_2} \right)^k)_{n \geq 1}$ in (24) converges to the matrix $\mathbf{x}_2^* \mathbf{v}_2$ where \mathbf{x}_2 is a positive right-eigenvector of matrix $\mathbf{P}_{2,2}$ (see e.g. [11]). Taking limit as n grows to infinity in (25), we have $\mathbf{w}_2 \mathbf{x}_2^* \mathbf{v}_2 = \mathbf{0}$. Because $\mathbf{x}_2, \mathbf{v}_2 > \mathbf{0}$, it implies $\mathbf{w}_2 = \mathbf{0}$, which is a contradiction.

When matrix $\mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2}$ is irreducible, let us rewrite Relation (23) as

$$\mathbf{0} \leq \mathbf{w}_2 \mathbf{U}_{\eta_2} \mathbf{U}_{\eta_2}^{-1} \mathbf{P}_{2,2} \mathbf{U}_{\eta_2} \leq (\lambda + \varepsilon) \mathbf{w}_2 \mathbf{U}_{\eta_2}.$$

One just has to remark that series $(\frac{1}{n} \sum_{k=1}^n \left(\frac{U_{\eta_2}^{-1} \mathbf{P}_{2,2} U_{\eta_2}}{\rho_2} \right)^k)_{n \geq 1}$ converges to the matrix $\tilde{\mathbf{x}}_2^* \tilde{\mathbf{v}}_2$ with $\tilde{\mathbf{x}}_2, \tilde{\mathbf{v}}_2 > \mathbf{0}$. The fact that $\mathbf{w}_2 U_{\eta_2} \tilde{\mathbf{x}}_2^* = \mathbf{0}$ implies $\mathbf{w}_2 = \mathbf{0}$, which is a contradiction.

Thus, we have equality in (22) when matrix $\mathbf{P}_{2,2}$ or $U_{\eta_2}^{-1} \mathbf{P}_{2,2} U_{\eta_2}$ is irreducible.

It remains to compare $\inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2)$ to $\inf_{\mathbf{c}_1 \geq \mathbf{0}, \mathbf{c}_2 \in \mathcal{D}_2} (\max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2)))$. It is clear that

$$\inf_{\mathbf{c}_1 \geq \mathbf{0}, \mathbf{c}_2 \in \mathcal{D}_2} (\max(G_1(\mathbf{c}_2), G_2(\mathbf{c}_1, \mathbf{c}_2))) \geq \inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2).$$

Now, since the left hand side of the inequality is less than ρ_2 and ρ_2 is $G_1(\mathbf{v}_2) = \inf_{\mathbf{c}_2 \in \mathcal{D}_2} G_1(\mathbf{c}_2)$ from the first part of the proof, the proposition is proved. \square

Remark 5.1 *In fact, it is easy to check that we can always select as matrix \hat{L} with the initial choice of vectors $\mathbf{c}_1, \mathbf{c}_2$*

$$\hat{L} = \begin{pmatrix} 0 & 1 \\ 1 - \rho_2 & \rho_2 \end{pmatrix}.$$

This shows that computation of $\hat{L}(1, 2)$ is always possible, that is $0 \leq \max(C_1, C_2) \leq 1$.

Remark 5.2 *Assumptions $U_{\eta_2}^{-1} \mathbf{P}_{2,2} U_{\eta_2}$ is irreducible and $\mathbf{P}_{2,2}$ is irreducible are not related. Indeed, the following matrix $\mathbf{P}_{2,2}$ is such that $U_{\eta_2}^{-1} \mathbf{P}_{2,2} U_{\eta_2} \geq 0$:*

$$\begin{pmatrix} 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 \\ 0 & 2/5 & 3/5 \end{pmatrix}.$$

Matrix $\mathbf{P}_{2,2}$ is irreducible but $U_{\eta_2}^{-1} \mathbf{P}_{2,2} U_{\eta_2}$ is not. Conversely, the following matrix $\mathbf{P}_{2,2}$ is reducible

$$\begin{pmatrix} 0 & 1/5 & 1/5 \\ 0 & 1/5 & 2/5 \\ 0 & 2/5 & 3/5 \end{pmatrix}$$

but $U_{\eta_2}^{-1} \mathbf{P}_{2,2} U_{\eta_2}$ is irreducible.

5.3 Illustrative example.

Let us consider the queue-size process, $X = (X_n)_{n \in \mathbb{N}}$ of a discrete time queue, departure first, with one server and capacity C (server included). Service duration is deterministic and equal to 1. The queue has i.i.d. batch arrivals specified by the row vector $\mathbf{b} = (b_0, \dots, b_a)$ with $a > 0$, b_i denoting the probability of i arrivals in time interval $[n, n+1)$, $n = 0, 1, 2, \dots$. The aim of this subsection is to study the

output process of such a queue and illustrate our method on a simple numerical example.

X is a DTMC on state space $S = \{1, \dots, C + 1\}$. State 1 means that queue is empty. Output process of such a queue is $X^\phi = (\phi(X_n))_{n \in \mathbb{N}}$ with $\phi(X_n) = 1 + 1_{\{X_n \geq 2\}}$ for every n . Indeed, if queue is not empty at time n there will be one departure at time $n + 1$ because service time is 1.

Let us consider the following numerical example: $a = C = 3$ and $\mathbf{b} = (0.35, 0.05, 0.4, 0.2)$. The t.p.m. \mathbf{P} of X is

$$\mathbf{P} = \left(\begin{array}{c|ccc} 0.35 & 0.05 & 0.4 & 0.2 \\ \hline 0.35 & 0.05 & 0.4 & 0.2 \\ 0.0 & 0.35 & 0.05 & 0.6 \\ 0.0 & 0.0 & 0.35 & 0.65 \end{array} \right).$$

\mathbf{P} is \leq_{st} -monotone and partition of S is $\phi^{-1}(1) = \{1\}$, $\phi^{-1}(2) = \{2, 3, 4\}$.

We have to choose vectors $\mathbf{c}_1, \mathbf{c}_2$. Because of cardinality of set $\phi^{-1}(1)$, vector \mathbf{c}_1 is (1). Condition (Γ) , Lemma 4.1 is rewritten as follows

$$\begin{aligned} (\Gamma_1) : \quad \mathbf{c}_1 \mathbf{P}_{1,2} \mathbf{U}_3 &\leq \mathbf{c}_2 \mathbf{U}_3 \quad \text{and} \quad (\Gamma_2) : \quad \mathbf{c}_2 \mathbf{P}_{2,2} \mathbf{U}_3 \leq \mathbf{c}_2 \mathbf{U}_3 \\ \text{with } \mathbf{P}_{1,2} &= (0.05, 0.4, 0.2) \quad \text{and} \quad \mathbf{P}_{2,2} = \begin{pmatrix} 0.05 & 0.4 & 0.2 \\ 0.35 & 0.05 & 0.6 \\ 0.0 & 0.35 & 0.65 \end{pmatrix}. \end{aligned}$$

Let us examine two cases. The first case corresponds to $\mathbf{c}_2 = \frac{1}{3} \mathbf{1}_3$ which clearly does not verify (Γ_2) (note that \mathbf{C} is then the Schweitzer's matrix). We will show that only lower bound can be obtained. The second case corresponds to $\mathbf{c}_2' = \frac{1}{10} (1, 2, 7)$. This vector satisfies conditions (Γ_1, Γ_2) and we will see that upper bound exists. So, we consider the two matrices $\mathbf{C} = \text{diag}((1), \frac{1}{3} \mathbf{1}_3)$ and $\mathbf{C}' = \text{diag}((1), \frac{1}{10} (1, 2, 7))$. Note that we have $\mathbf{C} \prec_U \mathbf{C}'$. Thus a \mathbf{C}' -lumpable upper bound exists but no \mathbf{C} -lumpable upper bound exists.

First case $\mathbf{c}_2 = \frac{1}{3} \mathbf{1}_3$. Using our method, we have to solve System (14) of inequalities (upper bound case). Writing down (14) (with $I = J = 2$), we get the following inequality

$$\frac{1}{3} (0.2 + 0.6 + 0.65) \leq \frac{1}{3} \widehat{L}(2, 2)$$

which can not be satisfied with $0 \leq \widehat{L}(2, 2) \leq 1$. Therefore, there is no solution and it is not possible to derive an upper bound from this vector \mathbf{c}_2 . Let us note that looking at lower bound, we have to reverse inequalities in System (14). Then we conclude to the existence of at least one lower \mathbf{C} -lumped matrix, $\widehat{\underline{L}}$. Take for instance:

$$\widehat{\underline{L}} = \begin{pmatrix} 0.4 & 0.6 \\ 0.117 & 0.883 \end{pmatrix}. \quad (26)$$

Second case $\mathbf{c}'_2 = \frac{1}{10} (1, 2, 7)$. System (14) reduces to

$$\begin{aligned} I = 1, J = 1 : \widehat{L}(1, 1) + \widehat{L}(1, 2) &= 1 & I = 1, J = 2 : \widehat{L}(1, 2) &\geq \frac{2}{3} \\ I = 2, J = 1 : \widehat{L}(2, 1) + \widehat{L}(2, 2) &= 1 & I = 2, J = 2 : \widehat{L}(2, 2) &\geq \frac{0.89}{0.90}. \end{aligned}$$

So, there exists at least an upper bounding \mathbf{C}' -lumped matrix. Let us take $\widehat{\mathbf{L}}$ as

$$\widehat{\mathbf{L}} = \begin{pmatrix} \frac{1}{90} & \frac{2}{90} \\ \frac{1}{90} & \frac{89}{90} \end{pmatrix}. \quad (27)$$

Note that we have the following conditions on \widehat{L} to derive a lower bound:

$$\begin{aligned} \widehat{L}(1, 1) + \widehat{L}(1, 2) &= 1, & \widehat{L}(1, 2) &\leq 0.65 \\ \widehat{L}(2, 1) + \widehat{L}(2, 2) &= 1, & \widehat{L}(2, 2) &\leq 0.85. \end{aligned}$$

Consider $\boldsymbol{\alpha} = (1, 0, 0, 0) \in \mathcal{A}(S)$ as initial distribution (that is queue is initially empty). Take $\boldsymbol{\beta} = (1, 0) \in \mathcal{A}(\Sigma)$. We obviously have that $\boldsymbol{\beta} \mathbf{C} \leq_{st} \boldsymbol{\alpha} \leq_{st} \boldsymbol{\beta} \mathbf{C}'$ since $\boldsymbol{\beta} \mathbf{C} = \boldsymbol{\beta} \mathbf{C}' = \boldsymbol{\alpha}$. Then we deduce from Corollary 3.2 that the DTMC $(\bar{Y}_n)_{n \in \mathbb{N}} = (\boldsymbol{\beta}, \widehat{\mathbf{L}})$, with $\widehat{\mathbf{L}}$ defined by (27), is an \leq_{st} -upper Markovian bound on aggregated process $(\phi(X_n))_{n \in \mathbb{N}}$. We also recall that, from Proposition 2.1, it is possible to obtain \leq_{st} -bound on RV X_n with the help of RV \bar{Y}_n . Indeed, the probability vector $\boldsymbol{\beta} (\widehat{\mathbf{L}})^n \mathbf{C}$ is an \leq_{st} -upper bound on the probability distribution of X_n . We report in Table 1, the computation, for $n = 0, 1, 5, 10$, of the probability distribution of the RVs X_n (column (1)), $\phi(X_n)$ (column (3)) and bound \bar{Y}_n (column (4)). In column (2), we compute the \leq_{st} -upper bound on the probability distribution of RV X_n . Table 1 clearly illustrates results of Corollary 3.2 and Proposition 2.1:

$$\begin{aligned} \forall n, \quad \boldsymbol{\alpha} \mathbf{P}^n \mathbf{V} &\leq_{st} \boldsymbol{\beta} (\widehat{\mathbf{L}})^n \text{ (compare column (3) with column (4))} \\ \forall n, \quad \boldsymbol{\alpha} \mathbf{P}^n &\leq_{st} \boldsymbol{\beta} (\widehat{\mathbf{L}})^n \mathbf{C} \text{ (compare column (1) with column (2)).} \end{aligned}$$

Results for lower bound $\underline{Y} = (\boldsymbol{\beta}, \underline{\mathbf{L}})$ with $\underline{\mathbf{L}}$ defined by (26) are given in Table 2.

In Table 3, we report numerical results concerning maximum deviation at each step n for the aggregated process and the original chain, $d(n)$ and $D(n)$, defined by (12) and (13), respectively. $d(n)$ is computed from columns (4) of Table 1 and Table 2, $D(n)$ from columns (2).

Remark 5.3 *Using generalized coupling method of Doisy [4], we get the following upper bounding matrix $\begin{pmatrix} 0.35 & 0.65 \\ 0.0 & 1.0 \end{pmatrix}$. It is easily seen that accuracy of this bound on the aggregated process is worse than ours when time grows up.*

n	(1) : $\alpha \mathbf{P}^n \sim X_n$	(2) : $\beta (\widehat{\mathbf{L}})^n \mathbf{C}'$
0	(1, 0, 0, 0)	(1, (0, 0, 0))
1	(0.35, 0.05, 0.4, 0.2)	(0.33, (0.067, 0.134, 0.469))
5	(0.0646, 0.1049, 0.2772, 0.5533)	(0.0194, = (0.09806, 0.19612, 0.68642))
10	(0.0564, 0.1044, 0.2754, 0.5638)	(0.0162, = (0.09838, 0.19676, 0.68866))

n	(3) : $\alpha \mathbf{P}^n \mathbf{V} \sim \phi(X_n)$	(4) : $\beta (\widehat{\mathbf{L}})^n \sim \bar{Y}_n$
0	(1, 0)	(1, 0)
1	(0.35, 0.65)	(0.33, 0.67)
5	(0.0646, 0.9354)	(0.0194, 0.9806)
10	(0.0564, 0.9436)	(0.0162, 0.9838)

Table 1: Upper bound with $\mathbf{C}' = \text{diag}((1), \frac{1}{10}(1, 2, 7))$.

n	(1) : $\alpha \mathbf{P}^n \sim X_n$	(2) : $\beta \widehat{\mathbf{L}}^n \mathbf{C}$
0	(1, 0, 0, 0)	(1, (0, 0, 0))
1	(0.35, 0.05, 0.4, 0.2)	(0.4, (0.2, 0.2, 0.2))
5	(0.0646, 0.1049, 0.2772, 0.5533)	(0.1647, (0.2785, 0.2784, 0.2784))
10	(0.0564, 0.1044, 0.2754, 0.5638)	(0.1632, (0.2790, 0.2789, 0.2789))

n	(3) : $\alpha \mathbf{P}^n \mathbf{V} \sim \phi(X_n)$	(4) : $\beta \widehat{\mathbf{L}}^n \sim \underline{Y}_n$
0	(1, 0)	(1, 0)
1	(0.35, 0.65)	(0.4, 0.6)
5	(0.0646, 0.9354)	(0.1647, 0.8353)
10	(0.0564, 0.9436)	(0.1632, 0.8368)

Table 2: Lower bound with $\mathbf{C} = \text{diag}((1), \frac{1}{3}(1, 1, 1))$.

6 Conclusion

In this paper we develop a new methodology to provide family of Markovian bounds on aggregated process defined as a surjective nondecreasing function of a monotone Markov chain. Namely, we investigate bounds on the finite dimensional distributions of the aggregated process with respect to the stochastic strong order.

Polyhedral cone is the central concept of our methodology. Polyhedral cone generated by \mathbf{U}_η induces strong order between S -valued RV and also comparison between stochastic processes. Rogers and Pitman's lumpability criterion for a Markov chain is nothing else but the invariance of a polyhedral cone generated by some matrix \mathbf{C} (see [8]). Thus, we have to choose a stochastic matrix $\mathbf{C} = \text{diag}(\mathbf{c}_I)$ which also ensures that an upper Markovian bound on aggregated process exists (lower bound always exists). In fact, the only constraint to obtain such a Markovian upper

n	$d(n)$	$D(n)$
0	0	0
1	0.07	0.269
5	0.1453	0.408
10	0.147	0.409

Table 3: Bounds on error made.

bound is to choose \mathbf{c}_N as a stochastic vector of the polyhedral cone generated by matrix $(\mathbf{I} - \mathbf{P}_{N,N}) \mathbf{U}_{\eta_N}$. For $I = 1, \dots, N-1$, \mathbf{c}_I may be any probability vector on $\phi^{-1}(I)$.

We emphasize that it is very difficult to know *a priori* if bounds are accurate. It mainly depends on choice of matrix \mathbf{C} . Paper proposes to take $\mathbf{C} = \text{diag}(\mathbf{v}_I)$, where \mathbf{v}_I is the stochastic left-eigenvector associated to spectral radius ρ_I of $\mathbf{P}_{I,I}$. This choice is motivated by two main arguments: (a) if \mathbf{P} is \mathbf{C} -lumpable then upper \mathbf{C} -lumpable bound \mathbf{L} coincides with \mathbf{P} ; (b) in the case of two aggregates, select \mathbf{v}_2 ensures that spectral radius ρ_2 of an irreducible matrix $\mathbf{P}_{2,2}$ is the optimum value for $\hat{L}(2,2)$ (see Section 5). Note that using stochastic eigenvectors is not new in bounding methodology. We can think about work of Courtois and Semal [2],[3]. This work is mainly concerned with componentwise bounds on the stationary distribution π of an irreducible Markov chain $X = (X_n)_{n \geq 0}$. We briefly recall the background to the derivation of such bounds. Let us write vector π as

$$\pi = \left(\frac{\pi_I}{\pi_I \mathbf{1}^*} \pi_I \mathbf{1}^* \right)_{I=1, \dots, N}$$

for any arbitrary partition of the state space S (π_I is the restriction of π to states class $\phi^{-1}(I)$). Vector $\pi^{(I)} = \frac{\pi_I}{\pi_I \mathbf{1}^*}$ is the stochastic left-eigenvector of the t.p.m. of Markov chain $(X_{T_n})_{n \geq 1}$ on the state space $\phi^{-1}(I)$, where T_n is the n th entrance epoch into subset $\phi^{-1}(I)$ by X . Vector $\hat{\pi} = (\pi_I \mathbf{1}^*)_{I=1}^N$ is the stochastic left-eigenvector of matrix \hat{L} defined by

$$\hat{L}(I, J) = \pi^{(I)} \mathbf{P}_{I,J} \mathbf{1}^*, \quad I, J = 1, \dots, N.$$

Then we have to obtain componentwise bounds on vector $\pi^{(I)}$ to get componentwise bounds on vector π . We see that their approach and ours are related in the sense that \hat{L} will be the t.p.m. of the aggregated process $(\phi(X_n))_{n \geq 0}$ if Markovian. But this the only connection between them.

References

- [1] BUCHHOLZ, P. (1994). Exact and ordinary lumpability in finite Markov chains. *J. Appl. Probab.* **31**, 59–75.
- [2] COURTOIS, P. J. AND SEMAL, P. (1984). Bounds for the positive eigenvectors of nonnegative matrices and for their approximations by decomposition. *JACM* **31**, 804–825.
- [3] COURTOIS, P. J. AND SEMAL, P. (1985). On polyhedra of Perron-Frobenius eigenvectors. *Lin. Alg. Appl.* **65**, 157–170.
- [4] DOISY, M. (2000). A coupling technique for comparison of functions of Markov processes. *J. Appl. Math. and Decision Sciences* **4**, 39–64.

- [5] KEILSON, J. AND KESTER, A. (1977). Monotone matrices and monotone Markov processes. *Stochastic Process. Appl.* **5**, 231–241.
- [6] KEMENY, J. AND SNELL, J. (1960). *Finite Markov Chains*. Princeton.
- [7] KIJIMA, M. (1997). *Markov Processes for Stochastic Modeling*. Chapman-Hall.
- [8] LEDOUX, J. (1997). A geometric invariant in weak lumpability of finite Markov chains. *J. Appl. Probab.* **34**, 847–858.
- [9] ROGERS, L. C. G. AND PITMAN, J. W. (1981). Markov functions. *Ann. Probab.* **9**, 573–582.
- [10] SCHWEITZER, P. Aggregation methods for large Markov chains. *Math. Comp. Perf. and Reliab.* Ed. G. Iazeola et al. Elsevier-North-Holland, Amsterdam, 1984.
- [11] SENETA, E. *Non-negative matrices and Markov chains*. Springer-Verlag.
- [12] STOYAN, D. (1976). *Comparison Methods for Queues and Other Stochastic Models*. J. Wiley and Sons.
- [13] TRUFFET, L. (2000). Reduction technique for discrete time Markov chains on totally ordered state space using stochastic comparisons. *J. Appl. Probab.* **37**, 795–806.